

Nomenclature des anticorps: d'IMGT-ONTOLOGY à la définition DCI

Marie-Paule Lefranc

IMGT® Fondateur et Directeur,
Professeur Université Montpellier 2,
IGH UPR CNRS 1142, Montpellier

GDR 3260 – ACCITH, Marseille, 25-26 Novembre 2010

From IMGT-ONTOLOGY to INN definition



1. Antibodies (mAb) and fusion proteins for immune applications (FPIA) represent a large number of the pharmaceutical substances submitted to the World Health Organization/International Nonproprietary Names (WHO/INN) Programme.

DCI: Dénomination Commune Internationale

World Health Organization. INN Working Document 09.251.Update 18/12/2009

2. The INN definition of antibodies is based on the IMGT-ONTOLOGY concepts of:

- **classification** (nomenclature)
- **description** (labels)
- **numerotation** (IMGT unique numbering).

Lefranc M-P. mAbs 3:1-2 (2011)

Monoclonal antibodies (INN -mab) and fusion proteins for immune applications (FPIA) (INN -cept) on the market



5 Human mAbs	adalimumab (2002) HUMIRA® / TRUDEXA®	panitumumab (2006) VECTIBIX™	golimumab (2009) SIMPONI™	canakinumab (2009) ILARIS®	ustekinumab (2009) STELARA™
--------------	--	--	---	--	---



5 Murine mAbs	muromonab-CD3 (1992) ORTHOCLONE OKT3®	edrecolomab (1995) PANOREX®	ibritumomab tiuxetan (2002) ZEVALIN®	tositumomab (2003) BEXXAR®	capromab * (1996) PROSTASCINT®
---------------	---	---	--	--	--



5 Chimeric mAbs	abciximab (1994) REOPRO®	rituximab (1997) MABTHERA® / RITUXAN®	basiliximab (1998) SIMULECT®	infliximab (1998) REMICADE®	cetuximab (2004) ERBITUX®
-----------------	--	---	--	---	---



11 Humanized mAbs	daclizumab (1997) ZENAPAX®	palivizumab (1998) SYNAGIS®	trastuzumab (1998) HERCEPTIN®	alemtuzumab (2001) CAMPATH® / MABCAMPATH®	omalizumab (2003) XOLAIR®	ranibizumab (2006) LUCENTIS®	eculizumab (2007) SOLIRIS™	natalizumab (2008) TYSABRI®	bevacizumab (2008) AVASTIN®	certolizumab pegol (2008) CIMZIA®	nimotuzumab (2008) THERACIM®
-------------------	--	---	---	---	---	--	--	---	---	---	--



4 FPIA	etanercept (1998) ENBREL®	alefacept (2003) AMEVIVE®	abatacept (2005) ORENCIA®	rilonacept (2008) ARCALYST™
--------	---	---	---	---

Conjugated and radiolabelled monoclonal antibodies

Monoclonal antibodies can be used:

conjugated with another molecule

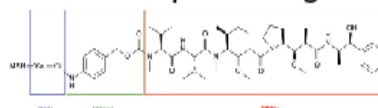
- Calicheamicin, a cytotoxic antitumor antibiotic
Ex: gentuzumab ozogamicin, MYLOTARG®
Scientists believe Alexander the Great was killed by calicheamicin, a dangerous compound produced by a bacteria (*Micromonospora echinospora*) found in water. (Squires N., Telegraph, 03 August 2010)



- Exotoxin A from *Pseudomonas Aeruginosa*
Ex: oportuzumab monatox, PROXINIUM™ VICINIUM™



- Monomethyl auristatin E (MMAE), a synthetic antineoplastic agent
Ex: glembatumumab vedotin

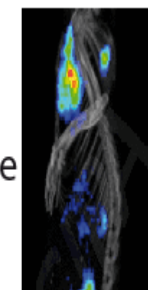


- Bouganin, from *Bougainvillea spectabilis* Willd toxin
Ex: citatuzumab bogatox

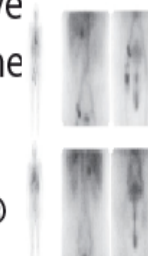


radiolabelled with an isotope

- Indium-111:
Half-Life 2.80 days, used in nuclear medicine to observe tumors and LCD manufacturing,
Ex: capromab, PROSTASCINT®



