



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 19, 2016 – 04:15 PM EST

PDB ID : 5GZN
Title : Structure of neutralizing antibody bound to Zika envelope protein
Authors : Wang, Q.; Yang, H.; Liu, X.; Dai, L.; Ma, T.; Qi, J.; Wong, G.; Peng, R.; Liu, S.; Li, J.; Li, S.; Song, J.; Liu, J.; He, J.; Yuan, H.; Xiong, Y.; Liao, Y.; Li, J.; Yang, J.; Tong, Z.; Griffin, B.; Bi, Y.; Liang, M.; Xu, X.; Cheng, G.; Wang, P.; Qiu, X.; Kobinger, G.; Shi, Y.; Yan, J.; Gao, G.F.
Deposited on : 2016-09-29
Resolution : 3.00 Å(reported)

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

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20028442
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20028442

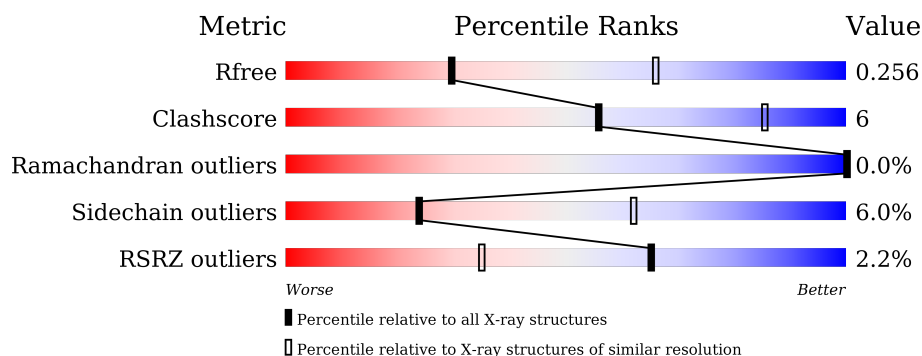
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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}	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$. 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	409	<div> <div>82%</div> <div>14%</div> <div>..</div> </div>
1	B	409	<div> <div>81%</div> <div>16%</div> <div>..</div> </div>
1	E	409	<div> <div>4%</div> <div>34%</div> <div>10%</div> <div>55%</div> </div>
1	G	409	<div> <div>4%</div> <div>35%</div> <div>8%</div> <div>55%</div> </div>
2	C	228	<div> <div>%</div> <div>78%</div> <div>19%</div> <div>..</div> </div>
2	H	228	<div> <div>%</div> <div>79%</div> <div>19%</div> <div>.</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	216	<div><div>%</div><div><div></div></div><div>81%17%..</div></div>
3	L	216	<div><div></div><div>87%11%..</div></div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 15537 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 Genome polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	402	Total	C	N	O	S	0	0	0
			3071	1915	537	594	25			
1	B	401	Total	C	N	O	S	0	0	0
			3061	1909	534	593	25			
1	E	184	Total	C	N	O	S	0	0	0
			1431	895	250	274	12			
1	G	183	Total	C	N	O	S	0	0	0
			1429	896	249	272	12			

- Molecule 2 is a protein called Antibody Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	224	Total	C	N	O	S	0	0	0
			1692	1070	286	330	6			
2	C	223	Total	C	N	O	S	0	0	0
			1686	1067	285	328	6			

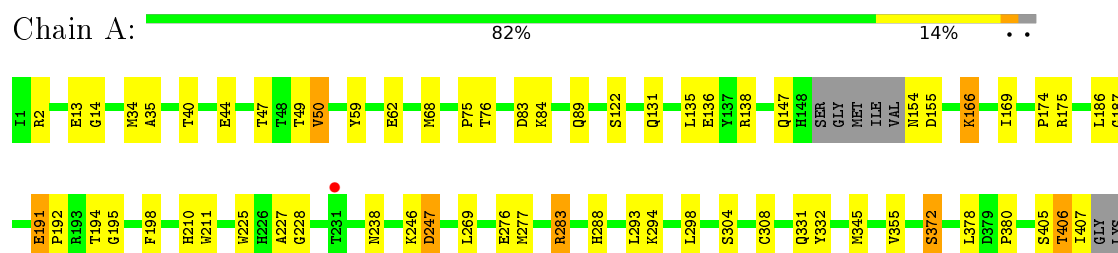
- Molecule 3 is a protein called Antibody light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	213	Total	C	N	O	S	0	0	0
			1579	991	262	322	4			
3	D	214	Total	C	N	O	S	0	0	0
			1588	996	263	325	4			

