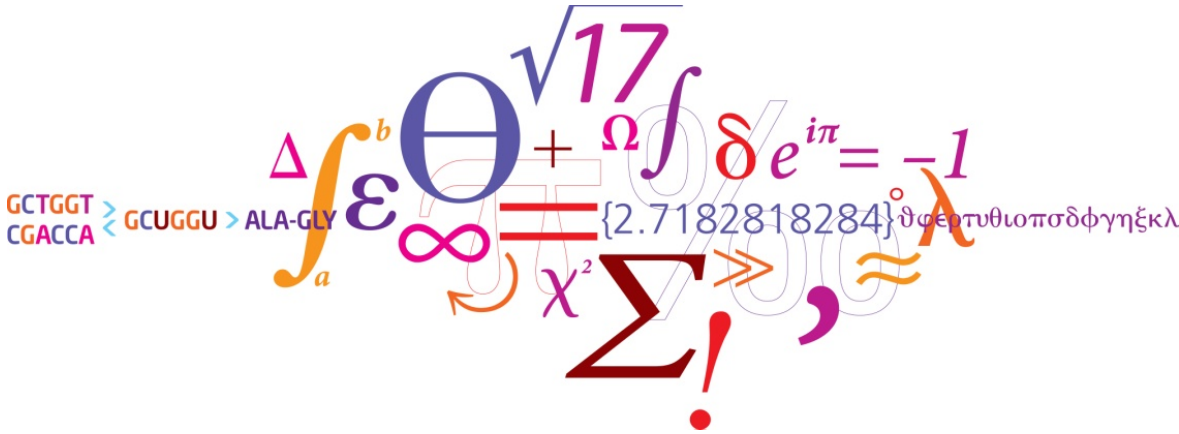
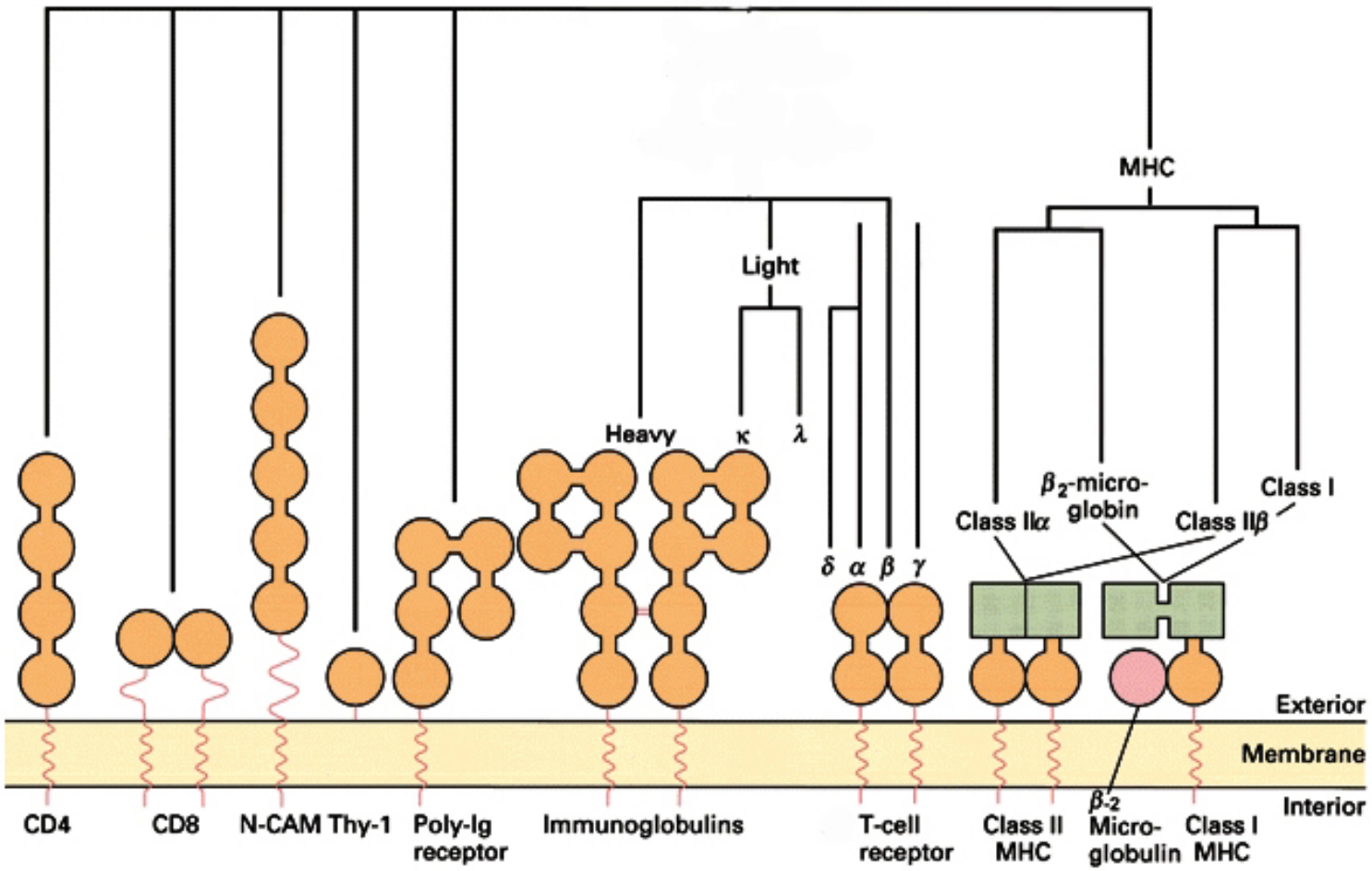
Antibody and TCR structure

