wwPDB X-ray Structure Validation Summary Report (i)

Oct 30, 2023 – 06:52 PM JST

PDB ID	:	4XTC
Title	:	Crystal structure of bacterial alginate ABC transporter in complex with alginate pentasaccharide-bound periplasmic protein
Authors	:	Kaneko, A.; Maruyama, Y.; Mizuno, N.; Baba, S.; Kumasaka, T.; Mikami, B.; Murata, K.; Hashimoto, W.
Deposited on	:	2015-01-23
Resolution	:	3.60 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 (i)) were used in the production of this report:

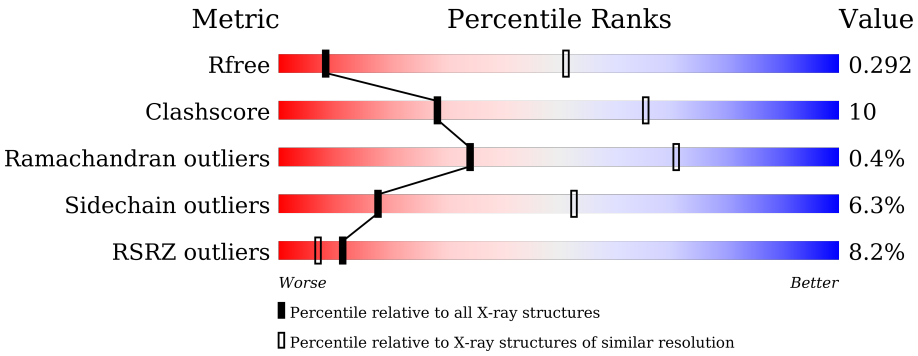
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.60 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	1257 (3.70-3.50)
Clashscore	141614	1353 (3.70-3.50)
Ramachandran outliers	138981	1307 (3.70-3.50)
Sidechain outliers	138945	1307 (3.70-3.50)
RSRZ outliers	127900	1161 (3.70-3.50)

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	M	301	6% 70% 24%
2	N	305	6% 72% 19% 6%
3	S	363	7% 68% 28%
3	T	363	8% 69% 28%
4	Q	516	11% 71% 23% 5%
5	A	5	60% 40%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 14264 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 AlgM1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	M	288	2323	1554	369	390	10	0	0	0

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	?	-	THR	deletion	UNP Q9KWT8
M	?	-	SER	deletion	UNP Q9KWT8
M	?	-	ALA	deletion	UNP Q9KWT8
M	?	-	THR	deletion	UNP Q9KWT8
M	?	-	LYS	deletion	UNP Q9KWT8
M	?	-	ALA	deletion	UNP Q9KWT8
M	?	-	GLN	deletion	UNP Q9KWT8
M	?	-	SER	deletion	UNP Q9KWT8
M	?	-	ILE	deletion	UNP Q9KWT8
M	?	-	PRO	deletion	UNP Q9KWT8
M	?	-	LEU	deletion	UNP Q9KWT8
M	?	-	PRO	deletion	UNP Q9KWT8
M	?	-	ALA	deletion	UNP Q9KWT8
M	?	-	ALA	deletion	UNP Q9KWT8
M	?	-	THR	deletion	UNP Q9KWT8
M	?	-	LEU	deletion	UNP Q9KWT8
M	?	-	ASP	deletion	UNP Q9KWT8
M	?	-	VAL	deletion	UNP Q9KWT8
M	?	-	ARG	deletion	UNP Q9KWT8
M	?	-	SER	deletion	UNP Q9KWT8
M	?	-	LYS	deletion	UNP Q9KWT8
M	?	-	PRO	deletion	UNP Q9KWT8
M	?	-	LEU	deletion	UNP Q9KWT8

- Molecule 2 is a protein called AlgM2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	N	287	Total	C	N	O	S	0	0	0
			2278	1521	362	382	13			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
N	294	LEU	-	expression tag	UNP Q9KWT7
N	295	GLU	-	expression tag	UNP Q9KWT7
N	296	HIS	-	expression tag	UNP Q9KWT7
N	297	HIS	-	expression tag	UNP Q9KWT7
N	298	HIS	-	expression tag	UNP Q9KWT7
N	299	HIS	-	expression tag	UNP Q9KWT7
N	300	HIS	-	expression tag	UNP Q9KWT7
N	301	HIS	-	expression tag	UNP Q9KWT7
N	302	HIS	-	expression tag	UNP Q9KWT7
N	303	HIS	-	expression tag	UNP Q9KWT7
N	304	HIS	-	expression tag	UNP Q9KWT7
N	305	HIS	-	expression tag	UNP Q9KWT7

