



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 17, 2020 – 06:39 AM EST

PDB ID : 6VBG  
Title : Lactose permease complex with thiodigalactoside and nanobody 9043  
Authors : Kumar, H.; Stroud, R.M.; Kaback, H.R.; Finer-Moore, J.; Smirnova, I.; Kasho, V.; Pardon, E.; Steyart, J.  
Deposited on : 2019-12-18  
Resolution : 2.80 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](mailto: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 : 2.14.6  
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) : 2.14.6

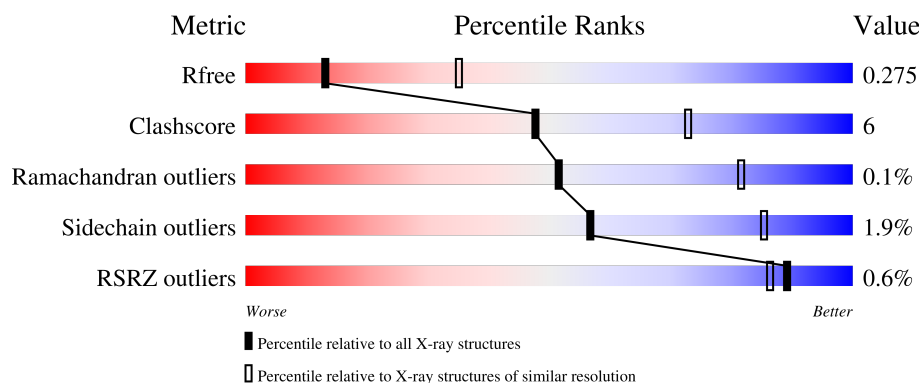
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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  $\geq 3$ , 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	417	
1	B	417	
2	C	121	
2	D	121	
3	E	2	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	2	 50%50%

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 16923 atoms, of which 8498 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 Galactoside permease.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	396	Total	C	H	N	O	S	0	0	0
			6365	2155	3201	478	509	22			
1	B	397	Total	C	H	N	O	S	0	0	0
			6379	2159	3207	479	512	22			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	46	TRP	GLY	engineered mutation	UNP C6FW78
A	262	TRP	GLY	engineered mutation	UNP C6FW78
B	46	TRP	GLY	engineered mutation	UNP C6FW78
B	262	TRP	GLY	engineered mutation	UNP C6FW78

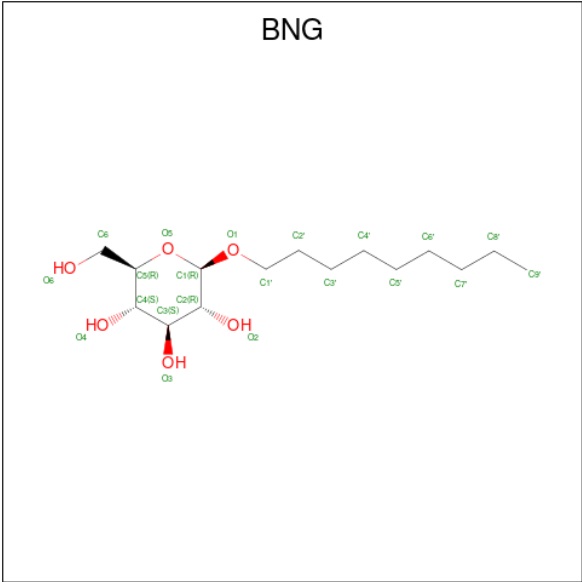
- Molecule 2 is a protein called nanobody 9043.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	121	Total	C	H	N	O	S	0	0	0
			1848	587	902	175	180	4			
2	D	117	Total	C	H	N	O	S	0	0	0
			1780	563	874	163	176	4			

- Molecule 3 is an oligosaccharide called beta-D-galactopyranose-(1-1)-1-thio-beta-D-galactopyranose.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	2	Total	C	H	O	S	0	0	0
			45	12	22	10	1			
3	F	2	Total	C	H	O	S	0	0	0
			45	12	22	10	1			

- Molecule 4 is nonyl beta-D-glucopyranoside (three-letter code: BNG) (formula: C<sub>15</sub>H<sub>30</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	H	O	0	0
			51	15	30	6		
4	A	1	Total	C	H	O	0	0
			51	15	30	6		
4	C	1	Total	C	H	O	0	0
			51	15	30	6		
4	C	1	Total	C	H	O	0	0
			51	15	30	6		
4	B	1	Total	C	H	O	0	0
			51	15	30	6		
4	B	1	Total	C	H	O	0	0
			51	15	30	6		
4	B	1	Total	C	H	O	0	0
			51	15	30	6		
4	B	1	Total	C	H	O	0	0
			51	15	30	6		
4	D	1	Total	C	H	O	0	0
			51	15	30	6		


- Molecule 5 is water.

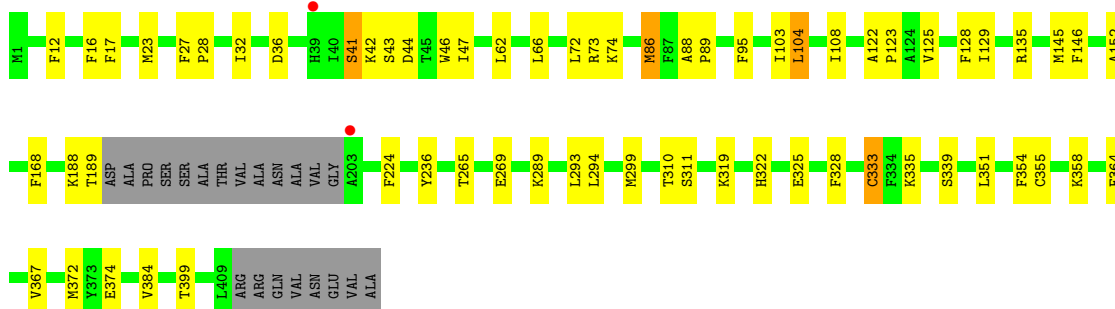
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total O	0	0
			1 1		
5	B	1	Total O	0	0
			1 1		

### 3 Residue-property plots


These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

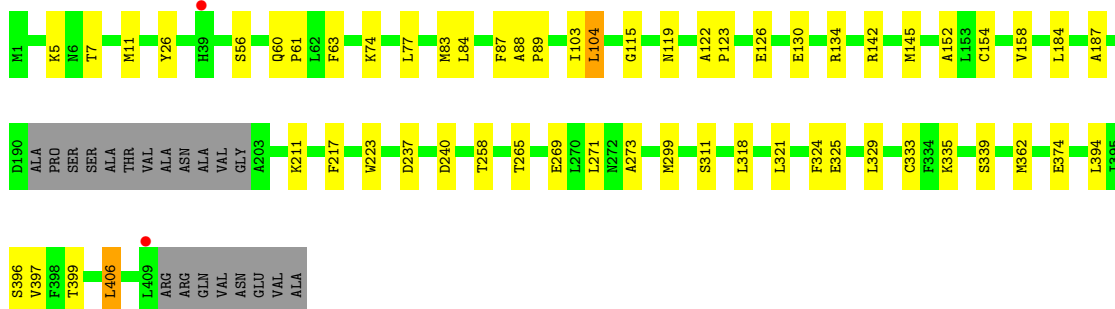
- Molecule 1: Galactoside permease

Chain A: 




- Molecule 1: Galactoside permease

Chain B: 




- Molecule 2: nanobody 9043

Chain C: 

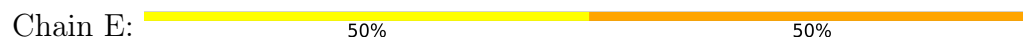


- Molecule 2: nanobody 9043

Chain D: 



- Molecule 3: beta-D-galactopyranose-(1-1)-1-thio-beta-D-galactopyranose



- Molecule 3: beta-D-galactopyranose-(1-1)-1-thio-beta-D-galactopyranose



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	151.27Å 151.27Å 182.48Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	39.06 – 2.80 39.06 – 2.80	Depositor EDS
% Data completeness (in resolution range)	96.9 (39.06-2.80) 81.8 (39.06-2.80)	Depositor EDS
$R_{merge}$	0.22	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.42 (at 2.81Å)	Xtriage
Refinement program	PHENIX 1.15.2_3472, PHENIX 1.15.2_3472	Depositor
R, $R_{free}$	0.239 , 0.275 0.238 , 0.275	Depositor DCC
$R_{free}$ test set	2002 reflections (3.50%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	49.7	Xtriage
Anisotropy	0.192	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 28.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.056 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	16923	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.70% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: YIO, GAL, BNG

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	0.28	0/3263	0.42	0/4421
1	B	0.28	0/3271	0.43	0/4432
2	C	0.30	0/968	0.52	0/1310
2	D	0.28	0/924	0.49	0/1250
All	All	0.28	0/8426	0.45	0/11413

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [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	3164	3201	3204	43	1
1	B	3172	3207	3208	34	0
2	C	946	902	902	21	1
2	D	906	874	874	9	0
3	E	23	22	21	1	0
3	F	23	22	21	1	0
4	A	42	60	60	2	0
4	B	84	120	120	1	0
4	C	42	60	60	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	21	30	30	0	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
All	All	8425	8498	8500	103	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

