



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 17, 2021 – 03:05 AM EDT

PDB ID : 1JTN  
Title : Alternative Structures of a Sequence Extended T4 Lysozyme Show that the Highly Conserved Beta-Sheet Region has weak intrinsic Folding Propensity  
Authors : Sagermann, M.; Matthews, B.W.  
Deposited on : 2001-08-21  
Resolution : 2.30 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
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
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)	:	2.23.2

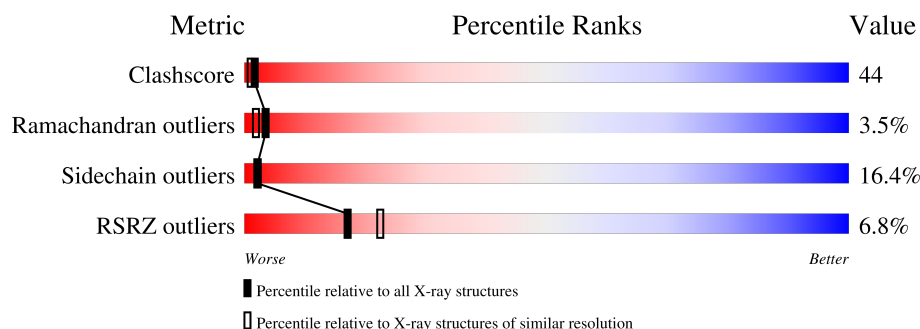
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	178	
1	B	178	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	SO4	B	502	-	-	X	-

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2951 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	177	Total	C	N	O	S	0	0	0
			1412	893	253	261	5			
1	B	174	Total	C	N	O	S	0	0	0
			1389	877	250	257	5			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	54	THR	CYS	engineered mutation	UNP P00720
A	97	ALA	CYS	engineered mutation	UNP P00720
A	165	GLU	-	insertion	UNP P00720
A	166	THR	-	insertion	UNP P00720
A	167	TYR	-	insertion	UNP P00720
A	168	TYR	-	insertion	UNP P00720
A	169	THR	-	insertion	UNP P00720
A	170	ILE	-	insertion	UNP P00720
A	171	GLY	-	insertion	UNP P00720
A	172	ILE	-	insertion	UNP P00720
A	173	GLY	-	insertion	UNP P00720
A	174	HIS	-	insertion	UNP P00720
A	175	LEU	-	insertion	UNP P00720
A	176	LEU	-	insertion	UNP P00720
A	177	THR	-	insertion	UNP P00720
A	178	LYS	-	insertion	UNP P00720
B	54	THR	CYS	engineered mutation	UNP P00720
B	97	ALA	CYS	engineered mutation	UNP P00720
B	165	GLU	-	insertion	UNP P00720
B	166	THR	-	insertion	UNP P00720
B	167	TYR	-	insertion	UNP P00720
B	168	TYR	-	insertion	UNP P00720
B	169	THR	-	insertion	UNP P00720
B	170	ILE	-	insertion	UNP P00720
B	171	GLY	-	insertion	UNP P00720

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	172	ILE	-	insertion	UNP P00720
B	173	GLY	-	insertion	UNP P00720
B	174	HIS	-	insertion	UNP P00720
B	175	LEU	-	insertion	UNP P00720
B	176	LEU	-	insertion	UNP P00720
B	177	THR	-	insertion	UNP P00720
B	178	LYS	-	insertion	UNP P00720

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		

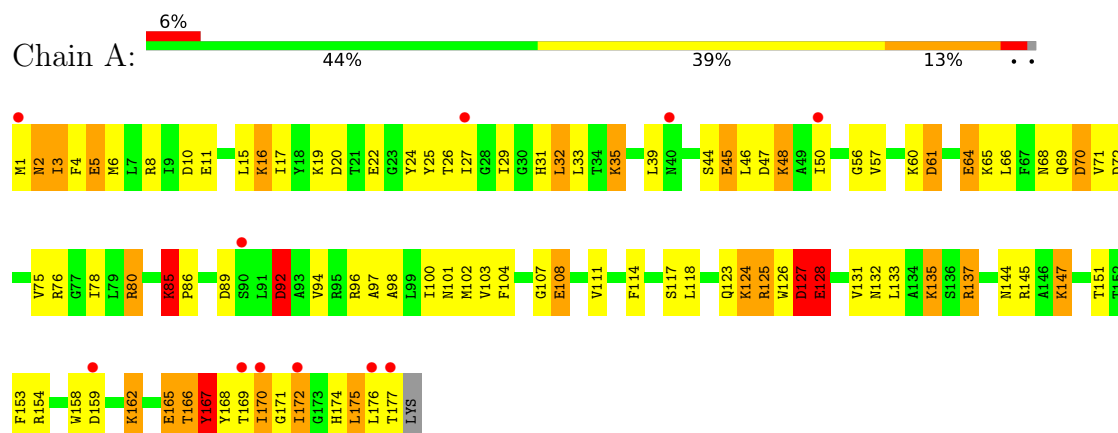
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	55	Total	O	0	0
			55	55		
3	B	85	Total	O	0	0
			85	85		

