



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2018 – 01:04 PM EST

PDB ID : 1TR1  
Title : CRYSTAL STRUCTURE OF E96K MUTATED BETA-GLUCOSIDASE A FROM BACILLUS POLYMYXA, AN ENZYME WITH INCREASED THERMORESISTANCE  
Authors : Sanz-Aparicio, J.; Hermoso, J.A.; Martinez-Ripoll, M.; Gonzalez-Perez, B.; Polaina, J.  
Deposited on : 1998-03-12  
Resolution : 2.20 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030736

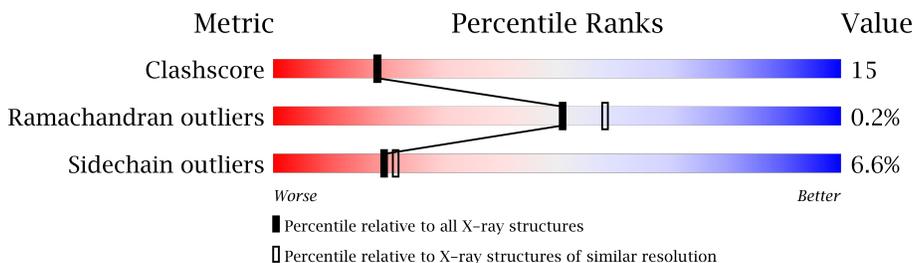
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	112137	4730 (2.20-2.20)
Ramachandran outliers	110173	4656 (2.20-2.20)
Sidechain outliers	110143	4657 (2.20-2.20)

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. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain	
1	A	447		73% 23% .
1	B	447		74% 23% .
1	C	447		72% 24% .
1	D	447		70% 26% .

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
2	GOL	A	2000	-	-	X	-
2	GOL	B	2000	-	-	X	-
2	GOL	C	2000	-	-	X	-
2	GOL	D	2000	-	-	X	-

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 16152 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.

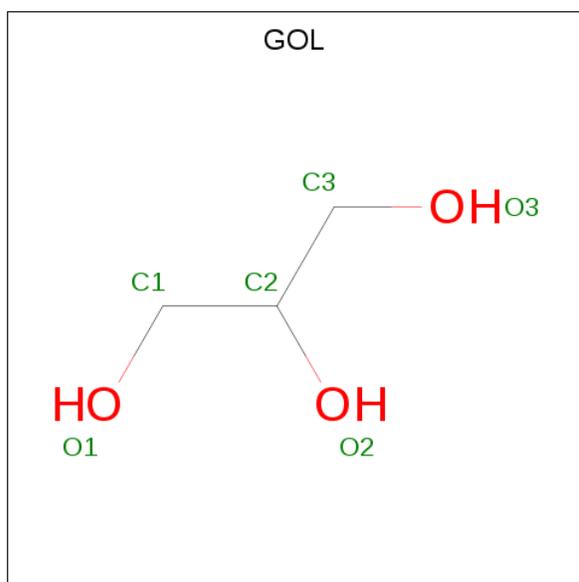
- Molecule 1 is a protein called BETA-GLUCOSIDASE A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	447	3648	2330	630	671	17	0	1	0
1	B	447	3648	2330	630	671	17	0	1	0
1	C	447	3648	2330	630	671	17	0	1	0
1	D	447	3648	2330	630	671	17	0	1	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	96	LYS	GLU	ENGINEERED MUTATION	UNP P22073
B	96	LYS	GLU	ENGINEERED MUTATION	UNP P22073
C	96	LYS	GLU	ENGINEERED MUTATION	UNP P22073
D	96	LYS	GLU	ENGINEERED MUTATION	UNP P22073

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

- Molecule 3 is water.

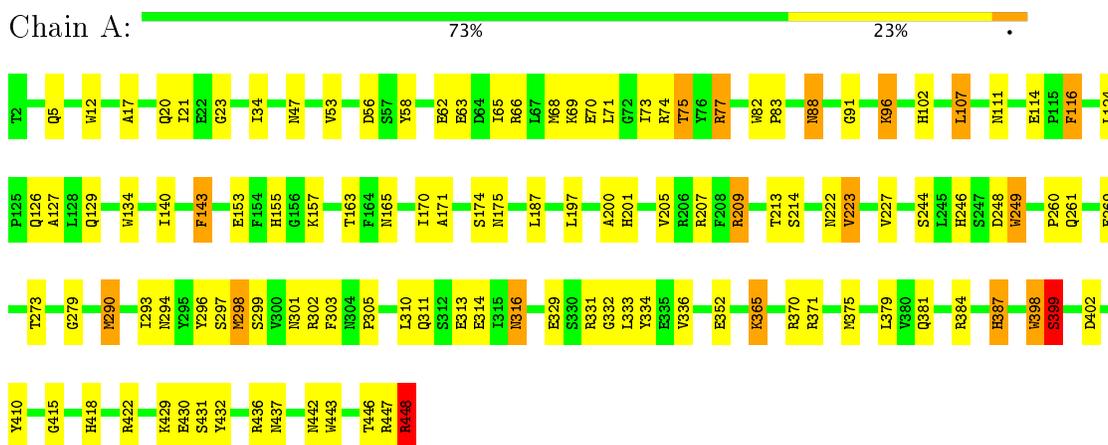
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	380	Total O 380 380	0	0
3	B	393	Total O 393 393	0	0
3	C	382	Total O 382 382	0	0
3	D	381	Total O 381 381	0	0

