

**A listing of human tumor antigens recognized by T cells: March 2004
update**

Luisa Novellino - Chiara Castelli - Giorgio Parmiani

Unit of Immunotherapy of Human Tumors, Istituto Nazionale Tumori,

Via G. Venezian 1, 20133 Milan (Italy)

Key words: Antigens, Tumor, T cells, Epitopes

✉ Corresponding author:

Dr Chiara Castelli, Unit of Immunotherapy of Human Tumors,

Istituto Nazionale Tumori, Via Venezian 1, 20133 Milan (Italy),

e-mail chiara.castelli@istitutotumori.mi.it

Complete list of abbreviations of tumor antigens **707-AP** = 707 alanine proline - **AFP** = alpha (α)-fetoprotein - **AIM-2** = interferon-inducible protein absent in melanoma 2 - **ART-4** = adenocarcinoma antigen recognized by T cells 4 - **BAGE** = B antigen - **Bcr-abl** = breakpoint cluster region-Abelson - **CAMEL** = CTL-recognized antigen on melanoma - **CAP-1** = carcinoembryonic antigen peptide-1 - **CASP-8** = caspase-8 - **CDC27** = cell-division-cycle 27 - **CDK4** = cyclin-dependent kinase 4 - **CEA** = carcino-embryonic antigen - **CLCA2** = calcium-activated chloride channel-2 - **CT** = cancer/testis (antigen) - **Cyp-B** = cyclophilin B - **DAM** = differentiation antigen melanoma (the epitopes of DAM-6 and DAM-10 are equivalent, but the gene sequences are different. DAM-6 is also called MAGE-B2 and DAM-10 is also called MAGE-B1) - **ELF2** = elongation factor 2 - **Ep-CAM** = epithelial cell adhesion molecule - **EphA2, 3** = Ephrin type-A receptor 2, 3 - **ETV6-AML1** = Ets variant gene 6/acute myeloid leukemia 1 gene ETS - **FGF-5** = Fibroblast growth factor-5 - **FN** = fibronectin - **G250** = glycoprotein 250 - **GAGE** = G antigen - **GnT-V** = N-acetylglucosaminyltransferase V - **Gp100** = glycoprotein 100 kD - **HAGE** = helicase antigen - **HER-2/neu** = human epidermal receptor-2/neurological - **HLA-A*0201-R170I** = arginine (R) to isoleucine (I) exchange at residue 170 of the α -helix of the α 2-domain in the HLA-A2 gene - **HSP70-2M** = heat shock protein 70-2 mutated - **HST-2** = human signet ring tumor-2 - **hTERT** = human telomerase reverse transcriptase - **iCE** = intestinal carboxyl esterase - **IL-13R α 2** = interleukin 13 receptor α 2 chain - **KIAA0205** = name of the gene as it appears in databases - **LAGE** = L antigen - **LDLR/FUT** = low density lipid receptor/GDP-L-fucose: β -D-galactosidase 2- α -L-fucosyltransferase - **MAGE** = melanoma antigen - **MART-1/Melan-A** = melanoma antigen recognized by T cells-1/Melanoma antigen A - **MART-2** = melanoma Ag recognized by T cells-2 - **MC1R** =

melanocortin 1 receptor - **M-CSF** = macrophage colony-stimulating factor gene - **MUC1, 2** = mucin 1, 2 - **MUM-1, -2, -3** = melanoma ubiquitous mutated 1, 2, 3 - **NA88-A** = NA cDNA clone of patient M88 - **Neo-PAP** = Neo-poly(A) polymerase - **NPM/ALK** = nucleophosmin/anaplastic lymphoma kinase fusion protein - **NY-ESO-1** = New York - esophageous 1 - **OA1** = ocular albinism type 1 protein - **OGT** = O-linked N-acetylglucosamine transferase gene - **OS-9** = name of the gene as it appears in databases - **P15** = protein 15 - **p190 minor bcr-abl** = protein of 190 KD bcr-abl - **Pml/RAR α** = promyelocytic leukemia/retinoic acid receptor α - **PRAME** = preferentially expressed antigen of melanoma - **PSA** = prostate-specific antigen - **PSMA** = prostate-specific membrane antigen - **PTPRK** = receptor-type protein-tyrosine phosphatase kappa - **RAGE** = renal antigen - **RU1, 2** = renal ubiquitous 1, 2 - **SAGE** = sarcoma antigen - **SART-1, -2, -3** = squamous antigen rejecting tumor 1, 2, 3 - **SSX-2** = synovial sarcoma, X breakpoint 2 - **Survivin-2B** = intron 2-retaining survivin - **SYT/SSX** = synaptotagmin I/synovial sarcoma, X fusion protein - **TEL/AML1** = translocation Ets-family leukemia/acute myeloid leukemia 1 - **TGF β RII** = transforming growth factor β receptor 2 - **TPI** = triosephosphate isomerase - **TRAG-3** = taxol resistant associated protein 3 - **TRG** = testin-related gene - **TRP-1** = tyrosinase related protein 1, or gp75 - **TRP-2** = tyrosinase related protein 2 - **TRP-2/INT2** = TRP-2/intron 2 - **TRP-2/6b** = TRP-2/novel exon 6b - **WT1** = Wilms' tumor gene.

