



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 10, 2020 – 08:36 AM BST

PDB ID : 6E7H  
Title : Crystal structure of H5 hemagglutinin mutant Y161A from A/Viet Nam/1203/2004 H5N1 influenza virus in complex with 3'-GcLN  
Authors : Tzarum, N.; Wilson, I.A.  
Deposited on : 2018-07-26  
Resolution : 3.30 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

---

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

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

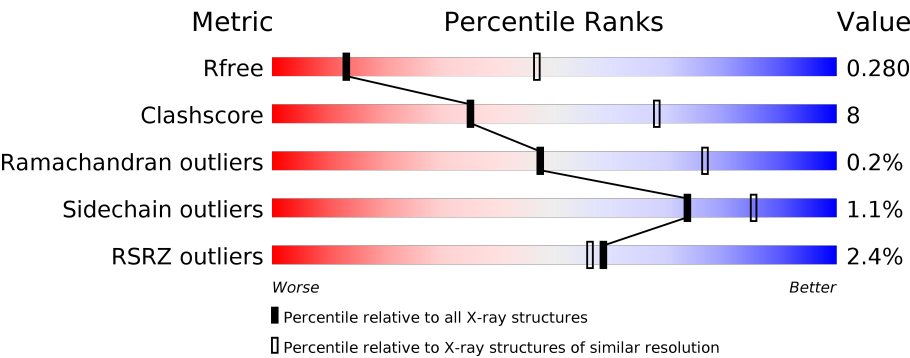
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.30 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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>72%</div><div>22%</div><div>7%</div></div></div>
1	C	334	<div><div>%</div><div><div></div><div>74%</div><div>19%</div><div>7%</div></div></div>
1	E	334	<div><div>%</div><div><div></div><div>70%</div><div>22%</div><div>8%</div></div></div>
2	B	177	<div><div>3%</div><div><div></div><div>72%</div><div>12%</div><div>16%</div></div></div>
2	D	177	<div><div>5%</div><div><div></div><div>69%</div><div>11%</div><div>20%</div></div></div>
2	F	177	<div><div>6%</div><div><div></div><div>71%</div><div>15%</div><div>14%</div></div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	2	<div><div></div><div>50%</div><div></div><div>50%</div></div>
3	J	2	<div><div></div><div>50%</div><div></div><div>50%</div></div>
4	H	3	<div><div></div><div>33%</div><div></div><div>67%</div></div>
5	I	2	<div><div></div><div>50%</div><div></div><div>50%</div></div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 11163 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	312	Total	C	N	O	S	0	0	0
			2466	1557	428	466	15			
1	C	310	Total	C	N	O	S	0	0	0
			2451	1548	426	462	15			
1	E	308	Total	C	N	O	S	0	0	0
			2435	1539	421	460	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	149	Total	C	N	O	S	0	0	0
			1214	750	211	245	8			

*Continued on next page...*

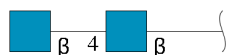
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	142	Total	C	N	O	S	0	0	0
			1175	728	204	235	8			
2	F	153	Total	C	N	O	S	0	0	0
			1266	778	223	257	8			

There are 9 discrepancies between the modelled and reference sequences:

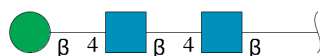
Chain	Residue	Modelled	Actual	Comment	Reference
B	175	SER	-	expression tag	UNP Q6DQ18
B	176	GLY	-	expression tag	UNP Q6DQ18
B	177	ARG	-	expression tag	UNP Q6DQ18
D	175	SER	-	expression tag	UNP Q6DQ18
D	176	GLY	-	expression tag	UNP Q6DQ18
D	177	ARG	-	expression tag	UNP Q6DQ18
F	175	SER	-	expression tag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



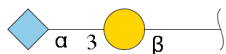
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	2	Total	C	N	O		0	0	0
			28	16	2	10				
3	J	2	Total	C	N	O		0	0	0
			28	16	2	10				

- 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.



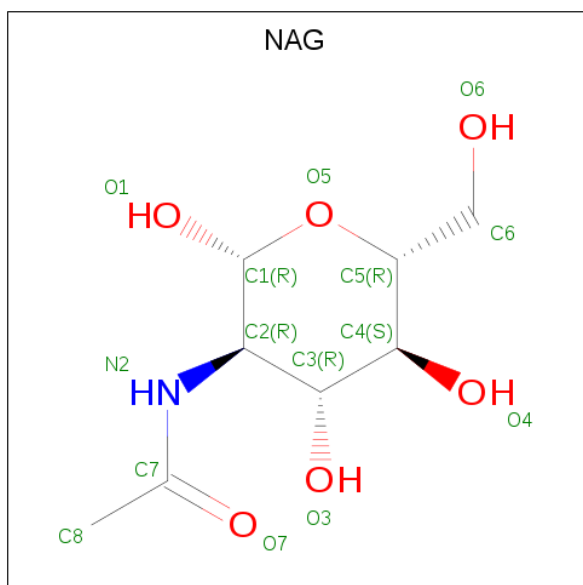
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	H	3	Total	C	N	O		0	0	0
			39	22	2	15				

- Molecule 5 is an oligosaccharide called N-glycolyl- $\alpha$ -neuraminic acid-(2-3)- $\beta$ -D-galactopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	I	2	Total	C	N	O	0	0	0
			33	17	1	15			

- Molecule 6 is 2-acetamido-2-deoxy- $\beta$ -D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).

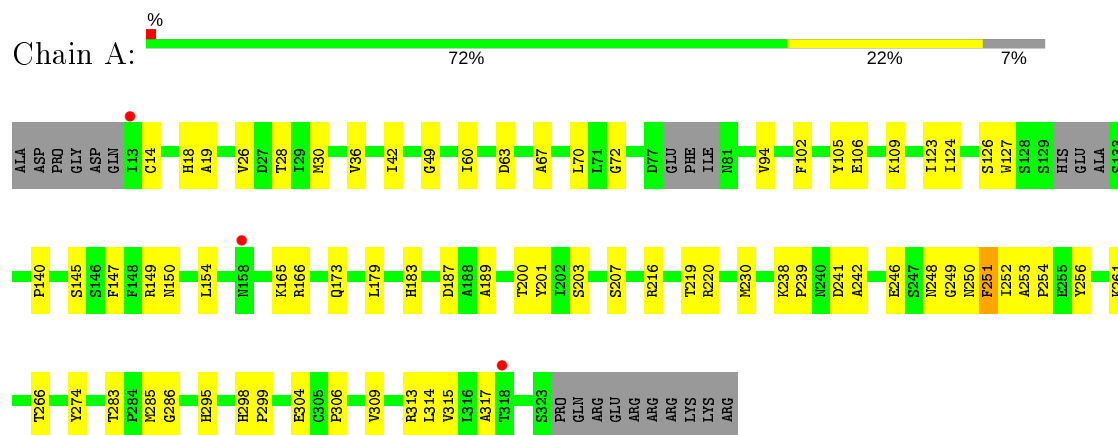


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	E	1	Total	C	N	O	0	0
			14	8	1	5		

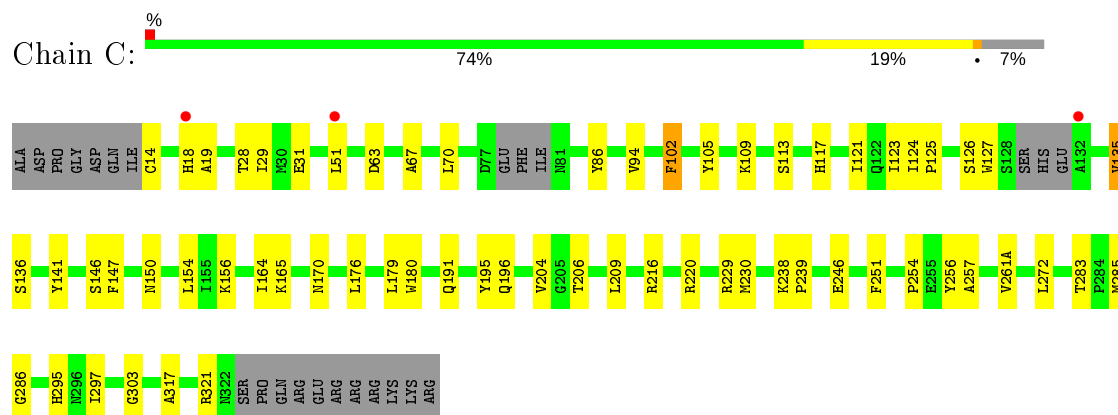
### 3 Residue-property plots [i](#)

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.

