Full wwPDB X-ray Structure Validation Report ⓘ

Mar 7, 2022 – 10:24 AM EST

PDB ID	:	7RXJ
Title	:	Fab236 in complex with the C-terminal alpha-TSR domain of P. falciparum
Authors	:	Pholcharee, T.; Oyen, D.; Wilson, I.A.
Deposited on	:	2021-08-23
Resolution	:	2.35 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 ⓘ) were used in the production of this report:

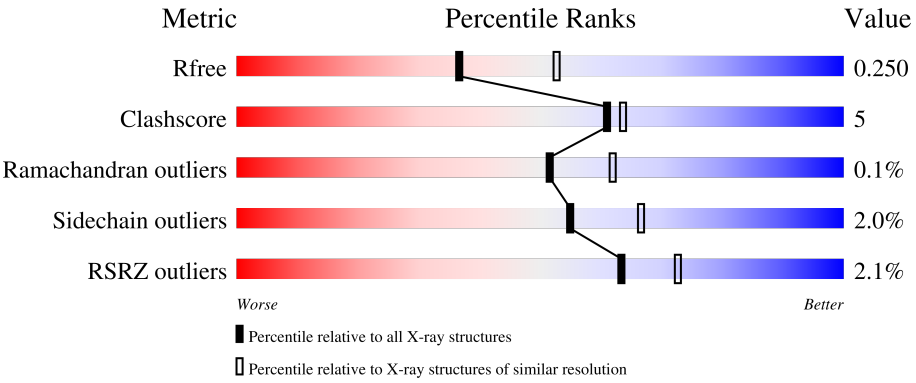
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.27
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

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 2.35 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.



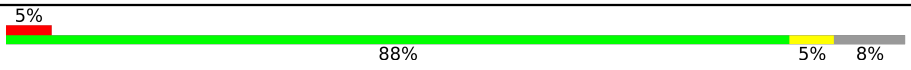
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2096 (2.36-2.32)
Clashscore	141614	2193 (2.36-2.32)
Ramachandran outliers	138981	2159 (2.36-2.32)
Sidechain outliers	138945	2160 (2.36-2.32)
RSRZ outliers	127900	2067 (2.36-2.32)

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	B	212	92% 6% •
1	L	212	% 90% 8% •
2	A	227	% 86% 12% •
2	H	227	4% 81% 13% 5%
3	G	66	3% 74% 17% • • 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	I	66	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 14317 atoms, of which 6918 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 Fab236 light chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	B	209	Total	C	H	N	O	S	0	0	0
			3032	975	1478	256	318	5			
1	L	210	Total	C	H	N	O	S	0	0	0
			3058	981	1491	260	321	5			

- Molecule 2 is a protein called Fab236 heavy chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	A	223	Total	C	H	N	O	S	0	0	0
			3200	1032	1565	276	321	6			
2	H	216	Total	C	H	N	O	S	0	0	0
			3107	1008	1515	268	310	6			

- Molecule 3 is a protein called Circumsporozoite protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	I	61	Total	C	H	N	O	S	0	0	0
			860	278	417	76	85	4			
3	G	62	Total	C	H	N	O	S	0	0	0
			921	291	452	81	93	4			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	31	Total	O	0	0
			31	31		
4	A	37	Total	O	0	0
			37	37		
4	L	36	Total	O	0	0
			36	36		
4	H	27	Total	O	0	0
			27	27		

Continued on next page...

Continued from previous page...

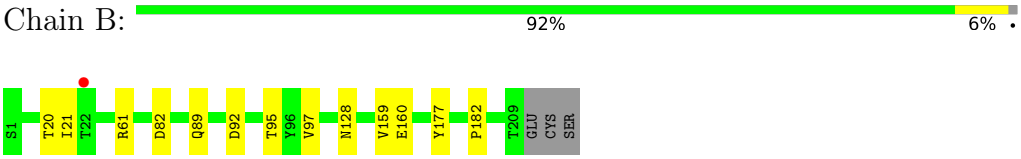
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	I	3	Total O 3 3	0	0
4	G	5	Total O 5 5	0	0

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.

