Full wwPDB X-ray Structure Validation Report (i)

Dec 17, 2023 – 12:19 am GMT

PDB ID	:	3ZNL
Title	:	H5 Haemagglutinin in Complex with 6-O-Sulfo-Sialyl-Lewis X (Sulfated Lewis X)
Authors	:	Xiong, X.; Tuzikov, A.; Coombs, P.; Martin, S.R.; Walker, P.A.; Gamblin, S.J.; Bovin, N.; Skehel, J.J.
Deposited on	:	2013-02-15
Resolution	:	2.50 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:

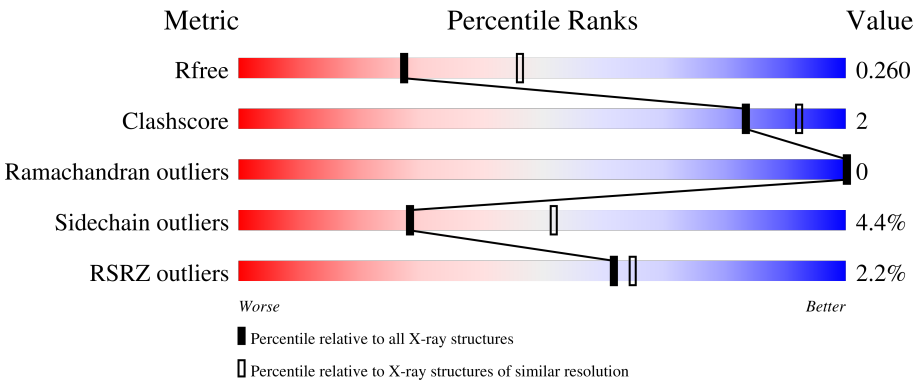
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	326	2% 88% 10% • •
1	C	326	3% 90% 9% •
1	E	326	3% 90% 9% •
2	B	166	% 87% 10% • •
2	D	166	% 90% 8% • •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	166	% 88% 10%
3	G	4	75% 25%
3	H	4	100%
3	I	4	75% 25%

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	B	1154	X	-	-	-
4	NAG	D	1154	X	-	-	-
4	NAG	F	1154	X	-	-	-

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 12211 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 HAEMAGGLUTININ.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	321	2547	1609	439	484	15	0	0	0
1	C	321	2547	1609	439	484	15	0	0	0
1	E	321	2547	1609	439	484	15	0	0	0

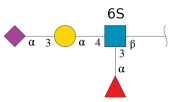
There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	325	THR	ARG	conflict	UNP Q6DQ34
C	325	THR	ARG	conflict	UNP Q6DQ34
E	325	THR	ARG	conflict	UNP Q6DQ34

- Molecule 2 is a protein called HAEMAGGLUTININ.

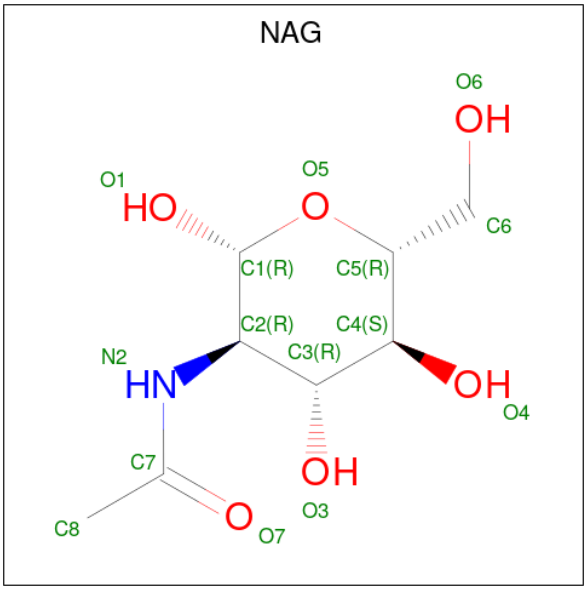
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	163	1318	822	227	261	8	0	0	0
2	D	163	1318	822	227	261	8	0	0	0
2	F	163	1318	822	227	261	8	0	0	0

- Molecule 3 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-3)-alpha-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-6-O-sulfo-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	G	4	60	31	2	26	1	0	0	0
			Total	C	N	O	S			
3	H	4	60	31	2	26	1	0	0	0
			Total	C	N	O	S			
3	I	4	60	31	2	26	1	0	0	0

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



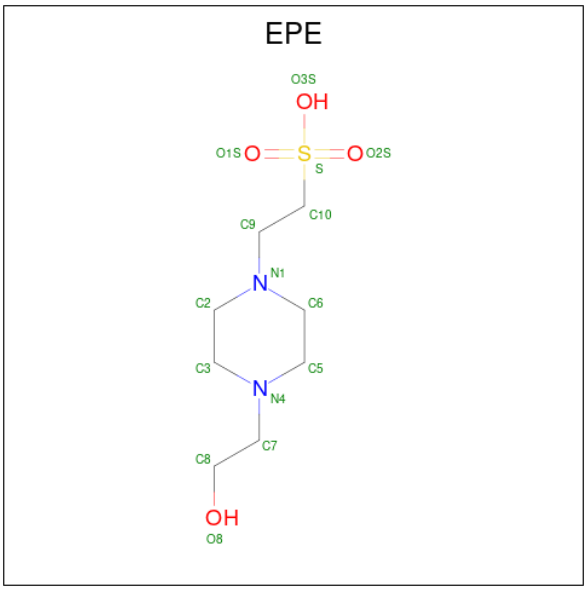
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	A	1	14	8	1	5	0	0
			Total	C	N	O		
4	A	1	14	8	1	5	0	0
			Total	C	N	O		
4	B	1	14	8	1	5	0	0
			Total	C	N	O		
4	C	1	14	8	1	5	0	0
			Total	C	N	O		
4	C	1	14	8	1	5	0	0
			Total	C	N	O		
4	D	1	14	8	1	5	0	0
			Total	C	N	O		
4	E	1	14	8	1	5	0	0
			Total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

Continued on next page...

Continued from previous page...

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

- Molecule 5 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
5	B	1	15	8	2	4	1	0	0
5	D	1	15	8	2	4	1	0	0
5	F	1	15	8	2	4	1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	A	60	60	60	0	0
6	B	37	37	37	0	0
6	C	52	52	52	0	0
6	D	38	38	38	0	0
6	E	49	49	49	0	0

Continued on next page...

Continued from previous page...

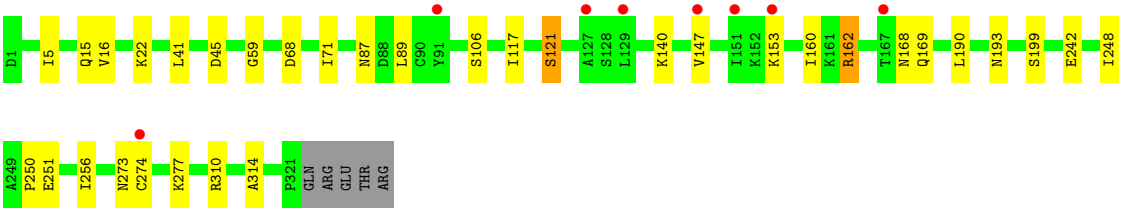
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	F	29	29	29	0	0

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.

