



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 9, 2023 – 02:05 pm GMT

PDB ID : 2JJB  
Title : Family 37 trehalase from Escherichia coli in complex with casuarine-6- O-alpha-glucopyranose  
Authors : Gloster, T.M.; Roberts, S.; Davies, G.J.; Cardona, F.; Parmeggiani, C.; Bonaccini, C.; Gratteri, P.; Sim, L.; Rose, D.R.; Goti, A.  
Deposited on : 2008-03-28  
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

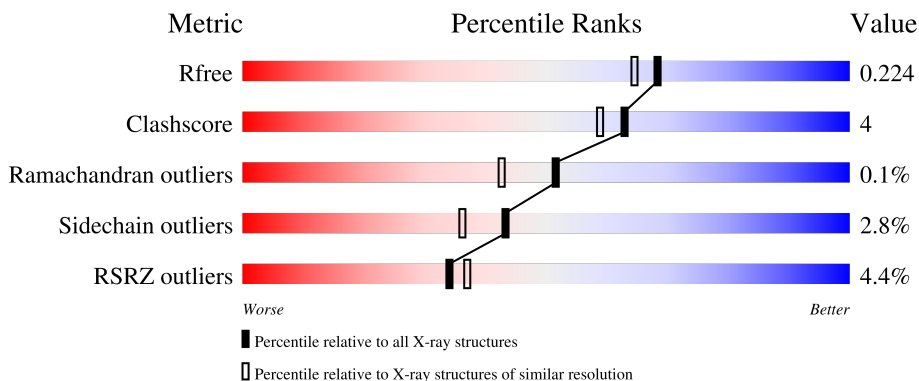
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

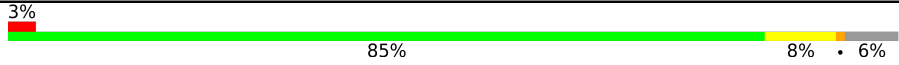



The reported resolution of this entry is 1.90 Å.

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	A	535	 3% 85% 8% • 6%
1	B	535	 3% 84% 7% • 8%
1	C	535	 4% 84% 9% • 7%
1	D	535	 7% 83% 10% 7%

## 2 Entry composition [i](#)

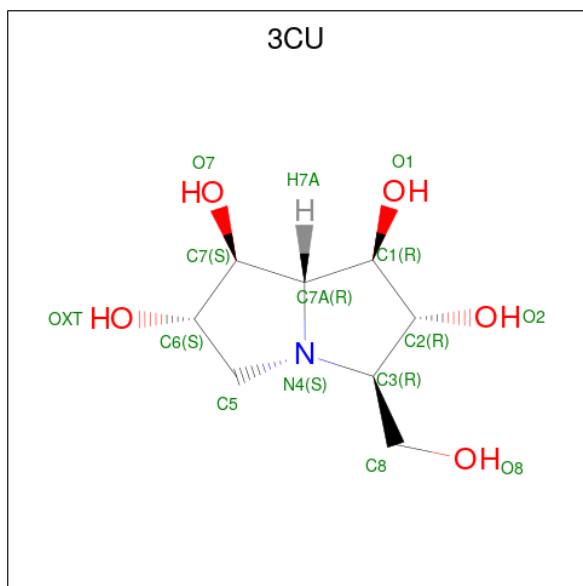
There are 6 unique types of molecules in this entry. The entry contains 18411 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 PERIPLASMIC TREHALASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	504	Total 4155	C 2652	N 704	O 781	S 18	0	26	0
1	B	493	Total 4045	C 2575	N 690	O 762	S 18	0	21	0
1	C	500	Total 4070	C 2589	N 693	O 771	S 17	0	16	0
1	D	500	Total 4054	C 2587	N 681	O 768	S 18	0	18	0

- Molecule 2 is CASUARINE (three-letter code: 3CU) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>5</sub>).