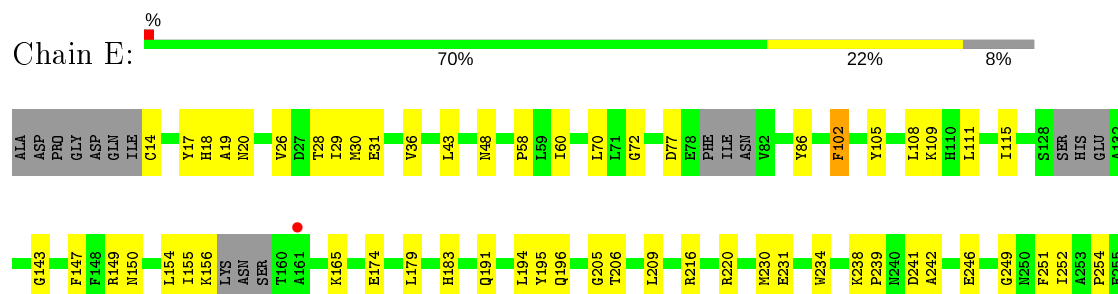
#### • Molecule 1: Hemagglutinin HA1 chain

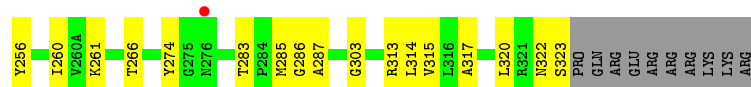


#### • Molecule 1: Hemagglutinin HA1 chain



#### • Molecule 1: Hemagglutinin HA1 chain





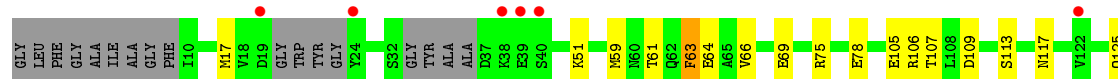
- Molecule 2: Hemagglutinin HA2 chain



- 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








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

Chain H:  33% 67%



- Molecule 5: N-glycolyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose

Chain I:  50% 50%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.56 Å 231.07 Å 71.30 Å 90.00° 97.04° 90.00°	Depositor
Resolution (Å)	47.07 – 3.30 47.07 – 3.30	Depositor EDS
% Data completeness (in resolution range)	97.1 (47.07-3.30) 97.1 (47.07-3.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.52 (at 3.33 Å)	Xtriage
Refinement program	PHENIX (1.12_2829: ???)	Depositor
R, $R_{free}$	0.239 , 0.280 0.239 , 0.280	Depositor DCC
$R_{free}$ test set	1656 reflections (5.32%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	94.6	Xtriage
Anisotropy	0.265	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 35.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	11163	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	100.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.88% 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: GAL, NGC, 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	0.24	0/2522	0.45	0/3421
1	C	0.24	0/2507	0.45	0/3401
1	E	0.24	0/2490	0.44	0/3377
2	B	0.24	0/1234	0.38	0/1656
2	D	0.24	0/1194	0.39	0/1601
2	F	0.23	0/1284	0.37	0/1719
All	All	0.24	0/11231	0.42	0/15175

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	2466	0	2424	53	0
1	C	2451	0	2409	45	0
1	E	2435	0	2388	51	0
2	B	1214	0	1125	22	0
2	D	1175	0	1084	21	0
2	F	1266	0	1177	28	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	J	28	0	25	0	0
4	H	39	0	34	0	0
5	I	33	0	19	2	0
6	A	14	0	13	0	0
6	E	14	0	13	0	0
All	All	11163	0	10736	180	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 (180) 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:C:14:CYS:HB2	2:D:25:HIS:HB3	1.59	0.83
2:D:51:LYS:NZ	2:D:107:THR:OG1	2.12	0.81
1:E:28:THR:HG22	1:E:30:MET:H	1.45	0.81
2:B:75:ARG:HH12	1:E:261:LYS:HE2	1.46	0.80
1:C:135:VAL:N	5:I:2:NGC:O11	2.13	0.79
1:A:283:THR:HG22	1:A:285:MET:H	1.51	0.75
2:B:151:SER:HB3	2:B:157:TYR:HA	1.67	0.74
2:F:151:SER:HB3	2:F:157:TYR:HA	1.69	0.73
1:C:283:THR:HG22	1:C:285:MET:H	1.54	0.71
1:E:283:THR:HG22	1:E:285:MET:H	1.56	0.70
1:C:19:ALA:HB3	2:D:15:GLN:HA	1.74	0.70
1:E:28:THR:HG23	2:F:105:GLU:HB2	1.74	0.70
1:E:150:ASN:HA	1:E:256:TYR:HD1	1.60	0.67
1:A:102:PHE:HB3	1:A:105:TYR:HB2	1.75	0.66
1:C:102:PHE:HB3	1:C:105:TYR:HB2	1.78	0.65
1:A:28:THR:HG22	1:A:30:MET:H	1.60	0.65
1:C:156:LYS:NZ	1:C:196:GLN:OE1	2.30	0.64
1:A:126:SER:O	1:A:166:ARG:NH1	2.29	0.64
1:C:18:HIS:HB2	2:D:21:TRP:HA	1.81	0.64
2:B:75:ARG:NH1	1:E:261:LYS:HE2	2.13	0.64
1:E:266:THR:HB	2:F:66:VAL:HB	1.79	0.63
1:A:200:THR:HG23	1:A:248:ASN:HB3	1.81	0.62
1:E:60:ILE:HD12	1:E:274:TYR:HB2	1.81	0.61
1:A:261:LYS:HZ1	2:D:75:ARG:NH1	1.99	0.60
1:E:156:LYS:NZ	1:E:196:GLN:OE1	2.28	0.60
1:A:123:ILE:HG13	1:A:124:ILE:HG13	1.85	0.59
1:A:261:LYS:NZ	2:D:75:ARG:NH1	2.51	0.59
1:E:191:GLN:O	1:E:195:TYR:N	2.33	0.58

