



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 17, 2024 – 06:18 AM EST

PDB ID : 4NAQ  
Title : Crystal structure of porcine aminopeptidase-N complexed with poly-alanine  
Authors : Chen, L.; Lin, Y.L.; Peng, G.; Li, F.  
Deposited on : 2013-10-22  
Resolution : 2.10 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 types of validation reports are described at

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

---

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

MolProbity : 4.02b-467  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.20.1  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

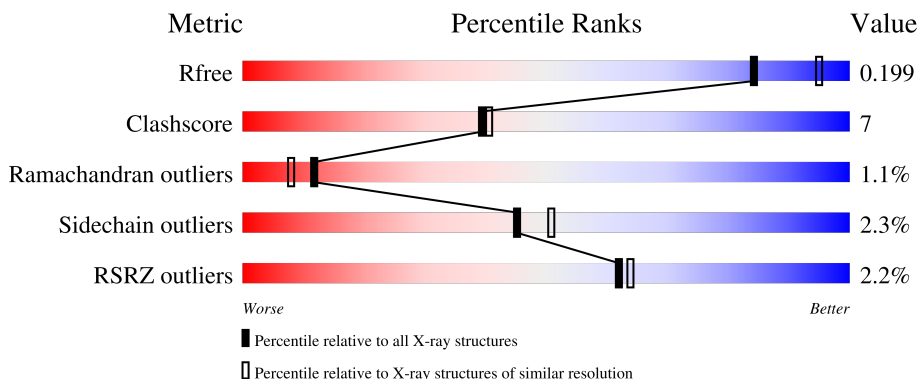
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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.



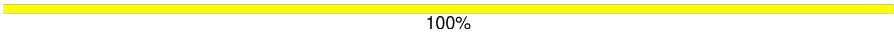

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	164625	6234 (2.10-2.10)
Clashscore	180529	6893 (2.10-2.10)
Ramachandran outliers	177936	6839 (2.10-2.10)
Sidechain outliers	177891	6840 (2.10-2.10)
RSRZ outliers	164620	6234 (2.10-2.10)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	907	
2	B	7	
3	C	2	
3	D	2	
3	E	2	

*Continued on next page...*

*Continued from previous page...*

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

## 2 Entry composition i

There are 7 unique types of molecules in this entry. The entry contains 8330 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 Aminopeptidase N.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	901	7240	4621	1210	1379	30	0	1	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	107	PHE	LEU	conflict	UNP P15145
A	384	GLN	GLU	conflict	UNP P15145
A	964	SER	-	expression tag	UNP P15145
A	965	HIS	-	expression tag	UNP P15145
A	966	HIS	-	expression tag	UNP P15145
A	967	HIS	-	expression tag	UNP P15145
A	968	HIS	-	expression tag	UNP P15145
A	969	HIS	-	expression tag	UNP P15145
A	970	HIS	-	expression tag	UNP P15145

- Molecule 2 is a protein called poly A peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	7	36	21	7	8	0	0	0

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



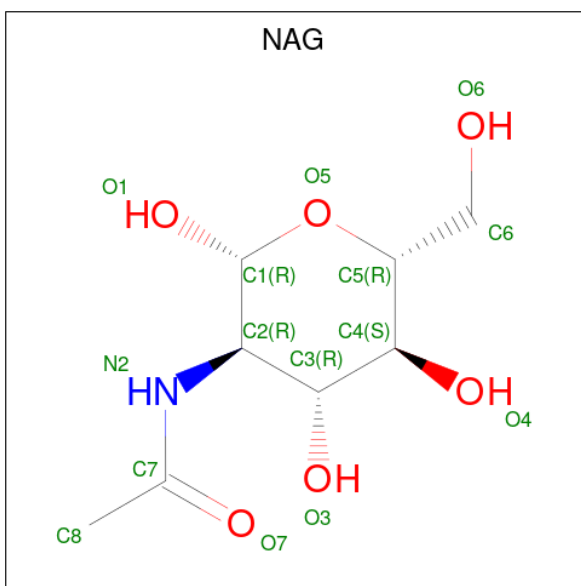
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	2	28	16	2	10	0	0	0

*Continued on next page...*

Continued from previous page...

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

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).



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

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Zn	0	0
			1	1		

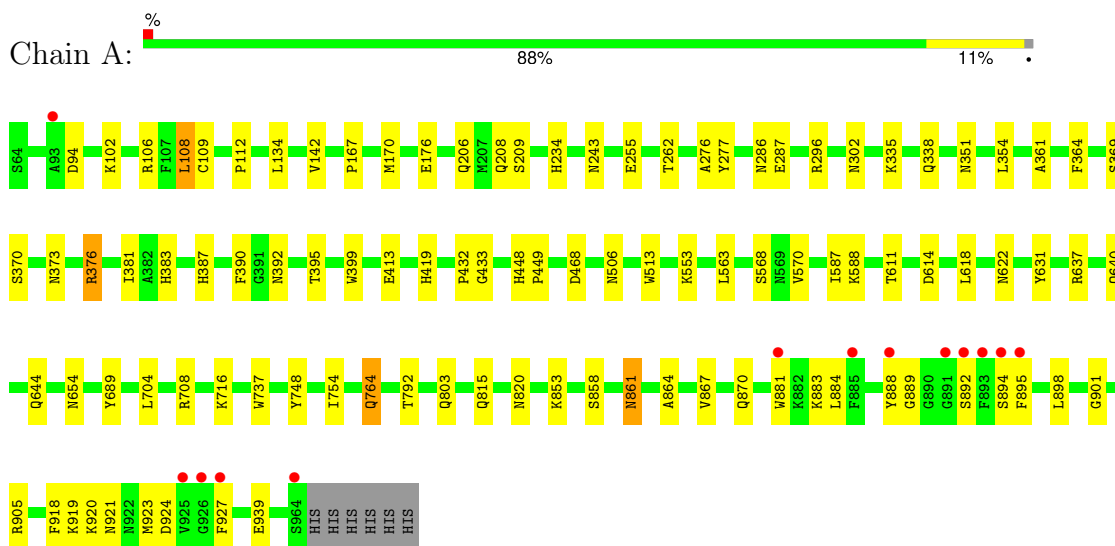
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	783	Total 783	O 783	0	0
7	B	5	Total 5	O 5	0	0

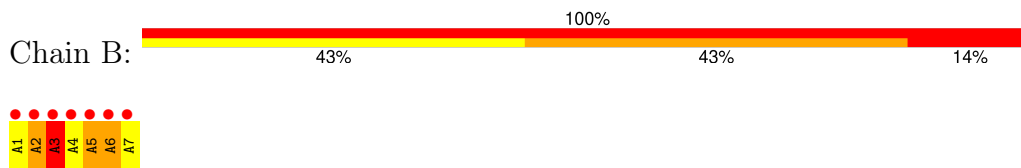
### 3 Residue-property plots [i](#)

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.

