



# Full wwPDB X-ray Structure Validation Report ⓘ

May 13, 2020 – 01:31 pm BST

PDB ID : 1VDP  
Title : The crystal structure of the monoclinic form of hen egg white lysozyme at 1.7 angstroms resolution in space  
Authors : Aibara, S.; Suzuki, A.; Kidera, A.; Shibata, K.; Yamane, T.; DeLucas, L.J.; Hirose, M.  
Deposited on : 2004-03-24  
Resolution : 1.70 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

<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

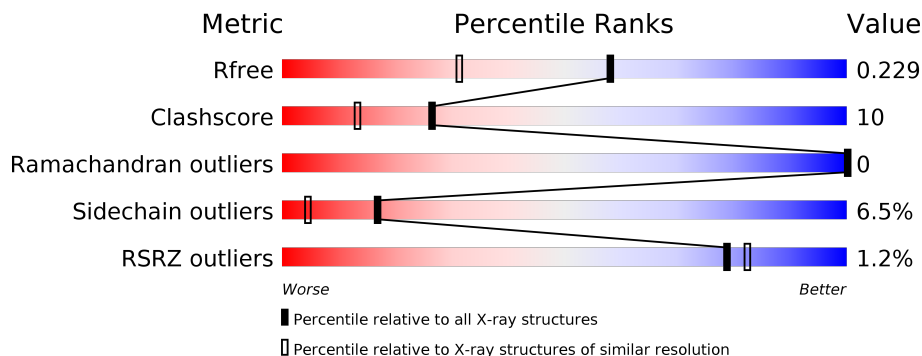
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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 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	129	 2% 76% 17% 5%
1	B	129	 77% 19%

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2673 atoms, of which 491 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 C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	129	1211	599	241	177	184	10	0	0	0
1	B	129	1232	605	250	183	184	10	0	0	0

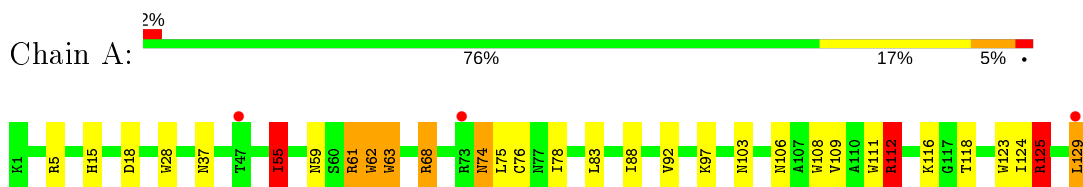
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	108	Total	O	0	0
			108	108		
2	B	122	Total	O	0	0
			122	122		

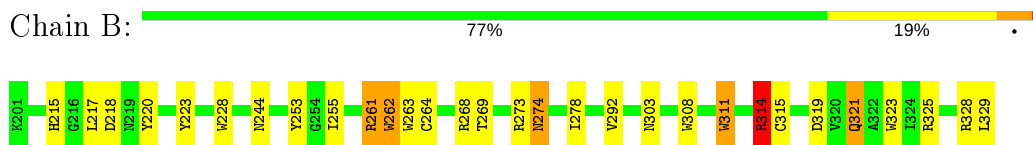
### 3 Residue-property plots [i](#)

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: Lysozyme C



- Molecule 1: Lysozyme C



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	28.05Å 62.88Å 60.65Å 90.00° 90.90° 90.00°	Depositor
Resolution (Å)	9.98 – 1.70 9.98 – 1.70	Depositor EDS
% Data completeness (in resolution range)	79.7 (9.98-1.70) 89.8 (9.98-1.70)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.62 (at 1.70Å)	Xtriage
Refinement program	CNS 1.0, X-PLOR 3.1	Depositor
R, $R_{free}$	0.175 , 0.221 0.189 , 0.229	Depositor DCC
$R_{free}$ test set	1003 reflections (4.83%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	16.2	Xtriage
Anisotropy	0.413	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 58.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.013 for -h,-l,-k 0.000 for -h,l,k 0.042 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	2673	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 34.30 % 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 6.9764e-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	0.92	0/990	1.75	32/1341 (2.4%)
1	B	0.92	0/1002	1.69	25/1356 (1.8%)
All	All	0.92	0/1992	1.72	57/2697 (2.1%)

There are no bond length outliers.

