



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 11:14 PM UTC

PDB ID : 3HXK / pdb_00003hxx
Title : Crystal Structure of a sugar hydrolase (YeeB) from *Lactococcus lactis*, Northeast Structural Genomics Consortium Target KR108
Authors : Forouhar, F.; Su, M.; Seetharaman, J.; Mao, M.; Xiao, R.; Ciccocanti, C.; Foote, E.L.; Maglaqui, M.; Zhao, L.; Everett, J.K.; Nair, R.; Acton, T.B.; Rost, B.; Montelione, G.T.; Tong, L.; Hunt, J.F.; Northeast Structural Genomics Consortium (NESG)
Deposited on : 2009-06-20
Resolution : 3.20 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

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

<https://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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)

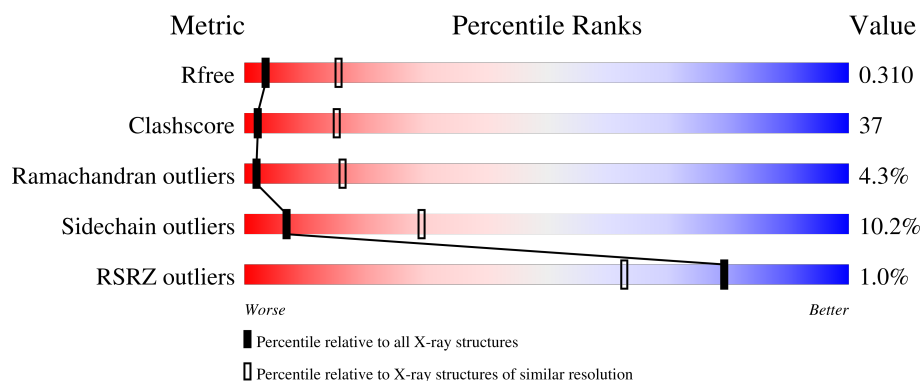
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.20 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))
R_{free}	180053	1466 (3.20-3.20)
Clashscore	190562	1573 (3.20-3.20)
Ramachandran outliers	187476	1548 (3.20-3.20)
Sidechain outliers	187428	1547 (3.20-3.20)
RSRZ outliers	180081	1466 (3.20-3.20)

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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 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	276	% 34% 48% 8% • 9%
1	B	276	% 33% 48% 9% • 9%
1	C	276	 32% 50% 8% • 9%
1	D	276	% 35% 47% 8% • 9%

Validation Pipeline (wwPDB-VP) : 2.49

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 8141 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 Sugar hydrolase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	250	Total 2020	C 1300	N 337	O 377	S 5	Se 1	0	0	0
1	B	250	Total 2020	C 1300	N 337	O 377	S 5	Se 1	0	0	0
1	C	250	Total 2020	C 1300	N 337	O 377	S 5	Se 1	0	0	0
1	D	250	Total 2020	C 1300	N 337	O 377	S 5	Se 1	0	0	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	269	LEU	-	expression tag	UNP Q9CID4
A	270	GLU	-	expression tag	UNP Q9CID4
A	271	HIS	-	expression tag	UNP Q9CID4
A	272	HIS	-	expression tag	UNP Q9CID4
A	273	HIS	-	expression tag	UNP Q9CID4
A	274	HIS	-	expression tag	UNP Q9CID4
A	275	HIS	-	expression tag	UNP Q9CID4
A	276	HIS	-	expression tag	UNP Q9CID4
B	269	LEU	-	expression tag	UNP Q9CID4
B	270	GLU	-	expression tag	UNP Q9CID4
B	271	HIS	-	expression tag	UNP Q9CID4
B	272	HIS	-	expression tag	UNP Q9CID4
B	273	HIS	-	expression tag	UNP Q9CID4
B	274	HIS	-	expression tag	UNP Q9CID4
B	275	HIS	-	expression tag	UNP Q9CID4
B	276	HIS	-	expression tag	UNP Q9CID4
C	269	LEU	-	expression tag	UNP Q9CID4
C	270	GLU	-	expression tag	UNP Q9CID4
C	271	HIS	-	expression tag	UNP Q9CID4
C	272	HIS	-	expression tag	UNP Q9CID4
C	273	HIS	-	expression tag	UNP Q9CID4

