



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 11:29 PM GMT

PDB ID : 1XJT  
Title : Crystal structure of active form of P1 phage endolysin Lyz  
Authors : Arockiasamy, A.; Sacchettini, J.C.  
Deposited on : 2004-09-24  
Resolution : 1.75 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  
<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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

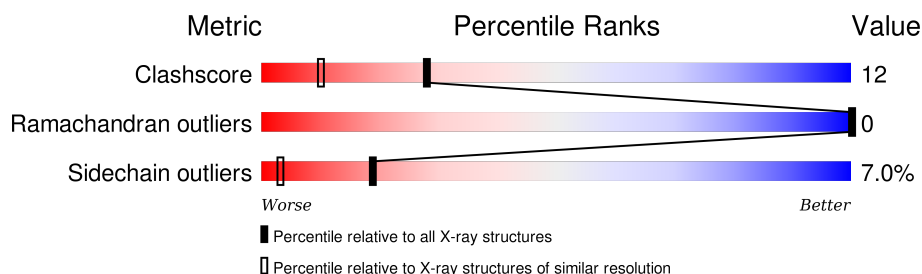
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.75 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))
Clashscore	102246	1730 (1.76-1.76)
Ramachandran outliers	100387	1711 (1.76-1.76)
Sidechain outliers	100360	1711 (1.76-1.76)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	191	 72% 17% • • 7%

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 1540 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.

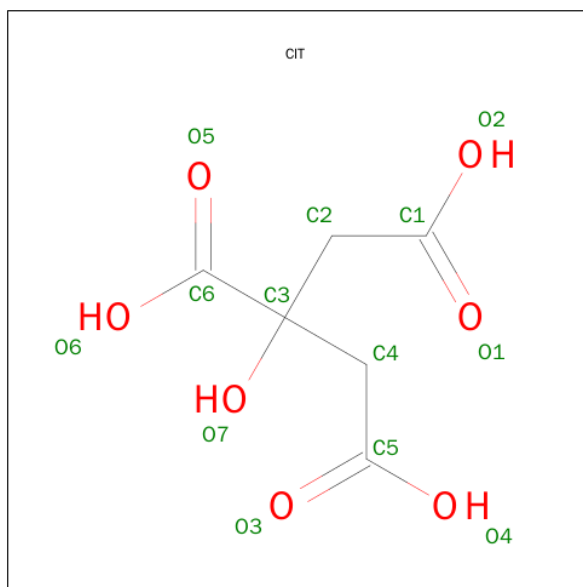
- Molecule 1 is a protein called Lysozyme.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	178	Total	C	N	O	S	Se	0	0	0
			1365	842	258	250	7	8			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	18	MSE	MET	MODIFIED RESIDUE	UNP Q37875
A	23	MSE	MET	MODIFIED RESIDUE	UNP Q37875
A	50	MSE	MET	MODIFIED RESIDUE	UNP Q37875
A	101	MSE	MET	MODIFIED RESIDUE	UNP Q37875
A	109	MSE	MET	MODIFIED RESIDUE	UNP Q37875
A	116	MSE	MET	MODIFIED RESIDUE	UNP Q37875
A	131	MSE	MET	MODIFIED RESIDUE	UNP Q37875
A	149	MSE	MET	MODIFIED RESIDUE	UNP Q37875
A	186	HIS	-	EXPRESSION TAG	UNP Q37875
A	187	HIS	-	EXPRESSION TAG	UNP Q37875
A	188	HIS	-	EXPRESSION TAG	UNP Q37875
A	189	HIS	-	EXPRESSION TAG	UNP Q37875
A	190	HIS	-	EXPRESSION TAG	UNP Q37875
A	191	HIS	-	EXPRESSION TAG	UNP Q37875

- Molecule 2 is CITRIC ACID (three-letter code: CIT) (formula: C<sub>6</sub>H<sub>8</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			13	6	7		
2	A	1	Total	C	O	0	0
			13	6	7		

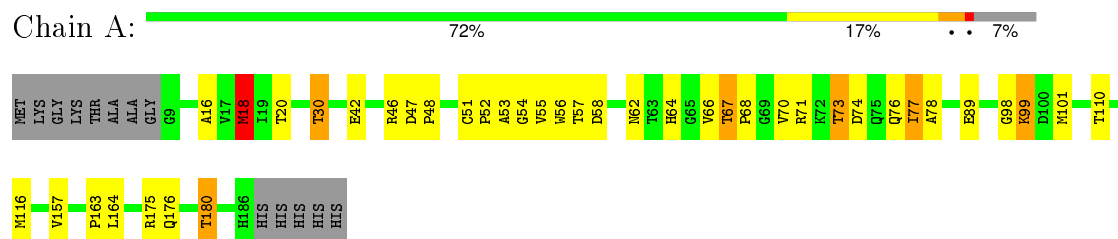
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	149	Total	O	0	0
			149	149		



Note EDS was not executed.

