



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 08:53 am BST

PDB ID : 5X1H  
Title : Structure of Legionella pneumophila DotN  
Authors : Kwak, M.J.; Oh, B.H.  
Deposited on : 2017-01-26  
Resolution : 3.00 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

<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

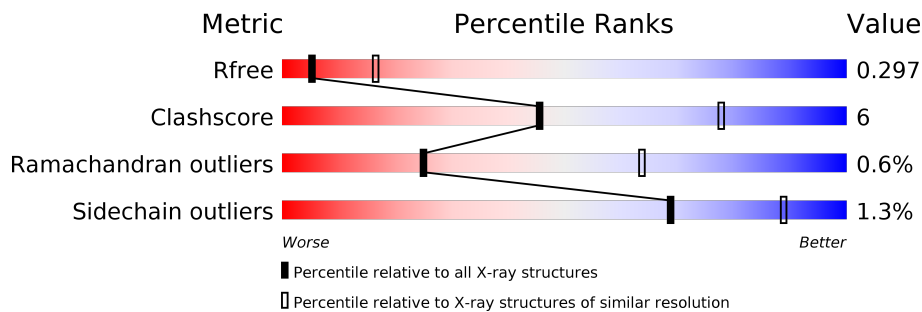
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)

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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	208	
1	C	208	
1	E	208	
1	G	208	
1	I	208	
1	K	208	
1	M	208	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	O	208	 72% 22% 6%
1	S	208	 77% 17% 6%
1	U	208	 76% 18% 6%
1	W	208	 77% 16% 6%
2	Q	195	 85% 14% .

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 18295 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 IcmJ (DotN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	199	1572	999	271	292	10	0	0	0
1	C	196	1546	983	270	283	10	0	0	0
1	E	197	1528	973	267	278	10	0	0	0
1	G	197	1492	949	259	274	10	0	0	0
1	I	197	1542	984	267	281	10	0	0	0
1	K	194	1498	953	255	281	9	0	0	0
1	M	195	1538	982	267	279	10	0	0	0
1	O	195	1458	923	254	272	9	0	0	0
1	S	195	1522	966	259	287	10	0	0	0
1	U	195	1522	965	264	283	10	0	0	0
1	W	196	1542	981	262	289	10	0	0	0

- Molecule 2 is a protein called IcmJ (DotN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	Q	195	1523	969	266	278	10	0	0	0

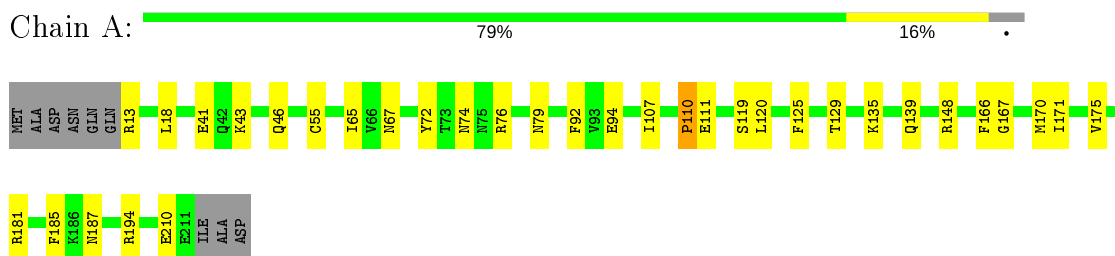
- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	G	1	Total Zn 1 1	0	0
3	Q	1	Total Zn 1 1	0	0
3	K	1	Total Zn 1 1	0	0
3	E	1	Total Zn 1 1	0	0
3	I	1	Total Zn 1 1	0	0
3	C	1	Total Zn 1 1	0	0
3	W	1	Total Zn 1 1	0	0
3	A	1	Total Zn 1 1	0	0
3	U	1	Total Zn 1 1	0	0
3	O	1	Total Zn 1 1	0	0
3	S	1	Total Zn 1 1	0	0
3	M	1	Total Zn 1 1	0	0

