Full wwPDB X-ray Structure Validation Report ⓘ

Aug 6, 2020 – 09:46 AM BST

PDB ID	:	6SDU
Title	:	Xyloglucanase domain of NopAA, a type three effector from Sinorhizobium fredii in complex with cellobiose
Authors	:	Dorival, D.; Philys, S.; Guintini, E.; Brailly, R.; de Ruyck, J.; Czjzek, M.; Biondi, E.
Deposited on	:	2019-07-29
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 following versions of software and data (see references ⓘ) were used in the production of this report:

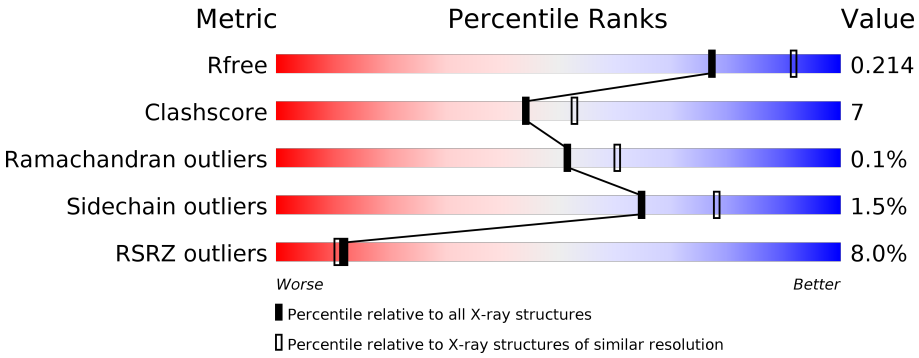
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

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 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 <=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	237	83% 10% 6%
1	B	237	25% 70% 19% • 10%
1	C	237	85% 8% 7%
1	D	237	3% 86% 7% 7%
2	E	2	100%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	BGC	E	1	-	-	-	X
2	BGC	E	2	-	-	X	X
3	BGC	A	301	-	-	X	X

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7158 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 Type III effector NopAA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	222	Total	C	N	O	S	0	0	0
			1711	1086	290	333	2			
1	B	213	Total	C	N	O	S	0	0	0
			1648	1044	280	323	1			
1	C	221	Total	C	N	O	S	0	0	0
			1707	1084	289	332	2			
1	D	221	Total	C	N	O	S	0	1	0
			1715	1089	292	332	2			

There are 68 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP M4PUR5
A	1	HIS	-	expression tag	UNP M4PUR5
A	2	HIS	-	expression tag	UNP M4PUR5
A	3	HIS	-	expression tag	UNP M4PUR5
A	4	HIS	-	expression tag	UNP M4PUR5
A	5	HIS	-	expression tag	UNP M4PUR5
A	6	HIS	-	expression tag	UNP M4PUR5
A	7	ILE	-	expression tag	UNP M4PUR5
A	8	THR	-	expression tag	UNP M4PUR5
A	9	SER	-	expression tag	UNP M4PUR5
A	10	LEU	-	expression tag	UNP M4PUR5
A	11	TYR	-	expression tag	UNP M4PUR5
A	12	LYS	-	expression tag	UNP M4PUR5
A	13	LYS	-	expression tag	UNP M4PUR5
A	14	ALA	-	expression tag	UNP M4PUR5
A	15	GLY	-	expression tag	UNP M4PUR5
A	16	MET	-	expression tag	UNP M4PUR5
B	0	MET	-	initiating methionine	UNP M4PUR5
B	1	HIS	-	expression tag	UNP M4PUR5
B	2	HIS	-	expression tag	UNP M4PUR5
B	3	HIS	-	expression tag	UNP M4PUR5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	4	HIS	-	expression tag	UNP M4PUR5
B	5	HIS	-	expression tag	UNP M4PUR5
B	6	HIS	-	expression tag	UNP M4PUR5
B	7	ILE	-	expression tag	UNP M4PUR5
B	8	THR	-	expression tag	UNP M4PUR5
B	9	SER	-	expression tag	UNP M4PUR5
B	10	LEU	-	expression tag	UNP M4PUR5
B	11	TYR	-	expression tag	UNP M4PUR5
B	12	LYS	-	expression tag	UNP M4PUR5
B	13	LYS	-	expression tag	UNP M4PUR5
B	14	ALA	-	expression tag	UNP M4PUR5
B	15	GLY	-	expression tag	UNP M4PUR5
B	16	MET	-	expression tag	UNP M4PUR5
C	0	MET	-	initiating methionine	UNP M4PUR5
C	1	HIS	-	expression tag	UNP M4PUR5
C	2	HIS	-	expression tag	UNP M4PUR5
C	3	HIS	-	expression tag	UNP M4PUR5
C	4	HIS	-	expression tag	UNP M4PUR5
C	5	HIS	-	expression tag	UNP M4PUR5
C	6	HIS	-	expression tag	UNP M4PUR5
C	7	ILE	-	expression tag	UNP M4PUR5
C	8	THR	-	expression tag	UNP M4PUR5
C	9	SER	-	expression tag	UNP M4PUR5
C	10	LEU	-	expression tag	UNP M4PUR5
C	11	TYR	-	expression tag	UNP M4PUR5
C	12	LYS	-	expression tag	UNP M4PUR5
C	13	LYS	-	expression tag	UNP M4PUR5
C	14	ALA	-	expression tag	UNP M4PUR5
C	15	GLY	-	expression tag	UNP M4PUR5
C	16	MET	-	expression tag	UNP M4PUR5
D	0	MET	-	initiating methionine	UNP M4PUR5
D	1	HIS	-	expression tag	UNP M4PUR5
D	2	HIS	-	expression tag	UNP M4PUR5
D	3	HIS	-	expression tag	UNP M4PUR5
D	4	HIS	-	expression tag	UNP M4PUR5
D	5	HIS	-	expression tag	UNP M4PUR5
D	6	HIS	-	expression tag	UNP M4PUR5
D	7	ILE	-	expression tag	UNP M4PUR5
D	8	THR	-	expression tag	UNP M4PUR5
D	9	SER	-	expression tag	UNP M4PUR5
D	10	LEU	-	expression tag	UNP M4PUR5
D	11	TYR	-	expression tag	UNP M4PUR5

Continued on next page...

