



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 13, 2016 – 03:33 AM EST

PDB ID : 5LEK  
Title : The Transcriptional Regulator PrfA-G145S mutant from *Listeria Monocytogenes* in complex with a 30-bp operator PrfA-box motif  
Authors : Hall, M.; Grundstrom, C.; Begum, A.; Lindberg, M.; Sauer, U.H.; Almqvist, F.; Johansson, J.; Sauer-Eriksson, A.E.  
Deposited on : 2016-06-29  
Resolution : 2.80 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/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
Mogul	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20028442
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20028442

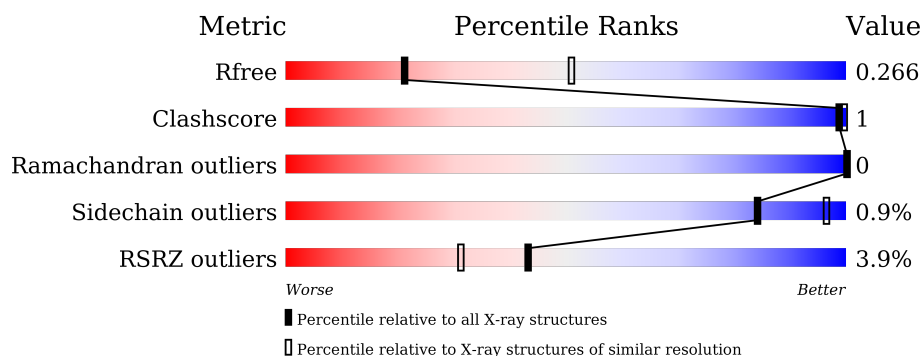
# 1 Overall quality at a glance

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))
$R_{free}$	91344	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	237	<div> <div>6%</div> <div>97%</div> <div>.</div> </div>
1	B	237	<div> <div>3%</div> <div>96%</div> <div>.</div> </div>
2	C	30	<div> <div>100%</div> </div>
3	D	30	<div> <div>100%</div> </div>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 9600 atoms, of which 4529 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 Listeriolysin regulatory protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	236	Total	C	H	N	O	S	0	0	0
			3844	1254	1922	301	360	7			
1	B	236	Total	C	H	N	O	S	0	0	0
			3845	1254	1923	301	360	7			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	145	SER	GLY	conflict	UNP P22262
B	145	SER	GLY	conflict	UNP P22262

- Molecule 2 is a DNA chain called DNA (30-MER).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	30	Total	C	H	N	O	P	0	0	0
			957	296	342	112	178	29			

- Molecule 3 is a DNA chain called DNA (30-MER).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	D	30	Total	C	H	N	O	P	0	0	0
			951	294	342	108	178	29			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	O	0	0
			1	1		
4	B	2	Total	O	0	0
			2	2		

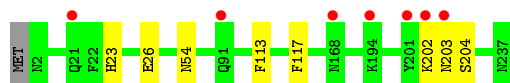
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Listeriolysin regulatory protein



- Molecule 1: Listeriolysin regulatory protein



- Molecule 2: DNA (30-MER)



There are no outlier residues recorded for this chain.

- Molecule 3: DNA (30-MER)



There are no outlier residues recorded for this chain.

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	79.08 Å 79.08 Å 265.98 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.30 – 2.80 47.30 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.2 (47.30-2.80) 99.4 (47.30-2.80)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.05 (at 2.81 Å)	Xtriage
Refinement program	PHENIX (1.10.1 _2155: ???)	Depositor
R, $R_{free}$	0.231 , 0.267 0.235 , 0.266	Depositor DCC
$R_{free}$ test set	1087 reflections (5.04%)	DCC
Wilson B-factor (Å <sup>2</sup> )	68.8	Xtriage
Anisotropy	0.454	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 42.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	9600	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	92.0	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 33.58 % 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 7.8143e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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	0.26	0/1966	0.42	0/2655
1	B	0.33	0/1966	0.60	0/2655
2	C	0.54	0/690	0.94	0/1064
3	D	0.57	0/682	0.96	0/1050
All	All	0.38	0/5304	0.67	0/7424

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	1922	1922	1922	4	0
1	B	1922	1923	1922	7	0
2	C	615	342	342	0	0
3	D	609	342	342	0	0
4	A	1	0	0	0	0
4	B	2	0	0	0	0
All	All	5071	4529	4528	8	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.

The worst 5 of 8 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:B:202:LYS:O	1:B:203:ASN:OD1	2.22	0.58
1:B:54:ASN:OD1	1:B:54:ASN:O	2.28	0.51
1:A:120:LEU:HD12	1:B:117:PHE:CE2	2.45	0.50
1:B:203:ASN:O	1:B:203:ASN:ND2	2.48	0.46
1:B:23:HIS:CD2	1:B:26:GLU:OE1	2.70	0.44

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	234/237 (99%)	224 (96%)	10 (4%)	0	100	100
1	B	234/237 (99%)	226 (97%)	8 (3%)	0	100	100
All	All	468/474 (99%)	450 (96%)	18 (4%)	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	212/213 (100%)	210 (99%)	2 (1%)	84	96
1	B	212/213 (100%)	210 (99%)	2 (1%)	84	96
All	All	424/426 (100%)	420 (99%)	4 (1%)	84	96

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

Mol	Chain	Res	Type
1	A	113	PHE
1	A	201	TYR
1	B	113	PHE
1	B	204	SER

Some 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 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	236/237 (99%)	0.53	14 (5%) 26 16	49, 78, 124, 162	0
1	B	236/237 (99%)	0.56	7 (2%) 54 41	50, 76, 118, 153	0
2	C	30/30 (100%)	0.12	0 100 100	65, 93, 113, 116	0
3	D	30/30 (100%)	0.07	0 100 100	66, 100, 111, 117	0
All	All	532/534 (99%)	0.49	21 (3%) 43 31	49, 80, 121, 162	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	203	ASN	5.4
1	B	203	ASN	4.4
1	B	201	TYR	3.9
1	B	202	LYS	3.7
1	A	202	LYS	3.3

### 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.