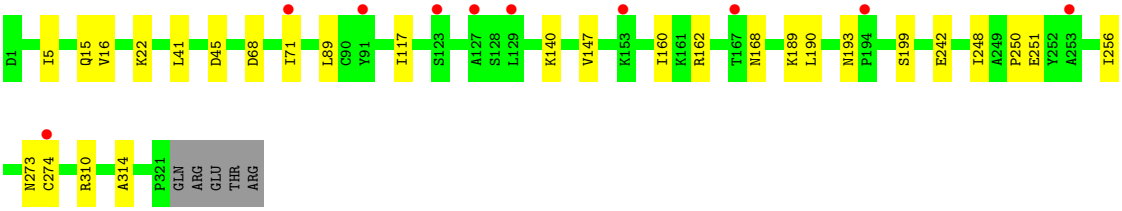
- Molecule 1: HAEMAGGLUTININ

Chain A:



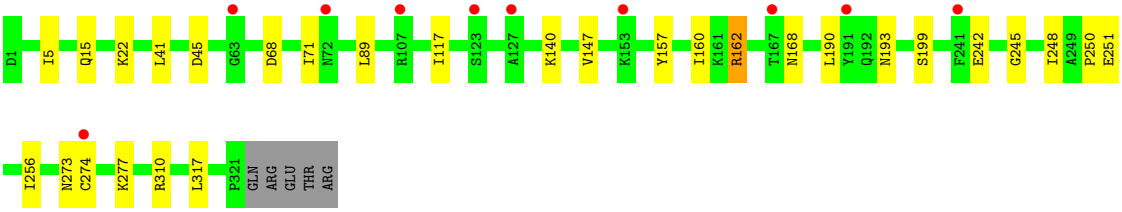
- Molecule 1: HAEMAGGLUTININ

Chain C:



- Molecule 1: HAEMAGGLUTININ

Chain E:

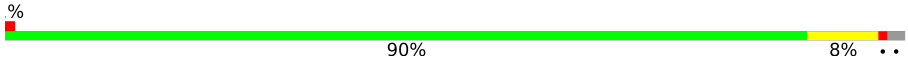


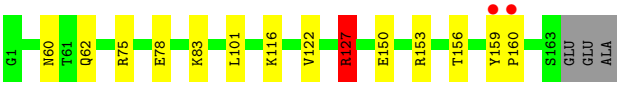
- Molecule 2: HAEMAGGLUTININ

Chain B:

• Molecule 2: HAEMAGGLUTININ

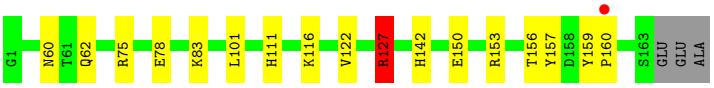
Chain D:





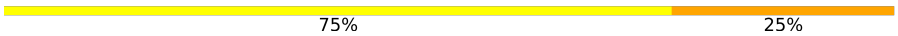
• Molecule 2: HAEMAGGLUTININ

Chain F:



• Molecule 3: N-acetyl-alpha-neuraminic acid-(2-3)-alpha-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-6-O-sulfo-beta-D-glucopyranose

Chain G:



• Molecule 3: N-acetyl-alpha-neuraminic acid-(2-3)-alpha-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-6-O-sulfo-beta-D-glucopyranose

Chain H:

• Molecule 3: N-acetyl-alpha-neuraminic acid-(2-3)-alpha-D-galactopyranose-(1-4)-[alpha-L-fucopyranose-(1-3)]2-acetamido-2-deoxy-6-O-sulfo-beta-D-glucopyranose

Chain I:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	175.98Å 101.60Å 161.37Å 90.00° 111.32° 90.00°	Depositor
Resolution (Å)	35.70 – 2.50 35.68 – 2.50	Depositor EDS
% Data completeness (in resolution range)	70.4 (35.70-2.50) 70.5 (35.68-2.50)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.71 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.237 , 0.256 0.239 , 0.260	Depositor DCC
R_{free} test set	3227 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	30.6	Xtriage
Anisotropy	0.045	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 10.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.458 for 1/2*h-3/2*k,-1/2*h-1/2*k,-1/2*h +1/2*k-l 0.467 for 1/2*h+3/2*k,1/2*h-1/2*k,-1/2*h-1/2*k-l	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	12211	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.29% 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: GLA, FUC, NAG, SIA, NGS, EPE

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	0.30	0/2609	0.50	0/3545
1	C	0.29	0/2609	0.50	0/3545
1	E	0.30	0/2609	0.52	1/3545 (0.0%)
2	B	0.32	0/1345	0.51	1/1810 (0.1%)
2	D	0.31	0/1345	0.51	1/1810 (0.1%)
2	F	0.31	0/1345	0.51	1/1810 (0.1%)
All	All	0.30	0/11862	0.51	4/16065 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	162	ARG	NE-CZ-NH1	6.05	123.32	120.30
2	B	127	ARG	NE-CZ-NH1	5.22	122.91	120.30
2	F	127	ARG	NE-CZ-NH1	5.17	122.88	120.30
2	D	127	ARG	NE-CZ-NH1	5.13	122.86	120.30

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	2547	0	2490	14	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2547	0	2490	11	0
1	E	2547	0	2490	11	0
2	B	1318	0	1223	8	0
2	D	1318	0	1223	6	0
2	F	1318	0	1223	8	0
3	G	60	0	44	1	0
3	H	60	0	44	3	0
3	I	60	0	44	1	0
4	A	28	0	26	1	0
4	B	14	0	13	0	0
4	C	28	0	26	1	0
4	D	14	0	13	0	0
4	E	28	0	26	1	0
4	F	14	0	13	0	0
5	B	15	0	18	0	0
5	D	15	0	18	0	0
5	F	15	0	18	0	0
6	A	60	0	0	3	0
6	B	37	0	0	0	0
6	C	52	0	0	0	0
6	D	38	0	0	0	0
6	E	49	0	0	0	0
6	F	29	0	0	0	0
All	All	12211	0	11442	55	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 2.

