wwPDB X-ray Structure Validation Summary Report (i)

Aug 26, 2023 – 10:16 PM EDT

PDB ID	:	3GLB
Title	:	Crystal structure of the effector binding domain of a CATM variant (R156H)
Authors	:	Ezezika, O.C.; Craven, S.H.; Neidle, E.L.; Momany, C.
Deposited on	:	2009-03-11
Resolution	:	2.80 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 (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:

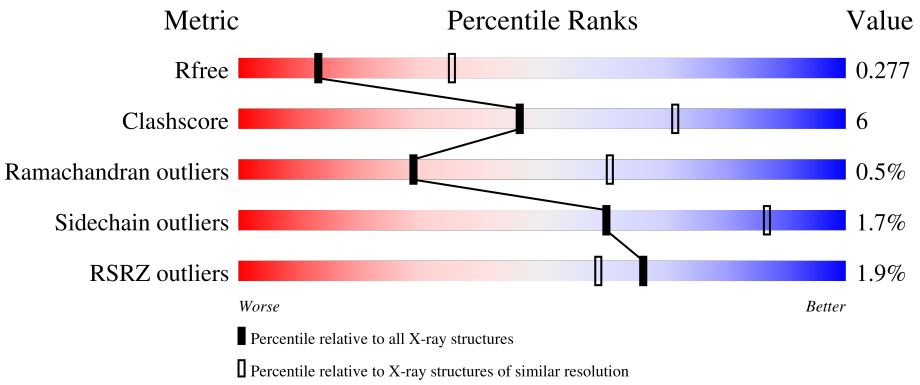
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	225	79% 16% • •
1	B	225	3% 84% 9% 6%
1	C	225	83% 11% 6%
1	D	225	4% 81% 16% •

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	CCU	B	2	-	-	X	-
3	GOL	B	5	-	-	-	X

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 7101 atoms, of which 2 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 HTH-type transcriptional regulator catM.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	215	1705	1103	2	290	304	6	0	1	0
			Total	C		N	O	S			
1	B	212	1671	1084		285	296	6	0	0	0
			Total	C		N	O	S			
1	C	212	1669	1081		285	297	6	0	0	0
			Total	C		N	O	S			
1	D	219	1725	1118		292	309	6	0	0	0

There are 44 discrepancies between the modelled and reference sequences:

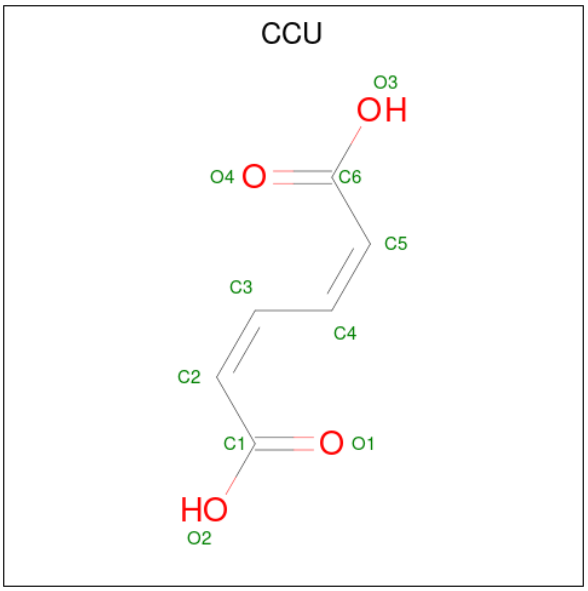
Chain	Residue	Modelled	Actual	Comment	Reference
A	156	HIS	ARG	engineered mutation	UNP P07774
A	304	SER	-	expression tag	UNP P07774
A	305	ILE	-	expression tag	UNP P07774
A	306	LEU	-	expression tag	UNP P07774
A	307	GLU	-	expression tag	UNP P07774
A	308	HIS	-	expression tag	UNP P07774
A	309	HIS	-	expression tag	UNP P07774
A	310	HIS	-	expression tag	UNP P07774
A	311	HIS	-	expression tag	UNP P07774
A	312	HIS	-	expression tag	UNP P07774
A	313	HIS	-	expression tag	UNP P07774
B	156	HIS	ARG	engineered mutation	UNP P07774
B	304	SER	-	expression tag	UNP P07774
B	305	ILE	-	expression tag	UNP P07774
B	306	LEU	-	expression tag	UNP P07774
B	307	GLU	-	expression tag	UNP P07774
B	308	HIS	-	expression tag	UNP P07774
B	309	HIS	-	expression tag	UNP P07774
B	310	HIS	-	expression tag	UNP P07774
B	311	HIS	-	expression tag	UNP P07774
B	312	HIS	-	expression tag	UNP P07774

Continued on next page...

