



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 22, 2024 – 10:07 AM EDT

PDB ID : 8TRR  
Title : T cell recognition of citrullinated vimentin peptide presented by HLA-DR4  
Authors : Loh, T.J.; Lim, J.J.; Reid, H.H.; Rossjohn, J.  
Deposited on : 2023-08-10  
Resolution : 2.65 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.37.1  
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) : 2.37.1

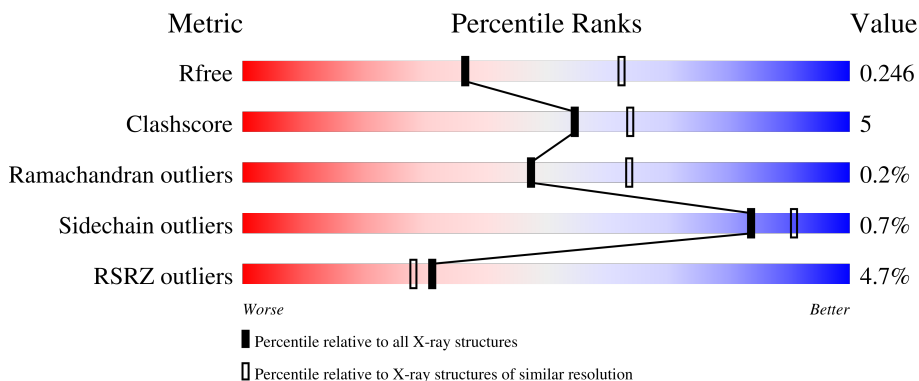
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.65 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.





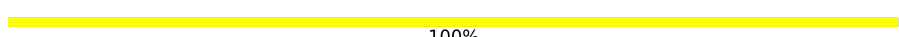
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1332 (2.68-2.64)
Clashscore	141614	1374 (2.68-2.64)
Ramachandran outliers	138981	1349 (2.68-2.64)
Sidechain outliers	138945	1349 (2.68-2.64)
RSRZ outliers	127900	1318 (2.68-2.64)

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  $\geq 3$ , 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	181	
1	F	181	
2	B	190	
2	G	190	
3	C	13	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	H	13	
4	D	209	
4	I	209	
5	E	243	
5	J	243	
6	K	2	
6	L	2	
7	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
8	GOL	F	203	-	-	-	X

## 2 Entry composition [i](#)

There are 11 unique types of molecules in this entry. The entry contains 13047 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 HLA class II histocompatibility antigen, DR alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	179	1473	954	239	275	5	0	0	0
1	F	179	1459	946	235	273	5	0	0	0

- Molecule 2 is a protein called HLA class II histocompatibility antigen, DRB1 beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	181	1488	943	257	283	5	0	0	0
2	G	182	1492	947	258	282	5	0	0	0

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	9	GLU	TRP	variant	UNP P01911
B	11	VAL	PRO	variant	UNP P01911
B	13	HIS	ARG	variant	UNP P01911
B	33	HIS	ASN	variant	UNP P01911
B	37	TYR	SER	variant	UNP P01911
B	47	TYR	PHE	variant	UNP P01911
B	67	LEU	ILE	variant	UNP P01911
B	71	LYS	ALA	variant	UNP P01911
B	86	GLY	VAL	variant	UNP P01911
B	96	TYR	GLN	variant	UNP P01911
B	98	GLU	LYS	variant	UNP P01911
B	104	ALA	SER	variant	UNP P01911
B	120	ASN	SER	variant	UNP P01911
B	133	ARG	LEU	variant	UNP P01911
B	140	THR	ALA	variant	UNP P01911
B	142	VAL	MET	variant	UNP P01911

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	180	LEU	VAL	variant	UNP P01911
G	9	GLU	TRP	variant	UNP P01911
G	11	VAL	PRO	variant	UNP P01911
G	13	HIS	ARG	variant	UNP P01911
G	33	HIS	ASN	variant	UNP P01911
G	37	TYR	SER	variant	UNP P01911
G	47	TYR	PHE	variant	UNP P01911
G	67	LEU	ILE	variant	UNP P01911
G	71	LYS	ALA	variant	UNP P01911
G	86	GLY	VAL	variant	UNP P01911
G	96	TYR	GLN	variant	UNP P01911
G	98	GLU	LYS	variant	UNP P01911
G	104	ALA	SER	variant	UNP P01911
G	120	ASN	SER	variant	UNP P01911
G	133	ARG	LEU	variant	UNP P01911
G	140	THR	ALA	variant	UNP P01911
G	142	VAL	MET	variant	UNP P01911
G	180	LEU	VAL	variant	UNP P01911

- Molecule 3 is a protein called Vimentin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	12	Total	C	N	O	0	0	0
			89	55	17	17			
3	H	13	Total	C	N	O	0	0	0
			94	58	18	18			

- Molecule 4 is a protein called A03 TCR alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	198	Total	C	N	O	S	0	0	0
			1483	923	243	309	8			
4	I	186	Total	C	N	O	S	0	0	0
			1382	862	226	287	7			

- Molecule 5 is a protein called A03 TCR beta chain.

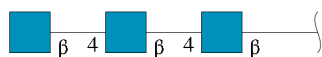
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	242	Total	C	N	O	S	0	0	0
			1878	1183	330	359	6			
5	J	241	Total	C	N	O	S	0	0	0
			1852	1170	319	357	6			

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



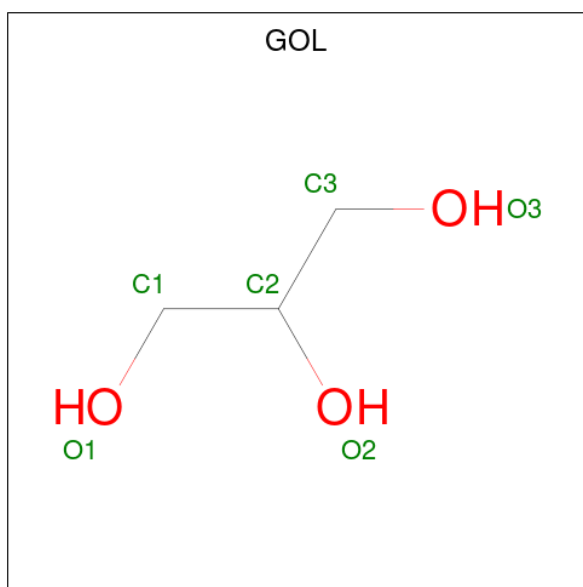
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
6	K	2	28	16	2	10	0	0	0
6	L	2	28	16	2	10	0	0	0

