Full wwPDB X-ray Structure Validation Report ⓘ

Oct 31, 2023 – 07:21 PM JST

PDB ID	:	5GY2
Title	:	Crystal structure of a complex between Bacillus subtilis flagellin and zebrafish Toll-like receptor 5
Authors	:	Song, W.S.; Yoon, S.I.
Deposited on	:	2016-09-21
Resolution	:	2.10 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

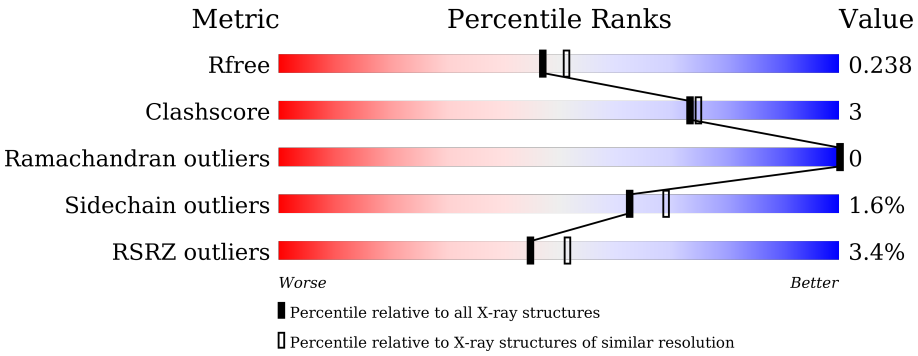
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	455	90% 7% •
1	B	455	88% 8% • •
2	C	183	83% 9% • 8%
2	D	183	82% 9% • 8%
3	E	2	100%
3	F	2	100%

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
4	NAG	A	803	-	-	-	X

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 10156 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 Tlr5b protein,Variable lymphocyte receptor B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	442	Total	C	N	O	S	0	6	0
			3526	2253	594	661	18			
1	B	441	Total	C	N	O	S	0	5	0
			3494	2238	586	652	18			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	18	ALA	-	expression tag	UNP B3DIN1
A	19	ASP	-	expression tag	UNP B3DIN1
A	20	PRO	-	expression tag	UNP B3DIN1
A	21	GLY	-	expression tag	UNP B3DIN1
A	22	THR	-	expression tag	UNP B3DIN1
A	24	GLU	VAL	engineered mutation	UNP B3DIN1
A	124	VAL	LEU	engineered mutation	UNP B3DIN1
A	159	LYS	GLN	engineered mutation	UNP B3DIN1
A	227	LYS	ARG	engineered mutation	UNP B3DIN1
A	229	THR	SER	engineered mutation	UNP B3DIN1
A	334	ASN	ASP	engineered mutation	UNP B3DIN1
A	466	SER	-	expression tag	UNP Q4G1L2
A	467	ALA	-	expression tag	UNP Q4G1L2
A	468	SER	-	expression tag	UNP Q4G1L2
A	469	LEU	-	expression tag	UNP Q4G1L2
A	470	VAL	-	expression tag	UNP Q4G1L2
A	471	PRO	-	expression tag	UNP Q4G1L2
A	472	ARG	-	expression tag	UNP Q4G1L2
B	18	ALA	-	expression tag	UNP B3DIN1
B	19	ASP	-	expression tag	UNP B3DIN1
B	20	PRO	-	expression tag	UNP B3DIN1
B	21	GLY	-	expression tag	UNP B3DIN1
B	22	THR	-	expression tag	UNP B3DIN1
B	24	GLU	VAL	engineered mutation	UNP B3DIN1
B	124	VAL	LEU	engineered mutation	UNP B3DIN1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	159	LYS	GLN	engineered mutation	UNP B3DIN1
B	227	LYS	ARG	engineered mutation	UNP B3DIN1
B	229	THR	SER	engineered mutation	UNP B3DIN1
B	334	ASN	ASP	engineered mutation	UNP B3DIN1
B	466	SER	-	expression tag	UNP Q4G1L2
B	467	ALA	-	expression tag	UNP Q4G1L2
B	468	SER	-	expression tag	UNP Q4G1L2
B	469	LEU	-	expression tag	UNP Q4G1L2
B	470	VAL	-	expression tag	UNP Q4G1L2
B	471	PRO	-	expression tag	UNP Q4G1L2
B	472	ARG	-	expression tag	UNP Q4G1L2

- Molecule 2 is a protein called Flagellin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	169	Total	C	N	O	S	0	5	0
			1250	761	217	268	4			
2	D	169	Total	C	N	O	S	0	3	0
			1255	764	219	268	4			

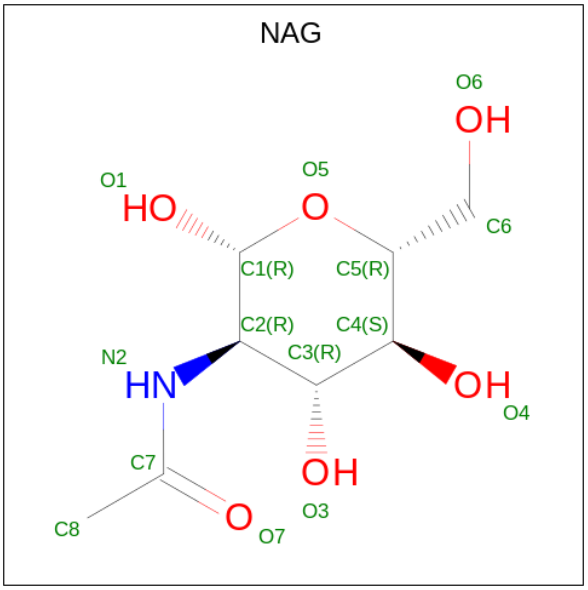
There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	48	GLY	-	expression tag	UNP E0U497
C	49	SER	-	expression tag	UNP E0U497
C	50	ALA	-	expression tag	UNP E0U497
C	51	LYS	-	expression tag	UNP E0U497
C	52	ASP	-	expression tag	UNP E0U497
C	53	PRO	-	expression tag	UNP E0U497
D	48	GLY	-	expression tag	UNP E0U497
D	49	SER	-	expression tag	UNP E0U497
D	50	ALA	-	expression tag	UNP E0U497
D	51	LYS	-	expression tag	UNP E0U497
D	52	ASP	-	expression tag	UNP E0U497
D	53	PRO	-	expression tag	UNP E0U497

