



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 7, 2020 – 07:39 AM BST

PDB ID : 3E80  
Title : Structure of Heparinase II complexed with heparan sulfate degradation disaccharide product  
Authors : Shaya, D.; Cygler, M.  
Deposited on : 2008-08-19  
Resolution : 2.35 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 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

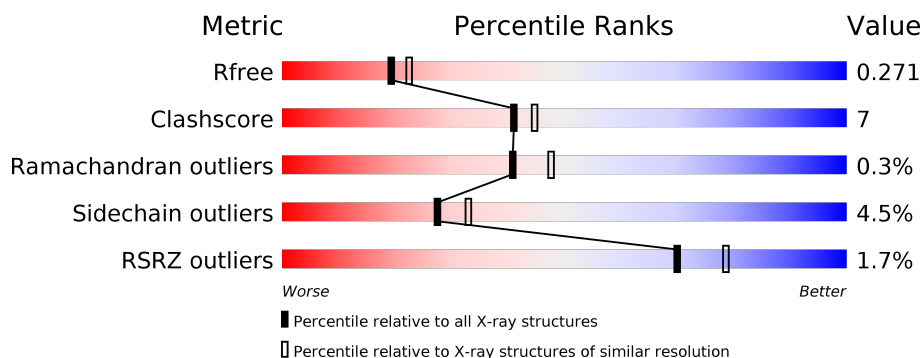
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.35 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.



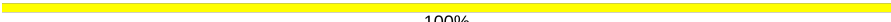


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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	749	<div> <div>0%</div> <div> <div></div> <div>81%</div> <div>17%</div> <div>.</div> </div> </div>
1	B	749	<div> <div>2%</div> <div> <div></div> <div>83%</div> <div>16%</div> <div>.</div> </div> </div>
1	C	749	<div> <div>2%</div> <div> <div></div> <div>79%</div> <div>19%</div> <div>.</div> </div> </div>
2	D	4	<div> <div></div> <div>100%</div> </div>
2	E	4	<div> <div></div> <div>100%</div> </div>
2	F	4	<div> <div></div> <div> <div>25%</div> <div>75%</div> </div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	2	 100%
3	H	2	 50% 50%
3	I	2	 50% 50%

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
2	XYS	E	3	X	-	-	-
2	GCU	F	2	X	-	-	-
5	PO4	C	6	-	-	X	-

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 18716 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.

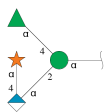
- Molecule 1 is a protein called Heparinase II protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	747	Total	C	N	O	S	0	0	0
			5978	3850	1008	1097	23			
1	B	747	Total	C	N	O	S	0	0	0
			5978	3850	1008	1097	23			
1	C	747	Total	C	N	O	S	0	0	0
			5975	3847	1008	1097	23			

There are 3 discrepancies between the modelled and reference sequences:

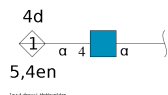
Chain	Residue	Modelled	Actual	Comment	Reference
A	758	ALA	PRO	SEE REMARK 999	UNP Q46080
B	758	ALA	PRO	SEE REMARK 999	UNP Q46080
C	758	ALA	PRO	SEE REMARK 999	UNP Q46080

- Molecule 2 is an oligosaccharide called alpha-D-xylopyranose-(1-4)-alpha-D-glucopyranuronic acid-(1-2)-[alpha-L-rhamnopyranose-(1-4)]alpha-D-mannopyranose.



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	D	4	Total	C	O	0	0	0
			41	23	18			
2	E	4	Total	C	O	0	0	0
			41	23	18			
2	F	4	Total	C	O	0	0	0
			41	23	18			

- Molecule 3 is an oligosaccharide called 4-deoxy-alpha-L-threo-hex-4-enopyranuronic acid-(1-4)-2-acetamido-2-deoxy-alpha-D-glucopyranose.

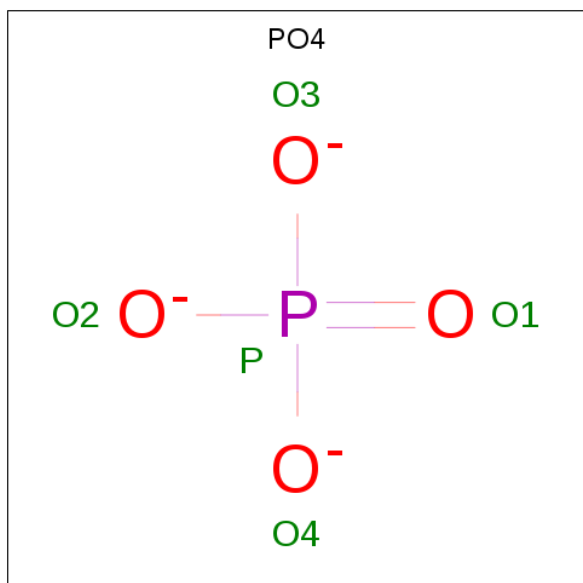


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	G	2	Total	C	N	O	0	0	0
			26	14	1	11			
3	H	2	Total	C	N	O	0	0	0
			26	14	1	11			
3	I	2	Total	C	N	O	0	0	0
			26	14	1	11			

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Zn	0	0
			1	1		
4	A	1	Total	Zn	0	0
			1	1		
4	C	1	Total	Zn	0	0
			1	1		

- Molecule 5 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



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

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	P	0	0
			5	4	1		
5	A	1	Total	O	P	0	0
			5	4	1		
5	A	1	Total	O	P	0	0
			5	4	1		
5	B	1	Total	O	P	0	0
			5	4	1		
5	B	1	Total	O	P	0	0
			5	4	1		
5	B	1	Total	O	P	0	0
			5	4	1		
5	C	1	Total	O	P	0	0
			5	4	1		
5	C	1	Total	O	P	0	0
			5	4	1		

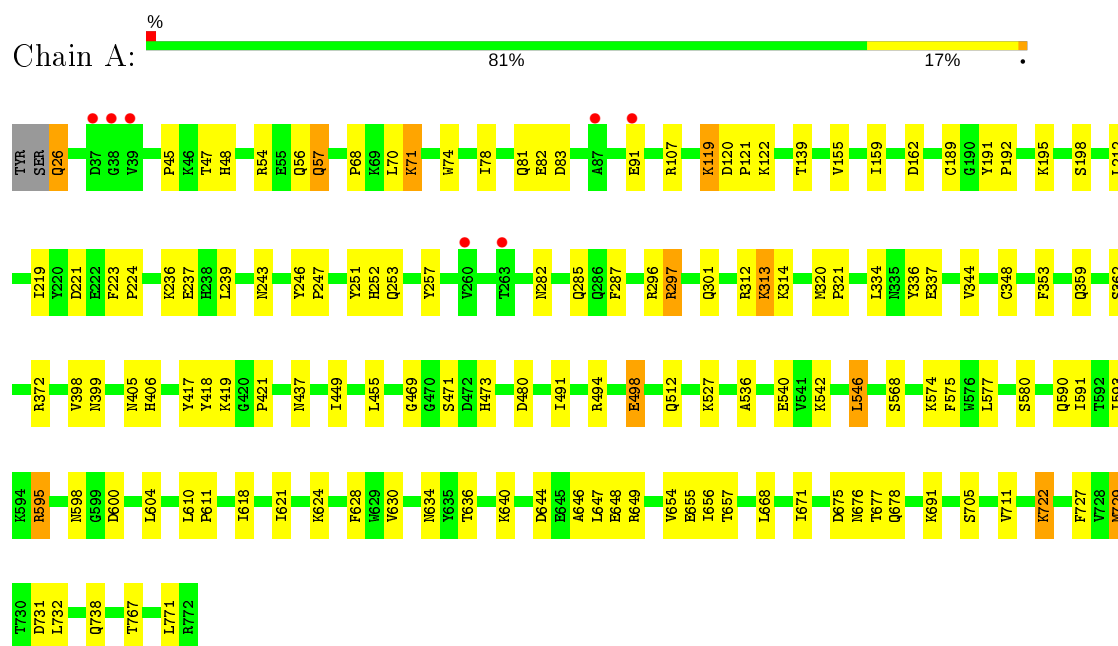
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	208	Total	O	0	0
			208	208		
6	B	151	Total	O	0	0
			151	151		
6	C	177	Total	O	0	0
			177	177		

