



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 10, 2026 – 01:15 AM UTC

PDB ID : 4A45 / pdb\_00004a45  
Title : CpGH89CBM32-5, from *Clostridium perfringens*, in complex with GalNAc-beta-1,3-galactose  
Authors : Ficko-Blean, E.; Stuart, C.P.; Suits, M.D.; Cid, M.; Tessier, M.; Woods, R.J.; Boraston, A.B.  
Deposited on : 2011-10-06  
Resolution : 1.75 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](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 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

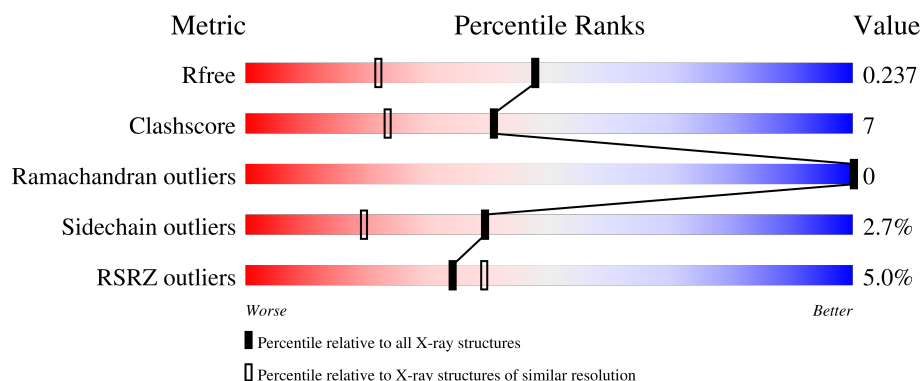
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.75 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	3183 (1.76-1.76)
Clashscore	190562	3299 (1.76-1.76)
Ramachandran outliers	187476	3274 (1.76-1.76)
Sidechain outliers	187428	3274 (1.76-1.76)
RSRZ outliers	180081	3183 (1.76-1.76)

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  $\geq 3$ , 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	161	 4% 46% 35% 14%
2	B	2	 50% 50%

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 1246 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.

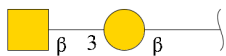
- Molecule 1 is a protein called ALPHA-N-ACETYLGLUCOSAMINIDASE FAMILY PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	139	Total	C	N	O	S	0	0	0
			1063	668	181	213	1			

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1333	MET	-	expression tag	UNP Q0TST1
A	1334	GLY	-	expression tag	UNP Q0TST1
A	1335	SER	-	expression tag	UNP Q0TST1
A	1336	SER	-	expression tag	UNP Q0TST1
A	1337	HIS	-	expression tag	UNP Q0TST1
A	1338	HIS	-	expression tag	UNP Q0TST1
A	1339	HIS	-	expression tag	UNP Q0TST1
A	1340	HIS	-	expression tag	UNP Q0TST1
A	1341	HIS	-	expression tag	UNP Q0TST1
A	1342	HIS	-	expression tag	UNP Q0TST1
A	1343	SER	-	expression tag	UNP Q0TST1
A	1344	SER	-	expression tag	UNP Q0TST1
A	1345	GLY	-	expression tag	UNP Q0TST1
A	1346	LEU	-	expression tag	UNP Q0TST1
A	1347	VAL	-	expression tag	UNP Q0TST1
A	1348	PRO	-	expression tag	UNP Q0TST1
A	1349	ARG	-	expression tag	UNP Q0TST1
A	1350	GLY	-	expression tag	UNP Q0TST1
A	1351	SER	-	expression tag	UNP Q0TST1
A	1352	HIS	-	expression tag	UNP Q0TST1
A	1353	MET	-	expression tag	UNP Q0TST1
A	1354	ALA	-	expression tag	UNP Q0TST1
A	1355	SER	-	expression tag	UNP Q0TST1

- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-galactopyranose-(1-3)-beta-D-galactopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	2	Total	C	N	O	0	0	0
			26	14	1	11			

- Molecule 3 is CALCIUM ION (CCD ID: CA) (formula: Ca).