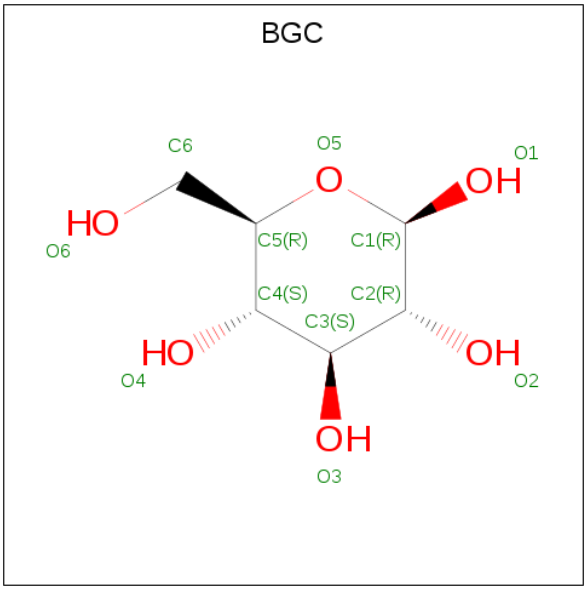
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	12	LYS	-	expression tag	UNP M4PUR5
D	13	LYS	-	expression tag	UNP M4PUR5
D	14	ALA	-	expression tag	UNP M4PUR5
D	15	GLY	-	expression tag	UNP M4PUR5
D	16	MET	-	expression tag	UNP M4PUR5

- Molecule 2 is an oligosaccharide called beta-D-glucopyranose-(1-4)-beta-D-glucopyranose.

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

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



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

- Molecule 4 is water.

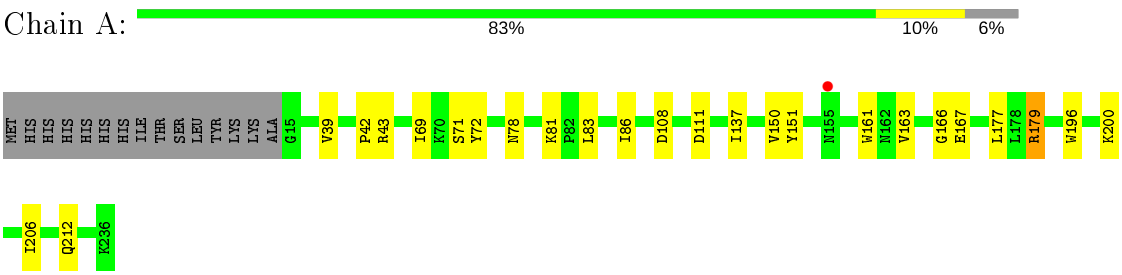
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	100	Total O 100 100	0	0
4	B	44	Total O 44 44	0	0
4	C	98	Total O 98 98	0	0
4	D	100	Total O 100 100	0	0

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.

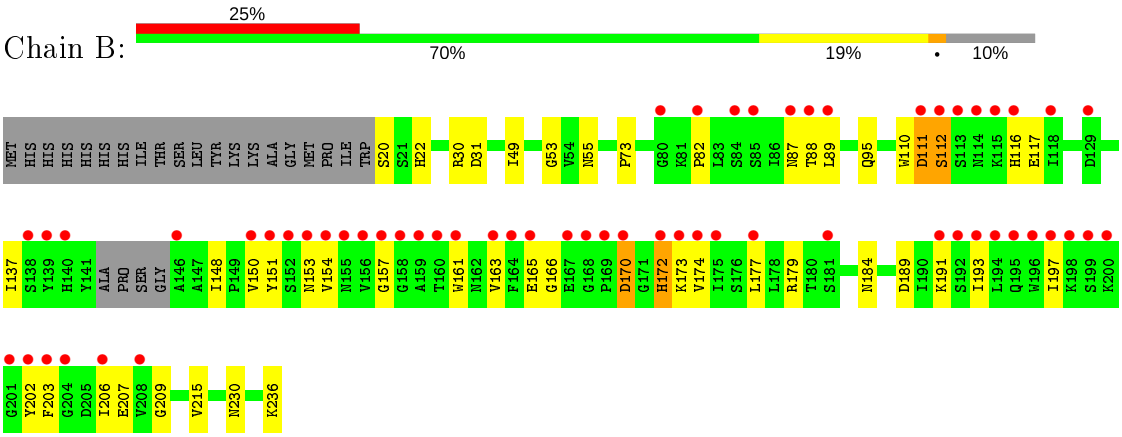
- Molecule 1: Type III effector NopAA

Chain A:



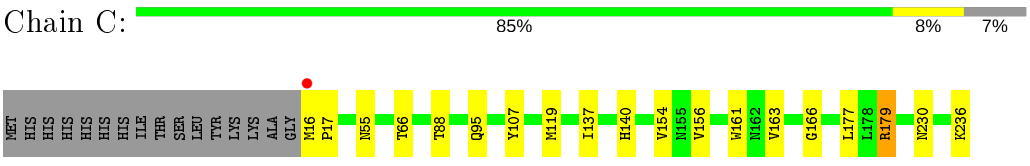
- Molecule 1: Type III effector NopAA

Chain B:



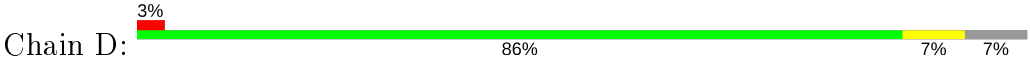
- Molecule 1: Type III effector NopAA

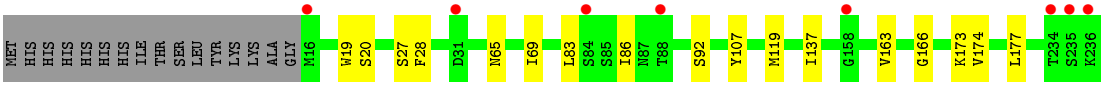
Chain C:



- Molecule 1: Type III effector NopAA

Chain D:





- Molecule 2: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain E:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	135.41Å 108.32Å 84.71Å 90.00° 113.11° 90.00°	Depositor
Resolution (Å)	44.47 – 2.20 44.47 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.9 (44.47-2.20) 99.9 (44.47-2.20)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.27 (at 2.20Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, R_{free}	0.181 , 0.214 0.183 , 0.214	Depositor DCC
R_{free} test set	2879 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	37.7	Xtriage
Anisotropy	0.277	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 64.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	7158	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.37% 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: BGC

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	0.51	0/1769	0.61	0/2418
1	B	0.50	0/1701	0.73	3/2322 (0.1%)
1	C	0.56	0/1765	0.64	0/2413
1	D	0.56	2/1776 (0.1%)	0.64	0/2428
All	All	0.53	2/7011 (0.0%)	0.65	3/9581 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	92	SER	CB-OG	8.03	1.52	1.42
1	D	27	SER	CB-OG	5.03	1.48	1.42

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	170	ASP	CB-CA-C	6.41	123.23	110.40
1	B	154	VAL	CG1-CB-CG2	6.36	121.07	110.90
1	B	31	ASP	CB-CA-C	5.80	122.01	110.40

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	A	1711	0	1600	22	0
1	B	1648	0	1539	41	0
1	C	1707	0	1597	11	0
1	D	1715	0	1610	14	0
2	E	23	0	20	7	0
3	A	12	0	11	8	0
4	A	100	0	0	1	0
4	B	44	0	0	2	0
4	C	98	0	0	0	0
4	D	100	0	0	2	0
All	All	7158	0	6377	87	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 7.

All (87) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:137:ILE:HD11	2:E:2:BGC:H4	1.31	1.11
1:B:165:GLU:HG3	1:B:173:LYS:NZ	1.73	1.03
1:B:165:GLU:CG	1:B:173:LYS:HZ1	1.77	0.97
1:B:165:GLU:CG	1:B:173:LYS:NZ	2.31	0.94
1:A:137:ILE:HD11	3:A:301:BGC:H1	1.51	0.90
1:B:165:GLU:HG3	1:B:173:LYS:HZ2	1.38	0.88
1:A:69:ILE:HD11	3:A:301:BGC:H4	1.66	0.78
1:B:148:ILE:HD12	1:B:148:ILE:O	1.85	0.77
1:D:69:ILE:HD12	2:E:2:BGC:H5	1.66	0.76
1:B:111:ASP:HB3	1:B:116:HIS:HB2	1.70	0.71
1:C:88:THR:HB	1:C:236:LYS:HG3	1.73	0.70
1:A:161:TRP:CZ3	1:A:179:ARG:HG3	2.28	0.69
1:D:69:ILE:HD11	2:E:2:BGC:H3	1.77	0.67
1:B:88:THR:OG1	1:B:236:LYS:HB2	1.97	0.64
1:B:165:GLU:CD	1:B:173:LYS:HZ1	2.00	0.64
1:B:20:SER:HA	1:B:49:ILE:O	1.97	0.64
1:B:89:LEU:N	4:B:301:HOH:O	2.16	0.62
1:A:137:ILE:HG13	3:A:301:BGC:H5	1.81	0.62
1:B:161:TRP:CZ3	1:B:179:ARG:HG2	2.34	0.62
1:B:30:ARG:NH1	1:B:53:GLY:HA2	2.15	0.62
3:A:301:BGC:H6C1	4:A:490:HOH:O	2.00	0.61
1:B:82:PRO:HA	1:B:207:GLU:HA	1.82	0.61
1:D:137:ILE:HD13	1:D:174:VAL:HG11	1.82	0.61
1:B:150:VAL:HG12	1:B:151:TYR:CE2	2.36	0.60

