Full wwPDB X-ray Structure Validation Report (i)

Dec 7, 2023 – 11:01 am GMT

PDB ID	:	1US2
Title	:	Xylanase10C (mutant E385A) from Cellvibrio japonicus in complex with xylopentaose
Authors	:	Pell, G.; Szabo, L.; Charnock, S.J.; Xie, H.; Gloster, T.M.; Davies, G.J.; Gilbert, H.J.
Deposited on	:	2003-11-17
Resolution	:	1.85 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

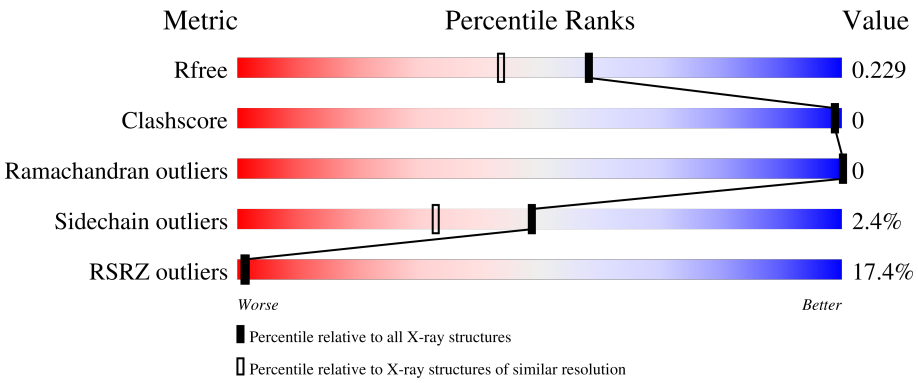
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.85 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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	530	17%, 91%, 5%
2	B	4	100%
2	C	4	100%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	XYP	C	1	-	-	-	X

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4854 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 ENDO-BETA-1,4-XYLANASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	507	3947	2476	665	788	18	21	9	0

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	85	MET	-	expression tag	UNP Q59675
A	186	ALA	GLY	SEE REMARK 999	UNP Q59675
A	536	ALA	GLY	SEE REMARK 999	UNP Q59675
A	594	ALA	GLY	SEE REMARK 999	UNP Q59675
A	385	ALA	GLU	engineered mutation	UNP Q59675
A	607	GLU	-	expression tag	UNP Q59675
A	608	LEU	-	expression tag	UNP Q59675
A	609	HIS	-	expression tag	UNP Q59675
A	610	HIS	-	expression tag	UNP Q59675
A	611	HIS	-	expression tag	UNP Q59675
A	612	HIS	-	expression tag	UNP Q59675
A	613	HIS	-	expression tag	UNP Q59675
A	614	HIS	-	expression tag	UNP Q59675

- Molecule 2 is an oligosaccharide called beta-D-xylopyranose-(1-4)-beta-D-xylopyranose-(1-4)-beta-D-xylopyranose-(1-4)-beta-D-xylopyranose.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
			Total	C	O			
2	B	4	37	20	17	0	0	0
2	C	4	37	20	17	0	0	0

- Molecule 3 is water.

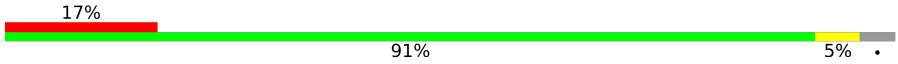
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	833	833	833	0	0

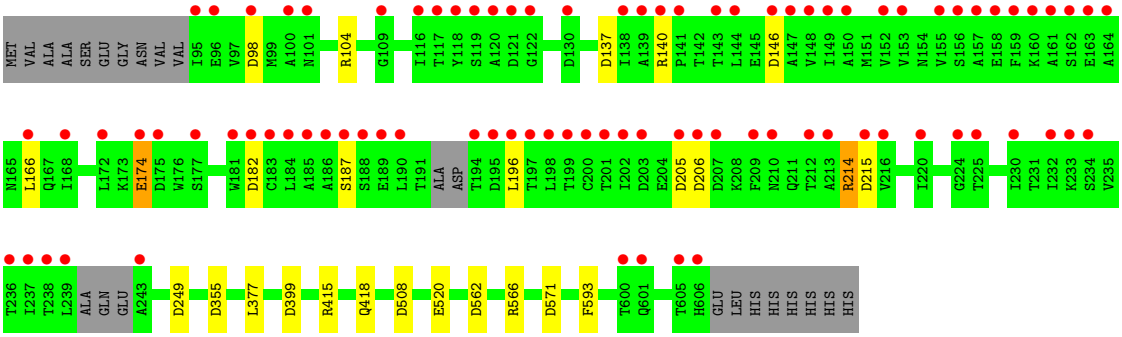
3 Residue-property plots (i)

