



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 7, 2020 – 03:23 PM BST

PDB ID : 6F2Y  
Title : Crystal structure of ectonucleotide phosphodiesterase/pyrophosphatase-3 (NPP3) in complex with Ap4A  
Authors : Dohler, C.; Zebisch, M.; Strater, N.  
Deposited on : 2017-11-27  
Resolution : 2.40 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

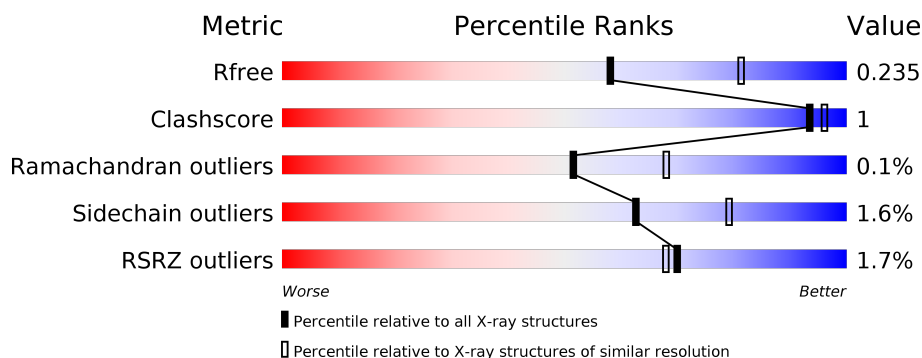
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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>2%</div> <div> <div></div> <div>90%</div> <div>6%</div> <div></div> </div> </div>
1	B	749	<div> <div>%</div> <div> <div></div> <div>90%</div> <div>6%</div> <div></div> </div> </div>
2	C	2	<div> <div></div> <div>100%</div> </div>
2	D	2	<div> <div></div> <div> <div>50%</div> <div>50%</div> </div> </div>
2	E	2	<div> <div></div> <div>100%</div> </div>
2	G	2	<div> <div></div> <div>100%</div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	3	 33% 67%

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 12376 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 Ectonucleotide pyrophosphatase/phosphodiesterase family member 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	720	Total	C	N	O	S	0	0	0
			5795	3712	976	1074	33			
1	B	719	Total	C	N	O	S	0	0	0
			5789	3708	974	1074	33			

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	136	ALA	-	expression tag	UNP P97675
A	137	GLU	-	expression tag	UNP P97675
A	138	THR	-	expression tag	UNP P97675
A	139	GLY	-	expression tag	UNP P97675
A	201	VAL	MET	variant	UNP P97675
A	206	ALA	THR	engineered mutation	UNP P97675
A	596	ASN	SER	variant	UNP P97675
A	597	ARG	GLY	variant	UNP P97675
A	876	GLY	-	expression tag	UNP P97675
A	877	THR	-	expression tag	UNP P97675
A	878	LYS	-	expression tag	UNP P97675
A	879	HIS	-	expression tag	UNP P97675
A	880	HIS	-	expression tag	UNP P97675
A	881	HIS	-	expression tag	UNP P97675
A	882	HIS	-	expression tag	UNP P97675
A	883	HIS	-	expression tag	UNP P97675
A	884	HIS	-	expression tag	UNP P97675
B	136	ALA	-	expression tag	UNP P97675
B	137	GLU	-	expression tag	UNP P97675
B	138	THR	-	expression tag	UNP P97675
B	139	GLY	-	expression tag	UNP P97675
B	201	VAL	MET	variant	UNP P97675
B	206	ALA	THR	engineered mutation	UNP P97675
B	596	ASN	SER	variant	UNP P97675

*Continued on next page...*

Continued from previous page...

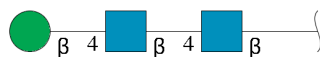
Chain	Residue	Modelled	Actual	Comment	Reference
B	597	ARG	GLY	variant	UNP P97675
B	876	GLY	-	expression tag	UNP P97675
B	877	THR	-	expression tag	UNP P97675
B	878	LYS	-	expression tag	UNP P97675
B	879	HIS	-	expression tag	UNP P97675
B	880	HIS	-	expression tag	UNP P97675
B	881	HIS	-	expression tag	UNP P97675
B	882	HIS	-	expression tag	UNP P97675
B	883	HIS	-	expression tag	UNP P97675
B	884	HIS	-	expression tag	UNP P97675

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



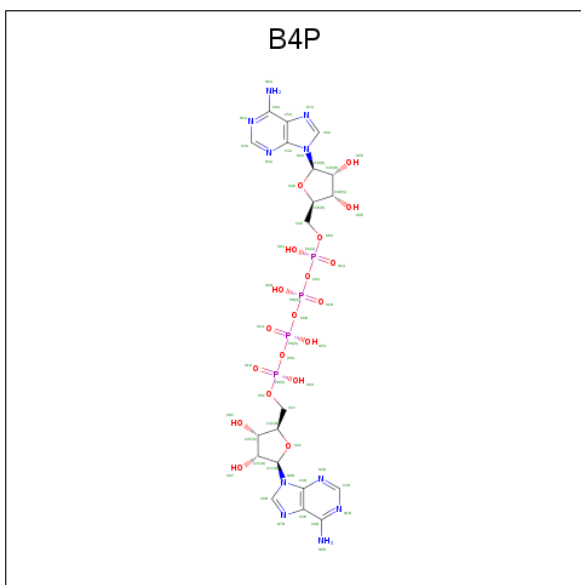
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	2	Total	C	N	O	0	0	0
			28	16	2	10			
2	D	2	Total	C	N	O	0	0	0
			28	16	2	10			
2	E	2	Total	C	N	O	0	0	0
			28	16	2	10			
2	G	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 3 is an oligosaccharide called beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



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

- Molecule 4 is BIS(ADENOSINE)-5'-TETRAPHOSPHATE (three-letter code: B4P) (formula: C<sub>20</sub>H<sub>28</sub>N<sub>10</sub>O<sub>19</sub>P<sub>4</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	0
			53	20	10	19	4	
4	B	1	Total	C	N	O	P	0
			53	20	10	19	4	

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	2	Total	Zn	0	0
			2	2		
5	A	2	Total	Zn	0	0
			2	2		

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total	Ca	0	0
			1	1		
6	A	1	Total	Ca	0	0
			1	1		

- Molecule 7 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
7	A	1	Total	C	N	O	0	0
			14	8	1	5		
7	A	1	Total	C	N	O	0	0
			14	8	1	5		
7	B	1	Total	C	N	O	0	0
			14	8	1	5		
7	B	1	Total	C	N	O	0	0
			14	8	1	5		
7	B	1	Total	C	N	O	0	0
			14	8	1	5		
7	B	1	Total	C	N	O	0	0
			14	8	1	5		

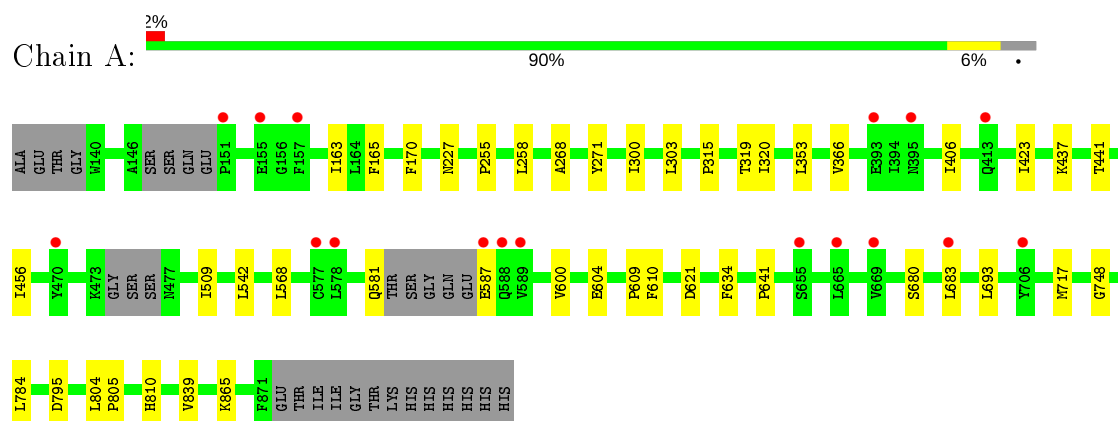
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	234	Total	O	0	0
			234	234		
8	B	197	Total	O	0	0
			197	197		

