



# wwPDB X-ray Structure Validation Summary Report

Jan 14, 2024 – 06:15 am GMT

PDB ID : 6YJQ  
Title : Crystal structure of unliganded MGAT5 (alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyltransferase V) luminal domain with a Lys329-Ile345 loop truncation  
Authors : Wu, L.; Darby, J.F.; Gilio, A.K.; Davies, G.J.  
Deposited on : 2020-04-04  
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](mailto: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 : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

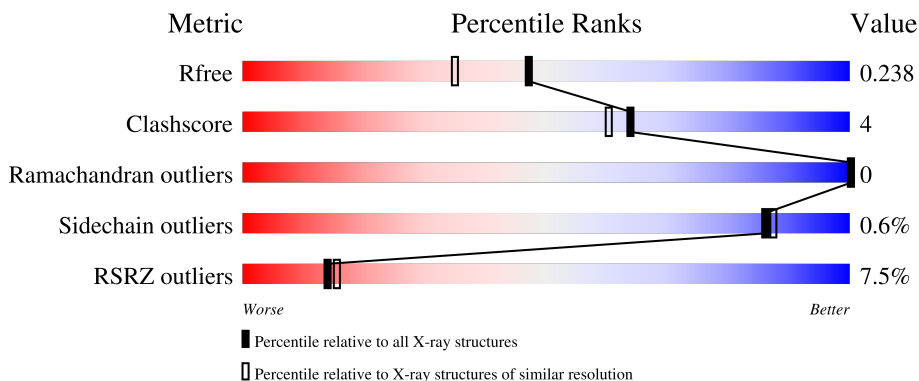
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

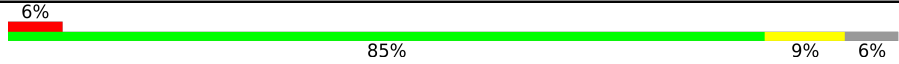
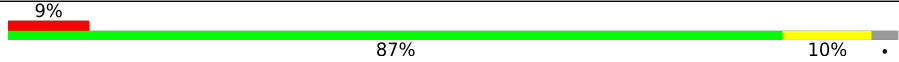
The reported resolution of this entry is 1.90 Å.

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))
$R_{free}$	130704	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (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  $\geq 3$ , 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	AAA	515	
1	BBB	515	

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
3	SO4	BBB	807	-	-	X	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 16444 atoms, of which 8072 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 Alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyl transferase A, Alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyltransferase A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	AAA	486	7884	2535	3949	677	697	26	195	0	0
1	BBB	500	8115	2612	4061	697	719	26	200	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	329	GLY	-	linker	UNP Q09328
AAA	330	GLY	-	linker	UNP Q09328
AAA	331	GLY	-	linker	UNP Q09328
AAA	332	GLY	-	linker	UNP Q09328
BBB	329	GLY	-	linker	UNP Q09328
BBB	330	GLY	-	linker	UNP Q09328
BBB	331	GLY	-	linker	UNP Q09328
BBB	332	GLY	-	linker	UNP Q09328

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	AAA	1	Total	C	H	O	1	0
			10	2	6	2		
2	AAA	1	Total	C	H	O	1	0
			10	2	6	2		
2	AAA	1	Total	C	H	O	1	0
			10	2	6	2		
2	BBB	1	Total	C	H	O	1	0
			10	2	6	2		
2	BBB	1	Total	C	H	O	1	0
			10	2	6	2		
2	BBB	1	Total	C	H	O	1	0
			10	2	6	2		
2	BBB	1	Total	C	H	O	1	0
			10	2	6	2		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
4	BBB	1	28	8	14	1	5	3	0

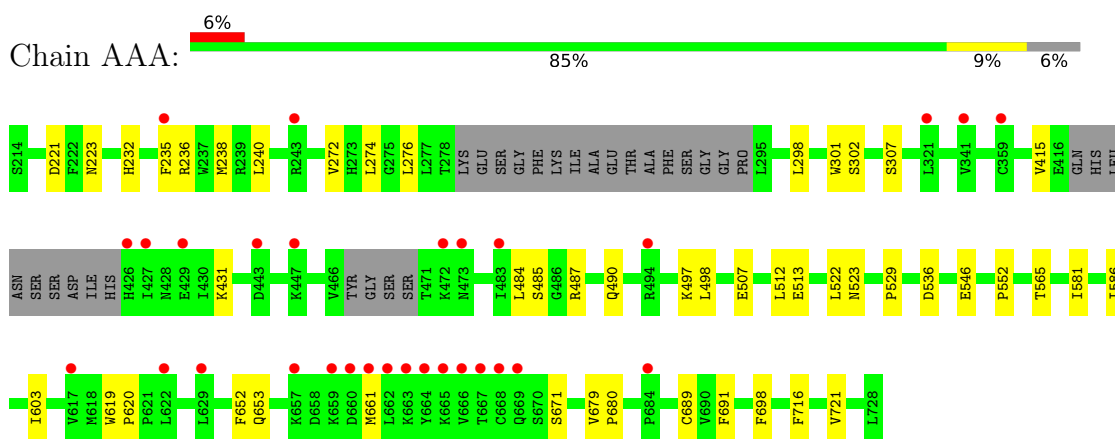
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	AAA	126	Total	O	0	0
			126	126		
5	BBB	161	Total	O	0	0
			161	161		

