



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 8, 2020 – 10:59 AM BST

PDB ID : 4APQ  
Title : Crystal structure of autoreactive-Valpha14-Vbeta6 NKT TCR in complex with CD1d-sulfatide  
Authors : Clarke, A.J.; Le Nours, J.; Rossjohn, J.  
Deposited on : 2012-04-05  
Resolution : 3.00 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
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) : 2.13.1

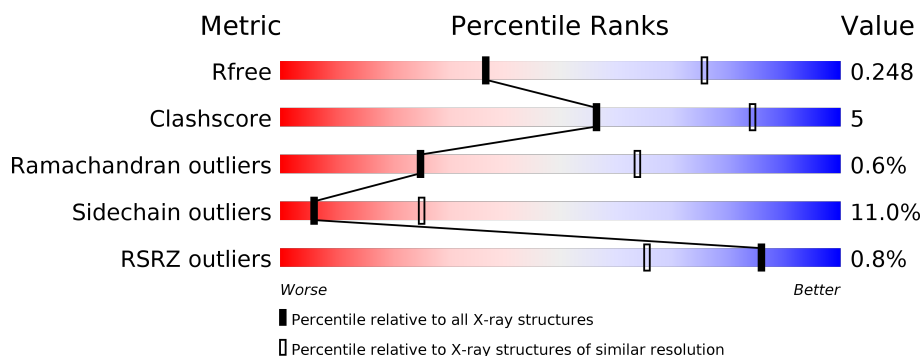
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$ . 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	302	
2	B	99	
3	C	207	
4	D	243	
5	E	2	
6	F	4	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
7	G	3	 <div>67%33%</div>

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 6533 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 ANTIGEN-PRESENTING GLYCOPROTEIN CD1D1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	290	Total	C	N	O	S	0	1	0
			2305	1471	402	418	14			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	201	HIS	ASP	variant	UNP P11609
A	280	GLY	-	expression tag	UNP P11609
A	281	SER	-	expression tag	UNP P11609
A	282	LEU	-	expression tag	UNP P11609
A	283	HIS	-	expression tag	UNP P11609
A	284	HIS	-	expression tag	UNP P11609
A	285	ILE	-	expression tag	UNP P11609
A	286	LEU	-	expression tag	UNP P11609
A	287	ASP	-	expression tag	UNP P11609
A	288	ALA	-	expression tag	UNP P11609
A	289	GLN	-	expression tag	UNP P11609
A	290	LYS	-	expression tag	UNP P11609
A	291	MET	-	expression tag	UNP P11609
A	292	VAL	-	expression tag	UNP P11609
A	293	TRP	-	expression tag	UNP P11609
A	294	ASN	-	expression tag	UNP P11609
A	295	HIS	-	expression tag	UNP P11609
A	296	ARG	-	expression tag	UNP P11609
A	297	HIS	-	expression tag	UNP P11609
A	298	HIS	-	expression tag	UNP P11609
A	299	HIS	-	expression tag	UNP P11609
A	300	HIS	-	expression tag	UNP P11609
A	301	HIS	-	expression tag	UNP P11609
A	302	HIS	-	expression tag	UNP P11609

- Molecule 2 is a protein called BETA-2-MICROGLOBULIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	99	Total	C	N	O	S	0	0	0
			806	514	136	149	7			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	85	ALA	ASP	variant	UNP P01887

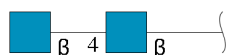
- Molecule 3 is a protein called MOUSE NKT TCR VALPHA14, HUMAN NKT TCR VALPHA14.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	196	Total	C	N	O	S	0	0	0
			1451	899	246	300	6			

- Molecule 4 is a protein called MOUSE NKT TCR AUTOREACTIVE-VBETA6, HUMAN NKT TCR AUTOREACTIVE-VBETA6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	240	Total	C	N	O	S	0	0	0
			1831	1159	317	348	7			

- Molecule 5 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



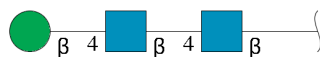
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	2	Total	C	N	O		0	0	0
			28	16	2	10				

- Molecule 6 is an oligosaccharide called alpha-D-mannopyranose-(1-4)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



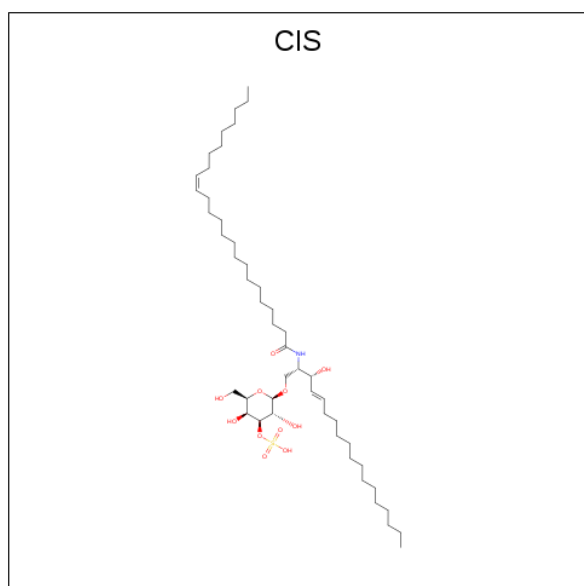
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
6	F	4	Total	C	N	O	0	0	0
			50	28	2	20			

