



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 10, 2026 – 07:44 AM UTC

PDB ID : 4IPN / pdb_00004ipn
Title : The complex structure of 6-phospho-beta-glucosidase BglA-2 with thiocellobiose-6P from *Streptococcus pneumoniae*
Authors : Yu, W.L.; Jiang, Y.L.; Andreas, P.; Cheng, W.; Bai, X.H.; Ren, Y.M.; Thompson, J.; Zhou, C.Z.; Chen, Y.X.
Deposited on : 2013-01-10
Resolution : 2.41 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 : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

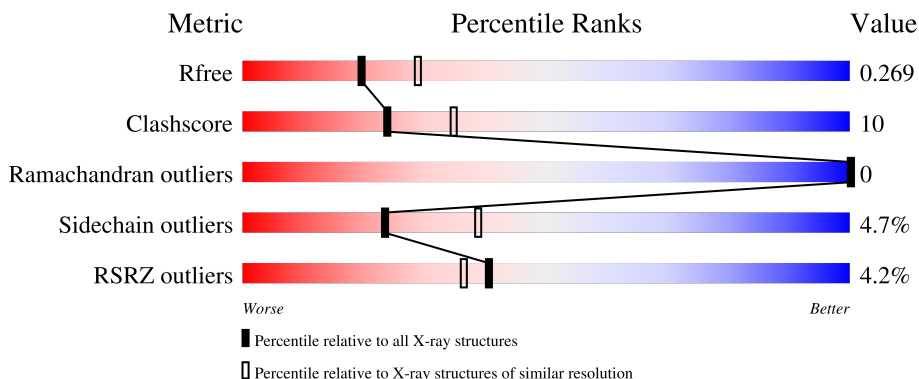
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

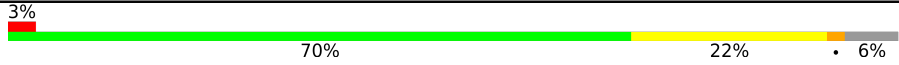



The reported resolution of this entry is 2.41 Å.

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	6062 (2.44-2.40)
Clashscore	190562	6562 (2.44-2.40)
Ramachandran outliers	187476	6481 (2.44-2.40)
Sidechain outliers	187428	6482 (2.44-2.40)
RSRZ outliers	180081	6066 (2.44-2.40)

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	B	487	
1	E	487	
2	A	2	
2	C	2	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 7778 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 6-phospho-beta-glucosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	457	3738	2409	604	712	13	0	0	0
1	E	463	3784	2436	612	723	13	0	0	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-15	HIS	-	expression tag	UNP Q97S37
B	-14	HIS	-	expression tag	UNP Q97S37
B	-13	HIS	-	expression tag	UNP Q97S37
B	-12	HIS	-	expression tag	UNP Q97S37
B	-11	HIS	-	expression tag	UNP Q97S37
B	-10	HIS	-	expression tag	UNP Q97S37
B	-9	SER	-	expression tag	UNP Q97S37
B	-8	SER	-	expression tag	UNP Q97S37
B	-7	GLY	-	expression tag	UNP Q97S37
B	-6	LEU	-	expression tag	UNP Q97S37
B	-5	VAL	-	expression tag	UNP Q97S37
B	-4	PRO	-	expression tag	UNP Q97S37
B	-3	ARG	-	expression tag	UNP Q97S37
B	-2	GLY	-	expression tag	UNP Q97S37
B	-1	SER	-	expression tag	UNP Q97S37
B	0	HIS	-	expression tag	UNP Q97S37
E	-15	HIS	-	expression tag	UNP Q97S37
E	-14	HIS	-	expression tag	UNP Q97S37
E	-13	HIS	-	expression tag	UNP Q97S37
E	-12	HIS	-	expression tag	UNP Q97S37
E	-11	HIS	-	expression tag	UNP Q97S37
E	-10	HIS	-	expression tag	UNP Q97S37
E	-9	SER	-	expression tag	UNP Q97S37
E	-8	SER	-	expression tag	UNP Q97S37
E	-7	GLY	-	expression tag	UNP Q97S37

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	-6	LEU	-	expression tag	UNP Q97S37
E	-5	VAL	-	expression tag	UNP Q97S37
E	-4	PRO	-	expression tag	UNP Q97S37
E	-3	ARG	-	expression tag	UNP Q97S37
E	-2	GLY	-	expression tag	UNP Q97S37
E	-1	SER	-	expression tag	UNP Q97S37
E	0	HIS	-	expression tag	UNP Q97S37

