Full wwPDB X-ray Structure Validation Report ⓘ

Aug 22, 2023 – 06:04 PM EDT

PDB ID	:	7UVQ
Title	:	Pfs230 domain 1 bound by RUPA-97 and 15C5 Fabs
Authors	:	Ivanochko, D.; Julien, J.P.
Deposited on	:	2022-05-02
Resolution	:	3.29 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:

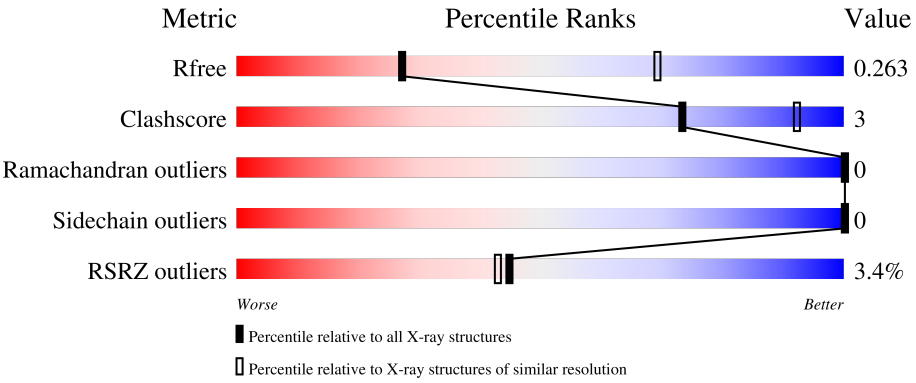
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 3.29 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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	199	78% 9% 13%
2	B	213	7% 88% 12%
3	C	221	5% 87% 13%
4	D	227	% 95%
5	E	215	3% 91% 8%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8062 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 Gametocyte surface protein P230.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	174	Total	C	N	O	S	0	0	0
			1384	886	215	279	4			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	585	GLN	ASN	conflict	UNP P68874
A	732	GLY	-	expression tag	UNP P68874
A	733	SER	-	expression tag	UNP P68874
A	734	LEU	-	expression tag	UNP P68874
A	735	LYS	-	expression tag	UNP P68874
A	736	GLU	-	expression tag	UNP P68874
A	737	ASN	-	expression tag	UNP P68874
A	738	LEU	-	expression tag	UNP P68874
A	739	TYR	-	expression tag	UNP P68874
A	740	PHE	-	expression tag	UNP P68874
A	741	GLN	-	expression tag	UNP P68874
A	742	GLY	-	expression tag	UNP P68874
A	743	TRP	-	expression tag	UNP P68874
A	744	SER	-	expression tag	UNP P68874
A	745	HIS	-	expression tag	UNP P68874
A	746	PRO	-	expression tag	UNP P68874
A	747	GLN	-	expression tag	UNP P68874
A	748	PHE	-	expression tag	UNP P68874
A	749	GLU	-	expression tag	UNP P68874
A	750	LYS	-	expression tag	UNP P68874

- Molecule 2 is a protein called Fab Kappa Light Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	213	Total	C	N	O	S	0	0	0
			1627	1021	266	331	9			

- Molecule 3 is a protein called Fab Heavy Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	221	Total	C	N	O	S	0	0	0
			1704	1081	278	339	6			

- Molecule 4 is a protein called Fab Heavy Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	227	Total	C	N	O	S	0	0	0
			1694	1069	283	335	7			

- Molecule 5 is a protein called Fab Kappa Light Chain.

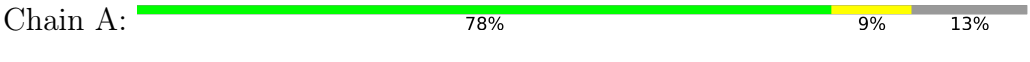
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	215	Total	C	N	O	S	0	0	0
			1653	1036	281	331	5			

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.

