Chapter 25

IMGT/3Dstructure-DB: T-Cell Receptor TR Paratope and Peptide/Major Histocompatibility pMH Contact Sites and Epitope

Marie-Paule Lefranc and Gérard Lefranc

Abstract

T-cell receptors (TR), the antigen receptors of T cells, specifically recognize peptides presented by the major histocompatibility (MH) proteins, as peptide/MH (pMH), on the cell surface. The structure characterization of the trimolecular TR/pMH complexes is crucial to the fields of immunology, vaccination, and immunotherapy. IMGT/3Dstructure-DB is the three-dimensional (3-D) structure database of IMGT®, the international ImMunoGenetics information system®. By its creation, IMGT® marks the advent of immunoinformatics, which emerged at the interface between immunogenetics and bioinformatics. The IMGT® immunoglobulin (IG) and TR gene and allele nomenclature (CLASSIFICATION axiom) and the IMGT unique numbering and IMGT/Collier-de-Perles (NUMEROTATION axiom) are the two founding breakthroughs of immunoinformatics. IMGT-ONTOLOGY concepts and IMGT Scientific chart rules generated from these axioms allowed IMGT® bridging genes, structures, and functions. IMGT/3Dstructure-DB contains 3-D structures of IG or antibodies, TR and MH proteins of the adaptive immune responses of jawed vertebrates (*gnathostomata*), IG or TR complexes with antigens (IG/Ag, TR/pMH), related proteins of the immune system of any species belonging to the IG and MH superfamilies, and fusion proteins for immune applications. The focus of this chapter is on the TR V domains and MH G domains and the contact analysis comparison in TR/pMH interactions. Standardized molecular characterization includes "IMGT pMH contact sites" for peptide and MH groove interactions and "IMGT paratopes and epitopes" for TR/pMH complexes. Data are available in the IMGT/3Dstructure database, at the IMGT Home page http://www.imgt.org.

Key words IMGT, T-cell receptor, CDR-IMGT, Major histocompatibility, Paratope, Epitope, TR/pMH, IMGT-ONTOLOGY, Immunoinformatics, IMGT/3Dstructure-DB

1 Introduction

The adaptive immune responses were acquired by jawed vertebrates (or *gnathostomata*) more than 450 million years ago and are found in all extant jawed vertebrate species from fishes to humans [1]. The adaptive immune responses are characterized by a remarkable specificity and memory, which are the properties of the B and T cells owing to an extreme diversity of their antigen receptors [1]. The

Anton W. Langerak (ed.), *Immunogenetics: Methods and Protocols*, Methods in Molecular Biology, vol. 2453, https://doi.org/10.1007/978-1-0716-2115-8_25, © The Author(s) 2022

specific antigen receptors comprise the immunoglobulins (IG) or antibodies of the B cells and plasma cells [2–5] and the T-cell receptors (TR) [6]. Whereas the IG recognize antigens in their native (unprocessed) form, the TR recognize processed antigens, which are presented as peptides by the highly polymorphic major histocompatibility (MH) proteins (in humans HLA for human leukocyte antigens, encoded by genes in the MHC locus) (Fig. 1). T cells are involved in cell-mediated immune response, against a stress of viral, bacterial, fungal, or tumoral origin and identify antigenic peptides presented by the MH proteins as

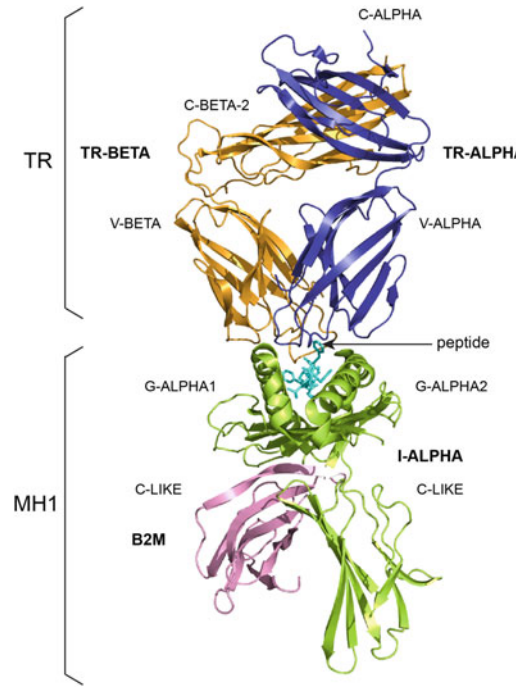


Fig. 1 A T-cell receptor (TR)/peptide-major histocompatibility 1 (pMH1) complex. A TR (here, TR-alpha_beta) is shown (on top, upside down) in complex with an MH (here, MH1) presenting a peptide in its groove [1]. In vivo, a TR is anchored in the membrane of a T cell as part of the signaling T-cell receptor (TcR = TR +CD3). A TR is made of two chains, each comprising a variable domain (V-DOMAIN) at the N-terminal end and a constant domain (C-DOMAIN) at the C-terminal end. The domains are V-ALPHA and C-ALPHA for the TR-ALPHA chain and V-BETA and C-BETA for the TR-BETA chain. An MH1 is made of the I-ALPHA chain with two G-DOMAIN (G-ALPHA1 and G-ALPHA2) and a C-LIKE-DOMAIN (C-LIKE), noncovalently associated with the B2M (a C-LIKE-DOMAIN). In this representation (with G-ALPHA1 on the left, G-ALPHA2 and B2M on the right), the peptide is oriented in the groove from front of the figure to back. The TR/pMH1 complex structure is 3qfj from IMGT/3Dstructure-DB (http:/www.imgt.org). (With permission from M-P. Lefranc and G. Lefranc, LIGM, Founders and Authors of IMGT®, the international ImMunoGeneTics information system®, http:/www.imgt.org)

peptide/MH (pMH) on cell surface [1]. The recognition and signal transduction are carried out by the multiprotein bifunctional T-cell receptor (TcR) assembly that comprises the TR responsible of the specific pMH recognition plus the associated transmembrane signaling CD3 proteins [6]. The TcR is itself associated, in the immunological synapse, with the CD4 or CD8 coreceptors, to the activating CD28 and inhibitory CTLA4 costimulatory proteins, to the CD2 adhesion molecule and to intracellular kinases. The CD8 expressed on most cytotoxic T cells binds the MH class I (MH1) that is expressed ubiquitously on cells of the organism [7]. The CD4 expressed on most helper T cells binds the MH class II (MH2) that is expressed by professional antigen presenting cells (dendritic cells, macrophages, monocytes, and B cells) [7].

IMGT®, the international ImMunoGeneTics information system®, http://www.imgt.org, is the global reference in immunogenetics and immunoinformatics [1], founded in 1989 by Marie-Paule Lefranc at Montpellier (Université de Montpellier and CNRS). It is a high quality integrated knowledge resource comprising 7 databases, 17 online tools, and more than 25,000 pages of web resources [8–11]. IMGT® is specialized in the sequences, structures, and genetic data of the IG, TR, and MH of human and other vertebrate species, in the immunoglobulin superfamily (IgSF) and the MH superfamily (MhSF) of vertebrates and invertebrates, and in related proteins of the immune system (RPI), fusion protein for immune applications, and composite proteins for clinical applications [1, 8–11]. IMGT/3Dstructure-DB [12–14] is the three-dimensional (3-D) structure database of IMGT®. This database provides the standardized IMGT annotation and analysis of the 3-D structures of the TR, pMH, and TR/pMH complexes and comprises detailed molecular characterization and description of their interactions [15–17]. The standardized analysis is based on the concepts of IMGT-ONTOLOGY, the first ontology in immunogenetics and immunoinformatics [18–24]. The IMGT-ONTOLOGY concepts are generated from seven axioms [25–31], of which the CLASSIFICATION axiom (IG and TR gene and allele nomenclature) (*see* **Note 1**) at the birth of IMGT® and immunoinformatics [1, 25] and the NUMEROTATION axiom (IMGT unique numbering [7, 26–27, 32–35] and IMGT Colliers de Perles [28, 29, 36–39]) allow bridging sequences, structures, and functions. The IMGT unique numbering for variable (V) domain includes the IG and TR V-DOMAIN and the V-like domains of IgSF other than IG and TR [32–34]. The IMGT unique numbering for constant (C) domain includes the IG and TR C-DOMAIN and the C-like domains of IgSF other than IG and TR [35]. The IMGT unique numbering for G domain includes the groove (G) domains of the MH G-DOMAIN and the G-like domains of MhSF other than MH (or RPI-MH1Like) [7]. The IMGT/DomainGapAlign tool [13, 40, 41] analyzes the amino acid

sequences of the V, C, and G domains using the IMGT unique numbering [7, 34, 35] and provides a direct link to the IMGT/Collier-de-Perles tool [39]. The IMGT Scientific chart rules provide a standardized description of the contact analysis [15–17] and comparison of TR/pMH complexes and their interactions, irrespective of the TR chains and domains, the MH class (MH1 or MH2), or the species (*Homo sapiens*, *Mus musculus,* etc.). Eleven "IMGT pMH contact sites" were defined for the comparison of pMH interactions, regardless of the peptide lengths [15–17]. The "IMGT pMH contact sites" visualize the interactions between the amino acids (AA) (*see* **Note 2**) of the peptide and those of the MH groove based on the contact analysis. They are a useful asset in peptide vaccine design and epitope prediction, and they precisely identify and visualize AA of the peptide located in the MH2 groove. The standardized "IMGT paratope and epitope" for TR/pMH complexes comprises the TR paratope and the pMH epitope, determined from contact analysis, in IMGT/3Dstructure-DB, at the IMGT Home page http://www.imgt.org.

2 TR and MH Standardized Description in IMGT/3Dstructure-DB

2.1 TR and MH Chains and Domains

2.1.1 TR Chains and Domains

The TR is made of two chains, an alpha chain (TR-ALPHA) and a beta chain (TR-BETA) for the TR-ALPHA_BETA receptor and a gamma chain (TR-GAMMA) and a delta chain (TR-DELTA) for the TR-GAMMA_DELTA receptor [6] (Table 1). Each complete TR chain comprises an extracellular region made up of a V-DOMAIN) (for instance, V-ALPHA for the alpha chain) and a C-DOMAIN) (for instance, C-ALPHA for the alpha chain), a

Table 1
IMGT standardized labels for the DESCRIPTION of the T-cell receptors (TR) and of their chains and domains. IMGT® labels (concepts of description) are written in capital letters [1]

IMGT receptor description	IMGT chain description	IMGT domain description	IMGT region labels
TR-ALPHA_BETA	TR-ALPHA	V-ALPHA	V-J-REGION
		C-ALPHA	Part of C-REGION[a]
	TR-BETA	V-BETA	V-D-J-REGION
		C-BETA	Part of C-REGION[a]
TR-GAMMA_DELTA	TR-GAMMA	V-GAMMA	V-J-REGION
		C-GAMMA	Part of C-REGION[a]
	TR-DELTA	V-DELTA	V-D-J-REGION
		C-DELTA	Part of C-REGION[a]

[a]The TR chain C-REGION also includes the CONNECTING-REGION (CO), the TRANSMEMBRANE-REGION (TM), and the CYTOPLASMIC-REGION (CY), which are not present in the 3-D structures (IMGT® http://www.imgt.org, IMGT Scientific chart >1. Sequence and 3D structure identification and description > Correspondence between labels for IG and TR domains in IMGT/3Dstructure-DB and IMGT/LIGM-DB)

connecting region (CONNECTING-REGION (CO)), a transmembrane region (TRANSMEMBRANE-REGION (TM)), and a short cytoplasmic region (CYTOPLASMIC-REGION (CY)) [6, 7] (Fig. 2, Table 1). The TR V domains that are directly involved in the TR/pMH interactions are described in Subheading 2.2.

2.1.2 MH Chains and Domains

The MH1 is formed by the association of a heavy chain (I-ALPHA) and a light chain (beta-2-microglobulin or B2M). The MH2 is an heterodimer formed by the association of an alpha chain (II-ALPHA) and a beta chain (II-BETA) [7] (Table 2) The I-ALPHA chain of the MH1 and the II-ALPHA and II-BETA chains of the MH2 comprise an extracellular region, made of three domains for the MH1 chains and of two domains for the MH2 chains, and CO, TM, and CY regions [7] (Fig. 2, Table 2). The I-ALPHA chain comprises two groove domains (G-DOMAIN), G-ALPHA1 [D1] and G-ALPHA2 [D2], and one C-LIKE domain [D3] [7]. The B2M corresponds to a single C-LIKE domain. The II-ALPHA chain and the II-BETA chain each comprises two domains, G-ALPHA [D1] and C-LIKE [D2], and G-BETA [D1] and C-LIKE [D2] [7] (Fig. 2). Only the extracellular region that corresponds to these domains has been crystallized. The MH G domains that are directly involved in the TR/pMH interactions are described in Subheading 2.3.

2.2 TR V Domains

2.2.1 Definition

A V domain [32–34] comprises about 100 AA and is made of nine antiparallel beta strands (A, B, C, C′, C″, D, E, F, and G) linked by beta turns (AB, CC′, C″D, DE, and EF) or loops (BC, C′C″, and FG) and forming a sandwich of two sheets (Table 3). The sheets are closely packed against each other through hydrophobic interactions giving a hydrophobic core and joined together by a disulfide bridge between first-CYS at position 23 in the B-STRAND in the first sheet and the second-CYS 104 in the F-STRAND in the second sheet [34]. The V domain type includes the V-DOMAIN of the TR (and IG), which corresponds to the V-J-REGION or V-D-J-REGION encoded by V-(D)-J rearrangements [1–6, 36], and the V-LIKE-DOMAIN of the IgSF other than IG and TR [37–44]. In a V-DOMAIN, the three hypervariable loops BC, C′C″, and FG involved in the ligand (antigen for IG or pMH for TR) recognition are designated as complementarity determining regions (CDR-IMGT) [1–6].

2.2.2 IMGT Unique Numbering for V Domain

The V domain strands and loops and their delimitations and lengths are based on the IMGT unique numbering for V domain (V-DOMAIN and V-LIKE-DOMAIN) [33, 34] (Table 3). In the IG and TR V-DOMAIN, the G-STRAND is the C-terminal part of the J-REGION, with J-PHE or J-TRP 118 and the canonical motif F/W-G-X-G at positions 118–121 [1]. The loop length (number of AA (or codons), which is the number of occupied positions, is a

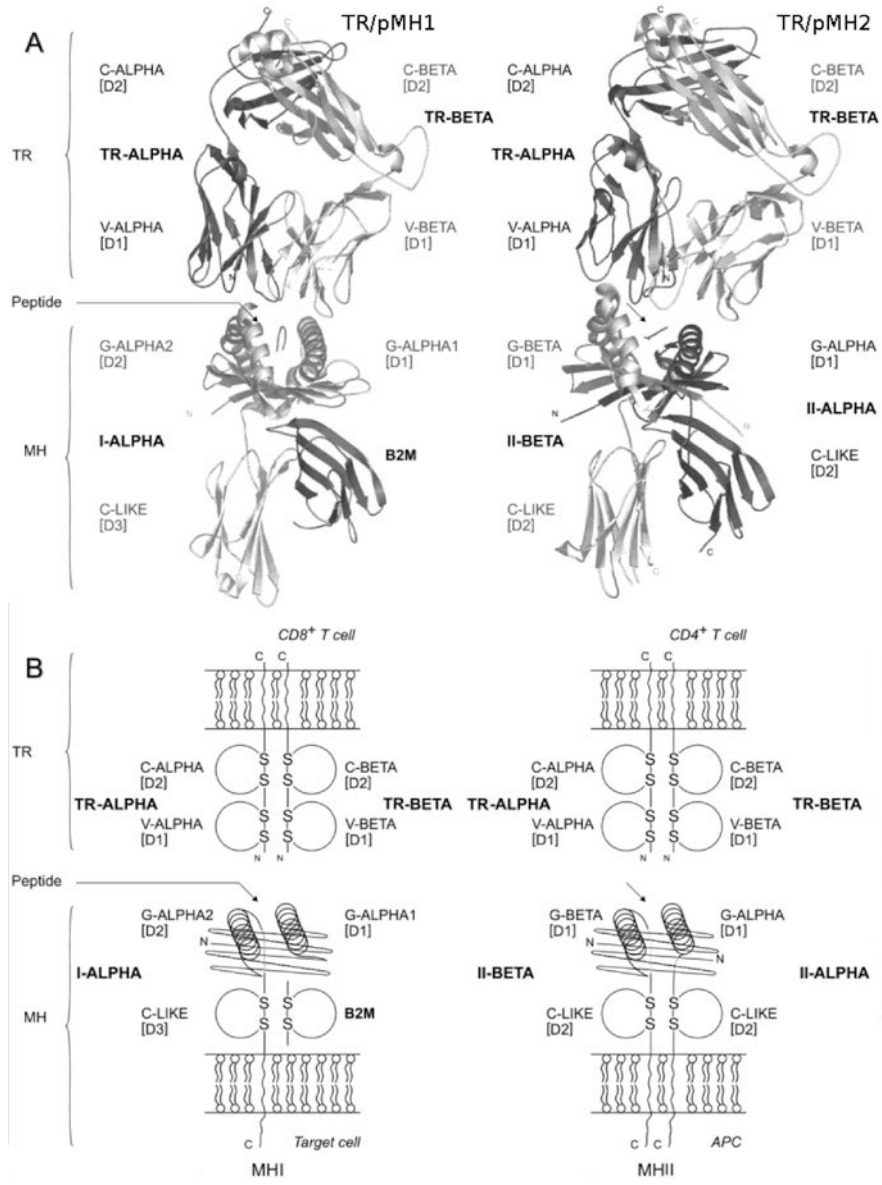


Fig. 2 T-cell receptor/peptide/MH complexes with MH class I (TR/pMH1) and MH class II (TR/pMH2). (**a**) 3-D structures of TR/pMH1 and TR/pMH2. (**b**) Schematic representation of TR/pMH1 and TR/pMH2. The TR (TR-ALPHA and TR-BETA chains), the MH1 (I-ALPHA and B2M chains), and the MH2 (II-ALPHA and II-BETA 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. [D1], [D2], and [D3] indicate the domains. Arrows indicate the peptide localization in the MH groove made of two G-DOMAIN [7]. In these representations (with G-ALPHA1 on the right, G-ALPHA2 and B2M on the left), the peptide is oriented in the groove from back of the figures to front. (With permission from M-P. Lefranc and G. Lefranc, LIGM, Founders and Authors of IMGT®, the international ImMunoGeneTics information system®, http:/www.imgt.org)

