W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Aug 29, 2023 – 12:00 AM EDT

PDB ID	:	3M6S
Title	:	Crystal structure of H1N1pdm Hemagglutinin
Authors	:	Yang, H.; Carney, P.J.; Stevens, J.
Deposited on	:	2010-03-16
Resolution	:	2.80 Å(reported)

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

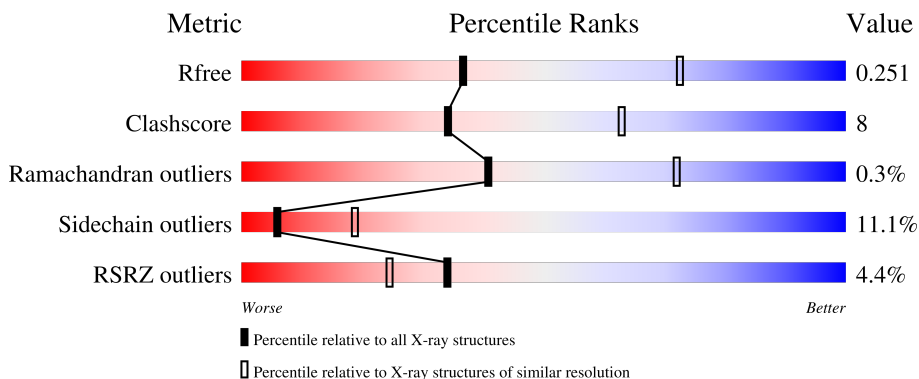|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	Percentile Ranks	Value
Rfree		0.251
Clashscore		8
Ramachandran outliers		0.3%
Sidechain outliers		11.1%
RSRZ outliers		4.4%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 >=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 <=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	331	3% / 79% 16% · ·
1	C	331	% / 76% 18% · ·
1	E	331	4% / 77% 17% · ·
1	G	331	3% / 78% 15% · · ·
1	I	331	5% / 79% 15% · ·

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	331	3% ... 74% ... 19% ...
2	B	181	2% ... 67% ... 22% ... 7%
2	D	181	70% ... 20% ... 6%
2	F	181	14% ... 70% ... 22% ... 6%
2	H	181	4% ... 73% ... 18% ... 6%
2	J	181	4% ... 66% ... 25% ... 6%
2	L	181	14% ... 67% ... 24% ... 6%
3	M	2	50% ... 50%
3	O	2	100%
3	P	2	50% ... 50%
3	Q	2	100%
4	N	3	67% ... 33%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	MAN	N	3	X	-	-	-
5	NAG	G	331	X	-	-	-

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 23587 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	325	Total	C	N	O	S	0	0	0
			2533	1602	435	485	11			
1	C	323	Total	C	N	O	S	0	0	0
			2518	1593	433	481	11			
1	E	323	Total	C	N	O	S	0	0	0
			2518	1593	433	481	11			
1	G	322	Total	C	N	O	S	0	0	0
			2514	1591	432	480	11			
1	I	322	Total	C	N	O	S	0	0	0
			2511	1588	432	480	11			
1	K	321	Total	C	N	O	S	0	0	0
			2506	1587	431	477	11			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	ALA	-	expression tag	UNP C5MV42
A	-2	ASP	-	expression tag	UNP C5MV42
A	-1	PRO	-	expression tag	UNP C5MV42
A	0	GLY	-	expression tag	UNP C5MV42
C	-3	ALA	-	expression tag	UNP C5MV42
C	-2	ASP	-	expression tag	UNP C5MV42
C	-1	PRO	-	expression tag	UNP C5MV42
C	0	GLY	-	expression tag	UNP C5MV42
E	-3	ALA	-	expression tag	UNP C5MV42
E	-2	ASP	-	expression tag	UNP C5MV42
E	-1	PRO	-	expression tag	UNP C5MV42
E	0	GLY	-	expression tag	UNP C5MV42
G	-3	ALA	-	expression tag	UNP C5MV42
G	-2	ASP	-	expression tag	UNP C5MV42
G	-1	PRO	-	expression tag	UNP C5MV42
G	0	GLY	-	expression tag	UNP C5MV42
I	-3	ALA	-	expression tag	UNP C5MV42

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
I	-2	ASP	-	expression tag	UNP C5MV42
I	-1	PRO	-	expression tag	UNP C5MV42
I	0	GLY	-	expression tag	UNP C5MV42
K	-3	ALA	-	expression tag	UNP C5MV42
K	-2	ASP	-	expression tag	UNP C5MV42
K	-1	PRO	-	expression tag	UNP C5MV42
K	0	GLY	-	expression tag	UNP C5MV42

- Molecule 2 is a protein called Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	169	Total 1368	C 860	N 232	O 270	S 6	5	0	0
2	D	171	Total 1386	C 870	N 234	O 276	S 6	0	0	0
2	F	171	Total 1386	C 870	N 234	O 276	S 6	0	0	0
2	H	170	Total 1377	C 865	N 233	O 273	S 6	0	0	0
2	J	170	Total 1377	C 865	N 233	O 273	S 6	0	0	0
2	L	171	Total 1386	C 870	N 234	O 276	S 6	0	0	0

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	175	SER	GLY	engineered mutation	UNP C5MV42
B	176	GLY	VAL	engineered mutation	UNP C5MV42
B	177	ARG	LYS	engineered mutation	UNP C5MV42
B	179	VAL	-	expression tag	UNP C5MV42
B	180	PRO	-	expression tag	UNP C5MV42
B	181	ARG	-	expression tag	UNP C5MV42
D	175	SER	GLY	engineered mutation	UNP C5MV42
D	176	GLY	VAL	engineered mutation	UNP C5MV42
D	177	ARG	LYS	engineered mutation	UNP C5MV42
D	179	VAL	-	expression tag	UNP C5MV42
D	180	PRO	-	expression tag	UNP C5MV42
D	181	ARG	-	expression tag	UNP C5MV42
F	175	SER	GLY	engineered mutation	UNP C5MV42
F	176	GLY	VAL	engineered mutation	UNP C5MV42
F	177	ARG	LYS	engineered mutation	UNP C5MV42

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	179	VAL	-	expression tag	UNP C5MV42
F	180	PRO	-	expression tag	UNP C5MV42
F	181	ARG	-	expression tag	UNP C5MV42
H	175	SER	GLY	engineered mutation	UNP C5MV42
H	176	GLY	VAL	engineered mutation	UNP C5MV42
H	177	ARG	LYS	engineered mutation	UNP C5MV42
H	179	VAL	-	expression tag	UNP C5MV42
H	180	PRO	-	expression tag	UNP C5MV42
H	181	ARG	-	expression tag	UNP C5MV42
J	175	SER	GLY	engineered mutation	UNP C5MV42
J	176	GLY	VAL	engineered mutation	UNP C5MV42
J	177	ARG	LYS	engineered mutation	UNP C5MV42
J	179	VAL	-	expression tag	UNP C5MV42
J	180	PRO	-	expression tag	UNP C5MV42
J	181	ARG	-	expression tag	UNP C5MV42
L	175	SER	GLY	engineered mutation	UNP C5MV42
L	176	GLY	VAL	engineered mutation	UNP C5MV42
L	177	ARG	LYS	engineered mutation	UNP C5MV42
L	179	VAL	-	expression tag	UNP C5MV42
L	180	PRO	-	expression tag	UNP C5MV42
L	181	ARG	-	expression tag	UNP C5MV42

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

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

- Molecule 4 is an oligosaccharide called alpha-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.
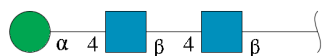
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	N	3	Total 39	C 22	N 2	O 15	0	0	0

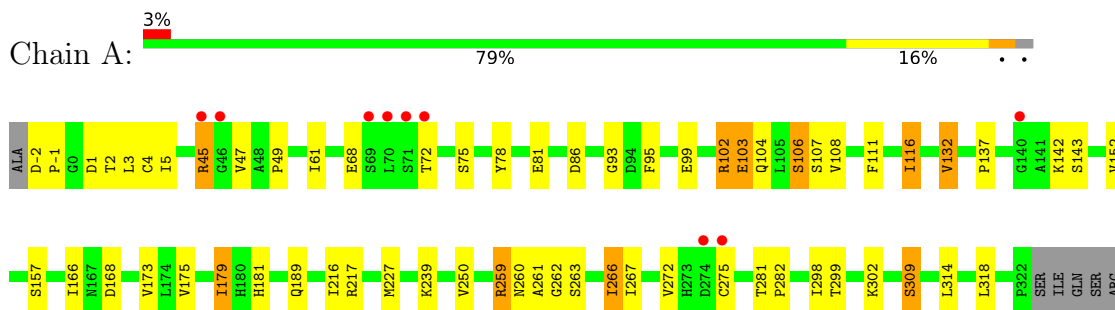- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

