Full wwPDB X-ray Structure Validation Report (i)

Aug 27, 2023 – 04:45 AM EDT

PDB ID	:	3GHH
Title	:	Structural insights into the catalytic mechanism of CD38: Evidence for a conformationally flexible covalent enzyme-substrate complex.
Authors	:	Egea, P.F.; Muller-Steffner, H.; Stroud, R.M.; Oppenheimer, N.J.; Kellenberger, E.; Schuber, F.
Deposited on	:	2009-03-03
Resolution	:	1.94 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

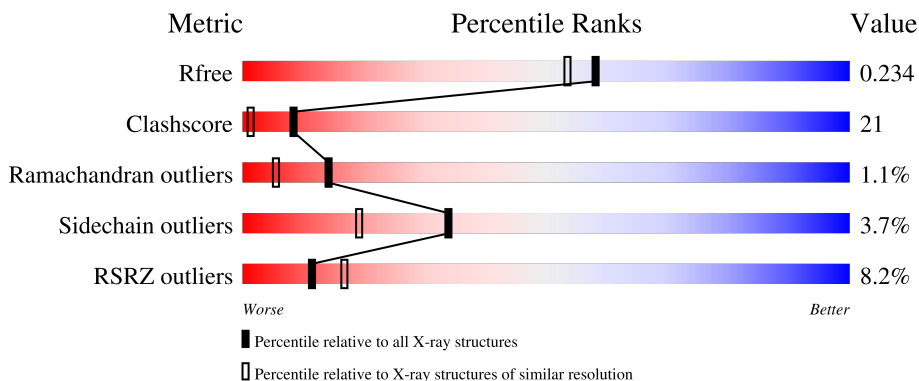|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 1.94 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	Percentile Ranks	Value
Rfree		0.234
Clashscore		21
Ramachandran outliers		1.1%
Sidechain outliers		3.7%
RSRZ outliers		8.2%

Worse *Better*

▌ Percentile relative to all X-ray structures

▐ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4310 (1.96-1.92)
Clashscore	141614	1023 (1.94-1.94)
Ramachandran outliers	138981	1007 (1.94-1.94)
Sidechain outliers	138945	1007 (1.94-1.94)
RSRZ outliers	127900	4250 (1.96-1.92)

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	A	247	% 77% 18% ..
1	B	247	15% 63% 29% ..
2	C	2	100%
2	D	2	50% 50%

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
3	SO4	A	5	-	-	X	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 4179 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 Ecto-NAD+ glycohydrolase (CD38 molecule).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	238	Total	C	N	O	S	0	3	0
			1909	1193	345	355	16			
1	B	237	Total	C	N	O	S	0	2	0
			1881	1177	343	347	14			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	218	GLN	GLU	engineered mutation	UNP Q9TTF5
B	218	GLN	GLU	engineered mutation	UNP Q9TTF5

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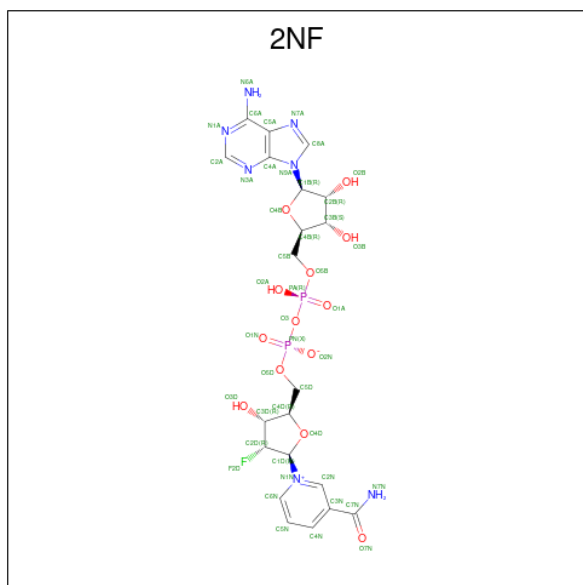

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	2	Total	C	N	O	0	0	0
			28	16	2	10			
2	D	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

SO4

O1

O3 O⁻ — S — O⁻ O4

S

O2

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 5	O 4	S 1	0	0
3	A	1	Total 5	O 4	S 1	0	0
3	A	1	Total 5	O 4	S 1	0	0
3	B	1	Total 5	O 4	S 1	0	0
3	B	1	Total 5	O 4	S 1	0	0

- Molecule 4 is [[(2R,3S,4R,5R)-5-(6-aminopurin-9-yl)-3,4-dihydroxyoxolan-2-yl]methoxy-hydroxyphosphoryl] [(2R,3S,4R,5R)-5-(3-carbamoylpyridin-1-ium-1yl)- 3-fluoro-,4- hydroxyoxolan-2-yl]methyl phosphate (three-letter code: 2NF) (formula: $C_{21}H_{26}FN_7O_{13}P_2$).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	B	1	Total	C	F	N	O	P	0	0
			44	21	1	7	13	2		

- Molecule 5 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	155	Total	O	0	0
			155	155		
5	B	109	Total	O	0	0
			109	109		

