WORLDWIDE PDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Nov 13, 2023 – 03:48 PM JST

PDB ID	:	5XL3
Title	:	Complex structure of H4 hemagglutinin from avian influenza H4N6 virus with LSTa
Authors	:	Song, H.; Qi, J.; Gao, G.F.
Deposited on	:	2017-05-10
Resolution	:	2.20 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	:	2.36
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)	:	2.36

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 2.20 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		0.244
Clashscore		8
Ramachandran outliers		1.1%
Sidechain outliers		1.1%
RSRZ outliers		6.3%

Worse *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
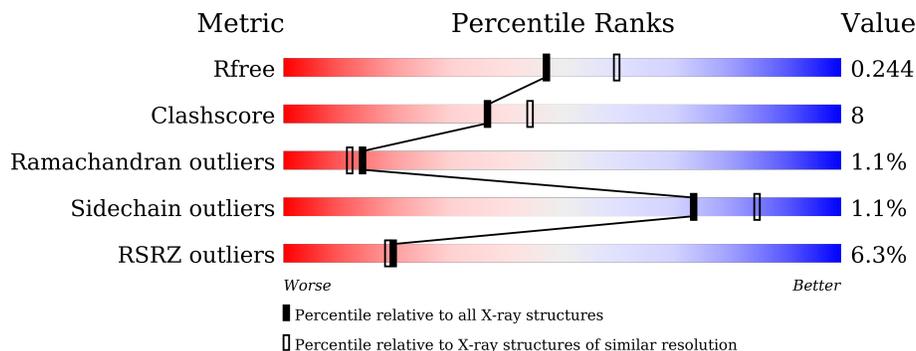
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	327	12% / 78% / 16% •••
1	B	327	6% / 88% / 9% ••
2	C	176	.% / 91% / 6% ••
2	D	176	.% / 90% / 7% •••
3	E	3	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
3	NAG	E	1	-	-	-	X

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 8217 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	319	Total 2457	C 1538	N 437	O 470	S 12	0	0	0
1	B	319	Total 2457	C 1538	N 437	O 470	S 12	0	0	0

- Molecule 2 is a protein called Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	172	Total 1404	C 871	N 250	O 279	S 4	0	0	0
2	D	172	Total 1404	C 871	N 250	O 279	S 4	0	0	0

- Molecule 3 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galacto pyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.
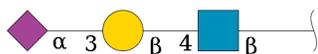
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	E	3	Total 46	C 25	N 2	O 19	0	0	0

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).
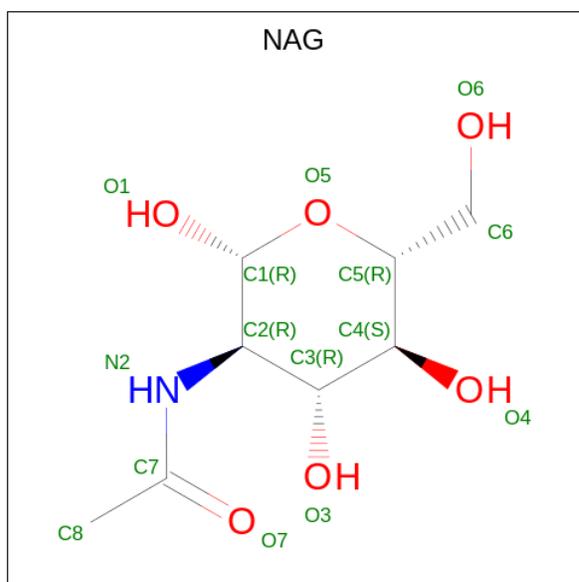
NAG

O6
OH

O5

O1
HO

O

C6

C1(R) C5(R)

C2(R) C4(S)
C3(R)

N2
HN

OH
O4

C7
OH
O3

C8 O
O7

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

- Molecule 5 is N-acetyl-alpha-neuraminic acid (three-letter code: SIA) (formula: $C_{11}H_{19}NO_9$).

SIA

O7
OH

O1B
HO

C9

C8(R)

C7(R)

O6   C1
O

O1A

HO
O9

C6(R)

