

# IMGT®: a paradigm for genetics, genome and 3D structure data integration towards Systems Biology

**Marie-Paule Lefranc**

University Montpellier 2  
Institute of Human Genetics CNRS

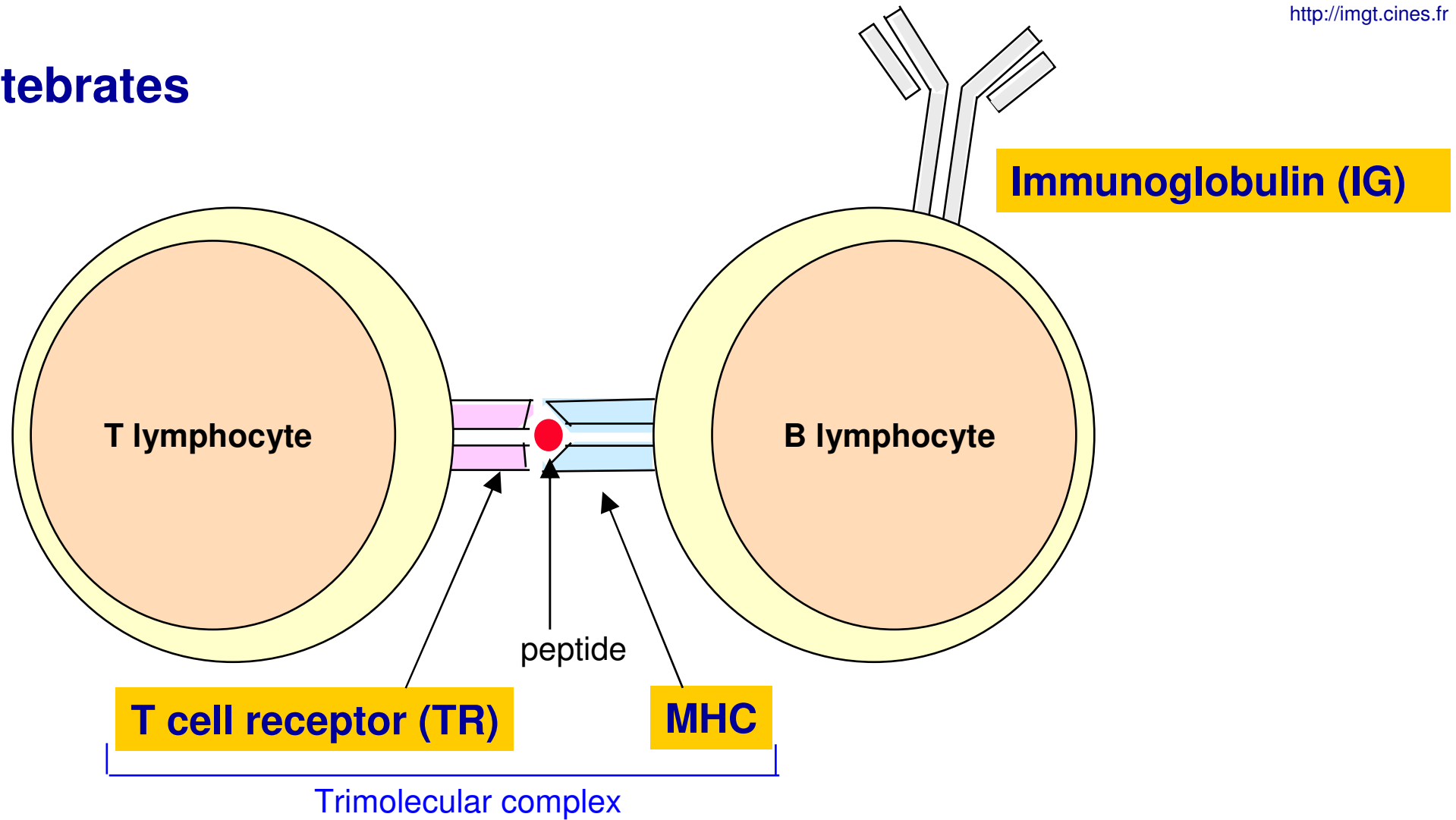
International Symposium on Biotechnology, ISB2008,  
4-8 May, 2008, SFAX, TUNISIA

# Outline

- What is the IMGT® domain of expertise?  
*Adaptive immune response*
- Why and how has IMGT® become a paradigm towards Systems Biology?  
*IMGT-ONTOLOGY axioms and concepts*
- Examples of IMGT tools based on the IMGT-ONTOLOGY concepts  
*IMGT/JunctionAnalysis, IMGT/V-QUEST, IMGT/3Dstructure-DB for antibody engineering and antibody humanization*
- Conclusions and Perspectives

