wwPDB X-ray Structure Validation Summary Report (i)

Aug 23, 2023 – 02:06 PM EDT

PDB ID	:	8EED
Title	:	Crystal structure of a NHP anti-ZIKV neutralizing antibody rhMZ107-B in complex with ZIKV E glycoprotein
Authors	:	Sankhala, R.S.; Joyce, M.G.
Deposited on	:	2022-09-07
Resolution	:	3.49 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
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:

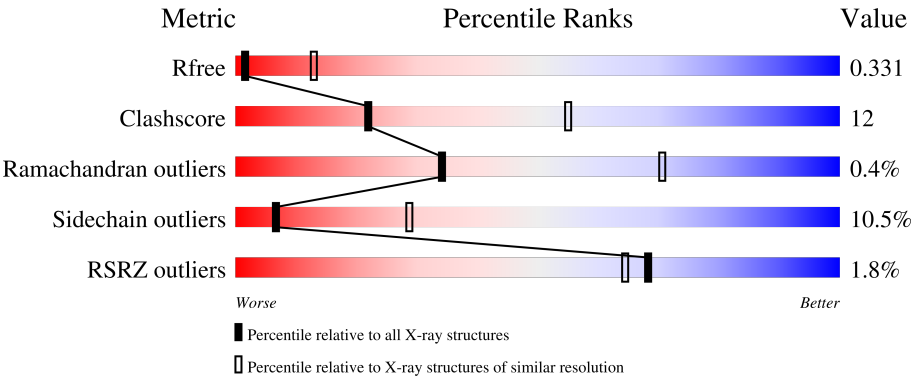
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 3.49 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	1379 (3.56-3.40)
Clashscore	141614	1461 (3.56-3.40)
Ramachandran outliers	138981	1424 (3.56-3.40)
Sidechain outliers	138945	1425 (3.56-3.40)
RSRZ outliers	127900	1289 (3.56-3.40)

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	405	70% 24%
1	B	405	67% 28%
1	C	405	64% 29%
1	D	405	61% 27% 5% 6%
2	E	226	4% 38% 15% 46%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	226	2% 39% 12% 44%
2	I	226	5% 34% 18% 45%
2	M	226	33% 17% 46%
3	F	220	3% 41% 11% 48%
3	J	220	40% 12% 48%
3	L	220	34% 18% 47%
3	N	220	% 41% 11% 47%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 19188 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
1	A	400	Total	C	N	O	S	0	0	0
			3030	1896	524	584	26			
1	C	391	Total	C	N	O	S	0	0	0
			2969	1857	517	570	25			
1	D	380	Total	C	N	O	S	0	0	0
			2841	1774	492	551	24			
1	B	403	Total	C	N	O	S	0	0	0
			3069	1916	534	593	26			

- Molecule 2 is a protein called rhMZ107-B antibody heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	126	Total	C	N	O	S	0	0	0
			962	612	157	191	2			
2	M	122	Total	C	N	O	S	0	0	0
			939	600	153	184	2			
2	E	123	Total	C	N	O	S	0	0	0
			945	603	154	186	2			
2	I	125	Total	C	N	O	S	0	0	0
			956	609	156	189	2			

- Molecule 3 is a protein called rhMZ107-B antibody light chain.