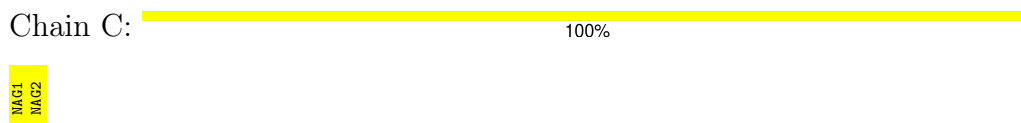
- Molecule 1: Aminopeptidase N



- Molecule 2: poly A peptide



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




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





MAG1  
MAG2

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

Chain E:  50% 50%MAG1  
MAG2

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

Chain F:  100%MAG1  
MAG2

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

Chain G:  50% 50%MAG1  
MAG2

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	260.16Å 62.74Å 81.49Å 90.00° 100.28° 90.00°	Depositor
Resolution (Å)	50.00 – 2.10 50.00 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.5 (50.00-2.10) 99.5 (50.00-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.36 (at 2.10Å)	Xtrriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.145 , 0.198 0.148 , 0.199	Depositor DCC
$R_{free}$ test set	3797 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	34.5	Xtrriage
Anisotropy	0.145	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 54.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	8330	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.73% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, NAG, SO4

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	0.43	0/7428	0.60	0/10123
2	B	0.97	0/35	1.38	0/46
All	All	0.43	0/7463	0.60	0/10169

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
2	B	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	3	ALA	Peptide

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	7240	0	6992	87	0
2	B	36	0	37	19	0
3	C	28	0	25	0	0
3	D	28	0	25	0	0
3	E	28	0	25	1	0
3	F	28	0	25	0	0
3	G	28	0	25	1	0
4	A	70	0	65	0	0
5	A	55	0	0	4	0
6	A	1	0	0	0	0
7	A	783	0	0	16	0
7	B	5	0	0	4	0
All	All	8330	0	7219	98	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.

All (98) 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:A:383:HIS:CE1	2:B:3:ALA:HB3	1.74	1.21
1:A:413:GLU:OE2	2:B:4:ALA:HA	1.42	1.17
1:A:881:TRP:CE3	1:A:918:PHE:HZ	1.87	0.91
1:A:870:GLN:HE21	1:A:905:ARG:HE	1.23	0.86
1:A:286:ASN:HB3	7:A:1799:HOH:O	1.75	0.86
1:A:383:HIS:ND1	2:B:3:ALA:HB3	1.91	0.85
1:A:376:ARG:HG3	7:A:1800:HOH:O	1.75	0.84
1:A:373:ASN:HD21	1:A:820:ASN:HD22	1.26	0.79
1:A:861:ASN:OD1	1:A:901:GLY:HA3	1.83	0.77
2:B:7:ALA:HB2	7:B:103:HOH:O	1.82	0.77
1:A:622:ASN:H	1:A:654:ASN:HD21	1.35	0.70
1:A:243:ASN:HD21	1:A:276:ALA:H	1.41	0.69
1:A:881:TRP:HZ2	1:A:927:PHE:HD1	1.40	0.69
1:A:864:ALA:O	1:A:905:ARG:NH1	2.26	0.68
1:A:208:GLN:HE21	1:A:209:SER:HA	1.59	0.68
1:A:413:GLU:CD	2:B:4:ALA:HA	2.14	0.68
1:A:373:ASN:HD21	1:A:820:ASN:ND2	1.91	0.67
2:B:3:ALA:HB1	2:B:4:ALA:O	1.93	0.67
1:A:881:TRP:CZ3	1:A:918:PHE:HZ	2.13	0.66
1:A:888:TYR:OH	1:A:895:PHE:CD1	2.47	0.65
1:A:881:TRP:CE3	1:A:918:PHE:CZ	2.78	0.65
1:A:176:GLU:HG3	7:A:1861:HOH:O	1.96	0.65

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:7:ALA:CB	7:B:103:HOH:O	2.44	0.64
1:A:884:LEU:HA	1:A:888:TYR:CE1	2.33	0.64
1:A:881:TRP:CZ2	1:A:927:PHE:HD1	2.16	0.63
1:A:468:ASP:HB2	5:A:1024:SO4:O2	1.99	0.63
1:A:553:LYS:HG3	7:A:1862:HOH:O	1.99	0.62
1:A:704:LEU:O	1:A:905:ARG:HG2	1.99	0.62
1:A:383:HIS:CE1	2:B:3:ALA:CB	2.68	0.61
1:A:764:GLN:H	1:A:764:GLN:HE21	1.48	0.61
1:A:921:ASN:HB2	7:A:1619:HOH:O	1.99	0.61
1:A:881:TRP:HZ2	1:A:927:PHE:CD1	2.18	0.60
1:A:881:TRP:CZ3	1:A:918:PHE:CZ	2.89	0.60
1:A:433:GLY:HA2	2:B:6:ALA:O	2.02	0.60
1:A:383:HIS:CD2	2:B:2:ALA:O	2.56	0.59
1:A:433:GLY:HA2	2:B:7:ALA:HA	1.84	0.58
1:A:448:HIS:HE1	1:A:468:ASP:OD2	1.86	0.58
1:A:234:HIS:HE1	1:A:262:THR:OG1	1.87	0.58
1:A:399:TRP:CZ3	5:A:1025:SO4:O1	2.59	0.55
7:A:1784:HOH:O	2:B:7:ALA:CB	2.55	0.55
1:A:708:ARG:HD2	1:A:905:ARG:HD2	1.89	0.55
1:A:618:LEU:HD12	1:A:631:TYR:OH	2.07	0.55
1:A:919:LYS:C	1:A:921:ASN:H	2.09	0.55
2:B:5:ALA:HB2	7:B:101:HOH:O	2.07	0.55
1:A:351:ASN:HB2	1:A:354:LEU:O	2.08	0.54
1:A:939:GLU:HG2	7:A:1611:HOH:O	2.07	0.53
1:A:919:LYS:O	1:A:921:ASN:N	2.42	0.53
1:A:563:LEU:HD11	1:A:570:VAL:CG2	2.39	0.53
1:A:622:ASN:H	1:A:654:ASN:ND2	2.06	0.53
1:A:881:TRP:CZ2	1:A:927:PHE:CD1	2.96	0.53
1:A:395:THR:O	1:A:506:ASN:HA	2.10	0.52
1:A:637:ARG:NH2	1:A:640:GLN:HE22	2.07	0.51
1:A:338:GLN:HE22	1:A:381:ILE:HG23	1.76	0.50
1:A:392:ASN:HD22	1:A:392:ASN:N	2.09	0.49
1:A:614:ASP:OD1	1:A:614:ASP:N	2.45	0.49
1:A:708:ARG:NH1	7:A:1312:HOH:O	2.46	0.49
1:A:419:HIS:NE2	5:A:1020:SO4:O2	2.44	0.49
1:A:373:ASN:ND2	1:A:820:ASN:HD22	2.03	0.49
1:A:881:TRP:CH2	1:A:895:PHE:CE2	3.01	0.49
1:A:208:GLN:HE21	1:A:209:SER:CA	2.27	0.48
1:A:888:TYR:CD1	1:A:888:TYR:N	2.81	0.48
2:B:1:ALA:O	2:B:2:ALA:HB3	2.15	0.47
1:A:764:GLN:H	1:A:764:GLN:NE2	2.12	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:754:ILE:HG22	1:A:792:THR:HG21	1.98	0.46
1:A:708:ARG:HD3	7:A:1165:HOH:O	2.15	0.46
1:A:861:ASN:HD21	1:A:898:LEU:HA	1.81	0.46
1:A:587:ILE:HG22	1:A:618:LEU:HB3	1.97	0.46
1:A:106:ARG:HG2	1:A:170:MET:HE1	1.99	0.45
1:A:302:ASN:ND2	5:A:1026:SO4:O4	2.39	0.45
1:A:737:TRP:CZ3	1:A:754:ILE:HD11	2.52	0.45
1:A:370:SER:O	1:A:373:ASN:HB3	2.17	0.45
1:A:102:LYS:HG3	7:A:1861:HOH:O	2.17	0.44
2:B:4:ALA:O	2:B:5:ALA:CB	2.66	0.44
1:A:208:GLN:HA	1:A:209:SER:HA	1.80	0.43
7:A:1571:HOH:O	2:B:1:ALA:HB3	2.17	0.43
7:A:1863:HOH:O	3:G:2:NAG:H82	2.18	0.43
1:A:112:PRO:HD3	1:A:167:PRO:HG3	2.00	0.43
1:A:134:LEU:HG	1:A:142:VAL:HG21	2.00	0.43
1:A:448:HIS:HD2	1:A:449:PRO:O	2.02	0.43
1:A:255:GLU:HB3	7:A:1704:HOH:O	2.18	0.43
1:A:884:LEU:HA	1:A:888:TYR:CD1	2.54	0.43
1:A:588:LYS:HE2	1:A:611:THR:HG22	2.02	0.42
1:A:881:TRP:CZ2	1:A:895:PHE:CE2	3.07	0.42
1:A:286:ASN:HB3	1:A:296:ARG:HG2	2.02	0.42
1:A:106:ARG:NH1	3:E:1:NAG:O3	2.41	0.42
1:A:364:PHE:CE1	1:A:369:SER:HB2	2.54	0.42
1:A:618:LEU:CD1	7:A:1215:HOH:O	2.68	0.42
1:A:689:TYR:CD1	1:A:748:TYR:HB3	2.54	0.42
1:A:108:LEU:HD12	1:A:109:CYS:N	2.34	0.42
1:A:881:TRP:CH2	1:A:895:PHE:CZ	3.08	0.42
1:A:243:ASN:ND2	1:A:276:ALA:H	2.12	0.41
2:B:4:ALA:O	2:B:5:ALA:HB2	2.21	0.41
1:A:919:LYS:C	1:A:921:ASN:N	2.74	0.41
1:A:387:HIS:HA	1:A:390:PHE:O	2.20	0.41
1:A:432:PRO:HB2	2:B:7:ALA:HB1	2.03	0.41
1:A:803:GLN:NE2	7:A:1768:HOH:O	2.54	0.41
1:A:361:ALA:HA	7:A:1800:HOH:O	2.21	0.40
2:B:1:ALA:C	7:B:104:HOH:O	2.60	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	900/907 (99%)	863 (96%)	31 (3%)	6 (1%)	19	16
2	B	5/7 (71%)	0	1 (20%)	4 (80%)	0	0
All	All	905/914 (99%)	863 (95%)	32 (4%)	10 (1%)	12	8

