Full wwPDB X-ray Structure Validation Report (i)

Oct 29, 2024 – 02:21 PM EDT

PDB ID	:	3V47
Title	:	Crystal structure of the N-terminal fragment of zebrafish TLR5 in complex with Salmonella flagellin
Authors	:	Yoon, S.I.; Hong, H.; Wilson, I.A.
Deposited on	:	2011-12-14
Resolution	:	2.47 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:

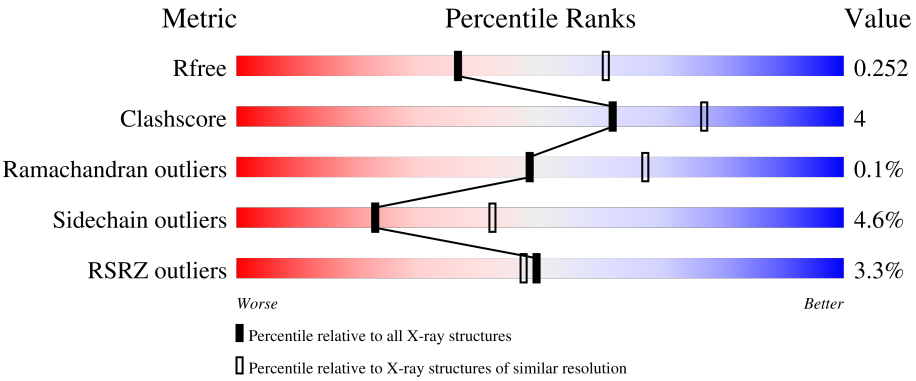
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.47 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}	164625	7106 (2.50-2.46)
Clashscore	180529	7991 (2.50-2.46)
Ramachandran outliers	177936	7888 (2.50-2.46)
Sidechain outliers	177891	7890 (2.50-2.46)
RSRZ outliers	164620	7106 (2.50-2.46)

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	86% 10%
1	B	455	85% 11%
2	C	425	5% 61% 7% 31%
2	D	425	5% 56% 9% 33%
3	E	2	100%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	2	100%
3	G	2	100%
3	H	2	100%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 11462 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 Toll-like receptor 5b and variable lymphocyte receptor B.61 chimeric protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	441	Total	C	N	O	S	0	0	0
			3459	2214	579	650	16			
1	B	441	Total	C	N	O	S	0	0	0
			3452	2208	576	652	16			

There are 34 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	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	24	GLU	VAL	engineered mutation	UNP B3DIN1
B	124	VAL	LEU	engineered mutation	UNP B3DIN1
B	159	LYS	GLN	engineered mutation	UNP B3DIN1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
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	295	Total	C	N	O	S	0	0	0
			2078	1258	368	450	2			
2	D	284	Total	C	N	O	S	0	0	0
			2028	1230	361	435	2			

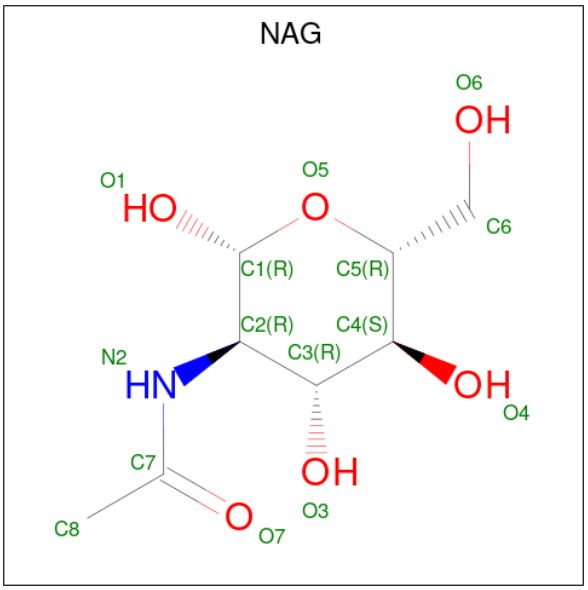
There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	41	GLY	-	expression tag	UNP Q06971
C	42	SER	-	expression tag	UNP Q06971
C	43	ALA	-	expression tag	UNP Q06971
C	44	LYS	-	expression tag	UNP Q06971
C	45	ASP	-	expression tag	UNP Q06971
C	46	PRO	-	expression tag	UNP Q06971
D	41	GLY	-	expression tag	UNP Q06971
D	42	SER	-	expression tag	UNP Q06971
D	43	ALA	-	expression tag	UNP Q06971
D	44	LYS	-	expression tag	UNP Q06971
D	45	ASP	-	expression tag	UNP Q06971
D	46	PRO	-	expression tag	UNP Q06971

- 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
3	G	2	Total C N O 28 16 2 10	0	0	0
3	H	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	B	1	Total C N O 14 8 1 5	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	111	Total O 111 111	0	0
5	B	110	Total O 110 110	0	0

Continued on next page...

Continued from previous page...

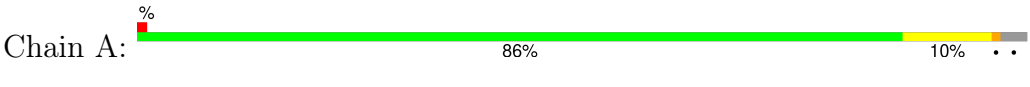
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	40	Total O 40 40	0	0
5	D	30	Total O 30 30	0	0

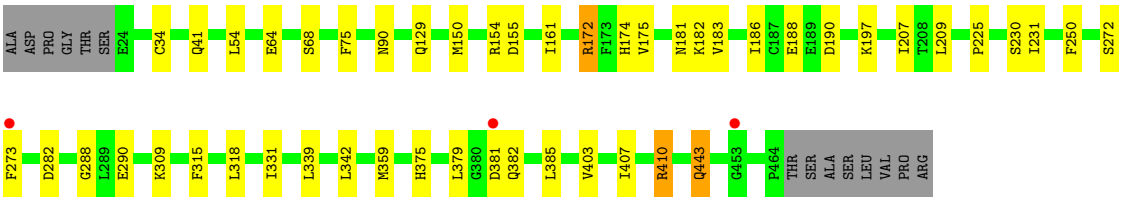
3 Residue-property plots (i)

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: Toll-like receptor 5b and variable lymphocyte receptor B.61 chimeric protein

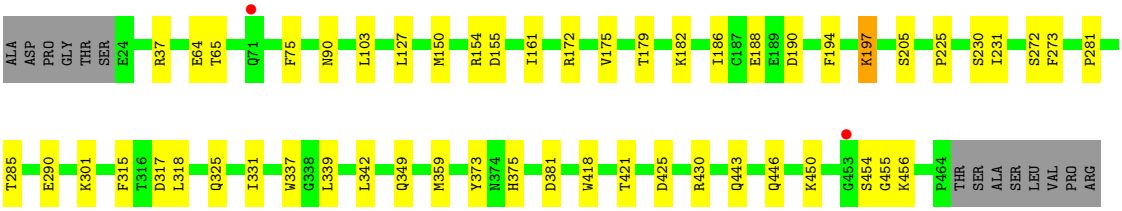
Chain A:





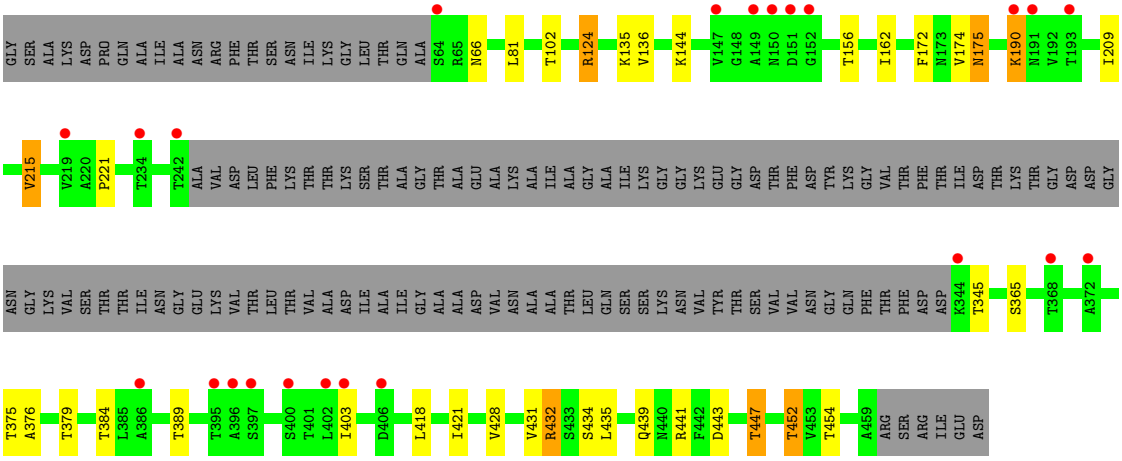
• Molecule 1: Toll-like receptor 5b and variable lymphocyte receptor B.61 chimeric protein

Chain B:



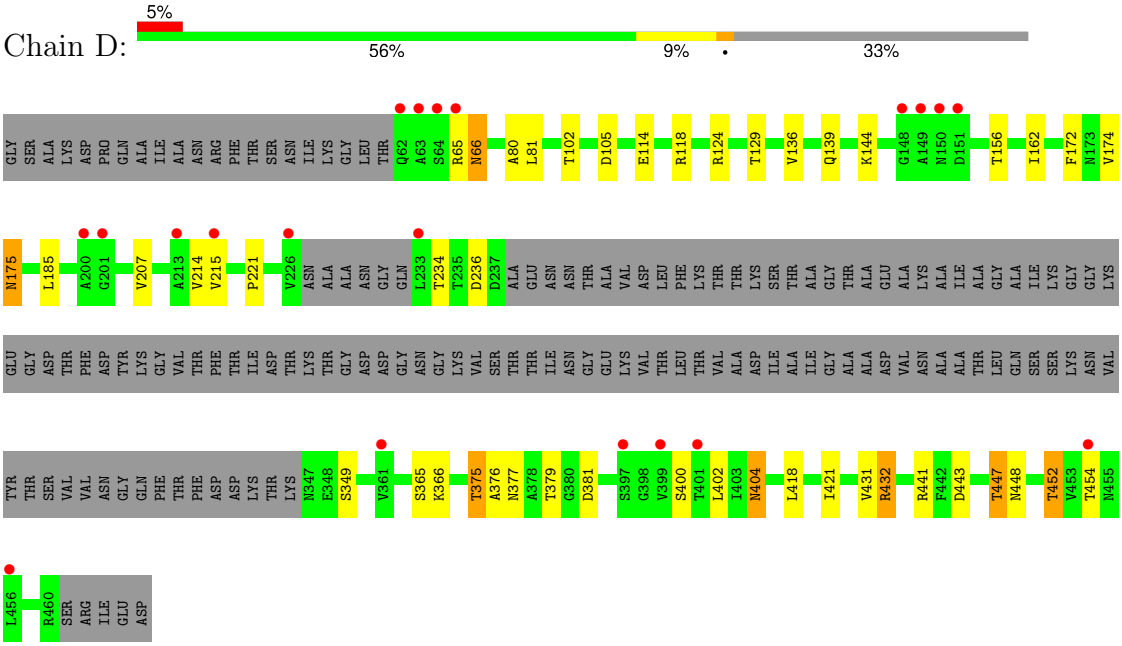
• Molecule 2: Flagellin

Chain C:



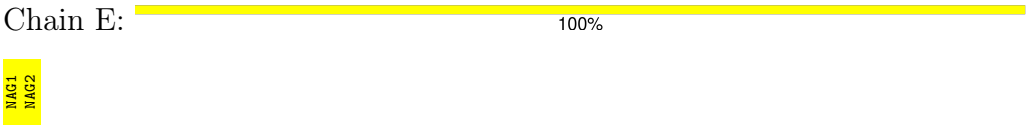
• Molecule 2: Flagellin

Chain D:



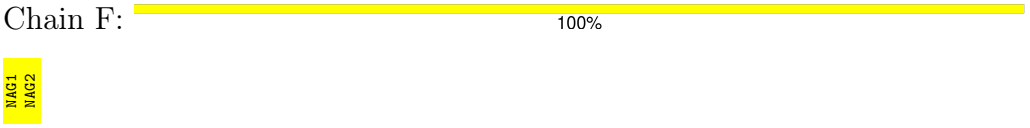
• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E:



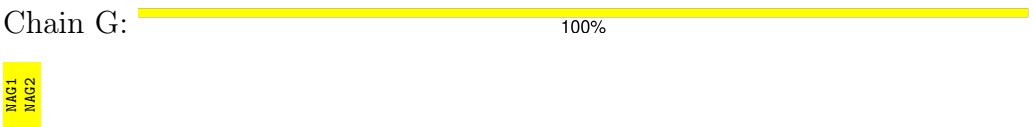
• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F:



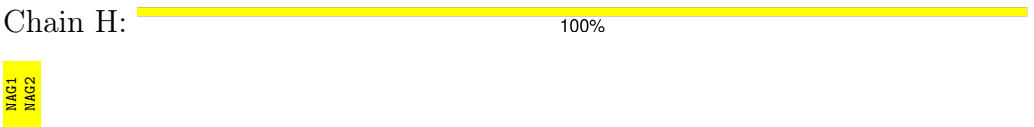
• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain G:



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

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	58.41Å 181.50Å 186.37Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.47 20.00 – 2.47	Depositor EDS
% Data completeness (in resolution range)	96.7 (20.00-2.47) 96.4 (20.00-2.47)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.81 (at 2.47Å)	Xtriage
Refinement program	REFMAC 5.5.0110	Depositor
R, R_{free}	0.221 , 0.259 0.215 , 0.252	Depositor DCC
R_{free} test set	3515 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	38.3	Xtriage
Anisotropy	0.313	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 33.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.032 for -h,l,k	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	11462	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.28% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0.72	0/3528	0.73	3/4775 (0.1%)
1	B	0.69	0/3521	0.72	2/4768 (0.0%)
2	C	0.64	2/2089 (0.1%)	0.75	6/2845 (0.2%)
2	D	0.62	2/2038 (0.1%)	0.77	6/2769 (0.2%)
All	All	0.68	4/11176 (0.0%)	0.74	17/15157 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	136	VAL	CB-CG2	-7.09	1.38	1.52
2	D	136	VAL	CB-CG1	-6.83	1.38	1.52
2	C	136	VAL	CB-CG1	-6.72	1.38	1.52
2	C	136	VAL	CB-CG2	-6.07	1.40	1.52