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

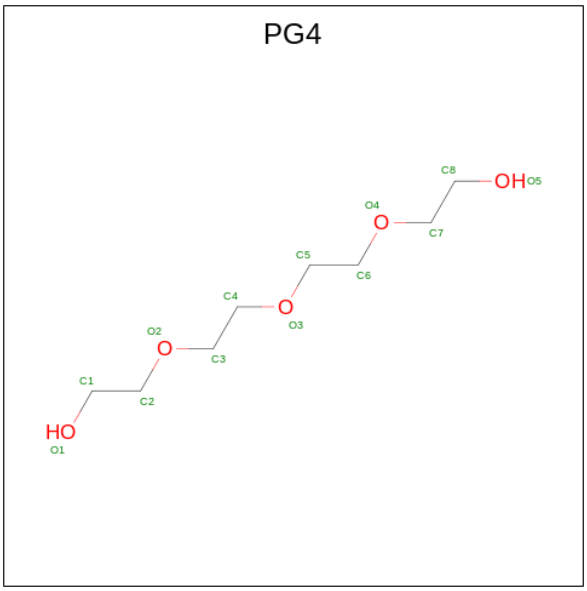
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
3	E	2	Total C N O 28 16 2 10	0	0	0
3	F	2	Total C N O 28 16 2 10	0	0	0

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



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

- Molecule 5 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 13 8 5	0	0
5	B	1	Total C O 13 8 5	0	0

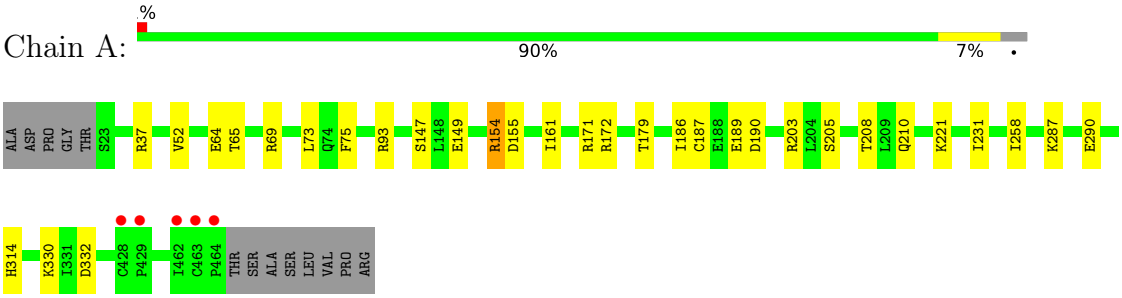
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	206	Total O 206 206	0	0
6	C	66	Total O 66 66	0	0
6	B	142	Total O 142 142	0	0
6	D	51	Total O 51 51	0	0

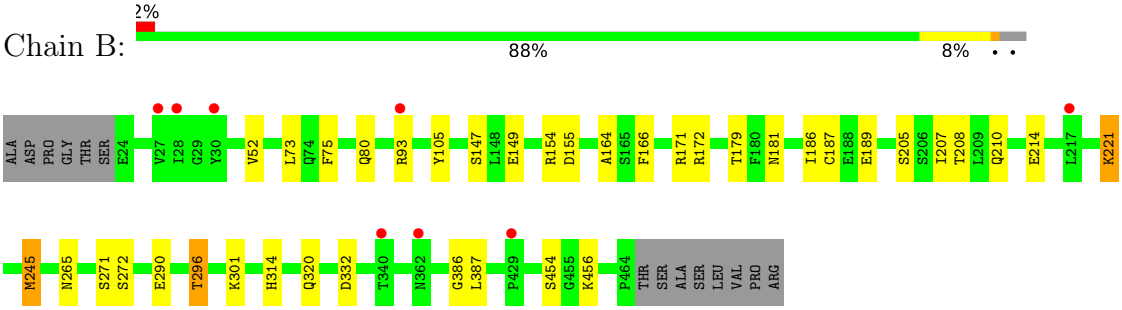
3 Residue-property plots (i)

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.

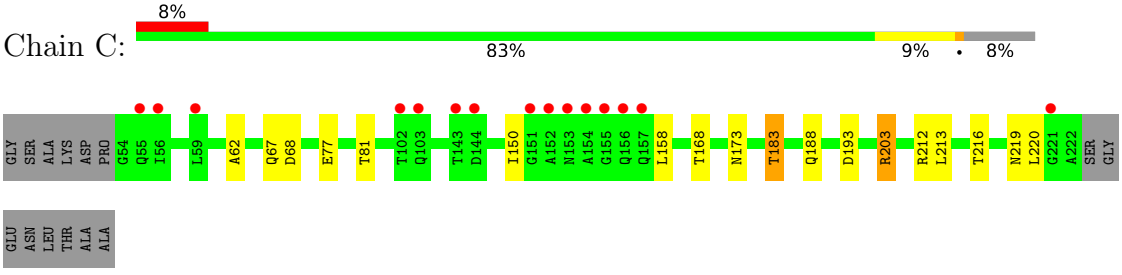
• Molecule 1: Tlr5b protein,Variable lymphocyte receptor B



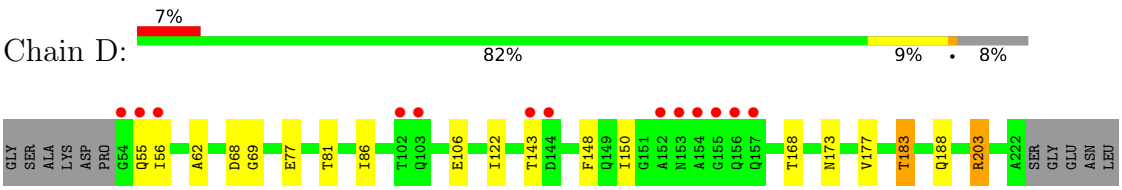
• Molecule 1: Tlr5b protein,Variable lymphocyte receptor B



