



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 7, 2020 – 01:52 AM BST

PDB ID : 4BL8  
Title : Crystal structure of full-length human Suppressor of fused (SUFU)  
Authors : Karlstrom, M.; Finta, C.; Cherry, A.L.; Toftgard, R.; Jovine, L.  
Deposited on : 2013-05-02  
Resolution : 3.04 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](mailto: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 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 : 2.13.1  
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) : 2.13.1

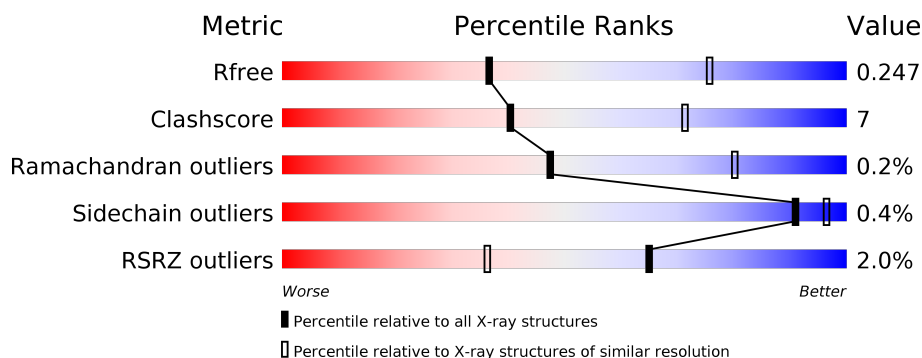
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.04 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))
$R_{free}$	130704	2752 (3.08-3.00)
Clashscore	141614	3096 (3.08-3.00)
Ramachandran outliers	138981	2986 (3.08-3.00)
Sidechain outliers	138945	2988 (3.08-3.00)
RSRZ outliers	127900	2636 (3.08-3.00)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$ . 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	831	<div> <div>%</div> <div> <div></div> <div>75%</div> <div>12%</div> <div>13%</div> </div> </div>
1	B	831	<div> <div>3%</div> <div> <div></div> <div>73%</div> <div>15%</div> <div>12%</div> </div> </div>
2	C	2	<div> <div></div> <div> <div>50%</div> <div>50%</div> </div> </div>
2	D	2	<div> <div></div> <div>100%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 11495 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 MALTOSE-BINDING PERIPLASMIC PROTEIN, SUPPRESSOR OF FUSED HOMOLOG.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	727	Total	C	N	O	S	0	0	0
			5694	3651	946	1079	18			
1	B	734	Total	C	N	O	S	0	0	0
			5755	3689	954	1093	19			

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	expression tag	UNP P0AEX9
A	3	THR	ILE	engineered mutation	UNP P0AEX9
A	360	ALA	GLU	engineered mutation	UNP P0AEX9
A	363	ALA	LYS	engineered mutation	UNP P0AEX9
A	364	ALA	ASP	engineered mutation	UNP P0AEX9
A	368	ASN	ARG	engineered mutation	UNP P0AEX9
A	369	ALA	-	linker	UNP P0AEX9
A	370	ALA	-	linker	UNP P0AEX9
A	371	ALA	-	linker	UNP P0AEX9
A	824	VAL	-	expression tag	UNP Q9UMX1
A	825	GLU	-	expression tag	UNP Q9UMX1
A	826	HIS	-	expression tag	UNP Q9UMX1
A	827	HIS	-	expression tag	UNP Q9UMX1
A	828	HIS	-	expression tag	UNP Q9UMX1
A	829	HIS	-	expression tag	UNP Q9UMX1
A	830	HIS	-	expression tag	UNP Q9UMX1
A	831	HIS	-	expression tag	UNP Q9UMX1
B	1	MET	-	expression tag	UNP P0AEX9
B	3	THR	ILE	engineered mutation	UNP P0AEX9
B	360	ALA	GLU	engineered mutation	UNP P0AEX9
B	363	ALA	LYS	engineered mutation	UNP P0AEX9
B	364	ALA	ASP	engineered mutation	UNP P0AEX9
B	368	ASN	ARG	engineered mutation	UNP P0AEX9
B	369	ALA	-	linker	UNP P0AEX9

