# Chapter 2

## IMGT Standardization for Molecular Characterization of the T Cell Receptor/Peptide/MHC Complexes

Quentin Kaas,[1] Elodie Duprat,[1] Guillaume Tourneur,[1] and Marie-Paule Lefranc[1,2]

[1] IMGT®, the International ImMunoGeneTics Information System®, Université Montpellier II, Laboratoire d'ImmunoGénétique Moléculaire LIGM, UPR CNRS 1142, Institut de Génétique Humaine IGH, 141 rue de la Cardonille, 34396 Montpellier Cedex 5, France

[2] Institut Universitaire de France, 103 Boulevard Saint-Michel, 75005 Paris, France, Marie-Paule.Lefranc@igh.cnrs.fr

**Abstract.** One of the key elements in the adaptive immune response is the presentation of peptides by the major histocompatibility complex (MHC) to the T cell receptors (TR) at the surface of T cells. The characterization of the TR/peptide/MHC trimolecular complexes (TR/pMHC) is crucial to the fields of immunology, vaccination, and immunotherapy. In order to facilitate data comparison and cross-referencing between experiments from different laboratories whatever the receptor, the chain type, the domain, or the species, IMGT®, the international ImMunoGeneTics information system (http://imgt.cines.fr), has developed IMGT-ONTOLOGY, the first ontology in immunogenetics and immunoinformatics. In IMGT/3Dstructure-DB, the IMGT three-dimensional structure database, the molecular characterization of the TR/pMHC is made according to the IMGT Scientific chart rules that are based on the IMGT-ONTOLOGY concepts. IMGT/3Dstructure-DB provides the standardized IMGT gene and allele names (CLASSIFICATION), the standardized IMGT labels (DESCRIPTION), and the IMGT unique numbering (NUMEROTATION). As the IMGT structural unit is the domain, amino acids at conserved positions always have the same number in the IMGT® databases, tools, and Web resources. For the TR α and β chains, the amino acids in contact with the peptide/MHC (pMHC) are defined according to the IMGT unique numbering for V-DOMAIN. The MHC chain cleft that binds the peptide is formed by two groove domains (G-DOMAIN), each one comprising four antiparallel β strands and one α helix. The IMGT unique numbering for G-DOMAIN applies both to the first two domains (G-ALPHA1 and G-ALPHA2) of the MHC class I α chain, and to the first domain (G-ALPHA and G-BETA) of the MHC class II α chain and β chain, respectively. Based on the IMGT unique numbering, we defined 11 contact sites for the analysis of the pMHC contacts. The TR/pMHC contact description, based on the IMGT numbering, can be queried in the IMGT/StucturalQuery tool, at http://imgt.cines.fr.

## 2.1 Introduction

T cells are implicated in the specific immune response against a stress of viral, bacterial, fungal, or tumoral origin. They identify antigenic peptides presented by the major histocompatibility complex (MHC) cell surface glycoproteins. The recognition is

carried out by the T cell receptor complex (TcR), a multisubunit transmembrane surface complex made up of a T cell receptor (TR) and of the CD3 chains, that is associated, in the immunological synapse, to the CD4 or CD8 coreceptors, to the CD28 and CTLA-4 costimulatory proteins, to the CD2 adhesion molecule, and to intracellular kinases (Lefranc and Lefranc 2001). The TR directly binds the peptide/MHC complex (pMHC), and activates the T cell through interactions with the CD3 and other components of the TcR (Vasmatzis, Cornette, Sezerman, and DeLisi 1996a; Sim, Zerva, Greene, and Gascoigne 1996; Kjer-Nielsen, Clements, Purcell, Brooks, Whisstock, Burrows, McCluskey, and Rossjohn 2003). Three-dimensional (3D) structures of the TR, pMHC, and TR/pMHC complexes provide an atomic description of their interactions (Kaas, Ruiz, and Lefranc 2004; Kaas and Lefranc 2005).

Since 1989, IMGT®, the international ImMunoGeneTics information system® (Lefranc, Giudicelli, Kaas, Duprat, Jabado-Michaloud, Scaviner, Ginestoux, Clément, Chaume, and Lefranc 2005c), http://imgt.cines.fr, has offered standardized genetic and structural data on immunoglobulins (IG), TR, and MHC, and on related proteins of the immune system (RPI) that belong to the immunoglobulin superfamily (IgSF) and to the MHC superfamily (MhcSF). In order to facilitate data comparison and cross-referencing between experiments from different laboratories whatever the receptor, the chain type, the domain, or the species, IMGT® has developed IMGT-ONTOLOGY (Giudicelli and Lefranc 1999), the first ontology in immunogenetics and immunoinformatics.

Based on the IMGT-ONTOLOGY concepts, the IMGT Scientific chart provides the controlled vocabulary and the annotation rules necessary for the identification, the description, the classification, and the numbering of the IG, TR, MHC, and RPI (Lefranc 2004a; Lefranc, Giudicelli, Ginestoux, Bosc, Folch, Guiraudou, Jabado-Michaloud, Magris, Scaviner, Thouvenin, Combres, Girod, Jeanjean, Protat, Monod, Duprat, Kaas, Pommié, Chaume, and Lefranc 2004b; Lefranc, Clément, Kaas, Duprat, Chastellan, Coelho, Combres, Ginestoux, Giudicelli, Chaume, and Lefranc 2005a). The IDENTIFICATION concept refers to the IMGT standardized keywords that are essential for the sequence and 3D structure assignments. The DESCRIPTION concept provides the IMGT standardized labels used to describe structural and functional regions that compose IG, TR, MHC, and RPI sequences and 3D structures. Standardized labels have also been defined to characterize the three-dimensional assembly of domains and chains. The CLASSIFICATION concept provides immunologists and geneticists with a standardized nomenclature per locus and per species. The human IG and TR gene nomenclature elaborated by IMGT was approved by the Human Genome Organisation (HUGO) Nomenclature Committee, HGNC (Wain, Bruford, Lovering, Lush, Wright, and Povey 2002), in 1999. The mouse IG and TR gene names with IMGT reference sequences were provided by IMGT to HGNC and to the Mouse Genome Database (MGD; Blake, Richardson, Bult, Kadin, and Eppig 2003) in July 2002. The NUMEROTATION concept provides the IMGT unique numbering for the IG and TR V-DOMAIN and the V-LIKE-DOMAIN of the IgSF proteins other than IG or TR (Lefranc, Pommié, Ruiz, Giudicelli, Foulquier, Truong, Thouvenin-Contet, and Lefranc 2003b), and for the IG and TR C-DOMAIN and the C-LIKE-DOMAIN of the IgSF proteins other than IG or TR (Lefranc, Pommié, Kaas, Duprat, Bosc,

Guiraudou, Jean, Ruiz, Da Piedade, Rouard, Foulquier, Thouvenin, and Lefranc 2005d). An IMGT unique numbering has also been set up for the MHC G-DOMAIN and the G-LIKE-DOMAIN of the MhcSF proteins other than MHC (Lefranc, Duprat, Kaas, Tranne, Thiriot, and Lefranc 2005b).

The IMGT standardization has allowed the construction of a unique frame for the comparison of the TR, peptide, and MHC interactions in the different resources provided by the IMGT® information system. IMGT/3Dstructure-DB (Kaas et al. 2004), the IMGT structural database, is used with the IMGT sequence databases, IMGT/LIGM-DB (Lefranc 2003a; Giudicelli, Ginestoux, Folch, Jabado-Michaloud, Chaume, and Lefranc 2006) and IMGT/MHC-DB (Robinson, Waller, Parham, de Groot, Bontrop, Kennedy, Stoehr, and Marsh 2003); the IMGT gene database, IMGT/GENE-DB (Giudicelli, Chaume, and Lefranc 2005); the IMGT tools for sequence analysis, IMGT/V-QUEST (Giudicelli, Chaume, and Lefranc 2004), IMGT/JunctionAnalysis (Yousfi Monod, Giudicelli, Chaume, and Lefranc 2004); and the IMGT tool for 3D structure analysis, IMGT/StructuralQuery (Kaas et al. 2004), to explore the TR and MHC conserved structural features. In this paper, we describe the IMGT standardized rules that have been set up for the molecular characterization of the TR/pMHC complexes. Coordinate files are from IMGT/3Dstructure-DB (Kaas et al. 2004), http://imgt.cines.fr, with original crystallographic data from the Protein Data Bank, PDB (Berman, Westbrook, Feng, Gilliland, Bhat, Weissig, Shindyalov, and Bourne 2000). Eleven IMGT pMHC contact sites were defined (C1 to C11) which can be used to compare pMHC interactions (Kaas and Lefranc 2005).

## 2.2 T Cell Receptor/Peptide/MHC 3D Structures and IMGT Standardization

IMGT/3Dstructure-DB (Table 1) contains 18 TR/pMHC structures: 14 (12 TR/pMHC-I and 2 TR/pMHC-II) with complete extracellular regions of the α-β TR (TR-ALPHA_BETA) and 4 structures with an Fv variable fragment (FV-ALPHA_BETA). The α-β TR chains, TR-ALPHA and TR-BETA, are described with standardized IMGT labels in Fig. 1.

The references for the 18 TR/pMHC 3D structures are: 1ao7 (Garboczi, Ghosh, Utz, Fan, Biddison, and Wiley 1996), 1qrn, 1qse, 1qsf (Ding, Baker, Garboczi, Biddison and Wiley 1999), 1bd2 (Ding, Smith, Garboczi, Utz, Biddison, and Wiley 1998), 1oga (Stewart-Jones, McMichael, Bell, Stuart, and Jones 2003), 1mi5 (Kjer-Nielsen et al. 2003), 1lp9 (Buslepp, Wang, Biddison, Appella, and Collins 2003), 1g6r (Degano, Garcia, Apostolopoulos, Rudolph, Teyton, and Wilson 2000), 1jtr, 1mwa (Luz, Huang, Garcia, Rudolph, Apostolopoulos, Teyton, and Wilson 2002), 2ckb (Garcia, Degano, Pease, Huang, Peterson, Teyton, and Wilson 1998), 1fo0 (Reiser, Darnault, Guimezanes, Gregoire, Mosser, Schmitt-Verhulst, Fontecilla-Camps, Malissen, Housset, and Mazza 2000), 1nam (Reiser, Darnault, Gregoire, Mosser, Mazza, Kearney, van der Merwe, Fontecilla-Camps, Housset, and Malissen 2003), 1kj2 (Reiser, Gregoire, Darnault, Mosser, Guimezanes, Schmitt-Verhulst, Fontecilla-Camps, Mazza, Malissen, and Housset 2002), 1fyt (Hennecke, Carfi, and

Wiley 2000), 1j8h (Hennecke and Wiley 2002), 1d9k (Reinherz, Tan, Tang, Kern, Liu, Xiong, Hussey, Smolyar, Hare, Zhang, Joachimiak, Chang, Wagner, and Wang 1999).

**Table 1.** TR/peptide/MHC complexes in IMGT/3Dstructure-DB (Kaas et al. 2004), http://imgt.cines.fr. Sp, species; Hs, *Homo sapiens*; Mm, *Mus musculus*; L, length in amino acids. Fourteen 3D structures (12 TR/pMHC-I and 2 TR/pMHC-II) correspond to TR receptors (TR-ALPHA_BETA). Four 3D structures (1d9k, 1fo0, 1kj2, and 1nam) correspond to an Fv variable fragment (FV-ALPHA_BETA). Gene and allele names are according to IMGT/GENE-DB (Giudicelli et al. 2005) for human and mouse TR, to IMGT/HLA-DB (Robinson et al. 2003) for human MHC, and to IMGT for mouse MHC. Amino acid sequences of the TR V-DOMAINs and MHC G-DOMAINs are reported in Figs. 3 and 4, respectively. H2-K1*01 encodes H2-K1b, H2-AB*02 and H2-AA*02 encode I-Abk and I-Aak, respectively. Lengths of the CDR-IMGT are according to Lefranc et al. (2003b).

