Full wwPDB X-ray Structure Validation Report ⓘ

Oct 3, 2023 – 05:44 AM EDT

PDB ID	:	6PHW
Title	:	SpAga D472N structure in complex with melibiose
Authors	:	Pluvinage, B.; Boraston, A.B.
Deposited on	:	2019-06-25
Resolution	:	2.20 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	:	FAILED
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.20 Å.

There are no overall percentile quality scores available for this entry.

MolProbity and EDS failed to run properly - the sequence quality summary graphics cannot be shown.

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
3	TLA	A	804	-	X	-	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 6329 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 Alpha-galactosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	721	5749	3646	975	1107	21	0	1	0

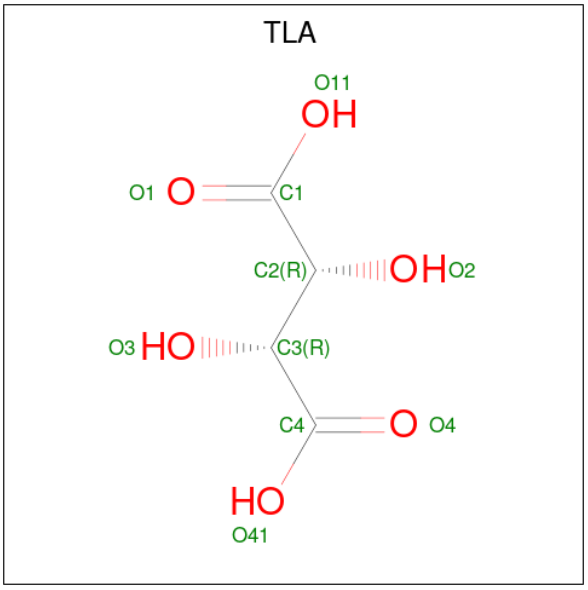
There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	expression tag	UNP A0A0H2URQ6
A	-18	GLY	-	expression tag	UNP A0A0H2URQ6
A	-17	SER	-	expression tag	UNP A0A0H2URQ6
A	-16	SER	-	expression tag	UNP A0A0H2URQ6
A	-15	HIS	-	expression tag	UNP A0A0H2URQ6
A	-14	HIS	-	expression tag	UNP A0A0H2URQ6
A	-13	HIS	-	expression tag	UNP A0A0H2URQ6
A	-12	HIS	-	expression tag	UNP A0A0H2URQ6
A	-11	HIS	-	expression tag	UNP A0A0H2URQ6
A	-10	HIS	-	expression tag	UNP A0A0H2URQ6
A	-9	SER	-	expression tag	UNP A0A0H2URQ6
A	-8	SER	-	expression tag	UNP A0A0H2URQ6
A	-7	GLY	-	expression tag	UNP A0A0H2URQ6
A	-6	LEU	-	expression tag	UNP A0A0H2URQ6
A	-5	VAL	-	expression tag	UNP A0A0H2URQ6
A	-4	PRO	-	expression tag	UNP A0A0H2URQ6
A	-3	ARG	-	expression tag	UNP A0A0H2URQ6
A	-2	GLY	-	expression tag	UNP A0A0H2URQ6
A	-1	SER	-	expression tag	UNP A0A0H2URQ6
A	0	HIS	-	expression tag	UNP A0A0H2URQ6
A	472	ASN	ASP	engineered mutation	UNP A0A0H2URQ6

- Molecule 2 is an oligosaccharide called alpha-D-galactopyranose-(1-6)-alpha-D-glucopyranose.

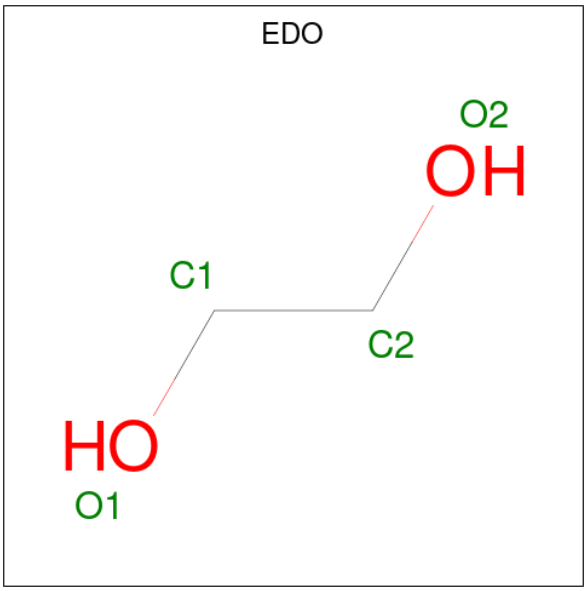
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
2	B	2	Total C O 23 12 11	0	0	0