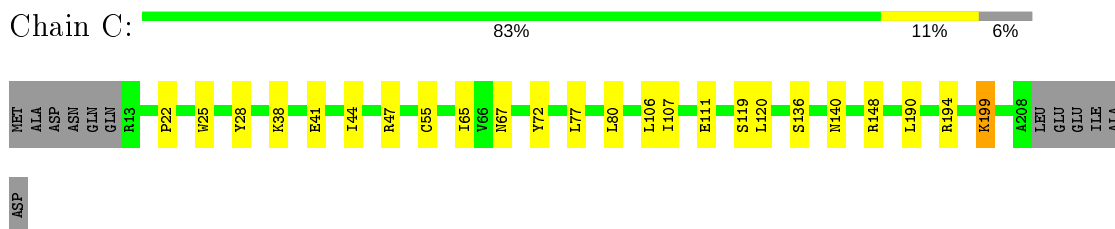
### 3 Residue-property plots [i](#)

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. 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. 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.

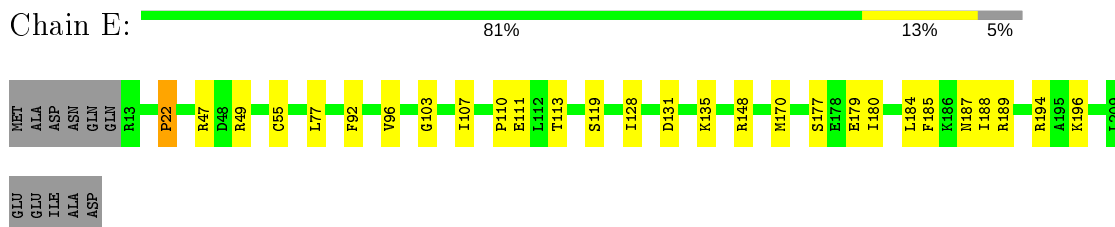
- Molecule 1: IcmJ (DotN)



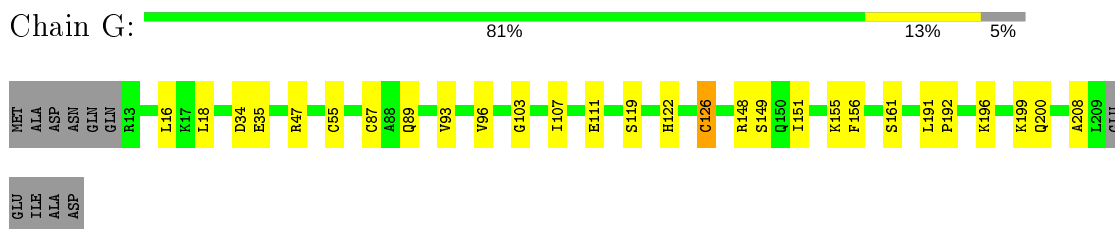
- Molecule 1: IcmJ (DotN)



- Molecule 1: IcmJ (DotN)

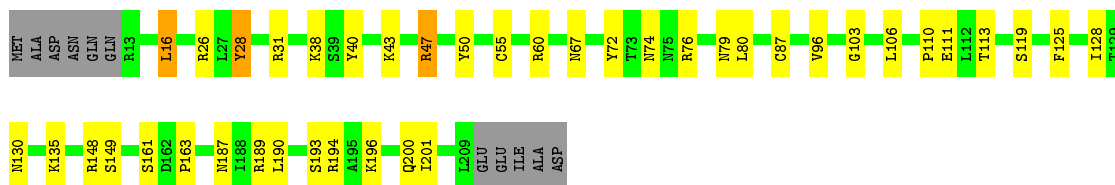


- Molecule 1: IcmJ (DotN)



- Molecule 1: IcmJ (DotN)