- Molecule 7 is an oligosaccharide called beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
7	G	3	Total	C	N	O	0	0	0
			39	22	2	15			

- Molecule 8 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<sub>48</sub>H<sub>91</sub>NO<sub>11</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
8	A	1	Total	C	N	O	S	0	0
			21	9	1	10	1		

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	1	Total	O	0	0
			1	1		

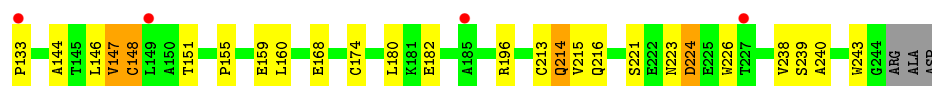
*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	B	1	Total	O	0	0
			1	1		



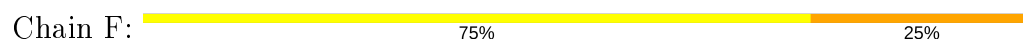




- Molecule 5: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



- Molecule 6: alpha-D-mannopyranose-(1-4)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



- Molecule 7: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	95.11Å 95.11Å 291.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.49 – 3.00 72.85 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.6 (36.49-3.00) 99.5 (72.85-3.00)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.03 (at 3.01Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, $R_{free}$	0.177 , 0.232 0.196 , 0.248	Depositor DCC
$R_{free}$ test set	1392 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	81.8	Xtriage
Anisotropy	0.034	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 70.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	6533	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	94.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.33% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: CIS, BMA, NAG, MAN

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	0.53	0/2379	0.69	0/3240
2	B	0.53	0/832	0.68	0/1132
3	C	0.55	0/1475	0.74	0/2015
4	D	0.52	0/1878	0.74	0/2559
All	All	0.53	0/6564	0.71	0/8946

There are no bond length outliers.

There are no bond angle outliers.

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	2305	0	2179	29	0
2	B	806	0	766	7	0
3	C	1451	0	1340	8	0
4	D	1831	0	1697	18	0
5	E	28	0	25	0	0
6	F	50	0	43	1	0
7	G	39	0	34	0	0
8	A	21	0	13	1	0
9	A	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	B	1	0	0	0	0
All	All	6533	0	6097	61	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 5.

All (61) 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)
4:D:148:CYS:HG	4:D:213:CYS:HG	1.36	0.72
4:D:20:THR:HG22	4:D:77:THR:HB	1.73	0.71
1:A:219:TRP:HB3	1:A:266:LYS:HB2	1.74	0.70
1:A:168:CYS:O	1:A:172:VAL:HG12	1.91	0.69
3:C:166:CYS:HB3	4:D:174:CYS:SG	2.40	0.61
4:D:68:ARG:HD2	4:D:73:SER:O	2.04	0.58
1:A:6:LYS:HA	1:A:107:TYR:HD1	1.68	0.57
1:A:100:LEU:HG	1:A:118:VAL:HG22	1.87	0.57
1:A:172:VAL:HA	1:A:175:LEU:HB2	1.86	0.57
6:F:2:NAG:H61	6:F:3:BMA:H2	1.88	0.56
1:A:189:PRO:HB3	1:A:213:PHE:HB3	1.88	0.56
1:A:89:SER:HB2	1:A:90:PRO:HD3	1.87	0.56
1:A:14:GLN:HB3	1:A:100:LEU:HB2	1.88	0.56
1:A:291:MET:HB3	2:B:15:PRO:O	2.06	0.55
1:A:224:ARG:NH2	1:A:259:ALA:O	2.40	0.55
1:A:89:SER:CB	1:A:90:PRO:HD3	2.37	0.55
4:D:159:GLU:HB2	4:D:216:GLN:HB3	1.90	0.54
4:D:31:THR:HA	4:D:49:SER:O	2.08	0.53
4:D:130:VAL:HG12	4:D:240:ALA:HB1	1.90	0.52
1:A:30:VAL:HB	1:A:38:HIS:HB2	1.92	0.52
1:A:163:LEU:HA	1:A:167:THR:HB	1.92	0.51
1:A:216:LYS:HD2	1:A:245:TRP:CZ2	2.46	0.51
4:D:36:ARG:HB2	4:D:46:ILE:HD11	1.93	0.51
1:A:224:ARG:HB3	1:A:229:GLN:NE2	2.26	0.50
8:A:1296:CIS:C1	8:A:1296:CIS:O6	2.59	0.50
4:D:131:PHE:HB2	4:D:147:VAL:HG12	1.94	0.49
3:C:201:PRO:HG2	3:C:204:THR:HG23	1.95	0.48
4:D:115:THR:HG21	4:D:155:PRO:HB3	1.95	0.48
1:A:107:TYR:HB2	1:A:111:ALA:HB3	1.95	0.48
1:A:92:GLU:HG2	1:A:120:PHE:HZ	1.80	0.47
1:A:127:ARG:NH1	1:A:134:GLN:OE1	2.47	0.47
1:A:124:TYR:CZ	1:A:136:VAL:HG11	2.49	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:191:ALA:HA	1:A:209:HIS:O	2.14	0.47
3:C:5:GLN:HG2	3:C:22:CYS:HB2	1.97	0.46
3:C:153:SER:H	3:C:198:SER:HB3	1.81	0.46
4:D:13:GLY:O	4:D:116:VAL:HA	2.16	0.45
3:C:22:CYS:CB	3:C:90:CYS:HG	2.27	0.45
1:A:67:GLN:O	1:A:71:GLN:HG3	2.17	0.45
2:B:21:ASN:HB3	2:B:70:PHE:CE1	2.51	0.45
3:C:63:SER:O	3:C:75:LEU:HD12	2.18	0.44
2:B:40:LEU:HD23	2:B:45:LYS:HA	2.00	0.44
3:C:13:ARG:O	3:C:16:GLU:HB2	2.18	0.44
4:D:133:PRO:HG3	4:D:144:ALA:HB1	2.00	0.43
1:A:99:GLN:HB2	1:A:119:ALA:HB3	2.01	0.43
4:D:44:ARG:HB3	4:D:60:LEU:HD13	2.01	0.43
1:A:188:LYS:HA	1:A:269:SER:OG	2.19	0.43
1:A:124:TYR:HB2	2:B:60:TRP:NE1	2.34	0.42
4:D:4:ILE:HG22	4:D:27:PHE:HE2	1.85	0.42
1:A:107:TYR:HB2	1:A:111:ALA:CB	2.48	0.42
4:D:160:LEU:HG	4:D:215:VAL:HG22	2.01	0.42
1:A:206:LEU:O	1:A:250:THR:HA	2.20	0.41
1:A:66:LEU:HD22	1:A:70:PHE:CE2	2.55	0.41
4:D:214:GLN:HA	4:D:239:SER:HB3	2.03	0.41
4:D:223:ASN:O	4:D:224:ASP:HB2	2.20	0.41
1:A:293:TRP:CZ3	1:A:295:HIS:HA	2.56	0.41
3:C:85:THR:HG23	3:C:115:ILE:HA	2.03	0.41
2:B:3:LYS:HB3	2:B:29:GLN:O	2.21	0.40
4:D:63:GLY:O	4:D:79:THR:HB	2.21	0.40
1:A:145:LEU:HB3	1:A:146:PRO:HD3	2.03	0.40
2:B:7:ILE:HB	2:B:93:VAL:HG21	2.03	0.40
2:B:24:ASN:HB3	2:B:65:LEU:HD11	2.03	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	289/302 (96%)	278 (96%)	10 (4%)	1 (0%)	41	76
2	B	97/99 (98%)	93 (96%)	3 (3%)	1 (1%)	15	53
3	C	192/207 (93%)	177 (92%)	14 (7%)	1 (0%)	29	68
4	D	238/243 (98%)	220 (92%)	16 (7%)	2 (1%)	19	57
All	All	816/851 (96%)	768 (94%)	43 (5%)	5 (1%)	25	64

