Full wwPDB X-ray Structure Validation Report ⓘ

Nov 22, 2023 – 04:46 PM JST

PDB ID	:	7EXQ
Title	:	Crystal structure of alkaline alpha-galactosidase D383A mutant from Arabidopsis thaliana complexed with product-galactose and sucrose.
Authors	:	Chuankhayan, P.; Guan, H.H.; Lin, C.C.; Chen, N.C.; Huang, Y.C.; Yoshimura, M.; Nakagawa, A.; Lee, R.H.; Chen, C.J.
Deposited on	:	2021-05-28
Resolution	:	2.20 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:

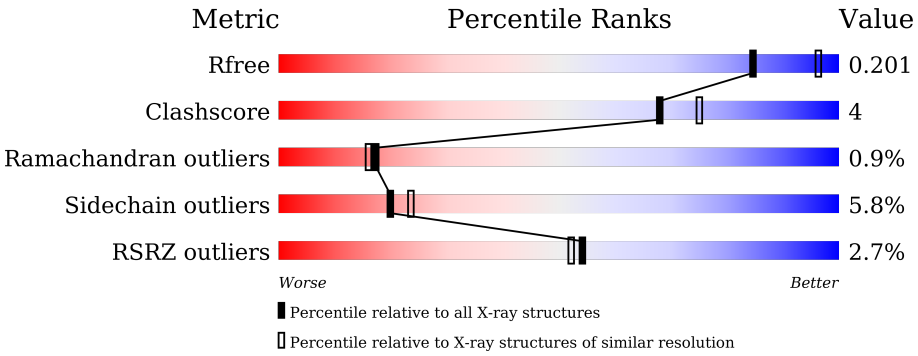
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	749	4% 83% 11%
1	B	749	% 84% 9%
2	C	2	50% 50%
2	D	2	50% 50%

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
3	GAL	A	801	X	-	X	-
3	GAL	B	801	X	-	-	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 11728 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.

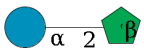
- Molecule 1 is a protein called Probable galactinol–sucrose galactosyltransferase 6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	718	Total	C	N	O	S	0	0	0
			5621	3573	969	1050	29			
1	A	718	Total	C	N	O	S	0	0	0
			5621	3573	969	1050	29			

There are 4 discrepancies between the modelled and reference sequences:

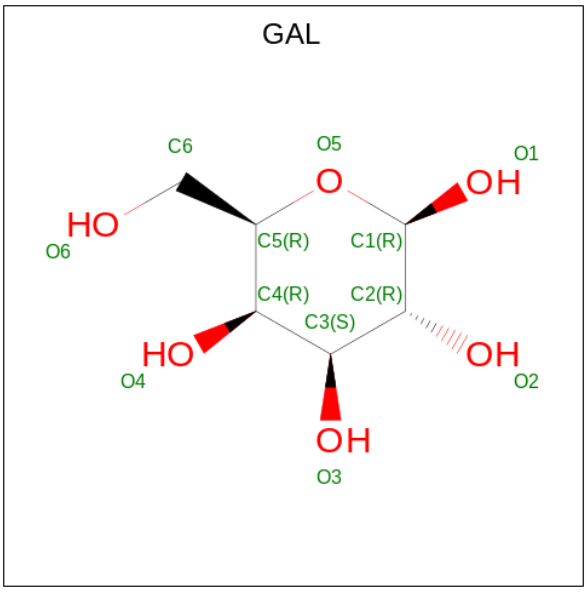
Chain	Residue	Modelled	Actual	Comment	Reference
B	302	ARG	LYS	conflict	UNP Q8RX87
B	383	ALA	ASP	engineered mutation	UNP Q8RX87
A	302	ARG	LYS	conflict	UNP Q8RX87
A	383	ALA	ASP	engineered mutation	UNP Q8RX87

- Molecule 2 is an oligosaccharide called beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	2	Total	C	O		0	0	0
			23	12	11				
2	D	2	Total	C	O		0	0	0
			23	12	11				

- Molecule 3 is beta-D-galactopyranose (three-letter code: GAL) (formula: $C_6H_{12}O_6$) (labeled as "Ligand of Interest" by depositor).



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

- Molecule 4 is water.

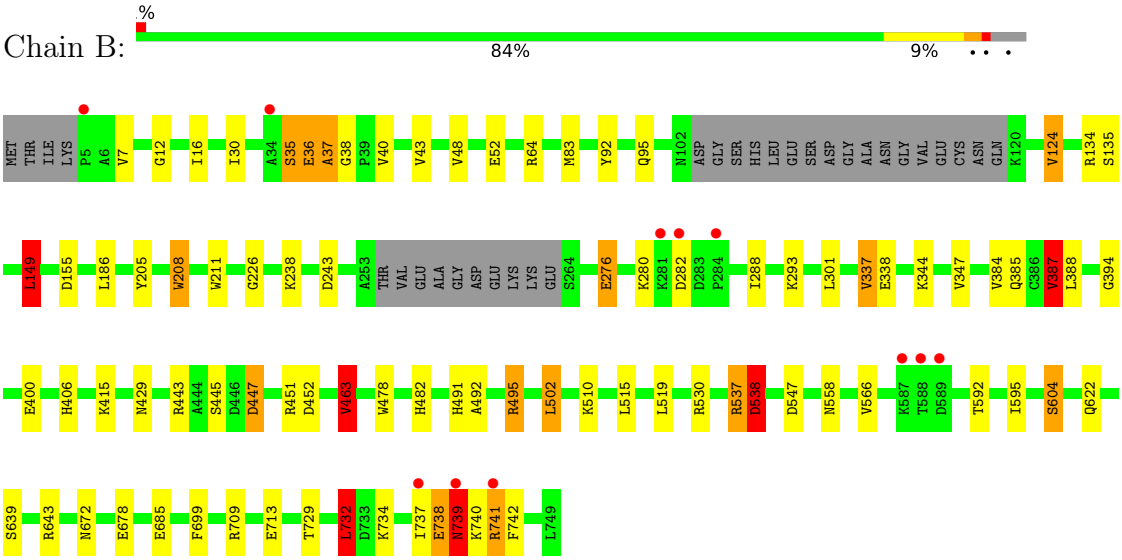
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	249	Total O 249 249	0	0
4	A	167	Total O 167 167	0	0

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.

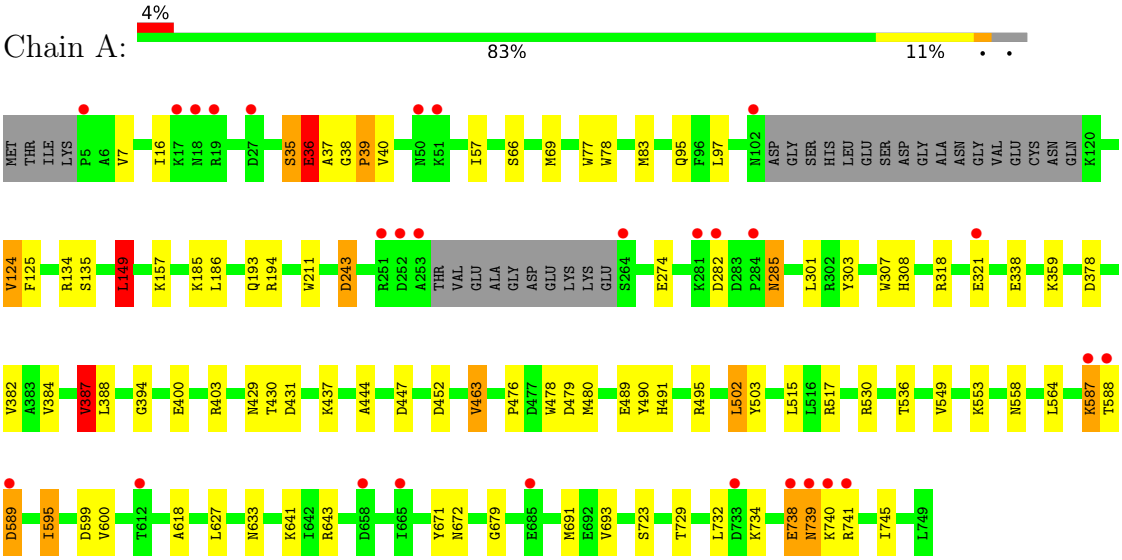
• Molecule 1: Probable galactinol–sucrose galactosyltransferase 6

