Full wwPDB X-ray Structure Validation Report ⓘ

Nov 24, 2024 – 12:08 am GMT

PDB ID	:	8R71
Title	:	Polysaccharide lyase BtPL33HA (BT4410) Y291A with HAdp4 collected at 1.22 A
Authors	:	Cartmell, A.
Deposited on	:	2023-11-23
Resolution	:	1.92 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:

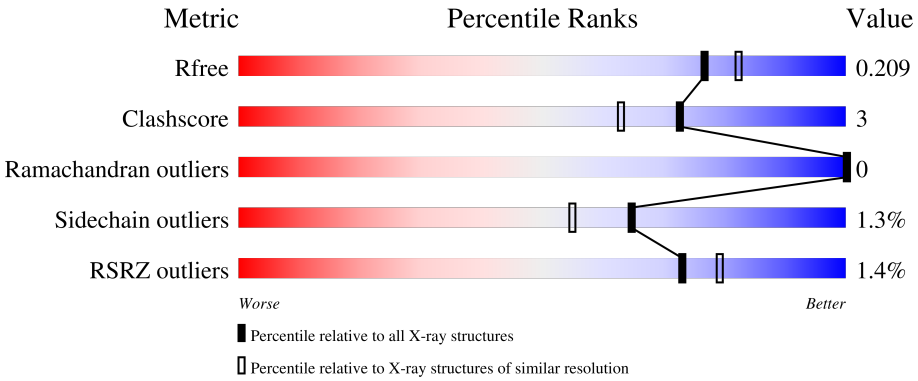
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

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 1.92 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}	164625	1028 (1.92-1.92)
Clashscore	180529	1100 (1.92-1.92)
Ramachandran outliers	177936	1087 (1.92-1.92)
Sidechain outliers	177891	1087 (1.92-1.92)
RSRZ outliers	164620	1028 (1.92-1.92)

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 <=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	644	89% 7% • •
1	B	644	87% 6% • 6%
2	C	3	100%
2	D	3	33% 67%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 20415 atoms, of which 9780 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 Heparinase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	618	Total	C	H	N	O	S	125	6	0
			9958	3197	4930	862	935	34			
1	B	603	Total	C	H	N	O	S	119	2	0
			9628	3093	4764	829	908	34			

There are 48 discrepancies between the modelled and reference sequences:

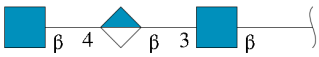
Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	MET	-	initiating methionine	UNP Q89ZG7
A	-6	GLY	-	expression tag	UNP Q89ZG7
A	-5	SER	-	expression tag	UNP Q89ZG7
A	-4	SER	-	expression tag	UNP Q89ZG7
A	-3	HIS	-	expression tag	UNP Q89ZG7
A	-2	HIS	-	expression tag	UNP Q89ZG7
A	-1	HIS	-	expression tag	UNP Q89ZG7
A	0	HIS	-	expression tag	UNP Q89ZG7
A	1	HIS	-	expression tag	UNP Q89ZG7
A	2	HIS	-	expression tag	UNP Q89ZG7
A	3	SER	-	expression tag	UNP Q89ZG7
A	4	SER	-	expression tag	UNP Q89ZG7
A	5	GLY	-	expression tag	UNP Q89ZG7
A	6	LEU	-	expression tag	UNP Q89ZG7
A	7	VAL	-	expression tag	UNP Q89ZG7
A	8	PRO	-	expression tag	UNP Q89ZG7
A	9	ARG	-	expression tag	UNP Q89ZG7
A	10	GLY	-	expression tag	UNP Q89ZG7
A	11	SER	-	expression tag	UNP Q89ZG7
A	12	HIS	-	expression tag	UNP Q89ZG7
A	13	MET	-	expression tag	UNP Q89ZG7
A	14	ALA	-	expression tag	UNP Q89ZG7
A	15	SER	-	expression tag	UNP Q89ZG7
A	283	ALA	TYR	engineered mutation	UNP Q89ZG7
B	-7	MET	-	initiating methionine	UNP Q89ZG7

Continued on next page...

Continued from previous page...

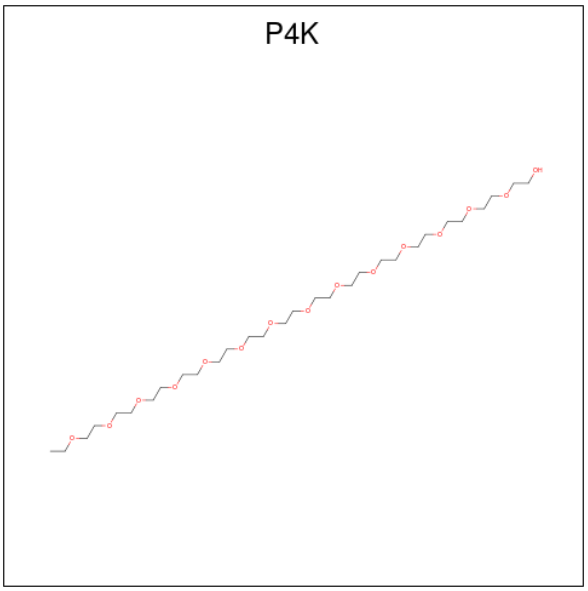
Chain	Residue	Modelled	Actual	Comment	Reference
B	-6	GLY	-	expression tag	UNP Q89ZG7
B	-5	SER	-	expression tag	UNP Q89ZG7
B	-4	SER	-	expression tag	UNP Q89ZG7
B	-3	HIS	-	expression tag	UNP Q89ZG7
B	-2	HIS	-	expression tag	UNP Q89ZG7
B	-1	HIS	-	expression tag	UNP Q89ZG7
B	0	HIS	-	expression tag	UNP Q89ZG7
B	1	HIS	-	expression tag	UNP Q89ZG7
B	2	HIS	-	expression tag	UNP Q89ZG7
B	3	SER	-	expression tag	UNP Q89ZG7
B	4	SER	-	expression tag	UNP Q89ZG7
B	5	GLY	-	expression tag	UNP Q89ZG7
B	6	LEU	-	expression tag	UNP Q89ZG7
B	7	VAL	-	expression tag	UNP Q89ZG7
B	8	PRO	-	expression tag	UNP Q89ZG7
B	9	ARG	-	expression tag	UNP Q89ZG7
B	10	GLY	-	expression tag	UNP Q89ZG7
B	11	SER	-	expression tag	UNP Q89ZG7
B	12	HIS	-	expression tag	UNP Q89ZG7
B	13	MET	-	expression tag	UNP Q89ZG7
B	14	ALA	-	expression tag	UNP Q89ZG7
B	15	SER	-	expression tag	UNP Q89ZG7
B	283	ALA	TYR	engineered mutation	UNP Q89ZG7

- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-beta-D-glucopyranuronic acid-(1-3)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	3	Total	C	H	N	O	6	0	0
			71	22	30	2	17			
2	D	3	Total	C	H	N	O	6	0	0
			71	22	30	2	17			

- Molecule 3 is polyethylene glycol (three-letter code: P4K) (formula: $C_{30}H_{62}O_{15}$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C H O 25 7 14 4	1	0
3	B	1	Total C H O 22 6 12 4	0	0

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Zn 1 1	0	0
4	B	1	Total Zn 1 1	0	0

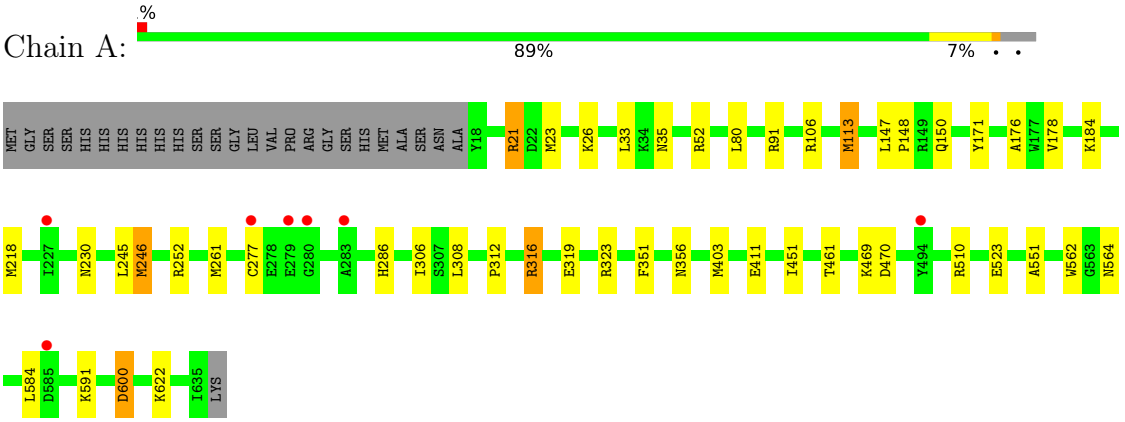
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	408	Total O 408 408	0	0
5	B	230	Total O 230 230	0	0

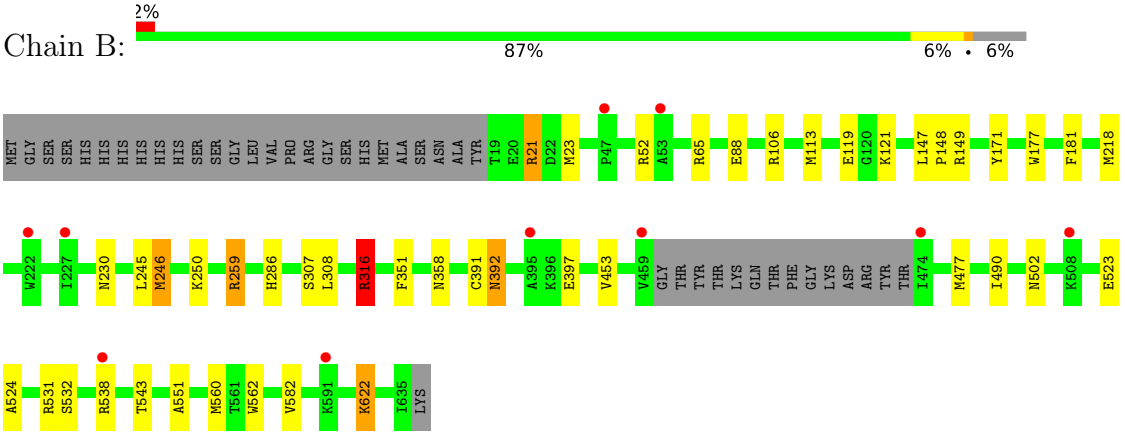
3 Residue-property plots (i)

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 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.

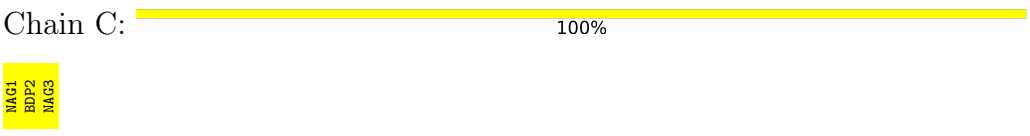
• Molecule 1: Heparinase



• Molecule 1: Heparinase



• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-beta-D-glucopyranuronic acid-(1-3)-2-acetamido-2-deoxy-beta-D-glucopyranose



• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-beta-D-glucopyranuronic acid-(1-3)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	53.63Å 137.35Å 202.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	60.66 – 1.92 60.66 – 1.92	Depositor EDS
% Data completeness (in resolution range)	100.0 (60.66-1.92) 100.0 (60.66-1.92)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.62 (at 1.92Å)	Xtriage
Refinement program	REFMAC 5.8.0419	Depositor
R, R_{free}	0.179 , 0.208 0.179 , 0.209	Depositor DCC
R_{free} test set	5759 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	31.8	Xtriage
Anisotropy	0.252	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 36.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	20415	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.97% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, P4K, BDP, ZN

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	0.51	0/5150	0.93	15/6969 (0.2%)
1	B	0.46	0/4981	0.87	13/6742 (0.2%)
All	All	0.48	0/10131	0.90	28/13711 (0.2%)

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	4
1	B	0	5
All	All	0	9

There are no bond length outliers.

