Full wwPDB X-ray Structure Validation Report ⓘ

May 31, 2022 – 01:39 am BST

PDB ID	:	7P1N
Title	:	Crystal structure of human acetylcholinesterase in complex with (2R,3R,4S,5S,6R)-2-{4-[1-(4-{5-hydroxy-6-[(E)-(hydroxyimino)methyl]pyridin-2-yl}butyl)-1H-1,2,3-triazol-4-yl]butoxy}-6-(hydroxymethyl)oxane-3,4,5-triol oxime
Authors	:	Da Silva, O.; Dias, J.; Nachon, F.
Deposited on	:	2021-07-02
Resolution	:	2.95 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see references ⓘ) were used in the production of this report:

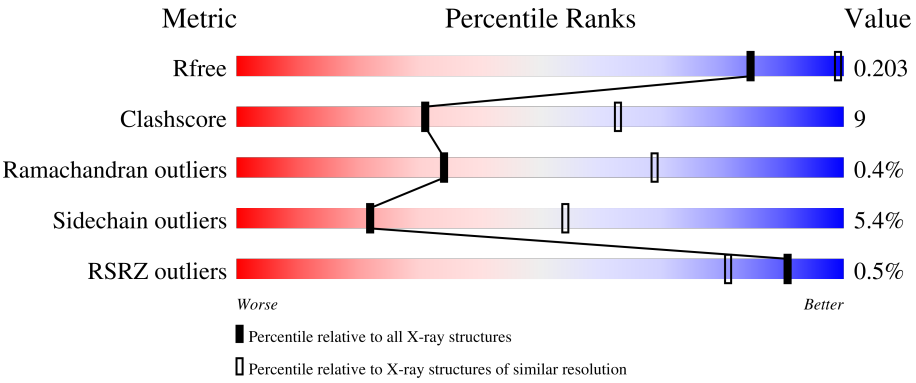
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.28.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.28.1

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.95 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	3104 (3.00-2.92)
Clashscore	141614	3462 (3.00-2.92)
Ramachandran outliers	138981	3340 (3.00-2.92)
Sidechain outliers	138945	3343 (3.00-2.92)
RSRZ outliers	127900	2986 (3.00-2.92)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	aa	257	95%
1	bb	257	% 93% 6%
2	A	282	80% 18%
2	B	282	77% 22%
3	C	4	50% 50%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	D	3	67% 33%
5	E	3	33% 67%

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	4J1	aa	601	-	-	-	X

2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 8831 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.

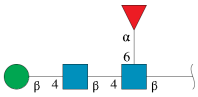
- Molecule 1 is a protein called Acetylcholinesterase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	aa	254	Total	C	N	O	S	0	0	0
			1936	1241	338	349	8			
1	bb	254	Total	C	N	O	S	0	0	0
			1936	1241	338	349	8			

- Molecule 2 is a protein called Acetylcholinesterase.

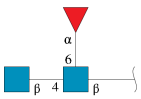
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	282	Total	C	N	O	S	0	0	0
			2229	1431	390	403	5			
2	B	279	Total	C	N	O	S	0	0	0
			2214	1423	387	399	5			

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



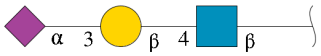
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	4	Total	C	N	O	0	0	0
			49	28	2	19			

- Molecule 4 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.



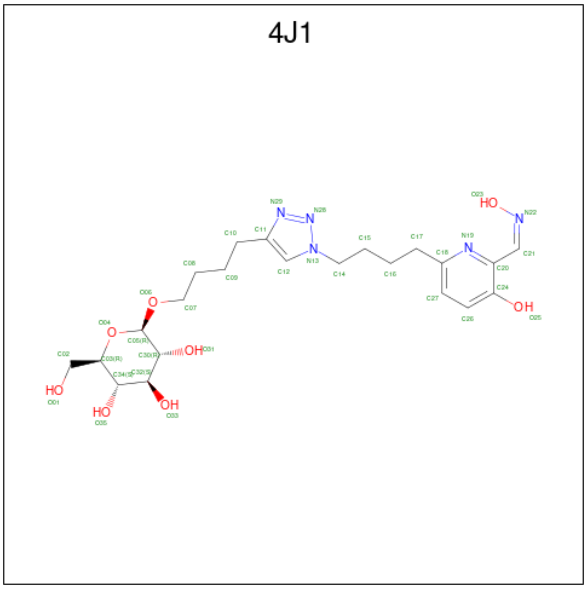
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	D	3	Total	C	N	O	0	0	0
			38	22	2	14			

- Molecule 5 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galacto pyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



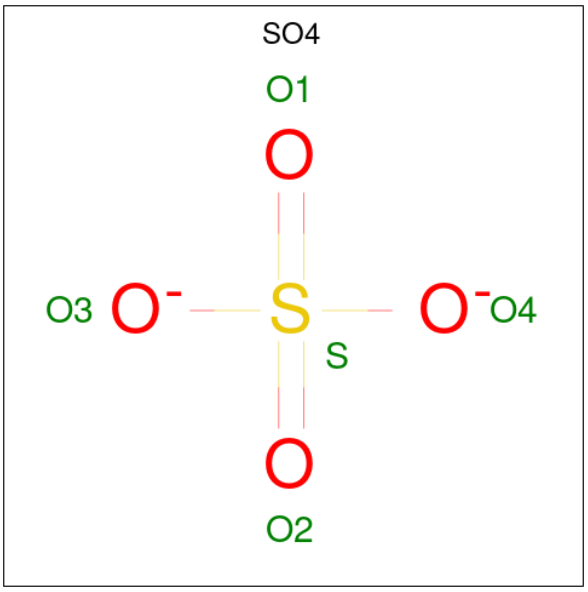
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	E	3	Total	C	N	O	0	0	0
			46	25	2	19			

- Molecule 6 is (2R,3R,4S,5S,6R)-2-[4-[1-[4-[6-[(Z)-hydroxyiminomethyl]-5-oxidanyl-pyridin-2-yl]butyl]-1,2,3-triazol-4-yl]butoxy]-6-(hydroxymethyl)oxane-3,4,5-triol (three-letter code: 4J1) (formula: $C_{22}H_{33}N_5O_8$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	aa	1	Total	C	N	O	0	0
			35	22	5	8		
6	aa	1	Total	C	N	O	0	0
			35	22	5	8		

- Molecule 7 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

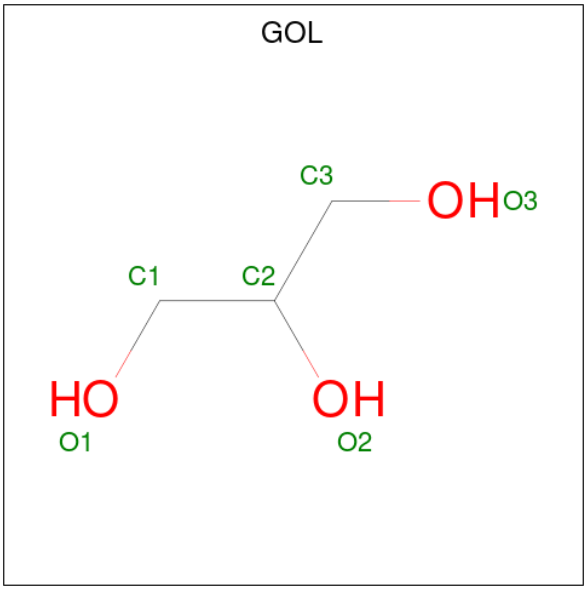


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	aa	1	Total O S 5 4 1	0	0
7	aa	1	Total O S 5 4 1	0	0
7	A	1	Total O S 5 4 1	0	0
7	A	1	Total O S 5 4 1	0	0
7	A	1	Total O S 5 4 1	0	0
7	A	1	Total O S 5 4 1	0	0
7	A	1	Total O S 5 4 1	0	0
7	A	1	Total O S 5 4 1	0	0
7	bb	1	Total O S 5 4 1	0	0
7	bb	1	Total O S 5 4 1	0	0
7	B	1	Total O S 5 4 1	0	0
7	B	1	Total O S 5 4 1	0	0
7	B	1	Total O S 5 4 1	0	0

- Molecule 8 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	aa	5	Total Cl 5 5	0	0
8	A	7	Total Cl 7 7	0	0
8	bb	5	Total Cl 5 5	0	0
8	B	9	Total Cl 9 9	0	0

- Molecule 9 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	aa	1	Total C O 6 3 3	0	0

- Molecule 10 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	A	2	Total Mg 2 2	0	0
10	bb	1	Total Mg 1 1	0	0
10	B	2	Total Mg 2 2	0	0

- Molecule 11 is water.

