



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 25, 2026 – 04:37 AM EDT

PDB ID : 2AO9 / pdb_00002ao9
Title : Structural Genomics, The crystal structure of a Phage protein (phBC6A51) from *Bacillus cereus* ATCC 14579
Authors : Zhang, R.; Joachimiak, G.; Collart, F.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)
Deposited on : 2005-08-12
Resolution : 1.90 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 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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

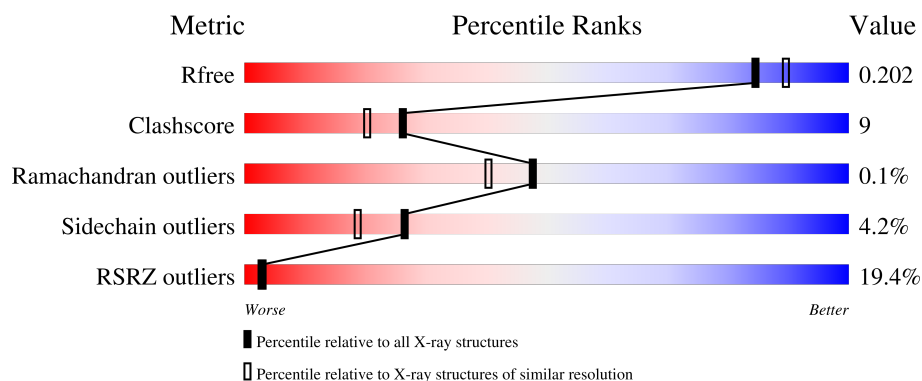
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.90 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}	180053	7789 (1.90-1.90)
Clashscore	190562	8410 (1.90-1.90)
Ramachandran outliers	187476	8333 (1.90-1.90)
Sidechain outliers	187428	8333 (1.90-1.90)
RSRZ outliers	180081	7790 (1.90-1.90)

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	155	15% 66% 9% • 25%
1	B	155	10% 61% 12% • 26%
1	C	155	8% 59% 14% • 25%
1	D	155	5% 64% 5% • 31%
1	E	155	11% 65% 8% • 26%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	155	
1	G	155	
1	H	155	
1	I	155	

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 9143 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 Phage protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	117	Total	C	N	O	S	0	0	0
			958	609	160	182	7			
1	B	114	Total	C	N	O	S	0	0	0
			933	593	157	177	6			
1	C	117	Total	C	N	O	S	0	0	0
			956	607	160	183	6			
1	D	107	Total	C	N	O	S	0	0	0
			877	558	147	167	5			
1	E	115	Total	C	N	O	S	0	0	0
			941	599	158	178	6			
1	F	110	Total	C	N	O	S	0	0	0
			902	575	153	169	5			
1	G	106	Total	C	N	O	S	0	0	0
			868	552	145	166	5			
1	H	114	Total	C	N	O	S	0	0	0
			932	592	157	178	5			
1	I	99	Total	C	N	O	S	0	0	0
			805	517	136	148	4			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	103	Total	O	0	0
			103	103		
2	B	119	Total	O	0	0
			119	119		
2	C	123	Total	O	0	0
			123	123		
2	D	123	Total	O	0	0
			123	123		
2	E	99	Total	O	0	0
			99	99		

Continued on next page...

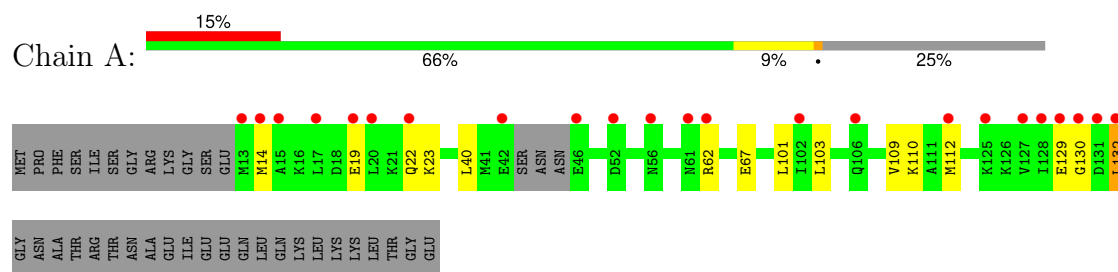
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	F	99	Total 99	O 99	0	0
2	G	102	Total 102	O 102	0	0
2	H	95	Total 95	O 95	0	0
2	I	108	Total 108	O 108	0	0

