



Full wwPDB X-ray Structure Validation Report ⓘ

Apr 22, 2025 – 12:36 AM EDT

PDB ID : 4YM1 / pdb_00004ym1
Title : Crystal structure of the human galectin-4 C-terminal carbohydrate recognition domain in complex with 2'-fucosyllactose
Authors : Bum-Erdene, K.; Blanchard, H.
Deposited on : 2015-03-06
Resolution : 2.00 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

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 : 2022.3.0, CSD as543be (2022)
Xtrriage (Phenix) : 2.0rc1
EDS : **FAILED**
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.42

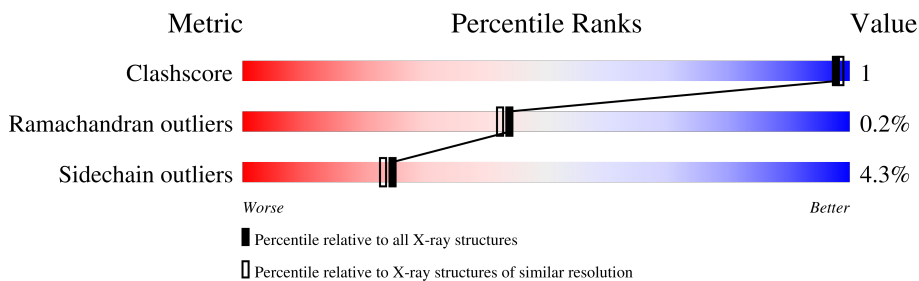
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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))
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (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 $\leq 5\%$

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	153	84% 6% 10%
1	B	153	87% 10%
1	C	153	88% 8%
1	D	153	83% 8% 8%
2	E	3	33% 67%
2	F	3	67% 33%

2 Entry composition [i](#)

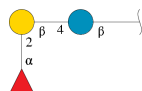
There are 4 unique types of molecules in this entry. The entry contains 4550 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Galectin-4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	138	1082	694	194	192	2	0	0	0
1	B	138	1082	694	194	192	2	0	0	0
1	C	140	1093	701	196	194	2	0	0	0
1	D	140	1093	701	196	194	2	0	0	0

- Molecule 2 is an oligosaccharide called alpha-L-fucopyranose-(1-2)-beta-D-galactopyranose-(1-4)-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
			Total	C	O			
2	E	3	33	18	15	0	0	0
2	F	3	33	18	15	0	0	0

- Molecule 3 is SULFATE ION (CCD ID: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0
3	C	1	Total O S 5 4 1	0	0

- Molecule 4 is water.


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	38	Total O 38 38	0	0
4	B	21	Total O 21 21	0	0
4	C	37	Total O 37 37	0	0
4	D	23	Total O 23 23	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. 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.


Note EDS failed to run properly.

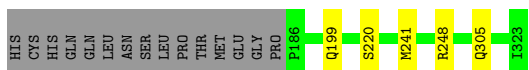
- Molecule 1: Galectin-4

Chain A:  84% 6% 10%



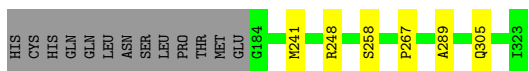
- Molecule 1: Galectin-4

Chain B:  87% 10%




- Molecule 1: Galectin-4

Chain C:  88% 8%



- Molecule 1: Galectin-4

Chain D:  83% 8% 8%



- Molecule 2: alpha-L-fucopyranose-(1-2)-beta-D-galactopyranose-(1-4)-beta-D-glucopyranose

Chain E:  33% 67%



- Molecule 2: alpha-L-fucopyranose-(1-2)-beta-D-galactopyranose-(1-4)-beta-D-glucopyranose



