



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 10, 2020 – 04:15 AM BST

PDB ID : 4WK2  
Title : Metal Ion and Ligand Binding of Integrin  
Authors : Xia, W.; Springer, T.A.  
Deposited on : 2014-10-01  
Resolution : 2.50 Å(reported)

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

We welcome your comments at [validation@mail.wwpdb.org](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  
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

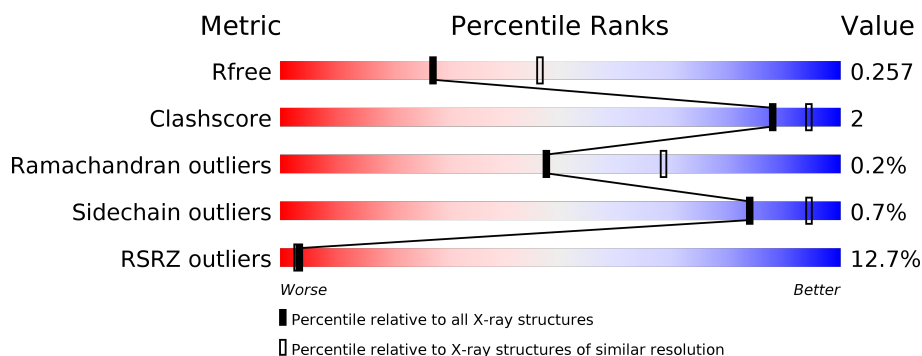
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.






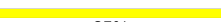
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments 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	452	<div> <div>8%</div> <div>96%</div> </div>
2	B	445	<div> <div>17%</div> <div>92%</div> </div>
3	C	6	<div> <div>17%</div> <div>100%</div> </div>
4	D	7	<div> <div>43%</div> <div>57%</div> </div>
5	E	2	<div> <div>100%</div> </div>
5	G	2	<div> <div>100%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	H	2	 100%
5	I	2	 100%
6	F	4	 75%  25%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	MAN	D	6	-	-	-	X
6	MAN	F	4	-	-	-	X
8	NAG	B	509	-	-	-	X

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 13981 atoms, of which 6767 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 Integrin alpha-5.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	448	Total	C	H	N	O	S	0	1	0
			6558	2149	3180	553	668	8			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	451	VAL	ILE	engineered mutation	UNP P08648

- Molecule 2 is a protein called Integrin beta-1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	427	Total	C	H	N	O	S	0	0	0
			6593	2082	3270	563	654	24			

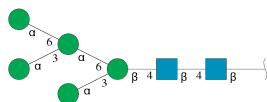
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	195	THR	SER	conflict	UNP P05556

- Molecule 3 is a protein called GLY-ARG-GLY-ASP-SER-PRO.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	6	Total	C	H	N	O		0	0	0
			75	22	34	9	10				

- Molecule 4 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]alpha-D-mannopyranose-(1-6)-[alpha-D-mannopyranose-(1-3)]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
4	D	7	Total	C	H	N	O	0	0	0
			157	46	74	2	35			

- 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	H	N	O	0	0	0
			55	16	27	2	10			
5	G	2	Total	C	H	N	O	0	0	0
			54	16	26	2	10			
5	H	2	Total	C	H	N	O	0	0	0
			55	16	27	2	10			
5	I	2	Total	C	H	N	O	0	0	0
			55	16	27	2	10			

- Molecule 6 is an oligosaccharide called alpha-D-mannopyranose-(1-6)-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	H	N	O	0	0	0
			96	28	46	2	20			

- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

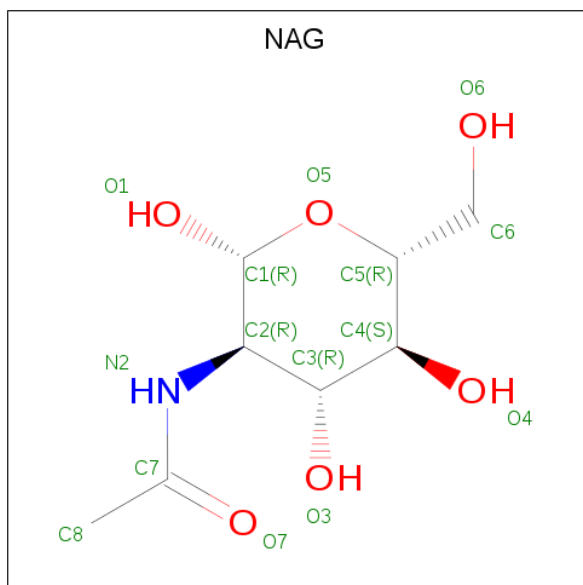
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	B	2	Total	Ca	0	0
			2	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	4	Total	Ca	0	0
			4	4		

- Molecule 8 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
8	A	1	Total	C	H	N	O	0	0
			28	8	14	1	5		
8	B	1	Total	C	H	N	O	0	0
			28	8	14	1	5		
8	B	1	Total	C	H	N	O	0	0
			28	8	14	1	5		
8	B	1	Total	C	H	N	O	0	0
			28	8	14	1	5		

- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

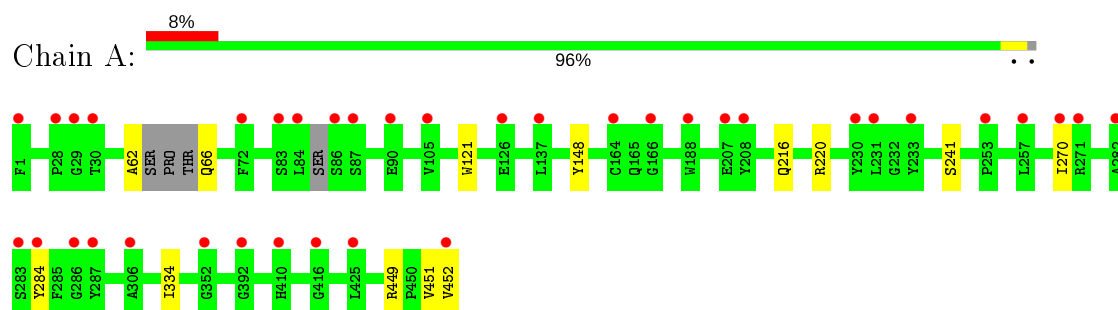
- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	84	Total 84	O 84	0	0
10	B	80	Total 80	O 80	0	0

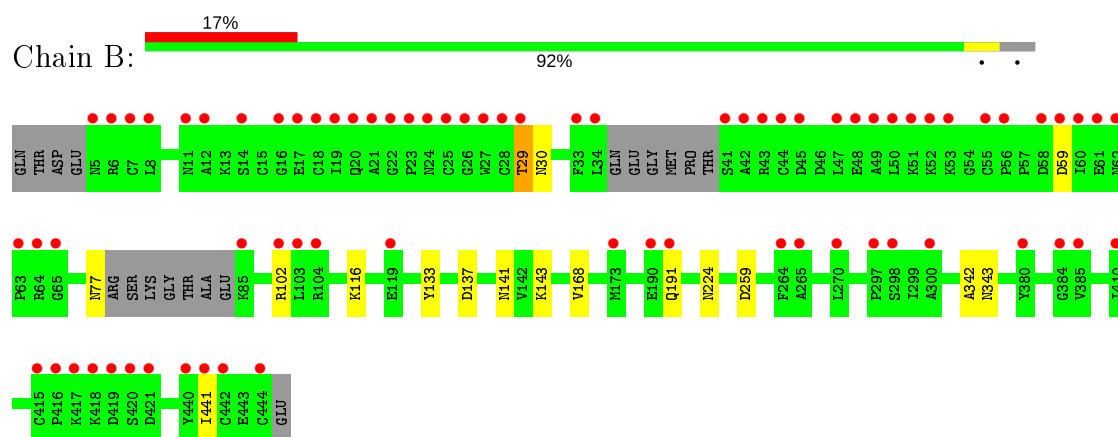
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

