



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 8, 2020 – 09:54 PM BST

PDB ID : 5XF1  
Title : Structure of the Full-length glucagon class B G protein-coupled receptor  
Authors : Zhang, H.; Qiao, A.; Yang, D.; Yang, L.; Dai, A.; de Graaf, C.; Reedtz-Runge, S.; Dharmarajan, V.; Zhang, H.; Han, G.W.; Grant, T.; Sierra, R.; Weierstall, U.; Nelson, G.; Liu, W.; Wu, Y.; Ma, L.; Cai, X.; Lin, G.; Wu, X.; Geng, Z.; Dong, Y.; Song, G.; Griffin, P.; Lau, J.; Cherezov, V.; Yang, H.; Hanson, M.; Stevens, R.; Jiang, H.; Wang, M.; Zhao, Q.; Wu, B.  
Deposited on : 2017-04-06  
Resolution : 3.19 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 : 2.13.1  
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

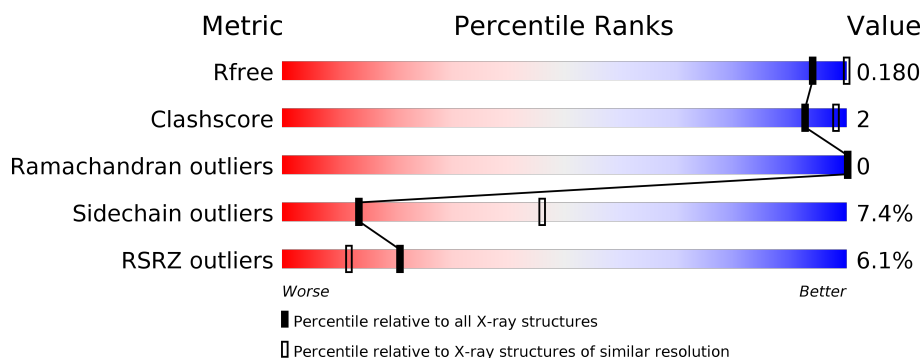
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.19 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.



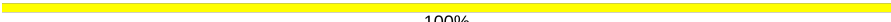


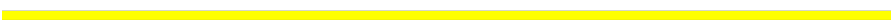


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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  $\geq 3$ , 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	574	<div> <div>8%</div> <div>88%</div> <div>7%</div> <div>5%</div> </div>
1	B	574	<div> <div>11%</div> <div>80%</div> <div>5%</div> <div>14%</div> </div>
2	C	231	<div> <div>%</div> <div>91%</div> <div>9%</div> </div>
2	H	231	<div> <div>%</div> <div>82%</div> <div>16%</div> </div>
3	D	214	<div> <div>%</div> <div>90%</div> <div>9%</div> </div>
3	L	214	<div> <div>%</div> <div>89%</div> <div>10%</div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	E	2	 100%
4	F	2	 50%50%
4	G	2	 50%50%
4	I	2	 100%
4	K	2	 100%
5	J	3	 100%

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 15041 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 Glucagon receptor,Endolysin,Glucagon receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	544	Total	C	N	O	S	0	0	0
			4222	2725	730	744	23			
1	B	496	Total	C	N	O	S	0	0	0
			3773	2434	653	664	22			

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	24	GLY	-	expression tag	UNP P47871
A	25	ALA	-	expression tag	UNP P47871
A	26	PRO	-	expression tag	UNP P47871
A	1054	THR	CYS	engineered mutation	UNP D9IEF7
A	1097	ALA	CYS	engineered mutation	UNP D9IEF7
A	433	GLU	-	expression tag	UNP P47871
A	434	PHE	-	expression tag	UNP P47871
A	435	LEU	-	expression tag	UNP P47871
A	436	GLU	-	expression tag	UNP P47871
A	437	VAL	-	expression tag	UNP P47871
A	438	LEU	-	expression tag	UNP P47871
A	439	PHE	-	expression tag	UNP P47871
A	440	GLN	-	expression tag	UNP P47871
B	24	GLY	-	expression tag	UNP P47871
B	25	ALA	-	expression tag	UNP P47871
B	26	PRO	-	expression tag	UNP P47871
B	1054	THR	CYS	engineered mutation	UNP D9IEF7
B	1097	ALA	CYS	engineered mutation	UNP D9IEF7
B	433	GLU	-	expression tag	UNP P47871
B	434	PHE	-	expression tag	UNP P47871
B	435	LEU	-	expression tag	UNP P47871
B	436	GLU	-	expression tag	UNP P47871
B	437	VAL	-	expression tag	UNP P47871
B	438	LEU	-	expression tag	UNP P47871
B	439	PHE	-	expression tag	UNP P47871

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	440	GLN	-	expression tag	UNP P47871

- Molecule 2 is a protein called Antibody mAb1 Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	231	Total	C	N	O	S	0	0	0
			1756	1107	297	344	8			
2	H	231	Total	C	N	O	S	0	0	0
			1749	1103	296	342	8			

- Molecule 3 is a protein called Antibody mAb1 Light chain.

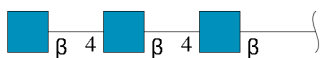
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	214	Total	C	N	O	S	0	0	0
			1629	1013	280	330	6			
3	L	214	Total	C	N	O	S	0	0	0
			1633	1015	280	332	6			

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



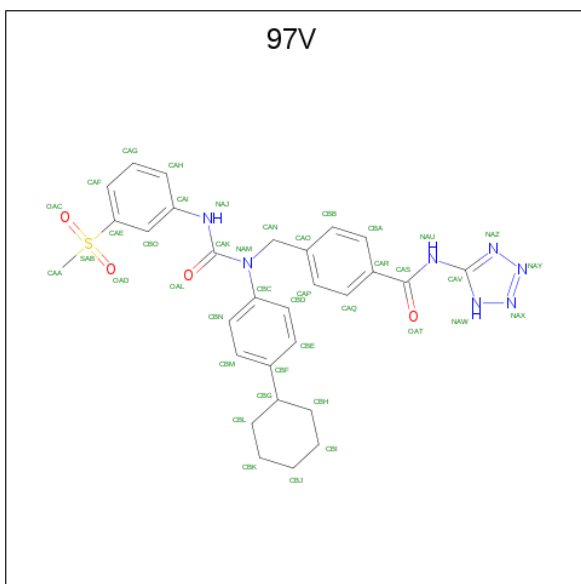
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	E	2	Total	C	N	O	0	0	0
			28	16	2	10			
4	F	2	Total	C	N	O	0	0	0
			28	16	2	10			
4	G	2	Total	C	N	O	0	0	0
			28	16	2	10			
4	I	2	Total	C	N	O	0	0	0
			28	16	2	10			
4	K	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 5 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	J	3	Total	C	N	O	0	0	0
			42	24	3	15			

- Molecule 6 is 4-{[(4-cyclohexylphenyl){[3-(methylsulfonyl)phenyl]carbamoyl}amino]methyl}-N-(1H-tetrazol-5-yl)benzamide (three-letter code: 97V) (formula: C<sub>29</sub>H<sub>31</sub>N<sub>7</sub>O<sub>4</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	A	1	Total	C	N	O	S	0	0
			41	29	7	4	1		
6	B	1	Total	C	N	O	S	0	0
			41	29	7	4	1		

- Molecule 7 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 8 is water.

