



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 12, 2024 – 10:44 pm BST

PDB ID : 6TVK  
Title : Alpha-L-fucosidase isoenzyme 2 from *Paenibacillus thiaminolyticus*  
Authors : Kovalova, T.; Koval, T.; Lipovova, P.; Dohnalek, J.  
Deposited on : 2020-01-09  
Resolution : 2.10 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](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 : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

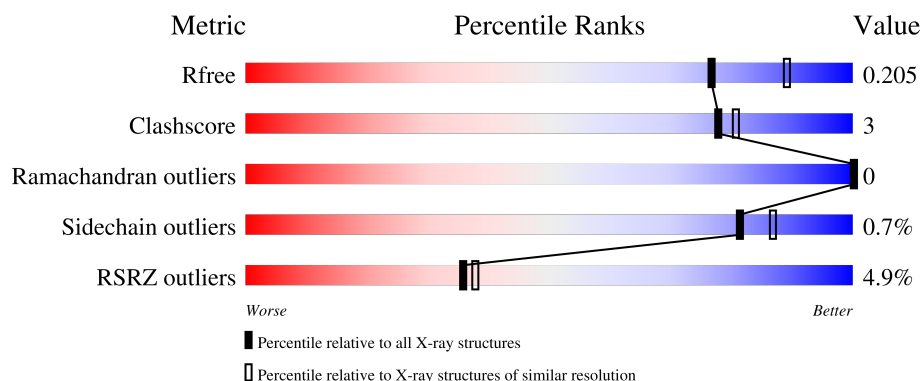
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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))
$R_{free}$	164625	6234 (2.10-2.10)
Clashscore	180529	6893 (2.10-2.10)
Ramachandran outliers	177936	6839 (2.10-2.10)
Sidechain outliers	177891	6840 (2.10-2.10)
RSRZ outliers	164620	6234 (2.10-2.10)

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	689	<div> <div>5%</div> <div>89%</div> <div>7%</div> <div>.</div> </div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 5648 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 Alpha-L-fucosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	660	Total	C	N	O	S	0	3	0
			5249	3335	907	975	32			

There are 29 discrepancies between the modelled and reference sequences:

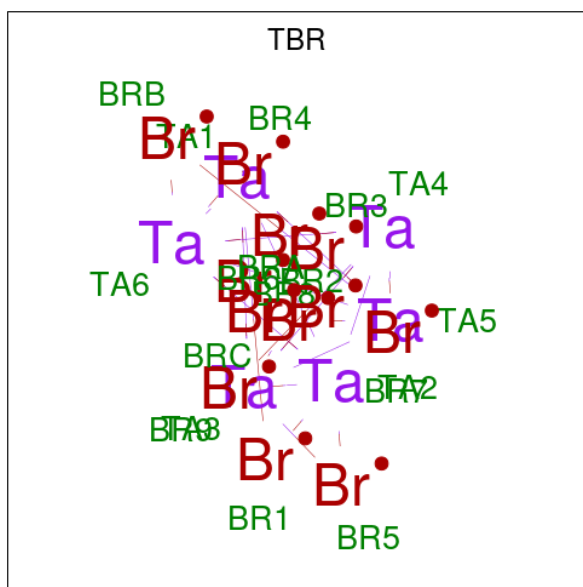
Chain	Residue	Modelled	Actual	Comment	Reference
AAA	-28	MET	-	initiating methionine	UNP K0JCW6
AAA	-27	GLY	-	expression tag	UNP K0JCW6
AAA	-26	HIS	-	expression tag	UNP K0JCW6
AAA	-25	HIS	-	expression tag	UNP K0JCW6
AAA	-24	HIS	-	expression tag	UNP K0JCW6
AAA	-23	HIS	-	expression tag	UNP K0JCW6
AAA	-22	HIS	-	expression tag	UNP K0JCW6
AAA	-21	HIS	-	expression tag	UNP K0JCW6
AAA	-20	HIS	-	expression tag	UNP K0JCW6
AAA	-19	HIS	-	expression tag	UNP K0JCW6
AAA	-18	HIS	-	expression tag	UNP K0JCW6
AAA	-17	HIS	-	expression tag	UNP K0JCW6
AAA	-16	SER	-	expression tag	UNP K0JCW6
AAA	-15	SER	-	expression tag	UNP K0JCW6
AAA	-14	GLY	-	expression tag	UNP K0JCW6
AAA	-13	HIS	-	expression tag	UNP K0JCW6
AAA	-12	ILE	-	expression tag	UNP K0JCW6
AAA	-11	GLU	-	expression tag	UNP K0JCW6
AAA	-10	GLY	-	expression tag	UNP K0JCW6
AAA	-9	ARG	-	expression tag	UNP K0JCW6
AAA	-8	HIS	-	expression tag	UNP K0JCW6
AAA	-7	GLY	-	expression tag	UNP K0JCW6
AAA	-6	GLU	-	expression tag	UNP K0JCW6
AAA	-5	ASN	-	expression tag	UNP K0JCW6
AAA	-4	LEU	-	expression tag	UNP K0JCW6
AAA	-3	TYR	-	expression tag	UNP K0JCW6
AAA	-2	PHE	-	expression tag	UNP K0JCW6

