Database of T cell-defined human tumor antigens: the 2013 update

Nathalie Vigneron1,2, Vincent Stroobant1,2, Benoît J. Van den Eynde1,2 and Pierre van der Bruggen1,2

1Ludwig Institute for Cancer Research, Brussels Branch, Brussels, Belgium
2WELBIO and de Duve Institute, Université catholique de Louvain, Brussels, Belgium

The plethora of tumor antigens that have been—and are still being—defined required systematization to provide a comprehensive overview of those tumor antigens that are the most relevant targets for cancer immunotherapy approaches. Here, we provide a new update of a peptide database resource that we initiated many years ago. This database compiles all human antigenic peptides described in the literature that fulfill a set of strict criteria needed to ascertain their actual “tumor antigen” nature, as we aim at guiding scientists and clinicians searching for appropriate cancer vaccine candidates (www.cancerimmunity.org/peptide). In this review, we revisit those criteria in light of recent findings related to antigen processing. We also introduce the 29 new tumor antigens that were selected for this 2013 update. Two of the new peptides show unusual features, which will be briefly discussed. The database now comprises a total of 403 tumor antigenic peptides.

Keywords: peptide, database, HLA, CD4, CD8, T cell

A tumor antigen database

Since the discovery of the first human T cell defined-tumor antigen in 1991 (1), a growing number of tumor antigens have been described at a regular pace, with dozens of new antigenic peptides reported in the literature every year. In 2001, we started compiling what we think are the most relevant human tumor antigens and created a peptide database, which was posted on the Cancer Immunity website and was regularly updated (2). This tool was primarily designed to provide the scientific community with an open view of the nature and diversity of the tumor antigens available to date for cancer immunotherapy. We selected a classification of peptides based on their tumor specificity because we think it is the most critical factor determining their usefulness for cancer immunotherapy. A first distinction is made between unique antigens, generally derived from point mutations, and shared antigens. The shared antigens are further divided into tumor-specific antigens, differentiation antigens, and overexpressed antigens (3). Various tables are provided on the Cancer Immunity website, recapitulating the sequences of the peptides included in these categories, their position in the protein sequence, a GeneCard link for the encoding gene and/or the parent protein, the HLA presenting molecule and its frequency in Caucasians, the method used to isolate the cytolytic T lymphocyte (CTL) recognizing the antigen, and a PubMed link to the relevant reference. Of note, a number of viruses, such as the Epstein-Barr virus (EBV) and human papilloma virus (HPV), are associated with human malignancies. The antigenic peptides encoded by viral genes have not been included in the database, despite their high potential as targets for immunotherapy.

Inclusion criteria for the peptide database

Because we wanted the database to be useful to clinicians contemplating the development of immunotherapy trials, we included only the peptides that were fully validated, i.e., whose characterization was comprehensive and fully demonstrated the existence, nature, immunogenicity and, most importantly, the natural presentation of these antigenic peptides by tumor cells. Those candidate peptides whose comprehensive characterization was not reported were listed in an additional category of “potential peptides” awaiting further characterization. The “validated peptides” we selected needed to meet the following six requirements:

1. Isolation of stable human T lymphocyte clones or lines recognizing the peptide
2. Identification of the peptide recognized by the T cells
3. Identification of the HLA presenting molecule
4. Evidence that the peptide is processed by tumor cells and presented to the specific CTL

This implies showing recognition of tumor cells expressing the relevant gene and HLA molecule by the T cells. When a polyclonal T cell line is used rather than a clone, it is essential to demonstrate that the CTLs that lyse the tumor cells are the same as those that recognize the peptide. This can be done by “cold target inhibition” experiments using peptide-pulsed cold targets (4). Other means of proof are also possible, such as the testing of stable transfectants of tumor cells with the sequence encoding the parental protein (5) or knocking down the gene encoding the peptide using siRNA or shRNA technology (6).

In some cases, unusual processing features of the peptide (e.g., by only some proteasome subtypes (7-12)) may explain the lack of recognition of some tumor cells by the CTLs, without precluding the “validation” of the peptide, provided it is explained.

When post-translational modifications are involved, characterization of the peptide should include elution of the peptide from the cell surface (13-16). Eluted fractions can then be tested for their ability to activate the CTLs. The CTL-sensitizing fraction should correspond to the fraction able to sensitize the CTLs when the synthetic peptide of interest is fractionated by HPLC in the exact same conditions. Alternatively, the presence of the peptide of interest in the CTL-sensitizing fraction could also be demonstrated by mass spectrometry.
Table 1
Tumor antigens resulting from mutations.