Chain B:



• Molecule 1: Probable galactinol–sucrose galactosyltransferase 6

Chain A:



• Molecule 2: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose

Chain C:

• Molecule 2: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	97.71Å 104.10Å 182.42Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.88 – 2.20 27.68 – 2.20	Depositor EDS
% Data completeness (in resolution range)	91.0 (29.88-2.20) 91.1 (27.68-2.20)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.32 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.193 , 0.255 0.201 , 0.201	Depositor DCC
R_{free} test set	4367 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	32.1	Xtriage
Anisotropy	0.012	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 32.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	11728	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.79% 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: FRU, GLC, GAL

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	0.81	1/5755 (0.0%)	0.91	6/7794 (0.1%)
1	B	0.93	3/5755 (0.1%)	1.05	20/7794 (0.3%)
All	All	0.87	4/11510 (0.0%)	0.98	26/15588 (0.2%)

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	6
All	All	0	9

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	604	SER	CB-OG	-6.64	1.33	1.42
1	B	643	ARG	CZ-NH1	6.63	1.41	1.33
1	A	36	GLU	CG-CD	5.92	1.60	1.51
1	B	208	TRP	CE3-CZ3	5.92	1.48	1.38

All (26) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	643	ARG	NE-CZ-NH2	-17.21	111.70	120.30
1	B	643	ARG	NE-CZ-NH1	12.16	126.38	120.30
1	B	530	ARG	NE-CZ-NH1	11.10	125.85	120.30
1	B	530	ARG	NE-CZ-NH2	-10.27	115.17	120.30
1	B	447	ASP	CB-CG-OD1	8.09	125.58	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	124	VAL	CB-CA-C	-7.75	96.67	111.40
1	B	149	LEU	CA-CB-CG	7.31	132.10	115.30
1	B	547	ASP	CB-CG-OD1	7.19	124.77	118.30
1	A	194	ARG	NE-CZ-NH1	7.17	123.88	120.30
1	A	124	VAL	CB-CA-C	-6.81	98.46	111.40
1	B	387	VAL	CB-CA-C	-6.47	99.10	111.40
1	B	155	ASP	CB-CG-OD2	-6.41	112.53	118.30
1	A	387	VAL	CB-CA-C	-6.23	99.57	111.40
1	B	538	ASP	CB-CG-OD1	-6.09	112.82	118.30
1	B	537	ARG	NE-CZ-NH1	6.09	123.35	120.30
1	A	318	ARG	NE-CZ-NH1	6.04	123.32	120.30
1	B	443	ARG	NE-CZ-NH2	-5.86	117.37	120.30
1	B	463	VAL	CG1-CB-CG2	5.82	120.21	110.90
1	B	447	ASP	CB-CG-OD2	-5.78	113.09	118.30
1	B	495	ARG	NE-CZ-NH2	-5.78	117.41	120.30
1	B	37	ALA	N-CA-C	5.57	126.05	111.00
1	B	709	ARG	NE-CZ-NH1	5.57	123.09	120.30
1	B	451	ARG	NE-CZ-NH2	-5.47	117.56	120.30
1	B	64	ARG	NE-CZ-NH1	5.30	122.95	120.30
1	A	530	ARG	CB-CA-C	-5.20	100.00	110.40
1	A	149	LEU	CA-CB-CG	5.10	127.02	115.30

There are no chirality outliers.

All (9) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	35	SER	Peptide
1	A	37	ALA	Peptide
1	A	394	GLY	Peptide
1	B	12	GLY	Peptide
1	B	35	SER	Peptide
1	B	37	ALA	Peptide
1	B	394	GLY	Peptide
1	B	732	LEU	Peptide
1	B	740	LYS	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	5621	0	5542	48	0
1	B	5621	0	5542	37	0
2	C	23	0	20	1	0
2	D	23	0	20	1	0
3	A	12	0	10	6	0
3	B	12	0	8	3	0
4	A	167	0	0	2	0
4	B	249	0	0	1	0
All	All	11728	0	11142	86	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.

