Full wwPDB X-ray Structure Validation Report ⓘ

Aug 22, 2023 – 06:05 PM EDT

PDB ID	:	7UVO
Title	:	Pfs230 domain 1 bound by RUPA-38 Fab
Authors	:	Ivanochko, D.; Newton, J.; Julien, J.P.
Deposited on	:	2022-05-02
Resolution	:	2.09 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

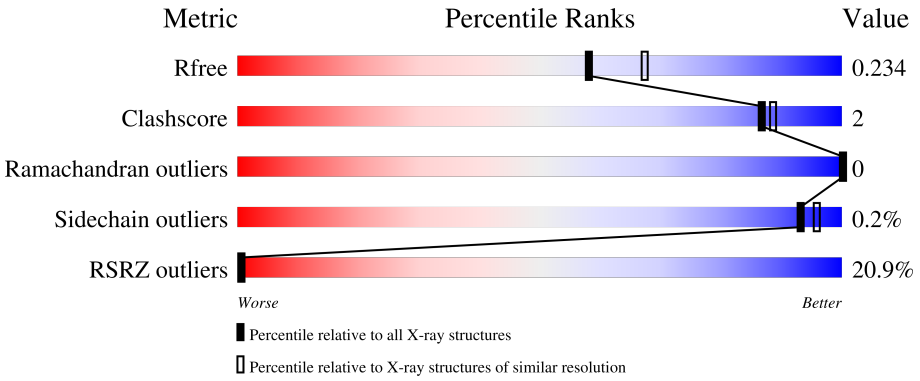
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 2.09 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	223	26% 96% .
2	B	218	27% 90% 8% .
3	C	199	6% 84% 7% 10%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 4891 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 RUPA-38 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	223	Total	C	N	O	S	0	0	0
			1698	1076	290	324	8			

- Molecule 2 is a protein called RUPA-38 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	215	Total	C	N	O	S	0	0	0
			1611	1008	268	327	8			

- Molecule 3 is a protein called Gametocyte surface protein P230.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	180	Total	C	N	O	S	0	0	0
			1434	918	222	290	4			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	585	GLN	ASN	conflict	UNP P68874
C	732	GLY	-	expression tag	UNP P68874
C	733	SER	-	expression tag	UNP P68874
C	734	LEU	-	expression tag	UNP P68874
C	735	LYS	-	expression tag	UNP P68874
C	736	GLU	-	expression tag	UNP P68874
C	737	ASN	-	expression tag	UNP P68874
C	738	LEU	-	expression tag	UNP P68874
C	739	TYR	-	expression tag	UNP P68874
C	740	PHE	-	expression tag	UNP P68874
C	741	GLN	-	expression tag	UNP P68874
C	742	GLY	-	expression tag	UNP P68874
C	743	TRP	-	expression tag	UNP P68874
C	744	SER	-	expression tag	UNP P68874

Continued on next page...

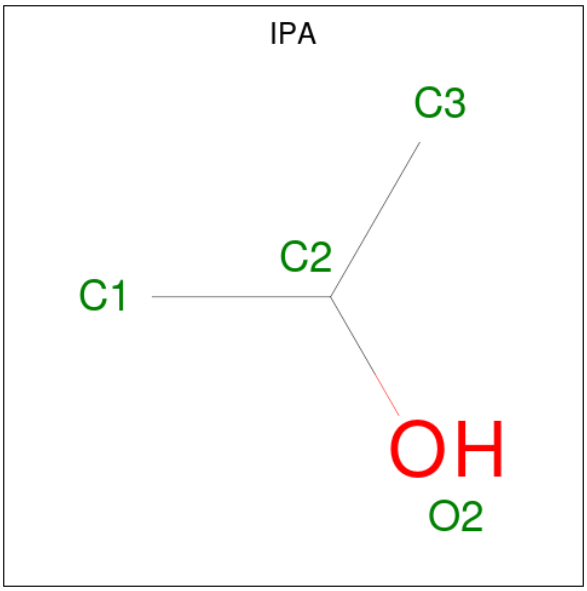
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	745	HIS	-	expression tag	UNP P68874
C	746	PRO	-	expression tag	UNP P68874
C	747	GLN	-	expression tag	UNP P68874
C	748	PHE	-	expression tag	UNP P68874
C	749	GLU	-	expression tag	UNP P68874
C	750	LYS	-	expression tag	UNP P68874

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Cl 1 1	0	0
4	B	1	Total Cl 1 1	0	0

- Molecule 5 is ISOPROPYL ALCOHOL (three-letter code: IPA) (formula: C_3H_8O).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total C O 4 3 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	42	Total O 42 42	0	0

Continued on next page...

