WORLDWIDE PDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 02:02 am BST

PDB ID	:	3I3U
Title	:	Crystal structure of lp_1913 protein from lactobacillus plantarum, northeast structural genomics Consortium target lpr140a
Authors	:	Seetharaman, J.; Chen, Y.; Forouhar, F.; Sahdev, S.; Janjua, H.; Xiao, R.; Ciccosanti, C.; Foote, E.L.; Acton, T.B.; Rost, B.; Montelione, G.T.; Hunt, J.F.; Tong, L.; Northeast Structural Genomics Consortium (NESG)
Deposited on	:	2009-06-30
Resolution	:	2.80 Å(reported)

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.
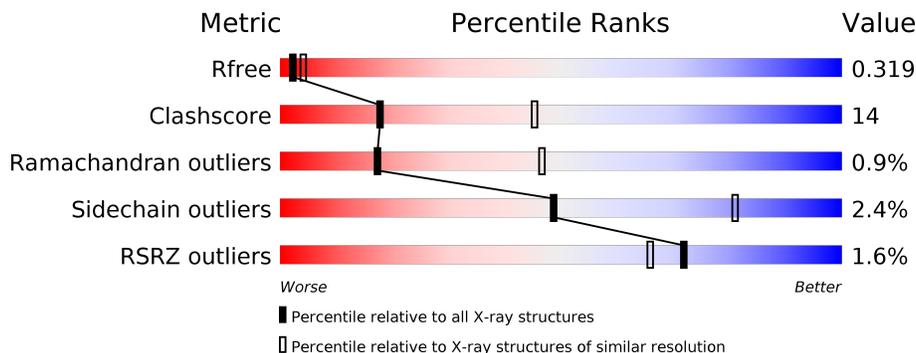
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 <=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	112	54% 19% • 27%
1	B	112	2% 54% 21% • 23%
1	C	112	52% 23% 25%
1	D	112	3% 60% 17% • 22%
1	E	112	2% 52% 19% • 27%
1	F	112	% 54% 22% • 23%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3970 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_1913.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	82	Total 629	C 408	N 100	O 119	Se 2	0	0	0
1	B	86	Total 666	C 433	N 105	O 126	Se 2	0	0	0
1	C	84	Total 655	C 427	N 104	O 122	Se 2	0	0	0
1	D	87	Total 670	C 437	N 106	O 125	Se 2	0	0	0
1	E	82	Total 635	C 416	N 97	O 120	Se 2	0	0	0
1	F	86	Total 665	C 432	N 105	O 126	Se 2	0	0	0