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:137:ILE:HG13	3:A:301:BGC:H3	1.82	0.60
1:A:111:ASP:HB3	1:A:206:ILE:HD11	1.85	0.59
1:D:69:ILE:CD1	2:E:2:BGC:H3	2.35	0.57
1:C:154:VAL:HG12	1:C:156:VAL:HG23	1.87	0.56
1:D:137:ILE:HG22	1:D:166:GLY:HA3	1.88	0.55
1:B:111:ASP:O	1:B:207:GLU:O	2.25	0.55
1:B:88:THR:HA	4:B:301:HOH:O	2.07	0.55
1:A:83:LEU:HA	1:A:86:ILE:HD12	1.88	0.54
1:A:69:ILE:CD1	3:A:301:BGC:H4	2.34	0.54
1:B:150:VAL:HG21	1:B:165:GLU:HB3	1.88	0.54
1:B:137:ILE:HG22	1:B:166:GLY:HA3	1.90	0.54
1:B:203:PHE:CD2	1:B:206:ILE:HG12	2.45	0.51
1:B:157:GLY:HA2	1:B:189:ASP:HB3	1.91	0.51
1:B:87:ASN:O	1:B:191:LYS:NZ	2.26	0.51
1:B:165:GLU:OE1	1:B:173:LYS:NZ	2.43	0.50
1:C:163:VAL:HG22	1:C:177:LEU:HG	1.93	0.50
1:A:137:ILE:HG22	1:A:166:GLY:HA3	1.93	0.50
1:B:172:HIS:C	1:B:172:HIS:CD2	2.85	0.49
1:B:150:VAL:HG12	1:B:150:VAL:O	2.12	0.49
1:C:16:MET:HB3	1:C:17:PRO:HD2	1.95	0.49
1:A:43:ARG:HH21	1:B:22:HIS:H	1.61	0.48
1:D:19:TRP:CE2	1:D:28:PHE:HB2	2.48	0.48
1:A:163:VAL:HG22	1:A:177:LEU:HG	1.96	0.48
1:C:137:ILE:HG22	1:C:166:GLY:HA3	1.96	0.48
1:A:108:ASP:HB3	1:A:212:GLN:HB2	1.95	0.47
1:B:112:SER:HB2	1:B:209:GLY:CA	2.44	0.47
1:D:137:ILE:CD1	2:E:2:BGC:H4	2.21	0.47
1:B:110:TRP:CZ3	1:B:117:GLU:HB2	2.49	0.47
1:C:161:TRP:CZ3	1:C:179:ARG:HG3	2.49	0.47
1:B:193:ILE:O	1:B:197:ILE:HG13	2.14	0.47
1:C:16:MET:CB	1:C:17:PRO:HD2	2.44	0.47
1:B:117:GLU:HB3	1:B:174:VAL:HG22	1.97	0.46
1:B:150:VAL:HG12	1:B:151:TYR:CD2	2.51	0.46
1:B:150:VAL:O	1:B:151:TYR:CG	2.69	0.46
1:A:137:ILE:CG1	3:A:301:BGC:H5	2.45	0.46
1:B:165:GLU:HG2	1:B:173:LYS:HZ1	1.75	0.46
1:B:150:VAL:HG11	1:B:202:TYR:OH	2.17	0.45
1:D:137:ILE:HG13	2:E:2:BGC:H2	1.99	0.45
1:A:167:GLU:OE2	1:C:140:HIS:HA	2.17	0.44
1:A:42:PRO:HB3	1:C:66:THR:HB	1.99	0.44
1:B:73:PRO:HD2	1:B:215:VAL:O	2.18	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:55:ASN:O	1:C:230:ASN:HA	2.19	0.43
1:A:39:VAL:HG13	1:A:71:SER:HB3	2.00	0.43
1:D:83:LEU:HA	1:D:86:ILE:HD12	1.99	0.43
1:A:196:TRP:CZ2	1:A:200:LYS:HE3	2.53	0.42
1:A:69:ILE:CD1	3:A:301:BGC:H2	2.49	0.42
1:D:19:TRP:CD2	1:D:28:PHE:HB2	2.54	0.42
1:A:72:TYR:OH	1:A:108:ASP:HB2	2.20	0.42
1:A:81:LYS:HE3	1:A:81:LYS:HB3	1.79	0.42
1:D:107:TYR:O	1:D:119:MET:HA	2.19	0.42
1:D:137:ILE:HD12	4:D:412:HOH:O	2.19	0.42
1:B:112:SER:HB2	1:B:209:GLY:HA3	2.01	0.42
1:B:95:GLN:NE2	1:B:184:ASN:HA	2.35	0.41
4:D:404:HOH:O	2:E:1:BGC:H6C2	2.19	0.41
1:C:107:TYR:O	1:C:119:MET:HA	2.21	0.41
1:D:163:VAL:HG22	1:D:177:LEU:HG	2.02	0.41
1:B:172:HIS:O	1:B:172:HIS:CD2	2.73	0.41
1:B:55:ASN:ND2	1:B:230:ASN:HD22	2.19	0.41
1:B:177:LEU:HA	1:B:177:LEU:HD23	1.67	0.41
1:A:150:VAL:HG23	1:A:151:TYR:CD2	2.55	0.40
1:A:43:ARG:HA	1:A:43:ARG:HD3	1.64	0.40
1:B:150:VAL:CG1	1:B:150:VAL:O	2.70	0.40
1:B:163:VAL:HG22	1:B:177:LEU:HD21	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	220/237 (93%)	216 (98%)	4 (2%)	0	100 100
1	B	209/237 (88%)	206 (99%)	2 (1%)	1 (0%)	29 31
1	C	219/237 (92%)	218 (100%)	1 (0%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	D	220/237 (93%)	217 (99%)	3 (1%)	0	100 100
All	All	868/948 (92%)	857 (99%)	10 (1%)	1 (0%)	51 60

