



# Full wwPDB X-ray Structure Validation Report ⓘ

Apr 4, 2022 – 10:06 AM EDT

PDB ID : 7MFH  
Title : Crystal structure of BIO-32546 bound mouse Autotaxin  
Authors : Chodaparambil, J.V.  
Deposited on : 2021-04-09  
Resolution : 2.30 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 : 2.27  
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) : 2.27

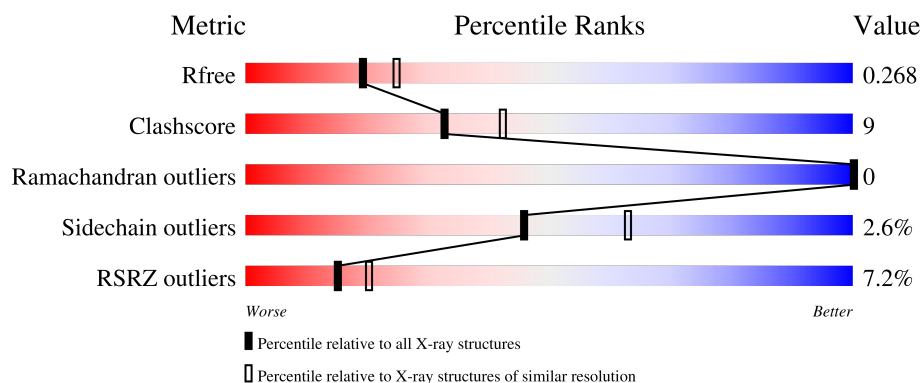
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	832	<div> <div>7%</div> <div>75%</div> <div>17%</div> <div>• 7%</div> </div>
2	B	6	<div> <div>50%</div> <div>50%</div> </div>
3	C	2	<div> <div>50%</div> <div>50%</div> </div>

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
6	EDO	A	914	-	-	X	-

## 2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 6618 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.

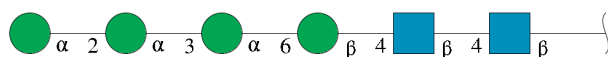
- Molecule 1 is a protein called Ectonucleotide pyrophosphatase/phosphodiesterase family member 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	776	Total	C	N	O	S	0	0	0
			6237	3956	1074	1159	48			

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	56	GLY	-	insertion	UNP Q9R1E6
A	57	SER	-	insertion	UNP Q9R1E6
A	?	-	GLY	deletion	UNP Q9R1E6
A	?	-	LYS	deletion	UNP Q9R1E6
A	?	-	VAL	deletion	UNP Q9R1E6
A	?	-	GLU	deletion	UNP Q9R1E6
A	?	-	PRO	deletion	UNP Q9R1E6
A	863	SER	-	expression tag	UNP Q9R1E6
A	864	ARG	-	expression tag	UNP Q9R1E6
A	865	GLU	-	expression tag	UNP Q9R1E6
A	866	ASN	-	expression tag	UNP Q9R1E6
A	867	LEU	-	expression tag	UNP Q9R1E6
A	868	TYR	-	expression tag	UNP Q9R1E6
A	869	PHE	-	expression tag	UNP Q9R1E6
A	870	GLN	-	expression tag	UNP Q9R1E6

- Molecule 2 is an oligosaccharide called alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-alpha-D-mannopyranose-(1-6)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



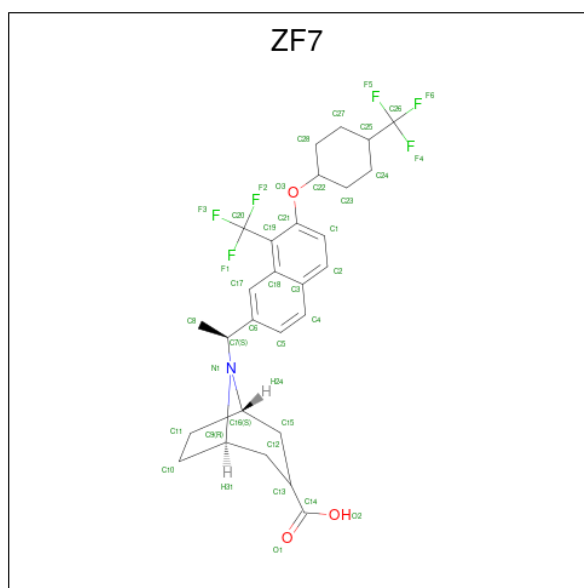
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	6	Total	C	N	O	0	0	0
			72	40	2	30			

- Molecule 3 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
3	C	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 4 is (1R,3S,5S)-8-{(1S)-1-[8-(trifluoromethyl)-7-{[(1s,4R)-4-(trifluoromethyl)cyclohexyl]oxy}naphthalen-2-yl]ethyl}-8-azabicyclo[3.2.1]octane-3-carboxylic acid (three-letter code: ZF7) (formula: C<sub>28</sub>H<sub>31</sub>F<sub>6</sub>NO<sub>3</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	F	N	O	0	0
			38	28	6	1	3		

- Molecule 5 is IODIDE ION (three-letter code: IOD) (formula: I).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	I	0	0
			1	1		

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



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

*Continued on next page...*

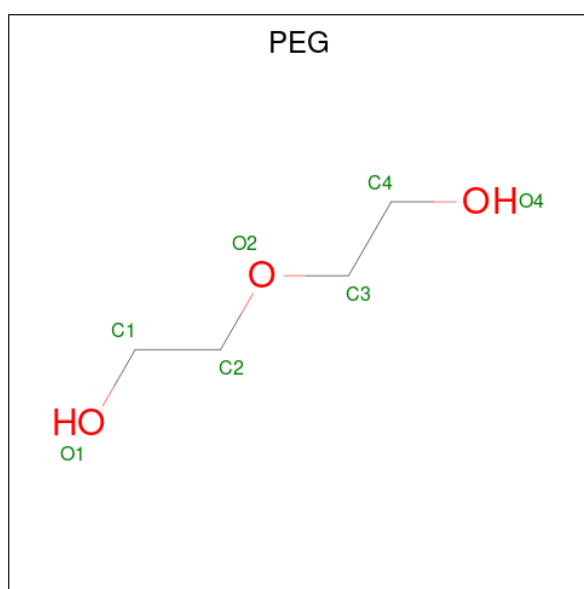
*Continued from previous page...*

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

- Molecule 7 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	7	Total	Zn	0	0
			7	7		