Continued from previous page...

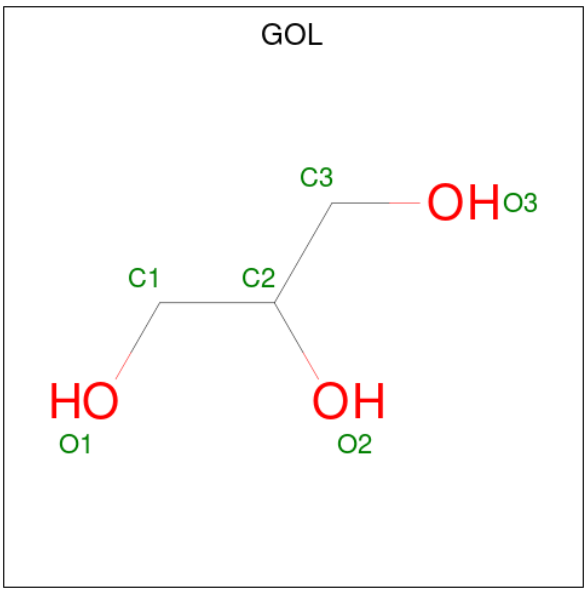
Chain	Residue	Modelled	Actual	Comment	Reference
B	313	HIS	-	expression tag	UNP P07774
C	156	HIS	ARG	engineered mutation	UNP P07774
C	304	SER	-	expression tag	UNP P07774
C	305	ILE	-	expression tag	UNP P07774
C	306	LEU	-	expression tag	UNP P07774
C	307	GLU	-	expression tag	UNP P07774
C	308	HIS	-	expression tag	UNP P07774
C	309	HIS	-	expression tag	UNP P07774
C	310	HIS	-	expression tag	UNP P07774
C	311	HIS	-	expression tag	UNP P07774
C	312	HIS	-	expression tag	UNP P07774
C	313	HIS	-	expression tag	UNP P07774
D	156	HIS	ARG	engineered mutation	UNP P07774
D	304	SER	-	expression tag	UNP P07774
D	305	ILE	-	expression tag	UNP P07774
D	306	LEU	-	expression tag	UNP P07774
D	307	GLU	-	expression tag	UNP P07774
D	308	HIS	-	expression tag	UNP P07774
D	309	HIS	-	expression tag	UNP P07774
D	310	HIS	-	expression tag	UNP P07774
D	311	HIS	-	expression tag	UNP P07774
D	312	HIS	-	expression tag	UNP P07774
D	313	HIS	-	expression tag	UNP P07774

- Molecule 2 is (2Z,4Z)-HEXA-2,4-DIENEDIOIC ACID (three-letter code: CCU) (formula: $C_6H_6O_4$).



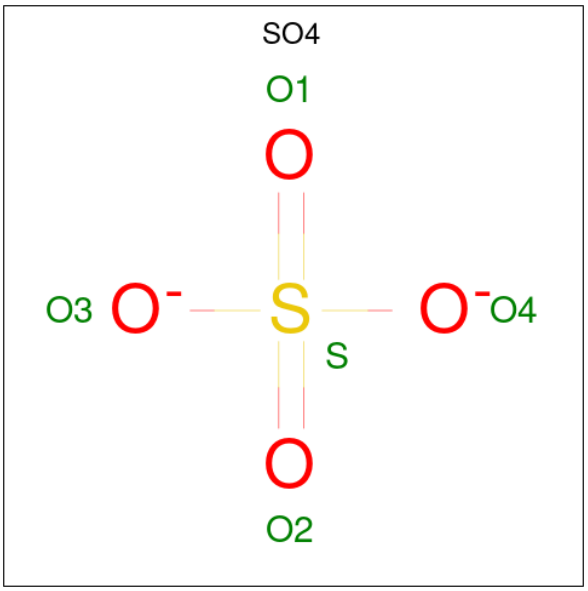
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	10	6	4	0	0
2	B	1	10	6	4	0	0
2	C	1	10	6	4	0	0
2	D	1	10	6	4	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	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

- Molecule 5 is water.

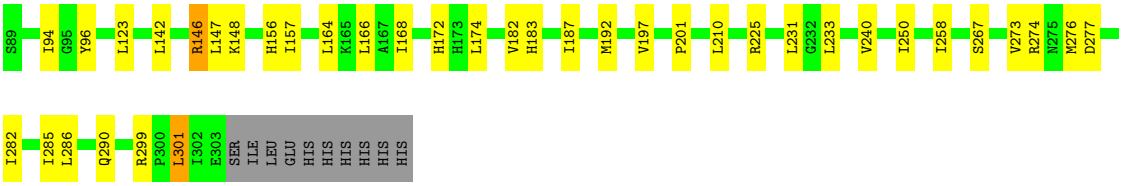
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	67	Total O 67 67	0	0
5	B	73	Total O 73 73	0	0
5	C	61	Total O 61 61	0	0
5	D	45	Total O 45 45	0	0

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.

