Full wwPDB X-ray Structure Validation Report (i)

Aug 9, 2020 – 07:51 PM BST

PDB ID	:	4MWT
Title	:	Crystal structure of human PPCA (trigonal crystal form 2)
Authors	:	Kolli, N.; Garman, S.C.
Deposited on	:	2013-09-25
Resolution	:	3.85 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

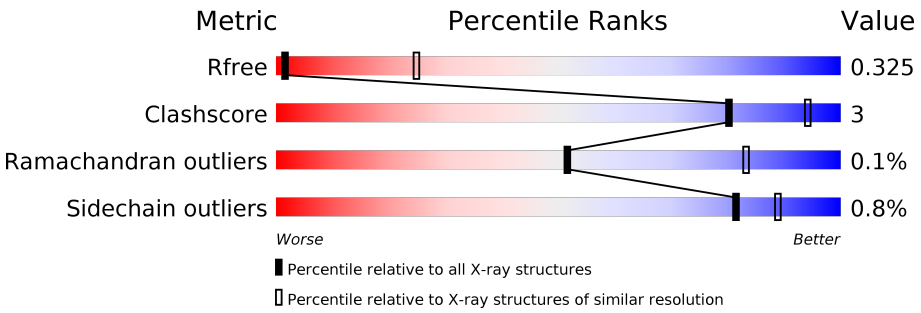
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 3.85 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	1048 (4.10-3.62)
Clashscore	141614	1015 (4.08-3.64)
Ramachandran outliers	138981	1069 (4.10-3.62)
Sidechain outliers	138945	1062 (4.10-3.62)

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	428	90% 7% •
1	B	428	89% 7% •
2	C	3	100%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6677 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 Lysosomal protective protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	413	3298	2112	547	620	19	0	0	0
1	B	413	3298	2112	547	620	19	0	0	0

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	TYR	deletion	UNP P10619
A	?	-	GLU	deletion	UNP P10619
A	?	-	LYS	deletion	UNP P10619
A	?	-	ASP	deletion	UNP P10619
A	?	-	THR	deletion	UNP P10619
A	?	-	VAL	deletion	UNP P10619
A	?	-	VAL	deletion	UNP P10619
A	?	-	VAL	deletion	UNP P10619
A	?	-	GLN	deletion	UNP P10619
A	?	-	ASP	deletion	UNP P10619
A	?	-	LEU	deletion	UNP P10619
A	?	-	GLY	deletion	UNP P10619
A	?	-	ASN	deletion	UNP P10619
A	?	-	ILE	deletion	UNP P10619
A	?	-	PHE	deletion	UNP P10619
A	?	-	THR	deletion	UNP P10619
A	?	-	ARG	deletion	UNP P10619
A	?	-	LEU	deletion	UNP P10619
A	?	-	PRO	deletion	UNP P10619
A	?	-	LEU	deletion	UNP P10619
A	?	-	LYS	deletion	UNP P10619
A	?	-	ARG	deletion	UNP P10619
A	?	-	MET	deletion	UNP P10619
A	?	-	TRP	deletion	UNP P10619
A	?	-	HIS	deletion	UNP P10619

Continued on next page...

Continued from previous page...

