wwPDB X-ray Structure Validation Summary Report (i)

Nov 5, 2026 – 06:12 PM JST

PDB ID	:	47SS / pdb_000047ss
Title	:	The structure of hemagglutinin T160A mutant (H3 numbering) from H9N2 influenza virus (A/chicken/Tianjin/CAS/2007) in complex with avian receptor analog 3'SLNLN
Authors	:	Wang, M.; Peng, R.C.; Qi, J.X.; Shi, Y.; Gao, G.F.
Deposited on	:	2026-10-30
Resolution	:	2.20 Å(reported)

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

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

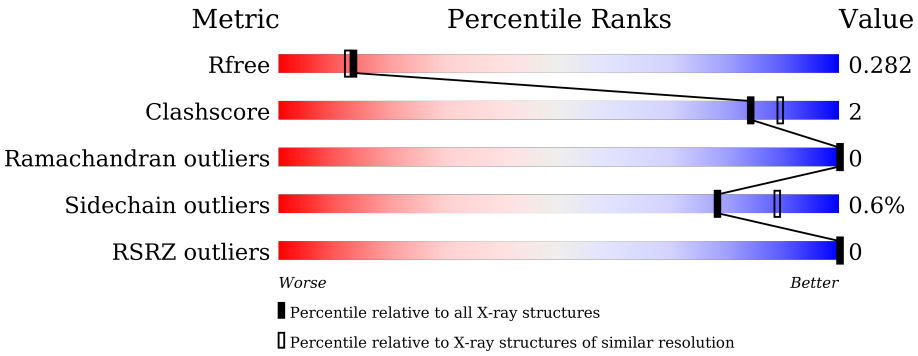
MolProbity	:	4-5-2 with Phenix2.2
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	2.2
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.51

1 Overall quality at a glance (i)

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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	6164 (2.20-2.20)
Clashscore	190562	6851 (2.20-2.20)
Ramachandran outliers	187476	6768 (2.20-2.20)
Sidechain outliers	187428	6769 (2.20-2.20)
RSRZ outliers	180081	6166 (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	320	91% 8% .
1	D	320	91% 8% .
2	B	176	89% . 8%
2	E	176	89% . 8%
3	C	2	50% 50%
3	F	2	50% 50%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 7900 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 (Fragment).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	316	Total	C	N	O	S	0	0	0
			2475	1553	437	473	12			
1	D	316	Total	C	N	O	S	0	1	0
			2485	1559	440	474	12			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	86	VAL	THR	conflict	UNP A7YL49
D	86	VAL	THR	conflict	UNP A7YL49

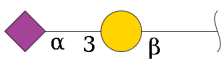
- Molecule 2 is a protein called Hemagglutinin (Fragment).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	162	Total	C	N	O	S	0	0	0
			1291	802	227	255	7			
2	E	162	Total	C	N	O	S	0	0	0
			1291	802	227	255	7			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	385	VAL	ILE	conflict	UNP A7YL88
E	385	VAL	ILE	conflict	UNP A7YL88

- Molecule 3 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galacto pyranose.



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
3	C	2	Total C N O 31 17 1 13	0	0	0
3	F	2	Total C N O 31 17 1 13	0	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	75	Total O 75 75	0	0
4	B	63	Total O 63 63	0	0
4	D	100	Total O 100 100	0	0
4	E	58	Total O 58 58	0	0

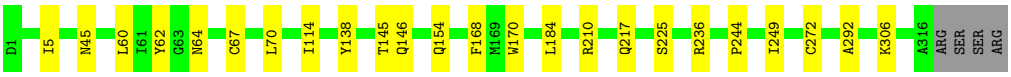
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

• Molecule 1: Hemagglutinin (Fragment)

Chain A:





• Molecule 1: Hemagglutinin (Fragment)

Chain D:

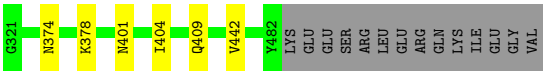




• Molecule 2: Hemagglutinin (Fragment)

Chain B:

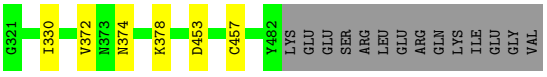




• Molecule 2: Hemagglutinin (Fragment)

Chain E:





• Molecule 3: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose

Chain C:





• Molecule 3: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose

Chain F:




GAL1
SIA2

4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α, β, γ	94.81Å 94.81Å 357.53Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.01 – 2.20 20.01 – 2.20	Depositor EDS
% Data completeness (in resolution range)	96.3 (20.01-2.20) 96.3 (20.01-2.20)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.60 (at 2.21Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.253 , 0.281 0.253 , 0.282	Depositor DCC
R_{free} test set	2996 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å^2)	28.4	Xtriage
Anisotropy	0.074	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 11.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.40$, $< L^2 > = 0.22$	Xtriage
Estimated twinning fraction	0.326 for -h-k,k,-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	7900	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.09% 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: GAL, SIA

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	0.10	0/2533	0.31	0/3453
1	D	0.11	0/2544	0.31	0/3469
2	B	0.11	0/1312	0.28	0/1767
2	E	0.12	0/1312	0.28	0/1767
All	All	0.11	0/7701	0.30	0/10456

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	2475	0	2417	14	0
1	D	2485	0	2424	15	0
2	B	1291	0	1230	3	0
2	E	1291	0	1230	5	0
3	C	31	0	26	2	0
3	F	31	0	26	2	0
4	A	75	0	0	1	0
4	B	63	0	0	0	0
4	D	100	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	E	58	0	0	1	0
All	All	7900	0	7353	34	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 2.

The worst 5 of 34 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:D:64:ASN:HB3	1:D:67:CYS:SG	2.46	0.55
1:A:60:LEU:HD12	1:A:70:LEU:HD13	1.89	0.55
1:D:311:LEU:HD23	1:D:311:LEU:H	1.72	0.54
1:D:67:CYS:HB3	1:D:70:LEU:HD12	1.91	0.53
1:A:64:ASN:HB3	1:A:67:CYS:SG	2.49	0.52

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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	314/320 (98%)	302 (96%)	12 (4%)	0	100	100
1	D	315/320 (98%)	302 (96%)	13 (4%)	0	100	100
2	B	160/176 (91%)	157 (98%)	3 (2%)	0	100	100
2	E	160/176 (91%)	158 (99%)	2 (1%)	0	100	100
All	All	949/992 (96%)	919 (97%)	30 (3%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

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	278/282 (99%)	276 (99%)	2 (1%)	76 87
1	D	279/282 (99%)	277 (99%)	2 (1%)	76 87
2	B	137/150 (91%)	136 (99%)	1 (1%)	76 87
2	E	137/150 (91%)	137 (100%)	0	100 100
All	All	831/864 (96%)	826 (99%)	5 (1%)	78 89

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	146	GLN
1	A	306	LYS
2	B	409	GLN
1	D	19	LEU
1	D	179	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	123	ASN
1	D	214	ASN
2	E	345	GLN
2	E	398	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

4 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	GAL	C	1	3	11,11,12	0.52	0	15,15,17	0.90	0
3	SIA	C	2	3	20,20,21	1.94	3 (15%)	24,28,31	1.76	5 (20%)
3	GAL	F	1	3	11,11,12	0.50	0	15,15,17	0.87	0
3	SIA	F	2	3	20,20,21	1.98	2 (10%)	24,28,31	1.65	3 (12%)

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	GAL	C	1	3	-	0/2/19/22	0/1/1/1
3	SIA	C	2	3	-	6/18/34/38	0/1/1/1
3	GAL	F	1	3	-	2/2/19/22	0/1/1/1
3	SIA	F	2	3	-	0/18/34/38	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	2	SIA	C2-C1	7.59	1.59	1.52
3	C	2	SIA	C2-C1	7.35	1.58	1.52
3	F	2	SIA	O6-C2	2.82	1.47	1.43
3	C	2	SIA	O6-C2	2.76	1.47	1.43
3	C	2	SIA	C7-C6	2.00	1.55	1.53

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	2	SIA	O1A-C1-C2	-4.53	111.88	122.57
3	F	2	SIA	O1A-C1-C2	-4.47	112.01	122.57

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	2	SIA	C6-O6-C2	4.26	120.45	111.34
3	F	2	SIA	C6-O6-C2	3.56	118.96	111.34
3	C	2	SIA	C3-C4-C5	-2.63	108.28	111.46

