



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 21, 2017 – 09:40 AM EST

PDB ID : 1LYS  
Title : X-RAY STRUCTURE OF A MONOCLINIC FORM OF HEN EGG-WHITE  
LYSOZYME CRYSTALLIZED AT 313K. COMPARISON OF TWO INDE-  
PENDENT MOLECULES  
Authors : Harata, K.  
Deposited on : unknown  
Resolution : 1.72 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
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030345
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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-20030345

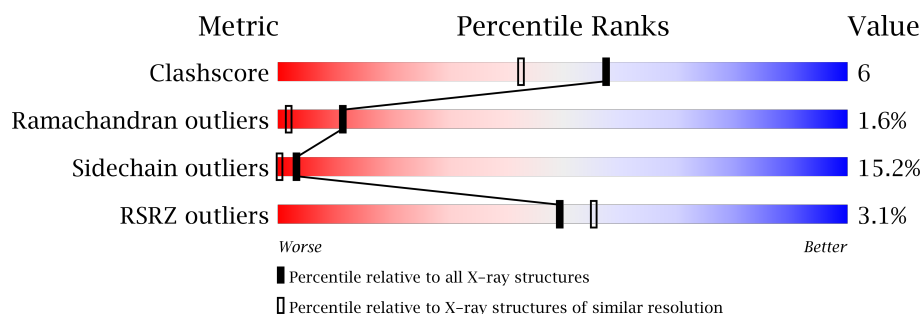
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.72 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	112137	4841 (1.74-1.70)
Ramachandran outliers	110173	4771 (1.74-1.70)
Sidechain outliers	110143	4771 (1.74-1.70)
RSRZ outliers	101464	4426 (1.74-1.70)

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	129	<div> <div>5%</div> <div> <div></div> <div>76%</div> <div>16%</div> <div>7%</div> <div>.</div> </div> </div>
1	B	129	<div> <div>2%</div> <div> <div></div> <div>69%</div> <div>26%</div> <div>5%</div> <div>.</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2217 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 HEN EGG WHITE LYSOZYME.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	129	Total	C	N	O	S	0	0	0
			1001	613	193	185	10			
1	B	129	Total	C	N	O	S	0	0	0
			1001	613	193	185	10			

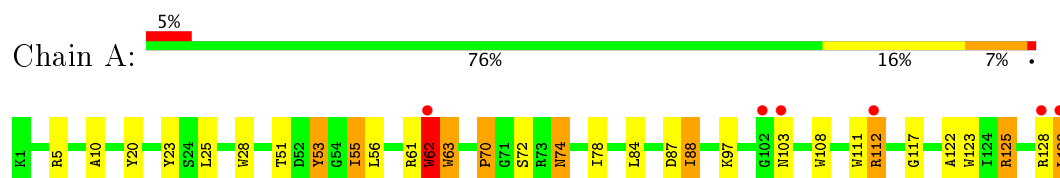
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	98	Total	O	0	0
			98	98		
2	B	117	Total	O	0	0
			117	117		

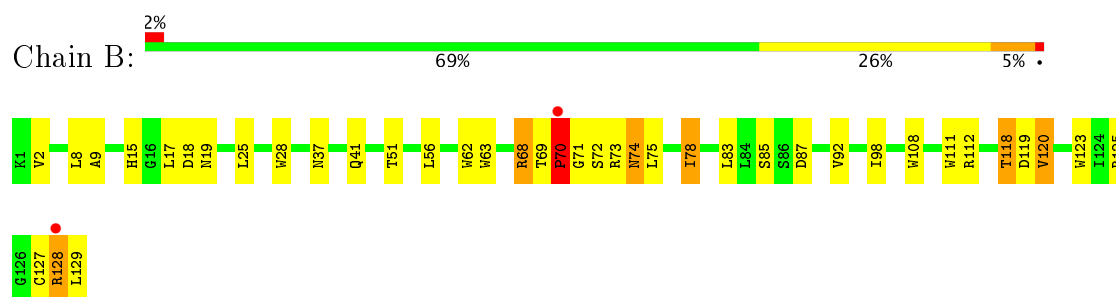
### 3 Residue-property plots [i](#)

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: HEN EGG WHITE LYSOZYME



- Molecule 1: HEN EGG WHITE LYSOZYME



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	27.23Å 63.66Å 59.12Å 90.00° 92.90° 90.00°	Depositor
Resolution (Å)	10.00 – 1.72 9.98 – 1.72	Depositor EDS
% Data completeness (in resolution range)	(Not available) (10.00-1.72) 85.3 (9.98-1.72)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.39 (at 1.71Å)	Xtriage
Refinement program	X-PLOR	Depositor
R, $R_{free}$	0.187 , (Not available) 0.182 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	11.5	Xtriage
Anisotropy	0.293	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 96.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.42$ , $\langle L^2 \rangle = 0.25$	Xtriage
Estimated twinning fraction	0.066 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	2217	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	17.0	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 75.59 % 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 1.2263e-06. 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

### 5.1 Standard geometry

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	0.85	0/1021	1.69	28/1379 (2.0%)
1	B	0.91	0/1021	1.74	31/1379 (2.2%)
All	All	0.88	0/2042	1.71	59/2758 (2.1%)

There are no bond length outliers.

