W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Oct 14, 2023 – 11:55 PM EDT

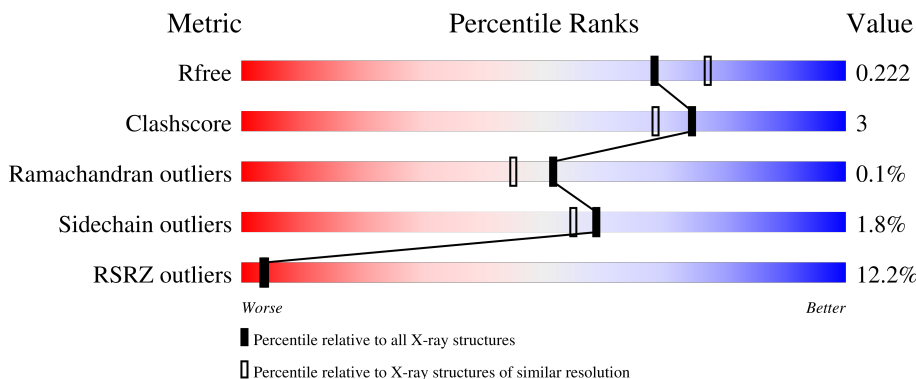|		
PDB ID	:	6MYA
Title	:	Crystal structure of InvbP.18715.a.KN11: Influenza hemagglutinin from strain A/Almaty/32/1998
Authors	:	Seattle Structural Genomics Center for Infectious Disease (SSGCID)
Deposited on	:	2018-11-01
Resolution	:	2.05 Å(reported)

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 2.05 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		0.222
Clashscore		3
Ramachandran outliers		0.1%
Sidechain outliers		1.8%
RSRZ outliers		12.2%

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	1692 (2.04-2.04)
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)
RSRZ outliers	127900	1672 (2.04-2.04)

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	499	10% / 89% / 8% ··
1	B	499	10% / 86% / 10% ·
1	C	499	4% / 88% / 9% ·
1	D	499	19% / 87% / 8% ··
1	E	499	14% / 91% / 7% ··

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	499	15% ▬▬▬ 88% ▬▬▬▬▬▬▬▬▬▬▬ 8% ▬
2	G	2	50% ▬▬▬▬▬ 50% ▬▬▬▬▬
2	H	2	100% ▬▬▬▬▬▬▬▬▬▬
2	I	2	100% ▬▬▬▬▬▬▬▬▬▬

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
2	NAG	G	2	-	-	-	X
2	NAG	H	2	-	-	-	X
2	NAG	I	1	-	-	-	X
2	NAG	I	2	-	-	-	X
3	NAG	A	502	-	-	-	X
3	NAG	A	503	-	-	-	X
3	NAG	A	504	-	-	-	X
3	NAG	B	501	-	-	-	X
3	NAG	B	504	-	-	-	X
3	NAG	C	501	-	-	-	X
3	NAG	C	503	-	-	-	X
3	NAG	F	501	-	-	-	X

2   Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 24594 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	488	Total 3784	C 2384	N 651	O 730	S 19	0	8	0
1	B	482	Total 3715	C 2345	N 640	O 711	S 19	0	4	0
1	C	486	Total 3802	C 2395	N 647	O 741	S 19	0	9	0
1	D	478	Total 3650	C 2300	N 626	O 705	S 19	0	7	0
1	E	489	Total 3769	C 2377	N 647	O 725	S 20	0	7	0
1	F	484	Total 3698	C 2328	N 637	O 713	S 20	0	5	0

- Molecule 2 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
2	G	2	Total 28	C 16	N 2	O 10	0	0	0
2	H	2	Total 28	C 16	N 2	O 10	0	0	0
2	I	2	Total 28	C 16	N 2	O 10	0	0	0

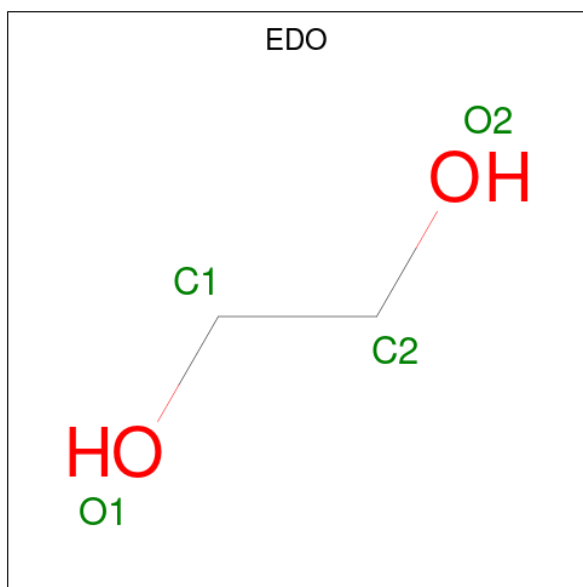- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

NAG

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

Continued on next page...

Continued from previous page...

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

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 4	C 2	O 2	0	0
4	A	1	Total 4	C 2	O 2	0	0
4	A	1	Total 4	C 2	O 2	0	0
4	C	1	Total 4	C 2	O 2	0	0
4	C	1	Total 4	C 2	O 2	0	0
4	C	1	Total 4	C 2	O 2	0	0
4	C	1	Total 4	C 2	O 2	0	0
4	C	1	Total 4	C 2	O 2	0	0
4	C	1	Total 4	C 2	O 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	C	1	Total 4	C 2	O 2	0	0
4	C	1	Total 4	C 2	O 2	0	0
4	C	1	Total 4	C 2	O 2	0	0
4	D	1	Total 4	C 2	O 2	0	0
4	D	1	Total 4	C 2	O 2	0	0
4	D	1	Total 4	C 2	O 2	0	0
4	D	1	Total 4	C 2	O 2	0	0
4	E	1	Total 4	C 2	O 2	0	0
4	E	1	Total 4	C 2	O 2	0	0
4	E	1	Total 4	C 2	O 2	0	0
4	E	1	Total 4	C 2	O 2	0	0
4	F	1	Total 4	C 2	O 2	0	0
4	F	1	Total 4	C 2	O 2	0	0

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).

PEG

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	E	1	Total 7	C 4	O 3	0	0

- Molecule 6 is water.

