Full wwPDB X-ray Structure Validation Report (i)

Aug 16, 2023 – 03:41 PM EDT

PDB ID	:	2AKR
Title	:	Structural basis of sulfatide presentation by mouse CD1d
Authors	:	Zajonc, D.M.; Halder, R.; Wu, D.; Maricic, I.; Roy, K.; Wong, C.-H.; Kumar, V.; Wilson, I.A.
Deposited on	:	2005-08-03
Resolution	:	1.90 Å(reported)

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

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

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

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

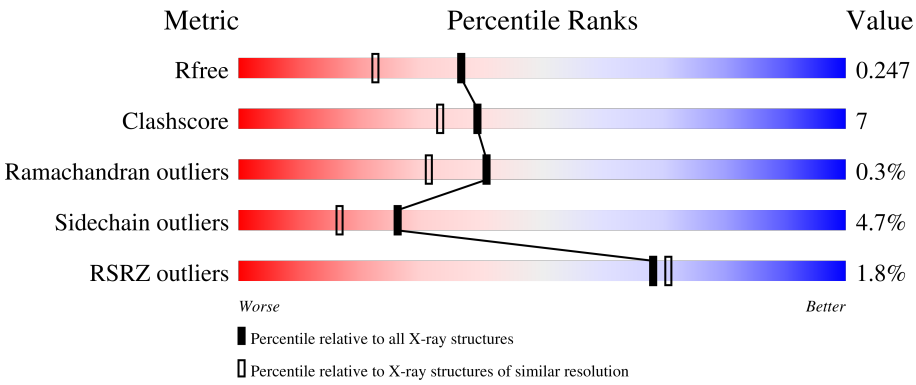
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber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

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	285	3%, 75%, 17%, • 6%
1	C	285	%, 80%, 13%, • 6%
2	B	99	83%, 16%, •
2	D	99	84%, 13%, • •
3	E	2	100%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 6918 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 T-cell surface glycoprotein CD1d1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	269	2165	1381	372	399	13	0	0	0
1	C	269	2165	1381	372	399	13	0	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	201	HIS	ASP	SEE REMARK 999	UNP P11609
A	280	HIS	-	expression tag	UNP P11609
A	281	HIS	-	expression tag	UNP P11609
A	282	HIS	-	expression tag	UNP P11609
A	283	HIS	-	expression tag	UNP P11609
A	284	HIS	-	expression tag	UNP P11609
A	285	HIS	-	expression tag	UNP P11609
C	201	HIS	ASP	SEE REMARK 999	UNP P11609
C	280	HIS	-	expression tag	UNP P11609
C	281	HIS	-	expression tag	UNP P11609
C	282	HIS	-	expression tag	UNP P11609
C	283	HIS	-	expression tag	UNP P11609
C	284	HIS	-	expression tag	UNP P11609
C	285	HIS	-	expression tag	UNP P11609

- Molecule 2 is a protein called Beta-2-microglobulin.

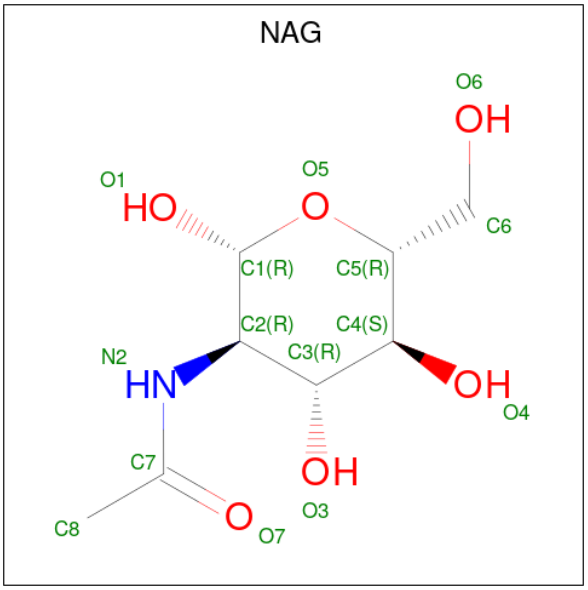
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	98	810	517	137	149	7	0	0	0
2	D	98	810	517	137	149	7	0	0	0

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

cetamido-2-deoxy-beta-D-glucopyranose.

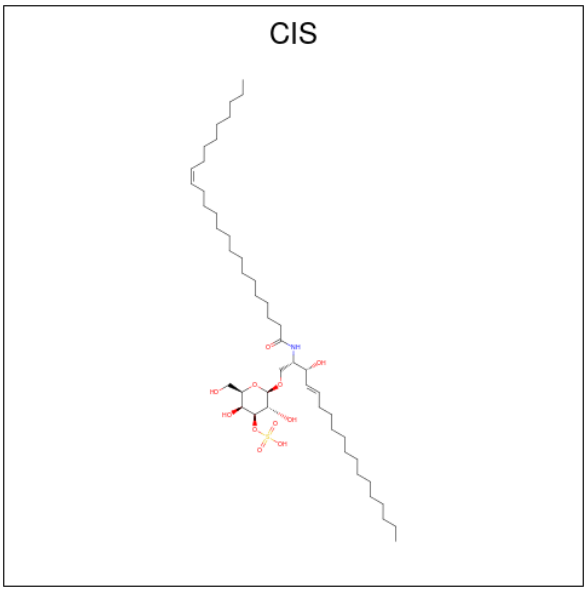
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	E	2	28	16	2	10	0	0	0

- 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
			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	A	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

- Molecule 5 is (15Z)-N-((1S,2R,3E)-2-HYDROXY-1-{[(3-O-SULFO-BETA-D-GALACTOPYRANOSYL)OXY]METHYL}HEPTADEC-3-ENYL)TETRACOS-15-ENAMIDE (three-letter code: CIS) (formula: $C_{48}H_{91}NO_{11}S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
5	A	1	61	48	1	11	1	0	0
			Total	C	N	O	S		
5	C	1	61	48	1	11	1	0	0