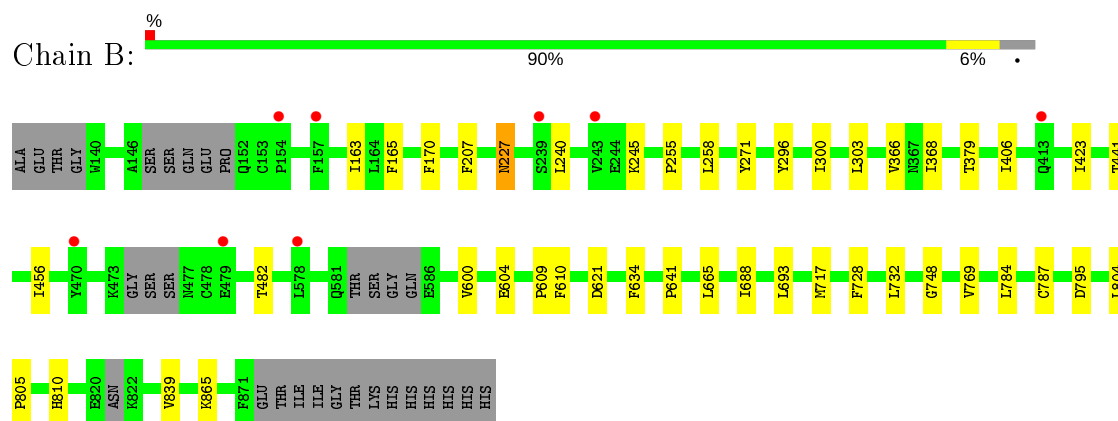
### 3 Residue-property plots

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.

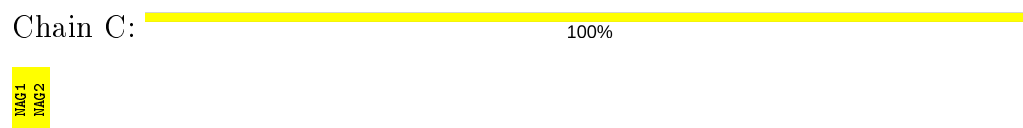
- Molecule 1: Ectonucleotide pyrophosphatase/phosphodiesterase family member 3



- Molecule 1: Ectonucleotide pyrophosphatase/phosphodiesterase family member 3



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



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



Chain D:  50% 50%

MAG1  
MAG2

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

Chain E:  100%


MAG1  
MAG2

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

Chain G:  100%

MAG1  
MAG2

- Molecule 3: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F:  33% 67%

MAG1  
MAG2  
E/A3

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.94Å 126.85Å 114.48Å 90.00° 97.54° 90.00°	Depositor
Resolution (Å)	48.24 – 2.40 48.24 – 2.40	Depositor EDS
% Data completeness (in resolution range)	98.7 (48.24-2.40) 98.7 (48.24-2.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.80 (at 2.39Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, $R_{free}$	0.204 , 0.217 0.226 , 0.235	Depositor DCC
$R_{free}$ test set	1615 reflections (1.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	45.6	Xtriage
Anisotropy	0.296	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 34.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	12376	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.91% 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: ZN, BMA, CA, B4P, 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	0.38	0/5968	0.57	0/8124
1	B	0.38	0/5960	0.58	0/8111
All	All	0.38	0/11928	0.58	0/16235

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	5795	0	5586	16	0
1	B	5789	0	5575	17	0
2	C	28	0	25	0	0
2	D	28	0	25	0	0
2	E	28	0	25	0	0
2	G	28	0	25	0	0
3	F	39	0	34	0	0
4	A	53	0	24	0	0
4	B	53	0	24	0	0
5	A	2	0	0	0	0
5	B	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	A	1	0	0	0	0
6	B	1	0	0	0	0
7	A	28	0	26	0	0
7	B	70	0	65	0	0
8	A	234	0	0	0	0
8	B	197	0	0	0	0
All	All	12376	0	11434	33	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

All (33) 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:255:PRO:HD2	1:B:258:LEU:HD12	1.85	0.58
1:A:255:PRO:HD2	1:A:258:LEU:HD12	1.85	0.57
1:B:634:PHE:HA	1:B:641:PRO:HA	1.89	0.55
1:A:634:PHE:HA	1:A:641:PRO:HA	1.89	0.54
1:B:163:ILE:HD12	1:B:366:VAL:HG11	1.91	0.52
1:A:163:ILE:HD12	1:A:366:VAL:HG11	1.91	0.51
1:B:688:ILE:HD11	1:B:769:VAL:HG13	1.93	0.49
1:B:379:THR:HG21	1:B:482:THR:HG22	1.94	0.49
1:A:717:MET:HA	1:A:810:HIS:NE2	2.28	0.48
1:A:839:VAL:HB	1:A:865:LYS:HA	1.95	0.48
1:B:839:VAL:HB	1:B:865:LYS:HA	1.95	0.48
1:A:784:LEU:HD12	1:A:804:LEU:HD23	1.96	0.47
1:B:784:LEU:HD12	1:B:804:LEU:HD23	1.97	0.47
1:A:268:ALA:HB2	1:A:315:PRO:HG2	1.98	0.46
1:B:296:TYR:O	1:B:300:ILE:HD12	2.15	0.46
1:B:717:MET:HA	1:B:810:HIS:NE2	2.30	0.46
1:B:240:LEU:HA	1:B:245:LYS:HG3	1.99	0.45
1:A:300:ILE:HD12	1:A:353:LEU:HB2	1.98	0.45
1:A:406:ILE:HG21	1:A:423:ILE:HG21	1.99	0.44
1:B:441:THR:HG21	1:B:456:ILE:HG22	1.99	0.44
1:A:441:THR:HG21	1:A:456:ILE:HG22	1.99	0.44
1:B:609:PRO:HB3	1:B:795:ASP:HA	2.00	0.44
1:A:609:PRO:HB3	1:A:795:ASP:HA	2.01	0.43
1:B:406:ILE:HG21	1:B:423:ILE:HG21	2.00	0.42
1:B:207:PHE:HB2	1:B:227:ASN:HD22	1.85	0.42
1:A:693:LEU:HD21	1:A:748:GLY:HA2	2.03	0.41
1:B:600:VAL:O	1:B:604:GLU:HG3	2.20	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:600:VAL:O	1:A:604:GLU:HG3	2.19	0.41
1:A:509:ILE:HD13	1:A:542:LEU:HD11	2.03	0.41
1:B:693:LEU:HD21	1:B:748:GLY:HA2	2.03	0.41
1:A:581:GLN:HE21	1:A:587:GLU:HB2	1.87	0.41
1:A:680:SER:HA	1:A:683:LEU:HD12	2.03	0.40
1:B:728:PHE:HA	1:B:732:LEU:HB2	2.03	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

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

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	712/749 (95%)	684 (96%)	27 (4%)	1 (0%)	51	68
1	B	709/749 (95%)	682 (96%)	26 (4%)	1 (0%)	51	68
All	All	1421/1498 (95%)	1366 (96%)	53 (4%)	2 (0%)	51	68

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	271	TYR
1	B	271	TYR

### 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	644/668 (96%)	633 (98%)	11 (2%)	60	78
1	B	643/668 (96%)	633 (98%)	10 (2%)	62	79
All	All	1287/1336 (96%)	1266 (98%)	21 (2%)	62	79