- Molecule 3 is a protein called AlgS.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	S	363	Total	C	N	O	S	0	0	0
			2777	1745	503	518	11			
3	T	363	Total	C	N	O	S	0	0	0
			2777	1745	503	518	11			

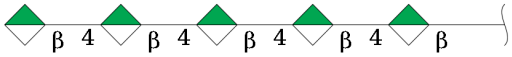
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
S	160	GLN	GLU	engineered mutation	UNP Q9KWT9
T	160	GLN	GLU	engineered mutation	UNP Q9KWT9

- Molecule 4 is a protein called AlgQ2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	Q	492	Total	C	N	O	S	0	0	0
			4048	2604	693	735	16			

- Molecule 5 is an oligosaccharide called beta-D-mannopyranuronic acid-(1-4)-beta-D-manno pyranuronic acid-(1-4)-beta-D-mannopyranuronic acid-(1-4)-beta-D-mannopyranuronic aci d-(1-4)-beta-D-mannopyranuronic acid.



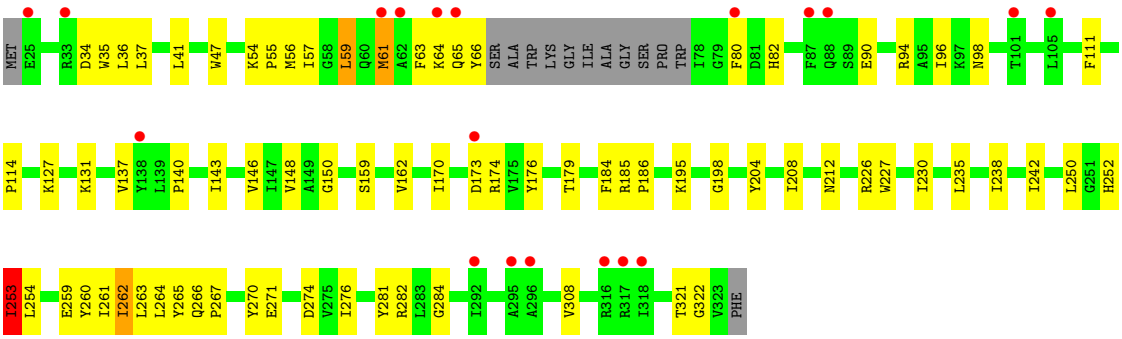
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
5	A	5	Total	C	O	0	0	0
			61	30	31			

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.

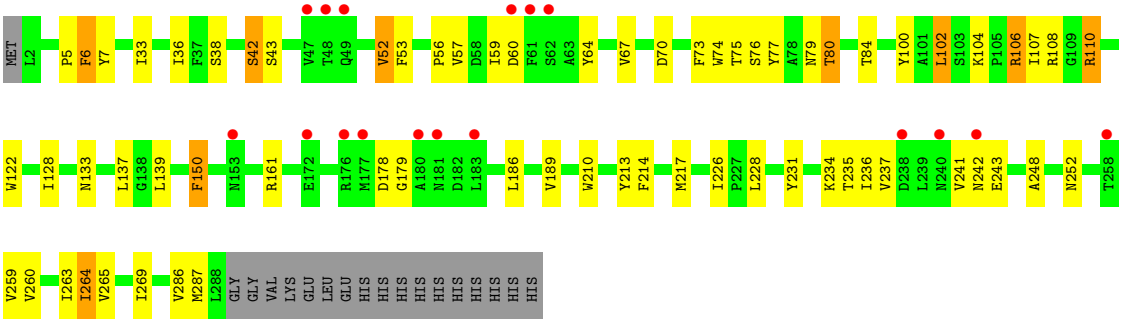
- Molecule 1: AlgM1

Chain M:



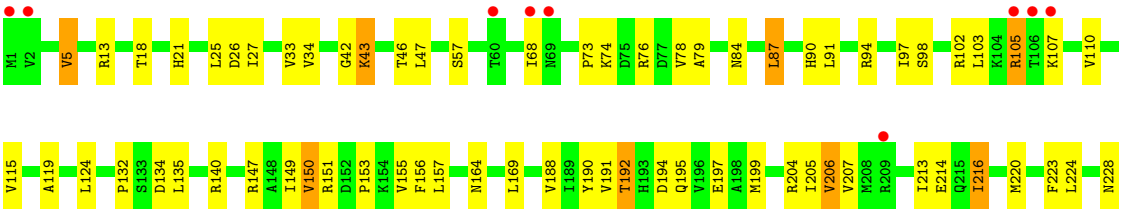
- Molecule 2: AlgM2

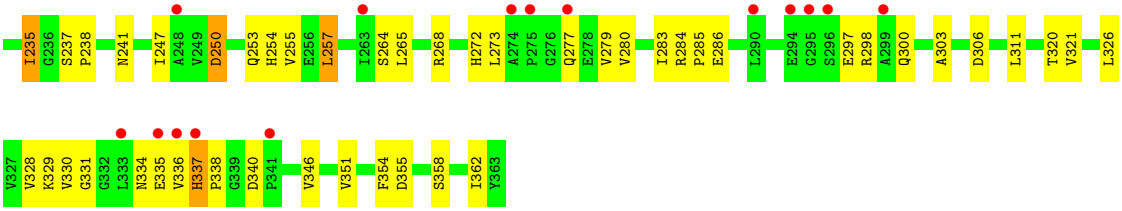
Chain N:



- Molecule 3: AlgS

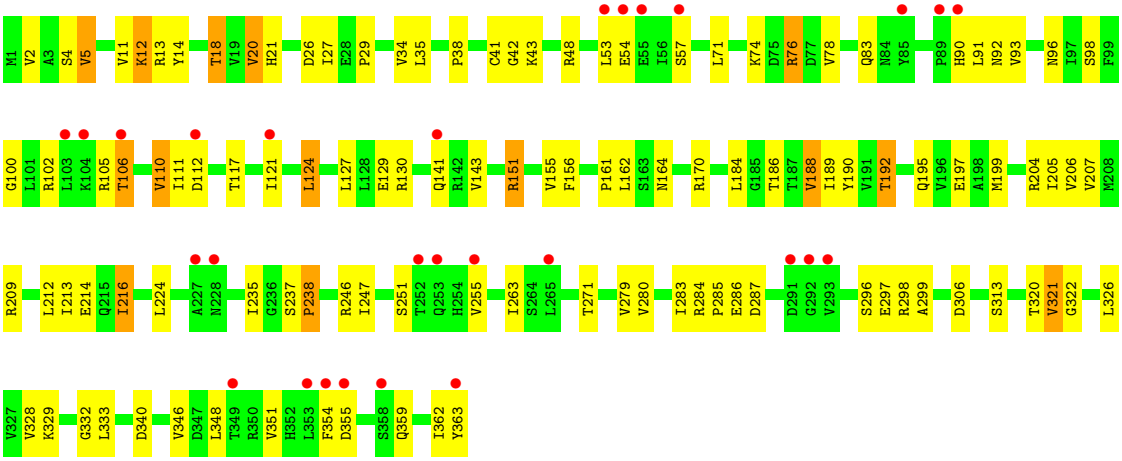
Chain S:





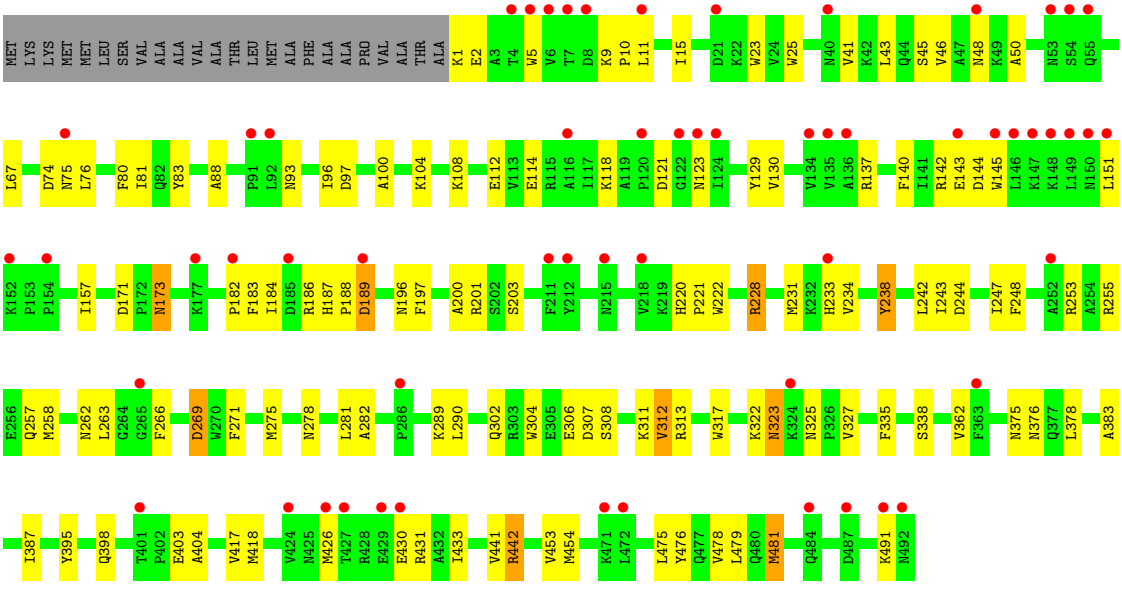
• Molecule 3: AlgS

Chain T:



• Molecule 4: AlgQ2

Chain Q:



• Molecule 5: beta-D-mannopyranuronic acid-(1-4)-beta-D-mannopyranuronic acid-(1-4)-beta-D-mannopyranuronic acid-(1-4)-beta-D-mannopyranuronic acid-(1-4)-beta-D-mannopyranuronic acid

Chain A:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	71.61Å 132.96Å 272.97Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	37.42 – 3.60 49.40 – 3.58	Depositor EDS
% Data completeness (in resolution range)	95.2 (37.42-3.60) 95.1 (49.40-3.58)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.45 (at 3.57Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.238 , 0.288 0.246 , 0.292	Depositor DCC
R_{free} test set	1516 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	108.1	Xtriage
Anisotropy	0.214	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 78.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	14264	wwPDB-VP
Average B, all atoms (Å^2)	122.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.89% 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: BEM

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	M	0.23	0/2384	0.42	0/3250
2	N	0.24	0/2337	0.42	0/3178
3	S	0.22	0/2822	0.48	1/3826 (0.0%)
3	T	0.21	0/2822	0.46	0/3826
4	Q	0.21	0/4168	0.37	0/5640
All	All	0.22	0/14533	0.43	1/19720 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	S	257	LEU	CA-CB-CG	5.34	127.58	115.30

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	M	2323	0	2412	46	0
2	N	2278	0	2347	51	0
3	S	2777	0	2854	60	0
3	T	2777	0	2854	67	0
4	Q	4048	0	3920	78	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	61	0	33	2	0
All	All	14264	0	14420	280	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 10.

The worst 5 of 280 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)
3:S:34:VAL:HG22	3:S:190:TYR:HB3	1.63	0.79
3:T:34:VAL:HG22	3:T:190:TYR:HB3	1.63	0.79
3:S:235:ILE:HD11	3:S:284:ARG:HH12	1.50	0.76
3:S:192:THR:HG22	3:S:194:ASP:H	1.55	0.72
4:Q:201:ARG:NH1	4:Q:418:MET:SD	2.64	0.71

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	M	284/301 (94%)	264 (93%)	19 (7%)	1 (0%)	34	71
2	N	285/305 (93%)	263 (92%)	20 (7%)	2 (1%)	22	61
3	S	361/363 (99%)	336 (93%)	23 (6%)	2 (1%)	25	64
3	T	361/363 (99%)	335 (93%)	24 (7%)	2 (1%)	25	64
4	Q	490/516 (95%)	464 (95%)	25 (5%)	1 (0%)	47	79
All	All	1781/1848 (96%)	1662 (93%)	111 (6%)	8 (0%)	34	71

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	M	253	ILE
3	S	362	ILE
3	T	111	ILE
4	Q	312	VAL
3	T	110	VAL

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	M	251/260 (96%)	238 (95%)	13 (5%)	23	58
2	N	242/258 (94%)	227 (94%)	15 (6%)	18	53
3	S	307/307 (100%)	282 (92%)	25 (8%)	11	43
3	T	307/307 (100%)	284 (92%)	23 (8%)	13	45
4	Q	424/440 (96%)	403 (95%)	21 (5%)	24	59
All	All	1531/1572 (97%)	1434 (94%)	97 (6%)	18	53

