



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 25, 2023 – 01:48 PM EDT

PDB ID : 3D6E  
Title : Crystal structure of the engineered 1,3-1,4-beta-glucanase protein from *Bacillus licheniformis*  
Authors : Fita, I.; Planas, A.; Calisto, B.M.; Addington, T.  
Deposited on : 2008-05-19  
Resolution : 2.40 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 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 : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

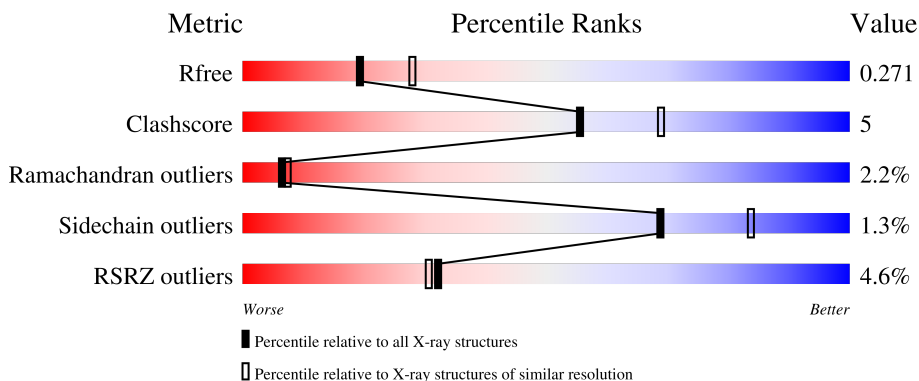
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 of 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	201	 4% (poor fit)   84% (0-1 outliers)   6% (2 outliers)   8% (3+ outliers)   0% (not modelled)
1	B	201	 4% (poor fit)   83% (0-1 outliers)   9% (2 outliers)   7% (3+ outliers)   0% (not modelled)

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3092 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-glucanase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	185	1472	946	243	279	4	0	0	0
1	B	186	1481	950	245	282	4	0	0	0

There are 62 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	ALA	deletion	UNP P27051
A	?	-	ASP	deletion	UNP P27051
A	?	-	GLY	deletion	UNP P27051
A	?	-	TYR	deletion	UNP P27051
A	?	-	SER	deletion	UNP P27051
A	?	-	ASN	deletion	UNP P27051
A	?	-	GLY	deletion	UNP P27051
A	?	-	ASN	deletion	UNP P27051
A	?	-	MET	deletion	UNP P27051
A	?	-	PHE	deletion	UNP P27051
A	?	-	ASN	deletion	UNP P27051
A	?	-	CYS	deletion	UNP P27051
A	?	-	THR	deletion	UNP P27051
A	22	ALA	ARG	engineered mutation	UNP P27051
A	23	PHE	ALA	engineered mutation	UNP P27051
A	24	ASP	ASN	engineered mutation	UNP P27051
A	25	HIS	ASN	engineered mutation	UNP P27051
A	48	GLY	CYS	engineered mutation	UNP P27051
A	50	GLY	GLU	engineered mutation	UNP P27051
A	52	GLN	ARG	engineered mutation	UNP P27051
A	77	ALA	SER	engineered mutation	UNP P27051
A	79	TYR	PHE	engineered mutation	UNP P27051
A	81	SER	TYR	engineered mutation	UNP P27051
A	167	TYR	MET	engineered mutation	UNP P27051
A	169	SER	ASN	engineered mutation	UNP P27051

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	172	ALA	ASN	engineered mutation	UNP P27051
A	190	TYR	SER	SEE REMARK 999	UNP P27051
A	191	ALA	ARG	SEE REMARK 999	UNP P27051
A	192	HIS	SER	SEE REMARK 999	UNP P27051
A	193	TYR	LEU	SEE REMARK 999	UNP P27051
A	194	ASN	HIS	SEE REMARK 999	UNP P27051
B	?	-	ALA	deletion	UNP P27051
B	?	-	ASP	deletion	UNP P27051
B	?	-	GLY	deletion	UNP P27051
B	?	-	TYR	deletion	UNP P27051
B	?	-	SER	deletion	UNP P27051
B	?	-	ASN	deletion	UNP P27051
B	?	-	GLY	deletion	UNP P27051
B	?	-	ASN	deletion	UNP P27051
B	?	-	MET	deletion	UNP P27051
B	?	-	PHE	deletion	UNP P27051
B	?	-	ASN	deletion	UNP P27051
B	?	-	CYS	deletion	UNP P27051
B	?	-	THR	deletion	UNP P27051
B	22	ALA	ARG	engineered mutation	UNP P27051
B	23	PHE	ALA	engineered mutation	UNP P27051
B	24	ASP	ASN	engineered mutation	UNP P27051
B	25	HIS	ASN	engineered mutation	UNP P27051
B	48	GLY	CYS	engineered mutation	UNP P27051
B	50	GLY	GLU	engineered mutation	UNP P27051
B	52	GLN	ARG	engineered mutation	UNP P27051
B	77	ALA	SER	engineered mutation	UNP P27051
B	79	TYR	PHE	engineered mutation	UNP P27051
B	81	SER	TYR	engineered mutation	UNP P27051
B	167	TYR	MET	engineered mutation	UNP P27051
B	169	SER	ASN	engineered mutation	UNP P27051
B	172	ALA	ASN	engineered mutation	UNP P27051
B	190	TYR	SER	SEE REMARK 999	UNP P27051
B	191	ALA	ARG	SEE REMARK 999	UNP P27051
B	192	HIS	SER	SEE REMARK 999	UNP P27051
B	193	TYR	LEU	SEE REMARK 999	UNP P27051
B	194	ASN	HIS	SEE REMARK 999	UNP P27051