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	432	ARG	NE-CZ-NH1	-10.95	114.83	120.30
2	C	124	ARG	NE-CZ-NH1	10.28	125.44	120.30
2	C	124	ARG	NE-CZ-NH2	-9.54	115.53	120.30
2	D	136	VAL	CG1-CB-CG2	-9.26	96.09	110.90
2	D	124	ARG	NE-CZ-NH2	8.78	124.69	120.30
2	C	432	ARG	NE-CZ-NH1	-8.69	115.95	120.30
2	D	124	ARG	NE-CZ-NH1	-8.55	116.02	120.30
2	C	432	ARG	NE-CZ-NH2	7.74	124.17	120.30
2	C	136	VAL	CG1-CB-CG2	-7.40	99.06	110.90
1	A	154	ARG	NE-CZ-NH2	-7.33	116.64	120.30
1	A	172	ARG	NE-CZ-NH1	7.25	123.92	120.30
2	D	432	ARG	NE-CZ-NH2	6.47	123.54	120.30
2	D	432	ARG	CG-CD-NE	-5.79	99.63	111.80

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	124	ARG	CD-NE-CZ	5.78	131.69	123.60
1	B	154	ARG	NE-CZ-NH2	-5.43	117.58	120.30
1	B	172	ARG	NE-CZ-NH1	5.40	123.00	120.30
1	A	282	ASP	CB-CG-OD1	5.04	122.84	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	3459	0	3416	26	0
1	B	3452	0	3397	29	1
2	C	2078	0	1933	16	0
2	D	2028	0	1917	25	1
3	E	28	0	25	0	0
3	F	28	0	25	0	0
3	G	28	0	25	0	0
3	H	28	0	25	0	0
4	A	28	0	26	0	0
4	B	14	0	13	0	0
5	A	111	0	0	0	0
5	B	110	0	0	0	0
5	C	40	0	0	0	0
5	D	30	0	0	0	0
All	All	11462	0	10802	87	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 4.

All (87) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:443:GLN:OE1	1:A:443:GLN:N	2.02	0.91
2:D:404:ASN:H	2:D:404:ASN:ND2	1.68	0.91

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:404:ASN:H	2:D:404:ASN:HD22	0.89	0.89
2:D:404:ASN:HD22	2:D:404:ASN:N	1.72	0.87
1:A:375:HIS:CE1	1:B:375:HIS:HB3	2.11	0.84
1:A:155:ASP:CG	2:C:432:ARG:HH12	1.83	0.81
2:D:447:THR:HA	2:D:452:THR:HG23	1.69	0.73
2:C:162:ILE:HD11	2:C:431:VAL:HG21	1.70	0.73
2:D:215:VAL:CG1	2:D:221:PRO:HB2	2.22	0.70
2:C:190:LYS:H	2:C:190:LYS:HE2	1.57	0.68
1:A:155:ASP:OD2	2:C:432:ARG:NH1	2.26	0.68
1:B:155:ASP:CG	2:D:432:ARG:HH12	1.97	0.68
1:A:375:HIS:HB3	1:B:375:HIS:CE1	2.29	0.67
1:A:331:ILE:HD12	1:A:359:MET:HE3	1.77	0.67
1:B:64:GLU:HG2	1:B:90:ASN:HB2	1.76	0.67
2:D:377:ASN:HD21	2:D:381:ASP:HB3	1.60	0.66
2:C:447:THR:HA	2:C:452:THR:HG23	1.77	0.66
2:D:175:ASN:H	2:D:175:ASN:HD22	1.41	0.66
2:C:365:SER:HB2	2:C:376:ALA:HB3	1.79	0.65
1:A:64:GLU:HG2	1:A:90:ASN:HB2	1.80	0.64
2:D:443:ASP:O	2:D:447:THR:HG23	1.98	0.63
2:C:172:PHE:CZ	2:C:421:ILE:HD11	2.34	0.63
1:B:150:MET:HG3	1:B:175:VAL:HB	1.80	0.62
2:C:443:ASP:O	2:C:447:THR:HG23	2.02	0.59
1:A:382:GLN:HG2	1:A:410:ARG:HH21	1.67	0.59
1:A:385:LEU:HA	1:A:410:ARG:HD3	1.85	0.58
1:A:315:PHE:HB3	1:A:318:LEU:HB2	1.86	0.57
2:D:215:VAL:HG11	2:D:221:PRO:HB2	1.86	0.57
1:B:454:SER:OG	1:B:456:LYS:HG2	2.06	0.56
1:B:315:PHE:HB3	1:B:318:LEU:HB2	1.87	0.55
2:D:162:ILE:HD11	2:D:431:VAL:HG21	1.87	0.55
1:A:385:LEU:HA	1:A:410:ARG:CD	2.36	0.55
2:D:365:SER:HB2	2:D:376:ALA:HB3	1.89	0.55
2:C:215:VAL:HG13	2:C:221:PRO:HB2	1.87	0.55
1:B:381:ASP:OD2	2:C:135:LYS:HE2	2.07	0.54
1:B:155:ASP:OD2	2:D:432:ARG:NH1	2.41	0.54
2:D:172:PHE:CZ	2:D:421:ILE:HD11	2.43	0.54
1:B:37:ARG:NH1	2:D:448:ASN:OD1	2.41	0.54
2:D:366:LYS:HE3	2:D:375:THR:HB	1.90	0.53
1:A:155:ASP:CG	2:C:432:ARG:NH1	2.58	0.53
1:B:450:LYS:HG3	1:B:455:GLY:O	2.09	0.52
2:D:215:VAL:HG12	2:D:221:PRO:HB2	1.92	0.52
1:A:181:ASN:O	1:A:207:ILE:HA	2.10	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:331:ILE:HD12	1:B:359:MET:HE3	1.93	0.51
2:D:207:VAL:HG22	2:D:214:VAL:HG22	1.93	0.51
1:A:129:GLN:HA	1:A:155:ASP:O	2.10	0.50
1:A:339:LEU:HD13	1:A:342:LEU:HD22	1.93	0.50
1:B:443:GLN:H	1:B:443:GLN:CD	2.15	0.49
1:B:339:LEU:HD13	1:B:342:LEU:HD22	1.95	0.48
1:B:161:ILE:O	1:B:190:ASP:HB3	2.12	0.48
2:D:66:ASN:HD22	2:D:66:ASN:HA	1.57	0.46
2:D:234:THR:OG1	2:D:236:ASP:HB2	2.16	0.46
1:B:230:SER:C	1:B:231:ILE:HG13	2.34	0.46
2:C:435:LEU:O	2:C:439:GLN:HG3	2.15	0.46
2:C:144:LYS:HG2	2:C:156:THR:HG22	1.97	0.46
1:B:349:GLN:HA	1:B:373:TYR:O	2.15	0.45
2:D:144:LYS:HG2	2:D:156:THR:HG22	1.98	0.45
1:A:34:CYS:HB2	1:A:54:LEU:HD23	1.97	0.45
1:A:150:MET:HG3	1:A:175:VAL:HB	1.97	0.45
1:B:325:GLN:HA	1:B:349:GLN:O	2.16	0.45
2:D:185:LEU:HD21	2:D:402:LEU:HD21	1.99	0.45
2:D:81:LEU:HD21	2:D:162:ILE:HD13	1.97	0.45
1:B:179:THR:HA	1:B:205:SER:O	2.17	0.44
1:B:281:PRO:HB3	1:B:285:THR:HG21	2.00	0.44
1:B:186:ILE:O	1:B:186:ILE:HG13	2.16	0.44
1:A:186:ILE:HG13	1:A:186:ILE:O	2.18	0.44
1:B:425:ASP:OD2	1:B:430:ARG:NH2	2.51	0.44
1:B:443:GLN:OE1	1:B:443:GLN:N	2.47	0.44
1:B:301:LYS:HA	1:B:325:GLN:O	2.18	0.43
2:D:114:GLU:O	2:D:118:ARG:HG2	2.19	0.43
1:B:337:TRP:CD1	1:B:337:TRP:C	2.92	0.43
1:A:230:SER:C	1:A:231:ILE:HG13	2.39	0.43
1:A:161:ILE:O	1:A:190:ASP:HB3	2.18	0.43
1:B:418:TRP:CD1	1:B:446:GLN:HB2	2.53	0.42
1:A:188:GLU:HA	1:A:225:PRO:HA	2.02	0.42
2:D:80:ALA:HB2	2:D:129:THR:HG21	2.02	0.42
1:A:172:ARG:O	1:A:174:HIS:HD2	2.02	0.42
1:A:403:VAL:HB	1:A:407:ILE:HG21	2.02	0.42
1:B:194:PHE:O	1:B:197:LYS:HB2	2.19	0.42
2:C:81:LEU:HD22	2:C:428:VAL:HG13	2.02	0.42
1:B:103:LEU:HB2	1:B:127:LEU:HD23	2.03	0.41
2:C:81:LEU:HD21	2:C:162:ILE:HD13	2.02	0.41
1:B:188:GLU:HA	1:B:225:PRO:HA	2.04	0.40
2:C:175:ASN:HD22	2:C:175:ASN:H	1.69	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:250:PHE:O	1:A:288:GLY:HA3	2.21	0.40
1:A:183:VAL:O	1:A:209:LEU:HA	2.21	0.40
1:A:379:LEU:HD12	1:A:403:VAL:HG12	2.03	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:317:ASP:OD1	2:D:139:GLN:NE2[1_455]	2.16	0.04