All (10) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	94	ASP
1	A	894	SER
1	A	920	LYS
2	B	2	ALA
2	B	5	ALA
2	B	6	ALA
2	B	3	ALA
1	A	889	GLY
1	A	924	ASP
1	A	892	SER

### 5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	797/802 (99%)	779 (98%)	18 (2%)	45	51

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	108	LEU
1	A	206	GLN
1	A	277	TYR
1	A	287	GLU
1	A	335	LYS
1	A	376	ARG
1	A	513	TRP
1	A	568	SER
1	A	644	GLN
1	A	716	LYS
1	A	764	GLN
1	A	815	GLN
1	A	853	LYS
1	A	858	SER
1	A	861	ASN
1	A	867	VAL
1	A	883	LYS
1	A	923	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (19) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	206	GLN
1	A	208	GLN
1	A	234	HIS
1	A	243	ASN
1	A	338	GLN
1	A	367	GLN
1	A	392	ASN
1	A	448	HIS
1	A	654	ASN
1	A	722	GLN
1	A	764	GLN
1	A	796	ASN
1	A	800	GLN
1	A	815	GLN
1	A	820	ASN
1	A	839	ASN
1	A	870	GLN
1	A	897	ASN
1	A	951	ASN



### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

10 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
3	NAG	C	1	3,1	14,14,15	0.56	0	17,19,21	0.98	1 (5%)
3	NAG	C	2	3	14,14,15	0.57	0	17,19,21	1.42	2 (11%)
3	NAG	D	1	3,1	14,14,15	0.49	0	17,19,21	1.39	2 (11%)
3	NAG	D	2	3	14,14,15	0.48	0	17,19,21	1.43	2 (11%)
3	NAG	E	1	3,1	14,14,15	0.64	0	17,19,21	0.97	1 (5%)
3	NAG	E	2	3	14,14,15	0.47	0	17,19,21	2.33	4 (23%)
3	NAG	F	1	3,1	14,14,15	0.52	0	17,19,21	2.11	4 (23%)
3	NAG	F	2	3	14,14,15	0.53	0	17,19,21	1.01	1 (5%)
3	NAG	G	1	3,1	14,14,15	0.46	0	17,19,21	1.12	1 (5%)
3	NAG	G	2	3	14,14,15	0.44	0	17,19,21	1.96	5 (29%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	C	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	C	2	3	-	0/6/23/26	0/1/1/1
3	NAG	D	1	3,1	-	1/6/23/26	0/1/1/1

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	D	2	3	-	2/6/23/26	0/1/1/1
3	NAG	E	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	E	2	3	-	1/6/23/26	0/1/1/1
3	NAG	F	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	F	2	3	-	3/6/23/26	0/1/1/1
3	NAG	G	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	G	2	3	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	2	NAG	C1-O5-C5	7.84	122.70	112.19
3	F	1	NAG	C1-O5-C5	6.00	120.23	112.19
3	G	2	NAG	C1-O5-C5	5.29	119.28	112.19
3	D	1	NAG	C1-O5-C5	4.63	118.39	112.19
3	D	2	NAG	C1-O5-C5	4.23	117.85	112.19
3	C	2	NAG	C4-C3-C2	4.07	116.98	111.02
3	G	1	NAG	C1-O5-C5	3.66	117.09	112.19
3	F	1	NAG	C1-C2-N2	-3.26	105.29	110.43
3	F	1	NAG	C2-N2-C7	-3.15	118.68	122.90
3	E	2	NAG	C3-C4-C5	3.11	115.88	110.23
3	G	2	NAG	C1-C2-N2	2.99	115.15	110.43
3	E	2	NAG	O5-C5-C4	2.87	117.81	110.83
3	C	1	NAG	C1-O5-C5	2.84	115.99	112.19
3	E	2	NAG	C6-C5-C4	-2.77	106.21	113.02
3	F	2	NAG	C8-C7-N2	2.62	120.46	116.12
3	G	2	NAG	C4-C3-C2	-2.60	107.21	111.02
3	D	2	NAG	C1-C2-N2	2.44	114.27	110.43
3	E	1	NAG	C2-N2-C7	-2.39	119.69	122.90
3	C	2	NAG	C2-N2-C7	-2.21	119.94	122.90
3	D	1	NAG	C4-C3-C2	2.18	114.21	111.02
3	G	2	NAG	O5-C1-C2	-2.12	108.01	111.29
3	G	2	NAG	O5-C5-C4	2.12	115.99	110.83
3	F	1	NAG	O3-C3-C4	-2.04	105.58	110.38

There are no chirality outliers.

