IMGT®, the international ImMunoGeneTics information system®: from sequence to molecular interactions

Patrice Duroux

IMGT, IGH, CNRS UPR 1142

4th Indo French Bioinformatics Meeting
November 19-21, 2007, Montpellier, France

Main objectives

1.Provide data and tools for immunoglobulins (IG) and T cell receptors (TR) sequences

=> data curation and annotation

=> specific bioinformatics development (databases, tools)

2.Ensure quality and consistancy with respect to a shared biological model as a standardize formal background

=> IMGT-ONTOLOGY

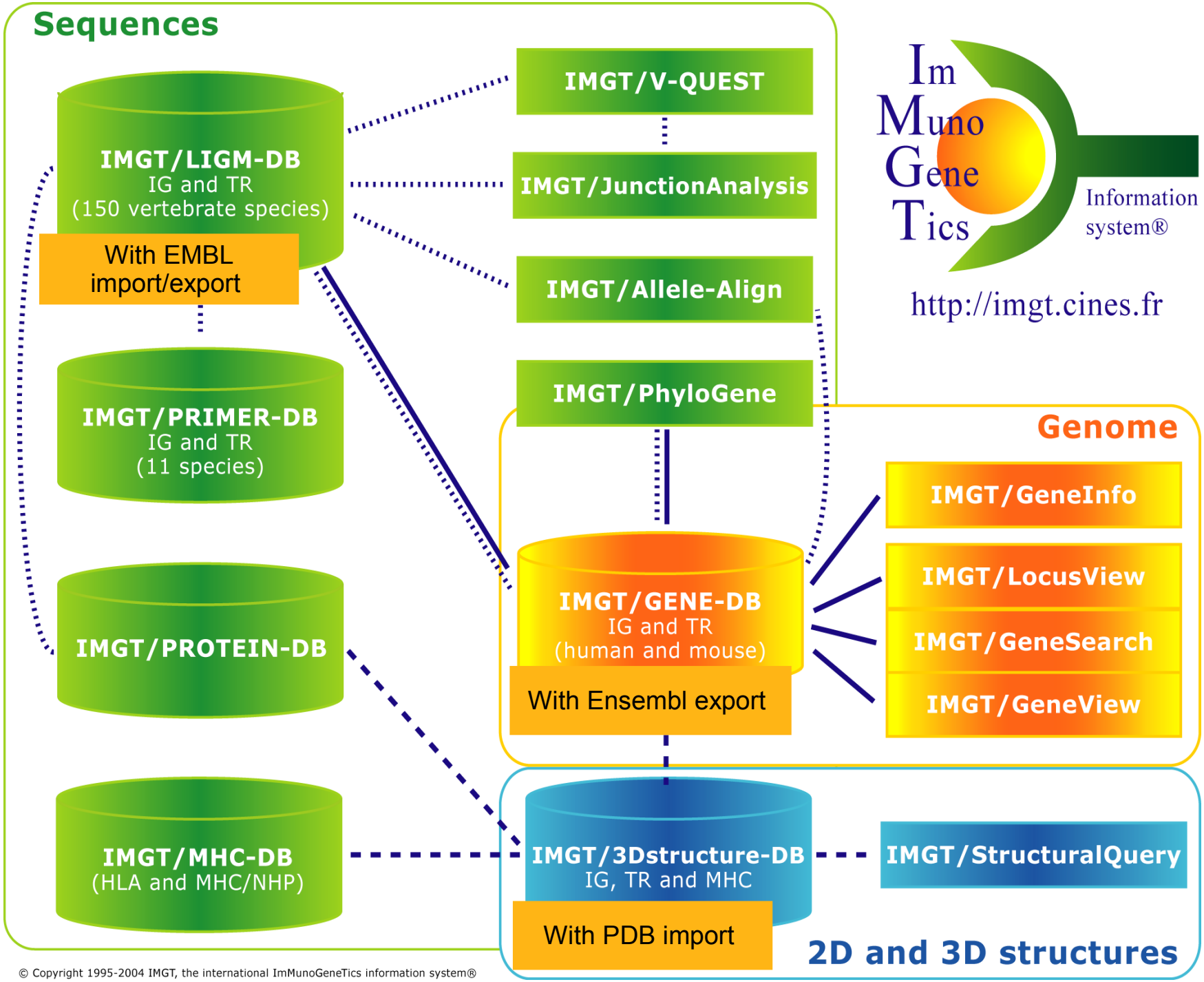
3.Reinforce relevance by integrating knowledge from different approaches

=> genetic,

=> genomic,

=> and more recently structure.

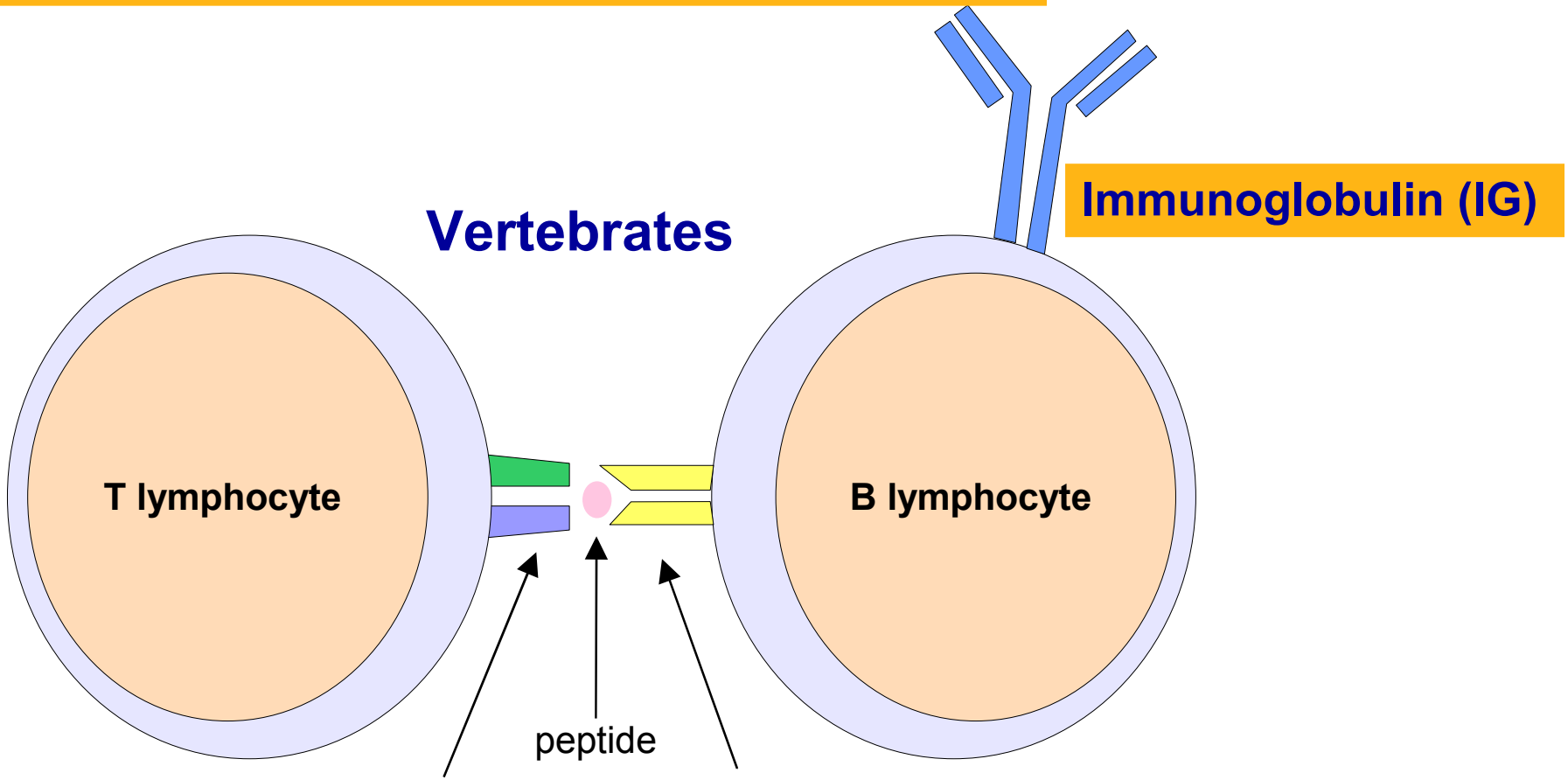
A story started in 1989!



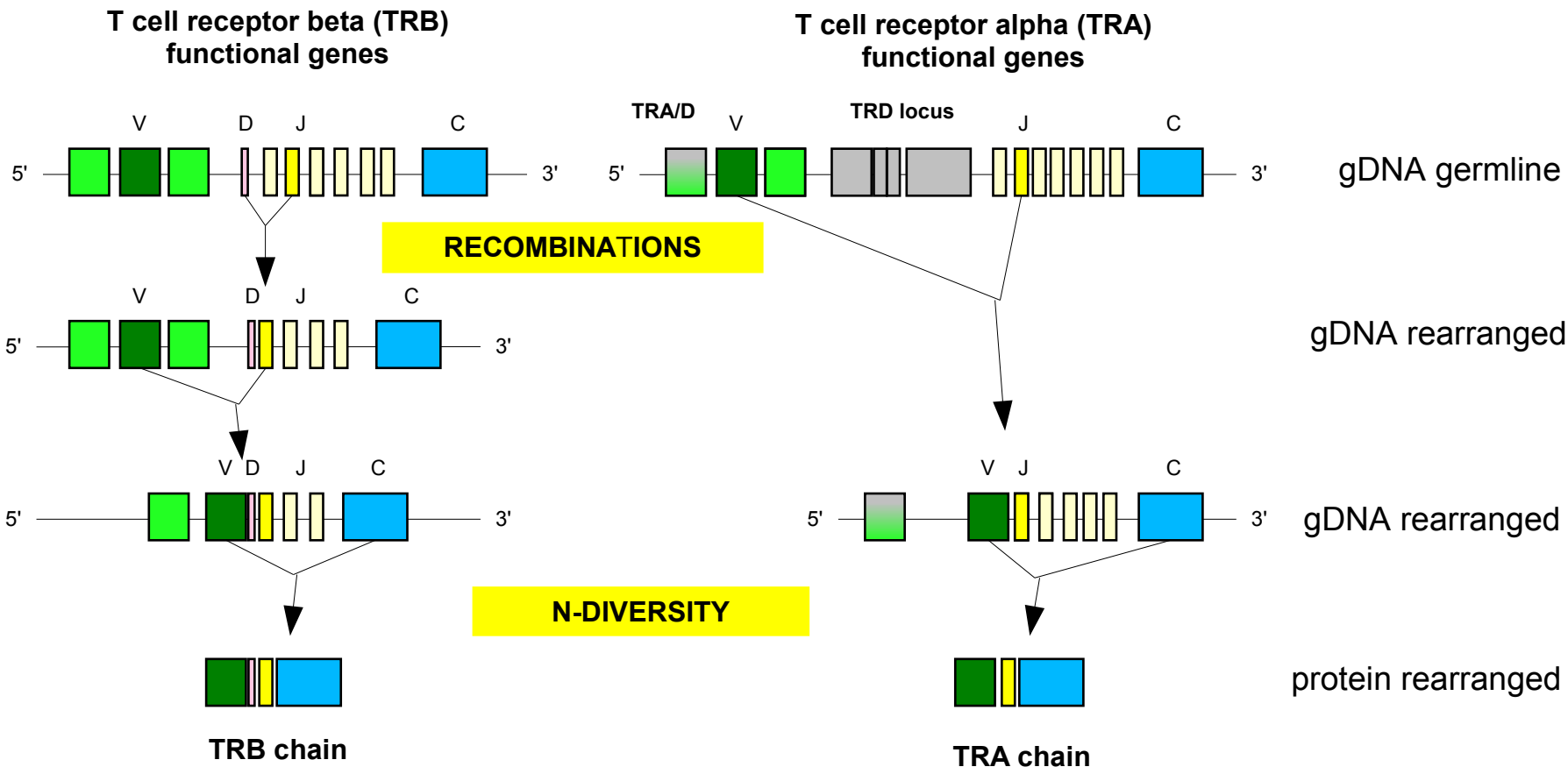
Adaptative immune response

Proteins which specifically recognize foreign antigens:
- in a native form for IG
- as processed as a peptide and presented by MHC for TR

Function of IG and TR is to bind specifically



T cell receptor diversity synthesis



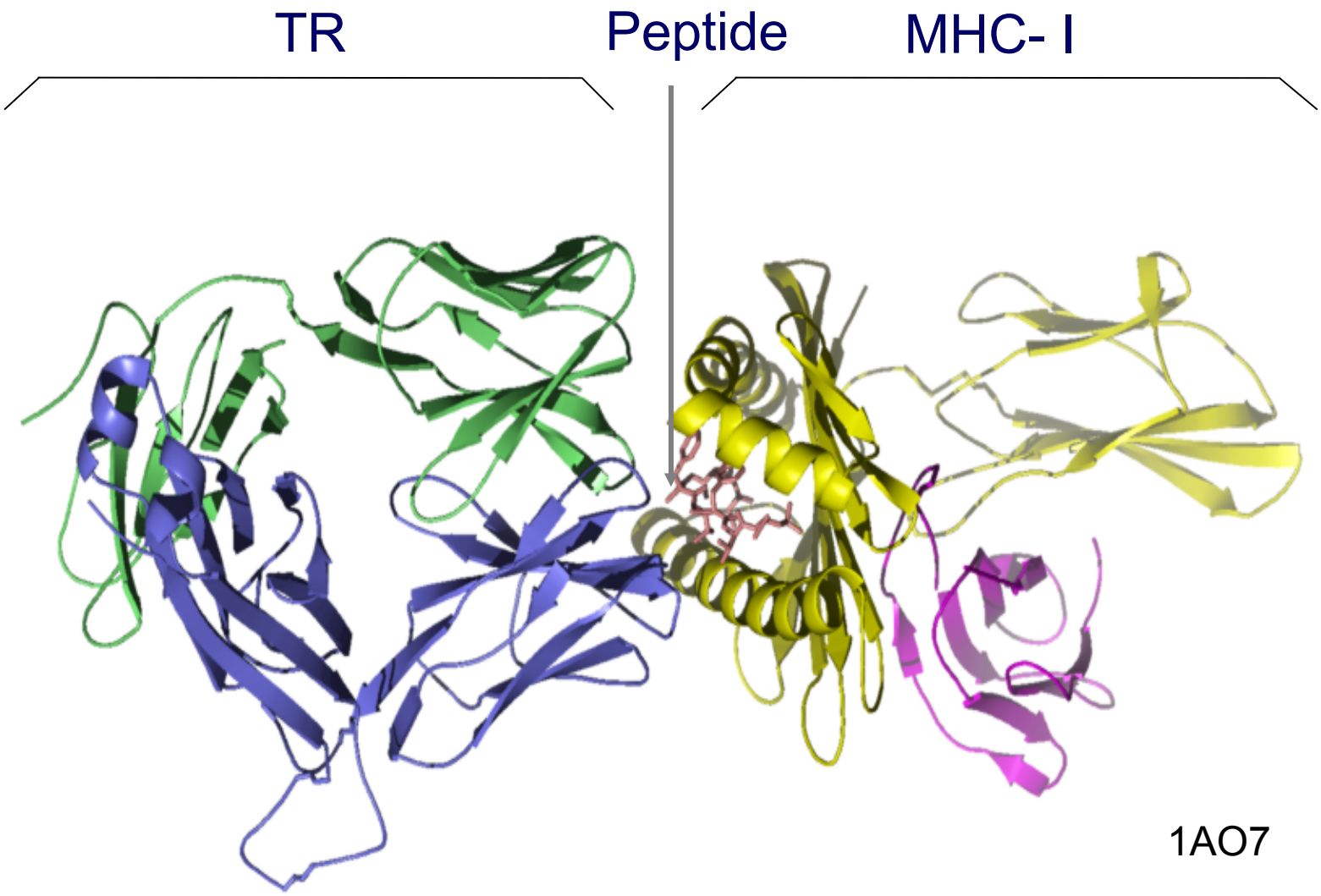
IMGT-ONTOLOGY provides:
- identification (molecule type, configuration type, chain type)
- description (prototypes, labels)
- gene and allele nomeclature by their classification (groups, subgroups, gene name, allele number, locus)