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.
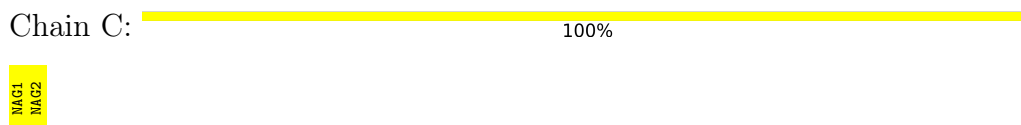
- Molecule 1: Ecto-NAD+ glycohydrolase (CD38 molecule)
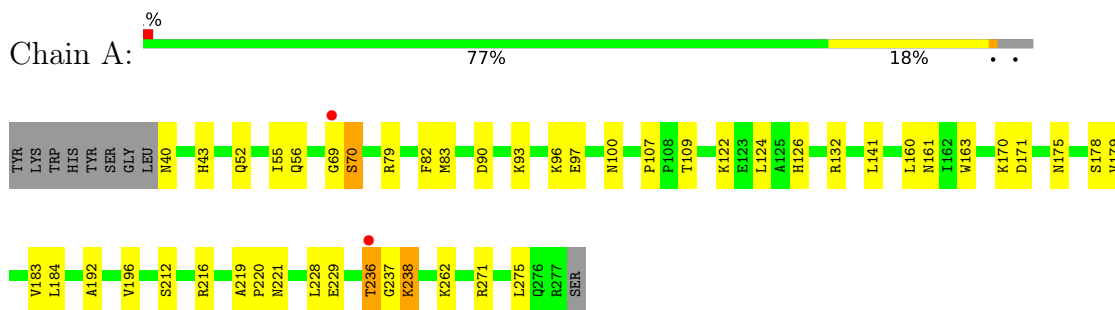
- Molecule 1: Ecto-NAD+ glycohydrolase (CD38 molecule)
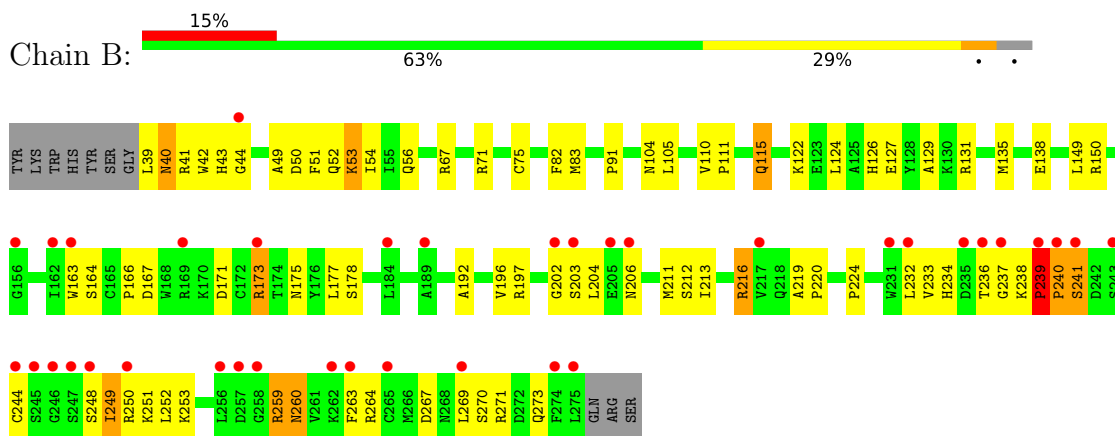
- Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

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

NAG1 NAG2

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	46.40Å 79.41Å 156.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.40 – 1.94 35.40 – 1.94	Depositor EDS
% Data completeness (in resolution range)	95.1 (35.40-1.94) 95.1 (35.40-1.94)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	1.93 (at 1.94Å)	Xtriage
Refinement program	PHENIX (phenix.refine), ELVES	Depositor
R, R_{free}	0.206 , 0.240 0.202 , 0.234	Depositor DCC
R_{free} test set	3242 reflections (7.74%)	wwPDB-VP
Wilson B-factor (Å^2)	28.8	Xtriage
Anisotropy	0.233	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 58.9	EDS
L-test for twinning[2]	$< |L| > = 0.46, < L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4179	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.79% 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: NAG, SO4, 2NF

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	0.50	0/1957	0.53	0/2645
1	B	0.43	0/1927	0.61	3/2610 (0.1%)
All	All	0.47	0/3884	0.57	3/5255 (0.1%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	273	GLN	CB-CA-C	6.81	124.01	110.40
1	B	216	ARG	O-C-N	5.38	131.31	122.70
1	B	239	PRO	C-N-CD	-5.11	109.36	120.60

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1909	0	1856	43	0
1	B	1881	0	1814	119	3
2	C	28	0	25	0	0
2	D	28	0	25	1	0
3	A	15	0	0	2	0
3	B	10	0	0	0	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	44	0	25	9	0
5	A	155	0	0	5	0
5	B	109	0	0	13	0
All	All	4179	0	3745	160	3

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 21.

All (160) 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:B:236:THR:HG21	1:B:270:SER:CB	1.28	1.56
1:B:236:THR:CG2	1:B:270:SER:HB2	1.01	1.45
1:B:248:SER:O	1:B:251:LYS:HB2	1.34	1.19
1:B:236:THR:CG2	1:B:270:SER:CB	1.95	1.19
1:B:236:THR:HG22	1:B:270:SER:HB2	1.23	1.14
1:A:96:LYS:NZ	5:A:410:HOH:O	1.81	1.13
1:B:167:ASP:HB2	5:B:380:HOH:O	1.42	1.13
1:B:202:GLY:HA2	1:B:204:LEU:H	0.97	1.11
1:B:202:GLY:HA2	1:B:204:LEU:N	1.67	1.07
1:A:69:GLY:HA3	1:A:70:SER:CB	1.87	1.02
1:A:237:GLY:C	1:A:238:LYS:HG3	1.78	1.01
1:A:237:GLY:C	1:A:238:LYS:CG	2.28	1.00
1:B:71:ARG:HD3	5:B:372:HOH:O	1.62	0.98
1:B:42:TRP:HB2	1:B:164[A]:SER:OG	1.64	0.95
1:B:202:GLY:CA	1:B:204:LEU:H	1.80	0.94
1:B:249:ILE:CD1	1:B:249:ILE:N	2.30	0.93
1:B:248:SER:HA	1:B:251:LYS:HD2	1.52	0.91
1:B:249:ILE:N	1:B:249:ILE:HD13	1.86	0.90
1:B:44:GLY:N	5:B:380:HOH:O	2.06	0.87
1:A:237:GLY:O	1:A:238:LYS:HG2	1.76	0.85
1:B:234:HIS:CE1	1:B:236:THR:HG22	2.12	0.84
1:B:129:ALA:HB1	5:B:375:HOH:O	1.78	0.82
1:B:238:LYS:CB	1:B:239:PRO:HD2	2.10	0.81
1:A:100:ASN:HD21	1:B:104:ASN:HD21	1.27	0.81
1:B:39:LEU:HD23	1:B:40:ASN:N	1.96	0.80
1:B:240:PRO:O	1:B:241:SER:CB	2.30	0.80
1:A:69:GLY:CA	1:A:70:SER:CB	2.61	0.78
1:B:238:LYS:CB	1:B:239:PRO:CD	2.60	0.78
1:B:44:GLY:CA	5:B:380:HOH:O	2.31	0.78
1:B:249:ILE:HD13	1:B:249:ILE:H	1.47	0.77