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

- Molecule 4 is SODIUM ION (CCD ID: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Na	0	0
			1	1		

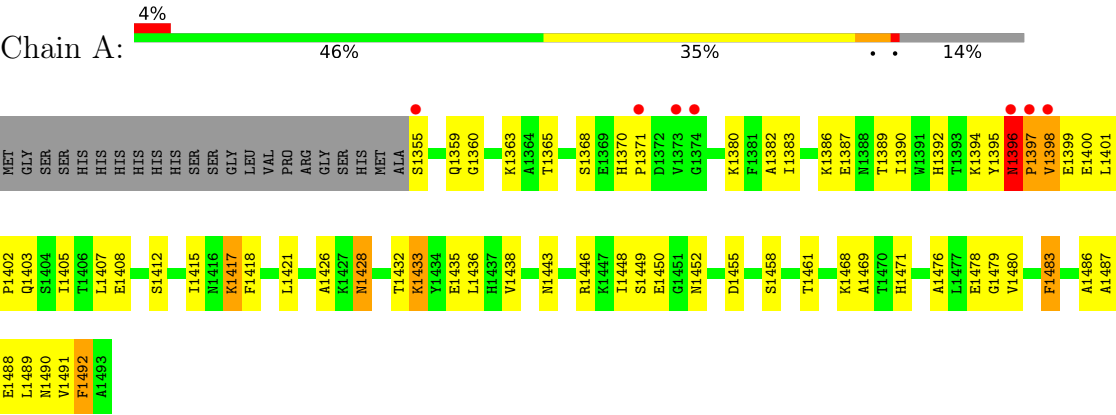
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	151	Total	O	0	4
			155	155		

### 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: ALPHA-N-ACETYLGLUCOSAMINIDASE FAMILY PROTEIN



• Molecule 2: 2-acetamido-2-deoxy-beta-D-galactopyranose-(1-3)-beta-D-galactopyranose



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	33.80Å 58.63Å 70.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.13 – 1.75 45.12 – 1.75	Depositor EDS
% Data completeness (in resolution range)	97.9 (45.13-1.75) 97.9 (45.12-1.75)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.92 (at 1.75Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.199 , 0.246 (Not available) , 0.237	Depositor DCC
$R_{free}$ test set	739 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.3	Xtriage
Anisotropy	0.511	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 50.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	1246	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.57% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GAL, NGA, CA, NA

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	2.13	54/1085 (5.0%)	1.10	5/1466 (0.3%)

All (54) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1412	SER	C-O	-7.49	1.15	1.24
1	A	1365	THR	C-O	-6.72	1.15	1.23
1	A	1359	GLN	C-O	-6.70	1.15	1.24
1	A	1415	ILE	C-O	-6.69	1.17	1.24
1	A	1405	ILE	C-O	-6.57	1.17	1.24
1	A	1476	ALA	C-O	-6.50	1.16	1.24
1	A	1471	HIS	C-O	-6.47	1.16	1.23
1	A	1487	ALA	C-O	-6.45	1.16	1.24
1	A	1452	ASN	C-O	-6.41	1.15	1.23
1	A	1491	VAL	C-O	-6.32	1.17	1.24
1	A	1433	LYS	C-O	-6.28	1.17	1.23
1	A	1452	ASN	CG-ND2	-6.22	1.20	1.33
1	A	1360	GLY	C-O	-6.11	1.16	1.23
1	A	1380	LYS	C-O	-6.09	1.16	1.24
1	A	1426	ALA	CA-CB	-6.05	1.43	1.53
1	A	1443	ASN	CG-ND2	-5.85	1.21	1.33
1	A	1408	GLU	C-O	-5.84	1.16	1.24
1	A	1458	SER	C-O	-5.83	1.16	1.23
1	A	1483	PHE	C-O	-5.79	1.16	1.23
1	A	1389	THR	C-O	-5.72	1.17	1.24
1	A	1403	GLN	C-O	-5.63	1.16	1.23
1	A	1443	ASN	C-O	-5.63	1.17	1.24
1	A	1436	LEU	C-O	-5.62	1.17	1.24
1	A	1396	ASN	CG-OD1	5.62	1.34	1.23
1	A	1386	LYS	C-O	-5.60	1.17	1.24
1	A	1448	ILE	C-O	-5.58	1.17	1.24