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

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.
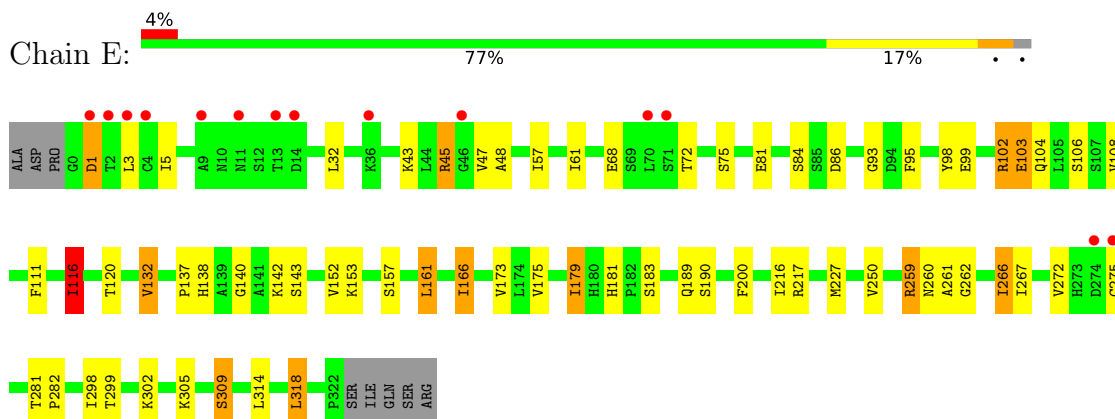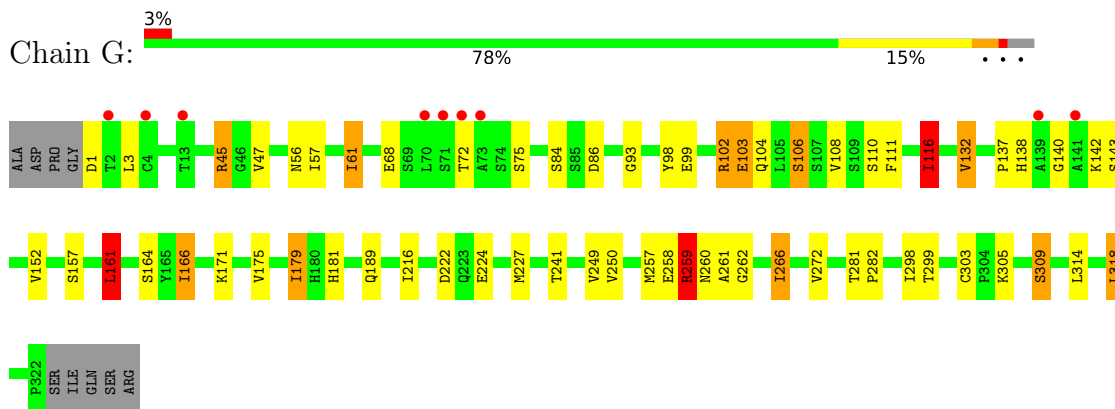
• Molecule 1: Hemagglutinin
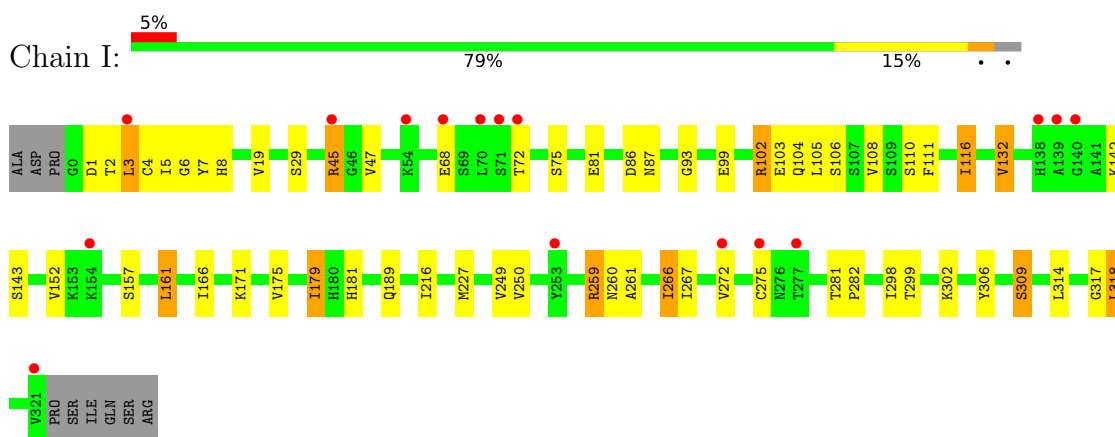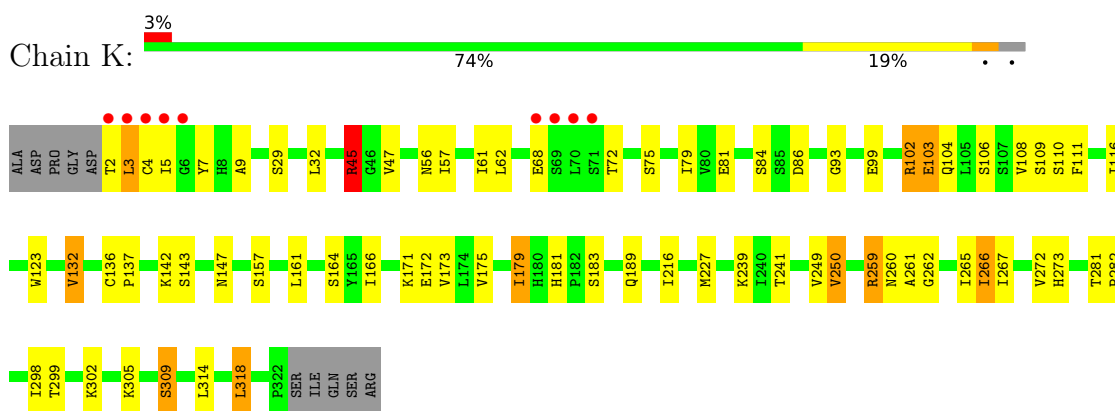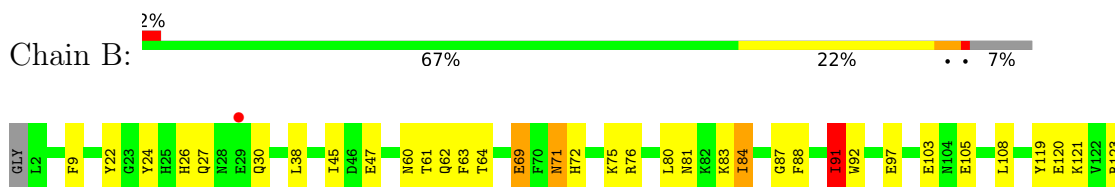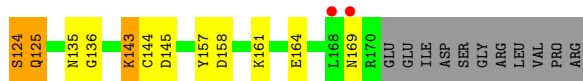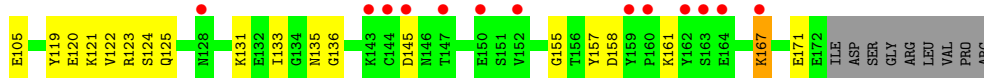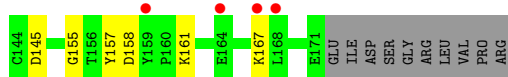