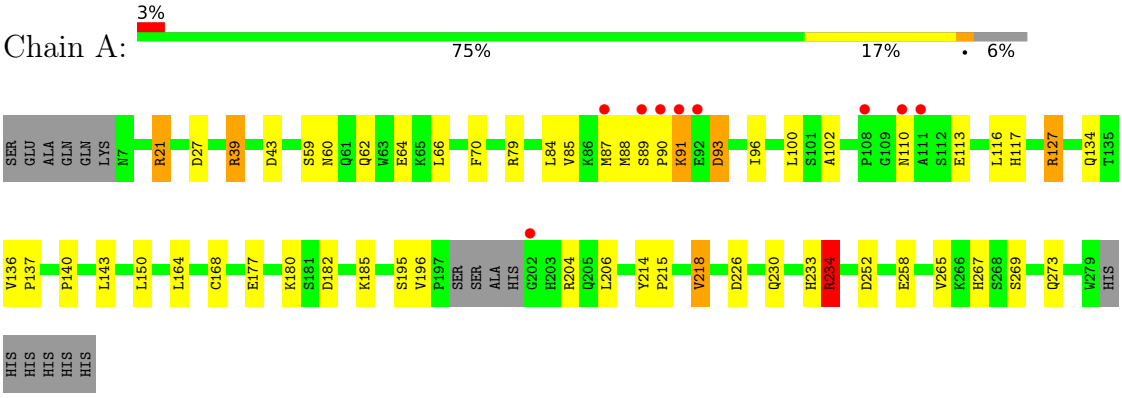
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	A	237	237	237	0	0
			Total	O		
6	B	132	132	132	0	0
			Total	O		
6	C	276	276	276	0	0
			Total	O		
6	D	103	103	103	0	0

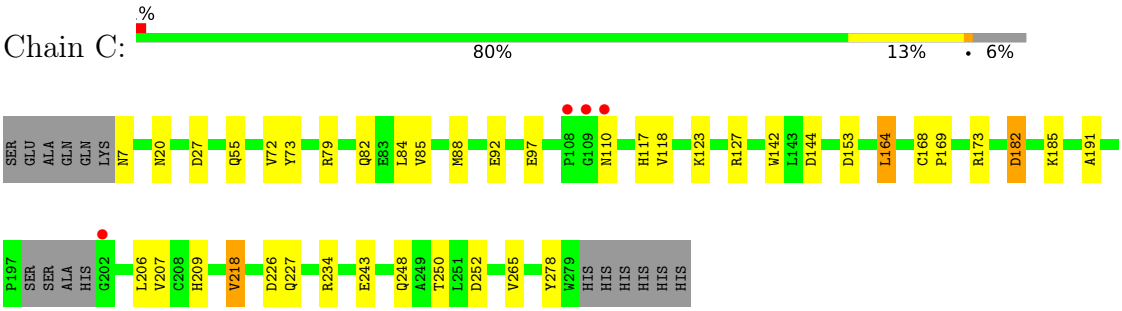
3 Residue-property plots (i)

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.

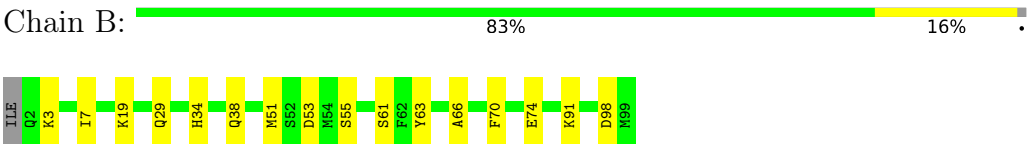
• Molecule 1: T-cell surface glycoprotein CD1d1



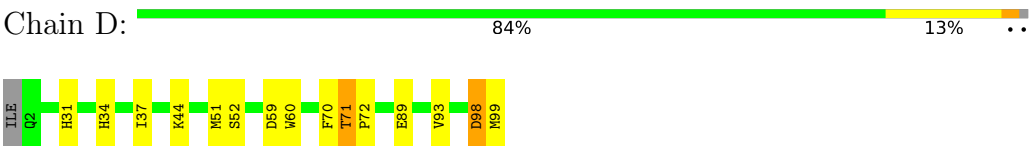
• Molecule 1: T-cell surface glycoprotein CD1d1



• Molecule 2: Beta-2-microglobulin



• Molecule 2: Beta-2-microglobulin



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

Chain E: 100%

NAG1
NAG2

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	58.86Å 74.84Å 101.56Å 90.00° 102.10° 90.00°	Depositor
Resolution (Å)	41.38 – 1.90 41.38 – 1.90	Depositor EDS
% Data completeness (in resolution range)	98.6 (41.38-1.90) 98.6 (41.38-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	1.90 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.2.0003	Depositor
R, R_{free}	0.186 , 0.249 0.187 , 0.247	Depositor DCC
R_{free} test set	1359 reflections (2.03%)	wwPDB-VP
Wilson B-factor (Å^2)	20.0	Xtriage
Anisotropy	0.082	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 55.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6918	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.80% 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, CIS

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	0.75	0/2228	0.91	6/3028 (0.2%)
1	C	0.77	0/2228	0.92	11/3028 (0.4%)
2	B	0.83	0/836	0.86	1/1133 (0.1%)
2	D	0.78	0/836	0.90	1/1133 (0.1%)
All	All	0.77	0/6128	0.91	19/8322 (0.2%)

There are no bond length outliers.

