



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 17, 2020 – 10:28 AM BST

PDB ID : 6E7G  
Title : Crystal structure of H5 hemagglutinin mutant Y161A from A/Viet Nam/1203/2004 H5N1 influenza virus  
Authors : Tzarum, N.; Wilson, I.A.  
Deposited on : 2018-07-26  
Resolution : 3.09 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

---

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

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

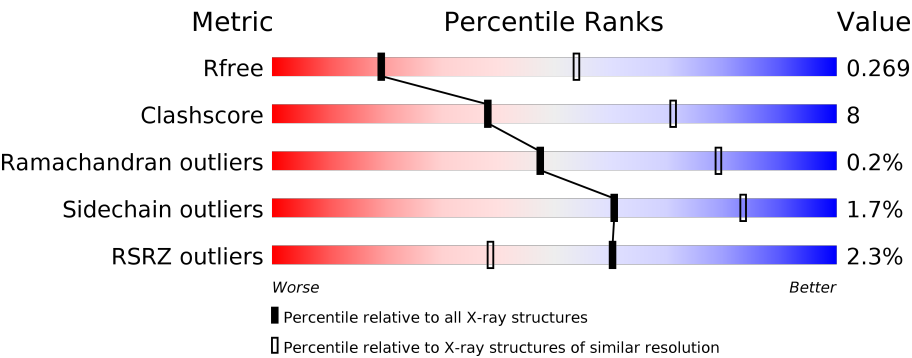
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.09 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.






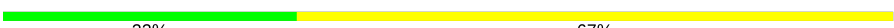
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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 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	334	<div><div>%</div><div><div></div><div></div><div></div><div></div><div></div></div><div>77%16%• 6%</div></div>
1	C	334	<div><div>%</div><div><div></div><div></div><div></div><div></div><div></div></div><div>73%21%6%</div></div>
1	E	334	<div><div>%</div><div><div></div><div></div><div></div><div></div><div></div></div><div>74%19%• 7%</div></div>
2	B	177	<div><div>3%</div><div><div></div><div></div><div></div><div></div><div></div></div><div>79%16%• •</div></div>
2	D	177	<div><div>4%</div><div><div></div><div></div><div></div><div></div><div></div></div><div>77%12%• 10%</div></div>
2	F	177	<div><div>5%</div><div><div></div><div></div><div></div><div></div><div></div></div><div>68%20%• 12%</div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	2	 50% 50%
3	H	2	 50% 50%
3	J	2	 50% 50%
4	I	3	 33% 67%

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 11548 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 Hemagglutinin HA1 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	313	Total	C	N	O	S	0	0	0
			2478	1565	430	468	15			
1	C	313	Total	C	N	O	S	0	0	0
			2473	1563	429	466	15			
1	E	311	Total	C	N	O	S	0	0	0
			2458	1552	426	465	15			

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	ALA	-	expression tag	UNP Q5EP31
A	8	ASP	-	expression tag	UNP Q5EP31
A	9	PRO	-	expression tag	UNP Q5EP31
A	10	GLY	-	expression tag	UNP Q5EP31
A	161	ALA	TYR	engineered mutation	UNP Q5EP31
C	7	ALA	-	expression tag	UNP Q5EP31
C	8	ASP	-	expression tag	UNP Q5EP31
C	9	PRO	-	expression tag	UNP Q5EP31
C	10	GLY	-	expression tag	UNP Q5EP31
C	161	ALA	TYR	engineered mutation	UNP Q5EP31
E	7	ALA	-	expression tag	UNP Q5EP31
E	8	ASP	-	expression tag	UNP Q5EP31
E	9	PRO	-	expression tag	UNP Q5EP31
E	10	GLY	-	expression tag	UNP Q5EP31
E	161	ALA	TYR	engineered mutation	UNP Q5EP31

- Molecule 2 is a protein called Hemagglutinin HA2 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	170	Total	C	N	O	S	0	0	0
			1383	858	241	276	8			

*Continued on next page...*

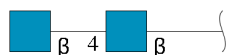
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	159	Total	C	N	O	S	0	0	0
			1319	818	230	263	8			
2	F	156	Total	C	N	O	S	0	0	0
			1286	794	227	257	8			

There are 9 discrepancies between the modelled and reference sequences:

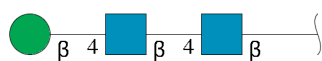
Chain	Residue	Modelled	Actual	Comment	Reference
B	175	SER	-	expression tag	UNP Q6DQ18
B	176	GLY	-	expression tag	UNP Q6DQ18
B	177	ARG	-	expression tag	UNP Q6DQ18
D	175	SER	-	expression tag	UNP Q6DQ18
D	176	GLY	-	expression tag	UNP Q6DQ18
D	177	ARG	-	expression tag	UNP Q6DQ18
F	175	SER	-	expression tag	UNP Q6DQ18
F	176	GLY	-	expression tag	UNP Q6DQ18
F	177	ARG	-	expression tag	UNP Q6DQ18

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



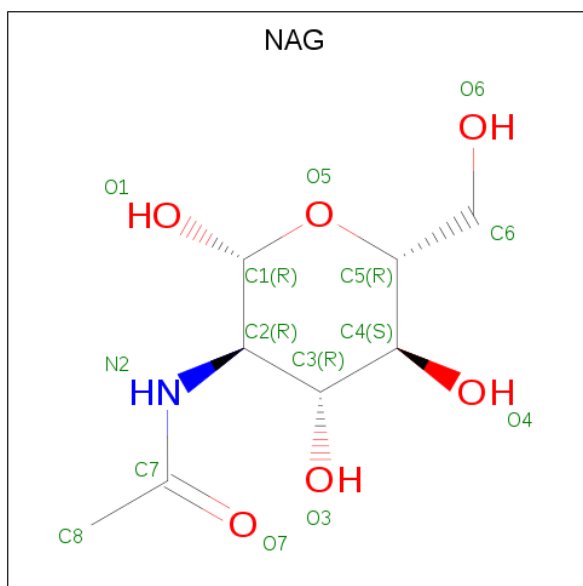
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	2	Total	C	N	O		0	0	0
			28	16	2	10				
3	H	2	Total	C	N	O		0	0	0
			28	16	2	10				
3	J	2	Total	C	N	O		0	0	0
			28	16	2	10				

- Molecule 4 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
4	I	3	Total	C	N	O	0	0	0
			39	22	2	15			

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).

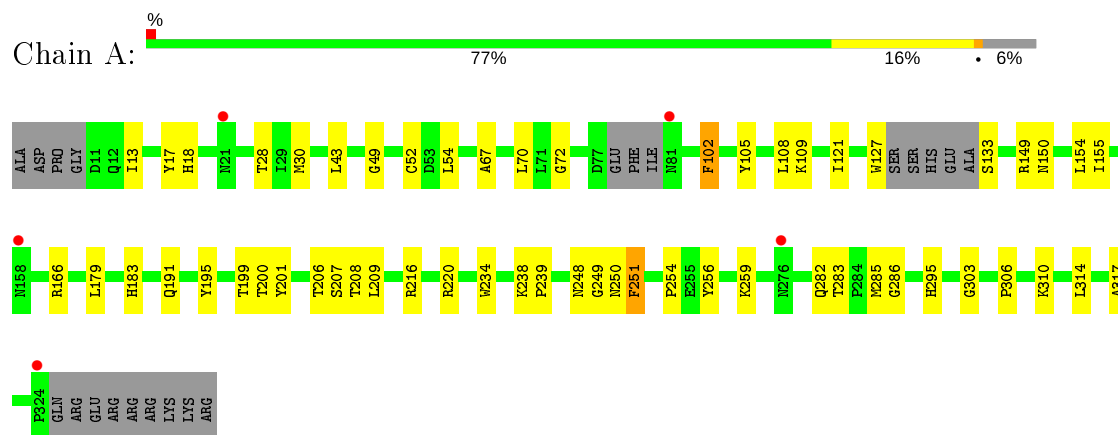


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	E	1	Total	C	N	O	0	0
			14	8	1	5		