(A) TR/pMHC-I

| | T cell receptor | | | | Peptide | | MHC |
|---|---|---|---|---|---|---|---|
| 3D | Name | Sp. | V-DOMAIN genes | CDR-IMGT | Sequence | L | Gene and allele |
| 1ao7 | A6 | Hs | TRAV12-2-TRAJ24 | [6.6.11] | LLFGYPVYV | 9 | HLA-A*0201 |
| | | | TRBV6-5-TRBD2-TRBJ2-7 | [5.6.14] | | | |
| 1qrn | A6 | | | | LLFGY<u>A</u>VYV | 9 | |
| 1qse | A6 | | | | LLFGYP<u>R</u>YV | 9 | |
| 1qsf | A6 | | | | LLFGYPV<u>A</u>V | 9 | |
| 1bd2 | B7 | Hs | TRAV29/DV5-TRAJ54 | [6.6.10] | LLFGYPVYV | 9 | HLA-A*0201 |
| | | | TRBV6-5-(TRBD2)-TRBJ2-7 | [5.6.13] | | | |
| 1oga | JM22 | Hs | TRAV27-TRAJ42 | [5.6.10] | GILGFVFTL | 9 | HLA-A*0201 |
| | | | TRBV19-(TRBD2)-TRBJ2-7 | [5.6.11] | | | |
| 1mi5 | LC13 | Hs | TRAV26-2-TRAJ52 | [7.4.14] | FLRGRAYGL | 9 | HLA-B*0801 |
| | | | TRBV7-8-(TRBD2)-TRBJ2-7 | [5.6.11] | | | |
| 1lp9 | 12.2 | Mm | TRAV12D-2-TRAJ50 | [6.6.13] | ALWGFFPVL | 9 | HLA-A*0201 |
| | | | TRBV13-3-(TRBD2)-TRBJ2-7 | [5.6.11] | | | |
| 1g6r | 2C | Mm | TRAV9-4-TRAJ35 | [6.7.10] | SIYRYYGL | 8 | H2-K1*01 |
| | | | TRBV13-2-(TRBD2)-TRBJ2-4 | [5.6.9] | | | |
| 1jtr | 2C | | | | EQYKFYSV | 8 | |
| 2dkb | 2C | | | | EQYKFYSV | 8 | |
| 1mwa | 2C | | | | EQYKFYSV | 8 | |

| | T cell receptor | | | | Peptide | | MHC |
|---|---|---|---|---|---|---|---|
| 3D | Name | Sp. | V-DOMAIN genes | CDR-IMGT | Sequence | L | Gene and allele |
| 1fo0 | BM3.3 | Mm | TRAV16-TRAJ32 | [7.7.14] | INFDFNTI | 8 | H2-K1*01 |
| | | | TRBV1-TRBD1-TRBJ1-3 | [6.6.12] | | | |
| 1nam | BM3.3 | | | | RGYVYQGL | 8 | |
| 1kj2 | KB5-C20 | Mm | TRAV14-1-TRAJ15 | [6.6.11] | KVITFIDL | 8 | H2-K1*01 |
| | | | TRBV1-TRBD2-TRBJ2-3 | [6.6.16] | | | |

(B) TR/pMHC-II

| | T cell receptor | | | | Peptide | | MHC |
|---|---|---|---|---|---|---|---|
| 3D | Name | Sp. | V-DOMAIN genes | CDR-IMGT | Sequence | L | Gene and allele |
| 1fyt | HA1.7 | Hs | TRAV8-4-TRAJ48 | [6.7.13] | PKYVKQNT LKLAT | 13 | HLA-DR A *0101 |
| | | Hs | TRBV28-TRBD1-TRBJ1-2 | [5.6.12] | | | HLA-DR B1*0101 |
| 1j8h | HA1.7 | | | | PKYVKQNTL KLAT | 13 | HLA-DR A*0101 |
| | | | | | | | HLA-DR B1*0401 |
| 1d9k | D10 | Mm | TRAV14D-2-TRAJ4 | [6.6.10] | GNSHRGAIE WEGIESG | 16 | H2-AA *02 |
| | | | TRBV13-2-TRBD2-TRBJ2-1 | [5.6.11] | | | H2-AB *02 |

Each complete TR chain comprises an extracellular region made up of a variable domain and a constant domain (V-ALPHA and C-ALPHA for the α chain, V-BETA and C-BETA for the β chain) (Fig. 1), a connecting region, a transmembrane region, and a very short intracytoplasmic region. The MHC-I is formed by the association of a heavy chain (I-ALPHA) and a light chain (β-2-microglobulin B2M, Fig. 1). The MHC-II is a heterodimer formed by the association of an α chain (II-ALPHA) and a β chain (II-BETA). The I-ALPHA chain of the MHC-I, and the II-ALPHA and II-BETA chains of the MHC-II comprise an extracellular region made of three domains for the MHC-I and of two domains for the MHC-II, a connecting region, a transmembrane region, and an intracytoplasmic region. The I-ALPHA chain comprises two groove domains (G-DOMAIN), G-ALPHA1 [D1] and G-ALPHA2 [D2], and a C-LIKE domain [D3]. The B2M corresponds to a single C-LIKE domain. The II-ALPHA chain and the II-BETA chain each comprise two domains, G-ALPHA [D1] and C-LIKE [D2],

and G-BETA [D1] and C-LIKE [D2]. Only the extracellular region that corresponds to these domains has been crystallized (Fig. 1). The TR V-DOMAINs and MHC G-DOMAINs that are directly involved in TR/pMHC interactions are described in the next sections.

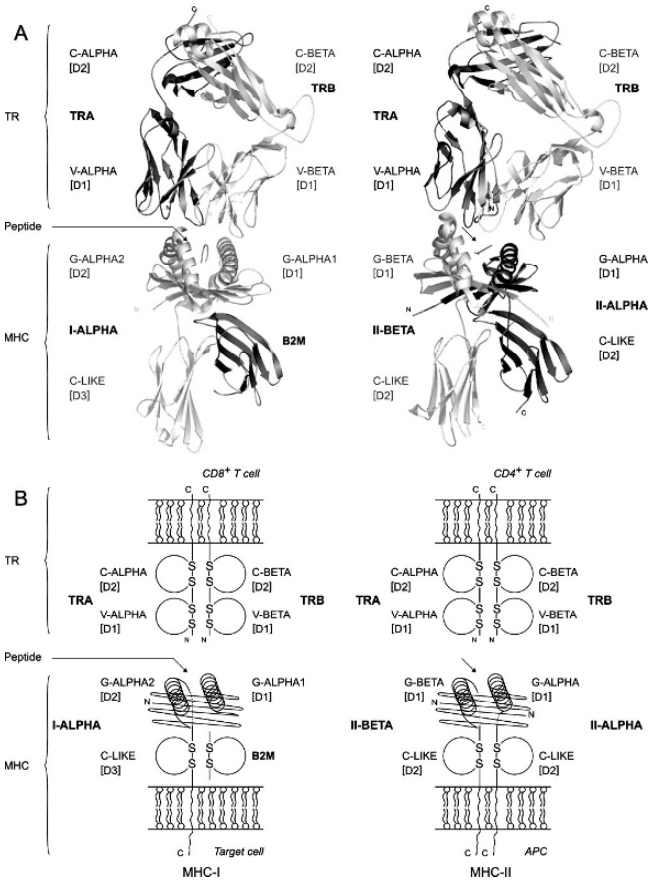


**Fig. 1.** T cell receptor/peptide/MHC complexes with MHC class I (TR/pMHC-I) and MHC class II (TR/pMHC-II). [D1], [D2] and [D3] indicate the domains. (A) 3D structures of TR/pMHC-I (1oga) and TR/pMHC-II (1j8h). (B) Schematic representation of TR/pMHC-I and TR/pMHC-II. The TR (TR-ALPHA and TR-BETA) chains, the MHC-I (I-ALPHA and β-2-microglobulin B2M) chains and the MHC-II (II-ALPHA and II-BETA) chains are shown with the extracellular domains (V-ALPHA and C-ALPHA for the TR-ALPHA chain; V-BETA and C-BETA for the TR-BETA chain; G-ALPHA1, G-ALPHA2 and C-LIKE for the I-ALPHA chain; C-LIKE for B2M; G-ALPHA and C-LIKE for the II-ALPHA chain; II-BETA and C-LIKE for the II-BETA chain), and the connecting, transmembrane and cytoplasmic regions. Arrows indicate the peptide localization in the G-DOMAIN groove. The MHC G-DOMAINs and TR V-DOMAINs are likely to be in a diagonal rather than in a vertical position relative to the cell surface (Wang, Meijers, Xiong, Liu, Sakihama, Zhang, Joachimiak and Reinherz 2001; Wang and Reinherz 2002).

### 2.2.1 TR V-DOMAINs

The V-DOMAINs have an immunoglobulin fold, that is an antiparallel β sheet sandwich structure with nine strands (Lefranc et al. 2003b; Lesk and Chothia 1982), the A, B, E, and D strands being on one sheet, and the G, F, C, C', and C" strands on the other sheet. These strands are indicated in the IMGT Colliers de Perles (Fig. 2) and in the IMGT Protein displays (Fig. 3).

IMGT Colliers de Perles are IMGT 2D graphical representations based on the IMGT unique numbering. The IMGT Colliers de Perles of TR V-DOMAINs are based on the IMGT unique numbering for V-DOMAIN and V-LIKE-DOMAIN (Lefranc et al. 2003b) and can be displayed on one layer or on two layers. IMGT Colliers de Perles of the V-ALPHA and V-BETA domains from 1ao7 (Garboczi et al. 1996) are shown as examples in Fig. 2. The IMGT Protein display (Fig. 3) shows the amino acid sequences of the different V-ALPHA and V-BETA domains found in TR/pMHC (Table 1).

The V-ALPHA and V-BETA domains share main conserved characteristics of the V-DOMAIN which are the disulfide bridge between cysteine 23 (1st-CYS) and cysteine 104 (2nd-CYS), and the other hydrophobic core residues tryptophan 41 (CONSERVED-TRP) and leucine (or hydrophobic) 89 (Lefranc et al. 2003b) (Figs. 2 and 3). The A strand comprises positions 1 to 15, B strand positions 16 to 26, C strand positions 39 to 46, C' strand positions 47 to 55, C" strand positions 66 to 74, D strand positions 75 to 84, E strand positions 85 to 96, F strand positions 97 to 104, and G strand positions 118 to 128 (Lefranc et al. 2003b). Compared to the general V-DOMAIN 3D structure, the V-ALPHA domains have shorter C" and D strands at the C" D turn.

The three hypervariable loops or complementarity determining regions (CDR) of each V-DOMAIN are involved in the pMHC recognition. The CDR1-IMGT comprises positions 27 to 38, the CDR2-IMGT positions 56 to 65, and the CDR3-IMGT positions 105 to 117 (Lefranc et al. 2003b). The CDR3-IMGT corresponds to the junction resulting from the V-J and V-D-J rearrangement, and is more variable in sequence and length than the CDR1-IMGT and CDR2-IMGT that are encoded by the V-REGION only (Lefranc and Lefranc 2001). Lengths of the CDR-IMGT are shown separated by dots between brackets (Lefranc et al. 2003b). Lengths of the CDR-IMGT from available TR/pMHC 3D structures are reported in Table 1, together with the names of the V, D, and J genes (Lefranc and Lefranc 2001).

For example, 1ao7 [6.6.11] V-ALPHA means that in the V-ALPHA domain of 1ao7, CDR1-IMGT has a length of 6 amino acids, CDR2-IMGT a length of amino acids, and CDR3-IMGT a length of 11 amino acids. The V-ALPHA CDR3-IMGT results from the TRAV12-2–TRAJ24 rearrangement (Table 1, Fig. 3). In the same way, 1ao7 [5.6.14] V-BETA means that in the V-BETA domain of 1ao7, CDR1-IMGT, CDR2-IMGT, and CDR3-IMGT have a length of 5, 6, and 14 amino acids, respectively (Lefranc et al. 2003b). The V-BETA CDR3-IMGT results from the TRBV6-5–TRBD2–TRBJ2-7 rearrangement (Table 1, Fig. 3).

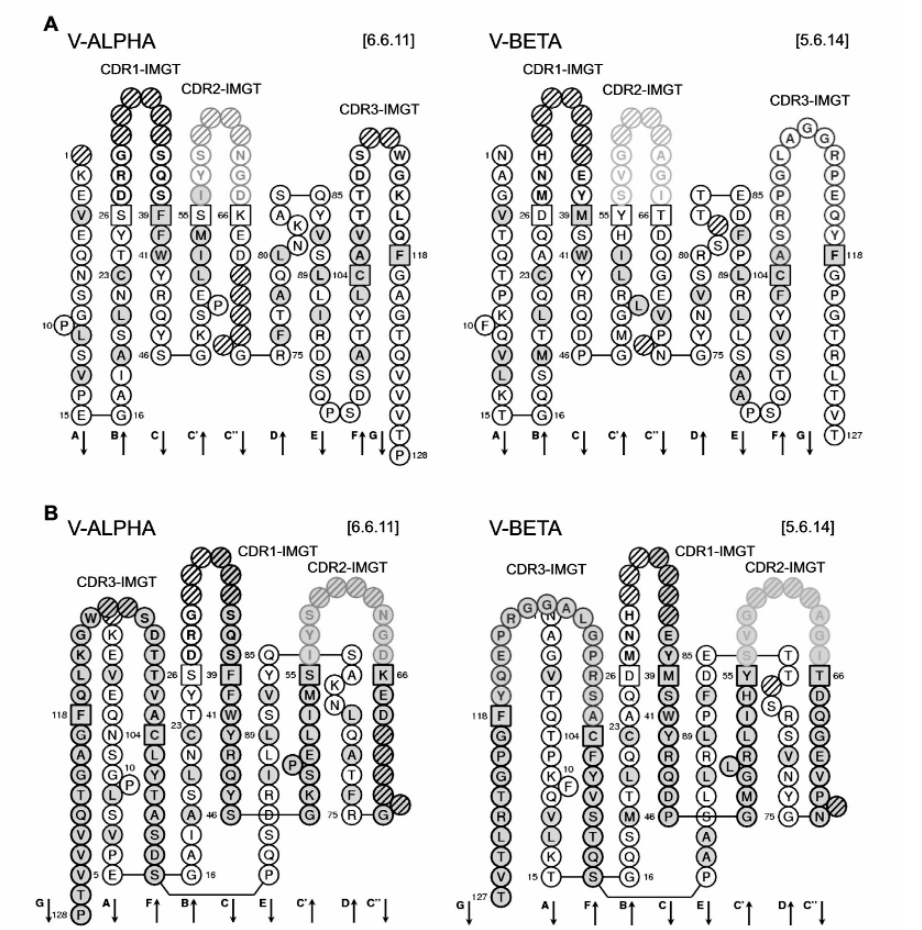


**Fig. 2.** IMGT Colliers de Perles of the V-ALPHA and V-BETA domains from 1ao7 (IMGT/ 3Dstructure-DB, http://imgt.cines.fr) (A) on one layer (B) on two layers. Amino acids are shown in the one-letter abbreviation. Hydrophobic amino acids (hydropathy index with positive value) and tryptophan (W) found at a given position in more than 50% of analysed IG and TR sequences are shown. The CDR-IMGTs are limited by amino acids shown in squares, which belong to the neighbouring FR-IMGT and represent anchor positions. Hatched circles correspond to missing positions according to the IMGT unique numbering (Lefranc et al. 2003b). Arrows indicate the direction of the $\beta$ sheets.