All (28) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	259	ARG	NE-CZ-NH1	10.69	125.64	120.30
1	B	259	ARG	NE-CZ-NH2	-9.31	115.65	120.30
1	A	113	MET	CG-SD-CE	-9.00	85.79	100.20
1	B	316	ARG	NE-CZ-NH2	-8.78	115.91	120.30
1	A	246	MET	CG-SD-CE	8.42	113.67	100.20
1	A	52[A]	ARG	NE-CZ-NH1	7.90	124.25	120.30
1	A	52[B]	ARG	NE-CZ-NH1	7.90	124.25	120.30
1	A	218	MET	CG-SD-CE	-7.79	87.73	100.20
1	B	246	MET	CG-SD-CE	7.54	112.26	100.20
1	B	52	ARG	CD-NE-CZ	6.98	133.38	123.60
1	B	52	ARG	NE-CZ-NH2	6.64	123.62	120.30
1	B	316	ARG	NE-CZ-NH1	6.55	123.58	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	184	LYS	CD-CE-NZ	6.04	125.60	111.70
1	B	218	MET	CG-SD-CE	-5.86	90.83	100.20
1	A	403	MET	CG-SD-CE	-5.74	91.01	100.20
1	A	21	ARG	NE-CZ-NH2	-5.58	117.51	120.30
1	B	560	MET	CG-SD-CE	-5.41	91.54	100.20
1	A	323	ARG	NE-CZ-NH2	-5.40	117.60	120.30
1	B	52	ARG	NE-CZ-NH1	-5.39	117.61	120.30
1	A	80	LEU	CB-CG-CD1	-5.35	101.91	111.00
1	A	91	ARG	NE-CZ-NH2	5.27	122.94	120.30
1	B	21	ARG	NE-CZ-NH2	-5.25	117.68	120.30
1	A	523	GLU	CB-CA-C	-5.18	100.04	110.40
1	B	622	LYS	N-CA-CB	-5.12	101.38	110.60
1	A	600	ASP	CB-CA-C	-5.07	100.26	110.40
1	B	477	MET	CG-SD-CE	5.03	108.25	100.20
1	A	52[A]	ARG	NE-CZ-NH2	-5.01	117.79	120.30
1	A	52[B]	ARG	NE-CZ-NH2	-5.01	117.79	120.30

There are no chirality outliers.

All (9) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	106	ARG	Sidechain
1	A	252	ARG	Sidechain
1	A	316[A]	ARG	Sidechain
1	A	510	ARG	Sidechain
1	B	106	ARG	Sidechain
1	B	21	ARG	Sidechain
1	B	259	ARG	Sidechain
1	B	316	ARG	Sidechain
1	B	531	ARG	Sidechain

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	5028	4930	4916	33	1
1	B	4864	4764	4753	27	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	41	30	33	0	0
2	D	41	30	33	2	0
3	A	11	14	0	1	0
3	B	10	12	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	408	0	0	13	0
5	B	230	0	0	7	1
All	All	10635	9780	9735	61	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 3.

All (61) 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:358:ASN:HB3	5:B:913:HOH:O	1.51	1.10
1:A:35:ASN:HB3	5:A:1071:HOH:O	1.55	1.04
1:B:113:MET:CE	1:B:181:PHE:HD2	1.76	0.97
1:B:113:MET:CE	1:B:181:PHE:CD2	2.54	0.90
1:B:113:MET:HE2	1:B:181:PHE:CD2	2.11	0.86
1:A:600:ASP:OD2	5:A:801:HOH:O	1.98	0.81
1:A:591:LYS:HE3	5:A:1143:HOH:O	1.86	0.74
1:B:245:LEU:HB3	1:B:246:MET:HE3	1.70	0.73
1:A:245:LEU:HB3	1:A:246:MET:CE	2.20	0.71
1:A:245:LEU:HB3	1:A:246:MET:HE3	1.75	0.69
1:A:356:ASN:HB3	5:A:817:HOH:O	1.92	0.68
1:A:451:ILE:HD11	5:A:819:HOH:O	1.95	0.67
1:A:35:ASN:CB	5:A:1071:HOH:O	2.28	0.67
1:A:564:ASN:ND2	5:A:804:HOH:O	2.28	0.66
1:A:246:MET:CE	5:A:1061:HOH:O	2.43	0.66
1:B:246:MET:HE2	5:B:922:HOH:O	1.95	0.66
1:B:23:MET:HE3	1:B:308:LEU:HD13	1.77	0.65
1:B:245:LEU:HB3	1:B:246:MET:CE	2.27	0.63
1:B:316:ARG:NH2	5:B:803:HOH:O	2.29	0.63
1:A:277:CYS:SG	5:A:1140:HOH:O	2.54	0.62
1:A:23:MET:HE3	1:A:308:LEU:HD13	1.82	0.60
1:A:246:MET:HE2	5:A:1061:HOH:O	2.00	0.60
1:B:246:MET:CE	5:B:922:HOH:O	2.50	0.60
1:A:23:MET:CE	1:A:308:LEU:HD13	2.32	0.59
1:B:88:GLU:OE1	5:B:801:HOH:O	2.16	0.59