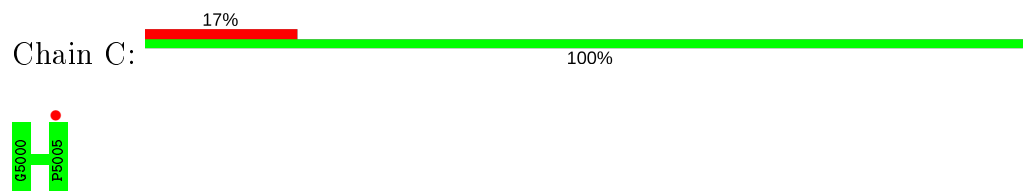
- Molecule 1: Integrin alpha-5



- Molecule 2: Integrin beta-1



- Molecule 3: GLY-ARG-GLY-ASP-SER-PRO



- Molecule 4: alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]alpha-D-mannopyranose-(1-6)-[alpha-D-mannopyranose-(1-3)]beta-D-mannopyranose-(1-4)-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 E:  100%



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

Chain G:  100%



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

Chain H:  100%



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

Chain I:  100%



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

Chain F:  75% 25%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	60.98Å 118.39Å 170.29Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.57 – 2.50 42.57 – 2.50	Depositor EDS
% Data completeness (in resolution range)	98.1 (42.57-2.50) 98.2 (42.57-2.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.32 (at 2.51Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.210 , 0.254 0.215 , 0.257	Depositor DCC
$R_{free}$ test set	1999 reflections (4.68%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	67.1	Xtriage
Anisotropy	0.389	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 51.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	13981	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	99.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.63% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, BMA, CA, 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	0.23	0/3473	0.39	0/4733
2	B	0.22	0/3379	0.39	0/4560
3	C	0.19	0/41	0.32	0/52
All	All	0.23	0/6893	0.39	0/9345

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

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	3378	3180	3177	8	0
2	B	3323	3270	3274	13	0
3	C	41	34	34	0	0
4	D	83	74	71	6	0
5	E	28	27	25	0	0
5	G	28	26	25	3	0
5	H	28	27	25	0	0
5	I	28	27	25	0	0
6	F	50	46	43	0	0
7	A	4	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	B	2	0	0	0	0
8	A	14	14	13	0	0
8	B	42	42	39	3	0
9	B	1	0	0	0	0
10	A	84	0	0	1	0
10	B	80	0	0	3	0
All	All	7214	6767	6751	26	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

All (26) 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:3:BMA:C6	4:D:4:MAN:C1	2.42	0.97
2:B:137:ASP:HB2	8:B:509:NAG:O7	1.75	0.86
1:A:241:SER:HB2	5:G:1:NAG:H82	1.60	0.83
2:B:137:ASP:CB	8:B:509:NAG:O7	2.33	0.76
2:B:259:ASP:OD2	10:B:614:HOH:O	2.08	0.72
4:D:3:BMA:O6	4:D:4:MAN:C2	2.37	0.72
1:A:449:ARG:NH2	10:A:668:HOH:O	2.30	0.63
2:B:30:ASN:N	2:B:59:ASP:OD1	2.33	0.62
4:D:3:BMA:H61	4:D:4:MAN:C1	2.31	0.58
2:B:29:THR:OG1	2:B:59:ASP:O	2.26	0.53
1:A:148:TYR:OH	1:A:216:GLN:NE2	2.42	0.53
2:B:133:TYR:OH	2:B:191:GLN:NE2	2.42	0.53
2:B:141:ASN:ND2	2:B:343:ASN:OD1	2.42	0.53
1:A:62:ALA:O	1:A:66:GLN:NE2	2.41	0.52
1:A:220:ARG:NH2	4:D:4:MAN:C1	2.73	0.52
1:A:241:SER:HB2	5:G:1:NAG:C8	2.38	0.48
1:A:451:VAL:HG22	1:A:452:VAL:H	1.79	0.47
2:B:29:THR:N	2:B:59:ASP:OD1	2.43	0.46
2:B:102:ARG:NH1	2:B:441:ILE:HD11	2.31	0.46
2:B:137:ASP:HB3	8:B:509:NAG:O7	2.16	0.45
2:B:77:ASN:HB3	2:B:116:LYS:NZ	2.33	0.44
2:B:143:LYS:NZ	10:B:633:HOH:O	2.51	0.43
2:B:342:ALA:O	10:B:680:HOH:O	2.21	0.43
1:A:220:ARG:NH2	4:D:6:MAN:O5	2.52	0.42
4:D:4:MAN:O5	4:D:6:MAN:C1	2.68	0.42
5:G:1:NAG:H61	5:G:2:NAG:N2	2.34	0.42

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	443/452 (98%)	427 (96%)	15 (3%)	1 (0%)	47	68
2	B	421/445 (95%)	398 (94%)	22 (5%)	1 (0%)	47	68
3	C	4/6 (67%)	4 (100%)	0	0	100	100
All	All	868/903 (96%)	829 (96%)	37 (4%)	2 (0%)	47	68

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	270	ILE
2	B	168	VAL

### 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	355/358 (99%)	352 (99%)	3 (1%)	81	93
2	B	380/395 (96%)	378 (100%)	2 (0%)	88	96
3	C	4/4 (100%)	4 (100%)	0	100	100
All	All	739/757 (98%)	734 (99%)	5 (1%)	84	94

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	121	TRP
1	A	284	TYR
1	A	334	ILE
2	B	29	THR
2	B	224	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	216	GLN
1	A	221	GLN
2	B	141	ASN
2	B	191	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 ⓘ

19 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	$\# Z  > 2$
4	NAG	D	1	1,4	14,14,15	0.37	0	17,19,21	0.38	0
4	NAG	D	2	4	14,14,15	0.24	0	17,19,21	0.46	0
4	BMA	D	3	4	11,11,12	0.68	0	15,15,17	0.78	0
4	MAN	D	4	4	11,11,12	0.26	0	15,15,17	0.58	0
4	MAN	D	5	4	11,11,12	0.25	0	15,15,17	0.61	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	MAN	D	6	4	11,11,12	0.27	0	15,15,17	0.64	0
4	MAN	D	7	4	11,11,12	0.76	0	15,15,17	1.02	2 (13%)
5	NAG	E	1	1,5	14,14,15	0.24	0	17,19,21	0.41	0
5	NAG	E	2	5	14,14,15	0.23	0	17,19,21	0.88	0
6	NAG	F	1	1,6	14,14,15	0.37	0	17,19,21	0.39	0
6	NAG	F	2	6	14,14,15	0.28	0	17,19,21	0.38	0
6	BMA	F	3	6	11,11,12	0.59	0	15,15,17	0.94	0
6	MAN	F	4	6	11,11,12	0.62	0	15,15,17	1.29	3 (20%)
5	NAG	G	1	1,5	14,14,15	0.30	0	17,19,21	0.61	0
5	NAG	G	2	5	14,14,15	0.31	0	17,19,21	0.64	0
5	NAG	H	1	2,5	14,14,15	0.30	0	17,19,21	0.59	0
5	NAG	H	2	5	14,14,15	0.28	0	17,19,21	0.58	0
5	NAG	I	1	2,5	14,14,15	0.28	0	17,19,21	0.36	0
5	NAG	I	2	5	14,14,15	0.20	0	17,19,21	0.40	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
4	NAG	D	1	1,4	-	2/6/23/26	0/1/1/1
4	NAG	D	2	4	-	0/6/23/26	0/1/1/1
4	BMA	D	3	4	-	2/2/19/22	0/1/1/1
4	MAN	D	4	4	-	0/2/19/22	0/1/1/1
4	MAN	D	5	4	-	0/2/19/22	0/1/1/1
4	MAN	D	6	4	-	0/2/19/22	0/1/1/1
4	MAN	D	7	4	-	2/2/19/22	0/1/1/1
5	NAG	E	1	1,5	-	2/6/23/26	0/1/1/1
5	NAG	E	2	5	-	0/6/23/26	0/1/1/1
6	NAG	F	1	1,6	-	0/6/23/26	0/1/1/1
6	NAG	F	2	6	-	0/6/23/26	0/1/1/1
6	BMA	F	3	6	-	2/2/19/22	0/1/1/1
6	MAN	F	4	6	-	1/2/19/22	0/1/1/1
5	NAG	G	1	1,5	-	0/6/23/26	0/1/1/1
5	NAG	G	2	5	-	1/6/23/26	0/1/1/1
5	NAG	H	1	2,5	-	0/6/23/26	0/1/1/1
5	NAG	H	2	5	-	1/6/23/26	0/1/1/1
5	NAG	I	1	2,5	-	2/6/23/26	0/1/1/1
5	NAG	I	2	5	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	F	4	MAN	C1-O5-C5	3.17	116.49	112.19
6	F	4	MAN	O5-C1-C2	2.22	114.20	110.77
6	F	4	MAN	O2-C2-C3	-2.19	105.75	110.14
4	D	7	MAN	O2-C2-C3	-2.17	105.78	110.14
4	D	7	MAN	C1-O5-C5	2.01	114.92	112.19

