



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 12:56 am BST

PDB ID : 3F4V  
Title : Semi-active E176Q mutant of rice BGlu1, a plant exoglucanase/beta-glucosidase  
Authors : Chuenchor, W.; Ketudat Cairns, J.R.; Pengthaisong, S.; Robinson, R.C.; Yuvaniyama, J.; Chen, C.-J.  
Deposited on : 2008-11-03  
Resolution : 1.65 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
Mogul	:	1.8.5 (274361), CSD as541be (2020)
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

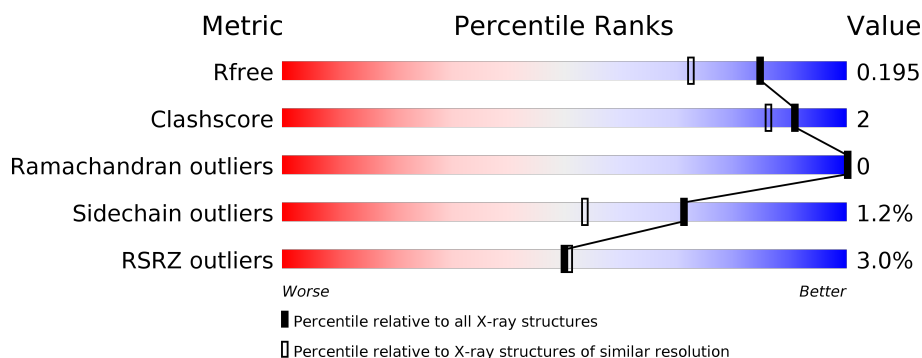
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.65 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	1827 (1.66-1.66)
Clashscore	141614	1931 (1.66-1.66)
Ramachandran outliers	138981	1891 (1.66-1.66)
Sidechain outliers	138945	1891 (1.66-1.66)
RSRZ outliers	127900	1791 (1.66-1.66)

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	481	<div> <div>2%</div> <div> <div></div> <div>93%</div> <div>5%</div> </div> </div>
1	B	481	<div> <div>3%</div> <div> <div></div> <div>94%</div> <div></div> </div> </div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 8572 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	472	Total	C	N	O	S	0	0	0
			3809	2450	656	690	13			
1	B	472	Total	C	N	O	S	0	0	0
			3809	2450	656	690	13			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	ALA	-	EXPRESSION TAG	UNP Q42975
A	-3	MET	-	EXPRESSION TAG	UNP Q42975
A	-2	ALA	-	EXPRESSION TAG	UNP Q42975
A	-1	ASP	-	EXPRESSION TAG	UNP Q42975
A	0	VAL	-	EXPRESSION TAG	UNP Q42975
A	176	GLN	GLU	ENGINEERED MUTATION	UNP Q42975
B	-4	ALA	-	EXPRESSION TAG	UNP Q42975
B	-3	MET	-	EXPRESSION TAG	UNP Q42975
B	-2	ALA	-	EXPRESSION TAG	UNP Q42975
B	-1	ASP	-	EXPRESSION TAG	UNP Q42975
B	0	VAL	-	EXPRESSION TAG	UNP Q42975
B	176	GLN	GLU	ENGINEERED MUTATION	UNP Q42975

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

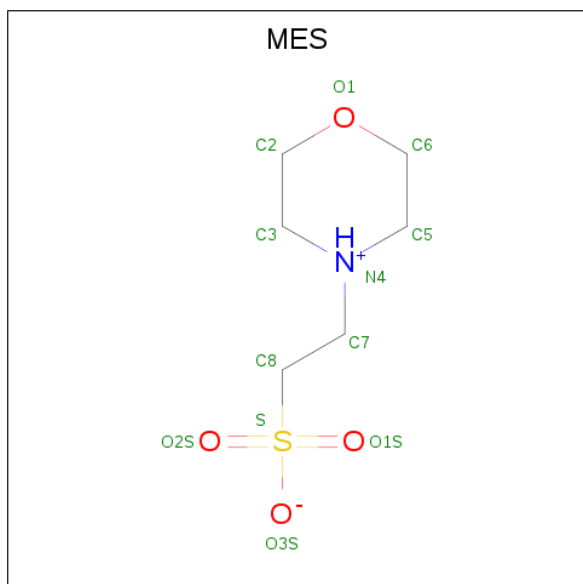
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Zn	0	0
			1	1		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula:  $C_6H_{13}NO_4S$ ).