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:206:THR:HB	1:E:209:LEU:H	1.68	0.58
1:C:105:TYR:CZ	1:C:109:LYS:HD2	2.38	0.58
1:A:105:TYR:CZ	1:A:109:LYS:HD2	2.40	0.57
1:C:14:CYS:HA	2:D:137:CYS:HA	1.87	0.57
2:D:106:ARG:NH1	2:F:106:ARG:HH11	2.03	0.56
2:D:68:ARG:HH12	2:D:85:GLU:CD	2.09	0.56
1:A:165:LYS:HE2	1:A:246:GLU:OE2	2.06	0.56
1:A:216:ARG:O	1:A:220:ARG:NH2	2.38	0.56
1:A:207:SER:HA	1:C:229:ARG:HH21	1.70	0.56
1:E:102:PHE:HB3	1:E:105:TYR:HB2	1.88	0.56
1:C:295:HIS:HD2	1:C:297:ILE:H	1.54	0.55
1:E:155:ILE:HD13	1:E:194:LEU:HD22	1.88	0.55
1:A:165:LYS:HG2	1:A:246:GLU:HG3	1.88	0.55
2:F:113:SER:O	2:F:117:ASN:ND2	2.35	0.55
2:D:106:ARG:HH11	2:F:106:ARG:NH1	2.04	0.55
1:E:111:LEU:O	1:E:261:LYS:NZ	2.28	0.55
2:D:106:ARG:NH1	2:F:106:ARG:NH1	2.55	0.55
1:A:313:ARG:HG3	1:A:315:VAL:HG23	1.89	0.54
1:A:70:LEU:HG	1:A:179:LEU:HD11	1.88	0.54
1:C:135:VAL:H	5:I:2:NGC:HO11	1.50	0.54
1:A:304:GLU:HG3	2:B:63:PHE:HB2	1.90	0.53
1:E:20:ASN:O	1:E:322:ASN:ND2	2.39	0.53
1:C:321:ARG:HH21	2:D:108:LEU:HB3	1.72	0.53
1:A:309:VAL:HG13	2:B:93:THR:HA	1.91	0.53
2:B:75:ARG:O	2:B:79:ASN:N	2.36	0.52
2:B:94:TYR:HB2	2:F:59:MET:HG3	1.92	0.52
2:B:106:ARG:HH22	2:D:105:GLU:HG2	1.75	0.52
1:A:261:LYS:NZ	2:D:75:ARG:HH12	2.07	0.52
1:E:322:ASN:OD1	1:E:323:SER:N	2.42	0.52
2:B:126:LEU:HD13	2:B:130:ALA:HB3	1.90	0.52
2:B:75:ARG:HA	2:B:78:GLU:HB3	1.91	0.52
1:C:154:LEU:HD13	1:C:164:ILE:HD13	1.90	0.52
2:F:125:GLN:HG2	2:F:157:TYR:HB3	1.92	0.51
2:B:125:GLN:HG2	2:B:157:TYR:HB3	1.92	0.51
1:A:238:LYS:HG3	1:A:239:PRO:HD2	1.92	0.51
1:A:26:VAL:HG22	1:A:36:VAL:HG21	1.92	0.50
1:C:179:LEU:O	1:C:254:PRO:HB3	2.11	0.50
1:E:43:LEU:HB2	1:E:314:LEU:HB2	1.93	0.50
1:E:26:VAL:HG21	1:E:317:ALA:HB2	1.94	0.50
2:F:126:LEU:HD13	2:F:130:ALA:HB3	1.93	0.50
1:E:165:LYS:HE2	1:E:246:GLU:OE2	2.12	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:150:ASN:HA	1:A:256:TYR:HD1	1.76	0.49
1:C:147:PHE:HZ	1:C:230:MET:HE1	1.76	0.49
1:C:125:PRO:O	1:C:126:SER:OG	2.30	0.49
1:A:295:HIS:CD2	1:A:306:PRO:HG2	2.48	0.49
1:E:18:HIS:ND1	2:F:17:MET:O	2.46	0.49
1:E:70:LEU:HG	1:E:179:LEU:HD11	1.93	0.49
1:E:179:LEU:O	1:E:254:PRO:HB3	2.13	0.49
2:F:51:LYS:NZ	2:F:107:THR:OG1	2.38	0.48
1:A:200:THR:HG21	1:A:250:ASN:H	1.78	0.48
1:E:154:LEU:HD12	1:E:251:PHE:HB3	1.96	0.48
1:C:123:ILE:HD13	1:C:257:ALA:HB2	1.96	0.48
1:A:241:ASP:OD1	1:A:242:ALA:N	2.47	0.47
1:A:313:ARG:NH1	1:A:315:VAL:HG21	2.29	0.47
1:A:28:THR:HG23	2:B:105:GLU:HB2	1.96	0.47
1:A:72:GLY:HA3	1:A:149:ARG:H	1.79	0.47
1:C:206:THR:HB	1:C:209:LEU:H	1.79	0.47
1:A:147:PHE:HZ	1:A:230:MET:HE1	1.79	0.46
1:C:123:ILE:HG13	1:C:124:ILE:HD12	1.98	0.46
1:C:150:ASN:HA	1:C:256:TYR:HD1	1.79	0.46
1:A:49:GLY:HA2	1:A:285:MET:O	2.15	0.46
1:E:238:LYS:HD2	1:E:238:LYS:HA	1.62	0.46
1:A:283:THR:HB	1:A:286:GLY:O	2.16	0.46
1:C:63:ASP:HB3	1:C:94:VAL:HG22	1.97	0.46
1:E:29:ILE:HG22	2:F:105:GLU:OE1	2.16	0.46
2:B:83:LYS:HD3	2:F:66:VAL:HG22	1.98	0.46
1:A:187:ASP:OD2	1:A:189:ALA:HB3	2.16	0.46
1:C:295:HIS:CD2	1:C:297:ILE:H	2.33	0.46
2:D:75:ARG:HA	2:D:78:GLU:HB3	1.97	0.45
1:A:14:CYS:N	2:B:25:HIS:O	2.48	0.45
1:C:283:THR:HG22	1:C:285:MET:N	2.28	0.45
1:A:203:SER:OG	1:A:246:GLU:HB3	2.17	0.45
1:E:183:HIS:HD1	1:E:195:TYR:HE2	1.63	0.45
1:E:241:ASP:OD1	1:E:242:ALA:N	2.46	0.45
2:F:132:GLU:HG2	2:F:138:PHE:HE2	1.82	0.45
1:C:121:ILE:O	1:C:257:ALA:N	2.50	0.45
2:D:106:ARG:HH11	2:F:106:ARG:HH11	1.62	0.45
1:E:216:ARG:O	1:E:220:ARG:NH2	2.50	0.45
1:A:18:HIS:CG	1:A:19:ALA:N	2.86	0.44
1:C:191:GLN:HG3	1:C:195:TYR:HD2	1.81	0.44
1:E:216:ARG:NH1	1:E:231:GLU:OE1	2.46	0.44
1:E:14:CYS:HA	2:F:137:CYS:HA	1.98	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:94:TYR:CD1	2:F:59:MET:HB2	2.52	0.44
1:C:117:HIS:HB3	1:C:261(A):VAL:HB	1.99	0.44
1:C:28:THR:HG23	1:C:31:GLU:H	1.82	0.44
2:F:75:ARG:HA	2:F:78:GLU:HB3	2.00	0.44
1:A:154:LEU:HD12	1:A:251:PHE:HB3	1.99	0.44
1:A:19:ALA:HB3	2:B:15:GLN:HA	1.99	0.44
1:E:115:ILE:HD12	1:E:260:ILE:HG23	1.99	0.44
1:E:147:PHE:HZ	1:E:230:MET:HE3	1.82	0.43
1:A:183:HIS:HB2	1:A:252:ILE:HD11	2.00	0.43
2:F:163:SER:O	2:F:167:ARG:HG3	2.19	0.43
1:E:313:ARG:NH1	1:E:315:VAL:HG21	2.34	0.43
1:C:180:TRP:NE1	1:C:204:VAL:HG21	2.34	0.43
2:D:150:GLU:HA	2:D:153:ARG:HH11	1.83	0.43
1:E:72:GLY:O	1:E:149:ARG:HG3	2.19	0.43
1:C:216:ARG:O	1:C:220:ARG:NH2	2.52	0.43
1:E:28:THR:HB	1:E:31:GLU:HB2	2.01	0.43
1:E:77:ASP:OD1	1:E:77:ASP:N	2.52	0.43
1:A:63:ASP:HB3	1:A:94:VAL:HG22	2.00	0.43
1:C:303:GLY:HA2	2:D:63:PHE:CD1	2.54	0.43
1:E:18:HIS:CG	1:E:19:ALA:N	2.87	0.42
1:E:48:ASN:HD21	1:E:287:ALA:HB3	1.84	0.42
1:C:283:THR:HB	1:C:286:GLY:O	2.20	0.42
2:F:105:GLU:HG2	2:F:109:ASP:OD1	2.18	0.42
1:A:67:ALA:HB2	1:A:105:TYR:CE1	2.54	0.42
1:C:70:LEU:HG	1:C:179:LEU:HD11	2.01	0.42
2:D:68:ARG:NH1	2:D:85:GLU:OE2	2.42	0.42
1:A:179:LEU:O	1:A:254:PRO:HB3	2.20	0.42
2:B:86:ASP:HB3	2:F:63:PHE:CD1	2.54	0.42
1:A:127:TRP:CZ2	1:A:253:ALA:HB1	2.54	0.42
1:A:42:ILE:HG12	1:A:314:LEU:O	2.19	0.42
1:E:58:PRO:HB3	1:E:86:TYR:CE2	2.54	0.42
2:B:51:LYS:HG3	1:C:29:ILE:HG12	2.01	0.42
2:B:89:LEU:HD22	2:F:63:PHE:HZ	1.85	0.42
1:C:147:PHE:CZ	1:C:230:MET:HE1	2.54	0.41
1:C:238:LYS:HD2	1:C:238:LYS:HA	1.68	0.41
1:A:219:THR:O	1:E:205:GLY:HA3	2.20	0.41
2:B:86:ASP:HB3	2:F:63:PHE:CE1	2.55	0.41
1:C:67:ALA:HB2	1:C:105:TYR:CE1	2.54	0.41
1:C:165:LYS:HG2	1:C:246:GLU:HG3	2.01	0.41
1:E:238:LYS:HG3	1:E:239:PRO:HD2	2.02	0.41
1:C:176:LEU:HD13	1:C:257:ALA:HB1	2.01	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:183:HIS:HB2	1:E:252:ILE:HD11	2.01	0.41
1:A:266:THR:HG22	2:B:65:ALA:HB1	2.01	0.41
1:A:26:VAL:HG21	1:A:317:ALA:HB2	2.03	0.41
1:A:295:HIS:HD2	1:A:306:PRO:HG2	1.86	0.41
2:D:120:ASP:HA	2:D:123:ARG:HB3	2.03	0.41
1:E:317:ALA:O	2:F:107:THR:HG21	2.21	0.41
1:A:140:PRO:HA	1:A:145:SER:HA	2.02	0.41
1:A:150:ASN:HA	1:A:256:TYR:CD1	2.56	0.41
2:F:167:ARG:HA	2:F:170:ARG:NH1	2.36	0.41
1:E:303:GLY:HA2	2:F:64:GLU:OE2	2.19	0.41
1:A:298:HIS:CG	1:A:299:PRO:HD2	2.56	0.41
1:C:86:TYR:HA	1:C:113:SER:O	2.20	0.41
1:C:141:TYR:HB2	1:C:146:SER:HB3	2.02	0.41
1:C:154:LEU:HD12	1:C:251:PHE:CD2	2.56	0.41
1:E:108:LEU:HB2	1:E:234:TRP:CE2	2.56	0.41
1:E:283:THR:HB	1:E:286:GLY:O	2.20	0.41
1:E:26:VAL:HG13	1:E:36:VAL:HG21	2.02	0.41
1:C:170:ASN:ND2	1:C:239:PRO:HA	2.36	0.41
1:E:174:GLU:N	1:E:174:GLU:OE2	2.53	0.41
1:E:105:TYR:CE2	1:E:109:LYS:HD2	2.55	0.41
1:E:17:TYR:HB2	1:E:320:LEU:HD22	2.03	0.40
1:E:266:THR:HA	2:F:69:GLU:OE2	2.21	0.40
1:A:106:GLU:N	1:A:106:GLU:OE1	2.51	0.40
1:A:201:TYR:CD2	1:A:248:ASN:HB2	2.56	0.40
1:C:51:LEU:HD13	1:C:272:LEU:HB2	2.02	0.40
1:A:127:TRP:HA	1:A:166:ARG:HH12	1.86	0.40
2:B:68:ARG:HA	2:B:68:ARG:HD3	1.96	0.40
1:A:60:ILE:HD12	1:A:274:TYR:HB2	2.02	0.40
1:C:317:ALA:O	2:D:107:THR:HG21	2.22	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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

