Full wwPDB X-ray Structure Validation Report (i)

Nov 15, 2023 – 01:12 PM JST

PDB ID	:	6J5G
Title	:	Complex structure of MAb 4.2-scFv with tick-borne encephalitis virus envelope protein
Authors	:	Yang, X.; Qi, J.; Peng, R.; Dai, L.; Gould, E.A.; Tien, P.; Gao, G.F.
Deposited on	:	2019-01-10
Resolution	:	3.29 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

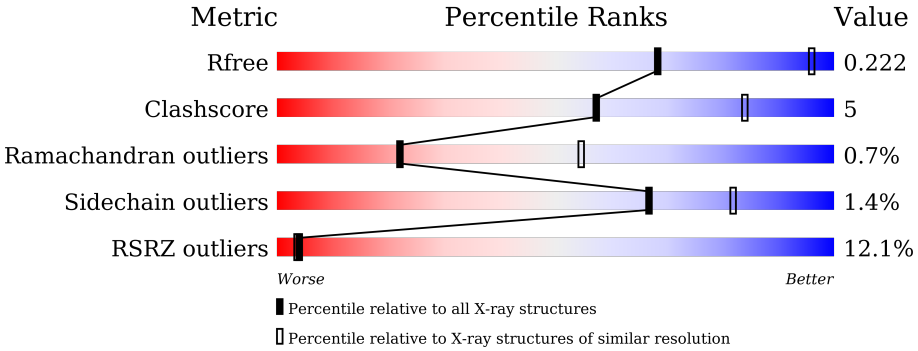
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.29 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	407	18%; 88%; • • 8%
2	H	120	95%; 5%
3	L	109	84%; 12%; • •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4579 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Envelope protein E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	373	2819	1765	501	533	20	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	402	HIS	-	expression tag	UNP P14336
A	403	HIS	-	expression tag	UNP P14336
A	404	HIS	-	expression tag	UNP P14336
A	405	HIS	-	expression tag	UNP P14336
A	406	HIS	-	expression tag	UNP P14336
A	407	HIS	-	expression tag	UNP P14336

- Molecule 2 is a protein called antibody heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	H	120	926	586	147	189	4	0	0	0

- Molecule 3 is a protein called antibody light chain.

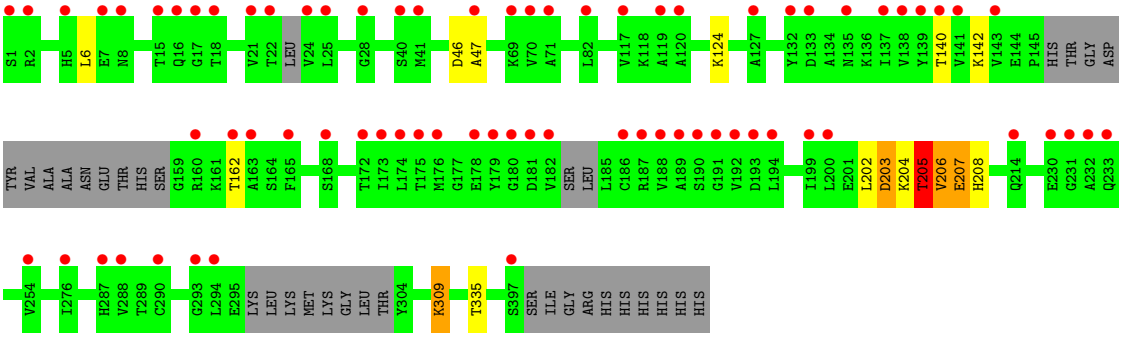
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	L	108	834	530	139	163	2	0	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

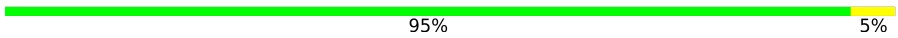
- Molecule 1: Envelope protein E

Chain A:



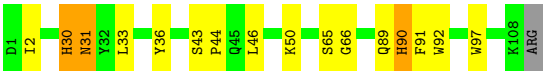
- Molecule 2: antibody heavy chain

Chain H:



- Molecule 3: antibody light chain

Chain L:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	135.73Å 100.69Å 140.24Å 90.00° 100.25° 90.00°	Depositor
Resolution (Å)	46.00 – 3.29 46.00 – 3.29	Depositor EDS
% Data completeness (in resolution range)	80.1 (46.00-3.29) 80.2 (46.00-3.29)	Depositor EDS
R_{merge}	0.26	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.89 (at 3.32Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.208 , 0.222 0.207 , 0.222	Depositor DCC
R_{free} test set	1136 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	65.8	Xtriage
Anisotropy	0.096	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 51.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	4579	wwPDB-VP
Average B, all atoms (Å^2)	102.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.10% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| >5$	RMSZ	$\#\|Z\| >5$
1	A	0.69	0/2879	0.65	0/3906
2	H	0.75	0/947	0.77	0/1284
3	L	0.83	0/857	0.75	0/1165
All	All	0.73	0/4683	0.69	0/6355

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2819	0	2718	18	0
2	H	926	0	883	6	0
3	L	834	0	802	24	0
All	All	4579	0	4403	44	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

All (44) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:106:ALA:HA	3:L:91:PHE:CE1	2.08	0.88

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:L:2:ILE:HG21	3:L:90:HIS:CD2	2.11	0.85
1:A:142:LYS:HE2	1:A:162:THR:OG1	1.78	0.83
1:A:124:LYS:HA	1:A:203:ASP:CG	2.00	0.82
2:H:106:ALA:HA	3:L:91:PHE:HE1	1.43	0.81
1:A:124:LYS:HA	1:A:203:ASP:OD2	1.85	0.75
3:L:90:HIS:ND1	3:L:90:HIS:O	2.21	0.74
3:L:2:ILE:HB	3:L:90:HIS:NE2	2.07	0.69
1:A:206:VAL:O	1:A:208:HIS:N	2.26	0.69
3:L:36:TYR:CE2	3:L:46:LEU:HD13	2.28	0.68
3:L:33:LEU:C	3:L:33:LEU:HD23	2.14	0.68
1:A:205:THR:O	1:A:207:GLU:N	2.27	0.67
1:A:204:LYS:O	1:A:205:THR:HG23	1.96	0.66
1:A:202:LEU:HD23	1:A:202:LEU:N	2.13	0.63
3:L:31:ASN:HD22	3:L:31:ASN:N	1.97	0.62
3:L:2:ILE:CB	3:L:90:HIS:NE2	2.65	0.59
3:L:2:ILE:CG2	3:L:90:HIS:CD2	2.85	0.58
3:L:33:LEU:HD23	3:L:33:LEU:O	2.05	0.56
1:A:335:THR:O	1:A:335:THR:HG23	2.06	0.55
2:H:106:ALA:CA	3:L:91:PHE:CE1	2.85	0.54
3:L:30:HIS:O	3:L:31:ASN:HB2	2.10	0.52
2:H:105:ILE:C	3:L:91:PHE:CD1	2.83	0.52
1:A:124:LYS:O	1:A:203:ASP:HB2	2.09	0.52
3:L:31:ASN:N	3:L:31:ASN:ND2	2.60	0.49
3:L:90:HIS:ND1	3:L:90:HIS:C	2.68	0.47
1:A:47:ALA:HB3	1:A:140:THR:HG22	1.97	0.47
3:L:90:HIS:O	3:L:90:HIS:CG	2.67	0.46
3:L:89:GLN:HG2	3:L:90:HIS:N	2.31	0.46
1:A:46:ASP:HB3	1:A:140:THR:HG23	1.97	0.46
1:A:6:LEU:HD23	1:A:6:LEU:N	2.30	0.46
3:L:90:HIS:HD1	3:L:92:TRP:H	1.65	0.45
1:A:140:THR:HG23	1:A:140:THR:O	2.16	0.45
1:A:47:ALA:HB3	1:A:140:THR:CG2	2.47	0.45
3:L:30:HIS:C	3:L:31:ASN:HD22	2.20	0.44
1:A:124:LYS:HD3	1:A:203:ASP:OD2	2.18	0.44
1:A:202:LEU:N	1:A:202:LEU:CD2	2.82	0.42
1:A:309:LYS:H	1:A:309:LYS:HG2	1.69	0.42
3:L:91:PHE:HA	3:L:97:TRP:CE3	2.55	0.42
3:L:33:LEU:C	3:L:33:LEU:CD2	2.85	0.42
2:H:51:ILE:O	2:H:51:ILE:HG23	2.21	0.41
3:L:43:SER:HA	3:L:44:PRO:HD3	1.97	0.41
3:L:65:SER:OG	3:L:66:GLY:N	2.53	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:12:VAL:HG21	2:H:86:LEU:HD12	2.03	0.40
1:A:205:THR:O	1:A:206:VAL:C	2.59	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	363/407 (89%)	338 (93%)	22 (6%)	3 (1%)	19	51
2	H	118/120 (98%)	111 (94%)	7 (6%)	0	100	100
3	L	106/109 (97%)	97 (92%)	8 (8%)	1 (1%)	17	48
All	All	587/636 (92%)	546 (93%)	37 (6%)	4 (1%)	22	54

