Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 07:21 am BST

PDB ID	:	6I9I
Title	:	Rift valley fever virus Gn in complex with a neutralizing antibody fragment
Authors	:	Allen, E.R.; Bowden, T.A.
Deposited on	:	2018-11-23
Resolution	:	1.98 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 ⓘ symbol.

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

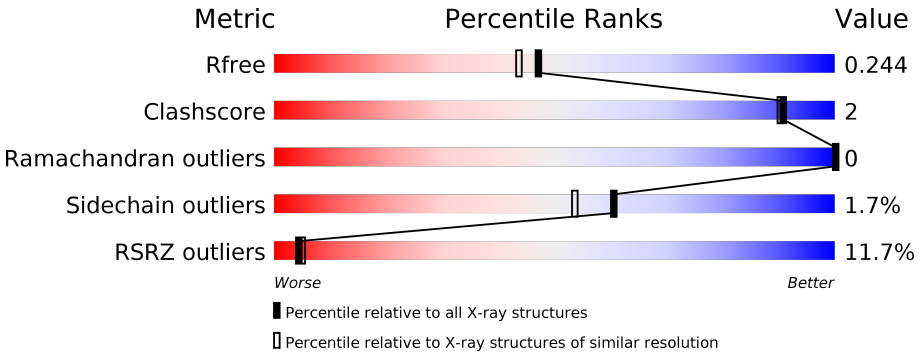
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 1.98 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	11647 (2.00-1.96)
Clashscore	141614	1014 (1.98-1.98)
Ramachandran outliers	138981	1006 (1.98-1.98)
Sidechain outliers	138945	1006 (1.98-1.98)
RSRZ outliers	127900	11410 (2.00-1.96)

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 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	223	10% 90% 7%
1	H	223	9% 92%
2	B	217	15% 95%
2	L	217	12% 93% 6%
3	C	325	2% 15% 84%
3	D	325	% 15% 84%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7491 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 RV-Gn1 Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	208	Total	C	N	O	S	0	0	0
			1538	969	254	308	7			
1	H	213	Total	C	N	O	S	0	0	0
			1570	987	259	317	7			

- Molecule 2 is a protein called RV-Gn1 Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	216	Total	C	N	O	S	0	0	0
			1592	988	261	334	9			
2	L	215	Total	C	N	O	S	0	0	0
			1585	984	260	332	9			

- Molecule 3 is a protein called Glycoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	52	Total	C	N	O	S	0	0	0
			385	238	66	75	6			
3	D	51	Total	C	N	O	S	0	0	0
			379	235	64	74	6			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	470	GLY	-	expression tag	UNP A2T080
C	471	THR	-	expression tag	UNP A2T080
C	472	LYS	-	expression tag	UNP A2T080
C	473	HIS	-	expression tag	UNP A2T080
C	474	HIS	-	expression tag	UNP A2T080
C	475	HIS	-	expression tag	UNP A2T080
C	476	HIS	-	expression tag	UNP A2T080
C	477	HIS	-	expression tag	UNP A2T080

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	478	HIS	-	expression tag	UNP A2T080
D	470	GLY	-	expression tag	UNP A2T080
D	471	THR	-	expression tag	UNP A2T080
D	472	LYS	-	expression tag	UNP A2T080
D	473	HIS	-	expression tag	UNP A2T080
D	474	HIS	-	expression tag	UNP A2T080
D	475	HIS	-	expression tag	UNP A2T080
D	476	HIS	-	expression tag	UNP A2T080
D	477	HIS	-	expression tag	UNP A2T080
D	478	HIS	-	expression tag	UNP A2T080

- Molecule 4 is water.

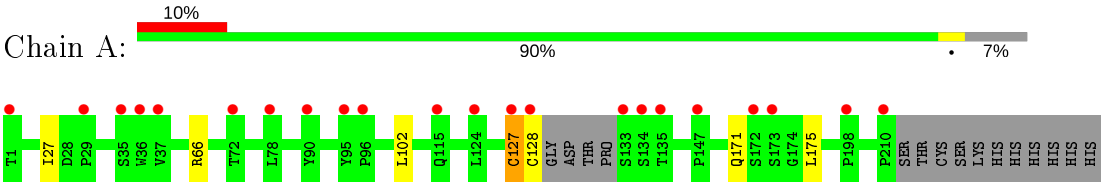
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	89	Total O 89 89	0	0
4	B	98	Total O 98 98	0	0
4	C	27	Total O 27 27	0	0
4	D	31	Total O 31 31	0	0
4	H	114	Total O 114 114	0	0
4	L	83	Total O 83 83	0	0

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 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.

