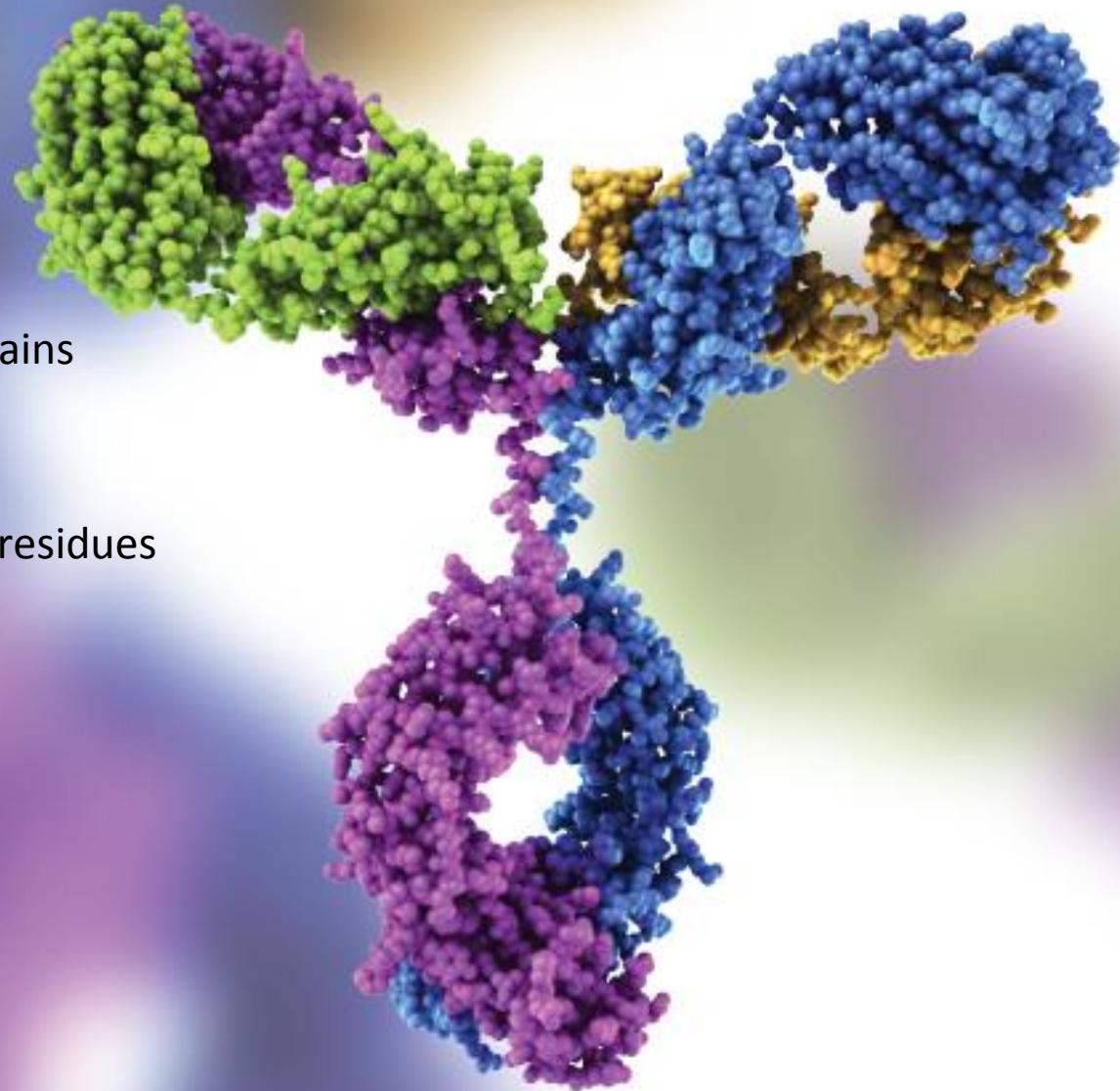


# IMMUNOGLOBULINS

Paula Galiana, Núria Isern, Andrea Martínez, Clara Ribas.

1. Introduction
  - immunoglobulin description
  - Ig constant and variable domains
2. Structural alignment
3. Sequence alignment
  - representation of conserved residues
4. Canonical structures
5. Ag-Ab interaction
6. Conclusions
7. References
8. Multiple choice questions



# Immunoglobulin description

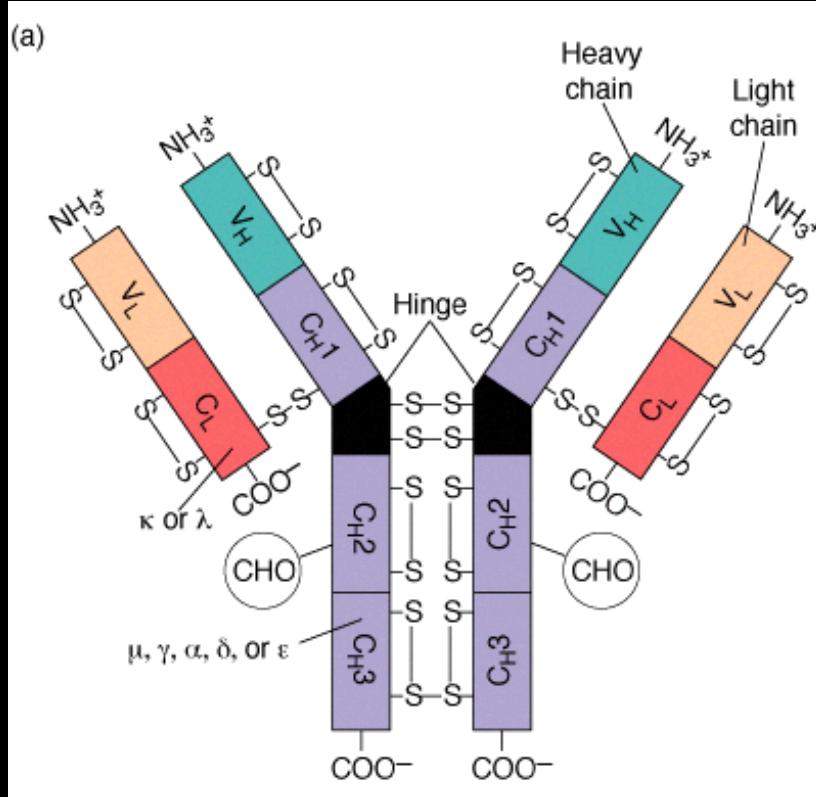
## SCOP Classification:

Class: All  $\beta$  protein

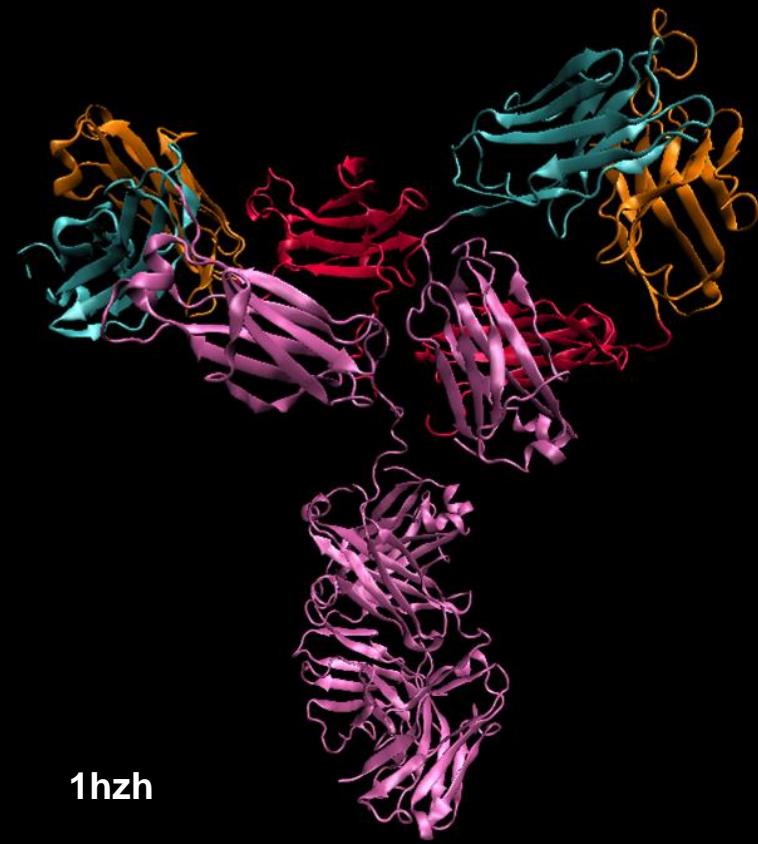
Fold: immunoglobulin-like  $\beta$ -sandwich

Superfamily: Immunoglobulin

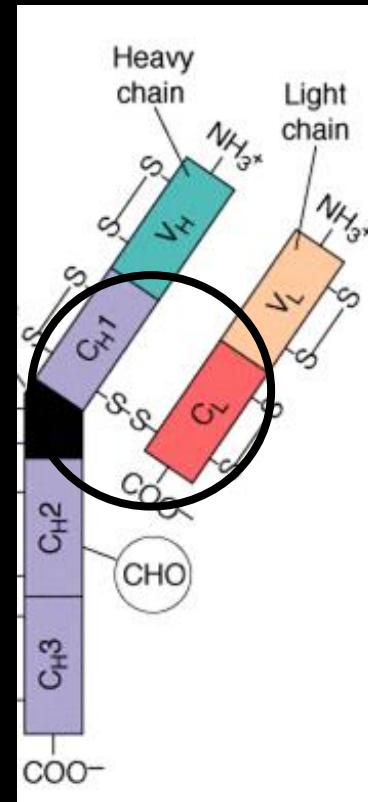
Family: Vset domains or C1 set domains



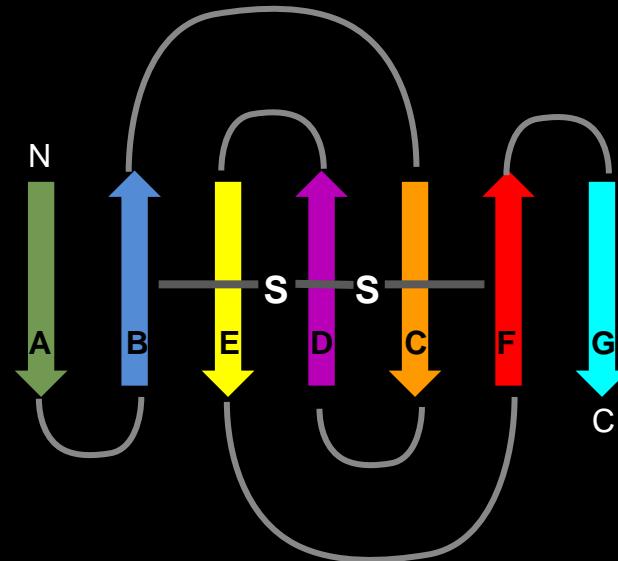
Fab



# Immunoglobulin description



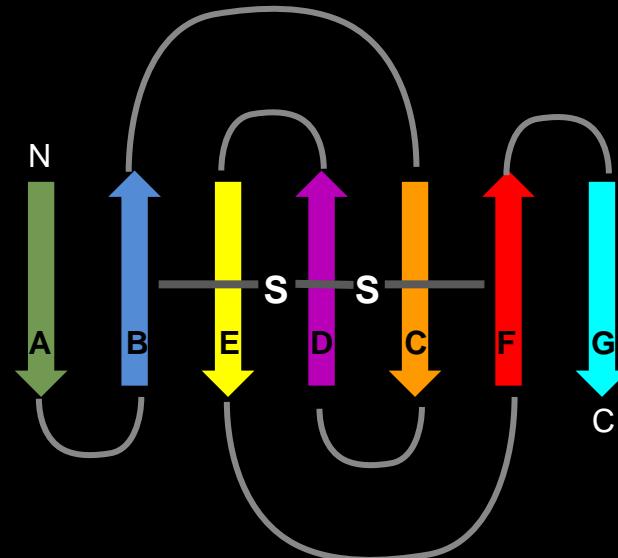
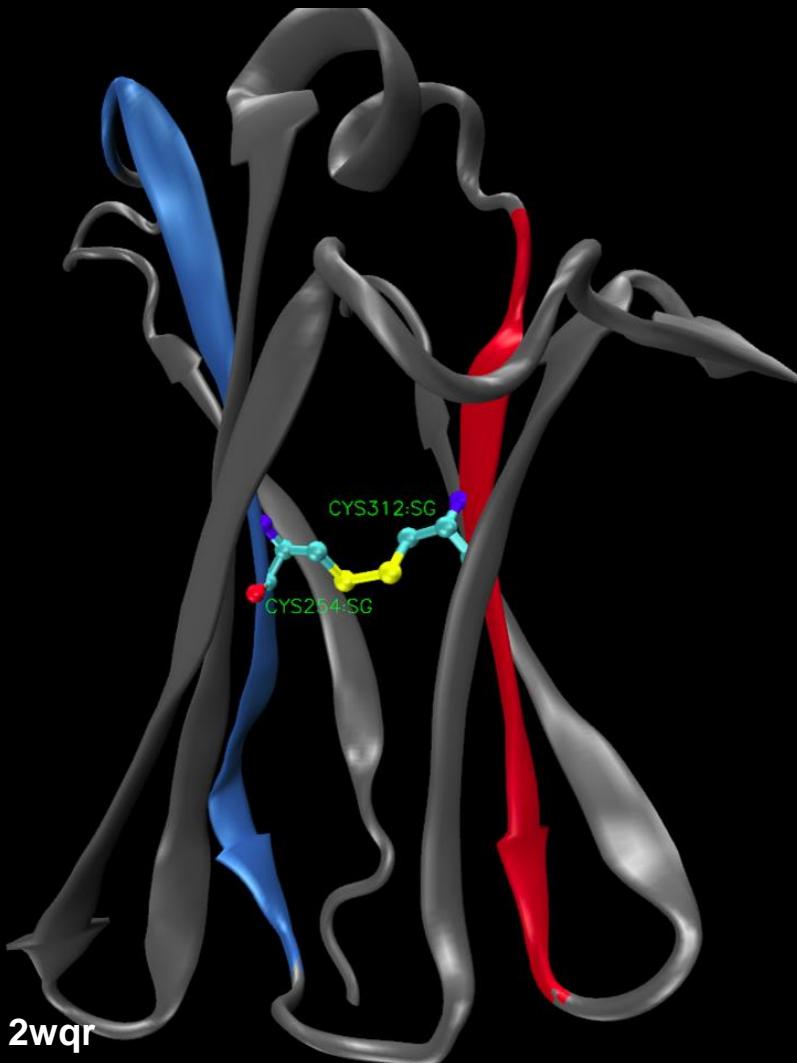
# Ig constant domain



- 7 antiparallel  $\beta$ -strands
- 2  $\beta$ -sheets:
  - 4-stranded sheet
  - 3-stranded sheet

Immunoglobulin-like  $\beta$ -sandwich

# Ig constant domain

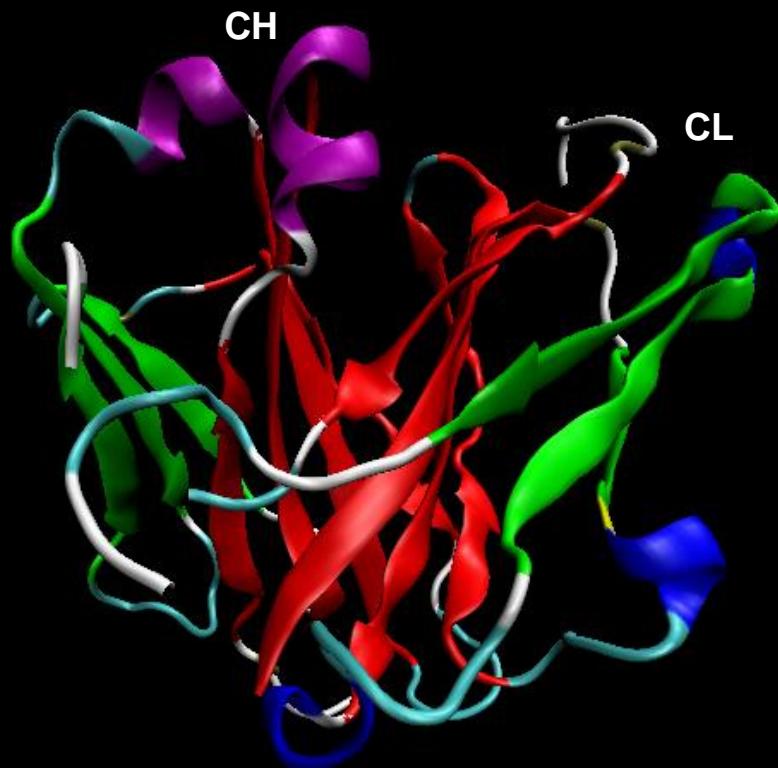


- 7 antiparallel  $\beta$ -strands
- 2  $\beta$ -sheets:
  - 4-stranded sheet
  - 3-stranded sheet

