Background and scope of the thesis

1. Prologue
2. Innate and adaptive immunity work together and are linked
3. The T cell response in a nutshell
4. Protein degradation pathways and the generation of T cell epitopes
   4.1. The autophagic pathway of antigen processing
   4.2. The endocytic pathway of antigen processing
   4.3. Cross-presentation: cross-talk between antigen processing pathways
   4.4. The ubiquitin-proteasome pathway for MHC class I antigen processing
      4.4.1. Structure of the proteasome
      4.4.2. Cleavage specificity of the proteasome
      4.4.3. Substrates for the UPS, rapid protein turnover and the DRiP model
      4.4.4. Generation of class I ligands, overview
      4.4.5. Generation of class I ligands, post-proteasomal processing
      4.4.6. Major unresolved questions in class I antigen processing
5. Immunity to cancer and immunoediting
6. T cell mediated immunotherapy for cancer, modalities, and basic requirements
   6.1. Adoptive transfer of undefined tumor-specific or defined epitope-specific T cells
   6.2. Vaccination strategies with undefined antigens
   6.3. Vaccination strategies with defined full length tumor associated antigens
   6.4. Vaccination strategies with defined T cell epitope containing synthetic peptides
7. Tumor associated antigens and their classification
   7.1. Strategies for the identification of tumor associated antigens
   7.2. Selection of tumor associated antigens for T cell immunotherapy
8. Identification of tumor-specific T cell epitopes
   8.1. Identification of CTL epitopes starting with CTL of unknown specificity
   8.2. Identification of CTL epitopes by reverse immunology
      8.2.1. Prediction phase of reverse immunology
      8.2.2. Improved CTL epitope prediction by verification of proteasomal processing and TAP translocation
      8.2.3. Validation phase of reverse immunology
   8.3. Identification of HLA class II presented T helper epitopes
   8.4. Identification of HLA class I and HLA class II ligands by tandem mass spectrometry
9. The need to define more T cell epitopes: multi-epitope-based cancer immunotherapy
10. Purpose and chronology of the thesis
1. Prologue

More than three decades after the discovery of MHC restriction [1], two decades after the first elucidation of MHC structure [2] and the peptide in the MHC-groove [3], and 18 years after both the finding of MHC-specific peptide binding motifs [4] and the identification of the first tumor-specific cytotoxic T lymphocyte epitope [5,6], T cell mediated immunotherapy of cancer has now outgrown its infancy. Nevertheless, trials in cancer patients have not shown consistent and high percentages of clinical successes [7-9], and immunotherapy of cancer is, with a single exception [10], not yet a standard (adjuvant) therapy.

Although the T cell arm of the immune system is exquisitely equipped to eradicate virally infected cells, the similar use of T cells for the destruction of cancer cells that (over)express tumor specific proteins has still to be exploited to its full clinical potential. Our rapidly accumulating understanding of the mechanisms involved in the adequate induction of anti-tumor immunity in patients is currently being used for the design of more effective immunotherapeutic treatments that will likely raise clinical success rates. For the development of effective T cell mediated cancer therapies it is crucially important that both an optimal immunostimulatory context is realized and that the targets of the CD8\(^+\) cytotoxic T lymphocytes (CTL) and CD4\(^+\) T helper (Th) lymphocytes are properly chosen. These targets are the tumor-associated antigens (TAA) expressed in the tumor cells and more specifically the T cell epitopes contained in these proteins.

The studies in this thesis address the epitopes recognized by CTL: the events leading to their generation and presentation and, based on these mechanisms, the prediction and identification of cancer-specific epitopes to be used as targets in cancer immunotherapy. This introductory chapter will therefore especially review current knowledge regarding T cell immunity in cancer, its induction by immunotherapy, the identity of TAA, and the generation, prediction and identification of T cell epitopes.

Because two of the studies in this thesis are at the basis of some reports in the literature, the outline and scope of each chapter is integrated, as boxed intermezzo, at its appropriate place in the text.

2. Innate and adaptive immunity work together and are linked

The human immune system is equipped with innate and adaptive (or acquired) arms that are defending the body against foreign pathogens. The innate system, composed of primarily macrophages, dendritic cells (DC), granulocytes, natural killer cells and the complement system, is evolutionary much older and constitutes a first line of defence that is already present at the start of the immune response and immediately interacts with pathogens. Foreign antigens, cells sensed as abnormal and conserved structures shared by large groups of micro-organisms (so-called pathogen-associated molecular patterns). On the other hand, the adaptive response, executed by B- and T lymphocytes, before it can exert its effector functions, first needs amplification and selection for which it uses clonal receptors with narrow specificity generated by gene rearrangements and somatic mutations. Thereby, the adaptive response, unlike the innate response, develops immunological memory. It is only since the last decade that the prescient prediction by Charles Janeway in 1989 that the innate immune system is driving the adaptive
immune system [11] has been firmly grounded and is dissected at the molecular level [12]. From a historical viewpoint, the model for adaptive immune response regulation has undergone strong changes in the last 50 years. In the Self-Nonself (SNS) model, proposed in 1959 by Burnet and Medawar, the induction of the response was completely defined at the level of the lymphocyte by the non-self nature of the antigen that is recognized (this so-called signal one is the self-nonself discriminator in this model). Under the pressure of accumulating incompatible observations that needed additional explanation, the SNS model has been strongly adapted and refined. In the 1970s, a helper cell (later found to be the T helper cell) providing help was proposed [13], and later the focus shifted to the stimulator cell that induces lymphocyte activation. The stimulator cell (now called antigen presenting cell; APC) was proposed to provide, next to the antigen-specific signal one, a necessary second costimulatory signal to the lymphocyte [14]. Ten years later, in 1986, this was confirmed empirically by Jenkins and Schwartz [15]. In 1989, Charles Janeway hypothesized that the costimulation provided by the APC first needs to be induced through ligation of so-called pattern recognition receptors (PRR) on the APC by conserved pathogen-associated molecular patterns (PAMPs) of bacteria [16], thereby linking innate and adaptive immunity. Thus, the PRR allow APC to discriminate between infectious-nonself and non-infectious-self. Therefore, this model has been coined either the infectious-Nonself model, PRR-model or Stranger-model [17,18]. In 1997, the first Toll like receptor (TLR) was identified [19] and was shown to act as PRR for components of bacteria. Since then numerous TLR that recognize a variety of conserved microbial-associated products, like lipopolysaccharide (LPS), have been identified [20], thereby further unravelling the linkage between innate and adaptive immunity. Meanwhile, in 1994, Polly Matzinger proposed the so-called Danger theory (of immune activation) [21]. This theory, initially purely theoretical, made ‘danger’ caused to cells and tissues the central concept and added still an extra level of cells and signals to the activation of the APC. Its activation could also be induced by danger/ alarm signals released from injured (necrotic) cells, thus including endogenous non-foreign signals in APC-activation and not putting the primacy on innate immunity. As Matzinger stated herself, “the danger theory, may seem to propose just one more step down the path of slowly increasingly complex cellular interactions, this small step drops us off a cliff, landing us in a totally different viewpoint, in which the ‘foreignness’ of a pathogen is not the important feature that triggers a response, and ‘selfness’ is no guarantee of tolerance” [22]. The Danger theory has been criticized by Janeway and his co-worker Medzhitov, because of its “inherent tautology”. As they state it (and rather ridicule the theory): “the (adaptive) immune response is induced by a danger signal, but the danger signal is defined as just about anything that can induce an immune response” [23]. Others, as well, have pointed to the flaws in this theory, being above all its conceptual emptiness and vagueness when the concept ‘danger’ is not specified and thus its metaphorical generalizing character [24]. Currently in 2009, studies have revealed several endogenous non-foreign alarm signals, like heat shock proteins, interferon-α (IFNα), interleukin-1β (IL-1β) and CD40-ligand (CD154) [25]. Furthermore, TLRs have been found to engage not only pathogenic components but also those from endogenous origin. Thus, one may argue that the Danger model is the most comprehensive theory because it incorporates also endogenous APC-ac-
tivation signals. However, this model may underestimate the importance of the exogenous pathogenic signals and it was Janeway who was the first to propose, and his group the first to identify, the important linkage between the innate and adaptive immune responses inducing the costimulatory signals of the APC.

As the immune system is a diverse collection of mechanisms that have come together during the course of evolution, it is impossible to explain its complexity by a too much restricted paradigm (the Stranger model) and not helpful to do so by a too much generalized metaphorical paradigm, like the Danger model). Indeed, later Matzinger [26,27] and others [28] have tried to reconcile both models. In any case, together innate and adaptive immune mechanisms can counteract the attack of in principle all pathogens ranging from viruses and bacteria to multi-cellular parasitic organisms.

3. The T cell response in a nutshell

Whereas B lymphocytes upon antigen-encounter produce antibodies (soluble B cell receptors) that recognize pathogen-derived native proteins, polysaccharides and lipids, T lymphocytes by their cell surface expressed T cell receptor (TCR) specifically recognize short linear protein sequences, i.e. peptides, derived from either endogenous or exogenous proteins. Upon proper activation, B- and T cells divide, expand in numbers, exert their effector functions, and memory is installed.

Two major subsets of T cells collaborate to mediate an effective immune response: CD8+ cytotoxic T lymphocytes (CTL) recognize short peptides of a defined length (8–12 aa) presented by HLA class I molecules on the cell surface and CD4+ T helper lymphocytes (Th cells) recognize longer peptides of less defined length (15–20 aa) that are presented by HLA class II molecules. Th cells are involved in the activation and regulation of B cells, CTL and APC through secreted cytokines and cell surface expressed molecules like CD40-ligand (CD154). Th type 2 (Th2) cells are mainly involved in B cell activation and Th type 1 (Th1) cells accomplish CTL activation via their stimulatory effect on APC and by secretion of cytokines. Moreover, regulatory CD4+ T cells (Treg) exist that down-regulate T cell responses, e.g. preventing autoimmunity but also suppressing anti-tumor responses [29]. CTL are the killer cells that lyse target cells expressing their cognate class I-presented peptide by perforin and/or Fas-mediated mechanisms. T cell activation is accomplished when the peptide (first signal) is presented in an appropriate costimulatory context by the APC, in particular the DC.

This second signal can be provided by any of the molecules within the B7-family [30] or TNF receptor (TNFR)-family [31] of proteins. Crucially important costimulatory molecules of the B7-family are CD80 (B7.1) and CD86 (B7.2) whose interaction with CD28 activates T cells [32]. Reversely, CD80/86 ligation of the T cell-expressed counter-regulatory receptor CTLA-4, whose expression is upregulated after T cell activation [33], attenuates T cell responses by feedback inhibition. For sustained T cell effector functions, survival and memory maintenance, additional signals are required. TNFR-family members that function after initial T cell activation to further costimulate and sustain T cell responses are CD27 [34], 4-1BB (CD137) [35], OX40 (CD134) and GITR, all expressed on (activated) T cells, interacting with CD70, 4-1BB-ligand, OX40-ligand and GITR-ligand, expressed by the APC [31]. Adhesion molecules like intercellular adhesion molecule 1 (ICAM-1) interacting with LFA-1 on T cells also contribute to T cell activation [36]. Apart from cell surface receptors that mediate co-
stimulation, the cytokine milieu composed of especially IL-2 (secreted by Th cells) [37], IL-7 [38], IL-12 (secreted by DC) [39-41], and IL-15 [42,43] has been shown to be decisive for T cell activation, function and memory. Next to positive signalling, negative regulation of T cell activation is accomplished by ligation of PD-1 (a member of the CD28-family expressed on activated T cells) with B7-family member PD-1 ligand (PD-L1 or B7-H1), which is sometimes overexpressed on tumor cells [30]. B7-family member ICOS-ligand (B7h) that engages with ICOS which is expressed on activated and resting memory T cells [44] may down regulate Th1-responses through the induction of IL-10 [45]. Together, the summation of positive and negative signals coming together in the immunological synapse between APC, in particular DC, and T cell [46] determines the activation, proliferation, effector functions and instalement of memory of T cells. Once activated, T cells loose receptors that are required for lymph node entry, in particular CD62L [47] and CCR7 [48], and accordingly can migrate from the lymph node via the blood into the peripheral tissues.

