



wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 12:22 am BST

PDB ID : 2PJQ
Title : Crystal structure of Q88U62_LACPL from Lactobacillus plantarum. Northeast Structural Genomics target LpR71
Authors : Benach, J.; Su, M.; Seetharaman, J.; Forouhar, F.; Chen, C.X.; Cunningham, K.; Ma, L.-C.; Owens, L.; Baran, M.; Acton, T.B.; Montelione, G.T.; Tong, L.; Hunt, J.F.; Northeast Structural Genomics Consortium (NESG)
Deposited on : 2007-04-16
Resolution : 2.80 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 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 : 2.11
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) : 2.11

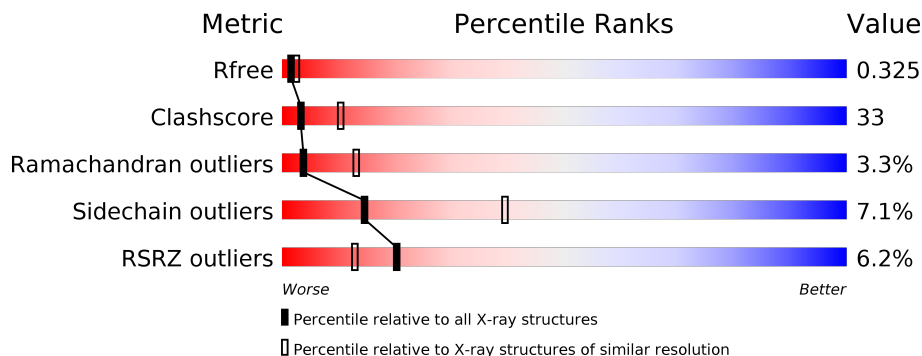
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 ≥ 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	231	
1	B	231	
1	C	231	
1	D	231	

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 6575 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 Uncharacterized protein lp_2664.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
1	A	206	Total 1625	C 1019	N 298	O 303	Se 5	0	0	0
1	B	212	Total 1667	C 1047	N 306	O 308	Se 6	0	0	0
1	C	212	Total 1667	C 1045	N 306	O 310	Se 6	0	0	0
1	D	203	Total 1604	C 1008	N 295	O 296	Se 5	0	0	0

There are 76 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	MSE	-	CLONING ARTIFACT	UNP Q88U62
A	-3	ALA	-	CLONING ARTIFACT	UNP Q88U62
A	-2	GLY	-	CLONING ARTIFACT	UNP Q88U62
A	-1	ASP	-	CLONING ARTIFACT	UNP Q88U62
A	0	PRO	-	CLONING ARTIFACT	UNP Q88U62
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
A	66	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
A	100	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
A	158	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
A	185	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
A	201	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
A	219	LEU	-	CLONING ARTIFACT	UNP Q88U62
A	220	GLU	-	CLONING ARTIFACT	UNP Q88U62
A	221	HIS	-	CLONING ARTIFACT	UNP Q88U62
A	222	HIS	-	CLONING ARTIFACT	UNP Q88U62
A	223	HIS	-	CLONING ARTIFACT	UNP Q88U62
A	224	HIS	-	CLONING ARTIFACT	UNP Q88U62
A	225	HIS	-	CLONING ARTIFACT	UNP Q88U62
A	226	HIS	-	CLONING ARTIFACT	UNP Q88U62
B	-4	MSE	-	CLONING ARTIFACT	UNP Q88U62
B	-3	ALA	-	CLONING ARTIFACT	UNP Q88U62