- Molecule 3 is L(+)-TARTARIC ACID (three-letter code: TLA) (formula: $C_4H_6O_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 10 4 6	0	0
3	A	1	Total C O 10 4 6	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	501	Total O 501 501	0	0

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

3 Data and refinement statistics (i)

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

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	90.91Å 127.19Å 151.53Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	24.75 – 2.20	Depositor
% Data completeness (in resolution range)	98.4 (24.75-2.20)	Depositor
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.60 (at 2.19Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.173 , 0.227	Depositor
Wilson B-factor (Å^2)	32.9	Xtriage
Anisotropy	0.022	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.52$, $< L^2 > = 0.36$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	6329	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.65% 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.

4 Model quality

4.1 Standard geometry

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

4.2 Too-close contacts

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

4.3 Torsion angles

4.3.1 Protein backbone

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

4.3.2 Protein sidechains

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

4.3.3 RNA

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

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

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

4.5 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	12,12,12	0.65	0	17,17,17	1.39	3 (17%)
2	GLA	B	2	2	11,11,12	0.77	0	15,15,17	1.86	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	GLC	B	1	2	-	0/2/22/22	0/1/1/1
2	GLA	B	2	2	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	2	GLA	C1-C2-C3	-5.06	103.45	109.67
2	B	1	GLC	O6-C6-C5	-3.19	100.33	111.29
2	B	1	GLC	O5-C5-C4	2.97	115.09	109.69
2	B	2	GLA	C1-O5-C5	2.16	115.11	112.19
2	B	1	GLC	C1-O5-C5	2.15	117.72	113.66
2	B	2	GLA	O5-C1-C2	-2.02	107.65	110.77

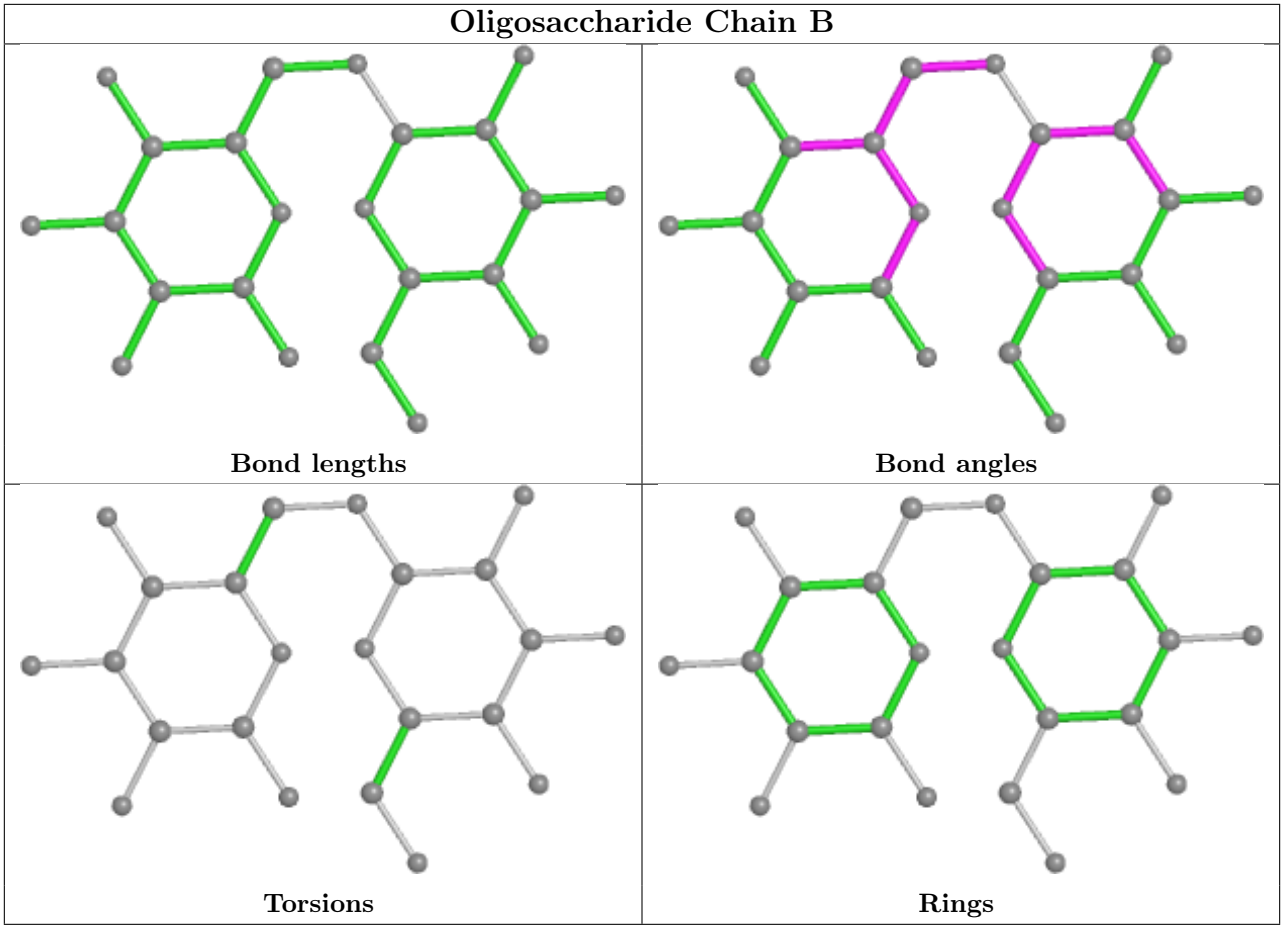
There are no chirality outliers.