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	F	1	GAL	C4-C5-C6-O6
3	C	2	SIA	C6-C7-C8-O8
3	F	1	GAL	O5-C5-C6-O6
3	C	2	SIA	C7-C8-C9-O9
3	C	2	SIA	C6-C7-C8-C9

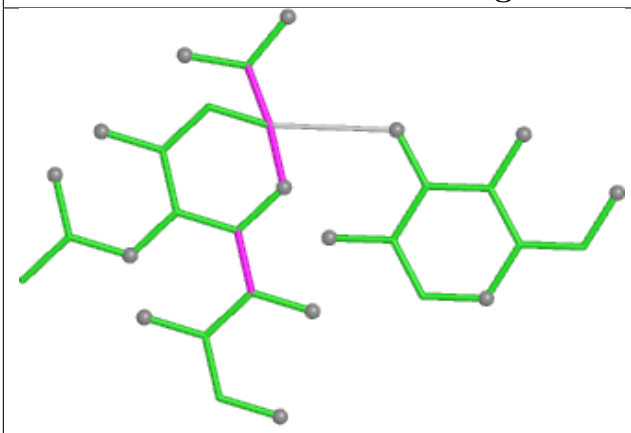
There are no ring outliers.

2 monomers are involved in 4 short contacts:

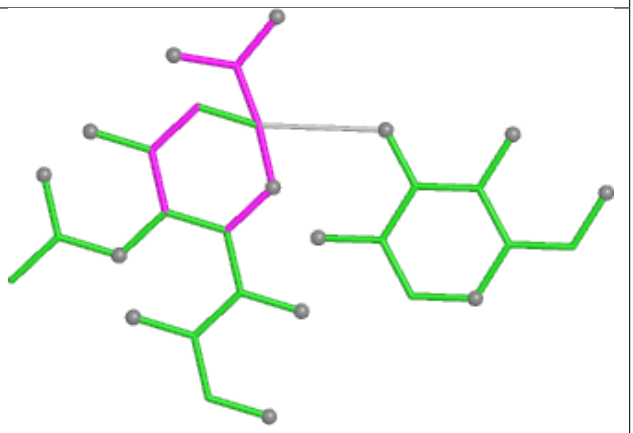
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	2	SIA	2	0
3	F	2	SIA	2	0