There are no chirality outliers.

All (15) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	F	3	BMA	O5-C5-C6-O6
4	D	3	BMA	C4-C5-C6-O6
5	I	1	NAG	C4-C5-C6-O6
5	I	1	NAG	O5-C5-C6-O6
4	D	1	NAG	O5-C5-C6-O6
6	F	3	BMA	C4-C5-C6-O6
5	E	1	NAG	O5-C5-C6-O6
4	D	1	NAG	C4-C5-C6-O6
5	E	1	NAG	C4-C5-C6-O6
6	F	4	MAN	O5-C5-C6-O6
4	D	3	BMA	O5-C5-C6-O6
4	D	7	MAN	O5-C5-C6-O6
4	D	7	MAN	C4-C5-C6-O6
5	H	2	NAG	O5-C5-C6-O6
5	G	2	NAG	O5-C5-C6-O6

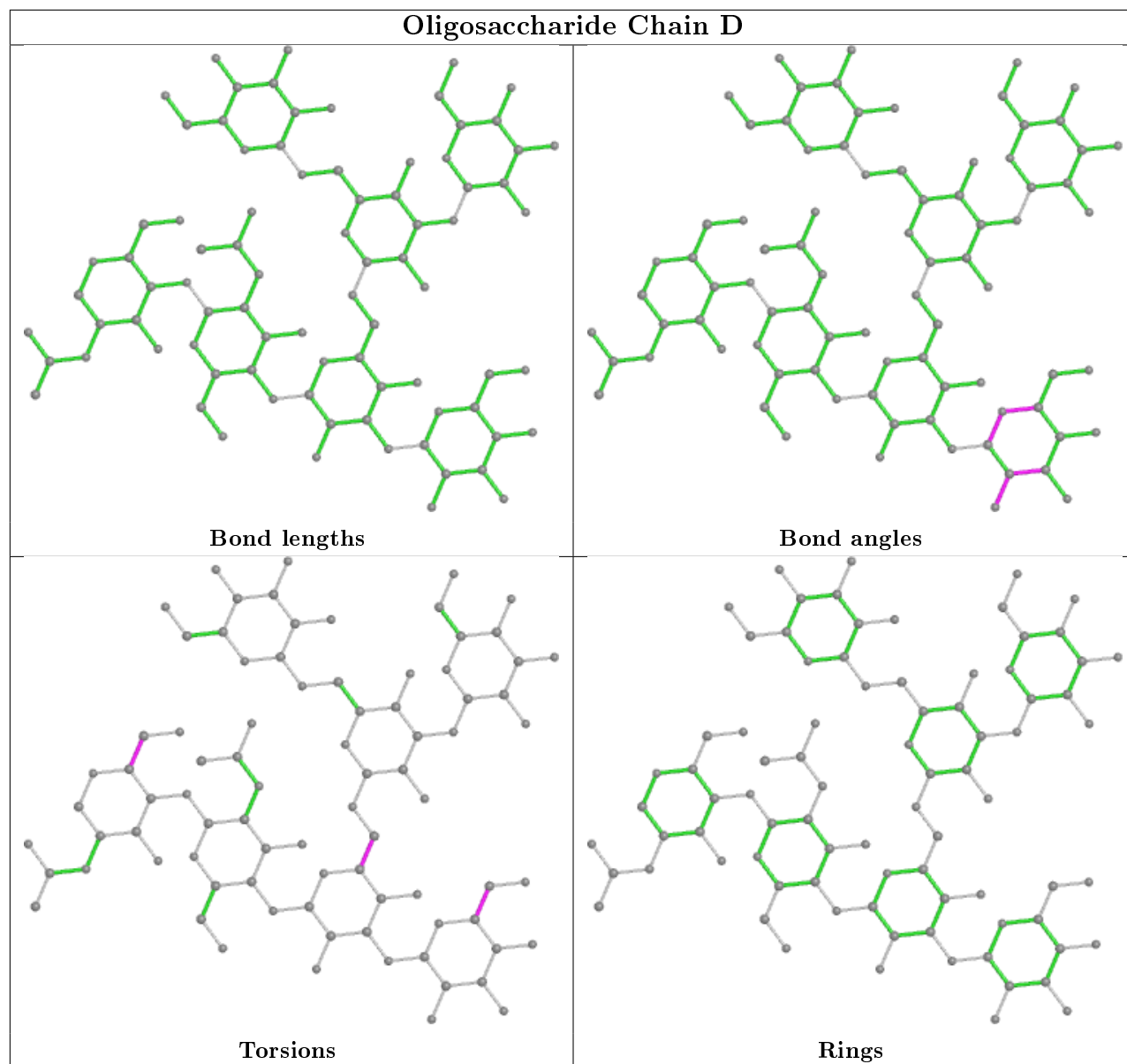
There are no ring outliers.

