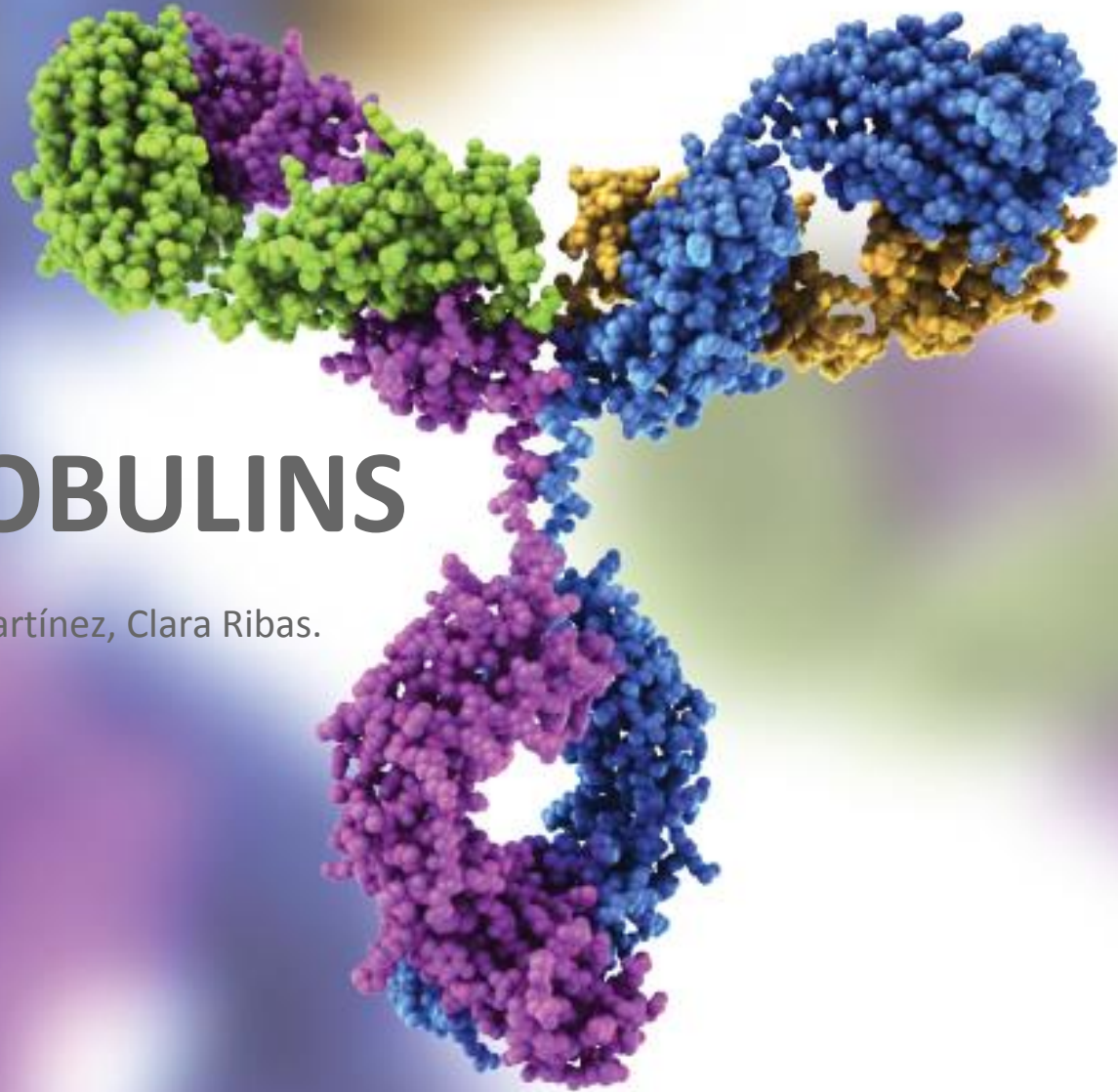


IMMUNOGLOBULINS

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

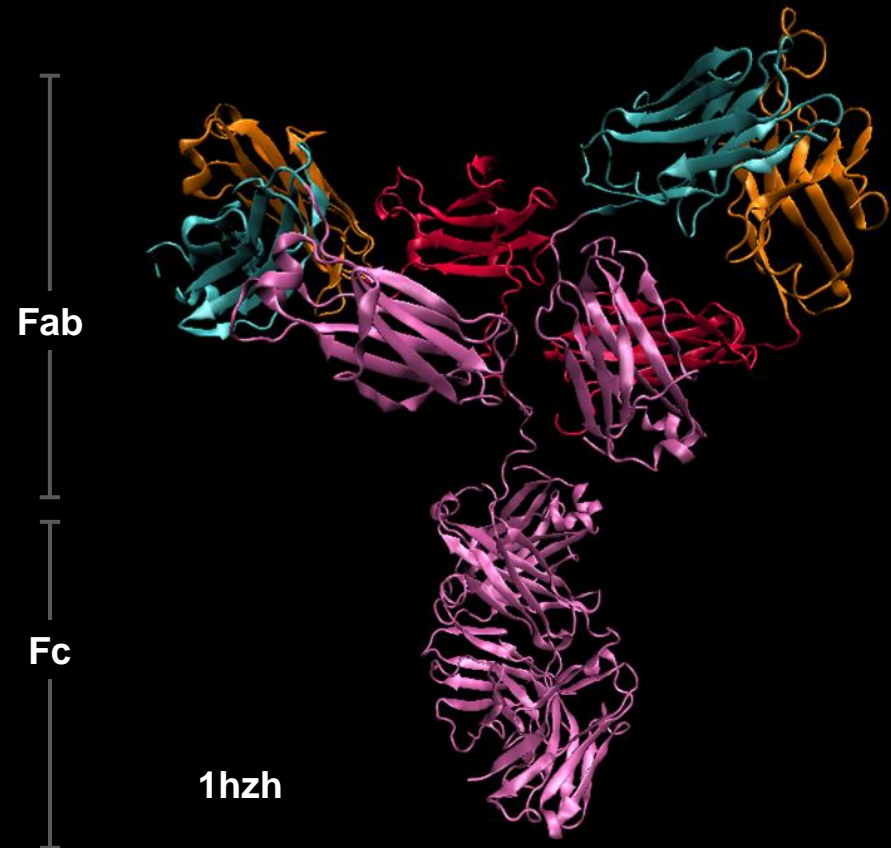
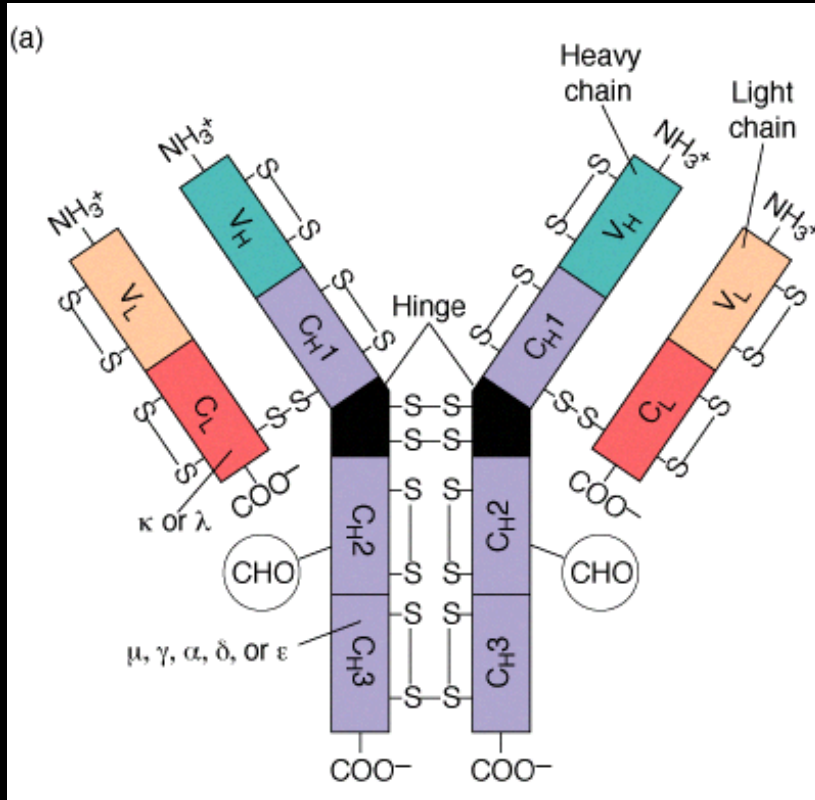
SCOP Classification:

Class: All β protein

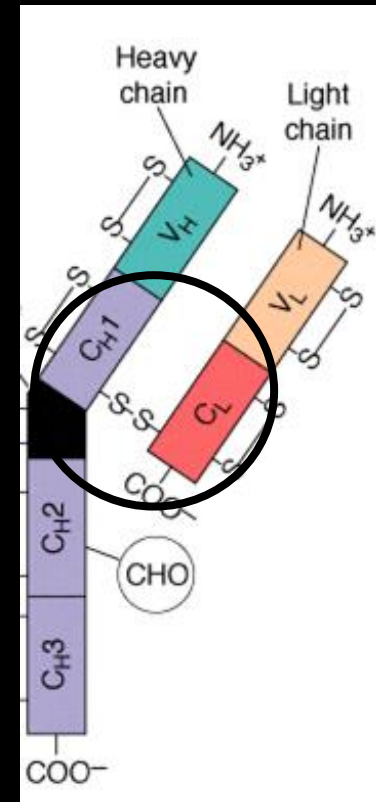
Fold: immunoglobulin-like β -sandwich

Superfamily: Immunoglobulin

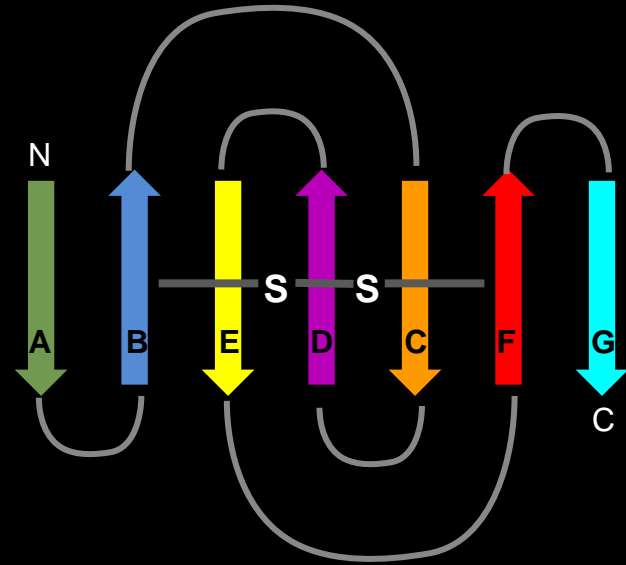
Family: Vset domains or C1 set domains



Immunoglobulin description



Ig constant domain

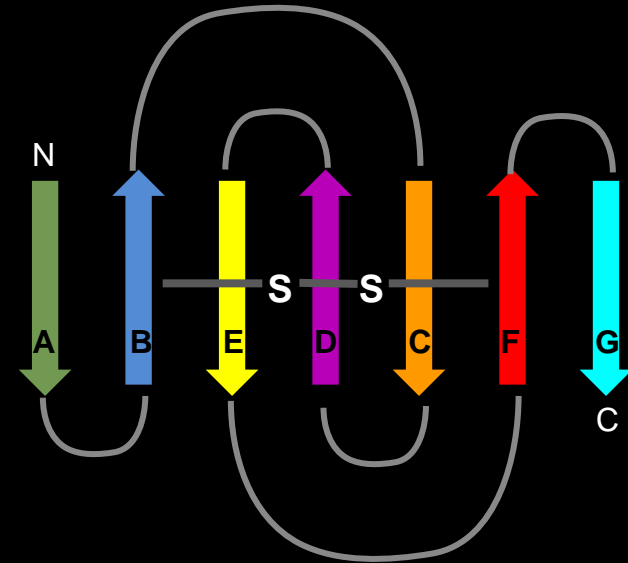
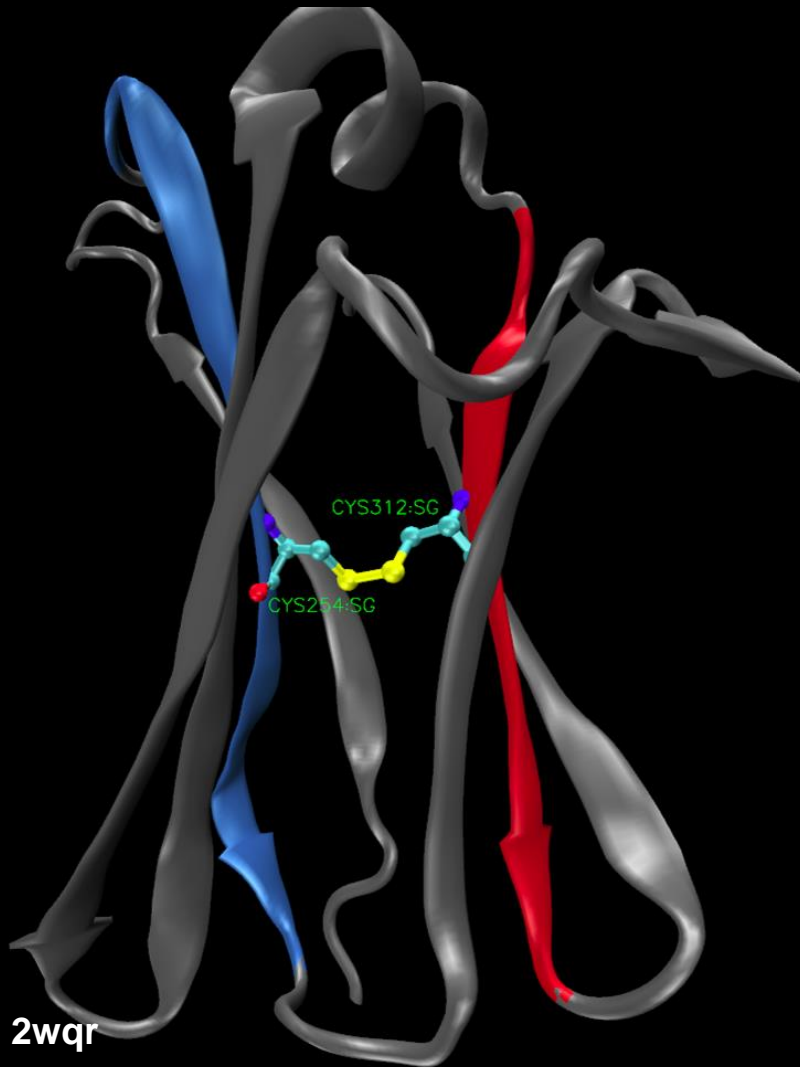


- 7 antiparallel β -strands
- 2 β -sheets:
 - 4-stranded sheet
 - 3-stranded sheet

Immunoglobulin-like β -sandwich

2wqr

Ig constant domain

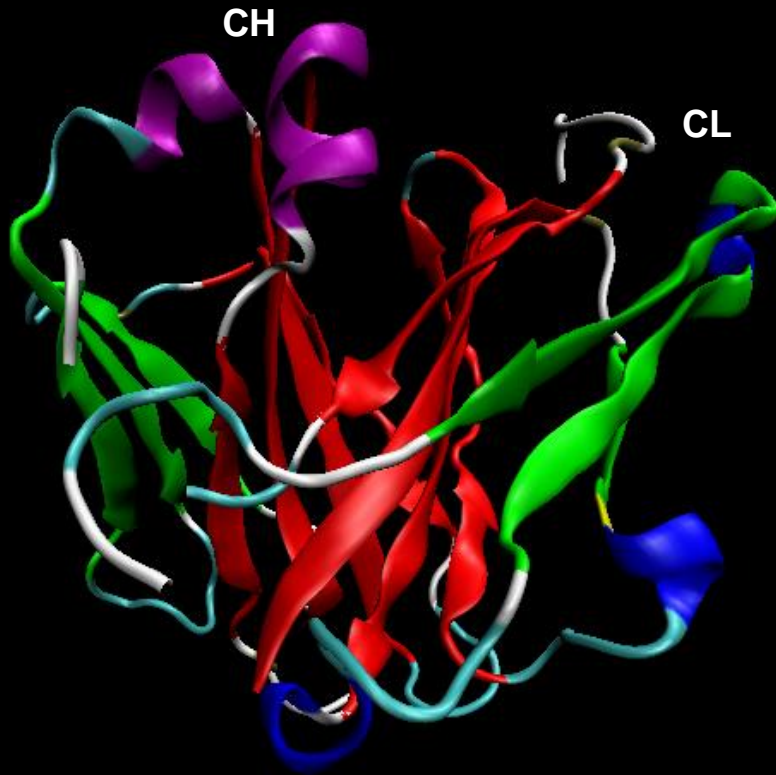


- 7 antiparallel β -strands
- 2 β -sheets:
 - 4-stranded sheet
 - 3-stranded sheet