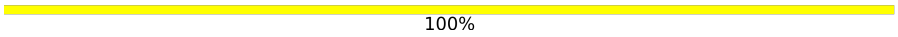
• Molecule 2: Flagellin



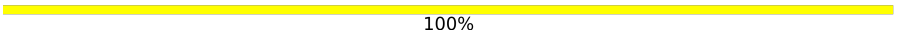
• Molecule 2: Flagellin



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

Chain E: 

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

Chain F: 

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	78.15Å 54.28Å 165.98Å 90.00° 98.62° 90.00°	Depositor
Resolution (Å)	30.00 – 2.10 28.92 – 2.09	Depositor EDS
% Data completeness (in resolution range)	98.0 (30.00-2.10) 97.5 (28.92-2.09)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.51 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.195 , 0.236 0.200 , 0.238	Depositor DCC
R_{free} test set	4023 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	24.3	Xtriage
Anisotropy	0.797	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 43.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	10156	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.39% 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: PG4, NAG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|>5$	RMSZ	#$\|Z\|>5$
1	A	0.77	0/3615	0.79	5/4890 (0.1%)
1	B	0.70	0/3580	0.74	3/4845 (0.1%)
2	C	0.81	0/1271	0.77	2/1719 (0.1%)
2	D	0.74	1/1267 (0.1%)	0.72	2/1712 (0.1%)
All	All	0.75	1/9733 (0.0%)	0.76	12/13166 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	177	VAL	CB-CG1	-5.04	1.42	1.52

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	154	ARG	NE-CZ-NH2	-9.96	115.32	120.30
1	B	154	ARG	NE-CZ-NH2	-9.12	115.74	120.30
1	B	245	MET	CG-SD-CE	-6.23	90.23	100.20
1	A	154	ARG	NE-CZ-NH1	5.91	123.25	120.30
1	A	69	ARG	NE-CZ-NH2	-5.85	117.38	120.30
2	C	203	ARG	NE-CZ-NH2	-5.71	117.45	120.30
2	D	203	ARG	NE-CZ-NH1	5.70	123.15	120.30
1	B	154	ARG	NE-CZ-NH1	5.52	123.06	120.30
1	A	69	ARG	NE-CZ-NH1	5.32	122.96	120.30
1	A	203	ARG	NE-CZ-NH2	-5.32	117.64	120.30
2	D	203	ARG	NE-CZ-NH2	-5.21	117.69	120.30
2	C	193	ASP	CB-CG-OD2	5.18	122.97	118.30

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	3526	0	3507	21	0
1	B	3494	0	3462	26	0
2	C	1250	0	1218	13	0
2	D	1255	0	1224	9	0
3	E	28	0	25	0	0
3	F	28	0	25	0	0
4	A	42	0	39	3	0
4	B	42	0	39	0	0
5	A	13	0	18	0	0
5	B	13	0	18	0	0
6	A	206	0	0	2	0
6	B	142	0	0	5	0
6	C	66	0	0	1	0
6	D	51	0	0	0	0
All	All	10156	0	9575	64	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 3.

All (64) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:212:ARG:O	2:C:216[A]:THR:HG23	1.74	0.87
1:B:296:THR:HG23	6:B:1039:HOH:O	1.78	0.84
1:A:65[A]:THR:HG23	4:A:801:NAG:HN2	1.43	0.82
1:B:245:MET:HE2	6:B:909:HOH:O	1.81	0.80
1:A:147:SER:O	1:A:172:ARG:NH1	2.17	0.77
6:A:966:HOH:O	2:C:216[A]:THR:HG22	1.89	0.73
1:B:149:GLU:OE2	1:B:172:ARG:NH2	2.23	0.71
1:A:149:GLU:OE2	1:A:172:ARG:NH2	2.24	0.69
1:A:290:GLU:OE2	1:A:314:HIS:ND1	2.23	0.68
2:D:183[A]:THR:HG23	2:D:188:GLN:NE2	2.10	0.67
1:B:147:SER:O	1:B:172:ARG:NH1	2.31	0.64
1:B:271:SER:O	1:B:272:SER:HB3	2.03	0.59
1:B:214:GLU:HG2	6:B:915:HOH:O	2.03	0.57

