



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 25, 2026 – 09:20 PM EDT

PDB ID : 7EW8 / pdb\_00007ew8  
Title : Legionella pneumophila effector AnkD  
Authors : Chen, T.T.; Lin, Y.L.  
Deposited on : 2021-05-24  
Resolution : 2.59 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 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 : 4-5-2 with Phenix2.0  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

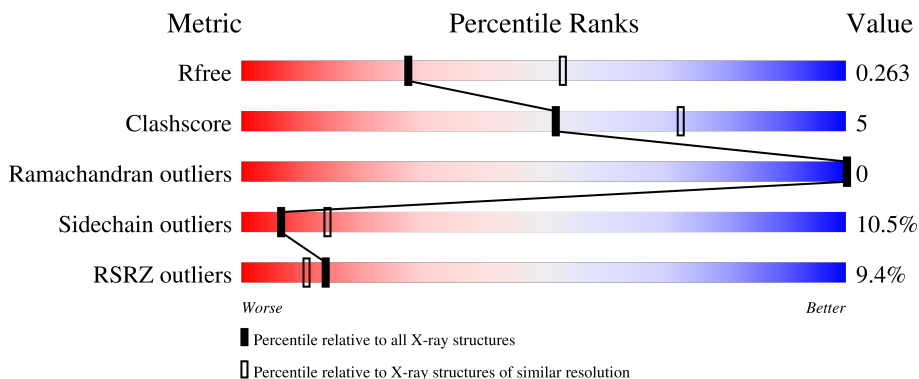
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.59 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))
$R_{free}$	180053	4008 (2.60-2.60)
Clashscore	190562	4347 (2.60-2.60)
Ramachandran outliers	187476	4277 (2.60-2.60)
Sidechain outliers	187428	4277 (2.60-2.60)
RSRZ outliers	180081	4008 (2.60-2.60)

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  $\geq 3$ , 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	494	 5% 71% 14% • 13%
1	B	494	 11% 68% 15% • • 13%

## 2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 6826 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 ANK\_REP\_REGION domain-containing protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	431	3348	2100	581	652	3	12	0	0	0
1	B	429	3328	2085	579	649	3	12	0	0	0

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-22	MSE	-	expression tag	UNP A0A2S6F2G5
A	-21	GLY	-	expression tag	UNP A0A2S6F2G5
A	-20	SER	-	expression tag	UNP A0A2S6F2G5
A	-19	SER	-	expression tag	UNP A0A2S6F2G5
A	-18	HIS	-	expression tag	UNP A0A2S6F2G5
A	-17	HIS	-	expression tag	UNP A0A2S6F2G5
A	-16	HIS	-	expression tag	UNP A0A2S6F2G5
A	-15	HIS	-	expression tag	UNP A0A2S6F2G5
A	-14	HIS	-	expression tag	UNP A0A2S6F2G5
A	-13	HIS	-	expression tag	UNP A0A2S6F2G5
A	-12	SER	-	expression tag	UNP A0A2S6F2G5
A	-11	SER	-	expression tag	UNP A0A2S6F2G5
A	-10	GLY	-	expression tag	UNP A0A2S6F2G5
A	-9	LEU	-	expression tag	UNP A0A2S6F2G5
A	-8	VAL	-	expression tag	UNP A0A2S6F2G5
A	-7	PRO	-	expression tag	UNP A0A2S6F2G5
A	-6	ARG	-	expression tag	UNP A0A2S6F2G5
A	-5	GLY	-	expression tag	UNP A0A2S6F2G5
A	-4	SER	-	expression tag	UNP A0A2S6F2G5
A	-3	HIS	-	expression tag	UNP A0A2S6F2G5
A	-2	MSE	-	expression tag	UNP A0A2S6F2G5
A	-1	ALA	-	expression tag	UNP A0A2S6F2G5
A	0	SER	-	expression tag	UNP A0A2S6F2G5
B	-22	MSE	-	expression tag	UNP A0A2S6F2G5
B	-21	GLY	-	expression tag	UNP A0A2S6F2G5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-20	SER	-	expression tag	UNP A0A2S6F2G5
B	-19	SER	-	expression tag	UNP A0A2S6F2G5
B	-18	HIS	-	expression tag	UNP A0A2S6F2G5
B	-17	HIS	-	expression tag	UNP A0A2S6F2G5
B	-16	HIS	-	expression tag	UNP A0A2S6F2G5
B	-15	HIS	-	expression tag	UNP A0A2S6F2G5
B	-14	HIS	-	expression tag	UNP A0A2S6F2G5
B	-13	HIS	-	expression tag	UNP A0A2S6F2G5
B	-12	SER	-	expression tag	UNP A0A2S6F2G5
B	-11	SER	-	expression tag	UNP A0A2S6F2G5
B	-10	GLY	-	expression tag	UNP A0A2S6F2G5
B	-9	LEU	-	expression tag	UNP A0A2S6F2G5
B	-8	VAL	-	expression tag	UNP A0A2S6F2G5
B	-7	PRO	-	expression tag	UNP A0A2S6F2G5
B	-6	ARG	-	expression tag	UNP A0A2S6F2G5
B	-5	GLY	-	expression tag	UNP A0A2S6F2G5
B	-4	SER	-	expression tag	UNP A0A2S6F2G5
B	-3	HIS	-	expression tag	UNP A0A2S6F2G5
B	-2	MSE	-	expression tag	UNP A0A2S6F2G5
B	-1	ALA	-	expression tag	UNP A0A2S6F2G5
B	0	SER	-	expression tag	UNP A0A2S6F2G5