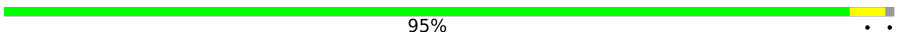
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	aa	60	Total O 60 60	0	0
11	A	55	Total O 55 55	0	0
11	bb	45	Total O 45 45	0	0
11	B	51	Total O 51 51	0	0

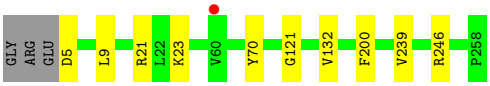
3 Residue-property plots (i)

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

• Molecule 1: Acetylcholinesterase

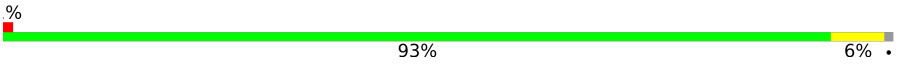
Chain aa:





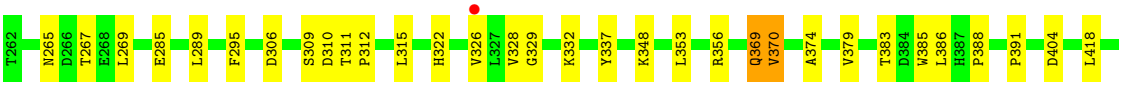
• Molecule 1: Acetylcholinesterase

Chain bb:



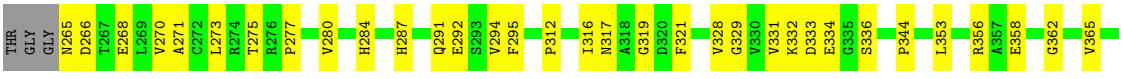
• Molecule 2: Acetylcholinesterase

Chain A:



• Molecule 2: Acetylcholinesterase

Chain B:



• Molecule 3: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

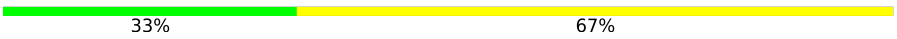
Chain C:

• Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain D:

• Molecule 5: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	212.29Å 212.29Å 116.08Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	78.33 – 2.95 183.85 – 2.95	Depositor EDS
% Data completeness (in resolution range)	99.9 (78.33-2.95) 94.6 (183.85-2.95)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.22 (at 2.96Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.172 , 0.205 0.171 , 0.203	Depositor DCC
R_{free} test set	3072 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å^2)	67.0	Xtriage
Anisotropy	0.788	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.021 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	8831	wwPDB-VP
Average B, all atoms (Å^2)	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.04% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GAL, CL, 4J1, GOL, BMA, SO4, FUC, NAG, SIA, MG

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	aa	0.46	0/1996	0.69	0/2729
1	bb	0.42	0/1996	0.69	0/2729
2	A	0.45	0/2296	0.64	0/3137
2	B	0.44	0/2281	0.66	0/3117
All	All	0.44	0/8569	0.67	0/11712

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	bb	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	bb	257	CYS	Peptide

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	aa	1936	0	1892	0	0
1	bb	1936	0	1892	0	0
2	A	2229	0	2155	37	0
2	B	2214	0	2142	31	0
3	C	49	0	43	0	0
4	D	38	0	34	0	0
5	E	46	0	40	0	0
6	aa	70	0	0	0	0
7	A	30	0	0	0	0
7	B	15	0	0	0	0
7	aa	10	0	0	0	0
7	bb	10	0	0	0	0
8	A	7	0	0	1	0
8	B	9	0	0	1	0
8	aa	5	0	0	0	0
8	bb	5	0	0	0	0
9	aa	6	0	8	0	0
10	A	2	0	0	0	0
10	B	2	0	0	0	0
10	bb	1	0	0	0	0
11	A	55	0	0	3	0
11	B	51	0	0	2	0
11	aa	60	0	0	0	0
11	bb	45	0	0	0	0
All	All	8831	0	8206	67	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 (67) 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:A:370:VAL:HG13	2:A:374:ALA:HB3	1.71	0.73
2:A:356:ARG:HH22	2:A:383:THR:HG21	1.53	0.72
2:A:452:GLU:HG3	11:A:704:HOH:O	1.94	0.66
2:B:370:VAL:HG13	2:B:374:ALA:HB3	1.79	0.65
2:A:456:GLY:HA3	2:A:477:MET:HE3	1.78	0.64
2:B:312:PRO:O	2:B:316:ILE:HG12	1.97	0.64
8:B:611:CL:CL	11:B:746:HOH:O	2.53	0.61
2:B:362:GLY:HA2	2:B:365:VAL:HG12	1.83	0.61
2:A:478:ARG:HG3	2:A:482:ASN:ND2	2.19	0.58
2:B:270:VAL:HA	2:B:273:LEU:HD12	1.85	0.58