Table 2
IMGT-standardized labels for the DESCRIPTION of the major histocompatibility (MH) and of their chains and domains. IMGT® labels (concepts of description) are written in capital letters [1]

IMGT receptor description	IMGT chain description	IMGT domain description	IMGT domain number
MHC-I-ALPHA_B2M	I-ALPHA	G-ALPHA1	[D1]
		G-ALPHA2	[D2]
		C-LIKE-DOMAIN	[D3][a]
	B2M	C-LIKE-DOMAIN	[D]
MHC-II-ALPHA_BETA	II-ALPHA	G-ALPHA	[D1]
		C-LIKE-DOMAIN	[D2][a]
	II-BETA	G-BETA	[D1]
		C-LIKE-DOMAIN	[D2][a]

[a]The I-ALPHA, II-ALPHA and II-BETA chains includes at the C-terminal end of the C-LIKE-DOMAIN, the CONNECTING-REGION (CO), the TRANSMEMBRANE-REGION (TM), and the CYTOPLASMIC-REGION (CY), which are not present in the 3-D structures

Table 3
V domain strands and loops, IMGT positions and lengths, based on the IMGT unique numbering for V domains (V-DOMAIN and V-LIKE-DOMAIN) [33, 34]

V domain strands and loops[a]	IMGT positions	Lengths[b]	Characteristic Residue@Position[c]	V-DOMAIN FR-IMGT and CDR-IMGT
A-STRAND	1–15	15 (14 if gap at 10)		FR1-IMGT
B-STRAND	16–26	11	1st-CYS 23	
BC-LOOP	27–38	12 (or less)		CDR1-IMGT
C-STRAND	39–46	8	CONSERVED-TRP 41	FR2-IMGT
C′-STRAND	47–55	9		
C′C″-LOOP	56–65	10 (or less)		CDR2-IMGT
C″-STRAND	66–74	9 (or 8 if gap at 73)		FR3-IMGT
D-STRAND	75–84	10 (or 8 if gaps at 81, 82)		
E-STRAND	85–96	12	Hydrophobic 89	
F-STRAND	97–104	8	2nd-CYS 104	
FG-LOOP	105–117	13 (or less, or more)		CDR3-IMGT
G-STRAND	118–128	11 (or 10)	V-DOMAIN J-PHE 118 or J-TRP 118[d]	FR4-IMGT

[a]IMGT® labels (concepts of description) are written in capital letters
[b]In number of AA (or codons)
[c]*See* Subheading 2.4
[d]In the IG and TR V-DOMAIN, the G-STRAND (or FR4-IMGT) is the C-terminal part of the J-REGION, with J-PHE or J-TRP 118 and the canonical motif F/W-G-X-G at positions 118–121. The JUNCTION refers to the CDR3-IMGT plus the two anchors second-CYS 104 and J-PHE or J-TRP 118 [1]

crucial and original concept of IMGT-ONTOLOGY. The lengths of the loops BC (or CDR1-IMGT), C′C″, (or CDR2-IMGT) and FG (or CDR3-IMGT) characterize the V-DOMAIN (Table 3). They are delimited by anchor positions (*see* **Note 3**). The BC loop (or CDR1-IMGT) comprises positions 27–38, the C′C″ (or CDR2-IMGT) positions 56–65, and the FG (or CDR3-IMGT) positions 105–117. In a V-DOMAIN, the CDR3-IMGT that encompasses the V-(D)-J junction resulting from V-J or V-D-J rearrangements [1] is more variable in sequence and length than the CDR1-IMGT and CDR2-IMGT that are encoded by the V-REGION only. The lengths of the three loops BC, C′C″, and FG are shown in number of AA (or codons), into brackets and separated by dots. For example, [9.6.9] means that the BC, C′C″, and FG loops (or CDR1-IMGT, CDR2-IMGT, and CDR3-IMGT for a V-DOMAIN) have a length of 9, 6, and 9 AA (or codons), respectively.

2.2.3 IMGT Colliers de Perles for V Domain

The V domain nine strands are indicated, with their orientation, in the IMGT Colliers de Perles [28, 29, 32–34, 36–39], which are IMGT 2D graphical representations based on the IMGT unique numbering. IMGT Colliers de Perles of the TR V-ALPHA and V-BETA domains from 1ao7 (*see* **Note 4**) a TR/pMH1 3-D structure complex are shown as examples (Fig. 3). The V-ALPHA and V-BETA domains share the main conserved characteristics of the V-DOMAIN, which are the disulfide bridge between cysteine 23 (first-CYS) and cysteine 104 (second-CYS), and the three other hydrophobic core residues tryptophan 41 (CONSERVED-TRP), leucine (or hydrophobic) 89, and phenylalanine 118 (J-PHE) (*see* **Note 5**). In Fig. 3, the V-ALPHA (1ao7_D chain; [6.6.11]) has a CDR1-IMGT and a.

CDR2-IMGT of 6 AA and a CDR3-IMGT of 11 AA, whereas the V-BETA (1ao7_E chain [5.6.14]) has a CDR1-IMGT, CDR2-IMGT, and CDR3-IMGT of 5, 6, and 14 AA, respectively (Subheading 2.2.2) (*see* **Note 6**). In IMGT/3Dstructure-DB, the IMGT genes and alleles that contribute to the V-DOMAIN are determined automatically by IMGT/DomainGapAlign [13, 40, 41], based on the standardized IMGT nomenclature [1, 2, 6] and IMGT unique numbering [34]. Thus, the V-ALPHA of 1ao7_D corresponds to *Homo sapiens* TRAV12-2*02-TRAJ24*02 and the V-BETA of 1ao7_E corresponds to *Homo sapiens* TRBV6-5*01-(TRBD2)-TRBJ2-7*01 [16, 17].

2.3 MH G Domains

2.3.1 Definition

A G domain [7] comprises about 90 AA and is made of a sheet of four antiparallel beta strands linked by turns and of a helix (Table 4); the helix sits on the beta strands, its axis forming an angle of about 40 degrees with the strands [16, 17]. Two G domains are needed to form the MhSF groove made of a “floor” and two “walls” [7]. Each G domain contributes by its four strands

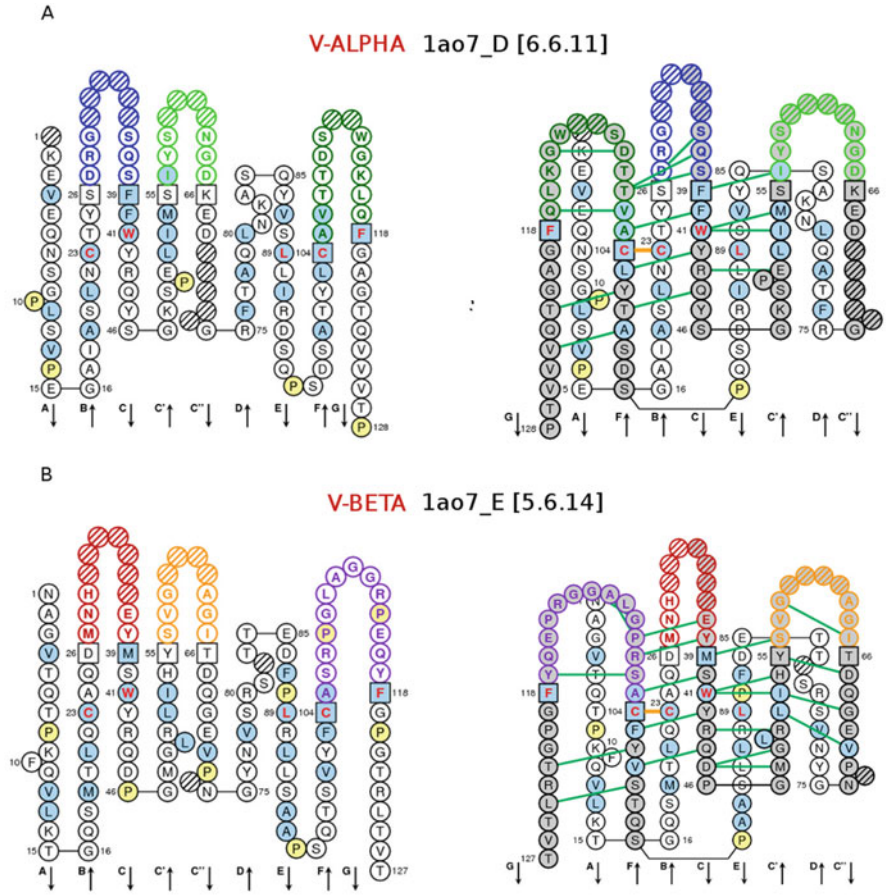


Fig. 3 IMGT/Collier-de-Perles for TR V domain (V-DOMAIN). (**a**) IMGT/Collier-de-Perles for TR V-ALPHA (chain 1ao7_D). The CDR-IMGT lengths are [6.6.11]. (**b**) IMGT/Collier-de-Perles for TR V-BETA (chain 1ao7_E). The CDR-IMGT lengths are [5.6.14]. AA ais shown in the one-letter abbreviation (*see* **Note 2**). Position at which hydrophobic AA (hydropathy index with positive value: I, V, L, F, C, M, A) and tryptophan (W) are found in more than 50% of analyzed sequences are shown in blue, online. All proline (P) are shown in yellow, online. Anchor positions are shown in squares (*see* **Note 3**). Arrows indicate the direction of the beta strands [28, 29]. Hatched circles correspond to missing positions according to the IMGT unique numbering for V domain [33, 34]. IMGT color menu for CDR1-IMGT, CDR2-IMGT, and CDR3-IMGT is blue, green, and greenblue, for V-ALPHA, and red, orange and purple, for V-BETA (*see* **Note 6**). IMGT/Collier-de-Perles are shown on one layer (on the left hand side) and two layers (on the right hand side). The IMGT Colliers de Perles on two layers show, in the forefront, the GFCC′C″ strands and, in the back, the ABED strands. Hydrogen bonds (from the IMGT/3Dstructure-DB entry) are show in green, online. Only those between the AA of the C, C′, C″, F, and G strands (in the forefront) and those of the CDR-IMGT are shown here. IMGT/Collier-de-Perles are from IMGT/3Dstructure-DB, http:/www.imgt.org. (With permission from M-P. Lefranc and G. Lefranc, LIGM, Founders and Authors of IMGT®, the international ImMunoGeneTics information system®, http:/www.imgt.org)

and turns to half of the groove floor and by its helix to one wall of the groove [7, 16, 17]. The G domain type includes the G-DOMAIN of the MH [7] and the G-LIKE-DOMAIN of the MhSF other than MH or RPI-MH1Like [7, 45, 46] (*see* **Note 7**).

Table 4
G domain strands, turns, and helix, IMGT positions and lengths, based on the IMGT unique numbering for G domains (G-DOMAIN and G-LIKE-DOMAIN) [7]

G domain strands, turns and helix[a]	IMGT positions	Lengths[b]	Characteristic Residue@Position[c] and additional positions[d]
A-STRAND	1–14	14	7A, CYS-11
AB-TURN	15–17	3 (or 2 or 0)	
B-STRAND	18–28	11 (or 10[e])	
BC-TURN	29–30	2	
C-STRAND	31–38	8	
CD-TURN	39–41	3 (or 1[f])	
D-STRAND	42–49	8	49.1 to 49.5
HELIX	50–92	43 (or less or more)	54A, 61A, 61B, 72A, CYS-74, 92A

[a]IMGT® labels (concepts of description) are written in capital letters
[b]In number of AA (or codons)
[c]*See* Subheading 2.4
[d]For details on the characteristic Residue@Position and additional positions, *see* Ref. [7]
[e]Or 9 in some G-BETA
[f]Or 0 in some G-ALPHA2-LIKE [7]

2.3.2 IMGT Unique Numbering for G Domain

The G domain strands, turns, and helix and their delimitations and lengths are detailed in Table 4, based on the IMGT unique numbering for G domain (G-DOMAIN and G-LIKE-DOMAIN) [7].

2.3.3 IMGT Colliers de Perles for G Domain

The MH groove in which the peptide binds is made of 2 G-DOMAIN, belonging to the same chain (I-ALPHA) for MH1 or to two different chains (II-ALPHA and II-BETA) for MH2 [7, 37–41]. For the RPI-MH1Like (*see* **Note 7**), the 2 G-LIKE-DOMAIN also belong, as for the MH1, to the same chain (I-ALPHA-LIKE) [7, 37–41]. The IMGT Colliers de Perles (Fig. 4) show, in the upper part of the groove representation, G-ALPHA1 ([D1] of I-ALPHA chain), G-ALPHA ([D1] of II-ALPHA chain) or G-ALPHA1-LIKE ([D1] of I-ALPHA-LIKE chain), and, respectively, in the lower part of the groove representation, G-ALPHA2 ([D2] of I-ALPHA chain), G-BETA ([D1] of II-BETA chain), or G-ALPHA2-LIKE ([D2] of I-ALPHA-LIKE chain). IMGT Colliers de Perles for the MH1 G-ALPHA1 and G-ALPHA2 (1ao7_A chain) are represented in Fig. 4a. IMGT Colliers de Perles for the MH2 G-ALPHA and G-BETA (1j8h_A and 1j8h_B chains, respectively) are represented in Fig. 4b. In IMGT/3Dstructure-DB, the IMGT genes and alleles that encode the G-DOMAIN are determined automatically by IMGT/

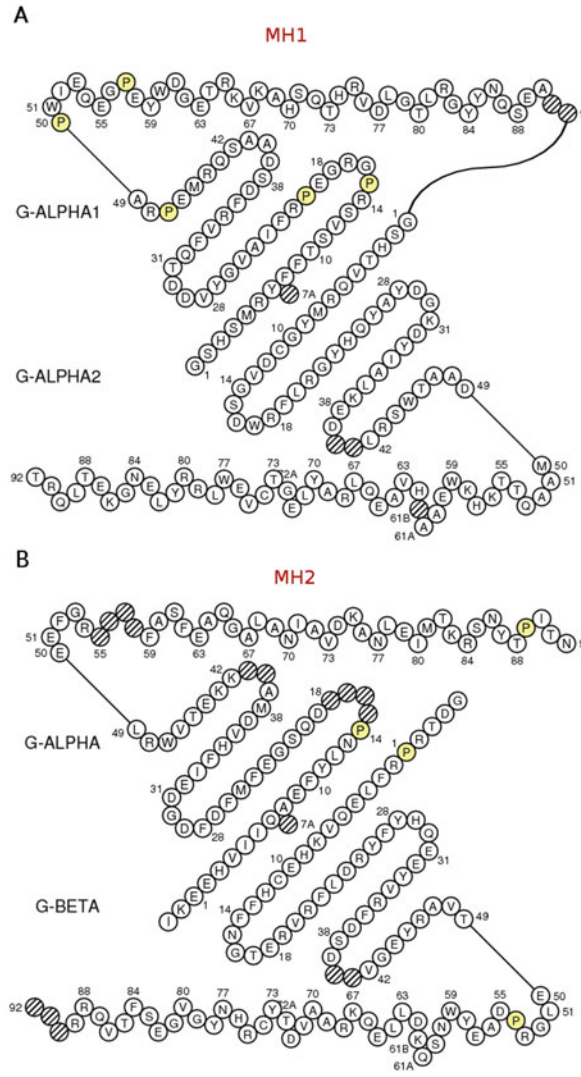


Fig. 4 IMGT/Collier-de-Perles of MH G domains (G-DOMAIN). (**a**) MH1 G-ALPHA1 and G-ALPHA2 domains from 1ao7 (I-ALPHA chain 1ao7_A). (**b**) MH2 G-ALPHA and G-BETA domains from 1j8h (II-ALPHA chain 1j8h_A and II-BETA 1j8h_B, respectively). AA positions and gaps (hatched positions) are according to the IMGT unique numbering for G domain [7]. 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/Collier-de-Perles) [7]. IMGT/Collier-de-Perles are from IMGT/3Dstructure-DB, http:/www.imgt.org. G-domain terminal hatched positions (MH1 G-ALPHA1 91 and 92 and MH2 G-BETA 90, 91 and 92) are not reported in online IMGT/Collier-de-Perles. The IMGT Colliers de Perles can also be obtained, with the sequences gapped by IMGT/DomainGapAlign [40, 41], using the IMGT/Collier-de-Perles tool [39]. (With permission from M-P. Lefranc and G. Lefranc, LIGM, Founders and Authors of IMGT®, the international ImMunoGeneTics information system®, http:/www.imgt.org)

DomainGapAlign [13, 40, 41], based on the standardized IMGT nomenclature and numbering [1, 7]. Thus the G-ALPHA1 and G-ALPHA2 of 1ao7_A are encoded by HLA-A*0201 [15–17].

