



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 4, 2024 – 06:31 am GMT

PDB ID : 5G62
Title : S.pneumoniae ABC-transporter substrate binding protein FusA EF-hand mutant in complex with fructo-nystose
Authors : Culurgioni, S.; Harris, G.; Singh, A.K.; King, S.J.; Walsh, M.A.
Deposited on : 2016-06-10
Resolution : 1.99 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 ⓘ 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 : 1.8.4, 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

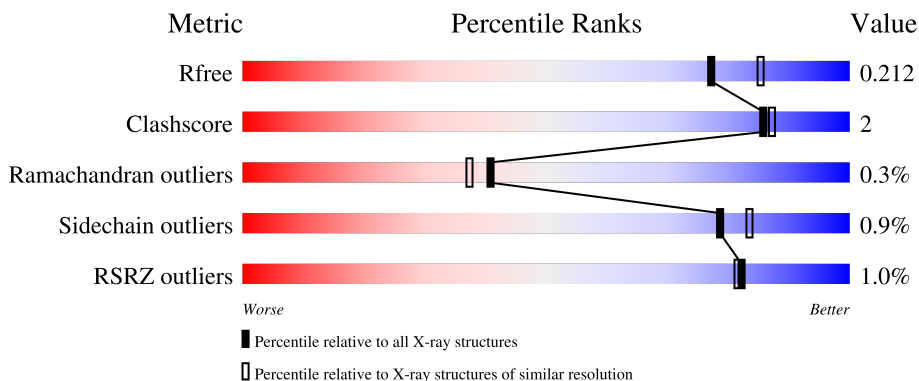
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.99 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 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 $\leq 5\%$. 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	494	96% (Poor fit: 0%, 0 types: 96%, 1 type: 3%, 2 types: 1%, 3 types: 0%)
1	B	494	2% (Poor fit: 2%, 0 types: 96%, 1 type: 3%, 2 types: 1%, 3 types: 0%)
2	C	5	60% (Poor fit: 0%, 0 types: 60%, 1 type: 40%)
2	D	5	60% (Poor fit: 0%, 0 types: 60%, 1 type: 40%)

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	ZEE	D	1	X	-	X	X
2	UEA	D	2	-	-	X	-

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 8913 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.

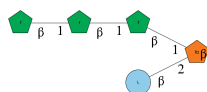
- Molecule 1 is a protein called ABC TRANSPORTER, SUBSTRATE-BINDING PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	491	3964	2522	653	777	12	0	3	0
1	B	491	3947	2512	653	770	12	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	45	GLY	-	expression tag	UNP A0A0H2URD
A	46	PRO	-	expression tag	UNP A0A0H2URD
A	223	ALA	ASP	engineered mutation	UNP A0A0H2URD
A	224	ALA	GLU	engineered mutation	UNP A0A0H2URD
B	45	GLY	-	expression tag	UNP A0A0H2URD
B	46	PRO	-	expression tag	UNP A0A0H2URD
B	223	ALA	ASP	engineered mutation	UNP A0A0H2URD
B	224	ALA	GLU	engineered mutation	UNP A0A0H2URD

- Molecule 2 is an oligosaccharide called beta-D-fructofuranose-(2-1)-beta-D-fructofuranose-(2-1)-beta-D-fructofuranose-(2-1)-beta-D-sorbofuranose-(2-1)-beta-L-talopyranose.



