WORLDWIDE PDB

PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Apr 21, 2024 – 05:23 am BST

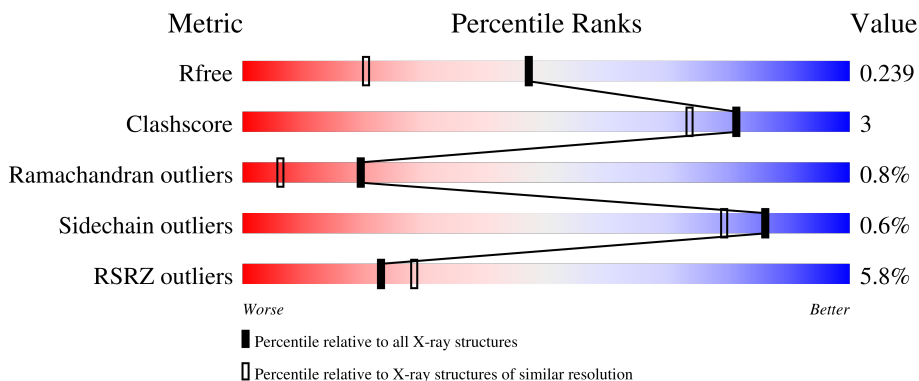|		
PDB ID	:	6ZN1
Title	:	Trehalose transferase bound to alpha-D-glucopyranosyl-beta-galactopyranose from Thermoproteus uzoniensis
Authors	:	Bento, I.; Mestrom, L.; Marsden, S.R.; van der Eijk, H.; Laustsen, J.U.; Jeffries, C.M.; Svergun, D.I.; Hagedoorn, P.-H.; Hanefeld, U.
Deposited on	:	2020-07-06
Resolution	:	1.75 Å(reported)

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 1.75 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	Percentile Ranks	Value
Rfree		0.239
Clashscore		3
Ramachandran outliers		0.8%
Sidechain outliers		0.6%
RSRZ outliers		5.8%

Worse   *Better*

▮ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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.

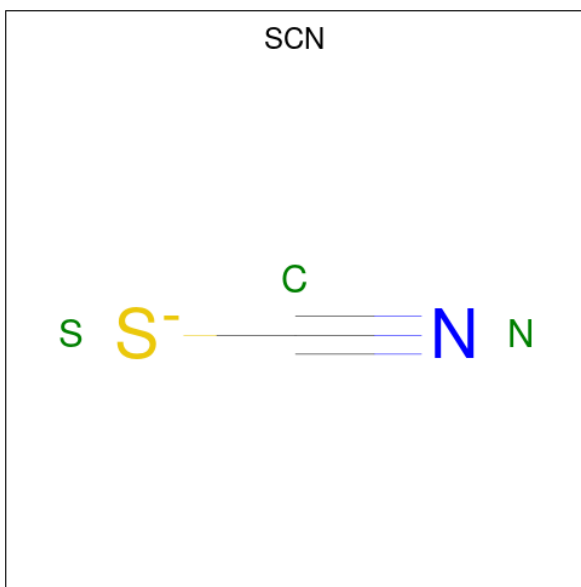| Mol	Chain	Length	Quality of chain
1	AAA	400	6% / 92% / 7%

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 6742 atoms, of which 3321 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 Trehalose phosphorylase/synthase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	400	Total 6473	C 2066	H 3258	N 562	O 578	S 9	72	1	0

- Molecule 2 is THIOCYANATE ION (three-letter code: SCN) (formula: CNS).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	AAA	1	Total 3	C 1	N 1	S 1	0	0
2	AAA	1	Total 3	C 1	N 1	S 1	0	0
2	AAA	1	Total 3	C 1	N 1	S 1	0	0
2	AAA	1	Total 3	C 1	N 1	S 1	0	0

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).

