



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 10, 2020 – 05:05 AM BST

PDB ID : 4AMX
Title : CRYSTAL STRUCTURE OF THE GRACILARIOPSIS LEMANEIFORMIS
ALPHA-1,4- GLUCAN LYASE Covalent Intermediate Complex with 5-fluoro-
glucosyl- fluoride
Authors : Rozeboom, H.J.; Yu, S.; Madrid, S.; Kalk, K.H.; Dijkstra, B.W.
Deposited on : 2012-03-14
Resolution : 2.10 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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 : 2.13.1
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) : 2.13.1

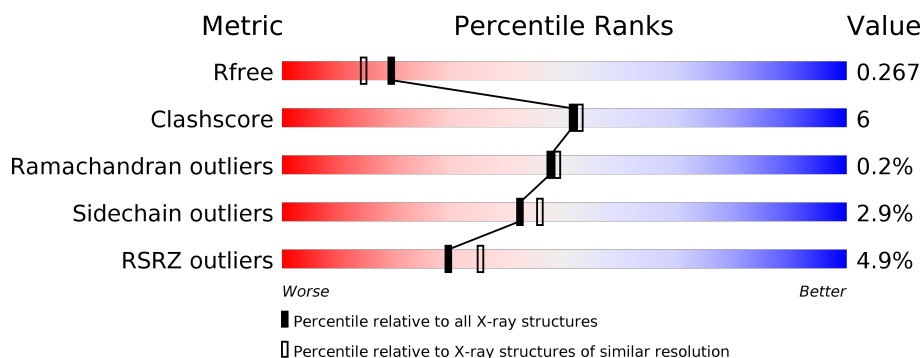
1 Overall quality at a glance

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}	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (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 on 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 $\leq 5\%$. 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	1027	<div> <div>3%</div> <div>86%</div> <div>13%</div> </div>
1	B	1027	<div> <div>6%</div> <div>85%</div> <div>14%</div> </div>
1	C	1027	<div> <div>7%</div> <div>87%</div> <div>13%</div> </div>
1	D	1027	<div> <div>4%</div> <div>85%</div> <div>14%</div> </div>

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	5GF	B	1039	X	-	-	-

2 Entry composition [i](#)

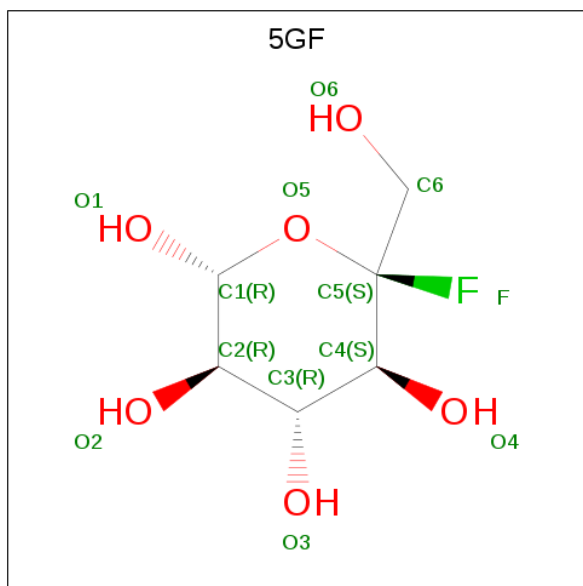
There are 5 unique types of molecules in this entry. The entry contains 33667 atoms, of which 0 are hydrogens and 0 are deuteriums.

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-1,4-GLUCAN LYASE ISOZYME 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1025	Total	C	N	O	S	0	1	0
			8171	5143	1385	1596	47			
1	B	1025	Total	C	N	O	S	0	0	0
			8162	5135	1385	1595	47			
1	C	1025	Total	C	N	O	S	0	0	0
			8162	5135	1385	1595	47			
1	D	1025	Total	C	N	O	S	0	1	0
			8171	5143	1385	1596	47			

- Molecule 2 is 5-fluoro-beta-D-glucopyranose (three-letter code: 5GF) (formula: C₆H₁₁FO₆).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	F	O	0	0
			12	6	1	5		
2	B	1	Total	C	F	O	0	0
			12	6	1	5		

Continued on next page...

Continued from previous page...