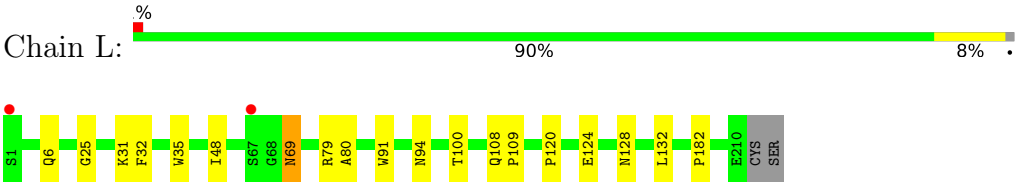
• Molecule 1: Fab236 light chain

Chain B:



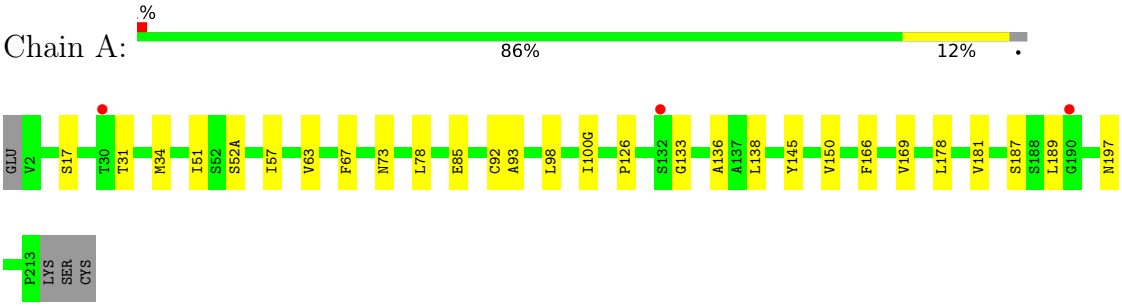
• Molecule 1: Fab236 light chain

Chain L:



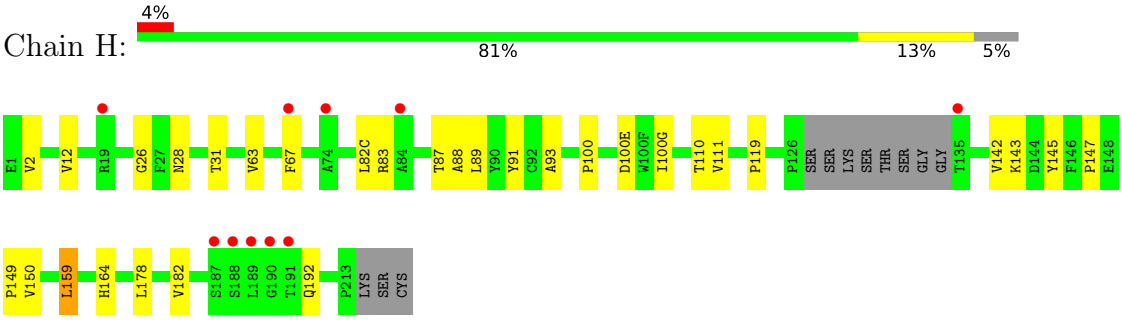
• Molecule 2: Fab236 heavy chain

Chain A:



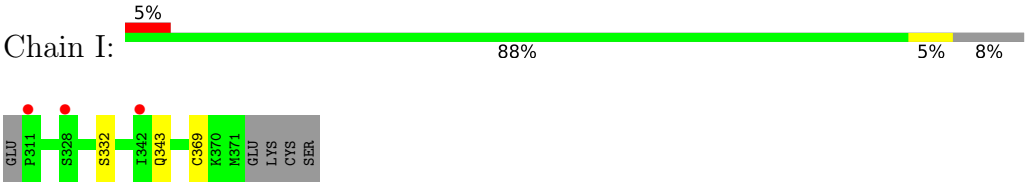
• Molecule 2: Fab236 heavy chain

Chain H:



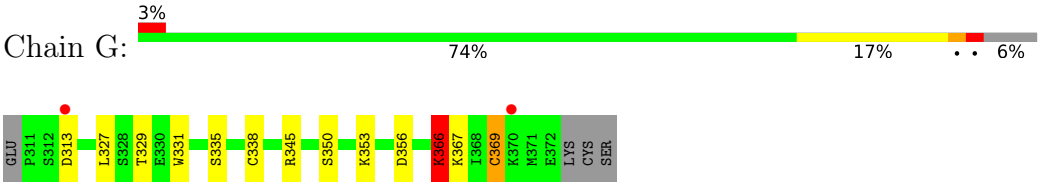
• Molecule 3: Circumsporozoite protein

Chain I:



• Molecule 3: Circumsporozoite protein

Chain G:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	130.14Å 45.56Å 201.09Å 90.00° 94.04° 90.00°	Depositor
Resolution (Å)	42.24 – 2.35 42.24 – 2.34	Depositor EDS
% Data completeness (in resolution range)	92.3 (42.24-2.35) 92.4 (42.24-2.34)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.12	Depositor
$< I/\sigma(I) >$ [1]	3.52 (at 2.34Å)	Xtriage
Refinement program	PHENIX (1.16_3549: ???)	Depositor
R, R_{free}	0.217 , 0.249 0.218 , 0.250	Depositor DCC
R_{free} test set	2438 reflections (5.26%)	wwPDB-VP
Wilson B-factor (Å^2)	38.8	Xtriage
Anisotropy	0.361	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 31.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	14317	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.31% 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	B	0.38	0/1593	0.54	0/2182
1	L	0.43	0/1606	0.56	0/2199
2	A	0.49	0/1677	0.59	0/2295
2	H	0.50	0/1633	0.60	0/2237
3	G	0.50	0/477	0.67	1/644 (0.2%)
3	I	0.45	0/451	0.60	0/611
All	All	0.46	0/7437	0.58	1/10168 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	G	366	LYS	CB-CA-C	5.96	122.31	110.40