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	127	ARG	NE-CZ-NH1	7.83	124.22	120.30
2	D	98	ASP	CB-CG-OD2	7.29	124.86	118.30
1	A	93	ASP	CB-CG-OD2	7.12	124.71	118.30
1	A	252	ASP	CB-CG-OD2	6.97	124.57	118.30
1	C	27	ASP	CB-CG-OD2	6.42	124.08	118.30
1	C	218	VAL	CB-CA-C	-6.24	99.54	111.40
1	C	218	VAL	CG1-CB-CG2	6.20	120.82	110.90
1	C	252	ASP	CB-CG-OD2	5.82	123.54	118.30
1	C	127	ARG	NE-CZ-NH2	-5.79	117.40	120.30
1	A	218	VAL	CB-CA-C	-5.75	100.47	111.40
1	C	164	LEU	CB-CG-CD1	5.69	120.67	111.00
1	C	144	ASP	CB-CG-OD2	5.67	123.41	118.30
1	A	27	ASP	CB-CG-OD2	5.49	123.24	118.30
1	A	226	ASP	CB-CG-OD2	5.46	123.21	118.30
1	A	234	ARG	NE-CZ-NH2	5.34	122.97	120.30
1	C	182	ASP	CB-CG-OD2	5.31	123.08	118.30
1	C	153	ASP	CB-CG-OD2	5.17	122.95	118.30
1	C	88	MET	CB-CA-C	-5.09	100.21	110.40
2	B	98	ASP	CB-CG-OD2	5.06	122.86	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	2165	0	2071	42	0
1	C	2165	0	2072	28	0
2	B	810	0	783	12	0
2	D	810	0	783	12	0
3	E	28	0	25	0	0
4	A	42	0	39	0	0
4	C	28	0	26	4	0
5	A	61	0	90	3	0
5	C	61	0	90	5	0
6	A	237	0	0	10	2
6	B	132	0	0	4	1
6	C	276	0	0	8	2
6	D	103	0	0	6	1
All	All	6918	0	5979	89	3

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 7.

All (89) 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:C:20:ASN:HD21	4:C:500:NAG:C1	1.01	1.57
1:C:226:ASP:HB2	6:C:1109:HOH:O	1.58	1.04
1:A:21:ARG:HE	1:A:21:ARG:H	1.03	0.98
2:D:52:SER:HB3	6:D:151:HOH:O	1.73	0.88
2:D:71:THR:HG21	6:D:156:HOH:O	1.77	0.85
1:C:97:GLU:OE2	2:D:34:HIS:HE1	1.61	0.83
1:A:88:MET:O	1:A:91:LYS:HA	1.83	0.79
1:A:60:ASN:O	1:A:64:GLU:HG2	1.83	0.77
1:A:89:SER:HA	1:A:91:LYS:H	1.52	0.75
1:A:59:SER:H	1:A:62:GLN:NE2	1.85	0.74

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:248:GLN:OE1	2:D:99:MET:SD	2.44	0.74
1:A:21:ARG:H	1:A:21:ARG:NE	1.82	0.73
1:A:59:SER:H	1:A:62:GLN:HE21	1.36	0.71
1:A:21:ARG:HE	1:A:21:ARG:N	1.85	0.71
1:C:20:ASN:HD21	4:C:500:NAG:C2	1.97	0.71
1:C:85:VAL:HG12	1:C:92:GLU:HG2	1.70	0.71
2:D:31:HIS:ND1	6:D:128:HOH:O	1.96	0.70
1:A:88:MET:CE	6:A:1099:HOH:O	2.39	0.68
1:C:243:GLU:HG2	6:C:1247:HOH:O	1.92	0.68
1:A:88:MET:HE1	6:A:1099:HOH:O	1.96	0.66
1:A:85:VAL:HG21	1:A:96:ILE:HD13	1.79	0.65
2:D:71:THR:HG22	6:D:123:HOH:O	1.96	0.64
1:A:215:PRO:O	1:A:267:HIS:HE1	1.81	0.64
1:C:85:VAL:HG13	1:C:92:GLU:HA	1.81	0.63
1:A:233:HIS:HE1	6:A:1035:HOH:O	1.81	0.62
1:C:85:VAL:CG1	1:C:92:GLU:HG2	2.31	0.61
1:A:127:ARG:NH1	1:A:134:GLN:OE1	2.34	0.60
1:A:230:GLN:NE2	1:C:55:GLN:H	2.02	0.57
1:C:117:HIS:HE1	6:C:1052:HOH:O	1.86	0.57
1:C:72:VAL:HG12	6:C:1272:HOH:O	2.07	0.55
2:D:59:ASP:O	2:D:60:TRP:HB2	2.07	0.55
1:A:136:VAL:HG13	1:A:137:PRO:HD2	1.88	0.54
1:C:84:LEU:HD23	5:C:1002:CIS:H162	1.90	0.54
1:A:267:HIS:CD2	1:A:269:SER:H	2.25	0.54
1:C:7:ASN:N	6:C:1228:HOH:O	2.41	0.53
1:C:168:CYS:HB3	1:C:169:PRO:HD3	1.89	0.53
2:B:61:SER:HA	6:B:150:HOH:O	2.09	0.53
1:A:84:LEU:O	1:A:88:MET:HG2	2.09	0.52
1:C:142:TRP:HH2	5:C:1002:CIS:H173	1.75	0.52
1:A:267:HIS:HD2	1:A:269:SER:OG	1.93	0.52
1:C:20:ASN:CG	4:C:500:NAG:C1	2.72	0.52
1:A:70:PHE:HE1	5:A:1001:CIS:H261	1.74	0.51
2:B:7:ILE:HD12	2:B:91:LYS:HD3	1.92	0.51
1:A:113:GLU:HG2	6:A:1201:HOH:O	2.10	0.51
1:A:39:ARG:HD3	2:B:53:ASP:OD2	2.12	0.50
2:B:74:GLU:H	2:B:74:GLU:CD	2.14	0.50
2:D:37:ILE:HD12	2:D:51:MET:HE1	1.91	0.50
1:A:66:LEU:HD21	5:A:1001:CIS:H371	1.94	0.50
1:A:140:PRO:HD2	1:A:143:LEU:HD12	1.94	0.50
1:C:182:ASP:O	1:C:185:LYS:HG2	2.11	0.50
1:A:89:SER:HA	1:A:91:LYS:N	2.23	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:204:ARG:HD2	1:A:258:GLU:OE2	2.13	0.48
6:A:1167:HOH:O	2:B:34:HIS:HB2	2.12	0.48
1:C:118:VAL:HG11	5:C:1002:CIS:H112	1.94	0.48
1:A:100:LEU:HD11	1:A:116:LEU:HD21	1.95	0.48
2:D:34:HIS:HD2	6:D:170:HOH:O	1.96	0.48
1:C:84:LEU:CD2	5:C:1002:CIS:H162	2.44	0.47
1:A:136:VAL:CG1	1:A:137:PRO:HD2	2.45	0.47
1:A:88:MET:HE2	6:A:1099:HOH:O	2.07	0.47
1:A:182:ASP:O	1:A:185:LYS:HG2	2.14	0.47
2:B:29:GLN:HA	2:B:61:SER:HB2	1.97	0.47
1:C:97:GLU:OE2	2:D:34:HIS:CE1	2.53	0.47
1:A:102:ALA:HB2	1:A:116:LEU:HG	1.96	0.46
1:A:177:GLU:OE1	1:A:177:GLU:HA	2.14	0.46
2:B:3:LYS:NZ	6:B:165:HOH:O	2.48	0.46
1:A:117:HIS:HE1	6:A:1041:HOH:O	2.00	0.45
1:C:20:ASN:ND2	4:C:500:NAG:C2	2.66	0.45
1:C:82:GLN:NE2	6:C:1193:HOH:O	2.46	0.45
1:C:191:ALA:HA	1:C:209:HIS:O	2.17	0.45
1:A:79:ARG:NH1	6:A:1170:HOH:O	2.29	0.45
2:B:19:LYS:HD3	6:B:148:HOH:O	2.17	0.44
2:B:55:SER:HB3	2:B:63:TYR:CZ	2.52	0.44
1:C:173:ARG:NH1	6:C:1066:HOH:O	2.51	0.44
1:A:59:SER:OG	1:A:62:GLN:HG3	2.17	0.44
2:D:98:ASP:OD1	6:D:187:HOH:O	2.20	0.43
1:C:73:TYR:CE1	5:C:1002:CIS:H232	2.53	0.43
2:D:71:THR:HA	2:D:72:PRO:HD2	1.78	0.43
2:B:38:GLN:HG3	6:B:166:HOH:O	2.19	0.42
1:A:204:ARG:CD	1:A:258:GLU:OE2	2.68	0.42
1:A:214:TYR:CG	1:A:215:PRO:HA	2.53	0.42
1:A:267:HIS:HD2	1:A:269:SER:H	1.67	0.41
1:A:168:CYS:HB2	5:A:1001:CIS:H402	2.02	0.41
1:A:195:SER:OG	1:A:204:ARG:NH2	2.53	0.41
1:A:79:ARG:HD3	6:A:1170:HOH:O	2.19	0.41
1:A:234:ARG:HD3	6:A:1054:HOH:O	2.19	0.41
1:C:278:TYR:HE2	6:C:1061:HOH:O	2.02	0.41
2:B:51:MET:SD	2:B:66:ALA:HB2	2.61	0.41
1:C:207:VAL:HG22	1:C:250:THR:HG22	2.03	0.40
2:B:3:LYS:HB3	2:B:29:GLN:O	2.22	0.40