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	439/455 (96%)	414 (94%)	24 (6%)	1 (0%)	44 62
1	B	439/455 (96%)	413 (94%)	26 (6%)	0	100 100
2	C	291/425 (68%)	281 (97%)	10 (3%)	0	100 100
2	D	278/425 (65%)	268 (96%)	10 (4%)	0	100 100
All	All	1447/1760 (82%)	1376 (95%)	70 (5%)	1 (0%)	48 67

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	381	ASP

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	386/409 (94%)	375 (97%)	11 (3%)	38 63
1	B	385/409 (94%)	377 (98%)	8 (2%)	48 72
2	C	207/339 (61%)	187 (90%)	20 (10%)	6 12
2	D	206/339 (61%)	190 (92%)	16 (8%)	10 19
All	All	1184/1496 (79%)	1129 (95%)	55 (5%)	23 42

All (55) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	41	GLN
1	A	68	SER
1	A	75	PHE
1	A	182	LYS
1	A	197	LYS
1	A	272	SER
1	A	273	PHE
1	A	290	GLU
1	A	309	LYS
1	A	410	ARG
1	A	443	GLN
1	B	65	THR
1	B	75	PHE
1	B	182	LYS
1	B	197	LYS
1	B	272	SER
1	B	273	PHE
1	B	290	GLU
1	B	421	THR
2	C	66	ASN
2	C	102	THR
2	C	124	ARG
2	C	174	VAL
2	C	175	ASN
2	C	190	LYS
2	C	209	ILE
2	C	215	VAL
2	C	345	THR
2	C	375	THR
2	C	379	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	C	384	THR
2	C	389	THR
2	C	403	ILE
2	C	418	LEU
2	C	434	SER
2	C	441	ARG
2	C	447	THR
2	C	452	THR
2	C	454	THR
2	D	65	ARG
2	D	66	ASN
2	D	102	THR
2	D	105	ASP
2	D	174	VAL
2	D	175	ASN
2	D	349	SER
2	D	375	THR
2	D	379	THR
2	D	400	SER
2	D	404	ASN
2	D	418	LEU
2	D	441	ARG
2	D	447	THR
2	D	452	THR
2	D	454	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	80	GLN
1	B	80	GLN
2	C	66	ASN
2	C	130	GLN
2	C	175	ASN
2	C	241	ASN
2	C	440	ASN
2	D	66	ASN
2	D	175	ASN
2	D	210	ASN
2	D	404	ASN
2	D	440	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

8 monosaccharides are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	NAG	E	1	3,1	14,14,15	0.58	0	17,19,21	1.62	3 (17%)
3	NAG	E	2	3	14,14,15	0.62	0	17,19,21	1.19	2 (11%)
3	NAG	F	1	3,1	14,14,15	0.66	0	17,19,21	1.32	2 (11%)
3	NAG	F	2	3	14,14,15	0.83	0	17,19,21	1.79	3 (17%)
3	NAG	G	1	3,1	14,14,15	0.67	0	17,19,21	2.07	5 (29%)
3	NAG	G	2	3	14,14,15	0.49	0	17,19,21	0.88	1 (5%)
3	NAG	H	1	3,1	14,14,15	0.74	0	17,19,21	1.15	3 (17%)
3	NAG	H	2	3	14,14,15	0.62	0	17,19,21	1.40	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
3	NAG	E	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	E	2	3	-	0/6/23/26	0/1/1/1
3	NAG	F	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	F	2	3	-	0/6/23/26	0/1/1/1
3	NAG	G	1	3,1	-	1/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	G	2	3	-	2/6/23/26	0/1/1/1
3	NAG	H	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	H	2	3	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	G	1	NAG	C1-O5-C5	5.58	119.67	112.19
3	F	2	NAG	C2-N2-C7	-4.55	116.81	122.90
3	E	1	NAG	C1-O5-C5	4.02	117.58	112.19
3	F	2	NAG	C1-O5-C5	3.65	117.08	112.19
3	E	2	NAG	C1-C2-N2	-3.32	105.19	110.43
3	G	1	NAG	O5-C1-C2	3.26	116.34	111.29
3	F	2	NAG	O4-C4-C5	2.94	116.55	109.32
3	F	1	NAG	C3-C4-C5	-2.91	104.96	110.23
3	G	1	NAG	O3-C3-C4	2.88	117.16	110.38
3	H	2	NAG	C8-C7-N2	-2.84	111.40	116.12
3	F	1	NAG	C2-N2-C7	-2.64	119.36	122.90
3	H	2	NAG	C3-C4-C5	-2.48	105.74	110.23
3	H	1	NAG	O5-C5-C4	-2.44	104.90	110.83
3	G	1	NAG	C4-C3-C2	-2.40	107.51	111.02
3	E	1	NAG	O7-C7-N2	2.12	125.72	121.98
3	G	2	NAG	C2-N2-C7	-2.10	120.09	122.90
3	H	1	NAG	O5-C1-C2	-2.06	108.11	111.29
3	G	1	NAG	O3-C3-C2	2.05	113.65	109.40
3	E	1	NAG	O3-C3-C4	2.03	115.17	110.38
3	E	2	NAG	C1-O5-C5	2.02	114.89	112.19
3	H	1	NAG	C3-C4-C5	-2.01	106.59	110.23

There are no chirality outliers.