4 Data and refinement statistics

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	45.29Å 129.96Å 45.21Å 90.00° 96.96° 90.00°	Depositor
Resolution (Å)	44.96 – 2.00	Depositor
% Data completeness (in resolution range)	99.2 (44.96-2.00)	Depositor
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	6.15 (at 2.00Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.163 , 0.213	Depositor
Wilson B-factor (Å ²)	29.6	Xtrriage
Anisotropy	0.228	Xtrriage
L-test for twinning ²	$\langle L \rangle = 0.37$, $\langle L^2 \rangle = 0.19$	Xtrriage
Estimated twinning fraction	0.437 for l,-k,h	Xtrriage
Reported twinning fraction	0.447 for H, K, L 0.553 for L, -K, H	Depositor
Outliers	0 of 34831 reflections	Xtrriage
Total number of atoms	4550	wwPDB-VP
Average B, all atoms (Å ²)	32.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.46% 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: BGC, SO4, FUC, GAL

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	0.39	0/1111	0.64	0/1502
1	B	0.40	0/1111	0.65	0/1502
1	C	0.41	0/1123	0.69	0/1520
1	D	0.39	0/1123	0.64	0/1520
All	All	0.40	0/4468	0.66	0/6044

There are no bond length outliers.

There are no bond angle outliers.

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	1082	0	1074	3	0
1	B	1082	0	1074	0	0
1	C	1093	0	1083	1	0
1	D	1093	0	1083	3	0
2	E	33	0	30	0	0
2	F	33	0	29	0	0
3	A	10	0	0	0	0
3	C	5	0	0	0	0
4	A	38	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	21	0	0	0	0
4	C	37	0	0	0	0
4	D	23	0	0	0	0
All	All	4550	0	4373	6	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.

All (6) 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:251:LEU:HD13	1:A:256:TRP:CE2	2.50	0.46
1:A:245:THR:HG22	1:D:265:HIS:HB3	2.00	0.43
1:A:287:VAL:HG21	1:A:295:PHE:CZ	2.55	0.41
1:D:236:HIS:NE2	1:D:238:ASN:HB2	2.35	0.41
1:D:235:LEU:HD13	1:D:297:PHE:CD1	2.55	0.41
1:C:267:PRO:HG3	1:C:289:ALA:CB	2.52	0.40

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	136/153 (89%)	134 (98%)	1 (1%)	1 (1%)	19	14
1	B	136/153 (89%)	131 (96%)	5 (4%)	0	100	100
1	C	138/153 (90%)	138 (100%)	0	0	100	100
1	D	138/153 (90%)	134 (97%)	4 (3%)	0	100	100
All	All	548/612 (90%)	537 (98%)	10 (2%)	1 (0%)	44	42

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	199	GLN

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	116/130 (89%)	113 (97%)	3 (3%)	41	44
1	B	116/130 (89%)	111 (96%)	5 (4%)	25	23
1	C	117/130 (90%)	113 (97%)	4 (3%)	32	32
1	D	117/130 (90%)	109 (93%)	8 (7%)	13	10
All	All	466/520 (90%)	446 (96%)	20 (4%)	25	23

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	241	MET
1	A	248	ARG
1	A	264	THR
1	B	199	GLN
1	B	220	SER
1	B	241	MET
1	B	248	ARG
1	B	305	GLN
1	C	241	MET
1	C	248	ARG
1	C	258	SER
1	C	305	GLN
1	D	199	GLN
1	D	202	LEU
1	D	241	MET
1	D	248	ARG
1	D	253	ASN
1	D	264	THR
1	D	286	LYS
1	D	301	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	253	ASN
1	D	305	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](#)

6 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	BGC	E	1	2	12,12,12	0.47	0	17,17,17	0.90	0
2	GAL	E	2	2	11,11,12	0.74	1 (9%)	15,15,17	2.27	3 (20%)
2	FUC	E	3	2	10,10,11	1.06	1 (10%)	14,14,16	3.09	2 (14%)
2	BGC	F	1	2	12,12,12	0.45	0	17,17,17	0.70	0
2	GAL	F	2	2	11,11,12	0.60	0	15,15,17	2.22	3 (20%)
2	FUC	F	3	2	10,10,11	0.40	0	14,14,16	0.66	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	BGC	E	1	2	-	0/2/22/22	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GAL	E	2	2	-	0/2/19/22	0/1/1/1
2	FUC	E	3	2	-	-	0/1/1/1
2	BGC	F	1	2	-	0/2/22/22	0/1/1/1
2	GAL	F	2	2	-	0/2/19/22	0/1/1/1
2	FUC	F	3	2	-	-	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	3	FUC	O2-C2	2.05	1.47	1.43
2	E	2	GAL	C2-C3	2.03	1.55	1.52