- Molecule 8 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			7	4	3		

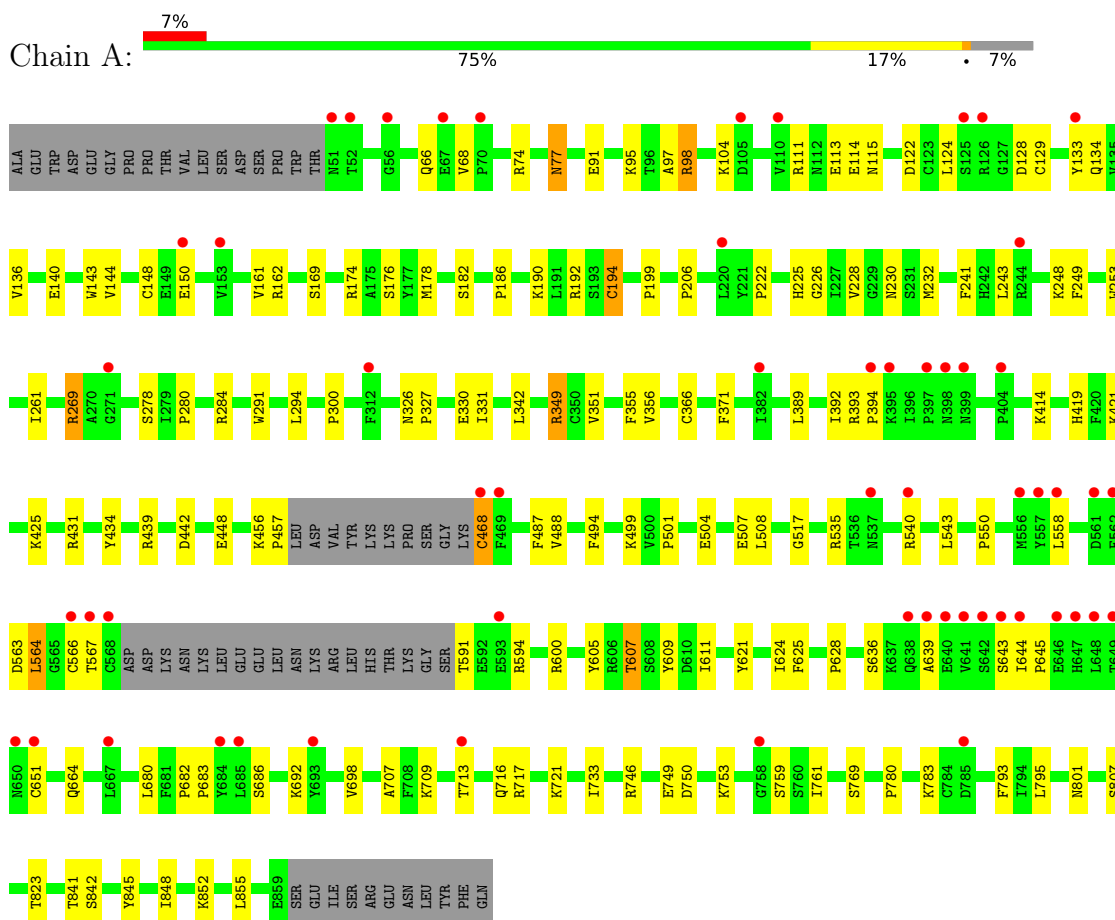
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	172	Total	O	0	0
			172	172		

### 3 Residue-property plots

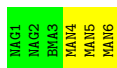
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Ectonucleotide pyrophosphatase/phosphodiesterase family member 2



- Molecule 2: alpha-D-mannopyranose-(1-2)-alpha-D-mannopyranose-(1-3)-alpha-D-mannopyranose-(1-6)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B:



- Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain C:  50% 50%

MAG1  
MAG2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	61.02Å 61.62Å 67.32Å 89.07° 74.20° 79.88°	Depositor
Resolution (Å)	46.13 – 2.30 46.13 – 2.30	Depositor EDS
% Data completeness (in resolution range)	98.0 (46.13-2.30) 98.0 (46.13-2.30)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.46 (at 2.29Å)	Xtriage
Refinement program	PHENIX 1.19.1_4122	Depositor
R, $R_{free}$	0.215 , 0.268 0.215 , 0.268	Depositor DCC
$R_{free}$ test set	2024 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	45.8	Xtriage
Anisotropy	0.416	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 45.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	6618	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.90% 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: NAG, BMA, IOD, EDO, ZN, PEG, MAN, ZF7

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	0.26	0/6415	0.50	0/8710

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	6237	0	5947	107	0
2	B	72	0	61	0	0
3	C	28	0	25	0	0
4	A	38	0	0	1	0
5	A	1	0	0	0	0
6	A	56	0	81	13	0
7	A	7	0	0	0	0
8	A	7	0	10	1	0
9	A	172	0	0	15	0
All	All	6618	0	6124	109	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 9.