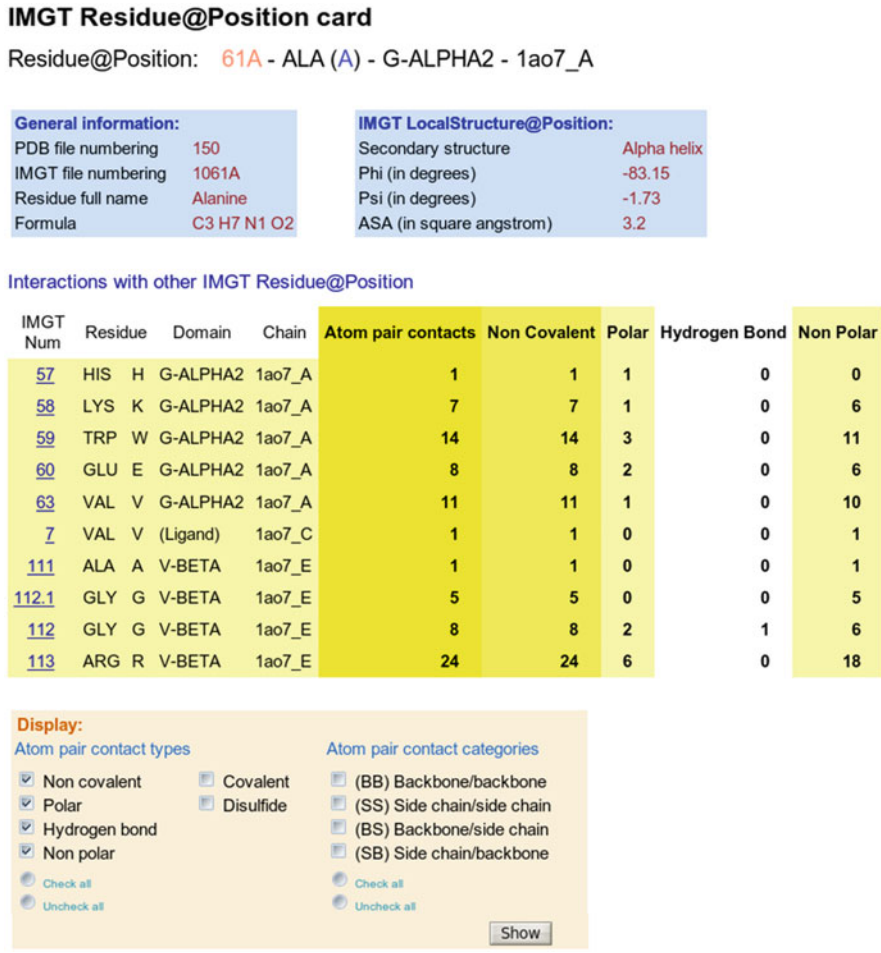
2.4 Residue@Position and Atom Pair Contacts

"Residue@Position" is an IMGT® concept of numerotation that numbers the position of a given residue (or by extension that of a conserved property AA class [47]), based on the IMGT unique numbering (*see* **Note 8**). A "Residue@Position" (R@P) is defined by the position numbering according to the IMGT unique numbering [7, 34, 35], the residue name (in the 3-letter abbreviation and/or in the one-letter abbreviation) (*see* **Note 2**), the IMGT domain label (Tables 1 and 2), and either the gene and allele name for AA sequences (*see* **Note 1**), or the "IMGT chain ID" for 3-D structures. In IMGT/3Dstructure-DB, a 'Residue@Position is described in a "Residue@Position card" (Fig. 5) that provides information on its characteristics (*see* **Note 9**) and the list of the other R@P with which it interacts [16, 17]. Each interaction is characterized by the total number of "atom pair contacts" (*see* **Note 10**) and, as selected by the user for display, the number of atom pair contacts per type ("noncovalent," "polar," "hydrogen bond," "non polar," "covalent," or "disulfide") and/or per category ("(BB) Backbone/backbone," "(SS) Side chain/side chain," "(BS) Backbone/side chain," and "(SB) Side chain/backbone") [16, 17].

3 IMGT pMH Contact Analysis

3.1 IMGT pMH Contact Sites Definition and Determination

"IMGT pMH contact sites" [15–17] highlight the contacts between the amino acids of a presented peptide and those of the floor and helix walls of the MH groove, in 3-D structures of pMH and TR/pMH complexes [12–14]. The "IMGT pMH contact sites" are visualized in IMGT Colliers de Perles for G-DOMAIN [7]. The "IMGT pMH contact sites" provide a standardized comparison of the interactions between a presented peptide and the MH, regardless of the MH class (MH1 or MH2), the G domain (G-ALPHA1, G-ALPHA2, G-ALPHA, and G-BETA), and the peptide length. The "IMGT pMH contact sites" also allow one to precisely identify the AA that is effectively bound in the MH groove. This is particularly informative for the peptides bound to MH2 as these peptides can be much longer than the actual groove length with the N-terminal and C-terminal ends extending outside the groove [7]. In order to deal with different peptide lengths in the groove, 11 standard "IMGT pMH contact sites" were defined (C1–C11) [15–17] (Fig. 6). They correspond to a theoretical maximum length of 11 AA in the groove. This means that, in 3-D structures, some (usually two or three) "IMGT pMH contact sites" are absent as peptides are shorter than 11 AA (usually nine or eight AA long).



IMGT Residue@Position card

Residue@Position: 61A - ALA (A) - G-ALPHA2 - 1ao7_A

General information:	
PDB file numbering	150
IMGT file numbering	1061A
Residue full name	Alanine
Formula	C3 H7 N1 O2

IMGT LocalStructure@Position:	
Secondary structure	Alpha helix
Phi (in degrees)	-83.15
Psi (in degrees)	-1.73
ASA (in square angstrom)	3.2

Interactions with other IMGT Residue@Position

IMGT Num	Residue		Domain	Chain	Atom pair contacts	Non Covalent	Polar	Hydrogen Bond	Non Polar
57	HIS	H	G-ALPHA2	1ao7_A	1	1	1	0	0
58	LYS	K	G-ALPHA2	1ao7_A	7	7	1	0	6
59	TRP	W	G-ALPHA2	1ao7_A	14	14	3	0	11
60	GLU	E	G-ALPHA2	1ao7_A	8	8	2	0	6
63	VAL	V	G-ALPHA2	1ao7_A	11	11	1	0	10
7	VAL	V	(Ligand)	1ao7_C	1	1	0	0	1
111	ALA	A	V-BETA	1ao7_E	1	1	0	0	1
112.1	GLY	G	V-BETA	1ao7_E	5	5	0	0	5
112	GLY	G	V-BETA	1ao7_E	8	8	2	1	6
113	ARG	R	V-BETA	1ao7_E	24	24	6	0	18

Display:

Atom pair contact types: Non covalent, Polar, Hydrogen bond, Non polar, Covalent, Disulfide; Check all, Uncheck all

Atom pair contact categories: (BB) Backbone/backbone, (SS) Side chain/side chain, (BS) Backbone/side chain, (SB) Side chain/backbone; Check all, Uncheck all

Show

Fig. 5 IMGT Residue@Position card. The "Residue@Position: 61A—ALA (A)—G-ALPHA2—1ao7_A" is defined by the position numbering ("61A") according to the IMGT unique numbering for G domain [7], the residue name in the three-letter abbreviation and in the one-letter abbreviation for AA ("ALA (A)") (*see* **Note 2**), the IMGT domain label (G-ALPHA2) (Table 2) and the IMGT chain ID (1ao7_A) (*see* **Note 4**). The list of atom pair contacts shows that this R@P interacts with 5 R@P of the same domain (G-ALPHA2) and, of interest for the TR/pMH interactions, with 4 R@P of the V-BETA and one of the peptide (Ligand). The "Residue@Position" card is from IMGT/3Dstructure-DB, http:/www.imgt.org. (With permission from M-P. Lefranc and G. Lefranc, LIGM, Founders and Authors of IMGT®, the international ImMunoGeneTics information system®, http:/www.imgt.org)

The peptide binding mode to MH1 is characterized by the N-terminal and C-terminal peptide ends docked deeply with the C1 and C11 contact sites (red and pink, respectively, in the IMGT/Collier-de-Perles) and by the peptide length that mechanically constrains the peptide conformation in the groove. Thus, for a peptide of 10 AA, one "IMGT pMH contact sites" is absent (C2), and for a peptide of 9 AA, two "IMGT pMH contact sites" are absent (C2 and C7), whereas for a peptide of 8 AA, three pMH contact sites are absent (C2, C7, and C8) [15–17] (*see* **Note 11**).

Standard 'IMGT pMH contact sites'

	MH1 bound peptides			MH2 bound peptides
	8-AA peptides	9-AA peptides	10-AA peptides	9 AA in the groove
C1	1	1	1	1
C2	-	-	-	2
C3	2	2	2	3
C4	3	3	3	4
C5	4	4	4	5
C6	5	5	5	6
C7	-	-	6	-
C8	-	6	7	-
C9	6	7	8	7
C10	7	8	9	8
C11	8	9	10	9

Fig. 6 Standard "IMGT pMH contact sites'. Eleven standard 'IMGT pMH contact sites' (C1 to C11) were defined for the standardized analysis and comparison of pMH interactions [16, 17]. (With permission from M-P. Lefranc and G. Lefranc, LIGM, Founders and Authors of IMGT®, the international ImMunoGeneTics information system®, http:/www.imgt.org)

The peptide binding mode to MH2 is different with the peptide lying in the groove. Thus, for nine amino acids lying in an MH2 groove, C2 is present but there are no C7 and C8. For a given 3-D structure in IMGT/3Dstructure-DB, the determination of the "IMGT pMH contact sites" combines contact analysis between the peptide and the MH (in a pMH or in a TR/pMH complex), with an interaction scoring function (*see* **Note 12**). The MH AA automatically selects the highest score that is listed and displayed in "IMGT/Collier-de-Perles with pMH contact sites." The characterization of the "IMGT pMH contact sites" based on contact analysis has superseded the previous identification of "pockets" in the MH groove (*see* **Note 13**).

3.2 Access to IMGT pMH Contact Sites

1. In the IMGT® Home page at http://www.imgt.org, click the link "IMGT/3Dstructure-DB and IMGT/2Dstructure-DB" to access the IMGT/3Dstructure-DB Welcome page [12–14].
2. In "IMGT complex type," select "pMH1" or "pMH2" (*see* **Note 14**), or "TR/pMH1" or "TR/pMH2" (*see* **Note 15**), to retrieve the corresponding IMGT/3Dstructure-DB entries.
3. Click on the "IMGT entry ID" to access an individual IMGT/3Dstructure-DB card [13, 14].
4. Click the "Contact analysis" section (in "Chain details" of the MH chain(s)) to access the "IMGT pMH contact sites."

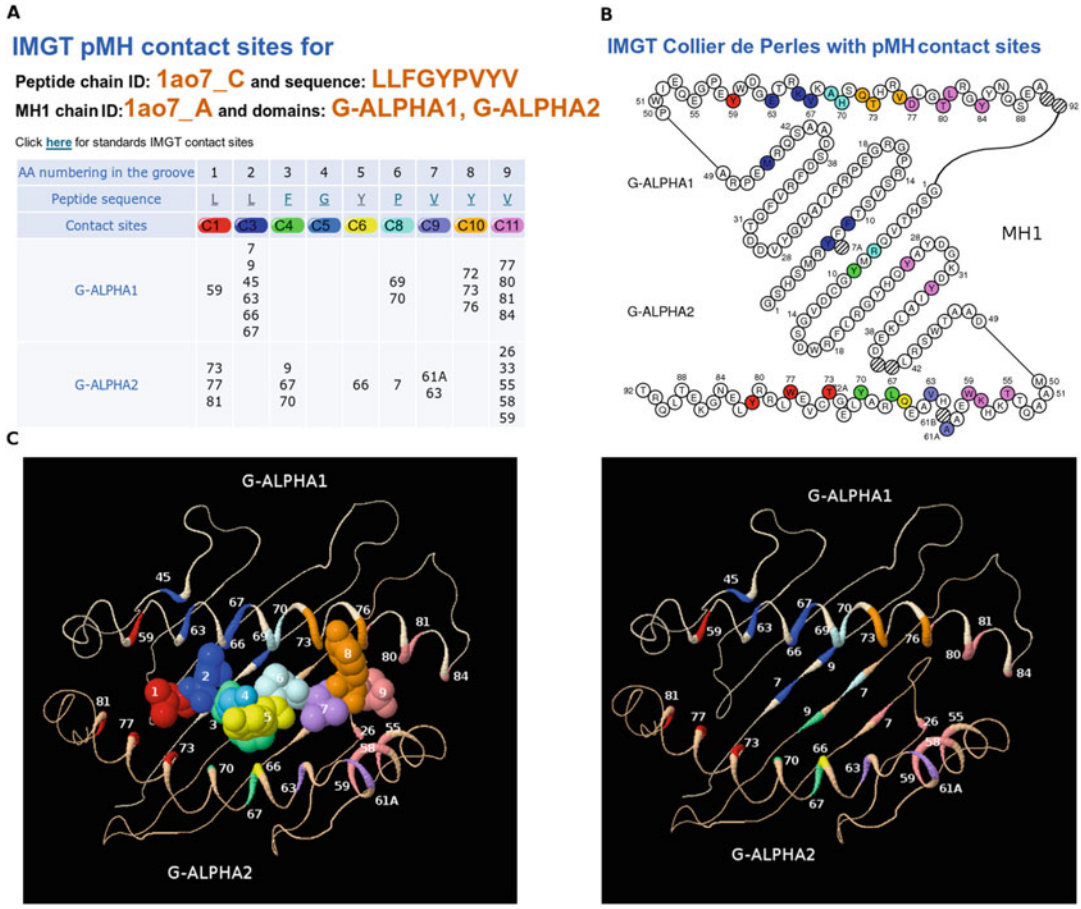


Fig. 7 "IMGT pMH contact sites" between MH1 and a 9-AA peptide. (**a**) "IMGT pMH contact sites" for MH1 (human HLA-A*0201, 1ao7_A) and peptide 1ao7_C. The numbers 1–9 refer to the peptide AA numbering (LLFGYPVYV). C1–C11 refer to the "IMGT pMH contact sites" (there are no C2 and C7 in agreement with MH1 binding a 9-AA peptide). In that 3-D structure, there is no C5 because the glycine G4 score is too low. The G-ALPHA1 and G-ALPHA2 AA positions assigned automatically to the "IMGT pMH contact sites" are listed. (**b**) "IMGT/Collier-de-Perles with pMH contact sites." View is from above the cleft, with G-ALPHA1 on top and G-ALPHA2 on bottom. (**c**) Groove 3-D structure. The groove is shown with and without the peptide (on the left and right hand side, respectively). The IMGT Color menu for "IMGT pMH contact sites" is used in (**a**), (**b**), and (**c**). (**a**) and (**b**) are from IMGT/3Dstructure-DB, http:/www.imgt.org. (With permission from M-P. Lefranc and G. Lefranc, LIGM, Founders and Authors of IMGT®, the international ImMunoGeneTics information system®, http:/www.imgt.org)

3.2.1 IMGT pMH Contact Sites for pMH1

An example of "IMGT pMH contact sites" for pMH1 is shown in Fig. 7. In that 3-D structure of a TR/pMH1 complex (1ao7), the groove made by the G-ALPHA1 and G-ALPHA2 of the I-ALPHA chain (1ao7_A) binds a 9-AA peptide (1ao7_C). "IMGT pMH contact sites" results provide first a table, which shows the positions 1–9 of the peptide AA (each AA is clickable, giving access to its Residue@Position card). Nine of the 11 C1–C11 contact sites are displayed, C2 and C7 being absent, in agreement with a 9-AA

peptide bound in a MH1 groove (*see* Subheading 3.1). The G-ALPHA1 and G-ALPHA2 AA positions that contribute to each "IMGT pMH contact site" are listed. For example, G-ALPHA1 59 and G-ALPHA2 73, 77, and 81 contribute to the "IMGT pMH contact site" C1 that predominantly interacts with leucine (L) 1 of the peptide (N-terminal end) (Fig. 7a). The "IMGT pMH contact sites" are displayed in "IMGT/Collier-de-Perles with pMH contact sites" (Fig. 7b). Clicking on one residue in the IMGT/Collier-de-Perles gives access to its "IMGT Residue@Position card" (*see* Subheading 2.4). The 3-D structure, with or without peptide, is shown in Fig. 7c.

3.2.2 IMGT pMH Contact Sites for pMH2

An example of "IMGT pMH contact sites" for pMH2 is shown in Fig. 8. In that 3-D structure of a TR/pMH2 complex (1j8h), the groove made by the G-ALPHA (of the II-ALPHA chain) (1j8h_A) and the G-BETA (of the II-BETA chain) binds a 13-AA peptide (1j8h_C). "IMGT pMH contact sites" results provide first a table, which shows the AA 1–9 in the groove (each AA is clickable, giving access to its Residue@Position card). However, in contrast to MH1 (*see* Subheading 3.2.1), the nine AA shown in Fig. 8a only correspond to the central part of the peptide. Indeed, the peptide bound to MH2 is longer than the length of the groove and extends outside its N-terminal and C-terminal ends, as the MH2 groove is "open" at both ends [7]. One major breakthrough of the "IMGT pMH contact sites" is the identification of the AA that is located in the MH2 groove [15–17]. Whereas the peptide (1j8h_C) is 13 AA long (PKYVKQNTLKLAT), the "IMGT pMH contact sites" results allow one to determine that the 9 AA in the MH2 groove are YVKQNTLKL (Fig. 8a). Nine of the 11 C1–C11 contact sites are displayed, C7 and C8 being absent, in agreement with 9 AA inside a MH2 groove (*see* Subheading 3.1). The G-ALPHA and G-BETA AA positions that contribute to each 'IMGT pMH contact sites' are listed. They are visualized in the "IMGT Collier de Perles with pMH contact sites" (Fig. 8b). Clicking on one residue in the IMGT Colliers de Perles gives access to its "IMGT Residue@Position card" (*see* Subheading 2.4). The 3-D structure, with or without peptide, is shown in Fig. 8c.

