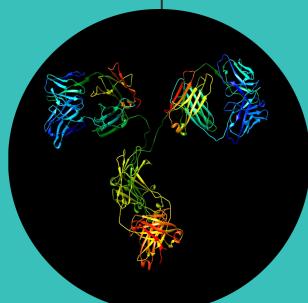


The structural biology of immunoglobulins: immunoglobulin G

Pau Forte, Laia López, Claudia Pallisé, Maria Quintana, Júlia Urgel

● List of contents

1. Basic concepts of immunoglobulins
2. Evolution of IgG chains
3. CDRs: Canonical Structures
4. Characterization of an antigen-antibody interaction
5. Conclusions
6. Multiple Choice Questions
7. References

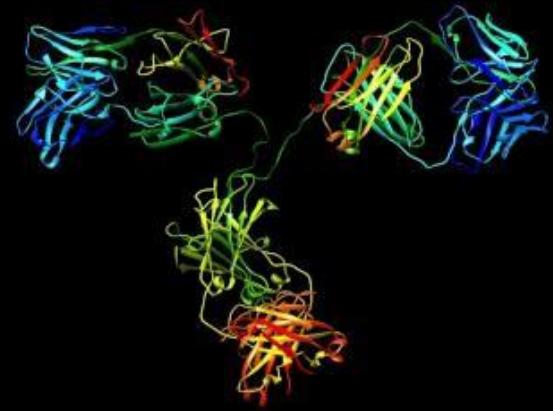


Basic concepts of immunoglobulins

● **Basic concepts:** Introduction

Immunoglobulins

Globular proteins



Functions

Conserved set of effector molecules

- **Activation** and **fixation** of the complement
- **Binding** to Fc receptors on the surface of granulocytes, monocytes, platelets and other components.

11GT

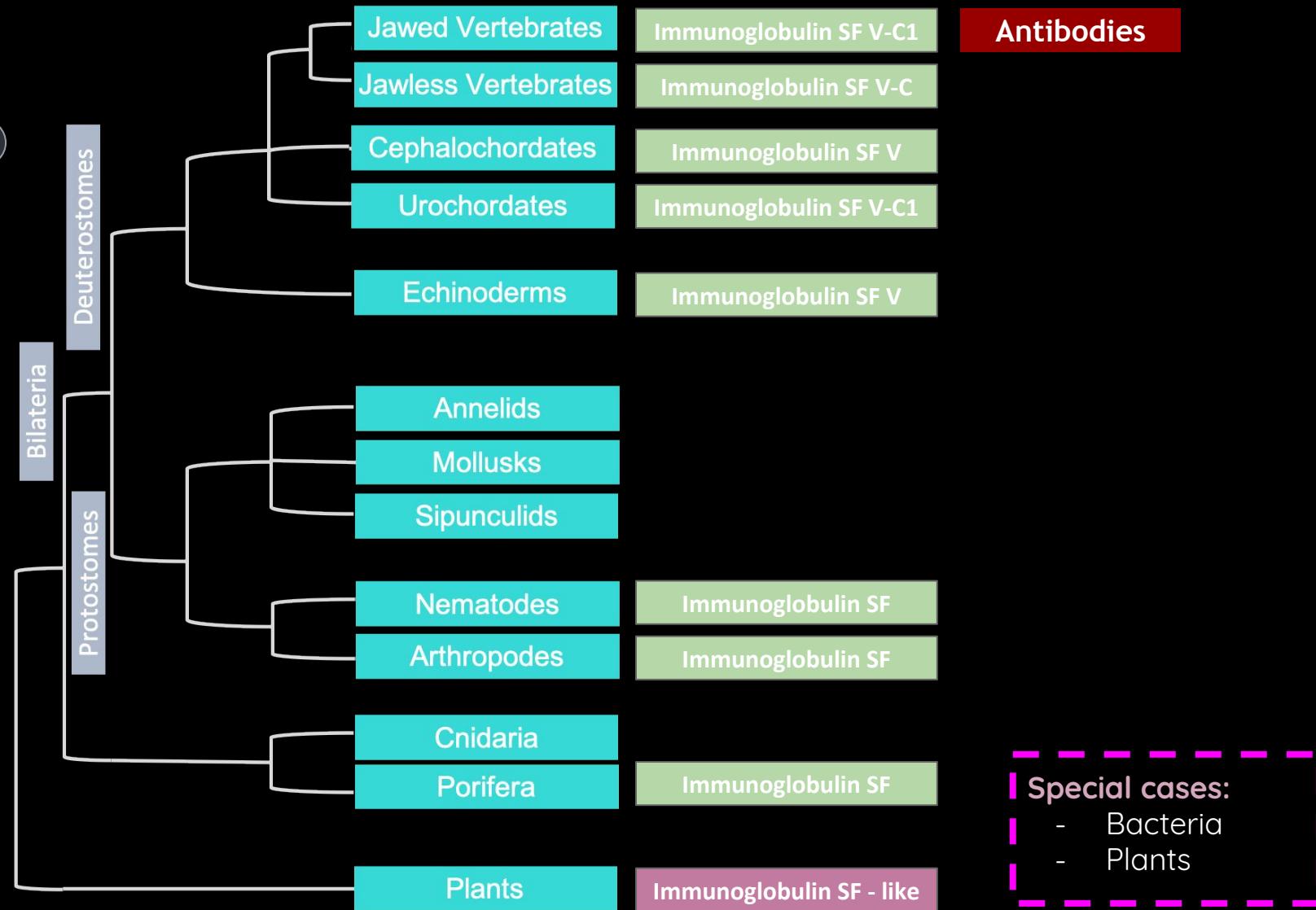
Inflammation

Polyclonal set of diverse ligand binding sites

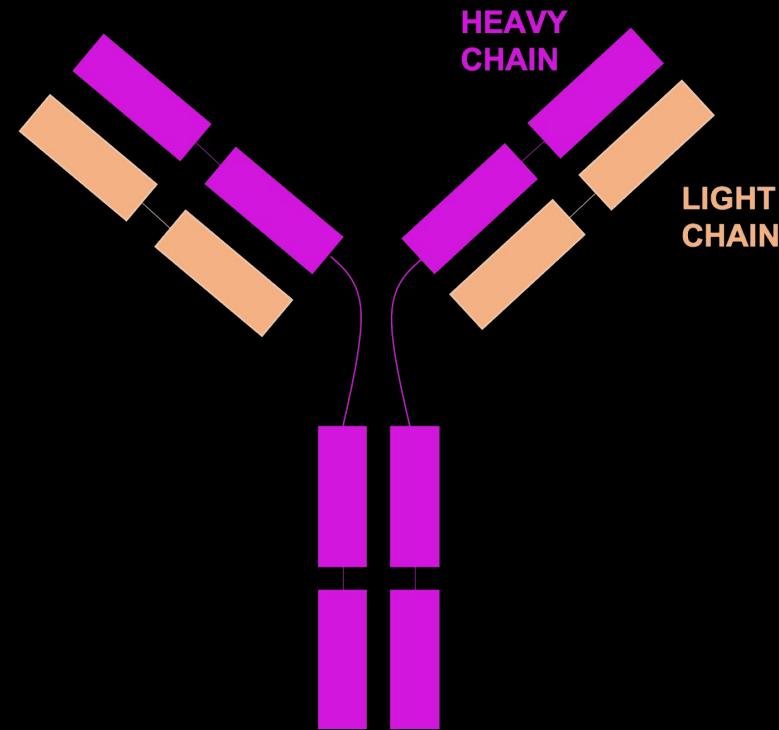
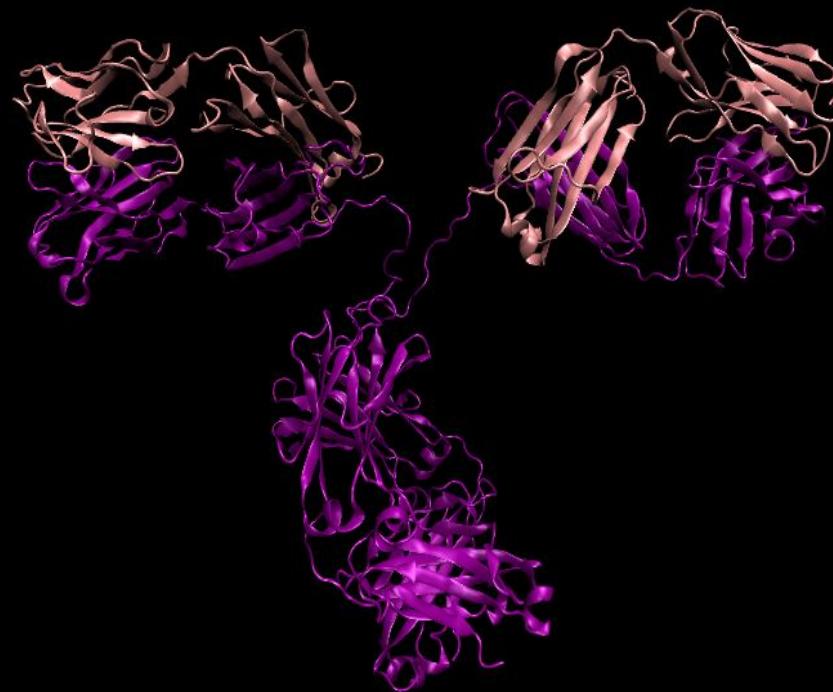
- Recognition of self and non-self-antigens



Basic concepts: evolutionary origin

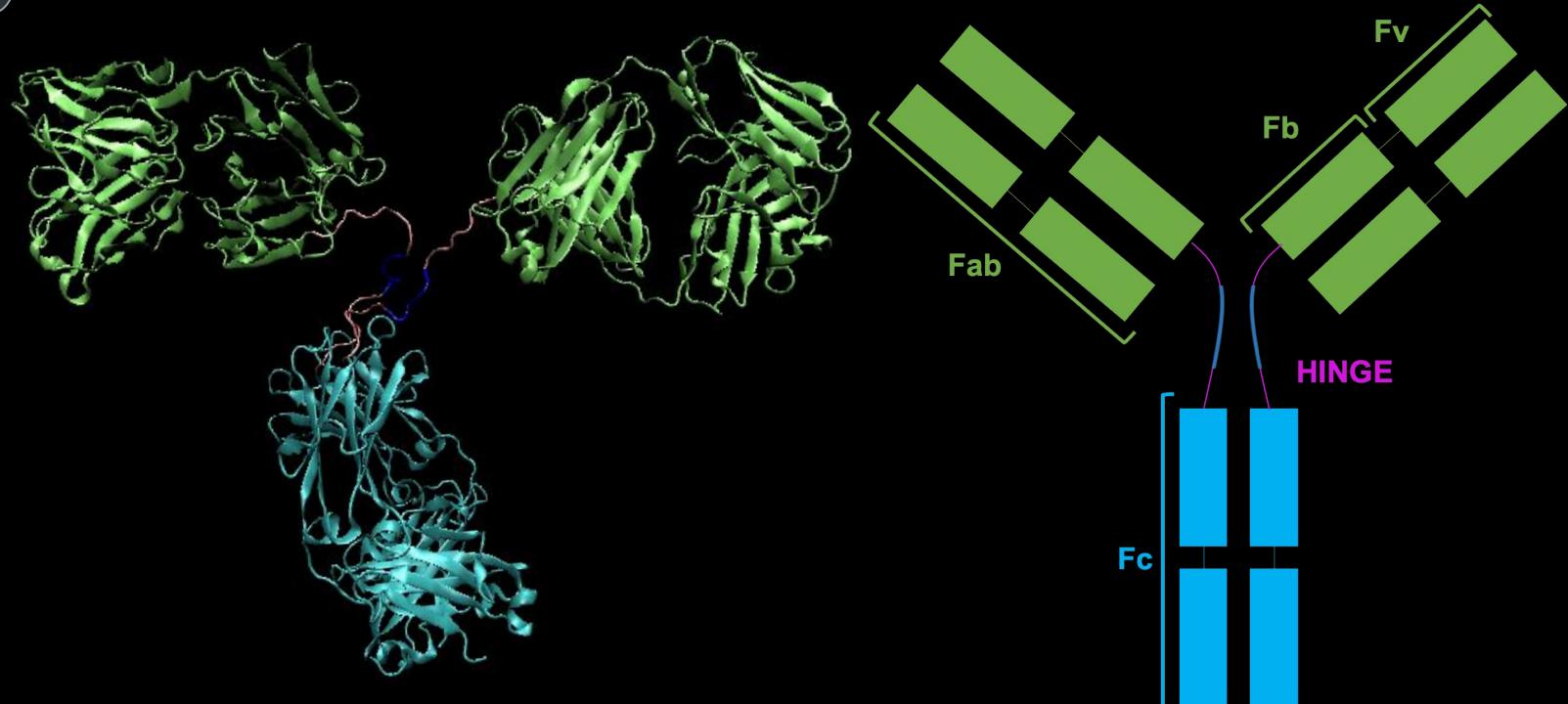


Basic concepts: Introduction - Immunoglobulins as Heterodimers



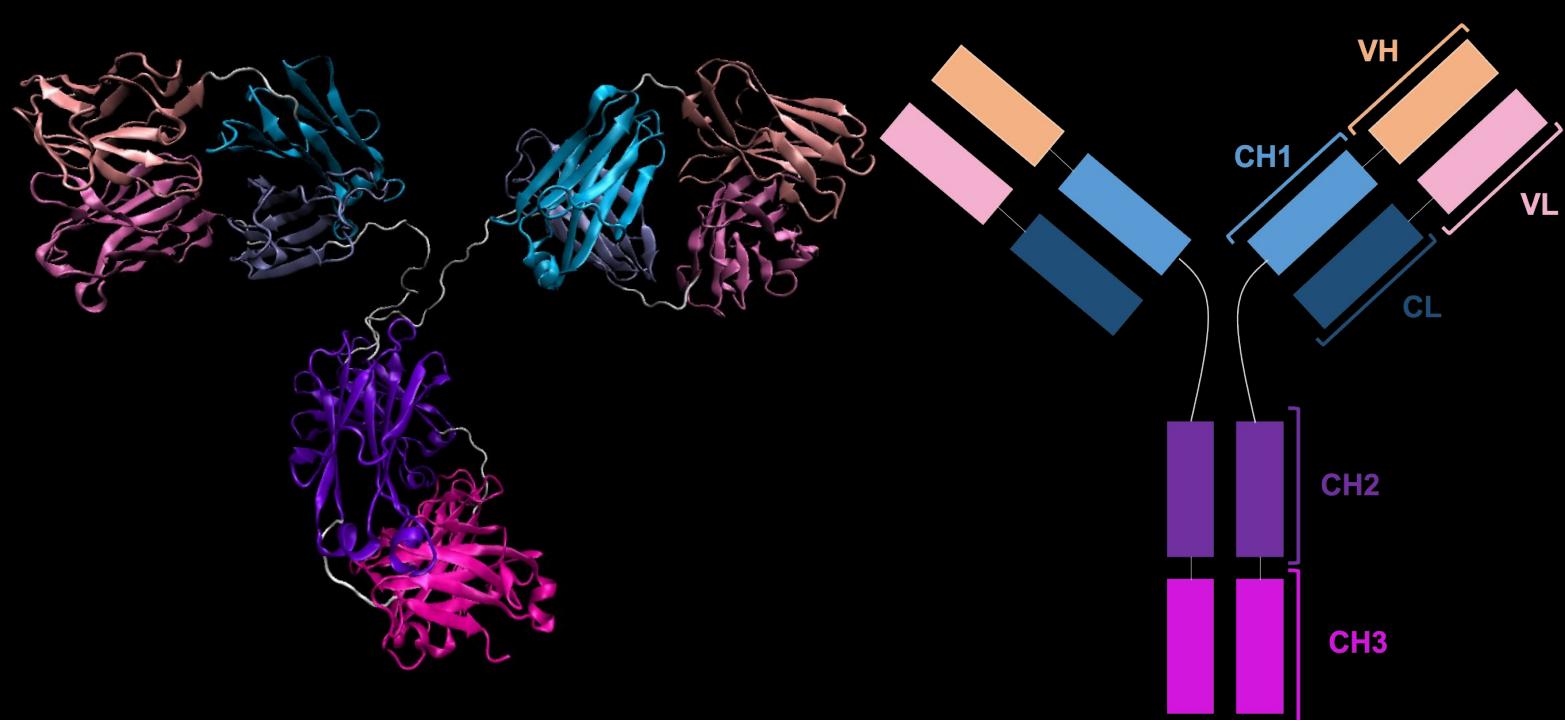
1IGT

Basic concepts: Introduction - Immunoglobulins as Heterodimers



1IGT

Basic concepts: Introduction - Immunoglobulins as Heterodimers



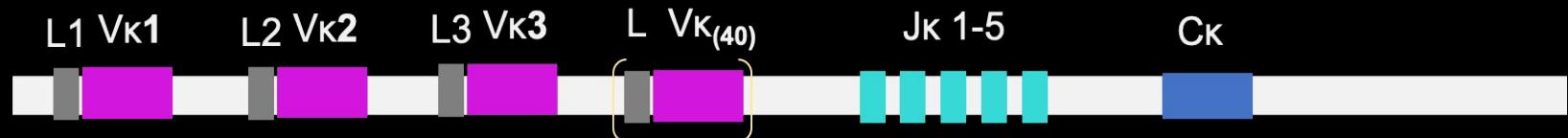
1IGT

● Basic concepts: Diversity

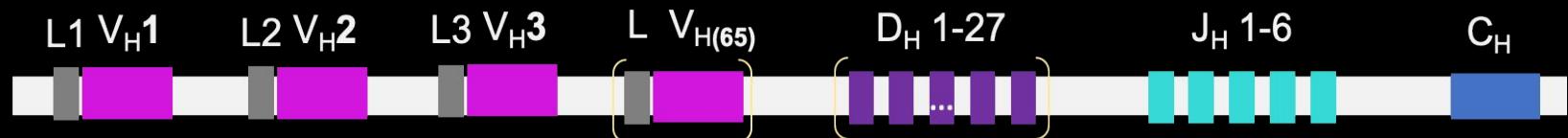
λ light chain locus - Chr 22



κ Light chain locus - Chr 2

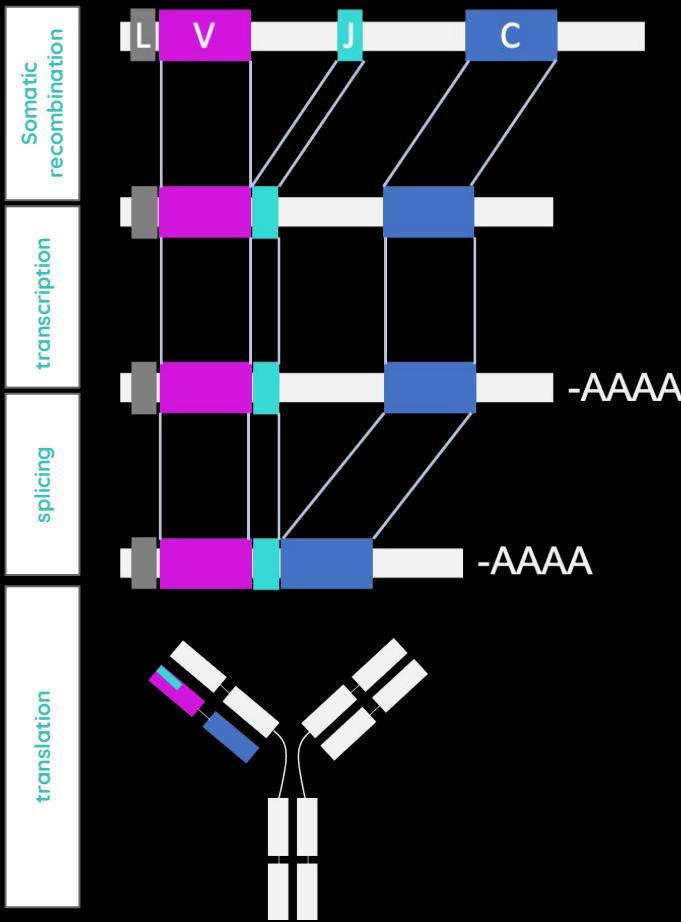


Heavy chain locus -Chr 14

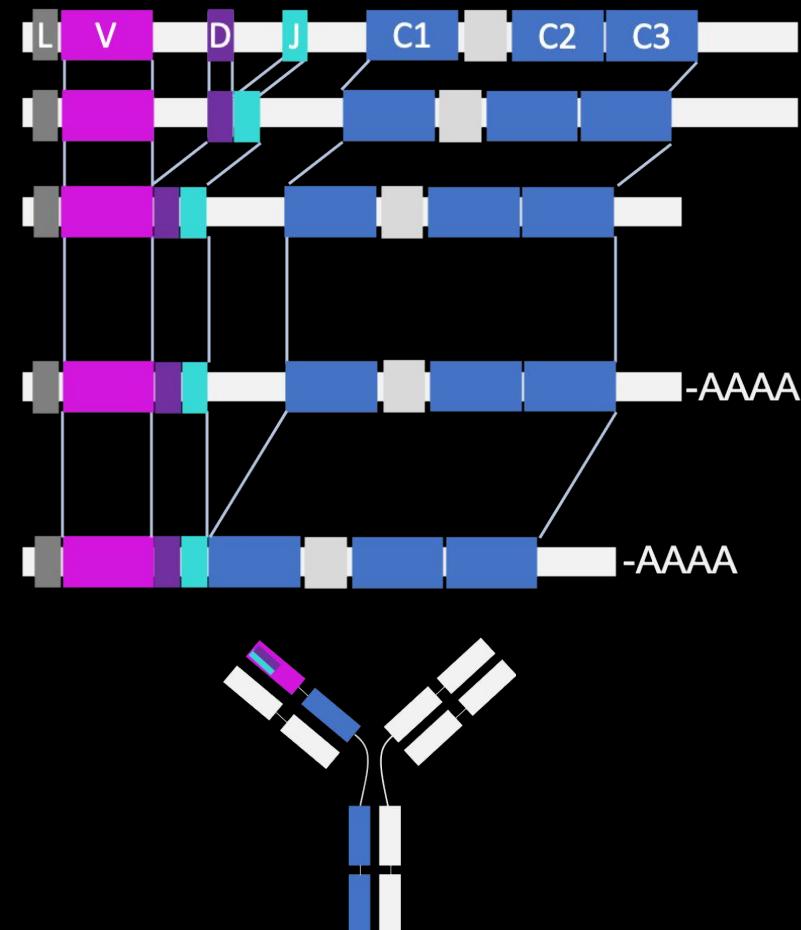


Basic concepts: Diversity

Light chain



Heavy chain



● **Basic concepts:** Immunoglobulin Structure

SCOP Classification

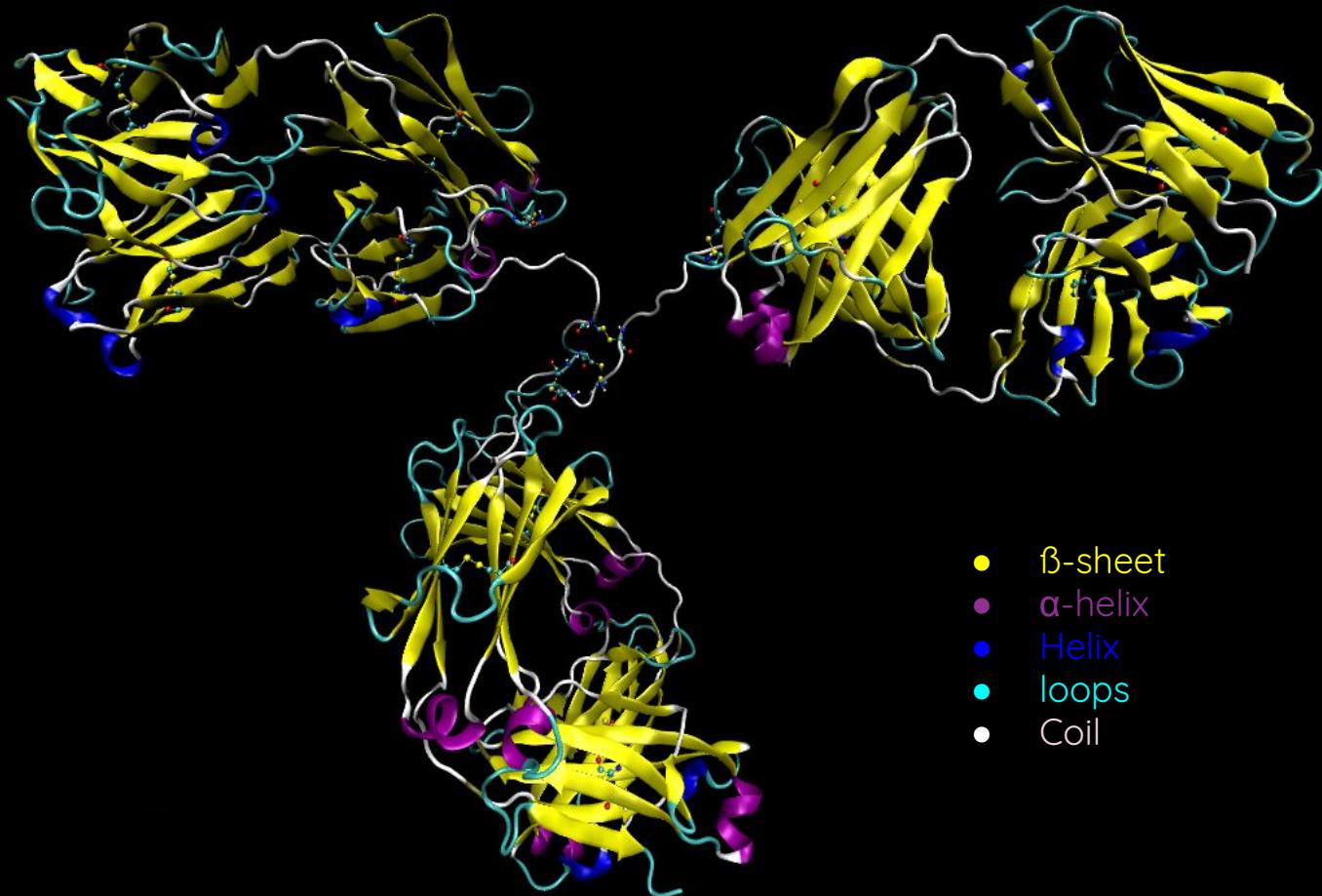
Lineage

- **Class:** All beta-proteins
- **Fold:** Immunoglobulin-like beta-sandwich
- **Superfamily:** Immunoglobulin

Families

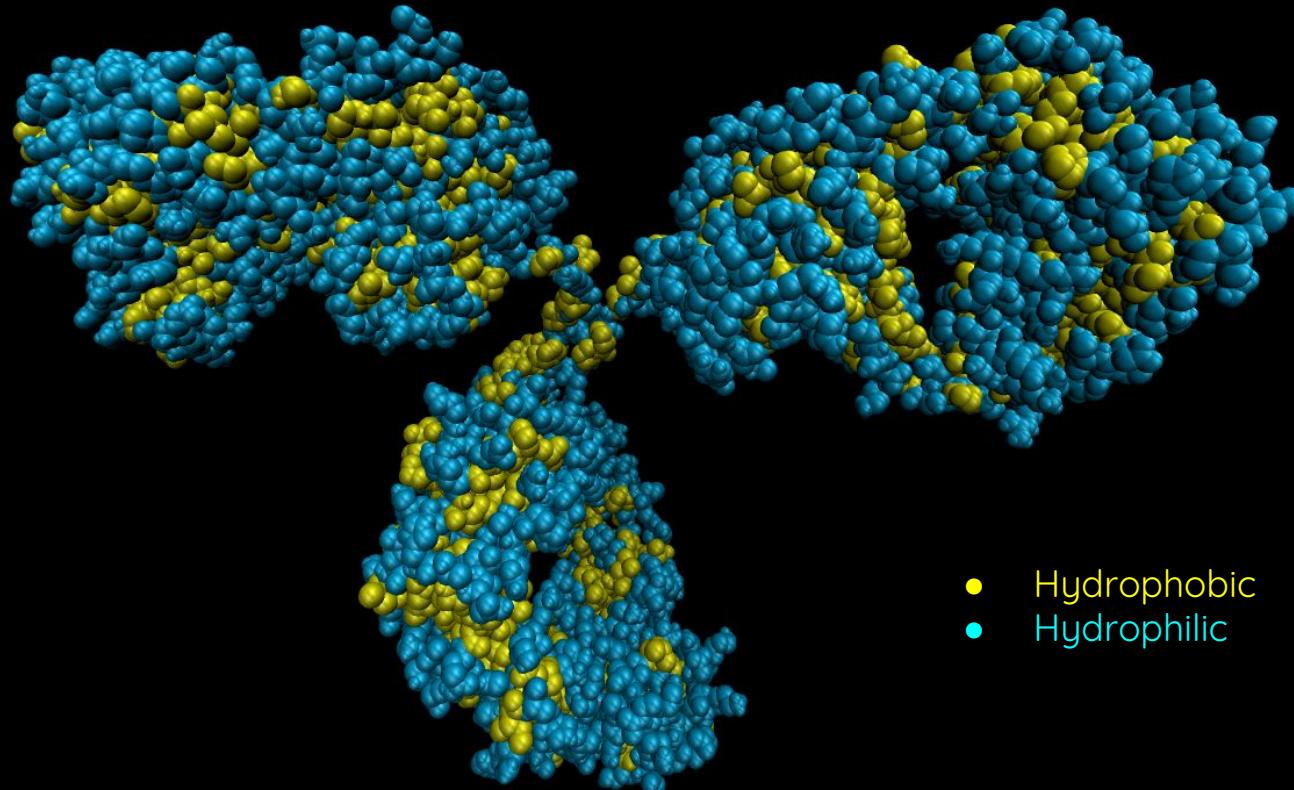
- V set domains (antibody variable domain-like)
- C1 set domains (antibody constant domain-like)
- C2 set domains
- I set domains

