



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 8, 2020 – 12:08 pm GMT

PDB ID : 6TJP
Title : Crystal structure of T7 bacteriophage portal protein, 13mer, closed valve - P212121
Authors : Fabrega-Ferrer, M.; Cuervo, A.; Fernandez, F.J.; Machon, C.; Perez-Luque, R.; Pous, J.; Vega, M.C.; Carrascosa, J.L.; Coll, M.
Deposited on : 2019-11-26
Resolution : 3.74 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 : 2.15.1
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) : 2.15.1

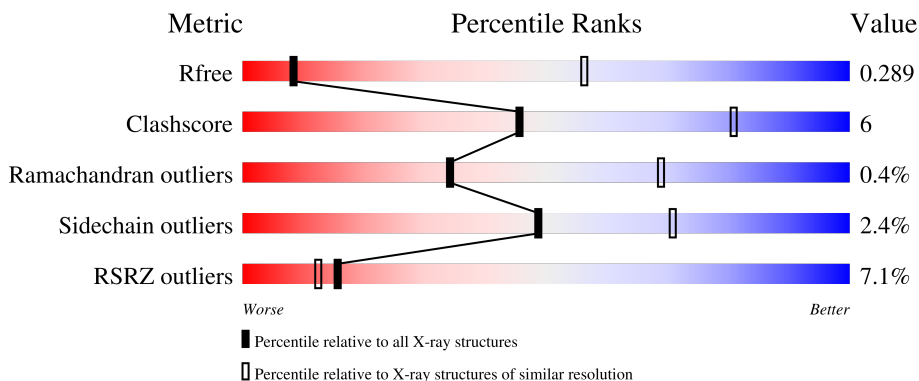
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.74 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	1001 (3.90-3.58)
Clashscore	141614	1063 (3.90-3.58)
Ramachandran outliers	138981	1027 (3.90-3.58)
Sidechain outliers	138945	1023 (3.90-3.58)
RSRZ outliers	127900	1006 (3.92-3.56)

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 $\leq 5\%$. 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	536	 6% 75% 13% • 12%
1	B	536	 5% 74% 14% 12%
1	C	536	 9% 72% 16% • 12%
1	D	536	 7% 73% 15% 12%
1	E	536	 9% 73% 15% • 12%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	536	<p>6% 70% 18% • 12%</p>
1	G	536	<p>5% 72% 16% • 12%</p>
1	H	536	<p>4% 71% 17% 12%</p>
1	I	536	<p>5% 71% 17% 12%</p>
1	J	536	<p>8% 70% 18% 12%</p>
1	K	536	<p>6% 75% 13% 12%</p>
1	L	536	<p>6% 73% 15% • 12%</p>
1	M	536	<p>5% 73% 15% 12%</p>

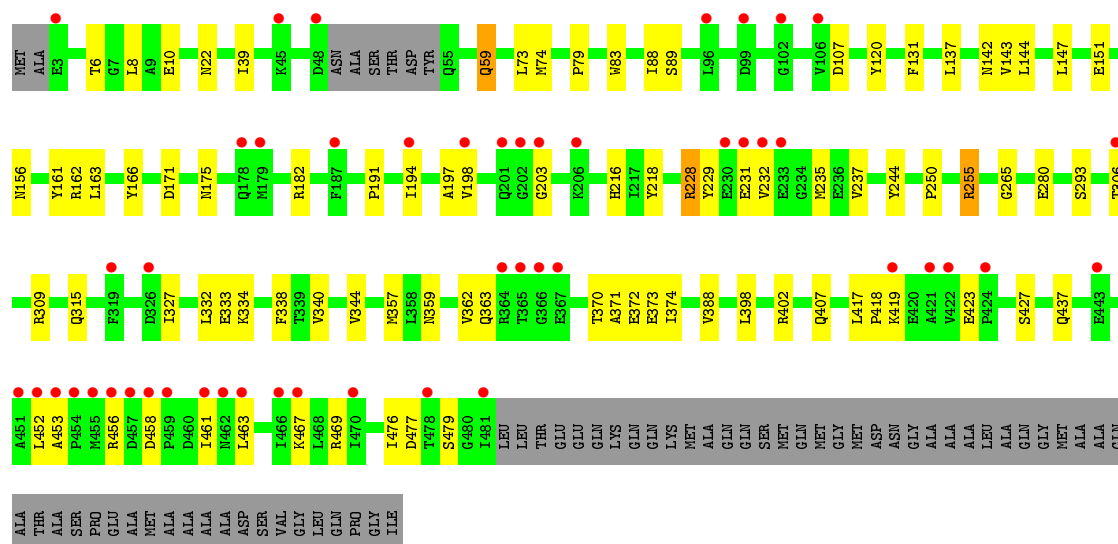
2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 48126 atoms, of which 0 are hydrogens and 0 are deuteriums.

