



wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 07:39 pm BST

PDB ID : 6A09
Title : Salmonella Typhi YfdX in the P222 space group
Authors : Ku, B.; Lee, H.S.; Kim, S.J.
Deposited on : 2018-06-05
Resolution : 2.29 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 : 2.11
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) : 2.11

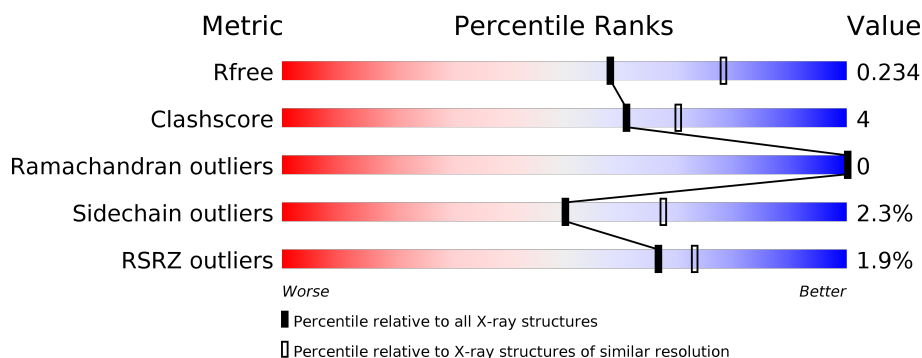
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.29 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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 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	185	<div> <div>0%</div> <div> <div></div> <div>85%</div> <div>9%</div> <div>6%</div> </div> </div>
1	B	185	<div> <div>0%</div> <div> <div></div> <div>85%</div> <div>8%</div> <div>7%</div> </div> </div>
1	C	185	<div> <div></div> <div> <div></div> <div>84%</div> <div>8%</div> <div>8%</div> </div> </div>
1	D	185	<div> <div>2%</div> <div> <div></div> <div>83%</div> <div>10%</div> <div>7%</div> </div> </div>
1	E	185	<div> <div>0%</div> <div> <div></div> <div>81%</div> <div>12%</div> <div>7%</div> </div> </div>
1	F	185	<div> <div>4%</div> <div> <div></div> <div>78%</div> <div>15%</div> <div>7%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	185	<div><div></div><div>3%</div><div>78%</div><div>15%</div><div>8%</div></div>
1	H	185	<div><div></div><div>3%</div><div>81%</div><div>10%</div><div>8%</div></div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 11255 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 YfdX protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	174	Total	C	N	O	S	0	0	0
			1352	837	235	275	5			
1	B	172	Total	C	N	O	S	0	0	0
			1335	827	233	271	4			
1	C	171	Total	C	N	O	S	0	0	0
			1328	823	232	269	4			
1	D	172	Total	C	N	O	S	0	0	0
			1335	827	233	271	4			
1	E	172	Total	C	N	O	S	0	0	0
			1335	827	233	271	4			
1	F	172	Total	C	N	O	S	0	0	0
			1335	827	233	271	4			
1	G	171	Total	C	N	O	S	0	0	0
			1328	823	232	269	4			
1	H	170	Total	C	N	O	S	0	0	0
			1320	819	231	266	4			