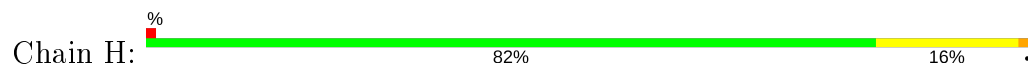
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	1	Total	O	0	0
			1	1		



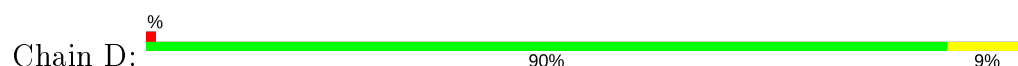




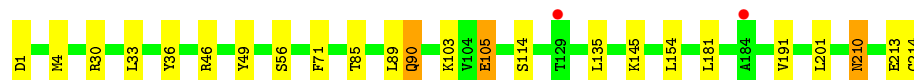
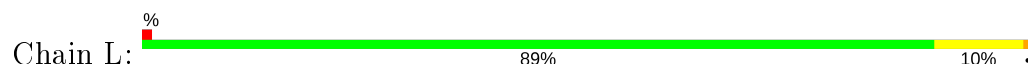
• Molecule 2: Antibody mAb1 Heavy chain



• Molecule 3: Antibody mAb1 Light chain



• Molecule 3: Antibody mAb1 Light chain



• Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



• Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose




• Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



MAG1  
MAG2

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

Chain I:  100%

MAG1  
MAG2

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

Chain K:  100%

MAG1  
MAG2

- Molecule 5: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain J:  100%

MAG1  
MAG2  
MAG3

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.46 Å   248.79 Å   93.33 Å 90.00°   90.16°   90.00°	Depositor
Resolution (Å)	46.92 – 3.19 46.92 – 3.19	Depositor EDS
% Data completeness (in resolution range)	94.1 (46.92-3.19) 98.5 (46.92-3.19)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.60 (at 3.19 Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
R, $R_{free}$	0.205   ,   0.231 0.161   ,   0.180	Depositor DCC
$R_{free}$ test set	2648 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	96.6	Xtriage
Anisotropy	0.140	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26   ,   49.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.109 for h,-k,-l	Xtriage
Reported twinning fraction	0.555 for H, K, L 0.445 for -h,-k,l	Depositor
Outliers	0 of 53171 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	15041	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	120.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.33% 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: 97V, NAG

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	0.48	0/4326	0.61	1/5886 (0.0%)
1	B	0.49	0/3869	0.61	2/5280 (0.0%)
2	C	0.62	0/1799	0.71	0/2448
2	H	0.68	0/1792	0.80	2/2440 (0.1%)
3	D	0.65	0/1662	0.68	0/2254
3	L	0.63	0/1666	0.69	0/2259
All	All	0.57	0/15114	0.66	5/20567 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	377	GLU	CB-CA-C	-7.48	95.44	110.40
1	B	111	ARG	NE-CZ-NH2	-5.57	117.52	120.30
1	A	173	ARG	NE-CZ-NH1	5.35	122.98	120.30
1	B	1154	ARG	NE-CZ-NH1	5.13	122.86	120.30
2	H	378	PRO	CA-N-CD	-5.08	104.39	111.50

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	4222	0	4075	5	0
1	B	3773	0	3553	7	0
2	C	1756	0	1704	1	1
2	H	1749	0	1691	19	1
3	D	1629	0	1586	6	0
3	L	1633	0	1590	9	0
4	E	28	0	25	0	0
4	F	28	0	25	0	0
4	G	28	0	25	0	0
4	I	28	0	25	0	0
4	K	28	0	25	0	0
5	J	42	0	37	0	0
6	A	41	0	0	0	0
6	B	41	0	0	0	0
7	B	14	0	13	0	0
8	A	1	0	0	0	0
All	All	15041	0	14374	44	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 2.

The worst 5 of 44 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:H:375:PHE:CE1	2:H:376:PRO:HB3	2.09	0.88
1:B:256:ALA:HB3	1:B:1002:ASN:HA	1.84	0.59
2:H:375:PHE:CD1	2:H:376:PRO:HB3	2.37	0.58
2:H:444:SER:O	2:H:445:CYS:HB2	2.06	0.55
2:H:374:TYR:CE2	2:H:379:VAL:HG22	2.42	0.54

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)
2:C:385:SER:OG	2:H:287:ASP:OD2[1_455]	2.19	0.01

## 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	540/574 (94%)	509 (94%)	31 (6%)	0	100	100
1	B	492/574 (86%)	470 (96%)	22 (4%)	0	100	100
2	C	229/231 (99%)	213 (93%)	16 (7%)	0	100	100
2	H	229/231 (99%)	210 (92%)	19 (8%)	0	100	100
3	D	212/214 (99%)	202 (95%)	10 (5%)	0	100	100
3	L	212/214 (99%)	200 (94%)	12 (6%)	0	100	100
All	All	1914/2038 (94%)	1804 (94%)	110 (6%)	0	100	100

There are no Ramachandran outliers to report.

### 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	426/493 (86%)	398 (93%)	28 (7%)	16	51
1	B	369/493 (75%)	346 (94%)	23 (6%)	18	53
2	C	196/196 (100%)	179 (91%)	17 (9%)	10	37
2	H	194/196 (99%)	172 (89%)	22 (11%)	6	25
3	D	186/187 (100%)	175 (94%)	11 (6%)	19	54
3	L	187/187 (100%)	173 (92%)	14 (8%)	13	45
All	All	1558/1752 (89%)	1443 (93%)	115 (7%)	13	46

5 of 115 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	232	LEU
2	C	425	CYS
3	L	105	GLU
2	C	234	LEU
2	C	285	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1132	ASN
2	H	421	GLN

### 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 ⓘ

13 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	$\# Z  > 2$
4	NAG	E	1	1,4	14,14,15	0.77	1 (7%)	17,19,21	1.73	2 (11%)
4	NAG	E	2	4	14,14,15	0.42	0	17,19,21	1.58	2 (11%)
4	NAG	F	1	1,4	14,14,15	0.64	0	17,19,21	1.50	3 (17%)
4	NAG	F	2	4	14,14,15	0.42	0	17,19,21	0.66	0
4	NAG	G	1	1,4	14,14,15	0.69	0	17,19,21	1.78	4 (23%)
4	NAG	G	2	4	14,14,15	0.31	0	17,19,21	1.00	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	NAG	I	1	1,4	14,14,15	0.55	0	17,19,21	1.56	2 (11%)
4	NAG	I	2	4	14,14,15	0.52	0	17,19,21	1.13	1 (5%)
5	NAG	J	1	1,5	14,14,15	0.84	1 (7%)	17,19,21	1.79	2 (11%)
5	NAG	J	2	5	14,14,15	0.77	1 (7%)	17,19,21	1.28	3 (17%)
5	NAG	J	3	5	14,14,15	1.17	1 (7%)	17,19,21	1.38	2 (11%)
4	NAG	K	1	1,4	14,14,15	0.70	0	17,19,21	1.48	3 (17%)
4	NAG	K	2	4	14,14,15	0.96	1 (7%)	17,19,21	1.14	1 (5%)

