



Full wwPDB X-ray Structure Validation Report ⓘ

Oct 22, 2023 – 12:56 PM JST

PDB ID : 8J51
Title : Crystal structure of Flavihumibacter petaseus GH31 alpha-galactosidase in complex with galactose
Authors : Ikegaya, M.; Miyazaki, T.
Deposited on : 2023-04-21
Resolution : 2.15 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

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 : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36
buster-report : 1.1.7 (2018)
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

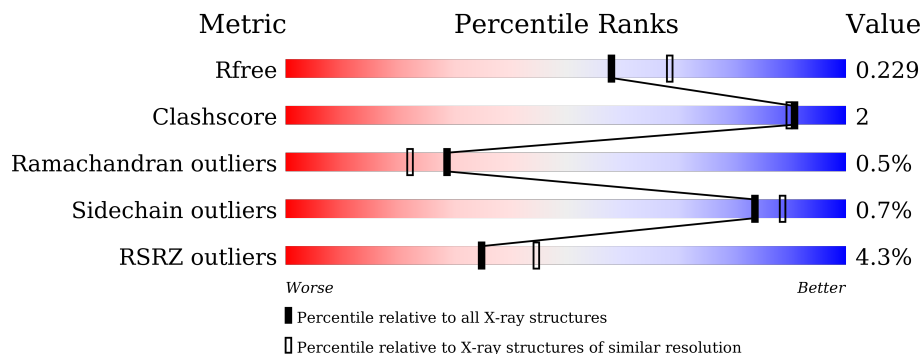
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.15 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	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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 $\leq 5\%$. 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	541	 3% 88% 6% 6%
1	B	541	 5% 88% 7% 6%

2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 16696 atoms, of which 8089 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 GH31 alpha-galactosidase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	510	8055	2608	3989	701	734	23	91	2	0
1	B	510	8037	2602	3980	701	731	23	90	0	0

There are 46 discrepancies between the modelled and reference sequences:

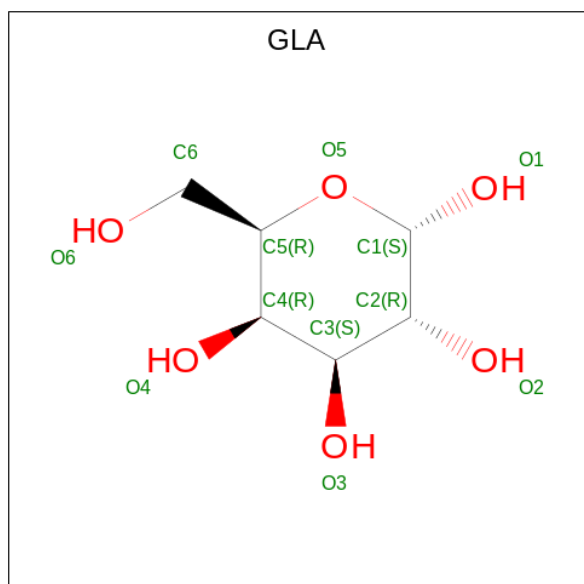
Chain	Residue	Modelled	Actual	Comment	Reference
A	6	MET	-	initiating methionine	UNP A0A0E9MUN5
A	7	GLY	-	expression tag	UNP A0A0E9MUN5
A	8	SER	-	expression tag	UNP A0A0E9MUN5
A	9	SER	-	expression tag	UNP A0A0E9MUN5
A	10	HIS	-	expression tag	UNP A0A0E9MUN5
A	11	HIS	-	expression tag	UNP A0A0E9MUN5
A	12	HIS	-	expression tag	UNP A0A0E9MUN5
A	13	HIS	-	expression tag	UNP A0A0E9MUN5
A	14	HIS	-	expression tag	UNP A0A0E9MUN5
A	15	HIS	-	expression tag	UNP A0A0E9MUN5
A	16	SER	-	expression tag	UNP A0A0E9MUN5
A	17	SER	-	expression tag	UNP A0A0E9MUN5
A	18	GLY	-	expression tag	UNP A0A0E9MUN5
A	19	LEU	-	expression tag	UNP A0A0E9MUN5
A	20	VAL	-	expression tag	UNP A0A0E9MUN5
A	21	PRO	-	expression tag	UNP A0A0E9MUN5
A	22	ARG	-	expression tag	UNP A0A0E9MUN5
A	23	GLY	-	expression tag	UNP A0A0E9MUN5
A	24	SER	-	expression tag	UNP A0A0E9MUN5
A	25	HIS	-	expression tag	UNP A0A0E9MUN5
A	26	MET	-	expression tag	UNP A0A0E9MUN5
A	27	ALA	-	expression tag	UNP A0A0E9MUN5
A	28	SER	-	expression tag	UNP A0A0E9MUN5
B	6	MET	-	initiating methionine	UNP A0A0E9MUN5
B	7	GLY	-	expression tag	UNP A0A0E9MUN5

Continued on next page...

Continued from previous page...

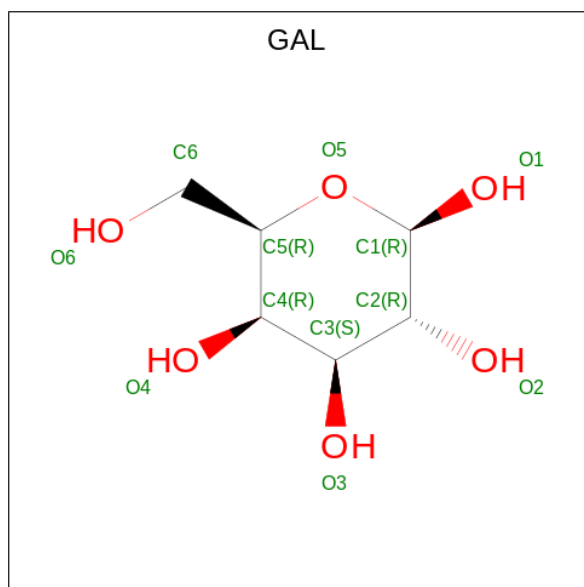
Chain	Residue	Modelled	Actual	Comment	Reference
B	8	SER	-	expression tag	UNP A0A0E9MUN5
B	9	SER	-	expression tag	UNP A0A0E9MUN5
B	10	HIS	-	expression tag	UNP A0A0E9MUN5
B	11	HIS	-	expression tag	UNP A0A0E9MUN5
B	12	HIS	-	expression tag	UNP A0A0E9MUN5
B	13	HIS	-	expression tag	UNP A0A0E9MUN5
B	14	HIS	-	expression tag	UNP A0A0E9MUN5
B	15	HIS	-	expression tag	UNP A0A0E9MUN5
B	16	SER	-	expression tag	UNP A0A0E9MUN5
B	17	SER	-	expression tag	UNP A0A0E9MUN5
B	18	GLY	-	expression tag	UNP A0A0E9MUN5
B	19	LEU	-	expression tag	UNP A0A0E9MUN5
B	20	VAL	-	expression tag	UNP A0A0E9MUN5
B	21	PRO	-	expression tag	UNP A0A0E9MUN5
B	22	ARG	-	expression tag	UNP A0A0E9MUN5
B	23	GLY	-	expression tag	UNP A0A0E9MUN5
B	24	SER	-	expression tag	UNP A0A0E9MUN5
B	25	HIS	-	expression tag	UNP A0A0E9MUN5
B	26	MET	-	expression tag	UNP A0A0E9MUN5
B	27	ALA	-	expression tag	UNP A0A0E9MUN5
B	28	SER	-	expression tag	UNP A0A0E9MUN5

