Full wwPDB X-ray Structure Validation Report ⓘ

Nov 23, 2023 – 01:58 AM JST

PDB ID	:	7WD1
Title	:	Crystal structure of R14 bound to SARS-CoV-2 RBD
Authors	:	Wang, Q.H.; Gao, G.F.; Qi, J.X.; Su, C.; Liu, H.H.; Wu, L.L.
Deposited on	:	2021-12-20
Resolution	:	2.50 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ 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 ⓘ) were used in the production of this report:

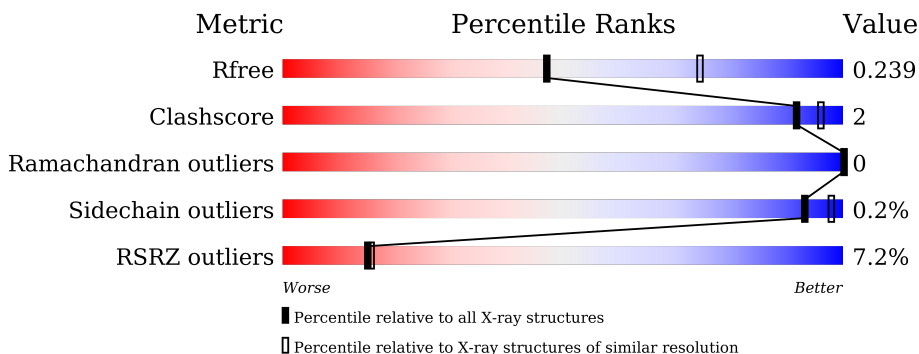|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 2.50 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		0.239
Clashscore		2
Ramachandran outliers		0
Sidechain outliers		0.2%
RSRZ outliers		7.2%

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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	196	18% / 93% / 6% ·
1	B	196	4% / 97% / ·
2	C	130	% / 93% / 7%
2	D	130	2% / 95% / 5%
3	F	3	67% / 33%
4	E	2	50% / 50%

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 5326 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 Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	194	Total	C	N	O	S	0	0	0
			1536	985	256	287	8			
1	B	196	Total	C	N	O	S	0	0	0
			1552	995	259	290	8			

- Molecule 2 is a protein called R14.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	130	Total	C	N	O	S	0	0	0
			988	617	166	199	6			
2	D	130	Total	C	N	O	S	0	0	0
			988	617	166	199	6			

- Molecule 3 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.
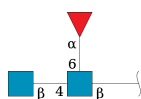
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	F	3	Total	C	N	O	0	0	0
			38	22	2	14			

- Molecule 4 is an oligosaccharide called alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	E	2	Total 24	C 14	N 1	O 9	0	0	0

- Molecule 5 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	46	Total 46	O 46	0	0
5	B	55	Total 55	O 55	0	0
5	C	40	Total 40	O 40	0	0
5	D	59	Total 59	O 59	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.
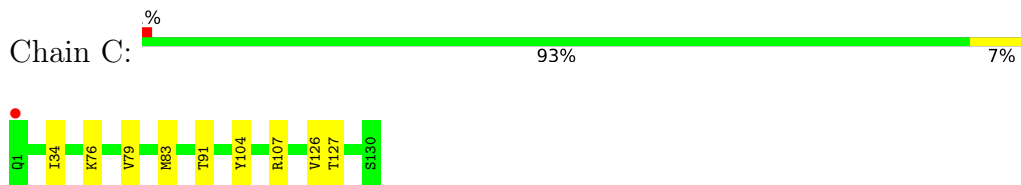
• Molecule 1: Spike protein S1

Chain A:

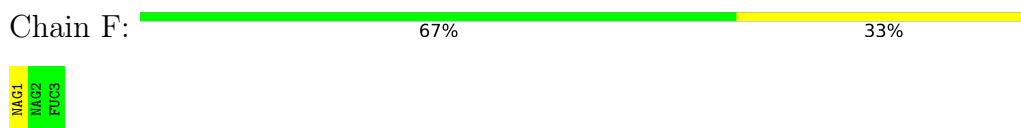• Molecule 1: Spike protein S1

Chain B:

• Molecule 2: R14

Chain C:

• Molecule 2: R14

Chain D:
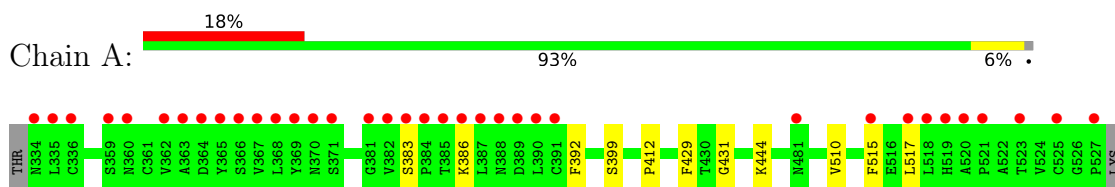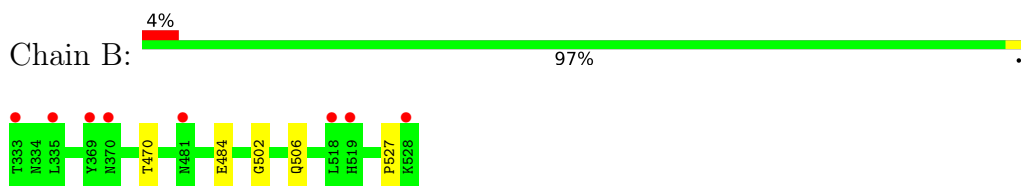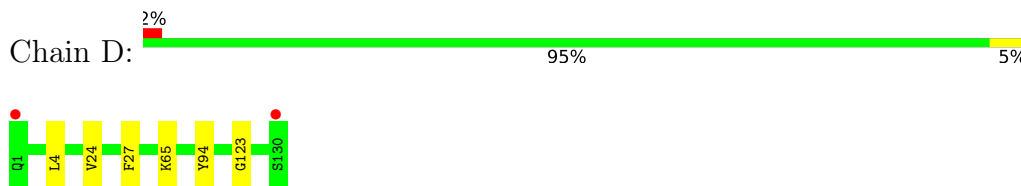
• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F:

• Molecule 4: alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	92.45Å　92.45Å　218.75Å 90.00°　90.00°　120.00°	Depositor
Resolution (Å)	46.22　–　2.50 46.23　–　2.50	Depositor EDS
% Data completeness (in resolution range)	93.9 (46.22-2.50) 93.9 (46.23-2.50)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.59 (at 2.51Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, R_{free}	0.201　,　0.239 0.201　,　0.239	Depositor DCC
R_{free} test set	1812 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	37.1	Xtriage
Anisotropy	0.107	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 36.2	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.027 for -h,-k,l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	5326	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 3.12% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: FUC, 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	0.27	0/1580	0.46	0/2151
1	B	0.27	0/1596	0.47	0/2172
2	C	0.29	0/1011	0.50	0/1371
2	D	0.30	0/1011	0.51	0/1371
All	All	0.28	0/5198	0.48	0/7065

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	B	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	527	PRO	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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1536	0	1452	6	0
1	B	1552	0	1472	3	0
2	C	988	0	928	6	0
2	D	988	0	928	3	0
3	F	38	0	34	0	0
4	E	24	0	22	0	0
5	A	46	0	0	0	0
5	B	55	0	0	0	0
5	C	40	0	0	0	0
5	D	59	0	0	0	0
All	All	5326	0	4836	16	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.