Human TR potential repertoire

	Sequences	Genes				Alleles			
		F	ORF	P	Total	F	ORF	P	Total
TRA	1511								
TRAV		43(+2)	1	8(+2)	54	101	1	10	112
TRAJ		50	8	3	61	57	8	3	68
TRAC		1			1	1			1
Total					116				181
TRB	5454								
TRBV		40-42(+6)	6(+1)	12-13(+5)	64-67	114	9	23	146
TRBD		2			2	3			3
TRBJ		12(+1)	1(+1)		14	14	2		16
TRBC		2			2	4			4
Total					82-85				169

So potentialy:
2150-2250 TRA chains
960-1248 TRB chains

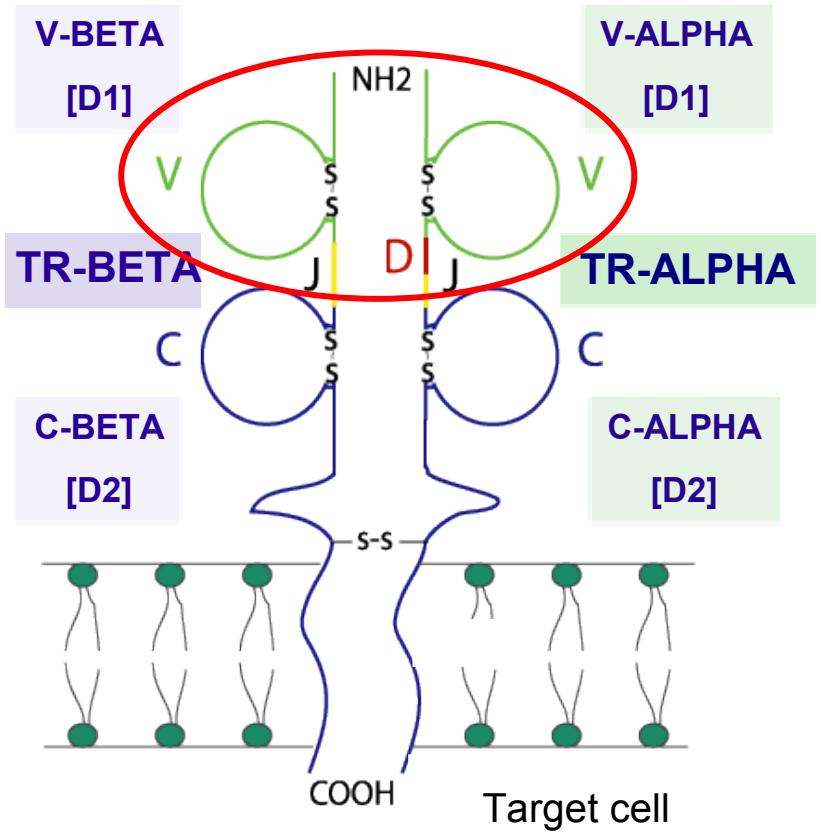
TR/peptide/MHC-I complex



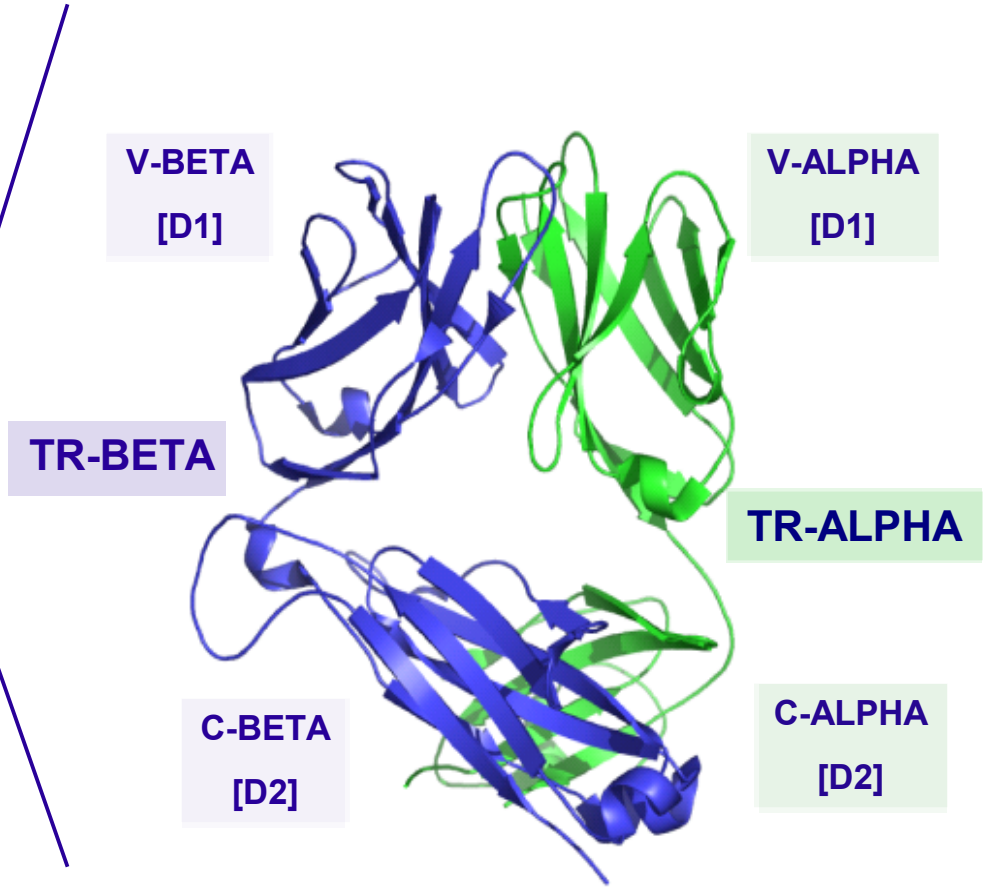
We don't have so much structural data!

TR chains and domains

Contribution of the
2 V-DOMAINs
to the antigen binding site



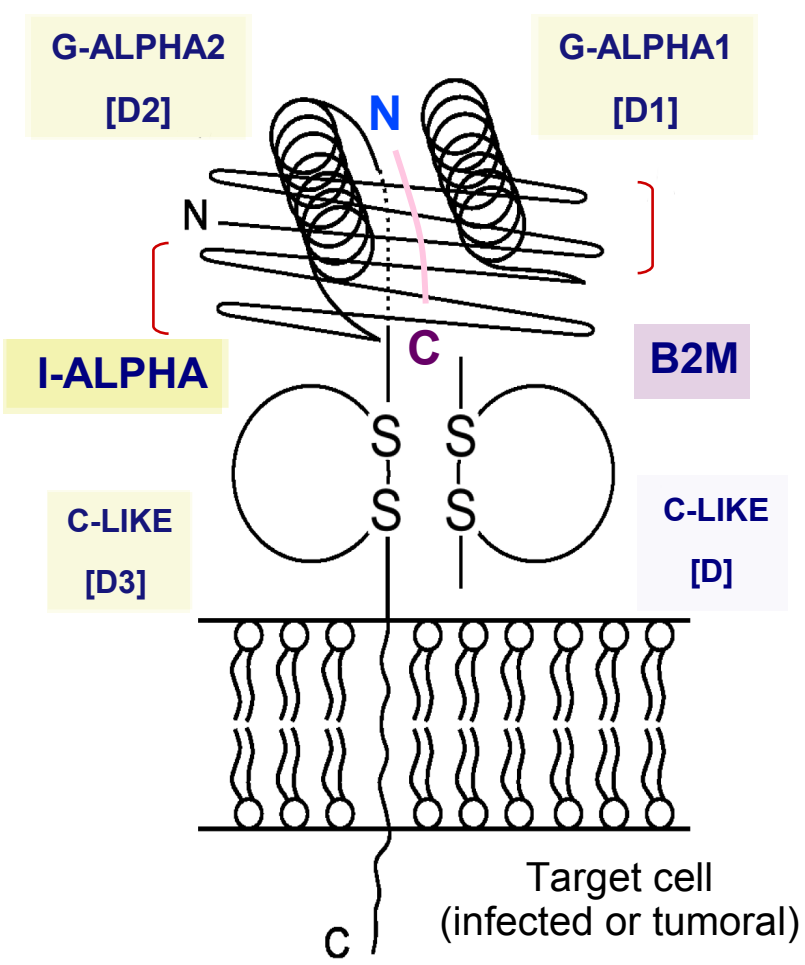
TR-ALPHA_BETA

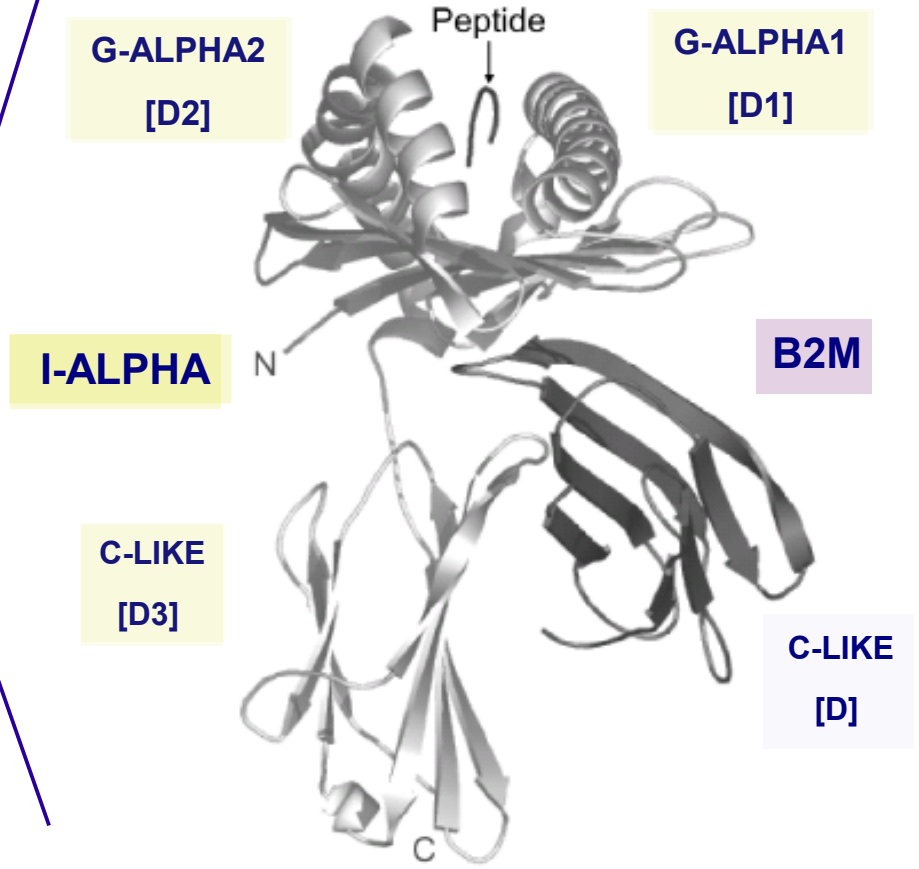


MHC-I chains and domains

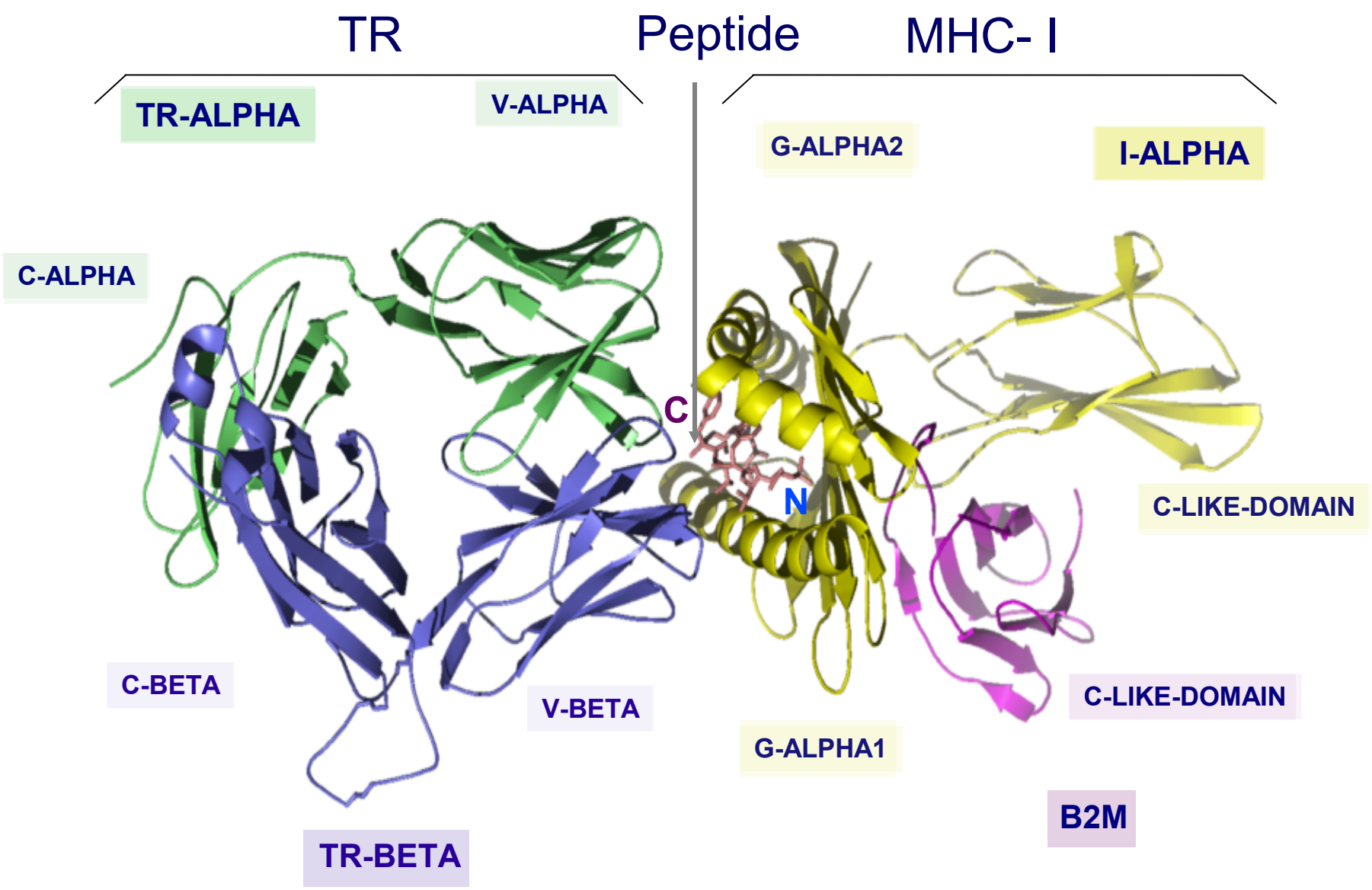
With 8/9/10 amino acid peptide in the G-DOMAIN groove