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	115	Total O 115 115	0	0
2	B	35	Total O 35 35	0	0



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.06Å 76.00Å 104.81Å 90.00° 90.81° 90.00°	Depositor
Resolution (Å)	45.21 – 2.59 45.21 – 2.59	Depositor EDS
% Data completeness (in resolution range)	92.8 (45.21-2.59) 92.7 (45.21-2.59)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.55 (at 2.58Å)	Xtrriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, $R_{free}$	0.201 , 0.268 0.228 , 0.263	Depositor DCC
$R_{free}$ test set	1485 reflections (4.59%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.2	Xtrriage
Anisotropy	0.128	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 41.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.031 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	6826	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.22% 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	0.76	1/3397 (0.0%)	1.20	29/4540 (0.6%)
1	B	0.70	2/3375 (0.1%)	1.16	29/4508 (0.6%)
All	All	0.73	3/6772 (0.0%)	1.18	58/9048 (0.6%)

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	1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	189	SER	C-N	-7.68	1.23	1.33
1	B	376	CYS	C-N	-7.41	1.24	1.33
1	A	239	PRO	C-O	-6.34	1.16	1.23

The worst 5 of 58 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	242	ASN	N-CA-C	17.33	136.64	113.77
1	B	53	ILE	N-CA-C	-10.26	101.40	110.74
1	A	433	LYS	CB-CA-C	-10.07	93.78	110.29
1	A	56	LYS	N-CA-C	-9.87	95.30	109.31
1	B	191	PHE	N-CA-C	-9.15	91.32	110.80

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	55	ASN	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	3348	0	3346	26	0
1	B	3328	0	3325	42	0
2	A	115	0	0	0	0
2	B	35	0	0	0	0
All	All	6826	0	6671	68	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 68 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:53:ILE:HG22	1:B:54:LEU:HG	1.57	0.83
1:B:53:ILE:HG22	1:B:54:LEU:H	1.48	0.77
1:B:201:MSE:O	1:B:205:VAL:HG23	1.86	0.76
1:B:265:CYS:HB3	1:B:270:MSE:HE1	1.67	0.76
1:B:53:ILE:HG22	1:B:54:LEU:N	2.01	0.75

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	425/494 (86%)	405 (95%)	20 (5%)	0	100	100
1	B	421/494 (85%)	400 (95%)	21 (5%)	0	100	100
All	All	846/988 (86%)	805 (95%)	41 (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	368/408 (90%)	335 (91%)	33 (9%)	9	20
1	B	366/408 (90%)	322 (88%)	44 (12%)	5	10
All	All	734/816 (90%)	657 (90%)	77 (10%)	6	14

5 of 77 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	296	LYS
1	B	441	ASN
1	B	336	SER
1	B	420	LYS
1	B	448	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	172	ASN
1	B	190	GLN
1	B	327	ASN
1	B	240	GLN
1	B	247	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 oligosaccharides 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	419/494 (84%)	0.30	27 (6%) 25 20	17, 34, 69, 93	0
1	B	417/494 (84%)	0.94	52 (12%) 8 6	30, 60, 83, 105	0
All	All	836/988 (84%)	0.62	79 (9%) 14 11	17, 50, 80, 105	0

The worst 5 of 79 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	53	ILE	7.3
1	B	442	ALA	5.3
1	A	444	PRO	5.1
1	A	242	ASN	5.0
1	A	426	PRO	4.8

### 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 oligosaccharides 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.