All (9) torsion outliers are listed below:

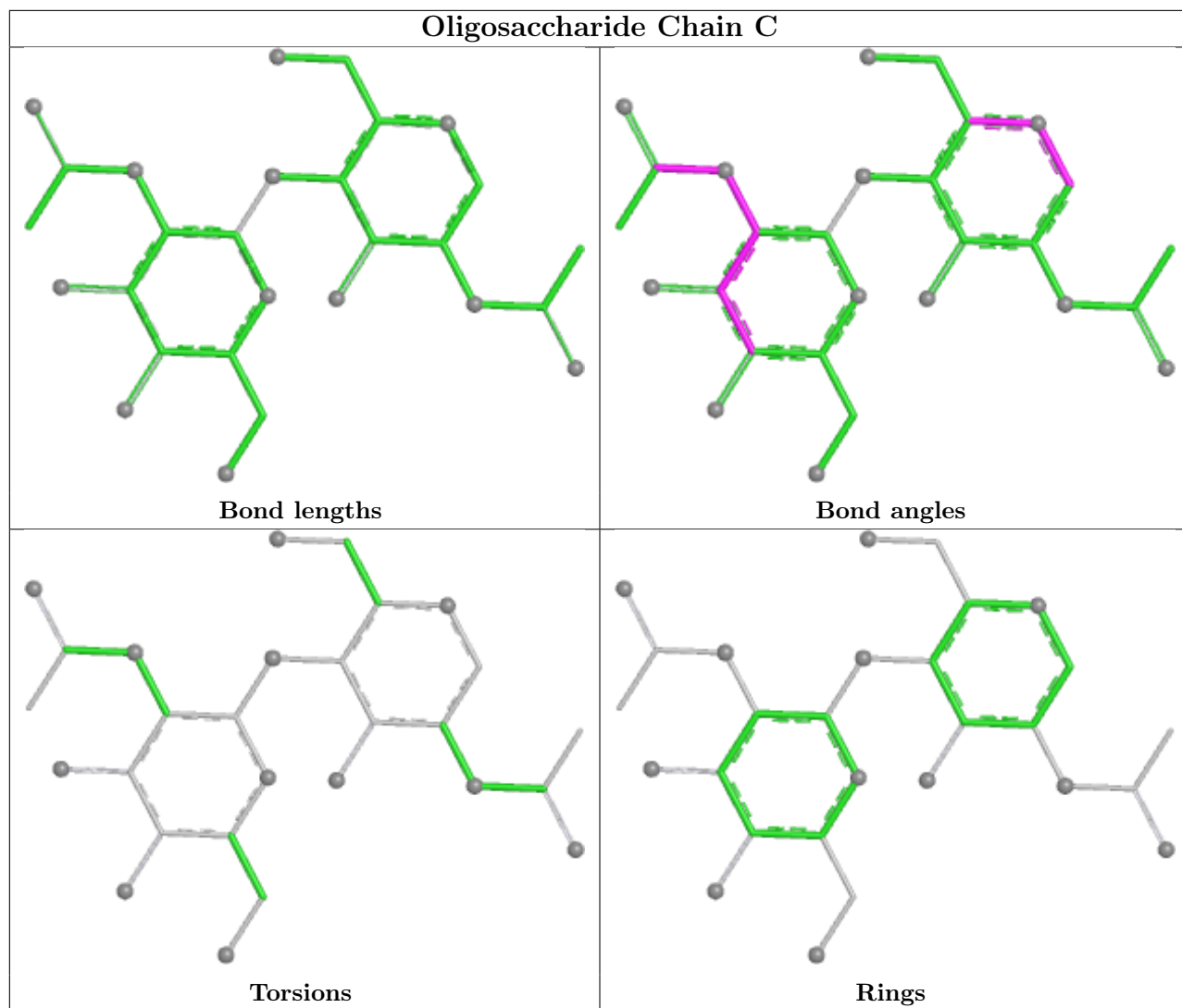
Mol	Chain	Res	Type	Atoms
3	D	2	NAG	C4-C5-C6-O6
3	G	2	NAG	C4-C5-C6-O6
3	F	2	NAG	C8-C7-N2-C2
3	F	2	NAG	O7-C7-N2-C2
3	G	2	NAG	O5-C5-C6-O6
3	D	2	NAG	O5-C5-C6-O6
3	E	2	NAG	C4-C5-C6-O6
3	D	1	NAG	C1-C2-N2-C7
3	F	2	NAG	C1-C2-N2-C7

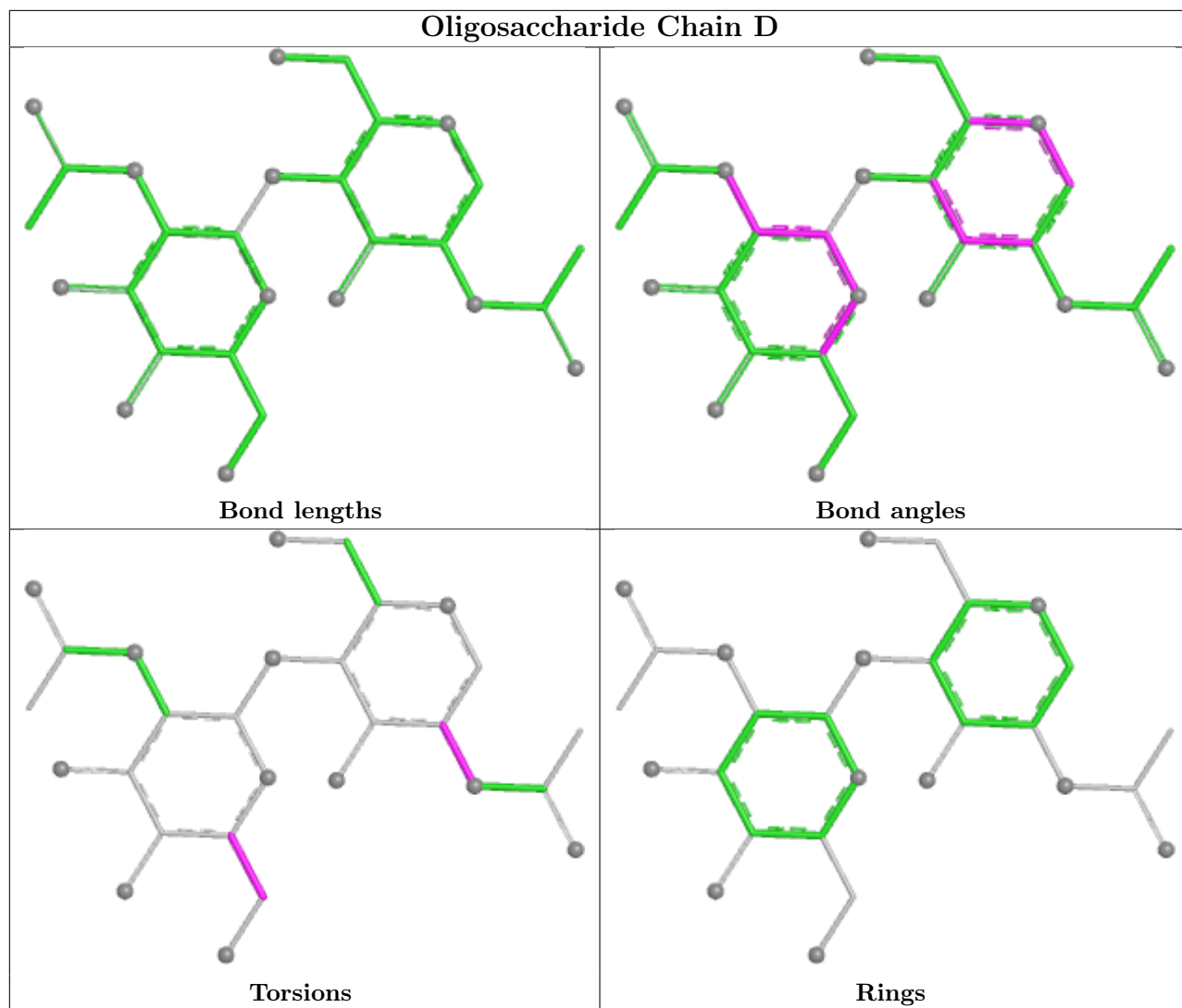
There are no ring outliers.