- Molecule 2 is alpha-D-galactopyranose (three-letter code: GLA) (formula: C₆H₁₂O₆) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	
2	A	1	Total	C	H	O	4	1
			24	6	12	6		
2	B	1	Total	C	H	O	4	1
			24	6	12	6		

- Molecule 3 is beta-D-galactopyranose (three-letter code: GAL) (formula: C₆H₁₂O₆) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	
3	A	1	Total	C	H	O	4	1
			24	6	12	6		
3	B	1	Total	C	H	O	4	1
			24	6	12	6		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



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

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Cl 1 1	0	0
5	B	1	Total Cl 1 1	0	0

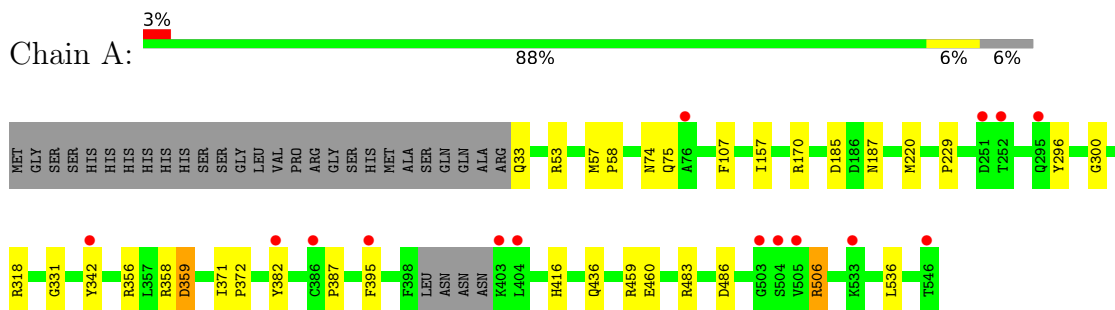
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	204	Total O 204 204	0	0
6	B	182	Total O 182 182	0	0

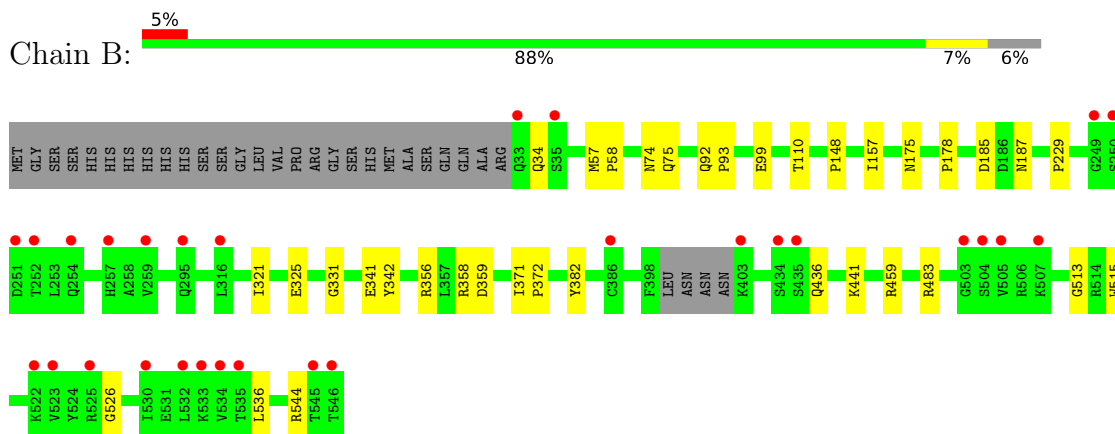
3 Residue-property plots [i](#)

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: GH31 alpha-galactosidase



- Molecule 1: GH31 alpha-galactosidase



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	71.36Å 72.37Å 80.88Å 101.82° 105.11° 103.67°	Depositor
Resolution (Å)	43.70 – 2.15 43.66 – 2.15	Depositor EDS
% Data completeness (in resolution range)	98.0 (43.70-2.15) 98.0 (43.66-2.15)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.49 (at 2.16Å)	Xtrriage
Refinement program	REFMAC 5.8.0405	Depositor
R, R_{free}	0.152 , 0.222 0.166 , 0.229	Depositor DCC
R_{free} test set	3948 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å ²)	25.1	Xtrriage
Anisotropy	0.479	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.46 , 48.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.004 for -k,-h,-l	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	16696	wwPDB-VP
Average B, all atoms (Å ²)	36.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.73% 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: GLA, EDO, GAL, CL

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	0.46	1/4183 (0.0%)	0.77	4/5669 (0.1%)
1	B	0.46	1/4168 (0.0%)	0.76	1/5648 (0.0%)
All	All	0.46	2/8351 (0.0%)	0.77	5/11317 (0.0%)

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	A	0	3
1	B	0	2
All	All	0	5

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	341	GLU	CD-OE1	5.28	1.31	1.25
1	A	460	GLU	CD-OE2	-5.05	1.20	1.25

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	356	ARG	NE-CZ-NH2	-6.65	116.98	120.30
1	B	356	ARG	NE-CZ-NH1	5.50	123.05	120.30
1	A	318	ARG	NE-CZ-NH2	-5.42	117.59	120.30
1	A	318	ARG	NE-CZ-NH1	5.37	122.98	120.30
1	A	296	TYR	CB-CG-CD1	5.36	124.21	121.00

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	459	ARG	Sidechain
1	A	483	ARG	Sidechain
1	A	506	ARG	Sidechain
1	B	459	ARG	Sidechain
1	B	483	ARG	Sidechain

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	4066	3989	3980	14	0
1	B	4057	3980	3969	15	0
2	A	12	12	12	0	0
2	B	12	12	12	0	0
3	A	12	12	12	0	0
3	B	12	12	12	0	0
4	A	20	30	30	0	0
4	B	28	42	42	0	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
6	A	204	0	0	2	0
6	B	182	0	0	0	0
All	All	8607	8089	8069	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 2.