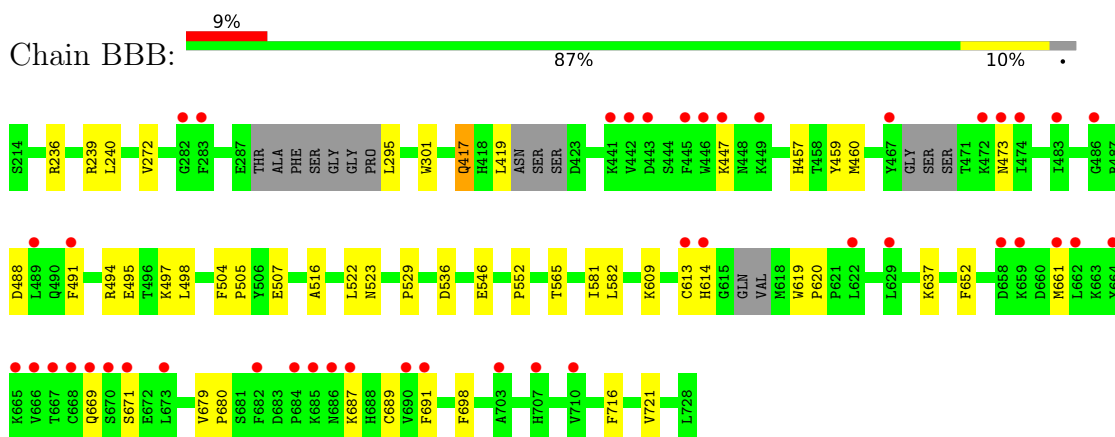
### 3 Residue-property plots [i](#)

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: Alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyltransferase A, Alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyltransferase A



- Molecule 1: Alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyltransferase A, Alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyltransferase A





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	47.00Å 68.17Å 90.99Å 108.36° 92.28° 106.86°	Depositor
Resolution (Å)	43.31 – 1.90 43.31 – 1.90	Depositor EDS
% Data completeness (in resolution range)	96.9 (43.31-1.90) 97.0 (43.31-1.90)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.71 (at 1.89Å)	Xtrriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.194 , 0.237 0.200 , 0.238	Depositor DCC
$R_{free}$ test set	3684 reflections (4.75%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	37.1	Xtrriage
Anisotropy	0.148	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.42 , 47.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	16444	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.73% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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](#)

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

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	AAA	0.67	0/4035	0.79	0/5448
1	BBB	0.68	0/4157	0.80	0/5609
All	All	0.68	0/8192	0.80	0/11057

There are no bond length outliers.

There are no bond angle outliers.

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	AAA	3935	3949	3925	35	0
1	BBB	4054	4061	4034	36	0
2	AAA	12	18	18	0	0
2	BBB	20	30	30	2	0
3	AAA	20	0	0	0	0
3	BBB	30	0	0	3	0
4	BBB	14	14	13	0	0
5	AAA	126	0	0	3	0
5	BBB	161	0	0	2	0
All	All	8372	8072	8020	67	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 4.

The worst 5 of 67 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:AAA:235:PHE:HD1	1:AAA:238:MET:SD	2.03	0.81
1:AAA:235:PHE:CD1	1:AAA:238:MET:SD	2.77	0.78
1:BBB:447:LYS:HA	1:BBB:473:ASN:HD22	1.57	0.68
1:AAA:235:PHE:HA	1:AAA:238:MET:SD	2.37	0.64
1:BBB:239:ARG:NH2	3:BBB:807:SO4:O2	2.30	0.64

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	AAA	478/515 (93%)	467 (98%)	11 (2%)	0	100	100
1	BBB	490/515 (95%)	478 (98%)	12 (2%)	0	100	100
All	All	968/1030 (94%)	945 (98%)	23 (2%)	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	AAA	433/456 (95%)	430 (99%)	3 (1%)	84	84
1	BBB	445/456 (98%)	443 (100%)	2 (0%)	91	91
All	All	878/912 (96%)	873 (99%)	5 (1%)	86	87

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	276	LEU
1	AAA	487	ARG
1	AAA	536	ASP
1	BBB	417	GLN
1	BBB	536	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