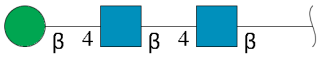
Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	GLN	deletion	UNP P10619
A	?	-	ALA	deletion	UNP P10619
A	?	-	LEU	deletion	UNP P10619
A	?	-	LEU	deletion	UNP P10619
A	?	-	ARG	deletion	UNP P10619
A	453	HIS	-	expression tag	UNP P10619
A	454	HIS	-	expression tag	UNP P10619
A	455	HIS	-	expression tag	UNP P10619
A	456	HIS	-	expression tag	UNP P10619
A	457	HIS	-	expression tag	UNP P10619
A	458	HIS	-	expression tag	UNP P10619
B	?	-	TYR	deletion	UNP P10619
B	?	-	GLU	deletion	UNP P10619
B	?	-	LYS	deletion	UNP P10619
B	?	-	ASP	deletion	UNP P10619
B	?	-	THR	deletion	UNP P10619
B	?	-	VAL	deletion	UNP P10619
B	?	-	VAL	deletion	UNP P10619
B	?	-	VAL	deletion	UNP P10619
B	?	-	GLN	deletion	UNP P10619
B	?	-	ASP	deletion	UNP P10619
B	?	-	LEU	deletion	UNP P10619
B	?	-	GLY	deletion	UNP P10619
B	?	-	ASN	deletion	UNP P10619
B	?	-	ILE	deletion	UNP P10619
B	?	-	PHE	deletion	UNP P10619
B	?	-	THR	deletion	UNP P10619
B	?	-	ARG	deletion	UNP P10619
B	?	-	LEU	deletion	UNP P10619
B	?	-	PRO	deletion	UNP P10619
B	?	-	LEU	deletion	UNP P10619
B	?	-	LYS	deletion	UNP P10619
B	?	-	ARG	deletion	UNP P10619
B	?	-	MET	deletion	UNP P10619
B	?	-	TRP	deletion	UNP P10619
B	?	-	HIS	deletion	UNP P10619
B	?	-	GLN	deletion	UNP P10619
B	?	-	ALA	deletion	UNP P10619
B	?	-	LEU	deletion	UNP P10619
B	?	-	LEU	deletion	UNP P10619
B	?	-	ARG	deletion	UNP P10619
B	453	HIS	-	expression tag	UNP P10619

Continued on next page...

Continued from previous page...

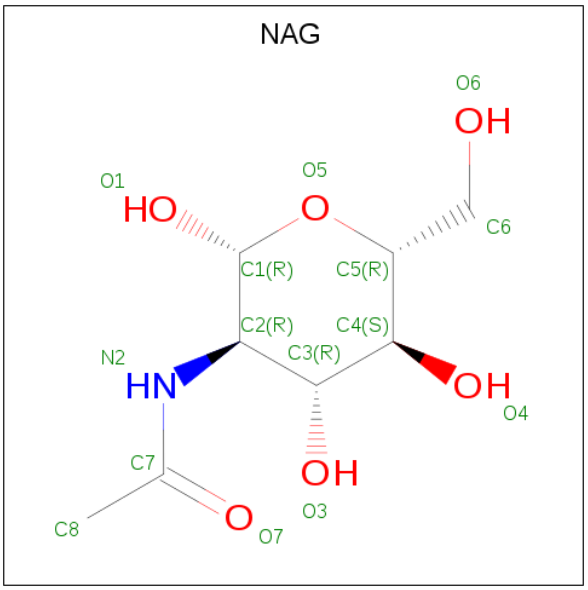
Chain	Residue	Modelled	Actual	Comment	Reference
B	454	HIS	-	expression tag	UNP P10619
B	455	HIS	-	expression tag	UNP P10619
B	456	HIS	-	expression tag	UNP P10619
B	457	HIS	-	expression tag	UNP P10619
B	458	HIS	-	expression tag	UNP P10619

- Molecule 2 is an oligosaccharide called beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	C	3	39	22	2	15	0	0	0

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	14	8	1	5	0	0
3	B	1	14	8	1	5	0	0

Continued on next page...

Continued from previous page...

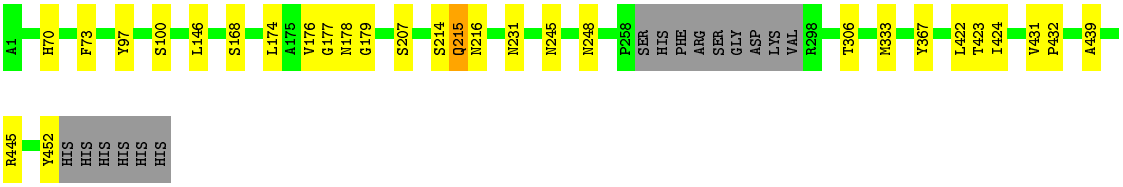
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	B	1	14	8	1	5	0	0

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.