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:A:53:ARG:HD3	6:A:722:HOH:O	1.80	0.80
1:B:57:MET:HA	1:B:58:PRO:C	2.26	0.56
1:A:57:MET:HA	1:A:58:PRO:C	2.29	0.53
1:A:74:ASN:HD21	1:A:358:ARG:H	1.59	0.51
1:B:436:GLN:HG3	1:B:536:LEU:HD21	1.92	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:387:PRO:HG2	1:A:416:HIS:CE1	2.47	0.50
1:A:436:GLN:HG3	1:A:536:LEU:HD21	1.94	0.49
1:A:331:GLY:HA2	1:A:342:TYR:CE2	2.48	0.49
1:A:229:PRO:HD2	1:B:229:PRO:HD2	1.95	0.48
1:A:486:ASP:O	1:A:506:ARG:NH2	2.44	0.48
1:B:321:ILE:HD12	1:B:325:GLU:HB3	1.94	0.48
1:B:331:GLY:HA2	1:B:342:TYR:CE2	2.49	0.47
1:B:75:GLN:HB3	1:B:382:TYR:CE2	2.50	0.46
1:B:74:ASN:ND2	1:B:358:ARG:H	2.14	0.46
1:B:175:ASN:O	1:B:441:LYS:HE2	2.16	0.46
1:B:92:GLN:HB3	1:B:93:PRO:CD	2.46	0.46
1:B:74:ASN:HD21	1:B:358:ARG:H	1.65	0.45
1:B:148:PRO:HD2	1:B:178:PRO:HG2	1.99	0.44
1:A:170:ARG:NH1	6:A:702:HOH:O	2.36	0.43
1:B:34:GLN:NE2	1:B:110:THR:HG22	2.33	0.43
1:B:515:TRP:CH2	1:B:544:ARG:HG3	2.54	0.43
1:B:513:GLY:O	1:B:526:GLY:HA3	2.19	0.43
1:A:74:ASN:ND2	1:A:358:ARG:H	2.17	0.43
1:A:75:GLN:HB3	1:A:382:TYR:CE2	2.54	0.43
1:A:220:MET:HA	1:A:300:GLY:O	2.20	0.42
1:B:371:ILE:N	1:B:372:PRO:HD2	2.35	0.41
1:A:33:GLN:HG3	1:A:107:PHE:O	2.22	0.41
1:A:371:ILE:N	1:A:372:PRO:HD2	2.36	0.41

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	508/541 (94%)	493 (97%)	12 (2%)	3 (1%)	25 18
1	B	506/541 (94%)	493 (97%)	11 (2%)	2 (0%)	34 29

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1014/1082 (94%)	986 (97%)	23 (2%)	5 (0%)	29 22

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	187	ASN
1	B	187	ASN
1	A	157	ILE
1	B	157	ILE
1	A	359	ASP

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	427/451 (95%)	424 (99%)	3 (1%)	84 89
1	B	425/451 (94%)	422 (99%)	3 (1%)	84 89
All	All	852/902 (94%)	846 (99%)	6 (1%)	84 89

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	185	ASP
1	A	359	ASP
1	A	395	PHE
1	B	99	GLU
1	B	185	ASP
1	B	359	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	74	ASN
1	B	34	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	74	ASN

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](#)

Of 18 ligands modelled in this entry, 2 are monoatomic - leaving 16 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	EDO	B	606	-	3,3,3	0.32	0	2,2,2	0.45	0
4	EDO	A	606	-	3,3,3	0.29	0	2,2,2	0.34	0
4	EDO	A	607	-	3,3,3	0.50	0	2,2,2	0.30	0
4	EDO	A	604	-	3,3,3	0.20	0	2,2,2	0.39	0
4	EDO	B	607	-	3,3,3	0.11	0	2,2,2	0.29	0
4	EDO	A	605	-	3,3,3	0.34	0	2,2,2	0.23	0
4	EDO	B	605	-	3,3,3	0.15	0	2,2,2	0.39	0
3	GAL	B	602[B]	-	12,12,12	0.57	0	17,17,17	0.71	0
4	EDO	B	608	-	3,3,3	0.23	0	2,2,2	0.35	0
4	EDO	B	603	-	3,3,3	0.14	0	2,2,2	0.08	0
4	EDO	B	609	-	3,3,3	0.25	0	2,2,2	0.58	0
2	GLA	B	601[A]	-	12,12,12	0.65	0	17,17,17	0.69	0
2	GLA	A	601[A]	-	12,12,12	0.46	0	17,17,17	0.53	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	EDO	B	604	-	3,3,3	0.28	0	2,2,2	0.33	0
3	GAL	A	602[B]	-	12,12,12	0.38	0	17,17,17	0.67	0
4	EDO	A	603	-	3,3,3	0.34	0	2,2,2	0.54	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	EDO	B	606	-	-	0/1/1/1	-
4	EDO	A	606	-	-	0/1/1/1	-
4	EDO	A	607	-	-	1/1/1/1	-
4	EDO	A	604	-	-	0/1/1/1	-
4	EDO	B	607	-	-	1/1/1/1	-
4	EDO	A	605	-	-	0/1/1/1	-
4	EDO	B	605	-	-	1/1/1/1	-
3	GAL	B	602[B]	-	-	0/2/22/22	0/1/1/1
4	EDO	B	608	-	-	0/1/1/1	-
4	EDO	B	603	-	-	1/1/1/1	-
4	EDO	B	609	-	-	1/1/1/1	-
2	GLA	B	601[A]	-	-	0/2/22/22	0/1/1/1
2	GLA	A	601[A]	-	-	0/2/22/22	0/1/1/1
4	EDO	B	604	-	-	1/1/1/1	-
3	GAL	A	602[B]	-	-	0/2/22/22	0/1/1/1
4	EDO	A	603	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