# IMGT® domain: the adaptive immune response

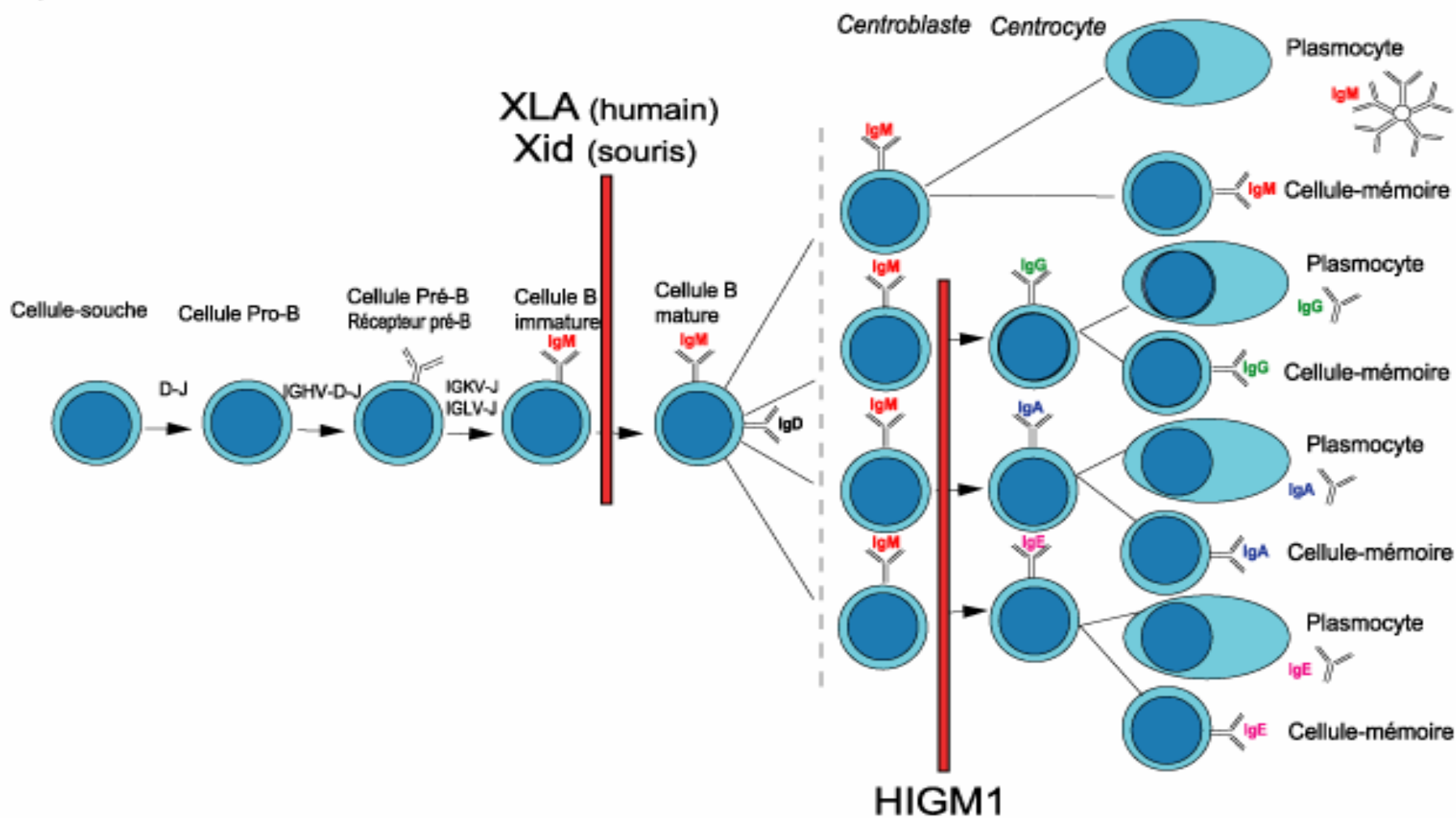
## Vertebrates



Bone marrow

Blood

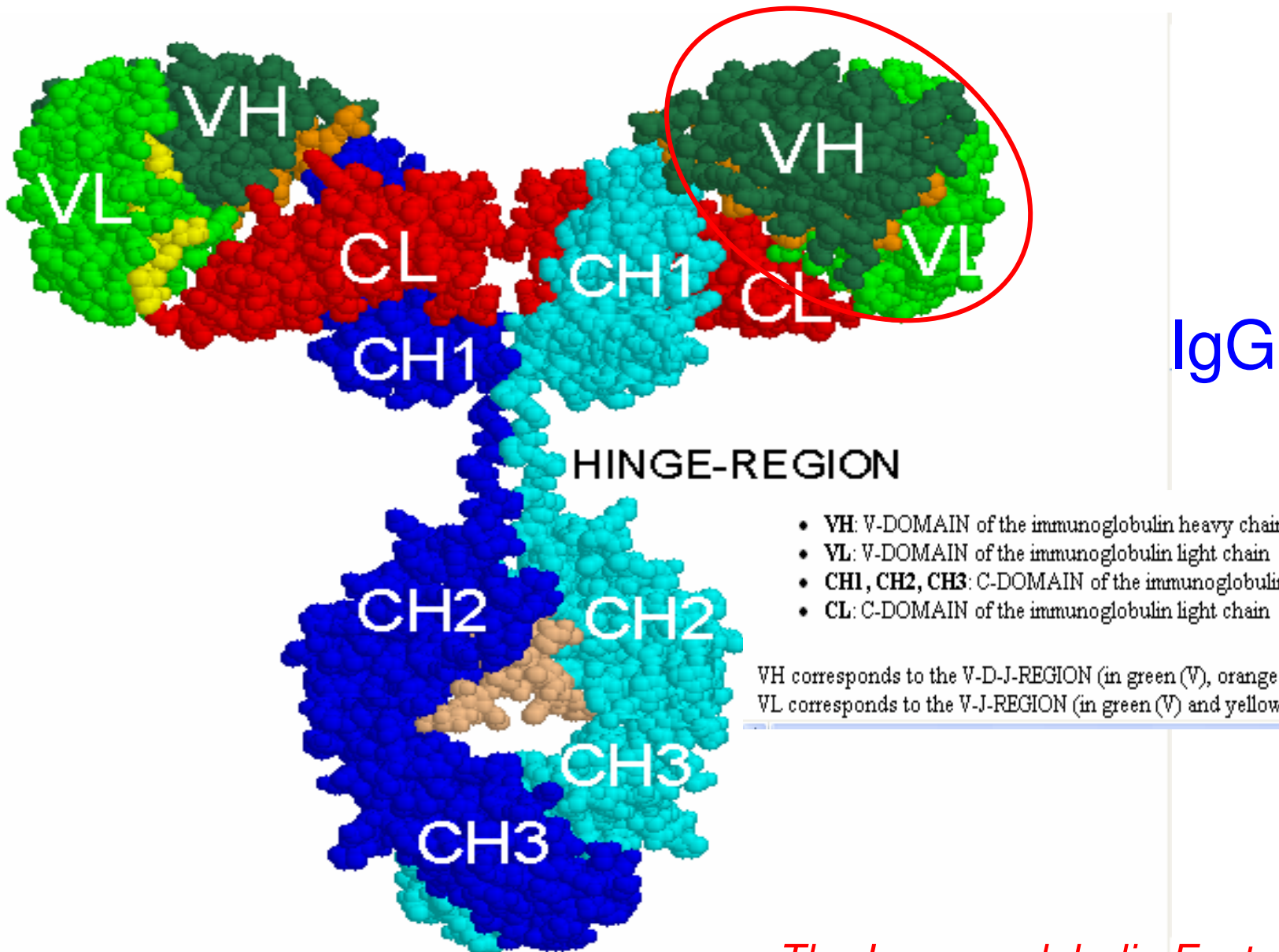
Lymph nodes, spleen



V-D-J and V-J rearrangements

Hypermutations, selection

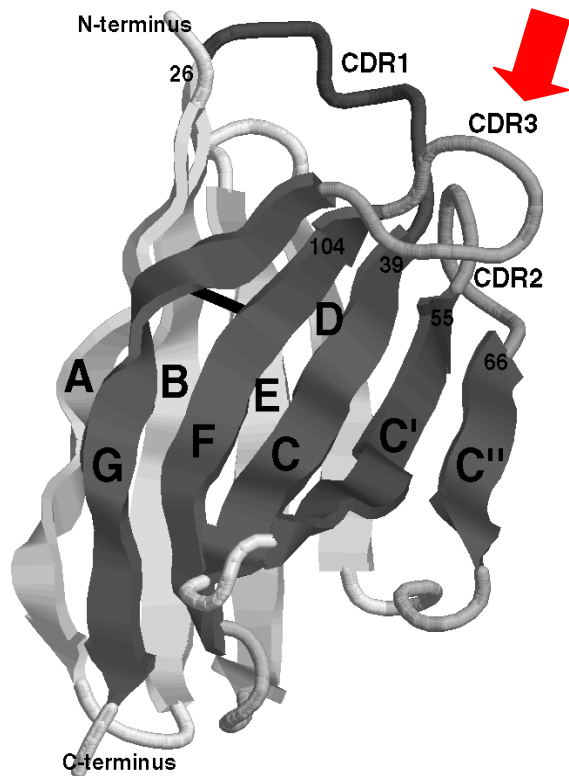
# Immunoglobulin or antibody



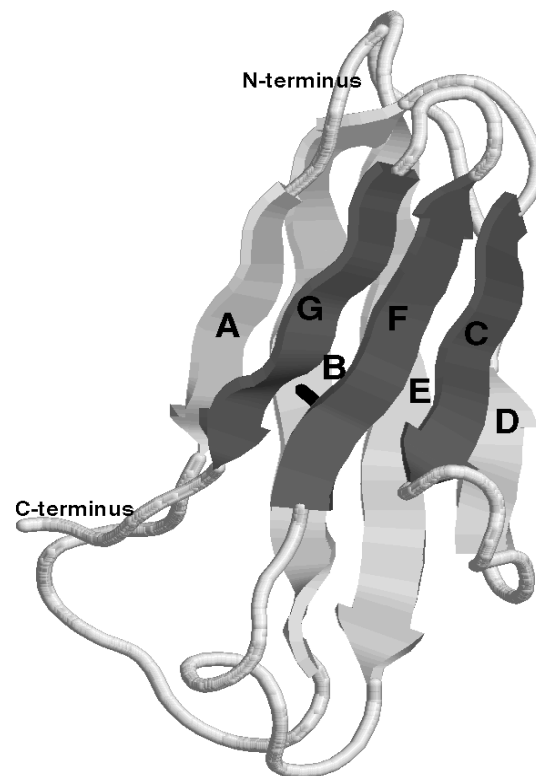
# Structural domains

## IG and TR

### V-DOMAIN



### C-DOMAIN

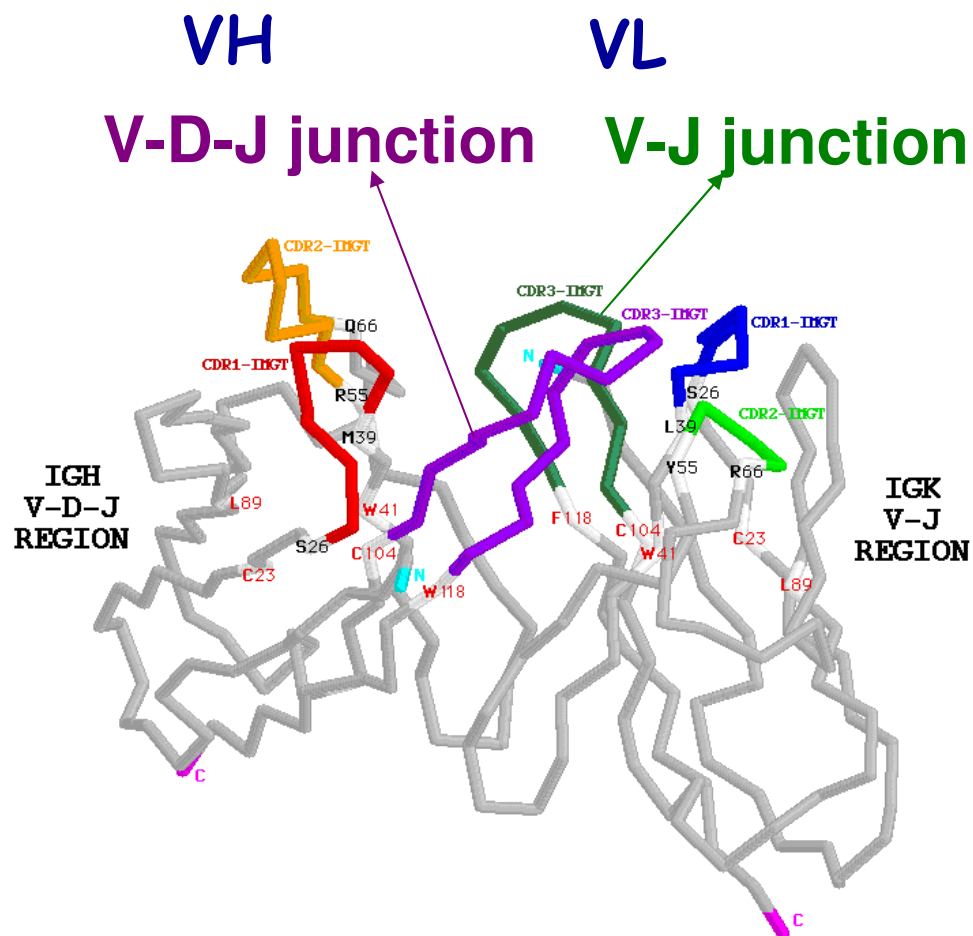


## MHC

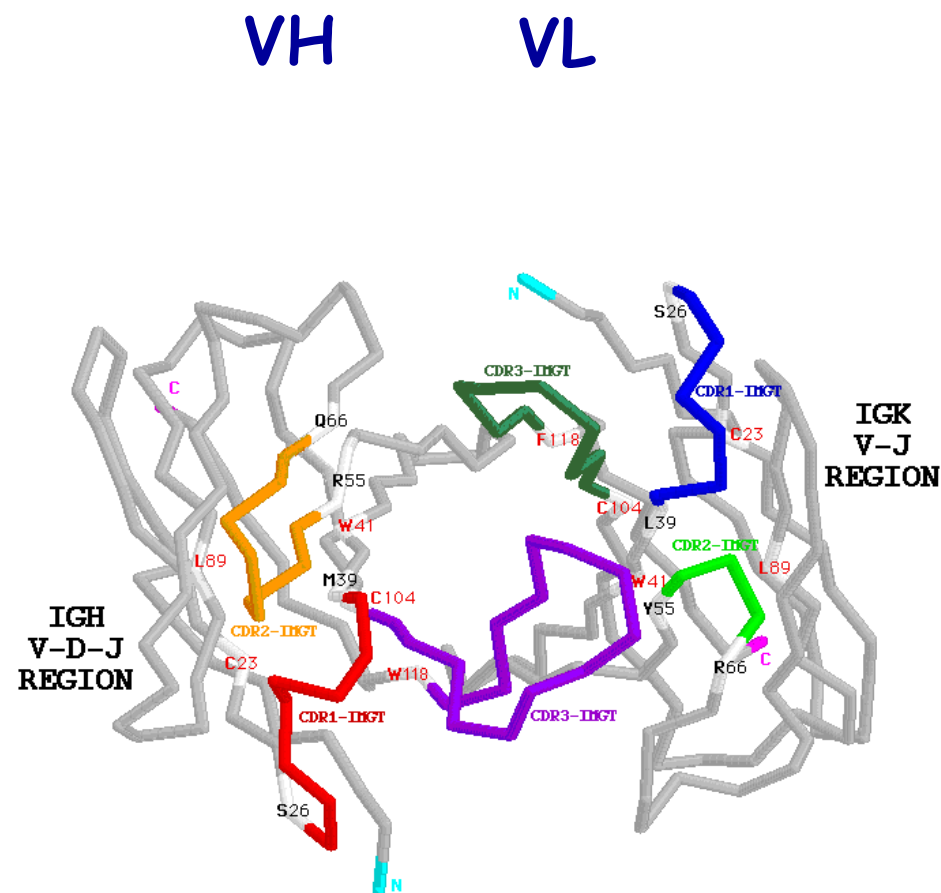
### G-DOMAINS



# V-DOMAINS: VH and VL



Side view of the V-DOMAINS



View from above the CDRs

Mouse (*Mus musculus*) E5.2Fv

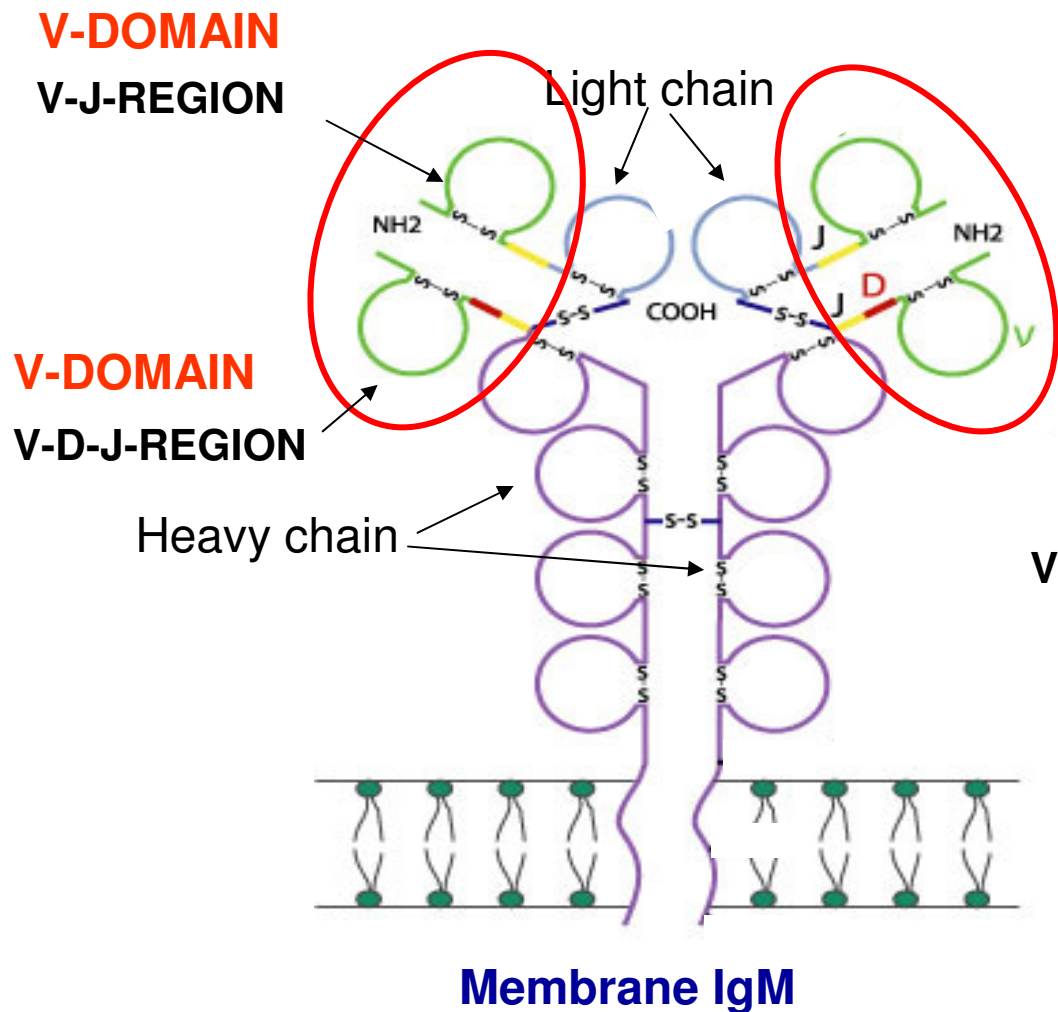
CDR3-IMGT= Complementarity determining region (105-117)

V-J junction (104-118)

V-D-J junction (104-118)

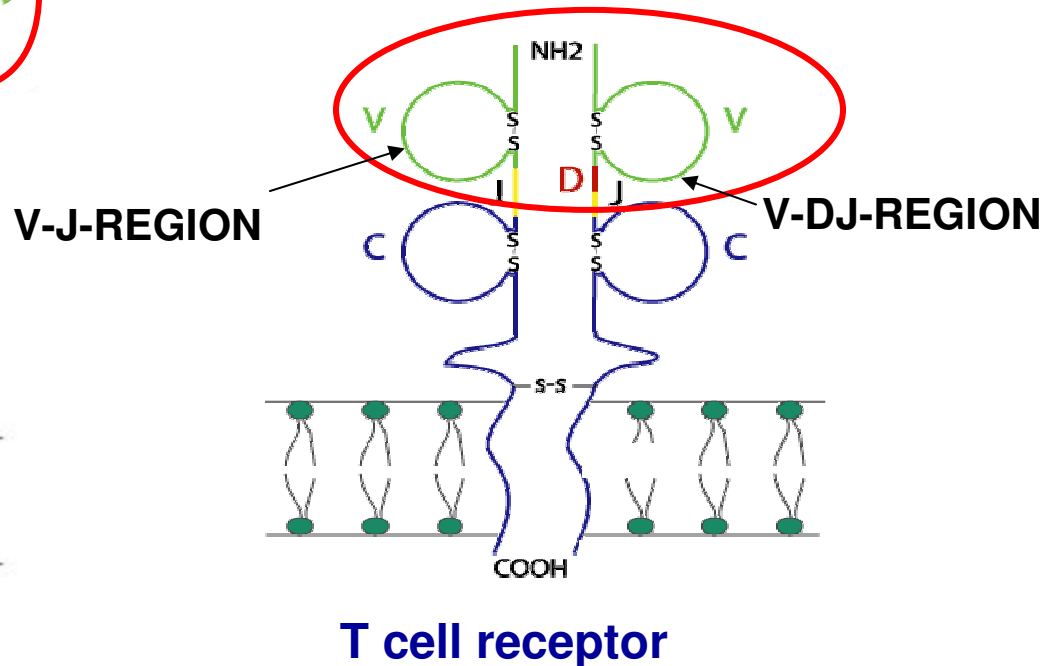
# Immunoglobulin (IG)