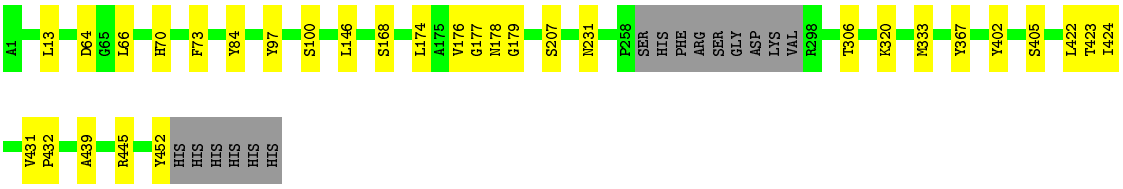
• Molecule 1: Lysosomal protective protein

Chain A:



• Molecule 1: Lysosomal protective protein

Chain B:



• Molecule 2: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain C:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	134.57Å 134.57Å 99.78Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.86 – 3.85 45.86 – 3.85	Depositor EDS
% Data completeness (in resolution range)	95.1 (45.86-3.85) 95.2 (45.86-3.85)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.06 (at 3.88Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.323 , 0.325 0.327 , 0.325	Depositor DCC
R_{free} test set	465 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å^2)	83.3	Xtriage
Anisotropy	0.181	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 43.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.41$, $< L^2 > = 0.23$	Xtriage
Estimated twinning fraction	0.358 for -h,-k,l	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	6677	wwPDB-VP
Average B, all atoms (Å^2)	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.80% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: BMA, NAG

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	0.44	0/3389	0.57	0/4607
1	B	0.42	0/3389	0.57	0/4607
All	All	0.43	0/6778	0.57	0/9214

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	3298	0	3154	17	0
1	B	3298	0	3154	18	0
2	C	39	0	34	0	0
3	A	14	0	13	0	0
3	B	28	0	26	0	0
All	All	6677	0	6381	35	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 3.

All (35) 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:B:367:TYR:HA	1:B:422:LEU:O	2.09	0.52
1:A:367:TYR:HA	1:A:422:LEU:O	2.10	0.51
1:B:445:ARG:NH1	1:B:452:TYR:O	2.43	0.50
1:A:445:ARG:NH1	1:A:452:TYR:O	2.45	0.50
1:B:431:VAL:HB	1:B:432:PRO:HD3	1.93	0.49
1:A:431:VAL:HB	1:A:432:PRO:HD3	1.93	0.49
1:B:13:LEU:HD22	1:B:84:TYR:CE1	2.48	0.48
1:A:207:SER:HB3	1:A:231:ASN:HD21	1.81	0.46
1:B:177:GLY:O	1:B:178:ASN:C	2.55	0.45
1:A:424:ILE:HD12	1:A:424:ILE:N	2.32	0.44
1:A:146:LEU:O	1:A:174:LEU:HA	2.17	0.44
1:B:424:ILE:HD12	1:B:424:ILE:N	2.33	0.43
1:A:306:THR:HG22	1:A:333:MET:CE	2.49	0.43
1:B:146:LEU:O	1:B:174:LEU:HA	2.19	0.43
1:A:245:ASN:HB3	1:A:248:ASN:O	2.19	0.43
1:A:423:THR:C	1:A:424:ILE:HD12	2.39	0.43
1:B:306:THR:HG22	1:B:333:MET:CE	2.49	0.43
1:A:215:GLN:O	1:A:216:ASN:HB2	2.18	0.42
1:B:13:LEU:HD22	1:B:84:TYR:HE1	1.83	0.42
1:B:423:THR:C	1:B:424:ILE:HD12	2.40	0.42
1:A:177:GLY:O	1:A:178:ASN:C	2.58	0.42
1:B:320:LYS:HB3	1:B:320:LYS:HE3	1.82	0.42
1:B:207:SER:HB3	1:B:231:ASN:HD21	1.84	0.42
1:A:207:SER:HB3	1:A:231:ASN:ND2	2.35	0.42
1:B:176:VAL:CG1	1:B:179:GLY:HA3	2.50	0.41
1:A:70:HIS:HD2	1:A:97:TYR:OH	2.03	0.41
1:B:402:TYR:HB2	1:B:405:SER:HB2	2.03	0.41
1:B:70:HIS:HD2	1:B:97:TYR:OH	2.04	0.41
1:A:431:VAL:HG12	1:A:439:ALA:HB2	2.03	0.41
1:A:176:VAL:CG1	1:A:179:GLY:HA3	2.51	0.41
1:B:431:VAL:HG12	1:B:439:ALA:HB2	2.02	0.41
1:B:66:LEU:HD23	1:B:66:LEU:C	2.42	0.40
1:B:64:ASP:OD1	1:B:333:MET:HE2	2.21	0.40
1:A:214:SER:C	1:A:215:GLN:HG2	2.41	0.40
1:A:306:THR:HG22	1:A:333:MET:HE3	2.03	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	409/428 (96%)	387 (95%)	21 (5%)	1 (0%)	47 78
1	B	409/428 (96%)	389 (95%)	20 (5%)	0	100 100
All	All	818/856 (96%)	776 (95%)	41 (5%)	1 (0%)	51 83