Thus, to accomplish Th and CTL activation, DC need to provide appropriate costimulation. This is induced after maturation (activation) via TLR signalling by PAMPS (e.g. lipopolysaccharide; LPS), and/or via ligation of cell surface expressed CD40 by CD40-ligand (expressed on activated Th cells) [49-52]. CD40 can be seen as a master switch for T cell costimulation because of its ability to induce B7-family ligands as well as several TNF family ligands on DC [31,53]. Optimal DC maturation is enhanced by proinflammatory cytokines like tumor necrosis factor (TNF)α, interferon (IFN)β, INFγ and IL-1β [54]. Immature DC, residing in peripheral tissues – particularly in barrier organs such as the skin and bowel – but also in the blood [55], are dedicated to capturing antigens, mostly by endocytosis. DC maturation via TLRs, together with partially unresolved mechanisms [56], induce migration of DC to lymph nodes [57,58] where they acquire the ‘mature’ phenotype specialized at presenting antigens and stimulating T cells through enhanced expression of CD80/86 and secretion of IL-12, a cytokine crucial for CTL effector and memory formation [41,59]. DC are believed to be at the crossroads of immunity and tolerance dependent on their maturation status [60]. On the other hand, as outlined, when in an immunogenic context DC activate naive anti-foreign T cells and on the other hand, when DC have an immature phenotype, they are capable of tolerizing autoreactive T cells – which have escaped the process of central tolerance – in the periphery (a process called peripheral tolerization) [61]. Several subsets of DC exist in vivo with distinct roles in immunity to infection and maintenance of self tolerance dependent on differences in their location and intrinsic abilities to capture, process and present antigens [62].

4. Protein degradation pathways and the generation of T cell epitopes

HLA class I molecules can be found on the surface of virtually all nucleated cells, whereas HLA class II molecules are mainly expressed by APC, but also by inflamed cells and thymus epithelial cells. Peptides presented by HLA class I and class II molecules are produced through proteolysis in one of the three major intracellular protein degradation and antigen processing systems that exist: (i) Ubiquitin (Ub)-mediated protein degradation that proceeds via the proteasome, which is called the ubiquitin-proteasome system (UPS); (ii) autophagy, which is an intracellular degradation system that delivers cytoplasmic constituents.
via the autophagosome to the lysosome. There are at least three different types of autophagy [63]: chaperone-mediated autophagy, microautophagy, and macroutrophagy, the latter being best characterized; (3) endocytosis-mediated lysosomal degradation of extracellular proteins and plasma membrane proteins. The classical but outdated doctrine holds that class I presented peptides are derived from endogenous proteins by UPS-mediated degradation, whereas class II presented peptides result from degradation of exogenous proteins in the endocytic pathway. However, as shown in table 1, current knowledge reveals exceptions to these rules and interconnections be-

Table 1. Common and uncommon antigen processing pathways for the generation of MHC class I and class II presented peptides.

<table>
<thead>
<tr>
<th>Class II presentation</th>
<th>Class I presentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>UPS pathway</td>
<td>exception (cytosolic antigens)&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Autophagic pathway</td>
<td>common (endogenous, often foreign, antigens)&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Endocytic pathway</td>
<td>common (exogenous antigens)&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> The UPS pathway has been shown to be connected to class II loading compartments in several studies [64–66], which are not discussed in the text and reviewed in ref. 67.

<sup>b</sup> The balance between the autophagic and endocytic pathways for class II presentation is not yet fully understood [68].

<sup>c</sup> Until now this pathway has been reported only twice, in conjunction with proteasomal degradation [69,70].

<sup>d</sup> Production of cross-presented peptides proceeds via the endocytic route mostly in conjunction with the UPS pathway, but can also proceed in a TAP- and proteasome-independent way [71] in the endocytic tract. Autophagy has been implicated in cross-presentation as well [72].

4.1. The autophagic pathway of antigen processing

Autophagy is in principle a nonselective process involved in removal of damaged or surplus organelles, turnover of long lived proteins, production of amino acids in nutrient emergency, and cell survival and death [73]. In autophagy part of the cytoplasm becomes surrounded by two concentric membranes. Fusion of the outer membrane of this so-called autophagosome with a lysosomal vesicle results in degradation of enclosed cytoplasmic structures and macromolecules. The autophagic process was already identified before the UPS, and its contribution to intracellular protein degradation is estimated to be as large as that of the UPS [74]. Lately, autophagy has attracted new research and its emerging roles in innate and adaptive immune responses are being unravelled [75,76]. In line with autophagy eliminating intracellular pathogens, this process has been shown to deliver viral genetic material to endosomal TLRs in plasmacytoid DC, thereby inducing IFNα secretion [77]. Importantly, autophagic sequestration of viral
components can fuel MHC class II presentation to CD4+ Th cells [78]. Peptide supply for class II presentation is even considered to depend significantly on autophagic degradation also of non-foreign (non-pathogenic) proteins [79]. Although from a protein trafficking perspective (autophagosomes being fused to late endosomes where class II loading occurs) it is easier to understand the role of autophagy in class II peptide-generation, autophagic degradation may also be involved in class I antigen processing, for instance by its clearance of ubiquitinilated cytoplasmic protein aggregates [80]. Recently, the first evidence for the involvement of autophagy in class I presentation, intricately linked to the proteasomal route, has been demonstrated in macrophages for endogenous antigens from herpes simplex virus type 1 [70]. This study suggests an intersection between the vacuolar and MHC class I presentation pathways.

4.2. The endocytic pathway of antigen processing

APC can internalize pathogens or parts thereof, dying virally infected cells and dying tumor cells into the endocytic pathway to provide a representation of the protein environment that they encounter in the periphery to T cells. The antigens can be endocytosed by a variety of mechanisms. Immature DC are highly efficient in all forms of endocytosis, being phagocytosis (for bacteria and cells, taken up in the phagosome), macro- and micro-pinocytosis and receptor-mediated uptake mechanisms [60]. Degradation of the antigens is accomplished in diverse endosomal-lysosomal compartments by a large collection of proteases with mostly broad substrate specificity and variable pH requirements [81]. Most of the endosomal proteases are known as cathepsins. In a late endosomal, early lysosomal compartment known as the MHC class II compartment (MIIC) the generated peptides encounter class II molecules and are loaded in the class II binding groove through exchange with the class I invariant chain peptide (so-called CLIP) [82]. Upon endocytosis of exogenous material, DC will be activated and migrate to T cell areas in the lymph nodes [83]. Thus, in general, pathogen-derived peptides produced in the endocytic pathway will allow the initiation of a CD4+ T cell response.

4.3. Cross-presentation: cross-talk between antigen processing pathways

For proper activation, naïve CD8+ T cells must be stimulated by signal one (the MHC-pep complex) together with costimulation provided by professional APC, such as DC. As DC are mostly not infected themselves by viruses (and neither often transformed), they must acquire the antigen exogenously in peripheral tissues and display it through a process termed cross-presentation to CD8+ T cells in the lymph nodes [84]. Thus, cross-presentation is the pathway by which the CD8+ T cell response can be initiated towards viral infections or mutations that exclusively occur in parenchymal cells [85]. DC are the principal cells endowed with the capacity to cross-present exogenous antigenic material. To this end, endocytosed antigenic material is transferred from the endosomal-lysosomal pathway into the cytosol for further processing by the UPS-pathway and MHC class I loading in the ER. Exogenous antigenic material can, therefore, end up stimulating either the CD4+ T cell response and/or the CD8+ T cell response, dependent on the intracellular antigen trafficking. The access to the cross-presentation pathway can occur already directly at the moment of endocytosis, as has been observed for antigen uptake via the mannose receptor [86], or antigens can divert...
to this pathway later. The escape of antigens from extracellular sources in the endocytic route to the cytosol may rely on diverse, as yet not completely resolved, mechanisms [68]. The delivery of proteins or peptides through a membrane pore, similar as the so-called ER-dislocon has been proposed. Likewise, fusion of the ER with phagosomes has been demonstrated explaining antigen transfer [87,88]. Alternatively, peptides and/or proteins may leak from the phagosomes into the cytosol or the phagolysosomal membrane may rupture.

Moreover, evidence exist for loading of class I molecules in the endo-lysosomal compartments themselves [89] allowing cross-presentation independent of cytosolic transit and without proteasomal involvement [71]. Spatial separation of cross-presentation and endogenous class I presentation may have the advantages of speed and absence of competition [90]. Importantly, recently also the autophagic pathway has been implicated in cross-presen-
Induction of (macro)autophagy in tumor cells was required for cross-presentation by DC in vitro and in vivo. Autophagosomes were suggested to be the antigen carriers in this study. Although the majority of proteasome-mediated and damaged proteins, but also in protein degradation, it has become clear since the continuous turnover of endogenous proteins is considered to start with substrates that are ubiquitinated, there is accumulating recent evidence that ubiquitin-independent degradation can be accomplished by both the 20S and 26S proteasome and may have been underestimated [101,102].

4.4.1. Structure of the proteasome
Proteasomes are complex multi-subunit proteases, ubiquitously expressed and abundantly present both in cytosol and the nucleus. Several types of proteasomes exist which share a proteolytically active core, the 20S proteasome [103]. This catalytic core unit is a cylindrical structure of four stacked rings. The two inner rings each consist of seven distinct β-subunits (β1-7) and the two outer rings are assembled from seven homologous but different α-subunits (α1-7) together forming a central channel in which proteolysis takes place. Thus of each subunit two copies are present in the 20S proteasome. The channel prevents unwarranted proteolysis of cellular proteins and its access is restricted to unfolded proteins or polypeptides [104]. Although the isolated 20S proteasome can degrade peptides in an ATP-independent manner, which can be used in vitro for the assessment of proteasomal cleavages in model polypeptides, it can not actively unfold native proteins.

The 26S proteasome consists of the 20S core unit capped – at one or both sites – with the 19S regulatory complex [103]. The 19S cap [105] recognizes multi-ubiquitinated proteins, unfolds these substrates by an ATP-dependent mechanism, removes ubiquitin chains, and provides a passageway for threading unfolded proteins into the 20S core complex by opening the gate of the channel that is otherwise blocked by N-termini of α-subunits. The proteasome activator 28 (PA28) [105,106] is another regulatory complex which forms a cap that, like the 19S cap, can associate with one or both ends of the 20S core particle, which may lead to hybrid proteasomes [107,108]. Recently, a thymus spe-
specific proteasome type has been discovered that incorporates a different \( \beta_5 \)-subunit [109] with a modulated cleavage specificity. Consequently, the class I-presented peptide repertoire in the thymus may be significantly different which could be important for proper positive selection of CD8\(^+\) T cells [110].