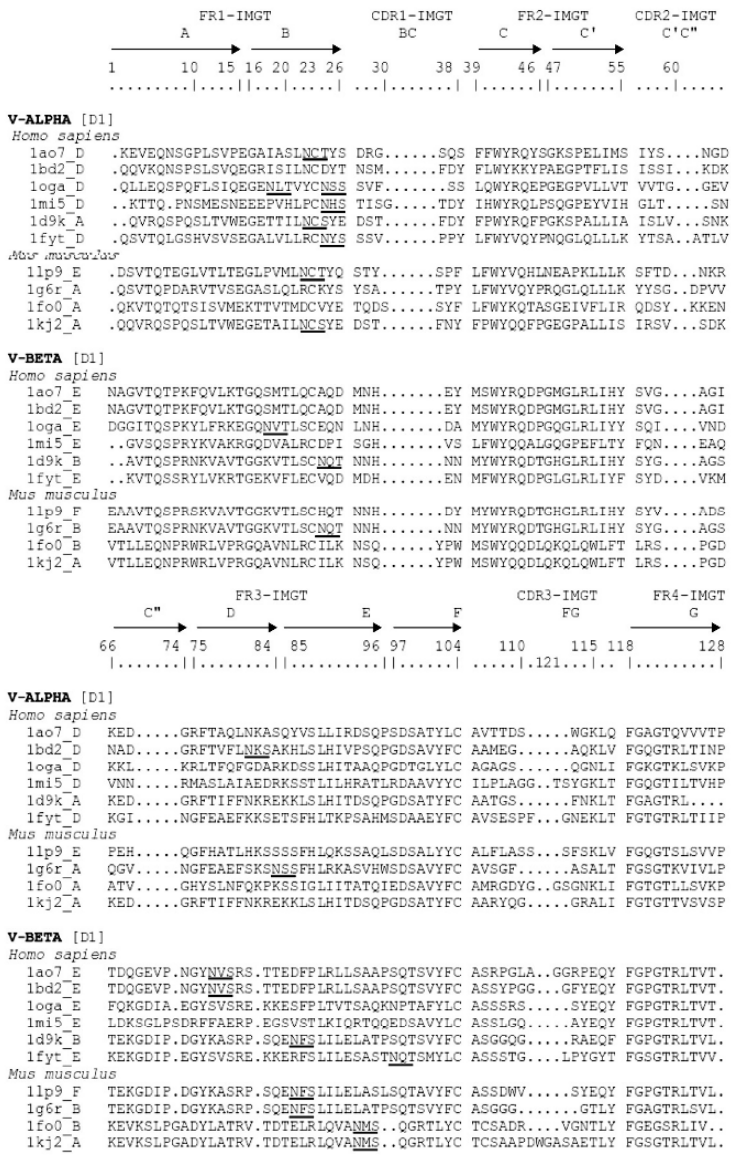


**Fig. 3.** Protein display of the TR V-ALPHA and V-BETA domains found in the TR/pMHC complexes in IMGT/3Dstructure-DB (Kaas et al. 2004), http://imgt.cines.fr. Amino acid sequences and gaps (shown by dots) are according to the IMGT unique numbering for V-DOMAIN (Lefranc et al. 2003b). The three additional positions in the CDR3-IMGT are 111.1, 112.2 and 112.1. Potential N-glycosylation sites are underlined. Assignments of the V, D and J genes are shown in Table 1.

### 2.2.2 MHC G-DOMAINs

The four G-DOMAINs, G-ALPHA1 and G-ALPHA2 of the MHC-I, and G-ALPHA and G-BETA of the MHC-II (Figs. 1, 4, and 5), have a similar groove 3D structure that consists of one sheet of four antiparallel β strands ("floor" of the groove or platform) and one long helical region ("wall" of the groove) (Lefranc et al. 2005b). For each G-DOMAIN (Figs. 4 and 5), the A strand comprises positions 1 to 14, B strand positions 18 to 28, C strand positions 31 to 38, and D strand positions 42 to 49 (Lefranc et al. 2005b). The helix (positions 50 to 92) seats on the β sheet and its axis forms an angle of about 40 degrees with the β strands. The helix is split into two parts separated by a kink, positions 58 of G-ALPHA1, 61 of G-ALPHA2, 63 of G-ALPHA, and 62 of G-BETA being the "highest" points on the floor groove. The G-ALPHA2 and G-BETA domains have a disulfide bridge between positions 11 and 74. The G-ALPHA1 and G-ALPHA domains have a conserved N-glycosylation site at position 86 (N-X-S/T, where N is asparagine, X any amino acid except proline, S is serine, and T is threonine), except for HLA-DMB and H2-DMB1. Asparagine 15 of the G-BETA domains also belongs to a conserved N-glycosylation site (Lefranc et al. 2005b).

## 2.3 TR/pMHC Contact Analysis

### 2.3.1 Peptide/MHC

The 3D structure of the MHC main chain is well conserved and the peptide binding groove specificity is due to side chain physicochemical characteristics (Reinherz et al. 1999). Both MHC-I and MHC-II grooves have pockets where side chains of bound peptides may anchor (Falk, Rotzschke, Stevanovic, Jung, and Rammensee 1991), the specificity of a peptide to a given MHC being controlled by the physicochemical properties of the pockets. Conversely, comparisons of peptide sequence alignments and pMHC 3D structures have revealed that some anchored peptide positions with conserved properties were needed to bind a peculiar MHC allele. Several databases, SYFPEITHI (Rammensee, Bachmann, Emmerich, Bachor, and Stevanovie 1999), JenPep (Blythe, Doytchinova, and Flower 2002), and MHCpep (Brusic, Rudy, and Harrison 1998), provide peptide sequences associated with MHC alleles together with anchor positions and experimental data on affinity. These observations have extensively been used in peptide/MHC binding prediction (Singh and Raghava 2003; Adams and Koziol 1995; Vasmatzis, Zhang, Cornette, and DeLisi 1996b). A list of prediction programs and servers is available at "The IMGT Immunoinformatics page" (http://imgt.cines.fr). Nevertheless, exceptions have been found (Mandelboim, Bar-Haim, Vadai, Fridkin, and Eisenbach 1997; Apostolopoulos, Yu, Corper, Teyton, Pieters, McKenzie, and Wilson 2002; Scott, Peterson, Teyton, and Wilson 1998) and it was noted that while only 30% of peptides with the expected pattern really bind, peptides without the expected pattern also bind (Gulukota, Sidney, Sette, and

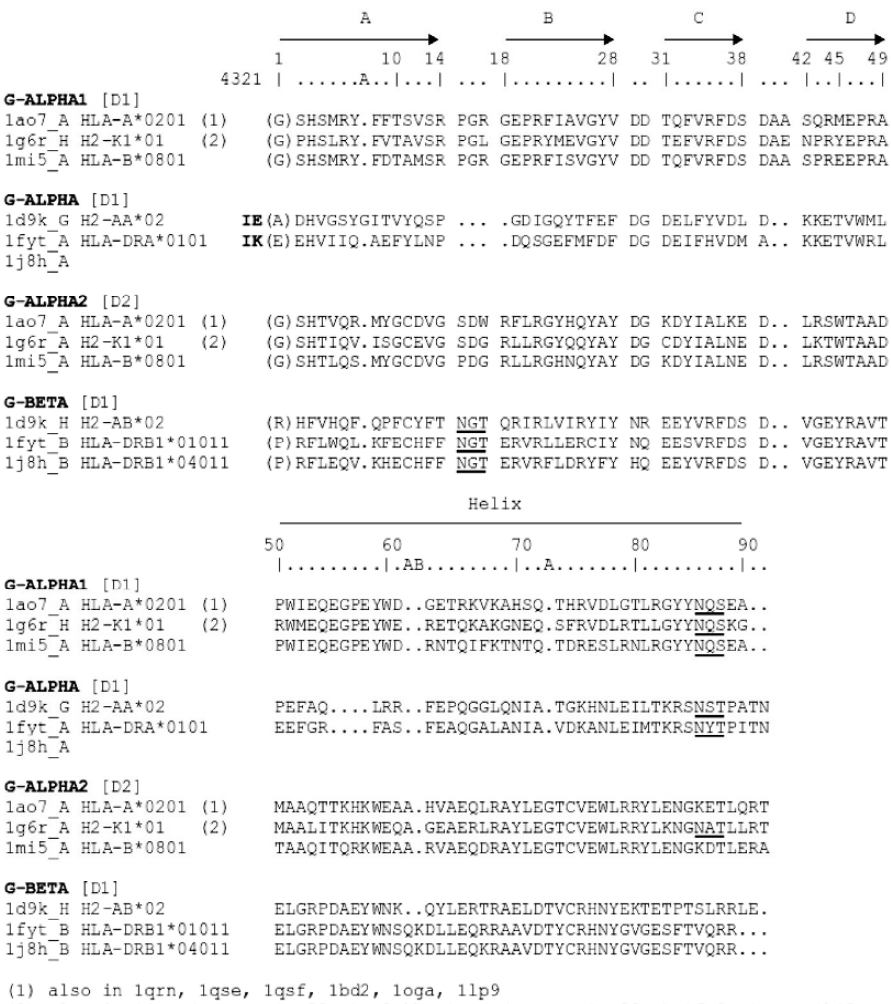


**Fig. 4.** Protein display of the G-DOMAINs found in the TR/pMHC complexes in IMGT/3Dstructure-DB (Kaas et al. 2004), http://imgt.cines.fr. Amino acid sequences and gaps (shown by dots) are according to the IMGT unique numbering for G-DOMAIN (Lefranc et al. 2005b). Amino acids resulting from the splicing with the preceding exon are shown within parentheses. Potential N-glycosylation sites are underlined. Positions 61A, 61B and 72A are characteristic of the G-ALPHA2 and G-BETA domains. The corresponding gaps in G-ALPHA1 and G-ALPHA shown in this IMGT Protein display are not reported in the IMGT Colliers de Perles as these gaps are shared by those two domains. H2-K1*01 encodes H2-K1b, H2-AB*02 and H2-AA*02 encode I-Abk and I-Aak, respectively.

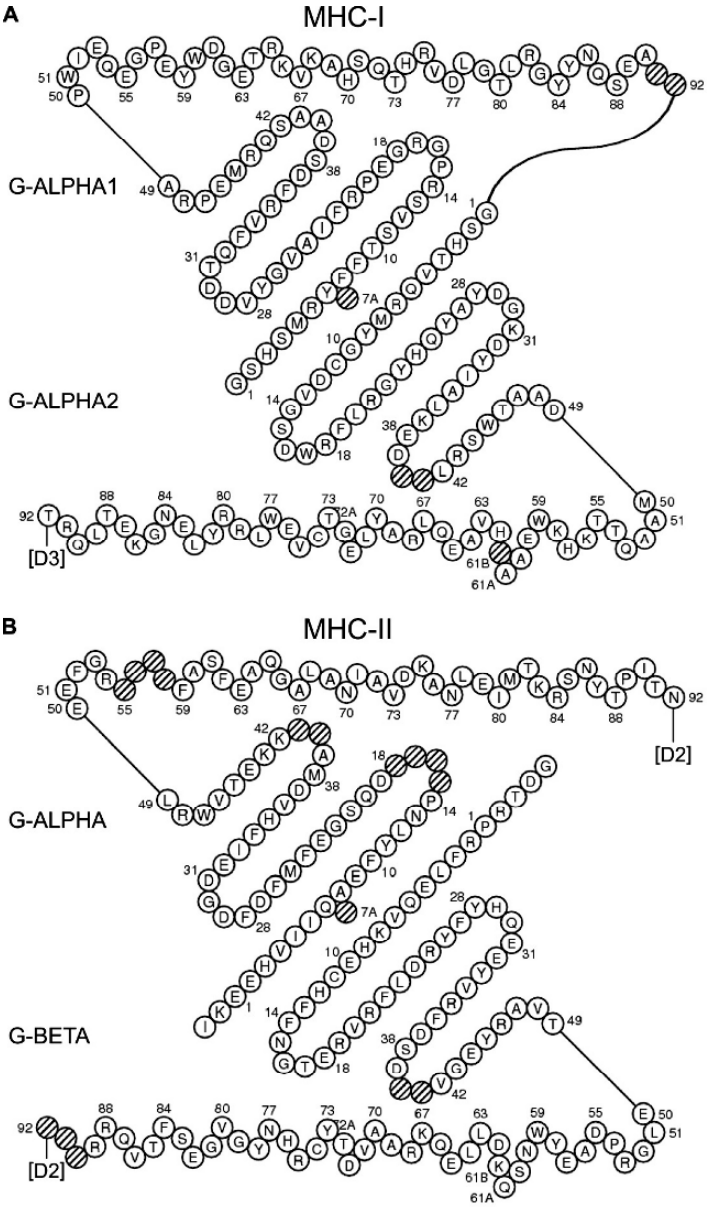


**Fig. 5.** IMGT Colliers de Perles of MHC G-DOMAINs. (A) MHC-I G-ALPHA1 and G-ALPHA2 domains from 1ao7 (B) MHC-II G-ALPHA and G-BETA domains from 1j8h (IMGT/3Dstructure-DB (Kaas et al. 2004), http://imgt.cines). Amino acids positions are according to the IMGT unique numbering for G-DOMAIN (Lefranc et al. 2005b). Positions 61A, 61B and 72A are characteristic of the G-ALPHA2 and G-BETA domains (and are not reported in the G-ALPHA1 and G-ALPHA IMGT Colliers de Perles).