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: ENDO-BETA-1,4-XYLANASE

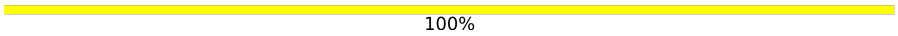
Chain A:





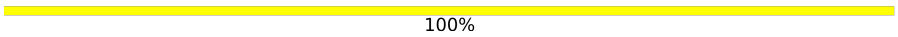
• Molecule 2: beta-D-xylopyranose-(1-4)-beta-D-xylopyranose-(1-4)-beta-D-xylopyranose-(1-4)-beta-D-xylopyranose

Chain B:



• Molecule 2: beta-D-xylopyranose-(1-4)-beta-D-xylopyranose-(1-4)-beta-D-xylopyranose-(1-4)-beta-D-xylopyranose

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	44.60Å 82.69Å 170.73Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	84.52 – 1.85 19.80 – 1.85	Depositor EDS
% Data completeness (in resolution range)	97.6 (84.52-1.85) 97.7 (19.80-1.85)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	7.86 (at 1.85Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.179 , 0.226 0.182 , 0.229	Depositor DCC
R_{free} test set	2682 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	14.3	Xtriage
Anisotropy	0.207	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 72.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4854	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 5.56% 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: XYP

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	0.78	0/4066	0.83	14/5540 (0.3%)

There are no bond length outliers.

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	249[A]	ASP	CB-CG-OD2	7.04	124.64	118.30
1	A	249[B]	ASP	CB-CG-OD2	7.04	124.64	118.30
1	A	205	ASP	CB-CG-OD2	6.06	123.75	118.30
1	A	562	ASP	CB-CG-OD2	6.04	123.74	118.30
1	A	146	ASP	CB-CG-OD2	5.86	123.58	118.30
1	A	206	ASP	CB-CG-OD2	5.50	123.25	118.30
1	A	508[A]	ASP	CB-CG-OD2	5.43	123.18	118.30
1	A	508[B]	ASP	CB-CG-OD2	5.43	123.18	118.30
1	A	399	ASP	CB-CG-OD2	5.36	123.13	118.30
1	A	137	ASP	CB-CG-OD2	5.35	123.11	118.30
1	A	182	ASP	CB-CG-OD2	5.34	123.10	118.30
1	A	215	ASP	CB-CG-OD2	5.23	123.01	118.30
1	A	355	ASP	CB-CG-OD2	5.11	122.90	118.30
1	A	98	ASP	CB-CG-OD2	5.07	122.86	118.30

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	3947	0	3763	3	0
2	B	37	0	0	0	0
2	C	37	0	0	0	0
3	A	833	0	0	0	0
All	All	4854	0	3763	3	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 0.

All (3) 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:566[B]:ARG:HG2	1:A:571:ASP:OD1	2.03	0.58
1:A:415:ARG:HA	1:A:418:GLN:HE21	1.72	0.54
1:A:174:GLU:HB3	1:A:214:ARG:HH11	1.83	0.43

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	510/530 (96%)	490 (96%)	20 (4%)	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	429/438 (98%)	419 (98%)	10 (2%)	50 34

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	104	ARG
1	A	140	ARG
1	A	166	LEU
1	A	174	GLU
1	A	187	SER
1	A	196	LEU
1	A	214	ARG
1	A	377	LEU
1	A	520	GLU
1	A	593	PHE

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

Mol	Chain	Res	Type
1	A	106	ASN
1	A	418	GLN
1	A	454	GLN
1	A	478	ASN
1	A	544	ASN
1	A	585	ASN

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)