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:452:GLU:HG3	11:B:702:HOH:O	2.04	0.56
2:B:490:ASN:HD21	2:B:500:TRP:H	1.53	0.56
2:A:328:VAL:O	2:A:427:ALA:HA	2.06	0.55
2:B:317:ASN:HA	2:B:421:GLN:HE22	1.70	0.55
2:B:284:HIS:HA	2:B:287:HIS:HD2	1.72	0.54
2:A:385:TRP:H	2:B:527:GLN:HE22	1.55	0.53
2:A:456:GLY:HA3	2:A:477:MET:CE	2.39	0.53
2:A:310:ASP:OD1	2:A:311:THR:N	2.34	0.52
2:A:356:ARG:HD2	2:A:388:PRO:O	2.10	0.52
2:A:326:VAL:HG21	2:A:418:LEU:HD13	1.92	0.51
2:A:329:GLY:HA3	2:A:428:TYR:CZ	2.45	0.51
2:A:404:ASP:OD1	2:A:525:ARG:NH1	2.43	0.51
2:A:475:ARG:NH1	2:A:479:TYR:OH	2.44	0.50
2:A:536:LEU:HB3	2:A:537:PRO:HD3	1.93	0.50
2:A:476:LEU:HA	2:A:479:TYR:HD1	1.77	0.49
2:B:317:ASN:HD22	2:B:417:ARG:CZ	2.26	0.49
2:A:285:GLU:HB2	11:A:734:HOH:O	2.12	0.48
2:B:470:LYS:O	2:B:474:GLN:HG3	2.13	0.48
2:B:319:GLY:HA3	2:B:321:PHE:CE2	2.48	0.48
2:B:265:ASN:HA	2:B:268:GLU:OE1	2.14	0.47
2:B:344:PRO:HB2	2:B:358:GLU:HG2	1.96	0.47
2:A:478:ARG:HG3	2:A:482:ASN:HD21	1.79	0.47
2:A:329:GLY:HA3	2:A:428:TYR:CE1	2.50	0.46
2:A:370:VAL:HG13	2:A:374:ALA:CB	2.44	0.46
2:B:458:PRO:HA	2:B:465:TYR:CD2	2.51	0.46
2:B:329:GLY:HA3	2:B:428:TYR:CZ	2.51	0.46
2:A:289:LEU:HD23	2:A:289:LEU:HA	1.76	0.46
2:A:337:TYR:HA	2:A:443:MET:CE	2.45	0.45
2:A:369:GLN:HE21	2:A:369:GLN:HB3	1.56	0.45
2:B:334:GLU:CD	2:B:407:VAL:HG11	2.37	0.45
2:A:379:VAL:O	2:A:383:THR:HG23	2.17	0.45
2:B:490:ASN:HD21	2:B:500:TRP:N	2.16	0.44
2:A:269:LEU:HD12	2:A:269:LEU:HA	1.77	0.44
2:A:458:PRO:HA	2:A:465:TYR:CD1	2.53	0.43
2:B:490:ASN:ND2	2:B:500:TRP:HB3	2.32	0.43
2:A:306:ASP:HB3	11:A:702:HOH:O	2.17	0.43
2:B:353:LEU:HB3	2:B:391:PRO:HB2	2.00	0.43
2:A:542:ALA:HB3	2:B:373:LEU:HD22	2.00	0.43
2:A:353:LEU:HB3	2:A:391:PRO:HB2	1.99	0.43
2:A:441:LEU:O	2:A:441:LEU:HD13	2.19	0.42
2:B:277:PRO:HG2	2:B:280:VAL:HG13	2.01	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:458:PRO:HG2	2:B:473:ALA:HB2	2.00	0.42
2:B:328:VAL:O	2:B:427:ALA:HA	2.20	0.41
2:B:271:ALA:O	2:B:275:THR:HG23	2.20	0.41
2:B:332:LYS:HD2	2:B:431:GLU:HB2	2.02	0.41
2:B:492:PRO:O	2:B:493:ARG:HB2	2.20	0.41
2:B:522:ARG:HE	2:B:522:ARG:HB3	1.70	0.41
2:B:333:ASP:OD2	2:B:444:GLY:HA3	2.20	0.41
2:A:476:LEU:HD21	2:A:513:LEU:HD13	2.03	0.41
2:A:524:LEU:HA	2:A:524:LEU:HD23	1.85	0.41
2:B:491:GLU:HA	2:B:492:PRO:HD3	1.90	0.41
2:A:312:PRO:HD2	8:A:611:CL:CL	2.58	0.41
2:A:459:LEU:HD12	2:A:459:LEU:HA	1.84	0.41
2:A:441:LEU:HD22	2:A:441:LEU:HA	1.93	0.40
2:B:331:VAL:HG22	2:B:334:GLU:OE2	2.21	0.40
2:A:310:ASP:HB3	2:A:315:LEU:HD13	2.04	0.40
2:A:322:HIS:HA	2:A:422:GLY:O	2.21	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	aa	252/257 (98%)	243 (96%)	8 (3%)	1 (0%)	34	69
1	bb	252/257 (98%)	240 (95%)	11 (4%)	1 (0%)	34	69
2	A	280/282 (99%)	264 (94%)	15 (5%)	1 (0%)	34	69
2	B	277/282 (98%)	257 (93%)	19 (7%)	1 (0%)	34	69
All	All	1061/1078 (98%)	1004 (95%)	53 (5%)	4 (0%)	34	69

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	265	ASN
1	bb	61	ASP
1	aa	121	GLY
2	B	291	GLN

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	aa	202/204 (99%)	193 (96%)	9 (4%)	27	61
1	bb	202/204 (99%)	188 (93%)	14 (7%)	15	44
2	A	232/232 (100%)	220 (95%)	12 (5%)	23	56
2	B	231/232 (100%)	219 (95%)	12 (5%)	23	56
All	All	867/872 (99%)	820 (95%)	47 (5%)	22	54

All (47) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	aa	5	ASP
1	aa	9	LEU
1	aa	21	ARG
1	aa	23	LYS
1	aa	70	TYR
1	aa	132	VAL
1	aa	200	PHE
1	aa	239	VAL
1	aa	246	ARG
2	A	267	THR
2	A	295	PHE
2	A	309	SER
2	A	332	LYS
2	A	348	LYS
2	A	369	GLN
2	A	370	VAL
2	A	386	LEU
2	A	441	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	A	459	LEU
2	A	468	GLU
2	A	522	ARG
1	bb	20	ILE
1	bb	22	LEU
1	bb	48	LEU
1	bb	64	THR
1	bb	70	TYR
1	bb	83	THR
1	bb	100	ASN
1	bb	132	VAL
1	bb	144	THR
1	bb	183	VAL
1	bb	187	VAL
1	bb	200	PHE
1	bb	226	VAL
1	bb	238	THR
2	B	266	ASP
2	B	292	GLU
2	B	294	VAL
2	B	295	PHE
2	B	336	SER
2	B	356	ARG
2	B	417	ARG
2	B	462	SER
2	B	468	GLU
2	B	476	LEU
2	B	534	ARG
2	B	539	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
2	A	387	HIS
2	A	464	ASN
2	A	474	GLN
2	A	533	ASN
1	bb	100	ASN
1	bb	181	GLN
1	bb	186	ASN
1	bb	228	GLN
2	B	287	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	413	GLN
2	B	432	HIS
2	B	490	ASN
2	B	527	GLN
2	B	533	ASN

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)

10 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	NAG	C	1	3,2	14,14,15	0.51	0	17,19,21	0.63	0
3	NAG	C	2	3	14,14,15	0.46	0	17,19,21	0.39	0
3	BMA	C	3	3	11,11,12	1.64	3 (27%)	15,15,17	0.94	0
3	FUC	C	4	3	10,10,11	1.32	2 (20%)	14,14,16	1.35	2 (14%)
4	NAG	D	1	4,2	14,14,15	0.62	0	17,19,21	0.52	0
4	NAG	D	2	4	14,14,15	0.52	0	17,19,21	0.58	0
4	FUC	D	3	4	10,10,11	1.40	2 (20%)	14,14,16	1.21	2 (14%)
5	NAG	E	1	5	15,15,15	0.36	0	21,21,21	0.73	0
5	GAL	E	2	5	11,11,12	1.64	4 (36%)	15,15,17	1.57	3 (20%)
5	SIA	E	3	5	17,20,21	1.40	4 (23%)	21,28,31	1.53	4 (19%)

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	NAG	C	1	3,2	-	0/6/23/26	0/1/1/1
3	NAG	C	2	3	-	1/6/23/26	0/1/1/1
3	BMA	C	3	3	-	0/2/19/22	0/1/1/1
3	FUC	C	4	3	-	-	0/1/1/1
4	NAG	D	1	4,2	-	2/6/23/26	0/1/1/1
4	NAG	D	2	4	-	4/6/23/26	0/1/1/1
4	FUC	D	3	4	-	-	0/1/1/1
5	NAG	E	1	5	-	3/6/26/26	0/1/1/1
5	GAL	E	2	5	-	2/2/19/22	0/1/1/1
5	SIA	E	3	5	-	6/14/34/38	0/1/1/1

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	3	BMA	C1-C2	3.59	1.60	1.52
5	E	3	SIA	C7-C6	2.68	1.56	1.53
5	E	2	GAL	C1-C2	2.58	1.58	1.52
5	E	3	SIA	C8-C7	2.44	1.58	1.53
5	E	2	GAL	C2-C3	2.41	1.56	1.52
3	C	3	BMA	C2-C3	2.35	1.56	1.52
5	E	3	SIA	C6-C5	2.32	1.56	1.53
4	D	3	FUC	C4-C5	2.28	1.57	1.52
5	E	2	GAL	O5-C5	2.28	1.48	1.43
3	C	4	FUC	C4-C3	2.25	1.58	1.52
4	D	3	FUC	O5-C5	2.24	1.48	1.43
3	C	3	BMA	C4-C5	2.20	1.57	1.53
5	E	3	SIA	O6-C2	2.13	1.49	1.43
5	E	2	GAL	O3-C3	2.11	1.47	1.43
3	C	4	FUC	C4-C5	2.01	1.57	1.52