Paolo Marcatili

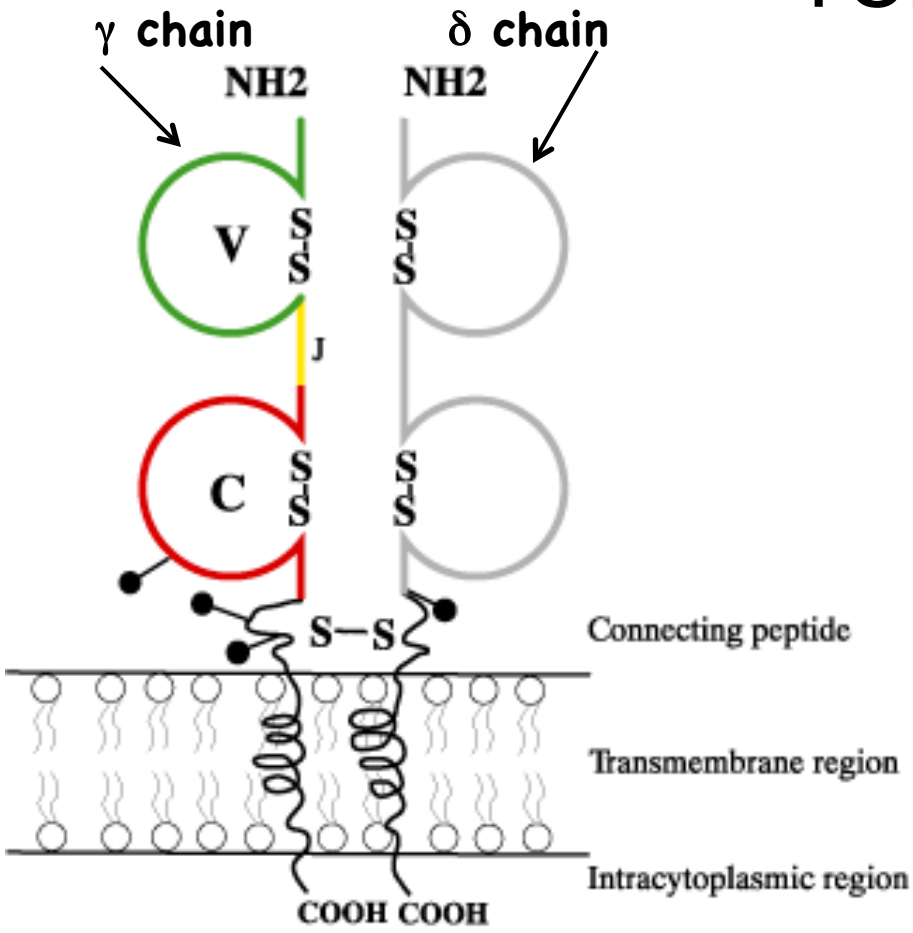


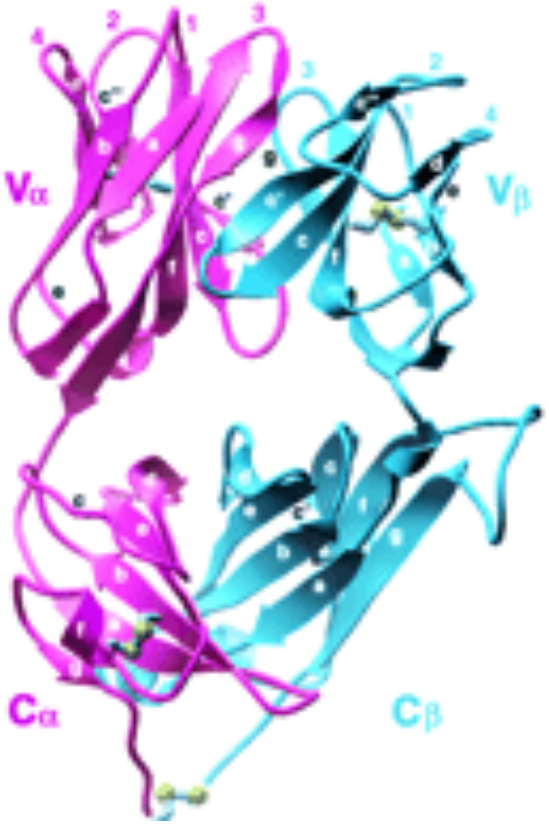
DTU Biosys
Department of Systems Biology

The Ig superfamily

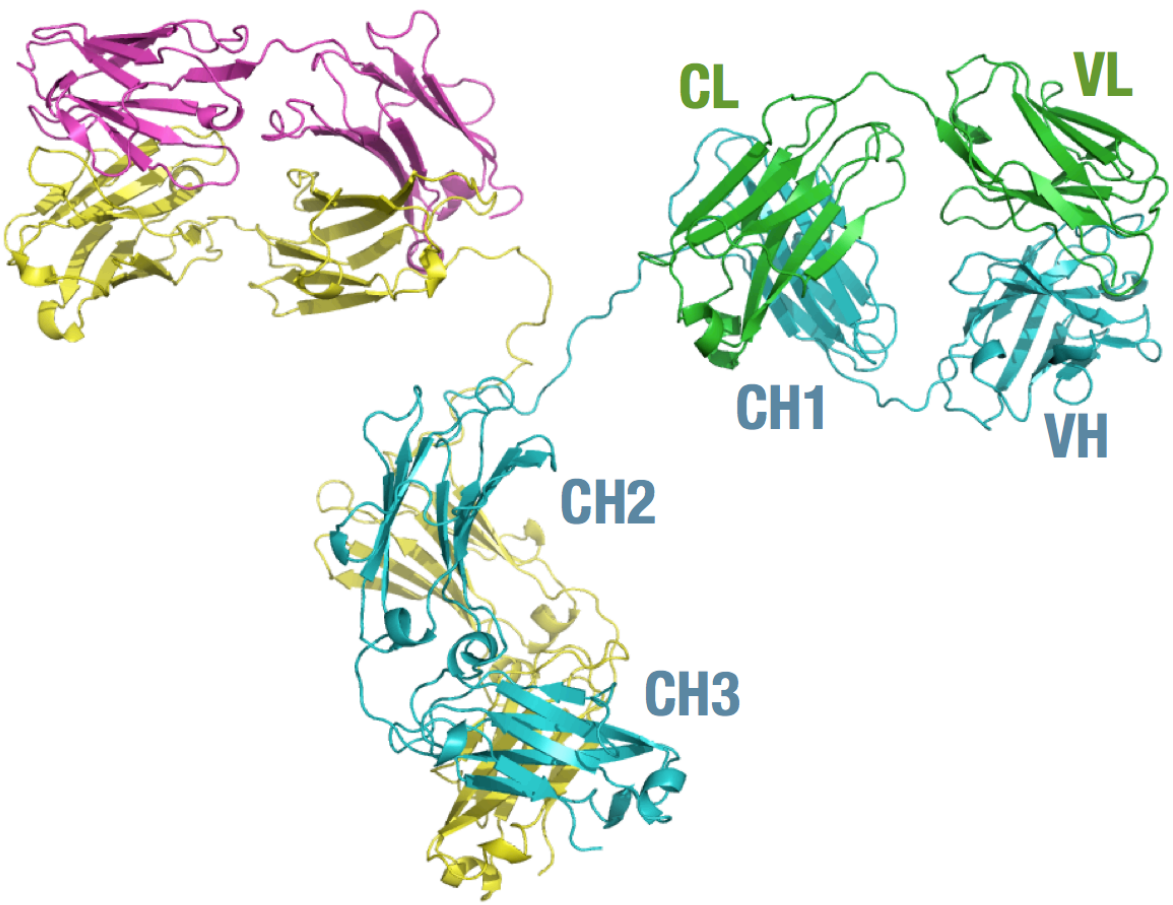


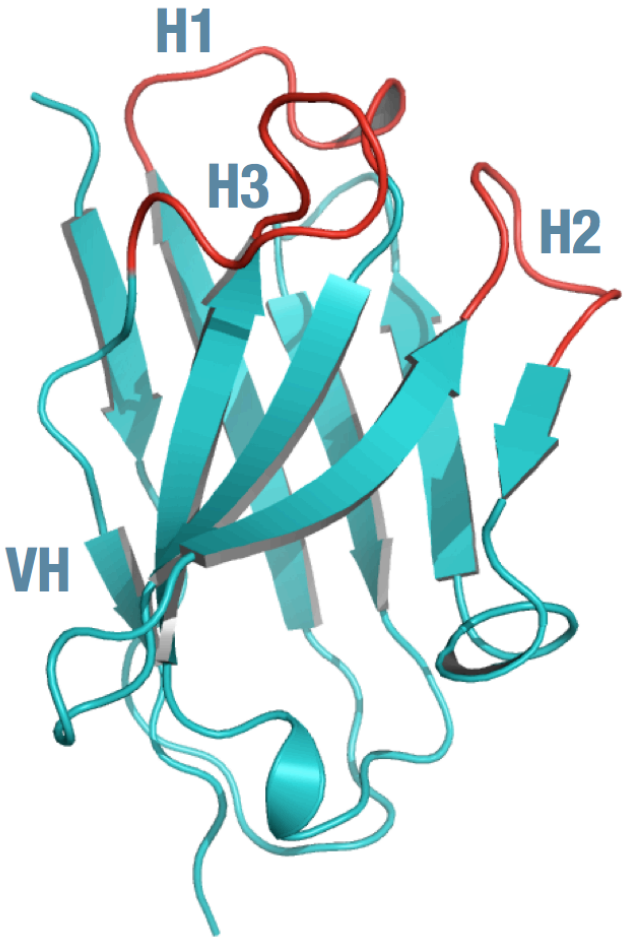
TCRs


γ chain
δ chain
NH2
NH2
V
S
S
S
S
J
C
S
S
S
S
S—S
Connecting peptide
Transmembrane region
Intracytoplasmic region
COOH COOH