All (57) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	125	ARG	NE-CZ-NH1	11.40	126.00	120.30
1	B	323	TRP	CD1-CG-CD2	9.95	114.26	106.30
1	A	123	TRP	CD1-CG-CD2	9.66	114.03	106.30
1	A	111	TRP	CD1-CG-CD2	9.49	113.89	106.30
1	B	323	TRP	CE2-CD2-CG	-9.27	99.88	107.30
1	A	62	TRP	CD1-CG-CD2	8.49	113.09	106.30
1	A	111	TRP	CE2-CD2-CG	-8.36	100.61	107.30
1	B	261	ARG	NE-CZ-NH1	8.36	124.48	120.30
1	A	63	TRP	CD1-CG-CD2	8.20	112.86	106.30
1	A	108	TRP	CD1-CG-CD2	8.16	112.83	106.30
1	B	263	TRP	CD1-CG-CD2	8.09	112.77	106.30
1	A	112	ARG	CB-CG-CD	7.81	131.90	111.60
1	A	62	TRP	CE2-CD2-CG	-7.60	101.22	107.30
1	A	123	TRP	CE2-CD2-CG	-7.56	101.25	107.30
1	B	262	TRP	CE2-CD2-CG	-7.43	101.36	107.30
1	B	261	ARG	NE-CZ-NH2	-7.43	116.59	120.30
1	B	311	TRP	CD1-CG-CD2	7.33	112.16	106.30
1	B	262	TRP	CD1-CG-CD2	7.29	112.13	106.30
1	B	308	TRP	CD1-CG-CD2	7.28	112.12	106.30
1	B	263	TRP	CE2-CD2-CG	-7.22	101.52	107.30
1	A	63	TRP	CE2-CD2-CG	-7.22	101.53	107.30
1	A	108	TRP	CE2-CD2-CG	-7.20	101.54	107.30
1	A	28	TRP	CD1-CG-CD2	7.09	111.97	106.30
1	B	228	TRP	CD1-CG-CD2	7.04	111.93	106.30
1	B	308	TRP	CE2-CD2-CG	-6.95	101.74	107.30

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	223	TYR	CB-CG-CD1	-6.94	116.84	121.00
1	B	228	TRP	CE2-CD2-CG	-6.66	101.97	107.30
1	A	68	ARG	NE-CZ-NH1	6.54	123.57	120.30
1	A	61	ARG	NE-CZ-NH1	-6.53	117.03	120.30
1	A	112	ARG	NE-CZ-NH1	6.42	123.51	120.30
1	B	220	TYR	CB-CG-CD1	-6.26	117.24	121.00
1	A	123	TRP	CG-CD1-NE1	-6.24	103.86	110.10
1	B	328	ARG	NE-CZ-NH1	6.23	123.42	120.30
1	A	5	ARG	NE-CZ-NH1	-6.21	117.19	120.30
1	A	5	ARG	NE-CZ-NH2	6.14	123.37	120.30
1	B	218	ASP	CB-CG-OD1	6.09	123.78	118.30
1	B	314	ARG	NE-CZ-NH1	6.03	123.31	120.30
1	A	28	TRP	CE2-CD2-CG	-5.97	102.53	107.30
1	B	253	TYR	CB-CG-CD2	-5.93	117.44	121.00
1	A	55	ILE	CB-CA-C	-5.91	99.78	111.60
1	A	111	TRP	CG-CD1-NE1	-5.91	104.19	110.10
1	B	311	TRP	CE2-CD2-CG	-5.86	102.61	107.30
1	A	111	TRP	CG-CD2-CE3	5.78	139.10	133.90
1	A	109	VAL	CG1-CB-CG2	-5.76	101.68	110.90
1	A	61	ARG	NE-CZ-NH2	5.74	123.17	120.30
1	A	18	ASP	CB-CG-OD1	5.73	123.45	118.30
1	A	63	TRP	CG-CD2-CE3	5.55	138.90	133.90
1	B	319	ASP	CB-CG-OD2	-5.55	113.31	118.30
1	A	63	TRP	CB-CG-CD1	-5.47	119.88	127.00
1	A	129	LEU	CA-CB-CG	5.45	127.84	115.30
1	A	63	TRP	CG-CD1-NE1	-5.39	104.70	110.10
1	B	323	TRP	CG-CD1-NE1	-5.38	104.72	110.10
1	B	263	TRP	CG-CD1-NE1	-5.14	104.96	110.10
1	A	108	TRP	CG-CD1-NE1	-5.10	105.00	110.10
1	B	325	ARG	NE-CZ-NH1	5.08	122.84	120.30
1	B	321	GLN	CA-CB-CG	-5.03	102.34	113.40
1	A	28	TRP	CG-CD1-NE1	-5.01	105.09	110.10

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	970	241	905	25	0
1	B	982	250	923	14	0
2	A	108	0	0	12	1
2	B	122	0	0	8	1
All	All	2182	491	1828	39	1

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 10.