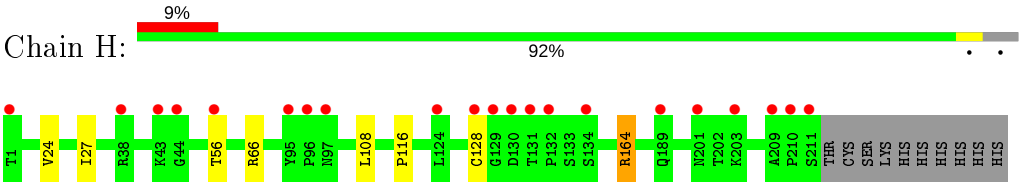
- Molecule 1: RV-Gn1 Heavy chain

Chain A:



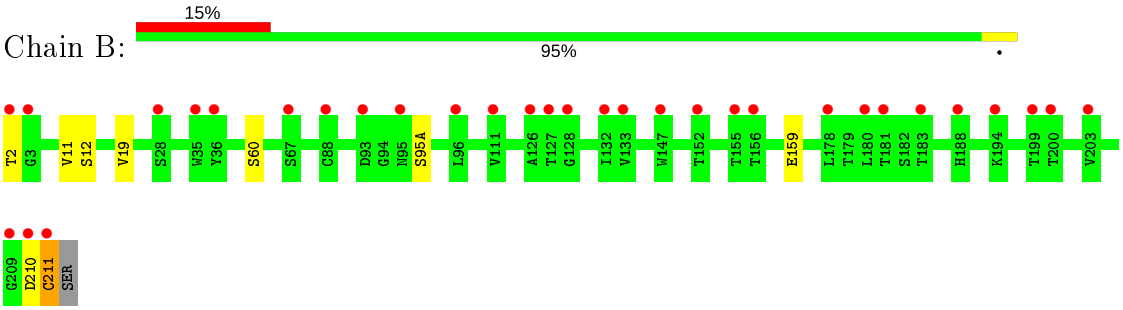
- Molecule 1: RV-Gn1 Heavy chain

Chain H:



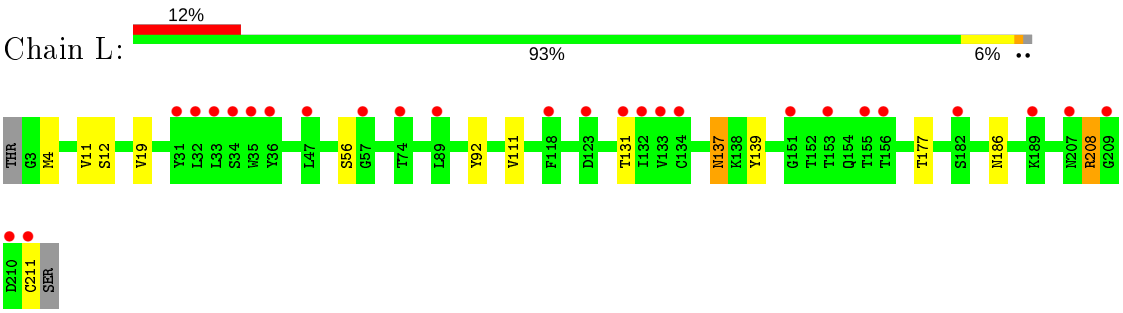
- Molecule 2: RV-Gn1 Light chain

Chain B:



