



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 7, 2020 – 04:00 PM BST

PDB ID : 2WNF
Title : Crystal Structure of a Mammalian Sialyltransferase in complex with Gal-beta-1-3GalNAc-ortho-nitrophenol
Authors : Rao, F.V.; Rich, J.R.; Raikic, B.; Wakarchuk, W.W.; Withers, S.G.; Strynadka, N.C.J.
Deposited on : 2009-07-09
Resolution : 1.25 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 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

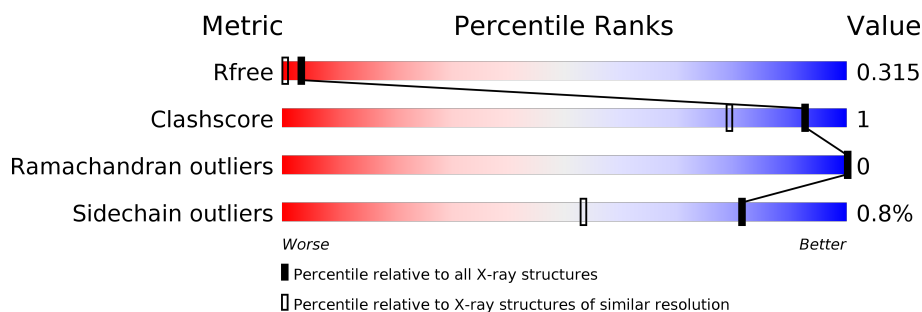
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.25 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	1023 (1.28-1.24)
Clashscore	141614	1060 (1.28-1.24)
Ramachandran outliers	138981	1029 (1.28-1.24)
Sidechain outliers	138945	1028 (1.28-1.24)

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 ≥ 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	298	 82% 8% 9%
2	B	2	 100%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 2699 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 CMP-N-ACETYLNEURAMINATE-BETA-GALACTOSAMIDE-ALPHA-2,3-SIALYLTRANSFERASE.

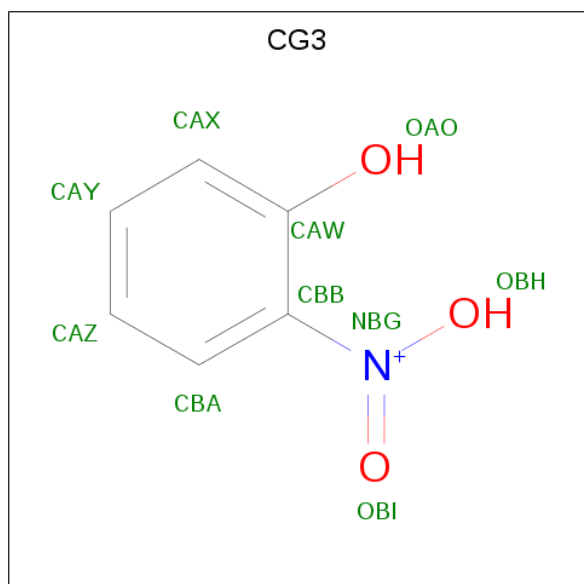
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	272	Total	C	N	O	S	Se	0	17	0
			2343	1493	415	424	8	3			

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



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	2	Total	C	N	O	0	0	0
			25	14	1	10			

- Molecule 3 is HYDROXY(2-HYDROXYPHENYL)OXOAMMONIUM (three-letter code: CG3) (formula: C₆H₆NO₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			10	6	1	3		

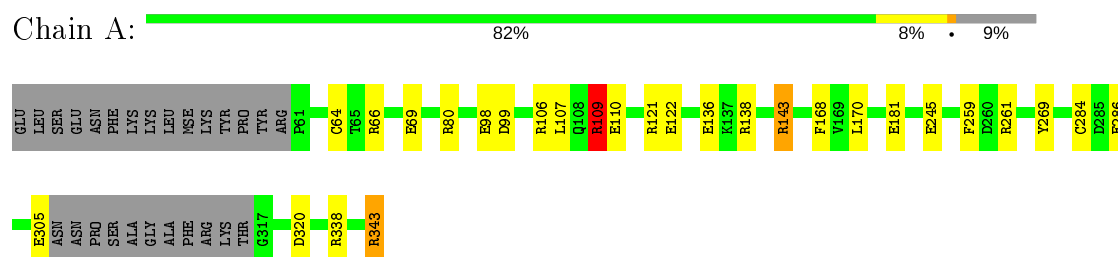
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	321	Total	O	0	0
			321	321		