All (86) 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)
2:D:2:FRU:C5	2:D:2:FRU:O5	1.67	1.17
2:C:2:FRU:C5	2:C:2:FRU:O5	1.69	1.08
1:A:558:ASN:HD21	1:A:672:ASN:HD21	1.16	0.92
1:B:738:GLU:HB3	1:B:739:ASN:HB3	1.62	0.81
1:B:95:GLN:HE21	1:B:134:ARG:HH21	1.30	0.78
1:B:738:GLU:HB3	1:B:739:ASN:CB	2.20	0.71
1:B:558:ASN:HD21	1:B:672:ASN:HD21	1.41	0.66
1:B:447:ASP:OD1	3:B:801:GAL:H2	1.97	0.64
1:B:211:TRP:CZ2	3:B:801:GAL:H62	2.33	0.64
1:A:558:ASN:ND2	1:A:672:ASN:HD21	1.93	0.62
1:B:463:VAL:HG13	1:B:478:TRP:CD2	2.35	0.61
1:A:95:GLN:HE21	1:A:134:ARG:HH21	1.49	0.60
1:B:35:SER:O	1:B:38:GLY:HA2	2.01	0.60
1:B:502:LEU:HD21	1:B:519:LEU:HD22	1.86	0.56
1:B:447:ASP:CG	3:B:801:GAL:H2	2.26	0.56
1:B:7:VAL:HG13	1:B:16:ILE:CD1	2.35	0.55
1:A:211:TRP:CZ2	3:A:801:GAL:H62	2.42	0.55
1:B:463:VAL:HG13	1:B:478:TRP:CE2	2.41	0.55
1:A:515:LEU:O	1:A:515:LEU:HD23	2.08	0.54
1:B:95:GLN:HE22	1:B:429:ASN:HA	1.72	0.54
1:B:491:HIS:HD1	1:B:495:ARG:HH12	1.56	0.54
1:A:95:GLN:HE22	1:A:429:ASN:HA	1.73	0.54
1:B:92:TYR:CE2	1:B:338:GLU:HG2	2.44	0.53
1:A:463:VAL:HG13	1:A:478:TRP:CE2	2.43	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:447:ASP:OD2	3:A:801:GAL:H2	2.09	0.53
1:A:564:LEU:HD23	1:A:595:ILE:HG13	1.91	0.52
1:A:564:LEU:CD2	1:A:595:ILE:HG13	2.40	0.51
1:A:7:VAL:HG22	1:A:16:ILE:HD12	1.93	0.51
1:A:447:ASP:CG	3:A:801:GAL:H2	2.30	0.51
1:B:276:GLU:H	1:B:276:GLU:CD	2.14	0.51
1:A:135:SER:OG	1:A:149:LEU:HD13	2.11	0.50
1:B:135:SER:OG	1:B:149:LEU:HD13	2.11	0.50
1:A:243:ASP:OD1	3:A:801:GAL:O4	2.28	0.50
1:A:495:ARG:HB3	1:A:502:LEU:HD22	1.93	0.50
1:A:589:ASP:N	1:A:589:ASP:OD1	2.44	0.50
1:A:463:VAL:HG13	1:A:478:TRP:CD2	2.47	0.49
1:A:97:LEU:HB3	1:A:125:PHE:HB2	1.94	0.49
1:A:517:ARG:NH1	4:A:903:HOH:O	2.45	0.49
1:A:641:LYS:NZ	4:A:902:HOH:O	2.45	0.49
1:A:211:TRP:CE2	3:A:801:GAL:H62	2.48	0.49
1:A:447:ASP:OD1	3:A:801:GAL:H2	2.13	0.49
1:B:699:PHE:HB2	1:B:732:LEU:HD22	1.94	0.48
1:A:444:ALA:HB2	1:A:476:PRO:HB3	1.94	0.48
1:A:679:GLY:O	1:A:691:MET:HA	2.14	0.48
1:B:592:THR:HG1	1:B:639:SER:HG	1.62	0.48
1:B:445:SER:HB3	1:B:463:VAL:HG23	1.97	0.47
1:A:303:TYR:HA	1:A:378:ASP:OD2	2.14	0.47
1:A:491:HIS:HD1	1:A:495:ARG:HH12	1.61	0.47
1:A:243:ASP:HA	1:A:307:TRP:HB2	1.96	0.46
1:B:492:ALA:CB	1:B:515:LEU:HD21	2.45	0.46
1:A:515:LEU:HD23	1:A:515:LEU:C	2.36	0.46
1:A:480:MET:HA	1:A:503:TYR:O	2.16	0.45
1:B:384:VAL:O	1:B:387:VAL:HG22	2.17	0.45
1:A:400:GLU:OE1	1:A:403:ARG:NH2	2.48	0.45
1:B:385:GLN:NE2	1:B:406:HIS:HE1	2.14	0.45
1:A:185:LYS:NZ	1:A:193:GLN:HE21	2.14	0.45
1:A:38:GLY:O	1:A:39:PRO:C	2.56	0.44
1:B:463:VAL:CG1	1:B:478:TRP:CD2	3.00	0.44
1:A:95:GLN:NE2	1:A:134:ARG:HH21	2.13	0.44
1:B:208:TRP:CH2	1:B:226:GLY:HA3	2.53	0.44
1:B:678:GLU:HG3	1:B:742:PHE:CE1	2.52	0.44
1:B:738:GLU:HB3	1:B:739:ASN:CA	2.47	0.44
1:B:482:HIS:HD2	4:B:1042:HOH:O	2.00	0.44
1:B:415:LYS:HD2	1:A:437:LYS:HD2	1.99	0.43
1:A:77:TRP:HB2	1:A:78:TRP:CD1	2.54	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:737:ILE:O	1:B:738:GLU:HB2	2.19	0.43
1:B:495:ARG:HB3	1:B:502:LEU:HD22	2.01	0.43
1:A:536:THR:HG21	1:A:599:ASP:HB3	1.99	0.43
1:A:693:VAL:HG21	1:A:745:ILE:HD12	2.01	0.43
1:B:385:GLN:HE22	1:B:406:HIS:HE1	1.66	0.42
1:A:384:VAL:O	1:A:387:VAL:HG22	2.19	0.42
1:A:95:GLN:NE2	1:A:430:THR:HG23	2.35	0.42
1:A:738:GLU:HB3	1:A:739:ASN:CB	2.49	0.42
1:B:95:GLN:NE2	1:B:134:ARG:HH21	2.08	0.42
1:A:553:LYS:HD3	1:A:595:ILE:HD12	2.01	0.42
1:B:566:VAL:HG21	1:B:595:ILE:HD11	2.01	0.42
1:B:337:VAL:HG13	1:B:344:LYS:HE3	2.02	0.42
1:A:489:GLU:OE1	1:A:671:TYR:OH	2.37	0.42
1:A:307:TRP:CE3	1:A:308:HIS:HA	2.54	0.42
1:B:205:TYR:HB3	1:B:238:LYS:HB2	2.01	0.41
1:A:308:HIS:O	1:A:382:VAL:HA	2.20	0.41
1:A:490:TYR:CE1	1:A:643:ARG:HA	2.55	0.41
1:B:685:GLU:H	1:B:685:GLU:CD	2.24	0.41
1:A:618:ALA:HB1	1:A:627:LEU:HD11	2.03	0.41
1:A:285:ASN:OD1	1:A:285:ASN:N	2.54	0.41
1:A:738:GLU:HB2	1:A:740:LYS:HB2	2.03	0.40

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	712/749 (95%)	671 (94%)	34 (5%)	7 (1%)	15	14
1	B	712/749 (95%)	676 (95%)	30 (4%)	6 (1%)	19	19
All	All	1424/1498 (95%)	1347 (95%)	64 (4%)	13 (1%)	17	16

All (13) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	36	GLU
1	B	741	ARG
1	A	36	GLU
1	A	587	LYS
1	B	282	ASP
1	B	739	ASN
1	A	282	ASP
1	B	738	GLU
1	B	538	ASP
1	A	39	PRO
1	A	738	GLU
1	A	479	ASP
1	A	600	VAL

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	615/640 (96%)	579 (94%)	36 (6%)	19	23
1	B	615/640 (96%)	580 (94%)	35 (6%)	20	24
All	All	1230/1280 (96%)	1159 (94%)	71 (6%)	20	23

All (71) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	30	ILE
1	B	36	GLU
1	B	40	VAL
1	B	43	VAL
1	B	48	VAL
1	B	52	GLU
1	B	83	MET
1	B	124	VAL
1	B	149	LEU
1	B	186	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	243	ASP
1	B	276	GLU
1	B	280	LYS
1	B	288	ILE
1	B	293	LYS
1	B	301	LEU
1	B	337	VAL
1	B	347	VAL
1	B	387	VAL
1	B	388	LEU
1	B	400	GLU
1	B	452	ASP
1	B	463	VAL
1	B	502	LEU
1	B	510	LYS
1	B	537	ARG
1	B	538	ASP
1	B	604	SER
1	B	622	GLN
1	B	713	GLU
1	B	729	THR
1	B	732	LEU
1	B	734	LYS
1	B	739	ASN
1	B	741	ARG
1	A	35	SER
1	A	36	GLU
1	A	40	VAL
1	A	57	ILE
1	A	66	SER
1	A	69	MET
1	A	83	MET
1	A	124	VAL
1	A	149	LEU
1	A	157	LYS
1	A	186	LEU
1	A	243	ASP
1	A	274	GLU
1	A	285	ASN
1	A	301	LEU
1	A	321	GLU
1	A	338	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	359	LYS
1	A	387	VAL
1	A	388	LEU
1	A	431	ASP
1	A	452	ASP
1	A	463	VAL
1	A	502	LEU
1	A	549	VAL
1	A	587	LYS
1	A	588	THR
1	A	589	ASP
1	A	595	ILE
1	A	633	ASN
1	A	723	SER
1	A	729	THR
1	A	732	LEU
1	A	734	LYS
1	A	739	ASN
1	A	741	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (17) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	95	GLN
1	B	138	GLN
1	B	168	HIS
1	B	175	GLN
1	B	193	GLN
1	B	385	GLN
1	B	456	HIS
1	B	482	HIS
1	B	558	ASN
1	B	584	HIS
1	A	95	GLN
1	A	168	HIS
1	A	175	GLN
1	A	193	GLN
1	A	385	GLN
1	A	456	HIS
1	A	558	ASN