- **Immunoglobulin Structure: Class - All beta proteins**



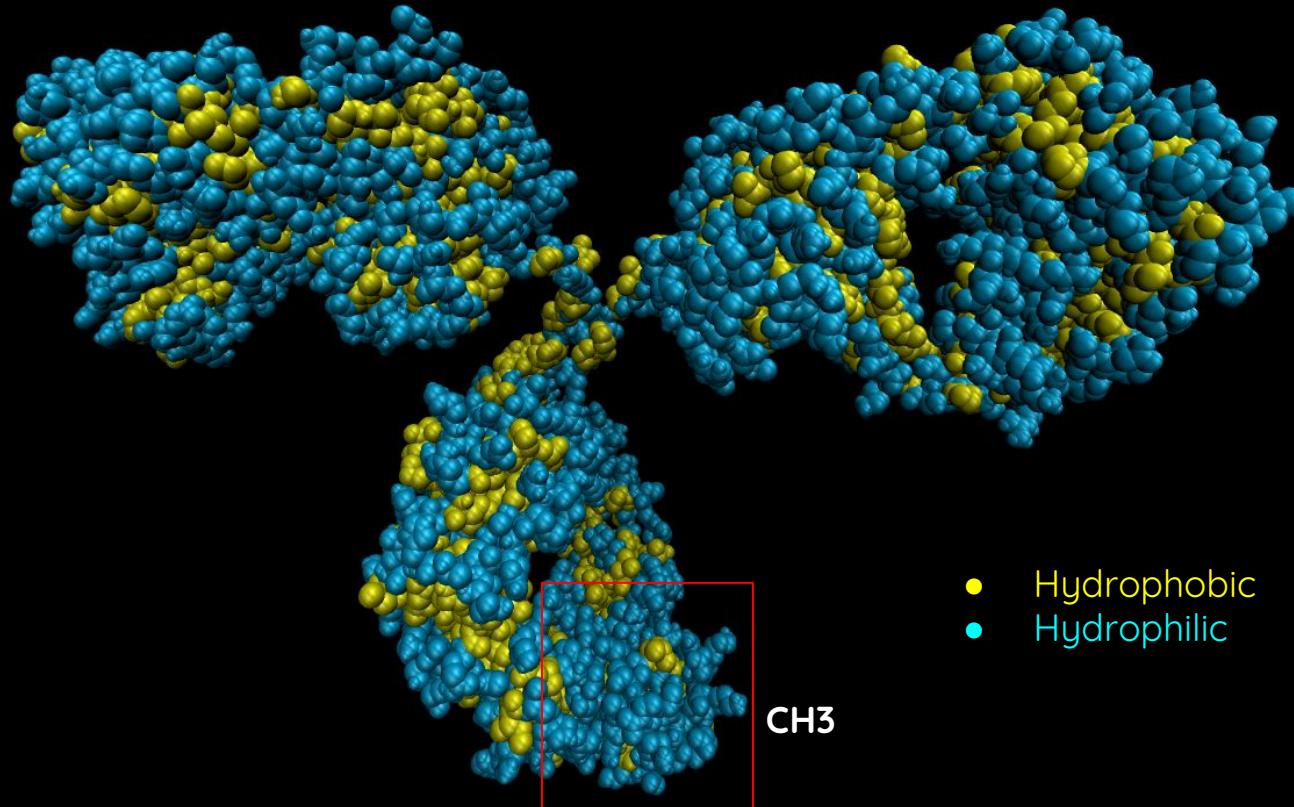
1IGT

- **Immunoglobulin Structure: Class - Hydrophobicity**



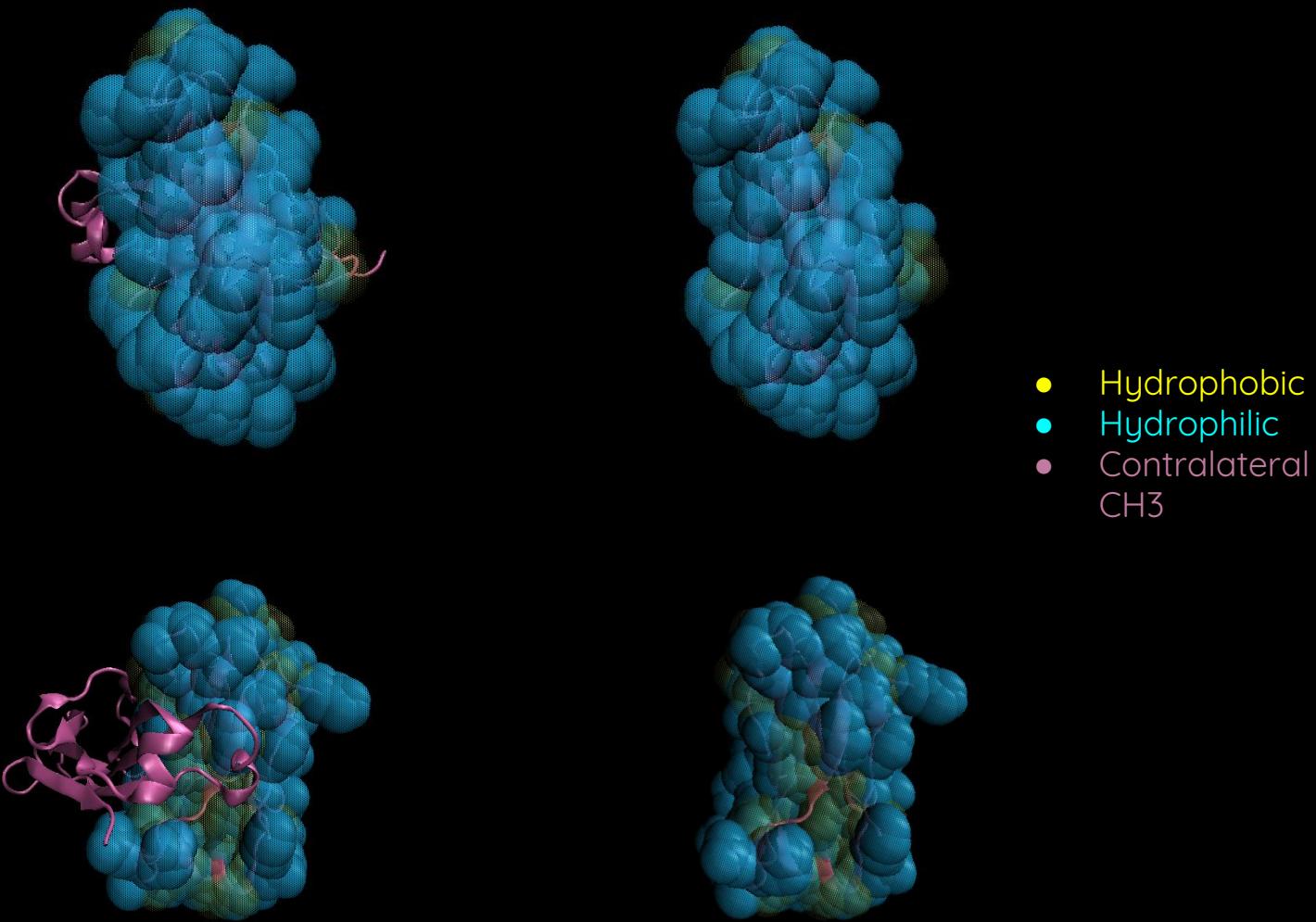
1IGT

- **Immunoglobulin Structure: Class - Hydrophobicity**



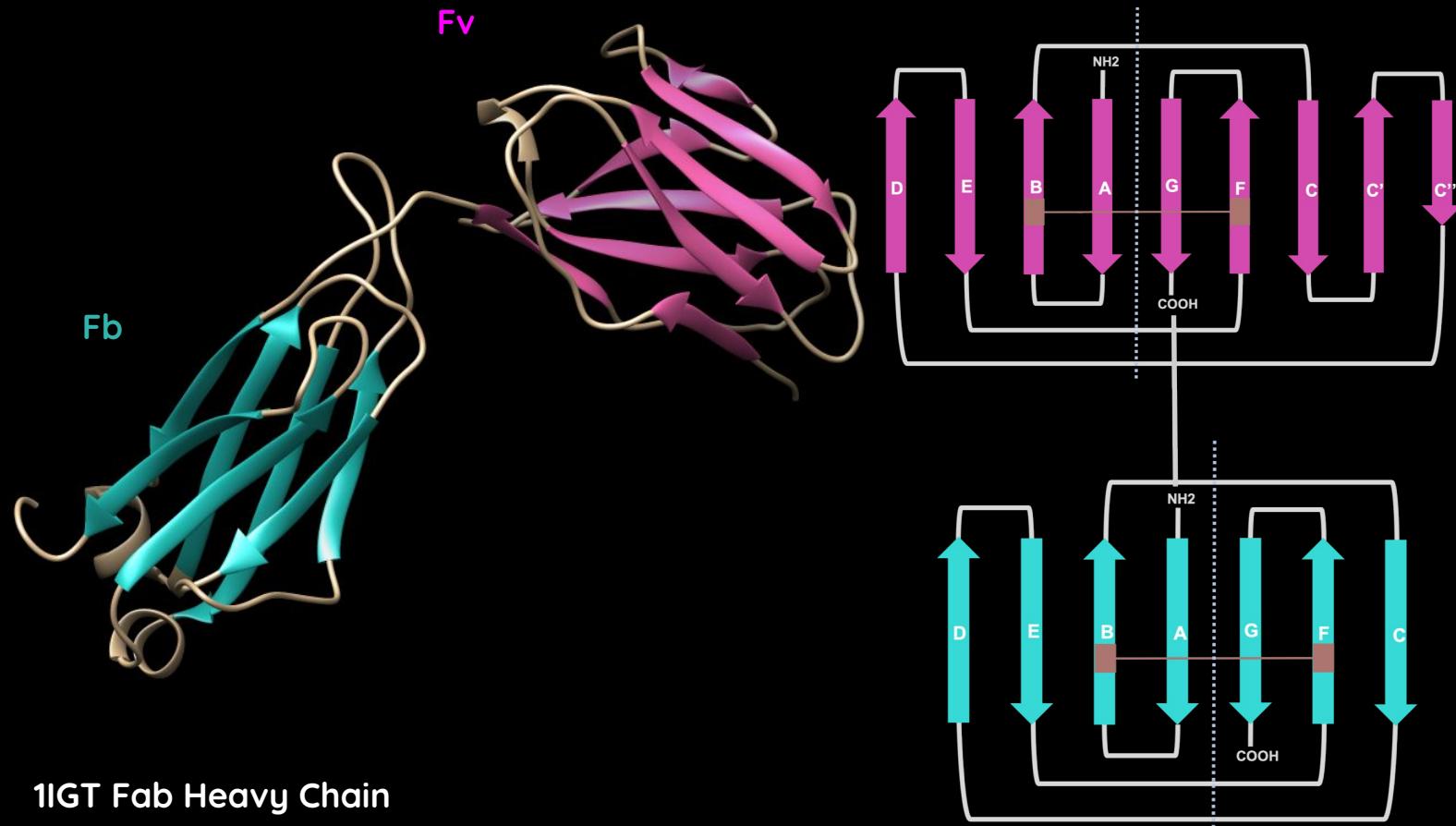
1IGT

- **Immunoglobulin Structure: Class - Hydrophobicity**



Immunoglobulin Structure: Fold -

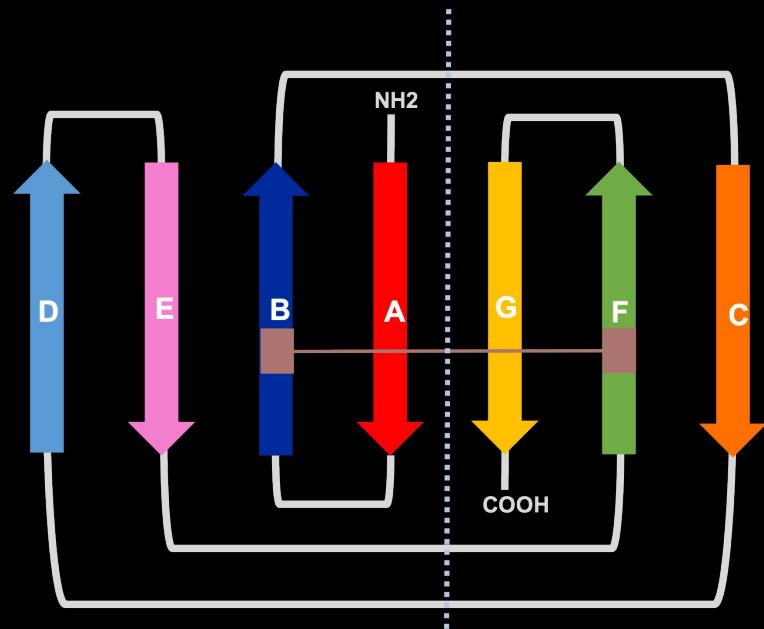
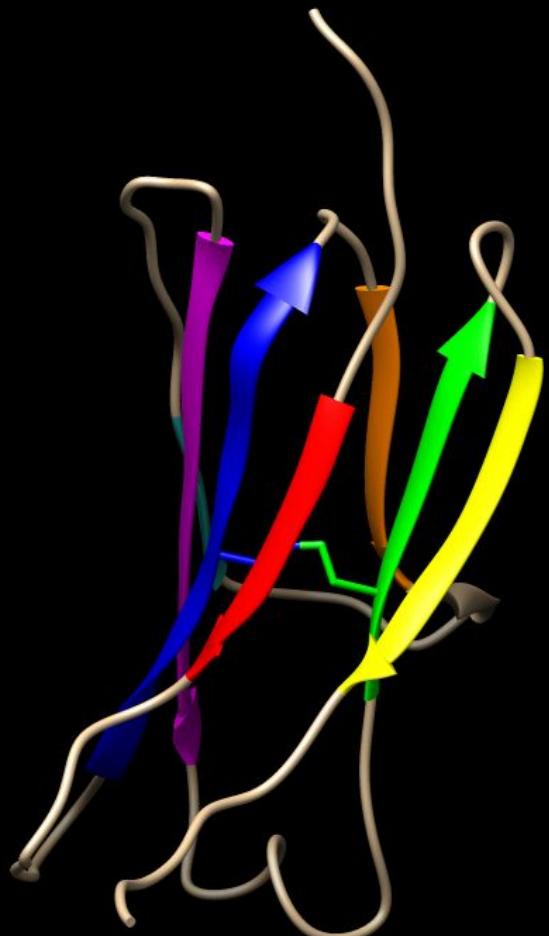
Ig like beta-sandwich



1IGT Fab Heavy Chain

Immunoglobulin Structure: Fold -

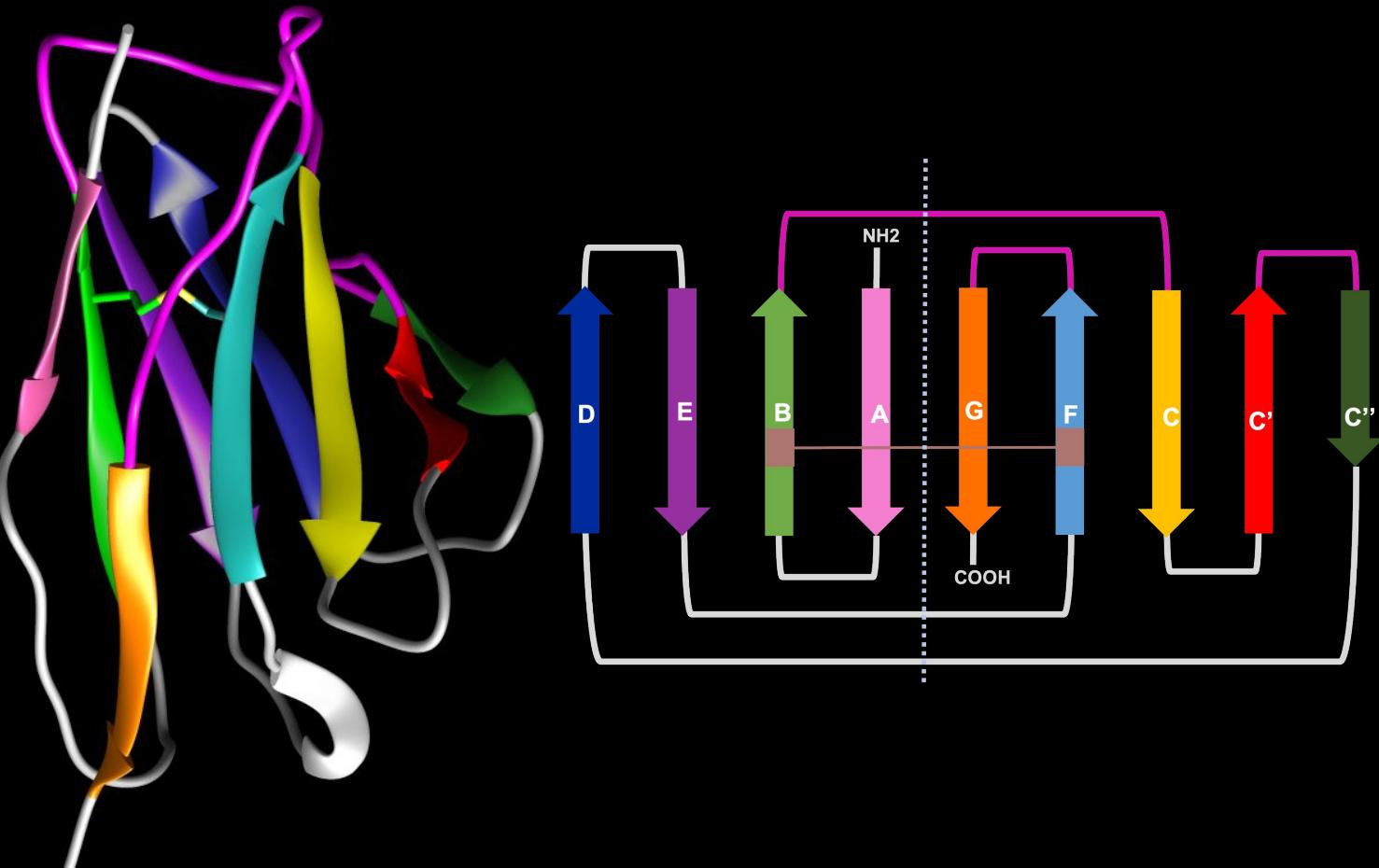
Ig like beta-sandwich



1IGT CL

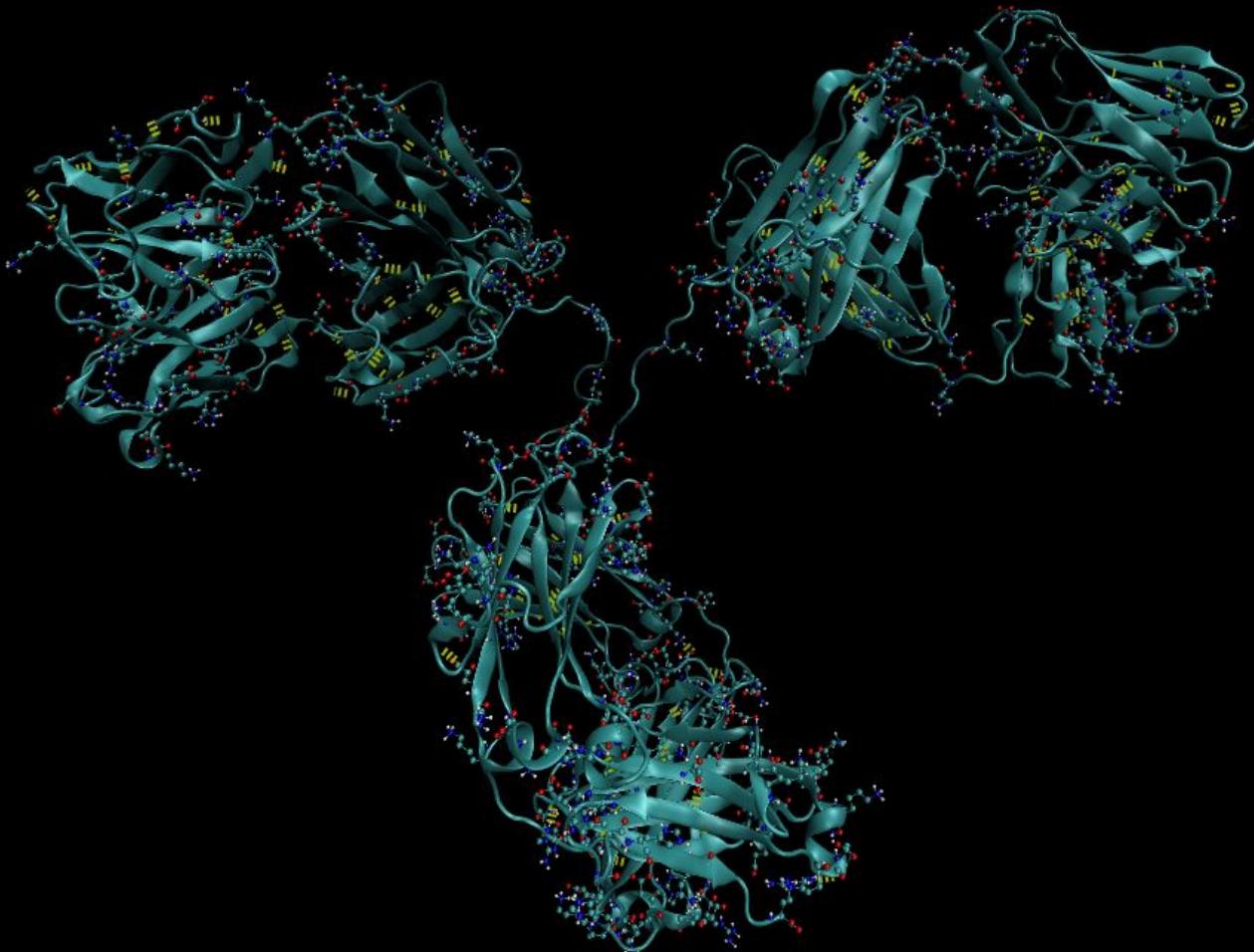
Immunoglobulin Structure: Fold -

- *Ig like beta-sandwich*



1IGT VL

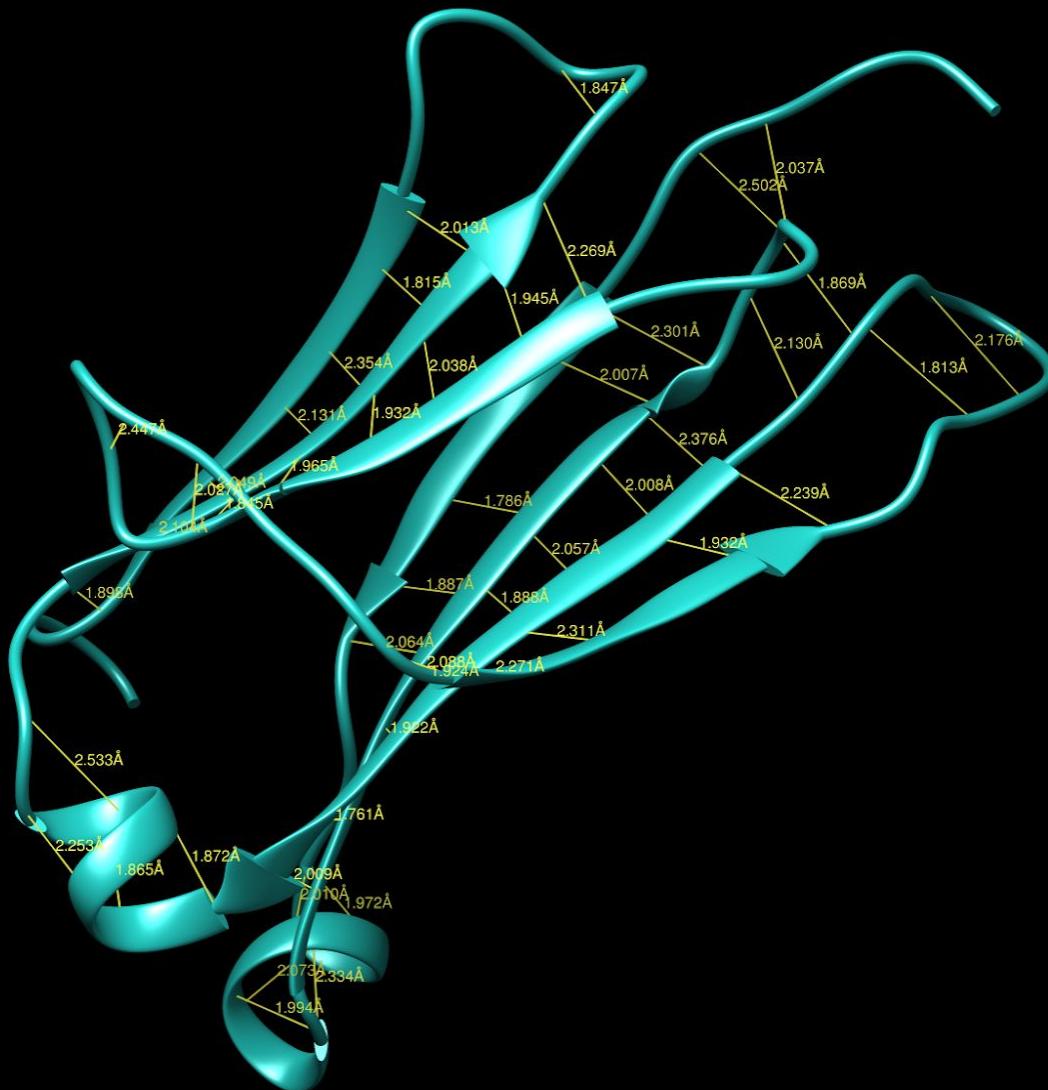
- **Immunoglobulin Structure: Hydrogen Bonds**



