W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

May 22, 2020 – 04:51 pm BST

PDB ID	:	5LJ8
Title	:	Structure of the E. coli MacB periplasmic domain (P21)
Authors	:	Crow, A.
Deposited on	:	2016-07-18
Resolution	:	1.95 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:

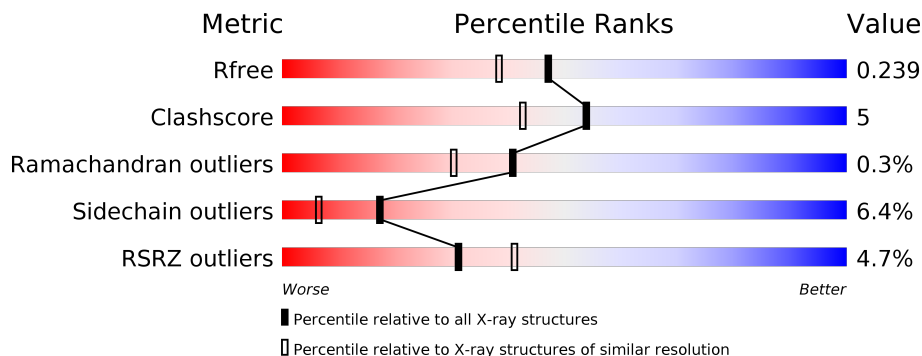|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 1.95 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	Percentile Ranks	Value
Rfree		0.239
Clashscore		5
Ramachandran outliers		0.3%
Sidechain outliers		6.4%
RSRZ outliers		4.7%

Worse   Better

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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	209	2% — 82% ... 13% · ·
1	B	209	7% — 81% ... 11% · 6%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3365 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 Macrolide export ATP-binding/permease protein MacB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	205	Total 1680	C 1053	N 301	O 320	S 6	0	6	0
1	B	197	Total 1586	C 996	N 280	O 304	S 6	0	4	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	308	MET	-	initiating methionine	UNP P75831
A	509	LEU	-	expression tag	UNP P75831
A	510	GLU	-	expression tag	UNP P75831
A	511	HIS	-	expression tag	UNP P75831
A	512	HIS	-	expression tag	UNP P75831
A	513	HIS	-	expression tag	UNP P75831
A	514	HIS	-	expression tag	UNP P75831
A	515	HIS	-	expression tag	UNP P75831
A	516	HIS	-	expression tag	UNP P75831
B	308	MET	-	initiating methionine	UNP P75831
B	509	LEU	-	expression tag	UNP P75831
B	510	GLU	-	expression tag	UNP P75831
B	511	HIS	-	expression tag	UNP P75831
B	512	HIS	-	expression tag	UNP P75831
B	513	HIS	-	expression tag	UNP P75831
B	514	HIS	-	expression tag	UNP P75831
B	515	HIS	-	expression tag	UNP P75831
B	516	HIS	-	expression tag	UNP P75831