Vα
Vβ
Cα
Cβ

Basic properties

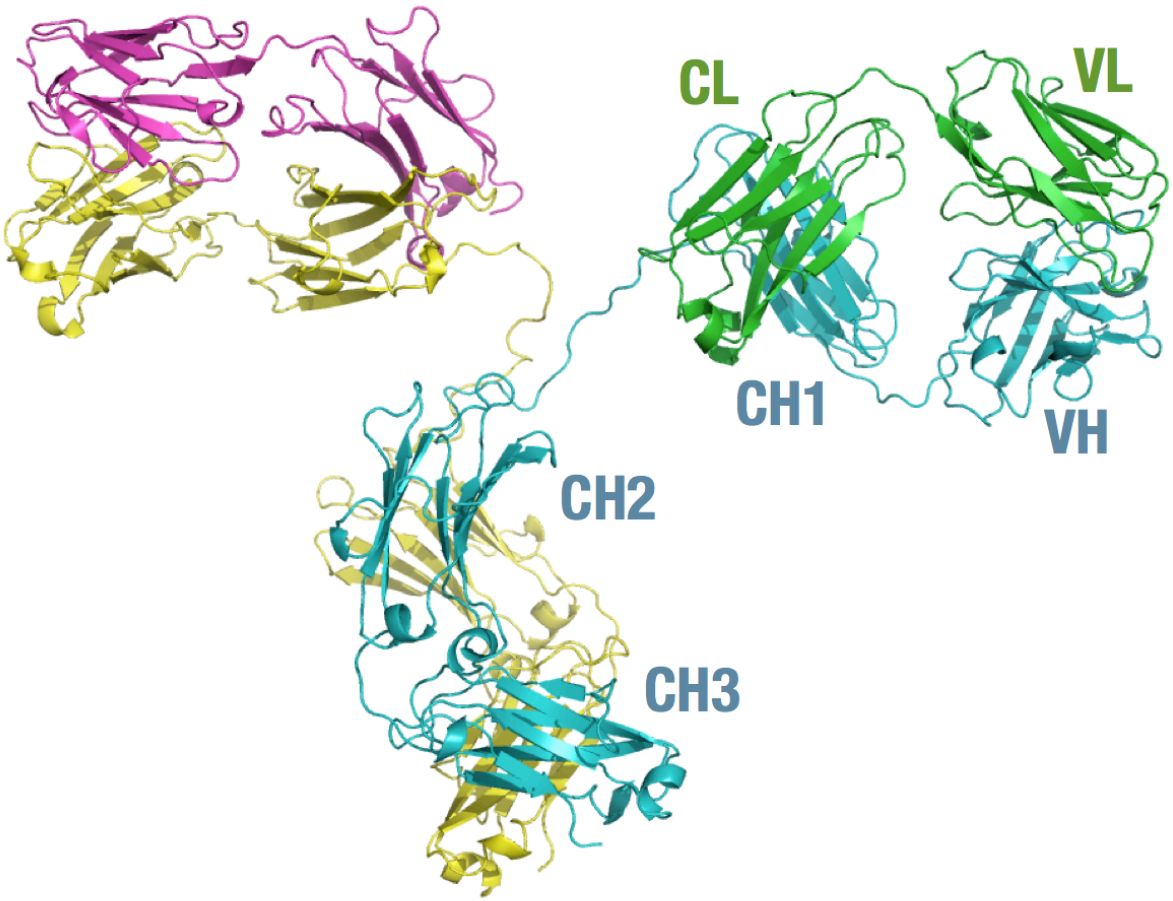


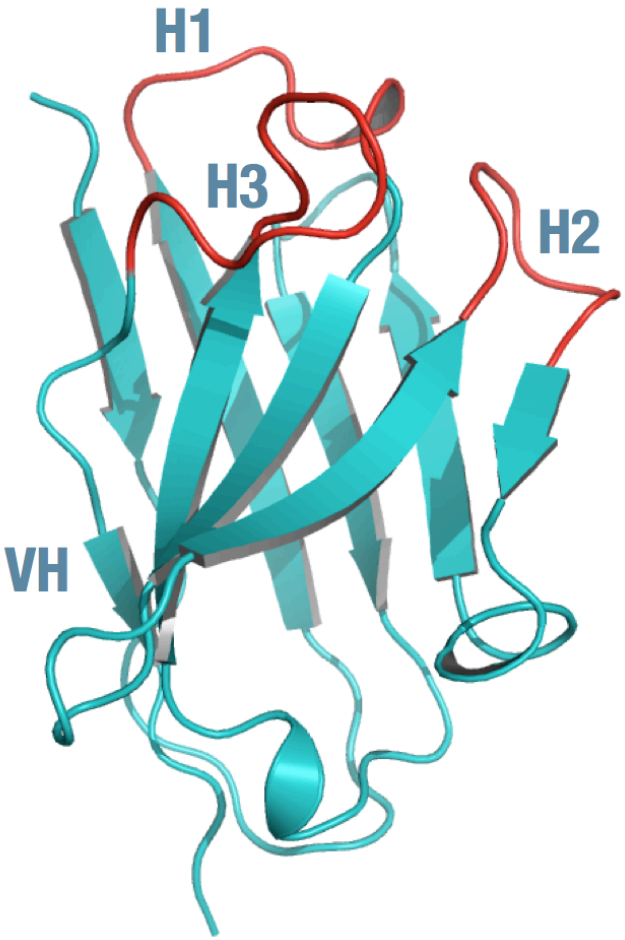


Basic properties

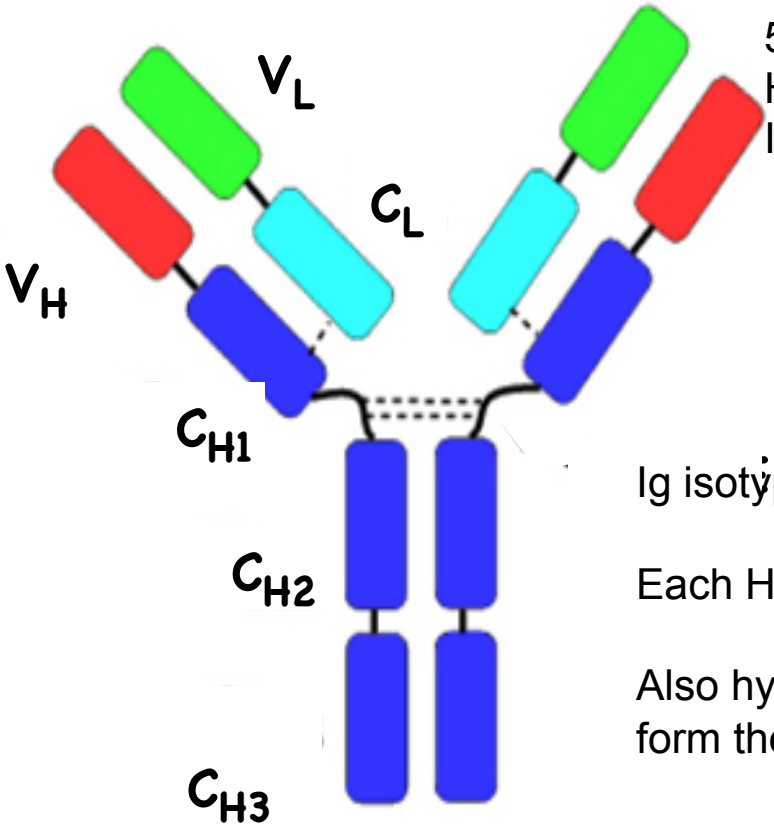
- Beta sandwich
- 2 light/alpha, 2 heavy/beta
- held together by disulphide

- 1 variable domain, 1 to 5 constant domains
- variable = antigen specificity
- constant = effector function





Ab STRUCTURE, cont.



5 types of heavy chain: μ, δ, γ (4 subtypes), ε, α.
H chain defines <u>class</u> (isotype) of Ig: IgM, IgD, IgG, IgE, IgA

Ig isotypes play different roles in immune responses.

Each H or L chain composed of a C and a V region.

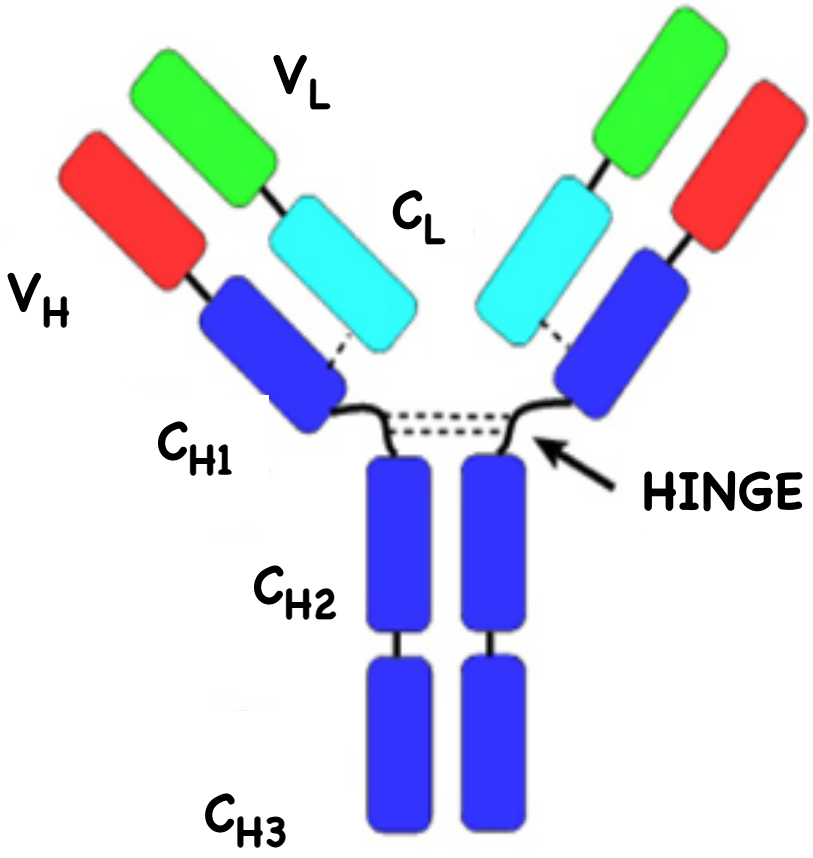
Also hypervariable regions interspersed throughout V. These form the 3-D structure of the Ag binding site.