C2(R)
OH O2

OH
O8 N5
HN

C5(R)
C4(S) C3

C10

OH
O4

C11 O
O10
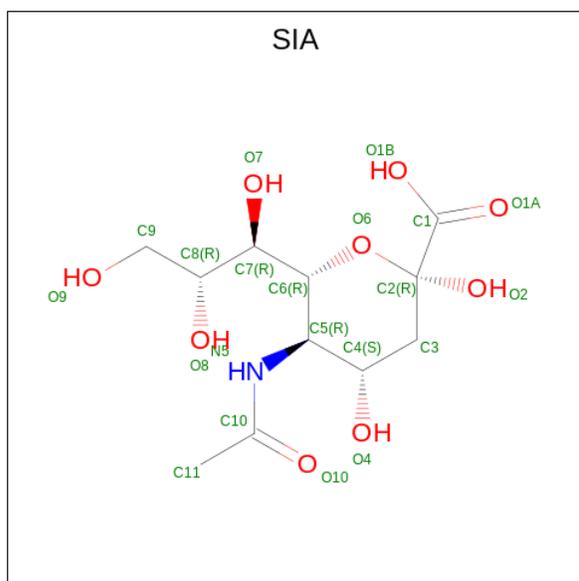
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total 21	C 11	N 1	O 9	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	58	Total 58	O 58	0	0
6	C	101	Total 101	O 101	0	0
6	B	103	Total 103	O 103	0	0
6	D	110	Total 110	O 110	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
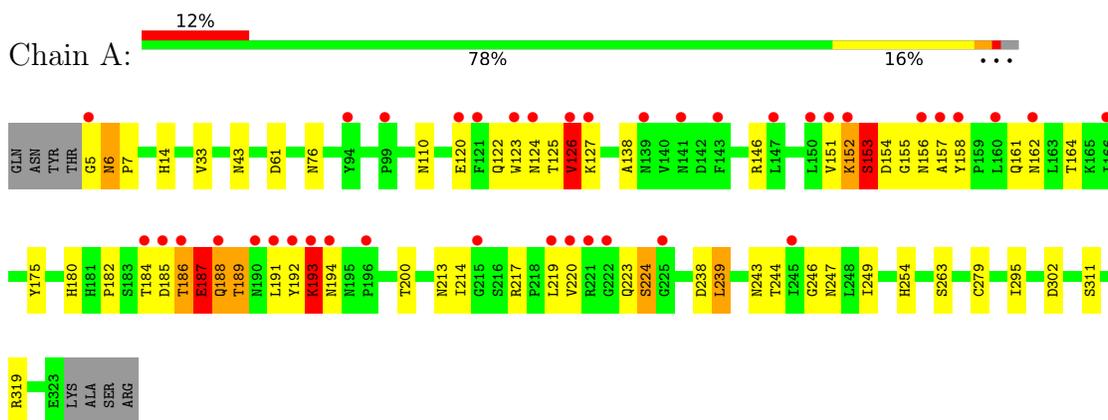
Chain A:

• Molecule 1: Hemagglutinin

Chain B:

• Molecule 2: Hemagglutinin

Chain C:

• Molecule 2: Hemagglutinin

Chain D:

G328 W348 E357 K366 K378 I383 E384 K385 T386 N387 E430 E447 R450 F468 D472 E477 Q499 ILE GLN GLY VAL
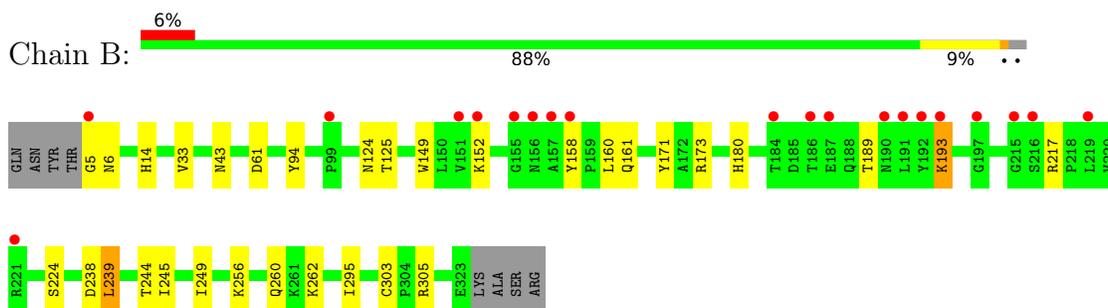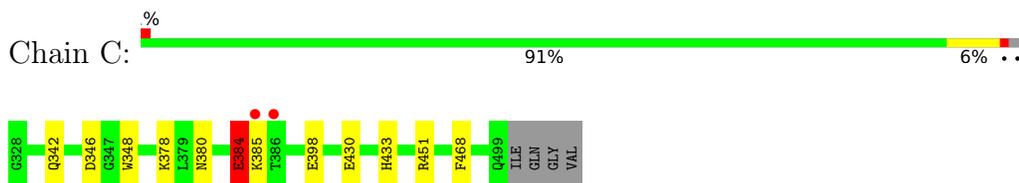
• Molecule 3: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose
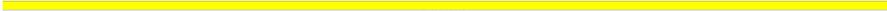
Chain E: 100%

NAG1 GAL2 SIA3

4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α, β, γ	100.53Å 100.53Å 685.08Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.09 – 2.20 49.09 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.7 (49.09-2.20) 91.9 (49.09-2.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.35 (at 2.20Å)	Xtriage
Refinement program	PHENIX 1.8.4_1496	Depositor
R, R_{free}	0.213 , 0.244 0.215 , 0.244	Depositor DCC
R_{free} test set	3468 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	35.7	Xtriage
Anisotropy	0.458	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 41.4	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	8217	wwPDB-VP
Average B, all atoms (Å^2)	54.0	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 49.07 % 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 7.8585e-05. 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: SIA, GAL, 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	0.36	0/2511	0.64	4/3418 (0.1%)
1	B	0.29	0/2511	0.48	0/3418
2	C	0.29	0/1428	0.51	1/1922 (0.1%)
2	D	0.31	0/1428	0.53	1/1922 (0.1%)
All	All	0.32	0/7878	0.55	6/10680 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	3
2	C	0	1
2	D	0	1
All	All	0	5

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	153	SER	CB-CA-C	-6.27	98.19	110.10
1	A	187	GLU	CA-CB-CG	6.09	126.80	113.40
1	A	187	GLU	N-CA-C	5.90	126.93	111.00
2	D	386	THR	N-CA-C	5.53	125.92	111.00
1	A	126	VAL	CB-CA-C	-5.41	101.12	111.40
2	C	384	GLU	CB-CA-C	-5.15	100.11	110.40

There are no chirality outliers.