Chain I:  75% 18% 5%




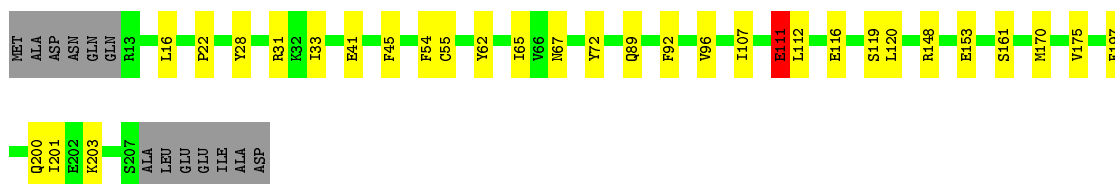
• Molecule 1: IcmJ (DotN)

Chain K:  76% 17% 7%



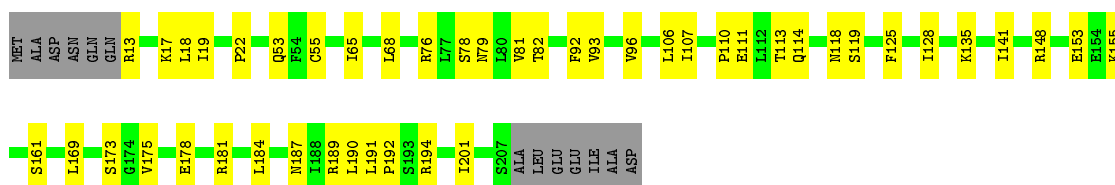
• Molecule 1: IcmJ (DotN)

Chain M:  79% 14% 6%




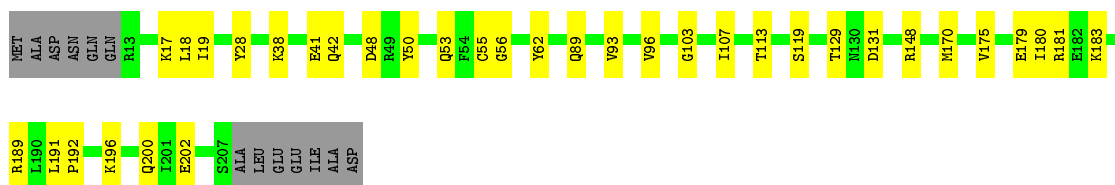
• Molecule 1: IcmJ (DotN)

Chain O:  72% 22% 6%



• Molecule 1: IcmJ (DotN)

Chain S:  77% 17% 6%



• Molecule 1: IcmJ (DotN)





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	155.36Å 155.36Å 527.71Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.93 – 3.00 49.93 – 3.01	Depositor EDS
% Data completeness (in resolution range)	77.0 (49.93-3.00) 76.7 (49.93-3.01)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.78 (at 3.01Å)	Xtrriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, $R_{free}$	0.256 , 0.297 0.256 , 0.297	Depositor DCC
$R_{free}$ test set	1788 reflections (3.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	51.7	Xtrriage
Anisotropy	0.030	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , -10.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.35$ , $\langle L^2 \rangle = 0.18$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.85	EDS
Total number of atoms	18295	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	63.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.89% 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](#)

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

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	0.24	0/1603	0.39	0/2158
1	C	0.24	0/1577	0.40	0/2124
1	E	0.24	0/1558	0.40	0/2102
1	G	0.24	0/1521	0.40	0/2056
1	I	0.44	4/1573 (0.3%)	0.40	0/2120
1	K	0.24	0/1528	0.40	0/2065
1	M	0.31	1/1569 (0.1%)	0.41	0/2113
1	O	0.24	0/1485	0.40	0/2010
1	S	0.25	0/1552	0.40	0/2094
1	U	0.24	0/1551	0.39	0/2091
1	W	0.25	0/1573	0.39	0/2121
2	Q	0.24	0/1554	0.39	0/2093
All	All	0.27	5/18644 (0.0%)	0.40	0/25147

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	I	47	ARG	CZ-NH2	-9.12	1.21	1.33
1	I	47	ARG	CZ-NH1	-7.04	1.23	1.33
1	I	47	ARG	CD-NE	-5.87	1.36	1.46
1	I	47	ARG	NE-CZ	-5.87	1.25	1.33
1	M	111	GLU	CD-OE1	-5.09	1.20	1.25