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	89	SER
4	D	83	LYS
4	D	224	ASP
2	B	47	PRO
3	C	39	THR

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	245/264 (93%)	224 (91%)	21 (9%)	10	37
2	B	90/93 (97%)	83 (92%)	7 (8%)	12	42
3	C	158/186 (85%)	142 (90%)	16 (10%)	7	29
4	D	187/209 (90%)	156 (83%)	31 (17%)	2	11
All	All	680/752 (90%)	605 (89%)	75 (11%)	6	25

All (75) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	7	ASN
1	A	16	SER
1	A	24	SER
1	A	36	GLN
1	A	41	SER
1	A	43	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	75	VAL
1	A	89	SER
1	A	101	SER
1	A	127	ARG
1	A	173	ARG
1	A	175	LEU
1	A	196	VAL
1	A	203	HIS
1	A	204	ARG
1	A	216	LYS
1	A	243	GLU
1	A	252	ASP
1	A	277	LEU
1	A	289	GLN
1	A	294	ASN
2	B	8	GLN
2	B	40	LEU
2	B	54	MET
2	B	64	ILE
2	B	69	GLU
2	B	70	PHE
2	B	89	GLU
3	C	20	LEU
3	C	21	GLN
3	C	22	CYS
3	C	39	THR
3	C	63	SER
3	C	68	LYS
3	C	115	ILE
3	C	116	VAL
3	C	127	VAL
3	C	133	SER
3	C	148	SER
3	C	166	CYS
3	C	167	VAL
3	C	168	LEU
3	C	200	ILE
3	C	204	THR
4	D	3	ILE
4	D	4	ILE
4	D	5	THR
4	D	7	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
4	D	21	LEU
4	D	25	GLN
4	D	31	THR
4	D	38	ASP
4	D	39	SER
4	D	59	ASP
4	D	61	SER
4	D	72	SER
4	D	77	THR
4	D	79	THR
4	D	82	GLN
4	D	84	ASN
4	D	86	MET
4	D	90	LEU
4	D	146	LEU
4	D	147	VAL
4	D	148	CYS
4	D	151	THR
4	D	168	GLU
4	D	180	LEU
4	D	182	GLU
4	D	196	ARG
4	D	214	GLN
4	D	221	SER
4	D	226	TRP
4	D	238	VAL
4	D	243	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	229	GLN
3	C	193	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 ⓘ