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:43:HIS:C	5:B:380:HOH:O	2.23	0.76
4:B:301:2NF:H8A	5:B:366:HOH:O	1.85	0.74
1:B:236:THR:CG2	1:B:270:SER:CA	2.66	0.74
1:A:124:LEU:HD13	1:A:271:ARG:HG3	1.67	0.74
1:B:124:LEU:HD22	2:D:1:NAG:H81	1.70	0.73
1:A:56:GLN:HE22	1:A:79:ARG:HH21	1.36	0.73
1:A:40:ASN:HB2	1:A:163:TRP:HB2	1.70	0.73
1:B:129:ALA:CB	5:B:375:HOH:O	2.35	0.72
1:B:213:ILE:HD12	4:B:301:2NF:H5N	1.72	0.71
1:B:203:SER:N	1:B:240:PRO:O	2.24	0.70
1:A:175:ASN:ND2	1:A:178:SER:H	1.89	0.70
1:B:236:THR:N	1:B:237:GLY:HA3	2.07	0.70
1:B:236:THR:HG21	1:B:270:SER:OG	1.92	0.69
3:A:5:SO4:O1	1:B:224:PRO:HD3	1.92	0.69
1:B:127:GLU:HG3	1:B:131:ARG:NH1	2.08	0.68
1:B:175:ASN:ND2	1:B:178:SER:H	1.92	0.68
1:B:240:PRO:O	1:B:241:SER:HB2	1.92	0.68
1:B:234:HIS:HE1	1:B:236:THR:HG22	1.60	0.66
1:B:249:ILE:C	1:B:251:LYS:N	2.44	0.66
1:B:122:LYS:O	1:B:126[B]:HIS:HD2	1.79	0.65
1:B:249:ILE:O	1:B:251:LYS:N	2.30	0.65
1:B:249:ILE:O	1:B:250:ARG:C	2.32	0.65
1:B:39:LEU:HD22	1:B:41:ARG:HB3	1.79	0.64
1:B:39:LEU:HD23	1:B:40:ASN:H	1.60	0.64
1:B:248:SER:HA	1:B:251:LYS:CD	2.28	0.64
1:B:51:PHE:CD2	1:B:83:MET:HE1	2.33	0.63
1:B:249:ILE:N	1:B:249:ILE:HD12	2.13	0.63
1:B:40:ASN:O	1:B:40:ASN:ND2	2.31	0.63
1:A:40:ASN:HB2	1:A:163:TRP:CB	2.29	0.63
1:B:206:ASN:HD22	1:B:248:SER:HB3	1.63	0.62
1:B:236:THR:HG21	1:B:270:SER:CA	2.22	0.62
1:A:175:ASN:HD21	1:A:178:SER:H	1.47	0.62
1:B:126[B]:HIS:HE1	1:B:138:GLU:HB2	1.65	0.62
1:B:213:ILE:HD12	4:B:301:2NF:C5N	2.30	0.61
1:B:219:ALA:HB3	1:B:220:PRO:HD3	1.82	0.61
1:A:83:MET:HE1	1:A:160:LEU:HD11	1.82	0.60
1:B:52:GLN:HG2	1:B:83:MET:CE	2.31	0.60
1:B:236:THR:HG22	1:B:270:SER:CB	1.99	0.60
1:A:83:MET:CE	1:A:160:LEU:HD11	2.32	0.60
1:B:212:SER:O	1:B:216:ARG:HG3	2.01	0.59
1:A:183:VAL:HG21	5:A:381:HOH:O	2.03	0.59

