Full wwPDB X-ray Structure Validation Report (i)

Mar 6, 2026 – 04:48 PM UTC

PDB ID	:	5MCM / pdb_00005mcm
Title	:	Radiation damage to GH7 Family Cellobiohydrolase from Daphnia pulex: Dose (DWD) 20.6 MGy
Authors	:	Bury, C.S.; McGeehan, J.E.; Ebrahim, A.; Garman, E.F.
Deposited on	:	2016-11-10
Resolution	:	2.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 (i) 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 (i)) were used in the production of this report:

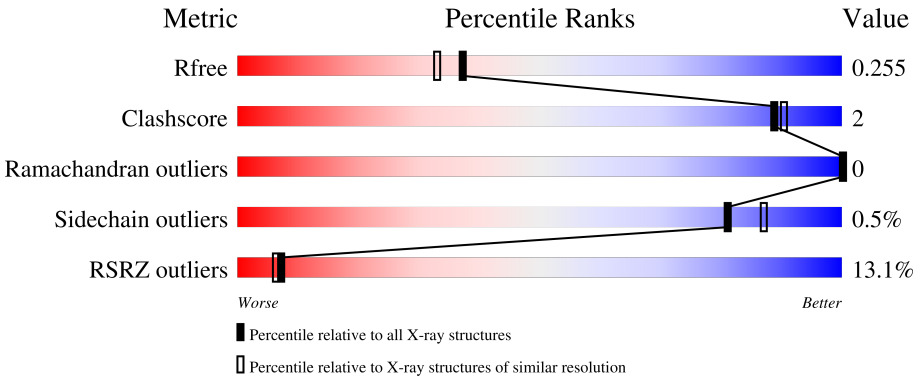
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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}	180053	10052 (2.00-2.00)
Clashscore	190562	11152 (2.00-2.00)
Ramachandran outliers	187476	11031 (2.00-2.00)
Sidechain outliers	187428	11029 (2.00-2.00)
RSRZ outliers	180081	10067 (2.00-2.00)

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 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 <=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	445	10% 97%
1	B	445	16% 94% 5%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GOL	B	501	-	-	-	X

2 Entry composition (i)

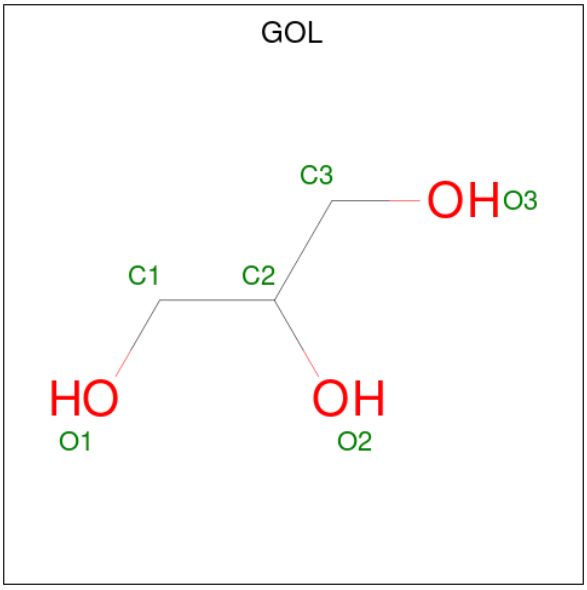
There are 4 unique types of molecules in this entry. The entry contains 7298 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Cellobiohydrolase CHBI.

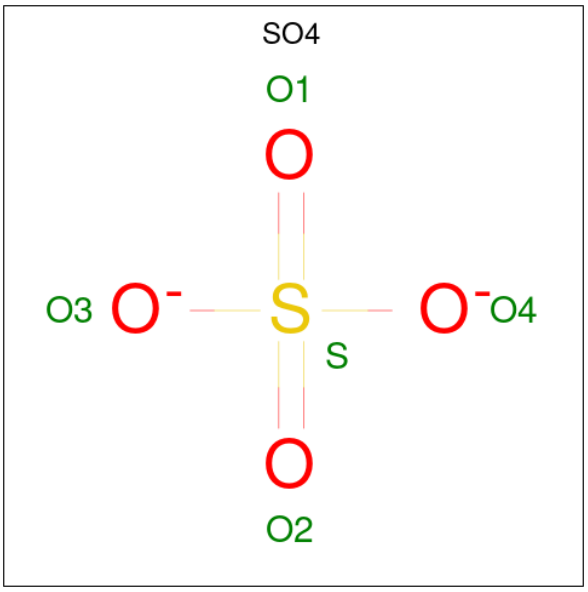
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	442	Total	C	N	O	S	0	0	0
			3353	2082	566	676	29			
1	B	441	Total	C	N	O	S	0	0	0
			3347	2079	565	674	29			

- Molecule 2 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			6	3	3		
2	B	1	Total	C	O	0	0
			6	3	3		
2	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is SULFATE ION (CCD ID: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total O S 5 4 1	0	0

- Molecule 4 is water.