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	NAG	E	1	1,4	-	0/6/23/26	0/1/1/1
4	NAG	E	2	4	-	3/6/23/26	0/1/1/1
4	NAG	F	1	1,4	-	0/6/23/26	0/1/1/1
4	NAG	F	2	4	-	0/6/23/26	0/1/1/1
4	NAG	G	1	1,4	-	1/6/23/26	0/1/1/1
4	NAG	G	2	4	-	0/6/23/26	0/1/1/1
4	NAG	I	1	1,4	-	2/6/23/26	0/1/1/1
4	NAG	I	2	4	-	2/6/23/26	0/1/1/1
5	NAG	J	1	1,5	-	1/6/23/26	0/1/1/1
5	NAG	J	2	5	-	0/6/23/26	0/1/1/1
5	NAG	J	3	5	-	0/6/23/26	0/1/1/1
4	NAG	K	1	1,4	-	1/6/23/26	0/1/1/1
4	NAG	K	2	4	-	0/6/23/26	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	J	3	NAG	C1-C2	3.83	1.58	1.52
4	K	2	NAG	C1-C2	2.73	1.56	1.52
5	J	2	NAG	O4-C4	2.48	1.48	1.43
5	J	1	NAG	O5-C1	-2.45	1.39	1.43
4	E	1	NAG	C1-C2	2.11	1.55	1.52

The worst 5 of 25 bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
4	E	1	NAG	C1-O5-C5	5.75	119.98	112.19
5	J	1	NAG	C1-O5-C5	5.36	119.45	112.19
4	E	2	NAG	C1-O5-C5	4.94	118.88	112.19
5	J	3	NAG	O5-C1-C2	-4.20	104.66	111.29
4	G	1	NAG	O5-C1-C2	-3.88	105.17	111.29

There are no chirality outliers.

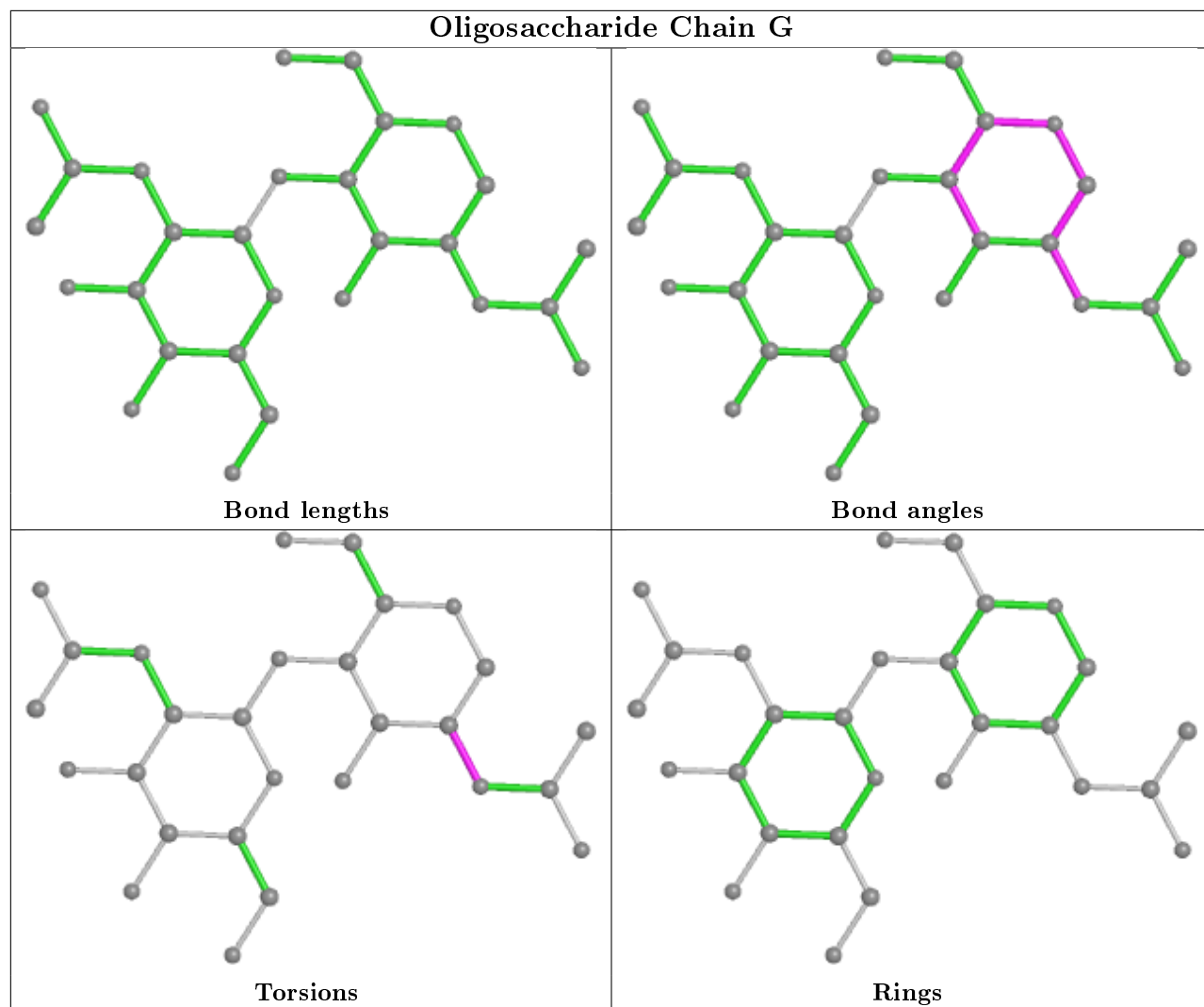
5 of 10 torsion outliers are listed below:

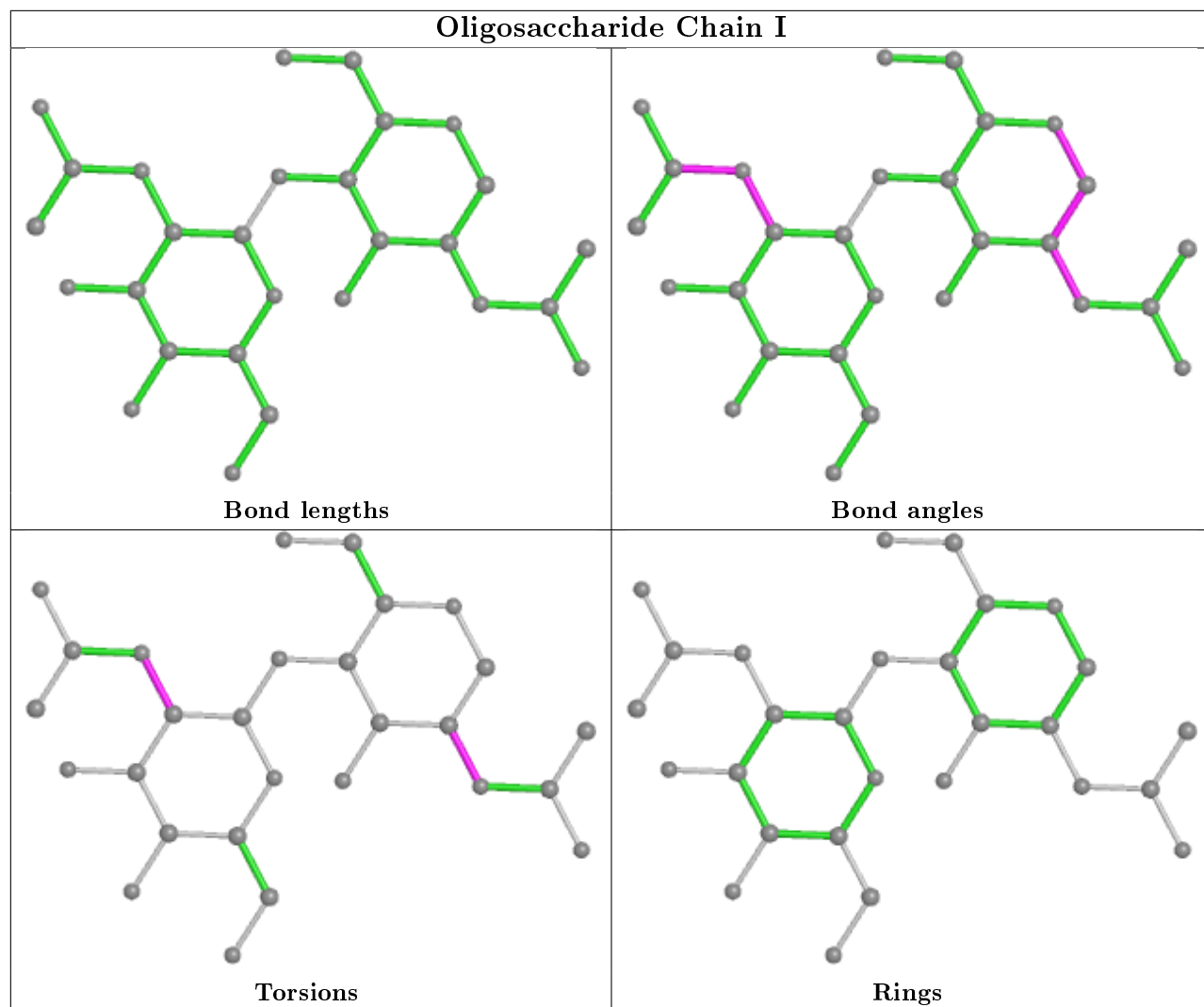
Mol	Chain	Res	Type	Atoms
4	E	2	NAG	O5-C5-C6-O6
4	E	2	NAG	C4-C5-C6-O6
5	J	1	NAG	C3-C2-N2-C7
4	E	2	NAG	C3-C2-N2-C7
4	I	2	NAG	C3-C2-N2-C7

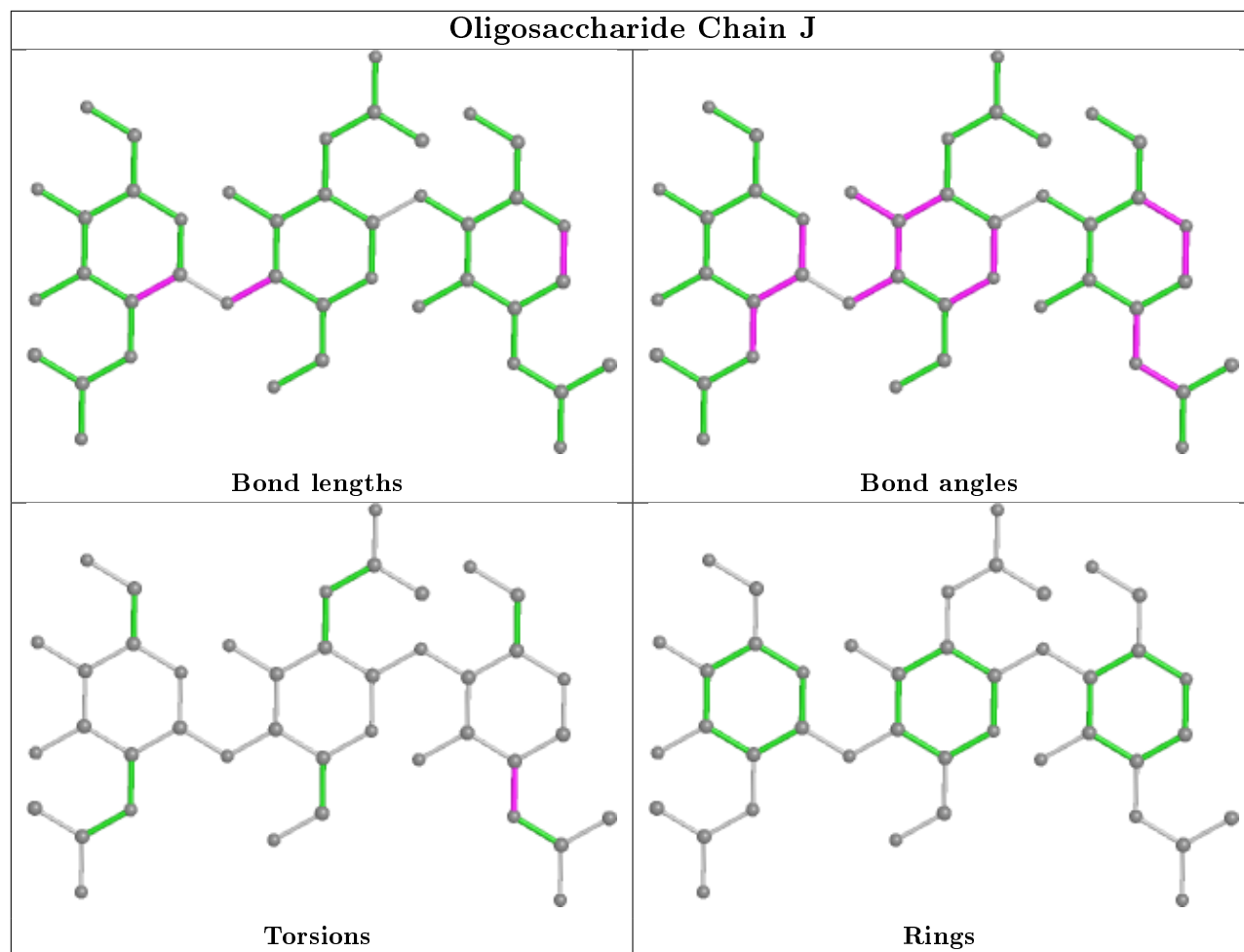
There are no ring outliers.