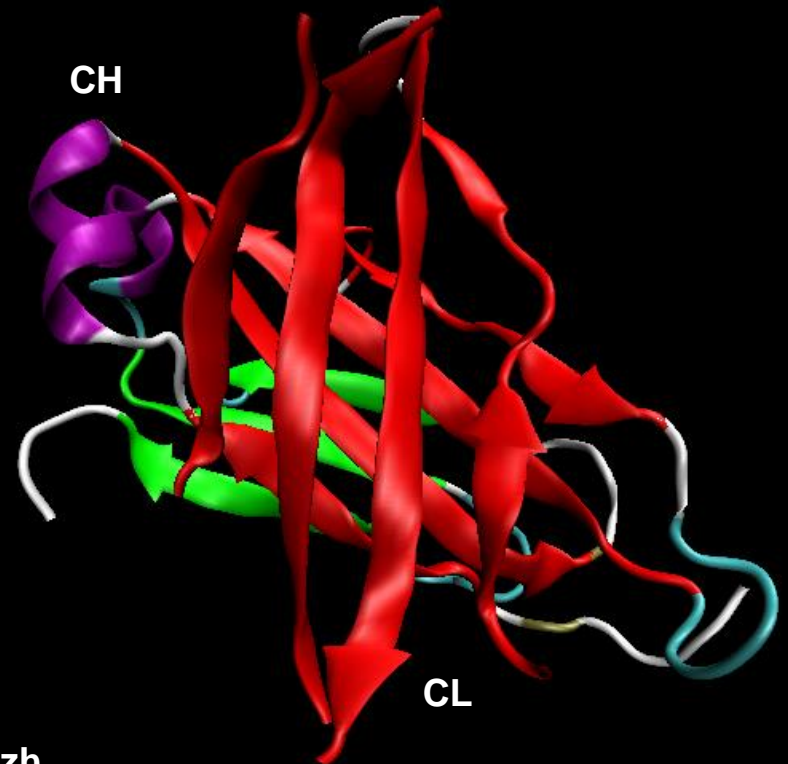
Immunoglobulin-like β -sandwich

Ig constant domain

4-stranded β -sheets interaction

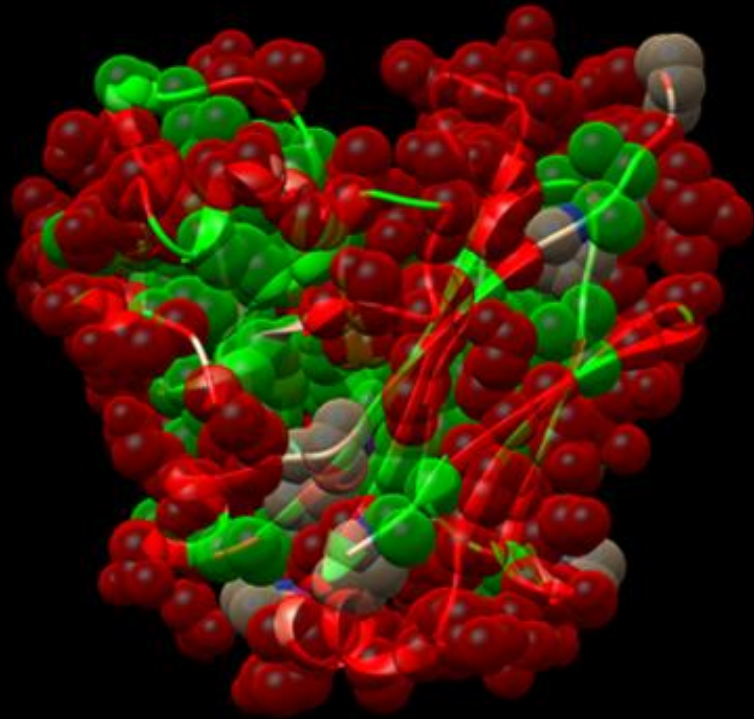


1hzh



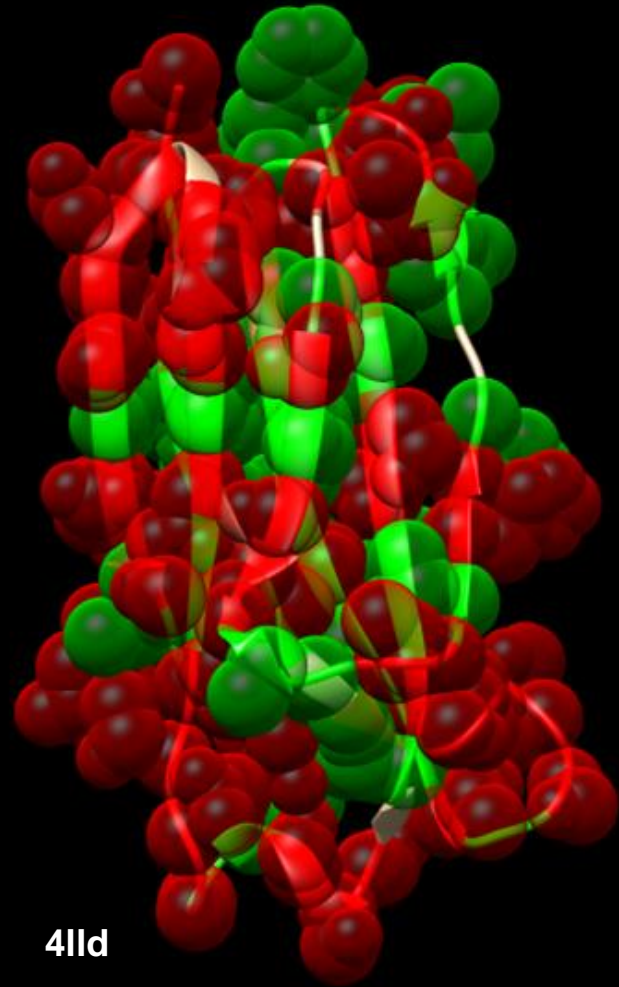
1hzh

Hydrophobic core



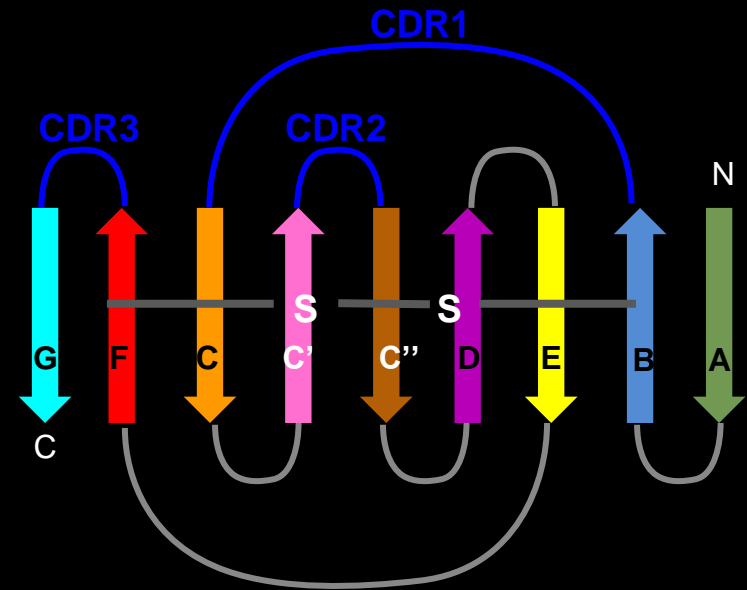
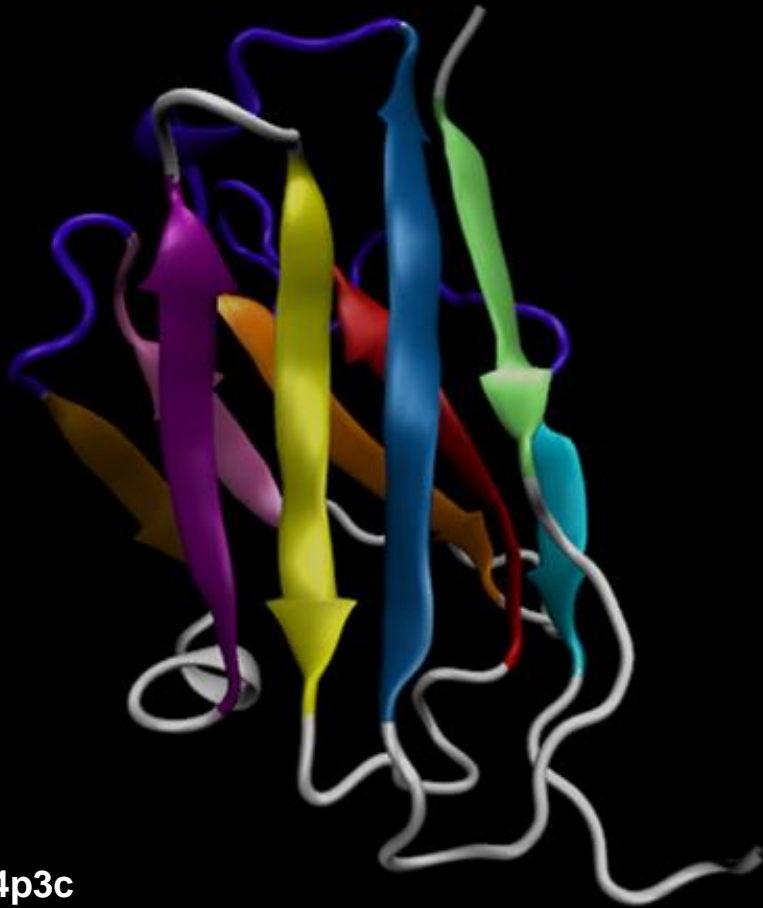
2wqr

● polar
● hydrophobic



4lld

Ig variable domain



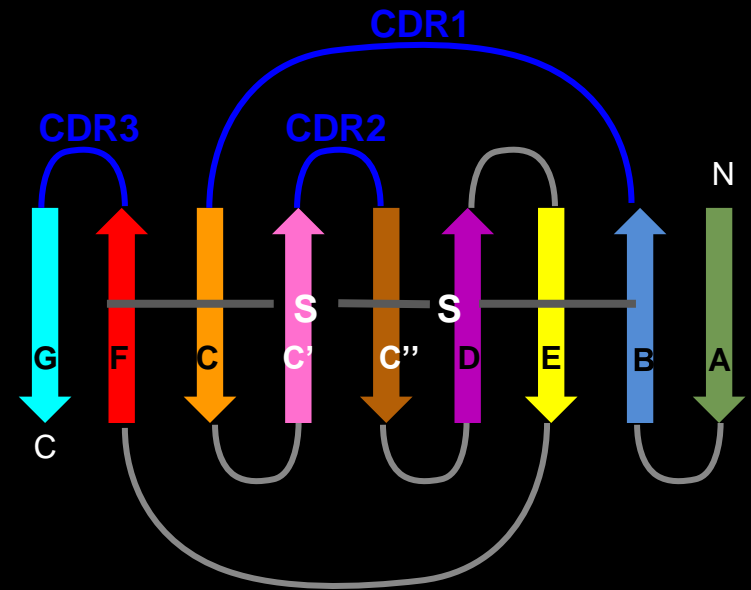
- 7 + 2 antiparallel β -strands: C' and C''
- 2 β -sheets:
 - 5-stranded sheet
 - 4-stranded sheet

Immunoglobulin-like β -sandwich

Ig variable domain



5ezj

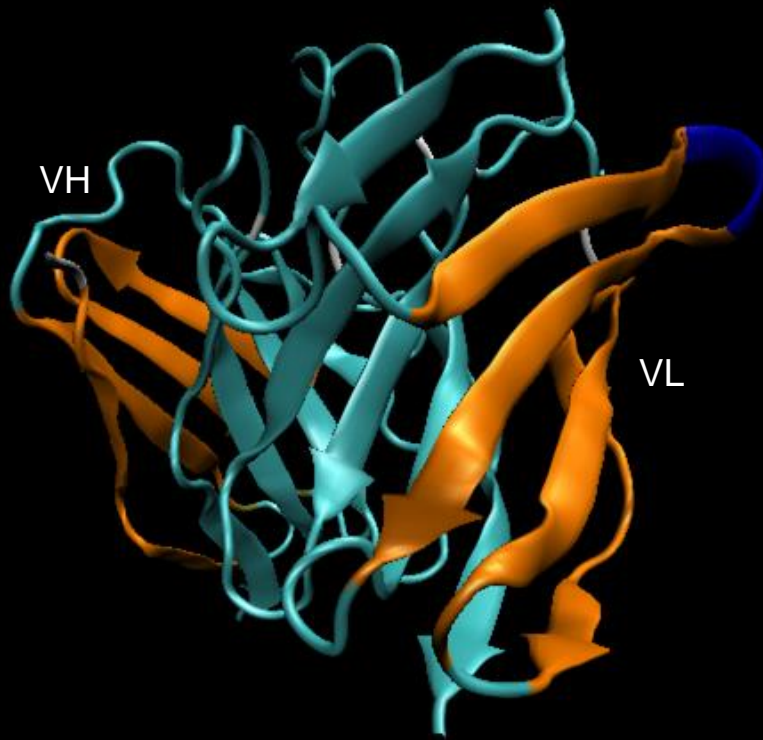


- 7 + 2 antiparallel β -strands: C' and C''
- 2 β -sheets:
 - 5-stranded sheet
 - 4-stranded sheet