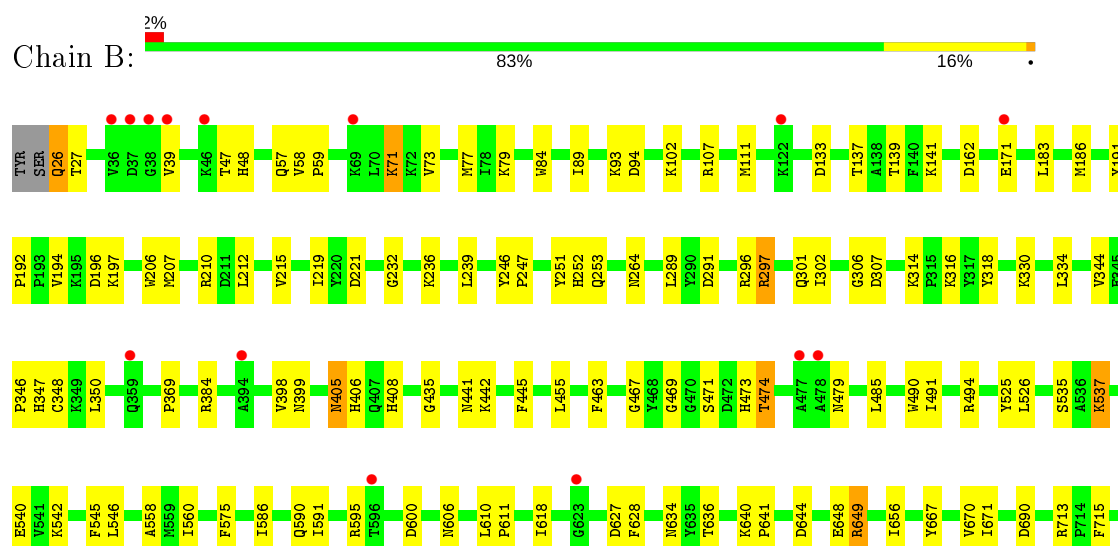
### 3 Residue-property plots [i](#)

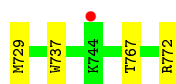
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Heparinase II protein

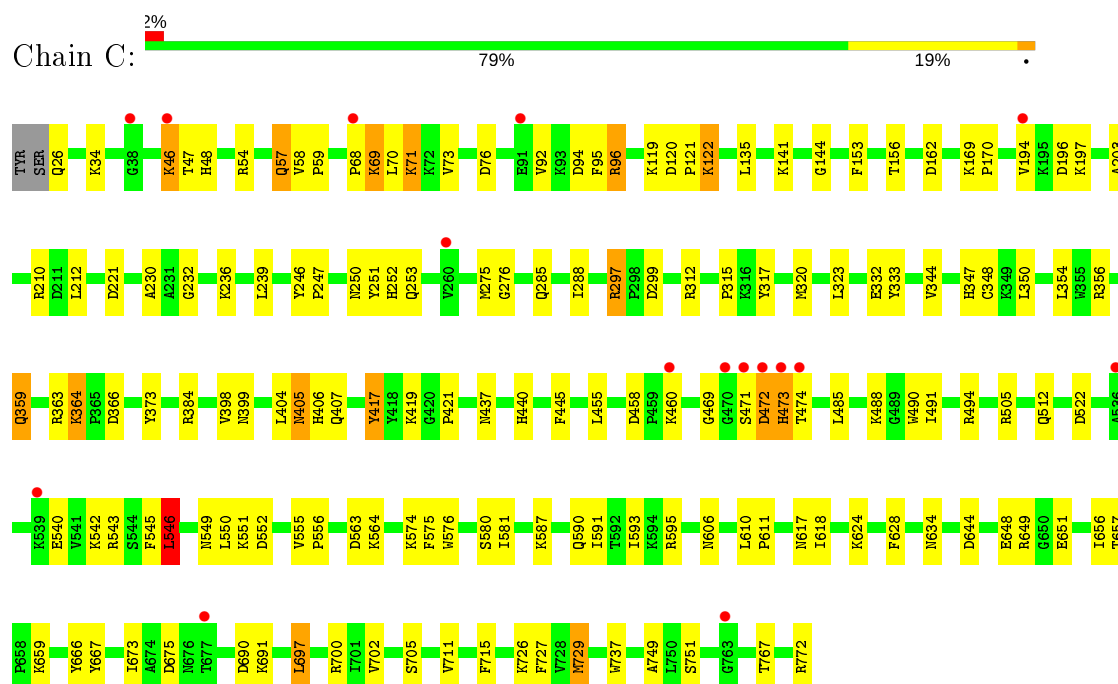


#### • Molecule 1: Heparinase II protein





- Molecule 1: Heparinase II protein



- Molecule 2: alpha-D-xylopyranose-(1-4)-alpha-D-glucopyranuronic acid-(1-2)-[alpha-L-rhamnopyranose-(1-4)]alpha-D-mannopyranose



- Molecule 2: alpha-D-xylopyranose-(1-4)-alpha-D-glucopyranuronic acid-(1-2)-[alpha-L-rhamnopyranose-(1-4)]alpha-D-mannopyranose



- Molecule 2: alpha-D-xylopyranose-(1-4)-alpha-D-glucopyranuronic acid-(1-2)-[alpha-L-rhamnopyranose-(1-4)]alpha-D-mannopyranose



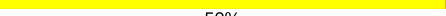

- Molecule 3: 4-deoxy-alpha-L-threo-hex-4-enopyranuronic acid-(1-4)-2-acetamido-2-deoxy-alpha-D-glucopyranose



Chain G:  100%


HDG1  
GCD2

- Molecule 3: 4-deoxy-alpha-L-threo-hex-4-enopyranuronic acid-(1-4)-2-acetamido-2-deoxy-alpha-D-glucopyranose

Chain H:  50%  50%

HDG1  
GCD2

- Molecule 3: 4-deoxy-alpha-L-threo-hex-4-enopyranuronic acid-(1-4)-2-acetamido-2-deoxy-alpha-D-glucopyranose

Chain I:  50%  50%

HDG1  
GCD2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	201.28Å 209.36Å 59.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	37.57 – 2.35 37.57 – 2.35	Depositor EDS
% Data completeness (in resolution range)	98.0 (37.57-2.35) 98.1 (37.57-2.35)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.85 (at 2.34Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.229 , 0.267 0.230 , 0.271	Depositor DCC
$R_{free}$ test set	5152 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	20.5	Xtriage
Anisotropy	0.274	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 38.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	18716	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 23.28 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 4.8089e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: ZN, XYS, GCD, PO4, RAM, NDG, GCU, PCA, MAN

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	0.53	0/6133	0.78	3/8303 (0.0%)
1	B	0.48	0/6133	0.73	3/8303 (0.0%)
1	C	0.51	0/6130	0.74	2/8299 (0.0%)
All	All	0.50	0/18396	0.75	8/24905 (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
1	A	0	1

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	26	PCA	O-C-N	-11.54	104.24	122.70
1	A	26	PCA	O-C-N	-10.11	106.52	122.70
1	A	26	PCA	CA-C-N	10.05	139.31	117.20
1	C	26	PCA	O-C-N	-8.42	109.23	122.70
1	B	26	PCA	C-N-CA	-6.93	104.37	121.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	26	PCA	Mainchain

## 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	5978	0	5889	86	0
1	B	5978	0	5889	72	0
1	C	5975	0	5880	98	0
2	D	41	0	29	0	0
2	E	41	0	30	1	0
2	F	41	0	29	0	0
3	G	26	0	16	0	0
3	H	26	0	16	1	0
3	I	26	0	16	2	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
5	A	20	0	0	0	0
5	B	15	0	0	0	0
5	C	10	0	0	3	0
6	A	208	0	0	9	0
6	B	151	0	0	4	0
6	C	177	0	0	7	0
All	All	18716	0	17794	257	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 7.