MHC-I-ALPHA_B2M

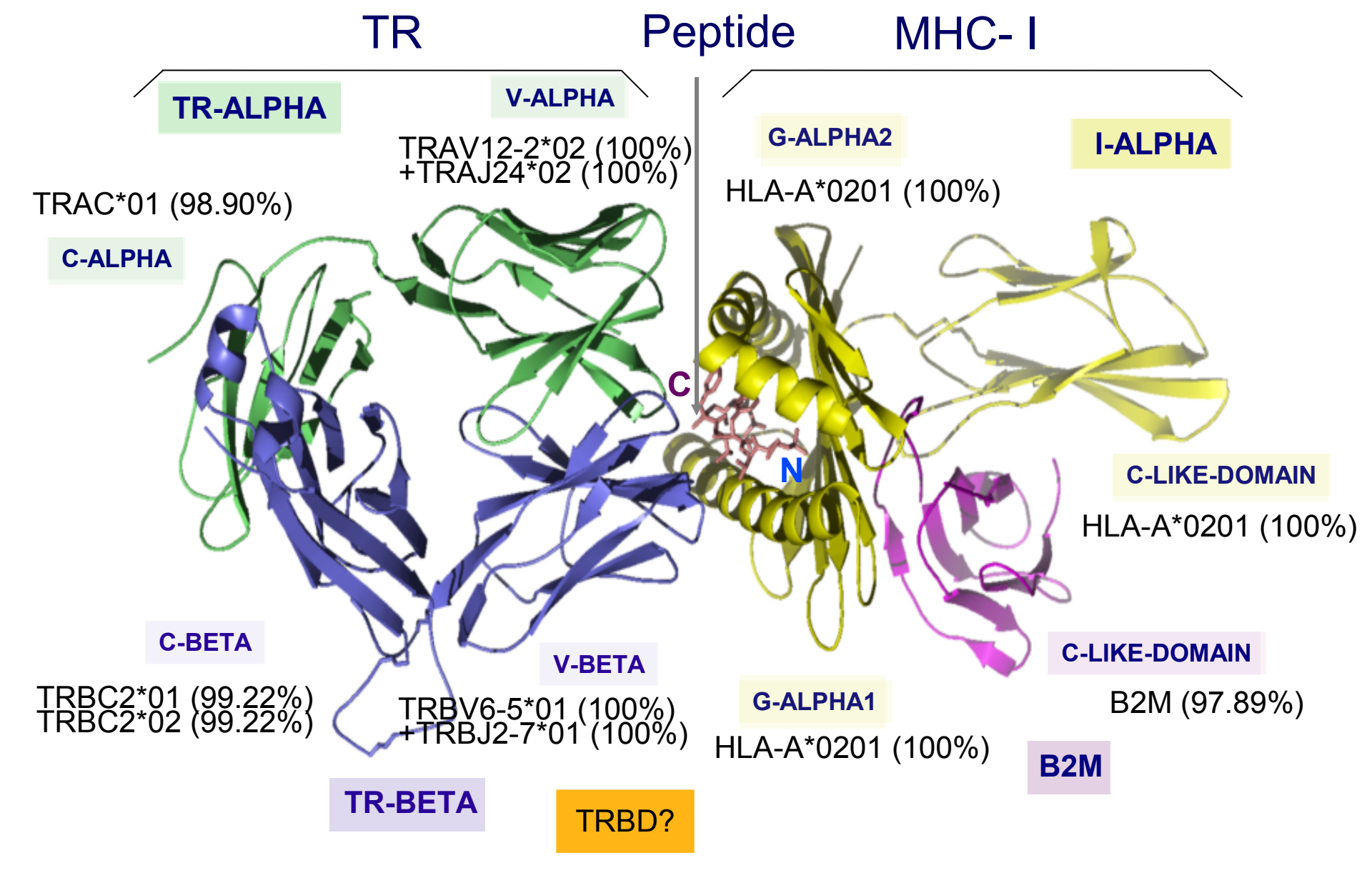




TR/peptide/MHC-I complex



TR/peptide/MHC-I complex



TRAV12-2*02 allele in sequence (M81774)

V-REGION		[67..345]
	/translation	QKEVEQNSGPLSVPEGAIASLNCTYSDRGSQSFFWYRQYSGKSPELIMSIYSNGDKEDGRFTAQLNKASQYVSLLIRDSQPSDSATYLCAVYH
		cagaaggaggtggagcagaattctggacccctcagtgttccagagggagccattgcctctctcaactgcacttacagtgaccgaggttcccagtccttcttctggta Q K E V E Q N S G P L S V P E G A I A S L N C T Y S D R G S Q S F F W Y R R R W S R I L D P S V F Q R E P L P L S T A L T V T E V P S P S S G E G G G A E F W T P Q C S R G S H C L S Q L H L Q * P R F P V L L L V **Caution:** translation of partial subregions can be erroneous.
	/gene	TRAV12-2
	/allele	TRAV12-2*02

Sequence

```
1      atgatgaaat ccttgagagt tttactagtg atcctgtggc ttcagttgag ctgggtttgg
61     agccaacaga aggaggtgga gcagaattct ggacccctca gtgttccaga gggagccatt
121    gcctctctca actgcactta cagtgaccga ggttcccagt ccttcttctg gtacagacaa
181    tattctggga aaagccctga gttgataatg tccatatact ccaatggtga caaagaagat
241    ggaaggttta cagcacagct caataaagcc agccagtatg tttctctgct catcagggac
301    tcccagccca gtgattcagc cacctacctc tgtgccgtgt accactctgg ttcaggaagg
361    caactgacct ttggatctgg gacacaattg actgttttac ctgatatcca gaaccctgac
421    cctgccgtgt ac
```

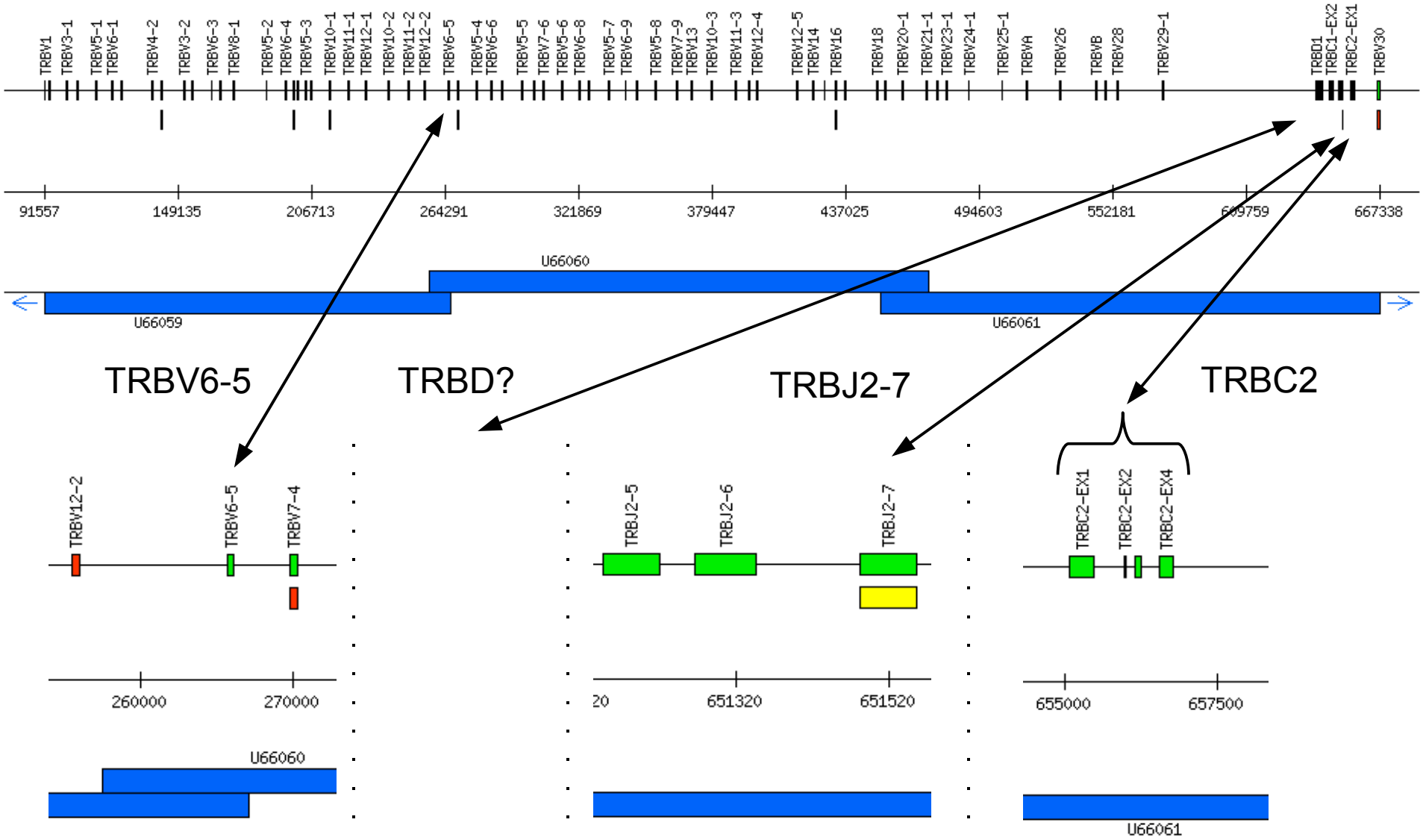
L-V-J-C-REGION [1..432>
L-V-J-REGION [1..402]
L-V-REGION [1..345]
V-J-C-REGION [67..432>
V-J-REGION [67..402]
V-REGION [67..345]
FR2-IMGT [163..213]
CONSERVED-TRP [169..171]

Literature References

[1] -1..-1 MEDLINE: 86253078
Yoshikai Y., Kimura N., Toyonaga B., Mak T.W.,
"Sequences and repertoire of human T cell receptor alpha chain variable region genes in mature T lymphocytes";
Journal: J. Exp. Med. 164(1) [1986]

A new forthcoming IMGT/LIGM-DB view

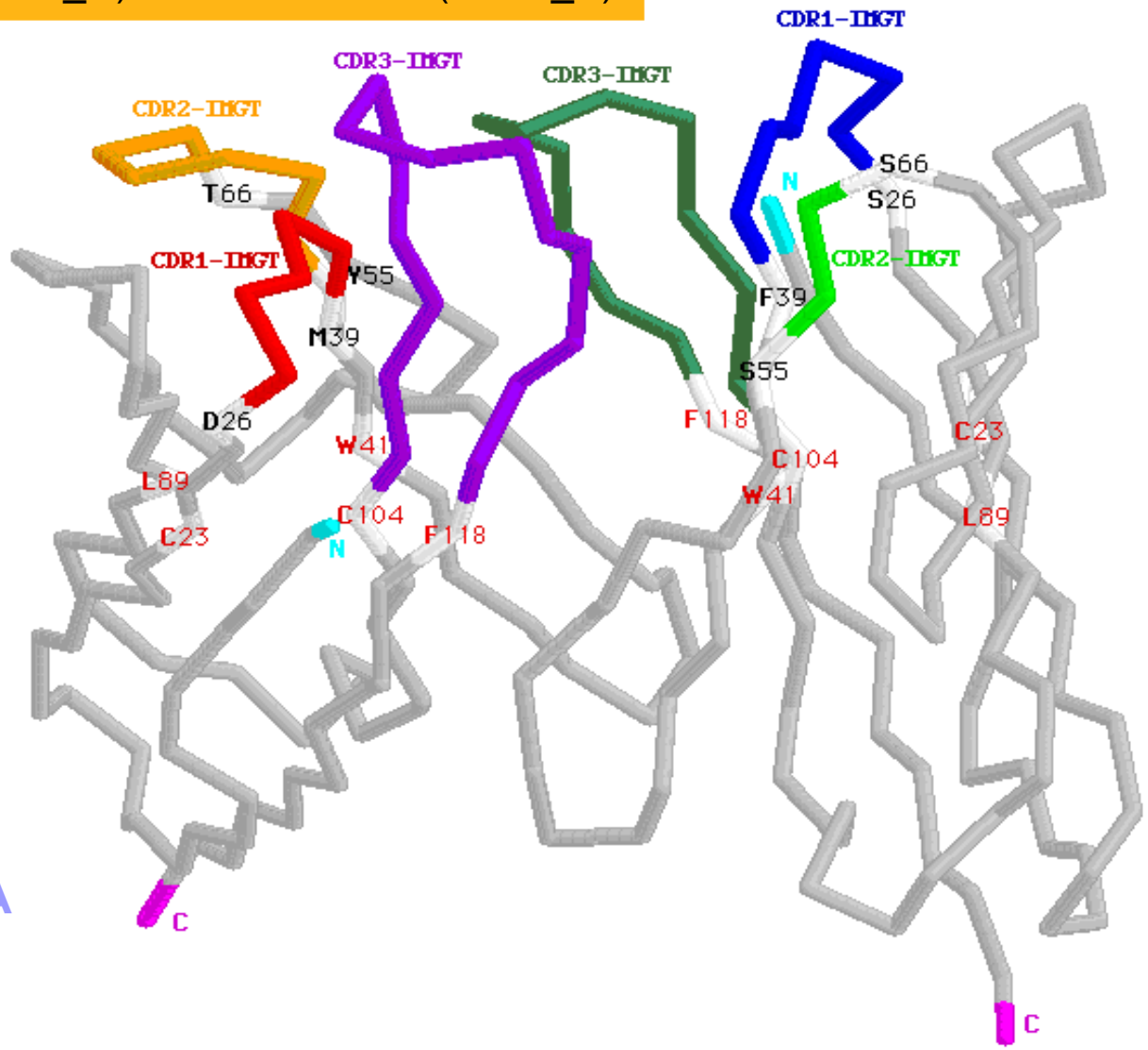
Human TRB locus at 7q34



Extracted from IMGT/LocusView

V-DOMAINs in Human TR αβ (A6)