All (103) 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:C:27:ARG:NH2	4:C:502:BNG:O2	2.02	0.92
1:A:355:CYS:SG	4:A:701:BNG:O3	2.38	0.79
1:B:374:GLU:OE2	2:D:27:ARG:NH1	2.16	0.78
2:D:67:ARG:NH2	2:D:90:ASP:OD2	2.27	0.68
3:E:2:GAL:HO4	3:E:2:GAL:HO6	1.45	0.65
2:C:67:ARG:NH1	2:C:90:ASP:OD2	2.33	0.61
2:D:52:ASN:ND2	2:D:56:THR:OG1	2.35	0.60
1:A:72:LEU:O	1:A:189:THR:HG21	2.02	0.59
2:D:52:ASN:ND2	2:D:54:SER:OG	2.35	0.59
2:C:52:ASN:ND2	2:C:56:THR:OG1	2.35	0.58
1:B:217:PHE:O	1:B:223:TRP:NE1	2.36	0.57
1:B:299:MET:HG3	1:B:325:GLU:HG3	1.85	0.57
1:B:406:LEU:HD23	1:B:406:LEU:H	1.69	0.57
1:A:73:ARG:O	1:A:74:LYS:HG2	2.05	0.57
1:A:88:ALA:HB3	1:A:89:PRO:HD3	1.85	0.56
1:A:299:MET:HG3	1:A:325:GLU:HG3	1.88	0.56
2:C:91:THR:HG23	2:C:114:THR:HA	1.88	0.56
1:B:126:GLU:OE2	3:F:1:YIO:O2	2.24	0.56
1:B:394:LEU:HA	1:B:397:VAL:HG23	1.88	0.55
1:A:41:SER:OG	1:A:43:SER:HB2	2.08	0.54
1:B:88:ALA:HB3	1:B:89:PRO:HD3	1.88	0.54
2:C:100:VAL:HG21	2:C:105:LYS:HD3	1.88	0.53
1:B:152:ALA:HB2	1:B:269:GLU:HB2	1.90	0.53
2:C:118:HIS:O	2:C:119:HIS:HB2	2.09	0.52
1:A:41:SER:OG	1:A:43:SER:N	2.43	0.52
1:A:86:MET:HE1	1:B:83:MET:SD	2.49	0.52
2:C:32:TYR:HH	4:C:502:BNG:HO3	1.52	0.52
1:A:32:ILE:O	1:A:36:ASP:HB2	2.11	0.51
1:B:60:GLN:N	1:B:61:PRO:HD2	2.25	0.51
2:D:52:ASN:OD1	2:D:55:GLY:N	2.43	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:396:SER:HA	1:B:399:THR:OG1	2.10	0.51
1:A:122:ALA:HB3	1:A:123:PRO:HD3	1.93	0.50
1:B:74:LYS:NZ	1:B:187:ALA:O	2.36	0.50
1:A:41:SER:HG	1:A:43:SER:HB2	1.76	0.50
2:D:71:SER:O	2:D:79:LEU:HD12	2.12	0.49
1:A:372:MET:HE1	1:A:384:VAL:HB	1.94	0.49
2:C:32:TYR:OH	4:C:502:BNG:O3	2.28	0.49
1:A:103:ILE:HG13	1:A:104:LEU:N	2.29	0.48
1:A:236:TYR:HH	1:A:322:HIS:HD1	1.61	0.48
1:A:46:TRP:CZ2	4:C:502:BNG:H5'2	2.49	0.48
2:C:36:TRP:HB2	2:C:48:VAL:HG23	1.96	0.48
1:B:318:LEU:HD23	1:B:321:LEU:HD12	1.96	0.48
2:C:65:LYS:O	2:C:65:LYS:HD3	2.14	0.47
1:A:152:ALA:HB2	1:A:269:GLU:HB2	1.95	0.47
1:A:364:PHE:O	1:A:367:VAL:HG22	2.14	0.47
1:A:374:GLU:O	2:C:28:PRO:HD3	2.15	0.47
1:A:44:ASP:HA	1:A:47:ILE:HG12	1.96	0.47
2:C:39:GLN:OE1	2:C:45:ARG:N	2.48	0.46
1:A:168:PHE:HD1	4:B:801:BNG:H4'2	1.80	0.46
1:B:145:MET:HG3	1:B:273:ALA:HB1	1.98	0.45
2:D:1:GLN:HB3	2:D:27:ARG:HE	1.81	0.45
1:B:115:GLY:HA2	1:B:119:ASN:HB2	1.98	0.45
1:A:41:SER:HB3	1:A:44:ASP:OD2	2.16	0.45
1:B:237:ASP:CB	1:B:362:MET:HE1	2.47	0.45
1:B:60:GLN:HA	1:B:63:PHE:HB2	1.99	0.45
2:C:39:GLN:OE1	2:C:45:ARG:CA	2.65	0.45
2:C:36:TRP:NE1	2:C:81:LEU:HB2	2.32	0.45
1:B:56:SER:O	1:B:60:GLN:HG2	2.17	0.44
1:A:16:PHE:CE2	1:A:129:ILE:HG21	2.52	0.44
1:B:5:LYS:HG3	1:B:5:LYS:O	2.18	0.44
1:B:335:LYS:O	1:B:339:SER:OG	2.21	0.44
2:C:14:ALA:HB3	2:C:120:HIS:HB3	2.00	0.44
1:A:188:LYS:HA	1:A:188:LYS:HD3	1.82	0.44
1:A:74:LYS:HG3	1:A:74:LYS:O	2.17	0.44
1:B:103:ILE:HG13	1:B:104:LEU:N	2.33	0.44
1:B:265:THR:O	1:B:269:GLU:HG2	2.17	0.44
1:A:27:PHE:HB3	1:A:28:PRO:HD3	2.00	0.43
1:B:122:ALA:HB3	1:B:123:PRO:HD3	2.00	0.43
1:A:62:LEU:O	1:A:66:LEU:HG	2.18	0.43
1:A:335:LYS:O	1:A:339:SER:OG	2.30	0.43
1:A:224:PHE:HB3	1:A:399:THR:HG22	2.01	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:7:THR:O	1:B:11:MET:HG2	2.19	0.43
1:B:299:MET:HE1	1:B:329:LEU:HD22	2.00	0.43
1:B:211:LYS:O	1:B:211:LYS:HD3	2.19	0.43
2:C:89:GLU:H	2:C:89:GLU:CD	2.21	0.43
1:B:142:ARG:HA	1:B:145:MET:HE3	2.01	0.42
1:A:12:PHE:HB3	1:A:129:ILE:HD12	2.01	0.42
1:B:77:LEU:HB3	1:B:184:LEU:HD21	2.01	0.42
2:C:105:LYS:O	2:C:105:LYS:HG2	2.20	0.42
1:A:265:THR:HG22	1:A:319:LYS:HE3	2.02	0.42
1:B:237:ASP:HB3	1:B:362:MET:HE1	2.01	0.42
1:A:294:LEU:HD23	1:A:328:PHE:CD1	2.54	0.42
1:A:333:CYS:HB2	1:A:354:PHE:CZ	2.55	0.42
1:B:84:LEU:O	1:B:87:PHE:HB2	2.19	0.42
1:A:289:LYS:O	1:A:293:LEU:HG	2.20	0.42
1:A:74:LYS:HB3	1:A:128:PHE:CZ	2.55	0.42
1:B:130:GLU:OE2	1:B:134:ARG:NH1	2.53	0.41
1:B:154:CYS:O	1:B:158:VAL:HG23	2.19	0.41
1:A:17:PHE:HB2	1:A:146:PHE:HB3	2.02	0.41
1:A:310:THR:O	2:C:56:THR:HG21	2.21	0.41
1:A:41:SER:HG	1:A:43:SER:N	2.19	0.41
1:A:104:LEU:O	1:A:108:ILE:HD12	2.20	0.41
1:A:351:LEU:HA	1:A:355:CYS:SG	2.60	0.41
2:C:105:LYS:O	2:C:105:LYS:CG	2.68	0.41
1:B:258:THR:HG21	2:D:102:VAL:HG21	2.03	0.41
1:A:125:VAL:O	1:A:129:ILE:HG12	2.20	0.41
2:D:85:THR:O	2:D:85:THR:HG22	2.21	0.41
1:A:44:ASP:CG	1:A:103:ILE:HD13	2.41	0.41
1:B:271:LEU:HD21	1:B:324:PHE:CE1	2.56	0.40
1:A:23:MET:CE	1:A:122:ALA:HB2	2.52	0.40
1:A:358:LYS:HE3	4:A:701:BNG:H5	2.04	0.40
2:C:30:SER:O	2:C:53:TRP:HB2	2.21	0.40
2:C:71:SER:O	2:C:79:LEU:HD22	2.22	0.40

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

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:135:ARG:HH21	2:C:109:GLN:O[6_654]	1.55	0.05

## 5.3 Torsion angles

### 5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	392/417 (94%)	377 (96%)	15 (4%)	0	100	100
1	B	393/417 (94%)	377 (96%)	16 (4%)	0	100	100
2	C	119/121 (98%)	113 (95%)	6 (5%)	0	100	100
2	D	115/121 (95%)	111 (96%)	3 (3%)	1 (1%)	17	46
All	All	1019/1076 (95%)	978 (96%)	40 (4%)	1 (0%)	51	81

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	14	ALA

### 5.3.2 Protein sidechains

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	333/348 (96%)	325 (98%)	8 (2%)	49	81
1	B	334/348 (96%)	328 (98%)	6 (2%)	59	86
2	C	98/98 (100%)	97 (99%)	1 (1%)	76	93
2	D	94/98 (96%)	93 (99%)	1 (1%)	73	92
All	All	859/892 (96%)	843 (98%)	16 (2%)	57	85

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	41	SER
1	A	42	LYS
1	A	86	MET
1	A	95	PHE
1	A	104	LEU
1	A	145	MET
1	A	311	SER
1	A	333	CYS
2	C	120	HIS
1	B	26	TYR
1	B	104	LEU
1	B	240	ASP
1	B	311	SER
1	B	333	CYS
1	B	406	LEU
2	D	21	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	205	HIS
1	A	219	GLN
1	A	241	GLN
1	A	340	GLN
1	B	38	ASN
1	B	167	GLN
1	B	219	GLN
1	B	359	GLN

### 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](#)

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	$\# Z  > 2$
3	YIO	E	1	3	11,12,12	2.29	1 (9%)	15,17,17	1.23	1 (6%)
3	GAL	E	2	3	11,11,12	1.92	3 (27%)	15,15,17	1.09	2 (13%)
3	YIO	F	1	3	11,12,12	2.20	1 (9%)	15,17,17	1.72	3 (20%)
3	GAL	F	2	3	11,11,12	1.85	3 (27%)	15,15,17	0.86	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	YIO	E	1	3	-	2/2/22/22	0/1/1/1
3	GAL	E	2	3	-	0/2/19/22	0/1/1/1
3	YIO	F	1	3	-	0/2/22/22	0/1/1/1
3	GAL	F	2	3	-	0/2/19/22	0/1/1/1

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	1	YIO	O5-C1	7.03	1.53	1.42
3	F	1	YIO	O5-C1	6.59	1.52	1.42
3	E	2	GAL	O5-C1	4.15	1.50	1.43
3	F	2	GAL	O5-C1	3.96	1.50	1.43
3	E	2	GAL	O3-C3	2.55	1.49	1.43
3	E	2	GAL	C2-C3	-2.41	1.49	1.52
3	F	2	GAL	O3-C3	2.40	1.48	1.43
3	F	2	GAL	C2-C3	-2.30	1.49	1.52

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	1	YIO	O5-C1-C2	5.21	116.87	110.31
3	E	1	YIO	O5-C1-C2	3.72	115.00	110.31

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	2	GAL	C2-C3-C4	2.58	115.35	110.89
3	F	1	YIO	O5-C5-C4	2.29	113.86	109.69
3	F	1	YIO	C1-O5-C5	2.20	116.62	112.57
3	E	2	GAL	C3-C4-C5	2.00	113.81	110.24

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	E	1	YIO	O5-C5-C6-O6
3	E	1	YIO	C4-C5-C6-O6

There are no ring outliers.