All (16) 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:B:484:GLU:OE2	2:C:107:ARG:NH1	2.33	0.61
2:C:91:THR:HG23	2:C:127:THR:HA	1.87	0.56
1:A:412:PRO:HG3	1:A:429:PHE:HB3	1.91	0.52
2:C:83:MET:HE1	2:C:126:VAL:HG21	1.95	0.49
1:B:502:GLY:O	1:B:506:GLN:HG3	2.16	0.46
1:A:392:PHE:HA	1:A:517:LEU:HD13	1.97	0.46
2:D:65:LYS:HB2	2:D:65:LYS:HE3	1.71	0.45
2:C:76:LYS:HE2	2:C:76:LYS:HB3	1.68	0.45
1:B:470:THR:HG21	2:C:104:TYR:CZ	2.52	0.43
1:A:444:LYS:HE2	1:A:444:LYS:HB3	1.64	0.42
2:C:34:ILE:HG21	2:C:79:VAL:HG21	2.02	0.42
1:A:383:SER:HB2	1:A:386:LYS:HD3	2.02	0.41
2:D:94:TYR:O	2:D:123:GLY:HA2	2.20	0.41
1:A:399:SER:HA	1:A:510:VAL:O	2.21	0.41
2:D:4:LEU:HD22	2:D:24:VAL:HG22	2.02	0.41
1:A:431:GLY:HA2	1:A:515:PHE:CD2	2.56	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	192/196 (98%)	182 (95%)	10 (5%)	0	100	100
1	B	194/196 (99%)	187 (96%)	7 (4%)	0	100	100
2	C	128/130 (98%)	127 (99%)	1 (1%)	0	100	100
2	D	128/130 (98%)	127 (99%)	1 (1%)	0	100	100
All	All	642/652 (98%)	623 (97%)	19 (3%)	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	167/169 (99%)	167 (100%)	0	100	100
1	B	169/169 (100%)	169 (100%)	0	100	100
2	C	104/104 (100%)	104 (100%)	0	100	100
2	D	104/104 (100%)	103 (99%)	1 (1%)	76	90
All	All	544/546 (100%)	543 (100%)	1 (0%)	93	98

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	27	PHE

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	487	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 ⓘ

5 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	#$|Z| > 2$
4	NAG	E	1	1,4	14,14,15	0.31	0	17,19,21	0.48	0
4	FUC	E	2	4	10,10,11	1.03	1 (10%)	14,14,16	0.75	0
3	NAG	F	1	1,3	14,14,15	0.38	0	17,19,21	0.70	1 (5%)
3	NAG	F	2	3	14,14,15	0.33	0	17,19,21	0.38	0
3	FUC	F	3	3	10,10,11	0.68	0	14,14,16	0.84	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	E	1	1,4	-	2/6/23/26	0/1/1/1
4	FUC	E	2	4	-	-	0/1/1/1
3	NAG	F	1	1,3	-	2/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	F	2	3	-	2/6/23/26	0/1/1/1
3	FUC	F	3	3	-	-	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	E	2	FUC	O5-C1	-2.18	1.40	1.43

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	1	NAG	C1-O5-C5	2.27	115.27	112.19

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	F	2	NAG	O5-C5-C6-O6
4	E	1	NAG	O5-C5-C6-O6
4	E	1	NAG	C4-C5-C6-O6
3	F	2	NAG	C4-C5-C6-O6
3	F	1	NAG	O5-C5-C6-O6
3	F	1	NAG	C4-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 F