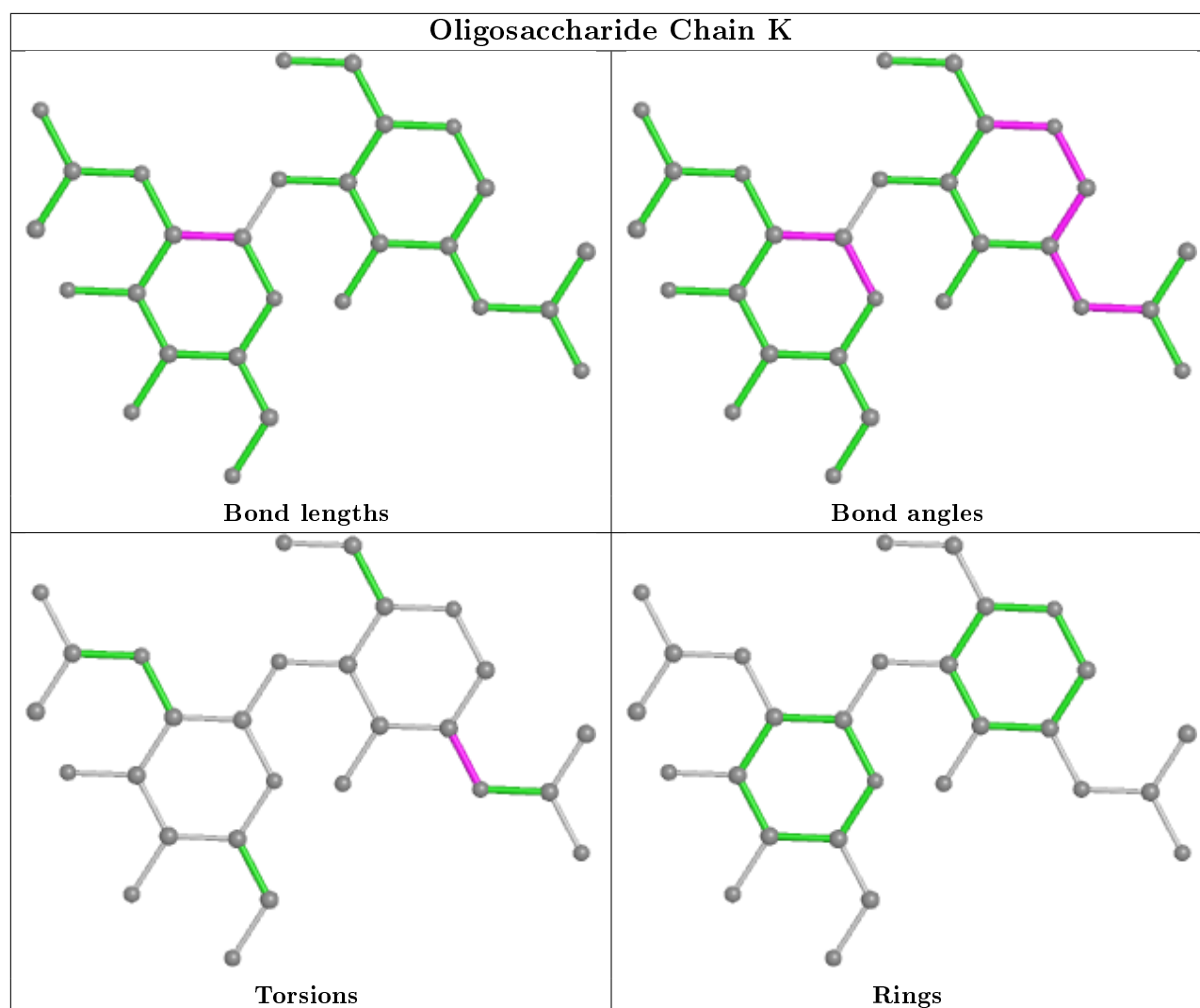
No monomer is involved in short contacts.

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









## 5.6 Ligand geometry ⓘ

3 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	$\# Z  > 2$
6	97V	B	1209	-	43,45,45	1.90	8 (18%)	55,63,63	1.58	8 (14%)
7	NAG	B	1206	1	14,14,15	0.31	0	17,19,21	1.37	2 (11%)
6	97V	A	1207	-	43,45,45	1.93	8 (18%)	55,63,63	1.43	7 (12%)

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
6	97V	B	1209	-	-	12/32/42/42	0/5/5/5
7	NAG	B	1206	1	-	2/6/23/26	0/1/1/1
6	97V	A	1207	-	-	10/32/42/42	0/5/5/5

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	1209	97V	CAN-CAO	-5.62	1.41	1.51
6	A	1207	97V	CAN-CAO	-5.55	1.41	1.51
6	A	1207	97V	CAR-CAS	-5.07	1.39	1.50
6	B	1209	97V	CAR-CAS	-5.02	1.39	1.50
6	B	1209	97V	CBF-CBG	-4.94	1.42	1.52

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	1209	97V	OAD-SAB-OAC	-5.73	108.05	117.92
6	A	1207	97V	OAD-SAB-OAC	-5.15	109.04	117.92
6	B	1209	97V	CAA-SAB-CAE	4.60	110.02	104.58
6	A	1207	97V	CAA-SAB-CAE	4.46	109.85	104.58
6	B	1209	97V	CAI-CBO-CAE	4.18	122.04	118.89

There are no chirality outliers.

5 of 24 torsion outliers are listed below:

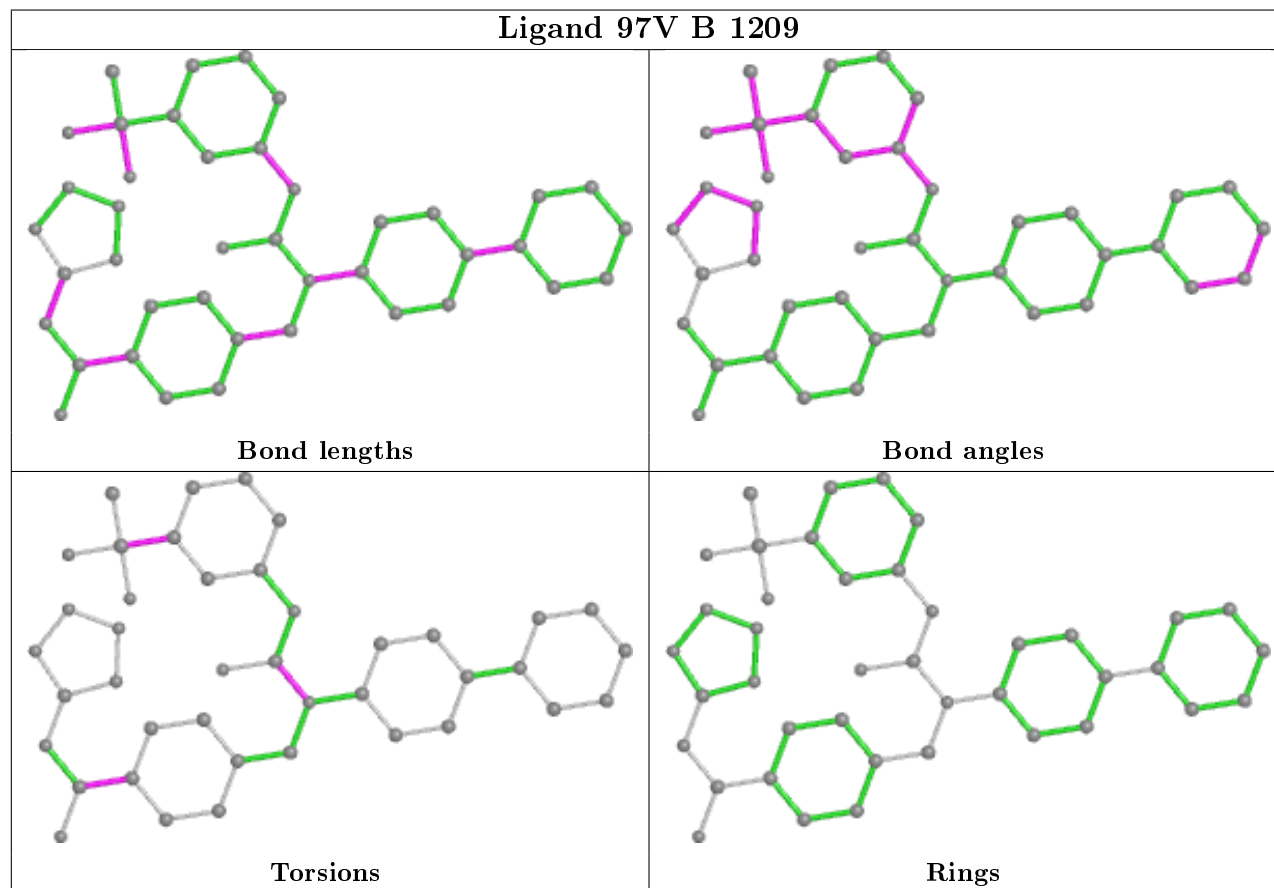
Mol	Chain	Res	Type	Atoms
6	B	1209	97V	NAJ-CAK-NAM-CAN
6	B	1209	97V	OAL-CAK-NAM-CAN
6	A	1207	97V	NAJ-CAK-NAM-CAN
6	A	1207	97V	OAL-CAK-NAM-CAN
6	A	1207	97V	CAQ-CAR-CAS-NAU