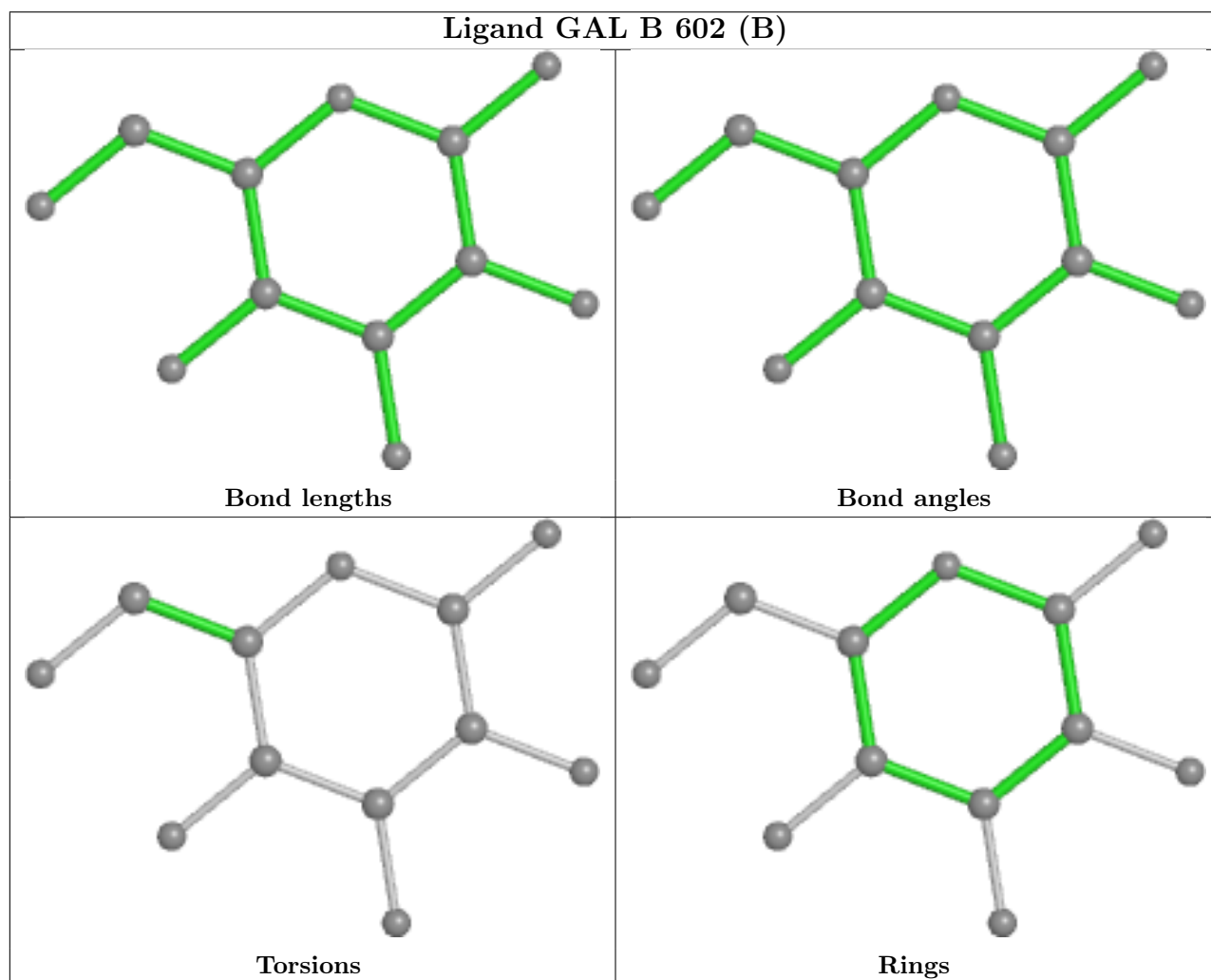
All (7) torsion outliers are listed below:

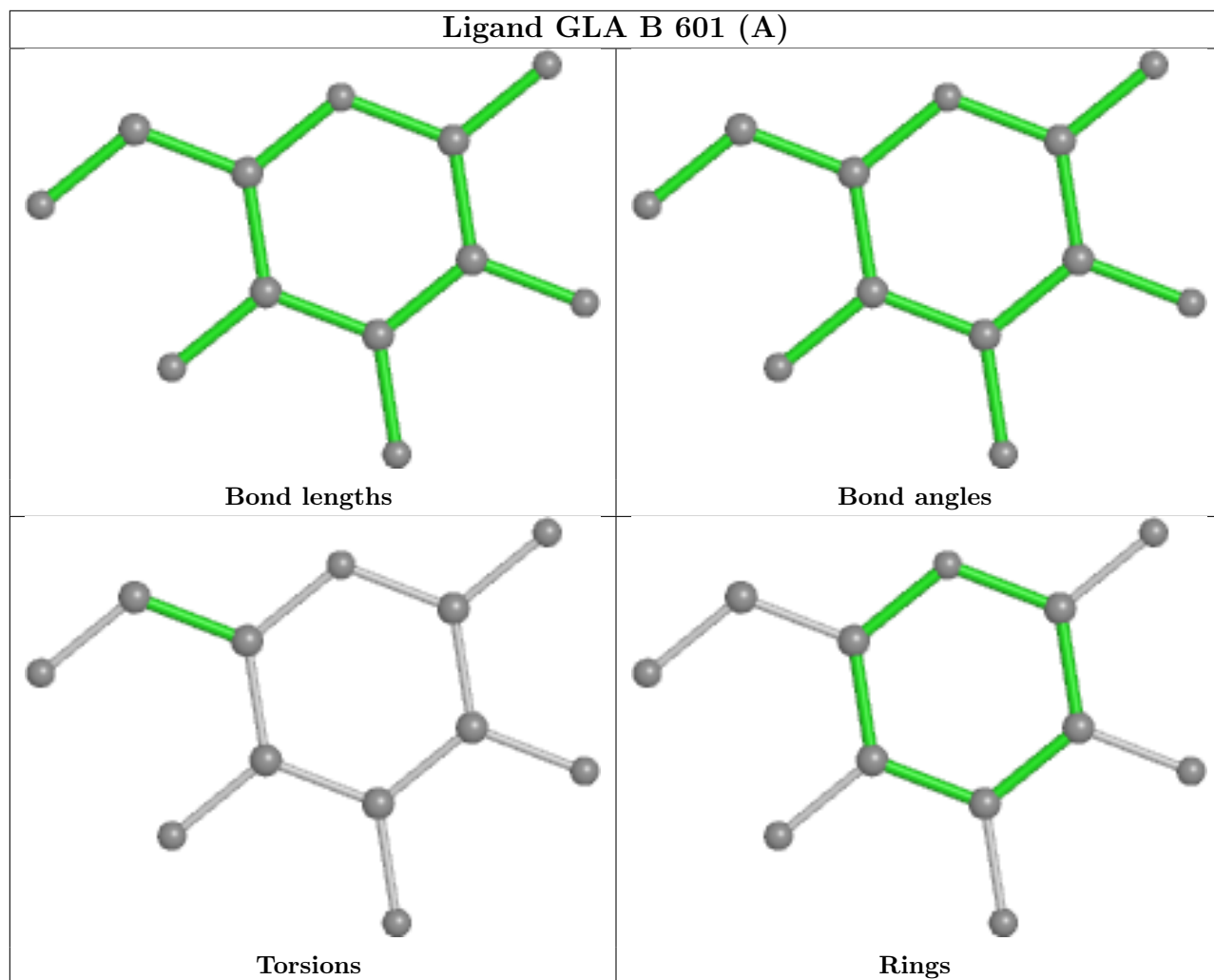
Mol	Chain	Res	Type	Atoms
4	A	603	EDO	O1-C1-C2-O2
4	A	607	EDO	O1-C1-C2-O2
4	B	605	EDO	O1-C1-C2-O2
4	B	607	EDO	O1-C1-C2-O2
4	B	609	EDO	O1-C1-C2-O2
4	B	603	EDO	O1-C1-C2-O2
4	B	604	EDO	O1-C1-C2-O2