4 IMGT/3Dstructure-DB Domain Pair Contacts

4.1 IMGT/3DStructure-DB Domain Pair Contacts (Overview)

"IMGT/3Dstructure-DB Domain pair contacts (overview)" (Fig. 9) is accessed by clicking on "Domain contacts (overview)" of "Contact analysis" in an IMGT/3Dstructure-DB card. The example shown in Fig. 9 is that of the TR/pMH1 structure 1ao7. Eight "Domain pair contacts" are of interest for TR/pMH interactions, two for pMH1 (*see* Subheading 4.1.1) and six for TR/pMH1 (*see* Subheading 4.1.2). Similar results are obtained for the

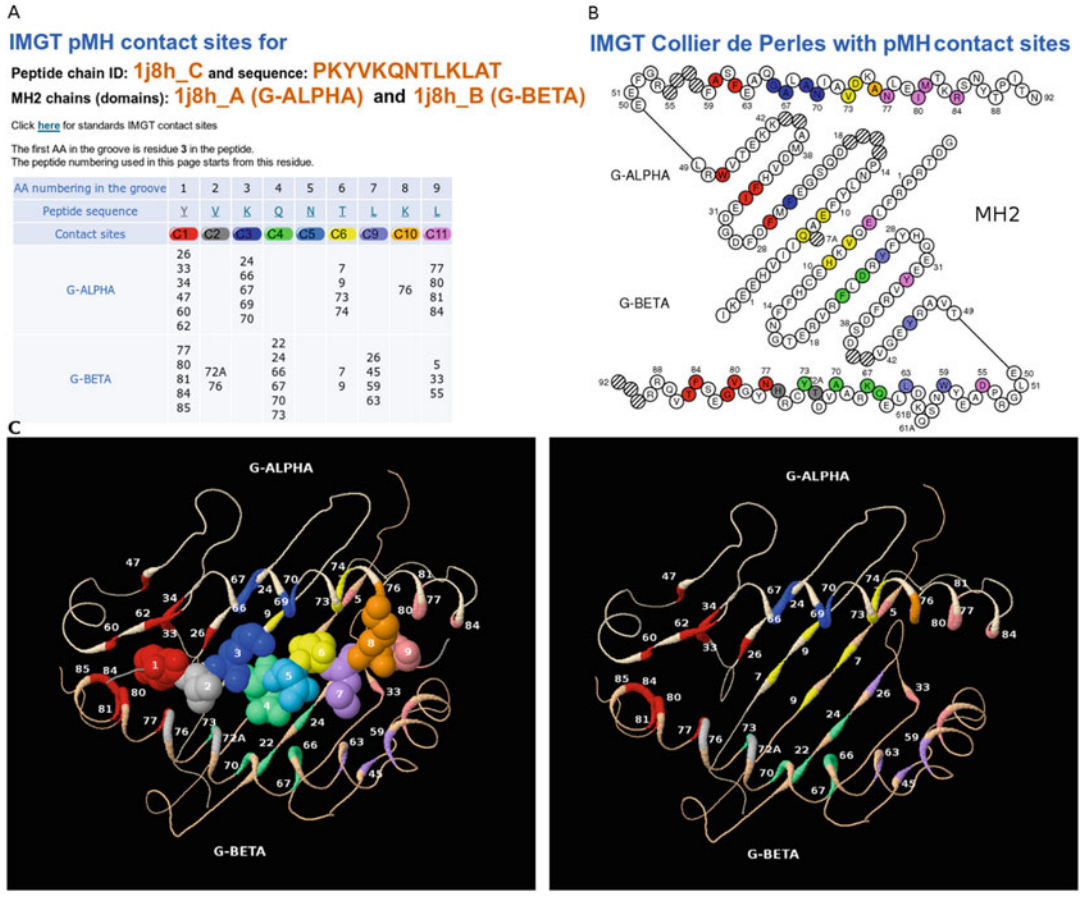


Fig. 8 "IMGT pMH contact sites" between MH2 and 9 AA in the groove. (**a**) "IMGT pMH contact sites" for MH2 (human HLA-DRA*0101_HLA-DRB1*0401) (1j8h_A-1j8h_B) and a 13 AA long peptide (1j8_C). The numbers 1–9 refer to the AA numbering in the groove (YVKQNTLKL) as determined by the "IMGT pMH contact sites." C1–C11 refer to the "IMGT pMH contact sites" (there are no C7 and C8 in agreement with MH2 binding 9 AA in the groove). In that 3-D structure, there is no C5 because the asparagine N5 score is too low. The G-ALPHA and G-BETA AA positions assigned automatically to the "IMGT pMH contact sites" are listed. (**b**) "IMGT/Collier-de-Perles with pMH contact sites." View is from above the cleft, with G-ALPHA on top and G-BETA on bottom. (**c**) Groove 3-D structure. The groove is shown with and without the peptide (on the left and right hand side, respectively). The IMGT Color menu for "IMGT pMH contact sites" is used in (**a**), (**b**), and (**c**). (**a**) and (**b**) are from IMGT/3Dstructure-DB, http:/www.imgt.org. (With permission from M-P. Lefranc and G. Lefranc, LIGM, Founders and Authors of IMGT®, the international ImMunoGeneTics information system®, http:/www.imgt.org)

TR/pMH2 structure, e.g., 1j8h (the only difference being the names of the G-DOMAIN, G-ALPHA, and G-BETA, instead of G-ALPHA1 and G-ALPHA2) (not further detailed here).

4.1.1 Domain Pair Contacts for pMH1 Interactions

The two domain pair contacts for pMH1 interactions are "(Ligand)/G-ALPHA1" and "(Ligand)/G-ALPHA2" (Fig. 9). Thus, for the pMH1 interactions in 1ao7, the domain pair "(Ligand)/G-ALPHA1" shows that 26 residues are involved,

IMGT/3Dstructure-DB Domain pair contacts (overview) of 1ao7

	Unit 1 Domain	Unit 1 Chain	Unit 2 Domain	Unit 2 Chain	Residue pair contacts	Number of residues Total	From 1	From 2	Atom pair contact types Total	Noncovalent	Polar	Hydrogen	Covalent	Disulfide
DomPair	V-ALPHA	1ao7_D	G-ALPHA1	1ao7_A	15	16	9	7	126	126	22	3	0	0
DomPair			G-ALPHA2	1ao7_A	12	15	7	8	105	105	17	2	0	0
DomPair			(Ligand)	1ao7_C	15	13	7	6	109	109	20	3	0	0
DomPair			C-ALPHA	1ao7_D	1	2	1	1	7	7	1	0	0	0
DomPair			V-BETA	1ao7_E	57	42	20	22	401	401	46	7	0	0
DomPair			C-BETA-2	1ao7_E	1	2	1	1	9	9	2	0	0	0
DomPair	C-ALPHA	1ao7_D	V-ALPHA	1ao7_D	1	2	1	1	7	7	1	0	0	0
DomPair	V-BETA	1ao7_E	G-ALPHA1	1ao7_A	3	4	1	3	23	23	0	0	0	0
DomPair			G-ALPHA2	1ao7_A	11	10	5	5	82	82	17	3	0	0
DomPair			(Ligand)	1ao7_C	14	13	9	4	119	119	9	2	0	0
DomPair			V-ALPHA	1ao7_D	57	42	22	20	401	401	46	7	0	0
DomPair			C-BETA-2	1ao7_E	32	27	12	15	236	236	30	1	0	0
DomPair	C-BETA-2	1ao7_E	V-ALPHA	1ao7_D	1	2	1	1	9	9	2	0	0	0
DomPair			V-BETA	1ao7_E	32	27	15	12	236	236	30	1	0	0
DomPair	G-ALPHA1	1ao7_A	G-ALPHA2	1ao7_A	119	77	36	41	961	961	137	22	0	0
DomPair			C-LIKE	1ao7_A	7	8	3	5	62	62	9	2	0	0
DomPair			C-LIKE	1ao7_B	18	18	11	7	153	153	18	6	0	0
DomPair			(Ligand)	1ao7_C	29	26	18	8	305	305	31	5	0	0
DomPair			V-ALPHA	1ao7_D	15	16	7	9	126	126	22	3	0	0
DomPair			V-BETA	1ao7_E	3	4	3	1	23	23	0	0	0	0
DomPair	G-ALPHA2	1ao7_A	G-ALPHA1	1ao7_A	119	77	41	36	961	961	137	22	0	0
DomPair			C-LIKE	1ao7_A	13	13	4	9	98	98	19	1	0	0
DomPair			C-LIKE	1ao7_B	25	20	11	9	246	246	20	3	0	0
DomPair			(Ligand)	1ao7_C	26	24	16	8	281	281	20	5	0	0
DomPair			V-ALPHA	1ao7_D	12	15	8	7	105	105	17	2	0	0
DomPair			V-BETA	1ao7_E	11	10	5	5	82	82	17	3	0	0
DomPair	C-LIKE	1ao7_A	G-ALPHA1	1ao7_A	7	8	5	3	62	62	9	2	0	0
DomPair			G-ALPHA2	1ao7_A	13	13	9	4	98	98	19	1	0	0
DomPair			C-LIKE	1ao7_B	31	25	12	13	310	310	43	8	0	0
DomPair	C-LIKE	1ao7_B		1ao7_1	6	7	6	1	64	64	0	0	0	0
DomPair				1ao7_2	2	3	2	1	6	6	0	0	0	0
DomPair			G-ALPHA1	1ao7_A	18	18	7	11	153	153	18	6	0	0
DomPair			G-ALPHA2	1ao7_A	25	20	9	11	246	246	20	3	0	0
DomPair			C-LIKE	1ao7_A	31	25	13	12	310	310	43	8	0	0

Fig. 9 IMGT/3Dstructure-DB Domain pair contacts (overview). The IMGT/3Dstructure-DB entry is the TR/pMH1 3-D structure 1ao7. The domain partners considered are designated as "Unit 1" and "Unit 2." The number of residue pair contacts, the number of residues involved (total, from Unit 1 and from Unit 2), the number of total atom pair contacts, and, as selected by the user for the display, the number of contacts per type and/or by category are provided. "(Ligand)" refers to the peptide. Two red frames highlight the domain pair contacts for pMH interactions. Two blue rectangles highlight the domain pair contacts for TR/pMH interactions, three for V-ALPHA and three for V-BETA. (With permission from M-P. Lefranc and G. Lefranc, LIGM, Founders and Authors of IMGT®, the international ImMunoGeneTics information system®, http:/www.imgt.org)

8 AA of the peptide (Ligand) interacting with 18 AA of G-ALPHA1 (creating 29 residue pair contacts with a total of 305 atom pair contacts). Similarly, the domain pair "(Ligand)/G-ALPHA2" shows that 24 residues are involved, 8 AA of the peptide interacting with 16 AA of G-ALPHA2 (creating 26 residue pair contacts with a total of 281 atom pair contacts).

4.1.2 Domain Pair Contacts for TR/pMH1 Interactions

The six domain pair contacts for TR/pMH1 interactions include three domain pairs involving V-ALPHA and three domain pairs involving V-BETA (Fig. 9). The TR/pMH1 interactions in 1ao7 are the following:

1. "V-ALPHA/G-ALPHA1": 9 AA of V-ALPHA interact with 7 AA of G-ALPHA1 (15 residue pair contacts with a total of 126 atom pair contacts).
2. "V-ALPHA/G-ALPHA2": 7 AA of V-ALPHA interact with 8 AA of G-ALPHA2 (12 residue pair contacts with a total of 105 atom pair contacts).
3. "V-ALPHA/(Ligand)": 7 AA of V-ALPHA interact with 6 AA of the peptide (15 residue pair contacts with a total of 109 atom pair contacts).
4. "V-BETA/G-ALPHA1": 1 AA of V-BETA interacts with 3 AA of G-ALPHA1 (3 residue pair contacts with a total of 23 atom pair contacts).
5. "V-BETA/G-ALPHA2": 5 AA of V-BETA interact with 5 AA of G-ALPHA2 (11 residue pair contacts with a total of 82 atom pair contacts).
6. "V-BETA/(Ligand)": 9 AA of V-BETA interact with 4 AA of the peptide (14 residue pair contacts with a total of 119 atom pair contacts).

4.2 IMGT/3Dstructure-DB Domain Pair Contacts (Per Pair)

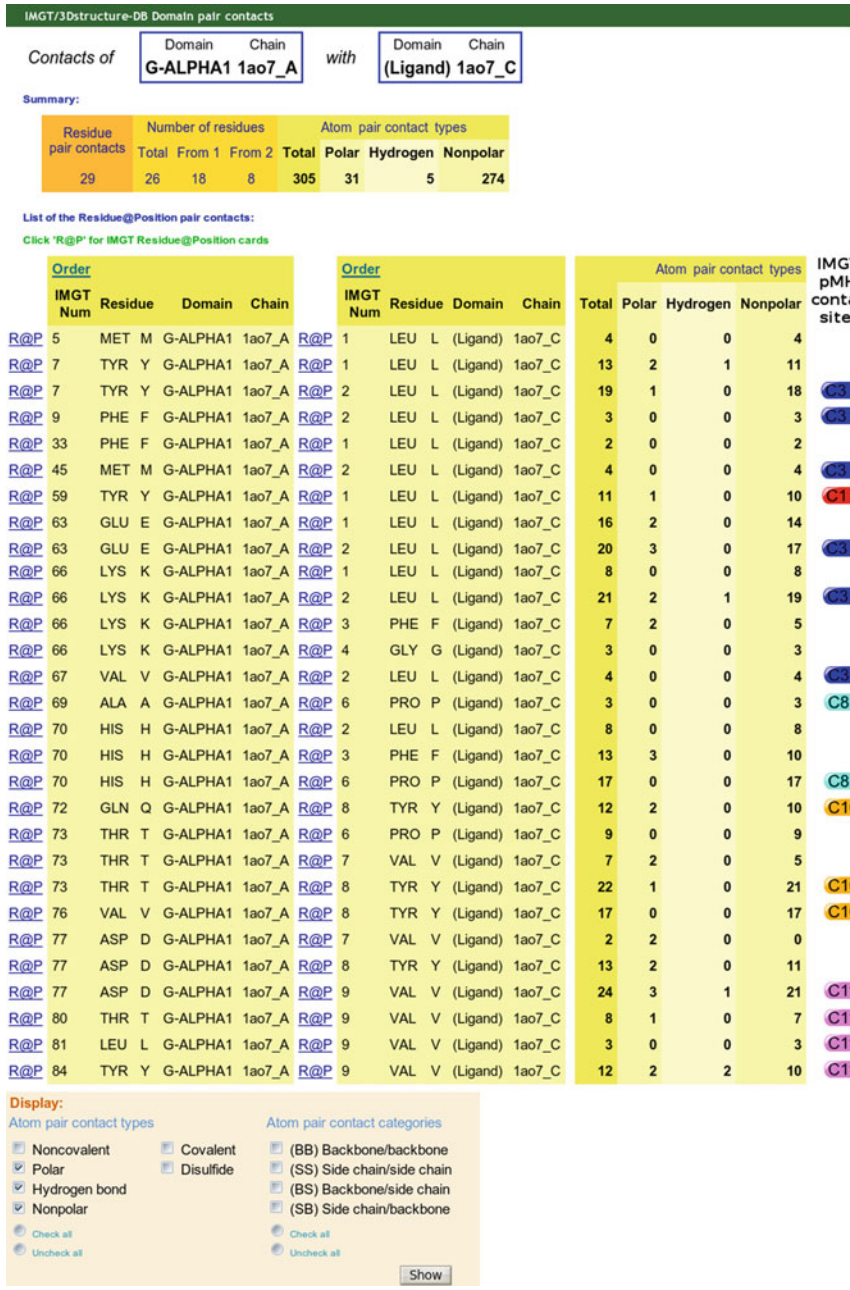
The corresponding detailed description of the "Domain pair contacts" (per pair) that characterize the interactions pMH and TR/pMH is accessed by clicking on "DomPair" (Fig. 9) [12–17].

4.2.1 pMH1 Interactions

The pMH1 interactions "G-ALPHA1/(Ligand)" (Fig. 10a) and "G-ALPHA2/(Ligand)" (Fig. 10b) provide the details of the residue pair contacts with the number of atom pair contact types (total, polar, hydrogen, and nonpolar) (identical results are obtained in the reciprocal queries "(Ligand)/G-ALPHA1" and "(Ligand)/G-ALPHA2"). In these tables, all the peptide—MH1 AA interactions—are listed, in contrast to the "IMGT pMH contact sites," that only visualize those with the highest score (Fig. 7) (*see* Subheading 3).

4.2.2 TR/pMH1 Interactions

The interactions "V-ALPHA/G-ALPHA1," "V-ALPHA/G-ALPHA2," and "V-ALPHA/(Ligand)" are shown in Fig. 11 ((A), (B), and (C), respectively). The interactions "V-BETA/G-ALPHA1," "V-BETA/G-ALPHA2," and "V-BETA/(Ligand)" are shown in Fig. 12((A), (B) and (C), respectively). Positions that belong to the CDR-IMGT are highlighted according to the IMGT color menu: blue (CDR1-IMGT), green (CDR2-IMGT), and greenblue (CDR3-IMGT) for V-ALPHA (Fig. 11) and red (CDR1-IMGT) and purple (CDR3-IMGT) for V-BETA (Fig. 12) (there is no contact with the CDR2-IMGT in 1ao7). Positions can be localized in the IMGT Colliers de Perles (Fig. 3, for V-ALPHA and B-BETA, and Fig. 4 for G-ALPHA1 and G-ALPHA2).



