Full wwPDB X-ray Structure Validation Report ⓘ

May 15, 2020 – 05:55 am BST

PDB ID	:	6AVF
Title	:	Crystal structure of the KFJ5 TCR-NY-ESO-1-HLA-B*07:02 complex
Authors	:	Gully, B.S.; Gras, S.; Rossjohn, J.
Deposited on	:	2017-09-02
Resolution	:	2.03 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:

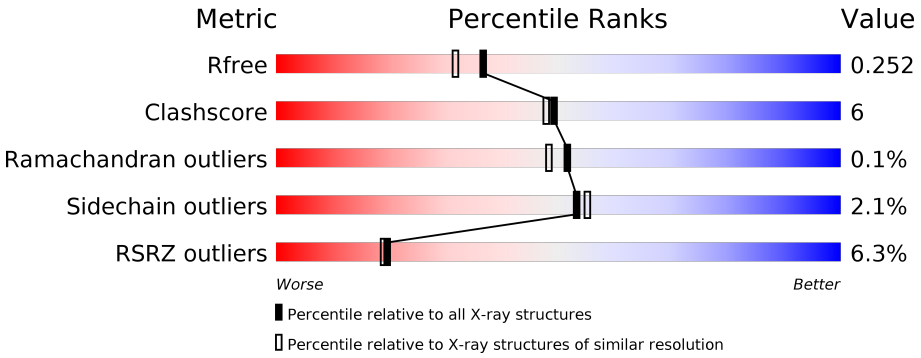
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.03 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	10434 (2.04-2.00)
Clashscore	141614	11643 (2.04-2.00)
Ramachandran outliers	138981	11493 (2.04-2.00)
Sidechain outliers	138945	11492 (2.04-2.00)
RSRZ outliers	127900	10220 (2.04-2.00)

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 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	M	99	10%, 76%, 17%, 6%
2	A	207	88%, 9%
3	B	244	91%, 8%
4	P	13	85%, 15%
5	H	362	10%, 59%, 15%, 26%

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 7027 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 Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	M	93	Total	C	N	O	S	0	0	0
			762	491	129	140	2			

- Molecule 2 is a protein called T-cell receptor alpha variable 4,TCR alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	202	Total	C	N	O	S	0	2	0
			1605	1014	261	322	8			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	MET	-	initiating methionine	UNP A0A0B4J268

- Molecule 3 is a protein called T-cell receptor beta variable 28,TCR beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	242	Total	C	N	O	S	0	1	0
			1964	1239	342	374	9			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	initiating methionine	UNP A0A5B6
B	95	GLN	LEU	conflict	UNP A0A5B6

- Molecule 4 is a protein called ALA-PRO-ARG-GLY-PRO-HIS-GLY-GLY-ALA-ALA-SER-GLY-LEU.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	P	13	Total	C	N	O	0	0	0
			81	48	18	15			

- Molecule 5 is a protein called HLA class I histocompatibility antigen, B-7 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	H	269	Total	C	N	O	S	0	0	0
			2168	1353	394	415	6			

- Molecule 6 is water.

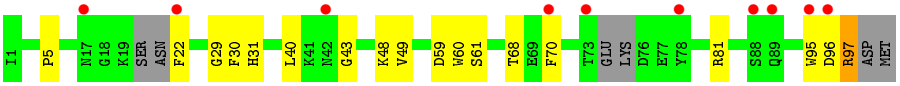
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	M	16	Total	O	0	0
			16	16		
6	A	136	Total	O	0	0
			136	136		
6	B	202	Total	O	0	0
			202	202		
6	P	4	Total	O	0	0
			4	4		
6	H	89	Total	O	0	0
			89	89		

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.