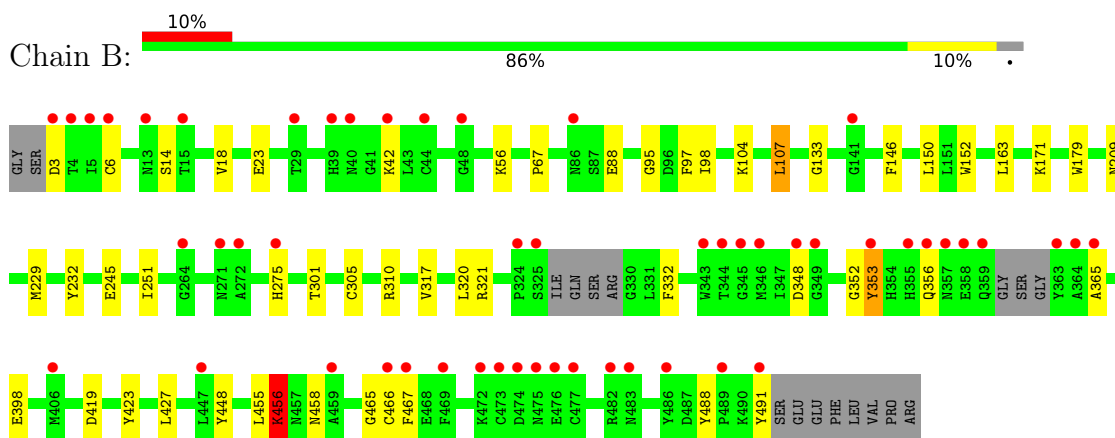| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	261	Total 264	O 264	0	3
6	B	295	Total 299	O 299	0	4
6	C	320	Total 323	O 323	0	3
6	D	309	Total 312	O 312	0	3
6	E	283	Total 284	O 284	0	1
6	F	289	Total 291	O 291	0	2

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.
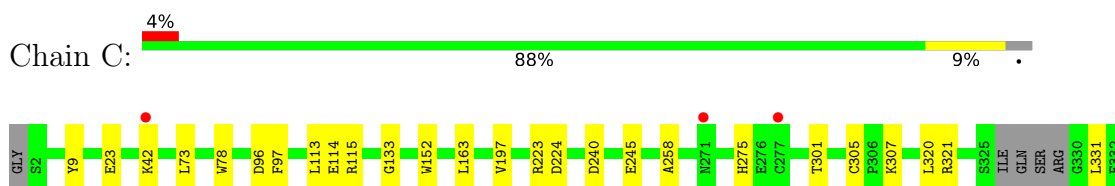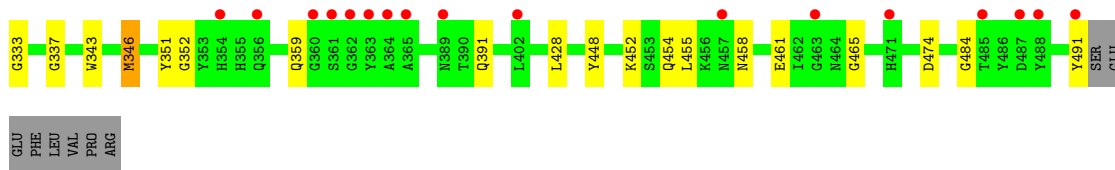
- Molecule 1: Hemagglutinin
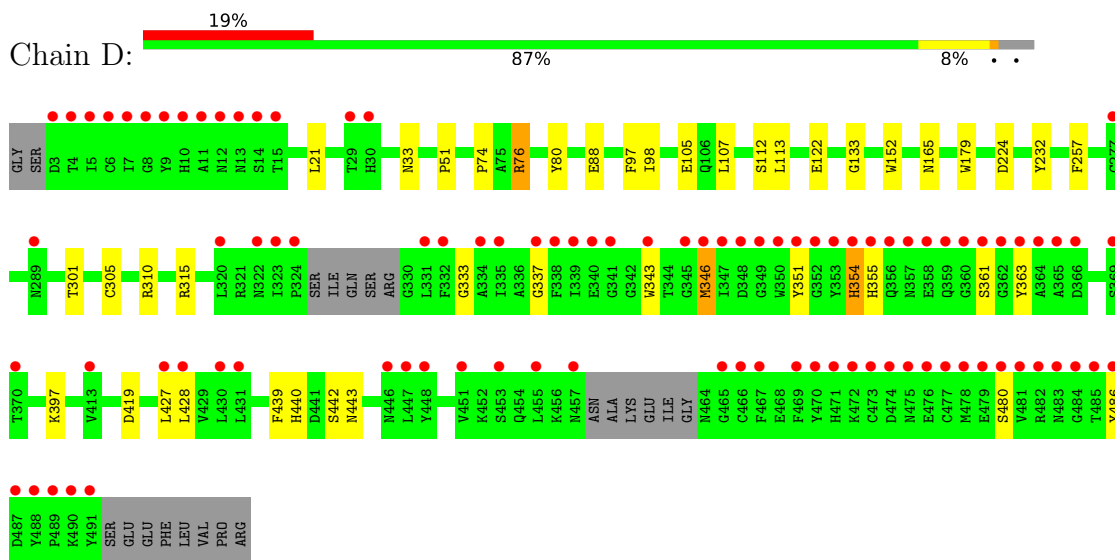

- Molecule 1: Hemagglutinin

- Molecule 1: Hemagglutinin

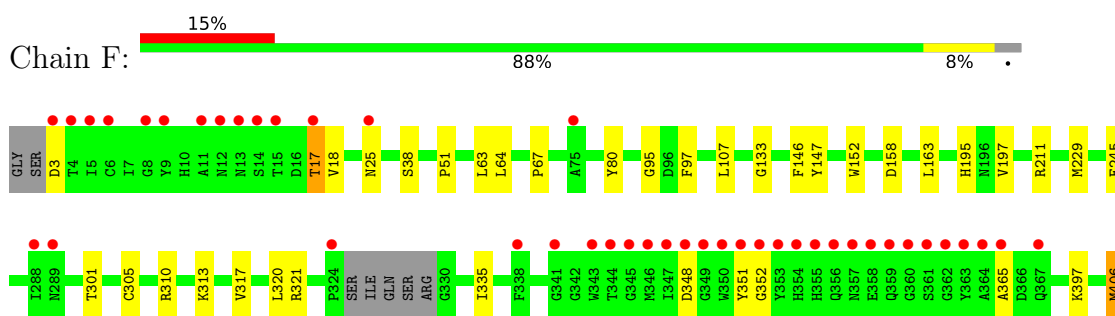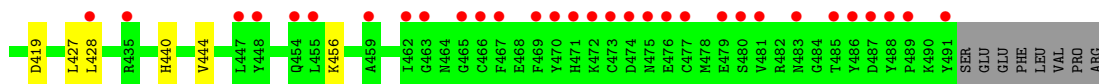G333 G337 W343 M346 Y351 G352 Y353 H354 H355 Q356 Q359 G360 S361 G362 Y363 A364 A365 N389 T390 Q391 L402 L428 Y448 K452 S453 Q454 L455 K456 N457 N458 E461 I462 G463 N464 G465 H471 D474 G484 T485 Y486 D487 Y488 Y491 SER GLU

GLU PHE LEU VAL PRO ARG

- **Molecule 1: Hemagglutinin**

Chain D:

GLY SER D3 T4 I5 C6 I7 G8 Y9 H10 A11 N12 N13 S14 T15 L21 T29 H30 N33 P51 P74 A75 R76 Y80 E88 F97 I98 E105 Q106 L107 S112 L113 E122 G133 W152 N165 W179 D224 Y232 F257 C277