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:51:PHE:HD2	1:B:83:MET:HE1	1.68	0.58
1:B:211:MET:O	1:B:216:ARG:NE	2.37	0.57
1:B:233:VAL:HG13	1:B:270:SER:O	2.04	0.57
1:B:224:PRO:HB3	1:B:260:ASN:HB3	1.86	0.57
1:A:237:GLY:CA	1:A:238:LYS:HG3	2.34	0.56
1:B:236:THR:HG21	1:B:270:SER:HB2	0.56	0.56
1:B:51:PHE:CE2	1:B:83:MET:HE2	2.41	0.55
1:B:149:LEU:C	1:B:150:ARG:HD2	2.28	0.55
1:A:40:ASN:ND2	1:A:161:ASN:HD21	2.04	0.54
1:A:43:HIS:HB3	1:A:170:LYS:HD2	1.90	0.54
1:B:216:ARG:HH22	4:B:301:2NF:HN6A	1.54	0.54
3:A:5:SO4:O1	1:B:224:PRO:CD	2.56	0.53
1:A:132:ARG:HH12	1:A:275:LEU:HA	1.72	0.53
1:B:175:ASN:HD21	1:B:178:SER:H	1.55	0.53
1:B:51:PHE:HE2	1:B:83:MET:HE2	1.74	0.52
1:B:248:SER:C	1:B:251:LYS:HB2	2.21	0.52
1:A:43:HIS:HD2	1:A:171:ASP:OD1	1.92	0.52
1:A:141:LEU:C	1:A:141:LEU:HD23	2.30	0.52
1:A:262:LYS:HD2	5:A:398:HOH:O	2.09	0.51
1:B:75:CYS:HB3	5:B:295:HOH:O	2.11	0.51
1:A:52:GLN:NE2	5:A:411:HOH:O	2.15	0.51
1:B:248:SER:OG	1:B:249:ILE:CD1	2.59	0.51
1:B:236:THR:CG2	1:B:271:ARG:H	2.24	0.51
1:B:39:LEU:HB3	1:B:41:ARG:HH21	1.76	0.51
1:B:43:HIS:HD2	1:B:171:ASP:OD1	1.93	0.51
1:B:240:PRO:O	1:B:241:SER:OG	2.29	0.50
1:A:56:GLN:NE2	1:A:79:ARG:HE	2.09	0.50
1:B:234:HIS:CE1	1:B:236:THR:CG2	2.92	0.50
1:B:49:ALA:O	1:B:50:ASP:HB2	2.12	0.49
1:B:236:THR:HA	1:B:237:GLY:O	2.12	0.49
1:B:216:ARG:NH2	4:B:301:2NF:HN6A	2.11	0.49
1:B:238:LYS:O	1:B:239:PRO:O	2.30	0.49
1:B:178:SER:OG	4:B:301:2NF:H2A	2.13	0.48
1:A:183:VAL:HG13	1:A:184:LEU:N	2.28	0.48
1:B:126[B]:HIS:CE1	1:B:138:GLU:HB2	2.47	0.48
1:B:126[A]:HIS:HD2	5:B:369:HOH:O	1.95	0.47
1:B:126[B]:HIS:CE1	1:B:138:GLU:OE1	2.68	0.47
1:B:122:LYS:O	1:B:126[B]:HIS:CD2	2.62	0.47
1:B:203:SER:H	1:B:240:PRO:C	2.17	0.46
1:A:221:ASN:ND2	1:B:71:ARG:NH2	2.64	0.46
1:A:212:SER:O	1:A:216:ARG:HB3	2.15	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:91:PRO:HB2	1:B:166:PRO:HG2	1.97	0.46
1:A:55:ILE:HD11	1:A:79:ARG:HG2	1.98	0.45
1:B:213:ILE:HG21	4:B:301:2NF:H6N	1.98	0.45
1:A:90:ASP:OD2	1:A:93[B]:LYS:HE2	2.16	0.45
1:B:51:PHE:CD2	1:B:83:MET:CE	2.99	0.45
1:B:248:SER:O	1:B:251:LYS:CB	2.30	0.45
1:A:219:ALA:HB3	1:A:220:PRO:HD3	1.98	0.45
1:B:39:LEU:HD13	1:B:41:ARG:NH2	2.31	0.45
1:A:122:LYS:O	1:A:126:HIS:HD2	1.99	0.45
1:B:44:GLY:HA2	5:B:380:HOH:O	2.06	0.45
1:B:249:ILE:O	1:B:252:LEU:N	2.50	0.44
1:A:107:PRO:HB2	1:A:109:THR:HG23	1.98	0.44
1:A:236:THR:N	1:A:237:GLY:HA3	2.32	0.44
1:A:97:GLU:HA	1:A:100:ASN:ND2	2.32	0.44
1:B:110:VAL:HB	1:B:111:PRO:HD2	2.00	0.44
1:B:50:ASP:HB3	1:B:53:LYS:HD3	1.99	0.43
1:A:179:VAL:O	1:A:183:VAL:HG12	2.18	0.43
1:B:51:PHE:CE2	1:B:83:MET:CE	3.00	0.43
1:B:249:ILE:C	1:B:251:LYS:H	2.20	0.43
1:B:39:LEU:CD2	1:B:41:ARG:HB3	2.46	0.43
1:B:213:ILE:HD12	4:B:301:2NF:C6N	2.48	0.43
1:B:232:LEU:HD12	1:B:232:LEU:N	2.33	0.43
1:A:40:ASN:CB	1:A:163:TRP:HB2	2.46	0.42
1:B:192:ALA:HB1	1:B:196:VAL:HG11	2.02	0.42
1:B:244:CYS:SG	1:B:267:ASP:HB3	2.59	0.42
1:A:122:LYS:HD2	1:A:126:HIS:CD2	2.54	0.42
1:B:178:SER:OG	4:B:301:2NF:C2A	2.67	0.42
1:B:53:LYS:HG2	1:B:54:ILE:N	2.34	0.42
1:A:228:LEU:C	1:A:228:LEU:HD23	2.39	0.42
1:B:239:PRO:HB2	1:B:240:PRO:HD2	2.01	0.42
1:B:40:ASN:HD22	1:B:164[B]:SER:HB3	1.85	0.42
1:B:115:GLN:OE1	1:B:192:ALA:HA	2.20	0.41
1:B:239:PRO:HA	1:B:240:PRO:HD3	1.76	0.41
1:A:79:ARG:O	1:A:83:MET:HG2	2.20	0.41
1:B:175:ASN:HD22	1:B:177:LEU:H	1.68	0.41
1:A:221:ASN:HD21	1:B:71:ARG:NH2	2.19	0.41
1:B:52:GLN:HG3	5:B:377:HOH:O	2.21	0.41
1:B:259:ARG:O	1:B:260:ASN:CB	2.67	0.41
1:B:253:LYS:HG3	1:B:263:PHE:HD2	1.86	0.41
1:B:40:ASN:HB2	1:B:163:TRP:CE3	2.56	0.40
1:B:67:ARG:CZ	1:B:135:MET:HE2	2.51	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:204:LEU:HB3	5:B:361:HOH:O	2.20	0.40
1:A:192:ALA:HB1	1:A:196:VAL:HG11	2.03	0.40
1:A:229:GLU:OE1	5:A:403:HOH:O	2.22	0.40
1:B:127:GLU:CG	1:B:131:ARG:NH1	2.82	0.40
1:B:203:SER:N	1:B:241:SER:HB2	2.37	0.40
1:B:39:LEU:HD22	1:B:41:ARG:HE	1.86	0.40
1:B:236:THR:CG2	1:B:270:SER:HA	2.50	0.40

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:163:TRP:CE3	1:B:173:ARG:NH2[4_545]	1.18	1.02
1:B:163:TRP:CZ3	1:B:173:ARG:NH2[4_545]	1.54	0.66
1:B:163:TRP:CD2	1:B:173:ARG:NH2[4_545]	1.99	0.21

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	239/247 (97%)	230 (96%)	8 (3%)	1 (0%)	34	24
1	B	237/247 (96%)	215 (91%)	18 (8%)	4 (2%)	9	2
All	All	476/494 (96%)	445 (94%)	26 (6%)	5 (1%)	14	5

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	70	SER
1	B	239	PRO
1	B	240	PRO
1	B	241	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	260	ASN

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	210/218 (96%)	207 (99%)	3 (1%)	67	58
1	B	204/218 (94%)	192 (94%)	12 (6%)	19	7
All	All	414/436 (95%)	399 (96%)	15 (4%)	34	20

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	82	PHE
1	A	236	THR
1	A	238	LYS
1	B	40	ASN
1	B	53	LYS
1	B	56	GLN
1	B	82	PHE
1	B	105	LEU
1	B	115	GLN
1	B	173	ARG
1	B	197	ARG
1	B	249	ILE
1	B	259	ARG
1	B	264	ARG
1	B	269	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	40	ASN
1	A	43	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	56	GLN
1	A	100	ASN
1	A	126	HIS
1	A	175	ASN
1	A	221	ASN
1	B	40	ASN
1	B	43	HIS
1	B	100	ASN
1	B	115	GLN
1	B	175	ASN
1	B	206	ASN

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 ⓘ

4 monosaccharides are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
2	NAG	C	1	1,2	14,14,15	0.51	0	17,19,21	0.87	1 (5%)
2	NAG	C	2	2	14,14,15	0.59	0	17,19,21	0.99	1 (5%)
2	NAG	D	1	1,2	14,14,15	0.44	0	17,19,21	0.95	1 (5%)
2	NAG	D	2	2	14,14,15	0.57	0	17,19,21	1.14	2 (11%)

