



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 7, 2026 – 05:06 AM UTC

PDB ID : 5XRC / pdb_00005xrc
Title : A Trimodular GH5_4 Subfamily Endoglucanase Structure with Large Unit Cell
Authors : Zhang, H.D.
Deposited on : 2017-06-08
Resolution : 2.90 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

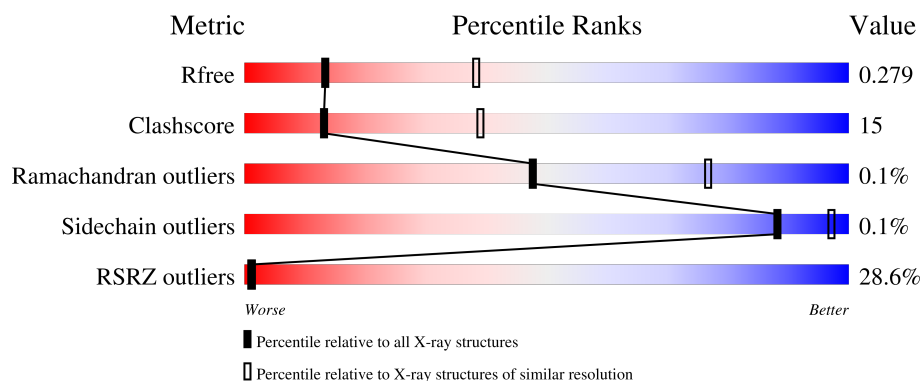
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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	2481 (2.90-2.90)
Clashscore	190562	2690 (2.90-2.90)
Ramachandran outliers	187476	2623 (2.90-2.90)
Sidechain outliers	187428	2625 (2.90-2.90)
RSRZ outliers	180081	2481 (2.90-2.90)

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	585	4% 71% 20% 9%
1	B	585	71% 54% 36% 9%
1	C	585	3% 69% 21% 9%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 12970 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 Cellulase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	531	Total	C	N	O	S	0	0	0
			4312	2724	727	848	13			
1	B	530	Total	C	N	O	S	0	0	0
			4303	2719	725	846	13			
1	C	532	Total	C	N	O	S	0	0	0
			4321	2730	729	849	13			