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

4 monosaccharides 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	GLC	C	1	2	11,11,12	1.94	4 (36%)	15,15,17	1.17	1 (6%)
2	FRU	C	2	2	11,12,12	5.46	6 (54%)	10,18,18	1.20	1 (10%)
2	GLC	D	1	2	11,11,12	1.81	3 (27%)	15,15,17	1.33	3 (20%)
2	FRU	D	2	2	11,12,12	5.48	6 (54%)	10,18,18	1.20	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	GLC	C	1	2	-	2/2/19/22	0/1/1/1
2	FRU	C	2	2	-	3/5/24/24	0/1/1/1
2	GLC	D	1	2	-	0/2/19/22	0/1/1/1
2	FRU	D	2	2	-	2/5/24/24	0/1/1/1

All (19) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	2	FRU	O5-C5	11.84	1.69	1.43

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	2	FRU	O5-C5	10.74	1.67	1.43
2	C	2	FRU	C4-C5	-9.06	1.29	1.53
2	D	2	FRU	C4-C5	-9.01	1.30	1.53
2	D	2	FRU	O5-C2	-8.45	1.30	1.43
2	C	2	FRU	O5-C2	-6.13	1.33	1.43
2	C	2	FRU	O2-C2	5.98	1.50	1.40
2	D	2	FRU	O3-C3	-5.80	1.31	1.42
2	C	2	FRU	O3-C3	-4.61	1.33	1.42
2	D	2	FRU	O2-C2	4.32	1.48	1.40
2	C	1	GLC	O5-C1	4.21	1.50	1.43
2	D	1	GLC	O3-C3	3.30	1.50	1.43
2	C	1	GLC	C2-C3	-3.10	1.47	1.52
2	D	1	GLC	O5-C1	2.84	1.48	1.43
2	D	2	FRU	O4-C4	2.66	1.49	1.43
2	C	2	FRU	O4-C4	2.64	1.49	1.43
2	D	1	GLC	O5-C5	2.58	1.48	1.43
2	C	1	GLC	O5-C5	2.21	1.47	1.43
2	C	1	GLC	O3-C3	2.15	1.48	1.43

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	2	FRU	O2-C2-O5	2.79	114.89	109.50
2	C	1	GLC	O5-C5-C6	2.66	111.38	107.20
2	D	1	GLC	C6-C5-C4	2.50	118.87	113.00
2	D	1	GLC	O3-C3-C4	2.34	115.77	110.35
2	D	1	GLC	O5-C5-C6	2.15	110.58	107.20

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	2	FRU	O1-C1-C2-O2
2	D	2	FRU	O1-C1-C2-O2
2	C	1	GLC	O5-C5-C6-O6
2	C	1	GLC	C4-C5-C6-O6
2	D	2	FRU	O1-C1-C2-O5
2	C	2	FRU	O1-C1-C2-C3
2	C	2	FRU	O5-C5-C6-O6

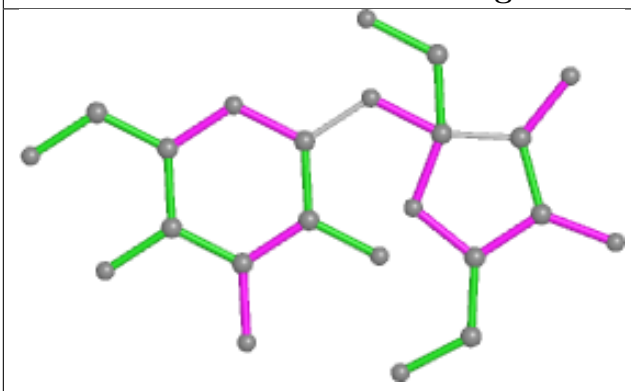
There are no ring outliers.

2 monomers are involved in 2 short contacts:

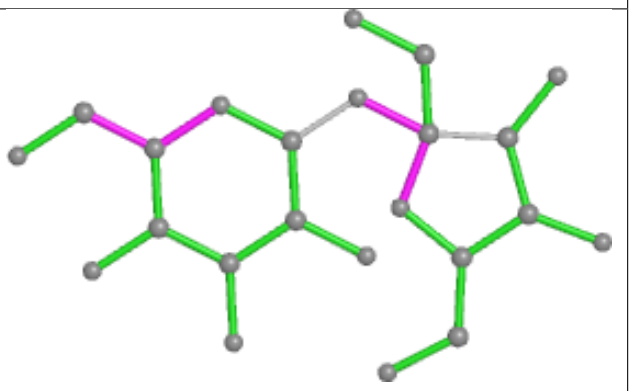
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	2	FRU	1	0
2	D	2	FRU	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