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	205	THR
1	A	206	VAL
1	A	207	GLU
3	L	30	HIS

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	296/337 (88%)	293 (99%)	3 (1%)	76	86
2	H	99/99 (100%)	98 (99%)	1 (1%)	76	86
3	L	91/92 (99%)	88 (97%)	3 (3%)	38	66
All	All	486/528 (92%)	479 (99%)	7 (1%)	67	82

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	203	ASP
1	A	205	THR
1	A	309	LYS
2	H	103	TYR
3	L	31	ASN
3	L	50	LYS
3	L	90	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
3	L	30	HIS
3	L	31	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	373/407 (91%)	0.85	73 (19%)	1	1	30, 138, 230, 282	0
2	H	120/120 (100%)	-0.34	0	100	100	26, 45, 74, 92	0
3	L	108/109 (99%)	-0.30	0	100	100	27, 43, 84, 113	0
All	All	601/636 (94%)	0.41	73 (12%)	4	3	26, 83, 221, 282	0

All (73) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	189	ALA	5.8
1	A	16	GLN	5.6
1	A	140	THR	5.4
1	A	5	HIS	5.3
1	A	15	THR	5.1
1	A	288	VAL	5.0
1	A	186	CYS	4.9
1	A	174	LEU	4.9
1	A	290	CYS	4.9
1	A	24	VAL	4.7
1	A	163	ALA	4.7
1	A	1	SER	4.6
1	A	187	ARG	4.5
1	A	162	THR	4.5
1	A	192	VAL	4.5
1	A	232	ALA	4.4
1	A	25	LEU	4.4
1	A	22	THR	4.4
1	A	18	THR	4.4
1	A	175	THR	4.3
1	A	141	VAL	4.2
1	A	119	ALA	4.2
1	A	179	TYR	4.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	173	ILE	4.2
1	A	40	SER	4.1
1	A	139	TYR	4.1
1	A	47	ALA	3.9
1	A	176	MET	3.8
1	A	287	HIS	3.7
1	A	137	ILE	3.6
1	A	28	GLY	3.5
1	A	294	LEU	3.5
1	A	188	VAL	3.4
1	A	172	THR	3.3
1	A	230	GLU	3.3
1	A	8	ASN	3.2
1	A	182	VAL	3.2
1	A	214	GLN	3.1
1	A	254	VAL	3.1
1	A	200	LEU	3.0
1	A	276	ILE	3.0
1	A	190	SER	3.0
1	A	233	GLN	3.0
1	A	194	LEU	2.9
1	A	181	ASP	2.9
1	A	138	VAL	2.8
1	A	7	GLU	2.7
1	A	160	ARG	2.7
1	A	127	ALA	2.6
1	A	82	LEU	2.6
1	A	21	VAL	2.5
1	A	133	ASP	2.5
1	A	199	ILE	2.4
1	A	71	ALA	2.4
1	A	293	GLY	2.4
1	A	397	SER	2.4
1	A	193	ASP	2.4
1	A	231	GLY	2.4
1	A	165	PHE	2.3
1	A	135	ASN	2.3
1	A	132	TYR	2.3
1	A	70	VAL	2.2
1	A	69	LYS	2.2
1	A	180	GLY	2.2
1	A	120	ALA	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	143	VAL	2.1
1	A	41	MET	2.1
1	A	191	GLY	2.1
1	A	178	GLU	2.1
1	A	168	SER	2.1
1	A	2	ARG	2.1
1	A	117	VAL	2.0
1	A	17	GLY	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.