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
			Total	C	O			
2	C	5	56	30	26	0	0	0
2	D	5	56	30	26	0	0	0

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	2	Total Ca 2 2	0	0
3	B	1	Total Ca 1 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	505	Total O 505 505	0	0
4	B	382	Total O 382 382	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: ABC TRANSPORTER, SUBSTRATE-BINDING PROTEIN

Chain A:  96%



- Molecule 1: ABC TRANSPORTER, SUBSTRATE-BINDING PROTEIN

Chain B:  96%



- Molecule 2: beta-D-fructofuranose-(2-1)-beta-D-fructofuranose-(2-1)-beta-D-fructofuranose-(2-1)-beta-D-sorbofuranose-(2-1)-beta-L-talopyranose

Chain C:  60% 40%



- Molecule 2: beta-D-fructofuranose-(2-1)-beta-D-fructofuranose-(2-1)-beta-D-fructofuranose-(2-1)-beta-D-sorbofuranose-(2-1)-beta-L-talopyranose

Chain D:  60% 40%



4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	122.58Å 122.58Å 169.47Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	77.17 – 1.99 86.68 – 1.99	Depositor EDS
% Data completeness (in resolution range)	100.0 (77.17-1.99) 100.0 (86.68-1.99)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.65 (at 1.98Å)	Xtrriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.171 , 0.208 0.176 , 0.212	Depositor DCC
R_{free} test set	4435 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å ²)	29.2	Xtrriage
Anisotropy	0.315	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 57.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	8913	wwPDB-VP
Average B, all atoms (Å ²)	35.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.34% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: ZEE, UEA, FRU, CA

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	0.60	0/4068	0.68	2/5508 (0.0%)
1	B	0.59	1/4043 (0.0%)	0.66	1/5475 (0.0%)
All	All	0.60	1/8111 (0.0%)	0.67	3/10983 (0.0%)

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

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	221	GLU	CG-CD	5.87	1.60	1.51

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	419	ARG	NE-CZ-NH1	7.89	124.25	120.30
1	A	419	ARG	NE-CZ-NH2	-6.83	116.88	120.30
1	B	218	GLY	N-CA-C	5.50	126.85	113.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	217	ASN	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	3964	0	3816	8	0
1	B	3947	0	3803	12	0
2	C	56	0	30	4	0
2	D	56	0	31	10	0
3	A	2	0	0	0	0
3	B	1	0	0	0	0
4	A	505	0	0	3	1
4	B	382	0	0	5	0
All	All	8913	0	7680	34	1

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)
2:D:1:ZEE:O4	2:D:2:UEA:C3	2.29	0.80
2:D:1:ZEE:C3	2:D:2:UEA:O3	2.30	0.80
2:C:1:ZEE:O4	2:C:2:UEA:C3	2.30	0.79
2:C:1:ZEE:O5	2:C:2:UEA:C3	2.29	0.75
2:D:1:ZEE:O5	2:D:2:UEA:C3	2.29	0.72

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:A:2065:HOH:O	4:A:2187:HOH:O[4_555]	2.07	0.13

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	491/494 (99%)	481 (98%)	10 (2%)	0	100	100
1	B	489/494 (99%)	479 (98%)	7 (1%)	3 (1%)	25	19
All	All	980/988 (99%)	960 (98%)	17 (2%)	3 (0%)	41	37

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	221	GLU
1	B	217	ASN
1	B	222	ALA

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	427/426 (100%)	423 (99%)	4 (1%)	78	83
1	B	424/426 (100%)	420 (99%)	4 (1%)	78	83
All	All	851/852 (100%)	843 (99%)	8 (1%)	78	83

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	470	PHE
1	B	406	LYS
1	B	217	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	486	MET
1	B	294	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	362	ASN
1	B	177	HIS
1	B	362	ASN
1	B	340	HIS
1	A	340	HIS

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](#)