All (55) 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:121:SER:OG	1:A:162:ARG:NH2	2.00	0.94
2:B:68:ARG:NH1	2:B:81:ASN:OD1	2.03	0.91
1:A:87:ASN:O	6:A:2027:HOH:O	2.10	0.70
1:A:22:LYS:HE2	4:A:1023:NAG:H82	1.74	0.69
1:A:59:GLY:HA3	6:A:2027:HOH:O	1.93	0.69
2:B:75:ARG:NH1	2:B:78:GLU:OE1	2.27	0.68
2:D:75:ARG:NH1	2:D:78:GLU:OE1	2.27	0.67
2:F:75:ARG:NH1	2:F:78:GLU:OE1	2.27	0.66
1:C:5:ILE:HD11	2:D:122:VAL:HG21	1.76	0.66
1:A:5:ILE:HD11	2:B:122:VAL:HG21	1.77	0.66

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:5:ILE:HD11	2:F:122:VAL:HG21	1.76	0.66
3:H:2:GLA:O5	3:H:4:FUC:H5	1.98	0.64
1:C:22:LYS:HE2	4:C:1023:NAG:H82	1.78	0.63
2:F:127:ARG:HH11	2:F:127:ARG:HG2	1.69	0.58
2:B:127:ARG:HH11	2:B:127:ARG:HG2	1.68	0.57
2:D:127:ARG:HG2	2:D:127:ARG:HH11	1.68	0.57
2:F:156:THR:O	2:F:156:THR:HG22	2.05	0.57
2:B:156:THR:O	2:B:156:THR:HG22	2.05	0.57
2:D:156:THR:HG22	2:D:156:THR:O	2.05	0.56
1:A:15:GLN:O	1:A:310:ARG:NH2	2.41	0.54
1:A:68:ASP:O	1:A:71:ILE:HG22	2.09	0.53
1:E:15:GLN:O	1:E:310:ARG:NH2	2.42	0.53
1:C:15:GLN:O	1:C:310:ARG:NH2	2.42	0.53
1:C:68:ASP:O	1:C:71:ILE:HG22	2.08	0.53
1:E:160:ILE:O	1:E:242:GLU:HA	2.10	0.52
1:E:68:ASP:O	1:E:71:ILE:HG22	2.08	0.52
1:C:160:ILE:O	1:C:242:GLU:HA	2.10	0.52
1:E:22:LYS:HE2	4:E:1023:NAG:H82	1.92	0.51
2:D:150:GLU:OE1	2:D:153:ARG:NH1	2.44	0.51
1:A:160:ILE:O	1:A:242:GLU:HA	2.10	0.51
2:B:150:GLU:OE1	2:B:153:ARG:NH1	2.44	0.51
2:F:150:GLU:OE1	2:F:153:ARG:NH1	2.44	0.51
1:A:117:ILE:HB	1:A:250:PRO:O	2.12	0.50
2:B:159:TYR:HB3	2:B:160:PRO:HD3	1.94	0.50
1:E:117:ILE:HB	1:E:250:PRO:O	2.12	0.50
2:D:159:TYR:HB3	2:D:160:PRO:HD3	1.93	0.50
1:C:117:ILE:HB	1:C:250:PRO:O	2.12	0.49
1:E:190:LEU:HD21	3:I:3:SIA:O10	2.12	0.49
2:F:159:TYR:HB3	2:F:160:PRO:HD3	1.94	0.49
1:A:190:LEU:HD21	3:G:3:SIA:O10	2.13	0.48
1:E:157:TYR:CE2	1:E:245:GLY:HA2	2.48	0.48
2:F:142:HIS:HE2	2:F:157:TYR:HH	1.62	0.47
1:C:190:LEU:HD21	3:H:3:SIA:O10	2.15	0.46
1:A:106:SER:HB3	6:A:2039:HOH:O	2.15	0.46
1:C:189:LYS:NZ	3:H:1:NGS:O7A	2.48	0.46
1:C:147:VAL:HB	1:C:248:ILE:HG22	1.98	0.45
1:A:147:VAL:HB	1:A:248:ILE:HG22	1.98	0.45
2:B:142:HIS:HE2	2:B:157:TYR:HH	1.63	0.45
1:E:147:VAL:HB	1:E:248:ILE:HG22	1.98	0.45
1:A:199:SER:OG	1:A:242:GLU:HB3	2.20	0.42
1:E:317:LEU:HB3	2:F:111:HIS:CG	2.55	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:199:SER:OG	1:E:242:GLU:HB3	2.20	0.41
1:C:199:SER:OG	1:C:242:GLU:HB3	2.20	0.41
1:A:16:VAL:HG21	1:A:314:ALA:HB2	2.03	0.41
1:C:16:VAL:HG21	1:C:314:ALA:HB2	2.03	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	319/326 (98%)	309 (97%)	10 (3%)	0	100	100
1	C	319/326 (98%)	309 (97%)	10 (3%)	0	100	100
1	E	319/326 (98%)	309 (97%)	10 (3%)	0	100	100
2	B	161/166 (97%)	157 (98%)	4 (2%)	0	100	100
2	D	161/166 (97%)	157 (98%)	4 (2%)	0	100	100
2	F	161/166 (97%)	157 (98%)	4 (2%)	0	100	100
All	All	1440/1476 (98%)	1398 (97%)	42 (3%)	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	289/294 (98%)	274 (95%)	15 (5%)	23	44
1	C	289/294 (98%)	278 (96%)	11 (4%)	33	58
1	E	289/294 (98%)	277 (96%)	12 (4%)	30	54
2	B	139/141 (99%)	133 (96%)	6 (4%)	29	53
2	D	139/141 (99%)	133 (96%)	6 (4%)	29	53
2	F	139/141 (99%)	133 (96%)	6 (4%)	29	53
All	All	1284/1305 (98%)	1228 (96%)	56 (4%)	28	52

All (56) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	41	LEU
1	A	45	ASP
1	A	89	LEU
1	A	121	SER
1	A	140	LYS
1	A	153	LYS
1	A	162	ARG
1	A	168	ASN
1	A	169	GLN
1	A	193	ASN
1	A	251	GLU
1	A	256	ILE
1	A	273	ASN
1	A	274	CYS
1	A	277	LYS
2	B	60	ASN
2	B	62	GLN
2	B	83	LYS
2	B	101	LEU
2	B	116	LYS
2	B	127	ARG
1	C	41	LEU
1	C	45	ASP
1	C	89	LEU
1	C	140	LYS
1	C	162	ARG
1	C	168	ASN
1	C	193	ASN
1	C	251	GLU
1	C	256	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	273	ASN
1	C	274	CYS
2	D	60	ASN
2	D	62	GLN
2	D	83	LYS
2	D	101	LEU
2	D	116	LYS
2	D	127	ARG
1	E	41	LEU
1	E	45	ASP
1	E	89	LEU
1	E	140	LYS
1	E	162	ARG
1	E	168	ASN
1	E	193	ASN
1	E	251	GLU
1	E	256	ILE
1	E	273	ASN
1	E	274	CYS
1	E	277	LYS
2	F	60	ASN
2	F	62	GLN
2	F	83	LYS
2	F	101	LEU
2	F	116	LYS
2	F	127	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	154	ASN
1	A	169	GLN
1	A	222	GLN
2	B	42	GLN
2	D	42	GLN
1	E	15	GLN
2	F	42	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

12 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	NGS	G	1	3	19,19,19	0.71	0	26,28,28	1.72	2 (7%)
3	GLA	G	2	3	11,11,12	0.49	0	15,15,17	1.66	3 (20%)
3	SIA	G	3	3	20,20,21	0.53	0	24,28,31	1.17	3 (12%)
3	FUC	G	4	3	10,10,11	0.47	0	14,14,16	1.50	2 (14%)
3	NGS	H	1	3	19,19,19	0.70	0	26,28,28	1.43	3 (11%)
3	GLA	H	2	3	11,11,12	0.46	0	15,15,17	2.03	3 (20%)
3	SIA	H	3	3	20,20,21	0.55	0	24,28,31	1.16	3 (12%)
3	FUC	H	4	3	10,10,11	0.41	0	14,14,16	1.04	1 (7%)
3	NGS	I	1	3	19,19,19	0.71	0	26,28,28	1.69	4 (15%)
3	GLA	I	2	3	11,11,12	0.48	0	15,15,17	1.59	2 (13%)
3	SIA	I	3	3	20,20,21	0.55	0	24,28,31	1.19	4 (16%)
3	FUC	I	4	3	10,10,11	0.48	0	14,14,16	1.55	1 (7%)

