



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 9, 2020 – 09:30 AM BST

PDB ID : 6MIY  
Title : Crystal structure of the mCD1d/xxa (JJ239)/iNKTCR ternary complex  
Authors : Zajonc, D.M.; Bitra, A.; Janssens, J.  
Deposited on : 2018-09-20  
Resolution : 2.75 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

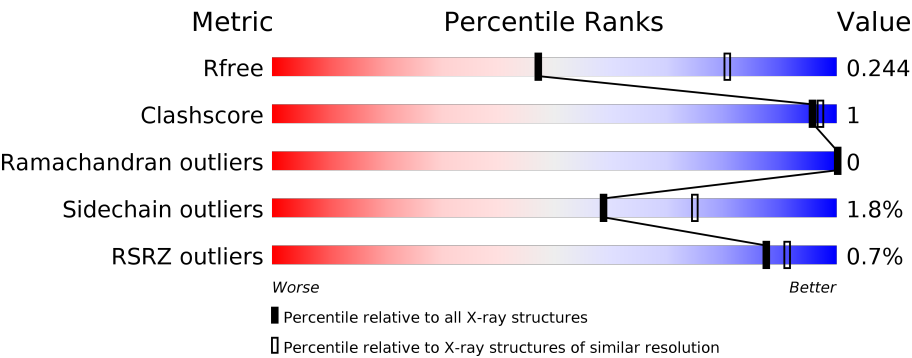
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.75 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	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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	C	209	<div><div>%</div><div><div></div><div>94%</div><div></div></div><div><div></div><div></div><div></div></div></div>
1	G	209	<div><div>%</div><div><div></div><div>93%</div><div></div></div><div><div></div><div></div><div></div></div></div>
2	D	241	<div><div></div><div><div></div><div>95%</div><div></div></div><div><div></div><div></div><div>5%</div></div></div>
2	H	241	<div><div></div><div><div></div><div>96%</div><div></div></div><div><div></div><div></div><div></div></div></div>
3	A	285	<div><div></div><div><div></div><div>90%</div><div></div></div><div><div></div><div>5%</div><div>5%</div></div></div>
3	E	285	<div><div>%</div><div><div></div><div>90%</div><div></div></div><div><div></div><div>6%</div><div></div></div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	B	99	
4	F	99	
5	I	2	
5	J	2	
5	L	2	
5	N	2	
6	K	3	
6	M	3	

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
7	NA	D	301	-	-	-	X

## 2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 13394 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 receptor alpha variable 11,T cell receptor alpha variable 11,T cell receptor alpha joining 18,Human nkt tcr alpha chain, CHIMERIC PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	200	Total	C	N	O	S	0	0	0
			1542	956	264	314	8			
1	G	200	Total	C	N	O	S	0	0	0
			1533	950	262	313	8			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	MET	-	initiating methionine	UNP A0A0B4J1J9
C	113	ILE	-	linker	UNP A0A0B4J1J9
G	0	MET	-	initiating methionine	UNP A0A0B4J1J9
G	113	ILE	-	linker	UNP A0A0B4J1J9

- Molecule 2 is a protein called Beta-chain,T cell receptor chain,T cell receptor beta constant 2, CHIMERIC PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	240	Total	C	N	O	S	0	0	0
			1877	1178	334	359	6			
2	H	239	Total	C	N	O	S	0	0	0
			1872	1175	334	357	6			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	95	ASP	-	linker	UNP A2NTY6
D	96	GLU	-	linker	UNP A2NTY6
D	97	GLY	-	linker	UNP A2NTY6
D	98	TYR	-	linker	UNP A2NTY6
D	130	ALA	ALA	linker	UNP A0N8J3
D	168	CYS	SER	variant	UNP A0A5B9

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	186	SER	CYS	variant	UNP A0A5B9
H	95	ASP	-	linker	UNP A2NTY6
H	96	GLU	-	linker	UNP A2NTY6
H	97	GLY	-	linker	UNP A2NTY6
H	98	TYR	-	linker	UNP A2NTY6
H	130	ALA	ALA	linker	UNP A0N8J3
H	168	CYS	SER	variant	UNP A0A5B9
H	186	SER	CYS	variant	UNP A0A5B9

- Molecule 3 is a protein called Antigen-presenting glycoprotein CD1d1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	271	Total	C	N	O	S	0	0	0
			2171	1386	371	401	13			
3	E	274	Total	C	N	O	S	0	0	0
			2193	1397	378	405	13			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	280	HIS	-	expression tag	UNP A0A0R4J090
A	281	HIS	-	expression tag	UNP A0A0R4J090
A	282	HIS	-	expression tag	UNP A0A0R4J090
A	283	HIS	-	expression tag	UNP A0A0R4J090
A	284	HIS	-	expression tag	UNP A0A0R4J090
A	285	HIS	-	expression tag	UNP A0A0R4J090
E	280	HIS	-	expression tag	UNP A0A0R4J090
E	281	HIS	-	expression tag	UNP A0A0R4J090
E	282	HIS	-	expression tag	UNP A0A0R4J090
E	283	HIS	-	expression tag	UNP A0A0R4J090
E	284	HIS	-	expression tag	UNP A0A0R4J090
E	285	HIS	-	expression tag	UNP A0A0R4J090

- Molecule 4 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	98	Total	C	N	O	S	0	0	0
			801	511	135	148	7			
4	F	98	Total	C	N	O	S	0	0	0
			801	511	135	148	7			

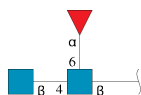
- Molecule 5 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-a

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



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

- Molecule 6 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
6	K	3	Total	C	N	O	0	0	0
			38	22	2	14			
6	M	3	Total	C	N	O	0	0	0
			38	22	2	14			

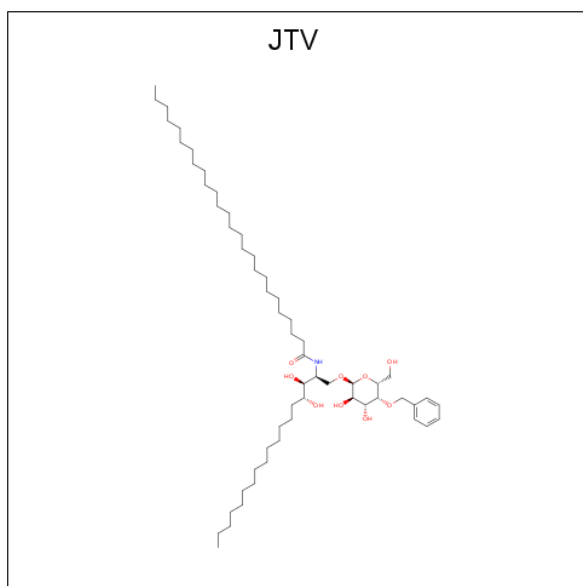
- Molecule 7 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	H	1	Total	Na	0	0
			1	1		
7	G	1	Total	Na	0	0
			1	1		
7	D	1	Total	Na	0	0
			1	1		
7	C	1	Total	Na	0	0
			1	1		

- Molecule 8 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	D	1	Total	Cl	0	0
			1	1		

- Molecule 9 is N-{(2S,3S,4R)-1-[(4-O-benzyl-alpha-D-galactopyranosyl)oxy]-3,4-dihydroxyoctadecan-2-yl}hexacosanamide (three-letter code: JTV) (formula: C<sub>57</sub>H<sub>105</sub>NO<sub>9</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	A	1	Total	C	N	O	0	0
			67	57	1	9		
9	E	1	Total	C	N	O	0	0
			67	57	1	9		

- Molecule 10 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	A	1	Total	C	O	0	0
			6	3	3		
10	A	1	Total	C	O	0	0
			6	3	3		
10	E	1	Total	C	O	0	0
			6	3	3		
10	G	1	Total	C	O	0	0
			6	3	3		

- Molecule 11 is water.

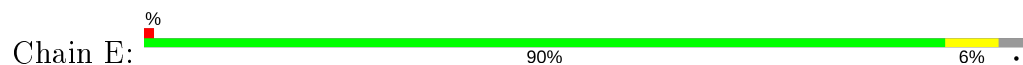
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	C	17	Total	O	0	0
			17	17		
11	D	35	Total	O	0	0
			35	35		
11	A	46	Total	O	0	0
			46	46		
11	B	14	Total	O	0	0
			14	14		
11	E	46	Total	O	0	0
			46	46		
11	G	34	Total	O	0	0
			34	34		
11	H	43	Total	O	0	0
			43	43		
11	F	18	Total	O	0	0
			18	18		







- Molecule 3: Antigen-presenting glycoprotein CD1d1



- Molecule 4: Beta-2-microglobulin



- Molecule 4: Beta-2-microglobulin



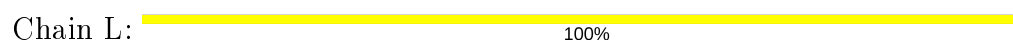
- Molecule 5: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