9 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	$\# Z  > 2$
5	NAG	E	1	1,5	14,14,15	0.25	0	17,19,21	0.80	1 (5%)
5	NAG	E	2	5	14,14,15	0.29	0	17,19,21	0.62	0
6	NAG	F	1	1,6	14,14,15	0.29	0	17,19,21	1.37	1 (5%)
6	NAG	F	2	6	14,14,15	0.35	0	17,19,21	1.13	2 (11%)
6	BMA	F	3	6	11,11,12	0.29	0	15,15,17	0.67	0
6	MAN	F	4	6	11,11,12	0.46	0	15,15,17	1.08	2 (13%)
7	NAG	G	1	1,7	14,14,15	0.30	0	17,19,21	0.57	0
7	NAG	G	2	7	14,14,15	0.33	0	17,19,21	1.33	2 (11%)
7	BMA	G	3	7	11,11,12	0.42	0	15,15,17	0.53	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	E	1	1,5	-	1/6/23/26	0/1/1/1
5	NAG	E	2	5	-	0/6/23/26	0/1/1/1
6	NAG	F	1	1,6	-	2/6/23/26	0/1/1/1
6	NAG	F	2	6	-	0/6/23/26	0/1/1/1
6	BMA	F	3	6	-	0/2/19/22	0/1/1/1
6	MAN	F	4	6	-	0/2/19/22	0/1/1/1
7	NAG	G	1	1,7	-	2/6/23/26	0/1/1/1
7	NAG	G	2	7	-	0/6/23/26	0/1/1/1
7	BMA	G	3	7	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	F	1	NAG	C1-O5-C5	5.05	119.03	112.19
7	G	2	NAG	O4-C4-C3	3.93	119.42	110.35
6	F	4	MAN	C1-O5-C5	2.81	116.00	112.19
6	F	2	NAG	C1-C2-N2	-2.62	106.01	110.49
6	F	2	NAG	O4-C4-C3	-2.55	104.45	110.35
6	F	4	MAN	C1-C2-C3	2.48	112.71	109.67
5	E	1	NAG	C1-O5-C5	2.43	115.48	112.19
7	G	2	NAG	C1-O5-C5	2.22	115.19	112.19

There are no chirality outliers.

All (5) torsion outliers are listed below:

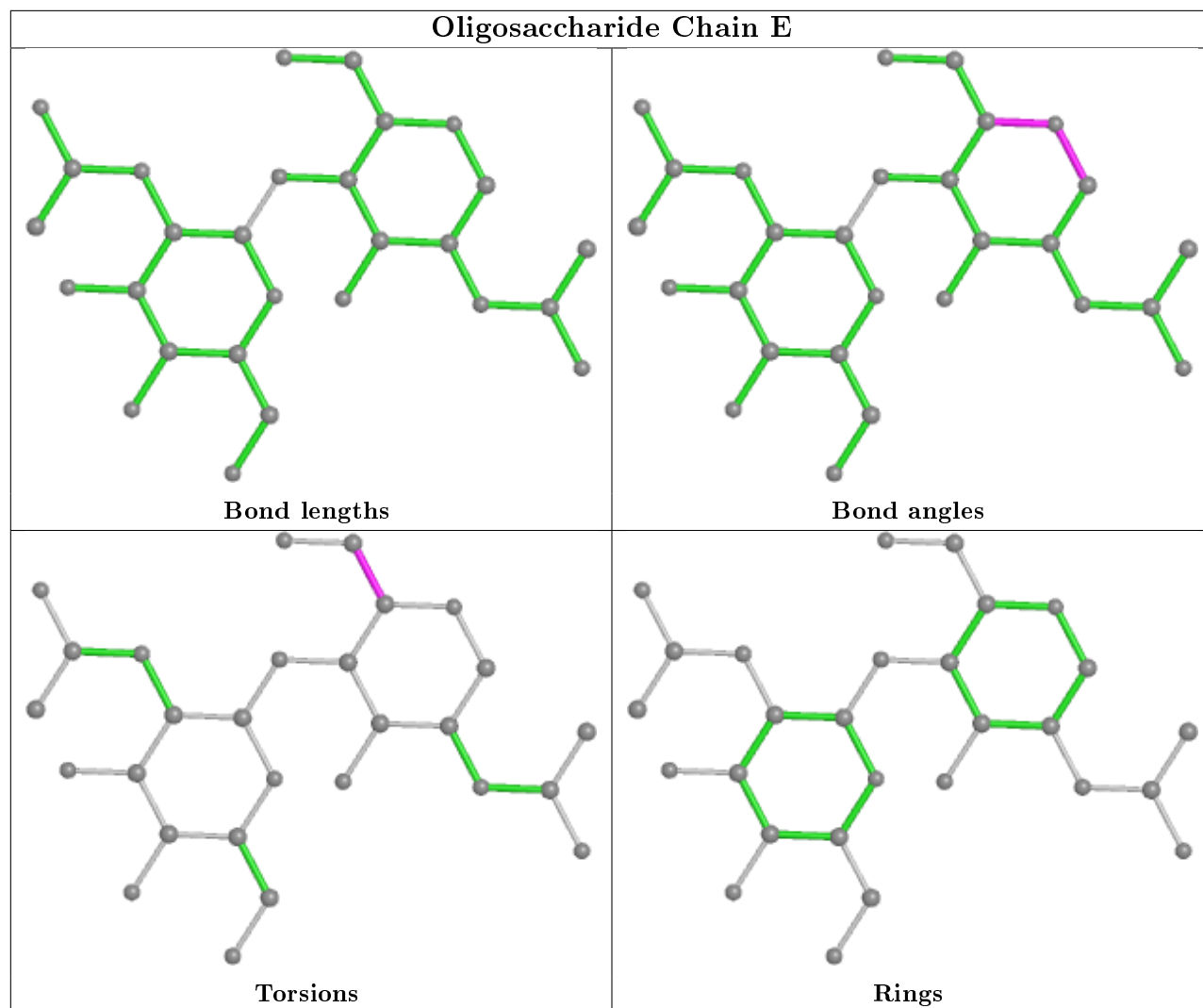
Mol	Chain	Res	Type	Atoms
6	F	1	NAG	O5-C5-C6-O6
6	F	1	NAG	C4-C5-C6-O6
7	G	1	NAG	C4-C5-C6-O6
7	G	1	NAG	O5-C5-C6-O6
5	E	1	NAG	C4-C5-C6-O6