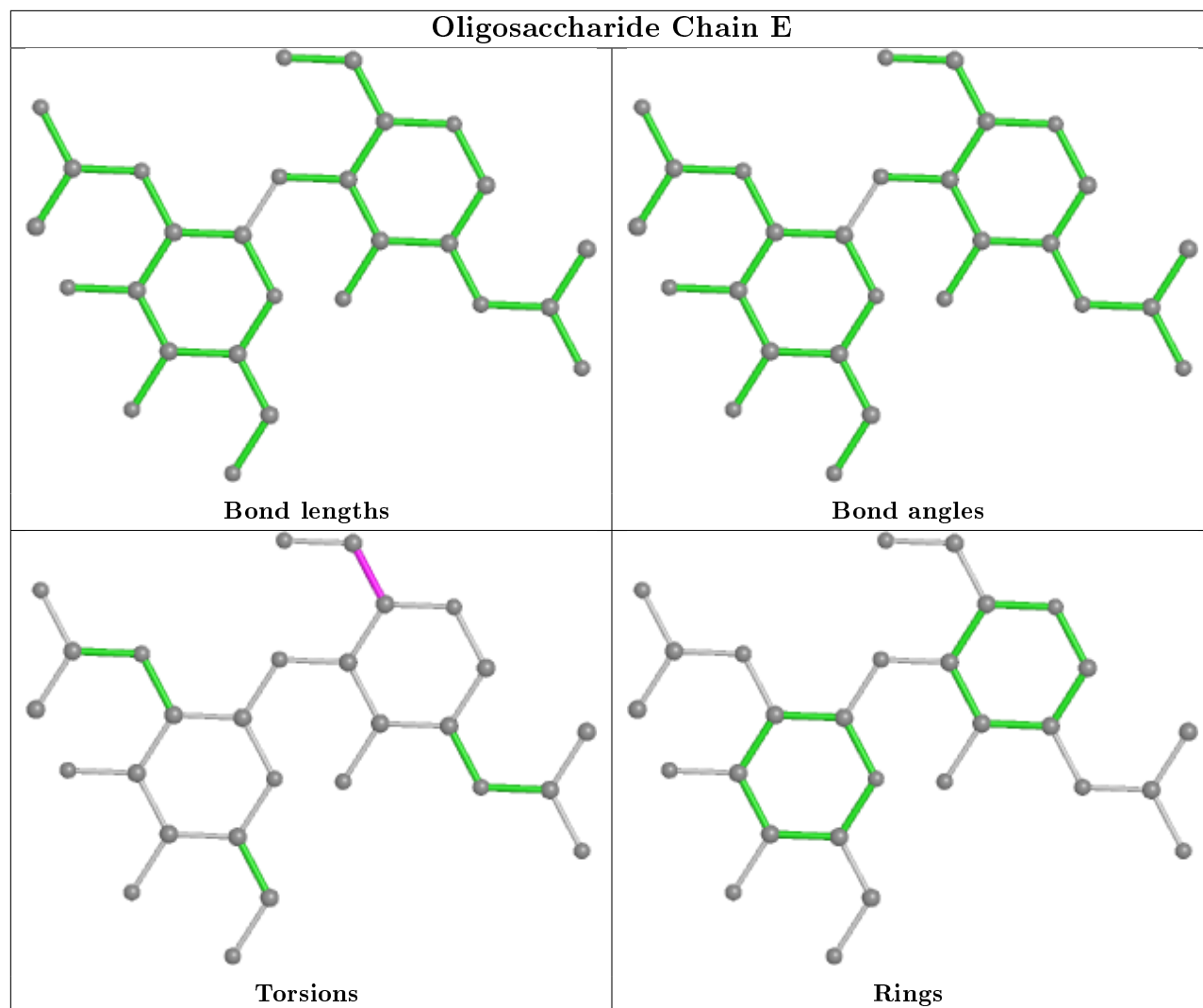
5 monomers are involved in 9 short contacts:

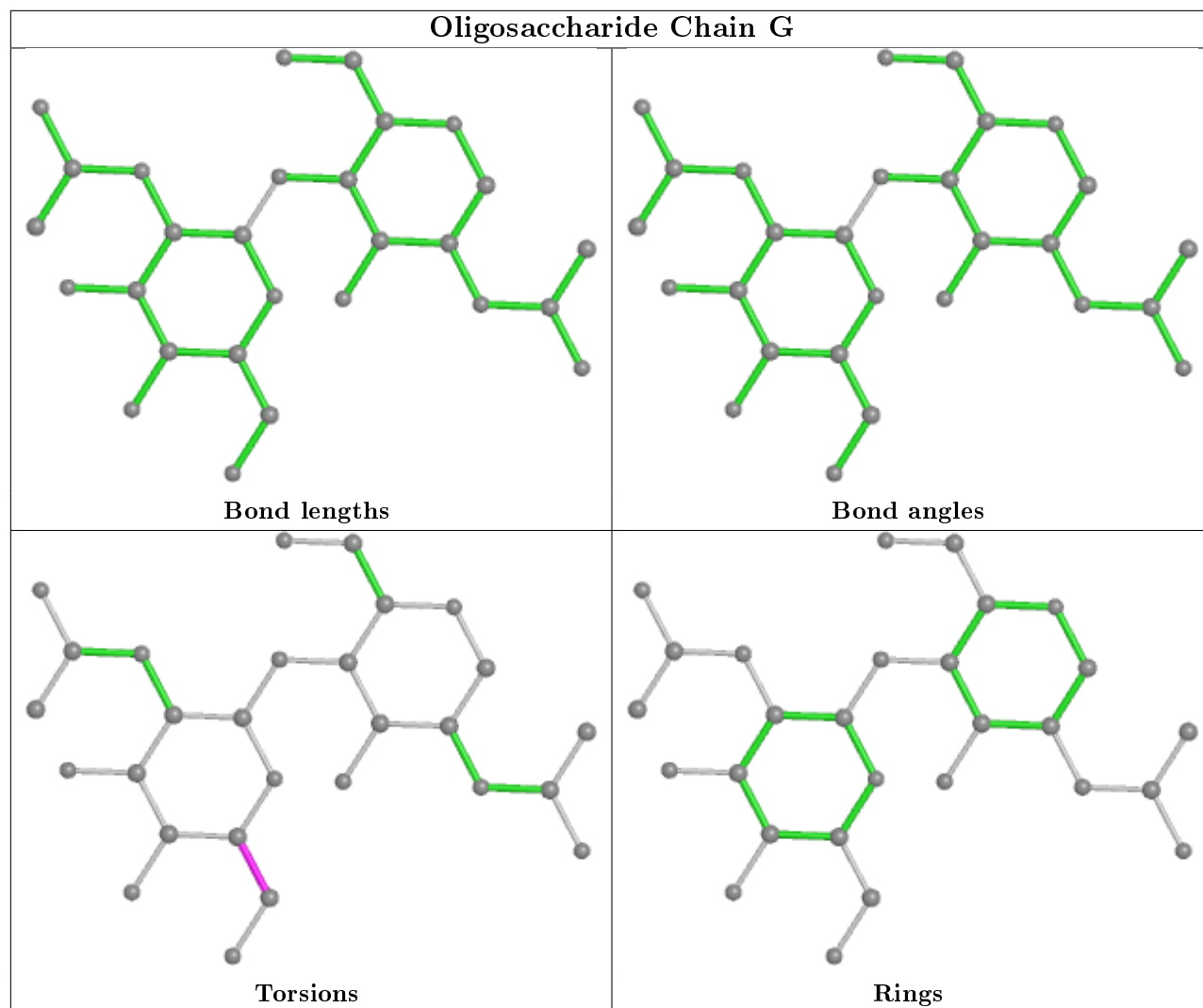
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	3	BMA	3	0
4	D	4	MAN	5	0
5	G	2	NAG	1	0
5	G	1	NAG	3	0
4	D	6	MAN	2	0