DeLisi 1997). Peptide/MHC binding prediction and epitope prediction remain a big challenge. In order to compare data from different MHC sequences and 3D structures, the IMGT unique numbering for G-DOMAIN has been set up (Lefranc et al. 2005b) (Figs. 4 and 5). This has allowed to graphically represent, in the IMGT Colliers de Perles for G-DOMAIN (Fig. 5), the MHC amino acid positions that have contacts with the peptide side chains. Eleven IMGT pMHC contact sites were defined (C1 to C11, in Figs. 6–8) which can be used to compare pMHC interactions (Kaas and Lefranc 2005). Examples of contact sites for an MHC-I binding an 8-mer peptide (1jtr), for an MHC-I binding a 9-mer peptide (1ao7), and for an MHC-II binding the nine amino acids of a peptide (1j8h) are shown in Figs. 6, 7, and 8, respectively.

In contrast to previous attempts to define pockets (Zhang, Anderson, and DeLisi 1998), structural data for defining the IMGT pMHC contact sites take into account the length of the peptides and are considered independently of the MHC class and sequence polymorphisms. The interactions between the peptide amino acid side chains and MHC amino acids were computed using an interaction scoring scheme based on true mean energy ratio (Kaas and Lefranc 2005). All direct contacts (defined with a cutoff equal to the sum of the atom van der Waals radii and of the diameter of a water molecule) and water-mediated hydrogen bonds were taken into account for the definition of the IMGT pMHC contact sites (Kaas and Lefranc 2005). The analysis was carried out for the pMHC available in IMGT/3Dstructure-DB (Kaas et al. 2004), http://imgt.cines.fr. One hundred fourteen 3D structures with peptides of 8, 9, and 10 amino acids bound to MHC-I and forty-four 3D structures of pMHC-II were identified. The contact analysis was performed for the peptide amino acid side chains of the 9 amino acids located in the groove. Results for MHC-I with 8-amino acid peptides (30 pMHC-I 3D structures), MHC-I with 9-amino acid peptides (74 pMHC-I 3D structures), and MHC-II for the 9 amino acids located in the groove (44 pMHC-II 3D structures) are reported in Table 2 (the results for the 10 pMHC-I with 10-amino acid peptides are not shown). These "IMGT reference pMHC contact sites" are also available as IMGT Colliers de Perles. They will be updated as the number of 3D structures increases. IMGT Colliers de Perles for IMGT pMHC contact sites are provided for each individual pMHC and TR/pMHC entry in IMGT/3Dstructure-DB. They allow easy identification of the amino acid contacts between the MHC and the peptide amino acid side chains and comparison of them with the "IMGT reference pMHC contact sites".

C1 to C11 refer to the 11 IMGT pMHC contact sites (Kaas and Lefranc 2005). 1 to 9 refer to the numbering of the peptide amino acids in the groove. The peptide binding mode to MHC-I is characterized by the N and C peptide ends docked deeply with C1 and C11 contact sites that correspond to the two conserved pockets A and F, and by the peptide length that mechanically constrains the peptide conformation in the groove. There are no C2, C7, and C8 contact sites for MHC-I with 8-amino acid peptides and no C2 and C7 contact sites for MHC-I with 9-amino acid peptides. In contrast, for MHC-II, C2 is present but there are no C7 and C8. Whereas C1 and C11 correspond to the conserved pockets A and F, respectively, the correspondence between the other

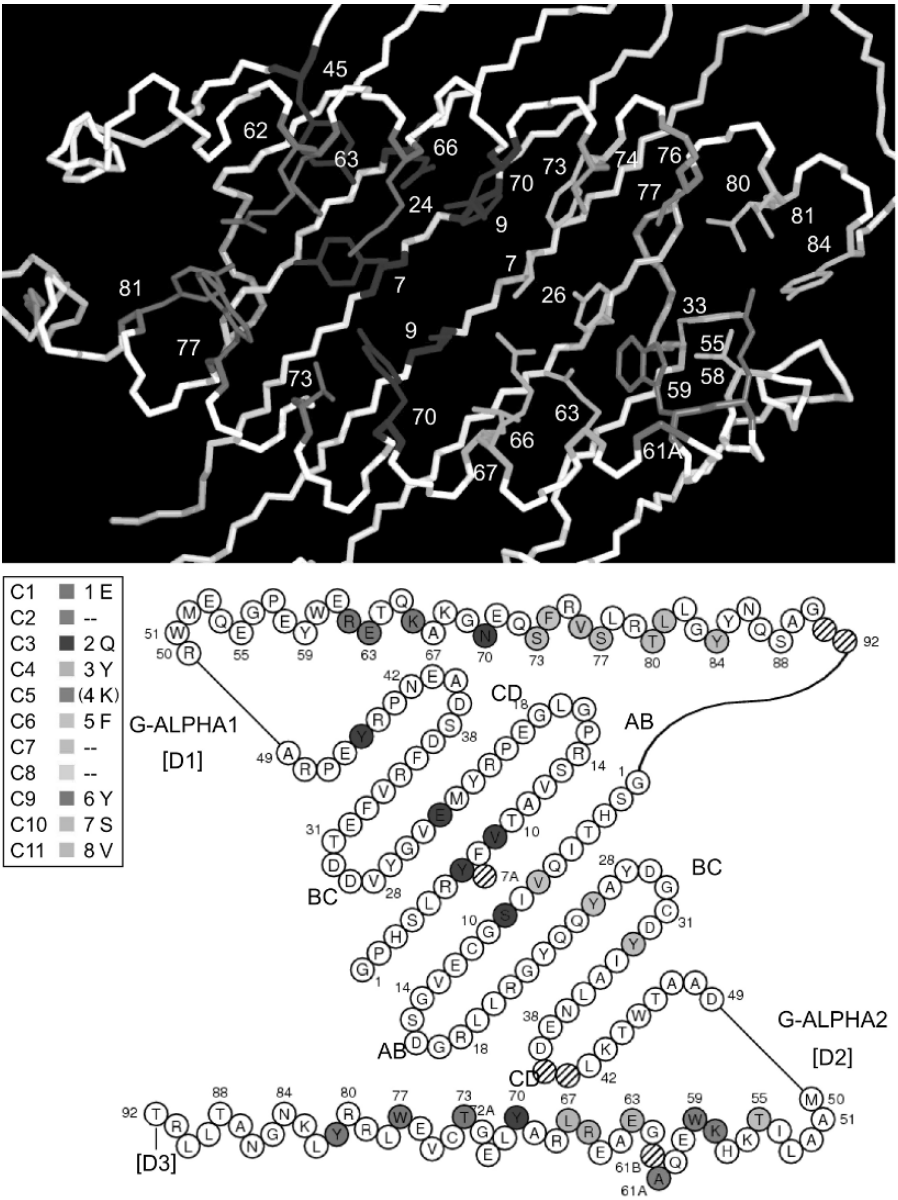


**Fig. 6.** IMGT pMHC contact sites of mouse H2-K1 MHC-I and a 8-amino acid peptide (1jtr). (A) 3D structure of the mouse H2-K1*01 groove. (B) IMGT pMHC contact sites IMGT Colliers de Perles. Both views are from above the cleft with G-ALPHA1 on top and G-ALPHA2 on bottom. In the box, C1 to C11 refer to contact sites (Kaas and Lefranc 2005), 1 to 8 refer to the numbering of the peptide amino acids P1 to P8. There are no C2, C7 and C8 in MHC-I 3D structures with 8-amino acid peptides. There is no C5 in this 3D structure as P4 does not contact MHC amino acids (4K is shown between parentheses in the box). (A color version of this figure appears between pages 76 and 77.)

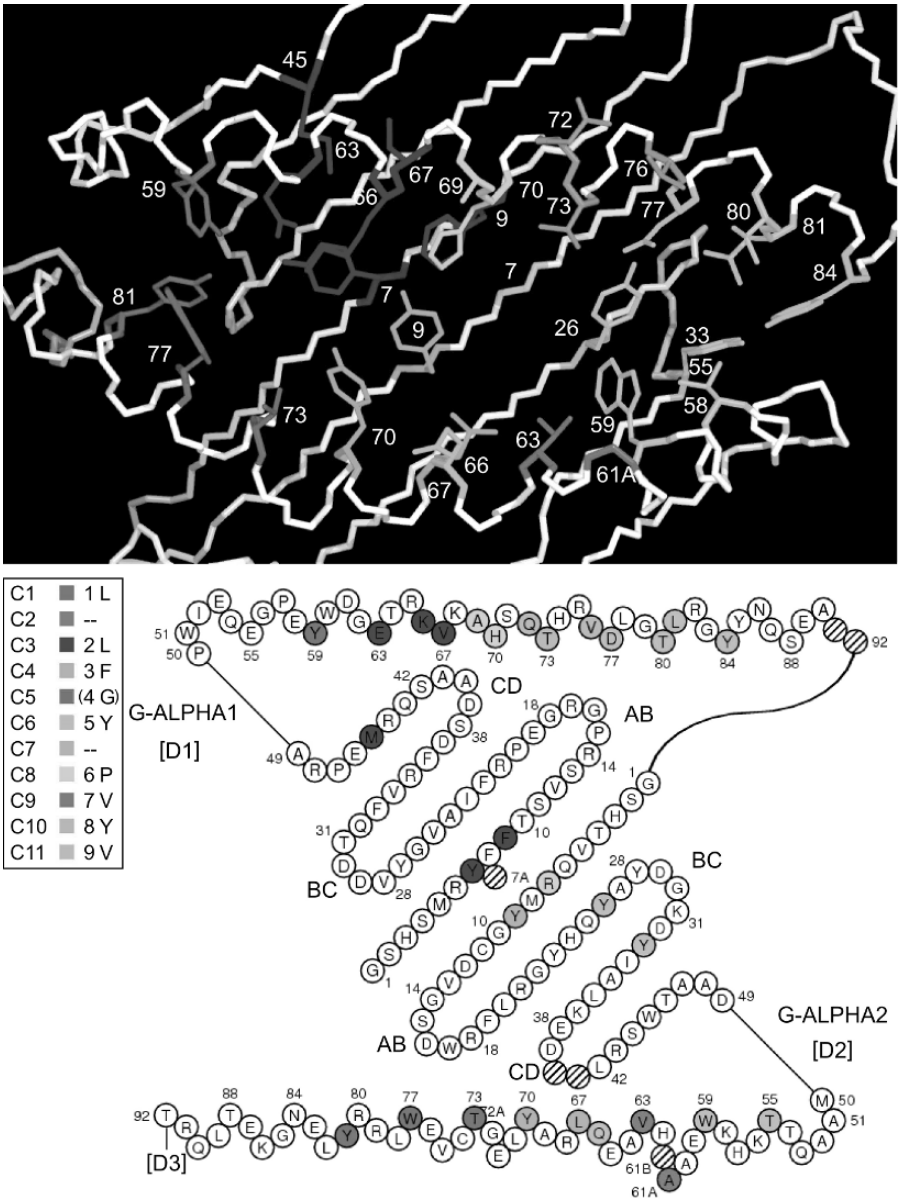


**Fig. 7.** IMGT pMHC contact sites of human HLA-A*0201 MHC-I and a 9-amino acid peptide (1ao7). (A) 3D structure of the human HLA-A*0201 groove. (B) IMGT pMHC contact sites IMGT Colliers de Perles. Both views are from above the cleft with G-ALPHA1 on top and G-ALPHA2 on bottom. In the box, C1 to C11 refer to contact sites (Kaas and Lefranc 2005). 1 to 9 refer to the numbering of the peptide amino acids P1 to P9. There are no C2 and C7 in MHC-I 3D structures with 9-amino acid peptides. There is no C5 in this 3D structure as P4 does not contact MHC amino acids (4G is shown between parentheses in the box). (A color version of this figure appears between pages 76 and 77.)

