HLA Class I Supertypes: A Revised and Updated Classification

Citation
Sidney, John, Bjoern Peters, Nicole Frahm, Christian Brander, and Alessandro Sette. 2008. HLA class I supertypes: a revised and updated classification. BMC Immunology 9: 1.

Published Version
doi://10.1186/1471-2172-9-1

Permanent link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:5333510

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story
The Harvard community has made this article openly available. Please share how this access benefits you. Submit a story.

Accessibility
Background
Class I major histocompatibility complex (MHC) molecules bind, and present to T cells, short peptides derived from the intracellular processing of proteins. The peptide repertoire of a specific molecule is to a large extent determined by the molecular structure accommodating so-called main anchor positions of the presented peptide. These receptors are extremely polymorphic, and much of the polymorphism influences the peptide-binding repertoire. However, despite this polymorphism, class I molecules can be clustered into sets of molecules that bind largely overlapping peptide repertoires. Almost a decade ago we introduced this concept of clustering human leukocyte antigen (HLA) alleles and defined nine different groups, denominated as supertypes, on the basis of their main anchor specificity. The utility of this original supertype classification, as well several other subsequent arrangements derived by others, has been demonstrated in a large number of epitope identification studies.

Results: Following our original approach, in the present report we provide an updated classification of HLA-A and -B class I alleles into supertypes. The present analysis incorporates the large amount of class I MHC binding data and sequence information that has become available in the last decade. As a result, over 80% of the 945 different HLA-A and -B alleles examined to date can be assigned to one of the original nine supertypes. A few alleles are expected to be associated with repertoires that overlap multiple supertypes. Interestingly, the current analysis did not identify any additional supertype specificities.

Conclusion: As a result of this updated analysis, HLA supertype associations have been defined for over 750 different HLA-A and -B alleles. This information is expected to facilitate epitope identification and vaccine design studies, as well as investigations into disease association and correlates of immunity. In addition, the approach utilized has been made more transparent, allowing others to utilize the classification approach going forward.
by class I molecules is accomplished by interaction of the peptide amino acid side chains with discrete pockets within the peptide-binding groove of the MHC molecule formed by the α1 and α2 domains of the heavy chain. Typically, in the case of human leukocyte antigen (HLA) class I, the main binding energy is provided by the interaction of residues in position 2 and the C-terminus of the peptide with the B and F binding pockets of the MHC molecule, respectively [1-8], although side chains throughout the ligand can have a positive or negative influence on binding capacity. The common chemical specificity of the peptide side chains amongst ligands bound by a specific MHC molecule is termed the binding motif [9].

MHC molecules are extremely polymorphic, and over a thousand allelic variants have already been described at the class I A and B loci. Most of the polymorphism is located in the peptide-binding region, and as a result each variant is believed to bind a unique repertoire of peptide ligands. Despite this polymorphism, HLA class I molecules can be clustered into groups, designated as supertypes, representing sets of molecules that share largely overlapping peptide binding specificity. Each supertype can be described by a supermotif that reflects the broad main anchor motif recognized by molecules within the corresponding supertype. For example, molecules of the A2-supertype share specificity for peptides with aliphatic hydrophobic residues in position 2 and at the C-terminus, while A3-supertype molecules recognize peptides with small or aliphatic residues in position 2 and basic residues at the C-terminus.

The first supertypes were described in the mid-1990’s by our group [10]. Using motifs derived from binding data or the sequencing of endogenously bound ligands, along with simple structural analyses, nine different supertypes were defined [11,12]. Initial peptide binding studies allowed the identification of several peptides with degenerate binding capacity for the A2- [13-15], A3- [16-18] and B7-supertypes [19,20]. In subsequent years, additional binding data have been published to confirm the B44- [21], A1- and A24-supertypes [22].

Over time, a number of different and more sophisticated computational approaches have been implemented by others to classify HLA class I alleles into clusters or supertypes [23-32]. While in general agreement, additional supertypes, representing either sub-clusters of the original nine or completely novel groupings, have been proposed [23,32]. Despite differences in various classification schemes, the concept of HLA supertypes has been effectively used to characterize and identify promiscuously recognized T cell epitopes from a variety of different disease targets, including measles-mumps-rubella [33], SARS [34], EBV [35,36], HIV [37-41], Kaposi’ Sarcoma Associated Human Herpesvirus [42,43], HCV [16,44], HBV [45,46], HPV [47], influenza [48], P. falciparum [49-52], LCMV [53], Lassa virus [54], F. tularensis [55], vaccinia [56-58], and also cancer antigens [59-68]. Supertypes have also been utilized as a component in several approaches and algorithms designed for predicting peptide candidates with degenerate HLA class I binding capacity [69-77]. Finally, supertypes have also been examined as a variable in studies of disease association, rates of susceptibility, and outcome [37,78-86].

The original classification proposed by Sette and Sidney [12] comprised about 100 different class I MHC alleles. However, over the last 10 years, substantially more binding data has been generated, and through the efforts of several seminal MHC-related databases, including SYFPEITHI [87], HLA Ligand [88], FIMM [89] and MHCBCN [90], MHC binding motif information is readily accessible. Up-to-date compilations of MHC sequence data are also readily available in the IMGT database [91]. Herein, we analyze the complete list of alleles available through IMGT (release 2.9) using a simple approach largely based on compilation of published motifs, binding data, and analyses of shared repertoires of binding peptides, in conjunction with clustering based on the primary sequence of the B and F peptide binding pockets. As a result, we now provide updated supertype assignments, with new assignments for about 700 different HLA-A and -B alleles. This will permit the use of our original classification scheme based on current data and its comparison to alternative supertype classification methodologies developed since then.

Results

MHC sequences and peptide binding pockets

α1 and 2 chain residues comprising the B and F peptide-binding pockets of 945 HLA A and B class I molecules were extracted and aligned. Based on crystallographic studies [1,4,6,8], residues 7, 9, 24, 34, 45, 63, 66, 67, 70, and 99 were considered as comprising the B pocket, which engages the peptide residue in position 2. The residues comprising the F pocket, which engages the peptide C-terminal residue, were defined as 74, 77, 80, 81, 84, 95, 97, 114, 116, 123, 133, 143, 146, and 147. Further, for the B pocket we defined the subset of residues 9, 45, 63, 66, 67, 70, and 99 as key residues. In the F pocket, the key residues are those in positions 77, 80, 81 and 116.

Pocket chemical specificity

Next, we defined the broad chemical specificity associated with the different B and F pockets. Table 1 summarizes the amino acid residues generally associated with each particular description of binding specificity. The particular set of residues associated with each use of a descriptor may vary...
with context. While these descriptions largely follow classical textbook definitions, they also consider our own historical use and experience with peptide binding studies. For example, in our analyses we often consider the polar residue Q along with more classically defined aliphatic residues based on observed behavior in MHC peptide binding studies [15,92,93]. It is also important to note that some residues are assigned to multiple categories, as their chemical specificity is compatible with different emphases. For example, L would be considered as both aliphatic and hydrophobic. The chemical specificities defining each supertype are listed in Table 2.