10 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	ZEE	C	1	2	11,11,12	2.05	4 (36%)	15,15,17	2.48	5 (33%)
2	UEA	C	2	2	11,12,12	4.06	2 (18%)	10,18,18	1.30	1 (10%)
2	FRU	C	3	2	11,11,12	4.17	4 (36%)	15,15,18	1.06	1 (6%)
2	FRU	C	4	2	11,11,12	4.34	5 (45%)	15,15,18	1.03	1 (6%)
2	FRU	C	5	2	11,11,12	4.06	4 (36%)	15,15,18	0.88	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	ZEE	D	1	2	11,11,12	1.45	1 (9%)	15,15,17	2.13	6 (40%)
2	UEA	D	2	2	11,12,12	3.70	3 (27%)	10,18,18	1.44	1 (10%)
2	FRU	D	3	2	11,11,12	3.96	4 (36%)	15,15,18	1.22	1 (6%)
2	FRU	D	4	2	11,11,12	4.35	5 (45%)	15,15,18	1.10	0
2	FRU	D	5	2	11,11,12	4.02	5 (45%)	15,15,18	0.85	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	ZEE	C	1	2	-	0/2/19/22	0/1/1/1
2	UEA	C	2	2	-	0/5/24/24	0/1/1/1
2	FRU	C	3	2	-	0/4/20/24	0/1/1/1
2	FRU	C	4	2	-	0/4/20/24	0/1/1/1
2	FRU	C	5	2	-	0/4/20/24	0/1/1/1
2	ZEE	D	1	2	1/1/4/5	1/2/19/22	0/1/1/1
2	UEA	D	2	2	-	3/5/24/24	0/1/1/1
2	FRU	D	3	2	-	0/4/20/24	0/1/1/1
2	FRU	D	4	2	-	0/4/20/24	0/1/1/1
2	FRU	D	5	2	-	0/4/20/24	0/1/1/1

The worst 5 of 37 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	2	UEA	O5-C2	10.64	1.59	1.43
2	C	3	FRU	C4-C3	-9.75	1.26	1.53
2	D	3	FRU	C4-C3	-9.73	1.26	1.53
2	D	4	FRU	O5-C2	-9.00	1.24	1.45
2	C	3	FRU	O5-C5	-8.89	1.25	1.45

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	1	ZEE	C1-C2-C3	6.12	117.18	109.67
2	D	1	ZEE	O2-C2-C1	-4.08	100.80	109.15
2	C	1	ZEE	O5-C1-C2	3.80	116.64	110.77
2	C	1	ZEE	O4-C4-C5	-3.41	100.83	109.30
2	D	1	ZEE	C1-O5-C5	3.32	116.69	112.19

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	D	1	ZEE	C3

All (4) torsion outliers are listed below:

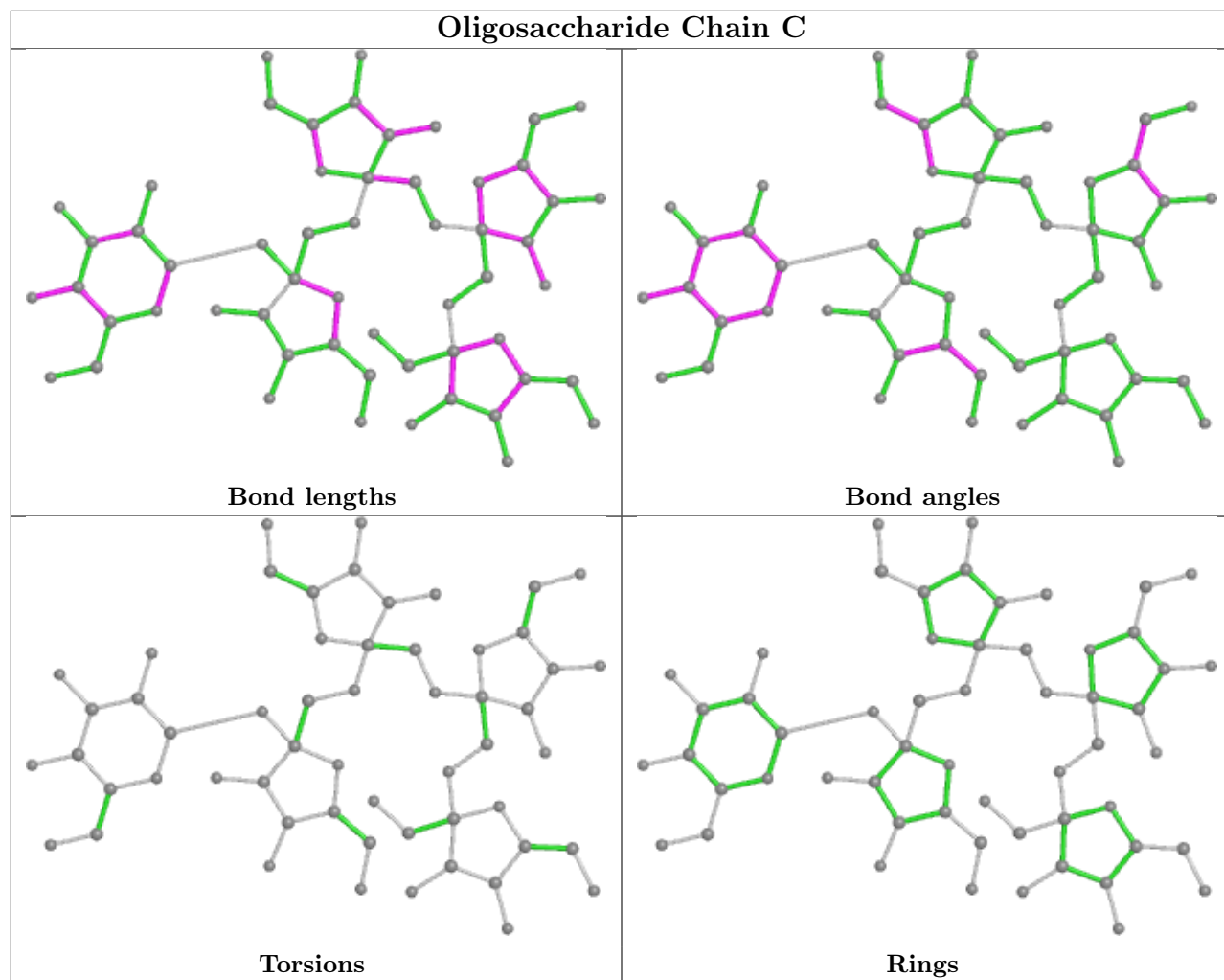
Mol	Chain	Res	Type	Atoms
2	D	2	UEA	O1-C1-C2-O2
2	D	2	UEA	O1-C1-C2-C3
2	D	2	UEA	O1-C1-C2-O5
2	D	1	ZEE	O5-C5-C6-O6