IMGT/3Dstructure-DB Domain pair contacts

Contacts of — Domain: G-ALPHA1, Chain: 1ao7_A — with — Domain: (Ligand), Chain: 1ao7_C

Summary:

Residue pair contacts	Number of residues			Atom pair contact types			
	Total	From 1	From 2	Total	Polar	Hydrogen	Nonpolar
29	26	18	8	305	31	5	274

List of the Residue@Position pair contacts:

Click 'R@P' for IMGT Residue@Position cards

	Order					Order				Atom pair contact types				IMGT pMH contact sites
	IMGT Num	Residue	Domain	Chain		IMGT Num	Residue	Domain	Chain	Total	Polar	Hydrogen	Nonpolar	
R@P	5	MET M	G-ALPHA1	1ao7_A	R@P	1	LEU L	(Ligand)	1ao7_C	4	0	0	4	
R@P	7	TYR Y	G-ALPHA1	1ao7_A	R@P	1	LEU L	(Ligand)	1ao7_C	13	2	1	11	
R@P	7	TYR Y	G-ALPHA1	1ao7_A	R@P	2	LEU L	(Ligand)	1ao7_C	19	1	0	18	C3
R@P	9	PHE F	G-ALPHA1	1ao7_A	R@P	2	LEU L	(Ligand)	1ao7_C	3	0	0	3	C3
R@P	33	PHE F	G-ALPHA1	1ao7_A	R@P	1	LEU L	(Ligand)	1ao7_C	2	0	0	2	
R@P	45	MET M	G-ALPHA1	1ao7_A	R@P	2	LEU L	(Ligand)	1ao7_C	4	0	0	4	C3
R@P	59	TYR Y	G-ALPHA1	1ao7_A	R@P	1	LEU L	(Ligand)	1ao7_C	11	1	0	10	C1
R@P	63	GLU E	G-ALPHA1	1ao7_A	R@P	1	LEU L	(Ligand)	1ao7_C	16	2	0	14	
R@P	63	GLU E	G-ALPHA1	1ao7_A	R@P	2	LEU L	(Ligand)	1ao7_C	20	3	0	17	C3
R@P	66	LYS K	G-ALPHA1	1ao7_A	R@P	1	LEU L	(Ligand)	1ao7_C	8	0	0	8	
R@P	66	LYS K	G-ALPHA1	1ao7_A	R@P	2	LEU L	(Ligand)	1ao7_C	21	2	1	19	C3
R@P	66	LYS K	G-ALPHA1	1ao7_A	R@P	3	PHE F	(Ligand)	1ao7_C	7	2	0	5	
R@P	66	LYS K	G-ALPHA1	1ao7_A	R@P	4	GLY G	(Ligand)	1ao7_C	3	0	0	3	
R@P	67	VAL V	G-ALPHA1	1ao7_A	R@P	2	LEU L	(Ligand)	1ao7_C	4	0	0	4	C3
R@P	69	ALA A	G-ALPHA1	1ao7_A	R@P	6	PRO P	(Ligand)	1ao7_C	3	0	0	3	C8
R@P	70	HIS H	G-ALPHA1	1ao7_A	R@P	2	LEU L	(Ligand)	1ao7_C	8	0	0	8	
R@P	70	HIS H	G-ALPHA1	1ao7_A	R@P	3	PHE F	(Ligand)	1ao7_C	13	3	0	10	
R@P	70	HIS H	G-ALPHA1	1ao7_A	R@P	6	PRO P	(Ligand)	1ao7_C	17	0	0	17	C8
R@P	72	GLN Q	G-ALPHA1	1ao7_A	R@P	8	TYR Y	(Ligand)	1ao7_C	12	2	0	10	C10
R@P	73	THR T	G-ALPHA1	1ao7_A	R@P	6	PRO P	(Ligand)	1ao7_C	9	0	0	9	
R@P	73	THR T	G-ALPHA1	1ao7_A	R@P	7	VAL V	(Ligand)	1ao7_C	7	2	0	5	
R@P	73	THR T	G-ALPHA1	1ao7_A	R@P	8	TYR Y	(Ligand)	1ao7_C	22	1	0	21	C10
R@P	76	VAL V	G-ALPHA1	1ao7_A	R@P	8	TYR Y	(Ligand)	1ao7_C	17	0	0	17	C10
R@P	77	ASP D	G-ALPHA1	1ao7_A	R@P	7	VAL V	(Ligand)	1ao7_C	2	2	0	0	
R@P	77	ASP D	G-ALPHA1	1ao7_A	R@P	8	TYR Y	(Ligand)	1ao7_C	13	2	0	11	
R@P	77	ASP D	G-ALPHA1	1ao7_A	R@P	9	VAL V	(Ligand)	1ao7_C	24	3	1	21	C11
R@P	80	THR T	G-ALPHA1	1ao7_A	R@P	9	VAL V	(Ligand)	1ao7_C	8	1	0	7	C11
R@P	81	LEU L	G-ALPHA1	1ao7_A	R@P	9	VAL V	(Ligand)	1ao7_C	3	0	0	3	C11
R@P	84	TYR Y	G-ALPHA1	1ao7_A	R@P	9	VAL V	(Ligand)	1ao7_C	12	2	2	10	C11

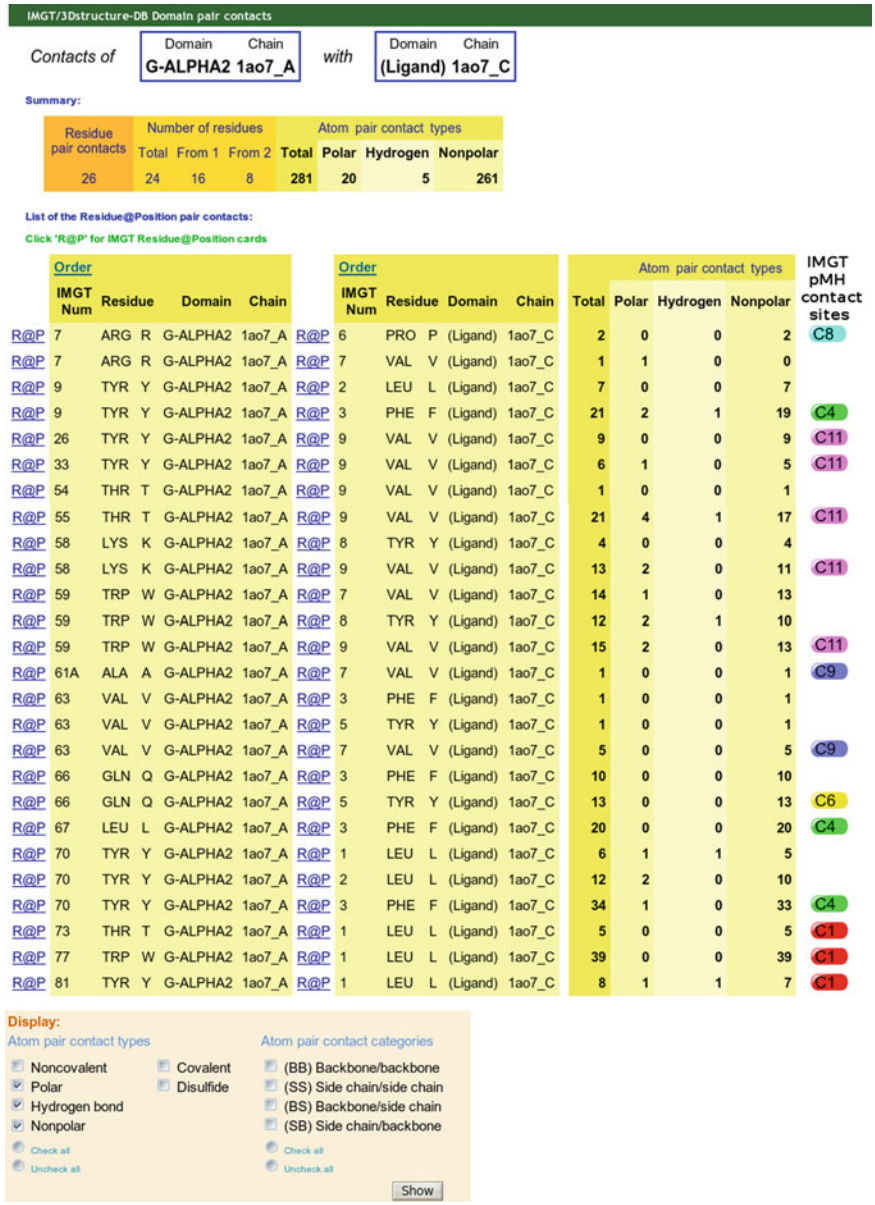
Display:

Atom pair contact types: Noncovalent, Polar, Hydrogen bond, Nonpolar; Covalent, Disulfide; Check all, Uncheck all

Atom pair contact categories: (BB) Backbone/backbone, (SS) Side chain/side chain, (BS) Backbone/side chain, (SB) Side chain/backbone; Check all, Uncheck all

Show

Fig. 10 pMH1 interactions. (**a**) Interactions "G-ALPHA1/(Ligand)" of 1ao7. (**b**) Interactions "G-ALPHA2/(Ligand)" of 1ao7. Clicking on a R@P link gives access to the corresponding IMGT Residue@Position card. "(Ligand)" refers to the peptide. The contact analysis of the TR/pMH 3-D structure 1ao7 is from IMGT/3Dstructure-DB, http:/www.imgt.org. The "IMGT pMH contact sites" for G-ALPHA1 (**a**) and G-ALPHA2 (**b**) were added on the right hand side of the figure, for a comparison with Fig. 7. (With permission from M-P. Lefranc and G. Lefranc, LIGM, Founders and Authors of IMGT®, the international ImMunoGeneTics information system®, http:/www.imgt.org)



IMGT/3Dstructure-DB Domain pair contacts

Contacts of — Domain: G-ALPHA2, Chain: 1ao7_A — with — Domain: (Ligand), Chain: 1ao7_C

Summary:

Residue pair contacts	Number of residues			Atom pair contact types			
	Total	From 1	From 2	Total	Polar	Hydrogen	Nonpolar
26	24	16	8	281	20	5	261

List of the Residue@Position pair contacts:

Click 'R@P' for IMGT Residue@Position cards

	Order						Order					Atom pair contact types				IMGT pMH contact sites
	IMGT Num	Residue		Domain	Chain		IMGT Num	Residue		Domain	Chain	Total	Polar	Hydrogen	Nonpolar	
R@P	7	ARG	R	G-ALPHA2	1ao7_A	R@P	6	PRO	P	(Ligand)	1ao7_C	2	0	0	2	C8
R@P	7	ARG	R	G-ALPHA2	1ao7_A	R@P	7	VAL	V	(Ligand)	1ao7_C	1	1	0	0	
R@P	9	TYR	Y	G-ALPHA2	1ao7_A	R@P	2	LEU	L	(Ligand)	1ao7_C	7	0	0	7	
R@P	9	TYR	Y	G-ALPHA2	1ao7_A	R@P	3	PHE	F	(Ligand)	1ao7_C	21	2	1	19	C4
R@P	26	TYR	Y	G-ALPHA2	1ao7_A	R@P	9	VAL	V	(Ligand)	1ao7_C	9	0	0	9	C11
R@P	33	TYR	Y	G-ALPHA2	1ao7_A	R@P	9	VAL	V	(Ligand)	1ao7_C	6	1	0	5	C11
R@P	54	THR	T	G-ALPHA2	1ao7_A	R@P	9	VAL	V	(Ligand)	1ao7_C	1	0	0	1	
R@P	55	THR	T	G-ALPHA2	1ao7_A	R@P	9	VAL	V	(Ligand)	1ao7_C	21	4	1	17	C11
R@P	58	LYS	K	G-ALPHA2	1ao7_A	R@P	8	TYR	Y	(Ligand)	1ao7_C	4	0	0	4	
R@P	58	LYS	K	G-ALPHA2	1ao7_A	R@P	9	VAL	V	(Ligand)	1ao7_C	13	2	0	11	C11
R@P	59	TRP	W	G-ALPHA2	1ao7_A	R@P	7	VAL	V	(Ligand)	1ao7_C	14	1	0	13	
R@P	59	TRP	W	G-ALPHA2	1ao7_A	R@P	8	TYR	Y	(Ligand)	1ao7_C	12	2	1	10	
R@P	59	TRP	W	G-ALPHA2	1ao7_A	R@P	9	VAL	V	(Ligand)	1ao7_C	15	2	0	13	C11
R@P	61A	ALA	A	G-ALPHA2	1ao7_A	R@P	7	VAL	V	(Ligand)	1ao7_C	1	0	0	1	C9
R@P	63	VAL	V	G-ALPHA2	1ao7_A	R@P	3	PHE	F	(Ligand)	1ao7_C	1	0	0	1	
R@P	63	VAL	V	G-ALPHA2	1ao7_A	R@P	5	TYR	Y	(Ligand)	1ao7_C	1	0	0	1	
R@P	63	VAL	V	G-ALPHA2	1ao7_A	R@P	7	VAL	V	(Ligand)	1ao7_C	5	0	0	5	C9
R@P	66	GLN	Q	G-ALPHA2	1ao7_A	R@P	3	PHE	F	(Ligand)	1ao7_C	10	0	0	10	
R@P	66	GLN	Q	G-ALPHA2	1ao7_A	R@P	5	TYR	Y	(Ligand)	1ao7_C	13	0	0	13	C6
R@P	67	LEU	L	G-ALPHA2	1ao7_A	R@P	3	PHE	F	(Ligand)	1ao7_C	20	0	0	20	C4
R@P	70	TYR	Y	G-ALPHA2	1ao7_A	R@P	1	LEU	L	(Ligand)	1ao7_C	6	1	1	5	
R@P	70	TYR	Y	G-ALPHA2	1ao7_A	R@P	2	LEU	L	(Ligand)	1ao7_C	12	2	0	10	
R@P	70	TYR	Y	G-ALPHA2	1ao7_A	R@P	3	PHE	F	(Ligand)	1ao7_C	34	1	0	33	C4
R@P	73	THR	T	G-ALPHA2	1ao7_A	R@P	1	LEU	L	(Ligand)	1ao7_C	5	0	0	5	C1
R@P	77	TRP	W	G-ALPHA2	1ao7_A	R@P	1	LEU	L	(Ligand)	1ao7_C	39	0	0	39	C1
R@P	81	TYR	Y	G-ALPHA2	1ao7_A	R@P	1	LEU	L	(Ligand)	1ao7_C	8	1	1	7	C1

Display:

Atom pair contact types: Noncovalent, Polar, Hydrogen bond, Nonpolar; Covalent, Disulfide; Check all, Uncheck all

Atom pair contact categories: (BB) Backbone/backbone, (SS) Side chain/side chain, (BS) Backbone/side chain, (SB) Side chain/backbone; Check all, Uncheck all

Show

Fig. 10 (continued)

4.3 IMGT Paratope and Epitope

IMGT paratope and epitope are concepts of the "SpecificityType" in IMGT-ONTOLOGY [48–50]. *Paratope*, or "antigen-binding site," identifies the part of the V-DOMAIN of an IG or antibody ("IG paratope") or of a TR ("TR paratope") that, respectively, recognizes (binds to) the antigen (Ag) or the peptide/major histocompatibility (pMH) ("epitope" or "antigenic determinant") [49]. *Epitope*, or "antigenic determinant," identifies the part of the antigen (Ag) or of the peptide/major histocompatibility

a

IMGT/3Dstructure-DB Domain pair contacts

Contacts of | Domain: V-ALPHA [D1] | Chain: 1ao7_D | with | Domain: G-ALPHA1 [D1] | Chain: 1ao7_A

Summary:

Residue pair contacts	Number of residues			Atom pair contact types			
	Total	From 1	From 2	Total	Polar	Hydrogen	Nonpolar
15	16	9	7	126	22	3	104

List of the Residue@Position pair contacts:

Click 'R@P' for IMGT Residue@Position cards

Order						Order						Atom pair contact types			
	IMGT Num	Residue		Domain	Chain		IMGT Num	Residue		Domain	Chain	Total	Polar	Hydrogen	Nonpolar
R@P	2	LYS	K	V-ALPHA [D1]	1ao7_D	R@P	58	GLU	E	G-ALPHA1 [D1]	1ao7_A	7	1	0	6
R@P	26	SER	S	V-ALPHA [D1]	1ao7_D	R@P	58	GLU	E	G-ALPHA1 [D1]	1ao7_A	3	2	0	1
R@P	27	ASP	D	V-ALPHA [D1]	1ao7_D	R@P	58	GLU	E	G-ALPHA1 [D1]	1ao7_A	24	6	1	18
R@P	28	ARG	R	V-ALPHA [D1]	1ao7_D	R@P	58	GLU	E	G-ALPHA1 [D1]	1ao7_A	1	1	0	0
R@P	37	GLN	Q	V-ALPHA [D1]	1ao7_D	R@P	66	LYS	K	G-ALPHA1 [D1]	1ao7_A	4	1	0	3
R@P	108	THR	T	V-ALPHA [D1]	1ao7_D	R@P	65	ARG	R	G-ALPHA1 [D1]	1ao7_A	5	2	1	3
R@P	108	THR	T	V-ALPHA [D1]	1ao7_D	R@P	66	LYS	K	G-ALPHA1 [D1]	1ao7_A	1	0	0	1
R@P	109	ASP	D	V-ALPHA [D1]	1ao7_D	R@P	62	GLY	G	G-ALPHA1 [D1]	1ao7_A	1	1	0	0
R@P	109	ASP	D	V-ALPHA [D1]	1ao7_D	R@P	65	ARG	R	G-ALPHA1 [D1]	1ao7_A	19	5	1	14
R@P	109	ASP	D	V-ALPHA [D1]	1ao7_D	R@P	66	LYS	K	G-ALPHA1 [D1]	1ao7_A	14	1	0	13
R@P	113	TRP	W	V-ALPHA [D1]	1ao7_D	R@P	65	ARG	R	G-ALPHA1 [D1]	1ao7_A	12	1	0	11
R@P	113	TRP	W	V-ALPHA [D1]	1ao7_D	R@P	68	LYS	K	G-ALPHA1 [D1]	1ao7_A	8	0	0	8
R@P	113	TRP	W	V-ALPHA [D1]	1ao7_D	R@P	69	ALA	A	G-ALPHA1 [D1]	1ao7_A	16	0	0	16
R@P	113	TRP	W	V-ALPHA [D1]	1ao7_D	R@P	72	GLN	Q	G-ALPHA1 [D1]	1ao7_A	4	0	0	4
R@P	114	GLY	G	V-ALPHA [D1]	1ao7_D	R@P	65	ARG	R	G-ALPHA1 [D1]	1ao7_A	7	1	0	6

b

IMGT/3Dstructure-DB Domain pair contacts

Contacts of | Domain: V-ALPHA [D1] | Chain: 1ao7_D | with | Domain: G-ALPHA2 [D2] | Chain: 1ao7_A