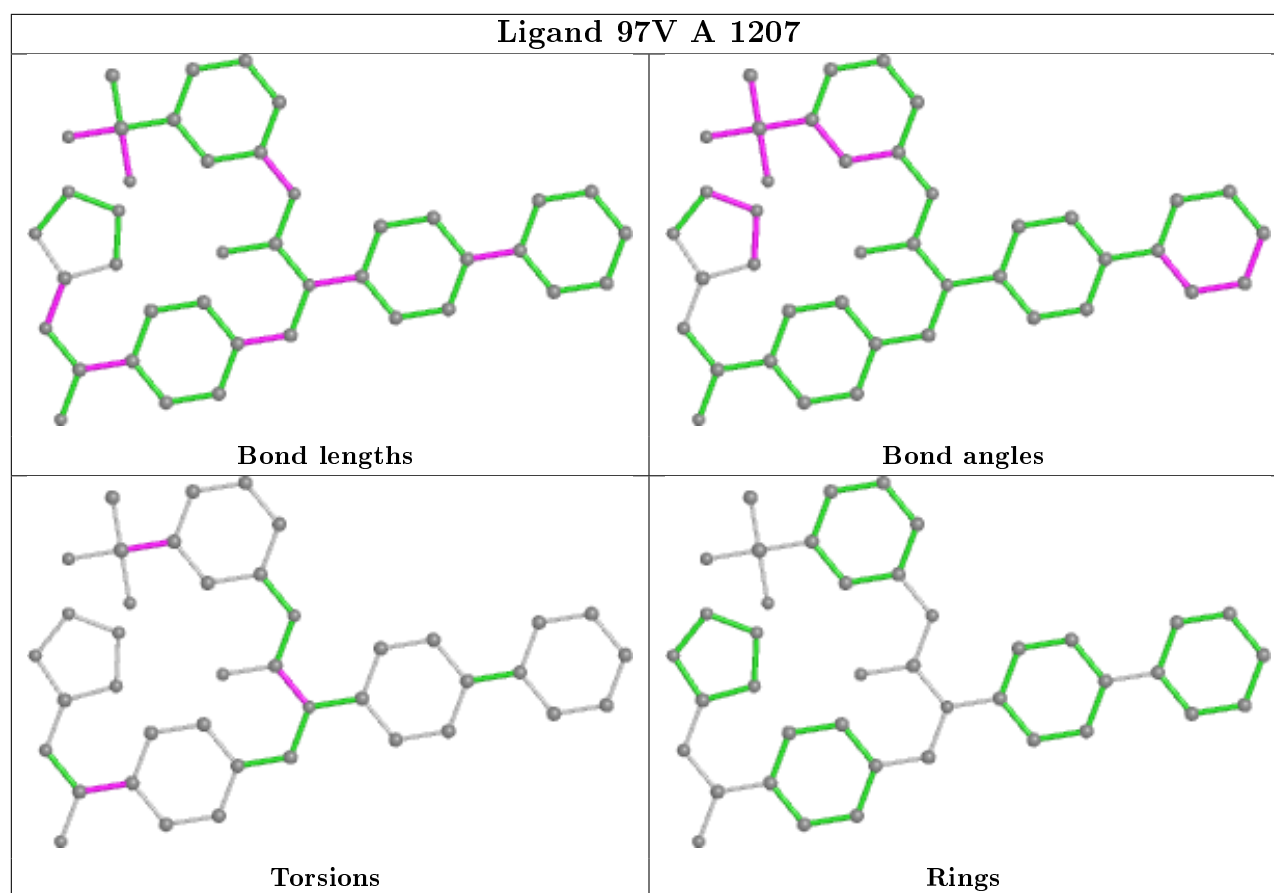
There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In

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





## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.



## 6 Fit of model and data [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	544/574 (94%)	0.42	47 (8%) 10 5	71, 143, 197, 259	0
1	B	496/574 (86%)	0.52	61 (12%) 4 2	69, 152, 197, 231	0
2	C	231/231 (100%)	-0.13	3 (1%) 77 65	62, 89, 123, 183	0
2	H	231/231 (100%)	-0.16	3 (1%) 77 65	53, 83, 118, 193	0
3	D	214/214 (100%)	-0.10	2 (0%) 84 75	58, 83, 112, 143	0
3	L	214/214 (100%)	-0.20	2 (0%) 84 75	58, 80, 119, 173	0
All	All	1930/2038 (94%)	0.18	118 (6%) 21 12	53, 114, 189, 259	0

The worst 5 of 118 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	359	SER	9.8
1	A	305	TRP	7.5
2	C	359	SER	7.0
1	A	1049	ALA	6.7
1	B	25	ALA	6.0