8 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	XYP	B	1	2	10,10,10	1.03	1 (10%)	14,14,14	1.28	2 (14%)
2	XYP	B	2	2	9,9,10	1.25	1 (11%)	10,12,14	1.15	1 (10%)
2	XYP	B	3	2	9,9,10	0.95	1 (11%)	10,12,14	0.90	0
2	XYP	B	4	2	9,9,10	1.32	1 (11%)	10,12,14	0.99	1 (10%)
2	XYP	C	1	2	10,10,10	1.54	1 (10%)	14,14,14	0.88	0
2	XYP	C	2	2	9,9,10	1.26	1 (11%)	10,12,14	1.53	2 (20%)
2	XYP	C	3	2	9,9,10	1.30	1 (11%)	10,12,14	0.81	0
2	XYP	C	4	2	9,9,10	1.22	1 (11%)	10,12,14	0.99	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	XYP	B	1	2	-	-	0/1/1/1
2	XYP	B	2	2	-	-	0/1/1/1
2	XYP	B	3	2	-	-	0/1/1/1
2	XYP	B	4	2	-	-	0/1/1/1
2	XYP	C	1	2	-	-	0/1/1/1
2	XYP	C	2	2	-	-	0/1/1/1
2	XYP	C	3	2	-	-	0/1/1/1
2	XYP	C	4	2	-	-	0/1/1/1

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1	XYP	O5-C1	-4.57	1.36	1.43
2	C	3	XYP	O5-C1	-3.51	1.36	1.42

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	4	XYP	O5-C1	-3.46	1.36	1.42
2	C	2	XYP	O5-C1	-3.30	1.36	1.42
2	C	4	XYP	O5-C1	-3.26	1.36	1.42
2	B	2	XYP	O5-C1	-3.23	1.36	1.42
2	B	3	XYP	O5-C1	-2.60	1.37	1.42
2	B	1	XYP	O5-C1	-2.34	1.39	1.43

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	2	XYP	C1-C2-C3	2.76	113.06	109.67
2	B	1	XYP	O4-C4-C5	2.54	114.35	109.15
2	B	1	XYP	C5-C4-C3	-2.53	106.56	109.67
2	B	4	XYP	O2-C2-C3	-2.19	105.75	110.14
2	B	2	XYP	O2-C2-C3	-2.02	106.10	110.14
2	C	2	XYP	C4-C3-C2	2.01	113.31	110.92

There are no chirality outliers.

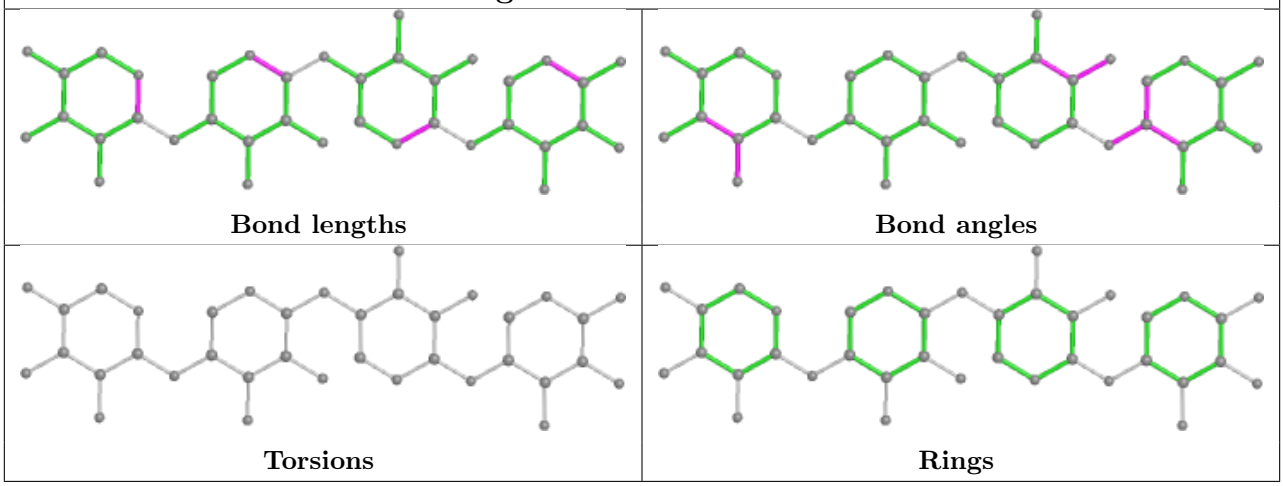
There are no torsion outliers.