All (109) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:494:PHE:O	9:A:1001:HOH:O	1.89	0.90
1:A:393:ARG:HG2	1:A:394:PRO:HD2	1.55	0.88
1:A:104:LYS:NZ	1:A:136:VAL:O	2.07	0.87
1:A:707:ALA:HA	8:A:910:PEG:H22	1.67	0.76
1:A:746:ARG:HD2	6:A:906:EDO:H21	1.67	0.76
1:A:161:VAL:HG13	1:A:162:ARG:HG3	1.67	0.75
1:A:456:LYS:HD2	1:A:457:PRO:HD2	1.70	0.74
1:A:232:MET:HG3	1:A:241:PHE:HB3	1.69	0.72
1:A:504:GLU:OE1	9:A:1002:HOH:O	2.08	0.71
1:A:162:ARG:NH2	1:A:517:GLY:HA3	2.05	0.70
1:A:68:VAL:HG13	1:A:74:ARG:HG2	1.74	0.70
1:A:150:GLU:OE2	1:A:535:ARG:NH2	2.23	0.70
1:A:129:CYS:H	6:A:916:EDO:H22	1.56	0.69
1:A:639:ALA:HB3	1:A:721:LYS:HD3	1.74	0.69
1:A:793:PHE:CD1	1:A:795:LEU:HG	2.28	0.69
1:A:841:THR:OG1	9:A:1003:HOH:O	2.11	0.68
1:A:564:LEU:H	1:A:564:LEU:HD23	1.57	0.67
1:A:698:VAL:HG22	9:A:1009:HOH:O	1.95	0.66
1:A:77:ASN:HD21	1:A:291:TRP:HH2	1.43	0.65
1:A:414:LYS:NZ	9:A:1013:HOH:O	2.29	0.61
1:A:504:GLU:OE2	6:A:914:EDO:O1	2.17	0.61
1:A:111:ARG:NH1	1:A:122:ASP:OD1	2.35	0.59
1:A:682:PRO:HB3	1:A:716:GLN:HB3	1.85	0.58
1:A:807:SER:OG	6:A:909:EDO:O1	2.22	0.58
1:A:713:THR:HG22	1:A:717:ARG:HD3	1.86	0.57
1:A:769:SER:HB2	1:A:793:PHE:CE2	2.40	0.57
1:A:66:GLN:OE1	9:A:1005:HOH:O	2.17	0.57
1:A:564:LEU:HD12	1:A:566:CYS:HB2	1.86	0.57
1:A:501:PRO:O	9:A:1004:HOH:O	2.17	0.56
1:A:753:LYS:NZ	9:A:1024:HOH:O	2.38	0.56
1:A:468:CYS:N	9:A:1023:HOH:O	2.37	0.56
1:A:801:ASN:O	6:A:909:EDO:O1	2.25	0.55
1:A:550:PRO:HB3	1:A:609:TYR:CZ	2.42	0.55
1:A:664:GLN:HG2	1:A:761:ILE:HG12	1.89	0.55
1:A:97:ALA:HB3	1:A:115:ASN:HA	1.88	0.54
1:A:230:ASN:ND2	9:A:1026:HOH:O	2.39	0.54
1:A:134:GLN:HE21	1:A:140:GLU:HG3	1.72	0.54
1:A:392:ILE:O	9:A:1006:HOH:O	2.18	0.54
1:A:625:PHE:CZ	6:A:906:EDO:H12	2.43	0.54
1:A:91:GLU:O	1:A:95:LYS:NZ	2.37	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:261:ILE:HD12	6:A:917:EDO:H21	1.88	0.53
1:A:128:ASP:OD1	1:A:128:ASP:N	2.40	0.53
1:A:222:PRO:HA	1:A:225:HIS:CE1	2.43	0.53
1:A:349:ARG:NH2	9:A:1022:HOH:O	2.37	0.52
1:A:563:ASP:OD1	1:A:563:ASP:N	2.41	0.51
1:A:95:LYS:HE2	1:A:114:GLU:HA	1.92	0.51
1:A:124:LEU:HD11	1:A:133:TYR:CE2	2.46	0.50
1:A:253:TRP:HA	1:A:439:ARG:HH22	1.76	0.50
1:A:550:PRO:HB2	1:A:611:ILE:HG13	1.93	0.50
1:A:594:ARG:NH2	1:A:636:SER:OG	2.44	0.50
1:A:226:GLY:O	1:A:232:MET:HE1	2.12	0.50
1:A:124:LEU:H	1:A:124:LEU:HD12	1.77	0.49
1:A:291:TRP:HA	1:A:294:LEU:HG	1.94	0.49
1:A:686:SER:OG	1:A:692:LYS:NZ	2.39	0.48
1:A:104:LYS:HD3	1:A:104:LYS:HA	1.63	0.48
1:A:355:PHE:HB3	1:A:488:VAL:HB	1.96	0.47
1:A:98:ARG:HG3	1:A:98:ARG:HH11	1.79	0.47
1:A:607:THR:HG22	1:A:624:ILE:HD13	1.96	0.47
1:A:780:PRO:HG2	1:A:783:LYS:HB2	1.97	0.47
6:A:914:EDO:O2	9:A:1002:HOH:O	2.20	0.47
1:A:95:LYS:HE2	1:A:113:GLU:O	2.14	0.47
1:A:823:THR:HG23	1:A:855:LEU:HD11	1.97	0.47
1:A:371:PHE:CG	1:A:457:PRO:HD3	2.50	0.46
1:A:143:TRP:O	1:A:190:LYS:NZ	2.44	0.46
1:A:269:ARG:HB3	1:A:300:PRO:HA	1.98	0.46
1:A:148:CYS:HB3	1:A:194:CYS:HB3	1.80	0.46
1:A:487:PHE:CE1	1:A:508:LEU:HD21	2.51	0.46
1:A:144:VAL:HG21	1:A:186:PRO:HB2	1.99	0.45
1:A:326:ASN:O	1:A:330:GLU:HG2	2.16	0.45
1:A:243:LEU:O	1:A:248:LYS:NZ	2.29	0.45
1:A:199:PRO:HG3	1:A:499:LYS:HE3	1.99	0.45
1:A:643:SER:O	1:A:645:PRO:HD3	2.17	0.45
1:A:419:HIS:HA	1:A:448:GLU:OE1	2.16	0.45
1:A:206:PRO:HD3	1:A:434:TYR:CE1	2.52	0.44
1:A:421:LYS:NZ	9:A:1038:HOH:O	2.49	0.44
1:A:178:MET:HE2	1:A:192:ARG:HD3	1.98	0.44
1:A:793:PHE:CE1	1:A:795:LEU:HG	2.51	0.44
1:A:543:LEU:HD12	1:A:845:TYR:CE1	2.53	0.44
1:A:567:THR:O	1:A:651:CYS:HA	2.18	0.44
1:A:77:ASN:OD1	1:A:77:ASN:N	2.51	0.43
1:A:605:TYR:CE2	6:A:906:EDO:H22	2.53	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:507:GLU:OE2	9:A:1007:HOH:O	2.21	0.43
1:A:124:LEU:HD11	1:A:133:TYR:HE2	1.83	0.43
1:A:431:ARG:HA	6:A:914:EDO:H22	2.00	0.43
1:A:431:ARG:HB2	6:A:914:EDO:H11	2.00	0.43
1:A:129:CYS:HB2	6:A:916:EDO:H11	2.01	0.43
1:A:366:CYS:N	1:A:468:CYS:SG	2.91	0.42
1:A:169:SER:HA	1:A:356:VAL:O	2.19	0.42
1:A:733:ILE:O	1:A:769:SER:HA	2.19	0.42
1:A:232:MET:CE	1:A:439:ARG:HD3	2.49	0.42
1:A:327:PRO:O	1:A:331:ILE:HG13	2.20	0.42
1:A:228:VAL:HG22	1:A:232:MET:HE3	2.01	0.42
1:A:564:LEU:HD12	1:A:566:CYS:CB	2.50	0.42
1:A:621:TYR:HA	1:A:628:PRO:HA	2.02	0.42
1:A:848:ILE:HG22	1:A:852:LYS:HD2	2.02	0.42
1:A:680:LEU:HD23	1:A:680:LEU:HA	1.89	0.42
4:A:901:ZF7:O3	4:A:901:ZF7:F1	2.26	0.41
1:A:644:ILE:HG13	1:A:717:ARG:HH21	1.85	0.41
1:A:280:PRO:O	1:A:284:ARG:HG3	2.21	0.41
1:A:550:PRO:HD3	1:A:600:ARG:NH1	2.36	0.41
1:A:342:LEU:HD13	1:A:351:VAL:HG11	2.03	0.41
1:A:558:LEU:HD12	1:A:558:LEU:HA	1.45	0.41
1:A:644:ILE:HG12	1:A:683:PRO:HB2	2.02	0.41
1:A:178:MET:HE2	1:A:178:MET:HB3	1.87	0.40
1:A:749:GLU:H	6:A:905:EDO:H21	1.86	0.40
1:A:550:PRO:HG3	1:A:600:ARG:HH11	1.86	0.40
1:A:248:LYS:HE2	1:A:249:PHE:CZ	2.57	0.40
1:A:425:LYS:HD3	1:A:442:ASP:HA	2.04	0.40
1:A:709:LYS:O	1:A:713:THR:OG1	2.23	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	768/832 (92%)	754 (98%)	14 (2%)	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	696/757 (92%)	678 (97%)	18 (3%)	46	63