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	165	PHE
1	A	170	PHE
1	A	227	ASN
1	A	303	LEU
1	A	319	THR
1	A	320	ILE
1	A	437	LYS
1	A	568	LEU
1	A	610	PHE
1	A	621	ASP
1	A	805	PRO
1	B	165	PHE
1	B	170	PHE
1	B	227	ASN
1	B	303	LEU
1	B	368	ILE
1	B	610	PHE
1	B	621	ASP
1	B	665	LEU
1	B	787	CYS
1	B	805	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	194	HIS
1	A	227	ASN
1	A	581	GLN
1	B	227	ASN
1	B	590	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

11 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	NAG	C	1	1,2	14,14,15	0.28	0	17,19,21	0.85	1 (5%)
2	NAG	C	2	2	14,14,15	0.29	0	17,19,21	0.77	1 (5%)
2	NAG	D	1	1,2	14,14,15	0.29	0	17,19,21	0.59	0
2	NAG	D	2	2	14,14,15	0.30	0	17,19,21	0.60	1 (5%)
2	NAG	E	1	1,2	14,14,15	0.30	0	17,19,21	0.75	0
2	NAG	E	2	2	14,14,15	0.30	0	17,19,21	0.56	0
3	NAG	F	1	1,3	14,14,15	0.29	0	17,19,21	0.86	1 (5%)
3	NAG	F	2	3	14,14,15	0.28	0	17,19,21	0.75	1 (5%)
3	BMA	F	3	3	11,11,12	0.32	0	15,15,17	0.52	0
2	NAG	G	1	1,2	14,14,15	0.28	0	17,19,21	0.75	0
2	NAG	G	2	2	14,14,15	0.30	0	17,19,21	0.57	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
2	NAG	C	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	C	2	2	-	0/6/23/26	0/1/1/1
2	NAG	D	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	D	2	2	-	0/6/23/26	0/1/1/1
2	NAG	E	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	E	2	2	-	0/6/23/26	0/1/1/1
3	NAG	F	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	F	2	3	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	BMA	F	3	3	-	0/2/19/22	0/1/1/1
2	NAG	G	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	G	2	2	-	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(°)
2	C	2	NAG	C1-O5-C5	2.82	116.02	112.19
3	F	1	NAG	O5-C1-C2	-2.82	106.84	111.29
2	C	1	NAG	O5-C1-C2	-2.81	106.86	111.29
3	F	2	NAG	C1-O5-C5	2.76	115.93	112.19
2	D	2	NAG	C1-O5-C5	2.09	115.02	112.19

There are no chirality outliers.

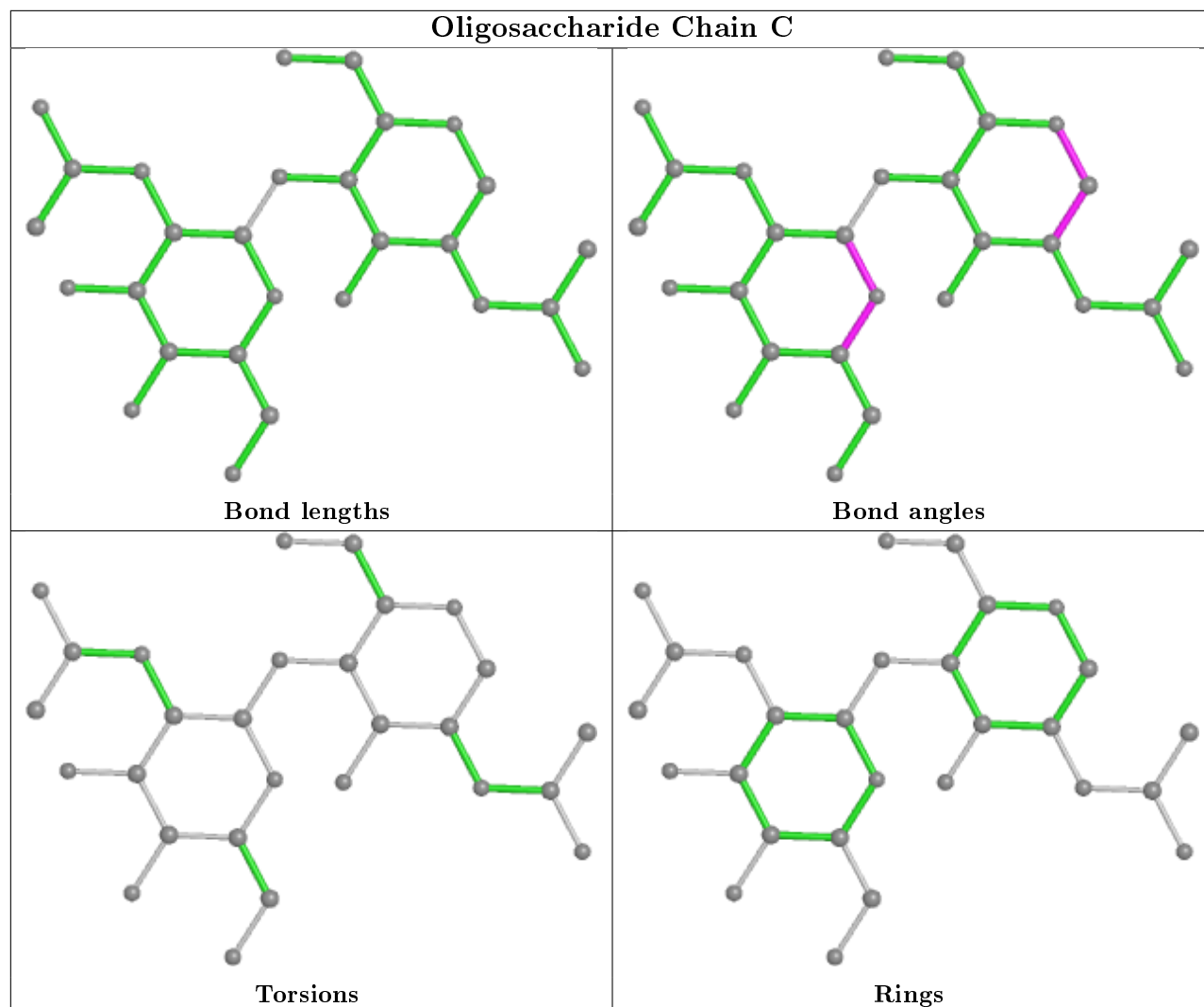
There are no torsion outliers.