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



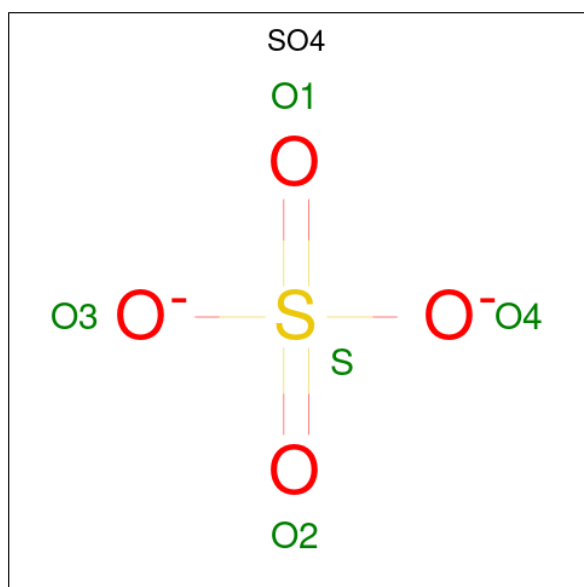
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
7	M	3	42	24	3	15	0	0	0

- Molecule 8 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
8	A	1	Total C O 6 3 3	0	0
8	A	1	Total C O 6 3 3	0	0
8	A	1	Total C O 6 3 3	0	0
8	B	1	Total C O 6 3 3	0	0
8	B	1	Total C O 6 3 3	0	0
8	D	1	Total C O 6 3 3	0	0
8	E	1	Total C O 6 3 3	0	0
8	F	1	Total C O 6 3 3	0	0
8	F	1	Total C O 6 3 3	0	0
8	G	1	Total C O 6 3 3	0	0

- Molecule 9 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	1	Total O S 5 4 1	0	0
9	A	1	Total O S 5 4 1	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	B	1	Total 5	O 4	S 1	0	0
9	D	1	Total 5	O 4	S 1	0	0
9	E	1	Total 5	O 4	S 1	0	0
9	E	1	Total 5	O 4	S 1	0	0
9	F	1	Total 5	O 4	S 1	0	0
9	G	1	Total 5	O 4	S 1	0	0
9	G	1	Total 5	O 4	S 1	0	0
9	G	1	Total 5	O 4	S 1	0	0
9	G	1	Total 5	O 4	S 1	0	0
9	G	1	Total 5	O 4	S 1	0	0
9	I	1	Total 5	O 4	S 1	0	0
9	I	1	Total 5	O 4	S 1	0	0
9	J	1	Total 5	O 4	S 1	0	0
9	J	1	Total 5	O 4	S 1	0	0

- Molecule 10 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).





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


- Molecule 11 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	17	Total	O	0	0
			17	17		
11	B	17	Total	O	0	0
			17	17		
11	C	2	Total	O	0	0
			2	2		
11	D	11	Total	O	0	0
			11	11		
11	E	7	Total	O	0	0
			7	7		
11	F	12	Total	O	0	0
			12	12		
11	G	6	Total	O	0	0
			6	6		
11	I	12	Total	O	0	0
			12	12		
11	J	7	Total	O	0	0
			7	7		

### 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: HLA class II histocompatibility antigen, DR alpha chain

Chain A: 




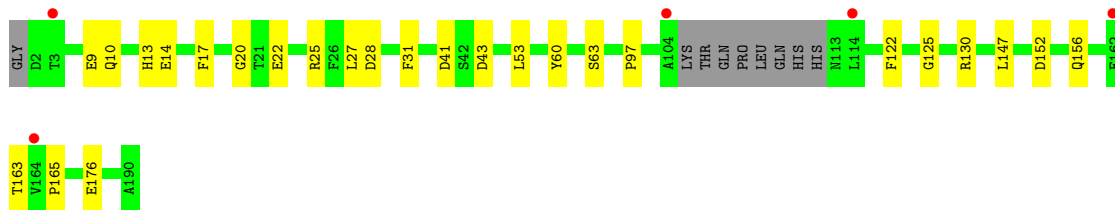
- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain

Chain F: 




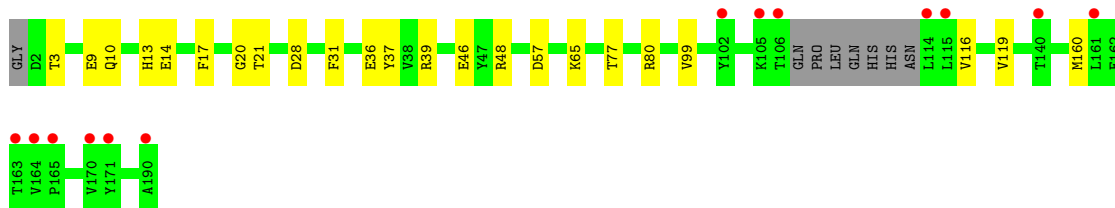
- Molecule 2: HLA class II histocompatibility antigen, DRB1 beta chain

Chain B: 



- Molecule 2: HLA class II histocompatibility antigen, DRB1 beta chain

Chain G: 



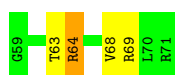
- Molecule 3: Vimentin

Chain C:  69% 23% 8%




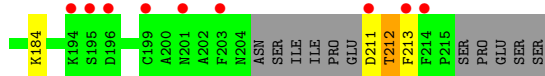
- Molecule 3: Vimentin

Chain H:  69% 23% 8%



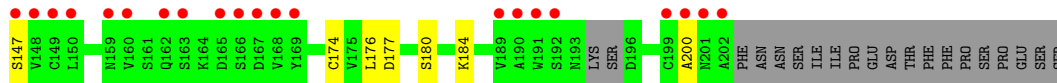
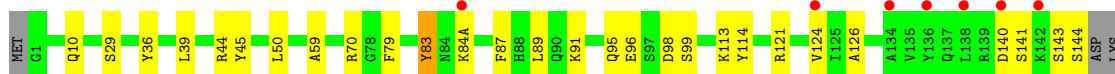
- Molecule 4: A03 TCR alpha chain