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	7	Total 7	O 7	0	0
2	B	9	Total 9	O 9	0	0
2	C	10	Total 10	O 10	0	0
2	D	9	Total 9	O 9	0	0
2	E	8	Total 8	O 8	0	0
2	F	7	Total 7	O 7	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA 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.
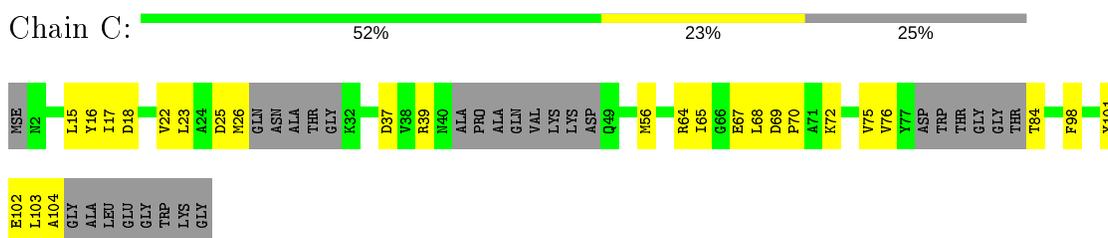
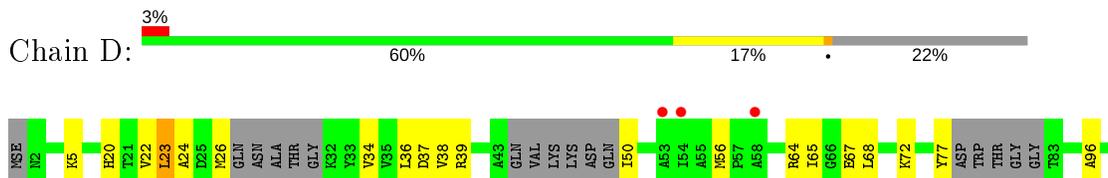- Molecule 1: uncharacterized protein lp_1913
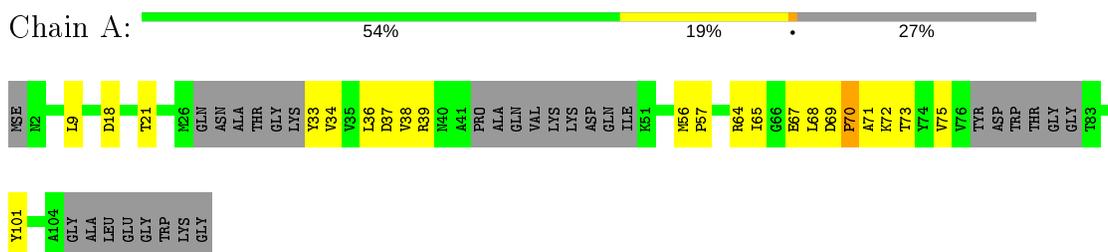
- Molecule 1: uncharacterized protein lp_1913
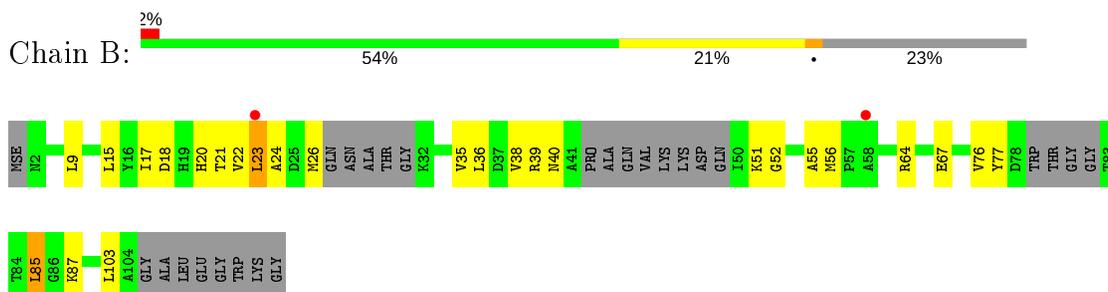
- Molecule 1: uncharacterized protein lp_1913

- Molecule 1: uncharacterized protein lp_1913

A104 GLY ALA LEU GLU GLY TRP LYS GLY

• Molecule 1: uncharacterized protein lp_1913

Chain E:

2%

52%　　19%　・　　27%

MSE N2 L8 L13 S14 L15 Y16 I17 ASP HIS HIS T21 V22 L23 A24 D25 M26 GLN ASN ALA THR GLY K32 V35 L36 R39 N40 A41 PRO ALA GLN VAL LYS LYS ASP GLN ILE LYS G52 A53 M56 R64 I65 G66 E67 L68 Y74 V75 W79 THR GLY GLY

T83 V92 F98 A104 GLY ALA LEU GLU GLY TRP LYS GLY

• Molecule 1: uncharacterized protein lp_1913

Chain F:

%

54%　　22%　・　　23%

MSE N2 T11 Y16 I17 D18 H19 H20 T21 V22 L23 M26 GLN ASN ALA T30 G31 K32 Y33 D37 V38 R39 N40 ALA PRO ALA GLN VAL LYS LYS ASP GLN I50 A53 M56 P57 R64 I65 L68 D69 K72 V76 Y77 D78 TRP THR GLY GLY THR T84