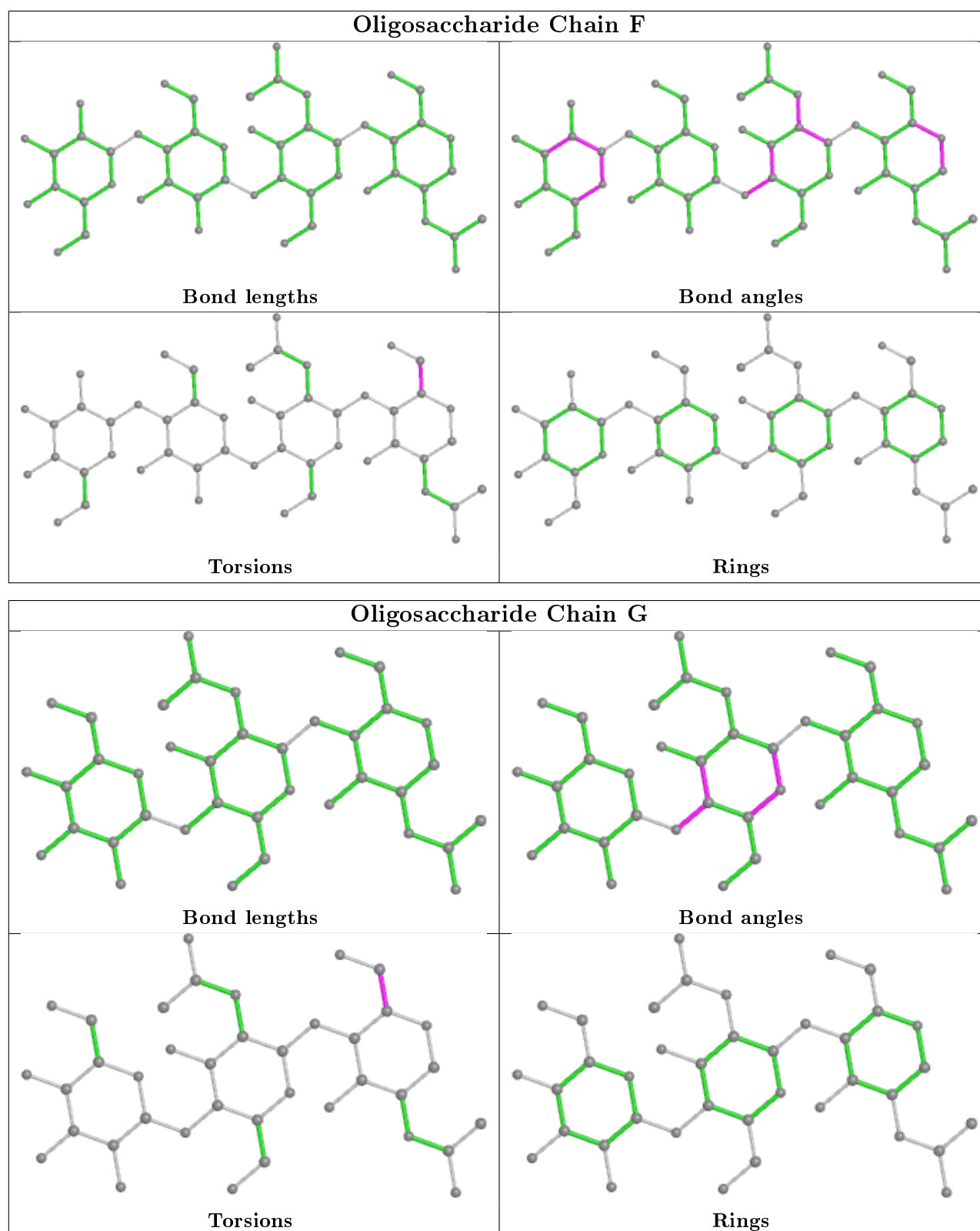
There are no ring outliers.

2 monomers are involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	F	2	NAG	1	0
6	F	3	BMA	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.