### Table 1: Physiochemical functionality of peptide side chains.

| Pocket specificity          | Associated amino acid residues       |
|-----------------------------|-------------------------------------|
| Acidic                      | DE                                  |
| Aliphatic                   | LIVMQ                               |
| Aromatic                    | FWY                                 |
| Aromatic and aliphatic      | FWYLVIMQ                            |
| Aromatic and basic          | YRK                                 |
| Aromatic and large hydrophobic | FWYLYIM                           |
| Basic                       | RHK                                 |
| Hydrophobic                 | LIVMFVYA                            |
| Large hydrophobic           | FLIM                                |
| Proline                     | P                                   |
| Small                       | AST                                 |
| Small and aliphatic         | ATSVLIMQ                            |
| Small hydrophobic           | AV                                  |
| Small, aliphatic and aromatic | ASTVLIMQFWYY                       |

### Table 2: HLA supertype specificity descriptions.

| Supertype(s) | B pocket specificity | F pocket specificity                     |
|--------------|----------------------|-----------------------------------------|
| A01          | Small and aliphatic  | Aromatic and large hydrophobic          |
| A01, A03     | Small and aliphatic  | Aromatic and basic                      |
| A01, A24     | Small and aliphatic  | Aromatic and large hydrophobic          |
| A02          | Small and aliphatic  | Aliphatic and small hydrophobic         |
| A03          | Small and aliphatic  | Basic                                   |
| A24          | Aromatic and aliphatic | Aromatic, aliphatic and hydrophobic     |
| B07          | Proline              | Aromatic, aliphatic and hydrophobic     |
| B08          | Undefined            | Aromatic, aliphatic and hydrophobic     |
| B27          | Basic                | Aromatic, aliphatic, basic and hydrophobic |
| B44          | Acidic               | Aromatic, aliphatic and hydrophobic     |
| B58          | Small                | Aromatic, aliphatic and hydrophobic     |
| B62          | Aliphatic            | Aromatic, aliphatic and hydrophobic     |

(e.g., A*2902 [22]). Also, because it utilizes a non-canonical main anchor spacing, and as a result appears to have a repertoire that overlaps with other specificities, we have presently separated HLA B*0801 and other alleles sharing sequence and serological antigen similarities, as a separate cluster. While these B*08 alleles utilize a unique mode of peptide binding, it is likely that in most cases their repertoires overlap significantly with other supertypes, especially the B07-supertype (Sidney, Frahm, Brander and Sette, unpublished observations). As a result, we have not defined this a separate supertype.

Next, the available peptide-binding motifs were compiled. In total, 88 different class I motifs were identified. Motif information was derived from our own peptide-binding studies, or from the published scientific literature as compiled in the SYFPEITHI database [87]. The basis for each motif assignment is listed in Tables 3 and 4 for HLA-A and -B alleles, respectively. Corresponding supertype associations were assigned based on the criteria listed in Table 2.

### Pocket analysis and supertype assignments

The amino acids comprising the B and F peptide binding pockets were compiled for all alleles for which complete sequence information was available. The positions of specific MHC residues forming the corresponding pockets are listed above. A reference panel of B and F pockets was generated to include all alleles for which the MHC-peptide binding specificity has been defined (see Tables 3 and 4). These B and F pocket structures, along with the associated alleles and corresponding binding specificity are shown in Tables 5 and 6, respectively.

Next, the B and F pocket structures of alleles whose peptide binding specificity was unknown were compared against the sequences in the reference panel. For each case, an attempt was made to find an exact match with the full set of residues in the corresponding pocket. If a match was identified, the allele was assigned the associated specific-
city shown in Tables 5 and 6. If an exact match could not be found, then a match with a key residue sequence was attempted. If no match was identified, the allele was considered unassigned. For each allele where a match for the full or, secondarily, key residue sequence could be identified for both the B and F pockets with any of the alleles indicated in Tables 3 and 4, a corresponding HLA supertype was assigned.

Of the 945 sequences analyzed, matches at both B and F pockets were found for 764 (80.8%) (Table 7). Notably, for the majority (57%) of alleles not in the reference panel full sequence matches were identified at both the B and F pockets. Figures 1 and 2 indicate the alleles associated with each HLA-A or -B supertype, respectively. Conversely, as an index, Additional file 1 provides the supertype assignment for each of the 945 HLA-A and -B alleles examined, along with their respective B and F pocket structures.

At the HLA-A locus, alleles were fairly evenly distributed amongst the four supertypes we have defined. Considering the alleles with broad, or dual, specificity (i.e., those assigned as A01–A03 or A01–A24), the minimum was 57

| Allele  | Basis for motif | Supertype association |
|---------|----------------|-----------------------|
| A*0101  | Binding assay  | A01                   |
| A*0201  | Binding assay  | A02                   |
| A*0202  | Binding assay  | A02                   |
| A*0203  | Binding assay  | A02                   |
| A*0204  | Pool sequencing/ligand elution | A02 |
| A*0205  | Binding assay  | A02                   |
| A*0206  | Binding assay  | A02                   |
| A*0207  | Binding assay  | A02                   |
| A*0214  | Pool sequencing/ligand elution | A02 |
| A*0217  | Pool sequencing/ligand elution | A02 |
| A*0301  | Binding assay  | A03                   |
| A*1101  | Binding assay  | A03                   |
| A*2301  | Binding assay  | A24                   |
| A*2402  | Binding assay  | A24                   |
| A*2601  | Binding assay  | A01                   |
| A*2602  | Pool sequencing/ligand elution | A01 |
| A*2603  | Pool sequencing/ligand elution | A01 |
| A*2902  | Binding assay  | A01 A24 B44          |
| A*3001  | Binding assay  | A01 A03              |
| A*3002  | Binding assay  | A01                   |
| A*3003  | Pool sequencing/ligand elution | A01 |
| A*3004  | Pool sequencing/ligand elution | A01 |
| A*3101  | Binding assay  | A03                   |
| A*3201  | Binding assay  | A01                   |
| A*3301  | Binding assay  | A03                   |
| A*3303  | Published motif | A03 |
| A*6601  | Pool sequencing/ligand elution | A03 |
| A*6801  | Binding assay  | A03                   |
| A*6802  | Binding assay  | A02                   |
| A*6901  | Pool sequencing/ligand elution | A02 |
| A*7401  | Binding assay  | A03                   |