There are 147 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-33	MET	-	expression tag	UNP D4P8C6
A	-32	GLY	-	expression tag	UNP D4P8C6
A	-31	SER	-	expression tag	UNP D4P8C6
A	-30	SER	-	expression tag	UNP D4P8C6
A	-29	HIS	-	expression tag	UNP D4P8C6
A	-28	HIS	-	expression tag	UNP D4P8C6
A	-27	HIS	-	expression tag	UNP D4P8C6
A	-26	HIS	-	expression tag	UNP D4P8C6
A	-25	HIS	-	expression tag	UNP D4P8C6
A	-24	HIS	-	expression tag	UNP D4P8C6
A	-23	SER	-	expression tag	UNP D4P8C6
A	-22	SER	-	expression tag	UNP D4P8C6
A	-21	GLY	-	expression tag	UNP D4P8C6
A	-20	LEU	-	expression tag	UNP D4P8C6
A	-19	VAL	-	expression tag	UNP D4P8C6
A	-18	PRO	-	expression tag	UNP D4P8C6
A	-17	ARG	-	expression tag	UNP D4P8C6
A	-16	GLY	-	expression tag	UNP D4P8C6
A	-15	SER	-	expression tag	UNP D4P8C6
A	-14	HIS	-	expression tag	UNP D4P8C6
A	-13	MET	-	expression tag	UNP D4P8C6
A	-12	ALA	-	expression tag	UNP D4P8C6
A	-11	SER	-	expression tag	UNP D4P8C6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-10	MET	-	expression tag	UNP D4P8C6
A	-9	THR	-	expression tag	UNP D4P8C6
A	-8	GLY	-	expression tag	UNP D4P8C6
A	-7	GLY	-	expression tag	UNP D4P8C6
A	-6	GLN	-	expression tag	UNP D4P8C6
A	-5	GLN	-	expression tag	UNP D4P8C6
A	-4	MET	-	expression tag	UNP D4P8C6
A	-3	GLY	-	expression tag	UNP D4P8C6
A	-2	ARG	-	expression tag	UNP D4P8C6
A	-1	GLY	-	expression tag	UNP D4P8C6
A	0	SER	-	expression tag	UNP D4P8C6
A	537	VAL	-	expression tag	UNP D4P8C6
A	538	ASP	-	expression tag	UNP D4P8C6
A	539	LYS	-	expression tag	UNP D4P8C6
A	540	LEU	-	expression tag	UNP D4P8C6
A	541	ALA	-	expression tag	UNP D4P8C6
A	542	ALA	-	expression tag	UNP D4P8C6
A	543	ALA	-	expression tag	UNP D4P8C6
A	544	LEU	-	expression tag	UNP D4P8C6
A	545	GLU	-	expression tag	UNP D4P8C6
A	546	HIS	-	expression tag	UNP D4P8C6
A	547	HIS	-	expression tag	UNP D4P8C6
A	548	HIS	-	expression tag	UNP D4P8C6
A	549	HIS	-	expression tag	UNP D4P8C6
A	550	HIS	-	expression tag	UNP D4P8C6
A	551	HIS	-	expression tag	UNP D4P8C6
B	-33	MET	-	expression tag	UNP D4P8C6
B	-32	GLY	-	expression tag	UNP D4P8C6
B	-31	SER	-	expression tag	UNP D4P8C6
B	-30	SER	-	expression tag	UNP D4P8C6
B	-29	HIS	-	expression tag	UNP D4P8C6
B	-28	HIS	-	expression tag	UNP D4P8C6
B	-27	HIS	-	expression tag	UNP D4P8C6
B	-26	HIS	-	expression tag	UNP D4P8C6
B	-25	HIS	-	expression tag	UNP D4P8C6
B	-24	HIS	-	expression tag	UNP D4P8C6
B	-23	SER	-	expression tag	UNP D4P8C6
B	-22	SER	-	expression tag	UNP D4P8C6
B	-21	GLY	-	expression tag	UNP D4P8C6
B	-20	LEU	-	expression tag	UNP D4P8C6
B	-19	VAL	-	expression tag	UNP D4P8C6
B	-18	PRO	-	expression tag	UNP D4P8C6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-17	ARG	-	expression tag	UNP D4P8C6
B	-16	GLY	-	expression tag	UNP D4P8C6
B	-15	SER	-	expression tag	UNP D4P8C6
B	-14	HIS	-	expression tag	UNP D4P8C6
B	-13	MET	-	expression tag	UNP D4P8C6
B	-12	ALA	-	expression tag	UNP D4P8C6
B	-11	SER	-	expression tag	UNP D4P8C6
B	-10	MET	-	expression tag	UNP D4P8C6
B	-9	THR	-	expression tag	UNP D4P8C6
B	-8	GLY	-	expression tag	UNP D4P8C6
B	-7	GLY	-	expression tag	UNP D4P8C6
B	-6	GLN	-	expression tag	UNP D4P8C6
B	-5	GLN	-	expression tag	UNP D4P8C6
B	-4	MET	-	expression tag	UNP D4P8C6
B	-3	GLY	-	expression tag	UNP D4P8C6
B	-2	ARG	-	expression tag	UNP D4P8C6
B	-1	GLY	-	expression tag	UNP D4P8C6
B	0	SER	-	expression tag	UNP D4P8C6
B	537	VAL	-	expression tag	UNP D4P8C6
B	538	ASP	-	expression tag	UNP D4P8C6
B	539	LYS	-	expression tag	UNP D4P8C6
B	540	LEU	-	expression tag	UNP D4P8C6
B	541	ALA	-	expression tag	UNP D4P8C6
B	542	ALA	-	expression tag	UNP D4P8C6
B	543	ALA	-	expression tag	UNP D4P8C6
B	544	LEU	-	expression tag	UNP D4P8C6
B	545	GLU	-	expression tag	UNP D4P8C6
B	546	HIS	-	expression tag	UNP D4P8C6
B	547	HIS	-	expression tag	UNP D4P8C6
B	548	HIS	-	expression tag	UNP D4P8C6
B	549	HIS	-	expression tag	UNP D4P8C6
B	550	HIS	-	expression tag	UNP D4P8C6
B	551	HIS	-	expression tag	UNP D4P8C6
C	-33	MET	-	expression tag	UNP D4P8C6
C	-32	GLY	-	expression tag	UNP D4P8C6
C	-31	SER	-	expression tag	UNP D4P8C6
C	-30	SER	-	expression tag	UNP D4P8C6
C	-29	HIS	-	expression tag	UNP D4P8C6
C	-28	HIS	-	expression tag	UNP D4P8C6
C	-27	HIS	-	expression tag	UNP D4P8C6
C	-26	HIS	-	expression tag	UNP D4P8C6
C	-25	HIS	-	expression tag	UNP D4P8C6