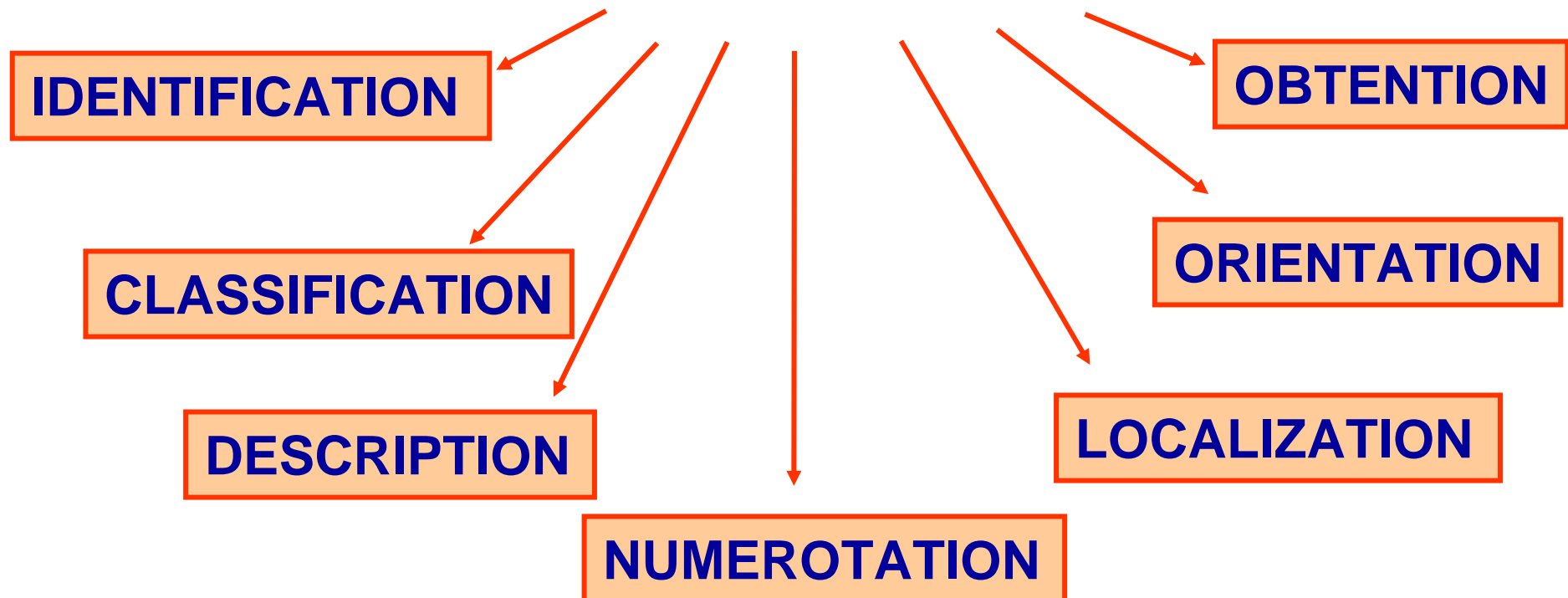
- Iodine-131:
Half-Life 8 days, used in nuclear medicine to observe a thyroid gland and radiotherapy on cancers such as lymphoma,
Ex: tositumomab, BEXXAR®



In November 2010, IMGT/mAb-DB contains: 17 INN **radiolabelled** and 11 INN **conjugated**.

IMGT-ONTOLOGY seven axioms:

To share, reuse and represent knowledge
in Immunogenetics and Life Sciences



Concepts of CLASSIFICATION



<http://www.imgt.org>

1. The IMGT-ONTOLOGY main concepts of classification
 - include 'group', 'subgroup', 'gene', 'allele'.
 - have allowed to set up the **nomenclature** of the immunoglobulin (IG) and T cell receptor (TR) genes (V, D, J, C genes).
2. **IMGT gene names** have been approved by the **HUGO Nomenclature Committee (HGNC)** in 1999.
3. **New alleles** are validated by the **WHO-IUIS/IMGT nomenclature committee** and entered in **IMGT/GENE-DB**.
4. **IMGT/GENE-DB** is the **international reference database** for IG and TR genes (direct links from **NCBI Entrez Gene**) and alleles.