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	3	FUC	O2-C2-C1	10.83	134.02	109.22
2	F	2	GAL	O3-C3-C2	6.21	122.73	110.05
2	E	2	GAL	C1-C2-C3	6.14	118.59	109.64
2	E	2	GAL	O2-C2-C1	4.94	120.52	109.22
2	F	2	GAL	C2-C3-C4	3.82	117.59	110.86
2	E	2	GAL	O2-C2-C3	2.92	116.19	110.15
2	E	3	FUC	C1-C2-C3	-2.42	106.12	109.64
2	F	2	GAL	O3-C3-C4	2.29	115.78	110.38

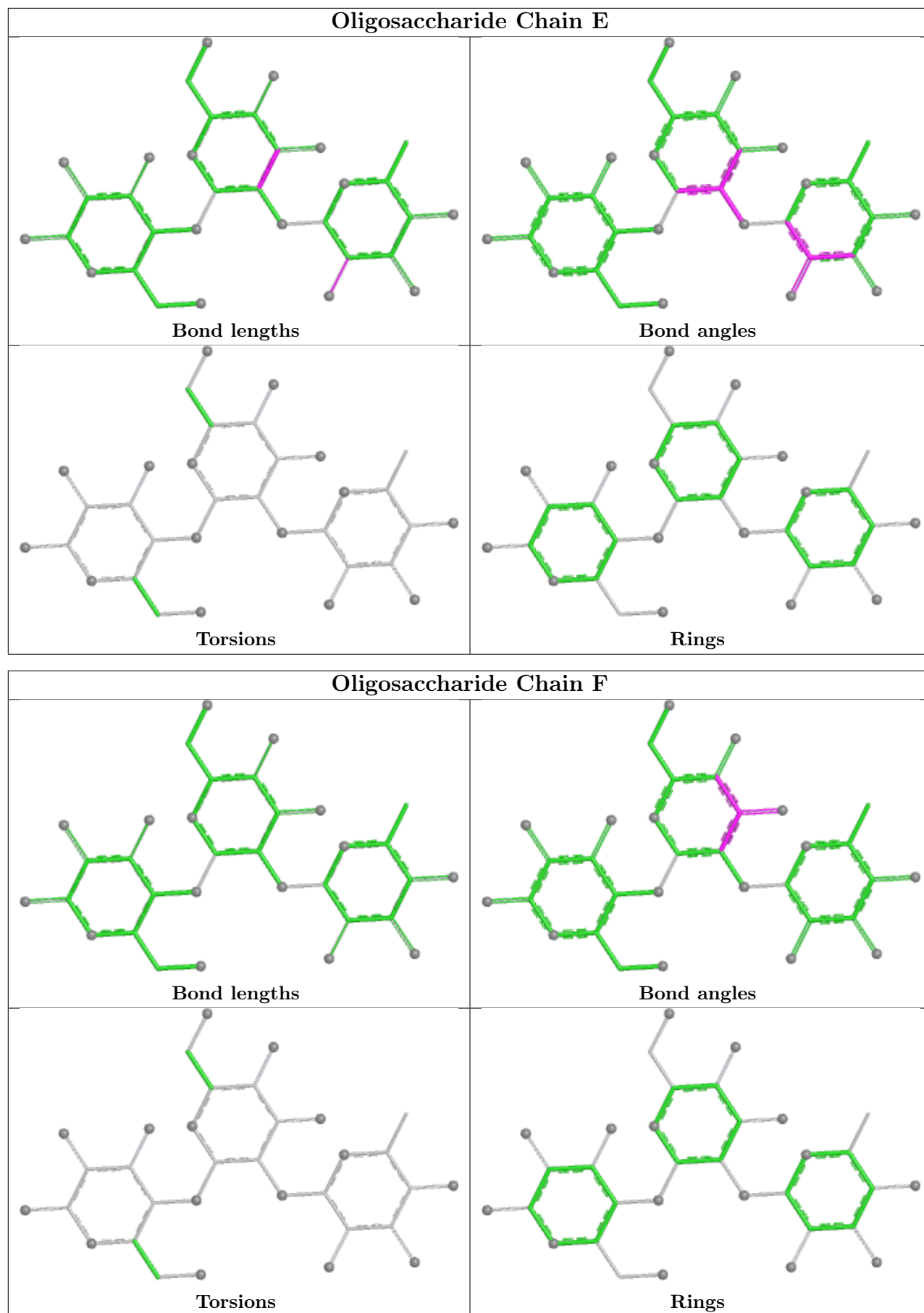
There are no chirality outliers.

