



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 24, 2024 – 12:13 am GMT

PDB ID : 8R73  
Title : Polysaccharide lyase BtPL33HA (BT4410) Apo form 1  
Authors : Cartmell, A.  
Deposited on : 2023-11-23  
Resolution : 2.11 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 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 : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.40

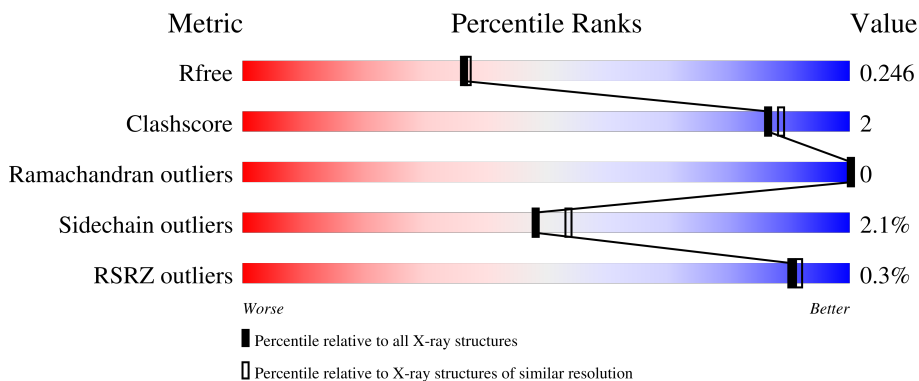
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

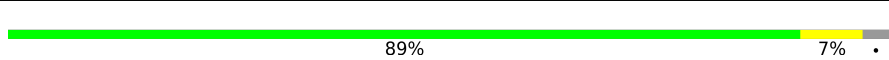
The reported resolution of this entry is 2.11 Å.

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))
$R_{free}$	164625	7689 (2.14-2.10)
Clashscore	180529	8431 (2.14-2.10)
Ramachandran outliers	177936	8366 (2.14-2.10)
Sidechain outliers	177891	8367 (2.14-2.10)
RSRZ outliers	164620	7689 (2.14-2.10)

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	636	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10134 atoms, of which 4840 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 Heparinase II/III-like protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	617	9821	3173	4840	852	932	24	133	0	0

There are 19 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	MET	-	initiating methionine	UNP A0A0P0GKX0
A	8	GLY	-	expression tag	UNP A0A0P0GKX0
A	9	SER	-	expression tag	UNP A0A0P0GKX0
A	10	SER	-	expression tag	UNP A0A0P0GKX0
A	11	HIS	-	expression tag	UNP A0A0P0GKX0
A	12	HIS	-	expression tag	UNP A0A0P0GKX0
A	13	HIS	-	expression tag	UNP A0A0P0GKX0
A	14	HIS	-	expression tag	UNP A0A0P0GKX0
A	15	HIS	-	expression tag	UNP A0A0P0GKX0
A	16	HIS	-	expression tag	UNP A0A0P0GKX0
A	17	SER	-	expression tag	UNP A0A0P0GKX0
A	18	SER	-	expression tag	UNP A0A0P0GKX0
A	19	GLY	-	expression tag	UNP A0A0P0GKX0
A	20	PRO	-	expression tag	UNP A0A0P0GKX0
A	21	GLN	-	expression tag	UNP A0A0P0GKX0
A	22	GLN	-	expression tag	UNP A0A0P0GKX0
A	23	GLY	-	expression tag	UNP A0A0P0GKX0
A	24	LEU	-	expression tag	UNP A0A0P0GKX0
A	25	ARG	-	expression tag	UNP A0A0P0GKX0

- 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		

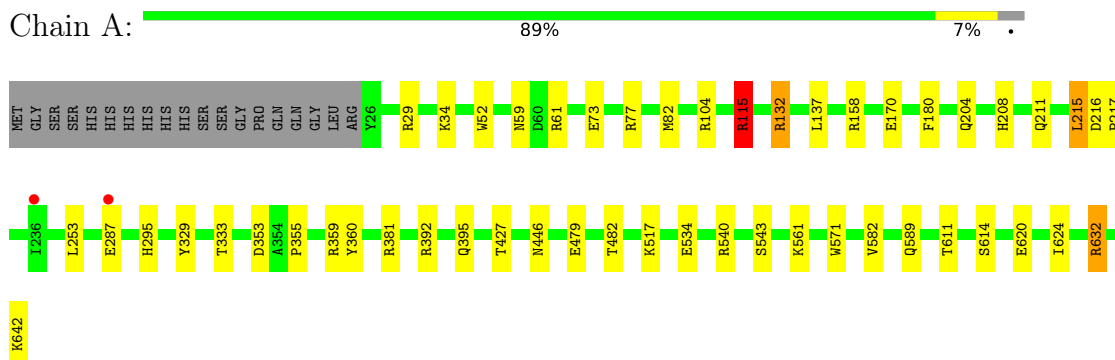
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	312	Total 312	O 312	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: Heparinase II/III-like protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	103.63Å 91.39Å 84.77Å 90.00° 109.28° 90.00°	Depositor
Resolution (Å)	40.01 – 2.11 40.01 – 2.11	Depositor EDS
% Data completeness (in resolution range)	99.7 (40.01-2.11) 92.3 (40.01-2.11)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.36 (at 2.12Å)	Xtrriage
Refinement program	REFMAC 5.8.0419	Depositor
R, $R_{free}$	0.189 , 0.250 0.186 , 0.246	Depositor DCC
$R_{free}$ test set	2098 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	34.4	Xtrriage
Anisotropy	0.354	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 39.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	10134	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	43.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.96% 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