All (5) planarity outliers are listed below:

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Chain	Res	Type	Group
1	A	123	TRP	Peptide
1	A	186	THR	Peptide
1	A	187	GLU	Peptide
2	C	384	GLU	Peptide
2	D	385	LYS	Peptide

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	2457	0	2402	70	0
1	B	2457	0	2402	27	0
2	C	1404	0	1320	11	0
2	D	1404	0	1320	19	0
3	E	46	0	40	2	0
4	A	28	0	26	1	0
4	B	28	0	26	3	0
5	A	21	0	18	2	0
6	A	58	0	0	5	0
6	B	103	0	0	6	0
6	C	101	0	0	3	0
6	D	110	0	0	2	0
All	All	8217	0	7554	122	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 8.

All (122) 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:A:153:SER:HB2	1:A:157:ALA:H	1.24	1.02
1:A:152:LYS:HE2	1:A:158:TYR:HB3	1.44	0.97
1:A:124:ASN:ND2	1:A:161:GLN:OE1	1.98	0.95
1:A:14:HIS:HD2	2:C:348:TRP:HA	1.32	0.94
1:B:14:HIS:HD2	2:D:348:TRP:HA	1.34	0.92
1:A:188:GLN:NE2	1:A:194:ASN:O	2.04	0.91
2:D:477:GLU:OE2	6:D:701:HOH:O	1.91	0.89

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:124:ASN:OD1	1:A:125:THR:N	2.04	0.88
1:A:244:THR:HG22	1:A:246:GLY:H	1.36	0.86
1:A:152:LYS:HE2	1:A:158:TYR:CB	2.05	0.85
1:A:152:LYS:HD2	1:A:193:LYS:N	1.92	0.84
2:D:385:LYS:HG3	2:D:386:THR:HB	1.58	0.83
1:A:152:LYS:HD2	1:A:193:LYS:H	1.43	0.82
1:A:14:HIS:CD2	2:C:348:TRP:HA	2.15	0.81
1:B:5:GLY:HA2	2:D:468:PHE:HA	1.63	0.81
1:A:188:GLN:O	1:A:192:TYR:N	2.15	0.78
1:B:14:HIS:CD2	2:D:348:TRP:HA	2.19	0.78
1:A:124:ASN:HB2	1:A:161:GLN:HE22	1.49	0.78
1:A:61:ASP:OD1	6:A:701:HOH:O	2.02	0.76
2:D:386:THR:OG1	2:D:387:ASN:N	2.17	0.75
1:A:311:SER:O	6:A:702:HOH:O	2.05	0.73
1:A:184:THR:O	1:A:214:ILE:HG21	1.89	0.73
1:A:127:LYS:HD3	1:A:151:VAL:HG13	1.72	0.70
1:A:5:GLY:HA2	2:C:468:PHE:HA	1.74	0.69
1:A:120:GLU:OE2	1:A:122:GLN:HG3	1.93	0.69
1:B:61:ASP:OD1	6:B:702:HOH:O	2.10	0.68
1:A:124:ASN:CG	1:A:125:THR:H	1.95	0.68
1:B:124:ASN:OD1	1:B:125:THR:N	2.27	0.68
1:A:185:ASP:HA	1:A:187:GLU:HG3	1.74	0.67
2:C:346:ASP:OD1	6:C:601:HOH:O	2.11	0.67
2:D:450:ARG:NH2	6:D:702:HOH:O	2.30	0.65
1:B:14:HIS:CE1	1:B:33:VAL:HG21	2.32	0.65
1:A:152:LYS:CE	1:A:158:TYR:HB3	2.25	0.64
1:A:319:ARG:NE	6:A:704:HOH:O	2.19	0.64
1:A:188:GLN:O	1:A:191:LEU:N	2.30	0.64
2:D:385:LYS:HG3	2:D:386:THR:CB	2.26	0.62
2:D:385:LYS:HE3	2:D:386:THR:HB	1.81	0.61
1:A:184:THR:O	1:A:187:GLU:HG3	2.00	0.61
4:A:601:NAG:O4	6:A:703:HOH:O	2.16	0.60
2:D:386:THR:CG2	2:D:387:ASN:H	2.15	0.59
1:A:124:ASN:HB3	1:A:126:VAL:CG2	2.32	0.59
1:A:152:LYS:NZ	1:A:193:LYS:HB2	2.18	0.59
4:B:602:NAG:O4	6:B:701:HOH:O	2.01	0.59
1:A:187:GLU:OE1	1:A:189:THR:N	2.36	0.58
1:B:149:TRP:CH2	3:E:3:SIA:H7	2.38	0.58
4:B:601:NAG:H83	4:B:601:NAG:H3	1.84	0.58
2:D:383:ILE:HG13	2:D:383:ILE:O	2.04	0.57
1:A:14:HIS:CE1	1:A:33:VAL:HG21	2.40	0.57

