



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 26, 2023 – 09:53 PM EDT

PDB ID : 3GEK
Title : Crystal structure of putative thioesterase yhdA from *Lactococcus lactis*.
Northeast Structural Genomics Consortium Target KR113
Authors : Kuzin, A.P.; Su, M.; Seetharaman, R.; Lee, D.; Foote, E.L.; Ciccocanti, C.;
Janjua, H.; Xiao, R.; Nair, R.; Rost, B.; Acton, T.B.; Everett, J.K.; Monte-
lione, G.T.; Hunt, J.F.; Tong, L.; Northeast Structural Genomics Consortium
(NESG)
Deposited on : 2009-02-25
Resolution : 2.24 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

<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
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.35
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)

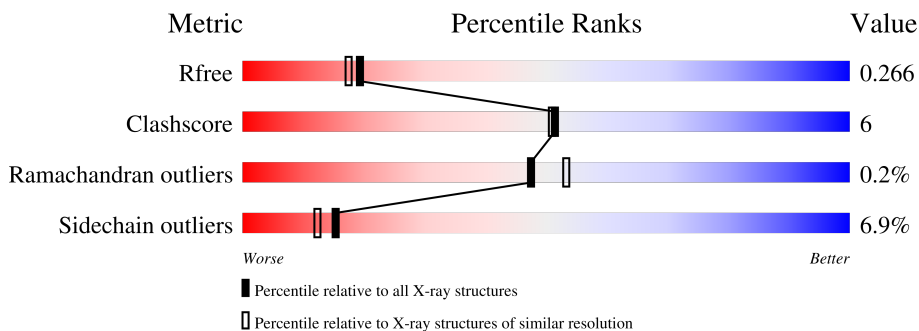
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.24 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	2391 (2.26-2.22)
Clashscore	141614	2539 (2.26-2.22)
Ramachandran outliers	138981	2489 (2.26-2.22)
Sidechain outliers	138945	2490 (2.26-2.22)

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	146	68% 18% • 10%
1	B	146	73% 10% • 14%
1	C	146	75% 11% • 13%
1	D	146	75% 8% • 16%

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 3986 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 Putative thioesterase yhdA.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	132	1031	654	177	195	1	4	0	2	0
1	B	125	965	613	166	181	1	4	0	0	0
1	C	127	983	624	169	185	1	4	0	0	0
1	D	123	958	609	165	179	1	4	0	1	0

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	23	PHE	SER	engineered mutation	UNP Q9CHK5
A	44	GLN	HIS	engineered mutation	UNP Q9CHK5
A	70	GLY	SER	engineered mutation	UNP Q9CHK5
A	75	PHE	LEU	engineered mutation	UNP Q9CHK5
A	80	ASN	SER	engineered mutation	UNP Q9CHK5
A	86	PRO	SER	engineered mutation	UNP Q9CHK5
A	96	ARG	CYS	engineered mutation	UNP Q9CHK5
A	139	LEU	-	expression tag	UNP Q9CHK5
A	140	GLU	-	expression tag	UNP Q9CHK5
A	141	HIS	-	expression tag	UNP Q9CHK5
A	142	HIS	-	expression tag	UNP Q9CHK5
A	143	HIS	-	expression tag	UNP Q9CHK5
A	144	HIS	-	expression tag	UNP Q9CHK5
A	145	HIS	-	expression tag	UNP Q9CHK5
A	146	HIS	-	expression tag	UNP Q9CHK5
B	23	PHE	SER	engineered mutation	UNP Q9CHK5
B	44	GLN	HIS	engineered mutation	UNP Q9CHK5
B	70	GLY	SER	engineered mutation	UNP Q9CHK5
B	75	PHE	LEU	engineered mutation	UNP Q9CHK5
B	80	ASN	SER	engineered mutation	UNP Q9CHK5
B	86	PRO	SER	engineered mutation	UNP Q9CHK5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	96	ARG	CYS	engineered mutation	UNP Q9CHK5
B	139	LEU	-	expression tag	UNP Q9CHK5
B	140	GLU	-	expression tag	UNP Q9CHK5
B	141	HIS	-	expression tag	UNP Q9CHK5
B	142	HIS	-	expression tag	UNP Q9CHK5
B	143	HIS	-	expression tag	UNP Q9CHK5
B	144	HIS	-	expression tag	UNP Q9CHK5
B	145	HIS	-	expression tag	UNP Q9CHK5
B	146	HIS	-	expression tag	UNP Q9CHK5
C	23	PHE	SER	engineered mutation	UNP Q9CHK5
C	44	GLN	HIS	engineered mutation	UNP Q9CHK5
C	70	GLY	SER	engineered mutation	UNP Q9CHK5
C	75	PHE	LEU	engineered mutation	UNP Q9CHK5
C	80	ASN	SER	engineered mutation	UNP Q9CHK5
C	86	PRO	SER	engineered mutation	UNP Q9CHK5
C	96	ARG	CYS	engineered mutation	UNP Q9CHK5
C	139	LEU	-	expression tag	UNP Q9CHK5
C	140	GLU	-	expression tag	UNP Q9CHK5
C	141	HIS	-	expression tag	UNP Q9CHK5
C	142	HIS	-	expression tag	UNP Q9CHK5
C	143	HIS	-	expression tag	UNP Q9CHK5
C	144	HIS	-	expression tag	UNP Q9CHK5
C	145	HIS	-	expression tag	UNP Q9CHK5
C	146	HIS	-	expression tag	UNP Q9CHK5
D	23	PHE	SER	engineered mutation	UNP Q9CHK5
D	44	GLN	HIS	engineered mutation	UNP Q9CHK5
D	70	GLY	SER	engineered mutation	UNP Q9CHK5
D	75	PHE	LEU	engineered mutation	UNP Q9CHK5
D	80	ASN	SER	engineered mutation	UNP Q9CHK5
D	86	PRO	SER	engineered mutation	UNP Q9CHK5
D	96	ARG	CYS	engineered mutation	UNP Q9CHK5
D	139	LEU	-	expression tag	UNP Q9CHK5
D	140	GLU	-	expression tag	UNP Q9CHK5
D	141	HIS	-	expression tag	UNP Q9CHK5
D	142	HIS	-	expression tag	UNP Q9CHK5
D	143	HIS	-	expression tag	UNP Q9CHK5
D	144	HIS	-	expression tag	UNP Q9CHK5
D	145	HIS	-	expression tag	UNP Q9CHK5
D	146	HIS	-	expression tag	UNP Q9CHK5