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1370	HIS	C-O	-5.56	1.16	1.24
1	A	1428	ASN	C-O	-5.54	1.17	1.23
1	A	1417	LYS	C-O	-5.51	1.17	1.23
1	A	1468	LYS	C-O	-5.49	1.17	1.23
1	A	1438	VAL	C-O	-5.44	1.17	1.23
1	A	1370	HIS	CD2-NE2	-5.43	1.31	1.37
1	A	1479	GLY	C-O	-5.42	1.18	1.24
1	A	1455	ASP	C-O	-5.36	1.16	1.24
1	A	1421	LEU	C-O	-5.34	1.17	1.24
1	A	1418	PHE	C-O	-5.32	1.17	1.23
1	A	1383	ILE	C-O	-5.31	1.17	1.24
1	A	1480	VAL	C-O	-5.31	1.18	1.24
1	A	1450	GLU	C-O	-5.30	1.17	1.23
1	A	1382	ALA	C-O	-5.29	1.17	1.24
1	A	1489	LEU	C-O	-5.27	1.17	1.24
1	A	1492	PHE	C-O	-5.27	1.17	1.23
1	A	1469	ALA	C-O	-5.26	1.17	1.23
1	A	1449	SER	C-O	-5.21	1.17	1.23
1	A	1363	LYS	C-O	-5.21	1.17	1.24
1	A	1435	GLU	C-O	-5.19	1.17	1.23
1	A	1488	GLU	C-O	-5.17	1.17	1.23
1	A	1446	ARG	C-O	-5.16	1.17	1.23
1	A	1461	THR	C-O	-5.14	1.17	1.24
1	A	1486	ALA	C-O	-5.12	1.18	1.23
1	A	1407	LEU	C-O	-5.12	1.18	1.24
1	A	1432	THR	C-O	-5.03	1.17	1.23
1	A	1490	ASN	C-O	-5.03	1.17	1.23
1	A	1390	ILE	C-O	-5.02	1.18	1.23

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1370	HIS	CB-CG-CD2	-6.57	122.66	131.20
1	A	1398	VAL	N-CA-C	5.53	117.59	109.63
1	A	1370	HIS	CB-CG-ND1	5.40	130.80	122.70
1	A	1400	GLU	N-CA-C	5.34	118.03	110.50
1	A	1397	PRO	N-CA-C	-5.12	98.80	112.10

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	A	1063	0	1030	14	0
2	B	26	0	24	1	0
3	A	1	0	0	0	0
4	A	1	0	0	0	0
5	A	155	0	0	4	0
All	All	1246	0	1054	14	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.

All (14) 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:1395:TYR:HE2	5:A:2046:HOH:O	1.74	0.69
1:A:1395:TYR:CE2	5:A:2046:HOH:O	2.49	0.65
1:A:1371:PRO:HD3	1:A:1394:LYS:HE2	1.78	0.65
1:A:1396:ASN:OD1	1:A:1396:ASN:N	2.29	0.64
1:A:1395:TYR:HA	1:A:1398:VAL:HA	1.90	0.53
1:A:1433:LYS:HB3	1:A:1478:GLU:HB2	1.90	0.53
1:A:1417:LYS:HB3	1:A:1492:PHE:HB2	1.91	0.53
1:A:1395:TYR:CG	2:B:2:NGA:H62	2.48	0.49
1:A:1428:ASN:HB2	5:A:2086:HOH:O	2.12	0.48
1:A:1368:SER:HB2	1:A:1399:GLU:OE1	2.17	0.44
1:A:1392:HIS:HB3	5:A:2046:HOH:O	2.17	0.43
1:A:1396:ASN:HA	1:A:1397:PRO:HA	1.58	0.43
1:A:1401:LEU:HB3	1:A:1402:PRO:HA	2.01	0.42
1:A:1428:ASN:ND2	1:A:1483:PHE:CE2	2.87	0.42

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	A	137/161 (85%)	135 (98%)	2 (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	A	113/131 (86%)	110 (97%)	3 (3%)	39	19

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1355	SER
1	A	1387	GLU
1	A	1396	ASN

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	1428	ASN

### 5.3.3 RNA [i](#)

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 ⓘ

2 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	GAL	B	1	2	12,12,12	0.54	0	17,17,17	1.12	2 (11%)
2	NGA	B	2	2	14,14,15	0.67	0	17,19,21	1.17	1 (5%)

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	GAL	B	1	2	-	2/2/22/22	0/1/1/1
2	NGA	B	2	2	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	2	NGA	C1-O5-C5	3.97	117.51	112.19
2	B	1	GAL	O3-C3-C2	2.51	116.30	110.38
2	B	1	GAL	O5-C5-C6	2.15	111.77	106.44