The worst 5 of 257 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:253:GLN:HE22	1:B:405:ASN:HB3	1.26	0.98
1:C:574:LYS:O	1:C:657:THR:HG22	1.74	0.87
1:C:253:GLN:HE22	1:C:405:ASN:HB3	1.46	0.79
1:A:54:ARG:H	1:A:57:GLN:HE21	1.26	0.79
1:B:628:PHE:H	1:B:634:ASN:HD21	1.33	0.74

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	745/749 (100%)	696 (93%)	47 (6%)	2 (0%)	41	47
1	B	745/749 (100%)	695 (93%)	49 (7%)	1 (0%)	51	63
1	C	745/749 (100%)	694 (93%)	48 (6%)	3 (0%)	34	38
All	All	2235/2247 (100%)	2085 (93%)	144 (6%)	6 (0%)	41	47

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	472	ASP
1	A	91	GLU
1	B	399	ASN
1	C	473	HIS
1	A	676	ASN

### 5.3.2 Protein sidechains

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	632/634 (100%)	609 (96%)	23 (4%)	35	43
1	B	632/634 (100%)	604 (96%)	28 (4%)	28	34
1	C	631/634 (100%)	597 (95%)	34 (5%)	22	25
All	All	1895/1902 (100%)	1810 (96%)	85 (4%)	27	33

5 of 85 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	314	LYS
1	B	649	ARG
1	C	590	GLN
1	B	346	PRO
1	B	537	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	252	HIS
1	B	473	HIS
1	C	437	ASN
1	B	253	GLN
1	B	347	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

3 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
1	PCA	B	26	1	7,8,9	2.19	1 (14%)	9,10,12	2.36	5 (55%)
1	PCA	A	26	1	7,8,9	2.24	1 (14%)	9,10,12	3.33	6 (66%)
1	PCA	C	26	1	7,8,9	2.27	2 (28%)	9,10,12	2.92	4 (44%)

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
1	PCA	B	26	1	-	0/0/11/13	0/1/1/1
1	PCA	A	26	1	-	0/0/11/13	0/1/1/1
1	PCA	C	26	1	-	0/0/11/13	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	26	PCA	CD-N	5.60	1.49	1.34
1	C	26	PCA	CD-N	5.46	1.49	1.34
1	B	26	PCA	CD-N	5.33	1.48	1.34
1	C	26	PCA	CA-N	2.02	1.48	1.46

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	26	PCA	OE-CD-CG	-5.36	117.42	126.76
1	C	26	PCA	CB-CG-CD	-4.97	96.39	104.40
1	A	26	PCA	CB-CG-CD	-4.73	96.78	104.40
1	A	26	PCA	O-C-CA	-4.65	112.60	124.78
1	C	26	PCA	CB-CA-C	4.45	118.83	112.70

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	26	PCA	1	0

## 5.5 Carbohydrates ⓘ

18 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
2	MAN	D	1	1,2	11,11,12	1.74	2 (18%)	15,15,17	3.27	3 (20%)
2	GCU	D	2	2	9,12,13	1.46	1 (11%)	12,17,19	3.36	4 (33%)
2	XYS	D	3	2	8,8,10	1.54	2 (25%)	8,10,14	2.76	2 (25%)
2	RAM	D	4	2	10,10,11	0.67	0	14,14,16	1.69	3 (21%)
2	MAN	E	1	1,2	11,11,12	1.79	2 (18%)	15,15,17	2.18	4 (26%)
2	GCU	E	2	2	9,12,13	1.56	1 (11%)	12,17,19	3.02	3 (25%)
2	XYS	E	3	2	8,8,10	1.35	1 (12%)	8,10,14	1.26	1 (12%)
2	RAM	E	4	2	10,10,11	0.58	0	14,14,16	0.99	0
2	MAN	F	1	1,2	11,11,12	1.81	2 (18%)	15,15,17	3.00	2 (13%)
2	GCU	F	2	2	9,12,13	1.45	1 (11%)	12,17,19	4.45	4 (33%)
2	XYS	F	3	2	8,8,10	1.51	2 (25%)	8,10,14	3.16	5 (62%)
2	RAM	F	4	2	10,10,11	0.69	0	14,14,16	1.15	0
3	NDG	G	1	3	15,15,15	2.65	2 (13%)	21,21,21	2.87	5 (23%)
3	GCD	G	2	3	7,11,12	2.31	1 (14%)	8,15,17	2.03	3 (37%)
3	NDG	H	1	3	15,15,15	2.83	2 (13%)	21,21,21	2.04	6 (28%)
3	GCD	H	2	3	7,11,12	2.34	1 (14%)	8,15,17	1.67	2 (25%)
3	NDG	I	1	3	15,15,15	2.70	2 (13%)	21,21,21	2.62	9 (42%)
3	GCD	I	2	3	7,11,12	2.57	1 (14%)	8,15,17	1.41	2 (25%)

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
2	MAN	D	1	1,2	-	0/2/19/22	0/1/1/1
2	GCU	D	2	2	-	0/0/21/24	0/1/1/1
2	XYS	D	3	2	-	-	0/1/1/1
2	RAM	D	4	2	-	-	0/1/1/1
2	MAN	E	1	1,2	-	0/2/19/22	0/1/1/1
2	GCU	E	2	2	-	0/0/21/24	0/1/1/1
2	XYS	E	3	2	1/1/2/4	-	0/1/1/1
2	RAM	E	4	2	-	-	0/1/1/1
2	MAN	F	1	1,2	-	2/2/19/22	0/1/1/1
2	GCU	F	2	2	1/1/5/6	0/0/21/24	0/1/1/1
2	XYS	F	3	2	-	-	0/1/1/1
2	RAM	F	4	2	-	-	0/1/1/1
3	NDG	G	1	3	-	2/6/26/26	0/1/1/1

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GCD	G	2	3	-	0/0/17/20	0/1/1/1
3	NDG	H	1	3	-	1/6/26/26	0/1/1/1
3	GCD	H	2	3	-	0/0/17/20	0/1/1/1
3	NDG	I	1	3	-	3/6/26/26	0/1/1/1
3	GCD	I	2	3	-	0/0/17/20	0/1/1/1

The worst 5 of 23 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	H	1	NDG	O7-C7	9.97	1.45	1.23
3	I	1	NDG	O7-C7	9.46	1.44	1.23
3	G	1	NDG	O7-C7	9.38	1.44	1.23
3	I	2	GCD	O5-C5	6.23	1.46	1.37
3	H	2	GCD	O5-C5	5.80	1.45	1.37

The worst 5 of 58 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	2	GCU	O4-C4-C5	14.62	137.99	110.05
2	D	1	MAN	O2-C2-C3	10.91	131.99	110.14
2	F	1	MAN	O2-C2-C3	9.65	129.48	110.14
2	D	2	GCU	O4-C4-C5	9.57	128.34	110.05
2	E	2	GCU	O4-C4-C5	9.44	128.09	110.05