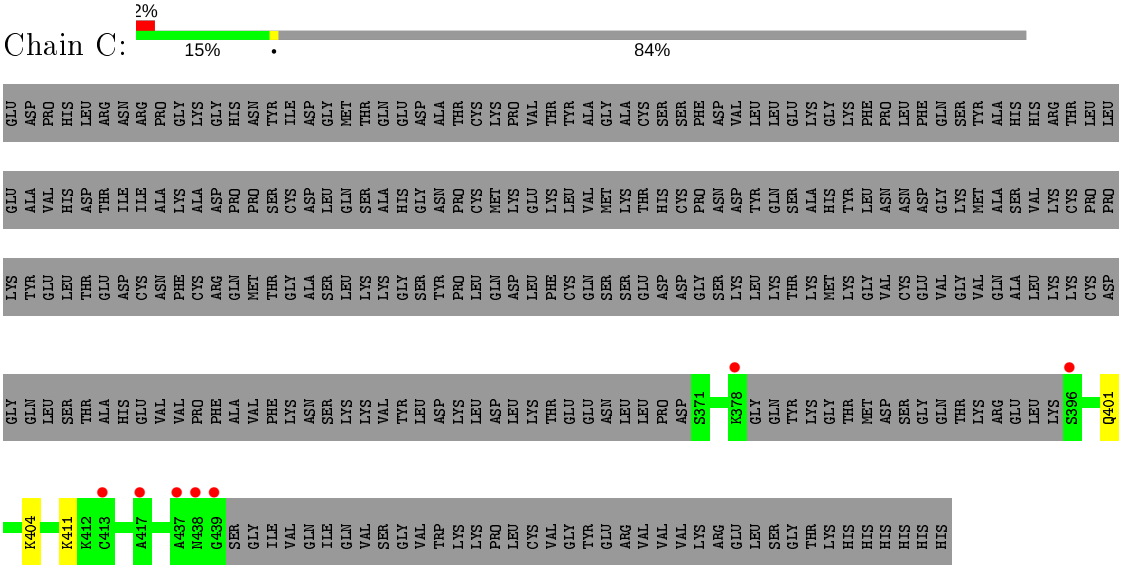
- Molecule 2: RV-Gn1 Light chain

Chain L:



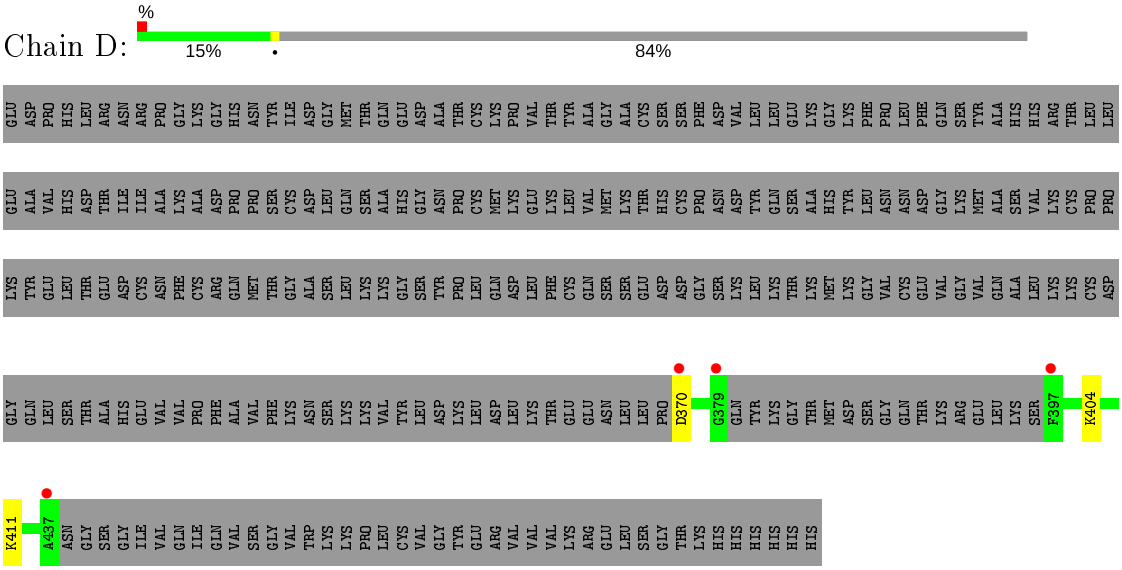
- Molecule 3: Glycoprotein

Chain C:



- Molecule 3: Glycoprotein

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	51.59Å 61.95Å 78.91Å 81.01° 77.52° 84.74°	Depositor
Resolution (Å)	23.60 – 1.98 23.64 – 1.98	Depositor EDS
% Data completeness (in resolution range)	96.2 (23.60-1.98) 96.3 (23.64-1.98)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.59 (at 1.98Å)	Xtriage
Refinement program	REFMAC 5.8.0189	Depositor
R, R_{free}	0.203 , 0.243 0.208 , 0.244	Depositor DCC
R_{free} test set	3141 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	36.4	Xtriage
Anisotropy	0.508	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 41.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	7491	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 6.26% 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	0.52	0/1577	0.74	2/2170 (0.1%)
1	H	0.55	0/1611	0.75	3/2219 (0.1%)
2	B	0.52	0/1624	0.69	0/2225
2	L	0.54	0/1617	0.69	2/2215 (0.1%)
3	C	0.54	0/394	0.63	0/527
3	D	0.52	0/388	0.62	0/519
All	All	0.53	0/7211	0.71	7/9875 (0.1%)