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



Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	306/334 (92%)	288 (94%)	17 (6%)	1 (0%)	41	71
1	C	304/334 (91%)	290 (95%)	14 (5%)	0	100	100
1	E	300/334 (90%)	288 (96%)	10 (3%)	2 (1%)	22	54
2	B	143/177 (81%)	137 (96%)	6 (4%)	0	100	100
2	D	132/177 (75%)	127 (96%)	5 (4%)	0	100	100
2	F	145/177 (82%)	135 (93%)	10 (7%)	0	100	100
All	All	1330/1533 (87%)	1265 (95%)	62 (5%)	3 (0%)	47	77

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	249	GLY
1	A	249	GLY
1	E	143	GLY

### 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	280/299 (94%)	278 (99%)	2 (1%)	84	90
1	C	277/299 (93%)	273 (99%)	4 (1%)	67	82
1	E	275/299 (92%)	274 (100%)	1 (0%)	91	95
2	B	131/151 (87%)	129 (98%)	2 (2%)	65	81
2	D	128/151 (85%)	127 (99%)	1 (1%)	81	89
2	F	139/151 (92%)	136 (98%)	3 (2%)	52	74
All	All	1230/1350 (91%)	1217 (99%)	13 (1%)	73	85

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	173	GLN
1	A	251	PHE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	28	ASN
2	B	126	LEU
1	C	102	PHE
1	C	127	TRP
1	C	135	VAL
1	C	136	SER
2	D	135	ASN
1	E	102	PHE
2	F	61	THR
2	F	63	PHE
2	F	148	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	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	1.06	2 (14%)	17,19,21	1.47	3 (17%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	NAG	G	2	3	14,14,15	0.22	0	17,19,21	0.52	0
4	NAG	H	1	1,4	14,14,15	0.89	1 (7%)	17,19,21	1.41	3 (17%)
4	NAG	H	2	4	14,14,15	0.21	0	17,19,21	0.53	0
4	BMA	H	3	4	11,11,12	1.46	2 (18%)	15,15,17	0.88	1 (6%)
5	GAL	I	1	5	12,12,12	0.64	0	17,17,17	2.81	7 (41%)
5	NGC	I	2	5	18,21,22	1.92	3 (16%)	22,29,32	2.35	7 (31%)
3	NAG	J	1	1,3	14,14,15	0.89	1 (7%)	17,19,21	1.42	3 (17%)
3	NAG	J	2	3	14,14,15	0.21	0	17,19,21	0.52	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
4	NAG	H	1	1,4	-	4/6/23/26	0/1/1/1
4	NAG	H	2	4	-	2/6/23/26	0/1/1/1
4	BMA	H	3	4	-	0/2/19/22	0/1/1/1
5	GAL	I	1	5	-	0/2/22/22	0/1/1/1
5	NGC	I	2	5	-	1/15/36/40	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 (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	I	2	NGC	C10-N5	5.24	1.45	1.34
5	I	2	NGC	C3-C2	-4.23	1.45	1.52
4	H	3	BMA	O5-C1	3.03	1.48	1.43
4	H	3	BMA	O5-C5	2.62	1.48	1.43
3	J	1	NAG	O5-C1	-2.46	1.39	1.43
3	G	1	NAG	O5-C1	-2.46	1.39	1.43
4	H	1	NAG	O5-C1	-2.45	1.39	1.43
5	I	2	NGC	O10-C10	-2.32	1.18	1.23
3	G	1	NAG	O7-C7	-2.11	1.18	1.23