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

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:A:1149:HOH:O	6:C:1101:HOH:O[2_746]	2.12	0.08
6:A:1128:HOH:O	6:B:110:HOH:O[1_455]	2.18	0.02
6:C:1095:HOH:O	6:D:187:HOH:O[1_455]	2.18	0.02

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	265/285 (93%)	255 (96%)	8 (3%)	2 (1%)	19 9
1	C	265/285 (93%)	262 (99%)	3 (1%)	0	100 100
2	B	96/99 (97%)	93 (97%)	3 (3%)	0	100 100
2	D	96/99 (97%)	95 (99%)	1 (1%)	0	100 100
All	All	722/768 (94%)	705 (98%)	15 (2%)	2 (0%)	41 31

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	90	PRO
1	A	91	LYS

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	235/249 (94%)	219 (93%)	16 (7%)	16 7
1	C	235/249 (94%)	226 (96%)	9 (4%)	33 24

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	92/93 (99%)	91 (99%)	1 (1%)	73	73
2	D	92/93 (99%)	87 (95%)	5 (5%)	22	13
All	All	654/684 (96%)	623 (95%)	31 (5%)	26	16

All (31) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	21	ARG
1	A	39	ARG
1	A	43	ASP
1	A	87	MET
1	A	93	ASP
1	A	110	ASN
1	A	127	ARG
1	A	150	LEU
1	A	164	LEU
1	A	180	LYS
1	A	196	VAL
1	A	206	LEU
1	A	218	VAL
1	A	234	ARG
1	A	265	VAL
1	A	273	GLN
2	B	70	PHE
1	C	79	ARG
1	C	110	ASN
1	C	123	LYS
1	C	164	LEU
1	C	206	LEU
1	C	218	VAL
1	C	227	GLN
1	C	234	ARG
1	C	265	VAL
2	D	44	LYS
2	D	70	PHE
2	D	71	THR
2	D	89	GLU
2	D	93	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	62	GLN
1	A	117	HIS
1	A	229	GLN
1	A	230	GLN
1	A	233	HIS
1	A	267	HIS
2	B	6	GLN
2	B	29	GLN
1	C	20	ASN
1	C	117	HIS
1	C	205	GLN
1	C	229	GLN
1	C	233	HIS
1	C	248	GLN
2	D	8	GLN
2	D	34	HIS

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)

2 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	0.82	0	17,19,21	1.18	1 (5%)
3	NAG	E	2	3	14,14,15	0.62	0	17,19,21	1.41	1 (5%)

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	-	1/6/23/26	0/1/1/1
3	NAG	E	2	3	-	4/6/23/26	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	E	2	NAG	C1-O5-C5	4.16	117.83	112.19
3	E	1	NAG	O5-C1-C2	-3.53	105.71	111.29

There are no chirality outliers.