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	NGS	G	1	3	-	3/10/30/30	0/1/1/1
3	GLA	G	2	3	-	2/2/19/22	0/1/1/1
3	SIA	G	3	3	-	0/18/34/38	0/1/1/1
3	FUC	G	4	3	-	-	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NGS	H	1	3	-	4/10/30/30	0/1/1/1
3	GLA	H	2	3	-	2/2/19/22	0/1/1/1
3	SIA	H	3	3	-	0/18/34/38	0/1/1/1
3	FUC	H	4	3	-	-	0/1/1/1
3	NGS	I	1	3	-	3/10/30/30	0/1/1/1
3	GLA	I	2	3	-	2/2/19/22	0/1/1/1
3	SIA	I	3	3	-	0/18/34/38	0/1/1/1
3	FUC	I	4	3	-	-	0/1/1/1

There are no bond length outliers.

All (31) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	G	1	NGS	C1-C2-N2	-6.54	103.15	110.73
3	I	1	NGS	C1-C2-N2	-6.00	103.78	110.73
3	H	2	GLA	C1-O5-C5	5.43	119.55	112.19
3	I	4	FUC	C1-C2-C3	4.83	115.61	109.67
3	H	1	NGS	C1-C2-N2	-4.74	105.23	110.73
3	I	2	GLA	C1-C2-C3	4.62	115.35	109.67
3	H	2	GLA	C1-C2-C3	4.61	115.34	109.67
3	G	2	GLA	C1-C2-C3	4.37	115.03	109.67
3	G	4	FUC	C1-C2-C3	4.29	114.94	109.67
3	H	1	NGS	O4-C4-C3	3.47	118.38	110.35
3	I	1	NGS	O4-C4-C3	3.33	118.06	110.35
3	G	1	NGS	O4-C4-C3	3.30	117.98	110.35
3	H	3	SIA	C4-C5-N5	-3.08	104.28	110.38
3	G	3	SIA	C4-C5-N5	-3.00	104.44	110.38
3	I	3	SIA	C4-C5-N5	-2.97	104.49	110.38
3	H	4	FUC	C1-O5-C5	2.69	118.88	112.78
3	G	2	GLA	C2-C3-C4	2.54	115.29	110.89
3	G	2	GLA	O3-C3-C2	-2.53	105.15	109.99
3	H	2	GLA	O3-C3-C2	-2.37	105.45	109.99
3	I	3	SIA	O6-C2-C1	2.35	112.32	107.70
3	H	1	NGS	C4-C3-C2	-2.33	106.92	110.34
3	I	2	GLA	O3-C3-C2	-2.31	105.58	109.99
3	G	3	SIA	O6-C2-C1	2.30	112.21	107.70
3	G	4	FUC	C2-C3-C4	2.21	114.72	110.89
3	H	3	SIA	C6-O6-C2	2.18	116.00	111.34
3	I	3	SIA	C6-O6-C2	2.17	115.99	111.34
3	I	1	NGS	C4-C3-C2	-2.16	107.18	110.34
3	H	3	SIA	O6-C2-C1	2.16	111.93	107.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	G	3	SIA	C6-O6-C2	2.11	115.85	111.34
3	I	3	SIA	C3-C4-C5	-2.07	108.96	111.46
3	I	1	NGS	C3-C4-C5	-2.05	106.58	110.24

There are no chirality outliers.

All (16) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	H	2	GLA	O5-C5-C6-O6
3	I	2	GLA	O5-C5-C6-O6
3	H	2	GLA	C4-C5-C6-O6
3	I	2	GLA	C4-C5-C6-O6
3	G	2	GLA	O5-C5-C6-O6
3	G	1	NGS	C6-O6-S-O9
3	I	1	NGS	C6-O6-S-O7A
3	I	1	NGS	C6-O6-S-O9
3	I	1	NGS	C6-O6-S-O8
3	G	1	NGS	C6-O6-S-O7A
3	G	1	NGS	C6-O6-S-O8
3	H	1	NGS	C6-O6-S-O7A
3	H	1	NGS	C4-C5-C6-O6
3	H	1	NGS	O5-C5-C6-O6
3	H	1	NGS	C6-O6-S-O9
3	G	2	GLA	C4-C5-C6-O6

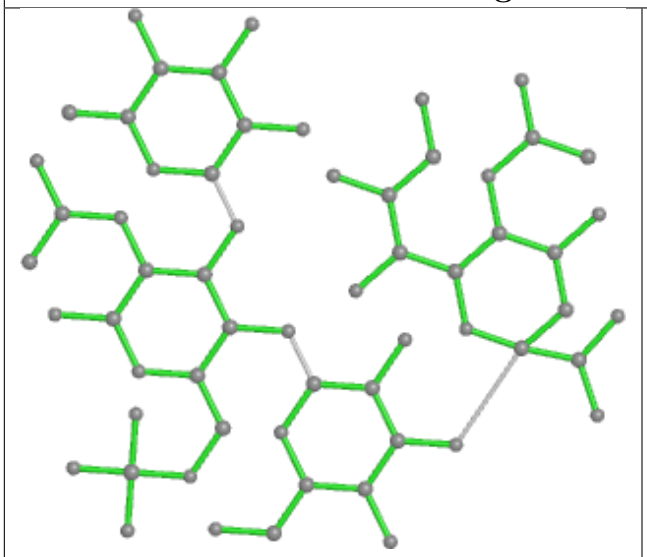
There are no ring outliers.

6 monomers are involved in 5 short contacts:

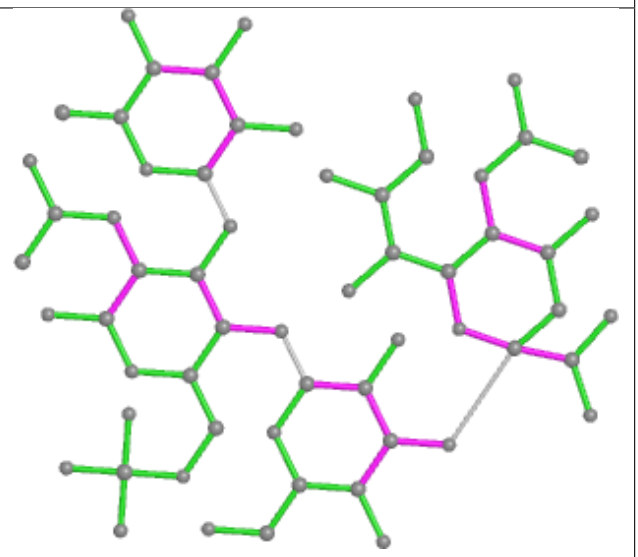
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	I	3	SIA	1	0
3	H	3	SIA	1	0
3	G	3	SIA	1	0
3	H	4	FUC	1	0
3	H	2	GLA	1	0
3	H	1	NGS	1	0