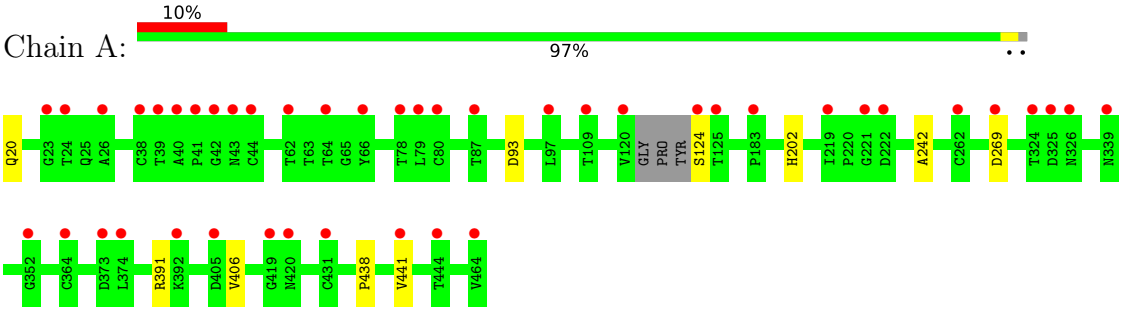
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	337	Total O 337 337	0	0
4	B	238	Total O 238 238	0	0

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.

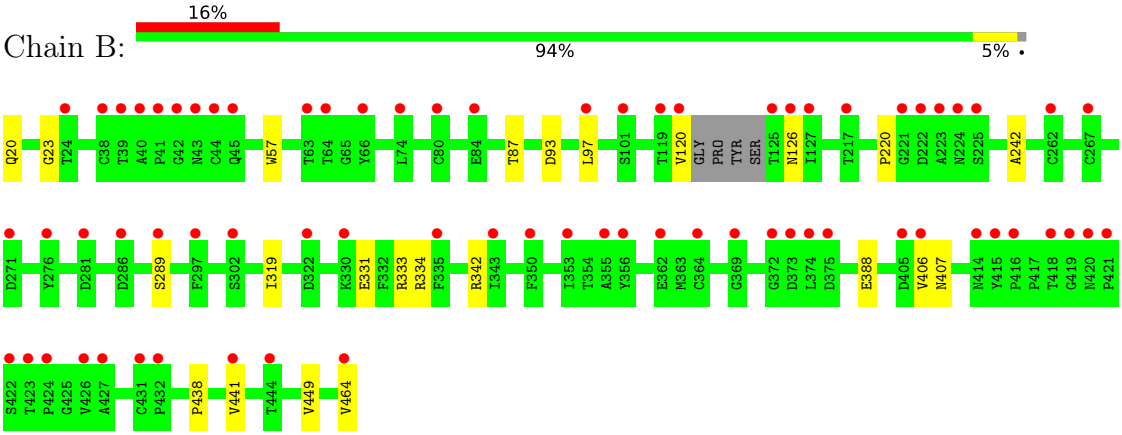
- Molecule 1: Cellobiohydrolase CHBI

Chain A:



- Molecule 1: Cellobiohydrolase CHBI

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	127.47Å 46.77Å 174.08Å 90.00° 108.20° 90.00°	Depositor
Resolution (Å)	82.69 – 2.00 82.69 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.9 (82.69-2.00) 99.6 (82.69-2.00)	Depositor EDS
R_{merge}	0.25	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.55 (at 2.00Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.232 , 0.252 0.235 , 0.255	Depositor DCC
R_{free} test set	3285 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å^2)	23.6	Xtriage
Anisotropy	0.343	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 49.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	7298	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 4.42% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, SO4, PCA

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	0.52	0/3427	0.72	1/4667 (0.0%)
1	B	0.50	0/3421	0.72	1/4659 (0.0%)
All	All	0.51	0/6848	0.72	2/9326 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	242	ALA	N-CA-C	6.07	117.74	108.52
1	B	242	ALA	N-CA-C	5.56	117.55	108.76

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	3353	0	3114	5	0
1	B	3347	0	3109	19	0
2	A	6	0	8	0	0
2	B	12	0	16	2	0
3	B	5	0	0	0	0
4	A	337	0	0	1	0
4	B	238	0	0	12	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	7298	0	6247	24	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 2.