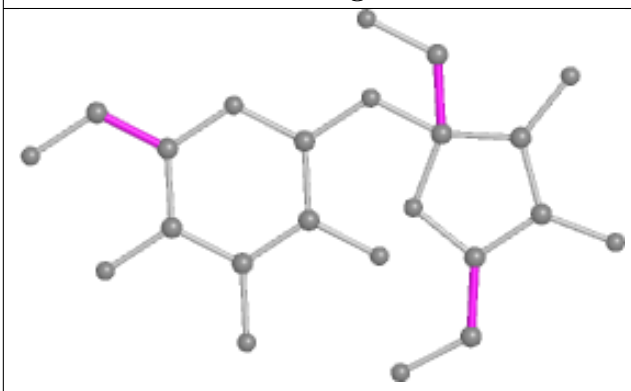
Oligosaccharide Chain C



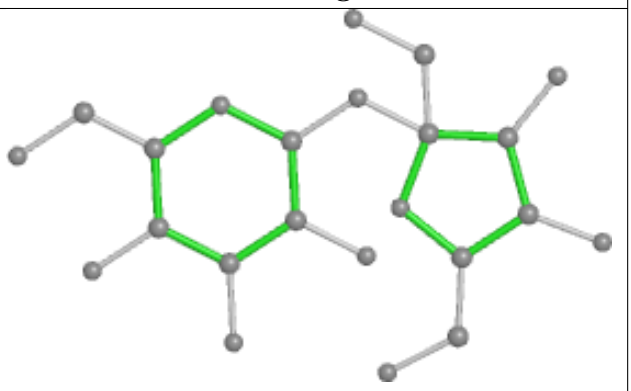
Bond lengths



Bond angles

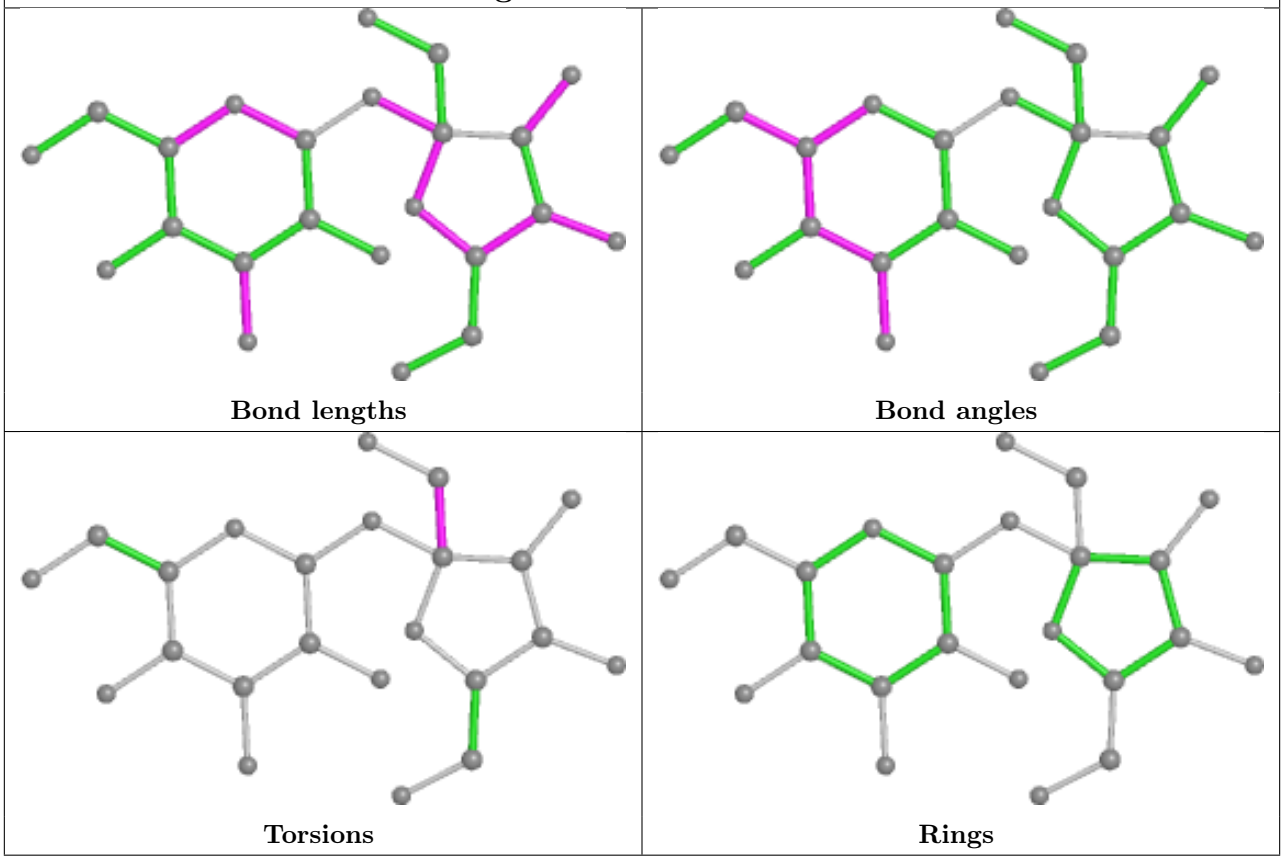


Torsions



Rings

Oligosaccharide Chain D



5.6 Ligand geometry (i)

2 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	GAL	A	801	-	12,12,12	1.66	3 (25%)	17,17,17	2.65	10 (58%)
3	GAL	B	801	-	12,12,12	1.21	2 (16%)	17,17,17	4.84	12 (70%)

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	GAL	A	801	-	3/3/5/5	0/2/22/22	0/1/1/1
3	GAL	B	801	-	3/3/5/5	0/2/22/22	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	801	GAL	C4-C3	-3.54	1.43	1.52
3	A	801	GAL	C3-C2	-3.13	1.44	1.52
3	B	801	GAL	C4-C3	-2.42	1.46	1.52
3	A	801	GAL	O4-C4	-2.07	1.38	1.43
3	B	801	GAL	O1-C1	-2.06	1.33	1.39

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	801	GAL	C3-C4-C5	9.71	127.55	110.24
3	B	801	GAL	O5-C5-C6	8.90	128.56	106.44
3	B	801	GAL	O5-C1-C2	7.52	123.71	110.28
3	B	801	GAL	O2-C2-C3	6.75	125.96	110.35
3	A	801	GAL	O5-C5-C6	5.68	120.55	106.44
3	B	801	GAL	O4-C4-C3	5.62	123.33	110.35
3	B	801	GAL	C4-C3-C2	-4.53	102.92	110.82
3	B	801	GAL	O3-C3-C4	4.29	120.26	110.35
3	A	801	GAL	O2-C2-C1	4.02	118.48	109.16
3	A	801	GAL	C6-C5-C4	3.88	122.10	113.00
3	B	801	GAL	O1-C1-C2	3.62	119.22	109.03
3	B	801	GAL	C1-O5-C5	-3.41	107.24	113.66
3	B	801	GAL	O6-C6-C5	3.28	122.53	111.29
3	A	801	GAL	C4-C3-C2	-3.13	105.35	110.82
3	A	801	GAL	O2-C2-C3	2.85	116.93	110.35
3	A	801	GAL	C1-O5-C5	-2.79	108.40	113.66
3	B	801	GAL	O2-C2-C1	2.49	114.92	109.16
3	B	801	GAL	C6-C5-C4	2.44	118.73	113.00
3	A	801	GAL	O1-C1-C2	2.42	115.85	109.03
3	A	801	GAL	C3-C4-C5	2.22	114.20	110.24
3	A	801	GAL	O6-C6-C5	-2.08	104.16	111.29
3	A	801	GAL	O5-C1-C2	2.05	113.94	110.28

All (6) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	B	801	GAL	C2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atom
3	B	801	GAL	C5
3	B	801	GAL	C4
3	A	801	GAL	C4
3	A	801	GAL	C2
3	A	801	GAL	C5

There are no torsion outliers.

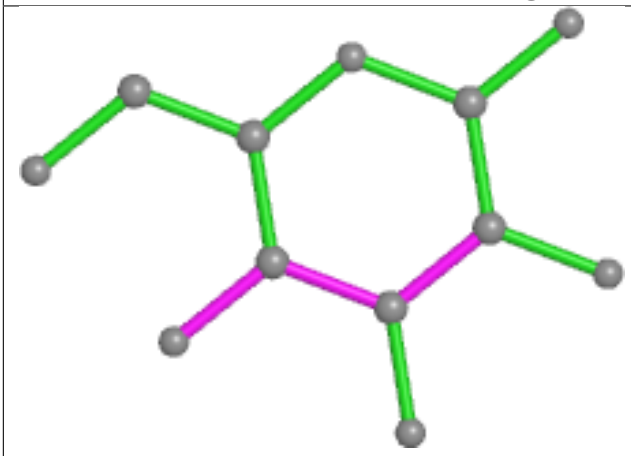
There are no ring outliers.

2 monomers are involved in 9 short contacts:

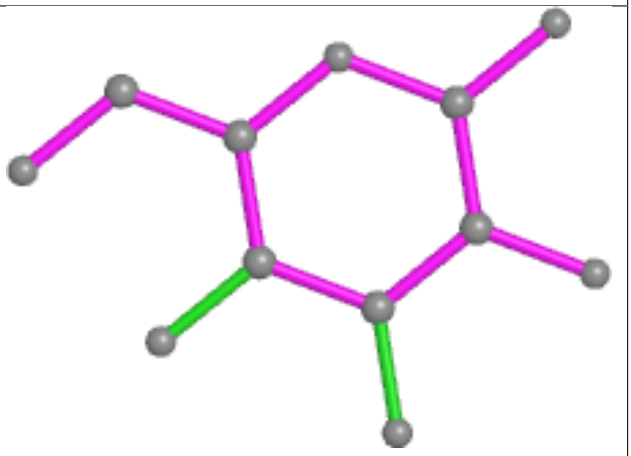
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	801	GAL	6	0
3	B	801	GAL	3	0

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.