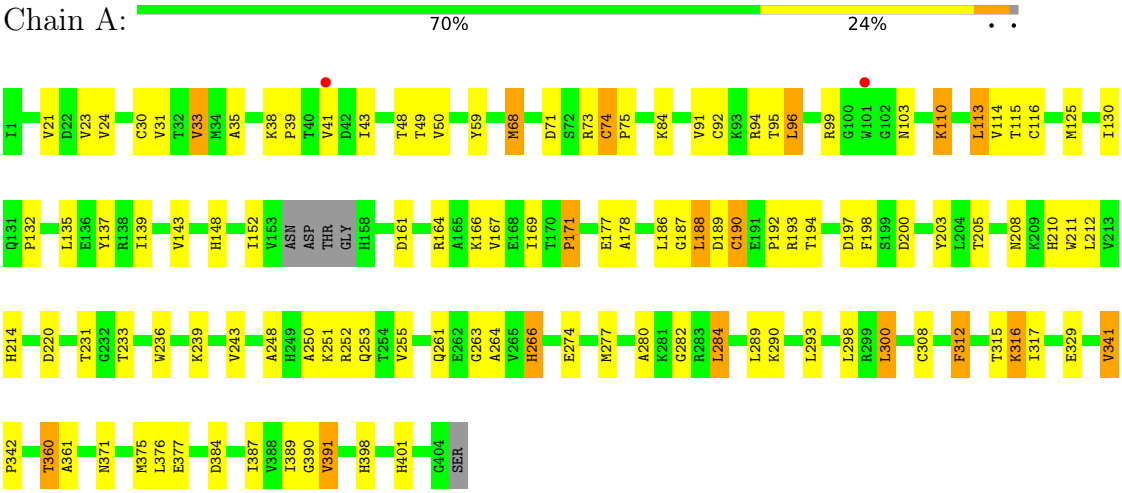
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	117	Total	C	N	O	S	0	0	0
			878	551	143	181	3			
3	N	116	Total	C	N	O	S	0	0	0
			869	546	141	179	3			
3	F	115	Total	C	N	O	S	0	0	0
			865	544	140	178	3			
3	J	115	Total	C	N	O	S	0	0	0
			865	544	140	178	3			

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.

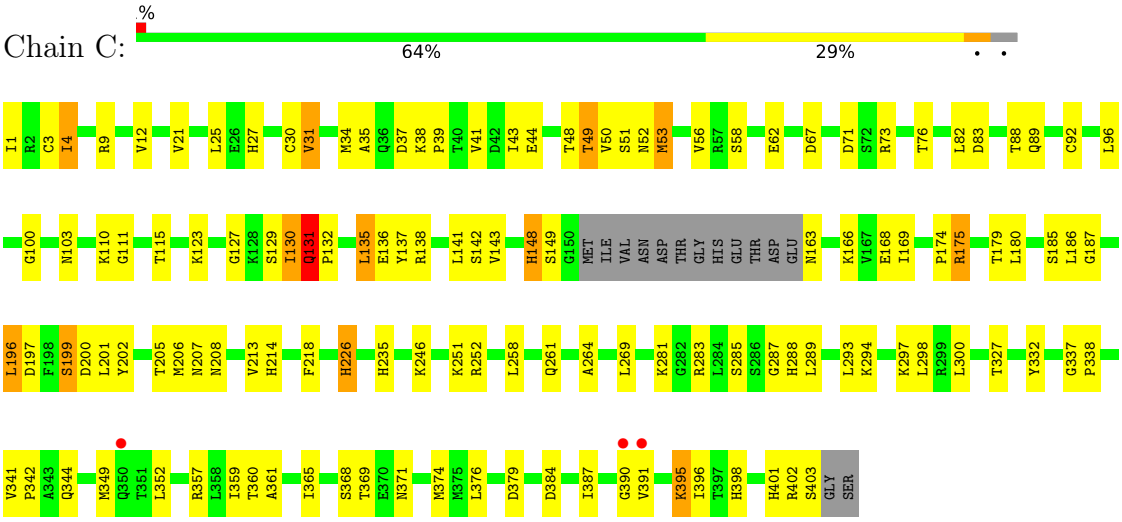
• Molecule 1: Envelope protein E

Chain A:



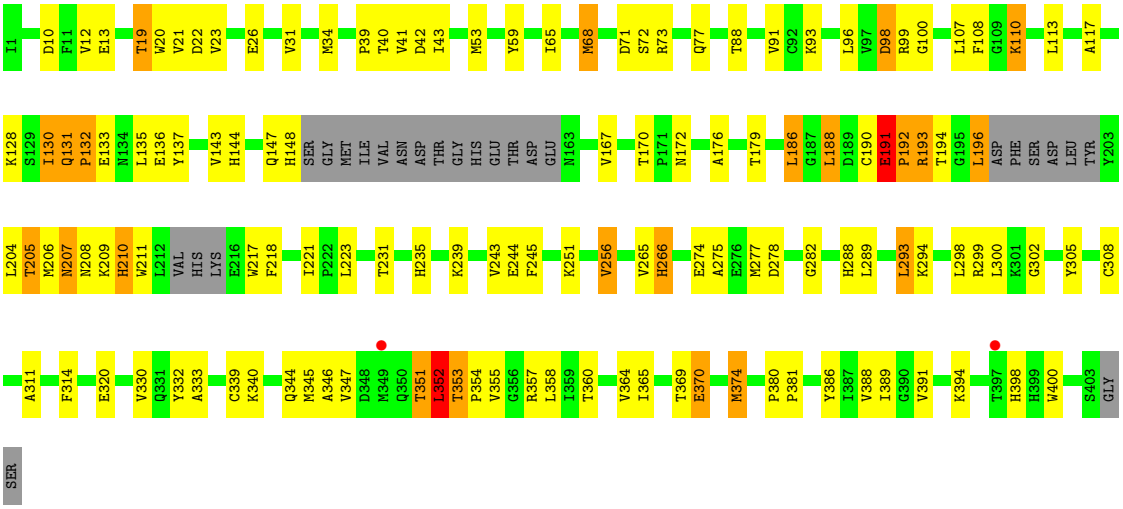
• Molecule 1: Envelope protein E

Chain C:



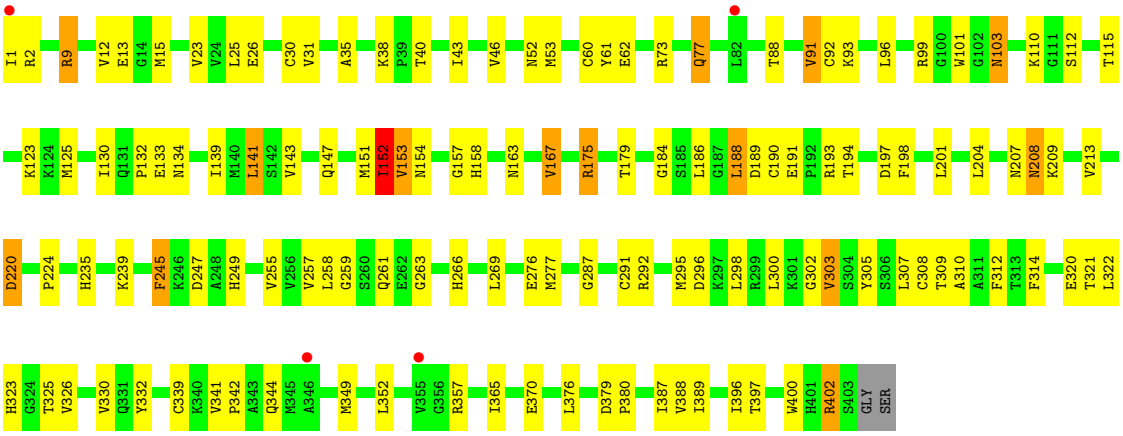
• Molecule 1: Envelope protein E

Chain D:



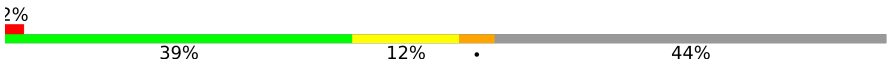
• Molecule 1: Envelope protein E

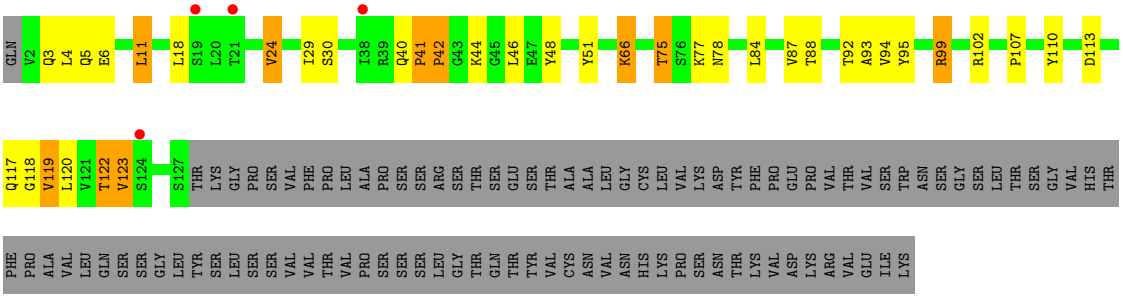
Chain B:



• Molecule 2: rhMZ107-B antibody heavy chain

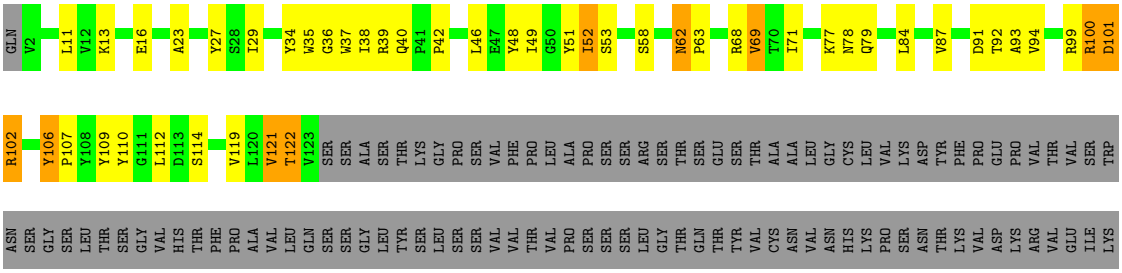
Chain H:





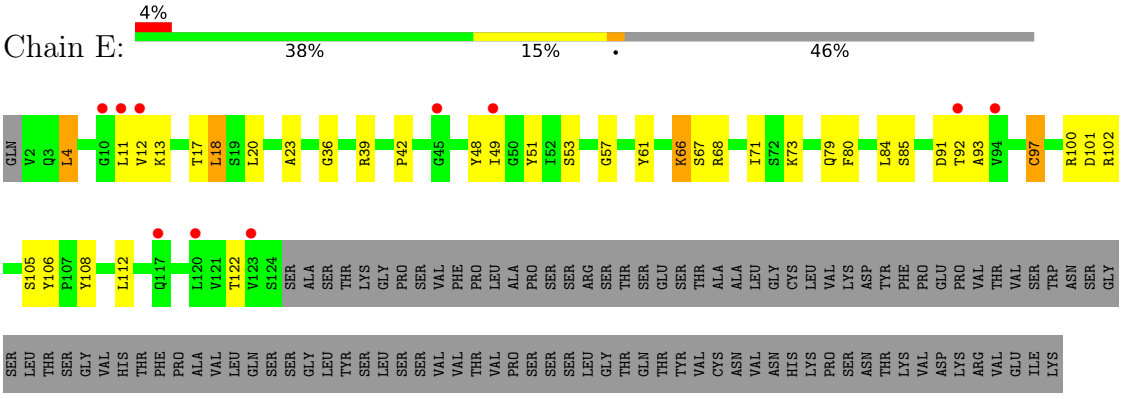
• Molecule 2: rhMZ107-B antibody heavy chain

Chain M:



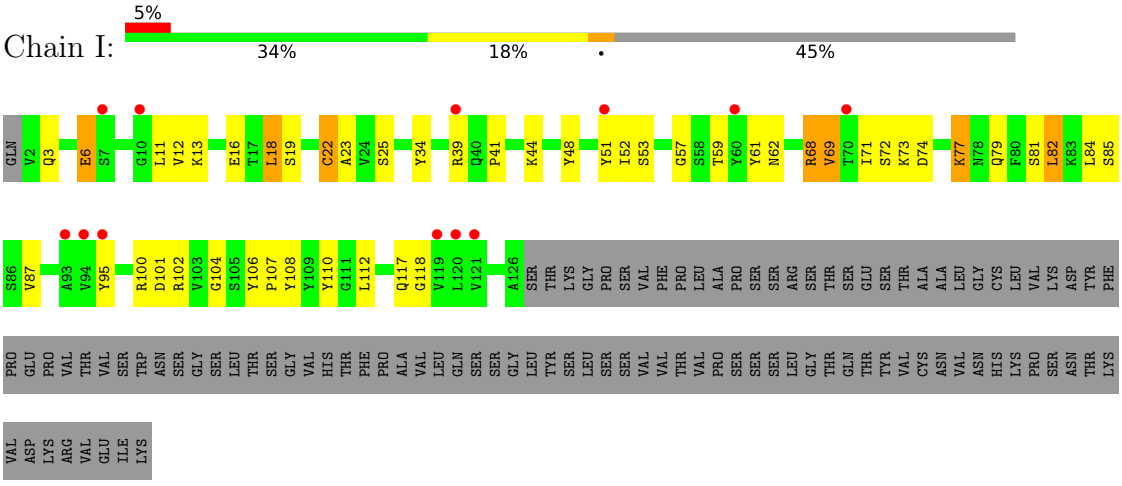
• Molecule 2: rhMZ107-B antibody heavy chain