All (24) 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:388:GLU:OE2	4:B:602:HOH:O	2.15	0.65
1:B:334:ARG:O	4:B:601:HOH:O	2.14	0.64
1:B:449:VAL:O	4:B:603:HOH:O	2.15	0.62
1:B:331:GLU:OE1	4:B:607:HOH:O	2.18	0.57
1:B:220:PRO:O	4:B:604:HOH:O	2.17	0.57
1:B:97:LEU:O	4:B:605:HOH:O	2.18	0.57
1:B:342:ARG:O	4:B:608:HOH:O	2.18	0.55
1:B:23:GLY:O	4:B:606:HOH:O	2.18	0.55
1:B:407:ASN:O	4:B:609:HOH:O	2.19	0.53
1:B:57:TRP:HD1	1:B:126:ASN:HD21	1.57	0.53
1:A:202:HIS:HE1	1:A:269:ASP:OD1	1.92	0.52
1:B:319:ILE:HG23	2:B:501:GOL:H11	1.91	0.51
1:B:57:TRP:CD1	1:B:126:ASN:HD21	2.30	0.50
1:B:87:THR:O	4:B:610:HOH:O	2.20	0.49
1:A:391:ARG:NH2	4:A:606:HOH:O	2.43	0.48
1:B:406:VAL:HG12	1:B:406:VAL:O	2.14	0.47
1:A:406:VAL:O	1:A:406:VAL:HG12	2.13	0.47
1:B:319:ILE:CG2	2:B:501:GOL:H11	2.47	0.44
1:B:333:ARG:HA	4:B:660:HOH:O	2.17	0.44
1:B:120:VAL:O	1:B:120:VAL:HG23	2.20	0.42
1:B:289:SER:HA	4:B:672:HOH:O	2.18	0.42
1:B:438:PRO:HA	1:B:441:VAL:HG22	2.02	0.42
1:A:202:HIS:CE1	1:A:269:ASP:OD1	2.71	0.42
1:A:438:PRO:HA	1:A:441:VAL:HG22	2.03	0.40

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	438/445 (98%)	432 (99%)	6 (1%)	0	100	100
1	B	437/445 (98%)	431 (99%)	6 (1%)	0	100	100
All	All	875/890 (98%)	863 (99%)	12 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	365/367 (100%)	363 (100%)	2 (0%)	81	87
1	B	364/367 (99%)	362 (100%)	2 (0%)	81	87
All	All	729/734 (99%)	725 (100%)	4 (0%)	81	87

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	93	ASP
1	A	124	SER
1	B	93	ASP
1	B	464	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	199	GLN
1	A	202	HIS
1	A	259	GLN
1	A	326	ASN
1	A	447	ASN
1	B	126	ASN
1	B	232	ASN
1	B	259	GLN
1	B	339	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)

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
1	PCA	A	20	1	7,8,9	0.51	0	9,10,12	1.48	2 (22%)
1	PCA	B	20	1	7,8,9	0.57	0	9,10,12	1.53	2 (22%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	PCA	A	20	1	-	0/0/11/13	0/1/1/1
1	PCA	B	20	1	-	0/0/11/13	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	20	PCA	CB-CA-C	-3.24	108.22	112.66
1	A	20	PCA	CB-CA-C	-3.01	108.53	112.66
1	B	20	PCA	OE-CD-CG	-2.21	122.78	126.72
1	A	20	PCA	OE-CD-CG	-2.16	122.86	126.72

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

4 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	GOL	B	502	-	5,5,5	0.33	0	5,5,5	0.27	0
2	GOL	A	501	-	5,5,5	0.36	0	5,5,5	0.32	0
2	GOL	B	501	-	5,5,5	0.31	0	5,5,5	0.46	0
3	SO4	B	503	-	4,4,4	0.42	0	6,6,6	0.07	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	B	502	-	-	1/4/4/4	-
2	GOL	A	501	-	-	2/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	B	501	-	-	1/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	501	GOL	C1-C2-C3-O3
2	B	502	GOL	C1-C2-C3-O3
2	A	501	GOL	O2-C2-C3-O3
2	B	501	GOL	O2-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	501	GOL	2	0

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

6.1 Protein, DNA and RNA chains

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, 95^{th} 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	441/445 (99%)	0.87	44 (9%) 12 11	15, 26, 40, 49	0
1	B	440/445 (98%)	1.16	71 (16%) 4 4	18, 31, 46, 55	0
All	All	881/890 (98%)	1.02	115 (13%) 7 6	15, 29, 44, 55	0