Abbreviations used **ALL** = acute lymphoblastic leukemia - **AML** = acute myeloid leukemia - **APL** = acute promyelocytic leukemia - **CML** = chronic myelogenous leukemia – **CTL** = cytotoxic T lymphocytes - **Ets** = E-26 transforming specific (family of transcription factors) - **H/N** = head and neck - **MHC** = major histocompatibility complex – **MSI** = microsatellite instability - **NSCLC** = non-small cell lung carcinoma - **ORF** = open reading frame - **RCC** = renal cell carcinoma - **SCC** = squamous cell carcinoma - **TAA** = tumor-associated antigen - **TSTA** = tumor-specific transplantation antigens.

Introduction

Since the cloning of *MAGE-1* [188], the first gene reported to encode a human tumor antigen recognized by T cells, molecular identification and characterization of novel tumor-associated antigens (TAAs) has rapidly evolved, in part due to the availability of new technology. Molecular cloning of single TAA by screening tumor-derived cDNA libraries with autologous tumor-specific T lymphocytes has been integrated with novel strategies such as 1) reverse immunology (epitope prediction on the basis of known HLA binding motifs performed by dedicated software and sometimes supported by proteasome-cleavage programs), 2) biochemical methods which elute and fractionate TAA peptides naturally expressed on tumor cells in the context of HLA molecules by chromatography and mass spectrometry, and 3) DNA microarray technology which allows comparison of gene expression profiles in tumor tissues and normal counterparts (RDA, representational difference analysis; DD, differential display; SSH, suppression subtractive hybridization; SAGE, serial analysis of gene expression).

Interestingly, these new technologies are shedding light on the involvement of a number of TAAs (both shared and unique) in the mechanisms of neoplastic transformation. This may allow novel tumor immunotherapeutic strategies based on administration of TAAs indispensable for maintaining the neoplastic state (e.g. N- and K-RAS), and/or the formulation of single patient-tailored vaccines which would comprise a large part of the individual patient's TAA repertoire, including strongly immunogenic unique tumor antigens.

Thus, it is important to categorize all these new antigens, particularly for the HLA allele restricting their recognition by T cells and for their tissue distribution. To this end,

here we survey TAAs so identified and briefly comment on each. The list presented in the tables below includes all T cell-defined epitopes encoded by TAAs and published by February 2004. Analogs or artificially modified epitopes are excluded from the list, as well as all viral encoded antigens. Only TAAs recognized by T cells (either CD8⁺ or CD4⁺) are listed, given their potential importance in the control of tumor growth. Antigens identified by antibodies are excluded, but a large collection of them, as detected by the Serex technology, can be found in the data base of the Institute for Cancer Research (<http://www.licr.org/SEREX.html>). It is of note that many tumor antigens (e.g. MAGE, NY-ESO-1) are now known to be recognized by both T cells and antibodies in the same cancer patient [30, 77].

In the tables herein, TAAs are listed in alphabetic order along with the epitope sequence and the HLA allele which restricts recognition by T cells. Furthermore, data on the tissue distribution of each antigen are provided, making this list an important source for easily retrieving data concerning human TAAs. Tables 1-4 collect different groups of class I HLA-restricted TAAs, whereas class II HLA-restricted counterparts are grouped under different subsets in Table 5. Table 6 assembles all characterized class I and class II HLA-restricted immunogenic fusion proteins. The separation of class I and class II HLA-restricted TAAs of corresponding groups into different tables is only justified by the fact that the number of the latter is still lower.

Moreover, some information is given to the reader in order to facilitate a comprehensive understanding of the data presented. All those TAAs in tables 1-3 which also include class II HLA-restricted immunogenic epitopes are shown in bold. Finally, splicing aberrations, point mutations and fusion junctions in epitopes listed in Tables 3-6 are underlined. The bibliography (alphabetically ordered) allows a rapid search for more

detailed information at the single antigen or epitope level. Overall, the updated list is intended to be a database tool for clinicians, scientists and students who have an interest in the field of tumor immunology and immunotherapy.