• Molecule 1: Hemagglutinin

• Molecule 1: Hemagglutinin

• Molecule 1: Hemagglutinin

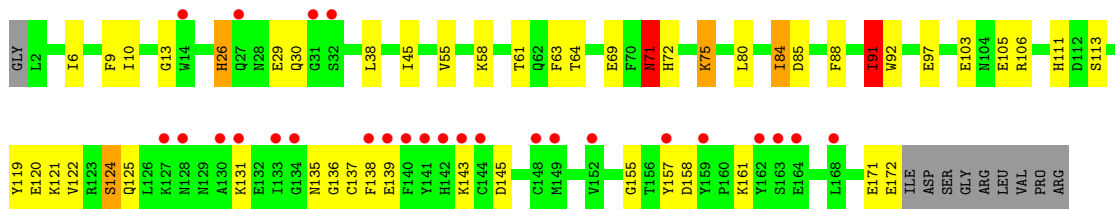Chain G:

• Molecule 1: Hemagglutinin

Chain I:

• Molecule 1: Hemagglutinin

Chain K:

• Molecule 2: Hemagglutinin

Chain B:

S124 Q125 N135 G136 K143 C144 D145 Y157 D158 K161 E164 L168 N169 R170 GLU ILE ASP SER GLY ARG LEU VAL PRO ARG

● Molecule 2: Hemagglutinin

Chain D:

70%   20%   6%

GLY L2 F9 Y22 H26 Q30 L38 I45 V55 T61 Q62 F63 T64 E69 F70 N71 H72 K75 R76 L80 N81 K82 K83 I84 G87 F88 L89 D90 I91 W92 E97 E103 N104 E105 Y119 E120 K121 S124 Q125 K131 E132 I133 G134

N135 G136 C137 F138 Y141 D145 G155 T156 Y157 D158 K161 K167 E171 E172 ILE ASP SER GLY ARG LEU VAL PRO ARG

● Molecule 2: Hemagglutinin

Chain F:

14%

70%   22%   6%

GLY L2 F9 G16 M17 V18 D19 G20 W21 Y22 G23 Y24 H25 H26 N28 E29 Q30 G31 S32 G33 Y34 A35 L38 I45 V55 T61 Q62 F63 T64 E69 F70 N71 K75 R76 L80 K83 I84 D85 D86 G87 F88 I91 W92 E97 E103 N104

E105 Y119 E120 K121 V122 R123 S124 Q125 N128 K131 E132 I133 G134 N135 G136 K143 C144 D145 N146 T147 E150 S151 V152 G155 T156 Y157 D158 Y159 P160 K161 Y162 S163 E164 K167 E171 E172 ILE ASP SER GLY ARG LEU VAL PRO ARG

● Molecule 2: Hemagglutinin

Chain H:

4%

73%   18%   6%

GLY L2 F9 D19 H26 Q27 N28 E29 Q30 L38 I45 N60 F63 T64 E69 F70 N71 H72 L73 E74 K75 L80 I84 I91 W92 E97 E103 N104 E105 Y119 E120 K121 S124 Q125 N128 K131 N135 G136 Y141 H142 K143

C144 D145 G155 T156 Y157 D158 Y159 P160 K161 E164 K167 L168 E171 GLU ILE ASP SER GLY ARG LEU VAL PRO ARG

● Molecule 2: Hemagglutinin

Chain J:

4%

66%   25%   6%

GLY L2 F9 W14 V18 D19 G20 W21 Y22 G23 Y24 H25 H26 Q27 Q30 G31 S32 G33 L38 I45 K58 F63 T64 E69 F70 N71 H72 L73 E74 K75 R76 L80 K83 I84 G87 F88 L89 D90 I91 E97 L102 E103 N104 E105 L108

L118 Y119 E120 K121 V122 R123 S124 Q125 N135 G136 C137 K143 C144 D145 K153 N154 G155 T156 Y157 D158 K161 E164 K167 L168 E171 GLU ILE ASP SER GLY ARG LEU VAL PRO ARG

● Molecule 2: Hemagglutinin

Chain L:

14%

67%   24%   6%

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

Chain M:

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

Chain O:

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

Chain P:

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

Chain Q:

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

Chain N:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	73.98Å 109.71Å 129.90Å 86.25° 74.68° 75.10°	Depositor
Resolution (Å)	50.00 – 2.80 48.87 – 2.80	Depositor EDS
% Data completeness (in resolution range)	98.1 (50.00-2.80) 98.1 (48.87-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	2.32 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.230 , 0.256 0.227 , 0.251	Depositor DCC
R_{free} test set	4608 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	50.1	Xtriage
Anisotropy	0.013	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 43.3	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.016 for -h,-k,-h+l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	23587	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.83% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

wwPDB X-ray Structure Validation Summary Report

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$ >5	RMSZ	#$|Z|$ >5
1	A	0.75	0/2597	0.83	1/3530 (0.0%)
1	C	0.79	1/2581 (0.0%)	0.84	2/3507 (0.1%)
1	E	0.71	0/2581	0.80	2/3507 (0.1%)
1	G	0.86	1/2577 (0.0%)	0.88	4/3502 (0.1%)
1	I	0.68	0/2573	0.78	1/3495 (0.0%)
1	K	0.75	0/2569	0.82	2/3491 (0.1%)
2	B	0.71	0/1396	0.79	3/1881 (0.2%)
2	D	0.77	0/1414	0.78	2/1905 (0.1%)
2	F	0.68	0/1414	0.73	1/1905 (0.1%)
2	H	0.75	1/1405 (0.1%)	0.72	1/1893 (0.1%)
2	J	0.77	1/1405 (0.1%)	0.75	2/1893 (0.1%)
2	L	0.73	0/1414	0.74	3/1905 (0.2%)
All	All	0.75	4/23926 (0.0%)	0.80	24/32414 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	258	GLU	CG-CD	6.18	1.61	1.51
1	C	4	CYS	CB-SG	-5.94	1.72	1.81
2	J	137	CYS	CB-SG	5.78	1.92	1.82
2	H	74	GLU	CG-CD	5.42	1.60	1.51

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	143	LYS	CB-CA-C	9.40	129.19	110.40
2	L	91	ILE	CG1-CB-CG2	-8.03	93.73	111.40
2	B	91	ILE	CG1-CB-CG2	-7.20	95.55	111.40
2	D	91	ILE	CG1-CB-CG2	-6.26	97.62	111.40
2	H	91	ILE	CG1-CB-CG2	-6.16	97.85	111.40

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	2533	0	2483	40	1
1	C	2518	0	2472	52	0
1	E	2518	0	2472	43	1
1	G	2514	0	2466	36	2
1	I	2511	0	2464	62	0
1	K	2506	0	2460	59	1
2	B	1368	0	1293	36	3
2	D	1386	0	1305	41	0
2	F	1386	0	1305	38	0
2	H	1377	0	1299	18	0
2	J	1377	0	1299	56	0
2	L	1386	0	1305	45	0
3	M	28	0	25	2	0
3	O	28	0	24	2	0
3	P	28	0	25	1	0
3	Q	28	0	25	0	0
4	N	39	0	34	1	0
5	A	14	0	13	1	0
5	E	14	0	13	1	0
5	G	28	0	26	0	0
All	All	23587	0	22808	390	4

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.

The worst 5 of 390 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:I:45:ARG:HG3	1:I:45:ARG:HH11	1.07	1.10
1:E:45:ARG:HG3	1:E:45:ARG:HH11	1.09	1.10
1:I:5:ILE:HD13	2:J:24:TYR:CD1	1.87	1.09

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:I:4:CYS:O	2:J:24:TYR:HA	1.52	1.09
1:A:45:ARG:HG3	1:A:45:ARG:HH11	1.14	1.08

All (4) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:143:LYS:CE	1:G:224:GLU:OE1[1_466]	1.91	0.29
1:A:168:ASP:O	1:K:273:HIS:CE1[1_655]	1.98	0.22
2:B:169:ASN:OD1	1:G:222:ASP:OD2[1_466]	2.13	0.07
2:B:164:GLU:OE1	1:E:120:THR:O[1_556]	2.19	0.01

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	323/331 (98%)	305 (94%)	16 (5%)	2 (1%)	25	56
1	C	321/331 (97%)	303 (94%)	16 (5%)	2 (1%)	25	56
1	E	321/331 (97%)	303 (94%)	16 (5%)	2 (1%)	25	56
1	G	320/331 (97%)	303 (95%)	16 (5%)	1 (0%)	41	72
1	I	320/331 (97%)	299 (93%)	19 (6%)	2 (1%)	25	56
1	K	319/331 (96%)	300 (94%)	18 (6%)	1 (0%)	41	72
2	B	167/181 (92%)	163 (98%)	4 (2%)	0	100	100
2	D	169/181 (93%)	163 (96%)	6 (4%)	0	100	100
2	F	169/181 (93%)	164 (97%)	5 (3%)	0	100	100
2	H	168/181 (93%)	164 (98%)	4 (2%)	0	100	100
2	J	168/181 (93%)	162 (96%)	6 (4%)	0	100	100
2	L	169/181 (93%)	162 (96%)	7 (4%)	0	100	100