| Allele  | Basis for motif | Supertype association |
|---------|----------------|-----------------------|
| B*0702  | Binding assay  | B07                   |
| B*0703  | Pool sequencing/ligand elution | B07 |
| B*0705  | Pool sequencing/ligand elution | B07 |
| B*0801  | Binding assay  | B08                   |
| B*0802  | Pool sequencing/ligand elution | B08 |
| B*1402  | Pool sequencing/ligand elution | B27 |
| B*1501  | Binding assay  | B62                   |
| B*1502  | Pool sequencing/ligand elution | B62 |
| B*1503  | Binding assay  | B27                   |
| B*1508  | Pool sequencing/ligand elution | B07 |
| B*1509  | Pool sequencing/ligand elution | B27 |
| B*1510  | Pool sequencing/ligand elution | B27 |
| B*1512  | Pool sequencing/ligand elution | B62 |
| B*1513  | Pool sequencing/ligand elution | B62 |
| B*1516  | Pool sequencing/ligand elution | B58 |
| B*1517  | Pool sequencing/ligand elution | B58 |
| B*1518  | Pool sequencing/ligand elution | B27 |
| B*1801  | Binding assay  | B44                   |
| B*2702  | Binding assay  | B58                   |
| B*2703  | Pool sequencing/ligand elution | B27 |
| B*2704  | Pool sequencing/ligand elution | B27 |
| B*2705  | Binding assay  | B27                   |
| B*2706  | Pool sequencing/ligand elution | B27 |
| B*2707  | Pool sequencing/ligand elution | B27 |
| B*2709  | Pool sequencing/ligand elution | B27 |
| B*3501  | Binding assay  | B07                   |
| B*3503  | Pool sequencing/ligand elution | B07 |
| B*3701  | Pool sequencing/ligand elution | B44 |
| B*3801  | Pool sequencing/ligand elution | B27 |
| B*3901  | Pool sequencing/ligand elution | B27 |
| B*3902  | Pool sequencing/ligand elution | B27 |
| B*3909  | Pool sequencing/ligand elution | B27 |
| B*4001  | Binding assay  | B44                   |
| B*4002  | Binding assay  | B44                   |
| B*4006  | Pool sequencing/ligand elution | B44 |
| B*4201  | Binding assay  | B07                   |
| B*4402  | Binding assay  | B44                   |
| B*4403  | Binding assay  | B44                   |
| B*4501  | Binding assay  | B44                   |
| B*4601  | Pool sequencing/ligand elution | B62 |
| B*4801  | Pool sequencing/ligand elution | B27 |
| B*5101  | Binding assay  | B07                   |
| B*5102  | Pool sequencing/ligand elution | B07 |
| B*5103  | Pool sequencing/ligand elution | B07 |
| B*5201  | Pool sequencing/ligand elution | B62 |
| B*5301  | Binding assay  | B07                   |
| B*5401  | Binding assay  | B07                   |
| B*5501  | Pool sequencing/ligand elution | B07 |
| B*5502  | Pool sequencing/ligand elution | B07 |
| B*5601  | Pool sequencing/ligand elution | B07 |
| B*5701  | Binding assay  | B58                   |
| B*5702  | Pool sequencing/ligand elution | B58 |
| B*5801  | Binding assay  | B58                   |
| B*5802  | Binding assay  | B58                   |
| B*6701  | Pool sequencing/ligand elution | B07 |
| B*7301  | Pool sequencing/ligand elution | B27 |
| B*7801  | Pool sequencing/ligand elution | B07 |
alleles for A24, and the maximum was 95 alleles for A03. 75 alleles were assigned to the A02 supertype, and 80 to the A01 supertype. At the HLA-B locus, the B07-supertype is the largest, with 165 members. 19 alleles were associated with the B08 pattern. As noted above, we hypothesize that alleles in this cluster will likely cross-react with other supertypes (especially B07), so we do not consider this group as a distinct supertype. The B58 supertype was assigned only 22 alleles, making it the smallest supertype. Novel supertypes could not be identified, but several alleles have specificities spanning two different supertypes From Tables 5 and 6, it is evident that not all possible combinations of B and F pocket specificities are present within the reference panel. For example, there are no alleles associated with a preference for acidic residues in position 2 and basic residues at the C-terminus. Hence, we were interested to see if combinations not found in the original supertype assignments were revealed in the present analysis. Interestingly, we did not identify any new combinations. It is possible that new specificities will be identified from amongst the set of 181 alleles for which pocket matches could not be obtained. However, we also note that in the majority of these cases, matches could be made if one allowed a single conservative residue change (e.g., E to D). Thus, based on the analyses done to date, it would appear that the set of supertype specificities currently identified will cover virtually all HLA-A and -B class I MHC alleles.

### Table 5: B pocket structures of reference panel alleles.

| Residues | Key residues | Reference panel allele(s) | Associated supertype(s) | Specificity |
|----------|--------------|---------------------------|-------------------------|-------------|
| YDSVENIFNY | DENIFNY | B*0801-2 | B08 | Undefined |
| YFAVMEKVC | A*0207 | A02 | Small and aliphatic |
| YFAVMEKVC | A*0217 | A02 | Small and aliphatic |
| YFAVMEKVC | A*0201-4 | A02 | Small and aliphatic |
| YFAVMEKHC | FMEKNH | A*0101 | A01 | Small and aliphatic |
| YFAVMEKHC | A*0101 | A01 | Small and aliphatic |
| YFAVMEKHC | A*3201, A*7401 | A01, A03 | Small and aliphatic |
| YFAVMEKHC | A*0301 | A03 | Small and aliphatic |
| YHSVEINS | HTEINS | B*3701 | B44 | Acidic |
| YHSVINS | HTINS | B*1801 | B44 | Acidic |
| YHTVEIC | HEEIC | B*2702-7, B*2709 | B27 | Basic |
| YHTVENC | HENC | B*2702-7 | B27 | Basic |
| YHTVEICS | HKEICS | B*4001-2, B*4006, B*4501 | B44 | Acidic |
| YSAYMEKHC | SMKHC | A*2301, A*2402 | A24 | Aromatic and aliphatic |
| YSAYMEKHC | A*3002-4 | A01 | Small and aliphatic |
| YSAYMEKHC | A*3001 | A01, A03 | Small and aliphatic |
| YTAVMEKHC | TMHC | A*3101 | A03 | Small and aliphatic |
| YTAVMEKHC | A*3301, A*3303 | A03 | Small and aliphatic |
| YTAVMEKHC | A*2902 | A01, A24 | Small, aliphatic and aromatic |
| YYAVNICH | YENICH | B*5001-2, B*5601 | B07 | Proline |
| YYAVNICH | YGNICH | B*5401 | B07 | Proline |
| YYAYMES | YMEIS | B*1501, B*1512 | B62 | Aliphatic |
| YYAYMES | YMEKHC | A*0205, A*0206, A*0214 | A02 | Small and aliphatic |
| YYAYMES | YMEKHC | B*4601 | B62 | Aliphatic |
| YYAVimens | YMENS | B*1516-17, B*5701-2 | B58 | Small |
| YYAVMINS | YMINS | A*1101 | A03 | Small and aliphatic |
| YYAVMINS | YMNINS | B*1508 | B07 | Proline |
| YYAVMINS | YMNINS | B*1502, B*1513 | B62 | Aliphatic |
| YYAVMINS | YMINS | A*2601-3 | B01 | Small and aliphatic |
| YYAVMINS | YMNINS | A*6601, A*6801-2, A*6901 | A02, A03 | Small and aliphatic |
| YYAYTEN | YTEINS | B*5201 | B62 | Aliphatic |
| YYAYTEN | YTNINS | B*5801-2 | B58 | Small |
| YYAYTEN | YTNINS | B*3501, B*3503, B*5101-3, B*5301, B*7801 | B07 | Proline |
| YSYEINS | YEINS | B*1503, B*3902, B*4801 | B27 | Basic |
| YSYEINS | YENIC | B*3909 | B27 | Basic |
| YSYEINS | YENIC | B*1402, B*1509-10, B*1518, B*3801, B*3901 | B27 | Basic |
| YSYEINS | YENIY | B*0703 | B07 | Proline |
| YSYEINS | YENIY | B*0702, B*0705, B*4201, B*6701 | B07 | Proline |
| YYVYKEINS | YKEINS | B*4402-3 | B44 | Acidic |