N289 T301 C305 R310 R315 L320 R321 N322 I323 P324 SER ILE GLN SER ARG G330 L331 F332 G333 A334 I335 A336 G337 F338 I339 E340 G341 G342 W343 T344 G345 M346 I347 D348 G349 W350 Y351 G352 Y353 H354 H355 N357 E358 G359 G360 S361 G362 Y363 A364 A365 D366 S369

T370 K397 V413 D419 L427 L428 V429 L430 L431 F439 H440 D441 S442 N443 N446 L447 Y448 V451 K452 S453 Q454 L455 K456 N457 ASN ALA LYS GLU ILE GLY N464 G465 C466 F467 E468 F469 Y470 H471 K472 C473 D474 N475 E476 C477 M478 E479 S480 V481 R482 G484 T485 Y486

D487 Y488 P489 K490 Y491 SER GLU GLU PHE LEU VAL PRO ARG

- **Molecule 1: Hemagglutinin**

Chain E:

GLY S2 D3 T4 I5 C6 I7 S14 T15 H30 L35 G41 N86 G95 D96 F97 I98 K104 E105 Q106 L107 G133 G142 W152 G157 W179 D224 M229 Y232 F263 N271 A272 C277 N278 T279 T301 C305 R310

L314 L320 I323 P324 SER ILE GLN SER ARG G330 A334 F338 W343 T344 G345 M346 I347 D348 Y351 G352 Y353 H354 H355 Q356 G362 Y363 A364 A365 G376 Q391 M406 D419 L427 L428 Y448 V451 K452 S453 N457 N458 A459 T462 G463 N464

G465 C466 F467 E468 F469 Y470 H471 K472 C473 D474 N475 E476 C477 M478 E479 S480 V481 R482 N483 G484 T485 Y486 D487 Y488 P489 K490 Y491 E494 F495 LEU VAL PRO ARG

- **Molecule 1: Hemagglutinin**

Chain F:

GLY SER D3 T4 I5 C6 I7 G8 Y9 H10 A11 N12 N13 S14 T15 D16 T17 V18 N25 S38 P51 L63 L64 P67 A75 Y80 G95 D96 F97 L107 G133 F146 Y147 W152 D158 L163 H195 N196 V197 R211 M229 E245

I288 N289 T301 C305 R310 K313 V317 L320 R321 P324 SER ILE GLN SER ARG G330 I335 F338 G341 G342 W343 T344 G345 M346 I347 D348 G349 W350 Y351 G352 Y353 H354 H355 Q356 N357 E358 Q359 G360 S361 G362 Y363 A364 A365 D366 Q367 K397 M406

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

Chain G:

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

Chain H:

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

Chain I:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	69.65Å 244.37Å 115.02Å 90.00° 90.09° 90.00°	Depositor
Resolution (Å)	45.93 – 2.05 48.11 – 2.05	Depositor EDS
% Data completeness (in resolution range)	96.5 (45.93-2.05) 96.5 (48.11-2.05)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.28 (at 2.05Å)	Xtriage
Refinement program	PHENIX dev_3283	Depositor
R, R_{free}	0.188 , 0.222 0.189 , 0.222	Depositor DCC
R_{free} test set	2003 reflections (0.87%)	wwPDB-VP
Wilson B-factor (Å^2)	30.7	Xtriage
Anisotropy	0.177	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 56.6	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.033 for h,-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	24594	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.99% 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.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

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	0.38	0/3899	0.55	0/5310
1	B	0.40	0/3817	0.59	1/5195 (0.0%)
1	C	0.41	0/3920	0.58	0/5332
1	D	0.43	0/3758	0.58	0/5118
1	E	0.40	0/3885	0.56	1/5294 (0.0%)
1	F	0.42	0/3804	0.58	0/5186
All	All	0.41	0/23083	0.57	2/31435 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	107	LEU	CA-CB-CG	-5.06	103.66	115.30
1	E	107	LEU	CA-CB-CG	-5.01	103.78	115.30

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	3784	0	3511	22	0
1	B	3715	0	3462	30	0
1	C	3802	0	3580	27	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	3650	0	3355	22	0
1	E	3769	0	3480	18	0
1	F	3698	0	3399	26	0
2	G	28	0	25	0	0
2	H	28	0	25	0	0
2	I	28	0	25	0	0
3	A	56	0	52	0	0
3	B	42	0	39	0	0
3	C	56	0	52	0	0
3	D	14	0	13	0	0
3	E	28	0	26	0	0
3	F	28	0	26	0	0
4	A	12	0	18	1	0
4	C	36	0	54	4	0
4	D	16	0	24	0	0
4	E	16	0	24	2	0
4	F	8	0	12	1	0
5	E	7	0	10	0	0
6	A	264	0	0	2	0
6	B	299	0	0	1	0
6	C	323	0	0	1	0
6	D	312	0	0	1	0
6	E	284	0	0	1	0
6	F	291	0	0	2	0
All	All	24594	0	21212	139	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 3.

The worst 5 of 139 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:23:GLU:OE1	1:C:321:ARG:NH2	2.13	0.77
1:A:399:PHE:CE1	1:A:406:MET:HG2	2.23	0.74
1:E:98[A]:ILE:HG13	1:E:232:TYR:CE2	2.31	0.65
1:F:335:ILE:HG13	1:F:444:VAL:HG21	1.79	0.64
1:F:351:TYR:OH	1:F:440:HIS:ND1	2.25	0.64

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	490/499 (98%)	480 (98%)	10 (2%)	0	100	100
1	B	480/499 (96%)	468 (98%)	11 (2%)	1 (0%)	47	39
1	C	491/499 (98%)	482 (98%)	9 (2%)	0	100	100
1	D	479/499 (96%)	468 (98%)	11 (2%)	0	100	100
1	E	492/499 (99%)	479 (97%)	11 (2%)	2 (0%)	34	24
1	F	485/499 (97%)	473 (98%)	11 (2%)	1 (0%)	47	39
All	All	2917/2994 (97%)	2850 (98%)	63 (2%)	4 (0%)	51	45

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	456	LYS
1	F	456	LYS
1	E	86	ASN
1	E	157	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	402/438 (92%)	392 (98%)	10 (2%)	47	40
1	B	392/438 (90%)	387 (99%)	5 (1%)	69	67
1	C	414/438 (94%)	409 (99%)	5 (1%)	71	70

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	378/438 (86%)	367 (97%)	11 (3%)	42	35
1	E	398/438 (91%)	390 (98%)	8 (2%)	55	50
1	F	386/438 (88%)	378 (98%)	8 (2%)	53	48
All	All	2370/2628 (90%)	2323 (98%)	47 (2%)	59	50

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	346	MET
1	E	351	TYR
1	D	354	HIS
1	E	107	LEU
1	E	464	ASN

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

Mol	Chain	Res	Type
1	D	443	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 ⓘ