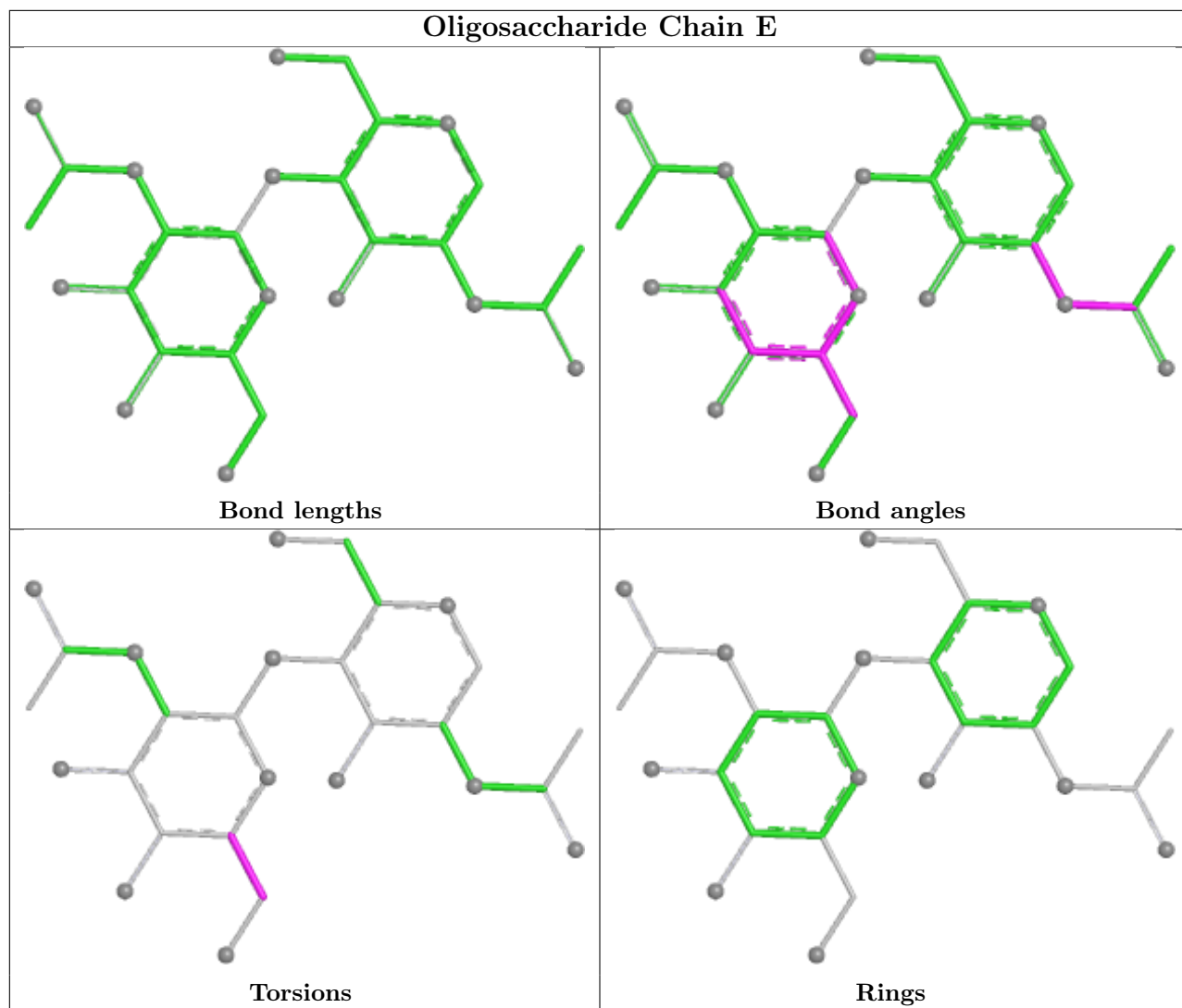
2 monomers are involved in 2 short contacts:

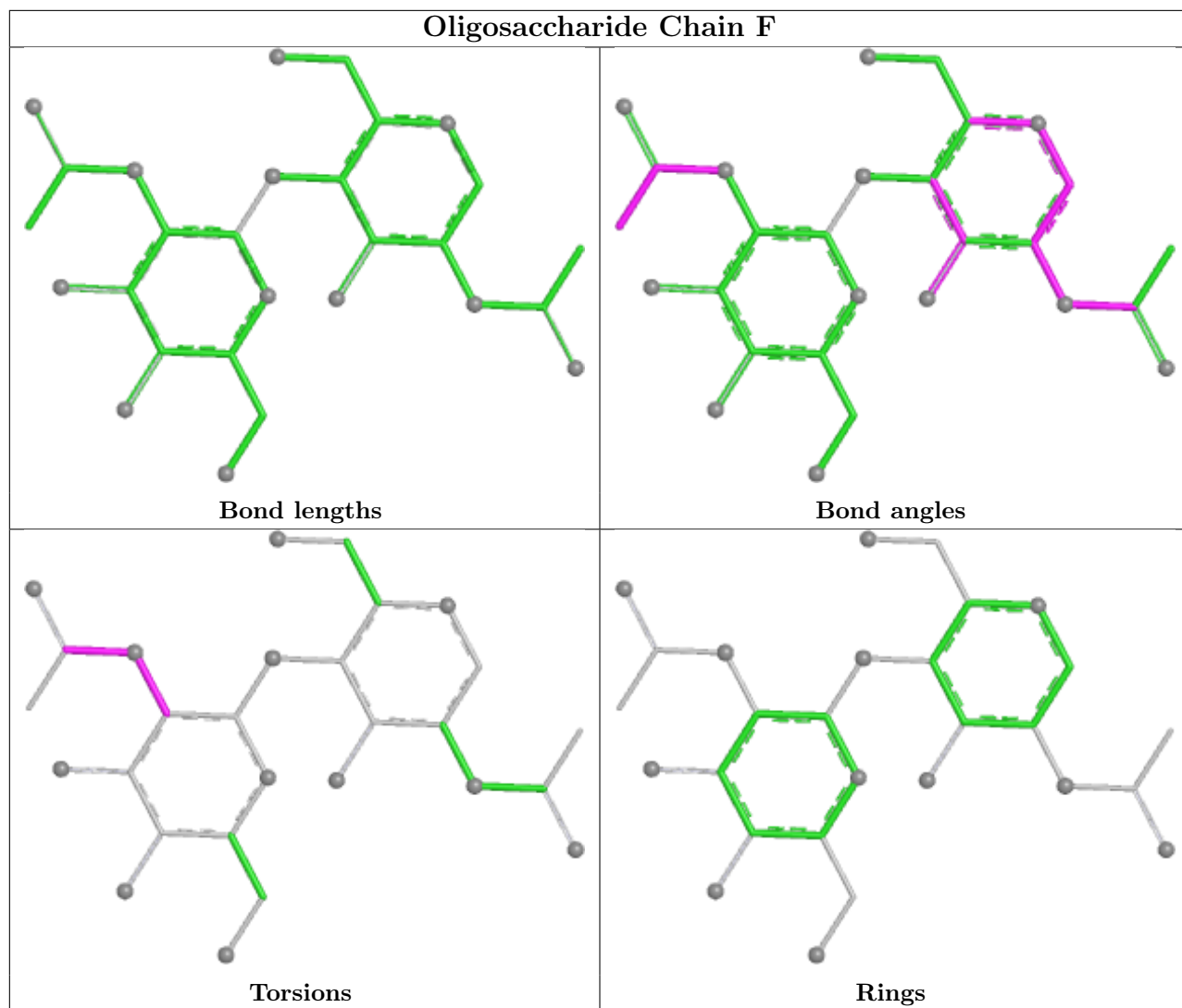
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	G	2	NAG	1	0
3	E	1	NAG	1	0

