



Full wwPDB X-ray Structure Validation Report ⓘ

Apr 6, 2023 – 10:17 am BST

PDB ID : 8AGG
Title : Structure of the Legionella phosphocholine hydrolase Lem3
Authors : Kaspers, M.S.; Itzen, A.; Pogenberg, V.
Deposited on : 2022-07-20
Resolution : 3.60 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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
Xtriage (Phenix) : 1.13
EDS : 2.32.2
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) : 2.32.2

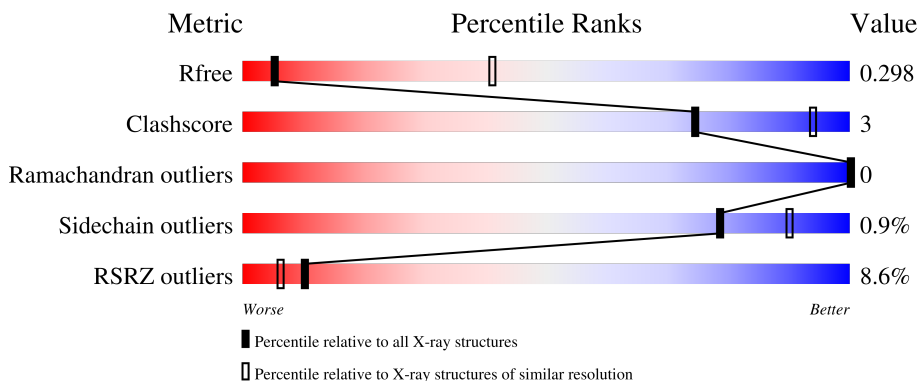
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.60 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	1257 (3.70-3.50)
Clashscore	141614	1353 (3.70-3.50)
Ramachandran outliers	138981	1307 (3.70-3.50)
Sidechain outliers	138945	1307 (3.70-3.50)
RSRZ outliers	127900	1161 (3.70-3.50)

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 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	572	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3843 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 Phosphocholine hydrolase Lem3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	489	3842	2417	656	756	13	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP Q5ZXN5
A	0	HIS	-	expression tag	UNP Q5ZXN5

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Mg	0	0
			1	1		

4 Data and refinement statistics

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α , β , γ	78.02Å 78.02Å 382.01Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	66.54 – 3.60 67.57 – 3.60	Depositor EDS
% Data completeness (in resolution range)	98.7 (66.54-3.60) 93.2 (67.57-3.60)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.48 (at 3.58Å)	Xtrriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.263 , 0.299 0.262 , 0.298	Depositor DCC
R_{free} test set	409 reflections (4.71%)	wwPDB-VP
Wilson B-factor (Å ²)	130.0	Xtrriage
Anisotropy	0.393	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 144.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.37$, $\langle L^2 \rangle = 0.20$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	3843	wwPDB-VP
Average B, all atoms (Å ²)	171.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.57% 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](#)

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

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	0.23	0/3900	0.44	0/5263

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 [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	3842	0	3775	24	0
2	A	1	0	0	0	0
All	All	3843	0	3775	24	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 3.

All (24) 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:394:ASP:OD1	1:A:395:CYS:N	2.30	0.65
1:A:245:GLU:OE2	1:A:406:HIS:ND1	2.39	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:338:ILE:HD11	1:A:357:LEU:HD21	1.88	0.55
1:A:521:LEU:HD11	1:A:554:ILE:HG21	1.89	0.54
1:A:514:LYS:HE3	1:A:517:ILE:HD11	1.92	0.52
1:A:555:GLU:HG3	1:A:558:LEU:HD12	1.90	0.52
1:A:107:LEU:HD12	1:A:165:GLU:HB3	1.92	0.51
1:A:252:MET:SD	1:A:257:TRP:HB3	2.52	0.48
1:A:71:GLN:NE2	1:A:387:LEU:O	2.46	0.48
1:A:379:PRO:HG2	1:A:382:MET:HB3	1.95	0.48
1:A:30:ILE:HD11	1:A:66:SER:HB3	1.96	0.47
1:A:386:ASP:O	1:A:390:ARG:HG3	2.16	0.45
1:A:28:LYS:HB2	1:A:80:ILE:HD12	1.99	0.44
1:A:35:VAL:HG12	1:A:319:ILE:HD13	1.98	0.44
1:A:499:LEU:HG	1:A:550:PHE:HB2	2.00	0.44
1:A:85:GLY:HA3	1:A:132:LEU:HD11	2.00	0.43
1:A:142:LEU:O	1:A:146:THR:HG22	2.19	0.43
1:A:50:ARG:HH22	1:A:69:SER:HB2	1.84	0.42
1:A:496:GLN:HB2	1:A:546:LEU:HD13	2.01	0.42
1:A:196:LEU:HB2	1:A:248:ILE:HB	2.03	0.41
1:A:100:ARG:HG2	1:A:172:ILE:HG12	2.01	0.41
1:A:450:VAL:HG13	1:A:466:ALA:HB2	2.02	0.40
1:A:190:ASP:HB3	1:A:254:ASP:HB3	2.03	0.40
1:A:225:ALA:O	1:A:229:THR:HG22	2.22	0.40

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	477/572 (83%)	455 (95%)	22 (5%)	0	100 100

There are no Ramachandran outliers to report.

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	424/513 (83%)	420 (99%)	4 (1%)	78 90

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	215	PHE
1	A	293	TYR
1	A	382	MET
1	A	408	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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](#)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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

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	489/572 (85%)	0.48	42 (8%) 10 6	121, 167, 221, 262	0

All (42) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	186	ALA	4.8
1	A	194	ILE	4.3
1	A	424	LEU	4.3
1	A	334	SER	4.1
1	A	192	LEU	3.9
1	A	467	PHE	3.8
1	A	193	ILE	3.7
1	A	168	MET	3.7
1	A	195	ILE	3.6
1	A	196	LEU	3.5
1	A	252	MET	3.3
1	A	266	ALA	3.2
1	A	84	SER	3.1
1	A	22	LYS	3.1
1	A	308	MET	3.0
1	A	47	LEU	3.0
1	A	267	GLN	3.0
1	A	242	LYS	2.9
1	A	260	LEU	2.9
1	A	423	ILE	2.9
1	A	398	ILE	2.8
1	A	481	HIS	2.7
1	A	281	PHE	2.7
1	A	425	ALA	2.7
1	A	345	PHE	2.7
1	A	248	ILE	2.7
1	A	325	ILE	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	33	GLN	2.6
1	A	250	ILE	2.5
1	A	204	HIS	2.5
1	A	183	GLY	2.5
1	A	411	ILE	2.4
1	A	458	ILE	2.3
1	A	442	PHE	2.3
1	A	299	ILE	2.3
1	A	401	THR	2.2
1	A	502	LEU	2.2
1	A	414	PHE	2.2
1	A	251	SER	2.2
1	A	304	THR	2.1
1	A	505	TYR	2.0
1	A	120	ILE	2.0

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 monosaccharides in this entry.

6.4 Ligands [i](#)

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, 95th 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	MG	A	601	1/1	0.57	0.15	196,196,196,196	0

6.5 Other polymers [i](#)

There are no such residues in this entry.