6 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
2	NAG	G	1	2,1	14,14,15	0.39	0	17,19,21	0.41	0
2	NAG	G	2	2	14,14,15	0.27	0	17,19,21	0.75	1 (5%)
2	NAG	H	1	2,1	14,14,15	0.31	0	17,19,21	0.57	0
2	NAG	H	2	2	14,14,15	0.75	0	17,19,21	0.70	0
2	NAG	I	1	2,1	14,14,15	0.29	0	17,19,21	0.57	0
2	NAG	I	2	2	14,14,15	0.34	0	17,19,21	0.81	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
2	NAG	G	1	2,1	-	0/6/23/26	0/1/1/1
2	NAG	G	2	2	-	1/6/23/26	0/1/1/1
2	NAG	H	1	2,1	-	0/6/23/26	0/1/1/1
2	NAG	H	2	2	-	3/6/23/26	0/1/1/1
2	NAG	I	1	2,1	-	2/6/23/26	0/1/1/1
2	NAG	I	2	2	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	G	2	NAG	C1-O5-C5	2.49	115.56	112.19

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

| Mol	Chain	Res	Type	Atoms
2	I	2	NAG	C1-C2-N2-C7
2	I	2	NAG	O5-C5-C6-O6
2	G	2	NAG	O5-C5-C6-O6
2	H	2	NAG	C4-C5-C6-O6
2	H	2	NAG	O5-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