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	2934/3072 (96%)	2791 (95%)	133 (4%)	10 (0%)	41　72

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	261	ALA
1	G	261	ALA
1	I	261	ALA
1	K	261	ALA
1	C	261	ALA

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	284/289 (98%)	254 (89%)	30 (11%)	6	20
1	C	282/289 (98%)	254 (90%)	28 (10%)	8	23
1	E	282/289 (98%)	249 (88%)	33 (12%)	5	16
1	G	282/289 (98%)	251 (89%)	31 (11%)	6	19
1	I	281/289 (97%)	250 (89%)	31 (11%)	6	19
1	K	281/289 (97%)	249 (89%)	32 (11%)	5	18
2	B	146/156 (94%)	128 (88%)	18 (12%)	4	15
2	D	148/156 (95%)	131 (88%)	17 (12%)	5	17
2	F	148/156 (95%)	133 (90%)	15 (10%)	7	22
2	H	147/156 (94%)	129 (88%)	18 (12%)	5	15
2	J	147/156 (94%)	131 (89%)	16 (11%)	6	19
2	L	148/156 (95%)	132 (89%)	16 (11%)	6	19
All	All	2576/2670 (96%)	2291 (89%)	285 (11%)	6	19

5 of 285 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	J	105	GLU
1	K	29	SER
1	K	239	LYS
1	E	68	GLU
1	E	3	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3   RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5   Carbohydrates ⓘ

11 monosaccharides are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
3	NAG	M	1	3,1	14,14,15	0.77	0	17,19,21	2.00	3 (17%)
3	NAG	M	2	3	14,14,15	0.91	1 (7%)	17,19,21	2.05	7 (41%)
4	NAG	N	1	4,1	14,14,15	0.81	0	17,19,21	1.68	4 (23%)
4	NAG	N	2	4	14,14,15	0.77	0	17,19,21	2.07	5 (29%)
4	MAN	N	3	4	11,11,12	0.89	0	15,15,17	2.85	7 (46%)
3	NAG	O	1	3,1	14,14,15	0.55	0	17,19,21	2.21	5 (29%)
3	NAG	O	2	3	14,14,15	1.02	1 (7%)	17,19,21	1.51	4 (23%)
3	NAG	P	1	3,1	14,14,15	1.10	1 (7%)	17,19,21	2.36	5 (29%)
3	NAG	P	2	3	14,14,15	0.63	0	17,19,21	1.30	4 (23%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
3	NAG	Q	1	3,1	14,14,15	1.74	2 (14%)	17,19,21	1.82	4 (23%)
3	NAG	Q	2	3	14,14,15	0.99	1 (7%)	17,19,21	2.05	4 (23%)

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	M	1	3,1	-	2/6/23/26	0/1/1/1
3	NAG	M	2	3	-	0/6/23/26	0/1/1/1
4	NAG	N	1	4,1	-	1/6/23/26	0/1/1/1
4	NAG	N	2	4	-	2/6/23/26	0/1/1/1
4	MAN	N	3	4	1/1/4/5	2/2/19/22	0/1/1/1
3	NAG	O	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	O	2	3	-	0/6/23/26	0/1/1/1
3	NAG	P	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	P	2	3	-	0/6/23/26	0/1/1/1
3	NAG	Q	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	Q	2	3	-	1/6/23/26	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	Q	1	NAG	C1-C2	5.10	1.60	1.52
3	P	1	NAG	C1-C2	3.80	1.58	1.52
3	O	2	NAG	C1-C2	2.56	1.56	1.52
3	Q	2	NAG	C1-C2	2.18	1.55	1.52
3	M	2	NAG	C1-C2	2.14	1.55	1.52

The worst 5 of 52 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	M	1	NAG	O5-C1-C2	-6.00	101.81	111.29
3	O	1	NAG	C1-O5-C5	5.92	120.22	112.19
4	N	3	MAN	C1-O5-C5	5.43	119.56	112.19
3	P	1	NAG	C1-O5-C5	5.05	119.03	112.19
4	N	2	NAG	C1-O5-C5	4.90	118.83	112.19

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	N	3	MAN	C1

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	N	3	MAN	C4-C5-C6-O6
3	M	1	NAG	O5-C5-C6-O6
4	N	3	MAN	O5-C5-C6-O6
3	M	1	NAG	C4-C5-C6-O6
4	N	2	NAG	C4-C5-C6-O6

There are no ring outliers.

5 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	M	1	NAG	2	0
3	O	1	NAG	2	0
3	P	1	NAG	1	0
4	N	1	NAG	1	0
3	O	2	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.