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	I	1	GAL	O1-C1-O5	8.18	134.93	110.38
5	I	2	NGC	C4-C5-N5	-7.18	96.16	110.38
5	I	1	GAL	O4-C4-C3	-5.17	98.39	110.35
5	I	2	NGC	C4-C3-C2	3.58	116.22	109.81
5	I	2	NGC	C3-C2-C1	-3.48	104.32	111.93
5	I	2	NGC	O9-C9-C8	-3.47	103.51	111.07
3	G	1	NAG	C1-C2-N2	-3.37	104.72	110.49
3	J	1	NAG	C2-N2-C7	-3.37	118.11	122.90
4	H	1	NAG	C2-N2-C7	-3.35	118.13	122.90
3	G	1	NAG	C6-C5-C4	-3.02	105.93	113.00
5	I	1	GAL	O4-C4-C5	-2.77	102.42	109.30
5	I	2	NGC	C11-C10-N5	2.76	121.17	116.37
3	J	1	NAG	C6-C5-C4	-2.69	106.71	113.00
4	H	1	NAG	C6-C5-C4	-2.66	106.78	113.00
5	I	1	GAL	C6-C5-C4	2.52	118.90	113.00
3	G	1	NAG	O3-C3-C2	2.49	114.62	109.47
5	I	1	GAL	O2-C2-C3	-2.48	104.61	110.35
3	J	1	NAG	C1-C2-N2	-2.43	106.33	110.49
4	H	1	NAG	C1-C2-N2	-2.41	106.37	110.49
5	I	1	GAL	O6-C6-C5	-2.40	103.06	111.29
5	I	2	NGC	C6-O6-C2	2.34	116.34	111.34
4	H	3	BMA	C1-O5-C5	2.16	115.11	112.19
5	I	1	GAL	O2-C2-C1	-2.13	104.22	109.16
5	I	2	NGC	O10-C10-N5	-2.11	119.39	122.95

There are no chirality outliers.

All (15) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	I	2	NGC	O10-C10-C11-O11
4	H	2	NAG	O5-C5-C6-O6
3	G	2	NAG	O5-C5-C6-O6
3	J	2	NAG	O5-C5-C6-O6
3	G	2	NAG	C4-C5-C6-O6
4	H	2	NAG	C4-C5-C6-O6
3	J	2	NAG	C4-C5-C6-O6
3	J	1	NAG	C8-C7-N2-C2
4	H	1	NAG	C8-C7-N2-C2
4	H	1	NAG	O7-C7-N2-C2
3	J	1	NAG	O7-C7-N2-C2
4	H	1	NAG	C4-C5-C6-O6
3	J	1	NAG	C4-C5-C6-O6
4	H	1	NAG	O5-C5-C6-O6

*Continued on next page...*

*Continued from previous page...*

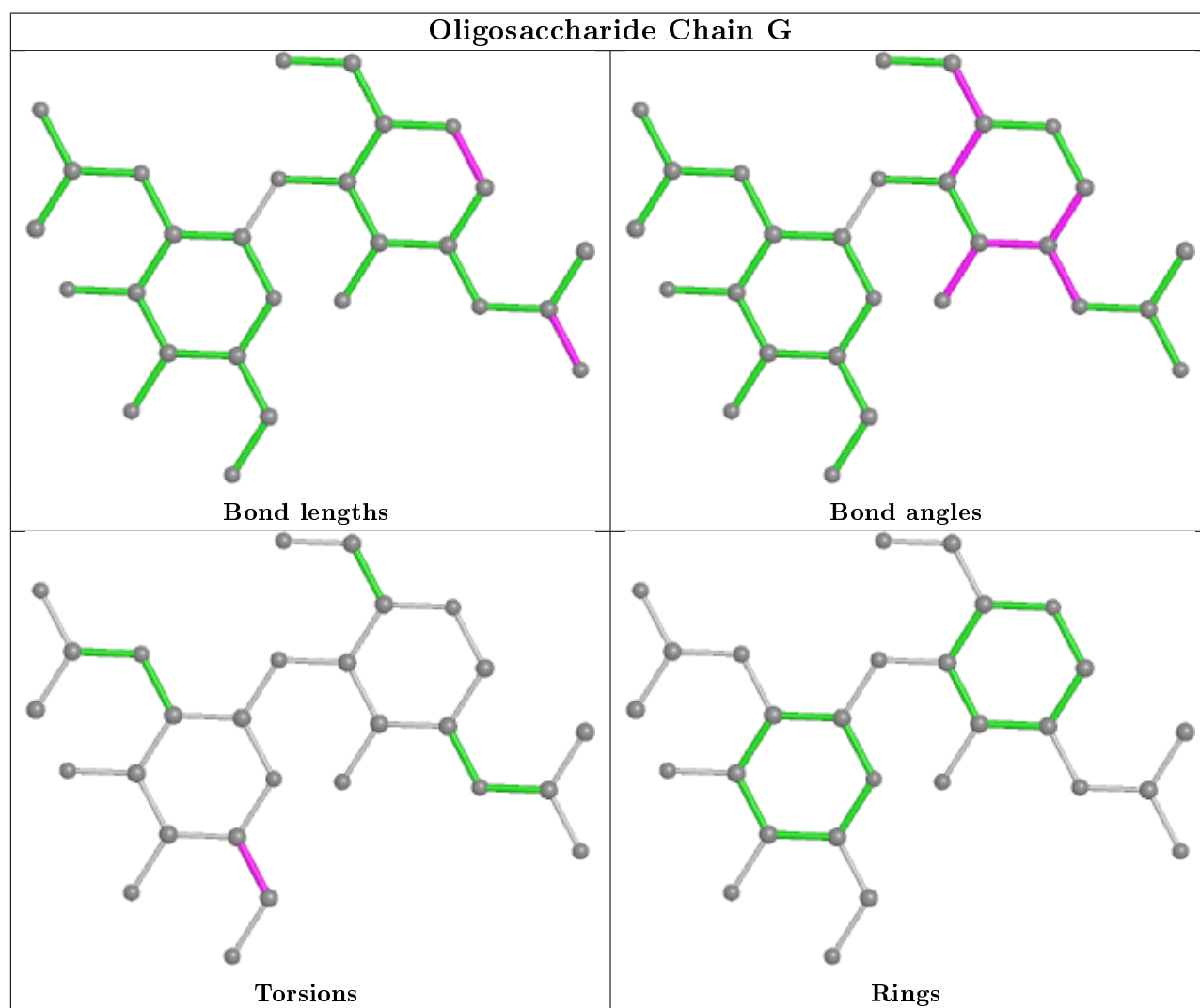
Mol	Chain	Res	Type	Atoms
3	J	1	NAG	O5-C5-C6-O6