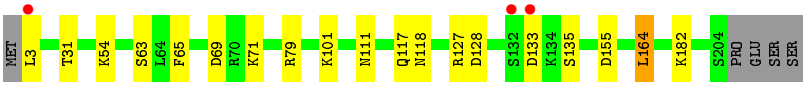
- Molecule 1: Beta-2-microglobulin

Chain M:



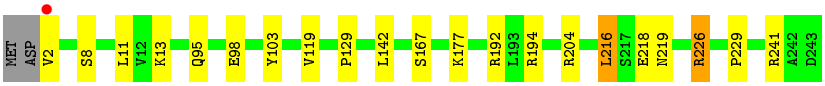
- Molecule 2: T-cell receptor alpha variable 4,TCR alpha chain

Chain A:



- Molecule 3: T-cell receptor beta variable 28,TCR beta chain

Chain B:

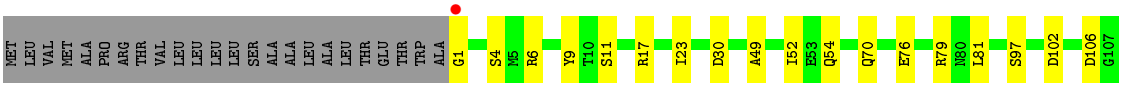


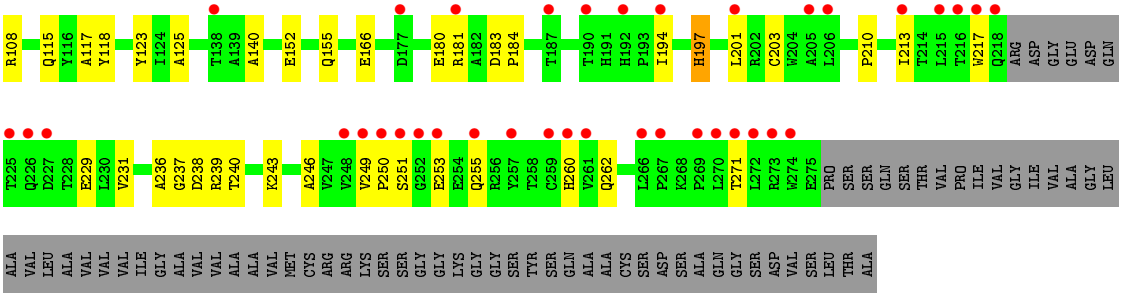
- Molecule 4: ALA-PRO-ARG-GLY-PRO-HIS-GLY-GLY-ALA-ALA-SER-GLY-LEU

Chain P:

- Molecule 5: HLA class I histocompatibility antigen, B-7 alpha chain

Chain H:




R108 Q115 Y116 A117 Y118 Y123 I124 A125 T138 A139 A140 E152 Q155 E166 D177 E180 R181 A182 D183 P184 T187 T190 H191 H192 P193 I194 H197 L201 R202 C203 W204 A205 L206 P210 I213 T214 L215 T216 W217 Q218 ARG ASP GLY GLU ASP GLN
T225 Q226 D227 T228 E229 L230 V231 A236 G237 D238 R239 T240 K243 A246 V247 V248 V249 P250 S251 G252 E253 E254 Q255 R256 Y257 T258 C259 H260 V261 Q262 L266 P267 K268 P269 L270 T271 L272 R273 W274 E275 PRO SER SER GLN SER THR VAL PRO ILE VAL GLY ILE VAL ALA GLY LEU
ALA VAL LEU ALA VAL VAL ILE GLY ALA VAL VAL ALA ALA VAL MET CYS ARG ARG LYS SER SER GLY GLY LYS GLY GLY SER TYR SER GLN ALA ALA CYS SER ASP SER ALA GLN GLY SER ASP VAL SER LEU THR ALA