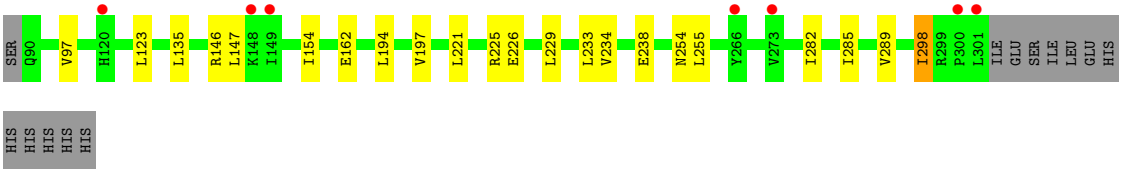
• Molecule 1: HTH-type transcriptional regulator catM

Chain A:



• Molecule 1: HTH-type transcriptional regulator catM

Chain B:

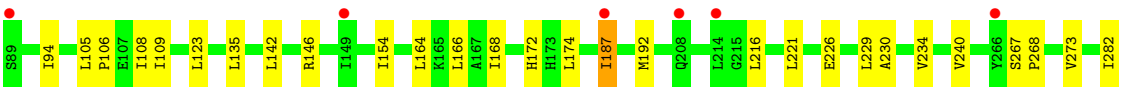


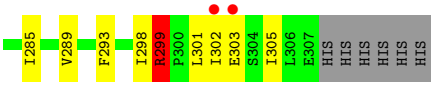
• Molecule 1: HTH-type transcriptional regulator catM

Chain C:

• Molecule 1: HTH-type transcriptional regulator catM

Chain D:




I285
V289
F293
I298
R299
P300
L301
I302
E303
S304
I305
L306
E307
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	110.69Å 115.41Å 76.07Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.98 – 2.80 45.97 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.8 (45.98-2.80) 99.8 (45.97-2.80)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.12 (at 2.81Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.205 , 0.279 0.209 , 0.277	Depositor DCC
R_{free} test set	1218 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å^2)	62.9	Xtriage
Anisotropy	0.460	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 54.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.018 for k,h,-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	7101	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.90% 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 (i)

5.1 Standard geometry (i)

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

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	0.32	0/1739	0.51	0/2362
1	B	0.35	0/1707	0.49	0/2319
1	C	0.32	0/1705	0.48	0/2316
1	D	0.32	0/1761	0.50	0/2392
All	All	0.33	0/6912	0.49	0/9389

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	D	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	299	ARG	Peptide

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	1703	2	1761	32	0
1	B	1671	0	1732	17	0
1	C	1669	0	1726	16	0
1	D	1725	0	1787	23	0
2	A	10	0	4	0	0
2	B	10	0	4	4	0
2	C	10	0	4	1	0
2	D	10	0	4	1	0
3	A	6	0	8	0	0
3	B	6	0	8	0	0
3	C	6	0	8	0	0
3	D	12	0	16	0	0
4	B	15	0	0	1	0
5	A	67	0	0	0	0
5	B	73	0	0	0	0
5	C	61	0	0	0	0
5	D	45	0	0	0	0
All	All	7099	2	7062	87	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 6.

The worst 5 of 87 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:298:ILE:CD1	1:B:298:ILE:CG1	1.82	1.56
1:A:168:ILE:HD13	1:A:258:ILE:HD11	1.56	0.88
1:A:290:GLN:HE22	1:A:301:LEU:HD22	1.38	0.84
1:A:197:VAL:HG22	1:A:225:ARG:HA	1.60	0.83
1:A:182:VAL:HG23	1:A:258:ILE:HG21	1.68	0.76

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	214/225 (95%)	209 (98%)	5 (2%)	0	100	100
1	B	210/225 (93%)	204 (97%)	5 (2%)	1 (0%)	29	61
1	C	210/225 (93%)	200 (95%)	9 (4%)	1 (0%)	29	61
1	D	217/225 (96%)	206 (95%)	9 (4%)	2 (1%)	17	46
All	All	851/900 (95%)	819 (96%)	28 (3%)	4 (0%)	29	61

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	299	ARG
1	C	275	ASN
1	D	187	ILE
1	B	298	ILE