Chain D:  8% 81% 13% 5%




- Molecule 4: A03 TCR alpha chain

Chain I:  13% 72% 16% 11%




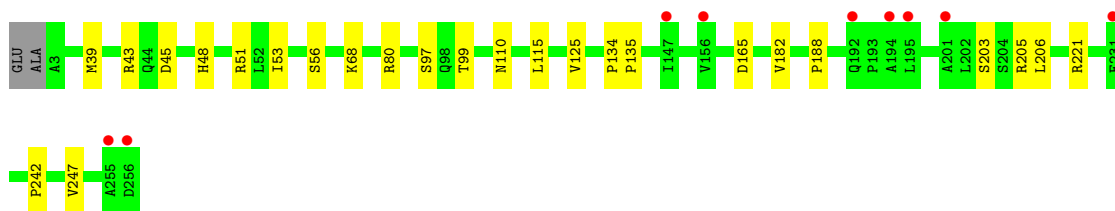
- Molecule 5: A03 TCR beta chain

Chain E:  87% 13%



- Molecule 5: A03 TCR beta chain

Chain J:  4% 89% 10%



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

Chain K: 100%

MAG1  
MAG2

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

Chain L: 100%

MAG1  
MAG2

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

Chain M: 33% 67%

MAG1  
MAG2  
MAG3

## 4 Data and refinement statistics i

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	98.93Å 59.97Å 379.16Å 90.00° 90.16° 90.00°	Depositor
Resolution (Å)	47.89 – 2.65 47.89 – 2.65	Depositor EDS
% Data completeness (in resolution range)	99.9 (47.89-2.65) 99.9 (47.89-2.65)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.14 (at 2.65Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487, PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.209 , 0.244 0.210 , 0.246	Depositor DCC
$R_{free}$ test set	3273 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	59.1	Xtrriage
Anisotropy	0.618	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 44.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.014 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	13047	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	79.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.95% 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: CIR, GOL, SO4, NAG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.49	0/1518	0.66	0/2070
1	F	0.48	0/1504	0.63	0/2054
2	B	0.52	0/1528	0.70	1/2078 (0.0%)
2	G	0.49	0/1532	0.65	0/2083
3	C	0.69	0/77	0.91	0/102
3	H	0.63	0/82	0.70	0/109
4	D	0.45	0/1510	0.64	0/2053
4	I	0.42	0/1406	0.62	0/1913
5	E	0.44	0/1929	0.60	0/2635
5	J	0.41	0/1903	0.58	1/2603 (0.0%)
All	All	0.47	0/12989	0.64	2/17700 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
3	C	0	1
3	H	0	2
All	All	0	3

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	152	ASP	CB-CG-OD1	5.64	123.38	118.30
5	J	68	LYS	C-N-CA	-5.34	111.08	122.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	63	THR	Mainchain
3	H	63	THR	Mainchain
3	H	64	CIR	Mainchain

## 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	1473	0	1405	12	0
1	F	1459	0	1379	15	0
2	B	1488	0	1386	16	0
2	G	1492	0	1397	16	0
3	C	89	0	87	2	0
3	H	94	0	89	3	0
4	D	1483	0	1372	16	0
4	I	1382	0	1273	25	0
5	E	1878	0	1758	19	0
5	J	1852	0	1716	20	0
6	K	28	0	25	0	0
6	L	28	0	25	0	0
7	M	42	0	37	2	0
8	A	18	0	23	1	0
8	B	12	0	16	0	0
8	D	6	0	8	1	0
8	E	6	0	8	0	0
8	F	12	0	16	2	0
8	G	6	0	8	1	0
9	A	10	0	0	0	0
9	B	5	0	0	0	0
9	D	5	0	0	0	0
9	E	10	0	0	0	0
9	F	5	0	0	0	0
9	G	25	0	0	0	0
9	I	10	0	0	0	0
9	J	10	0	0	0	0
10	B	14	0	13	3	0
10	F	14	0	13	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
11	A	17	0	0	0	0
11	B	17	0	0	0	0
11	C	2	0	0	0	0
11	D	11	0	0	0	0
11	E	7	0	0	0	0
11	F	12	0	0	1	0
11	G	6	0	0	0	0
11	I	12	0	0	0	0
11	J	7	0	0	0	0
All	All	13047	0	12054	129	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.

The worst 5 of 129 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)
5:E:250:GLU:OE2	5:J:221:ARG:NH1	1.99	0.96
1:F:46:GLU:OE2	1:F:50:ARG:NH1	2.11	0.83
2:G:10:GLN:HB2	2:G:31:PHE:HB2	1.70	0.72
4:I:29:SER:OG	4:I:113:LYS:HE2	1.92	0.70
4:I:79:PHE:HB3	4:I:89:LEU:HD11	1.76	0.68

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	177/181 (98%)	173 (98%)	4 (2%)	0	100 100
1	F	177/181 (98%)	173 (98%)	4 (2%)	0	100 100

Continued on next page...



Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	177/190 (93%)	171 (97%)	6 (3%)	0	100	100
2	G	178/190 (94%)	172 (97%)	6 (3%)	0	100	100
3	C	9/13 (69%)	9 (100%)	0	0	100	100
3	H	10/13 (77%)	9 (90%)	1 (10%)	0	100	100
4	D	194/209 (93%)	184 (95%)	9 (5%)	1 (0%)	29	43
4	I	180/209 (86%)	171 (95%)	7 (4%)	2 (1%)	14	21
5	E	240/243 (99%)	234 (98%)	6 (2%)	0	100	100
5	J	239/243 (98%)	236 (99%)	3 (1%)	0	100	100
All	All	1581/1672 (95%)	1532 (97%)	46 (3%)	3 (0%)	47	64

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	163	SER
4	I	200	ALA
4	I	143	SER

### 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	164/166 (99%)	164 (100%)	0	100	100
1	F	161/166 (97%)	158 (98%)	3 (2%)	57	74
2	B	161/171 (94%)	161 (100%)	0	100	100
2	G	161/171 (94%)	161 (100%)	0	100	100
3	C	8/9 (89%)	8 (100%)	0	100	100
3	H	8/9 (89%)	8 (100%)	0	100	100
4	D	161/184 (88%)	157 (98%)	4 (2%)	47	66
4	I	148/184 (80%)	146 (99%)	2 (1%)	67	81
5	E	200/209 (96%)	200 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	J	196/209 (94%)	195 (100%)	1 (0%)	88	94
All	All	1368/1478 (93%)	1358 (99%)	10 (1%)	84	91