Concepts of CLASSIFICATION



<http://www.imgt.org>

1. The IMGT-ONTOLOGY main concepts of classification
 - include 'group', 'subgroup', 'gene', 'allele'.
 - have allowed to set up the nomenclature of the immunoglobulin (IG) and T cell receptor (TR) genes (V, D, J, C genes).
2. IMGT gene names have been approved by the HUGO Nomenclature Committee (HGNC) in 1999.
3. New alleles are validated by the WHO-IUIS/IMGT nomenclature committee and entered in IMGT/GENE-DB.
4. IMGT/GENE-DB is the international reference database for IG and TR genes (direct links from NCBI Entrez Gene) and alleles.

Concepts of DESCRIPTION



<http://www.imgt.org>

1. The IMGT-ONTOLOGY concepts of description:
 - comprise the **standardized IMGT labels** and their **relations**.
 - have allowed **to describe** the IG (or antibody) and TR sequences and structures, **whatever the receptor type, the chain type or the species**.
2. **IMGT labels** are used in all IMGT® databases and tools for the description of:
 - **nucleotide and amino acid sequences (IMGT/LIGM-DB...)**
 - **2D and 3D structures (IMGT/3Dstructure-DB...)**.
3. Sequence Ontology (**SO**) includes **IMGT labels**.
4. **IMGT® databases** can be queried **using labels** (a big 'plus' compared to generalist databases).

Concepts of DESCRIPTION



<http://www.imgt.org>

1. The IMGT-ONTOLOGY concepts of description:
 - comprise the **standardized IMGT labels** and their **relations**.
 - have allowed **to describe** the IG (or antibody) and TR sequences and structures, **whatever the receptor type, the chain type or the species**.
2. **IMGT labels** are used in all IMGT® databases and tools for the description of:
 - **nucleotide and amino acid sequences (IMGT/LIGM-DB...)**
 - **2D and 3D structures (IMGT/3Dstructure-DB...)**.
3. Sequence Ontology (**SO**) includes **IMGT labels**.
4. **IMGT® databases** can be queried **using labels** (a big 'plus' compared to generalist databases).

Concepts of NUMEROTATION



<http://www.imgt.org>

1. The IMGT-ONTOLOGY concepts of numerotation include:
 - IMGT unique numbering
 - IMGT Collier de Perles.
2. The concepts bridge the gap between sequences and 3D structures, at the amino acid (and codon) level, for:
 - the variable domains (V-DOMAIN and V-LIKE-DOMAIN)
 - the constant domains (C-DOMAIN and C-LIKE-DOMAIN).
4. The concepts are used for:
 - mutations, polymorphisms
 - CDR-IMGT lengths
 - contact analysis, paratope definition.
5. WHO-INN programme requires the CDR-IMGT lengths for antibody.

Concepts of NUMEROTATION



<http://www.imgt.org>

1. The IMGT-ONTOLOGY concepts of numerotation include:

- IMGT unique numbering
- IMGT Collier de Perles.