PEG

C4

OH O4

O2

O

C3

C1

C2

HO

O1

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	AAA	1	Total C H O				2	1
			34 8 20 6					
3	AAA	1	Total C H O				1	0
			17 4 10 3					
3	AAA	1	Total C H O				1	0
			17 4 10 3					

• Molecule 4 is alpha-D-glucopyranose (three-letter code: GLC) (formula: $C_6H_{12}O_6$) (labeled as "Ligand of Interest" by depositor).

GLC

C6

O5

O1

HO

OH

O6

C5(R)

C1(S)

C4(S)

C2(R)

C3(S)

HO

OH

O4

O2

OH

O3
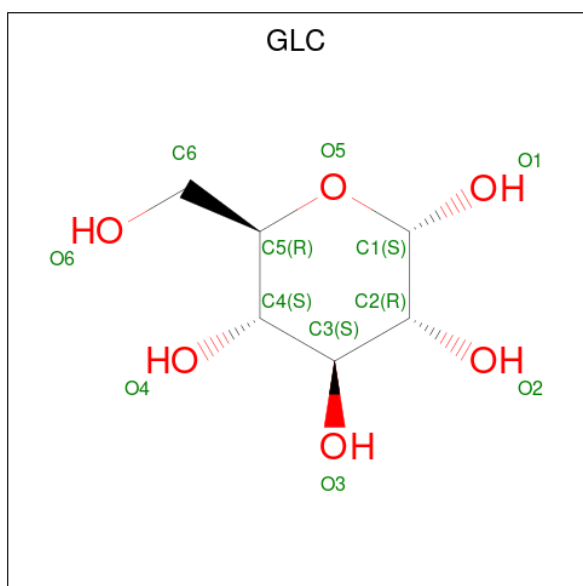
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	AAA	1	Total 22	C 6	H 11	O 5	4	0

- Molecule 5 is beta-L-galactopyranose (three-letter code: GIV) (formula: $C_6H_{12}O_6$) (labeled as "Ligand of Interest" by depositor).
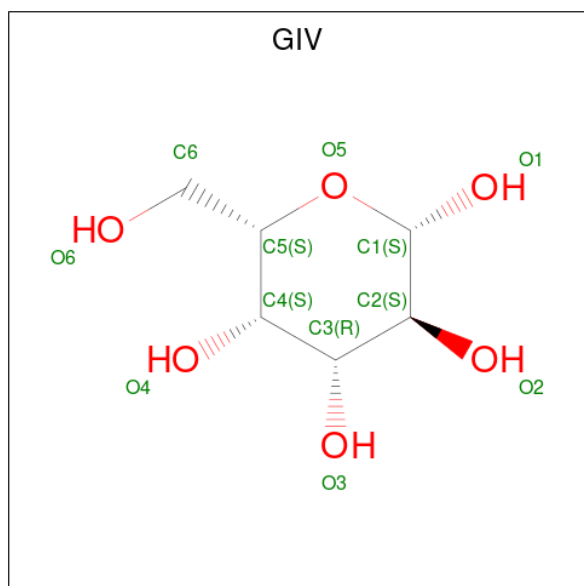
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	AAA	1	Total 24	C 6	H 12	O 6	4	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	AAA	143	Total 143	O 143	0	1

3 Residue-property plots ⓘ

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: Trehalose phosphorylase/synthase
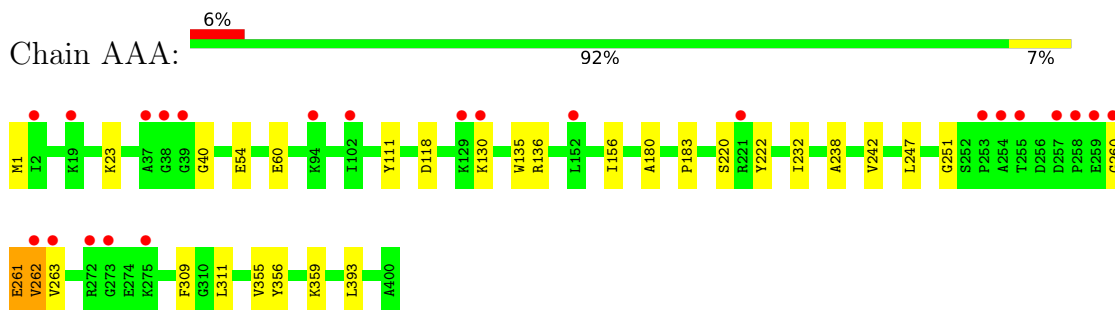
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	61.01Å 68.77Å 110.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	58.51 – 1.75 58.44 – 1.75	Depositor EDS
% Data completeness (in resolution range)	100.0 (58.51-1.75) 100.0 (58.44-1.75)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.66 (at 1.75Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.202 , 0.234 0.212 , 0.239	Depositor DCC
R_{free} test set	2329 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å^2)	34.9	Xtriage
Anisotropy	0.079	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.42 , 41.0	EDS
L-test for twinning[2]	$< |L| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6742	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.44% 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 ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GIV, GLC, SCN, PEG