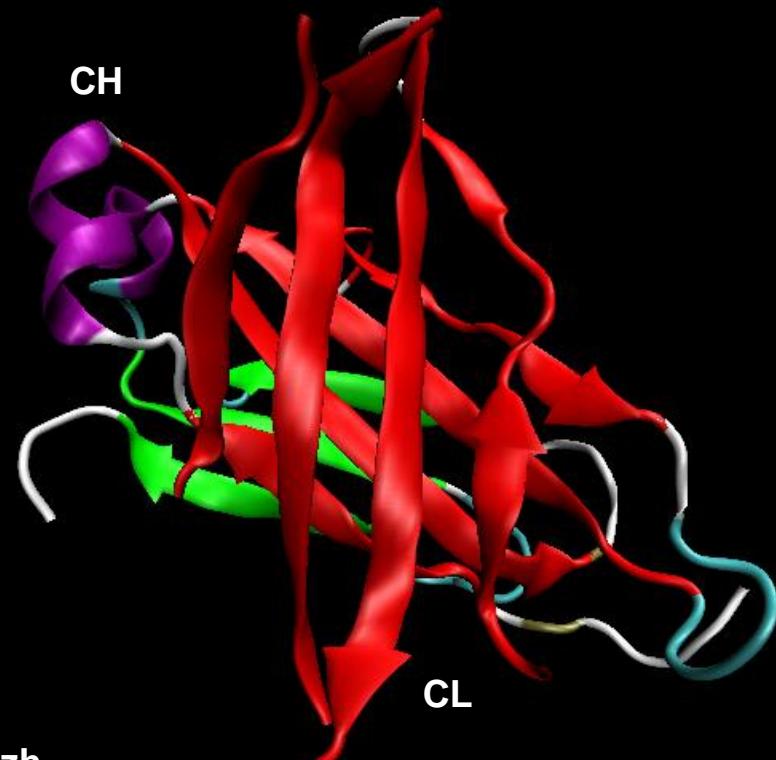
Immunoglobulin-like  $\beta$ -sandwich

# Ig constant domain

4-stranded  $\beta$ -sheets interaction

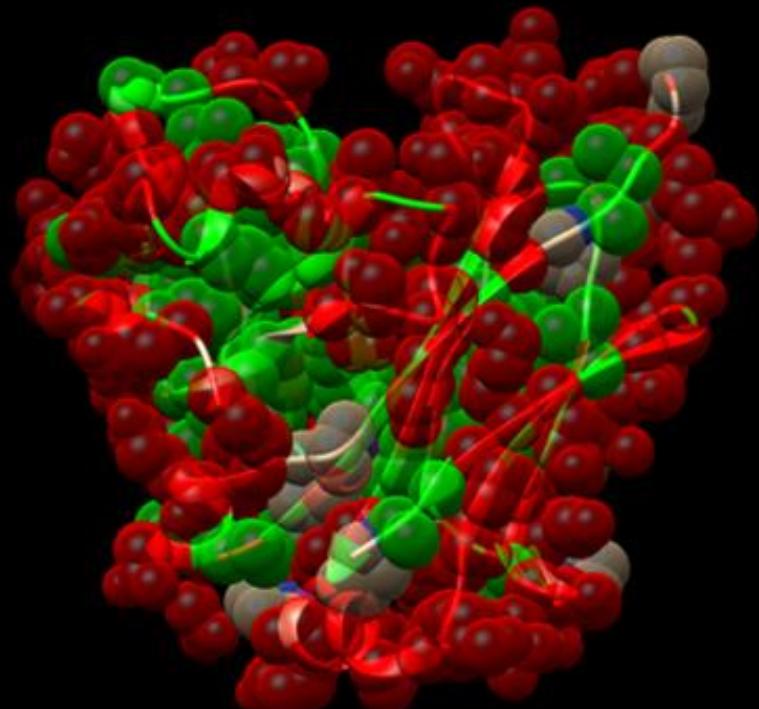


1hzh



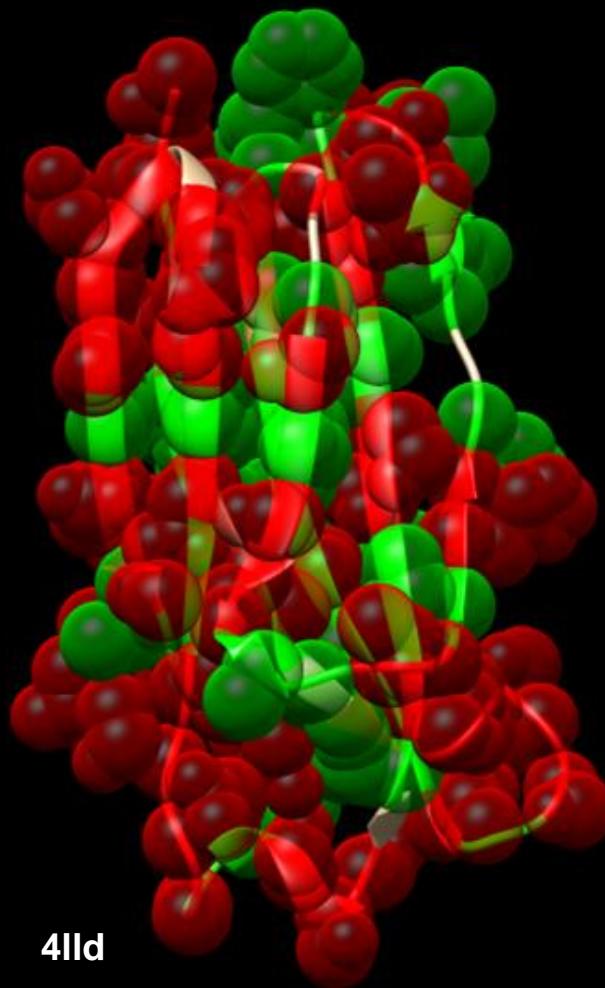
1hzh

# Hydrophobic core



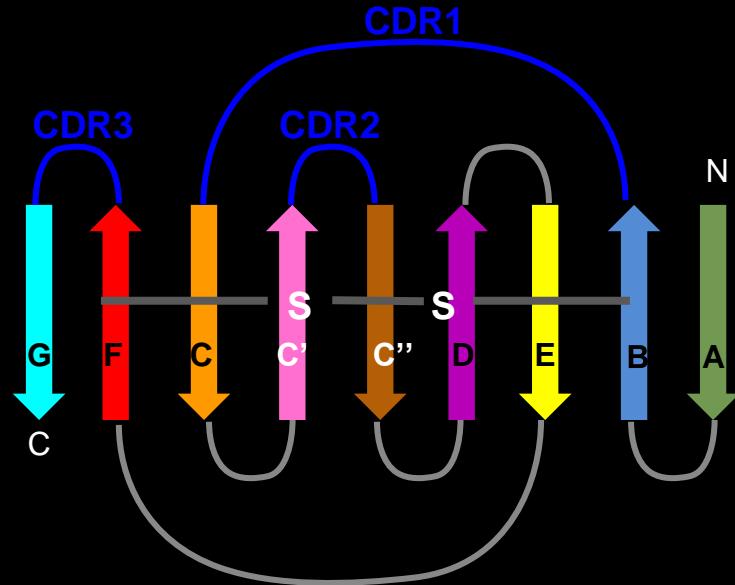
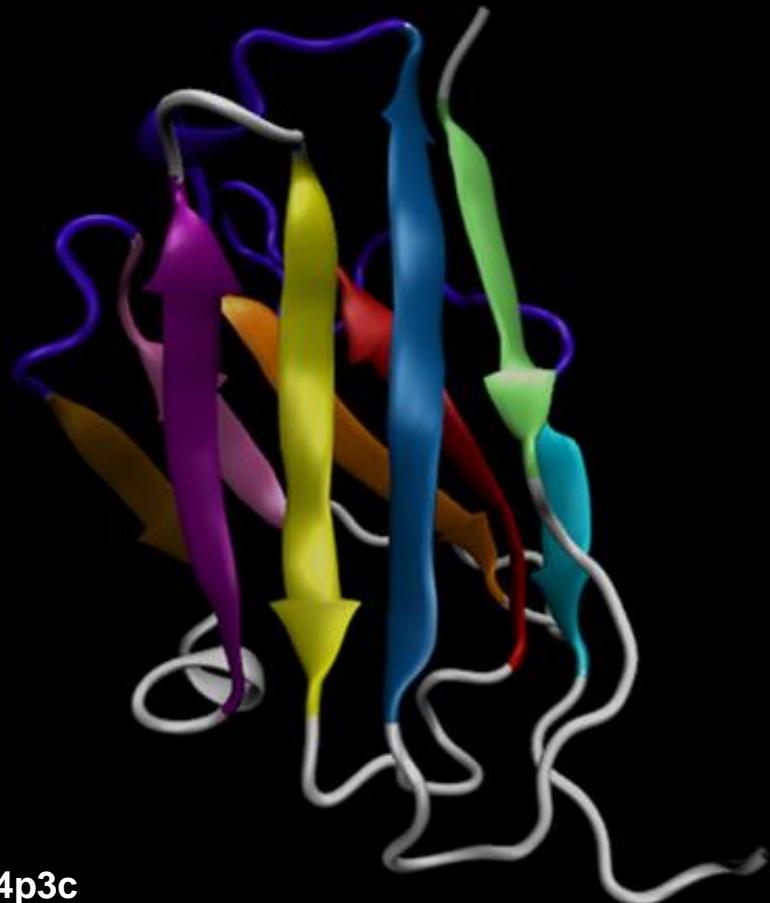
2wqr

- polar
- hydrophobic



4lld

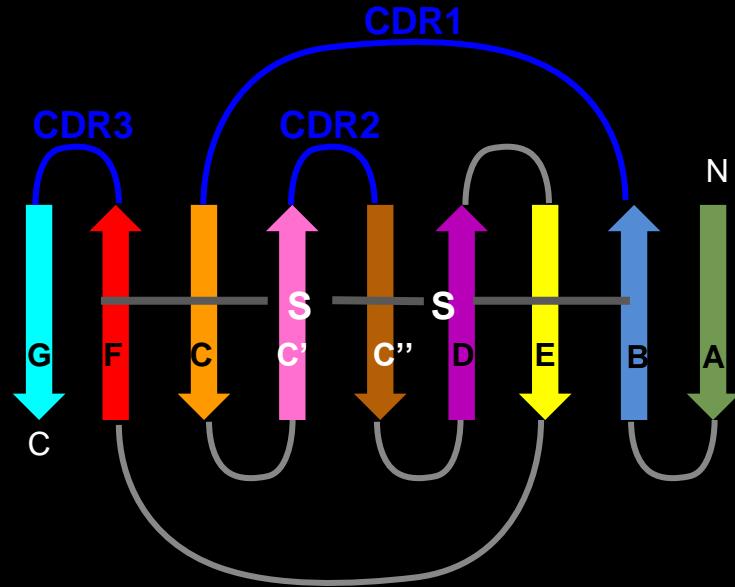
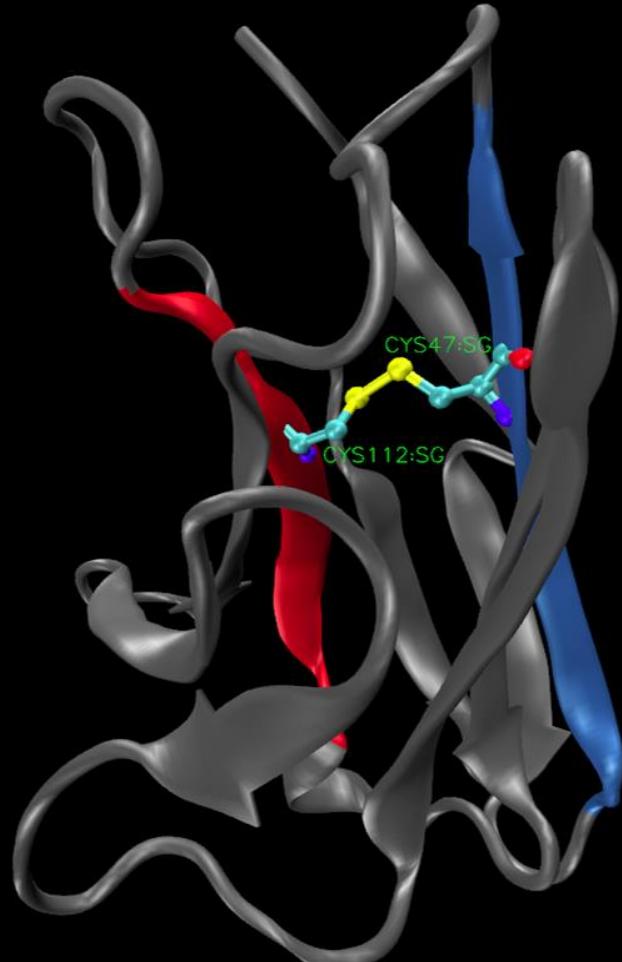
# Ig variable domain



- 7 + 2 antiparallel  $\beta$ -strands: C' and C''
- 2  $\beta$ -sheets:
  - 5-stranded sheet
  - 4-stranded sheet

Immunoglobulin-like  $\beta$ -sandwich

# Ig variable domain



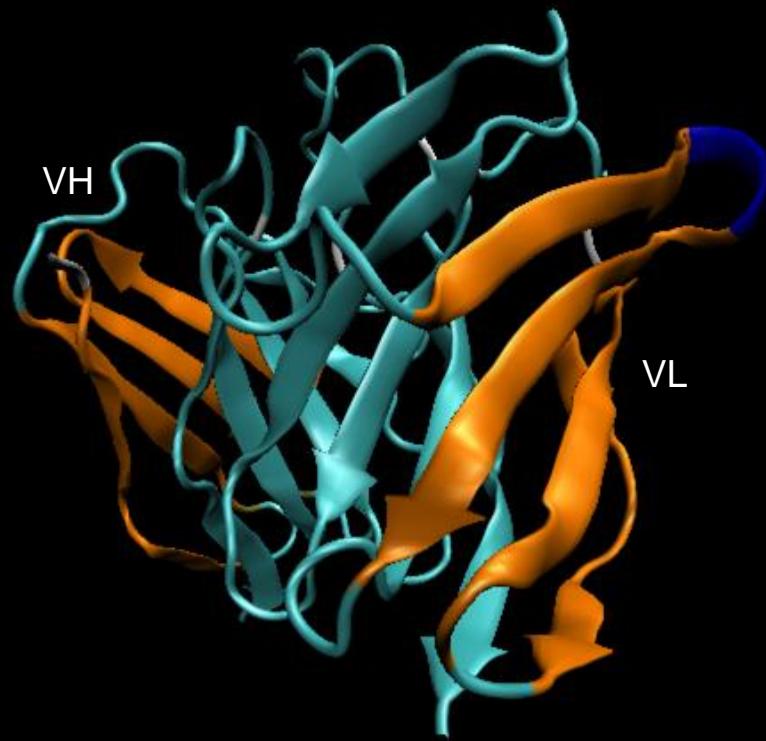
- 7 + 2 antiparallel  $\beta$ -strands: C' and C''
- 2  $\beta$ -sheets:
  - 5-stranded sheet
  - 4-stranded sheet

Immunoglobulin-like  $\beta$ -sandwich

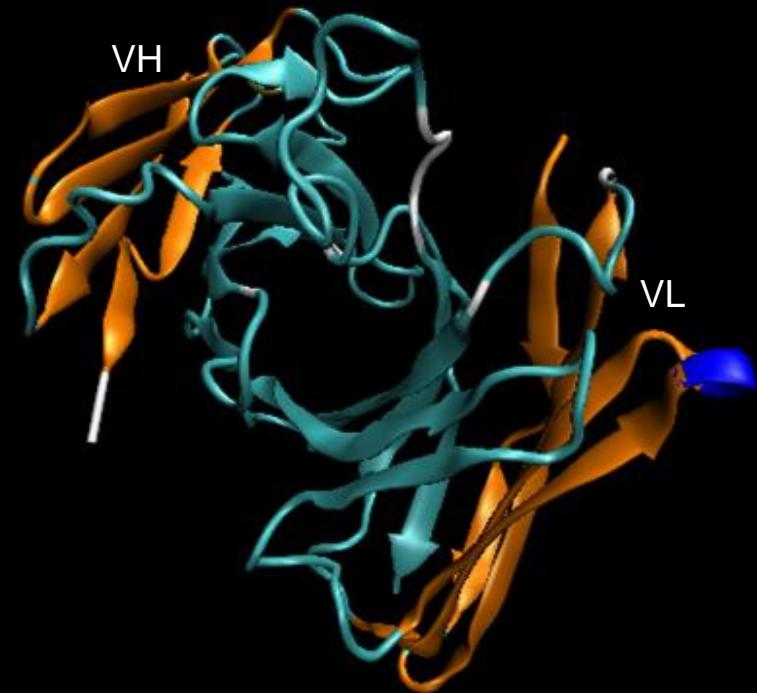
5ejj

# Ig variable domain

5-stranded  $\beta$ -sheets  $\rightarrow$   $\beta$ -barrel

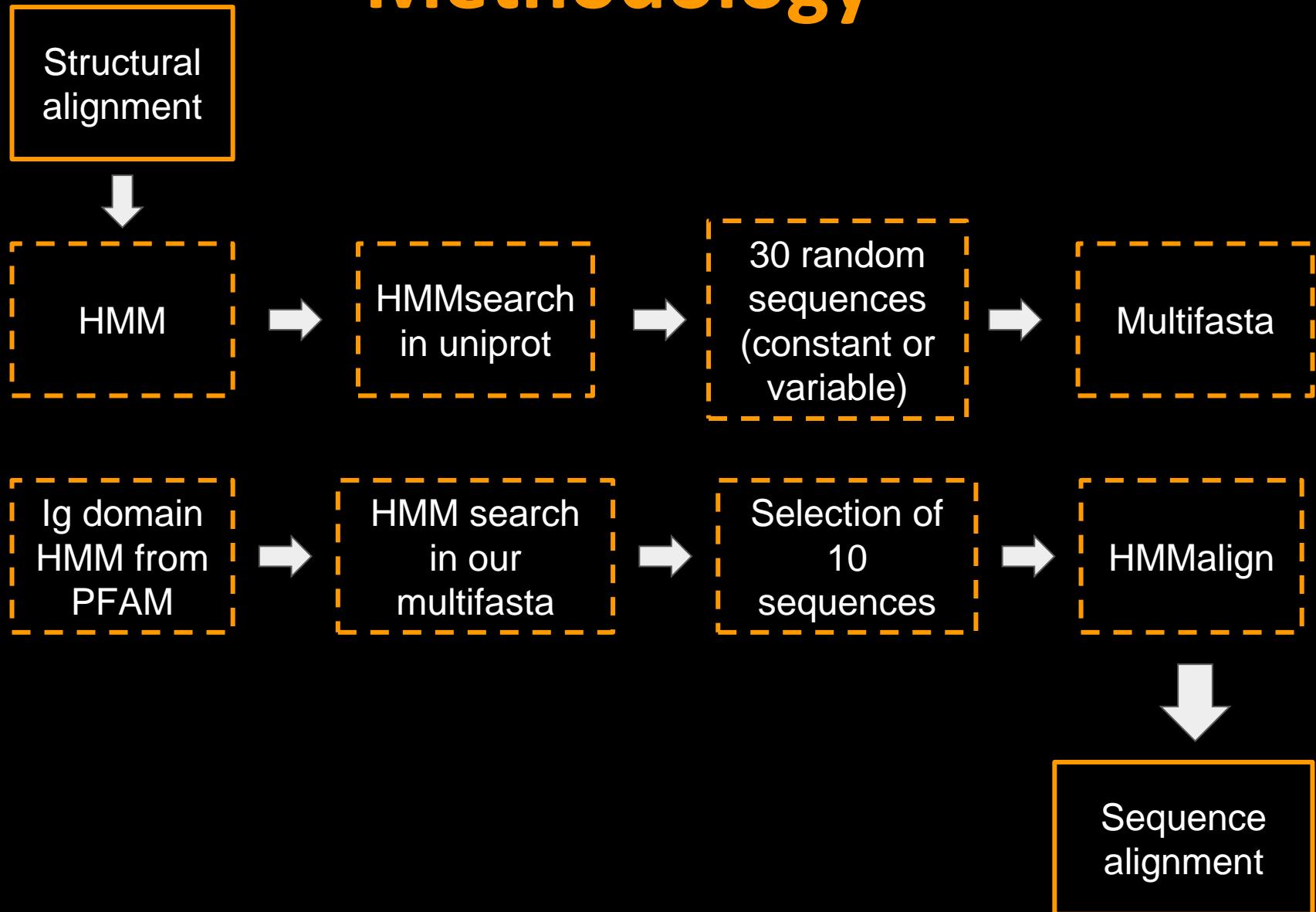


1hzh



1hzh

# Methodology



# Structural alignment - heavy chain

	A	B	C	C'
4jo4 Oryctolagus cuniculus	-QSLEESGGDLV	KPGASLT	L TCTAS	GFS - FTNNY
5ewi Homo sapiens	EVQLVESGGKL	VQPGGS	LRLSCEAS	GES - VGD -
5adp Homo sapiens	-VQLQESGPGLV	KPSET	LSLTCAVS	GYS - ISSGY
1zan Rattus norvegicus	QVQLKESGPGLV	QPSQT	LSLTCTVS	-GFSLTN - N
5dd3 Macaca mulatta	QVQLQESGPGLV	KPSET	LSVTCAVS	-GVSFSS - F
	<b>C''</b>	<b>D</b>	<b>E</b>	<b>CDR 1</b>
4jo4 Oryctolagus cuniculus	VFYATWAKGRFT	ISKTSS - T	TVTLQMTSLTAADT	ATYFCAREN - - FD - AV
5ewi Homo sapiens	TYVDAVKGRFT	VS RD	KGRN SYVLQM	KTLTVGDT
5adp Homo sapiens	TYVNP SLKSRV	TISVDT	TSKN QFSLKL	SSVTAADT
1zan Rattus norvegicus	TDYNSALKSR	L TITRD	TSKSQVFLKM	HSLQSEDT
5dd3 Macaca mulatta	GEYNPSLKSRT	TTISR - S	QISLEI	TSVTAADT
	<b>F</b>	<b>CDR 2</b>		
4jo4 Oryctolagus cuniculus	-----	Y - YF - DLWGPGLTV	IVSSGQPKAP	SVFPLAPCCGDT - PSA - -
5ewi Homo sapiens	Y -	YWGL - G	VWGRGTTTV	SSASTKGPSVFLAPS
5adp Homo sapiens	-----	TA	WNQGQTLTT	SSASTKGPSVFLAPS
1zan Rattus norvegicus	TL	----- Y - A	DAWGQGTTTV	SSASTTAPSVYPLAPG
5dd3 Macaca mulatta	-OPNGFSFTLTSY	-WF	DVWGPGVPTV	SSASTKGPSVFLA - P
	<b>G</b>	<b>A</b>	<b>CDR 3</b>	<b>B</b>
4jo4 Oryctolagus cuniculus	LVKGY	LPEPV	TVTWN	NSGTLTNG
5ewi Homo sapiens	LVKDY	FPEPV	TVSW	NSGALTSG
5adp Homo sapiens	LVKDY	FPEPV	TVSW	NSGALTSG
1zan Rattus norvegicus	LVKGY	FPEPV	TVTWN	NSGSLASG
5dd3 Macaca mulatta	LVKDY	FPEPV	TVSW	NSGSLTSG
	<b>CDR 3</b>	<b>C</b>	<b>D</b>	<b>E</b>
4jo4 Oryctolagus cuniculus	HPATN	TKVDKTV	-AP	---
5ewi Homo sapiens	HKPSN	TKVDKRV	EPK	---
5adp Homo sapiens	HKPSN	TKVDKKV	EPKSCL	
1zan Rattus norvegicus	HPASS	TKVDKKI	-VP - R -	
5dd3 Macaca mulatta	HKPSN	TKVDKRV	-EI - K -	
	<b>F</b>			

