wwPDB X-ray Structure Validation Summary Report (i)

Aug 10, 2020 – 04:47 AM BST

PDB ID	:	1CGU
Title	:	CATALYTIC CENTER OF CYCLODEXTRIN GLYCOSYLTRANSFERASE DERIVED FROM X-RAY STRUCTURE ANALYSIS COMBINED WITH SITE-DIRECTED MUTAGENESIS
Authors	:	Klein, C.; Hollender, J.; Bender, H.; Schulz, G.E.
Deposited on	:	1992-06-10
Resolution	:	2.50 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 following versions of software and data (see references (1)) were used in the production of this report:

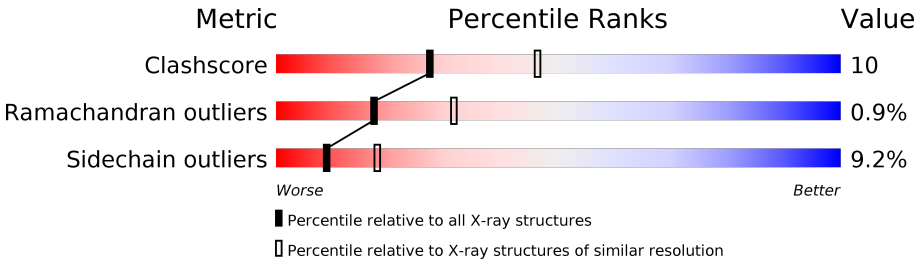
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.50 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	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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 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 <=5%

Mol	Chain	Length	Quality of chain
1	A	684	66% 28% 6% •
2	B	2	50% 50%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5766 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 CYCLODEXTRIN GLYCOSYL-TRANSFERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	684	Total	C	N	O	S	0	0	0
			5264	3322	891	1038	13			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	229	ALA	ASP	conflict	UNP P30920

- Molecule 2 is an oligosaccharide called alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	B	2	Total	C	O	0	0	0
			22	12	10			

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total	Ca	0	0
			2	2		

- Molecule 4 is water.

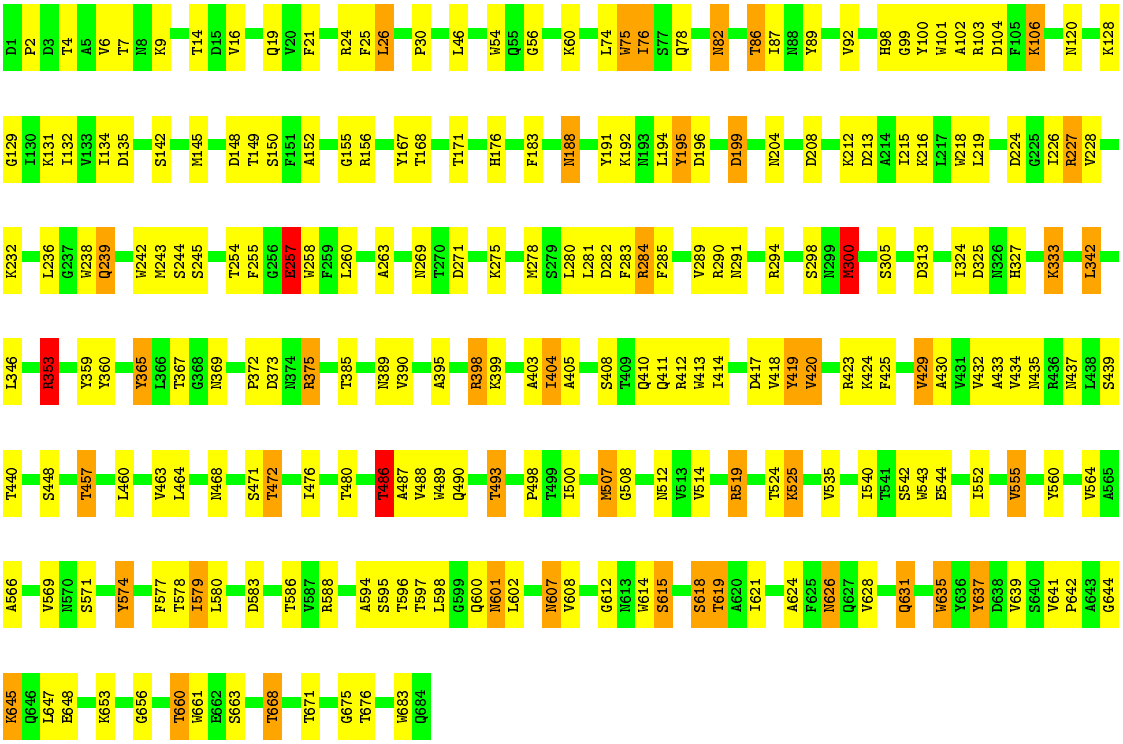
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	478	Total	O	0	0
			478	478		