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	GLY	-	CLONING ARTIFACT	UNP Q88U62
B	-1	ASP	-	CLONING ARTIFACT	UNP Q88U62
B	0	PRO	-	CLONING ARTIFACT	UNP Q88U62
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
B	66	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
B	100	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
B	158	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
B	185	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
B	201	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
B	219	LEU	-	CLONING ARTIFACT	UNP Q88U62
B	220	GLU	-	CLONING ARTIFACT	UNP Q88U62
B	221	HIS	-	CLONING ARTIFACT	UNP Q88U62
B	222	HIS	-	CLONING ARTIFACT	UNP Q88U62
B	223	HIS	-	CLONING ARTIFACT	UNP Q88U62
B	224	HIS	-	CLONING ARTIFACT	UNP Q88U62
B	225	HIS	-	CLONING ARTIFACT	UNP Q88U62
B	226	HIS	-	CLONING ARTIFACT	UNP Q88U62
C	-4	MSE	-	CLONING ARTIFACT	UNP Q88U62
C	-3	ALA	-	CLONING ARTIFACT	UNP Q88U62
C	-2	GLY	-	CLONING ARTIFACT	UNP Q88U62
C	-1	ASP	-	CLONING ARTIFACT	UNP Q88U62
C	0	PRO	-	CLONING ARTIFACT	UNP Q88U62
C	1	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
C	66	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
C	100	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
C	158	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
C	185	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
C	201	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
C	219	LEU	-	CLONING ARTIFACT	UNP Q88U62
C	220	GLU	-	CLONING ARTIFACT	UNP Q88U62
C	221	HIS	-	CLONING ARTIFACT	UNP Q88U62
C	222	HIS	-	CLONING ARTIFACT	UNP Q88U62
C	223	HIS	-	CLONING ARTIFACT	UNP Q88U62
C	224	HIS	-	CLONING ARTIFACT	UNP Q88U62
C	225	HIS	-	CLONING ARTIFACT	UNP Q88U62
C	226	HIS	-	CLONING ARTIFACT	UNP Q88U62
D	-4	MSE	-	CLONING ARTIFACT	UNP Q88U62
D	-3	ALA	-	CLONING ARTIFACT	UNP Q88U62
D	-2	GLY	-	CLONING ARTIFACT	UNP Q88U62
D	-1	ASP	-	CLONING ARTIFACT	UNP Q88U62
D	0	PRO	-	CLONING ARTIFACT	UNP Q88U62
D	1	MSE	MET	MODIFIED RESIDUE	UNP Q88U62

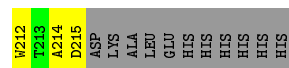
Continued on next page...

Continued from previous page...

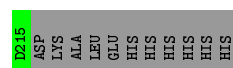
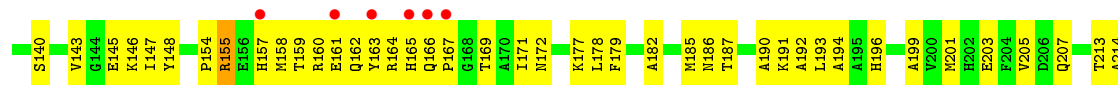
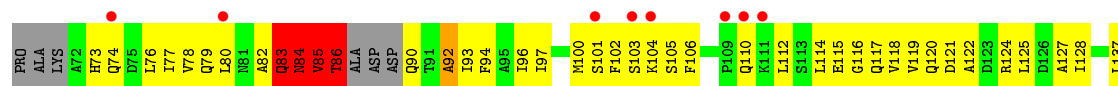
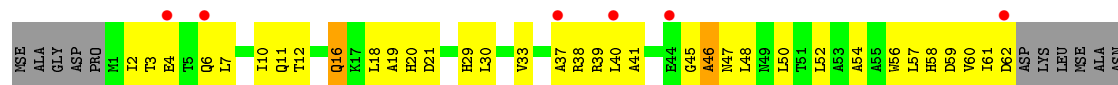
Chain	Residue	Modelled	Actual	Comment	Reference
D	66	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
D	100	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
D	158	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
D	185	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
D	201	MSE	MET	MODIFIED RESIDUE	UNP Q88U62
D	219	LEU	-	CLONING ARTIFACT	UNP Q88U62
D	220	GLU	-	CLONING ARTIFACT	UNP Q88U62
D	221	HIS	-	CLONING ARTIFACT	UNP Q88U62
D	222	HIS	-	CLONING ARTIFACT	UNP Q88U62
D	223	HIS	-	CLONING ARTIFACT	UNP Q88U62
D	224	HIS	-	CLONING ARTIFACT	UNP Q88U62
D	225	HIS	-	CLONING ARTIFACT	UNP Q88U62
D	226	HIS	-	CLONING ARTIFACT	UNP Q88U62

- Molecule 2 is water.