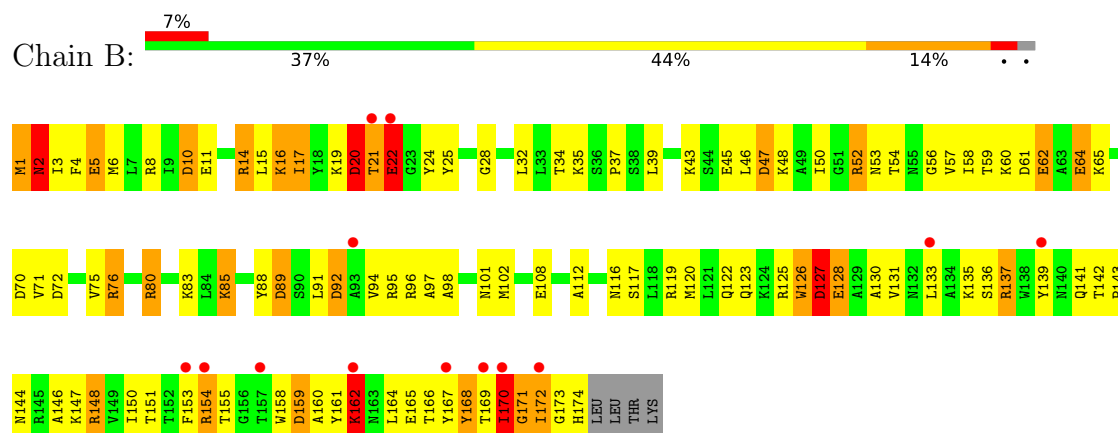
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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



#### • Molecule 1: LYSOZYME



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	60.12Å 32.31Å 85.92Å 90.00° 102.64° 90.00°	Depositor
Resolution (Å)	19.60 – 2.30 19.56 – 2.30	Depositor EDS
% Data completeness (in resolution range)	95.0 (19.60-2.30) 98.3 (19.56-2.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	4.40	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	9.86 (at 2.30Å)	Xtriage
Refinement program	TNT	Depositor
R, $R_{free}$	0.220 , 0.314 0.232 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	19.8	Xtriage
Anisotropy	0.299	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 75.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	2951	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	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 42.73 % 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 1.9541e-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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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	0.74	6/1435 (0.4%)	1.03	20/1936 (1.0%)
1	B	0.74	9/1412 (0.6%)	1.00	18/1904 (0.9%)
All	All	0.74	15/2847 (0.5%)	1.01	38/3840 (1.0%)

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

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	62	GLU	CD-OE2	5.79	1.32	1.25
1	A	108	GLU	CD-OE2	5.71	1.31	1.25
1	B	64	GLU	CD-OE2	5.69	1.31	1.25
1	A	128	GLU	CD-OE2	5.56	1.31	1.25
1	A	11	GLU	CD-OE2	5.54	1.31	1.25

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	127	ASP	CB-CG-OD2	-6.64	112.33	118.30
1	B	61	ASP	CB-CG-OD2	-6.62	112.34	118.30
1	A	70	ASP	CB-CG-OD2	-6.36	112.57	118.30
1	B	89	ASP	CB-CG-OD2	-6.36	112.57	118.30
1	A	72	ASP	CB-CG-OD2	-6.33	112.60	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	167	TYR	Sidechain

## 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	1412	0	1438	129	1
1	B	1389	0	1409	133	0
2	A	5	0	0	1	0
2	B	5	0	0	2	0
3	A	55	0	0	13	1
3	B	85	0	0	16	0
All	All	2951	0	2847	249	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 44.

The worst 5 of 249 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:47:ASP:OD1	1:B:54:THR:OG1	1.66	1.12
1:A:170:ILE:HB	1:B:4:PHE:HB3	1.40	1.03
1:B:116:ASN:HD22	1:B:119:ARG:HH12	1.04	0.94
1:A:174:HIS:HD2	1:B:8:ARG:HE	1.18	0.91
1:A:1:MET:HE3	1:A:5:GLU:HB2	1.56	0.87

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)
1:A:22:GLU:OE2	3:A:502:HOH:O[2_756]	2.18	0.02



## 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	175/178 (98%)	162 (93%)	10 (6%)	3 (2%)	9	8
1	B	172/178 (97%)	150 (87%)	13 (8%)	9 (5%)	2	1
All	All	347/356 (98%)	312 (90%)	23 (7%)	12 (4%)	3	2

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	ASN
1	B	2	ASN
1	B	20	ASP
1	B	170	ILE
1	B	171	GLY

### 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	148/149 (99%)	123 (83%)	25 (17%)	2	2
1	B	145/149 (97%)	122 (84%)	23 (16%)	2	2
All	All	293/298 (98%)	245 (84%)	48 (16%)	2	2

5 of 48 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	14	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	53	ASN
1	B	16	LYS
1	B	22	GLU
1	B	76	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	2	ASN
1	B	69	GLN
1	B	116	ASN
1	A	123	GLN
1	A	116	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

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	SO4	A	501	-	4,4,4	0.27	0	6,6,6	0.05	0
2	SO4	B	502	-	4,4,4	0.18	0	6,6,6	0.07	0

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.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501	SO4	1	0
2	B	502	SO4	2	0

## 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	177/178 (99%)	0.48	11 (6%) 20 26	10, 25, 48, 76	0
1	B	174/178 (97%)	0.62	13 (7%) 14 19	12, 26, 59, 100	0
All	All	351/356 (98%)	0.55	24 (6%) 17 22	10, 26, 55, 100	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	172	ILE	7.6
1	A	177	THR	7.5
1	B	21	THR	7.5
1	A	172	ILE	5.2
1	B	22	GLU	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 monosaccharides in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	SO4	B	502	5/5	0.94	0.25	11,15,37,91	5
2	SO4	A	501	5/5	0.95	0.27	4,11,45,63	5

## 6.5 Other polymers [i](#)

There are no such residues in this entry.