### 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](#)

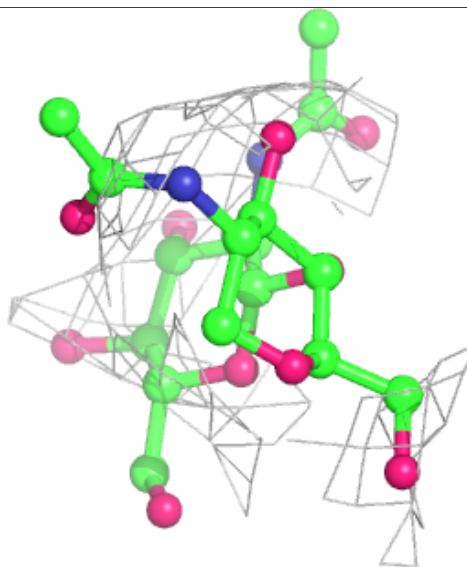
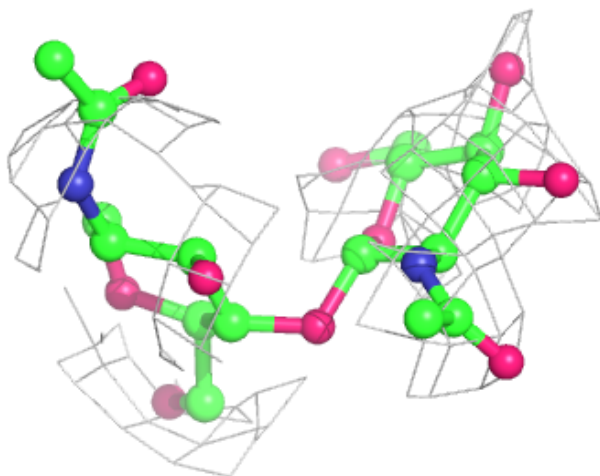
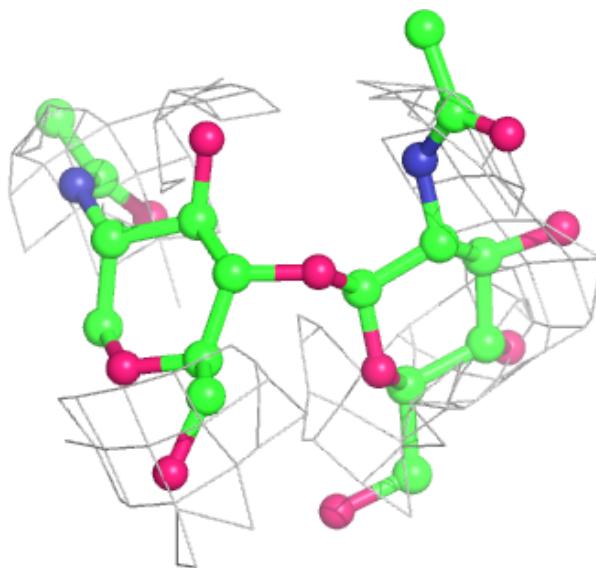
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAG	G	2	14/15	0.82	0.37	179,190,204,204	0
4	NAG	K	2	14/15	0.82	0.26	137,148,154,157	0
4	NAG	E	2	14/15	0.84	0.12	144,163,172,174	0
4	NAG	E	1	14/15	0.84	0.21	146,166,182,190	0
5	NAG	J	3	14/15	0.87	0.15	167,174,182,183	0
4	NAG	F	1	14/15	0.88	0.14	119,141,146,146	0
4	NAG	I	2	14/15	0.89	0.18	139,150,156,160	0
4	NAG	K	1	14/15	0.90	0.16	118,133,137,142	0
4	NAG	F	2	14/15	0.90	0.15	137,149,154,156	0
4	NAG	G	1	14/15	0.91	0.14	157,161,180,182	0
5	NAG	J	2	14/15	0.91	0.15	146,162,192,204	0
5	NAG	J	1	14/15	0.93	0.18	119,132,143,148	0
4	NAG	I	1	14/15	0.95	0.20	145,157,166,168	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.

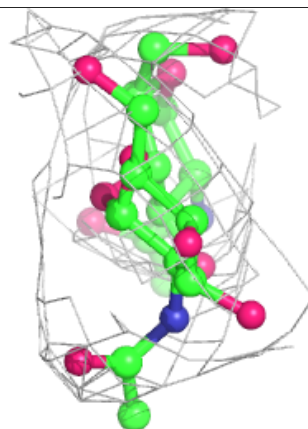
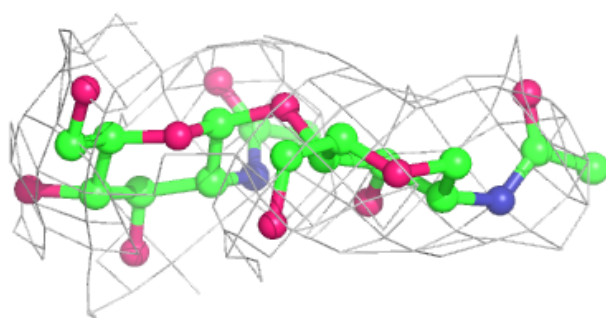
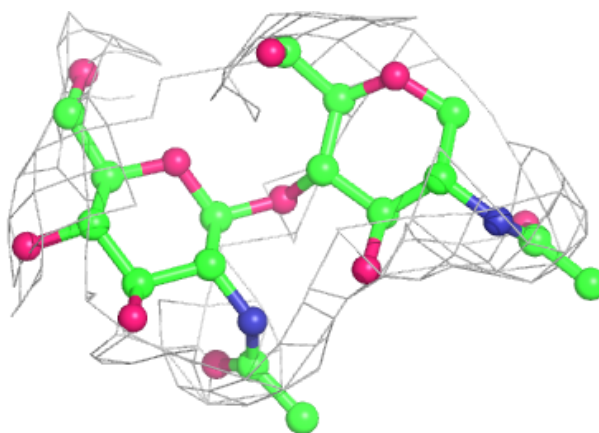
**Electron density around Chain G:**

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

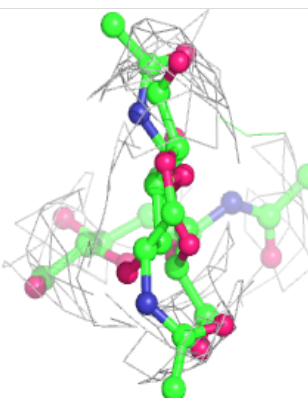
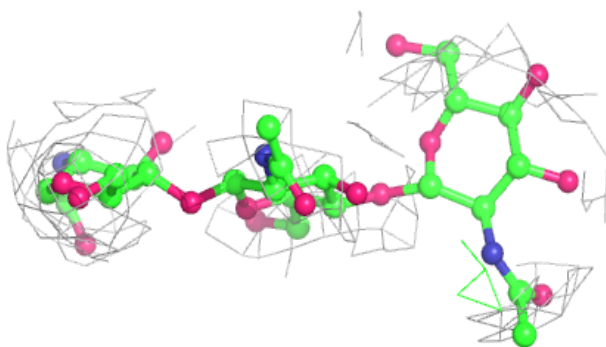
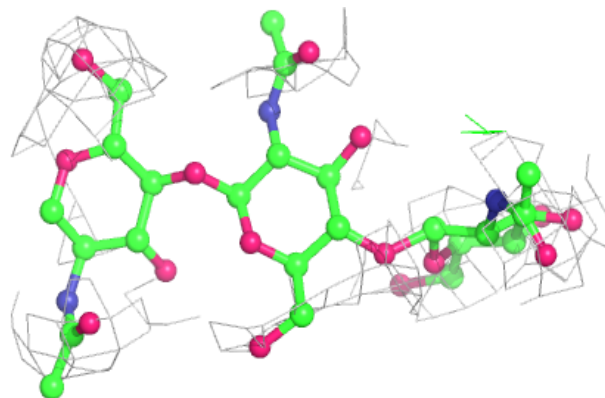