### 4.4.2. Cleavage specificity of the proteasome

The catalytic activities of the proteasome reside in three of the \( \beta \)-subunits, each in twofold present in the 20S unit. The \( \beta \)-subunit is responsible for the so-called caspase-like activity, cleaving after acidic or small hydrophobic residues, \( \beta_2 \) cleaves after basic or small hydrophobic amino acids (the trypsin-like activity) and the \( \beta_5 \)-subunit cuts after hydrophobic residues whether bulky or not (chemotryptsin-like activity) [111]. Consequently, and in line with its task to degrade a multitude of different substrates, the proteasome has a broad cleavage specificity with the capacity to cut in principle after or before all twenty amino acids [112-114]. However, not all amide (peptide) bonds are equally prone to be cleaved by the proteasome. For instance the proteasome does not readily cleave after a lysine [115]. Apart from the residues directly linking a scissile amide bond (indicated as P1 and P1’ residues), also residues in the N-terminal and C-terminal flanking regions, up to eight residues, contribute to the propensity of the proteasome to hydrolyze a specific bond. Consequently, single residue differences in epitope-flanking regions from two related viral proteins could lead to abrogation of CTL epitope production [116,117]. Similarly, carboxy-terminal liberation of a CTL epitope from hepatitis C virus was impaire due to a mutation in the flanking region [118]. The broad cleavage specificity of the proteasome makes it extremely difficult to deduce algorithms capable to reliably predict proteasomal cleavages in silico (see below). Proteasomes cleave in a processive manner [119], meaning that each substrate leaves the proteasome before the next one enters. The cleavage propensity of the proteasome is stochastic in nature [112]. Thus, a certain amide bond may be either cleaved or not in different copies of the same protein, mostly leading to a cleavage pattern with partially overlapping degradation fragments having different abundances and also resulting in differences in frequencies between cleavages. The length of proteasome degradation fragments varies between 3 and 23 amino acids, and the median length products of 7 to 9 aa comprise only \(~15\%\) of the products [112,113].

Under inflammatory circumstances when INF-\( \gamma \) is produced, and also constitutively in cells of lymphoid organs, proteasomes exchange the normal, so-called constitutive, catalytic subunits (\( \beta_1, \beta_2 \) and \( \beta_5 \)) with slightly different subunits, the immuno(i)-subunits \( \beta i_1 \) (also called LMP2), \( \beta i_2 \) (MECL1) and \( \beta i_5 \) (LMP7), giving rise to so-called immunoproteasomes [120-122]. Cleavage specificities of the immuno-subunits are quantitatively, and also slightly qualitatively, different compared to their constitutive counterparts [123]. Consequently, differences in epitope production have been found in cells expressing either of the subunit types. Some CTL epitopes were found dependent on immuno-subunits [124-126] while another epitope was only generated in cells expressing constitutive subunits [127]. Stimulation with INF-\( \gamma \) also has the capacity to induce PA28. The PA28 cap, together with the immuno-subunits, is up-regulated in DC upon maturation [128], although the effect on proteasome composition in DC is only very moderate due to low turnover of proteasomes [129]. The incorporation of the PA28 subunit reportedly enhanced the presentation of some
viral epitopes [126,130] but its precise influence on antigen processing is not yet completely resolved.

4.4.3. Substrates for the UPS, rapid protein turnover and the DRiP model
In 1996, it was proposed by Yewdell that a significant proportion of proteasomal substrates originate from so-called defective ribosomal products (DRiPs) [131]. This could explain the observation that after viral infection, epitopes derived from long-lived viral proteins are rapidly, within an hour, presented [132,133]. DRiPs include all proteins that fail to achieve a stable conformation due to defects in transcription, translation, post-translational modifications or protein folding [134]. Formal proof for the DRiPs hypothesis has not been provided yet [135,136], but strong support has been reported [133,137,138]. Two related studies showed that (1) blocking protein synthesis slowed the export of MHC class I molecules from the ER, indicating decreased supply of antigenic peptides [138] and (2) in acute influenza infection TAP becomes fully employed owing to the production and degradation of viral proteins [137]. Both studies indicate that an important proportion of MHC class I ligands are derived form newly synthesized proteins. The direct linkage of translation and antigen presentation would make perfect sense for immunity to acute virus infections, in which speed is of extreme importance to minimize viral replication [139]. Also because a DRiP itself has not yet been identified [140], refinements of the DRiP model, involving the existence of an immunoribosome, have been postulated [139,141]. An alternative model proposes that a subset of nascent polypeptides is stochastically delivered to the 20S proteasome owing to neglect by the protein folding machinery, which would also explain rapid peptide presentation [136]. In mature DC, in which antigen processing and presentation is optimized, DRiPs were found to be stored rapidly in intracellular aggregates that have been termed DALIS (for dendritic cell aggresome-like induced structures) [142]. DALIS only form when protein synthesis is ongoing and this is the place where DRiPs become ubiquitinated [143]. Thus, DC regulate the degradation of DRiPs by producing DALIS. Another special source of polypeptides for cytosolic proteolysis are the products of cryptic translation [144], which include peptides encoded by introns, intron/exon junctions, 5’- and 3’- untranslated regions, and alternate translational reading frames [145]. These polypeptides with no biological function may constitute a subcategory of DRiPs [144]. Several CTL epitopes have been demonstrated to result from cryptic translation [146-149].

4.4.4. Generation of class I ligands, overview
MHC class I antigen processing involves a process that can be divided in the three most defining events: proteolysis in the cytosol and ER, transport of peptides into the endoplasmic reticulum (ER), and the assembly of the class I peptide complex.
In principle all proteins that are tagged for destruction are prone to be degraded into single residues. Next to the first degradation step by the proteasome, cytosol-resident aminopeptidases and endopeptidases accomplish further protein degradation of proteasome-degradation products (length 2-22 aa, on average 7-9 aa [113]). Cytosolic aminopeptidases that have been implicated are puromycin sensitive aminopeptidase (PSA) [150], bleomycin hydrolase (BH) [150] and leucine aminopeptidase (LAP) [151]. Endopeptidases reportedly acting on proteasomal products and contributing to protein degradation are tripeptidyl peptidase II
(TPPII) [152] and thimet oligopeptidase (TOP) [153-155]. Various other cytosolic (endo)peptidases, like nardilysin [156], neurolysin [157], insulin degrading enzyme (IDE) [158], exist in the cytosol, but their involvement in protein degradation has not yet been demonstrated. The generation of immunogenic peptides presented by class I molecules can be considered as a by-product of protein degradation, because upon partial degradation a small proportion of intermediate degradation products (length on average 8–16 aa) escape from very rapid destruction in the cytosol [159] by transfer into the ER where class I loading takes place. This is accomplished by the transporter associated with antigen processing (TAP), a heterodimer that translocates peptides (on average 8–16 aa in length [160]) in an ATP-dependent fashion (reviewed in ref. 161). In the ER, peptides can be N-terminally trimmed by the ER-resident aminopeptidase ERAP1/ERAP2 [162,163] and, dependent on their binding affinity, assemble with MHC class I heavy chain and light chain (β2-microglobulin) in a complex folding mechanism assisted by the chaperones calnexin, calreticulin and Erp57 and the accessory protein tapasin (reviewed in ref. 164). Peptides need to fulfill the specific binding requirements of the class I molecule to which they bind. Seminal studies in the 1980s learned that residues in the peptide function as anchors, of which the side chains bind in pockets of the class I binding groove, enabling peptides to bind with high affinity to the class I molecule [4]. The combination of primary anchor residues (mostly at position two and the C-terminal position for the human class I molecules) and secondary anchors in the peptide that are required for efficient binding is defined in the peptide binding motif for each MHC/HLA class I molecule, which also allows the in silico prediction of peptide binding for any peptide with the appropriate length [165].

The extensively polymorphic HLA class I molecules – each individual expresses up to six HLA class I molecules, of which hundreds of variants are known [166]) – can be grouped in several HLA class I supertypes [167] with overlapping binding motifs, but each molecule may have its own fine specificity.

4.4.5. Generation of class I ligands, post-proteasomal processing

For class I binding, peptides need a defined length (8–12 aa, mostly 9 or 10 aa) with anchors at appropriate positions in the peptide. Seminal studies have shown the absence of C-terminal excision of model CTL epitopes in proteasome-inhibited cells [168-170]. Together with a failure to detect C-terminal trimming activities in the cytosol or ER [151,159], this has led to the current notion that the proteasome liberates the exact C-terminus of the vast majority of class I presented peptides [144,171]. Although some CTL epitopes are directly made by the proteasome, a significant fraction of class I ligands is made as N-terminally extended variant by the proteasome. The generation of the amino-terminus of these class I ligands is accomplished by aminopeptidases that reside either in the cytosol (PSA, BH, and LAP) or the ER (ERAP1/ERAP2). Redundancy in the function of these N-terminal trimming enzymes occurs [172]. For instance, the SI-INFEKL epitope from ovalbumin (OVA) that was dependent on LAP in one study [151] was normally presented in LAP knock-out mice [172]. It has been shown in knockout mice that especially the trimming in the ER by ERAP1/ERAP2 is required for the presentation of many class I ligands [173-175], although other peptides can be (partially) destroyed by ERAP activity [175,176]. Contradictory
results have been reported as to which extent ERAAP/ERAAP1 influences the anti-viral CD8+ T cell immunohierarchies [175,177]. From a mechanistic viewpoint, peptides in the ER may be (partially) protected from destruction by ERAAP/ERAAP1 through their binding to class I molecules [178].

Cytosolic endopeptidases of which the role in class I antigen processing has been studied are TPPII and TOP. TPPII is a very large homooligomer of 5–6 MDa consisting of subunits of ~138 kDa that are organized as two stacks of 10 dimers each that form a twisted, spindle shape structure [103,179,180]. It was first known for its (tripeptidyl) aminopeptidase activity removing tripeptides from the substrate’s N-terminus [180]. Indeed, N-terminal liberation of a CTL epitope (from RU) was reported to depend on TPPII (in conjunction with PSA) [181]. It was found that TPPII can substitute for partially impaired proteasome function when cells are cultured under prolonged periods with proteasome inhibitors [182-185]. TPPII was identified to exert a relatively low endoproteolytic activity of the trypsin-like type next to its amino-peptidase activity [183]. Accordingly, a CTL epitope from HIV Nef was found to be produced both at its N-terminus and its C-terminus in an endoproteolytic manner by TPPII [186], and a CTL epitope from influenza virus nucleoprotein was likewise suggested to be TPPII-dependent [187,188]. The HIV Nef epitope is the first epitope of which the C-terminus is known to be produced independently of the proteasome. As TPPII was found to be responsible for the degradation of the majority of cytosolic polypeptides (> 15 aa) in vivo [152], it was suggested that this enzyme may be necessary for the post-proteasomal generation of many class I ligands. However, subsequent studies [189-191] showed that in general TPPII seems not to be required for the (C-terminal) generation of CTL epitopes (reviewed in ref. 192).

TOP is a ubiquitously expressed cytosolic metallopeptidase of which the crystal structure revealed a deep substrate-binding channel [193]. TOP has great flexibility in substrate recognition [194,195], and prefers to release 3-5 residues from the C-terminus of substrates with a preferred length of 6–17 aa [196]. In principle, this endows TOP with the capacity to either destroy or generate class I ligands dependent on the specific substrate (either a minimal epitope or a C-terminal extended epitope-precursor). A positive effect of TOP overexpression on the presentation of a specific CTL epitope (from hsp65 M. tuberculosis) has been reported [197]. Most studies, however, demonstrate a destructive effect of TOP on the production of class I presented peptides [154,155,198]. Overexpression of TOP (5 to 16-fold the physiological level) reduced total class I expression, and RNAi-mediated silencing of TOP (modestly) enhanced class I expression [155]. At the level of defined epitopes, destruction by TOP was only shown for one epitope (namely SIINFEKL) [155]. A recent study, again points to the possible role of TOP in both antigen destruction and antigen generation, because, using a biochemical approach, in the cytosol both the substrates and products of TOP were demonstrated to include peptides of the length of class I ligands [199].