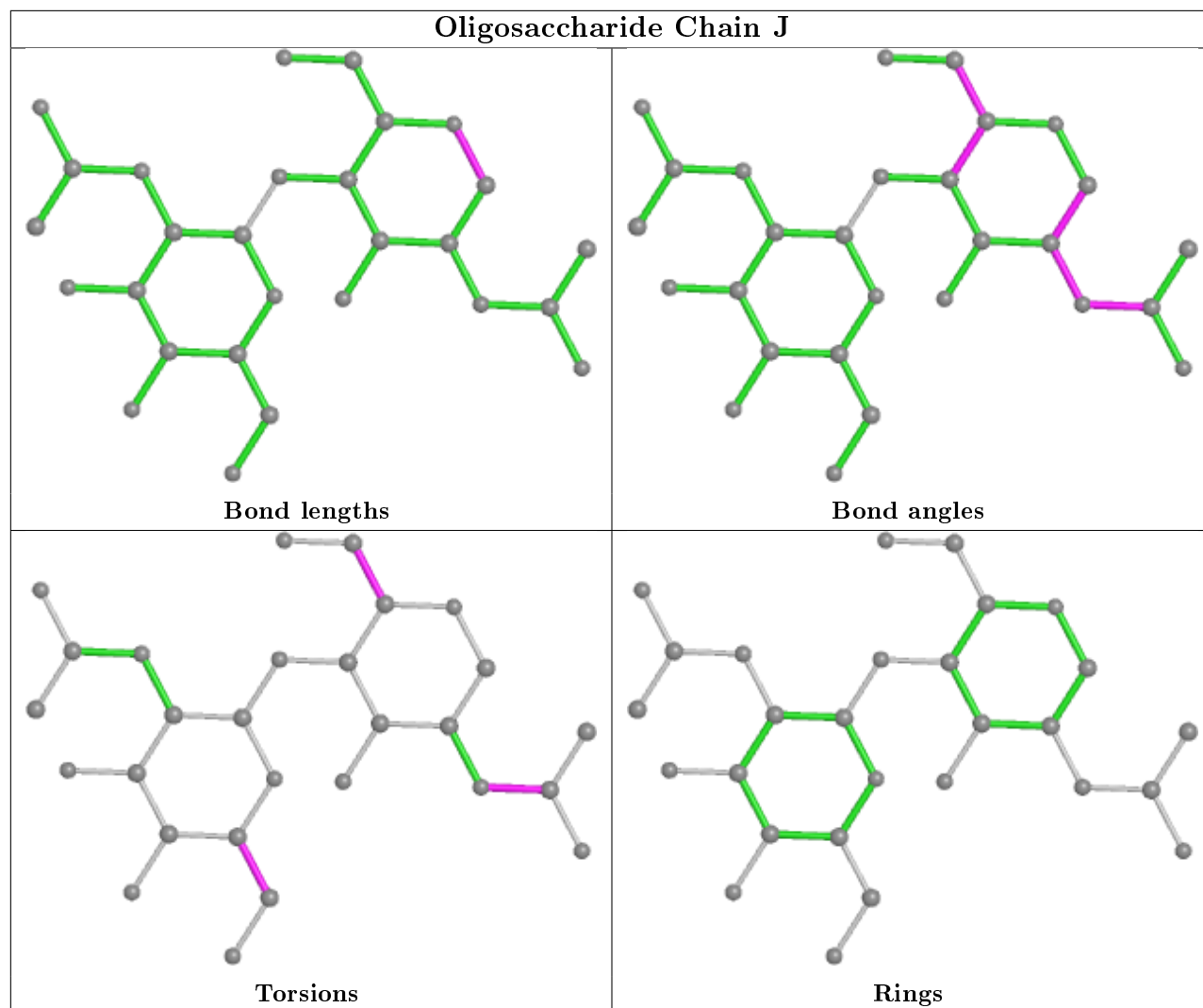
There are no ring outliers.

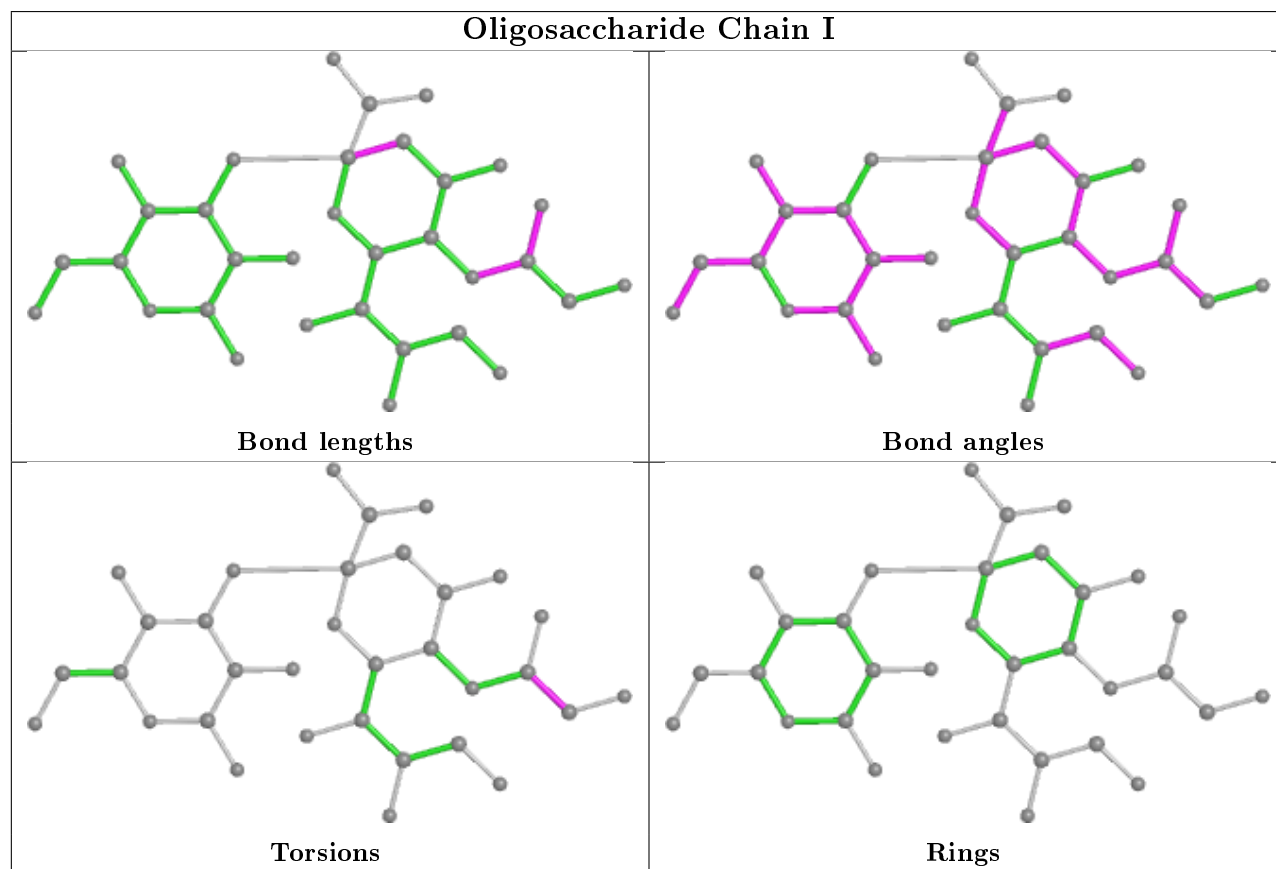
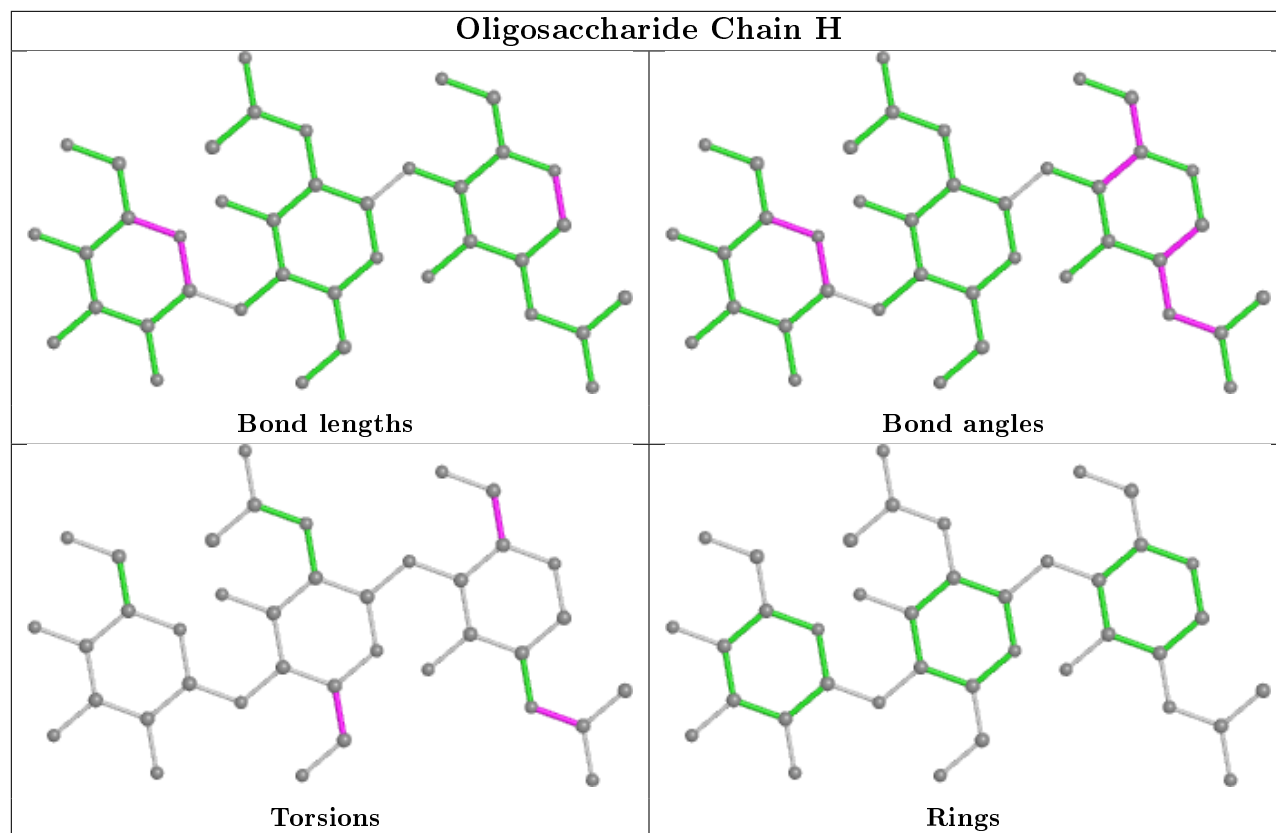
1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	I	2	NGC	2	0