Complementarity determining regions (CDR) in TR-ALPHA (1AO7_D) and TR-BETA-1 (1AO7_E)



V-BETA [D1]

V-ALPHA [D1]

TR/peptide/MHC-I available complexes

	IMGT entry ID	IMGT protein name	IMGT receptor description	Species	Ligand(s)	Experimental technique	Resolution	PDB release date
1	1ao7	A6 HLA-A*0201	TR-ALPHA_BETA-1 MHC-I-ALPHA_B2M	*Homo sapiens*	Tax peptide 11-19 (Q82235) [Human T lymphotropic virus type 1]	X-ray diffraction	2.6	17-SEP-97
2	1bd2	HLA-A*0201 B7	MHC-I-ALPHA_B2M TR-ALPHA_BETA-1	*Homo sapiens*	Tax peptide 11-19 (Q82235) [HTLV-1]	X-ray diffraction	2.5	19-AUG-98
3	1fo0	BM3.3 H2-K1b	FV-ALPHA_BETA MHC-I-ALPHA_B2M	*Mus musculus*	pBM1 peptide [Mouse]	X-ray diffraction	2.50	02-OCT-00
4	1g6r	2C H2-K1b	TR-ALPHA_BETA-1 MHC-I-ALPHA_B2M	*Mus musculus*	Superantagonist peptide SIYR [Chimeric]	X-ray diffraction	2.80	15-NOV-00
5	1jtr	H2-K1b 2C	MHC-I-ALPHA_B2M TR-ALPHA_BETA-1	*Mus musculus*	NADH-ubiquinone oxidoreductase MLRQ subunit peptide 61-68 (Q62425) [Mouse]	X-ray diffraction	2.40	15-MAY-02
6	1kj2	H2-K1b KB5-C20	MHC-I-ALPHA_B2M FV-ALPHA_BETA	*Mus musculus*	GTP-binding protein 1 peptide 161-168 pKB1 (O08582) [Mouse]	X-ray diffraction	2.71	27-MAR-02
7	1lp9	HLA-A*0201 12.2	MHC-I-ALPHA_B2M TR-ALPHA_BETA-1	*Homo sapiens* *Mus musculus*	Self peptide P1049 [Human]	X-ray diffraction	2.00	11-NOV-03
8	1mi5	LC13 HLA-B*0801	TR-ALPHA_BETA-1 MHC-I-ALPHA_B2M	*Homo sapiens*	EBNA-3A peptide 193-201 (P12977), I9>L [EBV]	X-ray diffraction	2.50	04-FEB-03
9	1mwa	H2-K1b 2C	MHC-I-ALPHA_B2M TR-ALPHA_BETA-1	*Mus musculus*	NADH-ubiquinone oxidoreductase MLRQ subunit peptide 61-68 (Q62425) [Mouse]	X-ray diffraction	2.40	27-NOV-02
10	1nam	BM3.3 H2-K1b	FV-ALPHA_BETA MHC-I-ALPHA_B2M	*Mus musculus*	Nucleocapsid protein VSV8 peptide 52-59 (P11212) [Stomatitis]	X-ray diffraction	2.70	11-MAR-03
11	1oga	JM22 HLA-A*0201	TR-ALPHA_BETA-1 MHC-I-ALPHA_B2M	*Homo sapiens*	Matrix protein M1 peptide 58-66 (Q66PA1) [Influenza A virus]	X-ray diffraction	1.4	11-JUL-03
12	1qrn	A6 HLA-A*0201	TR-ALPHA_BETA-1 MHC-I-ALPHA_B2M	*Homo sapiens*	Tax peptide 11-19 (Q82235), P6>A [HTLV-1]	X-ray diffraction	2.80	08-JUN-01
13	1qse	HLA-A*0201 A6	MHC-I-ALPHA_B2M TR-ALPHA_BETA-1	*Homo sapiens*	Tax peptide 11-19 (Q82235), V7>R [HTLV-1]	X-ray diffraction	2.80	21-DEC-99
14	1qsf	A6 HLA-A*0201	TR-ALPHA_BETA-1 MHC-I-ALPHA_B2M	*Homo sapiens*	Tax peptide 11-19 (Q82235), Y8>A [HTLV-1]	X-ray diffraction	2.80	21-DEC-99
15	1ypz	G8 H2-T22	TR-GAMMA-1_DELTA MHC-I-ALPHA_B2M	*Mus musculus* *Homo sapiens*		X-ray diffraction	3.40	12-APR-05
16	2ak4	SB27 HLA-B*3508	TR-ALPHA_BETA-1 MHC-I-ALPHA_B2M	*Homo sapiens*	BZLF1 trans-activator protein peptide 52-64 (P03206) [EBV]	X-ray diffraction	2.50	11-OCT-05
17	2ckb	H2-K1b 2C	MHC-I-ALPHA_B2M TR-ALPHA_BETA-1	*Mus musculus*	NADH-ubiquinone oxidoreductase MLRQ subunit peptide 61-68 (Q62425) [Mouse]	X-ray diffraction	3.2	09-SEP-98
18	2esv	KK50.4 HLA-E*0101	TR-ALPHA_BETA-1 MHC-I-ALPHA_B2M	*Homo sapiens*	PEPTIDE FROM CMV GPUL40 [Cytomegalovirus]	X-ray diffraction	2.60	21-MAR-06
19	2f53	1G4 mutant (c49c50) HLA-A*0201	TR-ALPHA_BETA-1 MHC-I-ALPHA_B2M	*Homo sapiens*	NY-ESO-1 tumor-associated antigen	X-ray diffraction	2.10	25-APR-06
20	2f54	1G4 (AV-wt) HLA-A*0201	TR-ALPHA_BETA-1 MHC-I-ALPHA_B2M	*Homo sapiens*	NY-ESO-1 tumor-associated antigen	X-ray diffraction	2.70	25-APR-06
21	2nx5	ELS4 HLA-B*3501	TR-ALPHA_BETA-1 MHC-I-ALPHA_B2M	*Homo sapiens*	EBV PEPTIDE	X-ray diffraction	2.70	27-FEB-07

Human	12
Mouse	7
Human/Mouse	2
Total	21

TR/peptide/MHC-II

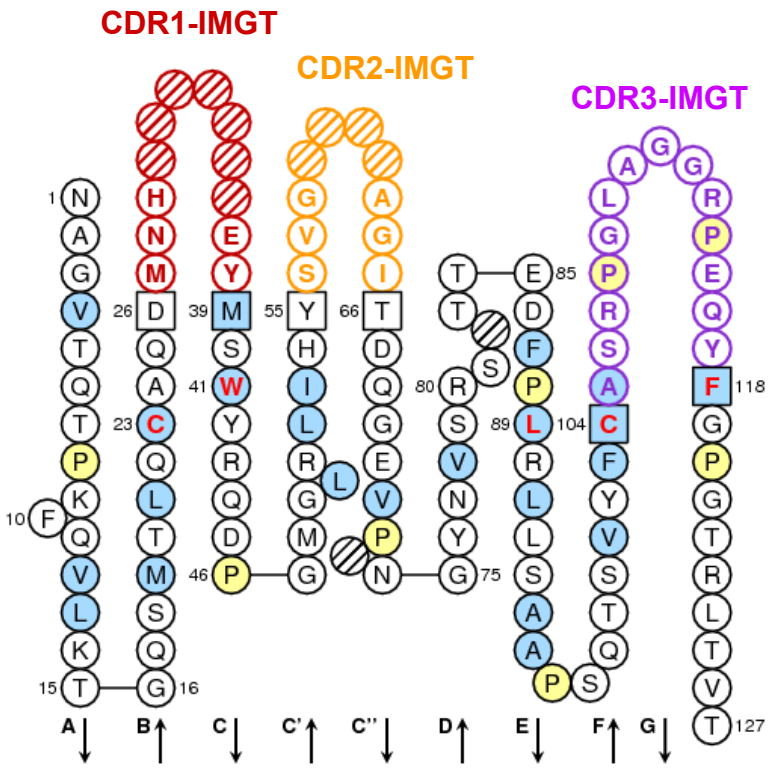
Human	4
Mouse	2
Human/Mouse	1
Total	7

FR and CDR lengths in V-DOMAINs

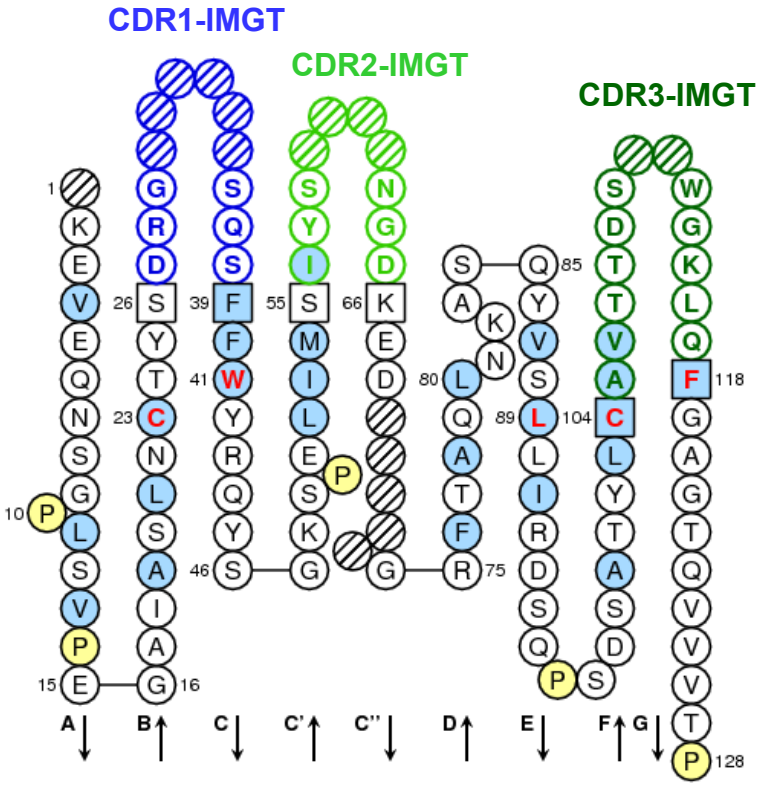
		V-BETA	V-ALPHA
A-STRAND		15	14
AB-TURN	(FR1-IMGT)	0	0
B-STRAND		11	11
BC-TURN	(CDR1-IMGT)	5	6
C-STRAND		8	8
CC'-TURN	(FR2-IMGT)	0	0
C'-STRAND		9	9
C'C''-TURN	(CDR2-IMGT)	6	6
C''-STRAND		8	4
C''D-TURN		0	0
D-STRAND		9	10
DE-TURN	(FR3-IMGT)	0	0
E-STRAND		11	11
EF-TURN		2	2
F-STRAND		7	7
FG-TURN	(CDR3-IMGT)	14	11
G-STRAND		10	11
Total		115	110

IMGT Colliers de Perles for V-DOMAIN

Based on IMGT unique numbering of V-DOMAIN



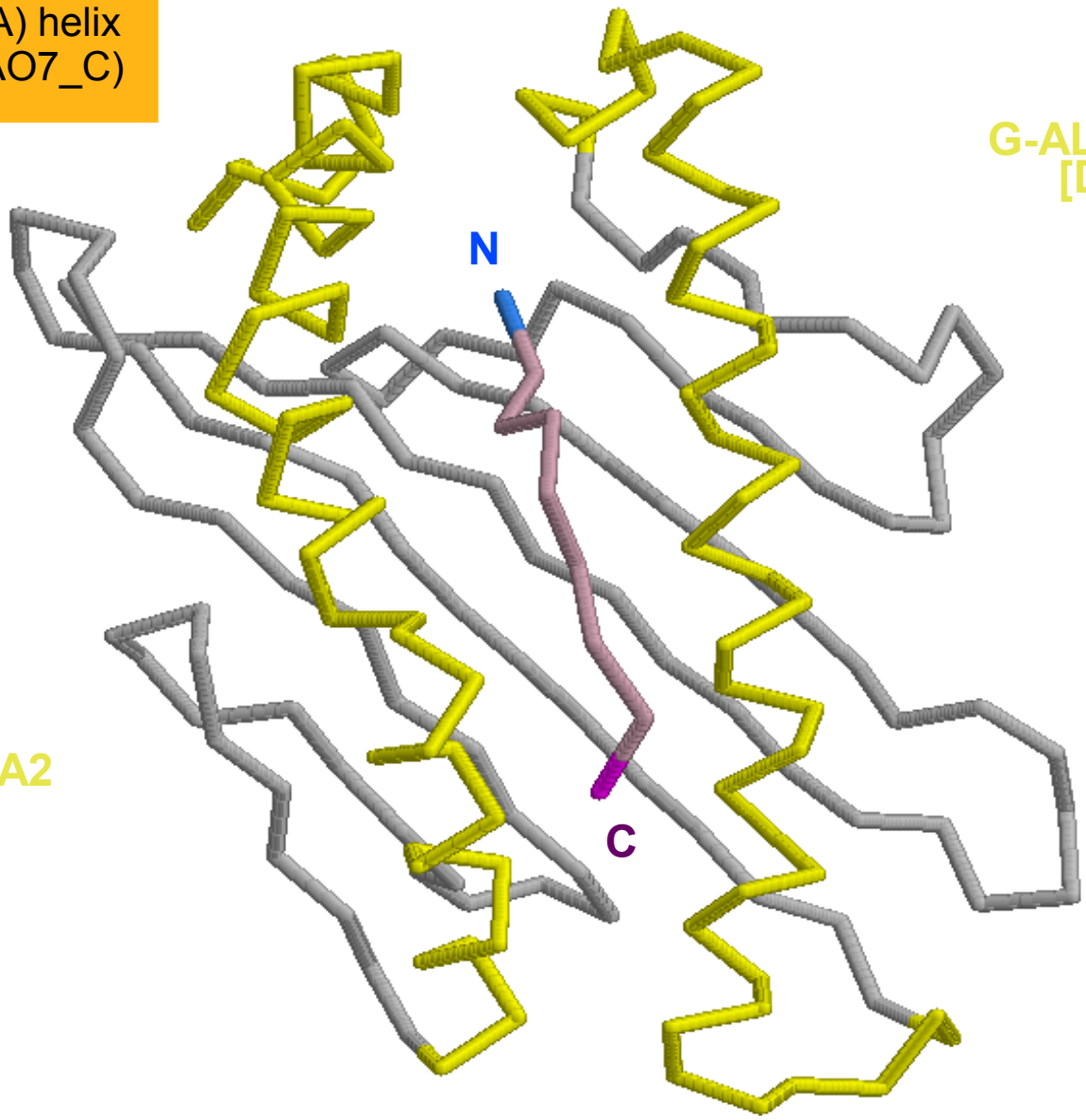
V-BETA
[5.6.14]