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	0.63	0/3296	0.72	0/4468

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 ⓘ

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	3215	3258	3247	14	1
2	AAA	12	0	0	0	0
3	AAA	28	40	40	1	1
4	AAA	11	11	10	3	0
5	AAA	12	12	11	3	0
6	AAA	143	0	0	0	0
All	All	3421	3321	3308	17	1

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 (17) 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)
4:AAA:708:GLC:C1	5:AAA:709:GIV:O1	1.64	1.46
4:AAA:708:GLC:C1	5:AAA:709:GIV:C1	2.57	0.82
4:AAA:708:GLC:C2	5:AAA:709:GIV:O1	2.47	0.60
1:AAA:232:ILE:HD13	1:AAA:247:LEU:CD2	2.35	0.56
1:AAA:261:GLU:O	1:AAA:262:VAL:HG12	2.04	0.56
1:AAA:222:TYR:O	1:AAA:263:VAL:HG21	2.05	0.56
1:AAA:1:MET:HB3	1:AAA:54:GLU:OE1	2.12	0.49
1:AAA:260:GLY:O	1:AAA:261:GLU:HB2	2.12	0.49
1:AAA:183:PRO:HG2	1:AAA:309:PHE:CD2	2.50	0.46
1:AAA:220:SER:O	1:AAA:251:GLY:HA2	2.17	0.45
1:AAA:356:TYR:HA	3:AAA:704:PEG:H11	1.98	0.44
1:AAA:135:TRP:CD2	1:AAA:156:ILE:HD12	2.55	0.42
1:AAA:60:GLU:OE1	1:AAA:111:TYR:OH	2.28	0.42
1:AAA:238:ALA:O	1:AAA:242:VAL:HG22	2.20	0.41
1:AAA:23:LYS:HD2	1:AAA:393:LEU:HD21	2.03	0.41
1:AAA:40:GLY:HA2	1:AAA:311:LEU:HD12	2.02	0.41
1:AAA:355:VAL:O	1:AAA:359:LYS:HG2	2.21	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:AAA:180:ALA:H	3:AAA:704:PEG:H21[3_545]	1.16	0.44

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	399/400 (100%)	382 (96%)	14 (4%)	3 (1%)	19	6

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	262	VAL
1	AAA	261	GLU
1	AAA	118	ASP

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	345/344 (100%)	343 (99%)	2 (1%)	86	79

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	130	LYS
1	AAA	136	ARG

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 ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