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	112	SER

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	188/202 (93%)	186 (99%)	2 (1%)	73 85
1	B	182/202 (90%)	178 (98%)	4 (2%)	52 65
1	C	188/202 (93%)	186 (99%)	2 (1%)	73 85
1	D	189/202 (94%)	186 (98%)	3 (2%)	62 76
All	All	747/808 (92%)	736 (98%)	11 (2%)	65 78

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	78	ASN
1	A	179	ARG
1	B	111	ASP
1	B	153	ASN
1	B	170	ASP
1	B	172	HIS
1	C	95	GLN
1	C	179	ARG
1	D	20	SER
1	D	65	ASN
1	D	173	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such

sidechains are listed below:

Mol	Chain	Res	Type
1	B	55	ASN
1	B	172	HIS
1	B	230	ASN
1	C	225	ASN
1	C	227	ASN
1	D	140	HIS

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

2 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	BGC	E	1	2	12,12,12	1.38	1 (8%)	17,17,17	1.06	0
2	BGC	E	2	2	11,11,12	1.66	2 (18%)	15,15,17	1.77	4 (26%)

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	BGC	E	1	2	-	2/2/22/22	0/1/1/1
2	BGC	E	2	2	-	0/2/19/22	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	2	BGC	O5-C1	4.27	1.50	1.43
2	E	1	BGC	O5-C1	3.48	1.51	1.42
2	E	2	BGC	O5-C5	2.67	1.48	1.43

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	2	BGC	C1-C2-C3	3.92	114.48	109.67
2	E	2	BGC	O3-C3-C2	2.38	114.55	109.99
2	E	2	BGC	O2-C2-C1	-2.12	104.81	109.15
2	E	2	BGC	O4-C4-C5	2.12	114.55	109.30

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	E	1	BGC	O5-C5-C6-O6
2	E	1	BGC	C4-C5-C6-O6

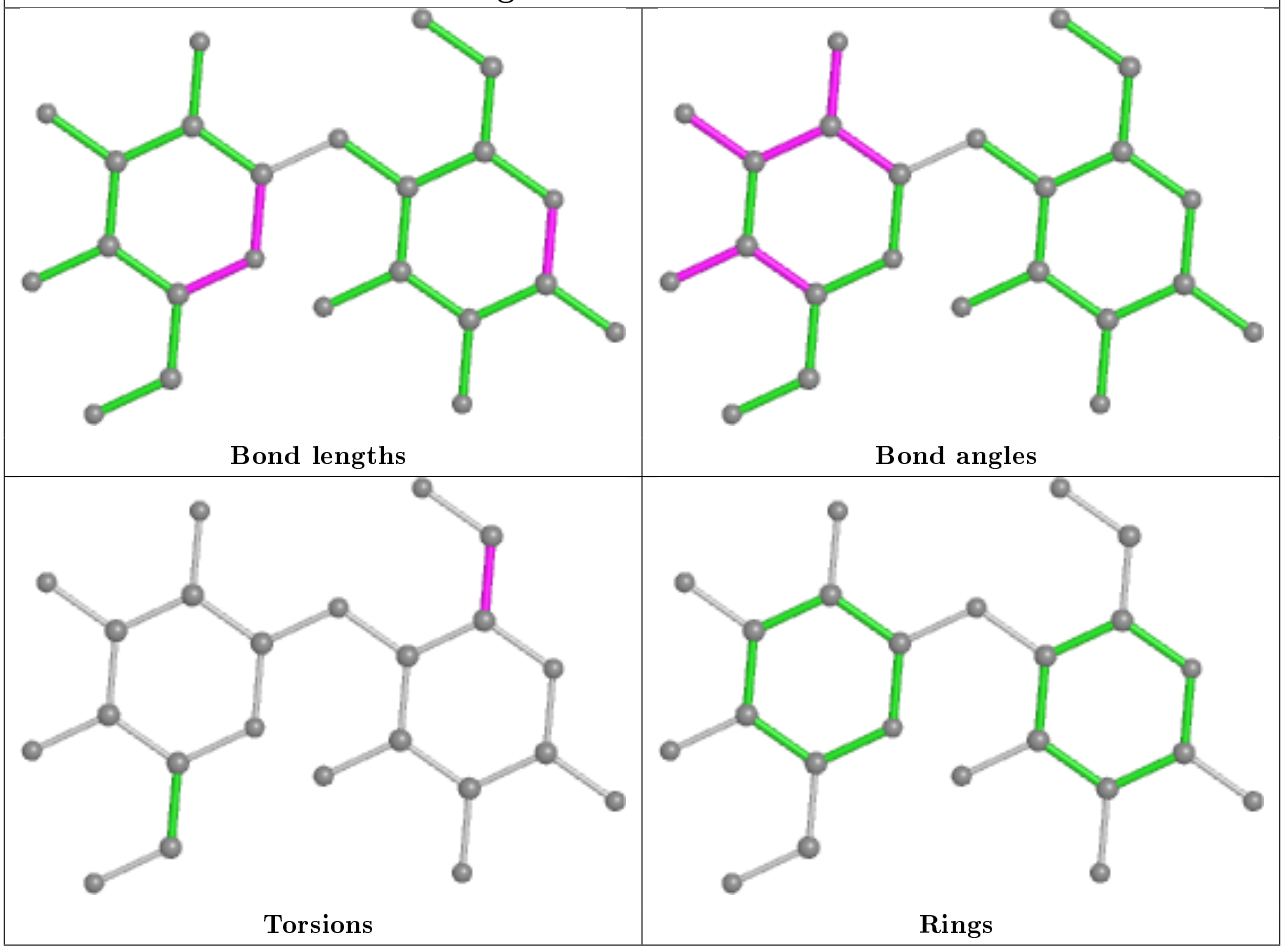
There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	2	BGC	6	0
2	E	1	BGC	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 E