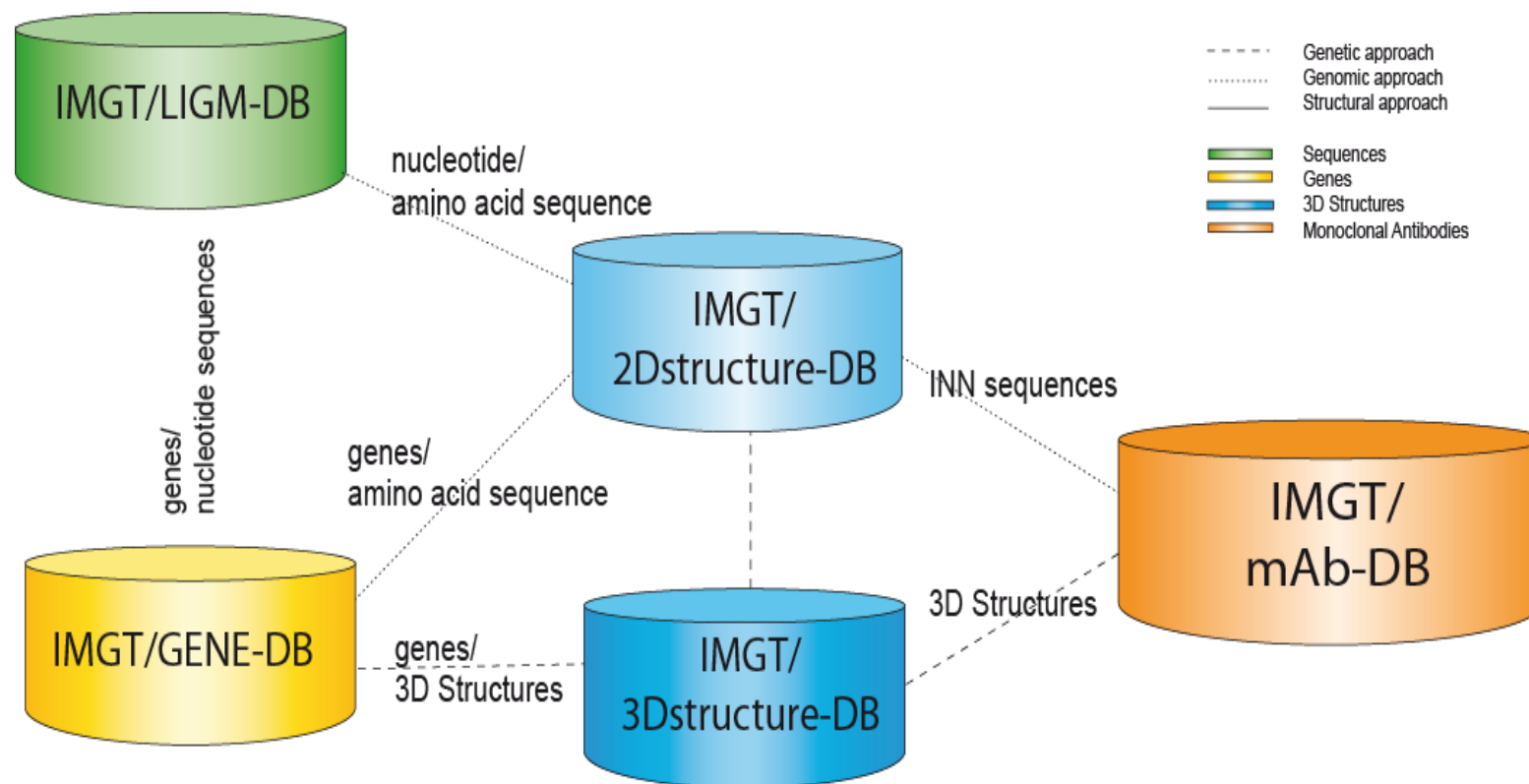
2. The concepts bridge the gap between sequences and 3D structures, at the amino acid (and codon) level, for:

- the variable domains (V-DOMAIN)
- the constant domains (C-DOMAIN).

4. The concepts are used for:

- Mutations, polymorphisms
- CDR-IMGT lengths
- contact analysis, paratope definition.

5. WHO-INN programme requires the CDR-IMGT lengths for antibody.



IMGT/mAb-DB



IMGT/mAb-DB is the clinical antibodies database of IMGT®.

IMGT®, the international ImMunoGeneTics information system®, is the global reference in immunogenetics and immunoinformatics.

IMGT/mAb-DB provides a unique expertised resource:

- on immunoglobulins (IG) or monoclonal antibodies (mAb) with clinical indications
- on fusion proteins for immune applications (FPIA).

Since 2008, amino acid sequences of mAb (INN suffix -**mab**) and of fusion proteins for immune applications FPIA (INN suffix -**cept**) from the WHO/INN Programme have been entered in IMGT®.