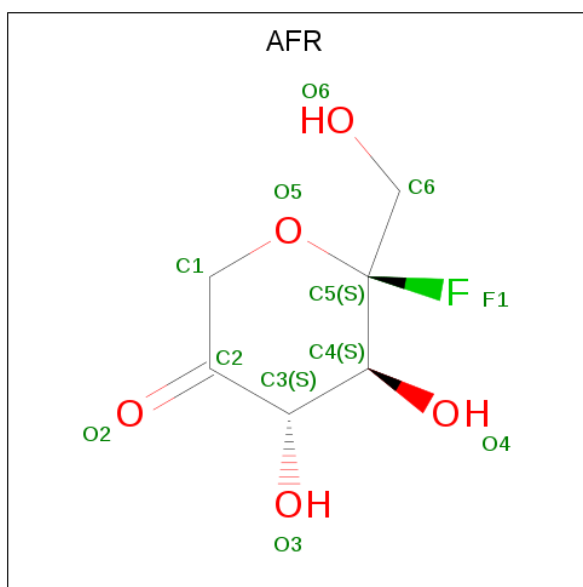
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	C	1	Total	C	F	O	0	0
			12	6	1	5		

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



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

- Molecule 4 is alpha-D-threo-hexo-2,5-diulo-5,1-pyranosyl fluoride (three-letter code: AFR) (formula: $C_6H_9FO_5$).



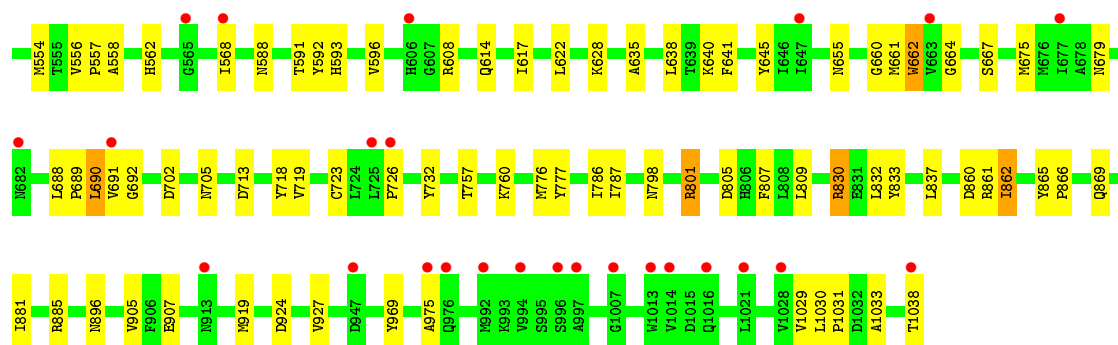
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	D	1	Total	C	F	O	0	0
			12	6	1	5		

- Molecule 5 is water.

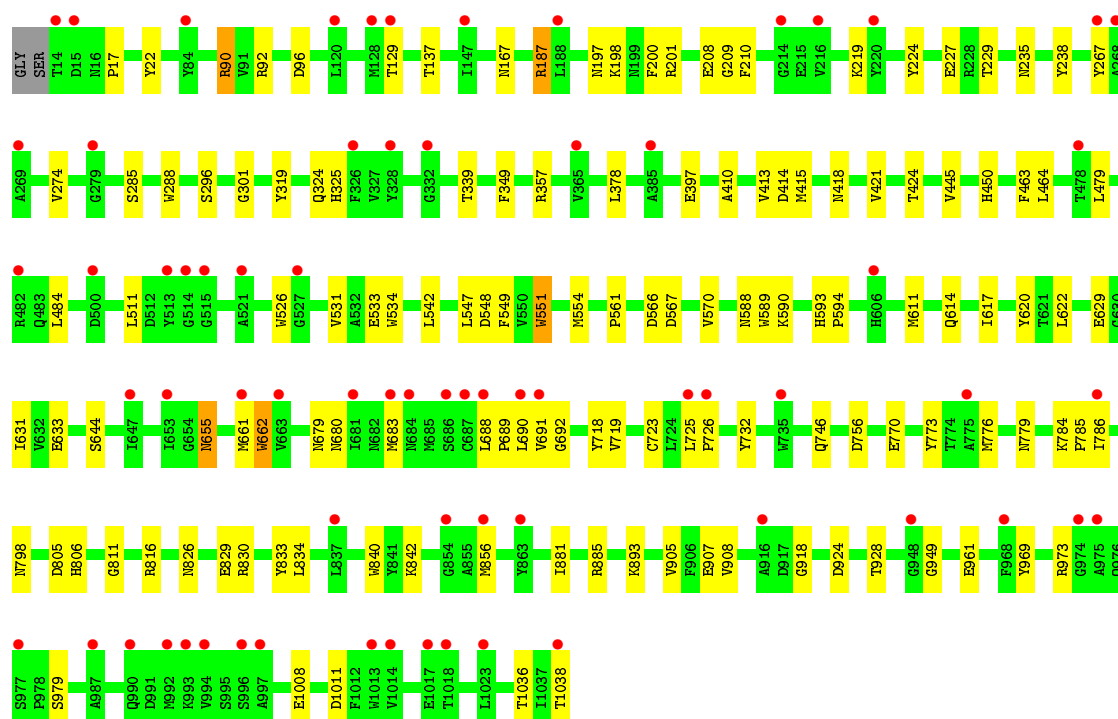
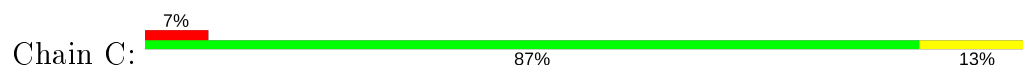
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	328	Total	O	0	0
			328	328		
5	B	147	Total	O	0	0
			147	147		
5	C	152	Total	O	0	0
			152	152		
5	D	296	Total	O	0	0
			296	296		