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	187	LEU	-	expression tag	UNP A0A0F7DJF1
A	188	GLU	-	expression tag	UNP A0A0F7DJF1
A	189	HIS	-	expression tag	UNP A0A0F7DJF1
A	190	HIS	-	expression tag	UNP A0A0F7DJF1
A	191	HIS	-	expression tag	UNP A0A0F7DJF1
A	192	HIS	-	expression tag	UNP A0A0F7DJF1
A	193	HIS	-	expression tag	UNP A0A0F7DJF1
A	194	HIS	-	expression tag	UNP A0A0F7DJF1
B	187	LEU	-	expression tag	UNP A0A0F7DJF1
B	188	GLU	-	expression tag	UNP A0A0F7DJF1
B	189	HIS	-	expression tag	UNP A0A0F7DJF1
B	190	HIS	-	expression tag	UNP A0A0F7DJF1
B	191	HIS	-	expression tag	UNP A0A0F7DJF1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	192	HIS	-	expression tag	UNP A0A0F7DJF1
B	193	HIS	-	expression tag	UNP A0A0F7DJF1
B	194	HIS	-	expression tag	UNP A0A0F7DJF1
C	187	LEU	-	expression tag	UNP A0A0F7DJF1
C	188	GLU	-	expression tag	UNP A0A0F7DJF1
C	189	HIS	-	expression tag	UNP A0A0F7DJF1
C	190	HIS	-	expression tag	UNP A0A0F7DJF1
C	191	HIS	-	expression tag	UNP A0A0F7DJF1
C	192	HIS	-	expression tag	UNP A0A0F7DJF1
C	193	HIS	-	expression tag	UNP A0A0F7DJF1
C	194	HIS	-	expression tag	UNP A0A0F7DJF1
D	187	LEU	-	expression tag	UNP A0A0F7DJF1
D	188	GLU	-	expression tag	UNP A0A0F7DJF1
D	189	HIS	-	expression tag	UNP A0A0F7DJF1
D	190	HIS	-	expression tag	UNP A0A0F7DJF1
D	191	HIS	-	expression tag	UNP A0A0F7DJF1
D	192	HIS	-	expression tag	UNP A0A0F7DJF1
D	193	HIS	-	expression tag	UNP A0A0F7DJF1
D	194	HIS	-	expression tag	UNP A0A0F7DJF1
E	187	LEU	-	expression tag	UNP A0A0F7DJF1
E	188	GLU	-	expression tag	UNP A0A0F7DJF1
E	189	HIS	-	expression tag	UNP A0A0F7DJF1
E	190	HIS	-	expression tag	UNP A0A0F7DJF1
E	191	HIS	-	expression tag	UNP A0A0F7DJF1
E	192	HIS	-	expression tag	UNP A0A0F7DJF1
E	193	HIS	-	expression tag	UNP A0A0F7DJF1
E	194	HIS	-	expression tag	UNP A0A0F7DJF1
F	187	LEU	-	expression tag	UNP A0A0F7DJF1
F	188	GLU	-	expression tag	UNP A0A0F7DJF1
F	189	HIS	-	expression tag	UNP A0A0F7DJF1
F	190	HIS	-	expression tag	UNP A0A0F7DJF1
F	191	HIS	-	expression tag	UNP A0A0F7DJF1
F	192	HIS	-	expression tag	UNP A0A0F7DJF1
F	193	HIS	-	expression tag	UNP A0A0F7DJF1
F	194	HIS	-	expression tag	UNP A0A0F7DJF1
G	187	LEU	-	expression tag	UNP A0A0F7DJF1
G	188	GLU	-	expression tag	UNP A0A0F7DJF1
G	189	HIS	-	expression tag	UNP A0A0F7DJF1
G	190	HIS	-	expression tag	UNP A0A0F7DJF1
G	191	HIS	-	expression tag	UNP A0A0F7DJF1
G	192	HIS	-	expression tag	UNP A0A0F7DJF1
G	193	HIS	-	expression tag	UNP A0A0F7DJF1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	194	HIS	-	expression tag	UNP A0A0F7DJF1
H	187	LEU	-	expression tag	UNP A0A0F7DJF1
H	188	GLU	-	expression tag	UNP A0A0F7DJF1
H	189	HIS	-	expression tag	UNP A0A0F7DJF1
H	190	HIS	-	expression tag	UNP A0A0F7DJF1
H	191	HIS	-	expression tag	UNP A0A0F7DJF1
H	192	HIS	-	expression tag	UNP A0A0F7DJF1
H	193	HIS	-	expression tag	UNP A0A0F7DJF1
H	194	HIS	-	expression tag	UNP A0A0F7DJF1

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	85	Total O 85 85	0	0
2	B	93	Total O 93 93	0	0
2	C	81	Total O 81 81	0	0
2	D	70	Total O 70 70	0	0
2	E	90	Total O 90 90	0	0
2	F	68	Total O 68 68	0	0
2	G	49	Total O 49 49	0	0
2	H	51	Total O 51 51	0	0

