



Full wwPDB X-ray Structure Validation Report ⓘ

Oct 2, 2023 – 01:07 AM EDT

PDB ID : 6MQ4
Title : GH5-4 broad specificity endoglucanase from *Hungateiclostridium cellulosyticum*
Authors : Bingman, C.A.; Smith, R.W.; Glasgow, E.M.; Fox, B.G.
Deposited on : 2018-10-09
Resolution : 1.40 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)
Xtrriage (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

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.40 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.

2 Entry composition i

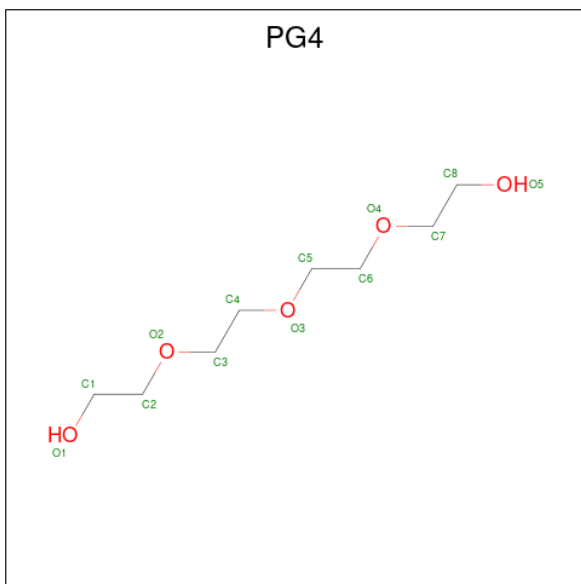
There are 5 unique types of molecules in this entry. The entry contains 6280 atoms, of which 2882 are hydrogens and 0 are deuteriums.

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 cellulase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	350	5618	1806	2768	475	556	13	14	17	0

- Molecule 2 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C₈H₁₈O₅).



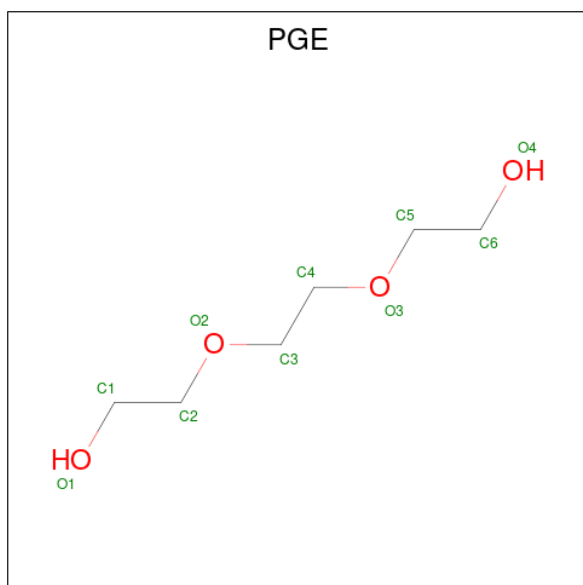
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
2	A	1	31	8	18	5	0	0

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C₄H₁₀O₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
3	A	1	17	4	10	3	0	0
3	A	1	17	4	10	3	0	0
3	A	1	17	4	10	3	0	0
3	A	1	17	4	10	3	0	0

- Molecule 4 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: $C_6H_{14}O_4$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	H	O	0	0
			24	6	14	4		
4	A	1	Total	C	H	O	0	0
			24	6	14	4		
4	A	1	Total	C	H	O	0	0
			24	6	14	4		
4	A	1	Total	C	H	O	0	0
			24	6	14	4		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	446	Total	O	0	21
			467	467		

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	P 31	Depositor
Cell constants a, b, c, α , β , γ	70.85Å 70.85Å 57.61Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.18 – 1.40	Depositor
% Data completeness (in resolution range)	100.0 (30.18-1.40)	Depositor
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.51 (at 1.27Å)	Xtrriage
Refinement program	PHENIX (1.14_3211: ???)	Depositor
R, R_{free}	0.121 , 0.138	Depositor
Wilson B-factor (Å ²)	11.9	Xtrriage
Anisotropy	0.398	Xtrriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.012 for -h,-k,l 0.034 for h,-h-k,-l 0.019 for -k,-h,-l	Xtrriage
Total number of atoms	6280	wwPDB-VP
Average B, all atoms (Å ²)	15.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.97% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

4 Model quality [i](#)

4.1 Standard geometry [i](#)

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

4.2 Too-close contacts [i](#)

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

4.3 Torsion angles [i](#)

4.3.1 Protein backbone [i](#)

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

4.3.2 Protein sidechains [i](#)

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

4.3.3 RNA [i](#)

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

4.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

4.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

4.6 Ligand geometry [i](#)

9 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	PGE	A	406	-	9,9,9	0.35	0	8,8,8	0.43	0
3	PEG	A	403	-	6,6,6	0.48	0	5,5,5	0.33	0
4	PGE	A	408	-	9,9,9	0.33	0	8,8,8	0.58	0
2	PG4	A	401	-	12,12,12	0.50	0	11,11,11	0.46	0
4	PGE	A	407	-	9,9,9	0.33	0	8,8,8	0.37	0
4	PGE	A	409	-	9,9,9	0.35	0	8,8,8	0.44	0
3	PEG	A	405	-	6,6,6	0.52	0	5,5,5	0.66	0
3	PEG	A	402	-	6,6,6	0.49	0	5,5,5	0.41	0
3	PEG	A	404	-	6,6,6	0.47	0	5,5,5	0.41	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	PGE	A	406	-	-	0/7/7/7	-
3	PEG	A	403	-	-	2/4/4/4	-
4	PGE	A	408	-	-	2/7/7/7	-
2	PG4	A	401	-	-	0/10/10/10	-
4	PGE	A	407	-	-	0/7/7/7	-
4	PGE	A	409	-	-	1/7/7/7	-
3	PEG	A	405	-	-	2/4/4/4	-
3	PEG	A	402	-	-	0/4/4/4	-
3	PEG	A	404	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	405	PEG	O1-C1-C2-O2
3	A	404	PEG	O2-C3-C4-O4
4	A	408	PGE	C1-C2-O2-C3
4	A	409	PGE	O3-C5-C6-O4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
3	A	403	PEG	C4-C3-O2-C2
3	A	405	PEG	C1-C2-O2-C3
3	A	404	PEG	C4-C3-O2-C2
4	A	408	PGE	C3-C4-O3-C5
3	A	403	PEG	O2-C3-C4-O4

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 [i](#)

5.1 Protein, DNA and RNA chains [i](#)

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

5.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.3 Carbohydrates [i](#)

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

5.4 Ligands [i](#)

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

5.5 Other polymers [i](#)

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