- Molecule 1: ALPHA-1,4-GLUCAN LYASE ISOZYME 1

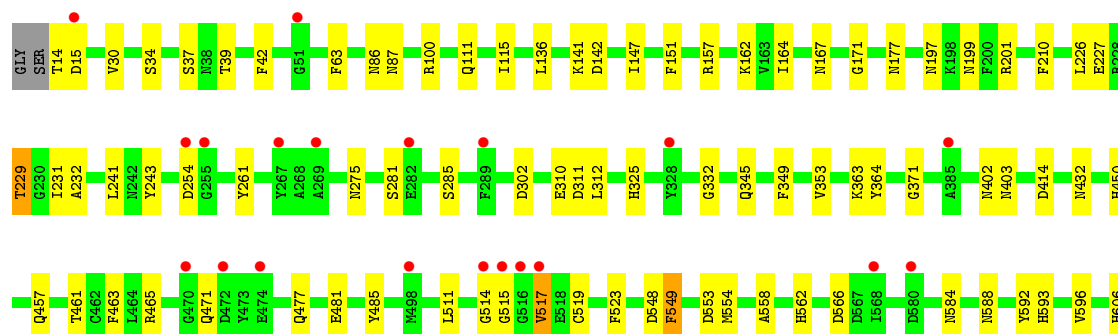
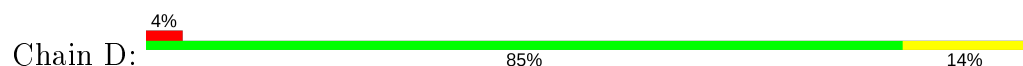


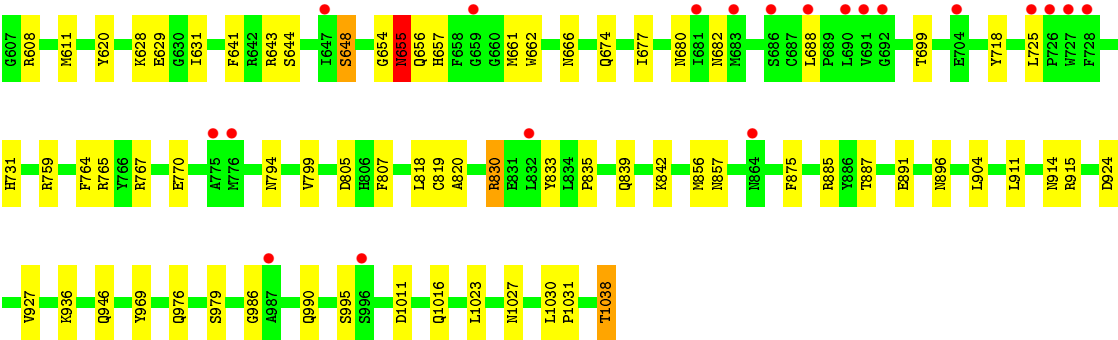


• Molecule 1: ALPHA-1,4-GLUCAN LYASE ISOZYME 1



• Molecule 1: ALPHA-1,4-GLUCAN LYASE ISOZYME 1





4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	91.64Å 97.01Å 135.71Å 80.38° 83.11° 85.22°	Depositor
Resolution (Å)	46.74 – 2.10 46.74 – 2.10	Depositor EDS
% Data completeness (in resolution range)	96.0 (46.74-2.10) 95.9 (46.74-2.10)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.77 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.220 , 0.267 0.222 , 0.267	Depositor DCC
R_{free} test set	12914 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	24.0	Xtriage
Anisotropy	0.533	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 36.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	33667	wwPDB-VP
Average B, all atoms (Å ²)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.73% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: GOL, CSO, 5GF, AFR