4.4.6. Major unresolved questions in class I antigen processing

Several important issues in class I antigen processing of endogenous proteins are still considerably unknown and debated. First, as discussed before, the precise origin of class I presented peptides has not yet been unravelled. Although it is apparent that rapid presentation of epitopes derived from long lived
proteins occurs, it has not yet been established that DRIPs [140] and/or immunoribosomes [135] really exist and are the major source of class I ligands. Alternative mechanisms may account for rapid antigenic presentation by class I molecules [136].

Second, although the major role of the proteasome in antigen processing is undisputed for the majority of class I ligands, a vast number of studies have indicated that there might be a significant fraction of class I-peptides that is generated independently – or partially independently – of the proteasome by other endopeptidases. Benham et al. [200] observed that proteasome inhibitor insensitivity was allele-specific. In particular HLA-A3 and -A1 matured efficiently, whereas other alleles tested were not resistant to inhibitor treatment, suggesting that a non-proteasomal protease or peptidase may preferentially generate peptides with basic C-termini binding in the acidic F-pocket of the HLA-A3/A1-binding groove. Likewise, Luckey et al. observed a broad resistance to proteasome inhibition of cell surface expression of 13 human class I alleles [201].

They also observed relatively high levels of re-expression of class I molecules accommodating basic C-termini (HLA-A3, -A68, -B705), but several other alleles, in particular those that bind a broad array of C-termini, displayed the same behaviour. In another study, comparison by mass spectrometry of the cell surface HLA-B2705-displayed peptide-repertoire under conditions with and without proteasome inhibition revealed that the repertoire was mainly unaffected and demonstrated the complete range of HLA-B2705 binding C-termini (including basic aa) suggesting a role for at least one non-proteasomal (endo)peptidase with a broad range of specificities [202]. It was shown that this peptidase is not TPII [191].

Proteasome inhibition has been shown to have indifferent effects on the presentation of defined epitopes [187,188,203]. Treatment of cells with proteasome inhibitors even led to enhanced presentation of several defined epitopes [204-208]. Taken together, these results strongly suggest the existence of endoproteolytic activities that complement the proteasome in the generation of the C-terminus of class I ligands. However, by using as primary tool proteasome inhibitors that are known to be leaky, especially for the tryptic-activity of the proteasome [114], these studies do not proof the existence of proteasome-independent generation of class I ligands. Only one defined epitope from HIV Nef was found to be made in a non-proteasomal manner by TPII [186]. However, subsequent studies rendered a broad role of this enzyme in the generation of class I ligands unlikely [192]. Thus, the extent of non-proteasomal processing and the identity of the alternative endopeptidases that are involved remain to be explored.

A third open issue is the extent to which cytosolic peptides are protected by chaperones from rapid destruction in the cytosol [144,209]. The group of Shastri has shown that post-proteasomal N-terminally extended variants of the OVA SIINFEKL epitope bind to the chaperone TRiC [210], which may increase the efficiency of presentation. In a more recent study, they showed that both N-terminal and C-terminal pre-proteasomal processing-intermediates from the same SIINFEKL-epitope are associated with the hsp90α chaperone [211]. These findings raise important questions. Because only one model CTL epitope was studied, the prevalence of the association of processing-intermediates with chaperones is not clear, and the nature of the pre-proteasomal intermediates is still unresolved as well [209].
5. Immunity to cancer and tumor immunooediting

From a historical viewpoint, the study of immunity to malignancies is deeply rooted in the treatment of tumors [212]. Based on a single observation in a patient that recovered from sarcoma after he had developed severe erysipelas, at the end of the 19th century, the New York surgeon William Coley started to treat cancer patients with bacterial vaccines resulting in sporadic regressions [213]. It was only much later that the underlying mechanisms, inducing innate immunity, were explained at the molecular level by the discovery of bacterial endotoxins [214] and tumor necrosis factor [214]. In the 1950s, studies of chemical induced tumors in syngeneic mice [215] definitively indicated the existence of tumor-specific antigens and, thus, recognisability of tumors by adaptive immunity. This raised hope for a well-grounded immunotherapy of cancer. These findings also instigated Burnet [216,217] and Thomas [218] to independently propose the theory of immunosurveillance of cancer, speculating that spontaneously arising cancer cells are often destroyed or kept in check by the immune system. A temporary setback again in the tumor immunology field was caused in the mid-1970s by a study of Hewitt and colleagues that showed absence of immunogenicity of spontaneously arising tumors [219]. However, several years later the group of Boon found that spontaneous murine leukemia cells possessed weak antigens that only led to rejection after the immune system was challenged with related more immunogenic tumor cells [220]. Based on this work and allowed by new insights in the basics of antigen presentation [1,2], in the early 1990s the same group identified by a laborious genetic approach the first mouse [221] and human [5,6] tumor antigens and their encoded CTL epitopes. Since then, the identification of tumor-specific antigens and insights in immunity to cancer in general has progressed tremendously. Current knowledge tells that by the time cancer is clinically detectable, it likely has already been adapted to the host immune recognition and attack, so that it effectively evades any immune response. The concept of immunosurveillance – which has been doubted for a long while because no differences in tumor development were found between athymic nude mice (that later were found to not completely lack functional T cells) and syngeneic wild-type mice [222,223] – has been substantiated only recently [224,225]. For instance, a deficiency in IFNγ enhanced host susceptibility to both chemically induced and spontaneous tumors [226]. In another mouse study a genetic trait was serendipitously found that conferred resistance to a highly aggressive sarcoma cell line [227]. This was dependent on innate immunity infiltrates of natural killer cells, macrophages, and neutrophils that independently killed the tumor cells [228]. Immunosurveillance is now seen as a phase
in a broader evolutionary process of the tumor as reaction to immune pressure, called cancer immunoediting [229]. Immunoediting ranges from tumor recognition and elimination (through immunosurveillance) to tumor sculpting (by immunoselection) and escape. Both innate and adaptive cellular immunity takes part in tumor suppression and tumor shaping in a complex process influenced by multiple variables such as the tumor’s type, anatomic location, stromal response, cytokine profile and inherent immunogenicity [224]. In immunoediting, before escape, an equilibrium phase is envisioned in which occult tumors are kept in check by the immune system during a period of latency, previously also called tumor dormancy [230]. A study reporting the occurrence of metastatic melanoma in two allograft recipients that had received kidneys from the same donor who had suffered from primary melanoma 16 years before her death, indeed strongly suggests an apparent equilibrium phase in the donor [231].

Often – it is not known how often – tumors escape from naturally induced immune pressure. Obviously, this also happens under circumstances of non-optimal therapeutically induced immunity. Numerous immune evasion mechanisms, all affecting the interplay of tumor and immune system, are known that contribute to escape from natural or therapeutically induced anti-cancer immunity (reviewed in ref. 232).

The tumor’s inherent low capacity to appropriately stimulate the immune system, primarily caused by the absence of a proinflammatory environment and costimulatory context [233], but also often by a low density of (cancer-specific) MHC-peptide complexes, will lead to T cell ignorance, anergy or deletion, together called peripheral tolerance [234].

Cross-presentation of tumor antigens and subsequent T cell activation, needed to induce robust tumor immunity, will fail when a too low number of tumor cells are dying (either by apoptosis [235] or necrosis) or when DC are not appropriately matured [236]. Although (dying) tumor cells may sometimes inherently express danger signals, such as uric acid [237], it has now become increasingly clear that mostly the tumor microenvironment both actively (e.g. by secretion of TGF-β and IL-10) and passively suppresses the induction of tumor immunity (reviewed in ref. 238). Especially the maturation of DC is mostly lacking or incomplete in the tumor environment [239,240]. Stimulation by immature DC will lead to T cell tolerization [236,241] and may induce regulatory T cells that often play an immune suppressive role in cancer immunity [29,242]. Another mode of suppression is accomplished by tumors that express high levels of PD-L1 interacting with inhibitory B7-family member PD-1 on activated and exhausted T cells [243]. Moreover, tumor sculpting, either caused by natural immunity or by therapeutically induced immunity, may result in loss of tumor antigens [232,244] and possibly even antigenic drift [245], a mechanism common in viral immunity. Finally, lesions in molecules of antigen processing and presentation pathways, such as class I and TAP downregulation, often occur in tumors, highly likely as a result of immune pressure [105,246].

6. T cell mediated immunotherapy for cancer, modalities, and basic requirements

Immunotherapy of cancer by T cells can be divided in passive adoptive T cell transfer and active immunostimulatory vaccination strate-
6.1. Adoptive transfer of undefined tumor-specific or defined epitope-specific T cells

The only routine immunotherapy for cancer in the clinic to date is the infusion of donor lymphocytes after allogeneic stem cell transplantation in leukemia. This therapy is curative in significant percentages of patients [10]. The broad donor-derived CD4+ and CD8+ T cell repertoire targeting a diversity of undefined (allogeneic) leukemia antigens is exploited in this setting. Remarkable clinical responses were observed in metastatic melanoma patients after adoptive transfer of autologous tumor-specific infiltrating lymphocytes (TIL) that were ex vivo expanded to high numbers [250]. The non-myeloablative conditioning regime in this trial may have contributed to the further expansion in vivo of the adoptively transferred T cells, by making space and also by the depletion of negative regulatory CD4+ T cells or so-called myeloid derived suppressor cells [251,252]. Furthermore, the CD4+ T cell component in the transferred TILs has likely helped the memory CD8+ T cell population [250]. Because it is often hard to expand tumor-specific CTL at high numbers ex vivo, efforts are undertaken to introduce tumor-epitope specific T cell receptors (TCR) in peripheral blood lymphocytes (PBL) of the patient [253]. Promising clinical results were reported in patients with metastatic melanoma who were given autologous PBL retrovirally transduced with the TCR specific for the well known MART-1(27-35) HLA-A2-presented epitope [254]. TCR gene transfer has the advantage that the problem of expanding enough tumor-specific T cells is bypassed. However, targeting a single epitope may lead to antigen loss variants. Therefore, adoptive transfer of PBL transduced with multiple ‘off the shelf’ TCRs targeting CTL epitopes in different TAA is a logical and promising next step.