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

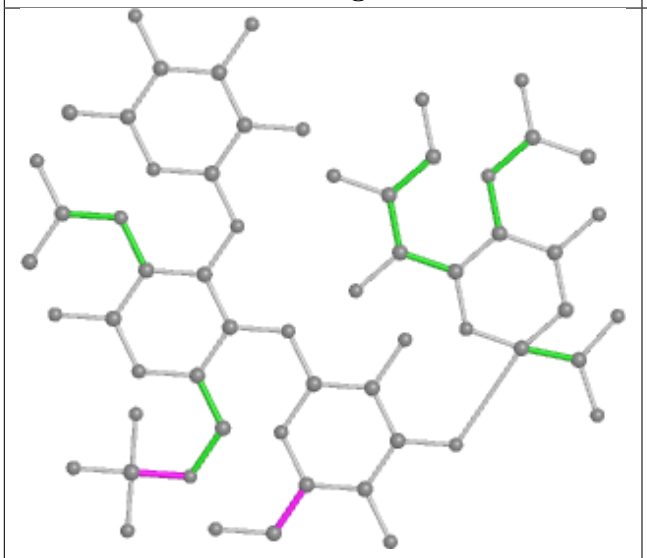
Oligosaccharide Chain G



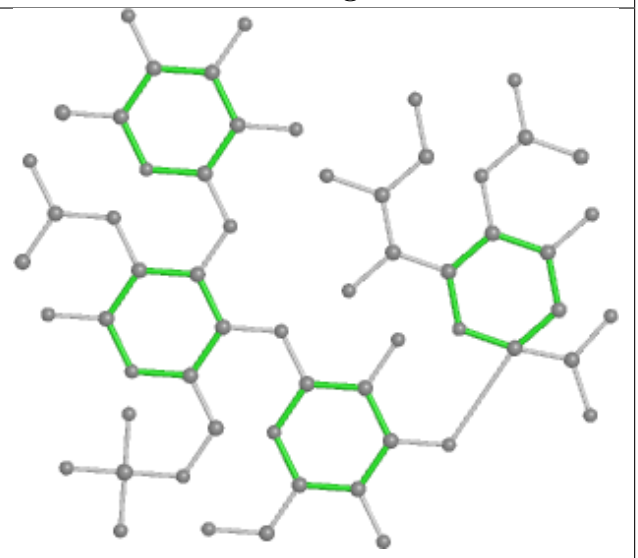
Bond lengths



Bond angles

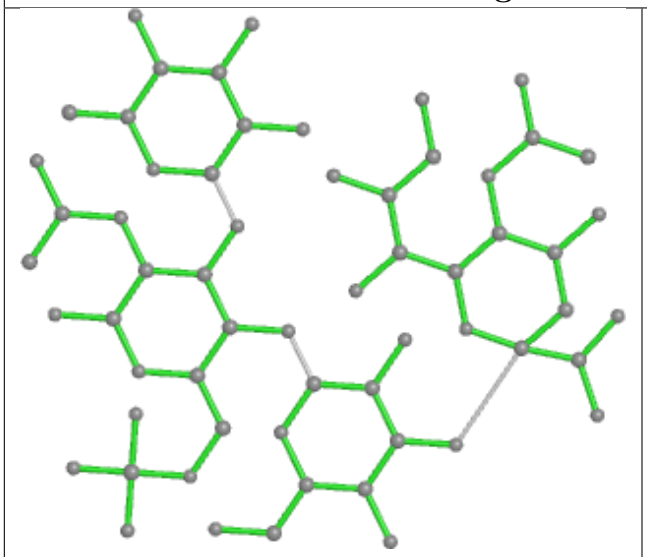


Torsions

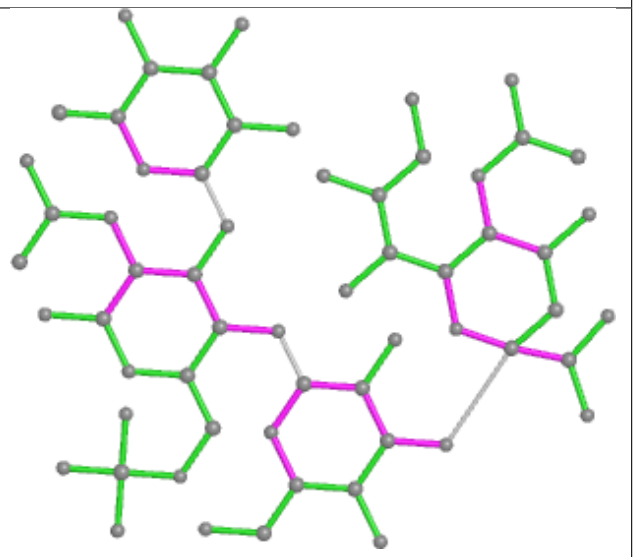


Rings

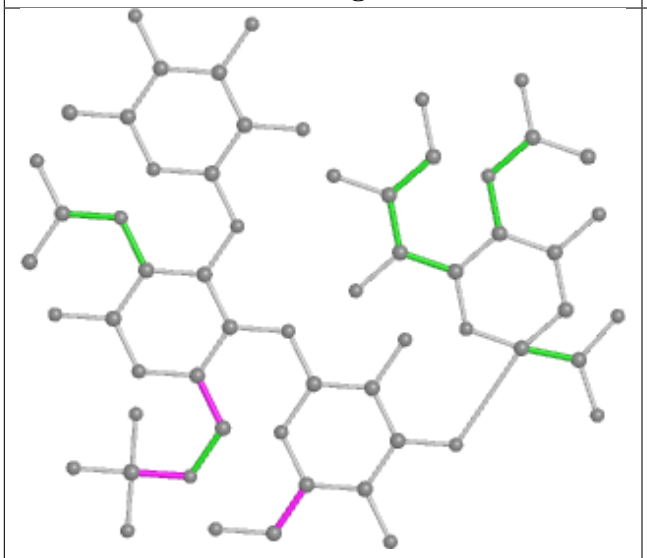
Oligosaccharide Chain H



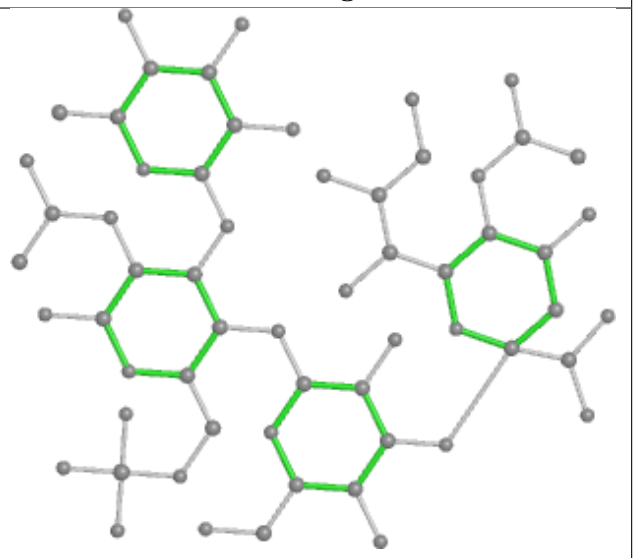
Bond lengths



Bond angles

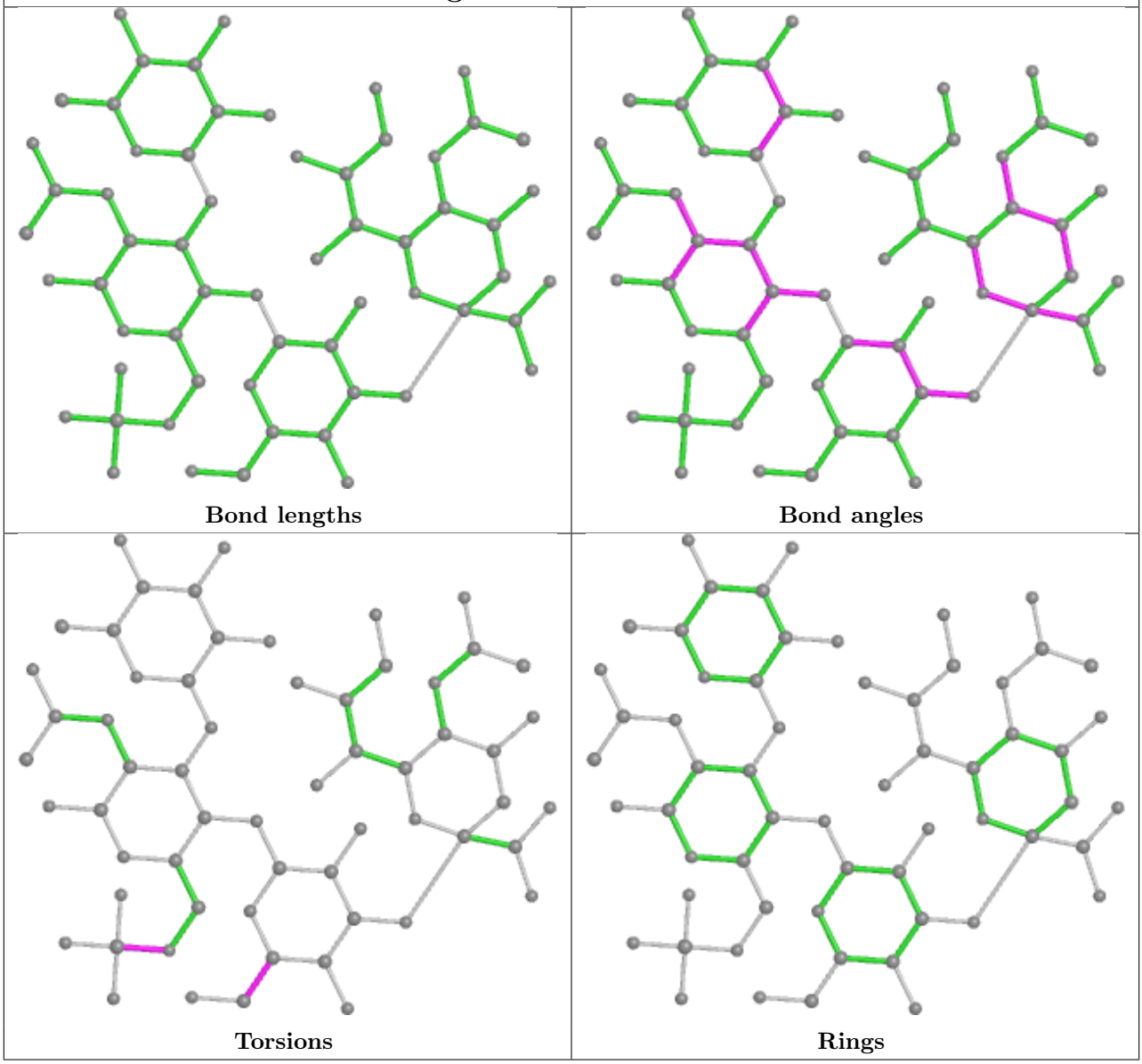


Torsions



Rings

Oligosaccharide Chain I



5.6 Ligand geometry (i)

12 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	1165	1	14,14,15	0.63	0	17,19,21	1.35	1 (5%)
4	NAG	E	1023	1	14,14,15	0.50	0	17,19,21	1.58	3 (17%)
4	NAG	A	1023	1	14,14,15	0.58	0	17,19,21	1.60	3 (17%)
5	EPE	F	1164	-	15,15,15	2.09	1 (6%)	18,20,20	2.71	9 (50%)
4	NAG	C	1165	1	14,14,15	0.62	0	17,19,21	1.60	6 (35%)
4	NAG	E	1165	1	14,14,15	0.64	0	17,19,21	1.46	3 (17%)