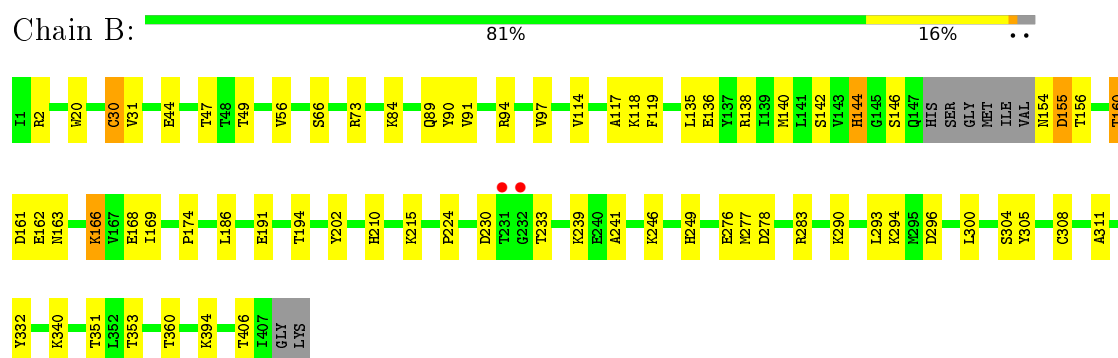
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

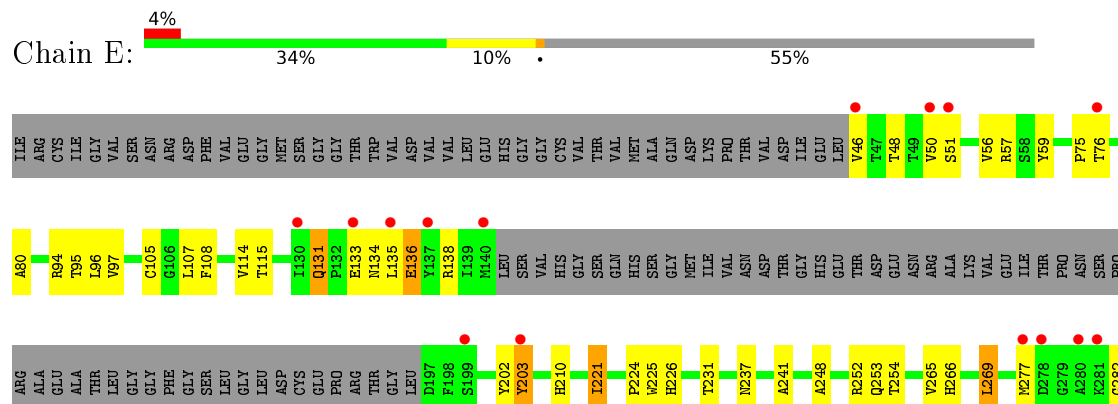
• Molecule 1: Genome polyprotein

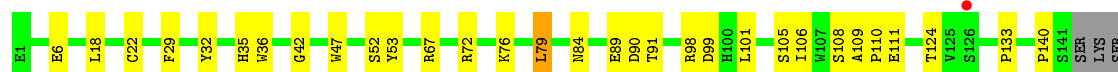


• Molecule 1: Genome polyprotein



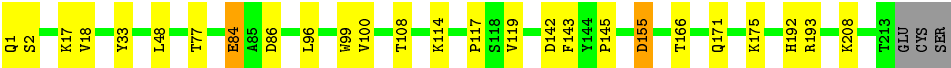
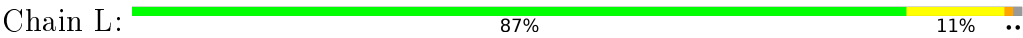
• Molecule 1: Genome polyprotein