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	77	ASN
1	A	98	ARG
1	A	174	ARG
1	A	176	SER
1	A	182	SER
1	A	194	CYS
1	A	269	ARG
1	A	278	SER
1	A	349	ARG
1	A	389	LEU
1	A	468	CYS
1	A	540	ARG
1	A	564	LEU
1	A	591	THR
1	A	607	THR
1	A	750	ASP
1	A	759	SER
1	A	842	SER

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

Mol	Chain	Res	Type
1	A	118	HIS
1	A	134	GLN
1	A	831	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 ⓘ

8 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	NAG	B	1	1,2	14,14,15	0.44	0	17,19,21	0.55	0
2	NAG	B	2	2	14,14,15	0.40	0	17,19,21	0.45	0
2	BMA	B	3	2	11,11,12	0.65	0	15,15,17	0.78	0
2	MAN	B	4	2	11,11,12	0.66	0	15,15,17	1.07	2 (13%)
2	MAN	B	5	2	11,11,12	0.70	0	15,15,17	1.03	2 (13%)
2	MAN	B	6	2	11,11,12	0.56	0	15,15,17	1.09	2 (13%)
3	NAG	C	1	1,3	14,14,15	0.32	0	17,19,21	0.44	0
3	NAG	C	2	3	14,14,15	1.03	1 (7%)	17,19,21	1.10	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
2	NAG	B	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	B	2	2	-	0/6/23/26	0/1/1/1
2	BMA	B	3	2	-	2/2/19/22	0/1/1/1
2	MAN	B	4	2	-	0/2/19/22	0/1/1/1
2	MAN	B	5	2	-	0/2/19/22	0/1/1/1
2	MAN	B	6	2	-	0/2/19/22	0/1/1/1
3	NAG	C	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	2	3	-	4/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	2	NAG	O5-C1	-3.21	1.38	1.43

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	6	MAN	C1-O5-C5	3.11	116.40	112.19
3	C	2	NAG	C4-C3-C2	2.77	115.07	111.02
2	B	4	MAN	C1-O5-C5	2.74	115.91	112.19
2	B	5	MAN	C1-O5-C5	2.46	115.53	112.19
2	B	5	MAN	O2-C2-C3	-2.37	105.40	110.14
2	B	6	MAN	O2-C2-C3	-2.29	105.54	110.14
2	B	4	MAN	O2-C2-C3	-2.27	105.59	110.14

There are no chirality outliers.