## 5.6 Ligand geometry [i](#)

1 ligand is 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	$\# Z  > 2$
8	CIS	A	1296	-	21,21,61	1.13	1 (4%)	23,29,72	1.65	3 (13%)

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
8	CIS	A	1296	-	-	9/13/33/78	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	A	1296	CIS	O-C43	3.27	1.45	1.40

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	A	1296	CIS	O10-S-O9	4.40	123.80	108.49
8	A	1296	CIS	O2-C18-N	-3.76	121.09	124.89
8	A	1296	CIS	C45-O4-S	-3.65	111.82	118.88

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
8	A	1296	CIS	C1-C-O-C43
8	A	1296	CIS	C46-C45-O4-S
8	A	1296	CIS	C45-O4-S-O10
8	A	1296	CIS	C45-O4-S-O8
8	A	1296	CIS	C46-C47-C48-O7
8	A	1296	CIS	O6-C47-C48-O7
8	A	1296	CIS	C45-O4-S-O9

*Continued on next page...*

*Continued from previous page...*

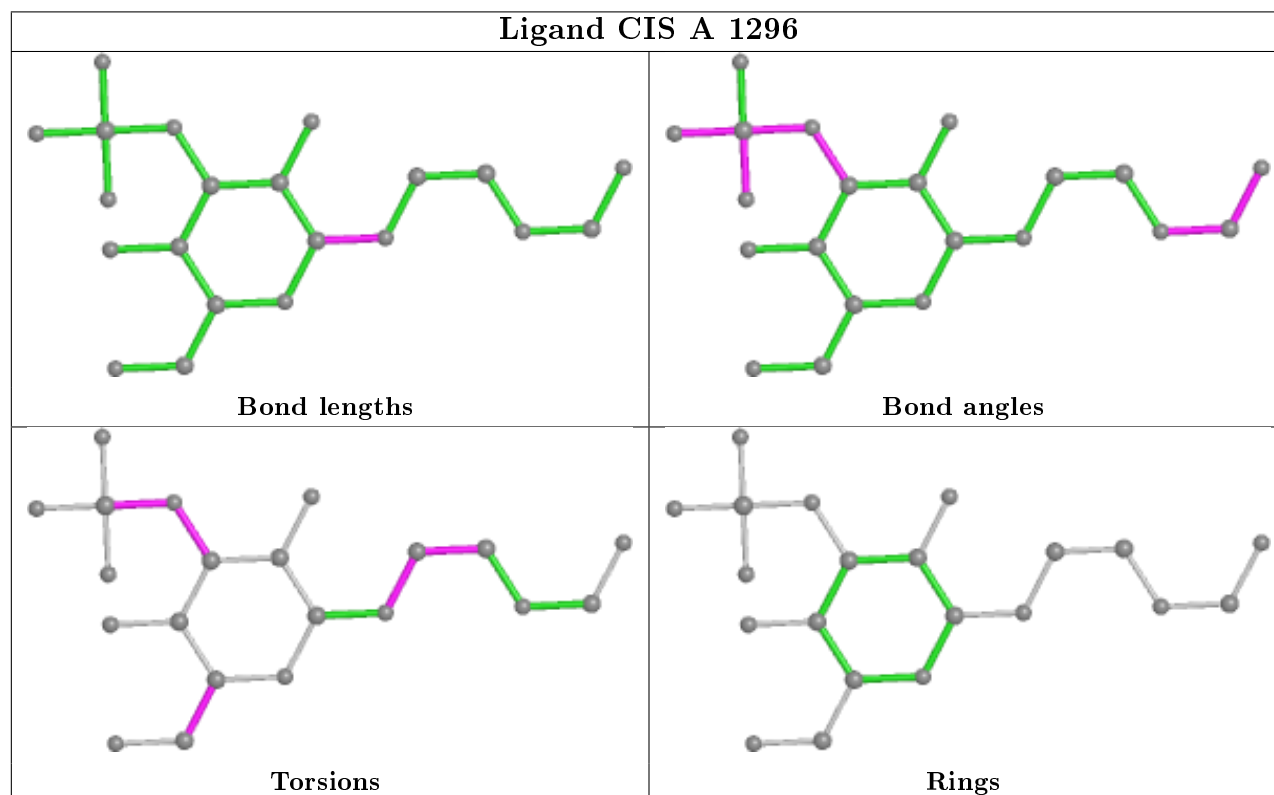
Mol	Chain	Res	Type	Atoms
8	A	1296	CIS	C44-C45-O4-S
8	A	1296	CIS	O-C-C1-N

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	A	1296	CIS	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 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 than 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.



## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	290/302 (96%)	-0.01	1 (0%) 94 84	57, 79, 113, 136	1 (0%)
2	B	99/99 (100%)	0.05	0 100 100	54, 76, 103, 112	0
3	C	196/207 (94%)	-0.09	1 (0%) 91 75	60, 98, 176, 185	0
4	D	240/243 (98%)	0.05	5 (2%) 63 34	62, 105, 153, 168	0
All	All	825/851 (96%)	-0.01	7 (0%) 86 65	54, 87, 156, 185	1 (0%)

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	149	LEU	2.6
4	D	133	PRO	2.6
4	D	227	THR	2.5
1	A	8	TYR	2.2
3	C	152	VAL	2.2
4	D	129	ALA	2.1
4	D	185	ALA	2.1