# T cell receptor (TR)



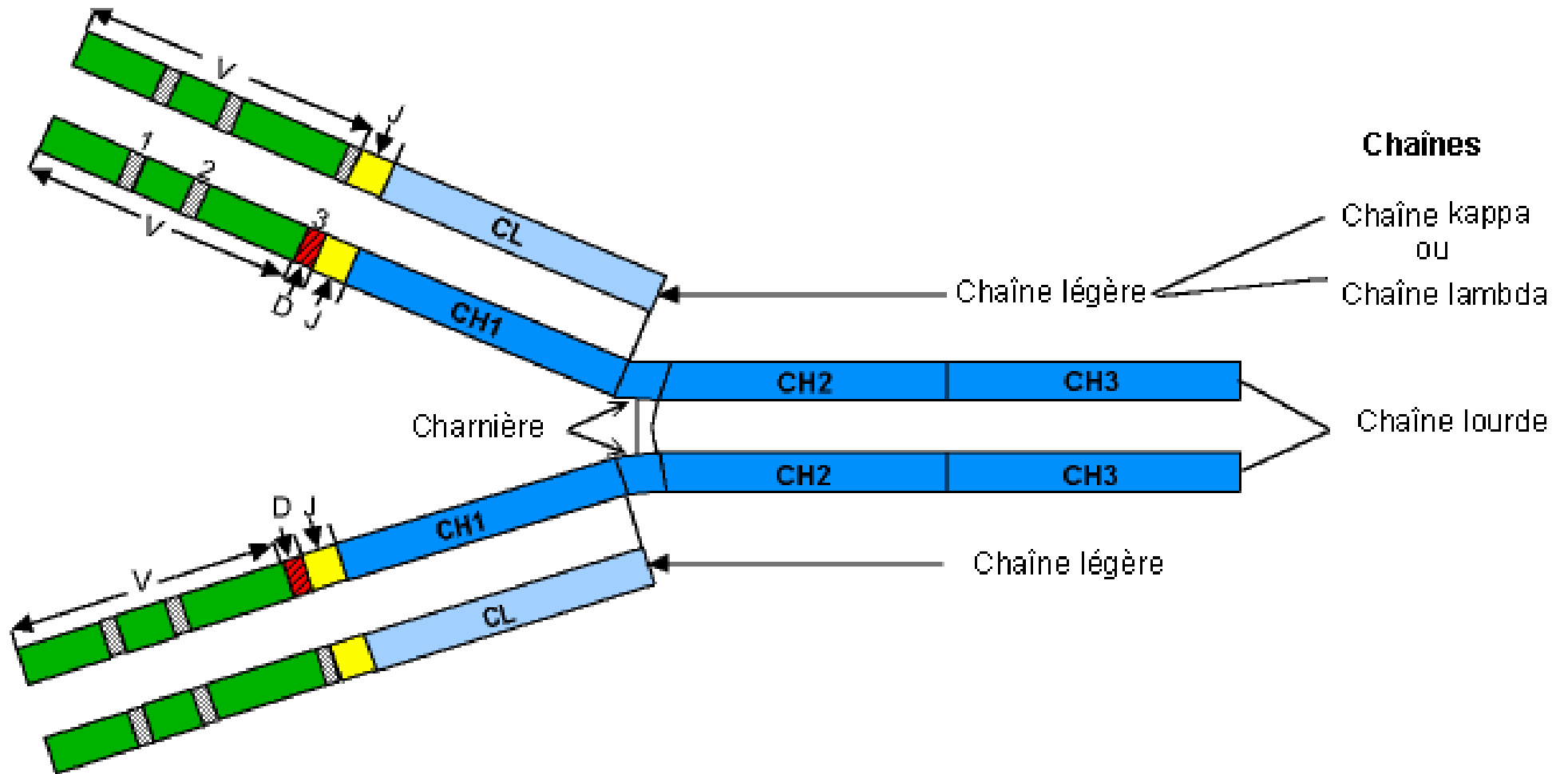
Contribution of the  
**2 V-DOMAINS**  
to the antigen binding site

Alpha - Beta  
Gamma - Delta



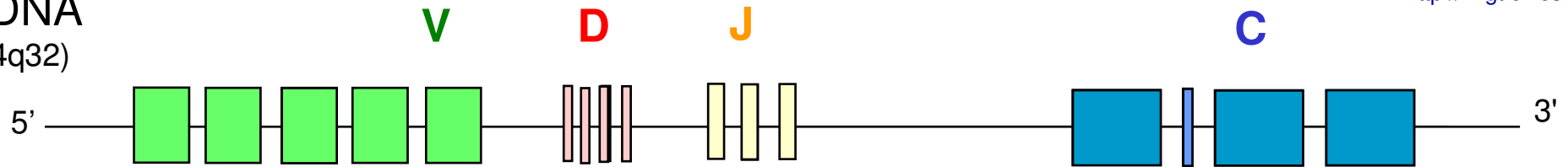


# Immunoglobulin IgG



# Immunoglobulin (IG) synthesis

genomic DNA  
(IGH Locus 14q32)

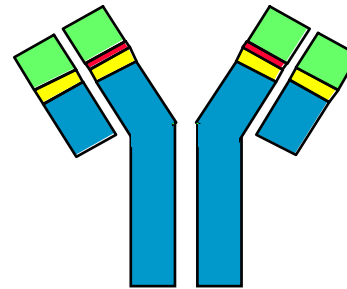


rearranged  
DNA



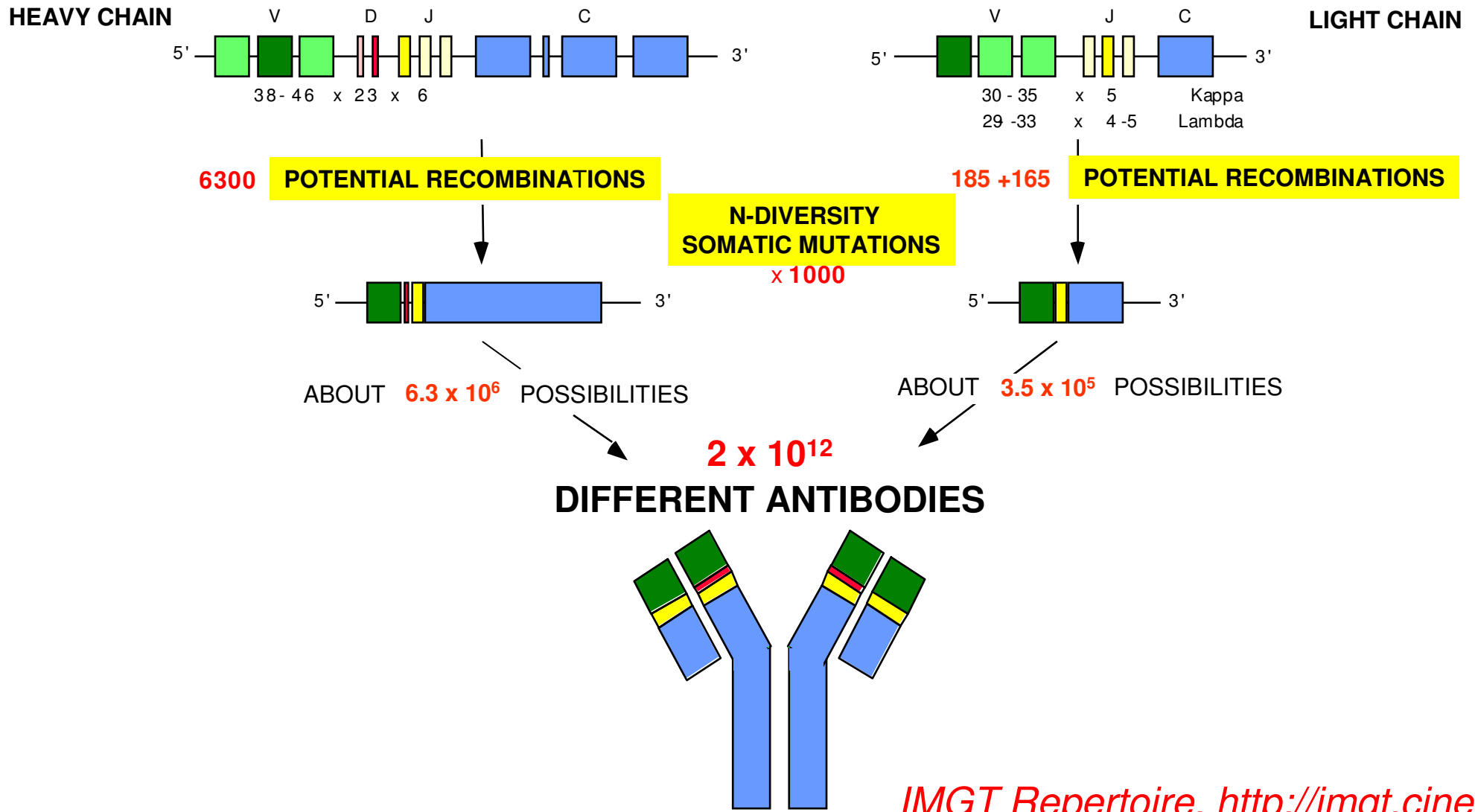
mRNA

$2 \times 10^{12}$  different IG  
per individual

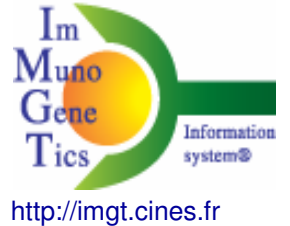


# Immunoglobulin (IG) synthesis

**150**  
**FUNCTIONAL IG GENES**



IMGT® <http://imgt.cines.fr>



IMGT®, the international ImMunoGeneTics information system®

Created in 1989 at Montpellier, France (University Montpellier 2 and CNRS)

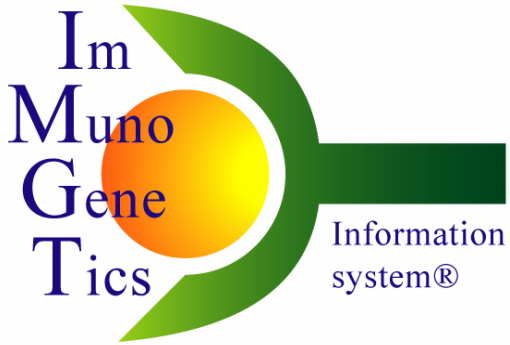
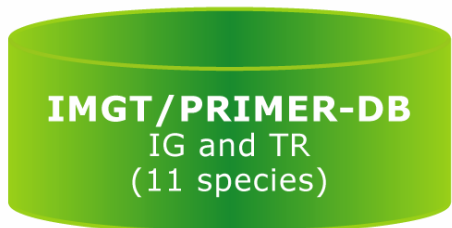
IMGT® is the international reference in immunogenetics and immunoinformatics.

IMGT® comprises:

- 6 databases
- 15 on-line tools
- more than 10,000 HTML pages of Web resources.

IMGT® receives 150.000 requests per month.

# Sequences



<http://imgt.cines.fr>

# Genome



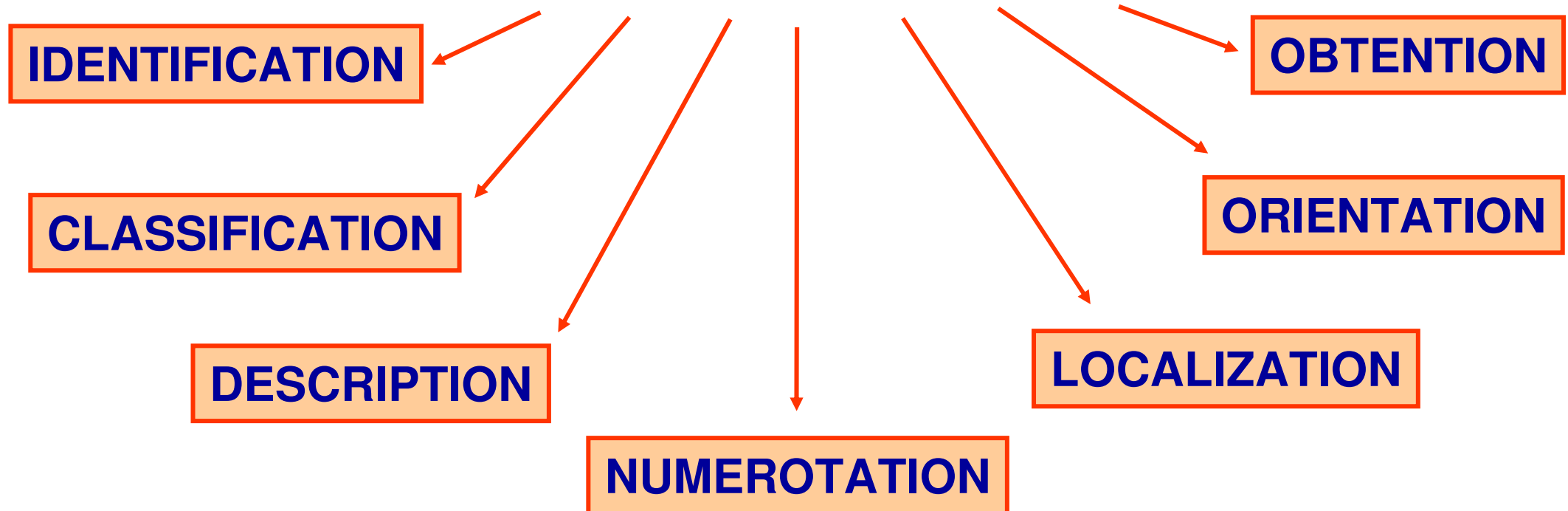
# 2D and 3D structures

Why and how has IMGT® become a paradigm towards Systems Biology?