1. Based on data presented in Saper et al, and Madden (see Methods), the MHC residues utilized as constituting the B pocket are: 7, 9, 24, 34, 45, 63, 66, 67, 70, and 99. Key residues are those in positions 9, 45, 63, 66, 67, 70, and 99.
Table 6: F pocket structures of reference panel alleles.

| Residues1 | Key residues2 | Reference panel allele(s) | Associated supertype(s) | Specificity |
|-----------|--------------|---------------------------|------------------------|-------------|
| DDTLYIIEHYKW | DTLH | A*3001 | A01 A03 | Aromatic and basic |
| DDTLYIIRDYW | DTLH | A*0301, A*1101 | A03 | Basic |
| DDTLYIIRDYW | DTLH | A*3101, A*3301, A*3303, A*7401 | A03 | Basic |
| DDTLYIIRDYW | DTLH | A*6801 | A03 | Basic |
| DDTLYIIRDYW | DTLH | A*6802 | A02 | Aliphatic |
| DDTLYIIRDYW | DTLH | A*6601 | A03 | Basic |
| DDTLYIIRDYW | DTLH | B*2703, B*2705 | B27 | Hydrophobic and basic |
| DDTLYIIRDYW | DTLH | B*2709 | B27 | Large hydrophobic |
| DDTLYIIRDYW | DTLH | B*2707 | B27 | Large hydrophobic |
| DDTLYIIRDYW | DTLH | A*6691 | A02 | Aliphatic |
| DNLNYWTNFYW | GNLF | B*7301 | B27 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | A*2301, A*2402 | A24 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | A*0101 | A01 | Aromatic |
| DNLNYWTNFYW | GNLF | A*2902 | A01 A24 | Aromatic and large hydrophobic |
| DNLNYWTNFYW | GNLF | A*2601 | A01 | Aromatic |
| DNLNYWTNFYW | GNLF | A*2602 | A01 | Aromatic |
| DNLNYWTNFYW | GNLF | A*3201 | A01 | Aromatic |
| DNLNYWTNFYW | GNLF | A*2402 | A24 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | B*0802 | B08 | Hydrophobic |
| DNLNYWTNFYW | GNLF | A*3002-4 | A01 | Aromatic |
| DNLNYWTNFYW | GNLF | A*0101 | A01 | Aromatic |
| DNLNYWTNFYW | GNLF | A*2902 | A01 A24 | Aromatic and large hydrophobic |
| DNLNYWTNFYW | GNLF | B*4601 | B62 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | B*3901-2, B*3909, B*6701 | B07 B27 | Hydrophobic |
| DNLNYWTNFYW | GNLF | B*0702-3 | B07 | Hydrophobic |
| DNLNYWTNFYW | GNLF | B*0705, B*0801, B*4201 | B07 | Hydrophobic |
| DNLNYWTNFYW | GNLF | B*1402 | B27 | Hydrophobic |
| DNLNYWTNFYW | GNLF | B*5401, B*5501-2, B*5601 | B07 | Small hydrophobic |
| DNLNYWTNFYW | GNLF | B*7801 | B07 | Hydrophobic |
| DNLNYWTNFYW | GNLF | B*2706 | B27 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | B*2704 | B27 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | B*2703, B*2705 | B27 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | B*2709 | B27 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | B*2707 | B27 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | A*6691 | A02 | Aliphatic |
| DNLNYWTNFYW | GNLF | A*2601 | A01 | Aromatic |
| DNLNYWTNFYW | GNLF | A*2602 | A01 | Aromatic |
| DNLNYWTNFYW | GNLF | A*3201 | A01 | Aromatic |
| DNLNYWTNFYW | GNLF | B*4601 | B62 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | B*3901-2, B*3909, B*6701 | B07 B27 | Hydrophobic |
| DNLNYWTNFYW | GNLF | B*0702-3 | B07 | Hydrophobic |
| DNLNYWTNFYW | GNLF | B*0705, B*0801, B*4201 | B07 | Hydrophobic |
| DNLNYWTNFYW | GNLF | B*1402 | B27 | Hydrophobic |
| DNLNYWTNFYW | GNLF | B*5401, B*5501-2, B*5601 | B07 | Small hydrophobic |
| DNLNYWTNFYW | GNLF | B*7801 | B07 | Hydrophobic |
| DNLNYWTNFYW | GNLF | B*2706 | B27 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | B*2704 | B27 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | B*2703, B*2705 | B27 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | B*2709 | B27 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | B*2707 | B27 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | A*6691 | A02 | Aliphatic |
| DNLNYWTNFYW | GNLF | A*2601 | A01 | Aromatic |
| DNLNYWTNFYW | GNLF | A*2602 | A01 | Aromatic |
| DNLNYWTNFYW | GNLF | A*3201 | A01 | Aromatic |
| DNLNYWTNFYW | GNLF | B*4601 | B62 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | A*2301, A*2402 | A24 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | A*0101 | A01 | Aromatic |
| DNLNYWTNFYW | GNLF | A*2902 | A01 A24 | Aromatic and large hydrophobic |
| DNLNYWTNFYW | GNLF | B*4601 | B62 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | A*2301, A*2402 | A24 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | A*0101 | A01 | Aromatic |
| DNLNYWTNFYW | GNLF | A*2902 | A01 A24 | Aromatic and large hydrophobic |
| DNLNYWTNFYW | GNLF | B*4601 | B62 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | A*2301, A*2402 | A24 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | A*0101 | A01 | Aromatic |
| DNLNYWTNFYW | GNLF | A*2902 | A01 A24 | Aromatic and large hydrophobic |
| DNLNYWTNFYW | GNLF | B*4601 | B62 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | A*2301, A*2402 | A24 | Large hydrophobic |
| DNLNYWTNFYW | GNLF | A*0101 | A01 | Aromatic |
| DNLNYWTNFYW | GNLF | A*2902 | A01 A24 | Aromatic and large hydrophobic |

1. Based on data presented in Saper et al, and Madden (see Methods), the MHC residues utilized as constituting the F pocket are: 74, 77, 80, 81, 84, 95, 97, 114, 116, 123, 133, 143, 146, and 147. Key residues are those in positions 77, 80, 81, and 116.