# Structural alignment - light chain

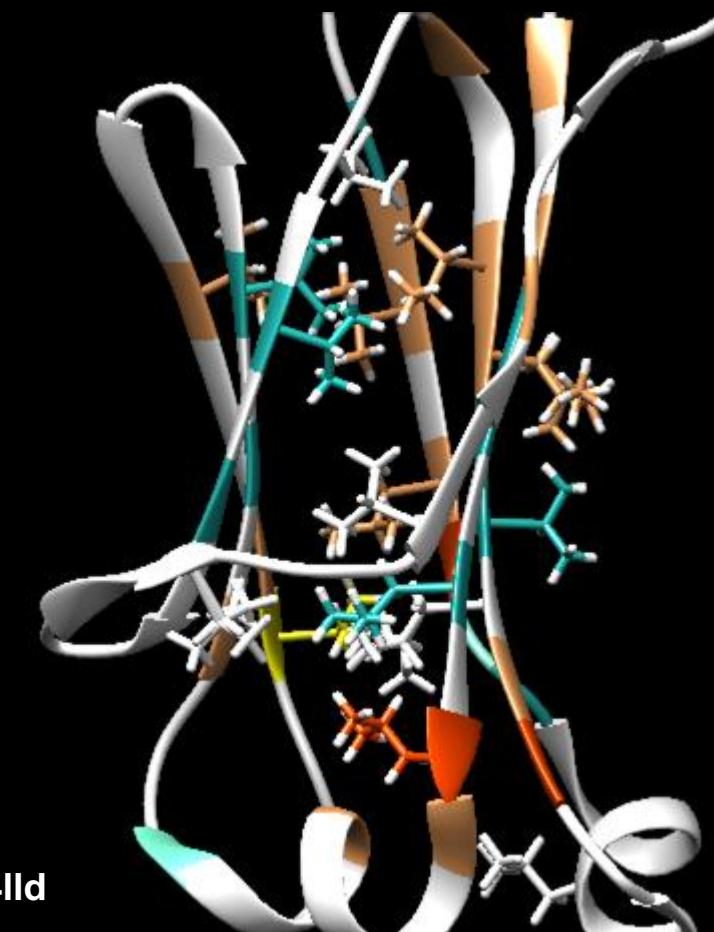
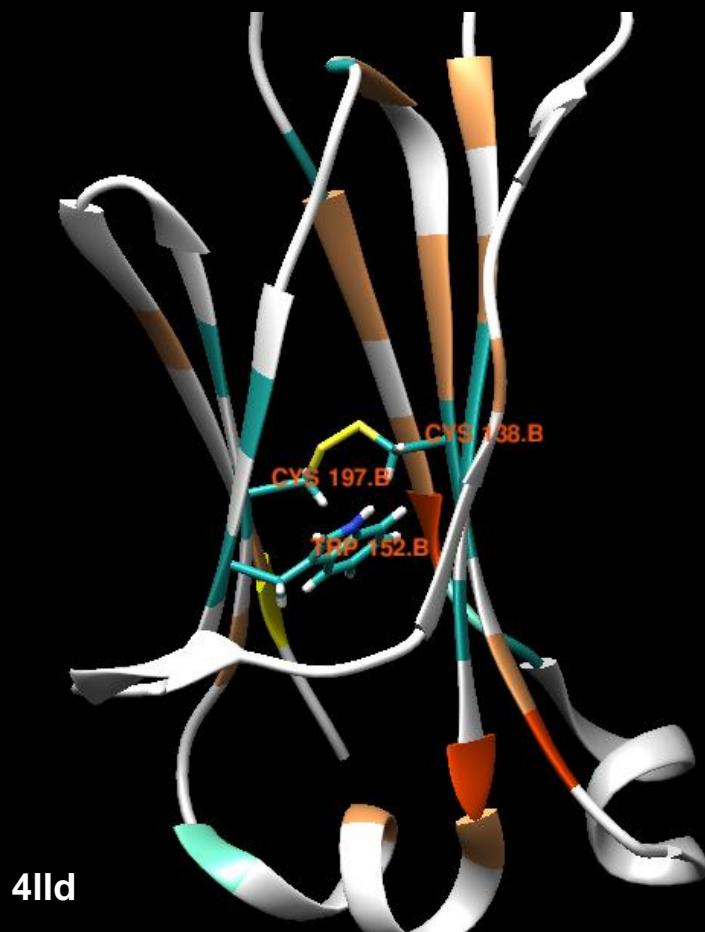
	A	B	C	C'		
1etz Mus Musculus	FAVVTQ--E-SALT	TSPGETVTLTCRS	STGAVTTSN-----YA	IWVQEKPDLFSGLIGG		
5dd3 Macaca mulatta	DIQVTQ-SP-SSL	SASVGDTVTISCRTS	QT-IS-----T-W	LAWYQVKPGKAPKLLIYT		
4jo4 Oryctolagus cuniculus	DIVMTQ-TPASVSAAV	-GGTVTINCQA	--S-ET--IS-N-Y	LAWYQQKPGQPPKLLIYK		
5i15 Homo sapiens	DIQMTQSPS-SLSASV	-GDRVATICRA	-----SQ---SISSY	LNWYQQKPGKAPKLLIYA		
5ihu Bos taurus	--VLNQ-PS-SVSGSL	-GQRVSITCSG	-----SSSNVNGNG-Y	VSWYQLIPGSAPRTLIG		
	C''	D	E	CDR 1	F	G
1etz Mus Musculus	TNNRVPGVPAR	FSGSLIG-D	KAALT	TVTGAQTEDE	AIYFCALWYS-----N-H	-WVFGGGT
5dd3 Macaca mulatta	ASSESGVPSR	FSGSG-SGT	DFTLT	I	STLQSEDFATYYCQQYIS-----L-P	-PTFGLGT
4jo4 Oryctolagus cuniculus	ASTASGVSSR	FKGSG-SGT	EYTL	TVTGVQCDDA	ATYYCQQGYS---ISDIDNSFGG-GT	
5i15 Homo sapiens	ASLQSGVPSR	FSGSG-SGT	DFTLT	SSLQPEDFATYYCQQSYS---T-PL	--TFGQ-GT	
5ihu Bos taurus	DT	ASGVVPDR	FSGSR-SGN	TATLTSSLQAEDEADYFCASA-EDSSS-NA--VFGS-GT		
	CDR 2	A	B	CDR 3	C	D
1etz Mus Musculus	KLTVLG-QPKSSPSVT	LTPSSEELETN	KATLVCTITDF	YPGVVTVDWKVDGTPV	TQGME	
5dd3 Macaca mulatta	KVEIK--RAVAAPS	SVFIEPPSEDQVKSG	TVSVVCLNNF	YPREASVWKVVDGVL	KTGNSQ	
4jo4 Oryctolagus cuniculus	EVVVK--GDPVAP	TVLIEPPAADQVATG	TVTIVCVANKYFPD-V	TVTWEVDGTTQITGIE		
5i15 Homo sapiens	KVEIK--RTVAAPS	VFIEPPSDEQLKSG	TASVVCLNNF	YPREAKVQWKVDNALQSGNSQ		
5ihu Bos taurus	TLT	VLG-QPKSSPSVT	ELNGNKATLVCLISDF	YPGSVTVVWKADGSTITRNVE		
	E	F	G			
1etz Mus Musculus	TTQPSKQS-NNK	YMASSYLT	TARAWERHS	SYSCQVTHEGH----TVE	KSLRAECS	
5dd3 Macaca mulatta	ESVTEQDSKDNT	YSLSS	TLTSNTDYQSHN	YVACEVTH-Q-GLSSP	VTKSFNRGE--	
4jo4 Oryctolagus cuniculus	NSKT	PQNSADCTY	NLSSTLTL	TSTQYNSHKEYTCKV	TQ-GT---TSV	VQSFNRGDC-
5i15 Homo sapiens	ESVT	EQDSKDSTY	SLSS	TLTSKADYEKHK	YVACEVTH-QGLSSP	VTKSFNR-GEC
5ihu Bos taurus	TTRASKOS-NSK	YAAASSYLSLT	SSDWKS	KGSYSCEVTH-EG---	STVKT	KPSEC-

# Sequence alignment - CH1

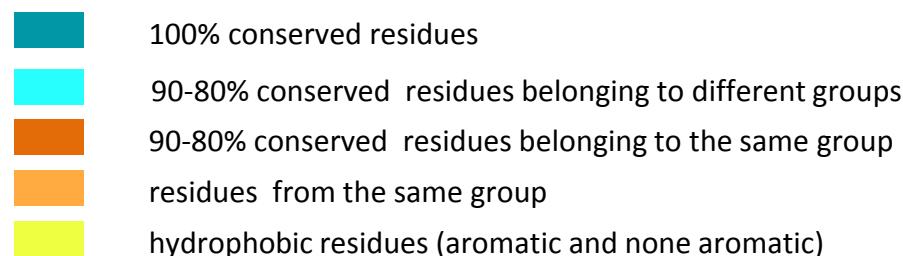
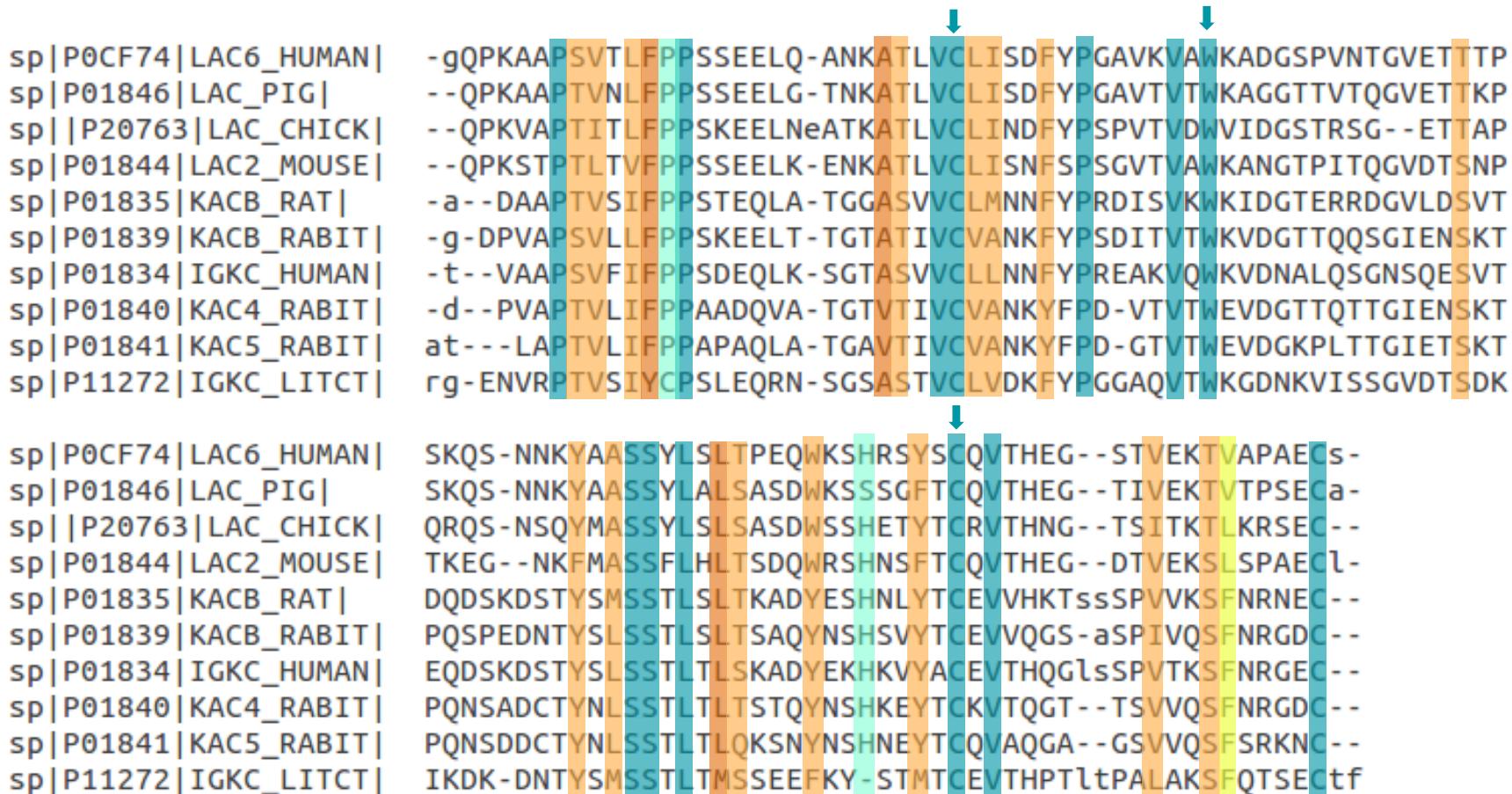


- 100% conserved residues
- 90-80% conserved residues belonging to different groups
- 90-80% conserved residues belonging to the same group
- residues from the same group
- hydrophobic residues (aromatic and none aromatic)

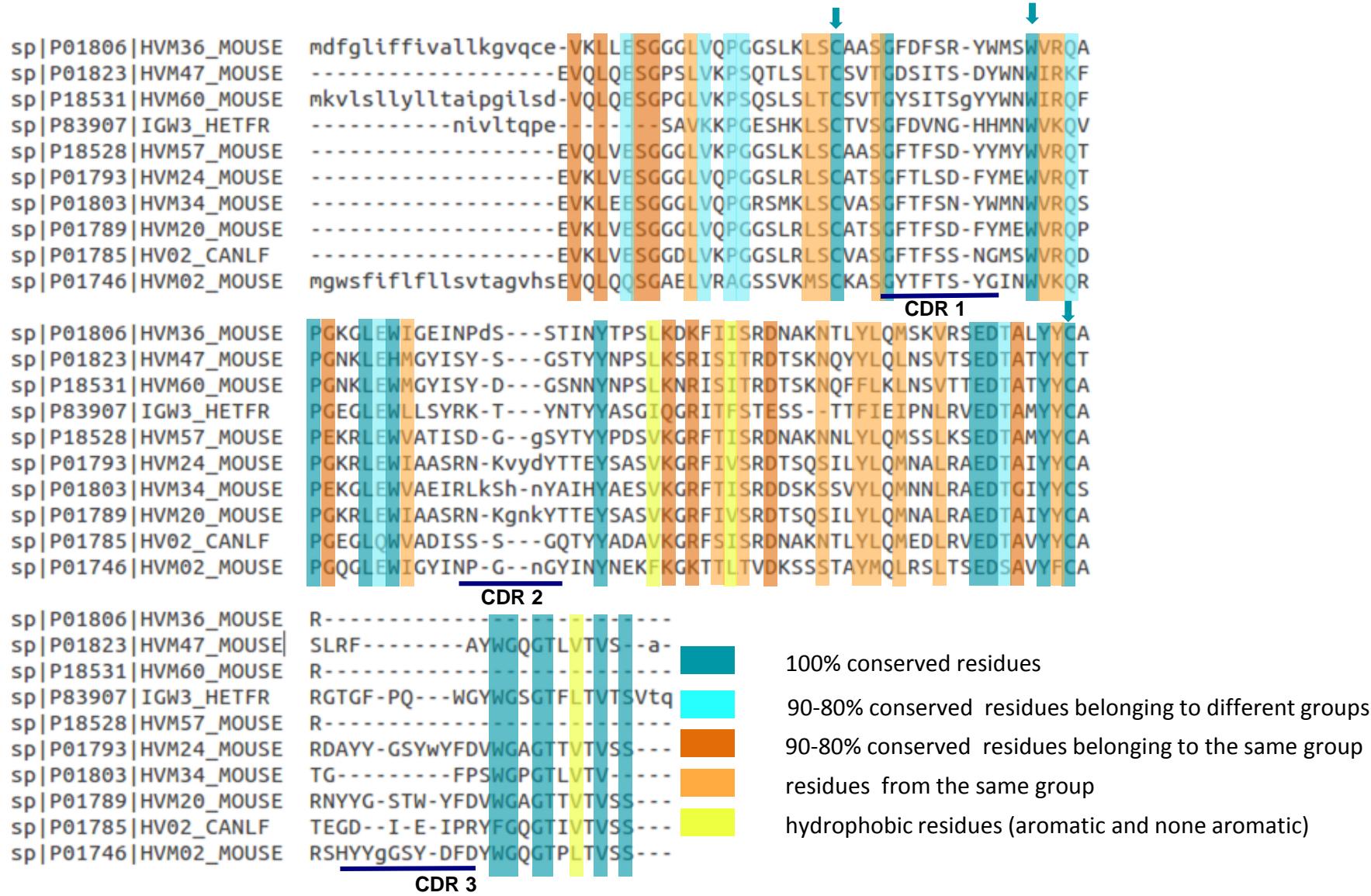
# Sequence alignment constant domain



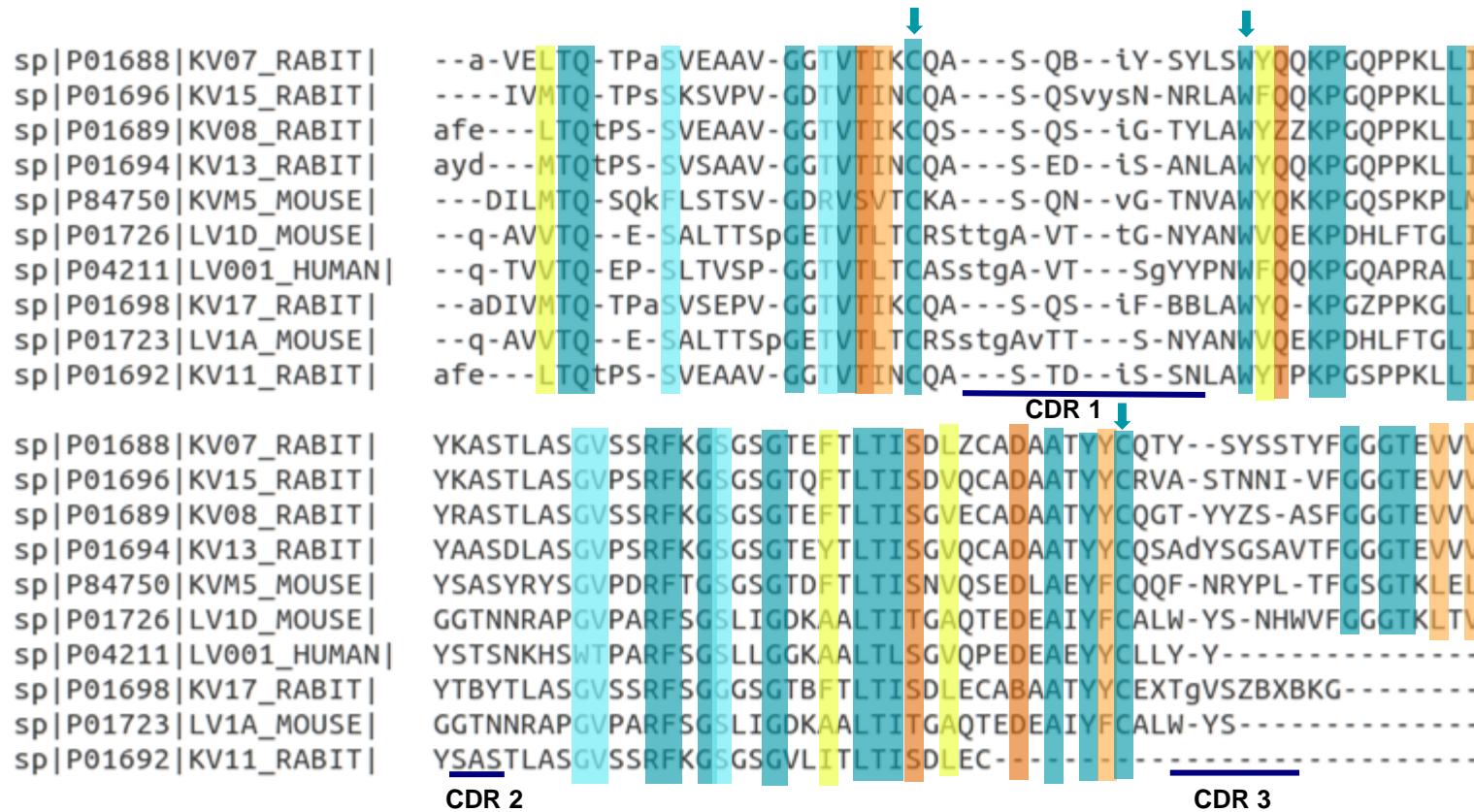
# Sequence alignment - CL



# Sequence alignment - VH



# Sequence alignment - VL



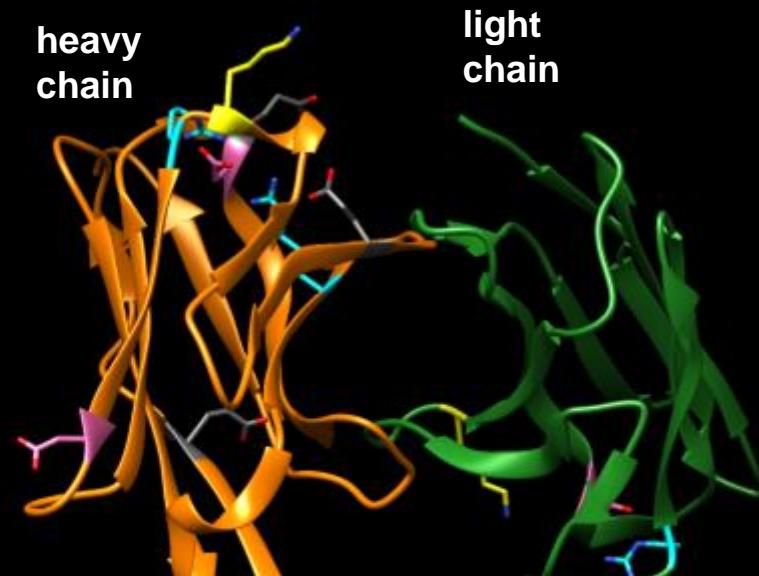
- 100% conserved residues
- 90-80% conserved residues belonging to different groups
- 90-80% conserved residues belonging to the same group
- residues from the same group
- hydrophobic residues (aromatic and none aromatic)