All (39) 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:268:ARG:CG	2:B:657:HOH:O	1.69	1.27
1:A:112:ARG:CG	2:A:532:HOH:O	2.20	0.88
1:A:97:LYS:HD2	2:A:591:HOH:O	1.78	0.83
1:B:314:ARG:NH2	2:B:578:HOH:O	2.14	0.80
1:A:112:ARG:HG3	2:A:532:HOH:O	1.80	0.79
1:B:215:HIS:HB3	1:B:292:VAL:HG11	1.72	0.72
1:B:274:ASN:HD21	1:B:278:ILE:H	1.49	0.60
1:A:15:HIS:HB3	1:A:92:VAL:HG11	1.86	0.57
1:A:112:ARG:HH21	1:A:116:LYS:HZ3	1.53	0.56
1:A:118:THR:HG22	2:B:466:HOH:O	2.07	0.55
1:A:112:ARG:CB	2:A:532:HOH:O	2.52	0.55
1:A:59:ASN:ND2	1:A:61:ARG:HB3	2.22	0.55
1:B:255:ILE:CG2	2:B:515:HOH:O	2.55	0.54
1:A:112:ARG:HD2	2:A:503:HOH:O	2.07	0.53
1:B:244:ASN:ND2	2:B:560:HOH:O	2.43	0.52
1:A:59:ASN:HD21	1:A:61:ARG:HB3	1.75	0.52
1:A:125:ARG:NH1	2:A:479:HOH:O	2.42	0.51
1:A:112:ARG:HE	1:A:116:LYS:HZ2	1.58	0.51
1:A:74:ASN:HD21	1:A:78:ILE:H	1.59	0.51
1:A:62:TRP:NE1	2:A:499:HOH:O	2.44	0.51
1:B:321:GLN:CG	2:B:611:HOH:O	2.59	0.49
1:A:125:ARG:CG	1:A:125:ARG:HH11	2.26	0.49
1:A:125:ARG:HG2	1:A:125:ARG:HH11	1.80	0.46
1:A:62:TRP:CD1	2:A:499:HOH:O	2.56	0.46
1:A:55:ILE:H	1:A:55:ILE:HG13	1.63	0.46
1:A:112:ARG:HD2	2:A:532:HOH:O	2.17	0.45
1:B:262:TRP:CZ3	1:B:273:ARG:HG3	2.51	0.45
1:B:215:HIS:CD2	2:B:522:HOH:O	2.69	0.45
1:B:261:ARG:HA	1:B:269:THR:HG21	1.99	0.45
1:A:112:ARG:CD	2:A:532:HOH:O	2.60	0.44

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:103:ASN:O	1:A:106:ASN:HB2	2.18	0.44
1:A:78:ILE:CD1	1:A:83:LEU:HD21	2.48	0.43
1:B:311:TRP:CD1	1:B:315:CYS:HB2	2.55	0.42
1:A:88:ILE:HA	2:A:410:HOH:O	2.19	0.41
1:B:264:CYS:HA	1:B:274:ASN:ND2	2.36	0.41
1:A:112:ARG:CD	2:A:503:HOH:O	2.68	0.41
1:B:314:ARG:HD2	2:B:586:HOH:O	2.21	0.41
1:A:63:TRP:O	1:A:76:CYS:HB2	2.21	0.41
1:B:217:LEU:HA	1:B:217:LEU:HD12	1.95	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:579:HOH:O	2:B:484:HOH:O[2_756]	2.15	0.05

## 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	127/129 (98%)	125 (98%)	2 (2%)	0	100	100
1	B	127/129 (98%)	126 (99%)	1 (1%)	0	100	100
All	All	254/258 (98%)	251 (99%)	3 (1%)	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	99/105 (94%)	90 (91%)	9 (9%)	9	2
1	B	101/105 (96%)	97 (96%)	4 (4%)	31	13
All	All	200/210 (95%)	187 (94%)	13 (6%)	17	4

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	37	ASN
1	A	55	ILE
1	A	68	ARG
1	A	74	ASN
1	A	75	LEU
1	A	112	ARG
1	A	124	ILE
1	A	125	ARG
1	A	129	LEU
1	B	274	ASN
1	B	303	ASN
1	B	314	ARG
1	B	329	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	37	ASN
1	A	44	ASN
1	A	59	ASN
1	A	74	ASN
1	A	93	ASN
1	A	106	ASN
1	A	113	ASN
1	B	239	ASN
1	B	274	ASN
1	B	293	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%)	-0.00	3 (2%) 60 65	10, 17, 35, 48	0
1	B	129/129 (100%)	-0.11	0 100 100	9, 16, 35, 42	0
All	All	258/258 (100%)	-0.05	3 (1%) 79 82	9, 17, 35, 48	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	129	LEU	4.8
1	A	47	THR	2.4
1	A	73	ARG	2.2

### 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.