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:124:ASN:O	6:B:703:HOH:O	2.17	0.57
1:A:120:GLU:CD	1:A:122:GLN:HG3	2.26	0.56
2:D:386:THR:HG23	2:D:387:ASN:H	1.70	0.56
2:D:386:THR:HG23	2:D:387:ASN:N	2.20	0.55
1:A:154:ASP:HA	6:A:738:HOH:O	2.06	0.55
1:A:124:ASN:HB3	1:A:126:VAL:HG23	1.88	0.54
4:B:602:NAG:H83	6:B:743:HOH:O	2.07	0.54
1:B:305:ARG:NH1	6:B:707:HOH:O	2.41	0.54
1:B:94:TYR:OH	3:E:3:SIA:H91	2.08	0.54
1:A:152:LYS:HD3	1:A:193:LYS:HG3	1.90	0.53
1:A:14:HIS:HD2	2:C:348:TRP:CA	2.13	0.53
1:B:14:HIS:HD2	2:D:348:TRP:CA	2.13	0.53
1:A:124:ASN:HB2	1:A:161:GLN:NE2	2.20	0.52
1:B:238:ASP:OD1	1:B:239:LEU:N	2.39	0.52
1:A:43:ASN:HB3	1:A:295:ILE:HD13	1.92	0.52
1:A:152:LYS:HE3	1:A:192:TYR:HA	1.91	0.52
2:D:378:LYS:NZ	2:D:430:GLU:O	2.42	0.52
1:A:153:SER:CB	1:A:157:ALA:H	2.10	0.51
1:A:200:THR:HB	1:A:243:ASN:HB3	1.94	0.49
1:B:152:LYS:HD3	1:B:193:LYS:H	1.76	0.49
1:B:124:ASN:OD1	1:B:124:ASN:C	2.51	0.49
1:A:156:ASN:O	1:A:193:LYS:NZ	2.45	0.49
1:B:171:TYR:CZ	1:B:256:LYS:HD3	2.46	0.49
1:B:152:LYS:NZ	1:B:189:THR:O	2.32	0.48
1:A:187:GLU:HG2	1:A:188:GLN:CB	2.42	0.48
1:A:153:SER:HB3	1:A:155:GLY:N	2.27	0.48
1:A:192:TYR:CZ	1:A:247:ASN:HA	2.49	0.48
1:B:260:GLN:NE2	1:B:262:LYS:HD2	2.28	0.48
2:C:378:LYS:NZ	2:C:430:GLU:O	2.47	0.48
1:A:152:LYS:HZ2	1:A:193:LYS:HB2	1.78	0.47
1:A:14:HIS:CE1	1:A:33:VAL:HG11	2.49	0.46
2:C:378:LYS:HZ1	2:C:433:HIS:HB3	1.80	0.46
1:A:152:LYS:CE	1:A:192:TYR:HA	2.46	0.46
1:B:160:LEU:HD23	1:B:245:ILE:HG23	1.97	0.46
1:A:152:LYS:H	1:A:152:LYS:HG2	1.39	0.46
2:D:357:GLU:OE2	2:D:472:ASP:HB2	2.15	0.46
1:B:158:TYR:HE2	1:B:245:ILE:HA	1.79	0.46
1:A:76:ASN:OD1	1:A:146:ARG:NH2	2.43	0.46
2:C:451:ARG:HD2	6:C:606:HOH:O	2.16	0.46
1:B:14:HIS:HE1	1:B:33:VAL:HG21	1.77	0.46
1:A:213:ASN:O	1:A:217:ARG:NH2	2.43	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:303:CYS:O	2:D:387:ASN:HB3	2.16	0.45
1:A:127:LYS:HB3	1:A:151:VAL:O	2.16	0.45
1:A:186:THR:H	1:A:187:GLU:HB3	1.82	0.45
1:A:185:ASP:CA	1:A:187:GLU:HG3	2.43	0.45
2:C:342:GLN:NE2	6:C:607:HOH:O	2.48	0.45
1:A:187:GLU:CG	1:A:188:GLN:N	2.80	0.44
1:A:187:GLU:CD	1:A:189:THR:H	2.21	0.44
1:B:217:ARG:O	1:B:224:SER:OG	2.33	0.44
1:A:6:ASN:H	1:A:7:PRO:CD	2.31	0.44
1:A:182:PRO:HG2	1:A:188:GLN:HB2	2.00	0.44
1:A:217:ARG:O	1:A:224:SER:HB2	2.18	0.43
2:C:380:ASN:O	2:C:384:GLU:HG3	2.18	0.43
1:A:175:TYR:CE2	1:A:254:HIS:HB3	2.54	0.43
1:B:161:GLN:HB2	1:B:244:THR:HG23	2.00	0.43
1:A:187:GLU:HG2	1:A:188:GLN:HB2	2.01	0.42
1:A:120:GLU:OE2	1:A:122:GLN:N	2.53	0.42
1:A:180:HIS:HB2	1:A:249:ILE:HD11	2.02	0.42
1:B:14:HIS:CE1	1:B:33:VAL:HG11	2.55	0.42
2:C:398:GLU:H	2:C:398:GLU:HG3	1.62	0.42
1:A:191:LEU:HD11	5:A:603:SIA:H92	2.01	0.41
1:A:14:HIS:HE1	1:A:33:VAL:HG11	1.85	0.41
1:A:219:LEU:HA	1:A:223:GLN:O	2.20	0.41
1:A:164:THR:HG23	1:A:239:LEU:HD23	2.02	0.41
2:D:366:LYS:HB2	2:D:366:LYS:HE3	1.68	0.41
2:D:447:GLU:OE1	2:D:450:ARG:NH1	2.54	0.41
1:A:238:ASP:OD1	1:A:239:LEU:N	2.48	0.41
1:A:110:ASN:HA	1:A:263:SER:O	2.21	0.41
1:B:173:ARG:NE	6:B:705:HOH:O	2.50	0.41
1:A:223:GLN:NE2	5:A:603:SIA:O1B	2.54	0.40
1:B:43:ASN:HB3	1:B:295:ILE:HD13	2.04	0.40
1:B:180:HIS:HB2	1:B:249:ILE:HD11	2.03	0.40
1:A:188:GLN:HG2	1:A:192:TYR:HB2	2.03	0.40
1:A:279:CYS:HB2	1:A:302:ASP:O	2.21	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	317/327 (97%)	290 (92%)	19 (6%)	8 (2%)	5	3
1	B	317/327 (97%)	301 (95%)	14 (4%)	2 (1%)	25	26
2	C	170/176 (97%)	158 (93%)	11 (6%)	1 (1%)	25	26
2	D	170/176 (97%)	159 (94%)	11 (6%)	0	100	100
All	All	974/1006 (97%)	908 (93%)	55 (6%)	11 (1%)	14	12

