Full wwPDB X-ray Structure Validation Report (i)

Apr 27, 2024 – 03:13 pm BST

PDB ID	:	5DFZ
Title	:	Structure of Vps34 complex II from S. cerevisiae.
Authors	:	Rostislavleva, K.; Soler, N.; Ohashi, Y.; Zhang, L.; Williams, R.L.
Deposited on	:	2015-08-27
Resolution	:	4.40 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 (i) 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 (1)) were used in the production of this report:

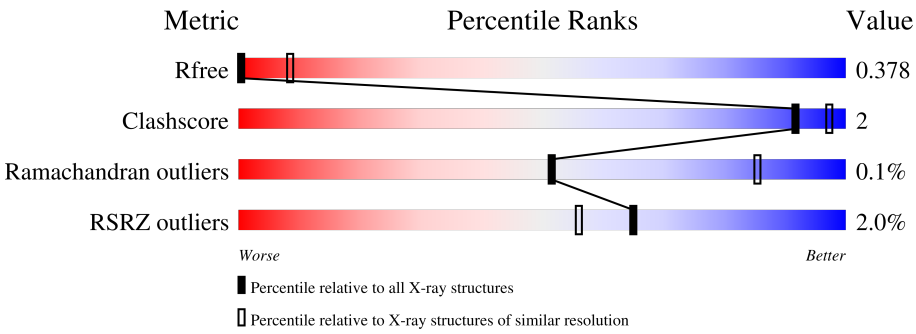
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 4.40 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	1043 (5.00-3.80)
Clashscore	141614	1111 (5.00-3.80)
Ramachandran outliers	138981	1059 (5.00-3.80)
RSRZ outliers	127900	1095 (5.08-3.70)

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	441	5%, 76%, •, 22%
2	C	875	%, 86%, •, 11%
3	B	1460	%, 82%, •, 16%
4	E	124	4%, 90%, •, 9%
5	D	559	61%, 39%
6	G	49	96%, •

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 14160 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 Vacuolar protein sorting-associated protein 38.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	A	343	1700	1014	343	343	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	440	GLY	-	expression tag	UNP Q05919
A	441	SER	-	expression tag	UNP Q05919

- Molecule 2 is a protein called Phosphatidylinositol 3-kinase VPS34.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	C	783	3888	2322	783	783	30	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	2	ALA	SER	conflict	UNP P22543
C	241	GLY	ASP	conflict	UNP P22543

- Molecule 3 is a protein called Serine/threonine-protein kinase VPS15.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	B	1224	6079	3631	1224	1224	0	0	0

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	SER	deletion	UNP P22219

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	MET	deletion	UNP P22219
B	?	-	ASP	deletion	UNP P22219
B	?	-	ASN	deletion	UNP P22219
B	?	-	ASN	deletion	UNP P22219
B	?	-	THR	deletion	UNP P22219
B	?	-	GLU	deletion	UNP P22219
B	?	-	ILE	deletion	UNP P22219
B	?	-	MET	deletion	UNP P22219
B	?	-	GLU	deletion	UNP P22219
B	1455	GLY	-	expression tag	UNP P22219
B	1456	SER	-	expression tag	UNP P22219
B	1457	SER	-	expression tag	UNP P22219
B	1458	ARG	-	expression tag	UNP P22219
B	1459	PRO	-	expression tag	UNP P22219
B	1460	THR	-	expression tag	UNP P22219
B	1461	THR	-	expression tag	UNP P22219
B	1462	ALA	-	expression tag	UNP P22219
B	1463	SER	-	expression tag	UNP P22219
B	1464	GLU	-	expression tag	UNP P22219
B	1465	ASN	-	expression tag	UNP P22219
B	1466	LEU	-	expression tag	UNP P22219
B	1467	TYR	-	expression tag	UNP P22219
B	1468	PHE	-	expression tag	UNP P22219
B	1469	GLN	-	expression tag	UNP P22219
B	1470	GLY	-	expression tag	UNP P22219

- Molecule 4 is a protein called Nanobody binding S. cerevisiae Vps34.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	E	113	552	326	113	113	0	0	0

- Molecule 5 is a protein called Vacuolar protein sorting-associated protein 30.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
5	D	341	1696	1014	341	341	0	0	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	77	ILE	LYS	conflict	UNP Q02948

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	467	ALA	THR	conflict	UNP Q02948
D	549	HIS	TYR	conflict	UNP Q02948
D	558	GLY	-	expression tag	UNP Q02948
D	559	SER	-	expression tag	UNP Q02948

- Molecule 6 is a protein called Putative N-terminal domain of S. cerevisiae Vps30.