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	66	ARG	NE-CZ-NH2	9.00	124.80	120.30
1	A	66	ARG	NE-CZ-NH1	-8.29	116.16	120.30
1	H	66	ARG	NE-CZ-NH1	7.53	124.06	120.30
2	L	208	ARG	NE-CZ-NH1	-6.78	116.91	120.30
2	L	208	ARG	NE-CZ-NH2	6.48	123.54	120.30
1	H	66	ARG	NE-CZ-NH2	-6.48	117.06	120.30
1	H	164	ARG	NE-CZ-NH2	-5.24	117.68	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	1538	0	1514	5	0
1	H	1570	0	1541	4	0
2	B	1592	0	1524	5	0
2	L	1585	0	1517	13	0
3	C	385	0	338	1	0
3	D	379	0	331	0	0
4	A	89	0	0	0	0
4	B	98	0	0	1	0
4	C	27	0	0	0	0
4	D	31	0	0	0	0
4	H	114	0	0	1	0
4	L	83	0	0	0	0
All	All	7491	0	6765	23	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 2.

All (23) 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:L:186:ASN:C	2:L:208:ARG:NH1	2.32	0.83
2:L:186:ASN:C	2:L:208:ARG:HH12	1.90	0.74
2:L:4:MET:SD	2:L:92:TYR:OH	2.46	0.73
2:L:186:ASN:HA	2:L:208:ARG:CZ	2.21	0.71
1:A:128:CYS:N	2:B:211:CYS:SG	2.68	0.66
2:L:111:VAL:HG23	2:L:139:TYR:HA	1.81	0.61
1:H:128:CYS:N	2:L:211:CYS:SG	2.74	0.58
2:L:186:ASN:CA	2:L:208:ARG:NH1	2.69	0.56
2:L:186:ASN:HA	2:L:208:ARG:NH1	2.21	0.54
1:A:127:CYS:SG	2:B:210:ASP:O	2.68	0.51
2:B:11:VAL:HG11	2:B:19:VAL:CG2	2.41	0.50
1:H:164:ARG:NH1	2:L:137:ASN:OD1	2.43	0.49
1:A:171:GLN:NE2	2:B:159:GLU:OE1	2.38	0.49
2:L:131:THR:HG23	2:L:177:THR:CG2	2.44	0.47
2:L:186:ASN:O	2:L:208:ARG:NH1	2.47	0.45
1:A:27:ILE:HD11	1:A:102:LEU:HD12	2.00	0.43
2:L:11:VAL:CG1	2:L:19:VAL:HG23	2.48	0.43
2:L:11:VAL:HG11	2:L:19:VAL:HG23	2.00	0.42
1:A:175:LEU:HD22	1:A:175:LEU:N	2.36	0.41
1:H:108:LEU:CD1	4:H:396:HOH:O	2.69	0.41
2:B:2:THR:N	4:B:308:HOH:O	2.54	0.41
3:C:401:GLN:HB3	3:C:411:LYS:HG2	2.03	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:H:24:VAL:CG1	1:H:27:ILE:HG13	2.51	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	204/223 (92%)	204 (100%)	0	0	100	100
1	H	211/223 (95%)	209 (99%)	2 (1%)	0	100	100
2	B	214/217 (99%)	209 (98%)	5 (2%)	0	100	100
2	L	213/217 (98%)	207 (97%)	6 (3%)	0	100	100
3	C	48/325 (15%)	48 (100%)	0	0	100	100
3	D	47/325 (14%)	47 (100%)	0	0	100	100
All	All	937/1530 (61%)	924 (99%)	13 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	178/192 (93%)	177 (99%)	1 (1%)	86	85
1	H	182/192 (95%)	180 (99%)	2 (1%)	73	70

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	B	182/183 (100%)	178 (98%)	4 (2%)	52 46
2	L	181/183 (99%)	178 (98%)	3 (2%)	60 53
3	C	40/279 (14%)	39 (98%)	1 (2%)	47 39
3	D	39/279 (14%)	36 (92%)	3 (8%)	13 4
All	All	802/1308 (61%)	788 (98%)	14 (2%)	60 53