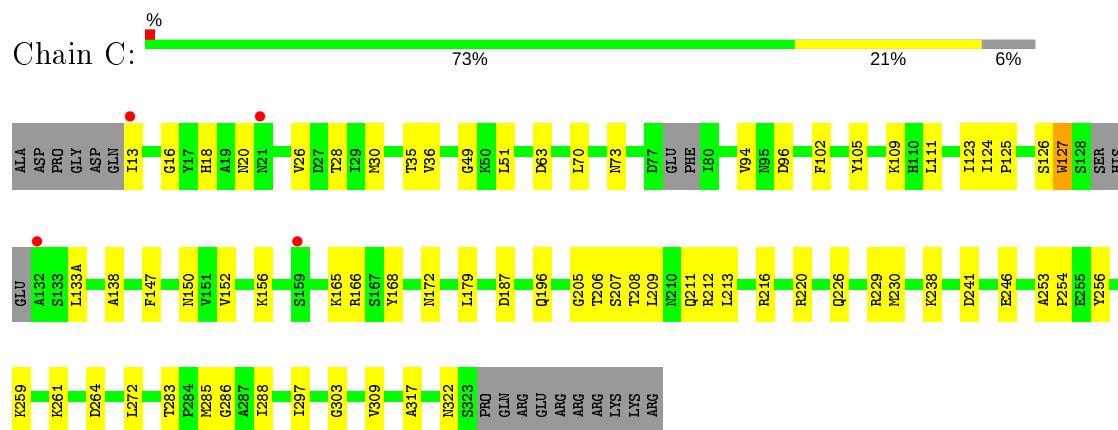
### 3 Residue-property plots

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.

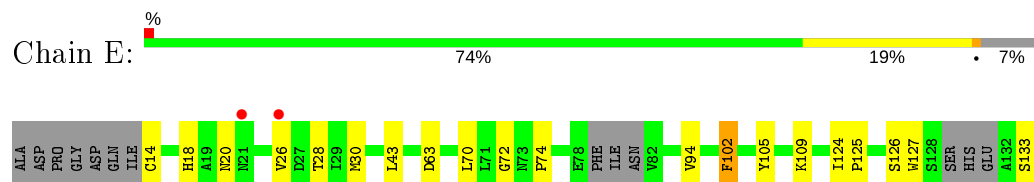
- Molecule 1: Hemagglutinin HA1 chain

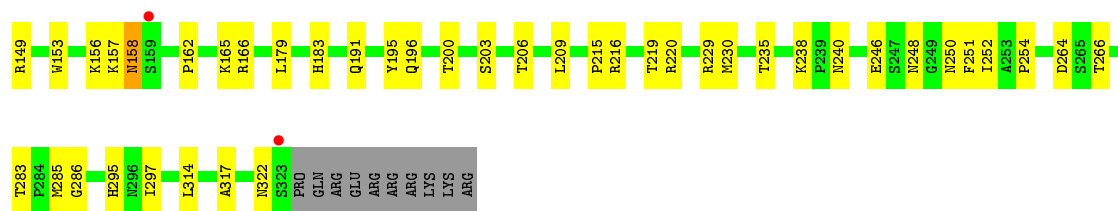


- Molecule 1: Hemagglutinin HA1 chain

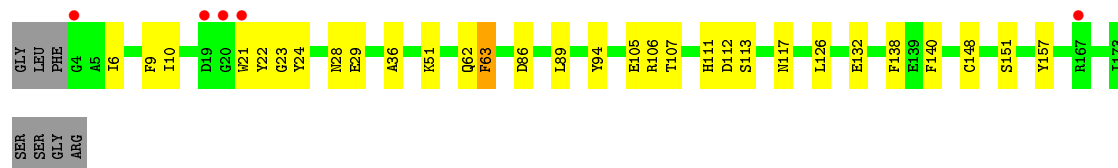
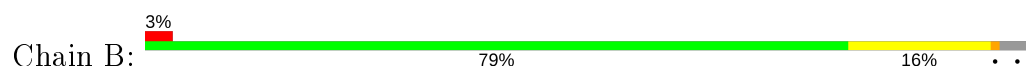


- Molecule 1: Hemagglutinin HA1 chain

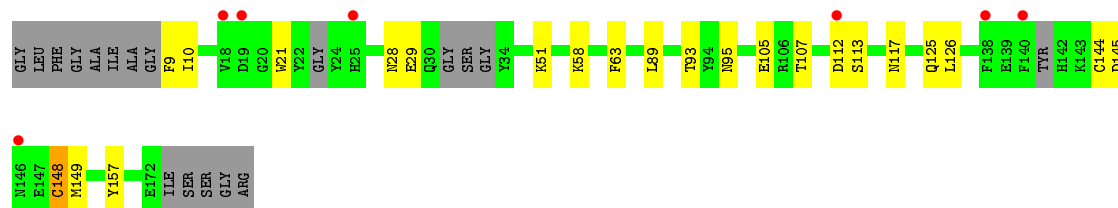
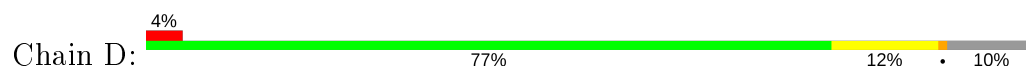




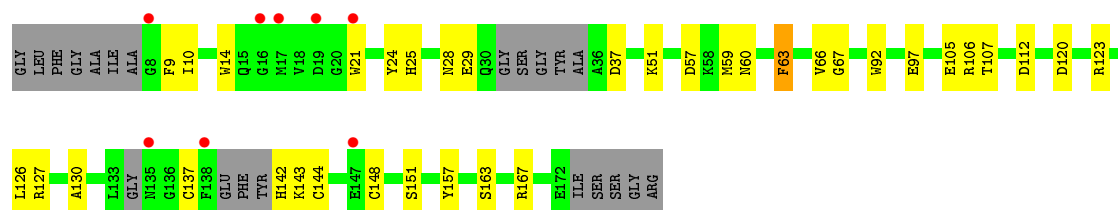
• Molecule 2: Hemagglutinin HA2 chain



• Molecule 2: Hemagglutinin HA2 chain



• Molecule 2: Hemagglutinin HA2 chain



• 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



Chain H:  50% 50%

  
MAG1  
MAG2

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

Chain J:  50% 50%

  
MAG1  
MAG2

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

Chain I:  33% 67%

  
MAG1  
MAG2  
BGL3

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.21Å 231.14Å 70.86Å 90.00° 96.88° 90.00°	Depositor
Resolution (Å)	37.81 – 3.09 37.81 – 3.09	Depositor EDS
% Data completeness (in resolution range)	99.6 (37.81-3.09) 99.6 (37.81-3.09)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.61 (at 3.12Å)	Xtriage
Refinement program	PHENIX (1.12_2829: ???)	Depositor
R, $R_{free}$	0.222 , 0.269 0.222 , 0.269	Depositor DCC
$R_{free}$ test set	1885 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	95.3	Xtriage
Anisotropy	0.073	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 53.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	11548	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	104.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.48% 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: BMA, 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	0.25	0/2535	0.45	0/3440
1	C	0.24	0/2529	0.44	0/3431
1	E	0.24	0/2514	0.44	0/3410
2	B	0.25	0/1409	0.40	0/1894
2	D	0.24	0/1341	0.39	0/1798
2	F	0.24	0/1306	0.38	0/1750
All	All	0.24	0/11634	0.43	0/15723

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 [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	2478	0	2433	43	0
1	C	2473	0	2435	49	0
1	E	2458	0	2413	47	0
2	B	1383	0	1288	25	0
2	D	1319	0	1224	15	0
2	F	1286	0	1198	25	0
3	G	28	0	25	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	H	28	0	25	0	0
3	J	28	0	25	0	0
4	I	39	0	34	0	0
5	A	14	0	13	0	0
5	E	14	0	13	0	0
All	All	11548	0	11126	173	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 8.