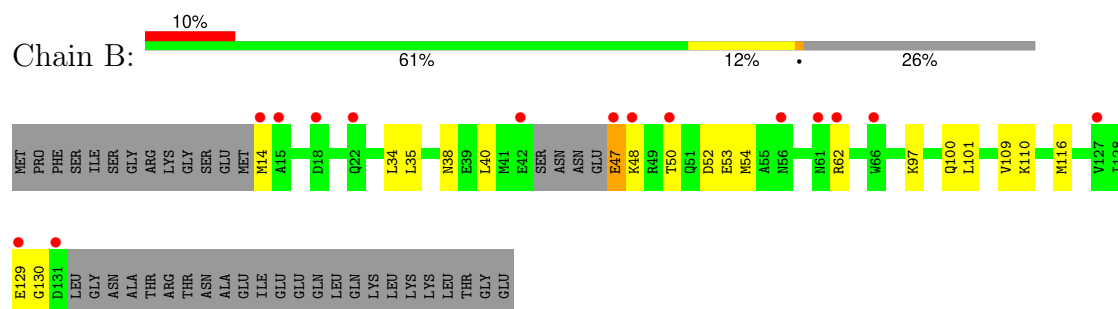
3 Residue-property plots [i](#)

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.

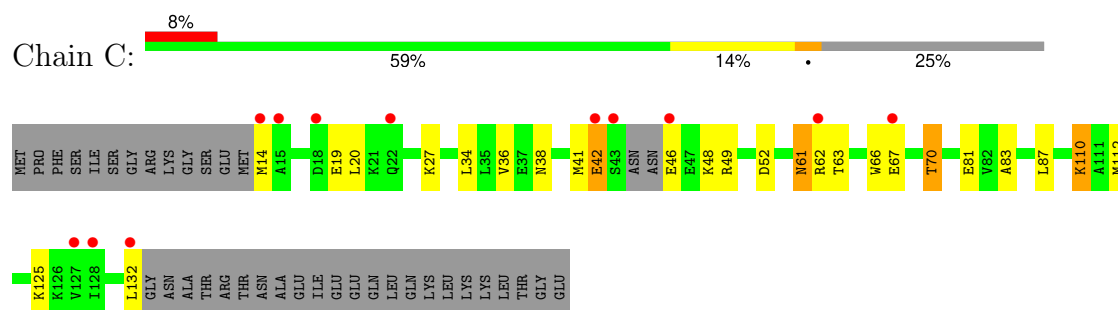
• Molecule 1: Phage protein



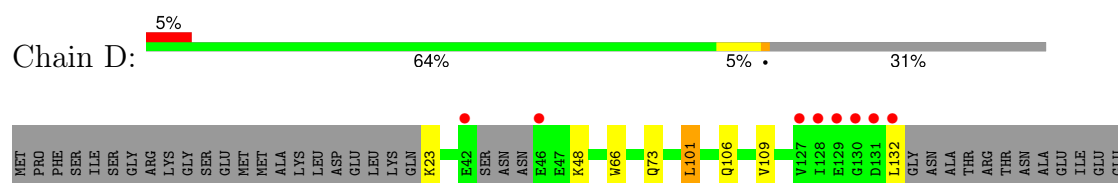
• Molecule 1: Phage protein



• Molecule 1: Phage protein



• Molecule 1: Phage protein



GLN
LEU
GLN
LYS
LEU
LYS
LEU
THR
GLY
GLU

• Molecule 1: Phage protein

Chain E: 11% 65% 8% 26%

MET PRO PHE SER ILE SER GLY ARG LYS SER MET M14 A15 K16 L17 D18 E19 L20 K21 N38 E42 SER ASN ASN GLU E47 K48 R49 N61 R62 F78 E89 L98 L101 P105 V109 K110 A111 M112 Q113 L114 D124 I128 E129 G130 D131 L132