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:113:MET:HE3	1:B:181:PHE:HD2	1.66	0.59
1:A:356:ASN:CB	5:A:817:HOH:O	2.47	0.59
1:B:113:MET:HE1	1:B:177:TRP:HB3	1.86	0.58
1:B:23:MET:CE	1:B:308:LEU:HD13	2.35	0.55
1:B:230:ASN:HB2	1:B:286:HIS:O	2.08	0.54
1:B:490:ILE:HD13	1:B:523:GLU:HG2	1.90	0.54
5:B:947:HOH:O	2:D:3:NAG:H61	2.08	0.53
1:A:245:LEU:HB3	1:A:246:MET:HE2	1.89	0.52
1:B:490:ILE:HD12	1:B:524:ALA:HA	1.93	0.51
1:A:246:MET:HE1	5:A:1061:HOH:O	2.09	0.51
1:A:461:THR:OG1	5:A:802:HOH:O	2.18	0.49
1:A:469:LYS:O	1:A:470:ASP:HB2	2.13	0.49
1:A:113:MET:HE1	1:A:178:VAL:N	2.28	0.48
1:A:21:ARG:O	1:A:26:LYS:HE2	2.13	0.47
1:B:245:LEU:CB	1:B:246:MET:HE3	2.43	0.47
1:A:147:LEU:O	1:A:150:GLN:HG2	2.14	0.47
1:B:551:ALA:HB1	1:B:622:LYS:HG2	1.96	0.47
1:B:532[C]:SER:OG	1:B:543:THR:OG1	2.08	0.47
1:A:113:MET:CE	1:A:178:VAL:HA	2.45	0.46
1:B:119:GLU:HG2	1:B:121:LYS:HG3	1.97	0.46
1:B:149:ARG:NH2	2:D:1:NAG:O7	2.41	0.46
1:A:176:ALA:HB1	1:A:246:MET:HE3	1.99	0.45
1:A:230:ASN:HB2	1:A:286:HIS:O	2.17	0.45
1:A:147:LEU:N	1:A:148:PRO:CD	2.81	0.44
1:A:551:ALA:HB1	1:A:622:LYS:HG2	1.99	0.43
1:B:453:VAL:HG11	1:B:582:VAL:HG21	1.99	0.43
1:B:147:LEU:N	1:B:148:PRO:CD	2.81	0.43
1:A:319:GLU:OE1	3:A:701:P4K:C12	2.67	0.43
1:B:392:ASN:OD1	1:B:392:ASN:N	2.51	0.42
1:A:312:PRO:HB2	1:A:316[A]:ARG:NH2	2.35	0.42
1:A:33[A]:LEU:HD11	1:A:261:MET:SD	2.60	0.41
1:B:307:SER:HA	5:B:927:HOH:O	2.20	0.41
1:A:469:LYS:O	1:A:470:ASP:CB	2.66	0.41
1:A:33[A]:LEU:HD23	1:A:306:ILE:HD12	2.03	0.41
1:A:584:LEU:HD23	1:A:584:LEU:C	2.42	0.40
1:B:113:MET:CE	1:B:181:PHE:CE2	3.02	0.40

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)
1:A:564:ASN:HB3	5:B:802:HOH:O[3_654]	1.45	0.15

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	622/644 (97%)	609 (98%)	13 (2%)	0	100	100
1	B	602/644 (94%)	585 (97%)	17 (3%)	0	100	100
All	All	1224/1288 (95%)	1194 (98%)	30 (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	536/551 (97%)	532 (99%)	4 (1%)	81	77
1	B	520/551 (94%)	510 (98%)	10 (2%)	52	39
All	All	1056/1102 (96%)	1042 (99%)	14 (1%)	65	55

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	171	TYR
1	A	351	PHE
1	A	411	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	562	TRP
1	B	65	ARG
1	B	171	TYR
1	B	250	LYS
1	B	351	PHE
1	B	391	CYS
1	B	392	ASN
1	B	397	GLU
1	B	502	ASN
1	B	538	ARG
1	B	562	TRP

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

Mol	Chain	Res	Type
1	B	401	HIS
1	B	478	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	NAG	C	1	2	15,15,15	1.23	1 (6%)	21,21,21	1.26	3 (14%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	BDP	C	2	2	12,12,13	1.30	2 (16%)	14,17,19	1.89	4 (28%)
2	NAG	C	3	2	14,14,15	2.72	4 (28%)	17,19,21	2.35	6 (35%)
2	NAG	D	1	2	15,15,15	0.82	0	21,21,21	1.48	2 (9%)
2	BDP	D	2	2	12,12,13	1.13	0	14,17,19	1.53	4 (28%)
2	NAG	D	3	2	14,14,15	2.67	2 (14%)	17,19,21	2.04	7 (41%)

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	2	-	0/6/26/26	0/1/1/1
2	BDP	C	2	2	-	0/4/21/24	0/1/1/1
2	NAG	C	3	2	-	2/6/23/26	0/1/1/1
2	NAG	D	1	2	-	0/6/26/26	0/1/1/1
2	BDP	D	2	2	-	0/4/21/24	0/1/1/1
2	NAG	D	3	2	-	2/6/23/26	0/1/1/1

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	3	NAG	C8-C7	-8.38	1.33	1.50
2	C	3	NAG	C8-C7	-7.95	1.34	1.50
2	D	3	NAG	O7-C7	4.52	1.33	1.23
2	C	3	NAG	O7-C7	4.06	1.32	1.23
2	C	1	NAG	C2-N2	3.14	1.50	1.45
2	C	3	NAG	C1-C2	2.28	1.55	1.52
2	C	2	BDP	O4-C4	2.13	1.48	1.43
2	C	3	NAG	O5-C5	2.06	1.47	1.43
2	C	2	BDP	C1-C2	2.04	1.56	1.52

All (26) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	3	NAG	C8-C7-N2	5.55	125.49	116.10
2	C	2	BDP	C1-C2-C3	4.69	115.43	109.67
2	D	1	NAG	C4-C3-C2	-4.65	103.54	110.34
2	D	3	NAG	C8-C7-N2	4.27	123.33	116.10
2	C	3	NAG	O7-C7-N2	-3.82	114.92	121.95