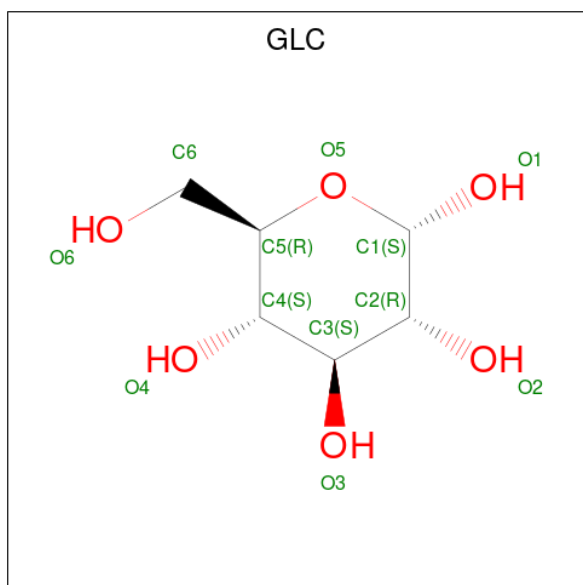
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	Total 13	C 8	N 1	O 4	0	0
2	B	1	Total 13	C 8	N 1	O 4	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	C	1	Total	C	N	O	0	0
			13	8	1	4		
2	D	1	Total	C	N	O	0	0
			13	8	1	4		

- Molecule 3 is alpha-D-glucopyranose (three-letter code: GLC) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			12	6	6		
3	B	1	Total	C	O	0	0
			12	6	6		
3	C	1	Total	C	O	0	0
			12	6	6		
3	D	1	Total	C	O	0	0
			12	6	6		

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



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

*Continued on next page...*

*Continued from previous page...*

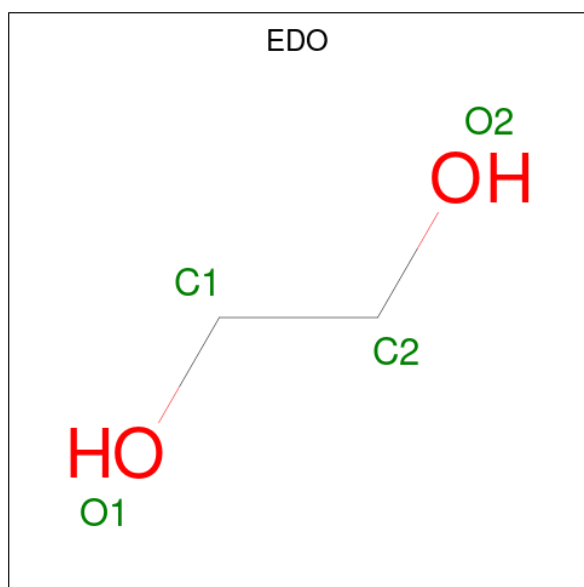
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
4	B	1	5	4	1	0	0
4	B	1	5	4	1	0	0
4	B	1	5	4	1	0	0
4	B	1	5	4	1	0	0
4	B	1	5	4	1	0	0
4	C	1	5	4	1	0	0
4	C	1	5	4	1	0	0
4	C	1	5	4	1	0	0
4	C	1	5	4	1	0	0
4	C	1	5	4	1	0	0
4	C	1	5	4	1	0	0
4	C	1	5	4	1	0	0
4	C	1	5	4	1	0	0
4	C	1	5	4	1	0	0
4	C	1	5	4	1	0	0
4	C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

*Continued on next page...*

Continued from previous page...

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

- Molecule 5 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
5	A	1	Total C O 4 2 2	0	0
5	B	1	Total C O 4 2 2	0	0
5	C	1	Total C O 4 2 2	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	517	Total O 523 523	0	26
6	B	460	Total O 461 461	0	1
6	C	454	Total O 458 458	0	6

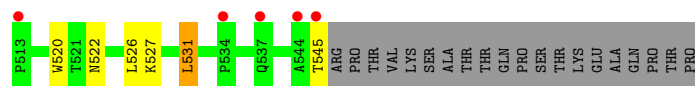
Continued on next page...