All (11) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	153	SER
1	A	188	GLN
1	A	193	LYS
1	B	6	ASN
1	B	193	LYS
1	A	138	ALA
1	A	162	ASN
1	A	187	GLU
1	A	189	THR
2	C	385	LYS
1	A	6	ASN

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	274/282 (97%)	267 (97%)	7 (3%)	46	58
1	B	274/282 (97%)	273 (100%)	1 (0%)	91	96
2	C	147/150 (98%)	147 (100%)	0	100	100
2	D	147/150 (98%)	146 (99%)	1 (1%)	84	91
All	All	842/864 (98%)	833 (99%)	9 (1%)	73	85

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	126	VAL
1	A	152	LYS
1	A	187	GLU
1	A	193	LYS
1	A	220	VAL
1	A	224	SER
1	A	239	LEU
1	B	239	LEU
2	D	385	LYS

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

Mol	Chain	Res	Type
1	A	260	GLN

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 ⓘ

3 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	E	1	3	15,15,15	0.40	0	21,21,21	0.67	1 (4%)
3	GAL	E	2	3	11,11,12	0.55	0	15,15,17	1.07	1 (6%)
3	SIA	E	3	3	20,20,21	0.58	0	24,28,31	0.78	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	E	1	3	-	4/6/26/26	0/1/1/1
3	GAL	E	2	3	-	2/2/19/22	0/1/1/1
3	SIA	E	3	3	-	2/18/34/38	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	E	2	GAL	O5-C5-C4	-2.34	105.14	110.83
3	E	1	NAG	O4-C4-C5	2.22	114.82	109.30

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	E	2	GAL	O5-C5-C6-O6
3	E	1	NAG	O5-C5-C6-O6
3	E	1	NAG	C8-C7-N2-C2
3	E	1	NAG	O7-C7-N2-C2
3	E	1	NAG	C4-C5-C6-O6
3	E	2	GAL	C4-C5-C6-O6
3	E	3	SIA	O8-C8-C9-O9
3	E	3	SIA	C7-C8-C9-O9

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	3	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 E