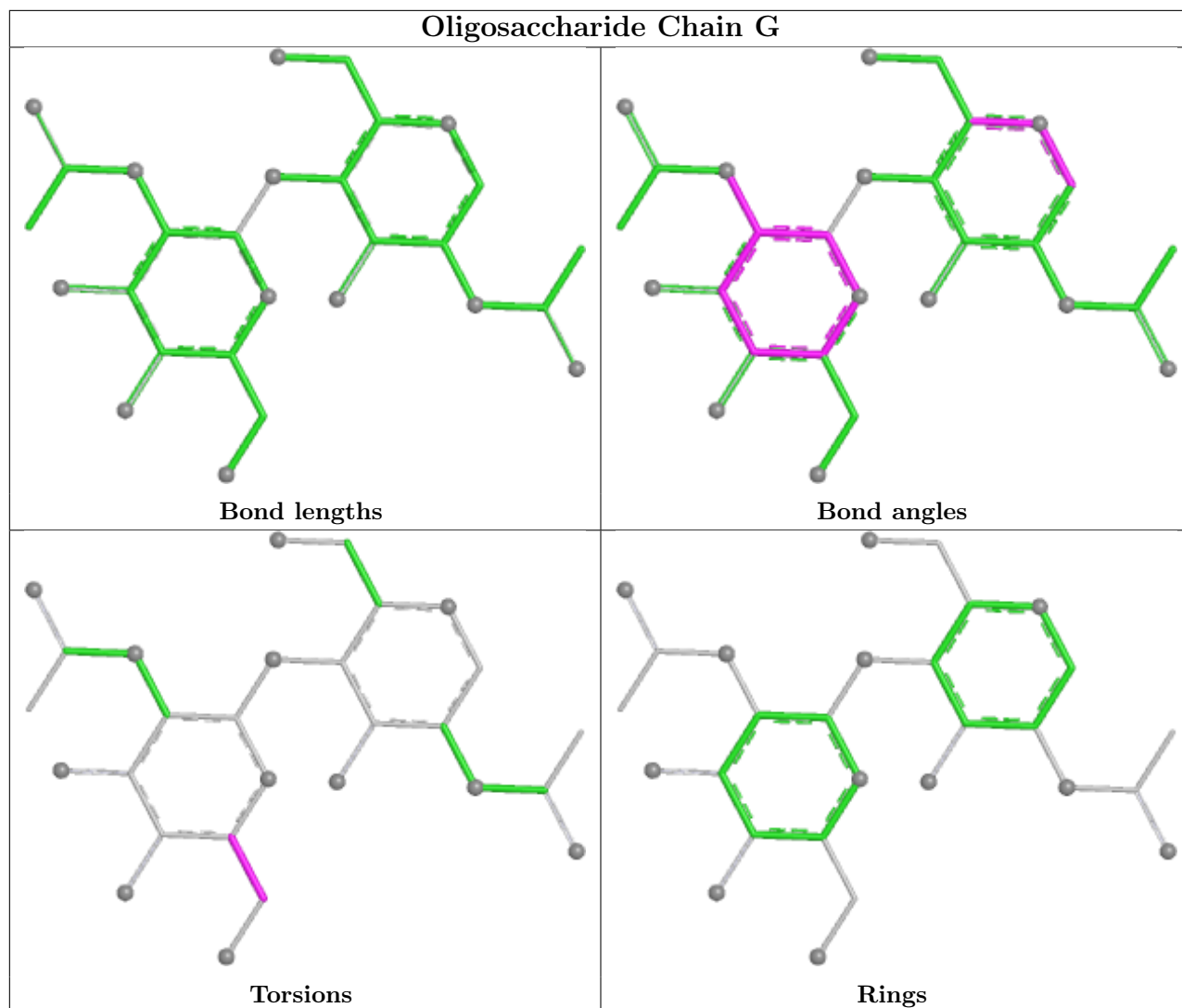
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.











## 5.6 Ligand geometry [i](#)

Of 17 ligands modelled in this entry, 1 is monoatomic - leaving 16 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	$\# Z  > 2$
4	NAG	A	1012	1	14,14,15	0.50	0	17,19,21	1.15	1 (5%)
5	SO4	A	1016	-	4,4,4	0.74	0	6,6,6	0.71	0
4	NAG	A	1003	1	14,14,15	0.36	0	17,19,21	1.64	5 (29%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	SO4	A	1021	-	4,4,4	0.47	0	6,6,6	0.16	0
5	SO4	A	1023	-	4,4,4	0.44	0	6,6,6	0.13	0
5	SO4	A	1024	-	4,4,4	0.41	0	6,6,6	0.15	0
5	SO4	A	1018	-	4,4,4	0.45	0	6,6,6	0.08	0
5	SO4	A	1026	-	4,4,4	0.44	0	6,6,6	0.17	0
4	NAG	A	1007	1	14,14,15	0.52	0	17,19,21	1.48	3 (17%)
4	NAG	A	1006	1	14,14,15	0.55	0	17,19,21	0.98	0
5	SO4	A	1017	-	4,4,4	0.45	0	6,6,6	0.27	0
5	SO4	A	1019	-	4,4,4	0.54	0	6,6,6	0.33	0
5	SO4	A	1020	-	4,4,4	0.42	0	6,6,6	0.17	0
4	NAG	A	1015	1	14,14,15	0.51	0	17,19,21	1.14	1 (5%)
5	SO4	A	1022	-	4,4,4	0.47	0	6,6,6	0.17	0
5	SO4	A	1025	-	4,4,4	0.52	0	6,6,6	0.85	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	A	1012	1	-	2/6/23/26	0/1/1/1
4	NAG	A	1003	1	-	1/6/23/26	0/1/1/1
4	NAG	A	1007	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1006	1	-	1/6/23/26	0/1/1/1
4	NAG	A	1015	1	-	1/6/23/26	0/1/1/1

There are no bond length outliers.

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1007	NAG	C1-O5-C5	4.57	118.31	112.19
4	A	1003	NAG	C1-O5-C5	4.09	117.67	112.19
4	A	1015	NAG	C1-O5-C5	3.69	117.12	112.19
4	A	1012	NAG	C1-O5-C5	3.19	116.46	112.19
4	A	1003	NAG	C1-C2-N2	2.34	114.13	110.43
4	A	1003	NAG	C2-N2-C7	2.29	125.97	122.90
4	A	1003	NAG	O7-C7-N2	2.20	125.87	121.98
4	A	1003	NAG	O5-C1-C2	-2.14	107.98	111.29
4	A	1007	NAG	C2-N2-C7	-2.03	120.17	122.90
4	A	1007	NAG	C1-C2-N2	2.02	113.62	110.43

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1012	NAG	C4-C5-C6-O6
4	A	1012	NAG	O5-C5-C6-O6
4	A	1003	NAG	O5-C5-C6-O6
4	A	1006	NAG	C1-C2-N2-C7
4	A	1015	NAG	C1-C2-N2-C7

There are no ring outliers.

