



# Full wwPDB X-ray Structure Validation Report ⓘ

May 21, 2020 – 08:43 am BST

PDB ID : 5KZT  
Title : *Listeria monocytogenes* OppA bound to peptide  
Authors : Light, S.H.; Whiteley, A.T.; Minasov, G.; Portnoy, D.A.; Anderson, W.F.;  
Center for Structural Genomics of Infectious Diseases (CSGID)  
Deposited on : 2016-07-25  
Resolution : 1.85 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](mailto: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
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

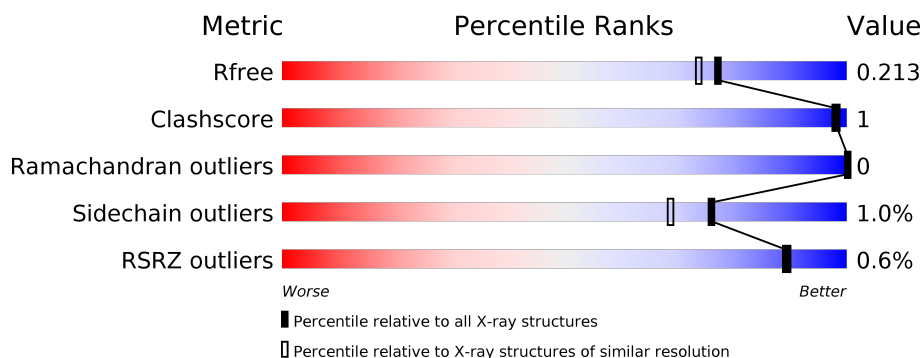
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.85 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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  $\geq 3$ , 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	536	<div> <div style="width: 93%;"></div> <div style="width: 5%;"></div> <div>93%</div> <div>• 5%</div> </div>
1	B	536	<div> <div style="width: 92%;"></div> <div style="width: 5%;"></div> <div>92%</div> <div>• 5%</div> </div>
2	C	6	<div> <div style="width: 17%;"></div> <div style="width: 83%;"></div> <div style="width: 17%;"></div> <div>17%</div> <div>83%</div> <div>17%</div> </div>
3	D	6	<div> <div style="width: 17%;"></div> <div style="width: 100%;"></div> <div>17%</div> <div>100%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 9391 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 Peptide/nickel transport system substrate-binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	509	Total	C	N	O	S	0	2	0
			4089	2588	676	821	4			
1	B	510	Total	C	N	O	S	0	4	0
			4108	2600	679	825	4			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	31	MET	-	initiating methionine	UNP A0A0H3GJB6
A	559	LEU	-	expression tag	UNP A0A0H3GJB6
A	560	GLU	-	expression tag	UNP A0A0H3GJB6
A	561	HIS	-	expression tag	UNP A0A0H3GJB6
A	562	HIS	-	expression tag	UNP A0A0H3GJB6
A	563	HIS	-	expression tag	UNP A0A0H3GJB6
A	564	HIS	-	expression tag	UNP A0A0H3GJB6
A	565	HIS	-	expression tag	UNP A0A0H3GJB6
A	566	HIS	-	expression tag	UNP A0A0H3GJB6
B	31	MET	-	initiating methionine	UNP A0A0H3GJB6
B	559	LEU	-	expression tag	UNP A0A0H3GJB6
B	560	GLU	-	expression tag	UNP A0A0H3GJB6
B	561	HIS	-	expression tag	UNP A0A0H3GJB6
B	562	HIS	-	expression tag	UNP A0A0H3GJB6
B	563	HIS	-	expression tag	UNP A0A0H3GJB6
B	564	HIS	-	expression tag	UNP A0A0H3GJB6
B	565	HIS	-	expression tag	UNP A0A0H3GJB6
B	566	HIS	-	expression tag	UNP A0A0H3GJB6

- Molecule 2 is a protein called Hexamer peptide: SER-ASP-GLU-SER-LYS-GLY.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	6	Total	C	N	O	0	0	0
			42	23	7	12			

- Molecule 3 is a protein called Hexamer peptide: SER-ASP-GLU-SER-SER-GLY.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	D	6	Total	C	N	O	0	0	0
			39	20	6	13			

- Molecule 4 is water.

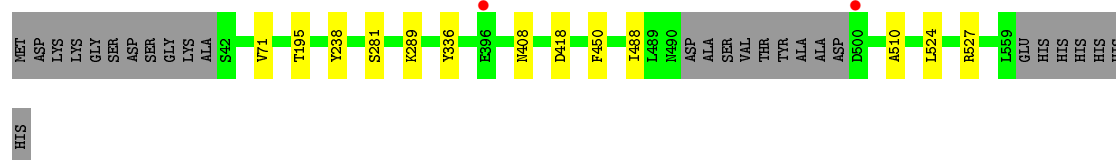
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	514	Total	O	0	0
			514	514		
4	B	597	Total	O	0	0
			597	597		
4	D	2	Total	O	0	0
			2	2		