# Additional conservations - charged residues

Alignment from the heavy chain:

```
mdfglifivallkgvqce-VKLE|SGGLVQPGGSLKLSCAASGFDTSR-YWMSWV|RQA
-----EVQLQE|SGPSLVKPSQTLSTCSTVGDTS-DYWNIW|RKF
mkvlslllylltaipgilsd-VQLQE|SGPGLVQPGSLSLTCSTVGDTSgYYWNW|RQF
-----nivltqpe-----SAVKKPGESHKLSCTVGFDVNG-HHMNWV|KQV
-----EVQLVE|SGGLVQPGGSLKLSCAASGFTFSD-YYMYWV|RQT
-----EVKLVE|SGGLVQPGGSLRLSCATSGFTLSD-FYMEWV|RQT
-----EVKLE|SGGLVQPGGRSMKLSCVASGFTFSN-YWMNWV|RQS
-----EVKLVE|SGGLVQPGGSLRLSCATSGFTFSD-FYMEWV|RQP
-----EVKLVE|SGGDLVQPGGSLRLSCVASGFTFSS-NGMSWV|RQD
mgwsfiflfllsvtagvhsEVQLQSGAELVRAGSSVKMSCKASGYFTS-YGINWV|KQR

PGKGLE|WIGEINPdS---STINYTPSL|DKFIISRDNAKNTLYLQMSKVRSEDTALYYCA
PGNKLE|HMGYISY-S---GSTYYNPSL|KSRISITRDTSKNQYYLQLNSVTSEDTATYYCT
PGNKLE|WMGYISY-D---GSNNYNPSL|KNRISITRDTSKNQFFLKLNSVTTEDTATYYCA
PGEGL|EWLLSYRK-T---YNTYYASG|QGRITFSTESS-TTFIEIPNLRVEDTAMYYCA
PEKRL|EWVATISD-G---gSYTYYPD|SVKGRTFTISRDNAKNNLYLQMSSLKSEDTAMYYCA
PGKRL|EWIAASRN-KvydYTTE|YASVKGRTFIVSRDTQSILYLQMNALRAEDTAIYYCA
PEKGLE|WVAEIRLkSh-nYAIHYA|ESVKGRTFTISRDSSKSSVYLQMNNLRAEDTGIYYCS
PGKRL|EWIAASRN-KgnkYTTE|YASVKGRTFIVSRDTQSILYLQMNALRAEDTAIYYCA
PGEGL|QWVADISS-S---GQTYYADAV|KGRTFSISRDNAKNTLYLQMEDLRVEDTAVYYCA
PGQGLE|EWIGYINP-G--nGYINYNEKF|GKTTLTVDKSSSTAYMQLRSLTS|EDSAVYFCA
```



- Arginine (R)
- Aspartic acid (D)
- Glutamic acid (E)
- Lysine (K)

} found in the surface

## Additional conservations - charged residues

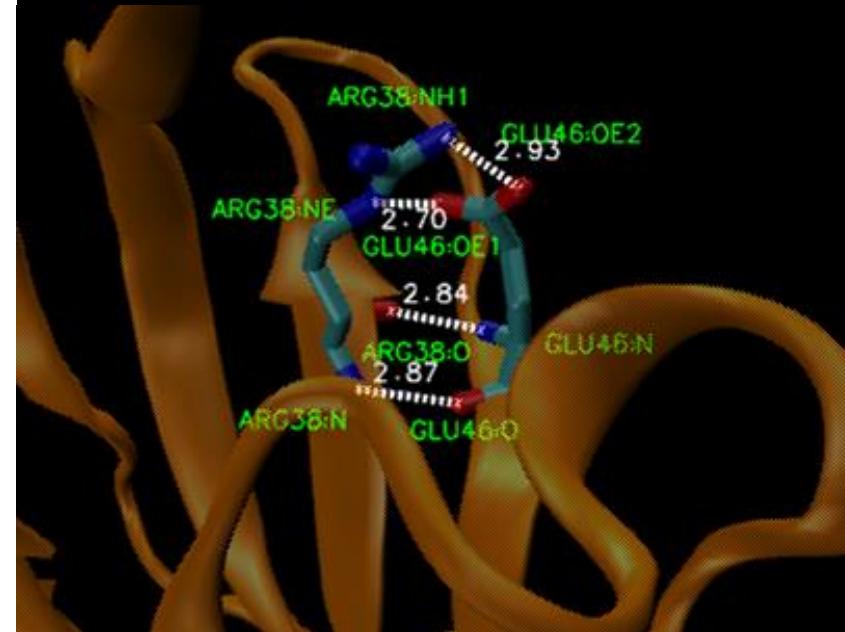
## Salt bridges between charged residues in the heavy chain

Arg 38

mdfgliffivallkgvqce-VKLLESGGGLVQPGGSLKLSCAASGFDTSR-YWMSWVQRQ  
-----EVQLQESGPSLVKPSQTLSTCSVTGDSITS-DYWNWIRKF  
mkvlslllylltaipgilsd-VQLQESGPGLVKPSQSLSTCSVTGYSITSgYYWNWIRQF  
-----nivltqpe-----SAVKKPGESHKLSCTVSGFDVNG-HHMNWKVQV  
-----EVQLVESGGGLVKPGGSLKLSCAASGFTFSD-YYMYWVRQT  
-----EVKLVESGGGLVQPGGSLRLSCATSGFTLSD-FYMEWVRQT  
-----EVKLEESGGGLVQPGRSMKLCVASGFTFSN-YWMNWVRQ  
-----EVKLVESGGGLVQPGGSLRLSCATSGFTFSD-FYMEWVRQ  
-----EVKLVESGGDLVKPGGSLRLSCAVSGFTFSS-NGMSWVRQD  
mgwsfiflfllsvtagvhEVQLQQSGAELVRAGSSVKMSCKASGYTFTS-YGINWVKQR  
PGKGLEWIGEINPds---STINYTPSLKDKFIISRDNAKNTLYLQMSKVRSEDTALYYCA  
PGNKLHEHMGYISY-S---GSTYYNPSLKSRSITRDTSKNQYYLQLNSVTSEDATYYCT  
PGNKLHEWMGYISY-D---GSNNYNPSLKNRISITRDTSKNQFFLKLKNSVTTEDATYYCA  
PGEGLEWLLSYRK-T---YNTYYASGIQGRITFSTESS--TTFIEIPNLRVEDTAMYYCA  
PEKRLHEWVATISD-G--gSYTYYPDVKGRFTISRDNAKNNLYLQMSLKSEDTAMYYCA  
PGKRLHEWIAASRN-KvydYTTEYSASVKGRFIVSRDTSQSILYLQMNALRAEDTAIYYCA  
PEKGLEWVAEIRLkSh-nYAIHYAESVKGRFTISRDDSKSSVYLYLQMNALRAEDTGIYYCS  
PGKRLHEWIAASRN-KgnkYTTEYSASVKGRFIVSRDTSQSILYLQMNALRAEDTAIYYCA  
PGEGLQWVADISS-S---GQTYYADAVVKGRFSISRDNAKNTLYLQMEDLRVEDTAVYYCA  
PGQGLEWIGYINP-G--nGYINYNEFKFKGKTTLTVDKSSSTAYMQLRSLTSEDSAVYFCA

Arg 38

## heavy chain



5mvz

Glu 46

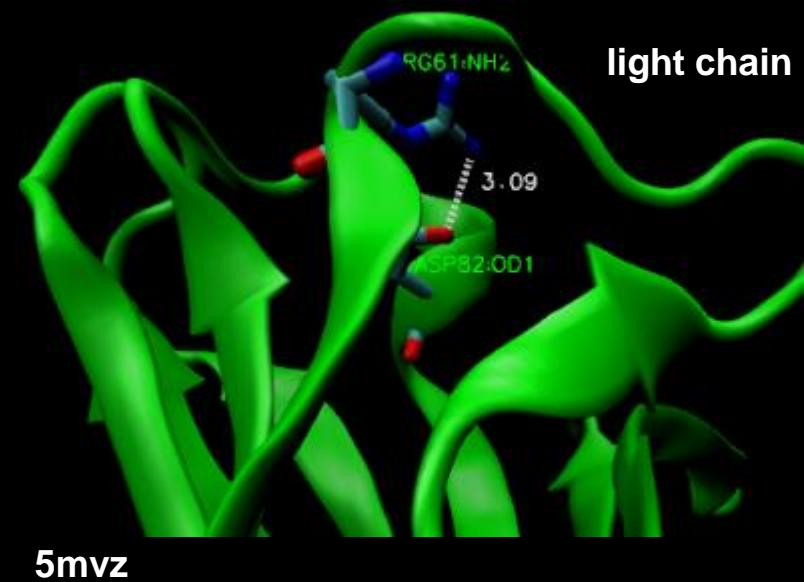
# Additional conservations - charged residues

Salt bridge between charged residues in the light chain

YKASTLASGVSSRFKGSGSGTEFTLTISDLZCADAAATYYCQTY  
YKASTLASGVPSRFKGSGSGTQFTLTISDVQCADAATYYCRVA  
YRASTLASGVSSRFKGSGSGTEFTLTISGVVECADAATYYCQGT  
YAASDLASGVPSRFKGSGSGTEYTLTISGVQCADAATYYCQSA  
YSASRYSGVPDRFTGSGSGTDFTLTISNVQSED LAEYFCQQF  
GGTNNRAPGVPARFSGSLIGDKAALTITGAQTED EAIYFCALW  
YSTSNKHSWTPARFSGSLLGGKAALTLSGVQPED EAEYYCLLY  
YTBTLASGVSSRFSGGGSGTBFTLTISDLECABAATYYCEXT  
GGTNNRAPGVPARFSGSLIGDKAALTITGAQTED EAIYFCALW  
YSASTLASGVSSRFKGSGSGVLITLTISDLEC-----

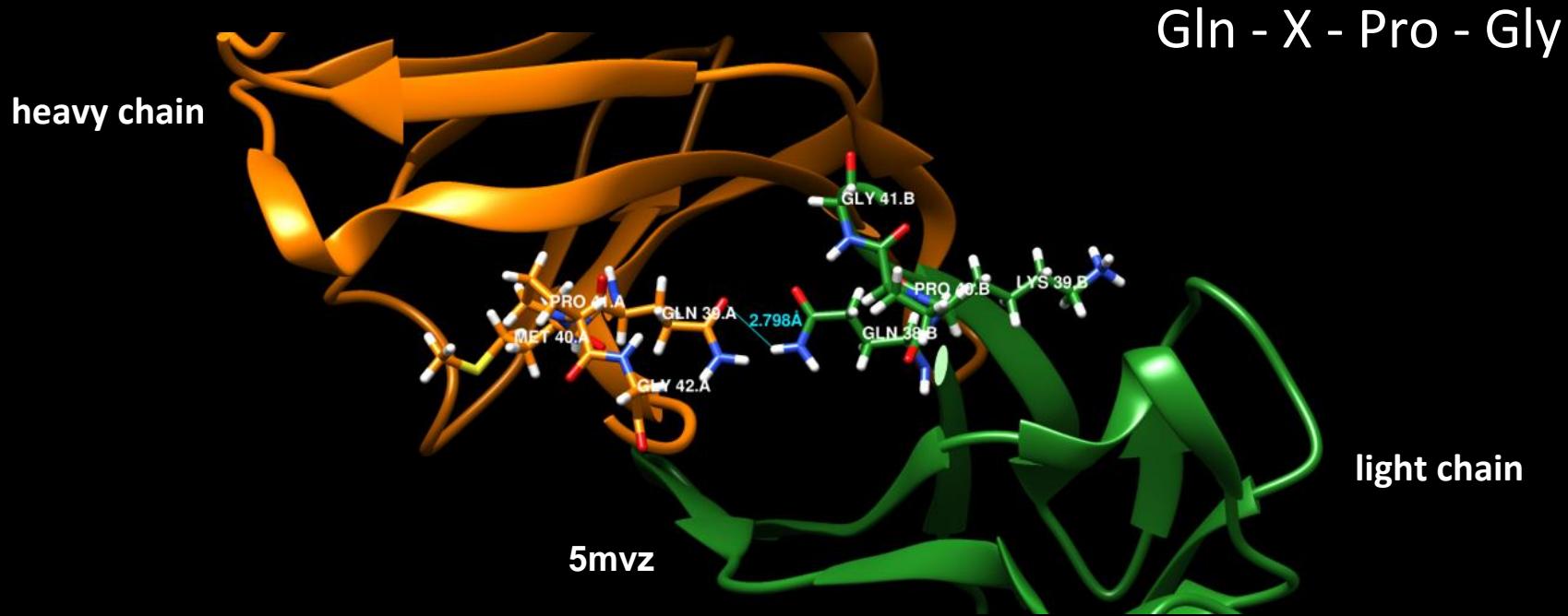
Arg 61

Asp 82

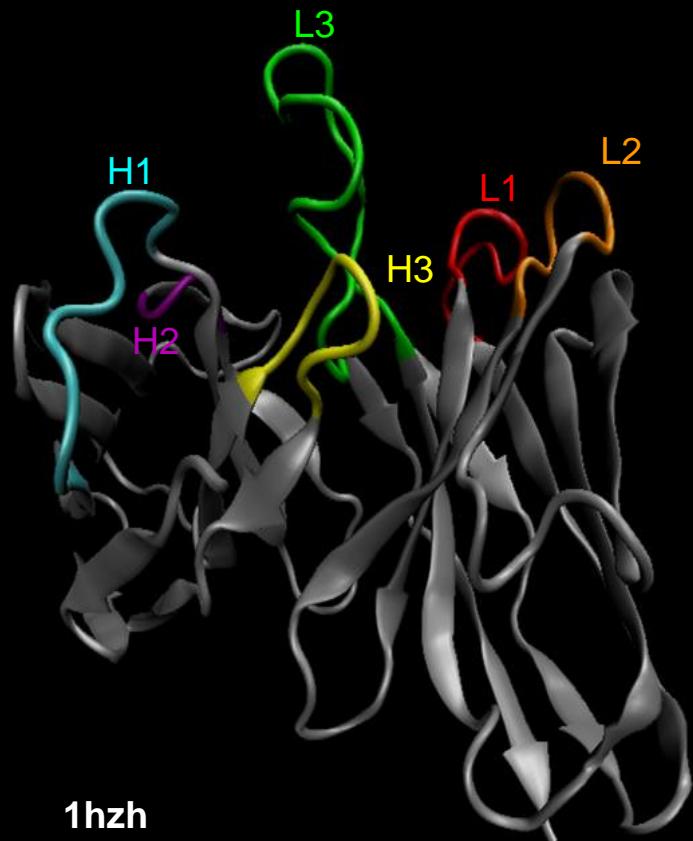


# Additional conservations - polar residues

sp P01806 HVM36_MOUSE	mdfglifflivallkgvqce-VKLLESGGLVQPGGSLKLSCAASGFDTSR-YWMSWVRQAPGKGLEWIGEINPds-
sp P01823 HVM47_MOUSE	-----EVQLQESGPGSLVKPSQTLSTCSVTGDSITS-DYWNWIRKFPGNPKLEHMGYISY-S-
sp P18531 HVM60_MOUSE	mkvlsllylltaipgilsd-VQLQESGPGGLVKPSQSLSTCSVTGYSITSgYYWNWIRQFPGNPKLEHMGYISY-D-
sp P83907 IGW3_HETFR	-----nivltqpe-----SAVKKPGESHKLSCTVSGFDVNG-HHMNWVKQVPGEGLEWLLSYRK-T-
sp P18528 HVM57_MOUSE	-----EVQLVESGGGLVQPGGSLKLSCAASGFTFSD-YYMYWVRQTPEKRLEWVATISD-G-
sp P01793 HVM24_MOUSE	-----EVKLVESGGGLVQPGGSLRLSCATSGFTLSD-FYMEWVRQTPEKRLEWIAASRN-K-
sp P01803 HVM34_MOUSE	-----EVKLEESGGGLVQPGRSMKLSVASGFTFSN-YWMNWVRQSPEKGLEWVAEIRLKS-
sp P01789 HVM20_MOUSE	-----EVKLVESGGGLVQPGGSLRLSCATSGFTFSD-FYMEWVRQPPEKRLEWIAASRN-K-
sp P01785 HV02_CANLF	-----EVKLVESGGDLVQPGGSLRLSCAVSGFTFSS-NGMSWVRQDPGEGLQWVADISS-S-
sp P01746 HVM02_MOUSE	mgwsfiflfls vtagvhsEVQLQQSGAELVRAGSSVKMSCKASGYTFTS-YGINWVKQRPGQGLEWIGYINP-G-