4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1024	SO4	1	0
5	A	1026	SO4	1	0
5	A	1020	SO4	1	0
5	A	1025	SO4	1	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	901/907 (99%)	-0.37	13 (1%) 73 74	22, 44, 81, 159	1 (0%)
2	B	7/7 (100%)	4.79	7 (100%) 0 0	31, 39, 65, 72	0
All	All	908/914 (99%)	-0.33	20 (2%) 62 64	22, 44, 81, 159	1 (0%)

All (20) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	6	ALA	7.1
2	B	2	ALA	6.0
2	B	3	ALA	5.3
2	B	5	ALA	4.9
2	B	4	ALA	4.2
1	A	888	TYR	4.0
1	A	885	PHE	3.9
1	A	895	PHE	3.9
2	B	1	ALA	3.4
1	A	925	VAL	3.4
1	A	892	SER	3.0
1	A	964	SER	3.0
1	A	893	PHE	3.0
1	A	881	TRP	2.9
1	A	93	ALA	2.8
1	A	894	SER	2.8
1	A	926	GLY	2.6
2	B	7	ALA	2.5
1	A	927	PHE	2.5
1	A	891	GLY	2.3

## 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](#)

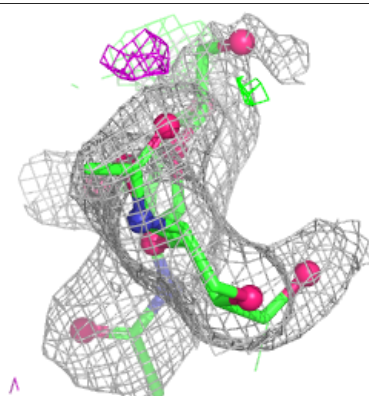
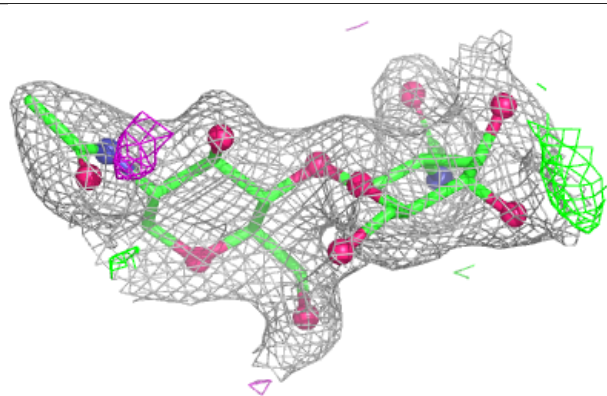
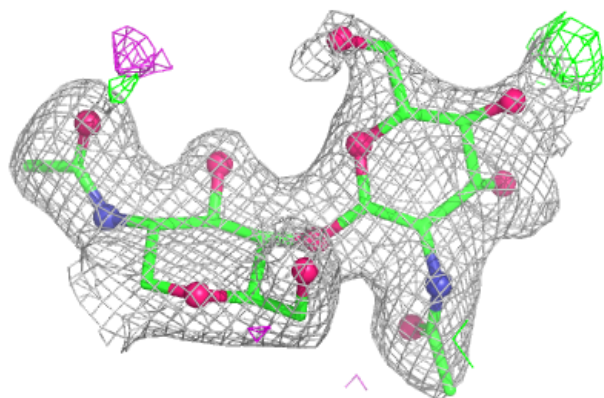
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
3	NAG	D	2	14/15	0.81	0.12	43,92,106,117	0
3	NAG	D	1	14/15	0.84	0.11	59,73,86,94	0
3	NAG	C	2	14/15	0.85	0.10	58,76,92,103	0
3	NAG	F	2	14/15	0.87	0.10	61,67,85,90	0
3	NAG	G	2	14/15	0.90	0.11	75,94,101,105	0
3	NAG	E	2	14/15	0.91	0.10	51,60,79,91	0
3	NAG	G	1	14/15	0.94	0.09	51,62,79,80	0
3	NAG	F	1	14/15	0.95	0.07	32,38,45,48	0
3	NAG	C	1	14/15	0.96	0.06	35,42,52,55	0
3	NAG	E	1	14/15	0.97	0.06	28,34,46,49	0

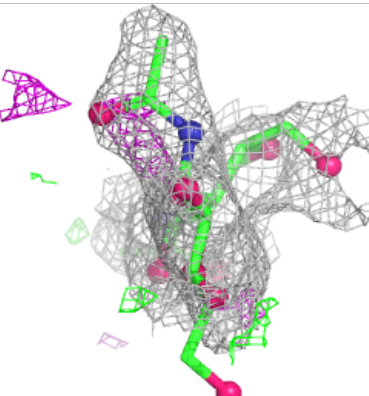
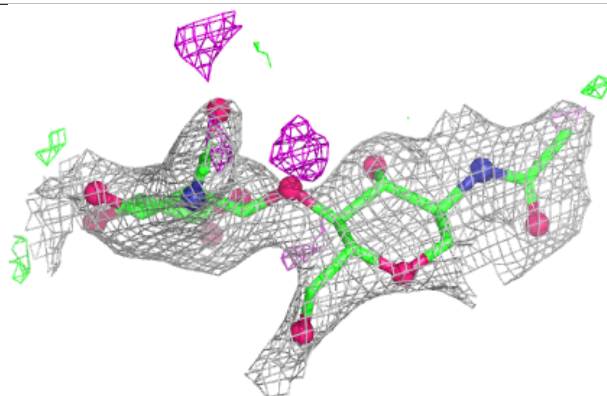
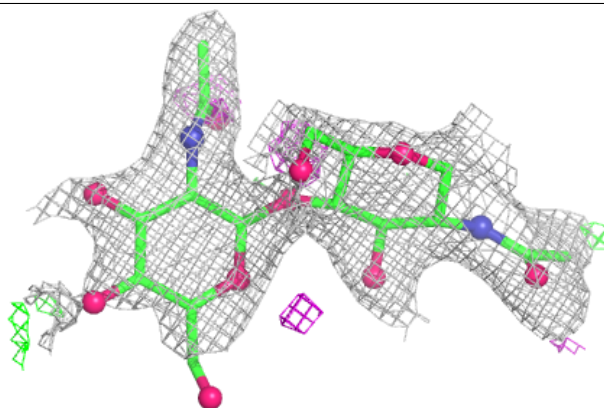
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

**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)

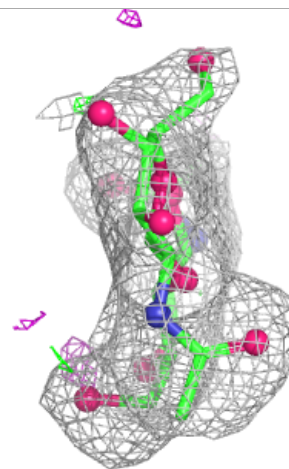
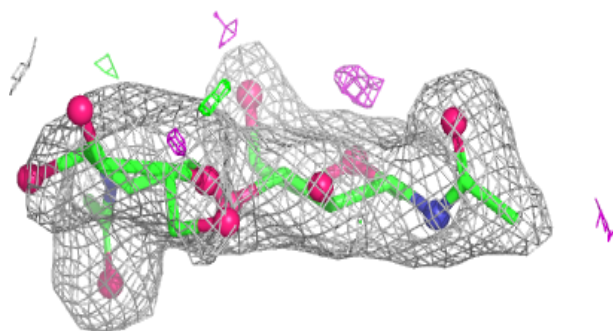
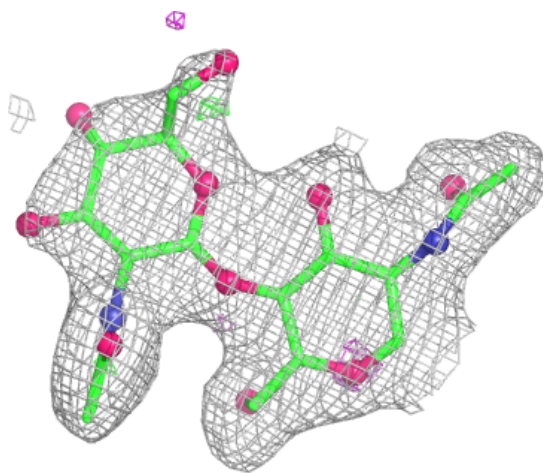
**Electron density around Chain D:**