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	0.62	0/8387	0.70	2/11407 (0.0%)
1	B	0.49	0/8374	0.60	3/11389 (0.0%)
1	C	0.48	0/8374	0.59	0/11389
1	D	0.61	0/8387	0.67	1/11407 (0.0%)
All	All	0.56	0/33522	0.64	6/45592 (0.0%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	B	975	ALA	O-C-N	-7.75	110.31	122.70
1	B	622	LEU	CA-CB-CG	5.44	127.81	115.30
1	A	805	ASP	CB-CG-OD1	5.16	122.95	118.30
1	A	717	ARG	NE-CZ-NH1	5.12	122.86	120.30
1	D	643	ARG	NE-CZ-NH2	-5.10	117.75	120.30

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	8171	0	7615	92	0
1	B	8162	0	7605	86	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	8162	0	7606	85	0
1	D	8171	0	7615	94	0
2	A	12	0	9	1	0
2	B	12	0	9	0	0
2	C	12	0	9	0	0
3	A	12	0	16	1	0
3	B	6	0	8	0	0
3	D	12	0	16	2	0
4	D	12	0	9	3	0
5	A	328	0	0	20	0
5	B	147	0	0	13	0
5	C	152	0	0	8	0
5	D	296	0	0	18	0
All	All	33667	0	30517	357	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 357 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:C:210:PHE:H	1:C:229:THR:HG22	1.26	1.00
1:A:529:PRO:O	1:A:530:ASP:HB2	1.65	0.96
1:D:584:ASN:HB3	5:D:2168:HOH:O	1.68	0.94
1:A:642:ARG:HD2	5:A:2210:HOH:O	1.69	0.92
1:C:210:PHE:H	1:C:229:THR:CG2	1.85	0.88

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	1023/1027 (100%)	966 (94%)	53 (5%)	4 (0%)	34	32
1	B	1022/1027 (100%)	962 (94%)	59 (6%)	1 (0%)	51	54
1	C	1022/1027 (100%)	974 (95%)	47 (5%)	1 (0%)	51	54
1	D	1023/1027 (100%)	960 (94%)	59 (6%)	4 (0%)	34	32
All	All	4090/4108 (100%)	3862 (94%)	218 (5%)	10 (0%)	47	49

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	239	ASP
1	A	530	ASP
1	A	664	GLY
1	D	655	ASN
1	C	949	GLY

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	A	878/878 (100%)	855 (97%)	23 (3%)	46	50
1	B	877/878 (100%)	852 (97%)	25 (3%)	42	46
1	C	877/878 (100%)	852 (97%)	25 (3%)	42	46
1	D	878/878 (100%)	850 (97%)	28 (3%)	39	41
All	All	3510/3512 (100%)	3409 (97%)	101 (3%)	42	46

5 of 101 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	860	ASP
1	C	418	ASN
1	D	839	GLN
1	B	1038	THR
1	C	187	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 63 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	325	HIS
1	C	812	HIS
1	D	934	ASN
1	C	450	HIS
1	C	655	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

4 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	CSO	A	336	1	3,6,7	0.54	0	0,6,8	0.00	-
1	CSO	D	336	1	3,6,7	0.47	0	0,6,8	0.00	-
1	CSO	B	336	1	3,6,7	0.69	0	0,6,8	0.00	-
1	CSO	C	336	1	3,6,7	0.56	0	0,6,8	0.00	-

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	A	336	1	-	0/1/5/7	-
1	CSO	D	336	1	-	0/1/5/7	-
1	CSO	B	336	1	-	1/1/5/7	-
1	CSO	C	336	1	-	0/1/5/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	B	336	CSO	N-CA-CB-SG