- Molecule 1: YfdX protein

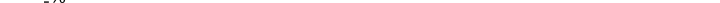


Chain B: %

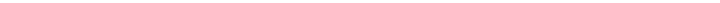


Chain C: 84% 8% 8%

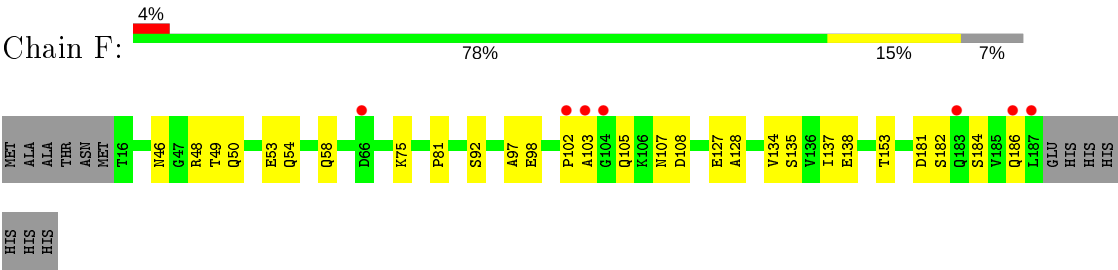


Chain D:  2% 83% 10% 7%

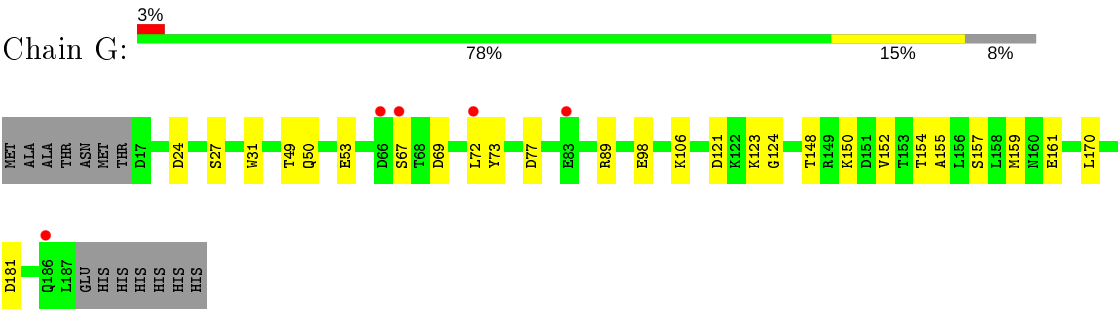


Chain E:  81% 12% 7%

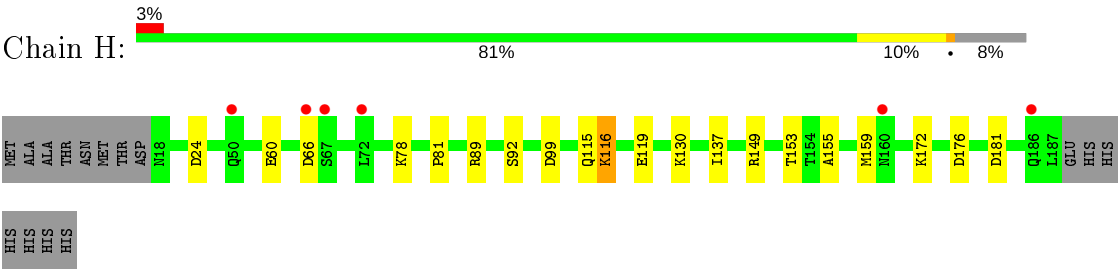




• Molecule 1: YfdX protein



• Molecule 1: YfdX protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	72.39Å 127.07Å 170.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.50 – 2.29 29.50 – 2.29	Depositor EDS
% Data completeness (in resolution range)	98.5 (29.50-2.29) 98.5 (29.50-2.29)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.64 (at 2.29Å)	Xtriage
Refinement program	PHENIX (1.10.1 _2155: ???)	Depositor
R, R_{free}	0.180 , 0.233 0.180 , 0.234	Depositor DCC
R_{free} test set	2000 reflections (2.85%)	wwPDB-VP
Wilson B-factor (Å ²)	31.3	Xtriage
Anisotropy	0.196	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 45.7	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	0.95	EDS
Total number of atoms	11255	wwPDB-VP
Average B, all atoms (Å ²)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.49% 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	0.45	0/1366	0.54	0/1840
1	B	0.42	0/1349	0.59	1/1818 (0.1%)
1	C	0.42	0/1342	0.58	1/1808 (0.1%)
1	D	0.41	0/1349	0.53	0/1818
1	E	0.42	0/1349	0.55	0/1818
1	F	0.40	0/1349	0.52	0/1818
1	G	0.38	0/1342	0.51	0/1808
1	H	0.37	0/1334	0.53	0/1797
All	All	0.41	0/10780	0.54	2/14525 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	170	LEU	CA-CB-CG	-8.24	96.36	115.30
1	B	48	ARG	NE-CZ-NH1	-5.27	117.67	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	1352	0	1352	11	0
1	B	1335	0	1337	10	0
1	C	1328	0	1330	15	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	1335	0	1337	11	0
1	E	1335	0	1337	12	0
1	F	1335	0	1337	20	0
1	G	1328	0	1330	16	0
1	H	1320	0	1326	12	0
2	A	85	0	0	1	1
2	B	93	0	0	4	1
2	C	81	0	0	3	0
2	D	70	0	0	2	0
2	E	90	0	0	1	1
2	F	68	0	0	2	1
2	G	49	0	0	0	1
2	H	51	0	0	2	0
All	All	11255	0	10686	96	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 4.