All (5) torsion outliers are listed below:

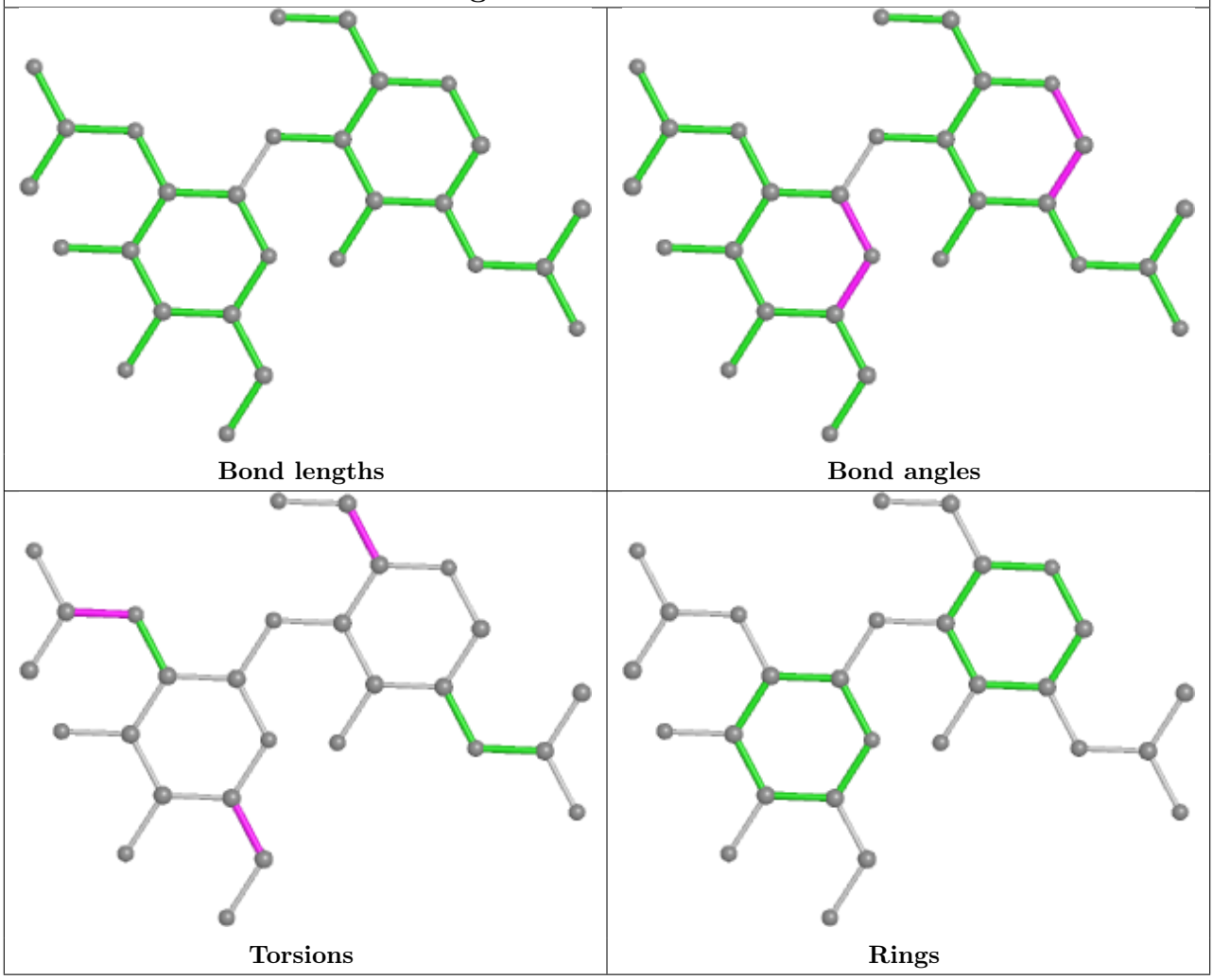
Mol	Chain	Res	Type	Atoms
3	E	2	NAG	O5-C5-C6-O6
3	E	2	NAG	C4-C5-C6-O6
3	E	2	NAG	C8-C7-N2-C2
3	E	1	NAG	C4-C5-C6-O6
3	E	2	NAG	O7-C7-N2-C2

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



5.6 Ligand geometry (i)

7 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	$\#\|Z\| > 2$
5	CIS	C	1002	-	60,61,61	1.08	4 (6%)	65,72,72	1.34	8 (12%)
4	NAG	A	501	1	14,14,15	0.66	0	17,19,21	1.51	2 (11%)
5	CIS	A	1001	-	60,61,61	0.99	2 (3%)	65,72,72	1.19	5 (7%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	NAG	C	501	1	14,14,15	0.79	0	17,19,21	1.35	2 (11%)
4	NAG	A	500	1	14,14,15	0.73	0	17,19,21	2.03	4 (23%)
4	NAG	C	500	1	14,14,15	0.83	1 (7%)	17,19,21	1.11	0
4	NAG	A	511	1	14,14,15	0.82	0	17,19,21	1.44	1 (5%)

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
5	CIS	C	1002	-	-	19/58/78/78	0/1/1/1
4	NAG	A	501	1	-	0/6/23/26	0/1/1/1
5	CIS	A	1001	-	-	20/58/78/78	0/1/1/1
4	NAG	C	501	1	-	0/6/23/26	0/1/1/1
4	NAG	A	500	1	-	4/6/23/26	0/1/1/1
4	NAG	C	500	1	-	0/6/23/26	0/1/1/1
4	NAG	A	511	1	-	0/6/23/26	0/1/1/1

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	1002	CIS	O4-S	-3.24	1.47	1.57
5	C	1002	CIS	O6-C47	2.94	1.51	1.44
5	A	1001	CIS	O4-S	-2.92	1.48	1.57
5	C	1002	CIS	O6-C43	2.74	1.48	1.41
5	C	1002	CIS	O-C43	2.30	1.44	1.40
4	C	500	NAG	O5-C1	-2.17	1.40	1.43
5	A	1001	CIS	O6-C43	2.11	1.47	1.41