5.6 Ligand geometry (i)

1 ligand is 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	BGC	A	301	-	12,12,12	1.52	2 (16%)	17,17,17	1.42	3 (17%)

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	BGC	A	301	-	-	2/2/22/22	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	301	BGC	O5-C1	4.05	1.53	1.42
3	A	301	BGC	O5-C5	2.06	1.49	1.44

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	301	BGC	O2-C2-C1	2.61	115.22	109.16
3	A	301	BGC	O5-C5-C6	2.49	112.62	106.44
3	A	301	BGC	O4-C4-C5	2.37	115.17	109.30

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	301	BGC	C4-C5-C6-O6
3	A	301	BGC	O5-C5-C6-O6

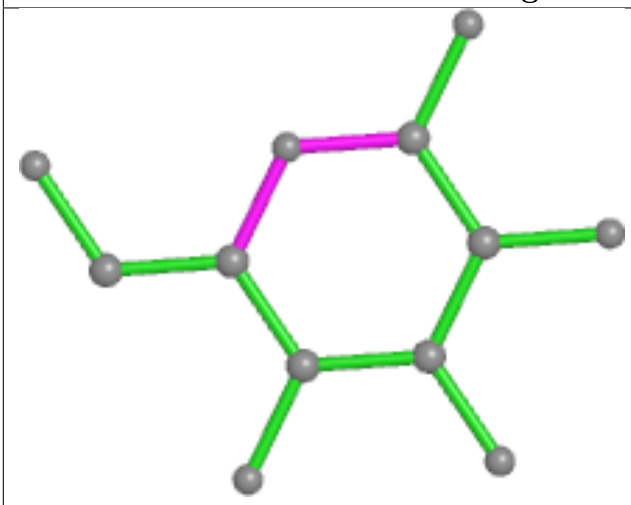
There are no ring outliers.

1 monomer is involved in 8 short contacts:

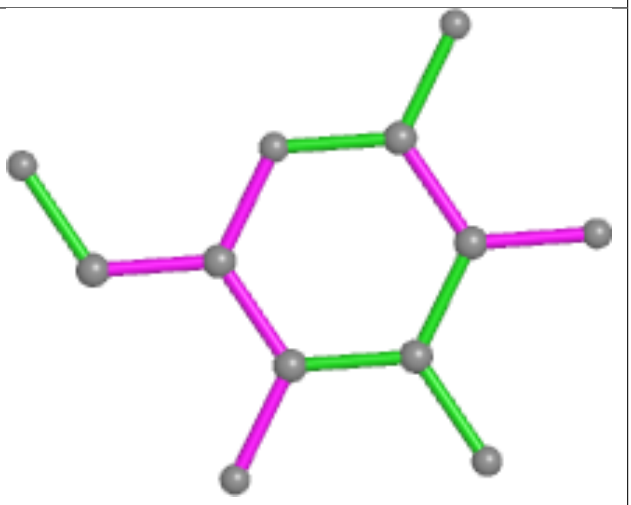
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	301	BGC	8	0

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.

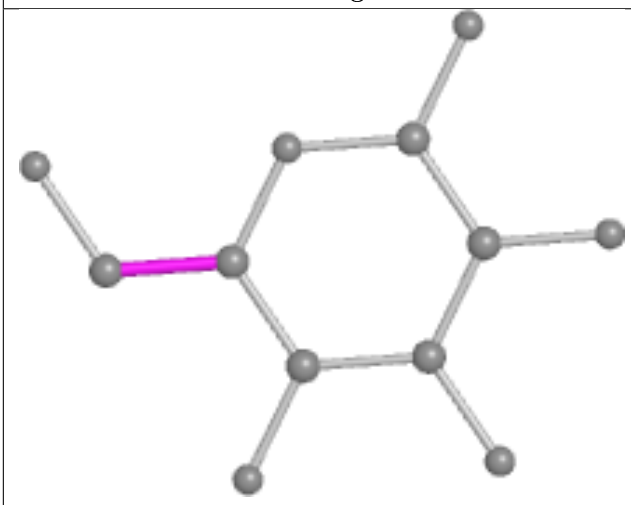
Ligand BGC A 301



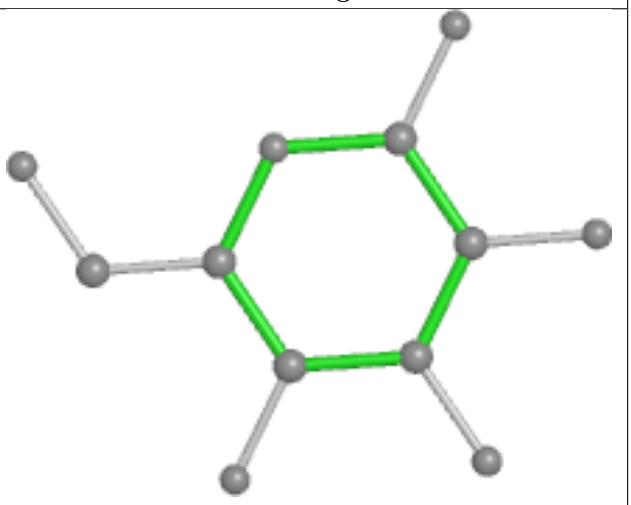
Bond lengths



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, 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	222/237 (93%)	-0.05	1 (0%) 91 90	25, 36, 55, 92	0
1	B	213/237 (89%)	1.13	60 (28%) 0 0	28, 63, 104, 121	0
1	C	221/237 (93%)	-0.07	1 (0%) 91 90	26, 37, 56, 68	0
1	D	221/237 (93%)	0.07	8 (3%) 42 41	23, 41, 65, 83	0
All	All	877/948 (92%)	0.26	70 (7%) 12 11	23, 41, 91, 121	0