All (115) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	120	VAL	6.1
1	A	44	CYS	5.3
1	B	364	CYS	5.1
1	A	66	TYR	4.5
1	A	364	CYS	4.3
1	A	26	ALA	4.3
1	A	120	VAL	4.2
1	B	101	SER	3.9
1	B	44	CYS	3.9
1	B	423	THR	3.7
1	B	419	GLY	3.7
1	B	125	THR	3.6
1	B	45	GLN	3.5
1	A	222	ASP	3.4
1	B	374	LEU	3.4
1	B	64	THR	3.4
1	A	64	THR	3.3
1	B	41	PRO	3.3
1	A	464	VAL	3.3
1	A	419	GLY	3.2
1	B	362	GLU	3.2
1	B	302	SER	3.2
1	B	418	THR	3.2
1	B	221	GLY	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	42	GLY	3.1
1	B	223	ALA	3.1
1	B	422	SER	3.1
1	B	426	VAL	3.1
1	B	350	PHE	3.0
1	A	392	LYS	3.0
1	B	97	LEU	3.0
1	B	464	VAL	3.0
1	A	43	ASN	3.0
1	B	415	TYR	3.0
1	A	262	CYS	2.9
1	B	66	TYR	2.9
1	B	40	ALA	2.9
1	A	124	SER	2.9
1	A	38	CYS	2.9
1	B	224	ASN	2.8
1	B	38	CYS	2.8
1	B	432	PRO	2.8
1	B	420	ASN	2.8
1	A	183	PRO	2.7
1	B	322	ASP	2.7
1	A	352	GLY	2.7
1	B	343	ILE	2.7
1	B	427	ALA	2.6
1	A	79	LEU	2.6
1	B	39	THR	2.6
1	A	405	ASP	2.6
1	B	267	CYS	2.6
1	A	23	GLY	2.6
1	A	78	THR	2.6
1	A	109	THR	2.6
1	A	24	THR	2.5
1	A	444	THR	2.5
1	B	441	VAL	2.5
1	B	356	TYR	2.5
1	B	80	CYS	2.5
1	B	271	ASP	2.5
1	B	421	PRO	2.5
1	B	289	SER	2.4
1	B	355	ALA	2.4
1	A	219	ILE	2.4
1	B	353	ILE	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	41	PRO	2.4
1	A	40	ALA	2.4
1	B	297	PHE	2.4
1	B	375	ASP	2.4
1	B	225	SER	2.4
1	A	80	CYS	2.4
1	B	217	THR	2.4
1	B	74	LEU	2.4
1	B	63	THR	2.4
1	A	374	LEU	2.3
1	A	221	GLY	2.3
1	B	373	ASP	2.3
1	B	405	ASP	2.3
1	A	441	VAL	2.3
1	B	262	CYS	2.3
1	B	335	PHE	2.3
1	B	119	THR	2.3
1	B	416	PRO	2.3
1	B	424	PRO	2.3
1	A	326	ASN	2.3
1	A	324	THR	2.2
1	A	420	ASN	2.2
1	A	373	ASP	2.2
1	B	222	ASP	2.2
1	B	431	CYS	2.2
1	A	39	THR	2.2
1	B	444	THR	2.2
1	B	126	ASN	2.2
1	B	24	THR	2.2
1	B	127	ILE	2.1
1	B	406	VAL	2.1
1	B	369	GLY	2.1
1	A	339	ASN	2.1
1	A	269	ASP	2.1
1	A	97	LEU	2.1
1	B	84	GLU	2.1
1	B	414	ASN	2.1
1	A	431	CYS	2.1
1	A	325	ASP	2.1
1	B	286	ASP	2.1
1	A	87	THR	2.1
1	A	125	THR	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	62	THR	2.0
1	B	276	TYR	2.0
1	B	330	LYS	2.0
1	B	43	ASN	2.0
1	B	281	ASP	2.0
1	B	42	GLY	2.0
1	B	372	GLY	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
1	PCA	B	20	8/9	0.86	0.11	21,24,26,27	0
1	PCA	A	20	8/9	0.91	0.10	28,30,31,31	0

6.3 Carbohydrates (i)

There are no oligosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	GOL	B	501	6/6	0.51	0.42	47,59,61,66	0
2	GOL	B	502	6/6	0.62	0.23	44,51,57,59	0
3	SO4	B	503	5/5	0.78	0.14	64,66,74,75	0
2	GOL	A	501	6/6	0.80	0.18	30,33,36,40	0

6.5 Other polymers (i)

There are no such residues in this entry.