All (3) torsion outliers are listed below:

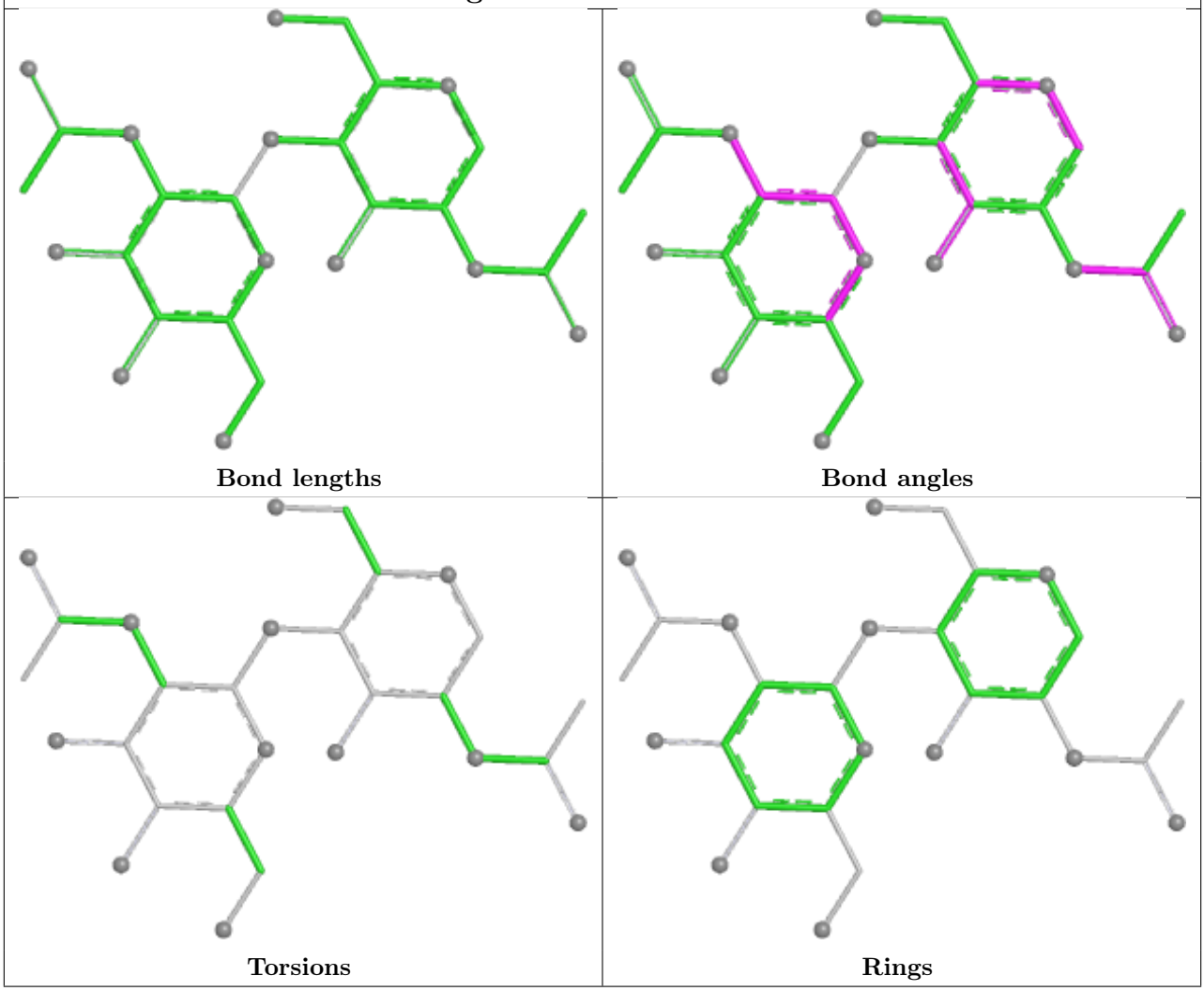
Mol	Chain	Res	Type	Atoms
3	G	2	NAG	C4-C5-C6-O6
3	G	2	NAG	O5-C5-C6-O6
3	G	1	NAG	C4-C5-C6-O6

There are no ring outliers.

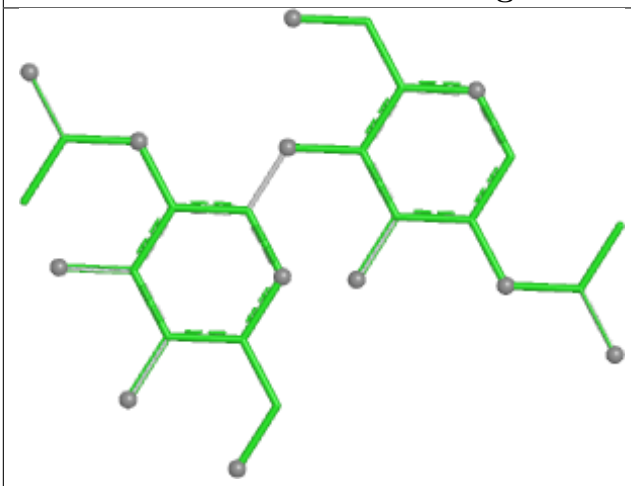
No monomer is involved in short contacts.