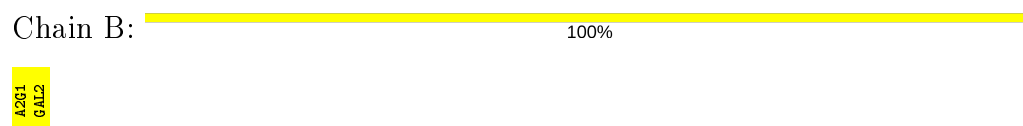
3 Residue-property plots

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. 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. 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: CMP-N-ACETYLNEURAMINATE-BETA-GALACTOSAMIDE-ALPHA-2,3-SIALYLTRANSFERASE



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



4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	77.99Å 78.35Å 99.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.11 – 1.25 32.93 – 1.24	Depositor EDS
% Data completeness (in resolution range)	99.8 (33.11-1.25) 99.3 (32.93-1.24)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.39 (at 1.24Å)	Xtriage
Refinement program	REFMAC 5.5.0088	Depositor
R, R_{free}	0.180 , 0.195 0.307 , 0.315	Depositor DCC
R_{free} test set	4261 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	9.1	Xtriage
Anisotropy	0.409	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 30.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.011 for -k,-h,-l	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	2699	wwPDB-VP
Average B, all atoms (Å ²)	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.85% 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: CG3, GAL, A2G

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	1.24	13/2423 (0.5%)	1.10	13/3271 (0.4%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	245	GLU	CD-OE1	8.41	1.34	1.25
1	A	110	GLU	CB-CG	-6.79	1.39	1.52
1	A	69	GLU	CD-OE2	-6.43	1.18	1.25
1	A	98	GLU	CD-OE1	6.14	1.32	1.25
1	A	69	GLU	CB-CG	-6.13	1.40	1.52

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	343	ARG	NE-CZ-NH1	10.48	125.54	120.30
1	A	143	ARG	NE-CZ-NH2	-8.15	116.23	120.30
1	A	99	ASP	CB-CG-OD1	6.66	124.30	118.30
1	A	80	ARG	NE-CZ-NH2	-6.62	116.99	120.30
1	A	138	ARG	NE-CZ-NH1	5.82	123.21	120.30

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	A	2343	0	2294	7	0
2	B	25	0	21	0	0
3	A	10	0	5	0	0
4	A	321	0	0	3	0
All	All	2699	0	2320	7	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 1.

The worst 5 of 7 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:122[A]:GLU:HG2	4:A:2093:HOH:O	1.60	1.02
1:A:106:ARG:O	1:A:109:ARG:HD2	2.01	0.61
1:A:122[A]:GLU:CG	4:A:2093:HOH:O	2.35	0.57
1:A:181:GLU:HB2	4:A:2167:HOH:O	2.08	0.54
1:A:143:ARG:HG2	1:A:284[B]:CYS:SG	2.59	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	285/298 (96%)	278 (98%)	7 (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	261/263 (99%)	259 (99%)	2 (1%)	81	53

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	109	ARG
1	A	343	ARG

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 ⓘ

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	A2G	B	1	3,2	14,14,15	1.26	2 (14%)	17,19,21	1.57	3 (17%)
2	GAL	B	2	2	11,11,12	1.49	2 (18%)	15,15,17	1.14	2 (13%)

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	A2G	B	1	3,2	-	0/6/23/26	0/1/1/1
2	GAL	B	2	2	-	0/2/19/22	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	A2G	C1-C2	3.11	1.57	1.52
2	B	2	GAL	O5-C5	2.98	1.49	1.43
2	B	2	GAL	O2-C2	2.27	1.48	1.43
2	B	1	A2G	O5-C5	2.06	1.47	1.43

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1	A2G	C1-O5-C5	2.94	116.17	112.19
2	B	1	A2G	O5-C1-C2	-2.80	106.87	111.29
2	B	2	GAL	C1-C2-C3	2.13	112.28	109.67
2	B	1	A2G	O3-C3-C2	-2.10	105.11	109.47
2	B	2	GAL	C2-C3-C4	-2.09	107.27	110.89