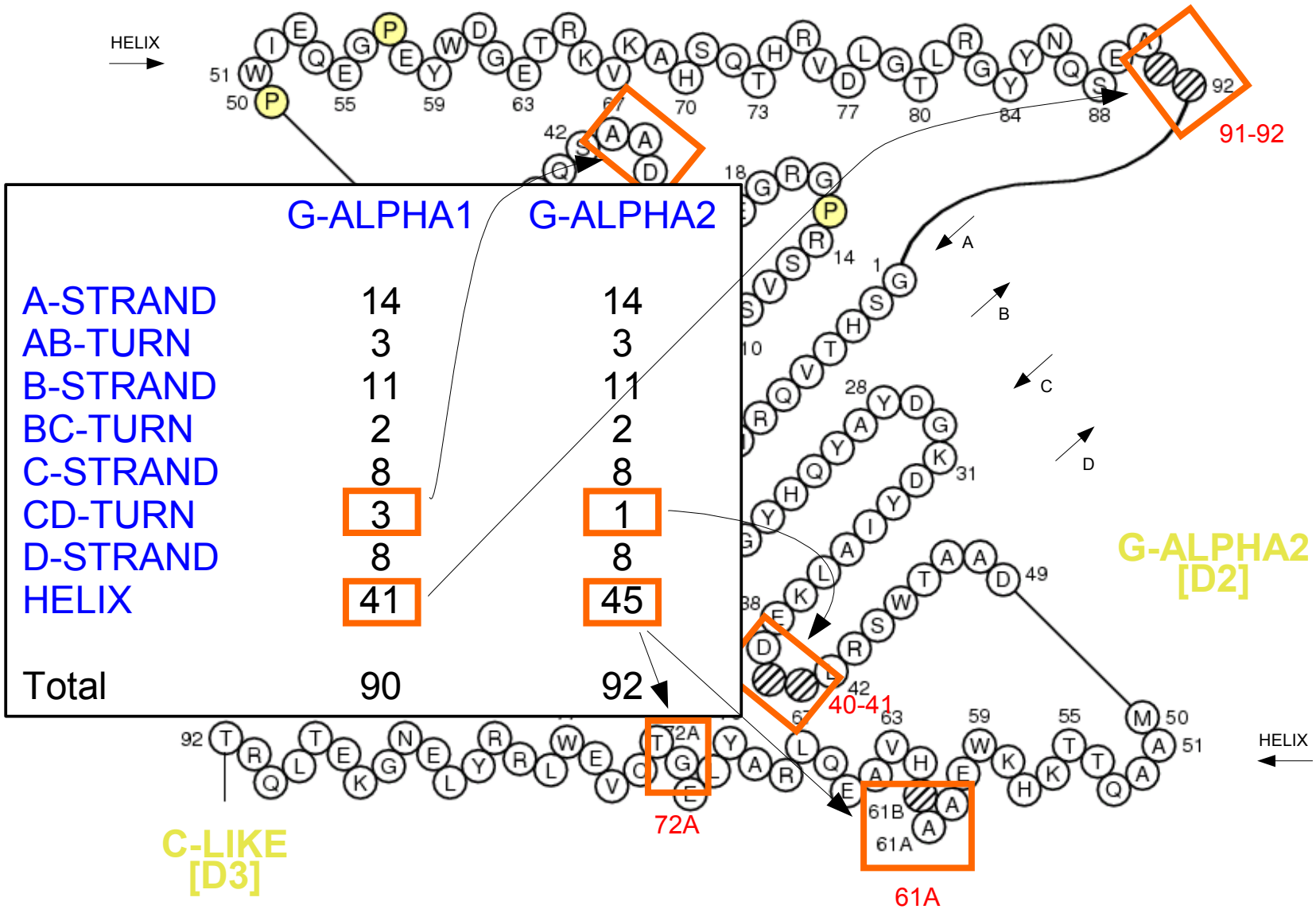
V-ALPHA
[6.6.11]

G-DOMAINs in Human MHC-I (HLA-A*0201)

I-ALPHA (1AO7_A) helix and a peptide (1AO7_C)

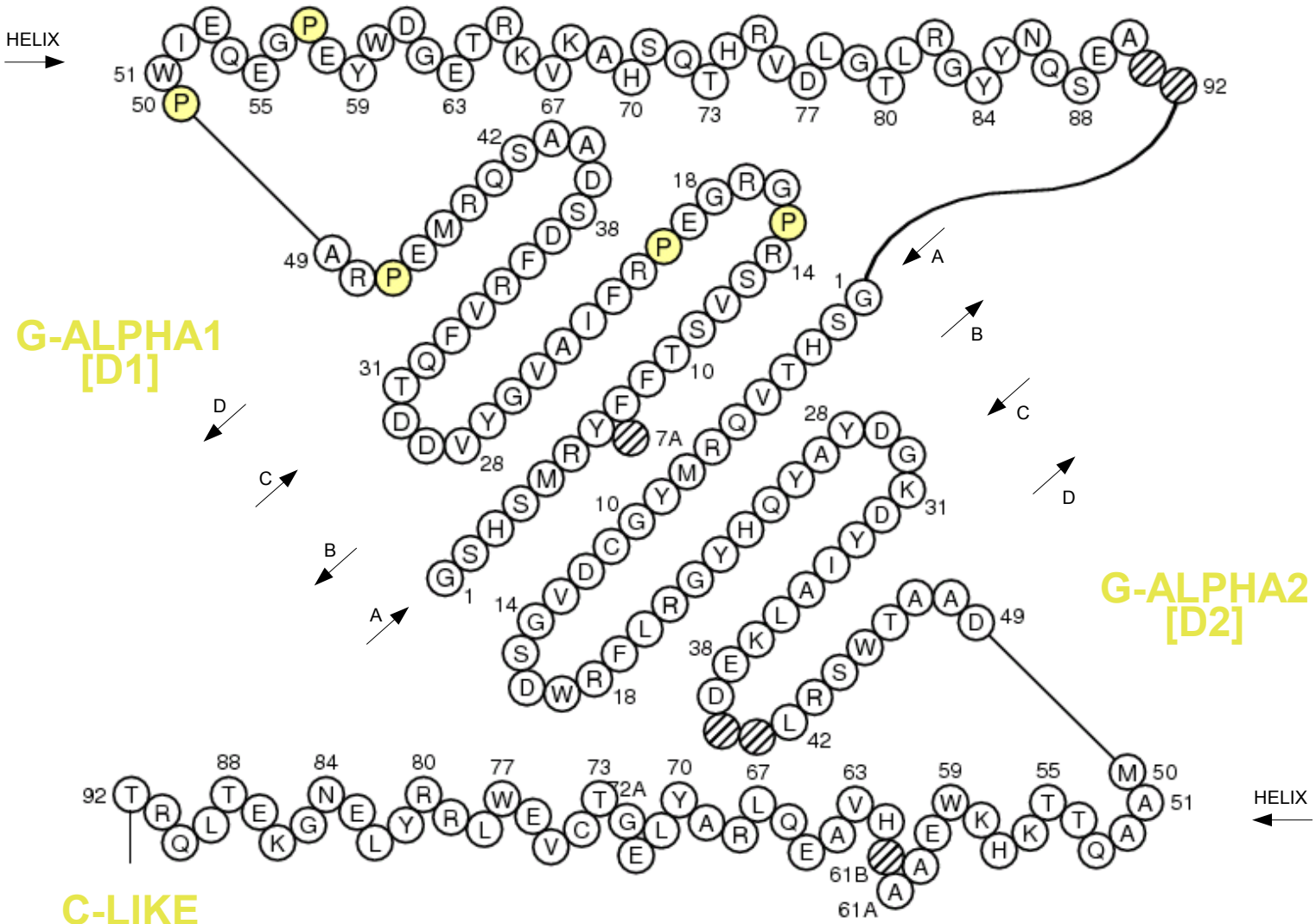


Strand, turn and helix lengths



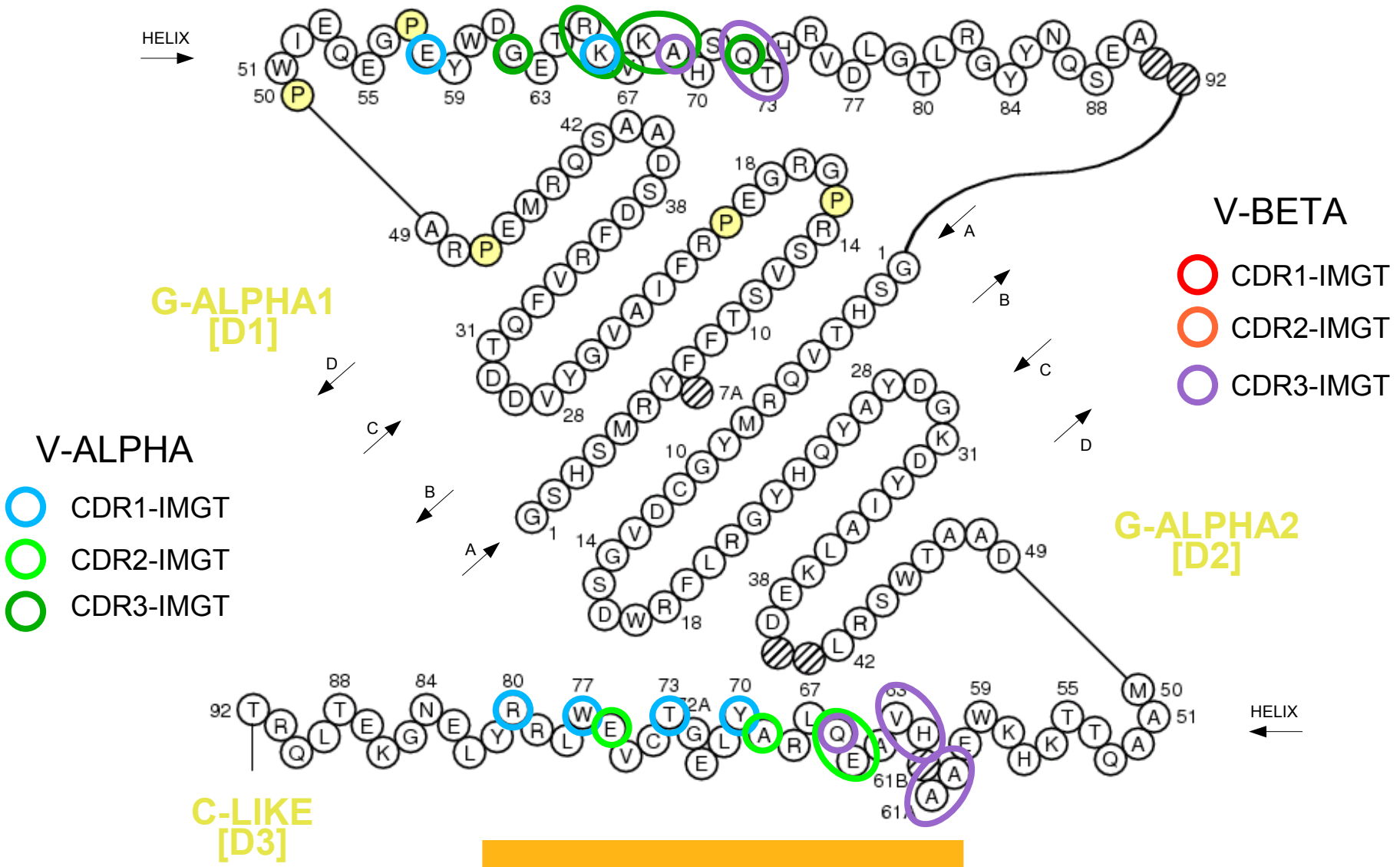
	G-ALPHA1	G-ALPHA2
A-STRAND	14	14
AB-TURN	3	3
B-STRAND	11	11
BC-TURN	2	2
C-STRAND	8	8
CD-TURN	3	1
D-STRAND	8	8
HELIX	41	45
Total	90	92