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	C	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	C	2	2	-	0/6/23/26	0/1/1/1
2	NAG	D	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	D	2	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

| Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	2	NAG	C1-O5-C5	2.91	116.14	112.19
2	D	2	NAG	C2-N2-C7	-2.65	119.13	122.90
2	C	2	NAG	O5-C5-C6	2.36	110.90	107.20
2	D	1	NAG	C1-O5-C5	2.30	115.31	112.19
2	C	1	NAG	C1-O5-C5	2.15	115.10	112.19

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	1	NAG	1	0

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

Oligosaccharide Chain C

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain D

Bond lengths | Bond angles | Torsions | Rings
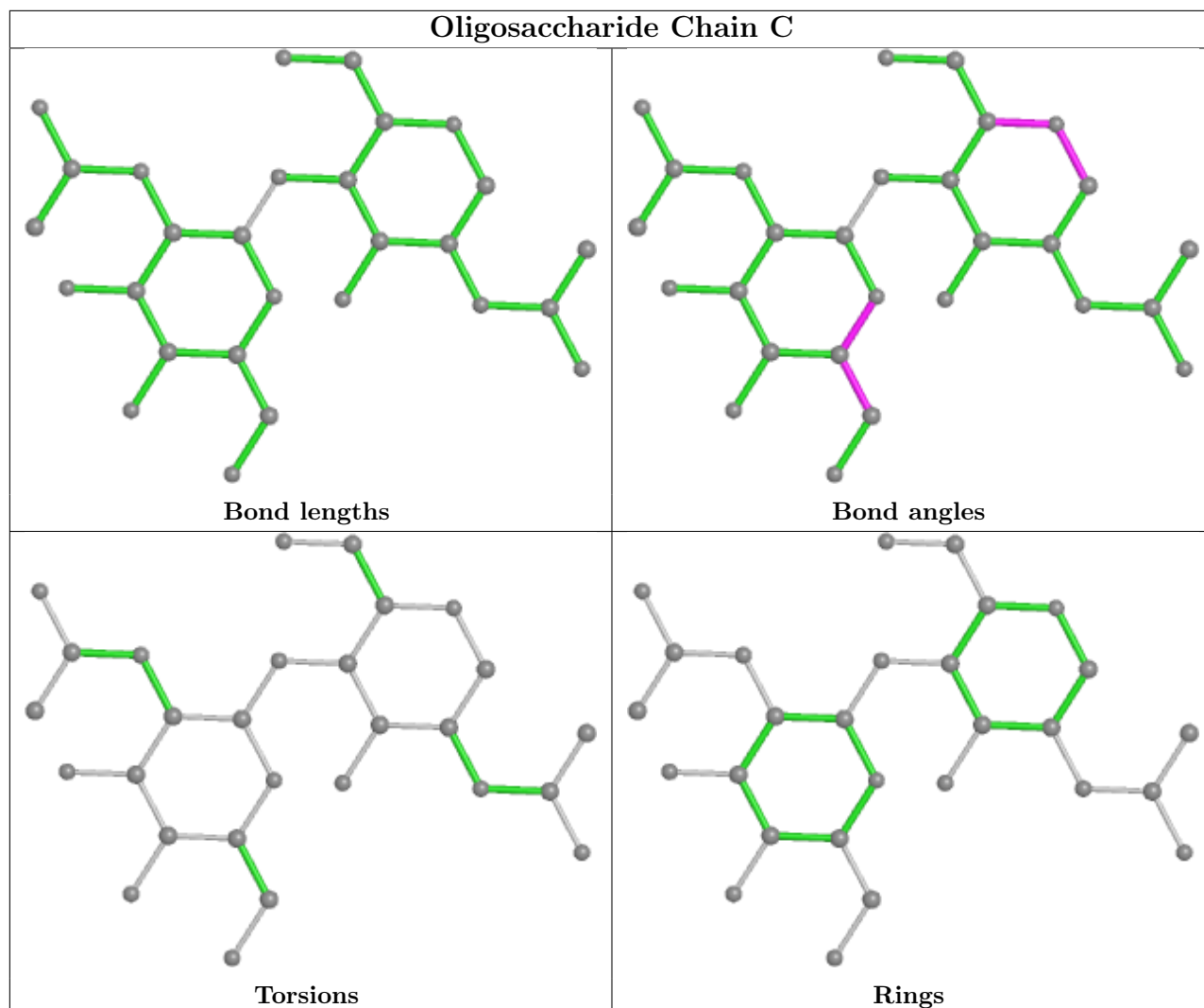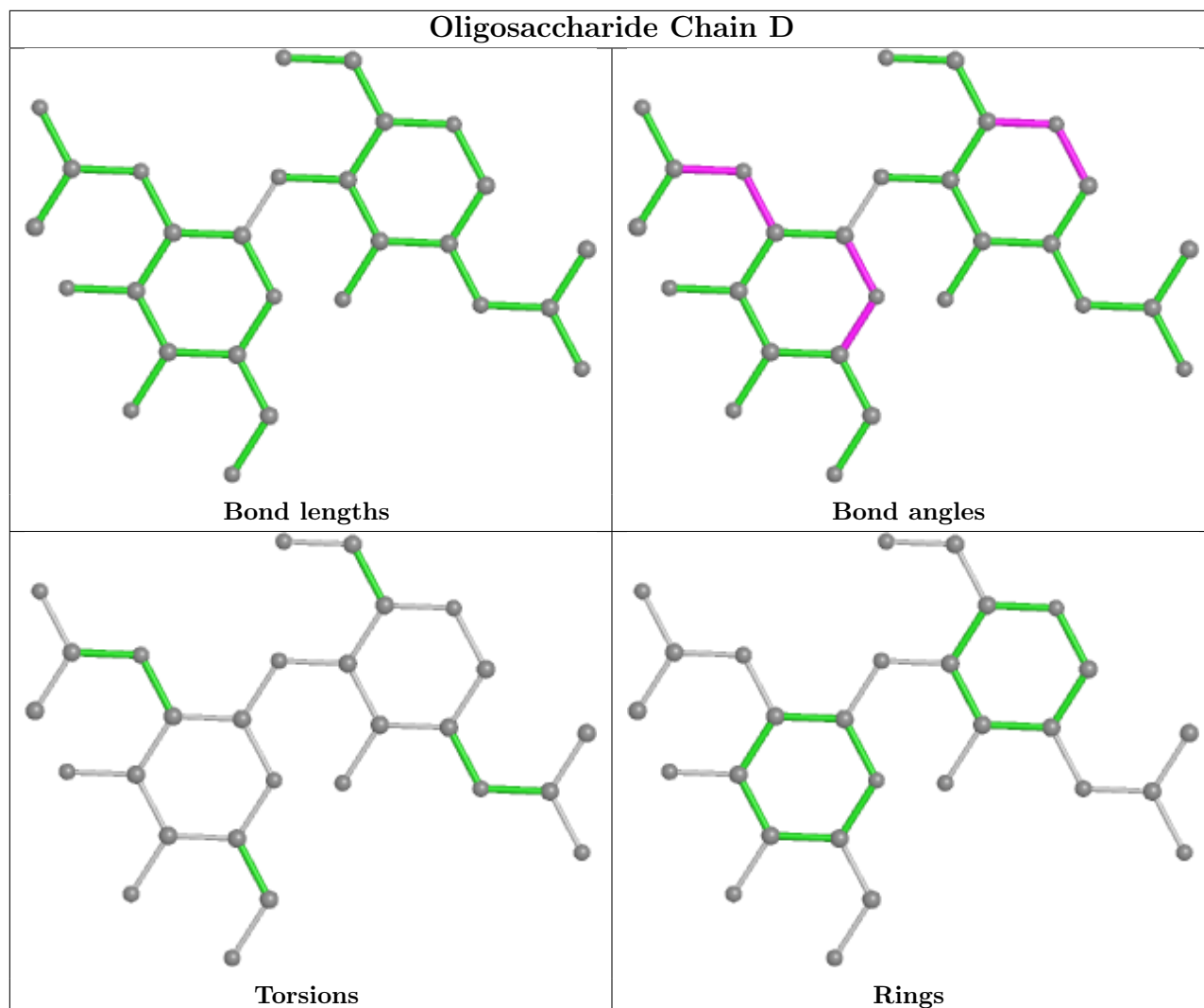
5.6 Ligand geometry ⓘ