5 of 97 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	Q	112	GLU
4	Q	442	ARG
4	Q	151	LEU
4	Q	269	ASP
3	T	12	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
4	Q	375	ASN
4	Q	450	GLN
3	T	359	GLN
3	S	193	HIS
4	Q	55	GLN

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)

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	$\#\|Z\| > 2$
5	BEM	A	1	5	13,13,13	0.72	0	18,19,19	0.69	0
5	BEM	A	2	5	12,12,13	0.68	0	14,17,19	0.69	0
5	BEM	A	3	5	12,12,13	0.66	0	14,17,19	0.77	0
5	BEM	A	4	5	12,12,13	0.67	0	14,17,19	0.63	0
5	BEM	A	5	5	12,12,13	0.70	0	14,17,19	0.69	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
5	BEM	A	1	5	-	0/4/24/24	0/1/1/1
5	BEM	A	2	5	-	0/4/21/24	0/1/1/1
5	BEM	A	3	5	-	0/4/21/24	0/1/1/1
5	BEM	A	4	5	-	0/4/21/24	0/1/1/1
5	BEM	A	5	5	-	1/4/21/24	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

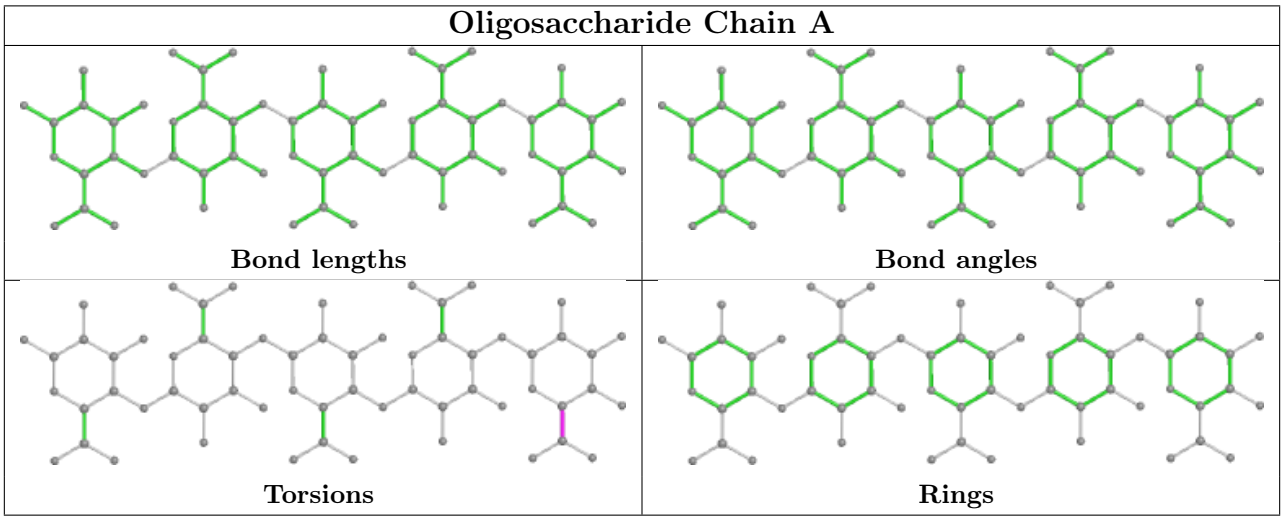
Mol	Chain	Res	Type	Atoms
5	A	5	BEM	C4-C5-C6-O6B

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1	BEM	1	0
5	A	3	BEM	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 oligosaccharide.



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, 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	M	288/301 (95%)	0.29	19 (6%)	18	10	58, 103, 179, 224	0
2	N	287/305 (94%)	0.17	17 (5%)	22	13	60, 99, 148, 213	0
3	S	363/363 (100%)	0.28	24 (6%)	18	10	59, 103, 178, 259	0
3	T	363/363 (100%)	0.44	28 (7%)	13	8	75, 126, 201, 281	0
4	Q	492/516 (95%)	0.64	59 (11%)	4	3	89, 137, 196, 252	0
All	All	1793/1848 (97%)	0.40	147 (8%)	11	7	58, 118, 186, 281	0

The worst 5 of 147 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	M	317	ARG	6.9
4	Q	150	ASN	6.7
3	S	1	MET	6.4
3	S	295	GLY	6.1
3	S	106	THR	5.4

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
5	BEM	A	4	12/13	0.82	0.34	100,116,124,125	0

Continued on next page...