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



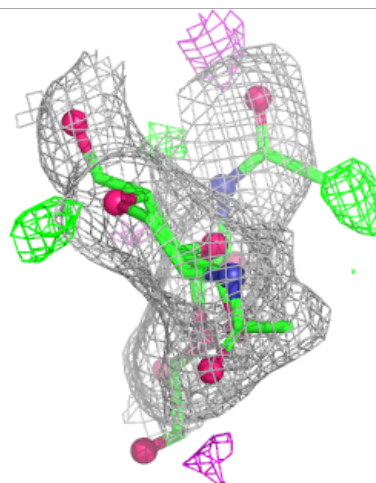
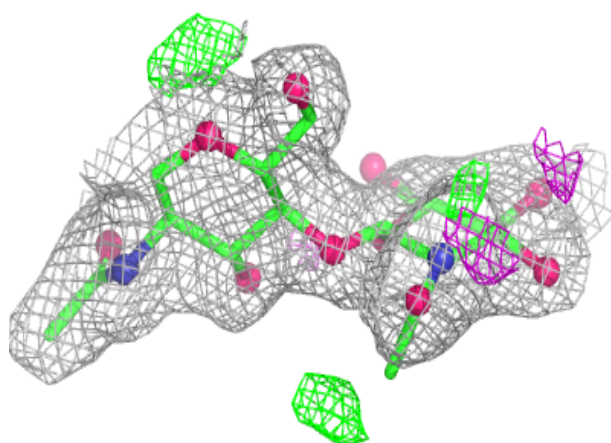
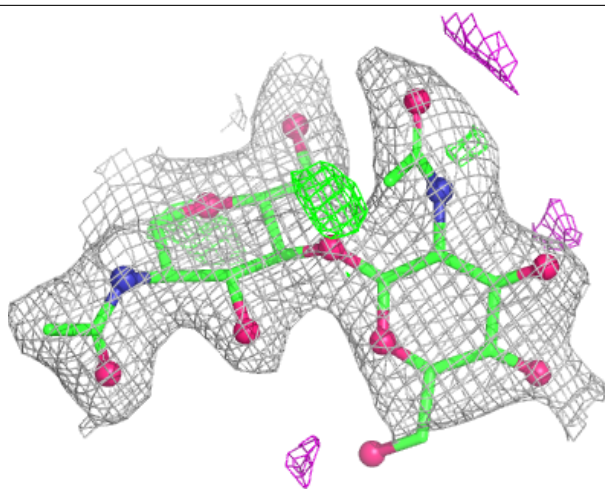
**Electron density around Chain E:**

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

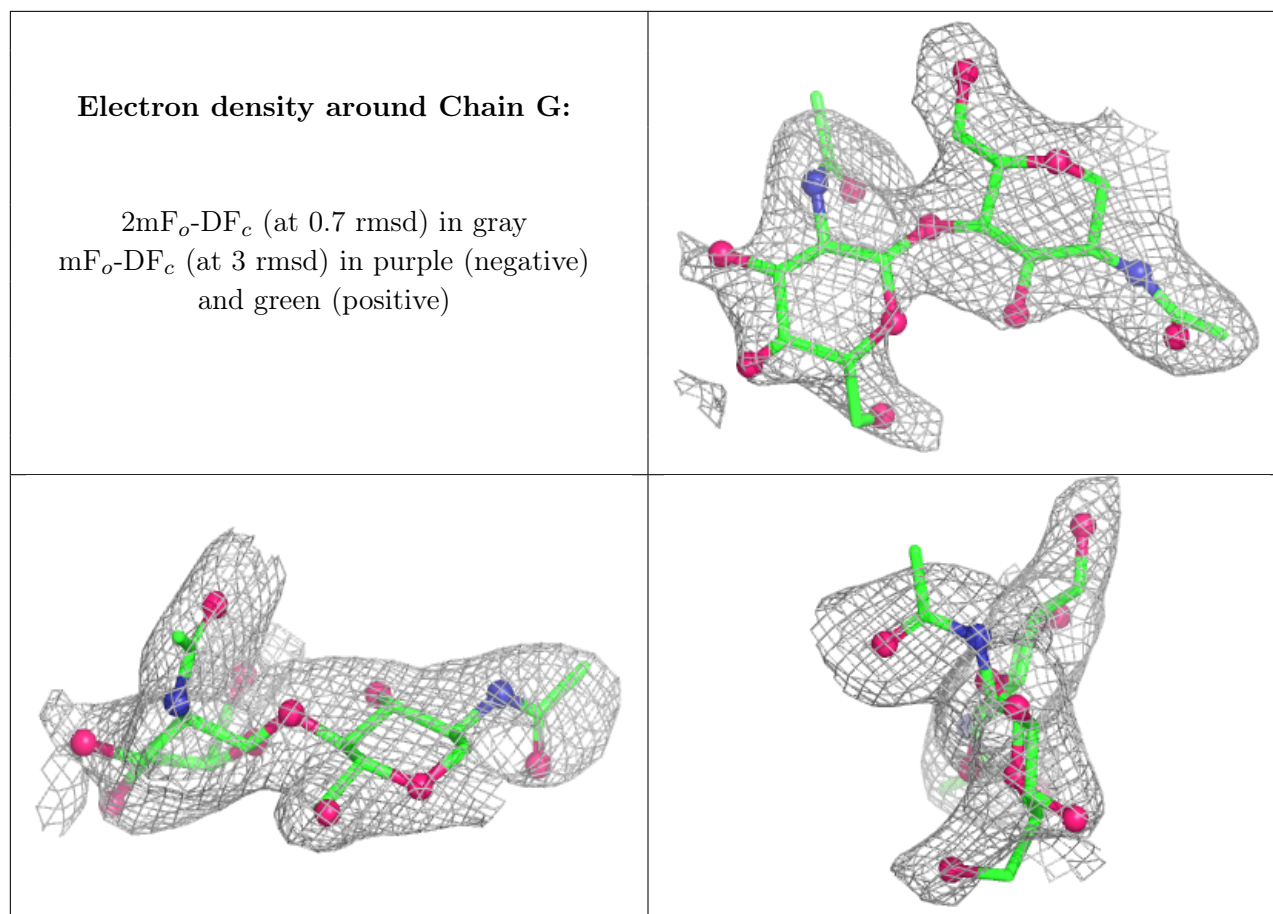


**Electron density around Chain F:**

$2mF_o-DF_c$  (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
4	NAG	A	1015	14/15	0.78	0.11	94,100,107,114	0
5	SO4	A	1024	5/5	0.78	0.10	70,88,101,112	0
4	NAG	A	1012	14/15	0.79	0.13	69,84,94,96	0
4	NAG	A	1006	14/15	0.85	0.11	49,59,76,81	0
5	SO4	A	1023	5/5	0.86	0.15	106,109,136,136	0
6	ZN	A	1027	1/1	0.86	0.10	29,29,29,29	0
5	SO4	A	1017	5/5	0.88	0.09	74,75,90,94	0
5	SO4	A	1026	5/5	0.89	0.08	82,88,94,98	0
5	SO4	A	1020	5/5	0.89	0.10	83,83,94,111	0
5	SO4	A	1021	5/5	0.90	0.09	89,99,104,110	0
4	NAG	A	1007	14/15	0.90	0.11	47,55,65,70	0
4	NAG	A	1003	14/15	0.92	0.07	43,53,68,72	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	SO4	A	1018	5/5	0.93	0.10	63,73,84,85	0
5	SO4	A	1022	5/5	0.93	0.11	84,85,106,107	0
5	SO4	A	1025	5/5	0.93	0.13	43,53,82,86	0
5	SO4	A	1019	5/5	0.94	0.10	70,70,84,89	0
5	SO4	A	1016	5/5	0.97	0.10	53,54,56,71	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.