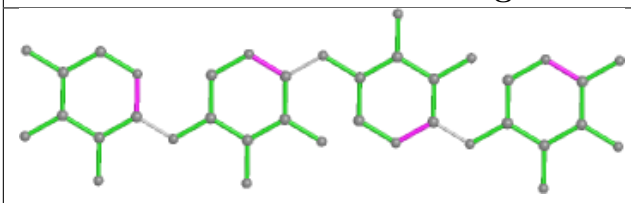
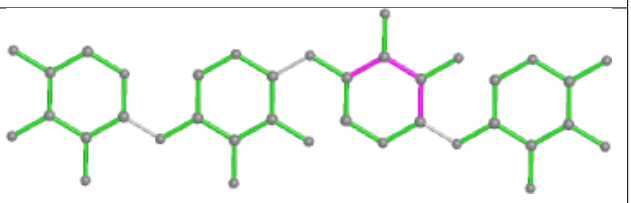
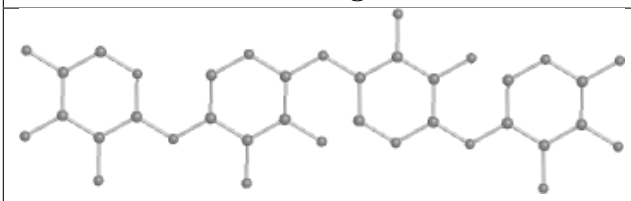
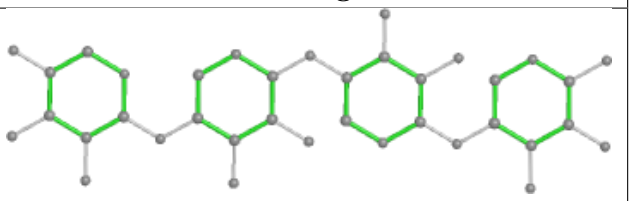
There are no ring outliers.

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.

Oligosaccharide Chain B



Oligosaccharide Chain C	
 Bond lengths	 Bond angles
 Torsions	 Rings

5.6 Ligand geometry (i)

There are no ligands in this entry.

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)

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, 95th 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	507/530 (95%)	0.44	88 (17%) 1 1	3, 9, 20, 38	5 (0%)

All (88) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	120	ALA	8.0
1	A	153	VAL	6.4
1	A	606	HIS	6.3
1	A	157	ALA	6.1
1	A	207	ASP	6.1
1	A	161	ALA	5.9
1	A	150	ALA	5.8
1	A	95	ILE	5.6
1	A	195	ASP	5.1
1	A	117	THR	5.0
1	A	194	THR	4.9
1	A	202	ILE	4.8
1	A	190	LEU	4.8
1	A	205	ASP	4.8
1	A	184	LEU	4.8
1	A	100	ALA	4.5
1	A	196	LEU	4.4
1	A	238	THR	4.4
1	A	188	SER	4.4
1	A	203	ASP	4.3
1	A	199	THR	4.3
1	A	116	ILE	4.2
1	A	232	ILE	4.0
1	A	224	GLY	4.0
1	A	166	LEU	3.8
1	A	206	ASP	3.8
1	A	155	VAL	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	225	THR	3.6
1	A	243	ALA	3.6
1	A	183	CYS	3.5
1	A	213	ALA	3.5
1	A	209	PHE	3.5
1	A	197	THR	3.4
1	A	162	SER	3.3
1	A	139	ALA	3.3
1	A	601	GLN	3.3
1	A	158	GLU	3.3
1	A	144	LEU	3.2
1	A	152	VAL	3.2
1	A	198	LEU	3.1
1	A	147	ALA	3.1
1	A	168	ILE	3.1
1	A	149	ILE	3.1
1	A	189	GLU	3.1
1	A	185	ALA	3.1
1	A	230	ILE	3.0
1	A	172	LEU	3.0
1	A	200	CYS	3.0
1	A	148	VAL	2.9
1	A	216	VAL	2.9
1	A	163	GLU	2.9
1	A	237	ILE	2.9
1	A	182	ASP	2.9
1	A	181	TRP	2.8
1	A	177	SER	2.8
1	A	143	THR	2.8
1	A	239	LEU	2.8
1	A	138	ILE	2.8
1	A	140	ARG	2.7
1	A	233	LYS	2.7
1	A	118	TYR	2.7
1	A	186	ALA	2.7
1	A	146	ASP	2.7
1	A	220	ILE	2.7
1	A	121	ASP	2.6
1	A	212	THR	2.6
1	A	187	SER	2.6
1	A	122	GLY	2.6
1	A	101	ASN	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	236	THR	2.5
1	A	605	THR	2.5
1	A	96	GLU	2.5
1	A	159	PHE	2.4
1	A	175	ASP	2.4
1	A	210	ASN	2.4
1	A	160	LYS	2.3
1	A	164	ALA	2.3
1	A	215	ASP	2.3
1	A	156	SER	2.3
1	A	234	SER	2.3
1	A	141	PRO	2.3
1	A	201	THR	2.3
1	A	174	GLU	2.2
1	A	130	ASP	2.1
1	A	600	THR	2.1
1	A	98	ASP	2.1
1	A	119	SER	2.1
1	A	109	GLY	2.1