### 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.

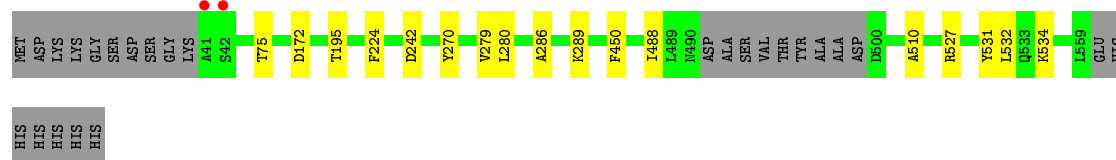
- Molecule 1: Peptide/nickel transport system substrate-binding protein

Chain A: 




- Molecule 1: Peptide/nickel transport system substrate-binding protein

Chain B: 



- Molecule 2: Hexamer peptide: SER-ASP-GLU-SER-LYS-GLY

Chain C: 



- Molecule 3: Hexamer peptide: SER-ASP-GLU-SER-SER-GLY

Chain D: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	100.28 Å   74.63 Å   155.06 Å 90.00°   109.38°   90.00°	Depositor
Resolution (Å)	30.00 – 1.85 29.63 – 1.83	Depositor EDS
% Data completeness (in resolution range)	97.0 (30.00-1.85) 97.1 (29.63-1.83)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.48 (at 1.83 Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, $R_{free}$	0.167   ,   0.205 0.173   ,   0.213	Depositor DCC
$R_{free}$ test set	4604 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.9	Xtriage
Anisotropy	0.208	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33   ,   40.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.088 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	9391	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.77% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	0.48	0/4171	0.68	1/5649 (0.0%)
1	B	0.50	0/4190	0.69	1/5676 (0.0%)
2	C	0.41	0/41	0.59	0/52
3	D	0.40	0/38	0.47	0/49
All	All	0.49	0/8440	0.68	2/11426 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	172	ASP	CB-CG-OD1	5.72	123.45	118.30
1	A	527	ARG	NE-CZ-NH2	5.20	122.90	120.30

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	4089	0	3990	4	0
1	B	4108	0	4009	6	0
2	C	42	0	38	1	0
3	D	39	0	30	0	0
4	A	514	0	0	0	0
4	B	597	0	0	0	0
4	D	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	9391	0	8067	10	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 1.

All (10) 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:336:TYR:OH	2:C:5:LYS:NZ	2.26	0.69
1:B:279:VAL:HG22	1:B:531:TYR:CE2	2.46	0.51
1:A:488:ILE:HG21	1:A:510:ALA:HB2	1.96	0.47
1:B:286:ALA:O	1:B:289:LYS:HG2	2.16	0.45
1:B:280:LEU:HD11	1:B:532:LEU:HG	1.99	0.44
1:B:75:THR:HA	1:B:224:PHE:O	2.19	0.42
1:B:488:ILE:HG21	1:B:510:ALA:HB2	2.03	0.41
1:A:281:SER:OG	1:A:408:ASN:CG	2.59	0.41
1:B:270:TYR:CZ	1:B:534:LYS:HE2	2.55	0.41
1:A:71:VAL:HA	1:A:238:TYR:OH	2.21	0.41

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	507/536 (95%)	496 (98%)	11 (2%)	0	100	100
1	B	510/536 (95%)	497 (98%)	13 (2%)	0	100	100
2	C	4/6 (67%)	4 (100%)	0	0	100	100
3	D	4/6 (67%)	4 (100%)	0	0	100	100
All	All	1025/1084 (95%)	1001 (98%)	24 (2%)	0	100	100



There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

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	448/467 (96%)	442 (99%)	6 (1%)	69	58
1	B	450/467 (96%)	446 (99%)	4 (1%)	78	72
2	C	5/5 (100%)	5 (100%)	0	100	100
3	D	5/5 (100%)	5 (100%)	0	100	100
All	All	908/944 (96%)	898 (99%)	10 (1%)	76	65

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	195	THR
1	A	289	LYS
1	A	418	ASP
1	A	450	PHE
1	A	524[A]	LEU
1	A	524[B]	LEU
1	B	195	THR
1	B	242	ASP
1	B	450	PHE
1	B	527	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	213	GLN
1	B	138	ASN
1	B	299	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	509/536 (94%)	-0.42	2 (0%) 92 92	13, 22, 33, 55	0
1	B	510/536 (95%)	-0.47	2 (0%) 92 92	12, 18, 28, 47	0
2	C	6/6 (100%)	0.93	1 (16%) 1 1	20, 25, 39, 43	0
3	D	6/6 (100%)	1.22	1 (16%) 1 1	20, 27, 39, 45	0
All	All	1031/1084 (95%)	-0.43	6 (0%) 89 89	12, 20, 31, 55	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	41	ALA	3.7
2	C	6	GLY	2.8
3	D	6	GLY	2.7
1	B	42	SER	2.4
1	A	500	ASP	2.4
1	A	396	GLU	2.1

### 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.