WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jun 25, 2024 – 11:23 PM EDT

PDB ID	:	6TY1
Title	:	Crystal structure of the haemagglutinin mutant (Gln226Leu, Gly228Ser) from an H10N7 seal influenza virus isolated in Germany in complex with human receptor analogue 6'-SLN
Authors	:	Zhang, J.; Xiong, X.; Purkiss, A.; Walker, P.; Gamblin, S.; Skehel, J.J.
Deposited on	:	2020-01-15
Resolution	:	3.20 Å(reported)

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

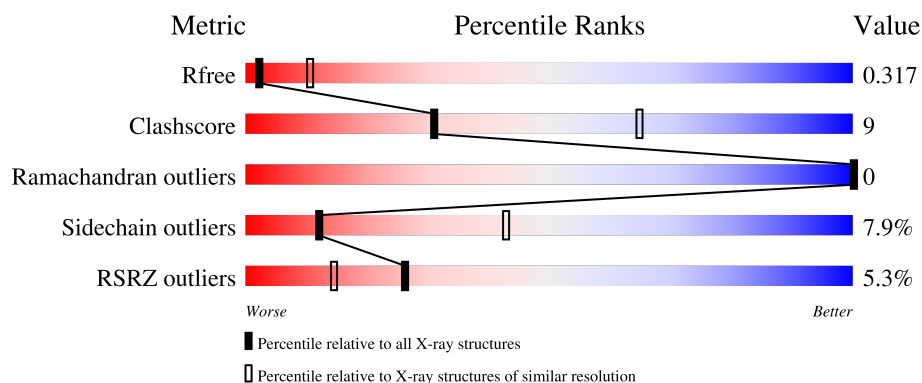|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.20 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		0.317
Clashscore		9
Ramachandran outliers		0
Sidechain outliers		7.9%
RSRZ outliers		5.3%

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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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	325	2% / 79% / 18% / ••
1	C	325	3% / 79% / 19% / ••
1	E	325	2% / 75% / 21% / ••
1	G	325	5% / 79% / 19% / ••

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	I	325	5% / 77% / 21%
1	K	325	10% / 74% / 24%
2	B	177	4% / 72% / 23%
2	D	177	5% / 74% / 21%
2	F	177	7% / 77% / 18%
2	H	177	5% / 68% / 25% / 5%
2	J	177	10% / 75% / 19%
2	L	177	8% / 76% / 20%
3	M	3	100%
3	P	3	67% / 33%
4	N	2	100%
4	O	2	50% / 50%

2 Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 23049 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	320	Total	C	N	O	S	0	0	0
			2443	1515	442	470	16			
1	C	320	Total	C	N	O	S	0	0	0
			2439	1512	441	470	16			
1	E	320	Total	C	N	O	S	0	0	0
			2402	1491	433	462	16			
1	G	320	Total	C	N	O	S	0	0	0
			2423	1504	436	467	16			
1	I	320	Total	C	N	O	S	0	0	0
			2412	1492	438	466	16			
1	K	320	Total	C	N	O	S	0	0	0
			2420	1504	435	465	16			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	ASP	-	expression tag	UNP A0A0A7HR51
A	0	PRO	-	expression tag	UNP A0A0A7HR51
A	221	SER	GLY	engineered mutation	UNP A0A0A7HR51
C	-1	ASP	-	expression tag	UNP A0A0A7HR51
C	0	PRO	-	expression tag	UNP A0A0A7HR51
C	221	SER	GLY	engineered mutation	UNP A0A0A7HR51
E	-1	ASP	-	expression tag	UNP A0A0A7HR51
E	0	PRO	-	expression tag	UNP A0A0A7HR51
E	221	SER	GLY	engineered mutation	UNP A0A0A7HR51
G	-1	ASP	-	expression tag	UNP A0A0A7HR51
G	0	PRO	-	expression tag	UNP A0A0A7HR51
G	221	SER	GLY	engineered mutation	UNP A0A0A7HR51
I	-1	ASP	-	expression tag	UNP A0A0A7HR51
I	0	PRO	-	expression tag	UNP A0A0A7HR51
I	221	SER	GLY	engineered mutation	UNP A0A0A7HR51
K	-1	ASP	-	expression tag	UNP A0A0A7HR51
K	0	PRO	-	expression tag	UNP A0A0A7HR51

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
K	221	SER	GLY	engineered mutation	UNP A0A0A7HR51

- Molecule 2 is a protein called Hemagglutinin HA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	172	Total 1374	C 850	N 238	O 278	S 8	0	0	0
2	D	172	Total 1366	C 846	N 239	O 273	S 8	0	0	0
2	F	172	Total 1373	C 849	N 238	O 278	S 8	0	0	0
2	H	172	Total 1369	C 847	N 236	O 278	S 8	0	0	0
2	J	172	Total 1364	C 842	N 240	O 274	S 8	0	0	0
2	L	172	Total 1366	C 845	N 238	O 275	S 8	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
B	177	LYS	-	expression tag	UNP A0A0A7HR51
D	177	LYS	-	expression tag	UNP A0A0A7HR51
F	177	LYS	-	expression tag	UNP A0A0A7HR51
H	177	LYS	-	expression tag	UNP A0A0A7HR51
J	177	LYS	-	expression tag	UNP A0A0A7HR51
L	177	LYS	-	expression tag	UNP A0A0A7HR51

- Molecule 3 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galacto pyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

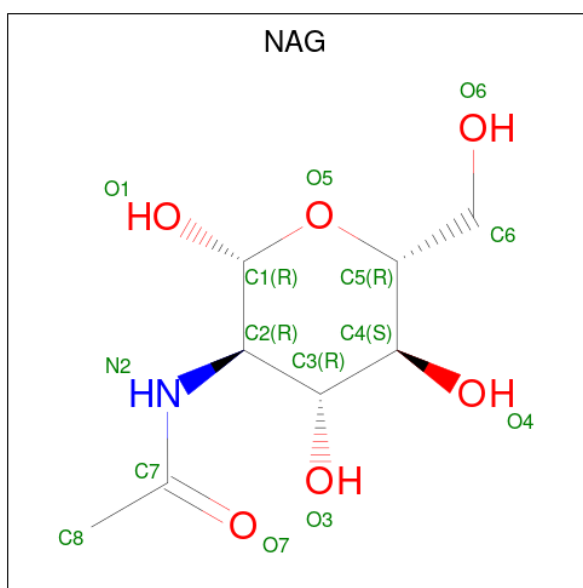| Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	M	3	Total 46	C 25	N 2	O 19	0	0	0
3	P	3	Total 46	C 25	N 2	O 19	0	0	0

- Molecule 4 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galacto
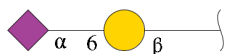
pyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	N	2	Total 32	C 17	N 1	O 14	0	0	0
4	O	2	Total 32	C 17	N 1	O 14	0	0	0

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

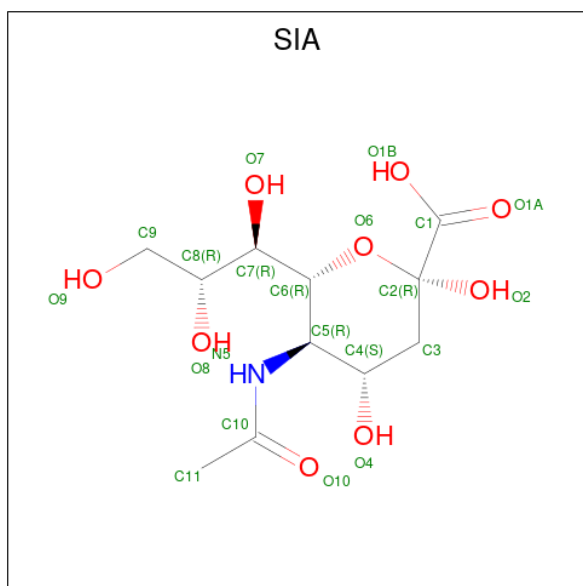| Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	D	1	Total 14	C 8	N 1	O 5	0	0
5	F	1	Total 14	C 8	N 1	O 5	0	0
5	H	1	Total 14	C 8	N 1	O 5	0	0
5	I	1	Total 14	C 8	N 1	O 5	0	0
5	J	1	Total 14	C 8	N 1	O 5	0	0
5	K	1	Total 14	C 8	N 1	O 5	0	0

Continued from previous page...

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

- Molecule 6 is N-acetyl-alpha-neuraminic acid (three-letter code: SIA) (formula: $C_{11}H_{19}NO_9$).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	I	1	Total 21	C 11	N 1	O 9	0	0
6	K	1	Total 21	C 11	N 1	O 9	0	0

- Molecule 7 is water.

