



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 13, 2016 – 12:25 AM EDT

PDB ID : 5C59
Title : Crystal structure of the periplasmic region of MacB from E. coli
Authors : Ha, N.C.; Kim, J.S.
Deposited on : 2015-06-19
Resolution : 3.00 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
<http://wwpdb.org/validation/2016/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
Mogul : unknown
Xtriage (Phenix) : 1.9-1692
EDS : rb-20027674
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027674

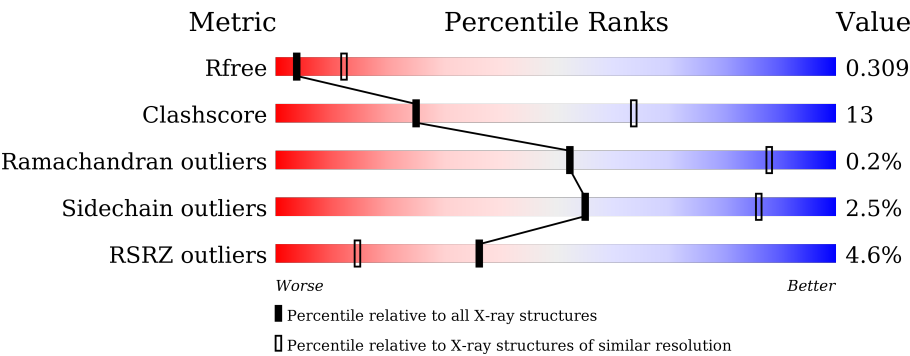
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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}	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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. 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	226	
1	B	226	
1	C	226	
1	D	226	
1	E	226	
1	F	226	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	226	 A horizontal bar chart showing the quality of chain G. The bar is divided into four segments: a small red segment at the beginning labeled '4%', followed by a green segment labeled '49%', a yellow segment labeled '17%', and a grey segment at the end labeled '33%'. A small black dot is located on the boundary between the yellow and grey segments.

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 8632 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.

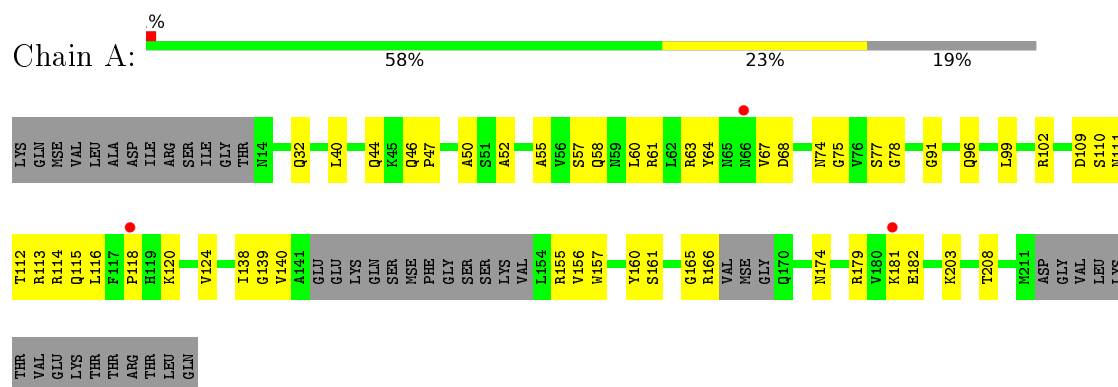
- Molecule 1 is a protein called Macrolide export ATP-binding/permease protein MacB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	183	Total	C	N	O	Se	0	0	0
			1449	912	257	276	4			
1	B	175	Total	C	N	O	Se	0	0	0
			1386	875	242	265	4			
1	C	173	Total	C	N	O	Se	0	0	0
			1371	866	242	260	3			
1	D	179	Total	C	N	O	Se	0	0	0
			1419	896	249	270	4			
1	E	91	Total	C	N	O	Se	0	0	0
			697	441	120	132	4			
1	F	144	Total	C	N	O	Se	0	0	0
			1130	716	196	214	4			
1	G	151	Total	C	N	O	Se	0	0	0
			1180	748	202	227	3			