Bond lengths

Bond angles

Torsions

Rings
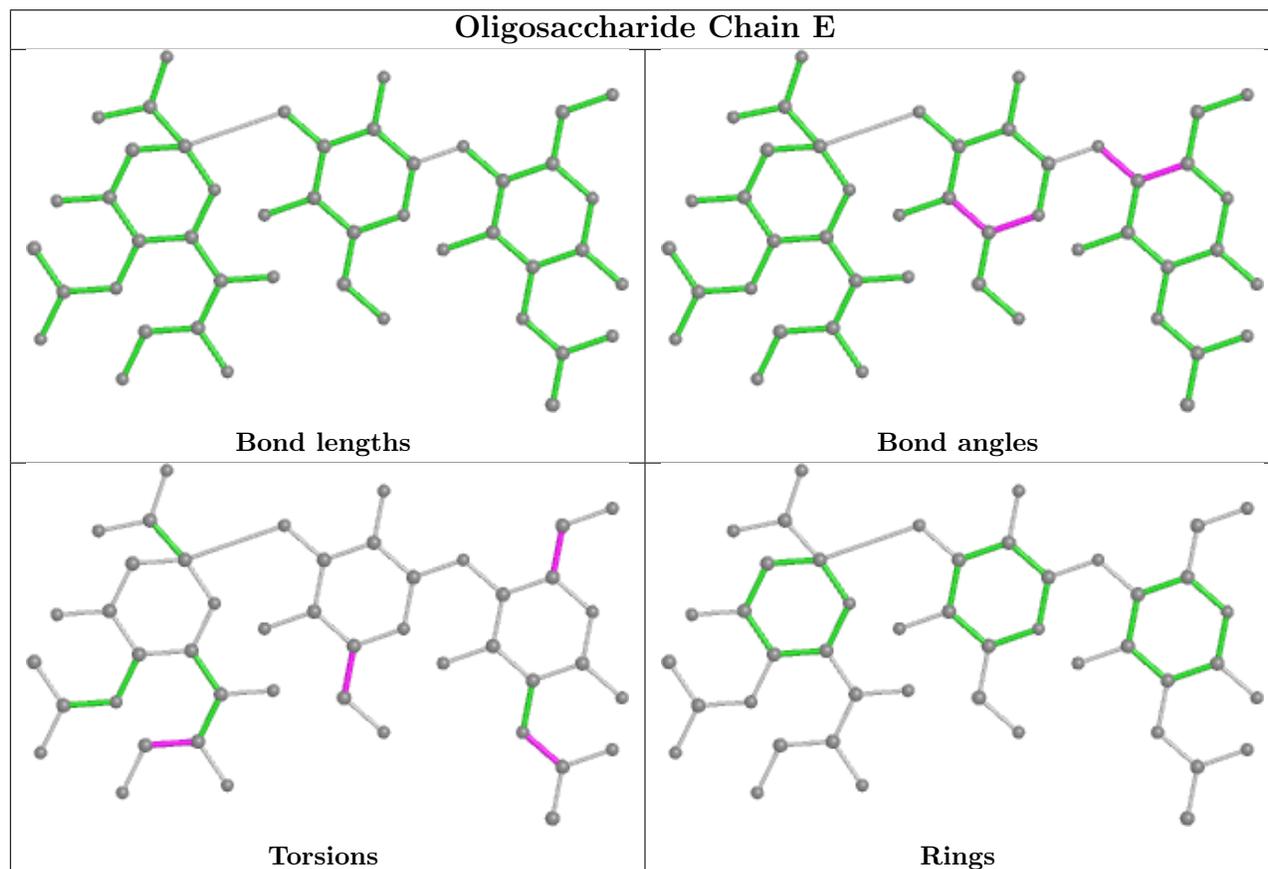
5.6 Ligand geometry (i)

5 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	NAG	A	602	1	14,14,15	0.26	0	17,19,21	0.53	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
5	SIA	A	603	-	21,21,21	0.87	1 (4%)	25,31,31	1.12	3 (12%)
4	NAG	B	602	1	14,14,15	0.36	0	17,19,21	0.52	0
4	NAG	B	601	1	14,14,15	0.52	0	17,19,21	1.28	2 (11%)
4	NAG	A	601	1	14,14,15	0.82	1 (7%)	17,19,21	0.48	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	NAG	A	602	1	-	2/6/23/26	0/1/1/1
5	SIA	A	603	-	-	3/20/38/38	0/1/1/1
4	NAG	B	602	1	-	3/6/23/26	0/1/1/1
4	NAG	B	601	1	-	5/6/23/26	0/1/1/1
4	NAG	A	601	1	-	2/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	603	SIA	O2-C2	3.06	1.43	1.39
4	A	601	NAG	O5-C1	-2.46	1.39	1.43

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	601	NAG	C2-N2-C7	4.28	129.00	122.90
5	A	603	SIA	O1A-C1-C2	-3.72	117.95	123.59
5	A	603	SIA	C3-C2-C1	-2.06	109.16	113.00
4	B	601	NAG	C1-C2-N2	2.05	113.99	110.49
5	A	603	SIA	C8-C7-C6	-2.02	109.21	113.03

There are no chirality outliers.

All (15) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	601	NAG	O5-C5-C6-O6
4	B	601	NAG	C4-C5-C6-O6
4	A	602	NAG	C8-C7-N2-C2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
4	A	602	NAG	O7-C7-N2-C2
4	B	601	NAG	C8-C7-N2-C2
4	B	601	NAG	O7-C7-N2-C2
4	B	602	NAG	C8-C7-N2-C2
4	B	602	NAG	O7-C7-N2-C2
4	A	601	NAG	O5-C5-C6-O6
4	A	601	NAG	C4-C5-C6-O6
4	B	602	NAG	O5-C5-C6-O6
5	A	603	SIA	C6-C7-C8-O8
5	A	603	SIA	O1A-C1-C2-O6
5	A	603	SIA	O1A-C1-C2-O2
4	B	601	NAG	C3-C2-N2-C7

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	603	SIA	2	0
4	B	602	NAG	2	0
4	B	601	NAG	1	0
4	A	601	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 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 A 603