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

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

6.3 Carbohydrates (i)

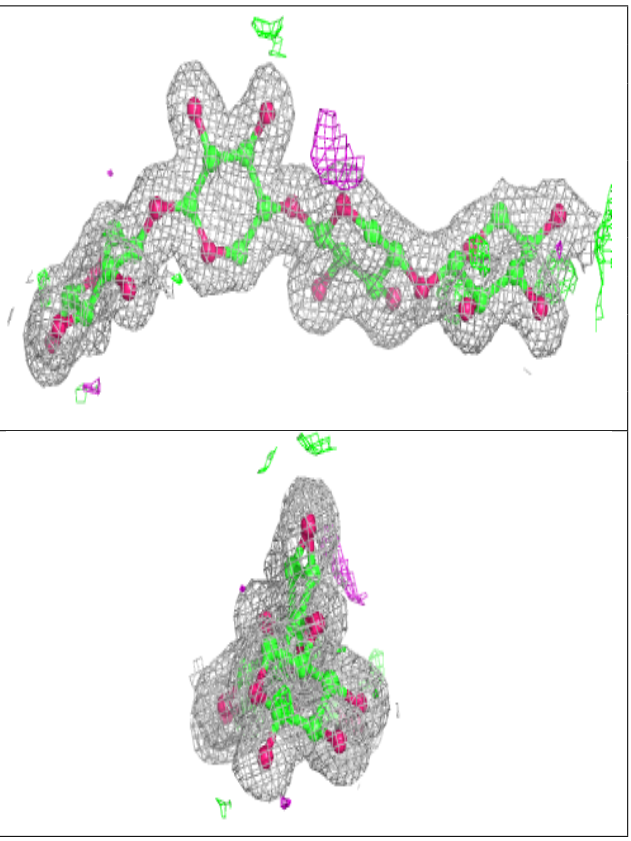
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists 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	XYP	C	4	9/10	0.66	0.39	61,62,63,64	0
2	XYP	C	1	10/10	0.73	0.42	56,60,61,62	0
2	XYP	C	2	9/10	0.79	0.16	50,52,53,54	0
2	XYP	B	4	9/10	0.87	0.17	23,26,28,29	0
2	XYP	C	3	9/10	0.89	0.19	54,55,57,59	0
2	XYP	B	3	9/10	0.97	0.08	16,18,21,22	0
2	XYP	B	1	10/10	0.98	0.06	8,9,11,11	0
2	XYP	B	2	9/10	0.99	0.06	10,12,13,14	0

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.

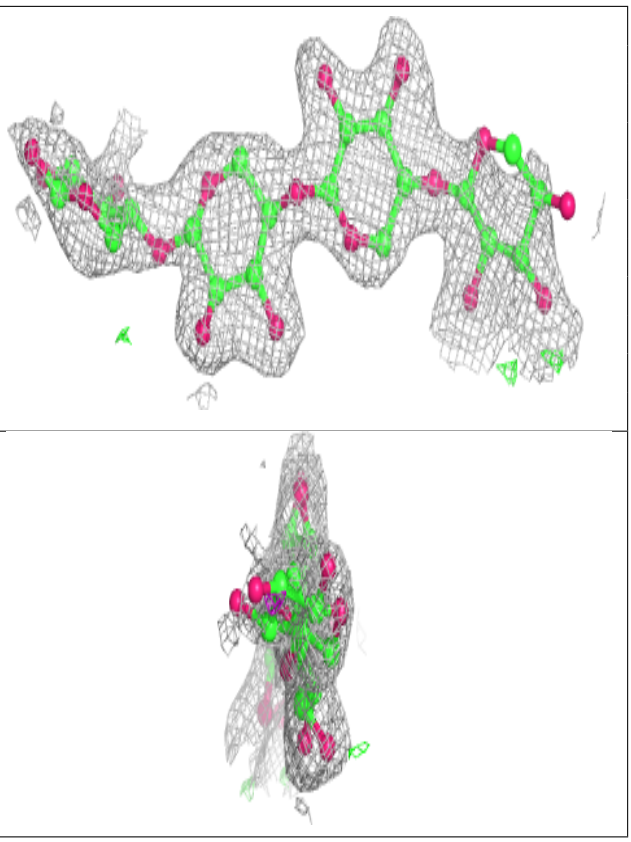
Electron density around Chain B:

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



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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.