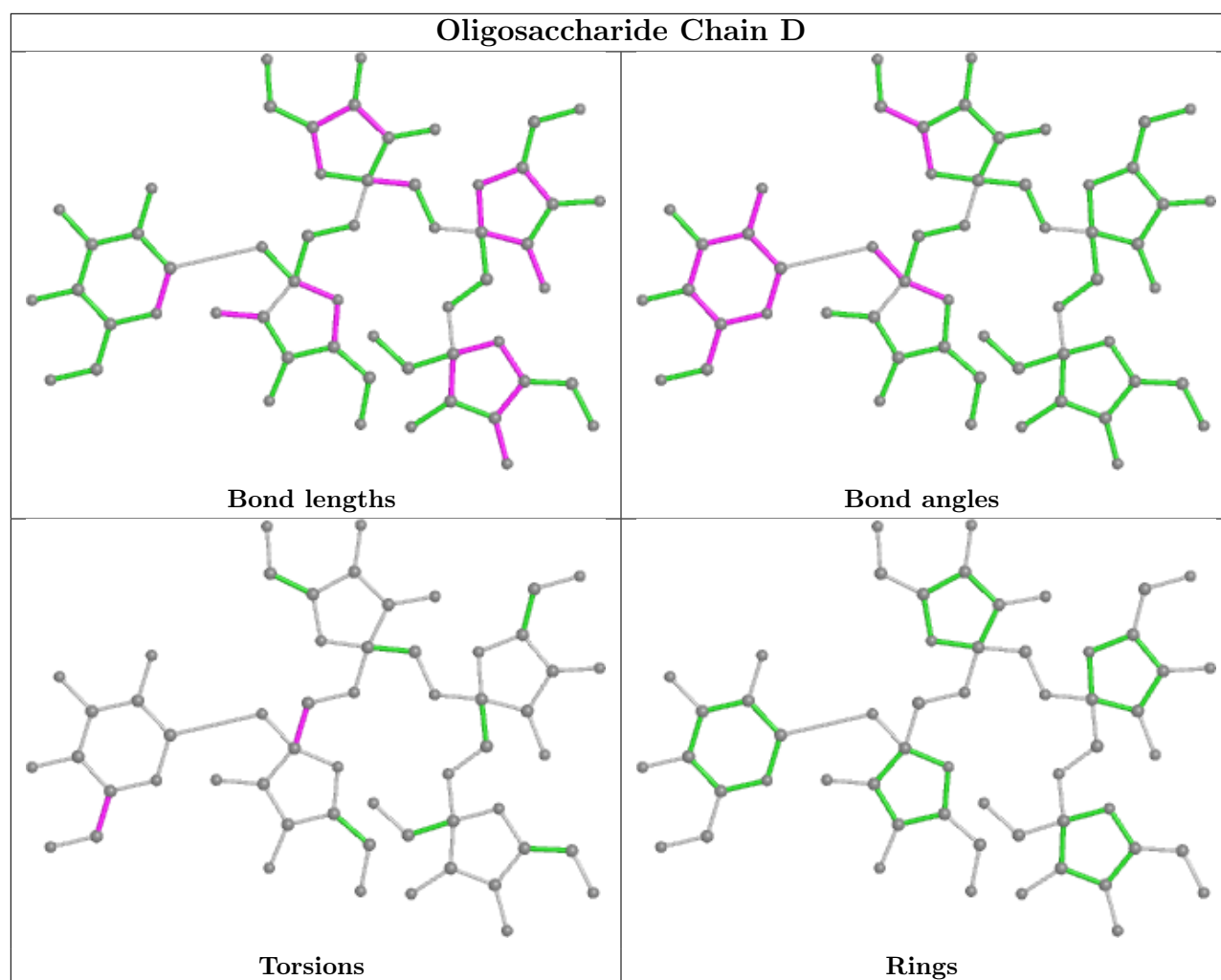
There are no ring outliers.

4 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	1	ZEE	3	0
2	D	2	UEA	10	0
2	D	1	ZEE	9	0
2	C	2	UEA	4	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.





5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	536:ASN	C	537:LYS	N	3.42

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, 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	491/494 (99%)	-0.36	1 (0%) 95 94	20, 30, 48, 114	0
1	B	491/494 (99%)	-0.22	9 (1%) 68 66	20, 33, 58, 150	0
All	All	982/988 (99%)	-0.29	10 (1%) 82 81	20, 31, 52, 150	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	222	ALA	8.1
1	A	537	LYS	5.4
1	B	221	GLU	4.8
1	B	218	GLY	4.6
1	B	219	ASN	4.4

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.