Bond lengths

Bond angles

Torsions

Rings
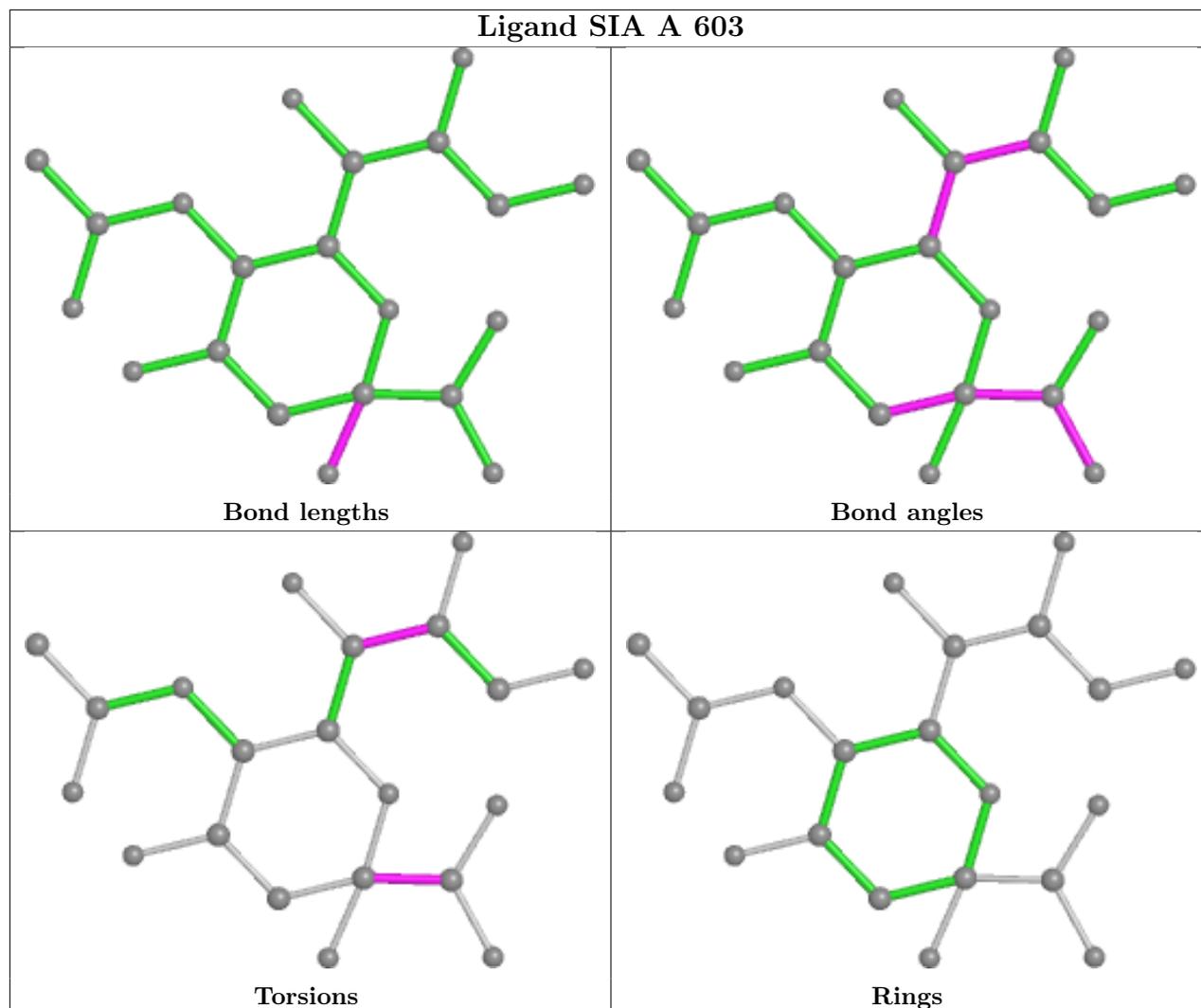
5.7 Other polymers (i)

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(Å^2)	Q<0.9
1	A	319/327 (97%)	0.42	39 (12%)	4	3	26, 68, 122, 168	0
1	B	319/327 (97%)	0.03	20 (6%)	20	19	23, 49, 92, 127	0
2	C	172/176 (97%)	-0.09	2 (1%)	79	77	27, 37, 64, 123	0
2	D	172/176 (97%)	-0.18	1 (0%)	89	88	23, 34, 56, 135	0
All	All	982/1006 (97%)	0.10	62 (6%)	20	19	23, 47, 107, 168	0

All (62) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	157	ALA	7.0
1	A	191	LEU	6.2
1	A	158	TYR	5.4
1	B	158	TYR	5.0
1	A	193	LYS	5.0
1	A	184	THR	4.8
1	A	192	TYR	4.7
1	A	123	TRP	4.4
1	B	157	ALA	4.4
1	A	221	ARG	4.3
1	A	127	LYS	4.3
2	C	386	THR	4.2
1	A	151	VAL	4.1
1	A	190	ASN	4.1
1	A	185	ASP	3.7
1	A	139	ASN	3.6
1	A	245	ILE	3.6
1	B	155	GLY	3.5
1	A	124	ASN	3.2
1	A	215	GLY	3.1
1	B	186	THR	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	192	TYR	3.0
1	A	160	LEU	3.0
1	A	219	LEU	3.0
1	A	156	ASN	3.0
1	A	121	PHE	3.0
1	B	193	LYS	3.0
1	A	143	PHE	2.9
1	A	162	ASN	2.9
1	B	216	SER	2.9
1	A	147	LEU	2.9
1	A	120	GLU	2.8
1	A	188	GLN	2.8
1	B	215	GLY	2.7
1	A	126	VAL	2.7
2	D	386	THR	2.7
1	A	99	PRO	2.7
1	B	184	THR	2.7
1	B	156	ASN	2.6
1	A	186	THR	2.6
1	B	219	LEU	2.6
1	A	5	GLY	2.6
1	A	94	TYR	2.6
1	A	222	GLY	2.5
2	C	385	LYS	2.4
1	A	166	ILE	2.4
1	B	221	ARG	2.3
1	B	191	LEU	2.3
1	B	152	LYS	2.3
1	B	151	VAL	2.3
1	B	187	GLU	2.2
1	A	225	GLY	2.2
1	A	150	LEU	2.2
1	B	197	GLY	2.2
1	A	196	PRO	2.2
1	B	5	GLY	2.2
1	A	152	LYS	2.2
1	A	220	VAL	2.1
1	A	141	ASN	2.1
1	B	99	PRO	2.1
1	B	190	ASN	2.0
1	A	194	ASN	2.0