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







## 5.6 Ligand geometry

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$
6	NAG	A	401	1	14,14,15	1.16	2 (14%)	17,19,21	1.21	2 (11%)
6	NAG	E	601	1	14,14,15	1.17	2 (14%)	17,19,21	1.21	2 (11%)

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
6	NAG	A	401	1	-	2/6/23/26	0/1/1/1
6	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)
6	E	601	NAG	O5-C1	-2.45	1.39	1.43
6	A	401	NAG	O5-C1	-2.41	1.39	1.43
6	A	401	NAG	C2-N2	-2.11	1.42	1.46
6	E	601	NAG	C2-N2	-2.10	1.42	1.46

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	401	NAG	O5-C5-C6	-2.20	103.75	107.20
6	E	601	NAG	O5-C5-C6	-2.17	103.81	107.20
6	A	401	NAG	O3-C3-C2	2.01	113.63	109.47
6	E	601	NAG	O3-C3-C2	2.00	113.61	109.47

There are no chirality outliers.

All (4) torsion outliers are listed below:



Mol	Chain	Res	Type	Atoms
6	A	401	NAG	O5-C5-C6-O6
6	E	601	NAG	O5-C5-C6-O6
6	A	401	NAG	C4-C5-C6-O6
6	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	312/334 (93%)	-0.02	3 (0%) 82 82	45, 87, 132, 161	0
1	C	310/334 (92%)	0.09	3 (0%) 82 82	55, 93, 135, 156	0
1	E	308/334 (92%)	-0.04	2 (0%) 89 90	57, 93, 126, 146	0
2	B	149/177 (84%)	0.29	6 (4%) 38 36	55, 124, 149, 165	0
2	D	142/177 (80%)	0.42	9 (6%) 20 20	57, 127, 155, 183	0
2	F	153/177 (86%)	0.42	10 (6%) 18 18	56, 129, 150, 166	0
All	All	1374/1533 (89%)	0.13	33 (2%) 59 56	45, 96, 144, 183	0

All (33) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	21	TRP	4.5
2	F	19	ASP	4.0
1	A	13	ILE	3.8
2	F	138	PHE	3.6
2	B	136	GLY	3.5
2	F	127	ARG	3.5
1	C	51	LEU	3.5
2	D	129	ASN	3.4
2	D	139	GLU	3.1
2	B	138	PHE	3.0
2	D	140	PHE	3.0
2	F	24	TYR	3.0
2	B	132	GLU	3.0
2	B	131	LYS	2.9
2	F	168	LEU	2.9
1	C	18	HIS	2.7
2	D	14	TRP	2.6
1	A	318	THR	2.6
2	D	143	LYS	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	D	19	ASP	2.4
1	A	158	ASN	2.3
2	F	39	GLU	2.3
1	C	132	ALA	2.3
2	D	36	ALA	2.3
2	D	136	GLY	2.2
2	F	38	LYS	2.2
2	B	140	PHE	2.1
2	B	141	TYR	2.1
2	F	122	VAL	2.1
2	F	40	SER	2.1
1	E	276	ASN	2.0
2	F	167	ARG	2.0
1	E	161	ALA	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](#)

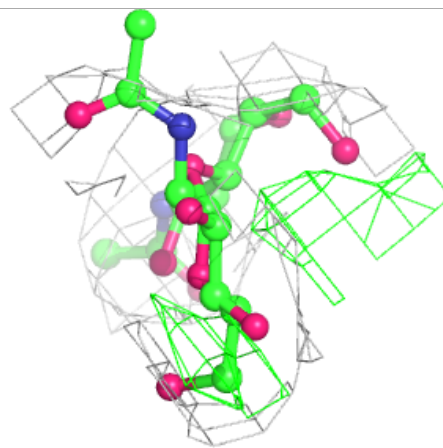
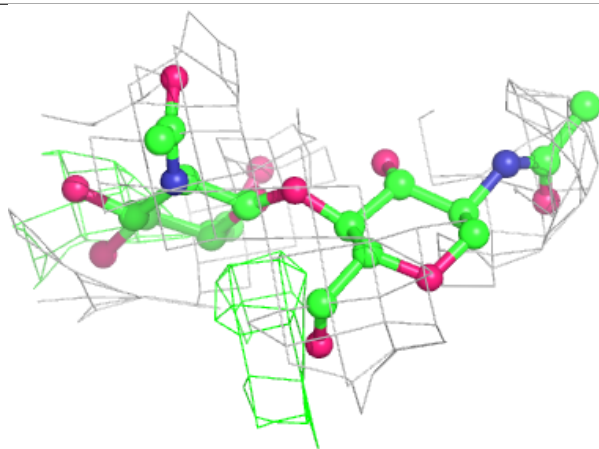
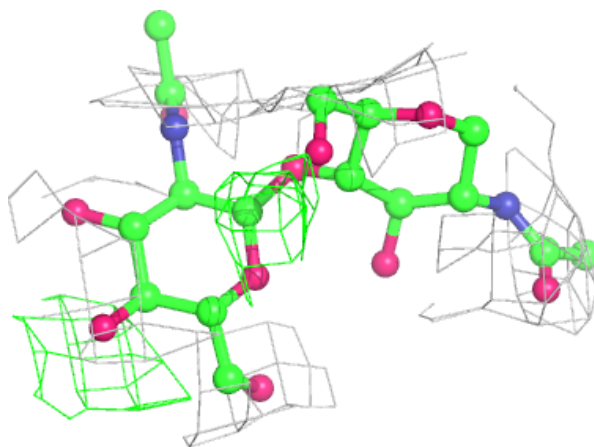
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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	H	3	11/12	0.67	0.34	156,162,173,176	0
3	NAG	J	2	14/15	0.74	0.33	128,145,151,152	0
3	NAG	G	2	14/15	0.77	0.24	106,121,136,139	0
3	NAG	J	1	14/15	0.80	0.23	88,119,133,143	0
4	NAG	H	2	14/15	0.81	0.17	119,144,156,161	0
5	GAL	I	1	12/12	0.82	0.33	130,137,151,158	0
3	NAG	G	1	14/15	0.82	0.19	92,110,118,119	0
5	NGC	I	2	21/22	0.88	0.37	110,121,131,140	1
4	NAG	H	1	14/15	0.89	0.15	100,114,126,134	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.