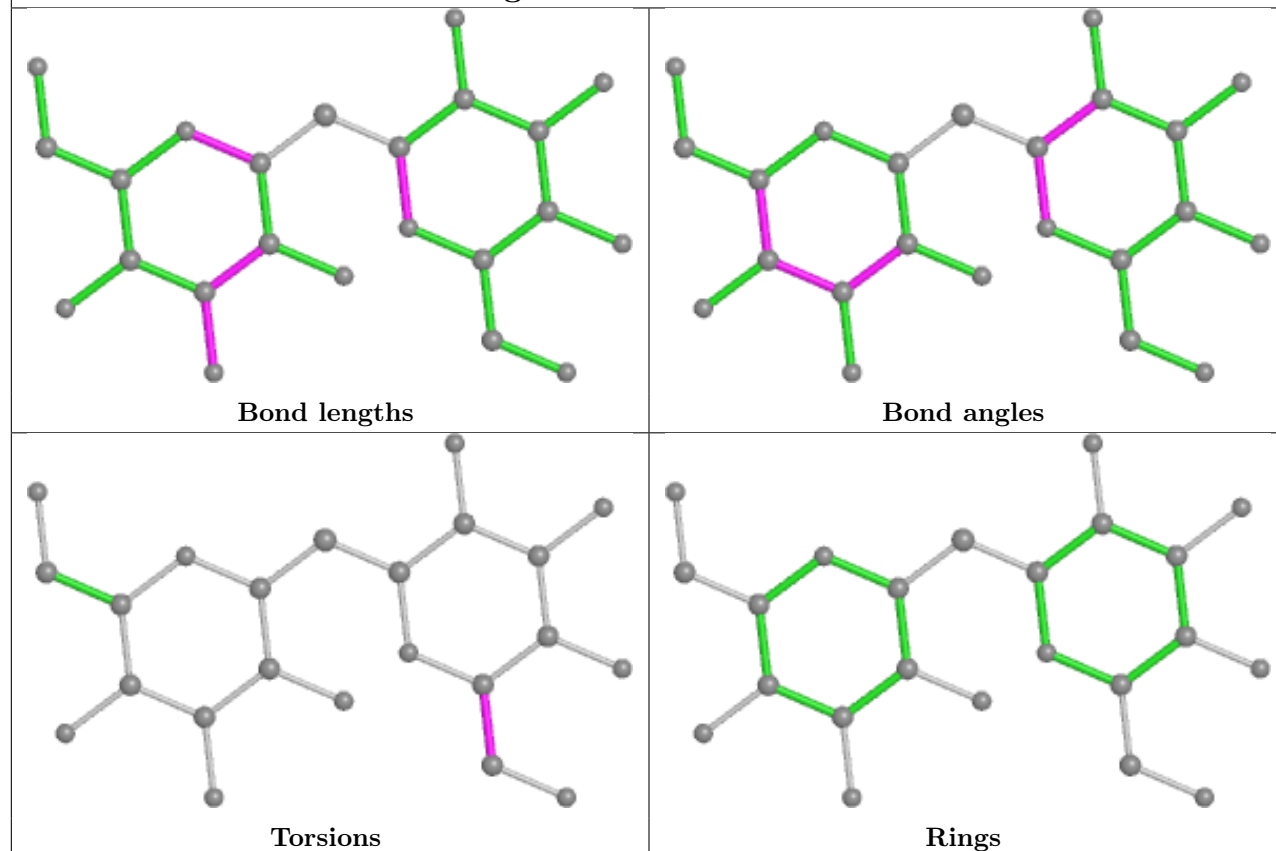
2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	F	1	YIO	1	0
3	E	2	GAL	1	0

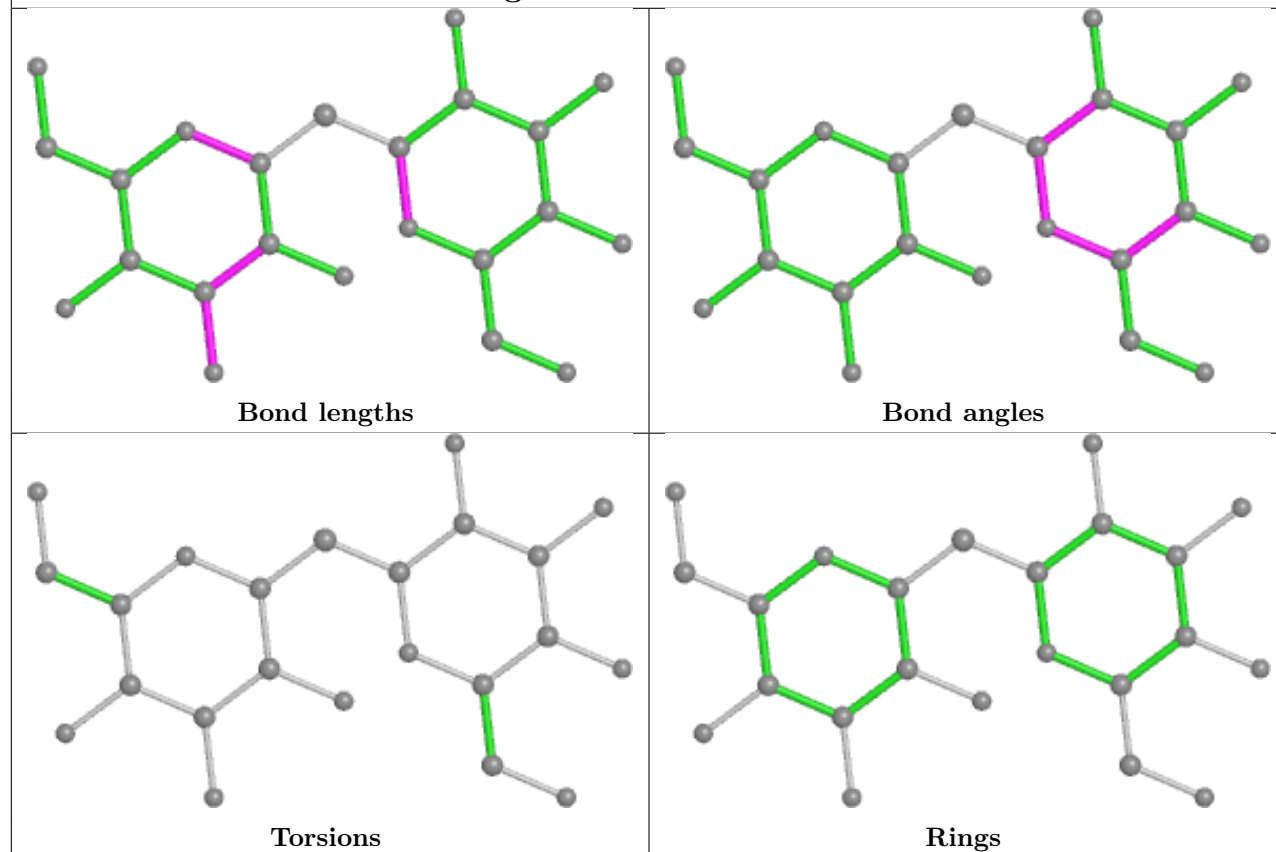
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.



## Oligosaccharide Chain E



## Oligosaccharide Chain F



## 5.6 Ligand geometry

9 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z  > 2$	Counts	RMSZ	# $ Z  > 2$
4	BNG	B	801	-	21,21,21	1.17	1 (4%)	26,26,26	1.51	4 (15%)
4	BNG	A	702	-	21,21,21	1.15	1 (4%)	26,26,26	1.04	0
4	BNG	D	601	-	21,21,21	1.17	2 (9%)	26,26,26	1.05	1 (3%)
4	BNG	A	701	-	21,21,21	1.11	2 (9%)	26,26,26	0.94	1 (3%)
4	BNG	B	803	-	21,21,21	1.15	2 (9%)	26,26,26	0.93	0
4	BNG	B	804	-	21,21,21	1.10	1 (4%)	26,26,26	0.85	0
4	BNG	C	501	-	21,21,21	1.06	1 (4%)	26,26,26	1.02	2 (7%)
4	BNG	B	802	-	21,21,21	1.25	2 (9%)	26,26,26	1.52	3 (11%)
4	BNG	C	502	-	21,21,21	1.21	2 (9%)	26,26,26	1.31	2 (7%)

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	BNG	B	801	-	-	5/12/32/32	0/1/1/1
4	BNG	A	702	-	-	1/12/32/32	0/1/1/1
4	BNG	D	601	-	-	4/12/32/32	0/1/1/1
4	BNG	A	701	-	-	8/12/32/32	0/1/1/1
4	BNG	B	803	-	-	7/12/32/32	0/1/1/1
4	BNG	B	804	-	-	5/12/32/32	0/1/1/1
4	BNG	C	501	-	-	6/12/32/32	0/1/1/1
4	BNG	B	802	-	-	7/12/32/32	0/1/1/1
4	BNG	C	502	-	-	8/12/32/32	0/1/1/1

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	801	BNG	O5-C1	4.10	1.52	1.41
4	C	502	BNG	O5-C1	4.08	1.52	1.41
4	B	803	BNG	O5-C1	3.90	1.51	1.41
4	A	701	BNG	O5-C1	3.60	1.51	1.41
4	D	601	BNG	O5-C1	3.49	1.50	1.41
4	A	702	BNG	O5-C1	3.47	1.50	1.41
4	B	804	BNG	O5-C1	3.45	1.50	1.41
4	B	802	BNG	O5-C1	3.31	1.50	1.41
4	C	501	BNG	O5-C1	3.12	1.49	1.41
4	B	802	BNG	O1-C1	-3.09	1.34	1.40
4	C	502	BNG	O1-C1	-2.52	1.35	1.40
4	D	601	BNG	O1-C1	-2.28	1.36	1.40
4	B	803	BNG	O1-C1	-2.06	1.36	1.40
4	A	701	BNG	O1-C1	-2.02	1.36	1.40

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	502	BNG	O1-C1-C2	4.69	115.63	108.30
4	B	802	BNG	C3-C4-C5	4.65	118.53	110.24
4	B	801	BNG	O1-C1-C2	3.59	113.90	108.30
4	B	802	BNG	O5-C5-C4	3.50	116.06	109.69
4	B	801	BNG	C1'-O1-C1	3.27	119.26	113.84
4	C	502	BNG	C1'-O1-C1	-2.91	109.02	113.84
4	B	801	BNG	O5-C1-O1	-2.88	103.17	109.97
4	D	601	BNG	C1'-O1-C1	2.51	118.00	113.84
4	B	802	BNG	C6-C5-C4	-2.48	107.20	113.00
4	B	801	BNG	O5-C1-C2	2.26	115.14	110.35
4	A	701	BNG	O5-C5-C4	2.25	113.78	109.69
4	C	501	BNG	C1-C2-C3	2.14	114.46	110.00
4	C	501	BNG	O1-C1-C2	-2.13	104.98	108.30

There are no chirality outliers.