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



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



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

Chain N:  100%

UAG1  
UAG2

- Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain K:  100%

UAG1  
UAG2  
FUC3

- Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain M:  67% 33%

UAG1  
UAG2  
FUC3

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	78.86Å 192.48Å 150.42Å 90.00° 90.79° 90.00°	Depositor
Resolution (Å)	47.42 – 2.75 47.37 – 2.75	Depositor EDS
% Data completeness (in resolution range)	96.9 (47.42-2.75) 93.2 (47.37-2.75)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.29 (at 2.77Å)	Xtriage
Refinement program	REFMAC 5.8.0222	Depositor
R, $R_{free}$	0.192 , 0.247 0.195 , 0.244	Depositor DCC
$R_{free}$ test set	2848 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	56.7	Xtriage
Anisotropy	0.023	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 14.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.360 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	13394	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.65% 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: GOL, NAG, CL, NA, JTV, FUC

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	C	0.28	0/1569	0.50	0/2132
1	G	0.28	0/1558	0.51	0/2115
2	D	0.28	0/1928	0.50	0/2629
2	H	0.29	0/1923	0.50	0/2622
3	A	0.29	0/2234	0.47	0/3037
3	E	0.30	0/2258	0.47	0/3071
4	B	0.28	0/827	0.50	0/1125
4	F	0.28	0/827	0.51	0/1125
All	All	0.29	0/13124	0.49	0/17856

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	C	1542	0	1477	2	0
1	G	1533	0	1452	4	0
2	D	1877	0	1779	4	0
2	H	1872	0	1776	2	0
3	A	2171	0	2071	7	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	E	2193	0	2090	8	0
4	B	801	0	761	0	0
4	F	801	0	761	1	0
5	I	28	0	25	0	0
5	J	28	0	25	0	0
5	L	28	0	25	0	0
5	N	28	0	25	0	0
6	K	38	0	34	0	0
6	M	38	0	34	0	0
7	C	1	0	0	0	0
7	D	1	0	0	0	0
7	G	1	0	0	0	0
7	H	1	0	0	0	0
8	D	1	0	0	1	0
9	A	67	0	0	0	0
9	E	67	0	0	0	0
10	A	12	0	16	0	0
10	E	6	0	8	0	0
10	G	6	0	8	0	0
11	A	46	0	0	0	0
11	B	14	0	0	0	0
11	C	17	0	0	0	0
11	D	35	0	0	0	0
11	E	46	0	0	0	0
11	F	18	0	0	0	0
11	G	34	0	0	0	0
11	H	43	0	0	0	0
All	All	13394	0	12367	28	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 1.

All (28) 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)
3:E:168:CYS:HB3	3:E:169:PRO:HD3	1.74	0.69
1:G:181:ASN:N	1:G:181:ASN:HD22	1.95	0.64
3:A:145:LEU:HB3	3:A:146:PRO:HD3	1.95	0.47
4:F:32:PRO:HB2	4:F:33:PRO:HD2	1.96	0.47
1:C:126:LEU:N	1:C:126:LEU:HD12	2.30	0.47
3:E:145:LEU:HB3	3:E:146:PRO:HD3	1.96	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:191:ALA:HA	3:A:209:HIS:O	2.16	0.46
3:E:191:ALA:HA	3:E:209:HIS:O	2.17	0.45
2:H:78:LEU:N	2:H:78:LEU:HD12	2.31	0.44
1:G:126:LEU:N	1:G:126:LEU:HD12	2.33	0.44
2:D:21:LEU:HD22	2:D:106:THR:HG21	2.00	0.43
2:D:46:ILE:HG22	2:D:47:HIS:CD2	2.53	0.43
3:A:168:CYS:HB3	3:A:169:PRO:CD	2.48	0.43
2:D:44:ARG:NE	8:D:302:CL:CL	2.78	0.43
3:E:35:LEU:HD12	3:E:183:LEU:HD23	2.00	0.43
3:A:168:CYS:HB3	3:A:169:PRO:HD3	2.01	0.43
1:C:114:PRO:HG3	1:C:163:VAL:HG21	2.01	0.43
3:E:216:LYS:N	3:E:217:PRO:CD	2.82	0.42
2:H:31:ASN:HA	2:H:49:SER:O	2.19	0.42
1:G:181:ASN:N	1:G:181:ASN:ND2	2.66	0.42
3:A:216:LYS:N	3:A:217:PRO:CD	2.83	0.42
2:D:78:LEU:N	2:D:78:LEU:HD12	2.34	0.42
1:G:50:LEU:HD23	1:G:57:THR:HG23	2.01	0.42
3:A:35:LEU:HD12	3:A:183:LEU:HD23	2.02	0.42
3:E:88:MET:CE	3:E:145:LEU:HD23	2.49	0.42
3:E:89:SER:HA	3:E:90:PRO:HA	1.86	0.41
3:E:219:TRP:HB3	3:E:266:LYS:HB2	2.02	0.41
3:A:89:SER:HA	3:A:90:PRO:HA	1.86	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	C	196/209 (94%)	183 (93%)	13 (7%)	0	100	100
1	G	195/209 (93%)	186 (95%)	9 (5%)	0	100	100
2	D	238/241 (99%)	236 (99%)	2 (1%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	H	237/241 (98%)	233 (98%)	4 (2%)	0	100	100
3	A	267/285 (94%)	257 (96%)	10 (4%)	0	100	100
3	E	272/285 (95%)	267 (98%)	5 (2%)	0	100	100
4	B	96/99 (97%)	95 (99%)	1 (1%)	0	100	100
4	F	96/99 (97%)	95 (99%)	1 (1%)	0	100	100
All	All	1597/1668 (96%)	1552 (97%)	45 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	C	178/188 (95%)	177 (99%)	1 (1%)	86	90
1	G	175/188 (93%)	171 (98%)	4 (2%)	50	69
2	D	202/208 (97%)	196 (97%)	6 (3%)	41	61
2	H	202/208 (97%)	197 (98%)	5 (2%)	47	67
3	A	235/249 (94%)	232 (99%)	3 (1%)	69	81
3	E	237/249 (95%)	234 (99%)	3 (1%)	69	81
4	B	90/93 (97%)	89 (99%)	1 (1%)	73	84
4	F	90/93 (97%)	87 (97%)	3 (3%)	38	58
All	All	1409/1476 (96%)	1383 (98%)	26 (2%)	59	75

All (26) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	181	ASN
2	D	25	GLN
2	D	168	CYS
2	D	172	GLN
2	D	184	ARG

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
2	D	190	ARG
2	D	235	GLU
3	A	15	MET
3	A	97	GLU
3	A	276	ILE
4	B	70	PHE
3	E	15	MET
3	E	173	ARG
3	E	276	ILE
1	G	50	LEU
1	G	155	ASP
1	G	181	ASN
1	G	193	ASN
2	H	25	GLN
2	H	75	SER
2	H	168	CYS
2	H	184	ARG
2	H	190	ARG
4	F	3	LYS
4	F	47	PRO
4	F	70	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (18) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	31	ASN
1	C	181	ASN
3	A	7	ASN
3	A	62	GLN
3	A	110	ASN
3	A	227	GLN
3	A	267	HIS
3	A	273	GLN
3	E	7	ASN
3	E	62	GLN
3	E	110	ASN
3	E	267	HIS
1	G	31	ASN
1	G	181	ASN
2	H	24	ASN
2	H	204	HIS
2	H	217	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	H	230	GLN

### 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 ⓘ

14 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	I	1	3,5	14,14,15	0.38	0	17,19,21	0.74	0
5	NAG	I	2	5	14,14,15	0.28	0	17,19,21	0.78	0
5	NAG	J	1	3,5	14,14,15	0.38	0	17,19,21	0.78	0
5	NAG	J	2	5	14,14,15	0.34	0	17,19,21	0.78	0
6	NAG	K	1	3,6	14,14,15	0.31	0	17,19,21	0.69	0
6	NAG	K	2	6	14,14,15	0.39	0	17,19,21	0.87	0
6	FUC	K	3	6	10,10,11	0.36	0	14,14,16	0.61	0
5	NAG	L	1	3,5	14,14,15	0.40	0	17,19,21	0.94	1 (5%)
5	NAG	L	2	5	14,14,15	0.34	0	17,19,21	0.90	1 (5%)
6	NAG	M	1	3,6	14,14,15	0.29	0	17,19,21	0.81	1 (5%)
6	NAG	M	2	6	14,14,15	0.41	0	17,19,21	0.78	0
6	FUC	M	3	6	10,10,11	0.36	0	14,14,16	0.59	0
5	NAG	N	1	3,5	14,14,15	0.36	0	17,19,21	0.69	0
5	NAG	N	2	5	14,14,15	0.30	0	17,19,21	0.79	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	I	1	3,5	-	0/6/23/26	0/1/1/1
5	NAG	I	2	5	-	2/6/23/26	0/1/1/1
5	NAG	J	1	3,5	-	0/6/23/26	0/1/1/1
5	NAG	J	2	5	-	2/6/23/26	0/1/1/1
6	NAG	K	1	3,6	-	0/6/23/26	0/1/1/1
6	NAG	K	2	6	-	2/6/23/26	0/1/1/1
6	FUC	K	3	6	-	-	0/1/1/1
5	NAG	L	1	3,5	-	0/6/23/26	0/1/1/1
5	NAG	L	2	5	-	0/6/23/26	0/1/1/1
6	NAG	M	1	3,6	-	0/6/23/26	0/1/1/1
6	NAG	M	2	6	-	2/6/23/26	0/1/1/1
6	FUC	M	3	6	-	-	0/1/1/1
5	NAG	N	1	3,5	-	0/6/23/26	0/1/1/1
5	NAG	N	2	5	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	L	1	NAG	C1-O5-C5	2.74	115.90	112.19
5	L	2	NAG	C4-C3-C2	2.31	114.41	111.02
6	M	1	NAG	O5-C5-C6	2.05	110.42	107.20

There are no chirality outliers.