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	215	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	360/374 (96%)	357 (99%)	3 (1%)	81 89
1	B	360/374 (96%)	357 (99%)	3 (1%)	81 89
All	All	720/748 (96%)	714 (99%)	6 (1%)	81 89

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	73	PHE
1	A	100	SER
1	A	168	SER
1	B	73	PHE
1	B	100	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	168	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	70	HIS
1	A	141	ASN
1	A	209	GLN
1	A	233	GLN
1	B	70	HIS
1	B	233	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)

3 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	NAG	C	1	1,2	14,14,15	0.79	0	17,19,21	1.49	4 (23%)
2	NAG	C	2	2	14,14,15	0.66	0	17,19,21	1.25	2 (11%)
2	BMA	C	3	2	11,11,12	0.53	0	15,15,17	1.72	3 (20%)

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	NAG	C	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	C	2	2	-	2/6/23/26	0/1/1/1
2	BMA	C	3	2	-	1/2/19/22	0/1/1/1

There are no bond length outliers.

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	3	BMA	C3-C4-C5	4.28	117.86	110.24
2	C	3	BMA	O5-C1-C2	-3.60	105.21	110.77
2	C	1	NAG	O5-C1-C2	-2.73	106.97	111.29
2	C	2	NAG	O5-C1-C2	-2.63	107.13	111.29
2	C	1	NAG	O5-C5-C6	2.46	111.07	107.20
2	C	1	NAG	O3-C3-C4	-2.43	104.74	110.35
2	C	1	NAG	C3-C4-C5	-2.41	105.93	110.24
2	C	2	NAG	O7-C7-C8	-2.32	117.75	122.06
2	C	3	BMA	C2-C3-C4	2.15	114.61	110.89

There are no chirality outliers.

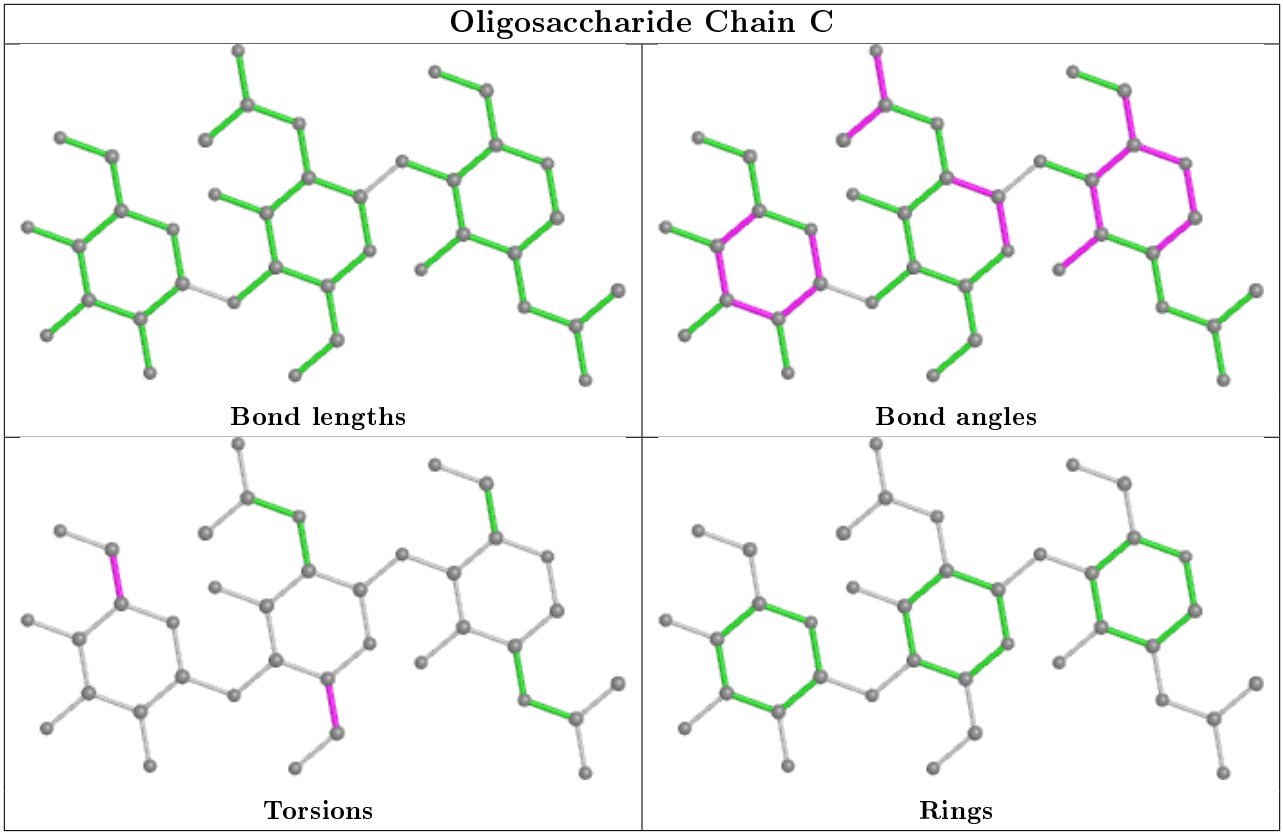
All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	2	NAG	O5-C5-C6-O6
2	C	3	BMA	O5-C5-C6-O6
2	C	2	NAG	C4-C5-C6-O6