Classification of tumor antigens

Cancer/testis antigens - Class I HLA-restricted antigens (Table 1) and class II HLA-restricted antigens (subset of Table 5)

A milestone in tumor immunology was certainly the cloning of *MAGE-1* [188] and the subsequent characterization of the first T cell-defined antigenic epitope a year later [181]. Those findings were rapidly followed by the identification of new members within this group of TAAs [16, 186]: the *MAGE*, *BAGE* and *GAGE* families of genes were established. The antigens belonging to this group, now including also NY-ESO-1 and its alternative ORF products (*LAGE*, *CAMEL*), were originally called cancer/testis (CT) antigens because of their expression in histologically different human tumors and, among normal tissues, only in spermatocytes/spermatogonia of testis and, occasionally, in placenta. An alternative but less popular designation of these TAAs is “germ-line antigens”.

These TAAs have represented one of the main components of the anti-tumor vaccines tested in the clinic during the last decade. CT antigens result from re-activation of genes which are normally silent in adult tissues [42], but are transcriptionally activated in different tumor histotypes [43]. Their expression in testis does not provide targets for an auto-immune reaction because cells of testis do not express class I and II HLA molecules [80]. Despite the fact that the CT antigens are probably the best characterized tumor targets, their physiological function remains largely unknown [135].

Considering that new genes of this group have been cloned (CT9 [157], CT10 [64], *LAGE* [107], *MAGE-B5*, *-B6*, *-C2*, *-C3* and *-D* [112, 113], *HAGE*, *SAGE* [118], *SSX-2*

[8] and TRAG-3 [218]), the question arises as to how many more genes encoding CT antigens remain to be discovered and how many epitopes may exist that could be of use in cancer immunotherapy.

Differentiation antigens - Class I HLA-restricted antigens (Table 2) and class II HLA-restricted antigens (subset of Table 5)

These TAA are shared between tumors and the normal tissue from which the tumor arose; most are found in melanomas and normal melanocytes [6]. Many of these melanocyte lineage-related proteins are involved in the biosynthesis of melanin. Interestingly, novel differentiation TAAs are being found in epithelial tissues and tumors such as prostate and breast carcinomas, providing new tools for immunotherapy specifically directed against these solid tumors.

This group of TAAs, despite representing self-antigens, has been and still is being commonly used in current cancer vaccination trials, often together with CT antigens.

Widely occurring, overexpressed TAAs - Class I HLA-restricted antigens (Table 3) and class II HLA-restricted antigens (subset of Table 5)

Genes encoding widely expressed TAAs have been detected in histologically different types of tumors (often with no preferential expression on a certain type of cancer) as well as in many normal tissues, generally with lower expression levels.

It is possible that many of the epitopes processed and potentially presented by normal tissues are below the threshold level for T cell recognition, while their

overexpression in tumor cells can trigger an anti-cancer response by breaking previously-established tolerance.

Interestingly, these widely-expressed gene products have revealed a broad spectrum of mechanisms involved in generating T cell-defined epitopes, such as splicing aberrations leading to cryptic epitopes encoded by non-spliced introns, alternative ORFs, and even a case of post-translational splicing (FGF-5, Table 3). Surprisingly, many of these aberrations are also found in normal tissues, although at low levels, thus revealing a possible as yet unknown role of alternative forms of these antigens [144].

It is worth noting that some of the widely expressed/overexpressed TAAs were discovered by DNA microarray technologies, combined with new immunological tools such as reverse immunology and tetramer staining [207]. Among the most interesting TAAs of this group are the anti-apoptotic proteins (livin, survivin), hTERT, and tumor suppressor proteins (e.g., p53).

Unique and shared tumor-specific antigens - Class I HLA-restricted antigens (Table 4) and class II HLA-restricted antigens (subset of Table 5)

Unique TAA arise from point mutations of normal genes (such as β -catenin, CDK4, etc.) [148, 209]. Some of these molecular changes are associated with neoplastic transformation and/or progression. In mouse models unique antigens have been shown to be more immunogenic than other groups of shared antigens [47]; because unique antigens are responsible of the rejection of tumor transplants in mice, they have been defined as tumor-specific transplantation antigens (TSTA). In humans, response to the unique TAAs appears to be associated with a good prognosis of the patient [12, 82, 127].

Unfortunately, the major drawback of these antigens is that they are generally expressed only in the tumor where they were first identified. Thus, unique TAAs are the most specific targets for immunotherapy, but this potential advantage must be balanced against the logistical difficulty of their widespread clinical use. However, novel immunotherapeutic strategies are pointing to single patient-tailored anti-tumor vaccinations, with the aim of designing in a short time personalized vaccines comprising all possible tumor antigens expressed by patient's own tumor (including unique TAAs) [13, 207].

