

Computational methods to develop potential neutralizing antibody Fab region against SARS-CoV-2 as therapeutic and diagnostic tool

Hemanth K. Manikyam^{1*}, Sunil K. Joshi^{2*}

¹Faculty of Science, North East Frontier Technical University, Arunachal Pradesh, India.

Email: phytochem2@gmail.com

²Division of Haematology / Oncology & Bone Marrow Transplantation, Department of Pediatrics, University of Miami Miller School of Medicine, Miami, FL USA

Email: sunil.joshi@med.miami.edu

***Corresponding Authors**

Running Title: Diagnosis and Immunotherapy of COVID-19.

Keywords: SARS-CoV-2, coronavirus, IMGT, Human antibodies, I-TASSER, IgBLAST, Molbiol, LYRA, Fab, Nabs and mAbs

Abstract

SARS-CoV-2, a global pandemic originated from Wuhan city of China in the month of December 2019. There is an urgency to identify potential antibodies to neutralize the virus and also as a diagnostic tool candidate. At present palliative treatments using existing antiviral drugs are under trails to treat SARS-CoV-2. Whole Genome sequence of Wuhan market sample of SARS-CoV-2 was obtained from NCBI Gene ID MN908947.3. Spike protein sequence PDB ID 6VSB obtained from RCSB database. Spike protein sequence had shown top V gene match with IGLV1-44*01, IGLV1-47*02 and has VL type chain. Whole Genome sequence had shown top V gene match with IGHV1-38-4*01 and has VH type chain. VD chain had shown link to allele HLA-A0206 80%, HLA-A0217 80%, HLA-A2301 75%, HLA-A0203 75%, HLA-A0202 70% and HLA-A0201 55% of binding levels. Some conserved regions of spike protein had shown strong binding affinity with HLA-A-0*201, HLA-A24, HLA-B-5701 and HLA-B-5703 alpha chains. Synthetic Fab construct BCR type antibody IgG (CR5840) had shown Polyspecific binding activity with spike glycoprotein when compared with available Anti-SARS antibody CR3022. Thus we propose CR5840 Fab constructed antibody as potential neutralizing antibody for SARS-CoV-2. Based on germline analysis we also propose cytotoxic T lymphocyte epitope peptide selective system as effective tool for the development of SARS-CoV-2 vaccine.

Introduction

The SARS-CoV-2 causing COVID-19 pandemic threatened global public health and economy. Neutralizing antibodies (NAbs) elicited by virus during infection may correlate to patient recovery. However characterisation of neutralizing antibodies elicited during viral infection has not been well studied in relation with clinical outcome in the patients. Currently, neutralizing antibodies (NAbs) versus this virus are expected to correlate with recovery and protection of this disease. However, the characteristics of these antibodies have not been well studied in association with the clinical manifestations in patients. NAbs play crucial role in controlling the viral infection. Monoclonal antibodies (mAbs) functional antigen-binding fragment (Fab) is the form of NAbs developed for SARS-CoV and MERS currently developed. These mAbs target Receptor binding domain and N terminal domains. mAbs thus inhibit membrane fusion and restrict viral host cell entry.

Palivizumab is a monoclonal antibody (mAb) produced by recombinant DNA technology used to reduce the risk of respiratory syncytial virus in children. Phage-display technology a new method of developing promising method in antibody development by mAbs that utilizes fragment antigen binding (Fab) molecules as an intermediate. This technology had proven the neutralization effect of Fabs against influenza virus. Phage display technology an alternative method of producing monoclonal antibodies that utilizes bacteriophages that display fragments of antibodies Fabs from viral surface. Fab fragment of an antibody contain light kappa chain and heavy chain that binds actively to specific regions of antigens. Human neutralizing mAbs like CR6261, Navivumab, Firivumab and Diridavumab are presently used to treat Influenza virus infections. Human neutralizing mAbs S230.15 and m396 isolated from SARS-COV-1 infected patients had shown competitive inhibition with ACE2 for binding RBD. Highly conserved Residue regions are found to key factors for cross reactivity.

Polyclonal antibodies from SARS-CoV-2-infected patients been used to treat SARS-CoV-2 infection, but no SARS-CoV-2-specific neutralizing mAbs or Fabs have been reported. Researchers around globe are working hard to identify monoclonal antibodies or their functional fragments Fabs as prophylactic, therapeutic and diagnostic agents against COVID-19 caused by SARS-CoV-2. Since SARS-CoV-2 closely related to SARS-Co-V-2 the conserved residues with epitopes may similar structural and functional properties. Hence there are chances of utilising mAbs or it active fragments Fabs against COVID-19 for neutralizing or cross neutralizing activity which to be proved through proper clinical studies.

As it takes time to identify proper mAbs against SARS-CoV-2, we suggest application of computational methods to predict the mAbs based on structural confirmations submitted like SARS-CoV-2 spike proteins. Once specific mAbs or Fab regions are predicted and their sequence been obtained, Phage display technology can be used to synthesize Fab fragments or Monoclonal antibody technology can be used to produce mAbs as a therapeutic and diagnostic agents against COVID-19.

Methods

IgBLAST and IMGT (Immunogenetics) Analysis:

IgBLAST was developed by NCBI for analysis of variable domain sequences of immunoglobulin^{1, 4, 5}. IMGT is a platform to access sequence, genome and structure Immunogenetics data based on gene ontology. After immunoglobulin germline prediction we

compare the V-J and V-D-J junction sequence with the extremities of the corresponding “V, J and D "germline" genes (see IMGT Repertoire > Alignments of alleles).^{8,9}

Binding Site and Ligand Prediction:

ProBis a computer based tool to predict binding site and corresponding Ligands for submitted protein structure. Spike protein sequence PDB ID 6VSB was obtained from RCSB database and submitted to ProBis to predict specific Ligands.^{6, 10}

I-TASSER (Iterative Threading Assembly Refinement):

I-TASSER is a computational approach to predict protein structure and function. It is an most accurate protein structure and function prediction algorithm.^{11, 12} After IMGT and ProBis analysis predicted V-D-J and Fab Ligands were obtained and submitted to I-TASSER for best functional model of ligand which can be obtained as PDB file.

Lymphocyte Receptor Automated Modelling (LYRA):

Predicted protein ligand sequences from ProBis algorithm obtained from RCSB and compared for best possible alignments and combinations using IgBLAST and LYRA predictions tools. LYRA server predicts either T cell or B cell receptors based on homology modelling. Based on BLOSUM score Framework templates are selected and complementary determining regions (CDR) are then selected if needed based on a canonical structure model and grafted onto the framework templates. The final output is a complete three-dimensional model of the target antibody.¹⁴

Stereochemical quality of a protein structure:

PROCHECK EMBL-EBI algorithm checks the Stereochemical quality of a protein structure and produces a number of postscript plots analysing residue-by-residue geometry¹³

FRDOCK and Pydock protein-protein docking analysis:

Constructed Fab fragment using I-TASSER and LYRA algorithms in PDB format was studied for Structural prediction of protein-protein interaction using FRDOCK and Pydock algorithms.^{15, 16} Spike protein structure obtained from RCSB database PDB ID: 6VSB. See table 1 and 2.

Background

Presently world is facing pandemic caused by SARS-CoV-2 and took more than 60,000 lives as per existing records. Any infections caused by microorganisms develop human neutralising antibodies after certain period of time. Microbial genome contains antigenic sequences that can trigger certain type of IMGT antibodies in the host which can be useful to design vaccines and also as diagnostic tools. IgBLAST and Molbiol tools are useful in detecting top gene matches and their germline antibodies using genome sequence of pathogens. We conducted IgBLAST analysis of whole genome and also Spike protein of SARS-CoV-2 coronavirus in order to identify possible antigen neutralizing human antibodies. Immune Epitope Database (IEDB) tools used to identify antigen sequence epitopes. Lymphocyte receptor automated modelling (LYRA) and I TASSER tools are used to construct the BCR antibody.

Results

IgBLAST and IMGT (Immunogenetics) Analysis:

Sequences producing significant alignments with germline gene and top gene alignment shown in figure 1&3 for whole genome and figure 2&4 for spike protein genome. Imgt Domain Classification was studied for given sequence. Minus strand of a V gene and has been converted to the plus strand during IgBLAST. V-D-J rearrangement summary of query sequence shown in figure 1 for whole genome and figure 2 for spike protein genome. Alignments with top V gene hits are shown in supplementary file 1, file 2 and file 3. Spike protein sequence had shown top V gene match with IGLV1-44*01, IGLV1-47*02 and has VL type chain. Whole Genome sequence had shown top V gene match with IGHV1-38-4*01 and has VH type chain. VD chain had shown link to allele HLA-A0206 80%, HLA-A0217 80%, HLA-A2301 75%, HLA-A0203 75%, HLA-A0202 70% and HLA-A0201 55% of binding levels. See supplementary file 3. Some conserved regions of spike protein had shown strong binding affinity with HLA-A-0*201, HLA-A24, HLA-B-5701 and HLA-B-5703 alpha chains.

ProBis, I-TASSER and LYRA analysis:

From epitope and germline data based on spike protein sequence, ProBis ligand prediction tools generated antibodies that are specific to spike glycoprotein. From the sequence of specific antibodies obtained, we generated synthetic IgG hybrid antibody Fab sequence. LYRA and I-

TASSER tools predicted possible structural conformation from heavy chain and kappa light chain and generated final hybrid antibody CR5840 see figure 5 and 6

Stereochemical quality of a protein structure:

Quality analysis of structurally confirmed Fab fragment annotated as CR5840 was conducted using PROCHECK EMBL-EBI algorithm. Ramchandran plot, residue to residue analysis, RMS distance analysis with main chain bond length, angle analysis and parameters obtained see figure 9, 10 and 11. Protein-protein interface of heavy chain and kappa light chain of Fab fragment shown in figure 12. Interacting chains are joined by coloured lines, each representing a different type of interaction, as per the key above. The area of each circle is proportional to the surface area of the corresponding protein chain. The extent of the interface region on each chain is represented by the black wedge whose size signifies the interface surface area. Statistics for this interface are given Table 1. Residue interactions shown in figure 13. The number of H-bond lines between any two residues indicates the number of potential hydrogen bonds between them. For non-bonded contacts, which can be plentiful, the width of the striped line is proportional to the number of atomic contacts see figure 13. Topology of Fab fragment shown in figure 8. Cleft, tunnel and Pores analysis had figure 14, 15, 16 and supplementary file 4. Binding site analysis shown in table. List of all matched Gene Ontology (GO) terms, Reverse template with certain matches (E-value < 1.00E-06) and list of predicted matches for CR5840 were presented in supplementary file 5. Most of the gene ontology predicted CR5840 structure for immunoglobulin IgG complex having cell surface interactions.

FRDOCK and Pydock protein-protein interaction:

FRDOCK and Pydockweb tools are used to run protein-protein interaction CR5840 as ligand and Spike glycoprotein as Receptor (PDB ID: 6VSB) in comparison with available antibody CR3022. CR5840 had shown high binding affinity see figure 7 when compared to CR3022 against spike glycoprotein of SARS-CoV-2. Predicted rank order for top models of binding pockets shown in table 2 and 3.

Discussion

There is an urgent necessary to address potential therapeutics for global pandemic created by SARS-CoV-2 infection. Viral antigens go through human innate immune system during initial or latent periods of incubation period and generate neutralizing antibodies against the

antigens. Viral genome carries the coding for viral antigens which can be used to predict neutralizing human antibodies using different genomic and Highthroughput screening tools. Here used IgBLAST, I-TASSER and Molbiol tools to predict and design the neutralizing human antibodies against SARS-CoV-2 infection. Whole genome and spike protein genome sequence of SARS-CoV-2 were obtained from NCBI and RCSB database and IgBLAST studies carried out. Whole Genome sequence had shown top V gene match with IGHV1-38-4*01 and has VH type chain. Spike protein sequence had shown top V gene match with IGLV1-44*01, IGLV1-47*02 and has VL type chain. Thus we propose possibility for usage of VH and VL type of antibody germline as per sequence obtained for V-D-J chain for whole genome and V chain sequence for Spike protein mentioned in this research work. We propose cytotoxic T lymphocyte epitope peptide selective system as effective tool for the development of SARS-CoV-2 vaccine. Top predicted structural conformation of V-D Chain of neutralizing antibody shown in figure 5 with C-score=-2.03 (Read more about C-score) , Estimated TM-score = 0.47 ± 0.15 , Estimated RMSD = $5.8 \pm 3.6 \text{ \AA}$. Fab fragment with Polyspecific activity IgG1 like properties CR5840 had shown highest affinity with spike protein PDN ID: 6VSB when compared with CR3302 an available Fab of antibody used in Rapid kit test for SARS-CoV-2. Thus we propose use of novel synthetic Fab active complex as therapeutic and diagnostic tool against SARS-CoV-2.

Acknowledgement:

This work was supported by the Bill and Melinda Gates Foundation Grand Challenge Exploration grant (GCE Round 4: OPP1015489) to SKJ.

Competing interests: The authors declare no competing interests.

References

1. <https://www.ncbi.nlm.nih.gov/igblast/igblast.cgi>
2. <https://zhanglab.ccmb.med.umich.edu/I-TASSER/>
3. A Roy, AKucukural, Y Zhang. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols*, 5: 725-738 (2010)
4. https://www.bioinformatics.org/sms2/rev_trans.html
5. <https://molbiol-tools.ca/>
6. http://probis.cmm.ki.si/browse.php?job_id=30032077587240
7. <http://tools.iedb.org/lyra/>
8. Lefranc M-P. Lefranc G. IMGT® and 30 years of Immunoinformatics Insight in Antibody V and C Domain Structure and Function. In Jefferis R; Strohl W. R., Kato K. *Antibodies* (Basel). 2019 Apr 11;8(2) 29. pii: E29. doi: 10.3390/antib8020029. PMID: 31544835
9. Ye J1, Ma N, Madden TL, Ostell JM. "IgBLAST: an immunoglobulin variable domain sequence analysis tool". *Nucleic Acids Res.* 2013 Jul; 41(Web Server issue):W34-40. doi: 10.1093/nar/gkt382. Epub 2013 May 13.
10. Janez Konc; Dušanka Janežič (2010). "ProBiS algorithm for detection of structurally similar protein binding sites by local structural alignment". *Bioinformatics*. 26 (9): 1160–1168. doi:10.1093/bioinformatics/btq100. PMC 2859123. PMID 20305268.
11. A Roy, A Kucukural, Y Zhang. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols*, 5: 725-738 (2010)
12. J Yang, R Yan, A Roy, D Xu, J Poisson, Y Zhang. The I-TASSER Suite: Protein structure and function prediction. *Nature Methods*, 12: 7-8 (2015)
13. Laskowski R A, MacArthur M W, Moss D S, Thornton J M (1993). PROCHECK - a program to check the stereochemical quality of protein structures. *J. App. Cryst.*, 26, 283-291.
14. Klausen MS, Anderson MV, Jespersen MC, Nielsen M, Marcatili P. LYRA, a webserver for lymphocyte receptor structural modeling. *Nucleic Acids Research*. 2015 Jul 1; 43(Web Server issue): W349-W355. doi:10.1093/nar/gkv535. PMID: 26007650
15. E. Ramírez-Aportela, J.R. López-Blanco, and P. Chacón (2016). FRODOCK 2.0: Fast Protein-Protein docking server. *Bioinformatics*, 32(15), 2386-2388.
16. Jimenez-Garcia B., Pons C. and Fernandez-Recio J. "pyDockWEB: a web server for rigid-body protein-protein docking using electrostatics and desolvation scoring". *Bioinformatics* (2013) 29(13):1698-1699.

Chain	No. of Interface Residues	Interface Area (Å ²)	No. of Salt Bridges	No. of Disulphide Bonds	No. of Hydrogen Bonds	No. of Non-Bonded Contacts
H	11	692	-	-	1	78
L	11	737				

Table 1: Protein - Protein interface statistics of heavy chain and kappa light chain of Fab region.

Conf	Electrostatics	Desolvation	VdW	Total	Rank
9721	-15.031	-24.564	-1.536	-39.748	1
7731	-8.751	-32.630	28.402	-38.541	2
5986	-15.974	-18.586	-17.699	-36.330	3
2796	-8.088	-32.259	42.454	-36.102	4
4113	-16.412	-24.967	56.981	-35.681	5
3627	-5.174	-29.850	-6.292	-35.653	6
502	-9.962	-27.967	29.594	-34.970	7
309	-17.792	-19.174	20.341	-34.932	8
6896	-3.943	-34.699	38.134	-34.828	9
414	-14.218	-21.921	15.601	-34.579	10

Table 2: Top 10 energies and Rank predicted for CR5840 interaction with SARS-CoV-2 Spike protein (PDB ID: 6VSB).

Conf	Electrostatics	Desolvation	VdW	Total	Rank
6007	-1.958	-28.188	44.860	- 25.659	1
1959	-20.303	-3.655	- 15.395	- 25.497	2
5463	-19.326	-4.472	- 11.838	- 24.982	3
669	-6.698	-18.374	1.197	- 24.953	4
3486	-10.300	-15.971	20.146	- 24.256	5
7414	-5.163	-18.277	-3.160	- 23.757	6
4729	-19.276	-6.367	19.584	- 23.684	7
6983	-18.379	-11.715	65.594	- 23.535	8
3956	-17.005	-11.086	55.340	- 22.557	9
7953	-8.637	-16.602	27.557	- 22.483	10

Table 3: Top 10 energies and Rank predicted for CR3022 interaction with SARS-CoV-2 Spike protein (PDB ID: 6VSB)

V-(D)-J rearrangement summary for query sequence (multiple equivalent top matches, if present, are separated by a comma):

Top V gene match	Top D gene match	Top J gene match	Chain type	stop codon	V-J frame	Productive	Strand
IGHV1-38-4*01	IGHD6-13*01	N/A	VH	Yes		No	-

V-(D)-J junction details based on top germline gene matches:

V region end	V-D junction*	D region	D-J junction*	J region start
TTTAT	CATTGTAGATGTCAAAAGCCCTGTATACGACATCAGTACTAGTGCCCTGTGCCGCACGGTGAAGACGGGCTGCACTTACACCGCAAACCCGTTTAAAAACGATTGT	GCATCAGCTG		

*: Overlapping nucleotides may exist at V-D-J junction (i.e, nucleotides that could be assigned to either rearranging gene). Such nucleotides are indicated inside a parenthesis (i.e., (TACAT)) but are not included under the V, D or J gene itself.

Alignment summary between query and top germline V gene hit:

	from	to	length	matches	mismatches	gaps	identity(%)
FR1-IMGT	16310	16319	10	8	2	0	80
CDR1-IMGT	16320	16343	24	17	7	0	70.8
FR2-IMGT	16344	16345	2	2	0	0	100
Total			36	27	9	0	75

Figure 1. Predicted V-D-J junctions and top germline V gene hits for the query sequence of Gene ID MN908947.3

V-(D)-J rearrangement summary for query sequence (multiple equivalent top matches, if present, are separated by a comma):

Top V gene match	Top J gene match	Chain type	stop codon	V-J frame	Productive	Strand
IGLV1-44*01,IGLV1-47*02	N/A	VL	No			-

V-(D)-J junction details based on top germline gene matches:

V region end	V-J junction*	J region start
TGCGG		

*: Overlapping nucleotides may exist at V-D-J junction (i.e., nucleotides that could be assigned to either rearranging gene). Such nucleotides are indicated inside a parenthesis (i.e., (TACAT)) but are not included under the V, D or J gene itself.

Alignment summary between query and top germline V gene hit:

	from	to	length	matches	mismatches	gaps	Identity (%)
FR2-IMGT	478	516	39	22	17	0	56.4
CDR2-IMGT	517	525	9	8	1	0	88.9
FR3-IMGT	526	531	6	5	1	0	83.3
Total			54	35	19	0	64.8

Figure 2. Predicted V-D-J junctions and top germline V gene hits for the query sequence of Spike protein side chain A PDB ID 6VSB

Query= MN908947.3

Length=29903

Sequences producing significant alignments:

	Score (Bits)	E Value
IGHV1-38-4*01germline gene	29.9	4.1
IGHV2-70*19germline gene	29.9	4.1
IGLV1-50*01germline gene	29.9	4.1
IGHD6-13*01germline gene	14.1	636
IGHD2-15*01germline gene	14.1	636
IGHD2-2*01germline gene	14.1	636

Domain classification requested: ingt

Figure 3 Sequences producing significant alignments for query sequence whole genome of SARS-CoV-2

Query= reverse translation of 6VSB:A|PDBID|CHAIN|SEQUENCE

Length=3864

Sequences producing significant alignments:

Score (Bits)	E Value
-----------------	------------

[IGLV1-44*01](#)germline gene

26.8	4.6
----------------------	-----

[IGLV1-47*02](#)germline gene

26.8	4.6
----------------------	-----

[IGLV3-22*01](#)germline gene

26.8	4.6
----------------------	-----

Domain classification requested: imgt

Figure 4 Sequences producing significant alignments for query sequence Spike protein chain A genome of SARS-CoV-2

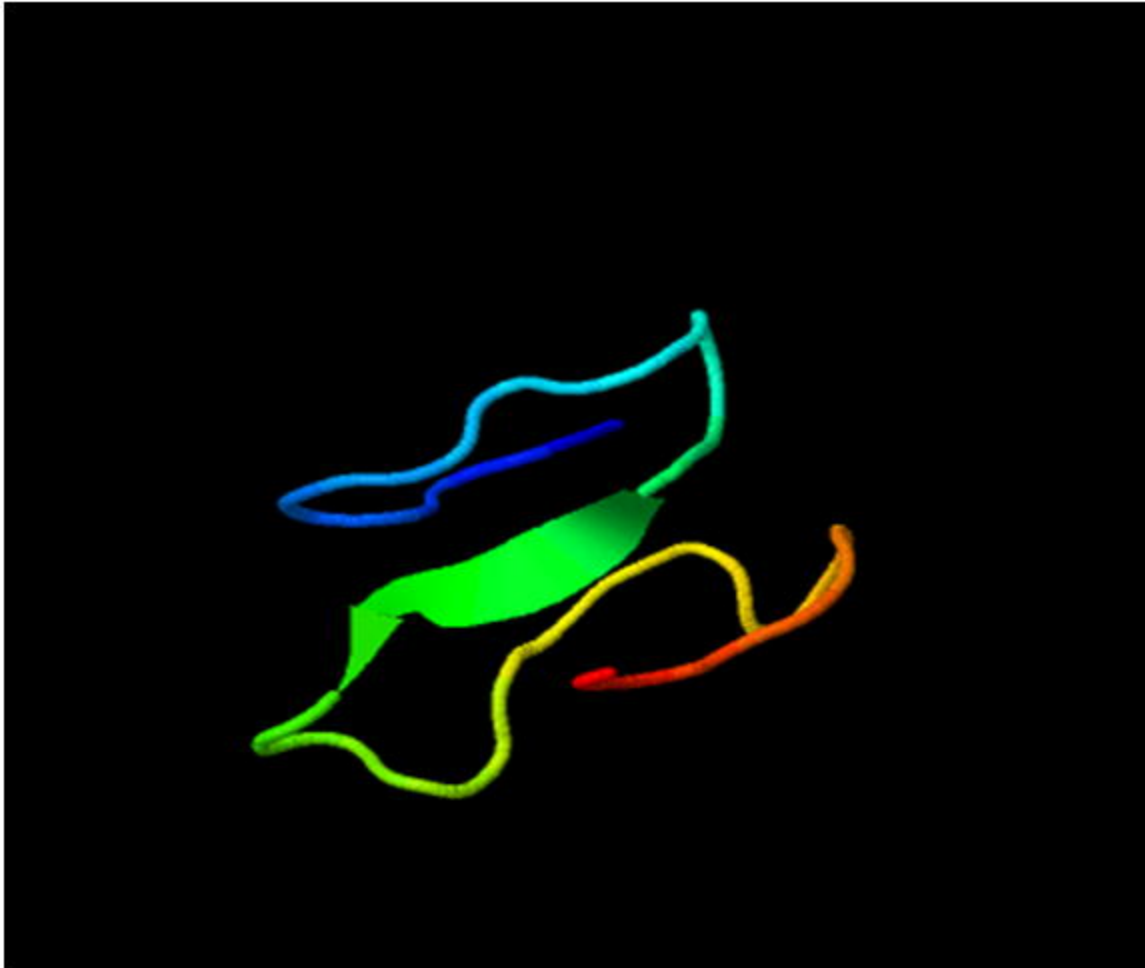


Figure 5. Top I-TASSER structural confirmation for VD chain for whole genome sequence of SARS-CoV-2



Figure 6 heavy and kappa light chains looped - final predicted structure (CR5840)

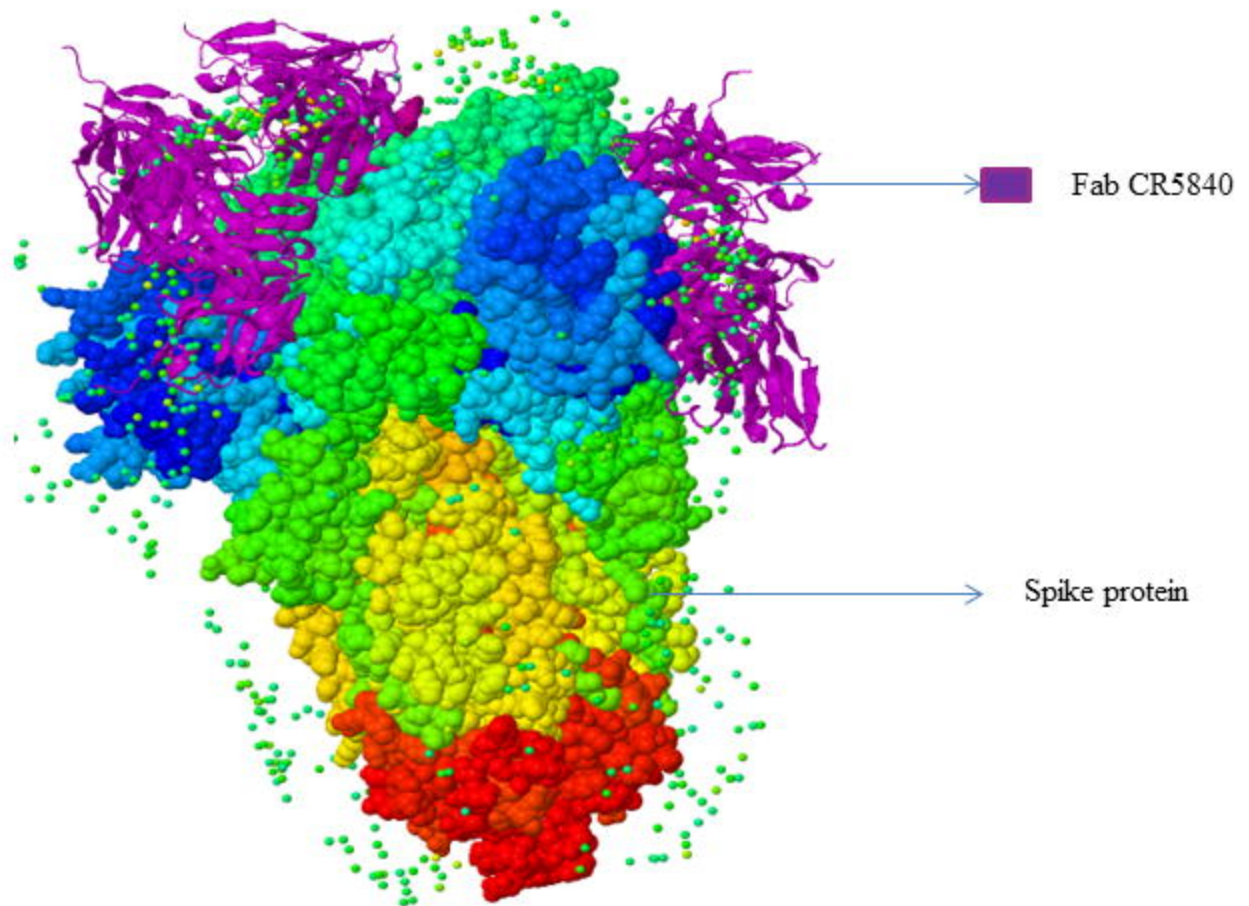


Figure 7 synthetic hybrid antibody confirmatory structures and interaction with spike glycoprotein of SARS-CoV-2