Summary:

Residue pair contacts	Number of residues			Atom pair contact types			
	Total	From 1	From 2	Total	Polar	Hydrogen	Nonpolar
12	15	7	8	105	17	2	88

List of the Residue@Position pair contacts:

Click 'R@P' for IMGT Residue@Position cards

Order						Order						Atom pair contact types			
	IMGT Num	Residue		Domain	Chain		IMGT Num	Residue		Domain	Chain	Total	Polar	Hydrogen	Nonpolar
R@P	28	ARG	R	V-ALPHA [D1]	1ao7_D	R@P	77	TRP	W	G-ALPHA2 [D2]	1ao7_A	14	1	0	13
R@P	28	ARG	R	V-ALPHA [D1]	1ao7_D	R@P	80	ARG	R	G-ALPHA2 [D2]	1ao7_A	13	5	0	8
R@P	29	GLY	G	V-ALPHA [D1]	1ao7_D	R@P	77	TRP	W	G-ALPHA2 [D2]	1ao7_A	6	0	0	6
R@P	37	GLN	Q	V-ALPHA [D1]	1ao7_D	R@P	70	TYR	Y	G-ALPHA2 [D2]	1ao7_A	8	0	0	8
R@P	37	GLN	Q	V-ALPHA [D1]	1ao7_D	R@P	73	THR	T	G-ALPHA2 [D2]	1ao7_A	10	2	0	8
R@P	57	TYR	Y	V-ALPHA [D1]	1ao7_D	R@P	65	GLU	E	G-ALPHA2 [D2]	1ao7_A	4	0	0	4
R@P	57	TYR	Y	V-ALPHA [D1]	1ao7_D	R@P	66	GLN	Q	G-ALPHA2 [D2]	1ao7_A	16	1	0	15
R@P	57	TYR	Y	V-ALPHA [D1]	1ao7_D	R@P	69	ALA	A	G-ALPHA2 [D2]	1ao7_A	7	1	0	6
R@P	58	SER	S	V-ALPHA [D1]	1ao7_D	R@P	69	ALA	A	G-ALPHA2 [D2]	1ao7_A	3	1	0	2
R@P	63	ASN	N	V-ALPHA [D1]	1ao7_D	R@P	76	GLU	E	G-ALPHA2 [D2]	1ao7_A	4	2	0	2
R@P	82	LYS	K	V-ALPHA [D1]	1ao7_D	R@P	73	THR	T	G-ALPHA2 [D2]	1ao7_A	5	2	1	3
R@P	82	LYS	K	V-ALPHA [D1]	1ao7_D	R@P	76	GLU	E	G-ALPHA2 [D2]	1ao7_A	15	2	1	13

Fig. 11 TR V-ALPHA/pMH1 interactions. (**a**) Interactions "V-ALPHA/G-ALPHA1" of 1ao7. (**b**) Interactions "V-ALPHA/G-ALPHA2" of 1ao7. (**c**) Interactions "V-ALPHA/(Ligand)" of 1ao7. Clicking on a R@P link gives access to the corresponding IMGT Residue@Position card. "(Ligand)" refers to the peptide. The contact analysis of the TR/pMH 3-D structure 1ao7 is from IMGT/3Dstructure-DB, http:/www.imgt.org. The IMGT color menu is blue, green, and greenblue for CDR1-IMGT, CDR2-IMGT, and CDR3-IMGT, respectively (*see* **Note 6**). (With permission from M-P. Lefranc and G. Lefranc, LIGM, Founders and Authors of IMGT®, the international ImMunoGeneTics information system®, http:/www.imgt.org)

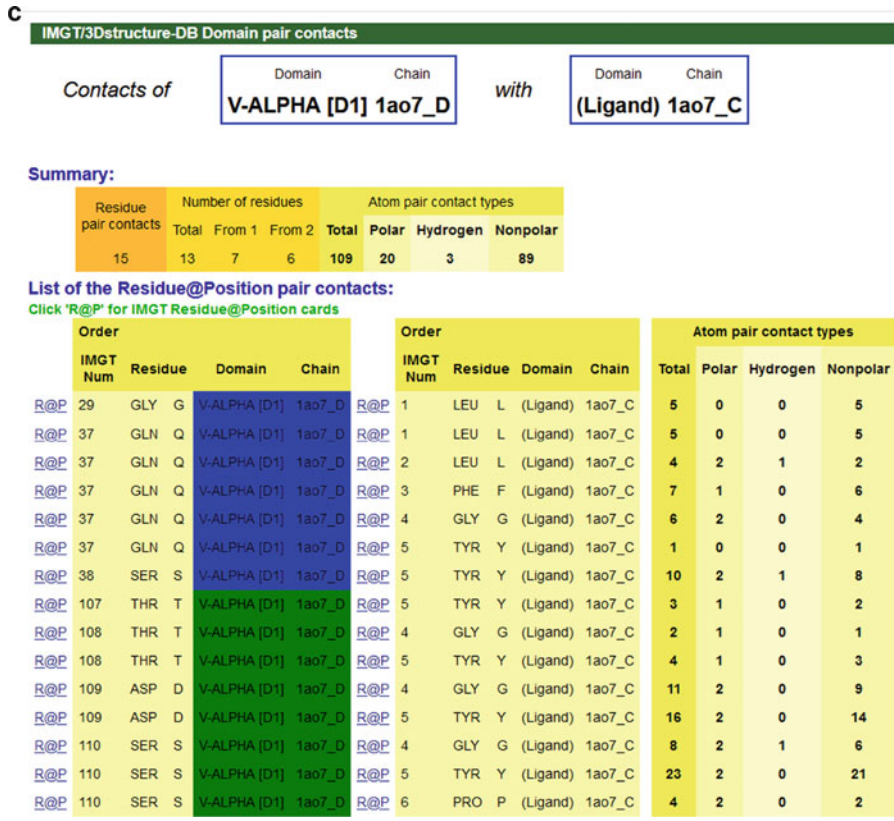


Fig. 11 (continued)

(pMH) that is recognized by the paratope of the V-DOMAIN of an IG or antibody or of a TR, respectively [50].

The amino acids that constitute the TR paratope belong to the paired V domains of a TR (V-alpha and V-beta for a TR-alpha_beta, V-gamma, and V-delta for a TR-gamma_delta), and more precisely to the CDR-IMGT [49]. Among the CDR-IMGT, the CDR3-IMGT that results from the V-J and V-D-J junction play the major role in TR/pMH interactions [15–17]. T-cell epitopes are usually identified as "linear" when referring to the processed peptide (p) presented in the groove of the MH proteins. However, in IMGT-ONTOLOGY, the "T-cell epitope" concept is identified as "discontinuous" as it comprises amino acids of the MH that bind to the TR V domains [50]. Thus, in a TR/pMH complex, the AA in contact at the interface between the TR and the pMH constitute the paratope on the TR surface and the epitope on the pMH surface (Fig. 13). In IMGT/3Dstructure-DB, the "IMGT paratope and epitope" for TR/pMH complexes are determined by combining contact analysis (Table 5) with an interaction scoring function, which roughly complies with the true mean energy ratio [15–17]. A standardized description of the "IMGT paratope and

a

IMGT/3Dstructure-DB Domain pair contacts

Contacts of | Domain: V-BETA [D1] | Chain: 1ao7_E | with | Domain: G-ALPHA1 [D1] | Chain: 1ao7_A

Lorem Ipsum

Summary:

Residue pair contacts	Number of residues			Atom pair contact types			
	Total	From 1	From 2	Total	Polar	Hydrogen	Nonpolar
3	4	1	3	23	0	0	23

List of the Residue@Position pair contacts:

Click 'R@P' for IMGT Residue@Position cards

Order						Order					Atom pair contact types			
	IMGT Num	Residue	Domain	Chain		IMGT Num	Residue	Domain	Chain	Total	Polar	Hydrogen	Nonpolar
R@P	110	LEU L	V-BETA [D1]	1ao7_E	R@P	69	ALA A	G-ALPHA1 [D1]	1ao7_A	9	0	0	9
R@P	110	LEU L	V-BETA [D1]	1ao7_E	R@P	72	GLN Q	G-ALPHA1 [D1]	1ao7_A	4	0	0	4
R@P	110	LEU L	V-BETA [D1]	1ao7_E	R@P	73	THR T	G-ALPHA1 [D1]	1ao7_A	10	0	0	10

Display:

Atom pair contact types: ☐ Noncovalent ☐ Covalent ☑ Polar ☑ Hydrogen bond ☑ Nonpolar ☐ Disulfide

Atom pair contact categories: ☐ (BB) Backbone/backbone ☐ (SS) Side chain/side chain ☐ (BS) Backbone/side chain ☐ (SB) Side chain/backbone

○ Check all ○ Uncheck all (for each group)

Show

b

IMGT/3Dstructure-DB Domain pair contacts

Contacts of | Domain: V-BETA [D1] | Chain: 1ao7_E | with | Domain: G-ALPHA2 [D2] | Chain: 1ao7_A

Summary:

Residue pair contacts	Number of residues			Atom pair contact types			
	Total	From 1	From 2	Total	Polar	Hydrogen	Nonpolar
11	10	5	5	82	17	3	65

List of the Residue@Position pair contacts:

Click 'R@P' for IMGT Residue@Position cards

Order						Order					Atom pair contact types			
	IMGT Num	Residue	Domain	Chain		IMGT Num	Residue	Domain	Chain	Total	Polar	Hydrogen	Nonpolar
R@P	111	ALA A	V-BETA [D1]	1ao7_E	R@P	61A	ALA A	G-ALPHA2 [D2]	1ao7_A	1	0	0	1
R@P	112.1	GLY G	V-BETA [D1]	1ao7_E	R@P	61A	ALA A	G-ALPHA2 [D2]	1ao7_A	5	0	0	5
R@P	112	GLY G	V-BETA [D1]	1ao7_E	R@P	61A	ALA A	G-ALPHA2 [D2]	1ao7_A	8	2	1	6
R@P	112	GLY G	V-BETA [D1]	1ao7_E	R@P	62	HIS H	G-ALPHA2 [D2]	1ao7_A	4	1	0	3
R@P	112	GLY G	V-BETA [D1]	1ao7_E	R@P	63	VAL V	G-ALPHA2 [D2]	1ao7_A	4	0	0	4
R@P	112	GLY G	V-BETA [D1]	1ao7_E	R@P	66	GLN Q	G-ALPHA2 [D2]	1ao7_A	10	2	1	8
R@P	113	ARG R	V-BETA [D1]	1ao7_E	R@P	61	ALA A	G-ALPHA2 [D2]	1ao7_A	5	2	1	3
R@P	113	ARG R	V-BETA [D1]	1ao7_E	R@P	61A	ALA A	G-ALPHA2 [D2]	1ao7_A	24	6	0	18
R@P	113	ARG R	V-BETA [D1]	1ao7_E	R@P	62	HIS H	G-ALPHA2 [D2]	1ao7_A	12	2	0	10
R@P	113	ARG R	V-BETA [D1]	1ao7_E	R@P	66	GLN Q	G-ALPHA2 [D2]	1ao7_A	2	1	0	1
R@P	114	PRO P	V-BETA [D1]	1ao7_E	R@P	66	GLN Q	G-ALPHA2 [D2]	1ao7_A	7	1	0	6

Fig. 12 TR V-BETA/pMH1 interactions. (**a**) Interactions "V-BETA/G-ALPHA1" of 1ao7. (**b**) Interactions "V-BETA/G-ALPHA2" of 1ao7. (**c**) Interactions "V-BETA/(Ligand)." Clicking on a R@P link gives access to the corresponding IMGT Residue@Position card. "(Ligand)" refers to the peptide. The contact analysis of the TR/pMH 3-D structure (1ao7) is from IMGT/3Dstructure-DB, http:/www.imgt.org. R@P belonging to CDR1-IMGT is in red and those belonging to the CDR3-IMGT are in purple according to the IMGT color menu (*see* **Note 6**). (With permission from M-P. Lefranc and G. Lefranc, LIGM, Founders and Authors of IMGT®, the international ImMunoGeneTics information system®, http:/www.imgt.org)

c

IMGT/3Dstructure-DB Domain pair contacts

Contacts of | Domain: V-BETA [D1] | Chain: 1ao7_E | with | Domain: (Ligand) | Chain: 1ao7_C

Summary:

Residue pair contacts	Number of residues			Atom pair contact types			
	Total	From 1	From 2	Total	Polar	Hydrogen	Nonpolar
14	13	9	4	119	9	2	110

List of the Residue@Position pair contacts:

Click 'R@P' for IMGT Residue@Position cards

	Order						Order					Atom pair contact types			
	IMGT Num	Residue		Domain	Chain		IMGT Num	Residue		Domain	Chain	Total	Polar	Hydrogen	Nonpolar
R@P	37	GLU	E	V-BETA [D1]	1ao7_E	R@P	8	TYR	Y	(Ligand)	1ao7_C	10	2	1	8
R@P	107	ARG	R	V-BETA [D1]	1ao7_E	R@P	5	TYR	Y	(Ligand)	1ao7_C	5	2	0	3
R@P	109	GLY	G	V-BETA [D1]	1ao7_E	R@P	6	PRO	P	(Ligand)	1ao7_C	1	1	0	0
R@P	110	LEU	L	V-BETA [D1]	1ao7_E	R@P	6	PRO	P	(Ligand)	1ao7_C	12	1	0	11
R@P	110	LEU	L	V-BETA [D1]	1ao7_E	R@P	7	VAL	V	(Ligand)	1ao7_C	10	1	0	9
R@P	110	LEU	L	V-BETA [D1]	1ao7_E	R@P	8	TYR	Y	(Ligand)	1ao7_C	32	1	1	31
R@P	111	ALA	A	V-BETA [D1]	1ao7_E	R@P	7	VAL	V	(Ligand)	1ao7_C	4	0	0	4
R@P	111	ALA	A	V-BETA [D1]	1ao7_E	R@P	8	TYR	Y	(Ligand)	1ao7_C	5	0	0	5
R@P	112.1	GLY	G	V-BETA [D1]	1ao7_E	R@P	5	TYR	Y	(Ligand)	1ao7_C	1	0	0	1
R@P	112.1	GLY	G	V-BETA [D1]	1ao7_E	R@P	7	VAL	V	(Ligand)	1ao7_C	8	0	0	8
R@P	112	GLY	G	V-BETA [D1]	1ao7_E	R@P	5	TYR	Y	(Ligand)	1ao7_C	9	0	0	9
R@P	112	GLY	G	V-BETA [D1]	1ao7_E	R@P	7	VAL	V	(Ligand)	1ao7_C	2	0	0	2
R@P	113	ARG	R	V-BETA [D1]	1ao7_E	R@P	5	TYR	Y	(Ligand)	1ao7_C	1	0	0	1
R@P	114	PRO	P	V-BETA [D1]	1ao7_E	R@P	5	TYR	Y	(Ligand)	1ao7_C	19	1	0	18

Fig. 12 (continued)

epitope" is provided. Thus, the pMH1 epitope of 1ao7 (Fig. 13) comprises AA of G-ALPHA1 and G-ALPHA2 (1ao7_A) (HLA-A*0201) and of the peptide (1ao7_C, Tax peptide 11–19). Twenty-four AA form the pMH1 epitope: sixteen from the MH1 (six from G-ALPHA1 and ten from G-ALPHA2) and eight from the peptide. Each AA that belongs to the epitope is characterized by its position according to the IMGT unique numbering for G domain [7] and by its position in the peptide.

The TR paratope of 1ao7 (T-cell receptor A6) (Fig. 13) comprises AA of V-ALPHA (1ao7_D chain) and of V-BETA (1ao7_E chain). Sixteen AA of the TR (11 from V-ALPHA and 5 from V-BETA) form the paratope. The IMGT/Collier-de-Perles (Fig. 3) show that nine out of the 11 AA of the V-ALPHA paratope belong to the CDR-IMGT (D27, R28 and G29, Q37 to the CDR1-IMGT, Y57 to the CDR2-IMGT, T108, D109, and W113 and G114 to the CDR3-IMGT) and that five AA of the V-BETA paratope belong to the CDR3-IMGT and are localized at the top of loop. Clicking on "Epitope IMGT Residue@Position cards" and "Paratope IMGT Residue@Position cards" (Fig. 13) provide detailed contacts for each AA belonging to the epitope and paratope, respectively. IMGT paratope and epitope are

Fig. 13 "IMGT paratope and epitope" of an IMGT TR/pMH complex. Each AA that belongs to the pMH epitope is characterized by its position in the peptide or in the G domains according to the IMGT unique numbering [7]. For examples, "E (58G1_A)" means that the glutamate (E) is at position 58 of the G-ALPHA1 domain (1ao7_A), "AAH (61G2-62G2_A)" means that the alanine (A), alanine (A), and histidine (H) are at positions 61, 61A, and 62 of the G-ALPHA2 domain (1ao7_A) (*see* also Fig. 4a). Each AA that belongs to the TR paratope is characterized by its position in the V domains according to the IMGT unique numbering [32–34]. Thus, "DRG (27 V1-29V1_D1)" means that the aspartate (D), arginine (R), and glycine (G) are at positions 27, 28, and 29 of the V domain 1 of 1ao7_D (V-ALPHA) (*see* also Fig. 3a). In the same way, "GGRP (112.1V1-114V1_E)" means that the glycine (G), glycine (G), arginine (R), and proline (P) are at positions 112.1, 112, 113, and 114 of the V domain 1 of 1ao7_E (V-BETA) (*see* also Fig. 3b). The "IMGT paratope and epitope" analysis of the TR/pMH1 3-D structure (1ao7) is from IMGT/3Dstructure-DB, http:/www.imgt.org. (With permission from M-P. Lefranc and G. Lefranc, LIGM, Founders and Authors of IMGT®, the international ImMunoGeneTics information system®, http:/www.imgt.org)

determined automatically for the TR/pMH 3D structures in IMGT/3Dstructure-DB (*see* **Note 15**). Clicking on the "References and links" tag in the IMGT/3Dstructure-DB card gives access to external links. Links to the Immune Epitope Database (IEDB) [51, 52] are provided. Clicking on the "IMGT numbering comparison" displays, per chain, a table providing the correspondence between the IMGT unique numbering per domain and the PDB numbering of the chain entry.