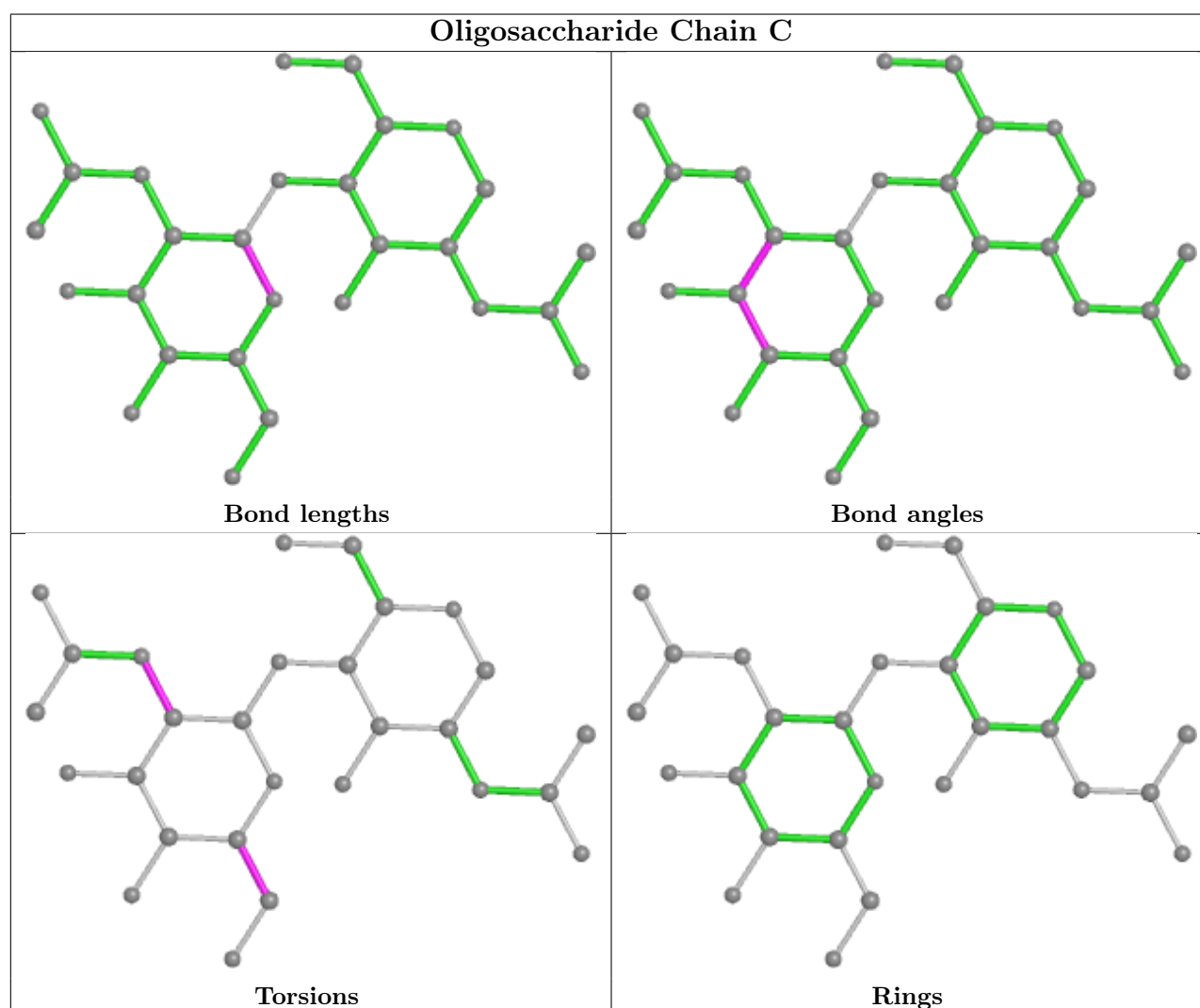
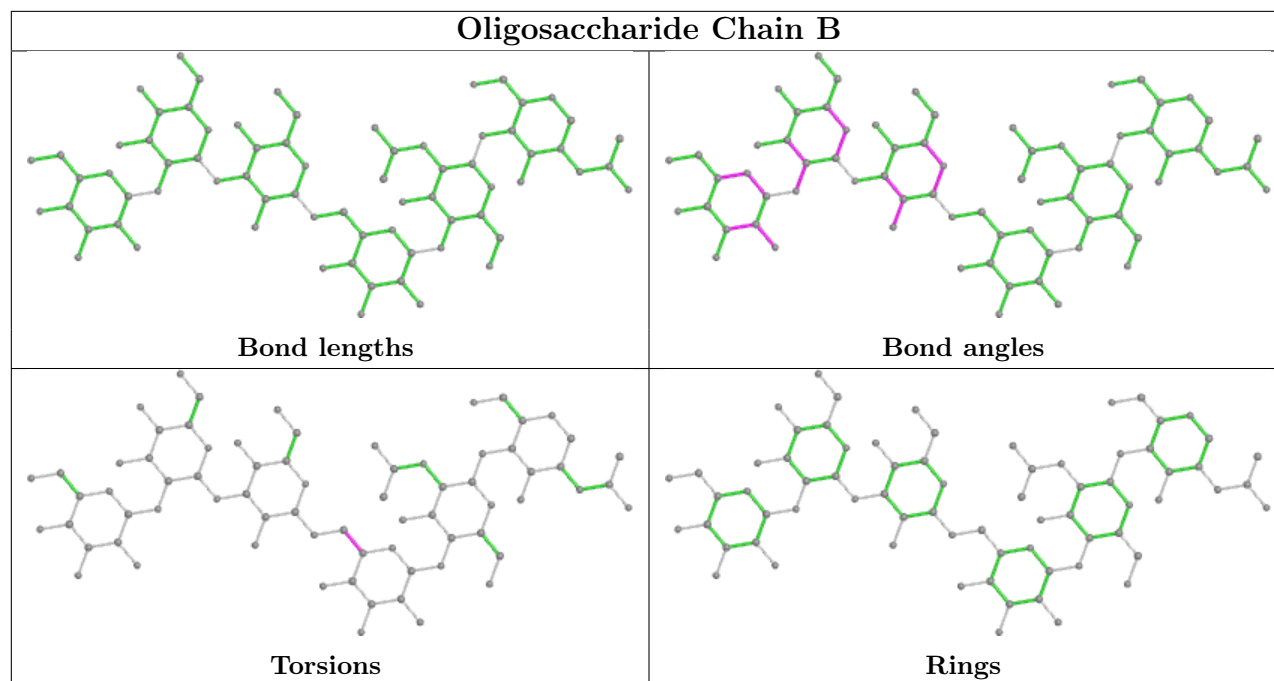
All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	2	NAG	O5-C5-C6-O6
3	C	2	NAG	C4-C5-C6-O6
2	B	3	BMA	O5-C5-C6-O6
3	C	2	NAG	C1-C2-N2-C7
2	B	3	BMA	C4-C5-C6-O6
3	C	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

Of 24 ligands modelled in this entry, 8 are monoatomic - leaving 16 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
6	EDO	A	915	-	3,3,3	0.59	0	2,2,2	0.39	0
6	EDO	A	923	-	3,3,3	0.44	0	2,2,2	0.40	0
6	EDO	A	907	-	3,3,3	0.46	0	2,2,2	0.32	0
6	EDO	A	924	-	3,3,3	0.47	0	2,2,2	0.37	0
6	EDO	A	908	-	3,3,3	0.44	0	2,2,2	0.35	0
6	EDO	A	916	-	3,3,3	0.32	0	2,2,2	0.84	0
6	EDO	A	917	-	3,3,3	0.48	0	2,2,2	0.27	0
6	EDO	A	905	-	3,3,3	0.45	0	2,2,2	0.35	0
6	EDO	A	906	-	3,3,3	0.45	0	2,2,2	0.21	0
6	EDO	A	914	-	3,3,3	0.46	0	2,2,2	0.06	0
6	EDO	A	918	-	3,3,3	0.46	0	2,2,2	0.40	0
6	EDO	A	903	-	3,3,3	0.47	0	2,2,2	0.35	0
8	PEG	A	910	-	6,6,6	0.09	0	5,5,5	0.12	0
4	ZF7	A	901	-	39,42,42	0.57	0	59,65,65	0.73	2 (3%)
6	EDO	A	909	-	3,3,3	0.45	0	2,2,2	0.39	0
6	EDO	A	911	-	3,3,3	0.45	0	2,2,2	0.34	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
6	EDO	A	915	-	-	0/1/1/1	-
6	EDO	A	923	-	-	0/1/1/1	-
6	EDO	A	907	-	-	0/1/1/1	-
6	EDO	A	924	-	-	0/1/1/1	-
6	EDO	A	908	-	-	0/1/1/1	-
6	EDO	A	916	-	-	1/1/1/1	-
6	EDO	A	917	-	-	0/1/1/1	-
6	EDO	A	905	-	-	0/1/1/1	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	EDO	A	906	-	-	0/1/1/1	-
6	EDO	A	914	-	-	0/1/1/1	-
6	EDO	A	918	-	-	0/1/1/1	-
6	EDO	A	903	-	-	0/1/1/1	-
8	PEG	A	910	-	-	1/4/4/4	-
4	ZF7	A	901	-	-	0/24/59/59	0/6/5/5
6	EDO	A	909	-	-	0/1/1/1	-
6	EDO	A	911	-	-	0/1/1/1	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	901	ZF7	C21-C19-C18	2.11	119.01	117.53
4	A	901	ZF7	O3-C21-C19	2.10	116.38	114.26