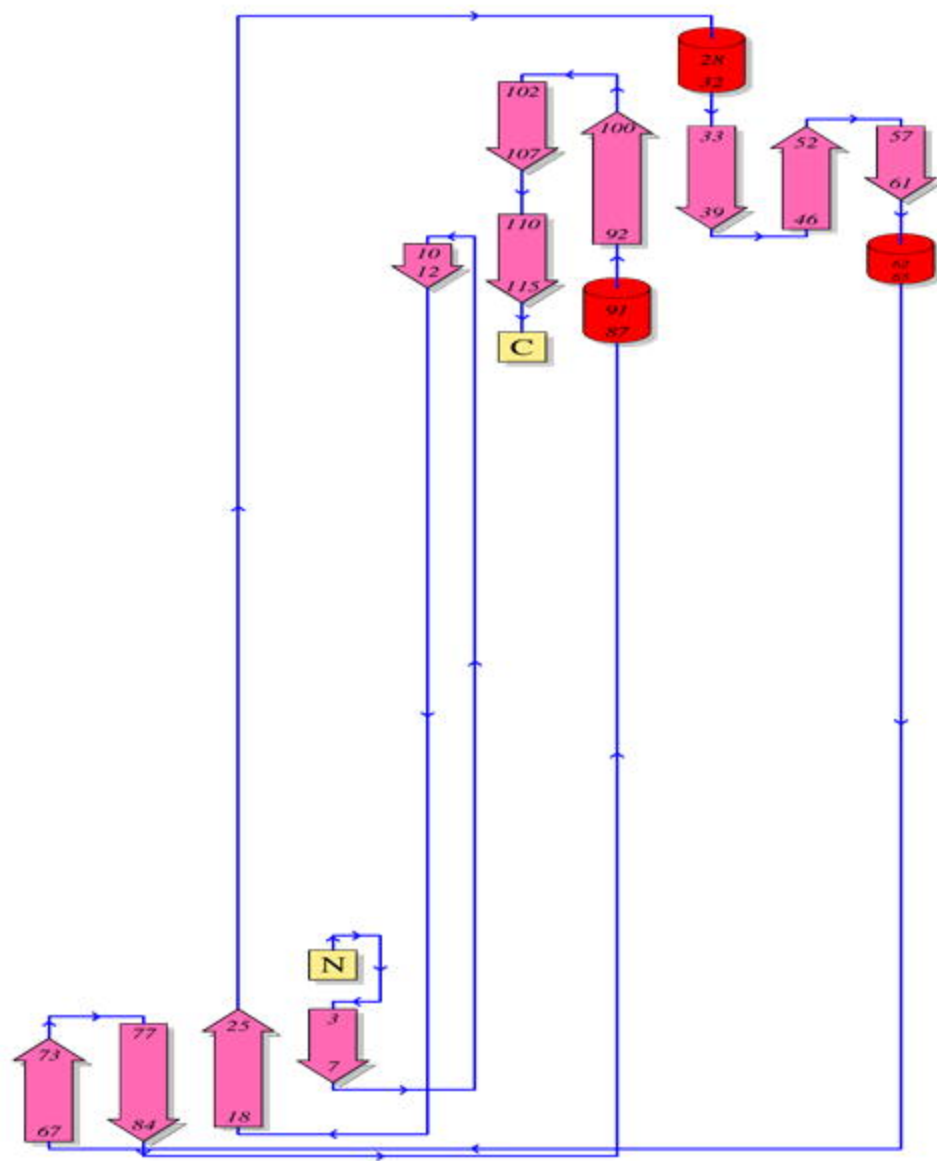


Figure 8 Topology of heavy and kappa light chain looping

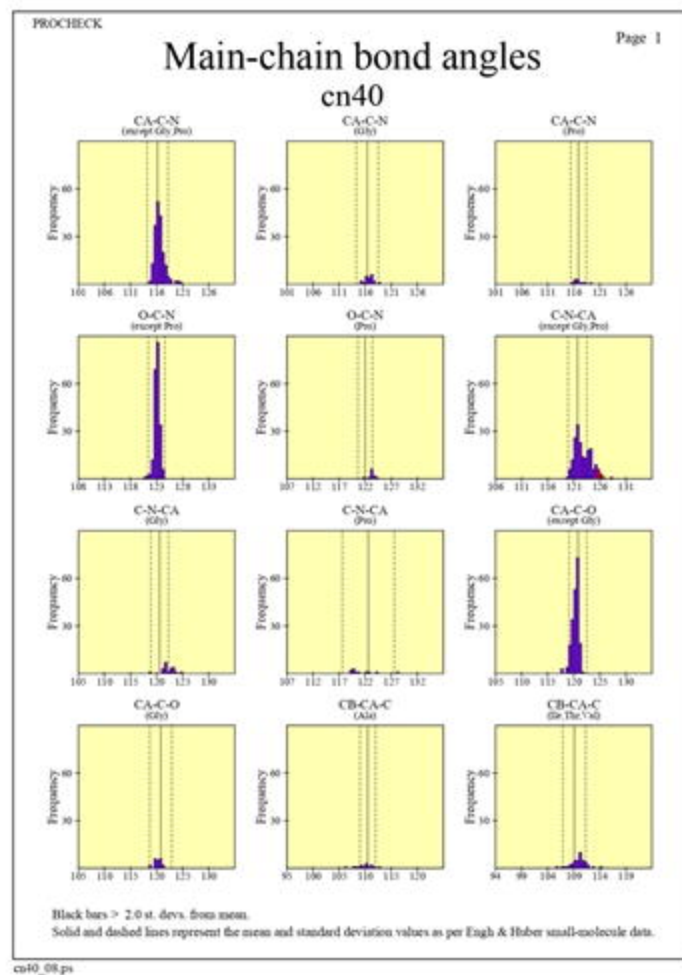
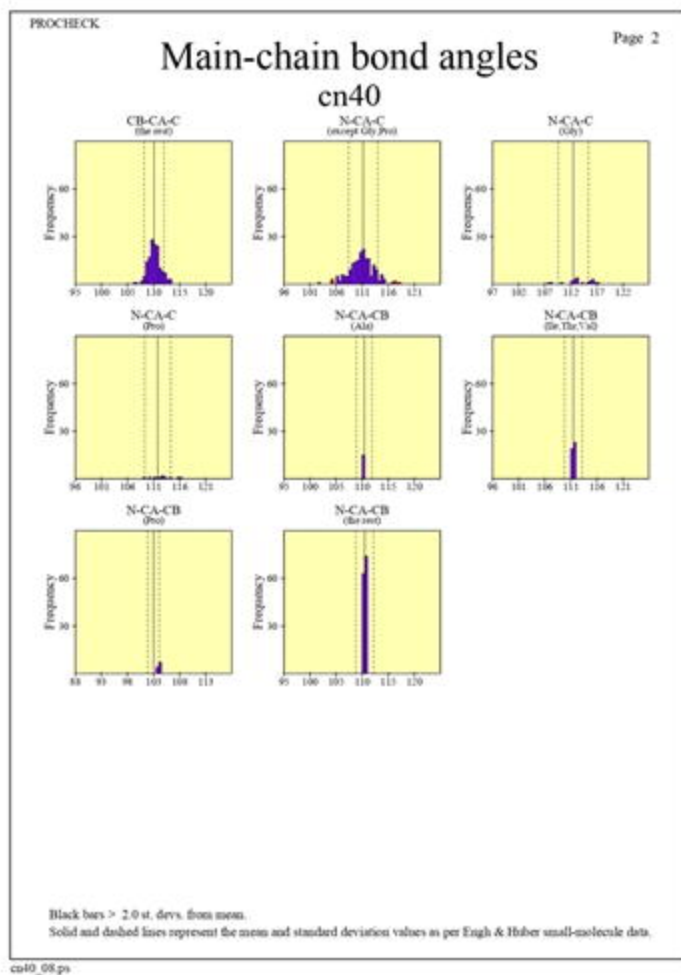
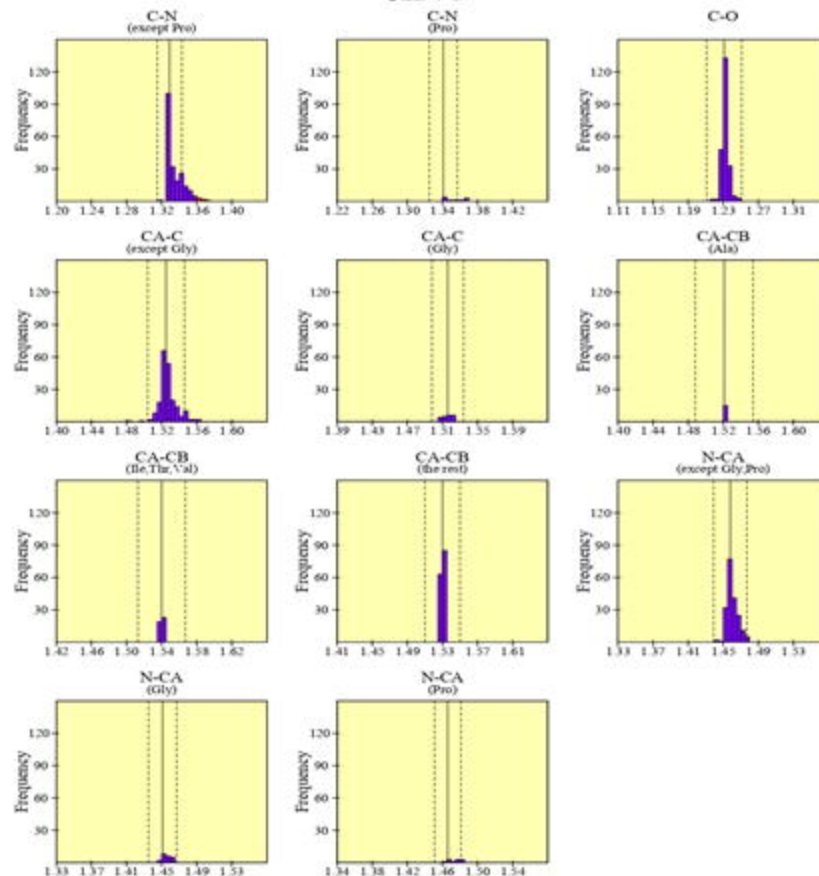


Figure 9 Main chain bond angles of CR5840 Fab region

Main-chain bond lengths cn40

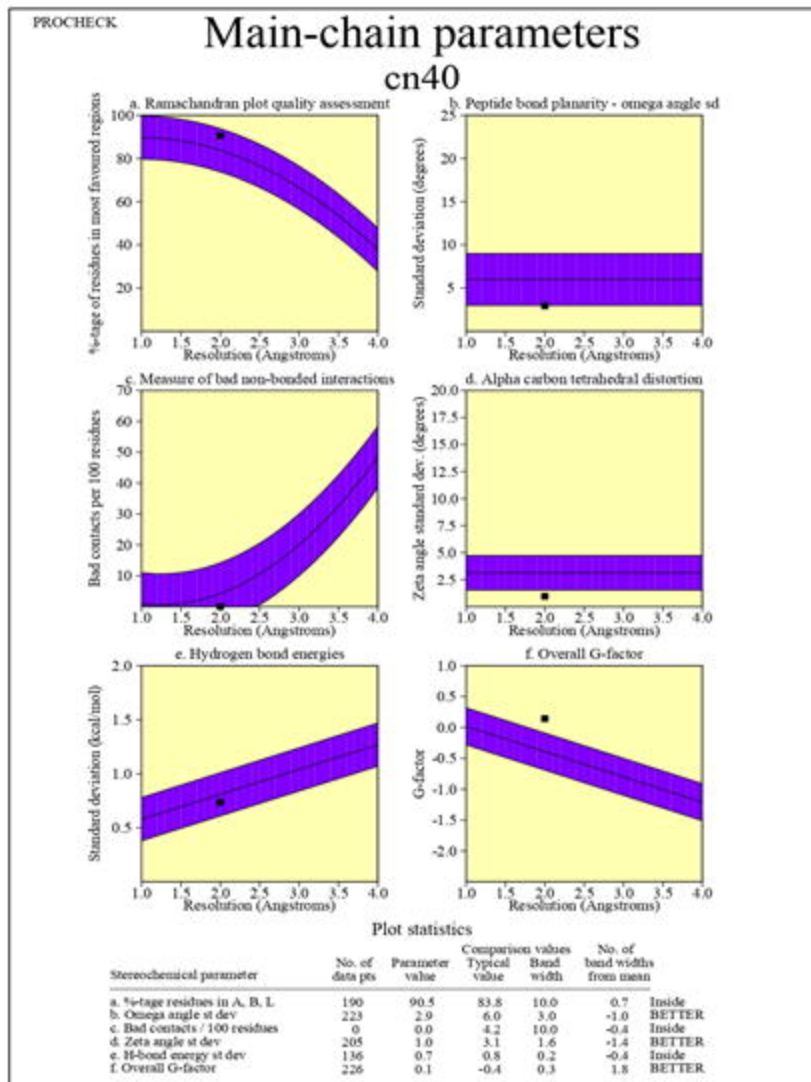


Black bars > 2.0 st. devs. from mean.

Solid and dashed lines represent the mean and standard deviation values as per Engh & Huber small-molecule data.

cn40_07.ps

Figure 10 Main chain bond Lengths of CR5840 Fab region



cn40_04.ps

Figure 11 Main chain Parameters of CR5840 Fab region

Protein - Protein interface of Heavy chain and Light Kappa chain of FAB region

Chain H Chain I

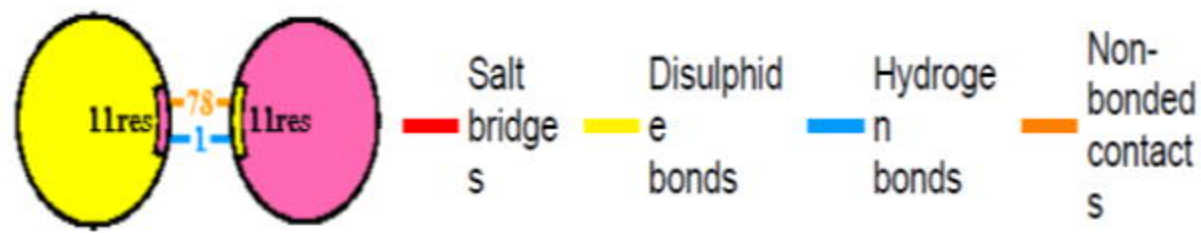


Figure 12 Protein-Protein interface of Heavy chain and Light Kappa chain Fab region