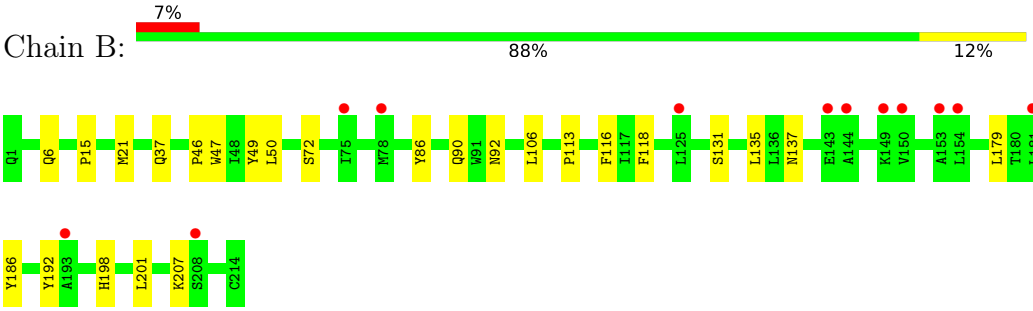
• Molecule 1: Gametocyte surface protein P230

Chain A:



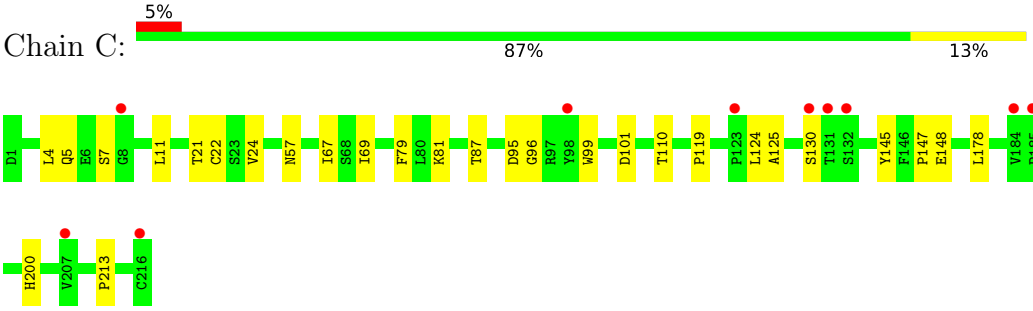
• Molecule 2: Fab Kappa Light Chain

Chain B:



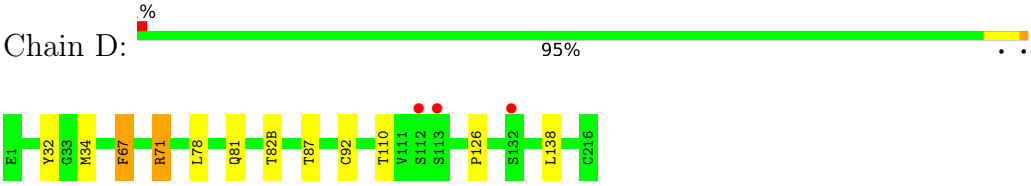
• Molecule 3: Fab Heavy Chain

Chain C:



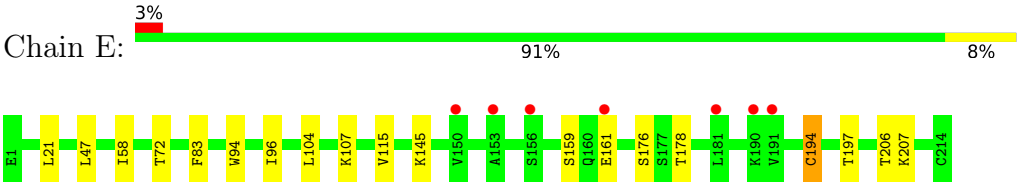
• Molecule 4: Fab Heavy Chain

Chain D:



• Molecule 5: Fab Kappa Light Chain

Chain E:


3%
91%
8%
E1
L21
L47
I58
T72
F83
W94
I96
L104
K107
V115
K145
V150
A153
S156
S159
Q160
E161
S176
S177
T178
L181
K190
V191
C194
T197
T206
K207
C214

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 62	Depositor
Cell constants a, b, c, α, β, γ	155.00Å 155.00Å 92.90Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.76 – 3.29 29.76 – 3.29	Depositor EDS
% Data completeness (in resolution range)	100.0 (29.76-3.29) 100.0 (29.76-3.29)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.87 (at 3.31Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.215 , 0.264 0.215 , 0.263	Depositor DCC
R_{free} test set	988 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	128.5	Xtriage
Anisotropy	0.132	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 80.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.031 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	8062	wwPDB-VP
Average B, all atoms (Å^2)	150.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.76% 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 (i)

