wwPDB X-ray Structure Validation Summary Report (i)

Jun 14, 2020 – 09:43 pm BST

PDB ID	:	3D3N
Title	:	Crystal structure of lipase/esterase (lp_2923) from Lactobacillus plantarum. Northeast Structural Genomics Consortium target LpR108
Authors	:	Forouhar, F.; Su, M.; Seetharaman, J.; Mao, L.; Janjua, H.; Xiao, R.; Ciccosanti, C.; Maglaqui, M.; Foote, E.L.; Zhao, L.; Everett, J.K.; Acton, T.B.; Montelione, G.T.; Tong, L.; Hunt, J.F.; Northeast Structural Genomics Consortium (NESG)
Deposited on	:	2008-05-12
Resolution	:	2.50 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

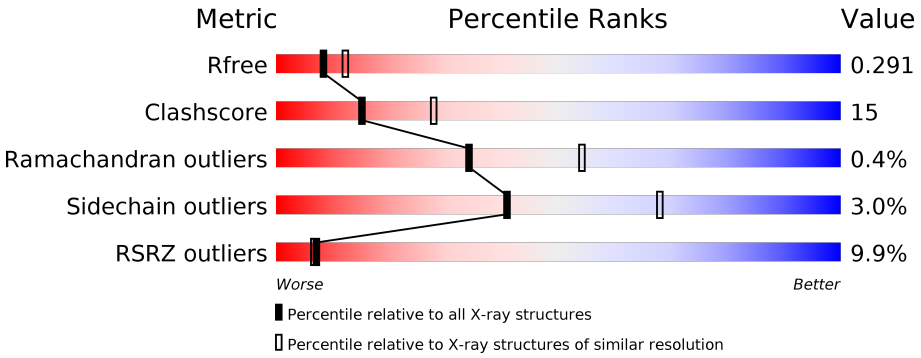
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	284	6% 60% 26% • 14%
1	B	284	11% 62% 25% • 11%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 4025 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.

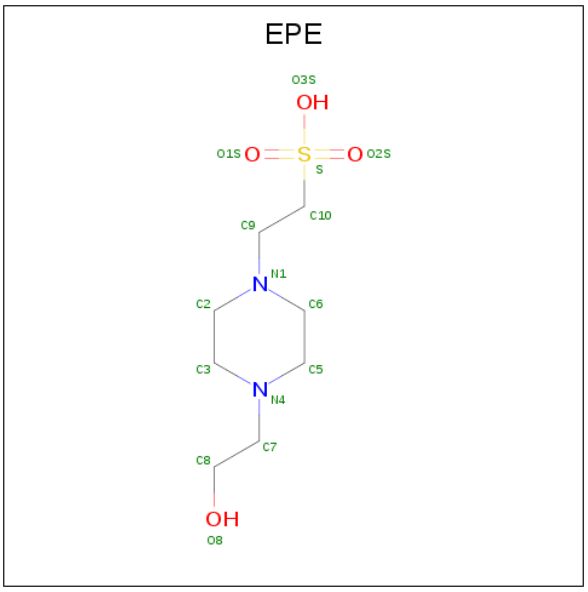
- Molecule 1 is a protein called Putative lipase/esterase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	245	Total	C	N	O	S	Se	0	0	0
			1909	1222	333	346	2	6			
1	B	254	Total	C	N	O	S	Se	0	0	0
			1987	1273	345	361	2	6			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	277	LEU	-	EXPRESSION TAG	UNP Q88TL9
A	278	GLU	-	EXPRESSION TAG	UNP Q88TL9
A	279	HIS	-	EXPRESSION TAG	UNP Q88TL9
A	280	HIS	-	EXPRESSION TAG	UNP Q88TL9
A	281	HIS	-	EXPRESSION TAG	UNP Q88TL9
A	282	HIS	-	EXPRESSION TAG	UNP Q88TL9
A	283	HIS	-	EXPRESSION TAG	UNP Q88TL9
A	284	HIS	-	EXPRESSION TAG	UNP Q88TL9
B	277	LEU	-	EXPRESSION TAG	UNP Q88TL9
B	278	GLU	-	EXPRESSION TAG	UNP Q88TL9
B	279	HIS	-	EXPRESSION TAG	UNP Q88TL9
B	280	HIS	-	EXPRESSION TAG	UNP Q88TL9
B	281	HIS	-	EXPRESSION TAG	UNP Q88TL9
B	282	HIS	-	EXPRESSION TAG	UNP Q88TL9
B	283	HIS	-	EXPRESSION TAG	UNP Q88TL9
B	284	HIS	-	EXPRESSION TAG	UNP Q88TL9

- Molecule 2 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C N O S 15 8 2 4 1	0	0
2	B	1	Total C N O S 15 8 2 4 1	0	0

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Ca 1 1	0	0

- Molecule 4 is water.