Residue interactions across interface
Coloured by residue type



Figure 13 Residue interactions across interface of heavy and light kappa

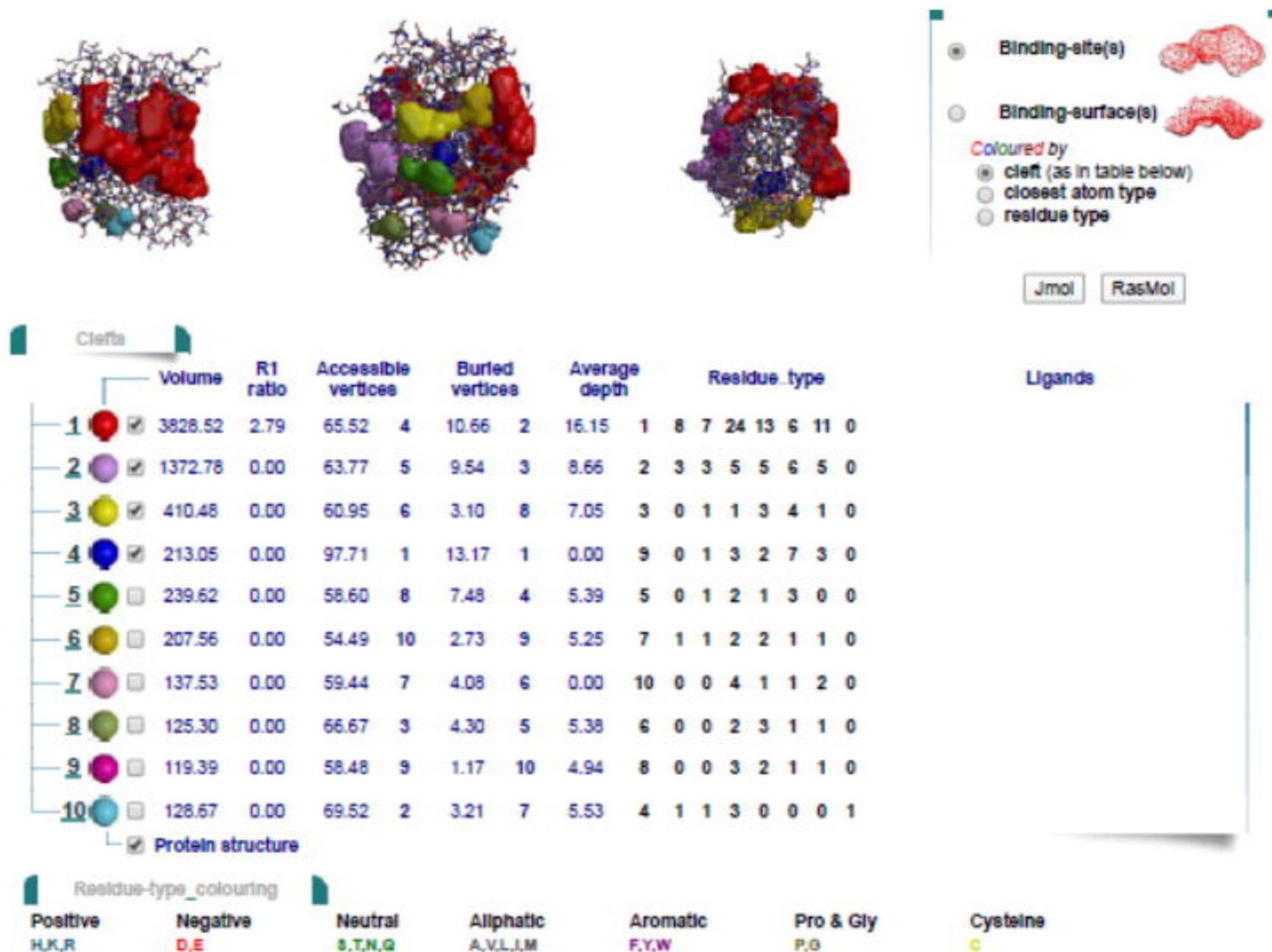


Figure 14 Cleft analysis

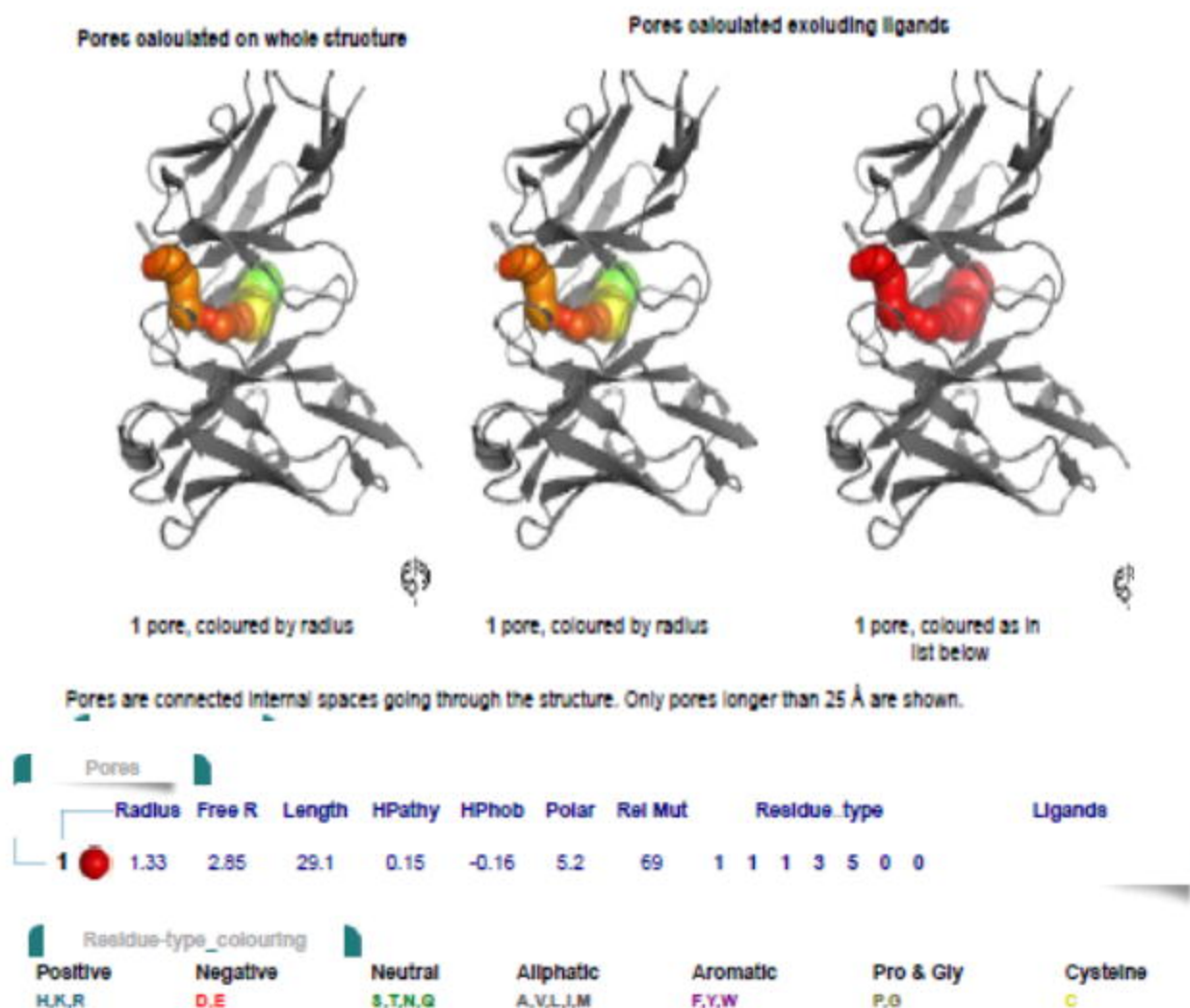


Figure 15 Pores analysis

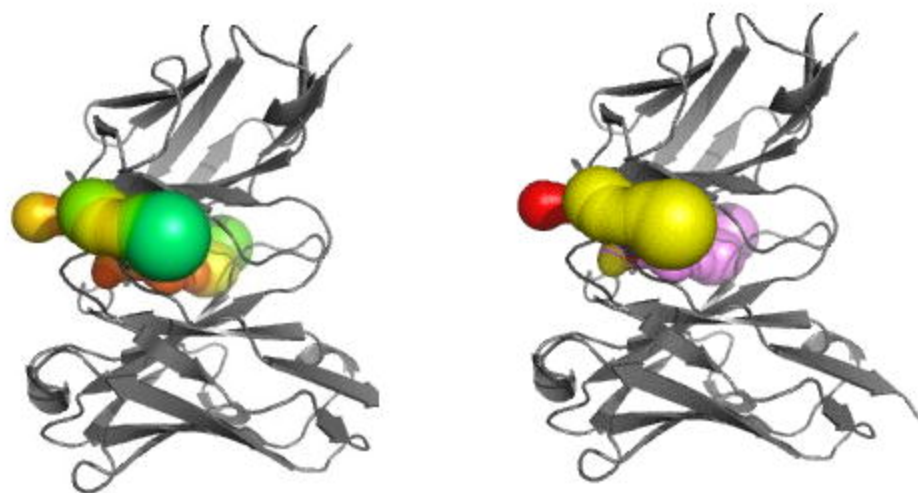


Figure 16 Tunnels in Fab chain

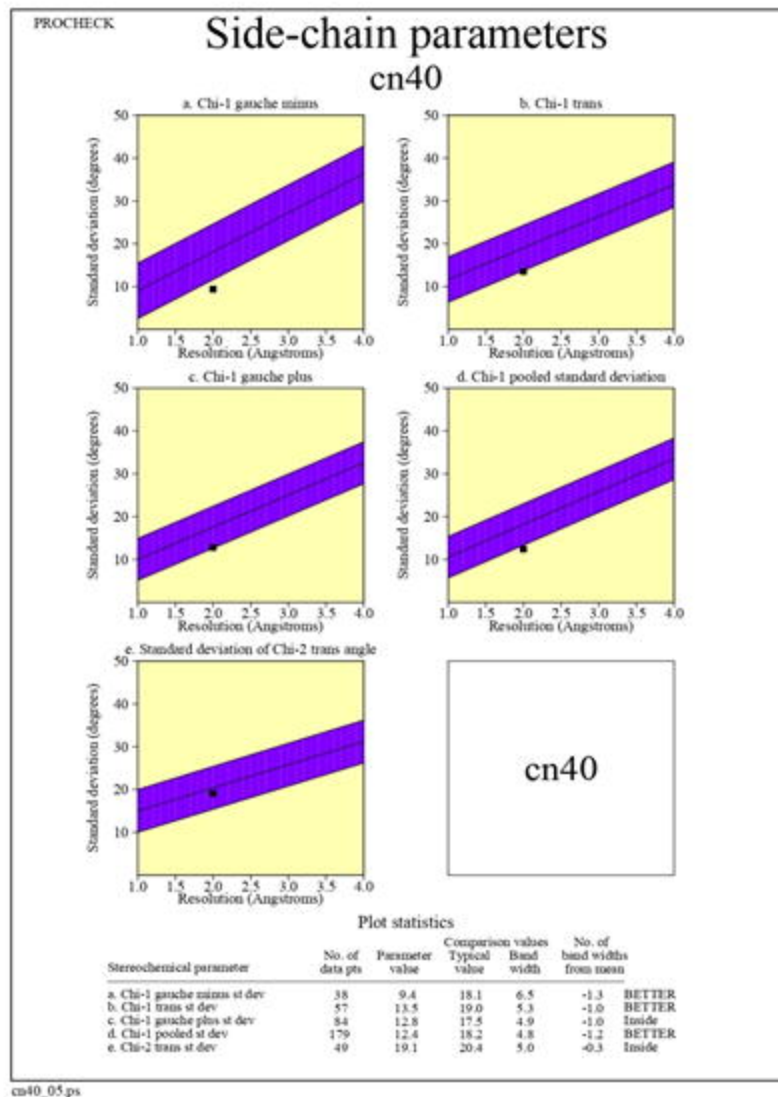


Figure 17 Side Chain parameters of Fab chain CR5840

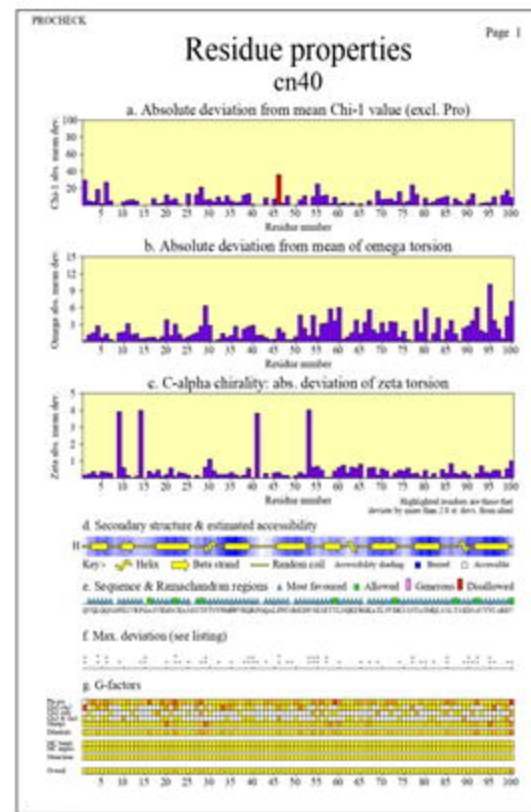
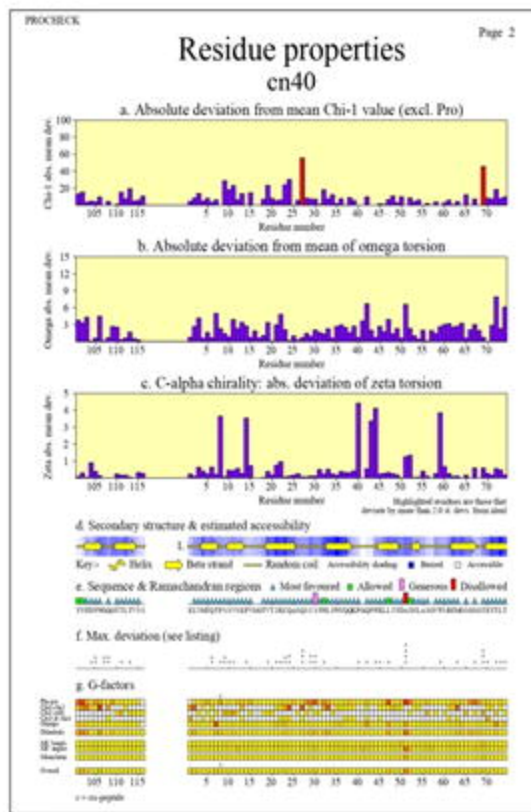
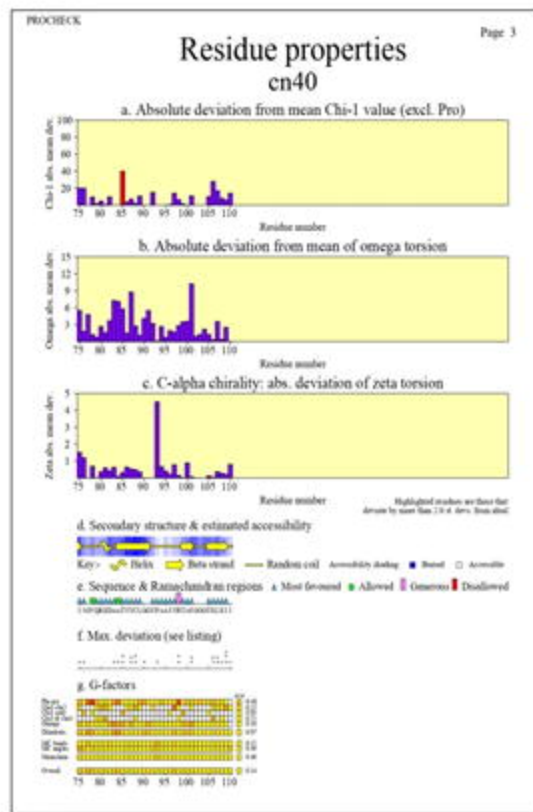
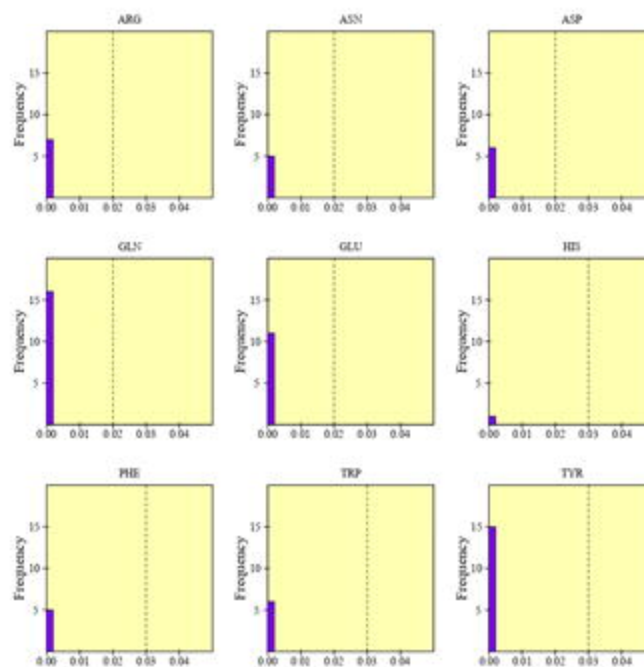


Figure 18 Residue properties of Fab chain CR5840

RMS distances from planarity cn40



Histograms showing RMS distances of planar atoms from best-fit plane.
Black bars indicate large deviations from planarity: RMS dist > 0.03 for rings, and > 0.02 otherwise.

cn40_09.ps

Figure 19 RMS distance of planarity of Fab chain CR5840

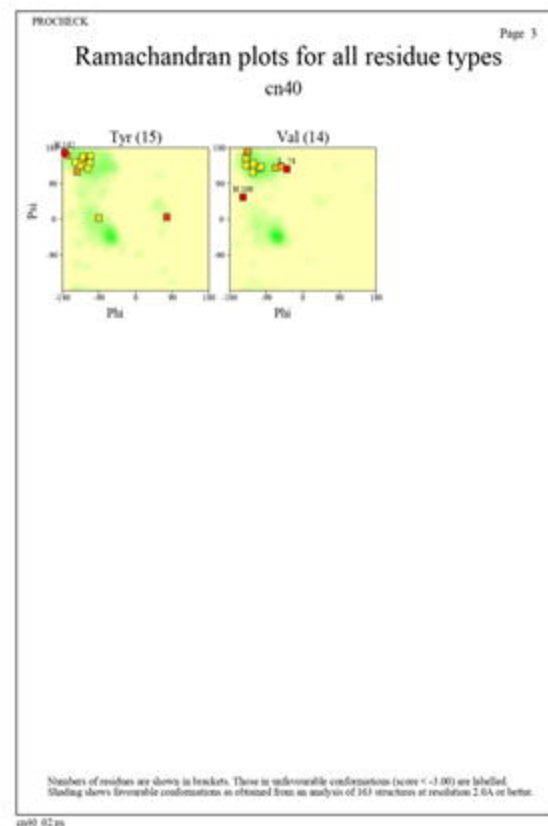
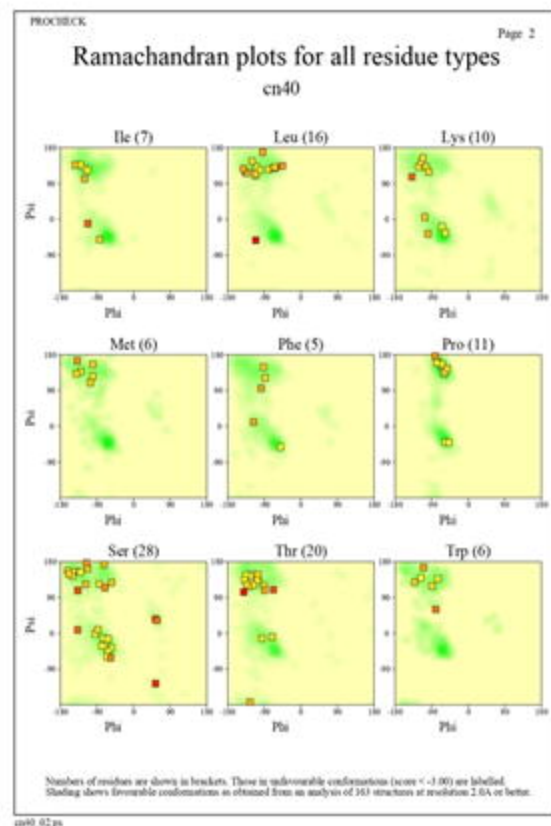
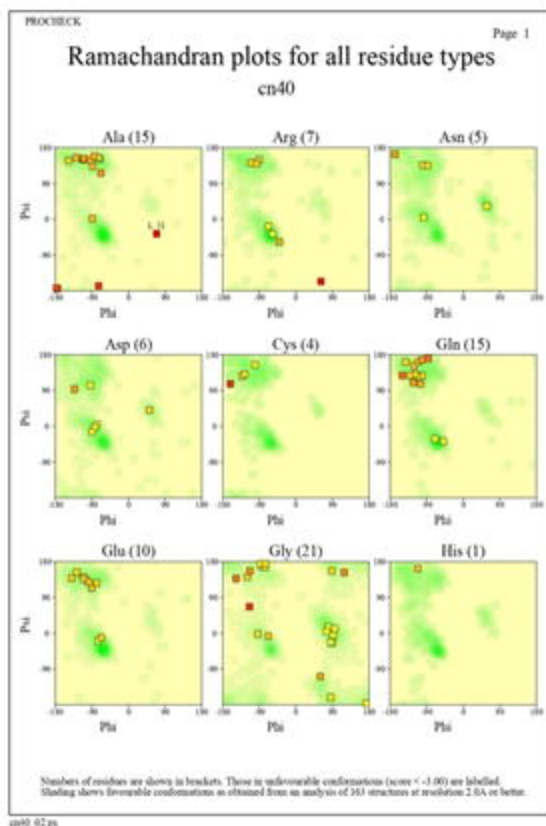


Figure 20 Ramachandran plots for all residue types of Fab chain CR5840

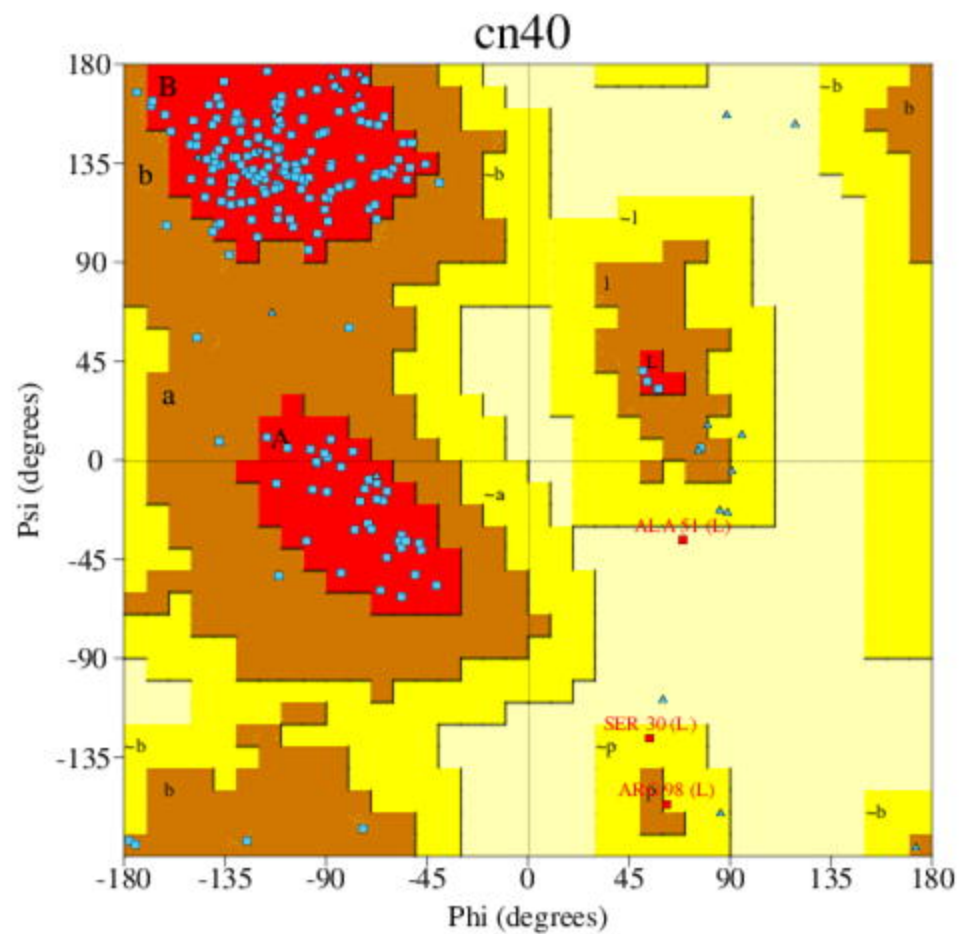


Figure 21 Ramachandran plots - Fab chain CR5840

- [NCBI Home](#)
- [Sign in to NCBI](#)
- [Skip to Main Content](#)
- [Skip to Navigation](#)
- [About NCBI Accesskeys](#)



[U.S. National Library of Medicine](#)



[NCBI National Center for Biotechnology Information](#)

-
- [My NCBI](#)
- [Sign in to NCBI](#)
- [Register](#)
- [Sign Out](#)

[IGBLAST](#)» JOB ID: v6FUBaoHEysgiyS9faYH23yvAspj

Formatting Results

Database: imgt.Homo_sapiens.V.f.orf.p; imgt.Homo_sapiens.D.f.orf;
imgt.Homo_sapiens.J.f.orf
597 sequences; 157,743 total letters

Query= MN908947.3

Length=29903

Sequences producing significant alignments:

	Score (Bits)	E Value
IGHV1-38-4*01 germlinegene	29.9	4.1
IGHV2-70*19 germlinegene	29.9	4.1
IGLV1-50*01 germlinegene	29.9	4.1
IGHD6-13*01 germlinegene	14.1	636
IGHD2-15*01 germlinegene	14.1	636
IGHD2-2*01 germlinegene	14.1	636

Domain classification requested: imgt

Note that your query represents the minus strand of a V gene and has been converted to the plus strand. The sequence positions refer to the converted sequence.

V-(D)-J rearrangement summary for query sequence (multiple equivalent top matches, if present, are separated by a comma):

Top V gene match	Top D gene match	Top J gene match	Chain type	stop codon	V-J frame	Productive	Strand
IGHV1-38-4*01	IGHD6-13*01	N/A	VH	Yes		No	-

V-(D)-J junction details based on top germline gene matches:

V region end	V-D junction*	D region	D-Jjunction*	J region start
TTTAT	CATTGTAGATGTCAAAGCCCTGTATACGACATCAGTACTAGTGCCTGTGCCGCACGGTGTAAAGACGGGCTGCACTTACACCGCAAACCCGTTTAAAAACGATTGT	GCATCAGCTG		

*: Overlapping nucleotides may exist at V-D-J junction (i.e., nucleotides that could be assigned to either rearranging gene). Such nucleotides are indicated inside a parenthesis (i.e., (TACAT)) but are not included under the V, D or J gene itself.

Alignment summary between query and top germline V gene hit:

	from	to	length	matches	mismatches	gaps	identity(%)
FR1-IMGT	16310	16319	10	8	2	0	80
CDR1-IMGT	16320	16343	24	17	7	0	70.8
FR2-IMGT	16344	16345	2	2	0	0	100
Total			36	27	9	0	75

Alignments

lcl|Query_1_reversed 11549 GTAAATTGGTACCAACAGCTTCTCTAGTAGCATGACACCCCTCGACATCGAAGCCAATCCATGCACGTACATGTCTTATAGCTTCTTCGC 11638
V 90.9% (20/22) [IGLV1-50*01](#) 103 ...C.....G.....----- 124

lcl|Query_1_reversed 11639 GGGTGATAAACATGTTAGGGTAACCATTAACTTGATAATTCATTTTAAACCCATCATAGAGATGAGTCTTCTATAGGTCATGTCCTTAG 11728

lcl|Query_1_reversed 11729 GTATGCCAGGTATGTCAACACATAAACCTTCAGTTTTGAATTTAGTGTCAACACTGAGGTGTGTAGGTGCCTGTGTAGGATGTAACCCAG 11818

lcl|Query_1_reversed 11819 TGATTACCTTACTACAATCTTTAAAGAGTCCTGTTACATTTTCAGCTTGTAAGTTGCCACATTCCTACGTGGAATTTCAAGACTTGTA 11908

lcl|Query_1_reversed 11909 ATTGCAACTTGTCAATAAGGTCTCTATCAGACATTATGCAAAGTATGCCTACTTTTGTCTGGTAATAGCAACATTAAATCTGTTTAC AT 11998

lcl Query_1_reversed	11999	TACAAGAGTGAGCTGTTTCAGTGGTTTGAGTGAATATGACATAGTCATATTCTGAGCCCTGTGATGAATCAACAGTTTGAGTTGGTAGTC	12088
lcl Query_1_reversed	12089	CCAAAATCTTTGAGGCTACAGCATTCTGTGAATTATAAGGTGAAATAAAGACAGCTTTTCTCCAAGCAGGGTTACGTGTAAGGAATTCTC	12178
lcl Query_1_reversed	12179	TTACCACGCCTATTTGTGGCCTGTTAATTGCAGATGAAACATCATGCGTGATAACACCCTTATAAAACATTTTAAAGCATTGAGCTGATT	12268
lcl Query_1_reversed	12269	TGTCTTTATGTGCTTTAAGCTTATTATCATAAACCAAAGCACTCACAGTGTCAACAATTTTCAGCAGGACAACGCCGACAAGTTCCGAGGA	12358
lcl Query_1_reversed	12359	ACATGTCTGGACCTATAGTTTTTCATAAGTCTACACACTGAATTGAAATATTCTGGTTCTAGTGTGCCCTTAGTTAGCAATGTGCGTGGTG	12448
lcl Query_1_reversed	12449	CAGGTAATTGAGCAGGGTCGCCAATGTACACATAGTGCTTAGCACGTAATCTGGCATTGACAACACTCAAATCATAATTTGTGGCCATTG	12538
lcl Query_1_reversed	12539	AAATTCATCAAAGACAACCTATATCTGCTGTCGTCTCAGGCAATGCATTTACAGTACAAAAGACATACTGTTCTAATGTTGAATTCACCT	12628
lcl Query_1_reversed	12629	TGAATTTATCAAAACACTCTACACGAGCACGTGCAGGTATAATTCTACTACATTTATCTATAGGCAAATATTTTAATGCCTTCTCACA TA	12718
lcl Query_1_reversed	12719	GTGCATCAACAGCGGCATGAGAGCAAGCTGTATACACTATGCGAGCAGAAGGGTAGTAGAGAGCTAGGCCAATAGCAAAATGACTCTTAC	12808
lcl Query_1_reversed	12809	CAGTACCAGGTGGTCCCTGGAGTGTAGAATACTTTTGCATACCAACCTTTTGATAATTTGCAACATTGCTAGAAAACCTCATCTGAGATAT	12898
lcl Query_1_reversed	12899	TGAGTGTTGGGTATAAGCCAGTAATTCTAACATAGTGCTCTTGTGGCACTAGTGTAGGTGCACTTAATGGCATTACTGTATGTGATGTCA	12988
lcl Query_1_reversed	12989	GCACAAAATAATCACCAACATTTAATTTGTAAGTTGTTGTACCTCGGTAAACAACAGCATCACCATAGTCACCTTTTTCAAAGGTGTAC T	13078
lcl Query_1_reversed	13079	CTCCTATTTGTACTTTACTGTTTTTAGTTACACGATAACCAGTAAAGACATAATTTTCGGTTAAGTGGTGGTCTAGGTTTACCAACTTCCC	13168

lc Query_1_reversed	13169	ATGAAAGATGTAATTCTCTGTCAGACAGCACTTCACGTACAGTAGCAATACCATAAGACAGTTTAAATGTCTCCTCAGTAGCTTTGAGCG	13258
lc Query_1_reversed	13259	TTTCTGCTGCAAAAAGCTTGAGTCTTTCAGTACAGGTGTTAGCTAAAATGTAATCACCAGCATTTGTCCAGTCACATGTTGCAATTGCAT	13348
lc Query_1_reversed	13349	TAAAGTCAGTAACATTATCGCTACCAACACATGTATTTTTATATAAACCAAAACTTGTCATTAGCACACAATGGAAACTAATGGGTG	13438
lc Query_1_reversed	13439	GTTTATGTGATTTACAATAATAGCTCATACCTCCTAAGTAAAGTTGAGTCACATCTGTGACATCACAACCTGGAGCATTGCAAACATACG	13528
lc Query_1_reversed	13529	GATTAACAGACAAGACTAATTTATGTGATGTTGATATGACATGGTCGTAACAGCATTTACAACATAAGAATGGTCTACGTATGCAAGCAC	13618
lc Query_1_reversed	13619	CACATCTTAATGAAGTCTGTGAATTGCAAAGAACACAAGCCCCAACAGCCTGTAAGACTGTATGCGGTGTGTACATAGCCTCATAAACT	13708
lc Query_1_reversed	13709	CAGGTTCCCAATACCTTGAAGTGTATCATTAGTAAGCATAACAGAATACATGTCTAACATGTGTCCTGTAACTCATCATGTAGCTTTC	13798
lc Query_1_reversed	13799	TTATGTATTGTAAGTACAAATGAAAGACATCAGCATACTCCTGATTAGGATGTTTAGTAAGTGGGTAAAGCATCTATAGCTAAAGACACGA	13888
lc Query_1_reversed	13889	ACCGTTCAATCATAAGTGTACCATCTGTTTTACGATATCATCTACAAAACAGCCGGCCCCCTAGGATTCTTGATGGATCTGGGTAAGGAA	13978
lc Query_1_reversed	13979	GGTACACATAATCATCACCTGTTTAACTAGCATTGTATGTTGAGAGCAAAATTCATGAGGTCCTTTAGTAAGGTCAGTCTCAGTCCAAC	14068
lc Query_1_reversed	14069	ATTTTGCTTCAGACATAAAAACATTGTTTTGATAATAAAGAACTGACTTAAAGTTCTTTATGCTAGCCACTAGACCTTGAGATGCATAAG	14158
lc Query_1_reversed	14159	TGCTATTGAAACACACAACAGCATCGTCAGAGAGTATCATCATTGAGAAATGTTTACGCAAATATGCGTAAAACTCATTACAAAAGTCTG	14248
lc Query_1_reversed	14249	TGTCAACATCTCTATTTCTATAGAGACACTCATAAAGTCTGTGTTGTAAATTGCGGACATACTTATCGGCAATTTTGTACCATCAGTAG	14338

lcl|Query_1_reversed 14339 ATAAAAGTGCATTAAACATTGGCCGTGACAGCTTGACAAATGTTAAAAACACTATTAGCATAAGCAGTTGTGGCATCTCCTGATGAGGTTC 14428

lcl|Query_1_reversed 14429 CACCTGGTTTAAACATATAGTGAACCGCCACACATGACCATTTCACTCAATACTTGAGCACACTCATTAGCTAATCTATAGAAACGGTGTG 14518

lcl|Query_1_reversed 14519 ACAAGCTACAACACGTTGTATGTTTGCGAGCAAGAACAAGTGAGGCCATAATTCTAAGCATGTTAGGCATGGCTCTATCACATTTAGGA T 14608

lcl|Query_1_reversed 14609 AATCCCAACCCATAAGGTGAGGGTTTTCTACATCACTATAAACAGTTTTTAACATGTTGTGCCAACCACCATAGAATTTGCTTGTTCOA 14698

lcl|Query_1_reversed 14699 TTACTACAGTAGCTCCTCTAGTGGCGGCTATTGATTTCAATAATTTTTGATGAACTGTCTATTGGTCATAGTACTACAGATAGAGACAC 14788

lcl|Query_1_reversed 14789 CAGCTACGGTGCGAGCTCTATTCTTTGCACTAATGGCATACTTAAGATTCATTTGAGTTATAGTAGGGATGACATTACGTTTTGTATATG 14878

lcl|Query_1_reversed 14879 CGAAAAGTGCATCTTGATCCTCATAACTCATTGAATCATAATAAAGTCTAGCCTTACCCATTTATTAATGGAAAACCAGCTGATTTGT 14968

lcl|Query_1_reversed 14969 CTAGGTTGTTGACGATGACTTGTTAGCATTAAACAGCCACCATCGTAACAATCAAAGTACTTATCAACAACCTCAACTACAAATAGTA 15058

lcl|Query_1_reversed 15059 GTTGTCTGATATCACACATTGTTGGTAGATTATAACGATAGTAGTCATAATCGCTGATAGCAGCATTACCATCCTGAGCAAAGAAGAAGT 15148

lcl|Query_1_reversed 15149 GTTTTAATTCAACAGAACTTCCTTCCTTAAAGAAACCCCTAGACACAGCAAAGTCATAGAAGTCTTTGTTAAAATTACCGGGT TTGACAG 15238

lcl|Query_1_reversed 15239 TTTGAAAAGCAACATTGTTAGTAAGTGCAGCTACTGAAAAGCACGTAGTGC GTTTATCTAGTAATAGATTACCAGAAGCAGCGTGCATAG 15328

lcl|Query_1_reversed 15329 CAGGGTCAGCAGCATACACAAGTAATTCCTTAAACTAAGTCTAGAGCTATGTAAGTTTACATCCTGATTATGTACAACACCTAGCTCTC 15418

lcl|Query_1_reversed 15419 TGAAGTGGTATCCAGTTGAACTACAAATGGAACACCATCAACAAATATTTTTCTCACTAGTGGTCCAAAACCTGTAGGTGGGAACACTG 15508

Ic|Query_1_reversed 15509 TAGAGAATAAAACATTAAAGTTTGCACAATGCAGAATGCATCTGTCATCCAAACAGTTAACACAATTTGGGTGGTATGTCTGATCCCAAT 15598

Ic|Query_1_reversed 15599 ATTTAAATAACGGTCAAAGAGTTTTAACCTCTCTCCGTGAAGTCATATTTTAACAAATCCCACTTAATGTAAGGCTTTGTTAAGT CAG 15688

Ic|Query_1_reversed 15689 TGTCAACATGTGACTCTGCAGTTAAAGCCCTGGTCAAGGTTAATATAGGCATTAACAATGAATAATAAGAATCTACAACAGGAACTCCAC 15778

Ic|Query_1_reversed 15779 TACCTGGCGTGGTTTGTATGAAATCACCGAAATCATACCAGTTACCATTGAGATCTTGATTATCTAATGTCAGTACACCAACAATACCAG 15868

Ic|Query_1_reversed 15869 CATTTGCGATGGCATCACAGAATTGTACTGTTTTTAACAAAGCTTGGCGTACACGTTACCTAAGTTGGCGTATACGCGTAATATATCTG 15958

Ic|Query_1_reversed 15959 GGTTTTCTACAAAATCATACCAGTCCTTTTTATTGAAATAATCATCATCACAACAATTGTATGTGACAAGTATTTCTTTTAATGTGTCAC 16048

Ic|Query_1_reversed 16049 AATTACCTTCATCAAAATGCCTTAAAGCATAGACGAGGTCTGCCATTGTGTATTTAGTAAGACGTTGACGTGATATATGTGGTACCATGT 16138

Ic|Query_1_reversed 16139 CACCGTCTATTCTAAACTTAAAGAAGTCATGTTTAGCAACAGCTGGACAATCCTTAAGTAAATTATAAATTGTTTCTTCATGTTGGTAGT 16228