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.

- Molecule 1: CYCLODEXTRIN GLYCOSYL-TRANSFERASE

Chain A:



- Molecule 2: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain B:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	94.10Å 105.00Å 113.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 2.50 27.27 – 2.56	Depositor EDS
% Data completeness (in resolution range)	(Not available) (10.00-2.50) 62.2 (27.27-2.56)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$	-	Xtriage
Refinement program	X-PLOR	Depositor
R, R_{free}	0.187 , (Not available) 0.248 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	8.3	Xtriage
Anisotropy	0.198	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 35.6	EDS
L-test for twinning[1]	$< \|L\| > = 0.31$, $< L^2 > = 0.15$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.74	EDS
Total number of atoms	5766	wwPDB-VP
Average B, all atoms (Å^2)	13.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.98% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]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: CA, GLC

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	0.97	0/5395	1.82	122/7362 (1.7%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

There are no bond length outliers.

The worst 5 of 122 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	353	ARG	NE-CZ-NH2	-13.62	113.49	120.30
1	A	398	ARG	NE-CZ-NH2	-13.28	113.66	120.30
1	A	412	ARG	NE-CZ-NH2	-12.82	113.89	120.30
1	A	284	ARG	CB-CG-CD	-11.52	81.65	111.60
1	A	365	TYR	CB-CG-CD2	-10.73	114.56	121.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	360	TYR	Sidechain

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	5264	0	5009	104	0
2	B	22	0	19	1	0
3	A	2	0	0	0	0
4	A	478	0	0	18	1
All	All	5766	0	5028	104	1

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 104 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:4:THR:HG23	1:A:399:LYS:HD2	1.61	0.82
1:A:226:ILE:HB	1:A:254:THR:HG23	1.71	0.73
1:A:76:ILE:HD11	1:A:134:ILE:HG22	1.70	0.70
1:A:227:ARG:NH1	1:A:257:GLU:HG2	2.08	0.69
1:A:300:MET:HG3	1:A:419:TYR:HB2	1.75	0.68

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:A:1160:HOH:O	4:A:1235:HOH:O[3_645]	2.11	0.09

5.3 Torsion angles

5.3.1 Protein backbone

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	682/684 (100%)	618 (91%)	58 (8%)	6 (1%)	17 31

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	626	ASN
1	A	607	ASN
1	A	300	MET
1	A	365	TYR
1	A	7	THR

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	565/565 (100%)	513 (91%)	52 (9%)	9 18

5 of 52 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	353	ARG
1	A	439	SER
1	A	645	LYS
1	A	372	PRO
1	A	404	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	98	HIS
1	A	204	ASN
1	A	416	ASN
1	A	570	ASN
1	A	575	ASN

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	GLC	B	1	2	11,11,12	0.96	1 (9%)	15,15,17	2.62	6 (40%)
2	GLC	B	2	2	11,11,12	0.99	1 (9%)	15,15,17	1.81	4 (26%)

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	GLC	B	1	2	-	0/2/19/22	0/1/1/1
2	GLC	B	2	2	-	2/2/19/22	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	GLC	C4-C3	2.01	1.57	1.52
2	B	2	GLC	C2-C3	-2.01	1.49	1.52

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1	GLC	O4-C4-C3	-5.26	98.20	110.35
2	B	1	GLC	C1-C2-C3	-5.19	103.29	109.67
2	B	2	GLC	C1-O5-C5	4.68	118.53	112.19
2	B	1	GLC	C1-O5-C5	4.31	118.03	112.19
2	B	1	GLC	O2-C2-C1	3.59	116.49	109.15

There are no chirality outliers.