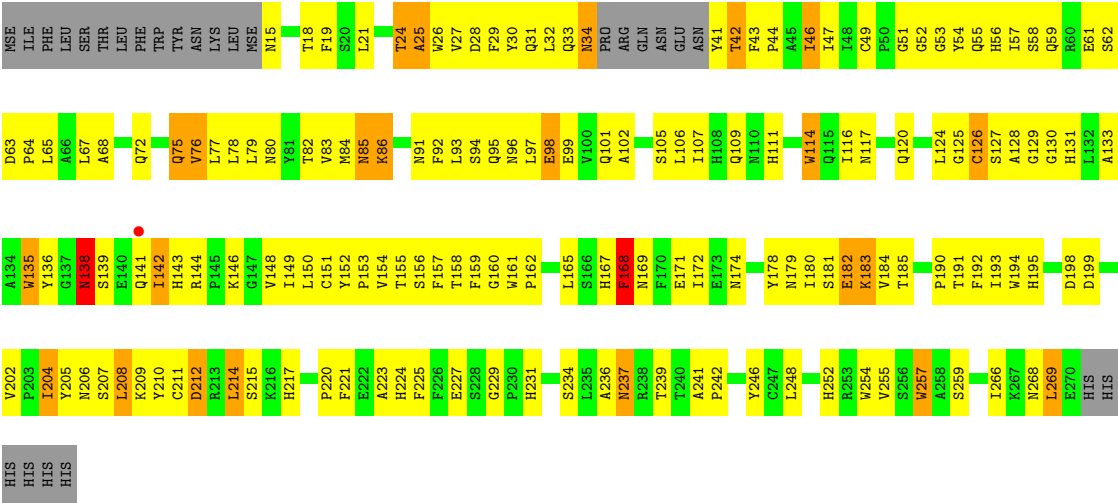
Continued on next page...

Continued from previous page...

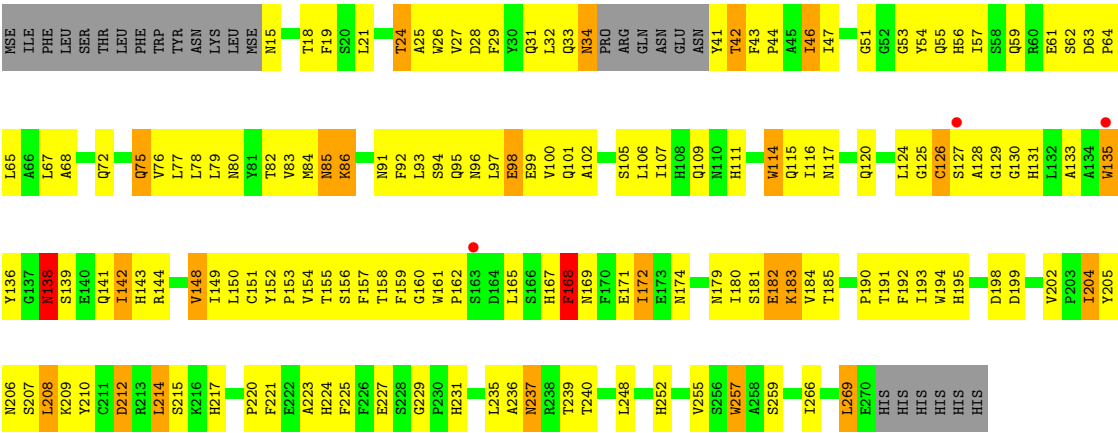
Chain	Residue	Modelled	Actual	Comment	Reference
C	274	HIS	-	expression tag	UNP Q9CID4
C	275	HIS	-	expression tag	UNP Q9CID4
C	276	HIS	-	expression tag	UNP Q9CID4
D	269	LEU	-	expression tag	UNP Q9CID4
D	270	GLU	-	expression tag	UNP Q9CID4
D	271	HIS	-	expression tag	UNP Q9CID4
D	272	HIS	-	expression tag	UNP Q9CID4
D	273	HIS	-	expression tag	UNP Q9CID4
D	274	HIS	-	expression tag	UNP Q9CID4
D	275	HIS	-	expression tag	UNP Q9CID4
D	276	HIS	-	expression tag	UNP Q9CID4

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	9	Total O 9 9	0	0
2	B	13	Total O 13 13	0	0
2	C	14	Total O 14 14	0	0
2	D	25	Total O 25 25	0	0