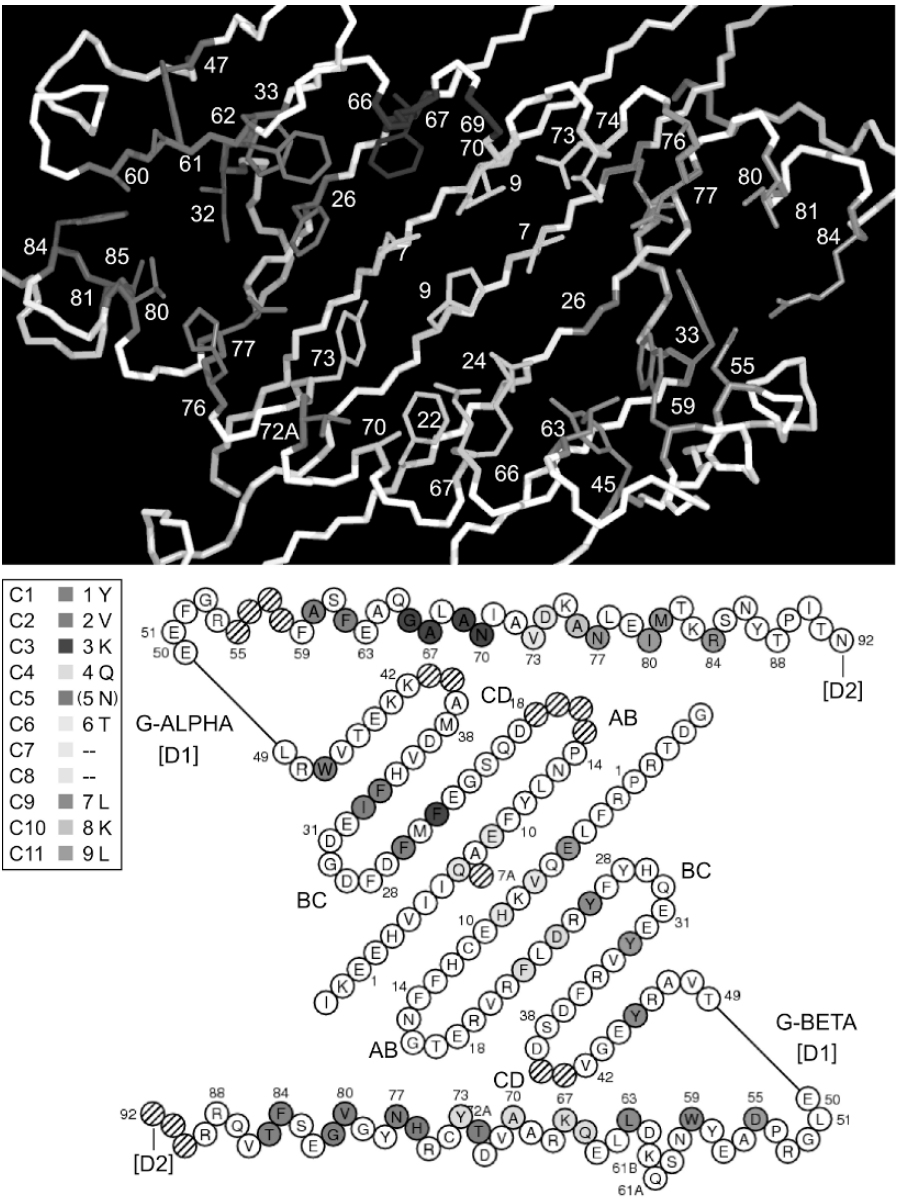


**Fig. 8.** IMGT pMHC contact sites of the human HLA-DRA*0101 and HLA-DRB1*0401 MHC-II and the peptide side chains (9-amino acids located in the groove). (A) 3D structure of the human HLA-DRA*0101 and HLA-DRB1*0401 groove (1j8h). (B) IMGT pMHC contact sites IMGT Colliers de Perles. Both views are from above the cleft with G-ALPHA on top and G-BETA on bottom. In the box, C1 to C11 refer to contact sites. 1 to 9 refer to the numbering of the peptide amino acids 1 to 9 located in the groove. There is no C5 and C7 in MHC-I 3D structures with 9-amino acid peptides. There is no C5 in this 3D structure as 5 does not contact MHC amino acids (5N is shown between parentheses in the box). (A color version of this figure appears between pages 76 and 77.)

contact sites and the previously defined pockets is more approximative. For MHC-I with a peptide of 8-amino acids, C3, C4, C6, and C9 correspond roughly to the B, D, C, and E pockets, and for MHC-I with a peptide of 9-amino acids C3, C4, and C9 correspond to the B, D, and E pockets.

**Table 2.** IMGT reference pMHC contact sites. (A) MHC-I. Results for 104 pMHC-I 3D structures (30 with 8-amino acid peptides and 74 with 9-amino acid peptides). (B) MHC-II. Results from 44 pMHC-II 3D structures with 9 amino acids in the groove.

| (A) MHC-I | | | |
|---|---|---|---|
| 8-amino acid peptides | | | |
| | | G-ALPHA1 | G-ALPHA2 |
| C1 | 1 | 59 62 63 66 | 73 77 81 |
| C3 | 2 | 7 24 45 | 9 |
| C4 | 3 | | 9 24 63 66 67 70 |
| C5 | 4 | | |
| C6 | 5 | 7 9 22 70 74 | 7 9 24 26 |
| C9 | 6 | | 59 61A 63 66 |
| C10 | 7 | 77 73 76 | |
| C11 | 8 | 77 80 81 84 | 5 26 33 34 55 59 |
| 9-amino acid peptides | | | |
| | | G-ALPHA1 | G-ALPHA2 |
| C1 | 1 | 5 59 62 63 66 | 73 77 81 |
| C3 | 2 | 7 9 22 24 34 45 63 66 67 70 | |
| C4 | 3 | | 7 9 24 66 67 70 |
| C5 | 4 | 65 66 | 66 |
| C6 | 5 | 70 73 74 | 7 26 66 67 |
| C8 | 6 | 66 69 70 73 74 | 7 24 62 66 |
| C9 | 7 | | 7 24 59 61A 63 66 |
| C10 | 8 | 72 73 76 80 | 58 |
| C11 | 9 | 77 80 81 84 | 5 26 33 34 55 59 |
| (B) MHC-II | | | |
| | | G-ALPHA | G-BETA |
| C1 | 1 | 26 33 34 47 60 61 62 | 77 80 81 82 84 85 |
| C2 | 2 | | 72A 73 76 |
| C3 | 3 | 7 24 62 63 66 67 69 | |
| C4 | 4 | 7 | 9 11 22 24 66 67 70 73 74 |
| C5 | 5 | | 66 |
| C6 | 6 | 9 69 70 73 74 | 7 26 |
| C9 | 7 | | 24 26 45 59 63 66 |
| C10 | 8 | 73 76 | |
| C11 | 9 | 77 80 81 84 | 5 33 55 |

### 2.3.2 TR/pMHC

The analysis of the pairwise contacts that occur at the TR/MHC and TR/peptide interfaces was carried out using the IMGT unique numbering for V-DOMAINs (Lefranc et al. 2003b) for the TR, and the IMGT unique numbering for G-DOMAINs for the MHC (Lefranc et al. 2005b). Table 3 shows the interactions of the TR V-ALPHA and TR V-BETA with MHC-I and the peptide, in nine TR/pMHC-I 3D structures. Table 4 shows the interactions of the TR V-ALPHA and TR V-BETA with MHC-II and the peptide, in two TR/pMHC-II 3D structures. The results show that positions implicated in the binding are well conserved but not the pairwise interactions. The MHC contact positions belong to the G-DOMAIN helices. The TR positions that are involved in the contacts belong mostly to the CDR-IMGT and to anchor positions (shown by squares in Fig. 2). The FR-IMGT positions involved in the contacts are positions 84 and 84A that are located at the DE turn (designated as "hypervariable 4" or HV4). The contact analysis confirms that the V-ALPHA CDR2-IMGT seats on top of the G-ALPHA2 (MHC-I) or G-BETA (MHC-II) helices and that the V-BETA CDR2-IMGT seats on top of the G-ALPHA1 (MHC-I) or G-ALPHA (MHC-II) helices (Tables 3 and 4). This agrees with data (Sim et al. 1996) which showed that most of the TR/MHC specificity comes from the CDR1 and CDR2 because mutations in these CDRs are able to change specificity between MHC-I and MHC-II. V-ALPHA and V-BETA CDR3-IMGT usually follow the same G-DOMAIN contact preference as the CDR2-IMGT but they can also have contacts with the other G-DOMAINs. For example, in the 1oga 3D structure, position 66 of G-ALPHA2 is contacted by the V-ALPHA CDR3-IMGT but also by the V-BETA CDR3-IMGT.

The diagonal orientation of the TR/pMHC complex puts the TR in a globally conserved position for a peptide "read-out" (Buslepp et al. 2003). V-ALPHA is on top of the peptide N terminus while V-BETA is on top of the peptide C terminus. TR positions implicated in the peptide recognition are in CDR3-IMGT and generally to a lesser extent in V-ALPHA CDR1-IMGT (Tables 3 and 4). Nearly every 3D structure shows different CDR3 conformations and binding mode. In the JM22/peptide/HLA-A complex (1oga) (Stewart-Jones et al. 2003), the V-BETA CDR3-IMGT extensively contacts the peptide and G-ALPHA2 through hydrogen bonds (Table 3), by inserting itself between the peptide and the G-ALPHA2. In contrast, the 2C/peptide/H2-K1complex (1jtr) (Degano et al. 2000) has comparatively fewer contacts between the V-BETA CDR3-IMGT and the peptide and the MHC; however the V-BETA CDR1-IMGT has more contacts and hydrogen bonds with the peptide and G-ALPHA2.

The TR LC13 and 2C were crystallized both alone and in complex with a pMHC. The structural superimposition of both V-DOMAIN scaffold $\alpha$ carbons reveals large movements of the CDR3 and of the CDR1, respectively. The V-ALPHA domains of LC13, in the 1mi5 and 1kgc 3D structures, have 3.5 Å root mean square (RMS) between their CDR3. The V-ALPHA domains of 2C, in the 2ckb and 1tcr 3D structures, have 2.3 Å RMS between their CDR1. The TR A6 was crystallized in complex with the same MHC but with different peptides. In these structures, the V-BETA CDR3 adopt different conformations to adapt to the different peptides (Rudolph, Luz, and Wilson 2002). The CDR3 conformational change does not increase the binding surface but gives a better shape complementarity to the interface (Lawrence and Colman 1993).

**Table 3.** TR V-ALPHA and V-BETA CDR interactions with pMHC-I. TR positions in bold indicate hydrogen bonds. 3D structures are from IMGT/3Dstructure-DB (Kaas et al. 2004), http://imgt.cines.fr. Lengths of the CDR-IMGT are shown within brackets. Amino acids are shown in the one-letter code. Sequences of the peptides are reported in Table 1, sequences of the TR V-ALPHA and V-BETA domains in Fig. 3 and sequences of the MHC-I G-ALPHA1 and G-ALPHA2 in Fig. 4. (A) V-ALPHA CDR-IMGT interactions. (B) V-BETA CDR-IMGT interactions. (C) V-ALPHA and V-BETA FR-IMGT interactions.

(A) V-ALPHA CDR-IMGT interactions

| V-ALPHA CDR1-IMGT | | | | |
|---|---|---|---|---|
| PDB | CDR1 | G-ALPHA1 | Peptide | G-ALPHA2 |
| 1ao7 [6.] | $27_D$ | **$58_E$** | | |
| | $28_R$ | $58_E$ | | $77_W$ $80_R$ |
| | $29_G$ | | $1_L$ | $77_W$ |
| | $37_Q$ | $66_K$ | $1_L$ **$2_L$** $3_F$ $4_G$ $5_Y$ | $70_Y$ $73_T$ |
| | $38_S$ | | $5_Y$ | |
| 1bd2 [6.] | $28_S$ | | $1_L$ | $76_E$ **$77_W$** |
| | $29_M$ | $58_E$ $59_Y$ $62_G$ $63_E$ $66_K$ | $1_L$ | $77_W$ |
| | $37_D$ | $66_K$ | $4_G$ **$5_Y$** | $66_Q$ $73_T$ |
| | $38_Y$ | | $5_Y$ | $66_Q$ |
| 1oga [5.] | $37_S$ | | | $65_E$ $66_Q$ |
| 1mi5 [7.] | $29_S$ | $62_R$ | | |
| | $30_G$ | | | $69_A$ |
| | $36_T$ | | $4_G$ | **$66_Q$** $70_Y$ $73_T$ |
| | $38_Y$ | | $7_Y$ | $61A_A$ $62_R$ $63_V$ $64_A$ $65_E$ $66_Q$ |
| 1lp9 [6.] | $28_T$ | | | $76_E$ |
| | $29_Y$ | | | $69_A$ $72A_G$ $73_T$ $76_E$ $77_W$ |
| | $36_S$ | | | $69_A$ |
| | $38_F$ | | | $65_E$ $66_Q$ $69_A$ |
| 1g6r [6.] | $27_Y$ | $62_R$ | | |
| | $28_S$ | $58_E$ **$62_R$** | | |
| | $29_A$ | $62_R$ | | |
| | $36_T$ | | | $76_E$ |
| | $38_Y$ | | $3_Y$ $4_R$ | $66_R$ |
| 1jtr [6.] | $27_Y$ | $62_R$ | | |
| | $28_S$ | $58_E$ $59_Y$ $62_R$ | | |
| | $29_A$ | $62_R$ | $1_E$ | $77_W$ |
| | $36_T$ | | $1_E$ | |
| | $38_Y$ | | $3_Y$ $4_K$ | $66_R$ |
| 1fo0 [7.] | $28_Q$ | $58_E$ $62_R$ | | |
| | $29_D$ | $62_R$ | | |
| | $30_S$ | | | **$73_T$** |
| | $36_S$ | | | $69_A$ |
| | $38_F$ | | | $66_R$ |
| 1kj2 [6.] | $27_D$ | $58_E$ **$62_R$** | | |
| | $29_T$ | $62_R$ | $1_K$ | $77_W$ |
| | $37_N$ | | | $73_T$ |

(continued)