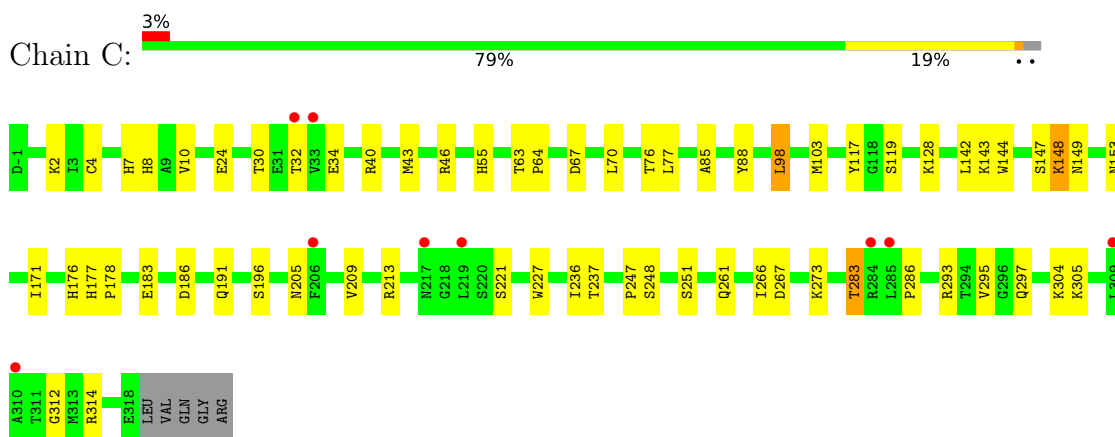| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	B	1	Total 1	O 1	0	0
7	I	1	Total 1	O 1	0	0

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.
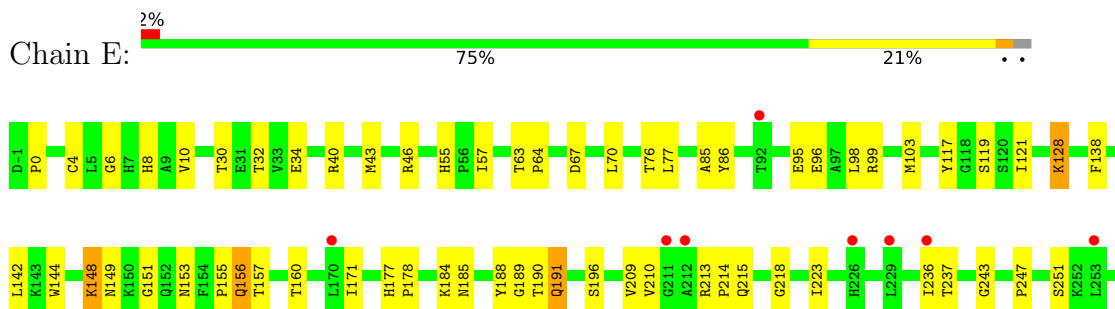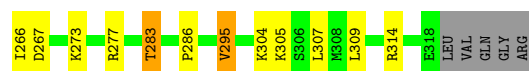
• Molecule 1: Hemagglutinin

• Molecule 1: Hemagglutinin

• Molecule 1: Hemagglutinin

I266 D267 K273 R277 T283 P286 V295 K304 K305 S306 L307 M308 L309 R314 E318 LEU VAL GLN ARG

• Molecule 1: Hemagglutinin

Chain G:

5%

79%   19%

D-1 C4 L5 G6 H7 H8 A9 V10 T14 L19 E24 E31 T30 T32 V33 E34 R40 M43 H55 P56 I57 I61 G62 T63 P64 D67 L70 T73 W74 D75 T76 L77 A85 Y88 T92 L98 M103 Y117 G118 S119 K128

Y139 W144 S147 K148 T157 T158 N159 T160 E168 I171 M172 H176 H177 P178 E183 K184 N185 D186 Q191 S196 K203 N204 N205 V209 A212 R213 P214 Q215 V216 N217 S221 H226 W227 T228 L229 I236 T237 F238 S239 S251 I266 D267
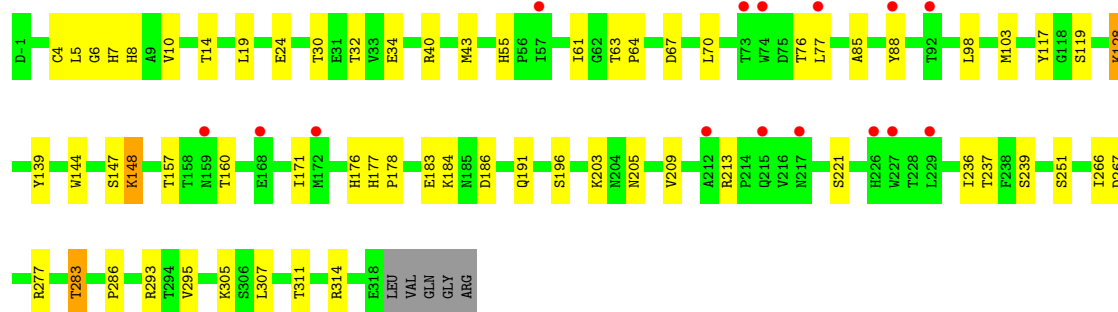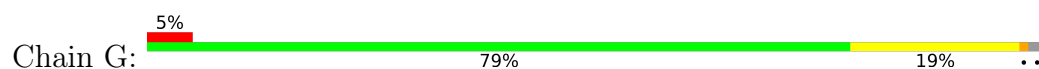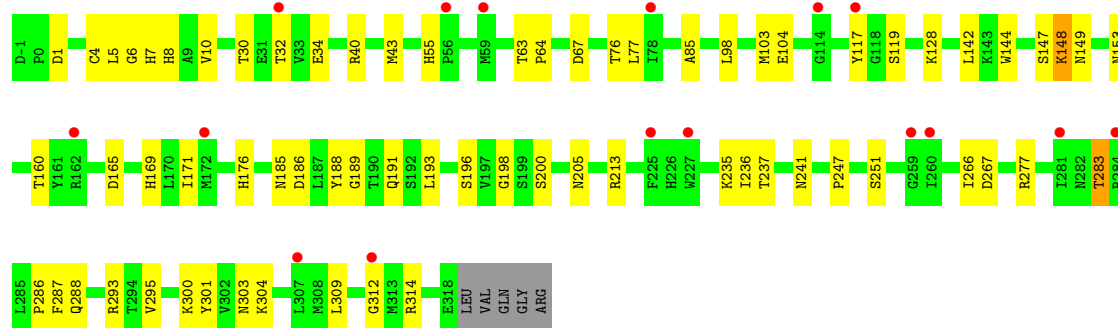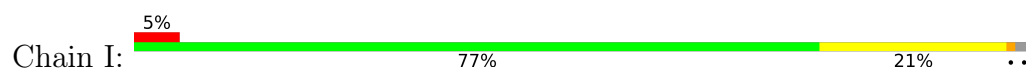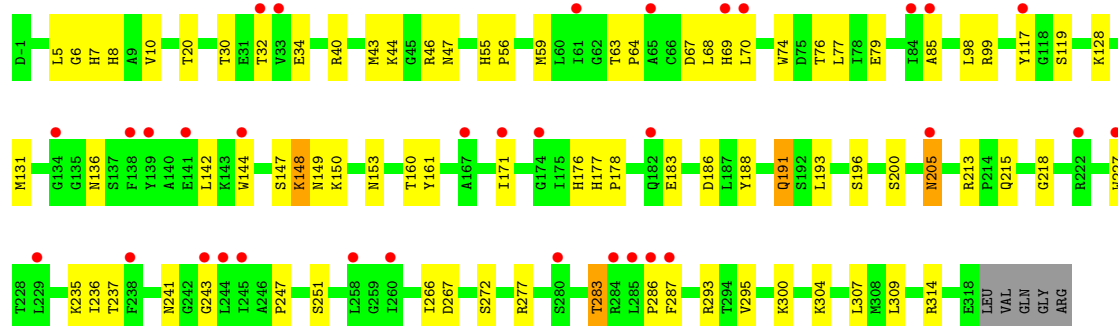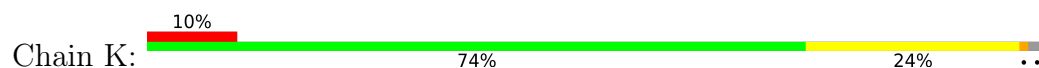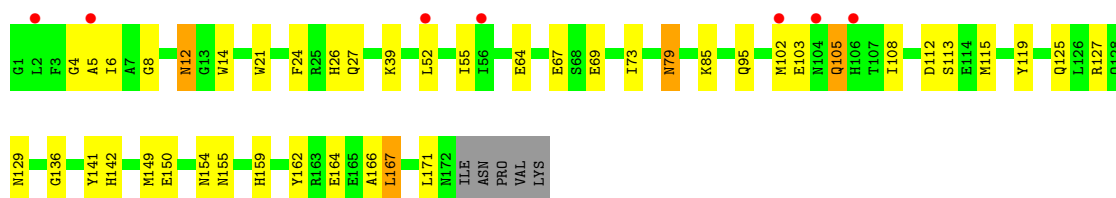
R277 T283 P286 R293 T294 V295 K305 S306 L307 T311 R314 E318 LEU VAL GLN ARG

• Molecule 1: Hemagglutinin

Chain I:

5%

77%   21%

D-1 P0 D1 C4 L5 G6 H7 H8 A9 V10 T30 E31 T32 V33 E34 R40 M43 H55 P56 M59 T63 P64 D67 T76 L77 I78 A85 L98 M103 E104 G114 Y117 G118 S119 K128 L142 K143 W144 S147 K148 N149 N153

T160 Y161 R162 D165 H169 L170 I171 M172 H176 N185 D186 L187 Y188 G189 T190 Q191 S192 L193 S196 V197 G198 S199 S200 N205 R213 F225 H226 W227 K235 I236 T237 N241 P247 S251 G259 I260 I266 D267 R277 I281 N282 T283 R284

L285 P286 F287 Q288 R293 T294 V295 K300 Y301 V302 N303 K304 L307 M308 L309 G312 M313 R314 E318 LEU VAL GLN ARG

• Molecule 1: Hemagglutinin

Chain K:

10%

74%   24%

D-1 L5 G6 H7 H8 A9 V10 T20 T30 E31 T32 V33 E34 R40 M43 K44 G45 R46 N47 H55 P56 M59 L60 I61 G62 T63 P64 A65 C66 D67 L68 H69 L70 W74 D75 T76 L77 I78 E79 I84 A85 L98 R99 Y117 G118 S119 K128

M131 G134 G135 N136 S137 F138 Y139 A140 E141 L142 K143 W144 S147 K148 N149 K150 N153 T160 Y161 A167 I171 G174 I175 H176 H177 P178 Q182 E183 D186 L187 Y188 Q191 S192 L193 S196 S200 N205 R213 P214 Q215 G218 R222 W227