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1	GAL	O5-C5-C6-O6

*Continued on next page...*

*Continued from previous page...*

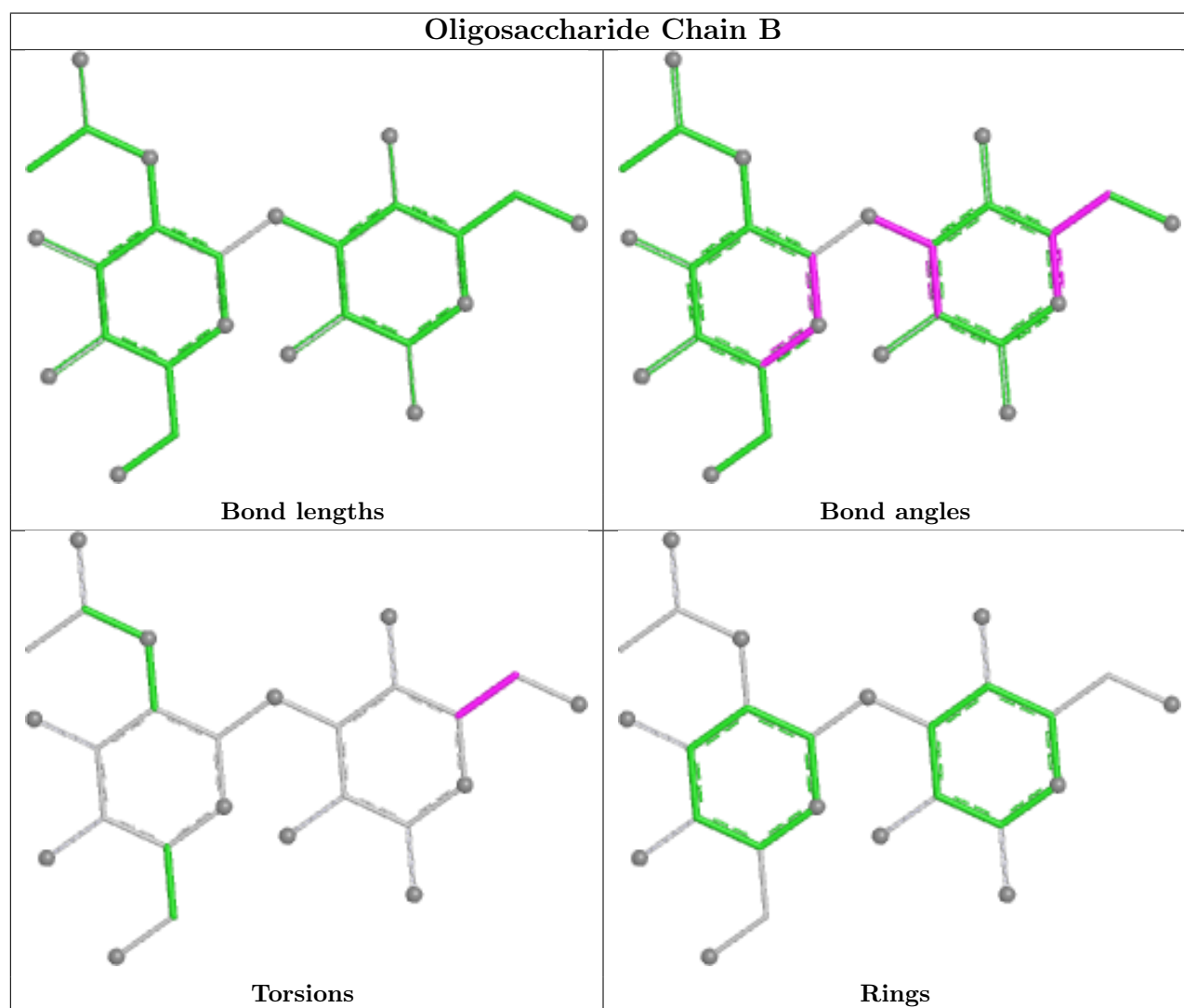
Mol	Chain	Res	Type	Atoms
2	B	1	GAL	C4-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	2	NGA	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 ⓘ

Of 2 ligands modelled in this entry, 2 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 [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	139/161 (86%)	0.10	7 (5%) 34 39	16, 22, 46, 54	0

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1396	ASN	4.7
1	A	1397	PRO	4.7
1	A	1398	VAL	3.3
1	A	1371	PRO	2.7
1	A	1373	VAL	2.7
1	A	1374	GLY	2.3
1	A	1355	SER	2.1