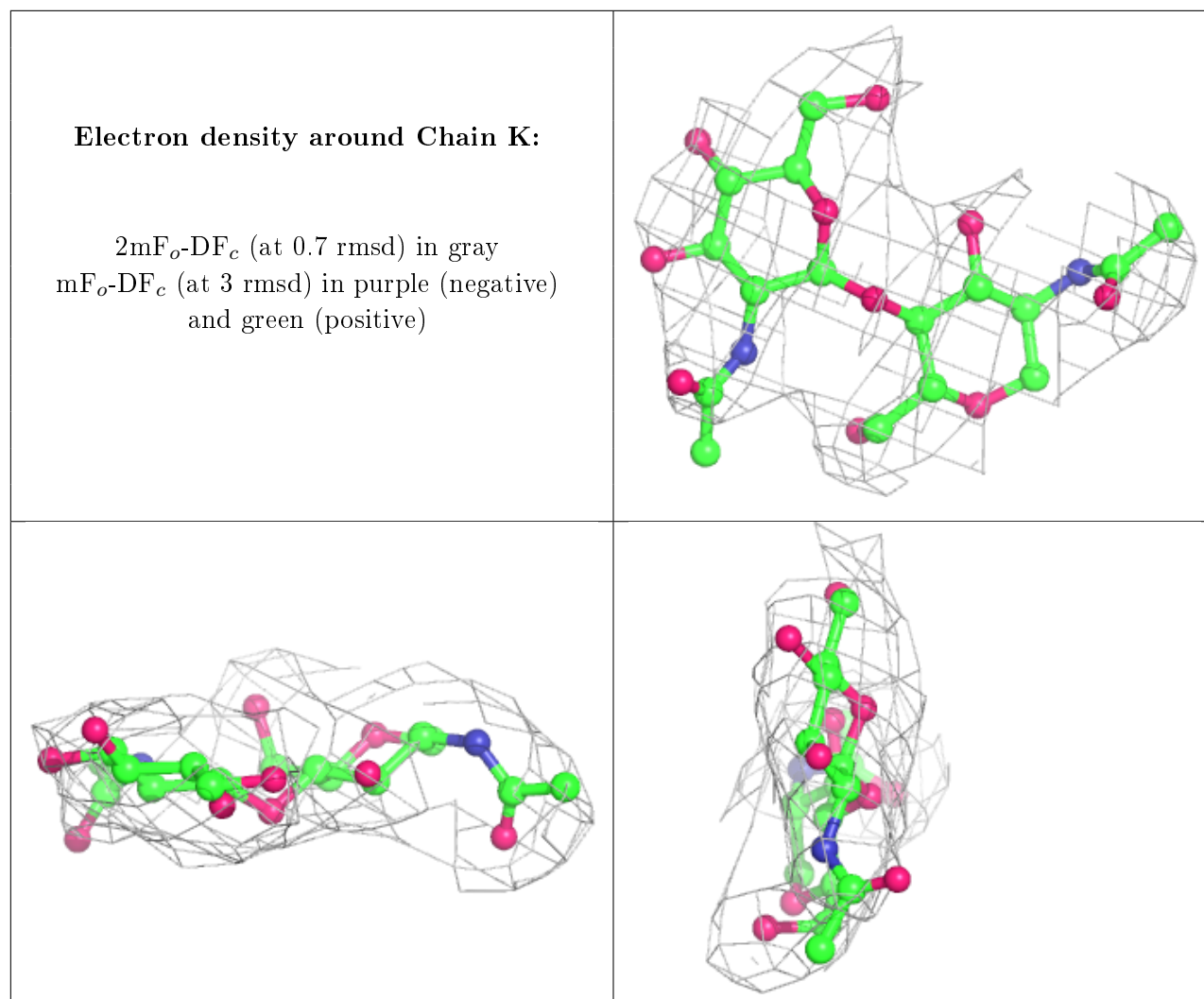
**Electron density around Chain I:**

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

**Electron density around Chain J:**

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





## 6.4 Ligands ⓘ

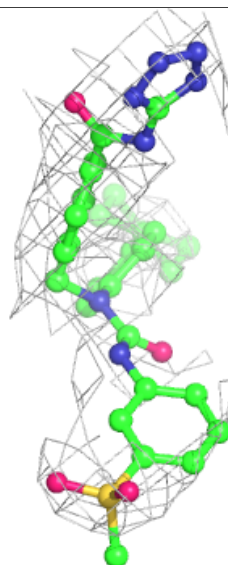
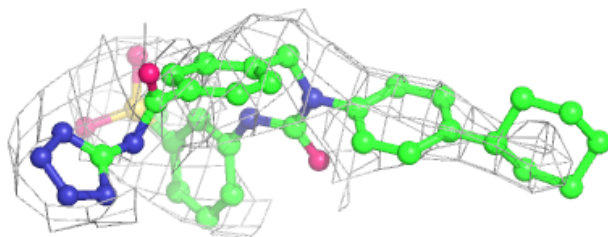
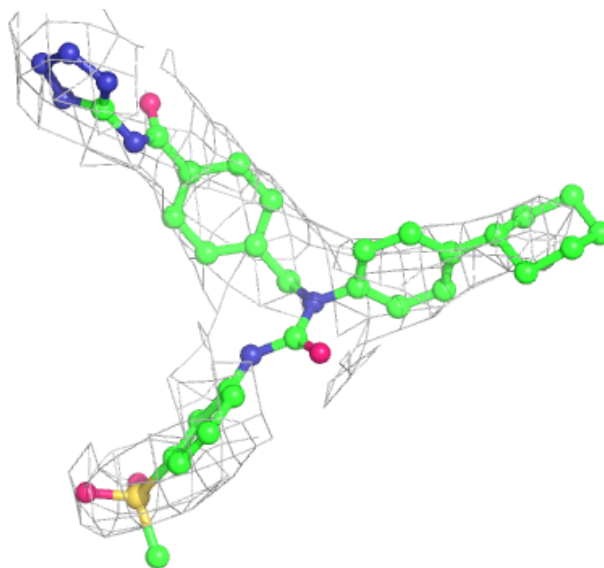
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
6	97V	A	1207	41/41	0.86	0.26	155,176,191,195	0
7	NAG	B	1206	14/15	0.90	0.17	123,131,137,149	0
6	97V	B	1209	41/41	0.91	0.23	157,173,187,188	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.

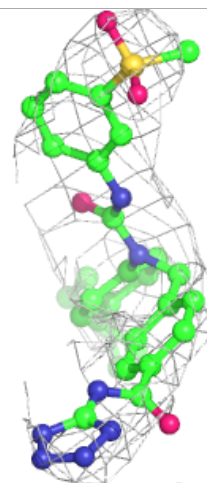
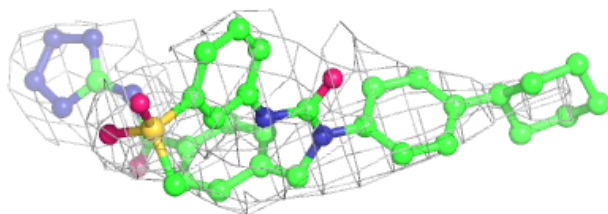
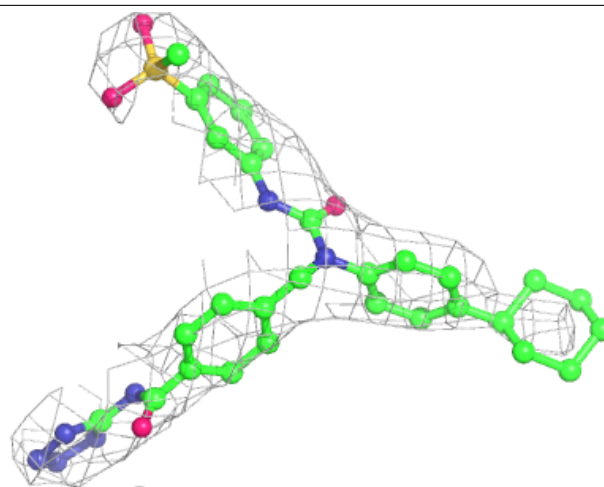
**Electron density around 97V A 1207:**

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



**Electron density around 97V B 1209:**

$2mF_o-DF_c$  (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.