6.2. Vaccination strategies with undefined antigens

Irradiated autologous tumor cells or allogeneic HLA-matched tumor cell lines that are modified to express GM-CSF, IL-2 and other cytokines or costimulatory molecules have been used as vaccines. In various clinical trials this type of vaccine has induced immune responses [255] and clinical responses have been reported [256-259]. However, several disadvantages are connected to this strategy like the suboptimal direct antigen presenting capacity of tumor cells, absence of HLA class II presentation, uncertain cross-presentation, and often the lack of autologous tumor samples needed for preparation of the vaccine. Other forms of vaccination with the full potential of undefined antigens from the targeted tumor are tumor lysates (loaded on DC [260]), heat shock proteins (HSPs) derived from the tumor and DC transfected with amplified tumor mRNA [261,262]. The main advantage of vaccination with autologous tumor cells, or tumor derived lysates, HSPs or mRNA, is the presence of the full undefined repertoire of relevant tumor antigens, including those with mutations that are unique in the individual tumor. In this sense the strategies applying autologous tumor material are all personalized non-standardized vaccines that have to be produced for each patient separately. These therapies aim to induce T cell responses against as much as possible (undefined) tumor-specific HLA class I (and in certain settings HLA class II) presented peptides.
6.3. Vaccination strategies with defined full length tumor associated antigens
Vaccinations with recombinant viral vectors or naked DNA plasmids encoding defined full length tumor associated antigens and vaccination with recombinant tumor proteins themselves have been applied in vaccines aiming to raise humoral and T cell responses against the tumor expressing the antigen. Likewise, DC electroporated with mRNA encoding full length TAA are currently being optimized for clinical testing [263]. Vaccination strategies aiming to raise immunity to a full length antigen have the advantage that the HLA haplotype of the individual patient does not need to be considered. On the other hand, and apart from the problems related to each mode of delivery (virus, DNA, mRNA, protein; reviewed in refs. 247 and 248), vaccination with single whole antigens has the important drawback that vaccine induced immune pressure may induce escape through antigen loss variants of the tumor. In principle this could be circumvented by vaccination with multiple full length defined antigens (either in the form of DNA, mRNA [263] or protein).

6.4. Vaccination strategies with defined T cell epitope containing synthetic peptides
Since the first identification of a defined tumor-specific CTL cell epitope [6], the concept of immunizing cancer patients with synthetic peptide epitopes has been elaborated. Numerous clinical peptide vaccine trials have been conducted with sometimes promising results. The relatively poor immunogenicity of peptides per se requires them to be injected either together with adjuvants or loaded on DC (reviewed in refs. 264 and 265). Further optimization of the peptide vaccination strategy is envisaged [266]. It is now firmly established that for robust and persistent CD8+ T cell responses a concomitant CD4+ T helper response is needed [52,267-269]. Therefore, HLA class II presented tumor-specific epitopes are preferably incorporated in peptide vaccines to promote the CTL mediated tumor destruction. Important advantages of peptide vaccination are its defined nature and the easy manner to synthesize peptides by good manufacturing practice (GMP), enabling peptide vaccines to be used as pre-fabricated ‘off the shelf’ vaccines. Furthermore, modifications aiming at increasing the immunostimulatory context of the vaccine – like conjugation with synthetic Toll-like receptor (TLR) ligands [270] – can easily be accomplished. Immunizations with a single (or only a few) CTL epitope(s) may induce outgrowth of antigen loss variants of the tumor. Therefore, peptide vaccines should preferably contain multiple HLA class I presented CTL epitopes derived from different target antigens together with a tumor-specific HLA class II presented CD4+ T helper epitope. The use of longer (e.g. 30-mer) epitope-containing vaccine peptides that require processing which can only be accomplished efficiently by professional antigen presenting cells (DC) has been shown beneficial [271-273].

7. Tumor associated antigens and their classification
For immunotherapeutic purposes the most important criteria to classify tumor associated antigens (TAA) are: (1) broadness of expression (shared between patients and/or cancer types), (2) tumor specificity (absence of expression in healthy tissues) and (3) the function of the TAA in the oncogenic process and/or cancer survival. Additionally, (4) possible changes in turnover kinetics of the TAA are important to consider [274], as e.g. in the case of p53 [275].
With respect to breadth of expression, there is a first rough division in unique tumor antigens that are restricted to only an individual tumor in one patient – which for obvious reasons restricts their immunotherapeutic applicability – and the antigens that are shared between cancer patients. When combined with the criterion of tumor-specificity, this results in the following often used classification. Unique tumor-specific antigens are resulting from mutations occurring in a single tumor of one patient. The first example of a unique point-mutation was found in the melanoma-associated mutated antigen-1 (MUM1) gene [276] (for more examples, see listing in ref. 277).

Shared lineage-specific differentiation antigens are expressed in both the tumor and its original healthy tissue. Examples are the melanoma/melanocyte antigens (MART-1/Melan-A, gp100, tyrosinase, TRP2) and prostate antigens (PSA, kallikrein 4).

Shared tumor-specific antigens are expressed in different tumors but not in healthy tissues. The most prominent among these TAA is the group of so-called cancer-testis antigens like the MAGE, BAGE and GAGE families and NY-ESO-1, which in normal tissues are only expressed in testis and/or placental tissues. Further examples are viral oncoproteins (e.g. HPV16 E6 and E7) and the fusion-proteins encoded by translocated genes (e.g. BCR-ABL).

Shared antigens overexpressed in tumors are formally not tumor-specific but have a much higher expression level in tumors. Often these TAA are widely expressed in different cancer types, like hTERT, survivin and PRAME. Others, like carcinoembryonic antigen (CEA) and MUC1, own a more restricted expression pattern. A special case in this category of TAA is p53 because this oncoprotein is mutated in a variety of tumors and apart from being overexpressed can also show enhanced turnover, rendering it possibly applicable for immunotherapy [275]. Moreover because of its rapid degradation in normal cells, there appears to be no tolerance of p53 at the level of CD4+ T cells [278].

Some antigens can be positioned in between two of the categories; e.g. PRAME is widely expressed in various cancer types and, in contrast, in healthy tissues only at very low levels in adrenals, ovaries and endometrium, next to its expression in testis and placenta [279]. Further extensive listings of TAAs can be found in the literature [280] or in databases (e.g. at www.cancerimmunity.org).

**7.1. Strategies for the identification of TAA**

Identification of TAA can be accomplished with different experimental strategies [281-284]. The discovery of MAGE-1 [5] in the early 1990’s as the gene encoding the first tumor-specific CTL epitope [6] is one of the pillars of tumor immunology. An autologous melanoma specific CTL line was used to find the tumor specific cDNA that encodes the recognized CTL epitope from a cDNA library derived from the melanoma. Subsequently, the minimal CTL epitope was identified by cDNA truncation and peptide recognition techniques. This classical strategy of expression profiling, which is often revered to as ‘direct immunology’ because it is based on natural immunity, has since then been applied for the identification of (among others) the MAGE, BAGE and GAGE families [285,286], Melan-A/MART-1 [287,288], tyrosinase [289] and gp100 [290]. In a biochemical strategy, the CTL clone can also be used to identify the HPLC-fraction of peptides isolated from the tumor cell surface that contain the epitope. Subsequently, mass spectrometry can identify the precise epitope sequence, and databank searches may lead to the identification of novel TAA [285-290].
A key characteristic of both strategies is the use of an autologous tumor specific CTL as the selection tool. The unknown tumor specific CTL epitope is used as handle to identify the source protein, and, therefore, tumor protein discovery and T cell epitope identification are intertwined in these strategies. The serological identification of antigens by recombinant expression cloning (SEREX) strategy defines putative tumor antigens using patient-derived serum IgG antibodies to screen proteins expressed from tumor-derived cDNA libraries [291]. Tumor antigens identified by SEREX will likely contain CD4+ T helper cell epitopes because isotype switching from IgM to IgG implies the presence of specific help from CD4+ T cells. The cancer-testis antigen NY-ESO-1 [292] is only one example of a large array of (putative) tumor associated antigens that were identified by SEREX methodology [282].

With the rise of genomics and in silico data mining techniques, transcriptome analysis is currently used to detect tumor-specific expression profiles directly at the genetic level without the need for patient derived T cells or serum (in more detail reviewed in ref. 293). Various methods are used, like classical mRNA/cDNA subtraction techniques, representational differential analysis (RDA) [294,295], differential PCR display and comparison of cDNA profiles obtained by serial analysis of gene expression (SAGE) [296], DNAChip/microarray analysis [297-299] and expressed sequence tag (EST) databases. These studies often aim to identify expression profiles that can be used for the improved diagnosis [299], classification [297,298] or prognosis [300] of cancer. Tumor-specific expression of identified transcripts has to be confirmed at the protein level before immunogenicity studies are planned. The advent of the complete human genome sequence has enabled recent studies identifying mutational profiles in various cancers [301,302] that reveal the number and uniqueness of cancer-related mutations (within and between classes of cancer). Importantly, each mutation may give rise to (unique) cancer-specific T cell epitopes [303].

7.2. Selection of TAA for T cell immunotherapy

Which tumor associated antigens are most suitable as immunotherapeutic target? And how many TAA should be targeted simultaneously? These important questions are still being debated. First, the ideal tumor antigen target is widely expressed in different tumor types, enabling ‘off the shelf’ vaccines that are applicable in broad patient populations. Secondly, the function of the targeted TAA in the oncogenic process is highly relevant, although for several TAA not yet known. The phenomenon of immune escape by selection of antigen loss variants of the tumor is far beyond only theoretical consideration [304-306]. Therefore, TAA that either play a role in the oncogenic process or promote cancer cell survival are favourable targets. In this respect, the lineage specific differentiation antigens are lower ranked tumor antigens than purely oncogenic proteins like the HPV16 derived E6 and E7 proteins and the BCR-ABL fusion protein [307,308]. Overexpressed anti-apoptotic proteins like survivin are interesting because down regulation or loss of such TAA would severely impede the growth potential of the tumor cell [309,310]. Likewise, the telomerase catalytic subunit (hTERT) is involved in the pathogenic process [311] and has a reported anti-apoptotic role [312]. Another tumor antigen for which a role in tumorigenesis and metastasis has been reported is PRAME [300,313-315]. These tumor antigens (survivin, hTERT
and PRAME) are widely expressed in different tumor types and constitute also for this reason attractive tumor antigens. However, it can not be excluded that therapy-induced immune pressure may give rise to selection of antigen loss variants, even of anti-apoptotic or tumor promoting proteins, as has been observed in patients with melanomas expressing the melanoma inhibitor of apoptosis protein (ML-IAP) [255]. A third important consideration is the immunogenicity of the targeted tumor antigen. Tolerance to the non-mutated lineage specific differentiation antigens (like gp100, tyrosinase and MART-1/Melan-A), which are self proteins, may severely hamper an effective immune response against these antigens [316,317]. Such tolerance is likely affecting the immunodominant epitopes more than sub-dominant T cell epitopes, which is a reason why the latter category of epitopes in these differentiation antigens has attracted considerable attention (see below). To circumvent both the selection of antigen loss variants and the tolerance to differentiation antigens, targeting of multiple antigens by polyvalent vaccines (or multi-specific adoptive transfer) is essential. An additional advantage is that the full potential of the anti-tumor response in the patient is better exploited. The rule here would be: targeting more antigens is better. Lately, driven by these problems, different groups have made an argument in favour of personalized immunotherapy targeting the unique antigens caused by mutations [301] that are often only present in the tumor of one patient [303,318,319]. These tumor antigens are purely tumor specific, and therefore not tolerogenic, and are believed to be often crucial to the oncogenic process [277,303]. Furthermore, the natural immune response in some patients was found to be stronger against the unique antigens than the response against shared antigens [320]. Immunotherapies against non-defined tumor antigens such as vaccination with DC transfected with tumor-derived mRNA [262] or tumor-lysate pulsed DC are in fact personalized therapies that target both the shared and the unique antigens of each patient. However, the application of a defined patient-tailored immunotherapy will require identification of the unique antigens at the epitope level separately for each patient and meets with tremendous technical and logistic difficulties [318].