All (173) 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:E:283:THR:HG22	1:E:285:MET:H	1.45	0.82
1:E:28:THR:HG22	1:E:30:MET:H	1.48	0.77
1:C:283:THR:HG22	1:C:285:MET:H	1.49	0.76
2:B:51:LYS:NZ	2:B:107:THR:OG1	2.20	0.75
1:C:28:THR:HG22	1:C:30:MET:H	1.53	0.74
1:A:283:THR:HG22	1:A:285:MET:H	1.53	0.72
1:A:28:THR:HG22	1:A:30:MET:H	1.53	0.72
1:C:28:THR:HG23	2:D:105:GLU:HB2	1.73	0.71
1:E:26:VAL:HG21	1:E:317:ALA:HB2	1.72	0.70
1:A:28:THR:HG23	2:B:105:GLU:HB2	1.75	0.69
1:C:13:ILE:HD11	2:D:149:MET:HG2	1.74	0.68
1:E:295:HIS:HD2	1:E:297:ILE:H	1.39	0.68
1:C:26:VAL:HG21	1:C:317:ALA:HB2	1.75	0.67
1:C:288:ILE:HD11	1:C:297:ILE:HG13	1.79	0.65
1:E:20:ASN:O	1:E:322:ASN:ND2	2.29	0.64
1:E:28:THR:HG23	2:F:105:GLU:HB2	1.80	0.64
1:E:191:GLN:NE2	1:E:250:ASN:HD21	1.96	0.64
1:C:211:GLN:HE21	1:C:213:LEU:HD21	1.61	0.64
2:D:51:LYS:NZ	2:D:107:THR:OG1	2.31	0.64
1:E:14:CYS:HB2	2:F:25:HIS:HB3	1.79	0.64
1:E:266:THR:HB	2:F:66:VAL:HB	1.80	0.64
1:A:216:ARG:O	1:A:220:ARG:NH2	2.32	0.63
1:C:206:THR:HB	1:C:209:LEU:HB3	1.81	0.62
1:C:102:PHE:HB3	1:C:105:TYR:HB2	1.82	0.62
1:A:154:LEU:HD12	1:A:251:PHE:HB3	1.83	0.61
1:E:156:LYS:NZ	1:E:196:GLN:OE1	2.30	0.61
2:F:28:ASN:ND2	2:F:144:CYS:O	2.33	0.61
1:E:200:THR:HG23	1:E:248:ASN:HB3	1.82	0.60

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:206:THR:HG22	1:C:208:THR:H	1.66	0.60
1:E:43:LEU:HB2	1:E:314:LEU:HB2	1.84	0.60
1:E:216:ARG:O	1:E:220:ARG:NH2	2.35	0.59
2:B:89:LEU:HD22	2:F:63:PHE:HZ	1.66	0.59
1:E:124:ILE:HD12	1:E:166:ARG:HH21	1.67	0.59
1:C:51:LEU:HD13	1:C:272:LEU:HB2	1.84	0.59
1:A:121:ILE:HG13	1:A:259:LYS:HD2	1.86	0.58
1:E:102:PHE:HB3	1:E:105:TYR:HB2	1.84	0.57
2:B:23:GLY:HA3	2:B:36:ALA:HA	1.87	0.57
2:F:126:LEU:HD13	2:F:130:ALA:HB3	1.85	0.57
1:E:191:GLN:HG3	1:E:195:TYR:HD2	1.70	0.56
2:F:151:SER:HB3	2:F:157:TYR:HA	1.87	0.56
1:A:206:THR:HB	1:A:209:LEU:H	1.70	0.56
1:A:206:THR:HG22	1:A:208:THR:H	1.70	0.56
1:C:18:HIS:HB2	2:D:21:TRP:HA	1.88	0.56
1:A:70:LEU:HG	1:A:179:LEU:HD11	1.87	0.56
2:B:151:SER:HB3	2:B:157:TYR:HA	1.88	0.56
1:E:165:LYS:HG2	1:E:246:GLU:HG3	1.88	0.55
1:E:200:THR:HB	1:E:250:ASN:HD22	1.72	0.55
1:C:147:PHE:HZ	1:C:230:MET:HE1	1.71	0.55
1:A:13:ILE:HD11	2:B:24:TYR:HB3	1.88	0.55
1:E:74:PRO:HB3	1:E:141:TYR:HD1	1.72	0.54
2:B:113:SER:O	2:B:117:ASN:ND2	2.34	0.54
1:C:303:GLY:HA2	2:D:63:PHE:CD1	2.43	0.54
1:A:200:THR:HG23	1:A:248:ASN:HB3	1.89	0.54
2:F:51:LYS:NZ	2:F:107:THR:OG1	2.37	0.54
2:D:28:ASN:ND2	2:D:144:CYS:O	2.38	0.54
1:C:216:ARG:O	1:C:220:ARG:NH2	2.41	0.53
2:F:120:ASP:HA	2:F:123:ARG:HB3	1.90	0.53
1:A:43:LEU:HB2	1:A:314:LEU:HB2	1.90	0.53
1:E:147:PHE:HZ	1:E:230:MET:HE1	1.74	0.53
1:C:111:LEU:O	1:C:261:LYS:NZ	2.36	0.52
1:C:123:ILE:HD11	1:C:254:PRO:HB2	1.91	0.52
1:A:238:LYS:HG3	1:A:239:PRO:HD2	1.91	0.52
2:D:125:GLN:HG2	2:D:157:TYR:HB3	1.90	0.52
2:D:113:SER:O	2:D:117:ASN:ND2	2.38	0.52
2:F:10:ILE:HD11	2:F:14:TRP:CZ2	2.45	0.51
1:E:105:TYR:CZ	1:E:109:LYS:HD2	2.45	0.51
2:D:58:LYS:HD2	2:F:97:GLU:HB3	1.92	0.51
1:E:14:CYS:HA	2:F:137:CYS:HA	1.92	0.51
1:A:191:GLN:O	1:A:195:TYR:N	2.31	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:70:LEU:HG	1:C:179:LEU:HD11	1.92	0.50
2:D:145:ASP:H	2:D:148:CYS:HB3	1.77	0.50
1:E:215:PRO:HG3	1:E:250:ASN:ND2	2.27	0.49
2:B:62:GLN:HE22	2:D:89:LEU:HD22	1.77	0.49
1:E:72:GLY:HA3	1:E:149:ARG:H	1.76	0.49
1:C:124:ILE:HD11	1:C:168:TYR:HB2	1.94	0.48
1:A:199:THR:O	1:A:248:ASN:ND2	2.46	0.48
2:B:132:GLU:HG2	2:B:138:PHE:HE2	1.77	0.48
2:B:94:TYR:CD1	2:F:59:MET:HB2	2.47	0.48
1:E:206:THR:HB	1:E:209:LEU:HB3	1.94	0.48
1:E:295:HIS:CD2	1:E:297:ILE:H	2.26	0.48
1:E:125:PRO:O	1:E:126:SER:OG	2.30	0.48
1:A:310:LYS:HB2	2:B:89:LEU:HD21	1.95	0.48
1:A:18:HIS:HB2	2:B:21:TRP:HA	1.96	0.48
2:F:142:HIS:ND1	2:F:143:LYS:O	2.47	0.47
2:B:126:LEU:HD22	2:B:140:PHE:HE2	1.78	0.47
1:C:138:ALA:HB2	1:C:226:GLN:HG3	1.96	0.47
1:E:206:THR:HB	1:E:209:LEU:H	1.78	0.47
1:C:211:GLN:NE2	1:C:213:LEU:HD21	2.29	0.47
1:E:70:LEU:HG	1:E:179:LEU:HD11	1.95	0.47
1:A:317:ALA:O	2:B:107:THR:HG21	2.14	0.47
1:C:179:LEU:O	1:C:254:PRO:HB3	2.14	0.47
1:C:283:THR:HB	1:C:286:GLY:O	2.14	0.47
1:C:73:ASN:ND2	1:C:96:ASP:O	2.46	0.47
1:A:179:LEU:O	1:A:254:PRO:HB3	2.15	0.47
1:C:156:LYS:NZ	1:C:196:GLN:OE1	2.43	0.47
1:E:18:HIS:HB2	2:F:21:TRP:HA	1.97	0.46
1:E:179:LEU:O	1:E:254:PRO:HB3	2.15	0.46
1:A:200:THR:HA	1:A:248:ASN:HD22	1.81	0.46
2:B:86:ASP:HB3	2:F:63:PHE:CE1	2.50	0.46
1:C:206:THR:HB	1:C:209:LEU:H	1.81	0.46
2:D:144:CYS:SG	2:D:149:MET:HG3	2.55	0.46
1:C:165:LYS:HG2	1:C:246:GLU:HG3	1.96	0.46
1:C:172:ASN:HB3	1:C:259:LYS:HE3	1.96	0.46
1:C:238:LYS:HD2	1:C:238:LYS:HA	1.68	0.46
1:C:20:ASN:O	1:C:322:ASN:ND2	2.49	0.46
2:F:28:ASN:OD1	2:F:29:GLU:N	2.49	0.46
1:A:52:CYS:H	1:A:282:GLN:HE22	1.64	0.45
2:B:151:SER:HB3	2:B:157:TYR:CA	2.46	0.45
1:A:49:GLY:HA2	1:A:285:MET:O	2.16	0.45
1:C:209:LEU:HD11	1:C:211:GLN:HB2	1.97	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:102:PHE:HB3	1:A:105:TYR:HB2	1.98	0.45
1:A:133:SER:N	1:A:155:ILE:O	2.50	0.45
2:B:94:TYR:HD2	2:F:92:TRP:HZ3	1.63	0.45
1:E:283:THR:HB	1:E:286:GLY:O	2.17	0.45
2:F:24:TYR:CE1	2:F:37:ASP:HB2	2.52	0.45
1:E:183:HIS:HB2	1:E:252:ILE:HD11	1.99	0.45
1:C:207:SER:HA	1:E:229:ARG:HH21	1.81	0.45
1:E:63:ASP:HB3	1:E:94:VAL:HG22	1.97	0.44
1:A:201:TYR:CD2	1:A:248:ASN:HB2	2.53	0.44
1:A:206:THR:HB	1:A:209:LEU:HB3	1.98	0.44
1:E:133:SER:HB2	1:E:157:LYS:HD3	2.00	0.44
2:F:57:ASP:O	2:F:60:ASN:HB2	2.18	0.44
2:D:28:ASN:OD1	2:D:29:GLU:N	2.51	0.44
1:A:127:TRP:CZ3	1:A:166:ARG:HG3	2.52	0.44
2:B:106:ARG:HH11	2:F:106:ARG:NH1	2.16	0.44
1:C:35:THR:OG1	1:C:322:ASN:O	2.34	0.44
1:A:207:SER:HA	1:C:229:ARG:HH21	1.83	0.43
1:A:183:HIS:O	1:A:250:ASN:ND2	2.51	0.43
2:B:28:ASN:OD1	2:B:29:GLU:N	2.52	0.43
2:B:9:PHE:CD1	2:B:10:ILE:HG13	2.54	0.43
1:E:238:LYS:HD2	1:E:238:LYS:HA	1.71	0.43
1:E:266:THR:HG21	2:F:67:GLY:N	2.34	0.43
2:B:22:TYR:OH	2:B:111:HIS:ND1	2.36	0.43
1:A:49:GLY:H	1:A:286:GLY:HA3	1.82	0.43
2:B:132:GLU:O	2:F:127:ARG:HD3	2.18	0.43
1:A:67:ALA:HB2	1:A:105:TYR:CE1	2.53	0.42
1:C:150:ASN:HA	1:C:256:TYR:HD1	1.83	0.42
1:E:158:ASN:O	1:E:158:ASN:ND2	2.46	0.42
1:A:150:ASN:HA	1:A:256:TYR:HD1	1.84	0.42
1:C:26:VAL:HG22	1:C:36:VAL:HG21	2.02	0.42
1:E:147:PHE:CZ	1:E:230:MET:HE1	2.55	0.42
1:A:108:LEU:HB2	1:A:234:TRP:CE2	2.54	0.42
1:A:283:THR:HB	1:A:286:GLY:O	2.19	0.42
1:E:203:SER:OG	1:E:246:GLU:HB3	2.19	0.42
1:A:238:LYS:HD2	1:A:238:LYS:HA	1.76	0.42
1:A:52:CYS:H	1:A:282:GLN:NE2	2.18	0.42
1:C:126:SER:C	1:C:166:ARG:HH12	2.23	0.42
1:A:303:GLY:HA2	2:B:63:PHE:CD1	2.54	0.42
1:C:105:TYR:CE2	1:C:109:LYS:HD2	2.55	0.42
1:C:16:GLY:HA2	2:D:10:ILE:HG21	2.02	0.42
1:C:63:ASP:HB3	1:C:94:VAL:HG22	2.01	0.42