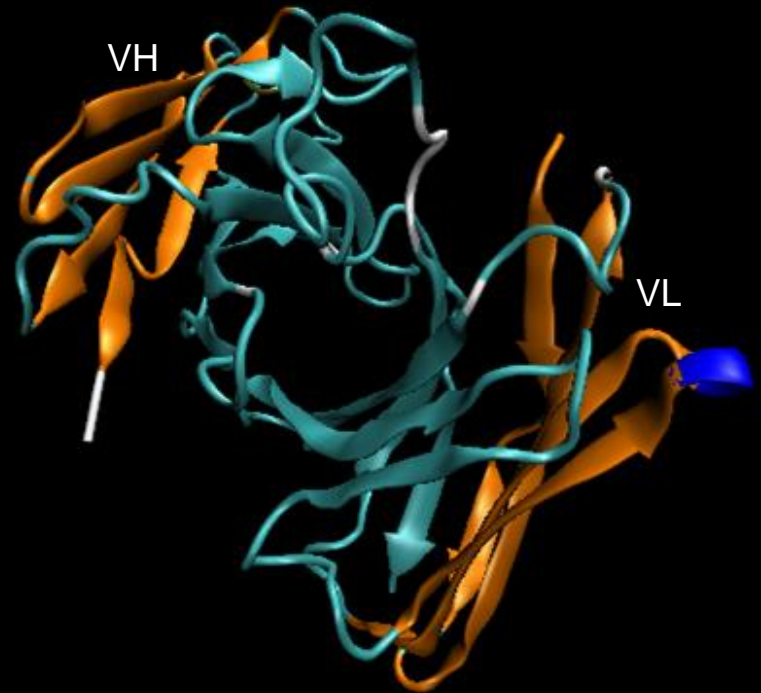
Immunoglobulin-like β -sandwich

Ig variable domain

5-stranded β -sheets \rightarrow β -barrel

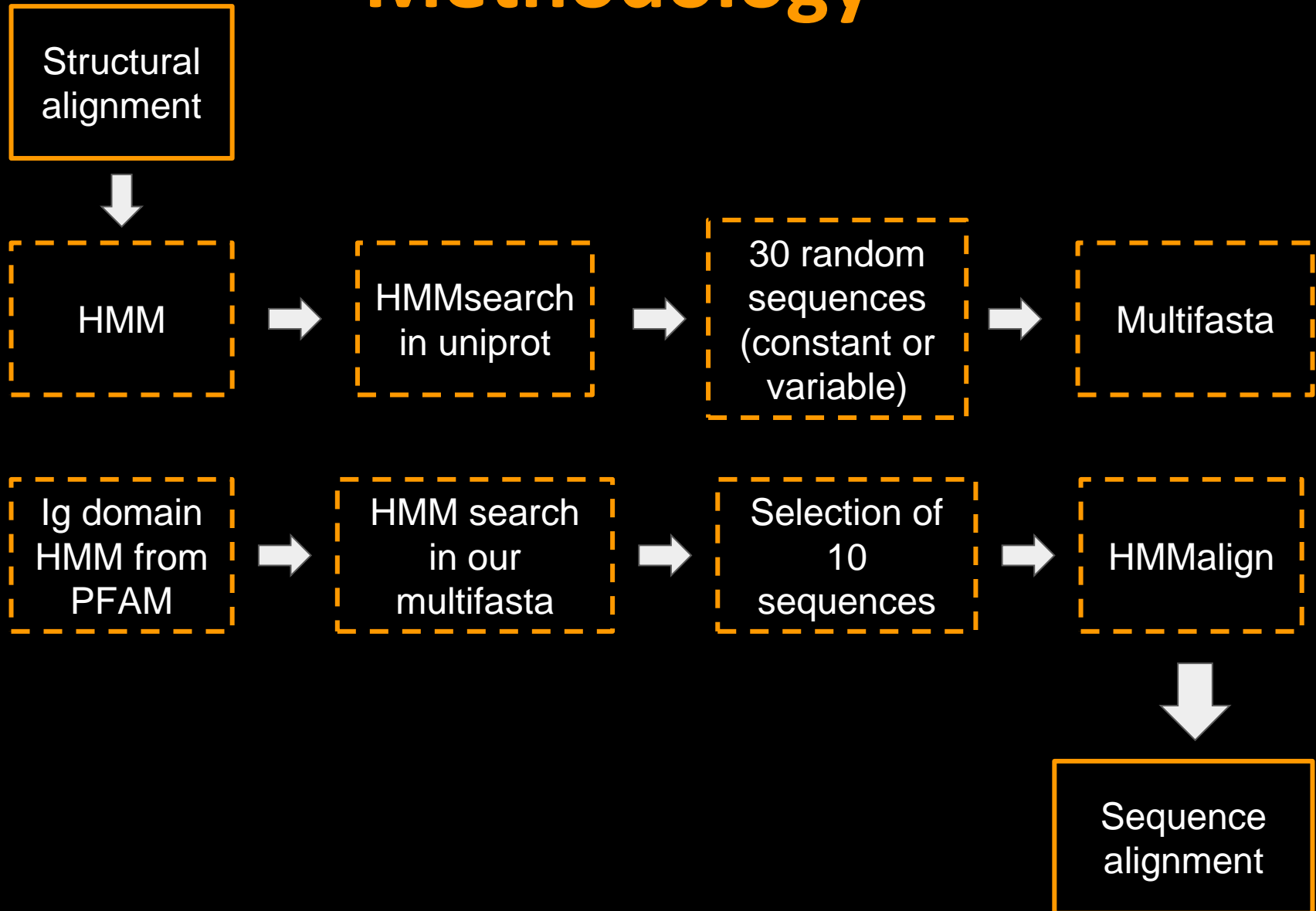


1hzh



1hzh

Methodology



Structural alignment - heavy chain

	A	B	C	C'					
4jo4 Oryctolagus cuniculus	-QSLEESGGDLVKPGASLTLTCTASGFS-FTNNY	YMCWVRQ	APGKGL	EWIACIYGGGRDI					
5ewi Homo sapiens	EVQLVESGGKLVQPGGSLRLSCEASGES-VGD-	NDMHWVRQ	VAG	KGLEWVSSIGS-S-GD					
5adp Homo sapiens	-VQLQESGPGLVKPSETLSLTCAVSGYS-ISSGY	YWGWIRO	PPGKGL	EWIGSIYH-S-GS					
1zan Rattus norvegicus	QVQLKESGPGLVQPSQTLTLTCTVSGFSLTN-N	NVNWVRQ	ATG	RGLEWMGGVWA-G-GA					
5dd3 Macaca mulatta	QVQLQESGPGLVKPSETLSVTCAVSGVSFSS-F	WWGWIRQ	SPGKG	LEWIGTIYG-SSGR					
	C''	D	E	CDR 1	F	CDR 2			
4jo4 Oryctolagus cuniculus	VFYATWAKGRFTISKTS-S-T	TVTLQM	TSLTAADT	ATYFCAREN--FD-AV	GVGGGT	YSTD			
5ewi Homo sapiens	TYVIDAVKGRFTVSRDKGRN	SVYLQM	KTLTVGDT	GVYFCVRGPE	SGW-F-----	YH			
5adp Homo sapiens	TYYNPSLKSRVTISVDTSKN	QFSLKL	SSVTAADT	AVYYCAGI	-----	---			
1zan Rattus norvegicus	TDYNSALKSR	TTITRD	TSKS	QVFLKM	HSLQSEDT	ATYYCARDG--GYSS-----S			
5dd3 Macaca mulatta	GEYNPSLKSR	TTISR-S---	QISLEI	TSVTAADT	AIYYCSRGL--F-----	---			
	G	A	CDR 3	B					
4jo4 Oryctolagus cuniculus	-----Y-YF-DLWGPGLVIVSSGQPKAPSVFPL	APCCGDT-PSA--	TVTLGC						
5ewi Homo sapiens	Y-----YWGL-GVWGRGTTVTVSSASTKGP	SVFPLAPS-----	SK-G	TAALGC					
5adp Homo sapiens	-----TANWGQGTLTTVSSASTKGP	SVFPLAPS-----	S-G	TAALGC					
1zan Rattus norvegicus	TL-----Y-AM-DAWGQGTTVTVSSASTTAP	SVYPLAPG----	S-----	MVTLCG					
5dd3 Macaca mulatta	-QPNGFSFTLTSY-WF-DVWGP	GVPTVSSASTKGP	SVFPLA-P-----	STSESTAALGC					
	B	CDR 3	C	D	E	F			
4jo4 Oryctolagus cuniculus	LVKGYLPEPV	TVTWN	NSGTLTNG	VRTFPSVR	QSSGL	YSLSSVSVTS---SSQP	VT	CNVA	
5ewi Homo sapiens	LVKDYFPEPV	TVSWN	SGALTSG	VHTFPAVL	QSSGL	YSLSSVVTVP	SSSL-GTQT	YI	CNVN
5adp Homo sapiens	LVKDYFPEPV	TVSWN	SGALTSG	VHTFPAVL	QSSGL	YSLSSVVTVP	SSSL-GTQT	YI	CNVN
1zan Rattus norvegicus	LVKGYFPEPV	TVTWN	NSGSLASG	VHTFPAVLQ	-SGL	LYTLSSSVTV	PASPWAS-EA	VT	CNVA
5dd3 Macaca mulatta	LVKDYFPEPV	TVSWN	SGSLTSG	VHTFPAVL	QSSGL	YSLSSVVTVP	SSSLGT-QT	YV	CNVN
	G								
4jo4 Oryctolagus cuniculus	HPATNTKVDKTV-AP---								
5ewi Homo sapiens	HKPSNTKVDKRV	EPK---							
5adp Homo sapiens	HKPSNTKVDKKV	EPKSCL							
1zan Rattus norvegicus	HPASSTKVDKKI	-VP-R-							
5dd3 Macaca mulatta	HKPSNTKVDKRV	-EI-K-							

Structural alignment - light chain

	A	B	C	C'		
1etz Mus Musculus	FAVVTQ--E-SALT	TSPGETVTLTCRS	STGAVTTSN-----YA	IWVQEKPDHLFSGLIIG		
5dd3 Macaca mulatta	DIQVTQ-SP-SSL	SASVGD	TVTISCRTS-QT-IS-----T-W	LAWYQVKPGKAPKLLIYT		
4jo4 Oryctolagus cuniculus	DIVMTQ-TPASV	SAAV-GG	TVTINCQA---S-ET--IS-N-Y	LAWYQQKPGQPPKLLIYK		
5i15 Homo sapiens	DIQMTQSPS-SLS	SASV-GDR	VTITCRA-----SQ---SISSY	LNWYQQKPGKAPKLLIYA		
5ihu Bos taurus	--VLNQ-PS-SV	SGSL-GQR	VSITCSG-----SSSNVGNG-Y	VSWEYQLIPGSAPRTLIIYG		
	C''	D	E	CDR 1	F	G
1etz Mus Musculus	TNNRVPGV	PARFSGSLIG-D	KAALTVTGA	QTEDE	AIYFCALWYS-----N-	H-WVFGGGT
5dd3 Macaca mulatta	AS	SL	ESGVPSRFSGSG-SGT	DFTLTI	STLQSEDF	ATYYCQYIS-----L-P-PTFGLGT
4jo4 Oryctolagus cuniculus	AS	TL	ASGVSSRFKSGSG-SGT	EYTLTI	SGVQCDDA	ATYYCQQGYS---ISDIDNSFGG-GT
5i15 Homo sapiens	AS	SL	QSGVPSRFSGSG-SGT	DFTLTI	SSLQPEDF	ATYYCQQSYS---T-PL--TFGQ-GT
5ihu Bos taurus	DT	SR	ASGVPDRFSGSR-SGN	TATLTI	SSLQAED	ADYFCASA-EDSSS-NA--VFGS-GT
	CDR 2	A	B	CDR 3	D	
1etz Mus Musculus	KLTVLG-QPKSSP	SVTLFTP	SSSELETN	KATLVCTITDF	YPGV	VTVDWKVDGTPVTQGME
5dd3 Macaca mulatta	KVEIK--RAVAAP	SVFIFPP	SEDQVKSG	TVSVVCLLNNE	YPRE	ASVKWKVDGVLKTGNSQ
4jo4 Oryctolagus cuniculus	EVVVK--GDPVAP	TVLIFPP	AADQVATG	TVTIVCVANKY	FPD-V	TVTWEVDGTTQTGTGIE
5i15 Homo sapiens	KVEIK--RTVAAP	SVFIFPP	SDEQLKSG	TASVVCLLNNE	YPRE	AKVQWKVDNALQSGNSQ
5ihu Bos taurus	TLTVL-GQPKSPP	SVTLFPP	STEELNGN	KATLVCLISDF	YPGSV	TVVWKADGSTITRNVE
	E	F	G			
1etz Mus Musculus	TTQPSKQS-NNKY	MASSYLTLT	ARAWERHS	SYSCQVTH	EGH----	TVEKSLSRAECS
5dd3 Macaca mulatta	ESVTEQDSKDNT	YSLSSTLTLS	NTDYQSHN	VYACEVTH	-Q-GLSSP	VTKSFNRGE--
4jo4 Oryctolagus cuniculus	NSKTPQNSADCT	YNLSSTLTLT	STQYNCHK	EYTCVKVTH	-GT---	TSVVQSFNRGDC-
5i15 Homo sapiens	ESVTEQDSKDS	YSLSSTLTLS	KADYEKHK	VYACEVTH	-QGLSSP	VTKSFNRGEC
5ihu Bos taurus	TTTRASKQS-NSKY	AASSYLSLT	SSDWKSKGS	YSCEVTH	-EG---	STVTKTVKPSEC-

Sequence alignment - CH1

```

sp|P01859|IGHG2_HUMAN  ---ASTKGP100%SV90-80%FLAP90-80%Csr-stSE-ST90-80%AAL90-80%GCLVKDYF-PEPVT100%VS90-80%W----N---S--GA
sp|P01860|IGHG3_HUMAN  ---ASTKGP100%SV90-80%FLAP90-80%Csr-stSG-GT90-80%AAL90-80%GCLVKDYF-PEPVT100%VS90-80%W----N---S--GA
sp|P20762|IGG2C_RAT    ---ARTTAP100%SV90-80%PLVP90-80%Gcs-gtSG-SL90-80%VT90-80%LGCLVKGYF-PEPVT100%VK90-80%W----N---S--GA
sp|P01868|IGHG1_MOUSE  ---AKTTP100%SV90-80%PLAP90-80%Gsa-aqTN-SM90-80%VT90-80%LGCLVKGYF-PEPVT100%VT90-80%W----N---S--GS
sp|P06337|IGHM_MESAU   --s--PSSPT100%VF90-80%PLVSCesplsDE-NL90-80%VAM90-80%GCLARDFL-PSSIS90-80%FS90-80%W----NyqnK--SE
sp|P03988|IGHM_RABIT   vsl---SSPT100%LY90-80%PLVSCeg-a90-80%lTDgNL90-80%VAM90-80%GCLARDFL-PSSVT90-80%FS90-80%WsfknN---S--EI
sp|P01871|IGHM_HUMAN   --g-SASAP100%TL90-80%FL90-80%LVSCenspsDT-SS90-80%VAV90-80%GCLAQDFL-PDSIT90-80%LS90-80%W----Ky-kN--NS
sp|P23087|HVCS_HETFR   -at--PSPPT100%LY90-80%GLCSCeq-pnTD-GS90-80%LAY90-80%GCLAMDYI-PQITS90-80%VS90-80%W----K---KdnEP
sp|P01872|MUCM_MOUSE    sqs----FPNV100%FL90-80%LVSCesplsDK-NL90-80%VAM90-80%GCLARDFL-PSTIS90-80%FT90-80%W----N---Y--QN
sp|P23735|MUCM ICTPU    -pk-----SLF100%PVWQCg--saSD-GL90-80%VT90-80%LGCVTRDLAS90-80%ADGLSFI90-80%W----Kd-aS--GS

sp|P01859|IGHG2_HUMAN  ---LTSGVHT100%FPAVLQSSGLYSLSS-VVTV90-80%PS90-80%SNF---GTQTY-T100%CN90-80%VDHKPSNTKVDKT
sp|P01860|IGHG3_HUMAN  ---LTSGVHT100%FPAVLQSSGLYSLSS-VVTV90-80%PS90-80%SSL---GTQTY-T100%CN90-80%VNHKPSNTKVDKR
sp|P20762|IGG2C_RAT    ---LSSGVHT100%FPAVLQ-SGLYTLSS-SVTV90-80%PS90-80%STW---SSQTV-T100%CS90-80%VAHPATKSNLIK90-80%R
sp|P01868|IGHG1_MOUSE  ---LSSGVHT100%FPAVLQ-SDLYTLSS-SVTV90-80%PS90-80%SPR---PSETV-T100%CN90-80%VAHPASSTKV90-80%DKK
sp|P06337|IGHM_MESAU   ---VNQGVRT100%FPTLRM-GEKYAATS-QVFL90-80%PPKSV90-80%legSDEYL-VCK90-80%VHHGNTNK-----
sp|P03988|IGHM_RABIT   ---SSRTVRT100%FPVVKR-GDKYMATS-QVLV90-80%PSKDV90-80%l-qGTEEY90-80%LVCK90-80%VQHSNSNR-----
sp|P01871|IGHM_HUMAN   ---DISSTRGF100%PSVLR-GGKYAATS-QVLL90-80%PSKDV90-80%m90-80%qgTDEHV-VCK90-80%VQHPNGNK-----
sp|P23087|HVCS_HETFR   ---ITTGLKTY100%PSVLNKKGT90-80%YTQSS-QLTITESEV---GSSKI-YCEVRR-----
sp|P01872|MUCM_MOUSE    nteVIQGI100%RT90-80%FPTLRT-GGKYLAT90-80%sqVLLSPKSI90-80%Le-gSDEYL-VCK90-80%IHYGG-----
sp|P23735|MUCM ICTPU    ---ALTDVVQY100%PAVQ-ATGGYTSVS-HVRVKASDW--nGNKKF-TCEVKN-----

```



100% conserved residues



90-80% conserved residues belonging to different groups



90-80% conserved residues belonging to the same group

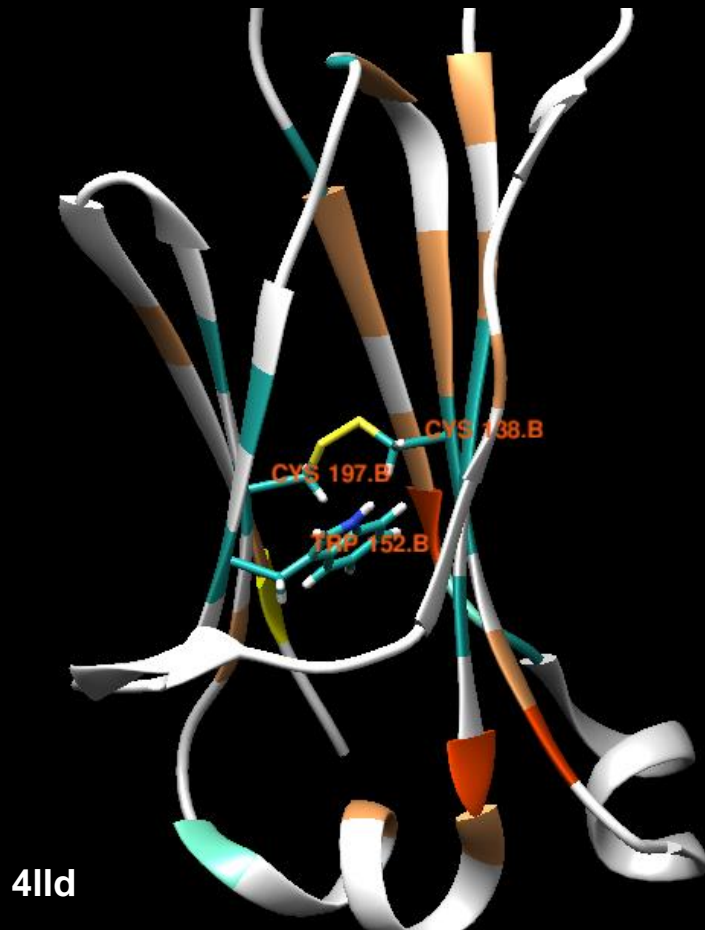


residues from the same group



hydrophobic residues (aromatic and none aromatic)

Sequence alignment constant domain



Cys - Trp - Cys triad








Conserved Valines, Isoleucines
and Leucines

Sequence alignment - CL

sp P0CF74 LAC6_HUMAN	-gQPKAAPSVTLFPPSSEELQ-ANKATLVCLISDFYPGAVKVAWKADGSPVNTGVETITP
sp P01846 LAC_PIG	--QPKAAPTVNLFPPSSEELG-TNKATLVCLISDFYPGAVTVTWKAGGTTVTQGVETTKP
sp P20763 LAC_CHICK	--QPKVAPTITLFPFSKEELNeATKATLVCLINDFYPSPTVDWVIDGSTRSG--ETTAP
sp P01844 LAC2_MOUSE	--QPKSTPTLTVFPFSSEELK-ENKATLVCLISNFSPPSGVTVAWKANGTPTITQGVDTSNP
sp P01835 KACB_RAT	-a--DAAPTVSIFPPSTEQLA-TGGASVVCMLNNFYPRDISVKWKIDGTERRDGVLDSTV
sp P01839 KACB_RABIT	-g-DPVAPSVLLFPFSKEELT-TGTATIVCVANKFYPSDITVTWKVDGTTQQSGIENSKT
sp P01834 IGKC_HUMAN	-t--VAAPSVFIFPPSDEQLK-SGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVT
sp P01840 KAC4_RABIT	-d--PVAPTVLIFPPAADQVA-TGTVTIVCVANKYFPD-VTVTWEVDGTTQTTGIENSKT
sp P01841 KAC5_RABIT	at---LAPTVLIFPPAPAQLA-TGAVTIVCVANKYFPD-GTVTWEVDGKPLTTGIETSKT
sp P11272 IGKC_LITCT	rg-ENVRPTVSIYCPISLEQRN-SGSASTVCLVDKFYPPGGAQVTWKGDNKVISSGVDTSDK

sp P0CF74 LAC6_HUMAN	SKQS-NNKYAASSYLSLTPEQWKS <hr/> SHRSYSCQVTHEG--STVEKTVAPAECS-
sp P01846 LAC_PIG	SKQS-NNKYAASSYLALSASDWKSSSGFTCQVTHEG--TIVEKTVTPSECa-
sp P20763 LAC_CHICK	QRQS-NSQYMASSYLSLSASDWSSHETYT <hr/> CRVTHNG--TSITKTLKRSEC--
sp P01844 LAC2_MOUSE	TKEG--NKFMASFFLHLSQDQWRSHNSFTCQVTHEG--DTVEKSLSPAEC-
sp P01835 KACB_RAT	DQDSKDSTYSMSSTLSLTKADYESHNLTYCEVVHKTssSPVVKSFNRNEC--
sp P01839 KACB_RABIT	PQSPEDNTYSLSSSTLSLTSAQYNSSSVYTCEVVQGS-aSPIVQSFNRGDC--
sp P01834 IGKC_HUMAN	EQDSKDSTYSLSSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC--
sp P01840 KAC4_RABIT	PQNSADCTYNLSSSTLTLTSTQYNSSHKEYTCKVTQGT--TSVVQSFNRGDC--
sp P01841 KAC5_RABIT	PQNSDDCTYNLSSSTLTLLQSNYNSSHNEYTCQVAQGA--GSVVQSFNRKNC--
sp P11272 IGKC_LITCT	IKDK-DNTYSMSSTLTMSSEEFKY-STMTCEVTHPTLTpALAKSFQTSECTf

-  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 - VH

```

sp|P01806|HVM36_MOUSE  mdfgliffivallkgvqce-VKLLESGGGLVQPGGSLKLSCAASGFDFSSR-YWMSWVRQA
sp|P01823|HVM47_MOUSE  -----EVQLQESSGPSLVKPSQTLSLTCSVTGSSITS-DYWNWVIRKF
sp|P18531|HVM60_MOUSE  mkvlslllylltaipgilsd-VQLQESSGPGLVKPSQSLSLTCSVTGSSITSgYYWNWVIRQF
sp|P83907|IGW3_HETFR   -----nivltqpe-----SAVKKPGESHKLSCTVSGFDVNG-HHMNWVKQV
sp|P18528|HVM57_MOUSE  -----EVQLVESSGGGLVKPGGSLKLSCAASGFTFSSD-YYMYWVRQT
sp|P01793|HVM24_MOUSE  -----EVKLVEESGGGLVQPGGSLRLSCATSGSFTLSD-FYMEWVRQT
sp|P01803|HVM34_MOUSE  -----EVKLEESSGGGLVQPGGRSMKLSCVASGSFTFSN-YWMNWVRQS
sp|P01789|HVM20_MOUSE  -----EVKLVEESGGGLVQPGGSLRLSCATSGSFTFSN-FYMEWVRQP
sp|P01785|HV02_CANLF   -----EVKLVEESGGDLVKPGGSLRLSCVASGSFTFSS-NGMSWVRQD
sp|P01746|HVM02_MOUSE  mgwsfiflflsvtagvhseVQLQESSGAELVRAGSSVKMSCKASGYTFSTSS-YGINWVKQR