All (51) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	801	BNG	O5-C1-O1-C1'
4	C	501	BNG	O5-C1-O1-C1'
4	C	501	BNG	C2'-C1'-O1-C1
4	B	802	BNG	C2'-C1'-O1-C1
4	C	502	BNG	C2-C1-O1-C1'
4	C	502	BNG	O5-C1-O1-C1'

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
4	C	502	BNG	C1'-C2'-C3'-C4'
4	A	701	BNG	C4-C5-C6-O6
4	A	701	BNG	O5-C5-C6-O6
4	B	803	BNG	C4-C5-C6-O6
4	B	803	BNG	O5-C5-C6-O6
4	C	502	BNG	O5-C5-C6-O6
4	D	601	BNG	C2-C1-O1-C1'
4	A	701	BNG	C2-C1-O1-C1'
4	C	502	BNG	C4-C5-C6-O6
4	A	701	BNG	O5-C1-O1-C1'
4	B	804	BNG	O5-C1-O1-C1'
4	B	802	BNG	O1-C1'-C2'-C3'
4	B	803	BNG	C2-C1-O1-C1'
4	A	701	BNG	C2'-C3'-C4'-C5'
4	C	502	BNG	O1-C1'-C2'-C3'
4	B	803	BNG	O5-C1-O1-C1'
4	B	804	BNG	C4-C5-C6-O6
4	C	501	BNG	C4'-C5'-C6'-C7'
4	C	501	BNG	O1-C1'-C2'-C3'
4	D	601	BNG	C1'-C2'-C3'-C4'
4	B	801	BNG	C2-C1-O1-C1'
4	B	804	BNG	O5-C5-C6-O6
4	B	802	BNG	O5-C5-C6-O6
4	C	502	BNG	C5'-C6'-C7'-C8'
4	A	701	BNG	C3'-C4'-C5'-C6'
4	B	802	BNG	C5'-C6'-C7'-C8'
4	C	501	BNG	C5'-C6'-C7'-C8'
4	B	803	BNG	C3'-C4'-C5'-C6'
4	B	804	BNG	C3'-C4'-C5'-C6'
4	A	702	BNG	C3'-C4'-C5'-C6'
4	B	802	BNG	C4'-C5'-C6'-C7'
4	C	502	BNG	C3'-C4'-C5'-C6'
4	A	701	BNG	C4'-C5'-C6'-C7'
4	B	802	BNG	C3'-C4'-C5'-C6'
4	D	601	BNG	C3'-C4'-C5'-C6'
4	C	501	BNG	C3'-C4'-C5'-C6'
4	B	802	BNG	C1'-C2'-C3'-C4'
4	B	804	BNG	C4'-C5'-C6'-C7'
4	B	801	BNG	C3'-C4'-C5'-C6'
4	B	803	BNG	C2'-C3'-C4'-C5'
4	B	801	BNG	C4'-C5'-C6'-C7'
4	B	801	BNG	C1'-C2'-C3'-C4'

*Continued on next page...*

*Continued from previous page...*

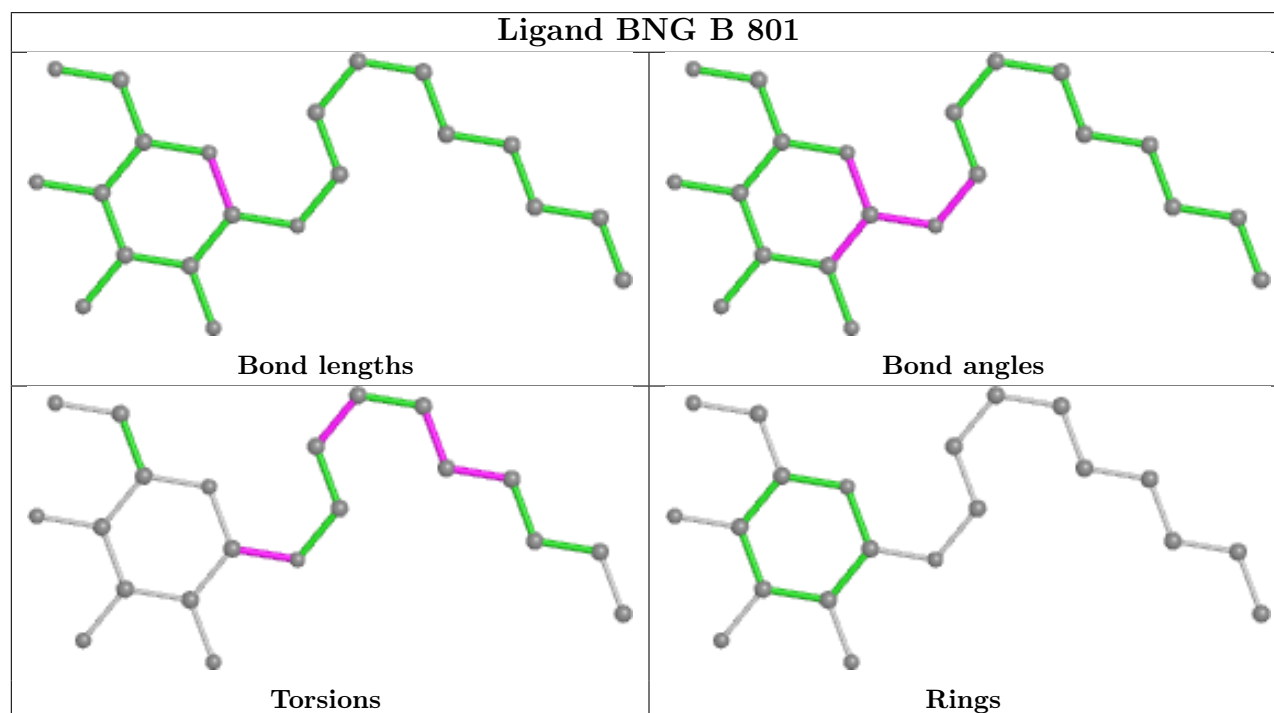
Mol	Chain	Res	Type	Atoms
4	D	601	BNG	O1-C1'-C2'-C3'
4	B	803	BNG	C4'-C5'-C6'-C7'
4	A	701	BNG	C6'-C7'-C8'-C9'

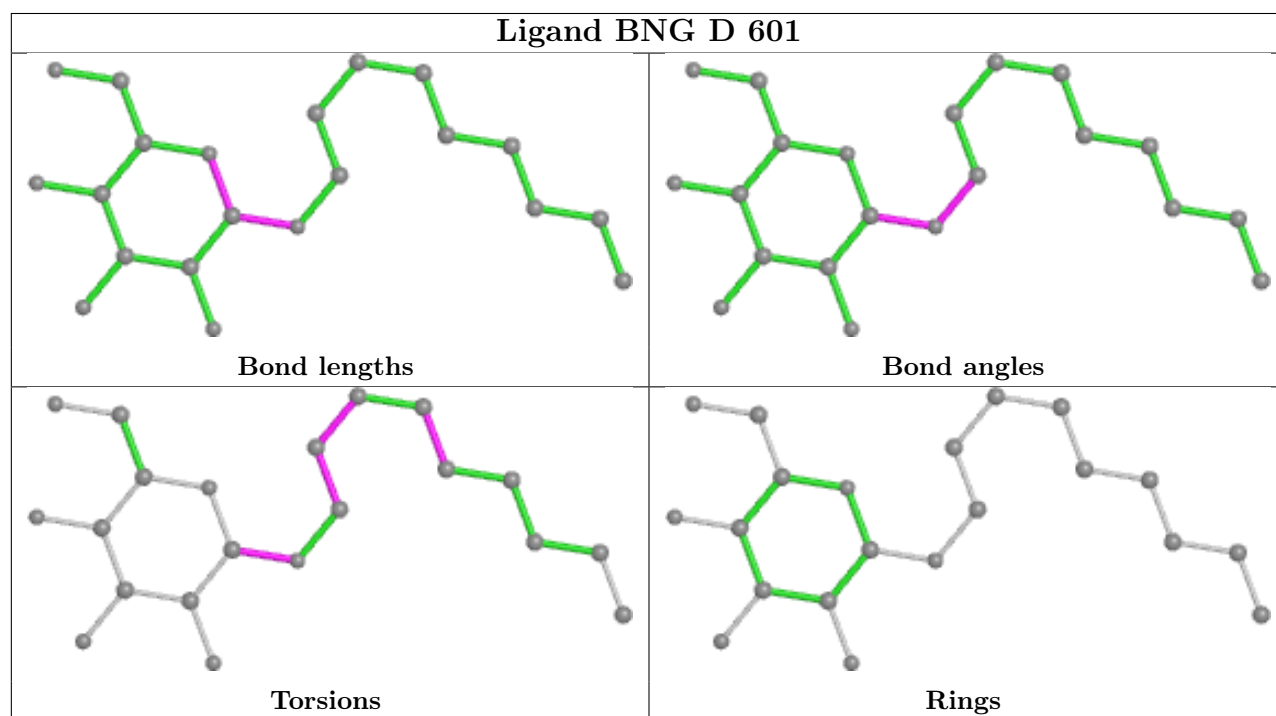
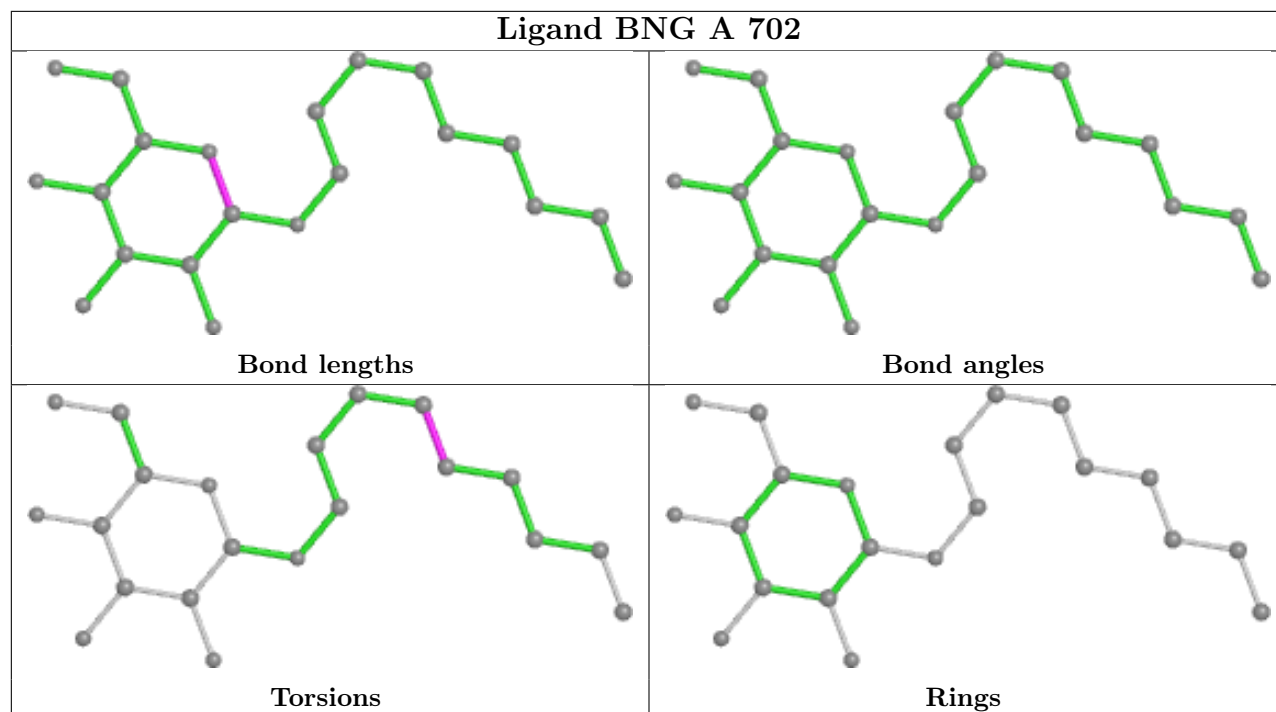
There are no ring outliers.