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	3	NAG	C1-O5-C5	3.63	117.12	112.19
2	D	3	NAG	O5-C1-C2	-3.50	105.77	111.29
2	C	3	NAG	O5-C5-C6	3.36	112.47	107.20
2	C	3	NAG	O5-C1-C2	-3.09	106.41	111.29
2	C	2	BDP	C3-C4-C5	3.00	114.38	109.25
2	D	3	NAG	C4-C3-C2	-2.94	106.71	111.02
2	C	1	NAG	C1-O5-C5	2.93	119.19	113.66
2	D	2	BDP	C1-C2-C3	2.83	113.15	109.67
2	D	2	BDP	C3-C4-C5	2.63	113.75	109.25
2	D	3	NAG	O5-C5-C6	2.57	111.23	107.20
2	C	1	NAG	C2-N2-C7	-2.51	117.07	123.18
2	D	3	NAG	O3-C3-C2	2.39	114.40	109.47
2	D	2	BDP	O5-C1-C2	2.38	114.44	110.77
2	D	3	NAG	C2-N2-C7	2.38	126.29	122.90
2	C	1	NAG	O1-C1-C2	2.35	114.09	109.22
2	C	2	BDP	O6B-C6-C5	2.34	122.22	113.65
2	D	2	BDP	O6B-C6-C5	2.34	122.20	113.65
2	D	1	NAG	C1-O5-C5	-2.22	109.47	113.66
2	C	2	BDP	O5-C1-C2	2.19	114.15	110.77
2	C	3	NAG	C3-C4-C5	2.18	114.13	110.24
2	D	3	NAG	O7-C7-N2	-2.01	118.25	121.95

There are no chirality outliers.

All (4) torsion outliers are listed below:

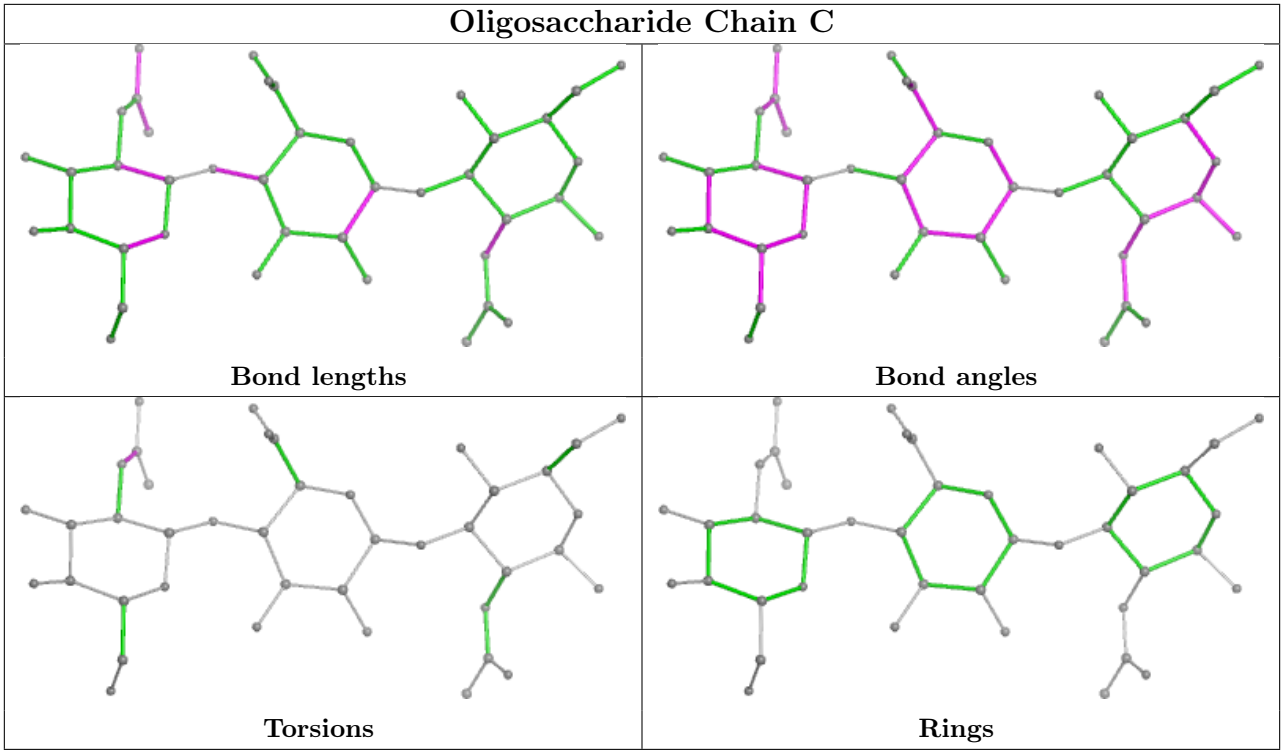
Mol	Chain	Res	Type	Atoms
2	C	3	NAG	C8-C7-N2-C2
2	C	3	NAG	O7-C7-N2-C2
2	D	3	NAG	C8-C7-N2-C2
2	D	3	NAG	O7-C7-N2-C2