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
8	A	910	PEG	C1-C2-O2-C3
6	A	916	EDO	O1-C1-C2-O2

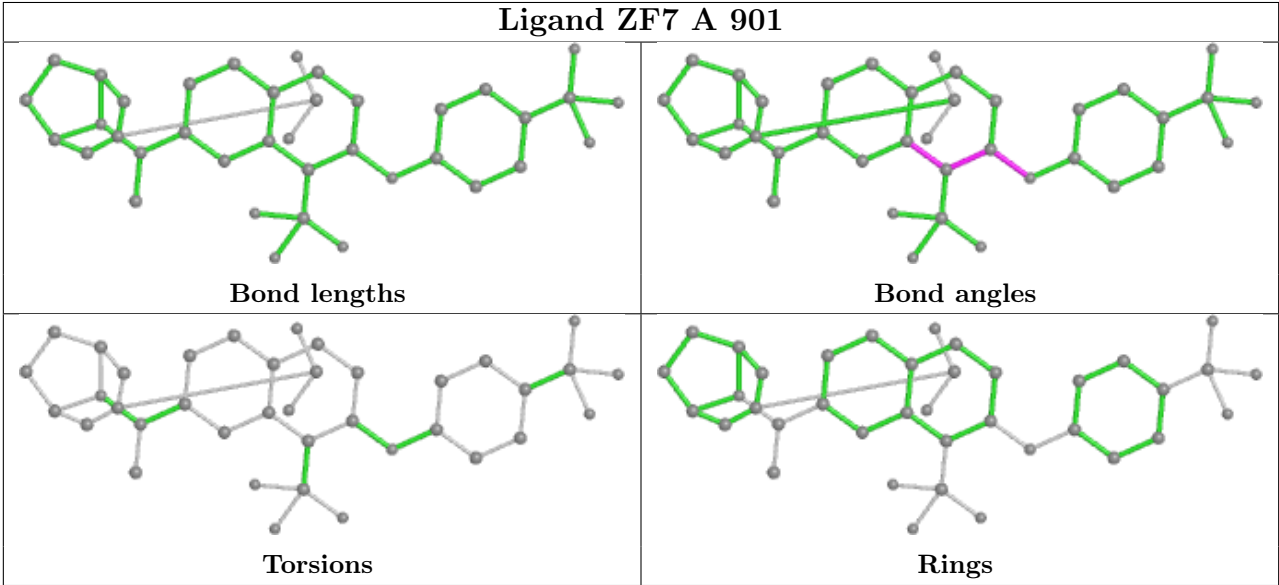
There are no ring outliers.

8 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	916	EDO	2	0
6	A	917	EDO	1	0
6	A	905	EDO	1	0
6	A	906	EDO	3	0
6	A	914	EDO	4	0
8	A	910	PEG	1	0
4	A	901	ZF7	1	0
6	A	909	EDO	2	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 ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	68:VAL	C	70:PRO	N	6.01

## 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<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	776/832 (93%)	0.61	56 (7%) 15 20	37, 57, 90, 115	0

All (56) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	641	VAL	8.2
1	A	642	SER	7.5
1	A	70	PRO	5.9
1	A	398	ASN	5.8
1	A	648	LEU	5.7
1	A	647	HIS	5.5
1	A	649	THR	4.7
1	A	562	PHE	4.5
1	A	693	TYR	4.3
1	A	271	GLY	4.2
1	A	758	GLY	4.1
1	A	566	CYS	4.0
1	A	561	ASP	4.0
1	A	638	GLN	3.9
1	A	651	CYS	3.9
1	A	639	ALA	3.8
1	A	568	CYS	3.7
1	A	468	CYS	3.7
1	A	537	ASN	3.6
1	A	650	ASN	3.6
1	A	558	LEU	3.6
1	A	556	MET	3.6
1	A	125	SER	3.6
1	A	469	PHE	3.4
1	A	244	ARG	3.4
1	A	110	VAL	3.2
1	A	126	ARG	3.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	153	VAL	3.1
1	A	640	GLU	3.1
1	A	713	THR	3.0
1	A	643	SER	3.0
1	A	67	GLU	2.9
1	A	382	ILE	2.8
1	A	56	GLY	2.8
1	A	646	GLU	2.8
1	A	557	TYR	2.8
1	A	312	PHE	2.7
1	A	785	ASP	2.7
1	A	105	ASP	2.6
1	A	397	PRO	2.5
1	A	684	TYR	2.5
1	A	667	LEU	2.4
1	A	540	ARG	2.4
1	A	404	PRO	2.4
1	A	51	ASN	2.3
1	A	150	GLU	2.3
1	A	685	LEU	2.2
1	A	593	GLU	2.2
1	A	399	ASN	2.2
1	A	644	ILE	2.2
1	A	220	LEU	2.2
1	A	395	LYS	2.1
1	A	52	THR	2.1
1	A	133	TYR	2.1
1	A	567	THR	2.1
1	A	394	PRO	2.1