5.1 Standard geometry (i)

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	0.82	1/1410 (0.1%)	0.94	0/1907
2	B	0.53	0/1667	0.72	0/2268
3	C	0.62	0/1752	0.90	3/2395 (0.1%)
4	D	0.68	1/1737 (0.1%)	0.88	4/2365 (0.2%)
5	E	0.60	0/1689	0.82	1/2297 (0.0%)
All	All	0.65	2/8255 (0.0%)	0.85	8/11232 (0.1%)

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	A	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	92	CYS	CB-SG	-6.62	1.71	1.82
1	A	706	CYS	CB-SG	-5.75	1.72	1.81

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	D	71	ARG	CG-CD-NE	8.17	128.96	111.80
4	D	67	PHE	CB-CG-CD1	6.36	125.25	120.80
4	D	67	PHE	CB-CG-CD2	-5.65	116.85	120.80
3	C	95	ASP	CB-CG-OD1	5.64	123.38	118.30
5	E	194	CYS	CA-CB-SG	5.51	123.92	114.00
3	C	95	ASP	CB-CG-OD2	-5.39	113.45	118.30
4	D	82(B)	THR	CA-CB-CG2	5.37	119.91	112.40
3	C	178	LEU	CA-CB-CG	5.12	127.07	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	731	GLY	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	1384	0	1392	10	0
2	B	1627	0	1567	14	0
3	C	1704	0	1631	17	0
4	D	1694	0	1653	5	0
5	E	1653	0	1620	10	0
All	All	8062	0	7863	52	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 3.

All (52) 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)
4:D:34:MET:HB3	4:D:78:LEU:HD22	1.88	0.56
2:B:116:PHE:HB2	2:B:135:LEU:HD23	1.87	0.56
4:D:32:TYR:O	4:D:71:ARG:NH2	2.39	0.55
3:C:67:ILE:HA	3:C:81:LYS:O	2.06	0.55
5:E:94:TRP:HB2	5:E:96:ILE:HG22	1.89	0.53
1:A:608:VAL:HG22	1:A:720:ARG:HB3	1.92	0.53
3:C:7:SER:HB3	3:C:21:THR:HB	1.91	0.51
3:C:96:GLY:HA2	3:C:101:ASP:HB2	1.91	0.51
5:E:194:CYS:O	5:E:206:THR:HA	2.11	0.51
1:A:701:VAL:HB	1:A:726:TYR:HD1	1.75	0.51
4:D:87:THR:HG23	4:D:110:THR:HA	1.93	0.51
1:A:604:SER:HA	1:A:716:LYS:HA	1.92	0.51
3:C:11:LEU:HD12	3:C:147:PRO:HD3	1.94	0.50
2:B:131:SER:HA	2:B:179:LEU:O	2.13	0.49
5:E:159:SER:HA	5:E:178:THR:O	2.12	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:571:ASP:N	1:A:571:ASP:OD1	2.46	0.48
3:C:125:ALA:HB1	3:C:213:PRO:HA	1.94	0.48
1:A:587:THR:OG1	1:A:588:ASN:N	2.47	0.48
3:C:119:PRO:HB3	3:C:145:TYR:HB3	1.94	0.48
3:C:5:GLN:O	3:C:22:CYS:HA	2.13	0.47
2:B:15:PRO:HG3	2:B:106:LEU:HD11	1.95	0.47
1:A:701:VAL:HA	1:A:725:VAL:O	2.15	0.46
3:C:87:THR:HG23	3:C:110:THR:HA	1.98	0.46
5:E:21:LEU:O	5:E:72:THR:HA	2.15	0.46
5:E:115:VAL:O	5:E:207:LYS:NZ	2.47	0.46
2:B:118:PHE:HB3	3:C:124:LEU:HD22	1.97	0.46
2:B:90:GLN:OE1	2:B:92:ASN:N	2.46	0.46
3:C:4:LEU:HD23	3:C:24:VAL:HG22	1.98	0.46
2:B:21:MET:O	2:B:72:SER:HA	2.16	0.46
2:B:46:PRO:HG3	3:C:101:ASP:HA	1.97	0.46
3:C:69:ILE:HA	3:C:79:PHE:O	2.15	0.45
4:D:126:PRO:HB3	4:D:138:LEU:HB3	1.98	0.45
1:A:629:LYS:HE3	1:A:629:LYS:HB2	1.85	0.44
2:B:37:GLN:HB3	2:B:47:TRP:NE1	2.32	0.44
5:E:47:LEU:HA	5:E:58:ILE:HG13	2.00	0.44
2:B:6:GLN:NE2	2:B:86:TYR:O	2.51	0.44
2:B:198:HIS:HB3	2:B:201:LEU:HG	2.00	0.44
5:E:161:GLU:HA	5:E:176:SER:O	2.18	0.43
5:E:145:LYS:HB3	5:E:197:THR:HB	2.01	0.43
2:B:207:LYS:NZ	3:C:130:SER:O	2.46	0.43
5:E:83:PHE:HD1	5:E:104:LEU:HB3	1.83	0.43
2:B:113:PRO:HA	2:B:137:ASN:O	2.19	0.43
2:B:49:TYR:CD2	2:B:50:LEU:HD22	2.54	0.42
3:C:119:PRO:HD3	3:C:200:HIS:ND1	2.35	0.42
3:C:57:ASN:HD22	3:C:57:ASN:HA	1.56	0.42
4:D:67:PHE:HA	4:D:81:GLN:O	2.21	0.41
2:B:186:TYR:O	2:B:192:TYR:OH	2.39	0.41
5:E:107:LYS:HE2	5:E:107:LYS:HB2	1.84	0.41
1:A:705:ILE:HG12	1:A:722:ILE:HG23	2.02	0.40
3:C:145:TYR:HE2	3:C:148:GLU:HA	1.86	0.40
1:A:669:GLY:HA3	3:C:99:TRP:CZ2	2.55	0.40
1:A:703:TYR:HA	1:A:723:VAL:O	2.21	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	172/199 (86%)	164 (95%)	8 (5%)	0	100	100
2	B	211/213 (99%)	207 (98%)	4 (2%)	0	100	100
3	C	219/221 (99%)	207 (94%)	12 (6%)	0	100	100
4	D	225/227 (99%)	219 (97%)	6 (3%)	0	100	100
5	E	213/215 (99%)	206 (97%)	7 (3%)	0	100	100
All	All	1040/1075 (97%)	1003 (96%)	37 (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	162/185 (88%)	162 (100%)	0	100	100
2	B	186/186 (100%)	186 (100%)	0	100	100
3	C	195/195 (100%)	195 (100%)	0	100	100
4	D	189/189 (100%)	189 (100%)	0	100	100
5	E	188/188 (100%)	188 (100%)	0	100	100
All	All	920/943 (98%)	920 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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