<-----FR1-IMGT----->

SF*

Ic|Query_1_reversed 16229 TAGAGAAAGTGTGTCTCTTAACCTACAAAGTAAGAATCAATTAAATTGTCATCTTCGTCCTTTTCTTGGAAGCGACAACAATTAGTTTTTA 16318
 V 75.0% (27/36) [IGHV1-38-4*01](#) 66 ----- C 74
SF S

><-----CDR1-IMGT-----><-----FR2-IMGT----->

EFSKTSYFIIVDVKSPVYDI STSACAARCK

Ic|Query_1_reversed 16319 GGAATTTAGCAAAACCAGCTACTTTATCATTGTAGATGTCAAAAGCCCTGTATACGACATCAGTACTAGTGCCTGTGCCGACGGTGTA 16408
 V 75.0% (27/36) [IGHV1-38-4*01](#) 75 T.GG...C..TC.....GG... 101
GFTITSYG

D 100.0% (7/7) [IGHD2-15*01](#) 6 ----- 12
 D 100.0% (7/7) [IGHD2-2*01](#) 6 ----- 12

TGCTYTANPFKND CASAD* SMGSR S* SQLQ

Ic|Query_1_reversed 16409 GACGGGCTGCACTTACACGCAAAACCGTTTAAAAACGATTGTGCATCAGCTGACTGAAGCATGGGTTCCGCGAGTTGATCACAAC TACA 16498
 D 90.0% (9/10) [IGHD6-13*01](#) 8 ----- ...G..... 17

P*PFHI PQTVQTVFLSVKPTGSLAQVVGIC

Icd Query_1_reversed	16499	GCCATAACCTTTCCACATACCGCAGACGGTACAGACTGTGTTTTTAAGTGTA AACCCACAGGGTCATTAGCACAGTTGTAGGTATTTG	16588
Icd Query_1_reversed	16589	TYLPFKSQNPLGFG*SMWQRQYRQHDAPPK TACATACTTACCTTTAAGTCACAAAATCCTTTAGGATTTGGATGATCTATGTGGCAACGGCAGTACAGACAACACGATGCACCACAAA	16678
Icd Query_1_reversed	16679	DS*SI LASGVTVIA*PVPVVCVHNI LTQLVI GGATTCTTGATCCATATTGGCTTCCGGTGTA ACTGTTATTGCCTGACCAGTACCAGTGTGTGTACACAACATCTTAACACAATTAGTGAT	16768
Icd Query_1_reversed	16769	GCPPLAR*SL*ALAASTAKAQKDNTVELAG TGGTTGTCCCCACTAGCTAGATAATCTTTGAAGCTTTAGCAGCATCTACAGCAAAAGCACAGAAAGATAATACAGTTGAATTGGCAGG	16858
Icd Query_1_reversed	16859	TSVALPACRRRTVAAKLPSTI PLFRLENPLI CACTTCTGTTGCATTACCAGCTGTAGACGTACTGTGGCAGCTAACTACCAAGTACCATACCTCTATTTAGGTTGTTAATCCTTTAAT	16948
Icd Query_1_reversed	16949	KYKYFTLGPLGVSVTNLQGGSSSV*IVPVP AAAGTATAAATACTTCACTTTAGGACCTTTAGGTGTGTCTGTAAACAACTACAAGGTGGTTCCAGTTCTGTATAGATAGTACCAGTTCC	17038
Icd Query_1_reversed	17039	SLLGNLAHFKSKSDNSASTNLPPFVVL** ATCACTCTTAGGGAATCTAGCCATTTCAAATCCTGTAAATCGGATAACAGTGCAAGTACAAACCTACCTCCCTTTGTTGTGTGTAGTA	17128
Icd Query_1_reversed	17129	ANALSSVQAVCVVPAAQD I CRSATGLSSLF AGCTAACGCATTGTCATCAGTGCAAGCAGTTTGTGTAGTACCGGCAGCACAGACATCTGTCGTAGTGCAACAGGACTAAGCTCATTATT	17218
Icd Query_1_reversed	17219	CNLTAELALKAVTI RGHAKLGELS I L I SLS CTGTAATTTGACAGCAGAATTGGCCCTTAAAGCTGTTACAATAAGAGGCCATGCTAAATTAGGTGAATTGTCCATACTAATTTCACTAAG	17308
Icd Query_1_reversed	17309	*T I LLSASTTCW I SHNADA*VNVVPSHVFL TTGAACAATTTTACTATCTGCATCTACAACCTGTTGGATTTCCACAATGCTGATGCATAAGTAAATGTTGTACCATCACACGTATTTTT	17398
Icd Query_1_reversed	17399	YVL*SGMTTI SLAAVVRG I MFKGTQPSLAL ATATGTGTTATAGTCTGGTATGACAACCATTAGTTTGGCTGCTGTTGTAAGAGGTATTATGTTCAAGGGAACACAACCATCTCTGCATT	17488
Icd Query_1_reversed	17489	L I MLLSASLSNFLS I VKS I VCI ALVTFALL GTTGATAATGTTGTTGAGTGCATCATTATCCAACCTTTCTAAGCATAGTGAAAAGCATTGTCTGCATAGCACTAGTAACCTTTGCCCTCTT	17578
Icd Query_1_reversed	17579	SSDLACL Y I WV I A*SA I FSNLRCMAASRSN GTCCTCAGATCTAGCCTGTTTATACATTTGGGTCATAGCTTGATCAGCCATCTTTTCCAACCTACGTTGCATGGCTGCATCACGGTCAAA	17668

SDLATFKDFFNFLRTTSESPLATACS*AS*

H*VLWTKTKSEVKI VSNNQWCVP LI VLFTA

Ic Query_1_reversed	18839	CCATTGAGTACTCTGGACTAAACTAAAAGTGAAGTCAAAATTGTGAGTAACAACCAGTGGTGTGTACCCTTGATTGTTCTTTTCACTGC	18928
Ic Query_1_reversed	18929	LWKVTPEHCLTTSKGVNSSSNKALPNMVRP ACTTTGGAAAGTAACACCTGAGCATTGTCTAACACATCAAAAGGTGTAAATTCATCTTCTAATAAAGCACTACCCAATATGGTACGTCC	19018
Ic Query_1_reversed	19019	FIPFCSNSFNEAHISKTAIPV*AERGPSMS ATTCATACCATTTTGCAGTAATTCTTTTAAATGAAGCACACATATCTAAACGGCAATCCAGTTTGAGCAGAAAGAGGTCCTAGTATGTC	19108
Ic Query_1_reversed	19109	TWSCVRGS*LYFIATRLKSLRVVVNRLRNH AACATGGTCTTGTTAGAGGTTACATAATTGACTTCATAGCCACAAGGTTAAAGTCATTAAGAGTTGTGGTAAATCGATTGAGAAACCA	19198
Ic Query_1_reversed	19199	LSPFI TAAYNQAKTLTVI VVSVPAAACAVCL CCTGTCTCCATTTATAACAGCAGCGTACAACCAAGCTAAACATTAAGTGAATAGTTGTGCCGTACCAGCTGCTTGTGCTGTTGCCT	19288
Ic Query_1_reversed	19289	STKGP*KLPSKSVPA*TPVGNSI WCM*QKE GTCAACAAAAGGTCCATAAAAGTTACCTTCTAAGTCTGTGCCAGCATGAAGTCCAGTTGGTAATTCATATGGTGCATGTAACAAAAGA	19378
Ic Query_1_reversed	19379	TQS*SMLKPTLPHEPLRNEPLI VKLGLI AH GACACAGTCATAATCTATGTTAAACCAACACTACCACATGAACCATTAAGGAATGAACCCCTAATAGTGAATTTGGGCCTCATAGCACA	19468
Ic Query_1_reversed	19469	W*TPDGEPL*QANTEKVC PG*MRTNLYLG V TTGGTAAACACCAGATGGTGAACCATTTGAACAAGCTAACACTGAAAAAGTCTGTCCTGGTTGAATGCGAACAACCTATACCTAGGTGT	19558
Ic Query_1_reversed	19559	LGLAVSTLSLSTQFCIECPI TLS*TL PACT CTTAGGATTGGCTGTATCAACCTTAAGCTTAAGTACACAATTTTCATAGAATGTCCAATAACCTGAGTTGAACATTACCAGCCTGTAC	19648
Ic Query_1_reversed	19649	KKL*LDLRMSKSS*LGLSMSSEVQITCLGQ CAAGAAATTATGATTAGACTTACGAATGAGTAAATCTTCATAATTAGGGTTAAGCATGTCTTCAGAGGTGCAGATCACATGTCTTGACA	19738
Ic Query_1_reversed	19739	*TTSSSRPLSVVVPQVTCTIQPSTLPDGN GTAAACTACGTCATCAAGCCAAAGACCGTTAAGTGTAGTTGTACCACAAGTACTTGTACCATACAACCTCAACTTTACCAGATGGGAA	19828
Ic Query_1_reversed	19829	AIFLKPLCKTAEVIEVCGGW*RTSEPELLK TGCCATTTTCTAAAACCACTCTGCAAAACAGCTGAGGTGATAGAGGTTTGTGGTGGTTGGTAAAGAACATCAGAACCTGAGTTACTGAA	19918
Ic Query_1_reversed	19919	SLRAFAR*QQAASL*LVVSI APLKYLYLL* GTCATTGAGAGCCTTTGCGAGATGACAACAAGCAGCTTCTCTGTAGCTAGTTGTATCCATTGCTCCACTAAAATACTTGTACTTATTATA	20008

RAKYLLYCVRGNSTSLRNFRYISLFNKKVH

Ic Query_1_reversed	20009	AAGAGCTAAGTATCTATTATATTGCGTAAGAGGTAATAGCACATCACTACGCAACTTTAGATACATTTCTTTATTTAACAAAAAGGTGCA	20098
Ic Query_1_reversed	20099	SAASSKVLKETPLKTTTRLFR*LLKNQ*KCF CAGCGCAGCTTCTTCAAAAGTACTAAAGGAAACACCATTAAAGACTACACGTCTCTTTAGGTAATTACTAAAGAACCAATAGAAATGCTT	20188
Ic Query_1_reversed	20189	VEIQMI*AIVIQKGTKGVNITH*ICAKKE TGTGGAATACAAATGATATAAGCAATTGTTATCCAGAAAGGTACTAAAGGTGTGAACATAACCATCCACTGAATATGTGCTAAAAAGA	20278
Ic Query_1_reversed	20279	TSLVR*NVKYK*ITE*TPGKNE*TG VKQST AACATCATTAGTAAGATAAAATGTCAAGTACAAGTAAATAACAGAATAAACACCAGGTAAGAATGAGTAAACTGGTGTTAAACAGAGTAC	20368
Ic Query_1_reversed	20369	VNDIRNSKVLKATT*LYSPKALLNLIK**A AGTGAATGACATAAGGAATAGTAAAGTATTAAGGCAACTACATGACTGTATTCACCAAAAGCTCTCTAAACCTCATAAAATAGTAGGC	20458
Ic Query_1_reversed	20459	RHVTTIATIPPATIDADMSKAPIG*ISGVN AAGGCATGTTACTACGATAGCTACAATACCACCAGCTACTATAGATGCTGATATGTCCAAAGCACCAATAGGTTGAATTAGTGGTGTA	20548
Ic Query_1_reversed	20549	ILVSKFTASTPQKTPGKDL**SLLSTHLPL CATATTAGTAAGTAAATTTACAGCATCTACACCACAGAAAACCTCTGGTAAAGATCTGTAATAATCATTGTTAAGTACCCATCTACCACT	20638
Ic Query_1_reversed	20639	VDTQTPASDLSQVPC LQYSESKVVTTLTEP AGTAGATACAAAAACCAAGCTTCTGATCTTTACAAAGTGCGTGCTACAGTACTCAGAATCAAAAGTTGTTACCACTCTAACAGAACC	20728
Ic Query_1_reversed	20729	SR*VLGN*IIEPSMST*RVSGRKLS*ATEP TTCAAGGTAGGTGTTAGGAAATTGAATAATAGAGCCATCCATGAGCACATAACGTGTGTGAGGGCGTAACTTTCATAAGCAACAGAACC	20818
Ic Query_1_reversed	20819	SSTLVS*QYGTGLPEASLKIVHSAAKTQAD TTCTAGTACATTGGTATCATAACAATATGGTACTGGCTTACCAGAAGCATCTTTAAAAATTGTACATTCAGCAGCCAAAACACAAGCTGA	20908
Ic Query_1_reversed	20909	VAKSVYSISFDGV*QMLPTALKTLGKKCKK TGTTGCAAAGTCAGTGTACTCTATAAGTTTGTGGTGTGTAACAGATGTTACCAACTGCACTAAAACTCTAGGTAAGAAATGCAAAAA	20998
Ic Query_1_reversed	20999	SPLVVRNIVPGKPGTTKPTSLVMTAAINGQ GTCACCATTAGTTGTGCGTAATATCGTGCCAGGCAAACAGGCACGACAAAACCACTTCTCTTGTATGACTGCAGCAATCAATGGGCA	21088
Ic Query_1_reversed	21089	ALSLV*LPPRWLNHVSKSACL LAKQVS VDA AGCTTTGTCAATTAGTATAACTACCACCAGCTGGCTAAACCATGTGTCAAATCAGCATGTTTGTAGCAAAACAAGTATCTGTAGATGC	21178

MSRVTPPSIALYPMISLEKSVCLDMT* TGV

Ic Query_1_reversed	21179	TATGTCACGAGTGACACCACCATCAATAGCCTTGTATCCTATGATTTCACTTGAAAAGTCAGTATGTTTAGACATGACATGAACAGGTGT	21268
Ic Query_1_reversed	21269	I K * K I A A T K R N T S V T L I N C F N Q L L T I L P P L TATTAAATAGAAAATAGCAGCAACAAAAAGGAACACAAGTGTAACCTTAATTAAGTCTCAACCAATTATTAACAATTTTACCACCCTT	21358
Ic Query_1_reversed	21359	S A I F V V T T L T T C L V V A H V N L K G K L F F L A A L AAGTGCTATCTTTGTTGTACAACTTAACAACCTTGCTAGTAGTTGCACATGTCAACTTAAAGGTAAGTTATTCTTTTAGCAGCACT	21448
Ic Query_1_reversed	21449	R I C F R S C S D N D M K S L T F H I K A M L * L F A T C A ACGTATTTGTTTTCGTAGTTGTTTCAGACAATGACATGAAATCTTTAACGTTCCATATCAAAGCAATGTTGTGACTTTTTGCTACCTGCGC	21538
Ic Query_1_reversed	21539	L I * R A L Q S I Q A P R S R G V M F S T L L * V S I * L L ATTAATATGACGCGCACTACAGTCAATACAAGCACCAAGGTCACGGGGTGTCATGTTTTCAACTTTGTTATAGGTGAGCATATAGTTATT	21628
Ic Query_1_reversed	21629	Q L S P V T S M S D * C D N L R H S T T S L V S T S E S T N ACAACATCGCCAGTAACCTTCTATGTCAGATTGATGTGACAATTAAGACATTCAACAACATCTTTAGTTTCTACATCTGAATCAACAAA	21718
Ic Query_1_reversed	21719	P C R A A E I K V D K T L S K D T F F A S S A S A V A T S V CCCTTGCCGAGCTGCTGAAATAAAGTAGATAAGACATTGTCTAAGGACACATTCTTTGCAAGTTCAGCTTCTGCAGTTGCAACTAGTGT	21808
Ic Query_1_reversed	21809	L S F S I G T L K V D E N V L T * A S N I L T A T S A L S P TTTGAGTTTTTCCATTGGTACGTTAAAAGTTGATGAAAACGTATTAACGTAAGCATCAAACATTTAACTGCAACTCCGCACTATCACC	21898
Ic Query_1_reversed	21899	T S D T N A * S S N S I G * H I S * L * * T D A D F A D D S AACATCAGACACTAATGCCTGATCTAGTAACAGTATAGGTTGACACATAAGCTGACTGTAGTAAACAGACGCTGATTTTGAGATGATTC	21988
Ic Query_1_reversed	21989	S H F D L P S K T I T L I G N E P L V L L A L R L S K L T K TTCACATTTTGATTACCATCAAAAACATAACATTAATAGGCAATGAACCTTTAGTGTTATTAGCTCTCAGGTTGTCTAAGTTAACAAA	22078
Ic Query_1_reversed	22079	* E R E C L S * V F * P A L S K * R W M E P F F T V T L S T ATGAGAGAGAGAATGTCTTTCATAAGTCTTTTGACCAGCTTTATCAAAGTAAAGATGGATGGAACCATTTCTACTGTAACACTATCAAC	22168
Ic Query_1_reversed	22169	M * E D W S V G F I G L L N C S D K S L A T S S L I N V L P GATGTAAGAAGACTGGTCAGTAGGATTTATTGGTCTTTTAACTGTAGTGACAAGTCTCTCGCAACTTCATCACTAATAAATGTACTACC	22258
Ic Query_1_reversed	22259	A Q N V S Q L T Q F Q L C S L Q K P L P P L A * T * K D L L AGCACAGAATGTATCAATTAACACAATTCCAATTGTGTAGTTTGCAAAAGCCTTTACCTCCATTAGCATAGACATAAAAGGACCTTCT	22348

T P L T I V V H S T L V A L L R L * H I I Q V D E L Q P S T

Icd Query_1_reversed	22349	AACACCATTAACAATAGTTGTACATTCGACTCTTGTGCTCTATTACGTTTGTAAACACATCATACAAGTTGATGAATTACAACCGTCTAC	22438
Icd Query_1_reversed	22439	TCT*LFHT**NDAKKMYILTI AEI GAI CTR AACATGCACATAACTTTCCATACATAATAAAATGATGCAAAGAAGATGTACATTCTAACCATAGCTGAAATCGGGGCCATTTGTACAAG	22528
Icd Query_1_reversed	22529	L I I N H I S Q E L L I K C T A K * L K N N C M I A A N P S ATTAATTATTAACACATAAGCCAAGAATTACTAATAAAATGTACTGCAAATAGCTGAAAAACAATTGCATGATTGCAGCCAATCCAAG	22618
Icd Query_1_reversed	22619	T * K N L V K R I Y A K N H S A T K P K A V K S H L K D E M TACATAGAAAAACCTAGTGAAAAGAATATATGCCAAAAACCACTCTGCAACTAAGCCAAAAGCAGTTAAATCCCATTTAAAAGATGAAAT	22708
Icd Query_1_reversed	22709	V I C I V S K E G * V S K E S K P L R Q T L Q G I E P V Q * GGTAATTTGTATAGTTTCTAAAGAAGGATAGGTGTCTAAAGAATCTAAACCACTAAGACAAACACTACAAGGTATAGAACCAGTACAGTA	22798
Icd Query_1_reversed	22799	V A I V T L V E F K * P S L * P V Q * E G M P K L D I K T P GGTTGCAATAGTGACATTAGTAGAGTTCAAATAGCCTTCTCTGTAAACCACTACAGTAAGAAGGCATGCCTAAATTAGACATTAACACACC	22888
Icd Query_1_reversed	22889	K A A V E * I K E P R Q T L N S K N Q I I I F I S L E K L G TAAAGCAGCGTTGAGTAGATTAAGAACCTAGGCCAAACACTTAATAGTAAAAACCAAAATTATAATATTATCAGTTTAGAAAAATTAGG	22978
Icd Query_1_reversed	22979	D F K * L N E A S R Q N L P T L L T V F F A I V V G I D A L TGACTTCAAATAATTAATGAAGCCTCTAGACAAAATTTACCGACACTCTTAACAGTATTCTTTGCTATAGTAGTCGGCATAGATGCTTT	23068
Icd Query_1_reversed	23069	I L E F V L L V K V H N C S N K V K K * G I * L V Q T R F K AATTCTAGAATTTGACTTCTAGTAAAGTACACAATTGTAGCAATAAAGTAAAGAAATAAGGCATATAATTAGTACAAACACGGTTTAA	23158
Icd Query_1_reversed	23159	H R V T M L V V V L T T L L R K G L A * L A I V S Q G T L L ACACCGTGTAAGTATGTTAGTAGTTGTACTAACAACCTTTGTTAAGAAAAGGCTTAGCATAATTAGCTATAGTATCCCAAGGGACACTATT	23248
Icd Query_1_reversed	23249	T A A K P * V A R V F K P N T L D N S L G F L I V R L E L S AACAGCAGCTAAACCATGAGTAGCAAGGGTTTCAAACCTAATACTCTAGATAATTCATTAGGTTTCTTAATAGTAAGACTAGAATTGTC	23338
Icd Query_1_reversed	23339	T * A A I R S V W P T S S V I F K L L F A G L S I M S P T T TACATAAGCAGCCATTAGATCTGTGTGGCCAACCTCTTCTGTAAATTTTAAACTATTATTTGCTGGTTTAAGTATAATGTCTCCTACAAC	23428
Icd Query_1_reversed	23429	S V V F T L H S R T S F C M V G F S T T S S E T G F R S S Q TTCGGTAGTTTTACATTACACTCAAGAACGTCTTTCTGTATGGTAGGATTTCCACTACTTCTCAGAGACTGGTTTTAGATCTTCGCA	23518

A R L S I P C A S S D F S T S N E F D V S T G F V L Q R Q R

Ic|Query_1_reversed 23519 GGCAAGATTATCCATTCCCTGCGCGTCTCTGACTTCAGTACATCAAACGAATTTGATGTTTCAACTGGTTTTGTGCTCCAAAGACAACG 23608

Ic|Query_1_reversed 23609 TATAACACCAGGTATTTGGTTTATACGTGGCTTTATTAGTTGCATTGTTAACATGCCAAACAATAGGTTTATGTAACAATTTAGCTCCTTT 23698
 V 78.1% (25/32)[IGHV2-70*19](#) 199 -----. 199

Ic|Query_1_reversed 23699 CTAAAAAGAGGGTGTGTAGTGTGTTTATAATCA 23729
 V 78.1% (25/32)[IGHV2-70*19](#) 198 ...C.G...T.TGC.....C..... 168

Lambda K H
 1.10 0.333 0.549

Gapped
 Lambda K H
 1.08 0.280 0.540

Effective search space used: 4073052391

Total queries = 1
 Total identifiable CDR3 = 0
 Total unique clonotypes = 0

Database: imgt.Homo_sapiens.V.f.orf.p
 Posted date: Feb 10, 2020 12:47 PM
 Number of letters in database: 155,489
 Number of sequences in database: 531

Database: imgt.Homo_sapiens.D.f.orf
 Posted date: May 17, 2012 12:56 PM
 Number of letters in database: 828
 Number of sequences in database: 34

Database: imgt.Homo_sapiens.J.f.orf
 Posted date: May 17, 2012 12:56 PM
 Number of letters in database: 1,426
 Number of sequences in database: 32

Matrix: blastn matrix 1 -1
 Gap Penalties: Existence: 4, Extension: 1

BLAST is a registered trademark of the National Library of Medicine

[Support center](#) [Mailing list](#) [YouTube](#)



- [National Library Of Medicine](#)



- [National Institutes Of Health](#)



- [U.S. Department of Health & Human Services](#)



- [USA.gov](#)

[NCBI](#)

[National Center for Biotechnology Information, U.S. National Library of Medicine](#) 8600 Rockville Pike, Bethesda MD, 20894 USA
[Policies and Guidelines](#) | [Contact](#)

- [NBI Home](#)
- [Sign in to NCBI](#)
- [Skip to Main Content](#)
- [Skip to Navigation](#)
- [About NCBI Accesskeys](#)



[U.S. National Library of Medicine](#)



[NBI National Center for Biotechnology Information](#)

-
- [MNCBI](#)
- [Sign in to](#)
- [NCBI Register](#)
- [Sign Out](#)

[GBLAST](#)» JOB ID: 1Mo-yQ3LtOeHR4Nx2mqgF9tjpQbE

Formatting Results

Database: imgt.Homo_sapiens.V.f.orf.p; imgt.Homo_sapiens.D.f.orf;
imgt.Homo_sapiens.J.f.orf
597 sequences; 157,743 total letters

Query= reverse translation of 6VSB:A|PDBID|CHAIN|SEQUENCE

Length=3864

Sequences producing significant alignments:

[GLV1-44*01](#)germline gene
[GLV1-47*02](#)germline gene
[GLV3-22*01](#)germlinegene

Score (Bits)	E Value
8.8	4.6
8.8	4.6
8.8	4.6

Domain classification requested: imgt

Note that your query represents the minus strand of a V gene and has been converted to the plus strand. The sequence positions refer to the converted sequence. V-

(D)-J rearrangement summary for query sequence (multiple equivalent top matches, if present, are separated by a comma):

Top V gene match	Top J gene match	Chain type	stop codon	V-J frame	Productive	Strand
------------------	------------------	------------	------------	-----------	------------	--------

IGLV1-44*01,IGLV1-47*02	N/A	VL	No			-
-------------------------	-----	----	----	--	--	---

V-(D)-J junction details based on top germline gene matches:

V region end	V-Jjunction*	J regionstart
TGCGG		

*: Overlapping nucleotides may exist at V-D-J junction (i.e, nucleotides that could be assigned to either rearranging gene). Suchnucleotides are indicated inside a parenthesis (i.e., (TACAT)) butarenotincludedundertheV,DorJgeneitself.