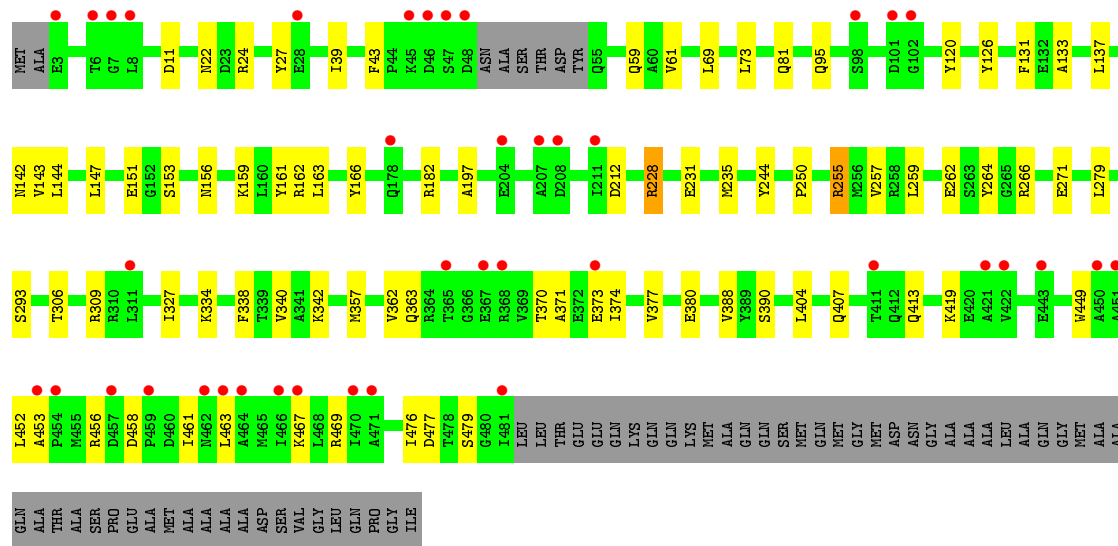
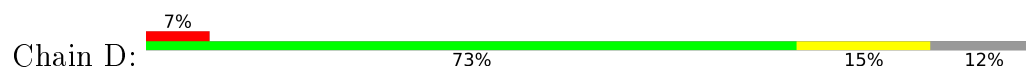
In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Portal protein.

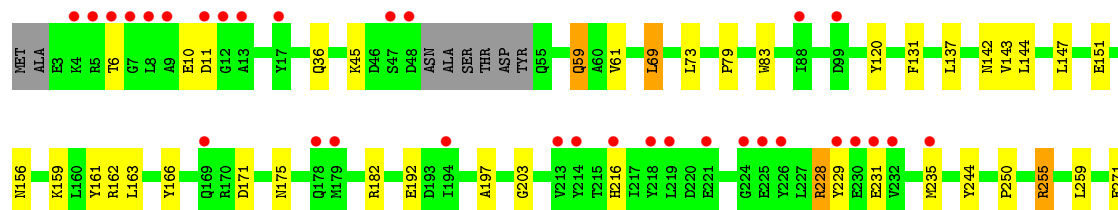
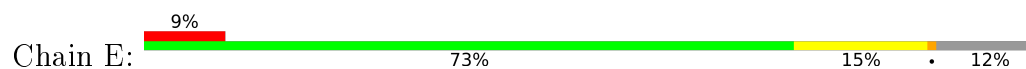
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	473	3702	2331	627	727	17	2	0	0
1	B	473	3702	2331	627	727	17	1	0	0
1	C	473	3702	2331	627	727	17	2	0	0
1	D	473	3702	2331	627	727	17	2	0	0
1	E	473	3702	2331	627	727	17	2	0	0
1	F	473	3702	2331	627	727	17	1	0	0
1	G	473	3702	2331	627	727	17	2	0	0
1	H	473	3702	2331	627	727	17	3	0	0
1	I	473	3702	2331	627	727	17	1	0	0
1	J	473	3702	2331	627	727	17	3	0	0
1	K	473	3702	2331	627	727	17	1	0	0
1	L	473	3702	2331	627	727	17	2	0	0
1	M	473	3702	2331	627	727	17	2	0	0