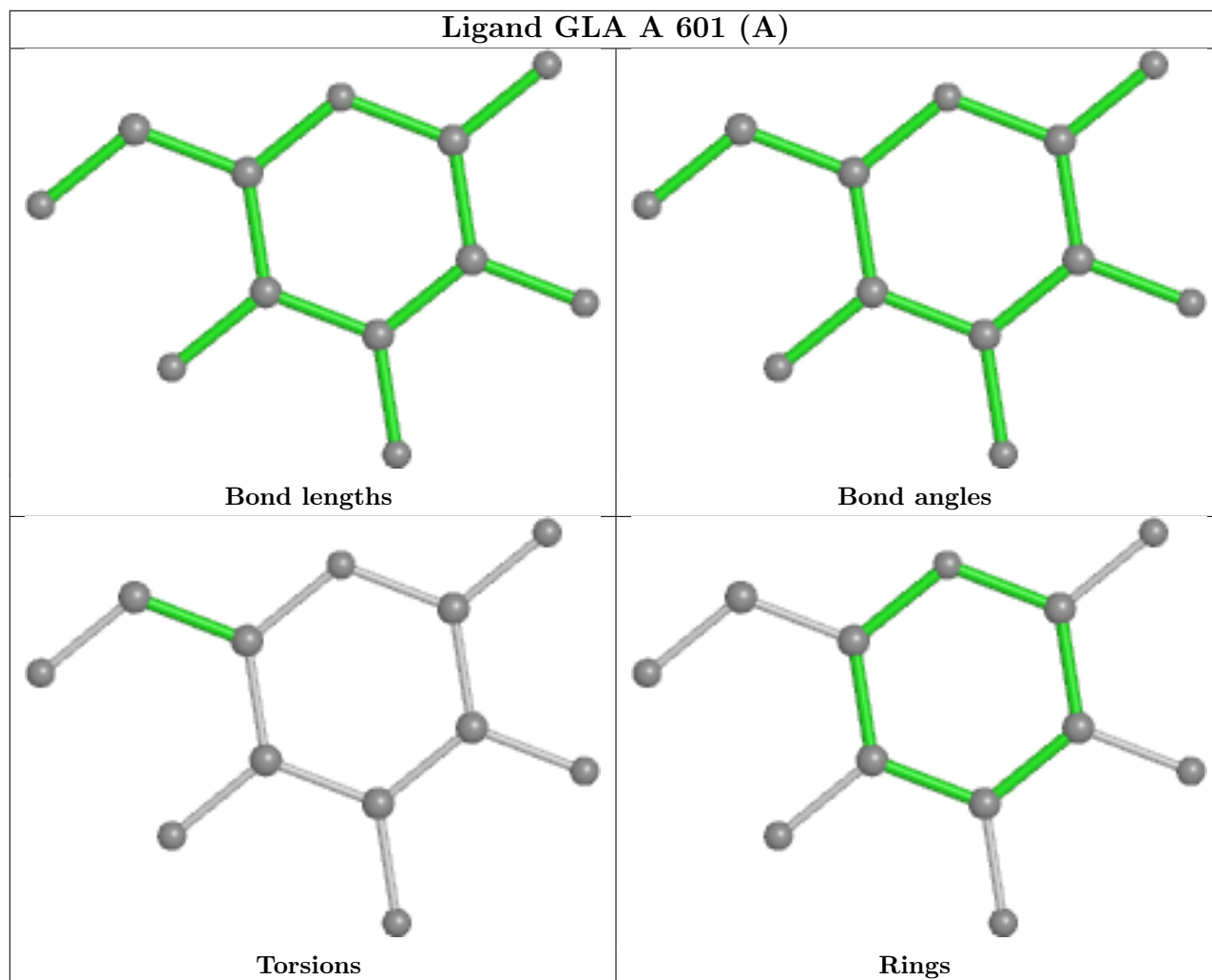
There are no ring outliers.