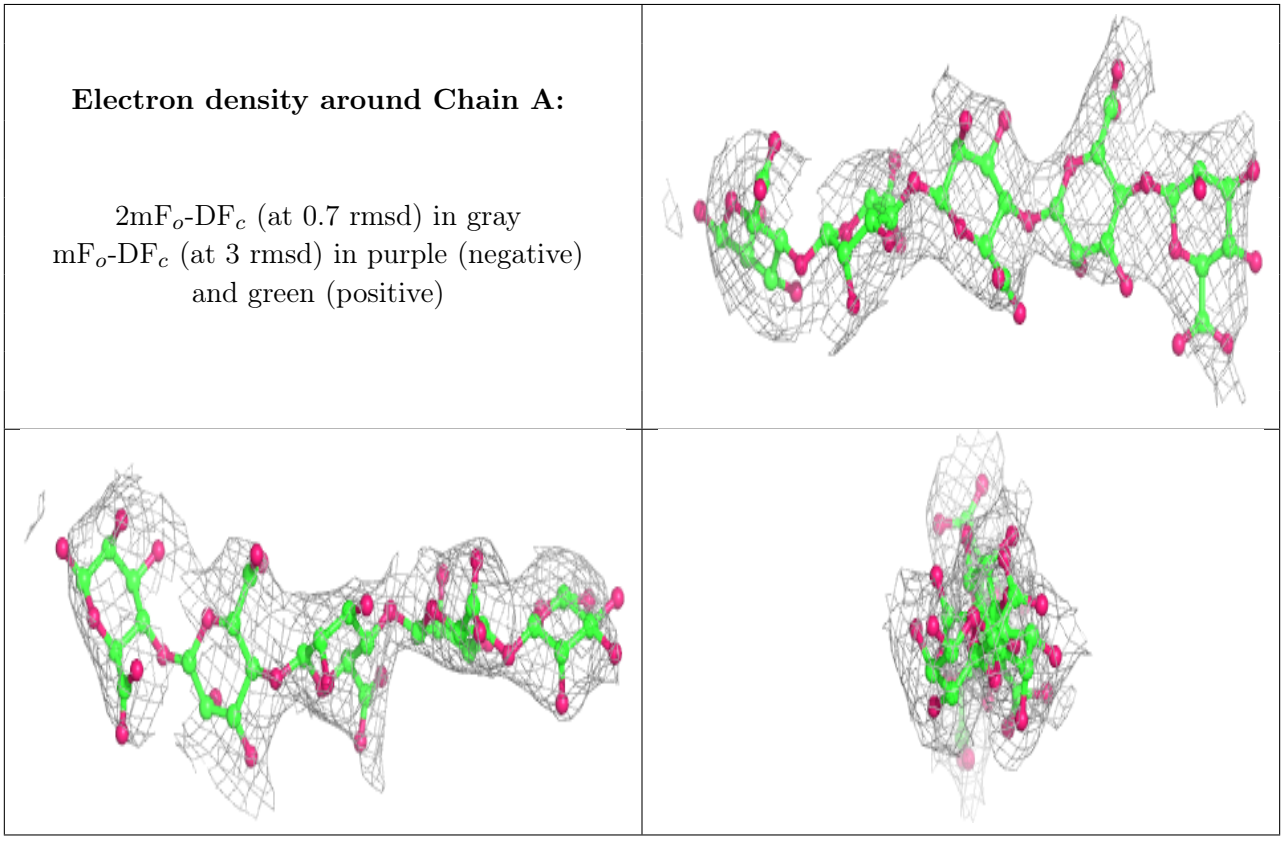
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	BEM	A	2	12/13	0.84	0.80	87,122,132,136	0
5	BEM	A	3	12/13	0.86	0.46	93,114,122,125	0
5	BEM	A	1	13/13	0.88	0.83	119,137,147,156	0
5	BEM	A	5	12/13	0.89	0.29	95,121,135,144	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 A:

$2mF_o$-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.