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

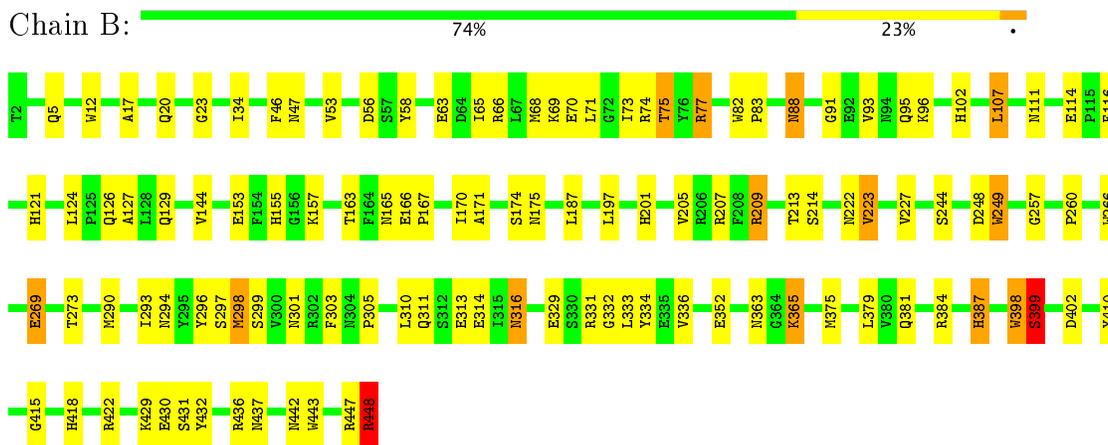
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). 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.

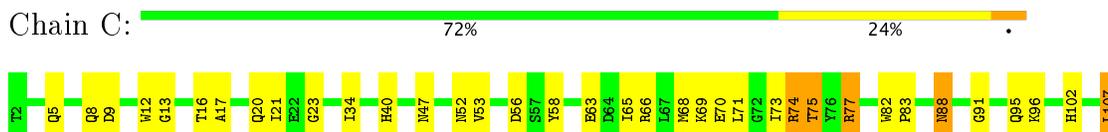
- Molecule 1: BETA-GLUCOSIDASE A

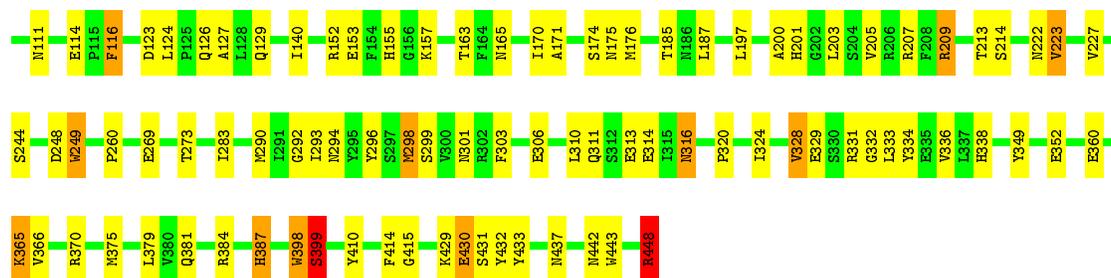


- Molecule 1: BETA-GLUCOSIDASE A



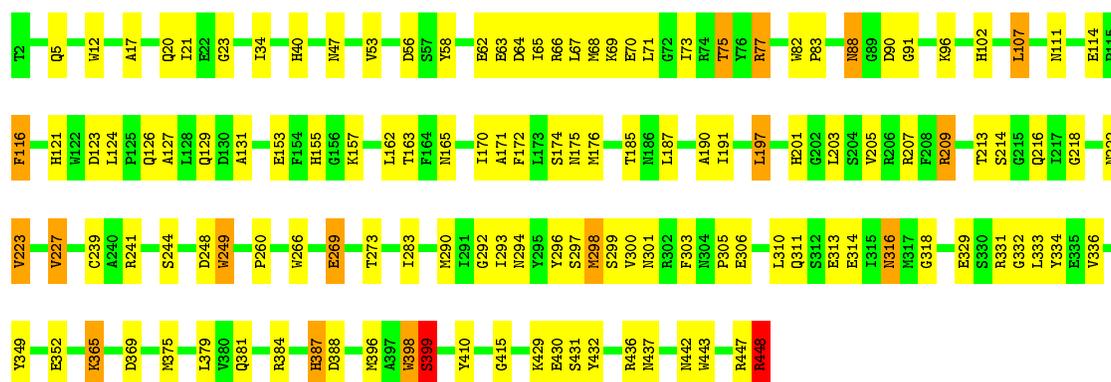
- Molecule 1: BETA-GLUCOSIDASE A





- Molecule 1: BETA-GLUCOSIDASE A

Chain D: 70% 26%



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	205.03Å 205.03Å 155.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 2.20	Depositor
% Data completeness (in resolution range)	75.6 (8.00-2.20)	Depositor
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	X-PLOR 3.843	Depositor
R, $R_{free}$	0.210 , 0.282	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	16152	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	12.0	wwPDB-VP