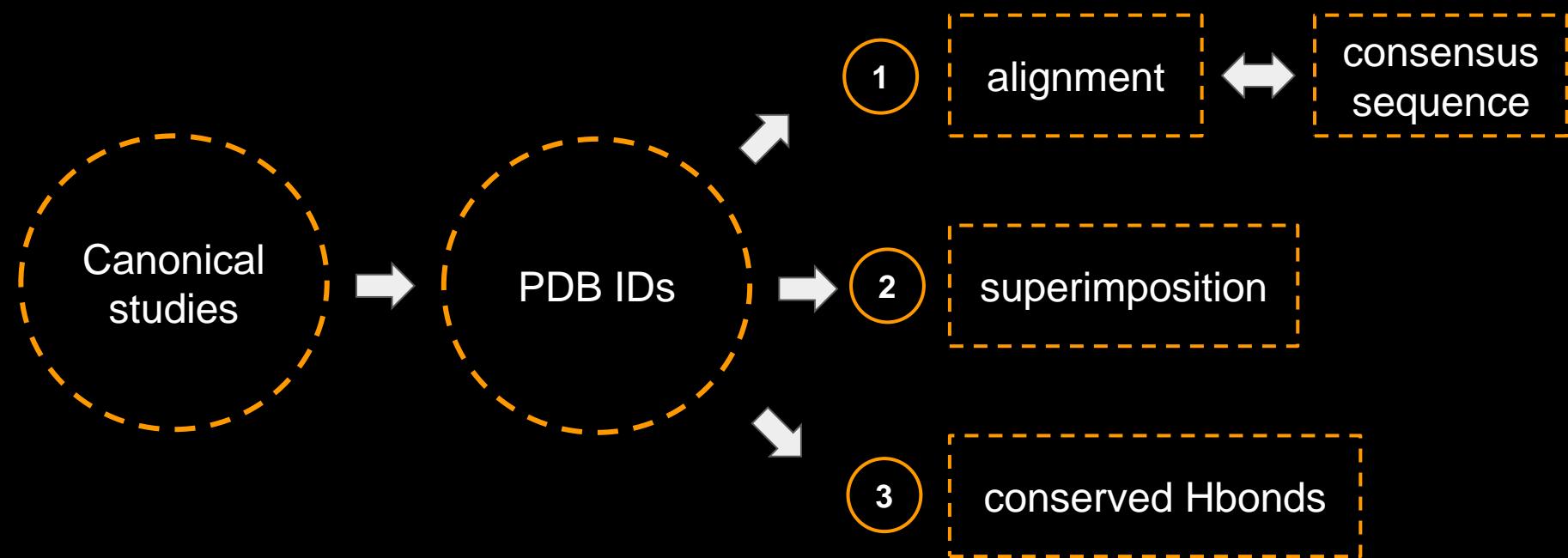
# Canonical structures



### *Chothia et al* classification:

LIGHT	L1	$\mathcal{K}$ 1 <b>2a</b> <b>2b</b> 3 <b>4</b> 5 6 $\Lambda$ 1 2 <b>3a</b> <b>3b</b> 4
	L2	<b>1</b>
	L3	$\mathcal{K}$ 1 2 3 4 5 $\Lambda$ <b>1a</b> <b>1b</b> <b>1c</b> 2
	H1	1 2 3
HEAVY	H2	1 <b>2a</b> 2b 2c 3a 3b <b>3c</b> 4

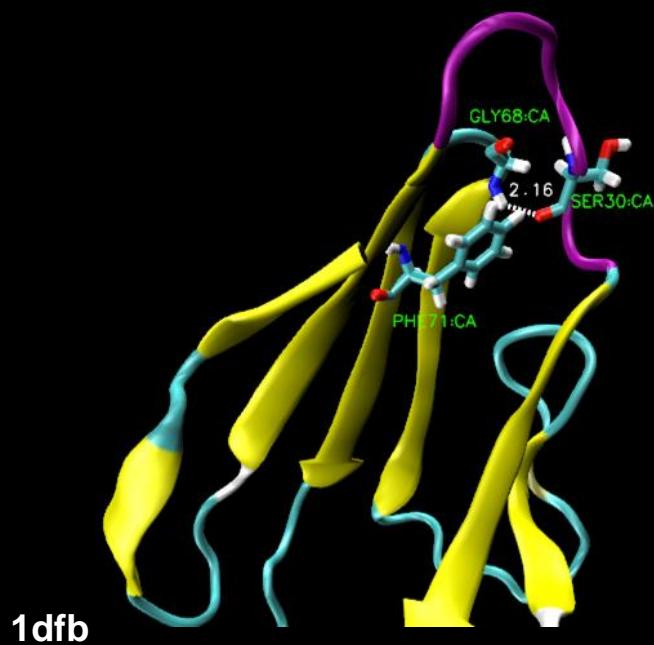
# Canonical structures - Methodology



# L1 - k2

L1: 24-34

## L1-k2a

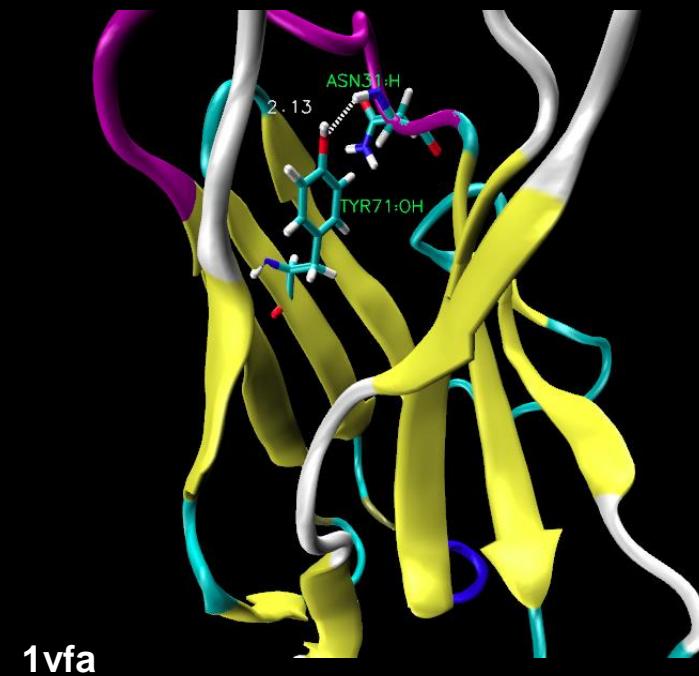


Position 71: Phe, Thr or Gly

### Hbond:

Ser 30 (O) --- Gly 68 (N)

## L1-k2b



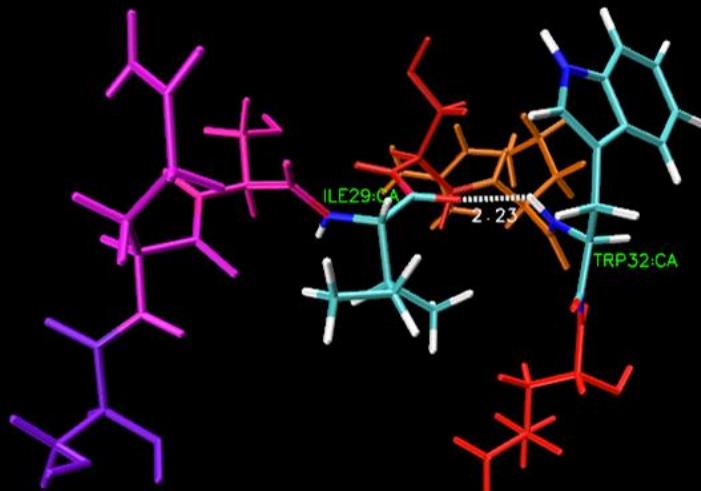
Position 71: Tyr

### Hbond:

Tyr 71 (O) --- Asn 31 (N)

# L1 - k2

## L1-k2a



1dfb

### Hbond:

Ile 29 (O) --- Trp 32 (N)

rASQdisnyla

1dfb RASQDISNYLAWYQQKPG  
1igm QASQDISNYLAWYQQKPG  
1lvl RASQSIGNRLFWYQQKSH  
1fvc RASQDVNTAVAWYQQKPG

RMS = 0,62

## L1-k2b



1vfa

### Hbond:

Ile 29 (O) --- Trp 32 (N)

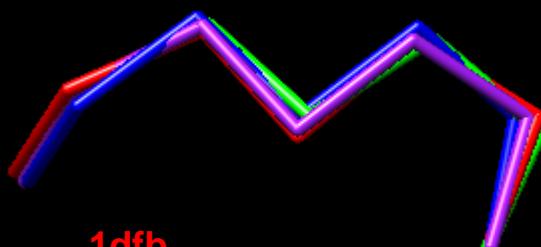
rASqdisnyLa

1zan CRASEDIYNALAW  
1vfa CRASGNIHNYLAW  
1fgv CRASQDINNNLNW

RMS = 0,51

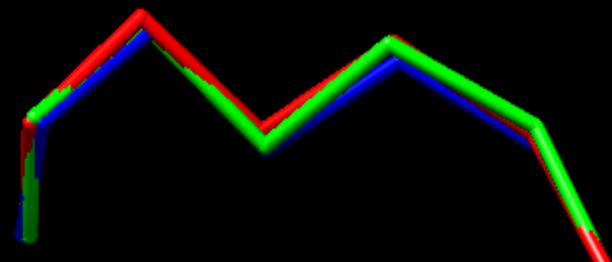
# L1 - k2

L1-k2a

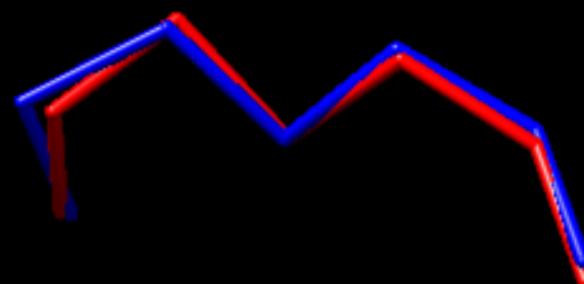


1dfb  
1igm  
1ivl  
7fvc

L1-k2b



1zan  
1vfa  
1fgv

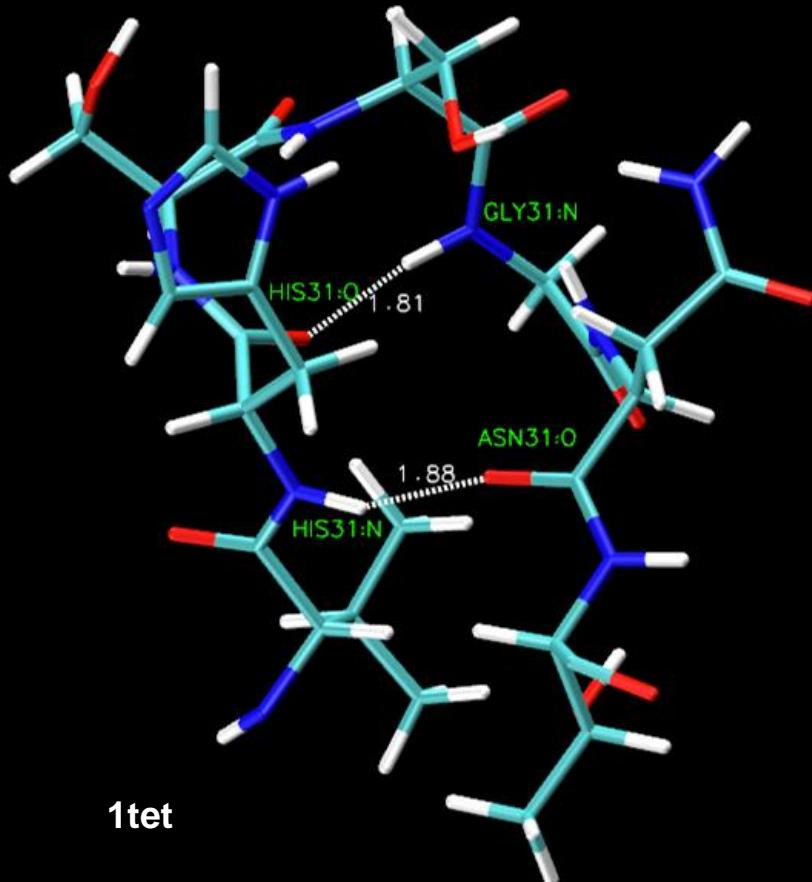


L1-K2A

L1-K2B

RMS = 0,70

# L1 - k4



## Hydrogen bonds:

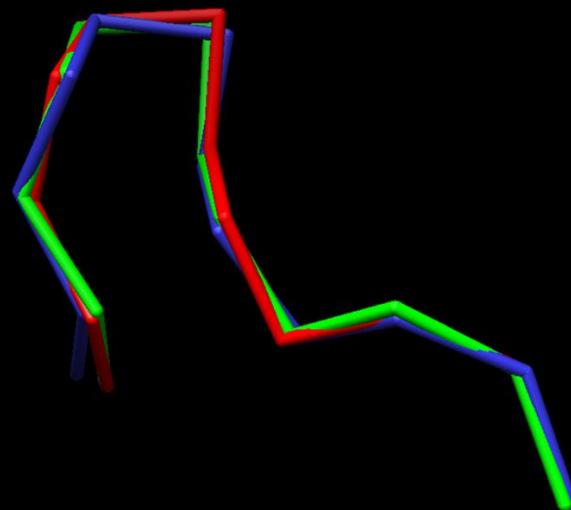
His 31 (O) --- Gly 31.3 (N)  
His 31 (N) --- Asn 31.4 (O)

RSSqslvhhsnGnTYle

1flr	SISCRSS	QSLVHSNGNTYLRW
1tet	SISCKSS	QSIIVHSSGNTYFEW
2cqr	SISCRPS	QSLVHSNGNTYLHW

24

RMS = 0,45



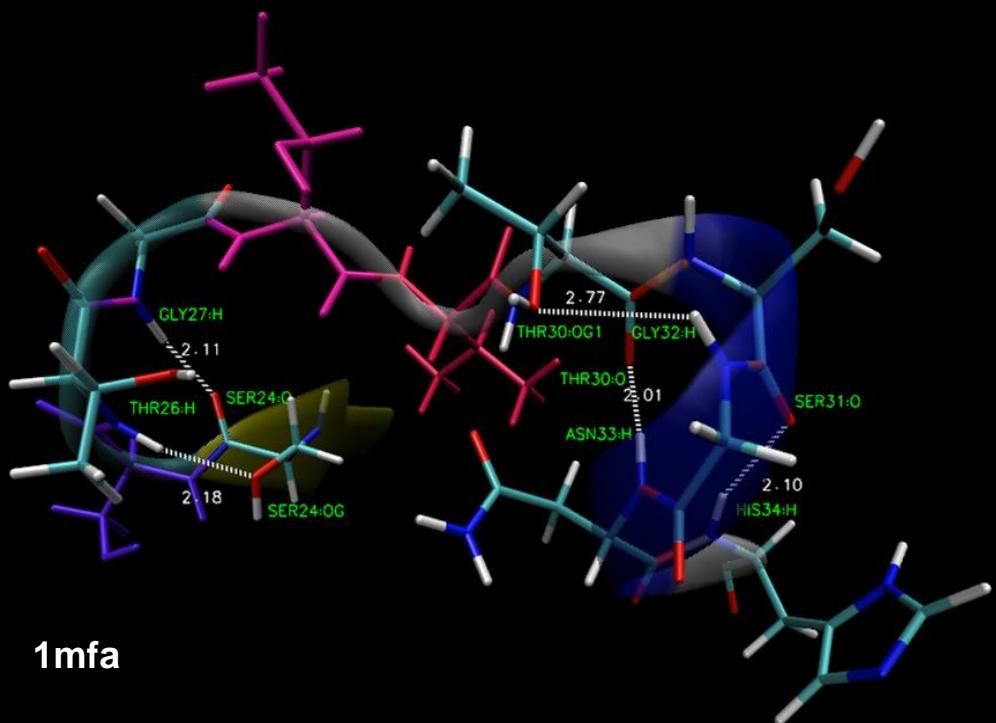
# L1 - λ3

## Hydrogen bonds:

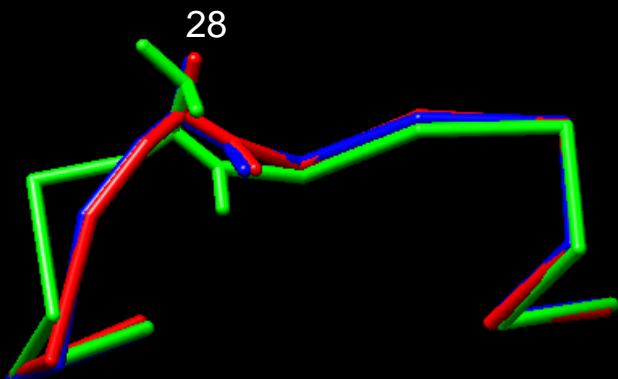
Ser	24	(OG)	---	Thr	26	(N)
Ser	24	(O)	---	Gly	27	(N)
Thr	30	(OG1)	---	Gly	32	(N)
Thr	30	(O)	---	Asn	33	(N)
Ser	31	(O)	---	His	34	(N)

**RSStGavTtsNYAn**

Sequence logo showing the alignment of three sequences: A, B, and C. The x-axis represents positions 1 to 34. The y-axis shows the frequency of each nucleotide (A, T, C, G) at each position. Position 28 is marked with an orange arrow and the number 28. Position 34 is marked with an orange arrow and the number 34.

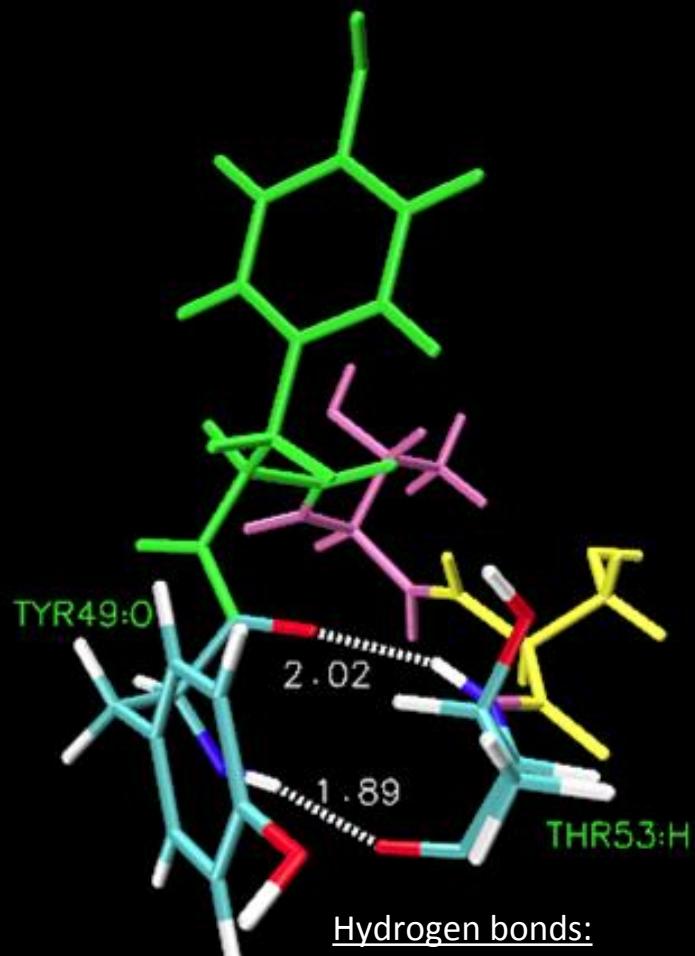


RSM = 0,58



# L2

positions: 50-52

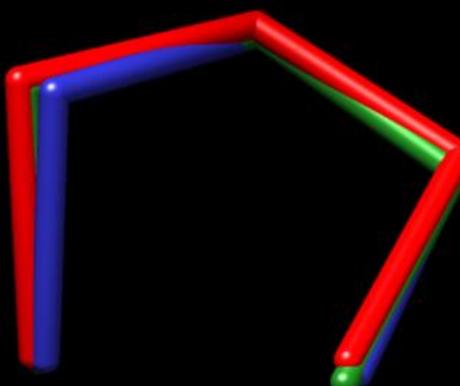


Tyr 49 (N) --- Thr 53 (O)  
Tyr 49 (O) --- Thr 53 (N)