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

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

6.3 Carbohydrates (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 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	E	1	15/15	0.33	0.41	133,146,155,157	0
3	GAL	E	2	11/12	0.81	0.22	92,101,115,120	0
3	SIA	E	3	20/21	0.88	0.17	72,80,86,87	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.
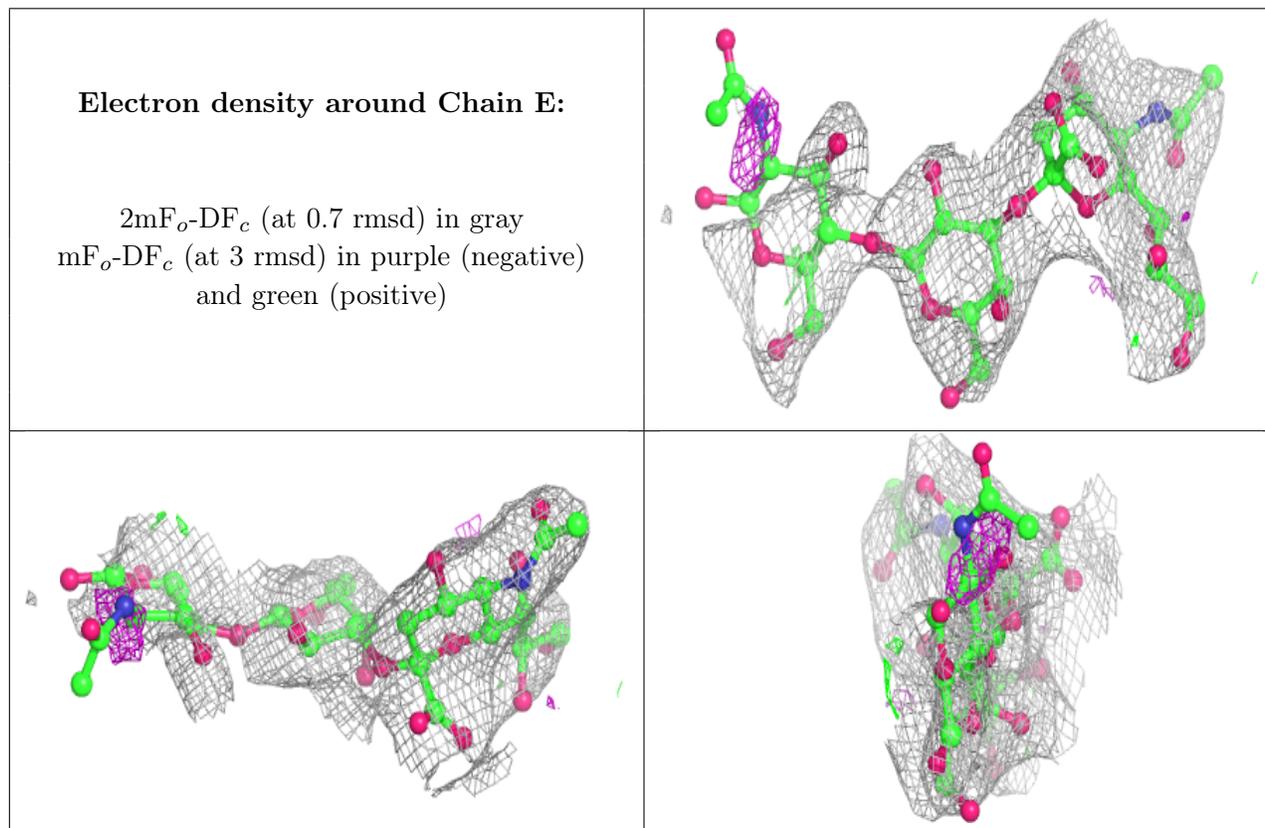

Electron density around Chain E:

$2mF_o\text{-}DF_c$ (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
4	NAG	A	601	14/15	0.79	0.30	116,123,128,133	0
4	NAG	B	601	14/15	0.79	0.28	108,114,117,118	0
5	SIA	A	603	21/21	0.90	0.19	92,96,99,106	0
4	NAG	B	602	14/15	0.94	0.10	30,36,40,50	0
4	NAG	A	602	14/15	0.94	0.12	40,45,48,60	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 A 603:

$2mF_o$-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 (i)

There are no such residues in this entry.