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

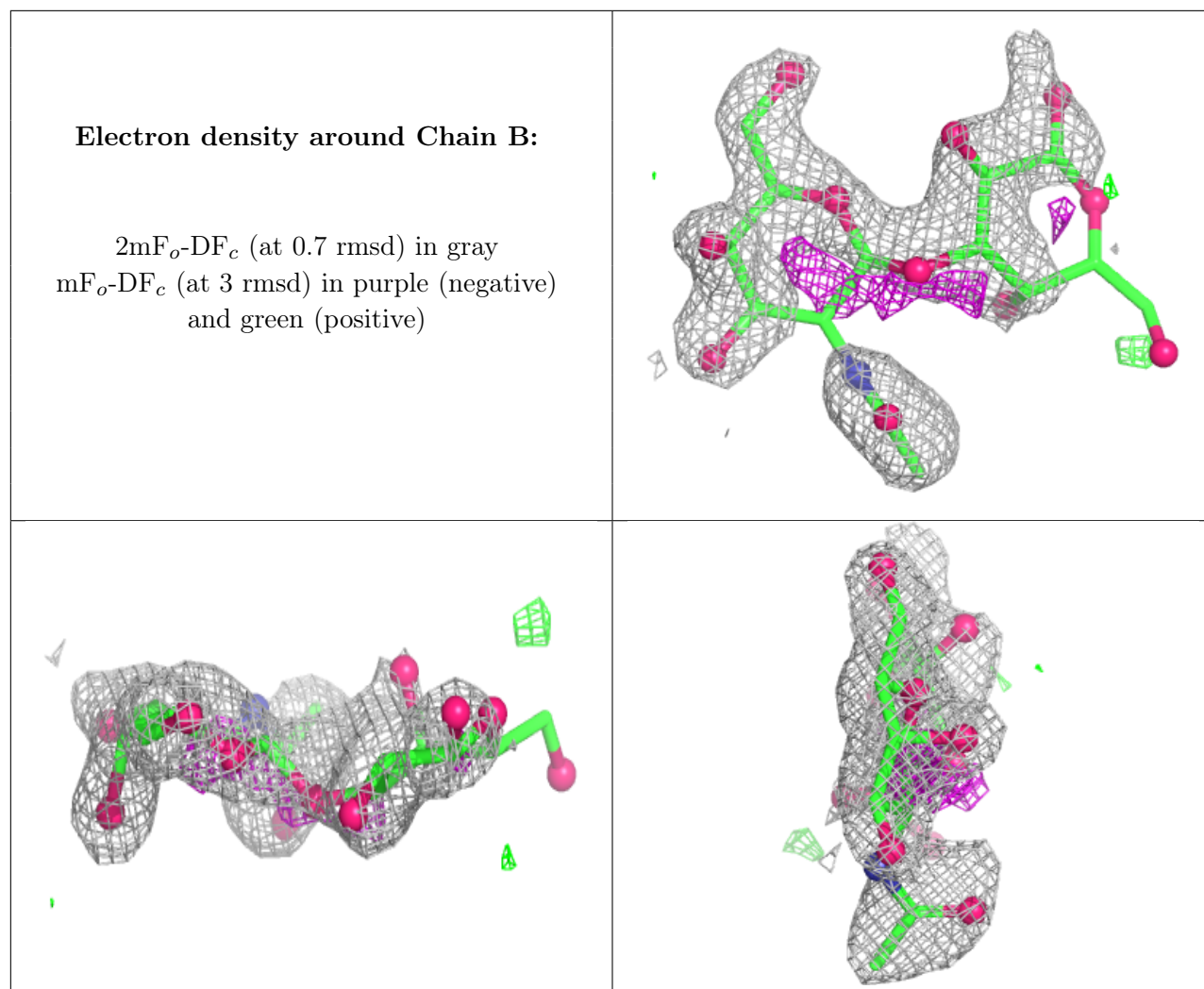
There are no non-standard protein/DNA/RNA residues in this entry.

### 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	GAL	B	1	12/12	0.59	0.14	60,63,65,66	0
2	NGA	B	2	14/15	0.71	0.14	48,53,56,57	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, 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( $\text{\AA}^2$ )	Q<0.9
4	NA	A	2497	1/1	0.98	0.22	5,5,5,5	0
3	CA	A	2496	1/1	0.99	0.03	19,19,19,19	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.