2. The key residue pockets DTLH, DTLH and DTLH are each associated with two different, although overlapping, specificity descriptions. In each case, one description is associated with HLA B27 alleles, and in the other with HLA A alleles. In assigning key residue matches, matches were assigned on the basis of corresponding locus.
In addition to A*2902 and A*3001, the present analysis has identified 17 alleles expected to have a specificity overlapping two supertypes. Nine alleles matched pocket preference patterns that would be compatible with both the A01- and A03-supertypes, and another 10 with the A01- and A24-supertypes. In the former case, the association was established on the basis of pocket matches with A*3001. Not surprisingly, several, but not all, of the alleles assigned the A01–A03 specificity also share the A30 antigen. The A01–A24 cross-reactivity pattern was established on the basis of pocket matches with A*2902. All of the alleles associated with the A01–A24 pattern also shared the A29 antigen.

Unexpected supertype assignments

Typically, the supertype assignment for a particular allele follows the predominant assignment for other alleles sharing the same serological antigen. With this in mind, we did identify several instances where the current classification differed from what was expected based on serology or previous classification (Table 8). Of these, the more striking ones involve alleles for which the assigned supertype represents a non-conservative change in the binding specificity from what would have been expected. For example, A*0265 and A*0280, which are associated with the A02 serological antigen, and thus might be expected to have an F pocket specificity for hydrophobic/alkiphatic residues, were found to have an F pocket matching that of A03 supertype alleles, which have a specificity for positively charged residues. Similarly, B*4012 and B*4440 were found to have B27 supertype B pockets, associated with a preference for positively charged residues, when a B44 supertype specificity for acidic residues might have been expected.

As noted above, in the present analysis, and contrary to what was done in the original analysis, we did not make supertype assignments for alleles without an exact or key match by considering conservative substitutions. As a result of this more conservative approach, we now no longer assign supertypes to some alleles previously given a supertype designation (Table 9). In the majority of these cases, however, the original designation would still seem to be a reasonable assumption.

Discussion

In the present study we have attempted to classify almost 1000 HLA-A and -B class I alleles into supertypes. This is nearly a 10-fold increase in the number of alleles compared to our original classification done about a decade ago [12]. Besides providing supertype assignments for considerably more alleles, the present report has attempted to make more transparent how the original "phenomenological" classifications were done. About 80% of the 945 alleles examined were classified into one of the nine supertypes identified previously. Analysis of B and F pocket specificity patterns did not suggest the existence of any novel supertypes.

HLA supertypes do not necessarily demarcate groups of alleles with completely non-overlapping repertoires. A binding repertoire overlapping multiple supertypes has been demonstrated previously, for example, in the cases of A*2902 [22] and A*3001 (see Hamdahl et al., IEDB submission 1000945 [94-96]). In the present study we have identified 17 other alleles that would appear to have specificities that bridge either the A01 and A03 supertypes, or the A01 and A24 supertypes. At the same time, individual peptides can be readily identified that bear a
Figure 1

Supertype classification of HLA-A alleles. The alleles associated with each HLA-A supertype, multiple supertypes, or that are unclassified, are shown. Under each supertype, alleles are group (by color) on the basis of the stringency of selection: experimentally established motif (i.e., reference panel) (green), exact match(es) in the B and F pockets (white), one exact and one key residue pocket match (yellow), key residue match(es) at B and F pockets (grey). Alleles with no match at one or both pockets are listed with red font.

particular supermotif, but that do not bind individual HLA allele members of the supertype, or that bind alleles of other supertypes, even supertypes associated with a different locus. Typically, in the first case, these phenomena result from differences in motif compatibility, perhaps at secondary positions. The second case likely reflects overlap(s) between the supertypes in terms of specificity, although in rare cases binding can be accomplished when no main anchor motif compatibility is apparent.

These observations are exemplified by a large scale analysis of the capacity of a non-redundant set of 252 known EBV and HIV derived epitopes to bind a panel of 30 different HLA class I A and B molecules (Sidney, Frahm, Brander and Sette, unpublished observations). It was found that about 21% of the peptides bearing a specific supermotif bound a given allele in the corresponding supertype with an affinity of 100 nM, or better. By contrast, only in 1% of the cases considered did an allele bind a peptide that did not have the corresponding supermotif. At the same time, it was noted that in the set of peptides utilized 62% (155/252) have motifs associated with 2 or more supertypes. The pattern of binding also followed this general promiscuity. It is also significant to note that when the same library of peptides was examined for recognition in HIV/EBV patients, it was found that ~95% of the epitopes were recognized in individuals not expressing the allele the epitope was originally reported to be
Figure 2
Supertype classification of HLA-B alleles. The alleles associated with each HLA-B supertype, multiple supertypes, or that are unclassified, are shown. Under each supertype, alleles are grouped (by color) on the basis of the stringency of selection as described in the legend to Figure 1.
Table 8: Unexpected and/or revised supertype assignments.

| Allele | Expected or previous | Current |
|--------|----------------------|---------|
| A*0252 | A02                  | A01 A03 |
| A*0265 | A02                  | A03     |
| A*0280 | A02                  | A03     |
| A*3001 | A24                  | A01 A03 |
| A*3002 | A24                  | A01     |
| A*3003 | A24                  | A01     |
| A*3004 | A03                  | A01     |
| A*7410 | A03                  | A01     |
| B*3528 | B07                  | B62     |
| B*3806 | B27                  | B07     |
| B*3807 | B27                  | B07     |
| B*3910 | B27                  | B07     |
| B*3916 | B27                  | B07     |
| B*3917 | B27                  | B07     |
| B*4012 | B44                  | B27     |
| B*4021 | B44                  | B62     |
| B*4406 | B44                  | B07     |
| B*4407 | B44                  | B62     |
| B*4440 | B44                  | B27     |
| B*5518 | B07                  | B27     |

1. Expected based on serological association.

Restricred by, and the promiscuity more often than not involved an allele outside of the supertype associated with originally described restricting allele [97]. Thus, it is apparent that the lines of demarcation between supertypes can be fuzzy from the perspective of both the allelic specificity and the peptide motif.

Restriction outside or across supertypes can also originate from overlaps in supermotifs (e.g., A02 and B62), or for alleles such as B*0801 which do not utilize the typical P2/C terminus anchor spacing. B*0801 utilizes P3 and P5, not P2, and as such may be compatible with several supertypes and alleles. Thus, an A02- or A24-supertype epitope cross-reacting with B*0801 is not an example of a motif “failure”, but merely reflects the fact that the specific peptide has both motifs.

It is important to emphasize that supertypes are based on MHC binding, and that MHC binding alone is not sufficient criteria for T cell recognition. Indeed, hundreds of examples of peptides that bind with remarkably high affinity, but that are not recognized by T cells, have been reported in the literature. We note that even in the best affinity ranges (i.e., IC50 <10 nM), rarely more than 10% of the peptides can be expected to be recognized [98]. Similarly, binding affinity is not necessarily correlative of frequency of recognition [99]. It is true that the trend is towards the most frequently recognized peptides being also the highest affinity binders [38,50,100], but that is not always the case, and there are clearly cases where the dominant epitope has an IC50 of ~100 nM, while several other non-recognized peptides have affinities in the <10 nM range.