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	C	1002	CIS	C45-O4-S	5.58	129.65	118.88
4	A	500	NAG	C4-C3-C2	5.03	118.39	111.02
5	A	1001	CIS	C-O-C43	-4.92	104.13	113.74
4	A	500	NAG	C1-O5-C5	4.65	118.49	112.19
5	C	1002	CIS	C46-C45-C44	-4.09	104.96	110.85
4	A	511	NAG	C1-O5-C5	4.08	117.72	112.19
5	A	1001	CIS	O-C43-C44	4.05	114.62	108.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	C	1002	CIS	O-C43-C44	3.43	113.66	108.30
4	A	500	NAG	C3-C4-C5	3.37	116.25	110.24
5	C	1002	CIS	C2-C3-C4	-3.26	117.52	124.79
4	A	501	NAG	O5-C1-C2	-3.14	106.33	111.29
4	C	501	NAG	C2-N2-C7	-2.95	118.70	122.90
4	A	501	NAG	C4-C3-C2	-2.62	107.18	111.02
5	A	1001	CIS	C2-C3-C4	-2.60	118.99	124.79
5	A	1001	CIS	O6-C47-C46	-2.51	105.14	109.69
5	C	1002	CIS	C5-C4-C3	-2.19	115.60	125.39
4	A	500	NAG	C6-C5-C4	-2.16	107.95	113.00
4	C	501	NAG	O7-C7-C8	-2.13	118.09	122.06
5	C	1002	CIS	C43-C44-C45	-2.13	106.32	110.07
5	C	1002	CIS	O3-C44-C43	2.10	115.15	110.05
5	C	1002	CIS	C45-C46-C47	-2.05	105.29	109.66
5	A	1001	CIS	O6-C43-C44	2.02	114.62	110.35

There are no chirality outliers.

All (43) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	500	NAG	C8-C7-N2-C2
4	A	500	NAG	O7-C7-N2-C2
5	C	1002	CIS	C46-C45-O4-S
5	C	1002	CIS	O6-C47-C48-O7
5	C	1002	CIS	C46-C47-C48-O7
4	A	500	NAG	C4-C5-C6-O6
4	A	500	NAG	O5-C5-C6-O6
5	C	1002	CIS	C12-C13-C14-C15
5	C	1002	CIS	C22-C23-C24-C25
5	C	1002	CIS	C13-C14-C15-C16
5	A	1001	CIS	C37-C38-C39-C40
5	A	1001	CIS	C9-C10-C11-C12
5	A	1001	CIS	C22-C23-C24-C25
5	C	1002	CIS	C29-C30-C31-C32
5	A	1001	CIS	C28-C29-C30-C31
5	A	1001	CIS	C5-C6-C7-C8
5	C	1002	CIS	C9-C10-C11-C12
5	A	1001	CIS	C34-C35-C36-C37
5	C	1002	CIS	C5-C6-C7-C8
5	A	1001	CIS	C36-C37-C38-C39
5	C	1002	CIS	C38-C39-C40-C41
5	A	1001	CIS	C26-C27-C28-C29

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
5	A	1001	CIS	C27-C28-C29-C30
5	C	1002	CIS	C14-C15-C16-C17
5	C	1002	CIS	C11-C12-C13-C14
5	A	1001	CIS	C25-C26-C27-C28
5	C	1002	CIS	C11-C10-C9-C8
5	A	1001	CIS	C38-C39-C40-C41
5	C	1002	CIS	C6-C7-C8-C9
5	A	1001	CIS	C10-C11-C12-C13
5	A	1001	CIS	C18-C19-C20-C21
5	C	1002	CIS	C21-C22-C23-C24
5	A	1001	CIS	C30-C31-C32-C33
5	A	1001	CIS	C11-C10-C9-C8
5	A	1001	CIS	C29-C30-C31-C32
5	A	1001	CIS	O6-C47-C48-O7
5	C	1002	CIS	C27-C28-C29-C30
5	A	1001	CIS	C23-C24-C25-C26
5	C	1002	CIS	C7-C8-C9-C10
5	C	1002	CIS	C30-C31-C32-C33
5	C	1002	CIS	C37-C38-C39-C40
5	A	1001	CIS	C32-C33-C34-C35
5	A	1001	CIS	C7-C8-C9-C10

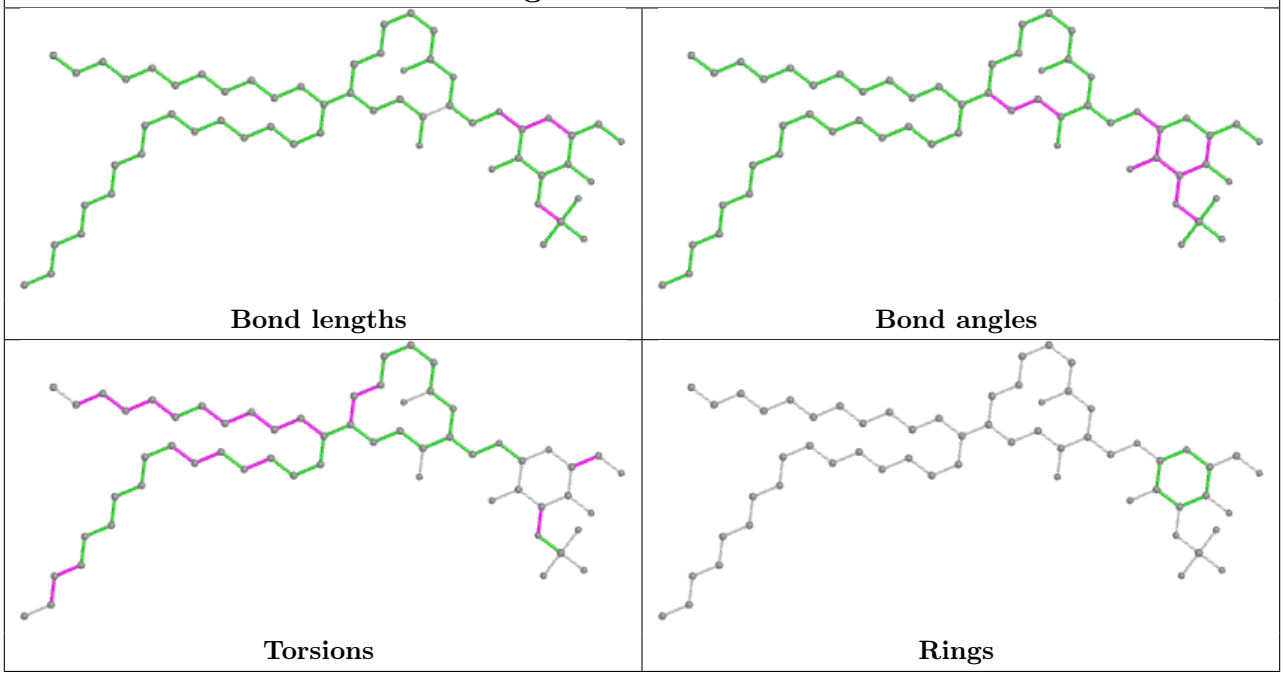
There are no ring outliers.