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

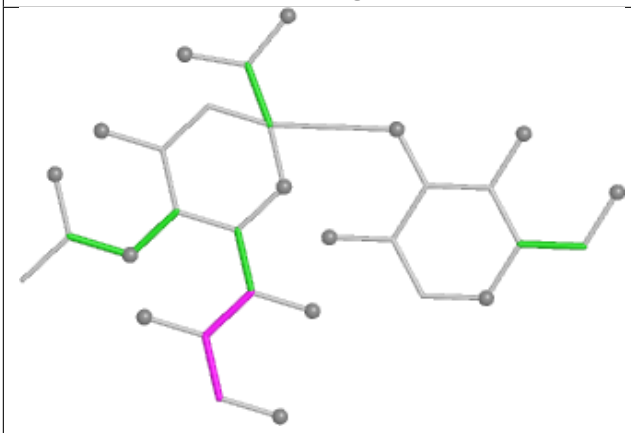
Oligosaccharide Chain C



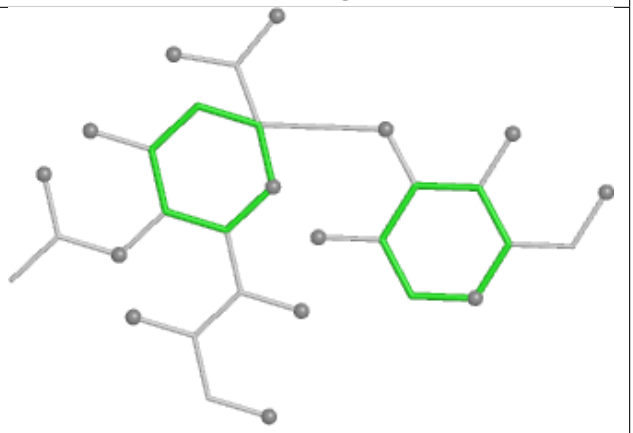
Bond lengths



Bond angles

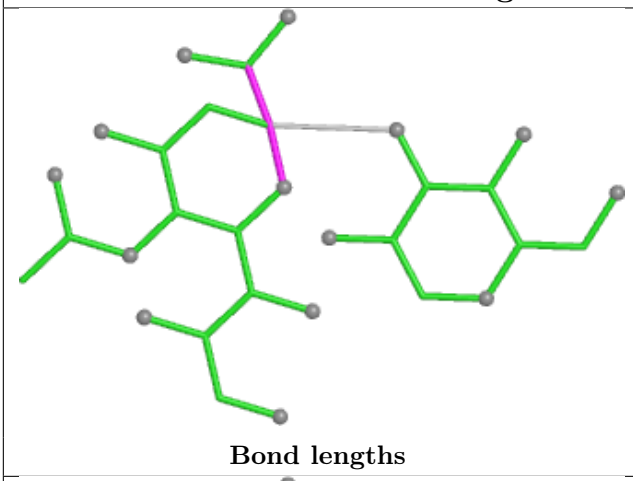


Torsions

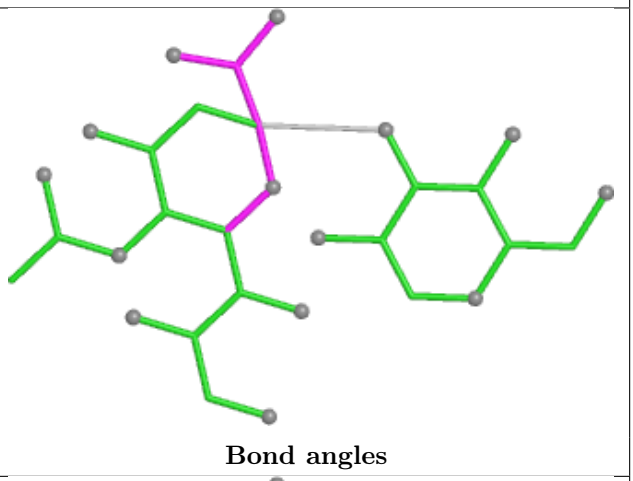


Rings

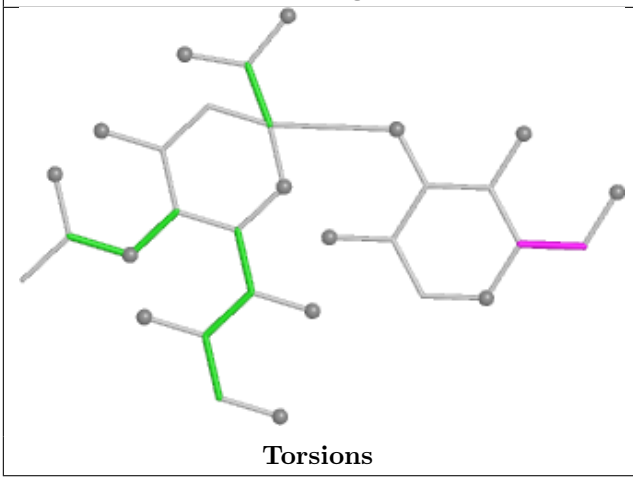
Oligosaccharide Chain F



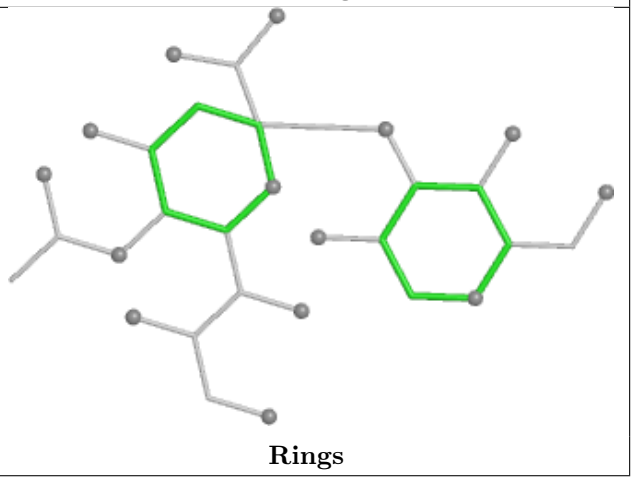
Bond lengths



Bond angles



Torsions



Rings

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} 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	316/320 (98%)	-1.59	0	100	100	19, 30, 47, 55	0
1	D	316/320 (98%)	-1.57	0	100	100	16, 31, 48, 60	1 (0%)
2	B	162/176 (92%)	-1.58	0	100	100	17, 29, 55, 66	0
2	E	162/176 (92%)	-1.49	0	100	100	18, 37, 64, 72	0
All	All	956/992 (96%)	-1.57	0	100	100	16, 31, 53, 72	1 (0%)

There are no RSRZ outliers to report.