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	I	83	TYR
4	I	141	SER
5	J	48	HIS
4	D	212	THR
1	F	46	GLU

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

Mol	Chain	Res	Type
5	E	54	HIS
5	E	84	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues 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	CIR	H	64	3	9,10,11	3.37	3 (33%)	6,11,13	1.50	1 (16%)
3	CIR	C	64	3	9,10,11	3.24	3 (33%)	6,11,13	1.53	1 (16%)

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	CIR	H	64	3	-	1/8/9/11	-
3	CIR	C	64	3	-	2/8/9/11	-

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	H	64	CIR	C7-N6	8.15	1.45	1.34
3	C	64	CIR	C7-N6	7.70	1.44	1.34
3	H	64	CIR	C7-N8	4.84	1.44	1.33
3	C	64	CIR	C7-N8	4.82	1.44	1.33
3	C	64	CIR	O7-C7	-2.98	1.19	1.24

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	H	64	CIR	C5-N6-C7	-2.87	119.32	122.73
3	C	64	CIR	O7-C7-N8	-2.33	119.21	123.22

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	64	CIR	CA-C3-C4-C5
3	H	64	CIR	CA-C3-C4-C5
3	C	64	CIR	C4-C5-N6-C7

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

7 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
6	NAG	K	1	6,1	14,14,15	0.52	0	17,19,21	1.25	2 (11%)
6	NAG	K	2	6	14,14,15	1.18	1 (7%)	17,19,21	1.46	2 (11%)
6	NAG	L	1	6,1	14,14,15	0.25	0	17,19,21	0.60	0
6	NAG	L	2	6	14,14,15	0.41	0	17,19,21	0.78	0
7	NAG	M	1	7,1	14,14,15	0.44	0	17,19,21	1.45	2 (11%)
7	NAG	M	2	7	14,14,15	1.22	1 (7%)	17,19,21	1.25	2 (11%)
7	NAG	M	3	7	14,14,15	3.28	4 (28%)	17,19,21	3.06	5 (29%)

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
6	NAG	K	1	6,1	-	2/6/23/26	0/1/1/1
6	NAG	K	2	6	-	2/6/23/26	0/1/1/1
6	NAG	L	1	6,1	-	0/6/23/26	0/1/1/1
6	NAG	L	2	6	-	3/6/23/26	0/1/1/1
7	NAG	M	1	7,1	-	3/6/23/26	0/1/1/1
7	NAG	M	2	7	-	1/6/23/26	0/1/1/1
7	NAG	M	3	7	-	2/6/23/26	0/1/1/1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	M	3	NAG	O5-C1	10.48	1.60	1.43
6	K	2	NAG	O5-C1	4.02	1.50	1.43
7	M	3	NAG	C8-C7	3.98	1.58	1.50
7	M	3	NAG	C1-C2	3.93	1.58	1.52
7	M	2	NAG	O5-C1	-3.73	1.37	1.43

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	M	3	NAG	C1-O5-C5	11.37	127.60	112.19
6	K	2	NAG	C1-O5-C5	4.87	118.78	112.19
7	M	1	NAG	C2-N2-C7	4.05	128.66	122.90

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	K	1	NAG	O4-C4-C3	-3.76	101.66	110.35
6	K	1	NAG	C1-O5-C5	2.69	115.83	112.19

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

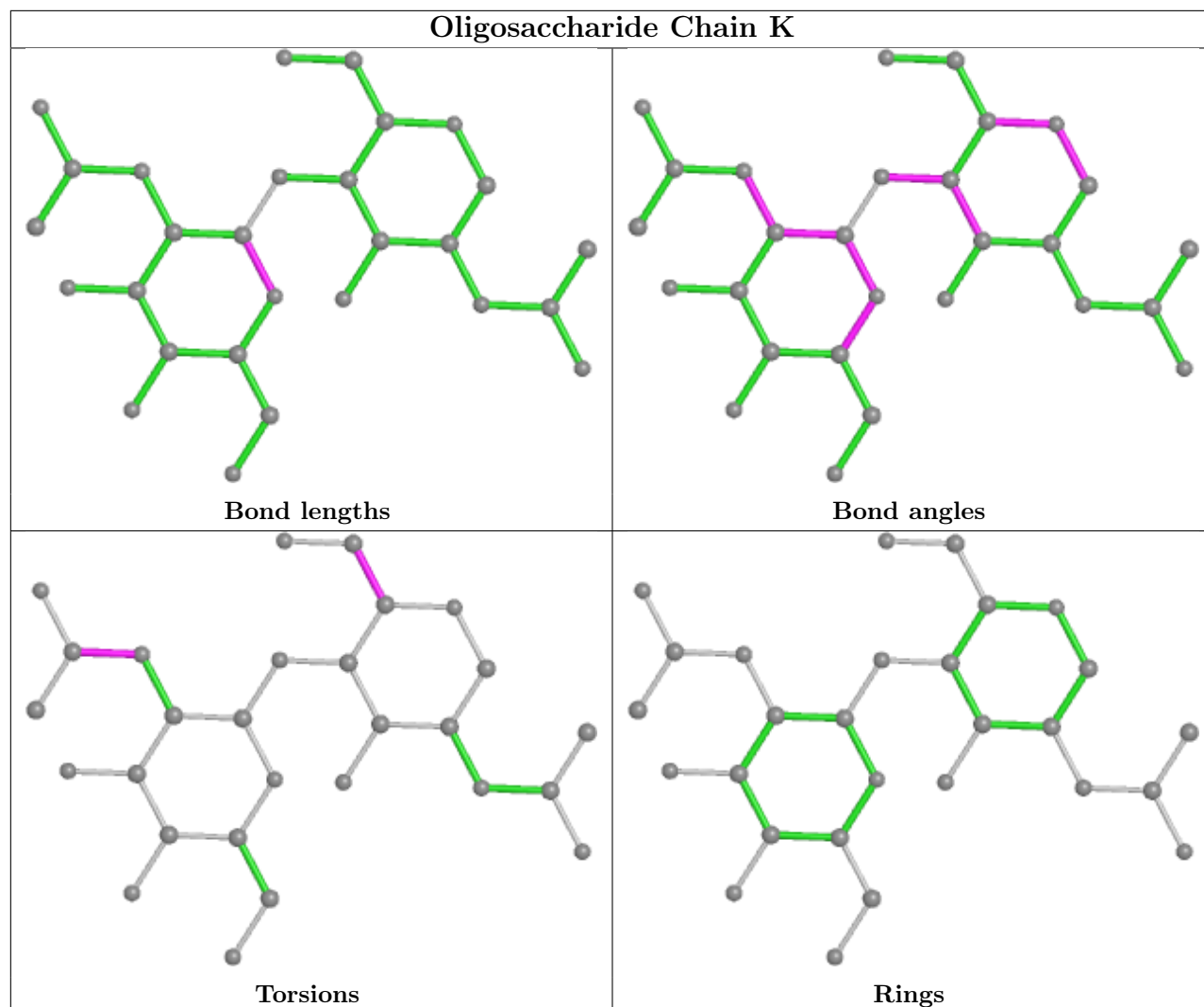
Mol	Chain	Res	Type	Atoms
6	K	1	NAG	O5-C5-C6-O6
7	M	2	NAG	C1-C2-N2-C7
6	K	1	NAG	C4-C5-C6-O6
6	K	2	NAG	C8-C7-N2-C2
6	K	2	NAG	O7-C7-N2-C2