Mol	Chain	Res	Type
5	E	53	ASN

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 monosaccharides 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	174/199 (87%)	-0.25	0 100 100	85, 109, 149, 172	0
2	B	213/213 (100%)	0.40	14 (6%) 18 18	106, 184, 248, 266	0
3	C	221/221 (100%)	0.30	12 (5%) 25 24	98, 165, 242, 255	0
4	D	227/227 (100%)	-0.08	3 (1%) 77 77	88, 140, 198, 249	0
5	E	215/215 (100%)	-0.01	7 (3%) 46 44	90, 149, 207, 225	0
All	All	1050/1075 (97%)	0.08	36 (3%) 45 43	85, 144, 234, 266	0

All (36) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	149	LYS	5.0
2	B	153	ALA	4.9
2	B	150	VAL	4.7
2	B	193	ALA	4.5
4	D	132	SER	4.5
3	C	130	SER	4.4
2	B	182	SER	4.0
3	C	131	THR	3.5
2	B	184	ALA	3.3
2	B	144	ALA	3.1
4	D	113	SER	3.1
2	B	154	LEU	3.0
3	C	132	SER	2.9
3	C	184	VAL	2.9
5	E	156	SER	2.8
2	B	78	MET	2.8
3	C	193	THR	2.6
2	B	181	LEU	2.6
3	C	8	GLY	2.5
3	C	123	PRO	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	143	GLU	2.4
3	C	207	VAL	2.4
5	E	153	ALA	2.3
5	E	190	LYS	2.2
3	C	98	TYR	2.2
5	E	161	GLU	2.2
4	D	112	SER	2.2
2	B	125	LEU	2.2
5	E	181	LEU	2.2
2	B	208	SER	2.2
5	E	150	VAL	2.1
5	E	191	VAL	2.1
3	C	185	PRO	2.1
3	C	216	CYS	2.1
3	C	188	SER	2.0
2	B	75	ILE	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.