Y-asnlas

1flr KVLIYKVSNRF  
1fgv KLLIYYYTSTLE  
1iqlm ELRIYDASNLE

RMS = 0,52

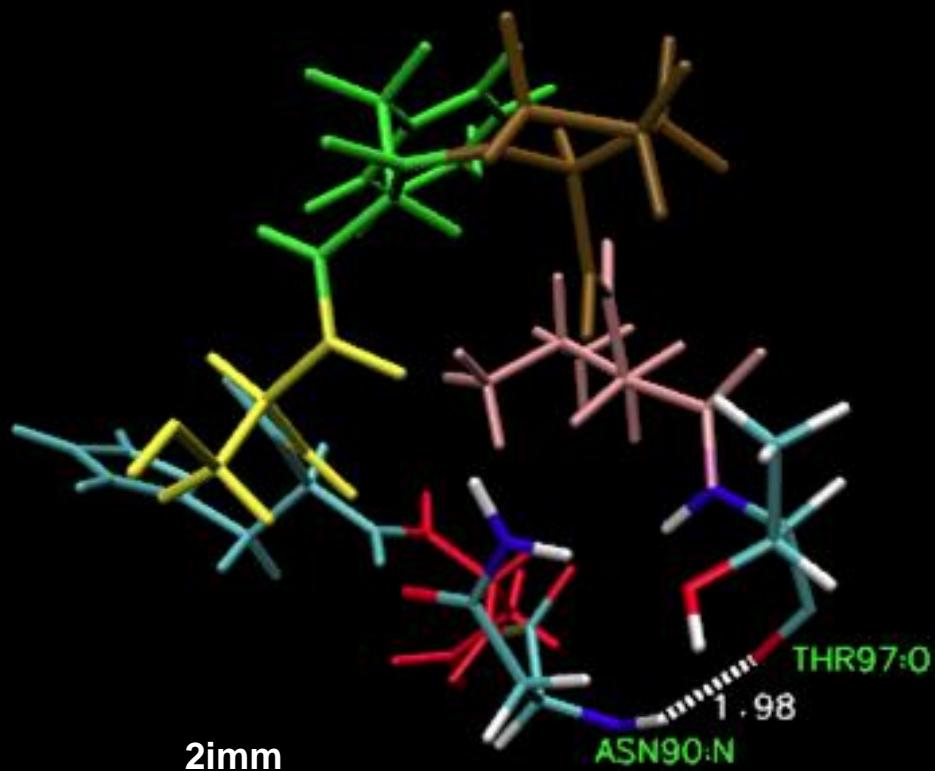


# L3 - k1

L3: 89-96

Hydrogen bond:

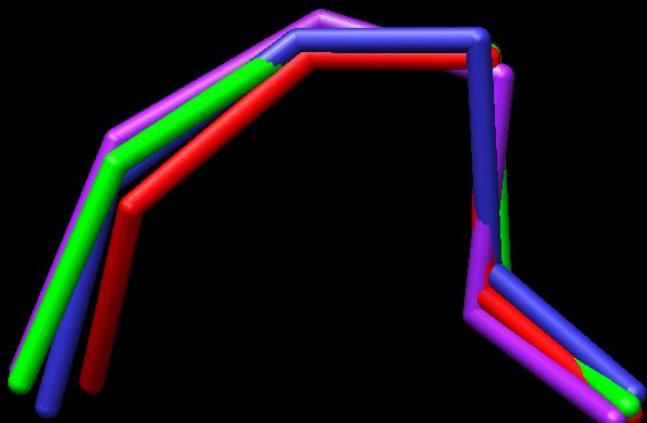
Asn 90 (N) --- Thr 97 (O)



qQgss-PIT

1hil	YCQ	NDY	SNP	LTF	GGG
2imm	YCQ	NDH	SYPL	TFG	GAG
1fgv	YCQ	QQG	GNTL	PPT	FGAG
1fvc	YCQ	QHY	TTP	PPT	FGQG

RMS = 0,70



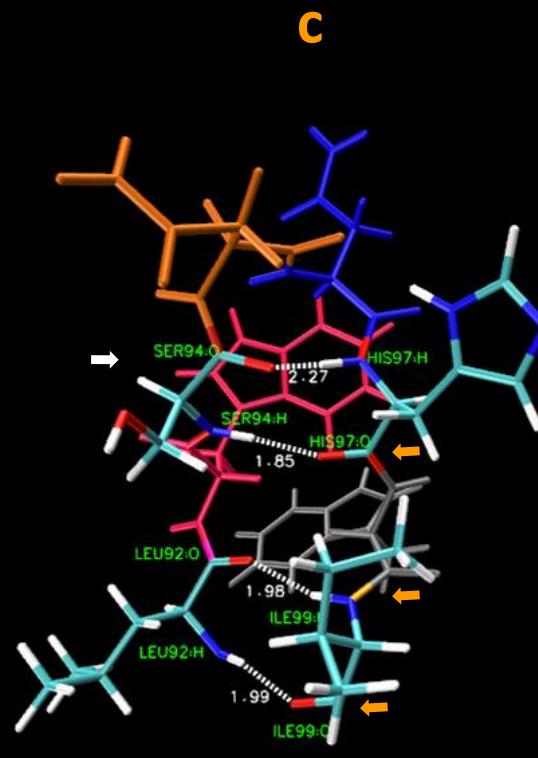
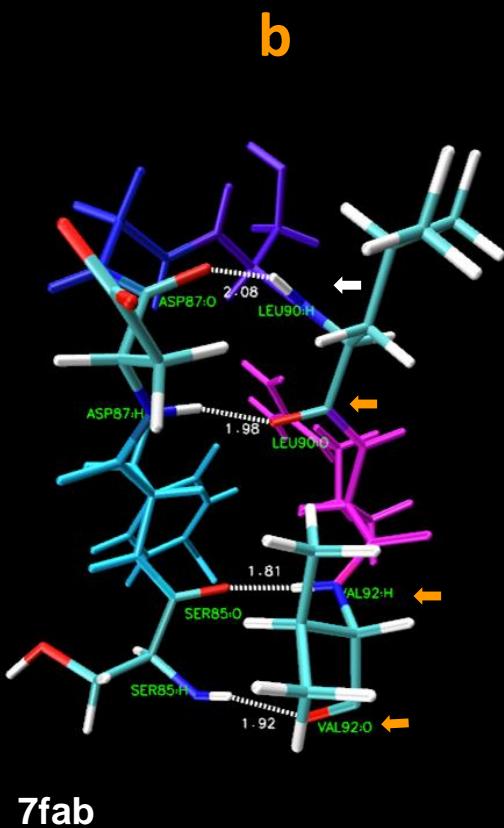
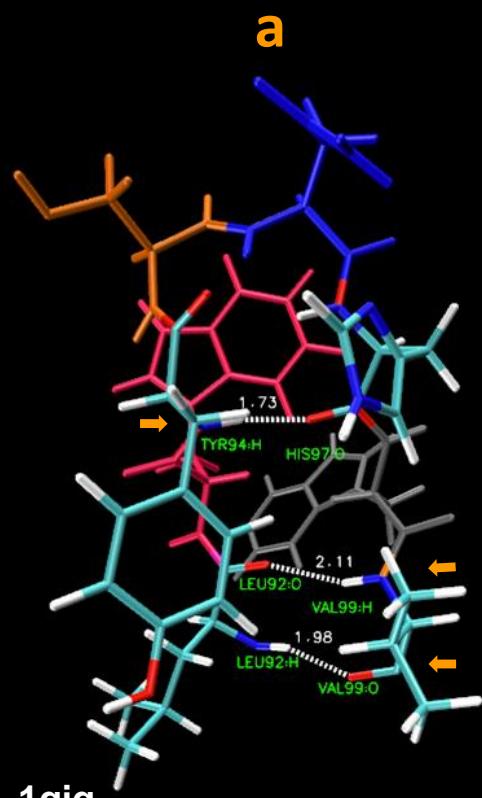
# L3 - $\lambda 1$

- Conserved Hbond between a, b and c
- Conserved Hbond between b and c

alw-snhwv

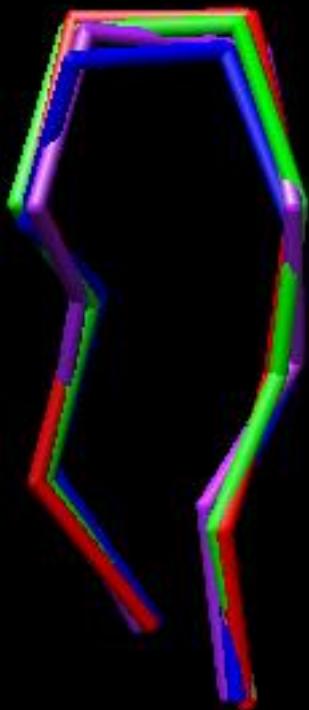
A	1ind	FCALWYSNLWVFGG
C	1mfa	FCALWSNNHWIFGG
A	1gig	FCALWYSNHWFGG
B	7fab	YQQSYDRSLRVLFGG

RMS = 0,57



Different orientation residues 93 and 94

# L3 - $\lambda 1$



**1ind** (a)  
**1mfa** (c)  
**1gig** (a)  
**7fab** (b)

# H1 - 1

H1: 26-33

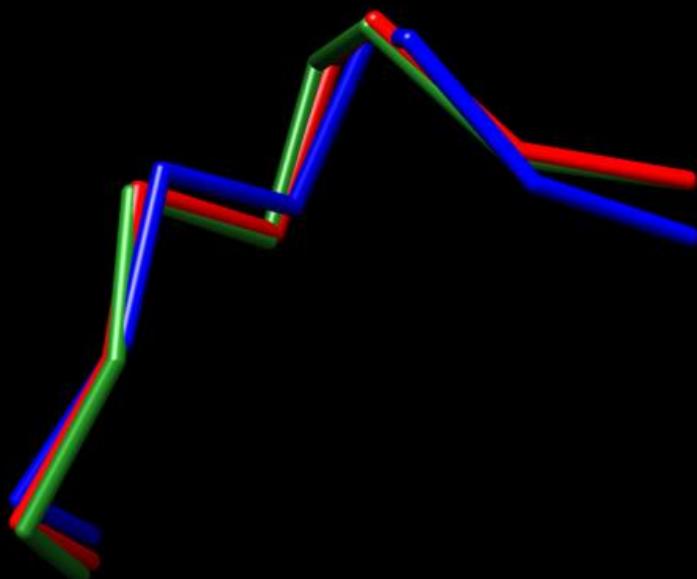
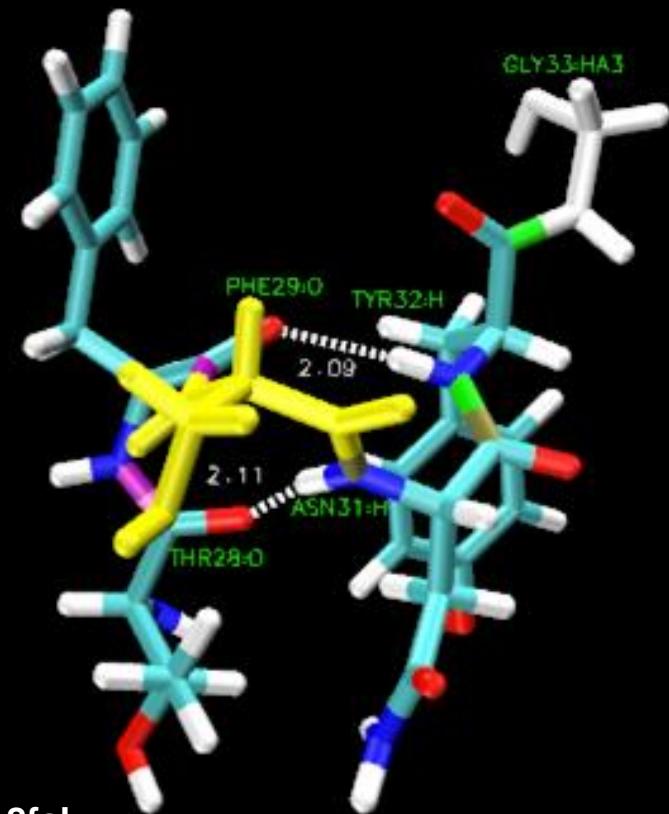
## Hydrogen bonds:

Thr 28 (O) --- Asn 31 (N)  
Phe 29 (O) --- Tyr 32 (N)

ka**SGftftdyymh**

**1hil** AAS**GF**SFSSYGM**S**  
**8fab** IA**S**GFTFSNYGM**H**  
**1mfa** KA**S**GYTFTNYWM**H**

RMS = 0,95



# H2 - 1

H2: 52-56

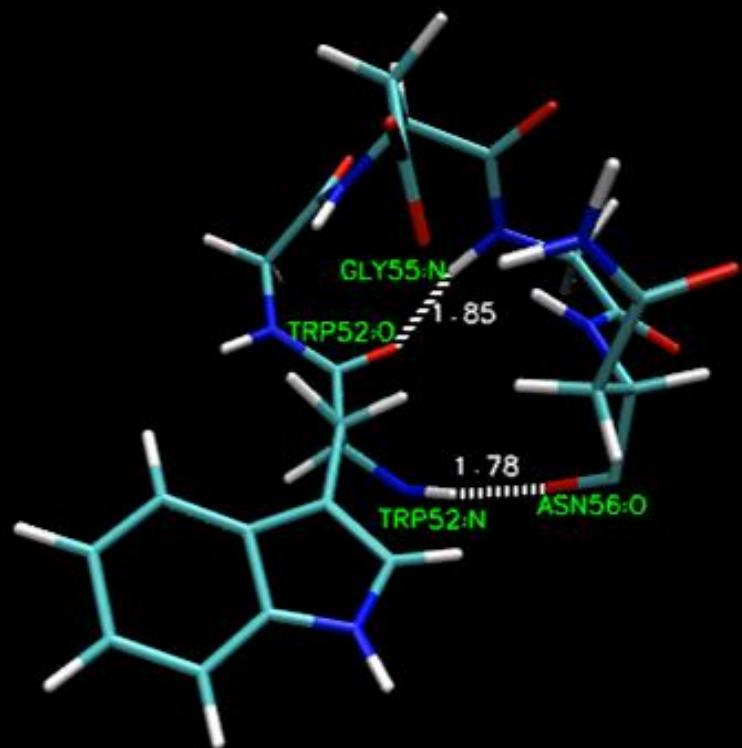
Hydrogen bonds:

Trp 52 (N) --- Gly 55 (N)  
Trp 52 (O) --- Asn 56 (O)

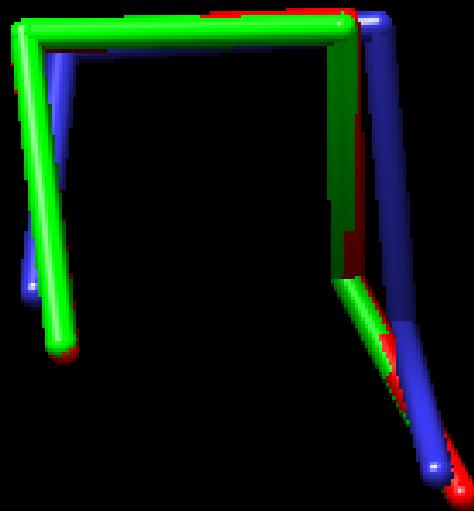
yIwys**G**sty

1vfa EWLGMIWGDGNTD  
7fab EWIGYVFYTGTTL  
1kip EWLGMIWGDGNTD

RMS = 0,41



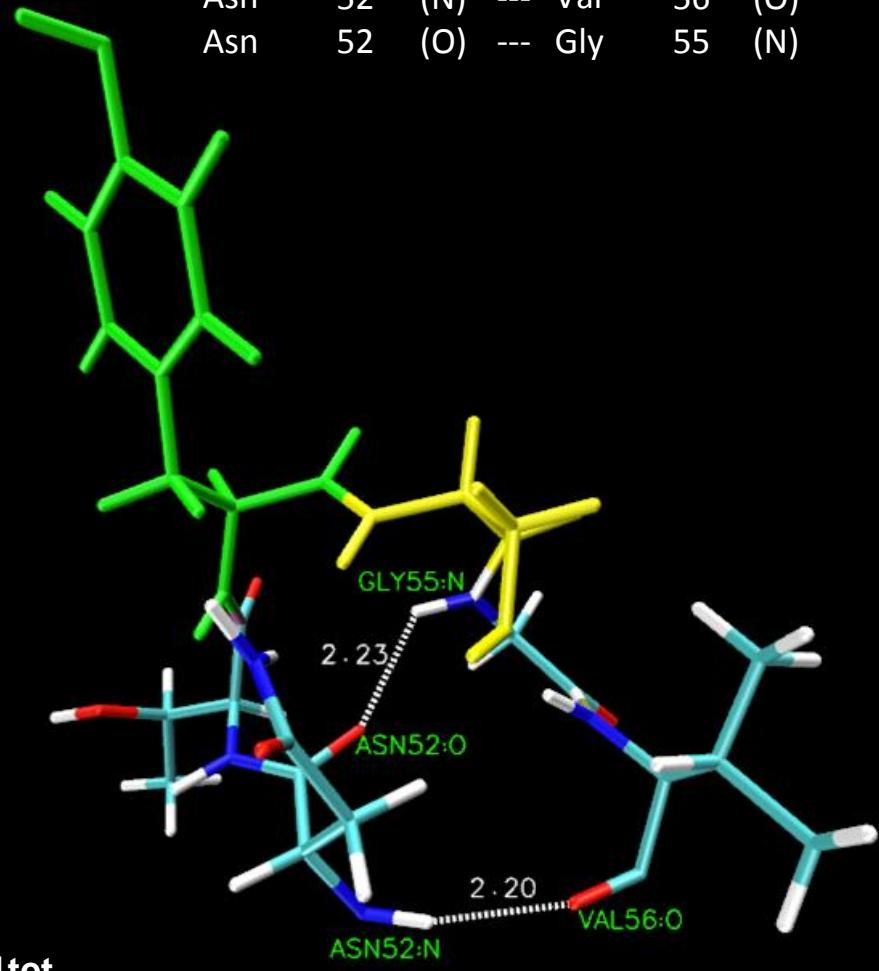
1vfa



# H2 - 2a

## Hydrogen bonds:

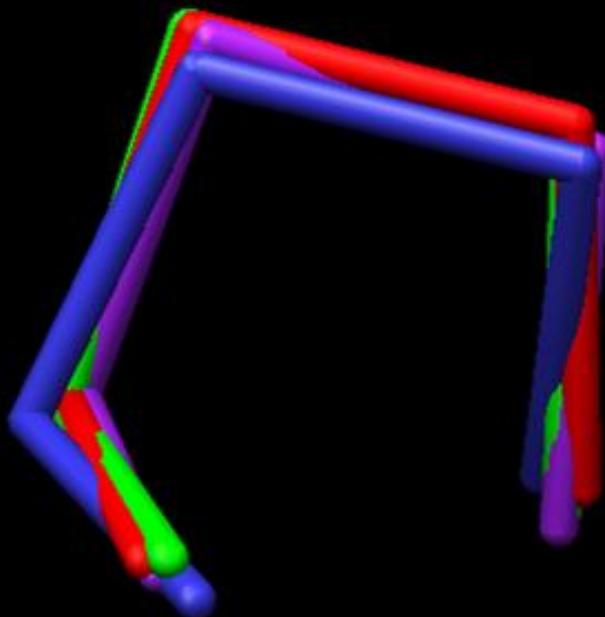
Asn 52 (N) --- Val 56 (O)  
Asn 52 (O) --- Gly 55 (N)