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

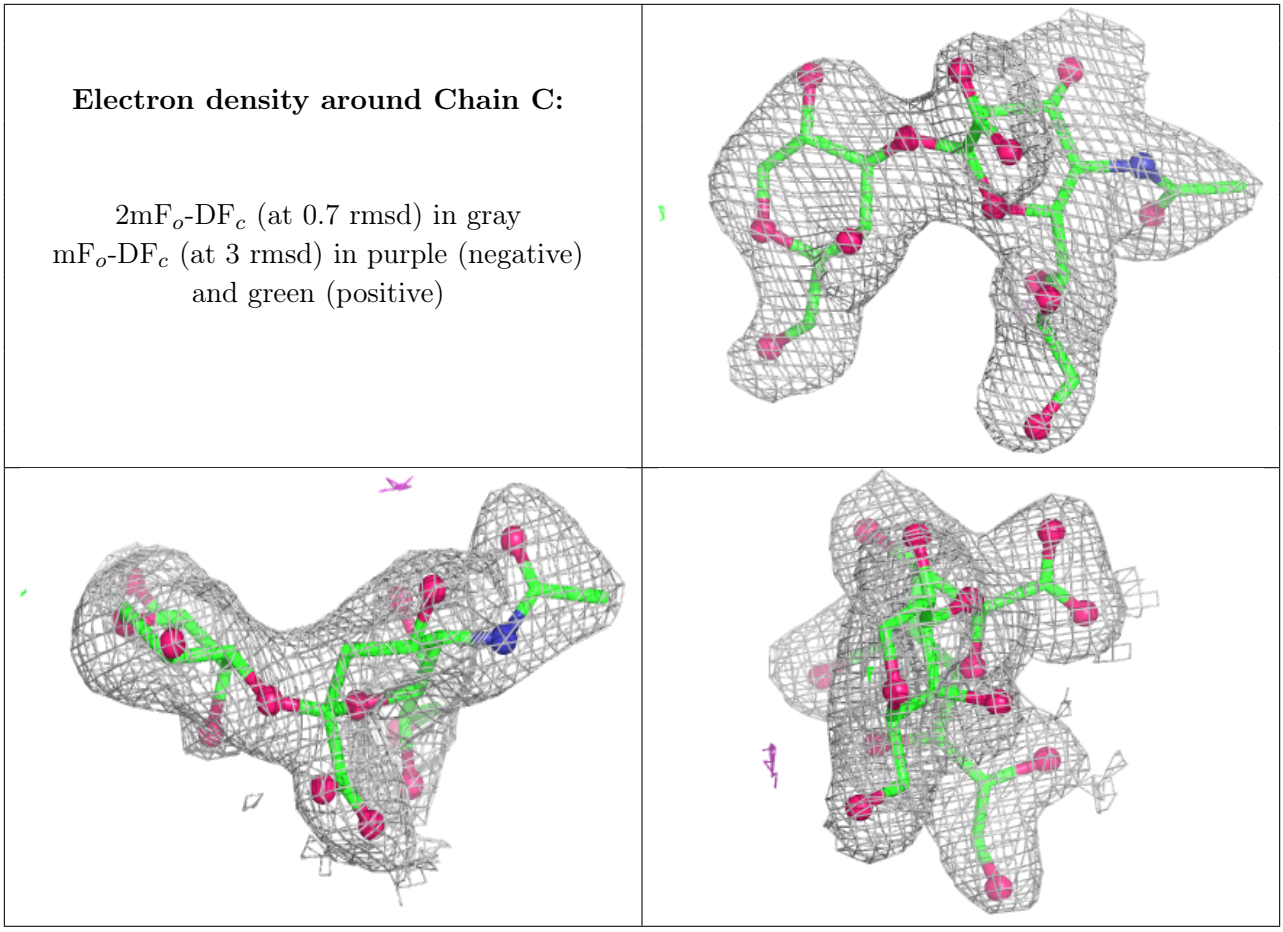
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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	GAL	C	1	11/12	-	-	37,49,52,57	0
3	SIA	C	2	20/21	-	-	32,35,43,45	0
3	GAL	F	1	11/12	0.99	0.04	34,46,49,54	0
3	SIA	F	2	20/21	1.00	0.02	32,35,43,45	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.