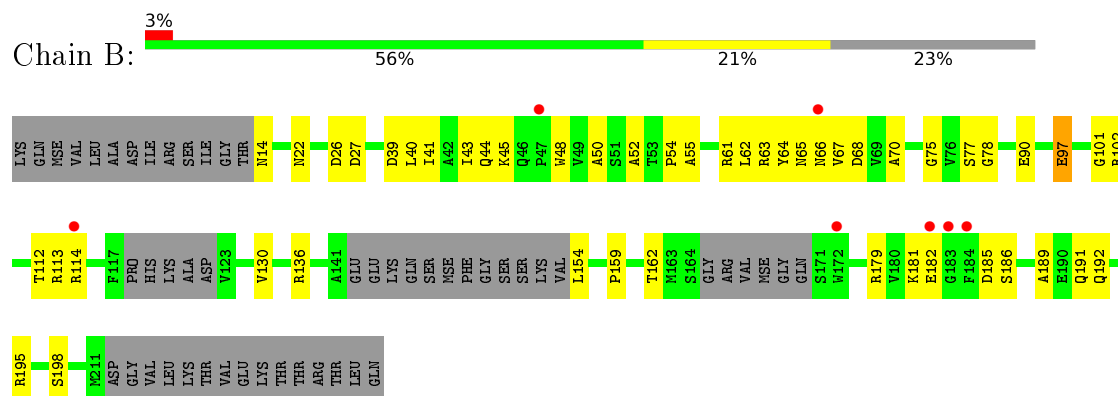
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

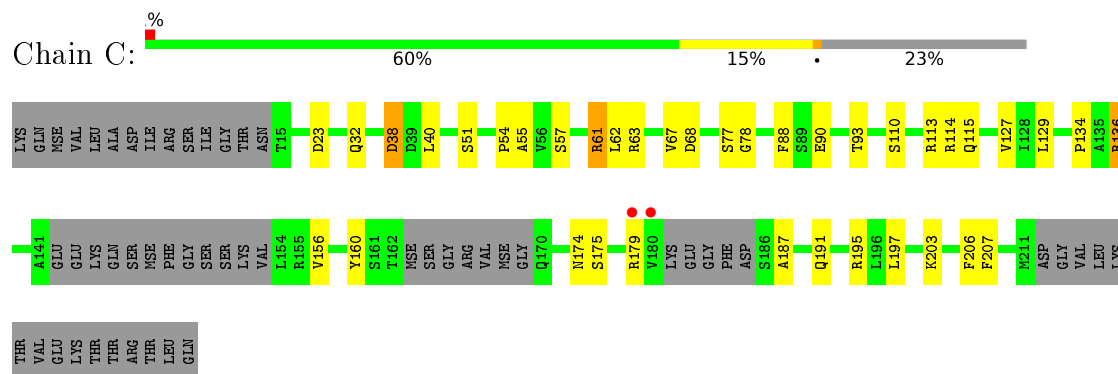
- Molecule 1: Macrolide export ATP-binding/permease protein MacB



- Molecule 1: Macrolide export ATP-binding/permease protein MacB


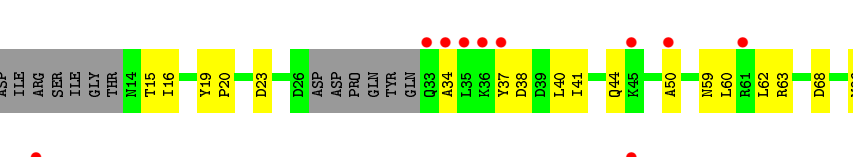


- Molecule 1: Macrolide export ATP-binding/permease protein MacB

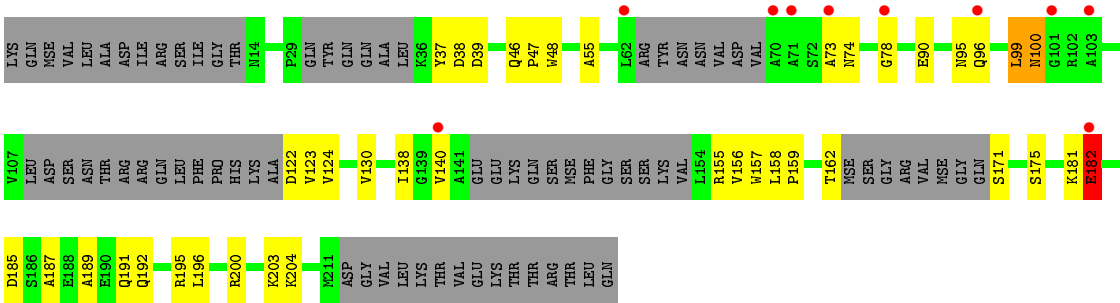


- Chain D:

- Chain E:
-
- | Residue | Count | Percentage |
|---------|-------|------------|
| Lys | 1 | 3% |
| Arg | 1 | 3% |
| Gln | 1 | 3% |
| Mse | 1 | 3% |
| Val | 1 | 3% |
| Asn | 1 | 3% |
| Leu | 1 | 3% |
| Ala | 1 | 3% |
| Asp | 1 | 3% |
| Ile | 1 | 3% |
| Arg | 1 | 3% |
| Ser | 1 | 3% |
| Ile | 1 | 3% |
| Gly | 1 | 3% |
| Thr | 1 | 3% |
| M14 | 1 | 3% |
| T15 | 1 | 3% |
| I16 | 1 | 3% |
| D17 | 1 | 3% |
| G21 | 1 | 3% |
| Asn | 1 | 3% |
| Asp | 1 | 3% |
| Phe | 1 | 3% |
| Gly | 1 | 3% |
| Asp | 1 | 3% |
| Asp | 1 | 3% |
| Asp | 1 | 3% |
| Pro | 1 | 3% |
| Gln | 1 | 3% |
| Tyr | 1 | 3% |
| Gln | 1 | 3% |
| Ala | 1 | 3% |
| Ala | 1 | 3% |
| Leu | 1 | 3% |
| Lys | 1 | 3% |
| Tyr | 1 | 3% |
| Asp | 1 | 3% |
| Asp | 1 | 3% |
| Leu | 1 | 3% |
| Ile | 1 | 3% |
| Ala | 1 | 3% |
| Ile | 1 | 3% |
| Gln | 1 | 3% |
| Lys | 1 | 3% |
| Gln | 1 | 3% |
| Pro | 1 | 3% |
| Trp | 1 | 3% |
| Val | 1 | 3% |
| Ala | 1 | 3% |
| Ser | 1 | 3% |
| Ala | 1 | 3% |
| Thr | 1 | 3% |
| Pro | 1 | 3% |
| Ala | 1 | 3% |
| Val | 1 | 3% |
| Gln | 1 | 3% |
| Asn | 1 | 3% |
| Leu | 1 | 3% |
| Arg | 1 | 3% |

- Chain F: 
- 
- Sequence logo for Chain F showing amino acid frequencies at 40 positions. The y-axis represents frequency from 0.00 to 0.10. The x-axis lists amino acids: VAL, LYS, GLN, MSE, VAL, LEU, ALA, ASP, SER, ILE, ARG, ILE, GLY, THR, N14, T15, I16, Y19, P20, D23, D26, ASP, PRQ, GLM, TYR, GLN, Q33, A34, L35, K36, Y37, D38, D39, L40, I41, Q44, K45, A50, N59, L60, R61, L62, R63, D68, M66, T67, F68, S69, E90, G91, N95, L99, R102, D109, S110, M111, R112, R114, Q115, L116, F117, PRQ, ARG, HIS, K120, A121, D122, V123, V130, G131, N132, G139, V140, A141, G149, LYS, GLN, LYS, SER, MSE, PHE, GLY, THR, SER, LYS, VAL, L154, A155, V156, W157, L158, P159, M163, SER, GLY, ARG, VAL, MSE, GLY, GLN, SER, TRP, L173, P170.

- Chain G: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	67.14Å 78.56Å 137.92Å 90.00° 99.72° 90.00°	Depositor
Resolution (Å)	20.00 – 3.00 33.46 – 2.99	Depositor EDS
% Data completeness (in resolution range)	96.1 (20.00-3.00) 96.0 (33.46-2.99)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	6.89 (at 3.00Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.247 , 0.306 0.255 , 0.309	Depositor DCC
R_{free} test set	1988 reflections (7.05%)	DCC
Wilson B-factor (Å ²)	70.7	Xtriage
Anisotropy	0.365	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 68.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	8632	wwPDB-VP
Average B, all atoms (Å ²)	76.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.45% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.39	0/1474	0.67	1/1994 (0.1%)
1	B	0.33	0/1408	0.64	1/1904 (0.1%)
1	C	0.31	0/1395	0.60	0/1891
1	D	0.39	0/1443	0.67	0/1952
1	E	0.31	0/701	0.59	0/939
1	F	0.36	0/1144	0.63	0/1545
1	G	0.32	0/1198	0.66	1/1619 (0.1%)
All	All	0.35	0/8763	0.64	3/11844 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	116	LEU	CA-CB-CG	7.72	133.06	115.30
1	G	99	LEU	CA-CB-CG	5.62	128.22	115.30
1	B	114	ARG	NE-CZ-NH2	-5.42	117.59	120.30