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

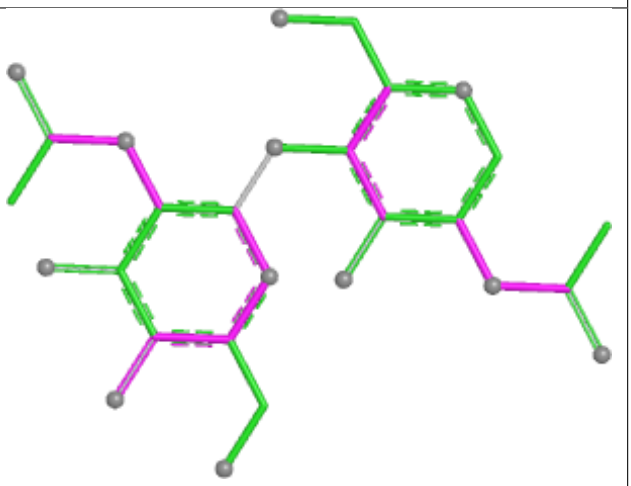
Oligosaccharide Chain E



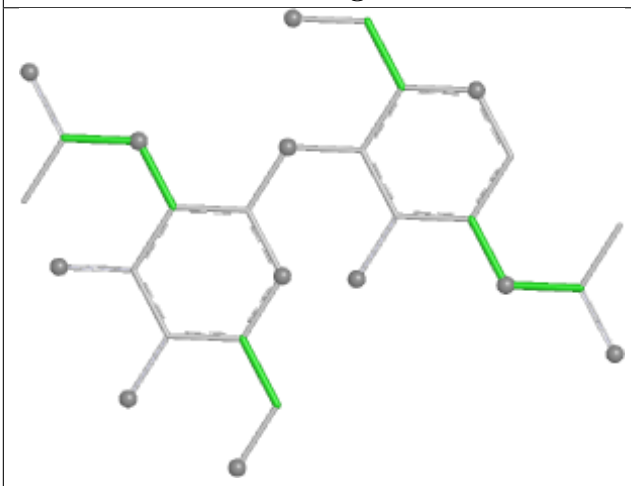
Oligosaccharide Chain F



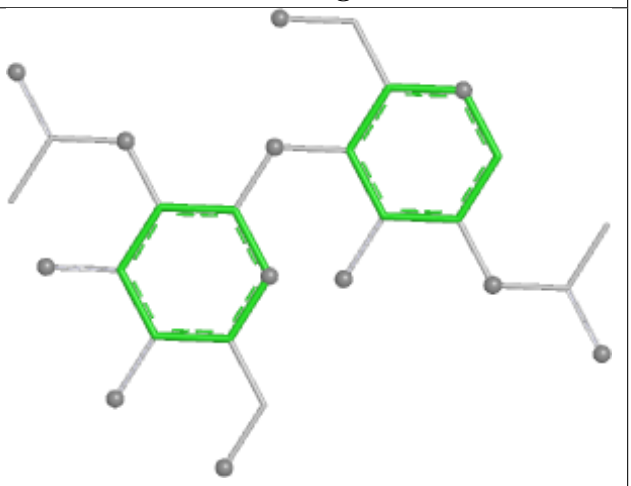
Bond lengths



Bond angles

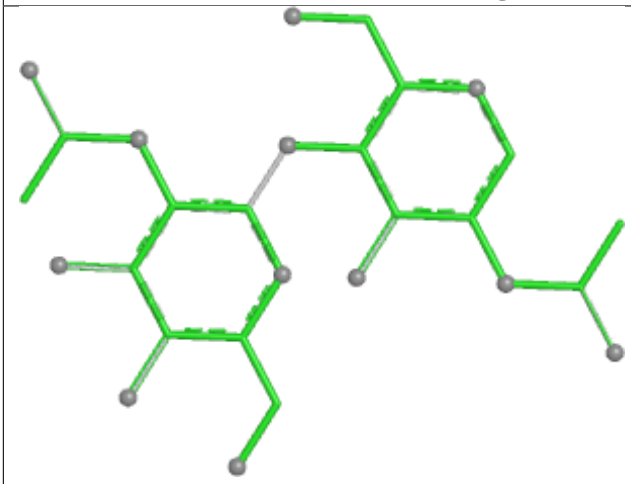


Torsions

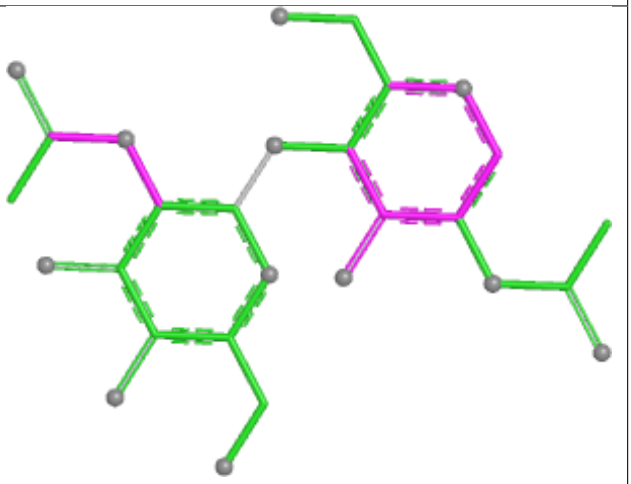


Rings

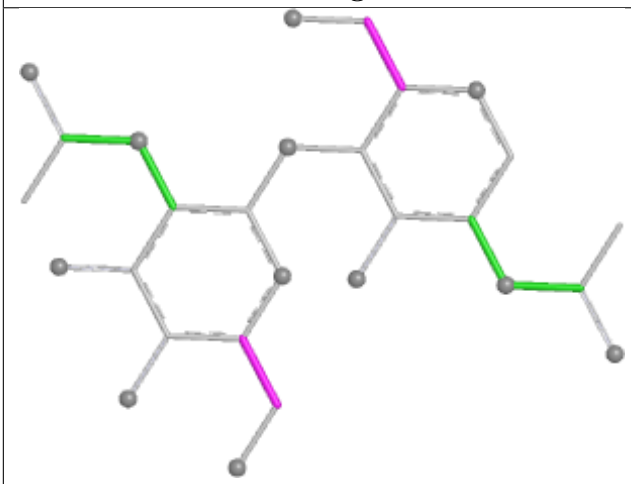
Oligosaccharide Chain G



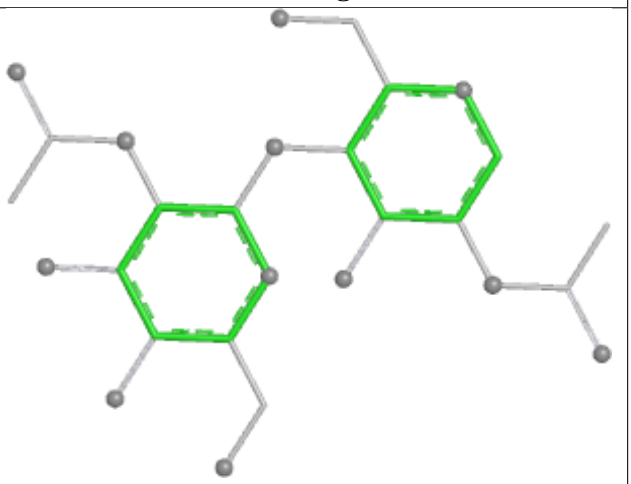
Bond lengths



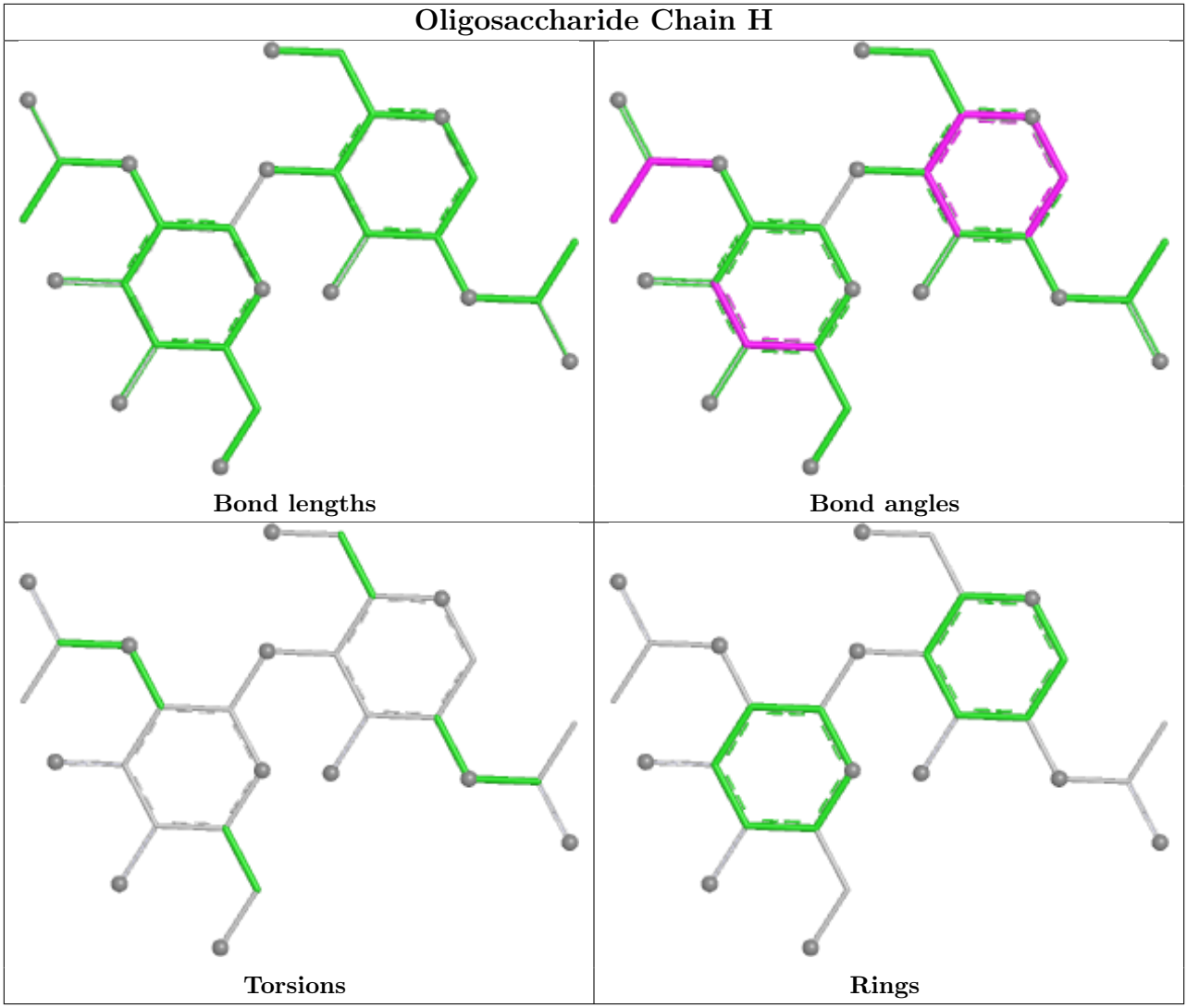
Bond angles



Torsions



Rings



5.6 Ligand geometry (i)

3 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	NAG	A	901	1	14,14,15	0.59	0	17,19,21	2.16	3 (17%)
4	NAG	A	801	1	14,14,15	0.49	0	17,19,21	1.50	1 (5%)
4	NAG	B	901	1	14,14,15	0.79	0	17,19,21	2.69	7 (41%)