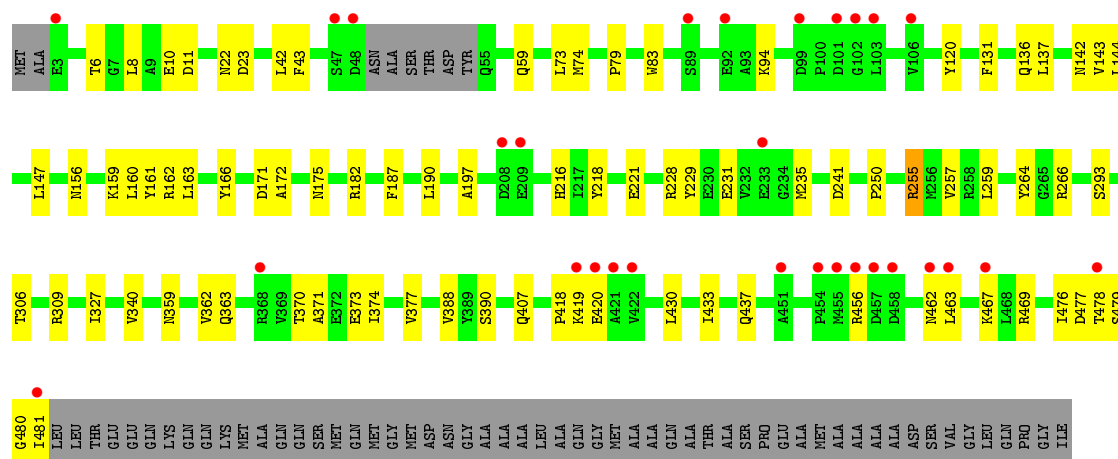


- Molecule 1: Portal protein



- Molecule 1: Portal protein





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	119.85Å 238.57Å 265.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.47 – 3.74 49.47 – 3.74	Depositor EDS
% Data completeness (in resolution range)	98.5 (49.47-3.74) 98.5 (49.47-3.74)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.58 (at 3.77Å)	Xtrriage
Refinement program	REFMAC 5.5, PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.244 , 0.286 0.257 , 0.289	Depositor DCC
R_{free} test set	3918 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	130.3	Xtrriage
Anisotropy	0.431	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 108.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.42$, $\langle L^2 \rangle = 0.25$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	48126	wwPDB-VP
Average B, all atoms (Å ²)	151.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.75% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.26	0/3760	0.46	0/5089
1	B	0.26	0/3760	0.46	0/5089
1	C	0.26	0/3760	0.47	0/5089
1	D	0.26	0/3760	0.46	0/5089
1	E	0.26	0/3760	0.47	0/5089
1	F	0.27	0/3760	0.47	0/5089
1	G	0.27	0/3760	0.47	0/5089
1	H	0.26	0/3760	0.46	0/5089
1	I	0.26	0/3760	0.47	0/5089
1	J	0.27	0/3760	0.47	0/5089
1	K	0.27	0/3760	0.47	0/5089
1	L	0.26	0/3760	0.47	0/5089
1	M	0.26	0/3760	0.47	0/5089
All	All	0.26	0/48880	0.47	0/66157

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	3702	0	3710	44	0
1	B	3702	0	3710	48	2
1	C	3702	0	3710	50	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	3702	0	3710	52	1
1	E	3702	0	3710	54	0
1	F	3702	0	3710	67	0
1	G	3702	0	3710	63	0
1	H	3702	0	3710	61	0
1	I	3702	0	3710	60	0
1	J	3702	0	3710	60	0
1	K	3702	0	3710	55	0
1	L	3702	0	3710	53	0
1	M	3702	0	3710	55	1
All	All	48126	0	48230	620	2

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.

The worst 5 of 620 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)
1:F:144:LEU:HB3	1:F:161:TYR:HB2	1.60	0.83
1:C:306:THR:HG21	1:C:327:ILE:HD11	1.66	0.78
1:C:479:SER:O	1:D:469:ARG:NH2	2.16	0.78
1:I:162:ARG:HH22	1:J:259:LEU:HD22	1.50	0.77
1:K:456:ARG:HA	1:K:463:LEU:HD11	1.66	0.76

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:15:SER:OG	1:D:95:GLN:OE1[4_555]	2.15	0.05
1:B:193:ASP:OD2	1:M:309:ARG:NH1[4_455]	2.17	0.03