<table>
<thead>
<tr>
<th>Gene/protein</th>
<th>Tumor</th>
<th>HLA</th>
<th>HLA Frequency (%)</th>
<th>Peptide</th>
<th>Position</th>
<th>Lymphocyte Stimulation Method</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLPP</td>
<td>melanoma</td>
<td>A2</td>
<td>44</td>
<td>ILDKVLVLHL</td>
<td>240-248</td>
<td>autologous tumor cells</td>
<td>(9)</td>
</tr>
</tbody>
</table>

Table 2
Shared tumor-specific antigens.

<table>
<thead>
<tr>
<th>Gene</th>
<th>HLA</th>
<th>HLA Frequency (%)</th>
<th>Peptide</th>
<th>Position</th>
<th>Lymphocyte Stimulation Method</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyclin-A1</td>
<td>A2</td>
<td>44</td>
<td>FLDRFLSCM</td>
<td>227-235</td>
<td>peptide library</td>
<td>(20)</td>
</tr>
<tr>
<td></td>
<td>A2</td>
<td>44</td>
<td>SUAAAACCLA</td>
<td>341-351</td>
<td>peptide library</td>
<td>(20)</td>
</tr>
<tr>
<td>MAGE-A1</td>
<td>B44</td>
<td>21</td>
<td>KEADPTGHSY</td>
<td>160-169</td>
<td>poxvirus - dendritic cells</td>
<td>(21)</td>
</tr>
<tr>
<td></td>
<td>Cw7</td>
<td>41</td>
<td>RVRFPPSL</td>
<td>289-298</td>
<td>peptide</td>
<td>(36)</td>
</tr>
<tr>
<td>MAGE-C1</td>
<td>A2</td>
<td>44</td>
<td>ILFGISLREV</td>
<td>959-968</td>
<td>peptide</td>
<td>(36)</td>
</tr>
<tr>
<td></td>
<td>A2</td>
<td>44</td>
<td>KVIEFLAML</td>
<td>1083-1091</td>
<td>peptide</td>
<td>(36)</td>
</tr>
<tr>
<td>MAGE-C2</td>
<td>B57</td>
<td>8</td>
<td>ASSTLYLVF</td>
<td>42-50</td>
<td>autologous tumor cells</td>
<td>(8)</td>
</tr>
<tr>
<td></td>
<td>DR15</td>
<td>20</td>
<td>SSSTLYLVFSFST</td>
<td>43-57</td>
<td>peptide</td>
<td>(40)</td>
</tr>
<tr>
<td>SSX2</td>
<td>DR1</td>
<td>18</td>
<td>FGRLGIGSPIK</td>
<td>101-111</td>
<td>peptide</td>
<td>(41)</td>
</tr>
<tr>
<td>XAGE1tvGAGED2a</td>
<td>A2</td>
<td>44</td>
<td>ROKKKIRKL</td>
<td>21-29</td>
<td>peptide</td>
<td>(42)</td>
</tr>
<tr>
<td></td>
<td>DR4</td>
<td>24</td>
<td>HLGSROKQRKIRIQSRQ</td>
<td>17-32</td>
<td>peptide</td>
<td>(42)</td>
</tr>
</tbody>
</table>

Table 3
Differentiation antigens.

<table>
<thead>
<tr>
<th>Gene/protein</th>
<th>Tumor</th>
<th>HLA</th>
<th>HLA Frequency (%)</th>
<th>Peptide</th>
<th>Position</th>
<th>Lymphocyte Stimulation Method</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Melan-A/MART-1</td>
<td>melanoma</td>
<td>DP5</td>
<td>3</td>
<td>YTTAEBAAGGLTVLGLVLLIGGWYCR</td>
<td>21-50</td>
<td>peptide</td>
<td>(43)</td>
</tr>
<tr>
<td>TRP-1</td>
<td>melanoma</td>
<td>DR17</td>
<td>21</td>
<td>SQWRVVDIDSLEVD</td>
<td>284-286</td>
<td>peptide</td>
<td>(44)</td>
</tr>
<tr>
<td>Tyrosinase</td>
<td>melanoma</td>
<td>A24</td>
<td>20</td>
<td>YVMQGTAQSF</td>
<td>368-373 and 335-340</td>
<td>autologous tumor cells</td>
<td>(13)</td>
</tr>
</tbody>
</table>

* The peptide is composed of two non-contiguous fragments that are spliced by the proteasome.