Oligosaccharide Chain M

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain O

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain P

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain Q

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain N

Bond lengths

Bond angles

Torsions

Rings
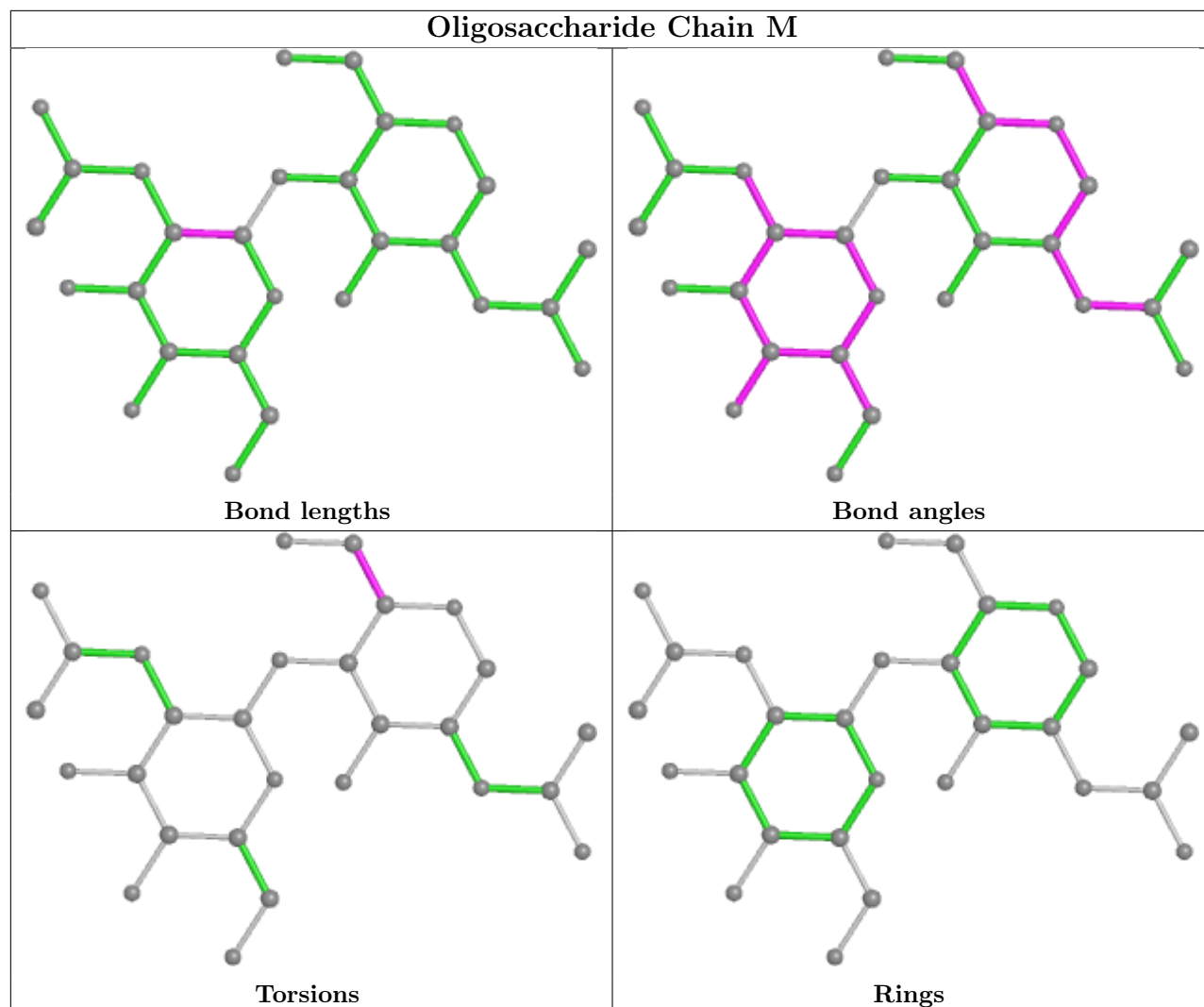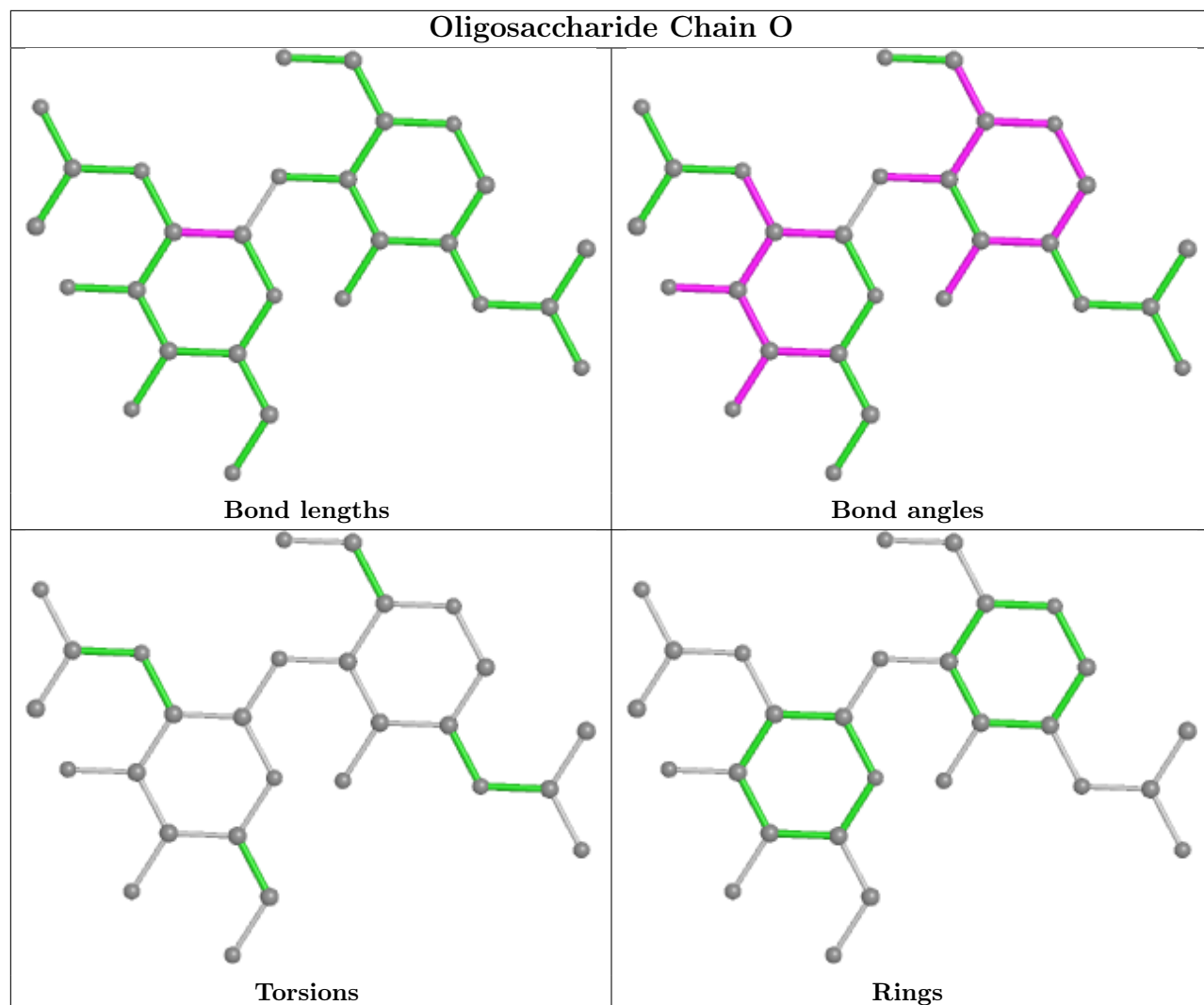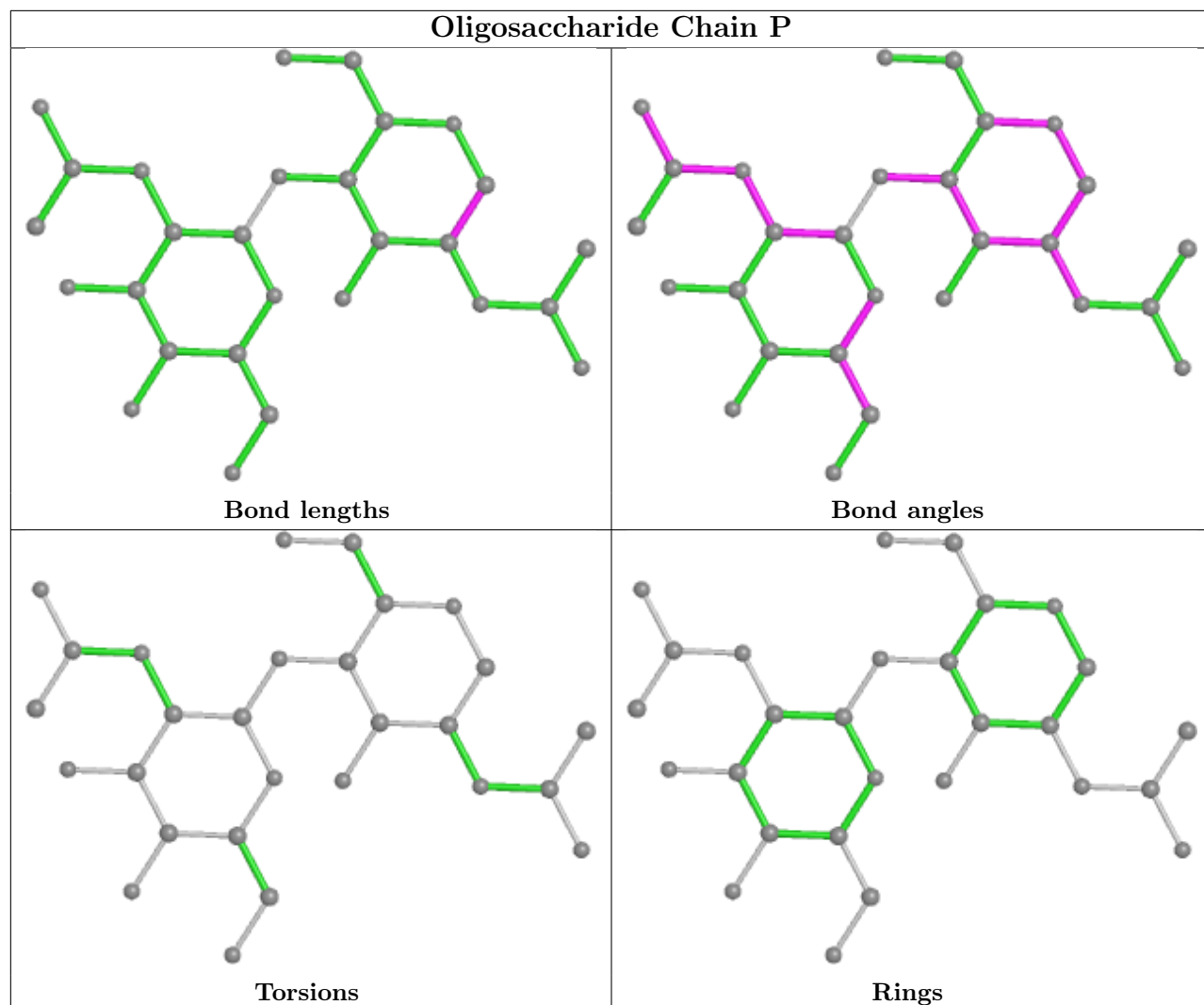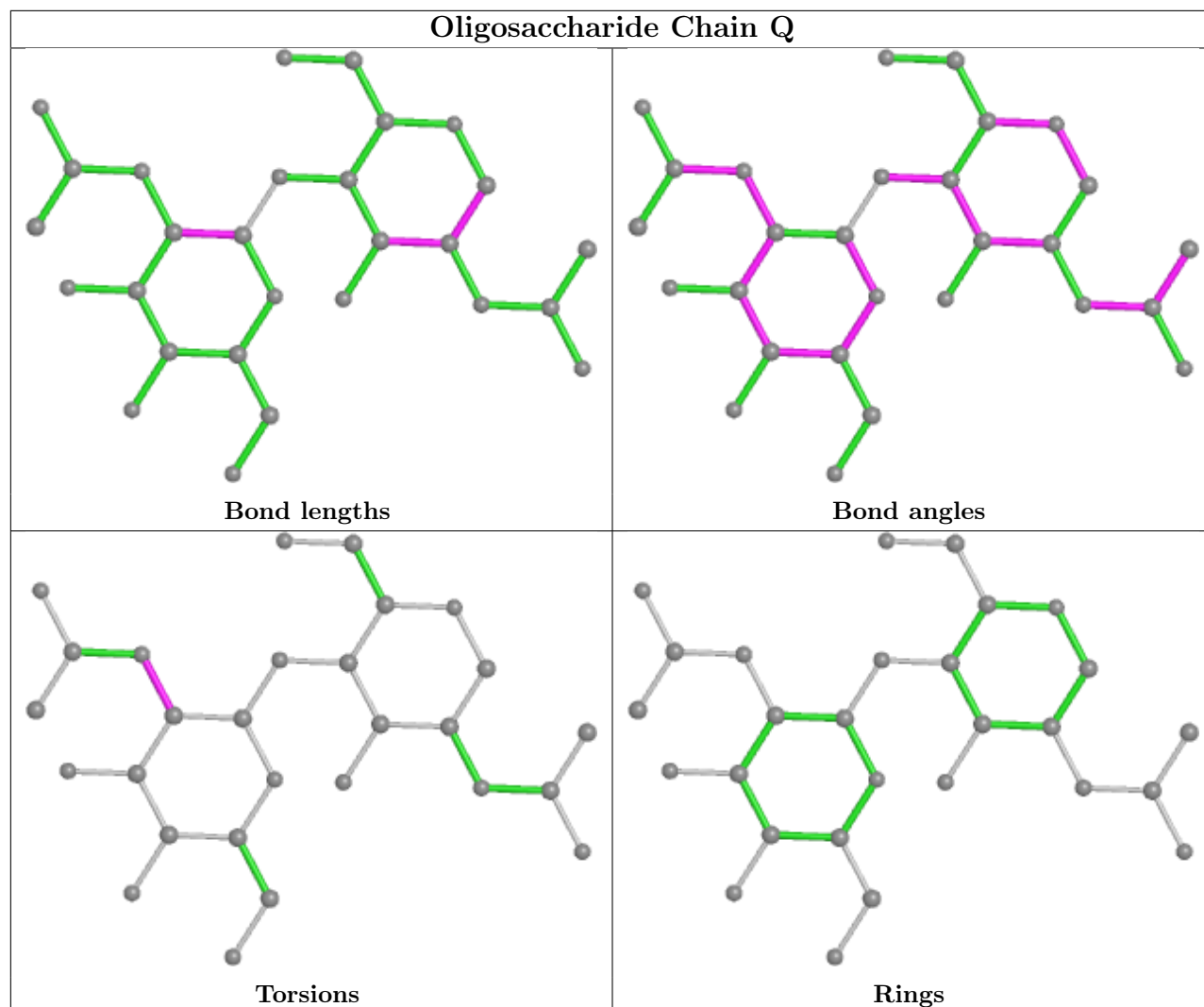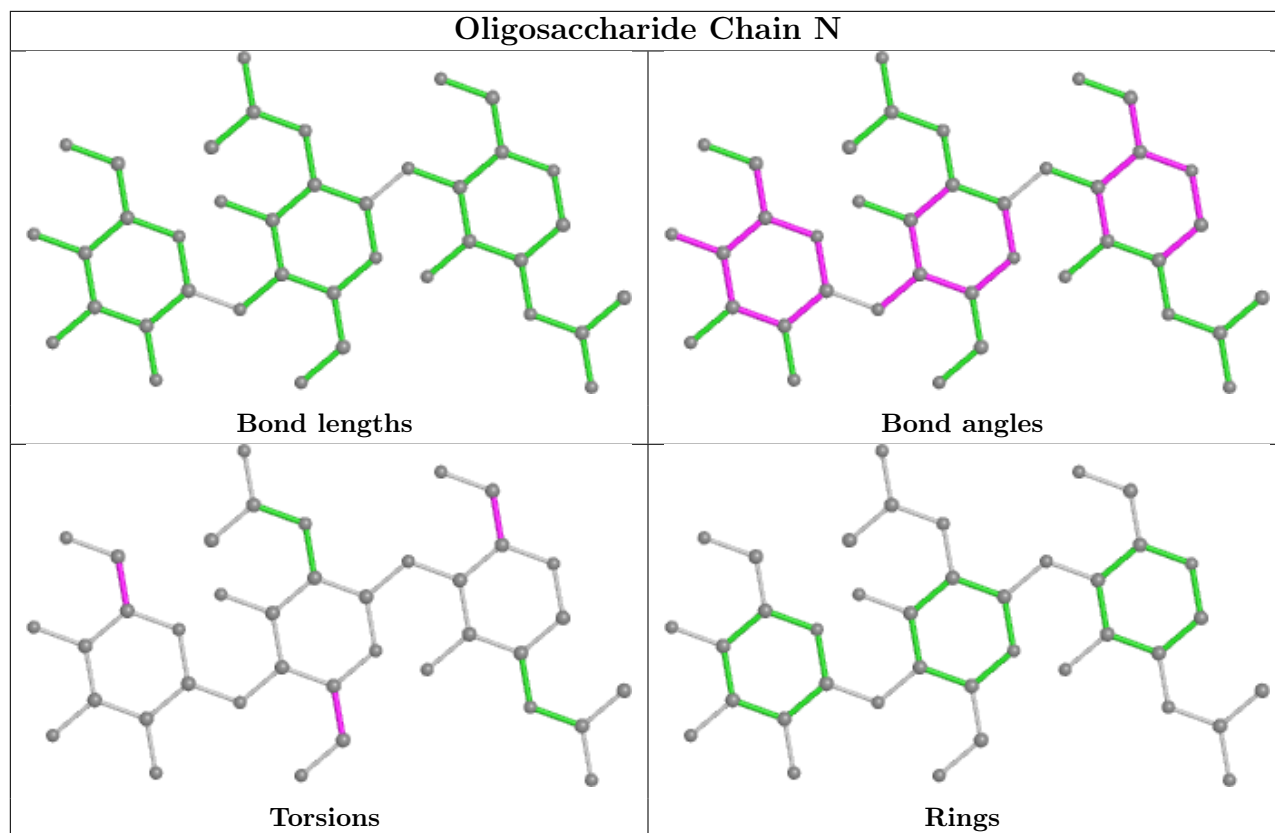
5.6 Ligand geometry ⓘ