Chain E:



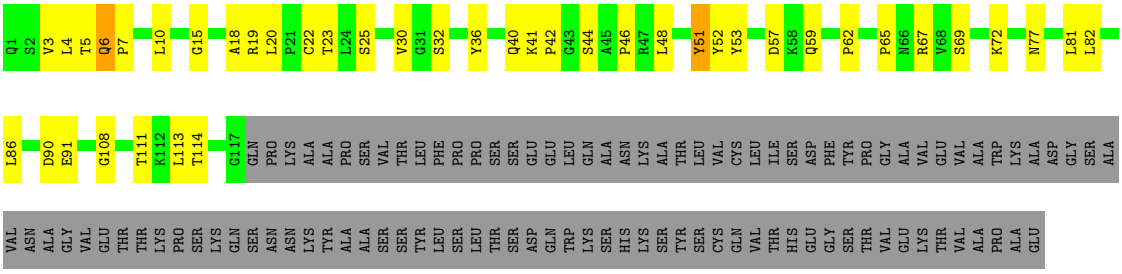
• Molecule 2: rhMZ107-B antibody heavy chain

Chain I:



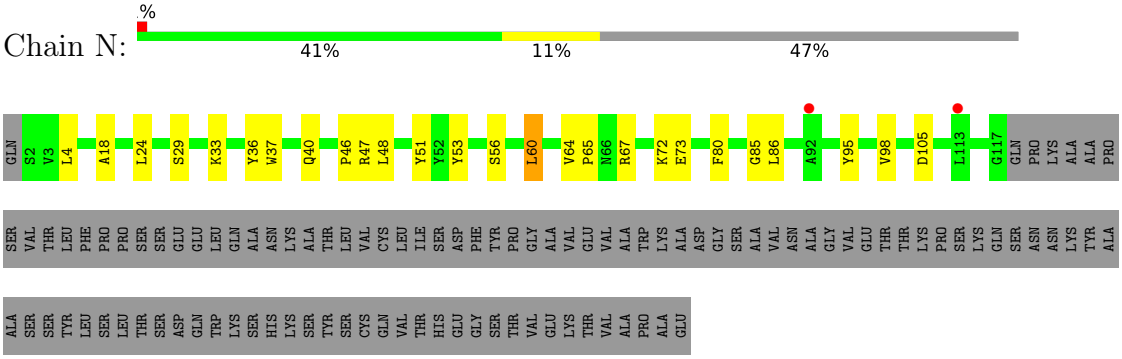
• Molecule 3: rhMZ107-B antibody light chain

Chain L:



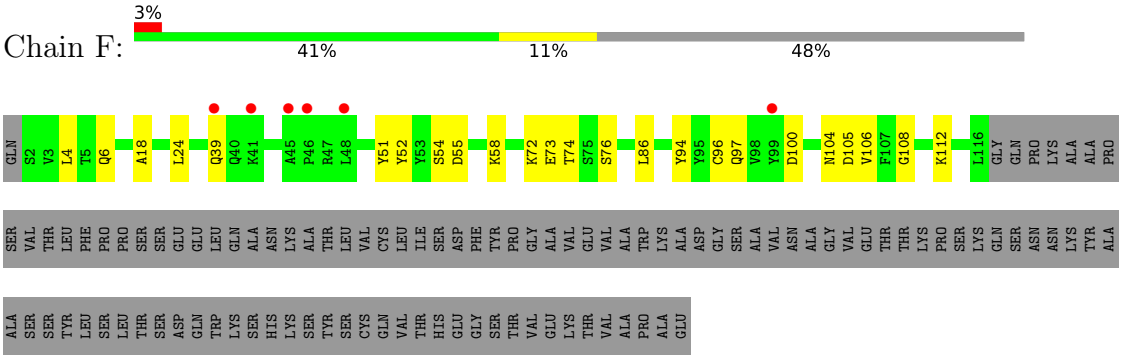
• Molecule 3: rhMZ107-B antibody light chain