1IGT

Immunoglobulin Structure: Hydrogen Bonds

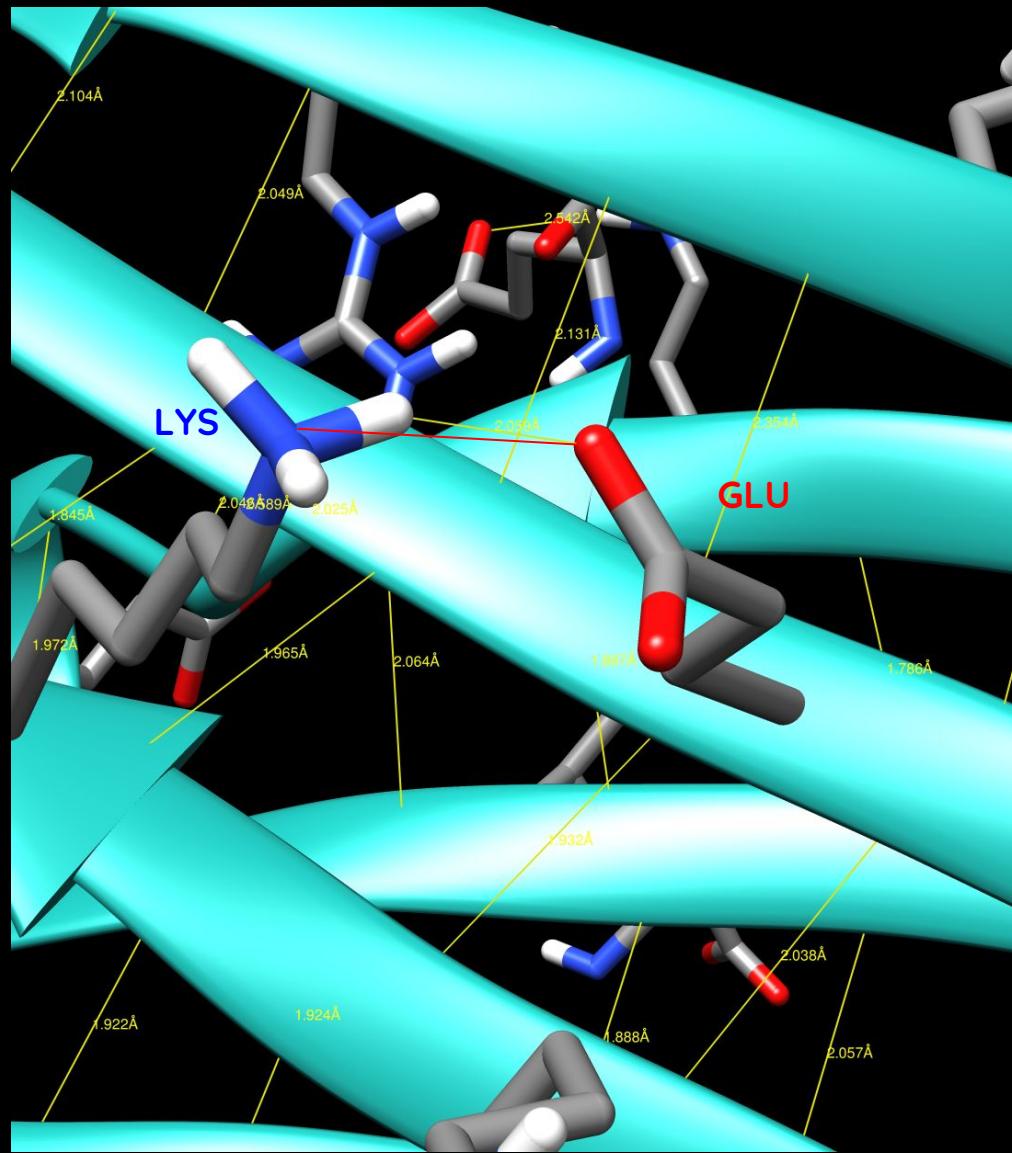
Ig like beta-sandwich



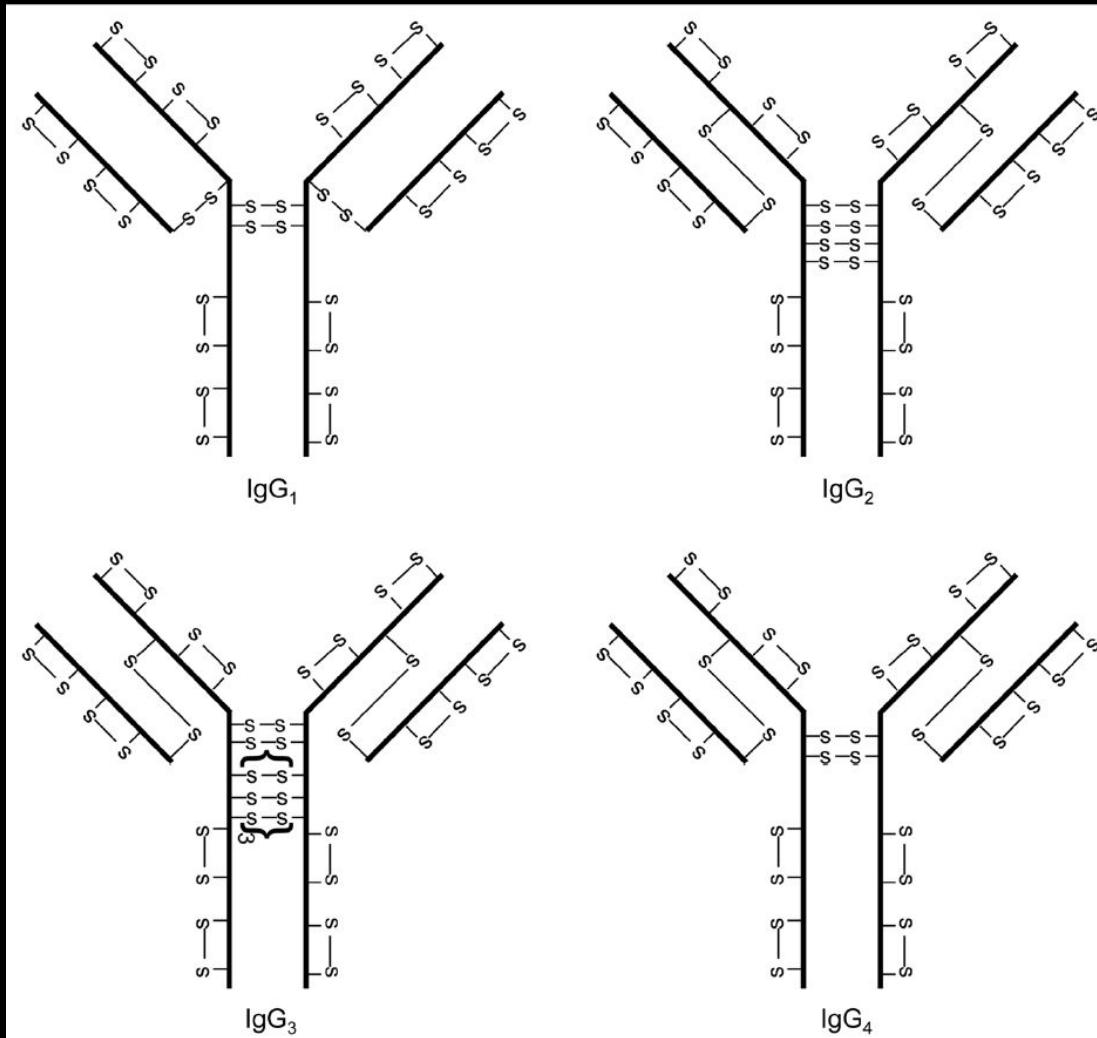
1IGT

Immunoglobulin Structure: Hydrogen Bonds

Ig like beta-sandwich

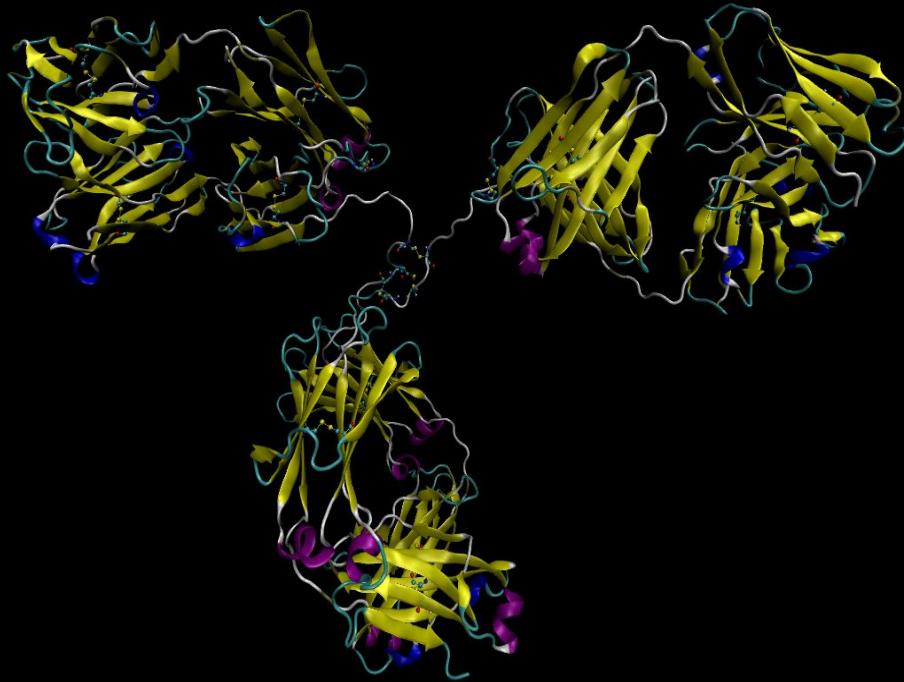


Immunoglobulin Structure: Classical IgG disulfide bond structures

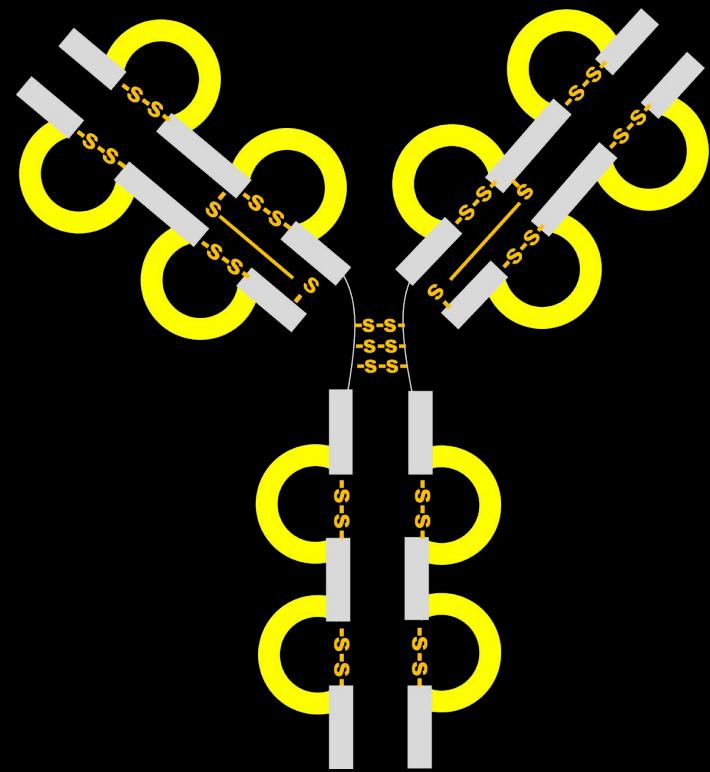




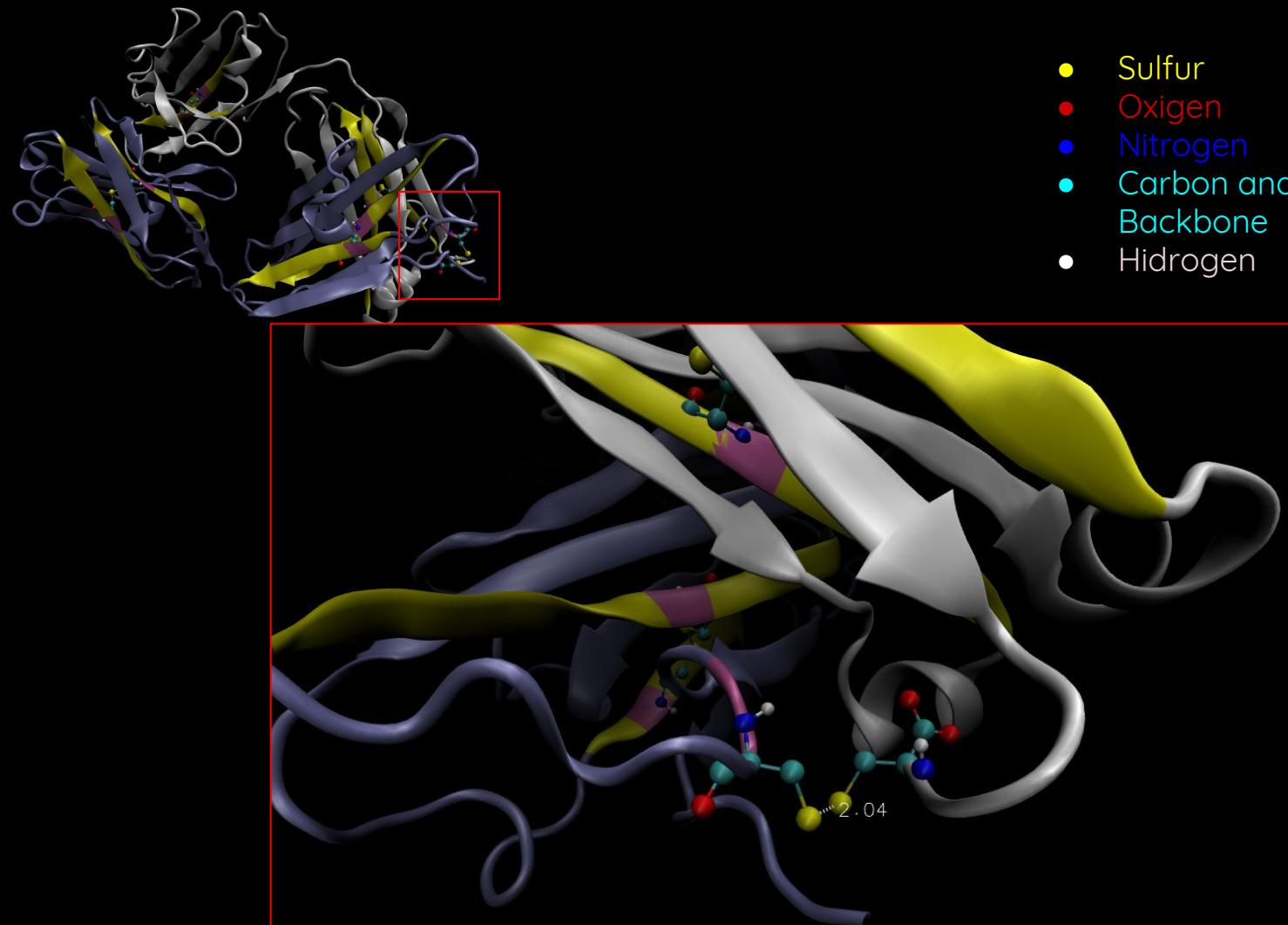
Immunoglobulin Structure: Disulfide bond structures



1IGT

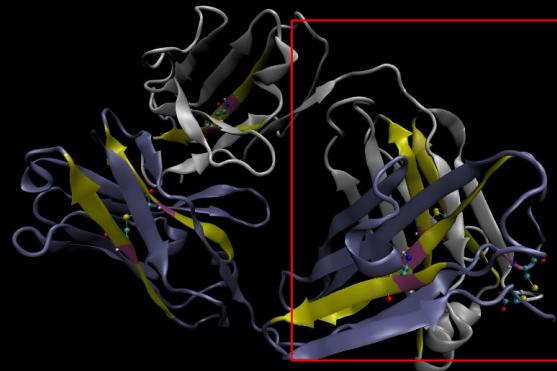


● Immunoglobulin Structure: Disulfide bond structures

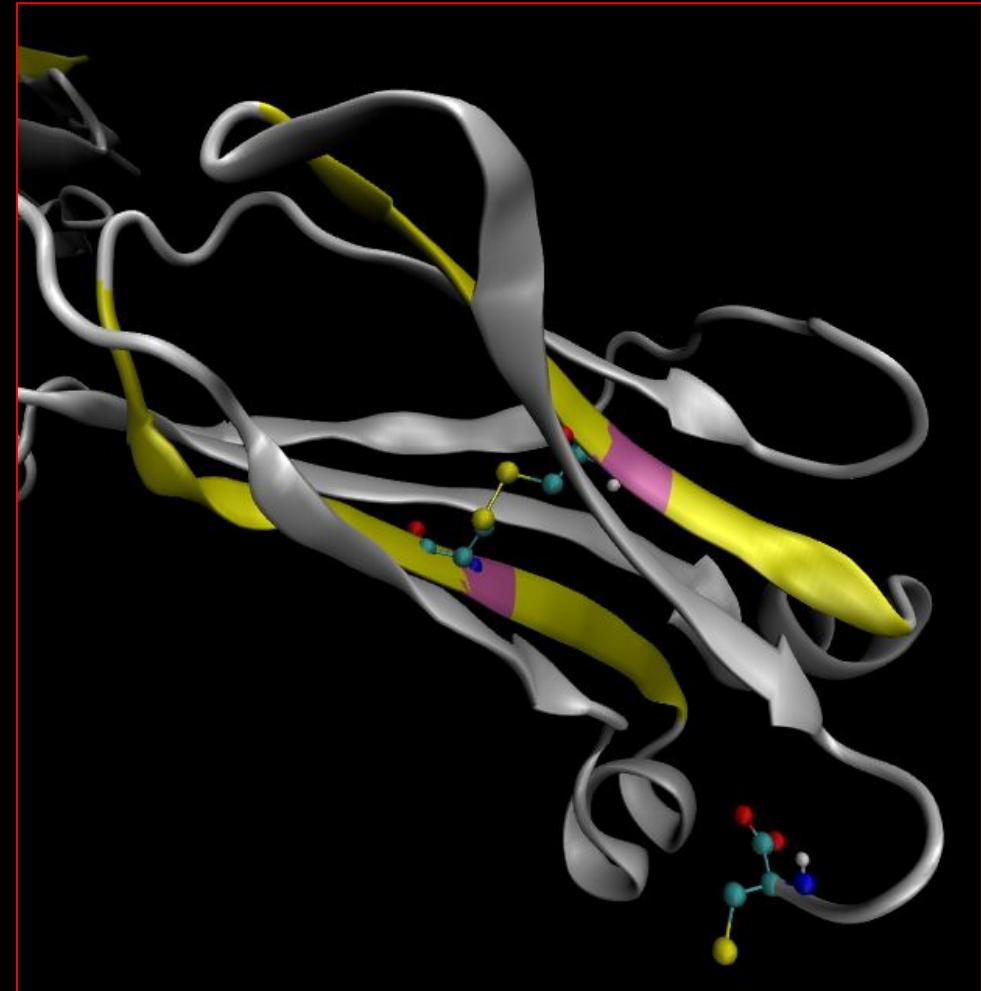


1IGT Fab

● Immunoglobulin Structure: Disulfide bond structures

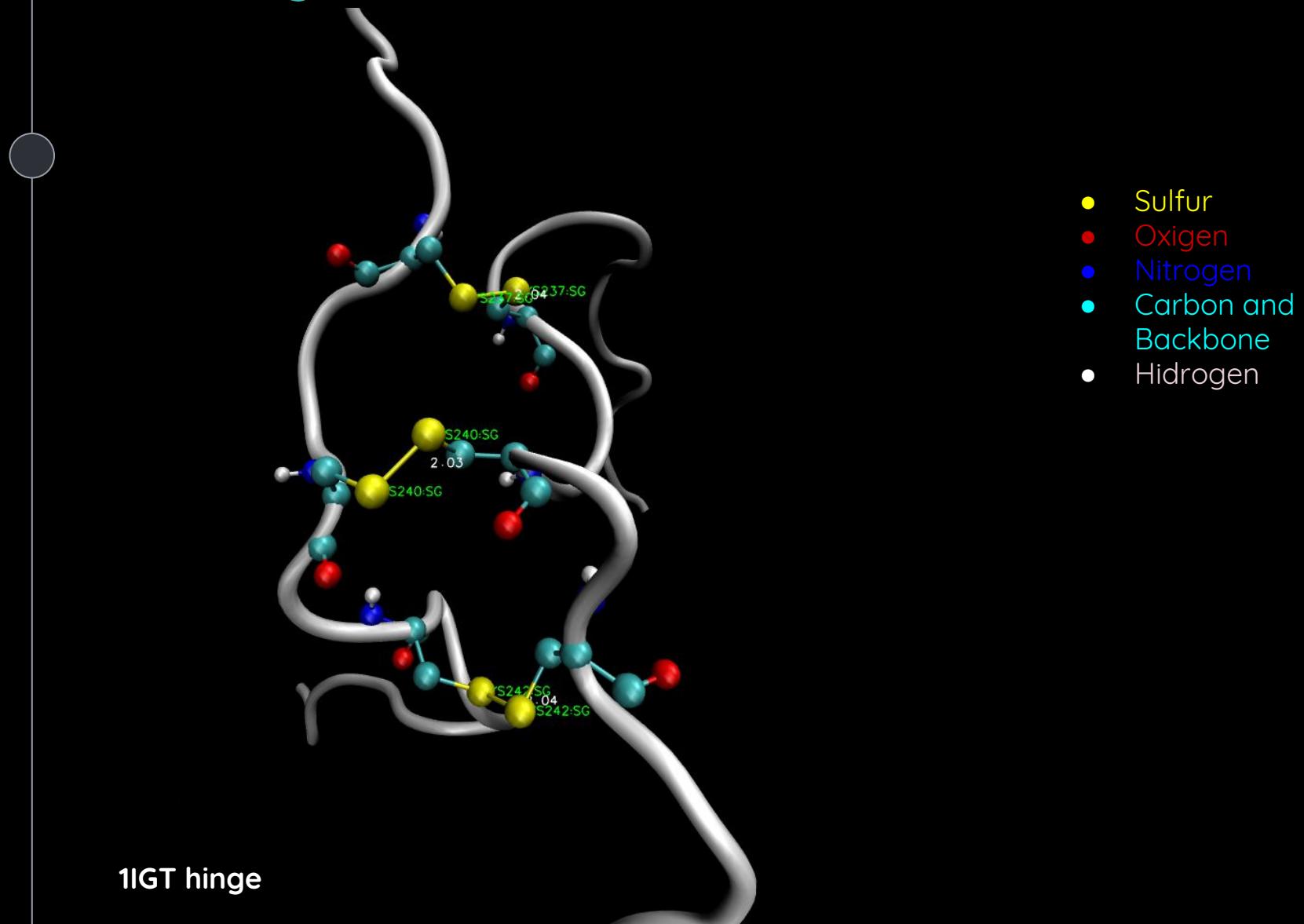


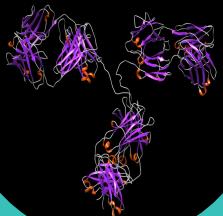
- Sulfur
- Oxigen
- Nitrogen
- Carbon and Backbone
- Hidrogen



1IGT Fab

● Immunoglobulin Structure: Disulfide bond structures

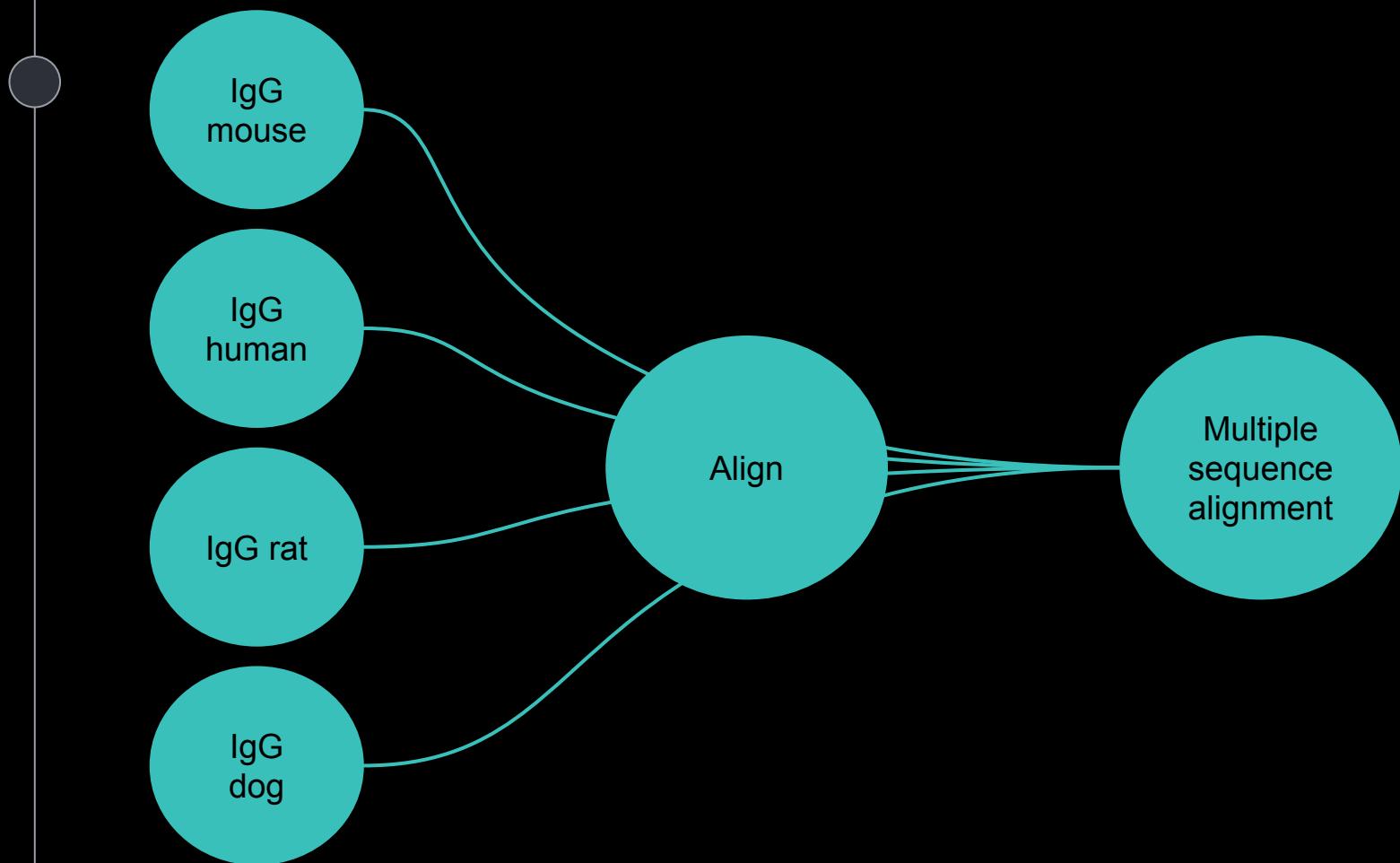




Evolution of IgG chains

Methodology

- **Objective of the evolutive study**



- **Objective of the evolutive study**

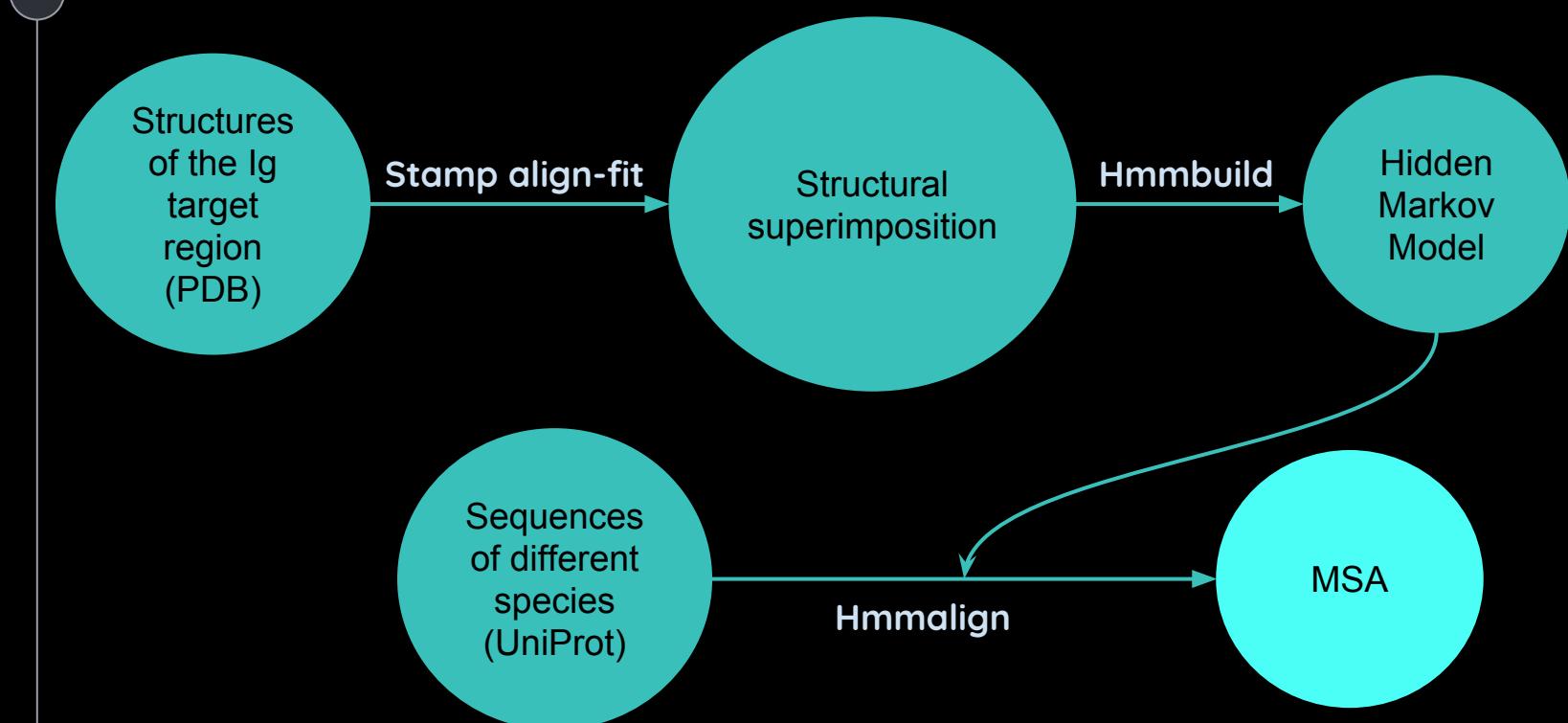


Very few different species with Ig structure in PDB

Solution: using UniProt sequences

However... UniProt is not awesome

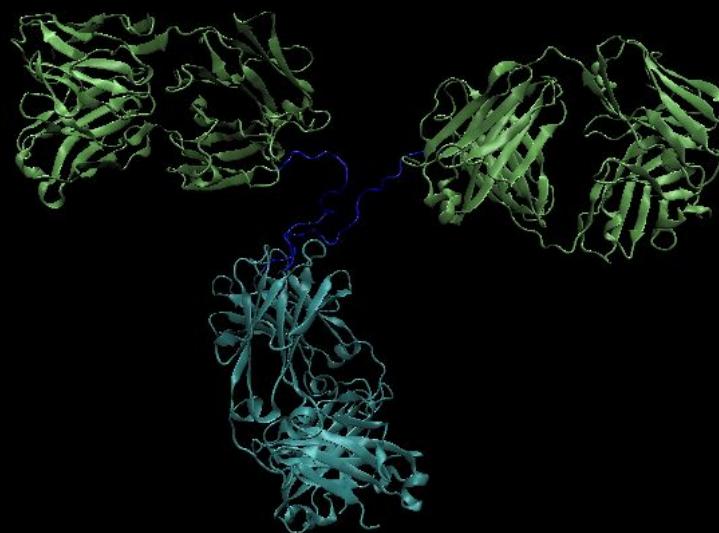
- **Objective of the evolutive study**



Methods of the evolutive study

Discovering database problems

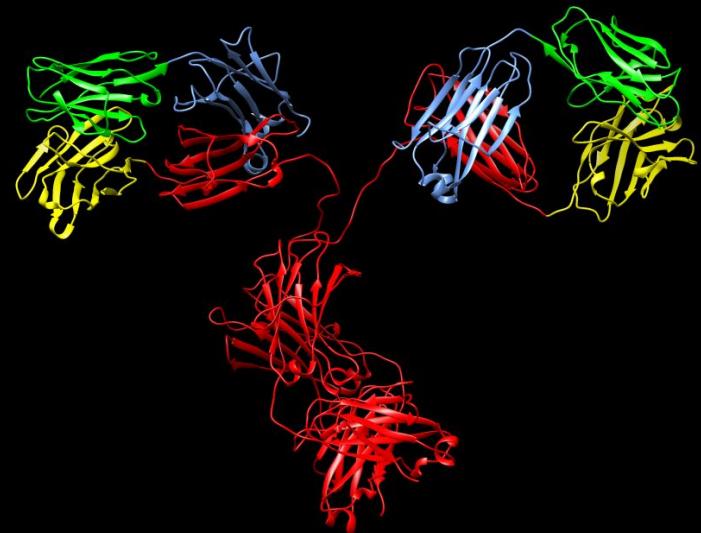
PDB



1IGT

- Fab
- Fc

UniProt

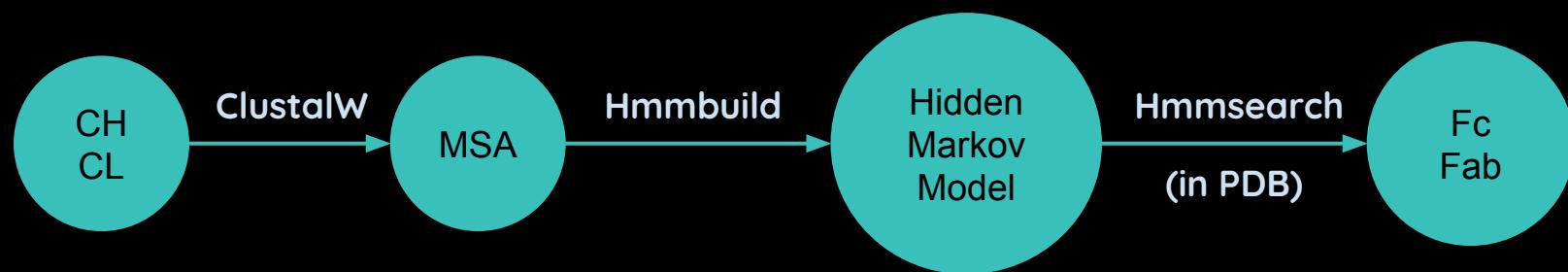


- CH
- VH
- CL
- VL

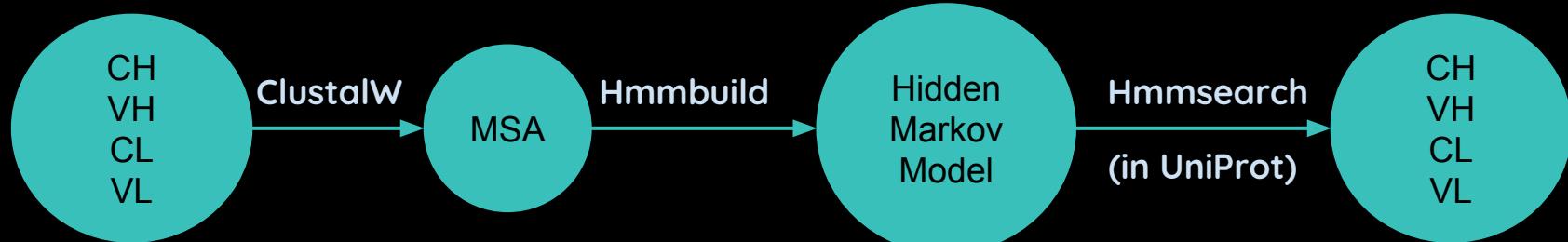
Methods of the evolutive study

Obtaining the data

Structures



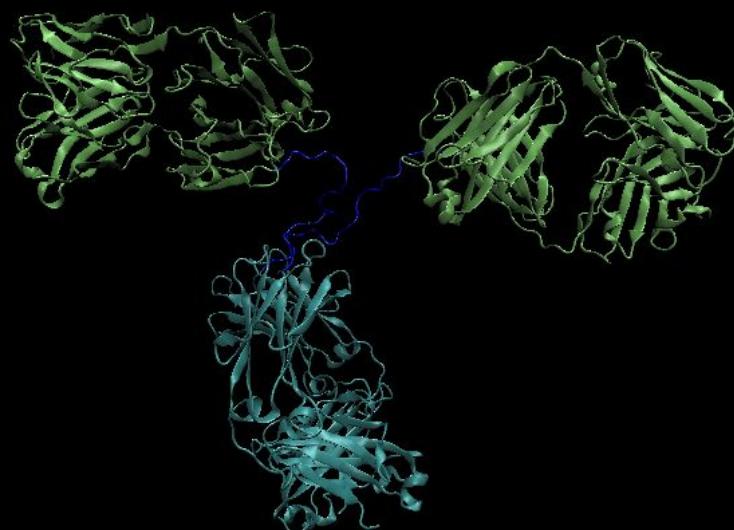
Sequences



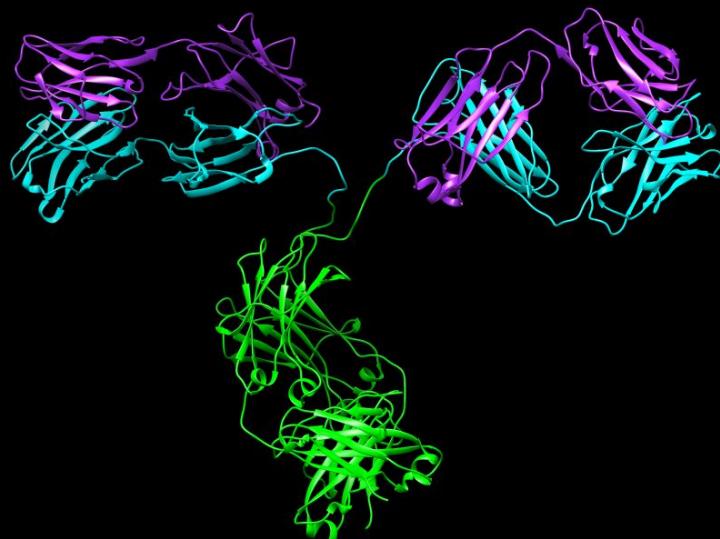
● Methods of the evolutive study

Solving database problems: obtained HMMs

PDB



HMM



- LC
- HC (Fab)
- HC (Fc)