- Molecule 2 is an oligosaccharide called 6-O-phosphono-alpha-L-idopyranose-(1-4)-4-thio-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
			Total	C	O	P				S
2	A	2	27	12	13	1	1	0	0	0
2	C	2	27	12	13	1	1	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	B	100	100	100	0	0
3	E	102	102	102	0	0



- Molecule 2: 6-O-phosphono-alpha-L-idopyranose-(1-4)-4-thio-beta-D-glucopyranose

Chain A:  100%

S661
R162

- Molecule 2: 6-O-phosphono-alpha-L-idopyranose-(1-4)-4-thio-beta-D-glucopyranose

Chain C:  50% 50%

S661
R162

4 Data and refinement statistics i

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	184.03Å 66.65Å 133.35Å 90.00° 136.33° 90.00°	Depositor
Resolution (Å)	45.95 – 2.41 45.95 – 2.41	Depositor EDS
% Data completeness (in resolution range)	97.8 (45.95-2.41) 98.0 (45.95-2.41)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.42 (at 2.42Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
R, R_{free}	0.174 , 0.251 0.198 , 0.269	Depositor DCC
R_{free} test set	2147 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å ²)	37.2	Xtrriage
Anisotropy	1.363	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 41.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.006 for -h-2*k,h+1 0.016 for h,-k,-h-l 0.018 for h+2*k,-k,-l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7778	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.16% 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: RTG, SGC

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	B	0.46	0/3837	0.87	5/5204 (0.1%)
1	E	0.47	0/3884	0.85	6/5268 (0.1%)
All	All	0.46	0/7721	0.86	11/10472 (0.1%)

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	237	THR	CA-C-N	7.40	127.07	119.82
1	E	237	THR	C-N-CA	7.40	127.07	119.82
1	B	237	THR	CA-C-N	6.76	126.45	119.82
1	B	237	THR	C-N-CA	6.76	126.45	119.82
1	E	265	TYR	CA-C-N	5.96	125.87	119.85

There are no chirality outliers.

There are no planarity outliers.

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.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	3738	0	3591	69	0
1	E	3784	0	3629	82	0
2	A	27	0	10	1	0
2	C	27	0	10	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	100	0	0	2	0
3	E	102	0	0	0	0
All	All	7778	0	7240	150	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 150 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:E:177:GLU:C	1:E:178:LEU:HD23	1.87	0.99
1:E:382:THR:HG22	1:E:383:VAL:H	1.51	0.75
1:E:364:GLU:OE1	2:C:2:RTG:O2	2.05	0.74
1:E:125:HIS:O	1:E:127:GLU:HG2	1.91	0.70
1:E:382:THR:HG23	1:E:451:LYS:O	1.91	0.69

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	B	453/487 (93%)	443 (98%)	10 (2%)	0	100	100
1	E	459/487 (94%)	449 (98%)	10 (2%)	0	100	100
All	All	912/974 (94%)	892 (98%)	20 (2%)	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	B	402/427 (94%)	383 (95%)	19 (5%)	23	39
1	E	407/427 (95%)	388 (95%)	19 (5%)	23	39
All	All	809/854 (95%)	771 (95%)	38 (5%)	23	39

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	228	VAL
1	E	377	GLU
1	E	248	GLU
1	E	311	LEU
1	E	446	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	440	ASN
1	E	380	ASN
1	E	115	HIS
1	B	380	ASN
1	E	357	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

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	SGC	A	1	2	11,12,12	2.51	3 (27%)	14,17,17	1.37	2 (14%)
2	RTG	A	2	2	15,15,16	2.29	3 (20%)	21,22,24	1.49	3 (14%)
2	SGC	C	1	2	11,12,12	2.86	4 (36%)	14,17,17	1.36	1 (7%)
2	RTG	C	2	2	15,15,16	2.45	3 (20%)	21,22,24	1.10	2 (9%)

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	SGC	A	1	2	-	2/2/22/22	0/1/1/1
2	RTG	A	2	2	-	3/6/23/26	0/1/1/1
2	SGC	C	1	2	-	2/2/22/22	0/1/1/1
2	RTG	C	2	2	-	2/6/23/26	0/1/1/1