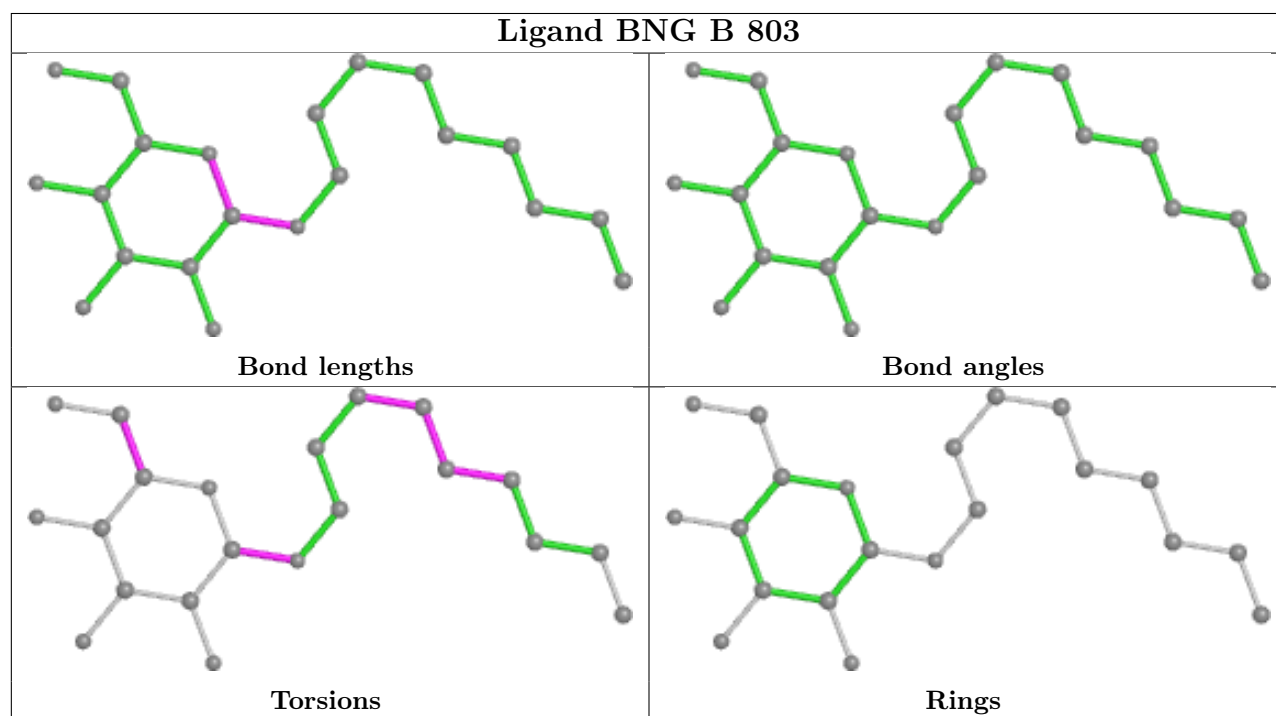
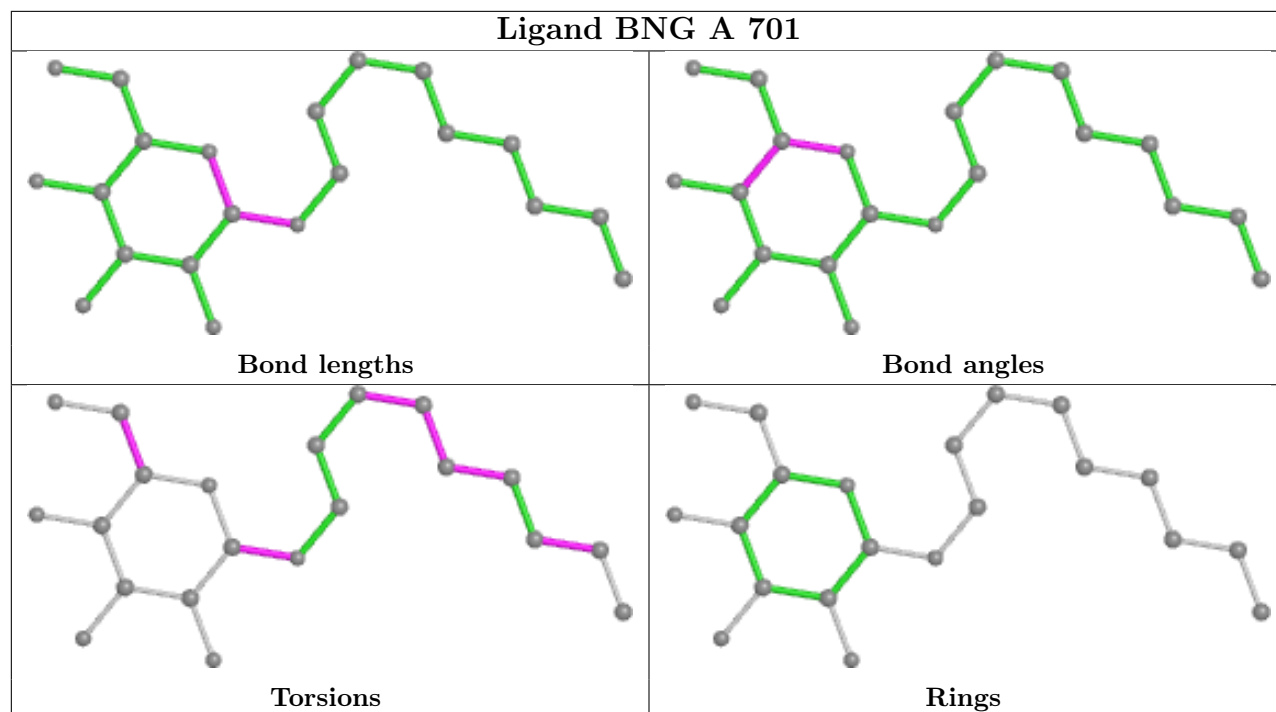
3 monomers are involved in 7 short contacts:

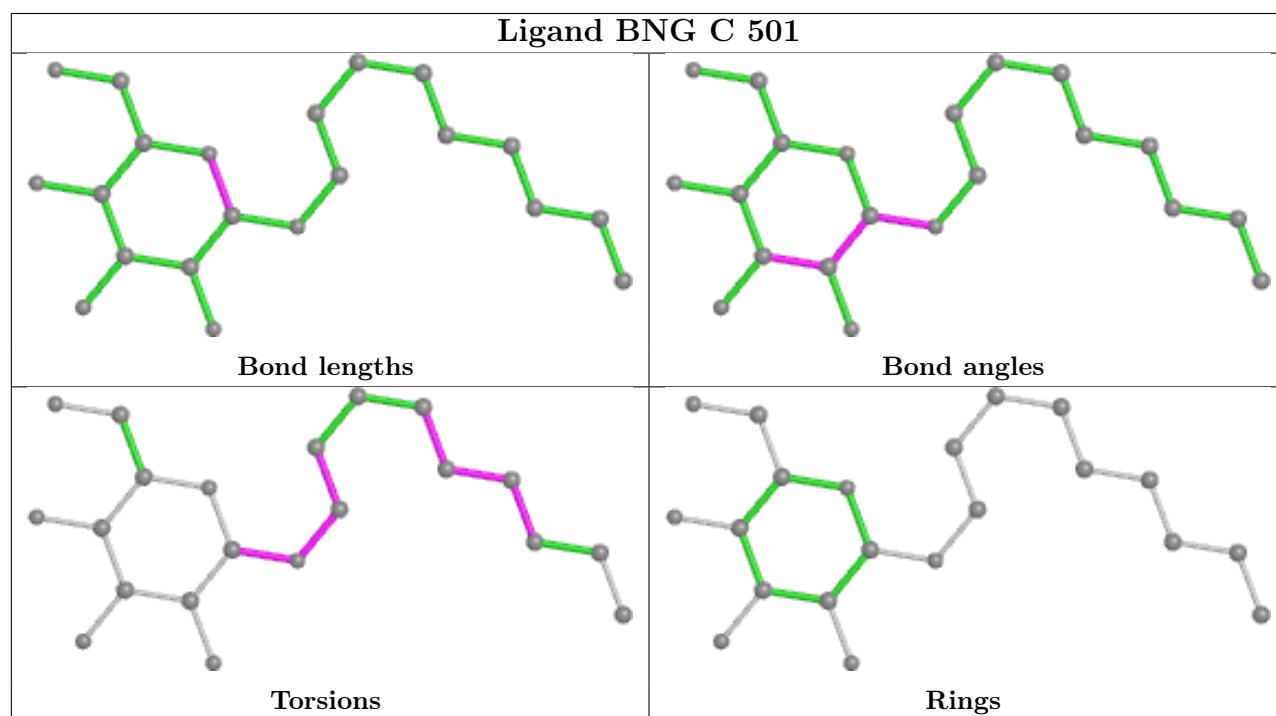
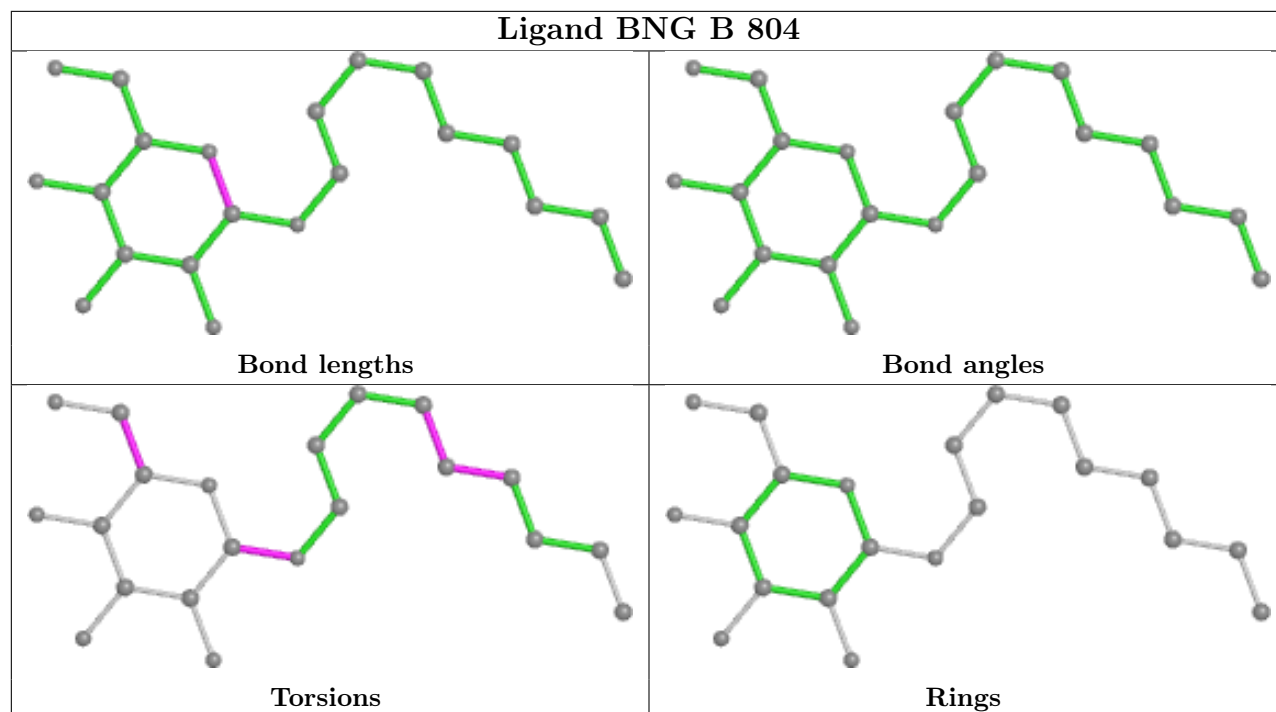
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	801	BNG	1	0
4	A	701	BNG	2	0
4	C	502	BNG	4	0