8. Identification of tumor-specific T cell epitopes

To date, a total of 180 HLA class I restricted CTL epitopes and 75 HLA class II restricted T helper epitopes in shared tumor associated antigens have been reported (according to the listing at www.cancerimmunity.org; update September 2006). Although this number of T cell epitopes seems a reasonable starting point for the design of defined immunotherapeutic vaccines, there is strong skewing to epitopes derived from antigens expressed primarily in melanoma and 75 (42%) of the HLA class I epitopes are presented in HLA-A2, leaving epitopes restricted by other HLA class I alleles severely underrepresented. This severely hampers the design and development of defined epitope-based vaccines targeting other tumors than melanoma, especially in patients lacking HLA-A2. Furthermore, the identification of HLA class II peptides recognized by T helper cells, which are indispensable as help to mount efficient CD8+ effector T cell responses [52,267–269], has lagged behind (Fig. 2).
8.1. Identification of CTL epitopes starting with CTL of unknown specificity

The discovery of CTL epitopes has proceeded along two different experimental lines: either starting with a pre-existing CTL clone with unknown specificity (direct immunology), or departing from a predicted epitope (reverse immunology). In the first years all epitopes were identified by the direct immunology approach of expression cloning (Fig. 2). A patient-derived autologous tumor-specific CTL clone recognizing an unknown epitope was used to screen a tumor derived cDNA library (mostly from melanoma) which is expressed in antigen-negative (tumor) cells. Subsequently, recognition of truncated variants of the epitope-encoding cDNA and mapping of synthetic peptides revealed the minimal epitope sequence. In this procedure, next to the unknown CTL epitope, the equally unknown source tumor antigen was often discovered together with the epitope (see above) [285-290]. The major drawback of this laborious strategy is the dependence on autologous tumor-specific T cells that are either generated in mixed lymphocyte tumor cultures (MLTC) or obtained as tumor infiltrated lymphocytes (TIL). Such T cell responses are generally scarce and the induction in MLTC is dependent on the availability of autologous tumor cell lines, which have been mainly obtained from melanomas. Furthermore, CTL responses from MLTC or TIL are per definition directed to immunodominant epitopes. In recent years, the direct approach has been adapted for the identification of epitopes in known antigens without the need for autologous tumor cell lines. As elaborated by Chaux et al., TAA artificially expressed in DC were used for the generation of autologous CTL clones specifically recognizing unknown epitopes derived from the transduced antigen. Peptide-mapping experiments then again revealed the exact epitope sequence [321]. By virtue of the natural CTL response that is used this method as well will result in the identification of mostly immunodominant epitopes, and a systematic search for novel epitopes is impossible. An alternative biochemical approach for defining the unknown specificity of tumor-reactive CTL, which are either induced against tumor cells [322,323] or e.g. against peptides eluted from tumor cells [324], starts with the immuno-affinity purification of the HLA class I – peptide complexes from the relevant tumor cell. The peptides are subsequently isolated and fractionated by (multiple rounds of) high-performance liquid chromatography (HPLC) to reduce the complexity of the peptide pool. Pinpointing the fraction that contained the epitope through recognition by the CTL together with tandem mass spectrometry (MS/MS) mediated sequencing of the peptides in that fraction then identifies the precise peptide sequence [322-327]. Subsequently, database searches may identify the unknown source antigen [324-327]. A particular advantage of this strategy is that it may identify post-translationally modified epitopes [328,329] (or special epitopes generated by protein/peptide splicing [283,284,330]). Still another way to analyse the specificity of pre-existing CTL clones is the application of synthetic peptide libraries to search for reactive mimicry epitopes. The natural epitope may subsequently be identified by screening recognition of substitution analogs, defining a recognition motif and database searching [331]. The combination of library screening-deduced T cell recognition motifs in the peptide and MS/MS sequencing of eluted peptides has also been exploited to identify a novel mouse CTL epitope [326].
8.2. Identification of CTL epitopes by reverse immunology

Since the first finding of HLA specific peptide binding motifs in the early 1990s [4], it is possible to screen known TAA for contained peptides that are predicted to be cell surface expressed. Predicted HLA class I ligands can be tested for their immunogenicity by raising CD8+ T cells against the exogenously loaded peptide. Subsequently, peptide-specific CTL are tested for their recognition of tumor cells expressing the relevant TAA and restriction element to prove the natural presentation of the CTL epitope. The basis of this indirect strategy for CTL epitope identification, which was coined ‘reverse immunology’, is that an initial epitope-prediction-phase is followed by an epitope-validation-phase (a flow scheme is presented in Fig. 3). Nowadays, approximately 40% of the CTL epitopes in shared tumor associated antigens (Fig. 2), and also numerous CTL epitopes in viral and microbial antigens, have been identified via the reverse immunology approach. The advantage of reverse immunology is that it is the only strategy that can be used to systematically search for novel epitopes, including subdominant ones, in known proteins and presented in any HLA molecule of interest. Both the prediction phase and the validation phase of reverse immunology have their own difficulties and weaknesses, although significant improvements have lately been implemented.

8.2.1. Prediction phase of reverse immunology

The prediction phase of the reverse immunol-
binding capacity. Various HLA class I binding algorithms have been developed of which the BIMAS algorithm [165] and the SYFPEITHI algorithm [334] are freely accessible and currently the most widely used (algorithms are listed in Table 2 and see refs. 335 and 336). The algorithms employ slightly different peptide binding motifs and different arithmetic methods, based on either the contribution to binding of each aa in a peptide independently or the overall peptide structure, but are all extremely valuable to select the small percentage of peptides with potential binding capacity [337]. Guidelines for validation and comparisons between the different algorithms have been made [337,338], and significant differences in the predictions often occur (exemplified in chapter 7). Therefore, from a practical point of view, the combination of two or more methods is advisable to reduce the number of non-selected peptides with binding capacity. Experimental verification of actual binding capacity is preferred, because the ranking of the predictions does not perfectly correlate with the actual binding measurements and false positive prediction of binding occurs (exemplified in chapter 7 and chapter 4 [339]).

HLA class I binding assays exist in various forms (briefly reviewed in chapter 4 and ref. 293), and can be divided on the one hand in cell free assays (using soluble HLA) versus cellular assays using HLA class I molecules on the cell surface (chapter 4 [339]), and on the other hand in competitive assays (chapter 4 [339]) (resulting in semi-quantitative data) versus assays that do not use a (labelled) reference peptide and are therefore quantitative. Next to verification of binding capacity, the stability of peptide binding can be measured. Highly stable peptides have been shown to be more immunogenic [340], because they allow a sustained interaction with the T cell.

Scope of the thesis, development of binding assays:
In chapter 4 [339], peptide binding assays were developed for 13 prevalent HLA class I molecules. Using B-LCL expressing the class I molecule of interest and a fluorescently-labeled class I ligand with proven high affinity as reference peptide to compete with, reliable competition-based cellular assay are now available for easy measurement of peptide binding.

8.2.2. Improved CTL epitope prediction by verification of proteasomal processing and TAP translocation
The tumor-specific CTL epitopes that were identified by reverse immunology in the first years (until 2001) were predicted by taking into account only the HLA class I peptide binding capacity [341-344]. However, it was observed that numerous CTL that were raised against high affinity binding peptides did not recognize tumor cells expressing the relevant TAA and restriction element [345-349]. A major reason for this was the lack of intracellular generation of predicted peptides by the processing machinery. Thus, a refining of the epitope prediction procedure was needed.

Scope of the thesis, improved prediction of CTL epitopes:
In chapter 2 [350], the in vitro proteasome-mediated excision of class I binding peptides from their flanking regions was incorporated in the epitope prediction procedure. This allowed the selection of four C-terminally liberated peptides from 19 high affinity HLA-A2 binding peptides in TAA PRAME. These four peptides were proven to be naturally presented epitopes. The other peptides were considered to be likely not produced intracellularly, avoiding laborious T cell induc-
tions against these peptides. The same strategy was followed in chapter 5 [307] to identify CTL epitopes in the fusion regions of the BCR-ABL fusion proteins expressed in CML and ALL. Putative CTL epitopes were found, of which one was proven to be expressed. Some published epitopes were made likely to be not expressed.

This optimization of the prediction procedure (chapter 2) has greatly enhanced the accuracy of epitope predictions and has since then been applied successfully in studies identifying CTL epitopes in various tumor associated antigens (see e.g. chapter 5 [307] and refs. 351-354) and autoimmune antigens [355-357].

Four computer algorithms, which are based on different computational methods, are currently freely available via the internet for the prediction of proteasomal cleavages: MAPPP/FragPredict [358], PAPoC [359], NetChop [360], and Pcleavage [361] (Table 2). Proteasomes cleave abundantly at certain sites and cleave much less abundant or do not cleave at other sites. However, due to the broad specificity of the proteasome, the stochastic nature of proteasomal digestion [112,360] (overlapping fragments are often found in the experimental systems [112,307,350]), and (partly) undefined influences of distant residues on cleavage efficiency, a qualitatively and quantitatively accurate prediction of proteasomal digestion sites is very complicated. In general, predictions may still result in a high number of improperly predicted cleavage sites [350]. Therefore, experimental determination of proteasome-mediated digestion is needed to reliably select peptides that are C-terminally liberated (see also below). An extra level of complexity here is the different forms in which proteasomes occur. Some CTL epitopes are preferentially made by immunoproteasomes [362], which are expressed in professional antigen presenting cells [128] and

![Prediction Phase Diagram]

In Silico predictions
- Proteas. proc. predic.
- Binding predic.
- TAP transloc. predic.

Experimental verifications
- Proteas. digestions
- Binding assays
- TAP transloc. assays

Predicted HLA class I ligands (CTL epitopes)

Validation Phase

Ligand confirmation by mass spectrometry
- Cell surface expression is tested by ‘Predict-calibrate-detect’ method

Proven HLA class I ligand

Epitope confirmation with CTL clone
- Immunogenicity and epitope expression is tested by T cell inductions and target recognition assays

Proven CTL epitope

Figure 3. Flow chart of the reverse immunology approach for CTL epitope identification.
contain variant catalytic subunits with slightly different catalytic activity [123], and other epitopes are preferentially made by constitutive proteasomes [127,363], although most epitopes are liberated by both types of proteasomes. To cover both categories of epitopes, predictions and experimental verifications should use (in silico and in vitro, respectively) both types of proteasomes. Instead of first performing HLA class I binding assays [350], the proteasomal digestion pattern can also be determined first [351]. This reflects the physiological mechanistic order and has the advantage that binding of only those peptides that are C-terminally liberated by a major cleavage site needs to be verified experimentally.

Translocation of peptides into the ER via TAP is also an important event in the class I antigen presentation pathway. However, the specificity of the TAP heterodimer for peptides is much less selective because peptides meant to bind in all possible HLA class I molecules should be translocated into the ER. Specificity of TAP even seems to have evolved to fit the specificity of the proteasome [364]. Despite that, differences in translocation efficiencies between peptides exist [160] and TAP affinity has impact on HLA class I presentation [365]. Thus reasoning, in silico TAP translocation prediction algorithms have been developed [366-368] (Table 2) to incorporate TAP translocation efficiency in the overall HLA class I ligand prediction. However, a problem related to determination of TAP translocation efficiencies (in silico or in vitro) is the amino-terminal trimming that can occur both in the cytosol and in the ER. Therefore it is not a priori known which peptides should be tested. In general, TAP translocation efficiency, either predicted or experimentally verified, has until now been incorporated in only few reverse immunology studies identifying novel CTL epitopes [356].

Recently, five integrated in silico CTL epitope prediction tools have been developed that combine predictions of proteasomal cleavages, TAP translocation and HLA class I binding [115,369-373] (see Table 2 and further discussed in chapter 7). In addition, an algorithm was developed that directly predicts CTL epitopes using large datasets of T cell epitopes and non-epitopes as training data for the algorithm (CTLPred) [374].