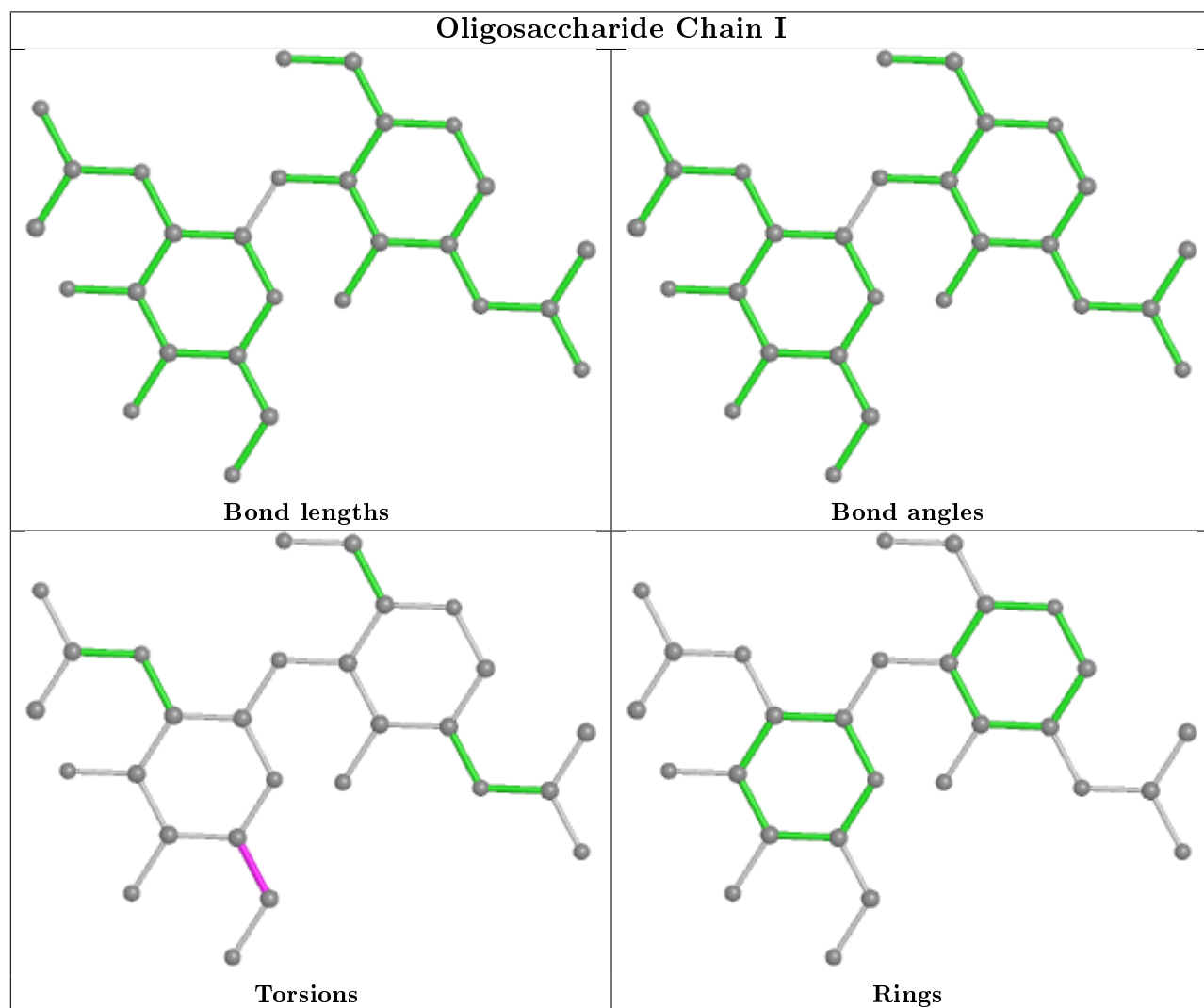
All (8) torsion outliers are listed below:

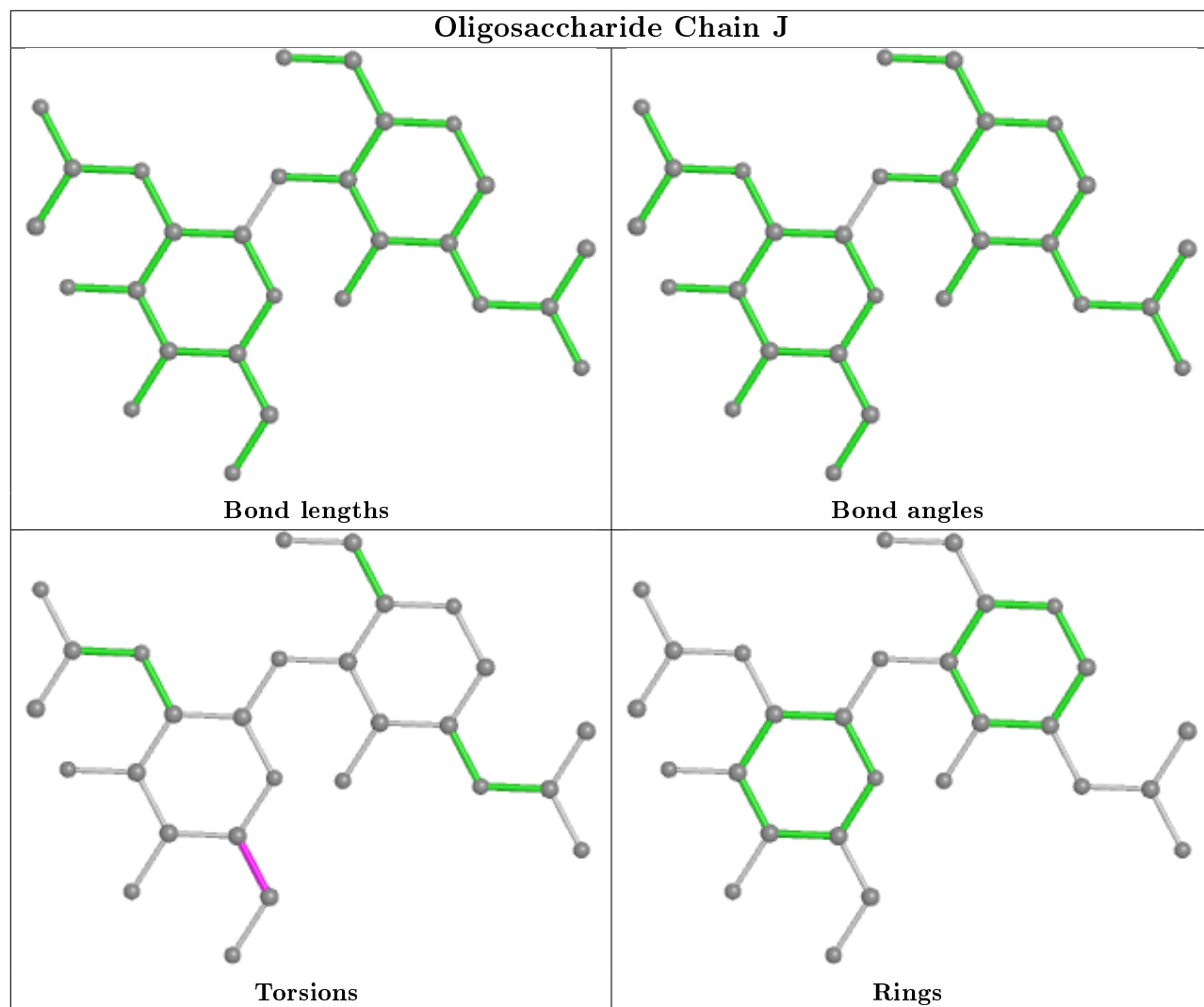
Mol	Chain	Res	Type	Atoms
5	I	2	NAG	C4-C5-C6-O6
6	K	2	NAG	O5-C5-C6-O6
6	K	2	NAG	C4-C5-C6-O6
5	I	2	NAG	O5-C5-C6-O6
5	J	2	NAG	C4-C5-C6-O6
5	J	2	NAG	O5-C5-C6-O6
6	M	2	NAG	O5-C5-C6-O6
6	M	2	NAG	C4-C5-C6-O6