Table 4
Antigens overexpressed in tumors.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Normal Tissue Expression</th>
<th>HLA</th>
<th>HLA Frequency (%)</th>
<th>Peptide</th>
<th>Position</th>
<th>Lymphocyte Stimulation Method</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD45</td>
<td>proliferating cells, tests, multiple tissues (low level)</td>
<td>A24</td>
<td>20</td>
<td>KFLDALISL</td>
<td>556-564</td>
<td>peptide</td>
<td>(6)</td>
</tr>
<tr>
<td>glypican-3</td>
<td>placenta and multiple tissues</td>
<td>A2</td>
<td>44</td>
<td>FVGEFFTDV</td>
<td>144-152</td>
<td>peptide</td>
<td>(45)</td>
</tr>
<tr>
<td></td>
<td>placenta and multiple tissues</td>
<td>A24</td>
<td>20</td>
<td>EYLISLEEL</td>
<td>298-306</td>
<td>peptide</td>
<td>(45)</td>
</tr>
<tr>
<td>IGF2B3</td>
<td>ubiquitous (low level)</td>
<td>A2</td>
<td>44</td>
<td>NLSMAEVV</td>
<td>515-523</td>
<td>peptide</td>
<td>(46)</td>
</tr>
<tr>
<td></td>
<td>ubiquitous (low level)</td>
<td>A3</td>
<td>44</td>
<td>RLLVPQTVF</td>
<td>199-207</td>
<td>peptide</td>
<td>(46)</td>
</tr>
<tr>
<td>Kallikrein-4</td>
<td>prostate and ovarian carcinoma</td>
<td>A2</td>
<td>44</td>
<td>FLGYILIGV</td>
<td>11-19</td>
<td>peptide</td>
<td>(47)</td>
</tr>
<tr>
<td>KIF20A</td>
<td>ubiquitous (low level)</td>
<td>A2</td>
<td>44</td>
<td>LSDDDDVVV</td>
<td>12-20</td>
<td>peptide</td>
<td>(48)</td>
</tr>
<tr>
<td></td>
<td>ubiquitous (low level)</td>
<td>A2</td>
<td>44</td>
<td>AQDTPALPV</td>
<td>284-283</td>
<td>peptide</td>
<td>(48)</td>
</tr>
<tr>
<td></td>
<td>ubiquitous (low level)</td>
<td>A2</td>
<td>44</td>
<td>CIAEQYHTV</td>
<td>809-817</td>
<td>peptide</td>
<td>(48)</td>
</tr>
<tr>
<td>Lengsin</td>
<td>eye lens and low level in multiple tissues</td>
<td>A2</td>
<td>44</td>
<td>FLPEFQISSA</td>
<td>270-279</td>
<td>peptide</td>
<td>(49)</td>
</tr>
<tr>
<td>Meloe</td>
<td>ubiquitous (low level)</td>
<td>Q6</td>
<td>63</td>
<td>RISSTLDECWPA</td>
<td>52-44</td>
<td>peptide</td>
<td>(50)</td>
</tr>
<tr>
<td></td>
<td>ubiquitous (low level)</td>
<td>DR11</td>
<td>25</td>
<td>CPPWPSERISSTL</td>
<td>24-37</td>
<td>peptide</td>
<td>(50)</td>
</tr>
<tr>
<td>MUC5AC</td>
<td>surface mucosal cells, respiratory tract and stomach epithelia</td>
<td>A24</td>
<td>20</td>
<td>TCOPTCRSL</td>
<td>716-724</td>
<td>peptide</td>
<td>(51)</td>
</tr>
<tr>
<td>survivin</td>
<td>ubiquitous</td>
<td>DR1</td>
<td>18</td>
<td>TLGEFLKLDLRAK</td>
<td>97-111</td>
<td>peptide</td>
<td>(52)</td>
</tr>
</tbody>
</table>
In the case of CD4 T lymphocytes, which may not recognize tumor cells directly, the fact that the peptide is processed can be shown by testing antigen-presenting cells (APCs) loaded with the recombinant protein or a control protein produced in the same organism (17, 18), or loaded with lysates of cells transfected or not with the relevant coding sequence.

5. Report of a peptide sensitization assay
Characterization of peptides recognized by CD8 T cells should generally include the identification of the shortest peptide recognized and a titration showing a clear recognition of this peptide at doses below 1 μM. Putting together this update, we realized that in recent years, fewer investigators perform this control. In the case of peptides that were identified by the “reverse immunology approach” (an approach that consists in raising T cells against specific peptides corresponding to fragments of conventional proteins), we considered that the peptide titration curves might not be mandatory as the CTLs were directly obtained against the peptide of interest. However, we believe that peptide titration curves remain crucial in order to determine the most adequate immunotherapeutic vaccination modalities (cfr. the B*4402-restricted peptide MAGE-A1 KEADPTGHSY example below) and should therefore always be included when describing new potential vaccine candidates.

6. Description of the pattern of antigen expression
A certain level of tumor- or tissue-specificity should be documented, as ubiquitous antigens do not qualify as tumor antigens. This can be done with gene expression, protein expression, or lymphocyte recognition data, which should ideally be corroborative.