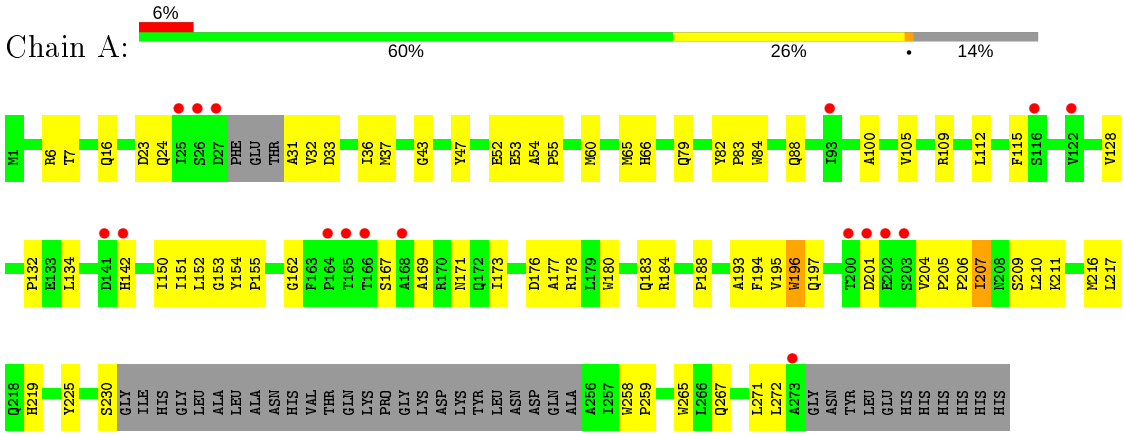
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	45	Total O 45 45	0	0
4	B	53	Total O 53 53	0	0

3 Residue-property plots (i)

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.

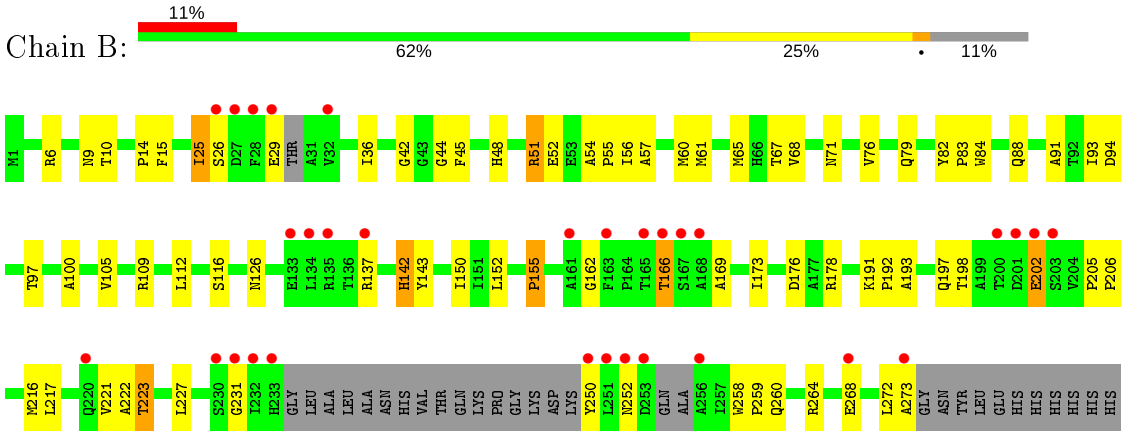
- Molecule 1: Putative lipase/esterase

Chain A:



- Molecule 1: Putative lipase/esterase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	66.07Å 93.02Å 95.77Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.94 – 2.50 29.61 – 2.50	Depositor EDS
% Data completeness (in resolution range)	66.1 (19.94-2.50) 88.6 (29.61-2.50)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	2.66 (at 2.51Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.216 , 0.263 0.245 , 0.291	Depositor DCC
R_{free} test set	3800 reflections (9.80%)	wwPDB-VP
Wilson B-factor (Å^2)	30.2	Xtriage
Anisotropy	0.439	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 46.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.000 for -h,l,k	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	4025	wwPDB-VP
Average B, all atoms (Å^2)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 92.47 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 9.0790e-09. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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

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

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	0.41	1/1956 (0.1%)	0.55	0/2670
1	B	0.38	0/2036	0.55	0/2777
All	All	0.39	1/3992 (0.0%)	0.55	0/5447

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	230	SER	CB-OG	6.38	1.50	1.42

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	1909	0	1865	58	0
1	B	1987	0	1930	60	0
2	A	15	0	17	1	0
2	B	15	0	17	1	0
3	B	1	0	0	0	0
4	A	45	0	0	1	0
4	B	53	0	0	1	0
All	All	4025	0	3829	116	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 15.

The worst 5 of 116 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:54:ALA:HB3	1:A:55:PRO:HD3	1.56	0.87
1:B:54:ALA:HB3	1:B:55:PRO:HD3	1.64	0.80
1:A:24:GLN:HE22	1:A:32:VAL:H	1.31	0.76
1:A:188:PRO:HD3	1:A:219:HIS:CD2	2.30	0.67
1:B:100:ALA:HA	1:B:105:VAL:HB	1.76	0.67

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	239/284 (84%)	222 (93%)	17 (7%)	0	100 100
1	B	246/284 (87%)	228 (93%)	16 (6%)	2 (1%)	19 35
All	All	485/568 (85%)	450 (93%)	33 (7%)	2 (0%)	34 54

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	25	ILE
1	B	76	VAL

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	196/222 (88%)	192 (98%)	4 (2%)	55 79
1	B	204/222 (92%)	196 (96%)	8 (4%)	32 57
All	All	400/444 (90%)	388 (97%)	12 (3%)	41 68

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	51	ARG
1	B	142	HIS
1	B	202	GLU
1	B	6	ARG
1	B	166	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	219	HIS
1	B	71	ASN
1	B	171	ASN
1	A	218	GLN
1	B	146	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)

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	EPE	A	285	-	15,15,15	0.82	0	18,20,20	1.02	1 (5%)
2	EPE	B	286	-	15,15,15	0.98	0	18,20,20	1.11	2 (11%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	EPE	A	285	-	-	4/9/19/19	0/1/1/1
2	EPE	B	286	-	-	1/9/19/19	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	286	EPE	C5-N4-C3	2.50	114.45	108.83
2	A	285	EPE	C5-N4-C3	2.22	113.83	108.83
2	B	286	EPE	C6-N1-C2	2.11	113.57	108.83

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	286	EPE	N4-C7-C8-O8
2	A	285	EPE	C10-C9-N1-C6
2	A	285	EPE	C8-C7-N4-C5
2	A	285	EPE	C8-C7-N4-C3
2	A	285	EPE	C10-C9-N1-C2

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	285	EPE	1	0
2	B	286	EPE	1	0

5.7 Other polymers

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

6.1 Protein, DNA and RNA chains

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^{th} 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	239/284 (84%)	0.56	17 (7%) 16 16	11, 28, 46, 58	3 (1%)
1	B	248/284 (87%)	0.75	31 (12%) 3 3	11, 29, 48, 58	3 (1%)
All	All	487/568 (85%)	0.65	48 (9%) 7 7	11, 28, 47, 58	6 (1%)

The worst 5 of 48 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	232	ILE	6.5
1	B	202	GLU	5.5
1	B	253	ASP	5.5
1	B	27	ASP	5.2
1	B	233	HIS	5.2

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

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^{th} 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^2)	Q<0.9
2	EPE	B	286	15/15	0.87	0.30	54,57,59,60	0
2	EPE	A	285	15/15	0.90	0.25	48,52,55,56	0
3	CA	B	285	1/1	0.91	0.08	25,25,25,25	0

6.5 Other polymers (i)

There are no such residues in this entry.