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

Oligosaccharide Chain G

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain H

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain I

Bond lengths

Bond angles

Torsions

Rings
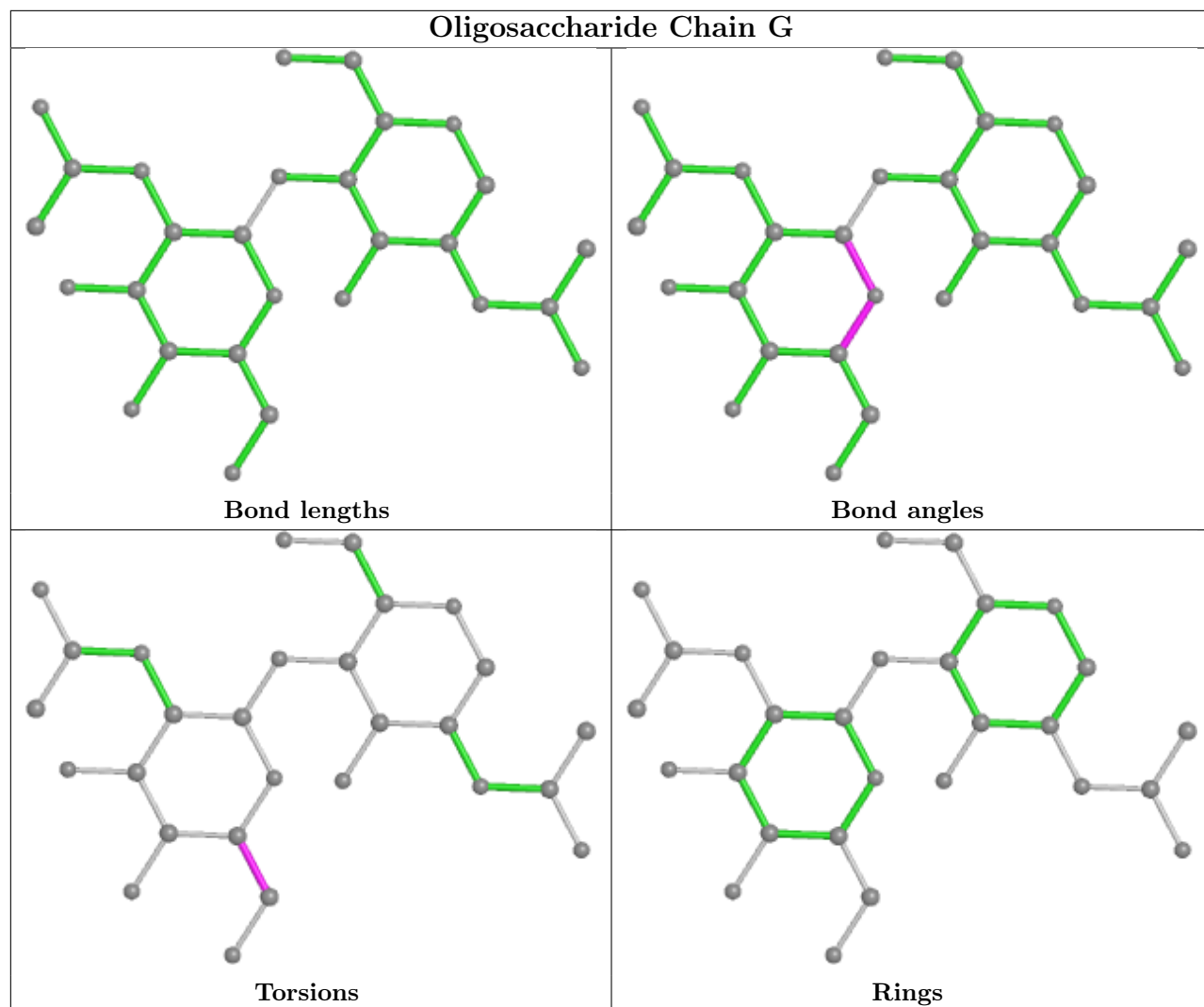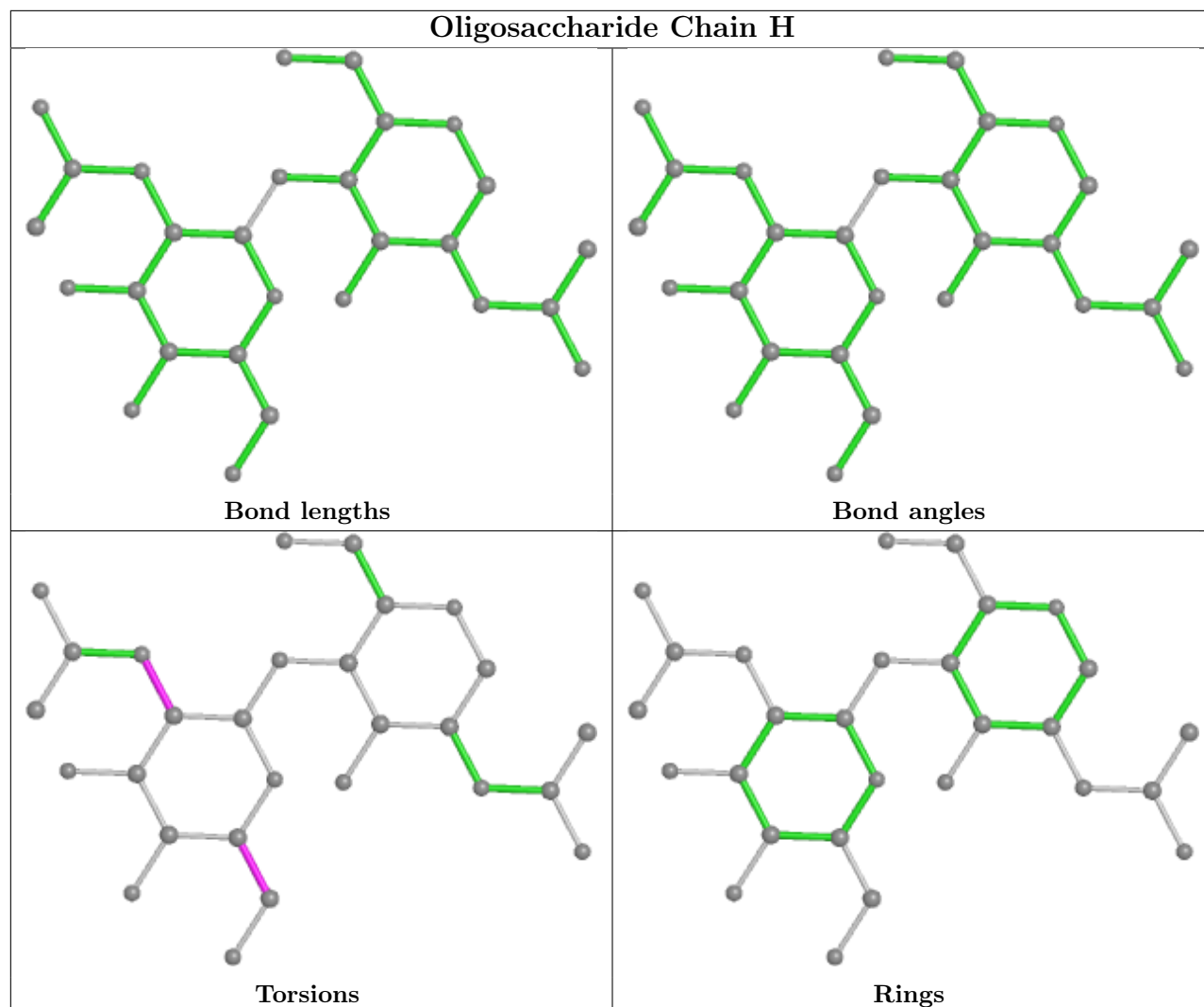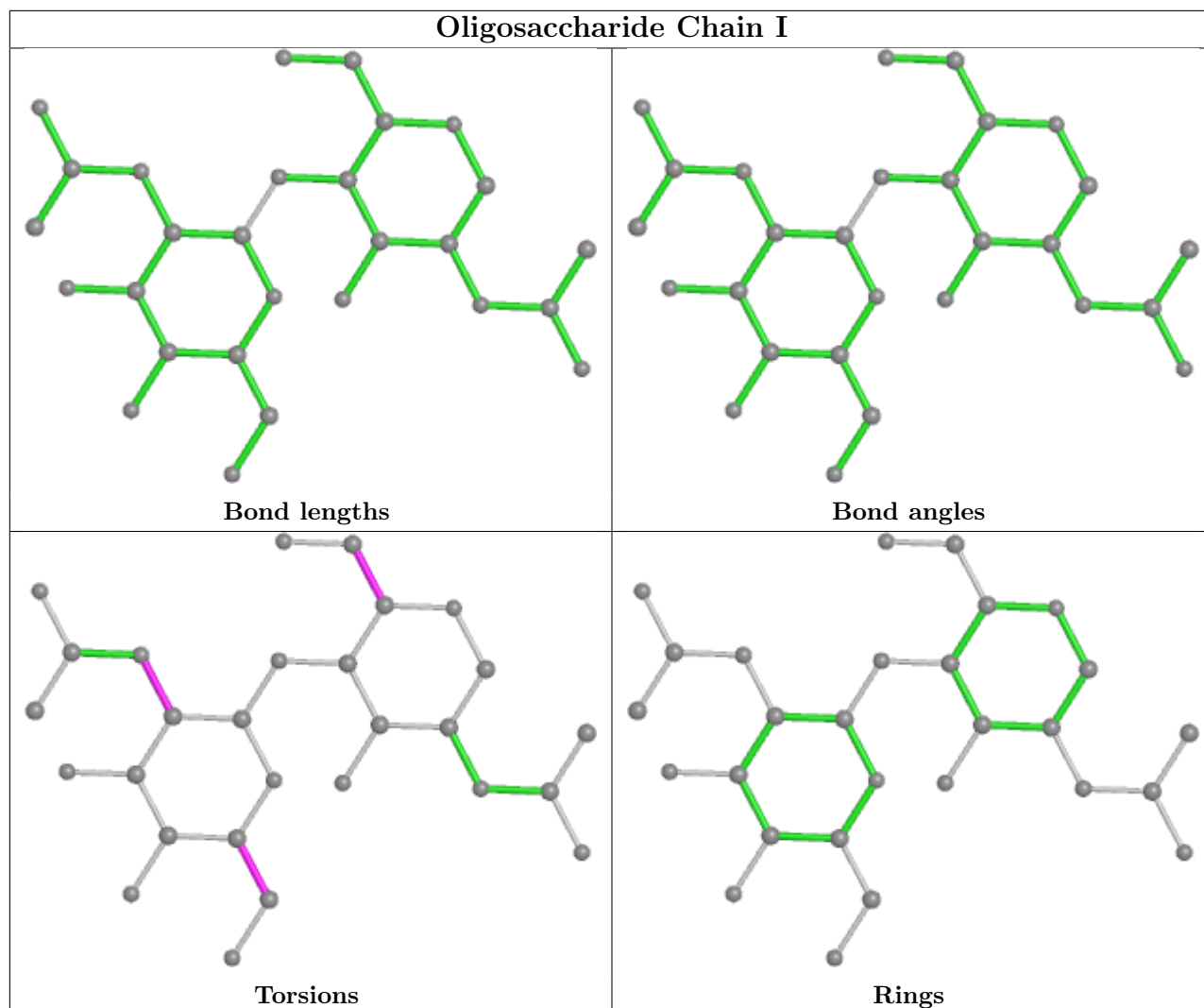
5.6 Ligand geometry ⓘ