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	64.37Å 67.68Å 105.72Å 90.00° 102.93° 90.00°	Depositor
Resolution (Å)	46.01 – 2.03 46.01 – 2.03	Depositor EDS
% Data completeness (in resolution range)	97.3 (46.01-2.03) 99.0 (46.01-2.03)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.75 (at 2.03Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.204 , 0.256 0.202 , 0.252	Depositor DCC
R_{free} test set	1997 reflections (3.50%)	wwPDB-VP
Wilson B-factor (Å^2)	25.8	Xtriage
Anisotropy	0.480	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 52.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.44$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	7027	wwPDB-VP
Average B, all atoms (Å^2)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.79% 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	M	0.35	0/783	0.55	0/1062
2	A	0.53	0/1648	0.65	1/2239 (0.0%)
3	B	0.54	0/2016	0.61	0/2732
4	P	0.41	0/83	0.67	0/110
5	H	0.39	0/2229	0.54	0/3034
All	All	0.47	0/6759	0.59	1/9177 (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})
2	A	164	LEU	CA-CB-CG	6.11	129.34	115.30

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	M	762	0	717	15	0
2	A	1605	0	1535	12	0
3	B	1964	0	1889	13	0
4	P	81	0	79	1	0
5	H	2168	0	1997	39	0
6	A	136	0	0	3	0
6	B	202	0	0	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	H	89	0	0	3	0
6	M	16	0	0	3	0
6	P	4	0	0	0	0
All	All	7027	0	6217	76	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.