The worst 5 of 96 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:E:60:GLU:OE1	2:E:201:HOH:O	1.85	0.92
1:B:81:PRO:HG3	1:C:137:ILE:HD11	1.58	0.85
1:F:103:ALA:O	1:F:107:ASN:ND2	2.10	0.83
1:F:137:ILE:HD11	1:H:81:PRO:HG3	1.64	0.79
1:F:75:LYS:HE2	1:F:181:ASP:OD2	1.82	0.79

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)
2:E:271:HOH:O	2:G:246:HOH:O[2_865]	2.10	0.10
2:F:252:HOH:O	2:F:264:HOH:O[3_756]	2.15	0.05
2:A:277:HOH:O	2:B:285:HOH:O[3_855]	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	172/185 (93%)	169 (98%)	3 (2%)	0	100	100
1	B	170/185 (92%)	169 (99%)	1 (1%)	0	100	100
1	C	169/185 (91%)	163 (96%)	6 (4%)	0	100	100
1	D	170/185 (92%)	166 (98%)	4 (2%)	0	100	100
1	E	170/185 (92%)	169 (99%)	1 (1%)	0	100	100
1	F	170/185 (92%)	167 (98%)	3 (2%)	0	100	100
1	G	169/185 (91%)	165 (98%)	4 (2%)	0	100	100
1	H	168/185 (91%)	165 (98%)	3 (2%)	0	100	100
All	All	1358/1480 (92%)	1333 (98%)	25 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	147/156 (94%)	144 (98%)	3 (2%)	55	72
1	B	145/156 (93%)	142 (98%)	3 (2%)	53	70
1	C	144/156 (92%)	140 (97%)	4 (3%)	43	60
1	D	145/156 (93%)	144 (99%)	1 (1%)	84	92
1	E	145/156 (93%)	138 (95%)	7 (5%)	25	36
1	F	145/156 (93%)	143 (99%)	2 (1%)	67	81

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	144/156 (92%)	141 (98%)	3 (2%)	53	70
1	H	143/156 (92%)	139 (97%)	4 (3%)	43	60
All	All	1158/1248 (93%)	1131 (98%)	27 (2%)	50	67

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	25	LYS
1	E	92	SER
1	H	92	SER
1	E	66	ASP
1	B	92	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	54	GLN
1	B	58	GLN
1	F	105	GLN
1	F	186	GLN
1	G	175	GLN

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 carbohydrates 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	174/185 (94%)	-0.40	2 (1%) 80 85	19, 28, 53, 67	0
1	B	172/185 (92%)	-0.27	1 (0%) 89 92	20, 31, 53, 64	0
1	C	171/185 (92%)	-0.38	0 100 100	22, 31, 49, 64	0
1	D	172/185 (92%)	-0.26	4 (2%) 60 67	24, 35, 51, 62	0
1	E	172/185 (92%)	-0.29	1 (0%) 89 92	21, 31, 53, 61	0
1	F	172/185 (92%)	-0.12	7 (4%) 37 44	23, 38, 60, 68	0
1	G	171/185 (92%)	-0.08	5 (2%) 51 58	30, 42, 59, 70	0
1	H	170/185 (91%)	-0.02	6 (3%) 44 51	29, 45, 60, 69	0
All	All	1374/1480 (92%)	-0.23	26 (1%) 66 73	19, 35, 57, 70	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	50	GLN	3.5
1	H	67	SER	3.4
1	F	104	GLY	3.3
1	A	188	GLU	3.0
1	H	66	ASP	3.0

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 carbohydrates 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.