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

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	0.48	0/5109	0.94	9/6927 (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	6

There are no bond length outliers.

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	82	MET	CG-SD-CE	6.83	111.12	100.20
1	A	392	ARG	NE-CZ-NH1	6.17	123.39	120.30
1	A	115	ARG	N-CA-CB	-6.10	99.62	110.60
1	A	61	ARG	NE-CZ-NH1	5.96	123.28	120.30
1	A	61	ARG	NE-CZ-NH2	-5.79	117.41	120.30
1	A	540	ARG	NE-CZ-NH2	-5.73	117.44	120.30
1	A	170	GLU	CB-CA-C	-5.26	99.89	110.40
1	A	215	LEU	CB-CG-CD1	5.25	119.92	111.00
1	A	132	ARG	N-CA-CB	-5.00	101.60	110.60

There are no chirality outliers.

All (6) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	104	ARG	Sidechain

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
1	A	115	ARG	Sidechain
1	A	158	ARG	Sidechain
1	A	381	ARG	Sidechain
1	A	446	ASN	Peptide
1	A	632	ARG	Sidechain

## 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	4981	4840	4831	20	0
2	A	1	0	0	0	0
3	A	312	0	0	7	0
All	All	5294	4840	4831	20	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 2.

All (20) 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:359:ARG:NH2	3:A:801:HOH:O	2.26	0.68
1:A:534:GLU:HG3	1:A:561:LYS:HD2	1.81	0.61
1:A:479:GLU:O	1:A:482:THR:OG1	2.19	0.59
1:A:582:VAL:HG11	1:A:624:ILE:HD13	1.87	0.57
1:A:517:LYS:HE3	3:A:1084:HOH:O	2.07	0.54
1:A:329:TYR:CZ	1:A:333:THR:HG21	2.44	0.52
1:A:29:ARG:O	1:A:34:LYS:HE3	2.13	0.49
1:A:59:ASN:OD1	1:A:59:ASN:C	2.53	0.47
1:A:611:THR:HA	1:A:614:SER:OG	2.16	0.46
1:A:211:GLN:HA	1:A:215:LEU:HB2	2.00	0.44
1:A:642:LYS:NZ	3:A:807:HOH:O	2.49	0.44
1:A:115:ARG:HG2	3:A:1031:HOH:O	2.19	0.43
1:A:295:HIS:HE1	3:A:831:HOH:O	2.01	0.42
1:A:359:ARG:HD3	1:A:395:GLN:HB3	2.02	0.42
1:A:52:TRP:CZ2	1:A:253:LEU:HB3	2.55	0.41

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:208:HIS:HB2	3:A:1077:HOH:O	2.20	0.41
1:A:353:ASP:OD1	1:A:355:PRO:HD2	2.21	0.41
1:A:132:ARG:HD2	3:A:962:HOH:O	2.21	0.41
1:A:216:ASP:N	1:A:217:PRO:CD	2.83	0.40
1:A:73:GLU:HB3	1:A:77:ARG:HH12	1.86	0.40

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	615/636 (97%)	595 (97%)	20 (3%)	0	<b>100</b> <b>100</b>

There are no Ramachandran outliers to report.

### 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	530/546 (97%)	519 (98%)	11 (2%)	48 54

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	137	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	180	PHE
1	A	204	GLN
1	A	287	GLU
1	A	360	TYR
1	A	427	THR
1	A	543	SER
1	A	571	TRP
1	A	589	GLN
1	A	620	GLU
1	A	632	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	204	GLN
1	A	295	HIS

### 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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 1 ligands modelled in this entry, 1 is 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	617/636 (97%)	-0.39	2 (0%) 90 91	26, 40, 62, 97	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	236	ILE	2.6
1	A	287	GLU	2.0

### 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(Å <sup>2</sup> )	Q<0.9
2	ZN	A	701	1/1	1.00	0.01	41,41,41,41	0

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

There are no such residues in this entry.