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

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

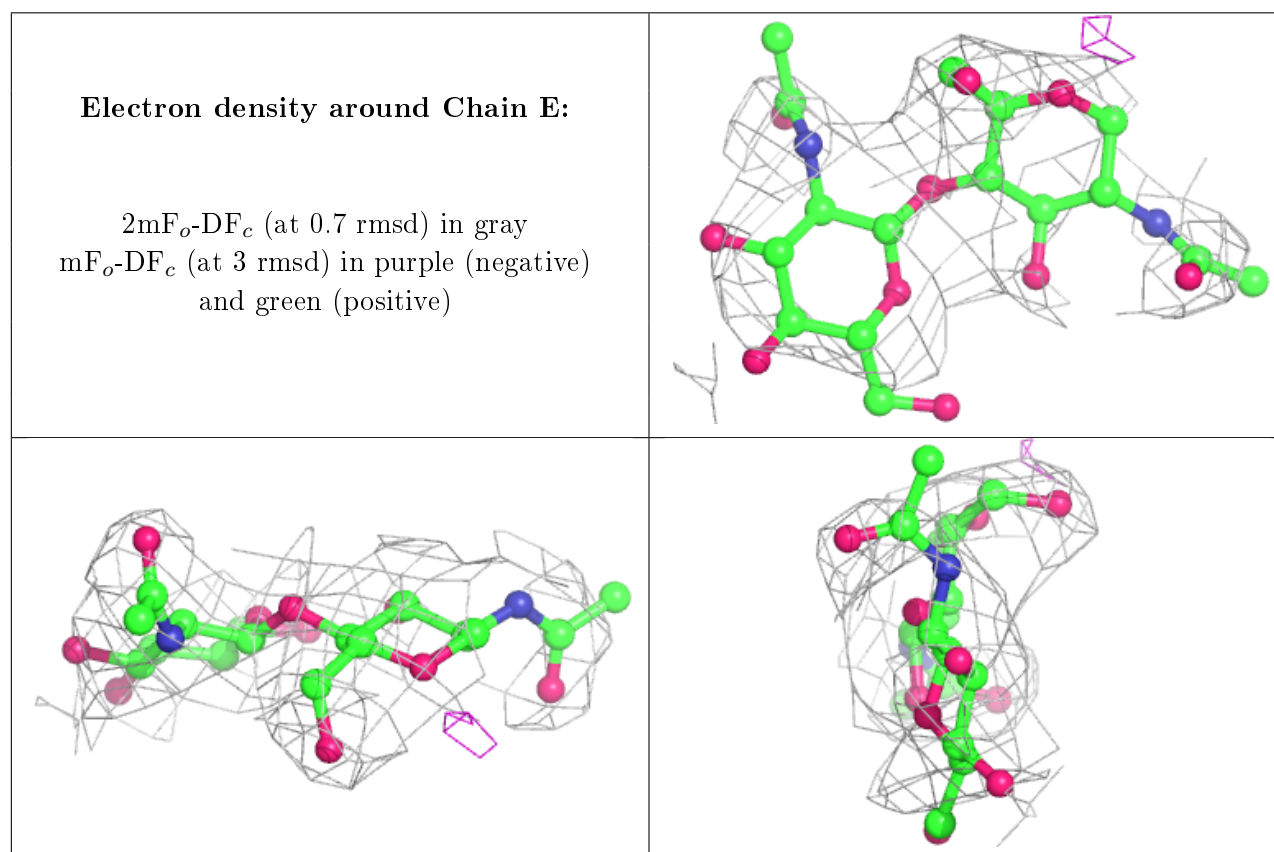
### 6.3 Carbohydrates ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> 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.



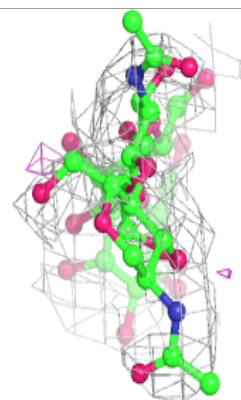
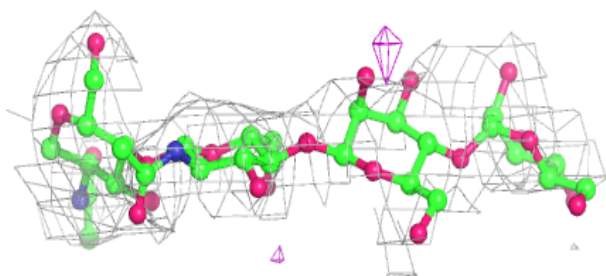
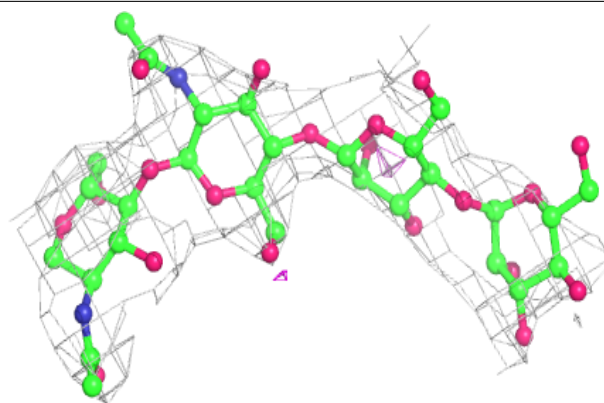
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
6	BMA	F	3	11/12	0.68	0.21	162,171,175,179	0
7	BMA	G	3	11/12	0.75	0.24	167,172,178,181	0
6	MAN	F	4	11/12	0.76	0.22	184,188,194,194	0
5	NAG	E	2	14/15	0.85	0.32	117,129,136,138	0
6	NAG	F	2	14/15	0.91	0.27	123,132,143,153	0
5	NAG	E	1	14/15	0.92	0.17	94,109,116,116	0
6	NAG	F	1	14/15	0.92	0.18	93,100,110,117	0
7	NAG	G	1	14/15	0.92	0.19	103,108,113,121	0
7	NAG	G	2	14/15	0.93	0.21	133,141,155,159	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.

