



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 24, 2020 – 02:21 am BST

PDB ID : 1E1E  
Title : Crystal structure of a Monocot (Maize ZMGlu1) beta-glucosidase  
Authors : Czjzek, M.; Cicek, M.; Bevan, D.R.; Henrissat, B.; Esen, A.  
Deposited on : 2000-05-03  
Resolution : 2.50 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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)	:	1.13
EDS	:	2.11
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)	:	2.11

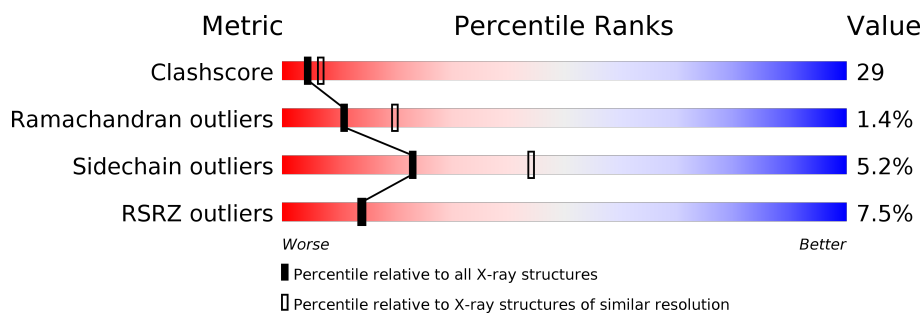
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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  $\leq 5\%$ . 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	512	<div> <div>5%</div> <div>58%</div> <div>34%</div> <div>• •</div> </div>
1	B	512	<div> <div>9%</div> <div>49%</div> <div>43%</div> <div>• •</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 8302 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 BETA-GLUCOSIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	490	Total	C	N	O	S	0	0	0
			3967	2544	656	749	18			
1	B	495	Total	C	N	O	S	0	0	0
			4004	2565	664	757	18			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	215	Total	O	0	0
			215	215		
2	B	116	Total	O	0	0
			116	116		



LYS	T417	G221	R303		
LYS	P418	L224	S304		
ILE	I419	V225	I305		
LEU		E226	R308		
THR	E422	P227	R309		
PRO			I310		
ALA	M426	A230	F311		
	D427	G231	F312		
		H232	F313		
	R430	N233	K314		
	L431	I234	D315		
	A439	L235	F316		
	T440	L236	Q317		
	L441	A237	F318		
	K442	H238	F319		
		A239	K320		
	I445	E240	L321		
		A241			
	V451	V242	Y325		
		D243	N326		
	F455	L244	F327		
	A456	Y245	L328		
	N457	N246	G329		
	S458	K247	I330		
	L459	H248	N331		
	L460	Y249			
	D461	K250	S335		
	M462	R251	R336		
	F463	D252			
		D253	S344		
	G468	T254	P345		
	F469	A255			
	T470	I256	S348		
	E471	G257	F349		
	R472	L258			
	Y473	A259	D355		
	G474	F260			
	I475	D261	K366		
	Y476	V262	F369		
	Y477	N263			
	Y478		N376		
		L273	P377		
	M482				
	M483	Q276	E384		
	C484	E279	G385		
	T485		L386		
	R486	N285			
			L389		
	K489	F289			
	E490	L290	K394		
	S491	E291			
	A492	P292	N399		
		V293	P400		
	T500	V294			
	A501		E406		
	LYS	Y298	N407		
	LYS	P299	G408		
	PRO				
	SER		E416		

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	62.88Å 118.34Å 77.10Å 90.00° 90.30° 90.00°	Depositor
Resolution (Å)	27.00 – 2.50 29.17 – 2.30	Depositor EDS
% Data completeness (in resolution range)	98.2 (27.00-2.50) 97.6 (29.17-2.30)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.08 (at 2.31Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.256 , 0.322 0.274 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	33.7	Xtriage
Anisotropy	0.283	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 47.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.029 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	8302	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.23% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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	0.45	0/4088	0.63	0/5550
1	B	0.47	0/4125	0.62	0/5600
All	All	0.46	0/8213	0.63	0/11150

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 ⓘ

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	3967	0	3759	175	0
1	B	4004	0	3793	273	0
2	A	215	0	0	33	0
2	B	116	0	0	36	0
All	All	8302	0	7552	443	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 29.

The worst 5 of 443 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:376:ASN:HB2	1:A:377:PRO:HD2	1.46	0.98

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:77:TYR:HA	1:B:121:TYR:CZ	2.00	0.96
1:B:146:PRO:HG2	1:B:149:LEU:HG	1.48	0.94
1:A:461:ASP:HA	2:A:2198:HOH:O	1.67	0.94
1:B:49:GLY:O	1:B:102:PRO:HB2	1.65	0.94

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	488/512 (95%)	430 (88%)	50 (10%)	8 (2%)	9	17
1	B	493/512 (96%)	425 (86%)	62 (13%)	6 (1%)	13	24
All	All	981/1024 (96%)	855 (87%)	112 (11%)	14 (1%)	11	20

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	155	GLY
1	A	162	LYS
1	A	307	ARG
1	B	180	GLY
1	A	15	PRO

### 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	423/441 (96%)	405 (96%)	18 (4%)	29	53
1	B	427/441 (97%)	401 (94%)	26 (6%)	18	36
All	All	850/882 (96%)	806 (95%)	44 (5%)	23	44

5 of 44 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	75	ASN
1	B	153	TYR
1	B	458	SER
1	B	86	LEU
1	B	119	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	52	ASN
1	B	115	ASN
1	B	481	ASN
1	B	75	ASN
1	B	127	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	490/512 (95%)	0.30	26 (5%)	26 28	7, 45, 75, 94	1 (0%)
1	B	495/512 (96%)	0.54	48 (9%)	7 7	8, 52, 86, 99	1 (0%)
All	All	985/1024 (96%)	0.42	74 (7%)	14 14	7, 48, 83, 99	2 (0%)

The worst 5 of 74 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	458	SER	7.8
1	B	141	PHE	4.3
1	B	458	SER	4.3
1	A	457	TRP	3.8
1	A	459	LEU	3.7

### 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](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.