- Molecule 2 is water.

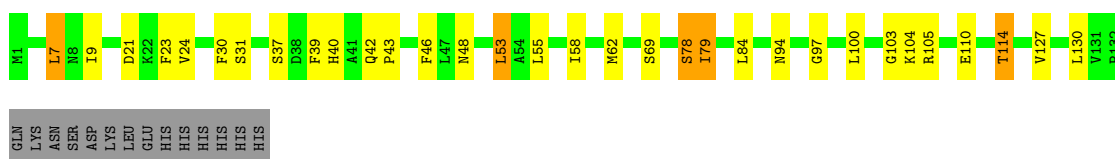
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	17	Total O 17 17	0	0
2	B	18	Total O 18 18	0	0
2	C	9	Total O 9 9	0	0
2	D	5	Total O 5 5	0	0

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. 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. 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.

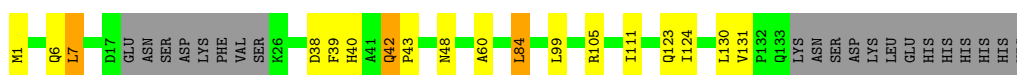
- Molecule 1: Putative thioesterase yhdA

Chain A:  68% 18% 10%




- Molecule 1: Putative thioesterase yhdA

Chain B:  73% 10% 14%




- Molecule 1: Putative thioesterase yhdA

Chain C:  75% 11% 13%



- Molecule 1: Putative thioesterase yhdA

Chain D:  75% 8% 16%



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	59.95Å 64.06Å 144.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.18 – 2.24 46.16 – 2.24	Depositor EDS
% Data completeness (in resolution range)	100.0 (46.18-2.24) 99.6 (46.16-2.24)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	7.95 (at 2.24Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.224 , 0.255 0.251 , 0.266	Depositor DCC
R_{free} test set	1385 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å ²)	46.3	Xtrriage
Anisotropy	0.043	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 43.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	3986	wwPDB-VP
Average B, all atoms (Å ²)	51.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.05% 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 [i](#)

5.1 Standard geometry [i](#)

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	0.49	1/1052 (0.1%)	0.54	0/1413
1	B	0.40	0/978	0.56	0/1313
1	C	0.57	1/996 (0.1%)	0.54	0/1336
1	D	0.50	1/970 (0.1%)	0.53	0/1300
All	All	0.49	3/3996 (0.1%)	0.54	0/5362

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	69	SER	CB-OG	13.12	1.59	1.42
1	D	69	SER	CB-OG	-6.70	1.33	1.42
1	A	23	PHE	C-N	6.01	1.47	1.34

There are no bond angle outliers.

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	1031	0	1031	21	0
1	B	965	0	963	12	0
1	C	983	0	982	9	0
1	D	958	0	960	10	0
2	A	17	0	0	0	0
2	B	18	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	9	0	0	0	0
2	D	5	0	0	0	0
All	All	3986	0	3936	44	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 6.

The worst 5 of 44 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:48:ASN:HD21	1:D:58:ILE:HD11	1.51	0.76
1:C:54:ALA:O	1:C:58:ILE:HG12	1.86	0.75
1:A:105[B]:ARG:HE	1:A:105[B]:ARG:H	1.34	0.73
1:B:1:MSE:HE2	1:B:6:GLN:CG	2.24	0.68
1:A:9:ILE:HD12	1:A:55:LEU:HD13	1.75	0.68

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	132/146 (90%)	128 (97%)	2 (2%)	2 (2%)	10	5
1	B	121/146 (83%)	119 (98%)	2 (2%)	0	100	100
1	C	123/146 (84%)	119 (97%)	4 (3%)	0	100	100
1	D	120/146 (82%)	116 (97%)	4 (3%)	0	100	100
All	All	496/584 (85%)	482 (97%)	12 (2%)	2 (0%)	47	35

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	69[A]	SER
1	A	69[B]	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	113/121 (93%)	102 (90%)	11 (10%)	8	4
1	B	104/121 (86%)	96 (92%)	8 (8%)	13	9
1	C	106/121 (88%)	101 (95%)	5 (5%)	26	27
1	D	103/121 (85%)	98 (95%)	5 (5%)	25	25
All	All	426/484 (88%)	397 (93%)	29 (7%)	15	13

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	84	LEU
1	D	53	LEU
1	B	130	LEU
1	D	7	LEU
1	B	105	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	106	ASN
1	D	6	GLN
1	D	106	ASN
1	D	77	GLN
1	D	94	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](#)

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

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.