GLY ASN ALA THR THR ASN ALA ILE SER GLU GLN MET LEU GLN LYS LEU LYS LEU THR GLY

• Molecule 1: Phage protein

Chain F: 13% 54% 15% 29%

MET PRO PHE SER ILE SER GLY ARG LYS SER MET MET ALA K16 L17 D18 E19 L20 K21 L34 L35 V36 E37 N38 M41 GLU SER ASN ASN GLU K48 R49 T50 E53 M54 A55 M56 M61 T63 T64 L65 W66 E67 V82 L87 K90 Y95 S96

K97 L101 G104 P105 V109 K110 D124 K125 V127 I128 E129 D131 LEU GLY ASN THR ARG THR ASN ALA ILE GLU SER GLN LEU LYS LYS LEU THR GLY

• Molecule 1: Phage protein

Chain G: 10% 55% 12% 32%

MET PRO PHE SER ILE SER GLY ARG LYS SER MET MET ALA L24 T25 A26 L34 E37 E39 L40 M41 E42 SER ASN ASN E46 E47 K48 R49 M54 G59 W56 E67 W68 K97 L98 M99 Q100 L101 V109 K110

Y115 D124 K125 V127 I128 E129 D131 L132 GLY ASN THR ARG THR ASN ALA ILE GLU GLN LEU GLN LYS LYS LEU THR GLY

• Molecule 1: Phage protein

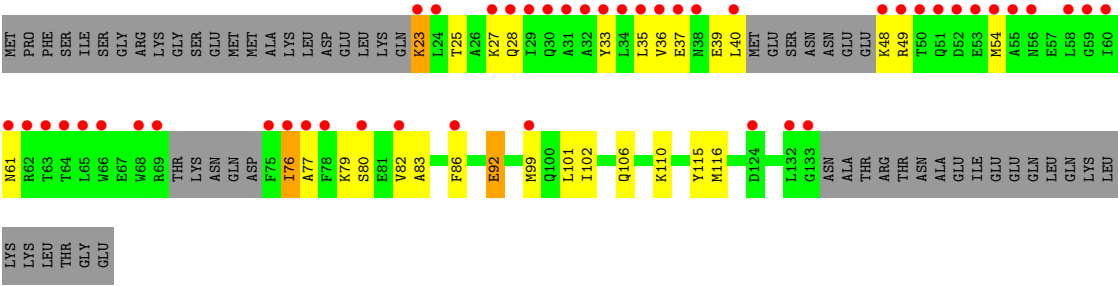
Chain H: 24% 67% 6% 26%

MET PRO PHE SER ILE SER GLY ARG LYS SER MET MET A15 K16 L17 D18 E19 L20 K21 Q22 K23 Q30 N38 E39 L40 M41 E42 S43 ASN E46 E47 K48 R49 T50 E53 M54 A55 E56 E57 L58 G59 I60 N61 R62 T63 T64 L65 W66 E67 W68 R69 T70 Q73

I102 V109 M112 L121 V127 I128 E129 G130 ASP LEU GLY ASN THR ARG THR ASN ALA GLU ILE GLU GLN LEU GLN LYS LEU LYS LEU THR GLY GLU

• Molecule 1: Phage protein

Chain I: 30% 45% 17% 36%



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	114.02Å 150.48Å 99.97Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.10 – 1.90 39.10 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.7 (39.10-1.90) 99.6 (39.10-1.90)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.75 (at 1.91Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.205 , 0.236 0.206 , 0.202	Depositor DCC
R_{free} test set	6730 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å ²)	25.9	Xtriage
Anisotropy	0.423	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 52.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	9143	wwPDB-VP
Average B, all atoms (Å ²)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.99% 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	0.59	0/969	0.77	0/1296
1	B	0.62	0/944	0.77	0/1263
1	C	0.64	0/967	0.80	0/1294
1	D	0.64	0/888	0.77	0/1190
1	E	0.65	1/952 (0.1%)	0.76	0/1274
1	F	0.56	0/913	0.81	0/1222
1	G	0.58	0/879	0.81	1/1179 (0.1%)
1	H	0.57	0/943	0.80	0/1262
1	I	0.74	1/815 (0.1%)	0.83	0/1091
All	All	0.62	2/8270 (0.0%)	0.79	1/11071 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	I	116	MET	SD-CE	-12.09	1.49	1.79
1	E	112	MET	SD-CE	-7.52	1.60	1.79

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	131	ASP	N-CA-C	6.24	118.88	108.96