4	NAG	B	1154	2	14,14,15	0.45	0	17,19,21	1.60	1 (5%)
5	EPE	D	1164	-	15,15,15	1.94	1 (6%)	18,20,20	2.73	8 (44%)
5	EPE	B	1164	-	15,15,15	1.89	1 (6%)	18,20,20	2.77	9 (50%)
4	NAG	D	1154	2	14,14,15	0.46	0	17,19,21	1.47	1 (5%)
4	NAG	F	1154	2	14,14,15	0.47	0	17,19,21	1.54	1 (5%)
4	NAG	C	1023	1	14,14,15	0.55	0	17,19,21	1.40	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
4	NAG	E	1023	1	-	2/6/23/26	0/1/1/1
4	NAG	A	1023	1	-	2/6/23/26	0/1/1/1
5	EPE	F	1164	-	-	4/9/19/19	0/1/1/1
4	NAG	C	1165	1	-	1/6/23/26	0/1/1/1
4	NAG	E	1165	1	-	2/6/23/26	0/1/1/1
4	NAG	B	1154	2	1/1/5/7	0/6/23/26	0/1/1/1
5	EPE	D	1164	-	-	8/9/19/19	0/1/1/1
5	EPE	B	1164	-	-	6/9/19/19	0/1/1/1
4	NAG	D	1154	2	1/1/5/7	0/6/23/26	0/1/1/1
4	NAG	C	1023	1	-	0/6/23/26	0/1/1/1
4	NAG	F	1154	2	1/1/5/7	0/6/23/26	0/1/1/1
4	NAG	A	1165	1	-	1/6/23/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	F	1164	EPE	C10-S	-7.74	1.66	1.77
5	D	1164	EPE	C10-S	-7.14	1.67	1.77
5	B	1164	EPE	C10-S	-6.93	1.67	1.77