Chain N:



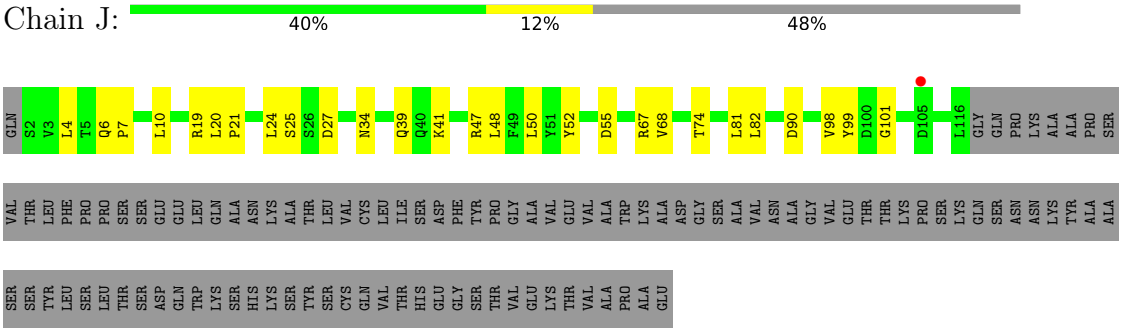
• Molecule 3: rhMZ107-B antibody light chain

Chain F:



• Molecule 3: rhMZ107-B antibody light chain

Chain J:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	92.03Å 105.63Å 132.38Å 82.52° 70.48° 81.04°	Depositor
Resolution (Å)	48.84 – 3.49 48.84 – 3.49	Depositor EDS
% Data completeness (in resolution range)	80.6 (48.84-3.49) 66.3 (48.84-3.49)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.99 (at 3.48Å)	Xtriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.289 , 0.331 0.289 , 0.331	Depositor DCC
R_{free} test set	1996 reflections (4.20%)	wwPDB-VP
Wilson B-factor (Å^2)	51.5	Xtriage
Anisotropy	0.193	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , -10.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.76	EDS
Total number of atoms	19188	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.18% 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

5.1 Standard geometry

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	0.25	0/3094	0.51	2/4194 (0.0%)
1	B	0.24	0/3134	0.49	0/4248
1	C	0.29	1/3032 (0.0%)	0.50	0/4110
1	D	0.26	0/2898	0.53	1/3934 (0.0%)
2	E	0.24	0/970	0.46	0/1320
2	H	0.26	0/987	0.49	0/1343
2	I	0.25	0/981	0.48	0/1335
2	M	0.26	0/964	0.50	0/1312
3	F	0.24	0/886	0.43	0/1207
3	J	0.25	0/886	0.43	0/1207
3	L	0.26	0/899	0.45	0/1224
3	N	0.24	0/890	0.43	0/1212
All	All	0.26	1/19621 (0.0%)	0.49	3/26646 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1
1	C	0	1
1	D	0	2
2	M	0	1
All	All	0	5

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	131	GLN	C-N	8.43	1.50	1.34

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	352	LEU	CA-CB-CG	6.36	129.93	115.30
1	A	280	ALA	CB-CA-C	-5.24	102.24	110.10
1	A	161	ASP	CB-CG-OD2	5.18	122.97	118.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	152	ILE	Peptide
1	C	130	ILE	Mainchain
1	D	191	GLU	Peptide
1	D	351	THR	Peptide
2	M	106	TYR	Peptide

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	3030	0	2947	63	0
1	B	3069	0	2991	70	0
1	C	2969	0	2891	81	0
1	D	2841	0	2730	71	0
2	E	945	0	917	20	0
2	H	962	0	932	27	0
2	I	956	0	927	29	0
2	M	939	0	912	32	0
3	F	865	0	829	16	0
3	J	865	0	829	16	0
3	L	878	0	843	29	0
3	N	869	0	832	16	0
All	All	19188	0	18580	444	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 12.