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	189/198 (96%)	184 (97%)	5 (3%)	46	79
1	B	185/198 (93%)	184 (100%)	1 (0%)	88	96
1	C	185/198 (93%)	183 (99%)	2 (1%)	73	92
1	D	192/198 (97%)	187 (97%)	5 (3%)	46	79
All	All	751/792 (95%)	738 (98%)	13 (2%)	60	87

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	299	ARG
1	D	216	LEU
1	D	305	ILE
1	D	299	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	301	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	C	208	GLN
1	D	133	ASN
1	D	228	GLN
1	D	137	GLN
1	A	290	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)

12 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	CCU	D	4	-	9,9,9	1.09	1 (11%)	10,10,10	1.08	1 (10%)
4	SO4	B	11	-	4,4,4	0.13	0	6,6,6	0.07	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	SO4	B	13	-	4,4,4	0.13	0	6,6,6	0.17	0
3	GOL	D	314	-	5,5,5	0.37	0	5,5,5	0.21	0
3	GOL	B	5	-	5,5,5	0.38	0	5,5,5	0.21	0
4	SO4	B	12	-	4,4,4	0.37	0	6,6,6	0.13	0
2	CCU	B	2	-	9,9,9	1.09	1 (11%)	10,10,10	0.87	0
2	CCU	A	3	-	9,9,9	1.07	0	10,10,10	1.48	2 (20%)
2	CCU	C	1	-	9,9,9	1.03	0	10,10,10	1.20	1 (10%)
3	GOL	C	314	-	5,5,5	0.37	0	5,5,5	0.19	0
3	GOL	A	2	-	5,5,5	0.38	0	5,5,5	0.32	0
3	GOL	D	3	-	5,5,5	0.36	0	5,5,5	0.26	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	CCU	D	4	-	-	0/7/7/7	-
3	GOL	D	314	-	-	2/4/4/4	-
3	GOL	B	5	-	-	2/4/4/4	-
2	CCU	B	2	-	-	0/7/7/7	-
2	CCU	A	3	-	-	0/7/7/7	-
2	CCU	C	1	-	-	0/7/7/7	-
3	GOL	C	314	-	-	2/4/4/4	-
3	GOL	A	2	-	-	4/4/4/4	-
3	GOL	D	3	-	-	2/4/4/4	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	2	CCU	O3-C6	-2.09	1.24	1.30
2	D	4	CCU	O2-C1	-2.01	1.25	1.30

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	3	CCU	C3-C2-C1	2.61	128.19	123.66
2	C	1	CCU	C3-C2-C1	2.24	127.55	123.66
2	D	4	CCU	C3-C2-C1	2.22	127.52	123.66
2	A	3	CCU	C4-C5-C6	2.22	127.52	123.66

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	314	GOL	O1-C1-C2-O2
3	C	314	GOL	O1-C1-C2-C3
3	D	3	GOL	O1-C1-C2-O2
3	A	2	GOL	O1-C1-C2-C3
3	B	5	GOL	C1-C2-C3-O3

There are no ring outliers.

4 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	4	CCU	1	0
4	B	13	SO4	1	0
2	B	2	CCU	4	0
2	C	1	CCU	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, 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	215/225 (95%)	0.16	0	100	100	43, 52, 56, 59	0
1	B	212/225 (94%)	0.22	7 (3%)	46	36	43, 51, 56, 60	0
1	C	212/225 (94%)	0.14	1 (0%)	91	88	43, 52, 56, 59	0
1	D	219/225 (97%)	0.22	8 (3%)	41	31	43, 52, 57, 61	0
All	All	858/900 (95%)	0.19	16 (1%)	66	59	43, 52, 56, 61	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	301	LEU	13.2
1	D	266	TYR	3.8
1	B	300	PRO	3.8
1	D	149	ILE	3.1
1	B	148	LYS	3.0

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, 95th 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
4	SO4	B	11	5/5	0.53	0.31	145,146,146,146	0
3	GOL	B	5	6/6	0.71	0.50	83,84,85,85	0
3	GOL	C	314	6/6	0.72	0.28	93,95,95,98	0
4	SO4	B	12	5/5	0.83	0.23	118,120,121,121	0
2	CCU	A	3	10/10	0.88	0.26	56,60,61,63	0
2	CCU	D	4	10/10	0.92	0.23	66,68,69,71	0
3	GOL	D	314	6/6	0.93	0.15	60,65,65,68	0
2	CCU	B	2	10/10	0.93	0.19	39,43,45,48	0
3	GOL	D	3	6/6	0.93	0.15	71,72,73,73	0
3	GOL	A	2	6/6	0.94	0.14	53,56,57,58	0
4	SO4	B	13	5/5	0.94	0.37	73,75,77,77	0
2	CCU	C	1	10/10	0.96	0.19	42,44,48,52	0

6.5 Other polymers (i)

There are no such residues in this entry.