*Table 3*. (continued) V-ALPHA CDR2-IMGT

| PDB | CDR2 | G-ALPHA1 | Peptide | G-ALPHA2 |
|---|---|---|---|---|
| 1ao7 [.6.] | $57_Y$ | | | $65_E$ $66_Q$ $69_A$ |
| | $58_S$ | | | $69_A$ |
| | $63_N$ | | | $76_E$ |
| 1bd2 [.6.] | $57_S$ | | | $65_E$ $66_Q$ $69_A$ |
| | $58_S$ | | | $69_A$ $70_Y$ $73_T$ |
| | $59_I$ | | | $68_R$ $69_A$ $72_E$ $72A_G$ |
| 1oga [.6.] | $57_V$ | | | $62_H$ $65_E$ |
| 1mi5 [.4.] | $56_G$ | | | $62_R$ |
| | $57_L$ | | | $65_E$ $66_Q$ $69_A$ |
| | $58_T$ | | | $65_E$ |
| | $64_S$ | | | $65_E$ |
| 1lp9 [.6.] | $57_F$ | | | $61A_A$ $62_H$ $65_E$ $66_Q$ |
| | $58_T$ | | | $62_H$ $65_E$ |
| | $64_K$ | | | $65_E$ |
| 1g6r [.7.] | $57_Y$ | | | $66_R$ $69_A$ |
| | $58_S$ | | | $69_A$ $72A_G$ $73_T$ $76_E$ |
| 1jtr [.7.] | $57_Y$ | | | $65_E$ $66_R$ $69_A$ |
| | $58_S$ | | | $69_A$ $76_E$ |
| 1fo0 [.7.] | $59_Y$ | | | $62_G$ $65_E$ $66_R$ $69_A$ |
| | $62_K$ | | | $65_E$ |
| 1kj2 [.6.] | $57_R$ | | | $69_A$ $72_E$ |
| | $58_S$ | | | $76_E$ |
| | $59_V$ | | | $72_E$ $72A_G$ $76_E$ |

V-ALPHA CDR3-IMGT

| PDB | CDR3 | G-ALPHA1 | Peptide | G-ALPHA2 |
|---|---|---|---|---|
| 1ao7 [.11] | $108_T$ | $\mathbf{65}_R$ $66_K$ | $4_G$ $5_Y$ | |
| | $109_D$ | $62_G$ $\mathbf{65}_R$ $66_K$ | $4_G$ $5_Y$ | |
| | $110_S$ | | $\mathbf{4}_G$ $5_Y$ $6_P$ | |
| | $113_W$ | $65_R$ $68_K$ $69_A$ $72_Q$ | | |
| | $114_G$ | $65_R$ | | |
| 1bd2 [.10] | $107_M$ | | $5_Y$ | |
| | $108_E$ | $58_E$ $62_G$ $\mathbf{65}_R$ $66_K$ | | |
| | $109_G$ | $65_R$ $66_K$ | $4_G$ $5_Y$ | |
| | $113_A$ | | $4_G$ $5_Y$ | |
| | $114_Q$ | $65_R$ $69_A$ | | |
| | $115_K$ | $65_R$ | | |
| 1oga [.10] | $107_A$ | | | $66_Q$ |
| | $108_G$ | | $5_F$ | $66_Q$ |
| | $109_S$ | | $4_G$ $5_F$ | $66_Q$ |
| | $113_Q$ | $66_K$ | $4_G$ $5_F$ | |
| | $114_G$ | | $4_G$ $5_F$ | |
| 1mi5 [.14] | $108_L$ | | $6_A$ $7_Y$ | $66_Q$ |
| | $109_A$ | $62_R$ | | |
| | $110_G$ | $62_R$ $66_I$ | | |
| | $111_G$ | $65_Q$ $66_I$ $\mathbf{69}_T$ | $4_G$ | |

| PDB | CDR3 | G-ALPHA1 | Peptide | G-ALPHA2 |
|---|---|---|---|---|
| | $112.1_T$ | $62_R$ $65_Q$ $66_I$ $69_T$ | | |
| | $112_S$ | $69_T$ | $6_A$ | |
| | $113_Y$ | $69_T$ $72_Q$ | $6_A$ | |
| 1lp9 [.13] | $107_F$ | | $5_F$ | $66_Q$ |
| | $109_A$ | | $3_W$ $4_G$ $5_F$ | $66_Q$ |
| | $110_S$ | | $2_L$ $3_W$ $4_G$ | $66_Q$ $69_A$ $70_Y$ $73_T$ |
| | $111_S$ | $63_E$ $\mathbf{66}_K$ | $2_L$ $4_G$ | $73_T$ $77_W$ |
| | $112_S$ | $66_K$ | $4_G$ $5_F$ | |
| | $113_F$ | $65_R$ $66_K$ $69_A$ | $4_G$ $6_F$ | |
| | $114_S$ | | $\mathbf{4}_G$ $5_F$ $6_F$ | |
| 1g6r [.10] | $107_S$ | | $\mathbf{4}_R$ | |
| | $108_G$ | | $\mathbf{4}_R$ | |
| | $109_F$ | $62_R$ $65_Q$ $66_K$ | $4_R$ | |
| | $113_A$ | | $4_R$ | |
| | $114_S$ | | $4_R$ | |
| 1jtr [.10] | $107_S$ | | $4_K$ | |
| | $108_G$ | | $4_K$ | |
| | $109_F$ | $62_R$ $65_Q$ $66_K$ $69_G$ | $4_K$ | |
| | $113_A$ | | $\mathbf{4}_K$ | |
| | $114_S$ | | $4_K$ | |
| 1fo0 [.14] | $110_Y$ | $65_Q$ | | |
| | $111_G$ | $65_Q$ | | |
| | $112.1_G$ | $65_Q$ | | |
| 1kj2 [.11] | $108_Y$ | $62_R$ | | |
| | $109_Q$ | $63_E$ $\mathbf{66}_K$ | $1_K$ $\mathbf{2}_V$ $3_I$ $4_T$ | $70_Y$ $73_T$ |
| | $110_G$ | $66_K$ | $4_T$ | |
| | $114_R$ | $\mathbf{65}_Q$ $68_K$ $69_G$ $72_Q$ | | |

(B) V-BETA CDR-IMGT interactions

| V-BETA CDR1-IMGT | | | | |
|---|---|---|---|---|
| PDB | CDR1 | G-ALPHA1 | Peptide | G-ALPHA2 |
| 1ao7 [5.] | $37_E$ | | $\mathbf{8}_Y$ | |
| 1oga [5.] | $37_D$ | | $\mathbf{8}_T$ | $58_K$ |
| 1mi5 [5.] | $37_V$ | $76_E$ $80_N$ | | |
| | $38_S$ | $76_E$ | | |
| 1lp9 [5.] | $37_D$ | $\mathbf{72}_Q$ $76_V$ | | |
| | $38_Y$ | $69_A$ $73_T$ | $6_F$ | |
| 1g6r [5.] | $28_N$ | | $6_Y$ | $58_K$ |
| | $29_H$ | | $6_Y$ | $61_Q$ $61A_A$ |
| | $37_N$ | | $\mathbf{6}_Y$ $7_G$ $8_L$ | $58_K$ |
| | $38_N$ | | $6_Y$ | |
| 1jtr [5.] | $27_N$ | | | $61_Q$ |
| | $28_N$ | | $6_Y$ | $\mathbf{58}_K$ $61_Q$ |
| | $29_H$ | | $6_Y$ | $58_K$ $61A_A$ |
| | $37_N$ | | $\mathbf{6}_Y$ $\mathbf{7}_S$ $8_V$ | $58_K$ |
| | $38_N$ | | $6_Y$ | |

(continued)

*Table 3*. (continued) V-BETA CDR1-IMGT

| PDB | CDR1 | G-ALPHA1 | Peptide | G-ALPHA2 |
|---|---|---|---|---|
| 1fo0 [6.] | $29_{2}$ | $61_{2}$ | | |
| | $38_{W}$ | $76_{V}$ | $7_{T}$ | |
| 1kj2 [6.] | $29_{Q}$ | | | $58_{K}$ $59_{W}$ $61_{Q}$ $61A_{A}$ |
| | $36_{Y}$ | | | $61A_{A}$ |
| | $37_{P}$ | | $7_{D}$ | |
| | $38_{W}$ | $69_{G}$ $72_{Q}$ | | |

V-BETA CDR2-IMGT

| PDB | CDR2 | G-ALPHA1 | Peptide | G-ALPHA2 |
|---|---|---|---|---|
| 1bd2 [.6.] | $65_{I}$ | $72_{Q}$ | | |
| 1oga [.6.] | $57_{Q}$ | $69_{A}$ | $\mathbf{4}_{G}$ $5_{F}$ $\mathbf{6}_{V}$ | |
| | $58_{I}$ | $69_{A}$ $72_{Q}$ $73_{T}$ $76_{V}$ | $6_{V}$ $8_{T}$ | |
| | $63_{V}$ | $72_{Q}$ $76_{V}$ | | |
| | $64_{N}$ | $72_{Q}$ $75_{R}$ | | |
| | $65_{D}$ | $65_{R}$ $68_{K}$ $69_{A}$ $72_{Q}$ | | |
| 1mi5 [.6.] | $57_{Q}$ | $72_{Q}$ $75_{R}$ $76_{E}$ | | |
| | $58_{N}$ | $\mathbf{79}_{R}$ | | |
| | $63_{E}$ | $\mathbf{79}_{R}$ | | |
| 1lp9 [.6.] | $57_{Y}$ | $65_{R}$ $68_{K}$ $69_{A}$ $72_{Q}$ | | |
| | $58_{V}$ | $72_{Q}$ | | |
| | $65_{S}$ | $68_{K}$ | | |
| 1g6r [.6.] | $57_{Y}$ | $69_{G}$ $70_{N}$ $72_{Q}$ $73_{S}$ $76_{V}$ | | |
| | $58_{G}$ | $76_{V}$ | | |
| | $63_{A}$ | $76_{V}$ $79_{R}$ | | |
| | $64_{G}$ | $79_{R}$ | | |
| | $65_{S}$ | $76_{V}$ | | |
| 1jtr [.6.] | $57_{Y}$ | $69_{G}$ $72_{Q}$ $73_{S}$ $76_{V}$ | $\mathbf{7}_{S}$ | |
| | $58_{G}$ | $76_{V}$ | | |
| | $63_{A}$ | $76_{V}$ $\mathbf{79}_{R}$ $80_{T}$ | | |
| | $64_{G}$ | $79_{R}$ | | |
| | $65_{S}$ | $72_{Q}$ $76_{V}$ $79_{R}$ | | |
| 1fo0 [.6.] | $57_{R}$ | $76_{V}$ $79_{R}$ $80_{T}$ | | |
| | $58_{S}$ | $76_{V}$ $\mathbf{79}_{R}$ | | |
| | $63_{P}$ | $79_{R}$ | | |
| 1kj2 [.6.] | $57_{R}$ | $72_{Q}$ $73_{S}$ $76_{V}$ | $\mathbf{7}_{\mathbf{D}}$ | |
| | $58_{S}$ | $72_{Q}$ | | |
| | $65_{D}$ | $72_{Q}$ | | |

| PDB | CDR3 | G-ALPHA1 | Peptide | G-ALPHA2 |
|---|---|---|---|---|
| 1ao7 [.14] | $107_R$ | | $5_Y$ | |
| | $109_G$ | | $6_P$ | |
| | $110_L$ | $69_A$ $72_Q$ $73_T$ | $6_P$ $7_V$ **$8_Y$** | |
| | $111_A$ | | $7_V$ $8_Y$ | $61A_A$ |
| | $112_G$ | | $5_Y$ $7_V$ | **$61A_A$** $62_H$ $63_V$ **$66_Q$** |
| | $112.1_G$ | | $5_Y$ $7_V$ | $61A_A$ |
| | $113_R$ | | $5_Y$ | **$61_A$** $61A_A$ $62_H$ $66_Q$ |
| | $114_P$ | | $5_Y$ | $66_Q$ |
| 1bd2 [.13] | $108_Y$ | | $8_Y$ | |
| | $109_P$ | | $6_P$ $7_V$ | |
| | $110_G$ | | $6_P$ $7_V$ **$8_Y$** | |
| | $111_G$ | | $7_V$ $8_Y$ | $61A_A$ |
| | $112_G$ | | $7_V$ | $61A_A$ |
| | $114_Y$ | | $5_Y$ $7_V$ | $61A_A$ $63_V$ $66_Q$ |
| 1oga [.11] | $108_S$ | | | $61A_A$ |
| | $109_R$ | | $5_F$ $6_V$ $7_F$ | **$61A_A$** **$62_H$** $63_V$ **$66_Q$** |
| | $110_S$ | | $5_F$ $6_V$ | $66_Q$ |
| | $113_S$ | | $5_F$ | $66_Q$ |
| | $114_Y$ | | | **$61_A$** $61A_A$ $62_H$ |
| 1mi5 [.11] | $108_L$ | $76_E$ | | $58_K$ |
| | $109_G$ | $76_E$ | | |
| | $110_Q$ | $69_T$ **$72_Q$** **$73_T$** **$76_E$** | $5_R$ **$6_A$** | |
| | $113_A$ | | $6_A$ $7_Y$ | |
| | $114_Y$ | $76_E$ | $7_Y$ **$8_G$** | $58_K$ $59_W$ $61A_A$ |
| 1lp9 [.11] | $109_W$ | | $5_F$ $6_F$ $7_P$ $8_V$ | $58_K$ $59_W$ $61A_A$ $63_V$ |
| | $110_V$ | | $5_F$ | $61A_A$ |
| | $113_S$ | | $5_F$ | |
| | $114_Y$ | | $5_F$ | $61_A$ $61A_A$ $62_H$ $66_Q$ |
| 1g6r [.9] | $107_G$ | | $6_Y$ | |
| | $108_G$ | | $6_Y$ | $61A_A$ $63_E$ |
| | $109_G$ | | $4_R$ $6_Y$ | $61A_A$ $66_R$ |
| | $114_G$ | | $4_R$ | $66_R$ |
| | $115_T$ | | | $61A_A$ |
| 1jtr [.9] | $107_G$ | | $6_Y$ | |
| | $108_G$ | | $6_Y$ | **$63_E$** $66_R$ |
| | $109_G$ | | $4_K$ $6_Y$ | $66_R$ |
| | $115_T$ | | | $61_A$ |
| 1fo0 [.12] | $108_A$ | | | $58_K$ |
| | $109_D$ | | $6_N$ **$7_T$** | $58_K$ $59_W$ |
| | $110_R$ | **$69_G$** $70_N$ $72_Q$ **$73_S$** | $4_D$ **$5_F$** $6_N$ | |
| | $112_V$ | | $4_D$ $5_F$ **$6_N$** | $66_R$ |
| | $113_G$ | | $6_N$ | |
| | $114_N$ | | $6_N$ | $61A_A$ |
| 1kj2 [.16] | $108_A$ | | $6_I$ | $66_R$ |
| | $109_A$ | | $4_T$ $6_I$ | $66_R$ |
| | $110_P$ | | $4_T$ | |