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	1572	0	1526	19	0
1	C	1546	0	1497	18	0
1	E	1528	0	1464	18	0
1	G	1492	0	1402	15	0
1	I	1542	0	1488	25	0
1	K	1498	0	1415	21	0
1	M	1538	0	1496	20	0
1	O	1458	0	1356	25	0
1	S	1522	0	1452	23	0
1	U	1522	0	1468	24	0
1	W	1542	0	1483	20	0
2	Q	1523	0	1456	16	0
3	A	1	0	0	0	0
3	C	1	0	0	0	0
3	E	1	0	0	0	0
3	G	1	0	0	0	0
3	I	1	0	0	0	0
3	K	1	0	0	0	0
3	M	1	0	0	0	0
3	O	1	0	0	0	0
3	Q	1	0	0	0	0
3	S	1	0	0	0	0
3	U	1	0	0	0	0
3	W	1	0	0	0	0
All	All	18295	0	17503	228	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 6.

The worst 5 of 228 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:K:111:GLU:HB3	1:M:148:ARG:HH12	1.44	0.81
1:S:107:ILE:HB	1:S:148:ARG:HB2	1.67	0.77
1:A:18:LEU:O	1:A:181:ARG:NH1	2.21	0.72

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:S:18:LEU:O	1:S:181:ARG:NH1	2.23	0.72
1:U:55:CYS:HB2	1:U:119:SER:HB3	1.72	0.72

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	197/208 (95%)	178 (90%)	17 (9%)	2 (1%)	15	53
1	C	194/208 (93%)	178 (92%)	15 (8%)	1 (0%)	29	68
1	E	195/208 (94%)	178 (91%)	16 (8%)	1 (0%)	29	68
1	G	195/208 (94%)	179 (92%)	15 (8%)	1 (0%)	29	68
1	I	195/208 (94%)	181 (93%)	13 (7%)	1 (0%)	29	68
1	K	192/208 (92%)	185 (96%)	7 (4%)	0	100	100
1	M	193/208 (93%)	183 (95%)	9 (5%)	1 (0%)	29	68
1	O	193/208 (93%)	176 (91%)	16 (8%)	1 (0%)	29	68
1	S	193/208 (93%)	175 (91%)	17 (9%)	1 (0%)	29	68
1	U	193/208 (93%)	176 (91%)	16 (8%)	1 (0%)	29	68
1	W	194/208 (93%)	178 (92%)	14 (7%)	2 (1%)	15	53
2	Q	193/195 (99%)	174 (90%)	16 (8%)	3 (2%)	9	40
All	All	2327/2483 (94%)	2141 (92%)	171 (7%)	15 (1%)	25	64

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	110	PRO
2	Q	175	VAL
1	O	22	PRO

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	22	PRO
1	M	22	PRO

### 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	166/179 (93%)	165 (99%)	1 (1%)	86	95
1	C	161/179 (90%)	160 (99%)	1 (1%)	86	95
1	E	155/179 (87%)	154 (99%)	1 (1%)	86	95
1	G	148/179 (83%)	146 (99%)	2 (1%)	67	88
1	I	159/179 (89%)	153 (96%)	6 (4%)	33	69
1	K	153/179 (86%)	151 (99%)	2 (1%)	69	89
1	M	160/179 (89%)	158 (99%)	2 (1%)	69	89
1	O	142/179 (79%)	140 (99%)	2 (1%)	67	88
1	S	159/179 (89%)	157 (99%)	2 (1%)	69	89
1	U	159/179 (89%)	158 (99%)	1 (1%)	86	95
1	W	162/179 (90%)	159 (98%)	3 (2%)	57	84
2	Q	155/169 (92%)	153 (99%)	2 (1%)	69	89
All	All	1879/2138 (88%)	1854 (99%)	25 (1%)	69	89

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	13	ARG
1	M	28	TYR
1	W	28	TYR
1	K	126	CYS
1	M	111	GLU

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	187	ASN
1	K	53	GLN
1	S	187	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](#)

Of 12 ligands modelled in this entry, 12 are 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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.