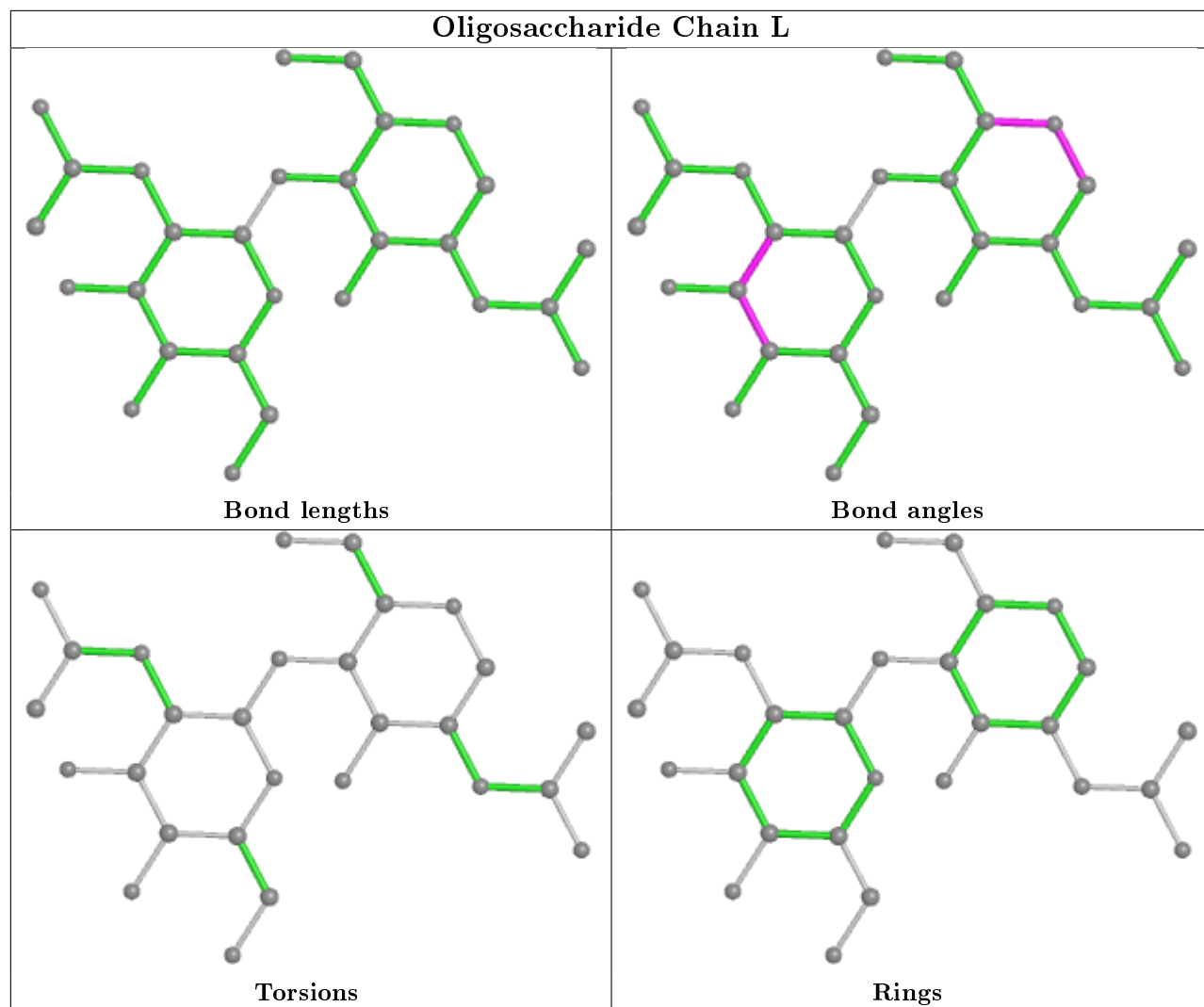
There are no ring outliers.

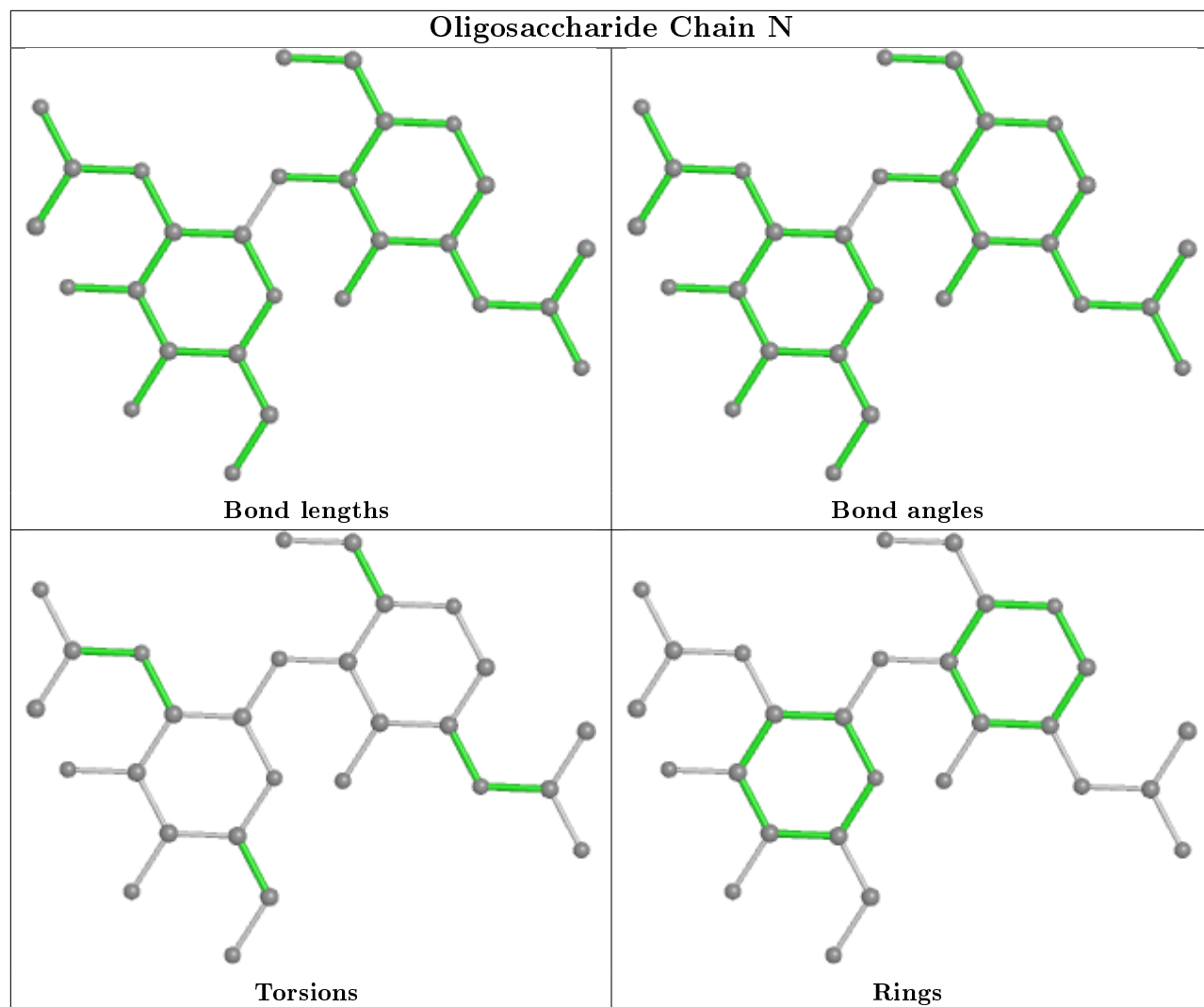
No monomer is involved in short contacts.

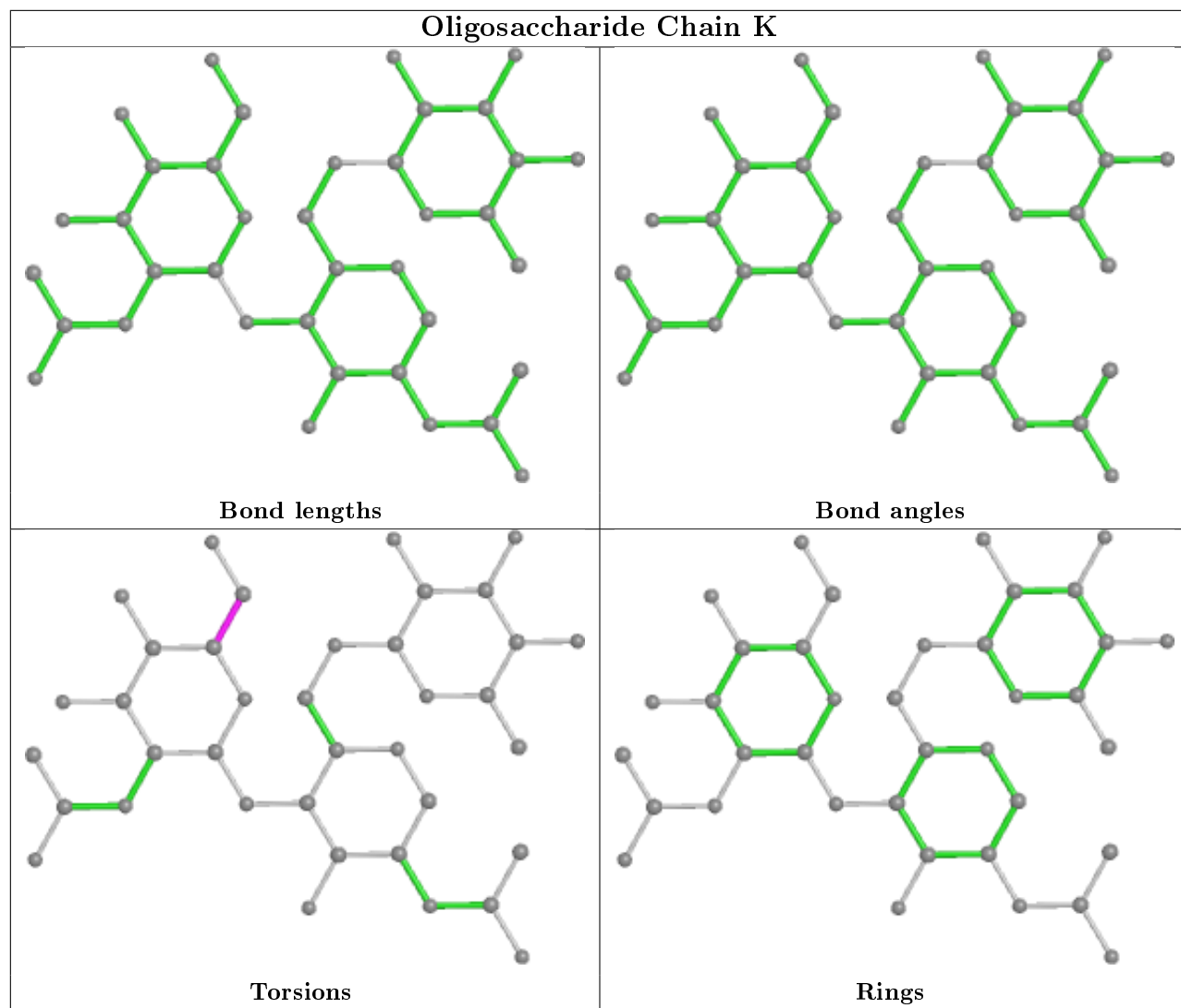
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.