IMGT/mAb-DB Query page



<http://www.imgt.org>

1. IMGT/mAb-DB ID

2. INN (International Nonproprietary Name) INN proposed list and before and after
INN number INN recommended list and before and after
Common name Proprietary name

3. IMGT/mAb-DB receptor type OR IMGT/mAb-DB species
Radiolabelled Conjugated
Isotype and format Entries in IMGT/2Dstructure-DB
FPIA or RPI format Entries in IMGT/3Dstructure-DB
Origin clone species Origin clone name

4. Specificity target name
Specificity target species

5. Clinical indication Development status
Regulatory agency Year
Company Expression system
Application Clinical domain

1. IMGT/mAb DB ID

2. INN and other names

3. Characteristics and structure

4. Specificity (target name and target species)

5. Clinical indication, development status and clinical domain

IMGT/mAb-DB, <http://www.imgt.org>

Your query: **IMGT/mAb-DB INN = alemtuzumab**



<http://www.imgt.org>

IMGT/mAb-DB ID	INN (International Nonproprietary Name)	INN Num.	INN Prop. list	INN Rec. list	Common name	Proprietary name
11	alemtuzumab	8005	<u>83</u> (2000)	<u>45</u> (2001)	Campath-1H, LDP-03	<u>CAMPATH®</u> (US) MABCAMPATH® (EU)

A query on a name allows to retrieve the International Nonproprietary Name (INN) and INN lists numbers as provided by the World Health Organization (WHO)/INN Programme, the Common names from literature, and the Proprietary name(s) if the antibody or fusion protein for immune applications (FPIA) is a registered trademark (symbol®)

IMGT/mAb-DB, <http://www.imgt.org>

Your query: IMGT/mAb-DB INN = alemtuzumab



<http://www.imgt.org>

IMGT/mAb-DB species	Isotype and format	IMGT/2Dstructure-DB	IMGT/3Dstructure-DB	Specificity (target) [origin]	Origin clone species	Origin clone name
Humanized	IgG1k	8005	1bey 1ce1	CD52 [<i>Homo sapiens</i>]	<i>Rattus norvegicus</i>	YTH34.5HL

For each entry, IMGT/mAb-DB provides the origin species (human, rat, murine, humanized or chimeric), the isotype and format, links to IMGT/2Dstructure-DB (amino acid sequences and IMGT Colliers de Perles), links to IMGT/3Dstructure-DB (3D structures), specificity (target), origin clone species and origin clone name.

IMGT/mAb-DB, <http://www.imgt.org>

Your query: IMGT/mAb-DB INN = alemtuzumab



<http://www.imgt.org>

Company	Clinical indication	Development status	Regulatory agency status and year	Application	Clinical domain	FDA / EMA / NCI number or ATC code
Berlex Inc. (Wayne NJ USA) (US) / Genzyme Corp. (Cambridge MA USA) (EU) / Millennium Pharmaceuticals Inc. (Cambridge MA USA) (EU/US)	Chronic lymphocytic leukemia (CLL)	Phase III	FDA approval May 2001	Therapeutic	Hematology, Oncology	FDA: (BLA) 103948
	Kidney transplant rejection	Phase III		Diagnostic	Immunology	NCI: C1681
	Multiple sclerosis (MS)	Phase III		Therapeutic	Immunology	Drugnum: 702

IMGT/mAb-DB also provides information on company, clinical indications (more than 200 in the database), development status, organizations that approved the drug such as Food and Drugs Administration (FDA) or European Medicines Agency (EMA), applications (diagnostic or therapeutic) and clinical domains.

IMGT/mAb-DB, <http://www.imgt.org>

INN mAbs (-mab) and FPIA (-cept) in IMGT/2Dstructure-DB



<http://www.imgt.org>

IMGT/2Dstructure-DB card for INN: 8005



IMGT molecule name	IMGT receptor type	IMGT receptor description	Ligand(s)	Species	CC	Chain ID
INN name alemtuzumab						
Common name CAMPATH-1H, LDP-03	IG	IG-GAMMA-1_KAPPA		Humanized (Humanized)	1	[8005_H 8005_L]
Commercial name MABCAMPATH®						

Proposed list L83 (2000) Recommended list R45 (2001)

IMGT Notes
 Alemtuzumab constant genes and alleles, and allotypes, based on sequence analysis are:
 IGHG1*01, *02 (100%) allotype G1m17,1
 IGKC*01 (100%) allotype Km3
 The allotypes have been confirmed serologically.
 Jefferis R. and Lefranc M.-P., mAbs, 1:4, 332-338 (2009).