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:52:VAL:CG2	1:B:73:LEU:HD21	2.35	0.56
1:A:64:GLU:HB3	4:A:801:NAG:H82	1.88	0.55
1:B:52:VAL:HG23	1:B:73:LEU:HD21	1.90	0.54
2:C:77:GLU:O	2:C:81:THR:HG23	2.08	0.54
2:D:77:GLU:O	2:D:81:THR:HG23	2.09	0.53
1:B:296:THR:HB	1:B:320:GLN:HB2	1.89	0.53
2:D:183[A]:THR:HG23	2:D:188:GLN:HE21	1.73	0.53
1:B:290:GLU:OE2	1:B:314:HIS:ND1	2.36	0.53
1:A:155:ASP:OD2	2:C:203:ARG:NH2	2.42	0.53
2:D:106[B]:GLU:CD	2:D:106[B]:GLU:H	2.13	0.52
1:B:187[B]:CYS:SG	1:B:189:GLU:HG2	2.50	0.51
1:B:296:THR:HG22	1:B:320:GLN:OE1	2.10	0.51
2:C:62:ALA:HB1	2:C:150:ILE:O	2.12	0.50
1:A:154:ARG:HD3	6:A:1088:HOH:O	2.12	0.49
1:A:155:ASP:CG	2:C:203:ARG:HH22	2.16	0.49
1:A:330:LYS:HE2	1:A:332:ASP:OD1	2.14	0.48
1:B:186:ILE:O	1:B:186:ILE:HG13	2.14	0.47
1:B:179:THR:HA	1:B:205:SER:O	2.14	0.47
1:A:161:ILE:O	1:A:190:ASP:HB3	2.15	0.47
1:B:386:GLY:O	1:B:387:LEU:HD23	2.15	0.47
1:B:52:VAL:HG23	1:B:73:LEU:CD2	2.45	0.47
1:B:155:ASP:OD2	2:D:203:ARG:NH2	2.48	0.47
1:A:65[A]:THR:HG23	4:A:801:NAG:N2	2.20	0.46
2:C:168:THR:OG1	2:C:173[A]:ASN:ND2	2.49	0.46
2:C:183[A]:THR:HG21	6:C:308:HOH:O	2.15	0.46
2:D:62:ALA:HB1	2:D:150:ILE:O	2.15	0.46
1:A:287:LYS:HA	1:A:314:HIS:CG	2.52	0.45
1:A:187[B]:CYS:SG	1:A:189:GLU:HG2	2.57	0.45
2:D:86:ILE:HG21	2:D:122:ILE:HG13	1.99	0.44
1:B:181:ASN:O	1:B:207:ILE:HA	2.16	0.44
1:B:332:ASP:CB	6:B:1031:HOH:O	2.65	0.44
2:C:183[A]:THR:HG23	2:C:188:GLN:NE2	2.33	0.44
1:B:265:ASN:HA	1:B:301:LYS:O	2.18	0.43
1:A:208:THR:HG22	1:A:210[B]:GLN:HB3	2.00	0.43
1:A:52:VAL:HG23	1:A:73:LEU:HD21	1.99	0.43
1:B:208:THR:HG22	1:B:210[B]:GLN:HB3	2.00	0.43
1:B:80:GLN:HA	1:B:105:TYR:O	2.19	0.42
1:A:155:ASP:CG	2:C:203:ARG:NH2	2.73	0.42
1:A:179:THR:HA	1:A:205:SER:O	2.20	0.42
1:A:52:VAL:CG2	1:A:73:LEU:HD21	2.49	0.42
2:C:158:LEU:HD23	2:C:213:LEU:HD21	2.01	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:231:ILE:O	1:A:258:ILE:HA	2.20	0.42
1:A:37:ARG:NH1	2:C:219:ASN:OD1	2.53	0.41
2:D:69:GLY:HA2	2:D:148:PHE:CZ	2.55	0.41
1:A:186:ILE:O	1:A:186:ILE:HG13	2.19	0.41
1:B:164:ALA:HB3	1:B:166:PHE:CE2	2.56	0.41
1:B:221:LYS:O	1:B:221:LYS:CG	2.68	0.41
1:B:454:SER:OG	1:B:456:LYS:HG3	2.21	0.41
2:D:168:THR:OG1	2:D:173[A]:ASN:ND2	2.53	0.40
1:B:296:THR:CG2	6:B:1039:HOH:O	2.51	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	446/455 (98%)	427 (96%)	19 (4%)	0	100	100
1	B	444/455 (98%)	419 (94%)	25 (6%)	0	100	100
2	C	172/183 (94%)	170 (99%)	2 (1%)	0	100	100
2	D	170/183 (93%)	168 (99%)	2 (1%)	0	100	100
All	All	1232/1276 (97%)	1184 (96%)	48 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	A	400/409 (98%)	396 (99%)	4 (1%)	76	82
1	B	392/409 (96%)	387 (99%)	5 (1%)	69	75
2	C	132/143 (92%)	128 (97%)	4 (3%)	41	44
2	D	132/143 (92%)	126 (96%)	6 (4%)	27	27
All	All	1056/1104 (96%)	1037 (98%)	19 (2%)	62	65

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	75	PHE
1	A	93	ARG
1	A	171	ARG
1	A	221	LYS
2	C	67	GLN
2	C	68	ASP
2	C	183[A]	THR
2	C	183[B]	THR
1	B	75	PHE
1	B	93	ARG
1	B	171	ARG
1	B	221	LYS
1	B	296	THR
2	D	55	GLN
2	D	56	ILE
2	D	68	ASP
2	D	143	THR
2	D	183[A]	THR
2	D	183[B]	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (15) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	107	GLN
2	C	67	GLN
2	C	132	ASN
2	C	149	GLN
2	C	157	GLN
2	C	210	GLN
1	B	80	GLN
1	B	107	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	D	67	GLN
2	D	132	ASN
2	D	149	GLN
2	D	157	GLN
2	D	159	ASN
2	D	210	GLN
2	D	211	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

2 non-standard protein/DNA/RNA residues are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	NAG	A	801	1	14,14,15	0.66	0	17,19,21	1.42	3 (17%)
4	NAG	B	801	1	14,14,15	0.35	0	17,19,21	1.34	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
4	NAG	A	801	1	-	0/6/23/26	0/1/1/1
4	NAG	B	801	1	-	2/6/23/26	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})
4	A	801	NAG	C2-N2-C7	3.01	127.19	122.90
4	A	801	NAG	C1-O5-C5	2.96	116.20	112.19
4	B	801	NAG	O5-C5-C6	2.92	111.77	107.20
4	A	801	NAG	O7-C7-C8	-2.36	117.67	122.06
4	B	801	NAG	C1-C2-N2	-2.25	106.64	110.49

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	801	NAG	O5-C5-C6-O6
4	B	801	NAG	C1-C2-N2-C7

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	801	NAG	3	0

5.5 Carbohydrates (i)

4 monosaccharides are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	NAG	E	1	1,3	14,14,15	1.13	1 (7%)	17,19,21	1.51	3 (17%)
3	NAG	E	2	3	14,14,15	0.75	0	17,19,21	1.12	1 (5%)
3	NAG	F	1	1,3	14,14,15	0.99	1 (7%)	17,19,21	1.11	1 (5%)
3	NAG	F	2	3	14,14,15	0.84	1 (7%)	17,19,21	2.22	3 (17%)