The worst 5 of 444 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)
1:B:53:MET:SD	1:B:130:ILE:HG22	1.36	1.65
1:B:53:MET:SD	1:B:130:ILE:CG2	2.25	1.24
1:C:53:MET:HE2	1:C:130:ILE:HB	1.52	0.91
1:C:131:GLN:OE1	1:B:77:GLN:NE2	2.12	0.83
1:C:342:PRO:HG2	1:C:391:VAL:HG22	1.63	0.80

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	396/405 (98%)	358 (90%)	38 (10%)	0	100 100
1	B	401/405 (99%)	369 (92%)	31 (8%)	1 (0%)	47 80
1	C	387/405 (96%)	351 (91%)	35 (9%)	1 (0%)	41 75
1	D	372/405 (92%)	338 (91%)	29 (8%)	5 (1%)	12 45
2	E	121/226 (54%)	113 (93%)	8 (7%)	0	100 100
2	H	124/226 (55%)	113 (91%)	9 (7%)	2 (2%)	9 41
2	I	123/226 (54%)	114 (93%)	8 (6%)	1 (1%)	19 57
2	M	120/226 (53%)	108 (90%)	12 (10%)	0	100 100
3	F	113/220 (51%)	108 (96%)	5 (4%)	0	100 100
3	J	113/220 (51%)	109 (96%)	4 (4%)	0	100 100
3	L	115/220 (52%)	107 (93%)	8 (7%)	0	100 100
3	N	114/220 (52%)	109 (96%)	5 (4%)	0	100 100
All	All	2499/3404 (73%)	2297 (92%)	192 (8%)	10 (0%)	34 70

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	192	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	153	VAL
1	D	193	ARG
2	H	42	PRO
2	H	41	PRO

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	329/338 (97%)	292 (89%)	37 (11%)	6 25
1	B	335/338 (99%)	294 (88%)	41 (12%)	5 22
1	C	322/338 (95%)	284 (88%)	38 (12%)	5 23
1	D	304/338 (90%)	260 (86%)	44 (14%)	3 16
2	E	104/196 (53%)	95 (91%)	9 (9%)	10 36
2	H	106/196 (54%)	92 (87%)	14 (13%)	4 19
2	I	105/196 (54%)	92 (88%)	13 (12%)	4 21
2	M	103/196 (53%)	91 (88%)	12 (12%)	5 24
3	F	98/185 (53%)	97 (99%)	1 (1%)	76 89
3	J	98/185 (53%)	95 (97%)	3 (3%)	40 70
3	L	99/185 (54%)	95 (96%)	4 (4%)	31 62
3	N	98/185 (53%)	94 (96%)	4 (4%)	30 62
All	All	2101/2876 (73%)	1881 (90%)	220 (10%)	7 29

5 of 220 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	369	THR
1	B	198	PHE
3	J	41	LYS
3	N	47	ARG
1	B	9	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
3	J	34	ASN
3	F	6	GLN
1	B	131	GLN
1	D	77	GLN
3	N	39	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 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

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, 95^{th} 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	400/405 (98%)	0.10	2 (0%) 91 88	16, 40, 70, 93	0
1	B	403/405 (99%)	0.11	4 (0%) 82 78	20, 41, 75, 101	0
1	C	391/405 (96%)	0.12	3 (0%) 86 82	20, 43, 72, 108	0
1	D	380/405 (93%)	0.11	2 (0%) 91 88	18, 40, 78, 104	0
2	E	123/226 (54%)	0.46	10 (8%) 12 14	38, 69, 85, 96	0
2	H	126/226 (55%)	0.18	4 (3%) 47 45	24, 59, 75, 84	0
2	I	125/226 (55%)	0.51	12 (9%) 8 10	39, 82, 102, 116	0
2	M	122/226 (53%)	0.21	0 100 100	33, 60, 79, 96	0
3	F	115/220 (52%)	0.31	6 (5%) 27 26	29, 53, 72, 98	0
3	J	115/220 (52%)	0.11	1 (0%) 84 80	34, 52, 79, 88	0
3	L	117/220 (53%)	-0.04	0 100 100	22, 41, 58, 67	0
3	N	116/220 (52%)	0.11	2 (1%) 70 66	23, 42, 62, 71	0
All	All	2533/3404 (74%)	0.16	46 (1%) 68 64	16, 47, 81, 116	0

The worst 5 of 46 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	390	GLY	5.4
1	C	391	VAL	5.0
2	E	120	LEU	3.8
2	E	123	VAL	3.8
2	E	11	LEU	3.5

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 (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.