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	2	RTG	C2-C3	-7.42	1.41	1.52
2	C	1	SGC	C3-C4	-6.91	1.47	1.53
2	A	2	RTG	C2-C3	-6.49	1.42	1.52
2	A	1	SGC	C3-C4	-5.73	1.48	1.53
2	C	1	SGC	C3-C2	-4.82	1.39	1.52

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1	SGC	O5-C5-C4	3.97	113.41	109.50

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	1	SGC	O5-C5-C4	3.83	113.27	109.50
2	A	2	RTG	C3-C4-C5	3.81	117.14	110.23
2	C	2	RTG	O6-C6-C5	2.62	117.93	108.99
2	C	2	RTG	C3-C4-C5	2.48	114.72	110.23

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

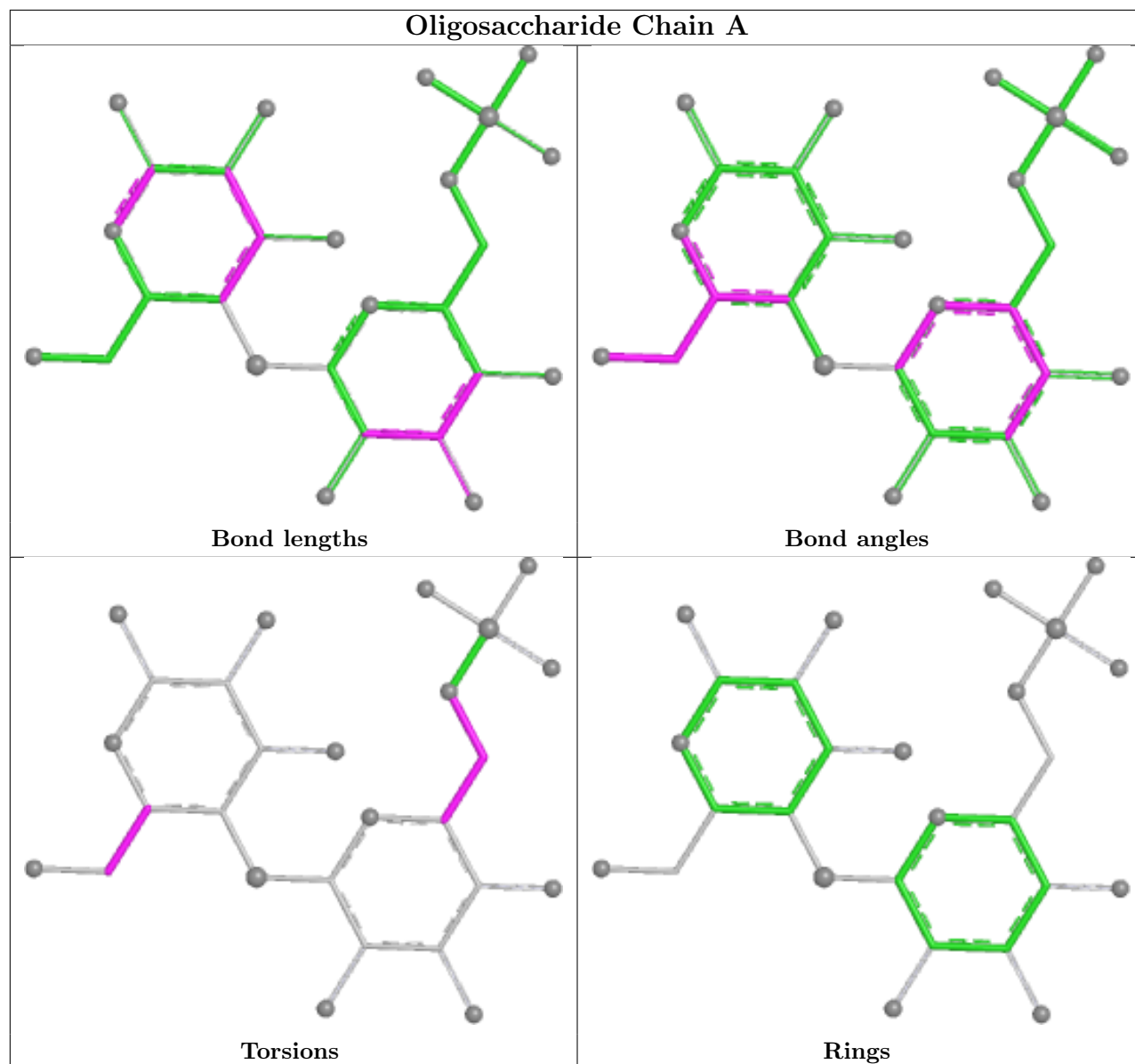
Mol	Chain	Res	Type	Atoms
2	A	2	RTG	C5-C6-O6-PBA
2	A	2	RTG	O5-C5-C6-O6
2	A	2	RTG	C4-C5-C6-O6
2	C	2	RTG	C6-O6-PBA-OAJ
2	C	2	RTG	C5-C6-O6-PBA