6 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	$\#|Z| > 2$
3	SO4	A	2	-	4,4,4	0.13	0	6,6,6	0.12	0
3	SO4	A	5	-	4,4,4	0.54	0	6,6,6	0.52	0
4	2NF	B	301	-	42,48,48	1.87	4 (9%)	50,73,73	1.94	12 (24%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	SO4	B	1	-	4,4,4	0.14	0	6,6,6	0.10	0
3	SO4	B	4	-	4,4,4	0.31	0	6,6,6	0.15	0
3	SO4	A	3	-	4,4,4	0.74	0	6,6,6	0.57	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
4	2NF	B	301	-	-	7/26/62/62	0/5/5/5

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	301	2NF	O7N-C7N	9.27	1.41	1.24
4	B	301	2NF	C2A-N1A	3.50	1.40	1.33
4	B	301	2NF	C2N-N1N	3.22	1.38	1.35
4	B	301	2NF	C2A-N3A	2.89	1.36	1.32

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	301	2NF	N3A-C2A-N1A	-7.43	117.06	128.68
4	B	301	2NF	C1B-N9A-C4A	-4.09	119.45	126.64
4	B	301	2NF	C6N-N1N-C2N	-3.96	118.36	121.97
4	B	301	2NF	O2A-PA-O1A	3.32	128.63	112.24
4	B	301	2NF	C3N-C7N-N7N	2.96	121.30	117.75
4	B	301	2NF	C2N-N1N-C1D	2.50	124.71	119.14
4	B	301	2NF	O4D-C1D-C2D	-2.47	103.24	105.79
4	B	301	2NF	C5B-C4B-C3B	-2.40	106.18	115.18
4	B	301	2NF	F2D-C2D-C3D	-2.33	104.34	109.22
4	B	301	2NF	O2N-PN-O1N	2.13	122.80	112.24
4	B	301	2NF	C2D-C3D-C4D	2.10	105.12	102.40
4	B	301	2NF	C3D-C2D-C1D	2.00	105.56	103.13

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	301	2NF	C2D-C1D-N1N-C6N

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
4	B	301	2NF	C2D-C1D-N1N-C2N
4	B	301	2NF	PN-O3-PA-O5B
4	B	301	2NF	PA-O3-PN-O1N
4	B	301	2NF	PA-O3-PN-O2N
4	B	301	2NF	C4D-C5D-O5D-PN
4	B	301	2NF	C3D-C4D-C5D-O5D

There are no ring outliers.

2 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	5	SO4	2	0
4	B	301	2NF	9	0