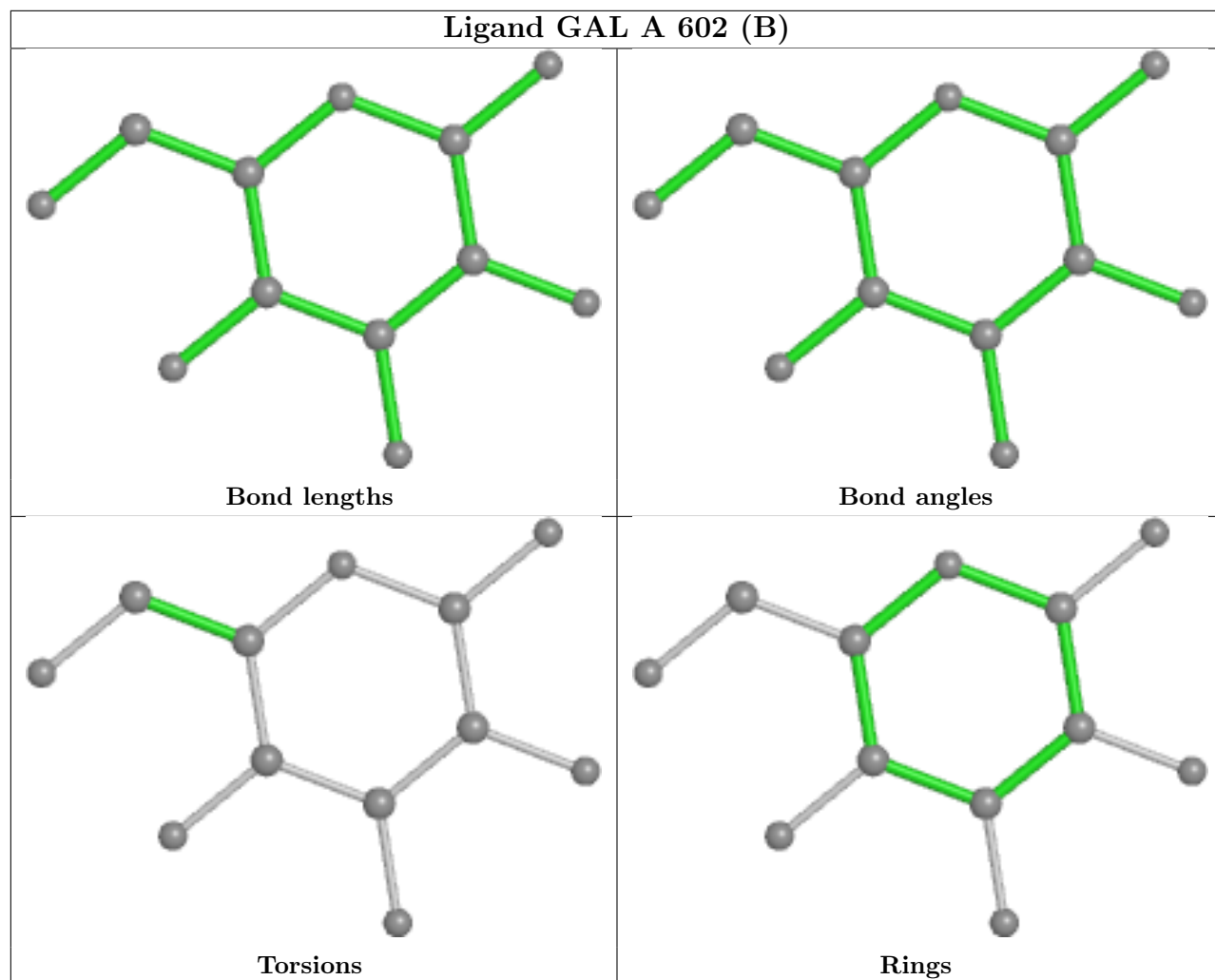
No monomer is involved in short contacts.

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 than 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.









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

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	A	510/541 (94%)	-0.10	15 (2%) 51 61	25, 33, 51, 77	0
1	B	510/541 (94%)	-0.07	29 (5%) 23 32	24, 33, 58, 89	0
All	All	1020/1082 (94%)	-0.09	44 (4%) 35 45	24, 33, 54, 89	0

All (44) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	546	THR	5.7
1	B	251	ASP	4.2
1	A	546	THR	4.2
1	B	252	THR	4.1
1	B	33	GLN	4.1
1	A	251	ASP	4.1
1	B	525	ARG	4.1
1	A	505	VAL	3.4
1	B	533	LYS	3.4
1	B	434	SER	3.2
1	A	252	THR	3.1
1	B	523	VAL	3.0
1	B	532	LEU	3.0
1	B	504	SER	3.0
1	A	403	LYS	3.0
1	B	505	VAL	2.9
1	B	403	LYS	2.9
1	B	545	THR	2.8
1	B	254	GLN	2.7
1	A	504	SER	2.7
1	B	530	ILE	2.7
1	A	295	GLN	2.6
1	A	386	CYS	2.6
1	B	316	LEU	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	295	GLN	2.6
1	B	522	LYS	2.6
1	B	503	GLY	2.5
1	B	535	THR	2.5
1	B	257	HIS	2.5
1	A	404	LEU	2.4
1	A	395	PHE	2.3
1	B	534	VAL	2.3
1	A	342	TYR	2.3
1	A	533	LYS	2.3
1	B	249	GLY	2.2
1	B	250	SER	2.2
1	B	259	VAL	2.1
1	B	435	SER	2.1
1	A	76	ALA	2.1
1	A	503	GLY	2.1
1	B	386	CYS	2.1
1	B	507	LYS	2.1
1	A	382	TYR	2.0
1	B	35	SER	2.0

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	EDO	B	609	4/4	0.86	0.29	59,63,64,64	1
4	EDO	A	607	4/4	0.89	0.19	44,47,49,50	1
4	EDO	B	607	4/4	0.91	0.12	42,50,64,64	1

Continued on next page...