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

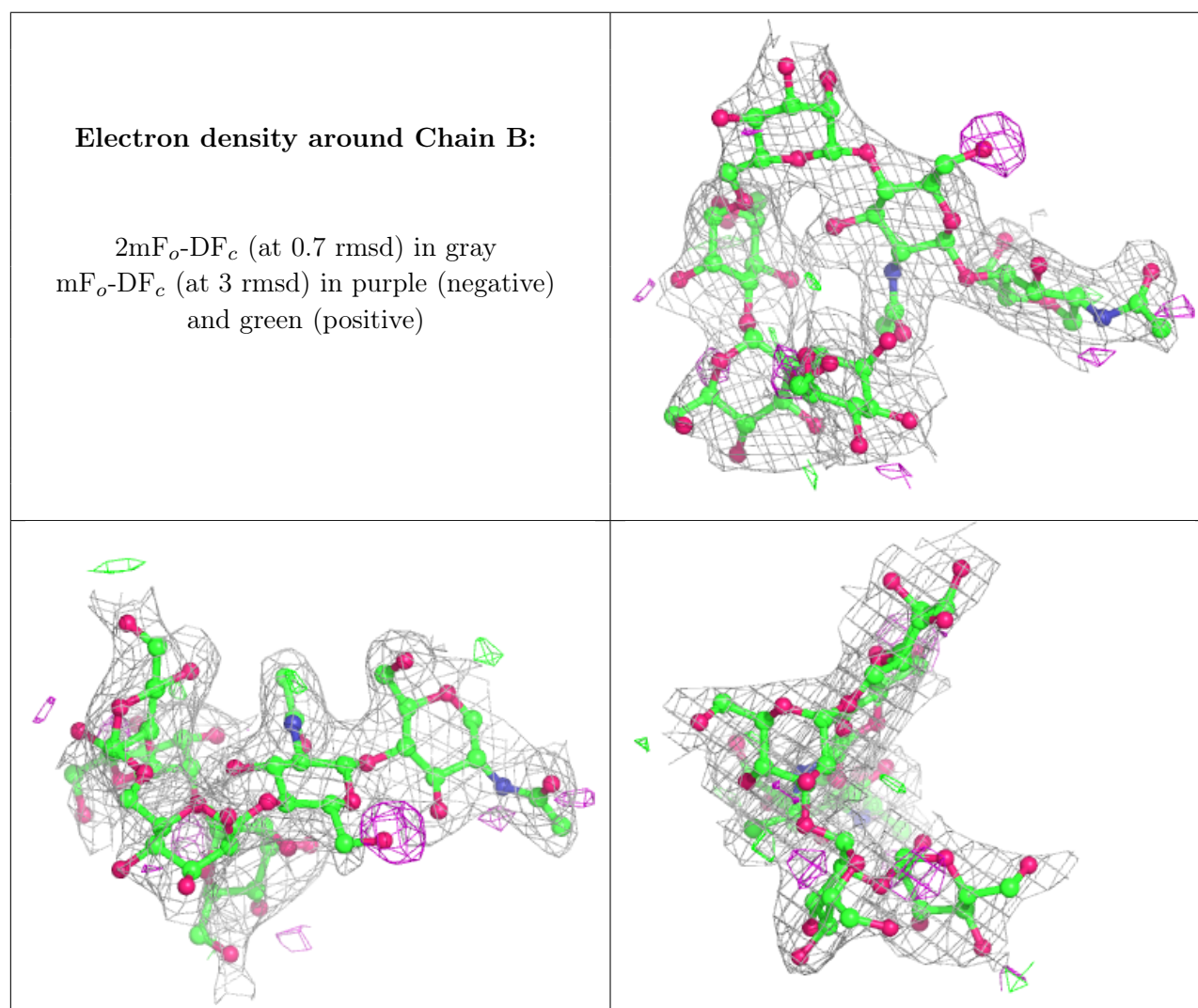
There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

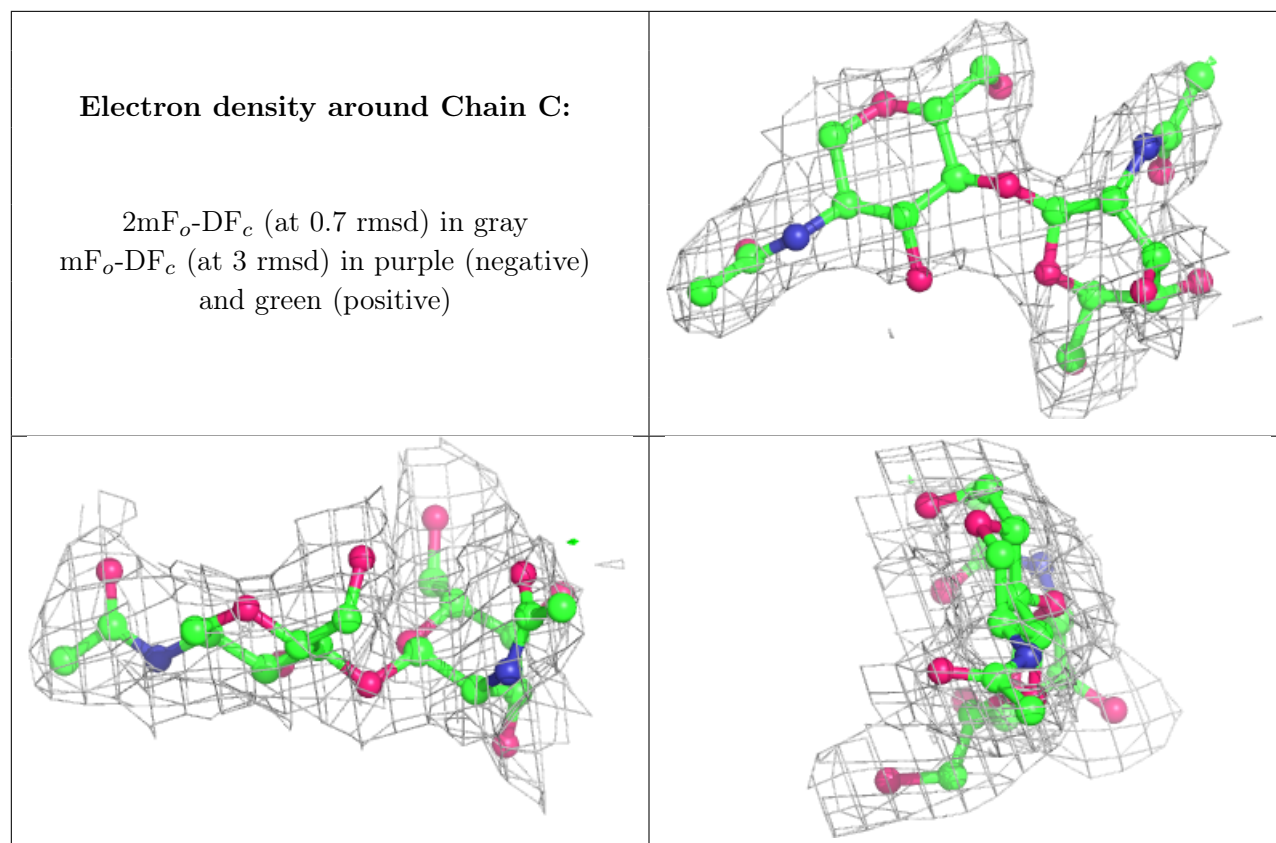
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
3	NAG	C	2	14/15	0.62	0.37	84,99,103,105	0
2	MAN	B	5	11/12	0.68	0.33	72,81,89,91	0
2	MAN	B	6	11/12	0.77	0.30	70,78,99,102	0
2	BMA	B	3	11/12	0.78	0.28	64,69,73,81	0
2	MAN	B	4	11/12	0.80	0.16	53,71,74,88	0
3	NAG	C	1	14/15	0.83	0.27	51,74,86,97	0
2	NAG	B	2	14/15	0.92	0.18	31,50,65,93	0
2	NAG	B	1	14/15	0.95	0.17	36,44,48,50	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.