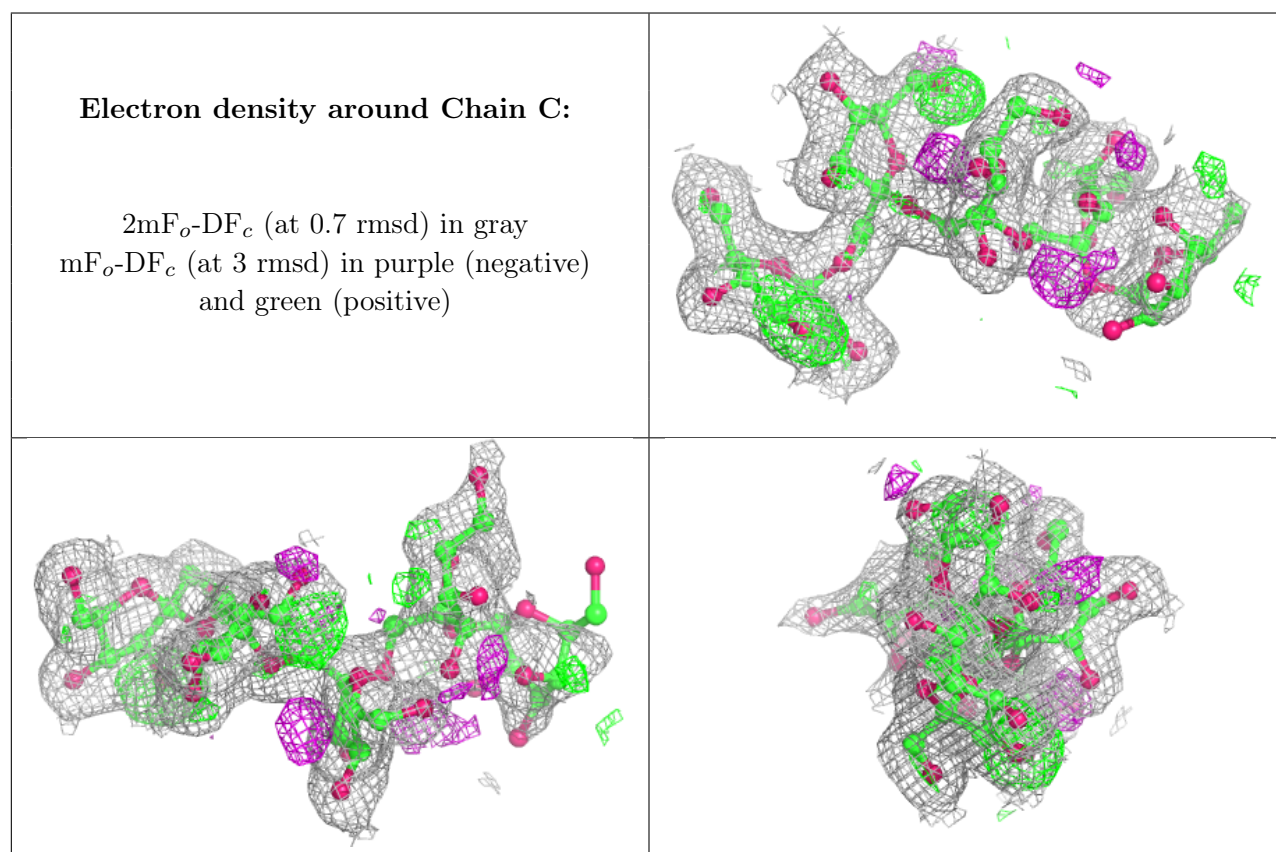
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	ZEE	D	1	11/12	0.57	0.45	100,138,145,149	0
2	ZEE	C	1	11/12	0.75	0.35	100,138,145,149	0
2	FRU	C	5	11/12	0.83	0.15	31,34,38,66	0
2	UEA	D	2	12/12	0.85	0.23	40,85,95,102	0
2	UEA	C	2	12/12	0.87	0.16	40,85,95,102	0