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 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.

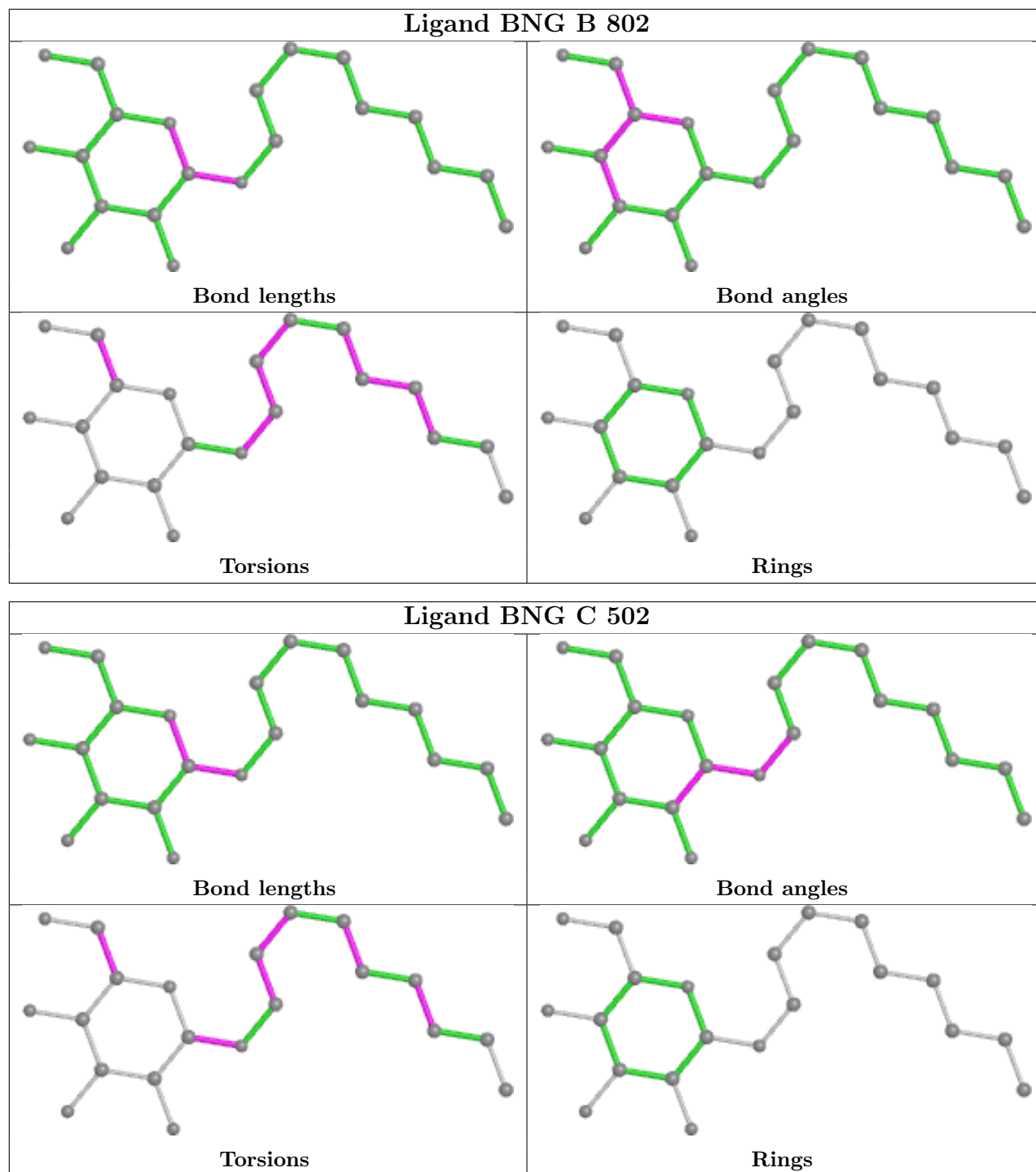












## 5.7 Other polymers [i](#)

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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	396/417 (94%)	-0.14	2 (0%) 91 88	53, 70, 100, 129	0
1	B	397/417 (95%)	-0.12	2 (0%) 91 88	52, 69, 101, 122	0
2	C	121/121 (100%)	-0.15	2 (1%) 70 63	54, 69, 106, 142	0
2	D	117/121 (96%)	-0.26	0 100 100	53, 68, 101, 116	0
All	All	1031/1076 (95%)	-0.15	6 (0%) 89 86	52, 69, 101, 142	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	39	HIS	3.6
2	C	120	HIS	3.1
1	A	203	ALA	2.9
1	B	409	LEU	2.9
2	C	121	HIS	2.6
1	B	39	HIS	2.2

### 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<sup>th</sup> 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.

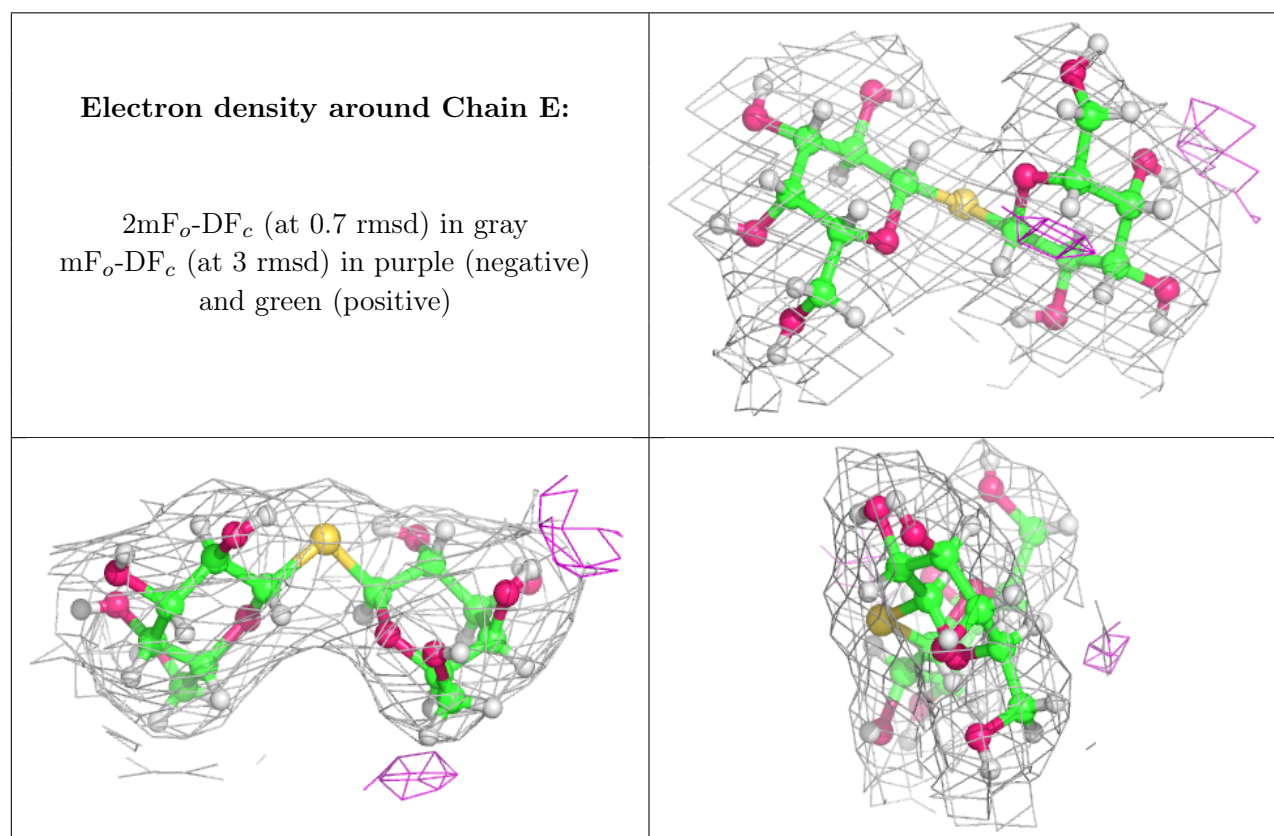
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	YIO	E	1	12/12	0.96	0.18	64,76,92,97	0