4 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	328	1	14,14,15	0.86	1 (7%)	17,19,21	1.56	3 (17%)
5	NAG	E	328	1	14,14,15	0.54	0	17,19,21	1.51	2 (11%)
5	NAG	G	331	1	14,14,15	1.51	2 (14%)	17,19,21	2.32	7 (41%)
5	NAG	G	332	1	14,14,15	0.98	1 (7%)	17,19,21	2.70	3 (17%)

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	328	1	-	<mark>2/6/23/26</mark>	0/1/1/1
5	NAG	E	328	1	-	<mark>2/6/23/26</mark>	0/1/1/1
5	NAG	G	331	1	1/1/5/7	<mark>2/6/23/26</mark>	0/1/1/1
5	NAG	G	332	1	-	0/6/23/26	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	G	331	NAG	C1-C2	4.46	1.59	1.52
5	G	332	NAG	C1-C2	2.42	1.56	1.52
5	G	331	NAG	C3-C2	2.35	1.57	1.52
5	A	328	NAG	C1-C2	2.09	1.55	1.52

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	G	332	NAG	C1-O5-C5	10.14	125.93	112.19
5	G	331	NAG	C1-O5-C5	6.51	121.01	112.19
5	A	328	NAG	C1-O5-C5	3.94	117.53	112.19
5	E	328	NAG	O5-C5-C6	3.92	113.35	107.20
5	G	331	NAG	C3-C4-C5	-3.15	104.63	110.24

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	G	331	NAG	C1

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	328	NAG	C4-C5-C6-O6
5	A	328	NAG	O5-C5-C6-O6
5	E	328	NAG	C4-C5-C6-O6
5	E	328	NAG	O5-C5-C6-O6
5	G	331	NAG	O5-C5-C6-O6

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	328	NAG	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	328	NAG	1	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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, 95th 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^2)	Q<0.9
1	A	325/331 (98%)	-0.13	9 (2%)	53	43	7, 20, 33, 49	0
1	C	323/331 (97%)	-0.27	3 (0%)	84	80	7, 20, 33, 49	0
1	E	323/331 (97%)	-0.07	14 (4%)	35	25	8, 20, 33, 48	0
1	G	322/331 (97%)	-0.21	9 (2%)	53	43	8, 20, 33, 49	0
1	I	322/331 (97%)	0.18	16 (4%)	28	19	8, 20, 33, 48	0
1	K	321/331 (96%)	-0.08	9 (2%)	53	43	8, 20, 33, 49	0
2	B	169/181 (93%)	0.07	3 (1%)	68	61	4, 31, 50, 69	1 (0%)
2	D	171/181 (94%)	-0.03	0	100	100	4, 32, 53, 69	0
2	F	171/181 (94%)	0.63	26 (15%)	2	1	4, 32, 50, 69	0
2	H	170/181 (93%)	0.13	7 (4%)	37	27	4, 32, 50, 69	0
2	J	170/181 (93%)	0.32	7 (4%)	37	27	4, 32, 50, 69	0
2	L	171/181 (94%)	0.67	26 (15%)	2	1	4, 32, 50, 69	0
All	All	2958/3072 (96%)	0.04	129 (4%)	34	24	4, 22, 44, 69	1 (0%)

The worst 5 of 129 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	71	SER	8.1
1	G	71	SER	7.6
1	I	139	ALA	7.4
2	L	164	GLU	6.9
1	K	3	LEU	6.2

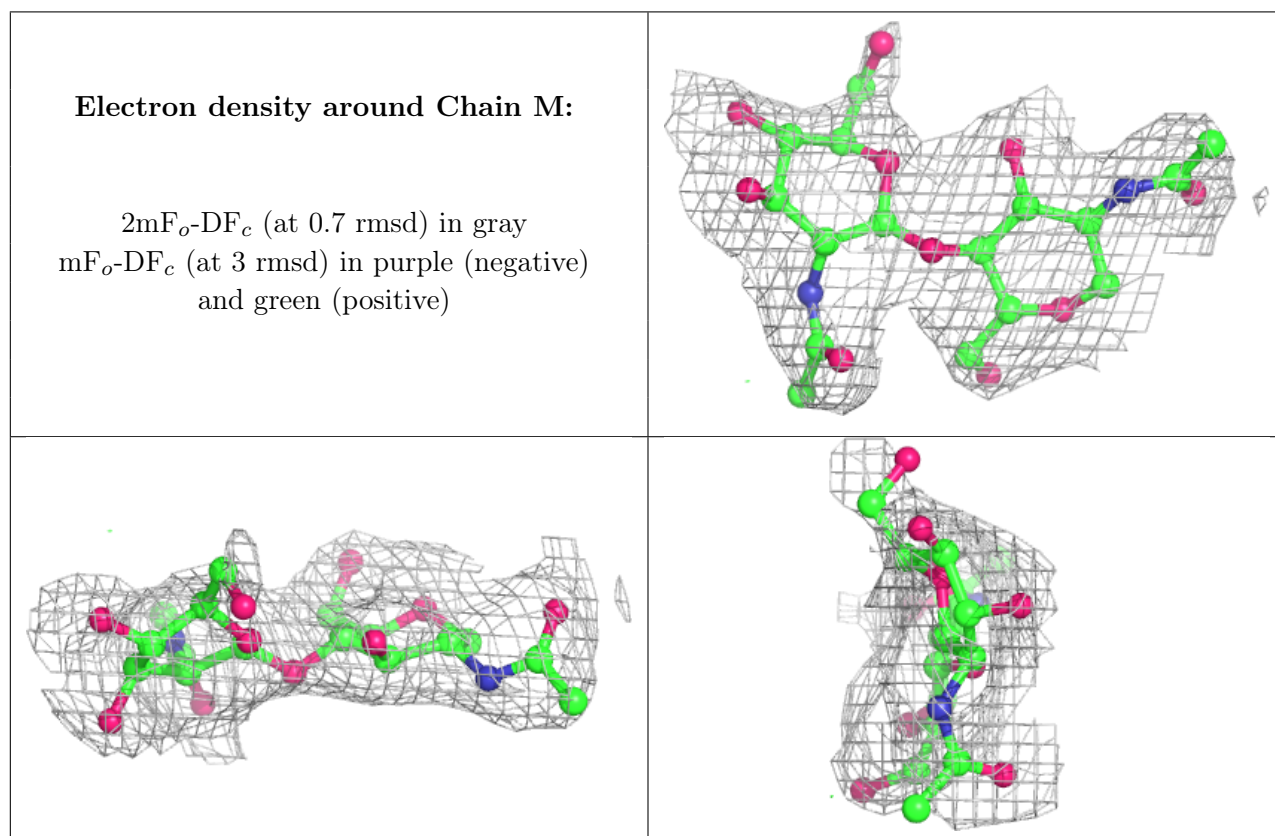## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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, 95th 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^2)	Q<0.9
4	MAN	N	3	11/12	0.69	0.24	77,83,85,86	0
3	NAG	Q	2	14/15	0.73	0.31	81,84,88,88	0
3	NAG	Q	1	14/15	0.73	0.26	56,67,74,78	0
3	NAG	O	2	14/15	0.76	0.26	76,80,82,82	0
3	NAG	M	2	14/15	0.83	0.17	73,77,78,79	0
3	NAG	O	1	14/15	0.86	0.18	52,60,68,69	0
3	NAG	P	2	14/15	0.88	0.23	89,94,101,101	0
4	NAG	N	2	14/15	0.92	0.18	45,58,63,71	0
3	NAG	P	1	14/15	0.93	0.15	61,68,73,82	0
4	NAG	N	1	14/15	0.94	0.20	48,53,55,57	0
3	NAG	M	1	14/15	0.95	0.13	47,52,60,66	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.
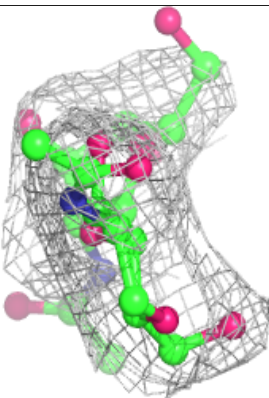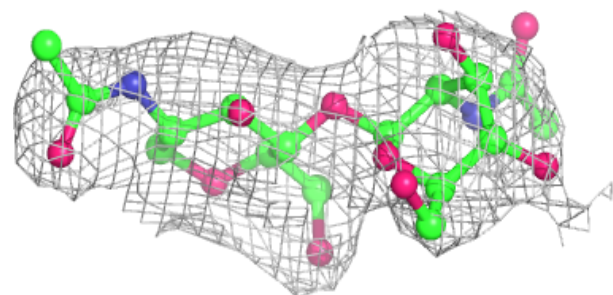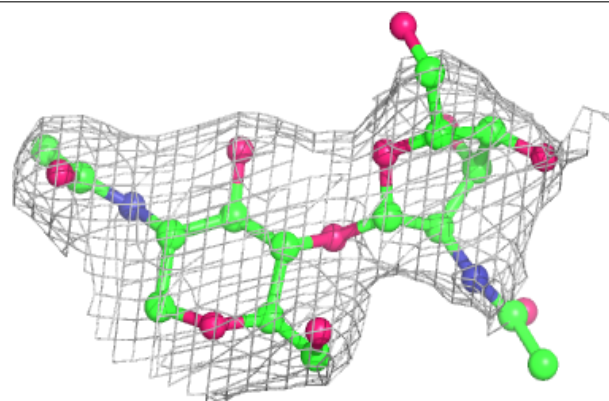
Electron density around Chain M:

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

Electron density around Chain O:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
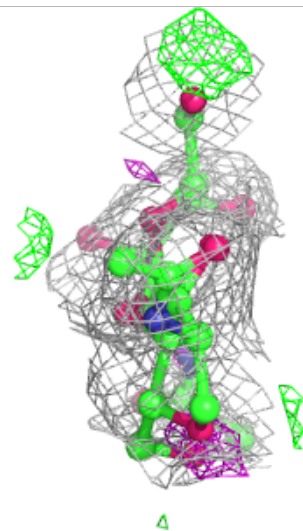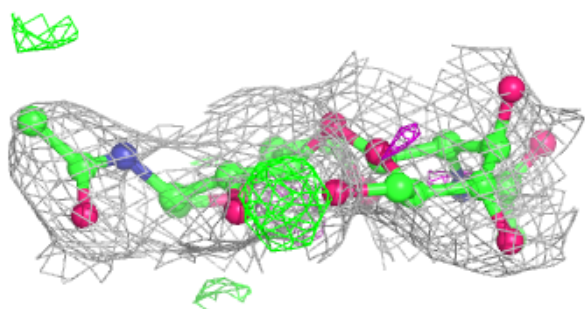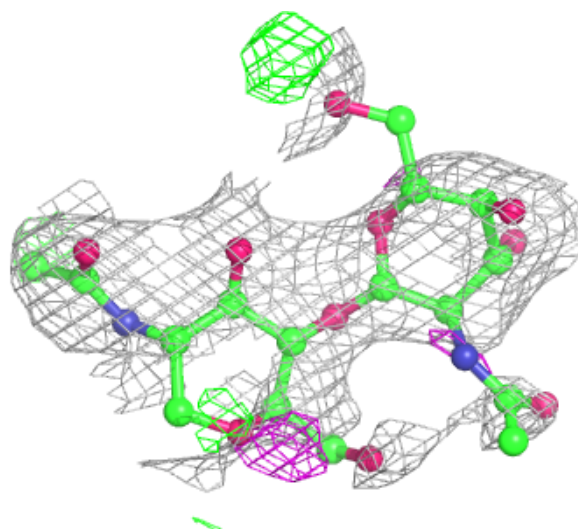mF_o-DF_c (at 3 rmsd) in purple (negative)
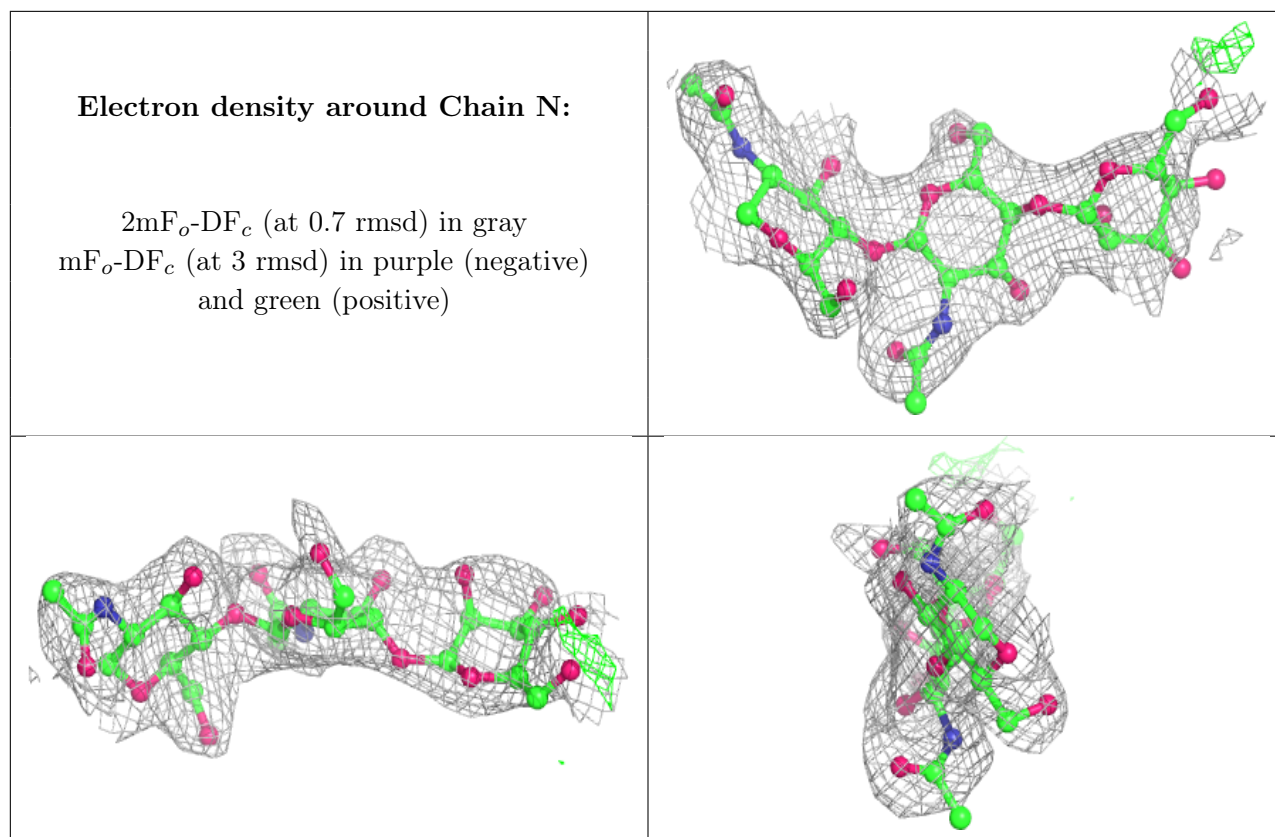and green (positive)
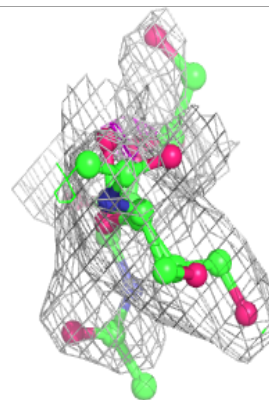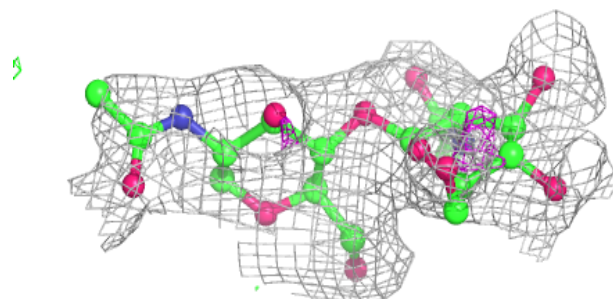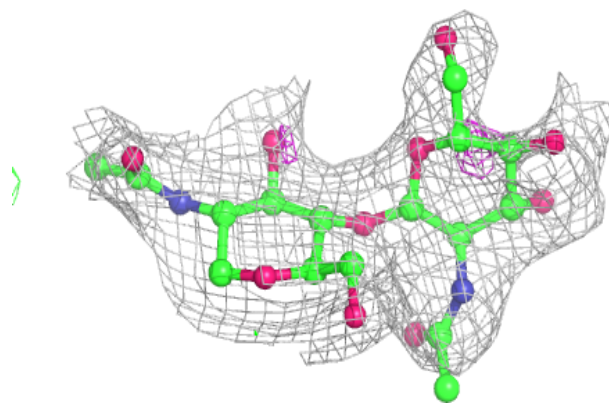
Electron density around Chain P:

$2mF_o$-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 Q:

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

Electron density around Chain N:

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

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th 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($Å^2$)	Q<0.9
5	NAG	G	331	14/15	0.63	0.27	59,72,74,76	0
5	NAG	G	332	14/15	0.73	0.21	64,74,80,81	0
5	NAG	A	328	14/15	0.90	0.17	49,55,59,59	0
5	NAG	E	328	14/15	0.95	0.19	46,51,53,55	0

6.5 Other polymers ⓘ

There are no such residues in this entry.