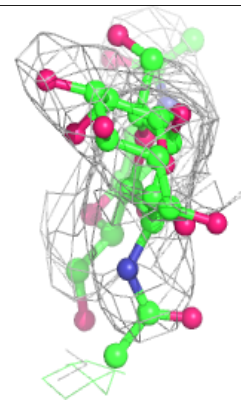
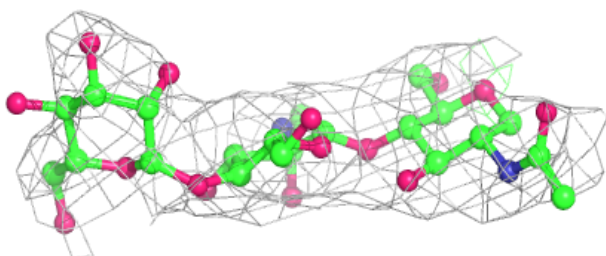
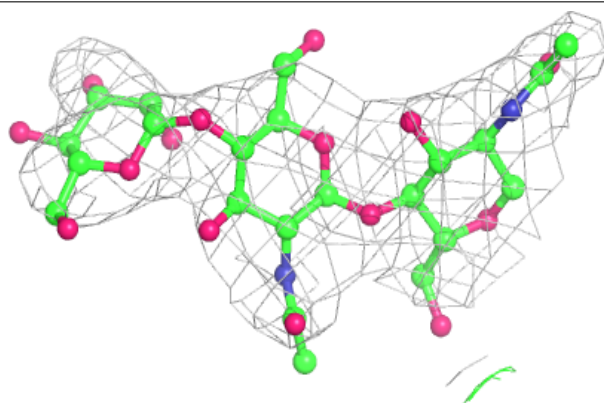


**Electron density around Chain F:**

$2mF_o-DF_c$  (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)

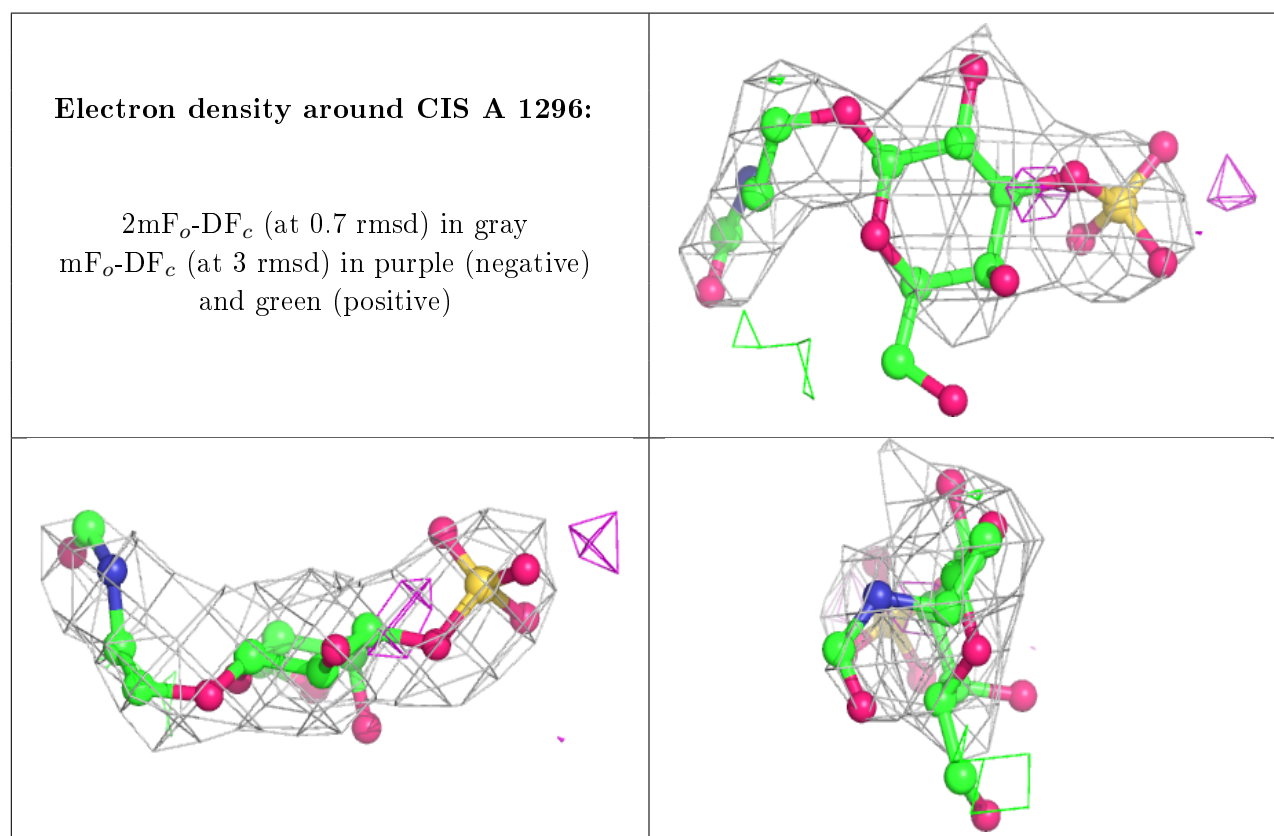


## 6.4 Ligands [i](#)

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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
8	CIS	A	1296	21/61	0.83	0.39	130,165,175,177	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.



## 6.5 Other polymers [i](#)

There are no such residues in this entry.