Few altered tumor-specific but shared antigenic epitopes have been identified, which are generated by different mechanisms occurring in tumor but not in normal cells, such as splicing aberrations (e.g. TRP-2/INT2, TRP-2/6b etc), and point mutations (N- and K-RAS) [59, 94, 109, 114]. Generally, these alterations are an obligatory step in neoplastic transformation, thus generating TAAs which are both widespread in different cancers and capable of inducing a true tumor-specific reaction mediated by the newly generated epitope. Also important is the spectrum of different tumors which could be targeted by these shared tumor-specific antigens, extending from melanoma (N-RAS-61m) [109] to pancreatic and colorectal adenocarcinomas (K-RAS) [59], and from MSI⁺ colon carcinomas to glioblastoma multiforme [94, 110].

Class II HLA-restricted antigens (Table 5)

As already mentioned above, the separation of class I and class II HLA-restricted TAAs of corresponding groups into separate tables is only justified by the fact that the number of class II HLA-restricted antigens is smaller than that of the class I HLA-restricted counterpart.

Stimulation of the CD4⁺ T helper cells is considered to be a pivotal step in raising an efficient and durable immune response to tumors. Therefore, the identification of TAA epitopes recognized by such lymphocytes is a crucial step in the long sought improvement of the anti-tumor immune response that could result in increased clinical efficacy.

The first epitope presented by a class II HLA molecule and capable of provoking a CD4⁺ T cell response was identified in 1994 as the melanoma differentiation antigen tyrosinase [176]. Then a gap of 4 years followed during which only one additional epitope was characterized [177], before other genes encoding class II HLA-restricted peptides were discovered. However, as technical and methodological approaches for identifying CD4⁺ T cell epitopes of tumor antigens have become available (among others, invariant chain-cDNA fusion libraries [204], humanized transgenic mice [217] and biochemical approaches [143]), an exponential increase in reporting such epitopes has been seen. Indeed, at the present time, class II HLA-restricted TAA epitopes have been identified which cover all the known types of TAA, from differentiation to CT antigens, and from widely expressed/overexpressed to tumor-specific unique antigens, as shown in Table 5.

It is of note that the mutated proteins subgroup also includes a shared tumor-specific antigen (TGFβRII) which is characteristic of MSI⁺ colon carcinomas [155].

Class I and class II HLA-restricted fusion proteins (Table 6)

In several malignancies, particularly in some forms of leukemias, the molecular mechanism of carcinogenesis involves translocation of chromosomes which results in

fusion of distant genes. This often causes the synthesis of fusion proteins which characterize each type of disease (e.g. BCR-ABL in CML, DEK-CAN and TEL/AML1 in AML, ETV6/AML and NPM/ALK in ALL, pml-RAR α in APL, SYT/SSX in synovial sarcomas) and generate new CD8⁺ and/or CD4⁺ T cell epitopes generally spanning the fusion junction. This provides new T cell epitopes falling within the group of non-self, shared, class I and class II HLA-restricted tumor-specific antigens, which can be employed in a large number of patients and tumor histologies [138].

Among these TAA only LDLR/FUT can be considered a unique antigen.

Frequency of epitope recognition by HLA-A, -B, -C, and -DR alleles (Table 7)

Table 7 summarizes the distribution of epitopes recognized in the context of different HLA loci. Data show that the majority of epitopes of a given group of TAAs is restricted by HLA-A, though in several cases (e.g. NY-ESO-1, tyrosinase) the percentage of HLA-DR restriction is equal or higher than class I HLA-restriction. This table suggests that a wide spectrum of tumor epitopes is available for the construction of anti-tumor vaccines potentially capable of stimulating both tumor-specific CD4⁺ and CD8⁺ T cells.

Conclusions

Several excellent and timely reviews on tumor antigens have been periodically published during the last few years [18, 87, 150]. The present contribution is a comprehensive list of all available TAAs, their T cell epitopes and HLA restriction, despite the fact that the features of each antigen can be easily found in the corresponding bibliography. A similar data base can be found in <http://www.cancerimmunity.org/peptidedatabase/Tcellepitopes.htm>. We hope that our work will be of interest to many tumor immunologists and students. Needless to say, we may have inadvertently missed information on some antigens despite our careful scrutiny of the published literature; therefore, we will be grateful to any reader who will provide us with any missing information.

The antigen list can also be found at the INT website (www.istitutotumori.mi.it).

Acknowledgements The authors would like to thank Ms Grazia Barp for editorial assistance. This work was supported in part by a grant of the European Community (QLK3-1999-00064) and the Italian Ministry of Education, University and Research (Rome, FIRB RBNE017B4C).