It must also be emphasized that membership of an epitope in a supertype is not sufficient to guarantee its recognition by T cells in the context of different MHC alleles. Peptide binding to MHC is an absolute requirement for an epitope to be recognized by T cells. At the same time, many other factors, including protein expression and processing, as well as T cell repertoire and the specific

Table 9: Alleles reclassified as "unassigned".

| Allele | Previous assignment | B pocket match | F pocket match | Comments |
|--------|---------------------|----------------|----------------|----------|
| A*0102 | A01                 | -              | A01 full       | Single difference with A*0101 (F to S) at position 9 in B pocket is non-conservative. |
| A*2404 | A24                 | A24 full       | -              | F pocket key sequence is similar to A02 and B27 (hydrophobic), so A24 is still likely. |
| A*3401 | A03                 | -              | A03 full       | N/K in 63/66 are relatively unique, otherwise similar to A03 in the B pocket. |
| A*4301 | A01                 | -              | A01 full       | Q to N in 63 is only difference with A*2601, so A01 is still likely. |
| B*1301 | B62                 | B62 key        | -              | Unusual F pocket. |
| B*1302 | B62                 | B62 key        | -              | Unusual F pocket. |
| B*1506 | B62                 | -              | B07 B27 B44 B62 full | Single conservative change (Y to F) at position 99 of B pocket, compared to B*1501; B62 still likely. |
| B*1521 | B62                 | -              | B07 B62 full   | S to C change in position 67 is the onlyB pocket difference with B*1501, so B62 is still likely. |
| B*2708 | B27                 | B27 full       | -              | Position 77 T to N change is only difference with b*2704, so B27 is still likely. |
| B*3802 | B27                 | B27 full       | -              | I to T in position 80 of the F pocket is the only difference with B*3801, so B27 is still likely. |
| B*4701 | B44                 | -              | B27 key        | Y to F at position 99 is the only difference with B*4402-3 in the B pocket, so B44 is still likely. |
| B*4901 | B44                 | B44 full       | -              | F full pocket is similar to B*3801 (hydrophobic), so B44 is still likely. |
MHC context, come into play in determining whether a peptide will be an epitope or not, or whether an epitope will be promiscuously recognized within a specific supertype. For example, Goulder and co-workers, studying B7-supertype epitopes, found that differential selection pressure exerted on HIV by CTL targeting identical epitopes, but restricted by distinct HLA alleles from the same supertype, can result in significant functional differences [101]. Macdonald et al., looking at two B44 subtypes described as members of the HLA B44-supertype, reported that a naturally selected dimorphism between the two molecules alters class I structure, peptide repertoire, and T cell recognition [102].

The intent of the current study was to derive an updated classification of HLA class I MHC alleles on the basis of primary anchor specificity. For the vast majority of HLA class I molecules whose binding specificity have been described by crystal structure, pool sequencing or peptide binding studies, the main anchor interactions of the peptide almost invariably involve the MHC B and F pockets, while other pockets likely dictate secondary interactions. This pattern also appears to be true for most macaque and chimpanzee class I alleles studied to date.

There are exceptions, however, and indeed we have not assigned B*08 alleles to a specific supertype in recognition of the fact that these alleles appear to utilize pockets other than the B pocket as a primary anchor contact. For HLA class I molecules in general, the B and F pockets are the most likely main anchor contacts, while other pockets likely dictate secondary interactions. High levels of cross-reactivity have been experimentally demonstrated in the case of 6 supertypes for alleles that vary at secondary pockets [15,17,20-22].

By contrast, in the murine system it is well recognized that other pockets are often the important primary peptide contacts [103-109]. Thus, to utilize the classification approach described here in the context of other species, additional pockets may need to be considered. It is also likely that the further parsing or sub-classifying of supertypes on the basis of secondary interactions can be accomplished.

In the case of HLA-B alleles, F pocket specificity is difficult to correlate with a specific sequence, as a diverse pattern of residues appears to be associated with similar binding specificity. Independent of the residues in the F pocket, most HLA-B alleles seem to bind hydrophobic residues. Thus, assignment of B alleles was primarily driven by the specificity exhibited by the B pocket. On the other hand, it is also possible that greater resolution in the F pocket could be achieved as more data become available to discriminate different preference patterns. For example, in the B7 supertype it is apparent that some alleles, like B*3501, prefer large hydrophobic residues, such as Y. Conversely, other B7 supertype alleles, such as B*5401, seem to prefer small hydrophobic residues, such as A, at the C-terminus. While we have noted these subtle differences in preference [20], in practice we have not found that they significantly impact cross-reactivity between the alleles. This perhaps suggests that the C-terminal anchor in some contexts is less important, and that shared secondary preferences can have a stronger influence on degenerate binding capacity than in other cases. At the same time, it may be necessary to consider additional key residues in the analysis of the F pocket. This is exemplified in the cases of A*2603 and A*0301 which have the same key F pocket residues, but which are associated with much different specificity.

The vast majority of HLA-A and -B alleles fall into one of the 9 supertypes we have described. There are likely reasons for this [110], which include evolutionary relationships, but also constraints and limitations inherent in the epitope processing infrastructure. For example, no allele has been identified to date that binds peptides with D, E, Q or P at the C-terminus, which is in congruence with the preferences of both proteasomal cleavage and TAP transport [111], and an observation that has been applied in the rational design of an in vitro test reagent tool (Pept-Gen) offered as a tool by the Los Alamos HIV Sequence Database [112].

Supertype classification should not be taken to necessarily imply an evolutionary relationship. In some cases this is largely true, as for example in the case of the A2-supertype, where most alleles are associated with the A2 serological antigen. In other cases the relationship is more complicated, such as the gene conversion relationship between the A2, A3 and A68 antigens. This latter example is somewhat of the exception that proves the rule. Supertype associations are based on shared binding specificity, which may result from both common ancestry and convergent evolution [110]. Thus, while alleles within a supertype may have a close evolutionary relationship, that is not a given. Also, alleles (supertypes) sharing specificity at one anchor position may be associated with very disparate specificities at the other.

Other groups have also utilized various methodologies to define supertypes. In general, our classification is in agreement with those derived by other approaches, as compiled by Hertz [32], Lund [23] and Tong [31]. This is not surprising given the good agreement observed between our initial dataset and other classifications, and that the methodology utilized here is not different from the one utilized to derive the original assignments. If there are variations, they usually represent the splitting of a supertype,
or reassignment of individual alleles. As in any classification problem of this kind, there is no absolute truth in supertype assignments. The practical application of supertype classification schemes to identify degenerately binding peptides will ultimately show what classification scheme has the most practical value.