(continued)

*Table 3*. (continued) V-BETA CDR3-IMGT

| PDB | CDR3 | G-ALPHA1 | Peptide | G-ALPHA2 |
|---|---|---|---|---|
| | $111_{D}$ | | $4_{T}$ | $66_{R}$ |
| | $111.1_{W}$ | | | $62_{G}$ $65_{E}$ $66_{R}$ $69_{A}$ |
| | $112.1_{A}$ | | | $61A_{A}$ |
| | $112_{S}$ | | | $61_{Q}$ $61A_{A}$ |
| | $114_{E}$ | | | $69_{A}$ |

(C) V-ALPHA and V-BETA FR-IMGT interactions

V-ALPHA FR-IMGT

| PDB | Position | G-ALPHA1 | Peptide | G-ALPHA2 |
|---|---|---|---|---|
| 1ao7 | $2_{K}$ | $58_{E}$ | | |
| | $26_{S}$ | $58_{E}$ | | |
| | $82_{K}$ | | | **$73_{T}$** **$76_{E}$** |
| 1bd2 | $2_{Q}$ | $58_{E}$ $65_{R}$ | | |
| | $82_{K}$ | | | $72A_{G}$ $73_{T}$ |
| 1oga | $84_{R}$ | | | $65_{E}$ |
| 1mi5 | $40_{H}$ | | $7_{Y}$ | |
| | $52_{Y}$ | | | $62_{R}$ |
| | $55_{H}$ | | $7_{Y}$ | **$61A_{A}$** **$62_{R}$** |
| | $66_{V}$ | | | $62_{R}$ |
| 1lp9 | $82_{K}$ | | | $65_{E}$ |
| 1g6r | $2_{Q}$ | | $4_{R}$ | |
| | $55_{K}$ | | | $65_{E}$ |
| 1kj2 | $82_{K}$ | | | $76_{E}$ |

V-BETA FR-IMGT

| PDB | Position | G-ALPHA1 | Peptide | G-ALPHA2 |
|---|---|---|---|---|
| 1bd2 | $55_{Y}$ | $65_{R}$ | | |
| | $67_{D}$ | $68_{K}$ | | |
| 1oga | $67_{Q}$ | $65_{R}$ | | |
| 1mi5 | $55_{Y}$ | $72_{Q}$ $76_{E}$ | | |
| | $66_{L}$ | $72_{Q}$ $75_{R}$ | | |
| 1lp9 | $55_{Y}$ | $65_{R}$ | | |
| | $67_{E}$ | **$65_{R}$** $68_{K}$ | | |
| 1g6r | $67_{E}$ | $72_{Q}$ | | |
| | $84_{Q}$ | | | $58_{K}$ |
| 1jtr | $67_{E}$ | $72_{Q}$ | | |
| | $84_{Q}$ | | | **$58_{K}$** |

**Table 4.** V-ALPHA and V-BETA CDR interactions with MHC-II. TR positions in bold indicate hydrogen bonds. Three dimensional (3D) structures are from IMGT/3Dstructure-DB (Kaas et al. 2004), http://imgt.cines.fr. Lengths of the CDR-IMGT are shown within brackets. Amino acids are shown in the one-letter code. Sequences of the peptides are reported in Table 1, sequences of the TR V-ALPHA and V-BETA domains in Fig. 3, and sequences of the MHC-II G-ALPHA and G-BETA in Fig. 4. (A) V-ALPHA CDR-IMGT interactions. (B) V-BETA CDR-IMGT interactions. (C) V-ALPHA and V-BETA FR-IMGT interactions.

(A) V-ALPHA CDR-IMGT interactions

V-ALPHA CDR1-IMGT

| PDB | Position | G-ALPHA | Peptide | G-BETA |
|---|---|---|---|---|
| 1j8h [6.] | $28_S$ | | $2_K$ | $76_H$ |
| | $29_V$ | | $2_K$ $4_V$ | $76_H$ |
| | $36_P$ | | $4_V$ | $72A_T$ $76_H$ |
| | $38_Y$ | | | $72A_T$ |
| 1d9k [6.] | $27_D$ | | $3_S$ | |
| | $28_S$ | | | $72A_T$ $76_H$ |
| | $29_T$ | | $3_S$ $4_H$ $\mathbf{5}_R$ | $72A_T$ $76_H$ |
| | $36_F$ | | $5_R$ | $72A_T$ |
| | $37_D$ | | $\mathbf{5}_R$ $8_I$ | $\mathbf{66}_R$ $69_A$ $72A_T$ |
| | $38_Y$ | | | $66_R$ |

V-ALPHA CDR2-IMGT

| PDB | Position | G-ALPHA | Peptide | G-BETA |
|---|---|---|---|---|
| 1j8h [.7.] | $57_T$ | | | $\mathbf{65}_E$ |
| | $58_S$ | | | $69_A$ $72A_T$ |
| | $59_A$ | | | $65_E$ |
| 1d9k [.6.] | $57_S$ | | | $65_E$ $66_R$ $69_A$ |
| | $58_L$ | | | $69_A$ $72_D$ $72A_T$ |
| | $59_V$ | | | $65_E$ $66_R$ $68_R$ $69_A$ |
| | $63_S$ | | | $65_E$ |

V-ALPHA CDR3-IMGT

| PDB | Position | G-ALPHA | Peptide | G-BETA |
|---|---|---|---|---|
| 1j8h [.13] | $108_E$ | $63_E$ | $\mathbf{2}_K$ $4_V$ | |
| | $110_P$ | | $7_N$ | $66_Q$ |
| | $111_F$ | | $7_N$ $9_L$ | $62_D$ $63_L$ $66_Q$ |
| | $114_E$ | $66_G$ $69_A$ $70_N$ | $5_K$ | |
| 1d9k [.10] | $107_T$ | | | $66_R$ |
| | $108_G$ | | $5_R$ $8_I$ | $66_R$ |
| | $109_S$ | $69_Q$ | $8_I$ | $66_R$ |
| | $113_F$ | $\mathbf{69}_Q$ $73_T$ | $8_I$ $9_E$ $10_W$ $11_E$ | $63_Y$ $66_R$ |
| | $114_N$ | $69_Q$ | | $66_R$ |
| | $115_K$ | $65_Q$ | | |

(continued)

*Table 4*. (continued)

(B) V-BETA CDR-IMGT interactions

V-BETA CDR1-IMGT

| PDB | Position | G-ALPHA | Peptide | G-BETA |
|---|---|---|---|---|
| 1j8h [5.] | $27_M$ | | $10_K$ | |
| | $28_D$ | $76_A$ | **$10_K$** | |
| | $29_H$ | | $10_K$ | |
| | $37_E$ | $72_A$ $73_V$ $76_A$ | **$10_K$** | |
| | $38_N$ | $69_A$ | | |
| 1d9k [5.] | $37_N$ | $76_H$ | | |
| | $38_N$ | **$69_Q$** | | |

V-BETA CDR2-IMGT

| PDB | Position | G-ALPHA | Peptide | G-BETA |
|---|---|---|---|---|
| 1j8h [.6.] | $57_Y$ | $65_Q$ $66_G$ $68_L$ $69_A$ $72_A$ | | |
| | $58_D$ | $68_L$ $72_A$ **$75_K$** | | |
| | $65_M$ | $43_K$ $68_L$ | | |
| 1d9k [.6.] | $57_Y$ | **$65_Q$** $66_G$ $68_L$ $69_Q$ $72_A$ | | |

V-BETA CDR3-IMGT

| PDB | Position | G-ALPHA | Peptide | G-BETA |
|---|---|---|---|---|
| 1j8h [.12] | $108_S$ | $73_V$ | $10_K$ | |
| | $109_T$ | $69_A$ $70_N$ $73_V$ | $5_K$ $7_N$ $8_T$ | |
| | $110_G$ | $73_V$ | $8_T$ $9_L$ **$10_K$** | |
| | $112_L$ | | $10_K$ | $58_Y$ |
| | $113_P$ | | | $61A_Q$ $62_D$ $63_L$ |
| 1d9k [.11] | $108_G$ | | $11_E$ | |
| | $109_Q$ | | $11_E$ | $58_Y$ $63_Y$ |
| | $110_G$ | | $10_W$ $11_E$ | **$63_Y$** $66_R$ |
| | $113_R$ | | | $61_K$ **$62_Q$** $63_Y$ $65_E$ $66_R$ |
| | $114_A$ | | | $66_R$ |

(C) V-ALPHA and V-BETA FR-IMGT interactions

V-ALPHA FR-IMGT

| PDB | Position | G-ALPHA | Peptide | G-BETA |
|---|---|---|---|---|
| 1j8h | $55_K$ | | | $62_D$ |
| 1d9k | $82_K$ | | | **$72_D$** |

V-BETA FR-IMGT

| PDB | Position | G-ALPHA | Peptide | G-BETA |
|---|---|---|---|---|
| 1j8h | $55_F$ | $65_Q$ | | |
| | $66_K$ | $43_K$ | | |
| | $67_E$ | **$43_K$** **$65_Q$** | | |
| | $84_K$ | $72_A$ $76_A$ | $10_K$ | |
| 1d9k | $55_Y$ | **$65_Q$** | | |
| | $66_T$ | **$43_K$** | | |
| | $67_E$ | **$43_K$** $65_Q$ $68_L$ | | |
| | $68_K$ | $65_Q$ | | |

## 2.4 Conclusions

With only 18 TR/pMHC 3D structures, the atomic details of TR/pMHC interactions already show a great deal of variability. IMGT standardization is a step toward a better understanding of the mechanisms ruling TR/pMHC recognition. It will help comparing new experimentally resolved 3D structures with published data. However, the TR/pMHC interactions are far from being unravelled and the study of the TR/pMHC interactions with the other proteins of the immunological synapse will be crucial. For example, the interaction between an MHC and the CD4 considerably enhances the pMHC/TR sensibility (Irvine, Purbhoo, Krosgaard, and Davis 2002; Davis 2002). The understanding of the T cell triggering early events is subject to active studies.

Although the TR/pMHC binding represents a necessary step for the TR recognition, many factors, the TR affinity for the pMHC, the relocation of surface proteins such as CD4 or CD8 in the immunological synapse are necessary for generating the T cell activation signal. Each of these steps needs to be described and characterized so that data from different experiments can be integrated. IMGT standardization will be further extended on the IMGT Web site at http://imgt.cines.fr as new parameters become available.

## 2.5 Citing IMGT/3Dstructure-DB

Users are requested to cite IMGT/3Dstructure-DB (Kaas et al. 2004) and this article, and to quote the IMGT home page URL, http://imgt.cines.fr.

## Acknowledgements

We are grateful to Vijay Garapati for his contribution to Tables 3 and 4 and to the IMGT® team for helpful discussion. E.D. was the holder of a doctoral grant from the Ministère de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche (MENESR). K.Q. was the recipient of a doctoral grant from the MENESR and was supported for one year by a grant from the Association pour la Recherche sur le Cancer (ARC). IMGT® is a registered Centre National de la Recherche Scientifique (CNRS) mark. IMGT® has been a National RIO Bioinformatics Platform since 2001 (CNRS, INSERM, CEA, INRA). IMGT® was funded in part by the BIOMED1 (BIOCT930038), Biotechnology BIOTECH2 (BIO4CT960037), and 5th PCRDT Quality of Life and Management of Living Resources (QLG2-2000-01287) programs of the European Union and received subventions from ARC and from the Génopole-Montpellier-Languedoc-Roussillon. IMGT® is currently supported by the CNRS, the MENESR (Université Montpellier II Plan Pluri-Formation), BIOSTIC-LR2004, Région Languedoc-Roussillon, ACI-IMPBIO IMP82-2004, the Réseau National des Génopoles RNG, GIS-AGENAE, Agence Nationale de la Recherche ANR (BIOSYS06_135457), and the European ImmunoGrid project (IST-2004-0280069).

Part of this work was carried out in the frame of the European Science Foundation Scientific Network Myelin Structure and its role in autoimmunity (MARIE).

## References

Adams, H.P., and Koziol, J.A. (1995) Prediction of binding to MHC class I molecules. J. Immunol. Methods 185:181-190.