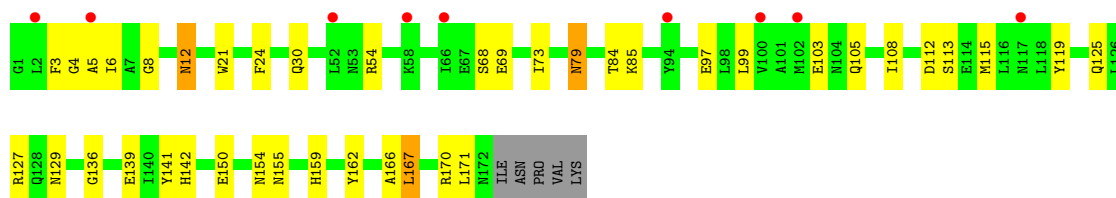T228 L229 K235 I236 T237 F238 N241 G242 G243 L244 I245 A246 P247 S251 L258 G259 I260 I266 D267 S272 R277 S280 T283 R284 L285 P286 F287 R293 T294 V295 K300 K304 L307 M308 L309 R314 E318 LEU VAL GLN ARG

• Molecule 2: Hemagglutinin HA2

Chain B:

4%

72%   23%

G1 L2 F3 G4 A5 I6 A7 G8 N12 G13 W14 W21 F24 R25 H26 Q27 K39 L52 I55 I56 E64 E67 S68 E69 I73 N79 K85 Q95 M102 E103 N104 Q105 H106 T107 I108 D112 S113 E114 M115 Y119 Q125 L126 R127 Q128

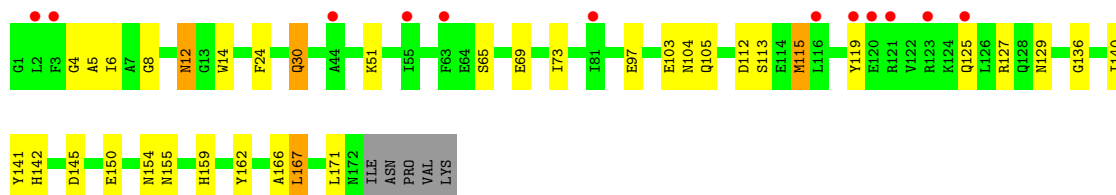N129 G136 Y141 H142 M149 E150 N154 N155 H159 Y162 R163 E164 E165 A166 L167 L171 N172 ILE ASN PRO VAL LYS

• Molecule 2: Hemagglutinin HA2

Chain D:

5%
74%   21%

G1 L2 F3 G4 A5 I6 A7 G8 N12 W21 F24 Q30 L52 N53 R54 K58 I66 E67 S68 E69 I73 N79 T84 K85 Y94 E97 L98 L99 V100 A101 M102 E103 N104 Q105 I108 D112 S113 E114 M115 L116 N117 L118 Y119 Q125 L126

R127 Q128 N129 G136 E139 I140 Y141 H142 E150 N154 N155 H159 Y162 A166 L167 R170 L171 N172 ILE ASN PRO VAL LYS

• Molecule 2: Hemagglutinin HA2

Chain F:

7%
77%   18%

G1 L2 F3 G4 A5 I6 A7 G8 N12 G13 W14 F24 Q30 A44 K51 I55 F63 E64 S65 E69 I73 I81 E97 E103 N104 Q105 D112 S113 E114 M115 L116 Y119 E120 R121 V122 R123 K124 Q125 L126 R127 Q128 N129 G136 I140

Y141 H142 D145 E150 N154 N155 H159 Y162 A166 L167 L171 N172 ILE ASN PRO VAL LYS

• Molecule 2: Hemagglutinin HA2

Chain H:

5%
68%   25%   5% •

G1 L2 F3 G4 A5 I6 A7 G8 N12 G13 W14 W21 F24 R25 H26 Q30 K39 I48 T49 G50 K51 I55 F63 E67 S68 E69 F70 S71 E72 I73 D74 I77 G78 N79 W83 T84 Y94 V100 E103 N104 Q105 I108 D112 S113

E114 M115 L116 Y119 E120 R121 V122 R123 K124 Q125 L126 R127 Q128 N129 G136 Y141 H142 M149 E150 N154 N155 H159 Y162 A166 L167 R170 L171 N172 ILE ASN PRO VAL LYS

• Molecule 2: Hemagglutinin HA2

Chain J:

10%
75%   19% • •

G1 G4 A5 I6 A7 G8 N12 G13 W14 D19 G20 W21 Y22 G23 F24 N28 A29 Q30 Q34 A35 A36 D37 I48 I55 I56 K57 K58 E62 E67 S68 E69 I73 N79 T84 D90 I91 W92 Q95 A96 E97 M102 E103 N104 Q105 H106

D112 S113 E114 M115 Y119 Q125 L126 R127 Q128 N129 F138 E139 H142 E150 S151 I152 R153 N154 N155 T156 H159 S160 Q161 Y162 L167 L171 N172 ILE ASN PRO VAL LYS
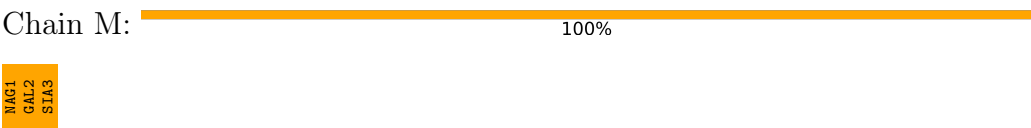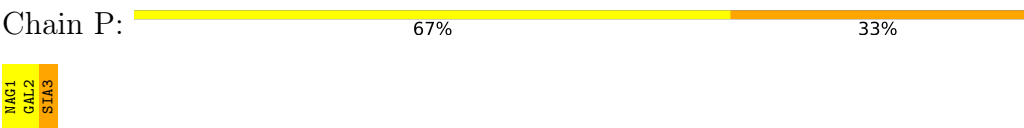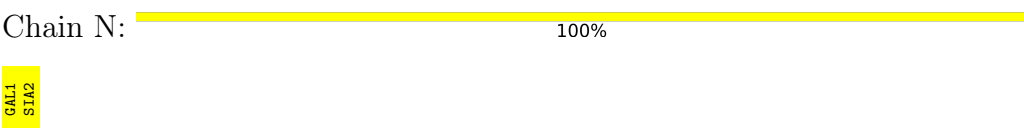
● Molecule 2: Hemagglutinin HA2

Chain L:

● Molecule 3: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain M:

● Molecule 3: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain P:

● Molecule 4: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose

Chain N:

● Molecule 4: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose

Chain O:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	68.99Å 212.67Å 154.90Å 90.00° 101.58° 90.00°	Depositor
Resolution (Å)	87.08 – 3.20 87.08 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.9 (87.08-3.20) 99.9 (87.08-3.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.85 (at 3.19Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.272 , 0.336 0.257 , 0.317	Depositor DCC
R_{free} test set	3549 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å^2)	77.7	Xtriage
Anisotropy	0.712	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 42.4	EDS
L-test for twinning[2]	$< |L| > = 0.45, < L^2 > = 0.27$	Xtriage
Estimated twinning fraction	0.014 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	23049	wwPDB-VP
Average B, all atoms (Å^2)	121.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 29.41 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.5567e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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.

5   Model quality ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GAL, SIA, 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	0.71	0/2493	0.85	0/3379
1	C	0.70	0/2489	0.84	0/3375
1	E	0.73	0/2451	0.85	0/3328
1	G	0.70	0/2473	0.83	0/3356
1	I	0.71	0/2460	0.82	0/3337
1	K	0.70	0/2470	0.84	0/3351
2	B	0.72	0/1399	0.84	0/1889
2	D	0.70	0/1391	0.82	0/1879
2	F	0.73	0/1398	0.84	0/1887
2	H	0.72	0/1394	0.83	0/1882
2	J	0.71	0/1388	0.83	0/1873
2	L	0.70	0/1390	0.85	0/1876
All	All	0.71	0/23196	0.84	0/31412

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2   Too-close contacts ⓘ

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2443	0	2399	44	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2439	0	2388	37	0
1	E	2402	0	2337	57	0
1	G	2423	0	2360	42	0
1	I	2412	0	2349	41	0
1	K	2420	0	2364	63	0
2	B	1374	0	1269	36	0
2	D	1366	0	1262	39	0
2	F	1373	0	1267	24	0
2	H	1369	0	1259	44	0
2	J	1364	0	1261	33	0
2	L	1366	0	1263	33	0
3	M	46	0	40	3	0
3	P	46	0	40	8	0
4	N	32	0	28	0	0
4	O	32	0	28	2	0
5	D	14	0	13	3	0
5	F	14	0	13	1	0
5	H	14	0	13	0	0
5	I	14	0	13	0	0
5	J	14	0	13	3	0
5	K	14	0	13	0	0
5	L	14	0	13	0	0
6	I	21	0	18	1	0
6	K	21	0	18	0	0
7	B	1	0	0	1	0
7	I	1	0	0	0	0
All	All	23049	0	22041	408	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 9.