- Molecule 1: Lysozyme



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.87Å 66.87Å 166.80Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 1.75	Depositor
% Data completeness (in resolution range)	96.4 (50.00-1.75)	Depositor
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	REFMAC 5.2.0000	Depositor
R, $R_{free}$	0.208 , 0.238	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1540	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

## 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: CIT

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	0.60	0/1381	0.80	4/1850 (0.2%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	175	ARG	NE-CZ-NH1	-6.11	117.25	120.30
1	A	175	ARG	NE-CZ-NH2	6.08	123.34	120.30
1	A	47	ASP	CB-CG-OD2	6.05	123.75	118.30
1	A	18	MSE	CG-SE-CE	5.43	110.84	98.90

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	1365	0	1341	32	0
2	A	26	0	10	0	0
3	A	149	0	0	3	0
All	All	1540	0	1351	32	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 12.

All (32) 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:18:MSE:HE2	1:A:78:ALA:HB2	1.12	1.08
1:A:18:MSE:CE	1:A:78:ALA:HB2	1.84	1.07
1:A:30:THR:HG23	1:A:89:GLU:OE2	1.66	0.94
1:A:18:MSE:HE2	1:A:78:ALA:CB	1.99	0.92
1:A:73:THR:O	1:A:77:ILE:HG23	1.77	0.84
1:A:73:THR:HB	1:A:76:GLN:H	1.56	0.70
1:A:30:THR:CG2	1:A:89:GLU:OE2	2.41	0.68
1:A:30:THR:HG21	1:A:110:THR:HG21	1.75	0.67
1:A:18:MSE:HA	1:A:18:MSE:HE3	1.77	0.65
1:A:18:MSE:HE1	1:A:74:ASP:O	1.96	0.65
1:A:16:ALA:O	1:A:20:THR:HG23	2.01	0.60
1:A:176:GLN:O	1:A:180:THR:HG22	2.00	0.60
1:A:58:ASP:OD1	1:A:66:VAL:HG23	2.02	0.59
1:A:42:GLU:HG2	3:A:382:HOH:O	2.00	0.59
1:A:51:CYS:O	1:A:54:GLY:HA2	2.07	0.54
1:A:67:THR:HG23	1:A:70:VAL:HB	1.89	0.54
1:A:116:MSE:CB	3:A:305:HOH:O	2.57	0.52
1:A:52:PRO:O	1:A:53:ALA:HB3	2.10	0.51
1:A:18:MSE:SE	1:A:77:ILE:HD11	2.61	0.51
1:A:30:THR:CG2	1:A:110:THR:HG21	2.42	0.50
1:A:48:PRO:HA	1:A:57:THR:O	2.13	0.48
1:A:58:ASP:HB2	1:A:62:ASN:HD22	1.78	0.48
1:A:46:ARG:HD2	1:A:71:ARG:HH21	1.79	0.48
1:A:18:MSE:SE	1:A:77:ILE:CG1	3.12	0.47
1:A:56:TRP:CD1	1:A:68:PRO:HG3	2.50	0.47
1:A:98:GLY:HA2	1:A:101:MSE:HE3	1.99	0.45
1:A:56:TRP:NE1	1:A:68:PRO:HG3	2.34	0.43
1:A:67:THR:HG23	1:A:70:VAL:CB	2.49	0.42
1:A:157:VAL:O	1:A:163:PRO:HA	2.20	0.41
1:A:99:LYS:HG3	3:A:282:HOH:O	2.20	0.41
1:A:18:MSE:HA	1:A:18:MSE:CE	2.49	0.40
1:A:18:MSE:SE	1:A:77:ILE:HG13	2.72	0.40

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	176/191 (92%)	167 (95%)	9 (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	143/145 (99%)	133 (93%)	10 (7%)	19	3

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

Mol	Chain	Res	Type
1	A	18	MSE
1	A	30	THR
1	A	55	VAL
1	A	64	HIS
1	A	67	THR
1	A	73	THR
1	A	77	ILE
1	A	99	LYS
1	A	164	LEU
1	A	180	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	CIT	A	192	-	3,12,12	1.07	0	3,17,17	1.35	1 (33%)
2	CIT	A	250	-	3,12,12	1.08	0	3,17,17	3.47	3 (100%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	CIT	A	192	-	-	0/6/16/16	0/0/0/0
2	CIT	A	250	-	-	0/6/16/16	0/0/0/0

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	A	250	CIT	C3-C2-C1	-4.41	107.90	114.96
2	A	250	CIT	C4-C3-C2	-3.53	101.35	109.81
2	A	192	CIT	C3-C2-C1	-2.22	111.41	114.96
2	A	250	CIT	C3-C4-C5	2.01	118.17	114.96

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 ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.