8.2.3. Validation phase of reverse immunology

In the validation phase of the reverse immunology approach, the natural presentation and immunogenicity of the putative epitope should be demonstrated. In principle, two roads are open (Fig. 3). First, the biochemical purification of HLA-peptide complexes – from cells expressing the relevant TAA and HLA class I molecule – followed by the mass spectrometric search for the predicted peptides in the eluted HLA class I bound ligands (the ‘predict-calibrate-detect’ method; see below) [375,376]. However, this method validates cell surface expression of predicted HLA class I ligands but not their immunogenicity. Therefore, in the vast majority of studies the prediction phase is directly followed by the induction of (naïve) T cells against the exogenously loaded predicted epitope. Peptide-specific T cells are then used as tool to test the natural presentation of the epitope. T cells have mostly been induced in vitro using human peripheral blood lymphocytes (PBL) from healthy donors [307,350]. Furthermore, PBL [353,377] or TIL [351] from patients with the relevant tumor antigen and restriction element have been used. An alternative approach is the induction of T cell responses in HLA class I transgenic mice [378]. Peptide specific T cells should be used at the clonal level to enhance specificity of the response and to reduce aspecific back-
ground recognition of target cells that lack tumor antigen or restriction element. Next to tumor cells expressing the relevant TAA and HLA class I molecule, target cells should preferably include transfected target pairs with or without the tumor antigen and lacking or expressing the relevant HLA molecule. This enables exclusion of aspecific recognition effects by T cell clones that may be cross-reactive to irrelevant antigens. The sensitivity of the CTL clone for the peptide should be high, which is to be determined with peptide titration, before a definitive judgement of the natural presentation of the epitope can be made. While CTL recognition of properly chosen target cells will prove the natural presentation of the predicted epitope, it should be noted that, in principle, sometimes a length variant of the predicted epitope (with comparable binding capacity) may be the actually recognized peptide. Is such a case suspected, then tandem mass spectrometry is required to determine the exact aa sequence of the epitope-variant after its isolation from the cell surface (Fig. 3).

8.3. Identification of HLA class II presented T helper epitopes

Like CTL epitope identification, the identification of HLA class II presented T helper cell epitopes can either start with a CD4+ T cell recognizing an unknown epitope or may depart from T helper cells which are induced against either predicted epitopes or a complete set of overlapping peptides in a reverse immunology setting (Fig. 2).

In vivo sensitized CD4+ T cells that recognize a tumor-antigen have been employed to identify epitopes by expression cloning [379,380]. Contrary to HLA class I ligands, peptides binding in HLA class II are relatively long and different length variants of a T helper epitope are often recognizable by a single T helper clone. This characteristic has often advantageously been exploited to screen the T helper cell reactivity against a complete set of overlapping peptides of a TAA using responder CD4+ T cells that were derived from either patients or healthy donors. Subsequently, peptide-specific T helper clones were tested for their recognition of endogenously processed TAA to validate the epitope [381-383]. The application of in silico algorithms for the prediction of HLA class II presented epitopes has lagged behind the use of predictions for CTL epitope identification. Binding requirements for peptides in HLA class II molecules are much less restricted than for class I molecules. Thus, the HLA class II binding motifs share a certain degree of degeneracy, and prediction of binding is less straightforward. Furthermore, the peptides bound in HLA class II have a broad length spectrum (9–25 aa) and class II antigen processing pathways are only incompletely defined making any assessment of processing uncertain. In addition to the classical endosomal-lysosomal pathway of exogenous and transmembrane proteins, alternate and partially overlapping routes for class II ligand generation exist in the cytosol [67]. Both proteasome-dependent and proteasome-independent cytosolic generation of class II ligands derived from either exogenous or endogenous sources have been reported [64-66,384,385]. Furthermore, HLA class II antigen processing can differ depending on the route of delivery of exogenous antigens [64] and the precise antigen presenting cells in which it takes place [66]. Despite this complexity in the processing, several HLA class II binding algorithms have been developed [334,386-388] (listed in Table 2). The degeneracy of HLA class II binding motifs allows the search and prediction of promiscuous pan-class II binding peptides, which are obviously more widely applicable
Table 2. Web-based algorithms for prediction of HLA class I and class II ligands and ligand/epitope databases.

<table>
<thead>
<tr>
<th>Name</th>
<th>URL</th>
<th>Possibilities / additional information</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLA Peptide Binding Prediction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIMAS</td>
<td><a href="http://bimas.dct.nih.gov/molbio/hla_bind">http://bimas.dct.nih.gov/molbio/hla_bind</a></td>
<td>MHC class I</td>
</tr>
<tr>
<td>SYFPEITHI</td>
<td><a href="http://www.imtech.res.in/raghava/syfpeithi">http://www.imtech.res.in/raghava/syfpeithi</a></td>
<td>MHC class I and II</td>
</tr>
<tr>
<td>NetMHC</td>
<td><a href="http://www.cbs.dtu.dk/services/NetMHC">http://www.cbs.dtu.dk/services/NetMHC</a></td>
<td>MHC class I</td>
</tr>
<tr>
<td>PREDEP</td>
<td><a href="http://margalit.huji.ac.il/">http://margalit.huji.ac.il/</a></td>
<td>MHC class I</td>
</tr>
<tr>
<td>ProPred-I</td>
<td><a href="http://www.imtech.res.in/raghava/propred1/index.html">http://www.imtech.res.in/raghava/propred1/index.html</a></td>
<td>MHC class I (with proteasome cleavage filter)</td>
</tr>
<tr>
<td>rhHLA-Pred</td>
<td><a href="http://www.imtech.res.in/raghava/rhlapred/">http://www.imtech.res.in/raghava/rhlapred/</a></td>
<td>MHC class I (two methods)</td>
</tr>
<tr>
<td>IEDB (HLA class I)</td>
<td><a href="http://tools.imtech.res.in/analyze/html/mhc_binding.html">http://tools.imtech.res.in/analyze/html/mhc_binding.html</a></td>
<td>HLA class I (different methods)</td>
</tr>
<tr>
<td>HLA-A2 (No name)</td>
<td><a href="http://zlab.edu/smm/">http://zlab.edu/smm/</a></td>
<td>HLA-A2, -A3 and -A2 supertypes</td>
</tr>
<tr>
<td>Multipred</td>
<td><a href="http://research.i2r.a-star.edu.sg/multipred/">http://research.i2r.a-star.edu.sg/multipred/</a></td>
<td>HLA-A2, -A3 and -A2 supertypes</td>
</tr>
<tr>
<td>MHCPred</td>
<td><a href="http://www.jenner.ac.uk/mhcppred">http://www.jenner.ac.uk/mhcppred</a></td>
<td>HLA class I and II (and TAP)</td>
</tr>
<tr>
<td>IEDB (MHC class II)</td>
<td><a href="http://tools.imtech.res.in/tools/matrix/iedb_input?matrixClass=II">http://tools.imtech.res.in/tools/matrix/iedb_input?matrixClass=II</a></td>
<td>MHC class II</td>
</tr>
<tr>
<td>MHC-BPS</td>
<td><a href="http://bidd.cz3.nus.edu.sg/mhc/">http://bidd.cz3.nus.edu.sg/mhc/</a></td>
<td>HLA class I and II (flexible length)</td>
</tr>
<tr>
<td>SVMHC</td>
<td><a href="http://www-bs.informatik.uni-tuebingen.de/SVMHC/">http://www-bs.informatik.uni-tuebingen.de/SVMHC/</a></td>
<td>MHC class I and HLA class II</td>
</tr>
<tr>
<td>SVRMHC</td>
<td><a href="http://svrmhc.unm.edu/SVRMHC/default.html">http://svrmhc.unm.edu/SVRMHC/default.html</a></td>
<td>HLA class I and class II</td>
</tr>
<tr>
<td>ProPred (class II)</td>
<td><a href="http://www.imtech.res.in/raghava/propred">http://www.imtech.res.in/raghava/propred</a></td>
<td>HLA class II</td>
</tr>
<tr>
<td>MHCP2Pred</td>
<td><a href="http://www.imtech.res.in/raghava/mhcppred/">http://www.imtech.res.in/raghava/mhcppred/</a></td>
<td>MHC class II (promiscuous binding)</td>
</tr>
<tr>
<td>Proteasomal Cleavage sites Prediction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NetChop</td>
<td><a href="http://www.cbs.dtu.dk/services/NetChop">http://www.cbs.dtu.dk/services/NetChop</a></td>
<td>Based on artificial neural network</td>
</tr>
<tr>
<td>PAProc</td>
<td><a href="http://paproc.de">http://paproc.de</a></td>
<td>Based on evolutionary algorithm</td>
</tr>
<tr>
<td>MAPPP (FragPred)</td>
<td><a href="http://www.mpib-berlin.mpg.de/MPAPP/cleavage.html">http://www.mpib-berlin.mpg.de/MPAPP/cleavage.html</a></td>
<td>Based on peptide cleavage data</td>
</tr>
<tr>
<td>MAPPP (combined)</td>
<td><a href="http://www.mpib-berlin.mpg.de/MPAPP/expertquery.html">http://www.mpib-berlin.mpg.de/MPAPP/expertquery.html</a></td>
<td>Combined with MHC class I binding</td>
</tr>
<tr>
<td>Pcleavage</td>
<td><a href="http://www.imtech.res.in/raghava/pcleavage/">http://www.imtech.res.in/raghava/pcleavage/</a></td>
<td>Based on cleavage data or ligands</td>
</tr>
<tr>
<td>TAP translocation Prediction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAP Pred</td>
<td><a href="http://www.imtech.res.in/raghava/tappred/">http://www.imtech.res.in/raghava/tappred/</a></td>
<td>TAP transporter affinity</td>
</tr>
<tr>
<td>MHCPred</td>
<td><a href="http://www.jenner.ac.uk/mhcppred">http://www.jenner.ac.uk/mhcppred</a></td>
<td>TAP transporter affinity</td>
</tr>
<tr>
<td>PRED</td>
<td><a href="http://antigen.i2r.a-star.edu.sg/predTAP/">http://antigen.i2r.a-star.edu.sg/predTAP/</a></td>
<td>TAP transporter affinity</td>
</tr>
<tr>
<td>Integrated Proteasomal cleavage, TAP translocation and Binding prediction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epplen</td>
<td><a href="http://www.jenner.ac.uk/Epplen/">http://www.jenner.ac.uk/Epplen/</a></td>
<td>Uses MHCPred</td>
</tr>
<tr>
<td>WAPP</td>
<td><a href="http://www-bs.informatik.uni-tuebingen.de/WAPP">http://www-bs.informatik.uni-tuebingen.de/WAPP</a></td>
<td>Uses SVMHC</td>
</tr>
<tr>
<td>NetCTL</td>
<td><a href="http://www.cbs.dtu.dk/services/NetCTL/">http://www.cbs.dtu.dk/services/NetCTL/</a></td>
<td>Uses NetMHC and NetChop</td>
</tr>
<tr>
<td>MHC-Pathway</td>
<td><a href="http://70.167.3.42/">http://70.167.3.42/</a></td>
<td>Older version of IEDB (combined)</td>
</tr>
<tr>
<td>IEDB (combined)</td>
<td><a href="http://tools.imtech.res.in/analyze/html/mhc_processing.html">http://tools.imtech.res.in/analyze/html/mhc_processing.html</a></td>
<td>Newer version of MHC-Pathway</td>
</tr>
<tr>
<td>CTLPred</td>
<td><a href="http://www.imtech.res.in/raghava/ctlpred/">http://www.imtech.res.in/raghava/ctlpred/</a></td>
<td>Directly based on epitope databases</td>
</tr>
<tr>
<td>Databases of HLA class I and II ligands and tumor T cell epitopes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SYFPEITHI</td>
<td><a href="http://www.imtech.res.in/raghava/syfpeithi">http://www.imtech.res.in/raghava/syfpeithi</a></td>
<td>HLA ligands and binding motifs, T cell epitopes</td>
</tr>
<tr>
<td>T-cell tumor epitopes</td>
<td><a href="http://www.cancerimmunity.org/peptidedatabase/Tcell">http://www.cancerimmunity.org/peptidedatabase/Tcell</a> epitopes.htm</td>
<td>T cell epitopes derived from TAA</td>
</tr>
<tr>
<td>IEDB</td>
<td><a href="http://beta.immuneepitope.org/home.do">http://beta.immuneepitope.org/home.do</a></td>
<td>T cell epitopes (among others)</td>
</tr>
<tr>
<td>MHCBN</td>
<td><a href="http://bioinformatics.usams.edu/minor/mhcbn/index.html">http://bioinformatics.usams.edu/minor/mhcbn/index.html</a></td>
<td>MHC binding and non-binding peptides</td>
</tr>
</tbody>
</table>

* Full name, institute and references can be found on webpage. Several references are also provided in the text of the current review.