The worst 5 of 408 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:156:GLN:HE21	1:E:156:GLN:HA	1.13	1.06
1:K:131:MET:HG2	1:K:136:ASN:HA	1.50	0.93
2:B:105:GLN:HE21	2:B:105:GLN:HA	1.40	0.86
2:D:99:LEU:HD22	2:H:94:TYR:OH	1.77	0.85
2:D:79:ASN:ND2	5:D:201:NAG:H81	1.92	0.85

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	318/325 (98%)	310 (98%)	8 (2%)	0	100	100
1	C	318/325 (98%)	310 (98%)	8 (2%)	0	100	100
1	E	318/325 (98%)	311 (98%)	7 (2%)	0	100	100
1	G	318/325 (98%)	310 (98%)	8 (2%)	0	100	100
1	I	318/325 (98%)	309 (97%)	9 (3%)	0	100	100
1	K	318/325 (98%)	310 (98%)	8 (2%)	0	100	100
2	B	170/177 (96%)	158 (93%)	12 (7%)	0	100	100
2	D	170/177 (96%)	156 (92%)	14 (8%)	0	100	100
2	F	170/177 (96%)	158 (93%)	12 (7%)	0	100	100
2	H	170/177 (96%)	157 (92%)	13 (8%)	0	100	100
2	J	170/177 (96%)	158 (93%)	12 (7%)	0	100	100
2	L	170/177 (96%)	159 (94%)	11 (6%)	0	100	100
All	All	2928/3012 (97%)	2806 (96%)	122 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	271/276 (98%)	249 (92%)	22 (8%)	11	42
1	C	270/276 (98%)	247 (92%)	23 (8%)	10	38

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	262/276 (95%)	240 (92%)	22 (8%)	11	39
1	G	266/276 (96%)	243 (91%)	23 (9%)	10	38
1	I	265/276 (96%)	244 (92%)	21 (8%)	12	43
1	K	266/276 (96%)	241 (91%)	25 (9%)	8	33
2	B	143/151 (95%)	132 (92%)	11 (8%)	13	44
2	D	141/151 (93%)	135 (96%)	6 (4%)	29	64
2	F	143/151 (95%)	136 (95%)	7 (5%)	25	61
2	H	142/151 (94%)	130 (92%)	12 (8%)	10	38
2	J	141/151 (93%)	128 (91%)	13 (9%)	9	33
2	L	141/151 (93%)	132 (94%)	9 (6%)	17	52
All	All	2451/2562 (96%)	2257 (92%)	194 (8%)	12	43

5 of 194 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	30	GLN
1	I	251	SER
2	H	74	ASP
1	I	67	ASP
2	J	73	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 52 such sidechains are listed below:

Mol	Chain	Res	Type
2	H	30	GLN
1	I	303	ASN
2	L	79	ASN
2	H	105	GLN
1	I	133	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

10 monosaccharides are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
3	NAG	M	1	3	15,15,15	0.58	0	21,21,21	1.91	5 (23%)
3	GAL	M	2	3	11,11,12	0.58	0	15,15,17	1.95	4 (26%)
3	SIA	M	3	3	20,20,21	1.19	1 (5%)	24,28,31	2.50	13 (54%)
4	GAL	N	1	4	12,12,12	0.79	0	17,17,17	1.73	6 (35%)
4	SIA	N	2	4	20,20,21	1.17	1 (5%)	24,28,31	1.94	6 (25%)
4	GAL	O	1	4	12,12,12	0.57	0	17,17,17	1.29	2 (11%)
4	SIA	O	2	4	20,20,21	0.75	1 (5%)	24,28,31	1.17	1 (4%)
3	NAG	P	1	3	15,15,15	0.77	0	21,21,21	1.75	6 (28%)
3	GAL	P	2	3	11,11,12	0.77	0	15,15,17	1.61	2 (13%)
3	SIA	P	3	3	20,20,21	0.73	0	24,28,31	1.75	6 (25%)

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	-	2/6/26/26	0/1/1/1
3	GAL	M	2	3	-	0/2/19/22	0/1/1/1
3	SIA	M	3	3	-	8/18/34/38	0/1/1/1
4	GAL	N	1	4	-	0/2/22/22	0/1/1/1
4	SIA	N	2	4	-	5/18/34/38	0/1/1/1
4	GAL	O	1	4	-	0/2/22/22	0/1/1/1
4	SIA	O	2	4	-	0/18/34/38	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	P	1	3	-	0/6/26/26	0/1/1/1
3	GAL	P	2	3	-	0/2/19/22	0/1/1/1
3	SIA	P	3	3	-	8/18/34/38	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	M	3	SIA	C8-C7	3.71	1.60	1.53
4	N	2	SIA	C2-C1	-3.70	1.49	1.52
4	O	2	SIA	C2-C1	-2.11	1.50	1.52

The worst 5 of 51 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	M	3	SIA	C3-C4-C5	-5.50	104.81	111.46
4	N	2	SIA	C6-C5-N5	-5.06	102.51	110.91
3	P	2	GAL	C1-O5-C5	4.25	117.95	112.19
3	M	1	NAG	C3-C4-C5	3.92	117.23	110.24
3	M	1	NAG	C1-C2-C3	-3.89	105.24	110.54

There are no chirality outliers.

5 of 23 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	M	3	SIA	O6-C6-C7-O7
3	M	3	SIA	C6-C7-C8-C9
3	M	3	SIA	C6-C7-C8-O8
3	M	3	SIA	O7-C7-C8-C9
3	M	3	SIA	O7-C7-C8-O8

There are no ring outliers.

5 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	M	2	GAL	1	0
3	P	3	SIA	8	0
3	M	3	SIA	2	0
3	M	1	NAG	1	0
4	O	2	SIA	2	0

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

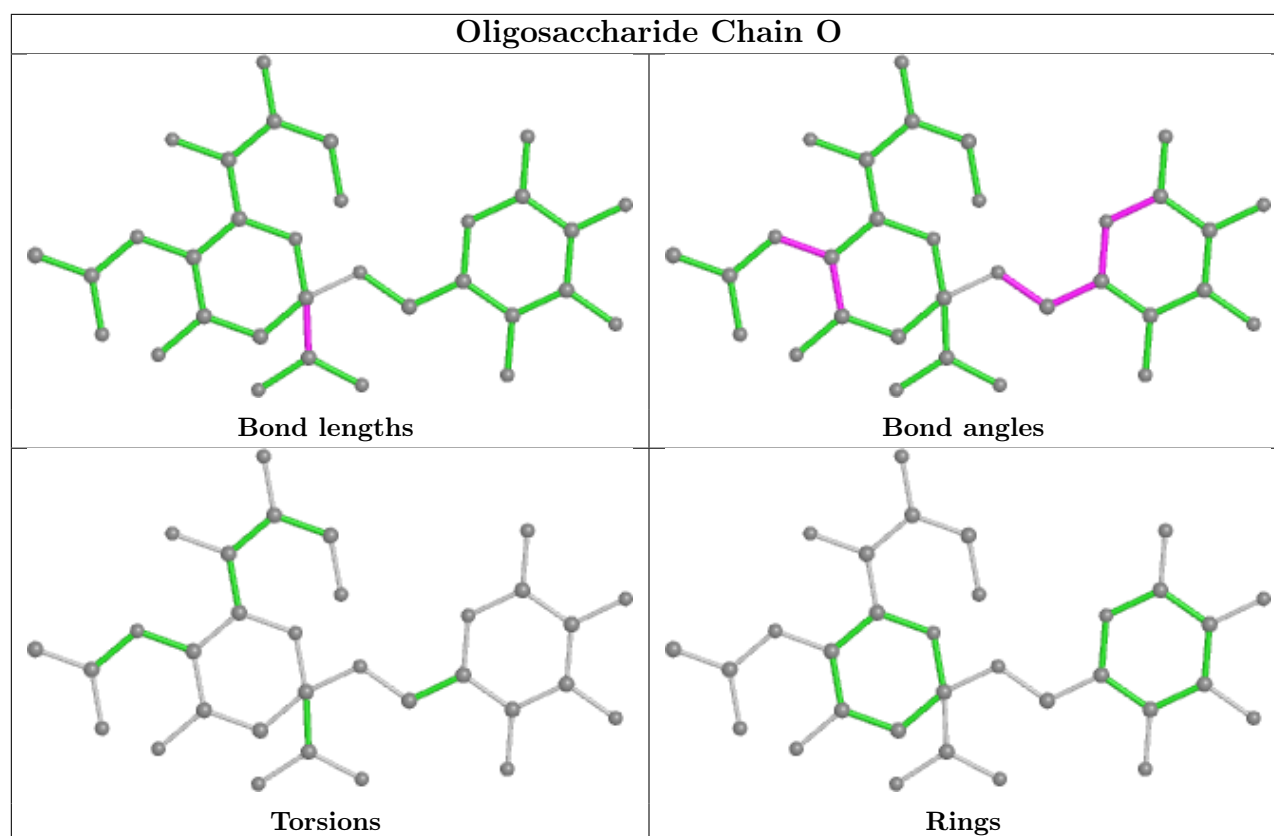bond angles, torsion angles, and ring geometry for oligosaccharide.