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	E	3	SIA	C6-O6-C2	3.35	118.51	111.34
5	E	2	GAL	O3-C3-C2	3.29	116.29	109.99
5	E	2	GAL	O5-C1-C2	-3.23	105.78	110.77
3	C	4	FUC	C1-C2-C3	-2.93	106.07	109.67
4	D	3	FUC	O5-C5-C4	2.81	114.56	109.52

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	E	3	SIA	C8-C7-C6	2.75	118.25	113.03
4	D	3	FUC	C1-O5-C5	2.48	118.41	112.78
5	E	3	SIA	C9-C8-C7	2.45	117.73	112.41
5	E	3	SIA	C4-C3-C2	2.37	114.06	109.81
5	E	2	GAL	O2-C2-C1	2.26	113.77	109.15
3	C	4	FUC	O5-C5-C4	2.20	113.46	109.52

There are no chirality outliers.

All (18) torsion outliers are listed below:

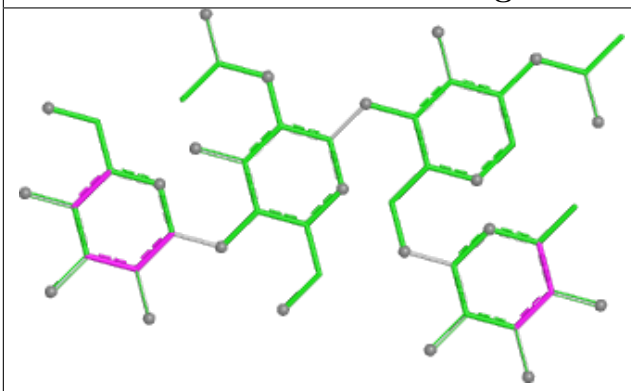
Mol	Chain	Res	Type	Atoms
5	E	3	SIA	C6-C7-C8-C9
5	E	3	SIA	O7-C7-C8-C9
5	E	3	SIA	O7-C7-C8-O8
4	D	2	NAG	O5-C5-C6-O6
4	D	2	NAG	C4-C5-C6-O6
5	E	2	GAL	O5-C5-C6-O6
4	D	2	NAG	C8-C7-N2-C2
4	D	2	NAG	O7-C7-N2-C2
5	E	1	NAG	C8-C7-N2-C2
5	E	1	NAG	O7-C7-N2-C2
5	E	3	SIA	C11-C10-N5-C5
5	E	3	SIA	O10-C10-N5-C5
5	E	2	GAL	C4-C5-C6-O6
5	E	1	NAG	O5-C5-C6-O6
5	E	3	SIA	C6-C7-C8-O8
4	D	1	NAG	C4-C5-C6-O6
3	C	2	NAG	C4-C5-C6-O6
4	D	1	NAG	O5-C5-C6-O6

There are no ring outliers.

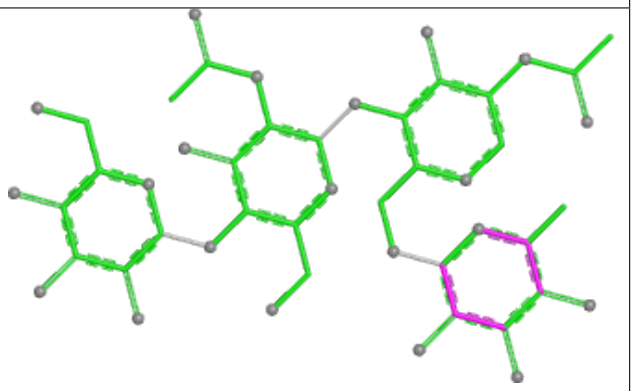
No monomer is involved in short contacts.