Continued from previous page...

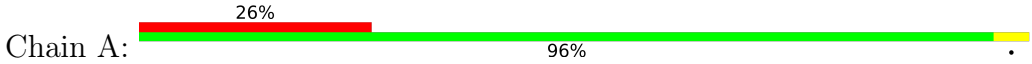
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	47	Total O 47 47	0	0
6	C	53	Total O 53 53	0	0

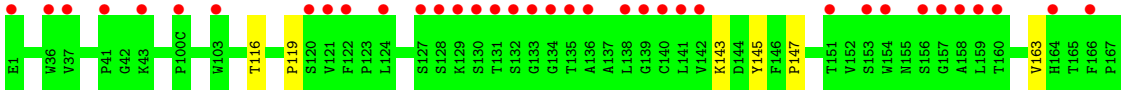
3 Residue-property plots (i)

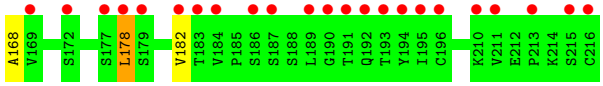
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: RUPA-38 Fab heavy chain

Chain A:

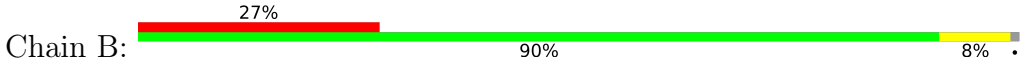


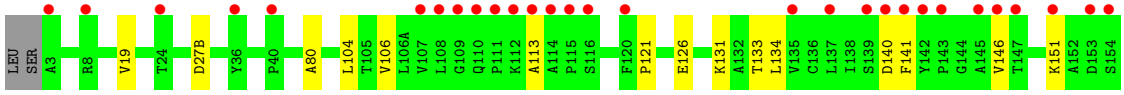


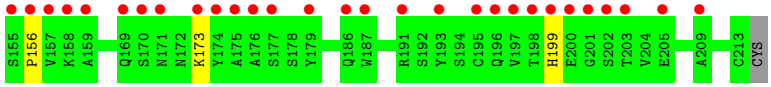


• Molecule 2: RUPA-38 Fab light chain

Chain B:

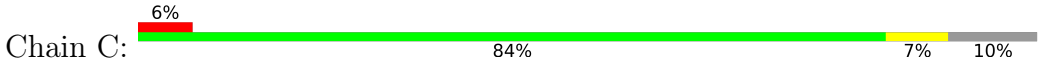


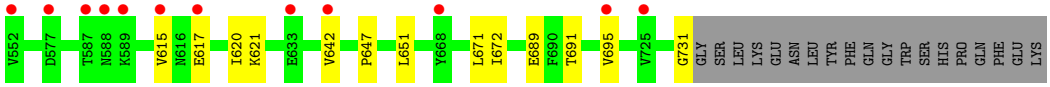




• Molecule 3: Gametocyte surface protein P230

Chain C:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	163.71Å 42.85Å 95.30Å 90.00° 102.49° 90.00°	Depositor
Resolution (Å)	29.06 – 2.09 29.06 – 2.09	Depositor EDS
% Data completeness (in resolution range)	99.8 (29.06-2.09) 99.8 (29.06-2.09)	Depositor EDS
R_{merge}	0.22	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.73 (at 2.10Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.217 , 0.235 0.217 , 0.234	Depositor DCC
R_{free} test set	1937 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	39.4	Xtriage
Anisotropy	0.194	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 64.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4891	wwPDB-VP
Average B, all atoms (Å^2)	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.53% 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