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	B	1554	1478	1480	8	0
1	L	1567	1491	1493	14	0
2	A	1635	1565	1565	19	0
2	H	1592	1515	1515	20	0
3	G	469	452	450	7	0
3	I	443	417	417	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	37	0	0	0	0
4	B	31	0	0	0	0
4	G	5	0	0	0	0
4	H	27	0	0	0	0
4	I	3	0	0	0	0
4	L	36	0	0	0	0
All	All	7399	6918	6920	64	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 (64) 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:G:369:CYS:SG	3:G:369:CYS:O	2.44	0.75
2:A:34:MET:HB3	2:A:78:LEU:HD22	1.79	0.64
2:H:63:VAL:HG21	2:H:67:PHE:CD2	2.37	0.60
1:L:120:PRO:HD3	1:L:132:LEU:CD2	2.32	0.59
2:A:150:VAL:HG22	2:A:178:LEU:HD21	1.85	0.59
1:L:108:GLN:HB2	1:L:109:PRO:HD2	1.87	0.57
2:A:31:THR:O	2:A:31:THR:OG1	2.18	0.56
1:L:91:TRP:CZ2	1:L:94:ASN:HA	2.40	0.56
3:G:331:TRP:CZ3	3:G:366:LYS:O	2.59	0.55
2:A:51:ILE:HD12	2:A:57:ILE:HG12	1.88	0.54
1:L:31:LYS:NZ	1:L:91:TRP:O	2.39	0.54
2:A:166:PHE:HE1	2:A:181:VAL:HG12	1.73	0.53
1:B:160:GLU:HB3	2:A:169:VAL:HG21	1.91	0.53
2:H:142:VAL:HG11	2:H:150:VAL:HG21	1.90	0.53
2:H:93:ALA:HB1	2:H:100(G):ILE:CG2	2.38	0.53
2:A:63:VAL:CG1	2:A:67:PHE:HB2	2.39	0.52
1:L:25:GLY:O	1:L:69:ASN:HB2	2.10	0.51
2:H:12:VAL:HG11	2:H:82(C):LEU:HD13	1.92	0.51
1:B:61:ARG:NH2	1:B:82:ASP:OD2	2.44	0.50
2:H:178:LEU:C	2:H:178:LEU:HD12	2.32	0.50
2:A:63:VAL:HG13	2:A:67:PHE:HB2	1.93	0.49
2:A:98:LEU:HA	3:G:356:ASP:HA	1.94	0.49
1:B:92:ASP:O	1:B:95:THR:HB	2.13	0.49
1:L:79:ARG:HD3	1:L:80:ALA:N	2.27	0.49
1:L:6:GLN:HB2	1:L:100:THR:OG1	2.13	0.49
2:A:78:LEU:HD23	2:A:92:CYS:SG	2.53	0.48
2:A:126:PRO:HG3	2:A:138:LEU:HB3	1.95	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:150:VAL:CG2	2:A:178:LEU:HD21	2.45	0.46
1:B:21:ILE:HD12	1:B:21:ILE:N	2.31	0.46
2:H:89:LEU:HD23	2:H:91:TYR:CZ	2.51	0.46
3:G:331:TRP:HZ3	3:G:366:LYS:O	1.97	0.45
2:H:150:VAL:HG22	2:H:178:LEU:HD21	1.99	0.45
1:B:89:GLN:HA	1:B:97:VAL:O	2.17	0.45
3:G:350:SER:HB3	3:G:353:LYS:HD2	1.99	0.45
2:H:28:ASN:O	2:H:31:THR:HG22	2.18	0.44
2:A:145:TYR:CE2	2:A:150:VAL:HG13	2.52	0.44
2:H:87:THR:HG23	2:H:110:THR:HA	1.98	0.44
1:L:91:TRP:CE2	2:H:100:PRO:HB2	2.53	0.43
2:A:52(A):SER:O	2:A:73:ASN:ND2	2.40	0.43
1:L:128:ASN:HA	1:L:182:PRO:HG2	2.00	0.43
2:H:2:VAL:HG12	2:H:26:GLY:O	2.19	0.43
2:H:145:TYR:CE1	2:H:150:VAL:HG13	2.53	0.43
1:L:32:PHE:HB2	2:H:100(E):ASP:OD2	2.19	0.43
2:H:119:PRO:HB3	2:H:145:TYR:HB3	2.01	0.43
1:L:132:LEU:N	1:L:132:LEU:HD23	2.34	0.42
1:B:128:ASN:HA	1:B:182:PRO:HG2	2.01	0.42
1:L:35:TRP:HB2	1:L:48:ILE:HB	2.01	0.42
2:H:83:ARG:C	2:H:111:VAL:HG11	2.40	0.42
2:A:136:ALA:CB	2:A:189:LEU:HD11	2.49	0.42
2:A:93:ALA:HB1	2:A:100(G):ILE:CG2	2.50	0.42
1:B:159:VAL:HA	1:B:177:TYR:O	2.20	0.42
1:L:124:GLU:OE2	2:H:143:LYS:HE2	2.20	0.41
2:H:87:THR:O	2:H:88:ALA:HB2	2.19	0.41
2:H:93:ALA:HB1	2:H:100(G):ILE:HG22	2.02	0.41
3:G:327:LEU:HD12	3:G:327:LEU:HA	1.89	0.41
2:A:178:LEU:C	2:A:178:LEU:HD12	2.40	0.41
1:L:79:ARG:HD3	1:L:80:ALA:H	1.83	0.41
1:B:20:THR:O	1:B:20:THR:HG23	2.20	0.41
3:G:329:THR:O	3:G:345:ARG:HD2	2.21	0.41
2:A:85:GLU:OE2	2:A:85:GLU:N	2.50	0.41
2:H:63:VAL:CG2	2:H:67:PHE:CG	3.03	0.41
2:A:63:VAL:HG13	2:A:67:PHE:CG	2.56	0.40
2:H:159:LEU:HD13	2:H:182:VAL:HG21	2.03	0.40
3:I:332:SER:OG	3:I:343:GLN:HA	2.22	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	B	207/212 (98%)	200 (97%)	7 (3%)	0	100 100
1	L	208/212 (98%)	200 (96%)	8 (4%)	0	100 100
2	A	221/227 (97%)	214 (97%)	6 (3%)	1 (0%)	29 31
2	H	212/227 (93%)	204 (96%)	8 (4%)	0	100 100
3	G	60/66 (91%)	56 (93%)	4 (7%)	0	100 100
3	I	59/66 (89%)	59 (100%)	0	0	100 100
All	All	967/1010 (96%)	933 (96%)	33 (3%)	1 (0%)	51 62