10 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	PEG	AAA	704	-	6,6,6	0.20	0	5,5,5	0.34	0
2	SCN	AAA	707	-	1,2,2	0.70	0	0,1,1	-	-
2	SCN	AAA	705	-	1,2,2	0.54	0	0,1,1	-	-
5	GIV	AAA	709	-	12,12,12	0.38	0	17,17,17	0.75	0
2	SCN	AAA	701	-	1,2,2	0.88	0	0,1,1	-	-
3	PEG	AAA	702[B]	-	6,6,6	0.13	0	5,5,5	0.08	0
3	PEG	AAA	703	-	6,6,6	0.10	0	5,5,5	0.09	0
4	GLC	AAA	708	-	11,11,12	0.47	0	15,15,17	1.67	2 (13%)
2	SCN	AAA	706	-	1,2,2	0.86	0	0,1,1	-	-
3	PEG	AAA	702[A]	-	6,6,6	0.12	0	5,5,5	0.07	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	PEG	AAA	704	-	-	3/4/4/4	-
5	GIV	AAA	709	-	-	2/2/22/22	0/1/1/1
3	PEG	AAA	702[B]	-	-	0/4/4/4	-
3	PEG	AAA	703	-	-	1/4/4/4	-
4	GLC	AAA	708	-	-	0/2/19/22	0/1/1/1
3	PEG	AAA	702[A]	-	-	1/4/4/4	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	AAA	708	GLC	C1-O5-C5	5.18	119.22	112.19
4	AAA	708	GLC	C1-C2-C3	2.57	112.83	109.67

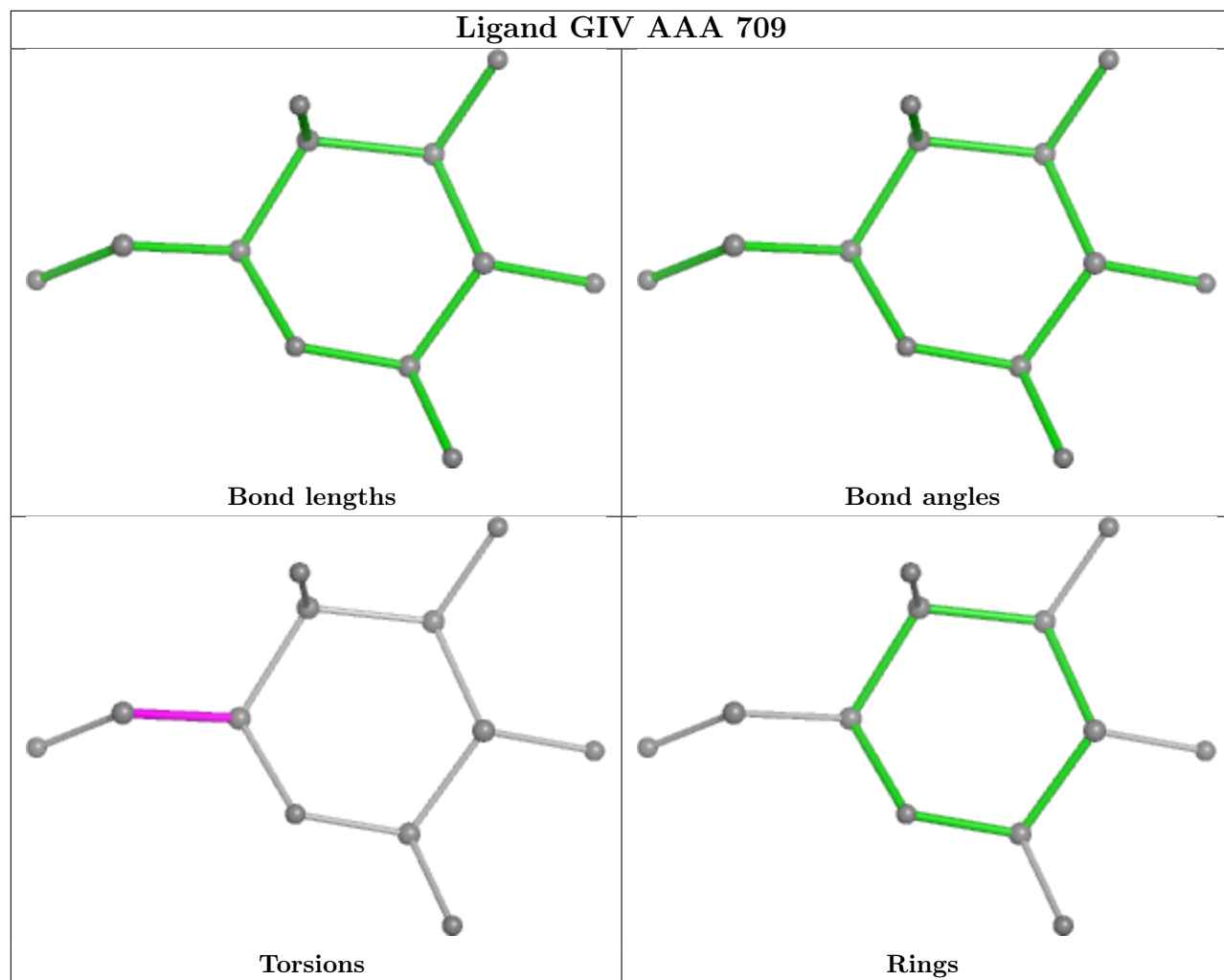There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	AAA	709	GIV	C4-C5-C6-O6
5	AAA	709	GIV	O5-C5-C6-O6
3	AAA	702[A]	PEG	O1-C1-C2-O2
3	AAA	704	PEG	O2-C3-C4-O4
3	AAA	703	PEG	O2-C3-C4-O4
3	AAA	704	PEG	C4-C3-O2-C2
3	AAA	704	PEG	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	AAA	704	PEG	1	1
5	AAA	709	GIV	3	0
4	AAA	708	GLC	3	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand GIV AAA 709