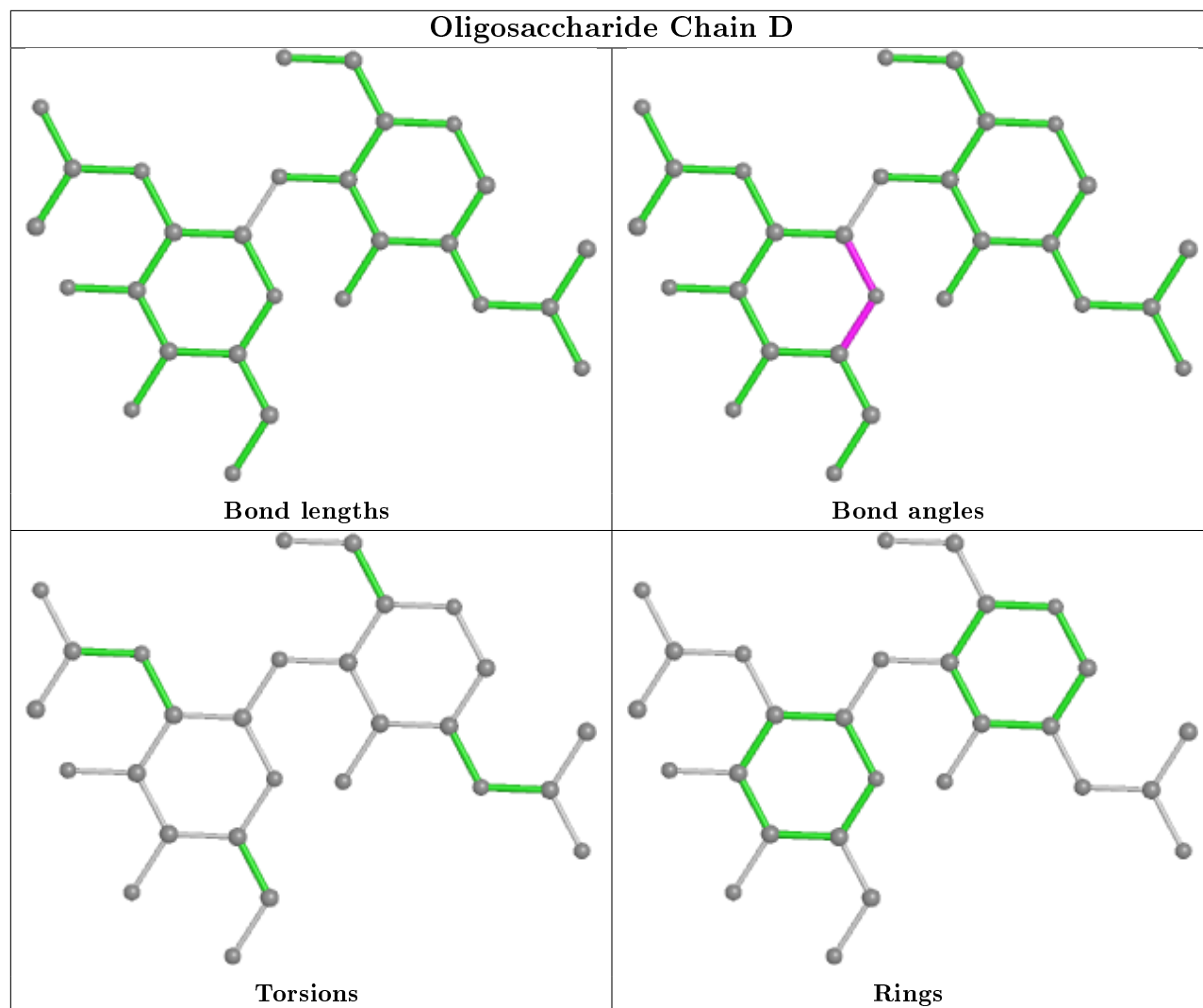
There are no ring outliers.

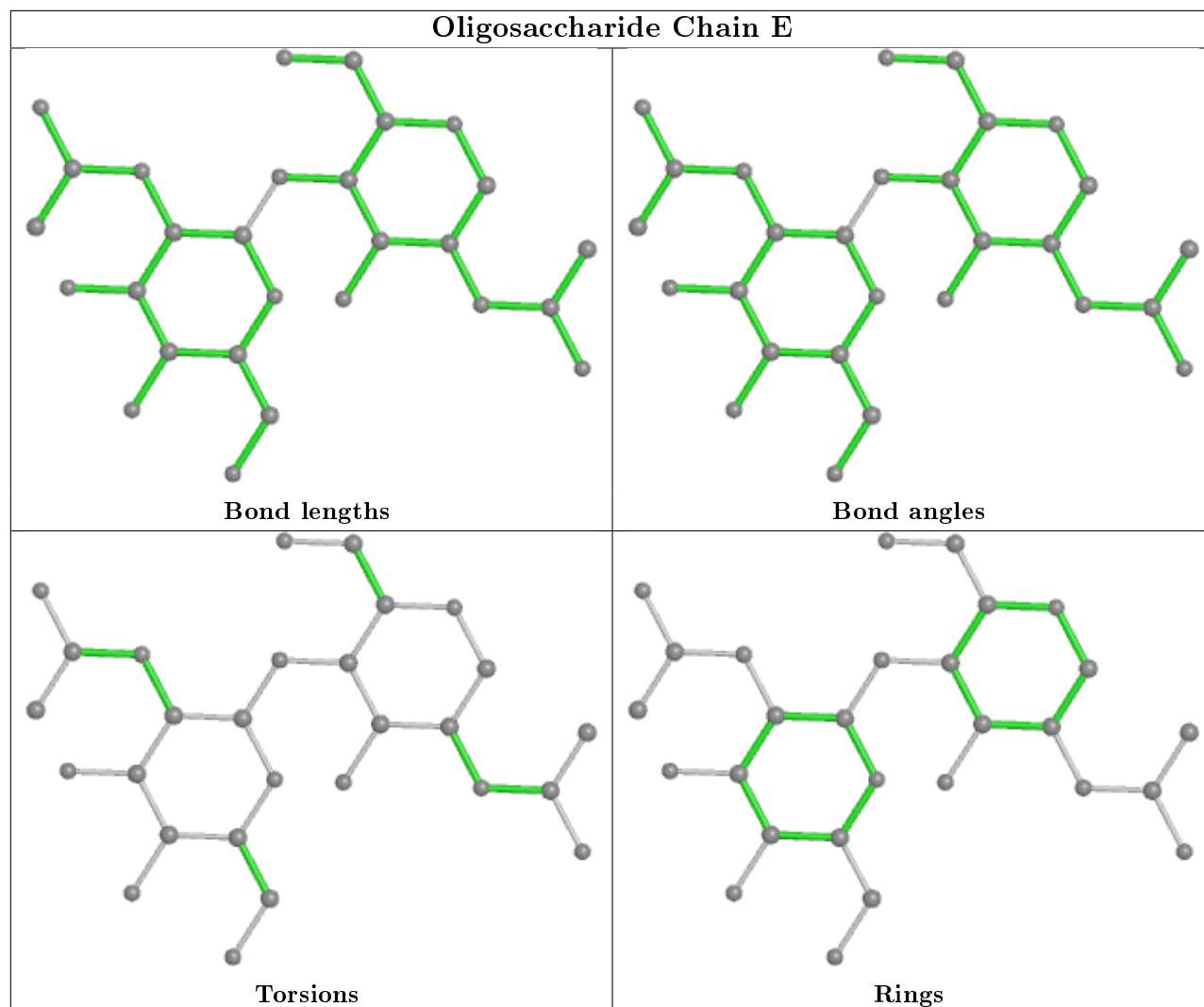
No monomer is involved in short contacts.

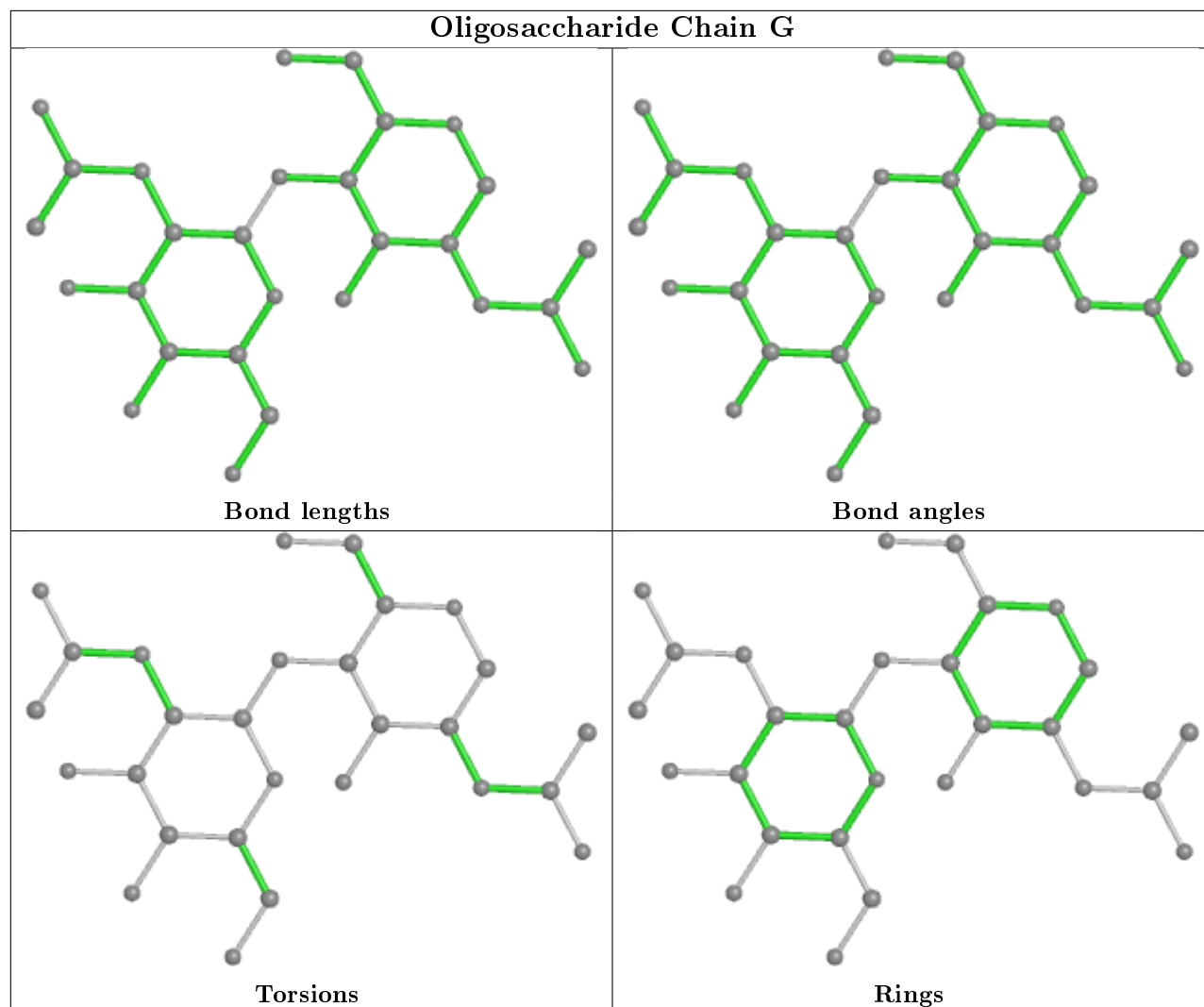
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