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
6	D	339	Total 343	O 343	0	4

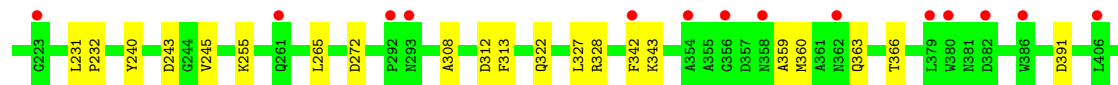
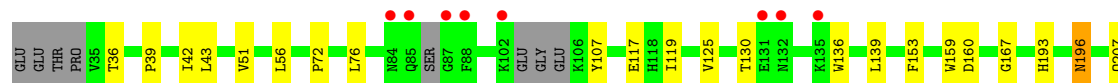






● Molecule 1: PERIPLASMIC TREHALASE

Chain D: 7% 83% 10% 7%



LYS  
SER  
ALA  
THR  
THR  
GLN  
PRO  
SER  
THR  
THR  
GLU  
LYS  
ALA  
GLN  
PRO  
THR  
PRO

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	92.48Å 117.86Å 203.45Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.92 – 1.90 19.92 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.5 (19.92-1.90) 99.5 (19.92-1.90)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.02 (at 1.90Å)	Xtrriage
Refinement program	REFMAC 5.4.0065	Depositor
R, $R_{free}$	0.177 , 0.224 0.177 , 0.224	Depositor DCC
$R_{free}$ test set	8658 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.0	Xtrriage
Anisotropy	0.578	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 74.3	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	18411	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.43% 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, GLC, EDO, 3CU

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	0.63	0/4340	0.68	0/5912
1	B	0.61	0/4215	0.65	0/5740
1	C	0.60	0/4229	0.66	0/5765
1	D	0.53	0/4216	0.62	0/5758
All	All	0.59	0/17000	0.65	0/23175

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	A	4155	0	4037	43	0
1	B	4045	0	3915	28	0
1	C	4070	0	3907	38	0
1	D	4054	0	3847	37	0
2	A	13	0	13	1	0
2	B	13	0	13	1	0
2	C	13	0	13	1	0
2	D	13	0	13	1	0
3	A	12	0	11	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	12	0	11	0	0
3	C	12	0	11	0	0
3	D	12	0	11	0	0
4	A	60	0	0	2	0
4	B	35	0	0	1	0
4	C	55	0	0	1	0
4	D	40	0	0	1	0
5	A	4	0	6	0	0
5	B	4	0	6	0	0
5	C	4	0	6	0	0
6	A	523	0	0	8	0
6	B	461	0	0	5	0
6	C	458	0	0	10	0
6	D	343	0	0	6	0
All	All	18411	0	15820	139	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 139 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:A:38[B]:GLN:OE1	6:A:2003[B]:HOH:O	1.57	1.22
1:A:176[B]:LYS:NZ	6:A:2203[B]:HOH:O	1.83	1.11
1:D:366[B]:THR:HG22	6:D:2254:HOH:O	1.51	1.07
1:A:38[B]:GLN:NE2	6:A:2003[B]:HOH:O	1.92	1.02
1:D:117[B]:GLU:OE2	6:D:2081:HOH:O	1.77	1.02

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	522/535 (98%)	507 (97%)	15 (3%)	0	100	100
1	B	508/535 (95%)	494 (97%)	13 (3%)	1 (0%)	47	38
1	C	512/535 (96%)	499 (98%)	13 (2%)	0	100	100
1	D	510/535 (95%)	495 (97%)	14 (3%)	1 (0%)	47	38
All	All	2052/2140 (96%)	1995 (97%)	55 (3%)	2 (0%)	51	42