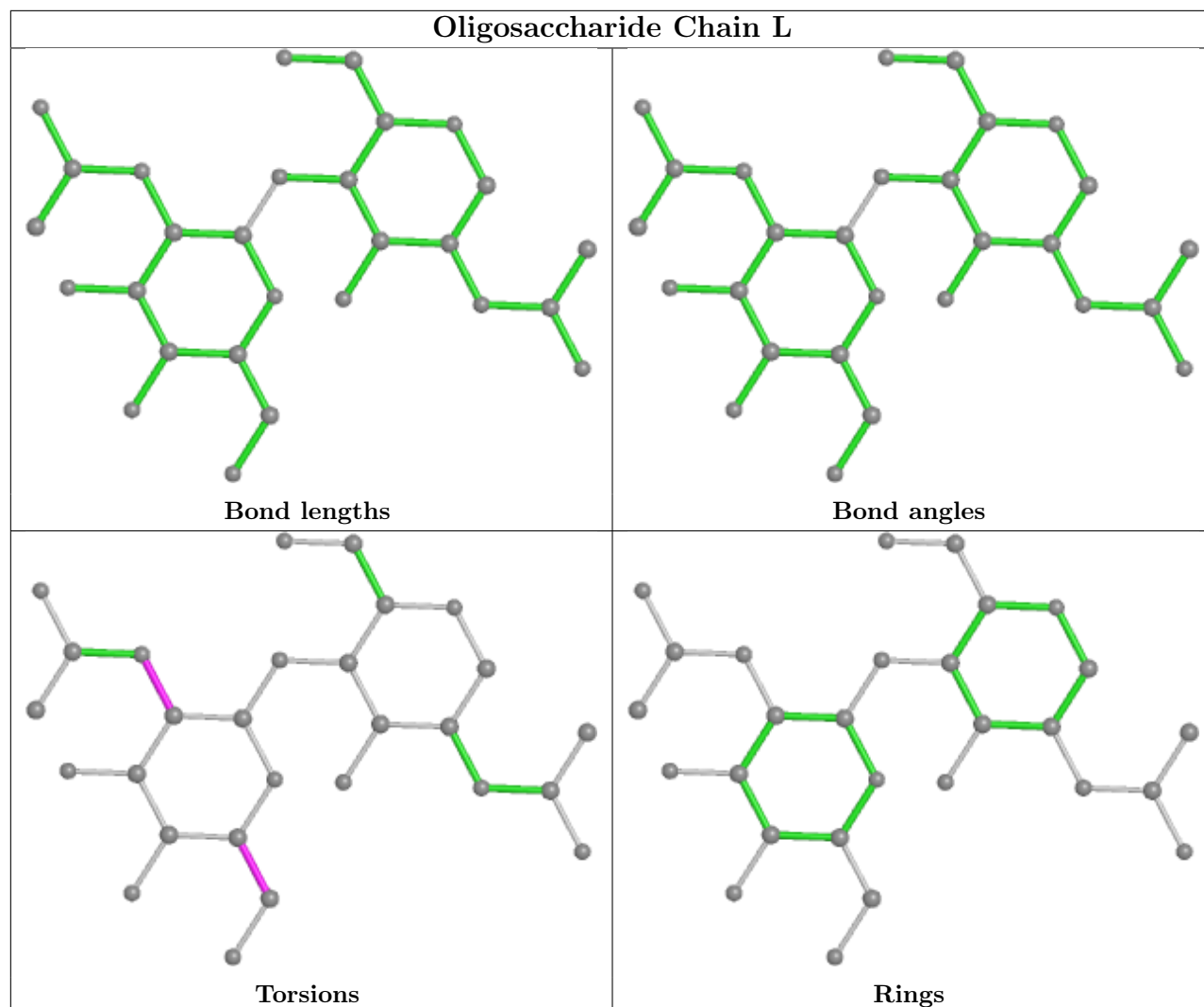
There are no ring outliers.