CLASSES AND SUBTYPES OF HUMAN Ig

property	Ig class or subtype							
	IgM	IgG1	IgG2	IgG3	IgG4	IgA1,2	IgE	IgD
form	pentamer	mono	mono	mono	mono	dimer	mono	mono
serum level (mg/ml)	1.5	9	3	1	0.5	3.5	0.00005 #	0.03
Complement activation	+++	+++	+	+++	-	-	-	-
Placental transfer	-	+	+	+	+	-	-	-
macrophage (Fc receptor) binding	-	+	-	+	-	-	-	-
present in external secretions	mucus* etc	milk *	milk *	milk *	milk *	mucus* etc	-	-
mast cell/ basophil degranulation	-	+*	-	-	-	-	+++*	-

*(species-dependent) # value is for non-allergic individuals

Ab STRUCTURE, cont.



H have 4 domains
Amino terminal variable domains (V_H) at the tips of the Y

3 constant domains: CH1, CH2, and the COOH terminal CH3, at the base of the stem

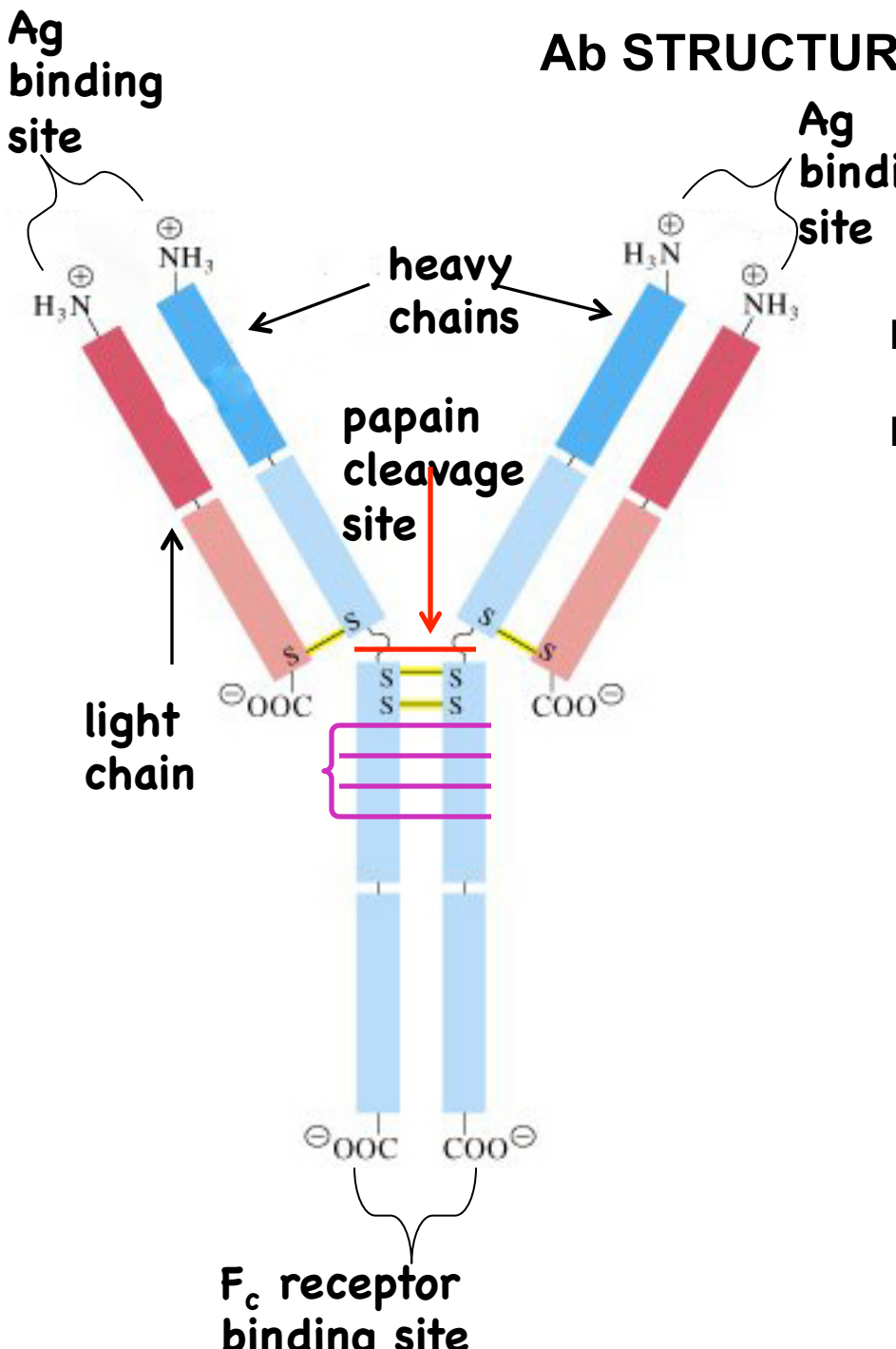
Switch connects the constant and variable regions

Hinge connects CH2 and CH3 (the F_c fragment) to the rest of Ig F_{ab} fragments).

L chains - 2 domains, variable (V_L) & (C_L), connected by a switch.

Individual B cells synthesize Abs of a single specificity (same H, same L).

Ab STRUCTURE, cont.

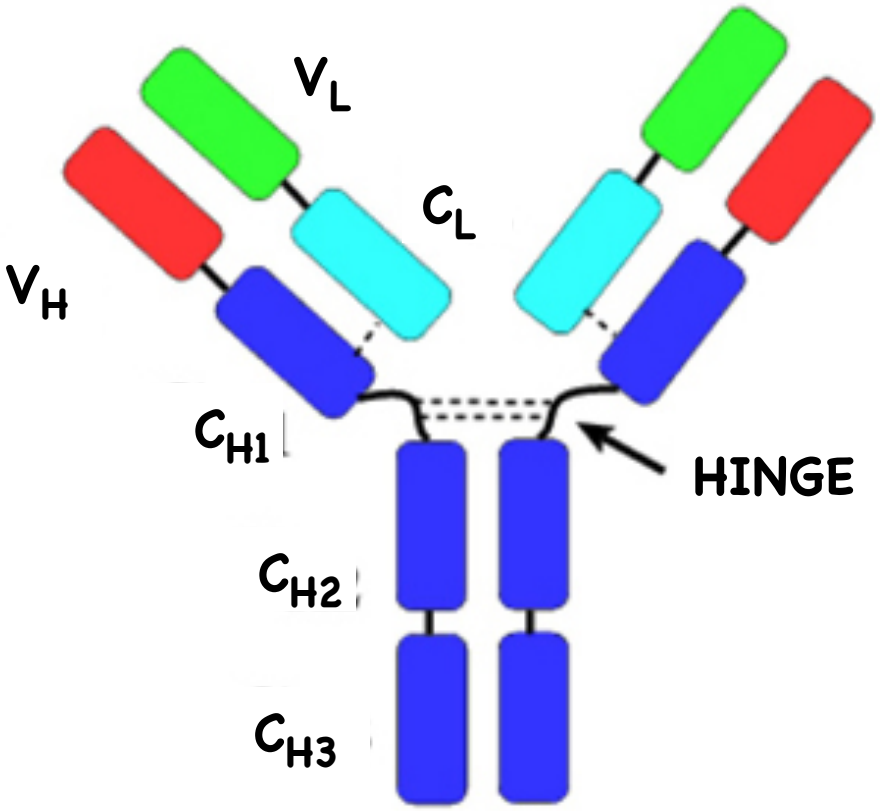


Cleavage of Ig with papain yields 3 fragments of 2 types:

Fab: has antigen (Ag) binding site.

Fc: C = crystal; this portion of Ig 1st to be crystallized. F_c binds to cell surface F_c receptors, augmenting immune responses (opsonization, e.g.).

Ab STRUCTURE, cont.



V of both L and H Ig chains contain 3 hypervariable regions, or complementarity-determining regions (CDRs).