1IGT

Methods of the evolutive study

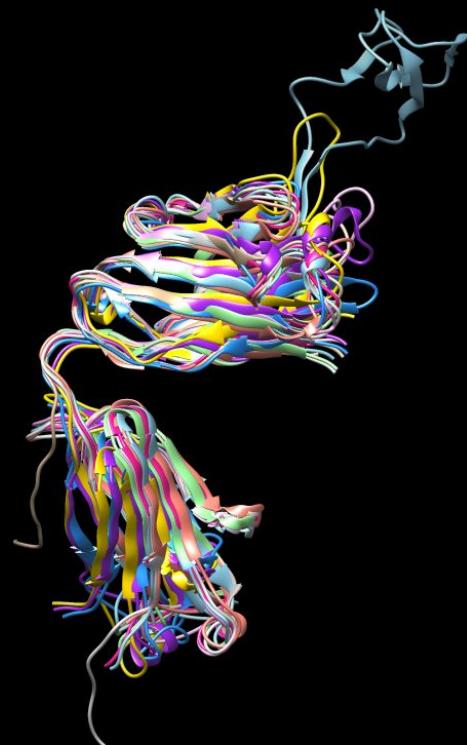
Solving database problems: obtained superimpositions

HC (Fc)



n=11

HC (Fab)



n=10

LC (Fab)



n=10

Methods of the evolutive study

Solving database problems: obtained MSA

Fab
(HC)

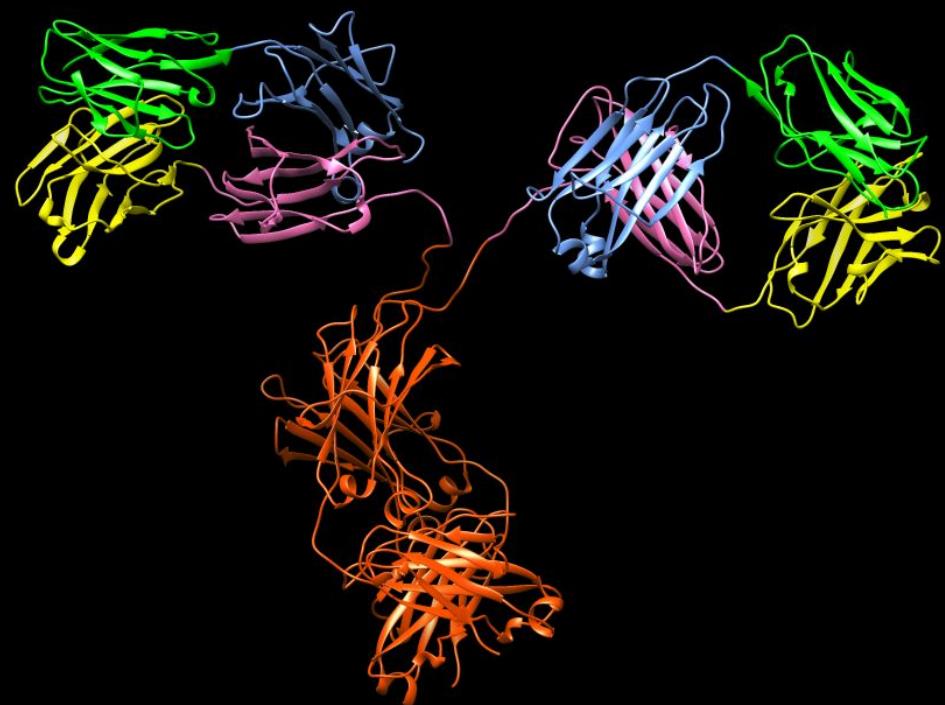
- VH
- CH1

Fab
(LC)

- VL
- CL

Fc

- CH2-3



1IGT

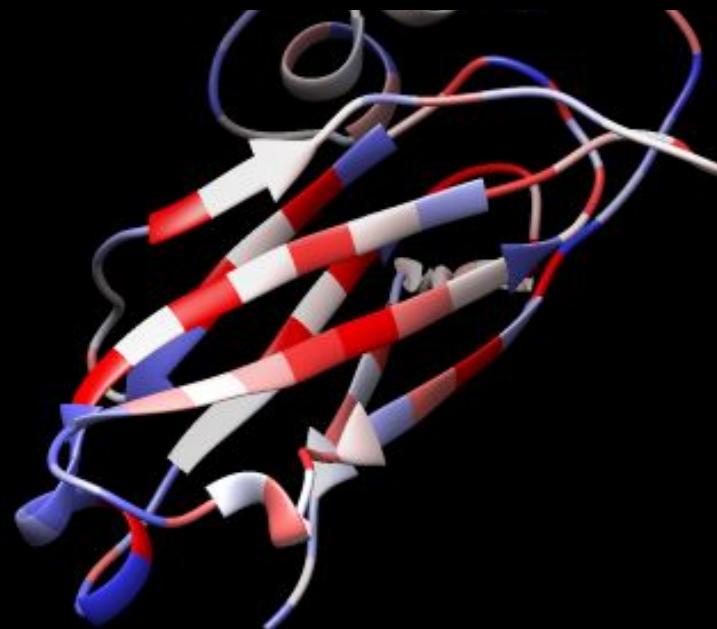
Visualization of the results



Clustal X Default Colouring			
Category	Colour	Residue at position	{ Threshold, Residue group }
Hydrophobic	BLUE	A,I,L,M,F,W,V	{>60%, WLVIMAFCHP}
		C	{>60%, WLVIMAFCHP}
Positive charge	RED	K,R	{>60%,KR}, {>80%, K,R,Q}
Negative charge	MAGENTA	E	{>60%,KR}, {>50%,QE}, {>85%,E,Q,D}
		D	{>60%,KR}, {>85%, K,R,Q}, {>50%,ED}
Polar	GREEN	N	{>50%, N}, {>85%, N,Y}
		Q	{>60%,KR}, {>50%,QE}, {>85%,Q,E,K,R}
		S,T	{>60%, WLVIMAFCHP}, {>50%, TS}, {>85%,S,T}
Cysteines	PINK	C	{>85%, C}
Glycines	ORANGE	G	{>0%, G}
Prolines	YELLOW	P	{>0%, P}
Aromatic	CYAN	H,Y	{>60%, WLVIMAFCHP}, {>85%, W,Y,A,C,P,Q,F,H,I,L,M,V}
Unconserved	WHITE	any / gap	If none of the above criteria are met

ClustalX coloring method

- **Visualization of the results**



Highly variable

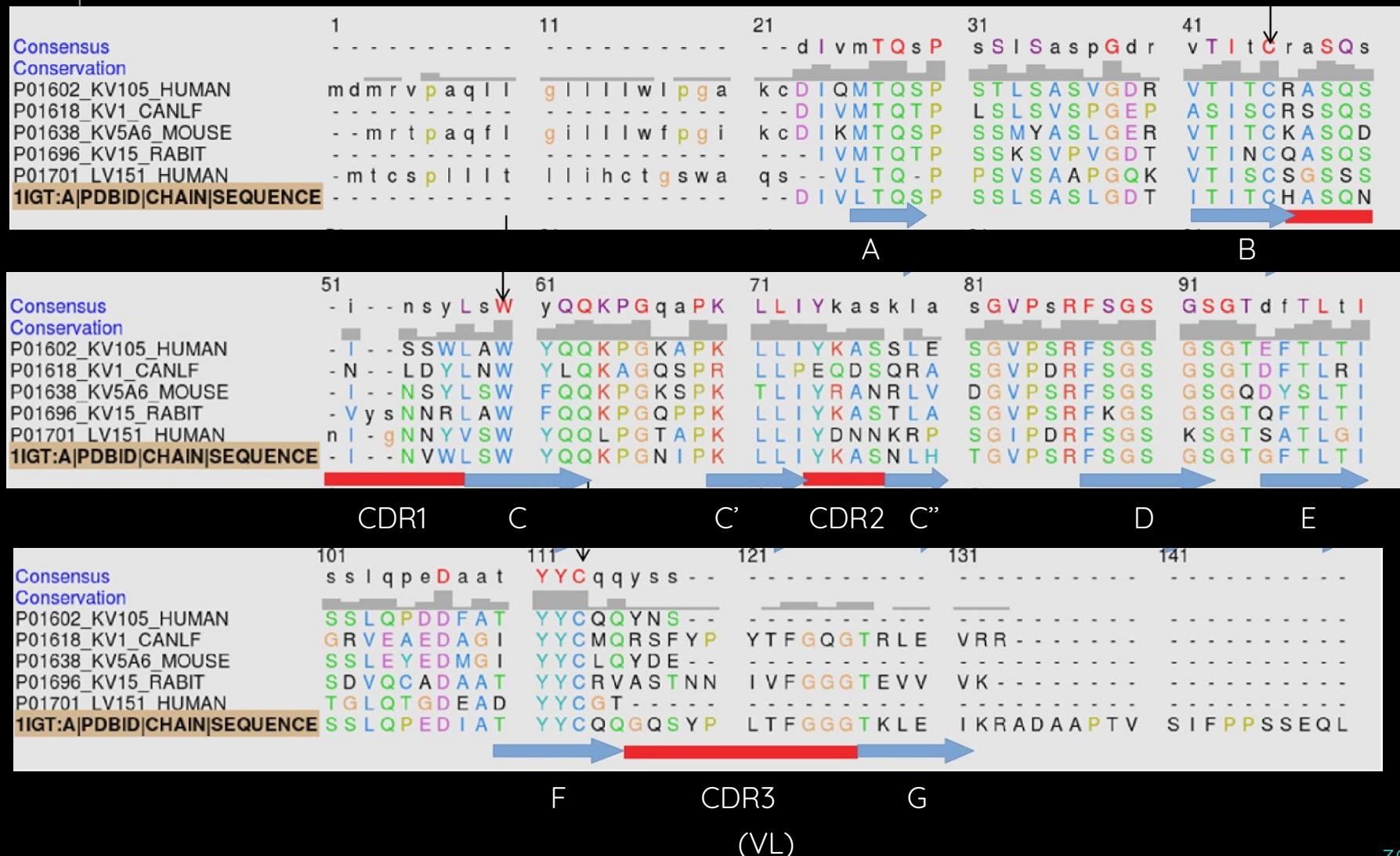
Highly conserved



Evolution of IgG chains

Results

CDR are highly variable regions



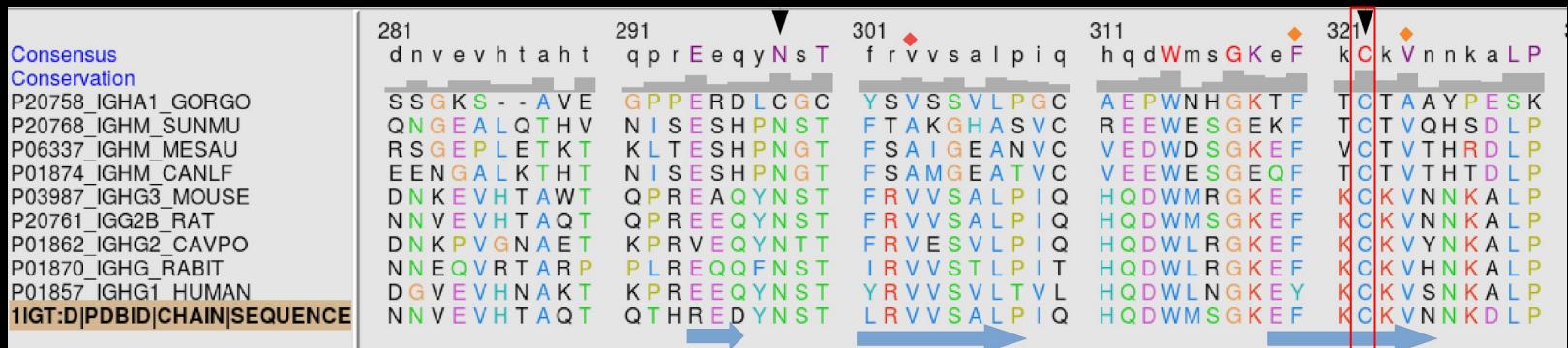
Disulphide bridge-forming Cys are totally conserved in constant and variable Ig domains



A

B

C

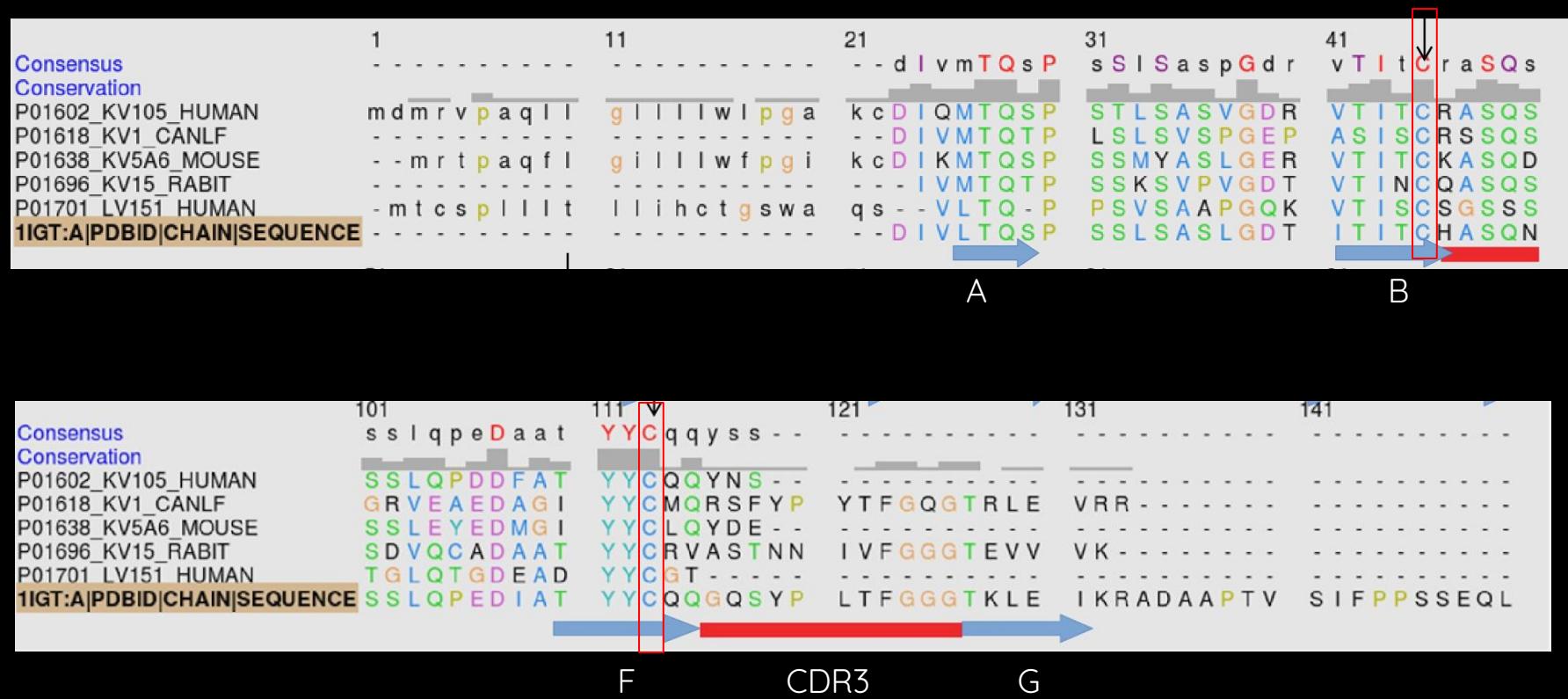


D

E

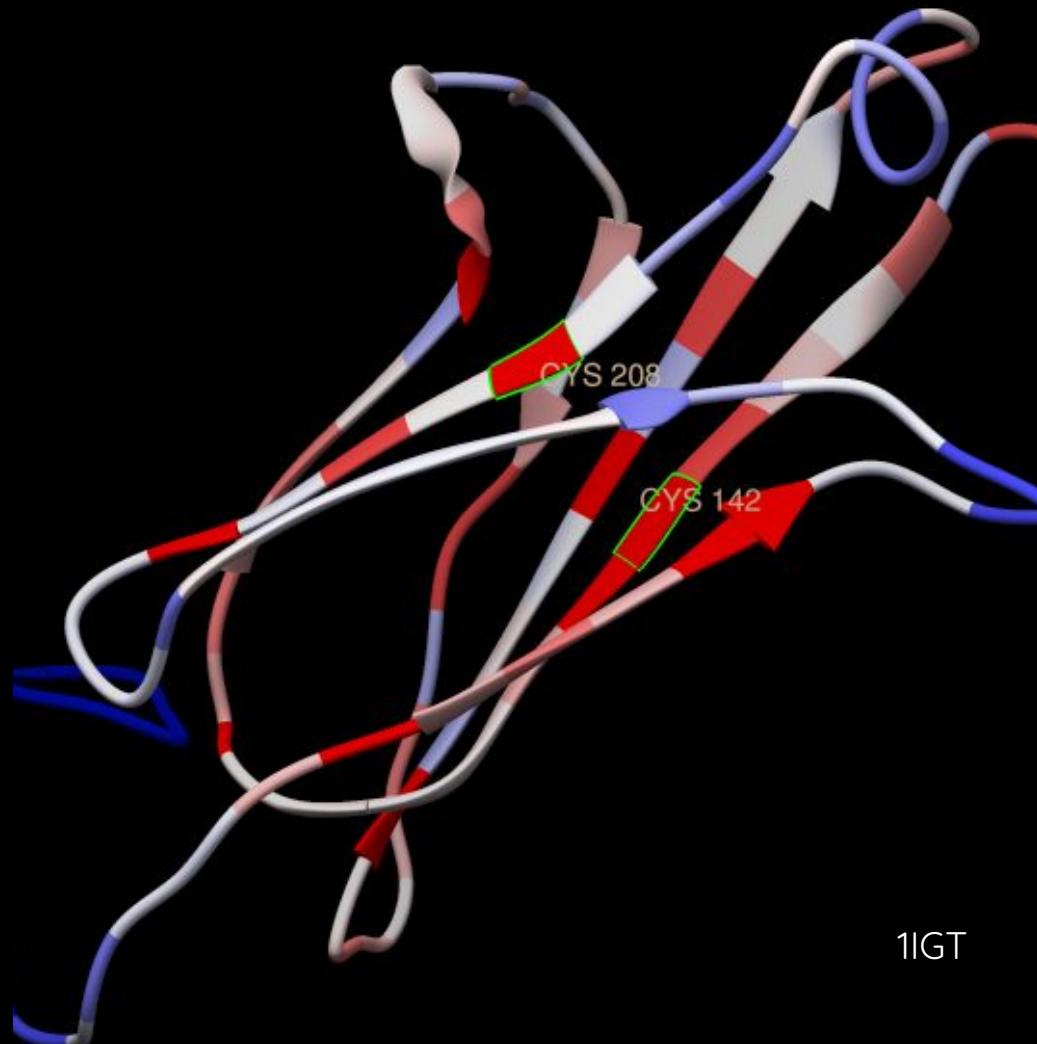
F

Disulphide bridge-forming Cys are totally conserved in constant and variable Ig domains



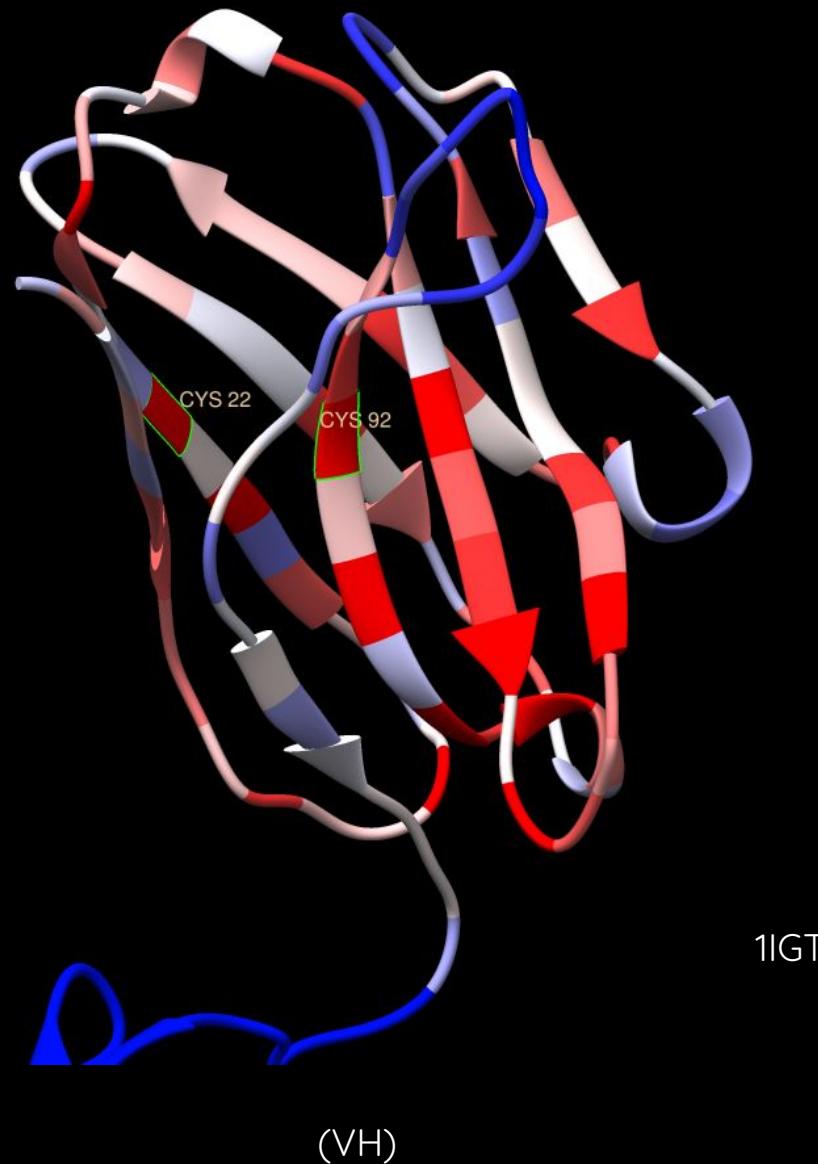
(VL)

- Cys residues are greatly conserved in constant and variable Ig domains

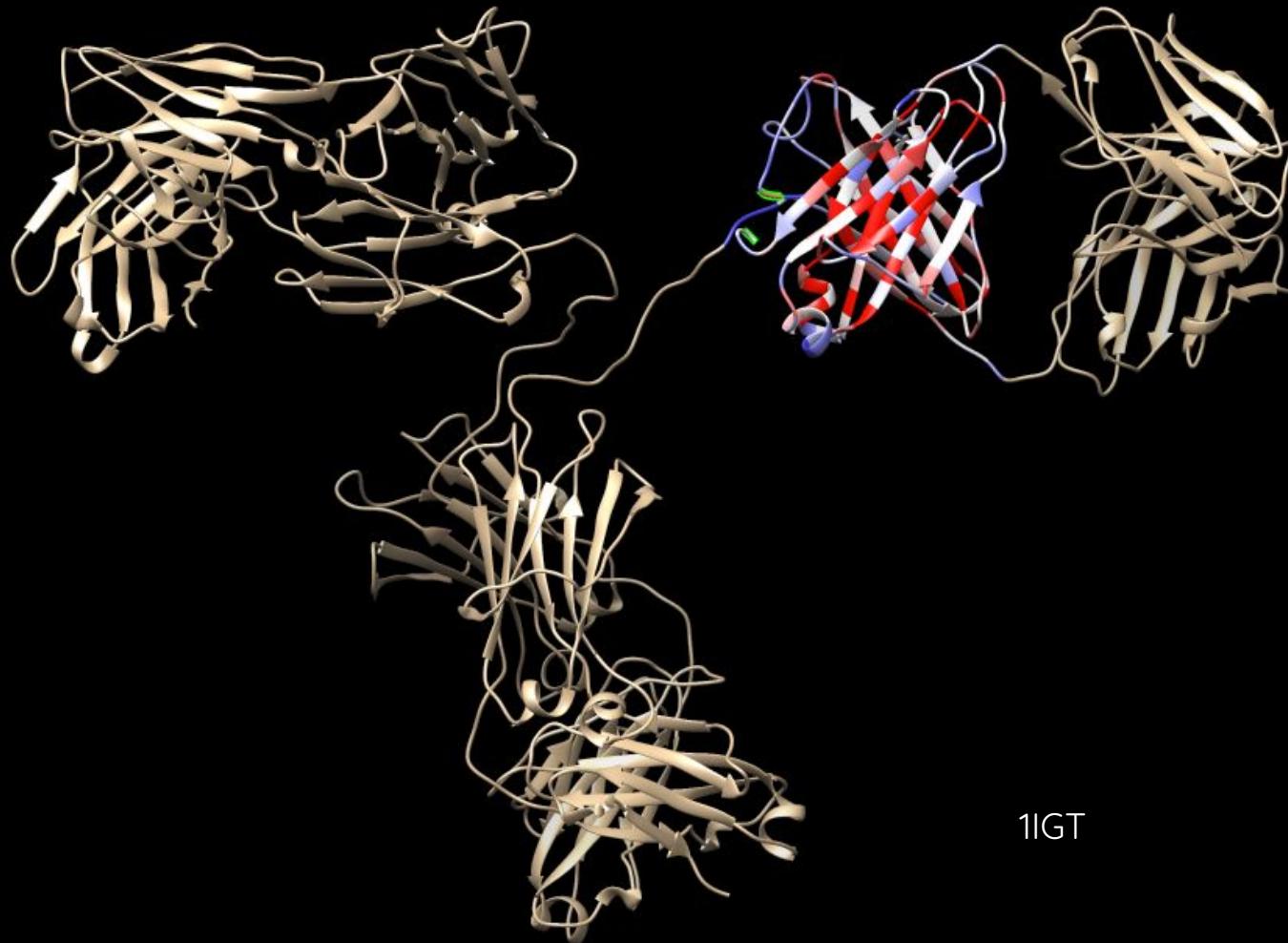


1IGT

Cys residues are greatly conserved in constant and variable Ig domains

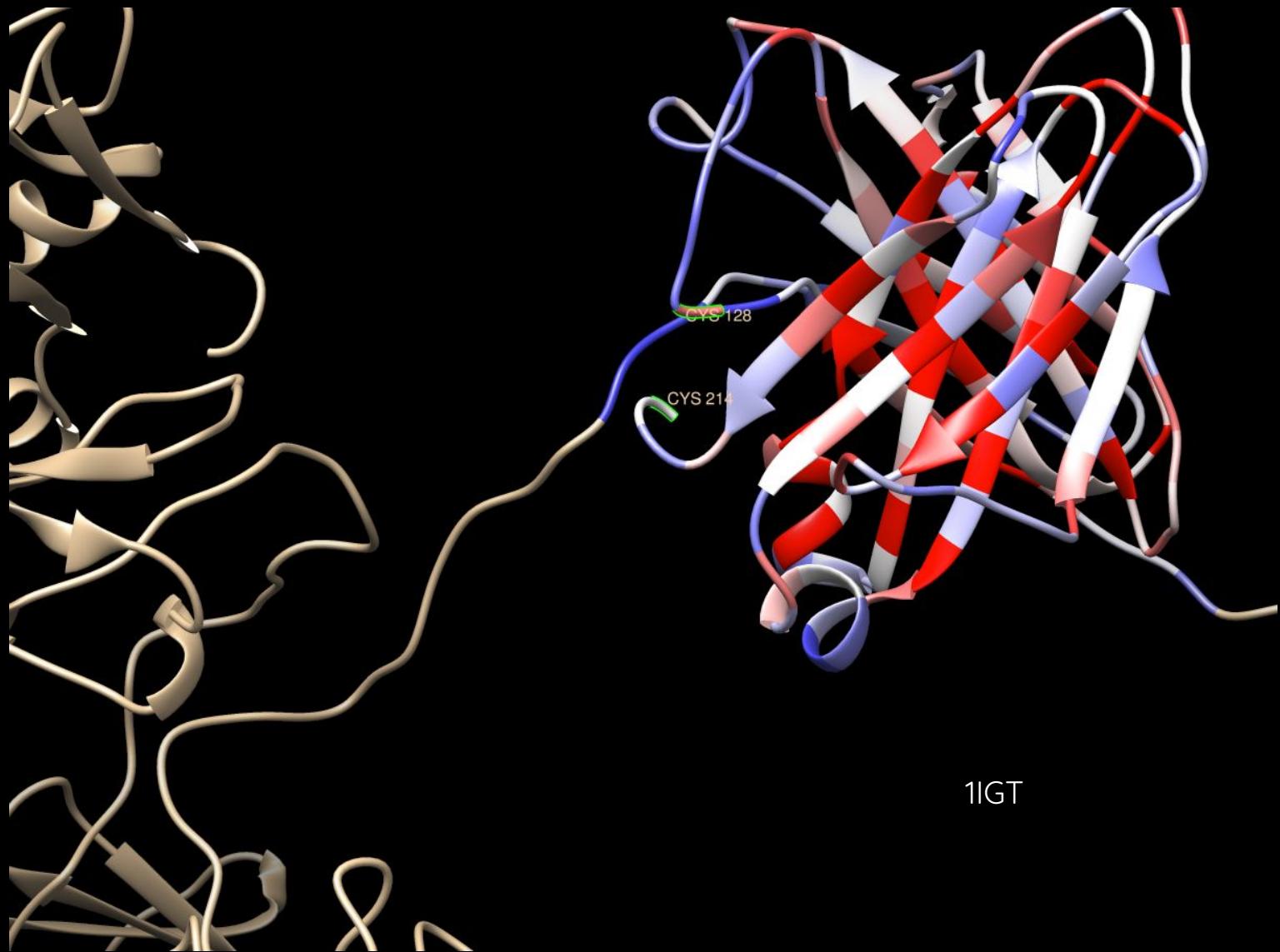


- Cys joining HC and LC are greatly conserved

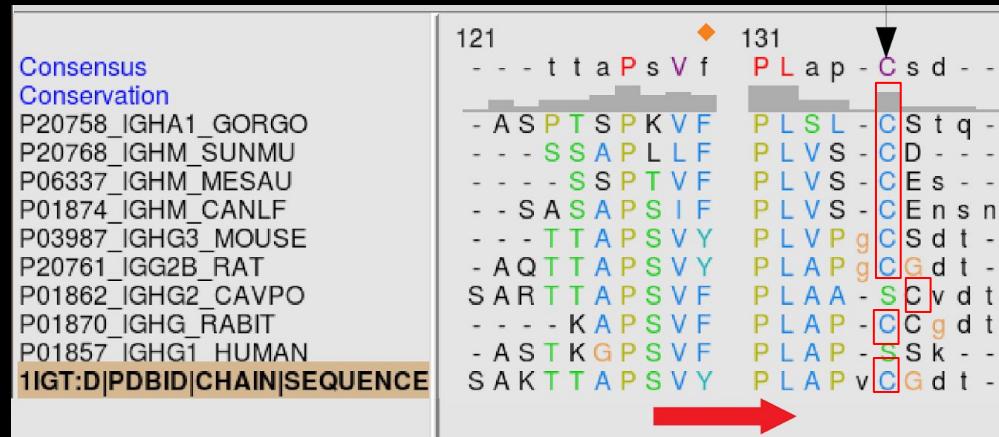


1IGT

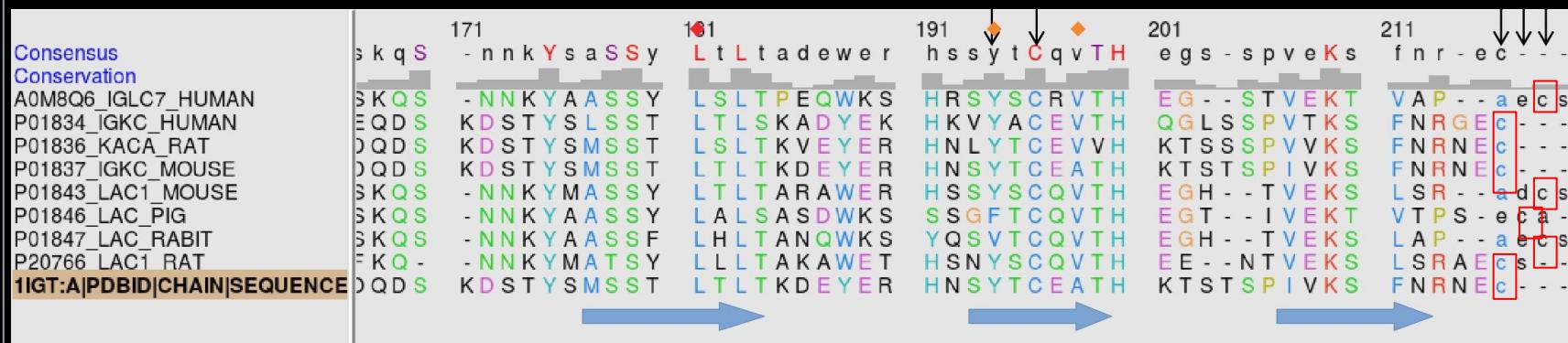
- Cys joining HC and LC are greatly conserved