Bond lengths and bond angles in the following residue types are not validated in this section: CL, IPA

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	0.38	0/1743	0.62	1/2375 (0.0%)
2	B	0.39	0/1651	0.60	0/2254
3	C	0.41	0/1460	0.60	0/1975
All	All	0.39	0/4854	0.61	1/6604 (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})
1	A	178	LEU	CA-CB-CG	5.29	127.47	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	A	1698	0	1657	5	0
2	B	1611	0	1554	9	0
3	C	1434	0	1443	7	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	B	4	0	8	0	0
6	A	42	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	B	47	0	0	0	0
6	C	53	0	0	0	0
All	All	4891	0	4662	20	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 (20) 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:C:642:VAL:HG12	3:C:651:LEU:HD11	1.87	0.56
1:A:163:VAL:HG22	1:A:182:VAL:HG22	1.88	0.55
3:C:617:GLU:HB2	3:C:620:ILE:HD11	1.91	0.52
2:B:121:PRO:HA	2:B:134:LEU:HD23	1.92	0.52
3:C:615:VAL:HG13	3:C:620:ILE:HD12	1.92	0.51
1:A:119:PRO:HB3	1:A:145:TYR:HB3	1.93	0.51
2:B:151:LYS:HG2	2:B:156:PRO:HA	1.93	0.50
1:A:168:ALA:HA	1:A:178:LEU:HB3	1.95	0.48
2:B:80:ALA:HA	2:B:106:VAL:HG11	1.96	0.48
2:B:140:ASP:HA	2:B:173:LYS:HB3	1.96	0.47
3:C:695:VAL:HG12	3:C:731:GLY:HA3	1.96	0.47
3:C:671:LEU:HB3	3:C:691:THR:HB	1.97	0.46
1:A:116:THR:HG22	1:A:147:PRO:HD3	1.96	0.46
2:B:146:VAL:HG21	2:B:199:HIS:HB2	1.97	0.46
2:B:19:VAL:HG21	2:B:104:LEU:HD11	1.97	0.46
2:B:113:ALA:HA	2:B:141:PHE:CE1	2.51	0.45
3:C:647:PRO:HG2	3:C:672:ILE:HD13	2.00	0.43
1:A:143:LYS:HZ3	2:B:131:LYS:HD2	1.84	0.42
3:C:621:LYS:HE2	3:C:689:GLU:HB3	2.03	0.41
2:B:126:GLU:OE2	2:B:133:THR:OG1	2.28	0.41

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	221/223 (99%)	211 (96%)	10 (4%)	0	100	100
2	B	213/218 (98%)	203 (95%)	10 (5%)	0	100	100
3	C	178/199 (89%)	173 (97%)	5 (3%)	0	100	100
All	All	612/640 (96%)	587 (96%)	25 (4%)	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	190/190 (100%)	190 (100%)	0	100	100
2	B	181/184 (98%)	180 (99%)	1 (1%)	86	90
3	C	168/185 (91%)	168 (100%)	0	100	100
All	All	539/559 (96%)	538 (100%)	1 (0%)	93	96

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	27(B)	ASP

Sometimes 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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
5	IPA	B	301	-	3,3,3	0.44	0	3,3,3	0.93	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	223/223 (100%)	1.47	59 (26%)	0	0	23, 56, 141, 156	0
2	B	215/218 (98%)	1.42	58 (26%)	0	0	27, 65, 131, 164	0
3	C	180/199 (90%)	0.55	12 (6%)	17	22	29, 50, 95, 135	0
All	All	618/640 (96%)	1.18	129 (20%)	1	0	23, 54, 132, 164	0