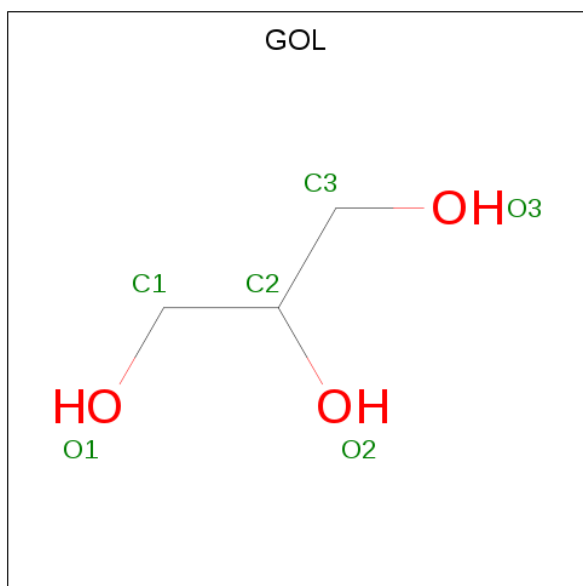
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	B	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



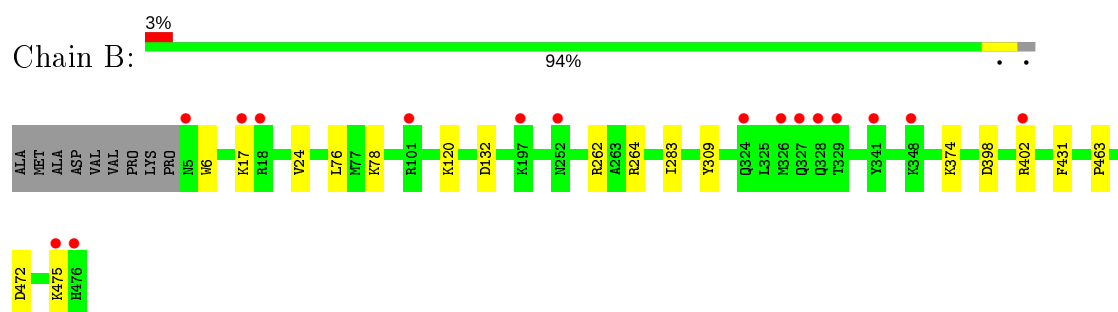
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			6	3	3		
5	A	1	Total	C	O	0	0
			6	3	3		
5	B	1	Total	C	O	0	0
			6	3	3		
5	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	461	Total	O	0	0
			461	461		
6	B	434	Total	O	0	0
			434	434		



- Molecule 1: Beta-glucosidase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.81Å 101.04Å 127.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.12 – 1.65 28.11 – 1.65	Depositor EDS
% Data completeness (in resolution range)	99.7 (28.12-1.65) 99.7 (28.11-1.65)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.02 (at 1.65Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.174 , 0.196 0.173 , 0.195	Depositor DCC
$R_{free}$ test set	6166 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	12.4	Xtriage
Anisotropy	0.028	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 50.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.36$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	8572	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	13.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 48.97 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 7.9735e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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 [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, ZN, SO4, MES

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	0.58	0/3925	0.65	1/5335 (0.0%)
1	B	0.59	0/3925	0.65	2/5335 (0.0%)
All	All	0.58	0/7850	0.65	3/10670 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	264	ARG	NE-CZ-NH2	-6.67	116.97	120.30
1	B	264	ARG	NE-CZ-NH2	-6.09	117.25	120.30
1	B	132	ASP	CB-CG-OD1	5.25	123.03	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	3809	0	3646	17	0
1	B	3809	0	3646	9	0
2	A	1	0	0	0	0
3	A	5	0	0	0	0
3	B	5	0	0	0	0
4	A	12	0	12	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	12	0	12	0	0
5	A	12	0	16	0	0
5	B	12	0	16	0	0
6	A	461	0	0	3	1
6	B	434	0	0	1	1
All	All	8572	0	7348	26	1

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 2.