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	958	0	980	15	0
1	B	933	0	954	30	0
1	C	956	0	976	24	0
1	D	877	0	891	10	0
1	E	941	0	965	12	0
1	F	902	0	928	35	0
1	G	868	0	878	30	0
1	H	932	0	952	12	0
1	I	805	0	828	28	0
2	A	103	0	0	1	0
2	B	119	0	0	5	0
2	C	123	0	0	5	0
2	D	123	0	0	3	0
2	E	99	0	0	4	0
2	F	99	0	0	5	0
2	G	102	0	0	3	0
2	H	95	0	0	3	0
2	I	108	0	0	5	0
All	All	9143	0	8352	156	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 9.

The worst 5 of 156 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:B:38:ASN:ND2	1:B:47:GLU:HG3	1.54	1.23
1:G:99:MET:SD	2:G:1417:HOH:O	2.00	1.17
1:F:34:LEU:HG	1:F:54:MET:HE2	1.24	1.16
1:G:34:LEU:HG	1:G:54:MET:HE2	1.22	1.15
1:B:52:ASP:OD1	1:B:62:ARG:HD3	1.51	1.09

There are no symmetry-related clashes.

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	113/155 (73%)	113 (100%)	0	0	100	100
1	B	110/155 (71%)	109 (99%)	1 (1%)	0	100	100
1	C	113/155 (73%)	112 (99%)	1 (1%)	0	100	100
1	D	103/155 (66%)	103 (100%)	0	0	100	100
1	E	111/155 (72%)	111 (100%)	0	0	100	100
1	F	106/155 (68%)	106 (100%)	0	0	100	100
1	G	102/155 (66%)	101 (99%)	0	1 (1%)	12	5
1	H	110/155 (71%)	109 (99%)	1 (1%)	0	100	100
1	I	93/155 (60%)	91 (98%)	2 (2%)	0	100	100
All	All	961/1395 (69%)	955 (99%)	5 (0%)	1 (0%)	48	40

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	130	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	104/136 (76%)	100 (96%)	4 (4%)	29	21
1	B	101/136 (74%)	97 (96%)	4 (4%)	28	20
1	C	104/136 (76%)	96 (92%)	8 (8%)	12	5
1	D	95/136 (70%)	93 (98%)	2 (2%)	47	44
1	E	102/136 (75%)	99 (97%)	3 (3%)	37	31
1	F	98/136 (72%)	95 (97%)	3 (3%)	35	29
1	G	94/136 (69%)	91 (97%)	3 (3%)	34	27
1	H	101/136 (74%)	97 (96%)	4 (4%)	28	20

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	I	86/136 (63%)	80 (93%)	6 (7%)	14 7
All	All	885/1224 (72%)	848 (96%)	37 (4%)	26 19

5 of 37 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	30	GLN
1	I	80	SER
1	H	63	THR
1	I	27	LYS
1	C	61	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	100	GLN
1	H	38	ASN
1	G	117	GLN
1	H	56	ASN
1	C	38	ASN

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 oligosaccharides 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	117/155 (75%)	1.09	23 (19%)	3 3	14, 26, 55, 65	0
1	B	114/155 (73%)	0.59	15 (13%)	7 7	13, 23, 45, 55	0
1	C	117/155 (75%)	0.39	12 (10%)	12 12	12, 22, 42, 46	0
1	D	107/155 (69%)	0.24	8 (7%)	20 21	12, 21, 39, 47	0
1	E	115/155 (74%)	0.67	17 (14%)	5 6	15, 24, 45, 64	0
1	F	110/155 (70%)	1.11	20 (18%)	3 3	16, 27, 45, 55	0
1	G	106/155 (68%)	0.87	16 (15%)	5 5	15, 25, 43, 55	0
1	H	114/155 (73%)	1.57	37 (32%)	1 1	13, 28, 61, 65	0
1	I	99/155 (63%)	2.25	46 (46%)	0 0	12, 34, 74, 77	0
All	All	999/1395 (71%)	0.96	194 (19%)	3 3	12, 25, 57, 77	0

The worst 5 of 194 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	47	GLU	7.3
1	I	65	LEU	7.2
1	I	76	ILE	7.1
1	I	34	LEU	6.9
1	I	60	ILE	6.7

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 oligosaccharides 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.