All (129) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	130	SER	15.2
2	B	113	ALA	14.7
1	A	132	SER	11.3
2	B	112	LYS	9.7
2	B	142	TYR	9.6
1	A	157	GLY	9.2
2	B	200	GLU	8.5
1	A	159	LEU	8.4
1	A	160	THR	8.0
2	B	174	TYR	7.9
1	A	191	THR	7.8
2	B	146	VAL	7.6
1	A	182	VAL	7.3
1	A	131	THR	6.9
2	B	141	PHE	6.3
1	A	189	LEU	6.3
2	B	202	SER	6.1
1	A	133	GLY	5.9
2	B	114	ALA	5.7
2	B	170	SER	5.7
1	A	216	CYS	5.5
1	A	136	ALA	5.5
2	B	171	ASN	5.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	134	GLY	5.4
2	B	201	GLY	5.4
2	B	175	ALA	5.4
2	B	154	SER	5.3
1	A	190	GLY	5.3
2	B	197	VAL	5.3
2	B	198	THR	5.2
1	A	129	LYS	5.1
2	B	158	LYS	5.1
1	A	135	THR	5.1
3	C	588	ASN	5.0
1	A	192	GLN	5.0
2	B	199	HIS	5.0
2	B	109	GLY	5.0
1	A	141	LEU	4.8
2	B	155	SER	4.6
1	A	156	SER	4.5
2	B	203	THR	4.5
1	A	194	TYR	4.5
3	C	552	VAL	4.5
2	B	111	PRO	4.5
2	B	159	ALA	4.4
1	A	178	LEU	4.3
2	B	110	GLN	4.0
1	A	215	SER	4.0
2	B	135	VAL	3.9
1	A	183	THR	3.9
1	A	184	VAL	3.9
2	B	196	GLN	3.9
2	B	3	ALA	3.8
1	A	140	CYS	3.8
2	B	40	PRO	3.8
2	B	145	ALA	3.7
1	A	122	PHE	3.7
1	A	179	SER	3.7
2	B	193	TYR	3.6
2	B	169	GLN	3.5
1	A	158	ALA	3.5
2	B	147	THR	3.4
1	A	164	HIS	3.3
2	B	176	ALA	3.3
1	A	153	SER	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	154	TRP	3.2
2	B	115	PRO	3.2
1	A	128	SER	3.2
1	A	151	THR	3.2
1	A	41	PRO	3.1
1	A	142	VAL	3.1
1	A	211	VAL	3.1
2	B	143	PRO	3.0
1	A	43	LYS	3.0
2	B	173	LYS	3.0
2	B	156	PRO	3.0
2	B	151	LYS	3.0
2	B	137	LEU	3.0
1	A	210	LYS	2.9
1	A	121	VAL	2.9
2	B	191	ARG	2.9
1	A	172	SER	2.8
1	A	138	LEU	2.8
3	C	633	GLU	2.7
1	A	177	SER	2.7
2	B	120	PHE	2.6
2	B	195	CYS	2.6
3	C	725	VAL	2.5
1	A	169	VAL	2.5
1	A	1	GLU	2.5
2	B	108	LEU	2.5
2	B	140	ASP	2.5
3	C	617	GLU	2.5
2	B	116	SER	2.5
1	A	100(C)	PRO	2.4
1	A	37	VAL	2.4
1	A	196	CYS	2.4
3	C	668	TYR	2.4
1	A	120	SER	2.3
1	A	186	SER	2.3
1	A	187	SER	2.3
3	C	589	LYS	2.3
1	A	166	PHE	2.3
1	A	127	SER	2.3
2	B	107	VAL	2.3
2	B	36	TYR	2.3
2	B	186	GLN	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	195	ILE	2.3
2	B	177	SER	2.2
3	C	577	ASP	2.2
2	B	209	ALA	2.2
1	A	103	TRP	2.2
2	B	187	TRP	2.2
1	A	213	PRO	2.2
2	B	8	ARG	2.2
1	A	124	LEU	2.1
2	B	205	GLU	2.1
3	C	642	VAL	2.1
2	B	179	TYR	2.1
3	C	615	VAL	2.1
2	B	24	THR	2.1
3	C	587	THR	2.1
1	A	139	GLY	2.0
2	B	139	SER	2.0
2	B	153	ASP	2.0
2	B	157	VAL	2.0
3	C	695	VAL	2.0
1	A	36	TRP	2.0
1	A	193	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)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	IPA	B	301	4/4	0.75	0.34	20,20,20,20	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	CL	B	302	1/1	0.96	0.04	63,63,63,63	0
4	CL	A	301	1/1	1.00	0.12	31,31,31,31	0

6.5 Other polymers (i)

There are no such residues in this entry.