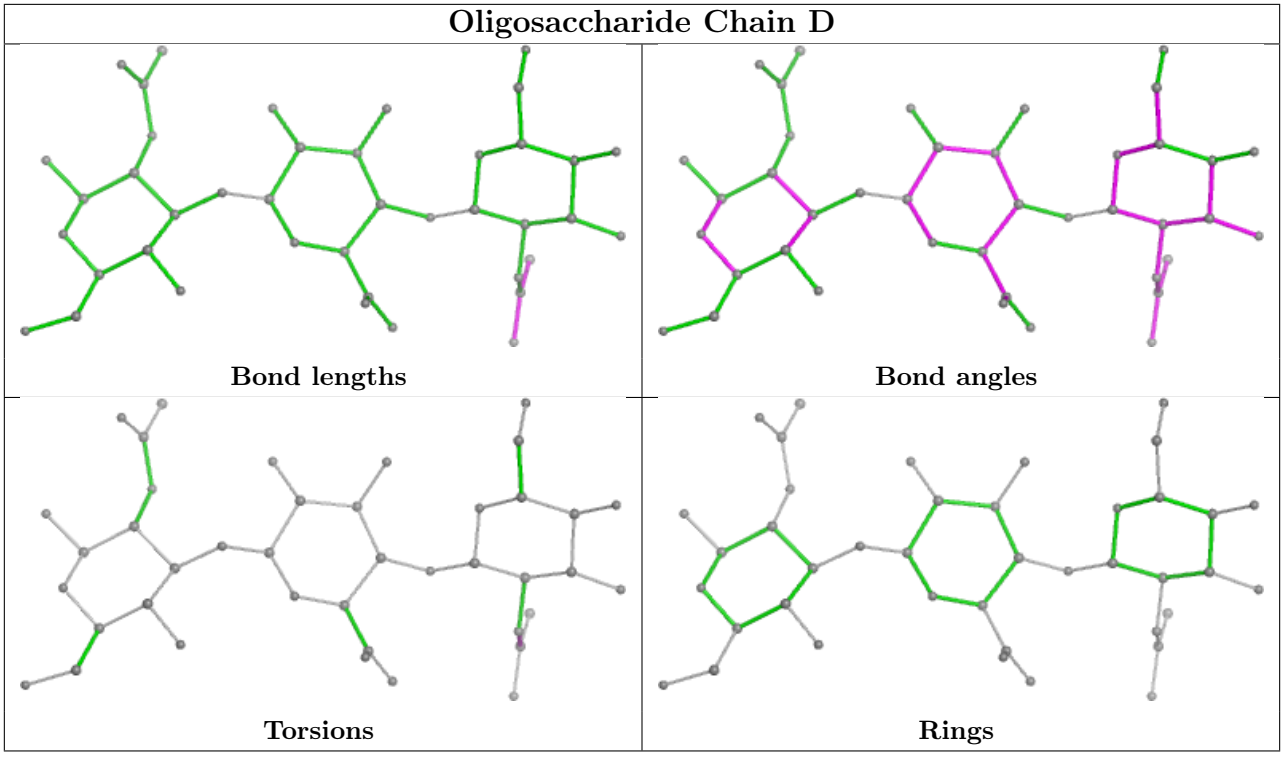
There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	3	NAG	1	0
2	D	1	NAG	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 (i)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 for Mogul analysis.

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	P4K	A	701	-	10,10,44	0.30	0	9,9,43	0.26	0
3	P4K	B	701	-	9,9,44	0.25	0	8,8,43	0.28	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	P4K	A	701	-	-	4/8/8/42	-
3	P4K	B	701	-	-	0/7/7/42	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	701	P4K	C16-C15-O8-C14
3	A	701	P4K	O9-C17-C18-O10
3	A	701	P4K	O7-C13-C14-O8
3	A	701	P4K	O8-C15-C16-O9

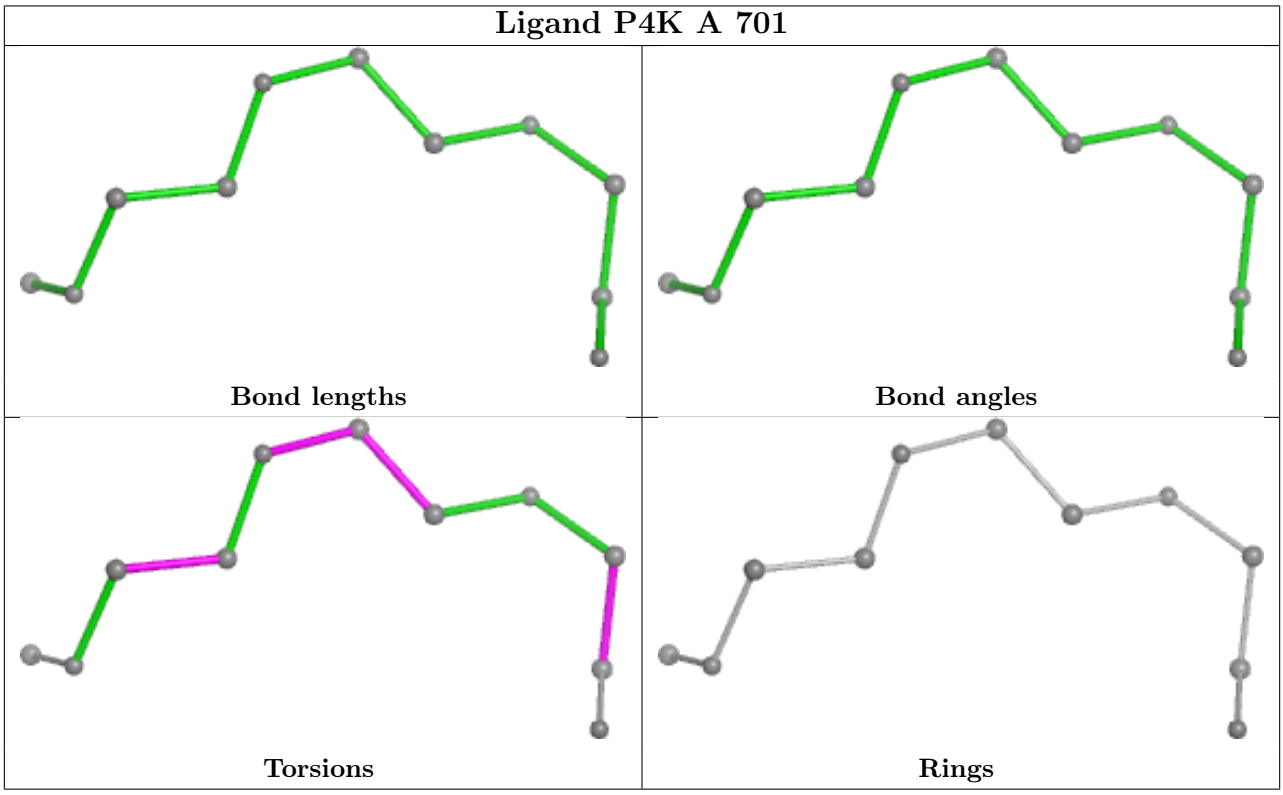
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	701	P4K	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 all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be

highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



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

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^{th} 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^2)	Q<0.9
1	A	618/644 (95%)	-0.43	7 (1%)	77	83	15, 30, 54, 77	6 (0%)
1	B	603/644 (93%)	-0.07	10 (1%)	69	74	8, 39, 63, 104	2 (0%)
All	All	1221/1288 (94%)	-0.25	17 (1%)	73	79	8, 35, 60, 104	8 (0%)