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

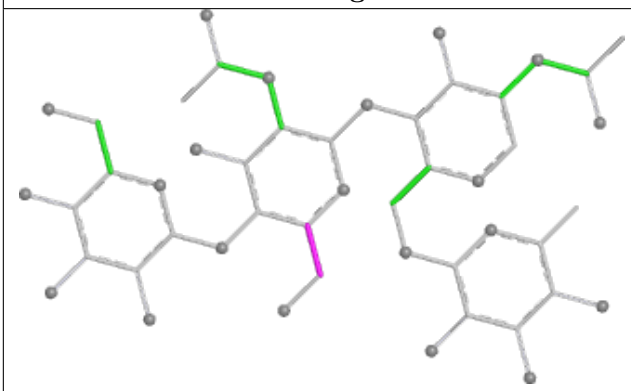
Oligosaccharide Chain C



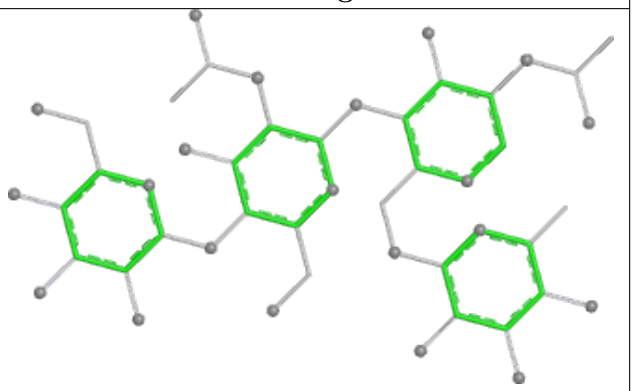
Bond lengths



Bond angles

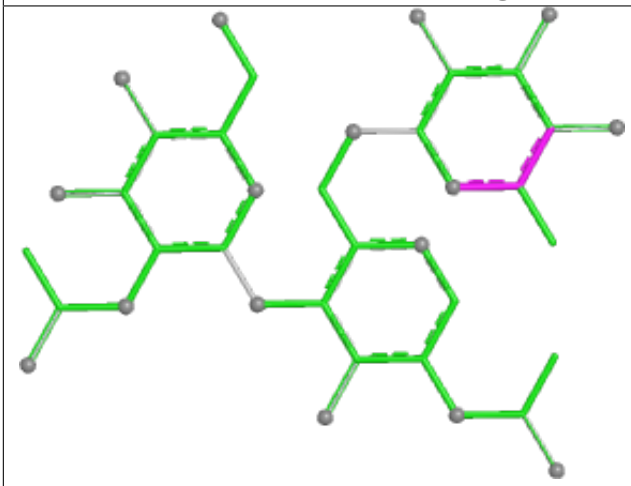


Torsions

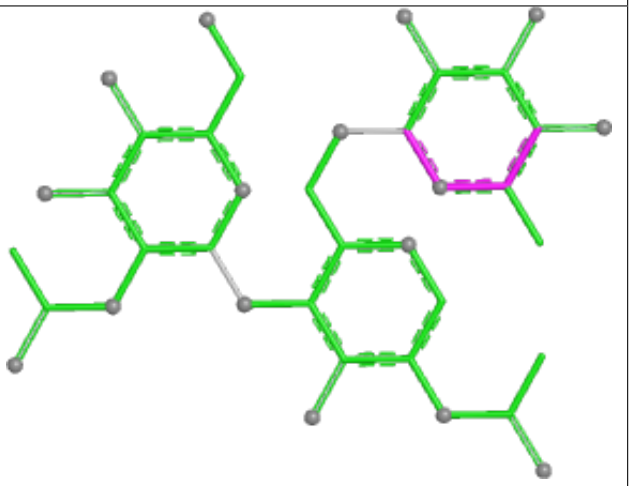


Rings

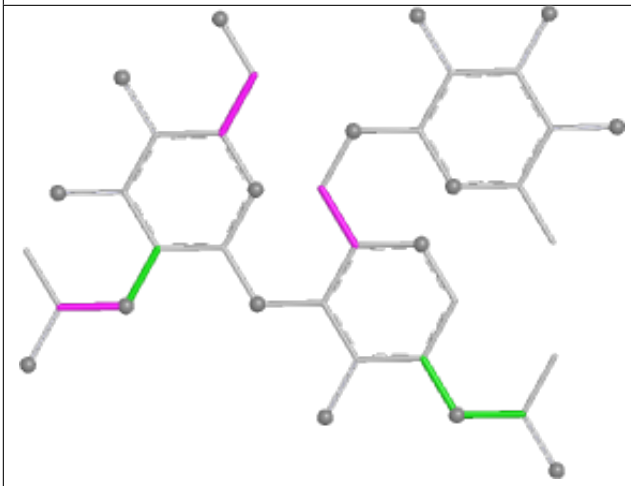
Oligosaccharide Chain D



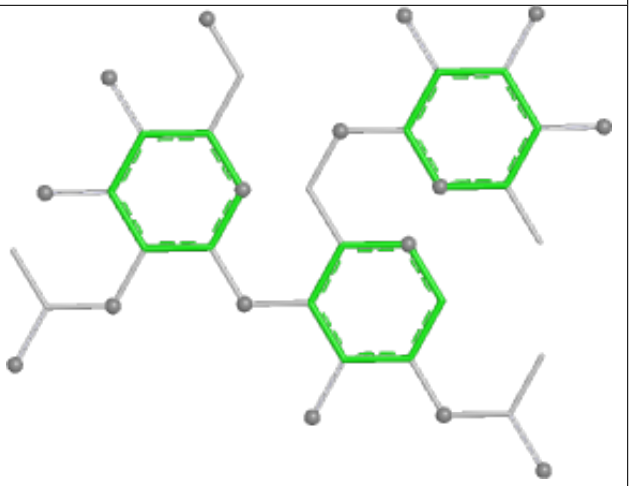
Bond lengths



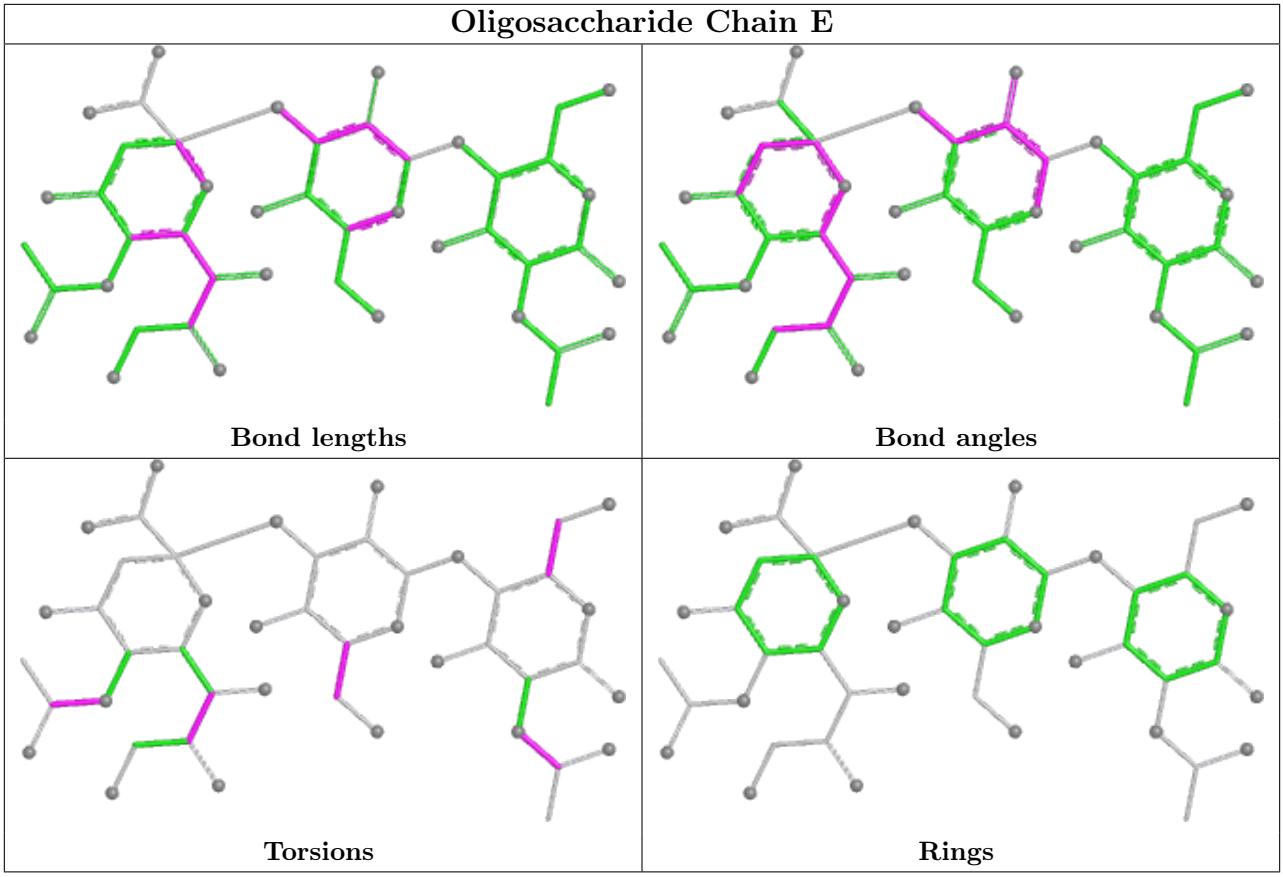
Bond angles



Torsions



Rings



5.6 Ligand geometry (i)

Of 47 ligands modelled in this entry, 31 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	$\#\|Z\| > 2$
7	SO4	bb	302	-	4,4,4	0.14	0	6,6,6	0.34	0
6	4J1	aa	602	-	37,37,37	1.42	7 (18%)	41,49,49	2.13	9 (21%)
7	SO4	A	601	-	4,4,4	0.26	0	6,6,6	0.55	0
7	SO4	A	602	-	4,4,4	0.10	0	6,6,6	0.34	0
6	4J1	aa	601	-	37,37,37	1.37	7 (18%)	41,49,49	1.53	5 (12%)
7	SO4	A	606	-	4,4,4	0.17	0	6,6,6	0.27	0
7	SO4	aa	604	-	4,4,4	0.23	0	6,6,6	0.20	0
7	SO4	B	603	-	4,4,4	0.24	0	6,6,6	0.35	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
7	SO4	bb	301	-	4,4,4	0.14	0	6,6,6	0.20	0
9	GOL	aa	610	-	5,5,5	1.28	1 (20%)	5,5,5	1.04	0
7	SO4	aa	603	-	4,4,4	0.33	0	6,6,6	0.24	0
7	SO4	A	604	-	4,4,4	0.19	0	6,6,6	0.25	0
7	SO4	B	601	-	4,4,4	0.15	0	6,6,6	0.52	0
7	SO4	A	605	-	4,4,4	0.26	0	6,6,6	0.31	0
7	SO4	A	603	-	4,4,4	0.19	0	6,6,6	0.19	0
7	SO4	B	602	10	4,4,4	0.18	0	6,6,6	0.20	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	4J1	aa	601	-	-	12/20/40/40	0/3/3/3
9	GOL	aa	610	-	-	0/4/4/4	-
6	4J1	aa	602	-	-	6/20/40/40	0/3/3/3