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



• Molecule 1: Uncharacterized protein lp_2664



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	79.93Å 167.44Å 63.77Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.80 47.78 – 2.80	Depositor EDS
% Data completeness (in resolution range)	87.1 (20.00-2.80) 96.4 (47.78-2.80)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.85 (at 2.81Å)	Xtrriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.236 , 0.307 0.252 , 0.325	Depositor DCC
R_{free} test set	3984 reflections (9.83%)	wwPDB-VP
Wilson B-factor (Å ²)	32.7	Xtrriage
Anisotropy	0.176	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 40.5	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	0.86	EDS
Total number of atoms	6575	wwPDB-VP
Average B, all atoms (Å ²)	32.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.09% 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	0.43	0/1653	0.69	0/2232
1	B	0.50	0/1695	0.74	6/2287 (0.3%)
1	C	0.70	5/1696 (0.3%)	0.78	5/2290 (0.2%)
1	D	0.62	5/1631 (0.3%)	0.86	11/2200 (0.5%)
All	All	0.57	10/6675 (0.1%)	0.77	22/9009 (0.2%)

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	D	0	4

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	60	VAL	CA-CB	-12.41	1.28	1.54
1	C	60	VAL	CB-CG1	-10.32	1.31	1.52
1	D	84	ASN	CA-C	-9.27	1.28	1.52
1	C	60	VAL	CB-CG2	-8.10	1.35	1.52
1	D	82	ALA	CA-CB	-7.50	1.36	1.52

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	83	GLN	CA-C-N	-9.64	95.99	117.20
1	C	61	ILE	CA-C-N	8.79	136.54	117.20
1	C	66	MSE	CG-SE-CE	-8.73	79.68	98.90
1	D	84	ASN	CA-C-N	-8.42	98.67	117.20
1	D	85	VAL	CA-CB-CG1	-8.40	98.31	110.90

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	83	GLN	Mainchain,Peptide
1	D	84	ASN	Mainchain
1	D	85	VAL	Mainchain

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	1625	0	1590	90	0
1	B	1667	0	1642	98	0
1	C	1667	0	1635	136	0
1	D	1604	0	1576	121	0
2	A	2	0	0	1	0
2	B	6	0	0	1	0
2	C	3	0	0	0	0
2	D	1	0	0	0	0
All	All	6575	0	6443	427	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 33.

The worst 5 of 427 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:C:60:VAL:O	1:C:61:ILE:HG22	1.45	1.12
1:B:18:LEU:HD12	1:B:30:LEU:HD11	1.42	0.99
1:D:38:ARG:HH12	1:D:48:LEU:HD21	1.27	0.99
1:C:187:THR:HG22	1:C:188:ASP:H	1.26	0.98
1:C:174:PHE:HA	1:C:178:LEU:HB2	1.44	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	202/231 (87%)	171 (85%)	23 (11%)	8 (4%)	3	9
1	B	208/231 (90%)	179 (86%)	24 (12%)	5 (2%)	6	20
1	C	208/231 (90%)	172 (83%)	24 (12%)	12 (6%)	1	4
1	D	197/231 (85%)	163 (83%)	32 (16%)	2 (1%)	15	44
All	All	815/924 (88%)	685 (84%)	103 (13%)	27 (3%)	4	13

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	85	VAL
1	A	97	ILE
1	A	113	SER
1	C	82	ALA
1	C	187	THR

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	165/178 (93%)	151 (92%)	14 (8%)	10	31
1	B	169/178 (95%)	156 (92%)	13 (8%)	13	35
1	C	169/178 (95%)	156 (92%)	13 (8%)	13	35
1	D	163/178 (92%)	156 (96%)	7 (4%)	29	62
All	All	666/712 (94%)	619 (93%)	47 (7%)	14	39

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	66	MSE
1	B	187	THR
1	D	86	THR
1	B	74	GLN
1	C	7	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	99	HIS
1	C	6	GLN
1	D	110	GLN
1	B	120	GLN
1	C	16	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 carbohydrates 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

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	201/231 (87%)	-0.14	5 (2%) 57 47	2, 17, 44, 63	0
1	B	206/231 (89%)	0.03	10 (4%) 29 20	1, 18, 82, 97	0
1	C	206/231 (89%)	0.37	15 (7%) 15 8	3, 36, 77, 101	0
1	D	198/231 (85%)	0.59	20 (10%) 7 4	3, 38, 83, 89	0
All	All	811/924 (87%)	0.21	50 (6%) 20 13	1, 26, 78, 101	0

The worst 5 of 50 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	164	ARG	8.0
1	C	108	GLY	7.9
1	C	107	ASN	6.0
1	C	110	GLN	6.0
1	B	157	HIS	5.4

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 carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.