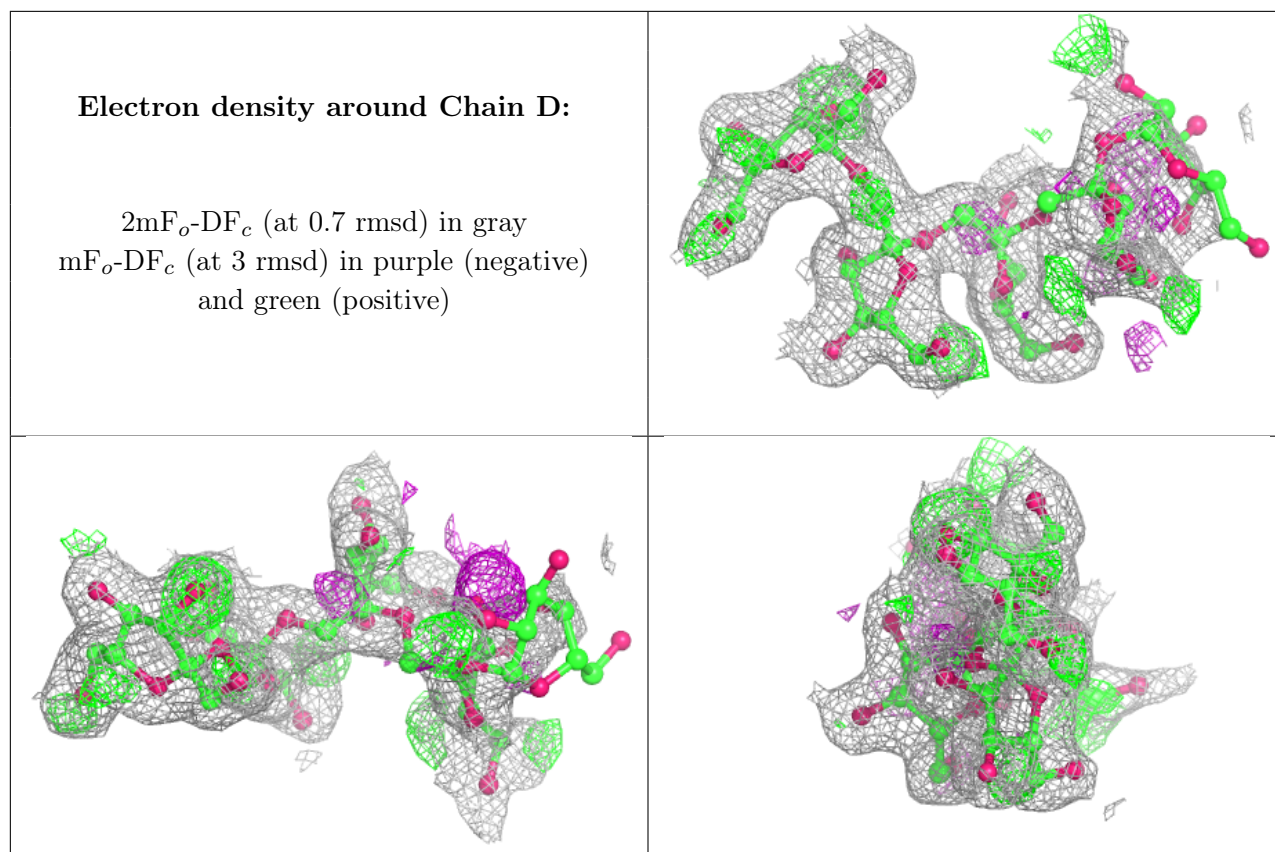
Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	FRU	D	3	11/12	0.87	0.11	30,38,42,45	0
2	FRU	D	5	11/12	0.87	0.14	31,34,38,66	0
2	FRU	C	4	11/12	0.91	0.13	32,36,42,48	0
2	FRU	D	4	11/12	0.94	0.10	32,36,42,48	0
2	FRU	C	3	11/12	0.94	0.11	30,38,42,45	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.





6.4 Ligands [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	CA	A	546	1/1	0.99	0.07	46,46,46,46	0
3	CA	A	545	1/1	1.00	0.18	31,31,31,31	0
3	CA	B	545	1/1	1.00	0.14	33,33,33,33	0

6.5 Other polymers [i](#)

There are no such residues in this entry.