New relevant tumor antigens

Among 46 potentially relevant papers considered for the 2013 update, 29 antigens fulfilled the criteria described above and were therefore included in the database, while the other relevant peptides found in the 25 remaining papers did not meet all the requirements and were therefore included in the “potential antigens” list. Table 1 through Table 4 provides the list of new antigens that will be included in the updated database. Most of these peptides were identified using the reverse immunology approach (19), involving either a specific peptide or peptide libraries. The latter technology led, for example, to the mapping of two new antigenic peptides derived from cyclin-A1 and probably representing the most truly leukemia-specific epitopes identified so far (20). Although the reverse immunology approach proved efficient for the identification of many tumor antigens, it will hardly lead to the identification of peptides originating from unexpected events relative to transcription, translation, post-translational modifications or processing. Below, we describe two such peptides, which were included in this year’s update. These peptides were recognized by CTLs and tumor-infiltrating lymphocytes (TILs) raised against autologous tumors and show unexpected properties.

The B*4402-restricted peptide MAGE-A1 KEADPTGHSY
This antigenic peptide is very unusual in the sense that, although it is clearly encoded by MAGE-A1, processed endogenously, and presented by tumor cells, the corresponding synthetic peptide is unable to sensitize target cells to CTL recognition, unless pulsed on paraformaldehyde or acid-treated targets (21). This was shown to result from the lack of binding of the peptide on surface HLA-B*4402, which could originate from the low peptide-receptiveness of this tapasin-dependent HLA molecule. This study exemplifies the importance of doing a peptide titration curve when identifying a potential peptide vaccine candidate. Indeed, peptides such as this MAGE-A1 peptide, which can only be recognized at concentrations higher that 1 μM, might not be adequate to use in a classical synthetic peptide vaccine but might rather require vaccination modalities relying on endogenous presentation such as RNA- or DNA-based vaccines.

An antigenic peptide produced after reverse splicing and double asparagine deamidation

While peptides recognized by CTLs are generally derived from fragments of conventional proteins, an increasing variety of non-classical events were shown to contribute to the production of these peptides. Antigenic peptides were produced by the aberrant transcription of intronic or reverse strand sequences (22-25). Other peptides arise from the translation of alternative open-reading frames or pseudogenes (26-32). Peptides produced from post-translational modifications such as deamidation of asparagine residues (15, 33), serine/threonine phosphorylation (34, 35), or peptide splicing were also described (14, 16, 36, 37). One of the peptides added in the current release of the database was found, surprisingly, to combine several post-translational modifications (13). It is first composed of two tyrosinase fragments that are spliced in the reverse order to that in which they appear in the parent protein. In addition, each of the splicing partners contains an aspartic acid residue, originating from the post-translational conversion of a genetically encoded asparagine. Studying the processing of this usual peptide, Dalet et al. showed that it requires translation of tyrosinase into the endoplasmic reticulum (ER) and retrotranslocation into the cytosol, where the two asparagines are deglycosylated by peptide-N-glycanase, resulting in their conversion into aspartates by deamidation. This is followed by cleavage and splicing of the appropriate fragments by the standard proteasome, and further transport of the resulting peptide into the ER through the TAP transporter.

Coda
We intend to keep updating the database on a regular basis, as a service to the scientific community. We welcome any suggestions regarding the inclusion of other peptides, or the revision of some inclusion criteria or any other aspect of the database.

Abbreviations

CTL, cytolytic T lymphocyte; TIL, tumor-infiltrating lymphocyte

Acknowledgements

N.V. is supported by a postdoctoral fellowship from the F.R.S.-FNRS.

References

recognized by cytolytic T lymphocytes on a human melanoma. *Science* 1991; **254**: 1643-1647. (PMID: 1840703)


5. Vigneron N, Ooms A, Morel S, Ma W, Degiovanni G, Van den Eynde B. A peptide derived from melanocytic protein gp100 and presented by HLA-B35 is recognized by autologous cytolytic T lymphocytes on melanoma cells. *Tissue Antigens* 2005; **65**: 156-162. (PMID: 15713214)


cyte antigen-A2-restricted cytotoxic T lymphocytes reactive to cancer cells. Cancer Sci 2011; 102: 71-78. (PMID: 21087352)


Contact
Address correspondence to:
Pierre van der Bruggen, Ph.D.
de Duve Institute, Université catholique de Louvain, and Ludwig Institute for Cancer Research
B1.74.03, Avenue Hippocrate 74
B-1200 Brussels
Belgium
Tel.: + 32 0 2 764 74 31
Fax: + 32 0 2 792 94 05
E-mail: pierre.vanderbruggen@bru.licr.org