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
-----	-------	----------	-------	---------	---------

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Ca 1 1	0	0
2	B	1	Total Ca 1 1	0	0


- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	69	Total O 69 69	0	0
3	B	68	Total O 68 68	0	0

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


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

- Molecule 1: Beta-glucanase

Chain A: 



- Molecule 1: Beta-glucanase

Chain B: 



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	39.55Å 54.81Å 54.91Å 61.38° 85.72° 86.10°	Depositor
Resolution (Å)	30.00 – 2.40 29.95 – 2.40	Depositor EDS
% Data completeness (in resolution range)	93.4 (30.00-2.40) 91.8 (29.95-2.40)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	0.04	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	8.08 (at 2.39Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.210 , 0.255 0.218 , 0.271	Depositor DCC
$R_{free}$ test set	731 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	22.0	Xtrriage
Anisotropy	0.423	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 18.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.466 for -h,-l,-k	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	3092	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.25% 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 [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: CA

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	0.29	0/1519	0.54	2/2069 (0.1%)
1	B	0.29	0/1528	0.53	1/2081 (0.0%)
All	All	0.29	0/3047	0.53	3/4150 (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	1	0
1	B	1	0
All	All	2	0

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	41	PRO	N-CA-CB	5.71	110.15	103.30
1	B	41	PRO	N-CA-CB	5.63	110.06	103.30
1	A	85	THR	N-CA-C	5.30	125.32	111.00

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	85	THR	CA
1	B	26	VAL	CA

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	1472	0	1350	14	0
1	B	1481	0	1356	15	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	69	0	0	0	0
3	B	68	0	0	0	0
All	All	3092	0	2706	29	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 5.

The worst 5 of 29 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:25:HIS:H	1:A:26:VAL:HG12	1.38	0.84
1:B:74:ILE:HD13	1:B:189:LEU:HD13	1.69	0.75
1:A:25:HIS:N	1:A:26:VAL:HG12	2.10	0.67
1:A:26:VAL:HG13	1:A:26:VAL:O	1.95	0.66
1:A:25:HIS:HB2	1:A:26:VAL:HB	1.81	0.63

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	181/201 (90%)	172 (95%)	5 (3%)	4 (2%)	<b>6</b> <b>7</b>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	182/201 (90%)	170 (93%)	8 (4%)	4 (2%)	6	7
All	All	363/402 (90%)	342 (94%)	13 (4%)	8 (2%)	6	7

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	24	ASP
1	A	26	VAL
1	A	41	PRO
1	B	26	VAL
1	B	41	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	150/169 (89%)	148 (99%)	2 (1%)	69	84
1	B	151/169 (89%)	149 (99%)	2 (1%)	69	84
All	All	301/338 (89%)	297 (99%)	4 (1%)	69	84

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

Mol	Chain	Res	Type
1	A	26	VAL
1	A	85	THR
1	B	26	VAL
1	B	47	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	51	ASN
1	B	71	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	122	ASN
1	A	19	GLN
1	A	11	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 monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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.

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	185/201 (92%)	0.01	9 (4%) 29 28	16, 22, 34, 40	0
1	B	186/201 (92%)	0.06	8 (4%) 35 33	16, 22, 34, 40	0
All	All	371/402 (92%)	0.04	17 (4%) 32 31	16, 22, 35, 40	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	40	SER	5.5
1	B	42	SER	4.9
1	A	40	SER	4.2
1	B	46	PHE	4.2
1	A	46	PHE	4.1

### 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 monosaccharides 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
2	CA	B	202	1/1	0.98	0.05	25,25,25,25	0
2	CA	A	202	1/1	0.99	0.04	30,30,30,30	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.