4.4 Bridging IMGT Clonotype (AA), TR-Mimic Antibody and Paratope

4.4.1 IMGT Clonotype (AA) Repertoire and TR Paratope

Next-generation sequencing (NGS) data, analyzed by IMGT/HighV-QUEST, provides a standardized characterization of the TR repertoire diversity and expression in normal (e.g., before and after vaccination) and pathological situations. The results include the IMGT variable (V), diversity (D), and joining (J) gene and allele names (identified at the nucleotide level) and the identification of the JUNCTION lengths and amino sequences, which, together, characterize the IMGT clonotypes (AA) [53]. TR V domain analysis, using IMGT nomenclature [1, 6] and IMGT unique numbering [34] for both NGS and 3-D structures of TR/pMH complexes

Table 5

Amino acids of the TR paratope (V-ALPHA, V-BETA) and of the pMH1 epitope (G-ALPHA1, peptide, and G-ALPHA2) of 1ao7 based on Contact analysis from IMGT/3Dstructure-DB, http://www.imgt.org [12–14]. (A) TR V-ALPHA paratope – pMH1 epitope. (B) TR V-BETA paratope—pMH1 epitope. Amino acid positions of the TR V-ALPHA and V-BETA are according to the IMGT unique numbering for V-DOMAIN [33, 34]. Amino acid positions of the TR V-ALPHA and V-BETA are according to the IMGT unique numbering for G-DOMAIN [7]. The list of contact analysis below is complete. Differences observed in visual displays are due to filters, based on contact types or scores

(A) TR V-ALPHA paratope – pMH1 epitope of 1ao7				
PARATOPE		**EPITOPE**		
TR V-ALPHA [6.6.11][a]		**G-ALPHA1**	**(Ligand) Peptide**	**G-ALPHA2**
[15] 16 (9/7)[b] G-ALPHA1 [15] 13 (7/6)[b] peptide [12] 15 (7/8)[b] G-ALPHA2	1ao7_D	1ao7_A 7 amino acids (126:22,3,104|604) 58 E, 62 G, 65 R, 66 K, 68 K, 69 A, 72 Q	1ao7_C 6 amino acids (109:20,3,89|549) 1 L, 2 L, 3 F, 4 G, 5 Y, 6 P	1ao7_A 8 amino acids (105:17,2,88|468) 65 E, 66 Q, 69 A, 70 Y, 73 T, 76 E,77 W, 80 R
FR1-IMGT	2 K	58 E (7:1,0,6|26)		
	26 S	58 E (3:2,0,1|41)		
CDR1-IMGT	27 D	58 E (24:6,1,18|158)		
	28 R	58 E (1:1,0,0|20)		77 W (14:1,0,13|33) 80 R (13:5,0,8|108)
	29 G		1 L (5:0,0,5|5)	77 W (6:0,0,6|6)
	37 Q	66 K (4:1,0,3|23)	1 L (5:0,0,5|5) 2 L (4:2,1,2|62) 3 F (7:1,0,6|26) 4 G (6:2,0,4|44) 5 Y (1:0,0,1|1)	70 Y (8:0,0,8|8) 73 T (10:2;0,8|48)
	38 S		5 Y (10:2,1,8|68)	
CDR2-IMGT	57 Y			65 E (4:0,0,4|4) 66 Q (16:1,0,15|35) 69 A (7:1,0,6|26)

(continued)

Table 5
(continued)

(A) TR V-ALPHA paratope – pMH1 epitope of 1ao7				
PARATOPE		**EPITOPE**		
TR V-ALPHA [6.6.11][a]		**G-ALPHA1**	**(Ligand) Peptide**	**G-ALPHA2**
	58 S			69 A (3:1,0,2|22)
	63 N			76 E (4:2:0,2|42)
FR3-IMGT	82 K			73 T (5:2,1,3|63) 76 E (15:2,1,13|73)
CDR3-IMGT	107 T		5 Y (3:1,0,2|22)	
	108 T	65 R (5:2,1,3|63) 66 K (1:0,0,1|1)	4 G (2:1,0,1|21) 5 Y (4:1,0,3|23)	
	109 D	62 G (1:1,0,0|20) 65 R (19:5,1,14|134) 66 K (14:1,0,13|33)	4 G (11:2,0,9|49) 5 Y (16:2,0,14|54)	
	110 S		4 G (8:2,1,6|66) 5 Y (23:2,0,21|61) 6 P (4:2,0,2|42)	
	113 W	65 R (12:1,0,11|31) 68 K (8:0,0,8|8) 69 A (16:0,0,16|16) 72 Q (4:0,0,4|4)		
	114G	65 R (7:1,0,6|26)		

(B) TR V-BETA paratope – pMH1 epitope of 1ao7				
PARATOPE			**EPITOPE**	
TR V-BETA [5.6.14][a]		G-ALPHA1	(Ligand) Peptide	G-ALPHA2
[3] 4 (1/3)[b] G-ALPHA1 [14] 13 (9/4)[b] peptide [11] 10 (5/5)[b] G-ALPHA2	1ao7_E	1ao7_A 3 amino acids (23:0,0,23|23) 69 A, 72 Q, 73 T	1ao7_C 4 amino acids (119:9,2110|330) 5 Y, 6 P, 7 V, 8 Y	1ao7_A 6 amino acids (97:19,4,78|538) 61 A, 61A A, 62 H, 63 V, 66 Q, 76 E
CDR1-IMGT	37 E		8 Y (10:2,1,8|68)	
CDR3-IMGT	107 R		5 Y (5:2,0,3|43)	76 E (15:2,1,13|73)
	109 G		6 P (1:1,0,0|20)	
	110 L	69 A (9:0,0,9|9) 72 Q (4:0,0,4|4) 73 T (10:0,0,10|10)	6 P (12:1,0,11|31) 7 V (10:1,0,9|29) 8 Y (32:1,1,31|71)	
	111 A		7 V (4:0,0,4|4) 8 Y (5:0,0,5|5)	61A A (1:0,0,1|1)
	112.1 G		5 Y (1:0,0,1|1) 7 V (8:0,0,8|8)	61A A (5:0,0,5|5)
	112 G		5 Y (9:0,0,9|9) 7 V (2:0,0,2|2)	61A A (8:2,1,6|66) 62 H (4:1,0,3|23) 63 V (4:0,0,4|4) 66 (155) Q (10:2,1,8|68)
	113 R		5 Y (1:0,0,1|1)	61 (149) A (5:2,1,3|63) 61A (150) A (24:6,0,18|138) 62 (151) H (12:2,0,10|50) 66 (155) Q (2:1,0,1|21)
	114 P		5 Y (19:1,0,18|38)	66 (155) Q (7:1,0,6|26)

[a]CDR-IMGT lengths
[b][Residue pair contacts] Number of residues Total (From paratope/From epitope)
[c]Atom pair contact types (Total: polar, hydrogen, nonpolar|score)

[7, 12–14], provides a paradigm for bridging IMGT clonotypes (AA) of NGS repertoires and TR paratope CDR-IMGT (particularly CDR3-IMGT) delimitations.

4.4.2 TR-Mimic Antibody Paratope

The 3-D structures of an engineered TR-mimic antibody and that of a TR targeting peptide-HLA were recently compared**:** the IG Fab 3M4E5 and the TR 1G4_a58b61 are receptors targeting the NY-ESO-1 peptide presented by HLA-A*02:01 [54]. The pMH contacts of the NY-ESO-1 peptide SLLMWITQC with the MH1 HLA-A*02:01 groove are similar in the two peptide-HLA complexes, as expected [55]**.** The paratope of the IG Fab (TR-mimic antibody) includes amino acids of VH [8.8.12] and V-LAMBDA [9.3.9], whereas the paratope of the TR, classically, includes amino acids of V-BETA [5.6.12] and V-ALPHA [6.7.13]. The IMGT unique numbering for V-DOMAIN [34] was used for the four domains in the description of the paratopes. Similarly, in both 3-D structures, the IMGT unique numbering for G-DOMAIN [7] was used for the description of the pMH epitope, which comprises the G-ALPHA1 helix, the peptide, and the G-ALPHA2 helix [55].

5 Availability and Citation

Authors who use IMGT® databases and tools are encouraged to cite this article and to quote the IMGT® Home page, http://www.imgt.org. Online access to IMGT® databases and tools is freely available for academics and under licenses and contracts for companies.

6 Notes

1. Since the creation of IMGT® in 1989, at New Haven during the tenth Human Genome Mapping Workshop (HGM10), the standardized classification and nomenclature of the IG and TR of human and other vertebrate species have been under the responsibility of the IMGT Nomenclature Committee (IMGT-NC). In 1995, following the first demonstration online of the nucleotide database IMGT/LIGM-DB at the ninth International Congress of Immunology in San Francisco, IMGT-NC has become the World Health Organization-International Union of Immunological Societies (WHO-IUIS)/IMGT Nomenclature SubCommittee for IG and TR. IMGT® gene and allele names are based on the concepts of classification of

"Group," "Subgroup," "Gene," and "Allele," generated from the IMGT-ONTOLOGY CLASSIFICATION axiom. The IMGT® gene nomenclature for IG and TR genes was approved at the international level by the Human Genome Organisation (HUGO) Nomenclature Committee (HGNC) in 1999 and by the WHO-IUIS [56, 57]. The IMGT® IG and TR gene names [2, 6, 58, 59] are the official reference for the vertebrate genome projects and, as such, have been entered in IMGT/GENE-DB, the IMGT® gene database [60], in National Center for Biotechnology Information (NCBI) Gene [61], in European Bioinformatics Institute (EBI) Ensembl, and in the Vega Genome Browser (Wellcome Trust Sanger Institute).

2. AA one-letter and three-letter abbreviations: A (Ala), alanine; C (Cys), cysteine; D (Asp), aspartic acid; E (Glu), glutamic acid; F (Phe), phenylalanine; G (Gly), glycine; H (His), histidine; I (Ileu), isoleucine; K (Lys), lysine; L (Leu), leucine; M (Met), methionine; N (Asn), asparagine; P (Pro), proline; Q (Gln), glutamine; R (Arg), arginine; S (Ser), serine; T (Thr), threonine; V (Val), valine; W (Trp), tryptophan; and Y (Tyr), tyrosine. In Residue@Position (Subheading 2.4), the AA three-letter abbreviation is in capital letters. AA physicochemical properties [47] are described in IMGT Aide-mémoire, in the section "Amino acids," http://www.imgt.org.
3. Anchor positions, first defined for V domains, belong to the strands (or FR-IMGT in V-DOMAIN) and represent "anchors" supporting the three BC, C'C", and FG loops (CDR1-IMGT, CDR2-IMGT, and CDR3-IMGT, respectively, in V-DOMAIN). Anchor positions for V domains (V-DOMAIN and V-LIKE-DOMAIN) are positions 26 and 39, 55 and 66, and 104 and 118 [34]. By analogy, anchor positions were defined in C domains at positions 26 and 39, 45 and 77 (delimiting the transverse CD strand), and 104 and 118 [35]. Anchor positions are shown in squares in the IMGT Colliers de Perles.
4. "1ao7" is the code of a 3-D structure entry in the *Research Collaboratory for Structural Bioinformatics* (RCSB) Protein Data Bank (PDB) [62], or "PDB code" (comprising four letters and/or numbers). IMGT® uses the "PDB code" as "IMGT entry ID" for the 3-D structures in IMGT/3Dstructure-DB, http://www.imgt.org [12–14]. An additional letter separated by a "_" identifies the different chains in a 3-D structure. For example, the 1ao7 entry (a TR/pMH1 3D structure) comprises the following chains: 1ao7_D (TR-ALPHA) and 1ao7_E (TR-BETA-1) for the TR, 1ao7_C for the peptide, and 1ao7_A (I-ALPHA) and 1ao7_B (B2M) for the MH1.

5. The other characteristic AA at position 118 of V-DOMAIN is tryptophan (J-TRP) (Table 3) that is found in the IG heavy (IGH) joining (J) regions [3] (and is also found in one human T-cell receptor alpha TRA J region [6]).
6. IMGT color menu for the CDR-IMGT of a V-DOMAIN indicates the type of rearrangement, V-J or V-D-J [1]. Thus, the IMGT color menu for CDR1-IMGT, CDR2-IMGT, and CDR3-IMGT is blue, green, and greenblue for V-ALPHA (encoded by a V-J-REGION resulting from a V-J rearrangement) and red, orange, and purple for V-BETA (encoded by a V-D-J-REGION resulting from a V-D-J rearrangement). The color menu red, orange, and purple is also used for the V-LIKE-DOMAIN BC, C′C″, and FG loops, respectively. The assignment is done automatically by IMGT/DomainGapAlign [13, 40, 41].
7. MhSF proteins other than MH only include RPI-MH1Like proteins (there is no "RPI-MH2Like" identified so far) [45, 46]. The RPI-MH1Like in humans comprise: AZGP1 (that regulates fat degradation in adipocytes), CD1A to CD1E proteins (that display phospholipid antigens to T cells and participate in immune defense against microbian pathogens), FCGRT (that transports maternal immunoglobulins through placenta and governs neonatal immunity), HFE (that interacts with transferring receptor and takes part in iron homeostasis by regulating iron transport through cellular membranes), MICA and MICB (that are induced by stress and involved in tumor cell detection), MR1 (that may regulate mucosal immunity), PROCR, previously EPCR, (that interacts with activated C protein and is involved in the blood coagulation pathway), and RAET1E, RAETG, and RAET1L (that are inducible by retinoic acid and stimulate cytokine/chemokine production and cytotoxic activity of NK cells) [45, 46].
8. The princeps references for the IMGT unique numbering, used to define Residue@Position, are available as pdf on the IMGT® web site (http://www.imgt.org) in the IMGT Scientific chart section: IMGT unique numbering for V domain (V-DOMAIN of IG and TR and V-LIKE-DOMAIN of IgSF other than IG and TR) [32–34], IMGT unique numbering for C domain (C-DOMAIN of IG and TR and C-LIKE-DOMAIN of IgSF other than IG and TR) [35], and IMGT unique numbering for G domain (G-DOMAIN of MH and G-LIKE-DOMAIN of MhSF other than MH) [7].
9. "Residue@Position" characteristics include general information (PDB file numbering, IMGT file numbering, residue full name and formula) and structural information "IMGT Local-Structure@Position" (secondary structure, Phi and Psi angles (in degrees), and accessible surface area (ASA) (in square angstrom)).

10. Atom pair contacts identify interactions between atoms of two "R@P." They are obtained in IMGT/3Dstructure-DB by a local program in which atoms are considered to be in contact when no water molecule can take place between them [16, 17].
11. The "absent" "IMGT pMH contact sites" are correlated to MH class and to peptide length in the groove (Fig. 6). However, it is worthwhile to note that the "IMGT pMH contact sites" were initially defined from statistical analysis using experimental data from IMGT/3Dstructure-DB, without a priori of the MH class and of the peptide lengths [16, 17].
12. For the determination of the "IMGT pMH contact sites" in IMGT/3Dstructure-DB, 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 are taken into account [16, 17]. Then, MH AA involved in the pMH binding interface are filtered and classified in "IMGT pMH contact sites" by combining contact analysis with an interaction scoring function. The score assigned to each contact is a constant value, independent on the distance between atoms [40 for direct hydrogen bond, 20 for water mediated hydrogen bond, 20 for polar interaction, and 1 for nonpolar interaction, which roughly complies with the true mean energy ratio] [16, 17].
13. The "IMGT pMH contact sites" C1 and C11 correspond approximatively to the MH1 "pockets" A and F, respectively. A correspondence between the "IMGT pMH contact sites" and the other "pockets" is much more approximative. Thus, for MH1 with a 8-AA peptide, the "IMGT pMH contact sites" C3, C4, C6, and C9 correspond roughly to the B, D, C, and E pockets, and for MH1 with a 9-AA peptide, "IMGT pMH contact sites" C3, C4, and C9 correspond roughly to the B, D, and E pockets. For MH2, the correspondence is not possible because the pockets are poorly defined.
14. In January 2021, IMGT/3Dstructure-DB [12–14] contained 847 pMH (of which 754 are pMH1 and 93 are pMH2). Two hundred thirty-nine of the pMH belong to trimolecular TR/pMH complexes) [15–17] (*see* **Note 15**).

	Number of pMH in IMGT/3Dstructure-DB			
Complex type	***Homo sapiens***	***Mus musculus***	**Other species**	**Total**
pMH1	576	170	8	754
pMH2	84	8	1	93
Total	660	178	9	847

15. In January 2021, IMGT/3Dstructure-DB [12–14] contained 239 TR/pMH complexes (of which 186 are TR/pMH1 and 53 are TR/pMH2) [15–17].