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	519	GLY
1	B	519	GLY

### 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	449/456 (98%)	436 (97%)	13 (3%)	42	35
1	B	436/456 (96%)	418 (96%)	18 (4%)	30	21
1	C	436/456 (96%)	423 (97%)	13 (3%)	41	33
1	D	428/456 (94%)	419 (98%)	9 (2%)	53	48
All	All	1749/1824 (96%)	1696 (97%)	53 (3%)	43	33

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	417	ASP
1	C	176	LYS
1	D	322[B]	GLN
1	B	520	TRP
1	C	76	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	85	GLN
1	D	249	GLN
1	D	369	ASN
1	D	293	ASN
1	B	113	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 monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

49 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
4	SO4	A	1560	-	4,4,4	0.13	0	6,6,6	0.20	0
4	SO4	C	1554	-	4,4,4	0.13	0	6,6,6	0.13	0
2	3CU	D	1547	3	14,14,15	0.82	0	18,21,23	3.39	7 (38%)
4	SO4	A	1553	-	4,4,4	0.17	0	6,6,6	0.23	0
4	SO4	C	1548	-	4,4,4	0.17	0	6,6,6	0.44	0
5	EDO	C	1559	-	3,3,3	0.55	0	2,2,2	0.07	0
2	3CU	A	1547	3	14,14,15	0.89	0	18,21,23	3.20	7 (38%)
3	GLC	D	1548	2	12,12,12	0.73	0	17,17,17	1.00	2 (11%)
2	3CU	B	1545	3	14,14,15	0.91	0	18,21,23	3.34	7 (38%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	SO4	A	1559	-	4,4,4	0.16	0	6,6,6	0.09	0
4	SO4	C	1553	-	4,4,4	0.13	0	6,6,6	0.18	0
5	EDO	A	1561	-	3,3,3	0.48	0	2,2,2	0.51	0
4	SO4	A	1556	-	4,4,4	0.13	0	6,6,6	0.15	0
4	SO4	B	1552	-	4,4,4	0.12	0	6,6,6	0.13	0
5	EDO	B	1554	-	3,3,3	0.48	0	2,2,2	0.92	0
4	SO4	B	1547	-	4,4,4	0.20	0	6,6,6	0.37	0
4	SO4	A	1558[A]	-	4,4,4	0.22	0	6,6,6	0.50	0
4	SO4	A	1554	-	4,4,4	0.13	0	6,6,6	0.17	0
4	SO4	B	1548	-	4,4,4	0.31	0	6,6,6	0.36	0
3	GLC	C	1547	2	12,12,12	0.62	0	17,17,17	1.07	1 (5%)
3	GLC	B	1546	2	12,12,12	0.57	0	17,17,17	1.09	1 (5%)
4	SO4	A	1552	-	4,4,4	0.17	0	6,6,6	0.20	0
4	SO4	C	1557	-	4,4,4	0.13	0	6,6,6	0.13	0
4	SO4	D	1549	-	4,4,4	0.19	0	6,6,6	0.36	0
4	SO4	A	1557	-	4,4,4	0.16	0	6,6,6	0.23	0
4	SO4	C	1556	-	4,4,4	0.13	0	6,6,6	0.21	0
4	SO4	C	1549	-	4,4,4	0.27	0	6,6,6	0.26	0
3	GLC	A	1548	2	12,12,12	0.66	0	17,17,17	1.07	1 (5%)
4	SO4	D	1556	-	4,4,4	0.16	0	6,6,6	0.19	0
4	SO4	A	1549	-	4,4,4	0.20	0	6,6,6	0.33	0
4	SO4	A	1551	-	4,4,4	0.10	0	6,6,6	0.35	0
4	SO4	A	1550	-	4,4,4	0.13	0	6,6,6	0.27	0
4	SO4	B	1549	-	4,4,4	0.15	0	6,6,6	0.35	0
4	SO4	B	1553	-	4,4,4	0.23	0	6,6,6	0.23	0
4	SO4	D	1551	-	4,4,4	0.11	0	6,6,6	0.30	0
4	SO4	A	1555	-	4,4,4	0.12	0	6,6,6	0.18	0
4	SO4	B	1551	-	4,4,4	0.12	0	6,6,6	0.23	0
4	SO4	D	1553	-	4,4,4	0.12	0	6,6,6	0.41	0
2	3CU	C	1546	3	14,14,15	0.92	1 (7%)	18,21,23	3.39	8 (44%)
4	SO4	C	1551	-	4,4,4	0.22	0	6,6,6	0.46	0
4	SO4	D	1550	-	4,4,4	0.18	0	6,6,6	0.38	0
4	SO4	C	1555	-	4,4,4	0.14	0	6,6,6	0.28	0
4	SO4	D	1555	-	4,4,4	0.14	0	6,6,6	0.13	0
4	SO4	C	1558	-	4,4,4	0.13	0	6,6,6	0.16	0
4	SO4	D	1554	-	4,4,4	0.14	0	6,6,6	0.10	0
4	SO4	C	1550	-	4,4,4	0.20	0	6,6,6	0.59	0
4	SO4	C	1552	-	4,4,4	0.16	0	6,6,6	0.41	0
4	SO4	D	1552	-	4,4,4	0.22	0	6,6,6	0.21	0
4	SO4	B	1550	-	4,4,4	0.13	0	6,6,6	0.24	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	3CU	D	1547	3	-	0/2/31/34	0/2/2/2
2	3CU	C	1546	3	-	0/2/31/34	0/2/2/2
5	EDO	C	1559	-	-	1/1/1/1	-
2	3CU	A	1547	3	-	0/2/31/34	0/2/2/2
3	GLC	A	1548	2	-	0/2/22/22	0/1/1/1
3	GLC	D	1548	2	-	0/2/22/22	0/1/1/1
5	EDO	B	1554	-	-	0/1/1/1	-
2	3CU	B	1545	3	-	0/2/31/34	0/2/2/2
3	GLC	C	1547	2	-	0/2/22/22	0/1/1/1
3	GLC	B	1546	2	-	0/2/22/22	0/1/1/1
5	EDO	A	1561	-	-	0/1/1/1	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1546	3CU	C7A-N4	-2.30	1.44	1.49