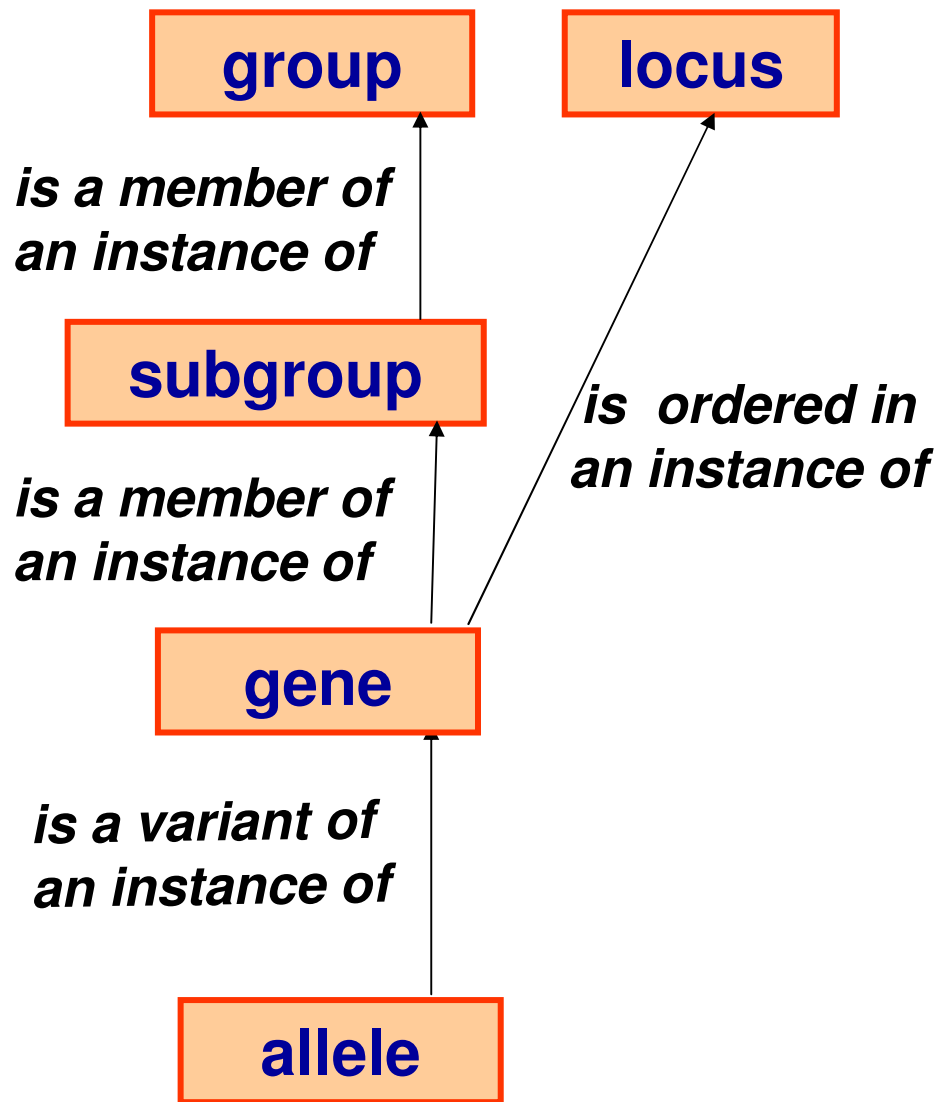
# IMGT-ONTOLOGY axioms and concepts

## IMGT-ONTOLOGY seven axioms:

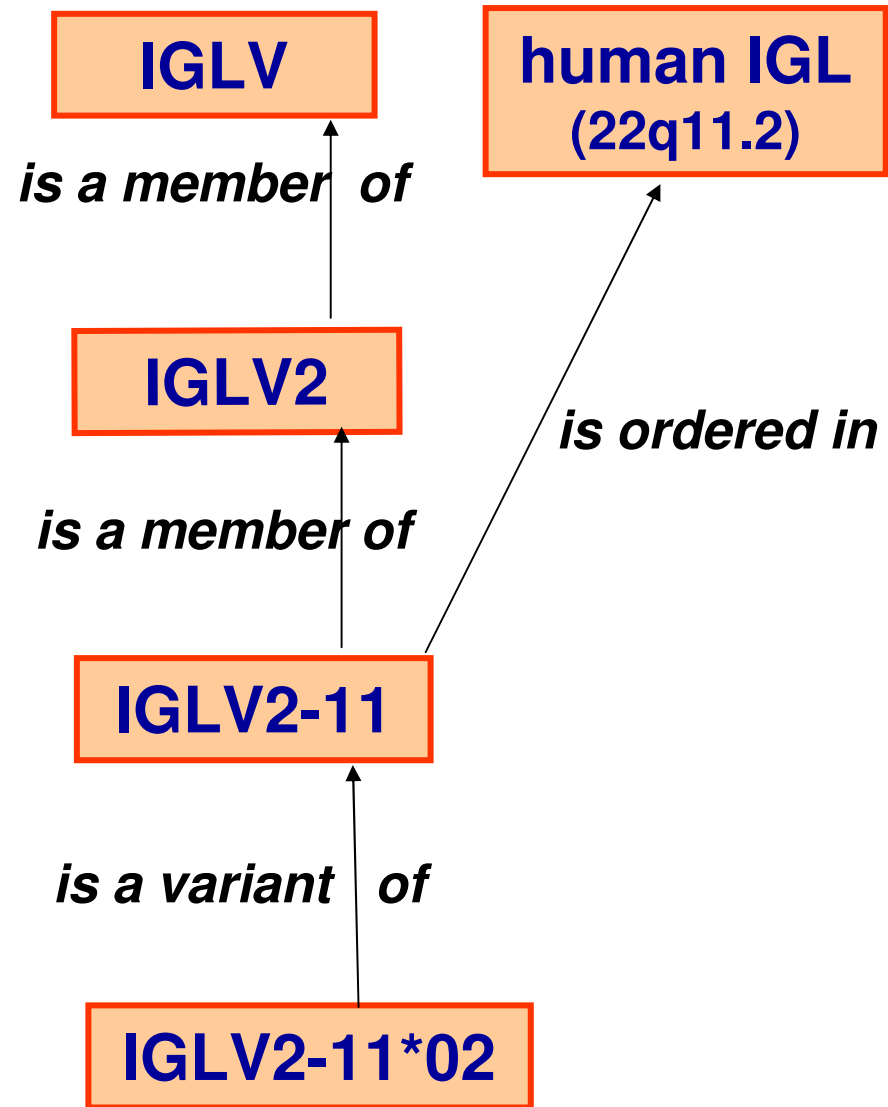
To share, reuse and represent knowledge  
in Immunogenetics and Life Sciences



# CLASSIFICATION axiom



« Concepts »



« Instances »





Locus representation: Human IGL

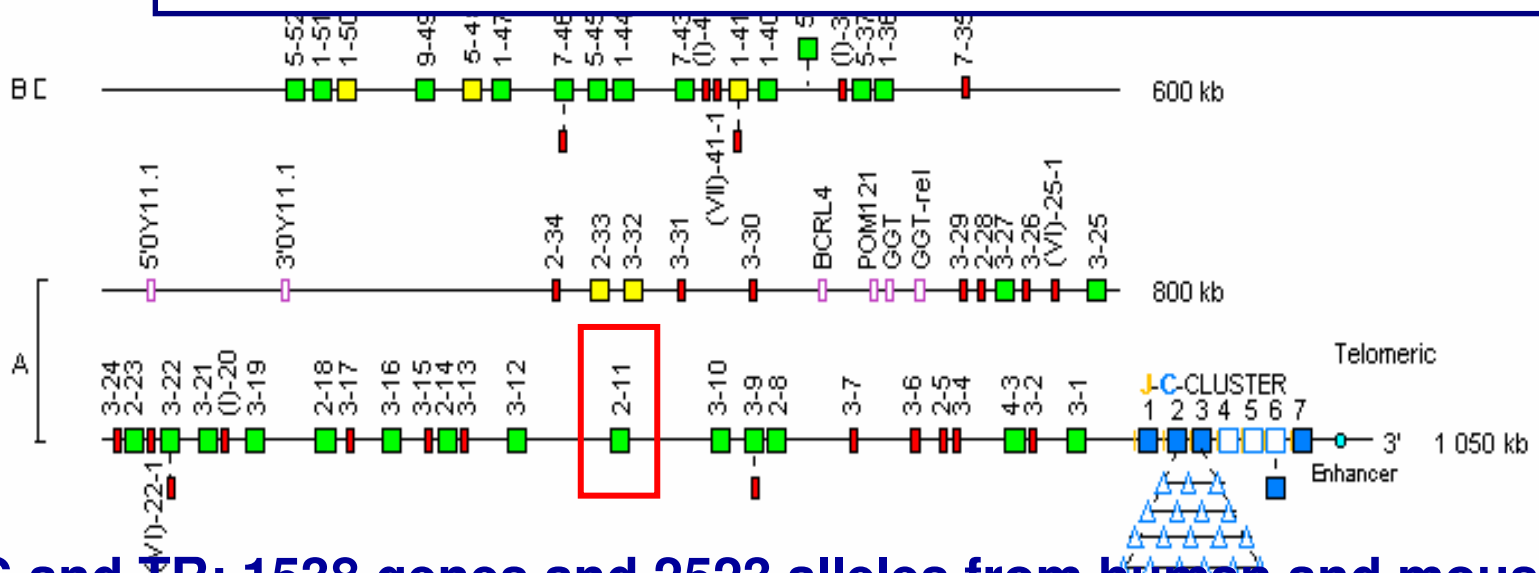
Human IGL 2

**WELCOME !**  
**to IMGT/GENE-DB**

THE  
 INTERNATIONAL  
 IMMUNOGENETICS  
 INFORMATION SYSTEM®

Information system®

<http://imgt.cines.fr>



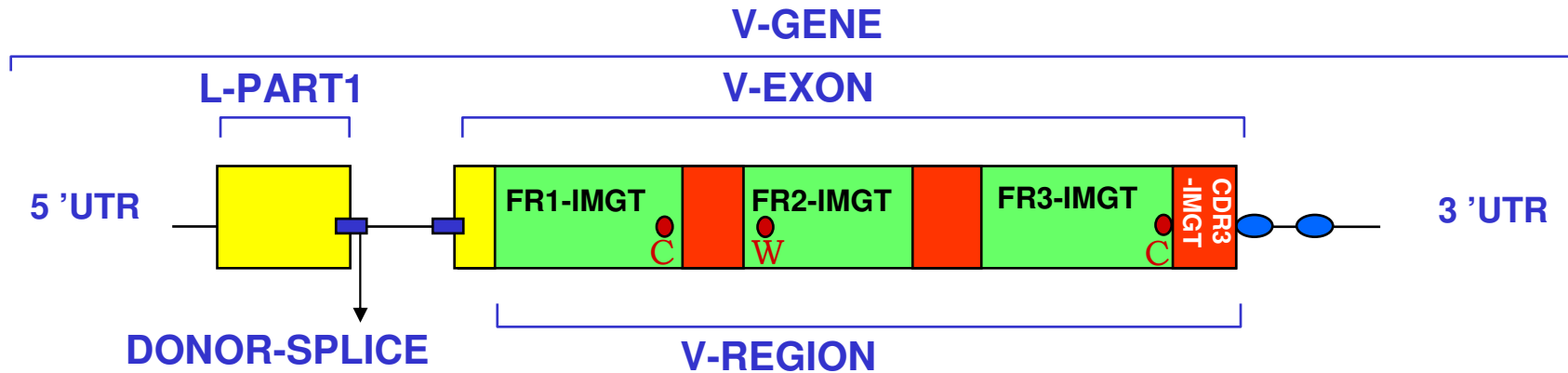
**IG and TR: 1538 genes and 2523 alleles from human and mouse**

# CLASSIFICATION axiom

- The IMGT-ONTOLOGY main concepts of classification include 'group', 'subgroup', 'gene', 'allele'.
- They allowed to set up the nomenclature for IG and TR genes (V, D, J, C genes).
- IMGT gene names were approved by HGNC in 1999 and entered in GDB, LocusLink and Entrez Gene (NCBI).
- IMGT/GENE-DB is the international reference database for IG and TR genes (direct links from Entrez Gene NCBI).
- WHO-IUIS/IMGT 2007 report (*Dev. Comp. Immunol., Immunogenetics*).

# DESCRIPTION axiom

## PROTOTYPE for a V-GENE



Label 1

Label 2

V-GENE

V-EXON

FR3-IMGT

CDR3-IMGT

L-PART1

DONOR-SPLICE

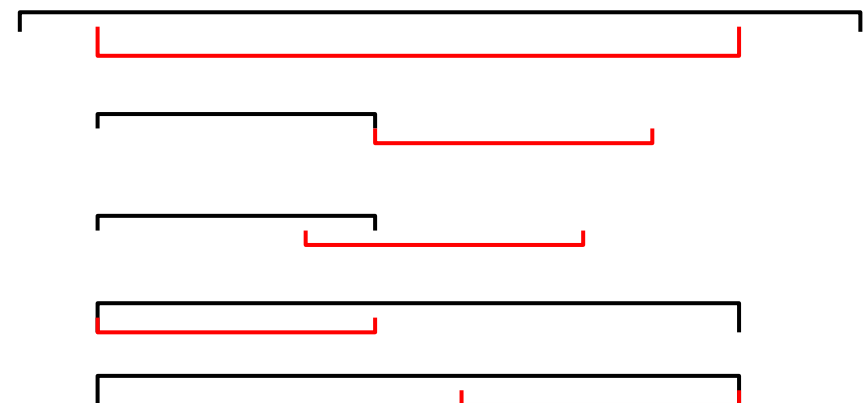
V-REGION

FR1-IMGT

V-REGION

CDR3-IMGT

Relations entre Labels

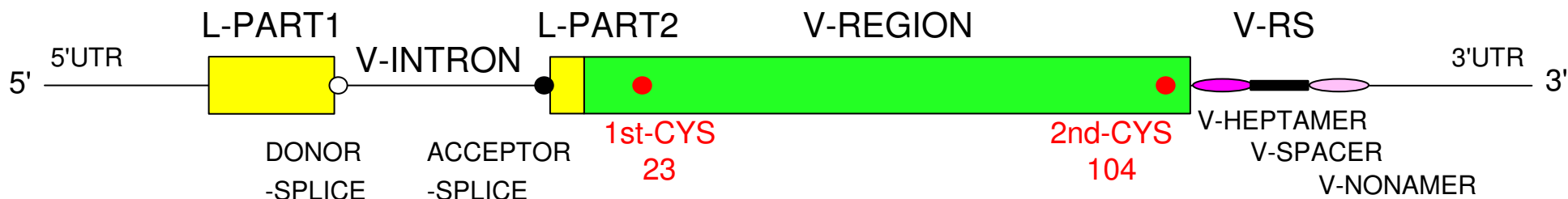


# An example of V-GENE

>X62106.0|HSVI2|*Homo sapiens* VI-2 gene for immunoglobulin heavy chain

```

tgagagctcc gttcctcacc atggactgga cctggaggat cctcttcttg gtggcagcag      60
ccacaggaa gaggctccct agtcccagtg atgagaaaga gattgagtcc agtccagggg      120
gatctcatcc acttctgtgt tctctccaca ggagccact ccaggtgca gctggtgcag      180
tctggggctg aggtgaagaa gcttggggcc tcagtgaagg tctctgcaa ggcttctgga      240
tacaccttca ccggctacta tatgcactgg gtgcgacagg ccctggaca agggcttgag      300
tggatgggat ggatcaacc taacagtggg ggcacaaact atgcacagaa gtttcagggc      360
agggtcacca tgaccagggg cacgtccatc agcacagcct acatggagct gagcaggctg      420
agatctgacg acacggccgt gtattactgt gcgagagaca cagtgtgaaa acccacatcc      480
tgagggtgtc agaaacccaa gggaggaggc ag
  