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](#)

3 ligands 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	$\# Z > 2$
3	SO4	C	401	-	4,4,4	0.43	0	6,6,6	0.14	0
3	SO4	A	404	-	4,4,4	0.45	0	6,6,6	0.10	0
3	SO4	A	403	-	4,4,4	0.45	0	6,6,6	0.16	0

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

6.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS failed to run properly - this section is therefore empty.

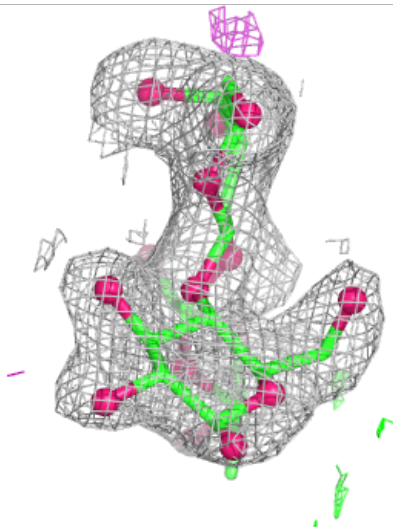
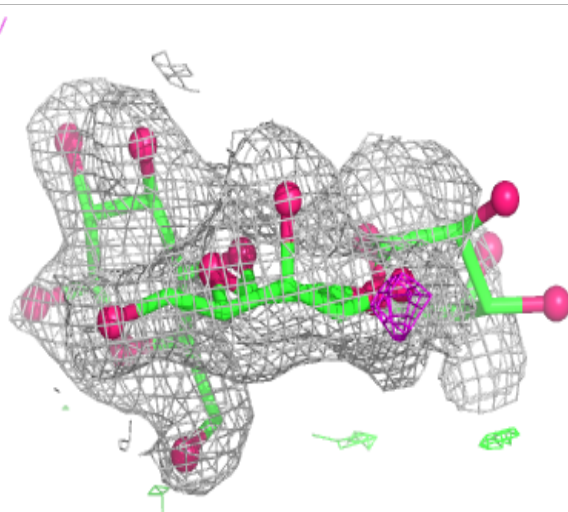
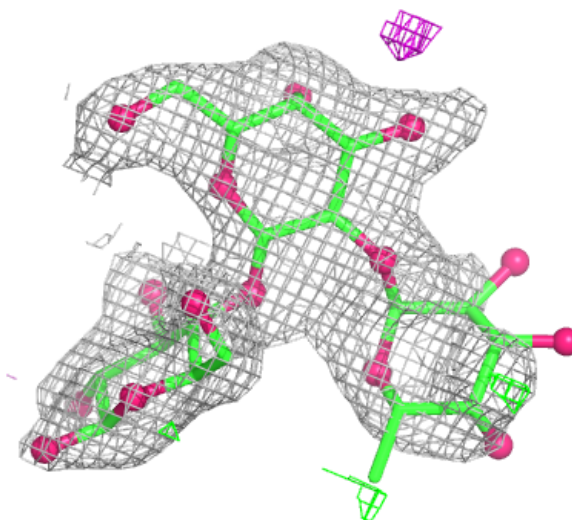
6.3 Carbohydrates