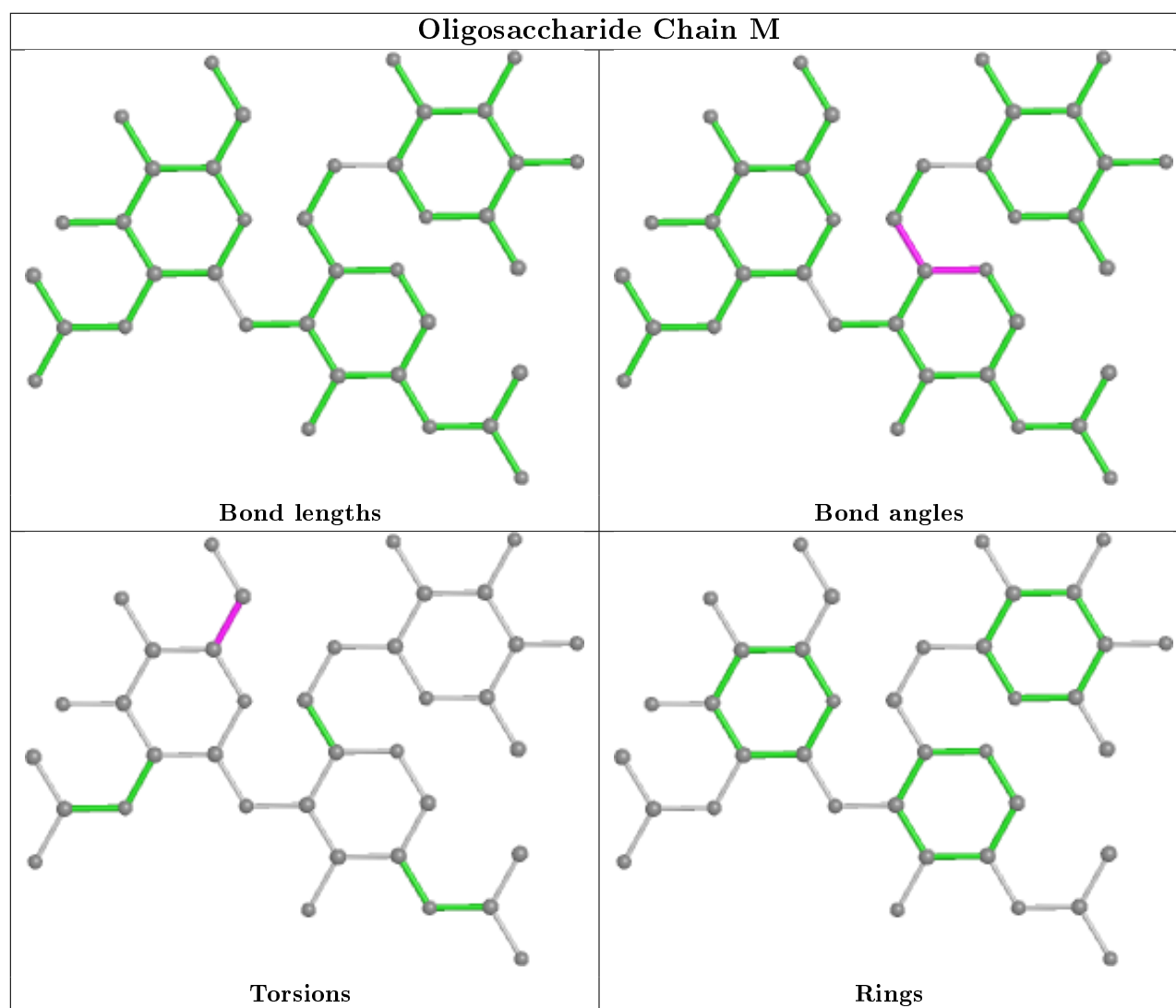












## 5.6 Ligand geometry [i](#)

Of 11 ligands modelled in this entry, 5 are monoatomic - leaving 6 for Mogul analysis.

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	$\# Z  > 2$
10	GOL	A	310	-	5,5,5	0.32	0	5,5,5	0.18	0
10	GOL	A	309	-	5,5,5	0.34	0	5,5,5	0.13	0
10	GOL	E	509	-	5,5,5	0.32	0	5,5,5	0.15	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
10	GOL	G	301	-	5,5,5	0.30	0	5,5,5	0.15	0
9	JTV	E	508	-	68,68,68	0.80	2 (2%)	75,79,79	0.74	0
9	JTV	A	308	-	68,68,68	0.81	2 (2%)	75,79,79	0.78	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
10	GOL	A	310	-	-	1/4/4/4	-
10	GOL	A	309	-	-	0/4/4/4	-
10	GOL	E	509	-	-	2/4/4/4	-
10	GOL	G	301	-	-	0/4/4/4	-
9	JTV	E	508	-	-	14/63/83/83	0/2/2/2
9	JTV	A	308	-	-	12/63/83/83	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	A	308	JTV	CAT-CAU	-4.18	1.40	1.50
9	E	508	JTV	CAT-CAU	-3.97	1.41	1.50
9	E	508	JTV	O1-C1	2.60	1.44	1.40
9	A	308	JTV	O1-C1	2.46	1.44	1.40

There are no bond angle outliers.

There are no chirality outliers.

All (29) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
10	E	509	GOL	O1-C1-C2-O2
10	E	509	GOL	O1-C1-C2-C3
9	E	508	JTV	CCK-CCL-CCM-CCN
9	A	308	JTV	O6-C2-C3-O2
9	A	308	JTV	CCK-CCL-CCM-CCN
9	E	508	JTV	O6-C2-C3-O2
9	E	508	JTV	CCG-CCH-CCI-CCJ
9	A	308	JTV	CCB-CCC-CCD-CCE
9	E	508	JTV	OAR-CAQ-CBG-CBH
9	A	308	JTV	CCA-CCB-CCC-CCD