```



# IMGT/LIGM-DB



Information system®

ines.fr

**D  
E  
S  
C  
R  
I  
P  
T  
I  
O  
N**

Key	Location/Qualifiers
L-V-D-J-C-SEQUENCE	<1..375> /partial /db_xref="taxon:9606" /cell_type="B-cell hybridoma 2F7" /IMGT_note="automatically annotated with IMGT tools" /organism="Homo sapiens"
V-D-J-REGION	1..375 RLSRAASGFTFSSYGMHWVRQAP NSKNTLYLQMNSLRAEDTAVYYC
V-REGION	1..296 /allele="IGHV3-33*01, putative" /gene="IGHV3-33" /CDR_length="[8.8.18]" /putative_limit="3' side" /translation="QVHLVESGGAVFHPGRSLRLSRAASGFTFSSYGMHWVRQAP AKGLEWVAVIWDGNSKYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYC AK"
FR1-IMGT	1..75 /AA_IMGT="1 to 26, AA 10 is missing" /translation="QVHLVESGGAVFHPGRSLRLSRAAS"
CDR1-IMGT	76..99 /AA_IMGT="27 to 34" /translation="GFTFSSYG"
FR2-IMGT	100..150 /AA_IMGT="39 to 55" /translation="MHWVRQAPAKGLEWVAV"
CONSERVED-TRP	106..108
CDR2-IMGT	151..174 /AA_IMGT="56 to 63" /translation="IWDGNSNK"
FR3-IMGT	175..288 /AA_IMGT="66 to 104, AA 73 is missing" /translation="YYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYC"

**CLASSIFICATION**

123 117 sequences from 223 species

**IMGT-ONTOLOGY:**

277 IMGT labels for sequences

285 IMGT labels for 3D structures

**SO (Sequence ontology):**

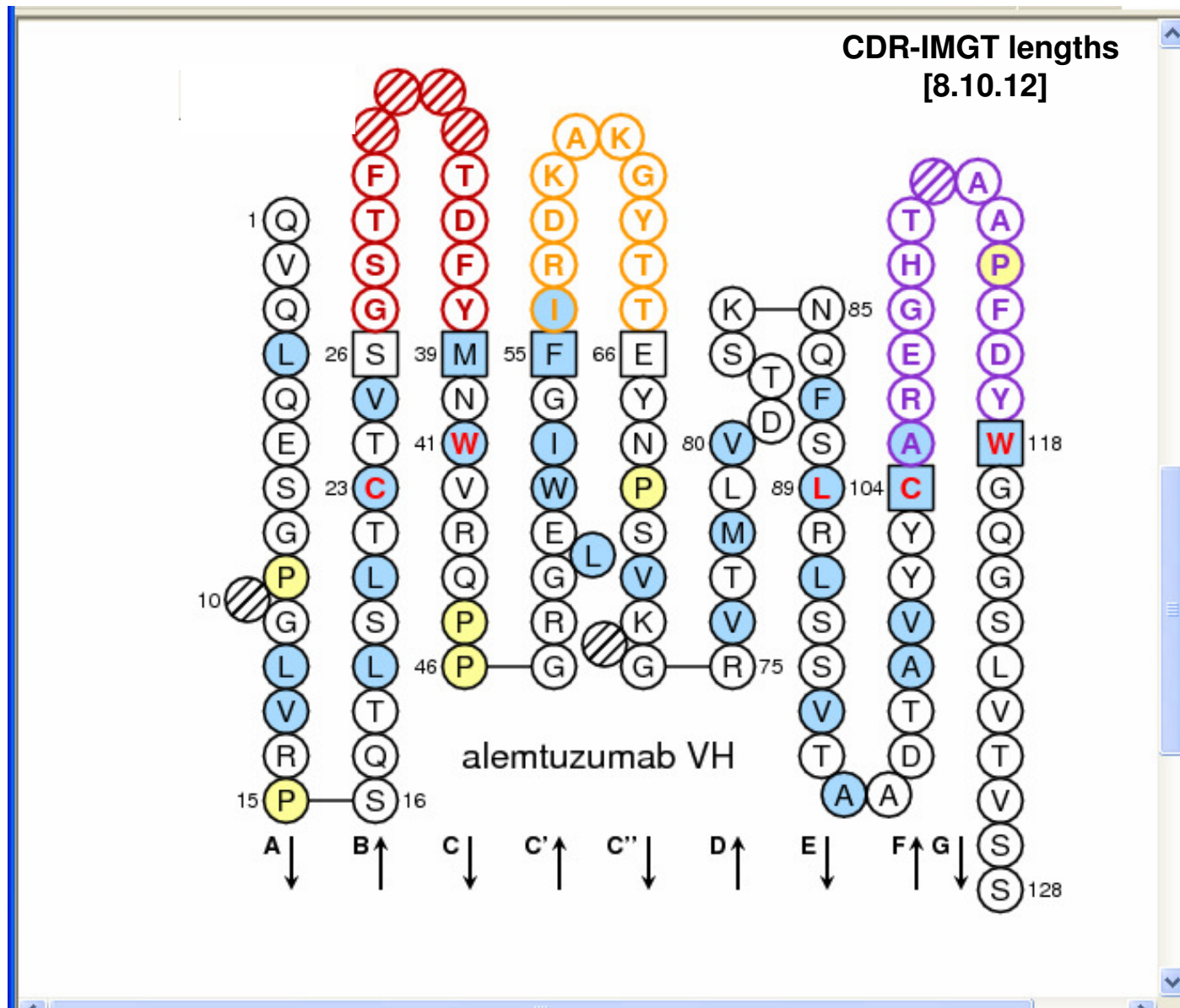
67 IMGT labels

# DESCRIPTION axiom

- The IMGT-ONTOLOGY concepts of description comprise the standardized **IMGT labels** and **relations**.
- They have allowed to describe the IG, TR and MHC sequences and 3D structures, **whatever the receptor type, the chain type, or the species**.
- They are particularly useful to describe IG, TR, and MHC and their complexes (**IG/antigen, TR/pMHC**).
- It is possible to query the IMGT® databases (**IMGT/LIGM-DB for sequences, IMGT/3Dstructure-DB for 3D structures**) with IMGT labels.
- Sequence Ontology (SO) includes IMGT labels.

# NUMEROTATION axiom

IMGT  
Collier  
de  
Perles



*Lefranc et al. Dev. Comp. Immunol. 27, 55-77 (2003)*

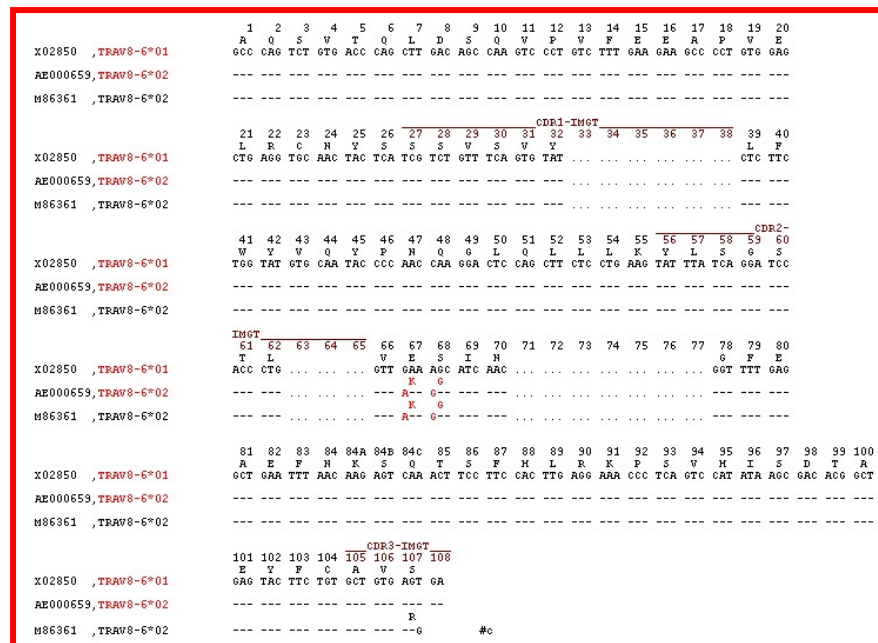
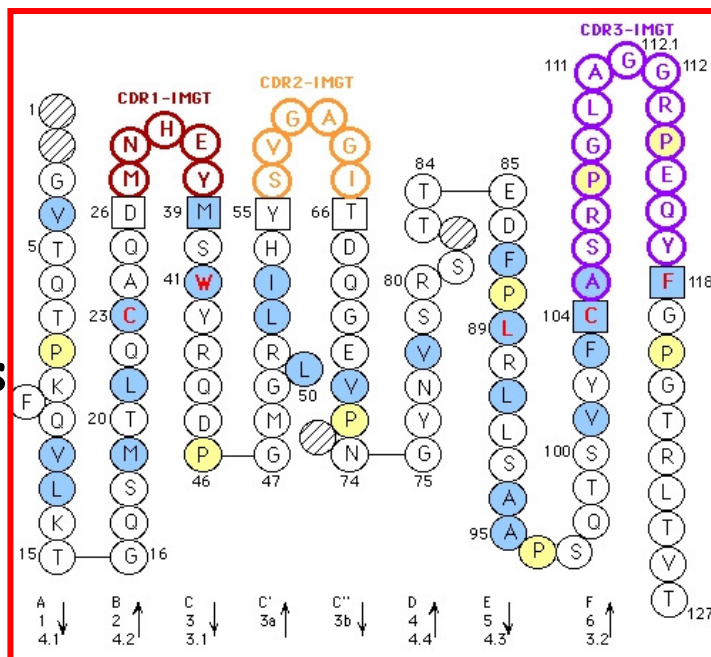


# IMGT Web resources: 10 000 pages HTML



<http://imgt.cines.fr>

IMGT  
Collier  
de Perles

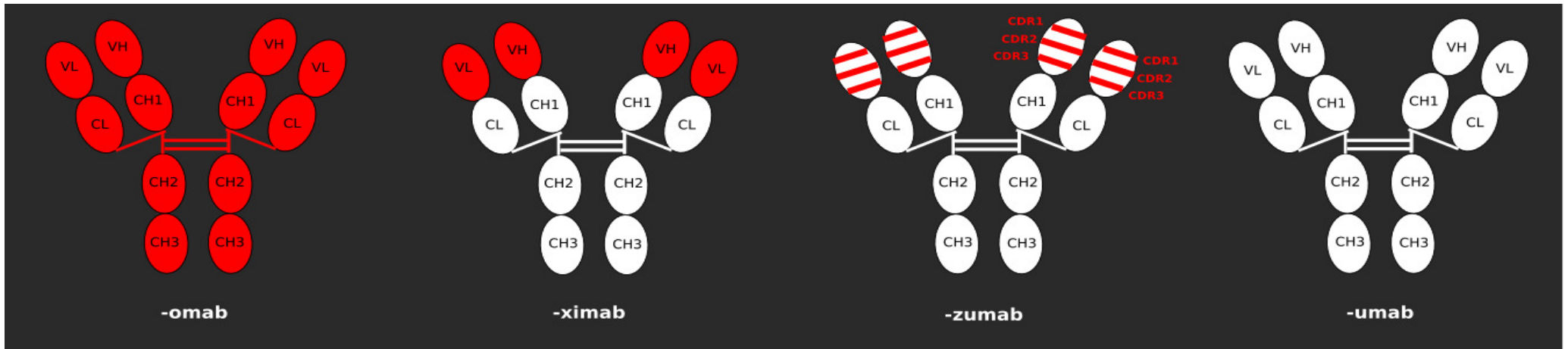


IMGT  
Alignment  
of alleles

IMGT  
Protein  
Display

TRAV gene	FR1-IMGT (1-26)	CDR1-IMGT (27-38)	FR2-IMGT (39-55)	CDR2-IMGT (56-65)	FR3-IMGT (66-104)	CDR3-IMGT (105-115)					
AE000658, TRAV1-1	GQSLEQ	PSEVTAVEGAI	VQINCTYQ	TSGFYG	LSWYQQHDGGAP	TFLSY NALDGL	LEETG	RFSSFLSRSDSYGYLLQLQMKDSASYFC	AVR		
AE000658, TRAV1-2	GQNIQD	PTEMTATEGAI	VQINCTYQ	TSGFNG	LFWYQQHAGEAPT	FLSY NVLDGL	LEEKG	RFSSFLSRSKGYSYLLKELQMKDSASYLC	AVR		
AE000658, TRAV2	KDQVFQ	PSTVASSEGAVVE	IFCNHS	VSNAVY	FFWYHLHPPG	CAPRLLVK GSK	PSQQG	RYNMTYER	FSSLLILQVREADAAVYFC	AVE	
AE000658, TRAV3	AQSVAQPEDQVNV	AEAGNPLT	VKCTYS	VSGNPY	LFWYVQYPN	RGLQFLK YITGDNL	VKGSY	GFEAEFNKSTSPHLKPPSALVSDSALYFC	AVRD		
AE000658, TRAV4	LAKITQ	PISMSDYE	QGEVNI	TCSHN	NIATNDY	ITWYQQPPSQGPRFIIQ	GYKT	KVINE	VASLFIPADRKSSTLSLPRVSLSDTAVYFC	LVGD	
AE000659, TRAV5	GEDWEQS	LFLSVREGDSS	VINCTY	DSSSTY	LYWYKQEPGAGL	QLLTY IFSNMD	MKQDQ	RLTVLLNKKDKHL	SLRIADTQTGDSAIYFC	AES	
AE000659, TRAV6	SQKIEQNSEAL	NIQEGK	TATLTCNYT	NYSPAY	LQWYRQDPGR	PVFLLL IRENEK	EKRKE	RLKVTFDITL	KQSLFHITASQPADSATYLC	ALD	
AE000659, TRAV7	ENQVEHSPHFL	GPQQGDV	ASMSCTYS	VSRFNN	LQWYRQNTGM	PKHLLS MYSAGY	EKQKG	RLNAILLK	NGSSLYITAVQPEDSATYFC	AVD	
AE000659, TRAV8-1	AQSVAQPEDQVNV	AEAGNPLT	VKCTYS	YGGTVN	LFWYVQYPN	RGLQFLK YFSGDPL	VKGIK	GFEAEF	IKSKPSPNLRKPSVQWSDTAEYFC	AVN	
AE000659, TRAV8-2	AQSVTQLD	SHSVSVEG	TPVLLRCNYS	SSYSPS	LFWYVQHPNK	GLQLLK YTSAATL	VKGIN	GFEAEFKK	SETSPHLTKPSAHMSDAAEYFC	VVS	
AE000659, TRAV8-3	AQSVTQPD	IHITVSE	GASLELRCNYS	YGATPY	LFWYVQSPG	QLLQLLK YFSGDTL	VQGIK	GFEAEFKRS	QSSFNLRKPSVHWSDAEYFC	AVG	
AE000659, TRAV8-4	AQSVTQLG	SHSVSVEG	ALVLLRCNYS	SSVPPY	LFWYVQYPN	QGLQLLK YTSAATL	VKGIN	GFEAEFKK	SETSPHLTKPSAHMSDAAEYFC	AVS	
X02850, TRAV8-6	AQSVTQLD	SQVPVFE	EAPVELRCNYS	SSVSVY	LFWYVQYPN	QGLQLLK YLSGDTL	YESIN	GFEAEFNK	SQTSFHLRKPSTHVSIDTAEYFC	AVS	
AE000660, TRAV8-7	TQSVTQLD	GHI	TVSEEA	PLELRCNYS	YSGVPS	..... DLTEATQ	VKGIK	GFEAEFKK	SETSPHLRKPSTHVSIDTAEYFC	AVGDR	
AE000659, TRAV9-1	GDSDVQTEG	QVPL	PSEAGSL	IVNCSYF	TTQVPS	LFWYVQYPG	EGDPLHLK AMKAMD	KGRNK	GFEAMV	KETTSFHLKPKSDVQESDASVYFC	ALS