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	E	3	XYS	C1
2	F	2	GCU	C4

5 of 8 torsion outliers are listed below:

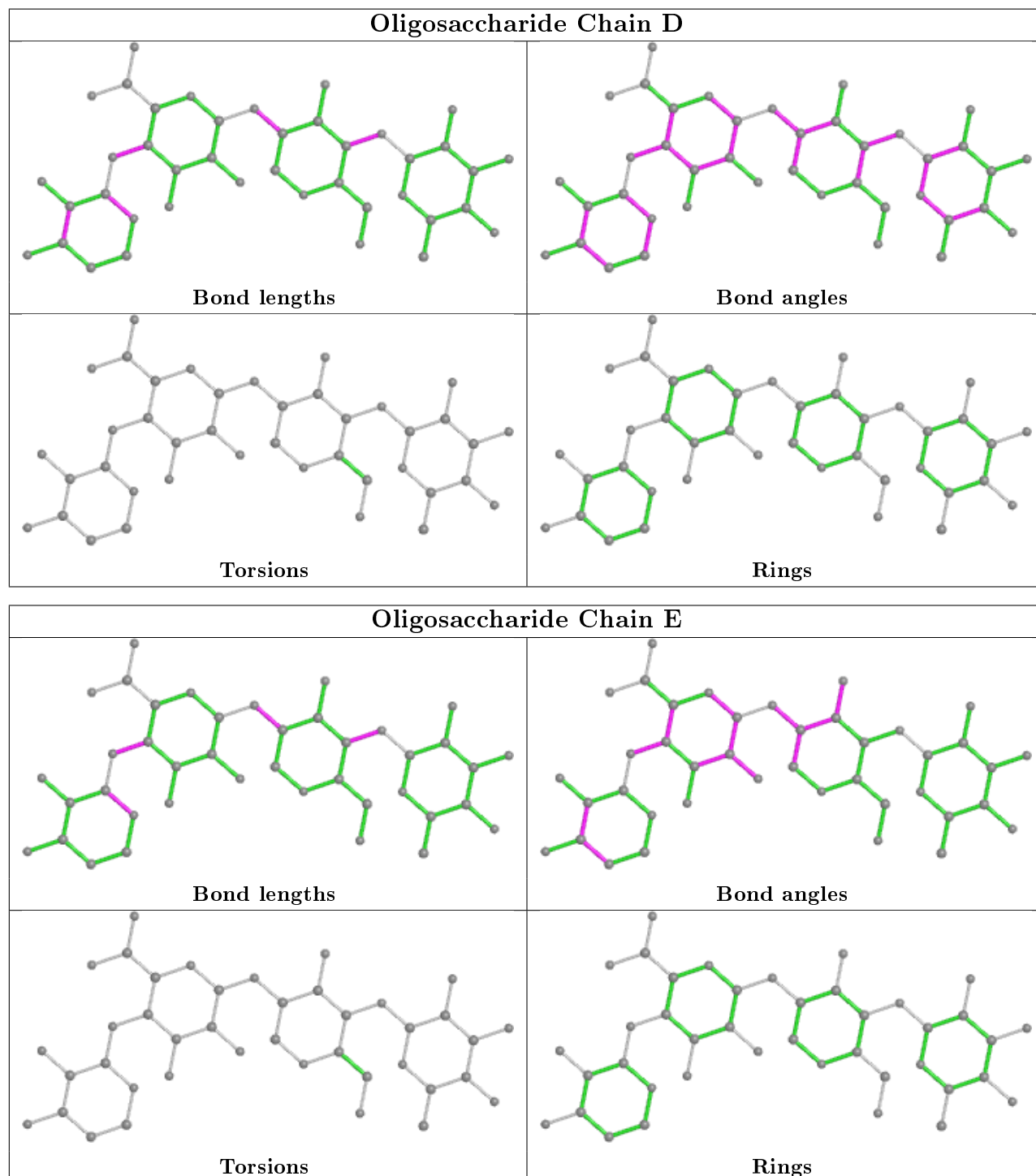
Mol	Chain	Res	Type	Atoms
3	G	1	NDG	O7-C7-N2-C2
3	I	1	NDG	C8-C7-N2-C2
3	H	1	NDG	O7-C7-N2-C2
3	I	1	NDG	O5-C5-C6-O6
3	I	1	NDG	C4-C5-C6-O6

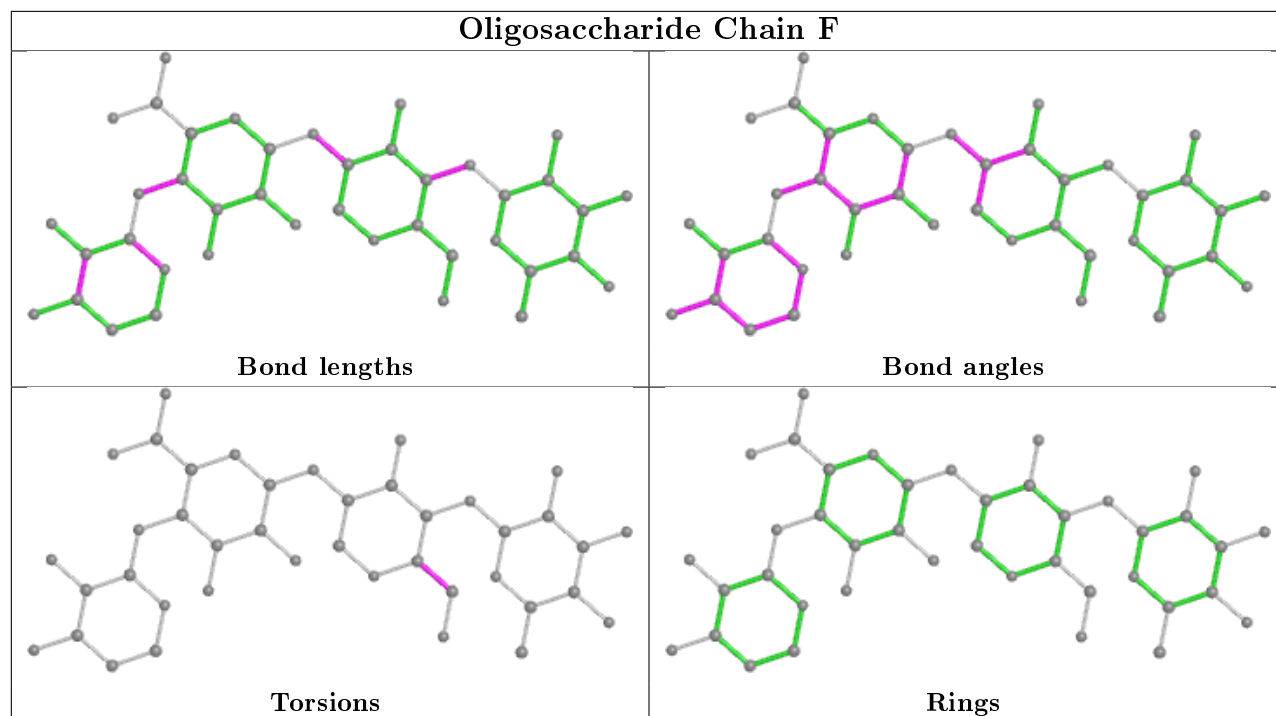
There are no ring outliers.