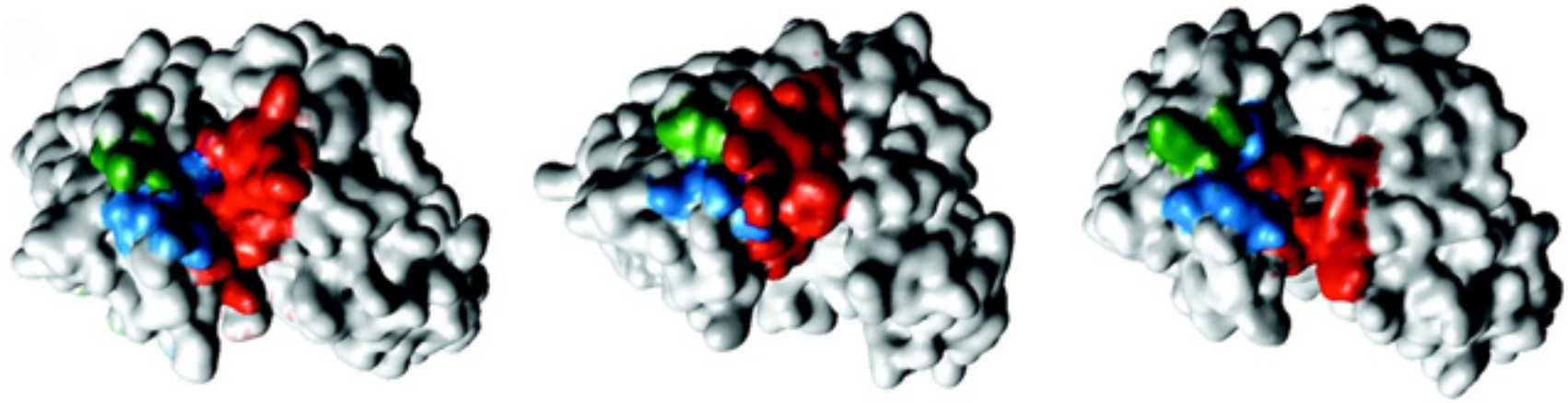
CDRs = loops connecting β strands in Ig fold

Residues in CDRs vary from one Ab to the next, **imparting Ag specificity to each Ab**.

V_L and V_H domains at tips of Abs packed so that 6 CDRs (3 on each) form surface for Ag-specific binding.

Residues from all 6 CDR's (V_L CDR1, CDR2, CDR3 & V_H CDR1, CDR2, CDR3) project from distal surface of Ab tip to bind Ag.

EFFECT OF SEQUENCE ON Ab BINDING SITE & Ag SPECIFICITY



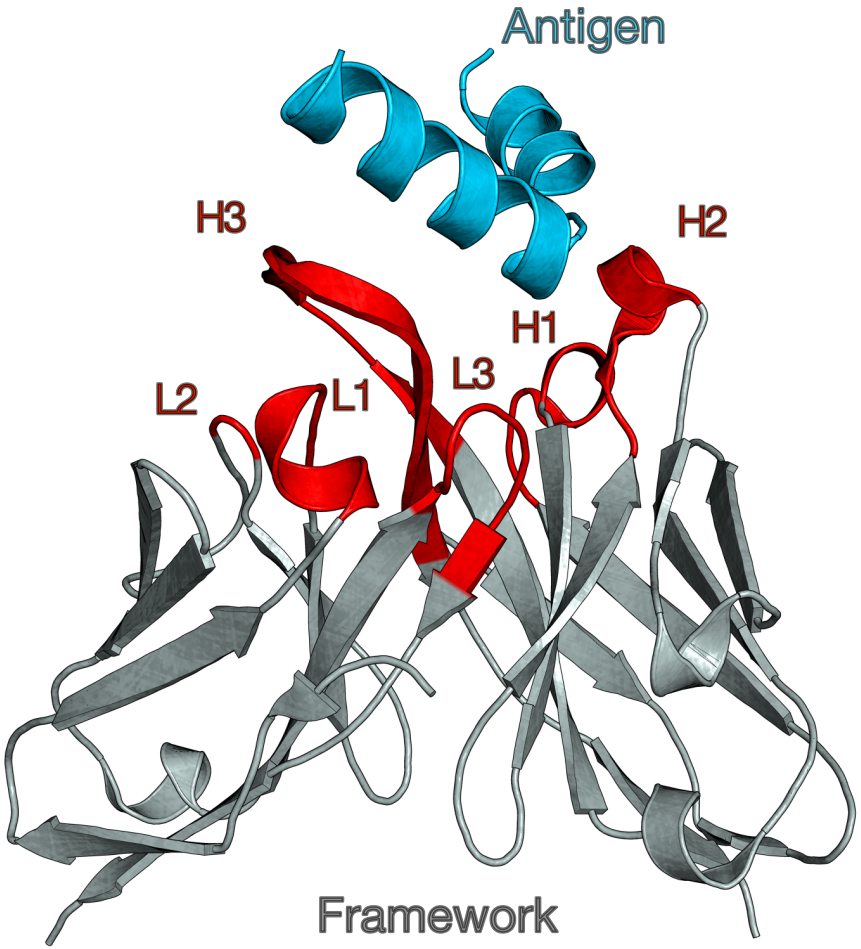
from Lantto, J., et al. J. Biol. Chem. 2002; 277: 45108-45114.

red = CDHR3; blue = CDHR2; green = CDHR1

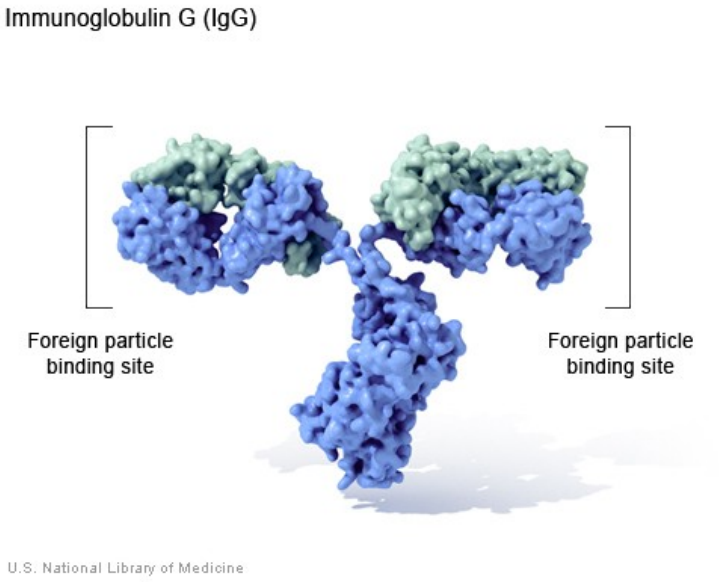
CDHR = complementarity-determining region of H chain (hypervariable region)