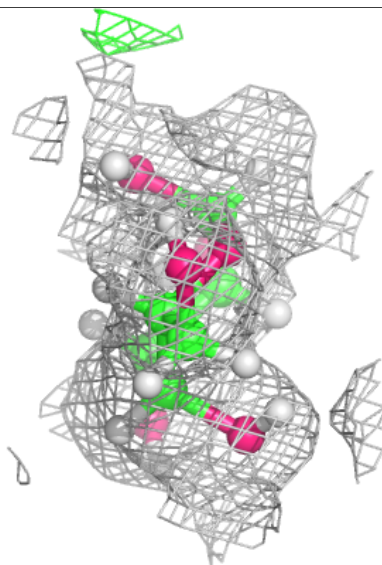
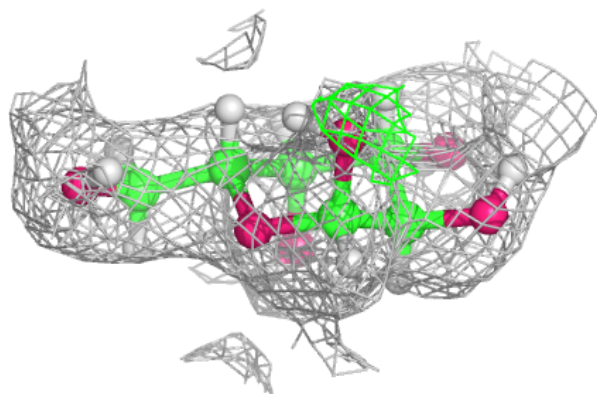
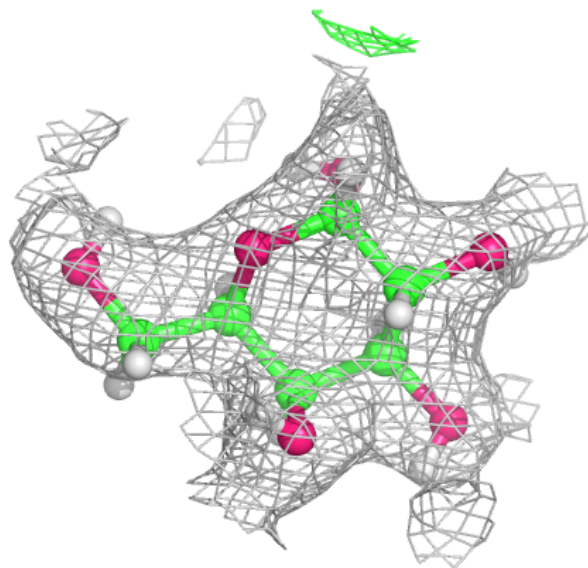
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	GLA	B	601[A]	12/12	0.91	0.12	28,31,33,33	24
4	EDO	B	605	4/4	0.92	0.12	44,45,46,46	1
3	GAL	B	602[B]	12/12	0.93	0.12	37,50,51,53	24
3	GAL	A	602[B]	12/12	0.94	0.12	40,50,52,53	24
4	EDO	A	606	4/4	0.94	0.11	38,39,43,43	1
2	GLA	A	601[A]	12/12	0.95	0.12	26,29,31,32	24
4	EDO	B	606	4/4	0.95	0.09	35,37,38,38	1
4	EDO	B	604	4/4	0.96	0.10	32,38,42,44	1
4	EDO	A	604	4/4	0.98	0.14	32,35,37,38	1
4	EDO	A	605	4/4	0.98	0.15	29,32,36,36	1
4	EDO	A	603	4/4	0.98	0.17	39,46,55,57	1
5	CL	B	610	1/1	0.98	0.08	48,48,48,48	0
4	EDO	B	603	4/4	0.99	0.13	32,34,37,37	1
4	EDO	B	608	4/4	0.99	0.08	36,36,41,42	1
5	CL	A	608	1/1	1.00	0.08	46,46,46,46	0

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.

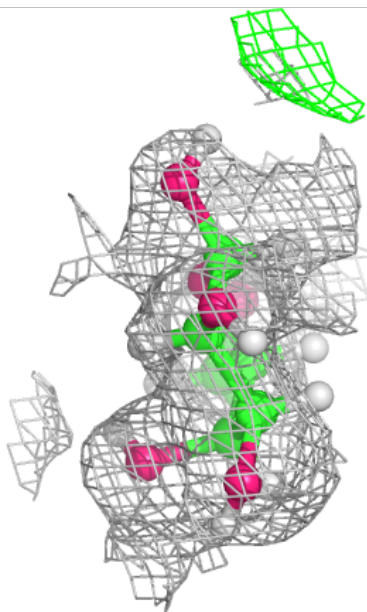
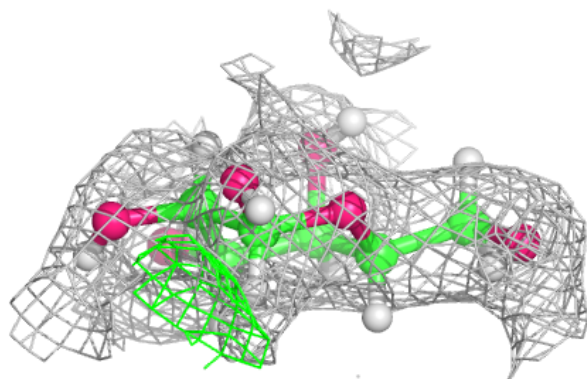
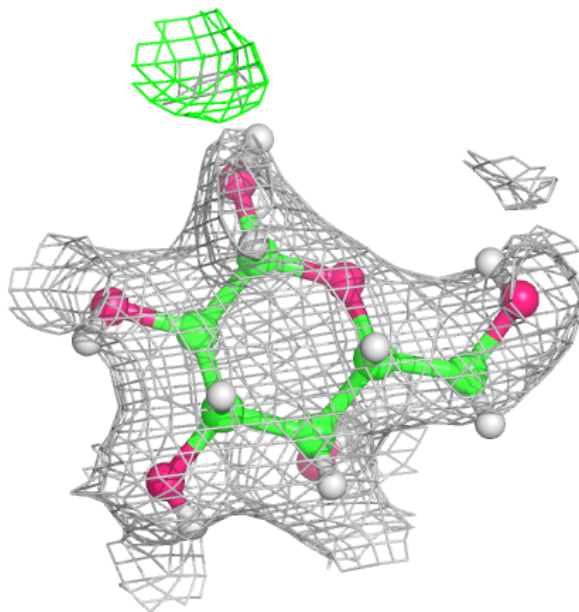
Electron density around GLA B 601 (A):