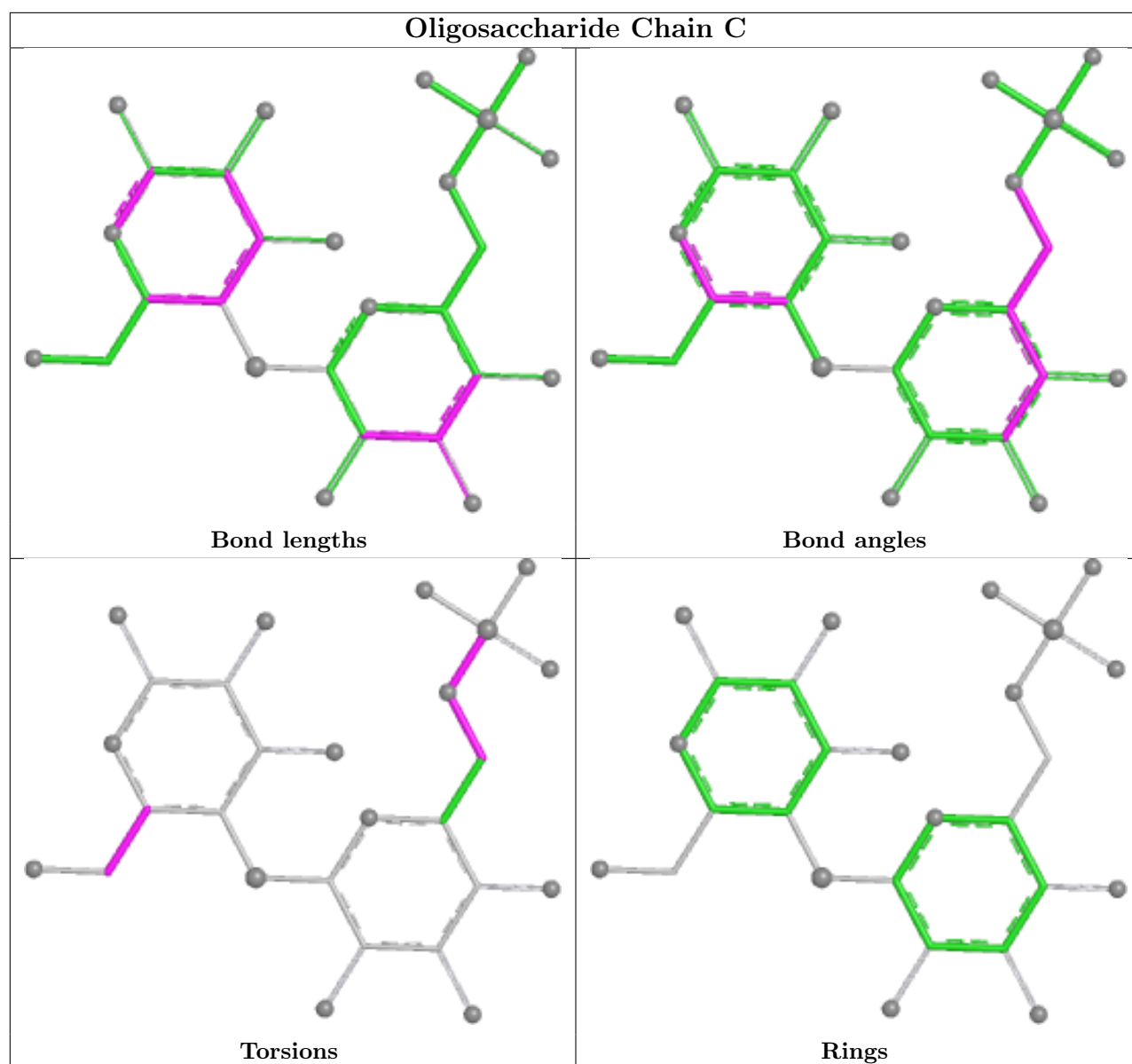
There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	2	RTG	3	0
2	A	1	SGC	1	0
2	A	2	RTG	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 [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, 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	B	457/487 (93%)	0.45	13 (2%) 55 51	29, 47, 67, 84	0
1	E	463/487 (95%)	0.63	26 (5%) 30 26	30, 50, 70, 97	0
All	All	920/974 (94%)	0.54	39 (4%) 40 36	29, 48, 69, 97	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	116	GLY	4.5
1	E	174	SER	3.3
1	E	442	ASP	3.1
1	E	45	THR	2.9
1	E	190	GLU	2.9