Oligosaccharide Chain M	
Bond lengths	Bond angles
Torsions	Rings

Oligosaccharide Chain P

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain N

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain O

Bond lengths | Bond angles

Torsions | Rings
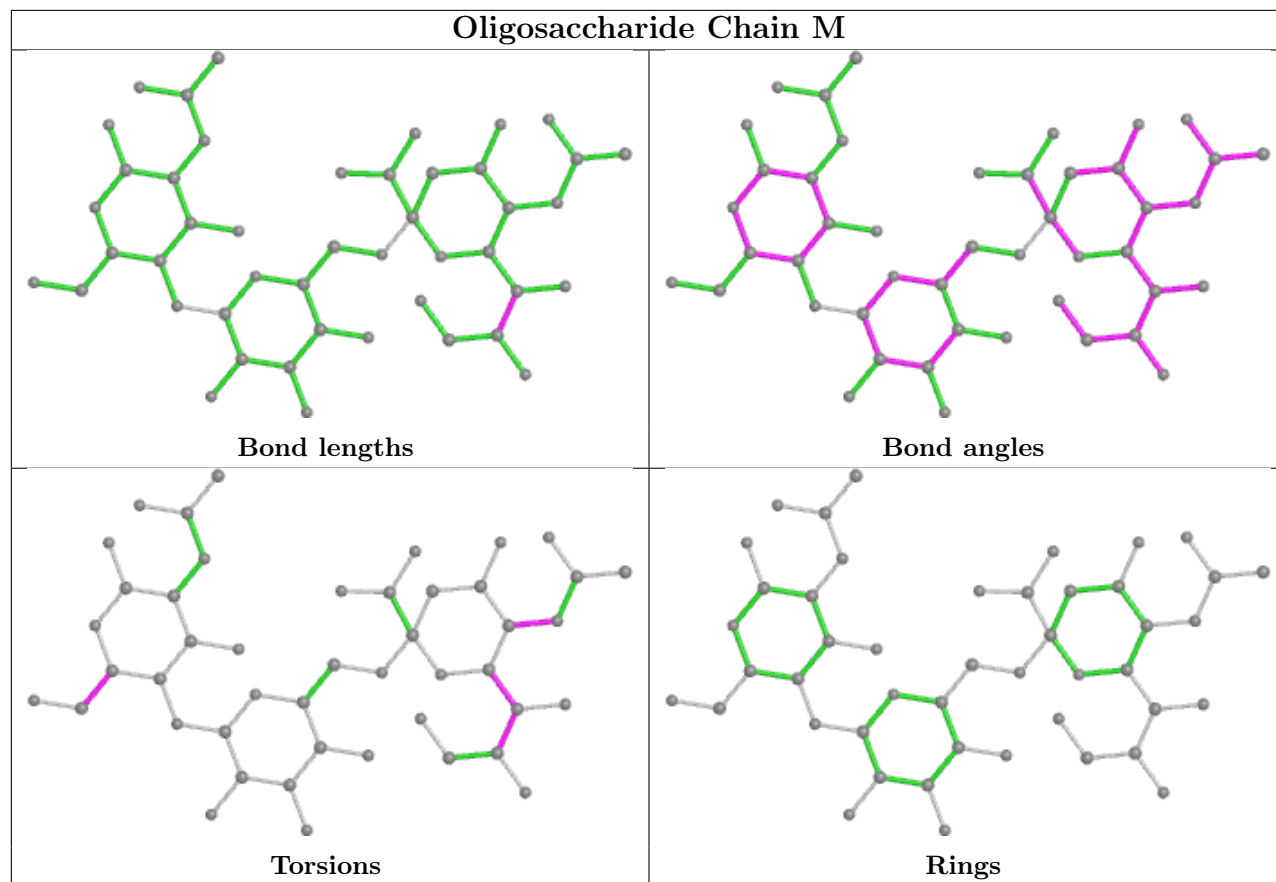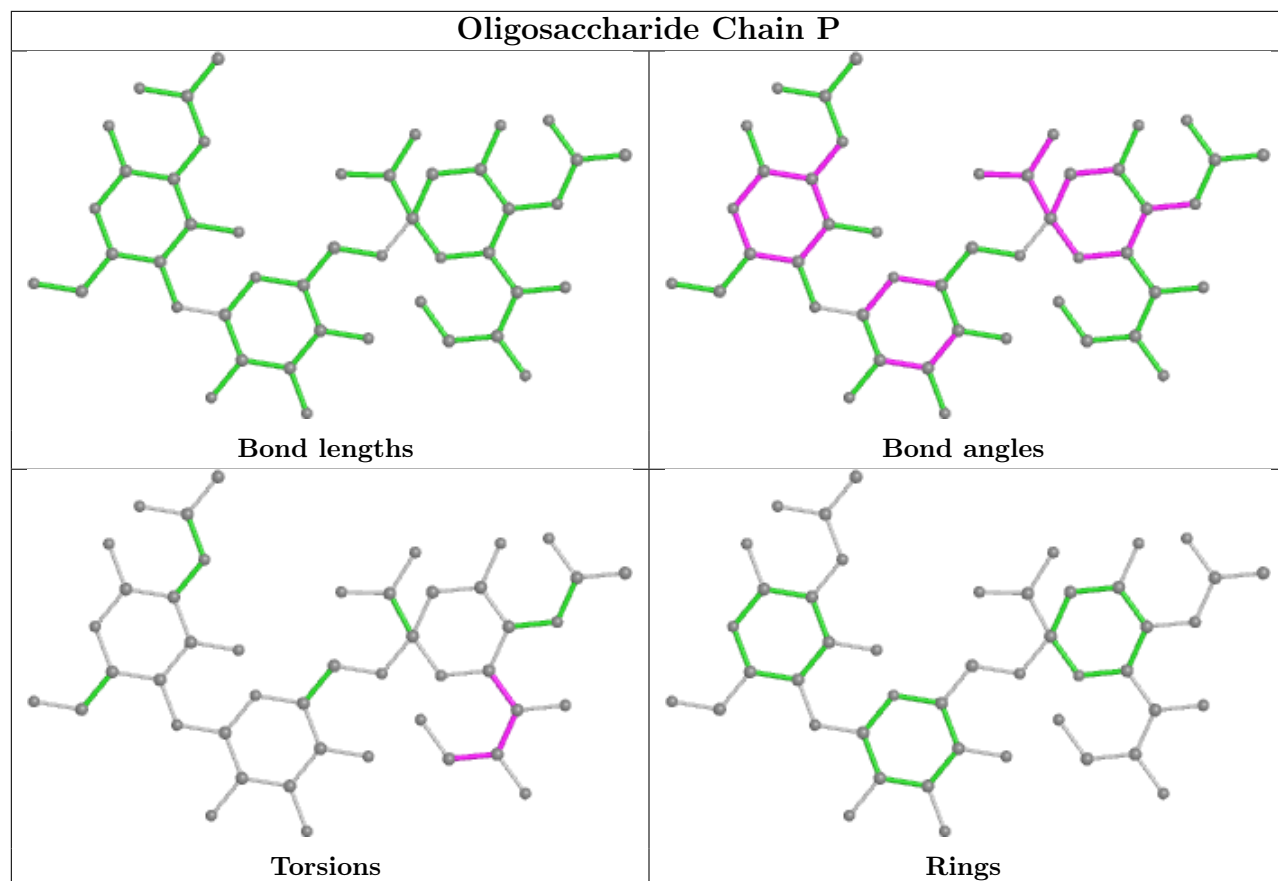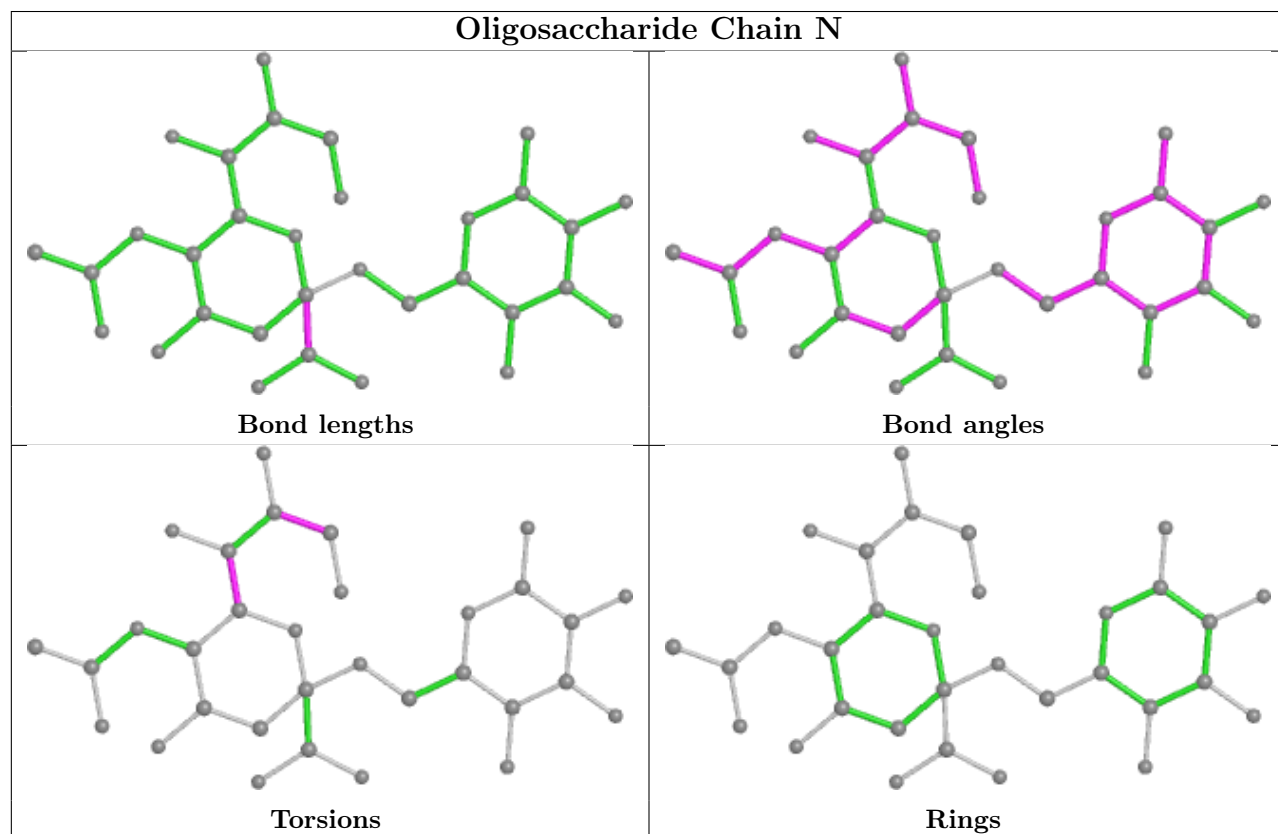
5.6 Ligand geometry ⓘ

9 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	J	201	2	14,14,15	1.04	0	17,19,21	2.74	8 (47%)
5	NAG	H	201	2	14,14,15	1.31	3 (21%)	17,19,21	3.10	7 (41%)
5	NAG	F	201	2	14,14,15	1.75	3 (21%)	17,19,21	3.34	9 (52%)
5	NAG	I	401	1	14,14,15	0.78	0	17,19,21	1.48	4 (23%)
5	NAG	L	201	2	14,14,15	0.70	0	17,19,21	1.71	5 (29%)
5	NAG	K	401	1	14,14,15	1.04	1 (7%)	17,19,21	1.50	4 (23%)
6	SIA	K	402	-	21,21,21	1.44	2 (9%)	25,31,31	1.64	6 (24%)
6	SIA	I	402	-	21,21,21	1.55	3 (14%)	25,31,31	2.13	8 (32%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
5	NAG	D	201	2	14,14,15	0.99	0	17,19,21	2.15	6 (35%)

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	J	201	2	-	2/6/23/26	0/1/1/1
5	NAG	H	201	2	-	2/6/23/26	0/1/1/1
5	NAG	F	201	2	-	1/6/23/26	0/1/1/1
5	NAG	I	401	1	-	0/6/23/26	0/1/1/1
5	NAG	L	201	2	-	0/6/23/26	0/1/1/1
5	NAG	K	401	1	-	0/6/23/26	0/1/1/1
6	SIA	K	402	-	-	9/20/38/38	0/1/1/1
6	SIA	I	402	-	-	7/20/38/38	0/1/1/1
5	NAG	D	201	2	-	1/6/23/26	0/1/1/1

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	K	402	SIA	O2-C2	5.13	1.46	1.39
6	I	402	SIA	O2-C2	4.88	1.45	1.39
5	F	201	NAG	C1-C2	3.31	1.57	1.52
5	F	201	NAG	O5-C5	3.23	1.50	1.43
6	I	402	SIA	O6-C2	3.01	1.46	1.43

The worst 5 of 57 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	F	201	NAG	O5-C5-C6	8.31	120.24	107.20
5	H	201	NAG	C1-O5-C5	7.42	122.25	112.19
5	F	201	NAG	C1-C2-N2	7.03	122.50	110.49
5	H	201	NAG	C8-C7-N2	-5.10	107.46	116.10
5	J	201	NAG	C8-C7-N2	-5.07	107.51	116.10

There are no chirality outliers.