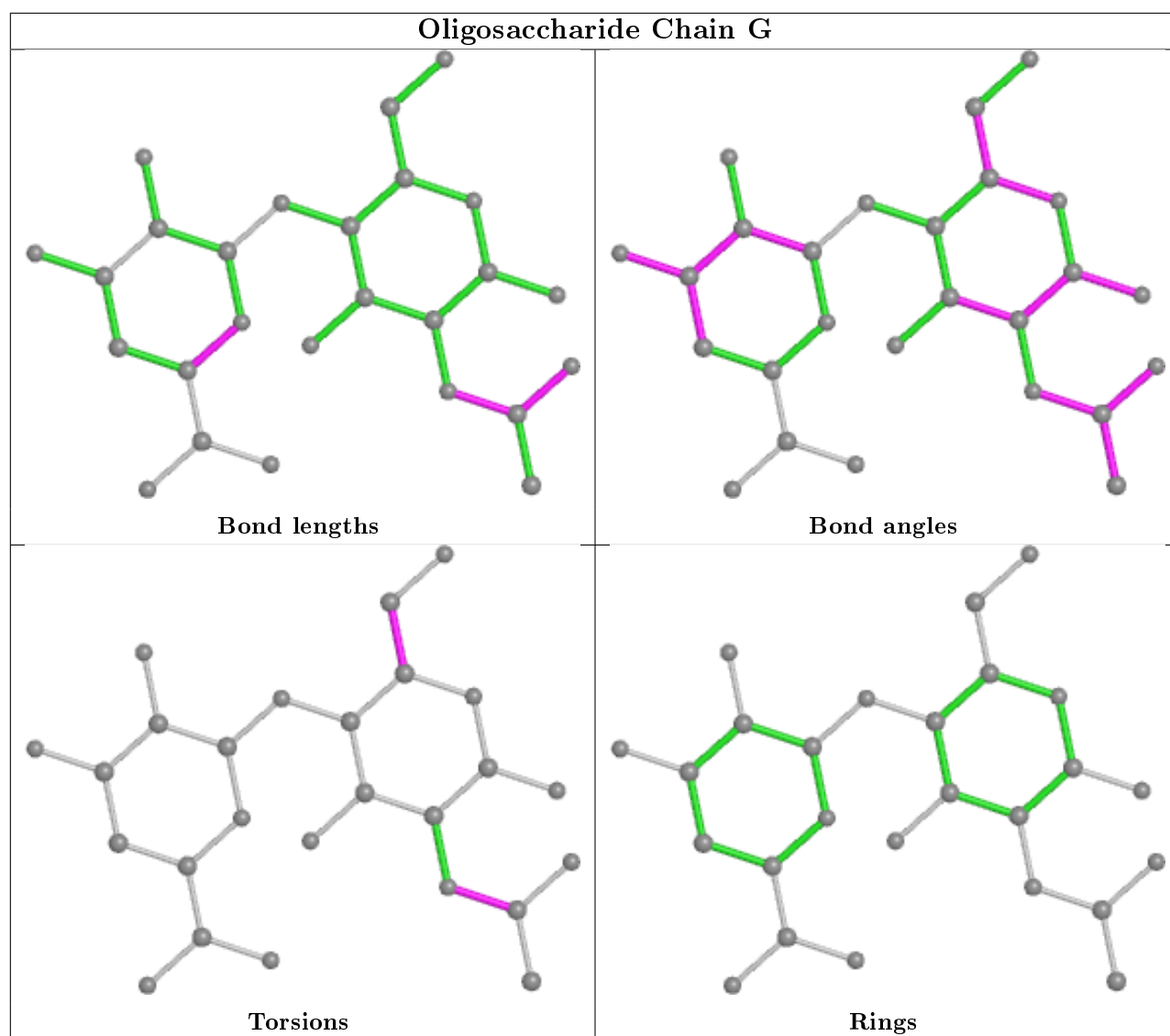
3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	4	RAM	1	0
3	H	1	NDG	1	0
3	I	2	GCD	2	0

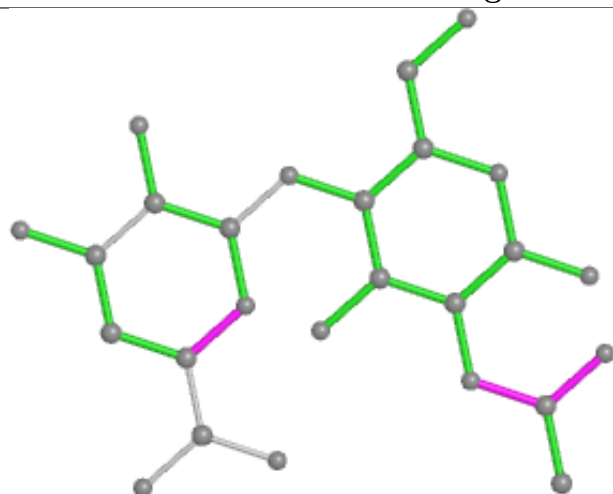
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.