All (48) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	F	1154	NAG	C1-O5-C5	5.57	119.73	112.19
4	B	1154	NAG	C1-O5-C5	5.48	119.62	112.19
4	D	1154	NAG	C1-O5-C5	5.14	119.16	112.19
5	F	1164	EPE	C5-N4-C3	5.12	120.34	108.83
5	D	1164	EPE	C9-N1-C2	5.08	124.22	111.23
5	D	1164	EPE	C5-N4-C3	4.89	119.83	108.83
5	B	1164	EPE	C6-N1-C2	4.65	119.29	108.83
5	F	1164	EPE	C6-N1-C2	4.54	119.04	108.83
5	B	1164	EPE	C5-N4-C3	4.39	118.71	108.83
5	B	1164	EPE	O1S-S-C10	4.36	112.16	106.92
5	D	1164	EPE	C6-N1-C2	4.28	118.46	108.83
4	A	1023	NAG	C1-O5-C5	4.25	117.96	112.19
4	E	1023	NAG	C1-O5-C5	4.18	117.85	112.19
5	B	1164	EPE	C7-N4-C5	4.09	121.69	111.23
5	F	1164	EPE	C9-N1-C2	4.07	121.65	111.23
5	D	1164	EPE	O1S-S-C10	3.99	111.72	106.92
5	B	1164	EPE	C9-N1-C6	3.99	121.43	111.23
5	F	1164	EPE	C7-N4-C3	3.72	120.75	111.23
5	D	1164	EPE	C7-N4-C3	3.71	120.73	111.23
5	B	1164	EPE	C7-N4-C3	3.27	119.60	111.23
4	C	1023	NAG	C1-O5-C5	3.24	116.58	112.19
5	F	1164	EPE	O3S-S-C10	3.23	111.00	105.77
5	D	1164	EPE	C7-N4-C5	3.21	119.43	111.23
5	F	1164	EPE	C9-N1-C6	3.15	119.30	111.23
5	F	1164	EPE	C5-C6-N1	-3.15	104.19	110.64
5	B	1164	EPE	C9-N1-C2	3.14	119.28	111.23
5	F	1164	EPE	C7-N4-C5	3.00	118.90	111.23
4	C	1165	NAG	C4-C3-C2	-2.76	106.97	111.02
5	B	1164	EPE	O3S-S-C10	2.70	110.13	105.77
5	D	1164	EPE	O3S-S-C10	2.68	110.10	105.77
4	E	1023	NAG	O5-C5-C6	2.65	111.36	107.20
5	F	1164	EPE	O1S-S-C10	2.60	110.05	106.92
4	C	1023	NAG	O5-C5-C6	2.57	111.24	107.20
4	A	1023	NAG	C3-C4-C5	-2.54	105.70	110.24
4	E	1165	NAG	C2-N2-C7	2.39	126.31	122.90
4	A	1023	NAG	O5-C5-C6	2.39	110.95	107.20
5	D	1164	EPE	C9-N1-C6	2.37	117.29	111.23
4	E	1165	NAG	C1-C2-N2	2.36	114.52	110.49
4	C	1165	NAG	C2-N2-C7	2.29	126.17	122.90
4	E	1165	NAG	C4-C3-C2	-2.27	107.69	111.02
4	C	1165	NAG	O5-C5-C6	2.18	110.62	107.20
4	C	1023	NAG	C3-C4-C5	-2.17	106.37	110.24

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	C	1165	NAG	C1-C2-N2	2.15	114.15	110.49
4	E	1023	NAG	C3-C4-C5	-2.15	106.41	110.24
4	A	1165	NAG	O5-C5-C6	2.06	110.43	107.20
4	C	1165	NAG	O7-C7-C8	-2.03	118.28	122.06
5	B	1164	EPE	C2-C3-N4	-2.03	106.48	110.64
4	C	1165	NAG	O3-C3-C2	2.02	113.64	109.47

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	B	1154	NAG	C1
4	D	1154	NAG	C1
4	F	1154	NAG	C1

All (26) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	D	1164	EPE	C10-C9-N1-C2
5	D	1164	EPE	S-C10-C9-N1
5	D	1164	EPE	C9-C10-S-O2S
5	D	1164	EPE	C9-C10-S-O3S
5	F	1164	EPE	N4-C7-C8-O8
5	F	1164	EPE	S-C10-C9-N1
5	D	1164	EPE	N4-C7-C8-O8
4	E	1023	NAG	O5-C5-C6-O6
4	E	1023	NAG	C4-C5-C6-O6
5	B	1164	EPE	N4-C7-C8-O8
5	B	1164	EPE	C9-C10-S-O3S
4	A	1023	NAG	O5-C5-C6-O6
4	A	1023	NAG	C4-C5-C6-O6
5	D	1164	EPE	C8-C7-N4-C3
5	F	1164	EPE	C8-C7-N4-C3
5	B	1164	EPE	C8-C7-N4-C5
5	B	1164	EPE	C10-C9-N1-C6
5	D	1164	EPE	C10-C9-N1-C6
5	F	1164	EPE	C10-C9-N1-C6
5	B	1164	EPE	C9-C10-S-O1S
5	B	1164	EPE	C9-C10-S-O2S
5	D	1164	EPE	C9-C10-S-O1S
4	A	1165	NAG	C3-C2-N2-C7
4	C	1165	NAG	C1-C2-N2-C7
4	E	1165	NAG	C1-C2-N2-C7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
4	E	1165	NAG	C3-C2-N2-C7

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	1023	NAG	1	0
4	A	1023	NAG	1	0
4	C	1023	NAG	1	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 (i)

6.1 Protein, DNA and RNA chains (i)

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(Å^2)	Q<0.9
1	A	321/326 (98%)	0.25	8 (2%)	57	61	19, 54, 82, 102	0
1	C	321/326 (98%)	0.30	10 (3%)	49	52	16, 54, 80, 102	0
1	E	321/326 (98%)	0.29	10 (3%)	49	52	16, 55, 80, 100	0
2	B	163/166 (98%)	0.07	1 (0%)	89	90	8, 29, 59, 80	0
2	D	163/166 (98%)	0.21	2 (1%)	79	80	8, 29, 60, 80	0
2	F	163/166 (98%)	0.20	1 (0%)	89	90	9, 32, 58, 75	0
All	All	1452/1476 (98%)	0.24	32 (2%)	62	65	8, 46, 78, 102	0

All (32) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	167	THR	3.8
1	A	274	CYS	3.5
2	D	160	PRO	3.2
1	C	153	LYS	3.2
1	E	167	THR	3.0
1	C	129	LEU	3.0
1	A	167	THR	2.9
1	A	129	LEU	2.6
1	C	253	ALA	2.6
1	C	127	ALA	2.6
1	A	91	TYR	2.6
1	E	123	SER	2.5
1	A	127	ALA	2.5
1	E	153	LYS	2.4
1	C	123	SER	2.3
1	C	71	ILE	2.3
1	E	107	ARG	2.3
1	E	191	TYR	2.3
1	C	194	PRO	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	E	127	ALA	2.2
1	E	72	ASN	2.2
1	E	63	GLY	2.2
2	B	160	PRO	2.2
1	A	151	ILE	2.1
1	A	153	LYS	2.1
1	E	274	CYS	2.1
1	A	147	VAL	2.0
2	F	160	PRO	2.0
1	C	91	TYR	2.0
2	D	159	TYR	2.0
1	E	241	PHE	2.0
1	C	274	CYS	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)