All (76) 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)
5:H:238:ASP:HB2	5:H:240:THR:HG22	1.59	0.85
5:H:181:ARG:HH22	5:H:239:ARG:HB3	1.42	0.83
5:H:6:ARG:NH2	5:H:102:ASP:OD1	2.17	0.77
1:M:29:GLY:HA2	1:M:61:SER:HB2	1.69	0.73
5:H:6:ARG:HH11	5:H:6:ARG:HG3	1.54	0.72
5:H:54:GLN:NE2	6:H:401:HOH:O	2.27	0.67
5:H:17:ARG:NH2	6:H:402:HOH:O	2.28	0.66
1:M:22:PHE:HA	1:M:68:THR:O	1.97	0.65
2:A:111[B]:ASN:OD1	6:A:301:HOH:O	2.15	0.64
2:A:3:LEU:N	6:A:303:HOH:O	2.32	0.63
1:M:31:HIS:ND1	6:M:102:HOH:O	2.22	0.62
5:H:6:ARG:NH1	5:H:6:ARG:HG3	2.13	0.62
5:H:54:GLN:HG3	6:H:433:HOH:O	2.02	0.60
1:M:22:PHE:N	1:M:70:PHE:H	2.01	0.59
5:H:76:GLU:CD	5:H:79:ARG:NH1	2.56	0.59
3:B:119:VAL:O	3:B:226:ARG:NH1	2.35	0.59
5:H:76:GLU:OE2	5:H:79:ARG:NH1	2.36	0.59
5:H:106:ASP:OD2	5:H:108:ARG:HG2	2.03	0.58
2:A:63:SER:HG	2:A:65[A]:PHE:HE1	1.53	0.56
3:B:241:ARG:NH2	6:B:309:HOH:O	2.37	0.56
5:H:236:ALA:HB3	5:H:240:THR:HG23	1.87	0.56
1:M:97:ARG:NH1	1:M:97:ARG:O	2.39	0.56
3:B:204:ARG:NH2	6:B:306:HOH:O	2.29	0.55
3:B:216:LEU:HD22	3:B:229:PRO:HG2	1.89	0.55
5:H:201:LEU:HD12	5:H:249:VAL:HG21	1.89	0.55
2:A:164:LEU:HD11	3:B:194:ARG:HB2	1.90	0.54
4:P:6:HIS:O	5:H:155:GLN:HG3	2.08	0.53
5:H:203:CYS:HB2	5:H:217:TRP:CZ2	2.44	0.53
5:H:213:ILE:HG13	5:H:262:GLN:O	2.09	0.53
5:H:250:PRO:O	5:H:253:GLU:HB3	2.09	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:101:LYS:HG2	2:A:101:LYS:O	2.10	0.52
3:B:218:GLU:O	3:B:219:ASN:HB2	2.09	0.52
2:A:128:ASP:OD2	2:A:133:ASP:OD1	2.28	0.52
5:H:181:ARG:HH22	5:H:239:ARG:CB	2.18	0.51
5:H:9:TYR:HB2	5:H:97:SER:HB2	1.92	0.51
1:M:95:TRP:CD2	1:M:96:ASP:HB3	2.46	0.51
1:M:43:GLY:O	6:M:101:HOH:O	2.18	0.51
3:B:2:VAL:HG13	3:B:103:TYR:HD2	1.76	0.50
5:H:76:GLU:CD	5:H:79:ARG:HH12	2.14	0.50
2:A:117:GLN:NE2	2:A:118:ASN:HB2	2.26	0.50
1:M:22:PHE:N	1:M:70:PHE:CD2	2.81	0.49
1:M:60:TRP:CE2	5:H:117:ALA:HB2	2.48	0.48
5:H:115:GLN:HG2	5:H:125:ALA:HB1	1.95	0.47
5:H:123:TYR:CZ	5:H:140:ALA:HA	2.49	0.47
6:M:113:HOH:O	5:H:237:GLY:HA3	2.14	0.47
3:B:177:LYS:NZ	6:B:318:HOH:O	2.47	0.47
5:H:6:ARG:HH11	5:H:6:ARG:CG	2.27	0.47
5:H:106:ASP:CG	5:H:108:ARG:HG2	2.35	0.47
5:H:30:ASP:OD2	5:H:210:PRO:HB2	2.15	0.46
3:B:2:VAL:HG13	3:B:103:TYR:CD2	2.50	0.46
2:A:54:LYS:HG3	6:A:313:HOH:O	2.15	0.46
1:M:40:LEU:HD11	1:M:81:ARG:HB2	1.97	0.46
3:B:226:ARG:NH1	3:B:229:PRO:HG3	2.30	0.46
5:H:81:LEU:HD13	5:H:118:TYR:CD1	2.51	0.46
3:B:129:PRO:HD3	3:B:142:LEU:HG	1.99	0.45
5:H:4:SER:HB2	5:H:6:ARG:NH2	2.32	0.45
5:H:76:GLU:OE1	5:H:79:ARG:NH1	2.49	0.45
5:H:260:HIS:CE1	5:H:271:THR:HG23	2.52	0.44
5:H:49:ALA:O	5:H:52:ILE:HG22	2.18	0.44
1:M:49:VAL:HG22	1:M:68:THR:HB	2.00	0.44
3:B:95:GLN:HB2	3:B:98:GLU:HG3	1.99	0.44
2:A:31:THR:HG23	5:H:166:GLU:HG2	2.00	0.43
1:M:22:PHE:C	1:M:70:PHE:HD2	2.21	0.43
5:H:1:GLY:N	5:H:180:GLU:OE2	2.51	0.43
3:B:11:LEU:HD22	3:B:13:LYS:HE2	2.00	0.43
5:H:9:TYR:CE2	5:H:70:GLN:HG2	2.54	0.43
2:A:155:ASP:HB2	2:A:182:LYS:CE	2.50	0.42
5:H:231:VAL:O	5:H:243:LYS:NZ	2.40	0.41
2:A:155:ASP:HB2	2:A:182:LYS:HE3	2.02	0.41
1:M:48:LYS:HA	1:M:48:LYS:HD3	1.46	0.41
2:A:69:ASP:CG	2:A:71:LYS:HG2	2.40	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:H:183:ASP:HA	5:H:184:PRO:HD2	1.96	0.41
1:M:59:ASP:OD1	1:M:59:ASP:O	2.38	0.41
1:M:5:PRO:HA	1:M:30:PHE:HB3	2.03	0.41
5:H:229:GLU:HB3	5:H:246:ALA:HB3	2.04	0.40
5:H:197:HIS:O	5:H:251:SER:N	2.48	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	M	87/99 (88%)	83 (95%)	4 (5%)	0	100 100
2	A	202/207 (98%)	191 (95%)	10 (5%)	1 (0%)	29 22
3	B	241/244 (99%)	236 (98%)	5 (2%)	0	100 100
4	P	11/13 (85%)	11 (100%)	0	0	100 100
5	H	265/362 (73%)	256 (97%)	9 (3%)	0	100 100
All	All	806/925 (87%)	777 (96%)	28 (4%)	1 (0%)	51 48