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	585[A]	ASP	3.7
1	B	222	TRP	3.0
1	B	538	ARG	2.8
1	B	459	VAL	2.7
1	B	227	ILE	2.5
1	B	474	ILE	2.5
1	B	53	ALA	2.4
1	A	280	GLY	2.4
1	B	395	ALA	2.3
1	A	283	ALA	2.3
1	B	508	LYS	2.3
1	B	47	PRO	2.1
1	A	227	ILE	2.1
1	B	591	LYS	2.0
1	A	279	GLU	2.0
1	A	494	TYR	2.0
1	A	277	CYS	2.0

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

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

6.3 Carbohydrates (i)

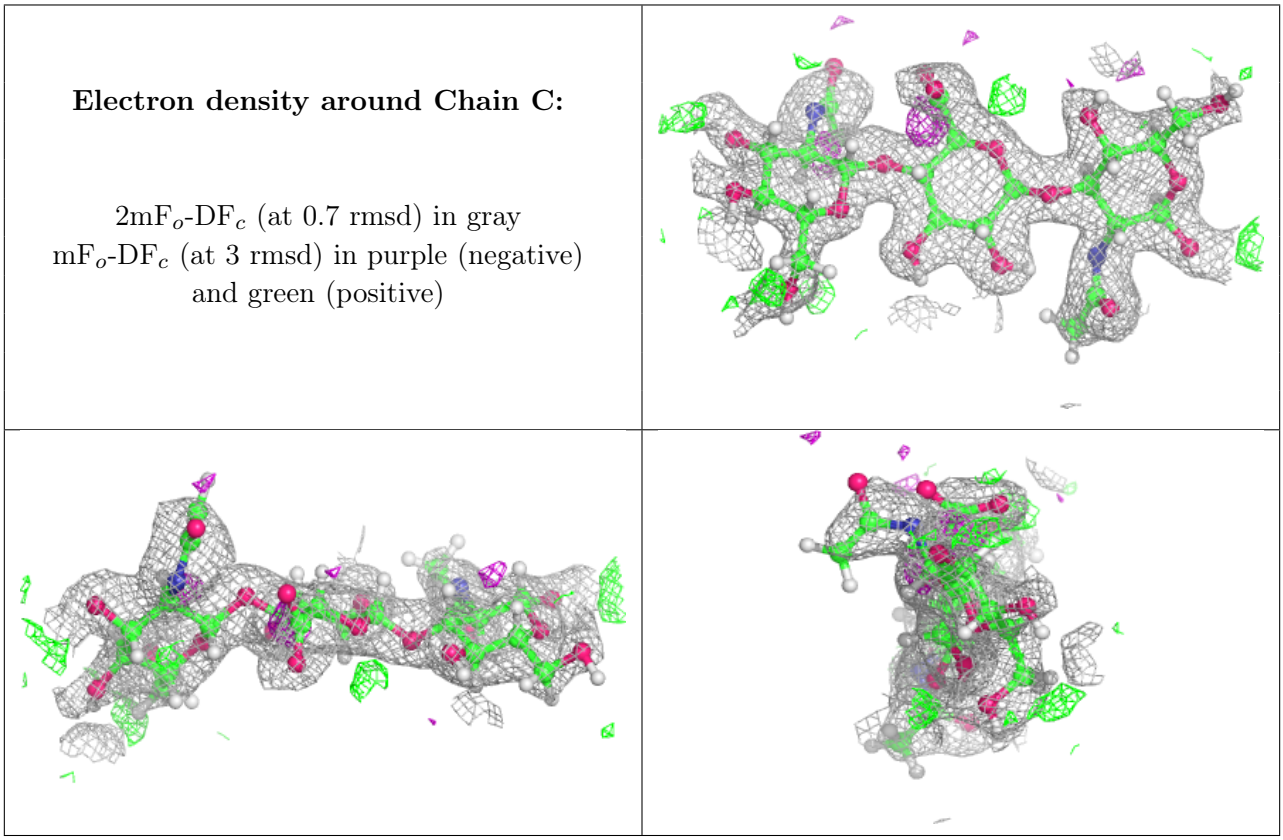
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NAG	C	3	14/15	0.77	0.14	37,54,61,76	2
2	NAG	D	3	14/15	0.78	0.17	37,49,56,59	25
2	BDP	D	2	12/13	0.86	0.13	33,42,49,56	19
2	NAG	C	1	15/15	0.89	0.10	36,44,54,60	2
2	NAG	D	1	15/15	0.90	0.14	37,47,51,54	27
2	BDP	C	2	12/13	0.91	0.09	28,40,55,65	2

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.