5 of 22 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	I	402	SIA	O8-C8-C9-O9
6	K	402	SIA	O1A-C1-C2-O2
6	K	402	SIA	O1B-C1-C2-O2
6	K	402	SIA	O1B-C1-C2-O6
6	K	402	SIA	C5-C6-C7-O7

There are no ring outliers.

4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	J	201	NAG	3	0
5	F	201	NAG	1	0
6	I	402	SIA	1	0
5	D	201	NAG	3	0

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

Ligand SIA K 402

Bond lengths

Bond angles

Torsions

Rings

Ligand SIA I 402

Bond lengths

Bond angles

Torsions

Rings
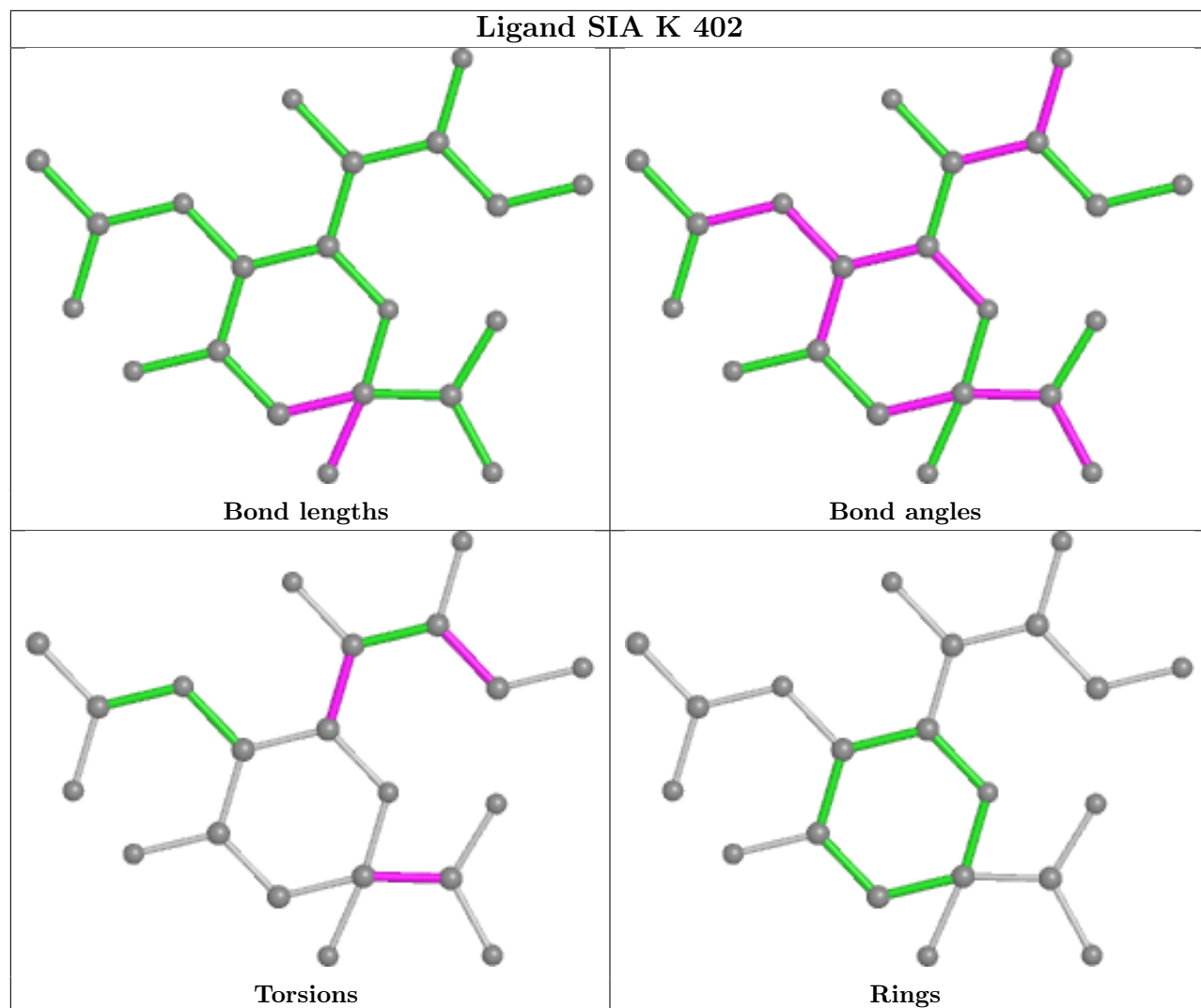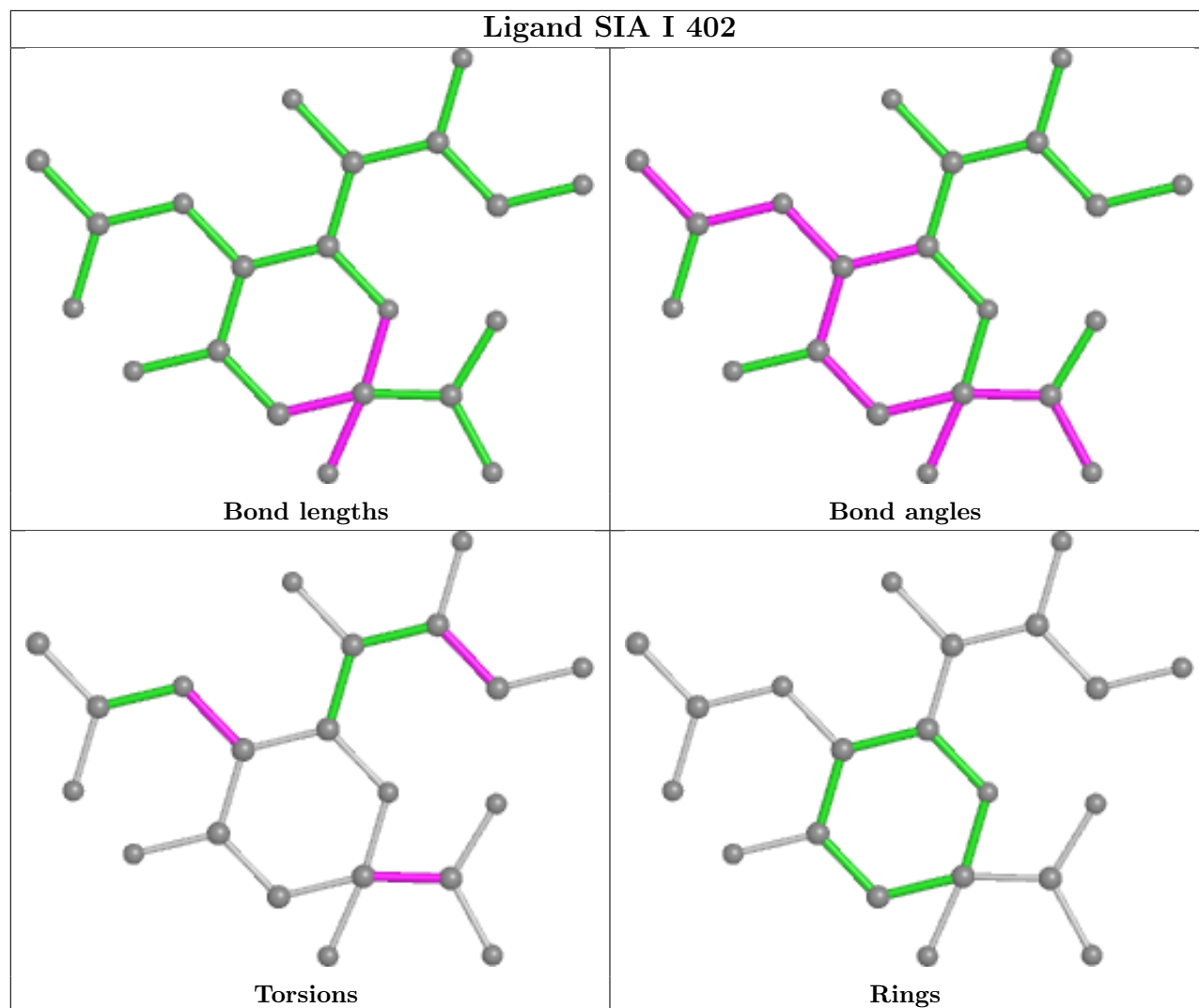
5.7 Other polymers (i)