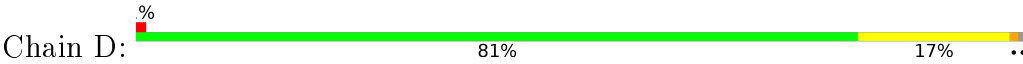




● Molecule 3: Antibody light chain



● Molecule 3: Antibody light chain



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	124.71Å 129.38Å 428.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.08 – 3.00 42.87 – 2.99	Depositor EDS
% Data completeness (in resolution range)	98.3 (42.08-3.00) 98.2 (42.87-2.99)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.28 (at 3.01Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.223 , 0.258 0.224 , 0.256	Depositor DCC
R_{free} test set	3480 reflections (5.04%)	DCC
Wilson B-factor (Å ²)	88.7	Xtriage
Anisotropy	0.325	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 53.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.052 for -k,-h,-l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	15537	wwPDB-VP
Average B, all atoms (Å ²)	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.90% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.24	0/3135	0.46	0/4247
1	B	0.26	0/3124	0.58	2/4232 (0.0%)
1	E	0.24	0/1463	0.45	0/1977
1	G	0.23	0/1460	0.45	0/1971
2	C	0.24	0/1730	0.49	1/2359 (0.0%)
2	H	0.26	0/1736	0.45	0/2367
3	D	0.24	0/1628	0.45	0/2226
3	L	0.25	0/1619	0.47	0/2214
All	All	0.25	0/15895	0.49	3/21593 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	146	SER	N-CA-CB	-21.96	77.56	110.50
2	C	160	PHE	C-N-CD	-7.01	105.17	120.60
1	B	144	HIS	CB-CA-C	-5.35	99.69	110.40

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	3071	0	2996	32	0
1	B	3061	0	2989	37	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	1431	0	1370	24	0
1	G	1429	0	1372	20	0
2	C	1686	0	1632	30	0
2	H	1692	0	1637	20	0
3	D	1588	0	1546	19	0
3	L	1579	0	1540	18	0
All	All	15537	0	15082	190	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 6.

The worst 5 of 190 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)
3:L:155:ASP:OD2	3:L:192:HIS:HB3	1.67	0.93
3:L:155:ASP:OD2	3:L:192:HIS:CB	2.27	0.82
2:C:160:PHE:HD2	2:C:161:PRO:HD3	1.47	0.80
2:C:160:PHE:CD2	2:C:161:PRO:HD3	2.18	0.78
3:L:155:ASP:CG	3:L:192:HIS:HB3	2.04	0.76

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	398/409 (97%)	366 (92%)	32 (8%)	0	100	100
1	B	397/409 (97%)	363 (91%)	34 (9%)	0	100	100
1	E	180/409 (44%)	168 (93%)	12 (7%)	0	100	100
1	G	177/409 (43%)	166 (94%)	11 (6%)	0	100	100
2	C	219/228 (96%)	192 (88%)	26 (12%)	1 (0%)	34	76

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	H	220/228 (96%)	194 (88%)	26 (12%)	0	100	100
3	D	212/216 (98%)	200 (94%)	12 (6%)	0	100	100
3	L	211/216 (98%)	200 (95%)	11 (5%)	0	100	100
All	All	2014/2524 (80%)	1849 (92%)	164 (8%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	161	PRO

5.3.2 Protein sidechains ⓘ

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	336/341 (98%)	314 (94%)	22 (6%)	21	58
1	B	335/341 (98%)	317 (95%)	18 (5%)	27	66
1	E	152/341 (45%)	140 (92%)	12 (8%)	15	48
1	G	152/341 (45%)	141 (93%)	11 (7%)	18	53
2	C	187/192 (97%)	179 (96%)	8 (4%)	35	75
2	H	188/192 (98%)	177 (94%)	11 (6%)	24	63
3	D	180/182 (99%)	169 (94%)	11 (6%)	23	61
3	L	179/182 (98%)	169 (94%)	10 (6%)	26	65
All	All	1709/2112 (81%)	1606 (94%)	103 (6%)	24	62

5 of 103 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	73	ARG
1	B	276	GLU
1	G	115	THR
1	B	84	LYS
1	B	160	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	158	HIS
1	E	226	HIS
1	E	249	HIS
1	G	253	GLN

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 carbohydrates 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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	402/409 (98%)	-0.17	1 (0%) 95 87	52, 84, 131, 224	0
1	B	401/409 (98%)	-0.10	2 (0%) 91 76	54, 85, 135, 203	0
1	E	184/409 (44%)	0.42	15 (8%) 14 5	76, 114, 186, 242	0
1	G	183/409 (44%)	0.29	18 (9%) 10 4	74, 108, 182, 219	0
2	C	223/228 (97%)	-0.06	3 (1%) 79 53	56, 97, 147, 179	0
2	H	224/228 (98%)	-0.11	3 (1%) 79 53	60, 98, 142, 192	0
3	D	214/216 (99%)	-0.14	3 (1%) 78 51	52, 79, 163, 202	0
3	L	213/216 (98%)	-0.17	0 100 100	48, 73, 144, 212	0
All	All	2044/2524 (80%)	-0.04	45 (2%) 65 35	48, 91, 154, 242	0

The worst 5 of 45 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	137	TYR	7.2
1	E	135	LEU	5.9
1	G	137	TYR	5.4
2	C	146	SER	5.3
1	G	135	LEU	4.3

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

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

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.