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	1449	0	1402	38	2
1	B	1386	0	1341	40	1
1	C	1371	0	1330	23	2

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	1419	0	1372	53	0
1	E	697	0	672	14	0
1	F	1130	0	1096	27	1
1	G	1180	0	1137	25	0
All	All	8632	0	8350	216	3

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 13.

The worst 5 of 216 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:C:90:GLU:OE1	1:C:113:ARG:NH1	2.03	0.91
1:A:113:ARG:HG2	1:A:113:ARG:HH11	1.36	0.90
1:A:61:ARG:NH1	1:A:68:ASP:OD2	2.05	0.90
1:C:63:ARG:HE	1:C:129:LEU:HD13	1.40	0.87
1:D:120:LYS:HD2	1:D:121:ALA:H	1.39	0.87

All (3) 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:A:44:GLN:OE1	1:B:198:SER:OG[1_655]	2.13	0.07
1:A:203:LYS:O	1:C:203:LYS:NZ[2_656]	2.16	0.04
1:C:195:ARG:NE	1:F:88:PHE:O[2_546]	2.18	0.02

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	177/226 (78%)	170 (96%)	7 (4%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	167/226 (74%)	161 (96%)	6 (4%)	0	100	100
1	C	165/226 (73%)	158 (96%)	7 (4%)	0	100	100
1	D	171/226 (76%)	162 (95%)	9 (5%)	0	100	100
1	E	77/226 (34%)	70 (91%)	6 (8%)	1 (1%)	15	53
1	F	132/226 (58%)	126 (96%)	6 (4%)	0	100	100
1	G	139/226 (62%)	133 (96%)	5 (4%)	1 (1%)	26	70
All	All	1028/1582 (65%)	980 (95%)	46 (4%)	2 (0%)	52	88

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	182	GLU
1	E	85	GLY

5.3.2 Protein sidechains ⓘ

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	156/187 (83%)	154 (99%)	2 (1%)	76	93
1	B	150/187 (80%)	148 (99%)	2 (1%)	76	93
1	C	148/187 (79%)	144 (97%)	4 (3%)	52	85
1	D	153/187 (82%)	147 (96%)	6 (4%)	39	77
1	E	76/187 (41%)	74 (97%)	2 (3%)	54	85
1	F	122/187 (65%)	117 (96%)	5 (4%)	37	76
1	G	127/187 (68%)	125 (98%)	2 (2%)	70	92
All	All	932/1309 (71%)	909 (98%)	23 (2%)	55	86

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	195	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	203	LYS
1	G	100	ASN
1	D	199	LEU
1	D	205	ASP

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

Mol	Chain	Res	Type
1	D	170	GLN
1	G	192	GLN
1	G	100	ASN
1	D	59	ASN
1	F	104	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 [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	179/226 (79%)	0.10	3 (1%) 73 45	33, 66, 98, 109	0
1	B	171/226 (75%)	0.16	7 (4%) 41 16	37, 60, 90, 104	0
1	C	170/226 (75%)	0.02	2 (1%) 81 55	44, 72, 101, 117	0
1	D	175/226 (77%)	0.33	9 (5%) 32 12	54, 83, 106, 119	0
1	E	87/226 (38%)	0.46	6 (6%) 20 7	63, 86, 115, 122	0
1	F	140/226 (61%)	0.47	12 (8%) 13 4	47, 81, 103, 108	0
1	G	148/226 (65%)	0.48	10 (6%) 20 7	56, 87, 108, 126	0
All	All	1070/1582 (67%)	0.26	49 (4%) 36 14	33, 76, 104, 126	0

The worst 5 of 49 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	173	LEU	4.9
1	B	183	GLY	4.9
1	F	112	THR	4.8
1	G	140	VAL	4.6
1	G	78	GLY	4.5

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.