*Continued on next page...*

*Continued from previous page...*

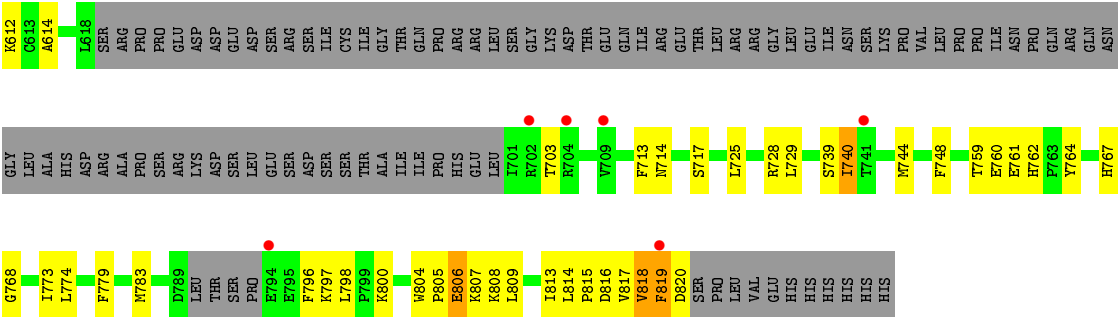
Chain	Residue	Modelled	Actual	Comment	Reference
B	370	ALA	-	linker	UNP P0AEX9
B	371	ALA	-	linker	UNP P0AEX9
B	824	VAL	-	expression tag	UNP Q9UMX1
B	825	GLU	-	expression tag	UNP Q9UMX1
B	826	HIS	-	expression tag	UNP Q9UMX1
B	827	HIS	-	expression tag	UNP Q9UMX1
B	828	HIS	-	expression tag	UNP Q9UMX1
B	829	HIS	-	expression tag	UNP Q9UMX1
B	830	HIS	-	expression tag	UNP Q9UMX1
B	831	HIS	-	expression tag	UNP Q9UMX1

- Molecule 2 is an oligosaccharide called alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose.



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	C	2	Total	C	O	0	0	0
			23	12	11			
2	D	2	Total	C	O	0	0	0
			23	12	11			





● Molecule 2: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose



● Molecule 2: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	97.32Å 99.55Å 192.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.67 – 3.04 48.66 – 3.04	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.67-3.04) 99.9 (48.66-3.04)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.80 (at 3.07Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.200 , 0.246 0.200 , 0.247	Depositor DCC
$R_{free}$ test set	2198 reflections (5.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	90.1	Xtriage
Anisotropy	0.089	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 56.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.024 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	11495	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	105.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.78% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: GLC

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	0.34	0/5845	0.47	0/7949
1	B	0.34	0/5907	0.47	0/8031
All	All	0.34	0/11752	0.47	0/15980

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
1	B	0	6
All	All	0	9

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	767	HIS	Peptide
1	A	768	GLY	Peptide
1	A	806	GLU	Peptide
1	B	370	ALA	Peptide
1	B	492	GLU	Peptide

## 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	5694	0	5558	68	0
1	B	5755	0	5615	97	0
2	C	23	0	21	0	0
2	D	23	0	21	0	0
All	All	11495	0	11215	164	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 7.

The worst 5 of 164 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:759:THR:O	1:A:760:GLU:HB3	1.80	0.79
1:A:374:LEU:HD21	1:A:378:TYR:CE2	2.17	0.78
1:B:759:THR:O	1:B:760:GLU:HB3	1.82	0.78
1:A:818:VAL:O	1:A:818:VAL:HG12	1.90	0.70
1:B:376:ALA:HB1	1:B:528:VAL:HG11	1.73	0.68

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	721/831 (87%)	700 (97%)	20 (3%)	1 (0%)	51 84
1	B	728/831 (88%)	708 (97%)	18 (2%)	2 (0%)	41 74

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1449/1662 (87%)	1408 (97%)	38 (3%)	3 (0%)	47 80

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	740	ILE
1	B	740	ILE
1	B	819	PHE

### 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	601/698 (86%)	600 (100%)	1 (0%)	93 98
1	B	608/698 (87%)	604 (99%)	4 (1%)	84 93
All	All	1209/1396 (87%)	1204 (100%)	5 (0%)	91 96

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

Mol	Chain	Res	Type
1	A	234	SER
1	B	171	LYS
1	B	345	ARG
1	B	380	GLU
1	B	491	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 ⓘ