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain E

Bond lengths

Bond angles

Torsions

Rings
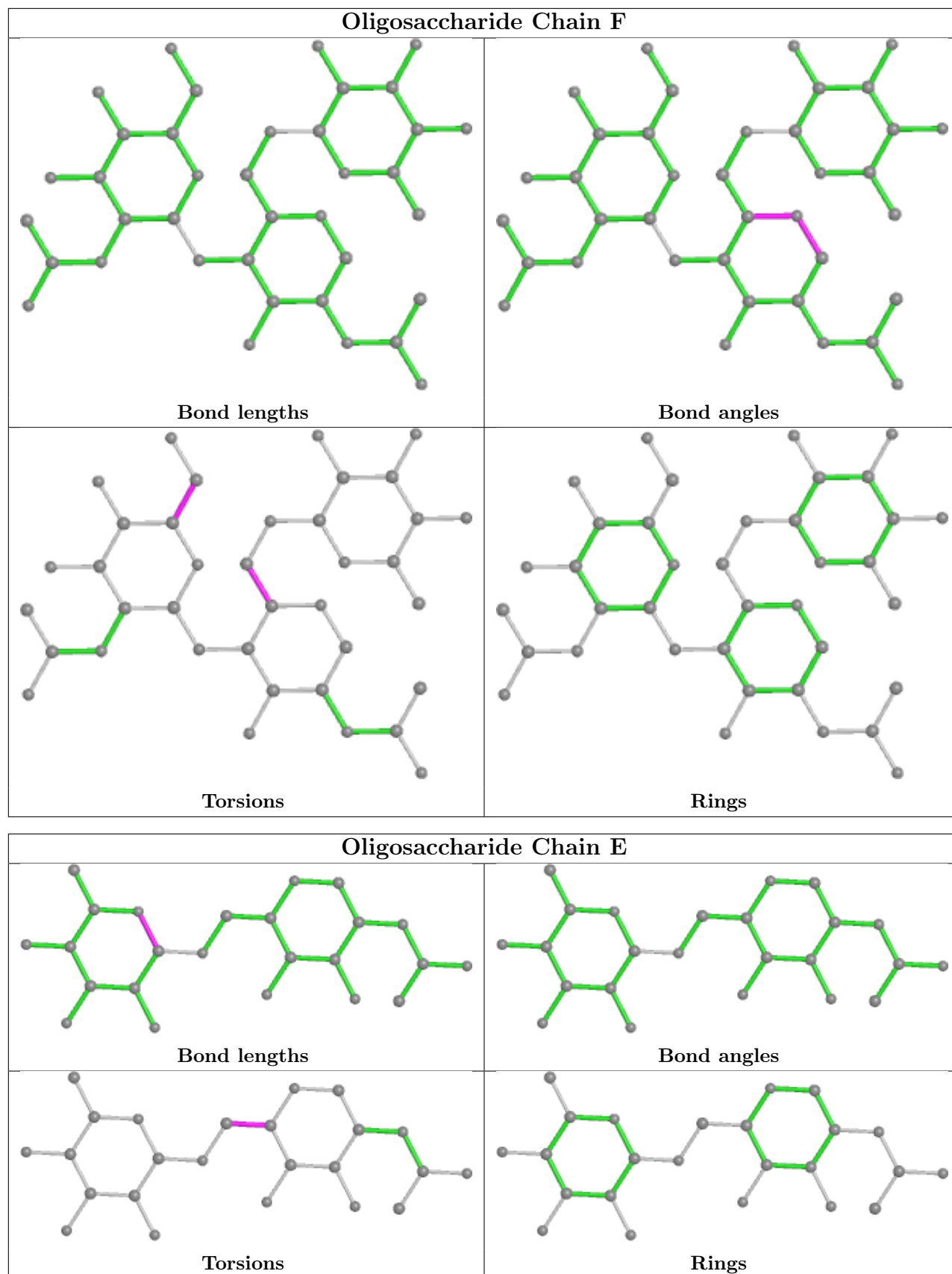
5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6　Fit of model and data ⓘ

6.1　Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	194/196 (98%)	0.65	36 (18%)	1	1	24, 50, 117, 136	0
1	B	196/196 (100%)	-0.04	8 (4%)	37	40	25, 41, 92, 113	0
2	C	130/130 (100%)	-0.28	1 (0%)	86	87	22, 33, 65, 84	0
2	D	130/130 (100%)	-0.35	2 (1%)	73	75	21, 32, 60, 76	0
All	All	650/652 (99%)	0.06	47 (7%)	15	16	21, 38, 104, 136	0

All (47) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	518	LEU	10.0
1	A	387	LEU	8.0
1	A	386	LYS	6.0
1	A	521	PRO	6.0
1	A	519	HIS	5.8
1	B	518	LEU	5.6
1	A	389	ASP	4.7
1	A	368	LEU	4.2
1	A	520	ALA	4.2
1	A	384	PRO	4.1
1	A	363	ALA	4.1
1	A	369	TYR	4.0
1	A	527	PRO	3.9
1	B	519	HIS	3.7
1	B	481	ASN	3.7
1	A	365	TYR	3.6
1	A	388	ASN	3.6
1	A	385	THR	3.4
1	A	517	LEU	3.4
2	D	130	SER	3.3
1	A	391	CYS	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	382	VAL	3.1
1	B	335	LEU	2.9
1	A	335	LEU	2.9
1	A	334	ASN	2.9
1	A	362	VAL	2.8
1	A	359	SER	2.8
1	A	383	SER	2.8
2	C	1	GLN	2.7
1	A	370	ASN	2.7
1	A	381	GLY	2.7
1	B	370	ASN	2.6
1	A	364	ASP	2.6
1	A	523	THR	2.6
1	A	515	PHE	2.5
1	B	333	THR	2.5
1	A	371	SER	2.5
1	B	369	TYR	2.5
1	A	525	CYS	2.5
1	B	528	LYS	2.5
1	A	390	LEU	2.4
1	A	481	ASN	2.4
1	A	366	SER	2.3
1	A	360	ASN	2.3
1	A	336	CYS	2.1
1	A	367	VAL	2.1
2	D	1	GLN	2.0