Alignment summary between query and top germline V gene hit:

	from	to	length	matches	mismatches	gaps	identity(%)
FR2-IMGT	478	516	39	22	17	0	56.4
CDR2-IMGT	517	525	9	8	1	0	88.9
FR3-IMGT	526	531	6	5	1	0	83.3
Total			54	35	19	0	64.8

Alignments

			<-----FR2-IMGT-----><CDR2-IM><-----			
			H H I A V A A H K G V I G G N N L R F I K V A L G H K P M G			
V 64.8% (35/54)	c Query_1_reversed	478	CACCACATCGCAGTTGCCGCTCACAAAGGTGTTATCGGTGGTAATAATCTGCGGTTTCATAAAAGTTGCGCTGGGTCACAAACCAATGGGT		567	
		112	..G..GC..C...GAA.G..C.C....CTCC.C...TA.A.....A.....		165	
			Q Q L P G T A P K L L I Y S N N Q R			
V 64.8% (35/54)	c Query_1_reversed	112	..G..GC..C...GAA.G..C.C....CTCC.C...TA.A.....A.....		165	
			-----FR3-IMGT----->			
			A V A H K H A F A R K M R F A I M A N R R R G G K V F F L R			
	c Query_1_reversed	568	GCCGTTGCTCACAAACACGCCTTCGCGCGGAAATGCGCTTGCCATCATGGCAAATCGCCGGCGCGGTGGTAAAGTTTTTCTCGCGC		657	
			R H I G H M Q K H H A M R R A L R K A H Q M I A F A A K I H			
	c Query_1_reversed	658	CGGCACATAGGTACATGCAGAAACACCACGCCATGCGGCGCGCTCTGCGGAAAGCTCATCAGATGATAGCCTTTGCCGCAAAATCCAC		747	
			A F A L A Q H A F A H F G R R Q V R A R A N F R R A N Q L L			
	c Query_1_reversed	748	GCGTTTGCTCTGGCCCAGCACGCATTGCTCATTTTGGTCGCCGCCAGGTTGCGCTCGCGCAATTTCCGCCGCGCAATCAGCTGCTG		837	
			G H I G L Q A L Q A A G N Q A I N L H F R F R R I Q A A Q N			
	c Query_1_reversed	838	GGTCACATAGGTCTGCAGGCTCTGCAGGCGGCCGGAATCAGGCGATCAATCTGCACTTCCGCTTCCGGCGGATCCAGGCGGCTCAGAAT		927	
			I V Q H A A N R A K V A A Q L F H Q G V Q R L R V L V H H I			

lcl Query_1_reversed	928	ATCGTTCAGCACGCTGCTAATCGCGCCAAAGTTGCTGCTCAGCTGTTTCACCAGGGTGTTACAGCGCCTGCGCGTTCTGGTTCACCACATC	1017
lcl Query_1_reversed	1018	LQFAQRARGAAQAI LNFANRAVKLVNQL CTGCAGTTTGCCAGCGCGCTCGCGGTGCTCAGGCTATCTGAATTTTGCCAATCGCGCTGTAACTGGTTCGCAATCAGTTTCTG	1107
lcl Query_1_reversed	1108	VFI QHVLGHANAVKAI RHLHRKRNLRAR GTTTTCATACAGCACGTTCTGGGTACGCCAATGCCGTTAAAGCGATACGCCATCTGCATCGCAAACGGAATCTGCAGCGCCGCGCCCGC	1197
lcl Query_1_reversed	1198	AKGPAAGNGARQQRAGI LRNHF I GQRRQH GCCAAAGGTCCAGCCGCTGGTAATGGTGCCCGCCAGCAGCGCGTGGTATACTGCGCAATCATTTTCATCGGTGAGCAGCGCGGCGAGCAC	1287
lcl Query_1_reversed	1288	GQAVKFLRANQI ARRN I AQAI A I LFNKARI GGTCAGGCCGTTAAATTTCTGCGCGCAAATCAGATCGCGCGCGCAATATCGCCAGGCAATCGCCATACTGTTTAATAAGCCCGCATC	1377
lcl Query_1_reversed	1378	RQGHFVKQQI FNKA AFARFARI RQNLAKVK CGCCAGGGTCACTTTGTTAAACAGCAGATCTTCAATAAGCTGCGTTTGCTCGGTTGCTCGGATCCGGCAGAATCTGGCTAAAGTTAAA	1467
lcl Query_1_reversed	1468	AAK I FNRRGFI NLFHLRKHF LGVFI LFHRN GCCGCCAAATCTTTAATCGCGGGGTTTTATAAATCTGTTTACCTGCGCAAACACTTCTGGGTGTTTTATCCTGTTCCACCGCAAT	1557
lcl Query_1_reversed	1558	AGQRAVQLGAKAAI LQQQVAAF GA I AANI H GCCGGTCAGCGCGCGGTTTCAGCTGGGTGCAAAGCTGCCATACTGCAGCAGCAGGTTGCTGCATTGGTGCTATCGCCGCAAATATACAT	1647
lcl Query_1_reversed	1648	GAI HAGFGHAHRQNFGGHANGKVGRNRNAV GGTGCAATCCACGCTGGTTTTGGTCATGCTACCGGCAGAAATTCGGTGGTCACGCTAATGGTAAAGTTGGTCGGAATCGCAATGCTGTT	1737
lcl Query_1_reversed	1738	VAI RHAVFRAQAHGI RNNALARHAARAARA GTTGCTATACGCCACGCTGTTTCCGCGCCAGGCTCATGGTATACGCAATAATGCTCTGGCTCGCCACGCTGCTCGCGCTGCCCGGGCT	1827
lcl Query_1_reversed	1828	VGLGLI ARANARANRN I AFI AVVHMFRAHQ GTTGGTCTGGGTCTGATAGCTCGCGCAAATGCCGCGCCAATCGGAATATCGCATTATAGCTGTTGTTACATGTTCCGCGCCAATCAG	1917
lcl Query_1_reversed	1918	AARAGLKHVAAGAI HAPGRGQLI RMNRHRH GCAGCCCGCGGGGTCTGAAACACGTTGCTGCCGGTGTATACAGCGCCAGGTCGGGGTCAGCTGATCCGCATGAATCGCCACCGGCAC	2007
lcl Query_1_reversed	2008	FGAVH I LI QHRHLVAGVGARGNHAHAAKAA TTCGGTGCAGTTCACATCTGTATACAGACCGCCACCTGGTTGCTGGTGTGGTGCCCGGGGTAATCAGCTCACGCCGCCAAAGCTGCA	2097

RGN I QNFQGLRI AHR I GG I RN I AAKLLKRQ

lc|Query_1_reversed 2098 CGGGGTAATATCCAGAATTTCCAGGGTCTGCGGATCGCGCACCGCATCGGTGGTATCCGCAATATCGCGGCCAACTGCTGAAACGGCAG 2187

			K F F V A F G Q H A G A G Q	
	lc Query_1_reversed	2188	AAATTTTTTGTGCTTTTCGGTCAGCACGCCGGTGCCGGTCAGG	2230
V 78.6% (22/28)	GLV3-22*01	240	-----G..C.T.C..AT.....	213

Lambda	K	H
1.10	0.333	0.549

Gapped Lambda	K	H
1.08	0.280	0.540

Effective search space used: 530720504

Total queries = 1
 Total identifiable CDR3 = 0
 Total unique clonotypes = 0

Database: imgt.Homo_sapiens.V.f.orf.p
 Posted date: Feb 10, 2020 12:47PM
 Number of letters in database: 155,489
 Number of sequences in database: 531

Database: imgt.Homo_sapiens.D.f.orf
 Posted date: May 17, 2012 12:56 PM
 Number of letters in database: 828
 Number of sequences in database: 34

Database: imgt.Homo_sapiens.J.f.orf
 Posted date: May 17, 2012 12:56 PM
 Number of letters in database: 1,426
 Number of sequences in database: 32

Matrix: blastn matrix 1 -1
 Gap Penalties: Existence: 4, Extension: 1

BLAST is a registered trademark of the National Library of Medicine
[Support center](#) [Mailing list](#) [YouTube](#)



- ♦ [National Library Of Medicine](#)



- ♦ [National Institutes Of Health](#)



- ♦ [U.S. Department of Health & Human Services](#)



- ♦ [USA.gov](#)

NBI

[National Center for Biotechnology Information, U.S. National Library of Medicine](#) 8600 Rockville Pike, Bethesda MD, 20894 USA
[Policies and Guidelines](#) | [Contact](#)



NetMHC 4.0 Server - prediction results

Technical University of Denmark - DTU

NetMHC version 4.0

Input is in FSA format

Peptide length 14

Rank Threshold for Strong binding peptides 0.500

Rank Threshold for Weakbindingpeptides 2.000

pos	HLA	peptide	CoreOffset	I_pos	I_len	D_pos	D_len	iCore	Identity	1-log50k(aff)	Affinity(nM)	%Rank	BindLevel
0	HLA-A0101	HCRCQKPCIRHQYC	CRCCIRHQY	1	0	0	3	3	CRCQKPCIRHQY	vdchain	0.047	30056.59	36.00
1	HLA-A0101	CRCQKPCIRHQYCL	PCIRHQYCL	5	0	0	0	0	PCIRHQYCL	vdchain	0.051	28883.99	32.00
2	HLA-A0101	RCQKPCIRHQYCLC	PCIRHQYCL	4	0	0	0	0	PCIRHQYCL	vdchain	0.042	31706.83	44.00
3	HLA-A0101	CQKPCIRHQYCLCR	PCIRHQYCL	3	0	0	0	0	PCIRHQYCL	vdchain	0.039	32883.22	50.00
4	HLA-A0101	QKPCIRHQYCLCRT	PCIRHQYCL	2	0	0	0	0	PCIRHQYCL	vdchain	0.035	34195.24	60.00
5	HLA-A0101	KPCIRHQYCLCRTV	CIYCLCRTV	2	0	0	2	3	CIRHQYCLCRTV	vdchain	0.054	27776.04	27.00
6	HLA-A0101	PCIRHQYCLCRTVD	PCIRHQYCL	0	0	0	0	0	PCIRHQYCL	vdchain	0.048	29669.84	35.00
7	HLA-A0101	CIRHQYCLCRTVDG	YCLCRTVDG	5	0	0	0	0	YCLCRTVDG	vdchain	0.048	29842.72	36.00
8	HLA-A0101	IRHQYCLCRTVDGL	YLCRTVDGL	4	0	0	1	1	YLCRTVDGL	vdchain	0.076	21886.22	13.00
9	HLA-A0101	RHQYCLCRTVDGLH	HQYCLVDGL	1	0	0	5	3	HQYCLCRTVDGL	vdchain	0.063	25204.78	20.00
10	HLA-A0101	HQYCLCRTVDGLHL	LCRTVDGLL	4	0	0	8	1	LCRTVDGLHL	vdchain	0.068	24008.29	17.00
11	HLA-A0101	QYCLCRTVDGLHLH	YLCRTVDGL	1	0	0	1	1	YLCRTVDGL	vdchain	0.058	26822.96	24.00
12	HLA-A0101	YCLCRTVDGLHLHR	YLCRTVDGL	0	0	0	1	1	YCLCRTVDGL	vdchain	0.062	25460.78	20.00
13	HLA-A0101	CLCRTVDGLHLHRK	TVDGLHLRK	4	0	0	7	1	TVDGLHLHRK	vdchain	0.049	29497.63	34.00
14	HLA-A0101	LCRTVDGLHLHRKP	LCRTVDGLL	0	0	0	8	1	LCRTVDGLHL	vdchain	0.040	32512.45	48.00
15	HLA-A0101	CRTVDGLHLHRKPV	TVDGLHLPV	2	0	0	7	3	TVDGLHLHRKPV	vdchain	0.101	16721.65	6.50
16	HLA-A0101	RTVDGLHLHRKPVK	TVDGLHLPV	1	0	0	7	3	TVDGLHLHRKPV	vdchain	0.060	26227.75	22.00
17	HLA-A0101	TVDGLHLHRKPVKR	TVDGLHLPV	0	0	0	7	3	TVDGLHLHRKPV	vdchain	0.048	29724.46	35.00
18	HLA-A0101	VDGLHLHRKPVKRL	LHRKPVKRL	5	0	0	0	0	LHRKPVKRL	vdchain	0.029	36598.32	75.00
19	HLA-A0101	DGLHLHRKPVKRLC	LHRKPVKRL	4	0	0	0	0	LHRKPVKRL	vdchain	0.025	38185.22	80.00

Protein vdchain. Allele HLA-A0101. Number of high binders 0. Number of weak binders 0. Number of peptides 20

Link to Allele Frequencies in Worldwide Populations [HLA-A0101](#)

Rank Threshold for Strong binding peptides 0.500

Rank Threshold for Weakbindingpeptides 2.000

pos	HLA	peptide	CoreOffset	I_pos	I_len	D_pos	D_len	iCore	Identity	1-log50k(aff)	Affinity(nM)	%Rank	BindLevel
0	HLA-A0201	HCRCQKPCIRHQYC	CQKPCIQYC	3	0	0	6	2 CQKPCIRHQYC	vdchain	0.041	32017.43	55.00	
1	HLA-A0201	CRCQKPCIRHQYCL	CQKPCYIYCL	2	0	0	6	3 CQKPCIRHQYCL	vdchain	0.063	25274.15	36.00	
2	HLA-A0201	RCQKPCIRHQYCLC	CQKPCYIYCL	1	0	0	6	3 CQKPCIRHQYCL	vdchain	0.055	27680.03	41.00	
3	HLA-A0201	CQKPCIRHQYCLCR	CQKPCYIYCL	0	0	0	6	3 CQKPCIRHQYCL	vdchain	0.052	28640.00	43.00	
4	HLA-A0201	QKPCIRHQYCLCRT	CIQYCLCRT	3	0	0	2	2 CIRHQYCLCRT	vdchain	0.050	29042.87	44.00	
5	HLA-A0201	KPCIRHQYCLCRTV	KQYCLCRTV	0	0	0	1	5 KPCIRHQYCLCRTV	vdchain	0.196	6003.43	11.00	
6	HLA-A0201	PCIRHQYCLCRTVD	CIYCLCRTV	1	0	0	2	3 CIRHQYCLCRTV	vdchain	0.120	13596.01	19.00	
7	HLA-A0201	CIRHQYCLCRTVDG	CIYCLCRTV	0	0	0	2	3 CIRHQYCLCRTV	vdchain	0.121	13515.05	19.00	
8	HLA-A0201	IRHQYCLCRTVDGL	YLCRTVDGL	4	0	0	1	1 YLCRTVDGL	vdchain	0.200	5768.28	11.00	
9	HLA-A0201	RHQYCLCRTVDGLH	YLCRTVDGL	3	0	0	1	1 YLCRTVDGL	vdchain	0.130	12267.80	18.00	
10	HLA-A0201	HQYCLCRTVDGLHL	HQYCLCLHL	0	0	0	6	5 HQYCLCRTVDGLHL	vdchain	0.201	5701.33	11.00	
11	HLA-A0201	QYCLCRTVDGLHLH	CLCRTVLHL	2	0	0	6	2 CLCRTVDGLHL	vdchain	0.132	11989.47	17.00	
12	HLA-A0201	YLCRTVDGLHLHR	CLCRTVLHL	1	0	0	6	2 CLCRTVDGLHL	vdchain	0.124	13092.32	19.00	
13	HLA-A0201	CLCRTVDGLHLHRK	CLCRTVLHL	0	0	0	6	2 CLCRTVDGLHL	vdchain	0.121	13545.79	19.00	
14	HLA-A0201	LCRTVDGLHLHRKP	LCRTVDGLHL	0	0	0	5	1 LCRTVDGLHL	vdchain	0.040	32439.02	55.00	
15	HLA-A0201	CRTVDGLHLHRKPV	RTVDGLHPV	1	0	0	7	4 RTVDGLHLHRKPV	vdchain	0.185	6747.66	12.00	
16	HLA-A0201	RTVDGLHLHRKPVK	RTVDGLHPV	0	0	0	7	4 RTVDGLHLHRKPV	vdchain	0.113	14712.15	21.00	
17	HLA-A0201	TVDGLHLHRKPVKR	TVDGLHLPV	0	0	0	7	3 TVDGLHLHRKPV	vdchain	0.070	23462.32	33.00	
18	HLA-A0201	VDGLHLHRKPVKRL	GLHLHVKRL	2	0	0	5	3 GLHLHRKPVKRL	vdchain	0.070	23522.06	33.00	
19	HLA-A0201	DGLHLHRKPVKRLC	GLHLVKRLC	1	0	0	4	4 GLHLHRKPVKRLC	vdchain	0.061	25950.27	38.00	

Protein vdchain. Allele HLA-A0201. Number of high binders 0. Number of weak binders 0. Number of peptides 20

Link to Allele Frequencies in Worldwide Populations [HLA-A0201](#)

Rank Threshold for Strong binding peptides 0.500

Rank Threshold for Weakbindingpeptides 2.000

pos	HLA	peptide	CoreOffset	I_pos	I_len	D_pos	D_len	iCore	Identity	1-log50k(aff)	Affinity(nM)	%Rank	BindLevel
0	HLA-A0202	HCRCQKPCIRHQYC	CQKPCIQYC	3	0	0	6	2 CQKPCIRHQYC	vdchain	0.024	38690.93	70.00	
1	HLA-A0202	CRCQKPCIRHQYCL	CQKPCYIYCL	2	0	0	6	3 CQKPCIRHQYCL	vdchain	0.098	17343.94	30.00	
2	HLA-A0202	RCQKPCIRHQYCLC	CQKPCYIYCL	1	0	0	6	3 CQKPCIRHQYCL	vdchain	0.069	23691.92	39.00	
3	HLA-A0202	CQKPCIRHQYCLCR	CQKPCYIYCL	0	0	0	6	3 CQKPCIRHQYCL	vdchain	0.068	24082.16	40.00	
4	HLA-A0202	QKPCIRHQYCLCRT	CIQYCLCRT	3	0	0	2	2 CIRHQYCLCRT	vdchain	0.036	33974.70	60.00	
5	HLA-A0202	KPCIRHQYCLCRTV	CIYCLCRTV	2	0	0	2	3 CIRHQYCLCRTV	vdchain	0.167	8243.69	19.00	
6	HLA-A0202	PCIRHQYCLCRTVD	CIYCLCRTV	1	0	0	2	3 CIRHQYCLCRTV	vdchain	0.128	12549.05	24.00	
7	HLA-A0202	CIRHQYCLCRTVDG	CIYCLCRTV	0	0	0	2	3 CIRHQYCLCRTV	vdchain	0.125	12896.05	24.00	
8	HLA-A0202	IRHQYCLCRTVDGL	YLCRTVDGL	4	0	0	1	1 YLCRTVDGL	vdchain	0.222	4515.09	13.00	
9	HLA-A0202	RHQYCLCRTVDGLH	YLCRTVDGL	3	0	0	1	1 YLCRTVDGL	vdchain	0.142	10779.18	22.00	
10	HLA-A0202	HQYCLCRTVDGLHL	HQYCLCLHL	0	0	0	6	5 HQYCLCRTVDGLHL	vdchain	0.207	5348.34	15.00	
11	HLA-A0202	QYCLCRTVDGLHLH	YLCRTVDGL	1	0	0	1	1 YLCRTVDGL	vdchain	0.120	13692.70	25.00	
12	HLA-A0202	YLCRTVDGLHLHR	YLCRTVDGL	0	0	0	1	1 YLCRTVDGL	vdchain	0.125	12885.17	24.00	
13	HLA-A0202	CLCRTVDGLHLHRK	CLCRTVLHL	0	0	0	6	2 CLCRTVDGLHL	vdchain	0.096	17613.61	30.00	
14	HLA-A0202	LCRTVDGLHLHRKP	LCRTVDLHL	0	0	0	6	1 LCRTVDGLHL	vdchain	0.032	35548.86	65.00	
15	HLA-A0202	CRTVDGLHLHRKPV	RTVDGLHLV	1	0	0	8	4 RTVDGLHLHRKPV	vdchain	0.135	11601.78	23.00	
16	HLA-A0202	RTVDGLHLHRKPVK	RTVDGLHLV	0	0	0	8	4 RTVDGLHLHRKPV	vdchain	0.089	19042.47	32.00	
17	HLA-A0202	TVDGLHLHRKPVKR	TVDGLHLPV	0	0	0	7	3 TVDGLHLHRKPV	vdchain	0.037	33461.05	60.00	

18	HLA-A0202VDGLHLHRKPVKRL	GLHLHVKRL	2	0	0	5	3 GLHLHRKPVKRL	vdchain	0.095	17971.87	31.00
19	HLA-A0202DGLHLHRKPVKRLC	GLHLHVKRL	1	0	0	5	3 GLHLHRKPVKRL	vdchain	0.067	24267.09	40.00

Protein vdchain. Allele HLA-A0202. Number of high binders 0. Number of weak binders 0. Number of peptides 20

Link to Allele Frequencies in Worldwide Populations[HLA-A0202](#)

Rank Threshold for Strong binding peptides 0.500

Rank Threshold for Weakbindingpeptides 2.000

pos	HLA	peptide	CoreOffset	I_pos	I_len	D_pos	D_len	iCore	Identity	1-log50k(aff)	Affinity(nM)	%Rank	BindLevel
0	HLA-A0203	HCRCQKPCIRHQYC	CQKPCIRYC	3	0	0	7	2 CQKPCIRHQYC	vdchain	0.028	36905.30	75.00	
1	HLA-A0203	CRCQKPCIRHQYCL	CQKPCIRHL	2	0	0	8	3 CQKPCIRHQYCL	vdchain	0.071	23223.14	41.00	
2	HLA-A0203	RCQKPCIRHQYCLC	CQKPCICYCL	1	0	0	6	3 CQKPCIRHQYCL	vdchain	0.048	29758.58	55.00	
3	HLA-A0203	CQKPCIRHQYCLCR	CQKPCICYCL	0	0	0	6	3 CQKPCIRHQYCL	vdchain	0.046	30320.85	60.00	
4	HLA-A0203	QKPCIRHQYCLCRT	CIRYCLCRT	3	0	0	3	2 CIRHQYCLCRT	vdchain	0.068	23939.02	43.00	
5	HLA-A0203	KPCIRHQYCLCRTV	KQYCLCRTV	0	0	0	1	5 KPCIRHQYCLCRTV	vdchain	0.281	2392.70	9.00	
6	HLA-A0203	PCIRHQYCLCRTVD	CIYCLCRTV	1	0	0	2	3 CIRHQYCLCRTV	vdchain	0.168	8090.73	18.00	
7	HLA-A0203	CIRHQYCLCRTVDG	CIYCLCRTV	0	0	0	2	3 CIRHQYCLCRTV	vdchain	0.162	8633.12	19.00	
8	HLA-A0203	IRHQYCLCRTVDGL	YLCRTVDGL	4	0	0	1	1 YLCRTVDGL	vdchain	0.140	11037.29	22.00	
9	HLA-A0203	RHQYCLCRTVDGLH	RQYCLCRTV	0	0	0	1	1 RHQYCLCRTV	vdchain	0.110	15247.28	28.00	
10	HLA-A0203	HQYCLCRTVDGLHL	HQYCLCLHL	0	0	0	6	5 HQYCLCRTVDGLHL	vdchain	0.170	7945.59	18.00	
11	HLA-A0203	QYCLCRTVDGLHLH	CLCRTVLHL	2	0	0	6	2 CLCRTVDGLHL	vdchain	0.096	17684.84	32.00	
12	HLA-A0203	YLCRTVDGLHLHR	CLCRTVLHL	1	0	0	6	2 CLCRTVDGLHL	vdchain	0.088	19290.07	34.00	
13	HLA-A0203	CLCRTVDGLHLHRK	CLCRTVLHL	0	0	0	6	2 CLCRTVDGLHL	vdchain	0.092	18453.24	33.00	
14	HLA-A0203	LCRTVDGLHLHRKP	LCRTVGLHL	0	0	0	5	1 LCRTVDGLHL	vdchain	0.037	33681.91	65.00	
15	HLA-A0203	CRTVDGLHLHRKPV	RTVDGLHPV	1	0	0	7	4 RTVDGLHLHRKPV	vdchain	0.290	2161.81	8.50	
16	HLA-A0203	RTVDGLHLHRKPVK	RTVDGLHPV	0	0	0	7	4 RTVDGLHLHRKPV	vdchain	0.176	7461.17	17.00	
17	HLA-A0203	TVDGLHLHRKPVKR	TLHLHRKPV	0	0	0	1	3 TVDGLHLHRKPV	vdchain	0.100	17024.33	31.00	
18	HLA-A0203	VDGLHLHRKPVKRL	GLHLHVKRL	2	0	0	5	3 GLHLHRKPVKRL	vdchain	0.136	11441.96	23.00	
19	HLA-A0203	DGLHLHRKPVKRLC	GLHLHVKRL	1	0	0	5	3 GLHLHRKPVKRL	vdchain	0.086	19762.01	35.00	

Protein vdchain. Allele HLA-A0203. Number of high binders 0. Number of weak binders 0. Number of peptides 20

Link to Allele Frequencies in Worldwide Populations[HLA-A0203](#)

Rank Threshold for Strong binding peptides 0.500

Rank Threshold for Weakbindingpeptides 2.000

pos	HLA	peptide	CoreOffset	I_pos	I_len	D_pos	D_len	iCore	Identity	1-log50k(aff)	Affinity(nM)	%Rank	BindLevel
0	HLA-A0205	HCRCQKPCIRHQYC	RCQKPCICYC	2	0	0	7	3 RCQKPCIRHQYC	vdchain	0.011	44559.96	43.00	
1	HLA-A0205	CRCQKPCIRHQYCL	RCQKPCICYC	1	0	0	7	3 RCQKPCIRHQYC	vdchain	0.011	44629.91	43.00	
2	HLA-A0205	RCQKPCIRHQYCLC	RCQKPCICYC	0	0	0	7	3 RCQKPCIRHQYC	vdchain	0.010	44757.60	44.00	
3	HLA-A0205	CQKPCIRHQYCLCR	KCIRHQYCL	2	0	0	1	1 KPCIRHQYCL	vdchain	0.008	45964.89	55.00	
4	HLA-A0205	QKPCIRHQYCLCRT	QKPRHQYCL	0	0	0	3	2 QKPCIRHQYCL	vdchain	0.008	45875.46	55.00	
5	HLA-A0205	KPCIRHQYCLCRTV	IRHQYCLTV	3	0	0	7	2 IRHQYCLCRTV	vdchain	0.053	28251.89	9.50	
6	HLA-A0205	PCIRHQYCLCRTVD	IRHQYCLTV	2	0	0	7	2 IRHQYCLCRTV	vdchain	0.047	29949.80	11.00	

7	HLA-A0205	CIRHQYCLCRTVDG	IRHQYCLTV	1	0	0	7	2	IRHQYCLCRTV	vdchain	0.039	32724.55	13.00
8	HLA-A0205	IRHQYCLCRTVDGL	IRHQYCLTV	0	0	0	7	2	IRHQYCLCRTV	vdchain	0.047	30047.18	11.00
9	HLA-A0205	RHQYCLCRTVDGLH	YLCRTVDGL	3	0	0	1	1	YLCRTVDGL	vdchain	0.025	38353.73	21.00
10	HLA-A0205	HQYCLCRTVDGLHL	YLCRTVDGL	2	0	0	1	1	YLCRTVDGL	vdchain	0.028	36967.64	18.00
11	HLA-A0205	QYCLCRTVDGLHLH	YLCRTVDGL	1	0	0	1	1	YLCRTVDGL	vdchain	0.025	38200.09	20.00
12	HLA-A0205	YLCRTVDGLHLHR	YLCRTVDGL	0	0	0	1	1	YLCRTVDGL	vdchain	0.025	38195.12	20.00
13	HLA-A0205	CLCRTVDGLHLHRK	CLVDGLHLH	0	0	0	2	3	CLCRTVDGLHLH	vdchain	0.014	42792.79	33.00
14	HLA-A0205	LCRTVDGLHLHRKP	LCRTVGLHL	0	0	0	5	1	LCRTVDGLHL	vdchain	0.006	47043.10	70.00
15	HLA-A0205	CRTVDGLHLHRKPV	CLHLHRKPV	0	0	0	1	5	CRTVDGLHLHRKPV	vdchain	0.035	34381.10	15.00
16	HLA-A0205	RTVDGLHLHRKPVK	RTVDGLHPV	0	0	0	7	4	RTVDGLHLHRKPV	vdchain	0.030	36091.83	17.00
17	HLA-A0205	TVDGLHLHRKPVKR	VLHLHRKPV	1	0	0	1	2	VDGLHLHRKPV	vdchain	0.025	38187.29	20.00
18	HLA-A0205	VDGLHLHRKPVKRL	VLHLHRKPV	0	0	0	1	2	VDGLHLHRKPV	vdchain	0.026	37810.68	20.00
19	HLA-A0205	DGLHLHRKPVKRLC	GLHLHRKPV	1	0	0	0	0	GLHLHRKPV	vdchain	0.022	39617.80	23.00

Protein vdchain. Allele HLA-A0205. Number of high binders 0. Number of weak binders 0. Number of peptides 20

Link to Allele Frequencies in Worldwide Populations[HLA-A0205](#)