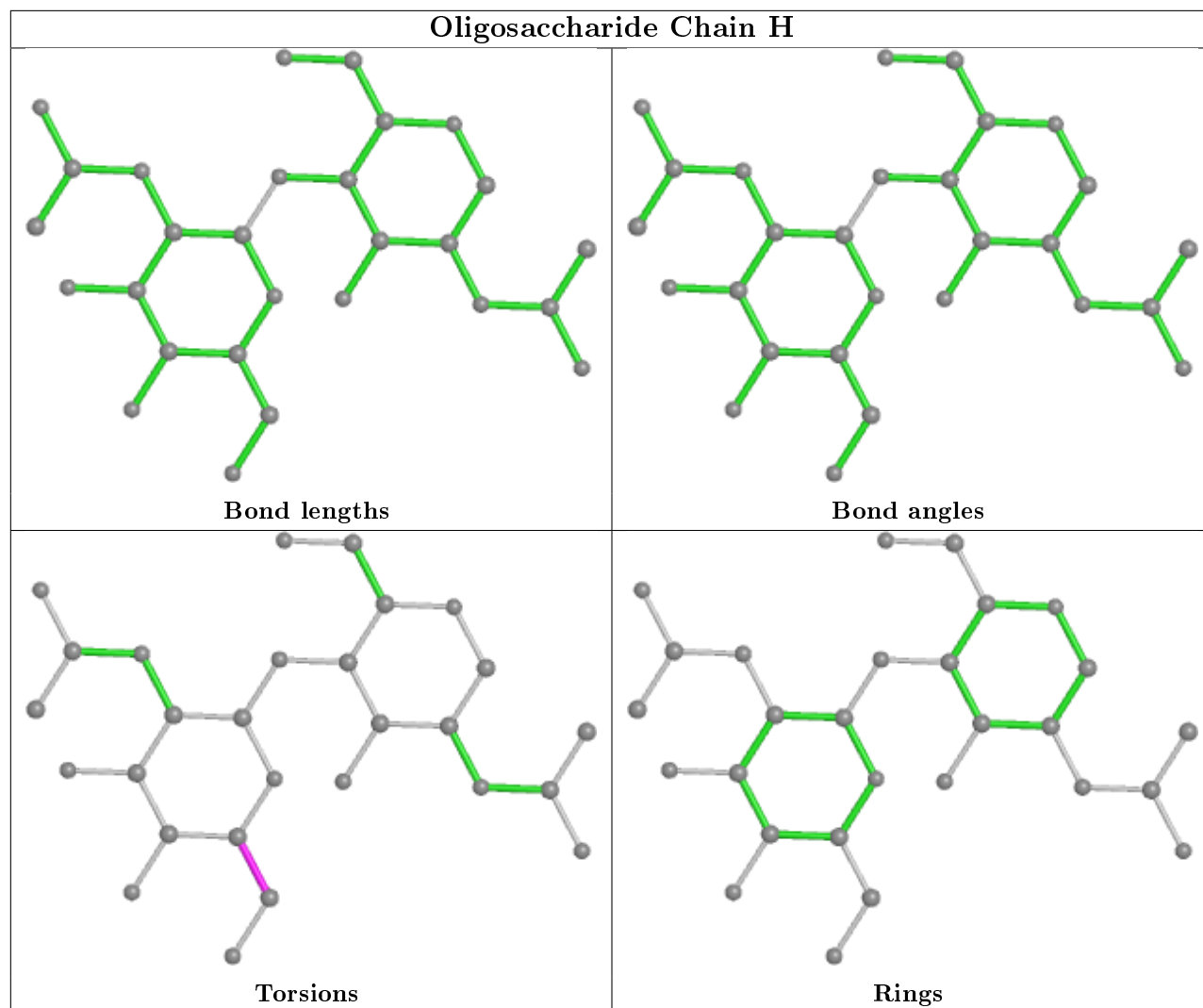
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