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	NAG	A	504	1	14,14,15	0.49	0	17,19,21	1.07	2 (11%)
3	NAG	B	502	1	14,14,15	0.77	0	17,19,21	0.73	0
3	NAG	B	501	1	14,14,15	0.66	0	17,19,21	0.89	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
3	NAG	A	504	1	-	2/6/23/26	0/1/1/1
3	NAG	B	502	1	-	1/6/23/26	0/1/1/1
3	NAG	B	501	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	504	NAG	O5-C1-C2	-2.41	107.48	111.29
3	A	504	NAG	O5-C5-C6	2.20	110.66	107.20

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	504	NAG	O5-C5-C6-O6
3	A	504	NAG	C4-C5-C6-O6
3	B	502	NAG	O5-C5-C6-O6

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.

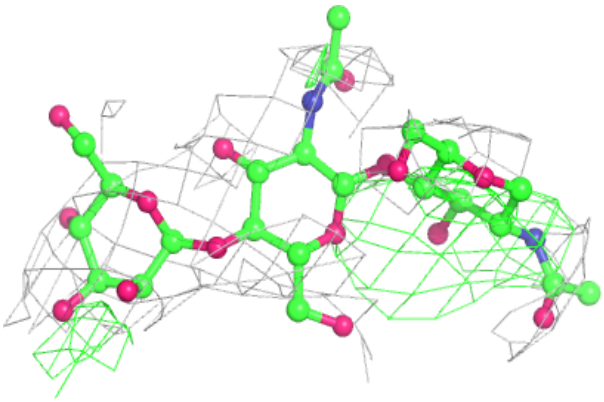
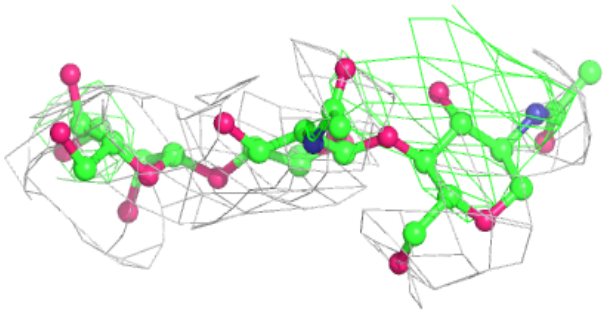
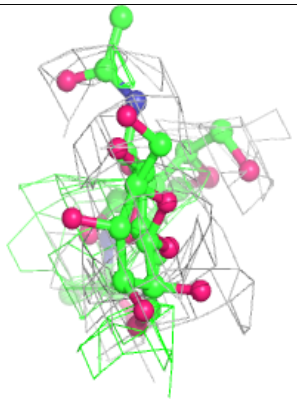
6.2 Non-standard residues in protein, DNA, RNA chains (i)

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.

Electron density around Chain C: $2mF_o$-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)

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.