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	E	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	E	2	3	-	2/6/23/26	0/1/1/1
3	NAG	F	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	F	2	3	-	2/6/23/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	1	NAG	C1-C2	3.02	1.56	1.52
3	E	1	NAG	O5-C5	-2.83	1.37	1.43
3	F	2	NAG	O5-C1	-2.20	1.40	1.43

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	2	NAG	C1-O5-C5	6.16	120.53	112.19
3	F	2	NAG	C4-C3-C2	-4.31	104.70	111.02
3	F	2	NAG	O5-C5-C6	4.14	113.70	107.20
3	E	1	NAG	O4-C4-C5	-2.98	101.90	109.30
3	E	1	NAG	O5-C1-C2	-2.80	106.87	111.29
3	F	1	NAG	C1-O5-C5	-2.75	108.47	112.19
3	E	2	NAG	O5-C5-C6	2.35	110.88	107.20
3	E	1	NAG	O6-C6-C5	-2.12	104.03	111.29

There are no chirality outliers.

All (4) torsion outliers are listed below:

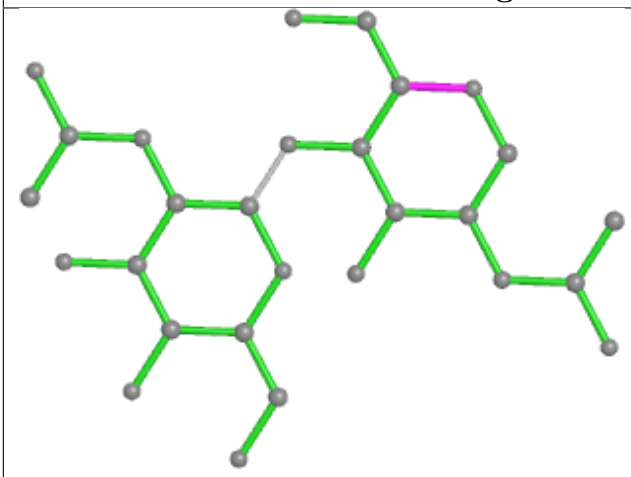
Mol	Chain	Res	Type	Atoms
3	F	2	NAG	C4-C5-C6-O6
3	F	2	NAG	O5-C5-C6-O6
3	E	2	NAG	O5-C5-C6-O6
3	E	2	NAG	C4-C5-C6-O6

There are no ring outliers.

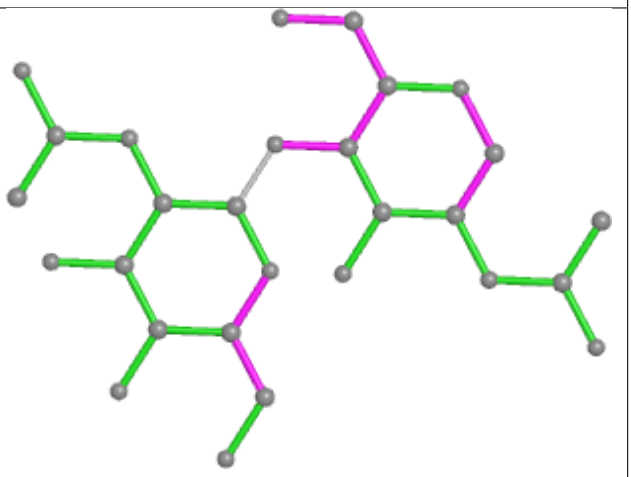
No monomer is involved in short contacts.