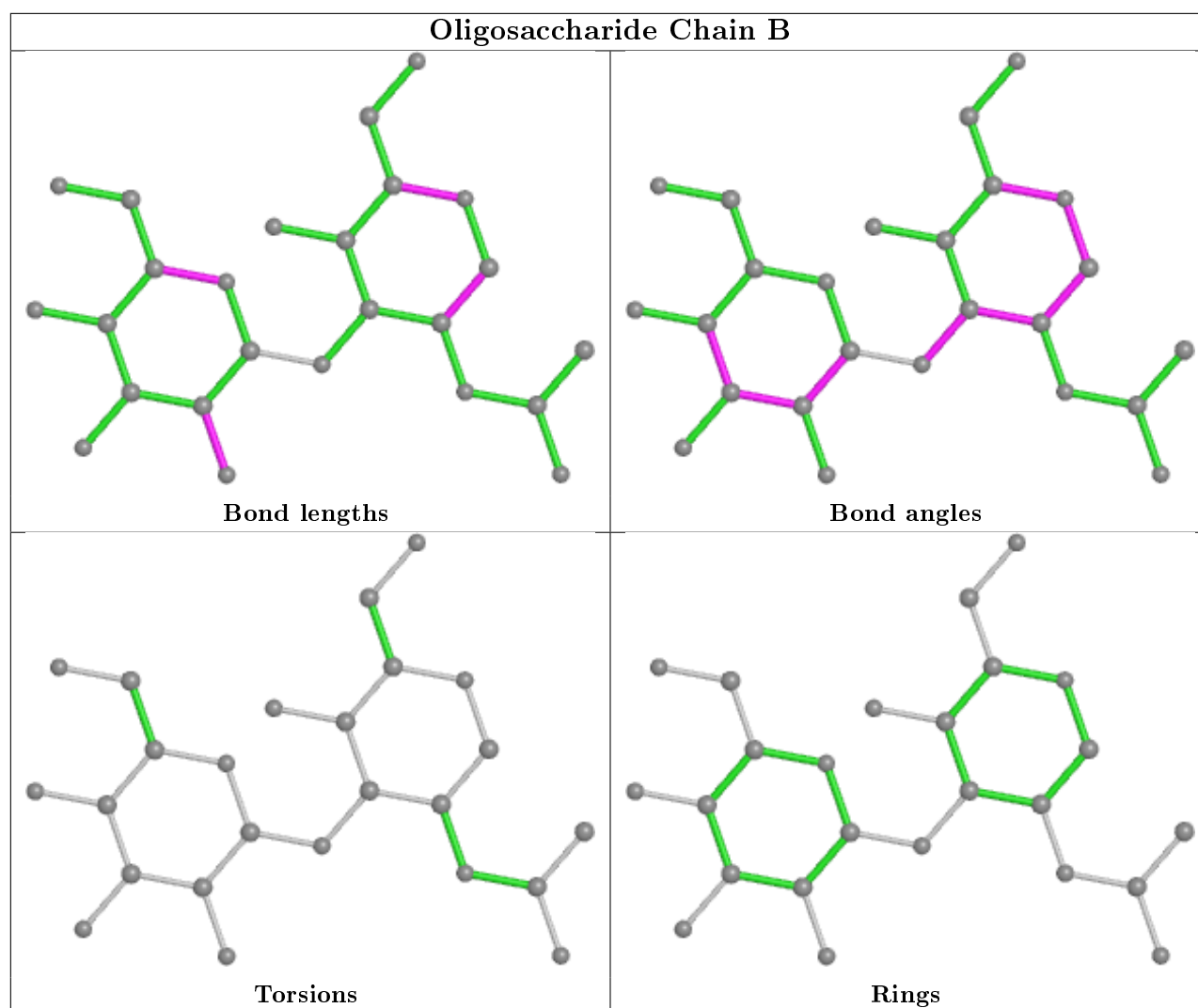
There are no chirality outliers.