## Immunogenicity

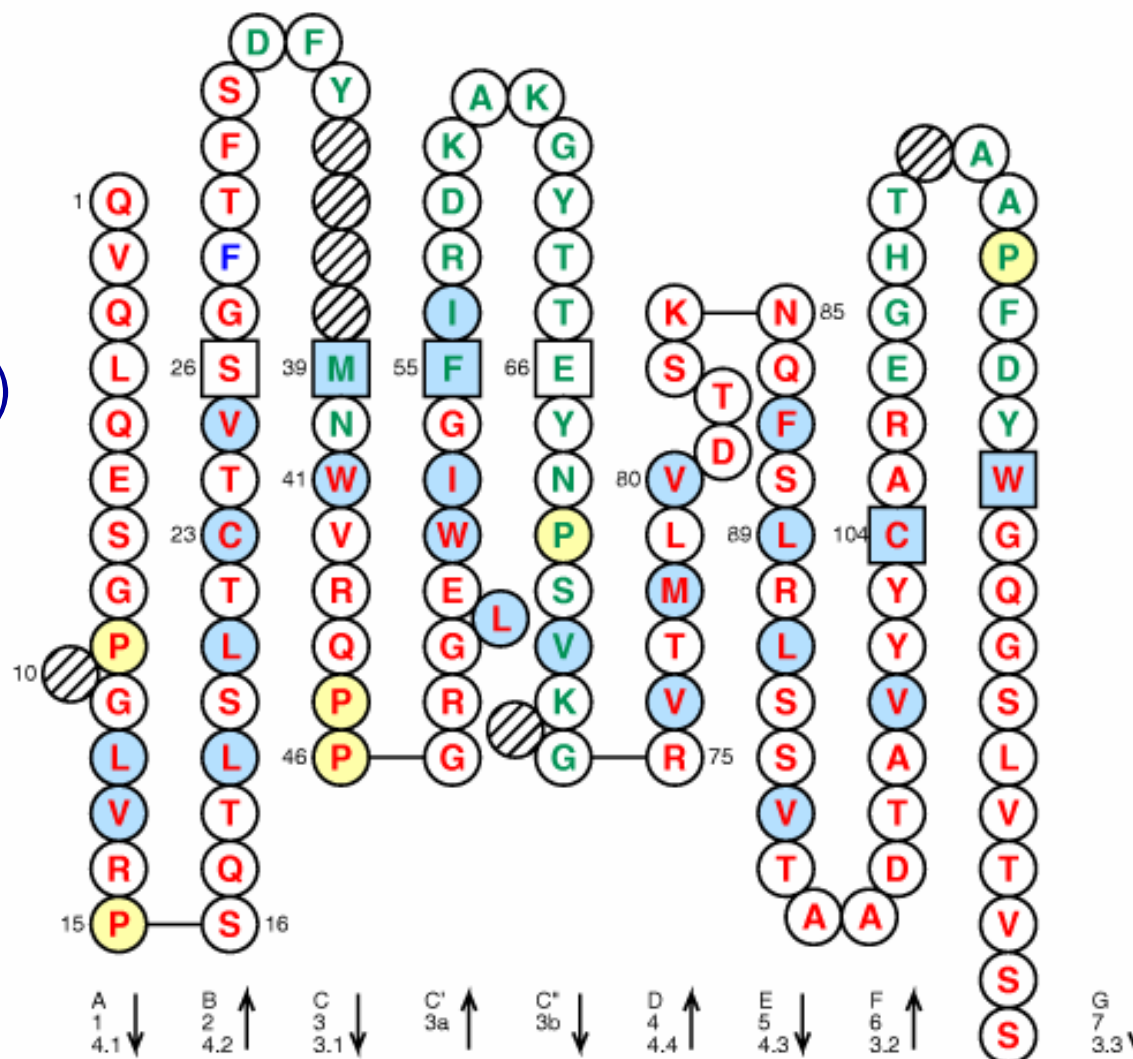
<b>-omab</b>	<b>-ximab</b>	<b>-zumab</b>	<b>-umab</b>
muromonab (1986)	abciximab (1994)	daclizumab (1997)	adalimumab (2002)
edrecolomab (1995)	rituximab (1997)	palivizumab (1998)	panitumumab (2006)
ibritumomab tiuxetan (2002)	basiliximab (1998)	trastuzumab (1998)	
tositumomab (2003)	infliximab (1998)	gemtuzumab ozogamicin (2000)	
	cetuximab (2004)	alemtuzumab (2001)	
		efalizumab (2003)	
		omalizumab (2003)	
		bevacizumab (2004)	
		natalizumab (2004)	
		nimotuzumab (2004)	
		ranibizumab (2006)	
		eculizumab (2007)	
		certolizumab pegol (2008)	

# Humanized CAMPATH-1H mutant 1

VH domain  
(V-D-J-REGION)

[8.10.12]

■ human  
■ rat



Mutant 1: **S28>F**

Mutant 2:  
alemtuzumab  
**S31>T**

# NUMEROTATION axiom

- The IMGT-ONTOLOGY concepts of numerotation include **IMGT unique numbering** and **IMGT Collier de Perles** for V-DOMAIN (IG and TR).
- They have been extended to the C-DOMAIN (IG and TR) and G-DOMAIN (MHC).
- They have allowed to bridge the gap between sequences and 3D structures in IMGT/3Dstructure-DB.
- They are used for mutations, polymorphisms, CDR-IMGT lengths, contact analysis, potential immunogenicity evaluation and paratope definition.
- WHO-INN programme requires the CDR-IMGT lengths for antibody.

# Examples of IMGT® tools based on the IMGT-ONTOLOGY concepts

**IMGT/JunctionAnalysis**  
**IMGT/V-QUEST**  
**IMGT/3Dstructure-DB**

# Immunoglobulin V-D-J generation of sequence diversity



JUNCTION

C A P Y R G D T Y D Y S W  
tgt gcg cca tac cgg ggt gac act tat gat tac tcc tgg

# IMGT/JunctionAnalysis: analysis of the IG and TR junctions

## IMGT/JunctionAnalysis Results

**Locus** IGH  
**Species** Homo sapiens  
**IMGT repertoire link** [Locus representation](#)

**Maximum number of mutations :**  
 V-REGION : 2; D-REGION : 4; J-REGION : 2  
**Deletion criterium :** Using patterns  
**Best D gene choice for a same score :** Less mutations

## Description of the JUNCTIONS

Click on mutated (underscored) nucleotid to see the original one:

Input	V name	V-REGION	P	N1	D-REGION	N2	P	J-REGION	J name	D name	Vmut	Dmut	Jmut	Ngc
#1	Z70256	<a href="#">IGHV2-26*01</a>		tgtgtgt	.....gcagcgcctggtac	ccaaatc		...actttgaccactgg	<a href="#">IGHJ4*02</a>	<a href="#">IGHD6-13*01</a>	1	2	1	5/15
#2	Z70257	<a href="#">IGHV3-7*02</a>		ggatgg	.....cagctcttatgcc	cgccc		ctactggacttcgatctctgg	<a href="#">IGHJ2*01</a>	<a href="#">IGHD2-2*01</a>	0	2	0	9/11
#3	Z70606	<a href="#">IGHV4-31*03</a>		c	.gactacg.....	cact		..atgcttttgatgtctgg	<a href="#">IGHJ3*01</a>	<a href="#">IGHD4-17*01</a>	0	0	0	3/5
#4	Z70608	<a href="#">IGHV4-39*05</a>		cagagta	....acgatttttgagtggttatt.....	ccccggggga		..atgcttttgatctctgg	<a href="#">IGHJ3*02</a>	<a href="#">IGHD3-3*01</a>	0	0	0	12/17
#5	Z70610	<a href="#">IGHV4-34*09</a>		tcgggag	.....cgatttttgagtggttatt.....	cccga	ca	tgatgcttttgatctctgg	<a href="#">IGHJ3*02</a>	<a href="#">IGHD3-3*01</a>	0	0	0	9/12
#6	Z70611	<a href="#">IGHV4-59*01</a>		ca	.....tggttaactataa.	tgccggcggtg		...actggttcgaccctgg	<a href="#">IGHJ5*02</a>	<a href="#">IGHD3-9*01</a>	0	2	0	9/13
#7	Z70613	<a href="#">IGHV4-59*01</a>			.....cagcagctggtac	ctccct		...ctttgactactgg	<a href="#">IGHJ4*02</a>	<a href="#">IGHD6-13*01</a>	0	0	0	4/6
#8	Z70614	<a href="#">IGHV4-59*01</a>		cactataa	.....ttcggggacttat.....	cccctc		.....gactactgg	<a href="#">IGHJ4*02</a>	<a href="#">IGHD3-16*01</a>	0	2	0	7/14
#9	Z70615	<a href="#">IGHV4-59*01</a>		ggctg	gtaaagaggg.....	tttcggaa		.tactggacttcgatctctgg	<a href="#">IGHJ2*01</a>	<a href="#">IGHD5-24*01</a>	0	2	0	7/13
#10	Z70616	<a href="#">IGHV4-34*01</a>		cgg	.....gtttggg.....	ttccc		...actggttcgaccctgg	<a href="#">IGHJ5*02</a>	<a href="#">IGHD3-16*01</a>	0	0	0	6/8
#11	Z70620	<a href="#">IGHV4-30-4*01</a>		cc	.....ggggcgggatggtt.....	cgg		.gatgcttttgatctctgg	<a href="#">IGHJ3*02</a>	<a href="#">IGHD3-16*01</a>	1	4	0	5/5
#12	Z70621	<a href="#">IGHV4-39*01</a>		ccacgatttatgg	.....ttcggggagtt.....	tgaccccc		.....ttgactactgg	<a href="#">IGHJ4*02</a>	<a href="#">IGHD3-16*01</a>	0	1	0	12/21
#13	Z70622	<a href="#">IGHV4-39*06</a>	t	tgccccgctcctgccaaaat	gtattactatggttcgggga.....	tatgtacg		.....tttggactactgg	<a href="#">IGHJ4*03</a>	<a href="#">IGHD3-10*01</a>	0	0	0	15/28

# The eleven IMG-T amino acid classes according to the physicochemical properties

'Volume' classes		'Hydropathy' classes						
	in Å <sup>3</sup>	Hydrophobic		Neutral		Hydrophilic		
Very large	189-228	<b>F</b>		<b>W</b>	<b>Y</b>			
Large	162-174	I	L	M		H	K R	
Medium	138-154	V					E Q	
Small	108-117			C	<b>P</b>		D N	
Very small	60-90	A			<b>G</b>			
		Aliphatic		Sulfur	Hydroxyl	Basic	Acidic	Amide
		Nonpolar			Uncharged	Charged	Uncharged	
					Polar			

# IMGT/JunctionAnalysis: analysis of the IG and TR junctions

## JUNCTION alignments with translation and IMGT AA classes

Click on mutated (underlined) amino acid to see the original one:

	104	105	106	107	108	109	110	111	111.1	111.2	111.3	112.3	112.2	112.1	112	113	114	115	116	117	118
#1 AY393054	C	S	P	G	G	S	<u>A</u>	Y						Y	<u>H</u>	E	<u>H</u>	F	Q	<u>Q</u>	W
	tgt	agt	ccc	ggg	ggt	agt	g <u>ct</u>	tat	...	...	...	...	...	tac	<u>ca</u> c	gaa	<u>ca</u> c	ttc	cag	<u>ca</u> g	tgg
#2 AY393055	C	<u>V</u>	K	P	T	D	D	D	G				<u>H</u>	R	A	E	Y	F	Q	<u>Y</u>	W
	tgt	g <u>tg</u>	aaa	ccc	acg	gat	gat	gat	ggc	...	...	...	<u>ca</u> c	cgg	gct	gaa	tac	ttc	cag	<u>ta</u> c	tgg
#3 AY393058	C	S	P	G	G	S	<u>A</u>	Y						Y	<u>H</u>	E	<u>D</u>	F	Q	<u>Q</u>	W
	tgt	agt	ccc	ggg	ggt	ag <u>c</u>	g <u>ct</u>	tat	...	...	...	...	...	tac	<u>ca</u> c	gaa	<u>ga</u> c	ttc	cag	<u>ca</u> g	tgg
#4 AY393072	C	S	P	G	G	S	<u>A</u>	Y						Y	<u>H</u>	E	<u>H</u>	F	Q	<u>Q</u>	W
	tgt	agt	ccc	ggg	ggt	agt	g <u>ct</u>	tat	...	...	...	...	...	tac	<u>ca</u> c	gaa	<u>ca</u> c	ttc	cag	<u>ca</u> g	tgg
#5 AY393088	C	A	R	Q	N	P	P	E	Y	S	G	A	Y	<u>H</u>	<u>D</u>	G	W	F	D	P	W
	tgt	gcg	aga	caa	aac	ccc	ccc	gag	tat	agt	ggc	gca	tat	<u>ca</u> t	<u>ga</u> t	ggg	tgg	ttc	gac	ccc	tgg
#6 AY393089	C	A	R	E	M	L	Y	G	S	G	<u>G</u>	Y	Y	P	P	D	A	F	<u>E</u>	<u>L</u>	W
	tgt	gcg	aga	gag	atg	ctc	tat	ggt	tcg	ggg	<u>gg</u> t	tat	tac	ccc	cct	gat	gca	ttt	<u>ga</u> g	<u>ct</u> c	tgg
#7 AY393091	C	A	R	Q	N	P	P	E	Y	S	G	A	Y	<u>H</u>	<u>D</u>	G	W	F	D	P	W
	tgt	gcg	aga	cag	aat	ccc	ccc	gag	tat	agt	ggc	gca	tat	<u>ca</u> t	<u>ga</u> t	ggg	tgg	ttc	gac	ccc	tgg
#8 AY393092	C	A	R	E	M	L	Y	G	S	G	<u>G</u>	Y	Y	P	P	D	A	F	<u>E</u>	V	W
	tgt	gcg	aga	gag	atg	ctc	tat	ggt	tcg	ggg	<u>gg</u> t	tat	tac	ccc	cct	gat	gca	ttt	<u>ga</u> g	gtc	tgg
#9 AY393094	C	A	R	Q	N	P	P	E	Y	S	G	A	Y	<u>H</u>	<u>D</u>	G	W	F	D	P	W
	tgt	gcg	aga	cag	aac	ccc	ccc	gag	tat	agt	ggc	gca	tat	<u>ca</u> t	<u>ga</u> t	ggg	tgg	ttc	gac	ccc	tgg

*Yousfi Monod et al. Bioinformatics 20, i379-i385 (2004)*  
*Pommié et al. J. Mol Recognit. 17, 17-32 (2004)*



# IMGT/V-QUEST: analysis of IG and TR sequences


## WELCOME ! to the IMGT/V-QUEST Search page

THE  
INTERNATIONAL  
IMMUNOGENETICS  
INFORMATION SYSTEM®



<http://imgt.cines.fr>

Citing IMGT/V-QUEST: Giudicelli, V. et al. Nucl. Acids Res. 2004, 32, W435-440 [PMID: 15215425](#) [PDF](#)

 You are in the new IMGT/V-QUEST, upgraded for multiple sequences and with new functionalities. **NEW!**

### Analyse your Immunoglobulin nucleotide sequences

-  [Human](#)
-  [Mouse](#)
-  [Chondrichthyes](#)
-  [Teleostei](#)
  - [Atlantic cod](#)
  - [Channel catfish](#)
  - [Rainbow trout](#)
-  [Sheep](#)

### Analyse your T cell Receptor nucleotide sequences

-  [Human](#)
-  [Mouse](#)
-  [Non-human primates](#)

## Analyse your Immunoglobulin sequences

### Your selection: Human

Your sequences are compared to the **Human IG set** from the [IMGTV-QUEST reference directory sets](#)

### Nucleotide sequences

Enter your sequence(s) in [FASTA format](#) (FASTA format is required):

Type (or copy/paste) your sequence(s) into the box below :

```
>AY393054
gctgggttttccttgttgetattttaaaaggtgtccaatgtgaggtgcagctggtggagtctgggggaggcttggtagcagccagggcgg
tccttgagactctcctgtgcagcttctggattgacctttggtgattactttatgagctggttcgccaggtccaggaagggactgga
gtgggtaggtttcattaaagagcgaacttatggtgggacaacagaatacggcgcgtctgtgaaaggcagattcatcatctcgagagatg
attccaaaagcatcgctatattgcaaatgaacagcctggaaccgaggacacagccatatattactgtagtcgggggtagtgcttat
taccacgaacacttccagcagtggggcccgggcaccatggtcaccgtctcctcagcctccaccaagggcccatcggtcttccccctggc
accctcctccaagagcacctctgggggcacagcggccctgggctgcttggtaaggactacttcccc
>AY393055
gctgggttttccttgttgetatttcaaaaggtgtccagtgtgaggtgcagctggtggagactggaggaggcttgatccagcctgggggg
tccttgagactctcctgtgcagcctctgggttaccgtcagtagcaactacatgagctgggttcgccaggtccaggaaggggctgga
```

Analysis by batches of up to 50 sequences in a single run

Or give the path access to a local file containing your sequence(s) in [FASTA format](#) (FASTA format is required):

<http://imgt.cines.fr>



Internet

# IMGT/3Dstructure-DB: analysis of the 3D structures

## THANK YOU

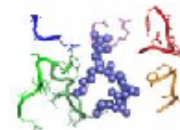
### for using [IMGT/3Dstructure-DB](#)

THE  
INTERNATIONAL  
IMMUNOGENETICS  
INFORMATION SYSTEM®



<http://imgt.cines.fr>

IMGT/3Dstructure-DB card for : **1ce1**



Entry code  Search

[Query page](#)

IMGT protein name	IMGT receptor type	IMGT receptor description	Ligand(s)	Species	CC	Chain ID
CAMPATH-1H, <a href="#">alemtuzumab</a> , MABCAMPATH®	IG	FAB-GAMMA-1_KAPPA		Humanized	1	[ <a href="#">1ce1_H</a> <a href="#">1ce1_L</a> ]
		Peptide	CD52 (synthetic peptide)	Synthetic	1	[ <a href="#">1ce1_P</a> ]

Experimental technique **X-ray diffraction**

Resolution (in angstrom) **1.90**

PDB release date **25-JUN-99**

[Epitope and Chain details](#)

[Contact analysis](#)

[3D visualization](#)  
Jmol or QuickPDB

[Renumbered IMGT file](#)

[IMGT numbering comparison](#)

[References and links](#)

[Printable card](#)

Contact analysis:

[IMGT/3Dstructure-DB Domain pair contacts](#) [\(overview\)](#) of **1ce1**

[Atom contact types](#)

Non covalent

Covalent

[Atom contact categories](#)

(BB) Backbone/backbone

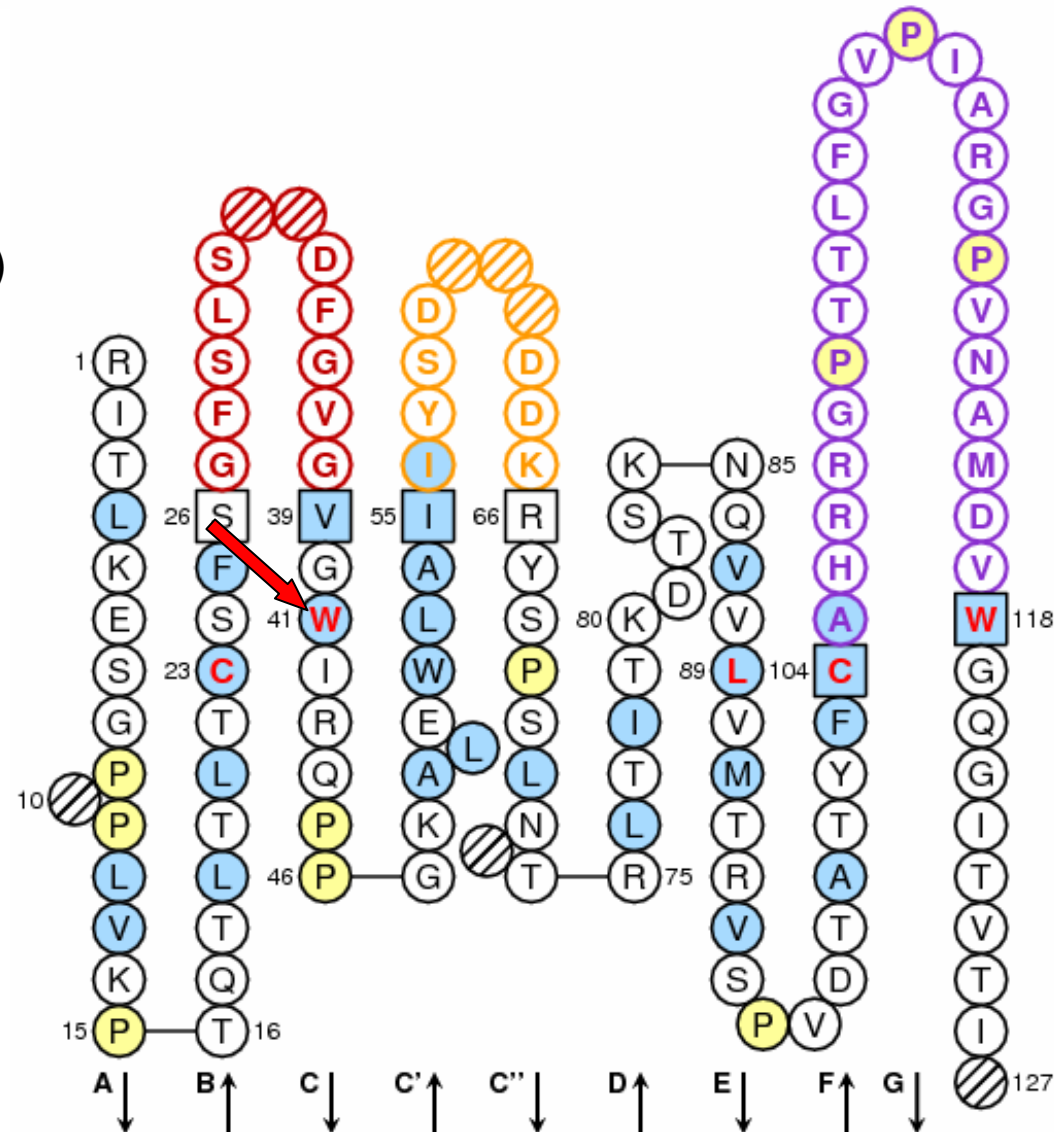
*Kaas Q. et al.*

# Access to atomic pair contacts in IMGT/3Dstructure-DB

IMGT Collier de Perles : *Homo sapiens* (Human) IGHV V-DOMAIN from 2F5 (1u8k\_B)

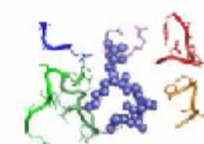
CDR-IMGT lengths [10.7.24]

Click on residue  
in IMGT Collier de Perles  
(or in amino acid sequence)



# Atomic pair contacts in IMGT/3Dstructure-DB

## IMGT Residue@Position card



Residue@Position: **41** - TRP (W) - VH - 1u8k\_B

### General information:

PDB file numbering 36  
 IMGT file numbering 41  
 Residue full name Tryptophan  
 Formula C11 H12 N2 O2

### IMGT LocalStructure@Position

Secondary structure Extended conformation  
 Phi (in degrees) -122.64  
 Psi (in degrees) 137.12  
 ASA (in square angstrom) 0.0

### Pair contacts:

#### Atom contact types

- Non covalent
- Polar
- Hydrogen bond
- Non polar

- Covalent
- Disulfide

Check all  
 Uncheck all

#### Atom contact categories

- (BB) Backbone/backbone
- (SS) Side chain/side chain
- (BS) Backbone/side chain
- (SB) Side chain/backbone

Check all  
 Uncheck all

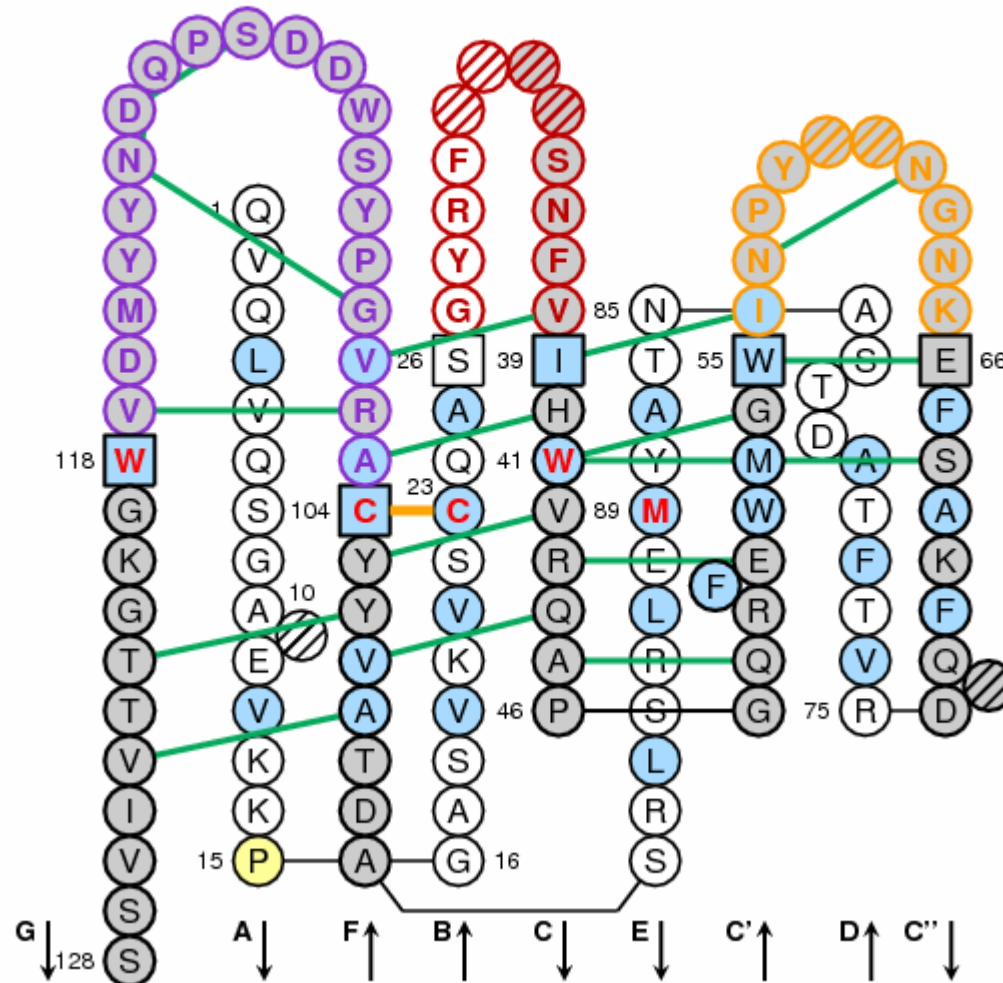
Show

IMGT Num	Residue	Domain	Chain	Atom contacts	Non Covalent	Polar	Hydrogen Bond	Non Polar
<a href="#">6</a>	GLU	E	VH	1u8k_B	6	6	0	6
<a href="#">21</a>	LEU	L	VH	1u8k_B	17	17	0	17
<a href="#">22</a>	THR	T	VH	1u8k_B	8	8	0	8
<a href="#">23</a>	CYS	C	VH	1u8k_B	10	10	0	10
<a href="#">39</a>	VAL	V	VH	1u8k_B	2	2	1	1
<a href="#">43</a>	ARG	R	VH	1u8k_B	2	2	1	1