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

9 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	GOL	D	1039	-	5,5,5	0.34	0	5,5,5	0.65	0
3	GOL	A	1041	-	5,5,5	0.30	0	5,5,5	0.44	0
4	AFR	D	1041	-	9,12,12	5.24	3 (33%)	10,18,18	0.95	0
3	GOL	B	1040	-	5,5,5	0.41	0	5,5,5	0.27	0
3	GOL	D	1040	-	5,5,5	0.40	0	5,5,5	0.63	0
2	5GF	B	1039	1	9,12,13	1.17	1 (11%)	14,18,20	0.86	0
2	5GF	A	1039	1	9,12,13	1.44	1 (11%)	14,18,20	1.07	1 (7%)
2	5GF	C	1039	1	9,12,13	1.28	1 (11%)	14,18,20	0.80	1 (7%)
3	GOL	A	1040	-	5,5,5	0.34	0	5,5,5	0.14	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
3	GOL	D	1039	-	-	4/4/4/4	-
3	GOL	A	1041	-	-	4/4/4/4	-
2	5GF	B	1039	1	1/1/4/5	2/2/23/26	0/1/1/1
3	GOL	B	1040	-	-	2/4/4/4	-
3	GOL	D	1040	-	-	0/4/4/4	-
4	AFR	D	1041	-	-	0/2/23/23	0/1/1/1
2	5GF	A	1039	1	-	0/2/23/26	0/1/1/1
2	5GF	C	1039	1	-	2/2/23/26	0/1/1/1
3	GOL	A	1040	-	-	2/4/4/4	-

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	1041	AFR	C1-C2	-15.10	1.35	1.51
2	A	1039	5GF	O5-C5	3.69	1.45	1.37
2	C	1039	5GF	O5-C5	3.55	1.44	1.37
2	B	1039	5GF	O5-C5	3.08	1.43	1.37
4	D	1041	AFR	O5-C1	2.99	1.46	1.43

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1039	5GF	O2-C2-C1	2.11	113.48	109.15
2	C	1039	5GF	C1-C2-C3	2.04	112.17	109.67

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	B	1039	5GF	C1

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	1039	GOL	C1-C2-C3-O3
3	A	1041	GOL	O1-C1-C2-C3
3	B	1040	GOL	C1-C2-C3-O3
2	B	1039	5GF	O5-C5-C6-O6
2	B	1039	5GF	C4-C5-C6-O6

There are no ring outliers.

4 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1041	GOL	1	0
4	D	1041	AFR	3	0
3	D	1040	GOL	2	0
2	A	1039	5GF	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	1024/1027 (99%)	0.26	26 (2%) 57 62	16, 30, 44, 68	0
1	B	1024/1027 (99%)	0.51	65 (6%) 20 24	24, 40, 58, 79	0
1	C	1024/1027 (99%)	0.56	69 (6%) 17 22	24, 40, 57, 78	0
1	D	1024/1027 (99%)	0.35	40 (3%) 39 45	17, 30, 48, 81	0
All	All	4096/4108 (99%)	0.42	200 (4%) 29 35	16, 35, 55, 81	0

The worst 5 of 200 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	987	ALA	6.2
1	C	837	LEU	4.9
1	B	514	GLY	4.6
1	A	988	GLY	4.5
1	C	514	GLY	4.4

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, 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
1	CSO	B	336	7/8	0.90	0.14	36,37,42,44	0
1	CSO	D	336	7/8	0.94	0.14	27,27,29,31	0
1	CSO	A	336	7/8	0.94	0.12	26,27,35,37	0
1	CSO	C	336	7/8	0.95	0.11	36,37,41,43	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, 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(\AA^2)	Q<0.9
3	GOL	A	1040	6/6	0.73	0.23	47,49,50,52	0
3	GOL	A	1041	6/6	0.83	0.19	45,46,46,47	0
3	GOL	D	1040	6/6	0.86	0.18	43,46,48,48	0
2	5GF	B	1039	12/13	0.88	0.15	42,44,46,47	0
4	AFR	D	1041	12/12	0.88	0.17	45,47,48,49	0
2	5GF	A	1039	12/13	0.89	0.16	35,36,38,39	0
3	GOL	B	1040	6/6	0.89	0.13	41,44,44,46	0
3	GOL	D	1039	6/6	0.90	0.17	34,35,37,37	0
2	5GF	C	1039	12/13	0.91	0.12	42,44,44,46	0

6.5 Other polymers [i](#)

There are no such residues in this entry.