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](#)

1 ligand is 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	$\# Z > 2$
3	CG3	A	1344	2	9,10,10	4.61	4 (44%)	10,13,13	3.17	6 (60%)

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
3	CG3	A	1344	2	-	1/2/4/4	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1344	CG3	OBI-NBG	11.46	1.42	1.22
3	A	1344	CG3	CBB-NBG	-6.11	1.34	1.45
3	A	1344	CG3	CAY-CAX	3.01	1.45	1.38
3	A	1344	CG3	OAO-CAW	2.74	1.42	1.36

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1344	CG3	CAY-CAZ-CBA	5.88	129.15	120.19
3	A	1344	CG3	CAX-CAW-CBB	4.70	125.28	118.95
3	A	1344	CG3	OBI-NBG-CBB	3.73	125.42	119.03
3	A	1344	CG3	CAY-CAX-CAW	-3.42	115.72	120.05
3	A	1344	CG3	CAZ-CBA-CBB	-2.86	113.75	118.61

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1344	CG3	CBA-CBB-NBG-OBI

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](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

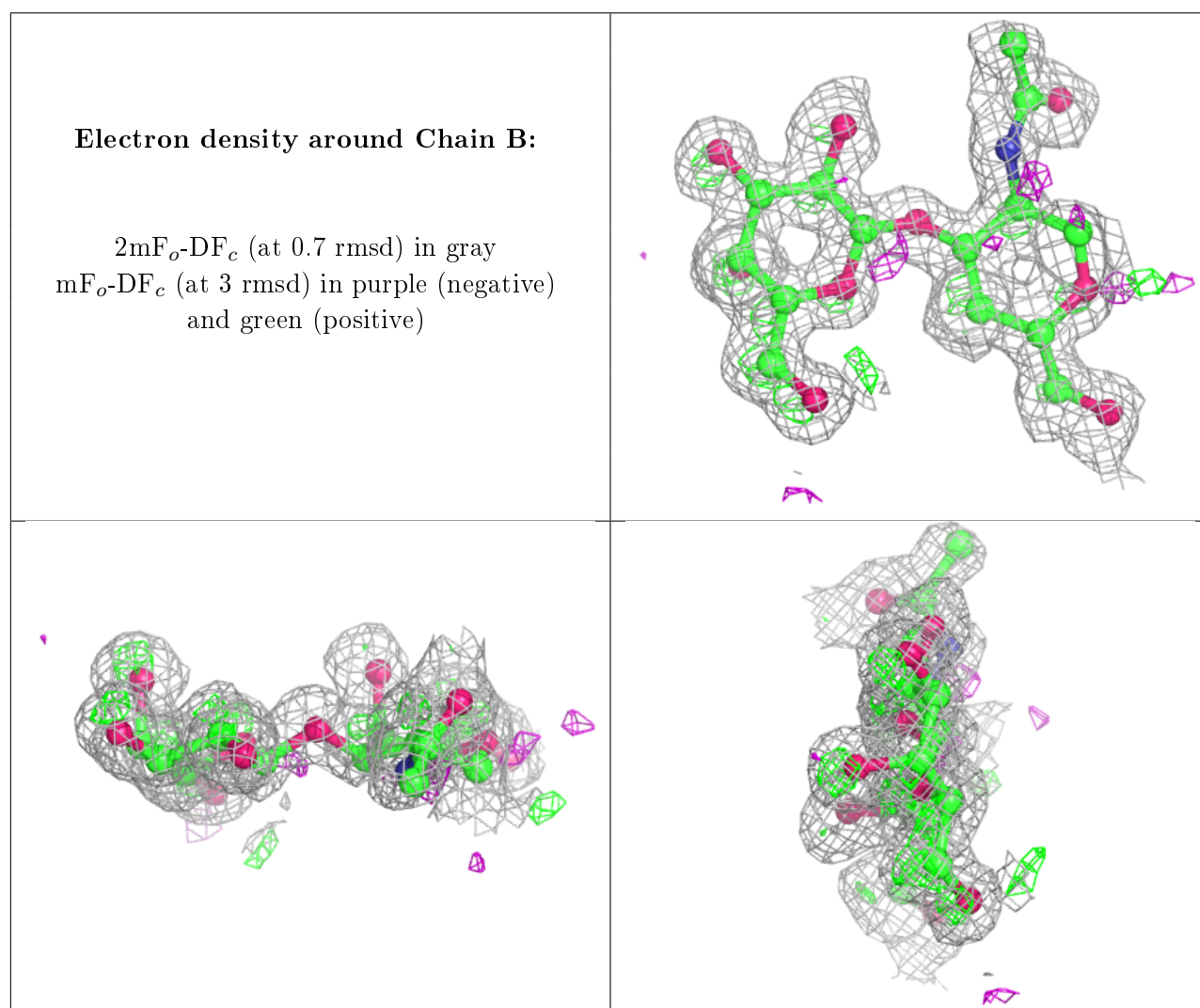
6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.