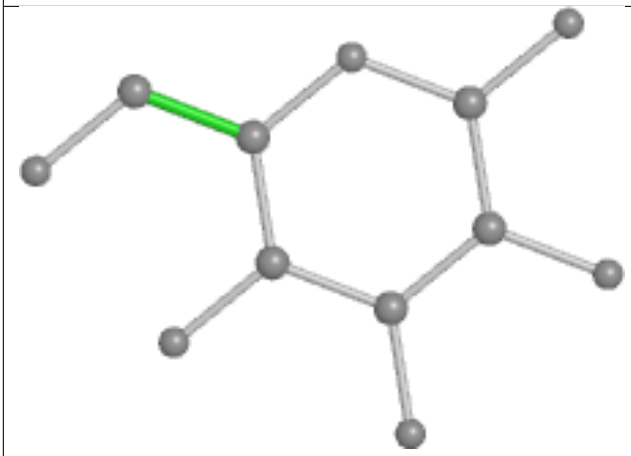
Ligand GAL A 801



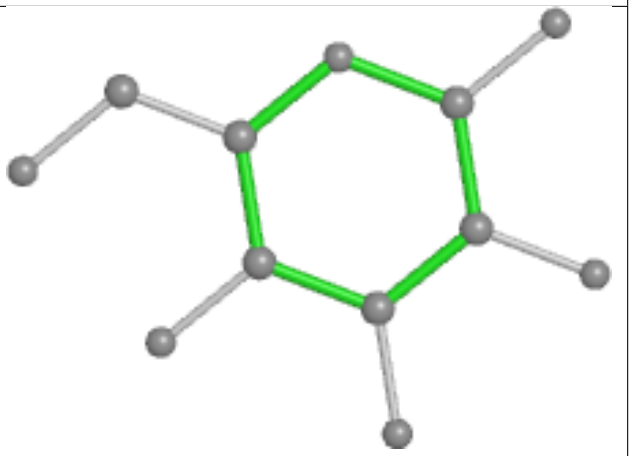
Bond lengths



Bond angles

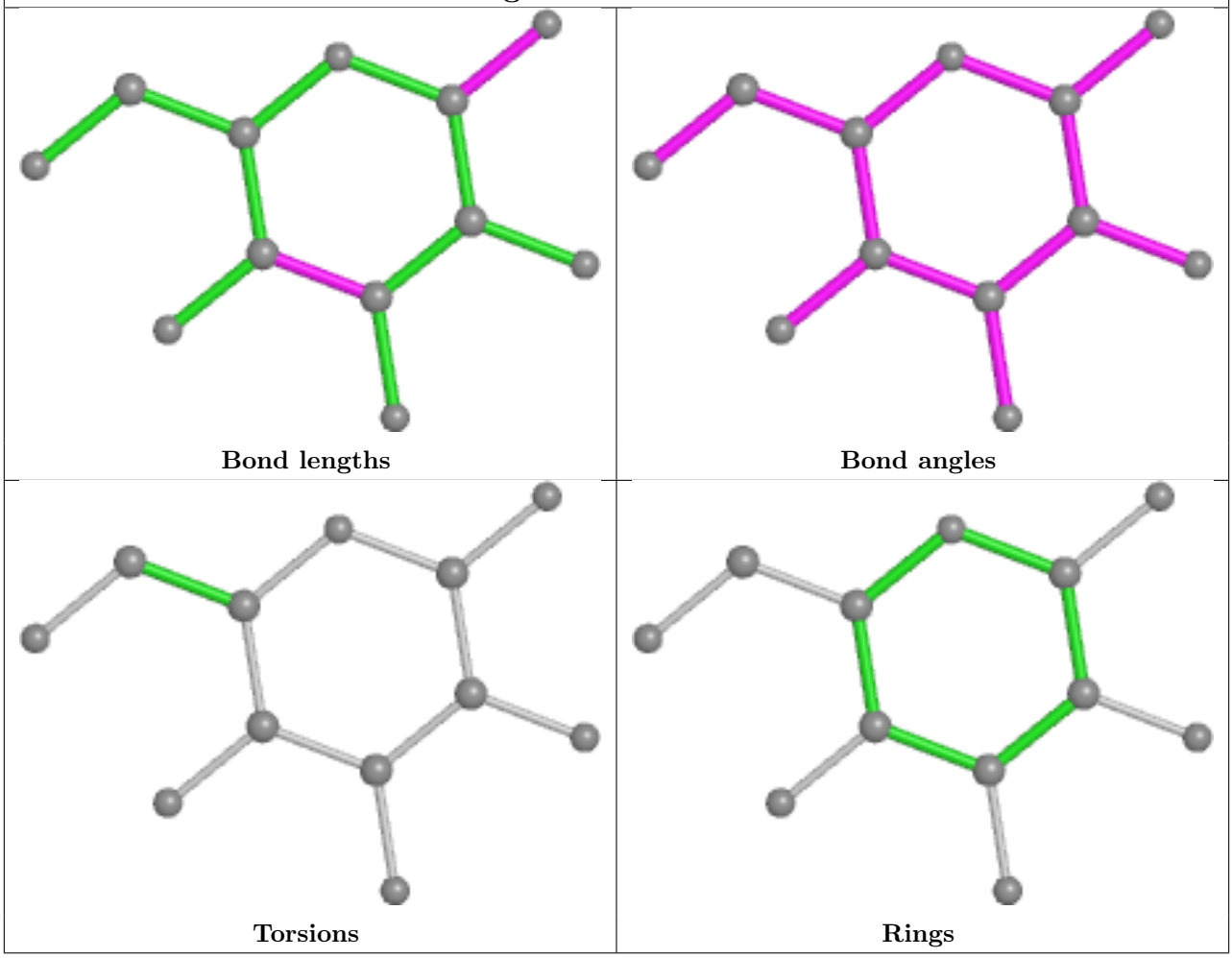


Torsions



Rings

Ligand GAL B 801



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, 95^{th} 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	718/749 (95%)	-0.20	28 (3%)	39	37	24, 42, 77, 109	0
1	B	718/749 (95%)	-0.47	11 (1%)	73	72	18, 30, 58, 102	0
All	All	1436/1498 (95%)	-0.33	39 (2%)	54	52	18, 36, 69, 109	0

All (39) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	253	ALA	7.1
1	A	739	ASN	6.4
1	A	282	ASP	4.2
1	A	658	ASP	4.1
1	B	739	ASN	4.0
1	A	588	THR	4.0
1	A	27	ASP	3.4
1	A	685	GLU	3.3
1	B	587	LYS	3.3
1	A	18	ASN	3.1
1	A	264	SER	3.0
1	A	102	ASN	2.9
1	A	740	LYS	2.9
1	A	284	PRO	2.7
1	A	51	LYS	2.7
1	A	50	ASN	2.7
1	B	5	PRO	2.7
1	A	281	LYS	2.6
1	B	737	ILE	2.6
1	A	665	ILE	2.6
1	A	252	ASP	2.6
1	A	5	PRO	2.6
1	A	741	ARG	2.6
1	A	587	LYS	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	589	ASP	2.5
1	A	612	THR	2.4
1	B	34	ALA	2.4
1	B	281	LYS	2.3
1	A	19	ARG	2.3
1	B	741	ARG	2.2
1	B	588	THR	2.2
1	A	251	ARG	2.2
1	B	284	PRO	2.1
1	B	589	ASP	2.1
1	A	17	LYS	2.1
1	B	282	ASP	2.0
1	A	321	GLU	2.0
1	A	738	GLU	2.0
1	A	733	ASP	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)