*Continued on next page...*

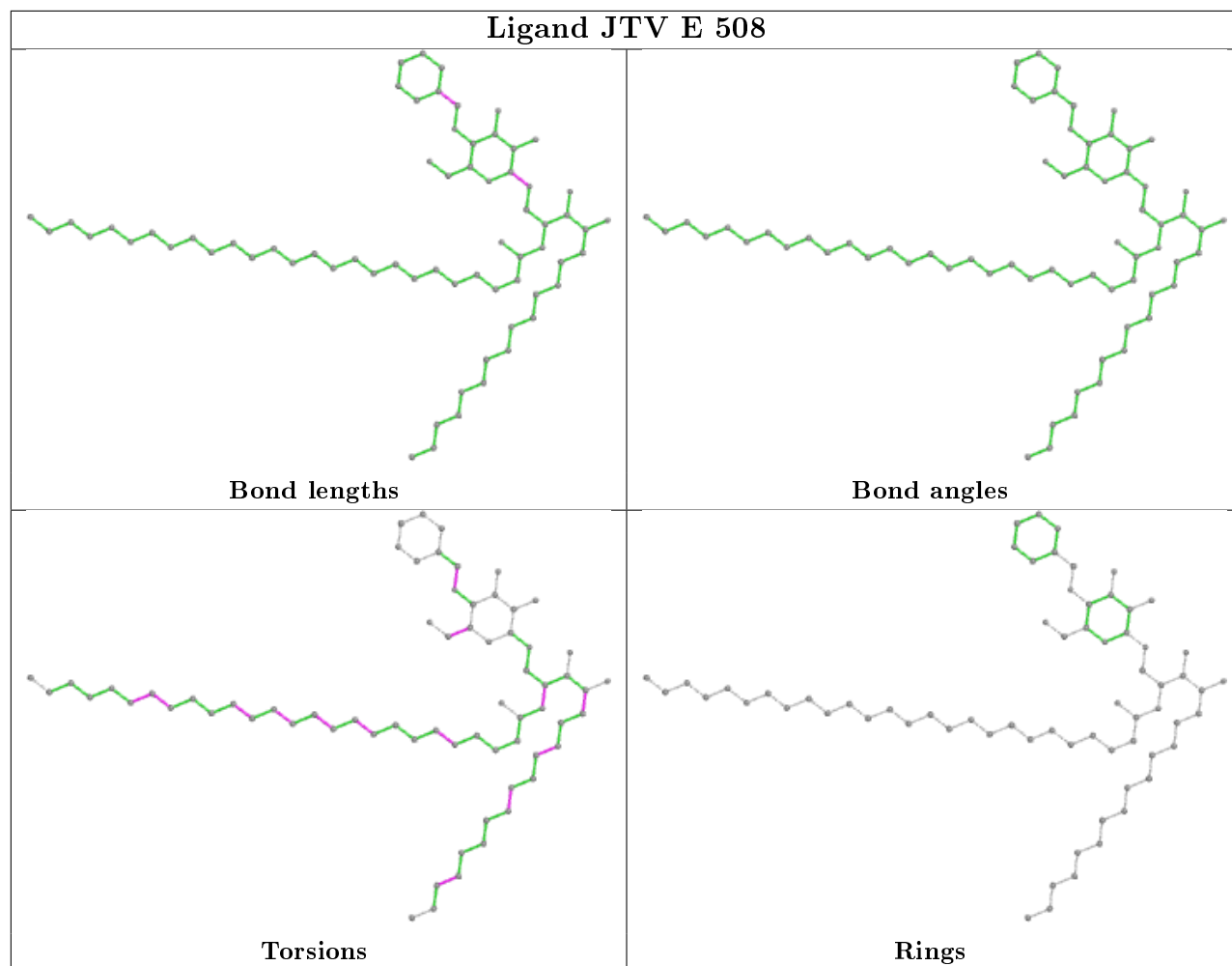
*Continued from previous page...*

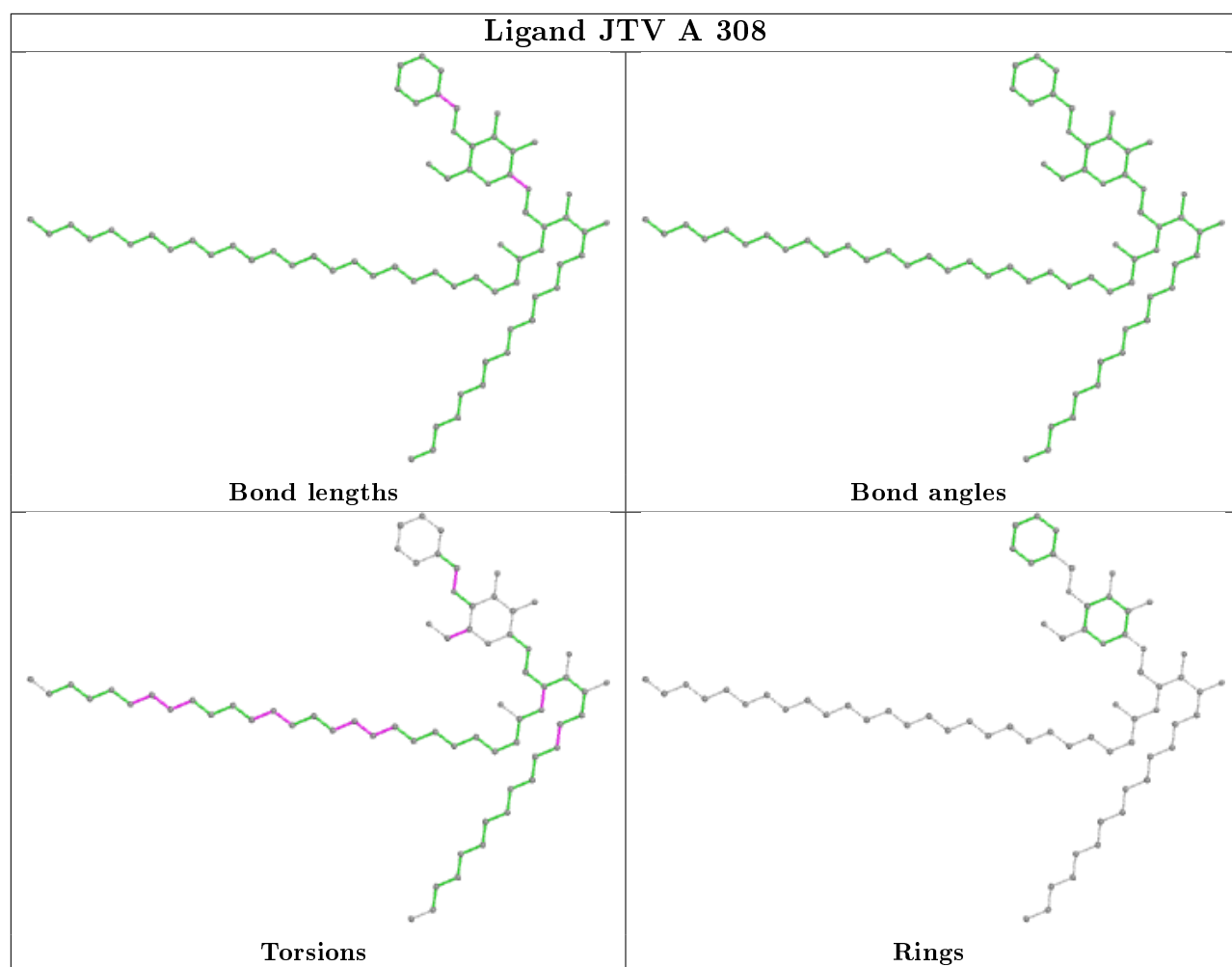
Mol	Chain	Res	Type	Atoms
9	A	308	JTV	CCE-CCF-CCG-CCH
9	E	508	JTV	CBP-CBQ-CBR-CBS
9	E	508	JTV	CBK-CBL-CBM-CBN
9	A	308	JTV	CCF-CCG-CCH-CCI
9	E	508	JTV	CCA-CCB-CCC-CCD
9	E	508	JTV	CBH-CBI-CBJ-CBK
9	A	308	JTV	CAU-CAT-O4-C4
9	E	508	JTV	CCE-CCF-CCG-CCH
9	E	508	JTV	CBW-CBX-CBY-CBZ
9	A	308	JTV	CBZ-CCA-CCB-CCC
9	E	508	JTV	CAU-CAT-O4-C4
9	E	508	JTV	CCL-CCM-CCN-CCO
9	A	308	JTV	CCJ-CCK-CCL-CCM
9	E	508	JTV	CAP-CAN-NAO-CBE
9	A	308	JTV	CAP-CAN-NAO-CBE
10	A	310	GOL	O2-C2-C3-O3
9	E	508	JTV	CCC-CCD-CCE-CCF
9	A	308	JTV	CBG-CBH-CBI-CBJ
9	A	308	JTV	CCL-CCM-CCN-CCO

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 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	C	200/209 (95%)	0.10	2 (1%) 82 87	39, 56, 91, 102	0
1	G	200/209 (95%)	0.09	2 (1%) 82 87	40, 56, 92, 101	0
2	D	240/241 (99%)	-0.01	0 100 100	40, 54, 78, 92	0
2	H	239/241 (99%)	-0.02	1 (0%) 92 95	40, 55, 74, 96	0
3	A	271/285 (95%)	0.13	1 (0%) 92 95	39, 57, 103, 131	0
3	E	274/285 (96%)	0.15	4 (1%) 73 81	41, 57, 108, 140	0
4	B	98/99 (98%)	-0.02	1 (1%) 82 87	46, 64, 85, 108	0
4	F	98/99 (98%)	-0.02	1 (1%) 82 87	44, 67, 88, 104	0
All	All	1620/1668 (97%)	0.06	12 (0%) 87 91	39, 57, 93, 140	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	B	99	MET	3.7
3	E	230	GLN	3.6
4	F	99	MET	3.3
3	E	231	GLY	3.2
1	G	195	ILE	3.0
1	C	185	PHE	2.9
3	A	231	GLY	2.5
3	E	203	HIS	2.4
1	G	169	MET	2.3
3	E	220	VAL	2.2
2	H	202	ARG	2.1
1	C	163	VAL	2.0