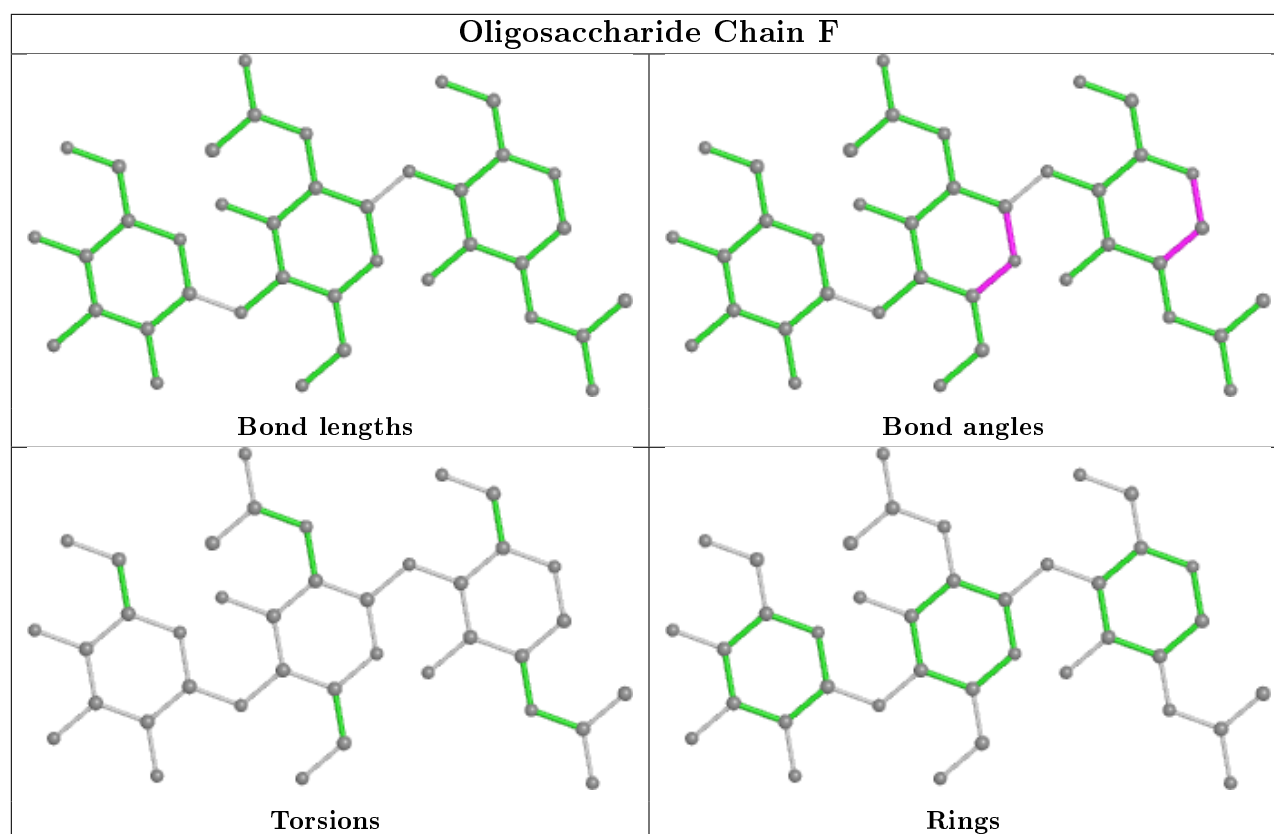












## 5.6 Ligand geometry [i](#)

Of 15 ligands modelled in this entry, 6 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	$\# Z  > 2$
7	NAG	B	1013	1	14,14,15	0.31	0	17,19,21	0.52	0
7	NAG	A	1012	1	14,14,15	0.29	0	17,19,21	0.62	1 (5%)
7	NAG	B	1008	1	14,14,15	0.29	0	17,19,21	0.65	1 (5%)
7	NAG	A	1007	1	14,14,15	0.28	0	17,19,21	0.72	1 (5%)
4	B4P	B	1001	5	46,58,58	0.60	0	49,91,91	0.78	3 (6%)
7	NAG	B	1012	1	14,14,15	0.31	0	17,19,21	0.59	0
4	B4P	A	1001	5	46,58,58	0.62	0	49,91,91	0.77	2 (4%)
7	NAG	B	1009	1	14,14,15	0.31	0	17,19,21	0.67	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
7	NAG	B	1014	1	14,14,15	0.29	0	17,19,21	0.50	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
7	NAG	B	1013	1	-	0/6/23/26	0/1/1/1
7	NAG	A	1012	1	-	0/6/23/26	0/1/1/1
7	NAG	B	1008	1	-	0/6/23/26	0/1/1/1
7	NAG	A	1007	1	-	0/6/23/26	0/1/1/1
4	B4P	B	1001	5	-	8/30/70/70	0/6/6/6
7	NAG	B	1012	1	-	0/6/23/26	0/1/1/1
4	B4P	A	1001	5	-	12/30/70/70	0/6/6/6
7	NAG	B	1009	1	-	0/6/23/26	0/1/1/1
7	NAG	B	1014	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	B	1009	NAG	C1-O5-C5	2.39	115.42	112.19
7	A	1007	NAG	C1-O5-C5	2.38	115.42	112.19
4	A	1001	B4P	C5A-C6A-N6A	2.26	123.79	120.35
4	B	1001	B4P	C5A-C6A-N6A	2.23	123.74	120.35
4	B	1001	B4P	C5B-C6B-N6B	2.23	123.74	120.35
4	A	1001	B4P	C5B-C6B-N6B	2.20	123.69	120.35
7	B	1008	NAG	C1-O5-C5	2.11	115.05	112.19
4	B	1001	B4P	PA-O5E-C5E	2.11	134.06	121.68
7	A	1012	NAG	C1-O5-C5	2.07	115.00	112.19

There are no chirality outliers.