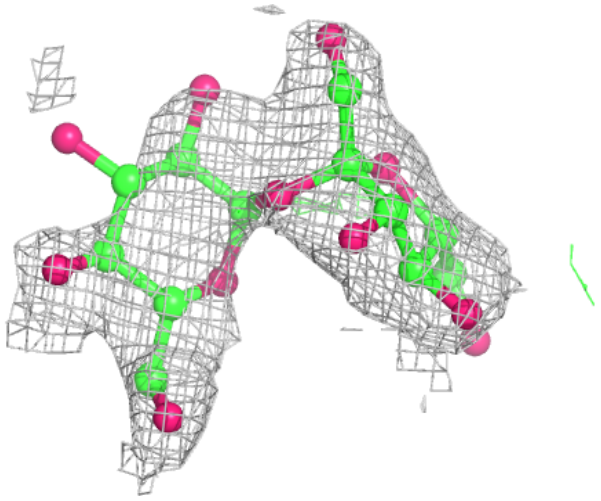
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

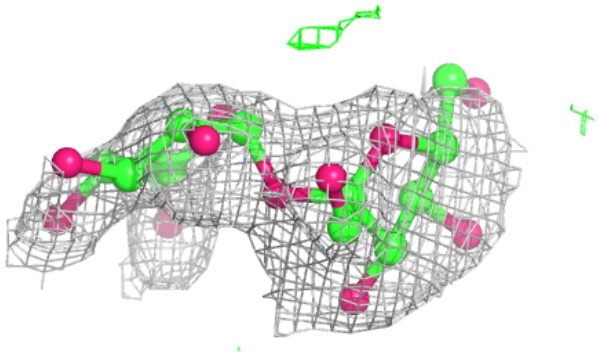
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	GLC	C	1	11/12	0.71	0.28	82,100,111,130	0
2	FRU	C	2	12/12	0.90	0.16	62,73,79,89	0
2	GLC	D	1	11/12	0.90	0.21	43,53,57,64	0
2	FRU	D	2	12/12	0.94	0.12	31,42,50,52	0

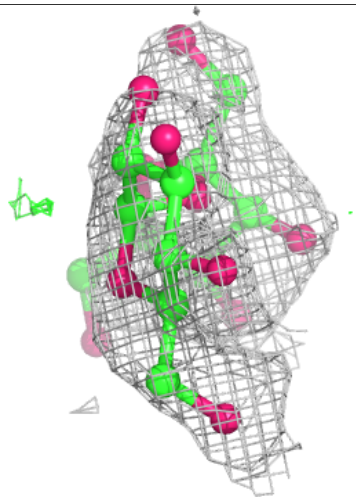
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around Chain C:

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

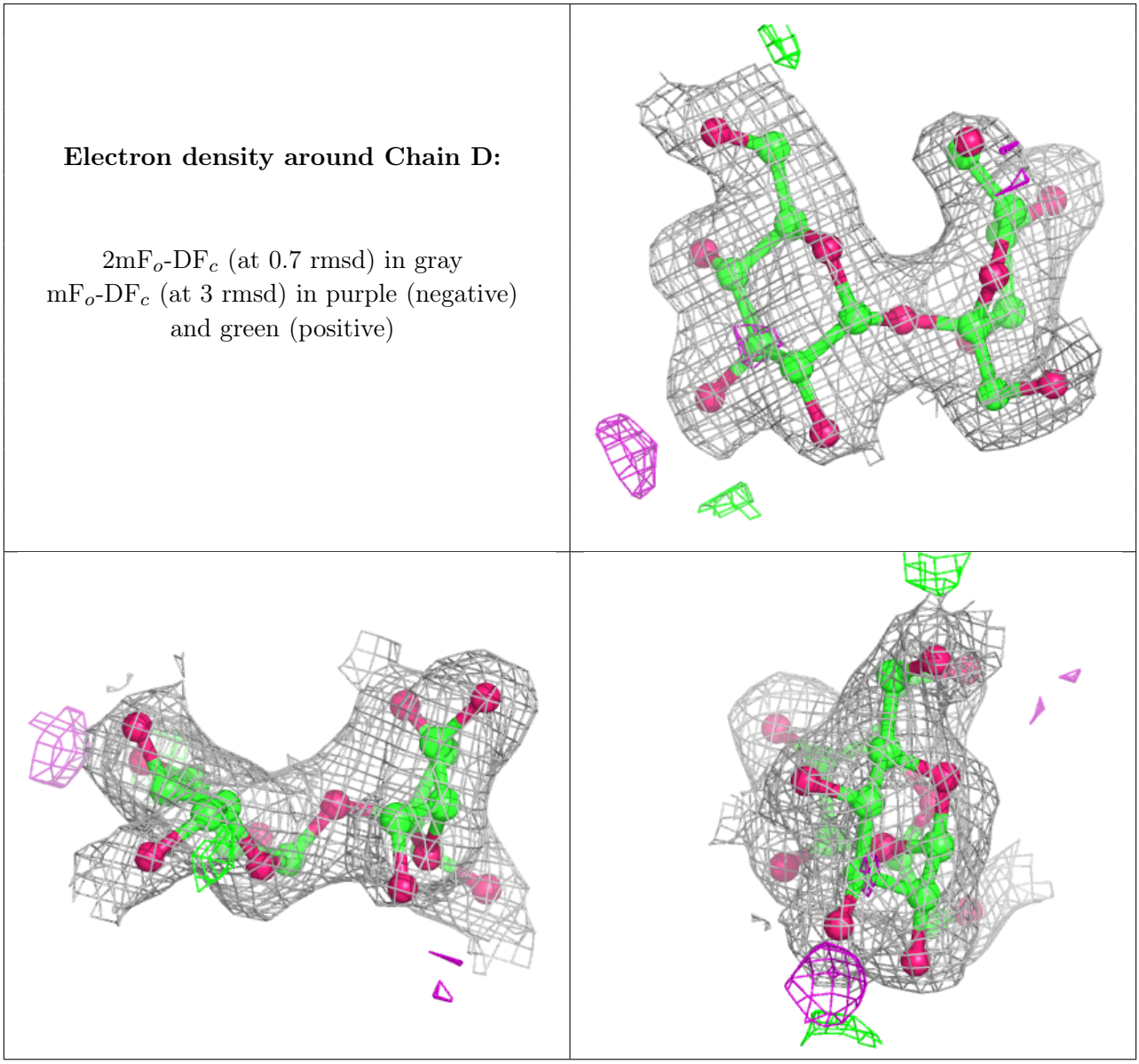






Electron density around Chain D:

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



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, 95^{th} 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.