Rank Threshold for Strong binding peptides 0.500
Rank Threshold for Weakbindingpeptides 2.000

pos	HLA	peptide	CoreOffset	I_pos	I_len	D_pos	D_len	iCore	Identity	1-log50k(aff)	Affinity(nM)	%Rank	BindLevel
0	HLA-A0206	HCRCQKPCIRHQYC	CQKPCIQYC	3	0	0	6	2	CQKPCIRHQYC	vdchain	0.020	40313.12	80.00
1	HLA-A0206	CRCQKPCIRHQYCL	CQKPCIQYCL	2	0	0	6	3	CQKPCIRHQYCL	vdchain	0.063	25390.63	43.00
2	HLA-A0206	RCQKPCIRHQYCLC	CQKPCIQYCL	1	0	0	6	3	CQKPCIRHQYCL	vdchain	0.044	31062.34	55.00
3	HLA-A0206	CQKPCIRHQYCLCR	CQKPCIQYCL	0	0	0	6	3	CQKPCIRHQYCL	vdchain	0.050	29116.79	49.00
4	HLA-A0206	QKPCIRHQYCLCRT	CIRHCLCRT	3	0	0	4	2	CIRHQYCLCRT	vdchain	0.028	37012.06	70.00
5	HLA-A0206	KPCIRHQYCLCRTV	KQYCLCRTV	0	0	0	1	5	KPCIRHQYCLCRTV	vdchain	0.300	1939.86	7.50
6	HLA-A0206	PCIRHQYCLCRTVD	IQYCLCRTV	2	0	0	1	2	IRHQYCLCRTV	vdchain	0.158	9026.86	19.00
7	HLA-A0206	CIRHQYCLCRTVDG	CQYCLCRTV	0	0	0	1	3	CIRHQYCLCRTV	vdchain	0.164	8484.58	18.00
8	HLA-A0206	IRHQYCLCRTVDGL	IQYCLCRTV	0	0	0	1	2	IRHQYCLCRTV	vdchain	0.183	6927.12	16.00
9	HLA-A0206	RHQYCLCRTVDGLH	RQYCLCRTV	0	0	0	1	1	RHQYCLCRTV	vdchain	0.144	10565.56	21.00
10	HLA-A0206	HQYCLCRTVDGLHL	HQYCLCLHL	0	0	0	6	5	HQYCLCRTVDGLHL	vdchain	0.267	2768.71	9.00
11	HLA-A0206	QYCLCRTVDGLHLH	YTV DGLHLH	1	0	0	1	4	YLCRTVDGLHLH	vdchain	0.050	28954.07	49.00
12	HLA-A0206	YLCRTVDGLHLHR	YVDGLHLHR	0	0	0	1	5	YLCRTVDGLHLHR	vdchain	0.057	27078.98	46.00
13	HLA-A0206	CLCRTVDGLHLHRK	CLCRTVLHL	0	0	0	6	2	CLCRTVDGLHL	vdchain	0.048	29623.65	50.00
14	HLA-A0206	LCRTVDGLHLHRKP	RTVDGLHLP	2	0	0	8	3	RTVDGLHLHRKP	vdchain	0.027	37347.17	70.00
15	HLA-A0206	CRTVDGLHLHRKPV	RTVDGLHPV	1	0	0	7	4	RTVDGLHLHRKPV	vdchain	0.400	658.29	4.00
16	HLA-A0206	RTVDGLHLHRKPVK	RTVDGLHPV	0	0	0	7	4	RTVDGLHLHRKPV	vdchain	0.318	1602.37	6.50
17	HLA-A0206	TVDGLHLHRKPVKR	TVDGLHLPV	0	0	0	7	3	TVDGLHLHRKPV	vdchain	0.139	11139.62	22.00
18	HLA-A0206	VDGLHLHRKPVKRL	VLHLHRKPV	0	0	0	1	2	VDGLHLHRKPV	vdchain	0.045	30772.97	55.00
19	HLA-A0206	DGLHLHRKPVKRLC	GLHLHRKPV	1	0	0	0	0	GLHLHRKPV	vdchain	0.027	37336.64	70.00

Protein vdchain. Allele HLA-A0206. Number of high binders 0. Number of weak binders 0. Number of peptides 20

Link to Allele Frequencies in Worldwide Populations[HLA-A0206](#)

Rank Threshold for Strongbindingpeptides 0.500

Rank Threshold for Weakbindingpeptides 2.000

pos	HLA	peptide	CoreOffset	I_pos	I_len	D_pos	D_len	iCore	Identity	1-log50k(aff)	Affinity(nM)	%Rank	BindLevel
0	HLA-A0207	HCRCQKPCIRHQYC	HCRCQKPCI	0	0	0	0	0	HCRCQKPCI	vdchain	0.014	42796.49	40.00
1	HLA-A0207	CRCQKPCIRHQYCL	QKPCIYCL	3	0	0	5	2	QKPCIRHQYCL	vdchain	0.015	42447.41	37.00
2	HLA-A0207	RCQKPCIRHQYCLC	QKPCIYCLC	2	0	0	5	3	QKPCIRHQYCLC	vdchain	0.016	42029.68	34.00
3	HLA-A0207	CQKPCIRHQYCLCR	QKPCIYCLC	1	0	0	5	3	QKPCIRHQYCLC	vdchain	0.015	42310.29	36.00
4	HLA-A0207	QKPCIRHQYCLCRT	QKPCIYCLC	0	0	0	5	3	QKPCIRHQYCLC	vdchain	0.015	42297.03	36.00
5	HLA-A0207	KPCIRHQYCLCRTV	RHQYCLCTV	4	0	0	7	1	RHQYCLCRTV	vdchain	0.025	37952.50	18.00
6	HLA-A0207	PCIRHQYCLCRTVD	RHQYCLCTV	3	0	0	7	1	RHQYCLCRTV	vdchain	0.025	38278.69	19.00
7	HLA-A0207	CIRHQYCLCRTVDG	RHQYCLCTV	2	0	0	7	1	RHQYCLCRTV	vdchain	0.024	38407.30	19.00
8	HLA-A0207	IRHQYCLCRTVDGL	RHQYCLCTV	1	0	0	7	1	RHQYCLCRTV	vdchain	0.021	39826.25	24.00
9	HLA-A0207	RHQYCLCRTVDGLH	RHQYCLCTV	0	0	0	7	1	RHQYCLCRTV	vdchain	0.021	39867.65	24.00
10	HLA-A0207	HQYCLCRTVDGLHL	YLCRTVDGL	2	0	0	1	1	YCLCRTVDGL	vdchain	0.017	41411.29	31.00
11	HLA-A0207	QYCLCRTVDGLHLH	YLCRTVDGL	1	0	0	1	1	YCLCRTVDGL	vdchain	0.016	42020.14	34.00
12	HLA-A0207	YLCRTVDGLHLHR	YVDGLHLHR	0	0	0	1	5	YCLCRTVDGLHLHR	vdchain	0.018	41013.52	29.00
13	HLA-A0207	CLCRTVDGLHLHRK	CLDGLHLHR	0	0	0	2	4	CLCRTVDGLHLHR	vdchain	0.015	42331.36	36.00
14	HLA-A0207	LCRTVDGLHLHRKP	RVDGLHLHR	2	0	0	1	1	RTVDGLHLHR	vdchain	0.011	44551.29	60.00
15	HLA-A0207	CRTVDGLHLHRKPV	TVDGLHLRPV	2	0	0	7	3	TVDGLHLHRKPV	vdchain	0.082	20630.50	2.00 <= WB
16	HLA-A0207	RTVDGLHLHRKPVK	RLHLHRKPV	0	0	0	1	4	RTVDGLHLHRKPV	vdchain	0.082	20557.64	1.90 <= WB
17	HLA-A0207	TVDGLHLHRKPVKR	GLHLHRKPV	3	0	0	0	0	GLHLHRKPV	vdchain	0.077	21756.83	2.50
18	HLA-A0207	VDGLHLHRKPVKRL	GLHLHRKPV	2	0	0	0	0	GLHLHRKPV	vdchain	0.057	27101.85	4.50
19	HLA-A0207	DGLHLHRKPVKRLC	GLHLHRKPV	1	0	0	0	0	GLHLHRKPV	vdchain	0.057	26987.73	4.50

Protein vdchain. Allele HLA-A0207. Number of high binders 0. Number of weak binders 2. Number of peptides 20

Link to Allele Frequencies in Worldwide Populations [HLA-A0207](#)

Rank Threshold for Strong binding peptides 0.500

Rank Threshold for Weakbindingpeptides 2.000

pos	HLA	peptide	CoreOffset	I_pos	I_len	D_pos	D_len	iCore	Identity	1-log50k(aff)	Affinity(nM)	%Rank	BindLevel
0	HLA-A0211	HCRCQKPCIRHQYC	CQCIHQYC	3	0	0	2	2	CQKPCIRHQYC	vdchain	0.011	44154.41	60.00
1	HLA-A0211	CRCQKPCIRHQYCL	CQKPCIYCL	2	0	0	6	3	CQKPCIRHQYCL	vdchain	0.019	40896.54	43.00
2	HLA-A0211	RCQKPCIRHQYCLC	CQKPCIYCL	1	0	0	6	3	CQKPCIRHQYCL	vdchain	0.017	41407.25	45.00
3	HLA-A0211	CQKPCIRHQYCLCR	CQKPCIYCL	0	0	0	6	3	CQKPCIRHQYCL	vdchain	0.017	41520.31	45.00
4	HLA-A0211	QKPCIRHQYCLCRT	CIQYCLCRT	3	0	0	2	2	CIRHQYCLCRT	vdchain	0.012	43809.42	60.00
5	HLA-A0211	KPCIRHQYCLCRTV	CIYCLCRTV	2	0	0	2	3	CIRHQYCLCRTV	vdchain	0.056	27232.36	19.00
6	HLA-A0211	PCIRHQYCLCRTVD	CIYCLCRTV	1	0	0	2	3	CIRHQYCLCRTV	vdchain	0.046	30246.79	22.00
7	HLA-A0211	CIRHQYCLCRTVDG	CIYCLCRTV	0	0	0	2	3	CIRHQYCLCRTV	vdchain	0.050	29222.21	21.00
8	HLA-A0211	IRHQYCLCRTVDGL	YLCRTVDGL	4	0	0	1	1	YLCRTVDGL	vdchain	0.082	20588.13	15.00
9	HLA-A0211	RHQYCLCRTVDGLH	YLCRTVDGL	3	0	0	1	1	YLCRTVDGL	vdchain	0.066	24456.35	17.00
10	HLA-A0211	HQYCLCRTVDGLHL	YLCRTVDGL	2	0	0	1	1	YLCRTVDGL	vdchain	0.082	20654.62	15.00
11	HLA-A0211	QYCLCRTVDGLHLH	YLCRTVDGL	1	0	0	1	1	YLCRTVDGL	vdchain	0.068	23830.49	17.00
12	HLA-A0211	YLCRTVDGLHLHR	YLCRTVDGL	0	0	0	1	1	YLCRTVDGL	vdchain	0.074	22533.56	16.00
13	HLA-A0211	CLCRTVDGLHLHRK	CLCRTVDGL	0	0	0	0	0	CLCRTVDGL	vdchain	0.071	23264.40	16.00
14	HLA-A0211	LCRTVDGLHLHRKP	RTVDGLHLP	2	0	0	8	3	RTVDGLHLHRKP	vdchain	0.008	45623.49	70.00
15	HLA-A0211	CRTVDGLHLHRKPV	RTVDGLHPV	1	0	0	7	4	RTVDGLHLHRKPV	vdchain	0.061	25902.02	18.00

16	HLA-A0211	RTVDGLHLHRKPVK	RTVDGLHPV	0	0	0	7	4	RTVDGLHLHRKPV	vdchain	0.059	26431.44	19.00
17	HLA-A0211	TVDGLHLHRKPVKR	TVDGLHLHV	0	0	0	8	3	TVDGLHLHRKPV	vdchain	0.028	37129.18	33.00
18	HLA-A0211	VDGLHLHRKPVKRL	VLHLHRKPV	0	0	0	1	2	VDGLHLHRKPV	vdchain	0.021	39933.26	40.00
19	HLA-A0211	DGLHLHRKPVKRLC	GLHLHRKPV	1	0	0	0	0	GLHLHRKPV	vdchain	0.017	41378.13	45.00

Protein vdchain. Allele HLA-A0211. Number of high binders 0. Number of weak binders 0. Number of peptides 20

Link to Allele Frequencies in Worldwide Populations [HLA-A0211](#)

Rank Threshold for Strong binding peptides 0.500

Rank Threshold for Weakbindingpeptides 2.000

pos	HLA	peptide	CoreOffset	I_pos	I_len	D_pos	D_len	iCore	Identity	1-log50k(aff)	Affinity(nM)	%Rank	BindLevel
0	HLA-A0212	HCRCQKPCIRHQYC	CQKPCIQYC	3	0	0	6	2	CQKPCIRHQYC	vdchain	0.014	42987.24	60.00
1	HLA-A0212	CRCQKPCIRHQYCL	CQIRHQYCL	2	0	0	2	3	CQKPCIRHQYCL	vdchain	0.046	30311.99	20.00
2	HLA-A0212	RCQKPCIRHQYCLC	CQIRHQYCL	1	0	0	2	3	CQKPCIRHQYCL	vdchain	0.038	33068.75	24.00
3	HLA-A0212	CQKPCIRHQYCLCR	CQIRHQYCL	0	0	0	2	3	CQKPCIRHQYCL	vdchain	0.039	32895.67	23.00
4	HLA-A0212	QKPCIRHQYCLCRT	KCIRHQYCL	1	0	0	1	1	KPCIRHQYCL	vdchain	0.014	42957.48	60.00
5	HLA-A0212	KPCIRHQYCLCRTV	HQYCLCRTV	5	0	0	0	0	HQYCLCRTV	vdchain	0.032	35403.76	28.00
6	HLA-A0212	PCIRHQYCLCRTVD	HQYCLCRTV	4	0	0	0	0	HQYCLCRTV	vdchain	0.027	37371.00	33.00
7	HLA-A0212	CIRHQYCLCRTVDG	CIYCLCRTV	0	0	0	2	3	CIRHQYCLCRTV	vdchain	0.027	37287.80	32.00
8	HLA-A0212	IRHQYCLCRTVDGL	CLCRTVDGL	5	0	0	0	0	CLCRTVDGL	vdchain	0.068	24076.43	14.00
9	HLA-A0212	RHQYCLCRTVDGLH	CLCRTVDGL	4	0	0	0	0	CLCRTVDGL	vdchain	0.055	27451.57	17.00
10	HLA-A0212	HQYCLCRTVDGLHL	YLCRTVDGL	2	0	0	1	1	YLCRTVDGL	vdchain	0.054	27739.39	17.00
11	HLA-A0212	QYCLCRTVDGLHLH	YLCRTVDGL	1	0	0	1	1	YLCRTVDGL	vdchain	0.053	28301.77	18.00
12	HLA-A0212	YLCRTVDGLHLHR	YLCRTVDGL	0	0	0	1	1	YLCRTVDGL	vdchain	0.056	27379.79	17.00
13	HLA-A0212	CLCRTVDGLHLHRK	CLCRTVDGL	0	0	0	0	0	CLCRTVDGL	vdchain	0.046	30311.99	20.00
14	HLA-A0212	LCRTVDGLHLHRKP	RTVDGLHLP	2	0	0	8	3	RTVDGLHLHRKP	vdchain	0.009	45472.18	80.00
15	HLA-A0212	CRTVDGLHLHRKPV	RTVDGLHPV	1	0	0	7	4	RTVDGLHLHRKPV	vdchain	0.053	28151.21	18.00
16	HLA-A0212	RTVDGLHLHRKPVK	RTVDGLHPV	0	0	0	7	4	RTVDGLHLHRKPV	vdchain	0.046	30362.54	20.00
17	HLA-A0212	TVDGLHLHRKPVKR	GLHLHRKPV	3	0	0	0	0	GLHLHRKPV	vdchain	0.023	38871.78	37.00
18	HLA-A0212	VDGLHLHRKPVKRL	GLHLHRKPV	2	0	0	0	0	GLHLHRKPV	vdchain	0.023	38940.38	37.00
19	HLA-A0212	DGLHLHRKPVKRLC	GLHLHRKPV	1	0	0	0	0	GLHLHRKPV	vdchain	0.023	38927.32	37.00

Protein vdchain. Allele HLA-A0212. Number of high binders 0. Number of weak binders 0. Number of peptides 20

Link to Allele Frequencies in Worldwide Populations [HLA-A0212](#)

Rank Threshold for Strong binding peptides 0.500

Rank Threshold for Weakbindingpeptides 2.000

pos	HLA	peptide	CoreOffset	I_pos	I_len	D_pos	D_len	iCore	Identity	1-log50k(aff)	Affinity(nM)	%Rank	BindLevel
0	HLA-A0216	HCRCQKPCIRHQYC	HCRCQKPCI	0	0	0	0	0	HCRCQKPCI	vdchain	0.012	43974.68	55.00
1	HLA-A0216	CRCQKPCIRHQYCL	CQKPCIQYC	2	0	0	6	3	CQKPCIRHQYCL	vdchain	0.033	35101.66	25.00
2	HLA-A0216	RCQKPCIRHQYCLC	CQKPCIQYC	1	0	0	6	3	CQKPCIRHQYCL	vdchain	0.029	36363.06	27.00
3	HLA-A0216	CQKPCIRHQYCLCR	CQKPCIQYC	0	0	0	6	3	CQKPCIRHQYCL	vdchain	0.033	35121.43	25.00
4	HLA-A0216	QKPCIRHQYCLCRT	KCIRHQYCL	1	0	0	1	1	KPCIRHQYCL	vdchain	0.014	43138.18	49.00

5	HLA-A0216	KPCIRHQYCLCRTV	KQYCLCRTV	0	0	0	1	5	KPCIRHQYCLCRTV	vdchain	0.025	38000.16	31.00
6	HLA-A0216	PCIRHQYCLCRTVD	HQYCLCRTV	4	0	0	0	0	HQYCLCRTV	vdchain	0.021	39768.55	35.00
7	HLA-A0216	CIRHQYCLCRTVDG	CQYCLCRTV	0	0	0	1	3	CIRHQYCLCRTV	vdchain	0.022	39381.05	34.00
8	HLA-A0216	IRHQYCLCRTVDGL	YLCRTVDGL	4	0	0	1	1	YLCRTVDGL	vdchain	0.035	34387.44	24.00
9	HLA-A0216	RHQYCLCRTVDGLH	YLCRTVDGL	3	0	0	1	1	YLCRTVDGL	vdchain	0.031	35861.00	26.00
10	HLA-A0216	HQYCLCRTVDGLHL	YLCRTVDGL	2	0	0	1	1	YLCRTVDGL	vdchain	0.031	35942.57	26.00
11	HLA-A0216	QYCLCRTVDGLHLH	YLCRTVDGL	1	0	0	1	1	YLCRTVDGL	vdchain	0.030	36329.63	27.00
12	HLA-A0216	YLCRTVDGLHLHR	YLCRTVDGL	0	0	0	1	1	YLCRTVDGL	vdchain	0.034	34533.58	24.00
13	HLA-A0216	CLCRTVDGLHLHRK	CLCRTVDGL	0	0	0	0	0	CLCRTVDGL	vdchain	0.027	37146.07	29.00
14	HLA-A0216	LCRTVDGLHLHRKP	LVDGLHLHR	0	0	0	1	3	LCRTVDGLHLHR	vdchain	0.006	46852.12	85.00
15	HLA-A0216	CRTVDGLHLHRKPV	RTVDGLHPV	1	0	0	7	4	RTVDGLHLHRKPV	vdchain	0.023	38917.23	33.00
16	HLA-A0216	RTVDGLHLHRKPVK	RTVDGLHPV	0	0	0	7	4	RTVDGLHLHRKPV	vdchain	0.023	38913.84	33.00
17	HLA-A0216	TVDGLHLHRKPVKR	TVDGLHLPV	0	0	0	7	3	TVDGLHLHRKPV	vdchain	0.019	40782.11	39.00
18	HLA-A0216	VDGLHLHRKPVKRL	GLHLHRKPV	2	0	0	0	0	GLHLHRKPV	vdchain	0.012	43820.78	55.00
19	HLA-A0216	DGLHLHRKPVKRLC	GLHLHRKPV	1	0	0	0	0	GLHLHRKPV	vdchain	0.011	44306.11	60.00

Protein vdchain. Allele HLA-A0216. Number of high binders 0. Number of weak binders 0. Number of peptides 20

Link to Allele Frequencies in Worldwide Populations [HLA-A0216](#)

Rank Threshold for Strong binding peptides 0.500

Rank Threshold for Weakbindingpeptides 2.000

pos	HLA	peptide	CoreOffset	I_pos	I_len	D_pos	D_len	iCore	Identity	1-log50k(aff)	Affinity(nM)	%Rank	BindLevel
0	HLA-A0217	HCRCQKPCIRHQYC	RCQKPCICYC	2	0	0	7	3	RCQKPCIRHQYC	vdchain	0.006	46770.57	80.00
1	HLA-A0217	CRCQKPCIRHQYCL	RCQKPCIRL	1	0	0	8	4	RCQKPCIRHQYCL	vdchain	0.017	41414.87	43.00
2	HLA-A0217	RCQKPCIRHQYCLC	RCQKPCIRL	0	0	0	8	4	RCQKPCIRHQYCL	vdchain	0.014	42747.91	48.00
3	HLA-A0217	CQKPCIRHQYCLCR	CIHQYCLCR	4	0	0	2	1	CIRHQYCLCR	vdchain	0.013	43347.34	55.00
4	HLA-A0217	QKPCIRHQYCLCRT	CIHQYCLCR	3	0	0	2	1	CIRHQYCLCR	vdchain	0.011	44362.23	60.00
5	HLA-A0217	KPCIRHQYCLCRTV	CIYCLCRTV	2	0	0	2	3	CIRHQYCLCRTV	vdchain	0.032	35387.67	28.00
6	HLA-A0217	PCIRHQYCLCRTVD	CIYCLCRTV	1	0	0	2	3	CIRHQYCLCRTV	vdchain	0.027	37416.31	32.00
7	HLA-A0217	CIRHQYCLCRTVDG	CIYCLCRTV	0	0	0	2	3	CIRHQYCLCRTV	vdchain	0.025	38231.94	34.00
8	HLA-A0217	IRHQYCLCRTVDGL	YLCRTVDGL	4	0	0	1	1	YLCRTVDGL	vdchain	0.091	18765.53	11.00
9	HLA-A0217	RHQYCLCRTVDGLH	YLCRTVDGL	3	0	0	1	1	YLCRTVDGL	vdchain	0.075	22271.77	13.00
10	HLA-A0217	HQYCLCRTVDGLHL	CLCRTVLHL	3	0	0	6	2	CLCRTVDGLHL	vdchain	0.085	20024.60	12.00
11	HLA-A0217	QYCLCRTVDGLHLH	CLCRTVLHL	2	0	0	6	2	CLCRTVDGLHL	vdchain	0.075	22268.88	13.00
12	HLA-A0217	YLCRTVDGLHLHR	YLCRTVDGL	0	0	0	1	1	YLCRTVDGL	vdchain	0.071	23140.63	14.00
13	HLA-A0217	CLCRTVDGLHLHRK	CLCRTVLHL	0	0	0	6	2	CLCRTVDGLHL	vdchain	0.060	26012.40	16.00
14	HLA-A0217	LCRTVDGLHLHRKP	LCRTVDLHL	0	0	0	6	1	LCRTVDGLHL	vdchain	0.009	45233.72	65.00
15	HLA-A0217	CRTVDGLHLHRKPV	TVDGLHLPV	2	0	0	7	3	TVDGLHLHRKPV	vdchain	0.028	36935.26	31.00
16	HLA-A0217	RTVDGLHLHRKPVK	TVDGLHLPV	1	0	0	7	3	TVDGLHLHRKPV	vdchain	0.023	38791.09	35.00
17	HLA-A0217	TVDGLHLHRKPVKR	TVDGLHLPV	0	0	0	7	3	TVDGLHLHRKPV	vdchain	0.022	39447.57	37.00
18	HLA-A0217	VDGLHLHRKPVKRL	HLHKPVKRL	4	0	0	3	1	HLHRKPVKRL	vdchain	0.065	24874.53	15.00
19	HLA-A0217	DGLHLHRKPVKRLC	HLHKPVKRL	3	0	0	3	1	HLHRKPVKRL	vdchain	0.053	28086.71	18.00

Protein vdchain. Allele HLA-A0217. Number of high binders 0. Number of weak binders 0. Number of peptides 20

Link to Allele Frequencies in Worldwide Populations [HLA-A0217](#)

Rank Threshold for Strong binding peptides 0.500

Rank Threshold for Weakbindingpeptides 2.000

pos	HLA	peptide	CoreOffset	I_pos	I_len	D_pos	D_len	iCore	Identity	1-log50k(aff)	Affinity(nM)	%Rank	BindLevel
0	HLA-A0219	HCRCQKPCIRHQYC	CQKPCIQYC	3	0	0	6	2 CQKPCIRHQYC	vdchain	0.014	43057.05	65.00	
1	HLA-A0219	CRCQKPCIRHQYCL	CQIRHQYCL	2	0	0	2	3 CQKPCIRHQYCL	vdchain	0.025	38003.44	33.00	
2	HLA-A0219	RCQKPCIRHQYCLC	CQIRHQYCL	1	0	0	2	3 CQKPCIRHQYCL	vdchain	0.025	38122.05	33.00	
3	HLA-A0219	CQKPCIRHQYCLCR	CQIRHQYCL	0	0	0	2	3 CQKPCIRHQYCL	vdchain	0.028	37124.37	30.00	
4	HLA-A0219	QKPCIRHQYCLCRT	PCIRHQYCL	2	0	0	0	0 PCIRHQYCL	vdchain	0.013	43289.68	65.00	
5	HLA-A0219	KPCIRHQYCLCRTV	CIYCLCRTV	2	0	0	2	3 CIRHQYCLCRTV	vdchain	0.017	41684.14	50.00	
6	HLA-A0219	PCIRHQYCLCRTVD	CIYCLCRTV	1	0	0	2	3 CIRHQYCLCRTV	vdchain	0.016	42094.32	55.00	
7	HLA-A0219	CIRHQYCLCRTVDG	CIYCLCRTV	0	0	0	2	3 CIRHQYCLCRTV	vdchain	0.018	41361.14	48.00	
8	HLA-A0219	IRHQYCLCRTVDGL	YLCRTVDGL	4	0	0	1	1 YLCRTVDGL	vdchain	0.048	29747.63	17.00	
9	HLA-A0219	RHQYCLCRTVDGLH	YLCRTVDGL	3	0	0	1	1 YLCRTVDGL	vdchain	0.047	29924.21	17.00	
10	HLA-A0219	HQYCLCRTVDGLHL	YLCRTVDGL	2	0	0	1	1 YLCRTVDGL	vdchain	0.047	29990.34	17.00	
11	HLA-A0219	QYCLCRTVDGLHLH	YLCRTVDGL	1	0	0	1	1 YLCRTVDGL	vdchain	0.049	29302.32	16.00	
12	HLA-A0219	YLCRTVDGLHLHR	YLCRTVDGL	0	0	0	1	1 YLCRTVDGL	vdchain	0.056	27407.95	14.00	
13	HLA-A0219	CLCRTVDGLHLHRK	CLCRTVDGL	0	0	0	0	0 CLCRTVDGL	vdchain	0.052	28503.37	15.00	
14	HLA-A0219	LCRTVDGLHLHRKP	CTVDGLHLH	1	0	0	1	1 CRTVDGLHLH	vdchain	0.009	45164.75	85.00	
15	HLA-A0219	CRTVDGLHLHRKPV	RTVDGLHPV	1	0	0	7	4 RTVDGLHLHRKPV	vdchain	0.028	37002.47	29.00	
16	HLA-A0219	RTVDGLHLHRKPVK	RTVDGLHPV	0	0	0	7	4 RTVDGLHLHRKPV	vdchain	0.030	36175.49	27.00	
17	HLA-A0219	TVDGLHLHRKPVKR	TVDGLHLPV	0	0	0	7	3 TVDGLHLHRKPV	vdchain	0.021	40046.64	41.00	
18	HLA-A0219	VDGLHLHRKPVKRL	VDGLHLPV	0	0	0	7	2 VDGLHLHRKPV	vdchain	0.021	39984.29	40.00	
19	HLA-A0219	DGLHLHRKPVKRLC	GLHLHRKPV	1	0	0	0	0 GLHLHRKPV	vdchain	0.017	41422.48	49.00	