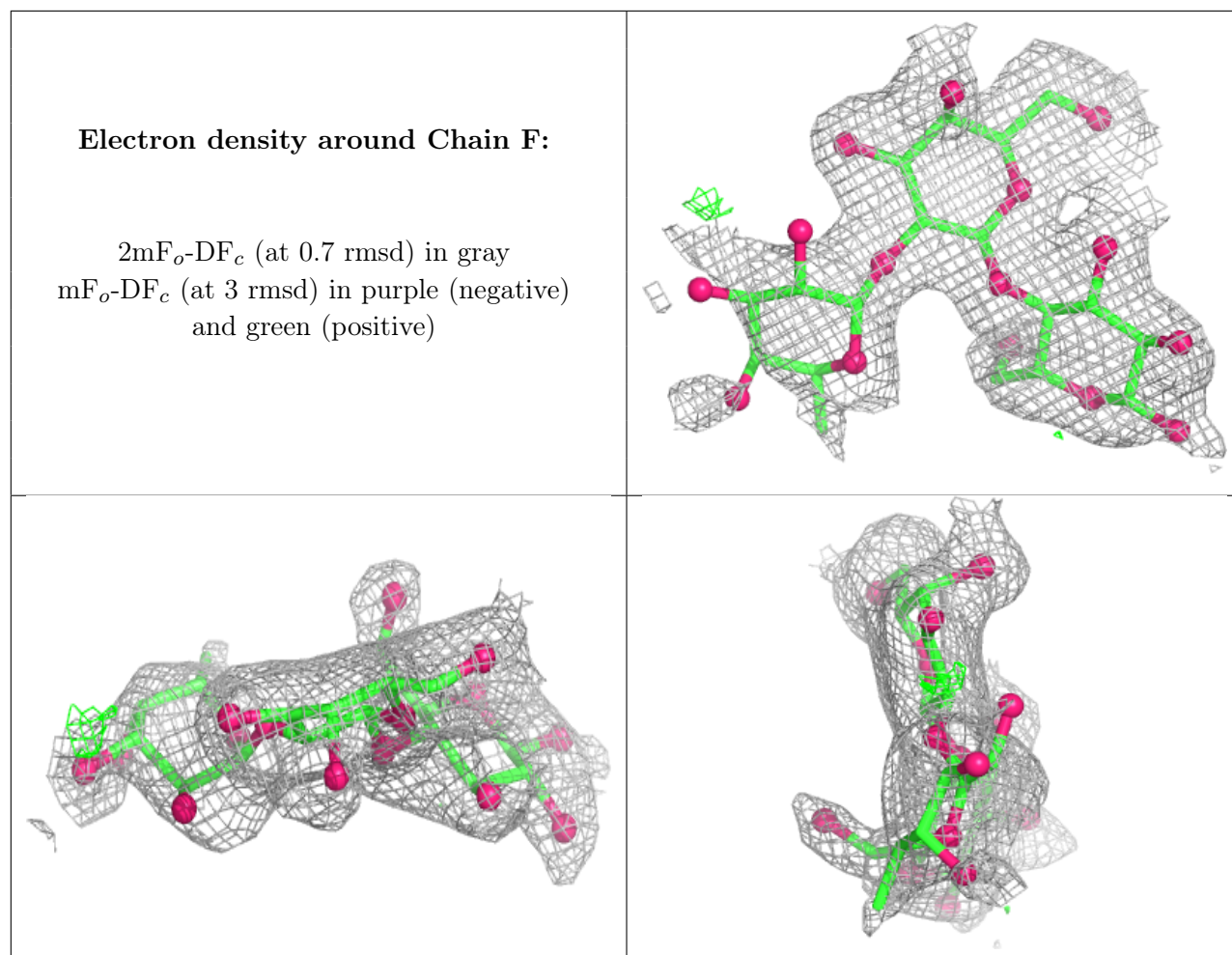
EDS failed to run properly - this section is therefore empty.

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.

Electron density around Chain E:

$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)





6.4 Ligands [i](#)

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers [i](#)

EDS failed to run properly - this section is therefore empty.