Cys joining HC and LC are greatly conserved

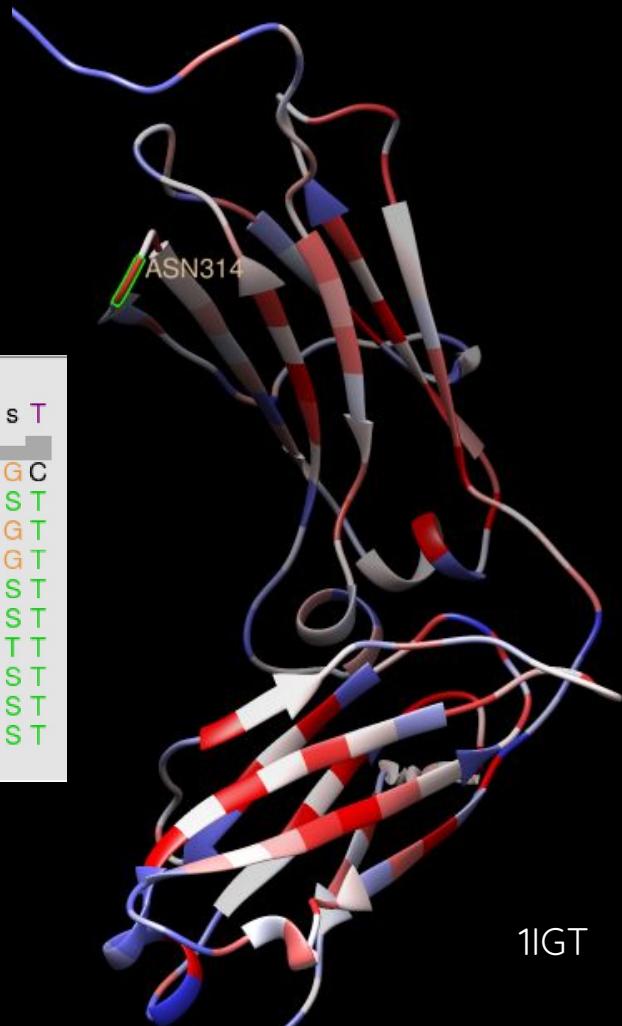
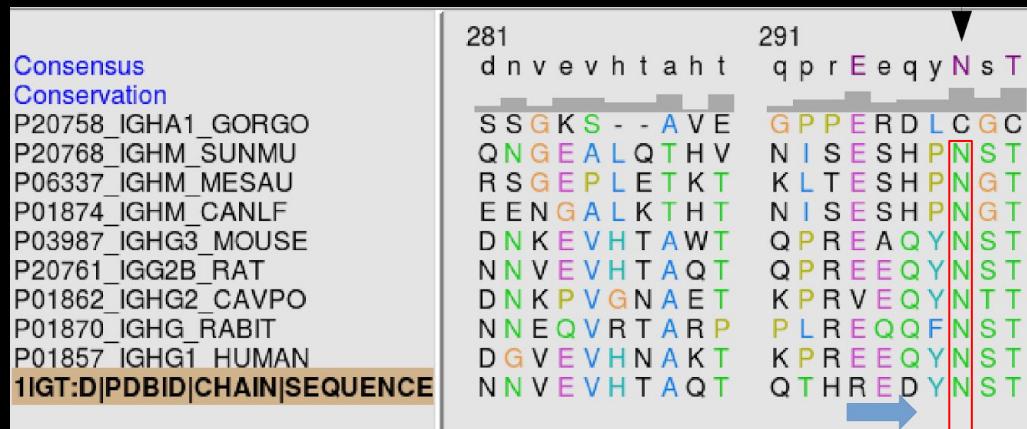


A
(CH1)

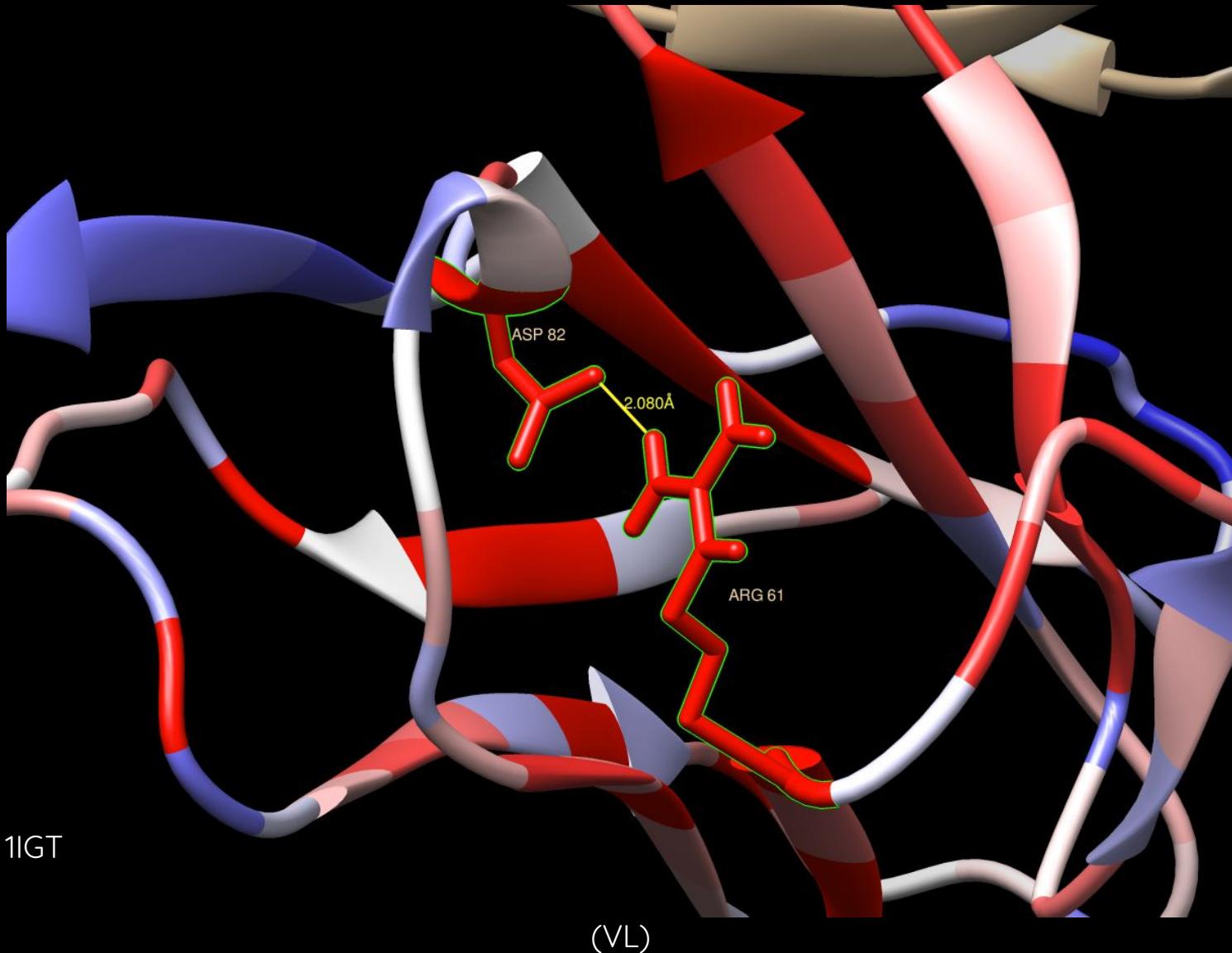


E (CL) F G

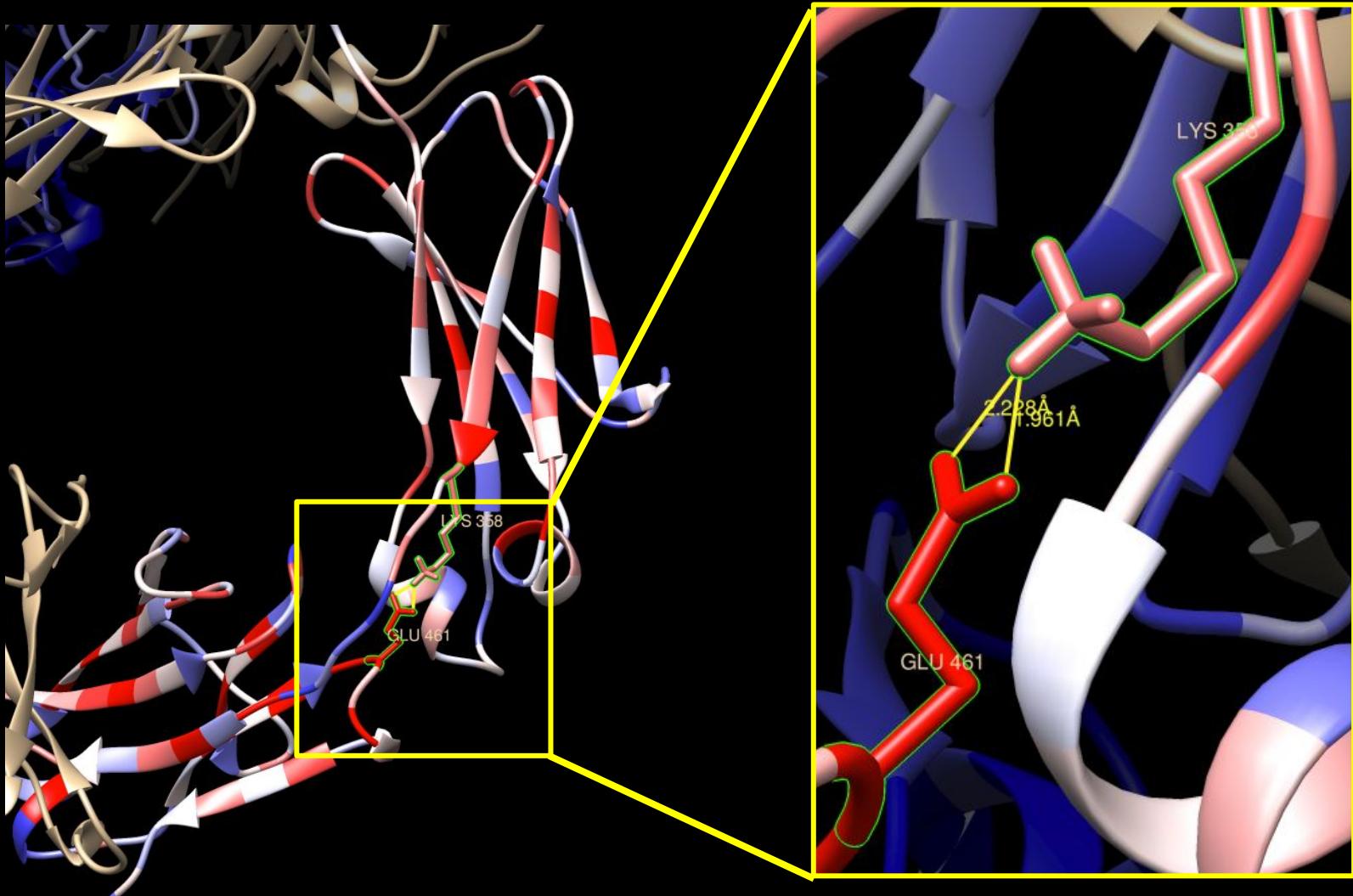
Asn314 is necessary for glycosylation



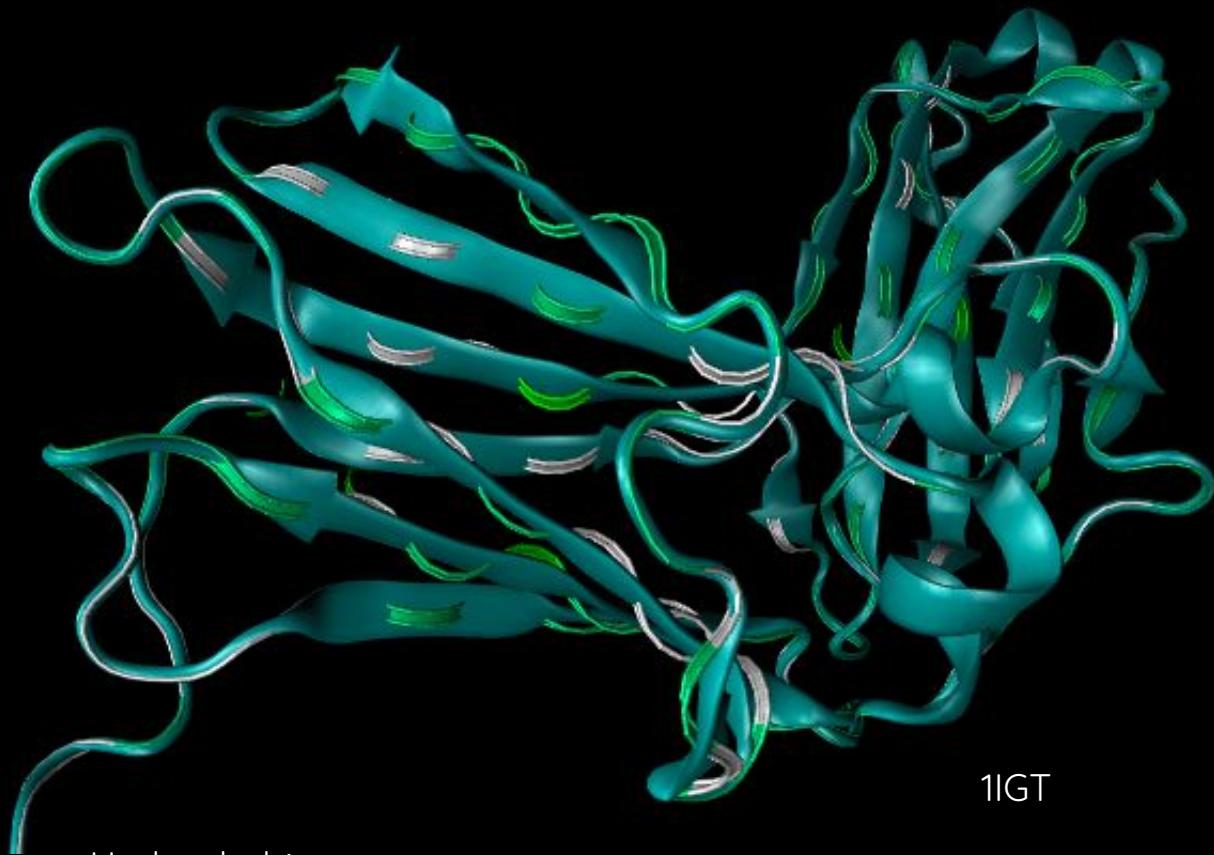
Conservation of charged residues: Within the same region



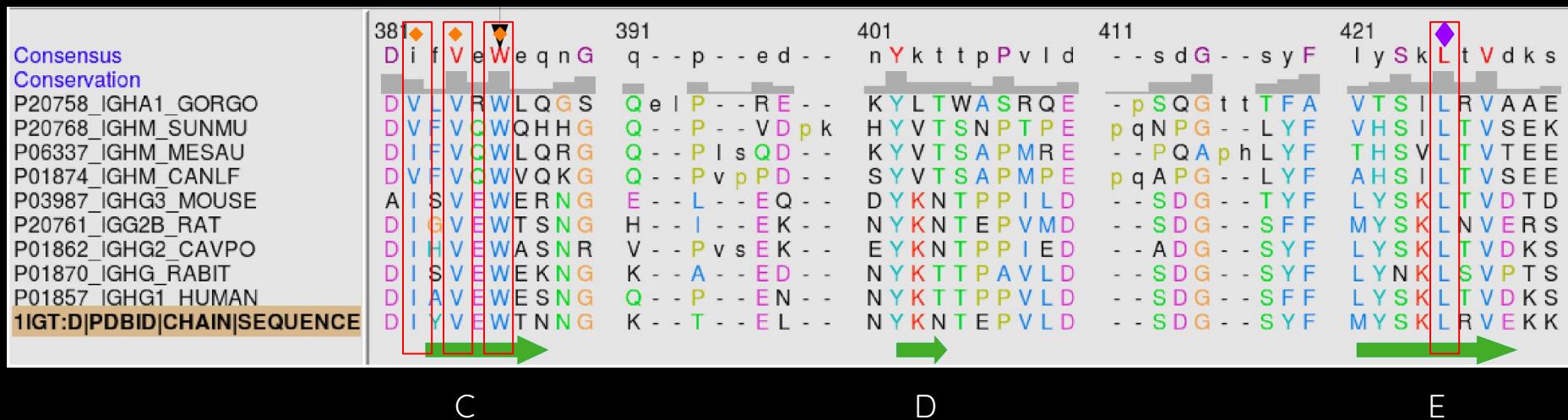
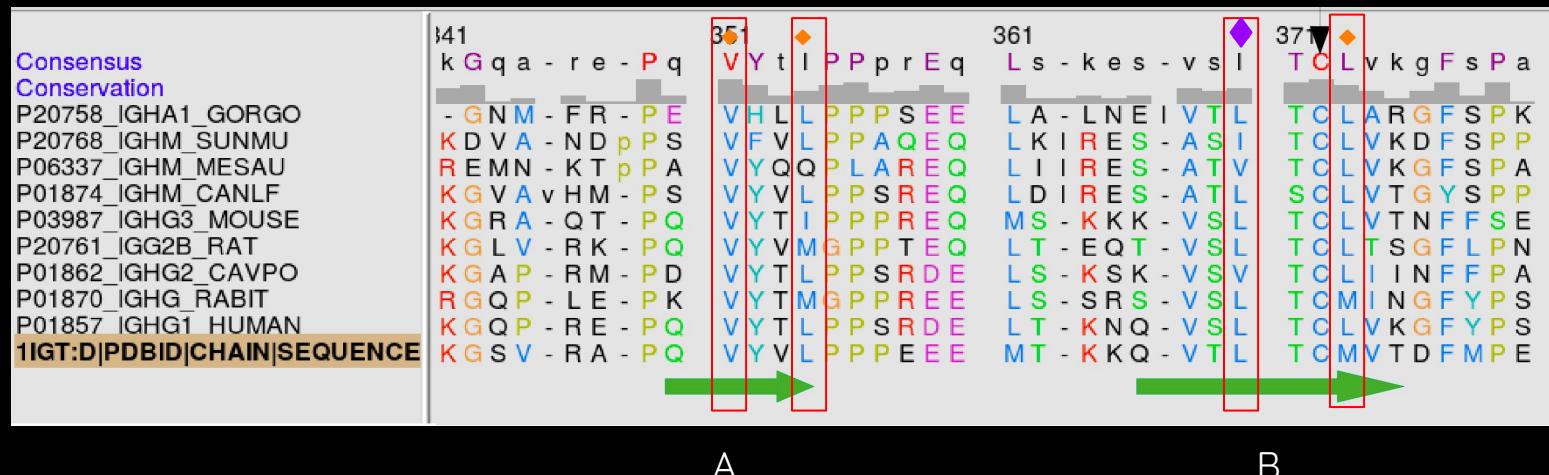
Conservation of charged residues: Between two regions



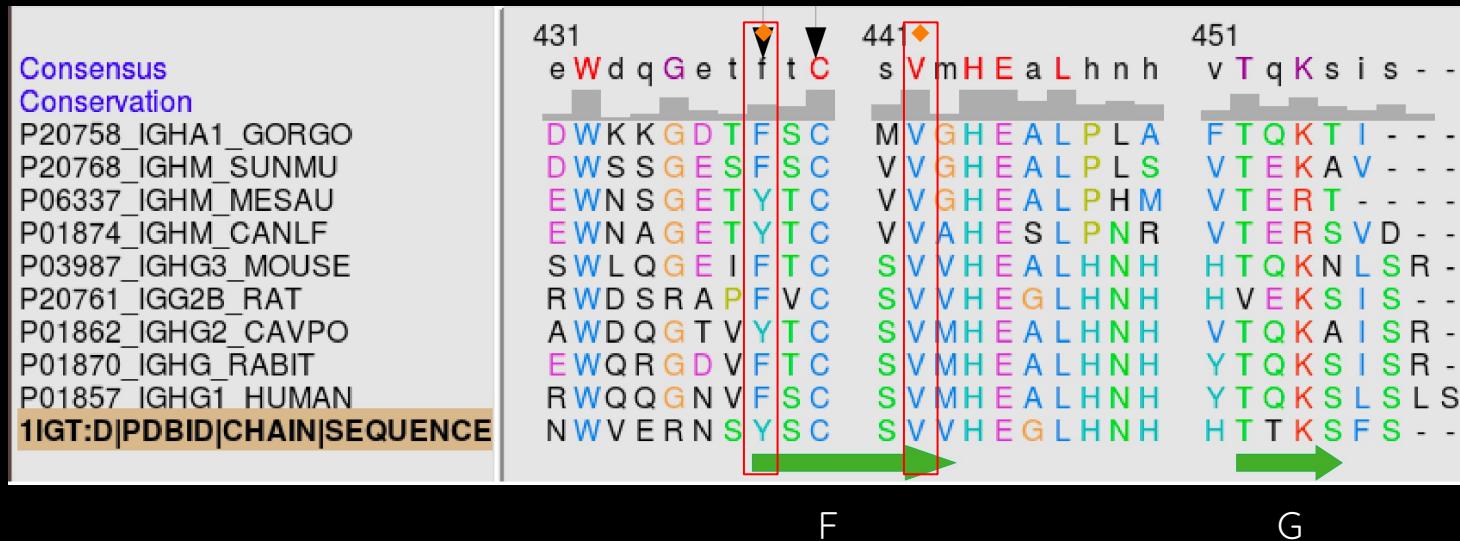
- Hydrophobic side-chains are faced inside of the beta chains



There are many conserved topohydrophobic amino acids

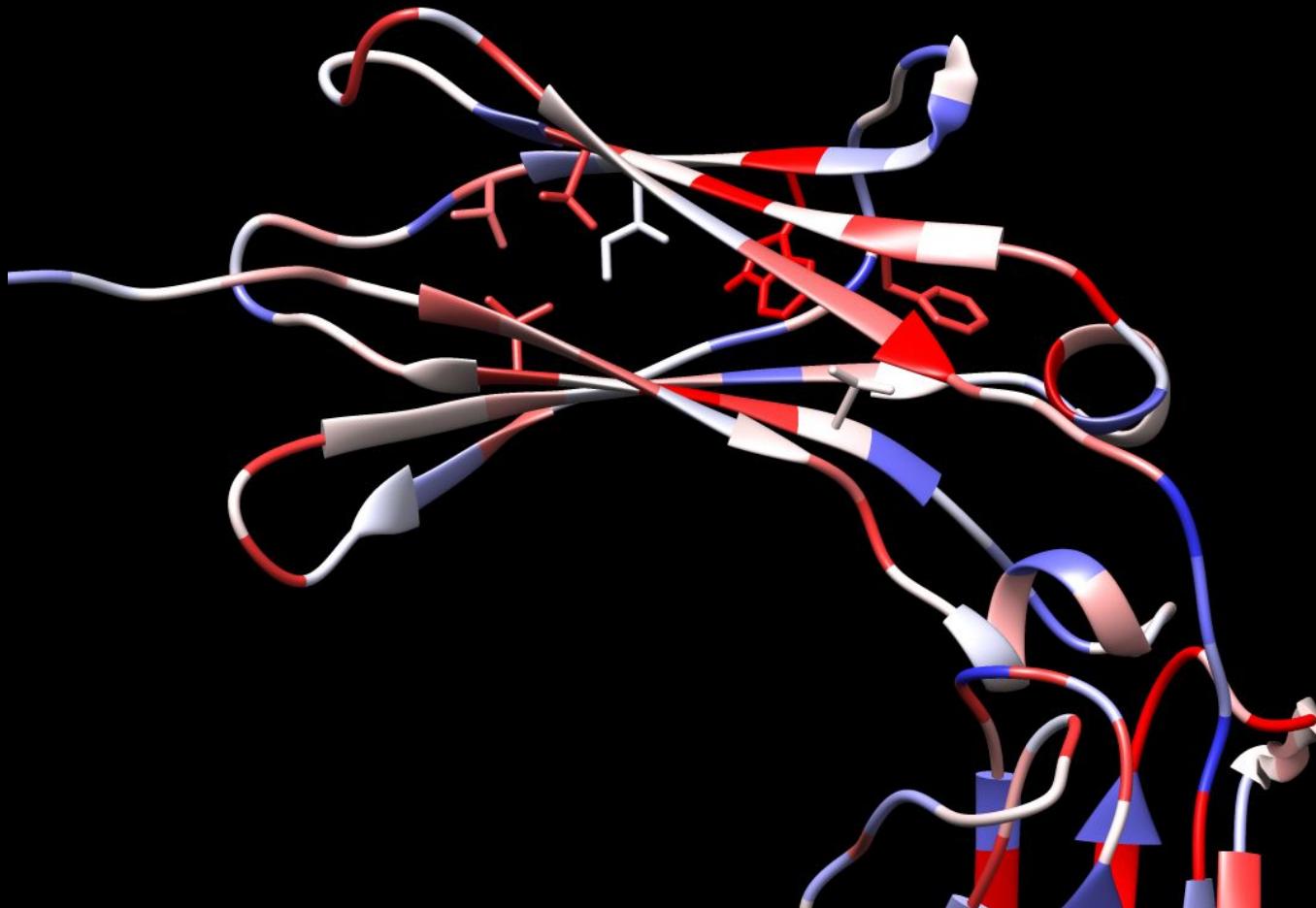


There are many conserved topohydrophobic amino acids



Type of amino acid	Position	Symbol
Topohydrophobic	A3, A7, B4, C1, C3, C5, F1, F5	◆
Aliphatic	B1, E5	◆