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	1546	3CU	C7A-N4-C3	9.80	119.67	107.17
2	B	1545	3CU	C7A-N4-C3	9.70	119.54	107.17
2	D	1547	3CU	C7A-N4-C3	9.40	119.16	107.17
2	A	1547	3CU	C7A-N4-C3	8.86	118.47	107.17
2	B	1545	3CU	C5-N4-C7A	5.65	116.03	106.10

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	1559	EDO	O1-C1-C2-O2

There are no ring outliers.

9 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	1547	3CU	1	0
2	A	1547	3CU	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1545	3CU	1	0
4	A	1552	SO4	1	0
4	A	1551	SO4	1	0
4	B	1549	SO4	1	0
2	C	1546	3CU	1	0
4	C	1552	SO4	1	0
4	D	1552	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	A	504/535 (94%)	-0.15	14 (2%) 53 56	14, 22, 36, 46	0
1	B	493/535 (92%)	-0.06	16 (3%) 47 50	16, 24, 39, 52	0
1	C	500/535 (93%)	-0.02	21 (4%) 36 39	16, 25, 40, 47	0
1	D	500/535 (93%)	0.36	37 (7%) 14 16	18, 32, 49, 54	0
All	All	1997/2140 (93%)	0.03	88 (4%) 34 37	14, 26, 43, 54	0

The worst 5 of 88 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	513	PRO	5.9
1	D	356	GLY	5.7
1	D	544	ALA	4.9
1	A	132	ASN	4.8
1	C	104	GLY	4.7