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.
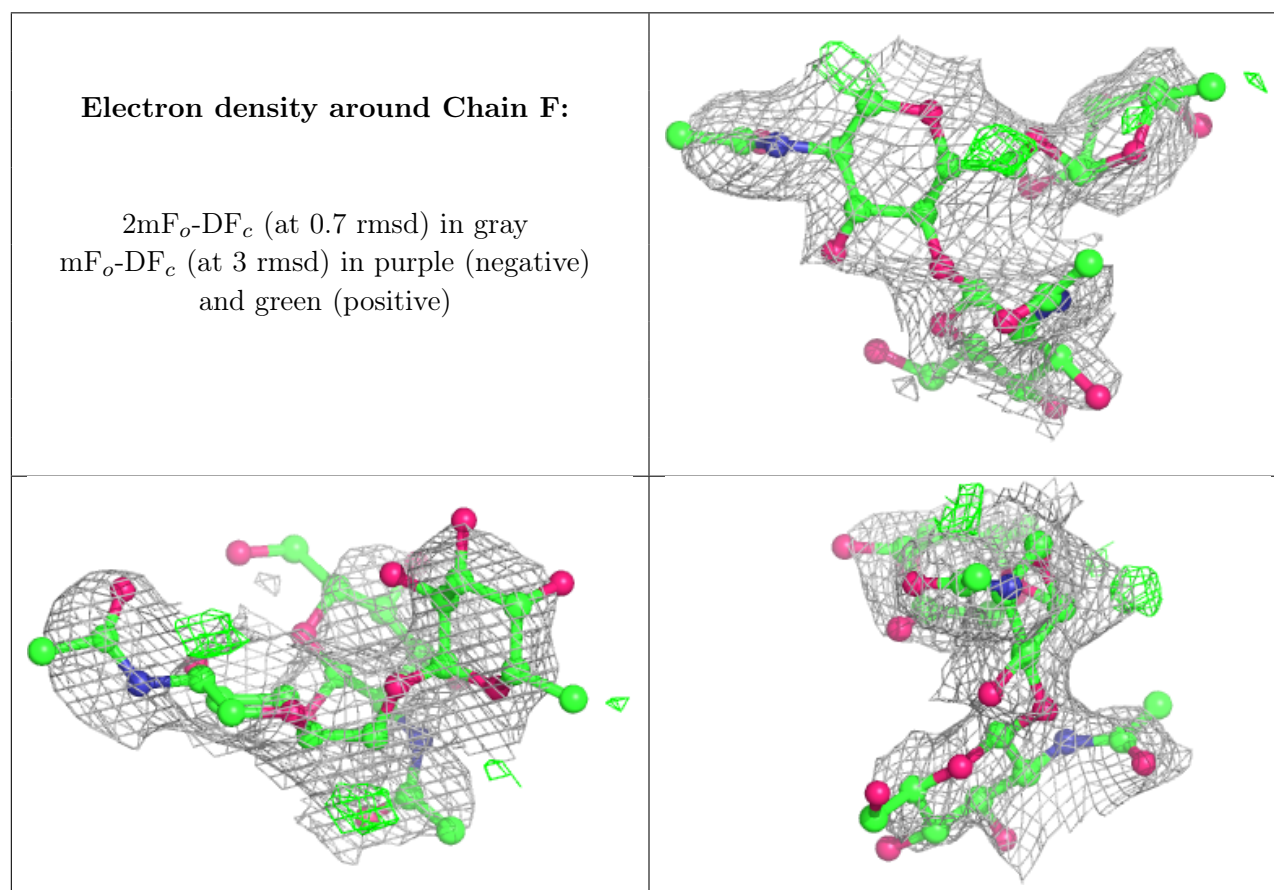
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	FUC	E	2	10/11	0.66	0.24	79,101,107,109	0
3	NAG	F	2	14/15	0.77	0.31	105,113,118,119	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\mathring{A}^2$)	Q<0.9
3	FUC	F	3	10/11	0.85	0.24	110,117,119,120	0
3	NAG	F	1	14/15	0.85	0.18	73,92,108,119	0
4	NAG	E	1	14/15	0.87	0.13	66,89,98,100	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 F:

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

Electron density around Chain E:

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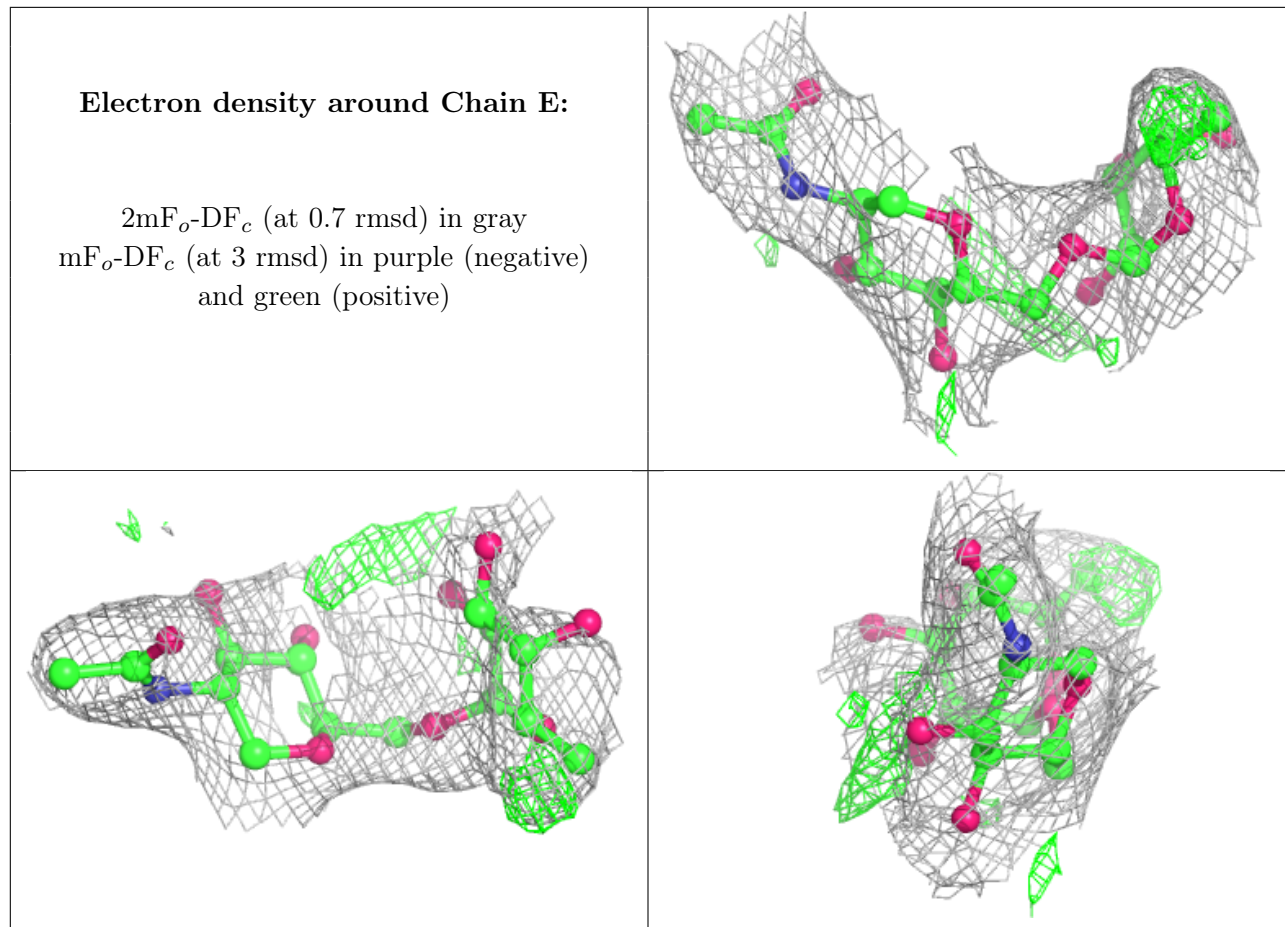
6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.