3 monomers are involved in 12 short contacts:

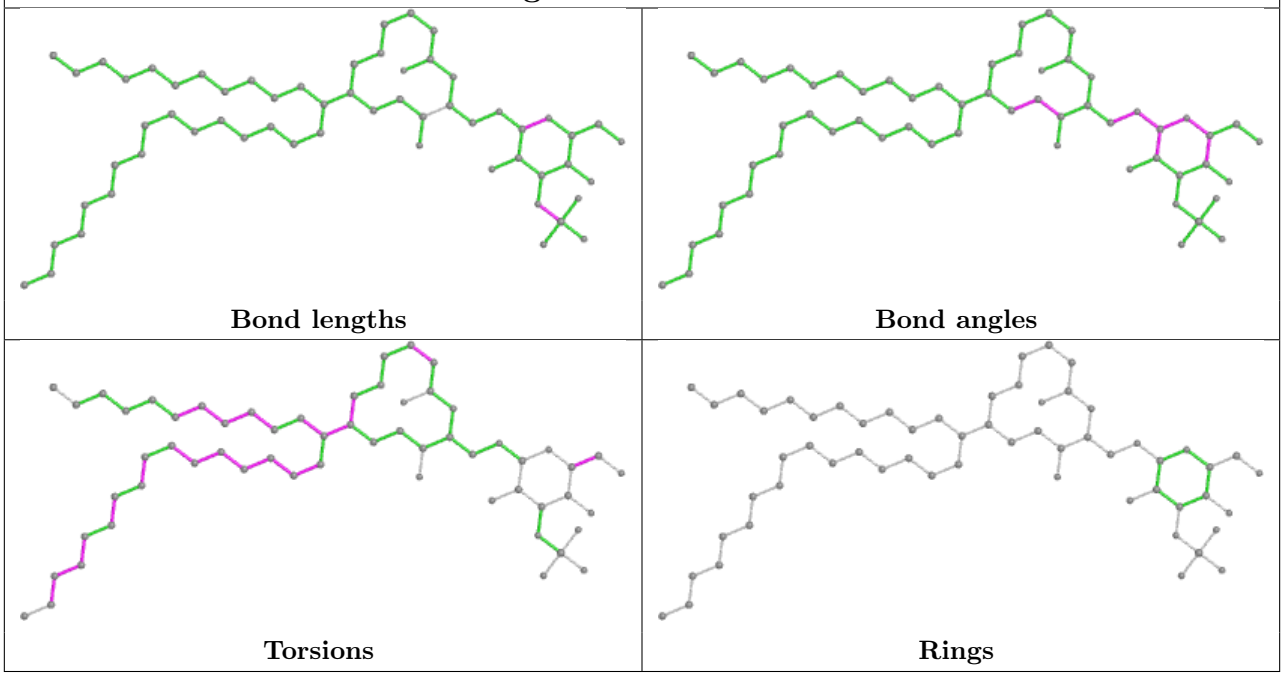
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	1002	CIS	5	0
5	A	1001	CIS	3	0
4	C	500	NAG	4	0

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

Ligand CIS C 1002



Ligand CIS A 1001



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	269/285 (94%)	-0.15	9 (3%) 46 49	13, 23, 49, 70	0
1	C	269/285 (94%)	-0.29	4 (1%) 73 76	12, 21, 40, 59	0
2	B	98/99 (98%)	-0.52	0 100 100	14, 21, 31, 38	0
2	D	98/99 (98%)	-0.34	0 100 100	14, 24, 38, 50	0
All	All	734/768 (95%)	-0.27	13 (1%) 68 71	12, 22, 42, 70	0

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	111	ALA	4.3
1	A	92	GLU	3.9
1	C	109	GLY	3.8
1	A	91	LYS	3.5
1	A	110	ASN	3.4
1	A	89	SER	3.1
1	A	87	MET	2.6
1	A	108	PRO	2.6
1	C	110	ASN	2.6
1	A	202	GLY	2.6
1	A	90	PRO	2.4
1	C	202	GLY	2.2
1	C	108	PRO	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)

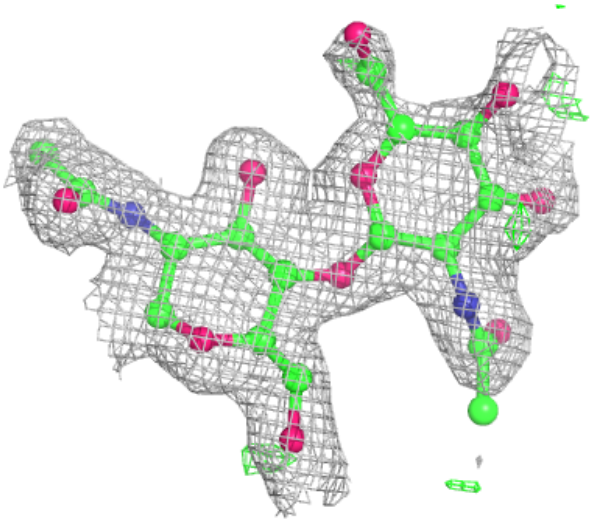
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