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

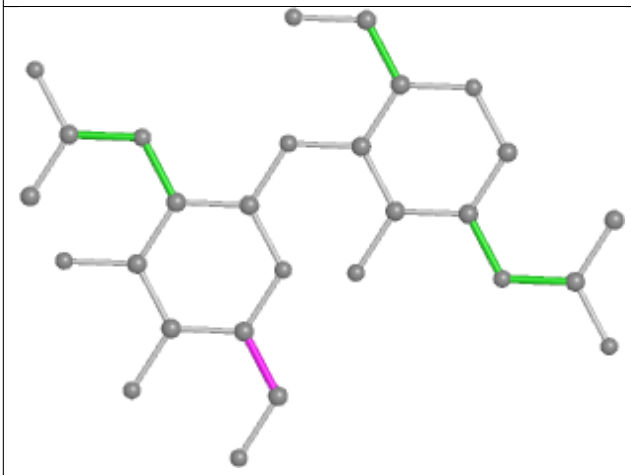
Oligosaccharide Chain E



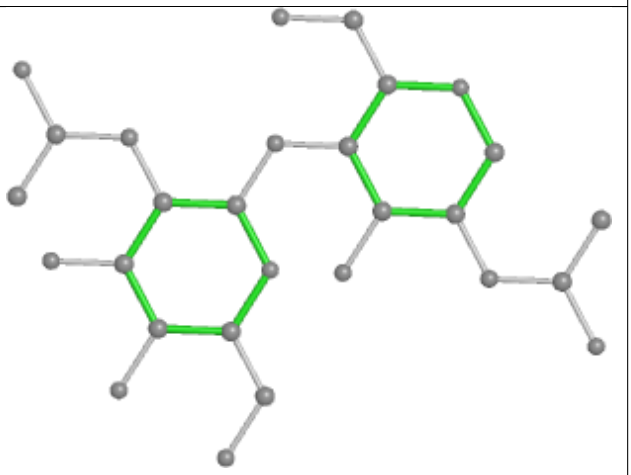
Bond lengths



Bond angles

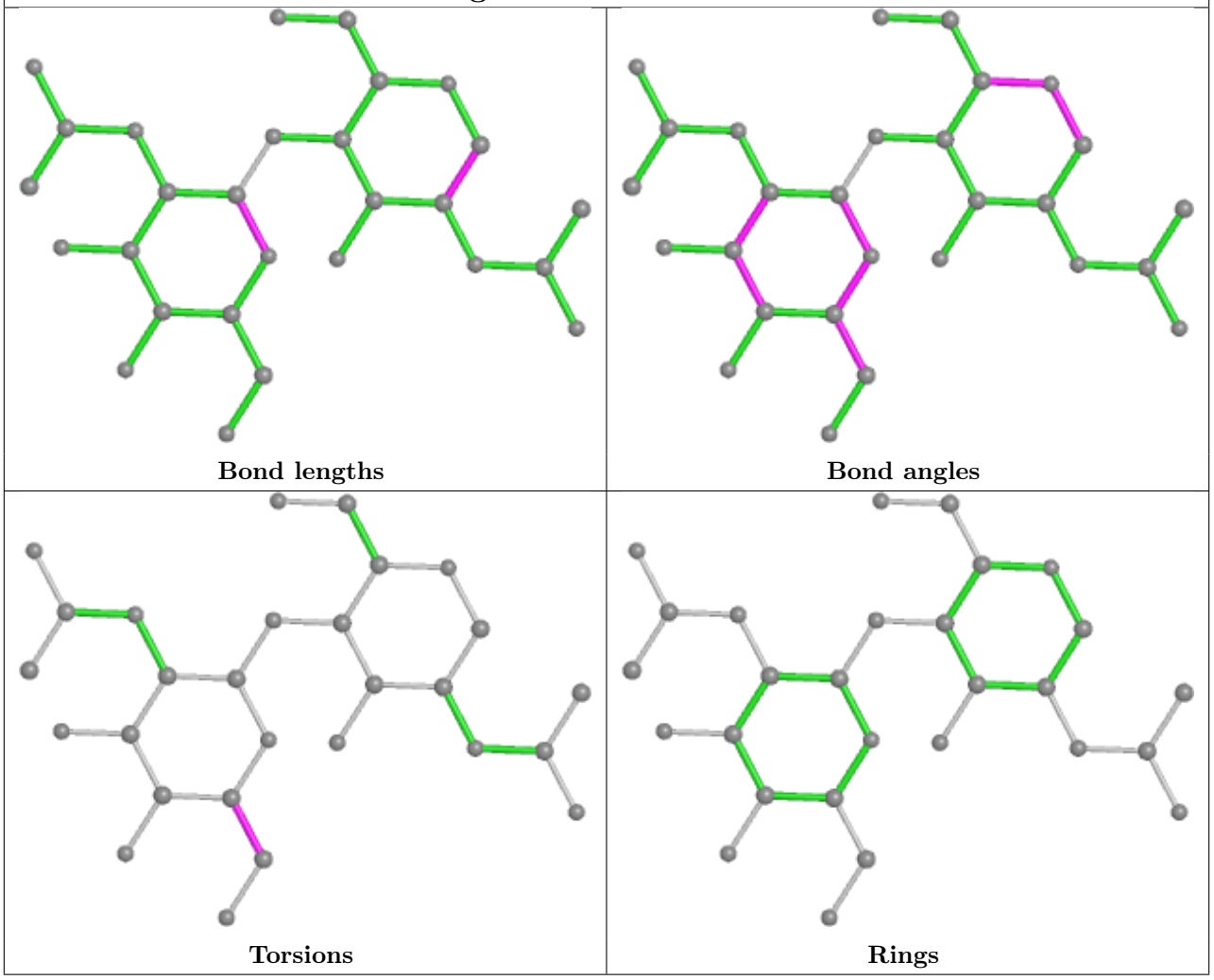


Torsions



Rings

Oligosaccharide Chain F



5.6 Ligand geometry (i)

8 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	NAG	B	802	1	14,14,15	0.62	0	17,19,21	1.44	2 (11%)
4	NAG	A	803	1	14,14,15	0.78	0	17,19,21	1.32	2 (11%)
4	NAG	A	802	1	14,14,15	0.59	0	17,19,21	1.71	4 (23%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	PG4	A	806	-	12,12,12	0.53	0	11,11,11	0.52	0
5	PG4	B	806	-	12,12,12	0.62	0	11,11,11	0.49	0
4	NAG	B	803	1	14,14,15	0.96	1 (7%)	17,19,21	1.45	2 (11%)
4	NAG	A	801	1	14,14,15	0.66	0	17,19,21	1.42	3 (17%)
4	NAG	B	801	1	14,14,15	0.35	0	17,19,21	1.34	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
4	NAG	B	802	1	-	2/6/23/26	0/1/1/1
4	NAG	A	803	1	-	0/6/23/26	0/1/1/1
4	NAG	A	802	1	-	2/6/23/26	0/1/1/1
5	PG4	A	806	-	-	6/10/10/10	-
5	PG4	B	806	-	-	5/10/10/10	-
4	NAG	B	803	1	-	0/6/23/26	0/1/1/1
4	NAG	A	801	1	-	0/6/23/26	0/1/1/1
4	NAG	B	801	1	-	2/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	803	NAG	C1-C2	2.73	1.56	1.52

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	803	NAG	C1-O5-C5	4.71	118.57	112.19
4	A	802	NAG	C4-C3-C2	3.70	116.45	111.02
4	B	802	NAG	O5-C5-C6	3.41	112.56	107.20
4	A	802	NAG	C1-O5-C5	3.20	116.53	112.19
4	A	802	NAG	O5-C5-C6	3.12	112.09	107.20
4	A	801	NAG	C2-N2-C7	3.01	127.19	122.90
4	A	801	NAG	C1-O5-C5	2.96	116.20	112.19
4	B	801	NAG	O5-C5-C6	2.92	111.77	107.20
4	A	803	NAG	O6-C6-C5	-2.66	102.17	111.29
4	B	802	NAG	C1-O5-C5	2.45	115.52	112.19
4	A	803	NAG	C1-O5-C5	2.38	115.42	112.19

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	801	NAG	O7-C7-C8	-2.36	117.67	122.06
4	B	801	NAG	C1-C2-N2	-2.25	106.64	110.49
4	A	802	NAG	O3-C3-C2	-2.16	105.00	109.47
4	B	803	NAG	O4-C4-C5	2.08	114.47	109.30

There are no chirality outliers.