Protein vdchain. Allele HLA-A0219. Number of high binders 0. Number of weak binders 0. Number of peptides 20

Link to Allele Frequencies in Worldwide Populations [HLA-A0219](#)

Rank Threshold for Strong binding peptides 0.500

Rank Threshold for Weakbindingpeptides 2.000

pos	HLA	peptide	CoreOffset	I_pos	I_len	D_pos	D_len	iCore	Identity	1-log50k(aff)	Affinity(nM)	%Rank	BindLevel
0	HLA-A0250	HCRCQKPCIRHQYC	CQKPCIQYC	3	0	0	6	2 CQKPCIRHQYC	vdchain	0.008	45773.79	55.00	
1	HLA-A0250	CRCQKPCIRHQYCL	QKPCIRHQL	3	0	0	8	2 QKPCIRHQYCL	vdchain	0.008	45948.47	60.00	
2	HLA-A0250	RCQKPCIRHQYCLC	CIRHQYCLC	5	0	0	0	0 CIRHQYCLC	vdchain	0.037	33576.37	24.00	
3	HLA-A0250	CQKPCIRHQYCLCR	CIRHQYCLC	4	0	0	0	0 CIRHQYCLC	vdchain	0.026	37535.95	29.00	
4	HLA-A0250	QKPCIRHQYCLCRT	CIRHQYCLC	3	0	0	0	0 CIRHQYCLC	vdchain	0.029	36544.50	28.00	
5	HLA-A0250	KPCIRHQYCLCRTV	CIRHQYCTV	2	0	0	7	3 CIRHQYCLCRTV	vdchain	0.039	32717.48	23.00	
6	HLA-A0250	PCIRHQYCLCRTVD	CIRHQYCTV	1	0	0	7	3 CIRHQYCLCRTV	vdchain	0.031	35601.58	26.00	
7	HLA-A0250	CIRHQYCLCRTVDG	CIRHQYCTV	0	0	0	7	3 CIRHQYCLCRTV	vdchain	0.033	34952.73	26.00	
8	HLA-A0250	IRHQYCLCRTVDGL	YLCRTVDGL	4	0	0	1	1 YLCRTVDGL	vdchain	0.037	33374.26	24.00	
9	HLA-A0250	RHQYCLCRTVDGLH	YLCRTVDGL	3	0	0	1	1 YLCRTVDGL	vdchain	0.029	36471.82	28.00	
10	HLA-A0250	HQYCLCRTVDGLHL	YLCRTVDGL	2	0	0	1	1 YLCRTVDGL	vdchain	0.028	37112.30	29.00	
11	HLA-A0250	QYCLCRTVDGLHLH	YLCRTVDGL	1	0	0	1	1 YLCRTVDGL	vdchain	0.037	33585.82	24.00	
12	HLA-A0250	YLCRTVDGLHLHR	YLCRTVDGL	0	0	0	1	1 YLCRTVDGL	vdchain	0.034	34676.61	25.00	
13	HLA-A0250	CLCRTVDGLHLHRK	CLCDGLHLH	0	0	0	3	3 CLCRTVDGLHLH	vdchain	0.026	37721.18	30.00	

14	HLA-A0250	LCRTVDGLHLHRKP	RTVDGLHLP	2	0	0	8	3	RTVDGLHLHRKP	vdchain	0.006	47060.91	65.00
15	HLA-A0250	CRTVDGLHLHRKPV	GLHLHRKPV	5	0	0	0	0	GLHLHRKPV	vdchain	0.035	34349.49	25.00
16	HLA-A0250	RTVDGLHLHRKPVK	RTVDGLHLV	0	0	0	8	4	RTVDGLHLHRKPV	vdchain	0.028	37094.65	29.00
17	HLA-A0250	TVDGLHLHRKPVKR	GLHLHRKPV	3	0	0	0	0	GLHLHRKPV	vdchain	0.013	43660.81	44.00
18	HLA-A0250	VDGLHLHRKPVKRL	HLHKPVKRL	4	0	0	3	1	HLHRKPVKRL	vdchain	0.015	42409.75	40.00
19	HLA-A0250	DGLHLHRKPVKRLC	GLHLHRKPV	1	0	0	0	0	GLHLHRKPV	vdchain	0.017	41493.36	38.00

Protein vdchain. Allele HLA-A0250. Number of high binders 0. Number of weak binders 0. Number of peptides 20

Link to Allele Frequencies in Worldwide Populations[HLA-A0250](#)

Rank Threshold for Strong binding peptides 0.500

Rank Threshold for Weakbindingpeptides 2.000

pos	HLA	peptide	CoreOffset	I_pos	I_len	D_pos	D_len	iCore	Identity	1-log50k(aff)	Affinity(nM)	%Rank	BindLevel
0	HLA-A0301	HCRCQKPCIRHQYC	RCQKPCIRY	2	0	0	8	2	RCQKPCIRHQY	vdchain	0.039	32848.37	47.00
1	HLA-A0301	CRCQKPCIRHQYCL	RCQKPCIRY	1	0	0	8	2	RCQKPCIRHQY	vdchain	0.038	32995.46	48.00
2	HLA-A0301	RCQKPCIRHQYCLC	RIRHQYCLC	0	0	0	1	5	RCQKPCIRHQYCLC	vdchain	0.071	23153.65	23.00
3	HLA-A0301	CQKPCIRHQYCLCR	CIRHQYCLR	4	0	0	8	1	CIRHQYCLCR	vdchain	0.112	14928.14	13.00
4	HLA-A0301	QKPCIRHQYCLCRT	CIRHQYCLR	3	0	0	8	1	CIRHQYCLCR	vdchain	0.081	20793.86	20.00
5	HLA-A0301	KPCIRHQYCLCRTV	CIRHQYCLR	2	0	0	8	1	CIRHQYCLCR	vdchain	0.076	22018.28	21.00
6	HLA-A0301	PCIRHQYCLCRTVD	CIRHQYCLR	1	0	0	8	1	CIRHQYCLCR	vdchain	0.074	22559.91	22.00
7	HLA-A0301	CIRHQYCLCRTVDG	CIRHQYCLR	0	0	0	8	1	CIRHQYCLCR	vdchain	0.081	20900.33	20.00
8	HLA-A0301	IRHQYCLCRTVDGL	RLCRTVDGL	1	0	0	1	4	RHQYCLCRTVDGL	vdchain	0.046	30399.69	39.00
9	HLA-A0301	RHQYCLCRTVDGLH	HQYCLCGLH	1	0	0	6	4	HQYCLCRTVDGLH	vdchain	0.076	22020.66	21.00
10	HLA-A0301	HQYCLCRTVDGLHL	HQYCLCGLH	0	0	0	6	4	HQYCLCRTVDGLH	vdchain	0.071	23266.15	23.00
11	HLA-A0301	QYCLCRTVDGLHLH	CLCRTLHLH	2	0	0	5	3	CLCRTVDGLHLH	vdchain	0.081	20782.84	20.00
12	HLA-A0301	YCLCRTVDGLHLHR	CLCRTHLHR	1	0	0	5	4	CLCRTVDGLHLHR	vdchain	0.172	7768.27	7.00
13	HLA-A0301	CLCRTVDGLHLHRK	CLCRHLHRK	0	0	0	4	5	CLCRTVDGLHLHRK	vdchain	0.383	790.12	1.70 <= WB
14	HLA-A0301	LCRTVDGLHLHRKP	RTVLHLHRK	2	0	0	3	2	RTVDGLHLHRK	vdchain	0.227	4306.71	5.00
15	HLA-A0301	CRTVDGLHLHRKPV	RTVLHLHRK	1	0	0	3	2	RTVDGLHLHRK	vdchain	0.215	4884.75	5.00
16	HLA-A0301	RTVDGLHLHRKPVK	RTLHRKPVK	0	0	0	2	5	RTVDGLHLHRKPVK	vdchain	0.458	350.91	1.10 <= WB
17	HLA-A0301	TVDGLHLHRKPVKR	TVLHRKPVK	0	0	0	2	4	TVDGLHLHRKPVK	vdchain	0.213	4986.15	5.00
18	HLA-A0301	VDGLHLHRKPVKRL	GLLHRKPVK	2	0	0	2	1	GLHLHRKPVK	vdchain	0.175	7526.60	7.00
19	HLA-A0301	DGLHLHRKPVKRLC	GLLHRKPVK	1	0	0	2	1	GLHLHRKPVK	vdchain	0.174	7571.93	7.00

Protein vdchain. Allele HLA-A0301. Number of high binders 0. Number of weak binders 2. Number of peptides 20

Link to Allele Frequencies in Worldwide Populations[HLA-A0301](#)

Rank Threshold for Strong binding peptides 0.500

Rank Threshold for Weakbindingpeptides 2.000

pos	HLA	peptide	CoreOffset	I_pos	I_len	D_pos	D_len	iCore	Identity	1-log50k(aff)	Affinity(nM)	%Rank	BindLevel
0	HLA-A1101	HCRCQKPCIRHQYC	CQKPCIRHQY	3	0	0	4	1	CQKPCIRHQY	vdchain	0.029	36394.57	55.00
1	HLA-A1101	CRCQKPCIRHQYCL	CQIRHQYCL	2	0	0	2	3	CQKPCIRHQYCL	vdchain	0.028	36735.97	55.00
2	HLA-A1101	RCQKPCIRHQYCLC	CQKPCIRHQY	1	0	0	6	1	CQKPCIRHQY	vdchain	0.026	37561.95	60.00

3	HLA-A1101	CQKPCIRHQYCLCR	CQKPCILCR	0	0	0	6	5	CQKPCIRHQYCLCR	vdchain	0.069	23807.80	21.00
4	HLA-A1101	QKPCIRHQYCLCRT	CIHQYCLCR	3	0	0	2	1	CIRHQYCLCR	vdchain	0.029	36578.93	55.00
5	HLA-A1101	KPCIRHQYCLCRTV	CIHQYCLCR	2	0	0	2	1	CIRHQYCLCR	vdchain	0.027	37379.10	60.00
6	HLA-A1101	PCIRHQYCLCRTVD	CIHQYCLCR	1	0	0	2	1	CIRHQYCLCR	vdchain	0.023	38822.18	70.00
7	HLA-A1101	CIRHQYCLCRTVDG	CIHQYCLCR	0	0	0	2	1	CIRHQYCLCR	vdchain	0.024	38490.09	65.00
8	HLA-A1101	IRHQYCLCRTVDGL	RLCRTVDGL	1	0	0	1	4	RHQYCLCRTVDGL	vdchain	0.025	38299.82	65.00
9	HLA-A1101	RHQYCLCRTVDGLH	HQYTVDGLH	1	0	0	3	4	HQYCLCRTVDGLH	vdchain	0.027	37136.41	60.00
10	HLA-A1101	HQYCLCRTVDGLHL	HQYCLGLHL	0	0	0	5	5	HQYCLCRTVDGLHL	vdchain	0.032	35452.82	50.00
11	HLA-A1101	QYCLCRTVDGLHLH	QTVDDLHLH	0	0	0	1	5	QYCLCRTVDGLHLH	vdchain	0.072	23051.67	20.00
12	HLA-A1101	YCLCRTVDGLHLHR	RTVDGLHLR	4	0	0	8	1	RTVDGLHLHR	vdchain	0.256	3117.33	4.50
13	HLA-A1101	CLCRTVDGLHLHRK	RTVDGLHLK	3	0	0	8	2	RTVDGLHLHRK	vdchain	0.499	226.59	1.20 <= WB
14	HLA-A1101	LCRTVDGLHLHRKP	RTVDGLHLK	2	0	0	8	2	RTVDGLHLHRK	vdchain	0.335	1330.23	3.00
15	HLA-A1101	CRTVDGLHLHRKPV	RTVDGLHLK	1	0	0	8	2	RTVDGLHLHRK	vdchain	0.292	2131.17	4.00
16	HLA-A1101	RTVDGLHLHRKPVK	RTVDGLHLK	0	0	0	8	5	RTVDGLHLHRKPVK	vdchain	0.558	119.50	0.80 <= WB
17	HLA-A1101	TVDGLHLHRKPVKR	TVDGLHLVK	0	0	0	7	4	TVDGLHLHRKPVK	vdchain	0.223	4455.01	5.50
18	HLA-A1101	VDGLHLHRKPVKRL	GLHLHRPVK	2	0	0	6	1	GLHLHRKPVK	vdchain	0.045	30760.33	34.00
19	HLA-A1101	DGLHLHRKPVKRLC	GLHLHRPVK	1	0	0	6	1	GLHLHRKPVK	vdchain	0.042	31793.05	37.00

Protein vdchain. Allele HLA-A1101. Number of high binders 0. Number of weak binders 2. Number of peptides 20

Link to Allele Frequencies in Worldwide Populations [HLA-A1101](#)

Rank Threshold for Strong binding peptides 0.500

Rank Threshold for Weakbindingpeptides 2.000

pos	HLA	peptide	CoreOffset	I_pos	I_len	D_pos	D_len	iCore	Identity	1-log50k(aff)	Affinity(nM)	%Rank	BindLevel
0	HLA-A2301	HCRCQKPCIRHQYC	CQKPCIQYC	3	0	0	6	2	CQKPCIRHQYC	vdchain	0.013	43376.42	75.00
1	HLA-A2301	CRCQKPCIRHQYCL	CQKPCIYCL	2	0	0	6	3	CQKPCIRHQYCL	vdchain	0.047	30122.36	26.00
2	HLA-A2301	RCQKPCIRHQYCLC	CQKPCIYCL	1	0	0	6	3	CQKPCIRHQYCL	vdchain	0.036	33754.49	33.00
3	HLA-A2301	CQKPCIRHQYCLCR	CQKPCIYCL	0	0	0	6	3	CQKPCIRHQYCL	vdchain	0.032	35209.31	36.00
4	HLA-A2301	QKPCIRHQYCLCRT	QCIRHQYCL	0	0	0	1	2	QKPCIRHQYCL	vdchain	0.020	40444.18	55.00
5	HLA-A2301	KPCIRHQYCLCRTV	RHQYCLCTV	4	0	0	7	1	RHQYCLCRTV	vdchain	0.045	30639.76	27.00
6	HLA-A2301	PCIRHQYCLCRTVD	RYCLCRTVD	3	0	0	1	2	RHQYCLCRTVD	vdchain	0.045	30776.98	27.00
7	HLA-A2301	CIRHQYCLCRTVDG	QYCLCRTVDG	4	0	0	2	1	QYCLCRTVDG	vdchain	0.053	28113.16	23.00
8	HLA-A2301	IRHQYCLCRTVDGL	QYCLCRTVL	3	0	0	8	2	QYCLCRTVDGL	vdchain	0.085	19960.15	14.00
9	HLA-A2301	RHQYCLCRTVDGLH	QYCLCRTVL	2	0	0	8	2	QYCLCRTVDGL	vdchain	0.060	26019.72	20.00
10	HLA-A2301	HQYCLCRTVDGLHL	QYCLCRLHL	1	0	0	6	4	QYCLCRTVDGLHL	vdchain	0.153	9569.98	7.00
11	HLA-A2301	QYCLCRTVDGLHLH	QYCLCRLHL	0	0	0	6	4	QYCLCRTVDGLHL	vdchain	0.114	14642.12	10.00
12	HLA-A2301	YCLCRTVDGLHLHR	YCLCRTLHL	0	0	0	6	3	YCLCRTVDGLHL	vdchain	0.045	30749.68	27.00
13	HLA-A2301	CLCRTVDGLHLHRK	CLCRTVLHL	0	0	0	6	2	CLCRTVDGLHL	vdchain	0.025	38287.40	46.00
14	HLA-A2301	LCRTVDGLHLHRKP	RTVDGLHLP	2	0	0	8	3	RTVDGLHLHRKP	vdchain	0.017	41600.80	65.00
15	HLA-A2301	CRTVDGLHLHRKPV	RTVDGLHLV	1	0	0	8	4	RTVDGLHLHRKPV	vdchain	0.030	36002.89	39.00
16	HLA-A2301	RTVDGLHLHRKPVK	RTVDGLHLV	0	0	0	8	4	RTVDGLHLHRKPV	vdchain	0.027	37405.00	43.00
17	HLA-A2301	TVDGLHLHRKPVKR	VHLHRKPVK	1	0	0	1	3	VDGLHLHRKPVK	vdchain	0.012	43705.71	75.00
18	HLA-A2301	VDGLHLHRKPVKRL	LHLHRKPVL	3	0	0	8	2	LHLHRKPVKRL	vdchain	0.028	36741.95	41.00
19	HLA-A2301	DGLHLHRKPVKRLC	LHLHRKPVL	2	0	0	8	2	LHLHRKPVKRL	vdchain	0.021	39789.64	55.00

Protein vdchain. Allele HLA-A2301. Number of high binders 0. Number of weak binders 0. Number of peptides 20

Link to Allele Frequencies in Worldwide Populations [HLA-A2301](#)

Rank Threshold for Strong binding peptides 0.500

Rank Threshold for Weakbindingpeptides 2.000

pos	HLA	peptide	CoreOffset	I_pos	I_len	D_pos	D_len	iCore	Identity	1-log50k(aff)	Affinity(nM)	%Rank	BindLevel
0	HLA-A2402	HCRCQKPCIRHQYC	RCQKRHQYC	2	0	0	4	3 RCQKPCIRHQYC	vdchain	0.016	41953.37	60.00	
1	HLA-A2402	CRCQKPCIRHQYCL	RCIRHQYCL	1	0	0	1	4 RCQKPCIRHQYCL	vdchain	0.059	26468.64	19.00	
2	HLA-A2402	RCQKPCIRHQYCLC	RCIRHQYCL	0	0	0	1	4 RCQKPCIRHQYCL	vdchain	0.053	28167.04	21.00	
3	HLA-A2402	CQKPCIRHQYCLCR	CQIRHQYCL	0	0	0	2	3 CQKPCIRHQYCL	vdchain	0.042	31623.57	26.00	
4	HLA-A2402	QKPCIRHQYCLCRT	KCIRHQYCL	1	0	0	1	1 KPCIRHQYCL	vdchain	0.024	38443.90	42.00	
5	HLA-A2402	KPCIRHQYCLCRTV	KCIRHQYCL	0	0	0	1	1 KPCIRHQYCL	vdchain	0.036	33951.20	30.00	
6	HLA-A2402	PCIRHQYCLCRTVD	RYCLCRTVD	3	0	0	1	2 RHQYCLCRTVD	vdchain	0.042	31629.04	26.00	
7	HLA-A2402	CIRHQYCLCRTVDG	RYCLCRTVD	2	0	0	1	2 RHQYCLCRTVD	vdchain	0.037	33659.67	29.00	
8	HLA-A2402	IRHQYCLCRTVDGL	QYCLCRTVL	3	0	0	8	2 QYCLCRTVDGL	vdchain	0.056	27335.38	20.00	
9	HLA-A2402	RHQYCLCRTVDGLH	RHQYCLCRL	0	0	0	8	4 RHQYCLCRTVDGL	vdchain	0.048	29720.61	23.00	
10	HLA-A2402	HQYCLCRTVDGLHL	QYCLCGLHL	1	0	0	5	4 QYCLCRTVDGLHL	vdchain	0.078	21430.44	14.00	
11	HLA-A2402	QYCLCRTVDGLHLH	QYCLCGLHL	0	0	0	5	4 QYCLCRTVDGLHL	vdchain	0.075	22245.53	15.00	
12	HLA-A2402	YCLCRTVDGLHLHR	YCLCRTLHL	0	0	0	6	3 YCLCRTVDGLHL	vdchain	0.036	33812.99	30.00	
13	HLA-A2402	CLCRTVDGLHLHRK	CLCRTVLHL	0	0	0	6	2 CLCRTVDGLHL	vdchain	0.025	38105.95	40.00	
14	HLA-A2402	LCRTVDGLHLHRKP	LCRTVGLHL	0	0	0	5	1 LCRTVDGLHL	vdchain	0.013	43335.15	65.00	
15	HLA-A2402	CRTVDGLHLHRKPV	RTVDGLHLV	1	0	0	8	4 RTVDGLHLHRKPV	vdchain	0.016	41839.13	60.00	
16	HLA-A2402	RTVDGLHLHRKPVK	RTVDGLHLV	0	0	0	8	4 RTVDGLHLHRKPV	vdchain	0.016	41945.20	60.00	
17	HLA-A2402	TVDGLHLHRKPVKR	TVDGLHLHV	0	0	0	8	3 TVDGLHLHRKPV	vdchain	0.011	44480.95	75.00	
18	HLA-A2402	VDGLHLHRKPVKRL	VDGLHLHRL	0	0	0	7	5 VDGLHLHRKPVKRL	vdchain	0.026	37784.09	39.00	
19	HLA-A2402	DGLHLHRKPVKRLC	LHLHRKPVL	2	0	0	8	2 LHLHRKPVKRL	vdchain	0.017	41438.16	55.00	

Protein vdchain. Allele HLA-A2402. Number of high binders 0. Number of weak binders 0. Number of peptides 20

Link to Allele Frequencies in Worldwide Populations [HLA-A2402](#)

[Explain](#) the output. Go [back](#).

Supplementary file 4

List of all matched Gene Ontology (GO) terms for cn40

Type	GO no.	Description	Score	BLAST	IPR	SUP	PDB	SSM	ENZ	DNA	LIG	REV
Cellular component												
	0019814	immunoglobulin complex	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0009897	external side of plasma membrane	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0042571	immunoglobulin complex , circulating	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0032991	macromolecular complex	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0043234	protein complex	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0005623	cell	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0044464	cell part	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0009986	cell surface	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0016020	membrane	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0044425	membrane part	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0044459	plasma membrane part	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0005886	plasma membrane	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0005576	extracellular region	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0044421	extracellular region part	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0005615	extracellular space	0.90	-	-	-	-	-	-	-	-	0.90 (1)
Biological process												
	0002250	adaptive immune response	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0002376	immune system process	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0006910	phagocytosis , recognition	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0006911	phagocytosis , engulfment	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0006958	complement activation , classical pathway	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0042742	defense response to bacterium	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0045087	innate immune response	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0050853	B cell receptor signaling pathway	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0050871	positive regulation of B cell activation	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0006955	immune response	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0050896	response to stimulus	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0016043	cellular component organization	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0016044	membrane organization	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0010324	membrane invagination	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0006897	endocytosis	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0006909	phagocytosis	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0009987	cellular process	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0008037	cell recognition	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0016192	vesicle -mediated transport	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0051234	establishment of localization	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0006810	transport	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0051179	localization	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0065007	biological regulation	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0050789	regulation of biological process	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0048518	positive regulation of biological process	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0002684	positive regulation of immune system process	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0050778	positive regulation of immune response	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0002253	activation of immune response	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0006956	complement activation	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0048584	positive regulation of response to stimulus	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0002682	regulation of immune system process	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0050776	regulation of immune response	0.90	-	-	-	-	-	-	-	-	0.90 (1)
	0048583	regulation of response to stimulus	0.90	-	-	-	-	-	-	-	-	0.90 (1)

0002252	immune effector process	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0002443	leukocyte mediated immunity	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0002449	lymphocyte mediated immunity	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0019724	B cell mediated immunity	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0016064	immunoglobulin mediated immune response	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0002455	humoral immune response mediated by circulating immunoglobulin	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0002460	adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0006959	humoral immune response	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0008152	metabolic process	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0043170	macromolecule metabolic process	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0043283	biopolymer metabolic process	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0019538	protein metabolic process	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0051604	protein maturation	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0016485	protein processing	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0051605	protein maturation by peptide bond cleavage	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0002541	activation of plasma proteins involved in acute inflammatory response	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0010467	gene expression	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0044238	primary metabolic process	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0009605	response to external stimulus	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0009611	response to wounding	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0006954	inflammatory response	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0002526	acute inflammatory response	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0006950	response to stress	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0006952	defense response	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0051704	multi-organism process	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0051707	response to other organism	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0009617	response to bacterium	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0009607	response to biotic stimulus	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0002757	immune response\activating signal transduction	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0002429	immune response\activating cell surface receptor signaling pathway	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0050851	antigen receptor\mediated signaling pathway	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0050794	regulation of cellular process	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0007165	signal transduction	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0007166	cell surface receptor linked signal transduction	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0002768	immune response\regulating cell surface receptor signaling pathway	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0002764	immune response\regulating signal transduction	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0007154	cell communication	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0048522	positive regulation of cellular process	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0050867	positive regulation of cell activation	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0002696	positive regulation of leukocyte activation	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0051251	positive regulation of lymphocyte activation	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0050865	regulation of cell activation	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)
0002694	regulation of leukocyte	0.90	-	-	-	-	-	-	-	-	-	0.90	(1)

activation										
0051249	regulation of lymphocyte activation	0.90	-	-	-	-	-	-	-	0.90 (1)
0050864	regulation of B cell activation	0.90	-	-	-	-	-	-	-	0.90 (1)
0001775	cell activation	0.90	-	-	-	-	-	-	-	0.90 (1)
0045321	leukocyte activation	0.90	-	-	-	-	-	-	-	0.90 (1)
0046649	lymphocyte activation	0.90	-	-	-	-	-	-	-	0.90 (1)
0042113	B cell activation	0.90	-	-	-	-	-	-	-	0.90 (1)

Biochemical function

0003823	antigen binding	0.90	-	-	-	-	-	-	-	0.90 (1)
0034987	immunoglobulin receptor binding	0.90	-	-	-	-	-	-	-	0.90 (1)
0005488	binding	0.90	-	-	-	-	-	-	-	0.90 (1)
0005515	protein binding	0.90	-	-	-	-	-	-	-	0.90 (1)
0005102	receptor binding	0.90	-	-	-	-	-	-	-	0.90 (1)

The score in red is a measure of how strongly the term is predicted from the hits obtained by the different methods. The scores in blue show each method's contribution to the total score (with the number of relevant sequences/structures shown in brackets in grey). Click on any score to see the list of hits from the given program and the source of the given term.