* Method and additional information to be found on webpage.

for vaccine development. HLA class II peptide binding predictions have successfully led to the identification of promiscuous T helper epitopes in among others NY-ESO-1 [389], TRP-2 [390], and hTERT [391], and Melan-A/MART-1 [392].

8.4. Identification of HLA class I and HLA class II ligands by tandem mass spectrometry

As discussed above, starting from pre-existing CTL with unknown or known specificity, the combination of microscale liquid chromatography coupled to tandem mass spectrometry (MS/MS) with functional immunoassays has been used for the identification or confirma-
tion of tumor-specific CTL epitopes, respectively [322-328]. Without the availability of an epitope-specific T cell, tandem mass spectrometry can also be applied for the validation of predicted HLA class I ligands (Fig 3). This alternative reverse immunology strategy was coined the 'predict-calibrate-detect' (PCD) method [376,378]. Predicted ligands are synthesized and used for the calibration of an HPLC – mass spectrometry system to identify co-eluting natural ligands of identical mass of which the precise identity is then verified by MS/MS sequencing. In this manner, Stevanovic and colleagues identified class I ligands from p53, CEA and MAGE-A1 in peptides extracted from tumor tissue or tumor cell lines [375,378]. Although technically demanding, the PCD method advantageously does not depend on the often cumbersome generation of peptide-specific CTL clones and allows the identification of low abundant peptides. It is, therefore, a secure intermediate station in the identification of CTL epitopes. The immunogenicity of the ligands still needs to be determined by T cell inductions [378,393]. With the aim to identify as many novel tumor-specific HLA ligands as possible in a single tumor sample, mass spectrometric ligand identification has been coupled to gene expression profiling to reveal antigens (over)expressed in the tumor but not in healthy tissue from the same patient [319]. This method has recently been used to identify HLA class I ligands derived from both universal and novel renal cell carcinoma (RCC) associated antigens [394]. These HLA class I ligands were in part unique for the patient and can therefore in principle be used in a personalized therapy.

Subtractive mass spectrometric approaches have been used for the direct identification of differentially expressed HLA class I and class II ligands with the aim to identify disease-related (e.g. TAA derived) ligands. Peptides in HLA-DR4 from an diabetes auto-antigen were identified by mass spectrometric comparison of HPLC-fractionated peptides purified from either untreated cells or cells that endogenously processed the antigen after it was delivered via a lectin-based method [395]. Lemmel et al. were the first to use differential stable isotope labelling (e.g. differential acetylation) of HLA class I bound peptides extracted from colon carcinoma versus regular colon tissue to quantify their ratio by mass spectrometry [396]. A variant of this subtractive analysis applied differential stable isotope labelling to two isoforms of a meningococcal outer membrane protein before their uptake by DC, again to create a so-called mass-tag (resulting in spectral doublets) that simplified the comparisons between HLA class II ligands extracted from the two sources [397]. The same group applied metabolic labelling by culturing virus-infected cells with stable isotope-labelled amino acids. Comparisons between the labelled peptides extracted from infected cells and the unlabeled peptides from non-infected cells have led to the identification of viral and infection-induced HLA class I ligands [398]. These pairwise comparative methods of mass spectrometric analysis could also be applied to tumor antigens. As most tumor cells do not express HLA class II, the direct isolation of class II bound peptides from tumor cells was not considered feasible. Instead, several strategies have been used to target antigens into antigen presenting cells [395,399,400]. A recent study, however, reports the successful identification of (tumor-associated) HLA class II ligands from dissected primary tumor samples that were found to express HLA class II [401].

The advent of proteomics in the identification of HLA bound ligands is typically technology-
driven [402,403]. With the development of improved mass spectrometry technologies that are more sensitive and more accurate the identification by MS/MS of many more peptides in one sample is now possible. High throughput MS/MS analyses in an automated data-dependent mode followed by database searches allows the identification of a significant percentage of the full HLA bound ‘ligandome’ of a sample without prior focusing on predicted epitopes. It is already now possible to identify up to 3000 peptides per allele from one cell line (personal communication P.A. van Veelen, unpublished data). Further improvements in sample preparation and separation techniques, and data analysis will still boost the results [404]. This will lead to the identification of novel tumor-specific HLA class I ligands (likely immunogenic T cell epitopes) derived from both known and as yet unknown universal tumor associated antigens and also from unique mutated antigens. Basic insights and the development of anti-cancer immunotherapies, which may include individualized vaccinations with defined tumor-specific epitopes that are partly unique for the patient [318], will greatly benefit from these developments.

9. The need to define more T cell epitopes

Several reasons exist to further broaden the repertoire of defined tumor specific CTL and T helper cell epitopes. First of all, the immunomonitoring of cancer patients treated with any T cell mediated therapy is required to assess the effectiveness of the treatment. Monitoring can also be performed to determine the existence of precursors in healthy donors against a specific candidate target CTL epitope. Monitoring of T cell responses is mostly accomplished ex vivo by measuring T cell populations with tetrarmers of the tumor-epitope or by measuring specific cytokine (e.g. IFNγ) production upon stimulation with defined T cell epitopes [405]. In the case of vaccination with undefined antigens (tumor cells, tumor lysates or e.g. tumor cell derived mRNA) or full length TAA (irrespective of the vehicle), the monitoring of T cell responses against multiple T cell epitopes will better reveal the effects induced by the therapy. Even in the case of vaccination with a single minimal epitope, immunomonitoring of a T cell response induced by so-called ‘antigenic spread’ against a non-vaccine CTL epitope that is expressed on the tumor may be necessary to correctly assess the efficacy of the therapy [406].

**Scope of the thesis, immunomonitoring:**

In chapter 3 [407], we screened HLA-A*0201-subtyped healthy individuals and advanced melanoma patients for the existence of CD8+ T cells directed against the four HLA-A0201-restricted CTL epitopes from PRAME that were identified in chapter 2. IFNγ enzyme-linked immunosorbent spot assays and tetramer staining were used to detect CTL reactivity. T cell reactivity was found to be directed especially towards the PRA100-108 epitope.

Recent studies, as well, have assessed above mentioned PRAME epitopes as candidate targets for immunotherapy. It was shown that CD8+ T cells against all four HLA-A2 restricted epitopes were detectable in healthy donors [408,409], and in patients suffering from CML [408-410]. A strong argument can be made for a defined multi-epitope and multi-TAA directed T cell immunotherapeutic approach, either by using adoptive transfer of PBL transduced with multiple TCR or by applying vaccination strat-
egies. Targeting multiple TAA will enhance the barrier against escape of antigen loss variants of the tumor and will exploit more fully the anti-tumor T cell potential of the patient (in the case of vaccination). Loss of HLA class I molecules on tumor cells, which can be another reason for immune escape, is often restricted to only one or a few alleles [411]. Targeting multiple epitopes restricted by different class I molecules of the patient will circumvent such an escape mechanism. Where the latter goals may also be reached by vaccination with multiple full length TAAs (expressed in the tumor), it has been shown that the use of optimal epitopes can induce immune responses with increased potency compared with the response induced by the same epitopes in the context of the full length protein [412]. Given the pivotal role of CD4⁺ T cells in promoting the primary and secondary CD8⁺ T cell responses through the induction of DC maturation and the production of cytokines [52,267-269], the inclusion of T helper epitopes in a multi-epitope based vaccine will have strong beneficial effects. Furthermore, vaccination with minimal CTL peptide epitopes, unless administered on DC, may cause T cell tolerance through their systemic spread and presentation on non-professional antigen presenting cells [413,414]. To circumvent this, vaccines should contain longer epitope-containing peptides that require processing which can only be accomplished efficiently by professional antigen presenting cells (DC). When enough CTL and T helper epitopes – derived from different TAA and presented in various prevalent HLA molecules – are identified, it would be feasible to combine these epitopes in a defined epitope-based vaccine (as peptide vaccine or e.g. recombinant ‘string-of-bead’ viral delivery system [415]) that is tailored to the TAA expression pattern and HLA haplotype of each patient. Importantly, vaccination with longer (35-mer) peptides, containing both CTL and T helper epitopes in their natural protein context, leads to a far more robust CD8⁺ T cell response and therapeutic immunity in a mouse model [271]. Both the induction of a concurrent CD4⁺ T cell response and the restricted processing and presentation of the long peptides only by professional APC contributed to this enhanced efficacy [416]. Vaccines based on defined epitopes have the additional advantage that the binding and TCR recognition characteristics of the epitopes can be optimized by aa replacements. In the case of differentiation TAA, tolerance against the immunodominant epitopes is expected, and these are therefore not first choice. The subdominant epitopes, however, mostly have a lower binding capacity rendering them less immunogenic. Designing modified analogs of the epitope, also called altered peptide ligands, with improved binding characteristics can be used to efficiently recruit a non-tolerized T cell repertoire [417,418]. However, care should be taken that vaccination with epitope analogs does not induce CTL that are incapable of recognizing tumor cells as has been observed in patients vaccinated with optimized variants of MART and gp100 CTL epitopes [419]. Obviously, any epitope contained in epitope-based vaccines should be thoroughly checked for its natural processing and cell surface presentation to avoid responses against so-called cryptic epitopes that are not presented on the tumor cells [420,421].

10. Purpose and chronology of the thesis

The research presented in this thesis has been initiated to identify CTL epitopes in the leve-
mia specific BCR-ABL fusion regions (chapter 5). As these fusion regions contain only a low number of epitopes, the priority shifted to epitope identification in tumor associated antigen PRAME using an optimized epitope-prediction procedure (chapter 2). To facilitate further epitope discovery, HLA class I binding assays were developed (chapter 4), and the potential usefulness of the epitopes that were identified was examined in healthy donors and patients (chapter 3). Meanwhile a study was started to further unravel the proteolytic mechanisms involved in CTL epitope generation (chapter 6).
References

59. Chang J, Cho JH, Lee SW, Choi SY, Ha SJ, Sung YC: IL-12 priming during in vitro antigenic stimulation changes properties of CD8 T cells and increases


381. Hural JA, Friedman RS, McNabb A, Steen SS, Henderson RA, Kalos M: Identification of naturally processed CD4(+) T cell epitopes from the prostate-