- Chain details
- Contact analysis
- 3D visualization Jmol or QuickPDB
- Renumbered IMGT file
- INN definitions
- References and links

Chain details

Differences with the closest IMGT allele sequence are in orange.

Chain details of [alemtuzumab](#), CAMPATH-1H, LDP-03, **IG**, IG-GAMMA-1_KAPPA Humanized (Humanized) [8005_H,8005_L]

Chain ID	INN 8005_H
Chain length	451
IMGT chain description	H-GAMMA-1 = VH (1-121) [D1] + CH1 (122-219) [D2] + HINGE-REGION (122-219) + CH2 (235-344) [D3] + CH3 (345-451) [D4]
Chain sequence	<pre> [VH (1-121) [D1] QVQLQESGPGCLVPSQTLNLTCTVSCFTFDYMNWVROPPEGCLEWIGFIRDKARGYTTETNPSSVRCRVTLMLVDTSKNQFSLRLSSVTA]N-AND-D[J-REGION] [ADTAVYYCAREGHTAAPPDYWGQCSLVTVSSASTKCPVFPPLAPSSKSTSGCTAALGCLVRDYPPEPVTVSWSGALTSQVHTFPAVLQS] [HINGE-REGION] [SGLYSLSSVWVTPSSSLGTQTYICNVNHRKPSNTRKVDKRVKPKSCDKRHTCPCCPAPELLGCGPSVFLFPPKPKDTLMISRTPEVTCVVDV] CH2 (235-344) [D3] SHEDPEVKFNWYVDGVEVHNARTKPREPQYNSTYRVVSVLTVLHQDWLNGKEYKCRVSNKALPAPIEKTIISKAKGQPREPQVYTLPPSRD] CH3 (345-451) [D4] ELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSPFLYSKLTVDKSRWQQGNVFPSCVMHEALHNHYTQKSLSLSPG] K </pre>
	Sequence in FASTA format Sequence in IMGT format

Ehrenmann, F. et al. Nucleic Acids Res., 38:D301-307 (2010)

An example of INN definition for a mAb



<http://www.imgt.org>

CLASSIFICATION axiom:
IMGT gene and allele nomenclature

Definitions for INN 9134, <i>intetumumab</i>	
English	Immunoglobulin G1-kappa, anti-[<i>Homo sapiens</i> integrin alpha-V (CD51, ITGAV, subunit of alphaV-beta3 or CD51/CD61 or vitronectin receptor or VNR, subunit of alphaV beta5)], <i>Homo sapiens</i> monoclonal antibody gamma1 heavy chain (1-449) [<i>Homo sapiens</i> VH (IGHV3-30*01 (91.80%) - (IGHD)-IGHJ3*02 [8.8.12] (1-19) - IGHG1*01 (120-449)), (222-215')-disulfide with kappa light chain (1'-215') [<i>Homo sapiens</i> V-KAPPA (IGKV3-11*01 (100.00%) - IGKJ3*01 [6.3.10] (1'-108') - IGKC*01 (109'-215')) (228-228":231-231")-bisdissulfide dimer antineoplastic

NUMEROTATION axiom:
CDR-IMGT

DEFINITION axiom
IMGT labels

From IMGT-ONTOLOGY to INN definition



<http://www.imgt.org>

The INN definition of antibodies provides:

1. For the V domains

- the closest V and J genes and alleles corresponding to the amino acid (AA) sequences
- the percentage of identity of the V regions for the humanized (INN -zumab) and human (INN -umab) antibodies
- AA changes of the FR4-IMGT, by comparison with the closest germline J genes and alleles
- CDR-IMGT lengths

From IMGT-ONTOLOGY to INN definition



The **INN definition** of antibodies provides:

2. For the C domains

- the **closest C genes and alleles** corresponding to the amino acid (AA) sequences
- **AA changes of the constant region**, by comparison with the reference alleles (and expected allotypes)

AA changes are indicated based on the **IMGT unique numbering** per domain with, between parentheses, positions in the sequence.