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, 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(Å²)	Q<0.9
1	A	320/325 (98%)	0.15	7 (2%)	62	48	66, 108, 154, 190	0
1	C	320/325 (98%)	0.18	9 (2%)	53	37	66, 103, 151, 179	0
1	E	320/325 (98%)	0.20	8 (2%)	57	43	70, 125, 201, 246	0
1	G	320/325 (98%)	0.30	15 (4%)	31	19	68, 126, 176, 211	0
1	I	320/325 (98%)	0.22	16 (5%)	28	16	82, 139, 192, 235	0
1	K	320/325 (98%)	0.39	33 (10%)	6	4	84, 150, 204, 253	0
2	B	172/177 (97%)	0.23	7 (4%)	37	24	61, 102, 144, 165	0
2	D	172/177 (97%)	0.26	9 (5%)	27	15	58, 100, 134, 164	0
2	F	172/177 (97%)	0.38	12 (6%)	16	9	62, 106, 135, 154	0
2	H	172/177 (97%)	0.33	8 (4%)	31	19	63, 103, 127, 145	0
2	J	172/177 (97%)	0.62	18 (10%)	6	3	72, 117, 152, 180	0
2	L	172/177 (97%)	0.49	15 (8%)	10	5	71, 113, 154, 175	0
All	All	2952/3012 (98%)	0.29	157 (5%)	26	14	58, 116, 179, 253	0

The worst 5 of 157 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	J	34	GLN	5.5
1	K	182	GLN	5.0
1	K	134	GLY	4.9
1	C	32	THR	4.8
1	E	212	ALA	4.4

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

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

WORLDWIDE
wwPDB
PROTEIN DATA BANK

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
3	NAG	P	1	15/15	0.83	0.18	123,157,169,182	0
3	NAG	M	1	15/15	0.87	0.25	133,176,188,205	0
4	GAL	O	1	12/12	0.87	0.28	117,146,156,158	0
4	GAL	N	1	12/12	0.89	0.18	100,127,139,149	0
3	GAL	P	2	11/12	0.90	0.13	135,159,172,177	0
3	GAL	M	2	11/12	0.92	0.12	142,155,165,165	0
4	SIA	N	2	20/21	0.93	0.21	63,92,148,161	0
3	SIA	P	3	20/21	0.93	0.18	109,127,171,186	0
4	SIA	O	2	20/21	0.94	0.20	120,137,153,154	0
3	SIA	M	3	20/21	0.95	0.17	81,107,135,138	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.
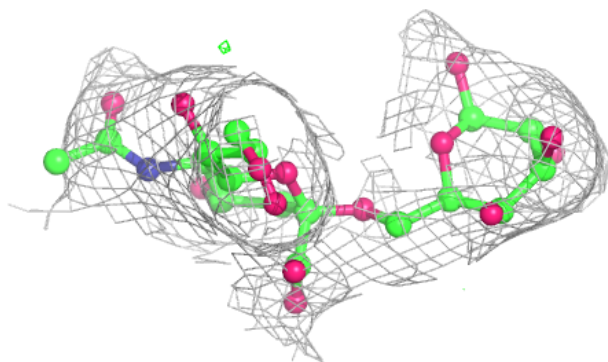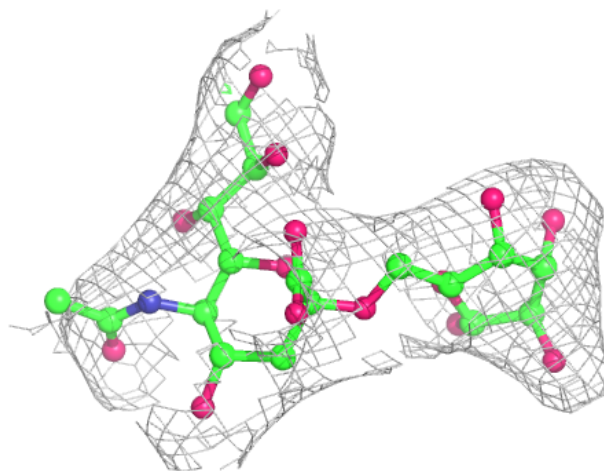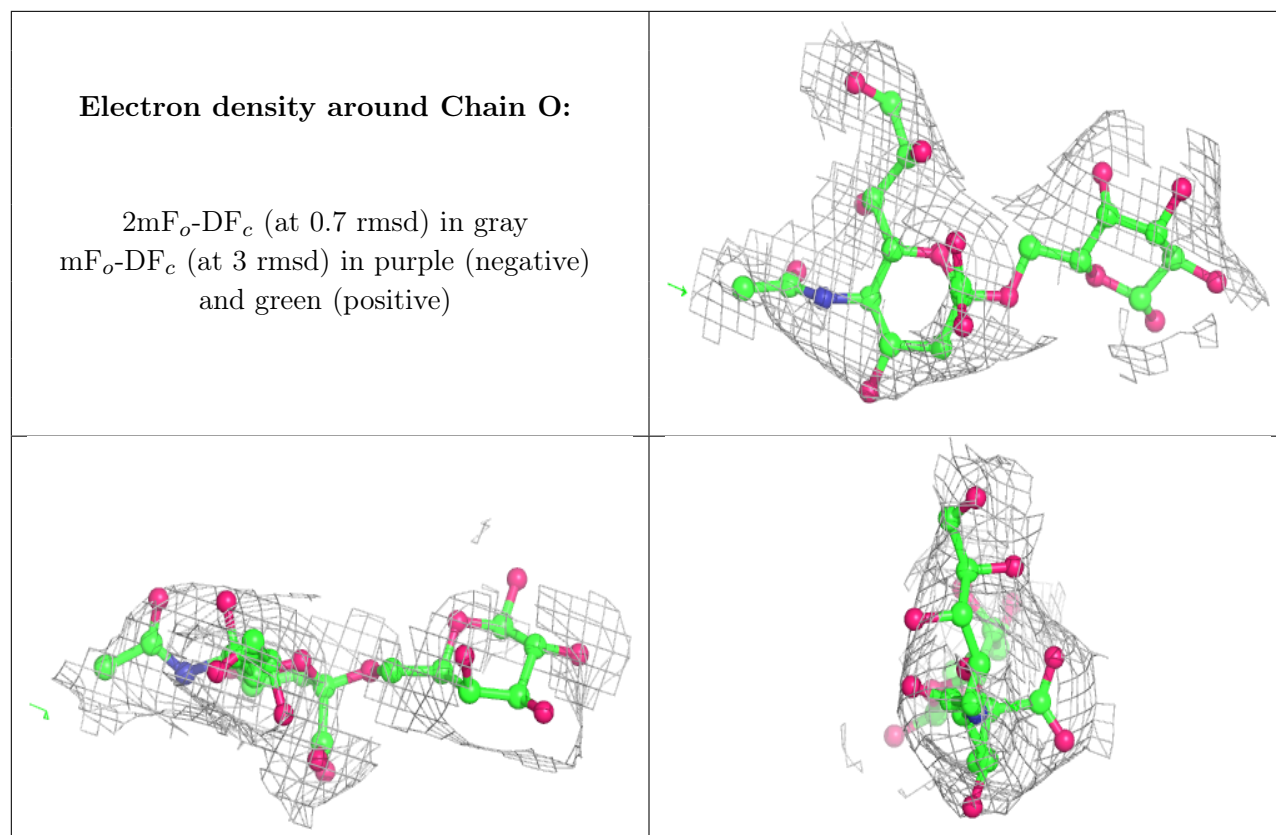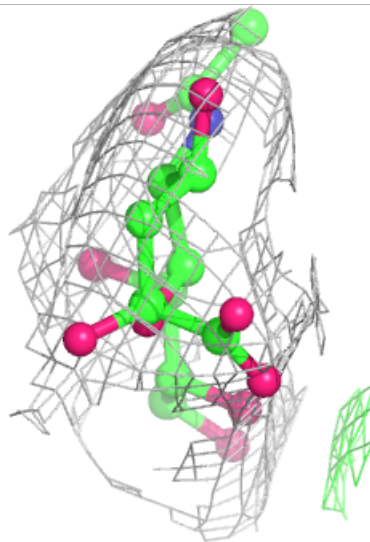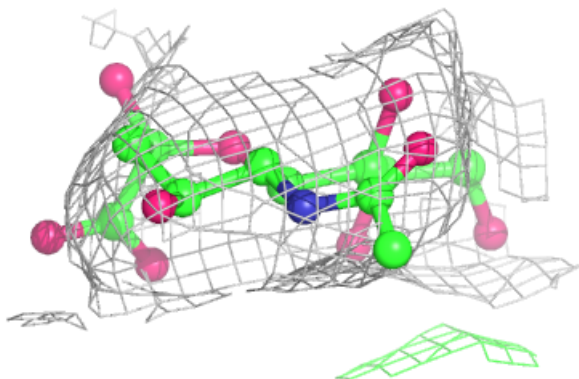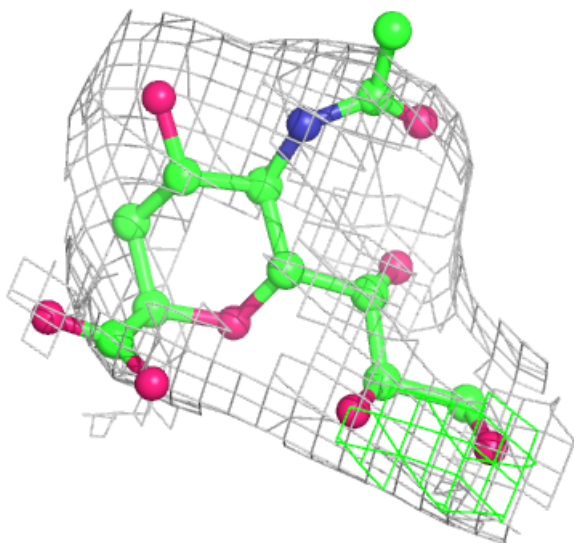

Electron density around Chain M:

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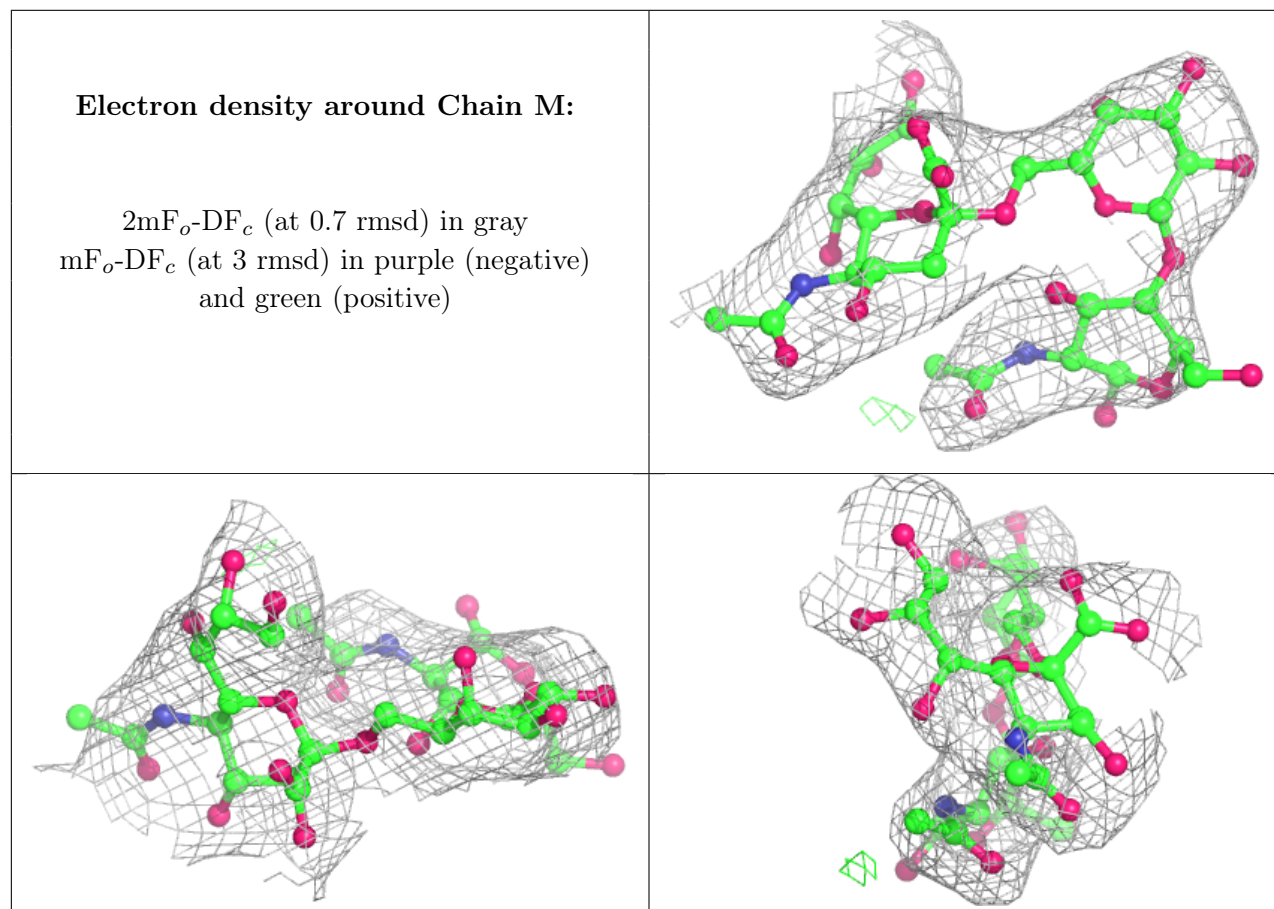
Electron density around Chain P:

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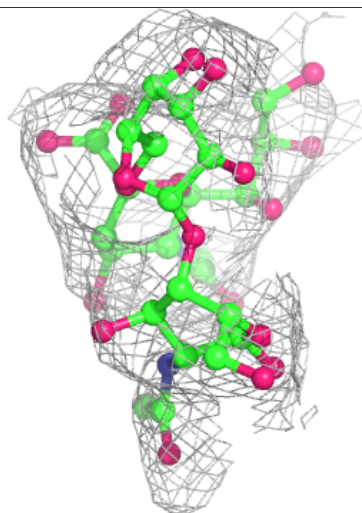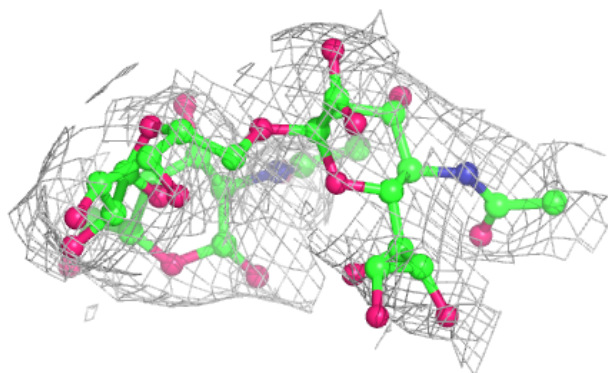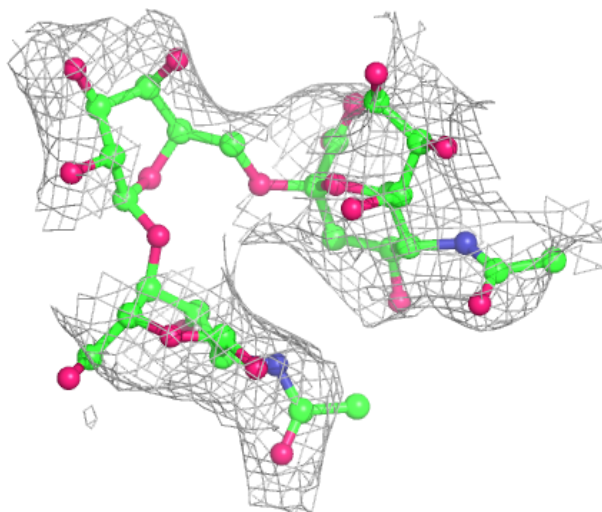
Electron density around Chain N:

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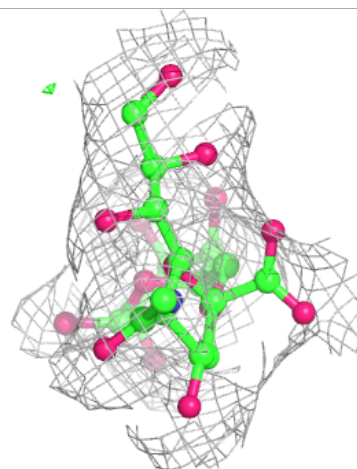
Electron density around Chain O:

$2mF_o\text{-}DF_c$ (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	D	201	14/15	0.72	0.23	124,155,166,167	0
5	NAG	I	401	14/15	0.73	0.18	123,184,208,208	0
5	NAG	J	201	14/15	0.80	0.14	92,124,130,131	0
5	NAG	F	201	14/15	0.82	0.18	82,100,142,145	0
5	NAG	K	401	14/15	0.83	0.17	131,162,179,182	0
5	NAG	L	201	14/15	0.83	0.14	100,136,151,155	0
5	NAG	H	201	14/15	0.84	0.15	88,108,126,127	0
6	SIA	I	402	21/21	0.84	0.20	124,153,186,205	0
6	SIA	K	402	21/21	0.84	0.34	112,151,180,211	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.
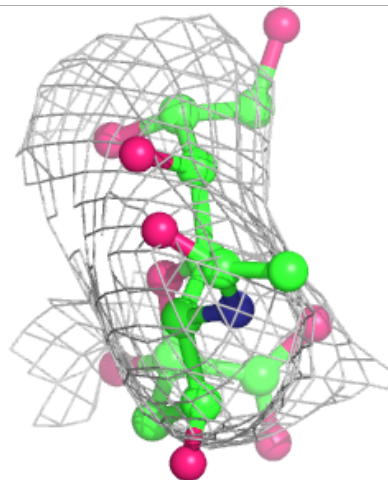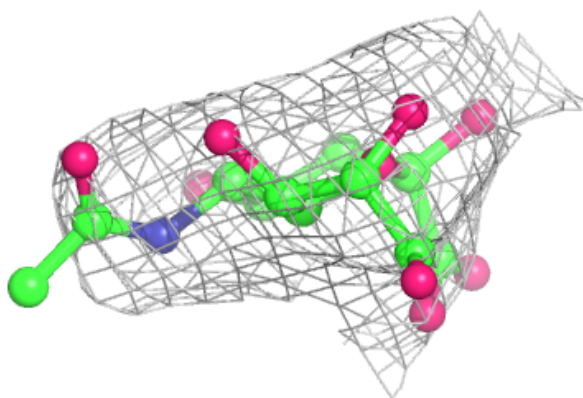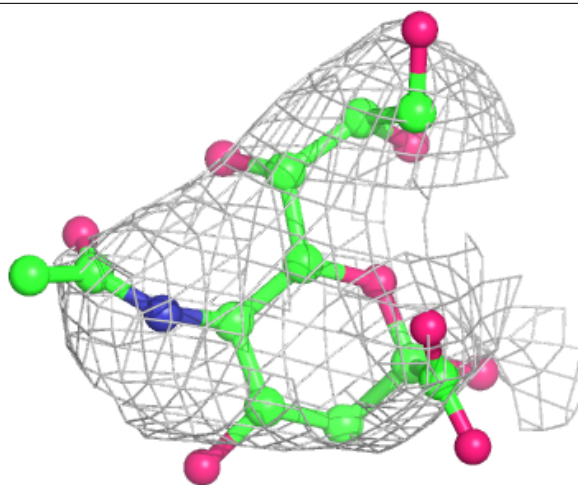
Electron density around SIA I 402:

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

Electron density around SIA K 402:

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

6.5 Other polymers ⓘ

There are no such residues in this entry.