## 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 ⓘ

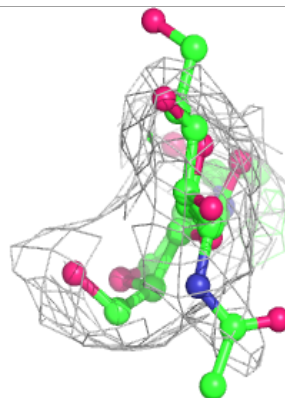
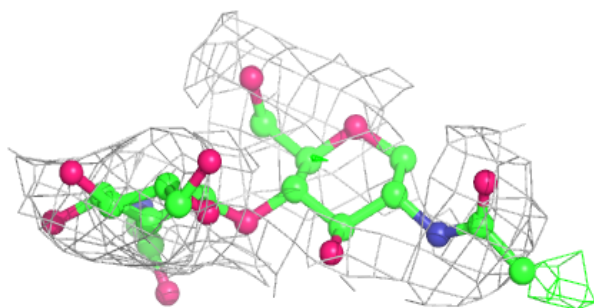
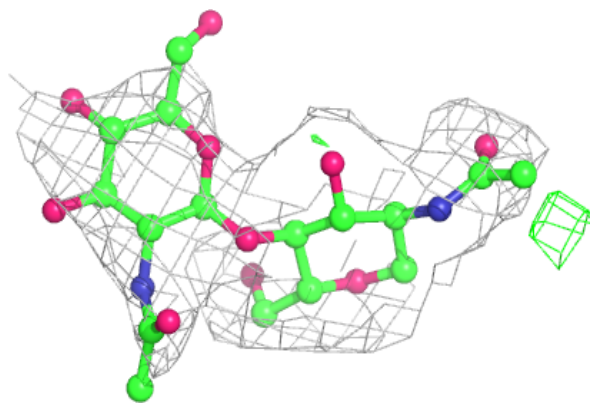
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	NAG	N	1	14/15	0.87	0.14	88,93,97,104	0
5	NAG	I	2	14/15	0.87	0.29	109,111,112,112	0
5	NAG	J	2	14/15	0.87	0.20	86,92,98,99	0
5	NAG	L	2	14/15	0.87	0.18	82,87,90,91	0
5	NAG	N	2	14/15	0.88	0.24	106,111,111,111	0
6	FUC	K	3	10/11	0.88	0.15	80,82,82,83	0
6	FUC	M	3	10/11	0.89	0.12	80,82,84,84	0
6	NAG	K	2	14/15	0.90	0.13	71,76,79,82	0
5	NAG	I	1	14/15	0.90	0.13	83,89,93,102	0
5	NAG	L	1	14/15	0.93	0.14	59,61,68,75	0
6	NAG	M	2	14/15	0.93	0.18	70,77,79,81	0
6	NAG	K	1	14/15	0.95	0.12	61,63,72,77	0
5	NAG	J	1	14/15	0.95	0.17	59,63,68,76	0
6	NAG	M	1	14/15	0.96	0.14	55,60,69,75	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 I:**

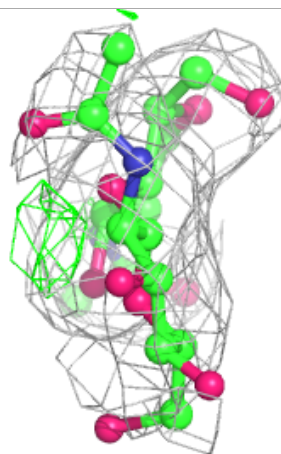
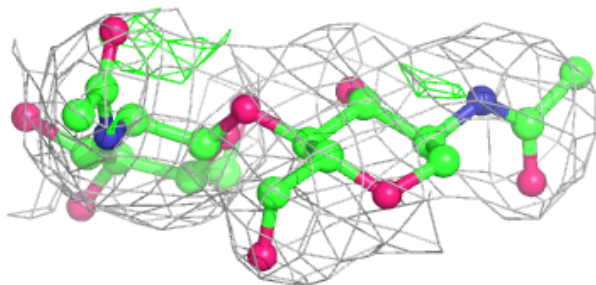
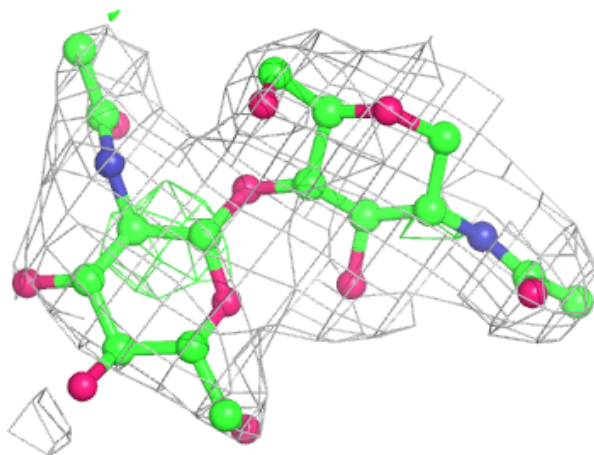
$2mF_o - DF_c$  (at 0.7 rmsd) in gray  
 $mF_o - DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





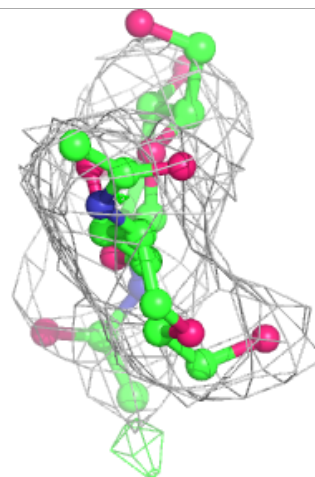
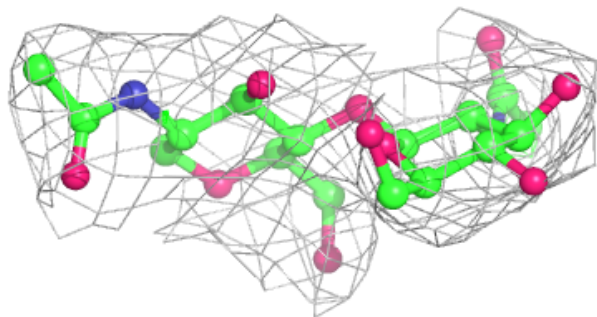
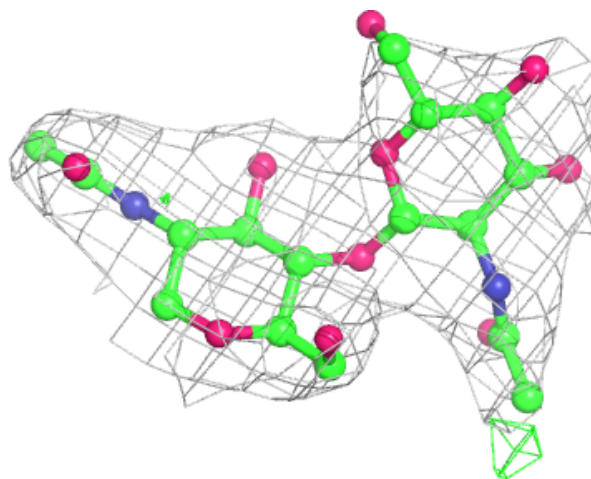
**Electron density around Chain J:**

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



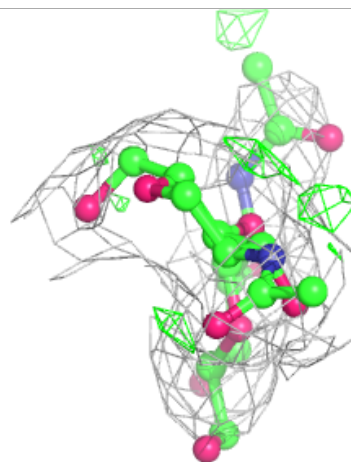
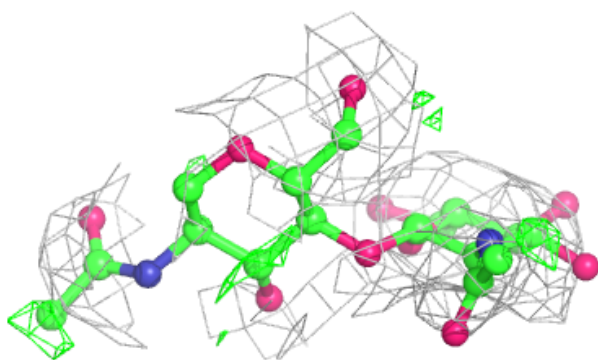
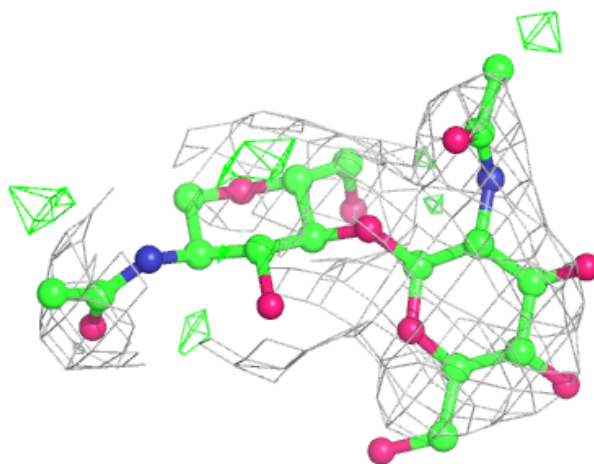
**Electron density around Chain L:**