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

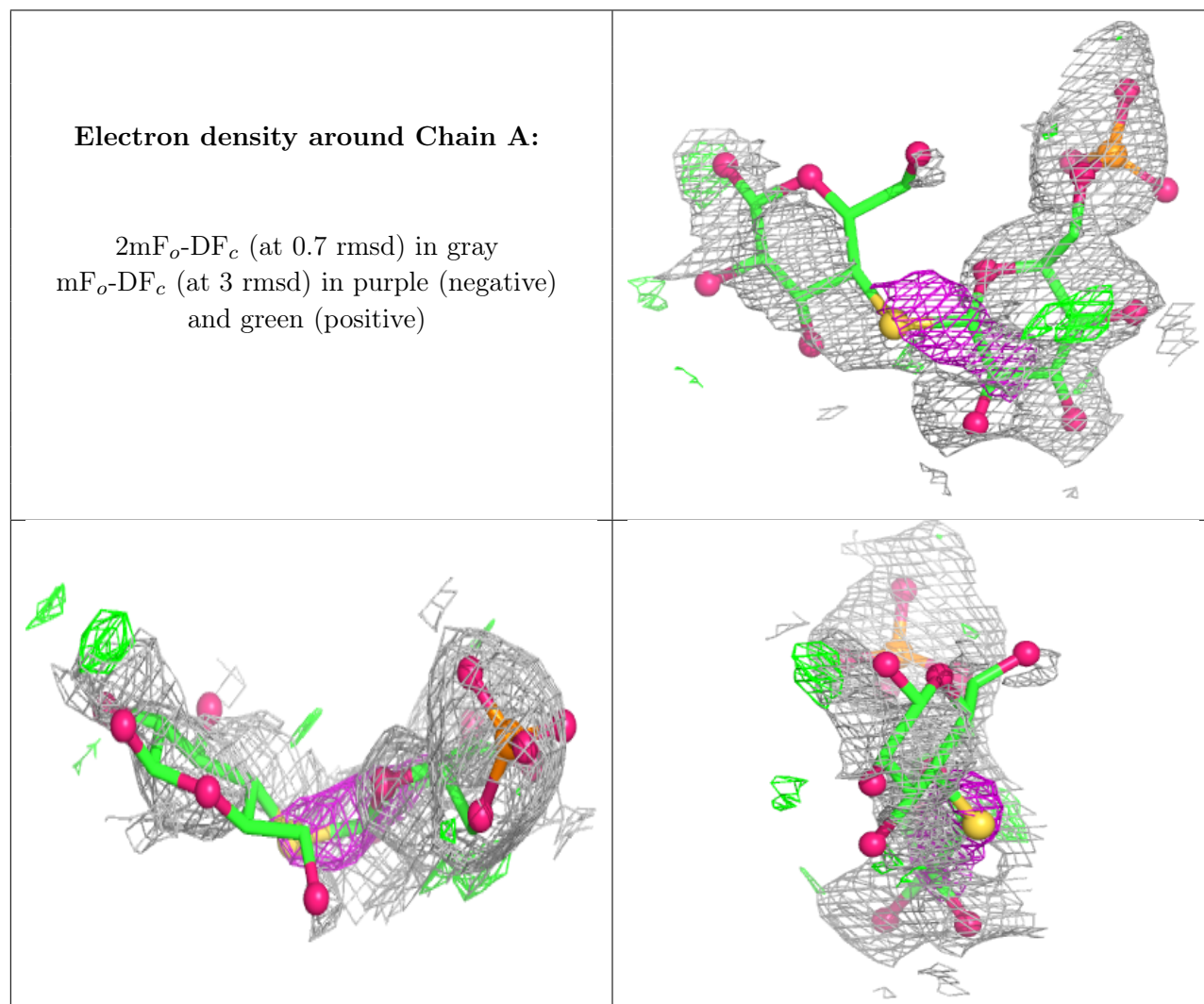
There are no non-standard protein/DNA/RNA residues in this entry.