• Molecule 1: Sugar hydrolase



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	68.39Å 150.43Å 218.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.96 – 3.20 19.96 – 3.20	Depositor EDS
% Data completeness (in resolution range)	91.7 (19.96-3.20) 95.6 (19.96-3.20)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.60 (at 3.05Å)	Xtriage
Refinement program	CNS 1.2 & XtalView, REFMAC	Depositor
R, R_{free}	0.256 , 0.296 0.273 , 0.310	Depositor DCC
R_{free} test set	3229 reflections (8.95%)	wwPDB-VP
Wilson B-factor (Å ²)	64.3	Xtriage
Anisotropy	0.442	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.25 , 42.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.39$, $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.86	EDS
Total number of atoms	8141	wwPDB-VP
Average B, all atoms (Å ²)	71.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.39% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.57	0/2085	0.88	6/2845 (0.2%)
1	B	0.55	0/2085	0.89	6/2845 (0.2%)
1	C	0.58	0/2085	0.88	6/2845 (0.2%)
1	D	0.58	0/2085	0.87	5/2845 (0.2%)
All	All	0.57	0/8340	0.88	23/11380 (0.2%)

There are no bond length outliers.

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	214	LEU	N-CA-C	-6.81	106.86	114.62
1	D	214	LEU	N-CA-C	-6.55	107.16	114.62
1	C	214	LEU	N-CA-C	-6.18	107.57	114.62
1	C	76	VAL	N-CA-C	5.95	116.50	108.17
1	A	257	TRP	N-CA-C	-5.87	104.79	111.07

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	2020	0	1897	149	0
1	B	2020	0	1897	144	0
1	C	2020	0	1897	157	0
1	D	2020	0	1897	152	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	9	0	0	3	0
2	B	13	0	0	3	0
2	C	14	0	0	6	0
2	D	25	0	0	9	0
All	All	8141	0	7588	576	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 37.

The worst 5 of 576 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:193:ILE:HB	2:A:283:HOH:O	1.50	1.12
1:D:126:CYS:HB3	1:D:151:CYS:HB2	1.35	1.02
1:B:126:CYS:HB3	1:B:151:CYS:HB2	1.39	1.01
1:A:126:CYS:HB3	1:A:151:CYS:HB2	1.45	0.99
1:C:126:CYS:HB3	1:C:151:CYS:HB2	1.43	0.96

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	246/276 (89%)	180 (73%)	56 (23%)	10 (4%)	2	17
1	B	246/276 (89%)	179 (73%)	56 (23%)	11 (4%)	2	15
1	C	246/276 (89%)	182 (74%)	52 (21%)	12 (5%)	1	13
1	D	246/276 (89%)	179 (73%)	58 (24%)	9 (4%)	2	18
All	All	984/1104 (89%)	720 (73%)	222 (23%)	42 (4%)	2	16

5 of 42 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	84	MSE
1	A	171	GLU
1	A	174	ASN
1	A	198	ASP
1	B	171	GLU

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	221/244 (91%)	200 (90%)	21 (10%)	8	32
1	B	221/244 (91%)	198 (90%)	23 (10%)	7	28
1	C	221/244 (91%)	199 (90%)	22 (10%)	7	30
1	D	221/244 (91%)	197 (89%)	24 (11%)	6	26
All	All	884/976 (91%)	794 (90%)	90 (10%)	7	29

5 of 90 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	168	PHE
1	D	75	GLN
1	C	204	ILE
1	C	269	LEU
1	D	100	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 49 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	108	HIS
1	C	268	ASN
1	C	110	ASN
1	C	217	HIS
1	D	33	GLN

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](#)

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, 95th 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 ²)	Q<0.9
1	A	249/276 (90%)	0.15	3 (1%) 76 58	14, 68, 106, 123	1 (0%)
1	B	249/276 (90%)	0.06	3 (1%) 76 58	18, 69, 105, 121	1 (0%)
1	C	249/276 (90%)	0.24	1 (0%) 88 79	21, 72, 105, 122	1 (0%)
1	D	249/276 (90%)	0.24	3 (1%) 76 58	19, 69, 106, 120	1 (0%)
All	All	996/1104 (90%)	0.17	10 (1%) 79 63	14, 70, 106, 123	4 (0%)

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	141	GLN	4.1
1	B	127	SER	3.8
1	D	127	SER	3.7
1	A	127	SER	2.7
1	B	180	ILE	2.6

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 oligosaccharides 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.