Apostolopoulos, V., Yu, M., Corper, A.L., Teyton, L., Pieters, G.A., McKenzie, I.F.C., and Wilson, I.A. (2002) Crystal structure of a non-canonical low-affinity peptide complexed with MHC class I: A new approach for vaccine design. J. Mol. Biol. 318:1293-1305.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000) The Protein Data Bank. Nucleic Acids Res. 28:235-242.

Blake, J.A., Richardson, J.E., Bult, C.J., Kadin, J.A., and Eppig, J.T. (2003) MGD: The Mouse Genome Database. Nucleic Acids Res. 31:193-195.

Blythe, M.J., Doytchinova, I.A., and Flower, D.R. (2002) JenPep: A database of quantitative functional peptide data for immunology. Bioinformatics 18:434-439.

Brusic, V., Rudy, G., and Harrison, L.C. (1998) MHCPEP, a database of MHC-binding peptides: Update 1997. Nucleic Acids Res. 26:368-371.

Buslepp, J., Wang, H., Biddison, W.E., Appella, E., and Collins, E.J. (2003) A correlation between TCR Vα docking on MHC and CD8 dependence: Implications for T cell selection. Immunity 19:595-606.

Davis, M.M. (2002) A new trigger for T cells. Cell 110:285-287.

Degano, M., Garcia, K.C., Apostolopoulos, V., Rudolph, M.G., Teyton, L., and Wilson, I.A. (2000) A functional hot spot for antigen recognition in a superagonist TCR/MHC complex. Immunity 12:251-261.

Ding, Y.H., Smith, K.J., Garboczi, D.N., Utz, U., Biddison, W.E., and Wiley, D.C. (1998) Two human T cell receptors bind in a similar diagonal mode to the HLA-A2/Tax peptide complex using different TCR amino acids. Immunity 8:403-411.

Ding, Y.H., Baker, B.M., Garboczi, D.N., Biddison, W.E., and Wiley, D.C. (1999) Four A6-TCR/peptide/HLA-A2 structures that generate very different T cell signals are nearly identical. Immunity 11:45-56.

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

Garboczi, D.N., Ghosh, P., Utz, U., Fan, Q.R., Biddison, W.E., and Wiley, D.C. (1996) Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. Nature 384:134-141.

Garcia, K.C., Degano, M., Pease, L.R., Huang, M., Peterson, P.A., Teyton, L., and Wilson, I.A. (1998) Structural basis of plasticity in T cell receptor recognition of a self peptide-MHC. Antigen Sci. 279:1166-1172.

Giudicelli, V., and Lefranc, M.-P. (1999) Ontology for immunogenetics: The IMGT-ONTOLOGY. Bioinformatics 15:1047-1054.

Giudicelli, V., Chaume, D., and Lefranc, M.-P. (2004) IMGT/V-QUEST, an integrated software program for immunoglobulin and T cell receptor V-J and V-D-J rearrangement analysis. Nucleic Acids Res. 32:435-440.

Giudicelli, V., Chaume, D., and Lefranc, M.-P. (2005) IMGT/GENE-DB: A comprehensive database for human and mouse immunoglobulin and T cell receptor genes. Nucleic Acids Res. 33:256-261.

Giudicelli, V., Ginestoux, C., Folch, G., Jabado-Michaloud, J., Chaume, D., and Lefranc, M.-P. (2006) IMGT/LIGM-DB, the IMGT® comprehensive database of immunoglobulin and T cell receptor nucleotide sequences. Nucleic Acids Res. 34:D781-D784.

Gulukota, K., Sidney, J., Sette, A., and DeLisi, C. (1997) Two complementary methods for predicting peptides binding major histocompatibility complex molecules. J. Mol. Biol. 267:1258-1267.

Hennecke, J., Carfi, A., and Wiley, D.C. (2000) Structure of a covalently stabilized complex of a human αβ T-cell receptor, influenza HA peptide and MHC class II molecule, HLA-DR1. EMBO J. 19:5611-5624.

Hennecke, J., and Wiley, D.C. (2002) Structure of a complex of the human alpha/beta T cell receptor (TCR) HA1.7, influenza hemagglutinin peptide, and major histocompatibility complex class II molecule, HLA-DR4 (DRA*0101 and DRB1*0401): Insight into TCR cross-restriction and alloreactivity. J. Exp. Med. 195:571-581.

Irvine, D.J., Purbhoo, M.A., Krosgaard, M., and Davis, M.M. (2002) Direct observation of ligand recognition by T cells. Nature 419:845-849.

Kaas, Q., Ruiz, M., and Lefranc, M.-P. (2004) IMGT/3Dstructure-DB and IMGT/StructuralQuery, a database and a tool for immunoglobulin, T cell receptor and MHC structural data. Nucleic Acids Res. 32:208-210.

Kaas, Q., and Lefranc, M.-P. (2005) T cell receptor/peptide/MHC molecular characterization and standardized pMHC contact sites in IMGT/3Dstructure-DB. In Silico Biol. 5:505-528.

Kjer-Nielsen, L., Clements, C.S., Purcell, A.W., Brooks, A.G., Whisstock, J.C., Burrows, S.R., McCluskey, J., and Rossjohn, J. (2003) A structural basis for the selection of dominant αβ T cell receptors in antiviral immunity. Immunity 18:53-64.

Lawrence, M.C., and Colman, P.M. (1993) Shape complementarity at protein/protein interfaces. J. Mol. Biol. 234:946-950.

Lefranc, M.-P., and Lefranc, G. (2001) *The T cell receptor FactsBook*. Academic Press, London, 398.

Lefranc, M.-P. (2003a) IMGT, the international ImMunoGeneTics information system® (http://imgt.cines.fr). In: B.K.C. Lo (Ed.), *Antibody Engineering: Methods and Protocols,* 2nd edition. Methods in Molecular Biology. Humana Press, Totowa, NJ, 248, pp. 27-49.

Lefranc, M.-P., Pommié, C., Ruiz, M., Giudicelli, V., Foulquier, E., Truong, L., Thouvenin-Contet, V., and Lefranc, G. (2003b) IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. Dev. Comp. Immunol. 27:55-77.

Lefranc, M.-P. (2004a) IMGT-ONTOLOGY and IMGT databases, tools and Web resources for immunogenetics and immunoinformatics. Mol. Immunol. 40:647-660.

Lefranc, M.-P., Giudicelli, V., Ginestoux, C., Bosc, N., Folch, G., Guiraudou, D., Jabado-Michaloud, J., Magris, S., Scaviner, D., Thouvenin, V., Combres, K., Girod, D., Jeanjean, S., Protat, C., Monod, Y.M., Duprat, E., Kaas, Q., Pommié, C., Chaume, D., and Lefranc, G. (2004b) IMGT-ONTOLOGY for immunogenetics and immunoinformatics. In Silico Biol. 4:17-29.

Lefranc, M.-P., Clément, O., Kaas, Q., Duprat, E., Chastellan, P., Coelho, I., Combres, K., Ginestoux, C., Giudicelli, V., Chaume, D., and Lefranc, G. (2005a) IMGT-Choreography for Immunogenetics and Immunoinformatics. In Silico Biol. 5:6.

Lefranc, M.-P., Duprat, E., Kaas, Q., Tranne, M., Thiriot, A., and Lefranc, G. (2005b) IMGT unique numbering for MHC groove G-DOMAIN and MHC superfamily (MhcSF) G-LIKE-DOMAIN. Dev. Comp. Immunol. 29:917-938.

Lefranc, M.-P., Giudicelli, V., Kaas Q., Duprat, E., Jabado-Michaloud, J., Scaviner, D., Ginestoux, C., Clément, O., Chaume, D., and Lefranc, G. (2005c) IMGT, the international ImMunoGeneTics information system. Nucleic Acids Res. 33:D593-D597.

Lefranc, M.-P., Pommié, C., Kaas, Q., Duprat, E., Bosc, N., Guiraudou, D., Jean, C., Ruiz, M., Da Piedade, L., Rouard, M., Foulquier, E., Thouvenin, V., and Lefranc, G. (2005d) IMGT unique numbering for immunoglobulin and T cell receptor constant domains and Ig superfamily C-like domains. Dev. Comp. Immunol. 29:185-203.

Lesk, A.M., and Chothia, C. (1982) Evolution of proteins formed by β-sheets. II. The core of the immunoglobulin domains. J. Mol. Biol. 160:325-342.

Luz, J.G., Huang, M., Garcia, K.C., Rudolph, M.G., Apostolopoulos, V., Teyton, L., and Wilson, I.A. (2002) Structural comparison of allogeneic and syngeneic T cell receptor-peptide-major histocompatibility complex complexes: A buried alloreactive mutation subtly alters peptide presentation substantially increasing V(β) interactions. J. Exp. Med. 195:1175-1186.

Mandelboim, O., Bar-Haim, E., Vadai, E., Fridkin, M., and Eisenbach, L. (1997) Identification of shared tumor-associated antigen peptides between two spontaneous lung carcinomas. J. Immunol. 159:6030-6036.

Rammensee, H.G., Bachmann, J., Emmerich, N.P.N., Bachor, O.A., and Stevanovie, S. (1999) SYFPEITHI: Database for MHC ligands and peptide motifs. Immunogenetics 50:213-219.

Reinherz, E.L., Tan, K., Tang, L., Kern, P., Liu, J., Xiong, Y., Hussey, R.E., Smolyar, A., Hare, B., Zhang, R., Joachimiak, A., Chang, H.C., Wagner, G., and Wang, J. (1999) The crystal structure of a T cell receptor in complex with peptide and MHC class II. Science 286:1913-1921.

Reiser, J.B., Darnault, C., Guimezanes, A., Gregoire, C., Mosser, T., Schmitt-Verhulst, A.M., Fontecilla-Camps, J.C., Malissen, B., Housset, D., and Mazza, G. (2000) Crystal structure of a T cell receptor bound to an allogeneic MHC molecule. Nat. Immunol. 1:291-297.

Reiser, J.B., Gregoire, C., Darnault, C., Mosser, T., Guimezanes, A., Schmitt-Verhulst, A.M., Fontecilla-Camps, J.C., Mazza, G., Malissen, B., and Housset, D. (2002) A T cell receptor CDR3β loop undergoes conformational changes of unprecedented magnitude upon binding to a peptide/MHC class I complex. Immunity 16:345-354.

Reiser, J.B., Darnault, C., Gregoire, C., Mosser, T., Mazza, G., Kearney, A., van der Merwe, P.A., Fontecilla-Camps, J.C., Housset, D., and Malissen, B. (2003) CDR3 loop flexibility contributes to the degeneracy of TCR recognition. Nat. Immunol. 4:241-247.

Robinson, J., Waller, M.J., Parham, P., de Groot, N., Bontrop, R., Kennedy, L.J., Stoehr, P., and Marsh, S.G. (2003) IMGT/HLA and IMGT/MHC: Sequence databases for the study of the major histocompatibility complex. Nucleic Acids Res. 31:311-314.

Rudolph, M.G., Luz, J.G., and Wilson, I.A. (2002) Structural and thermodynamic correlates of T cell signaling. Annu. Rev. Biophys. Biomol. Struct. 31:121-149.

Scott, C.A., Peterson, P.A., Teyton, L., and Wilson, I.A. (1998) Crystal structures of two I-Ad-peptide complexes reveal that high affinity can be achieved without large anchor residues. Immunity 8:319-329.

Sim, B.C., Zerva, L., Greene, M.I., and Gascoigne, N.R. (1996) Control of MHC restriction by TCR Vα CDR1 and CDR2. Science 273:963-964.

Singh, H., and Raghava, G.P. (2003) ProPred1: Prediction of promiscuous MHC class-I binding sites. Bioinformatics 19:1009-1014.

Stewart-Jones, G.B.E., McMichael, A.J., Bell, J.I., Stuart, D.I., and Jones, E.Y. (2003) A structural basis for immunodominant human T cell receptor recognition. Nat. Immunol. 4:657-663.

Vasmatzis, G., Cornette, J., Sezerman, U., and DeLisi, C. (1996a) TcR recognition of the MHC-peptide dimer: Structural properties of a ternary complex. J. Mol. Biol. 261:72-89.

Vasmatzis, G., Zhang, C., Cornette, J.L., and DeLisi, C. (1996b) Computational determination of side chain specificity for pockets in class I MHC molecules. Mol. Immunol. 33: 1231-1239.

Wain, H.M., Bruford, E.A., Lovering, R.C., Lush, M.J., Wright, M.W., and Povey, S. (2002) Guidelines for human gene nomenclature. Genomics 79:464-470.

Wang, J.H., Meijers, R., Xiong, Y., Liu, J.H., Sakihama, T., Zhang, R., Joachimiak, A., and Reinherz, E.L. (2001) Crystal structure of the human CD4 N-terminal two-domain fragment complexed to a class II MHC molecule. Proc. Natl. Acad. Sci. USA 98:10799-10804.

Wang, J.H., and Reinherz, E.L. (2002) Structural basis of T cell recognition of peptides bound to MHC molecules. Mol. Immunol. 38:1039-1049.

Yousfi Monod, M., Giudicelli, V., Chaume, D., and Lefranc, M.-P. (2004) IMGT/JunctionAnalysis: The first tool for the analysis of the immunoglobulin and T cell receptor complex V-J and V-D-J junctions. Bioinformatics 20:I379-I385.

Zhang, C., Anderson, A., and DeLisi, C. (1998) Structural principles that govern the peptide-binding motifs of class I MHC molecules. J. Mol. Biol. 281:929-947.