# Hydrogen bonds (IMGT Collier de Perles on 2 layers)

IMGT Collier de Perles : *Homo sapiens* (Human) IGHV V-DOMAIN from **b12** (1hzh\_H)

CDR-IMGT lengths [8.8.20]



# Contacts VH-(Ligand), V-KAPPA-(Ligand)

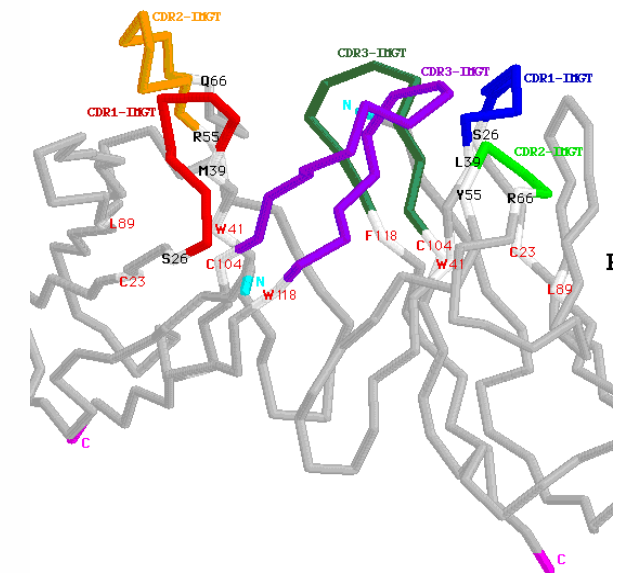
IMGT molecule name	IMGT description	Chain ID	IMGT chain description	Domain number	IMGT domain description
CAMPATH-1H, <a href="#">alemtezumab</a> , MABCAMPATH®	FAB-GAMMA-1_KAPPA	1ce1_H	VH-CH1	[D1]	VH
				[D2]	CH1
		1ce1_L	L-KAPPA	[D1]	V-KAPPA
				[D2]	C-KAPPA
CD52 (synthetic peptide)	Peptide	1ce1_P	Peptide		

	Unit 1		Unit 2		Residue contacts	Number of residues			Atom contact types		
	Domain	Chain	Domain	Chain		Total	From 1	From 2	Total	Polar	Hydrogen
<a href="#">DomPair</a>	VH	1ce1_H	CH1	1ce1_H	19	17	8	9	125	9	1
<a href="#">DomPair</a>			V-KAPPA	1ce1_L	63	45	24	21	532	61	6
<a href="#">DomPair</a>			(Ligand)	1ce1_P	25	19	12	7	216	40	9
<a href="#">DomPair</a>	CH1	1ce1_H	VH	1ce1_H	19	17	9	8	125	9	1
<a href="#">DomPair</a>			C-KAPPA	1ce1_L	68	58	28	30	498	40	6
<a href="#">DomPair</a>	V-KAPPA	1ce1_L	VH	1ce1_H	63	45	21	24	532	61	6
<a href="#">DomPair</a>			C-KAPPA	1ce1_L	18	18	8	10	137	19	2
<a href="#">DomPair</a>			(Ligand)	1ce1_P	16	14	7	7	171	37	5
<a href="#">DomPair</a>	C-KAPPA	1ce1_L	CH1	1ce1_H	68	58	30	28	498	40	6
<a href="#">DomPair</a>			V-KAPPA	1ce1_L	18	18	10	8	137	19	2



# Contacts VH-(Ligand)

	IMGT Num	Residue	Domain	Chain		IMGT Num	Residue	Domain	Chain	Total	Polar	Hydrogen	
<a href="#">R@P</a>	38	TYR	Y	VH	1ce1_H	<a href="#">R@P</a>	2	THR	T	1ce1_P	4	0	0
<a href="#">R@P</a>	38	TYR	Y	VH	1ce1_H	<a href="#">R@P</a>	7	ALA	A	1ce1_P	13	1	0
<a href="#">R@P</a>	38	TYR	Y	VH	1ce1_H	<a href="#">R@P</a>	8	ASP	D	1ce1_P	14	2	2
<a href="#">R@P</a>	55	PHE	F	VH	1ce1_H	<a href="#">R@P</a>	6	SER	S	1ce1_P	5	0	0
<a href="#">R@P</a>	55	PHE	F	VH	1ce1_H	<a href="#">R@P</a>	7	ALA	A	1ce1_P	16	0	0
<a href="#">R@P</a>	55	PHE	F	VH	1ce1_H	<a href="#">R@P</a>	8	ASP	D	1ce1_P	1	0	0
<a href="#">R@P</a>	57	ARG	R	VH	1ce1_H	<a href="#">R@P</a>	7	ALA	A	1ce1_P	9	3	2
<a href="#">R@P</a>	57	ARG	R	VH	1ce1_H	<a href="#">R@P</a>	8	ASP	D	1ce1_P	20	6	1
<a href="#">R@P</a>	61	LYS	K	VH	1ce1_H	<a href="#">R@P</a>	8	ASP	D	1ce1_P	11	2	1
<a href="#">R@P</a>	66	GLU	E	VH	1ce1_H	<a href="#">R@P</a>	7	ALA	A	1ce1_P	1	0	0
<a href="#">R@P</a>	107	GLU	E	VH	1ce1_H	<a href="#">R@P</a>	2	THR	T	1ce1_P	13	2	1
<a href="#">R@P</a>	107	GLU	E	VH	1ce1_H	<a href="#">R@P</a>	4	SER	S	1ce1_P	5	2	0
<a href="#">R@P</a>	107	GLU	E	VH	1ce1_H	<a href="#">R@P</a>	7	ALA	A	1ce1_P	5	0	0
<a href="#">R@P</a>	108	GLY	G	VH	1ce1_H	<a href="#">R@P</a>	1	GLY	G	1ce1_P	2	1	0
<a href="#">R@P</a>	108	GLY	G	VH	1ce1_H	<a href="#">R@P</a>	2	THR	T	1ce1_P	9	2	0
<a href="#">R@P</a>	109	HIS	H	VH	1ce1_H	<a href="#">R@P</a>	1	GLY	G	1ce1_P	24	4	0
<a href="#">R@P</a>	109	HIS	H	VH	1ce1_H	<a href="#">R@P</a>	2	THR	T	1ce1_P	21	5	0
<a href="#">R@P</a>	109	HIS	H	VH	1ce1_H	<a href="#">R@P</a>	3	SER	S	1ce1_P	9	2	1
<a href="#">R@P</a>	110	THR	T	VH	1ce1_H	<a href="#">R@P</a>	1	GLY	G	1ce1_P	1	1	0
<a href="#">R@P</a>	110	THR	T	VH	1ce1_H	<a href="#">R@P</a>	3	SER	S	1ce1_P	11	4	1
<a href="#">R@P</a>	112	ALA	A	VH	1ce1_H	<a href="#">R@P</a>	3	SER	S	1ce1_P	3	1	0
<a href="#">R@P</a>	113	ALA	A	VH	1ce1_H	<a href="#">R@P</a>	2	THR	T	1ce1_P	3	0	0
<a href="#">R@P</a>	113	ALA	A	VH	1ce1_H	<a href="#">R@P</a>	3	SER	S	1ce1_P	7	2	0
<a href="#">R@P</a>	113	ALA	A	VH	1ce1_H	<a href="#">R@P</a>	4	SER	S	1ce1_P	4	0	0
<a href="#">R@P</a>	114	PRO	P	VH	1ce1_H	<a href="#">R@P</a>	4	SER	S	1ce1_P	5	0	0





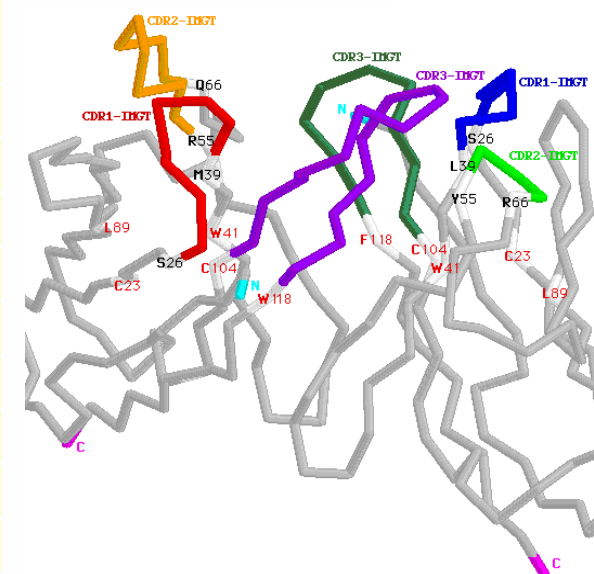
Summary:

Residue contacts	Number of residues			Atom contact types		
	Total	From 1	From 2	Total	Polar	Hydrogen
16	14	7	7	171	37	5

List of the Residue@Position pair contacts:

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

R@P	Order	Residue	Domain	Chain	R@P	Order	Residue	Domain	Chain	Atom contacts			
	IMGT Num					IMGT Num				Total	Polar	Hydrogen	
<a href="#">R@P</a>	<a href="#">38</a>	TYR	Y	V-KAPPA	1ce1_L	<a href="#">R@P</a>	3	SER	S	1ce1_P	1	0	0
<a href="#">R@P</a>	<a href="#">38</a>	TYR	Y	V-KAPPA	1ce1_L	<a href="#">R@P</a>	5	PRO	P	1ce1_P	21	0	0
<a href="#">R@P</a>	<a href="#">56</a>	ASN	N	V-KAPPA	1ce1_L	<a href="#">R@P</a>	3	SER	S	1ce1_P	3	2	0
<a href="#">R@P</a>	<a href="#">107</a>	HIS	H	V-KAPPA	1ce1_L	<a href="#">R@P</a>	4	SER	S	1ce1_P	20	4	1
<a href="#">R@P</a>	<a href="#">107</a>	HIS	H	V-KAPPA	1ce1_L	<a href="#">R@P</a>	5	PRO	P	1ce1_P	12	2	0
<a href="#">R@P</a>	<a href="#">107</a>	HIS	H	V-KAPPA	1ce1_L	<a href="#">R@P</a>	6	SER	S	1ce1_P	14	3	1
<a href="#">R@P</a>	<a href="#">108</a>	ILE	I	V-KAPPA	1ce1_L	<a href="#">R@P</a>	5	PRO	P	1ce1_P	12	1	0
<a href="#">R@P</a>	<a href="#">108</a>	ILE	I	V-KAPPA	1ce1_L	<a href="#">R@P</a>	6	SER	S	1ce1_P	12	3	0
<a href="#">R@P</a>	<a href="#">109</a>	SER	S	V-KAPPA	1ce1_L	<a href="#">R@P</a>	6	SER	S	1ce1_P	11	2	0
<a href="#">R@P</a>	<a href="#">114</a>	ARG	R	V-KAPPA	1ce1_L	<a href="#">R@P</a>	6	SER	S	1ce1_P	18	3	1
<a href="#">R@P</a>	<a href="#">114</a>	ARG	R	V-KAPPA	1ce1_L	<a href="#">R@P</a>	7	ALA	A	1ce1_P	4	2	0
<a href="#">R@P</a>	<a href="#">114</a>	ARG	R	V-KAPPA	1ce1_L	<a href="#">R@P</a>	8	ASP	D	1ce1_P	6	2	0
<a href="#">R@P</a>	<a href="#">116</a>	ARG	R	V-KAPPA	1ce1_L	<a href="#">R@P</a>	2	THR	T	1ce1_P	1	1	0
<a href="#">R@P</a>	<a href="#">116</a>	ARG	R	V-KAPPA	1ce1_L	<a href="#">R@P</a>	4	SER	S	1ce1_P	9	4	1
<a href="#">R@P</a>	<a href="#">116</a>	ARG	R	V-KAPPA	1ce1_L	<a href="#">R@P</a>	6	SER	S	1ce1_P	20	6	1
<a href="#">R@P</a>	<a href="#">116</a>	ARG	R	V-KAPPA	1ce1_L	<a href="#">R@P</a>	7	ALA	A	1ce1_P	7	2	0



# CONCLUSIONS and PERSPECTIVES



<http://imgt.cines.fr>

1. The **IMGT-ONTOLOGY** axioms and concepts: **CLASSIFICATION** (nomenclature), **DESCRIPTION** (labels), **NUMEROTATION** (IMGT unique numbering, IMGT Colliers de Perles)...are acknowledged as the international standards in immunogenetics and immunoinformatics.
2. The WHO-INN programme requires the CDR-IMGT lengths.
3. American and European companies (Centocor Johnson and Johnson USA, Merck USA,..) have adopted the IMGT® tools for **antibody engineering** and **antibody humanization**.
4. The IMGT-ONTOLOGY axioms are used for a multiscale and systemic approach (**system immunobiology**). Concepts are currently described at the **cell level** (**EU ImmunoGrid IST projet**).





**Many thanks to the IMGT® team at Montpellier, France**