IMGT Colliers de Perles for G-DOMAIN



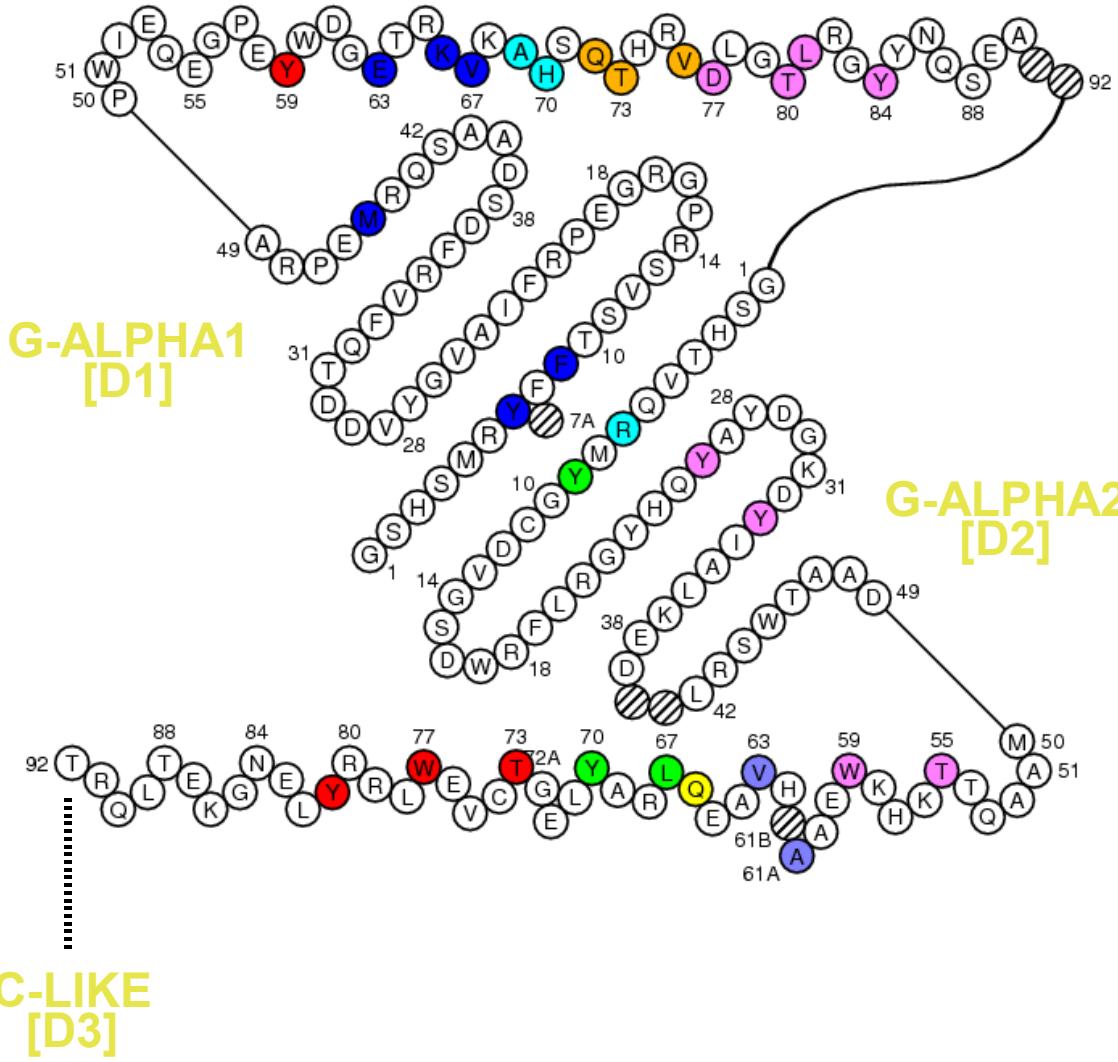
Based on IMGT unique numbering of G-DOMAIN

HLA-A*0201 regions in contact with A6



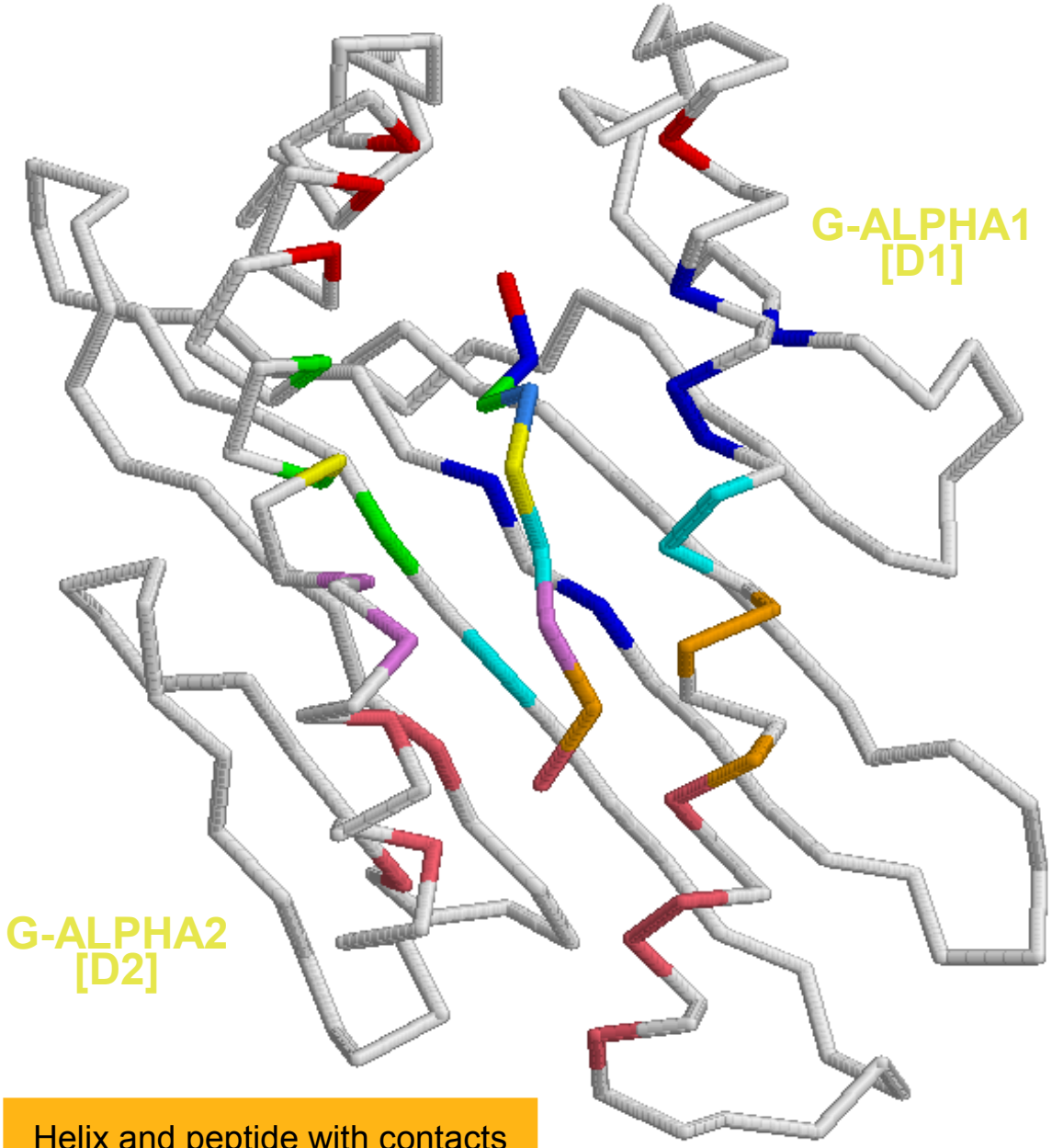
No contact with CDR1-IMGT and CDR2-IMGT of V-BETA

HLA-A*0201 contacts with 9 AA peptide



	9-amino acid peptides	Peptide
C1	1	1 L
C2	-	--
C3	2	2 L
C4	3	3 F
C5	4	(4 G)
C6	5	5 Y
C7	-	--
C8	6	6 P
C9	7	7 V
C10	8	8 Y
C11	9	9 V

HLA-A*0201 contacts with 9 AA peptide



	9-amino acid peptides	Peptide
C1	1	1 L
C2	-	--
C3	2	2 L
C4	3	3 F
C5	4	(4 G)
C6	5	5 Y
C7	-	--
C8	6	6 P
C9	7	7 V
C10	8	8 Y
C11	9	9 V

Helix and peptide with contacts

One residue contact analysis

IMGT Residue@Position cards
as found in IMGT/3Dstructure-DB

IMGT Residue@Position card

Residue@Position: 113 - ARG (R) - V-BETA - 1ao7_E CDR3-IMGT

General information:

PDB file numbering	102
IMGT file numbering	113
Residue full name	Arginine
Formula	C6 H15 N4 O2 1+

IMGT LocalStructure@Position

Secondary structure	Coil
Phi (in degrees)	-89.71
Psi (in degrees)	111.56
ASA (in square angstrom)	73.2

IMGT Num	Residue		Domain	Chain	Atom contacts	Polar	Hydrogen Bond	Non Polar
61	ALA	A	G-ALPHA2	1ao7_A	5	2	1	3
61A	ALA	A	G-ALPHA2	1ao7_A	24	6	0	18
62	HIS	H	G-ALPHA2	1ao7_A	12	2	0	10
66	GLN	Q	G-ALPHA2	1ao7_A	2	1	0	1
5	TYR	Y		1ao7_C	1	0	0	1
108	PRO	P	V-BETA	1ao7_E	15	1	0	14
111	ALA	A	V-BETA	1ao7_E	6	2	0	4
112.1	GLY	G	V-BETA	1ao7_E	24	5	0	19
115	GLU	E	V-BETA	1ao7_E	17	3	0	14

G-ALPHA2

Peptide

V-BETA

Human cDNA TRAV12-2 available junctions

IMGT/V-QUEST to look at some specificities

(cut has no D)

Sequence	V-GENE and allele	Functionality	V Score	V Identity	J-GENE and allele	CDR-IMGT lengths	AA JUNCTION	JUNCTION frame
M27369|M27369|Homo sapiens (cl	TRAV12-2*01	Unproductive (stop codons)	791	80,99% (196/242 nt)	TRAJ52*01	[12,10,14]	CAVKPAGGTSYGKLTF	in frame
S60781|S60781|Homo sapiens T-c	TRAV12-2*01, or TRAV12-2*02	No rearrangement found	636	99,22% (128/129 nt)	-	[12,X,X]	-	-
S82064|S82064|V alpha 2.1-J al	TRAV12-2*01	Productive	1075	100,00% (216/216 nt)	TRAJ39*01	[12,10,12]	CAVNAGNAGNMLTF	in frame
S82066|S82066|V alpha 2.1-J al	TRAV12-2*01	Productive	955	100,00% (192/192 nt)	TRAJ45*01	[12,10,10]	CAVNEGADGLTF	in frame
X92783|HSXPMS2A|H.sapiens mRNA	TRAV12-2*01	Productive	1142	99,14% (231/233 nt)	TRAJ57*01	[12,10,15]	CAVNIVGTQGGSEKLVF	in frame
AF020651|AF020651|Homo sapiens	TRAV12-2*01	Productive	1325	99,63% (266/267 nt)	TRAJ29*01	[12,10,12]	CSVMSNSGNTPLVF	in frame
AF327017|AF327017|Homo sapiens	TRAV12-2*01	Productive	1000	100,00% (201/201 nt)	TRAJ43*01	[12,10,11]	CAVDAADNNDMRF	in frame
M17652|HSTCAYN|Human T-cell re	TRAV12-2*02	Productive	1245	100,00% (250/250 nt)	TRAJ15*01	[12,10,13]	CAVNIPNQAGTALIF	in frame
M17653|HSTCAYO|Human T-cell re	TRAV12-2*02	Productive	1245	100,00% (250/250 nt)	TRAJ15*01	[12,10,10]	CAPKPGGTALIF	in frame
M81774|HSIGTCACA|Homo sapiens	TRAV12-2*02	Productive	1330	100,00% (267/267 nt)	TRAJ22*01	[12,10,12]	CAVYHSGSARQLTF	in frame
U40464|HS404641|Human T cell r	TRAV12-2*02, or TRAV12-2*03	Productive	875	100,00% (176/176 nt)	TRAJ41*01	[12,10,13]	CALKGRSNSGYALNF	in frame
X58746|X58746|Human mRNA for T	TRAV12-2*02	No rearrangement found	1330	100,00% (267/267 nt)	-	[12,10,X]	-	-
X92883|HSPHC46A1|H.sapiens mRN	TRAV12-2*02	Productive	1142	99,14% (231/233 nt)	TRAJ34*01	[12,10,11]	CAVPFYNTDKLIF	in frame
AF532854|AF532854|Homo sapiens	TRAV12-2*02	Productive	1161	99,57% (234/235 nt)	TRAJ32*02	[12,10,10]	CADGGATNKLIF	in frame
M13724|HSTCAXH|Human T-cell re	TRAV12-2*03	Rearranged sequence (but no junction found)	1210	100,00% (243/243 nt)	TRAJ16*01	[12,10,X]	-	-
M13725|HSTCAXI|Human T-cell re	TRAV12-2*03	Productive	1210	100,00% (243/243 nt)	TRAJ18*01	[12,10,14]	CAVNYPRGTTLGRLYF	in frame
X04946|HSTCRA12|Human mRNA for	TRAV12-2*03	Productive	1210	100,00% (243/243 nt)	TRAJ18*01	[12,10,14]	CAVNYPRGTTLGRLYF	in frame