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



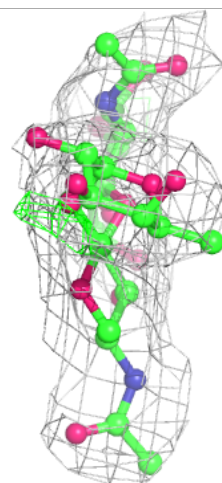
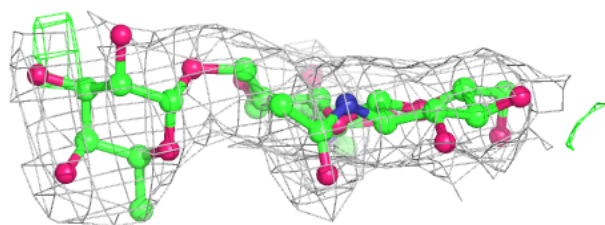
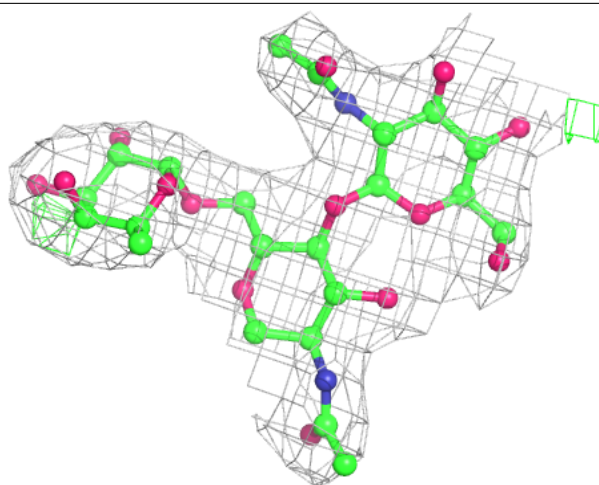
**Electron density around Chain N:**

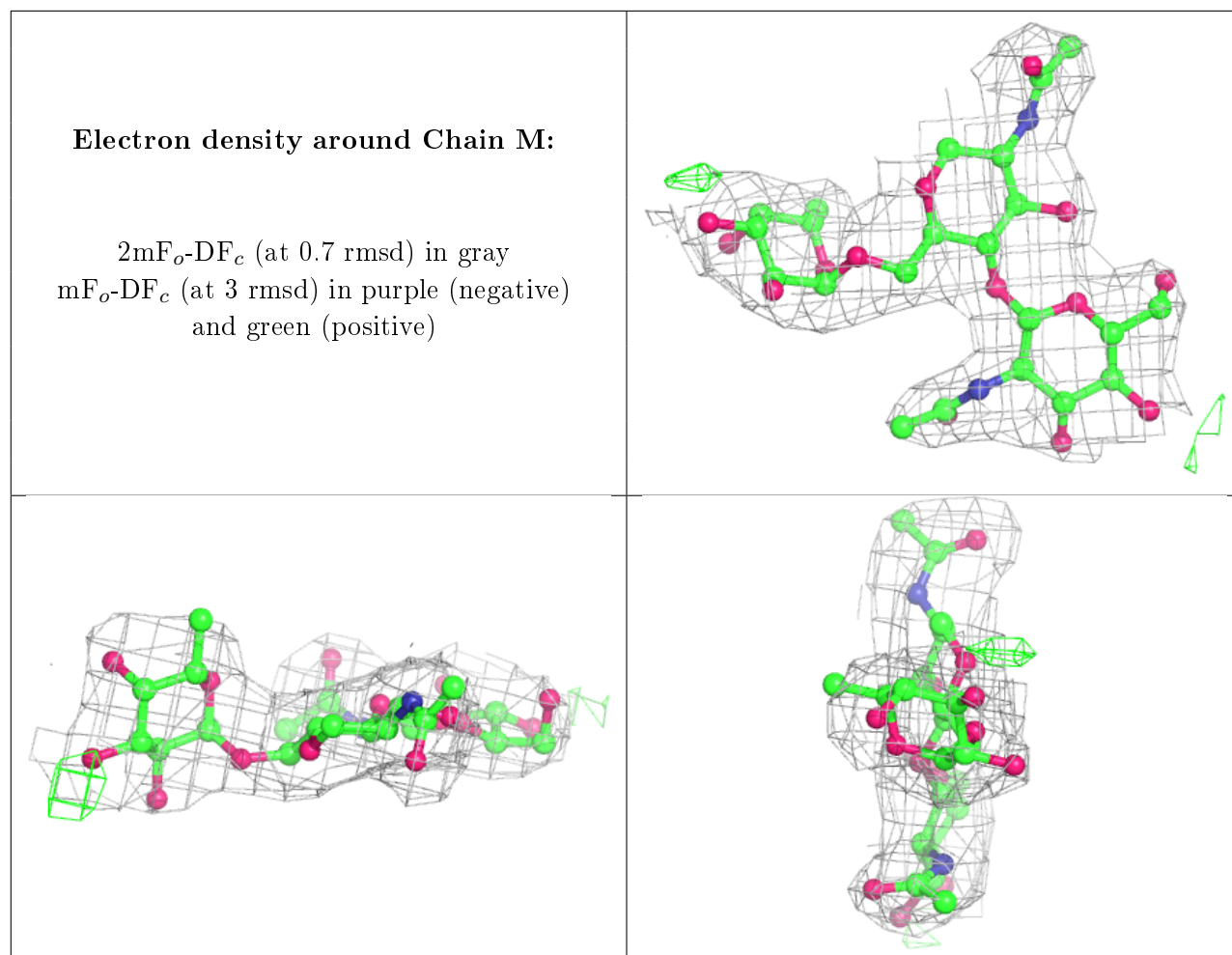
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around Chain K:**

$2mF_o-DF_c$  (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(Å <sup>2</sup> )	Q<0.9
10	GOL	A	310	6/6	0.71	0.17	100,102,104,106	0
10	GOL	A	309	6/6	0.71	0.16	86,87,89,89	0
10	GOL	E	509	6/6	0.76	0.20	94,96,97,98	0
7	NA	D	301	1/1	0.77	0.42	71,71,71,71	0
7	NA	C	301	1/1	0.82	0.75	75,75,75,75	0
10	GOL	G	301	6/6	0.87	0.26	76,78,79,81	0
8	CL	D	302	1/1	0.94	0.24	75,75,75,75	0
9	JTV	E	508	67/67	0.94	0.28	45,55,61,62	0
7	NA	H	301	1/1	0.94	0.35	54,54,54,54	0
7	NA	G	302	1/1	0.95	0.29	63,63,63,63	0

*Continued on next page...*

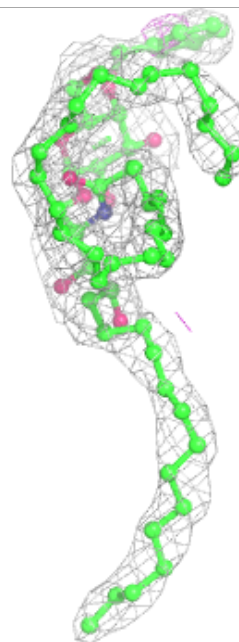
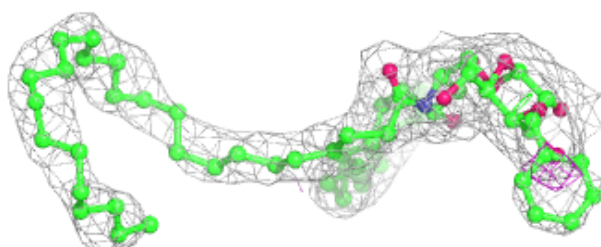
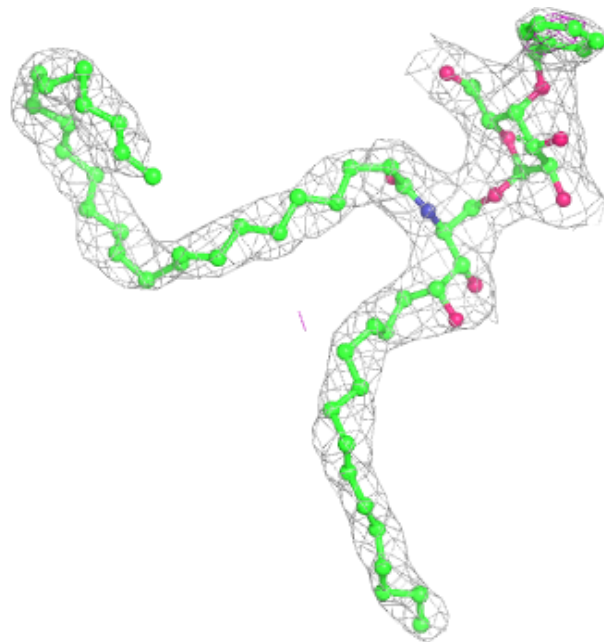
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
9	JTV	A	308	67/67	0.96	0.28	45,51,59,63	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.

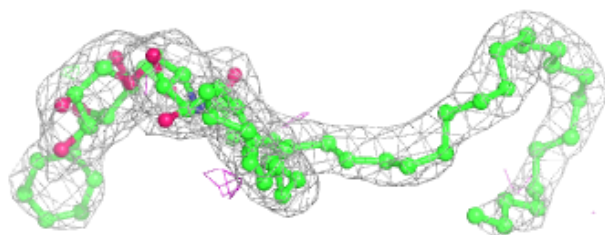
**Electron density around JTV E 508:**

2mF<sub>o</sub>-DF<sub>c</sub> (at 0.7 rmsd) in gray  
mF<sub>o</sub>-DF<sub>c</sub> (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around JTV A 308:**

2mF<sub>o</sub>-DF<sub>c</sub> (at 0.7 rmsd) in gray  
mF<sub>o</sub>-DF<sub>c</sub> (at 3 rmsd) in purple (negative)  
and green (positive)



## 6.5 Other polymers [i](#)

There are no such residues in this entry.