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	133	GLY

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	B	171/181 (94%)	171 (100%)	0	100 100
1	L	173/181 (96%)	172 (99%)	1 (1%)	86 92
2	A	179/190 (94%)	176 (98%)	3 (2%)	60 72
2	H	172/190 (90%)	167 (97%)	5 (3%)	42 52
3	G	52/61 (85%)	46 (88%)	6 (12%)	5 4
3	I	46/61 (75%)	45 (98%)	1 (2%)	52 63

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	793/864 (92%)	777 (98%)	16 (2%)	55 66

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	17	SER
2	A	187	SER
2	A	197	ASN
1	L	69	ASN
2	H	147	PRO
2	H	149	PRO
2	H	159	LEU
2	H	164	HIS
2	H	192	GLN
3	I	369	CYS
3	G	313	ASP
3	G	335	SER
3	G	338	CYS
3	G	366	LYS
3	G	367	LYS
3	G	369	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
3	G	343	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	B	209/212 (98%)	0.19	1 (0%)	91	95	23, 39, 52, 64	0
1	L	210/212 (99%)	0.19	2 (0%)	82	88	25, 38, 54, 71	0
2	A	223/227 (98%)	0.36	3 (1%)	77	83	24, 39, 56, 66	0
2	H	216/227 (95%)	0.40	10 (4%)	32	43	26, 42, 58, 75	0
3	G	62/66 (93%)	0.48	2 (3%)	47	58	39, 55, 72, 81	0
3	I	61/66 (92%)	0.42	3 (4%)	29	40	40, 54, 65, 68	0
All	All	981/1010 (97%)	0.31	21 (2%)	63	73	23, 41, 61, 81	0

All (21) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	191	THR	3.8
1	L	1	SER	3.2
3	I	311	PRO	3.1
2	H	74	ALA	3.0
2	A	190	GLY	2.7
2	H	190	GLY	2.7
1	B	22	THR	2.5
3	G	370	LYS	2.5
1	L	67	SER	2.5
2	H	135	THR	2.5
2	H	189	LEU	2.5
2	H	67	PHE	2.3
3	G	313	ASP	2.3
3	I	342	ILE	2.2
2	H	188	SER	2.2
2	H	19	ARG	2.2
2	A	132	SER	2.1
3	I	328	SER	2.1
2	H	84	ALA	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	H	187	SER	2.1
2	A	30	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 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.