Bond lengths

Bond angles

Torsions

Rings

Ligand GLC AAA 708

Bond lengths

Bond angles

Torsions

Rings
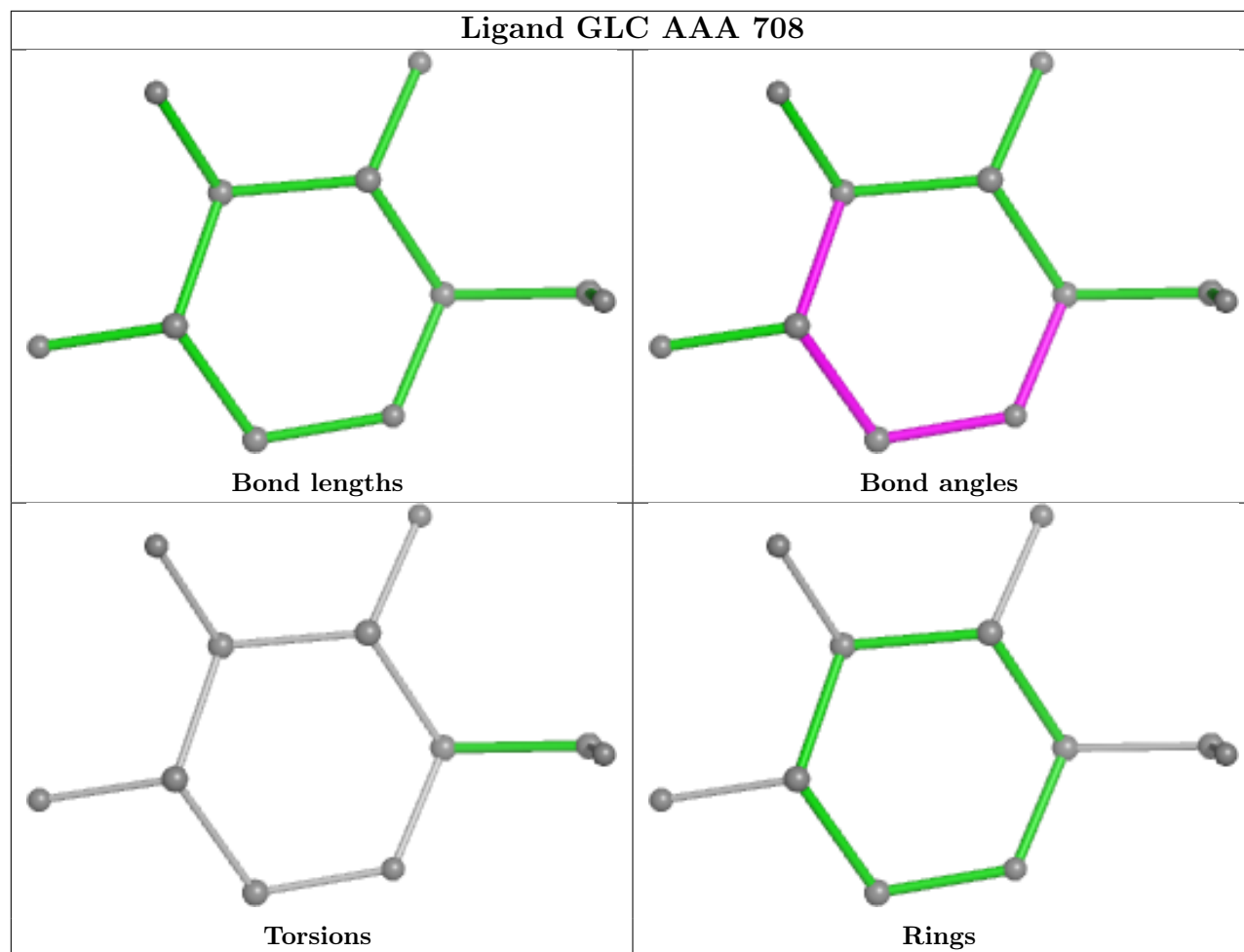
5.7 Other polymers ⓘ

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, 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	AAA	400/400 (100%)	0.45	23 (5%)	23	28	24, 39, 69, 103	0

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	130	LYS	7.8
1	AAA	254	ALA	4.7
1	AAA	39	GLY	4.7
1	AAA	38	GLY	4.4
1	AAA	273	GLY	4.2
1	AAA	37	ALA	3.9
1	AAA	253	PRO	3.9
1	AAA	272	ARG	3.7
1	AAA	255	THR	3.5
1	AAA	258	PRO	3.2
1	AAA	260	GLY	3.1
1	AAA	262	VAL	3.0
1	AAA	129	LYS	2.8
1	AAA	152	LEU	2.7
1	AAA	259	GLU	2.6
1	AAA	94	LYS	2.5
1	AAA	102	ILE	2.5
1	AAA	263	VAL	2.4
1	AAA	2	ILE	2.4
1	AAA	19	LYS	2.3
1	AAA	257	ASP	2.1
1	AAA	275	LYS	2.1
1	AAA	221	ARG	2.1

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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($\mathring{A}^2$)	Q<0.9
3	PEG	AAA	702[A]	7/7	0.75	0.17	47,49,52,53	17
3	PEG	AAA	702[B]	7/7	0.75	0.17	49,62,64,64	17
3	PEG	AAA	704	7/7	0.78	0.26	40,44,47,47	17
5	GIV	AAA	709	12/12	0.84	0.16	28,43,50,54	4
2	SCN	AAA	701	3/3	0.85	0.27	50,50,52,61	3
3	PEG	AAA	703	7/7	0.87	0.15	46,48,49,49	17
2	SCN	AAA	707	3/3	0.88	0.14	55,55,56,64	3
2	SCN	AAA	706	3/3	0.92	0.11	51,51,53,55	0
2	SCN	AAA	705	3/3	0.95	0.16	57,57,61,62	3
4	GLC	AAA	708	11/12	0.97	0.09	24,27,29,30	4