L85 G86 K87 E102 L103 A104 GLY ALA LEU GLU GLY TRP LYS GLY

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	43.71Å 60.29Å 60.32Å 120.04° 89.96° 90.03°	Depositor
Resolution (Å)	26.12 – 2.80 26.12 – 2.39	Depositor EDS
% Data completeness (in resolution range)	94.5 (26.12-2.80) 95.1 (26.12-2.39)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	0.04	Depositor
$< I/\sigma(I) >$ [1]	2.83 (at 2.39Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.264 , 0.291 0.273 , 0.319	Depositor DCC
R_{free} test set	1250 reflections (3.12%)	wwPDB-VP
Wilson B-factor (Å^2)	39.7	Xtriage
Anisotropy	0.219	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 44.0	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.467 for h,k+l,-k 0.467 for h,-l,k+l 0.457 for h,l,-k-l 0.457 for h,-k-l,k 0.457 for h,-k,-l 0.056 for -h,l,k 0.059 for -h,k,-k-l 0.057 for -h,-k-l,l 0.059 for -h,k+l,-l 0.063 for -h,-l,-k 0.057 for -h,-k,k+l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	3970	wwPDB-VP
Average B, all atoms (Å^2)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.97% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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	0.42	0/633	0.62	0/853
1	B	0.39	0/671	0.61	0/904
1	C	0.41	0/660	0.63	0/888
1	D	0.39	0/676	0.60	0/912
1	E	0.44	0/639	0.61	0/861
1	F	0.40	0/670	0.62	0/902
All	All	0.41	0/3949	0.62	0/5320

There are no bond length outliers.

There are no bond angle outliers.

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	629	0	654	16	0
1	B	666	0	691	23	0
1	C	655	0	683	20	0
1	D	670	0	699	15	0
1	E	635	0	658	24	0
1	F	665	0	689	16	0
2	A	7	0	0	0	0
2	B	9	0	0	0	0
2	C	10	0	0	1	0
2	D	9	0	0	0	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	E	8	0	0	1	0
2	F	7	0	0	1	0
All	All	3970	0	4074	109	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 14.

The worst 5 of 109 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:D:34:VAL:HG21	1:D:72:LYS:HD3	1.48	0.93
1:C:37:ASP:OD1	1:C:39:ARG:HG2	1.79	0.82
1:D:36:LEU:HD13	1:D:56:MSE:HG2	1.70	0.74
1:D:37:ASP:OD1	1:D:39:ARG:HD3	1.87	0.74
1:A:64:ARG:HB3	1:A:67:GLU:OE2	1.90	0.71

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	74/112 (66%)	67 (90%)	6 (8%)	1 (1%)	11	34
1	B	78/112 (70%)	71 (91%)	6 (8%)	1 (1%)	12	36
1	C	76/112 (68%)	74 (97%)	2 (3%)	0	100	100
1	D	79/112 (70%)	73 (92%)	6 (8%)	0	100	100
1	E	72/112 (64%)	65 (90%)	6 (8%)	1 (1%)	11	34
1	F	78/112 (70%)	75 (96%)	2 (3%)	1 (1%)	12	36
All	All	457/672 (68%)	425 (93%)	28 (6%)	4 (1%)	17	46

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	31	GLY
1	B	39	ARG
1	A	70	PRO
1	E	39	ARG

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	67/85 (79%)	67 (100%)	0	100	100
1	B	71/85 (84%)	67 (94%)	4 (6%)	21	51
1	C	70/85 (82%)	69 (99%)	1 (1%)	67	90
1	D	71/85 (84%)	70 (99%)	1 (1%)	67	90
1	E	67/85 (79%)	64 (96%)	3 (4%)	27	60
1	F	71/85 (84%)	70 (99%)	1 (1%)	67	90
All	All	417/510 (82%)	407 (98%)	10 (2%)	49	81

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	84	THR
1	D	23	LEU
1	E	40	ASN
1	B	87	LYS
1	E	23	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	2	ASN
1	C	2	ASN
1	C	49	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	F	2	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 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 (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(Å²)	Q<0.9
1	A	80/112 (71%)	-0.19	0	100 100	20, 48, 68, 71	0
1	B	84/112 (75%)	0.05	2 (2%)	59 49	24, 50, 75, 79	0
1	C	82/112 (73%)	-0.12	0	100 100	20, 49, 72, 74	0
1	D	85/112 (75%)	-0.12	3 (3%)	44 34	22, 49, 64, 68	0
1	E	80/112 (71%)	0.02	2 (2%)	57 47	23, 48, 69, 78	0
1	F	84/112 (75%)	-0.18	1 (1%)	79 73	22, 50, 72, 76	0
All	All	495/672 (73%)	-0.09	8 (1%)	72 66	20, 50, 71, 79	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	53	ALA	4.3
1	B	58	ALA	3.7
1	D	53	ALA	2.8
1	E	53	ALA	2.8
1	D	58	ALA	2.8

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 (i)

There are no such residues in this entry.