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

Ligand 2NF B 301

Bond lengths

Bond angles

Torsions

Rings
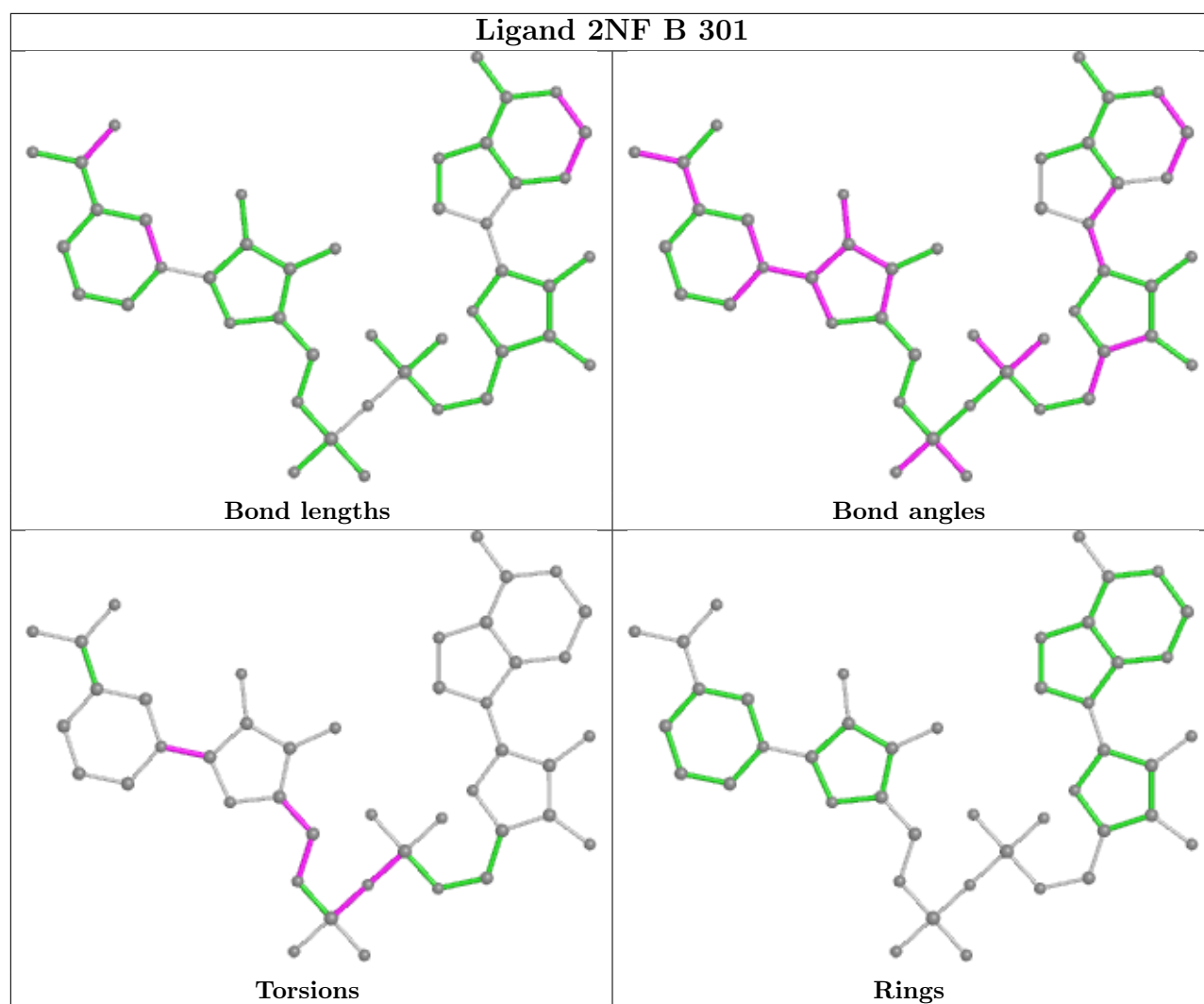
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6   Fit of model and data   ⓘ

6.1   Protein, DNA and RNA chains   ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th 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(Å²)	Q<0.9
1	A	238/247 (96%)	0.11	2 (0%)	86	89	16, 28, 51, 79	0
1	B	237/247 (95%)	0.90	37 (15%)	2	2	18, 39, 79, 110	0
All	All	475/494 (96%)	0.50	39 (8%)	11	17	16, 31, 75, 110	0

All (39) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	202	GLY	6.3
1	B	205	GLU	5.6
1	B	240	PRO	5.4
1	B	246	GLY	5.3
1	B	250	ARG	5.3
1	B	236	THR	5.1
1	B	237	GLY	4.3
1	B	247	SER	4.3
1	B	269	LEU	4.2
1	B	245	SER	4.0
1	B	156	GLY	3.9
1	A	69	GLY	3.8
1	B	163	TRP	3.8
1	B	262	LYS	3.6
1	B	275	LEU	3.5
1	B	244	CYS	3.4
1	B	44	GLY	3.4
1	B	206	ASN	3.3
1	B	232	LEU	3.1
1	B	239	PRO	2.8
1	B	274	PHE	2.7
1	B	162	ILE	2.7
1	B	265	CYS	2.6
1	B	235	ASP	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	257	ASP	2.6
1	B	184	LEU	2.5
1	B	241	SER	2.5
1	B	248	SER	2.5
1	B	243	SER	2.4
1	B	231	TRP	2.4
1	B	258	GLY	2.3
1	B	173	ARG	2.3
1	A	236	THR	2.2
1	B	203	SER	2.2
1	B	189	ALA	2.2
1	B	263	PHE	2.2
1	B	256	LEU	2.0
1	B	169	ARG	2.0
1	B	217	VAL	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, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.
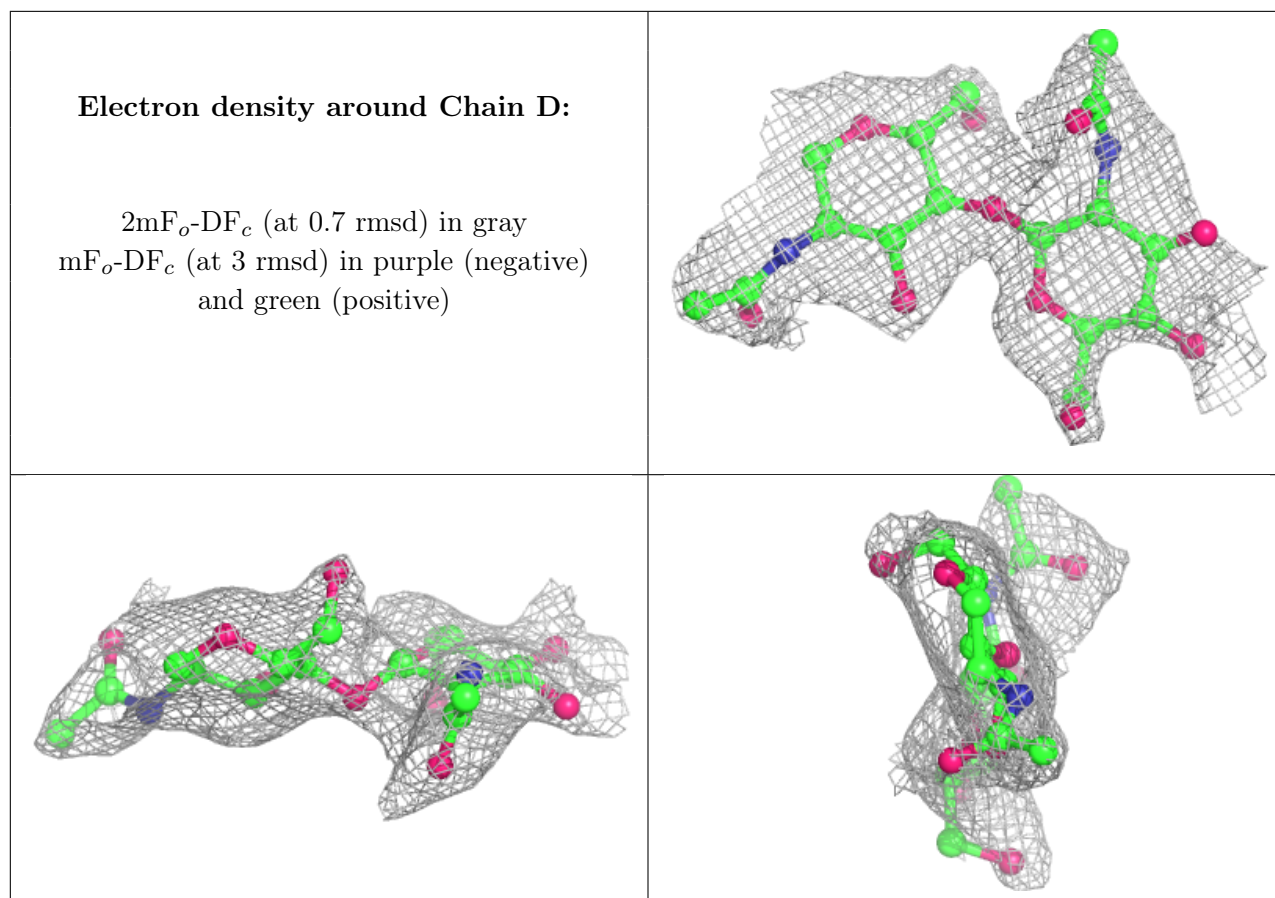
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NAG	D	2	14/15	0.78	0.28	75,89,94,95	0
2	NAG	D	1	14/15	0.86	0.17	52,70,78,80	0
2	NAG	C	2	14/15	0.88	0.26	50,57,68,72	0
2	NAG	C	1	14/15	0.92	0.12	26,34,43,48	0