```

CDR 1

```

sp|P01806|HVM36_MOUSE  PGKGLEWIGGEINPdS---STINYTPSLKDKFTISSRNAKNTLYLQMSKVRSEDTALYYCA
sp|P01823|HVM47_MOUSE  PGNKLEHMGYISY-S---GSTYYNPSLKSRISITRDTSKNQYYLQLNSVTSEDTATYYCT
sp|P18531|HVM60_MOUSE  PGNKLEWMGYISY-D---GSNNYNPSLKNRISITRDTSKNQFFLKLNSVTTEDTATYYCA
sp|P83907|IGW3_HETFR   PGEGLEWLLLSYRK-T---YNTYYASGIQGRITFSTESS--TFIEIPNLRVEDTAMYYCA
sp|P18528|HVM57_MOUSE  PEKRLEWVATISD-G--gSYTYYPDSVKGRFITSRNAKNNLYLQMSSLKSEDTAMYYCA
sp|P01793|HVM24_MOUSE  PGKRLEWIAASRN-KvydYTTEYSASVKGRFIVSRDTSQSILYLQMNALRAEDTAIYYCA
sp|P01803|HVM34_MOUSE  PEKGLEWVAEIRLkSh-nYAIHYAESVKGRFITSRDTSQSILYLQMNALRAEDTGIYYCS
sp|P01789|HVM20_MOUSE  PGKRLEWIAASRN-KgnkYTTEYSASVKGRFIVSRDTSQSILYLQMNALRAEDTAIYYCA
sp|P01785|HV02_CANLF   PGEGLQWVADISS-S---GQTYYADAVKGRFSISRONAKNTLYLQMEDLRVEDTAVYYCA
sp|P01746|HVM02_MOUSE  PGQGLEWIGYINP-G--nGYINYNEKFKGKTTLTVOKSSSTAYMQLRSLTSEDSAVYFCA

```

CDR 2

```

sp|P01806|HVM36_MOUSE  R-----
sp|P01823|HVM47_MOUSE  SLRF-----AYWGQGTLTVTVS--a-
sp|P18531|HVM60_MOUSE  R-----
sp|P83907|IGW3_HETFR   RGTGF-PQ--WGYWGSGTFLTVTSVTq
sp|P18528|HVM57_MOUSE  R-----
sp|P01793|HVM24_MOUSE  RDAYY-GSYwYFDVWAGTTVVSS--
sp|P01803|HVM34_MOUSE  TG-----FPSWGPGTLTTV----
sp|P01789|HVM20_MOUSE  RNYYG-STW-YFDVWAGTTVVSS--
sp|P01785|HV02_CANLF   TEGD--I-E-IPRYFGQGTIVVSS--
sp|P01746|HVM02_MOUSE  RSHYYgGSY-DFDYWQGTPLVSS--

```

CDR 3

100% conserved residues

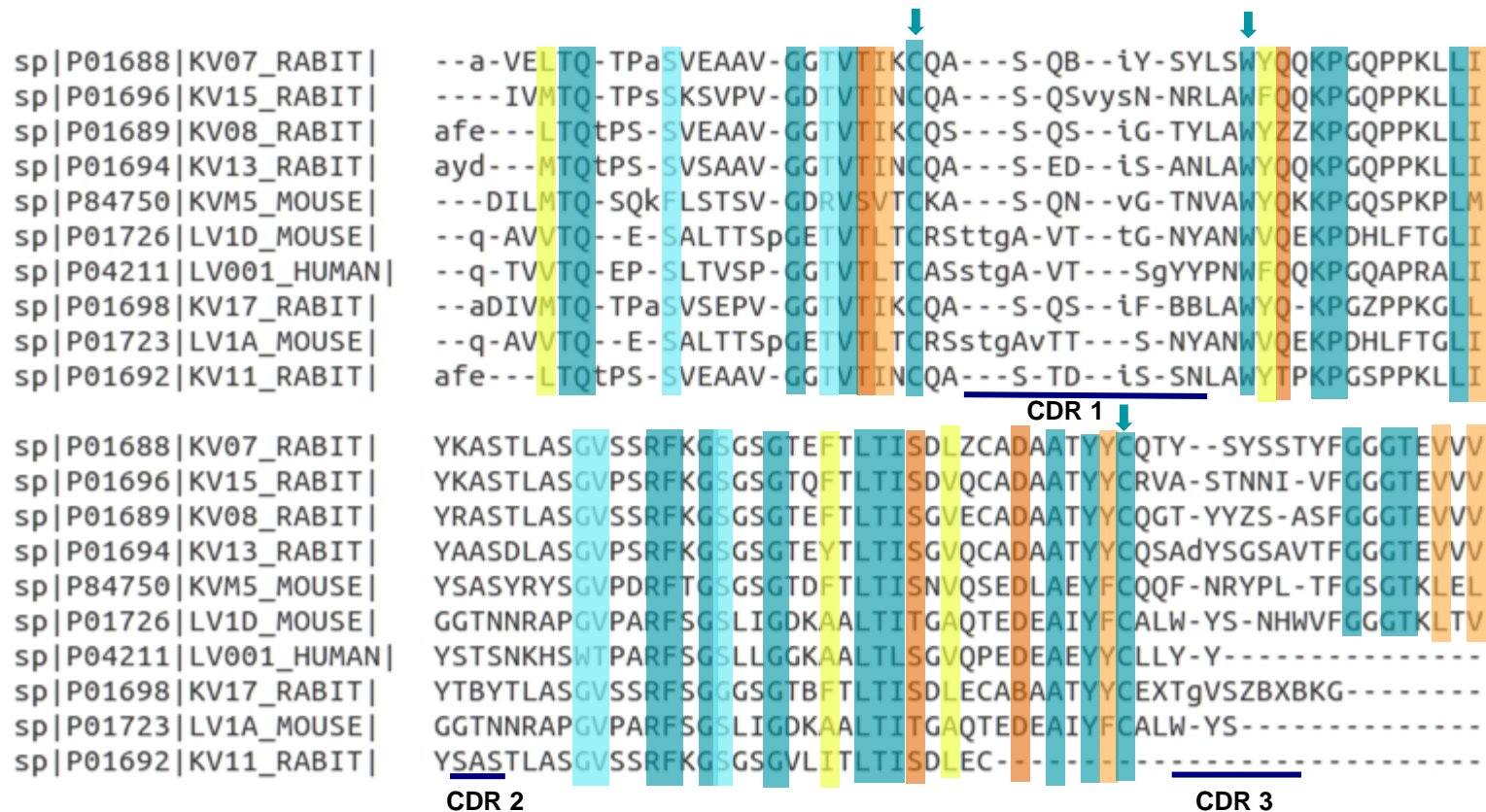
90-80% conserved residues belonging to different groups

90-80% conserved residues belonging to the same group

residues from the same group

hydrophobic residues (aromatic and none aromatic)

Sequence alignment - 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)

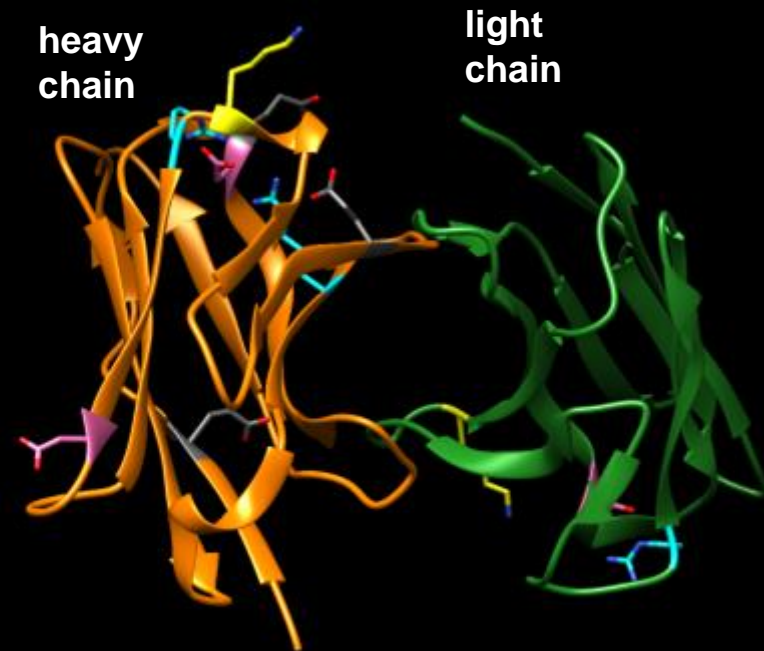
Additional conservations - charged residues

Alignment from the heavy chain:

```

mdfgliffivallkgvqce-VKLL-ESGGGLVQPGGSLKLSCAASGFDFSR-YWMSWVRQA
-----EVQLQ-ESGPSLVKPSQTLSTCSVTGDSITS-DYWNWIRKF
mkvlslllylltaipgilsd-VQLQ-ESGPGLVKPSQSLSTCSVTGYSITSgYYWNWIRQF
-----nivltqpe-----SAVKKPGESHKLCTVSGFDVNG-HHMNWVKQV
-----EVQLVESGGGLVKPGGSLKLSCAASGFTFSD-YYMYWVRQT
-----EVKLVESGGGLVQPGGSLRLSCATSGFTLSD-FYMEWVRQT
-----EVKLEESGGGLVQPGRSMKLSCVASGFTFSN-YWMNWVRQS
-----EVKLVESGGGLVQPGGSLRLSCATSGFTFSD-FYMEWVRQP
-----EVKLVESGGDLVKPGGSLRLSCVASGFTFSS-NGMSWVRQD
mgwsfiflflsvtagvhsEVQLQ-ESGAELVRAGSSVKMSCKASGYTFTS-YGINWVKQR

PGKGLEWIGEINPdS---STINYTPSLKDKFIISRDNAKNTLYLQMSKVRSEDATALLYCA
PGNKLEHMGYISY-S---GSTYYNPSLKSRIISITRDTSKNQYYLQLNSVTSEDATATYYCT
PGNKLEWMGYISY-D---GSNNYNPSLKNRISITRDTSKNQFFLKLNSVTTEDATATYYCA
PGEGLEWLLSYRK-T---YNTYYASGIQGRITFSTESS--TTFIEIPNLRVEDATAMYYCA
PEKRLWVATISD-G--gSYTYYPDSVKGRFTISRDNKNNLYLQMSSLKSEDATAMYYCA
PGKRLWIAASRN-KvydYTTEYSASVKGRFIVSRDTSQSILYLQMNALRAEDTAIYYCA
PEKGLEWVAEIRLkSh-nYAIHYAESVKGRFTISRDDSKSSVYLQMNALRAEDTGIYYCS
PGKRLWIAASRN-KgnKYTTEYSASVKGRFIVSRDTSQSILYLQMNALRAEDTAIYYCA
PGEGLQWVADISS-S---GQTYADAVKGRFSISRDNKNTLYLQMEDLRVEDTAVYYCA
PGQGLEWIGYINP-G--nGYINYNEKFKGKTTTLTVKSSSTAYMQLRSLTSEDASVYFCA
    
```



- 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

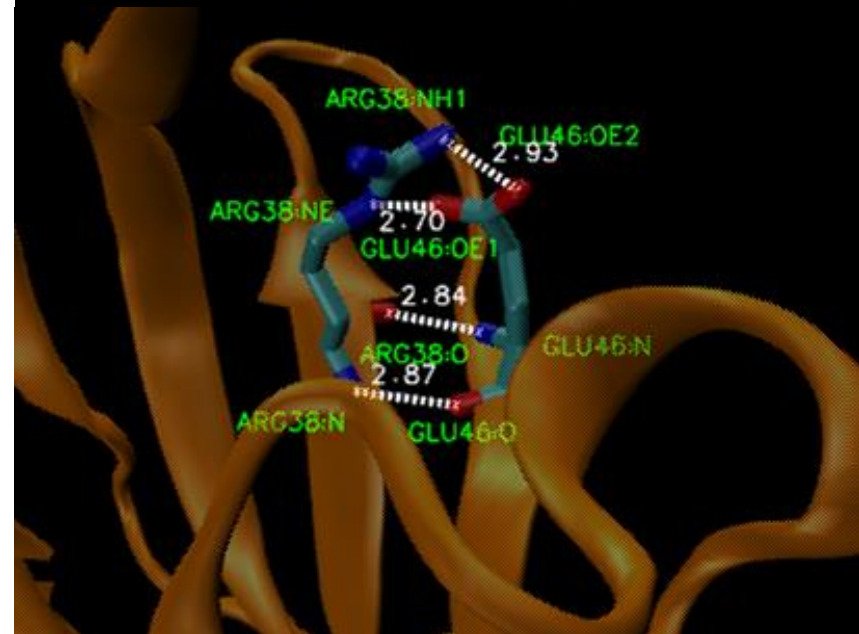
Arg 38

```
mdfgliffivallkgvqce-VKLLESGGGLVQPGGSLKLSAASGFDfsr-YWMSWVRQA
-----EVQLQESGpsLVKPSQTLsLTCSVTGDSITS-DYWNWIRKF
mkvlslllylltaipgilsd-VQLQESGPGLVKPSQSLsLTCSVTGYSITsGYYWNWIRQF
-----nivltqpe-----SAVKKPGESHKLCTVSGFDVNG-HHMNWVKQV
-----EVQLVESGGGLVKPGGSLKLSAASGFTFSd-YMYWVRQT
-----EVKLVESGGGLVQPGGSLRLSCATSGFTLSd-FYMEWVRQT
-----EVKLEESGGGLVQPGRSMKLSCVASGFTFSN-YWMNWVRQS
-----EVKLVESGGGLVQPGGSLRLSCATSGFTFSd-FYMEWVRQP
-----EVKLVESGGDLVKPGGSLRLSCVASGFTFSs-NGMSWVRQD
mgwsfiflflsvtagvhsEVQLQQSGAELVRAGSSVKMSCKASGYTFTS-YGINWVKQR
```

```
PGKGLEWIGEINpds---STINYTPSLKDKFIISRDNAKNTLYLQMSKVRSEDtALYYCA
PGNKLEHMGYISY-S---GSTYYNPslKSRISITRDTSKNQYYLQLNSVTSEDtATYYCT
PGNKLEWMGYISY-D---GSNNYNPSLKNRISITRDTSKNQFFLKLNSVTTEDtATYYCA
PGEGLWLLSYRK-T---YNTYYASGIQGRITFSTESS--TTFIEIPNLRVEDtAMYYCA
PEKRLWVATISD-G--gSYTYYPDSVKGRFTISRDNAKNNLYLQMSSLKSEdTAIYYCA
PGKRLWIAASRN-KvydYTTEYSASVKGRFIVSRDTSQSILYLQMNALRAEDtAIYYCA
PEKGLEWVAEIRLksh-nYAIHYAESVKGRFTISRDDSKSSVYLQMNNLRAEDtGIYYCS
PGKRLWIAASRN-KgnkYTTEYSASVKGRFIVSRDTSQSILYLQMNALRAEDtAIYYCA
PGEGLQWVADISS-S---GQTYADAVKGRFSISRDNakNTLYLQMEDLRVEDtAVYYCA
PGQGLEWIGYINP-G--nGYINYNEKFKGKTTLTVDKSSSTAYMQLRSLTSEDsAVYFCA
```

Glu 46

heavy chain



5mvz

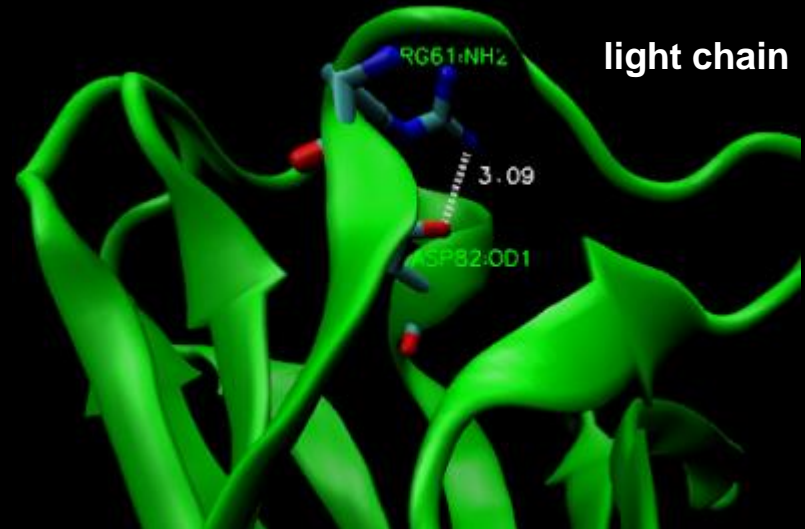
Additional conservations - charged residues

Salt bridge between charged residues in the light chain

```
YKASTLASGVSSRFKSGSGSGTEFTLTISDLZCADAATYYCQTY
YKASTLASGVPSRFKSGSGSGTQFTLTISDVQCA DAATYYCRVA
YRASTLASGVSSRFKSGSGSGTEFTLTISGVECA DAATYYCQGT
YAASDLASGVPSRFKSGSGSGTEYTLTISGVQCA DAATYYCQSA
YSASYRYSGVDPDRFTGSGSGTDFTLTISNVQSEDLAEYFCQQF
GGTNNRAPGVPARFSGSLIGDKAALTITGAQTEDEAIYFCALW
YSTSNKHSWTPARFSGSLLGGKAALTLSGVQPEDEAEYYCLLY
YTBYTLASGVSSRFSGGSGTBFTLTISDLECA BAATYYCEXT
GGTNNRAPGVPARFSGSLIGDKAALTITGAQTEDEAIYFCALW
YSASTLASGVSSRFKSGSGVLITLTISDLEC-----
```