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:49:GLY:HA2	1:C:285:MET:O	2.20	0.41
1:E:127:TRP:CZ3	1:E:166:ARG:HD3	2.56	0.41
1:A:200:THR:HG21	1:A:250:ASN:HB2	2.01	0.41
1:A:17:TYR:CZ	2:B:6:ILE:HG23	2.55	0.41
1:C:127:TRP:CZ2	1:C:253:ALA:HB1	2.56	0.41
1:E:200:THR:HB	1:E:250:ASN:ND2	2.35	0.41
1:C:212:ARG:HB3	1:E:216:ARG:HG2	2.01	0.41
2:F:163:SER:O	2:F:167:ARG:HG3	2.21	0.41
1:A:54:LEU:HD11	1:A:282:GLN:HB2	2.02	0.41
1:A:72:GLY:HA3	1:A:149:ARG:H	1.85	0.41
1:C:309:VAL:HG13	2:D:93:THR:HA	2.03	0.41
1:C:70:LEU:HA	1:C:70:LEU:HD12	1.90	0.41
2:F:10:ILE:HD11	2:F:14:TRP:CH2	2.55	0.41
2:B:132:GLU:HG2	2:B:138:PHE:CE2	2.55	0.41
1:C:207:SER:OG	1:C:241:ASP:OD2	2.28	0.41
1:E:209:LEU:HD22	1:E:235:THR:HG21	2.03	0.41
1:A:105:TYR:CZ	1:A:109:LYS:HD2	2.56	0.41
1:A:295:HIS:CD2	1:A:306:PRO:HG2	2.55	0.41
1:C:124:ILE:HA	1:C:125:PRO:HD3	1.96	0.40
1:C:133(A):LEU:HG	1:C:152:VAL:HG13	2.02	0.40
1:E:153:TRP:NE1	1:E:195:TYR:OH	2.50	0.40
1:C:205:GLY:HA3	1:E:219:THR:O	2.21	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	307/334 (92%)	293 (95%)	13 (4%)	1 (0%)	41	73
1	C	307/334 (92%)	290 (94%)	17 (6%)	0	100	100
1	E	305/334 (91%)	294 (96%)	10 (3%)	1 (0%)	41	73

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	168/177 (95%)	159 (95%)	8 (5%)	1 (1%)	25	59
2	D	151/177 (85%)	146 (97%)	5 (3%)	0	100	100
2	F	148/177 (84%)	141 (95%)	7 (5%)	0	100	100
All	All	1386/1533 (90%)	1323 (96%)	60 (4%)	3 (0%)	47	79

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	249	GLY
2	B	63	PHE
1	E	162	PRO

### 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	281/299 (94%)	279 (99%)	2 (1%)	84	93
1	C	280/299 (94%)	277 (99%)	3 (1%)	73	89
1	E	278/299 (93%)	273 (98%)	5 (2%)	59	82
2	B	146/151 (97%)	144 (99%)	2 (1%)	67	86
2	D	142/151 (94%)	137 (96%)	5 (4%)	36	68
2	F	139/151 (92%)	135 (97%)	4 (3%)	42	72
All	All	1266/1350 (94%)	1245 (98%)	21 (2%)	60	83