39 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	$\#|Z| > 2$
4	EDO	D	505	-	3,3,3	0.40	0	2,2,2	0.50	0
4	EDO	C	506	-	3,3,3	0.54	0	2,2,2	0.39	0
3	NAG	B	501	1	14,14,15	0.65	0	17,19,21	0.61	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
4	EDO	E	508	-	3,3,3	0.44	0	2,2,2	0.22	0
4	EDO	D	506	-	3,3,3	0.44	0	2,2,2	0.41	0
4	EDO	C	511	-	3,3,3	0.55	0	2,2,2	0.23	0
3	NAG	A	504	1	14,14,15	0.46	0	17,19,21	0.49	0
4	EDO	F	503	-	3,3,3	0.49	0	2,2,2	0.23	0
3	NAG	E	503	1	14,14,15	0.55	0	17,19,21	0.71	1 (5%)
4	EDO	C	508	-	3,3,3	0.44	0	2,2,2	0.32	0
3	NAG	C	501	1	14,14,15	0.33	0	17,19,21	0.55	0
4	EDO	C	505	-	3,3,3	0.82	0	2,2,2	0.57	0
3	NAG	B	505	1	14,14,15	0.34	0	17,19,21	0.53	0
4	EDO	E	505	-	3,3,3	0.50	0	2,2,2	0.23	0
3	NAG	E	504	1	14,14,15	0.32	0	17,19,21	0.47	0
4	EDO	D	507	-	3,3,3	0.63	0	2,2,2	0.07	0
3	NAG	A	501	1	14,14,15	0.53	0	17,19,21	0.63	0
3	NAG	D	501	1	14,14,15	0.71	0	17,19,21	0.61	1 (5%)
4	EDO	F	504	-	3,3,3	0.35	0	2,2,2	0.67	0
4	EDO	E	509	-	3,3,3	0.54	0	2,2,2	0.11	0
4	EDO	D	504	-	3,3,3	0.44	0	2,2,2	0.32	0
3	NAG	F	501	1	14,14,15	0.48	0	17,19,21	0.48	0
5	PEG	E	507	-	6,6,6	0.24	0	5,5,5	0.10	0
4	EDO	C	510	-	3,3,3	0.48	0	2,2,2	0.15	0
4	EDO	A	505	-	3,3,3	0.56	0	2,2,2	0.16	0
3	NAG	C	503	1	14,14,15	0.66	1 (7%)	17,19,21	0.57	0
4	EDO	A	507	-	3,3,3	0.45	0	2,2,2	0.42	0
3	NAG	A	503	1	14,14,15	0.42	0	17,19,21	0.46	0
4	EDO	E	506	-	3,3,3	0.40	0	2,2,2	0.37	0
4	EDO	C	507	-	3,3,3	0.35	0	2,2,2	0.60	0
3	NAG	A	502	1	14,14,15	0.46	0	17,19,21	0.55	0
3	NAG	F	502	1	14,14,15	0.47	0	17,19,21	0.49	0
3	NAG	C	504	1	14,14,15	0.44	0	17,19,21	0.60	0
4	EDO	C	513	-	3,3,3	0.54	0	2,2,2	0.25	0
4	EDO	A	506	-	3,3,3	0.48	0	2,2,2	0.25	0
3	NAG	C	502	1	14,14,15	0.37	0	17,19,21	0.40	0
4	EDO	C	509	-	3,3,3	0.51	0	2,2,2	0.20	0
4	EDO	C	512	-	3,3,3	0.50	0	2,2,2	0.23	0
3	NAG	B	504	1	14,14,15	0.46	0	17,19,21	0.47	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	EDO	D	505	-	-	0/1/1/1	-
4	EDO	C	506	-	-	0/1/1/1	-
3	NAG	B	501	1	-	2/6/23/26	0/1/1/1
4	EDO	E	508	-	-	1/1/1/1	-
4	EDO	D	506	-	-	1/1/1/1	-
4	EDO	C	511	-	-	1/1/1/1	-
3	NAG	A	504	1	-	3/6/23/26	0/1/1/1
4	EDO	F	503	-	-	1/1/1/1	-
3	NAG	E	503	1	-	0/6/23/26	0/1/1/1
4	EDO	C	508	-	-	1/1/1/1	-
3	NAG	C	501	1	-	0/6/23/26	0/1/1/1
4	EDO	C	505	-	-	1/1/1/1	-
3	NAG	B	505	1	-	1/6/23/26	0/1/1/1
4	EDO	E	505	-	-	1/1/1/1	-
3	NAG	E	504	1	-	1/6/23/26	0/1/1/1
4	EDO	D	507	-	-	1/1/1/1	-
3	NAG	A	501	1	-	3/6/23/26	0/1/1/1
3	NAG	D	501	1	-	0/6/23/26	0/1/1/1
4	EDO	F	504	-	-	0/1/1/1	-
4	EDO	E	509	-	-	0/1/1/1	-
4	EDO	D	504	-	-	0/1/1/1	-
3	NAG	F	501	1	-	4/6/23/26	0/1/1/1
5	PEG	E	507	-	-	1/4/4/4	-
4	EDO	C	510	-	-	0/1/1/1	-
4	EDO	A	505	-	-	1/1/1/1	-
3	NAG	C	503	1	-	0/6/23/26	0/1/1/1
4	EDO	A	507	-	-	0/1/1/1	-
3	NAG	A	503	1	-	2/6/23/26	0/1/1/1
4	EDO	E	506	-	-	0/1/1/1	-
4	EDO	C	507	-	-	0/1/1/1	-
3	NAG	A	502	1	-	0/6/23/26	0/1/1/1
3	NAG	F	502	1	-	2/6/23/26	0/1/1/1
3	NAG	C	504	1	-	2/6/23/26	0/1/1/1
4	EDO	C	513	-	-	0/1/1/1	-
4	EDO	A	506	-	-	1/1/1/1	-
3	NAG	C	502	1	-	2/6/23/26	0/1/1/1
4	EDO	C	509	-	-	1/1/1/1	-
4	EDO	C	512	-	-	1/1/1/1	-
3	NAG	B	504	1	-	1/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	503	NAG	O5-C1	2.16	1.47	1.43

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	E	503	NAG	C1-O5-C5	2.58	115.68	112.19
3	B	501	NAG	C1-O5-C5	2.11	115.05	112.19
3	D	501	NAG	C1-O5-C5	2.08	115.01	112.19

There are no chirality outliers.

5 of 36 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	501	NAG	C4-C5-C6-O6
3	A	503	NAG	O5-C5-C6-O6
3	C	502	NAG	O5-C5-C6-O6
3	A	503	NAG	C4-C5-C6-O6
3	C	502	NAG	C4-C5-C6-O6

There are no ring outliers.

7 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	508	EDO	2	0
4	C	511	EDO	1	0
4	F	503	EDO	1	0
4	C	505	EDO	1	0
4	C	510	EDO	1	0
4	C	513	EDO	1	0
4	A	506	EDO	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

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	488/499 (97%)	0.61	49 (10%)	7	7	20, 48, 91, 110	0
1	B	482/499 (96%)	0.60	52 (10%)	5	6	21, 42, 96, 124	0
1	C	486/499 (97%)	0.32	20 (4%)	37	40	20, 39, 70, 101	0
1	D	478/499 (95%)	1.05	93 (19%)	1	0	18, 42, 122, 154	0
1	E	489/499 (97%)	0.76	68 (13%)	2	2	19, 46, 96, 123	0
1	F	484/499 (96%)	0.80	74 (15%)	2	1	16, 43, 104, 130	0
All	All	2907/2994 (97%)	0.69	356 (12%)	4	3	16, 43, 99, 154	0

The worst 5 of 356 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	451	VAL	9.3
1	D	362	GLY	9.0
1	F	345	GLY	8.7
1	D	476	GLU	8.7
1	D	6	CYS	8.4

