



# Full wwPDB X-ray Structure Validation Report ⓘ

May 26, 2020 – 11:13 pm BST

PDB ID : 2JEM  
Title : Native family 12 xyloglucanase from *Bacillus licheniformis*  
Authors : Gloster, T.M.; Ibatullin, F.M.; Macauley, K.; Eklof, J.M.; Roberts, S.; Turkenburg, J.P.; Bjornvad, M.E.; Jorgensen, P.L.; Danielsen, S.; Johansen, K.S.; Borchert, T.V.; Wilson, K.S.; Brumer, H.; Davies, G.J.  
Deposited on : 2007-01-18  
Resolution : 1.78 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](mailto: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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
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)	:	2.11

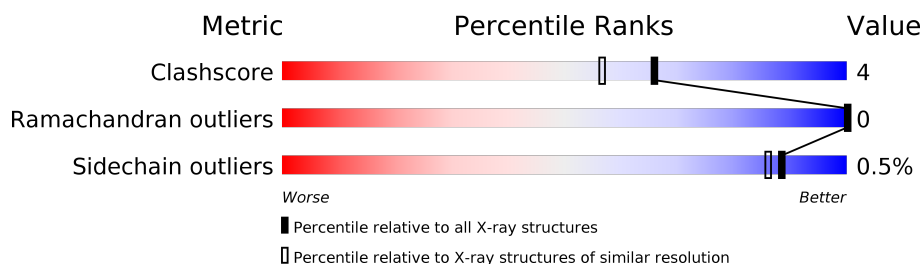
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.78 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))
Clashscore	141614	10184 (1.80-1.76)
Ramachandran outliers	138981	10051 (1.80-1.76)
Sidechain outliers	138945	10050 (1.80-1.76)

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  $\geq 3$ , 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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	261	 82% 6% • 11%
1	B	261	 83% 6% 11%

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4547 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-GLUCANASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	232	Total	C	N	O	S	0	14	0
			1916	1228	319	367	2			
1	B	232	Total	C	N	O	S	0	16	0
			1922	1223	320	377	2			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	155	GLN	GLU	engineered mutation	UNP Q7X4S4
B	155	GLN	GLU	engineered mutation	UNP Q7X4S4

- Molecule 2 is water.


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	364	Total	O	0	0
			364	364		
2	B	345	Total	O	0	0
			345	345		

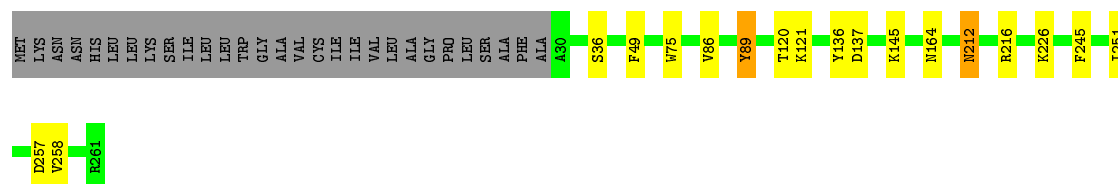
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.


Note EDS was not executed.

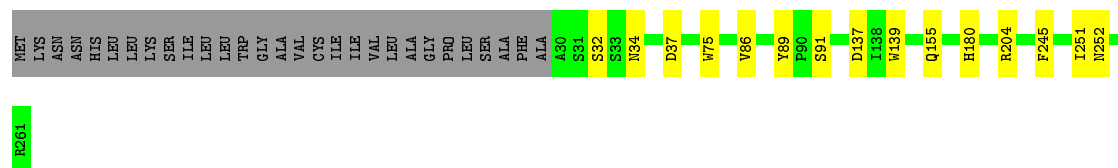
#### • Molecule 1: ENDO-BETA-1,4-GLUCANASE

Chain A: 



#### • Molecule 1: ENDO-BETA-1,4-GLUCANASE

Chain B: 



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	48.75Å 78.13Å 69.10Å 90.00° 109.80° 90.00°	Depositor
Resolution (Å)	64.55 – 1.78	Depositor
% Data completeness (in resolution range)	94.4 (64.55-1.78)	Depositor
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	REFMAC 5.3.0011	Depositor
R, $R_{free}$	0.155 , 0.200	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4547	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0.66	0/2020	0.66	0/2757
1	B	0.67	0/2019	0.67	0/2758
All	All	0.66	0/4039	0.67	0/5515

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	1916	0	1818	18	0
1	B	1922	0	1800	12	0
2	A	364	0	0	6	0
2	B	345	0	0	2	0
All	All	4547	0	3618	30	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 4.