List of all matched protein name terms for cn40

Terms extracted from protein names	Score	BLAST	IPR	SUP	PDB	SSM	ENZ	DNA	LIG	REV
fab	6.31	-	-	-	-	-	-	-	-	6.31 (11)
antibody	4.99	-	-	-	-	-	-	-	-	4.99 (12)
fragment	4.80	-	-	-	-	-	-	-	-	4.80 (9)
fab fragment	3.77	-	-	-	-	-	-	-	-	3.77 (6)
human	1.91	-	-	-	-	-	-	-	-	1.91 (3)
ige	1.80	-	-	-	-	-	-	-	-	1.80 (2)
neutralizing antibody	1.57	-	-	-	-	-	-	-	-	1.57 (2)
peptide	1.57	-	-	-	-	-	-	-	-	1.57 (2)
specific	1.54	-	-	-	-	-	-	-	-	1.54 (2)
rabbit	1.47	-	-	-	-	-	-	-	-	1.47 (3)
murine	1.19	-	-	-	-	-	-	-	-	1.19 (4)
conformation	1.18	-	-	-	-	-	-	-	-	1.18 (2)
germline	1.14	-	-	-	-	-	-	-	-	1.14 (4)
germline antibody	1.00	-	-	-	-	-	-	-	-	1.00 (2)
anti- 4-hydroxy-3- nitrophenyl -acetyl	0.97	-	-	-	-	-	-	-	-	0.97 (3)
4-hydroxy-3- nitrophenyl -acetyl murine	0.97	-	-	-	-	-	-	-	-	0.97 (3)
nitrophenyl -acetyl murine germline	0.97	-	-	-	-	-	-	-	-	0.97 (3)
antibody bbe6 12h3	0.97	-	-	-	-	-	-	-	-	0.97 (3)
-acetyl murine germline monoclonal	0.95	-	-	-	-	-	-	-	-	0.95 (2)
murine germline monoclonal antibody	0.95	-	-	-	-	-	-	-	-	0.95 (2)
germline monoclonal antibody bbe6	0.95	-	-	-	-	-	-	-	-	0.95 (2)
monoclonal antibody bbe6 12h3	0.95	-	-	-	-	-	-	-	-	0.95 (2)
antibody bbe6 12h3 fab	0.95	-	-	-	-	-	-	-	-	0.95 (2)
bbe6 12h3 fab fragment	0.95	-	-	-	-	-	-	-	-	0.95 (2)
phage display derived dodecapeptide	0.95	-	-	-	-	-	-	-	-	0.95 (2)
implications	0.90	-	-	-	-	-	-	-	-	0.90 (1)
hiv-1 rev dimer	0.90	-	-	-	-	-	-	-	-	0.90 (1)
multimeric binding response element	0.90	-	-	-	-	-	-	-	-	0.90 (1)
calcium-activated chloride channel bestrophin-1	0.90	-	-	-	-	-	-	-	-	0.90 (1)
chicken	0.90	-	-	-	-	-	-	-	-	0.90 (1)
fab antibody fragments	0.90	-	-	-	-	-	-	-	-	0.90 (1)
free conformation ab1	0.90	-	-	-	-	-	-	-	-	0.90 (1)
ige spe-7	0.90	-	-	-	-	-	-	-	-	0.90 (1)
trivalent antibody fragment	0.90	-	-	-	-	-	-	-	-	0.90 (1)

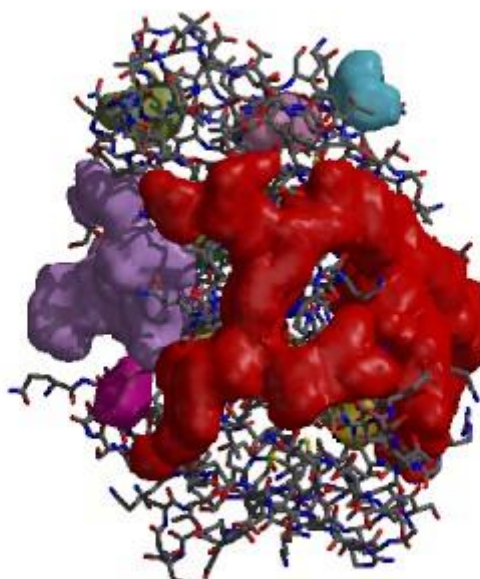
[illegible]

human igg1 kappa autoantibody	0.90	-	-	-	-	-	-	-	0.90 (1)
----------------------------------	------	---	---	---	---	---	---	---	----------

The score in red is a measure of how strongly the term is predicted from the hits obtained by the different methods. The scores in blue show each method's contribution to the total score (with the number of relevant sequences/structures shown in brackets in grey). Click on any score to see the list of hits from the given program and the source of the given term.

Binding sites

Binding site analysis



*Clefts and cavities in protein structure,
the colours corresponding to the entries in the table below.*

The following table shows the "gap" regions (clefts and cavities) in the protein's surface. The gaps are ordered by decreasing volume (in Å³), and the table includes various parameters that are indicative of binding sites. (Ignore data shown in grey).

Gap region	Volume	R1 ratio	Accessible vertices		Buried vertices		Ave. depth		Residue type*	Residue conservation**	Ligands
☒ 1	3825.56	2.80	65.37%	4	10.66%	2	13.70	1	211 874471 .		

☒ 2	1364.3 4	-	63.66 %	5	9.44%	3	9.75	2	336575 .		
☒ 3	410.48	-	61.17 %	6	3.21%	8	8.06	3	.11341 .		
☒ 4	210.94	-	97.70 %	1	13.03 %	1	0.00	9	.23263 .		
☐ 5	239.62	-	58.44 %	9	7.48%	4	5.71	4	.1224. .		
☐ 6	211.36	-	54.39 %	1 0	2.67%	9	5.47	5	112211 .		
☐ 7	137.53	-	59.18 %	8	4.08%	6	5.31	6	.14112 .		
☐ 8	125.30	-	66.67 %	3	4.30%	5	0.00	1 0	..3311 .		
☐ 9	119.39	-	59.36 %	7	1.17%	1 0	5.06	8	..3211 .		
☐ 10	128.67	-	69.52 %	2	3.21%	7	5.27	7	113... 1		

Total number of gap regions analysed: 10

Key to residue colour codes

* Residue-type colouring:						
Positive	Negative	Neutral	Aliphatic	Aromatic	Pro & Gly	Cysteine
H,K,R	D,E	S,T,N,Q	A,V,L,I,M	F,Y,W	P,G	C

** Colouring by residue-conservation score:									
blue	purple	skyblue	cyan	grey	green	yellow	orange	pink	Red
0.0-0.1	0.1-0.2	0.2-0.3	0.3-0.4	0.4-0.5	0.5-0.6	0.6-0.7	0.7-0.9	0.8-0.9	0.9-1.0

Reverse template search results for cn40

Below are the results reported by the template search program after scanning 400 auto-generated templates from the query structure against representative structures in the PDB. Click on the coloured ball or hit number at the left of each hit to see further details of the template match, together with a sequence alignment based on it.

The top 20 hits are listed below.

a. Certain matches (E-value < 1.00E-06)

Hit no.	E-value	Similarity score	Neighbours ident/simil [equiv]	Template id	Matched PDB entry	Longest fitted segment	Seq lengths query/target	Overlap	%-tage seq id	Structural similarity
 1	0.00E+00	940.00	48/0 [48]	TMP00336	2x7l ×53: Implications of the HIV-1 rev dimer structure at 3.2a resolution for multimeric binding to the rev response element	109/109	110/217	110	100.00%	99.1%
 2	0.00E+00	920.00	47/0 [47]	TMP00351	4rdq ×64: Calcium-activated chloride channel bestrophin-1, from chicken, in complex with fab antibody fragments, chloride and calcium	115/115	116/211	116	100.00%	99.1%
 3	0.00E+00	835.56	42/2 [44]	TMP00351	1oag ×53: Free conformation ab1 of the ige spe-7	115/115	116/120	120	81.03%	99.1%
 4	0.00E+00	835.56	42/2 [44]	TMP00351	1nqb ×35: Trivalent antibody fragment	114/114	116/232	119	80.17%	99.1%
 5	0.00E+00	834.31	41/4 [45]	TMP00351	2xzq ×56: Crystal structure analysis of the anti-(4-hydroxy-3-nitrophenyl)-acetyl murine germline monoclonal antibody bbe6.12h3 fab fragment in complex with a phage display derived dodecapeptide yqlrpnatlr	115/115	116/215	120	81.03%	99.1%
 6	0.00E+00	824.94	41/3 [44]	TMP00351	2y07 ×58: Crystal structure analysis of the anti-(4-hydroxy-3-nitrophenyl) -acetyl murine germline monoclonal antibody bbe6.12h3 fab fragment in complex with a phage display derived dodecapeptide ppypawhapgni	115/115	116/215	120	79.31%	99.1%
 7	0.00E+00	824.94	41/3 [44]	TMP00351	4a6y ×30: Crystal structure of fab fragment of anti-(4-hydroxy-3-nitrophenyl)-acetyl murine germline antibody bbe6.12h3	115/115	116/215	120	79.31%	99.1%
 8	0.00E+00	824.94	41/3 [44]	TMP00351	1a6v ×29: B1-8 fv fragment complexed with a (4-hydroxy-3-nitrophenyl) acetate	113/113	116/118	118	78.45%	99.1%

9	0.00E+00	821.88	39/6 [45]	TMP00251	compound 2c1p×63: Fab-fragment of enantioselective antibody complexed with finrozole	114/114	116/217	115	75.00%	99.1%
10	0.00E+00	798.00	39/4 [43]	TMP00071	4bh8×58: Crystal structure of germline antibody 36-65 in complex with peptide gdprpsyishll	115/115	116/218	121	78.45%	99.1%
11	0.00E+00	793.91	39/3 [42]	TMP00372	4jo4×62: Crystal structure of rabbit mab r20 fab	109/109	110/214	110	71.82%	99.1%
12	0.00E+00	788.62	39/3 [42]	TMP00071	1jfq×71: Antigen-binding fragment of the murine anti-phenylarsonate antibody 36-71, "fab 36-71"	114/114	116/220	120	76.72%	99.1%
13	0.00E+00	786.38	40/1 [41]	TMP00251	3qum×115: Crystal structure of human prostate specific antigen (psa) in fab sandwich with a high affinity and a pca selective antibody	115/115	116/219	117	76.72%	99.1%
14	0.00E+00	782.16	38/4 [42]	TMP00372	4hbc×58: Crystal structure of a conformation-dependent rabbit igg fab specific for amyloid prefibrillar oligomers	108/108	110/213	110	70.91%	99.1%
15	0.00E+00	781.22	38/4 [42]	TMP00372	4ht1×63: Human tweak in complex with the fab fragment of a neutralizing antibody	107/107	110/217	113	70.91%	99.1%
16	0.00E+00	780.19	36/8 [44]	TMP00336	4o4y×53: Crystal structure of the anti-hinge rabbit antibody 2095-2 in complex with ides hinge peptide	108/108	110/219	112	68.18%	99.1%
17	0.00E+00	776.94	37/6 [43]	TMP00251	3tt3×56: Crystal structure of leut in the inward-open conformation in complex with fab	115/115	116/218	118	73.28%	99.1%
18	0.00E+00	772.97	36/7 [43]	TMP00336	2r56×53: Crystal structure of a recombinant ige fab fragment in complex with bovine beta-lactoglobulin allergen	106/106	110/211	110	66.36%	99.1%
19	0.00E+00	768.50	37/5 [42]	TMP00071	1eo8×63: Influenza virus hemagglutinin complexed with a neutralizing antibody	115/115	116/217	118	64.66%	99.1%
20	0.00E+00	758.75	36/6 [42]	TMP00336	1vqe×55: Tr1.9 fab fragment of a human igg1 kappa autoantibody	106/106	110/214	110	66.36%	99.1%

Supplementary file 5

1. Ramachandran Plot statistics

	No. of residues	%-tage
	-----	-----
Most favoured regions [A,B,L]	172	90.5%
Additional allowed regions [a,b,l,p]	15	7.9%
Generously allowed regions [~a,~b,~l,~p]	2	1.1%
Disallowed regions [XX]	1	0.5%*
	----	-----
Non-glycine and non-proline residues	190	100.0%
End-residues (excl. Gly and Pro)	4	
Glycine residues	21	
Proline residues	11	

Total number of residues	226	

Based on an analysis of **118** structures of resolution of at least **2.0** Angstroms and *R*-factor no greater than **20.0** a good quality model would be expected to have over **90%** in the most favoured regions [A,B,L].

2. G-Factors

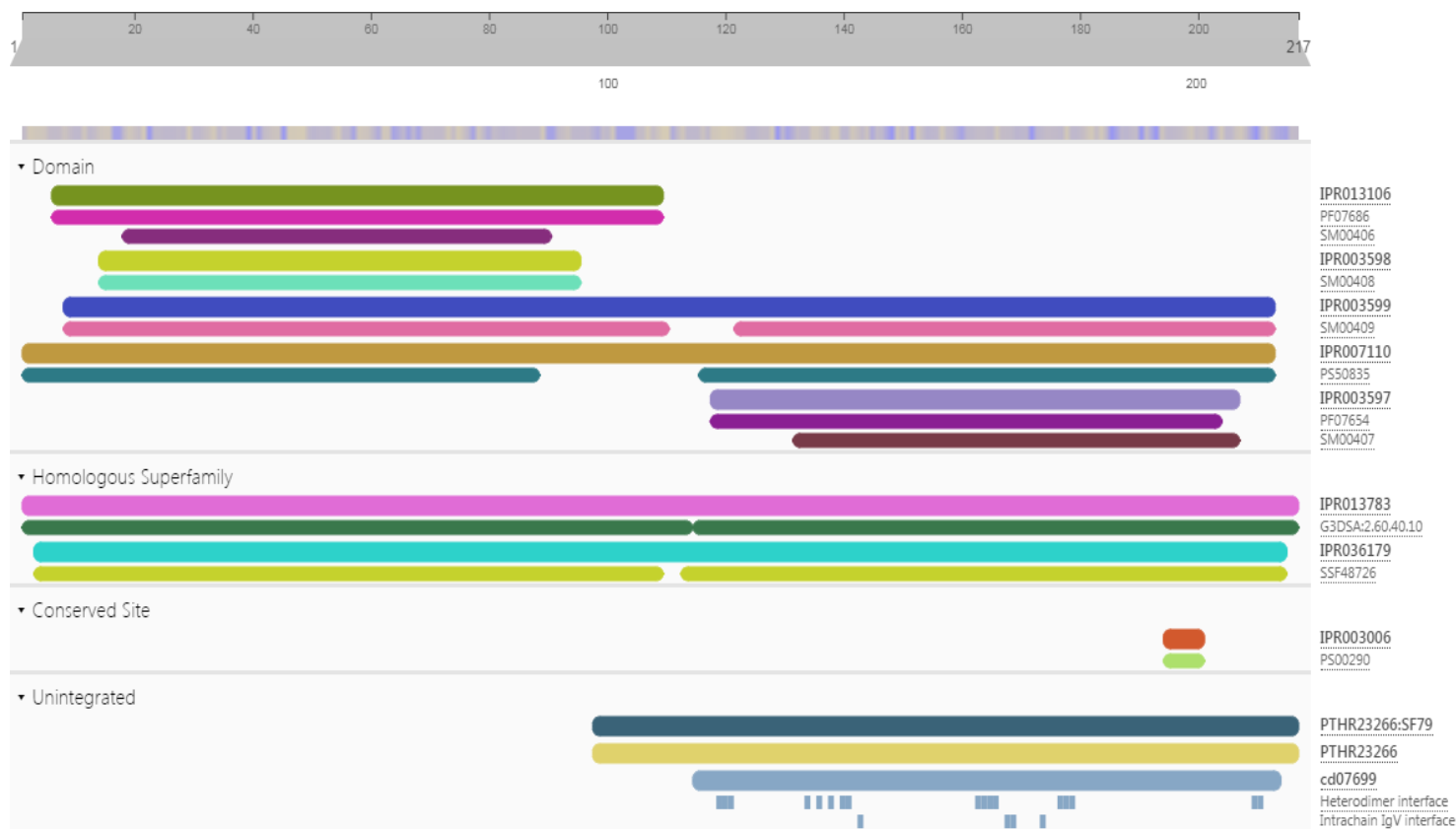
Parameter	Score	Average Score
-----	-----	-----
Dihedral angles:-		
Phi-psi distribution	-0.48	
Chi1-chi2 distribution	-0.15	
Chi1 only	0.03	
Chi3 & chi4	0.55	
Omega	0.16	
		-0.07
		=====
Main-chain covalent forces:-		
Main-chain bond lengths	0.55	
Main-chain bond angles	0.39	
		0.46
		=====
OVERALL AVERAGE		0.14
		=====

G-factors provide a measure of how **unusual**, or out-of-the-ordinary, a property is.

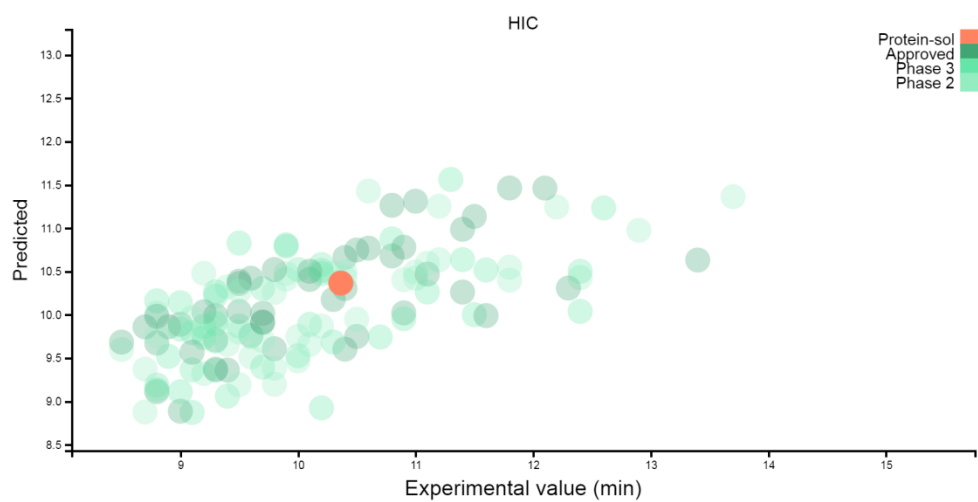
Values below -0.5* - unusual

Values below -1.0** - highly unusual

Important note: The main-chain bond-lengths and bond angles are compared with the Engh & Huber (1991) ideal values derived from small-molecule data. Therefore, structures refined using different restraints may show apparently large deviations from normality.



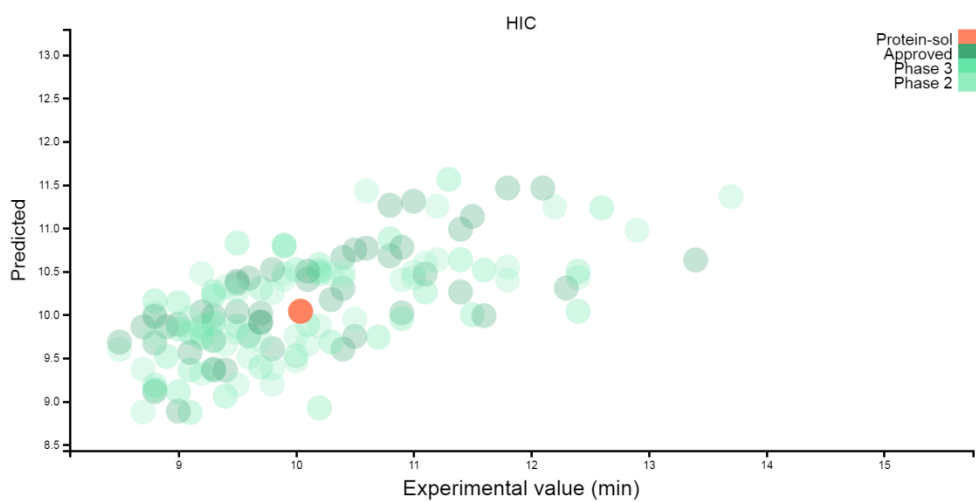
Heavy chain superfamily analysis



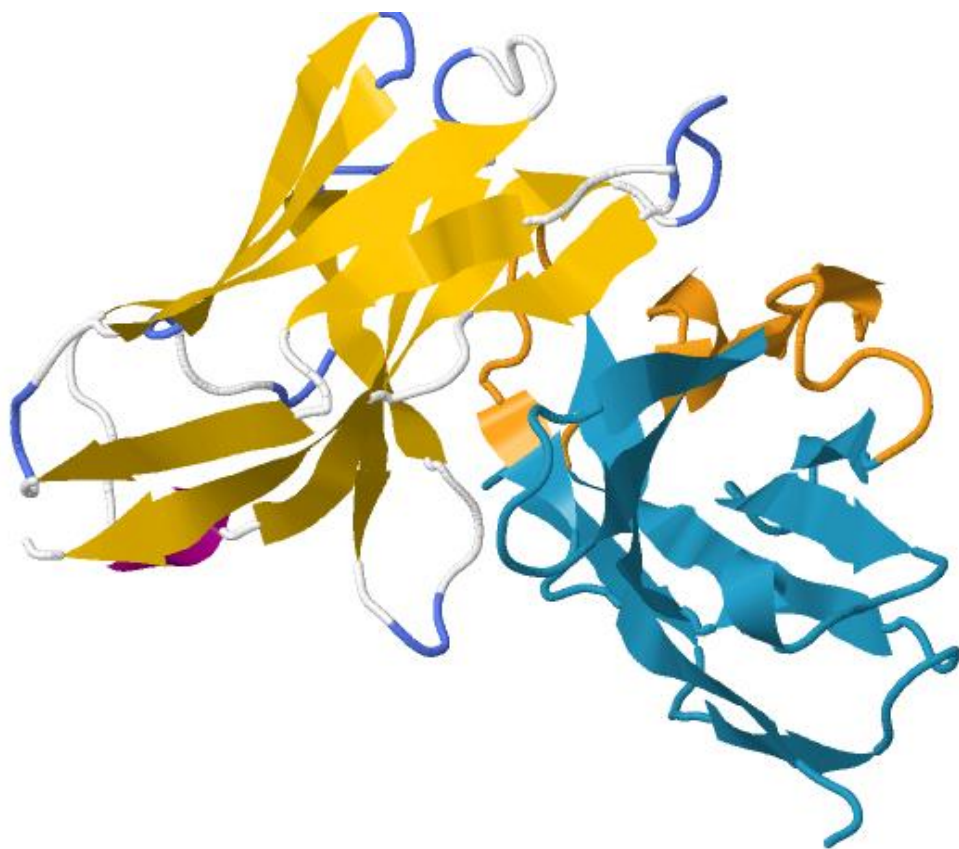
HIC PLOT FOR HEAVY CHAIN



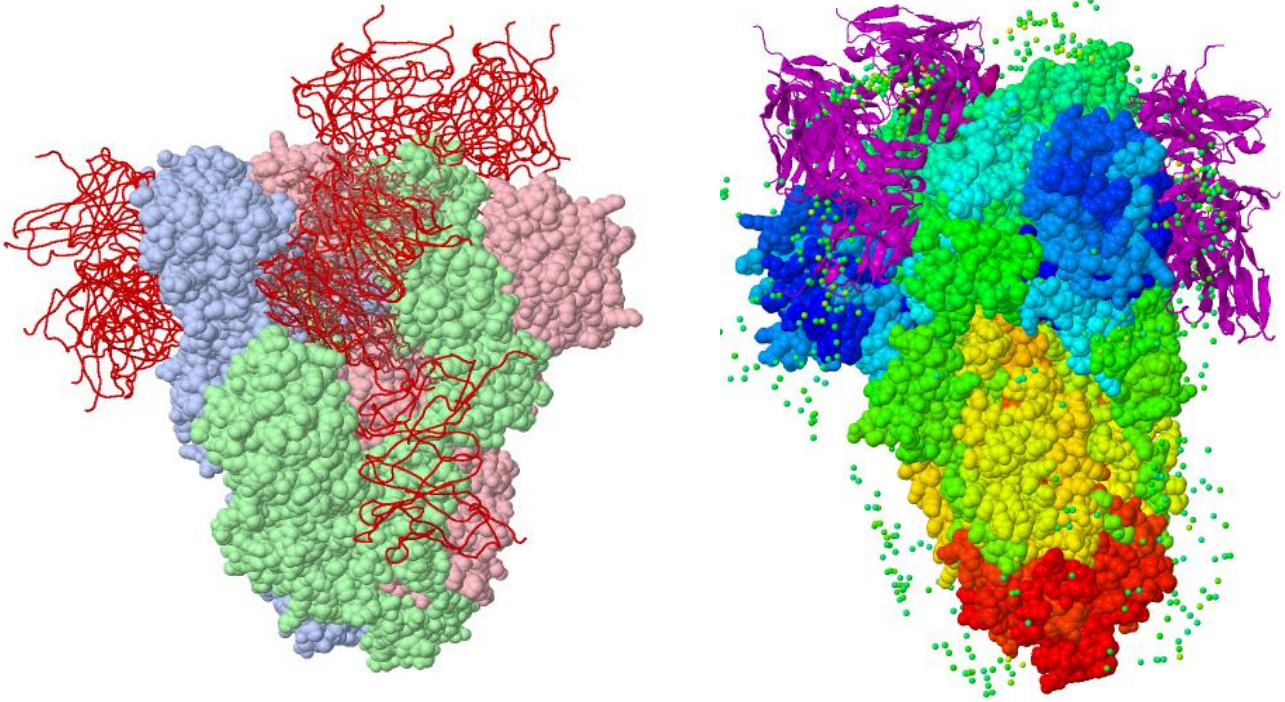
Light chain superfamily analysis



LIGHT CHAIN HIC PLOT



Final confirmed structure of FAB region of antibody



Polyspecific region binding of Fab- antibody to Spike glycoprotein of SARS-CoV-2

Closest reference gene and allele(s) from the IMGT V domain directory: *Homo sapiens* (human)

Species	Gene and allele	Domain	Domain label	Smith-Waterman score	% identity	Overlap
Homo sapiens	<u>IGHV1-46*01</u>	1	VH	464	66.3	98
Homo sapiens	<u>IGHV1-46*03</u>	1	VH	464	66.3	98
Homo sapiens	<u>IGHV1-46*02</u>	1	VH	459	65.3	98
Homo sapiens	<u>IGHV1-46*04</u>	1	VH	457	65.3	98
Homo sapiens	<u>IGHV1-3*01</u>	1	VH	443	65.3	98

Species	Gene and allele	Domain	Domain label	Smith-Waterman score	% identity	Overlap
Homo sapiens	<u>IGHJ4*01</u>	1		95	86.7	15
Homo sapiens	<u>IGHJ4*02</u>	1		95	86.7	15
Homo sapiens	<u>IGHJ4*03</u>	1		95	86.7	15

Heavy chain IMGT

Light chain IMGT

Closest reference gene and allele(s) from the IMGT V domain directory: *Homo sapiens* (human)

Species	Gene and allele	Domain	Domain label	Smith-Waterman score	% identity	Overlap
Homo sapiens	<u>IGKV1-5*01</u>	1	V-KAPPA	435	72.7	88
Homo sapiens	<u>IGKV1-5*02</u>	1	V-KAPPA	433	72.7	88
Homo sapiens	<u>IGKV1-12*01</u>	1	V-KAPPA	427	72.7	88
Homo sapiens	<u>IGKV1-12*02</u>	1	V-KAPPA	427	72.7	88
Homo sapiens	<u>IGKV1D-12*01</u>	1	V-KAPPA	427	72.7	88
Species	Gene and allele	Domain	Domain label	Smith-Waterman score	% identity	Overlap
Homo sapiens	<u>IGKJ4*01</u>	1		50	77.8	9
Homo sapiens	<u>IGKJ4*02</u>	1		50	77.8	9
Homo sapiens	<u>IGKJ2*01</u>	1		44	77.8	9
Homo sapiens	<u>IGKJ2*02</u>	1		44	77.8	9
Homo sapiens	<u>IGKJ5*01</u>	1		43	77.8	9

Here is a list of examples some FDA-approved monoclonal antibody drugs in market as examples

- abciximab (Reopro)
- adalimumab (Humira, Amjevita)
- alefacept (Amevive)
- alemtuzumab (Campath)
- basiliximab (Simulect)
- belimumab (Benlysta)
- bezlotoxumab (Zinplava)
- canakinumab (Ilaris)