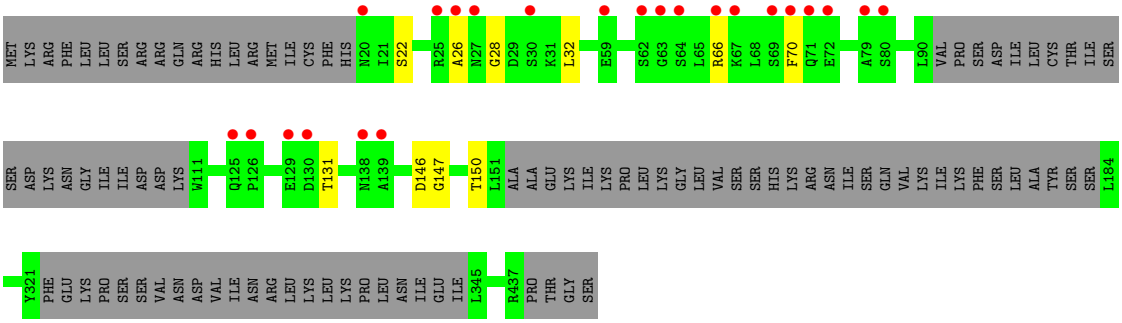
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
6	G	49	245	147	49	49	0	0	0

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.

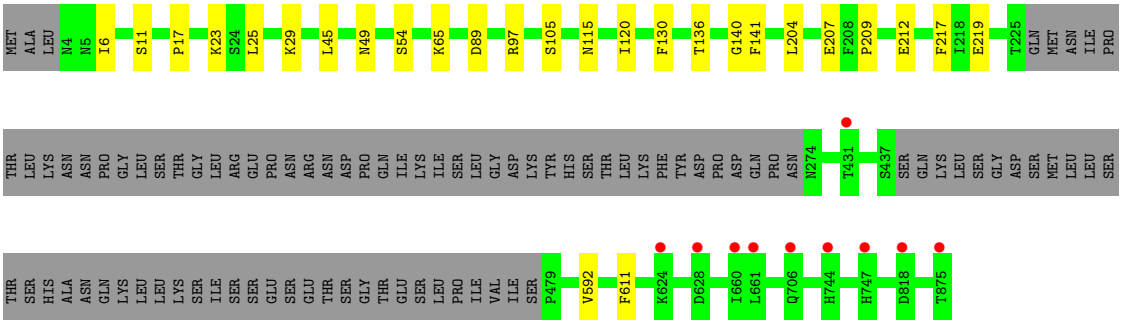
- Molecule 1: Vacuolar protein sorting-associated protein 38

Chain A:



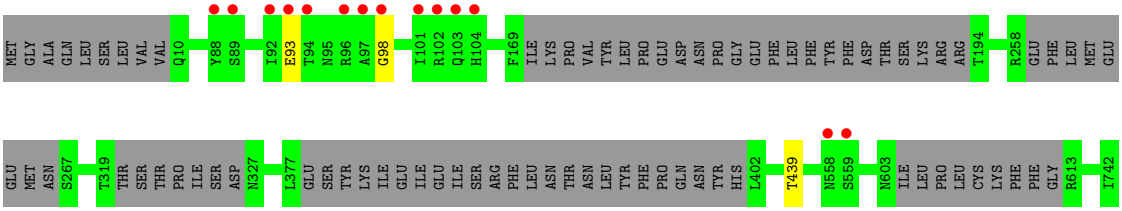
- Molecule 2: Phosphatidylinositol 3-kinase VPS34

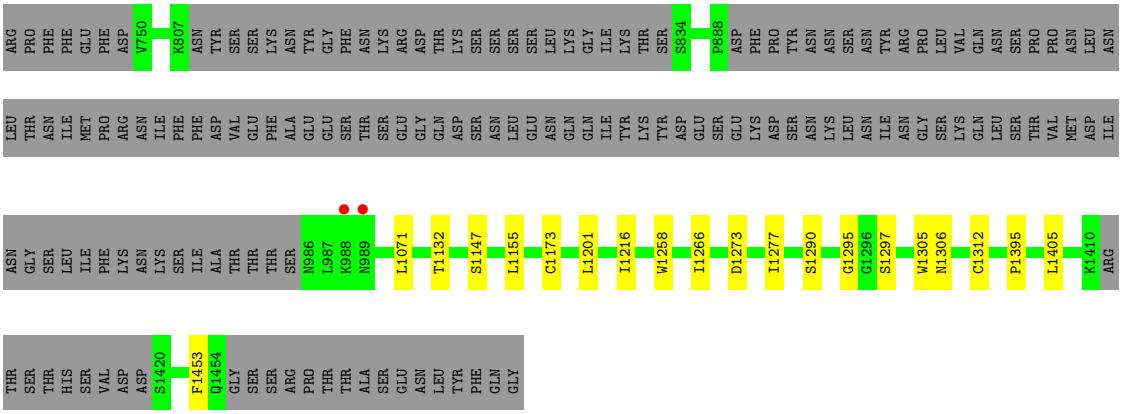
Chain C:



- Molecule 3: Serine/threonine-protein kinase VPS15

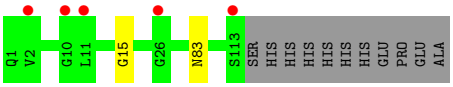
Chain B:





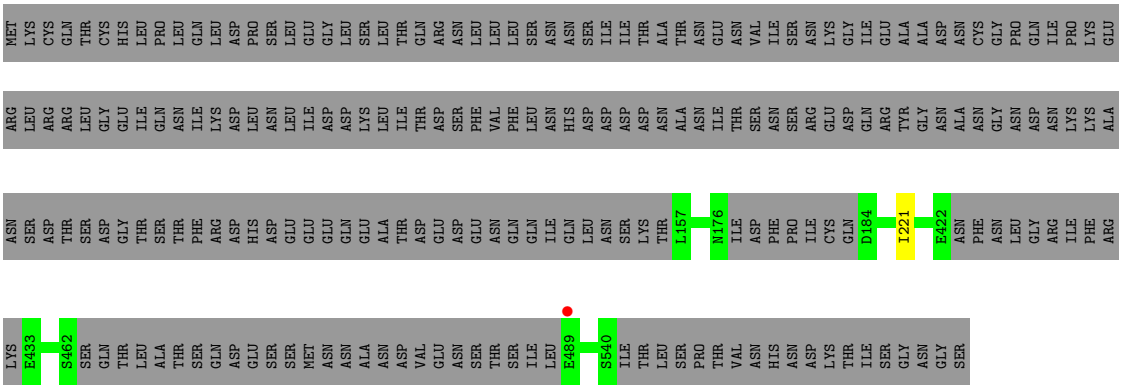
- Molecule 4: Nanobody binding S. cerevisiae Vps34

Chain E:



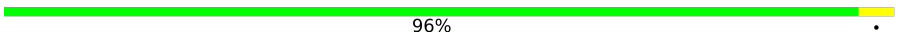
- Molecule 5: Vacuolar protein sorting-associated protein 30

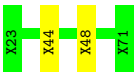
Chain D:



- Molecule 6: Putative N-terminal domain of S. cerevisiae Vps30

Chain G:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	215.18Å 226.84Å 127.21Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.36 – 4.40 50.36 – 4.31	Depositor EDS
% Data completeness (in resolution range)	100.0 (50.36-4.40) 98.2 (50.36-4.31)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.31	Depositor
$< I/\sigma(I) >$ [1]	1.03 (at 4.29Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.368 , 0.376 0.371 , 0.378	Depositor DCC
R_{free} test set	2043 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å^2)	254.5	Xtriage
Anisotropy	0.299	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.14 , 607.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.42$, $< L^2 > = 0.24$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	14160	wwPDB-VP
Average B, all atoms (Å^2)	258.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.19% 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	0.18	0/1696	0.37	0/2359
2	C	0.20	0/3885	0.47	0/5418
3	B	0.18	0/6069	0.41	0/8456
4	E	0.18	0/551	0.37	0/762
5	D	0.19	0/1692	0.39	0/2357
All	All	0.19	0/13893	0.42	0/19352