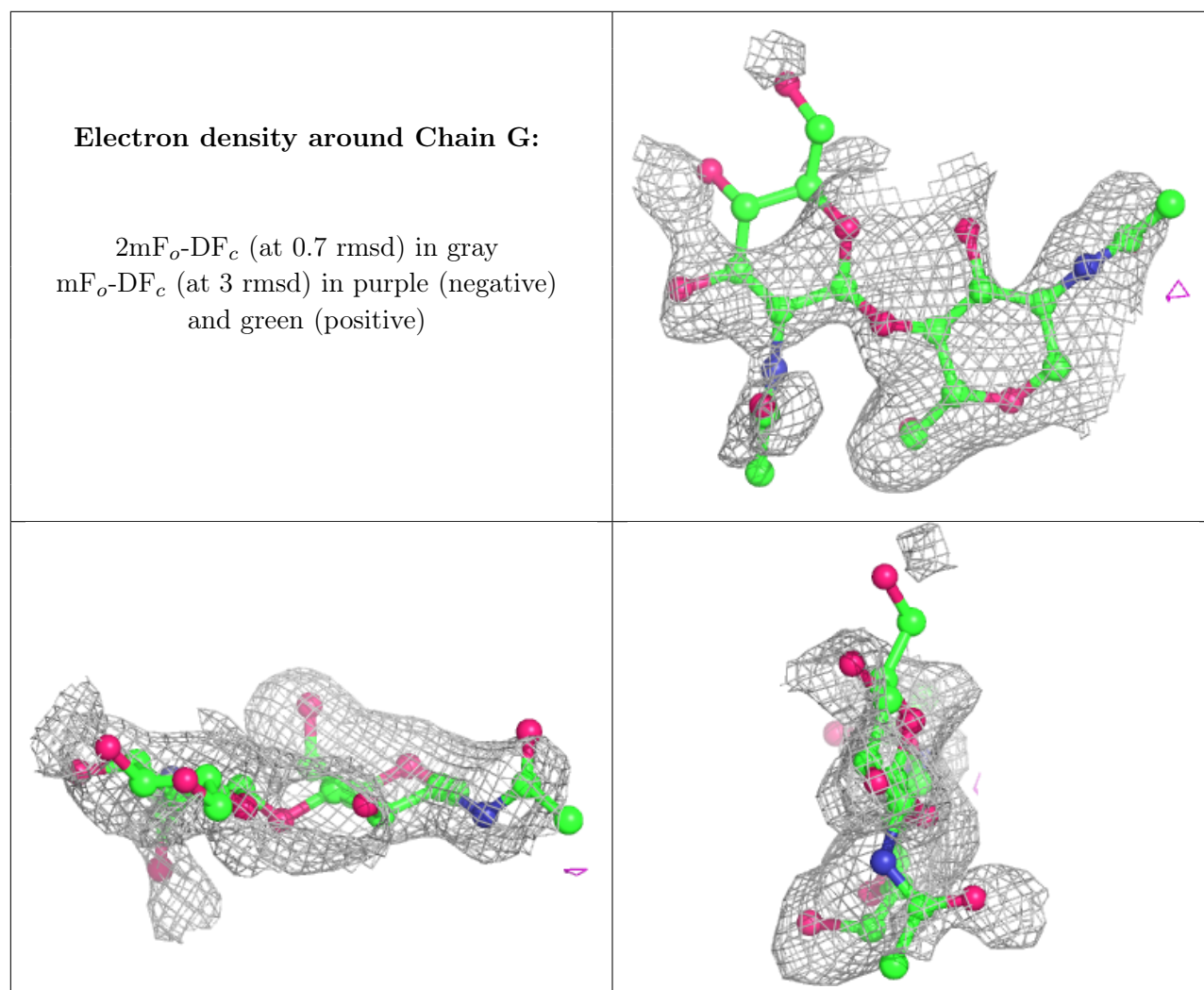## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