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	901	1	-	0/6/23/26	0/1/1/1
4	NAG	A	801	1	-	2/6/23/26	0/1/1/1
4	NAG	B	901	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	901	NAG	C1-O5-C5	8.51	123.59	112.19
4	A	901	NAG	C1-O5-C5	6.94	121.48	112.19
4	A	801	NAG	C1-O5-C5	4.37	118.04	112.19
4	A	901	NAG	C3-C4-C5	-3.59	103.73	110.23
4	B	901	NAG	C3-C4-C5	-3.34	104.17	110.23
4	B	901	NAG	O3-C3-C4	2.86	117.12	110.38
4	B	901	NAG	O3-C3-C2	2.68	114.98	109.40
4	B	901	NAG	O4-C4-C3	2.56	116.41	110.38
4	B	901	NAG	C4-C3-C2	-2.54	107.29	111.02
4	B	901	NAG	C2-N2-C7	-2.21	119.94	122.90
4	A	901	NAG	O4-C4-C5	2.20	114.73	109.32

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	801	NAG	O5-C5-C6-O6
4	A	801	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

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	441/455 (96%)	-0.46	3 (0%)	84	82	18, 30, 46, 58	0
1	B	441/455 (96%)	-0.43	2 (0%)	87	85	18, 31, 45, 57	0
2	C	295/425 (69%)	0.40	23 (7%)	20	19	20, 57, 98, 152	0
2	D	284/425 (66%)	0.33	20 (7%)	24	22	22, 57, 89, 117	0
All	All	1461/1760 (83%)	-0.12	48 (3%)	49	47	18, 34, 80, 152	0

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	62	GLN	4.6
2	C	150	ASN	4.3
2	C	242	THR	3.9
1	A	453	GLY	3.8
2	D	397	SER	3.6
2	C	397	SER	3.4
2	C	147	VAL	3.3
2	C	151	ASP	3.3
2	C	396	ALA	3.3
2	D	456	LEU	3.2
2	D	401	THR	3.1
2	D	150	ASN	3.1
2	C	386	ALA	3.0
2	C	191	ASN	2.9
2	D	63	ALA	2.9
1	A	273	PHE	2.9
1	B	71	GLN	2.9
2	D	151	ASP	2.9
1	B	453	GLY	2.8
2	C	402	LEU	2.7
2	C	344	LYS	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	64	SER	2.7
2	C	149	ALA	2.6
2	D	213	ALA	2.6
2	D	399	VAL	2.5
2	C	219	VAL	2.5
2	D	215	VAL	2.5
2	D	233	LEU	2.4
2	C	234	THR	2.4
2	C	152	GLY	2.3
1	A	381	ASP	2.3
2	C	403	ILE	2.3
2	D	149	ALA	2.3
2	D	148	GLY	2.2
2	D	201	GLY	2.1
2	C	406	ASP	2.1
2	C	372	ALA	2.1
2	D	226	VAL	2.1
2	C	368	THR	2.1
2	C	400	SER	2.1
2	D	200	ALA	2.1
2	C	190	LYS	2.1
2	C	395	THR	2.0
2	C	64	SER	2.0
2	D	361	VAL	2.0
2	C	193	THR	2.0
2	D	454	THR	2.0
2	D	65	ARG	2.0

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

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

6.3 Carbohydrates (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{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	F	2	14/15	0.71	0.13	40,50,58,58	0

Continued on next page...

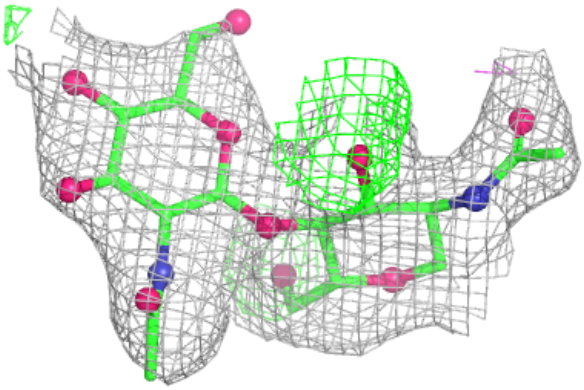
Continued from previous page...

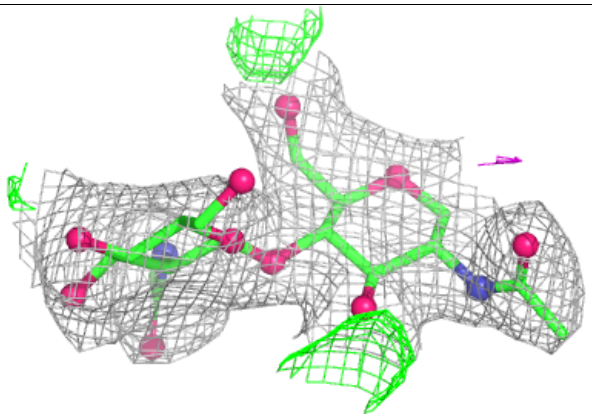
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	NAG	G	1	14/15	0.75	0.12	37,47,49,55	0
3	NAG	E	1	14/15	0.80	0.10	34,40,45,47	0
3	NAG	G	2	14/15	0.80	0.11	60,63,68,69	0
3	NAG	H	2	14/15	0.82	0.10	44,47,53,54	0
3	NAG	E	2	14/15	0.86	0.10	51,54,59,61	0
3	NAG	H	1	14/15	0.94	0.06	20,26,31,39	0
3	NAG	F	1	14/15	0.94	0.07	23,27,33,39	0

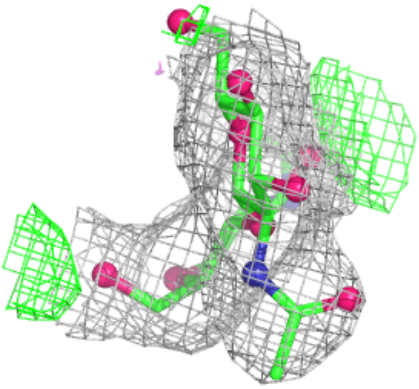
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

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)