### 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	SO4	A	1556	5/5	0.74	0.20	85,85,86,86	0
4	SO4	D	1554	5/5	0.76	0.18	92,92,93,93	0
4	SO4	A	1558[A]	5/5	0.82	0.28	26,27,32,32	5
4	SO4	A	1560	5/5	0.83	0.16	88,89,89,89	0
4	SO4	A	1557	5/5	0.83	0.25	70,71,71,72	0
4	SO4	C	1557	5/5	0.84	0.26	87,88,88,88	0
4	SO4	C	1558	5/5	0.88	0.17	88,88,88,88	0
4	SO4	C	1556	5/5	0.88	0.18	64,64,65,65	0
4	SO4	D	1551	5/5	0.91	0.15	49,50,51,52	0
4	SO4	A	1555	5/5	0.91	0.22	71,71,72,72	0
4	SO4	A	1554	5/5	0.92	0.23	64,64,65,65	0
5	EDO	B	1554	4/4	0.92	0.14	27,28,29,33	0
4	SO4	A	1559	5/5	0.93	0.19	84,85,85,85	0
4	SO4	B	1553	5/5	0.93	0.20	53,55,56,56	0
4	SO4	C	1553	5/5	0.93	0.21	69,70,70,71	0
4	SO4	B	1551	5/5	0.94	0.12	66,66,67,67	0
4	SO4	A	1553	5/5	0.94	0.23	57,58,59,60	0
4	SO4	D	1555	5/5	0.94	0.13	77,77,78,78	0
5	EDO	A	1561	4/4	0.94	0.12	26,30,32,33	0
4	SO4	A	1551	5/5	0.94	0.14	36,40,41,44	0
4	SO4	C	1554	5/5	0.95	0.11	72,72,73,73	0
4	SO4	D	1553	5/5	0.95	0.16	53,53,55,55	0
4	SO4	B	1552	5/5	0.95	0.19	63,63,64,64	0
4	SO4	B	1549	5/5	0.96	0.10	37,38,38,39	0
2	3CU	D	1547	13/14	0.96	0.08	18,20,22,22	0
4	SO4	C	1555	5/5	0.96	0.21	55,56,57,58	0
2	3CU	A	1547	13/14	0.97	0.09	12,15,16,16	0
2	3CU	C	1546	13/14	0.97	0.09	15,17,19,19	0
4	SO4	A	1552	5/5	0.97	0.10	44,45,45,46	0
4	SO4	C	1552	5/5	0.97	0.09	38,38,40,40	0
4	SO4	D	1556	5/5	0.97	0.13	61,61,63,63	0
4	SO4	B	1550	5/5	0.97	0.07	67,67,67,68	0
4	SO4	D	1550	5/5	0.97	0.10	53,53,55,56	0
5	EDO	C	1559	4/4	0.97	0.08	23,25,25,26	0
4	SO4	C	1549	5/5	0.98	0.07	31,32,33,35	0
4	SO4	C	1551	5/5	0.98	0.09	41,42,43,46	0
4	SO4	D	1552	5/5	0.98	0.08	43,44,45,46	0
4	SO4	B	1547	5/5	0.98	0.10	29,31,33,35	0
2	3CU	B	1545	13/14	0.98	0.10	16,19,21,21	0
3	GLC	A	1548	12/12	0.98	0.11	12,15,16,16	0
3	GLC	B	1546	12/12	0.98	0.09	16,19,20,23	0
3	GLC	C	1547	12/12	0.98	0.07	16,18,20,22	0
4	SO4	A	1550	5/5	0.98	0.07	26,29,32,33	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	SO4	C	1548	5/5	0.98	0.07	30,32,33,35	0
4	SO4	A	1549	5/5	0.99	0.07	23,25,25,27	0
4	SO4	B	1548	5/5	0.99	0.05	26,27,28,29	0
3	GLC	D	1548	12/12	0.99	0.07	18,19,20,20	0
4	SO4	C	1550	5/5	0.99	0.04	20,21,22,23	0
4	SO4	D	1549	5/5	0.99	0.05	29,32,33,33	0

## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.