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	$\# Z  > 2$
2	GLC	C	1	2	12,12,12	0.56	0	17,17,17	0.61	0
2	GLC	C	2	2	11,11,12	0.64	0	15,15,17	0.98	1 (6%)
2	GLC	D	1	2	12,12,12	0.51	0	17,17,17	0.63	0
2	GLC	D	2	2	11,11,12	0.69	0	15,15,17	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
2	GLC	C	1	2	-	0/2/22/22	0/1/1/1
2	GLC	C	2	2	-	0/2/19/22	0/1/1/1
2	GLC	D	1	2	-	0/2/22/22	0/1/1/1
2	GLC	D	2	2	-	0/2/19/22	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(°)	Ideal(°)
2	C	2	GLC	C2-C3-C4	-2.05	107.34	110.89

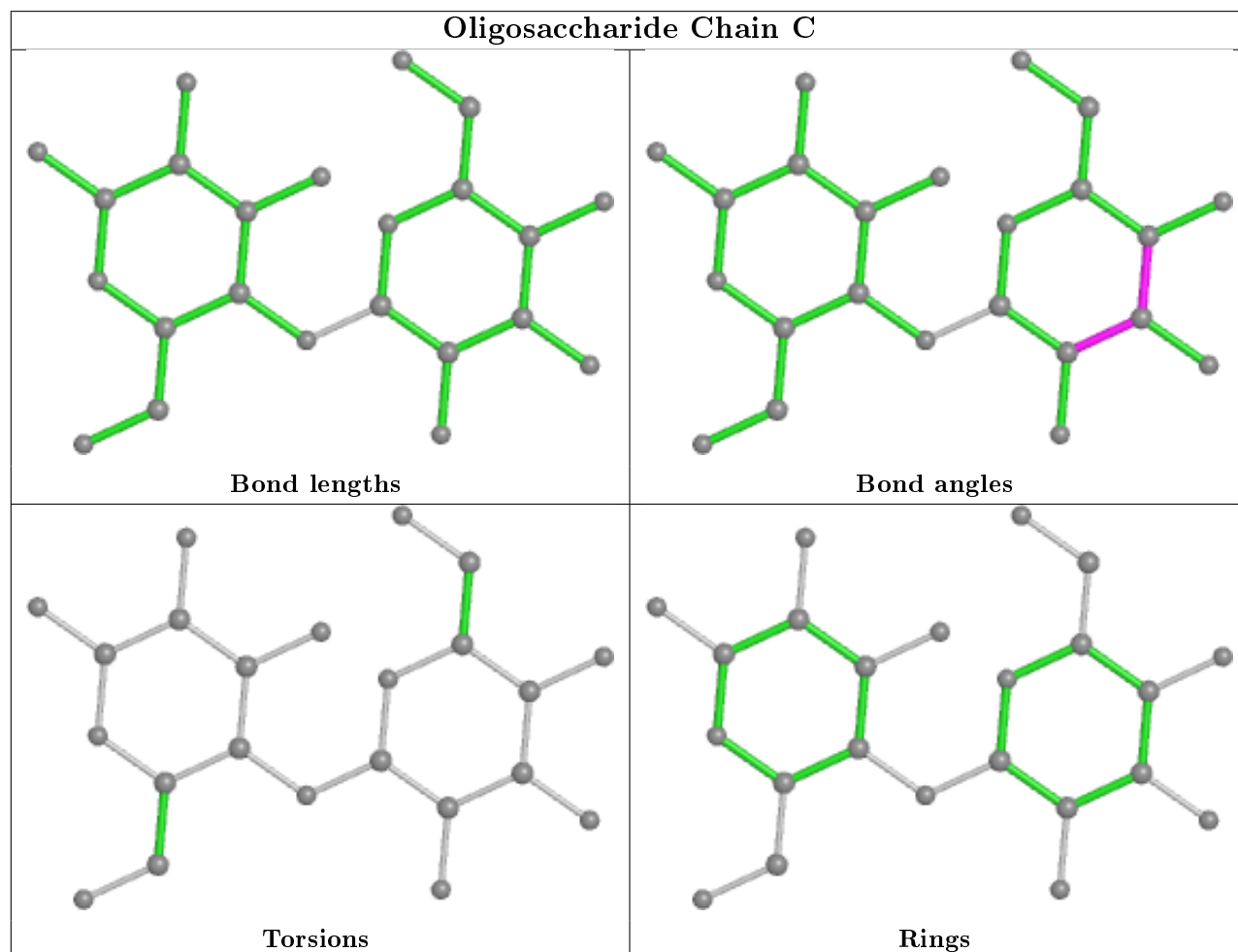
There are no chirality outliers.