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	-1	GLN	-	expression tag	UNP K0JCW6
AAA	0	GLY	-	expression tag	UNP K0JCW6

- Molecule 2 is HEXATANTALUM DODECABROMIDE (three-letter code: TBR) (formula:  $\text{Br}_{12}\text{Ta}_6$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	AAA	1	Total	Br	Ta	0	0
			18	12	6		

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula:  $\text{Zn}$ ).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	AAA	1	Total	Zn	0	0
			1	1		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula:  $\text{C}_3\text{H}_8\text{O}_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	AAA	1	Total	C	O	0	0
			6	3	3		
4	AAA	1	Total	C	O	0	0
			6	3	3		
4	AAA	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

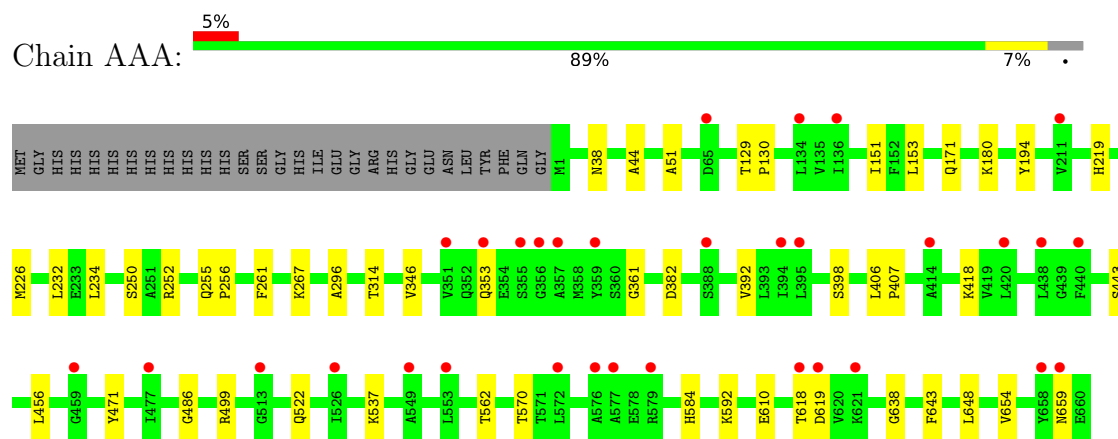
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	AAA	357	Total 357	O 357	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: Alpha-L-fucosidase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 64 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	125.12Å 125.12Å 231.91Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.08 – 2.10 49.08 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.6 (49.08-2.10) 99.6 (49.08-2.10)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.28 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.8.0253	Depositor
R, $R_{free}$	0.189 , 0.225 0.196 , 0.205	Depositor DCC
$R_{free}$ test set	3129 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.2	Xtriage
Anisotropy	0.576	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 50.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.36$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	5648	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.29% 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: SO4, TBR, GOL, CSO, ZN

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	0/5376	0.82	1/7277 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	AAA	38	ASN	CB-CA-C	5.15	120.70	110.40

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	5249	0	5119	28	0
2	AAA	18	0	0	1	0
3	AAA	1	0	0	0	0
4	AAA	18	0	24	0	0
5	AAA	5	0	0	0	0
6	AAA	357	0	0	5	0
All	All	5648	0	5143	28	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 3.