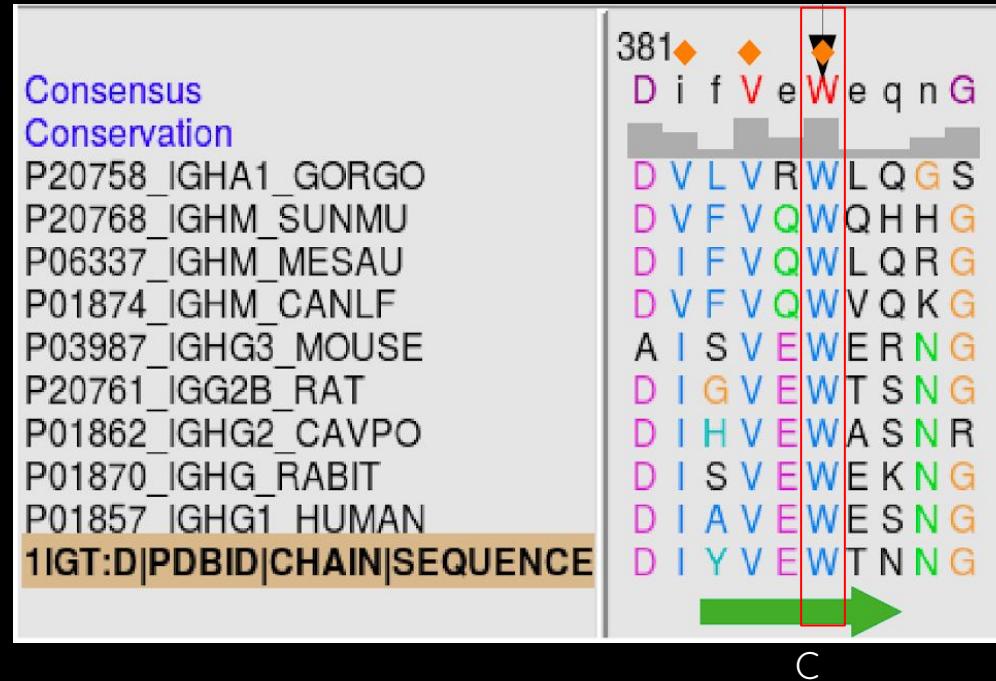
Topohydrophobic side-chains are faced inside of the beta chains



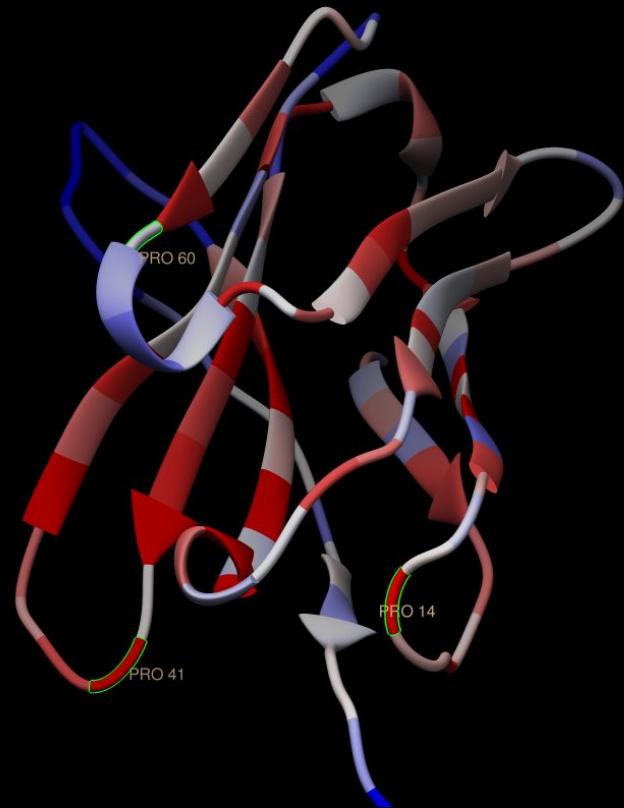
(CH₂)

1IGT

- Trp (in C5) is highly conserved



The role of proline

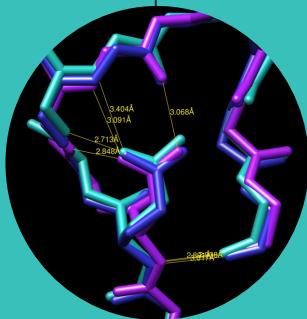


(VH)



(hinge, CH2 and CH3)

CDRs: Canonical Structures



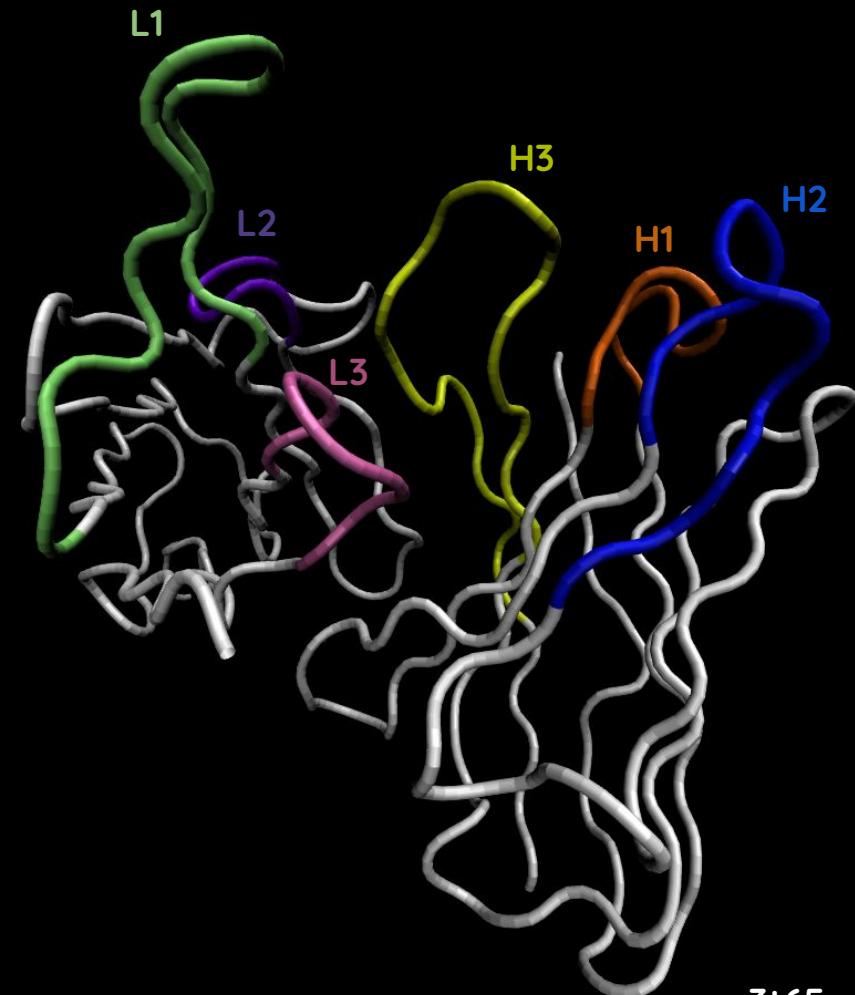
Canonical structures for the hypervariable regions of immunoglobulins

Light chain

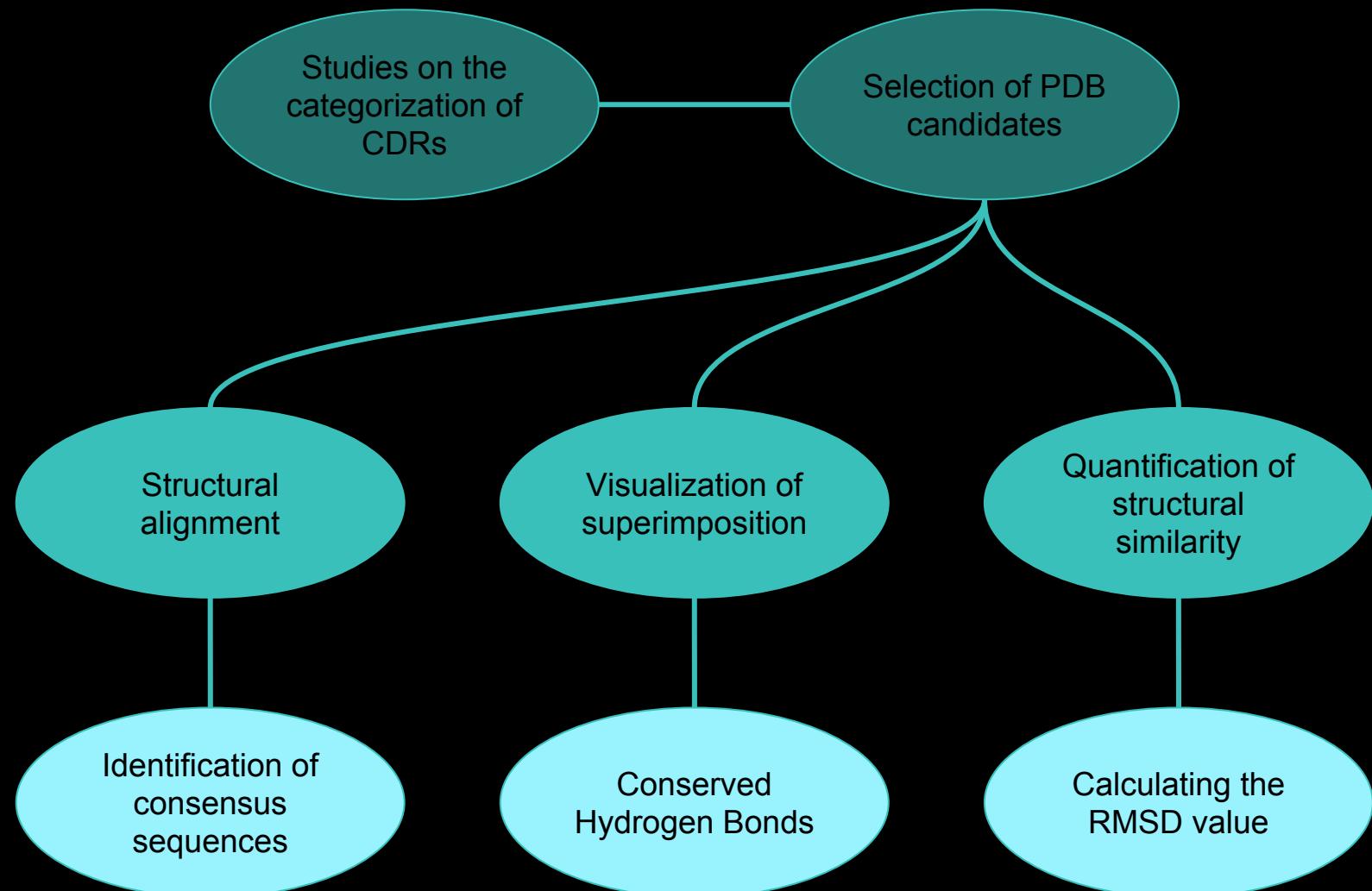
L1	$\kappa \rightarrow 1, 2\alpha, 2\beta, 3, 4, 5, 6$ $\lambda \rightarrow 1, 2, 3\alpha, 3\beta, 4$
L2	1
L3	$\kappa \rightarrow 1, 2, 3, 4, 5$ $\lambda \rightarrow 1\alpha, 1\beta, 1\gamma, 2$

Heavy chain

H1	1, 2, 3
H2	1, 2 α , 2b, 2c, 3 α , 3 β , 3c, 4
H3	No canonical structures



Methodology



L1-κ2

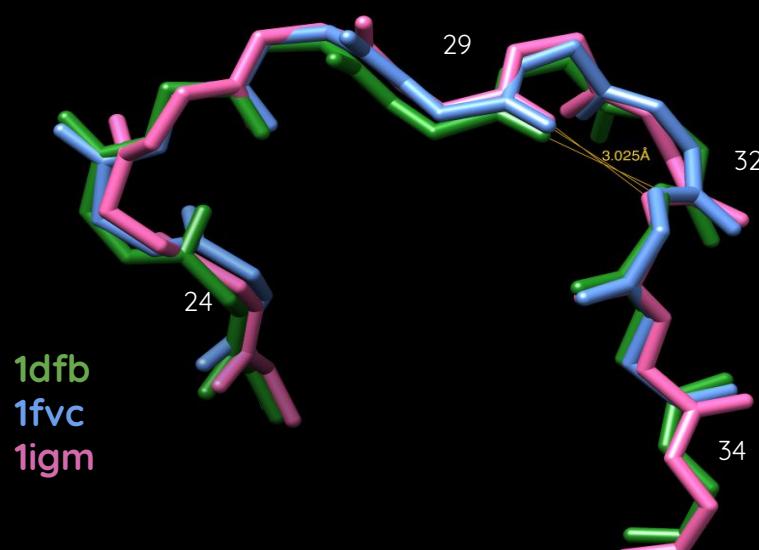
L1-κ2a

HBond: 29 (O) - 32 (N)

Consensus: rASQdisnyla

1dfb	GDRVTTITC	RASQ	SISRWLA	WYQQKP
1fvc	GDRVTTITC	RASQ	DVN	WYQQKP
1igm	GDRVTTITC	QASQ	DISNYLA	WYQQKP

RMSD = 0,721



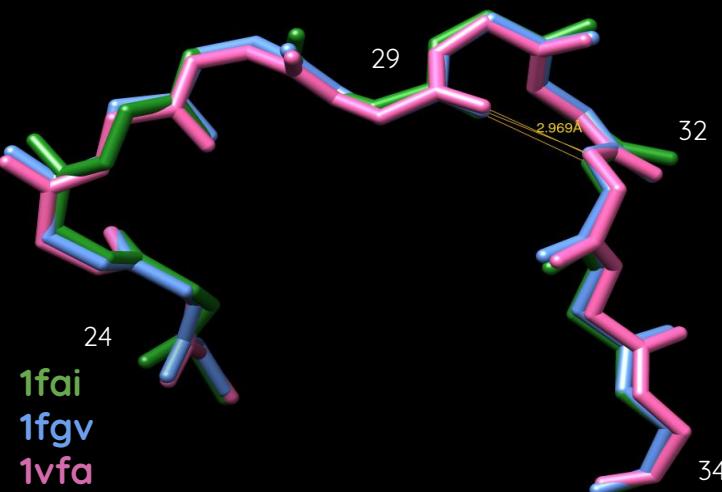
L1-κ2b

HBond: 29 (O) - 32 (N)

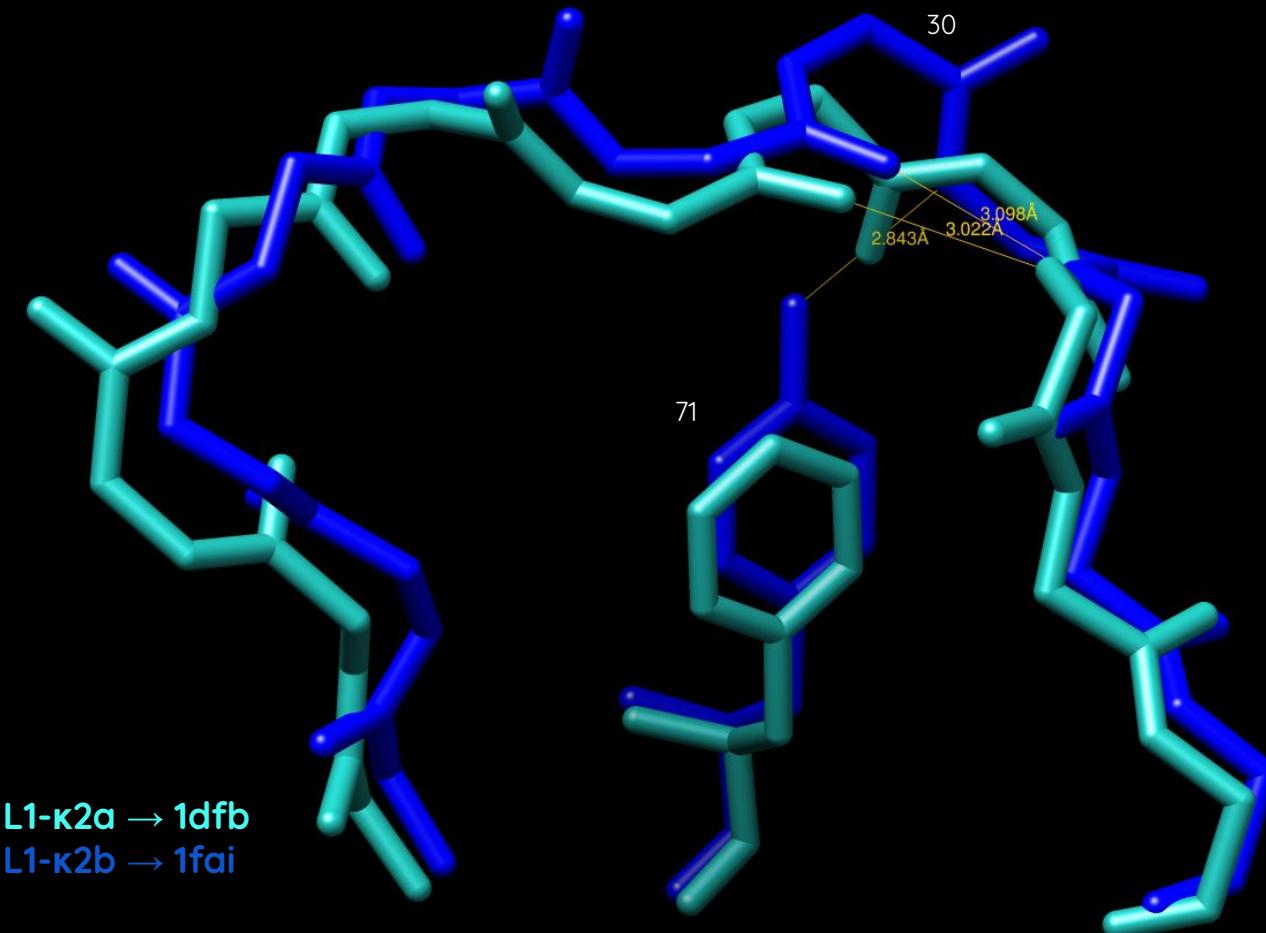
Consensus: rASqdlsnyLn

1fai	ISC	RASQ	DISNYLN	WYQQKP	DGTVK
1fgv	ITC	RASQ	DINNYLN	WYQQKP	PGKAPK
1vfa	ITC	RASG	NIHNYLA	WYQQK	QGKSPQ

RMSD = 0,807



L1-κ2



L1-κ2a
L1-κ2b

Phe71
Tyr71

HBond: 71 (O) --- 31 (N)

● L1- λ 3a

HBond:

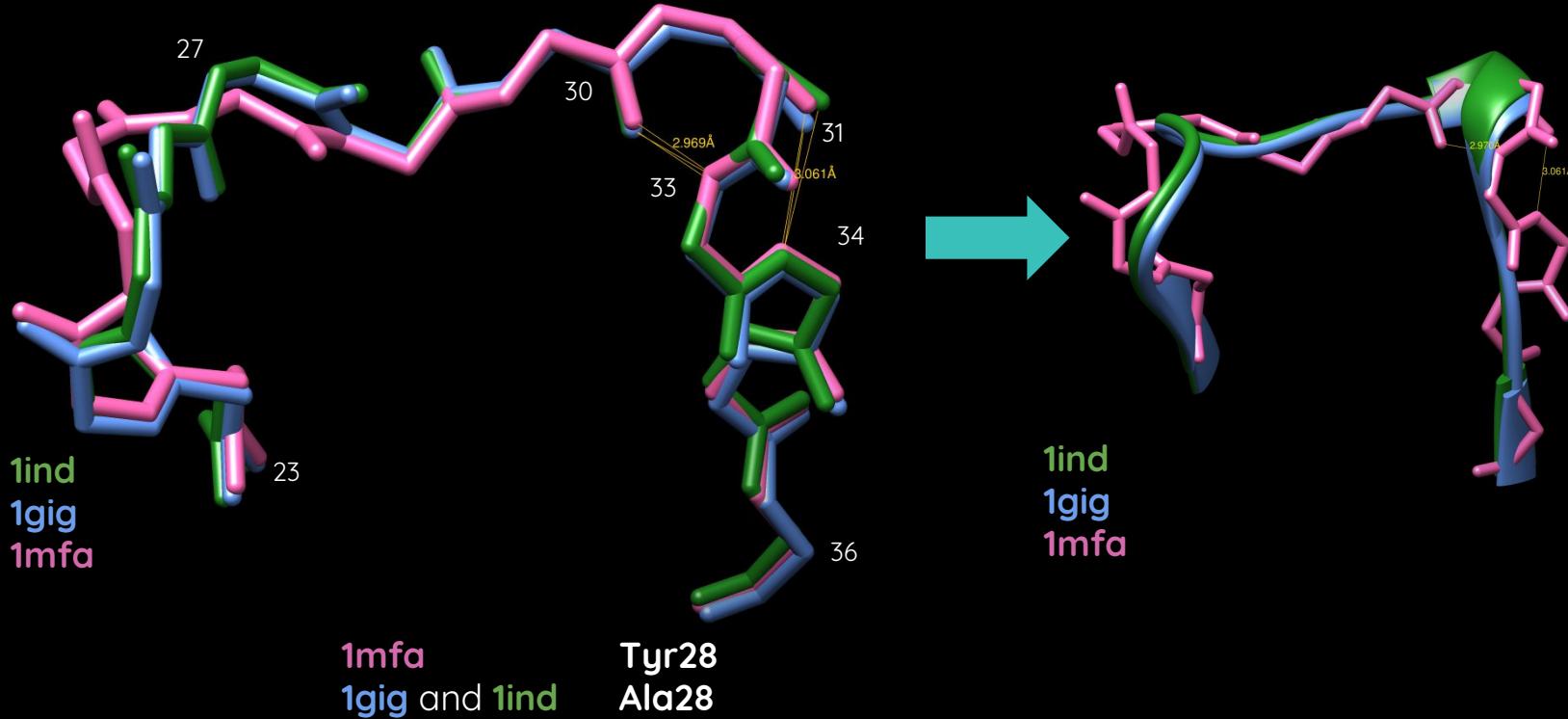
30 (O) - 33 (N)
31 (O) - 34 (N)

Consensus: RSStGavTtsNYAn

1ind
1gig
1mfa

VTLTCRSSTGAVTTSNYANWVQEKP
VTLTCRSSTGAVTTSNYANWVQEKP
VTLTCRSSTGTVTSGNHANWVQEKP

RMSD = 0,890



L2-1

HBond:

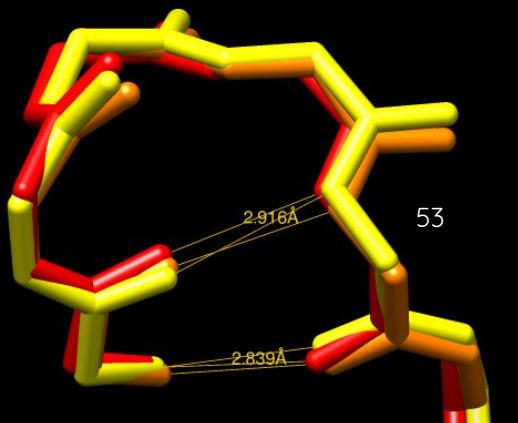
49 (N) - 53 (O)

49 (O) - 53 (N)

Consensus: **Y-asnlas**

1fgv	KLLI	YYTSTLES	GVPSRFGSGS
1flr	KVLI	YKVSNRFS	GVPDRFGSGS
1hil	KVLI	YWASTRES	GVPDRFTGSGS

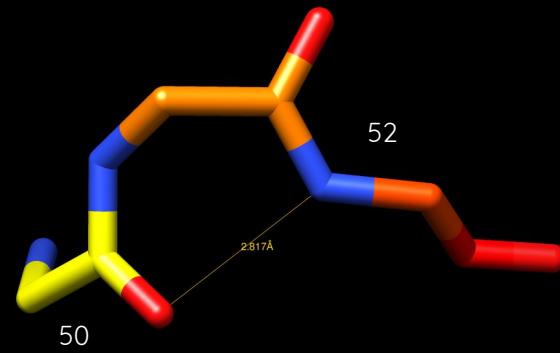
RMSD = 0,597



1fgv

1flr

1hil



HBond: 50 (O) - 52 (N)

● L3-K1

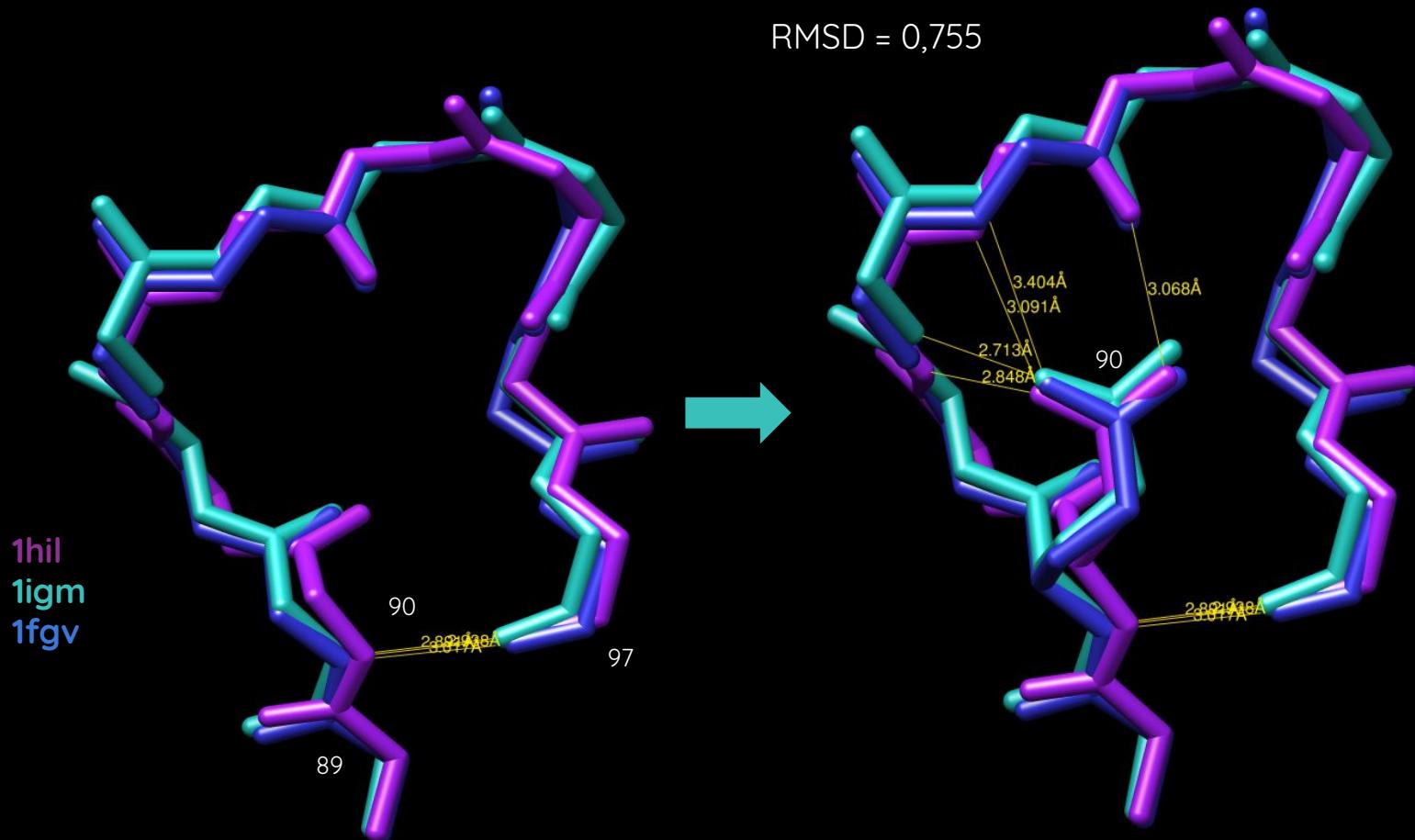
HBond:

90 (N) - 97 (O)

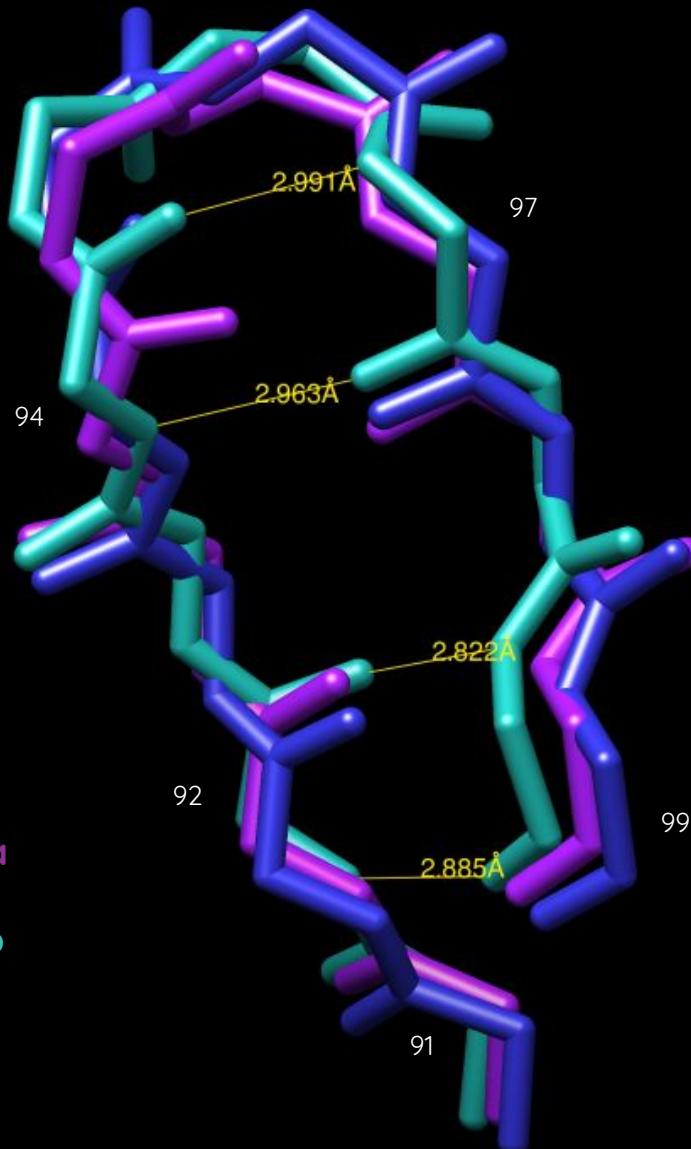
Consensus: qQgss-PIT

1hil	/YYCQNDYSNPLT	FGGGTKLELK
1fgv	TYYCQQGNTLPP	FGAGTKVEIK
1igm	TYYCQQYQNLPLT	FGPGTKVDIK

RMSD = 0,755



L3-λ1



HBond:
92 (N) - 99 (O)
92 (O) - 99 (N)
94 (N) - 97 (O)
94 (O) - 97 (N)