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	aa	602	4J1	C20-C21	3.43	1.54	1.47
6	aa	602	4J1	C12-N13	-3.23	1.32	1.35
6	aa	601	4J1	C12-N13	-3.21	1.32	1.35
6	aa	601	4J1	C20-C21	3.07	1.53	1.47
6	aa	602	4J1	C12-C11	2.89	1.40	1.36
6	aa	601	4J1	O04-C03	2.73	1.51	1.44
6	aa	602	4J1	O04-C03	2.53	1.50	1.44
6	aa	601	4J1	O04-C05	2.53	1.48	1.41
6	aa	601	4J1	O25-C24	2.36	1.41	1.36
6	aa	602	4J1	O04-C05	2.33	1.47	1.41
9	aa	610	GOL	C1-C2	2.27	1.61	1.51
6	aa	602	4J1	O25-C24	2.11	1.40	1.36
6	aa	602	4J1	O33-C32	2.08	1.47	1.43
6	aa	601	4J1	N29-N28	2.07	1.38	1.34
6	aa	601	4J1	C12-C11	2.02	1.39	1.36

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	aa	602	4J1	O23-N22-C21	8.37	126.65	111.86
6	aa	601	4J1	O23-N22-C21	6.06	122.57	111.86
6	aa	602	4J1	C21-C20-N19	4.22	124.19	115.69
6	aa	602	4J1	C20-N19-C18	4.13	122.62	117.91
6	aa	601	4J1	C20-N19-C18	3.81	122.25	117.91
6	aa	602	4J1	C24-C20-C21	-3.70	116.06	121.55
6	aa	602	4J1	O04-C03-C34	3.61	116.24	109.69
6	aa	602	4J1	C32-C34-C03	3.04	115.67	110.24
6	aa	601	4J1	C14-N13-C12	-2.83	123.16	129.82
6	aa	602	4J1	C27-C26-C24	-2.80	117.63	120.50
6	aa	602	4J1	C12-C11-N29	-2.62	107.45	111.34
6	aa	602	4J1	C17-C18-N19	2.43	119.57	115.95
6	aa	601	4J1	C12-C11-N29	-2.37	107.81	111.34
6	aa	601	4J1	C34-C32-C30	2.29	114.82	110.82

There are no chirality outliers.

All (18) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	aa	601	4J1	O04-C05-O06-C07
6	aa	601	4J1	C15-C16-C17-C18
6	aa	601	4J1	N19-C20-C21-N22
6	aa	601	4J1	N13-C14-C15-C16
6	aa	601	4J1	C14-C15-C16-C17
6	aa	602	4J1	C08-C07-O06-C05
6	aa	601	4J1	C30-C05-O06-C07
6	aa	601	4J1	O01-C02-C03-O04
6	aa	601	4J1	O06-C07-C08-C09
6	aa	602	4J1	C14-C15-C16-C17
6	aa	601	4J1	C20-C21-N22-O23
6	aa	602	4J1	C09-C10-C11-C12
6	aa	602	4J1	C20-C21-N22-O23
6	aa	601	4J1	C15-C14-N13-N28
6	aa	601	4J1	C24-C20-C21-N22
6	aa	601	4J1	C15-C14-N13-C12
6	aa	602	4J1	N19-C20-C21-N22
6	aa	602	4J1	C15-C16-C17-C18

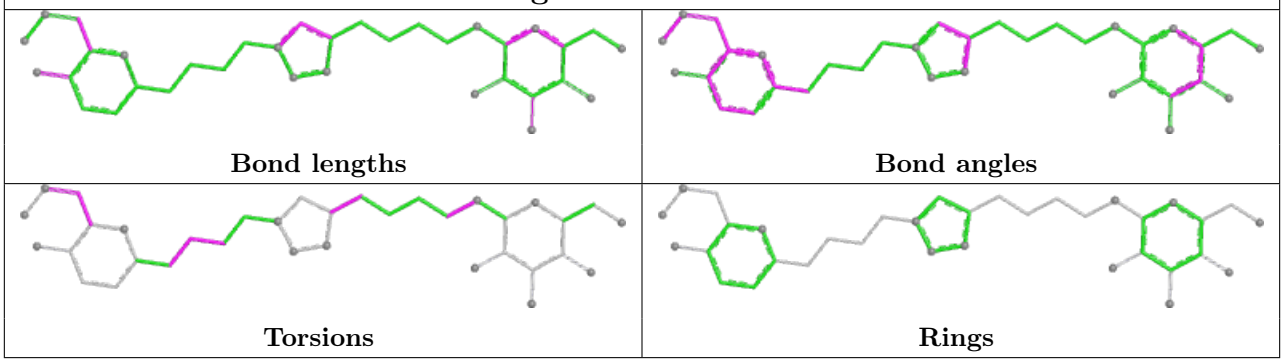
There are no ring outliers.

No monomer is involved in short contacts.

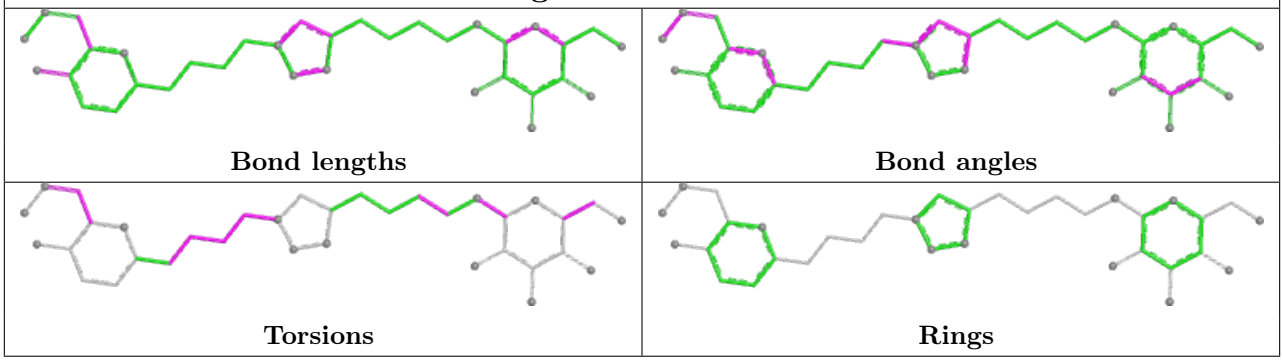
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In

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

Ligand 4J1 aa 602



Ligand 4J1 aa 601



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	aa	254/257 (98%)	0.49	1 (0%)	92	84	57, 72, 103, 168	0
1	bb	254/257 (98%)	0.46	2 (0%)	86	73	61, 83, 115, 152	0
2	A	282/282 (100%)	0.48	1 (0%)	92	84	59, 77, 107, 147	0
2	B	279/282 (98%)	0.48	1 (0%)	92	84	60, 84, 125, 188	0
All	All	1069/1078 (99%)	0.48	5 (0%)	91	81	57, 78, 114, 188	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	326	VAL	2.5
1	bb	102	TRP	2.2
1	bb	32	PHE	2.1
1	aa	60	VAL	2.0
2	B	515	LEU	2.0

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^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	BMA	C	3	11/12	0.55	0.30	177,183,192,192	0
5	NAG	E	1	15/15	0.81	0.33	149,157,171,172	0

Continued on next page...

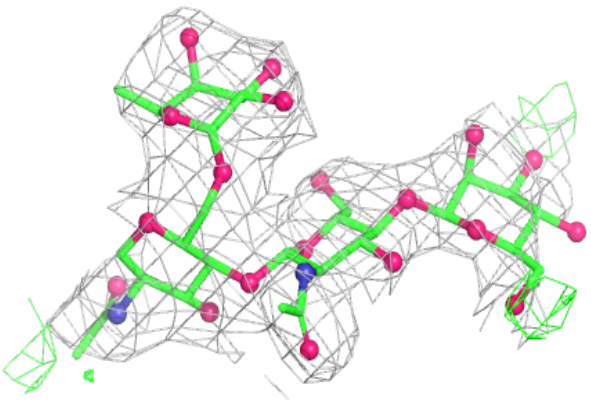
Continued from previous page...