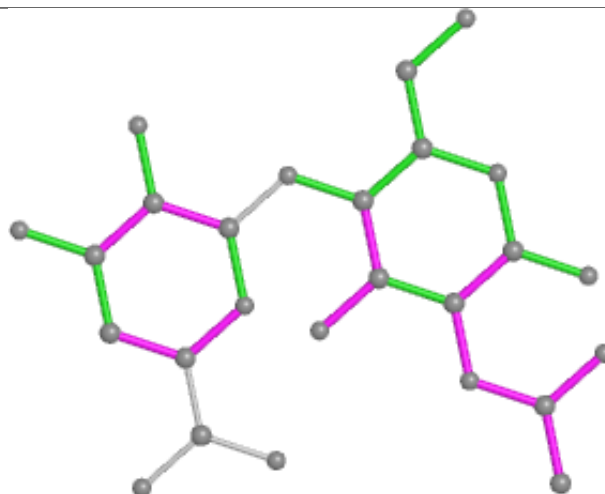




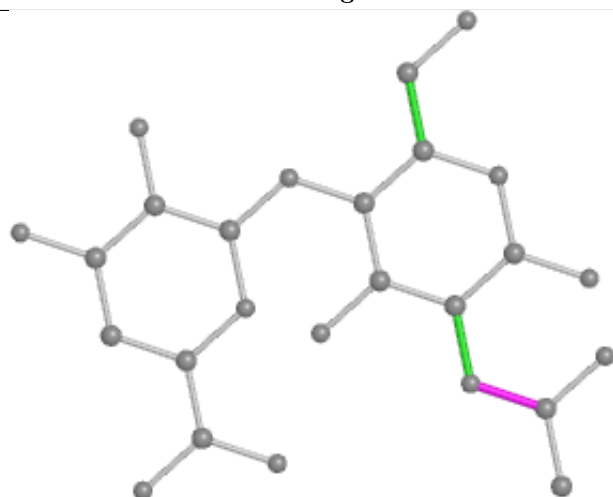
## Oligosaccharide Chain H



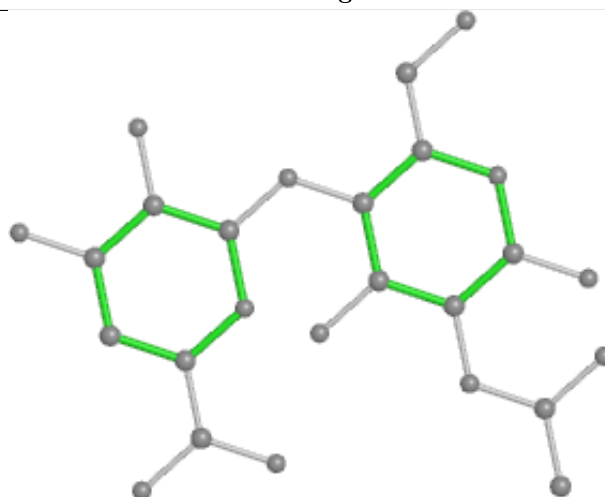
Bond lengths



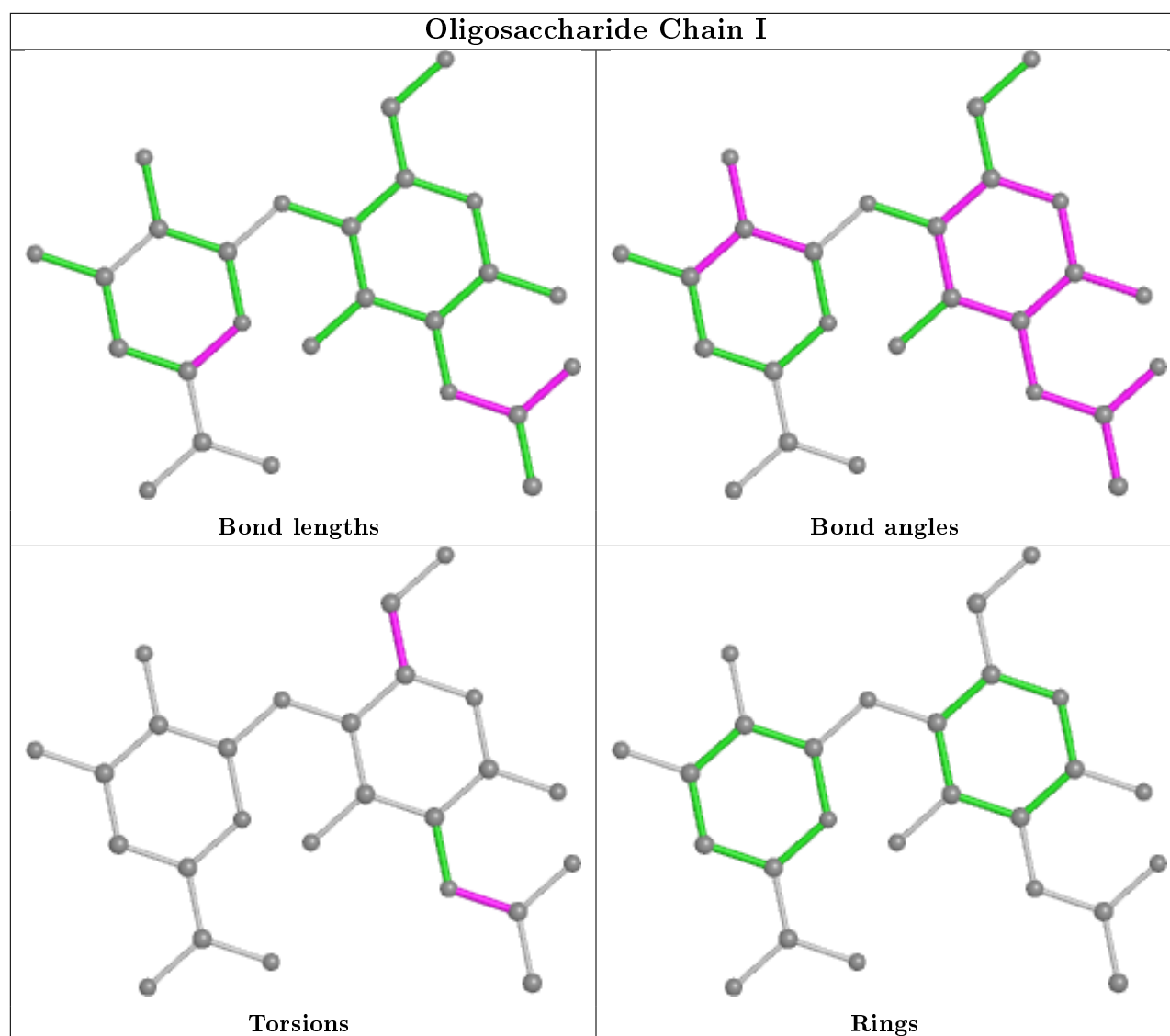
Bond angles



Torsions



Rings



## 5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 3 are monoatomic - leaving 9 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	$\# Z  > 2$
5	PO4	A	5	-	4,4,4	0.58	0	6,6,6	0.86	0
5	PO4	A	9	-	4,4,4	0.84	0	6,6,6	1.45	1 (16%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	PO4	C	4	-	4,4,4	0.81	0	6,6,6	1.30	1 (16%)
5	PO4	A	3	-	4,4,4	0.76	0	6,6,6	0.92	0
5	PO4	B	7	-	4,4,4	0.83	0	6,6,6	0.75	0
5	PO4	A	2	-	4,4,4	0.71	0	6,6,6	1.06	0
5	PO4	B	8	-	4,4,4	0.55	0	6,6,6	0.93	0
5	PO4	C	6	-	4,4,4	0.46	0	6,6,6	1.64	2 (33%)
5	PO4	B	1	-	4,4,4	1.02	0	6,6,6	1.63	2 (33%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	1	PO4	O4-P-O1	-2.70	101.00	110.89
5	A	9	PO4	O3-P-O2	2.49	115.96	107.97
5	C	6	PO4	O4-P-O3	2.40	115.68	107.97
5	C	4	PO4	O3-P-O2	2.33	115.45	107.97
5	B	1	PO4	O3-P-O1	2.30	119.30	110.89

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	6	PO4	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 [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	746/749 (99%)	0.03	7 (0%) 84 90	19, 24, 27, 31	0
1	B	746/749 (99%)	0.20	15 (2%) 65 75	19, 24, 27, 31	0
1	C	746/749 (99%)	0.20	16 (2%) 63 74	19, 24, 28, 34	0
All	All	2238/2247 (99%)	0.14	38 (1%) 70 78	19, 24, 28, 34	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	38	GLY	5.5
1	C	472	ASP	5.5
1	C	470	GLY	5.3
1	C	471	SER	4.7
1	C	473	HIS	4.6

### 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
1	PCA	A	26	8/9	0.80	0.23	28,29,29,31	0
1	PCA	B	26	8/9	0.82	0.16	30,31,32,33	0
1	PCA	C	26	8/9	0.85	0.17	29,29,30,30	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,