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	49	Total 49	O 49	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	50	Total 50	O 50	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: Macrolide export ATP-binding/permease protein MacB
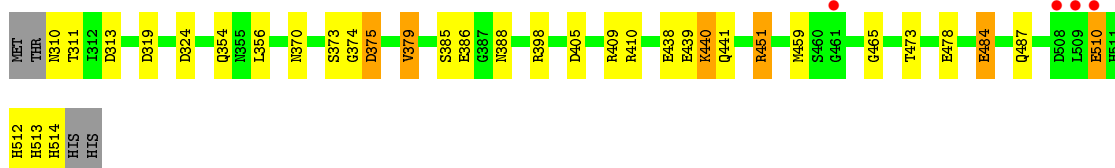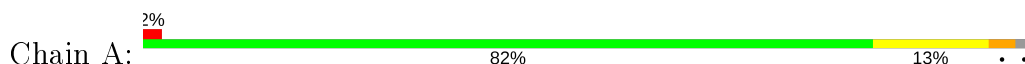
- Molecule 1: Macrolide export ATP-binding/permease protein MacB
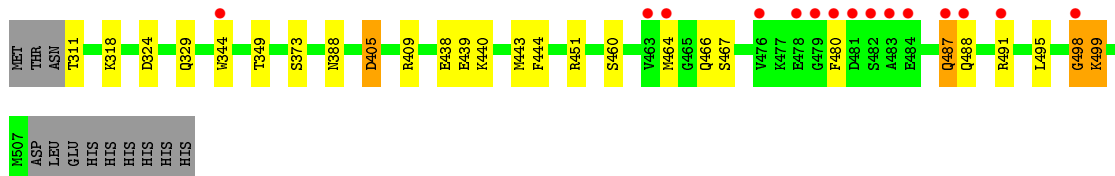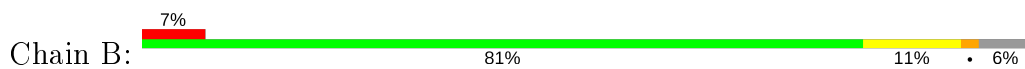
4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	50.11Å 54.86Å 66.89Å 90.00° 92.86° 90.00°	Depositor
Resolution (Å)	32.87 – 1.95 32.87 – 1.95	Depositor EDS
% Data completeness (in resolution range)	99.8 (32.87-1.95) 99.8 (32.87-1.95)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	1.95 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.8.0069	Depositor
R, R_{free}	0.178 , 0.228 0.190 , 0.239	Depositor DCC
R_{free} test set	1309 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å^2)	21.0	Xtriage
Anisotropy	0.031	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 56.1	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.047 for h,-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3365	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.06% 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 (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	0.87	0/1717	0.93	6/2327 (0.3%)
1	B	0.83	0/1618	0.88	3/2194 (0.1%)
All	All	0.85	0/3335	0.91	9/4521 (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	A	0	3
1	B	0	1
All	All	0	4

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	451	ARG	NE-CZ-NH2	-11.30	114.65	120.30
1	A	451	ARG	NE-CZ-NH2	8.66	124.63	120.30
1	A	451	ARG	NE-CZ-NH1	-7.83	116.38	120.30
1	A	405	ASP	CB-CG-OD1	6.61	124.25	118.30
1	B	405	ASP	CB-CG-OD1	5.90	123.61	118.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	310	ASN	Peptide
1	A	373	SER	Peptide
1	A	374	GLY	Peptide
1	B	498	GLY	Peptide

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	1680	0	1614	20	0
1	B	1586	0	1541	17	0
2	A	49	0	0	3	0
2	B	50	0	0	1	0
All	All	3365	0	3155	34	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 5.

The worst 5 of 34 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:398:ARG:HH11	1:A:398:ARG:HG3	1.51	0.74
1:A:375:ASP:OD1	1:B:373:SER:OG	2.07	0.73
1:B:498:GLY:HA3	1:B:499:LYS:HD2	1.76	0.68
1:A:386:GLU:OE1	2:A:601:HOH:O	2.14	0.65
1:B:405:ASP:HB2	1:B:439:GLU:OE2	1.99	0.62

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	209/209 (100%)	206 (99%)	2 (1%)	1 (0%)	29	17
1	B	199/209 (95%)	197 (99%)	2 (1%)	0	100	100

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	408/418 (98%)	403 (99%)	4 (1%)	1 (0%)	41 38

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	512	HIS

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	182/180 (101%)	170 (93%)	12 (7%)	16 6
1	B	172/180 (96%)	160 (93%)	12 (7%)	15 5
All	All	354/360 (98%)	330 (93%)	24 (7%)	17 5

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	484	GLU
1	B	318	LYS
1	B	495	LEU
1	A	510	GLU
1	B	311	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	514	HIS
1	B	487	GLN
1	B	355	ASN
1	A	378	ASN
1	B	342	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 (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	205/209 (98%)	0.15	4 (1%)	65	73	11, 22, 45, 57	0
1	B	197/209 (94%)	0.34	15 (7%)	13	21	12, 23, 51, 62	0
All	All	402/418 (96%)	0.24	19 (4%)	31	41	11, 22, 48, 62	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	480	PHE	5.4
1	B	479	GLY	5.1
1	B	483	ALA	4.5
1	B	498	GLY	4.3
1	A	461	GLY	4.3

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.