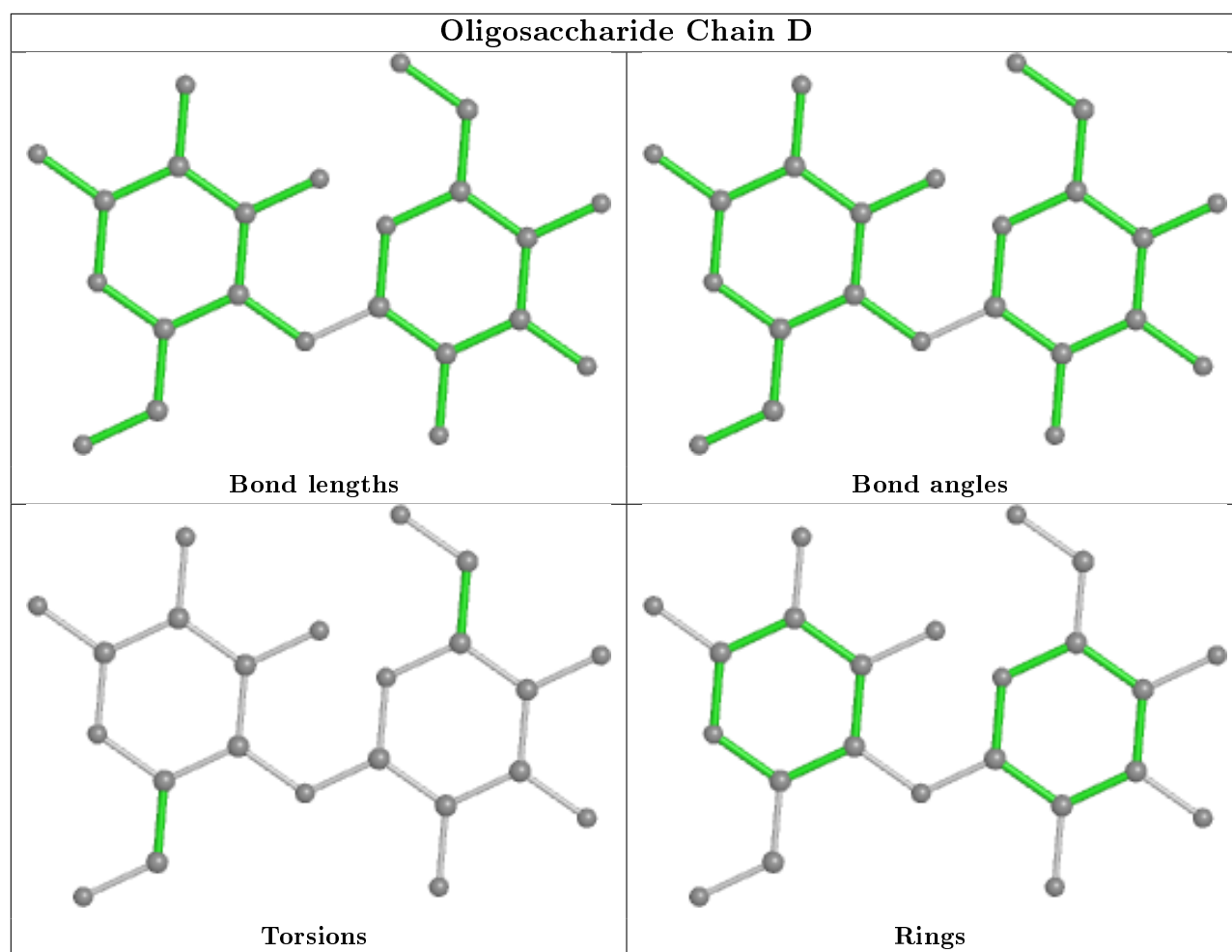
There are no torsion outliers.

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.