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.



4.6 Ligand geometry (i)

11 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	$\#\|Z\| > 2$
4	EDO	A	807	-	3,3,3	0.08	0	2,2,2	0.29	0
4	EDO	A	812	-	3,3,3	0.14	0	2,2,2	0.03	0
4	EDO	A	810	-	3,3,3	0.33	0	2,2,2	1.17	0
4	EDO	A	805	-	3,3,3	0.23	0	2,2,2	0.38	0
3	TLA	A	803	-	9,9,9	1.12	0	12,12,12	1.17	1 (8%)
4	EDO	A	813	-	3,3,3	0.39	0	2,2,2	0.59	0
4	EDO	A	808	-	3,3,3	0.13	0	2,2,2	0.15	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	EDO	A	809	-	3,3,3	0.22	0	2,2,2	0.56	0
3	TLA	A	804	-	9,9,9	1.77	3 (33%)	12,12,12	1.99	4 (33%)
4	EDO	A	811	-	3,3,3	0.17	0	2,2,2	0.51	0
4	EDO	A	806	-	3,3,3	0.14	0	2,2,2	0.32	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
4	EDO	A	807	-	-	1/1/1/1	-
4	EDO	A	812	-	-	0/1/1/1	-
4	EDO	A	810	-	-	1/1/1/1	-
4	EDO	A	805	-	-	0/1/1/1	-
3	TLA	A	803	-	-	0/12/12/12	-
4	EDO	A	813	-	-	1/1/1/1	-
4	EDO	A	808	-	-	0/1/1/1	-
4	EDO	A	809	-	-	1/1/1/1	-
3	TLA	A	804	-	-	8/12/12/12	-
4	EDO	A	811	-	-	1/1/1/1	-
4	EDO	A	806	-	-	1/1/1/1	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	804	TLA	O3-C3	2.73	1.47	1.42
3	A	804	TLA	O4-C4	2.41	1.29	1.22
3	A	804	TLA	C2-C1	2.17	1.55	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	804	TLA	O3-C3-C2	4.17	118.52	110.23
3	A	804	TLA	C3-C2-C1	3.10	116.81	109.87
3	A	804	TLA	O11-C1-C2	3.02	121.42	113.27
3	A	804	TLA	O11-C1-O1	-2.67	118.03	124.09
3	A	803	TLA	O11-C1-C2	2.37	119.67	113.27

There are no chirality outliers.

All (14) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	804	TLA	C1-C2-C3-O3
3	A	804	TLA	O2-C2-C3-O3
3	A	804	TLA	O2-C2-C3-C4
3	A	804	TLA	C1-C2-C3-C4
3	A	804	TLA	O1-C1-C2-O2
3	A	804	TLA	O11-C1-C2-O2
3	A	804	TLA	C2-C3-C4-O41
4	A	806	EDO	O1-C1-C2-O2
4	A	810	EDO	O1-C1-C2-O2
3	A	804	TLA	C2-C3-C4-O4
4	A	809	EDO	O1-C1-C2-O2
4	A	813	EDO	O1-C1-C2-O2
4	A	807	EDO	O1-C1-C2-O2
4	A	811	EDO	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

5 Fit of model and data

5.1 Protein, DNA and RNA chains

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

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

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

5.3 Carbohydrates

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

5.4 Ligands

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

5.5 Other polymers

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