Arg 61

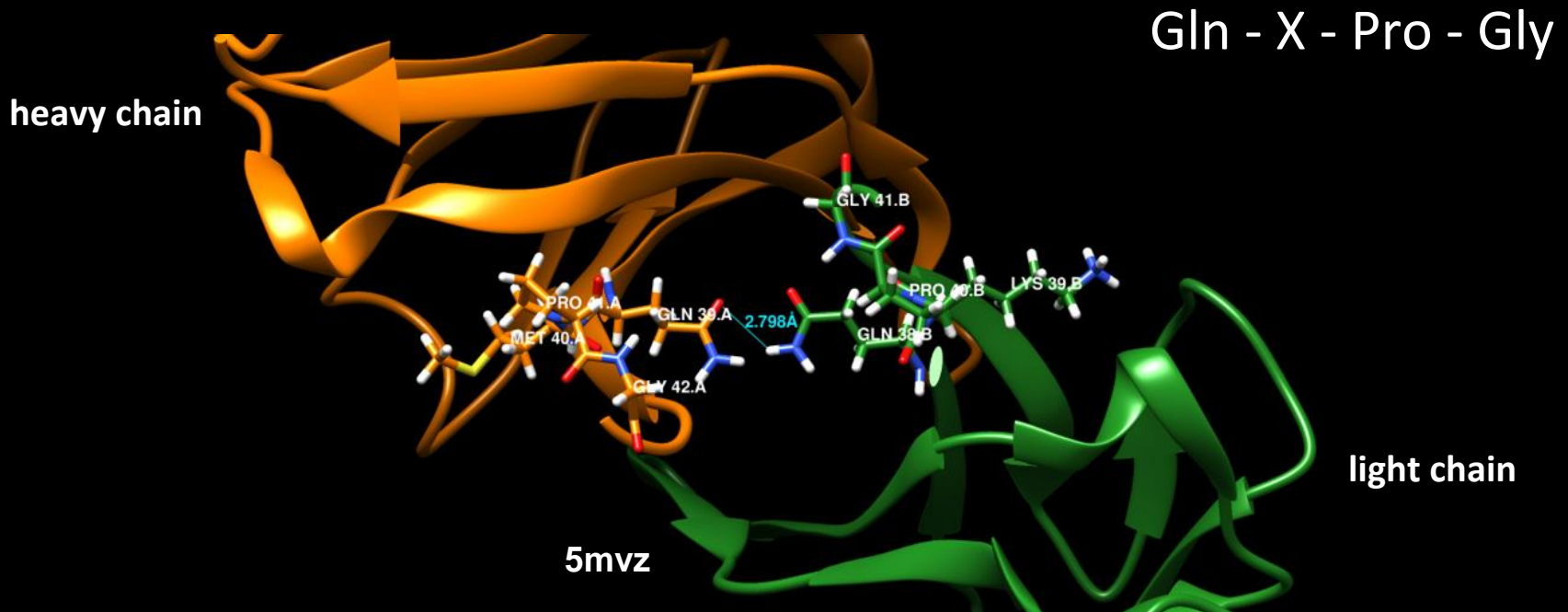
Asp 82



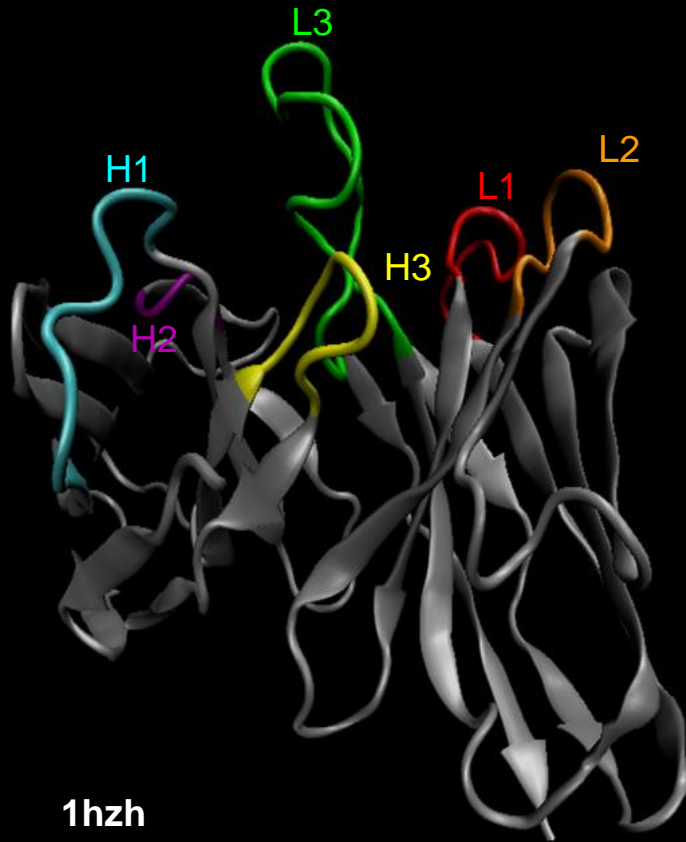
5mvz

Additional conservations - polar residues

sp P01806 HVM36_MOUSE	mdfgliffivallkgvqce-VKLEESGGGLVQPGGSLKLSCAASGDFFSR-YWMSWVRQAPGKGLEWIGEINPdS
sp P01823 HVM47_MOUSE	-----EVQLQESGPSLVKPSQTLSTCSVTGDSITS-DYWNWIRKFPGNKLEHMGYISY-S
sp P18531 HVM60_MOUSE	mkvlslllylltaipgilsd-VQLQESGPGLVKPSQSLSLTCSVTGYSITSgYYWNWIRQFPGNKLEWMGYISY-D
sp P83907 IGW3_HETFR	-----nivltqpe-----SAVKKPGESHKLCTVSGFDVNG-HHMNWVKQVPGEGLEWLLSYRK-T
sp P18528 HVM57_MOUSE	-----EVQLVESGGGLVKPGGSLKLSCAASGFTFSY-YYMYWVRQTPEKRLEWVATISD-G
sp P01793 HVM24_MOUSE	-----EVKLVESGGGLVQPGGSLRLSCATSGFTLSD-FYMEWVRQT PGKRLEWIAASRN-KV
sp P01803 HVM34_MOUSE	-----EVKLEESGGGLVQPGSRMKLSCVASGFTFSN-YWMNWVRQS PEKGLEWVAEIRLkSH
sp P01789 HVM20_MOUSE	-----EVKLVESGGGLVQPGGSLRLSCATSGFTFSY-FYMEWVRQP PGKRLEWIAASRN-KQ
sp P01785 HV02_CANLF	-----EVKLVESGGDLVKPGGSLRLSCVASGFTFSS-NGMSWVRQD PGEGQLQWVADISS-S
sp P01746 HVM02_MOUSE	mgwsfiflflsvtagvhsEVQLQQSGAELVRAGSSVKMSCKASGYTFTS-YGINWVKQR PGQGLEWIGYINP-G



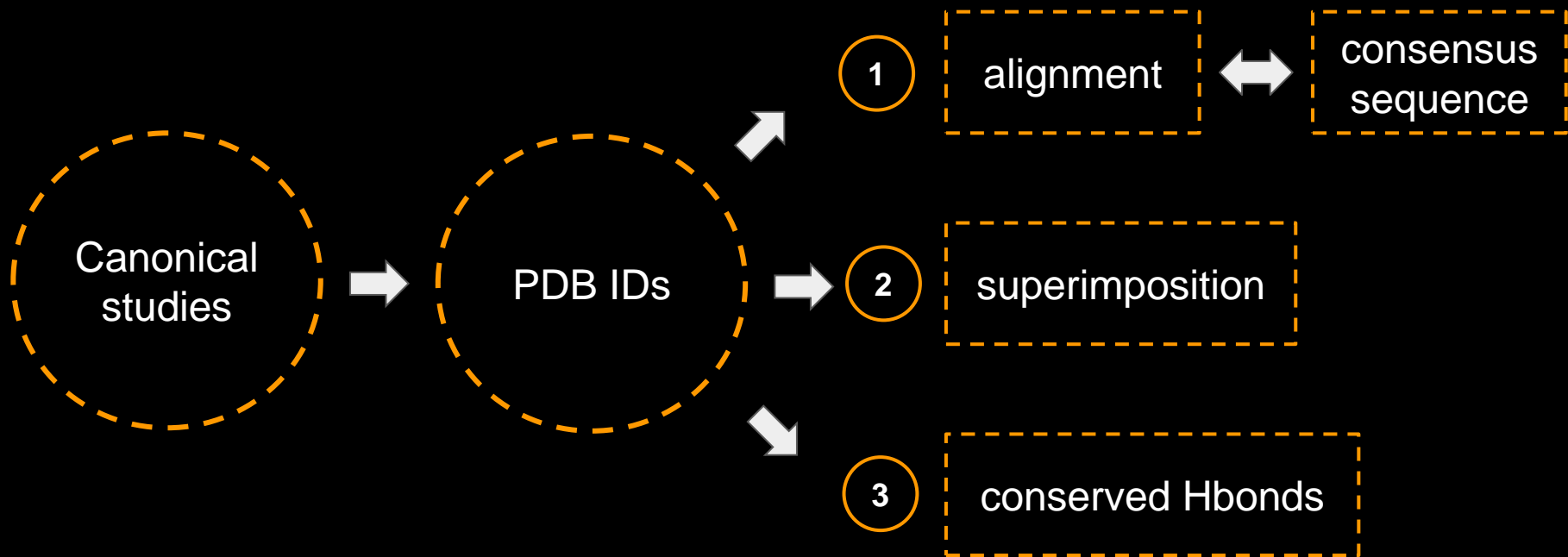
Canonical structures



Chothia et al classification:

LIGHT	L1	\mathcal{K}	1	2a	2b	3	4	5	6
		λ	1	2	3a	3b	4		
	L2		1						
HEAVY	L3	\mathcal{K}	1	2	3	4	5		
		λ	1a	1b	1c	2			
	H1		1	2	3				
	H2		1	2a	2b	2c	3a	3b	3c 4

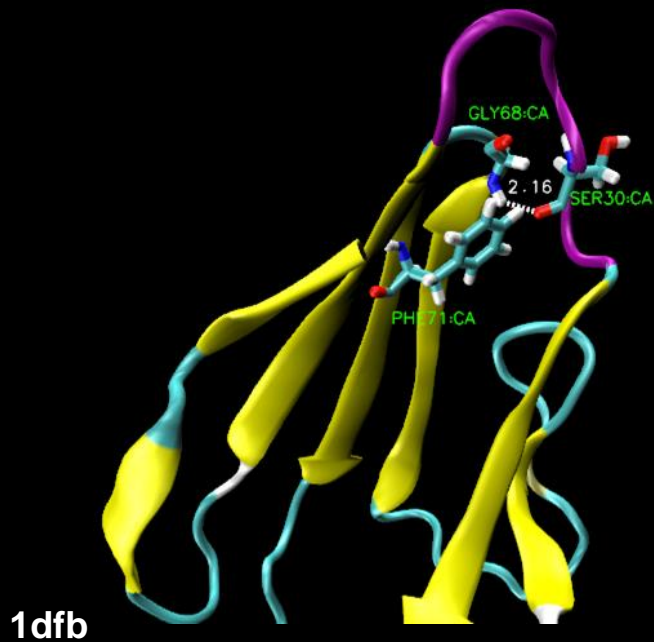
Canonical structures - Methodology



L1 - k2

L1: 24-34

L1-k2a

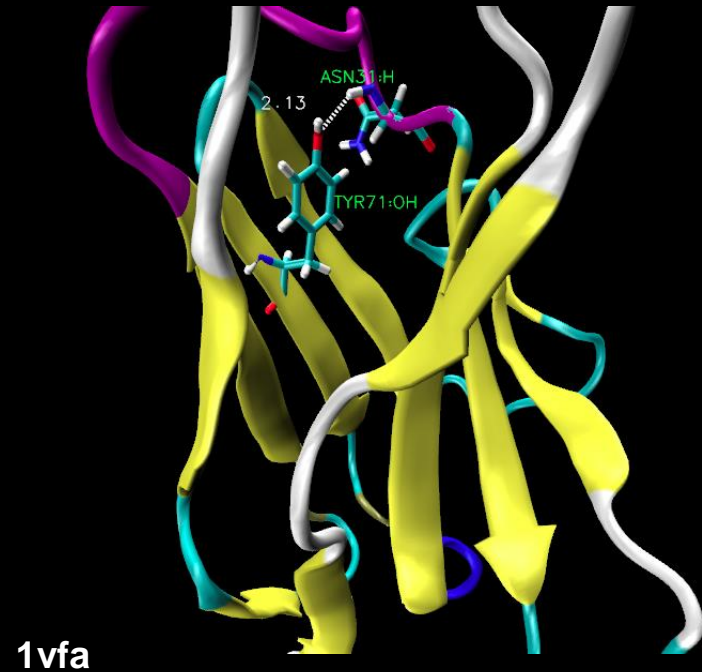


Position 71: Phe, Thr or Gly

Hbond:

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

L1-k2b



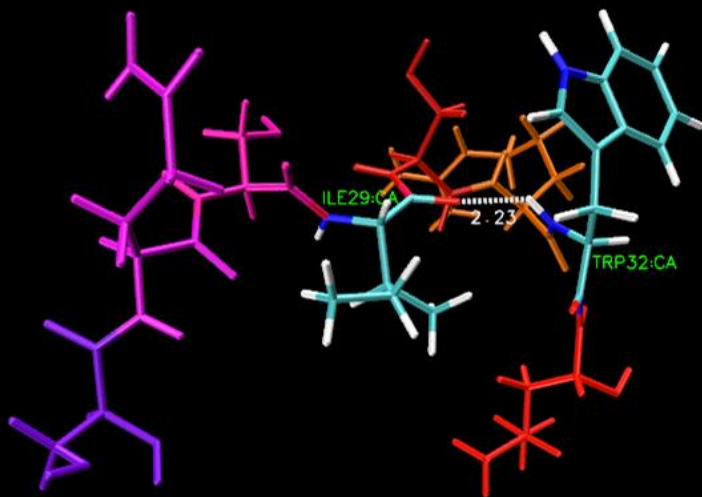
Position 71: Tyr

Hbond:

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

L1 - k2

L1-k2a



1dfb

Hbond:

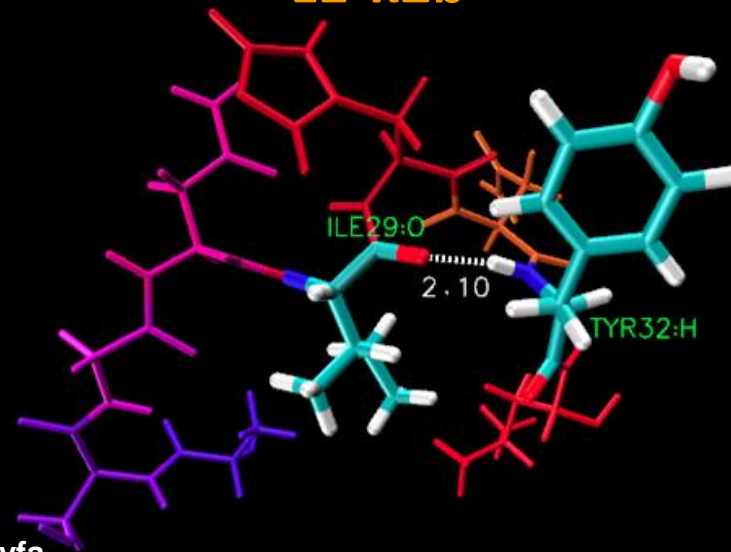
Ile 29 (O) --- Trp 32 (N)

rASQdisnyla

1dfb	RASQ	SISRWLAWYQQKPG
1igm	QASQ	DISNYLAWYQQKPG
1ivl	RASQ	SIGNRLFYQQKSH
1fvc	RASQ	DVNTAVAWYQQKPG

RMS = 0,62

L1-k2b



1vfa

Hbond:

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

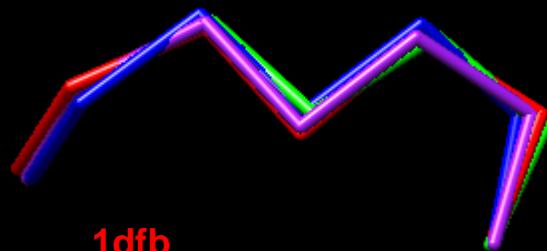
rASqdisnyLa

1zan	CRAS	EDIYNALAW
1vfa	CRAS	GNIHNYLAW
1fgv	CRAS	QDINNYLNW

RMS = 0,51

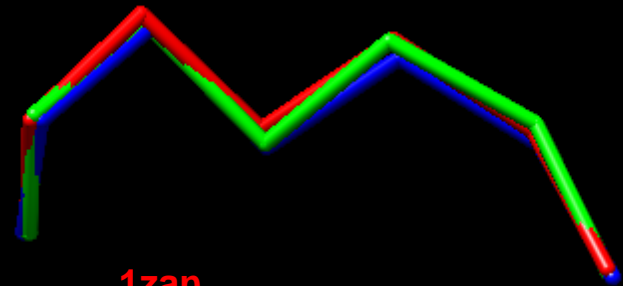
L1 - k2

L1-k2a

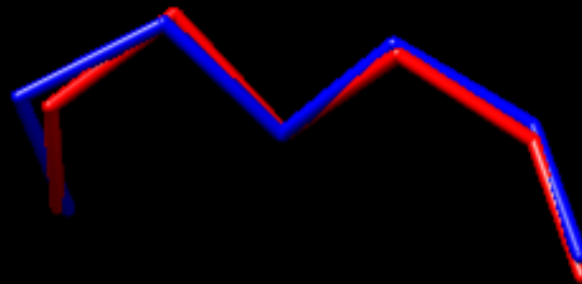


1dfb
1igm
1ivl
7fvc

L1-k2b



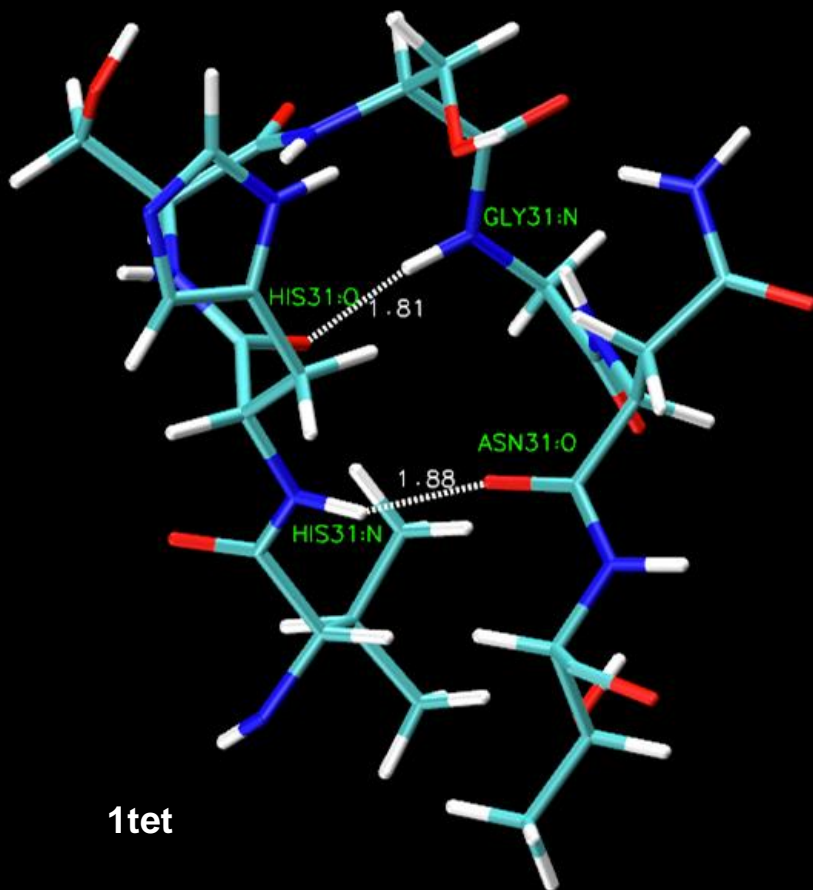
1zan
1vfa
1fgv



L1-K2A
L1-K2B

RMS = 0,70

L1 - k4



Hydrogen bonds:

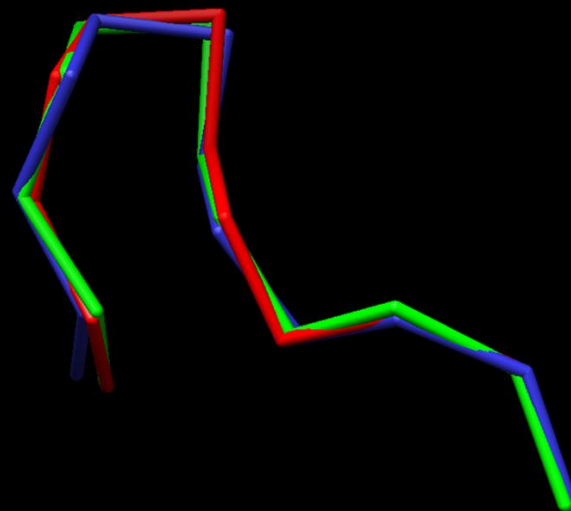
His	31	(O)	---	Gly	31.3	(N)
His	31	(N)	---	Asn	31.4	(O)

RSSqslvhsnGnTYle

1flr	SISC	RSS	QSLVHSNGNTYLRW
1tet	SISC	KSS	QSIVHSSGNTYFEW
2cqr	SISC	RPS	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)

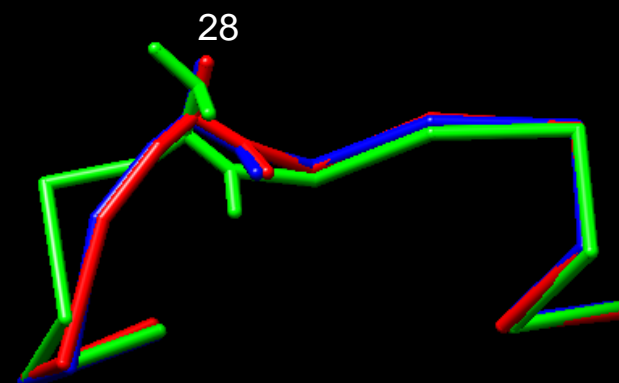
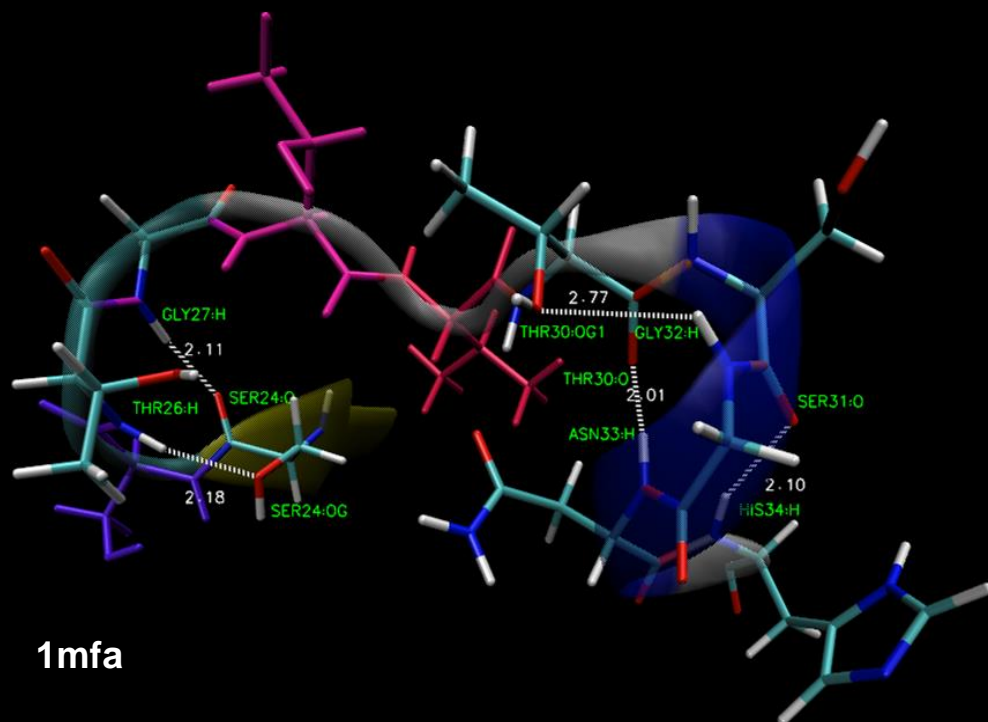
RSStGavTtsNYAn

A	<u>1ind</u>	CRSST	GAVTT	TSNYAN	WV
A	<u>1gig</u>	CRSST	GAVTT	TSNYAN	WV
B	<u>1mfa</u>	CRSST	GTVT	SGNHAN	WV

↑
28

↑
34

RSM = 0,58



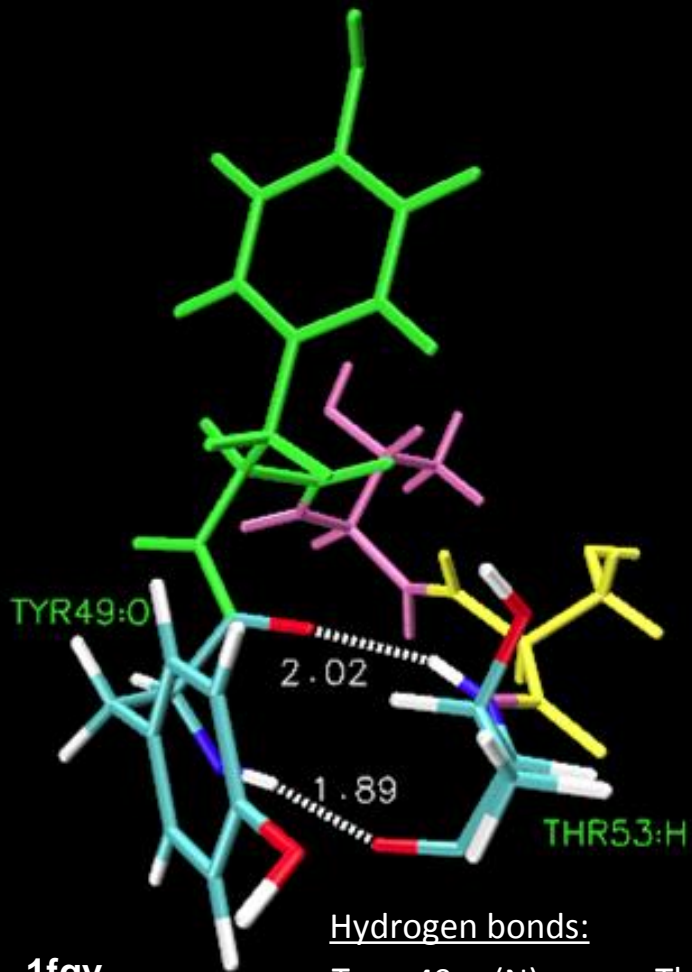
L2

positions: 50-52

Y-asnlas

<u>1flr</u>	KVLIYKVS	NRF
<u>1fgv</u>	KLLIYYT	STLE
<u>1lqm</u>	ELRIYD	ASNLE

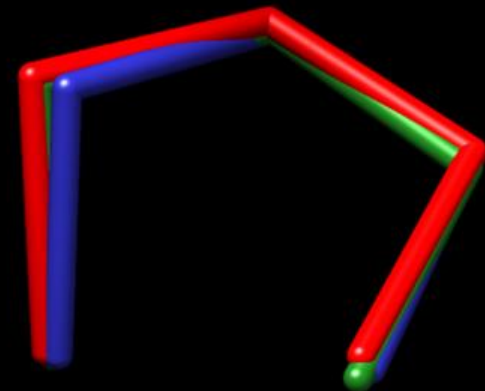
RMS = 0,52



1fgv

Hydrogen bonds:

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

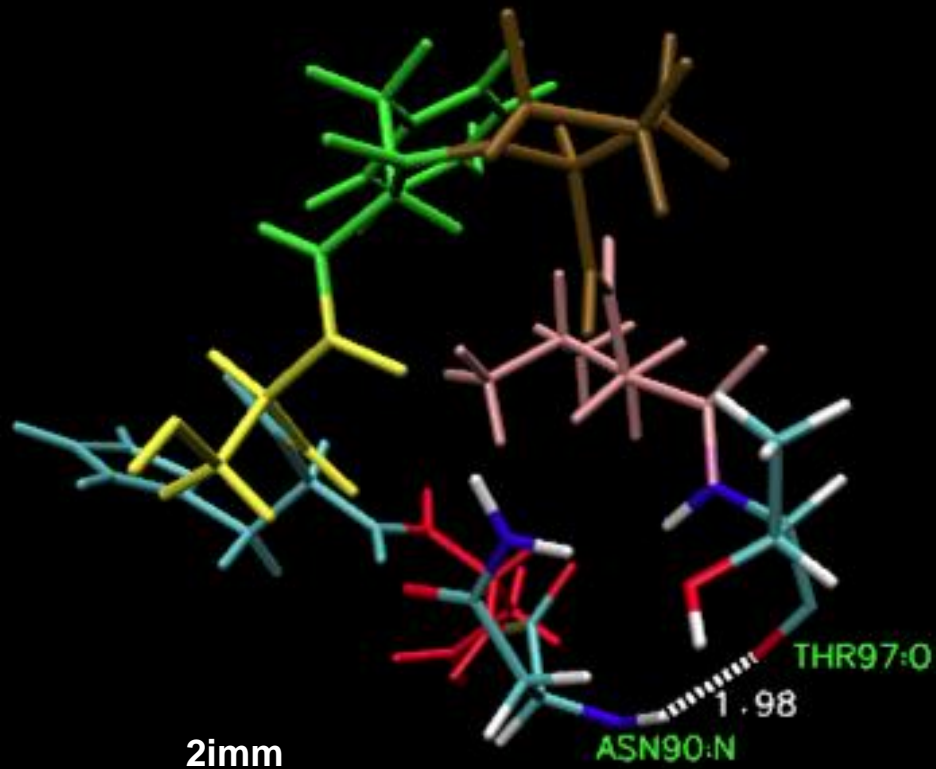


L3 - k1

L3: 89-96

Hydrogen bond:

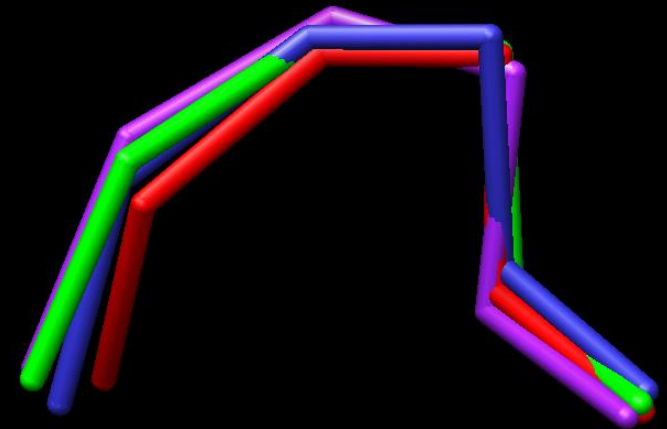
Asn 90 (N) --- Thr 97 (O)



qQgss-PIT

1hi	YCQ	NDYSN	PL	TFGGG
2im	YCQ	NDHSY	PL	TFGAG
1fqv	YCQ	QGNTL	PPT	TFGAG
1fvc	YCQ	QHYTT	PPT	TFGQG

RMS = 0,70



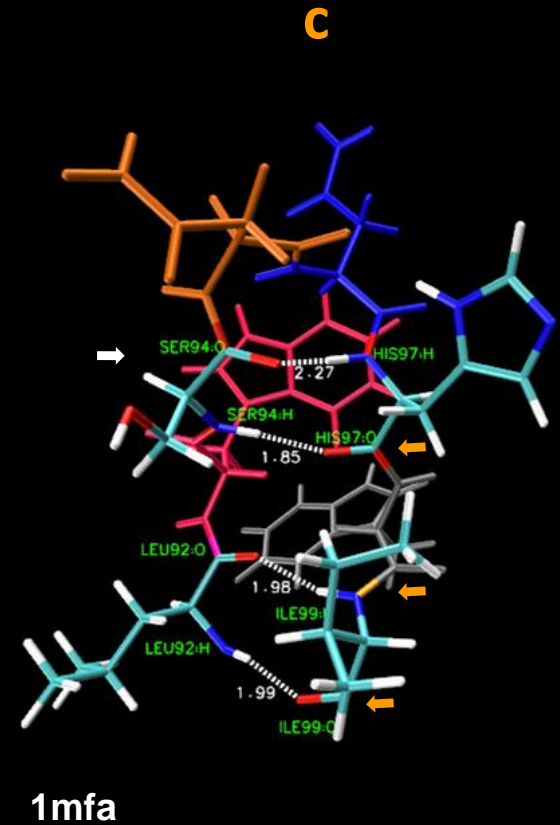
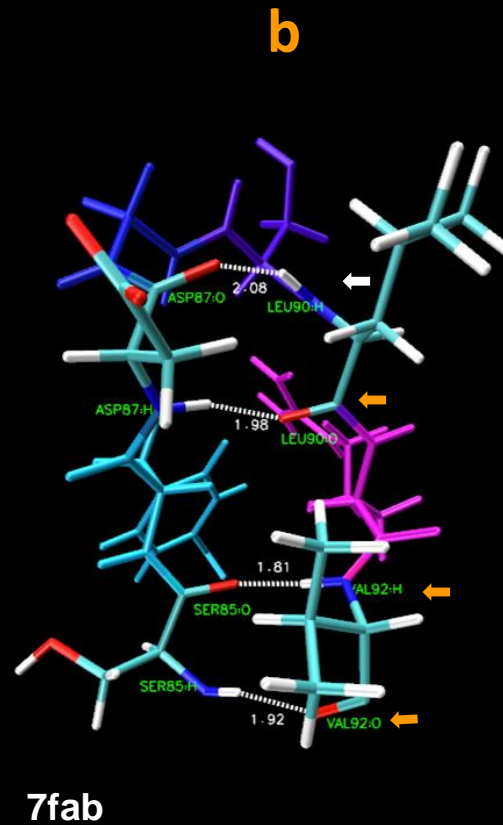
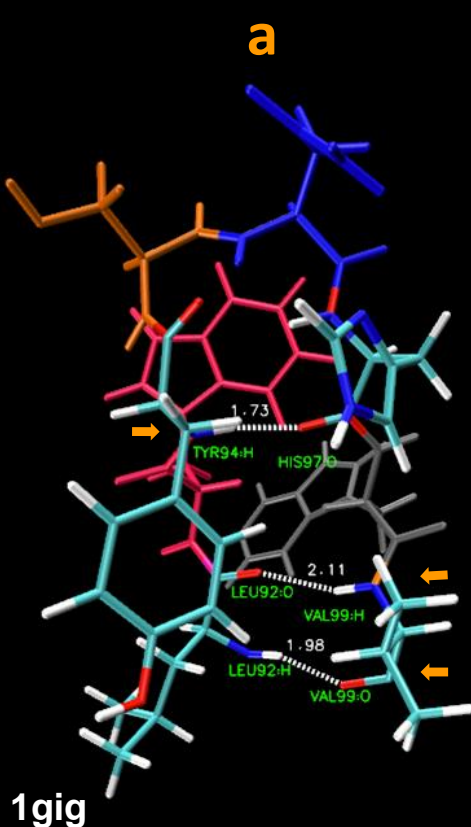
L3 - $\lambda 1$

alw-snhwv

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

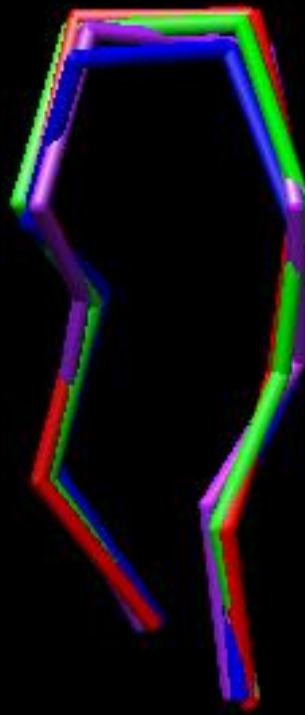
	1ind	FCALWYSNLWF	FGG
A	1mfa	FCALWSNNHWI	FGG
C	1gig	FCALWYSNHWV	FGG
A	7fab	YCQSYDRSLRV	FGG
B			