V-REGIONs *versus* V-ALPHA alignment

but:
1. no protein only sequence input
2. no match with the other JUNCTIONs

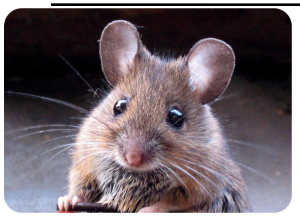
6. V-REGION protein display (mutations displayed)

```
                                  FR1-IMGT          CDR1-IMGT       FR2-IMGT      CDR2-IMGT              FR3-IMGT
                                  (1-26)            (27-38)         (39-55)       (56-65)                (66-104)
                                  ____________________ _____________ _____________ _________ ____________________________________
                                  1       10        20        30        40        50        60        70        80   ABC  90        100
                                  .........|.........|..... ...|..... .|.........|..... ....|..... ....|.........|.....|.........|....
AE000659 TRAV12-2*01              QKEVEQNSGPLSVPEGAIASLNCTYS DRGSQS..... FFWYRQYSGKSPELIMF IYSNG..... DKEDG.......RFTAQLNKASQYVSLLIRDSQPSDSATYLC AVN
M13724|HSTCAXH|Human T-cell r ........------------------- --V---..... ----------------S ----..... -----.......------------------------------ -LRDGQKLLFARGTMLKVDL
M13725|HSTCAXI|Human T-cell r ........------------------- --V---..... ----------------S ----..... -----.......------------------------------ ---YPRGTTLGRLYFGRGTQLTVWP
M17652|HSTCAYN|Human T-cell r .......------------------- ------..... ----------------S ----..... -----.......------------------------------ ---IPNQAGTALIFGKGTTLSVRS
M17653|HSTCAYO|Human T-cell r .......------------------- ------..... ----------------S ----..... -----.......------------------------------ -PKPGGTALIFGKGTTLSVSS
M27369|M27369|Homo sapiens (c .........TPQCSR-SHCLSQLHLQ *PS---..... ---------------- ----..... -----.......------------------------------ --KPAGGTSYGKLTFGQGTILTVHP
M81774|HSIGTCACA|Homo sapiens ------------------------- ------..... ----------------S ----..... -----.......------------------------------ --YHSGSARQLTFGSGTQLTVLP
S60781|S60781|Homo sapiens T- ------------------------- --V---..... ----------
S82064|S82064|V alpha 2.1-J a ................---------- ------..... ---------------- ----..... -----.......------------------------------ ---AGNAGNMLTFGGGTRLMVKP
S82066|S82066|V alpha 2.1-J a ........................- ------..... ---------------- ----..... -----.......------------------------------ ---EGADGLTFGKGTHLIIQP
U40464|HS404641|Human T cell  ......................... .....-..... ----------------S ----..... -----.......------------------------------ -LKGRSNSGYALNFGKGTSLLVTP
X04946|HSTCRA12|Human mRNA fo ........------------------- --V---..... ----------------S ----..... -----.......------------------------------ ---YPRGTTLGRLYFGRGTQLTVWP
X58746|X58746|Human mRNA for  ------------------------- ------..... ----------------S ----..... -----.......------------------------------ -
X92783|HSXPMS2A|H.sapiens mRN ............-----------S-- ------..... ---------------- ----..... -----.......----------------M------------- ---IVGTQGGSEKLVFGKGTKLTVNP
X92883|HSPHC46A1|H.sapiens mR ............-------------- ------..... --------R-------T ----..... -----.......------------------------------ --PFYNTDKLIFGTGTRLQVFP
AF020651|AF020651|Homo sapien ------------------------- ------..... ---------------- ----..... -X---.......------------------------------ S-MSNSGNTPLVFGKGTRLSVIX
AF327017|AF327017|Homo sapien ......................---- ------..... ---------------- ----..... -----.......------------------------------ --DAADNNDMRFGAGTRLTVKP
AF532854|AF532854|Homo sapien ..........-------V------- ------..... ----------------S ----..... -----.......------------------------------ -DGGATNKLIFGTGTLLAVQP

1ao7_D|TR-ALPHA .KEVEQNSGPLSVPEGAIASLNCTYS DRGSQS..... FFWYRQYSGKSPELIMS IYSNG..... DKEDG.......RFTAQLNKASQYVSLLIRDSQPSDSATYLC AVTTDSWGKLQFGAGTQVVVTP
```

Why are IMGT Collier de Perles so useful?

- bridge the gaps between sequences and structures,
- are used whatever the MHC and whatever the species,

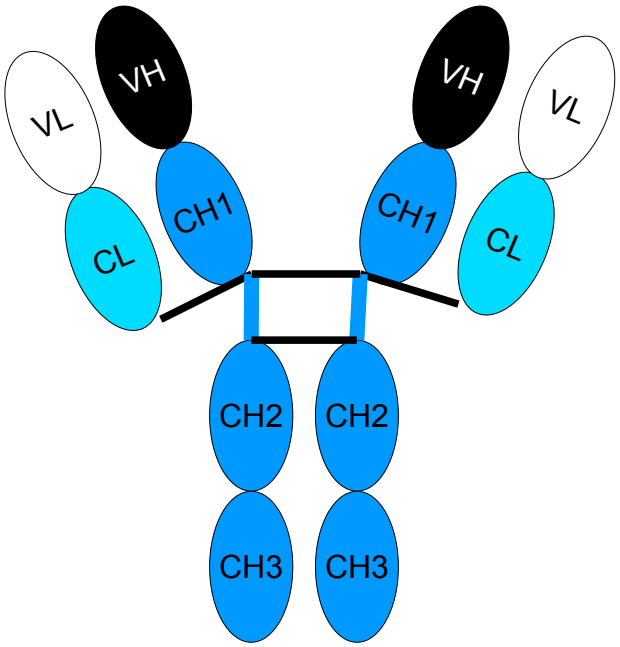
	MHC-Ia	MHC-Ib	MHC-IIa	MHC-IIb
	HLA-A,-B,-C	HLA-E,-F,-G	HLA-DPA,-DQA,- DRA HLA-DPB,-DQB,-DRB	HLA-DMA,-DOA HLA-DMB,-DOB
	H2-D,-K,-L	H2-M,-Q,-T	H2-AA,-EA H2-AB,-EB	H2-DMA,-DOA H2-DMB,-DOB

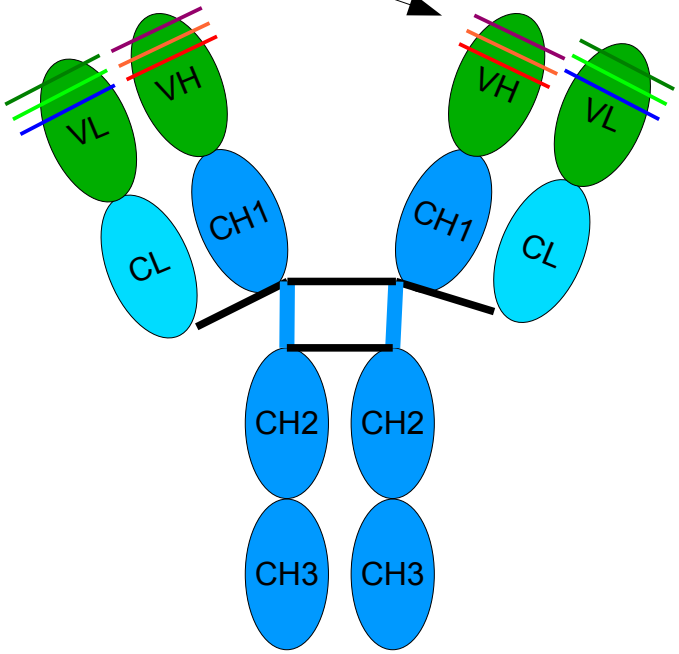
- have been extended to the MHC-I-like proteins (CD1, FCRN, RAET, HFE, MICA, AZGP1,…)

Interestingly, only one additional position **54A** in **G-ALPHA1-LIKE** was needed to extend the IMGT unique numbering for G-DOMAIN to the G-LIKE-DOMAIN

Go towards antibody engineering

Standardize.
A more clear view?





Chimeric
(-ximab)

Humanized
(-zumab)

Some perspectives

- Improve consistency and quality by considering structural approache and may be others

- Increase by the 3 approaches the data and tools to handle MHC and RPI sequences

- Extend consequently IMGT-ONTOLOGY

- Develop better interoperability between components inside the platform with IMGT-Choreography

IMGT® team thanks you!

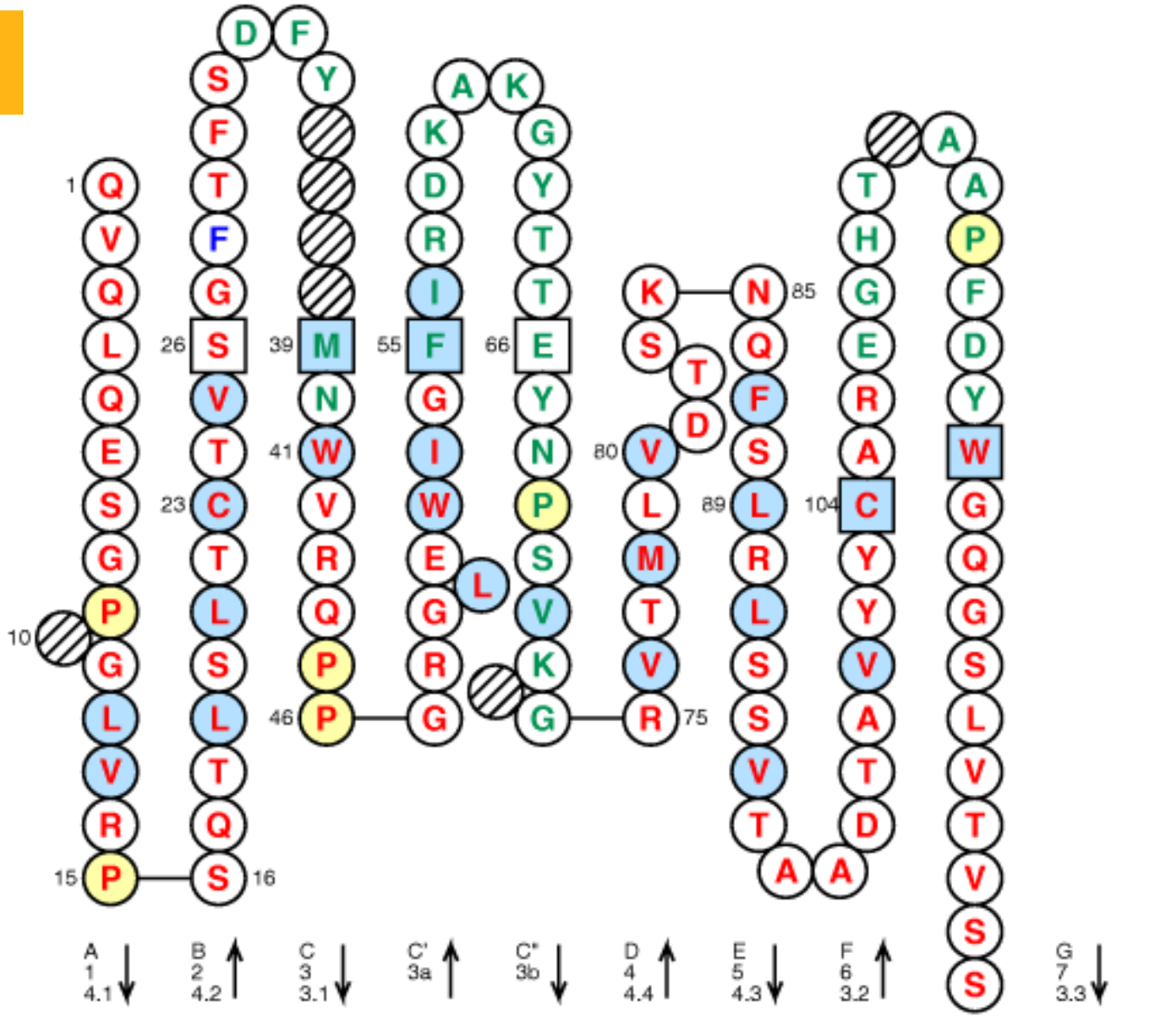
Humanized CAMPATH-1H mutant 1

Mutant 1: S28>F
Mutant 2: S31>T
(alemtuzumab)

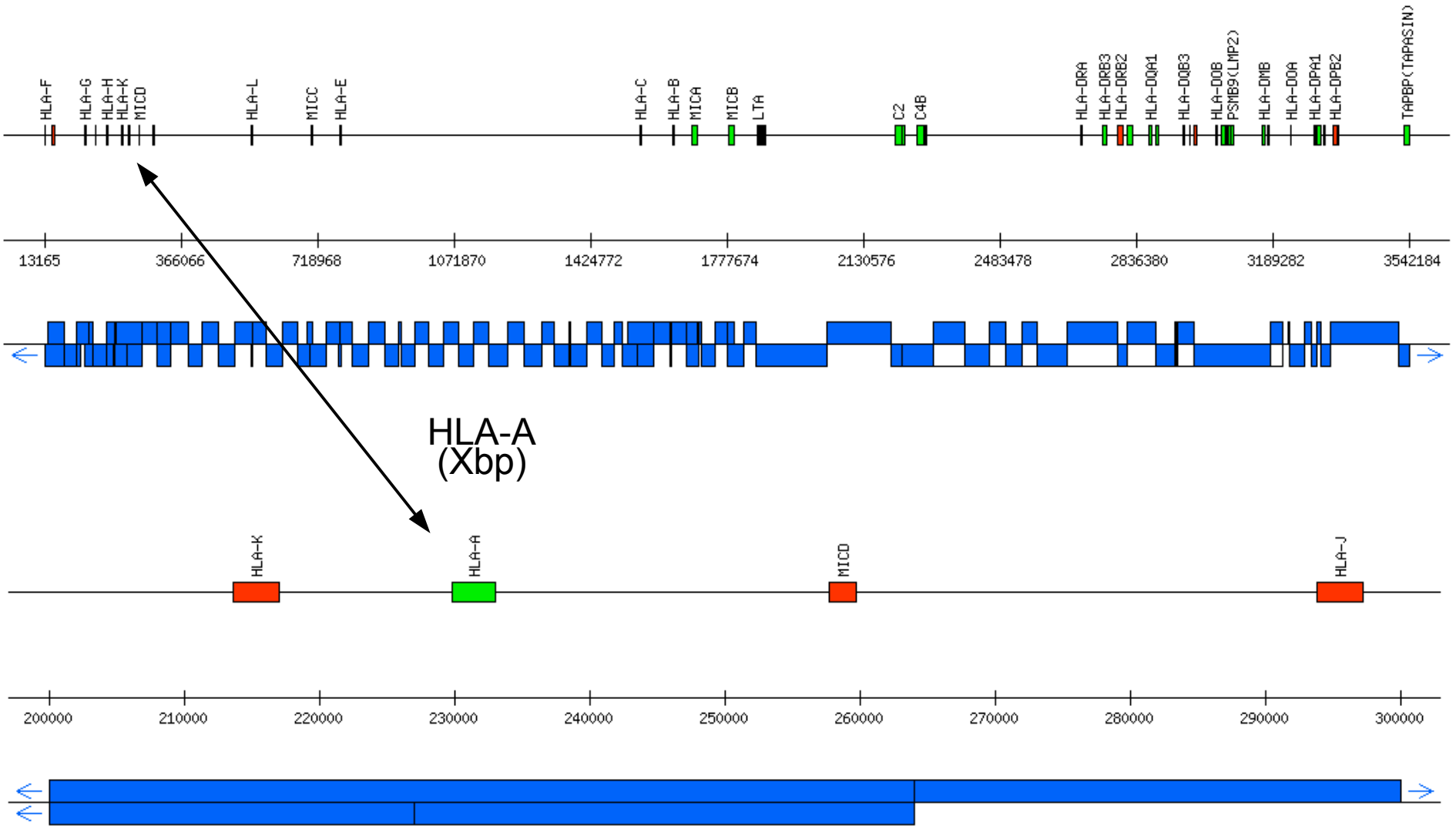
VH domain
(V-D-J-REGION)

[8.10.12]