-I ypgng- t -

1tet	KWMGWI	NTYSGVPTY
1mfa	EWIGAI	YPGNSATFY
1fvc	EWVARI	YPTNGYTRY
1fgv	EWVAGI	NPKNNGTSY

RMS = 0,47



# H2 - 3c

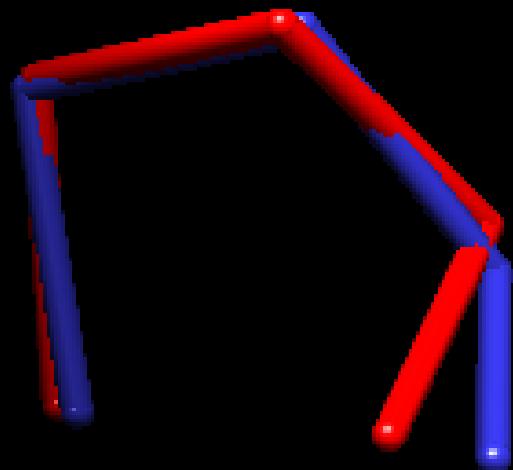
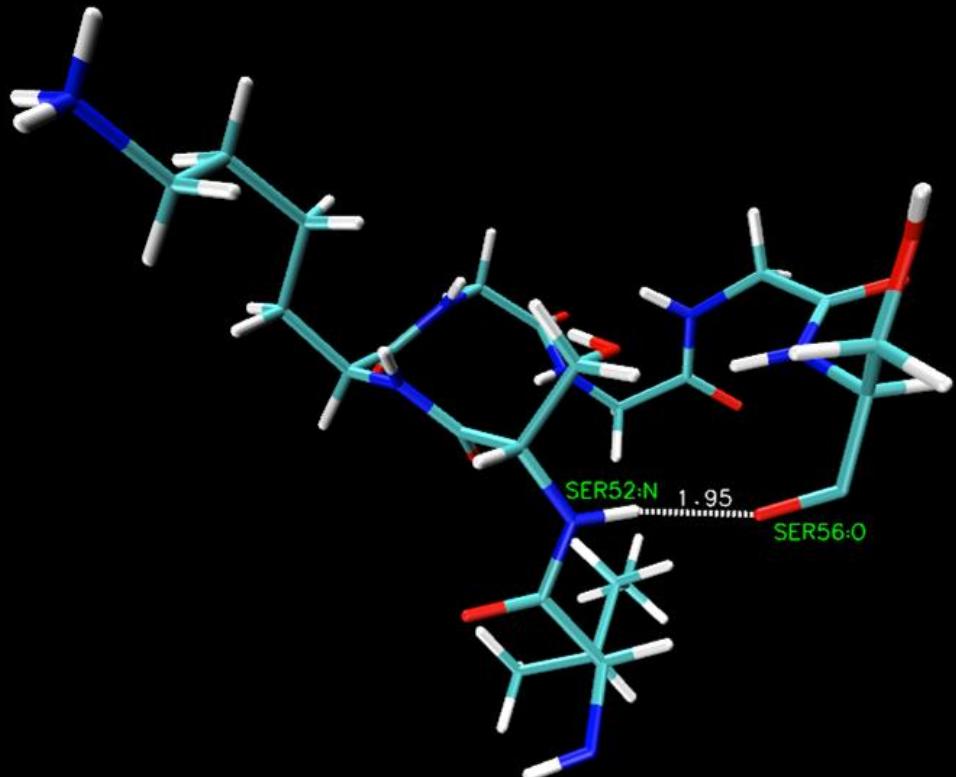
Hydrogen bonds:

Ser 52 (N) --- Ser 56 (O)

- Issgggnty

1igm SGVFGSGGNTDY  
1seq AYISKGGGSTYY

RMS = 0,51



1seq

# Canonical structures - Conclusions

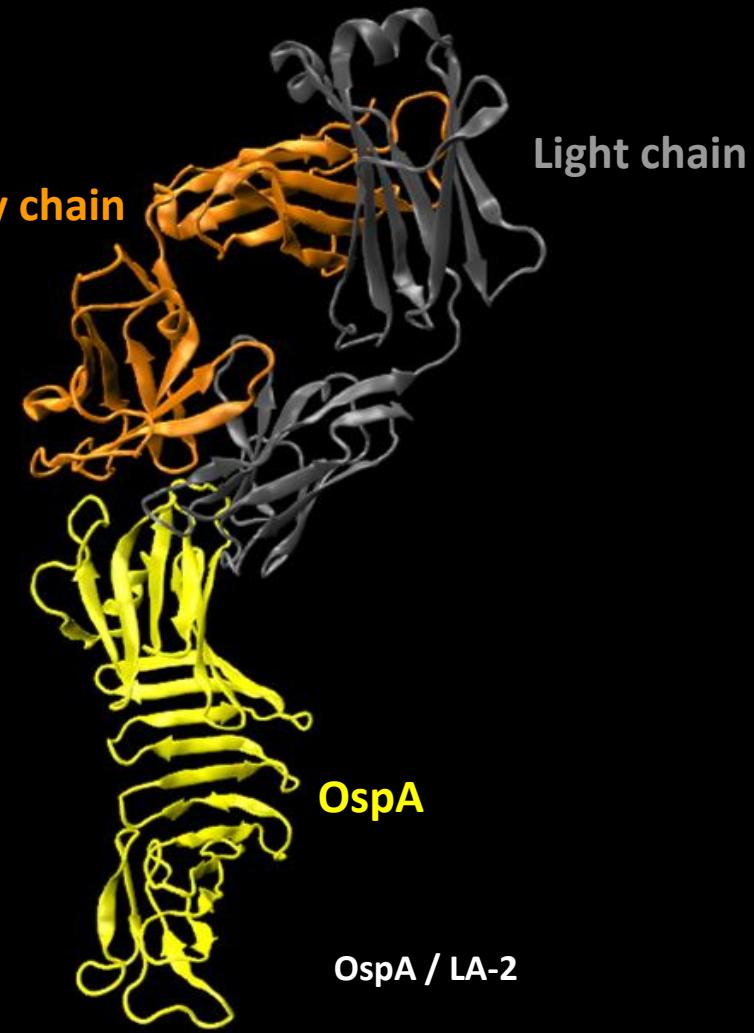
- Even though the sequence is different, the structure is conserved
- In the spatial conformation of each loop may be also involved residues from the  $\beta$ -strands
- Canonical structures may not be involved in determining specificity but affinity

# Ag-Ab interaction: ospA/LA-2



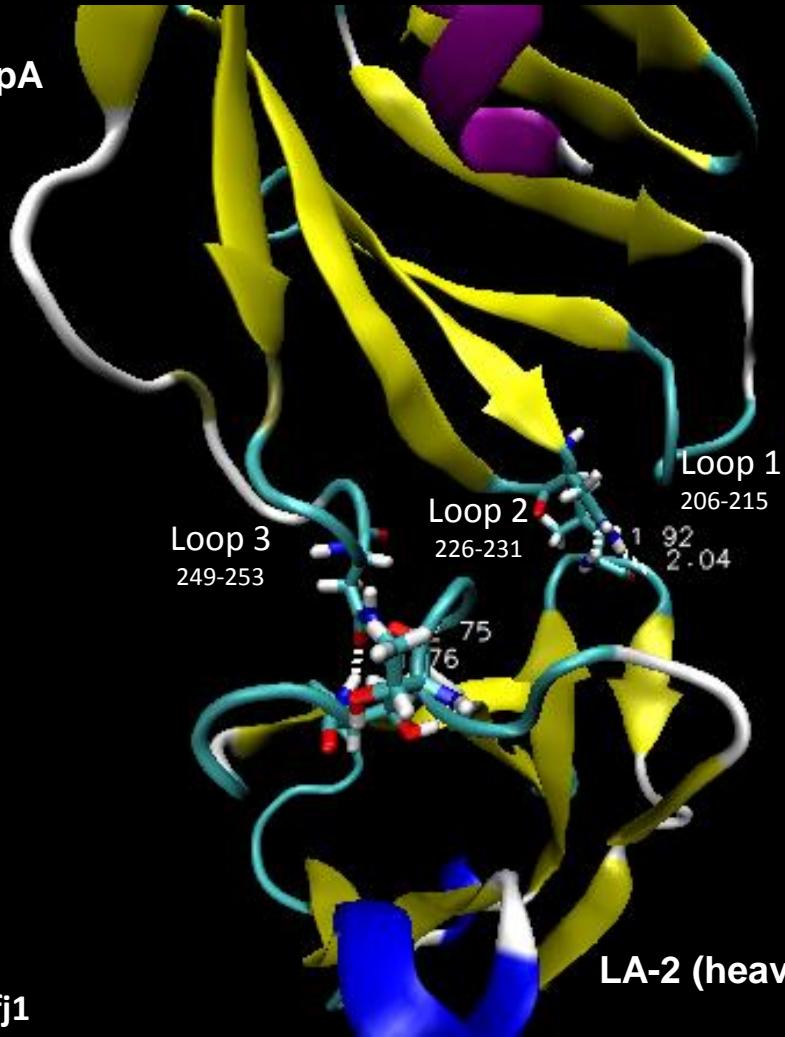
1fj1

Heavy chain



# OspA

OspA



*Borrelia burgdorferi*

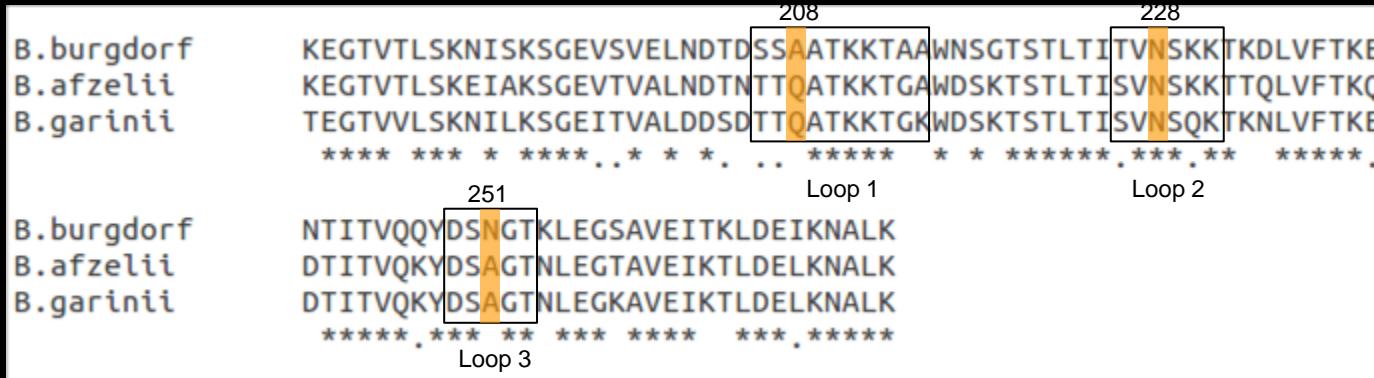
*Borrelia afzelii*

*Borrelia garinii*

→ CONSERVED EPITOPE

Why are the three species  
**not** recognised by LA-2?

# Important residues from OspA



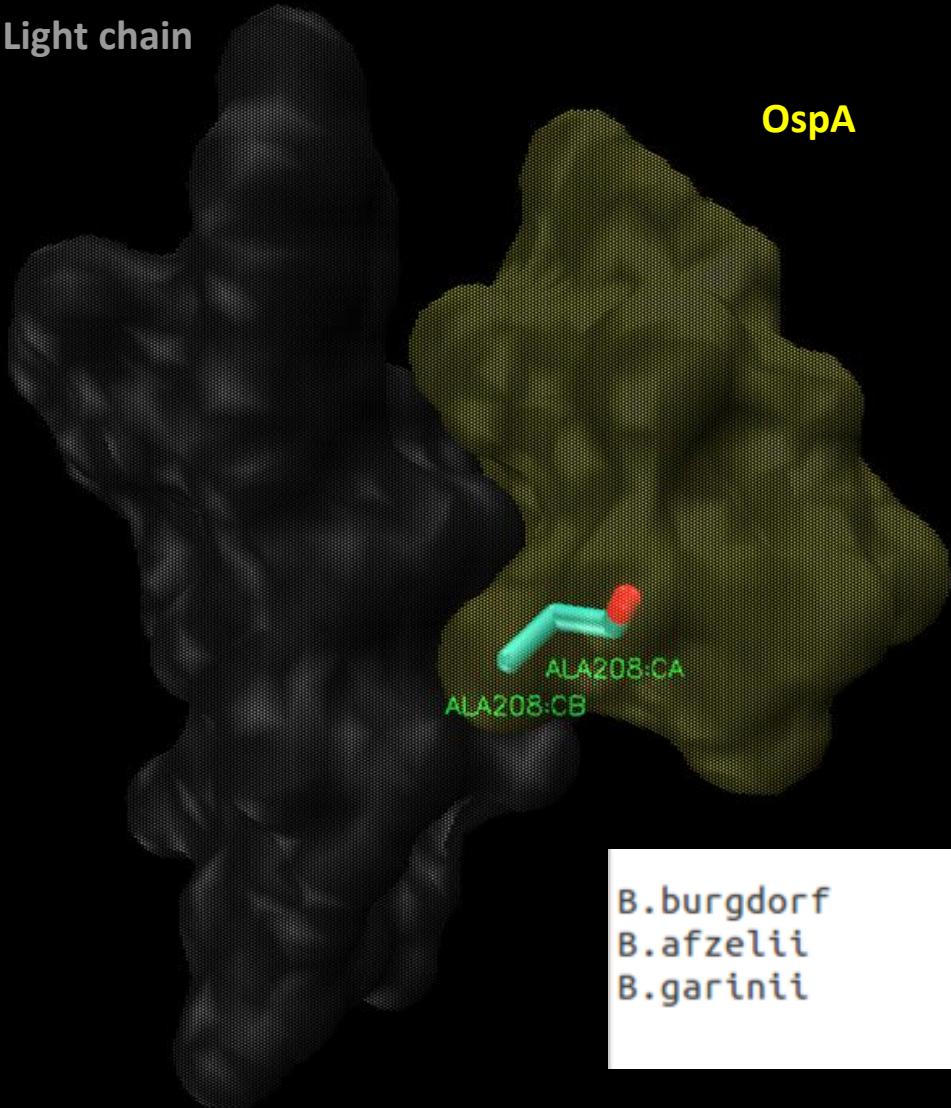
Any mutation in these three residues from OspA abolish binding to LA-2 antibody



WHY?

# A208 from OspA (loop 1)

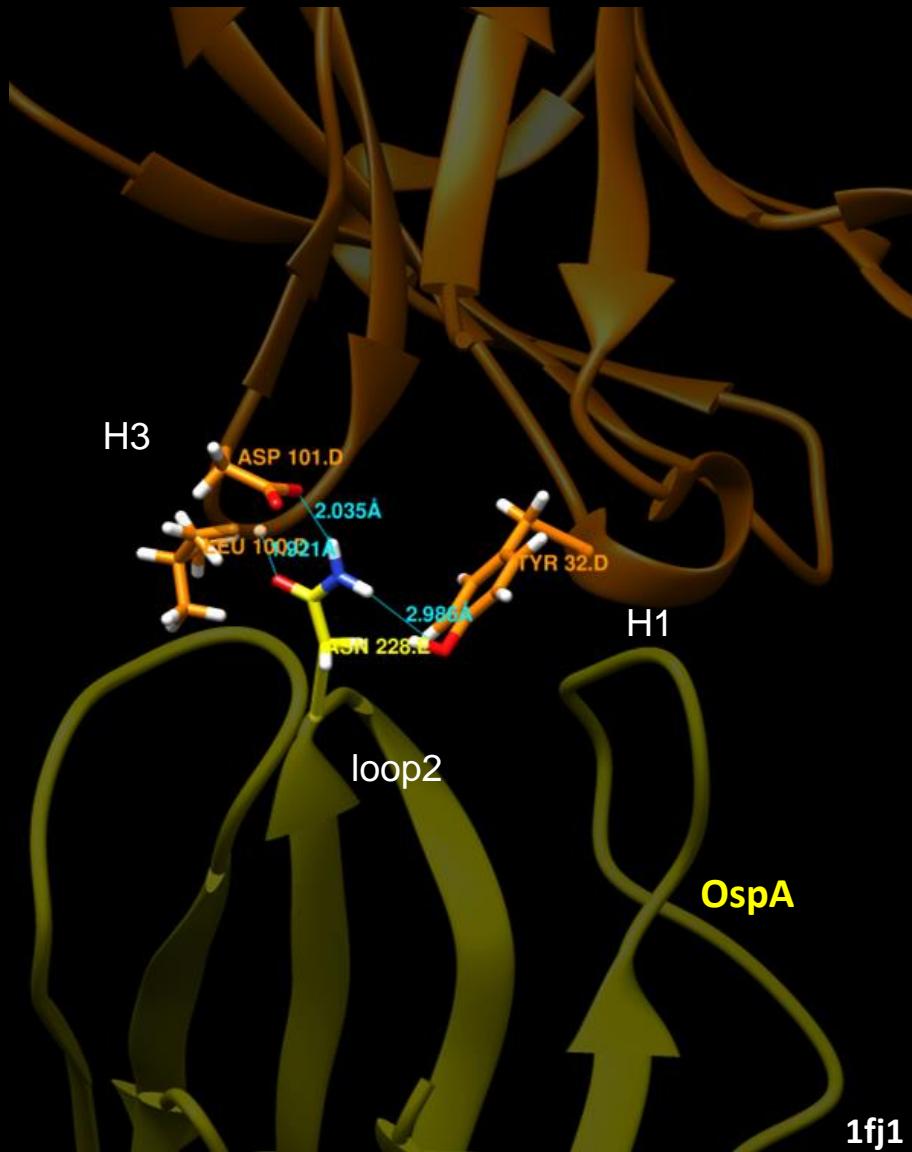
## Light chain



- B.burgdorffii
- B.afzelii
- B.garinii

KEGTVTLSKNISKSGEVVELNDTSSAATKKTAAWNS  
KEGTVTLSKEIAKSGEVTVLNDTNTTQATKKTGAWDS  
TEGTVVVL SKNILKSGEITVALDDSDTTQATKKTGKWDS  
\*\*\*\*\* \*\*\* \* \*\*\*\*\* \* \* \* \*\*\*\*\* \* \*

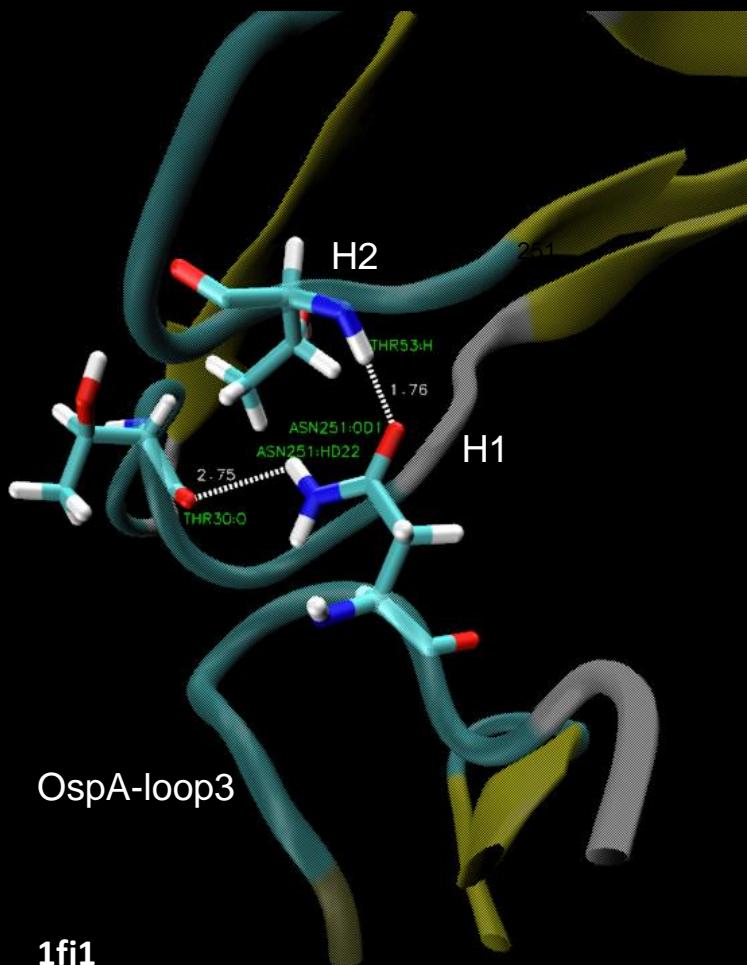
# N228 from OspA (loop 2)



## Hydrogen bonds:

<b>OspA</b>			<b>LA-2</b>		
Asn	228	(ND2)	---	Asp	101 (OD1)
Asn	228	(OD1)	---	Leu	100 (N)
Asn	228	(ND2)	---	Tyr	32 (OH)

# N251 from OspA (loop 3)



## Hydrogen bonds:

	OspA			LA-2		
Asn	251	(ND2)	---	Thr	30	(O)
Asn	251	(OD1)	---	Thr	53	(N)