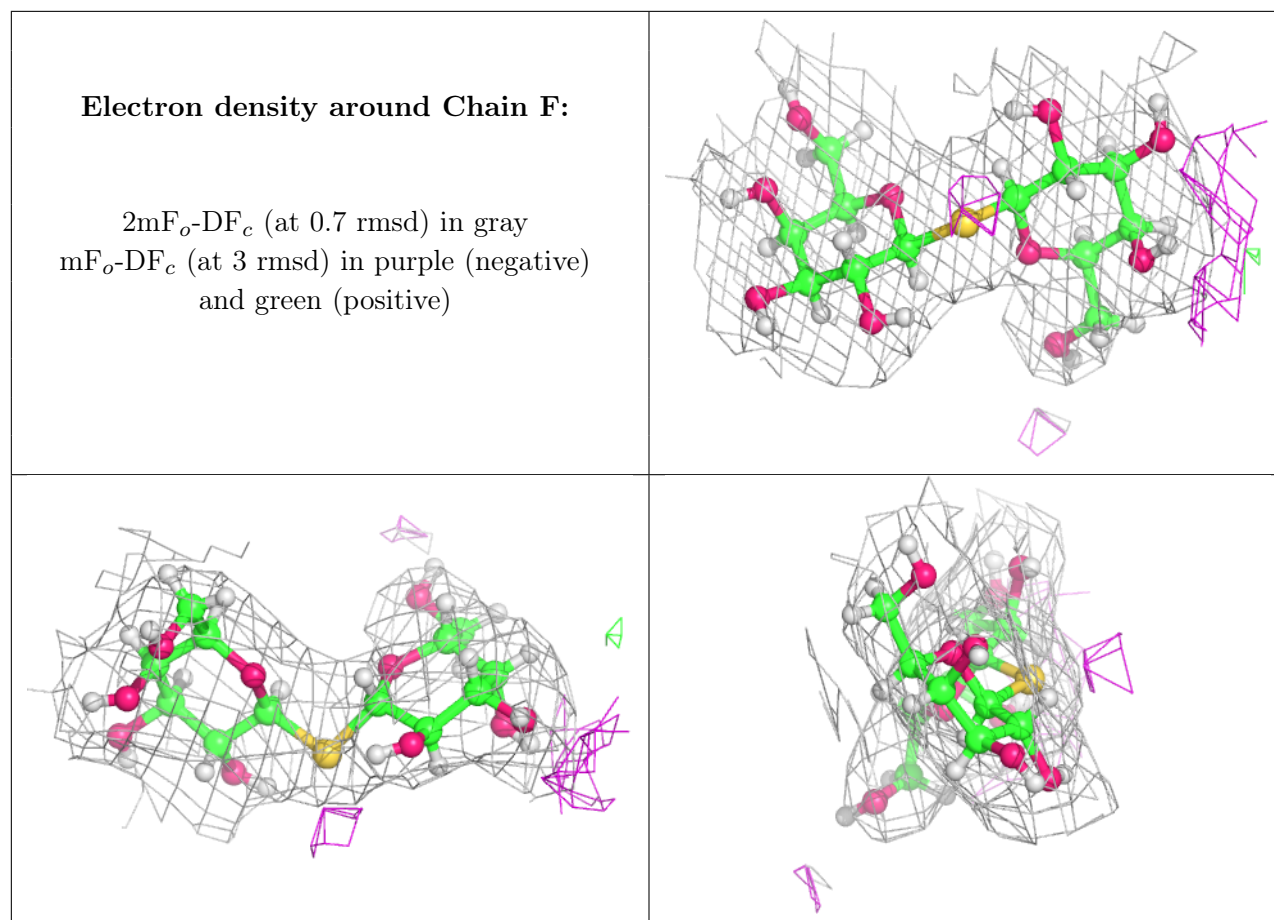
*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	GAL	F	2	11/12	0.98	0.16	47,67,80,92	0
3	YIO	F	1	12/12	0.98	0.21	59,76,96,97	0
3	GAL	E	2	11/12	0.98	0.18	46,63,75,81	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.





## 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, 95<sup>th</sup> 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.

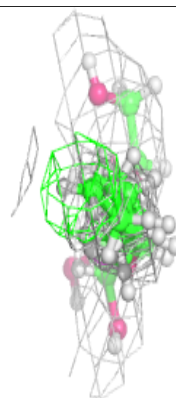
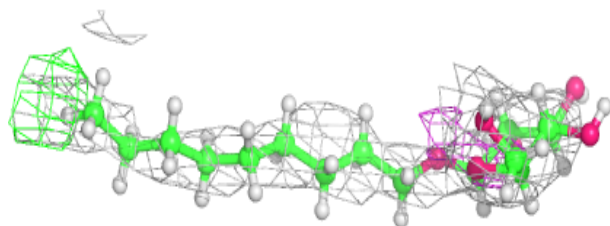
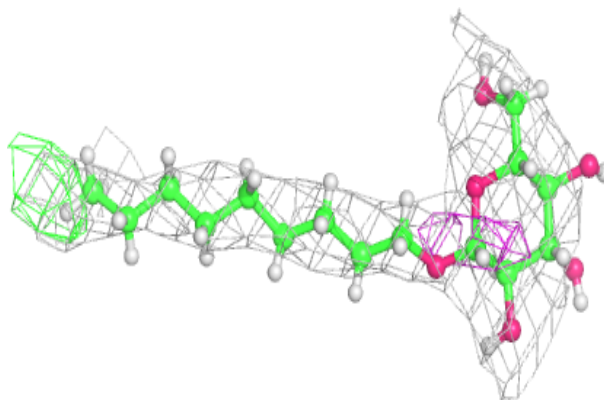
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	BNG	B	804	21/21	0.74	0.36	78,109,141,150	0
4	BNG	B	803	21/21	0.78	0.31	58,102,118,121	0
4	BNG	A	702	21/21	0.78	0.32	80,106,149,155	0
4	BNG	A	701	21/21	0.81	0.25	76,102,125,134	0
4	BNG	C	501	21/21	0.82	0.28	75,126,162,173	0
4	BNG	B	801	21/21	0.83	0.19	86,117,141,144	0
4	BNG	D	601	21/21	0.83	0.17	81,115,138,149	0
4	BNG	C	502	21/21	0.83	0.21	84,125,153,164	0
4	BNG	B	802	21/21	0.87	0.44	74,111,137,150	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.

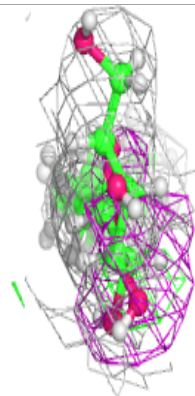
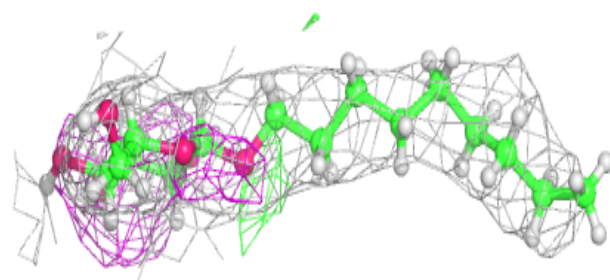
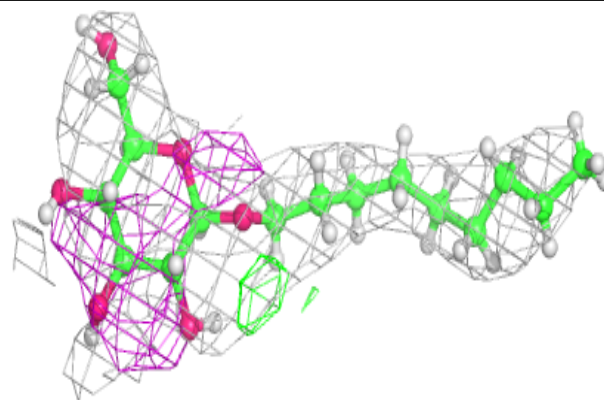
**Electron density around BNG B 804:**

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



**Electron density around BNG B 803:**

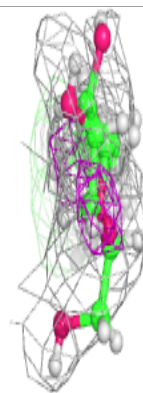
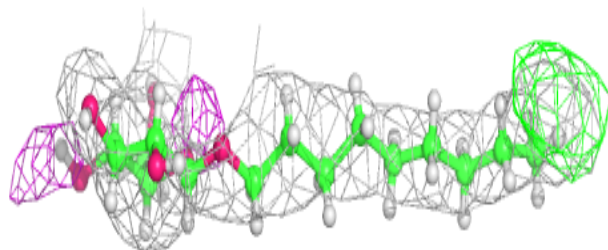
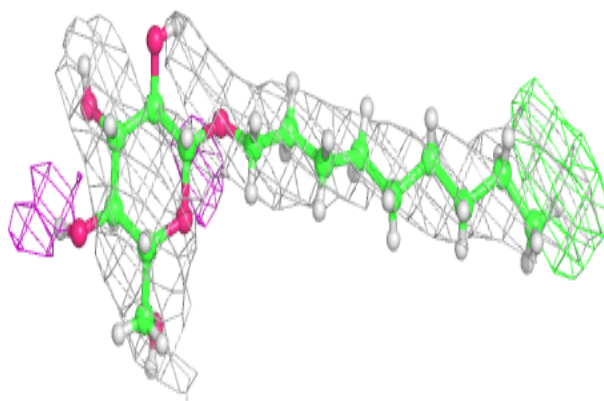
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



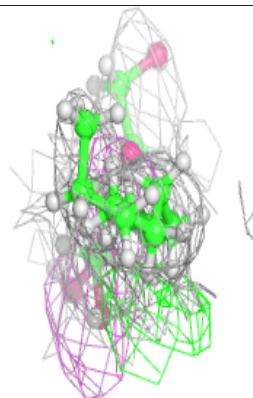
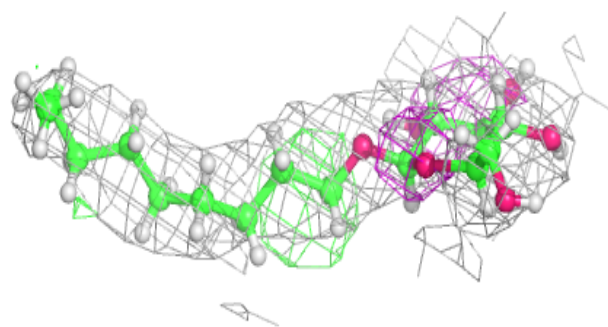
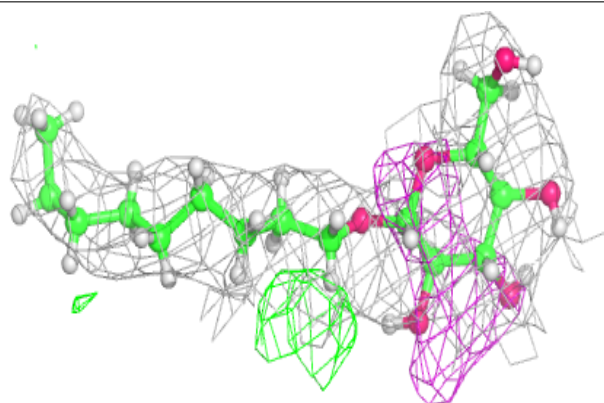