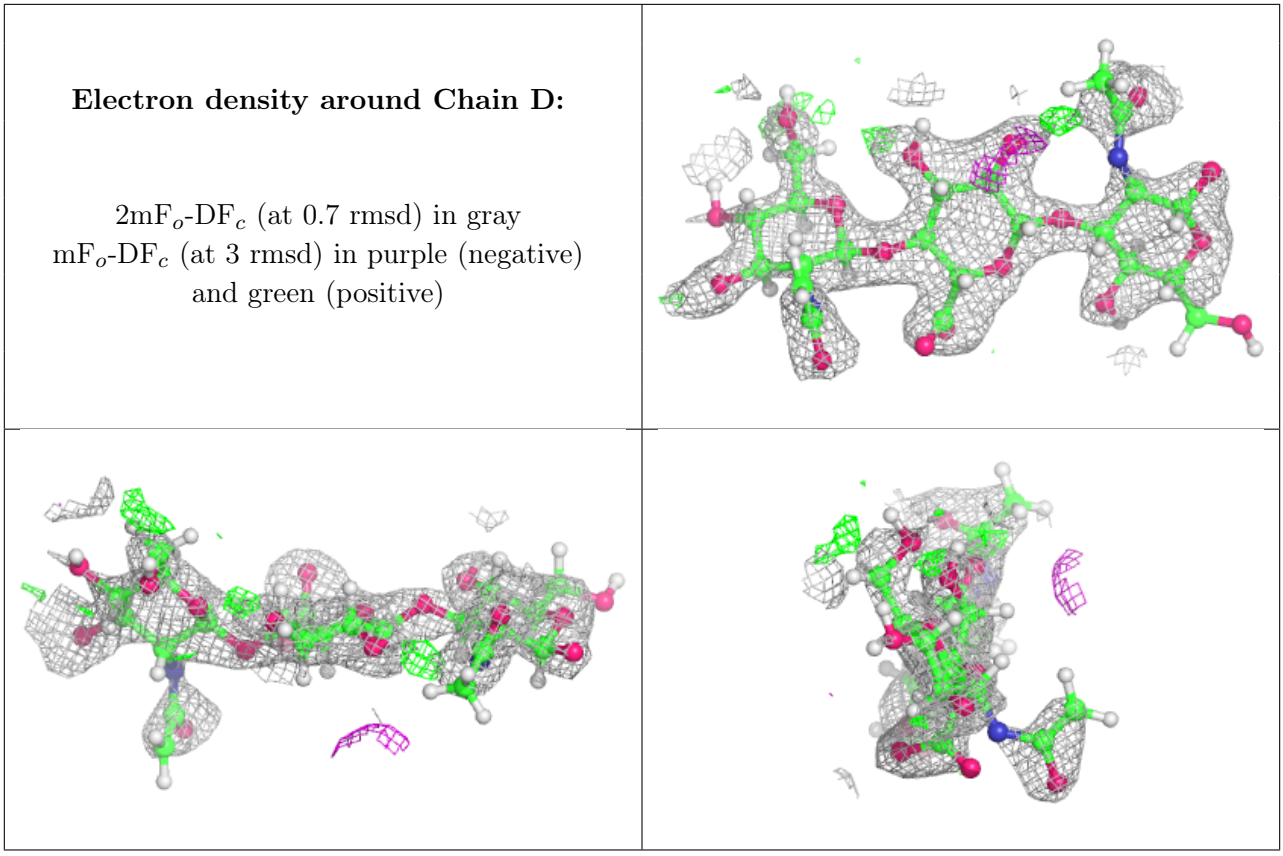
Electron density around Chain C:

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



Electron density around Chain D:

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



6.4 Ligands

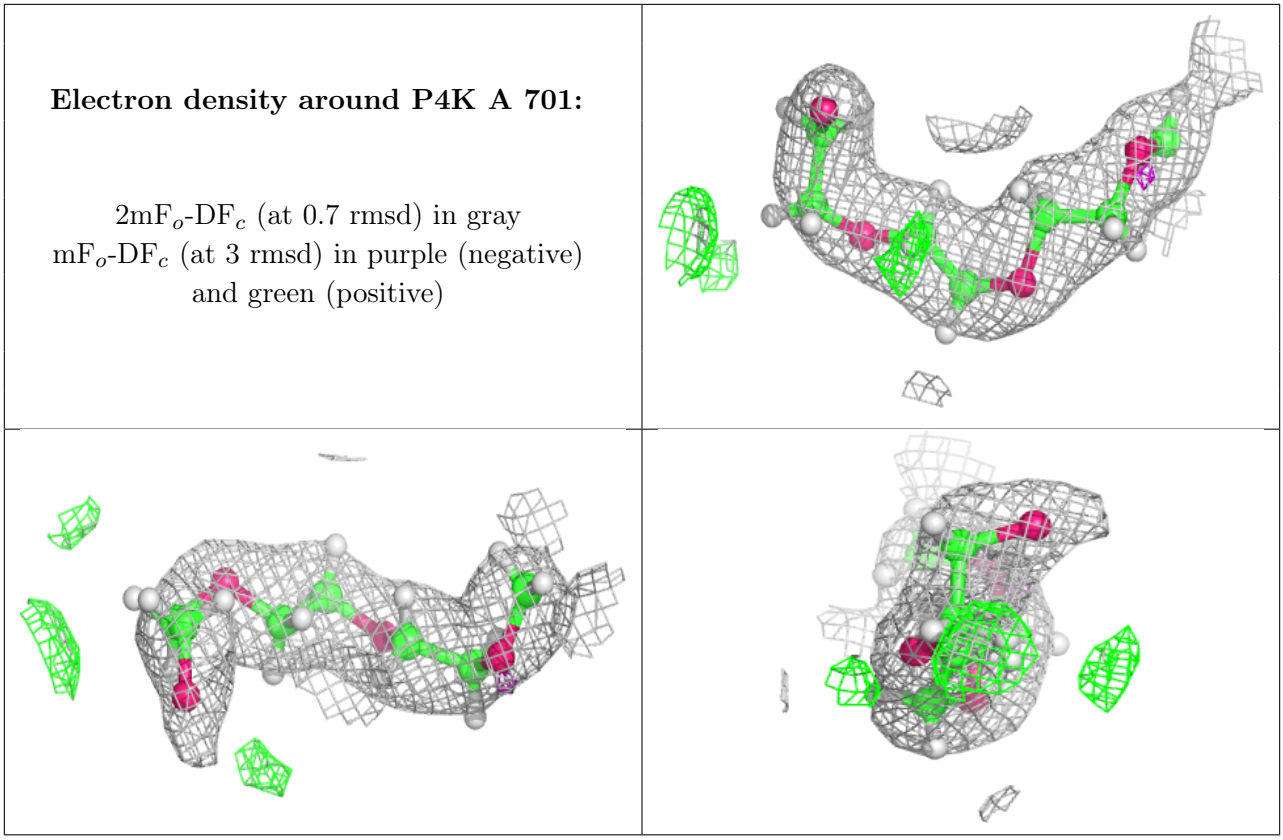
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	P4K	A	701	11/45	0.94	0.08	25,35,47,56	1
3	P4K	B	701	10/45	0.96	0.08	42,46,53,57	0
4	ZN	B	702	1/1	0.99	0.03	34,34,34,34	0
4	ZN	A	702	1/1	1.00	0.02	29,29,29,29	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around P4K A 701:

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



6.5 Other polymers (i)

There are no such residues in this entry.