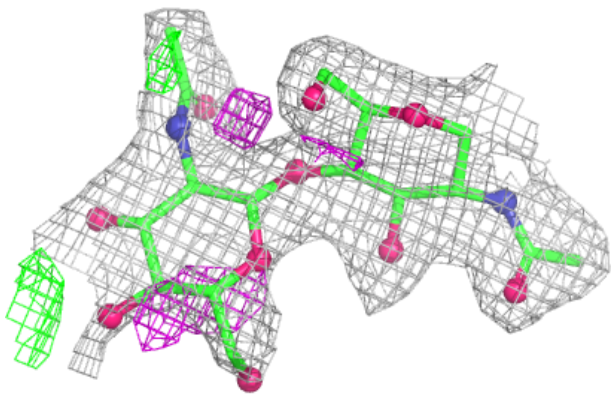


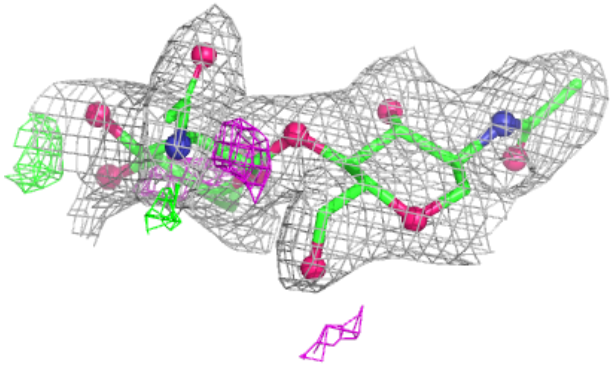


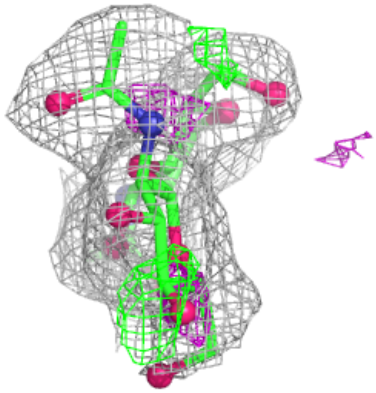


Electron density around Chain F:

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

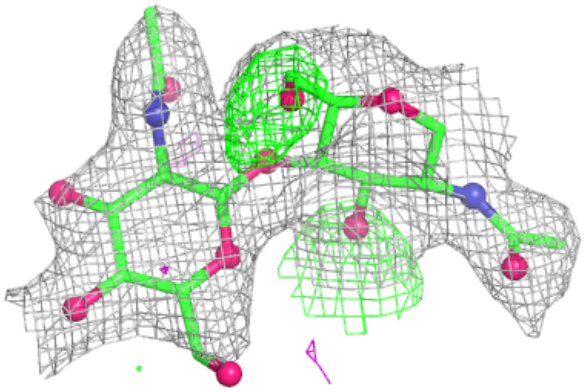


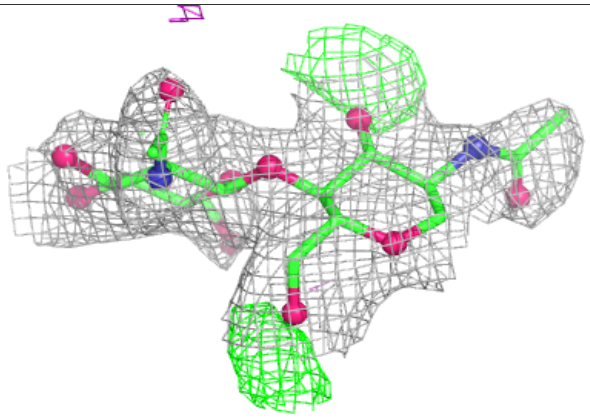


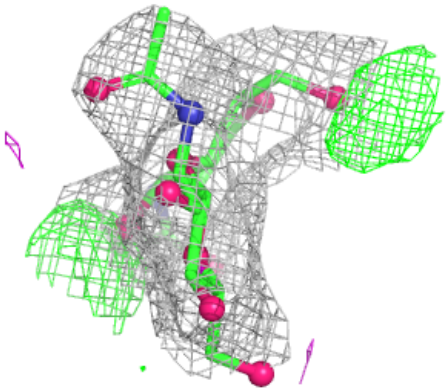


Electron density around Chain G:

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

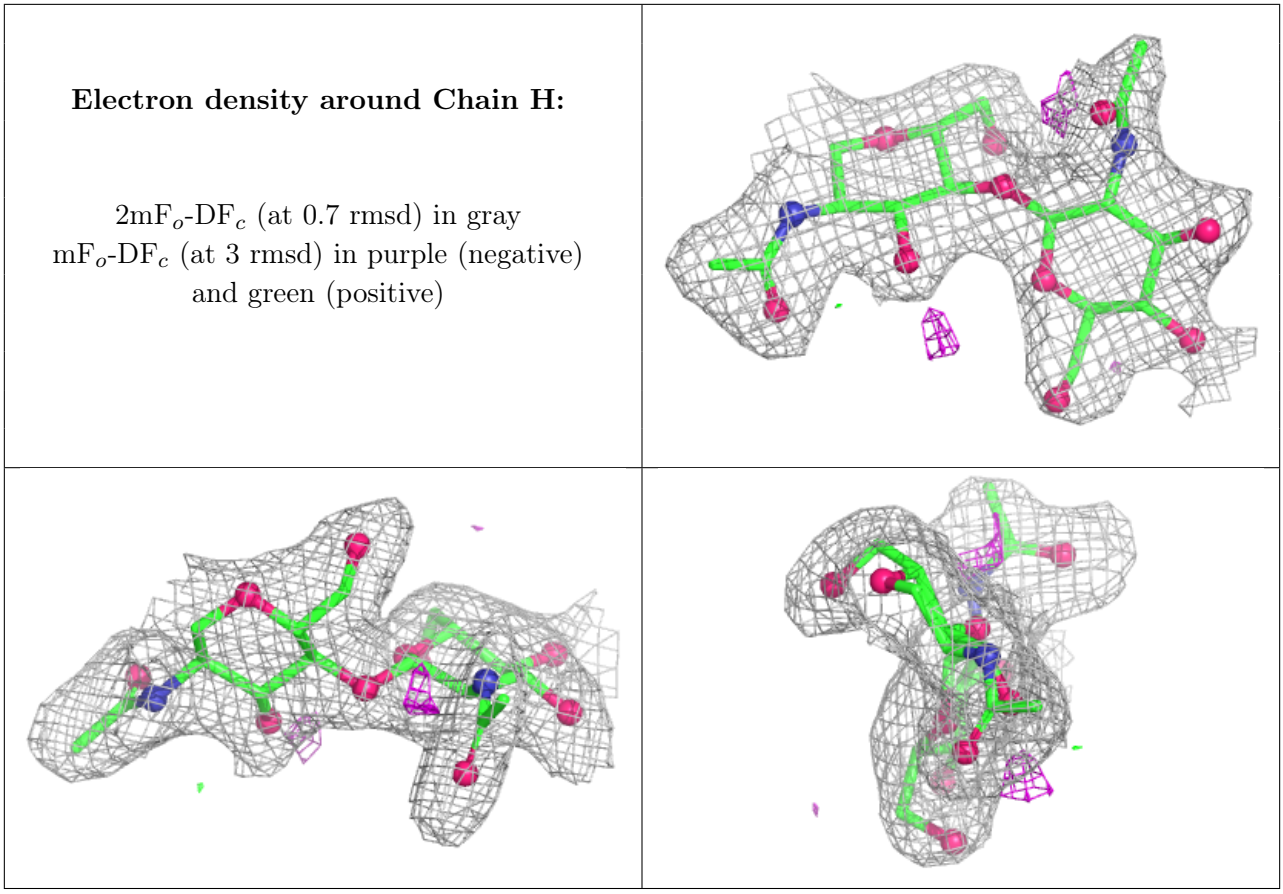






Electron density around Chain H:

$2mF_o$-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	801	14/15	0.76	0.14	51,61,64,64	0
4	NAG	A	901	14/15	0.78	0.11	37,47,51,52	0
4	NAG	B	901	14/15	0.78	0.11	35,43,49,52	0

6.5 Other polymers

There are no such residues in this entry.