**Electron density around BNG A 702:**

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

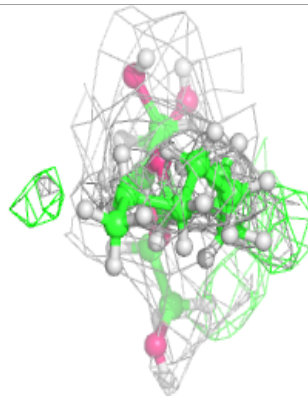
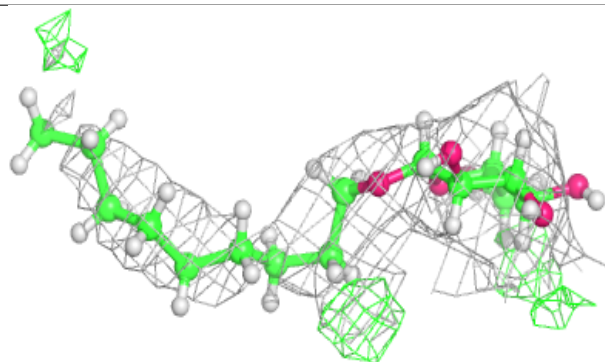
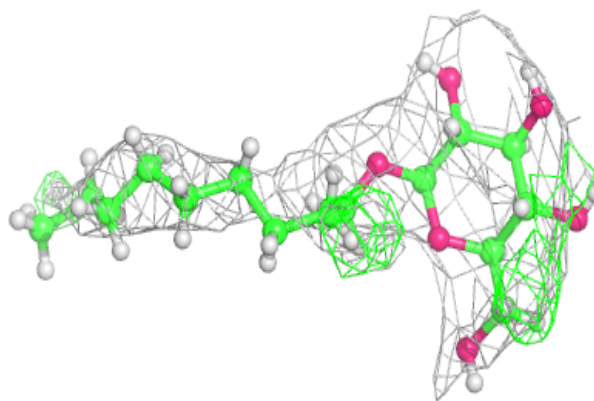
**Electron density around BNG A 701:**

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

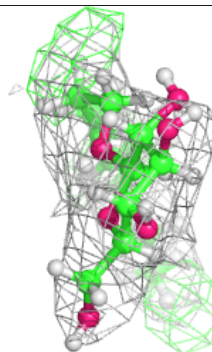
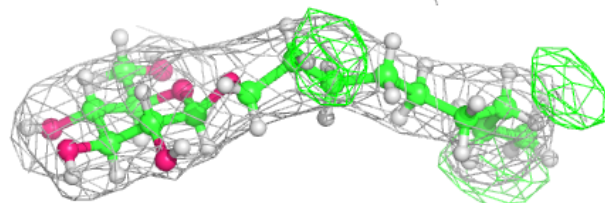
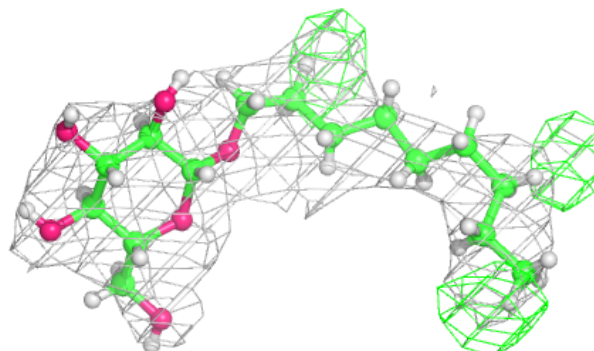


**Electron density around BNG C 501:**

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

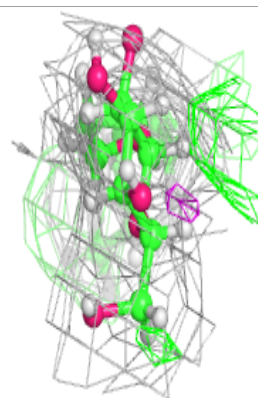
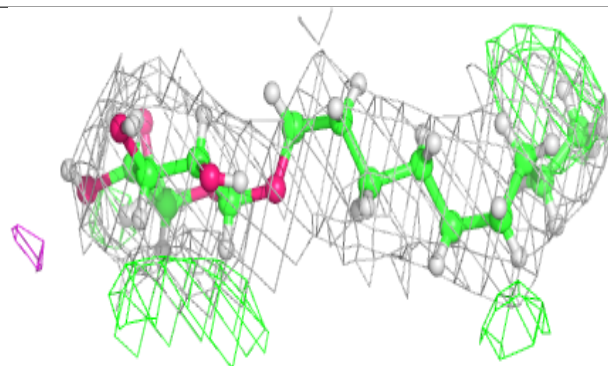
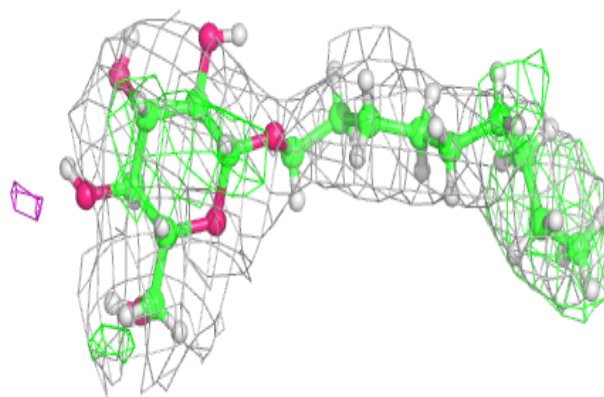
**Electron density around BNG B 801:**

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

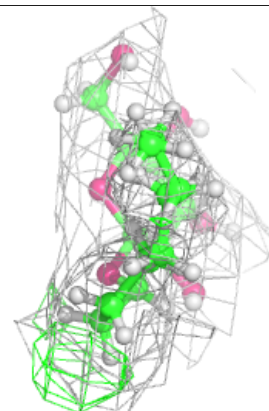
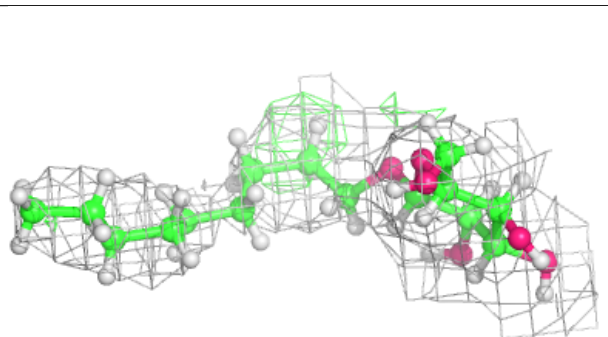
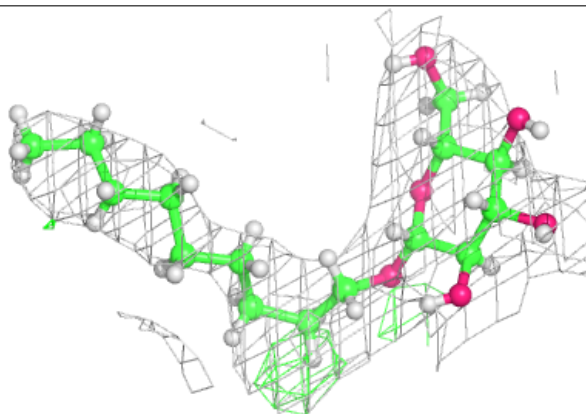


**Electron density around BNG D 601:**

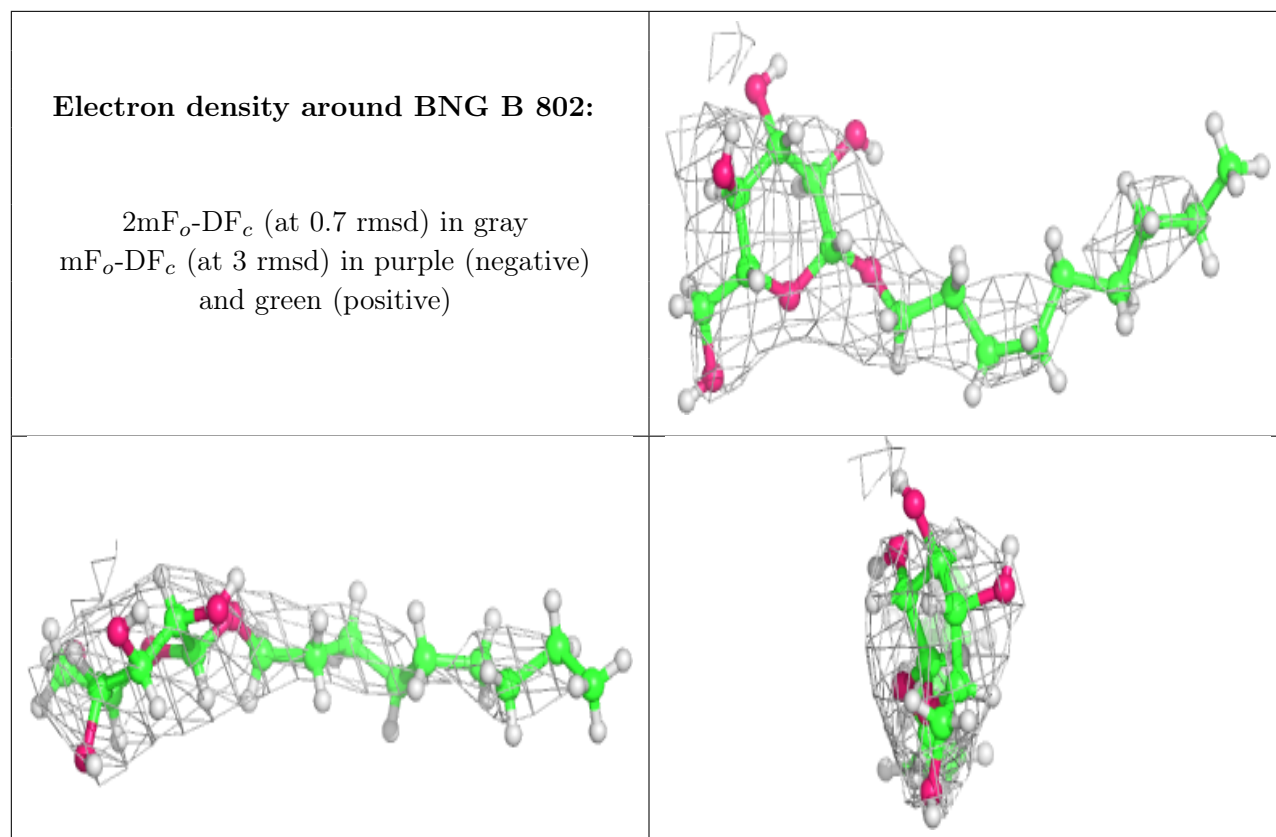
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

**Electron density around BNG C 502:**

$2mF_o-DF_c$  (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.