All (17) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	806	PG4	O2-C3-C4-O3
4	A	802	NAG	O5-C5-C6-O6
5	A	806	PG4	O3-C5-C6-O4
5	A	806	PG4	O2-C3-C4-O3
4	B	802	NAG	O5-C5-C6-O6
4	A	802	NAG	C4-C5-C6-O6
4	B	802	NAG	C4-C5-C6-O6
5	A	806	PG4	C3-C4-O3-C5
5	B	806	PG4	C3-C4-O3-C5
5	B	806	PG4	C1-C2-O2-C3
5	A	806	PG4	C1-C2-O2-C3
5	A	806	PG4	C8-C7-O4-C6
4	B	801	NAG	O5-C5-C6-O6
4	B	801	NAG	C1-C2-N2-C7
5	B	806	PG4	O1-C1-C2-O2
5	A	806	PG4	O4-C7-C8-O5
5	B	806	PG4	O4-C7-C8-O5

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	801	NAG	3	0

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	A	442/455 (97%)	-0.15	5 (1%) 80 84	18, 28, 42, 55	0
1	B	441/455 (96%)	-0.02	8 (1%) 68 72	19, 33, 56, 67	0
2	C	169/183 (92%)	0.09	15 (8%) 9 12	16, 30, 67, 82	0
2	D	169/183 (92%)	0.20	13 (7%) 13 17	17, 32, 69, 83	0
All	All	1221/1276 (95%)	-0.02	41 (3%) 45 51	16, 30, 56, 83	0

All (41) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	152	ALA	9.7
2	D	154	ALA	9.3
2	D	152	ALA	6.6
1	A	428	CYS	5.6
2	C	154	ALA	5.3
2	C	151	GLY	5.1
2	D	155	GLY	4.8
2	D	153	ASN	4.4
1	B	30	TYR	4.2
1	B	28	ILE	4.1
2	C	59	LEU	3.9
2	D	103	GLN	3.9
2	C	153	ASN	3.9
2	D	54	GLY	3.6
2	C	56	ILE	3.6
1	B	362	ASN	3.6
2	C	102	THR	3.5
2	D	157	GLN	3.4
1	A	464	PRO	3.3
2	C	143	THR	3.3
1	A	429	PRO	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	55	GLN	3.0
2	D	143	THR	2.9
2	D	144	ASP	2.9
2	C	156	GLN	2.6
2	C	155	GLY	2.5
2	C	221	GLY	2.5
2	D	56	ILE	2.5
1	B	27	VAL	2.5
2	D	102	THR	2.4
2	C	55	GLN	2.4
1	B	217	LEU	2.4
2	D	156	GLN	2.4
1	A	463	CYS	2.3
1	B	93	ARG	2.3
2	C	157	GLN	2.2
1	B	340	THR	2.2
2	C	103	GLN	2.1
1	B	429	PRO	2.1
1	A	462	ILE	2.0
2	C	144	ASP	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (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^{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
4	NAG	B	801	14/15	0.68	0.22	48,57,60,61	0
4	NAG	A	801	14/15	0.81	0.26	42,50,53,56	0

6.3 Carbohydrates (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{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	NAG	E	2	14/15	0.75	0.30	40,45,50,51	0

Continued on next page...

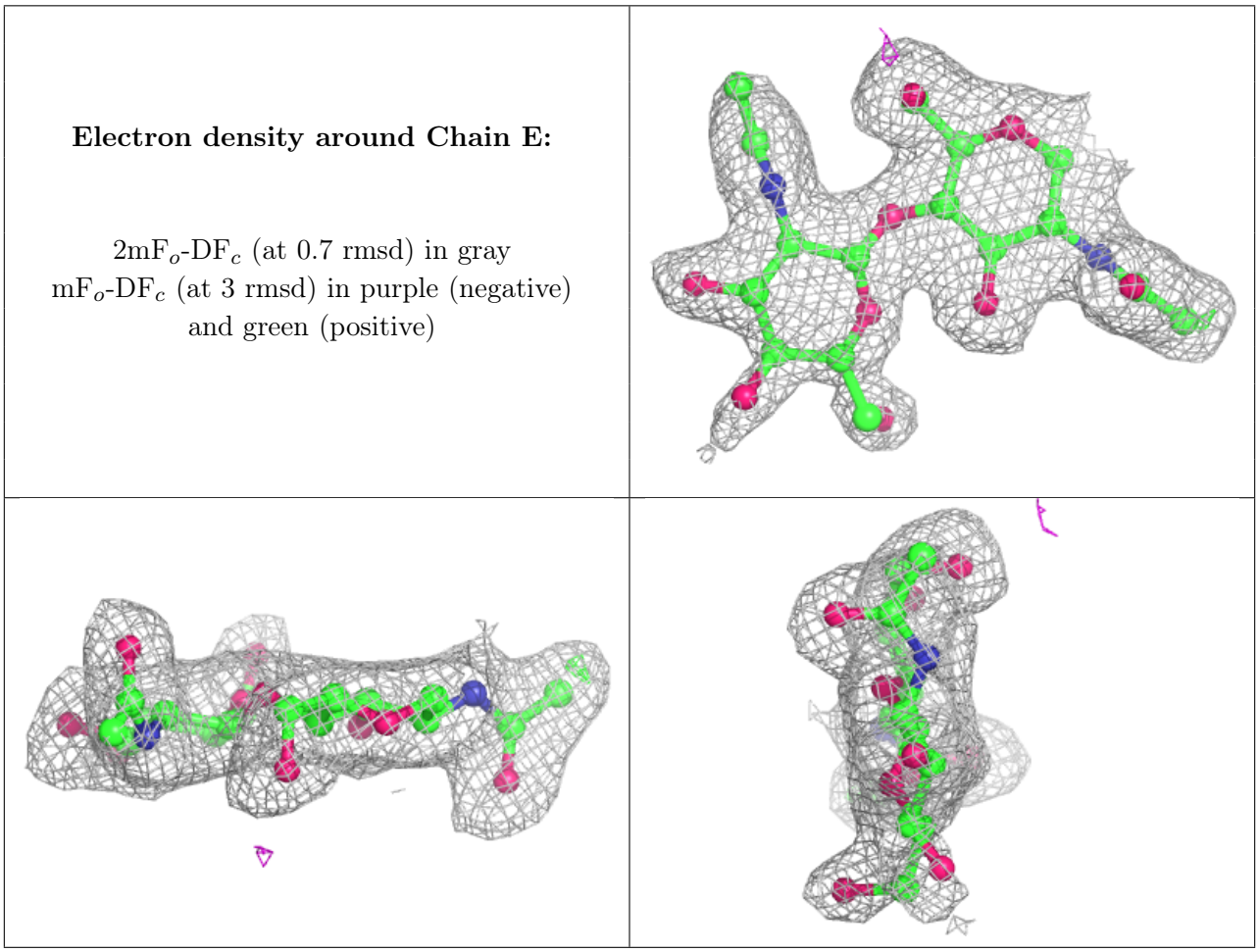
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	NAG	F	2	14/15	0.76	0.34	46,51,57,57	0
3	NAG	F	1	14/15	0.93	0.09	26,31,34,40	0
3	NAG	E	1	14/15	0.95	0.08	19,22,29,33	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.