	Number of TR/pMH complexes in IMGT/3Dstructure-DB			
Complex type	***Homo sapiens***	***Mus musculus***	**Other species**	**Total**
TR/pMH1	135	50	1	186
TR/pMH2	31	21	1	53
Total	166	71	2	239

Acknowledgements

We thank Patrice Duroux for the IMGT/3Dstructure-DB database and associated tools computing management, Anjana Kushwaha for the IMGT/3Dstructure-DB entries biocuration, and François Ehrenmann for help with the 3-D structure figures. We are grateful to the IMGT® team for its expertise and constant motivation. We thank Cold Spring Harbor Protocol Press for the pdf of the IMGT Booklet available in IMGT references. IMGT® is a registered trademark of CNRS. IMGT® is member of the International Medical Informatics Association (IMIA) and a member of the Global Alliance for Genomics and Health (GA4GH). All figures are used with permission from M-P. Lefranc and G. Lefranc, LIGM, Founders and Authors of IMGT®, the international ImMunoGeneTics information system®, http://www.imgt.org).

Funding: IMGT® was funded in part by the BIOMED1 (BIOCT930038), Biotechnology BIOTECH2 (BIO4CT960037), fifth PCRDT Quality of Life and Management of Living Resources (QLG2-2000-01287), and sixth PCRDT Information Science and Technology (ImmunoGrid, FP6 IST-028069) programs of the European Union (EU). IMGT® received financial support from the GIS IBiSA, the Agence Nationale de la Recherche (ANR) Labex MabImprove (ANR-10-LABX-53-01), the Région Occitanie Languedoc-Roussillon (Grand Plateau Technique pour la Recherche (GPTR), and BioCampus Montpellier. IMGT® is currently supported by the Centre National de la Recherche Scientifique (CNRS), the Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation (MESRI), and the University of Montpellier.

References

1. Lefranc M-P (2014) Immunoglobulin (IG) and T cell receptor genes (TR): IMGT® and the birth and rise of immunoinformatics. Front Immunol 5:22. https://doi.org/10.3389/fimmu.2014.00022
2. Lefranc M-P, Lefranc G (2001) The immunoglobulin FactsBook. Academic Press, London, UK, pp 1–458
3. Lefranc M-P, Lefranc G (2020) Immunoglobulins or antibodies: IMGT® bridging genes, structures and functions. Biomedicines 8(9): E319. https://doi.org/10.3390/biomedicines8090319
4. Lefranc LM-P, G. (2019) IMGT® and 30 years of Immunoinformatics Insight in Antibody V and C Domain Structure and Function. In Jefferis R; Strohl W. R., Kato K. Antibodies (Basel) 8(2):29. https://doi.org/10.3390/antib8020029
5. Lefranc M-P (2014) Immunoglobulins: 25 years of Immunoinformatics and IMGT-ONTOLOGY. Biomol Ther 4(4):1102–1139. https://doi.org/10.3390/biom4041102
6. Lefranc M-P, Lefranc G (2001) The T cell receptor FactsBook. Academic Press, London, UK, pp 1–398
7. Lefranc M-P, Duprat E, Kaas Q, Tranne M, Thiriot A, Lefranc G (2005) IMGT unique numbering for MHC groove G-DOMAIN and MHC superfamily (MhcSF) G-LIKE-DOMAIN. Dev Comp Immunol 29:917–938
8. Lefranc M-P, Clément O, Kaas Q, Duprat E, Chastellan P, Coelho I, Combres K, Ginestoux C, Giudicelli V, Chaume D, Lefranc G (2005) IMGT-choreography for immunogenetics and immunoinformatics. In Silico Biol 5(1):45–60
9. Lefranc M-P (2011) IMGT, the International ImMunoGeneTics Information System. Cold Spring Harb Protoc 2011(6):595–603. https://doi.org/10.1101/pdb.top115
10. Lefranc M-P (2013) IMGT® information system. In: Dubitzky W, Wolkenhauer O, Cho K-H, Yokota H (eds) Encyclopedia of systems biology. Springer Science+Business Media, LLC, New York, pp 959–964
11. Lefranc M-P, Giudicelli V, Duroux P, Jabado-Michaloud J, Folch G, Aouinti S, Carillon E, Duvergey H, Houles A, Paysan-Lafosse T, Hadi-Saljoqi S, Sasorith S, Lefranc G, Kossida S (2015) IMGT®, the international ImMunoGeneTics information system® 25 years on. Nucleic Acids Res 43:D413–D422. . Epub 2014 Nov 5. https://doi.org/10.1093/nar/gku1056
12. Kaas Q, Ruiz M, 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:D208–D210. https://doi.org/10.1093/nar/gkh042
13. Ehrenmann F, Kaas Q, Lefranc M-P (2010) IMGT/3Dstructure-DB and IMGT/DomainGapAlign: a database and a tool for immunoglobulins or antibodies, T cell receptors, MHC, IgSF and MhcSF. Nucleic Acids Res 38: D301–D307. https://doi.org/10.1093/nar/gkp946. Epub 2009 Nov 9
14. Ehrenmann F, Lefranc M-P (2011) IMGT/3Dstructure-DB: querying the IMGT database for 3D structures in immunology and Immunoinformatics (IG or antibodies, TR, MH, RPI, and FPIA). Cold Spring Harb Protoc 2011(6):750–761. https://doi.org/10.1101/pdb.prot5637
15. Lefranc M-P. (2007) MHC, what do we learn from IMGT colliers de Perles? ImmunoRIO2007, 13th international congress of immunology. Main symposia "Immunogenetics of MHC". A tribute to Jean Dausset. Rio de Janeiro, Brazil, 21 August 2007
16. Kaas Q, 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
17. Kaas Q, Duprat E, Tourneur G, Lefranc M-P (2008) IMGT standardization for molecular characterization of the T cell receptor/peptide/MHC complexes. In: Schoenbach C, Ranganathan S, Brusic V (eds) Immunoinformatics. Immunomics reviews, series of springer science and business media LLC, Springer, New York, USA, pp 19–49
18. Giudicelli V, Lefranc M-P (1999) Ontology for Immunogenetics: IMGT-ONTOLOGY. Bioinformatics 15:1047–1054
19. 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, Yousfi Monod M, Duprat E, Kaas Q, Pommié C, Chaume D, Lefranc G (2004) IMGT-ONTOLOGY for Immunogenetics and Immunoinformatics, http://www.imgt.org. In Silico Biol 4(1):17-29
20. Lefranc M-P (2004) IMGT-ONTOLOGY and IMGT databases, tools and web resources for immunogenetics and immunoinformatics. Mol Immunol 40:647–660

21. Lefranc M-P, Giudicelli V, Regnier L, Duroux P (2008) IMGT, a system and an ontology that bridge biological and computational spheres in bioinformatics. Brief Bioinform 9(4):263–275. https://doi.org/10.1093/bib/bbn014. Epub 2008 Apr 19
22. Duroux P, Kaas Q, Brochet X, Lane J, Ginestoux C, Lefranc M-P, Giudicelli V (2008) IMGT-kaleidoscope, the formal IMGT-ONTOLOGY paradigm. Biochimie 90:570–583. https://doi.org/10.1016/j.biochi.2007.09.003. Epub 2007 Sep11
23. Giudicelli V, Lefranc M-P (2012) IMGT-ONTOLOGY 2012. Frontiers in bioinformatics and computational biology. Front Genet 3: 79. Epub 2012 May 23
24. Giudicelli V, Lefranc M-P (2013) IMGT-ONTOLOGY. In: Dubitzky W, Wolkenhauer O, Cho K-H, Yokota H (eds) Encyclopedia of systems biology. Springer Science+Business Media, LLC, New York, pp 964–972
25. Lefranc M-P (2011) From IMGT-ONTOLOGY CLASSIFICATION axiom to IMGT standardized gene and allele nomenclature: for immunoglobulins (IG) and T cell receptors (TR). Cold Spring Harb Protoc 6: 627–632. https://doi.org/10.1101/pdb.ip84
26. Lefranc M-P (2011) IMGT unique numbering for the Variable (V), Constant (C), and Groove (G) domains of IG, TR, MH, IgSF, and MhSF. Cold Spring Harb Protoc 2011(6):633–642. https://doi.org/10.1101/pdb.ip85
27. Lefranc M-P (2013) IMGT unique numbering. In: Dubitzky W, Wolkenhauer O, Cho K-H, Yokota H (eds) Encyclopedia of systems biology. Springer Science+Business Media, LLC, New York, pp 952–959
28. Lefranc M-P (2011) IMGT collier de Perles for the variable (V), constant (C), and groove (G) domains of IG, TR, MH, IgSF, and MhSF. Cold Spring Harb Protoc 6:643–651. https://doi.org/10.1101/pdb.ip86
29. Lefranc M-P (2013) IMGT Collier de Perles. In: Dubitzky W, Wolkenhauer O, Cho K-H, Yokota H (eds) Encyclopedia of systems biology. Springer Science+Business Media, LLC, New York, pp 944–952
30. Lefranc M-P (2011) (2011) from IMGT-ONTOLOGY DESCRIPTION axiom to IMGT standardized labels: for immunoglobulin (IG) and T cell receptor (TR) sequences and structures. Cold Spring Harb Protoc 6: 614–626. https://doi.org/10.1101/pdb.ip83
31. Lefranc M-P (2011) (2011) from IMGT-ONTOLOGY IDENTIFICATION axiom to IMGT standardized keywords: for immunoglobulins (IG), T cell receptors (TR), and conventional genes. Cold Spring Harb Protoc 6:604–613. https://doi.org/10.1101/pdb.ip82
32. Lefranc M-P (1997) Unique database numbering system for immunogenetic analysis. Immunol Today 18:509
33. Lefranc M-P (1999) The IMGT unique numbering for immunoglobulins, T cell receptors and Ig-like domains. Immunologist 7: 132–136
34. Lefranc M-P, Pommié C, Ruiz M, Giudicelli V, Foulquier E, Truong L, Thouvenin-Contet V, Lefranc G (2003) IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. Dev Comp Immunol 27:55–77
35. Lefranc M-P, Pommié C, Kaas Q, Duprat E, Bosc N, Guiraudou D, Jean C, Ruiz M, Da Piedade I, Rouard M, Foulquier E, Thouvenin V, Lefranc G (2005) IMGT unique numbering for immunoglobulin and T cell receptor constant domains and Ig superfamily C-like domains. Dev Comp Immunol 29: 185–203
36. Ruiz M, Lefranc M-P (2002) IMGT gene identification and colliers de Perles of human immunoglobulins with known 3D structures. Immunogenetics 53:857–883
37. Kaas Q, Lefranc M-P (2007) IMGT colliers de Perles: standardized sequence-structure representations of the IgSF and MhcSF superfamily domains. Curr Bioinforma 2:21–30
38. Kaas Q, Ehrenmann F, Lefranc M-P (2007) IG, TR and IgSf, MHC and MhcSF: what do we learn from the IMGT colliers de Perles? Brief Funct Genomic Proteomic 6(4): 253–264. https://doi.org/10.1093/bfgp/elm032. Epub 2008 Jan 21
39. Ehrenmann F, Giudicelli., V, Duroux, P., Lefranc, M.-P. (2011) IMGT/collier de Perles: IMGT standardized representation of domains (IG, TR, and IgSF variable and constant domains, MH and MhSF groove domains). Cold Spring Harb Protoc 2011(6):726–736
40. Ehrenmann F, Lefranc M-P (2011) IMGT/DomainGapAlign: IMGT standardized analysis of amino acid sequences of variable, constant, and groove domains (IG, TR, MH, IgSF, MhSF). Cold Spring Harb Protoc 2011(6): 737–749. https://doi.org/10.1101/pdb.prot5636
41. Ehrenmann F, Lefranc M-P (2012) IMGT/DomainGapAlign: the IMGT® tool for the analysis of IG, TR, MH, IgSF, and MhSF domain amino acid polymorphism. Methods

Mol Biol 882:605–633. https://doi.org/10.1007/978-1-61779-842-9_33

42. Duprat E, Kaas Q, Garelle V, Lefranc G, Lefranc M-P (2004) IMGT standardization for alleles and mutations of the V-LIKE-DOMAINs and C-LIKE-DOMAINs of the immunoglobulin superfamily. In: Recent Research Developments in Human Genetics, vol 2. Research Signpost, Trivandrum, Kerla, pp 111–136
43. Bernard D, Hansen JD, du Pasquier L, Lefranc M-P, Benmansour A, Boudinot P (2005) Costimulatory receptors in jawed vertebrates: conserved CD28, odd CTLA4 and multiple BTLAs. Dev Comp Immunol 31:255–271
44. Garapati VP, Lefranc M-P (2007) IMGT colliers de Perles and IgSF domain standardization for T cell costimulatory activatory (CD28, ICOS) and inhibitory (CTLA4, PDCD1 and BTLA) receptors. Dev Comp Immunol 31: 1050–1072
45. Frigoul A, Lefranc M-P (2005) MICA: standardized IMGT allele nomenclature, polymorphisms and diseases. In: Pandalai SG (ed) Recent research developments in human genetics, vol 3. Research Signpost, Trivandrum, Kerala, pp 95–145
46. Duprat E, Lefranc M-P, Gascuel O (2006) A simple method to predict protein binding from aligned sequences - application to MHC superfamily and beta2-microglobulin. Bioinformatics 22:453–459
47. Pommié C, Levadoux S, Sabatier R, Lefranc M-P (2004) IMGT standardized criteria for statistical analysis of immunoglobulin V-REGION AA properties. J Mol Recognit 17:17–32
48. Lefranc M-P (2013) IMGT-ONTOLOGY, SpecificityType. In: Dubitzky W, Wolkenhauer O, Cho KH, Yokota H (eds) Encyclopedia of systems biology. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-9863-7_677
49. Lefranc M-P (2013) Paratope. In: Dubitzky W, Wolkenhauer O, Cho KH, Yokota H (eds) Encyclopedia of systems biology. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-9863-7_673
50. Lefranc M-P (2013) Epitope. In: Dubitzky W, Wolkenhauer O, Cho KH, Yokota H (eds) Encyclopedia of systems biology. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-9863-7_663
51. Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, Wheeler DK, Sette A, Peters B (2019) The immune Epitope database (IEDB): 2018 update. Nucleic Acids Res 47: D339–D343. https://doi.org/10.1093/nar/gky1006
52. Mahajan S, Vita R, Shackelford D, Lane J, Schulten V, Zarebski L, Jespersen MC, Marcatili P, Nielsen M, Sette A, Peters B (2018) Epitope specific antibodies and T cell receptors in the immune Epitope database. Front Immunol 9:2688. https://doi.org/10.3389/fimmu.2018.02688
53. Li S, Lefranc M-P, Miles JJ, Alamyar E, Giudicelli V, Duroux P, Freeman JD, Corbin VDA, Scheerlinck J-P, Frohman MA, Cameron PU, Plebanski M, Loveland B, Burrows SR, Papenfuss AT, Gowans EJ (2013) IMGT/HighV-QUEST paradigm for T cell receptor IMGT clonotype diversity and next generation repertoire immunoprofiling. Nat Commun 4: 2333. https://doi.org/10.1038/ncomms3333
54. Holland CJ, Crean RM, Pentier JM, de Wet B, Lloyd A, Srikannathasan V, Lissin N, Lloyd KA, Blicher TH, Conroy PJ, Hock M, Pengelly RJ, Spinner TE, Cameron B, Potter EA, Jeyanthan A, Molloy PE, Sami M, Aleksic M, Liddy N, Robinson RA, Harper S, Lepore M, Pudney CR, van der Kamp MW, Rizkallah PJ, Jakobsen BK, Vuidepot A, Cole DK (2020) Specificity of bispecific T cell receptors and antibodies targeting peptide-HLA. J Clin Invest 130(5):2673–2688. https://doi.org/10.1172/JCI130562
55. Lefranc M-P, Lefranc G. (2021) Antibody sequence and structure analyses using IMGT®: 30 years of immunoinformatics. Computer aided antibody design: methods and protocols.In: Tsumoto K, Kuroda D, eds. Methods Mol biol, in press
56. Lefranc M-P (2007) WHO-IUIS nomenclature subcommittee for immunoglobulins and T cell receptors report. Immunogenetics 59: 899–902
57. Lefranc M-P (2008) WHO-IUIS nomenclature subcommittee for immunoglobulins and T cell receptors report august 2007, 13th international congress of immunology, Rio de Janeiro, brazil. Dev Comp Immunol 32: 461–463
58. Lefranc M-P (2000) Nomenclature of the human immunoglobulin genes. In: Coligan JE, Bierer BE, Margulies DE, Shevach EM, Strober W (eds) Current protocols in immunology. John Wiley and Sons, Hoboken N.J, pp A.1P.1–A.1P.37
59. Lefranc M-P (2000) Nomenclature of the human T cell receptor genes. In: Coligan JE, Bierer BE, Margulies DE, Shevach EM, Strober W (eds) Current protocols in immunology.

John Wiley and Sons, Hoboken N.J, pp A.1O.1–A.1O.23
60. Giudicelli V, Chaume D, Lefranc M-P (2005) IMGT/GENE-DB: a comprehensive database for human and mouse immunoglobulin and T cell receptor genes. Nucleic Acids Res 33: D256–D261
61. Maglott D, Ostell J, Pruitt KD, Tatusova T (2007) Entrez gene: gene-centered information at NCBI. Nucleic Acids Res 35:D26–D31
62. Rose PW, Beran B, Bi C, Bluhm WF, Dimitropoulos D, Goodsell DS, Prlic A, Quesada M, Quinn GB, Westbrook JD, Young J, Yukich B, Zardecki C, Berman HM, Bourne PE (2011) The RCSB protein data Bank: redesigned web site and web services. Nucleic Acids Res 39:D392–D401