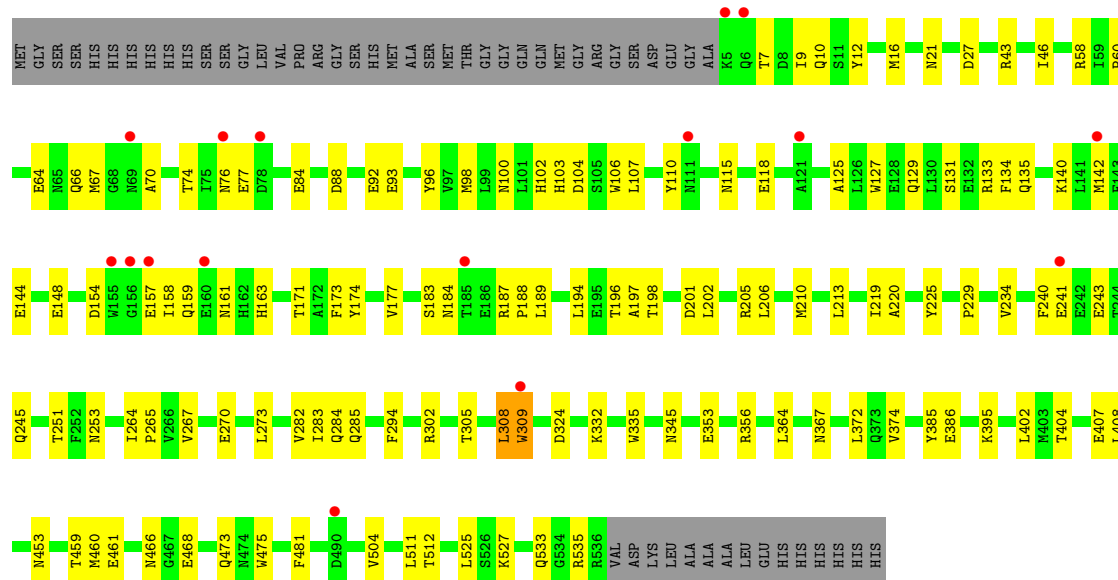
Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-24	HIS	-	expression tag	UNP D4P8C6
C	-23	SER	-	expression tag	UNP D4P8C6
C	-22	SER	-	expression tag	UNP D4P8C6
C	-21	GLY	-	expression tag	UNP D4P8C6
C	-20	LEU	-	expression tag	UNP D4P8C6
C	-19	VAL	-	expression tag	UNP D4P8C6
C	-18	PRO	-	expression tag	UNP D4P8C6
C	-17	ARG	-	expression tag	UNP D4P8C6
C	-16	GLY	-	expression tag	UNP D4P8C6
C	-15	SER	-	expression tag	UNP D4P8C6
C	-14	HIS	-	expression tag	UNP D4P8C6
C	-13	MET	-	expression tag	UNP D4P8C6
C	-12	ALA	-	expression tag	UNP D4P8C6
C	-11	SER	-	expression tag	UNP D4P8C6
C	-10	MET	-	expression tag	UNP D4P8C6
C	-9	THR	-	expression tag	UNP D4P8C6
C	-8	GLY	-	expression tag	UNP D4P8C6
C	-7	GLY	-	expression tag	UNP D4P8C6
C	-6	GLN	-	expression tag	UNP D4P8C6
C	-5	GLN	-	expression tag	UNP D4P8C6
C	-4	MET	-	expression tag	UNP D4P8C6
C	-3	GLY	-	expression tag	UNP D4P8C6
C	-2	ARG	-	expression tag	UNP D4P8C6
C	-1	GLY	-	expression tag	UNP D4P8C6
C	0	SER	-	expression tag	UNP D4P8C6
C	537	VAL	-	expression tag	UNP D4P8C6
C	538	ASP	-	expression tag	UNP D4P8C6
C	539	LYS	-	expression tag	UNP D4P8C6
C	540	LEU	-	expression tag	UNP D4P8C6
C	541	ALA	-	expression tag	UNP D4P8C6
C	542	ALA	-	expression tag	UNP D4P8C6
C	543	ALA	-	expression tag	UNP D4P8C6
C	544	LEU	-	expression tag	UNP D4P8C6
C	545	GLU	-	expression tag	UNP D4P8C6
C	546	HIS	-	expression tag	UNP D4P8C6
C	547	HIS	-	expression tag	UNP D4P8C6
C	548	HIS	-	expression tag	UNP D4P8C6
C	549	HIS	-	expression tag	UNP D4P8C6
C	550	HIS	-	expression tag	UNP D4P8C6
C	551	HIS	-	expression tag	UNP D4P8C6