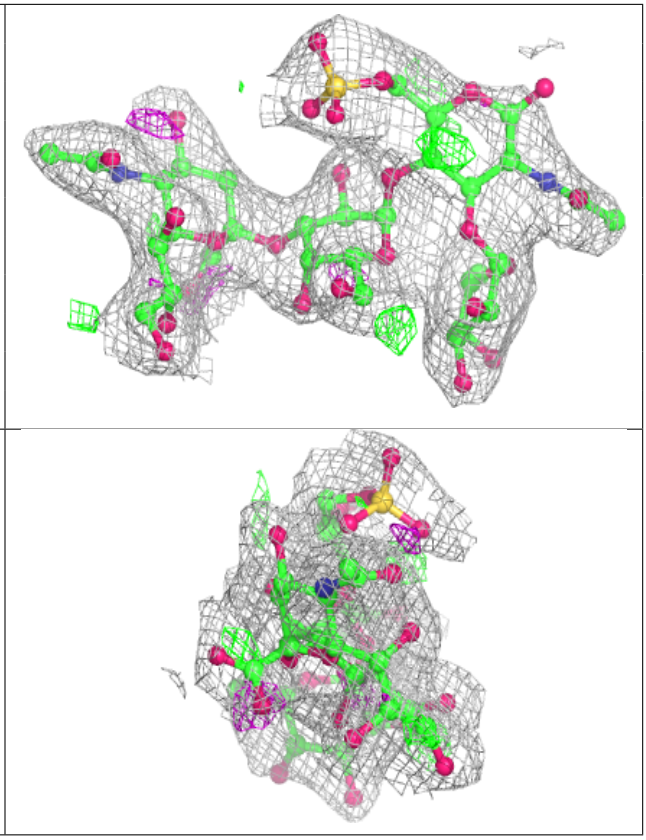
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists 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	GLA	H	2	11/12	0.81	0.14	50,54,57,59	0
3	NGS	H	1	19/19	0.85	0.23	65,71,79,79	0
3	NGS	G	1	19/19	0.86	0.19	66,72,81,82	0
3	NGS	I	1	19/19	0.86	0.22	62,70,75,76	0
3	GLA	I	2	11/12	0.87	0.12	47,51,55,56	0
3	GLA	G	2	11/12	0.89	0.11	50,54,58,60	0
3	SIA	I	3	20/21	0.89	0.19	39,40,47,47	0
3	FUC	H	4	10/11	0.90	0.13	63,66,67,68	0
3	SIA	G	3	20/21	0.90	0.17	38,41,48,49	0
3	FUC	G	4	10/11	0.90	0.23	66,70,71,72	0
3	SIA	H	3	20/21	0.90	0.19	40,42,48,49	0
3	FUC	I	4	10/11	0.94	0.15	64,68,70,71	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.

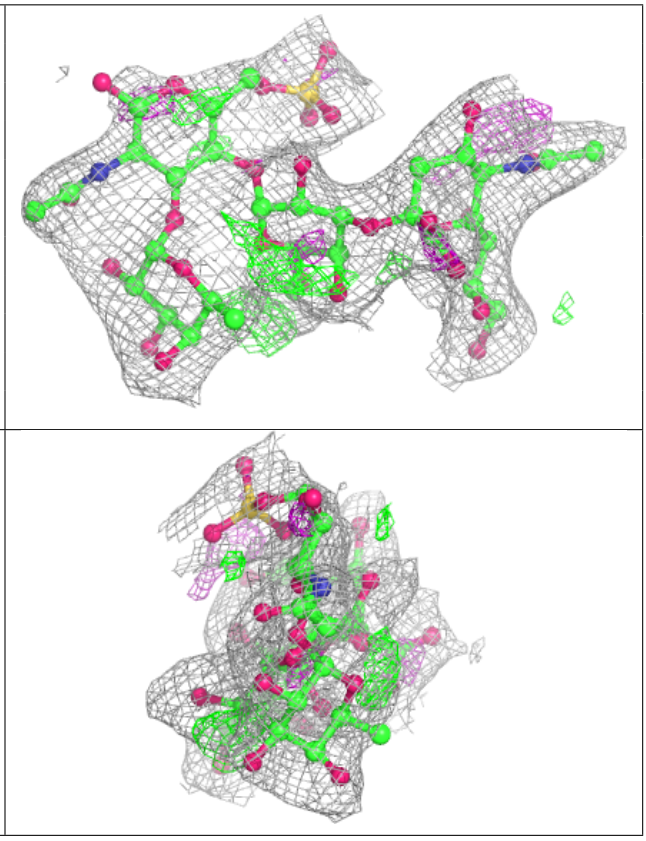
Electron density around Chain G:

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



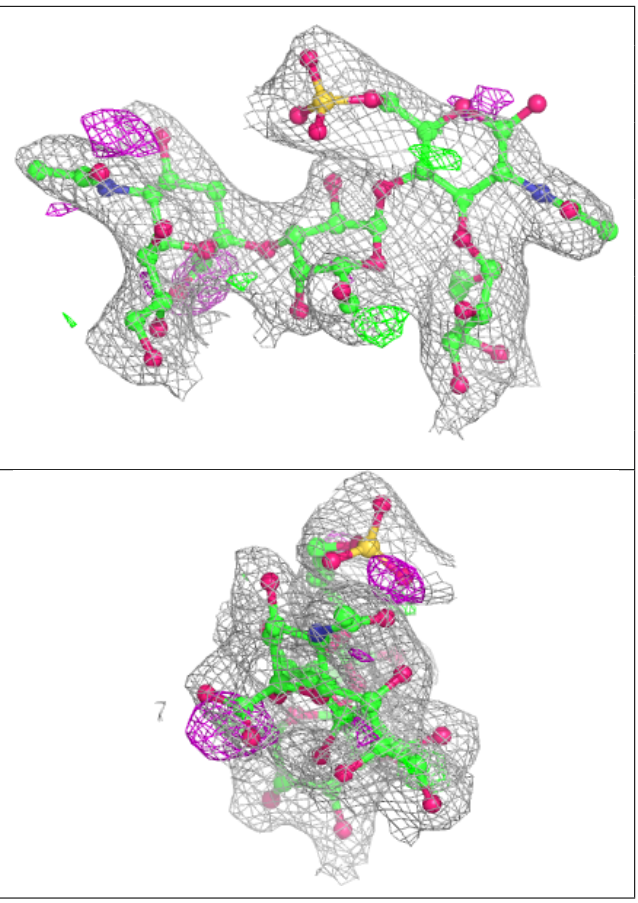
Electron density around Chain H:

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



Electron density around Chain I:

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



6.4 Ligands (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, 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
4	NAG	E	1165	14/15	0.82	0.16	24,27,31,32	0
5	EPE	D	1164	15/15	0.84	0.30	42,43,53,54	0
4	NAG	D	1154	14/15	0.85	0.23	33,38,40,40	0
5	EPE	B	1164	15/15	0.86	0.27	39,41,52,52	0
4	NAG	A	1165	14/15	0.86	0.17	26,29,33,34	0
4	NAG	C	1165	14/15	0.87	0.16	23,26,29,30	0
5	EPE	F	1164	15/15	0.87	0.22	38,39,47,47	0
4	NAG	F	1154	14/15	0.88	0.20	33,37,39,40	0
4	NAG	C	1023	14/15	0.88	0.14	24,29,30,30	0
4	NAG	A	1023	14/15	0.89	0.16	22,27,28,28	0
4	NAG	B	1154	14/15	0.90	0.20	34,38,41,42	0
4	NAG	E	1023	14/15	0.91	0.14	23,27,28,29	0

6.5 Other polymers (i)

There are no such residues in this entry.