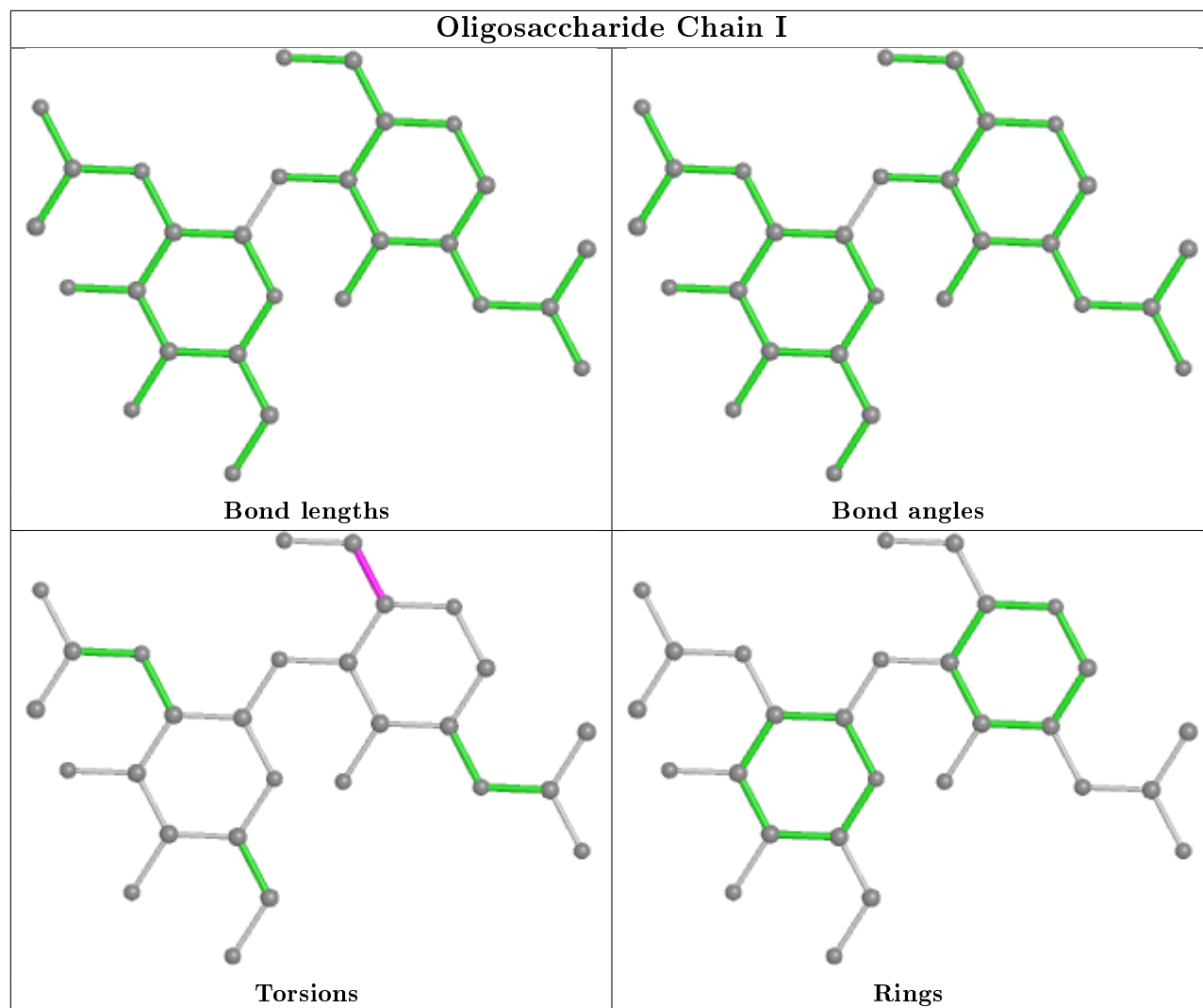


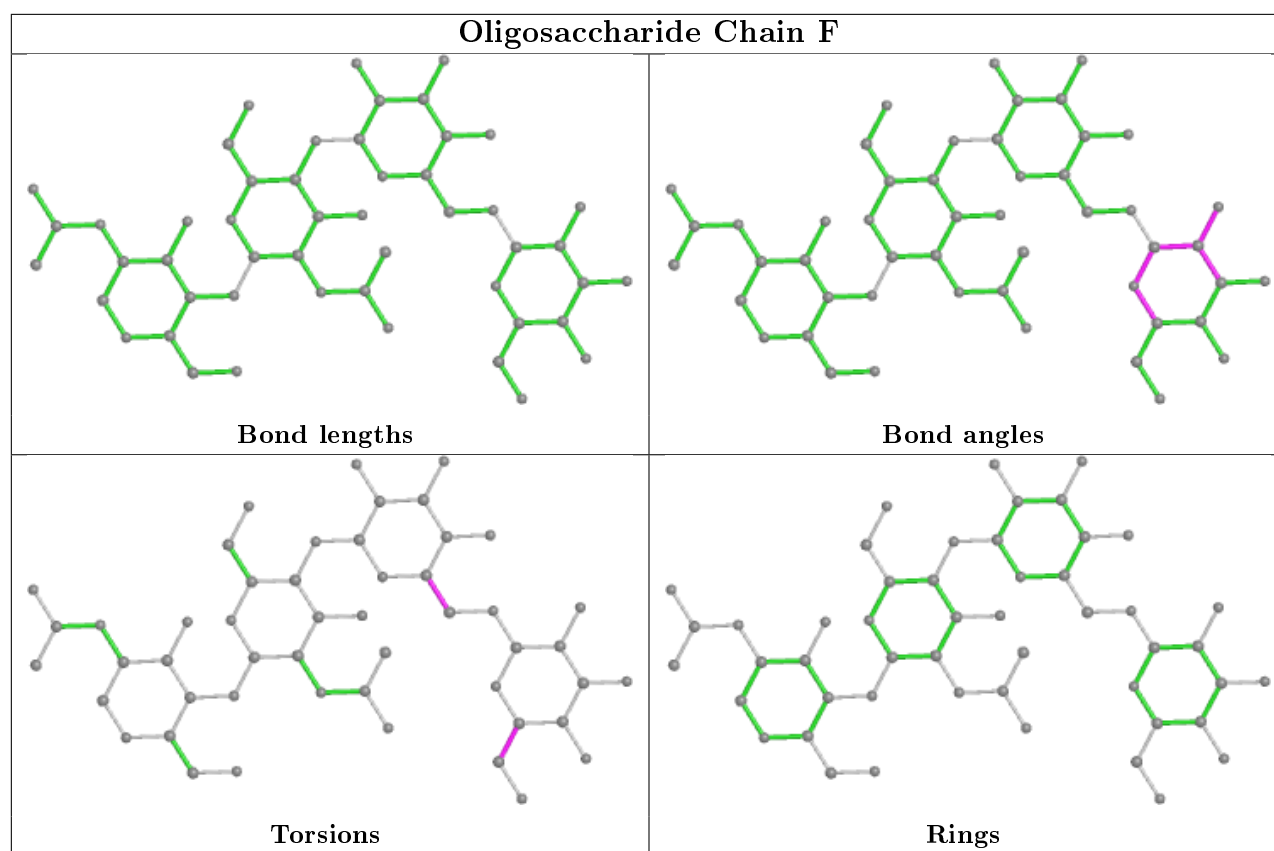












## 5.6 Ligand geometry [i](#)

Of 11 ligands modelled in this entry, 7 are monoatomic - leaving 4 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	$\# Z  > 2$
8	NAG	A	520	1	14,14,15	0.29	0	17,19,21	0.61	0
8	NAG	B	510	2	14,14,15	0.30	0	17,19,21	0.60	0
8	NAG	B	508	2	14,14,15	0.24	0	17,19,21	0.41	0
8	NAG	B	509	2	14,14,15	0.33	0	17,19,21	0.45	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	NAG	A	520	1	-	0/6/23/26	0/1/1/1
8	NAG	B	510	2	-	2/6/23/26	0/1/1/1
8	NAG	B	508	2	-	2/6/23/26	0/1/1/1
8	NAG	B	509	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
8	B	508	NAG	O5-C5-C6-O6
8	B	508	NAG	C4-C5-C6-O6
8	B	510	NAG	O5-C5-C6-O6
8	B	510	NAG	C4-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	B	509	NAG	3	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

### 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	448/452 (99%)	0.60	37 (8%) 11 11	49, 79, 130, 216	0
2	B	427/445 (95%)	0.97	74 (17%) 1 1	46, 83, 173, 286	0
3	C	6/6 (100%)	0.57	1 (16%) 1 1	92, 101, 127, 138	0
All	All	881/903 (97%)	0.78	112 (12%) 3 3	46, 81, 163, 286	0

All (112) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	60	ILE	9.2
1	A	84	LEU	9.1
1	A	86	SER	6.0
2	B	416	PRO	6.0
2	B	7	CYS	5.7
2	B	21	ALA	5.6
2	B	417	LYS	5.4
2	B	42	ALA	5.4
2	B	19	ILE	5.4
2	B	6	ARG	5.3
2	B	44	CYS	5.2
2	B	418	LYS	5.1
2	B	50	LEU	5.0
2	B	24	ASN	4.9
2	B	12	ALA	4.9
2	B	8	LEU	4.8
2	B	56	PRO	4.8
2	B	119	GLU	4.7
2	B	22	GLY	4.6
2	B	419	ASP	4.6
2	B	20	GLN	4.5
1	A	29	GLY	4.3
2	B	421	ASP	4.3

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	25	CYS	4.3
1	A	87	SER	4.1
2	B	53	LYS	4.1
2	B	62	ASN	3.9
2	B	33	PHE	3.8
2	B	59	ASP	3.7
1	A	30	THR	3.7
2	B	410	ILE	3.6
1	A	90	GLU	3.6
2	B	45	ASP	3.6
2	B	34	LEU	3.6
2	B	48	GLU	3.5
2	B	55	CYS	3.5
2	B	415	CYS	3.5
2	B	441	ILE	3.5
2	B	298	SER	3.4
2	B	52	LYS	3.4
2	B	270	LEU	3.4
2	B	442	CYS	3.4
1	A	126	GLU	3.3
2	B	103	LEU	3.3
1	A	283	SER	3.3
1	A	137	LEU	3.3
2	B	104	ARG	3.2
2	B	64	ARG	3.2
2	B	85	LYS	3.2
2	B	23	PRO	3.2
2	B	26	GLY	3.1
1	A	271	ARG	3.1
2	B	27	TRP	3.1
1	A	28	PRO	3.1
2	B	63	PRO	3.0
2	B	43	ARG	3.0
2	B	14	SER	3.0
2	B	384	GLY	3.0
2	B	190	GLU	2.9
1	A	287	TYR	2.8
1	A	286	GLY	2.7
1	A	282	ALA	2.7
2	B	264	PHE	2.7
2	B	29	THR	2.7
2	B	300	ALA	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	47	LEU	2.7
2	B	58	ASP	2.7
2	B	265	ALA	2.6
1	A	230	TYR	2.6
1	A	83	SER	2.6
2	B	5	ASN	2.6
1	A	284	TYR	2.6
1	A	164	CYS	2.5
2	B	18	CYS	2.5
2	B	28	CYS	2.5
2	B	61	GLU	2.5
1	A	1	PHE	2.4
1	A	392	GLY	2.4
1	A	352	GLY	2.4
2	B	65	GLY	2.4
2	B	385	VAL	2.4
1	A	253	PRO	2.4
1	A	270	ILE	2.4
1	A	188	TRP	2.4
1	A	425	LEU	2.4
3	C	5005	PRO	2.3
2	B	444	CYS	2.3
2	B	297	PRO	2.3
2	B	49	ALA	2.3
2	B	173	MET	2.3
2	B	102	ARG	2.3
2	B	420	SER	2.2
2	B	380	TYR	2.2
2	B	11	ASN	2.2
1	A	207	GLU	2.2
1	A	208	TYR	2.2
2	B	440	TYR	2.2
1	A	416	GLY	2.1
2	B	51	LYS	2.1
1	A	105	VAL	2.1
1	A	233	TYR	2.1
2	B	16	GLY	2.1
1	A	231	LEU	2.1
1	A	452	VAL	2.0
1	A	257	LEU	2.0
2	B	17	GLU	2.0
1	A	166	GLY	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	306	ALA	2.0
1	A	72	PHE	2.0
1	A	410	HIS	2.0
2	B	41	SER	2.0
2	B	191	GLN	2.0