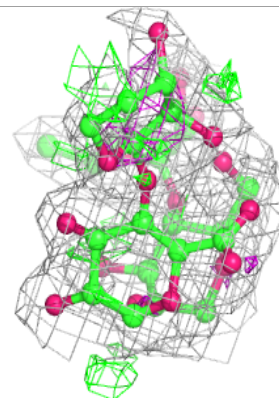
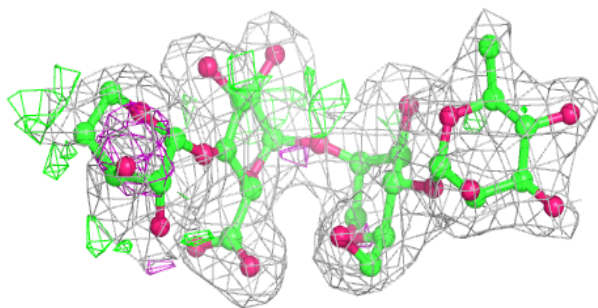
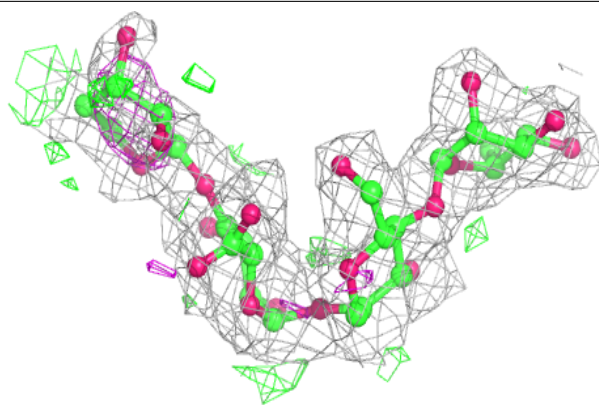
median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	XYS	D	3	8/10	0.69	0.38	45,47,47,47	0
2	XYS	F	3	8/10	0.72	0.38	49,49,50,50	0
2	XYS	E	3	8/10	0.80	0.34	48,48,49,49	0
3	GCD	I	2	11/12	0.83	0.21	34,34,37,37	0
3	NDG	I	1	15/15	0.87	0.25	32,35,38,38	0
2	GCU	E	2	12/13	0.90	0.16	39,40,42,45	0
2	GCU	F	2	12/13	0.90	0.15	33,35,37,44	0
3	NDG	G	1	15/15	0.91	0.18	26,30,31,33	0
3	GCD	H	2	11/12	0.92	0.14	21,24,26,27	0
3	GCD	G	2	11/12	0.92	0.14	24,30,31,33	0
2	RAM	E	4	10/11	0.93	0.17	28,29,30,30	0
2	GCU	D	2	12/13	0.93	0.14	33,35,36,42	0
2	MAN	D	1	11/12	0.94	0.13	24,27,31,35	0
3	NDG	H	1	15/15	0.94	0.12	23,26,28,28	0
2	MAN	F	1	11/12	0.95	0.13	23,27,29,32	0
2	MAN	E	1	11/12	0.95	0.12	26,27,29,36	0
2	RAM	D	4	10/11	0.95	0.15	23,27,28,30	0
2	RAM	F	4	10/11	0.97	0.17	25,26,28,28	0

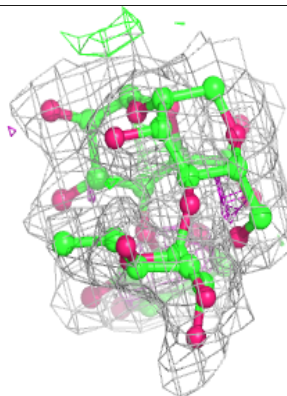
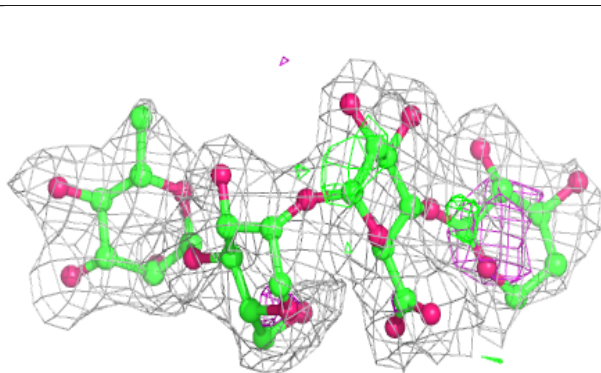
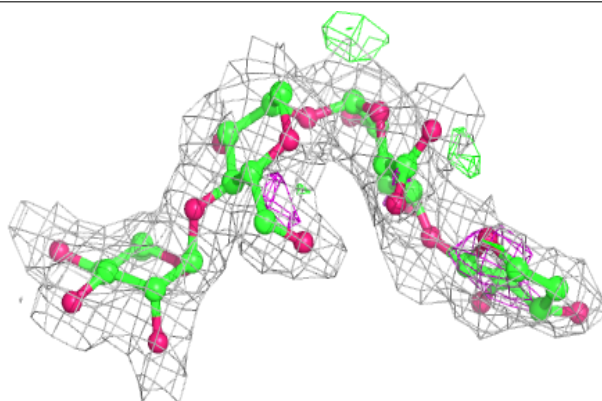
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

**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)

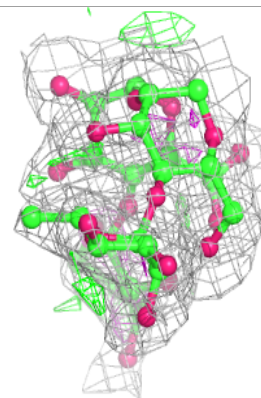
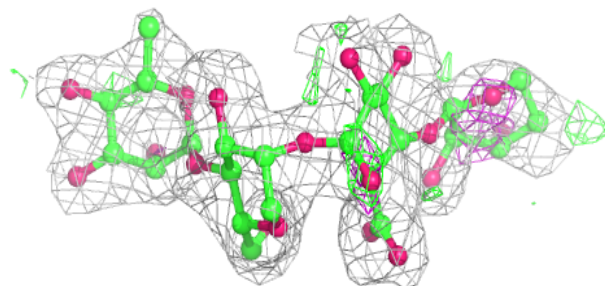
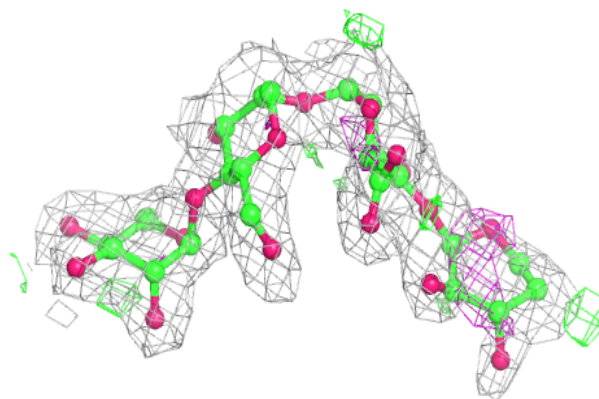
**Electron density around Chain E:**

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



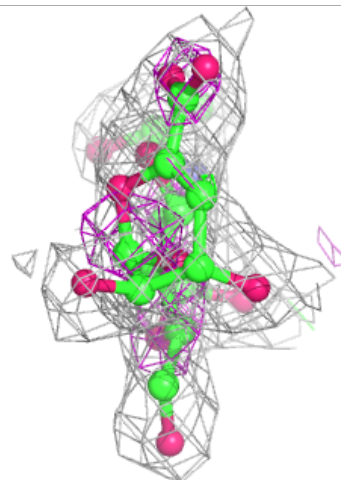
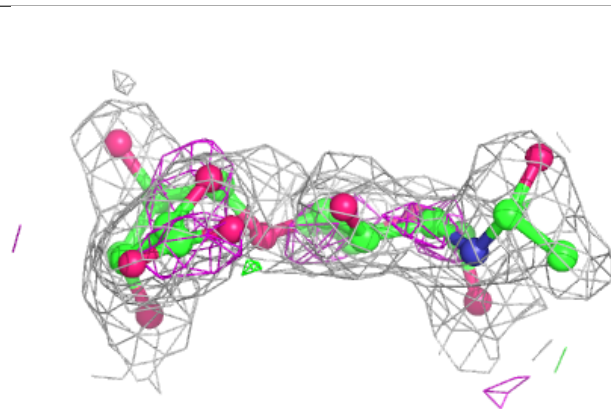
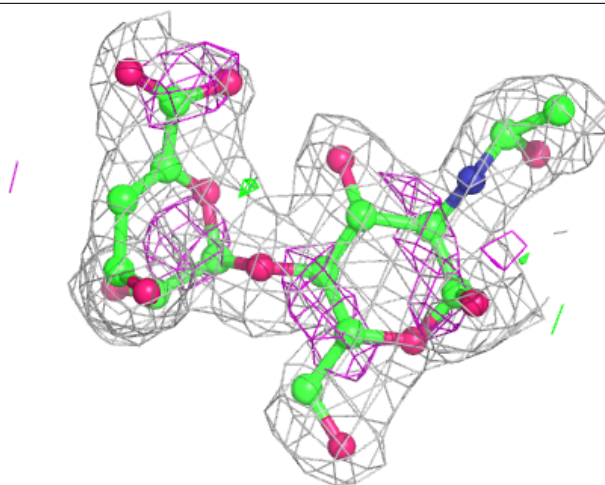
**Electron density around Chain F:**