All (70) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	203	PHE	7.2
1	B	197	ILE	6.7
1	B	170	ASP	6.6
1	D	16	MET	5.5
1	B	196	TRP	5.3
1	B	150	VAL	5.0
1	C	16	MET	4.9
1	B	151	TYR	4.7
1	B	172	HIS	4.6
1	B	168	GLY	4.6
1	B	200	LYS	4.5
1	B	156	VAL	4.5
1	B	161	TRP	4.4
1	B	160	THR	4.4
1	B	89	LEU	4.3
1	D	236	LYS	4.2
1	B	167	GLU	4.2
1	B	169	PRO	4.2
1	B	140	HIS	4.2
1	B	174	VAL	4.2
1	B	195	GLN	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	202	TYR	4.0
1	B	82	PRO	3.9
1	B	198	LYS	3.9
1	B	163	VAL	3.8
1	B	111	ASP	3.7
1	B	84	SER	3.6
1	B	154	VAL	3.4
1	B	159	ALA	3.4
1	B	85	SER	3.3
1	B	165	GLU	3.3
1	B	146	ALA	3.2
1	B	201	GLY	3.2
1	B	138	SER	3.2
1	B	199	SER	3.2
1	B	194	LEU	3.0
1	B	152	SER	2.9
1	D	234	THR	2.8
1	B	204	GLY	2.8
1	A	155	ASN	2.8
1	B	181	SER	2.8
1	D	235	SER	2.8
1	B	206	ILE	2.8
1	B	157	GLY	2.7
1	B	153	ASN	2.7
1	B	191	LYS	2.7
1	B	193	ILE	2.6
1	B	88	THR	2.6
1	B	113	SER	2.6
1	B	175	ILE	2.6
1	B	158	GLY	2.6
1	D	88	THR	2.6
1	B	139	TYR	2.6
1	B	112	SER	2.6
1	B	155	ASN	2.5
1	B	173	LYS	2.5
1	B	177	LEU	2.5
1	D	84	SER	2.4
1	B	115	LYS	2.4
1	B	116	HIS	2.3
1	B	114	ASN	2.3
1	B	87	ASN	2.3
1	B	164	PHE	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	208	VAL	2.2
1	D	158	GLY	2.2
1	B	129	ASP	2.2
1	B	118	ILE	2.2
1	B	192	SER	2.2
1	B	80	GLY	2.1
1	D	31	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)

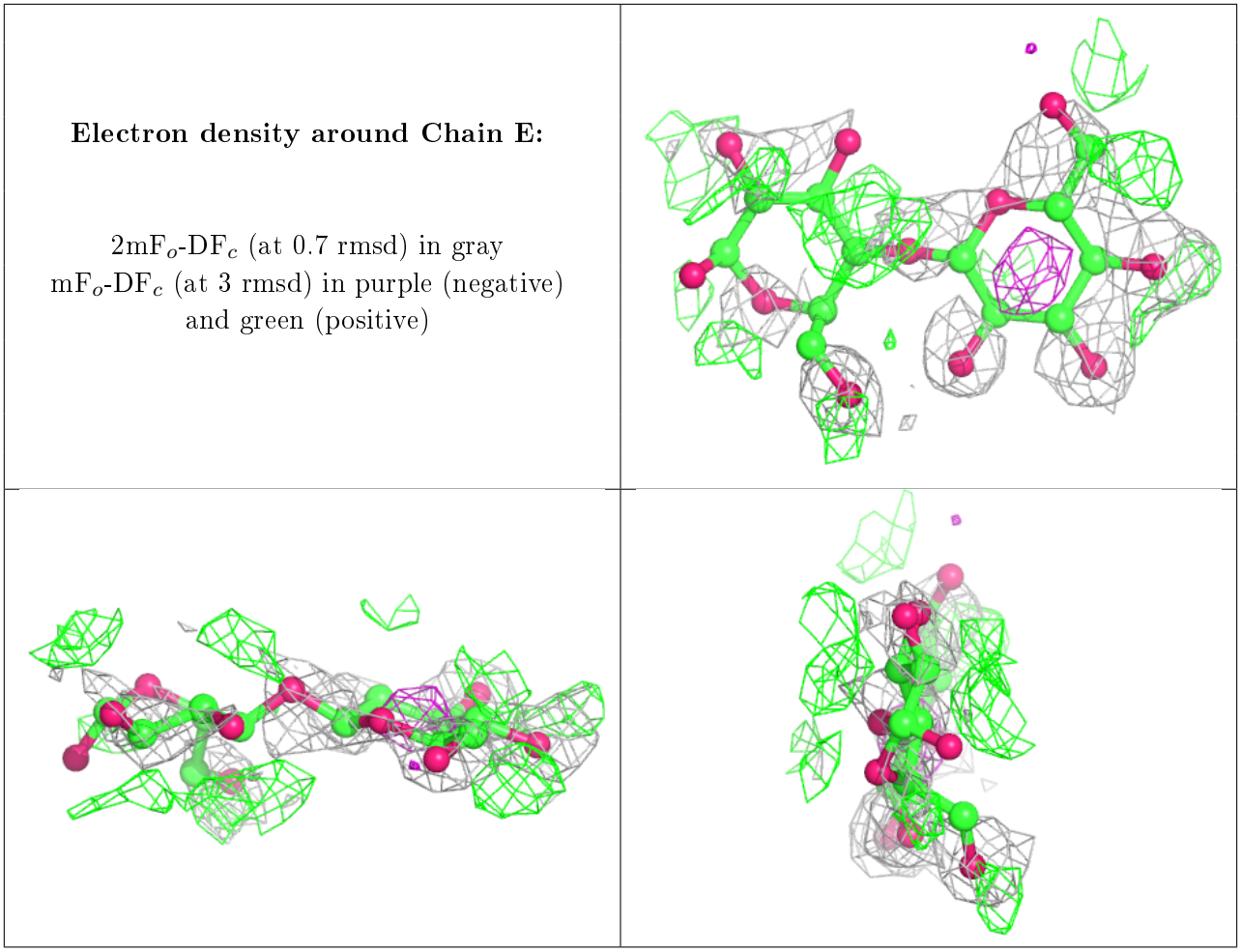
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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	BGC	E	1	12/12	0.59	0.46	24,43,47,49	12
2	BGC	E	2	11/12	0.77	0.45	4,27,36,37	11