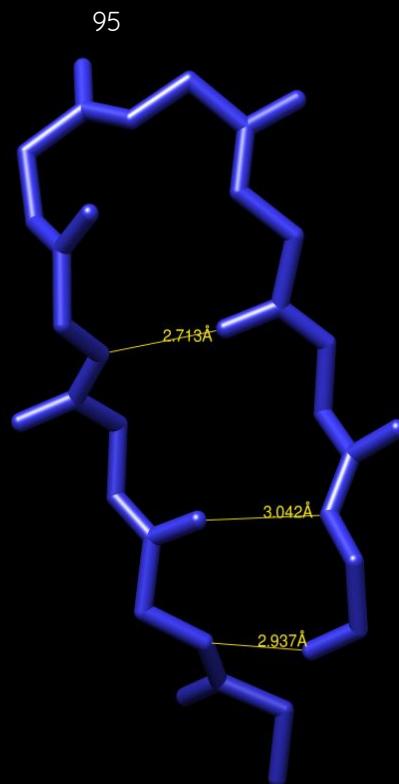
Consensus: alw-snhwv

1mfa	YFCALWSNNHWI	FGGGT
1gig	YFCALWYSNHWV	FGGGT
7fab	YYCQSYDRSLRV	FGGGT

RMSD = 1,045

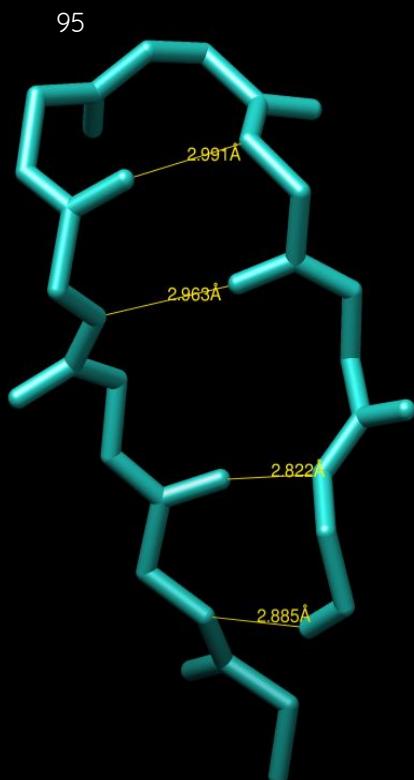
L3-λ1

L3-λ1a



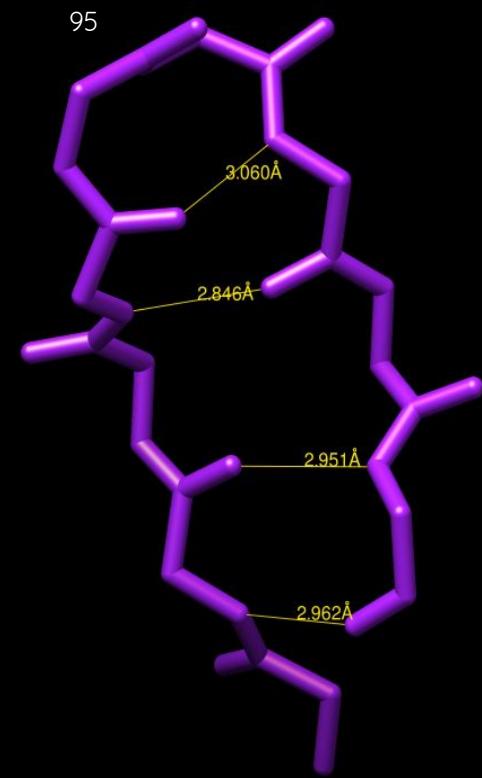
1gig

L3-λ1b



7fab

L3-λ1c



1mfa

H1-1

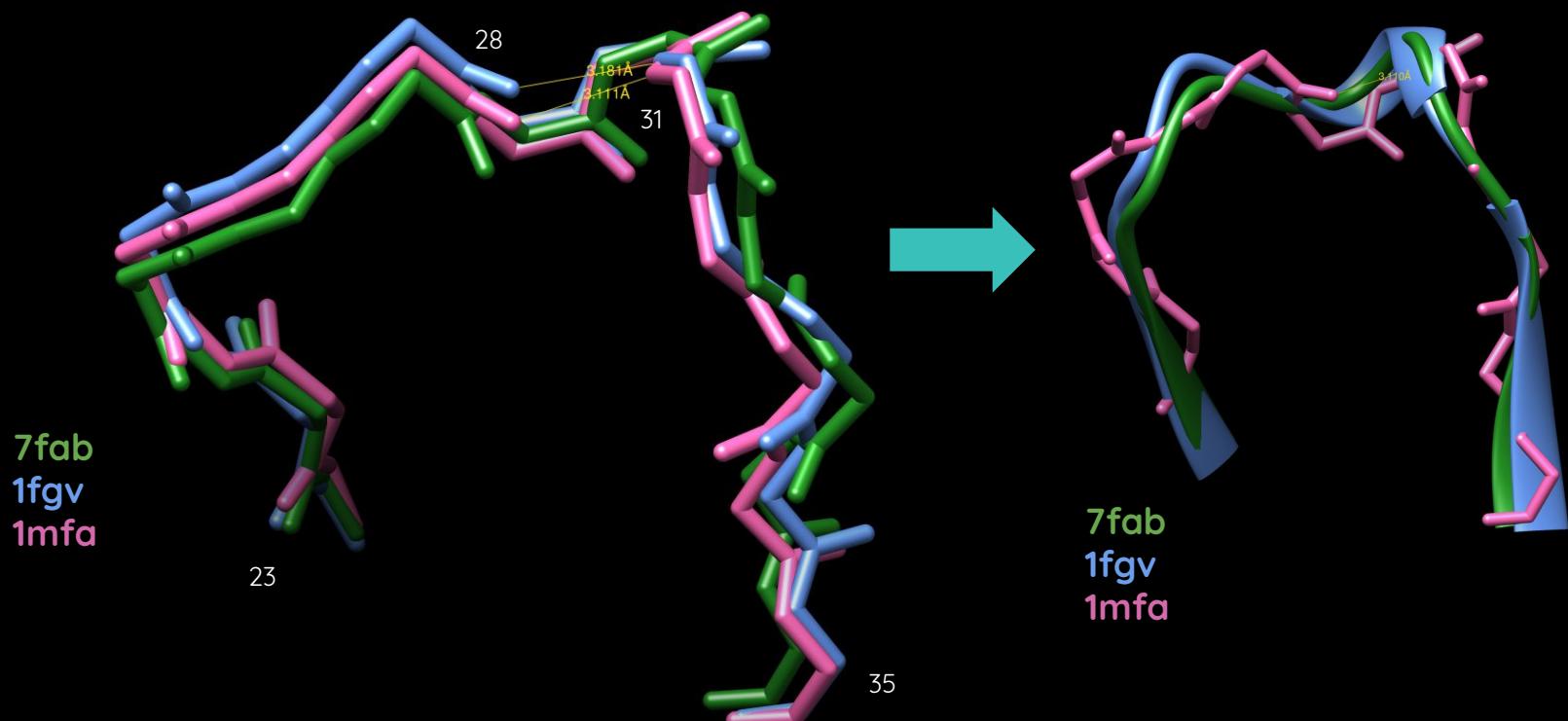
HBond:

28 (O) - 31 (N)

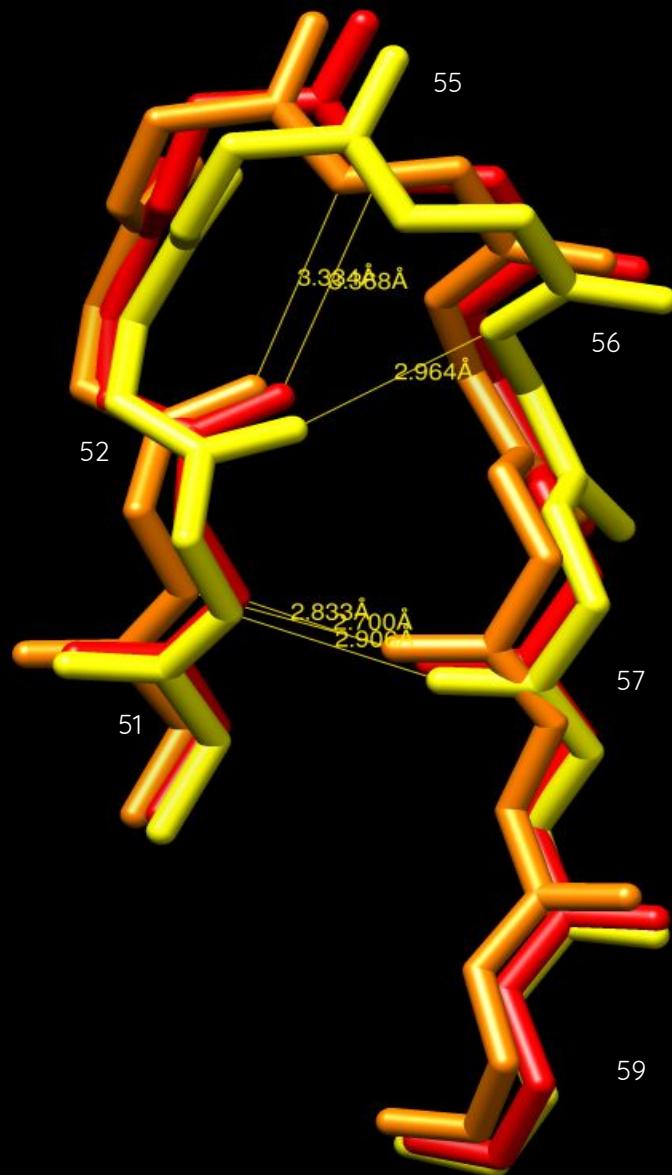
Consensus: ka**SG**ftftdyymh

7fab	LTCTVSGTSFDDYYWT	WVRQP
1fgv	LSCATSGYTFTEYTMH	WMRQA
1mfa	MSCKASGYTFTNYWMH	WIKQR

RMSD = 2,008



• H2-2a



HBond:
52 (N) - 57 (O)
52 (O) - 55 (N)
52 (O) - 56 (N)

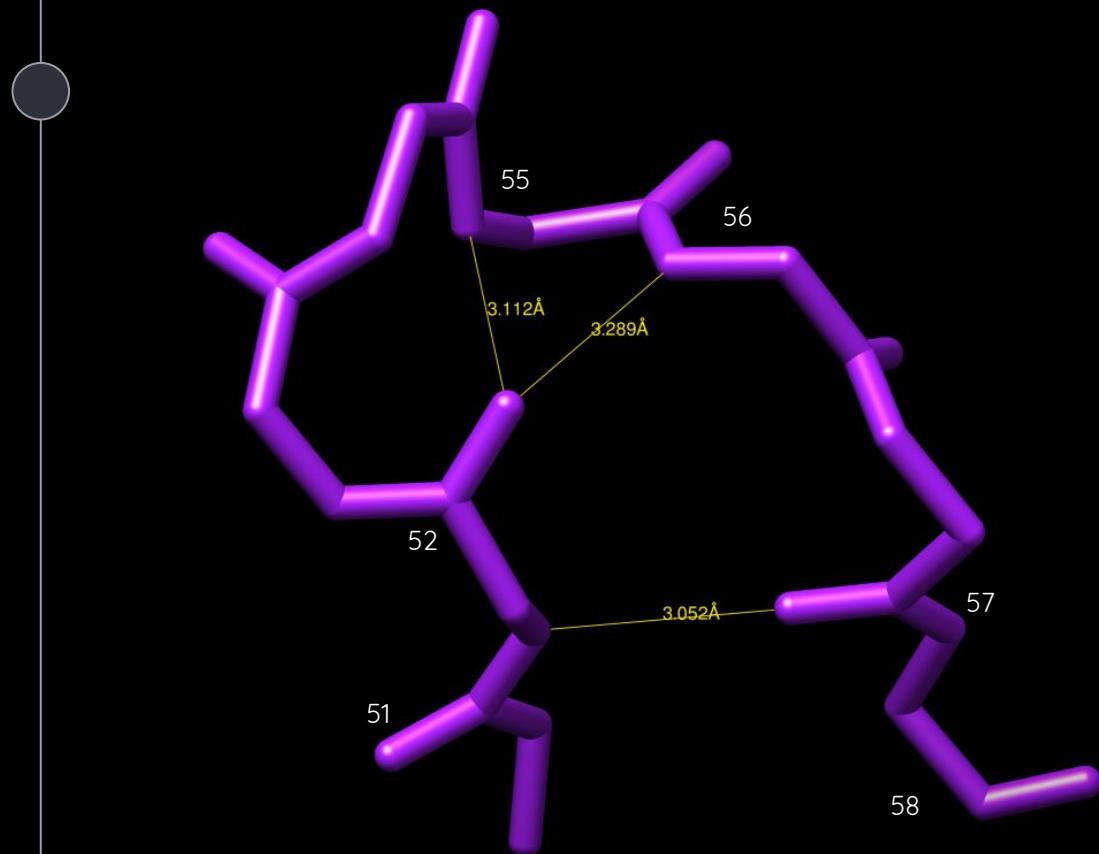
Consensus: -Iypgng-t-

1mfa
1fgv
1fvc

GLEWIGAIYPGNSATFY
GLEWVAGINPKNGGTSY
GLEWVARIYPTNGYTRY

RMSD = 0,931

H2-3b



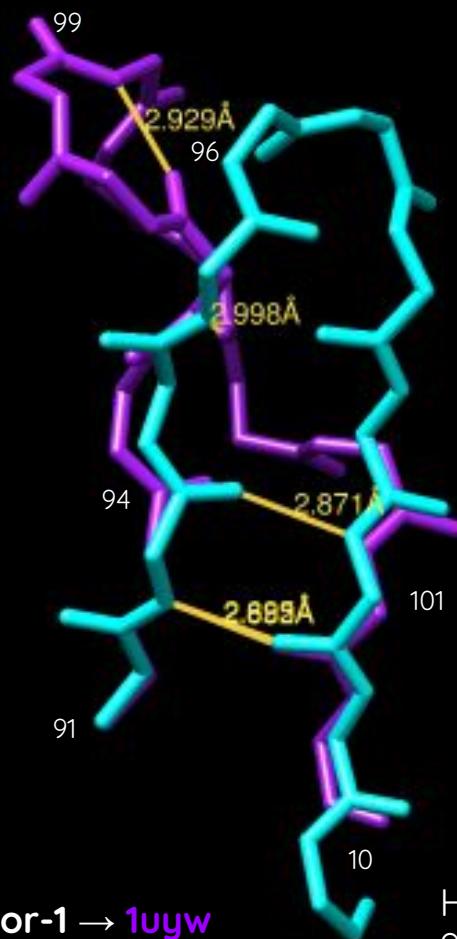
1igm

Gly53
Gly55
Gly56

HBond:
52 (N) - 57 (O)
52 (O) - 55 (N)
52 (O) - 56 (N)

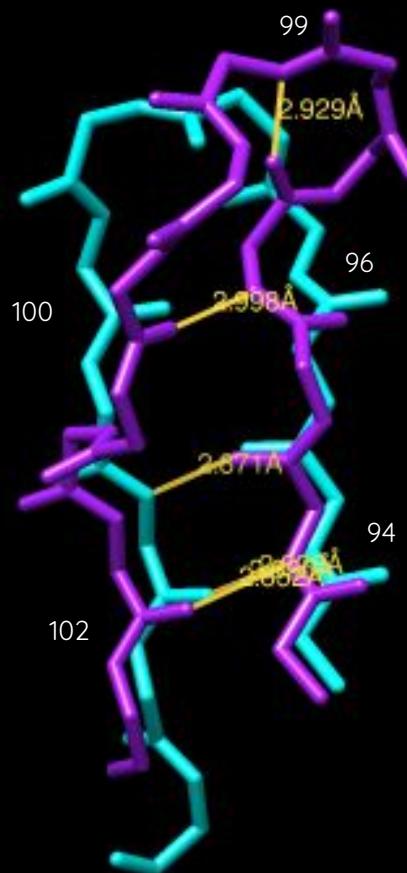
Consensus: -lssgggnty

● H3



H3-anchor-1 → 1uyw
H3-anchor-2 → 2j88

HBond 1uyw:
96 (O) - 99 (N)
96 (O) - 100 (N)
94 (N) - 102 (O)



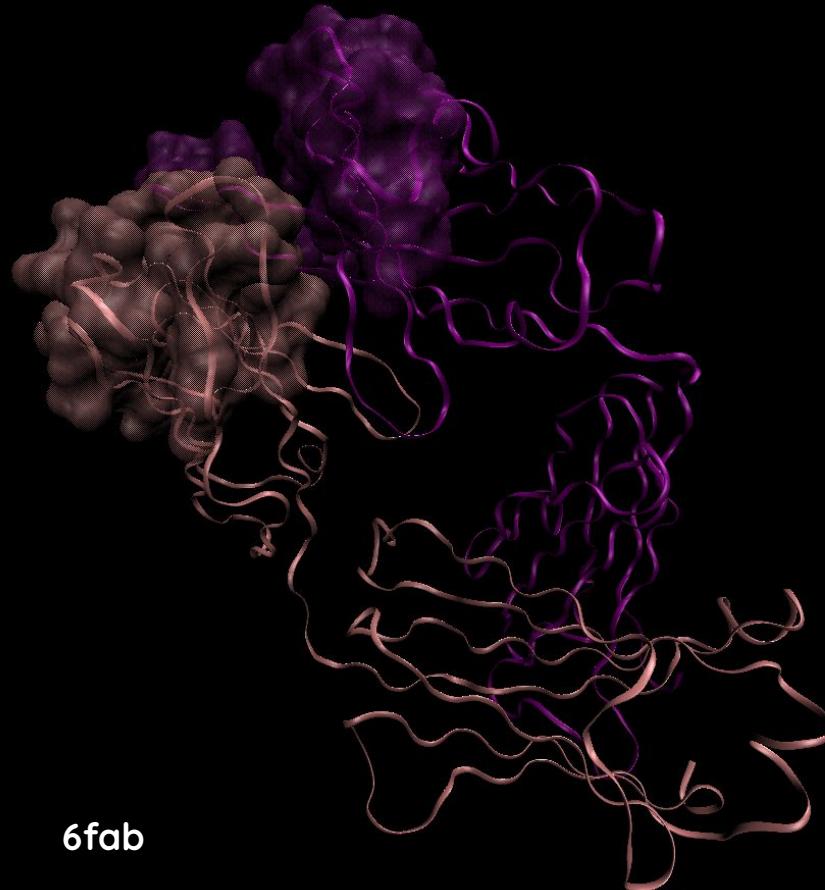
HBond 2j88:
94 (N) - 101 (O)
94 (O) - 101 (N)

Strong Geometrical Restrictions Associated to the Mechanism of Immune Recognition

Nomenclature: class H1-H2-L1-L2-L3
H1, L2 & L3 are usually type 1, resulting
in 1-X-Y-1-1

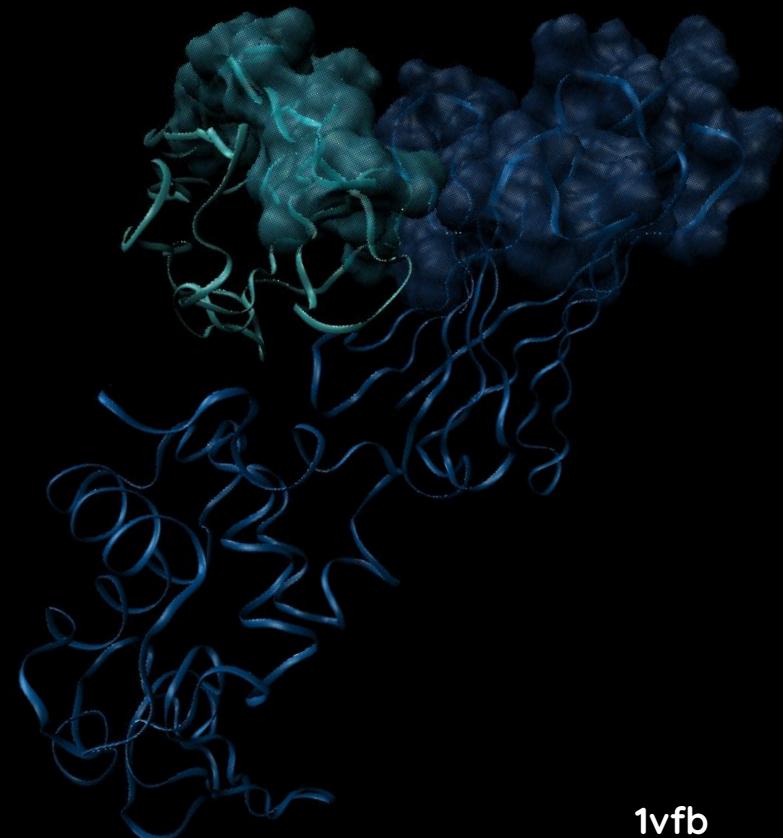
Class 1-4-4-1-1

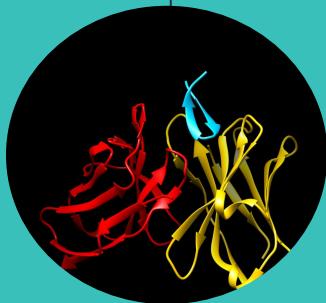
- Long loops
- Small molecules (haptens)
- Cleft at the antigen-binding sites



Class 1-1-2-1-1

- Short loops
- Large antigens (proteins)
- Flat antigen-binding sites





Characterization of an antigen-antibody interaction: antibody AP33 and E2 antigen from HCV

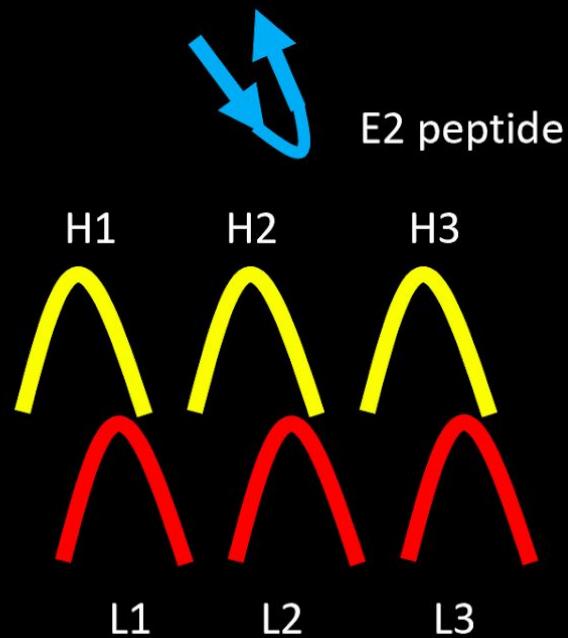
- The crystallographic structure



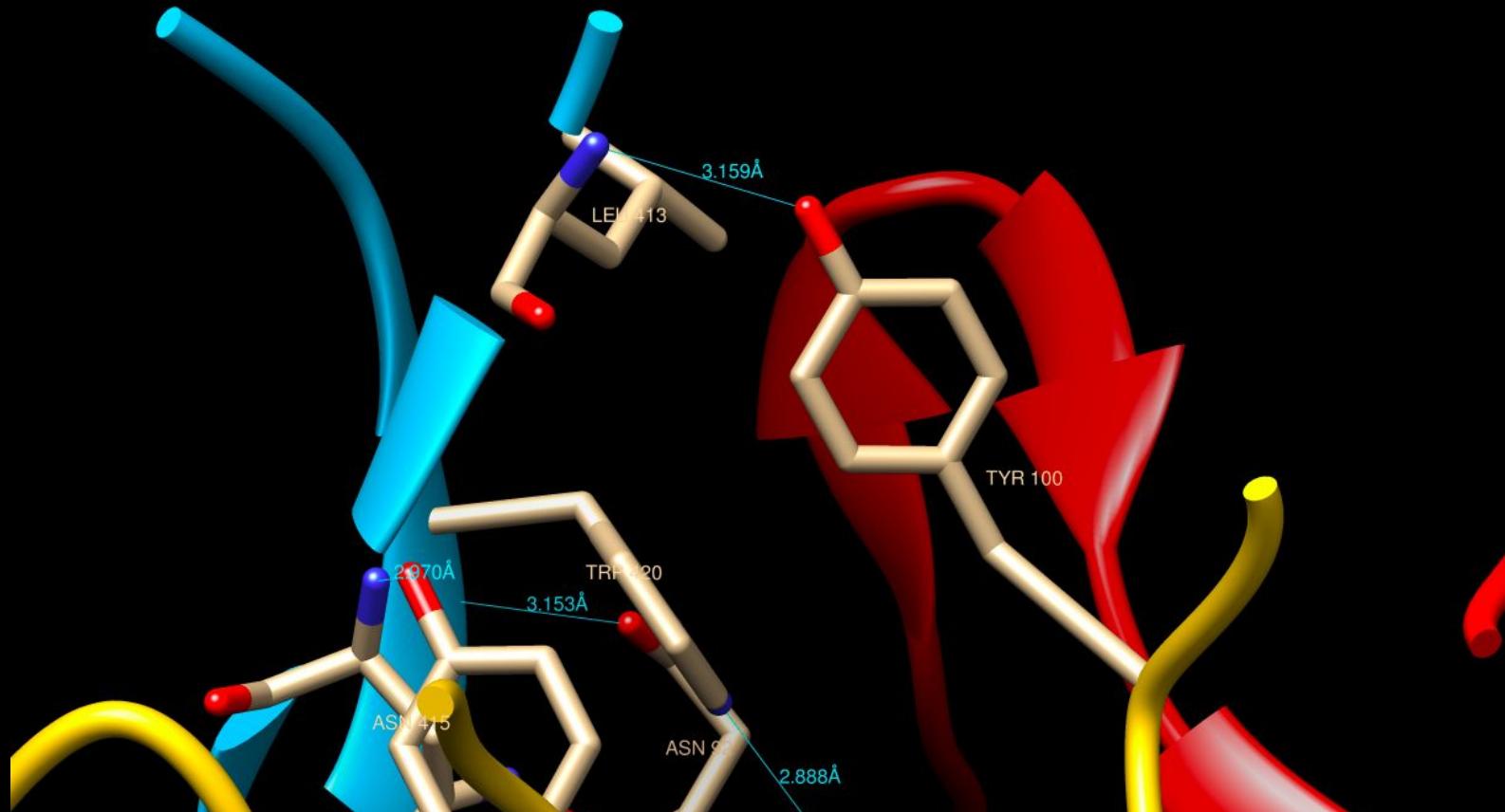
LC
HC
Antigen

(Fab)