B.burgdorff  
B.afzelii  
B.garinii

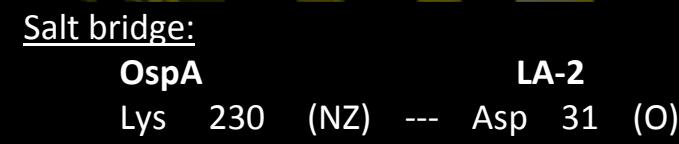
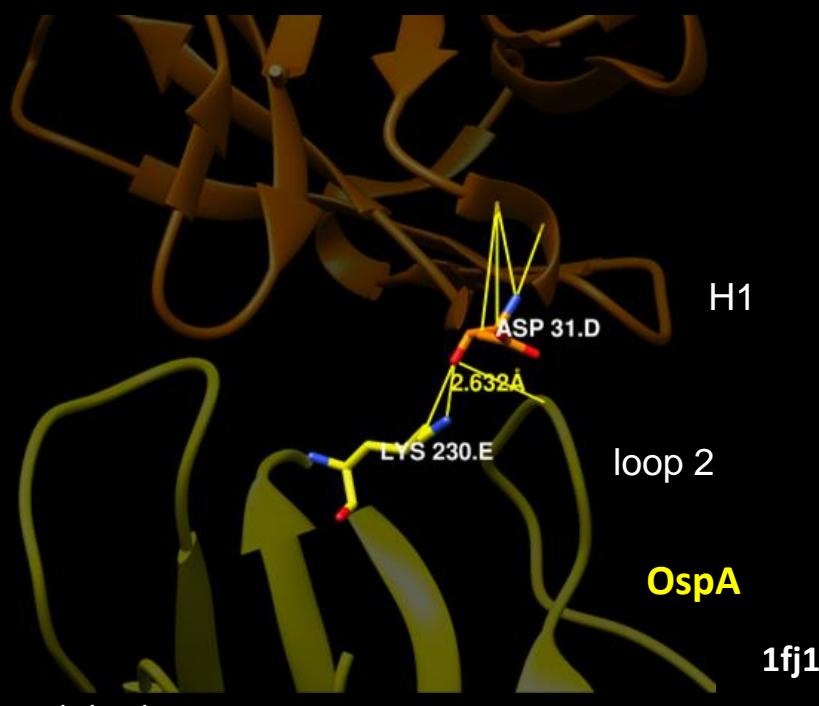
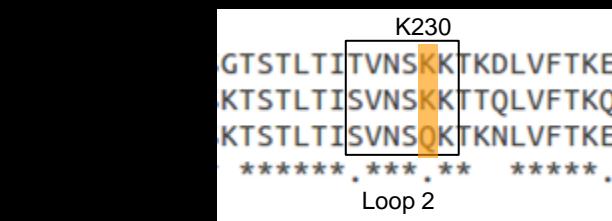
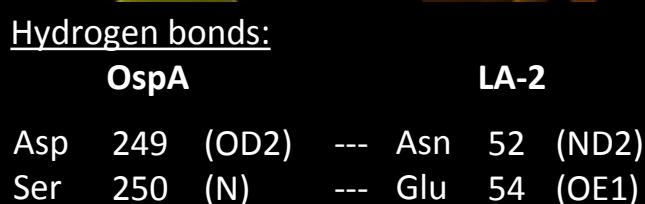
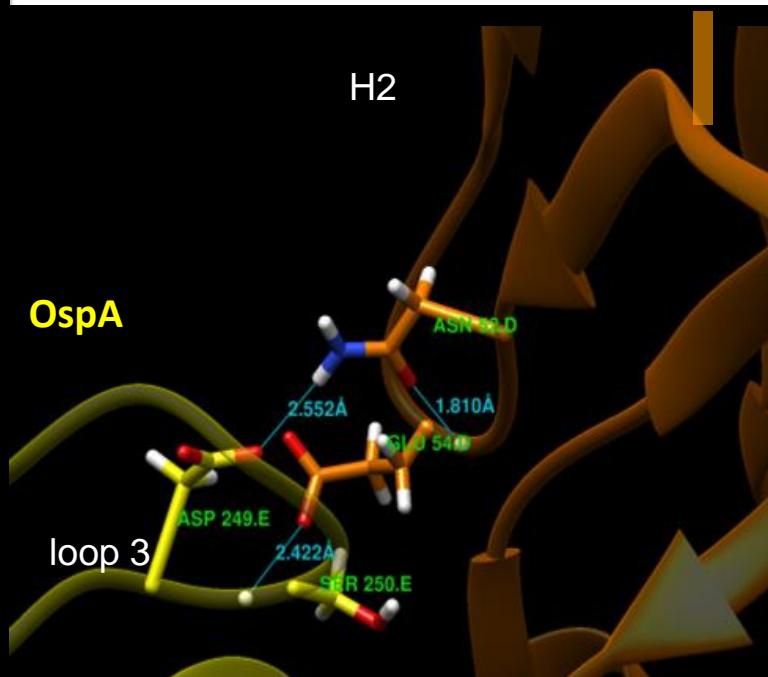
N251

NTITVQQYDSNGTTLEGSAVEITKLDEIKNALK  
DTITVQKYDSAGTNLEGTAVEIKTLDELKNALK  
DTITVQKYDSAGTNLEGKAVEIKTLDELKNALK

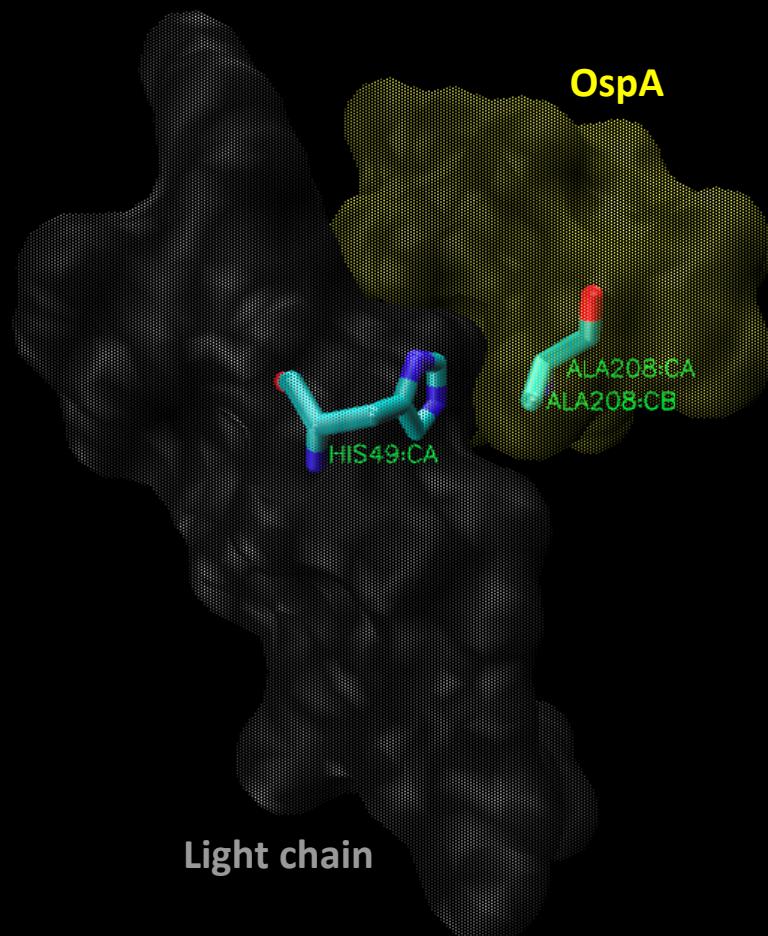
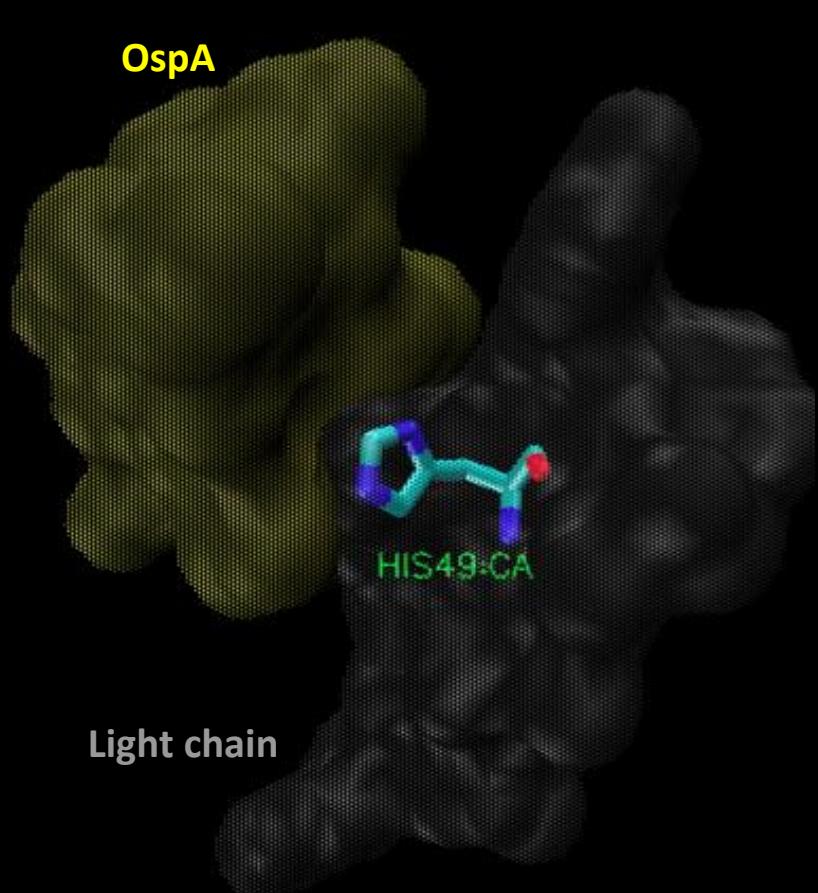
\*\*\*\*\*.\*\*\*\*\*.\*\*\*.\*\*\*\*\*.\*\*\*.\*\*\*\*\*

Loop 3

# Other interactions between Ag-Ab



# Important residues from LA-2



# Ag-Ab: Conclusions

- A208, N228 and N251 from OspA are key for binding to LA-2 being the reason why LA-2 only recognises OspA from *B.burgdorferi*
- H49 from the light chain is essential for the function of the antibody

# General conclusions

- The sequence conservation between variable and constant domains is important for the conservation of the structure, but not necessarily for the function
- Even though the CDRs are hypervariable regions there are some structures that are more frequently found than others
- There are key residues for the antigen-antibody interaction but they are different depending on the antigen the antibody has to recognise

# References

- Al-Lazikani B, Lesk A, Chothia C. Standard conformations for the canonical structures of immunoglobulins. *J Mol Biol.* 1997;273(4):927-948.
- Chothia C, Lesk A, Tramontano A, Levitt M, Smith-Gill S, Air G et al. Conformations of immunoglobulin hypervariable regions. *Nature.* 1989;342(6252):877-883.
- Chothia C, Lesk AM. Canonical structures for the hypervariable regions of immunoglobulins. *J Mol Biol.* 1987 Aug 20;196(4):901-17.
- Chothia C, Gelfand I, Kister A. Structural determinants in the sequences of the immunoglobulin variable domain. *J Mol Biol.* 1998;278:475-479.
- Ding W, Huang X, Yang X, Dunn J, Luft B, Koide S et al. Structural identification of a key protective B-cell epitope in lyme disease antigen OspA. *J Mol Bio.* 2000;302(5):1153-1164.
- Haard H, Kazemier B, Bent A, Oudshoorn P, Boender P, Germen B, et al. Absolute conservation of residue 6 of immunoglobulin heavy chain variable regions of class IIA is required for correct folding. *Protein Engineering.* 1998;11(12):1267-1276.
- Halaby DM, Pupon A, Mornon JP. The immunoglobulin fold family:sequence analysis and 3D structure comparisons. *Protein Eng Des Sel.* 1999;12(7):563-571.
- Harris LJ, Larson SB, Hasel KW, Day J, Greenwood A, McPherson A. The three-dimensional structure of an intact monoclonal antibody for canine lymphoma. *Nature.* 1992 Nov 26;360(6402):369-72.
- Harry W, Schroeder J, Cavacini L. Structure and function of immunoglobulins. *J allergy clin immunol.* 2010 Feb; 125:41-52.

# References

- Joerger TR, Du C, Linthicum DS. Mol immunol. 1999;36(6):373-386.
- Mosca R, Céol A, Aloy P. Interactome3D: adding structural details to protein networks. Nat Methods. 2012;10(1):47-53.
- North B, Lehmann A, Dunbrack R. A New Clustering of Antibody CDR Loop Conformations. J Mol Biol. 2011;406(2):228-256.
- Putnam FW, Takahashi N, Tetaert D, Debuire B, Lien LC. Amino acid sequence of the first constant region domain and the hinge region of the  $\delta$  heavy chain of the human IgD. Proc Natl Acad. Sci. 1981;78(10):6168-6172.
- Shandilya S, Kurt Yilmaz N, Sadowski A, Monir E, Schiller Z, Thomas W et al. Structural and molecular analysis of a protective epitope of Lyme disease antigen OspA and antibody interactions. J Mol Recogit. 2016. Doi: 10.1002/jmr.2595
- Sheriff S, Jeffrey PD, Bajorath J. Comparison of CH1 domains in different classes of murine antibodies. J Mol Biol. 1996;263:385-389.
- Stephen R. Campion. Conserved aromatic residues as determinants in the folding and assembly of immunoglobulin variable domains. Mol Immunol. 2016;70:63-71.
- Vargas-Madrazo E, Lara-Ochoa F, Carlos Almagro J. 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.

# Multiple choice questions

1. Which of the CDR is the most hypervariable and does NOT follow the canonical structures?:

- a) L1
- b) L3
- c) a and b are correct
- d) H3
- e) a, b and d are correct.

2. What is a CDR ?

- a) One of the sheets of a beta barrel in the immunoglobulins
- b) all the loops between the beta strands of the immunoglobulin
- c) **hypervariable region that plays an important role in the antigen recognition**
- d) A domain of the immunoglobulin
- e) a type of immunoglobulin

3. Immunoglobulins in SCOP classification are considered

- a) All alpha
- b) **All beta**
- c) alpha/beta
- d) alpha-beta
- e) this group of proteins is not present in the SCOP classification

# Multiple choice questions

4. The immunoglobulin structure has:

- a) Two variable domains, one from the heavy chain (VH) and another from the light chain (VL)
- b) A hinge region that connects CH1 with CH2 and the two heavy chains
- c) a and b are correct
- d) Only one constant domain in the heavy chain (CH)
- e) a, b and d are correct

5. The intra-domain disulfide bridge:

- a) Is essential for the immunoglobulins folding.
- b) Is located between sheet B and F in all domains.
- c) a and b are correct.
- d) is essential for the functionality of the immunoglobulins.
- e) all options are correct.

6. Choose the true sentence about canonical structures:

- a) imply a 100% sequence conservation
- b) there are only 2 different canonical structures
- c) they are different for the heavy and the light chain
- d) H3 has canonical structures
- e) any of the above is correct

# Multiple choice questions

**7. The Ig like domain:**

- a) is only found in proteins of the immune system
- b) **is found in all proteins in the immunoglobulin superfamily**
- c) do not require the disulfide bridge
- d) has a  $\beta$  - barrel conformation
- e) all options are false

**8. A canonical structure is:**

- a) A structure that the  $\beta$ -sheets adopt
- b) A structure that all CDR adopt
- c) **A structure that some CDRs adopt more frequently**
- d) A and B are true
- e) A, B and C are true

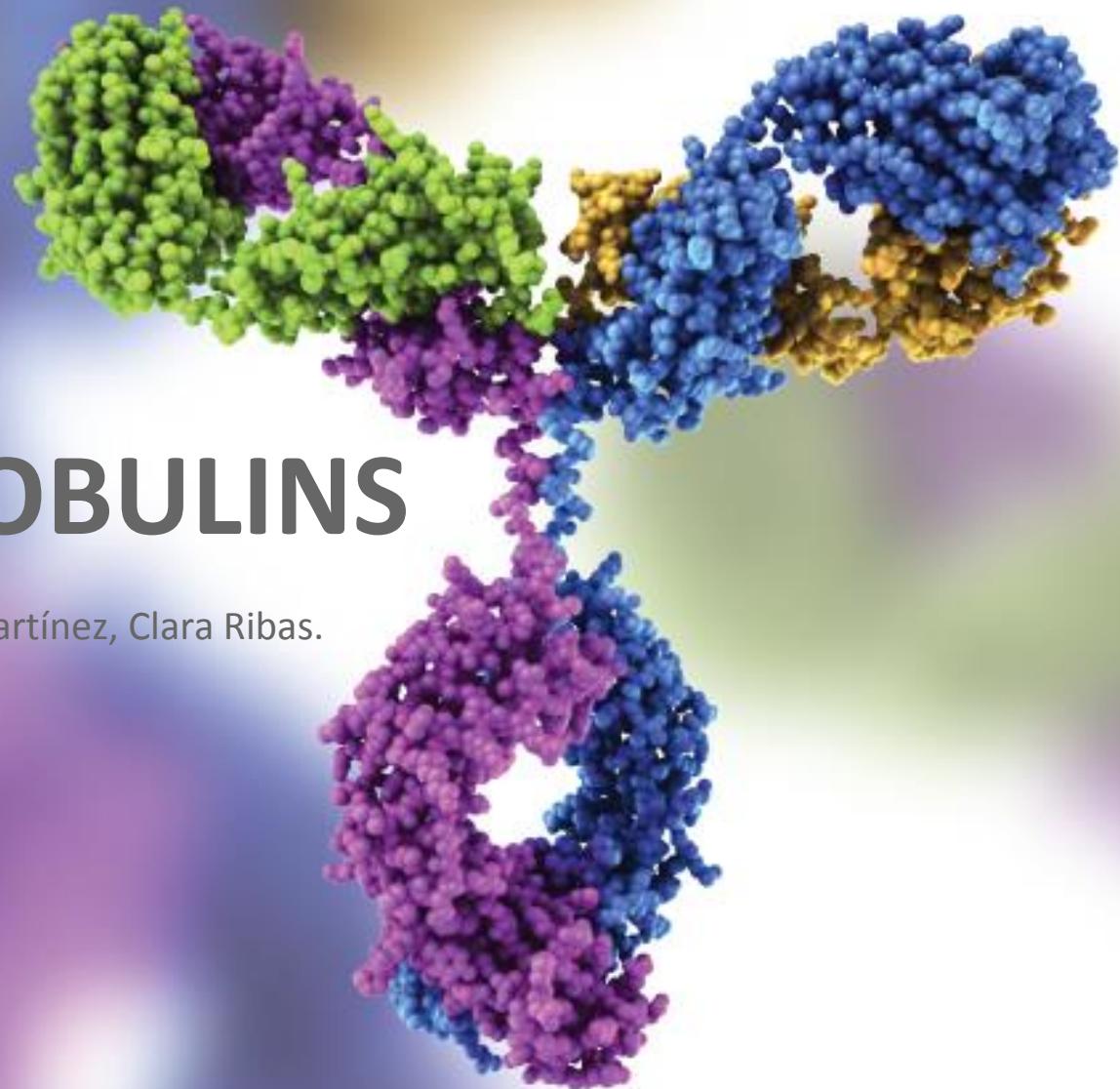
**9. Choose the correct sentence:**

- a) LA-2 antibody is able to recognize the OspA antigen only from *Borrelia burgdorferi*
- b) Between OspA and the LA-2 antibody there are hydrogen bonds only between the heavy chain and the antigen.
- c) There are conserved key residues between all the antigen antibody interactions
- d) **A, B and C are true**
- e) A, B and C are false

# Multiple choice questions

## 10. In immunoglobulins:

- a) the disulfide bond is made between cysteines that are not conserved.
- b) the triad is made between 2 cysteines and one alanine
- c) Hydrophobic residues are in the surface
- d) **Polar residues are located in the surface**
- e) All of them are false



# IMMUNOGLOBULINS

Paula Galiana, Núria Isern, Andrea Martínez, Clara Ribas.