All (28) 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:267:LYS:HG3	1:AAA:314:THR:HG23	1.83	0.61
1:AAA:219:HIS:HB3	6:AAA:1072:HOH:O	2.02	0.59
1:AAA:619:ASP:OD2	1:AAA:659:ASN:ND2	2.39	0.54
1:AAA:261:PHE:CE2	1:AAA:296:ALA:HB2	2.44	0.53
1:AAA:232:LEU:O	1:AAA:261:PHE:HA	2.09	0.52
1:AAA:499:ARG:HG3	6:AAA:852:HOH:O	2.09	0.51
1:AAA:643:PHE:CE1	1:AAA:654:VAL:HG11	2.46	0.51
1:AAA:234:LEU:HD13	1:AAA:250:SER:HB3	1.92	0.50
1:AAA:648:LEU:C	1:AAA:648:LEU:HD23	2.33	0.49
1:AAA:44:ALA:O	1:AAA:51:ALA:HA	2.12	0.49
1:AAA:180:LYS:HE3	6:AAA:1124:HOH:O	2.13	0.49
1:AAA:353:GLN:CD	1:AAA:361:GLY:HA3	2.33	0.49
1:AAA:570:THR:HA	1:AAA:584:HIS:O	2.14	0.47
1:AAA:171:GLN:HG2	2:AAA:701:TBR:BR1	2.69	0.47
1:AAA:456:LEU:HG	1:AAA:486:GLY:HA2	1.98	0.46
1:AAA:522:GLN:NE2	6:AAA:808:HOH:O	2.44	0.46
1:AAA:618:THR:O	1:AAA:638:GLY:HA2	2.17	0.45
1:AAA:562:THR:HA	1:AAA:610:GLU:O	2.17	0.44
1:AAA:346:VAL:HG11	1:AAA:353:GLN:HB3	2.01	0.43
1:AAA:471:TYR:CD2	1:AAA:537:LYS:HG3	2.54	0.42
1:AAA:255:GLN:HB3	1:AAA:256:PRO:HD3	2.01	0.42
1:AAA:406:LEU:HB3	1:AAA:407:PRO:HD3	2.00	0.42
1:AAA:382:ASP:HB2	6:AAA:905:HOH:O	2.20	0.41
1:AAA:392:VAL:HA	1:AAA:418:LYS:O	2.21	0.41
1:AAA:151:ILE:HG22	1:AAA:153:LEU:HG	2.01	0.41
1:AAA:226:MET:HE2	1:AAA:232:LEU:HD11	2.01	0.41
1:AAA:252:ARG:HA	1:AAA:252:ARG:CZ	2.51	0.41
1:AAA:129:THR:HB	1:AAA:130:PRO:CD	2.51	0.40

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	AAA	660/689 (96%)	630 (96%)	30 (4%)	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	AAA	553/574 (96%)	549 (99%)	4 (1%)	81	87

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

Mol	Chain	Res	Type
1	AAA	194	TYR
1	AAA	398	SER
1	AAA	443	SER
1	AAA	592	LYS

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

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

1 non-standard protein/DNA/RNA residue is 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	CSO	AAA	457	1	3,6,7	0.62	0	0,6,8	-	-

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	CSO	AAA	457	1	-	0/1/5/7	-

There are no bond length outliers.

There are no bond angle outliers.

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](#)

Of 6 ligands modelled in this entry, 1 is monoatomic - leaving 5 for Mogul analysis.

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	$\# Z  > 2$
4	GOL	AAA	704	-	5,5,5	0.19	0	5,5,5	0.49	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	TBR	AAA	701	-	0,36,36	-	-	-		
5	SO4	AAA	706	-	4,4,4	0.29	0	6,6,6	0.11	0
4	GOL	AAA	705	-	5,5,5	0.13	0	5,5,5	0.35	0
4	GOL	AAA	703	-	5,5,5	0.14	0	5,5,5	0.38	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	GOL	AAA	704	-	-	1/4/4/4	-
4	GOL	AAA	703	-	-	2/4/4/4	-
4	GOL	AAA	705	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	AAA	703	GOL	C1-C2-C3-O3
4	AAA	703	GOL	O2-C2-C3-O3
4	AAA	704	GOL	O1-C1-C2-C3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	AAA	701	TBR	1	0

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

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<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	659/689 (95%)	0.41	32 (4%) 36 38	31, 53, 75, 116	3 (0%)

All (32) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	357	ALA	5.1
1	AAA	577	ALA	4.9
1	AAA	359	TYR	4.3
1	AAA	394	ILE	3.8
1	AAA	356	GLY	3.5
1	AAA	420	LEU	3.4
1	AAA	513	GLY	3.2
1	AAA	136	ILE	3.2
1	AAA	353	GLN	3.1
1	AAA	576	ALA	3.0
1	AAA	579	ARG	2.9
1	AAA	549	ALA	2.9
1	AAA	355	SER	2.7
1	AAA	440	PHE	2.7
1	AAA	477	ILE	2.6
1	AAA	659	ASN	2.5
1	AAA	134	LEU	2.5
1	AAA	459	GLY	2.5
1	AAA	395	LEU	2.5
1	AAA	658	TYR	2.3
1	AAA	618	THR	2.2
1	AAA	553	LEU	2.2
1	AAA	388	SER	2.2
1	AAA	414	ALA	2.2
1	AAA	65	ASP	2.2
1	AAA	438	LEU	2.1
1	AAA	526	ILE	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	AAA	572	LEU	2.1
1	AAA	619	ASP	2.1
1	AAA	621	LYS	2.0
1	AAA	211	VAL	2.0
1	AAA	351	VAL	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<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
1	CSO	AAA	457	7/8	0.92	0.11	70,73,81,82	0

## 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
5	SO4	AAA	706	5/5	0.76	0.14	93,95,118,118	0
2	TBR	AAA	701	18/18	0.77	0.21	112,122,137,139	18
4	GOL	AAA	703	6/6	0.91	0.11	64,73,76,77	0
4	GOL	AAA	704	6/6	0.93	0.11	50,53,56,63	0
4	GOL	AAA	705	6/6	0.94	0.12	63,74,78,78	0
3	ZN	AAA	702	1/1	1.00	0.04	49,49,49,49	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.