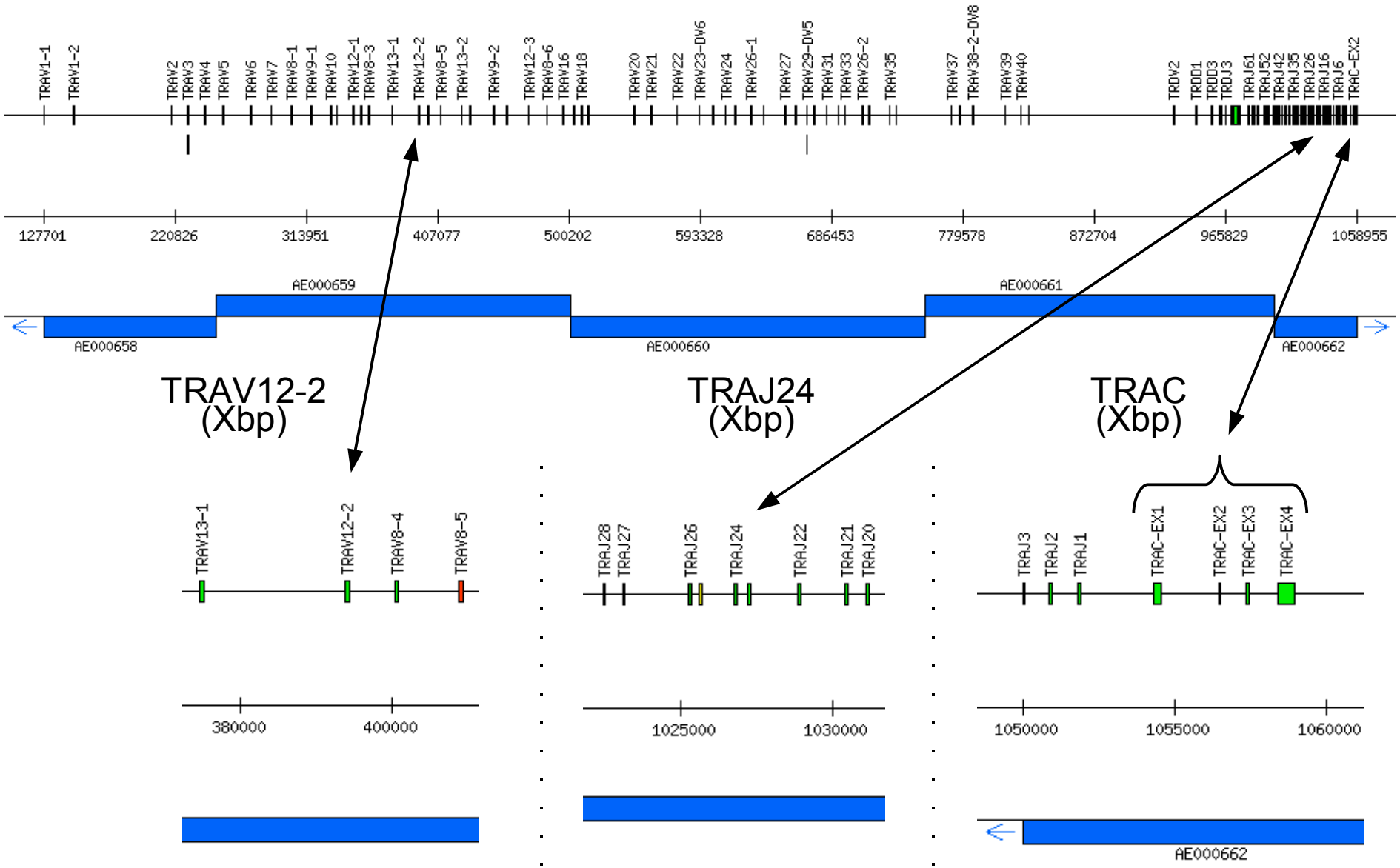
X human
X rat



Human MHC locus at 6p21.3

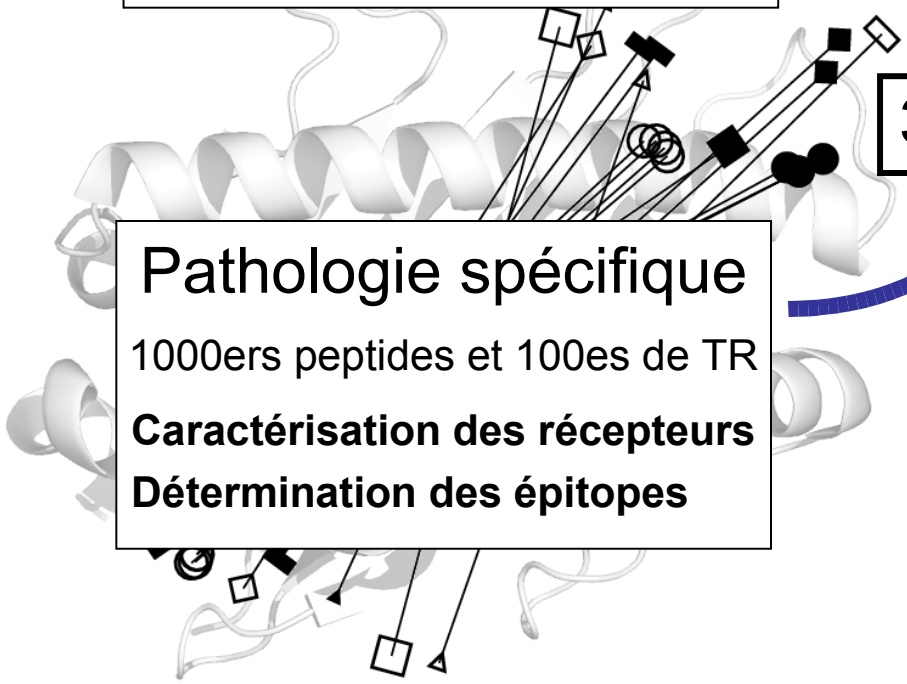


Human TRA-TRD locus at 14q11.2



Structural variability analysis

TR/peptide/MHC
21 3D-structures



Pathologie spécifique

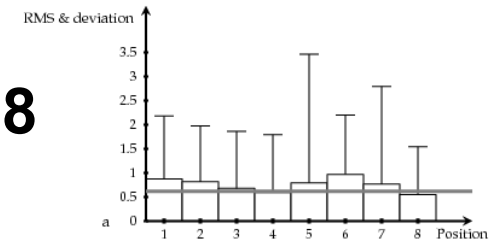
1000ers peptides et 100es de TR

Caractérisation des récepteurs

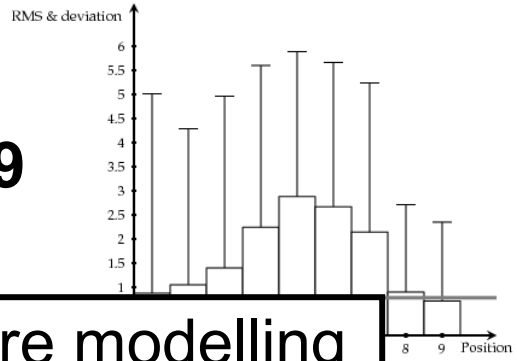
Détermination des épitopes

3D-structure modelling

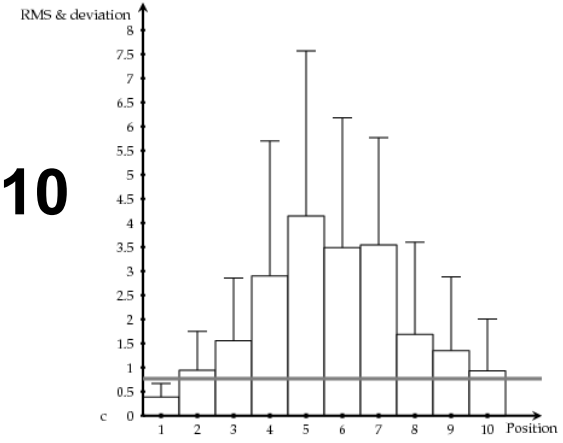
MHC-I 8



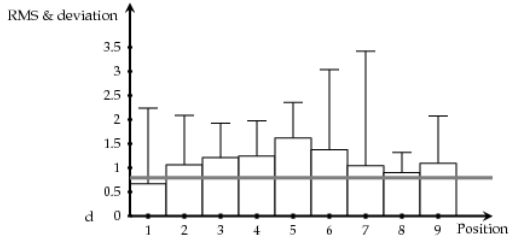
MHC-I 9



MHC-I 10

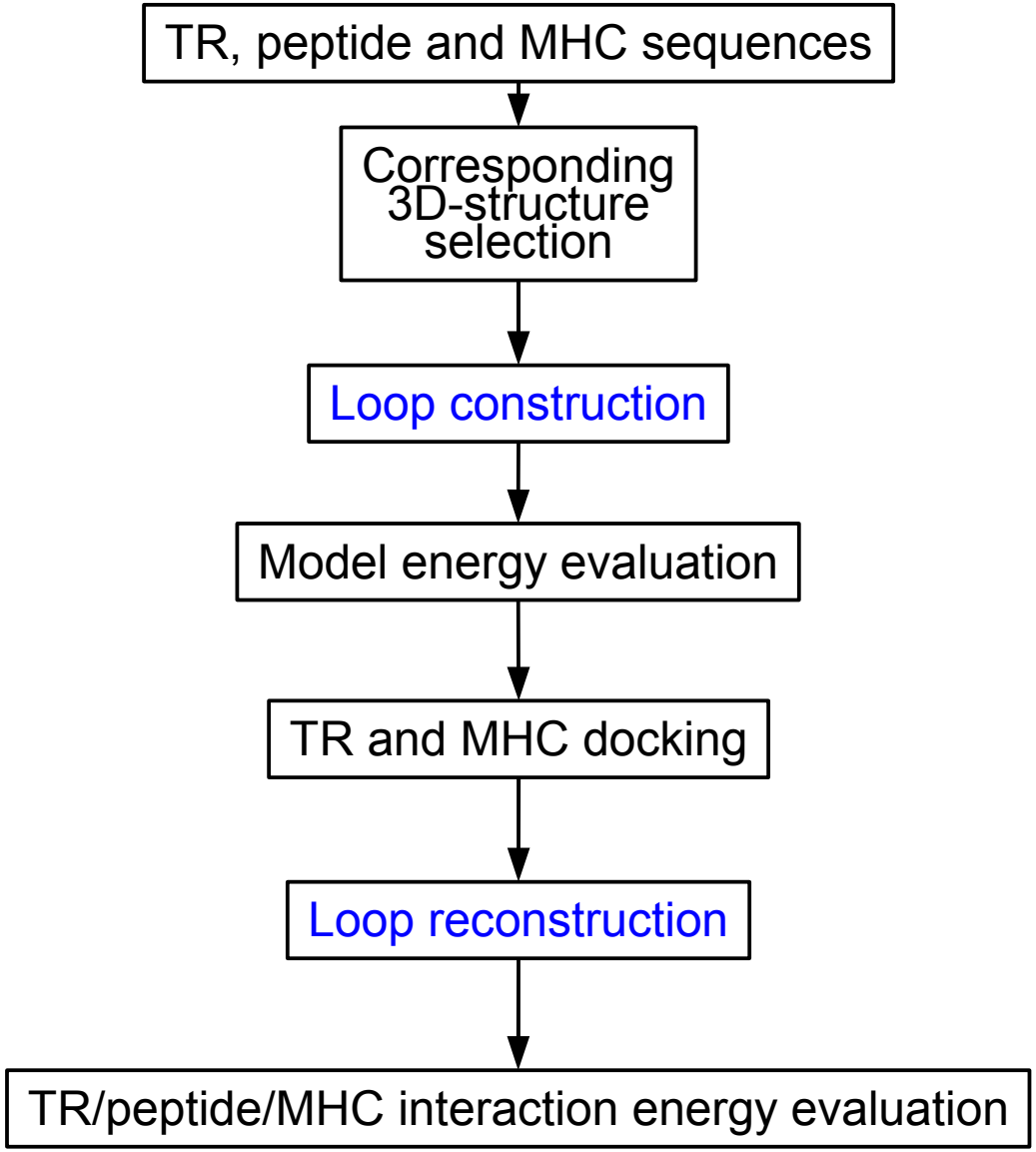


MHC-II



Im Muno Gene Tics Information system®

TR/peptide/MHC interaction model protocol

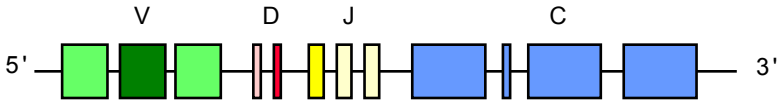


IG synthesis: a rich diversity

150

FUNCTIONAL IG GENES

HEAVY CHAIN

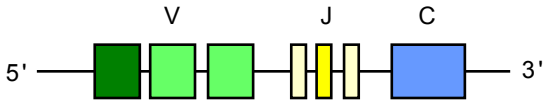


38-46 x 23 x 6

POTENTIAL RECOMBINATIONS

6300

LIGHT CHAIN



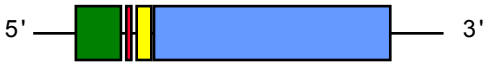
30-35 x 5 (Kappa)
29-33 x 4-5 (Lambda)

POTENTIAL RECOMBINATIONS

185 + 165

N-DIVERSITY SOMATIC MUTATIONS

x1000

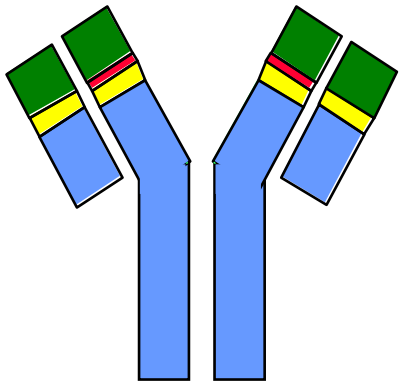


ABOUT 6.3×10^6 POSSIBILITIES

ABOUT 3.5×10^5 POSSIBILITIES

2×10^{12}

DIFFERENT ANTIBODIES



IG and TR not dealt in generalist databases owing to their particular structure

Different ways to consider:

- synthesis, we know the allele components and try to imagine the possible resulting structure(s)

- analytic, we try to solve the protein history as we know the structure (also in terms of AA sequence) and try to see what may be the possible origin alleles.

But we have to consider model for both to frame the search. IMGT-ONTOLOGY is a way to have a formal construction (in the sens of formal language) "describing" such we can talk about sequence in different approaches () in term of combination of components to see how it fit.

What can be wrong with a structure complex?
May be due to 'bad' alleles or synthesis?

Chiffres :

Combien de TR et de MHC pour l'homme ?

Quand utilise-t-on la structure exacte vs. les reconstructions ?