## 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	727/831 (87%)	-0.19	6 (0%) 86 65	46, 80, 135, 201	0
1	B	734/831 (88%)	0.10	23 (3%) 49 22	67, 118, 189, 248	0
All	All	1461/1662 (87%)	-0.05	29 (1%) 65 36	46, 99, 174, 248	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	144	GLY	4.6
1	B	32	THR	3.4
1	B	421	SER	2.9
1	B	362	LEU	2.9
1	A	494	THR	2.8

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

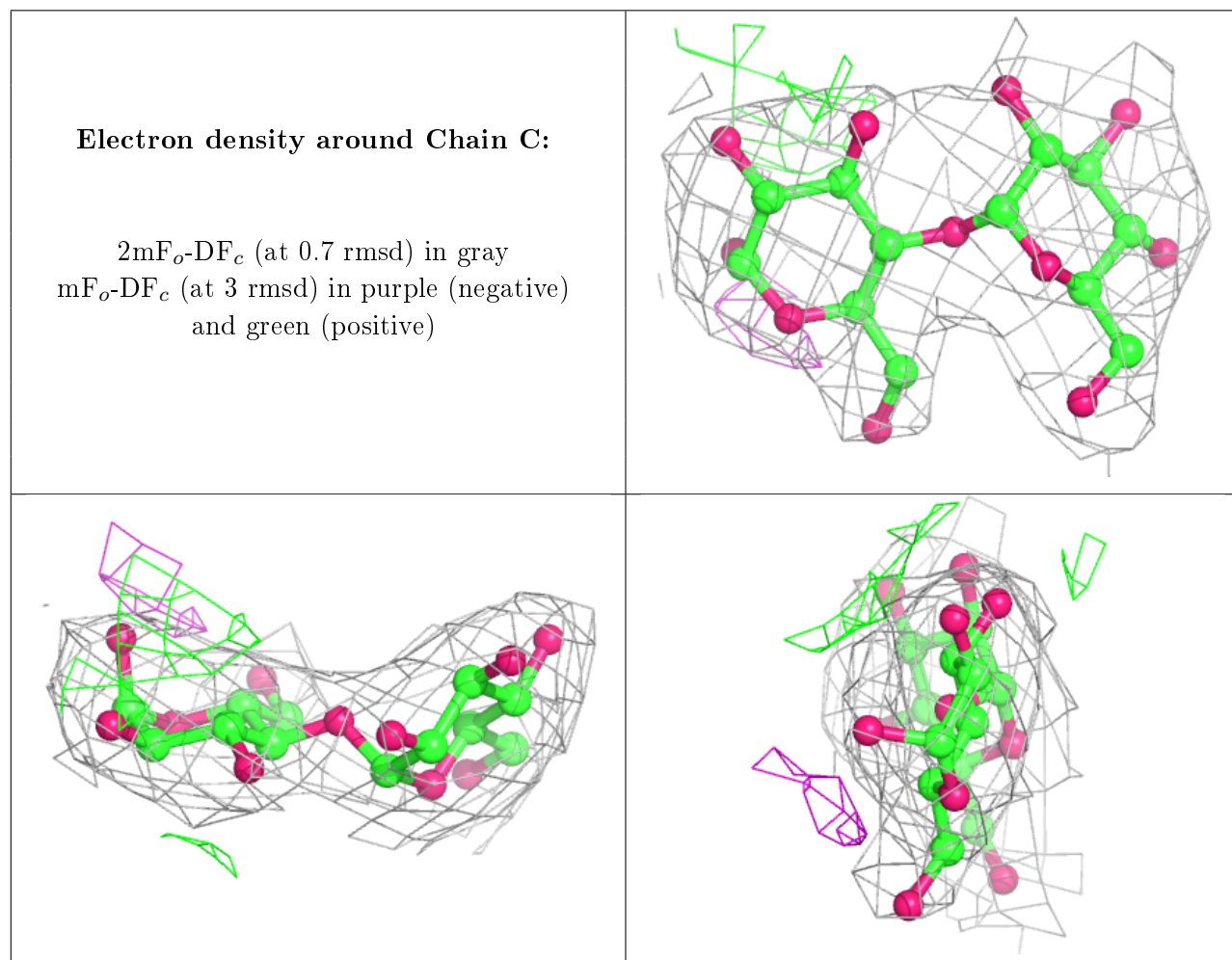
There are no non-standard protein/DNA/RNA residues in this entry.