## 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

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	1.13	7/3762 (0.2%)	0.69	4/5114 (0.1%)
1	B	0.77	6/3763 (0.2%)	0.73	4/5117 (0.1%)
1	C	0.92	6/3763 (0.2%)	0.71	4/5117 (0.1%)
1	D	1.03	5/3763 (0.1%)	0.72	5/5117 (0.1%)
All	All	0.97	24/15051 (0.2%)	0.71	17/20465 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	3
1	B	0	2
1	C	0	3
1	D	0	2
All	All	0	10

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	448	ARG	C-OXT	49.72	2.17	1.23
1	A	398	TRP	C-O	-42.61	0.42	1.23
1	C	448	ARG	C-OXT	40.93	2.01	1.23
1	A	448	ARG	C-OXT	35.51	1.90	1.23
1	B	448	ARG	C-OXT	25.66	1.72	1.23

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	398	TRP	O-C-N	17.16	150.15	122.70
1	B	398	TRP	O-C-N	16.87	149.69	122.70
1	D	398	TRP	O-C-N	16.39	148.93	122.70
1	A	398	TRP	CA-C-O	14.52	150.58	120.10
1	C	398	TRP	CA-C-N	-12.58	89.52	117.20

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	143	PHE	Sidechain
1	A	387	HIS	Sidechain
1	A	410	TYR	Sidechain
1	B	387	HIS	Sidechain
1	B	410	TYR	Sidechain

## 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	3648	0	3433	116	0
1	B	3648	0	3434	95	0
1	C	3648	0	3434	103	0
1	D	3648	0	3434	115	0
2	A	6	0	8	7	0
2	B	6	0	8	7	0
2	C	6	0	8	6	0
2	D	6	0	8	7	0
3	A	380	0	0	28	0
3	B	393	0	0	15	0
3	C	382	0	0	23	0
3	D	381	0	0	30	0
All	All	16152	0	13767	425	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 15.

The worst 5 of 425 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:398:TRP:CA	1:A:399:SER:N	1.84	1.41
1:A:398:TRP:CA	1:A:398:TRP:C	1.91	1.38
1:B:448:ARG:C	1:B:448:ARG:OXT	1.72	1.28
1:A:448:ARG:CA	1:A:448:ARG:OXT	1.94	1.15
1:A:398:TRP:HA	1:A:399:SER:N	1.46	1.13

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	444/447 (99%)	431 (97%)	13 (3%)	0	100	100
1	B	446/447 (100%)	429 (96%)	16 (4%)	1 (0%)	51	58
1	C	446/447 (100%)	430 (96%)	15 (3%)	1 (0%)	51	58
1	D	446/447 (100%)	429 (96%)	16 (4%)	1 (0%)	51	58
All	All	1782/1788 (100%)	1719 (96%)	60 (3%)	3 (0%)	51	58

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	399	SER
1	C	399	SER
1	D	399	SER

### 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	384/383 (100%)	359 (94%)	25 (6%)	20	22
1	B	384/383 (100%)	360 (94%)	24 (6%)	21	23
1	C	384/383 (100%)	358 (93%)	26 (7%)	18	20
1	D	384/383 (100%)	358 (93%)	26 (7%)	18	20
All	All	1536/1532 (100%)	1435 (93%)	101 (7%)	19	21

5 of 101 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	365	LYS
1	C	187	LEU
1	D	298	MET
1	B	437	ASN
1	C	75	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 88 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	367	GLN
1	C	121	HIS
1	D	311	GLN
1	B	376	GLN
1	B	442	ASN

### 5.3.3 RNA [i](#)

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

4 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
2	GOL	A	2000	-	5,5,5	0.72	0	5,5,5	0.60	0
2	GOL	B	2000	-	5,5,5	0.97	0	5,5,5	0.56	0
2	GOL	C	2000	-	5,5,5	0.87	0	5,5,5	0.43	0
2	GOL	D	2000	-	5,5,5	0.78	0	5,5,5	0.44	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
2	GOL	A	2000	-	-	0/4/4/4	0/0/0/0
2	GOL	B	2000	-	-	0/4/4/4	0/0/0/0
2	GOL	C	2000	-	-	0/4/4/4	0/0/0/0
2	GOL	D	2000	-	-	0/4/4/4	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 27 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	2000	GOL	7	0
2	B	2000	GOL	7	0
2	C	2000	GOL	6	0
2	D	2000	GOL	7	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	398:TRP	C	399:SER	N	2.18
1	A	96:LYS	C	97:GLY	N	1.17

## 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.