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	469/536 (88%)	423 (90%)	45 (10%)	1 (0%)	47	78
1	B	469/536 (88%)	424 (90%)	43 (9%)	2 (0%)	34	69
1	C	469/536 (88%)	420 (90%)	46 (10%)	3 (1%)	25	61
1	D	469/536 (88%)	420 (90%)	47 (10%)	2 (0%)	34	69
1	E	469/536 (88%)	421 (90%)	45 (10%)	3 (1%)	25	61
1	F	469/536 (88%)	421 (90%)	45 (10%)	3 (1%)	25	61
1	G	469/536 (88%)	420 (90%)	48 (10%)	1 (0%)	47	78
1	H	469/536 (88%)	421 (90%)	45 (10%)	3 (1%)	25	61
1	I	469/536 (88%)	421 (90%)	46 (10%)	2 (0%)	34	69
1	J	469/536 (88%)	421 (90%)	45 (10%)	3 (1%)	25	61
1	K	469/536 (88%)	422 (90%)	46 (10%)	1 (0%)	47	78
1	L	469/536 (88%)	420 (90%)	48 (10%)	1 (0%)	47	78
1	M	469/536 (88%)	423 (90%)	44 (9%)	2 (0%)	34	69
All	All	6097/6968 (88%)	5477 (90%)	593 (10%)	27 (0%)	34	69

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	M	419	LYS
1	A	419	LYS
1	B	419	LYS
1	C	419	LYS
1	D	419	LYS

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	400/442 (90%)	391 (98%)	9 (2%)	50	71
1	B	400/442 (90%)	391 (98%)	9 (2%)	50	71

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	400/442 (90%)	390 (98%)	10 (2%)	47	70
1	D	400/442 (90%)	389 (97%)	11 (3%)	43	67
1	E	400/442 (90%)	391 (98%)	9 (2%)	50	71
1	F	400/442 (90%)	391 (98%)	9 (2%)	50	71
1	G	400/442 (90%)	392 (98%)	8 (2%)	55	75
1	H	400/442 (90%)	389 (97%)	11 (3%)	43	67
1	I	400/442 (90%)	388 (97%)	12 (3%)	41	65
1	J	400/442 (90%)	390 (98%)	10 (2%)	47	70
1	K	400/442 (90%)	391 (98%)	9 (2%)	50	71
1	L	400/442 (90%)	390 (98%)	10 (2%)	47	70
1	M	400/442 (90%)	391 (98%)	9 (2%)	50	71
All	All	5200/5746 (90%)	5074 (98%)	126 (2%)	49	71

5 of 126 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	390	SER
1	H	231	GLU
1	L	390	SER
1	G	147	LEU
1	G	277	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	184	GLN
1	M	136	GLN
1	J	216	HIS
1	B	331	GLN
1	J	184	GLN

5.3.3 RNA ⓘ

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	473/536 (88%)	0.30	31 (6%) 18 14	68, 139, 224, 386	1 (0%)
1	B	473/536 (88%)	0.27	26 (5%) 25 21	72, 141, 242, 294	1 (0%)
1	C	473/536 (88%)	0.40	49 (10%) 6 6	76, 144, 247, 328	1 (0%)
1	D	473/536 (88%)	0.43	40 (8%) 10 9	90, 160, 263, 353	1 (0%)
1	E	473/536 (88%)	0.57	46 (9%) 7 7	89, 164, 258, 310	1 (0%)
1	F	473/536 (88%)	0.32	32 (6%) 17 13	81, 164, 240, 285	1 (0%)
1	G	473/536 (88%)	0.19	29 (6%) 21 17	78, 149, 242, 341	1 (0%)
1	H	473/536 (88%)	0.23	22 (4%) 31 27	76, 137, 229, 294	1 (0%)
1	I	473/536 (88%)	0.19	26 (5%) 25 21	75, 129, 232, 316	1 (0%)
1	J	473/536 (88%)	0.32	42 (8%) 9 8	67, 128, 233, 363	1 (0%)
1	K	473/536 (88%)	0.30	32 (6%) 17 13	65, 132, 236, 321	1 (0%)
1	L	473/536 (88%)	0.33	34 (7%) 15 12	68, 138, 237, 287	1 (0%)
1	M	473/536 (88%)	0.21	29 (6%) 21 17	66, 141, 231, 329	1 (0%)
All	All	6149/6968 (88%)	0.31	438 (7%) 16 12	65, 145, 240, 386	13 (0%)

The worst 5 of 438 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	48	ASP	10.9
1	A	48	ASP	9.8
1	B	454	PRO	9.8
1	C	454	PRO	9.3
1	B	463	LEU	9.2

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.