Effect of insertions/deletions on Ag binding

Antibody Binding

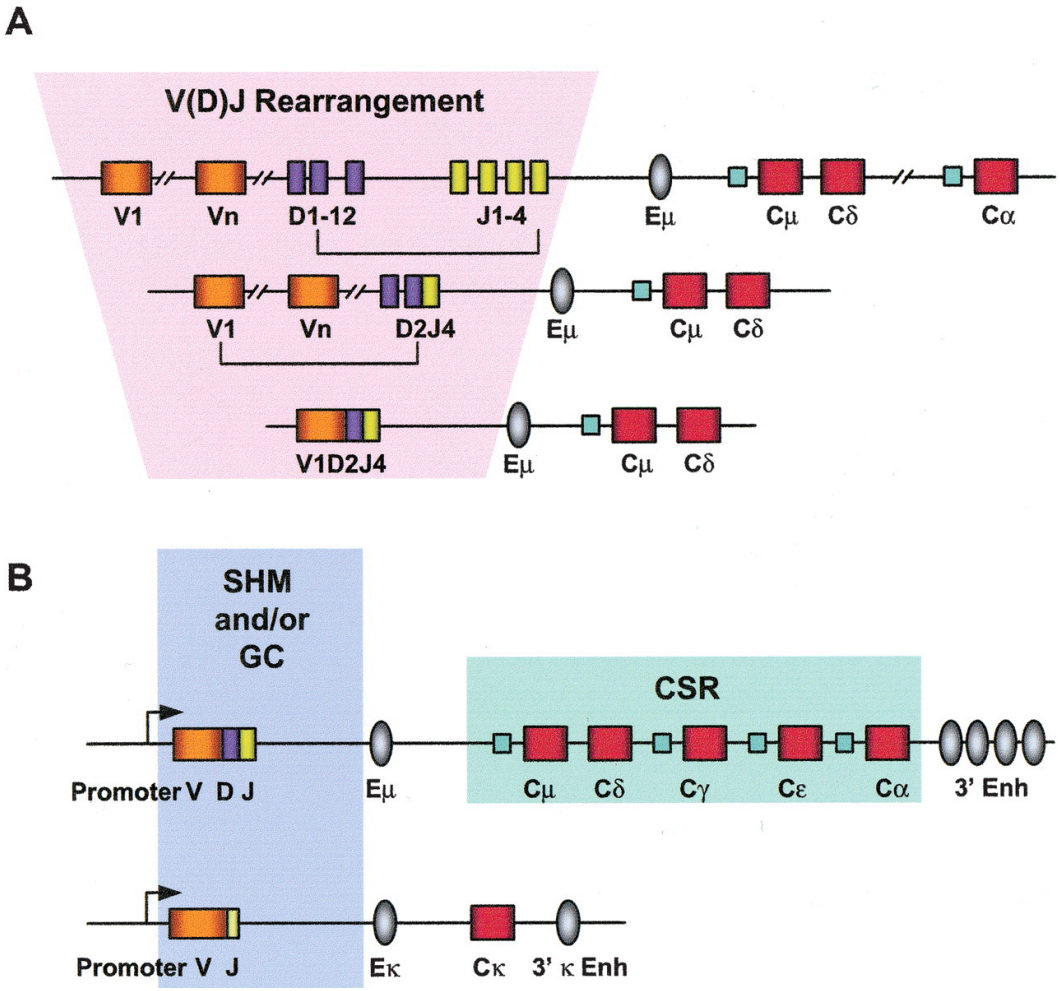


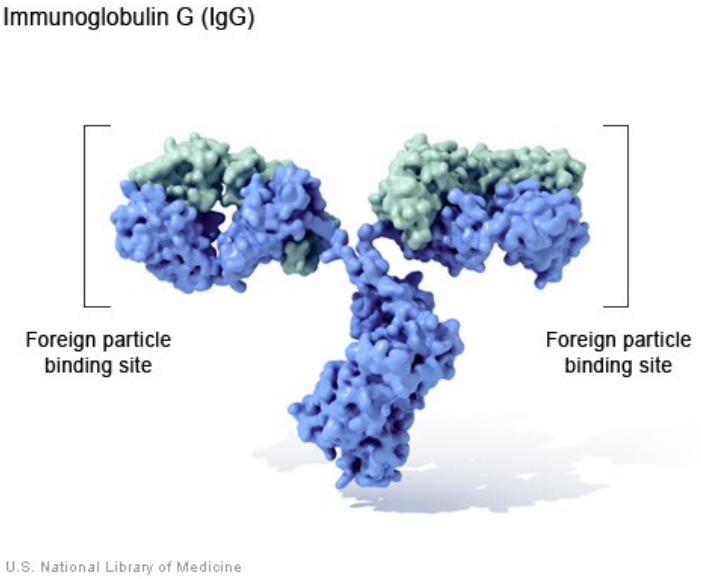
Immunoglobulins



The generation of antibody diversity through somatic hypermutation and class switch recombination
Ziqiang et al., G&D 2004

Immunoglobulins





The generation of antibody diversity through somatic hypermutation and class switch recombination
Ziqiang et al., G&D 2004

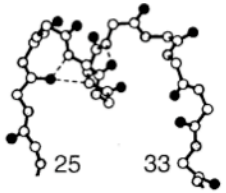
Hypervariable regions

```
EVQLVESGGGLVQPGGSLRLSCAASGFTFSN----YEMNWVRQAPGKGLEWISYISN----GDNTIYYADSVKGRFTISRDSAKNSLYLHMHSLRAEDTAVYYCARGDYGGNGYFYYYAM----------DVWGQGTTVTVSS
EVQLVESGGGLVQPGRSLRLSCTASGFTFGD----YAMSWFRQAPGKGLEWVGFIRSKA--YGGTTEYAASVKGRFTISRDDSKSIAYLQMNSLKTEDTAVYYCTRDRIGNYDFWSGYYTGV---------GYWGQGTLVTVSS
-VQLVQSGAEVKKPGSSVKVSCKASGGTFSS----YAISWVRQAPGQGLEWMGGIIP----IFGTANYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARVWGGSGSYYIWF-------------DPWGQGTLVTVSS
QVQLVQSGAEVKKPGASVKVSCKASGYTFTG----YYMHWVRQAPGQGLEWMGWINP----NSGGTNYAQKFQGRVTMTRDTSISTAYMELSRLRSDDTAVYYCAREQWLVLEHYF--------------DYWGQGTLVTVSS
-VQLVQSGAEVKKPGSSVKVSCKASGGTFSS----YAISWVRQAPGQGLEWMGGIIP----IFGTANYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCATKNDFWSGYYEGYYYYYYM-------DVWGKGTTVTVSS
QVQLVQSGAEVKKPGSSVKVSCKASGGTFSS----YTISWVRQAPGQGLEWMGRIIP----ILGIANYAQKFQGRVTITADKSTSTAYMELSSLRSEDTAVYYCATNYDFWSGY-----------------PYWGQGTLVT---
EVHLVESGGGLVQPGGSLRLSCAASEFTFDR----YWMSWVRQAPGKGLEWVANIKE----DGSEKKYVDSVRGRFTISRDNAKNSLVLQMNSLRAEDTAVYYCARGAYYF-------------------GYWGQGALVTVSS
```

HV loops (ABS) model unreliable?

Canonical Structures

Repertoire of discrete conformations
Key residues



Torsion angles:

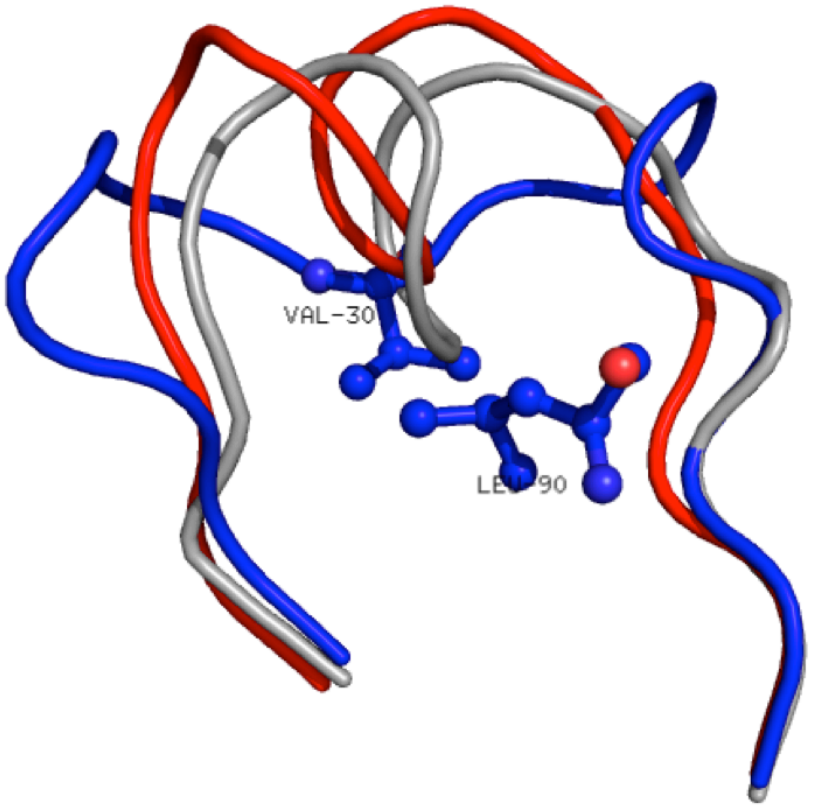
Residue	Angles			
25	Φ	-120	Ψ	-135
26	Φ	-149	Ψ	-180
27	Φ	-55	Ψ	-23
28	Φ	-84	Ψ	-19
29	Φ	-127	Ψ	-88
30	Φ	-61	Ψ	-38
30a	Φ	-74	Ψ	7
30b	Φ	-116	Ψ	26
30c	Φ	70	Ψ	37
31	Φ	-109	Ψ	148
32	Φ	-71	Ψ	157
33	Φ	-119	Ψ	125

Hydrogen Bonds:

26	O	- - -	N	29
26	O	- - -	N	30a
29	O	- - -	N	30b

Structure	PDB file	Residues
NEW	7fab	GSSSNIGAGHNV

Figure 7. λ L1 canonical structure 2.