All (2) torsion outliers are listed below:

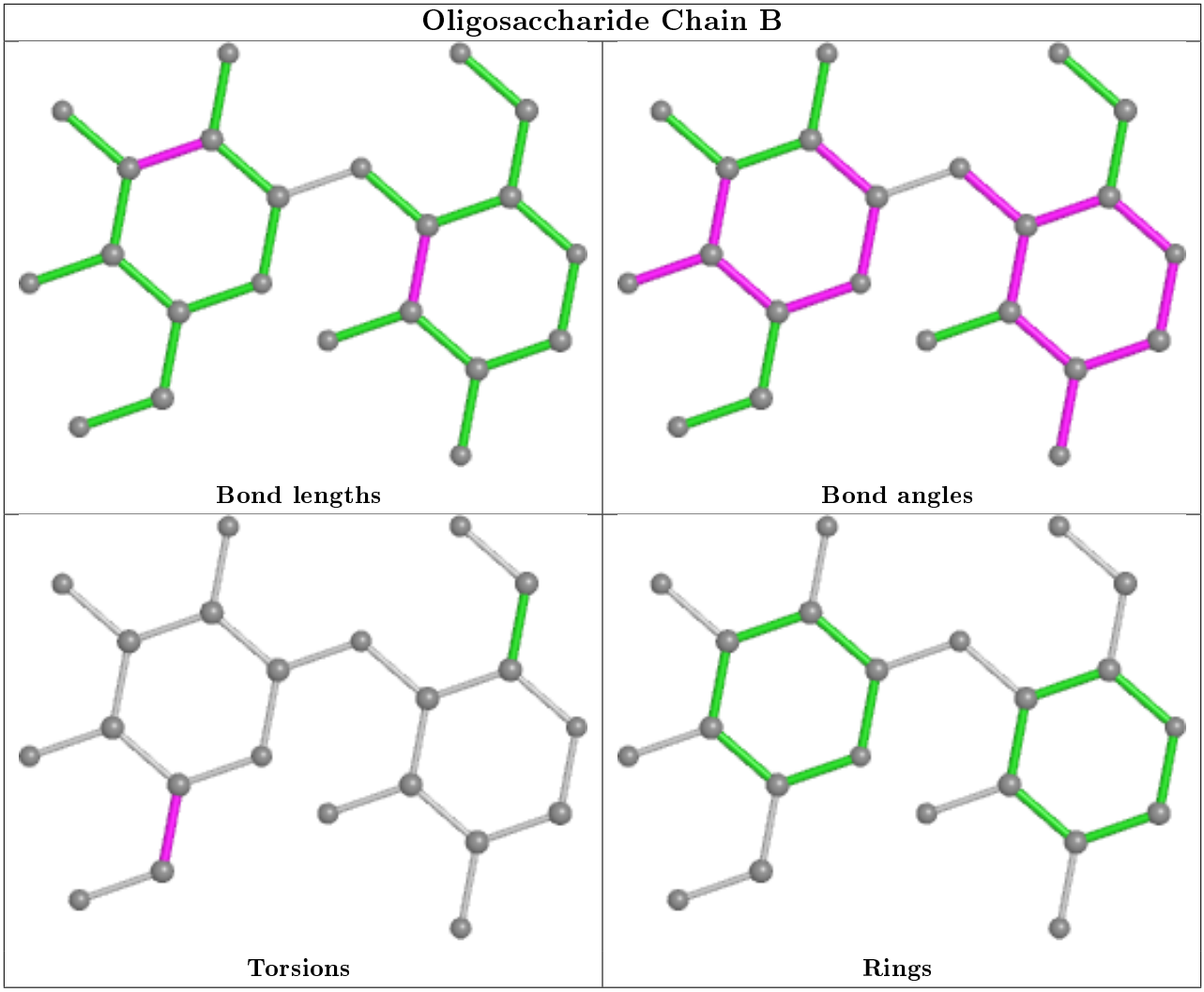
Mol	Chain	Res	Type	Atoms
2	B	2	GLC	O5-C5-C6-O6
2	B	2	GLC	C4-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1	GLC	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