**Conclusion**

The present study represents an update to the HLA class I supertype classification originally described almost a decade ago. Using MHC peptide binding motif data and MHC sequence information that has since become available, supertype associations have now been provided for over 750 HLA-A and -B alleles. In addition, the approach utilized has been made more transparent, allowing others to utilize the classification approach going forward.

**Methods**

**MHC sequences**

The sequences of HLA A and B class I alleles were obtained from the IMGT/HLA Database [91,113], release 2.9. Alleles with incomplete sequences were removed from further analysis. The residues forming the B and F peptide binding pockets were aligned as described in the Results section.

**Peptide binding motif reference panel**

The main anchor peptide binding motifs recognized by HLA-A and -B molecules were compiled from our own data, or as reported at the SYFPEITHI database [87]. The HLA supertype association with each motif was assigned as defined previously [11,12], and/or on the basis of the broad chemical specificity of each supertype indicated in Tables 1 and 2.

**Pocket analysis**

The residues forming the B and F pockets of each allele in the motif reference panel were compiled as a lookup table in Microsoft Excel. Pockets were defined more stringently by considering all residues forming the pocket, or less stringently by considering only a subset of residues, denominated as key residues, hypothesized to be most directly involved in peptide binding. To assign a B and F pocket specificity for each of the remaining alleles, the reference panel was scanned to identify exact pocket sequence matches. If an exact match could be identified, the allele was assigned the corresponding B or F pocket specificity. Supertype assignments were then made by matching the B and F pocket specificity pattern with the supertype descriptions indicated in Table 2. If an exact match for either the full pocket sequence or key residue sequence could not be made at both the B and F pockets, the allele was considered unassigned.

**Abbreviations**

EBV (Epstein-Barr virus); HBV (hepatitis B virus); HCV (hepatitis C virus); HIV (human immunodeficiency virus); HLA (human leukocyte antigen); HPV (human papilloma virus); IEDB (Immune Epitope Database); LCMV (lymphocytic choriomeningitis virus); MHC (major histocompatibility complex); SARS (severe acute respiratory syndrome).

**Authors’ contributions**

JS performed the sequence analyses, determined the supertype assignments and drafted the manuscript. BP participated in the conceptualization and design of the study, and assisted in the preparation of the manuscript. NF and CB assisted in the study design and data interpretation. AS participated in conceptualization of the study and its design, provided interpretation of the data, and helped to draft the manuscript. All authors read and approved the final manuscript.

**Additional material**

**Additional file 1**

[Click here for file](http://www.biomedcentral.com/content/supplementary/1471-2172-9-1-S1.xls)

**HLA supertype classification of HLA A and B class I alleles.** This table provides a complete listing of the HLA A and B alleles examined in the present study. The table is sorted in alphanumeric order to facilitate searching for a specific allele. For each allele, the supertype assignment is indicated, and its corresponding B and F pocket sequences are shown.

**Acknowledgements**

Efforts for this study were supported by funds from NIH-NIAID contracts N01-AI-40023, N01-AI-40024 and HHSN26620040006C to AS.

**References**

1. Guo HC, Jardetzky TS, Garrett TP, Lane WS, Strominger JL, Wiley DC. Different length peptides bind to HLA-Aw68 similarly at their ends but bulge out in the middle. *Nature* 1992, 360:364-6.
2. Silver ML, Guo HC, Strominger JL, Wiley DC. Atomic structure of a human MHC molecule presenting an influenza virus peptide. *Nature* 1992, 360:367-9.
3. Gorga JC, Madden DR, Prendergast JK, Wiley DC, Strominger JL. Crystallization and preliminary X-ray diffraction studies of the human major histocompatibility antigen HLA-B27. *Proteins* 1992, 12:87-90.
4. Madden DR. The three-dimensional structure of peptide-MHC complexes. *Annu Rev Immunol* 1995, 13:587-622.
5. Madden DR, Garbocki DN, Wiley DC. The antigenic identity of peptide-MHC complexes: a comparison of the conformations of five viral peptides presented by HLA-A2. *Cell* 1993, 75:693-708.
6. Madden DR, Gorga JC, Strominger JL, Wiley DC. The three-dimensional structure of HLA-B27 at 2.1 Å resolution suggests a general mechanism for tight peptide binding to MHC. *Cell* 1992, 70:1035-48.
7. Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, Strominger JL, Wiley DC. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature* 1987, 329:512-8.
Doytchinova IA, Guan P, Flower DR: Identifying human MHC supertypes using bioinformatic methods. J Immunol 2004, 172:4314-23.

Reche PA, Reinherz EL: Definition of MHC supertypes through clustering of MHC peptide binding repertoires. LNCS, ICAIRS 2004, 3239:189-196.

Chelvanayagam G: A roadmap for HLA-A, HLA-B, and HLA-C peptide binding specificities. Immunogenetics 1996, 45:15-26.

Kangueane P, Sakharkar MK, Rajaseeger G, Bolisettty S, Sivasekari B, Zhao B, Ravichandran M, Shapiah P, Subbiah S: A framework to sub-type HLA supertypes. Front Biosci 2005, 10:879-86.

Zhang C, Anderson A, DeLisi C: Structural principles that govern the peptide-binding motifs of class I MHC molecules. J Mol Biol 1999, 288:12-22.

Canbäck B, Kajander K, Nurminen J, Nokkala J, Esko T, Ylä-Herttuala S, Sjöblom T, Fagerberg L, Virtanen I, Lundegård C, Wernersson R, Pöysti A, Ovaskainen S: Supertypes of HLA class I molecules and the impact on the immunogenetic landscape. Clin Exp Immunol 2004, 138:90-97.

Stensballe A, Christiansen P, Kjeldsen LR, Aksnes G, Hansen HH, Skovgaard T, Møhl-Nielsen K, Gjorevski N, Nielsen JP, Jørgensen SB, Madsen EM, Borch-Johnsen K, Brunak S: The human MHC class I superfamily: the structure, evolution, and disease associations of the superfamily. J Mol Biol 2004, 339:631-672.

Zhao B, Deng AE, Ren EC, Kato K, Mathura VS, Sakharkar MK, Kangeue P: Compression of functional space in HLA-A sequence diversity. Hum Immunol 2004, 65:647-654.

Tong J, Tan TW, Ranganathan S: In silico grouping of peptide/HLA class I complexes using structural interaction characteristics. Bioinformatics 2007, 23:177-83.

Hertz T, Yanover C: Identifying HLA supertypes by learning distance functions. Bioinformatics 2007, 23(1):48-55.

Ovsyannikova IG, Jacobson RM, Vierkant RA, Rimm AK, Rimm EB, Fraser GL, Allison J, Rimm AF: HLA supertypes and immune responses to measles-mumps-rubella viral vaccine: findings and implications for vaccine design. Vaccine 2007, 25:3090-100.

Sylvester-Hvid C, Nielsen M, Lambeth K, Roder G, Justesen S, Lundegård C, Worning P, Thomsen H, Lund O, Brunak S, et al.: SARS CTL vaccine candidates; HLA supertypes, genome-wide scanning and biochemical validation. Tissue Antigens 2004, 63:395-400.