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



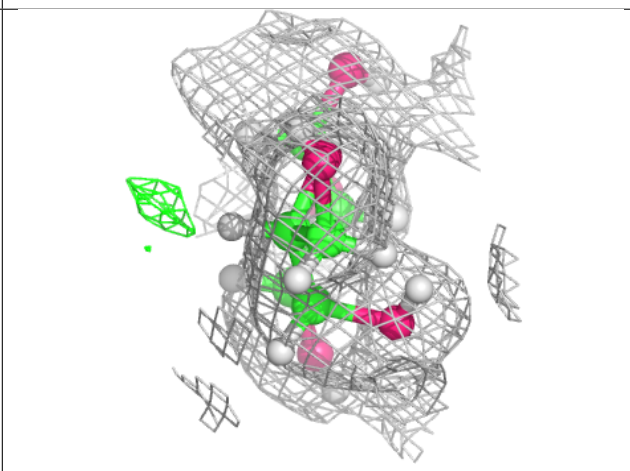
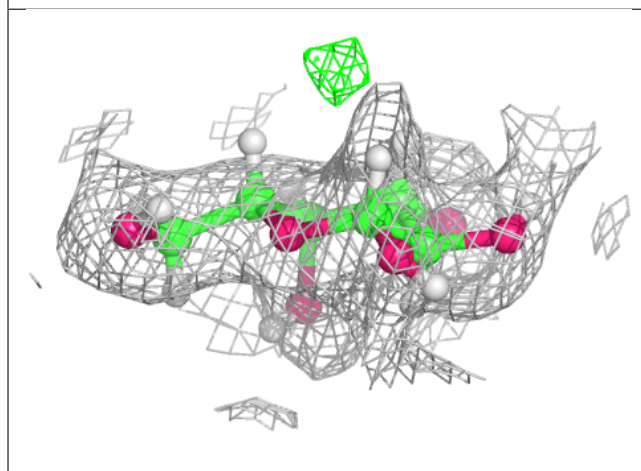
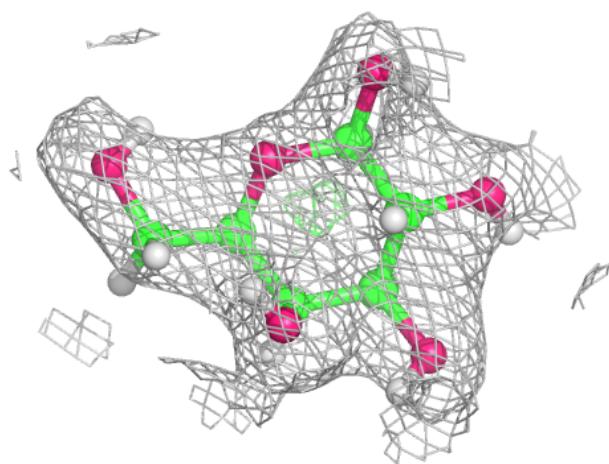
Electron density around GAL B 602 (B):

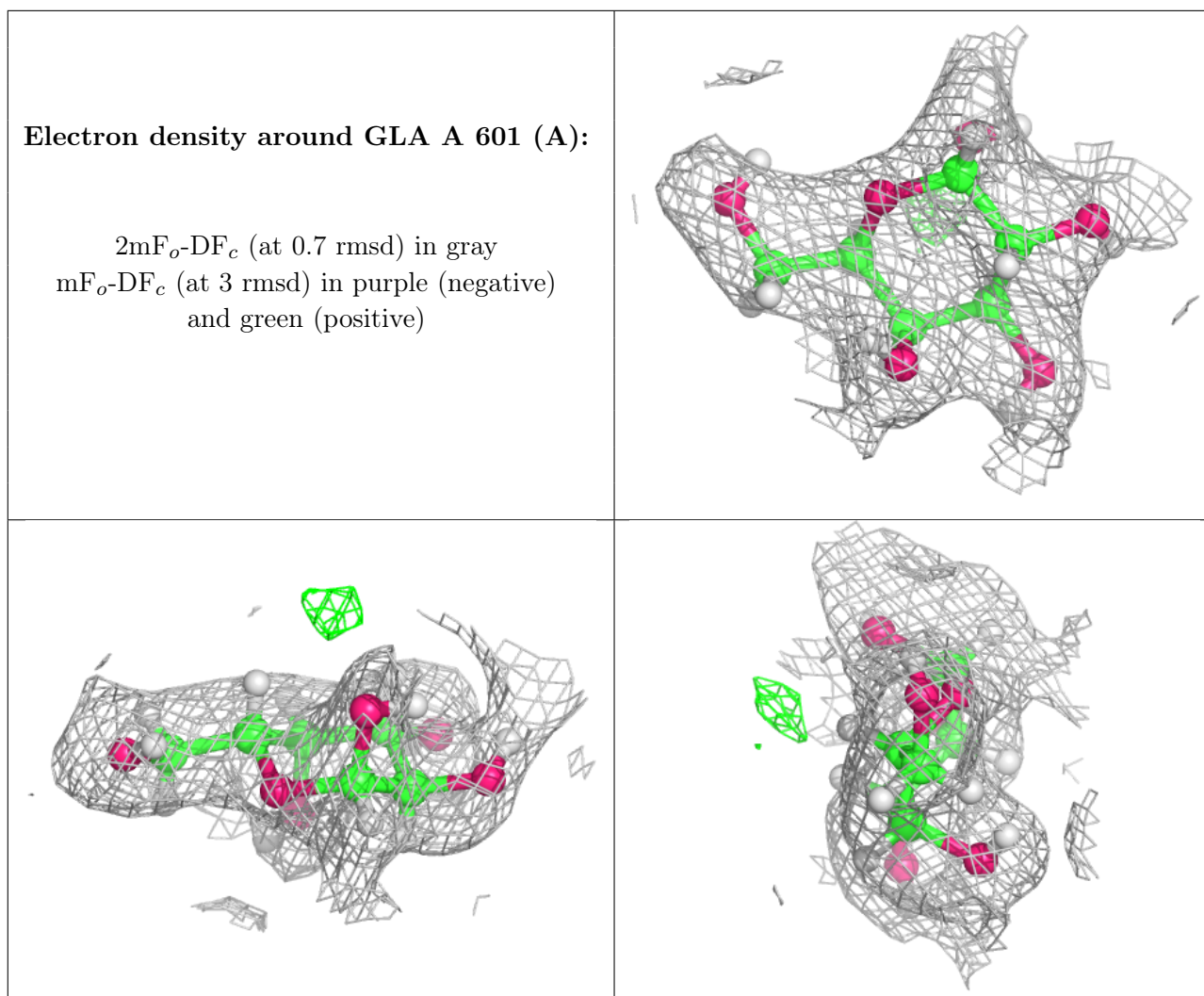
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around GAL A 602 (B):

$2mF_o-DF_c$ (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.