All (30) 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:164[B]:ASN:ND2	2:A:2249:HOH:O	1.72	1.23

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:212[A]:ASN:ND2	2:A:2314:HOH:O	1.76	1.16
1:A:120[B]:THR:HG21	1:A:136:TYR:HE1	1.28	0.97
1:A:120[B]:THR:HG21	1:A:136:TYR:CE1	2.01	0.96
1:A:120[B]:THR:HG23	1:A:258:VAL:HG22	1.44	0.96
1:B:180:HIS:HD2	1:B:204:ARG:HE	1.12	0.89
1:A:121[B]:LYS:HG2	1:A:212[B]:ASN:CG	1.93	0.89
1:A:212[A]:ASN:ND2	2:A:2313:HOH:O	1.99	0.86
1:A:216[A]:ARG:NH2	2:A:2316:HOH:O	2.17	0.78
1:A:145[A]:LYS:HE3	2:A:2227:HOH:O	1.84	0.78
1:B:180:HIS:CD2	1:B:204:ARG:HE	2.03	0.70
1:A:121[B]:LYS:HG2	1:A:212[B]:ASN:ND2	2.08	0.67
1:B:89:TYR:CE2	1:B:91[B]:SER:OG	2.54	0.59
1:A:120[A]:THR:HG22	1:A:258:VAL:HG22	1.85	0.58
1:B:252[B]:ASN:ND2	2:B:2334:HOH:O	2.36	0.58
1:B:86:VAL:HG22	1:B:245:PHE:CE2	2.38	0.58
1:A:216[A]:ARG:NH2	2:A:2319:HOH:O	2.22	0.56
1:B:34:ASN:HB2	1:B:37:ASP:OD2	2.08	0.53
1:A:86:VAL:HG22	1:A:245:PHE:CE2	2.47	0.50
1:A:121[A]:LYS:HE2	1:A:257[A]:ASP:HB2	1.96	0.48
1:A:36:SER:HA	1:A:49:PHE:CE1	2.49	0.47
1:B:139:TRP:CZ3	1:B:155:GLN:HB2	2.51	0.46
1:B:75:TRP:CE2	1:B:251:ILE:HG21	2.52	0.45
1:B:32[A]:SER:HB2	1:B:37:ASP:OD2	2.18	0.44
1:A:89[B]:TYR:OH	1:A:137:ASP:HB2	2.18	0.44
1:B:89:TYR:CE1	2:B:2137:HOH:O	2.68	0.43
1:B:89:TYR:CZ	1:B:91[B]:SER:OG	2.72	0.43
1:B:89:TYR:OH	1:B:137[A]:ASP:HB2	2.20	0.42
1:A:75:TRP:CE2	1:A:251:ILE:HG21	2.55	0.41
1:A:226:LYS:HD2	1:A:226:LYS:HA	1.95	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	244/261 (94%)	237 (97%)	7 (3%)	0	100	100
1	B	246/261 (94%)	241 (98%)	5 (2%)	0	100	100
All	All	490/522 (94%)	478 (98%)	12 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

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	210/219 (96%)	206 (98%)	4 (2%)	57	43
1	B	212/219 (97%)	212 (100%)	0	100	100
All	All	422/438 (96%)	418 (99%)	4 (1%)	88	72

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	89[A]	TYR
1	A	89[B]	TYR
1	A	212[A]	ASN
1	A	212[B]	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	227	GLN
1	B	34	ASN
1	B	180	HIS

### 5.3.3 RNA ⓘ

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](#)

There are no carbohydrates in this entry.

## 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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.