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

Mol	Chain	Res	Type
1	A	102	PHE
1	A	251	PHE
2	B	112	ASP
2	B	148	CYS
1	C	127	TRP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	187	ASP
1	C	264	ASP
2	D	9	PHE
2	D	95	ASN
2	D	112	ASP
2	D	126	LEU
2	D	148	CYS
1	E	102	PHE
1	E	158	ASN
1	E	240	ASN
1	E	251	PHE
1	E	264	ASP
2	F	9	PHE
2	F	63	PHE
2	F	112	ASP
2	F	148	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	191	GLN
1	E	250	ASN
1	E	295	HIS

### 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 ⓘ

9 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	G	1	1,3	14,14,15	1.07	2 (14%)	17,19,21	1.47	3 (17%)
3	NAG	G	2	3	14,14,15	0.22	0	17,19,21	0.52	0
3	NAG	H	1	1,3	14,14,15	1.17	2 (14%)	17,19,21	1.20	1 (5%)
3	NAG	H	2	3	14,14,15	0.35	0	17,19,21	0.77	0
4	NAG	I	1	1,4	14,14,15	0.89	1 (7%)	17,19,21	1.41	3 (17%)
4	NAG	I	2	4	14,14,15	0.22	0	17,19,21	0.53	0
4	BMA	I	3	4	11,11,12	1.45	2 (18%)	15,15,17	0.89	1 (6%)
3	NAG	J	1	1,3	14,14,15	0.89	1 (7%)	17,19,21	1.41	3 (17%)
3	NAG	J	2	3	14,14,15	0.20	0	17,19,21	0.53	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
3	NAG	G	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	G	2	3	-	2/6/23/26	0/1/1/1
3	NAG	H	1	1,3	-	2/6/23/26	0/1/1/1
3	NAG	H	2	3	-	0/6/23/26	0/1/1/1
4	NAG	I	1	1,4	-	4/6/23/26	0/1/1/1
4	NAG	I	2	4	-	2/6/23/26	0/1/1/1
4	BMA	I	3	4	-	0/2/19/22	0/1/1/1
3	NAG	J	1	1,3	-	4/6/23/26	0/1/1/1
3	NAG	J	2	3	-	2/6/23/26	0/1/1/1

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	I	3	BMA	O5-C1	2.98	1.48	1.43
4	I	3	BMA	O5-C5	2.60	1.48	1.43
3	G	1	NAG	O5-C1	-2.48	1.39	1.43
3	J	1	NAG	O5-C1	-2.48	1.39	1.43
4	I	1	NAG	O5-C1	-2.47	1.39	1.43
3	H	1	NAG	O5-C1	-2.46	1.39	1.43
3	H	1	NAG	C2-N2	-2.13	1.42	1.46

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	G	1	NAG	O7-C7	-2.11	1.18	1.23

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	I	1	NAG	C2-N2-C7	-3.38	118.09	122.90
3	J	1	NAG	C2-N2-C7	-3.37	118.10	122.90
3	G	1	NAG	C1-C2-N2	-3.36	104.75	110.49
3	G	1	NAG	C6-C5-C4	-3.00	105.97	113.00
4	I	1	NAG	C6-C5-C4	-2.68	106.72	113.00
3	J	1	NAG	C6-C5-C4	-2.67	106.74	113.00
3	G	1	NAG	O3-C3-C2	2.49	114.63	109.47
3	J	1	NAG	C1-C2-N2	-2.43	106.33	110.49
4	I	1	NAG	C1-C2-N2	-2.43	106.34	110.49
3	H	1	NAG	O5-C5-C6	-2.20	103.76	107.20
4	I	3	BMA	C1-O5-C5	2.14	115.09	112.19

There are no chirality outliers.

All (16) torsion outliers are listed below:

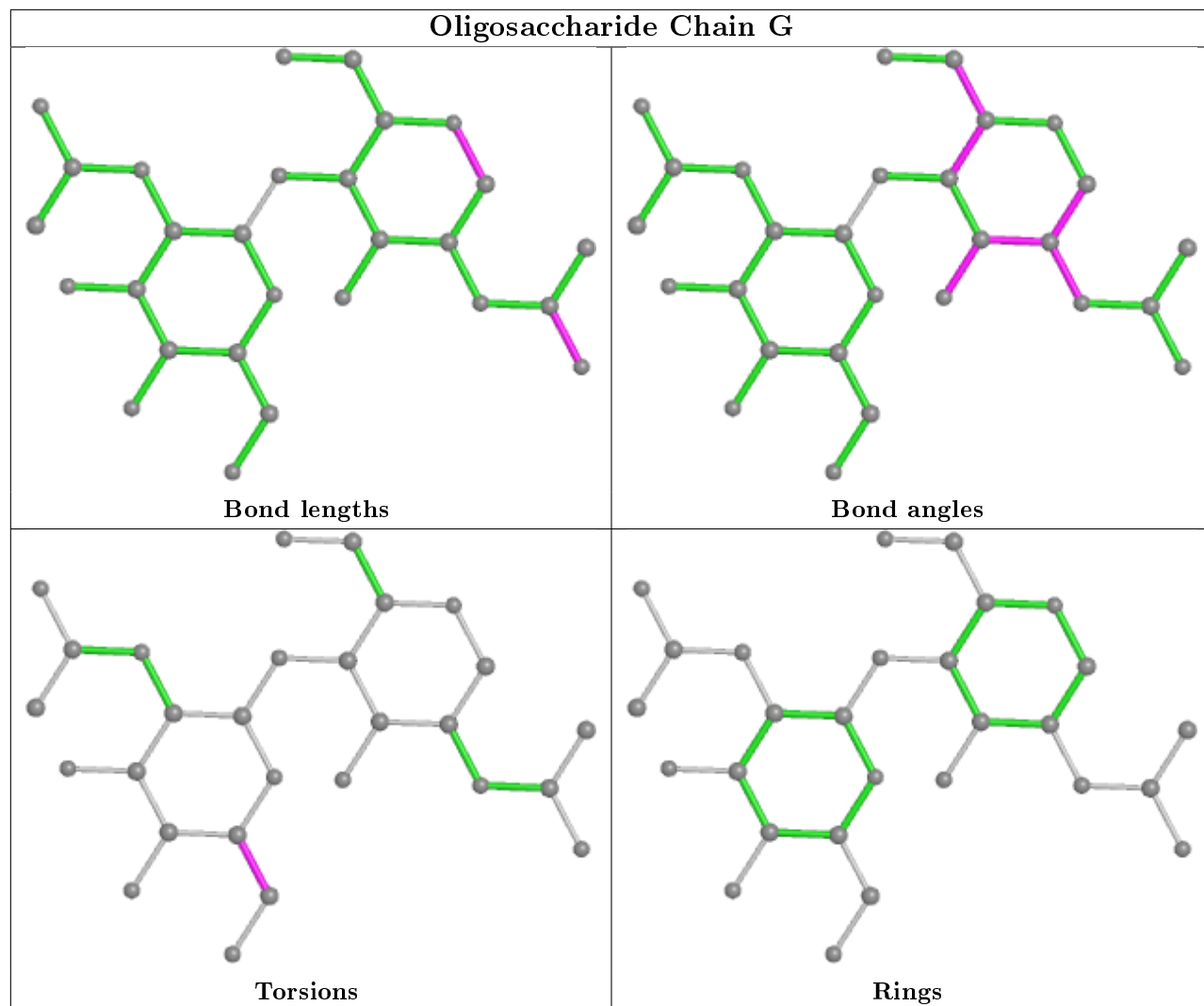
Mol	Chain	Res	Type	Atoms
3	H	1	NAG	O5-C5-C6-O6
3	H	1	NAG	C4-C5-C6-O6
3	J	2	NAG	O5-C5-C6-O6
3	G	2	NAG	O5-C5-C6-O6
4	I	2	NAG	O5-C5-C6-O6
3	G	2	NAG	C4-C5-C6-O6
3	J	2	NAG	C4-C5-C6-O6
4	I	2	NAG	C4-C5-C6-O6
4	I	1	NAG	C8-C7-N2-C2
3	J	1	NAG	C8-C7-N2-C2
4	I	1	NAG	O7-C7-N2-C2
3	J	1	NAG	O7-C7-N2-C2
3	J	1	NAG	C4-C5-C6-O6
4	I	1	NAG	C4-C5-C6-O6
3	J	1	NAG	O5-C5-C6-O6
4	I	1	NAG	O5-C5-C6-O6