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



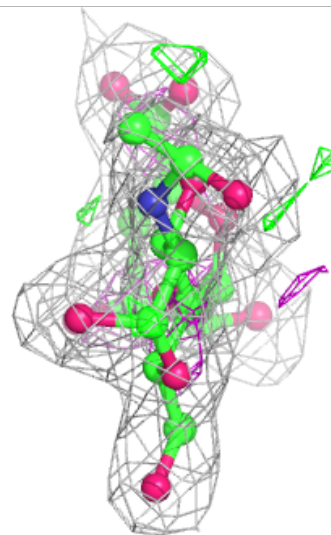
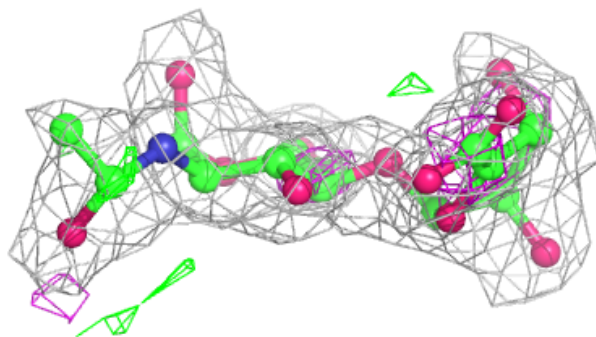
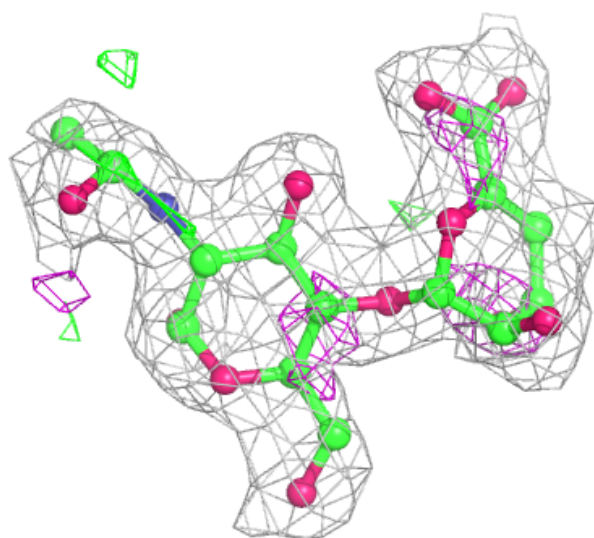
**Electron density around Chain G:**

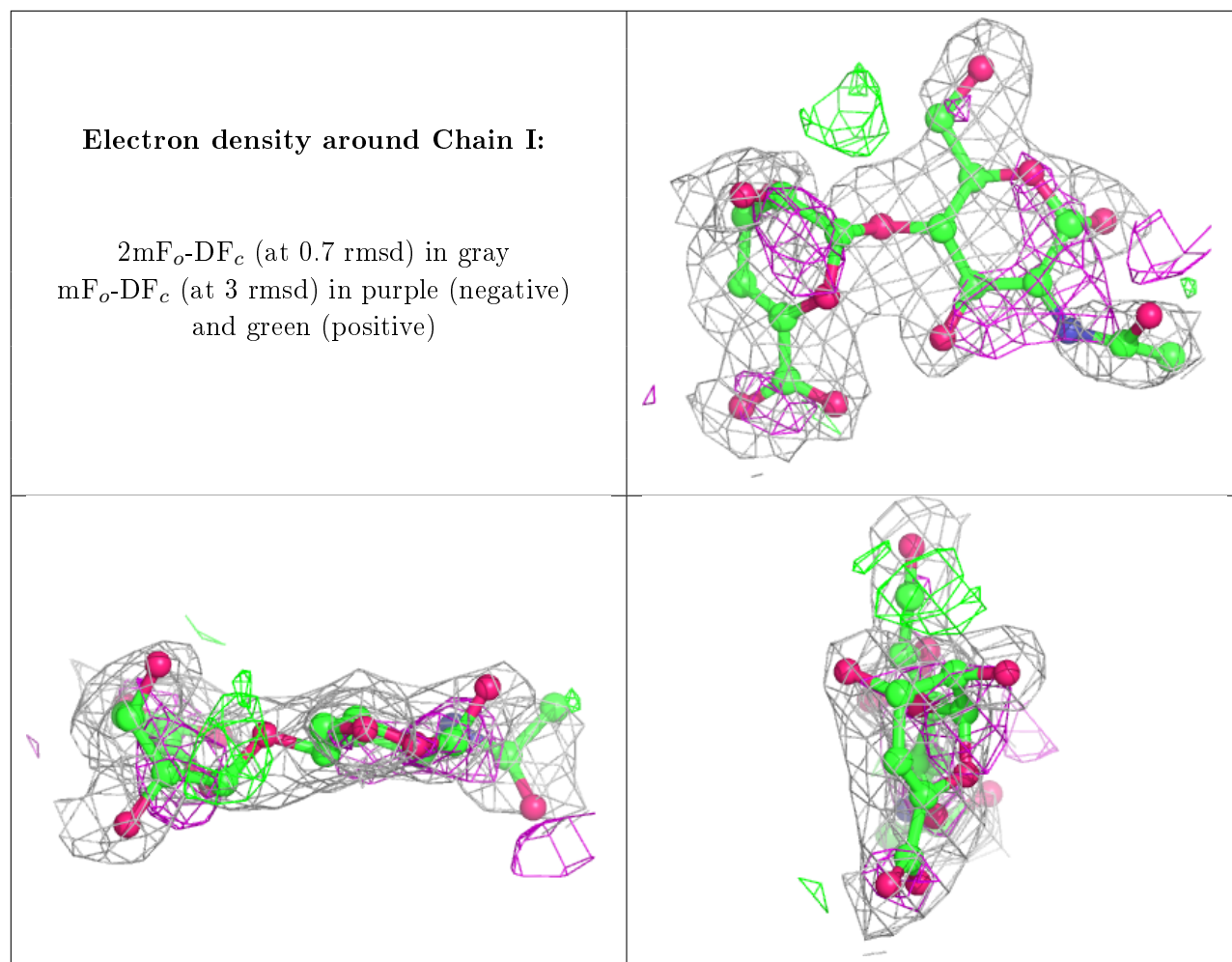
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**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)





## 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
5	PO4	C	6	5/5	0.80	0.33	57,57,59,59	0
5	PO4	A	5	5/5	0.87	0.26	52,53,54,54	0
5	PO4	B	7	5/5	0.92	0.21	52,52,53,54	0
5	PO4	A	2	5/5	0.93	0.20	44,45,45,45	0
5	PO4	B	8	5/5	0.94	0.28	48,49,49,50	0
5	PO4	A	3	5/5	0.95	0.22	41,41,42,43	0
5	PO4	C	4	5/5	0.97	0.19	31,31,31,31	0
5	PO4	A	9	5/5	0.98	0.09	26,26,27,27	0
5	PO4	B	1	5/5	0.98	0.10	27,27,28,29	0
4	ZN	A	1	1/1	0.99	0.04	27,27,27,27	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	ZN	B	2	1/1	1.00	0.04	26,26,26,26	0
4	ZN	C	3	1/1	1.00	0.02	25,25,25,25	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.