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

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

## 6.3 Carbohydrates [i](#)

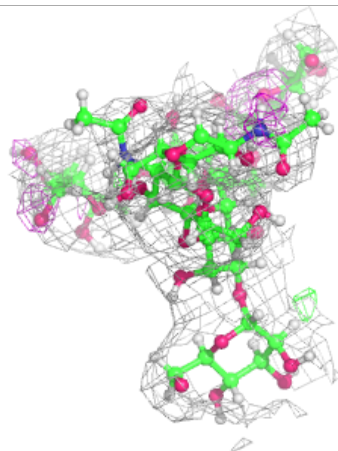
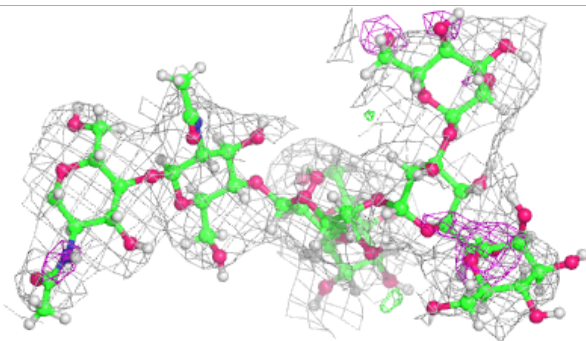
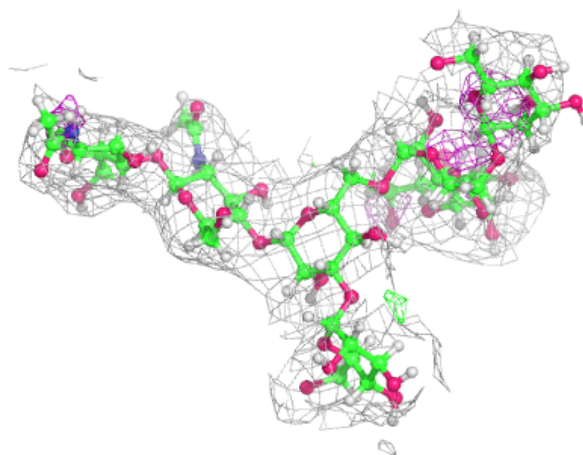
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
4	MAN	D	4	11/12	0.68	0.21	134,141,169,169	0
5	NAG	E	2	14/15	0.68	0.28	119,129,154,154	0
6	BMA	F	3	11/12	0.70	0.37	139,146,174,175	0
5	NAG	I	2	14/15	0.71	0.40	124,132,156,156	0
5	NAG	G	2	14/15	0.73	0.37	69,110,135,143	0
4	MAN	D	6	11/12	0.74	0.53	145,151,181,182	0
6	MAN	F	4	11/12	0.74	0.41	149,151,181,181	0
4	BMA	D	3	11/12	0.81	0.11	112,125,150,150	0
5	NAG	H	1	14/15	0.82	0.15	58,102,129,134	0
5	NAG	H	2	14/15	0.82	0.23	105,130,157,165	0
4	MAN	D	5	11/12	0.82	0.28	77,105,127,140	0
6	NAG	F	2	14/15	0.83	0.20	115,126,147,151	0
4	MAN	D	7	11/12	0.85	0.14	119,124,149,150	0
5	NAG	G	1	14/15	0.86	0.13	148,169,200,203	0
5	NAG	E	1	14/15	0.88	0.11	97,107,125,128	0
5	NAG	I	1	14/15	0.90	0.21	99,115,135,138	0
4	NAG	D	1	14/15	0.91	0.19	66,79,93,95	0
6	NAG	F	1	14/15	0.94	0.20	85,103,116,125	0
4	NAG	D	2	14/15	0.96	0.13	87,95,110,114	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.

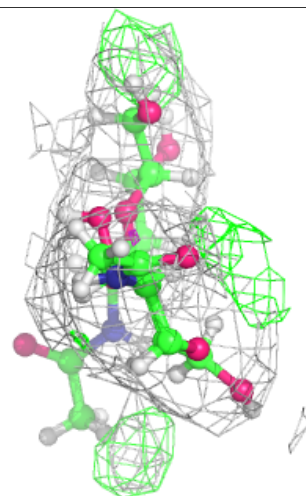
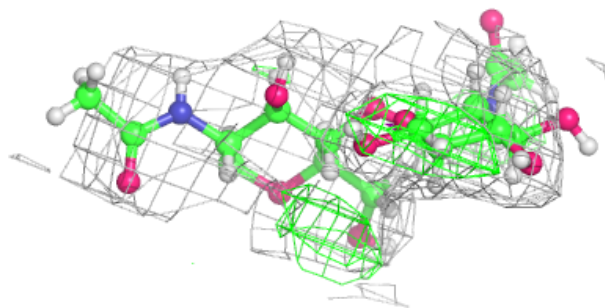
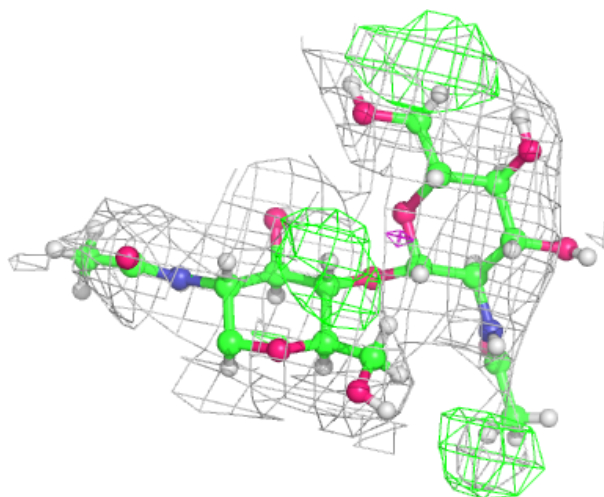
**Electron density around Chain D:**

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



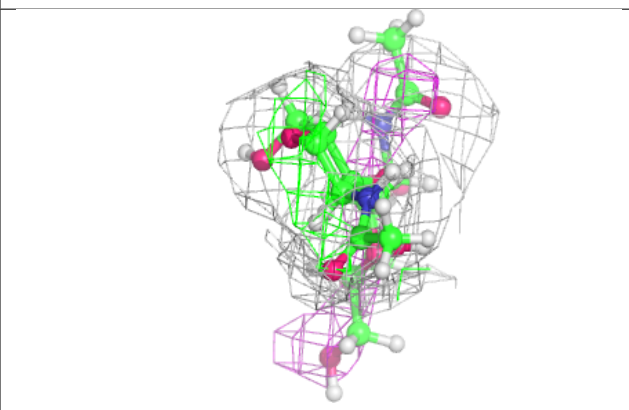
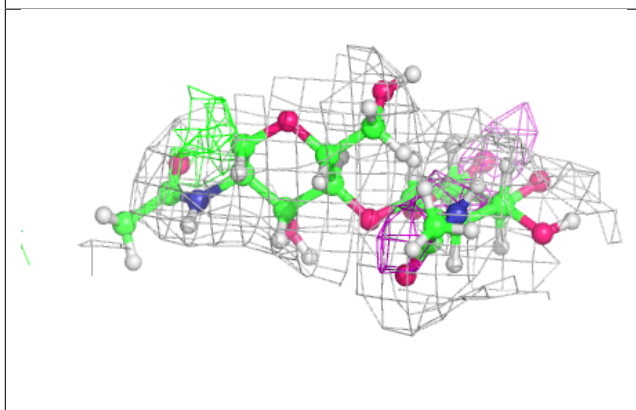
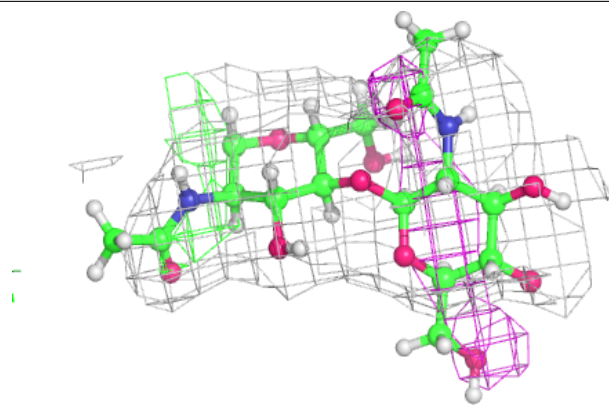
**Electron density around Chain E:**