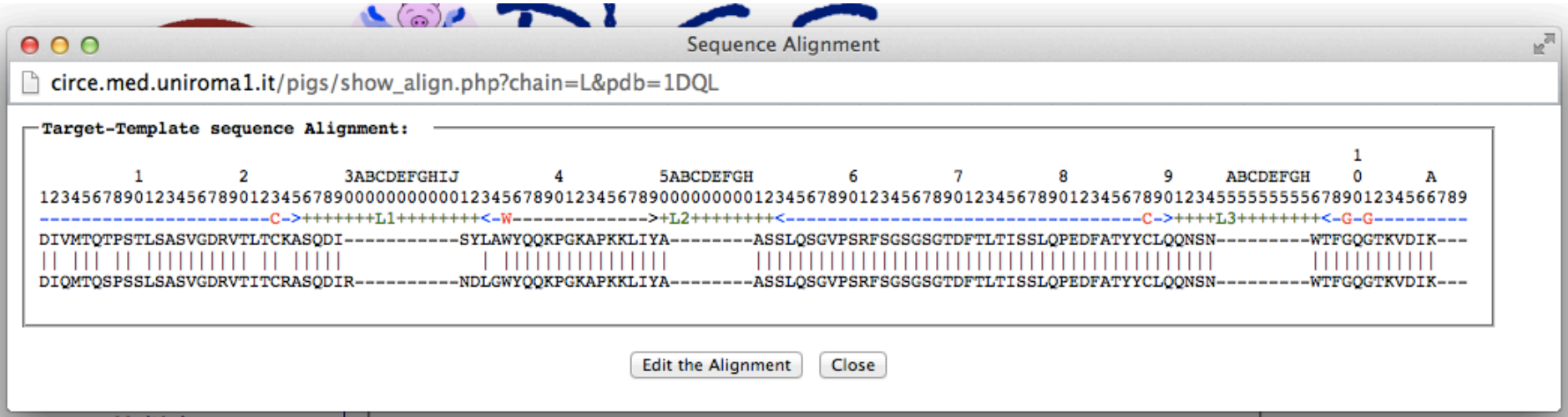


Loop	CS	Length	Constraints
κL1	1	6	29 VIL
	2	7	29 VIL
	3	13	29 VIL
	4	12	29 VIL
	5	11	29 VIL
	6	8	29 VIL
κL2	1	3	
κL3	1	6	90 QNH 95 P
	2	6	90 Q 94 P
	3A	5	90 Q 91/92 G 96 P
	3B	5	90 Q 91/92 ST 96 P
	4	4	90 Q
	5	7	90 Q 95A P
	6	5	90 Q 94 L
	7	5	94 P
	8	8	90 Q 95A P
λL1	1	10	25 G
	2	11	25 G 31 FHY 66 K 90 S
	3	11	66 L 90 L
	4	8	28 VIL 66 ST
	5	10	25 R 28 G
	6	11	25 G 31 ND 66 K 90 S
	7	8	28 VIL 66 N
	8	9	
λL2	1	3	
	2	7	
λL3	1 form A	6	
	1 form B	6	
	1 form C	6	
	1 form D	6	
	2	8	
	3	7	92 D 95 ST
	4	7	95 not ST

H1	1	7	
	2	8	
	3	9	
H2	1	3	
	2	4	71 AVL
	3	4	71 RK
	4	6	
H3	Bulged	>10	94 RK 101D
	Non-bulged	>10	94 not RK
	Short	10	

Canonical Structure prediction

Given an antibody sequence
we can predict the CS of each loop (H3 is kind of an exception)



Multiple Sequences

Authors

BioComputing Group

External links

IMGT Home Page
IG Numberings
Collier de Perles
RosettaAntibody

Light Chain: ?

Loops	Sequence	Structure
11o ?	u14162u1fm	
21o	91.51%	1DQL
21l	82.73%	2R56
21o	82.73%	3NA9
21o	82.73%	3NAA
21o	82.73%	3NAB
21o	82.73%	3NAC
21l	82.73%	3NCJ
21l	79.09%	1JPS
21l	79.09%	1JPT
21l	80%	1BJ1
21l	81.82%	1HEZ
21l	80.91%	3BN9

Heavy Chain: ?

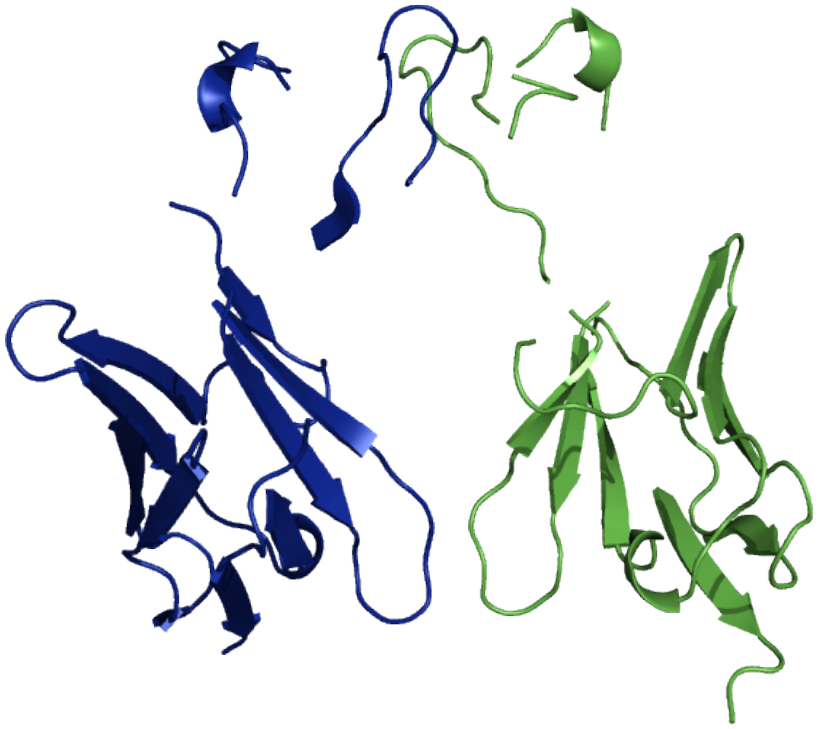
Loops	Sequence	Structure
14B ?	u14162u1fm	
14B	93.39%	1DEE
14B	92.56%	1HEZ
14B	80.99%	2GHW
14B	73.98%	1AQK
14B	77.69%	1NL0
14B	72.66%	3IDX
14B	75.61%	1DQL
13B	74.19%	2QQK
13B	76.03%	2FJG
14B	69.84%	1DFB
14B	74.38%	2GCY
14B	69.53%	3BN9