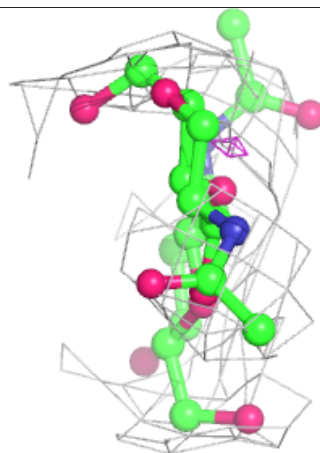
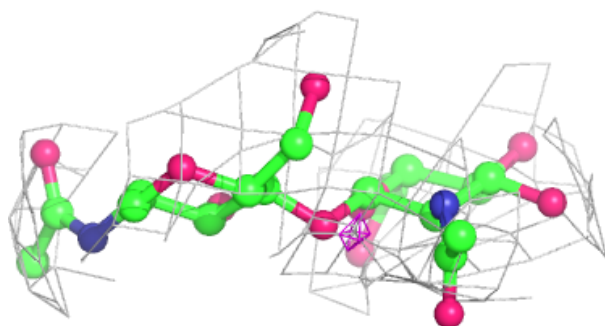
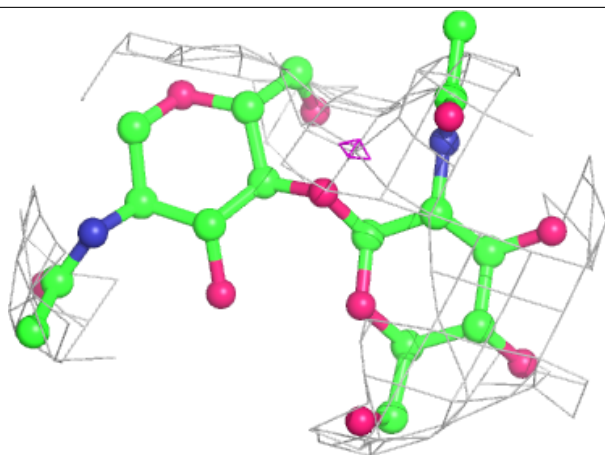
**Electron density around Chain G:**

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

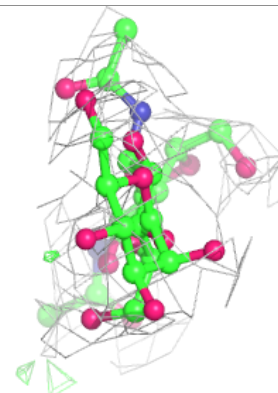
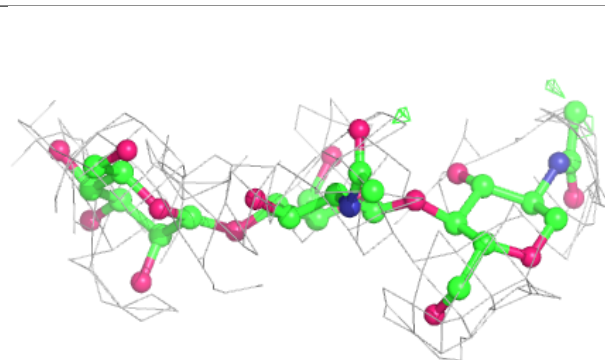
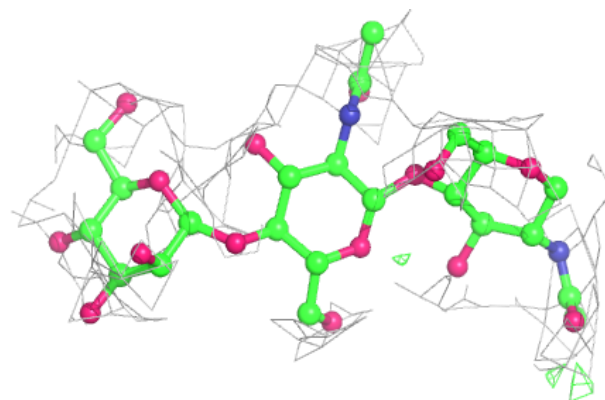


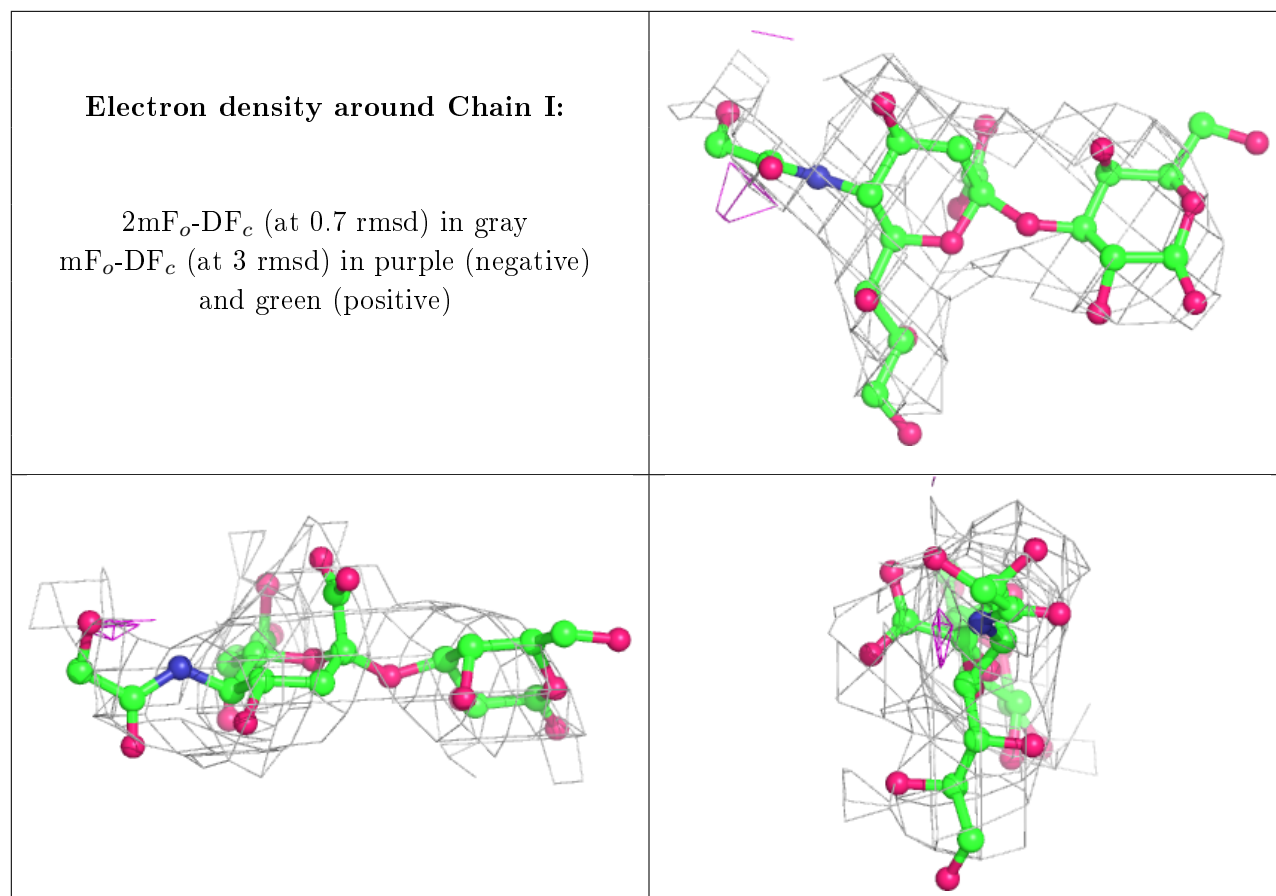
**Electron density around Chain J:**

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

**Electron density around Chain H:**

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





## 6.4 Ligands [i](#)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
6	NAG	A	401	14/15	0.72	0.21	140,157,163,167	0
6	NAG	E	601	14/15	0.75	0.34	148,163,173,175	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.