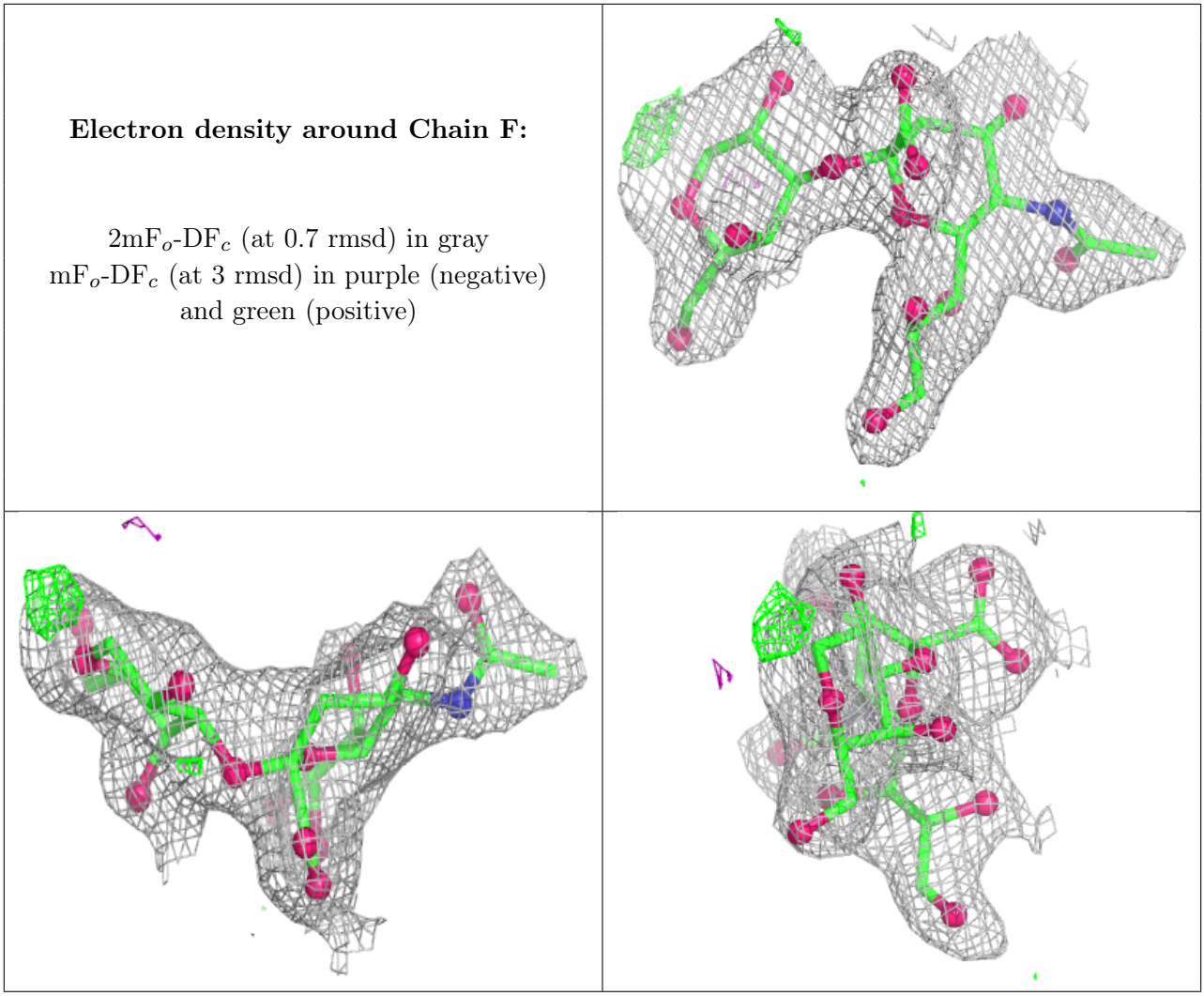
Electron density around Chain C:

$2mF_o$-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 F:

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



6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.