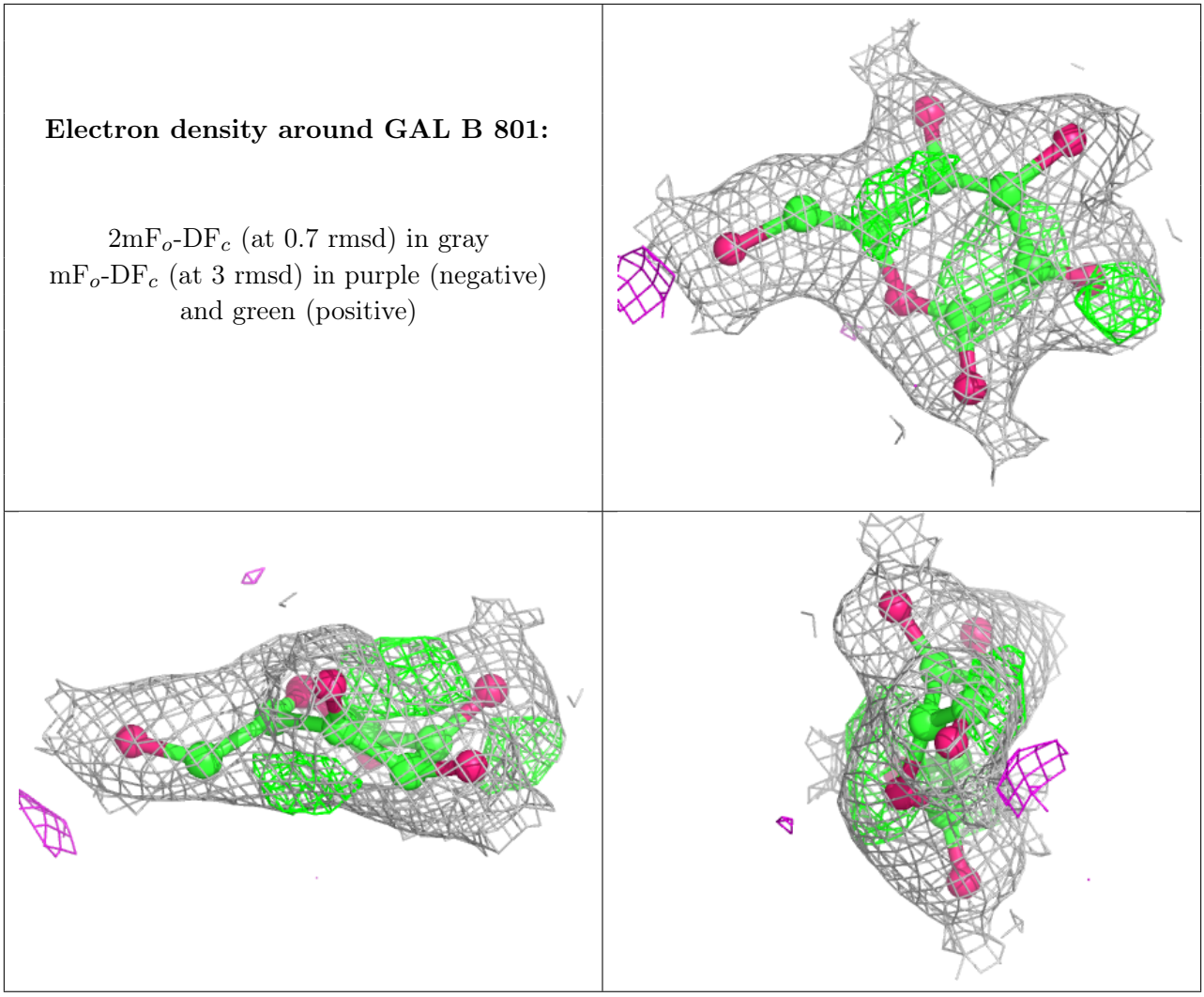
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	GAL	B	801	12/12	0.90	0.14	24,26,31,31	0
3	GAL	A	801	12/12	0.93	0.13	28,36,39,40	0

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.

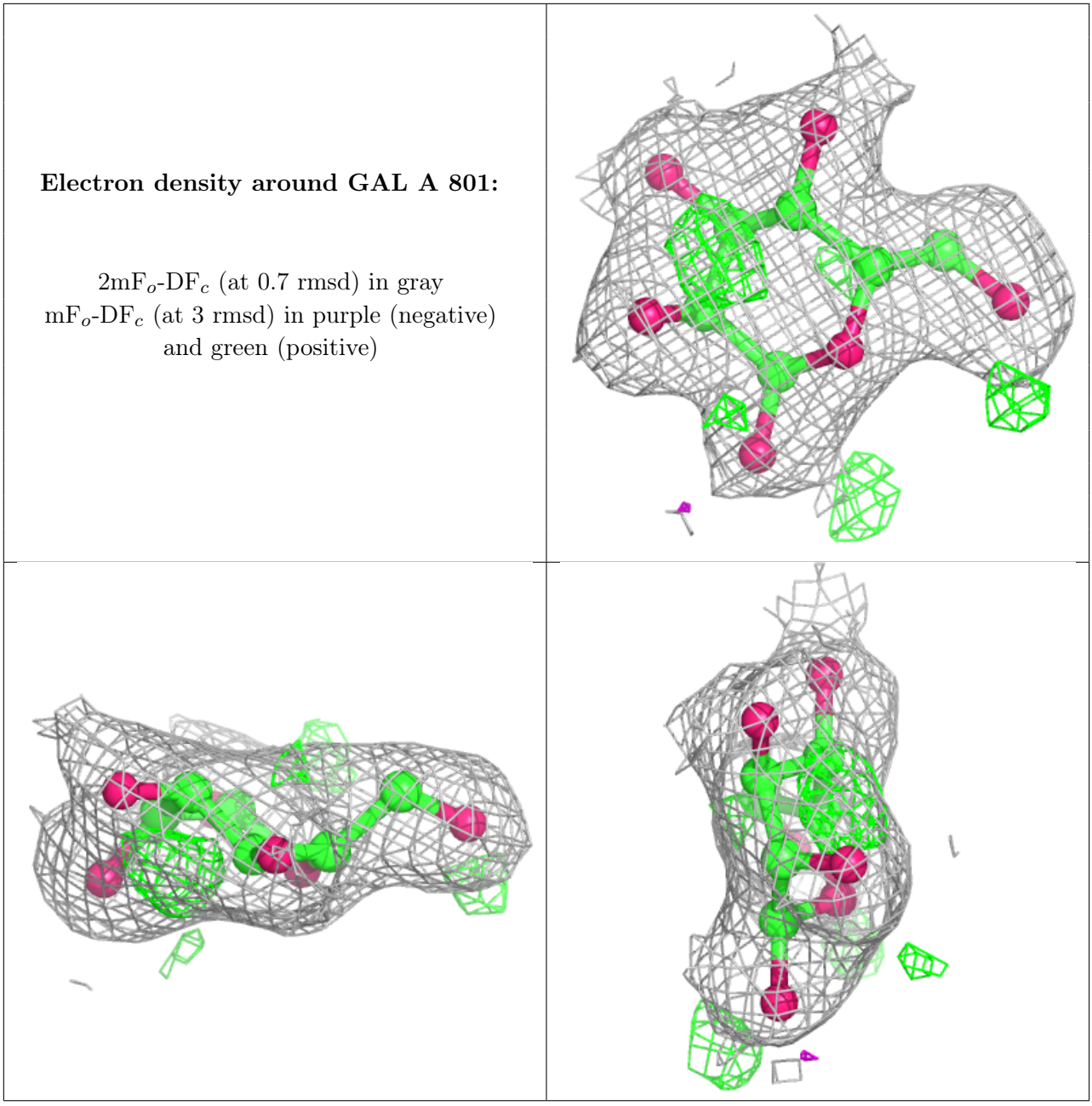
Electron density around GAL B 801:

$2mF_o$-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 801:

$2mF_o$-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

There are no such residues in this entry.