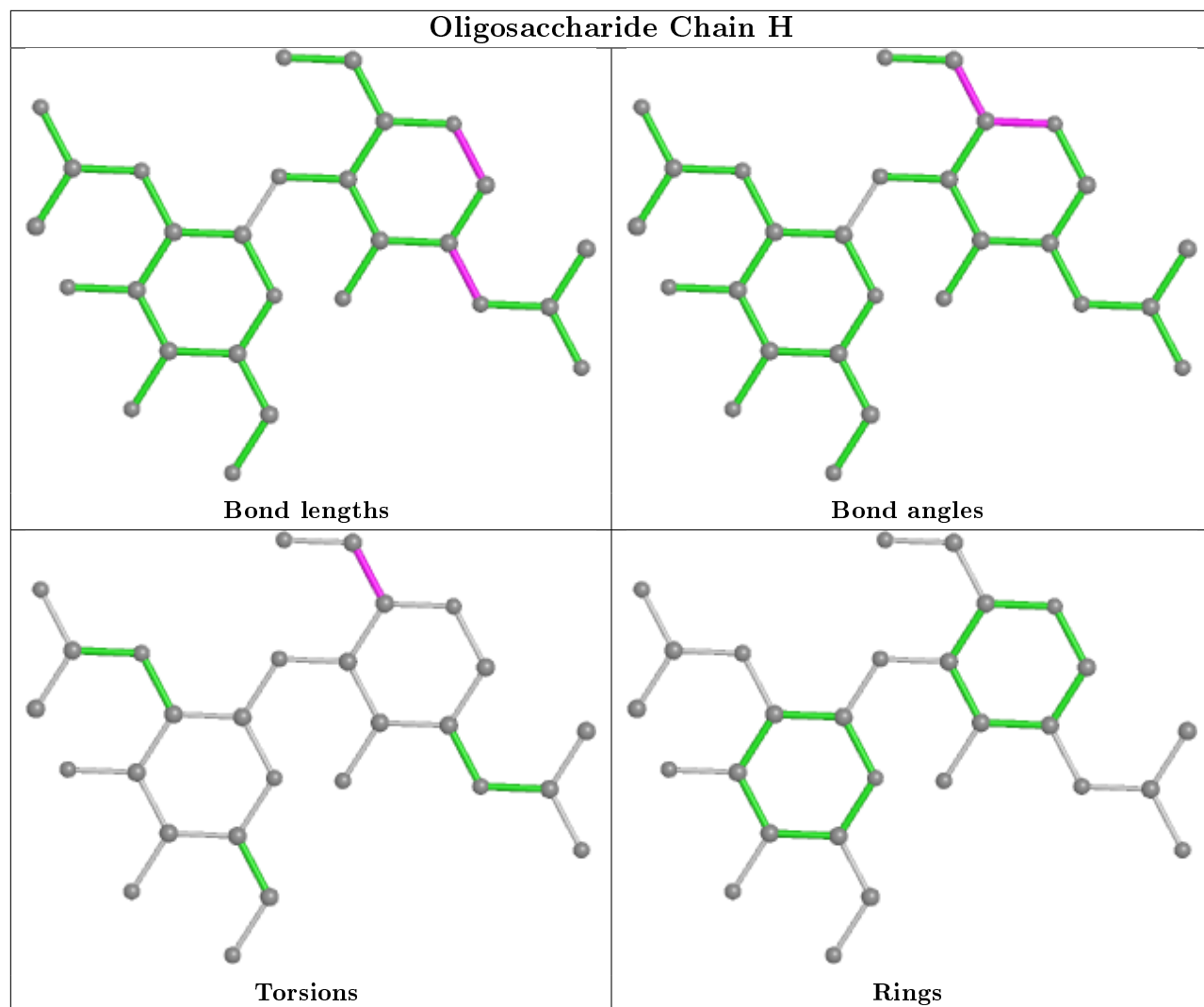
There are no ring outliers.

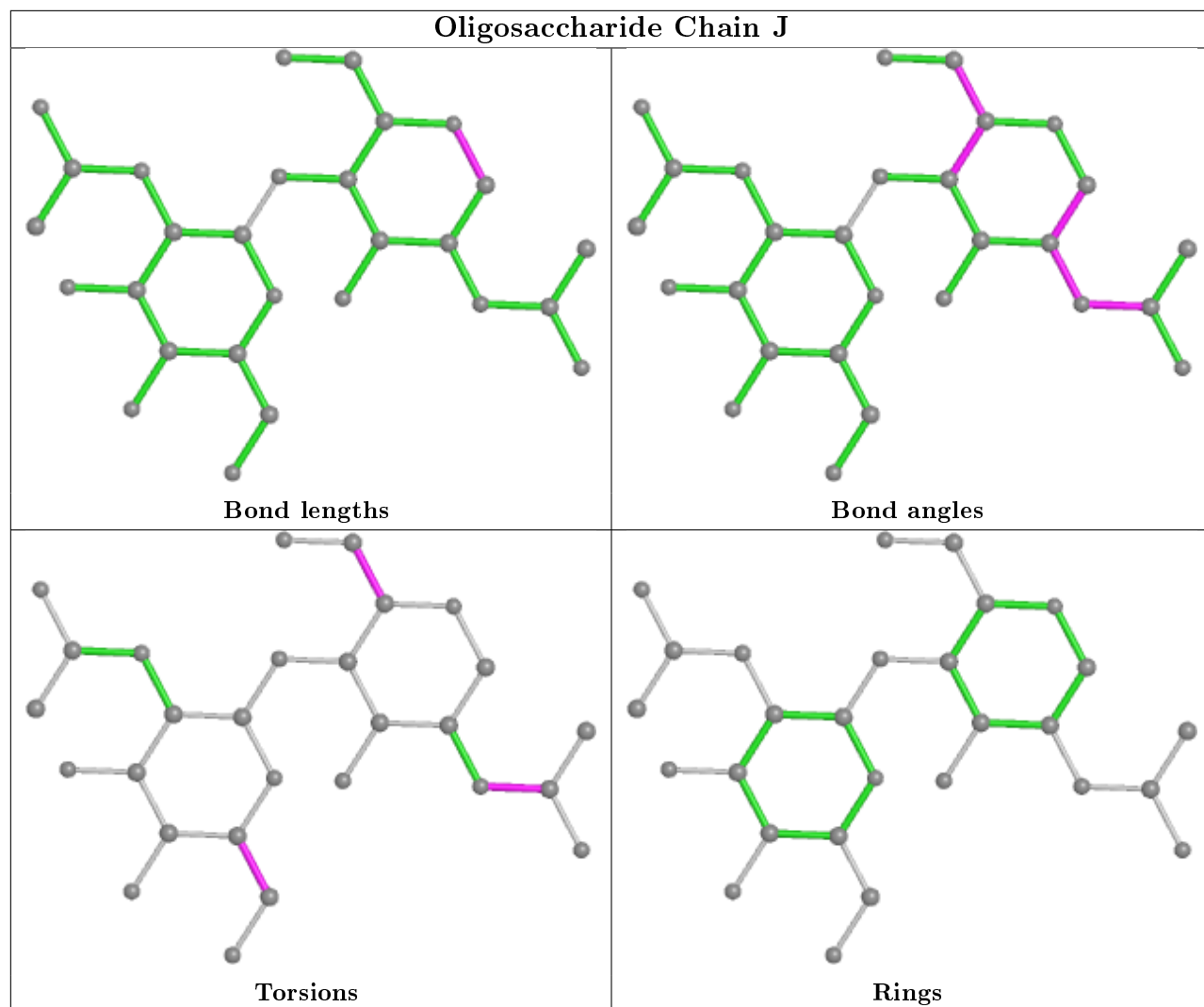
No monomer is involved in short contacts.