PIGS

Frameworks (L & H) selection based on %id

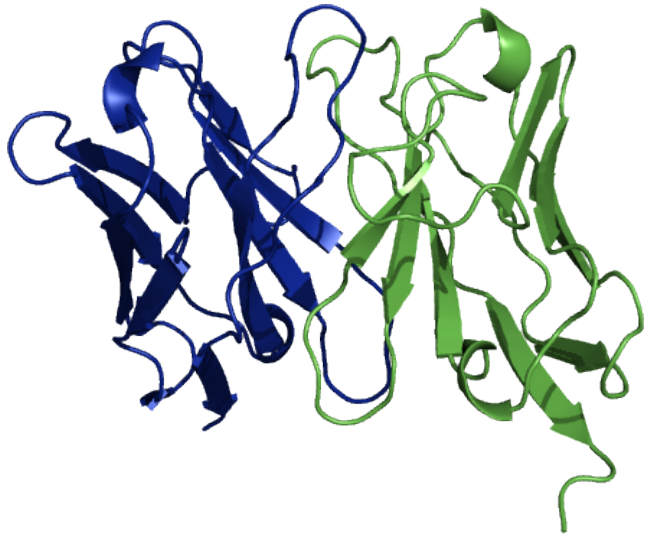
PIGS

Loop selection (L1-3, H1-3) based on CS + %id



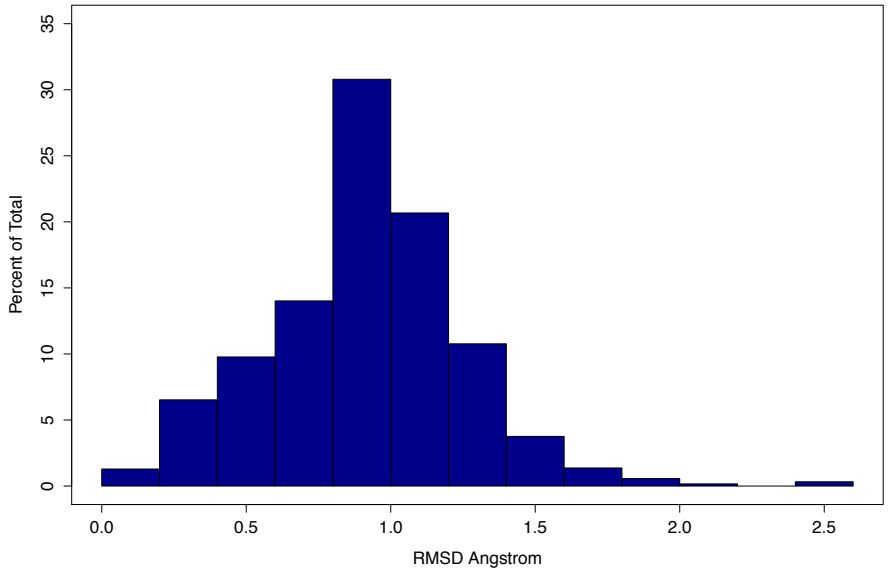
PIGS

VL-VH packing
Superpose loop stems
Side Chains (SCWRL)

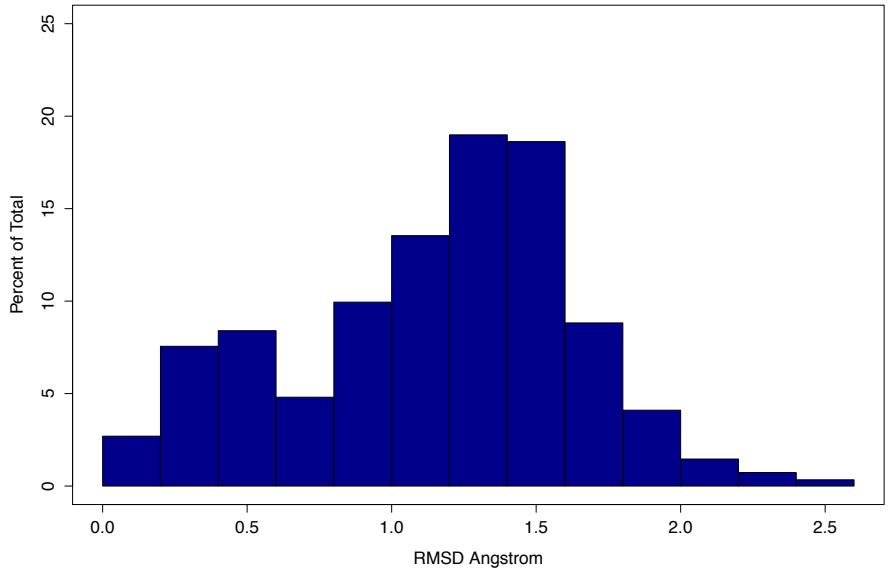


Assessment

Whole Fv

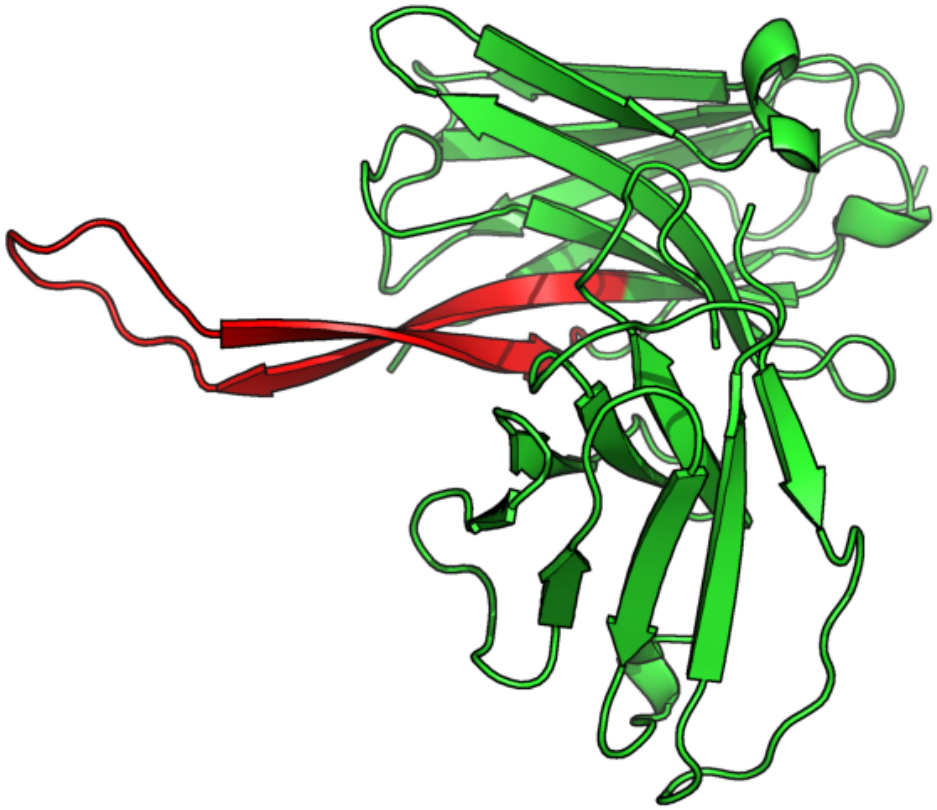


ABS (all CDRs)



H3 Prediction

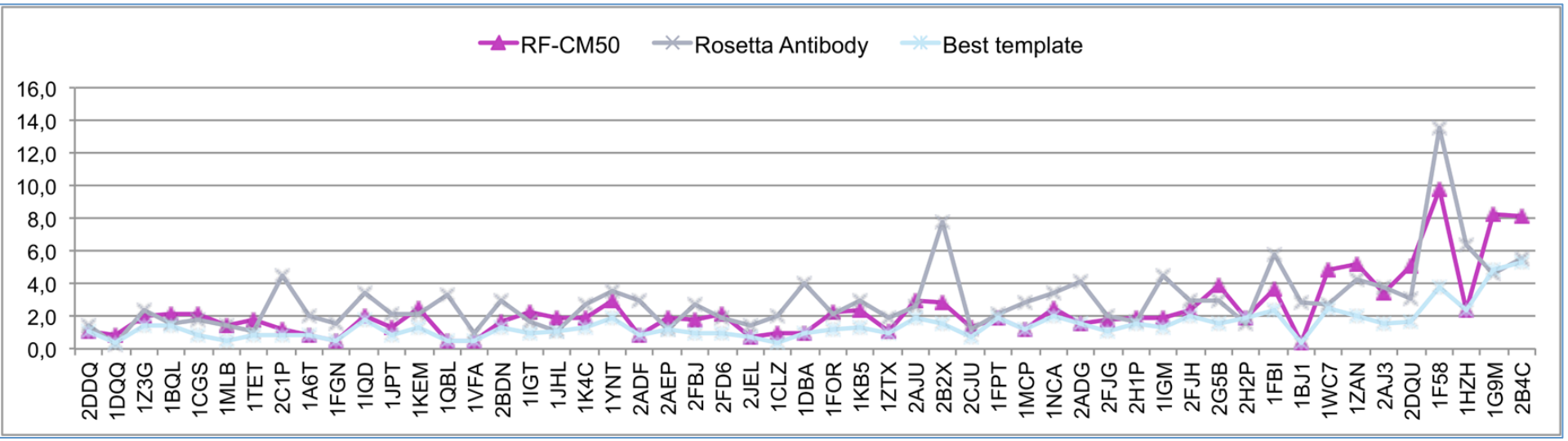
H3 HCDR3) is:
The longest loop (up to 25 residues)
(Almost) no CS
The most flexible
The most important for recognition
The hardest to predict



H3 accuracy

Good (~1Å) for short loops

Bad (up to 6Å and more) for long loops



TCR models

We have > 1500 solved antibodies

TCR: ~50

No method currently available!

Problems in modeling

H3 models are unreliable

Bias on Human and Mouse antibodies

Role of glycosilation

Signal transduction

Take home

The antibody fold is one of the most common
beta-sandwich

CDRs are located on loop regions
extremely variable in sequence, not in structure (backbone!!!)

Canonical structure method

Models are extremely reliable (with the exception of H3)