All (59) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	125	ARG	NE-CZ-NH2	-9.25	115.67	120.30
1	B	68	ARG	NE-CZ-NH1	8.92	124.76	120.30
1	B	62	TRP	CD1-CG-CD2	8.83	113.36	106.30
1	A	111	TRP	CD1-CG-CD2	8.42	113.04	106.30
1	B	123	TRP	CD1-CG-CD2	8.34	112.97	106.30
1	A	108	TRP	CD1-CG-CD2	8.04	112.73	106.30
1	B	28	TRP	CD1-CG-CD2	7.93	112.64	106.30
1	A	63	TRP	CD1-CG-CD2	7.88	112.61	106.30
1	A	62	TRP	CG-CD2-CE3	7.70	140.83	133.90
1	A	63	TRP	CE2-CD2-CG	-7.60	101.22	107.30
1	A	123	TRP	CD1-CG-CD2	7.58	112.37	106.30
1	B	63	TRP	CD1-CG-CD2	7.52	112.31	106.30
1	B	123	TRP	CE2-CD2-CG	-7.44	101.35	107.30
1	B	108	TRP	CD1-CG-CD2	7.43	112.24	106.30
1	B	71	GLY	N-CA-C	-7.29	94.88	113.10
1	A	63	TRP	CG-CD2-CE3	7.27	140.44	133.90
1	B	111	TRP	CD1-CG-CD2	7.19	112.05	106.30
1	A	28	TRP	CD1-CG-CD2	7.16	112.03	106.30
1	A	123	TRP	CE2-CD2-CG	-7.16	101.58	107.30
1	A	62	TRP	CE2-CD2-CG	-7.03	101.68	107.30
1	A	111	TRP	CE2-CD2-CG	-6.94	101.75	107.30
1	B	28	TRP	CE2-CD2-CG	-6.89	101.79	107.30
1	B	62	TRP	CG-CD1-NE1	-6.84	103.26	110.10
1	B	111	TRP	CE2-CD2-CG	-6.80	101.86	107.30
1	A	62	TRP	CB-CG-CD1	-6.68	118.32	127.00

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	28	TRP	CE2-CD2-CG	-6.37	102.20	107.30
1	B	112	ARG	NE-CZ-NH2	-6.18	117.21	120.30
1	B	108	TRP	CE2-CD2-CG	-6.13	102.39	107.30
1	A	108	TRP	CE2-CD2-CG	-6.12	102.40	107.30
1	A	62	TRP	CD1-CG-CD2	6.11	111.19	106.30
1	B	62	TRP	CE2-CD2-CG	-6.08	102.43	107.30
1	A	23	TYR	CB-CG-CD2	-5.98	117.41	121.00
1	B	63	TRP	CE2-CD2-CG	-5.97	102.52	107.30
1	A	63	TRP	CB-CG-CD1	-5.96	119.25	127.00
1	B	2	VAL	N-CA-C	-5.89	95.10	111.00
1	B	108	TRP	CG-CD1-NE1	-5.86	104.24	110.10
1	B	70	PRO	CA-N-CD	-5.78	103.40	111.50
1	A	108	TRP	CG-CD1-NE1	-5.77	104.33	110.10
1	A	55	ILE	CB-CA-C	-5.70	100.19	111.60
1	B	120	VAL	CA-CB-CG2	-5.67	102.40	110.90
1	A	108	TRP	CB-CG-CD1	-5.65	119.66	127.00
1	A	20	TYR	CB-CG-CD2	-5.61	117.63	121.00
1	A	111	TRP	CG-CD1-NE1	-5.55	104.55	110.10
1	B	118	THR	N-CA-CB	-5.47	99.91	110.30
1	B	120	VAL	N-CA-CB	-5.41	99.61	111.50
1	B	68	ARG	NE-CZ-NH2	-5.39	117.61	120.30
1	B	111	TRP	CG-CD2-CE3	5.38	138.74	133.90
1	A	53	TYR	CB-CG-CD2	-5.35	117.79	121.00
1	A	63	TRP	CG-CD1-NE1	-5.30	104.80	110.10
1	B	123	TRP	CG-CD1-NE1	-5.30	104.80	110.10
1	B	125	ARG	NE-CZ-NH1	5.25	122.93	120.30
1	A	88	ILE	N-CA-CB	-5.25	98.74	110.80
1	B	127	CYS	O-C-N	5.14	130.92	122.70
1	B	28	TRP	CG-CD1-NE1	-5.12	104.98	110.10
1	A	112	ARG	NE-CZ-NH1	5.09	122.84	120.30
1	B	111	TRP	CG-CD1-NE1	-5.01	105.09	110.10
1	A	123	TRP	CB-CG-CD1	-5.01	120.49	127.00
1	A	108	TRP	CG-CD2-CE3	5.00	138.41	133.90
1	B	41	GLN	CA-CB-CG	-5.00	102.39	113.40

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts ⓘ

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	1001	0	959	14	0
1	B	1001	0	959	10	0
2	A	98	0	0	1	0
2	B	117	0	0	1	0
All	All	2217	0	1918	22	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.