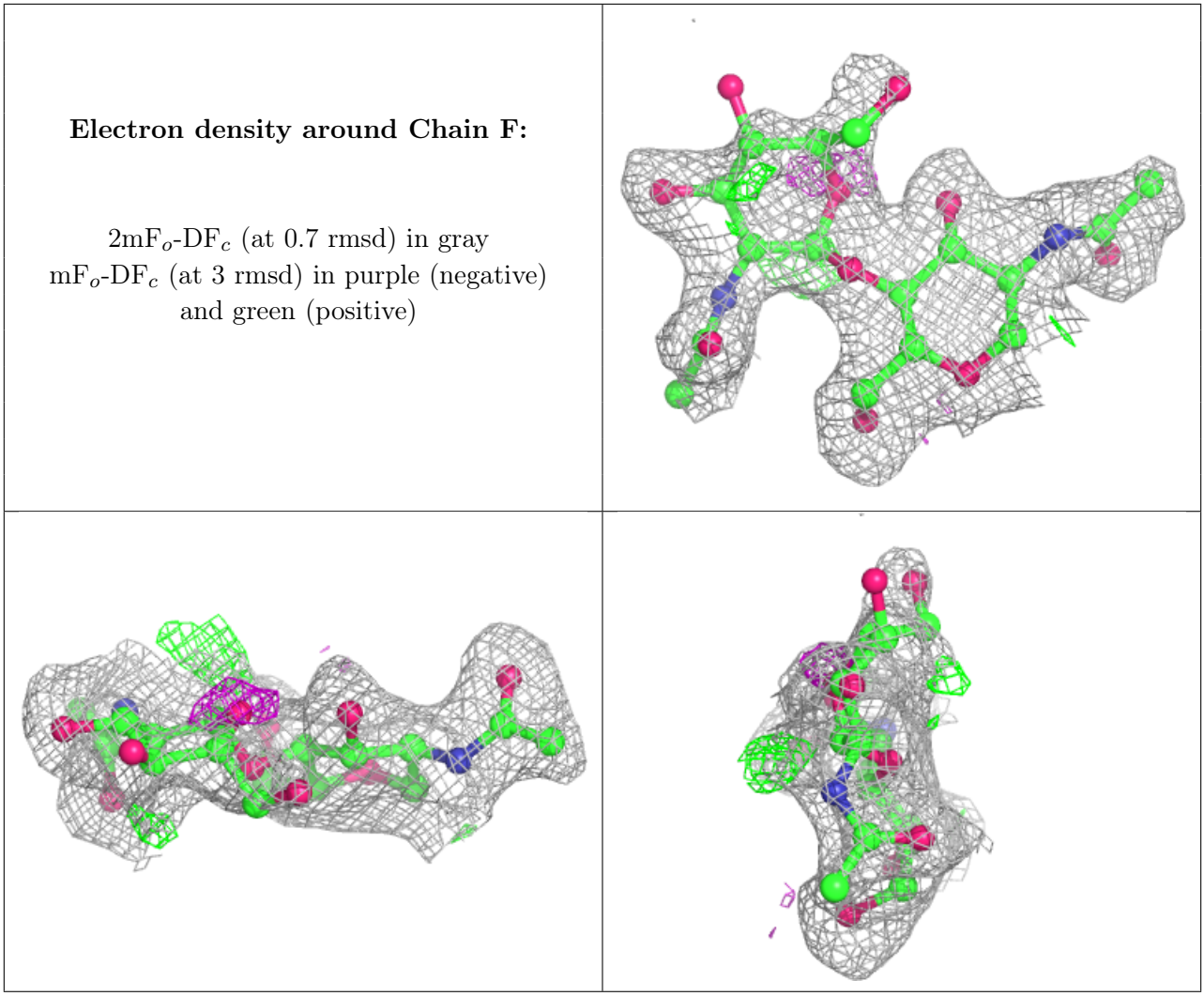
Electron density around Chain E:

$2mF_o$-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 F:

$2mF_o$-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
4	NAG	A	803	14/15	0.67	0.45	45,53,57,57	0
4	NAG	B	801	14/15	0.68	0.22	48,57,60,61	0
4	NAG	B	803	14/15	0.70	0.38	44,53,57,57	0
4	NAG	A	802	14/15	0.76	0.31	46,57,59,60	0
4	NAG	A	801	14/15	0.81	0.26	42,50,53,56	0
5	PG4	A	806	13/13	0.86	0.15	46,49,53,53	0
5	PG4	B	806	13/13	0.87	0.16	41,45,49,50	0
4	NAG	B	802	14/15	0.88	0.15	46,56,58,60	0

6.5 Other polymers (i)

There are no such residues in this entry.