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

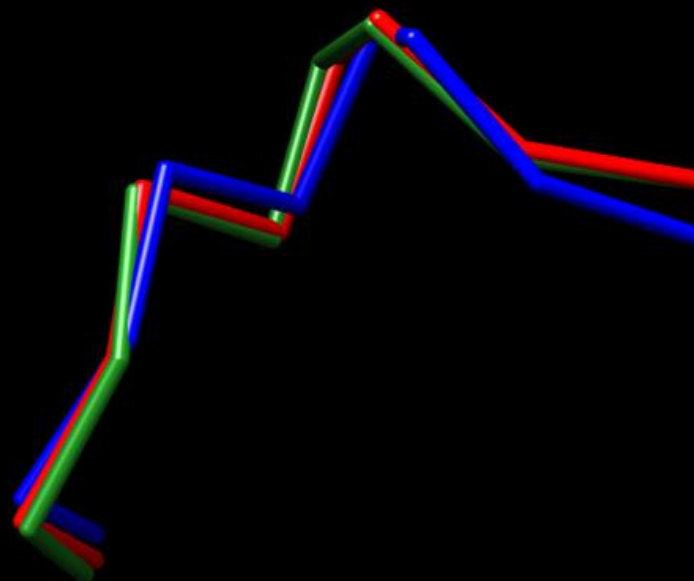
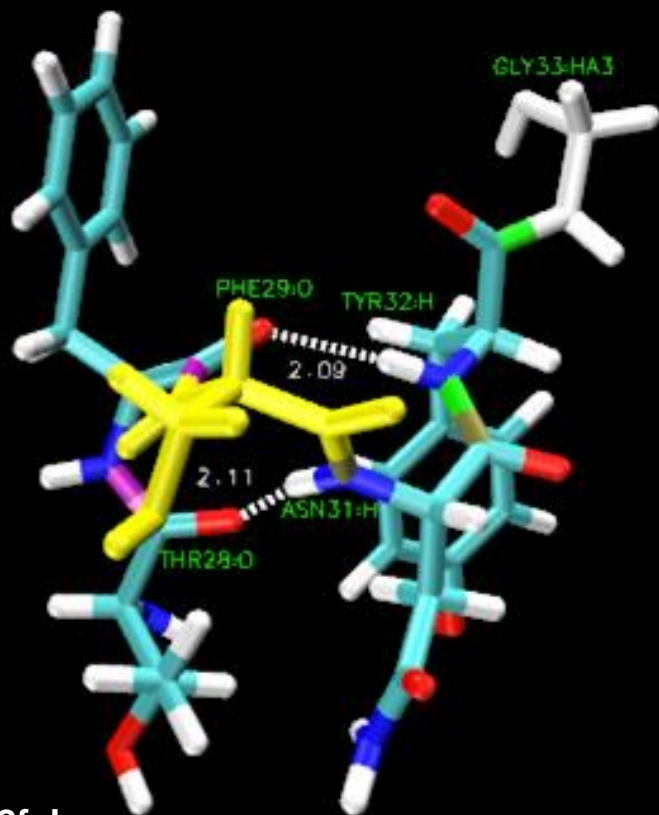
Hydrogen bonds:

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

ka**SG**ftftdyymh

<u>1h1l</u>	AASGFSFSYGM
<u>8fab</u>	IASGFTFSNYGMH
<u>1mfa</u>	KASGYTFTNYWMH

RMS = 0,95



H2 - 1

H2: 52-56

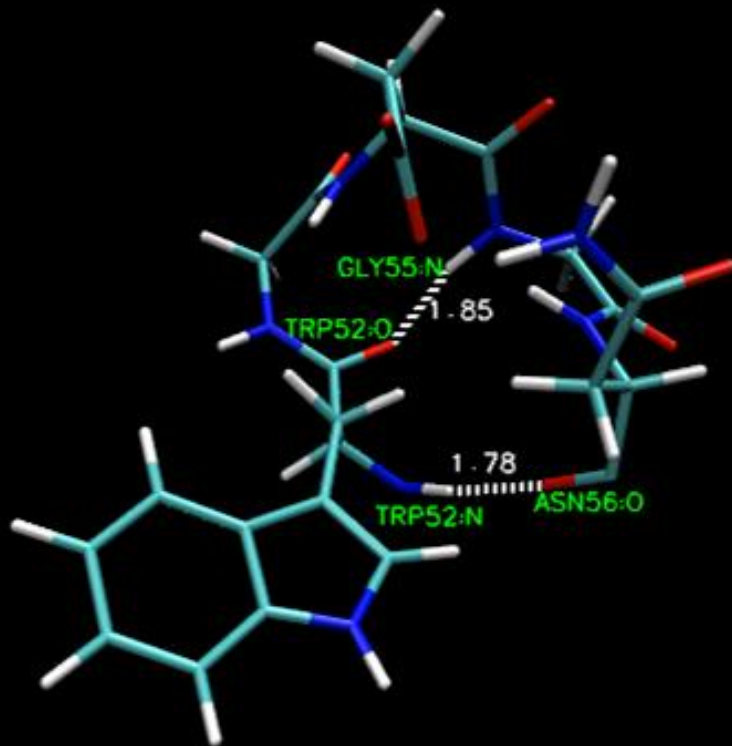
Hydrogen bonds:

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

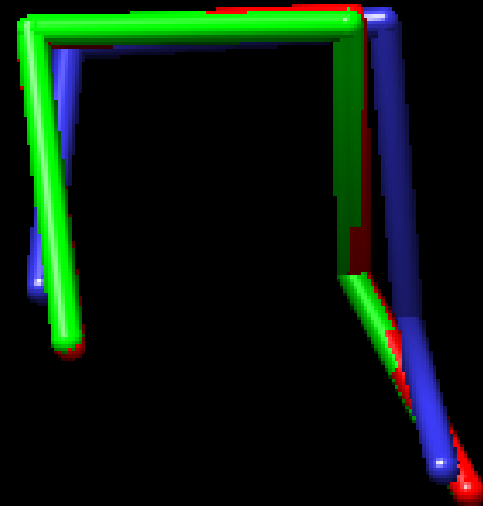
ylwysGsty

1vfa	EWLGM	I	WGD	GNTD
7fab	EWIGY	V	FYT	GTTL
1kip	EWLGM	I	WGD	GNTD

RMS = 0,41



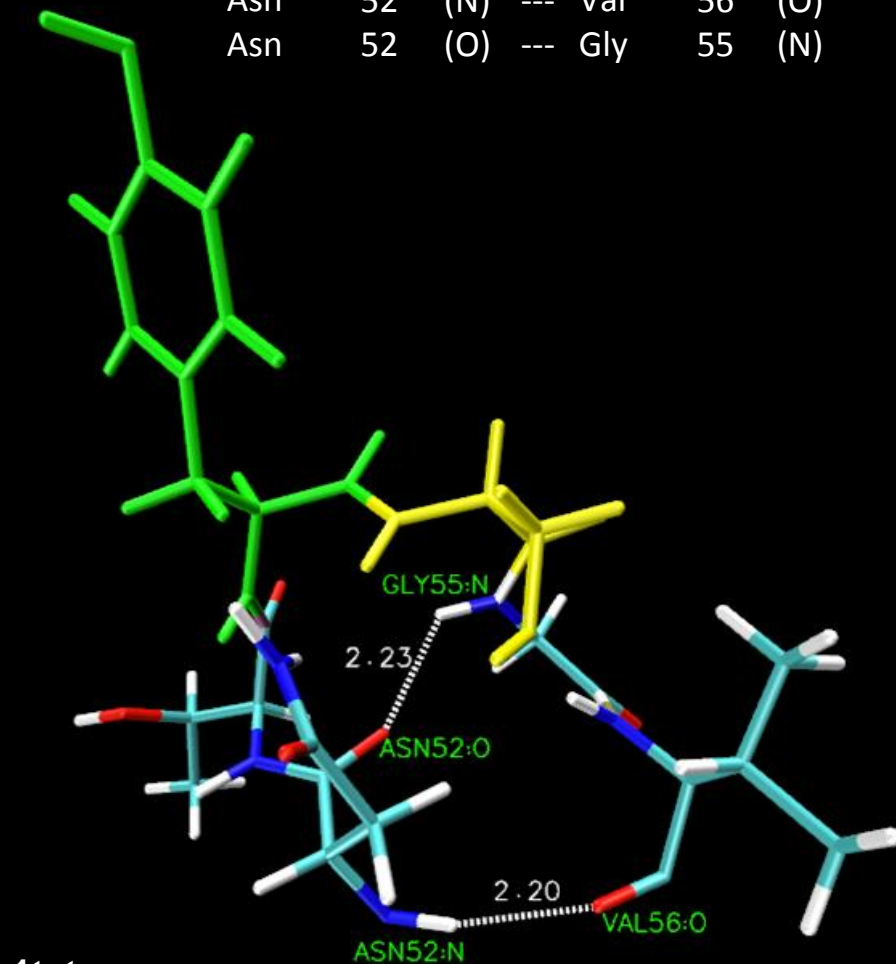
1vfa



H2 - 2a

Hydrogen bonds:

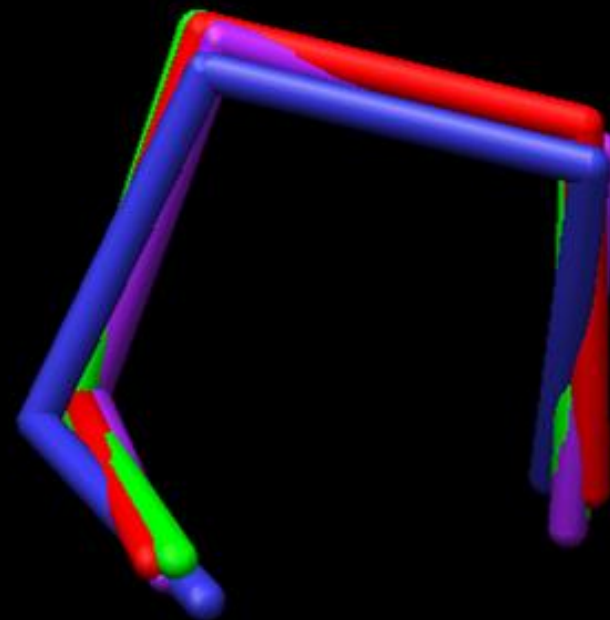
Asn	52	(N)	---	Val	56	(O)
Asn	52	(O)	---	Gly	55	(N)



-I ypngng- t -

1tet	KWMGW	I	NTYSGVPTY
1mfa	EWIGAT	I	YPGNSATFY
1fvc	EWVAR	I	YPTNGYTRY
1fgv	EWVAG	I	NPKNGGTSY

RMS = 0,47



H2 - 3c

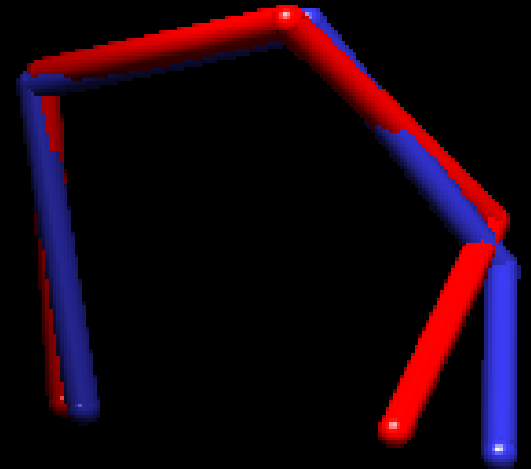
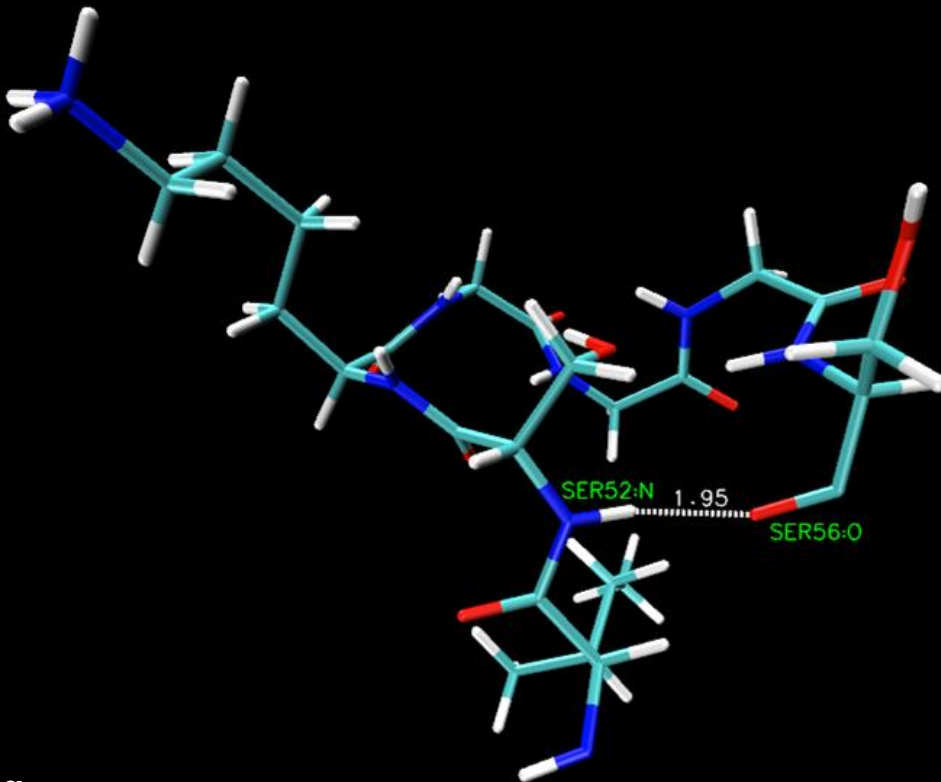
Hydrogen bonds:

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

- lssgggnty

1igm	SGV	FGSGGNTDY
1seq	AYI	SKGGGSTYY

RMS = 0,51

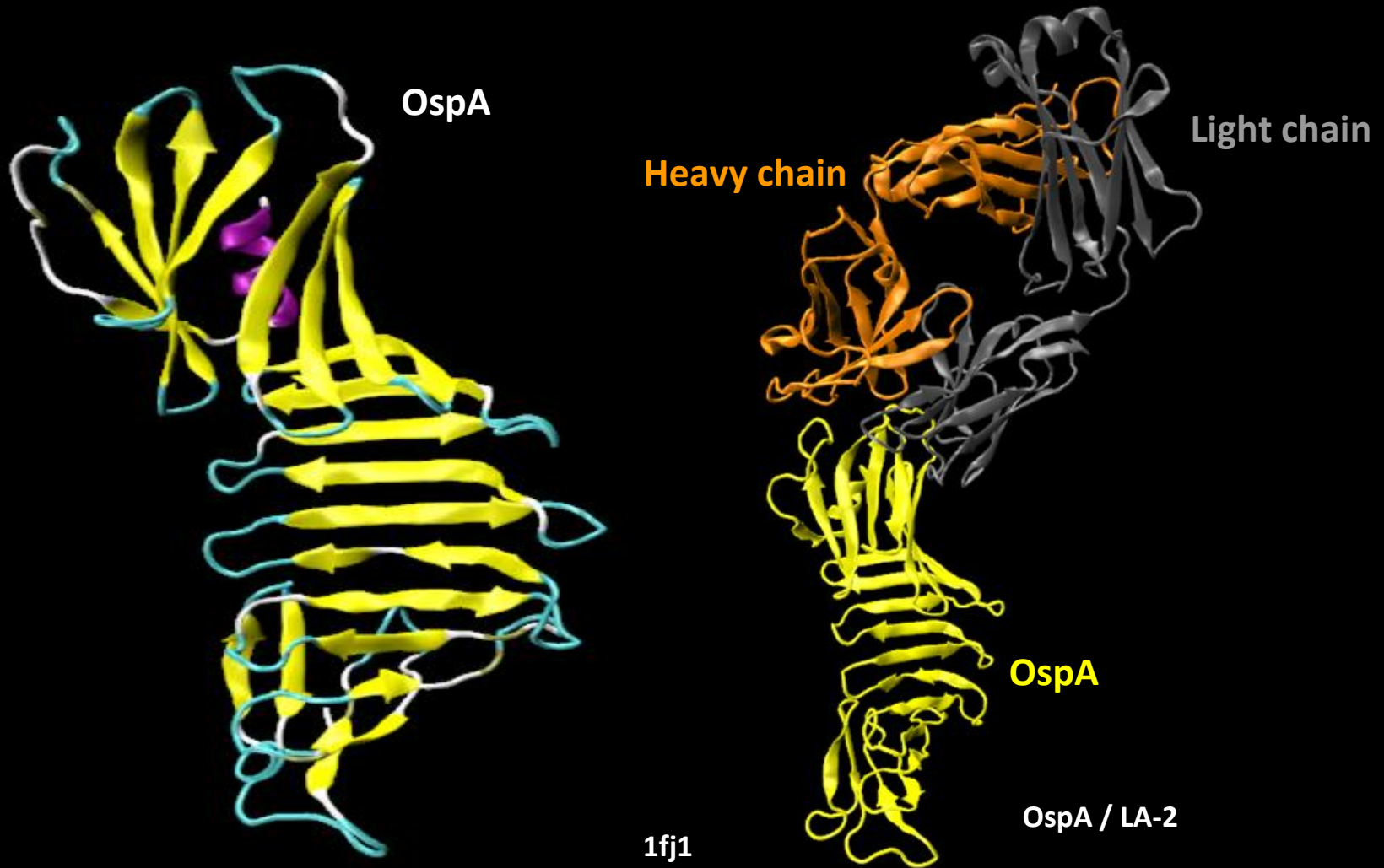


1seq

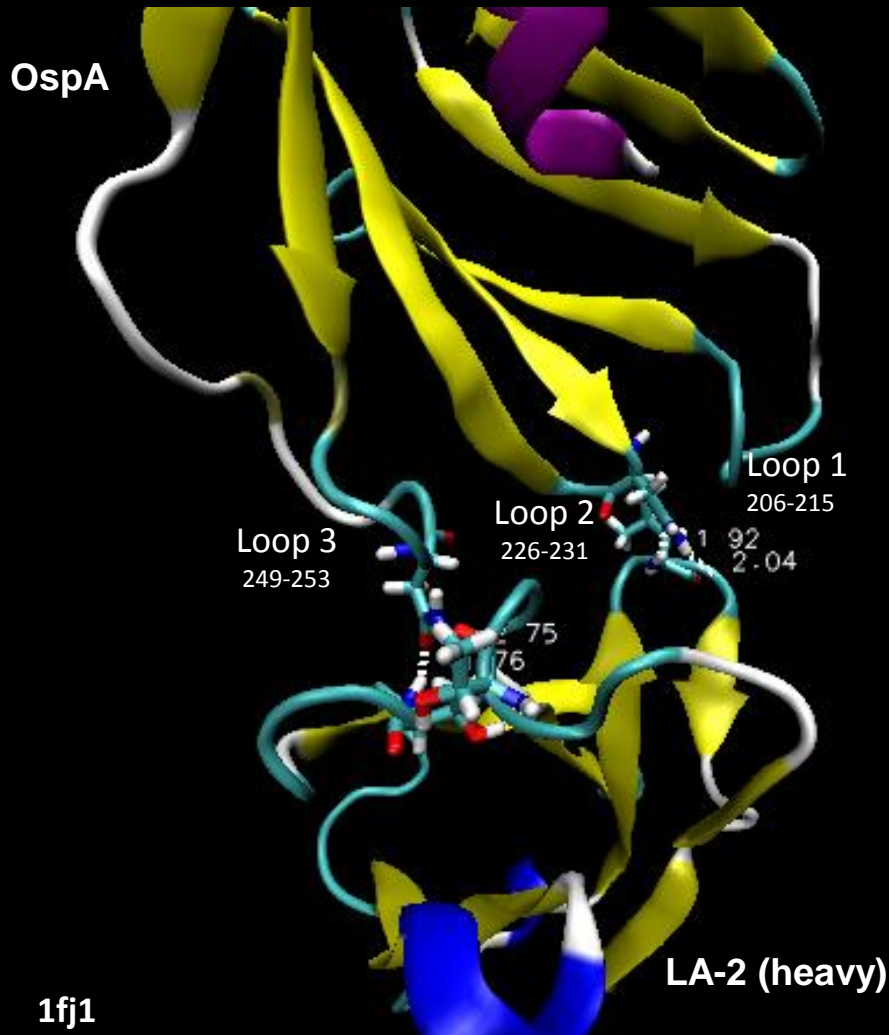
Canonical structures - Conclusions

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

Ag-Ab interaction: ospA/LA-2



OspA



Borrelia burgdorferi

Borrelia afzelii

Borrelia garinii

→ **CONSERVED EPITOPE**

↓

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

Important residues from OspA

	208	228
B.burgdorf	KEGTVTL SKNISKSGEVSVELNDTDS SSA ATKKTAAWNSGTSTLTITVNSKKT KDLVFTKE	
B.afzelii	KEGTVTL SKIEIAKSGEVTVALNDNTT Q ATKKTGAWDSKTSTLTISVNSKKT TQLVFTKQ	
B.garinii	TEGTVVL SKNILKSGEITVALDDSDT T ATKKTGKWDSKTSTLTISVNSQKT KNLVFTKE	
	***** **	***** **
	Loop 1	Loop 2
B.burgdorf	NTITVQQYDS N GTKLEGS AVEITKLDEIKNALK	
B.afzelii	DTITVQKYDS A GTNLEGT AVEIKTLDELKNALK	
B.garinii	DTITVQKYDS A GTNLEGT AVEIKTLDELKNALK	
	***** **	***** **
	Loop 3	

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



WHY?

A208 from OspA (loop 1)

Light chain

OspA



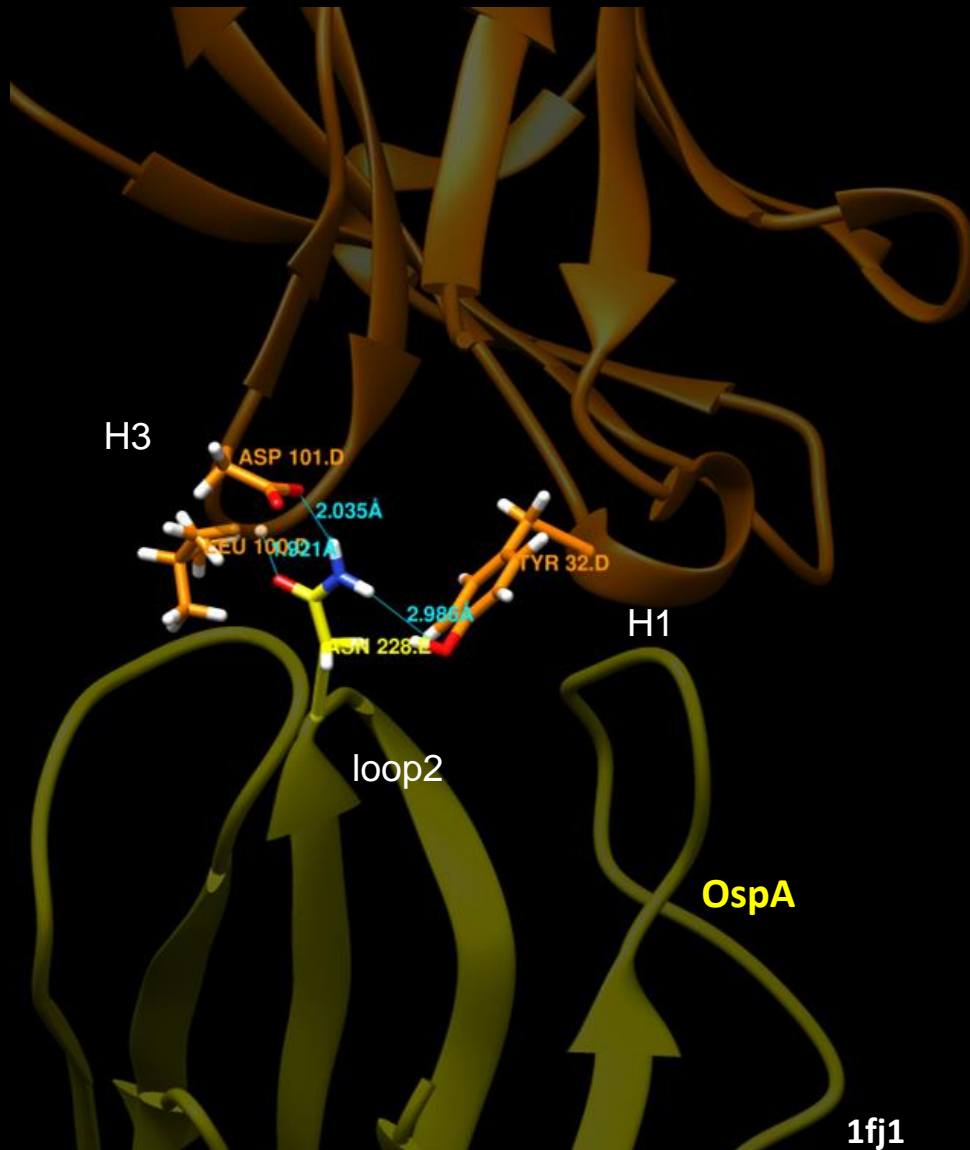
A208 adopts a specific conformation that allows to be bound by LA-2

A208

B.burgdorf
B.afzelii
B.garini

```
KEGTVTL SKNISKSGEVSVELNDTDSSAATKKTAAWNS  
KEGTVTL SKEIAKSGEVTVALNDTNTTQATKKTGAWDS  
TEGTVVL SKNILKSGEITVALDDSDTTQATKKTGKWD  
***** ** * *****..* * *. ..***** * *
```

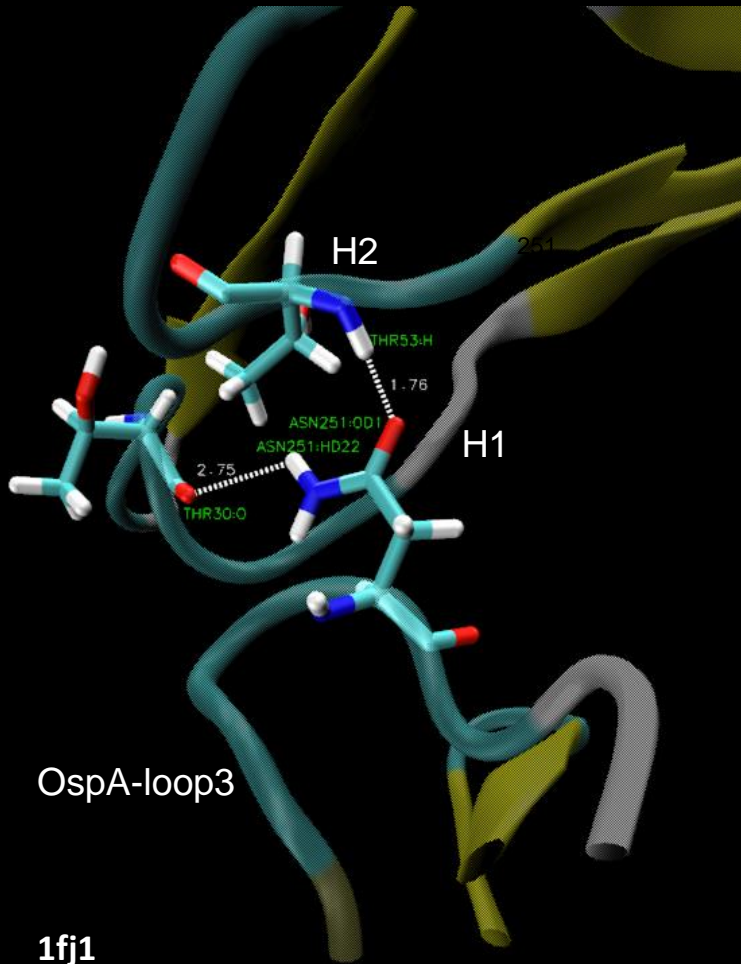
N228 from OspA (loop 2)



Hydrogen bonds:

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

N251 from OspA (loop 3)

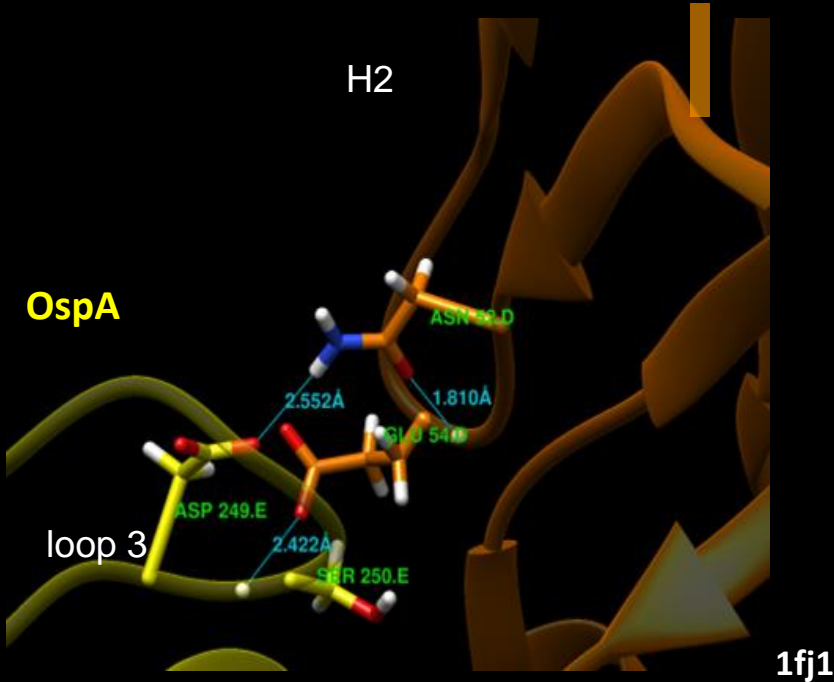


Hydrogen bonds:

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

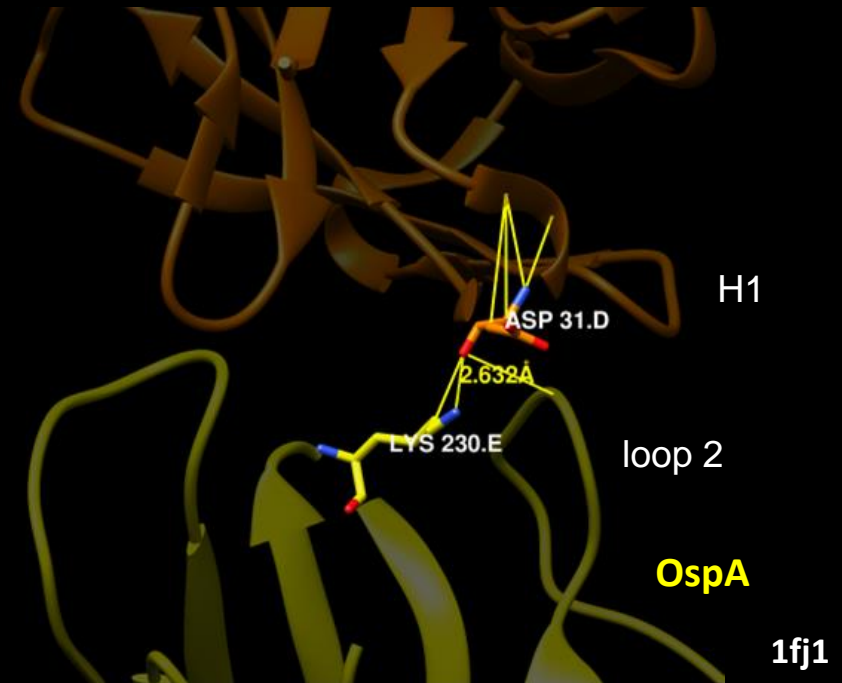
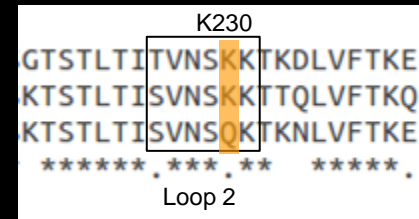
	N251	
B.burgdorf	NTITVQQYDSNGTKLEGS	AVEITKLDEIKNALK
B.afzelii	DTITVQKYDSAGTNLEGT	AVEIKTLDELKNALK
B.garini	DTITVQKYDSAGTNLEGT	AVEIKTLDELKNALK
	*****	*****
	Loop 3	

Other interactions between Ag-Ab



Hydrogen bonds:

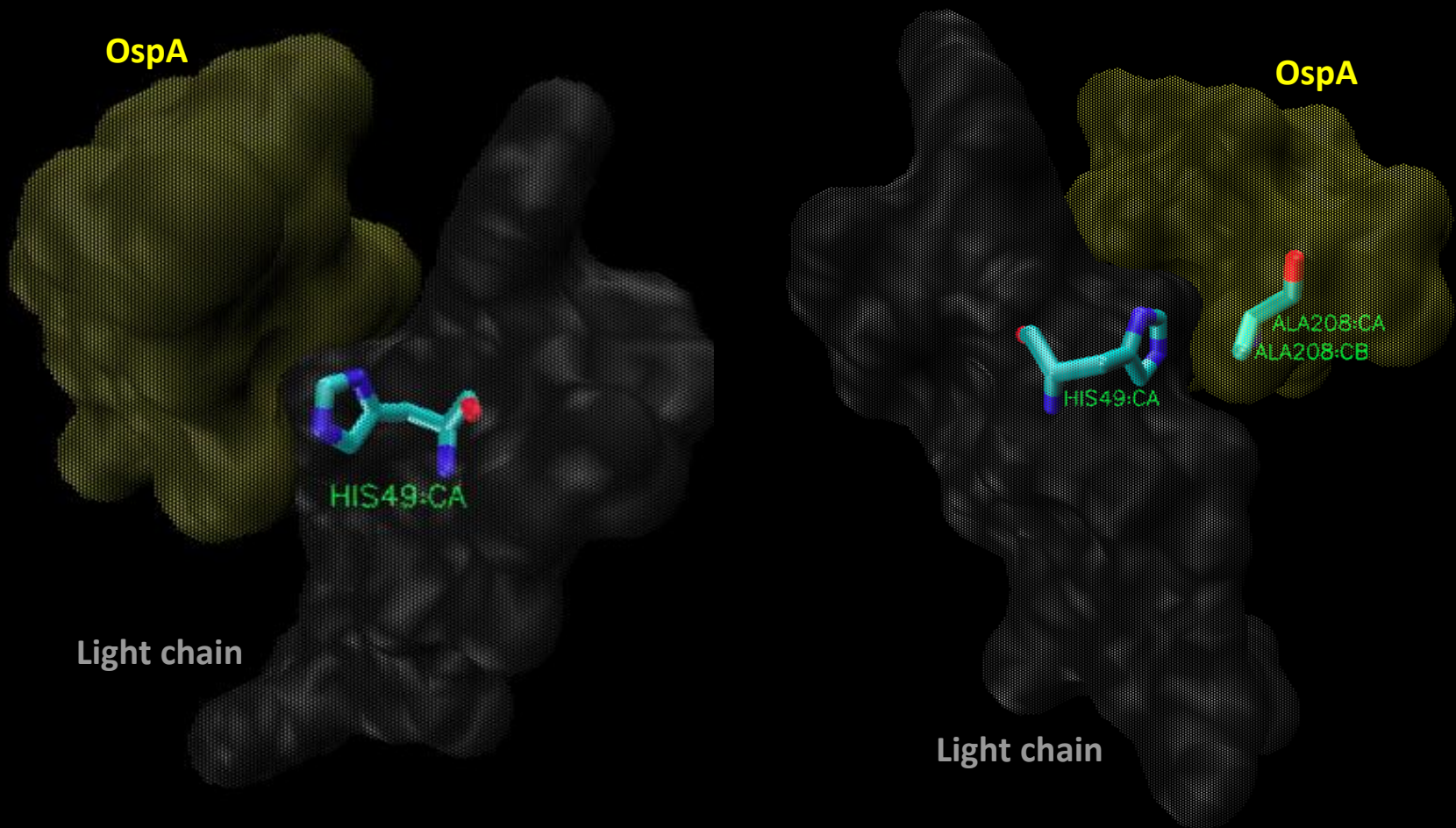
OspA				LA-2		
Asp	249	(OD2)	---	Asn	52	(ND2)
Ser	250	(N)	---	Glu	54	(OE1)



Salt bridge:

OspA				LA-2		
Lys	230	(NZ)	---	Asp	31	(O)

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 δ 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 Recong. 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-Madrado 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 β - barrel conformation
- e) all options are false

8. A canonical structure is:

- a) A structure that the b-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.