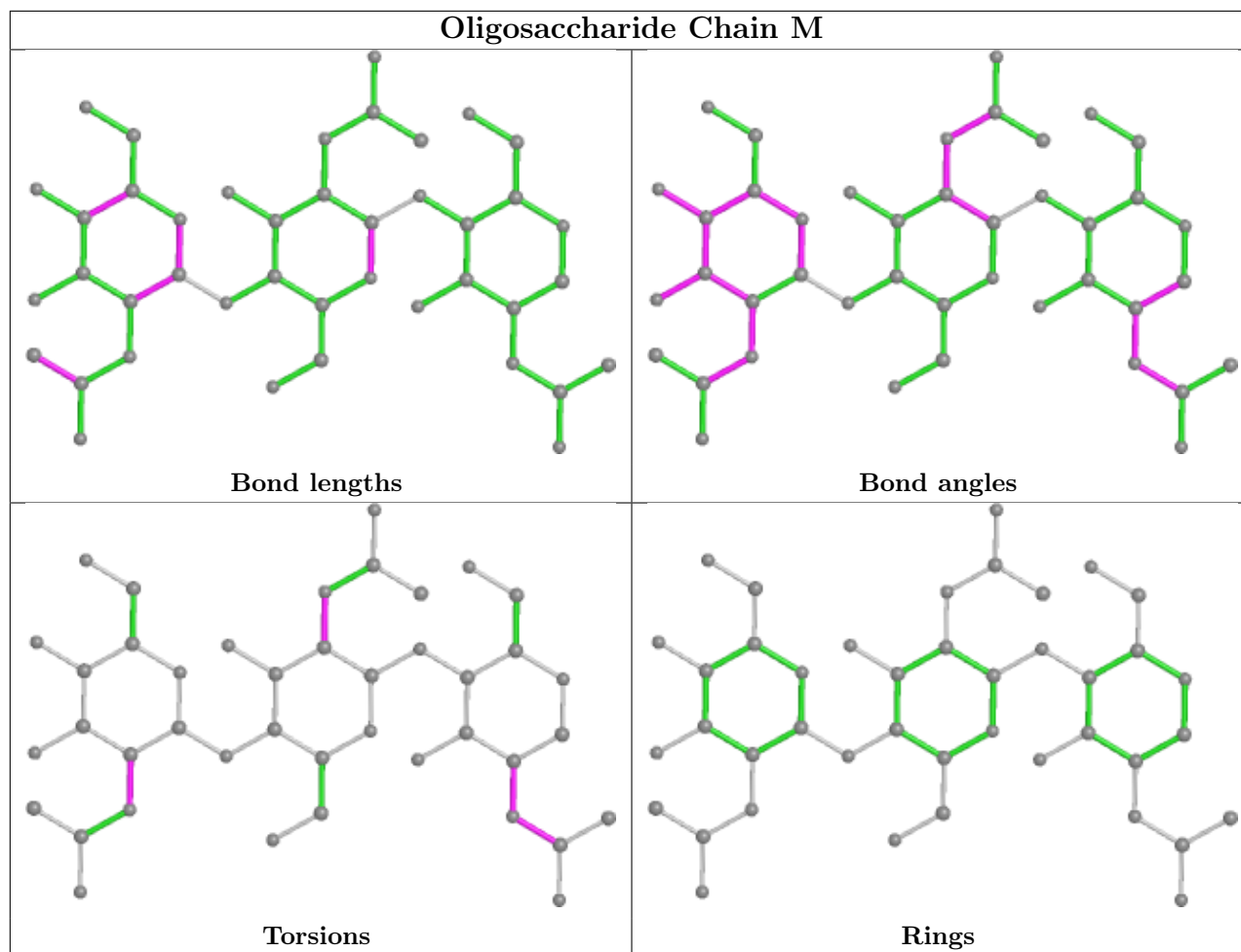
2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	M	1	NAG	1	0
7	M	3	NAG	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](#)

28 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
9	SO4	E	303	-	4,4,4	0.18	0	6,6,6	0.27	0
9	SO4	A	204	-	4,4,4	0.14	0	6,6,6	0.05	0
9	SO4	F	204	-	4,4,4	0.14	0	6,6,6	0.19	0
9	SO4	G	304	-	4,4,4	0.15	0	6,6,6	0.23	0
8	GOL	B	202	-	5,5,5	1.05	1 (20%)	5,5,5	0.84	0
9	SO4	E	302	-	4,4,4	0.14	0	6,6,6	0.37	0



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
9	SO4	G	305	-	4,4,4	0.13	0	6,6,6	0.11	0
8	GOL	A	202	-	5,5,5	1.15	1 (20%)	5,5,5	0.85	0
9	SO4	I	302	-	4,4,4	0.12	0	6,6,6	0.19	0
8	GOL	F	202	-	5,5,5	0.94	0	5,5,5	0.93	0
9	SO4	B	204	-	4,4,4	0.16	0	6,6,6	0.16	0
10	NAG	B	201	2	14,14,15	0.74	1 (7%)	17,19,21	1.51	2 (11%)
8	GOL	G	301	-	5,5,5	1.02	0	5,5,5	0.88	0
10	NAG	F	201	1	14,14,15	0.46	0	17,19,21	0.81	1 (5%)
9	SO4	I	301	-	4,4,4	0.14	0	6,6,6	0.15	0
9	SO4	J	301	-	4,4,4	0.16	0	6,6,6	0.07	0
8	GOL	E	301	-	5,5,5	1.14	1 (20%)	5,5,5	0.82	0
9	SO4	J	302	-	4,4,4	0.15	0	6,6,6	0.14	0
9	SO4	D	302	-	4,4,4	0.12	0	6,6,6	0.33	0
8	GOL	D	301	-	5,5,5	1.04	0	5,5,5	0.88	0
9	SO4	A	205	-	4,4,4	0.15	0	6,6,6	0.12	0
9	SO4	G	303	-	4,4,4	0.18	0	6,6,6	0.25	0
9	SO4	G	306	-	4,4,4	0.16	0	6,6,6	0.15	0
8	GOL	B	203	-	5,5,5	0.99	0	5,5,5	0.91	0
8	GOL	A	201	-	5,5,5	1.13	0	5,5,5	1.16	1 (20%)
9	SO4	G	302	-	4,4,4	0.11	0	6,6,6	0.19	0
8	GOL	A	203	-	5,5,5	0.87	0	5,5,5	1.14	0
8	GOL	F	203	-	5,5,5	1.05	0	5,5,5	0.90	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
8	GOL	D	301	-	-	0/4/4/4	-
8	GOL	B	202	-	-	0/4/4/4	-
8	GOL	B	203	-	-	2/4/4/4	-
8	GOL	A	202	-	-	2/4/4/4	-
8	GOL	F	202	-	-	2/4/4/4	-
8	GOL	E	301	-	-	0/4/4/4	-
10	NAG	F	201	1	-	2/6/23/26	0/1/1/1
8	GOL	A	201	-	-	1/4/4/4	-
10	NAG	B	201	2	-	2/6/23/26	0/1/1/1
8	GOL	G	301	-	-	2/4/4/4	-
8	GOL	A	203	-	-	3/4/4/4	-

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	GOL	F	203	-	-	0/4/4/4	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	B	201	NAG	O5-C1	2.43	1.47	1.43
8	A	202	GOL	O2-C2	-2.27	1.36	1.43
8	E	301	GOL	O2-C2	-2.14	1.37	1.43
8	B	202	GOL	O2-C2	-2.03	1.37	1.43

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
10	B	201	NAG	C1-O5-C5	5.39	119.49	112.19
10	F	201	NAG	C1-O5-C5	2.85	116.05	112.19
8	A	201	GOL	C3-C2-C1	-2.31	102.73	111.70
10	B	201	NAG	C3-C4-C5	2.17	114.11	110.24

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
8	A	203	GOL	O1-C1-C2-C3
8	A	203	GOL	C1-C2-C3-O3
8	B	203	GOL	O1-C1-C2-C3
10	F	201	NAG	O5-C5-C6-O6
10	B	201	NAG	O5-C5-C6-O6

There are no ring outliers.