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 E:

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



6.4 Ligands

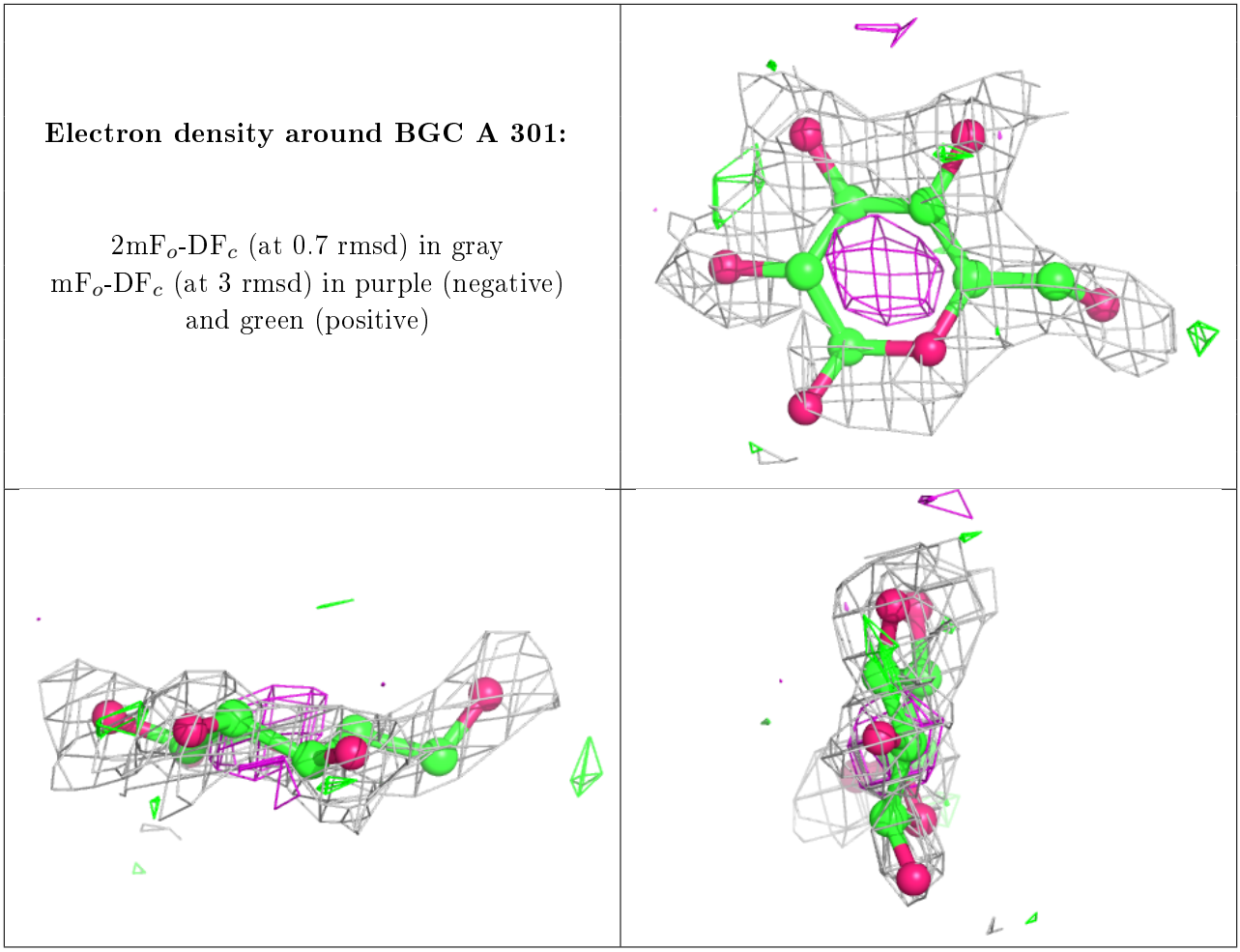
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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	BGC	A	301	12/12	0.76	0.53	7,20,30,35	12

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 BGC A 301:

$2mF_o$-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.