19 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
3	SO4	BBB	807	-	4,4,4	0.41	0	6,6,6	0.18	0
3	SO4	AAA	804	-	4,4,4	0.38	0	6,6,6	0.08	0
2	EDO	AAA	802	-	3,3,3	0.18	0	2,2,2	0.11	0
2	EDO	BBB	802	-	3,3,3	0.27	0	2,2,2	0.11	0
2	EDO	BBB	806	-	3,3,3	0.30	0	2,2,2	0.21	0
3	SO4	BBB	811	-	4,4,4	0.37	0	6,6,6	0.15	0
2	EDO	AAA	801	-	3,3,3	0.02	0	2,2,2	0.03	0
4	NAG	BBB	805	1	14,14,15	0.62	0	17,19,21	1.42	4 (23%)
3	SO4	AAA	807	-	4,4,4	0.46	0	6,6,6	0.10	0
3	SO4	BBB	809	-	4,4,4	0.36	0	6,6,6	0.14	0
2	EDO	BBB	803	-	3,3,3	0.37	0	2,2,2	0.30	0
2	EDO	AAA	803	-	3,3,3	0.05	0	2,2,2	0.11	0
3	SO4	AAA	806	-	4,4,4	0.39	0	6,6,6	0.11	0
3	SO4	AAA	805	-	4,4,4	0.38	0	6,6,6	0.07	0
3	SO4	BBB	808	-	4,4,4	0.41	0	6,6,6	0.12	0
2	EDO	BBB	804	-	3,3,3	0.21	0	2,2,2	0.17	0
3	SO4	BBB	810	-	4,4,4	0.39	0	6,6,6	0.21	0
2	EDO	BBB	801	-	3,3,3	0.10	0	2,2,2	0.07	0
3	SO4	BBB	812	-	4,4,4	0.36	0	6,6,6	0.04	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
4	NAG	BBB	805	1	-	2/6/23/26	0/1/1/1
2	EDO	BBB	803	-	-	1/1/1/1	-
2	EDO	AAA	802	-	-	0/1/1/1	-
2	EDO	BBB	801	-	-	1/1/1/1	-
2	EDO	AAA	803	-	-	0/1/1/1	-
2	EDO	BBB	802	-	-	1/1/1/1	-
2	EDO	BBB	804	-	-	0/1/1/1	-
2	EDO	BBB	806	-	-	0/1/1/1	-
2	EDO	AAA	801	-	-	1/1/1/1	-

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	BBB	805	NAG	C4-C3-C2	2.77	115.08	111.02

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	BBB	805	NAG	O5-C5-C6	2.58	111.24	107.20
4	BBB	805	NAG	C1-C2-N2	2.43	114.64	110.49
4	BBB	805	NAG	C2-N2-C7	2.20	126.03	122.90

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	BBB	805	NAG	O5-C5-C6-O6
4	BBB	805	NAG	C4-C5-C6-O6
2	BBB	802	EDO	O1-C1-C2-O2
2	AAA	801	EDO	O1-C1-C2-O2
2	BBB	803	EDO	O1-C1-C2-O2

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	BBB	807	SO4	2	0
2	BBB	803	EDO	2	0
3	BBB	810	SO4	1	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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	AAA	486/515 (94%)	0.46	30 (6%) 20 23	31, 47, 79, 102	0
1	BBB	500/515 (97%)	0.52	44 (8%) 10 11	30, 44, 78, 101	0
All	All	986/1030 (95%)	0.49	74 (7%) 14 15	30, 46, 79, 102	0

The worst 5 of 74 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	445	PHE	7.4
1	AAA	429	GLU	5.8
1	AAA	666	VAL	5.7
1	BBB	685	LYS	5.5
1	BBB	467	TYR	5.5

### 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 monosaccharides 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<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
4	NAG	BBB	805	14/15	0.83	0.17	60,82,89,91	3
3	SO4	AAA	804	5/5	0.84	0.24	62,63,70,70	5
3	SO4	BBB	812	5/5	0.88	0.20	54,59,63,66	5
3	SO4	BBB	808	5/5	0.91	0.13	48,50,59,61	5
2	EDO	BBB	803	4/4	0.91	0.17	42,56,69,71	1
3	SO4	AAA	806	5/5	0.91	0.15	57,73,74,76	5
2	EDO	AAA	801	4/4	0.92	0.18	57,66,67,68	1
2	EDO	BBB	804	4/4	0.93	0.12	47,52,60,64	1
2	EDO	AAA	803	4/4	0.94	0.07	64,65,68,71	1
2	EDO	BBB	801	4/4	0.94	0.17	59,60,62,65	1
3	SO4	AAA	807	5/5	0.94	0.17	43,55,58,65	5
2	EDO	AAA	802	4/4	0.95	0.07	52,54,57,58	1
2	EDO	BBB	802	4/4	0.95	0.20	48,55,56,59	1
2	EDO	BBB	806	4/4	0.95	0.21	44,47,49,52	1
3	SO4	BBB	811	5/5	0.96	0.10	76,78,82,86	0
3	SO4	AAA	805	5/5	0.96	0.09	61,64,68,71	5
3	SO4	BBB	810	5/5	0.96	0.07	55,59,63,70	5
3	SO4	BBB	809	5/5	0.97	0.07	61,64,67,69	5
3	SO4	BBB	807	5/5	0.99	0.10	51,54,60,62	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.