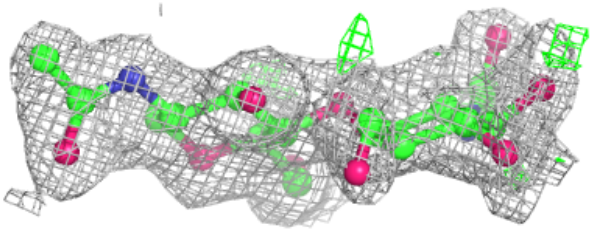
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	NAG	E	2	14/15	0.82	0.16	42,46,53,55	0
3	NAG	E	1	14/15	0.94	0.09	28,31,34,36	0

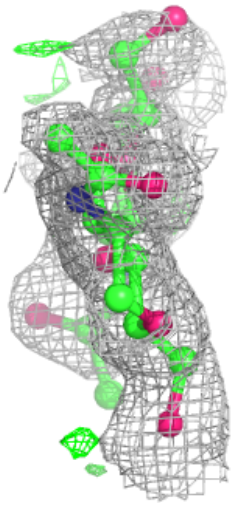
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

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)







6.4 Ligands (i)

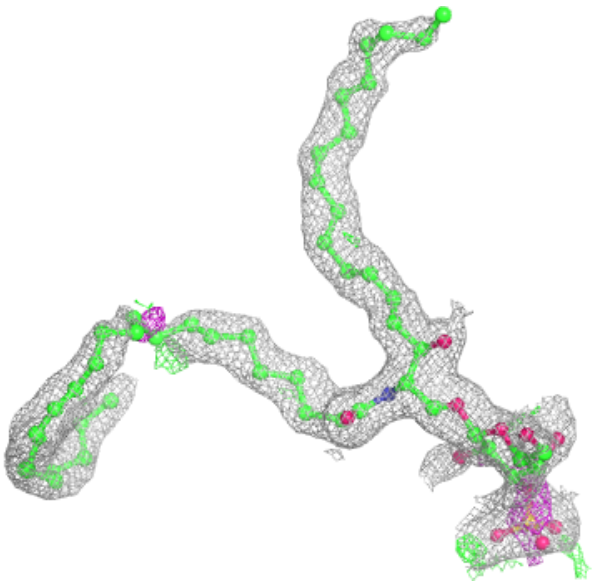
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

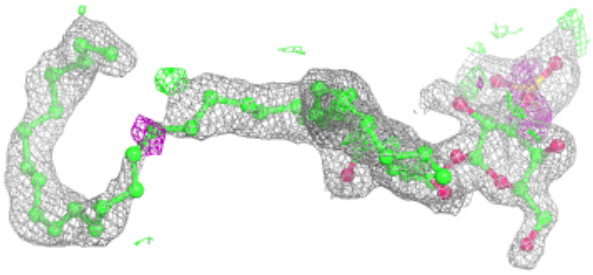
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	CIS	A	1001	61/61	0.79	0.18	27,37,64,70	0
5	CIS	C	1002	61/61	0.81	0.17	25,37,68,75	0
4	NAG	A	500	14/15	0.85	0.19	57,63,77,77	0
4	NAG	C	500	14/15	0.93	0.11	27,34,46,48	0
4	NAG	A	511	14/15	0.94	0.09	25,28,30,30	0
4	NAG	A	501	14/15	0.96	0.07	30,34,37,40	0
4	NAG	C	501	14/15	0.97	0.07	19,21,25,26	0

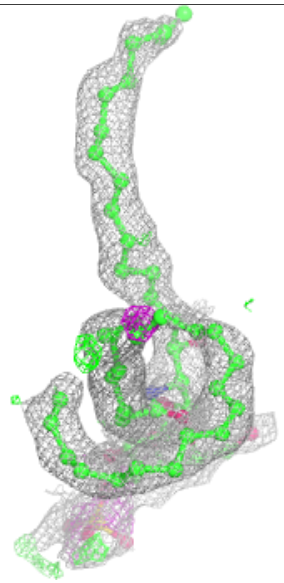
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around CIS A 1001:

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

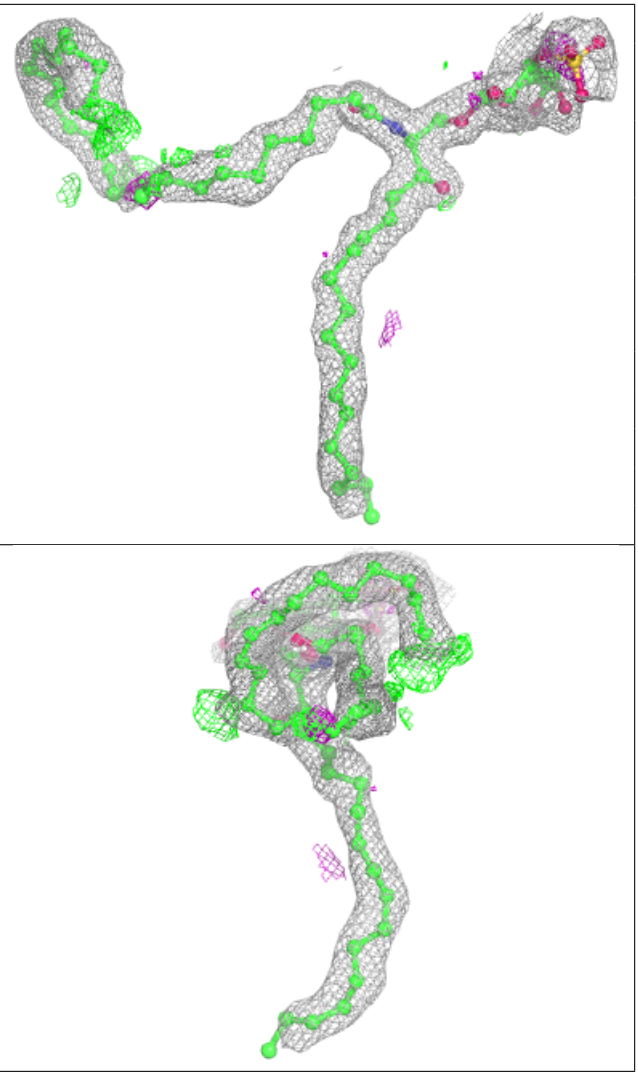






Electron density around CIS C 1002:

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



6.5 Other polymers (i)

There are no such residues in this entry.