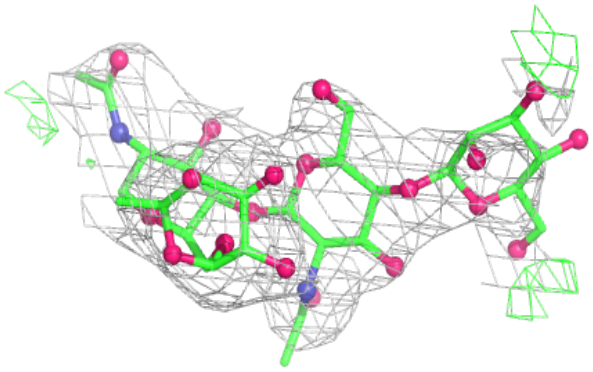
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	NAG	D	2	14/15	0.84	0.28	149,156,161,166	0
4	FUC	D	3	10/11	0.87	0.33	144,155,161,162	0
3	FUC	C	4	10/11	0.90	0.16	117,133,138,138	0
5	SIA	E	3	20/21	0.91	0.27	112,117,124,125	0
3	NAG	C	2	14/15	0.92	0.21	113,137,155,165	0
4	NAG	D	1	14/15	0.93	0.17	131,138,145,152	0
5	GAL	E	2	11/12	0.94	0.18	96,118,128,133	0
3	NAG	C	1	14/15	0.94	0.20	97,107,115,124	0

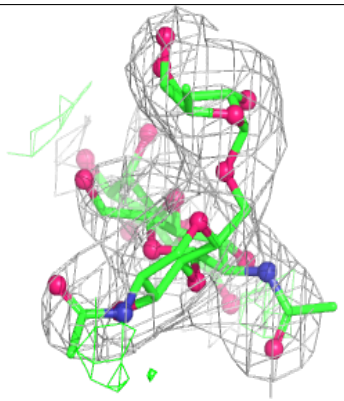
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around Chain C:

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

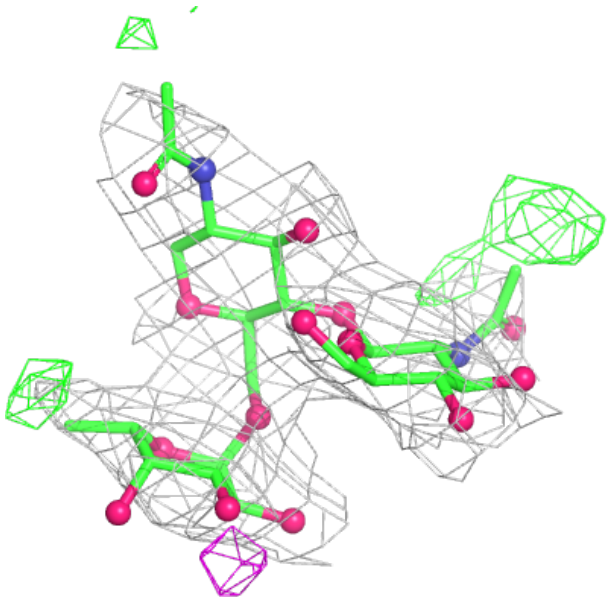


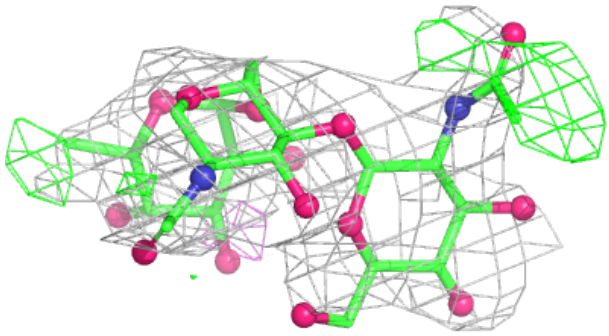


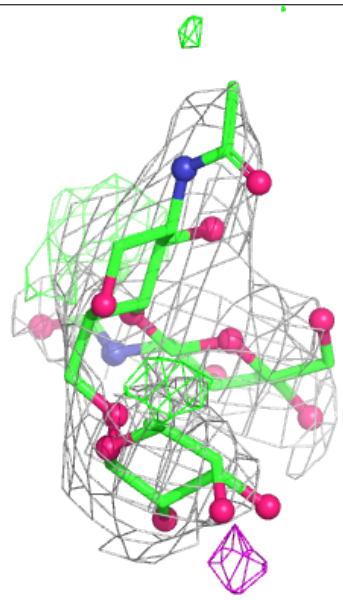


Electron density around Chain D:

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

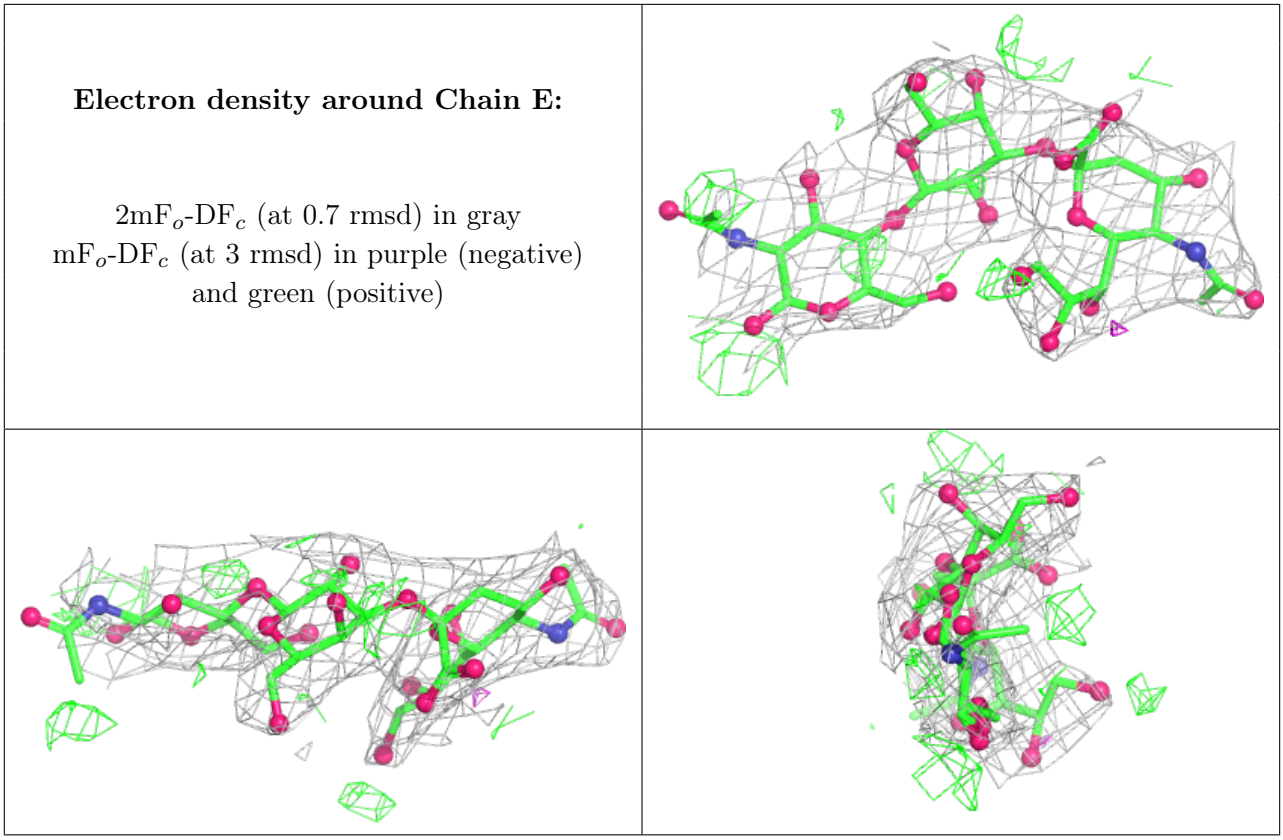






Electron density around Chain E:

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



6.4 Ligands

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
8	CL	A	615	1/1	0.50	0.26	120,120,120,120	0
8	CL	B	609	1/1	0.56	0.13	111,111,111,111	0
8	CL	bb	307	1/1	0.61	0.18	115,115,115,115	0
8	CL	B	613	1/1	0.67	0.15	119,119,119,119	0
8	CL	aa	607	1/1	0.69	0.14	97,97,97,97	0
8	CL	A	611	1/1	0.69	0.21	95,95,95,95	0
8	CL	aa	605	1/1	0.69	0.26	107,107,107,107	0
9	GOL	aa	610	6/6	0.70	0.29	95,98,103,104	0
8	CL	bb	308	1/1	0.72	0.14	110,110,110,110	0
8	CL	A	613	1/1	0.75	0.12	111,111,111,111	0
8	CL	aa	606	1/1	0.76	0.19	97,97,97,97	0
8	CL	B	614	1/1	0.76	0.14	94,94,94,94	0
6	4J1	aa	601	35/35	0.76	0.46	114,127,146,148	0
10	MG	B	605	1/1	0.76	0.15	75,75,75,75	0

Continued on next page...

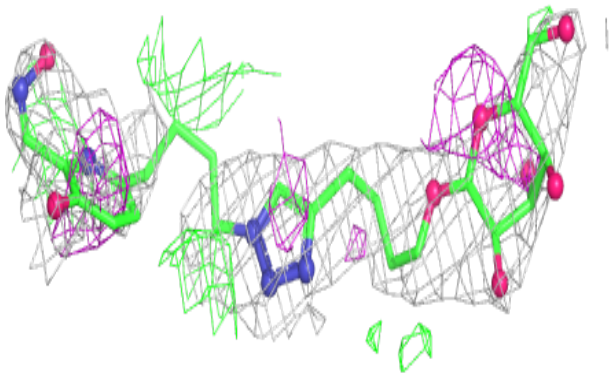
Continued from previous page...

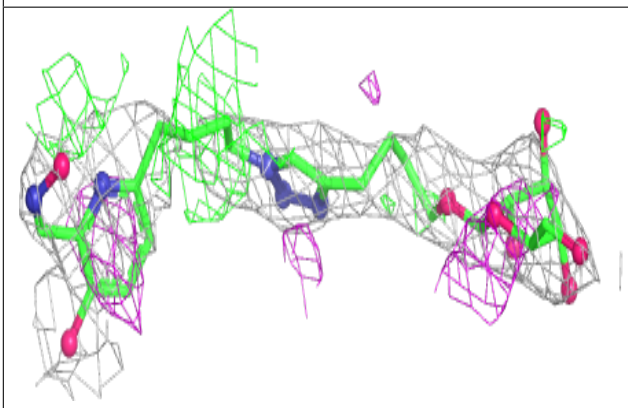
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
8	CL	B	607	1/1	0.78	0.16	88,88,88,88	0
8	CL	aa	609	1/1	0.78	0.16	101,101,101,101	0
8	CL	B	608	1/1	0.80	0.16	101,101,101,101	0
7	SO4	A	602	5/5	0.81	0.23	107,109,116,119	5
10	MG	B	604	1/1	0.82	0.33	98,98,98,98	0
8	CL	A	612	1/1	0.84	0.11	98,98,98,98	0
8	CL	B	610	1/1	0.85	0.13	106,106,106,106	0
6	4J1	aa	602	35/35	0.85	0.40	104,117,132,133	0
8	CL	bb	304	1/1	0.86	0.13	104,104,104,104	0
7	SO4	bb	302	5/5	0.86	0.17	132,133,139,144	0
10	MG	A	608	1/1	0.87	0.39	91,91,91,91	0
7	SO4	A	606	5/5	0.89	0.24	109,115,129,132	0
7	SO4	bb	301	5/5	0.89	0.12	130,134,140,143	0
7	SO4	aa	604	5/5	0.89	0.17	125,127,140,145	0
8	CL	A	610	1/1	0.90	0.17	104,104,104,104	0
8	CL	bb	305	1/1	0.91	0.31	106,106,106,106	0
8	CL	B	612	1/1	0.91	0.10	101,101,101,101	0
7	SO4	A	604	5/5	0.91	0.11	128,134,143,146	0
8	CL	A	609	1/1	0.91	0.17	85,85,85,85	0
8	CL	A	614	1/1	0.93	0.19	83,83,83,83	0
8	CL	B	606	1/1	0.94	0.13	99,99,99,99	0
8	CL	bb	306	1/1	0.94	0.18	80,80,80,80	0
7	SO4	A	603	5/5	0.94	0.18	91,92,98,101	0
7	SO4	A	601	5/5	0.94	0.13	83,85,103,104	0
7	SO4	aa	603	5/5	0.95	0.14	93,95,106,109	0
10	MG	bb	303	1/1	0.95	0.19	81,81,81,81	0
8	CL	aa	608	1/1	0.95	0.13	118,118,118,118	0
7	SO4	B	602	5/5	0.95	0.23	94,94,99,101	5
8	CL	B	611	1/1	0.96	0.15	82,82,82,82	0
7	SO4	B	601	5/5	0.96	0.15	75,76,84,85	5
7	SO4	B	603	5/5	0.96	0.18	90,92,104,107	0
7	SO4	A	605	5/5	0.97	0.16	86,90,105,106	5
10	MG	A	607	1/1	0.98	0.34	74,74,74,74	0

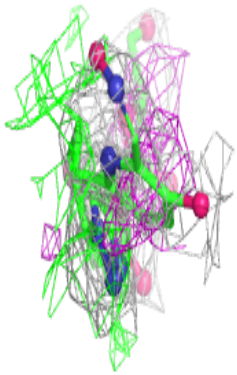
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around 4J1 aa 601:

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

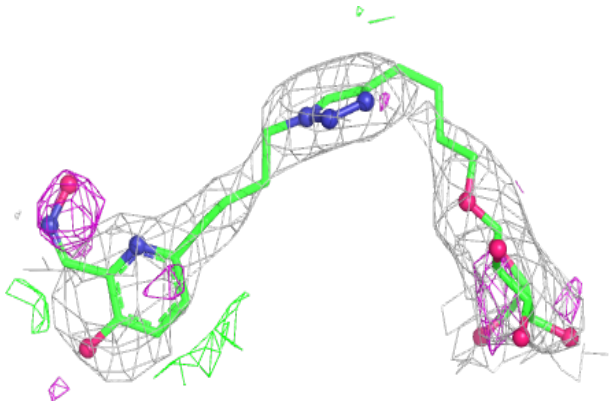


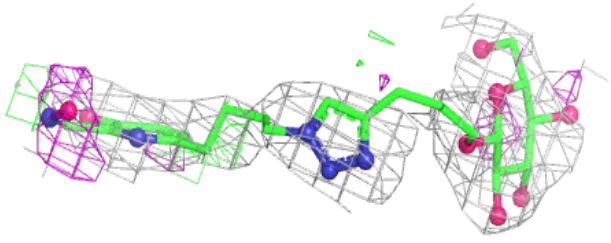




Electron density around 4J1 aa 602:

$2mF_o$-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.