5 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	A	202	GOL	1	0
10	B	201	NAG	3	0
8	G	301	GOL	1	0
8	D	301	GOL	1	0
8	F	203	GOL	2	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	179/181 (98%)	0.04	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	45, 61, 88, 118	0
1	F	179/181 (98%)	0.03	2 (1%) <span style="border: 1px solid blue; padding: 2px;">80</span> <span style="border: 1px solid blue; padding: 2px;">79</span>	49, 65, 90, 121	0
2	B	181/190 (95%)	0.15	5 (2%) <span style="border: 1px solid gray; padding: 2px;">53</span> <span style="border: 1px solid gray; padding: 2px;">49</span>	45, 65, 105, 127	0
2	G	182/190 (95%)	0.29	13 (7%) <span style="border: 1px solid red; padding: 2px;">16</span> <span style="border: 1px solid red; padding: 2px;">12</span>	48, 71, 134, 158	0
3	C	11/13 (84%)	0.07	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	47, 50, 63, 70	0
3	H	12/13 (92%)	0.20	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	49, 56, 76, 87	0
4	D	198/209 (94%)	0.38	17 (8%) <span style="border: 1px solid red; padding: 2px;">10</span> <span style="border: 1px solid red; padding: 2px;">8</span>	47, 79, 147, 178	0
4	I	186/209 (88%)	0.57	28 (15%) <span style="border: 1px solid red; padding: 2px;">2</span> <span style="border: 1px solid red; padding: 2px;">1</span>	53, 94, 140, 158	0
5	E	242/243 (99%)	0.14	2 (0%) <span style="border: 1px solid blue; padding: 2px;">86</span> <span style="border: 1px solid blue; padding: 2px;">85</span>	50, 80, 122, 140	0
5	J	241/243 (99%)	0.25	9 (3%) <span style="border: 1px solid red; padding: 2px;">41</span> <span style="border: 1px solid red; padding: 2px;">38</span>	57, 87, 125, 141	0
All	All	1611/1672 (96%)	0.23	76 (4%) <span style="border: 1px solid red; padding: 2px;">31</span> <span style="border: 1px solid red; padding: 2px;">28</span>	45, 75, 130, 178	0

The worst 5 of 76 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	194	LYS	4.3
4	D	195	SER	4.3
4	I	159	ASN	4.2
4	I	166	SER	4.2
4	D	145	ASP	4.0

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

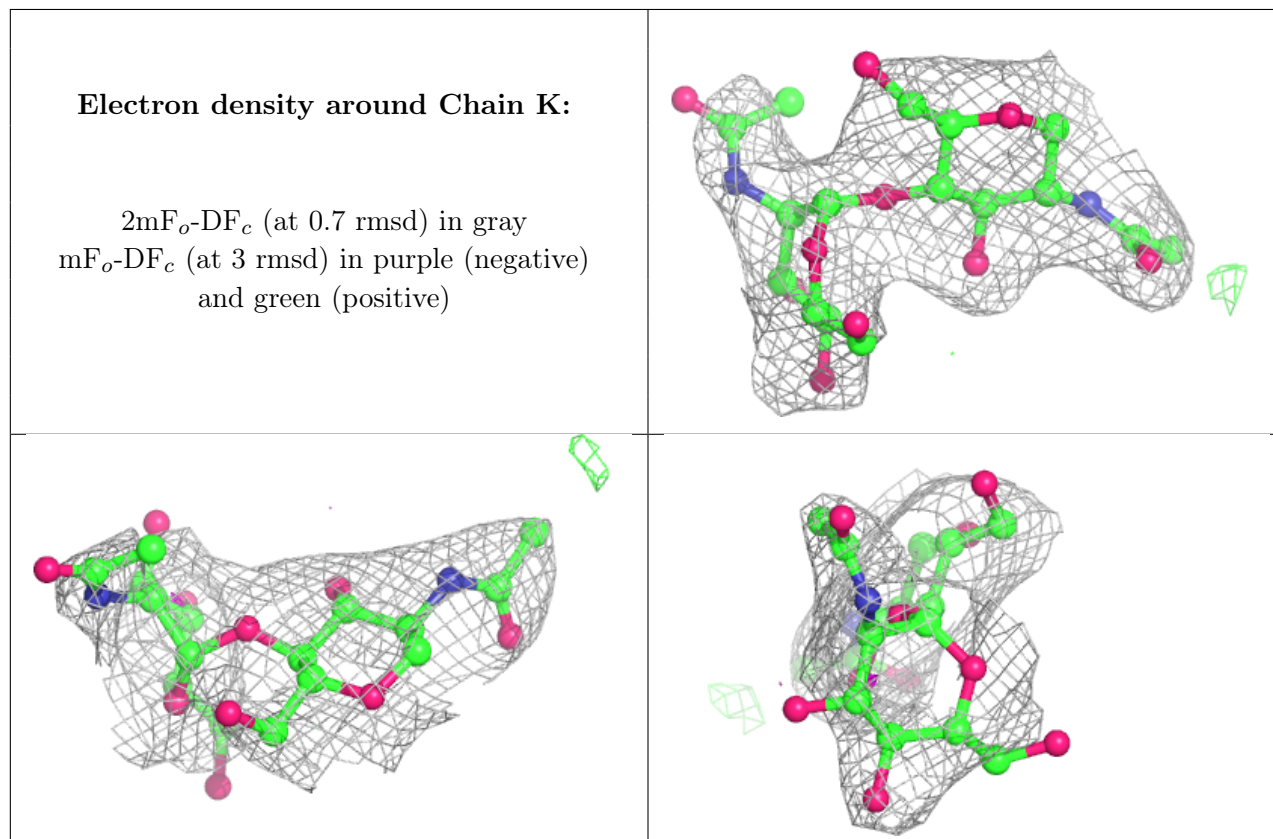
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	CIR	H	64	11/12	0.96	0.21	50,54,62,62	0
3	CIR	C	64	11/12	0.97	0.23	46,48,56,61	0