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.
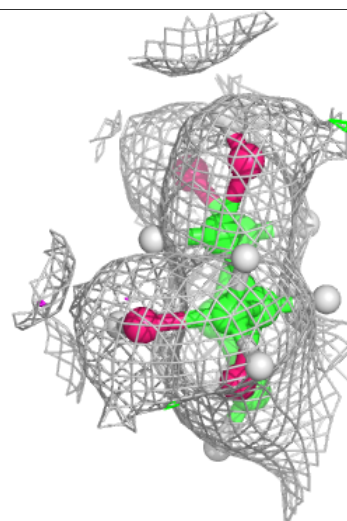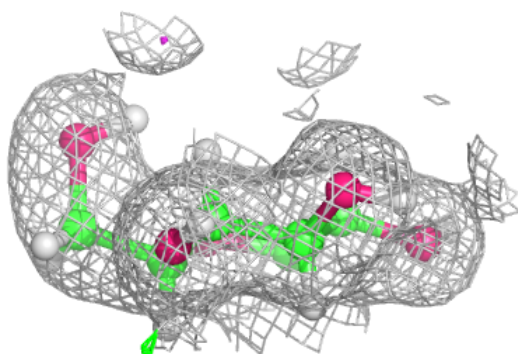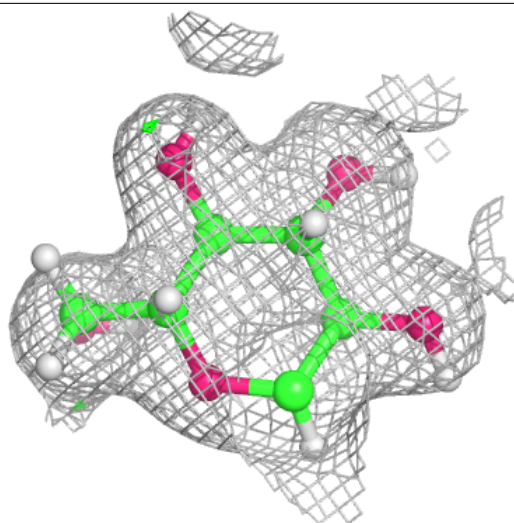
Electron density around GIV AAA 709:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
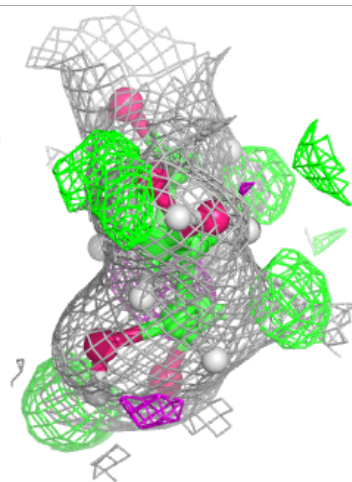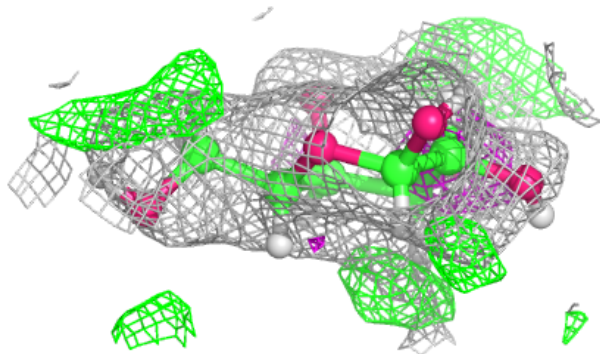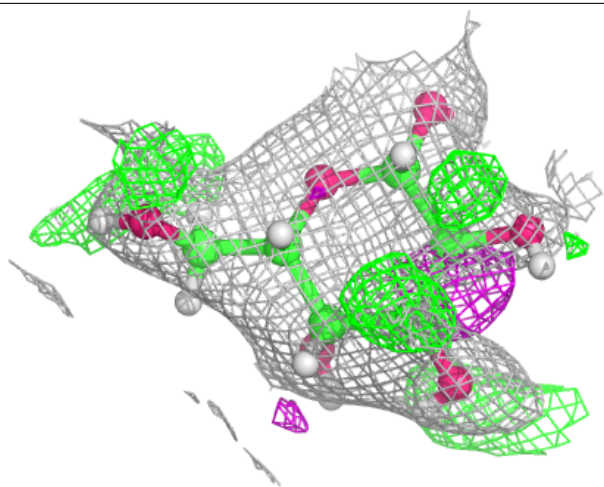
Electron density around GLC AAA 708:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

6.5 Other polymers (i)

There are no such residues in this entry.