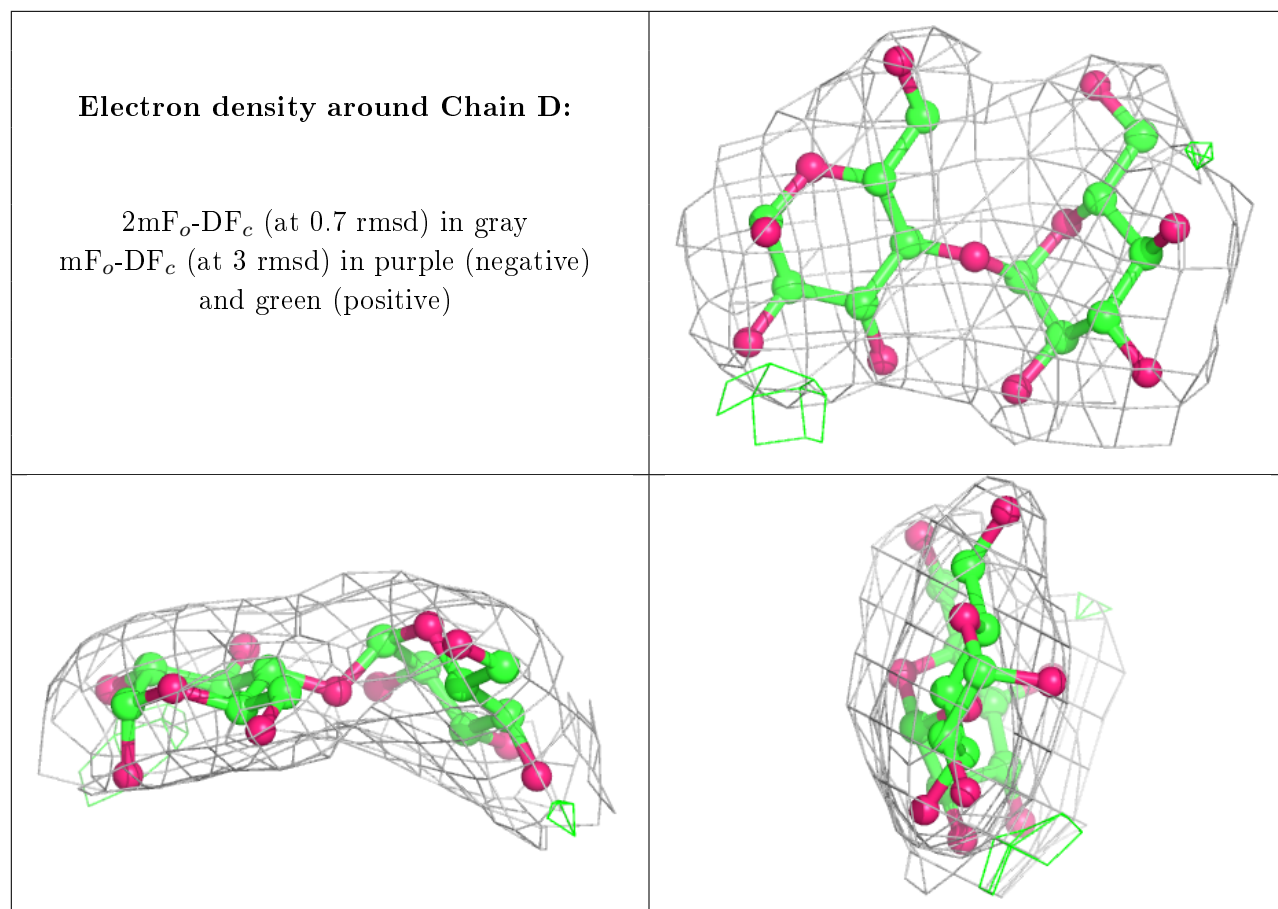
### 6.3 Carbohydrates [i](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
2	GLC	D	2	11/12	0.94	0.22	75,80,86,86	0
2	GLC	C	1	12/12	0.96	0.20	45,59,66,73	0
2	GLC	D	1	12/12	0.98	0.17	81,90,95,97	0
2	GLC	C	2	11/12	0.98	0.15	48,61,64,65	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.





## 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.