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	8	Total 8	O 8	0	0
2	B	1	Total 1	O 1	0	0
2	C	25	Total 25	O 25	0	0



4 Data and refinement statistics

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, α , β , γ	120.14Å 120.14Å 615.34Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	85.97 – 2.90 85.97 – 2.90	Depositor EDS
% Data completeness (in resolution range)	98.7 (85.97-2.90) 99.0 (85.97-2.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.48 (at 2.91Å)	Xtrriage
Refinement program	PHENIX (1.11.1_2575)	Depositor
R, R_{free}	0.276 , 0.332 (Not available) , 0.279	Depositor DCC
R_{free} test set	2408 reflections (4.74%)	wwPDB-VP
Wilson B-factor (Å ²)	33.3	Xtrriage
Anisotropy	0.030	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 61.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	12970	wwPDB-VP
Average B, all atoms (Å ²)	58.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 79.99 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 4.9383e-07. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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	0.22	1/4430 (0.0%)	0.47	4/6032 (0.1%)
1	B	0.29	0/4421	0.63	8/6020 (0.1%)
1	C	0.26	1/4439 (0.0%)	0.49	6/6043 (0.1%)
All	All	0.26	2/13290 (0.0%)	0.54	18/18095 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	308	LEU	C-N	-13.78	1.13	1.33
1	A	211	LYS	CG-CD	5.37	1.68	1.52

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	309	TRP	O-C-N	13.58	138.70	123.22
1	A	211	LYS	CA-CB-CG	7.91	129.93	114.10
1	B	239	ARG	CG-CD-NE	-7.16	96.25	112.00
1	C	308	LEU	CA-C-N	7.15	132.28	122.77
1	C	308	LEU	C-N-CA	7.15	132.28	122.77

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	4312	0	3976	90	0
1	B	4303	0	3968	195	0
1	C	4321	0	3988	87	0
2	A	8	0	0	1	0
2	B	1	0	0	0	0
2	C	25	0	0	0	0
All	All	12970	0	11932	370	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 15.

The worst 5 of 370 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:379:LEU:HD21	1:B:384:ASP:CG	1.51	1.35
1:B:360:ILE:HD11	1:B:426:PHE:CE2	1.69	1.28
1:B:424:TRP:CD1	1:B:426:PHE:CE1	2.20	1.27
1:B:379:LEU:HD21	1:B:384:ASP:OD2	1.09	1.19
1:B:381:LEU:HD23	1:B:382:GLY:N	1.57	1.17

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	529/585 (90%)	507 (96%)	21 (4%)	1 (0%)	43	72
1	B	528/585 (90%)	500 (95%)	28 (5%)	0	100	100
1	C	530/585 (91%)	505 (95%)	24 (4%)	1 (0%)	43	72
All	All	1587/1755 (90%)	1512 (95%)	73 (5%)	2 (0%)	48	77

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	70	ALA
1	A	160	GLU

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	462/503 (92%)	461 (100%)	1 (0%)	87	96
1	B	461/503 (92%)	461 (100%)	0	100	100
1	C	463/503 (92%)	463 (100%)	0	100	100
All	All	1386/1509 (92%)	1385 (100%)	1 (0%)	88	97

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	211	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	411	ASN
1	C	417	GLN
1	C	533	GLN
1	B	217	ASN
1	B	162	HIS

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	308:LEU	C	309:TRP	N	1.13

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	531/585 (90%)	0.76	23 (4%) 40 31	26, 55, 82, 119	0
1	B	530/585 (90%)	2.80	416 (78%) 0 0	30, 90, 108, 126	0
1	C	532/585 (90%)	0.58	16 (3%) 52 43	14, 28, 49, 83	0
All	All	1593/1755 (90%)	1.38	455 (28%) 1 1	14, 56, 101, 126	0

The worst 5 of 455 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	380	ALA	6.6
1	B	389	GLY	5.9
1	B	372	LEU	5.9
1	B	521	VAL	5.3
1	B	306	HIS	5.2

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 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.