This information is obtained with the **IMGT/DomainGapAlign** tool (<http://www.imgt.org>).

IMGT/DomainGapAlign

Sequence name: [alemumuzab_H](#)

Move your mouse over the amino acids in bold for the characterization of AA class changes

Closest reference gene and allele(s) from the IMGT domain directory


V gene and allele	Species	Domain	Smith-Waterman Score	Identity percentage	Overlap
IGHV4-59*01	Homo sapiens	1	494	73.0	100
J gene and allele	Species	Domain	Smith-Waterman Score	Identity percentage	Overlap
IGHJ4*01	Homo sapiens	1	94	92.9	14
IGHJ4*02	Homo sapiens	1	94	92.9	14
IGHJ4*03	Homo sapiens	1	94	92.9	14

Alignment with the closest genes and alleles from the IMGT domain directory

	FR1-IMGT (1-26)	CDR1-IMGT (27-38)	FR2-IMGT (39-55)	CDR2-IMGT (56-65)	FR3-IMGT (66-104)	CDR3-IMGT (105-117)	FR4-IMGT (118-128)																									
	A (1-15)	B (16-26)	BC (27-38)	C (39-46)	C' (47-55)	C'' (56-65)	C'' (66-74)	D (75-84)	E (85-96)	F (97-104)	FG (105-117)	G (118-128)																				
	1	10	15	16	23	26	27	38	39	41	46	47	55	56	65	66	74	75	80	84	85	89	96	97	104	105	111	112	117	118	128	
alemumuzab_H	QVQLQESGP	GLVLRP	SQTLISL	CTVSG	FTF	...TDFY	MNWRQPP	GRGLEWIGF	IRDKAKGYTT	EYNPSVK	GRVTMLVDTSK	NQFSLRLSSVIA	ADTAVYYC	AREGHTAAPFDYW	GQGSLVTIVSS																	
IGHV4-59*01	QVQLQESGP	GLVLRP	SEITLSL	CTVSG	GGSI	...SSYY	WSWIRQPP	GKGLEWIGY	IYYS	...GST	NYNPSLK	SRVTIISVDTSK	NQFSLRLSSVIA	ADTAVYYC	AR																	
<i>(Homo sapiens)</i>			R	Q			FTF	TDF	MN	V	R	F	RDK	YT	E	V	G	ML					R								S	
	-----V-REGION-----												---N---		-----J-REGION-----																	

Region(s) and domain(s) identified in your sequence (corresponding to the closest genes and alleles)

Without gaps [Sequence in FASTA format](#)


 Download

With gaps [Sequence in FASTA format](#)

 Download

```

QVQLQESGPGGLVLRPSQTLISLCTVSGFTTFIDFYMNWRQPPGRGLEWIGF
IRDKAKGYTTIEYNPSVKGRVTMLVDTSKNQFSLRLSSVIAADTAVYYCAR
EGHTAAPFDYWQGSLVTIVSSASTKGPSVFLPAPSSKSTSGGTAALGCLV
KDYFPEPVTVSWNSGALTSQVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQ
TYICNVNHKPSNTRVDDKKEAPELLGGPSVFLFPPKPKDTLMISRTPEVT
CVVVDVSHEDPEVKFNWYVDGVEVHNAKTRPREEQYNSTYRVVSVLTVLH
QDWLNGKEYRCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTK
NQVSLTCLVKGFPYPSDIAVEWESNGQPENNYKTTTPVLDSDGSEFFLYSKL
TVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK
    
```

 IMGT Collier de Perles

Acknowledgements



<http://www.imgt.org>

BioSTIC-LR

ACI IMPbio

GIS AGENAE

Plan Pluri-Formation Université Montpellier 2

ANR FLAVORES

ANR BIOSYS

GIS IBiSA

Grand Plateau Technique Régional Languedoc-Roussillon GPTR
«ImmunoGrid», 6th PCRDT, STREPS IST



and the GDR ACCITH, the national and international institutions and companies that support the IMGT efforts of standardization.



Many thanks to the IMGT® team at Montpellier, France



Im
Muno
Gene
Tics

Information
system®

<http://www.imgt.org>