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

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

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

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

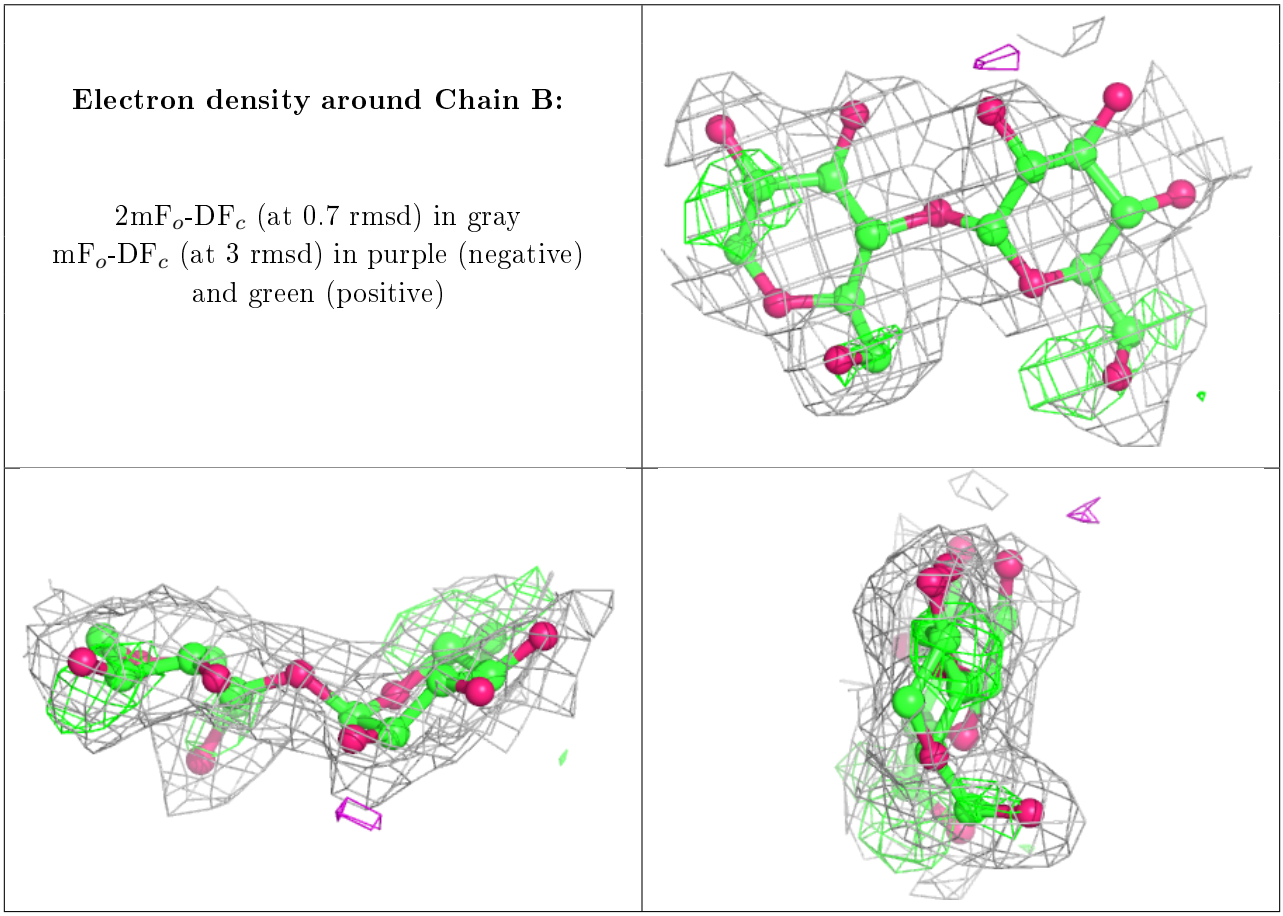
6.3 Carbohydrates

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.

Electron density around Chain B:

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



6.4 Ligands

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

6.5 Other polymers (i)

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