All (22) 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:69:THR:HG22	1:B:70:PRO:HD2	1.49	0.92
1:B:74:ASN:HD21	1:B:78:ILE:H	1.25	0.84
1:B:69:THR:CG2	1:B:70:PRO:HD2	2.15	0.74
1:B:70:PRO:HG2	1:B:72:SER:HB3	1.74	0.69
1:A:72:SER:HB2	2:A:216:HOH:O	1.92	0.69
1:A:117:GLY:HA3	1:B:78:ILE:HD11	1.82	0.61
1:A:10:ALA:HA	1:A:129:LEU:HD11	1.84	0.59
1:A:74:ASN:HD21	1:A:78:ILE:H	1.52	0.58
1:B:15:HIS:HB3	1:B:92:VAL:HG11	1.86	0.57
1:A:62:TRP:HE3	1:A:63:TRP:HE1	1.60	0.48
1:A:62:TRP:HE3	1:A:63:TRP:NE1	2.14	0.46
1:A:122:ALA:HA	1:A:125:ARG:HH11	1.81	0.46
1:B:128:ARG:HB3	1:B:129:LEU:H	1.51	0.43
1:B:9:ALA:HB1	1:B:129:LEU:HD11	2.01	0.43
1:A:51:THR:HB	1:A:53:TYR:CE1	2.54	0.42
1:A:5:ARG:NH2	1:A:125:ARG:O	2.52	0.41
1:B:129:LEU:HB2	2:B:243:HOH:O	2.20	0.41
1:A:97:LYS:NZ	1:A:97:LYS:HB2	2.36	0.41
1:A:117:GLY:CA	1:B:78:ILE:HD11	2.48	0.40
1:A:61:ARG:HH22	1:A:70:PRO:HB2	1.86	0.40
1:A:74:ASN:HD21	1:A:78:ILE:N	2.16	0.40
1:A:74:ASN:ND2	1:A:78:ILE:H	2.16	0.40

There are no symmetry-related clashes.



## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	127/129 (98%)	122 (96%)	4 (3%)	1 (1%)	22	7
1	B	127/129 (98%)	120 (94%)	4 (3%)	3 (2%)	7	1
All	All	254/258 (98%)	242 (95%)	8 (3%)	4 (2%)	11	2

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	70	PRO
1	B	128	ARG
1	A	103	ASN
1	B	18	ASP

### 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	105/105 (100%)	92 (88%)	13 (12%)	5	1
1	B	105/105 (100%)	86 (82%)	19 (18%)	2	0
All	All	210/210 (100%)	178 (85%)	32 (15%)	3	0

All (32) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	25	LEU
1	A	55	ILE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	56	LEU
1	A	62	TRP
1	A	70	PRO
1	A	74	ASN
1	A	84	LEU
1	A	87	ASP
1	A	88	ILE
1	A	112	ARG
1	A	125	ARG
1	A	128	ARG
1	A	129	LEU
1	B	8	LEU
1	B	17	LEU
1	B	19	ASN
1	B	25	LEU
1	B	37	ASN
1	B	51	THR
1	B	56	LEU
1	B	68	ARG
1	B	73	ARG
1	B	74	ASN
1	B	75	LEU
1	B	78	ILE
1	B	83	LEU
1	B	85	SER
1	B	87	ASP
1	B	98	ILE
1	B	118	THR
1	B	119	ASP
1	B	120	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	44	ASN
1	A	57	GLN
1	A	59	ASN
1	A	74	ASN
1	A	103	ASN
1	B	37	ASN
1	B	74	ASN
1	B	103	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	129/129 (100%)	-0.05	6 (4%) 32 37	6, 15, 33, 47	0
1	B	129/129 (100%)	-0.35	2 (1%) 72 77	4, 11, 26, 46	0
All	All	258/258 (100%)	-0.20	8 (3%) 49 55	4, 13, 33, 47	0

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	102	GLY	9.7
1	B	70	PRO	4.2
1	A	62	TRP	4.2
1	A	112	ARG	2.4
1	A	128	ARG	2.2
1	B	128	ARG	2.2
1	A	103	ASN	2.2
1	A	129	LEU	2.0

### 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.