### 6.3 Carbohydrates [i](#)

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

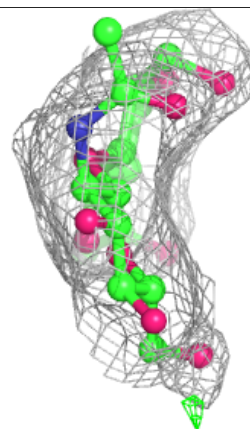
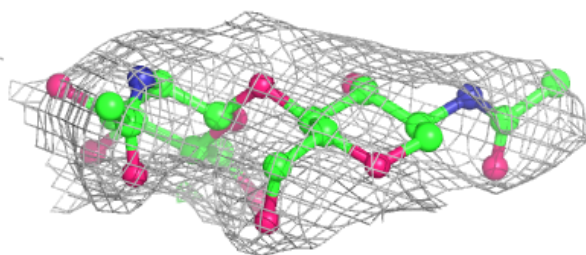
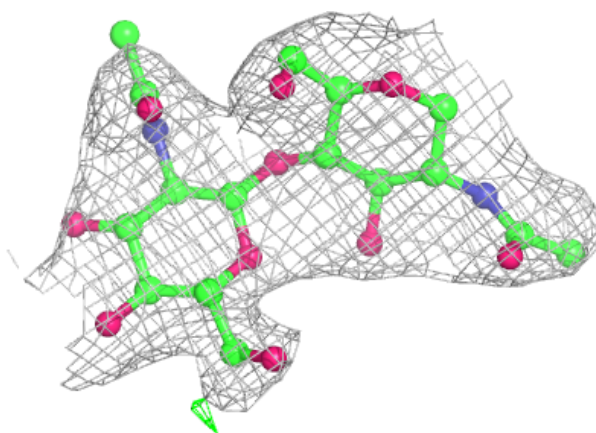
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
7	NAG	M	3	14/15	0.62	0.26	124,129,130,130	0
6	NAG	L	2	14/15	0.78	0.20	97,105,108,109	0
7	NAG	M	2	14/15	0.86	0.29	108,116,121,124	0
6	NAG	K	2	14/15	0.88	0.37	108,112,114,115	0
7	NAG	M	1	14/15	0.89	0.13	78,89,109,116	0
6	NAG	K	1	14/15	0.91	0.14	81,88,102,113	0
6	NAG	L	1	14/15	0.96	0.12	66,84,89,99	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.

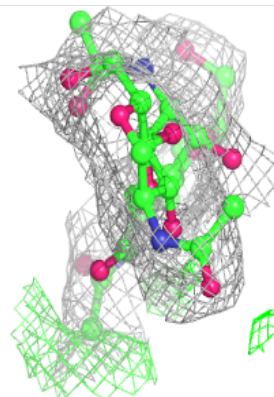
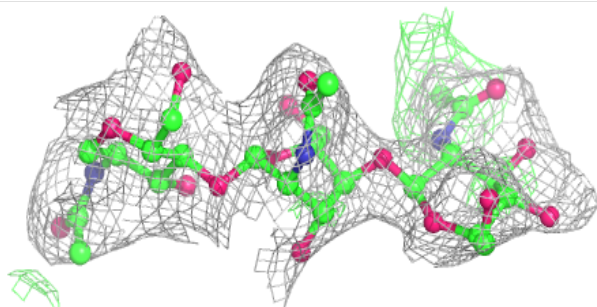
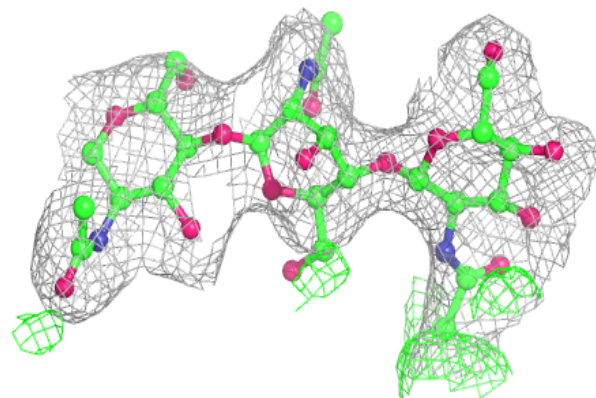


**Electron density around Chain L:**

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

**Electron density around Chain M:**

$2mF_o-DF_c$  (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
8	GOL	F	202	6/6	0.59	0.28	96,97,98,98	0
8	GOL	B	202	6/6	0.64	0.33	81,82,83,84	0
10	NAG	B	201	14/15	0.64	0.36	105,122,134,139	0
9	SO4	J	302	5/5	0.68	0.25	132,132,133,134	0
8	GOL	F	203	6/6	0.74	0.52	89,90,91,91	0
8	GOL	E	301	6/6	0.75	0.32	87,88,89,91	0
9	SO4	G	303	5/5	0.75	0.19	121,121,122,124	0
9	SO4	G	305	5/5	0.76	0.21	183,183,184,185	0
9	SO4	I	302	5/5	0.78	0.36	116,117,118,118	0
8	GOL	D	301	6/6	0.78	0.36	79,81,82,82	0
8	GOL	A	203	6/6	0.78	0.38	78,84,85,86	0
9	SO4	I	301	5/5	0.80	0.20	111,112,113,113	0
8	GOL	A	202	6/6	0.82	0.47	66,69,70,71	0
9	SO4	E	302	5/5	0.83	0.39	85,88,89,92	0
9	SO4	A	204	5/5	0.83	0.27	142,142,143,143	0
9	SO4	G	304	5/5	0.83	0.52	100,101,101,102	0
9	SO4	D	302	5/5	0.83	0.32	99,100,101,103	0
9	SO4	B	204	5/5	0.85	0.47	96,97,98,98	0
9	SO4	F	204	5/5	0.87	0.24	108,110,110,112	0
9	SO4	G	306	5/5	0.87	0.39	129,129,130,130	0
8	GOL	B	203	6/6	0.88	0.15	78,79,79,80	0
9	SO4	A	205	5/5	0.88	0.20	121,121,121,121	0
8	GOL	G	301	6/6	0.89	0.12	79,83,84,85	0
9	SO4	G	302	5/5	0.89	0.24	88,89,89,91	0
10	NAG	F	201	14/15	0.93	0.17	76,95,101,106	0
9	SO4	E	303	5/5	0.94	0.27	86,86,87,94	0
8	GOL	A	201	6/6	0.94	0.23	64,65,66,68	0
9	SO4	J	301	5/5	0.97	0.12	103,103,104,105	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.