All (20) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	1001	B4P	C5E-O5E-PA-O1A
4	B	1001	B4P	C5E-O5E-PA-O3A
4	B	1001	B4P	C5F-O5F-PD-O3G
4	A	1001	B4P	C5E-O5E-PA-O1A
4	A	1001	B4P	C5F-O5F-PD-O3G

*Continued on next page...*

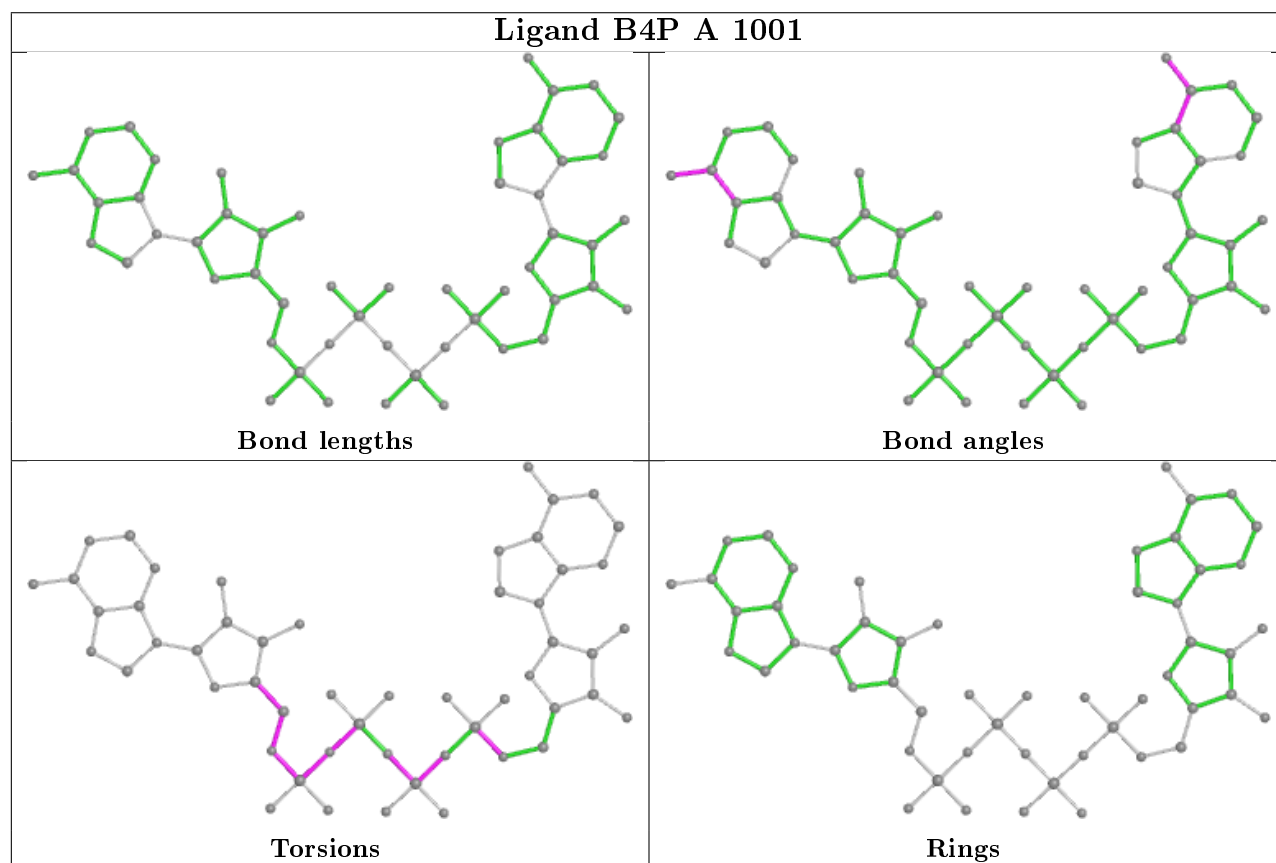
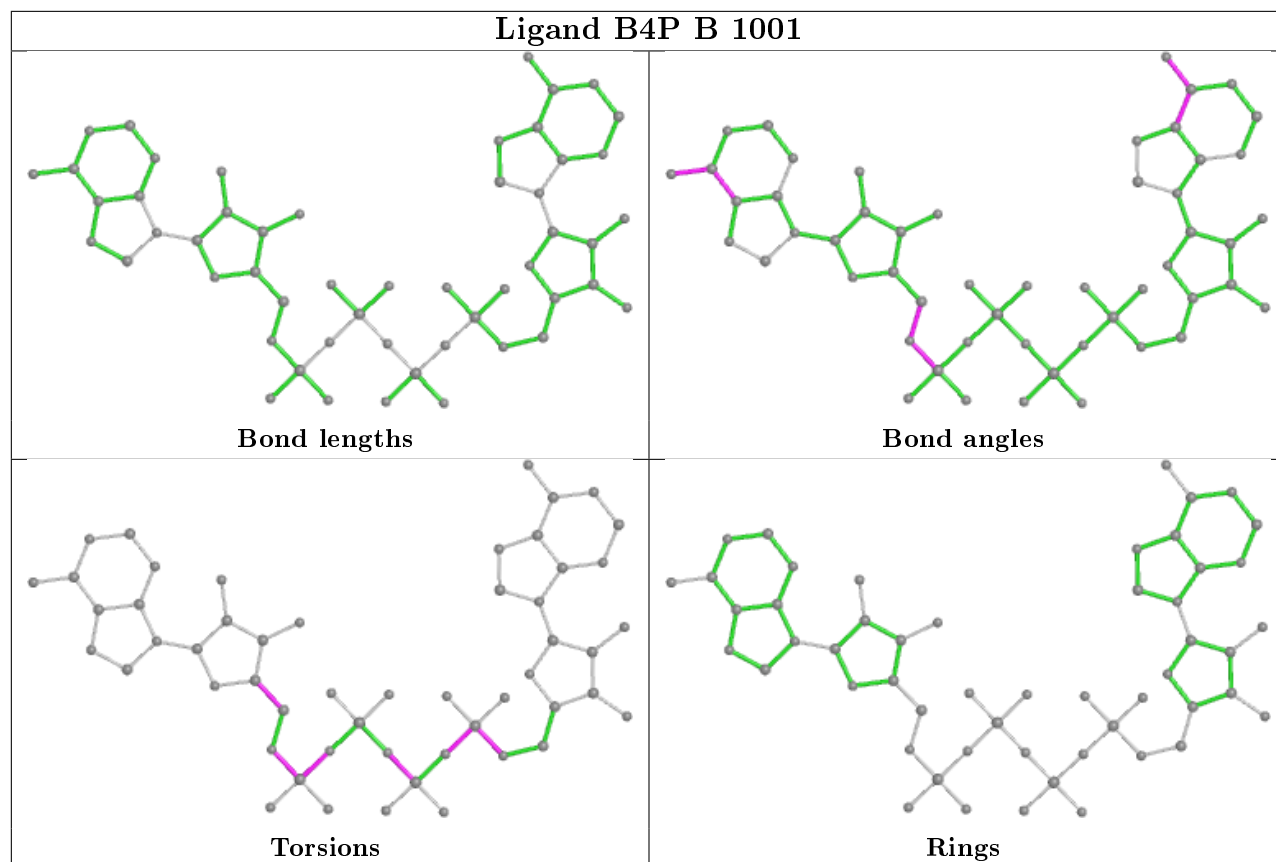
*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
4	B	1001	B4P	PB-O3B-PG-O1G
4	B	1001	B4P	PG-O3G-PD-O1D
4	A	1001	B4P	PB-O3A-PA-O1A
4	B	1001	B4P	PB-O3A-PA-O5E
4	A	1001	B4P	PB-O3A-PA-O5E
4	A	1001	B4P	C5E-O5E-PA-O3A
4	B	1001	B4P	C5F-O5F-PD-O2D
4	A	1001	B4P	C5F-O5F-PD-O1D
4	A	1001	B4P	C5F-O5F-PD-O2D
4	A	1001	B4P	C3E-C4E-C5E-O5E
4	B	1001	B4P	C3E-C4E-C5E-O5E
4	A	1001	B4P	PA-O3A-PB-O2B
4	A	1001	B4P	PB-O3B-PG-O2G
4	A	1001	B4P	PD-O3G-PG-O2G
4	A	1001	B4P	C4E-C5E-O5E-PA

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.





## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	720/749 (96%)	-0.03	17 (2%) 59 57	33, 53, 89, 138	1 (0%)
1	B	719/749 (95%)	-0.06	8 (1%) 80 79	36, 54, 83, 122	1 (0%)
All	All	1439/1498 (96%)	-0.04	25 (1%) 70 68	33, 53, 86, 138	2 (0%)

All (25) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	578	LEU	4.3
1	B	243	VAL	4.1
1	A	589	VAL	3.4
1	B	157	PHE	3.4
1	A	395	ASN	3.4
1	A	665	LEU	3.3
1	A	470	TYR	3.3
1	B	470	TYR	3.3
1	A	683	LEU	3.2
1	B	479	GLU	3.1
1	A	587	GLU	3.0
1	B	578	LEU	3.0
1	A	151	PRO	2.9
1	A	655	SER	2.8
1	B	154	PRO	2.7
1	A	413	GLN	2.5
1	A	669	VAL	2.4
1	B	239	SER	2.3
1	A	155	GLU	2.3
1	A	706	TYR	2.2
1	A	577	CYS	2.2
1	A	588	GLN	2.1
1	A	393	GLU	2.1
1	A	157	PHE	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	413	GLN	2.1

## 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](#)

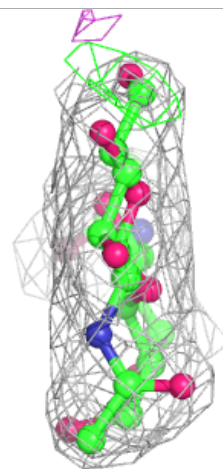
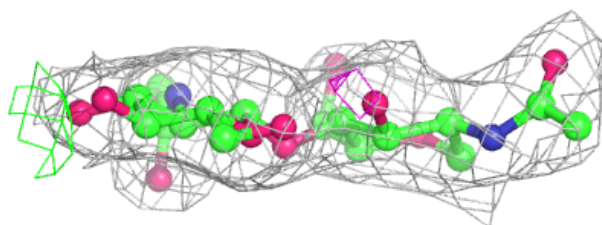
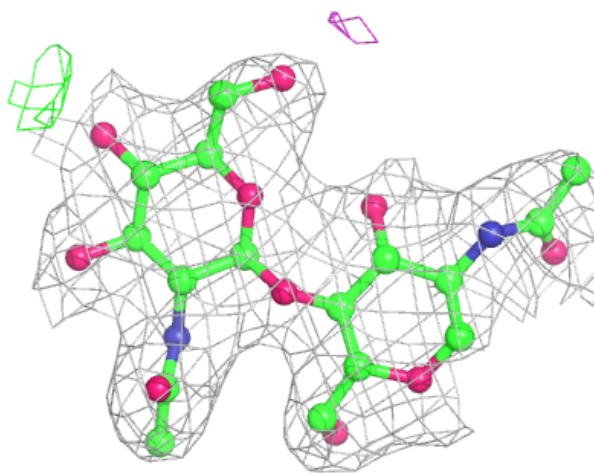
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	NAG	D	2	14/15	0.88	0.21	69,69,71,71	0
2	NAG	D	1	14/15	0.88	0.17	63,64,67,67	0
2	NAG	G	2	14/15	0.89	0.15	60,62,65,65	0
3	BMA	F	3	11/12	0.90	0.13	86,88,88,88	0
2	NAG	C	1	14/15	0.92	0.17	59,61,62,64	0
2	NAG	C	2	14/15	0.92	0.21	60,64,66,66	0
2	NAG	E	2	14/15	0.92	0.12	55,58,59,60	0
3	NAG	F	1	14/15	0.93	0.15	75,77,78,80	0
3	NAG	F	2	14/15	0.94	0.18	81,82,84,85	0
2	NAG	G	1	14/15	0.95	0.19	49,52,55,58	0
2	NAG	E	1	14/15	0.96	0.15	43,45,47,51	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.

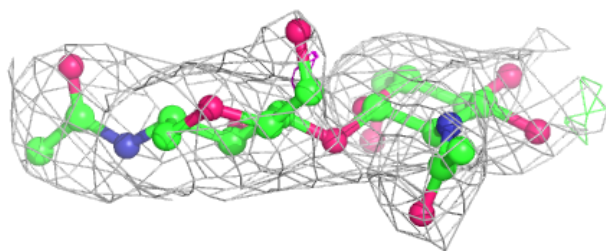
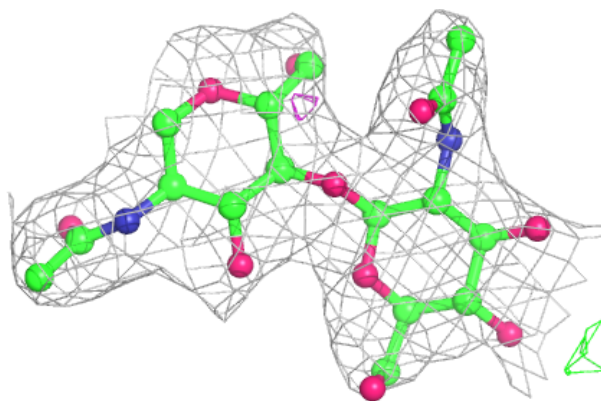
**Electron density around Chain C:**

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



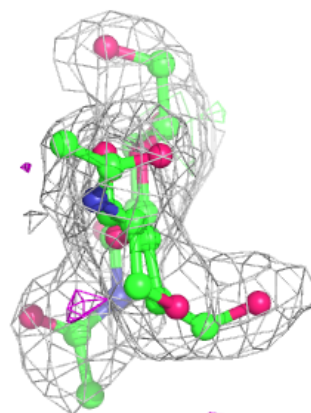
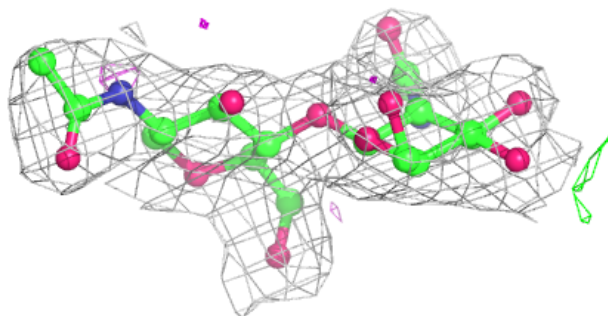
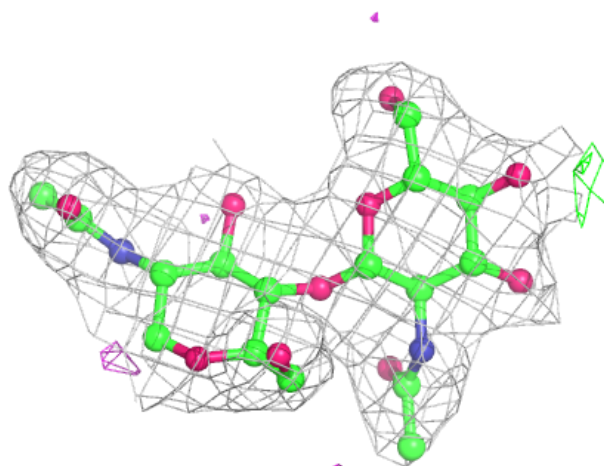
**Electron density around Chain D:**