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.
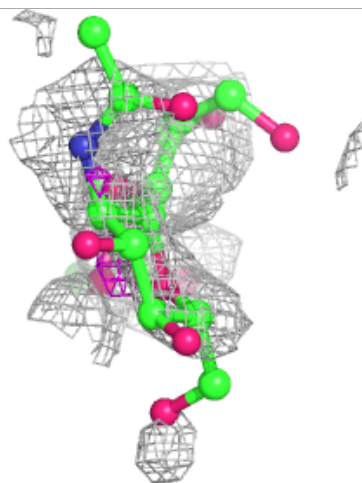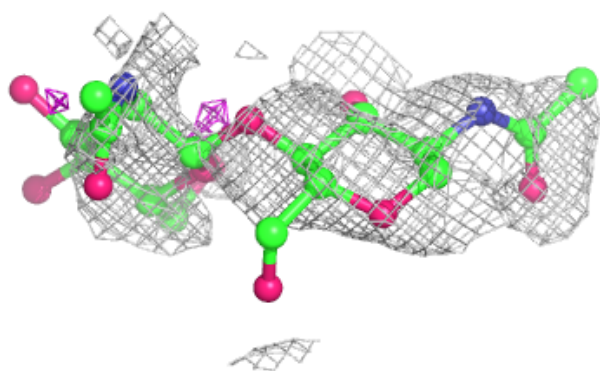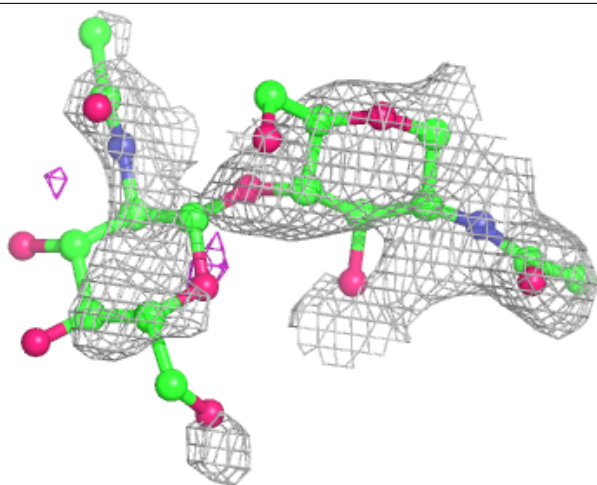
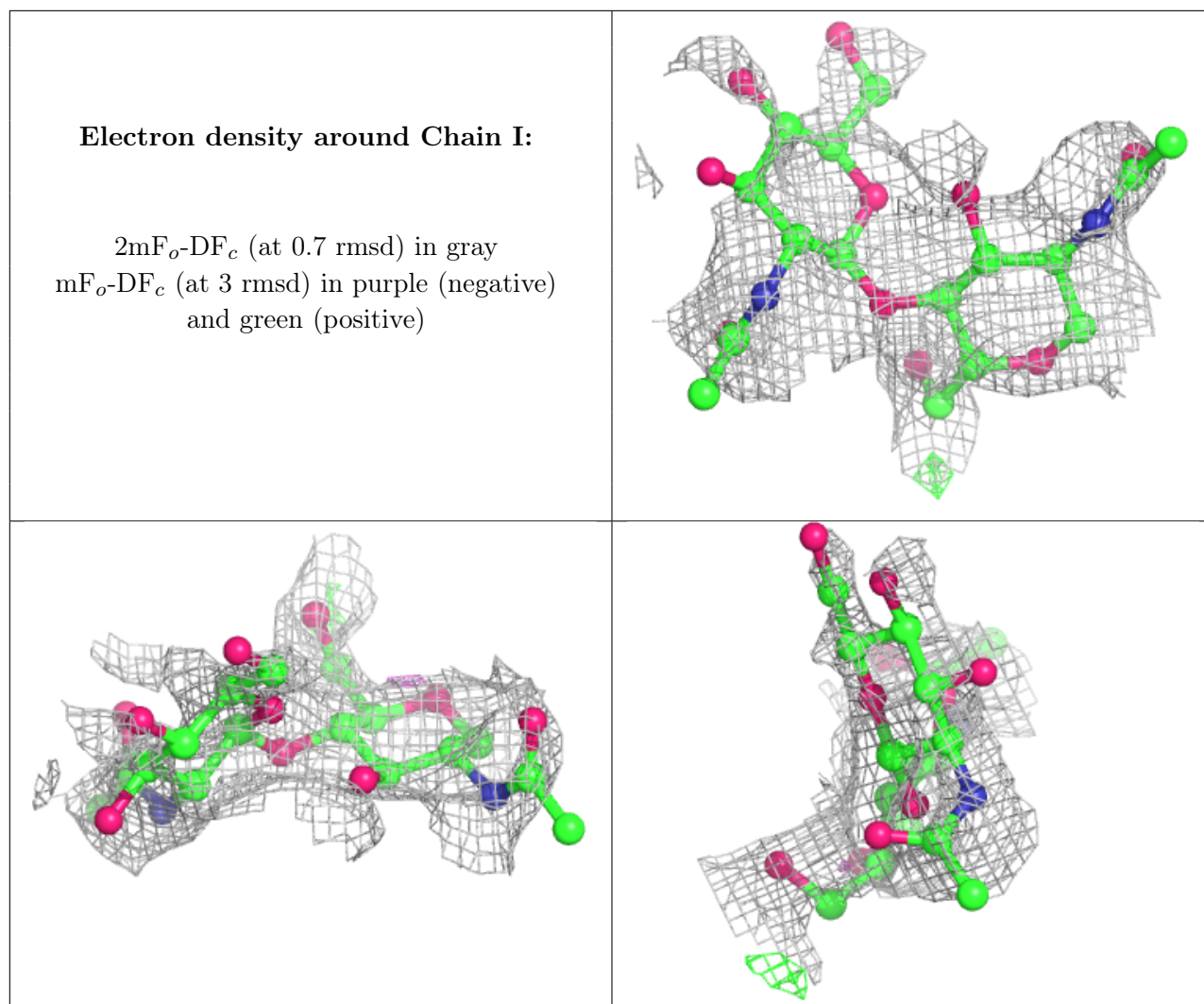| Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NAG	G	2	14/15	0.45	0.47	106,112,117,119	0
2	NAG	H	2	14/15	0.47	0.64	97,121,126,127	0
2	NAG	H	1	14/15	0.61	0.35	78,92,108,112	0
2	NAG	I	2	14/15	0.66	0.45	107,114,122,125	0
2	NAG	I	1	14/15	0.76	0.42	71,91,98,105	0
2	NAG	G	1	14/15	0.77	0.29	66,86,95,102	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:

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

Electron density around Chain H:

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

Electron density around Chain I:

$2mF_o\text{-}DF_c$ (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
3	NAG	A	503	14/15	0.41	0.53	95,111,120,122	0
3	NAG	A	504	14/15	0.57	0.55	76,100,106,111	0
3	NAG	C	503	14/15	0.58	0.50	77,86,92,94	0
3	NAG	D	501	14/15	0.59	0.29	87,102,109,112	0
4	EDO	D	507	4/4	0.61	0.29	46,54,57,58	0
3	NAG	A	502	14/15	0.65	0.54	87,102,107,109	0
3	NAG	F	501	14/15	0.66	0.48	82,100,109,109	0
3	NAG	B	501	14/15	0.68	0.41	84,100,106,110	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	EDO	A	505	4/4	0.69	0.20	60,62,62,64	0
3	NAG	F	502	14/15	0.70	0.35	85,100,111,112	0
3	NAG	E	503	14/15	0.70	0.30	78,91,101,101	0
3	NAG	C	501	14/15	0.70	0.49	79,99,104,107	0
4	EDO	E	509	4/4	0.70	0.27	63,67,69,70	0
5	PEG	E	507	7/7	0.72	0.16	51,58,63,69	0
3	NAG	B	504	14/15	0.73	0.47	82,95,104,104	0
4	EDO	C	506	4/4	0.74	0.25	44,46,52,55	0
4	EDO	A	506	4/4	0.74	0.20	71,74,75,75	0
3	NAG	E	504	14/15	0.78	0.39	77,99,104,105	0
3	NAG	B	505	14/15	0.80	0.41	74,87,96,97	0
4	EDO	C	509	4/4	0.81	0.15	52,52,54,55	0
3	NAG	C	502	14/15	0.81	0.30	69,88,99,99	0
4	EDO	C	512	4/4	0.82	0.17	65,67,67,70	0
4	EDO	A	507	4/4	0.83	0.38	65,65,68,69	0
3	NAG	A	501	14/15	0.84	0.23	65,81,96,101	0
4	EDO	C	505	4/4	0.84	0.26	33,43,46,50	0
3	NAG	C	504	14/15	0.85	0.32	67,78,91,92	0
4	EDO	E	505	4/4	0.85	0.20	47,48,52,60	0
4	EDO	D	505	4/4	0.86	0.22	61,66,69,71	0
4	EDO	F	503	4/4	0.87	0.13	64,70,70,72	0
4	EDO	C	513	4/4	0.87	0.17	54,56,59,61	0
4	EDO	C	508	4/4	0.90	0.25	64,66,69,74	0
4	EDO	C	510	4/4	0.91	0.18	47,59,64,66	0
4	EDO	D	504	4/4	0.91	0.16	50,54,57,67	0
4	EDO	D	506	4/4	0.92	0.12	57,59,62,62	0
4	EDO	C	511	4/4	0.92	0.30	38,46,57,60	0
4	EDO	F	504	4/4	0.93	0.12	46,48,56,57	0
4	EDO	E	506	4/4	0.93	0.26	60,61,63,63	0
4	EDO	C	507	4/4	0.94	0.12	40,45,47,52	0
4	EDO	E	508	4/4	0.97	0.17	37,39,42,43	0

6.5 Other polymers ⓘ

There are no such residues in this entry.