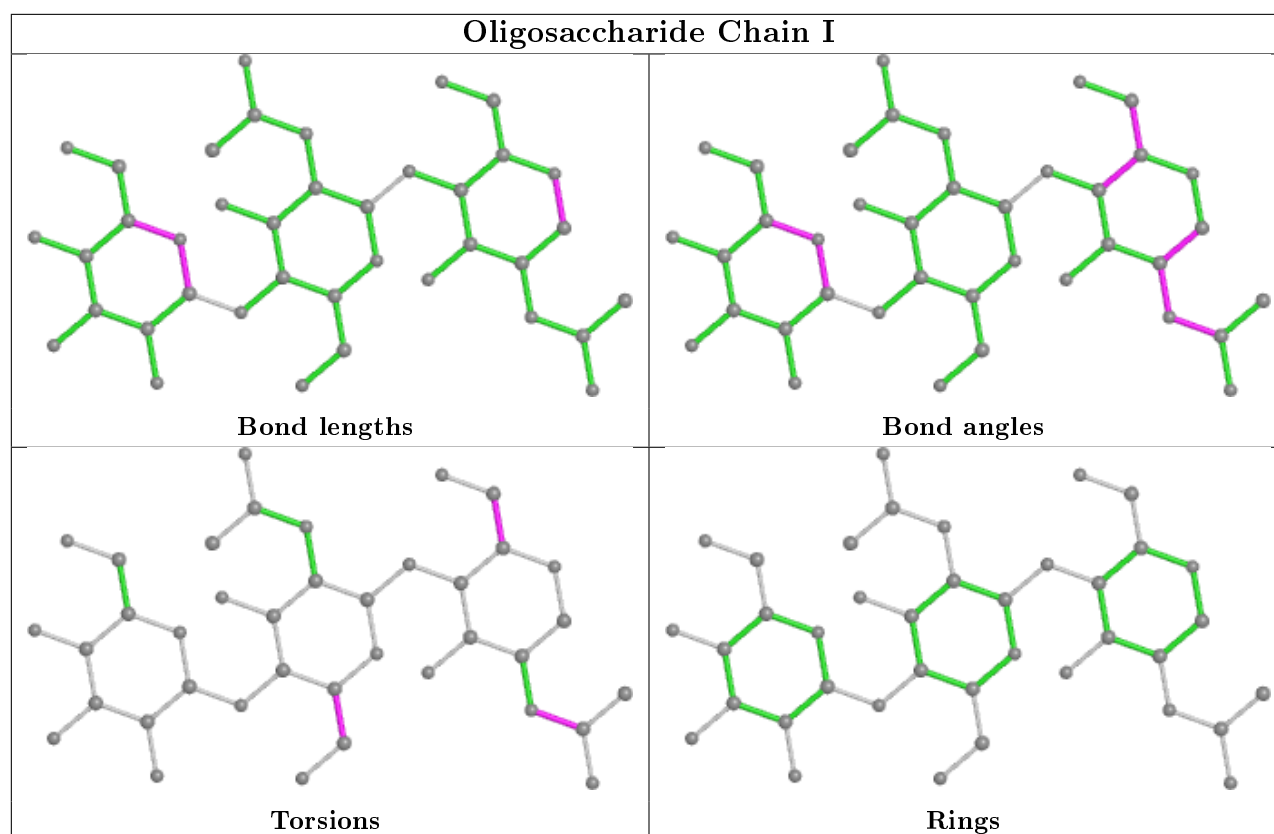
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths,

bond angles, torsion angles, and ring geometry for oligosaccharide.









## 5.6 Ligand geometry [i](#)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
5	NAG	A	401	1	14,14,15	1.16	2 (14%)	17,19,21	1.21	1 (5%)
5	NAG	E	601	1	14,14,15	1.17	2 (14%)	17,19,21	1.21	1 (5%)

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
5	NAG	A	401	1	-	2/6/23/26	0/1/1/1
5	NAG	E	601	1	-	2/6/23/26	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	601	NAG	O5-C1	-2.43	1.39	1.43
5	A	401	NAG	O5-C1	-2.38	1.39	1.43
5	E	601	NAG	C2-N2	-2.13	1.42	1.46
5	A	401	NAG	C2-N2	-2.10	1.42	1.46

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	401	NAG	O5-C5-C6	-2.21	103.74	107.20
5	E	601	NAG	O5-C5-C6	-2.20	103.76	107.20

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	401	NAG	O5-C5-C6-O6
5	E	601	NAG	O5-C5-C6-O6
5	A	401	NAG	C4-C5-C6-O6
5	E	601	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

## 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	313/334 (93%)	-0.11	5 (1%)	72 51	61, 95, 126, 169	0
1	C	313/334 (93%)	-0.12	4 (1%)	77 59	59, 94, 132, 170	0
1	E	311/334 (93%)	-0.08	4 (1%)	77 59	61, 95, 129, 159	0
2	B	170/177 (96%)	0.19	5 (2%)	51 28	59, 117, 151, 171	0
2	D	159/177 (89%)	0.33	7 (4%)	34 17	58, 126, 169, 180	0
2	F	156/177 (88%)	0.39	8 (5%)	28 13	57, 126, 173, 189	0
All	All	1422/1533 (92%)	0.03	33 (2%)	60 39	57, 99, 154, 189	0

All (33) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	8	GLY	5.4
1	C	13	ILE	3.9
2	D	19	ASP	3.8
2	B	4	GLY	3.4
2	D	140	PHE	2.9
1	A	324	PRO	2.7
2	B	21	TRP	2.7
1	A	158	ASN	2.6
1	C	159	SER	2.6
1	E	159	SER	2.5
2	B	20	GLY	2.5
2	D	18	VAL	2.5
1	A	21	ASN	2.5
1	A	276	ASN	2.5
2	B	19	ASP	2.4
1	E	21	ASN	2.4
2	F	147	GLU	2.3
2	B	167	ARG	2.3
2	F	21	TRP	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	F	19	ASP	2.3
2	D	138	PHE	2.3
2	F	138	PHE	2.3
2	F	17	MET	2.3
2	D	146	ASN	2.3
1	C	132	ALA	2.3
1	A	81	ASN	2.2
1	C	21	ASN	2.2
2	D	25	HIS	2.1
1	E	323	SER	2.1
1	E	26	VAL	2.1
2	F	16	GLY	2.1
2	D	112	ASP	2.0
2	F	135	ASN	2.0