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



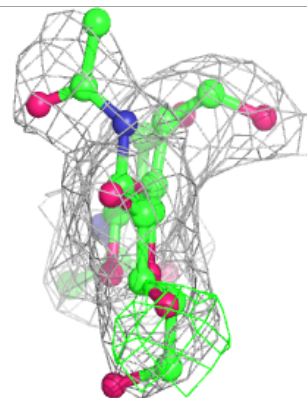
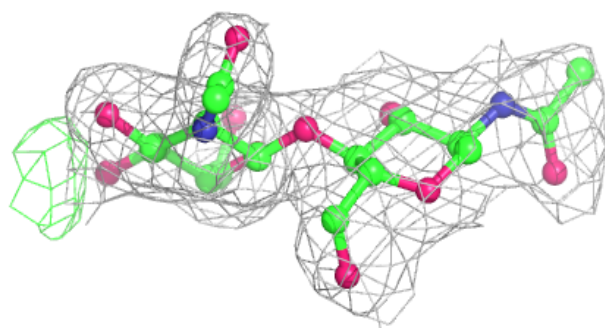
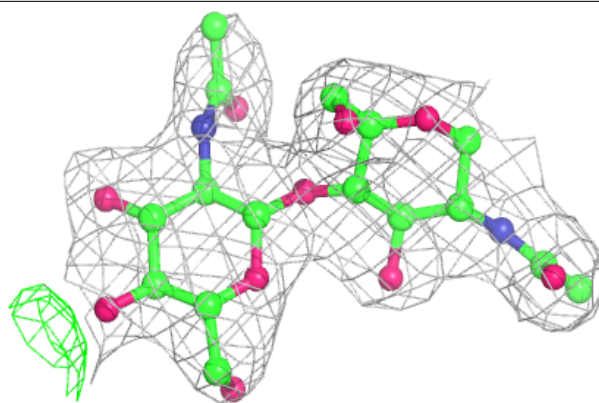
**Electron density around Chain E:**

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

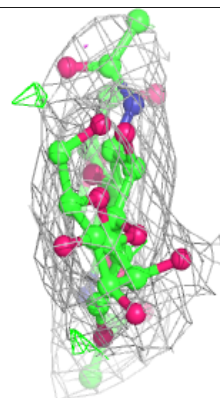
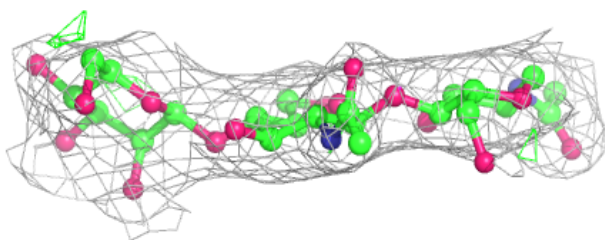
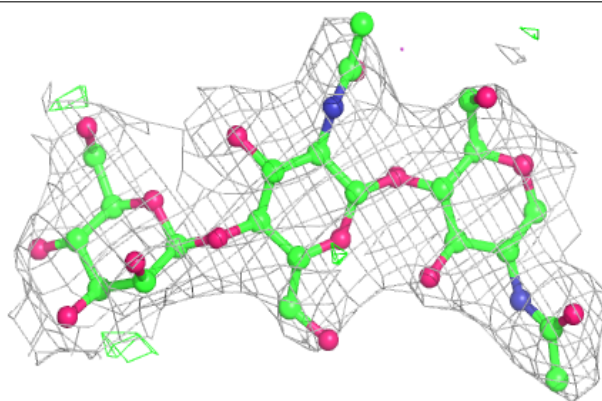


**Electron density around Chain G:**

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

**Electron density around Chain 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

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.

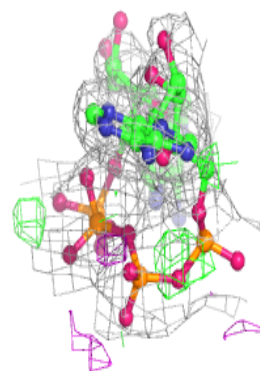
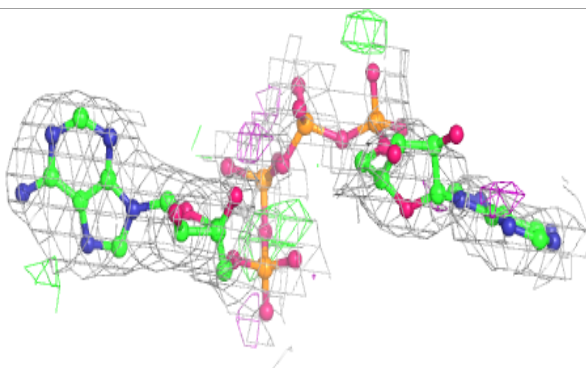
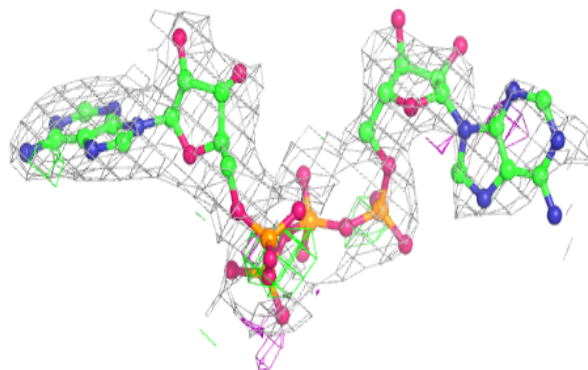
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
7	NAG	B	1013	14/15	0.71	0.31	96,97,99,99	0
7	NAG	B	1012	14/15	0.79	0.20	89,90,92,92	0
7	NAG	B	1014	14/15	0.85	0.18	81,81,82,82	0
7	NAG	A	1012	14/15	0.86	0.18	89,90,91,91	0
7	NAG	B	1008	14/15	0.86	0.18	67,70,73,73	0
7	NAG	B	1009	14/15	0.89	0.21	74,76,77,77	0
4	B4P	B	1001	53/53	0.90	0.20	58,88,112,112	0
7	NAG	A	1007	14/15	0.91	0.16	59,61,63,64	0
4	B4P	A	1001	53/53	0.93	0.18	49,69,81,82	0
6	CA	A	1004	1/1	0.94	0.12	51,51,51,51	0
5	ZN	B	1002	1/1	0.98	0.08	67,67,67,67	0
5	ZN	A	1002	1/1	0.99	0.07	55,55,55,55	0
5	ZN	B	1003	1/1	0.99	0.10	44,44,44,44	0
6	CA	B	1004	1/1	0.99	0.12	38,38,38,38	0
5	ZN	A	1003	1/1	1.00	0.13	44,44,44,44	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

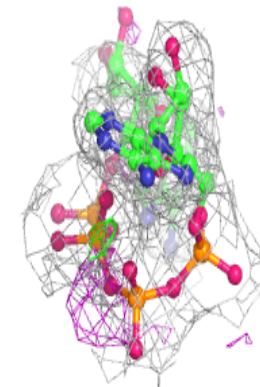
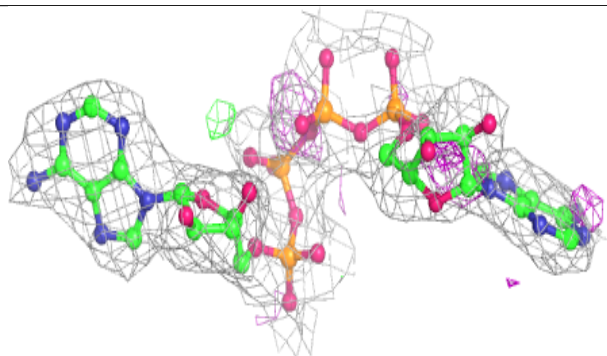
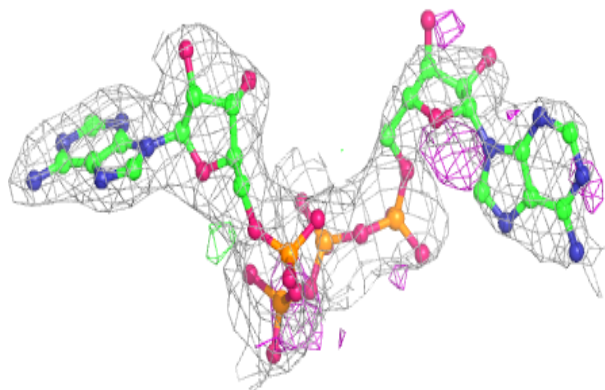


**Electron density around B4P B 1001:**

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

**Electron density around B4P A 1001:**

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



## 6.5 Other polymers

There are no such residues in this entry.