Boissonnas MC, McKinney D, Anders M, MaWhinney S, Forster J, Crimi C, Southwood S, Sette A, Chernev R, Newman MJ, et al.: Development of a DNA vaccine designed to induce cytotoxic T lymphocyte responses to multiple conserved epitopes in HIV-1. J Virol 2003, 77:15428-15433.

Alfors MA, Livingston B, Rishamwala N, Nguyen PT, Addo MM, Shea A, Newman M, Files J, Sidney J, Wentworth P, et al.: Identification of novel HLA-A2-restricted human immunodeficiency virus type 1-specific cytotoxic T-lymphocyte epitopes predicted by the HLA-A2 supertype peptide-binding motif. J Virol 2001, 75:1301-11.

Lichterfeld KL, Williams M, Mui SK, Shah SS, Mothe BR, Sette A, Kim HS, Chen B, Livingston B, Sette A, et al.: The HLA-A3-like supermotif demonstrates the overlapping peptide-binding repertoire of common HLA molecules. Hum Immunol 1996, 45:79-93.

Sylvester-Hvid C, Nielsen M, Lambeth K, Roder G, Justesen S, et al.: Definition of an HLA-A3-like supermotif: implications for vaccine development. J Infect Dis 1998, 179:1579-1590.
Page 14 of 15

BMC Immunology 2008, 9:1
http://www.biomedcentral.com/1471-2172/9/1

related protein-2 by CTL in the context of HLA-A31 and -A33. J Immunol 1998, 160:8907-9.

47. Iwashima I, Tsai Y, Nakamura S, Takesako K, Sette A, Celis E: Identification of HLA-A3-restricted cytotoxic T lymphocyte epitopes from carcinoembryonic antigen and HER-2/neu by primary in vitro immunization with peptide-pulsed dendritic cells. Cancer Res 1999, 59:431-5.

48. Matsuoka S, Takedatsu H, Saada T, Azuma K, Ishihara Y, Komohara Y, Noguchi M, Shichijo S, Itoh K, Harada M: New peptide vaccine candidates for epithelial cancer patients with HLA-A3 superfamily alleles. J Immunother (1997) 2002, 25:274-81.

49. Mimmó T, Matsuoka S, Takedatsu H, Tanaka M, Noguchi M, Uemura H, Itoh K, Harada M: Identification of SART3-derived peptides having the potential to induce cancer-reactive cytotoxic T lymphocytes from prostate cancer patients with HLA-A3 superfamily alleles. Cancer Immunol Immunother 2007, 56:689-98.

50. Takedatsu H, Shichijo K, Sawai Y, Sata M, Itoh K: Identification of HLA-A1, -A11+, and -A33+ cancer patients. Clin Cancer Res 2004, 10:1112-20.

51. Matsuoka S, Takedatsu H, Yao A, Tanaka M, Noguchi M, Itoh K, Harada M: Identification of peptide vaccine candidates for prostate cancer patients with HLA-A3 superfamily alleles. Clin Cancer Res 2005, 11:6933-43.

52. Brusic V, Petrovsky N, Zhang G, Bajic VB: Prediction of promiscuous peptides that bind HLA class I molecules. Immunol Cell Biol 2002, 80:280-5.

53. Khan AM, Miotti O, Heinig AT, Salmon J, Srinivasan KN, Nascimento Ej, Marques ET Jr, Brusic V, Tan TW, August J: A systematic bioinformatics approach for selection of epitope-based vaccine targets. Cell Immunol 2006, 244:1-17.

54. Reche PA, Reinherz EL: PEVPAC: a web server for multi-epitope vaccine design and development based on the prediction of supertypic HLA ligands. Nucleic Acids Res 2005, 33:W138-42.

55. Zhe D, Udaka K, Vazquez A, Rottnek F: Comparison of HIV-1 superpeptide binding affinity for healthy controls. Antigenic Peptides. Bioinformatics 2006, 22:1648-55.

56. Zhang GL, Khan AM, Srinivasan KN, August JT, Brusic V: MULTI-PRED: a computational system for prediction of promiscuous HLA binding peptides. Nucleic Acids Res 2005, 33:W172-9.

57. Hogan KT, Inchon JH, Chen TF, Babcock JA, Shabaker CJ, Hunt DF, Singhoff CL Jr: Use of selected reaction monitoring mass spectrometry for the detection of specific MHC class I peptide antigens on A3 supertype family members. Cancer Immunol Immunother 2003, 54:359-71.

58. Toyoshinova I, Flower DR: The HLA-A2-supermotif: a Qsar definition. Origi Biomol Chem 2003, 1:2648-54.

59. Gn P, Toyoshinova I, Flower DR: A comparative molecular similarity indices (CoMSIA) study of peptide binding to the HLA-A3 superefamily. Bioorg Med Chem 2003, 11:2307-11.

60. Gu P, Toyoshinova I, Flower DR: HLA-A3 supermotif defined by quantitative structure-activity relationship analysis. Protein Eng 2003, 16:11-8.

61. MacDonald KS, Fowke KR, Kimani J, Dunand VA, Nagelkerke NJ, Ball TB, Ouyang J, Njagi E, Gaur LK, Brunham RC, et al.: Influence of HLA superpeptide on susceptibility and resistance to human immunodeficiency virus type 1 infection. J Infect Dis 2000, 181:581-9.

62. MacDonald KS, Embree JE, Nagelkerke NJ, Castillo J, Ramhadin S, Sjenga O, Ouyang J, Ndinya-Achola J, Barber BH, Bwalya JJ, et al.: The HLA A2B602 superpeptide is associated with reduced risk of perinatal human immunodeficiency virus type 1 transmission. J Infect Dis 2001, 183:503-6.

63. MacDonald KS, Matukas L, Embree JE, Fowke K, Kimani J, Nagelkerke NJ, Ouyang J, Kiama P, Kaul R, Lusher MA, et al.: Human leukocyte antigen superpeptides and immune susceptibility to HIV-1 infection. Immunol Lett 2001, 79:151-7.

64. Stephens HA: HIV-1 diversity versus HLA class I polymorphism. Trends Immunol 2005, 26:41-7.

65. Balamurugan A, Sharma SK, Mehra NK: Human leukocyte antigen class I superpeptides influence susceptibility and severity of tuberculosis. J Infect Dis 2004, 189:805-11.

66. Lopez-Vazquez A, Mina-Blanco A, Martinez-Borra J, Njoo PB, Suarez-Alvarez B, Blanco-Gelaz MA, Gonzalez S, Rodrigo L, Lopez-Larrea C: Interaction between KIR3DL1 and HLA-B*57 superpeptide.

67. Wang RF, Johnston SL, Southwood S, Sette A, Rosenberg SA: Recognition of an antigenic peptide derived from tyrosinase-related protein-2 by CTL in the context of HLA-A31 and -A33. J Immunol 1998, 160:8907-9.