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

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

## 6.3 Carbohydrates [i](#)

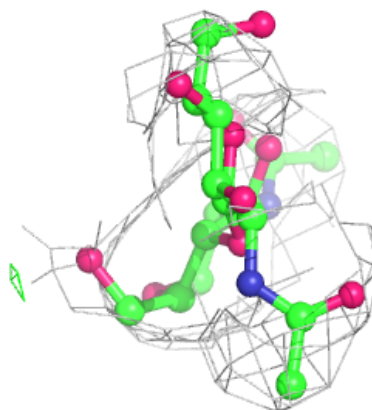
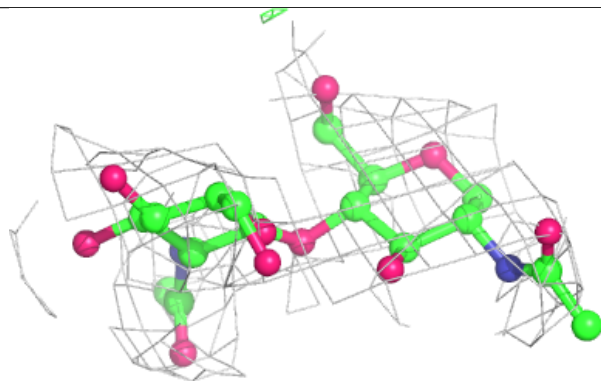
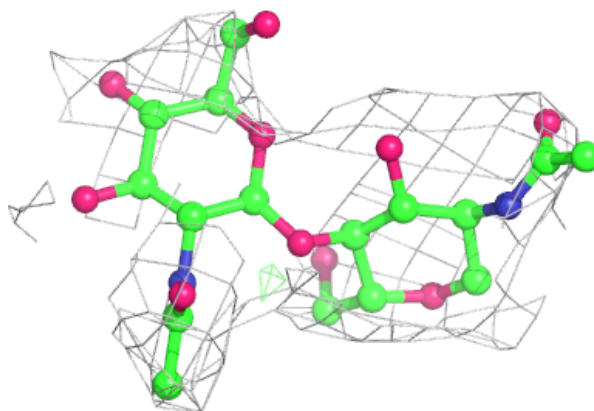
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	BMA	I	3	11/12	0.80	0.15	121,138,157,164	0
3	NAG	H	1	14/15	0.83	0.46	170,179,199,205	0
3	NAG	G	2	14/15	0.83	0.32	117,148,162,167	0
4	NAG	I	1	14/15	0.83	0.21	90,115,125,131	0
3	NAG	J	1	14/15	0.86	0.17	102,116,133,140	0
3	NAG	H	2	14/15	0.87	0.37	131,172,182,184	0
4	NAG	I	2	14/15	0.88	0.17	132,149,157,164	0
3	NAG	J	2	14/15	0.90	0.23	108,133,151,156	0
3	NAG	G	1	14/15	0.91	0.24	117,123,134,135	0

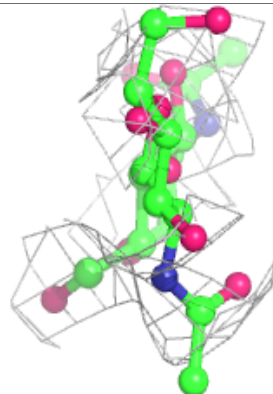
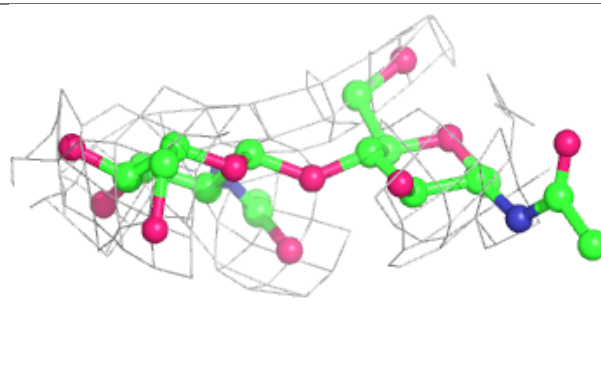
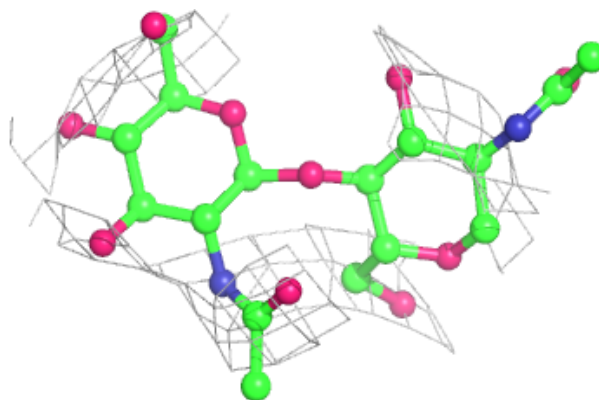
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around Chain G:**

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

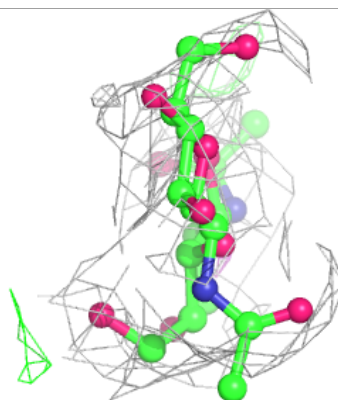
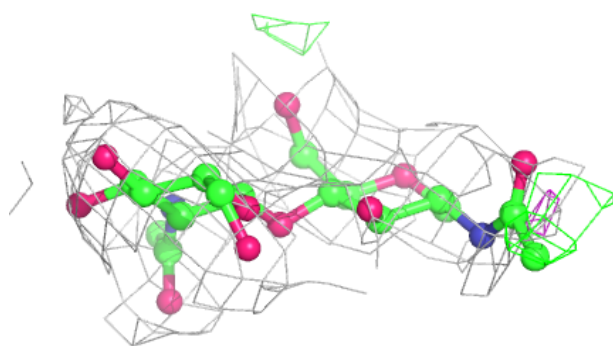
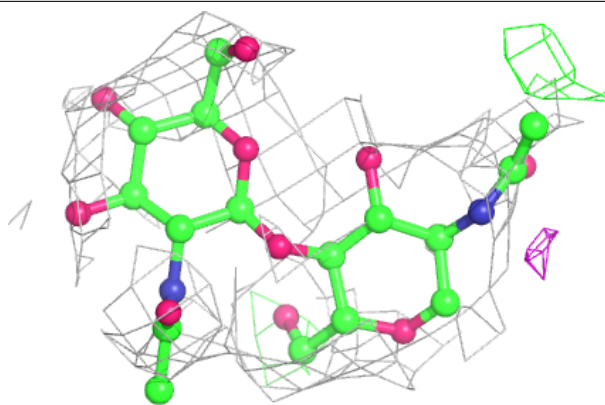
**Electron density around Chain H:**

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

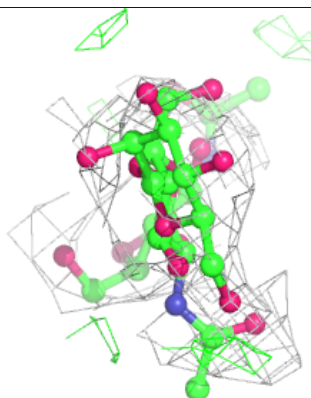
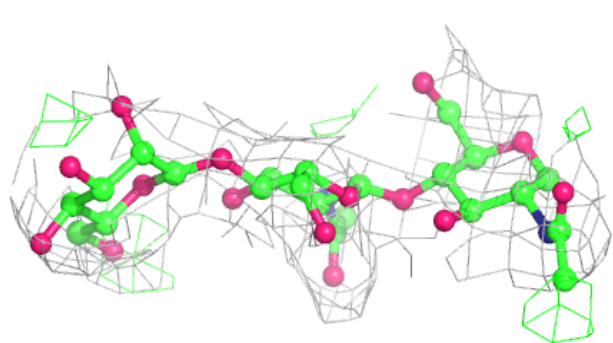
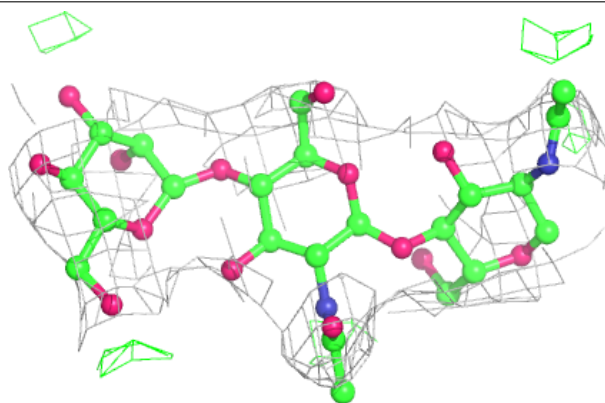


**Electron density around Chain J:**

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

**Electron density around Chain I:**

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



## 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	NAG	A	401	14/15	0.76	0.26	123,148,155,159	0
5	NAG	E	601	14/15	0.78	0.28	131,152,165,169	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.