The worst 5 of 26 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:235:GLN:HG3	6:A:919:HOH:O	1.49	1.13
1:A:398:ASP:OD1	6:A:677:HOH:O	2.12	0.67
1:A:398:ASP:OD1	1:A:402:ARG:NH1	2.28	0.65
1:B:24:VAL:CG2	1:B:431:PHE:HB3	2.36	0.56
1:B:398:ASP:OD1	1:B:402:ARG:NH1	2.41	0.54

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:A:937:HOH:O	6:B:745:HOH:O[2_455]	2.14	0.06

## 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	470/481 (98%)	456 (97%)	14 (3%)	0	100	100
1	B	470/481 (98%)	458 (97%)	12 (3%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	940/962 (98%)	914 (97%)	26 (3%)	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	390/397 (98%)	384 (98%)	6 (2%)	65	44
1	B	390/397 (98%)	387 (99%)	3 (1%)	81	70
All	All	780/794 (98%)	771 (99%)	9 (1%)	71	53

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	309	TYR
1	B	309	TYR
1	B	6	TRP
1	A	17	LYS
1	A	324	GLN

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	119	GLN
1	B	119	GLN
1	B	387	ASN

### 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](#)

Of 9 ligands modelled in this entry, 1 is monoatomic - leaving 8 for Mogul analysis.

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$
5	GOL	A	1005	-	5,5,5	0.31	0	5,5,5	0.28	0
5	GOL	B	1004	-	5,5,5	0.41	0	5,5,5	0.20	0
3	SO4	B	1002	-	4,4,4	0.17	0	6,6,6	0.23	0
3	SO4	A	1002	-	4,4,4	0.19	0	6,6,6	0.26	0
5	GOL	A	1004	-	5,5,5	0.38	0	5,5,5	0.40	0
5	GOL	B	1005	-	5,5,5	0.39	0	5,5,5	0.32	0
4	MES	B	1003	-	12,12,12	1.51	1 (8%)	14,16,16	5.08	8 (57%)
4	MES	A	1003	-	12,12,12	1.64	1 (8%)	14,16,16	5.26	7 (50%)

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
5	GOL	A	1005	-	-	0/4/4/4	-
5	GOL	B	1004	-	-	0/4/4/4	-
5	GOL	A	1004	-	-	0/4/4/4	-
5	GOL	B	1005	-	-	0/4/4/4	-
4	MES	B	1003	-	-	1/6/14/14	0/1/1/1
4	MES	A	1003	-	-	2/6/14/14	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1003	MES	C8-S	-5.31	1.70	1.77
4	B	1003	MES	C8-S	-4.90	1.70	1.77

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	1003	MES	O1S-S-C8	-11.35	93.25	106.92
4	A	1003	MES	O1S-S-C8	-11.02	93.64	106.92
4	A	1003	MES	O2S-S-C8	-9.40	95.60	106.92
4	B	1003	MES	O2S-S-C8	-9.01	96.07	106.92
4	A	1003	MES	O3S-S-C8	-8.83	91.49	105.77

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	1003	MES	C8-C7-N4-C5
4	A	1003	MES	C8-C7-N4-C5
4	A	1003	MES	C8-C7-N4-C3

There are no ring outliers.

No monomer is involved in short contacts.

## 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	472/481 (98%)	-0.18	12 (2%) 57 58	6, 11, 20, 33	0
1	B	472/481 (98%)	-0.17	16 (3%) 45 45	7, 11, 21, 33	0
All	All	944/962 (98%)	-0.18	28 (2%) 50 51	6, 11, 21, 33	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	476	HIS	5.3
1	A	324	GLN	4.2
1	A	476	HIS	4.0
1	B	197	LYS	3.9
1	A	5	ASN	3.8

### 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](#)

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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
5	GOL	A	1005	6/6	0.87	0.12	20,21,22,23	0
5	GOL	B	1005	6/6	0.88	0.12	23,24,25,25	0
4	MES	A	1003	12/12	0.95	0.15	15,22,24,25	0
4	MES	B	1003	12/12	0.96	0.11	15,21,21,22	0
5	GOL	B	1004	6/6	0.97	0.08	9,12,13,15	0
5	GOL	A	1004	6/6	0.98	0.07	10,11,13,14	0
3	SO4	B	1002	5/5	0.99	0.14	22,24,26,26	0
3	SO4	A	1002	5/5	0.99	0.13	20,23,24,25	0
2	ZN	A	1001	1/1	1.00	0.05	9,9,9,9	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.