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	79	ARG

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	M	83/94 (88%)	82 (99%)	1 (1%)	71 75
2	A	187/190 (98%)	185 (99%)	2 (1%)	73 77
3	B	216/218 (99%)	211 (98%)	5 (2%)	50 51
4	P	6/6 (100%)	5 (83%)	1 (17%)	2 1
5	H	219/297 (74%)	213 (97%)	6 (3%)	44 44
All	All	711/805 (88%)	696 (98%)	15 (2%)	53 55

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	M	97	ARG
2	A	127	ARG
2	A	135	SER
3	B	8	SER
3	B	167	SER
3	B	192	ARG
3	B	216	LEU
3	B	226	ARG
4	P	11	SER
5	H	11	SER
5	H	23	ILE
5	H	152	GLU
5	H	194	ILE
5	H	197	HIS
5	H	255	GLN

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

Mol	Chain	Res	Type
5	H	54	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

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

5.5 Carbohydrates

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	M	93/99 (93%)	0.59	10 (10%) 5 5	34, 64, 110, 143	0
2	A	202/207 (97%)	-0.10	3 (1%) 73 73	17, 32, 68, 127	0
3	B	242/244 (99%)	-0.24	1 (0%) 92 92	15, 25, 51, 81	0
4	P	13/13 (100%)	-0.08	0 100 100	20, 23, 31, 32	0
5	H	269/362 (74%)	0.61	38 (14%) 2 2	17, 50, 104, 172	0
All	All	819/925 (88%)	0.17	52 (6%) 20 19	15, 35, 94, 172	0

All (52) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	M	70	PHE	7.6
2	A	132	SER	5.7
5	H	1	GLY	5.3
5	H	272	LEU	4.7
5	H	248	VAL	4.2
5	H	270	LEU	3.9
5	H	194	ILE	3.9
3	B	2	VAL	3.9
1	M	78	TYR	3.8
5	H	273	ARG	3.8
5	H	252	GLY	3.8
5	H	271	THR	3.7
1	M	96	ASP	3.6
5	H	206	LEU	3.4
2	A	133	ASP	3.4
5	H	226	GLN	3.4
2	A	3	LEU	3.3
5	H	227	ASP	3.2
5	H	217	TRP	3.2
5	H	261	VAL	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	M	22	PHE	3.1
5	H	213	ILE	3.0
5	H	250	PRO	3.0
5	H	249	VAL	2.9
5	H	251	SER	2.9
5	H	225	THR	2.9
5	H	205	ALA	2.8
5	H	259	CYS	2.8
5	H	257	TYR	2.8
1	M	89	GLN	2.8
5	H	267	PRO	2.7
1	M	88	SER	2.7
1	M	95	TRP	2.7
5	H	216	THR	2.7
5	H	215	LEU	2.7
5	H	187	THR	2.6
5	H	192	HIS	2.5
5	H	218	GLN	2.4
5	H	181	ARG	2.3
5	H	269	PRO	2.3
1	M	73	THR	2.3
5	H	190	THR	2.3
5	H	260	HIS	2.3
5	H	253	GLU	2.2
5	H	274	TRP	2.2
5	H	255	GLN	2.1
5	H	138	THR	2.1
5	H	266	LEU	2.1
1	M	17	ASN	2.1
1	M	42	ASN	2.1
5	H	177	ASP	2.1
5	H	201	LEU	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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.