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



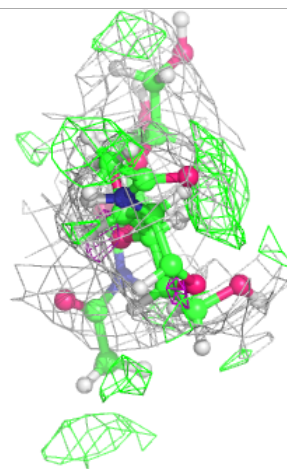
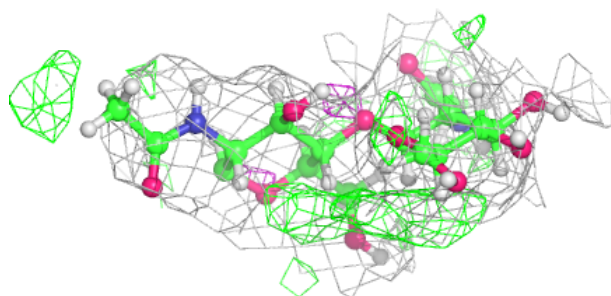
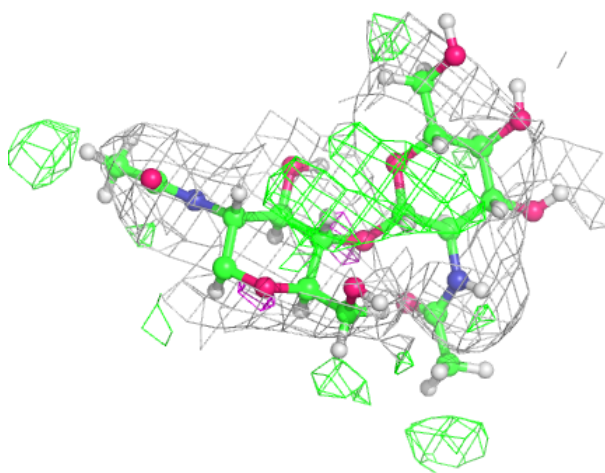
**Electron density around Chain G:**

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



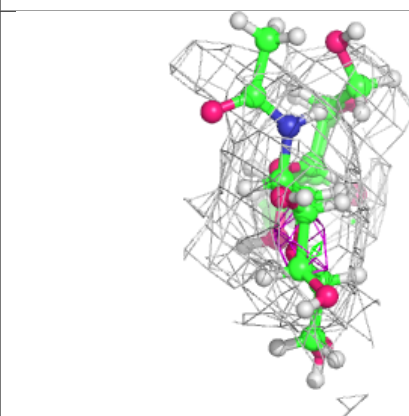
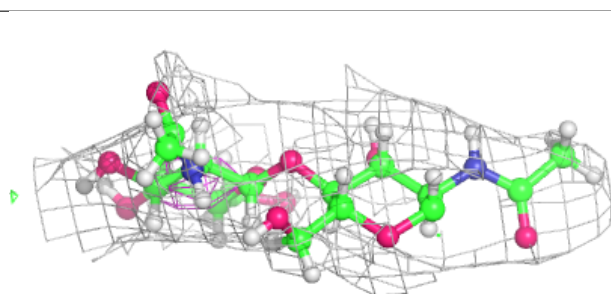
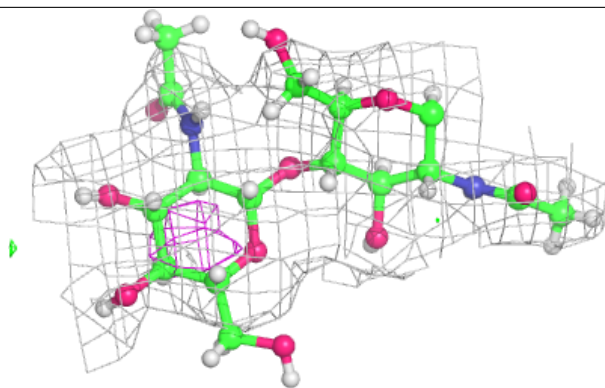
**Electron density around Chain H:**

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

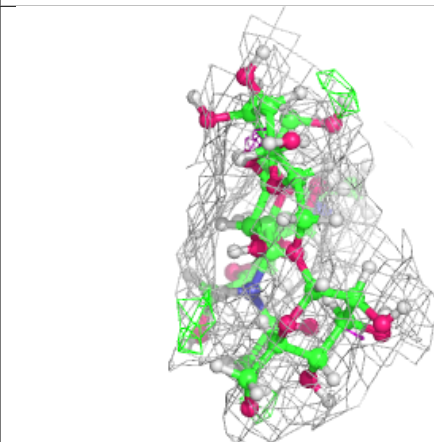
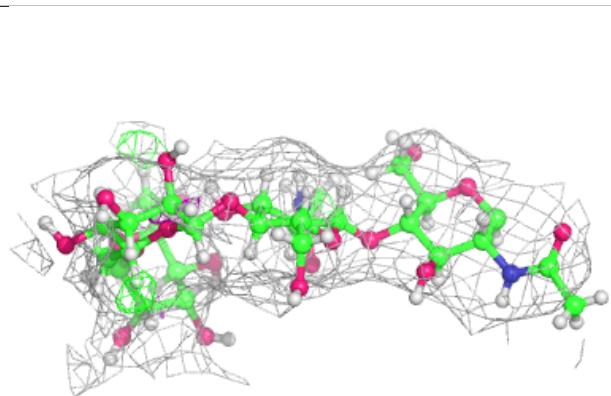
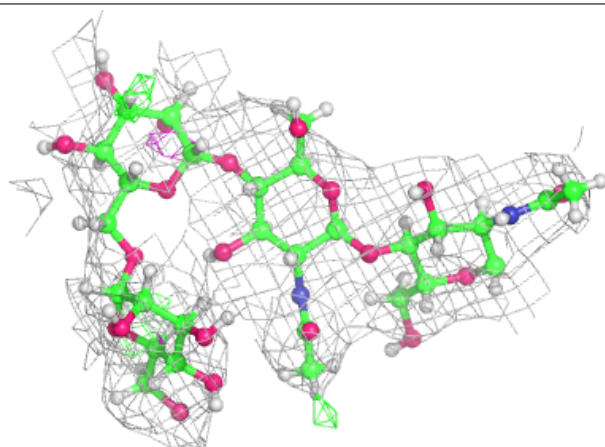


**Electron density around Chain I:**

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

**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)





## 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
7	CA	B	502	1/1	0.63	0.13	156,156,156,156	0
8	NAG	B	509	14/15	0.68	0.40	156,159,190,191	0
8	NAG	B	510	14/15	0.81	0.33	129,138,165,165	0
8	NAG	B	508	14/15	0.83	0.36	121,135,161,162	0
8	NAG	A	520	14/15	0.88	0.22	110,120,142,144	0
9	MG	B	501	1/1	0.95	0.20	59,59,59,59	0
7	CA	A	501	1/1	0.96	0.06	87,87,87,87	0
7	CA	A	504	1/1	0.96	0.06	81,81,81,81	0
7	CA	A	503	1/1	0.96	0.09	66,66,66,66	0
7	CA	B	503	1/1	0.97	0.22	66,66,66,66	0
7	CA	A	502	1/1	0.98	0.10	73,73,73,73	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.