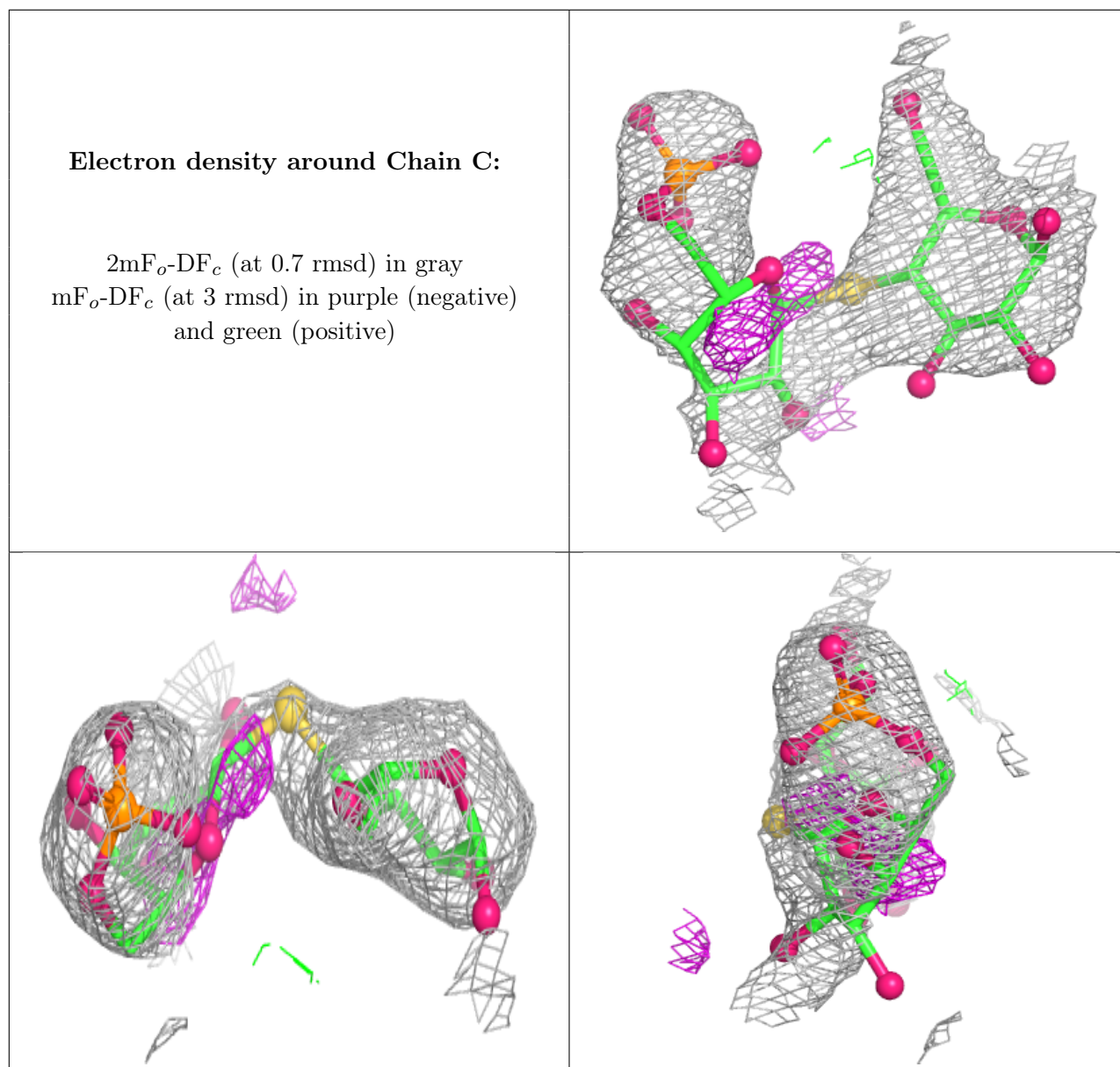
6.3 Carbohydrates [i](#)

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	SGC	A	1	12/12	-	-	63,72,79,81	12
2	RTG	A	2	15/16	-	-	36,53,59,60	15
2	SGC	C	1	12/12	0.51	0.20	73,76,82,87	12
2	RTG	C	2	15/16	0.75	0.18	60,64,69,72	15

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.