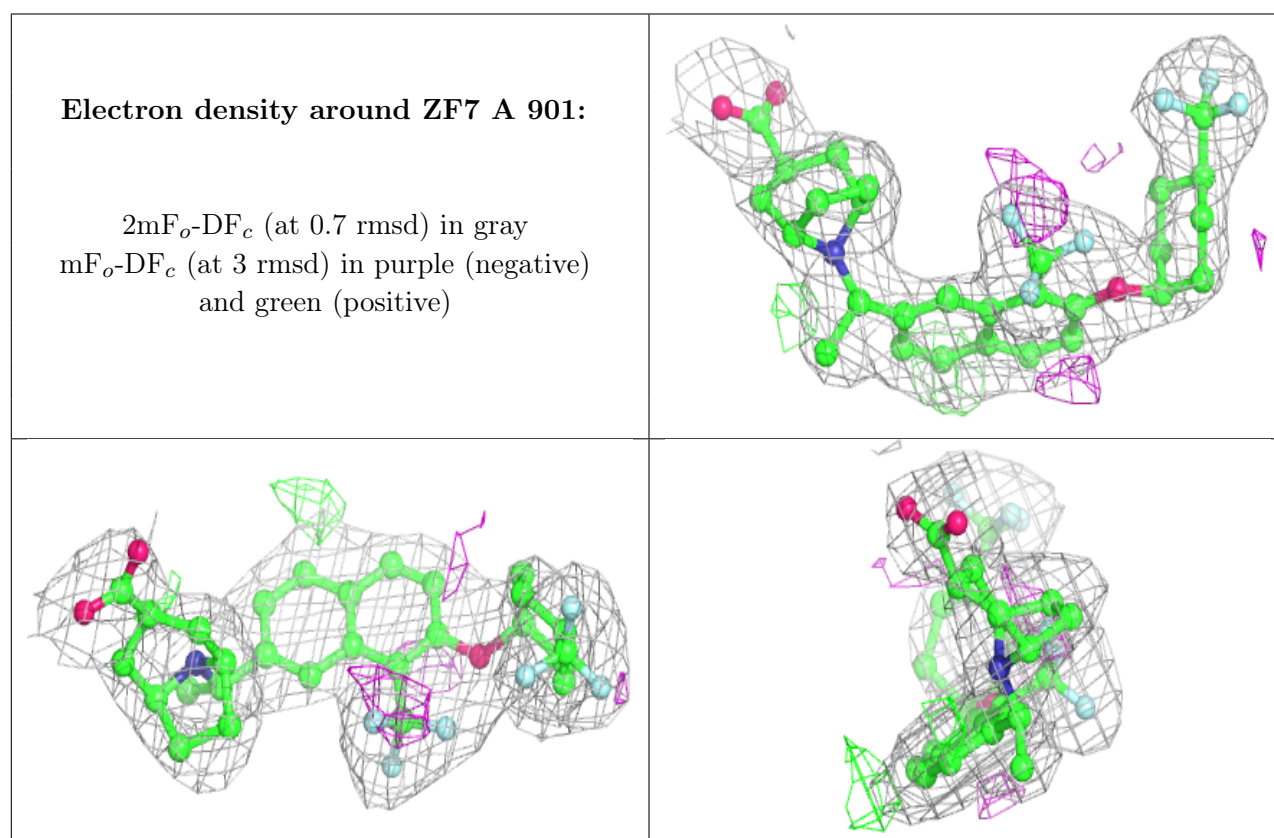
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
7	ZN	A	919	1/1	0.55	0.20	90,90,90,90	0
6	EDO	A	906	4/4	0.58	0.30	52,56,62,67	0
6	EDO	A	915	4/4	0.61	0.27	66,72,77,86	0
6	EDO	A	917	4/4	0.82	0.24	52,55,79,79	0
6	EDO	A	924	4/4	0.86	0.23	90,91,96,97	0
6	EDO	A	914	4/4	0.88	0.26	54,55,55,60	0
6	EDO	A	916	4/4	0.90	0.36	97,102,103,108	0
6	EDO	A	918	4/4	0.91	0.21	37,37,54,71	0
6	EDO	A	923	4/4	0.91	0.29	44,49,52,72	0
5	IOD	A	902	1/1	0.91	0.09	100,100,100,100	0
6	EDO	A	908	4/4	0.91	0.20	39,48,50,68	0
4	ZF7	A	901	38/38	0.92	0.17	35,47,62,71	0
8	PEG	A	910	7/7	0.92	0.21	27,49,63,65	0
7	ZN	A	913	1/1	0.94	0.06	68,68,68,68	0

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
6	EDO	A	911	4/4	0.94	0.17	37,48,56,63	0
6	EDO	A	907	4/4	0.94	0.17	40,57,60,67	0
6	EDO	A	909	4/4	0.95	0.32	47,53,60,75	0
6	EDO	A	903	4/4	0.95	0.31	33,36,43,53	0
7	ZN	A	921	1/1	0.97	0.13	63,63,63,63	0
7	ZN	A	912	1/1	0.97	0.18	50,50,50,50	0
7	ZN	A	920	1/1	0.98	0.19	44,44,44,44	0
7	ZN	A	904	1/1	0.98	0.08	66,66,66,66	0
6	EDO	A	905	4/4	0.98	0.26	26,54,59,61	0
7	ZN	A	922	1/1	0.99	0.27	27,27,27,27	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.



## 6.5 Other polymers ⓘ

There are no such residues in this entry.