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	127	CYS
2	B	12	SER
2	B	60	SER
2	B	95(A)	SER
2	B	211	CYS
3	C	404	LYS
3	D	370	ASP
3	D	404	LYS
3	D	411	LYS
1	H	56	THR
1	H	116	PRO
2	L	12	SER
2	L	56	SER
2	L	137	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	208/223 (93%)	0.74	22 (10%) 6 7	31, 48, 69, 107	1 (0%)
1	H	213/223 (95%)	0.70	21 (9%) 7 8	26, 44, 68, 107	2 (0%)
2	B	216/217 (99%)	0.90	32 (14%) 2 2	29, 46, 86, 112	2 (0%)
2	L	215/217 (99%)	0.81	26 (12%) 4 4	32, 49, 79, 103	3 (1%)
3	C	52/325 (16%)	0.76	7 (13%) 3 3	33, 42, 69, 100	0
3	D	51/325 (15%)	0.50	4 (7%) 13 14	32, 38, 61, 82	0
All	All	955/1530 (62%)	0.77	112 (11%) 4 5	26, 46, 79, 112	8 (0%)

All (112) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	211	CYS	10.7
1	A	128	CYS	9.5
1	H	128	CYS	9.5
2	B	67	SER	7.1
1	H	56	THR	6.9
1	H	132	PRO	6.8
3	D	370	ASP	6.8
1	H	211	SER	6.7
1	A	1	THR	6.1
1	H	38	ARG	6.0
3	C	439	GLY	5.8
2	L	133	VAL	5.6
1	H	1	THR	5.4
2	B	209	GLY	5.3
1	H	129	GLY	5.3
2	L	132	ILE	5.0
2	L	31	TYR	4.7
2	L	211	CYS	4.7
2	B	183	THR	4.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	D	437	ALA	4.4
2	B	132	ILE	4.3
2	L	153	THR	4.3
1	H	130	ASP	4.1
2	L	151	GLY	4.1
2	B	152	THR	4.1
1	A	127	CYS	3.9
1	H	203	LYS	3.8
2	B	200	THR	3.8
2	L	34	SER	3.8
1	H	209	ALA	3.8
2	L	155	THR	3.8
2	B	199	THR	3.6
1	H	210	PRO	3.6
3	D	379	GLY	3.5
2	L	74	THR	3.5
1	H	131	THR	3.5
1	H	134	SER	3.5
2	B	133	VAL	3.5
2	B	210	ASP	3.5
2	L	47	LEU	3.4
2	L	210	ASP	3.4
1	A	135	THR	3.4
2	L	33	LEU	3.4
3	C	396	SER	3.3
2	B	156	THR	3.2
2	L	209	GLY	3.2
2	L	36	TYR	3.2
1	A	36	TRP	3.2
2	B	93	ASP	3.1
2	B	128	GLY	3.1
2	L	89	LEU	3.1
1	A	172	SER	3.1
2	B	95	ASN	3.1
2	L	35	TRP	3.1
2	B	181	THR	3.0
2	B	126	ALA	3.0
2	L	207	ASN	3.0
3	C	438	ASN	2.9
3	D	397	PHE	2.9
1	A	95	TYR	2.9
1	H	201	ASN	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	35	TRP	2.9
1	A	210	PRO	2.8
1	A	133	SER	2.8
1	A	35	SER	2.8
2	B	2	THR	2.8
2	L	32	LEU	2.8
3	C	437	ALA	2.8
1	A	173	SER	2.8
1	A	124	LEU	2.7
1	H	43	LYS	2.7
1	A	96	PRO	2.7
3	C	378	LYS	2.7
1	H	96	PRO	2.7
2	B	147	TRP	2.6
2	L	156	THR	2.6
2	B	178	LEU	2.6
1	A	29	PRO	2.6
2	L	134	CYS	2.6
2	B	111	VAL	2.6
3	C	413	CYS	2.5
1	H	44	GLY	2.5
1	A	37	VAL	2.5
2	B	28	SER	2.4
1	H	124	LEU	2.4
2	B	188	HIS	2.4
2	B	155	THR	2.4
1	H	97	ASN	2.3
2	L	123	ASP	2.3
1	A	134	SER	2.3
1	A	147	PRO	2.3
2	L	131	THR	2.3
1	A	78	LEU	2.3
2	L	57	GLY	2.3
2	B	3	GLY	2.2
2	B	96	LEU	2.2
1	A	115	GLN	2.2
1	H	95	TYR	2.2
2	B	180	LEU	2.2
2	B	36	TYR	2.2
2	L	118	PHE	2.2
3	C	417	ALA	2.1
2	L	182	SER	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	H	189	GLN	2.1
2	L	189	LYS	2.1
1	A	90	TYR	2.1
1	A	198	PRO	2.1
1	A	72	THR	2.1
2	B	203	VAL	2.1
2	B	88	CYS	2.0
2	B	194	LYS	2.0
2	B	127	THR	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 carbohydrates 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.