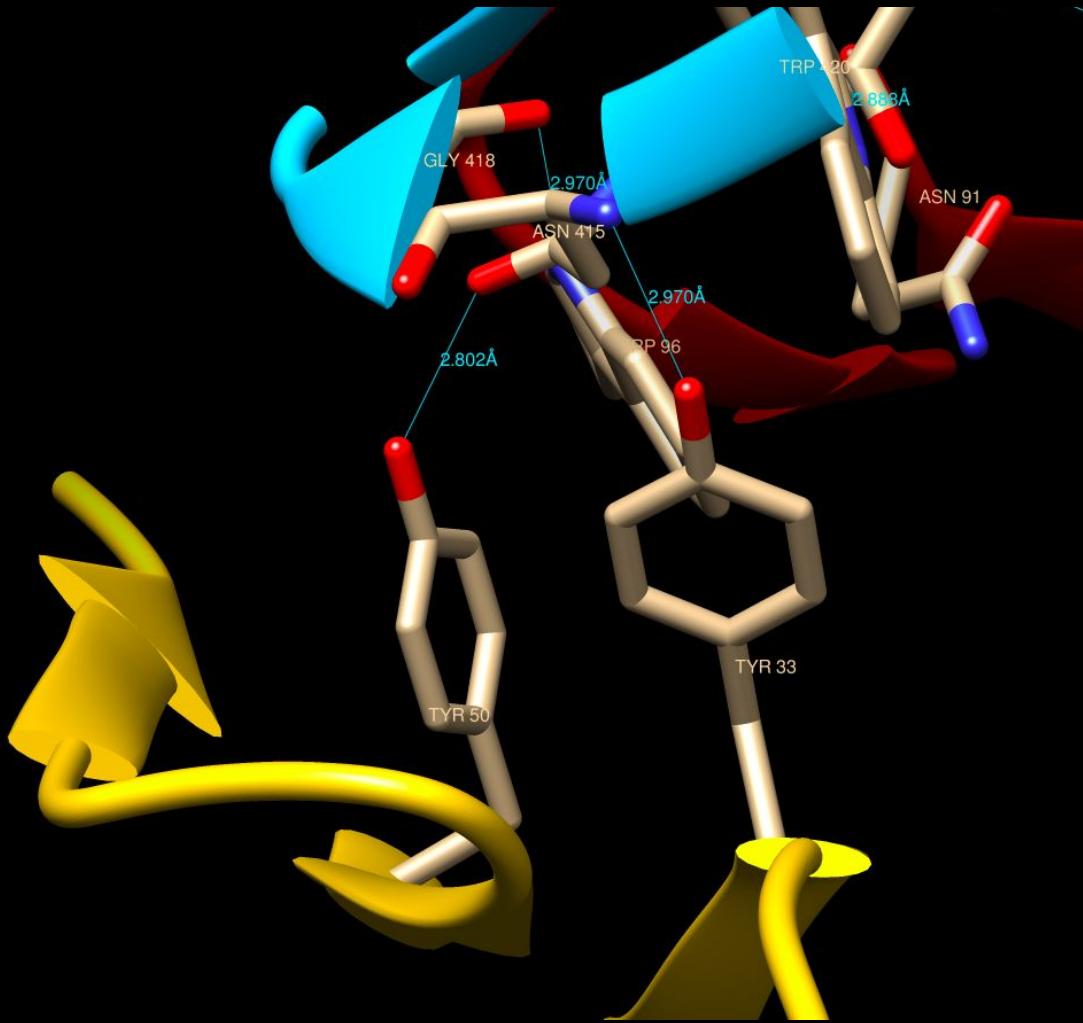
Hydrogen bonds



- Hydrogen bonds

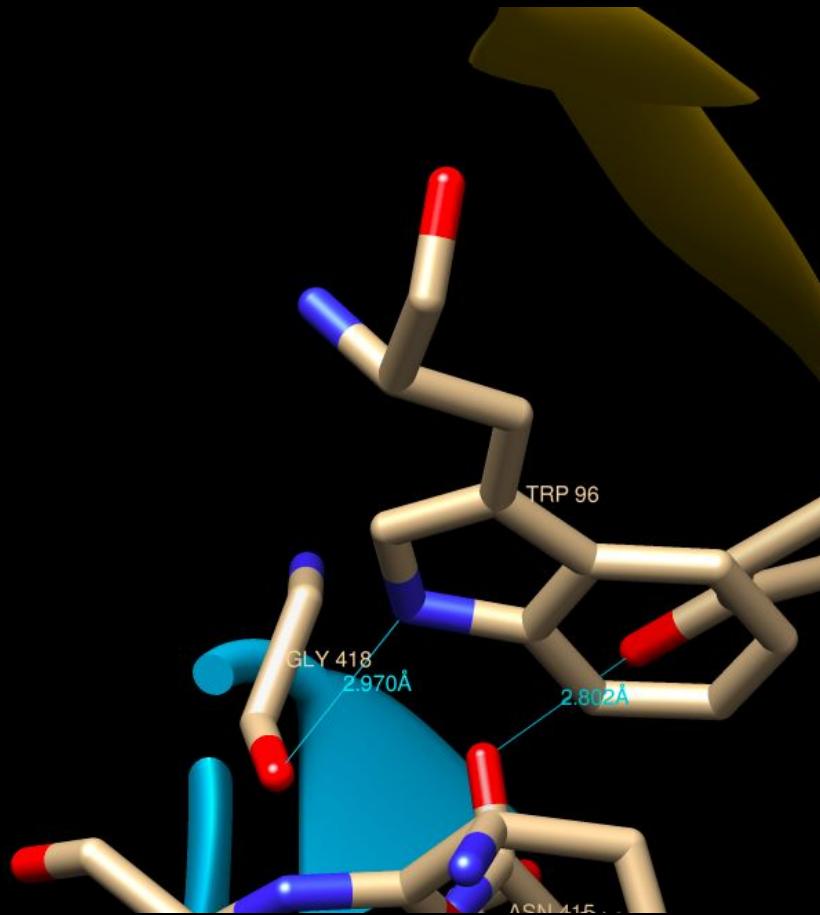


Hydrogen bonds



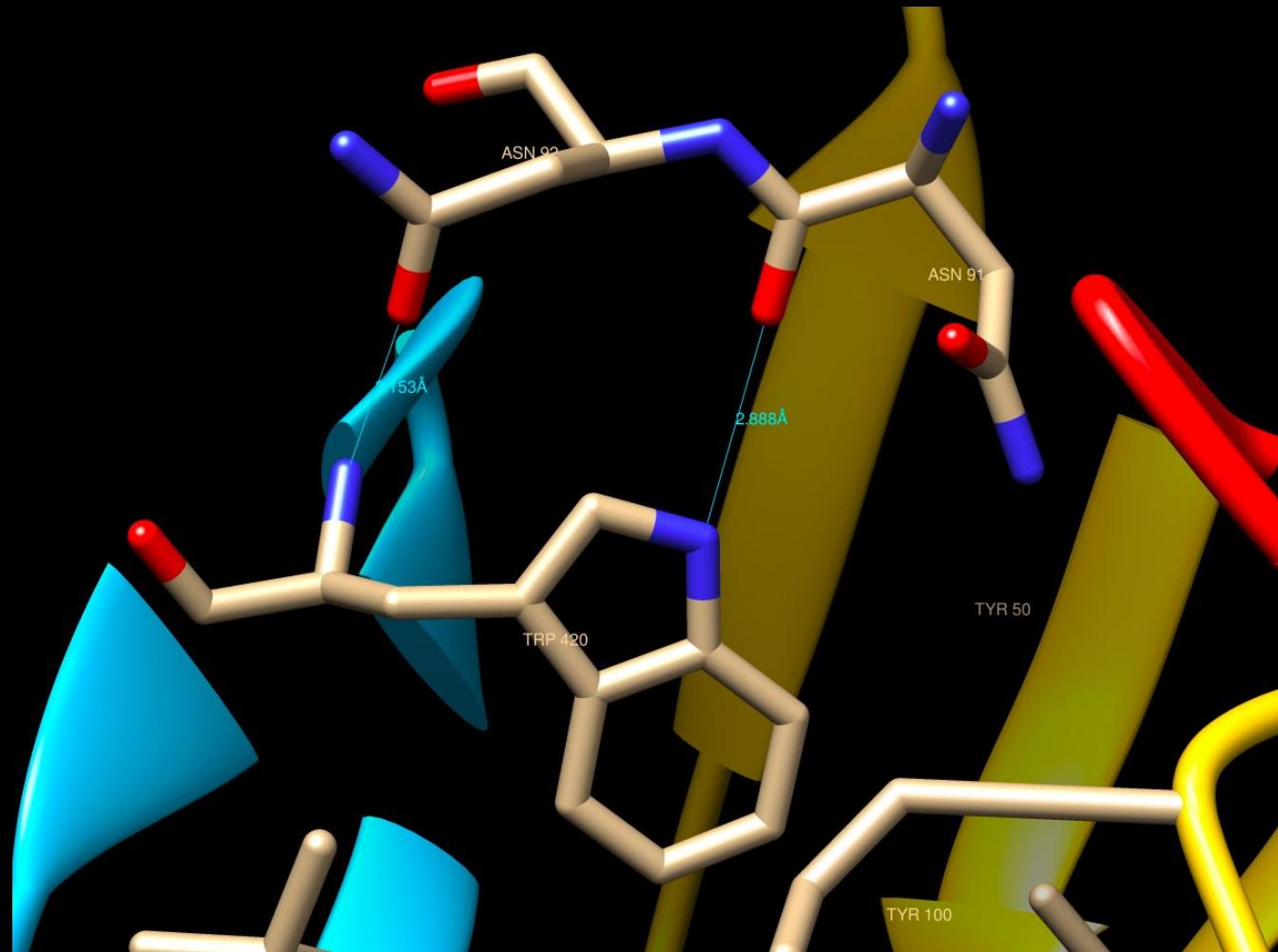
ASN 415 - TYR 50
ASN 415 - TYR 33

- Hydrogen bonds



GLY 418 - TRP 96

- Hydrogen bonds



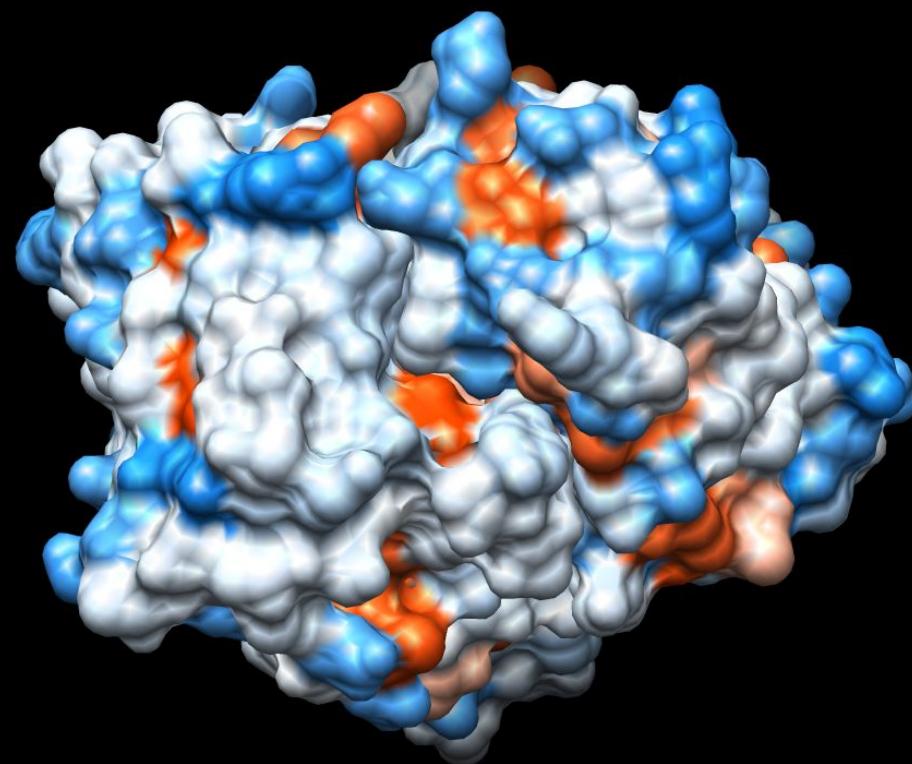
TRP 420 - ASN 91
TRP 420 - ASN 92

Hydrophobicity surface



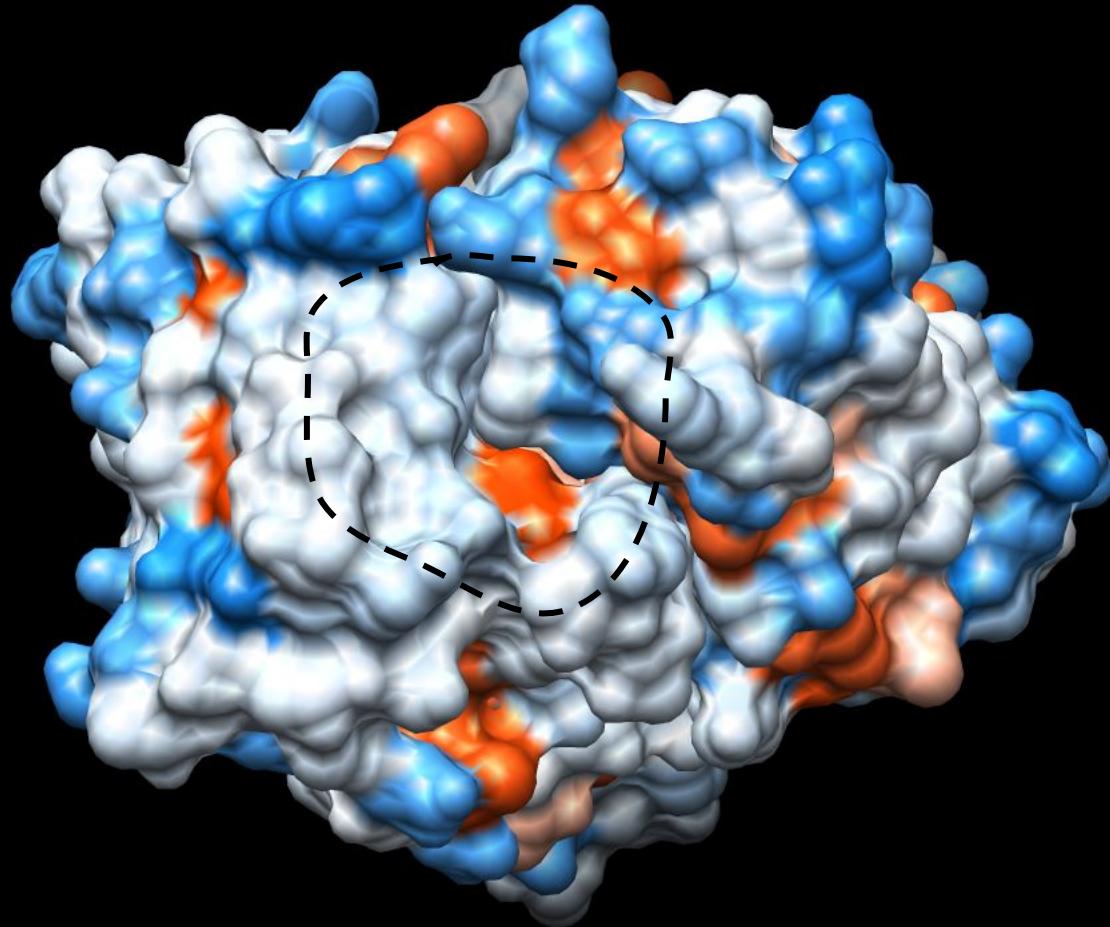
Hydrophilic

Hydrophobic



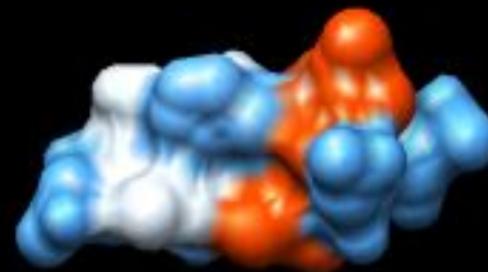
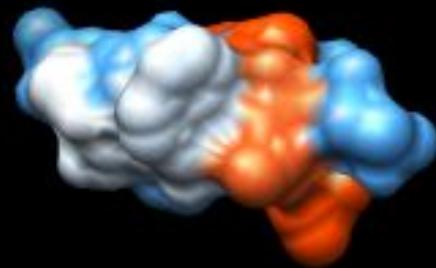
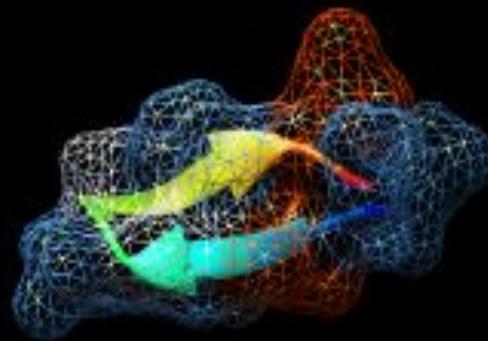
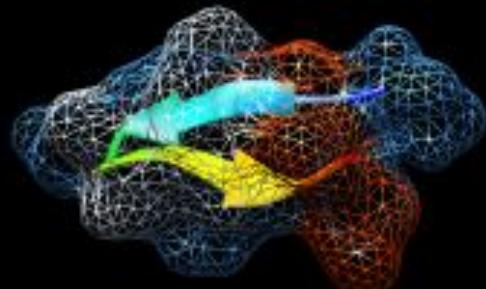
Fab (top)

- **Hydrophobicity surface**



Fab (top)

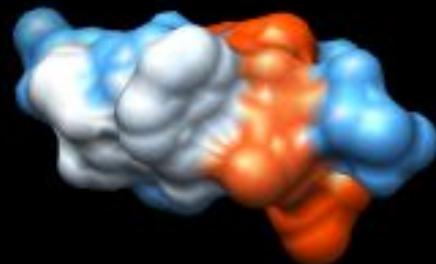
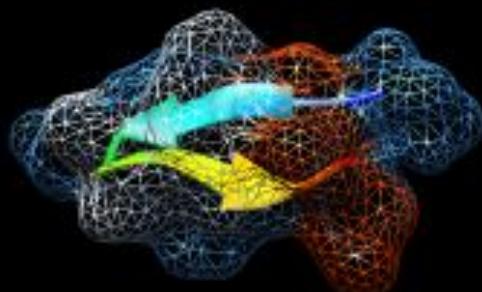
- **Hydrophobicity surface**



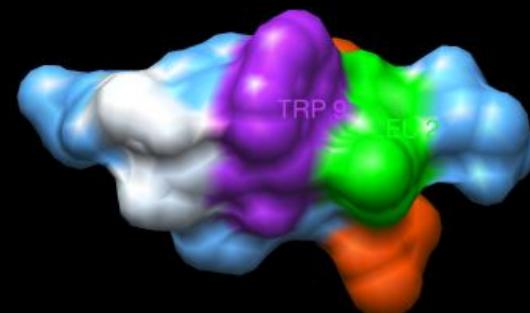
Ligand (bottom)

Ligand (top)

- **Hydrophobicity surface**



Ligand (bottom)



Trp426
Leu414

- **Hydrophobicity surface**



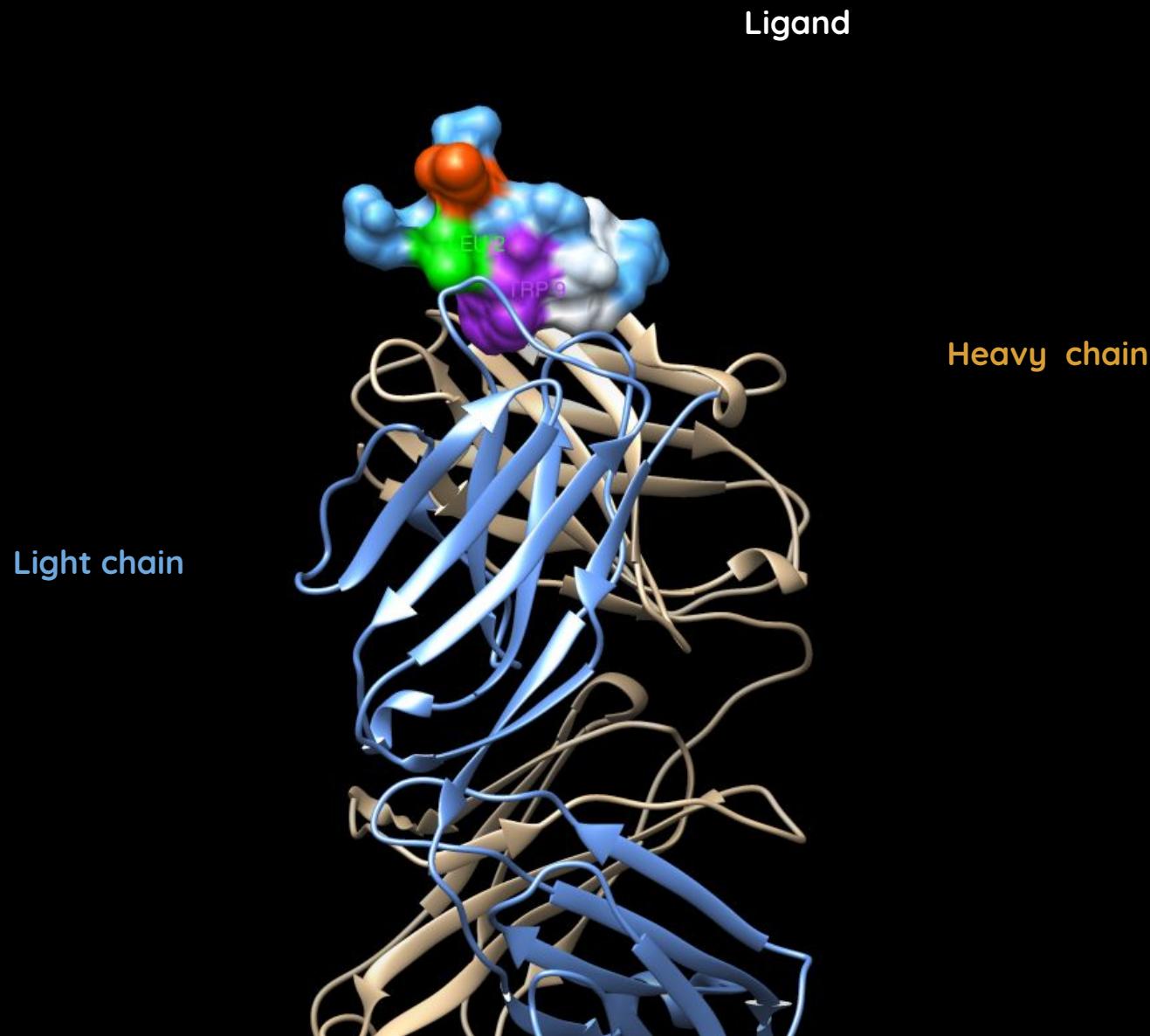
Fab + ligand

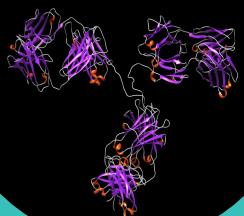
- **Hydrophobicity surface**



Fab + ligand

- **Hydrophobicity surface**

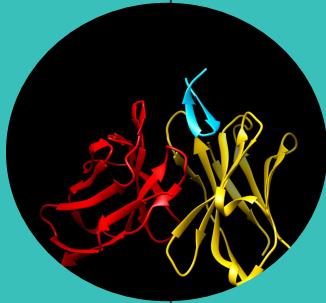




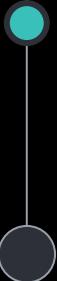
Conclusions

Conclusions

- Immunoglobulins are heterodimers whose main motif is a greek-key, which commonly consists of seven beta strands conforming 2 beta-sheets.
- Hydrogen bonds, disulfide bonds and hydrophobic bonds are the main responsible of the maintenance of the fold, the secondary structure and the tertiary and quaternary structure, respectively.
- Besides Igs being very variable proteins, there are some highly conserved residues: Cys-forming disulfide bonds, topohydrophobic positions, and residues with special roles (Asn, Pro, charged residues...).
- Despite CDRs being hypervariable regions, there are some chain conformations that are more frequently found, defining canonical structures.
- CDRs are a clear example of the fact that structure is generally more conserved than sequence.
- Correlation between the architecture of the antigen-binding site and their recognised antigen provides insight into the general mechanism of the molecular recognition process mediated by Igs.
- In the antigen-antibody binding, both hydrophobic interactions and hydrogen bonds are formed between the CDRs of the Fab and the epitope.



Multiple Choice Questions



1. About the topohydrophobic amino acids of the immunoglobulins:
 - a) They create hydrophobic interactions that permit the fold of the two beta-sheet domains
 - b) They are located on the surface of Ig's, solubilizing them
 - c) The two of answers above are correct
 - d) They are not conserved along evolution
 - e) All of the above are correct
2. About the interaction between the antigen E2 and the antibody AP33:
 - a) The antigen interacts with the Fc region of the antibody
 - b) Hydrogen bonds are not involved in this interaction, only hydrophobic interactions and Van der Waals forces.
 - c) When both proteins interact, hydrophobic surfaces are joined and only the polar parts remain exposed.
 - d) All of the CDR's of AP33 are involved in the interaction with E2.
 - e) All of the above are correct
3. Choose the correct option regarding cysteine residues in immunoglobulins:
 - a) Cysteine residues are not conserved through evolution.
 - b) Cysteine residues do not form hydrogen bonds.
 - c) Cysteine residues are located in a particular position.
 - d) There is not any Cys joining the heavy chain with the light chain.
 - e) Each immunoglobulin fold consist of just 1 Cys residue.

4. About the role of proline in immunoglobulins:

- a) Proline is not favored neither in beta-sheets nor alpha-helix
- b) Proline is the amino acids that forms most hydrogen bonds.
- c) Both are correct.
- d) Proline is not important for the structure of a protein, and thus is not conserved at all.
- e) All are correct.

5. About canonical structures in CDRs, find the correct answer:

- a) Because of its variability, canonical structures are hard to establish in H3 CDR.
- b) The first classification was done by Chothia et al in 1989.
- c) The two of the above are correct.
- d) Canonical structures only exist in the light chain CDRs.
- e) All kind of CDRs share the same consensus sequence.

6. Regarding CDR H3 loop, find the INCORRECT answer:

- a) Cluster H3-anchor-1 covers about two-thirds of the H3 loop structures.
- b) CDR H3 loop is known to have a substantial impact on the antigen binding.
- c) According to North et al, H3 can be split in an anchor region and a “head” or apex region.
- d) CDR H3 loop structures are very easy to predict.
- e) The main reason of its variability is the full loop structure.

7. Regarding the Immunoglobulin-like beta-sandwich fold:

- a) It always consists of sandwich which contains 7 strands in 2 sheets: greek-key.
- b) Some members of the fold have additional strands.**
- c) It always consists of sandwich which contains 9 strands in 2 sheets: greek-key.
- d) Only V-set domains are considered to have immunoglobulin-like beta-sandwich fold.
- e) This fold is exclusive of Deuterostomes.

8. Which of the following statements about immunoglobulins are FALSE:

- 1. Are all beta proteins, which means that the structure is composed mainly of beta-sheets, although the stabilization of the structure is possible thanks to other structures as alpha-helices.
- 2. The stabilization of beta strands forming one sheet (greek-key) is obtained thanks to the hydrogen bonds between them.
- 3. CH3 has only hydrophobic residues in the surface.
- 4. Disulfide bonds distribution may change according to the Immunoglobulin type and subtype.

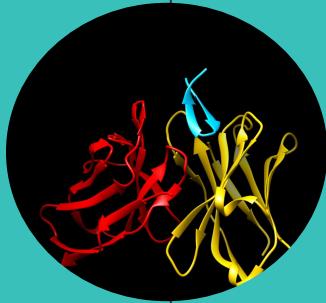
- a) 1, 2, 3.
- b) 1, 3.**
- c) 2, 4.
- d) 4.
- e) 1, 2, 3, 4.

9. Identify the FALSE sentence about canonical structures:

- a) Refer to similar chain conformations.
- b) Can be found in CDRs.
- c) Present a 100% conserved sequence.
- d) Have consensus sequences in most of the cases.
- e) Are involved in the binding of antigens

10. Which is the most variable CDR?

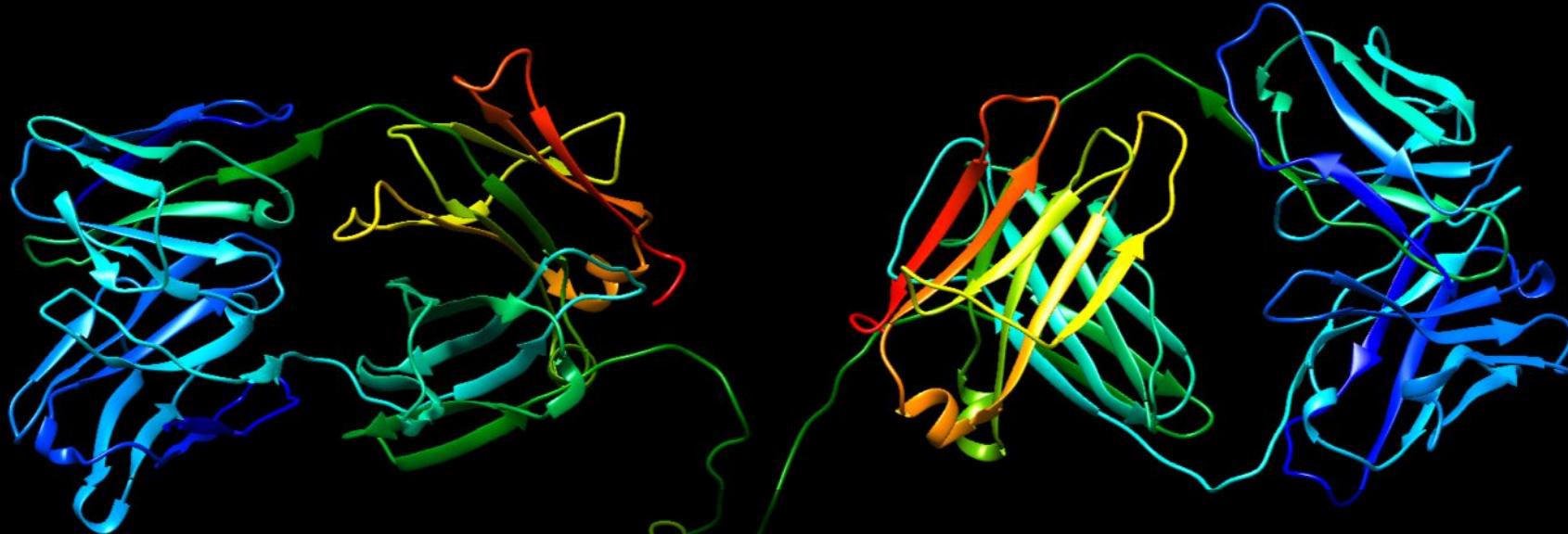
- a) L1
- b) L3
- c) K2
- d) H1
- e) H3



References

- Al-Lazikani B, Lesk A, Chothia C. Standard conformations for the canonical structures of immunoglobulins. *J Mol Biol* 1997;273(4):927-948.
- Benham C, Saleet JM. Disulfide bonding patterns and protein topologies. *Protein Sci* 2008;2(1):41-54.
- Berg JM, Tymoczko JL, Stryer L. Biochemistry. 5th edition. New York: W H Freeman; 2002. Section 33.2, The Immunoglobulin Fold Consists of a Beta-Sandwich Framework with Hypervariable Loops. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK22461/>
- Buljan, M, Bateman, A.. The evolution of protein domain families. *Biochem Soc Trans* 2009; 37(4): 751–755.
- Davies D, Padlan E, Segal D. Three-Dimensional Structure of Immunoglobulins. *Annu Rev Biochem*. 1975;44(1):639-667.
- Fernández-Quintero M, Loeffler J, Kraml J, Kahler U, Kamenik A, Liedl K. Characterizing the Diversity of the CDR-H3 Loop Conformational Ensembles in Relationship to Antibody Binding Properties. *Front Imm* 2019;9.
- Halaby D, Poupon A, Mornon J. The immunoglobulin fold family: sequence analysis and 3D structure comparisons. *Protein Eng Des Sel* 1999;12(7):563-571.
- Halaby DM, Mornon JP. The immunoglobulin superfamily: an insight on its tissular, species, and functional diversity. *J Mol Evol* 1998;46(4):389-400.
- Harris L, Larson S, Hasel K, McPherson A. Refined Structure of an intact IgG2a Monoclonal Antibody. *Biochem* 1997;36(7):1581-1597.
- Klein J, Nikolaidis N. The descent of the antibody-based immune system by gradual evolution. *Proc Natl Acad Sci USA* 2004;102(1):169-174.
- Kong L, Giang E, Nieusma T, Robbins JB, Deller MC, Stanfield RL, Wilson IA, Law M. Structure of hepatitis C virus envelope glycoprotein E2 antigenic site 412 to 423 in complex with antibody AP33. *J Virol* 2012; 86(23):13085-13088.
- Letunic I. SMART: IG domain annotation [Internet]. Smart.embl.de. 2019 [cited 10 February 2019]. Available from: http://smart.embl.de/smart/do_annotation.pl?DOMAIN=SM00409

- Liu H, May K. Disulfide bond structures of IgG molecules. *mAbs* 2012;4(1):17-23.
- Mix E, Goertsches R, Zettl U. Immunoglobulins: Basic Considerations. *ChemInform* 2007;38(23).
- Nezlin R. The immunoglobulins. San Diego: Academic Press; 1998.
- North B, Lehmann A, Dunbrack R. A New Clustering of Antibody CDR Loop Conformations. *J Mol Biol* 2011;406(2):228-256.
- Paul W. Fundamental Immunology. Philadelphia: Wolters Kluwer; 2013.
- Regep C, Georges G, Shi J, Popovic B, Deane C. The H3 loop of antibodies shows unique structural characteristics. *Proteins: Struct Funct Bioin* 2017;85(7):1311-1318.
- Schroeder H, Cavacini L. Structure and function of immunoglobulins. *J Allergy Clin Immunol* 2010;125(2):S41-S52.
- SCOP: Superfamily: Immunoglobulin [Internet]. Scop.mrc-lmb.cam.ac.uk. 2019 [cited 26 February 2019]. Available from: <http://scop.mrc-lmb.cam.ac.uk/scop/data/scop.b.c.b.b.html>
- Steward A, Adhya S, Clarke J. Sequence Conservation in Ig-like Domains: The Role of Highly Conserved Proline Residues in the Fibronectin Type III Superfamily. *J Mol Biol* 2002; 318(4): 935-940.
- Vargas-Madrazo E, Lara-Ochoa F, Carlos Almagro JC. Canonical Structure Repertoire of the Antigen-binding Site of Immunoglobulins Suggests Strong Geometrical Restrictions Associated to the Mechanism of Immune Recognition. *J Mol Biol* 1995;254(3):497-504.
- Vidarsson G, Dekkers G, Rispens T. IgG Subclasses and Allotypes: From Structure to Effector Functions. *Front Immunol* 2014;5.
- Wang N, Smith W, Miller B, Aivazian D, Lugovskoy A, Reff M et al. Conserved amino acid networks involved in antibody variable domain interactions. *Proteins* 2009;76(1):99-114.



• Thank you for your
attention