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	1700	0	720	5	0
2	C	3888	0	1633	14	0
3	B	6079	0	2576	12	0
4	E	552	0	253	1	0
5	D	1696	0	715	0	0
6	G	245	0	53	1	0
All	All	14160	0	5950	32	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 (32) 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:C:130:PHE:HA	2:C:136:THR:HA	1.68	0.74
2:C:6:ILE:HA	2:C:212:GLU:HA	1.68	0.73
2:C:65:LYS:H	2:C:105:SER:HA	1.60	0.66
2:C:23:LYS:HA	2:C:89:ASP:H	1.60	0.64
2:C:140:GLY:HA2	2:C:209:PRO:HA	1.83	0.61
2:C:115:ASN:H	2:C:120:ILE:HA	1.67	0.59
2:C:54:SER:HA	2:C:115:ASN:HA	1.86	0.57
2:C:592:VAL:HA	2:C:611:PHE:HA	1.87	0.56
2:C:25:LEU:HA	2:C:204:LEU:HA	1.88	0.55
3:B:1277:ILE:HA	3:B:1295:GLY:HA3	1.89	0.55
3:B:1290:SER:HA	3:B:1306:ASN:HA	1.88	0.55
3:B:1071:LEU:HA	3:B:1453:PHE:HA	1.89	0.54
3:B:1132:THR:H	3:B:1147:SER:HA	1.73	0.54
3:B:1258:TRP:HA	3:B:1266:ILE:H	1.72	0.53
4:E:15:GLY:HA2	4:E:83:ASN:HA	1.91	0.53
3:B:1305:TRP:HA	3:B:1312:CYS:HA	1.90	0.53
1:A:32:LEU:HA	1:A:66:ARG:HA	1.91	0.52
2:C:45:LEU:O	2:C:49:ASN:N	2.33	0.52
1:A:131:THR:HA	1:A:147:GLY:HA2	1.93	0.51
3:B:1155:LEU:HA	3:B:1173:CYS:HA	1.94	0.49
3:B:1395:PRO:HA	3:B:1405:LEU:HA	1.96	0.47
1:A:28:GLY:HA3	1:A:70:PHE:HA	1.96	0.47
1:A:22:SER:HA	1:A:150:THR:HA	1.98	0.45
2:C:17:PRO:HA	2:C:97:ARG:HA	1.98	0.45
2:C:11:SER:CB	2:C:217:PHE:H	2.30	0.45
2:C:219:GLU:H	3:B:439:THR:CB	2.30	0.44
2:C:141:PHE:HA	2:C:207:GLU:HA	2.00	0.44
1:A:26:ALA:HA	1:A:146:ASP:HA	1.99	0.43
3:B:1201:LEU:HA	3:B:1216:ILE:H	1.84	0.42
3:B:1273:ASP:H	3:B:1297:SER:HA	1.86	0.41
6:G:44:UNK:HA	6:G:48:UNK:HA	2.03	0.41
3:B:93:GLU:HA	3:B:98:GLY:HA2	2.03	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	335/441 (76%)	311 (93%)	24 (7%)	0	100	100
2	C	777/875 (89%)	720 (93%)	56 (7%)	1 (0%)	51	85
3	B	1204/1460 (82%)	1109 (92%)	95 (8%)	0	100	100
4	E	111/124 (90%)	107 (96%)	4 (4%)	0	100	100
5	D	333/559 (60%)	317 (95%)	15 (4%)	1 (0%)	41	76
All	All	2760/3459 (80%)	2564 (93%)	194 (7%)	2 (0%)	51	85

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	29	LYS
5	D	221	ILE

5.3.2 Protein sidechains (i)

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

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 (i)

6.1 Protein, DNA and RNA chains (i)

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, 95th 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	343/441 (77%)	-0.37	23 (6%)	17	14	144, 275, 387, 469	0
2	C	777/875 (88%)	-0.70	10 (1%)	77	68	19, 259, 386, 550	14 (1%)
3	B	1224/1460 (83%)	-0.75	16 (1%)	77	68	135, 227, 330, 441	0
4	E	113/124 (91%)	-0.29	5 (4%)	34	28	210, 305, 385, 407	0
5	D	341/559 (61%)	-0.80	1 (0%)	94	90	143, 247, 439, 499	0
6	G	0/49	-	-			-	-
All	All	2798/3508 (79%)	-0.68	55 (1%)	65	56	19, 247, 380, 550	14 (0%)

All (55) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	71	GLN	11.5
1	A	70	PHE	8.5
3	B	89	SER	8.1
1	A	139	ALA	7.3
1	A	20	ASN	7.2
1	A	63	GLY	6.8
3	B	97	ALA	6.2
3	B	93	GLU	5.9
1	A	138	ASN	5.6
1	A	72	GLU	4.9
1	A	66	ARG	4.9
1	A	126	PRO	4.8
3	B	558	ASN	4.7
3	B	94	THR	4.5
3	B	103	GLN	4.5
3	B	92	ILE	4.4
1	A	26	ALA	4.2
3	B	104	HIS	4.1
2	C	660	ILE	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	64	SER	3.8
1	A	129	GLU	3.7
3	B	989	ASN	3.5
1	A	62	SER	3.3
3	B	559	SER	3.3
4	E	2	VAL	3.3
1	A	79	ALA	3.3
2	C	818	ASP	3.1
5	D	489	GLU	3.0
1	A	80	SER	2.8
2	C	628	ASP	2.8
2	C	431	THR	2.7
2	C	875	THR	2.7
2	C	706	GLN	2.7
3	B	102	ARG	2.7
4	E	113	SER	2.5
3	B	88	TYR	2.5
1	A	27	ASN	2.4
1	A	67	LYS	2.4
2	C	661	LEU	2.3
3	B	98	GLY	2.3
1	A	69	SER	2.3
3	B	96	ARG	2.2
2	C	744	HIS	2.2
1	A	25	ARG	2.2
2	C	624	LYS	2.2
1	A	125	GLN	2.2
3	B	101	ILE	2.2
4	E	11	LEU	2.2
4	E	10	GLY	2.2
1	A	59	GLU	2.2
3	B	988	LYS	2.2
2	C	747	HIS	2.1
4	E	26	GLY	2.1
1	A	30	SER	2.1
1	A	130	ASP	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.