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)
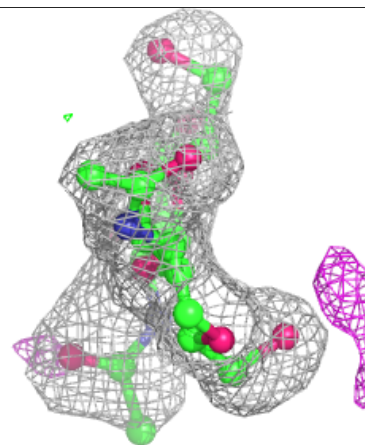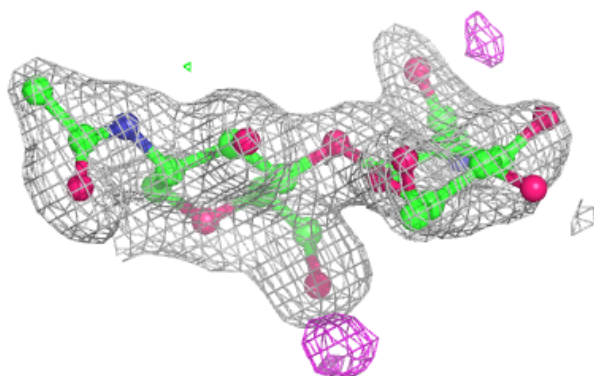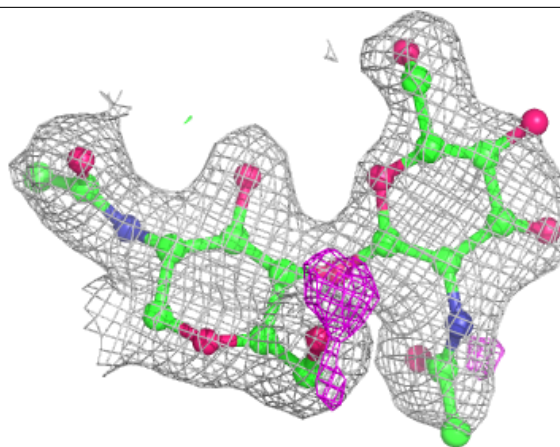
Electron density around Chain D:

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

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, 95th 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	SO4	A	5	5/5	0.84	0.43	70,70,70,70	0
4	2NF	B	301	44/44	0.92	0.13	33,45,59,66	0
3	SO4	B	4	5/5	0.94	0.19	86,90,91,94	0
3	SO4	A	3	5/5	0.94	0.22	69,70,73,78	0
3	SO4	A	2	5/5	0.97	0.20	57,61,65,70	0
3	SO4	B	1	5/5	0.97	0.15	64,67,71,76	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 2NF B 301:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
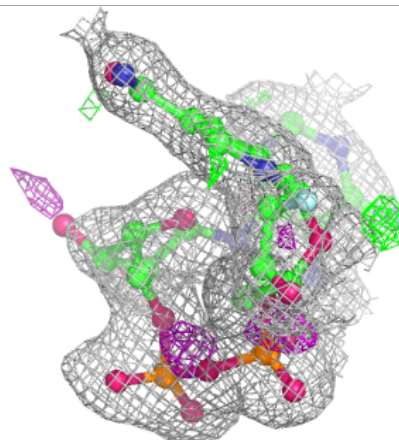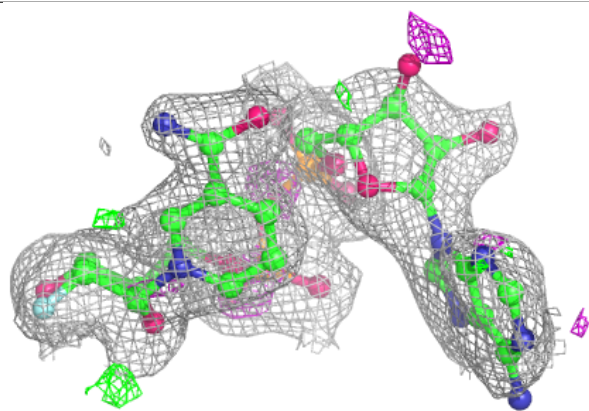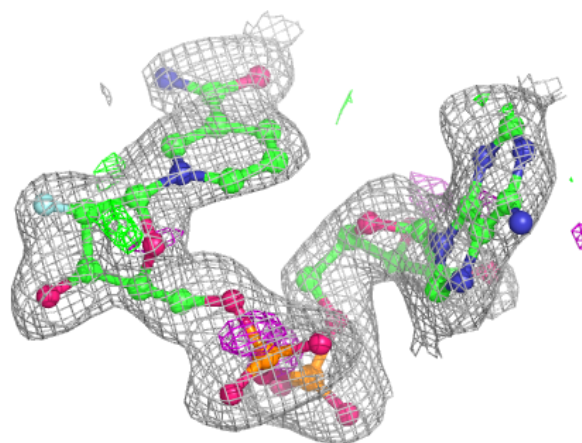
6.5 Other polymers (i)

There are no such residues in this entry.