



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 01:12 PM UTC

PDB ID : 5W0V / pdb_00005w0v
Title : Crystal structure of full-length Kluyveromyces lactis Kap123 with histone H4 1-34
Authors : An, S.; Yoon, J.; Song, J.-J.; Cho, U.-S.
Deposited on : 2017-05-31
Resolution : 2.82 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

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
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

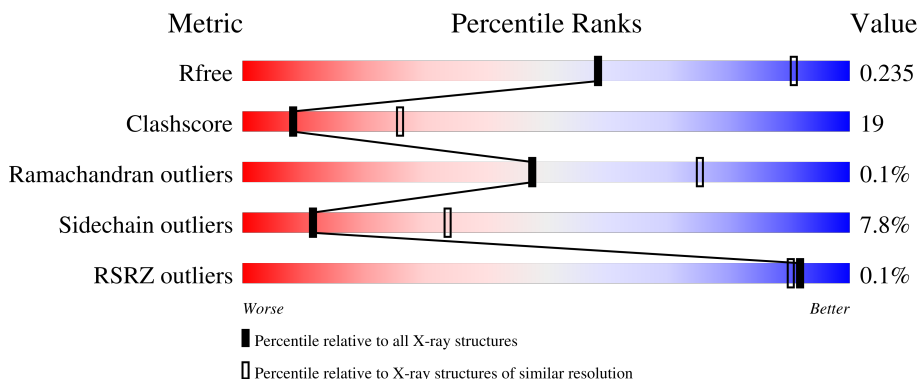
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.82 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	4591 (2.84-2.80)
Clashscore	190562	5010 (2.84-2.80)
Ramachandran outliers	187476	4916 (2.84-2.80)
Sidechain outliers	187428	4918 (2.84-2.80)
RSRZ outliers	180081	4594 (2.84-2.80)

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 ≥ 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	1116	
1	B	1116	
2	C	34	
2	D	34	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 15787 atoms, of which 38 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 Kap123.

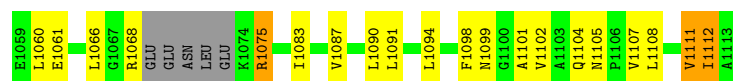
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	1017	Total	C	H	N	O	S	0	0	0
			7910	5020	38	1288	1543	21			
1	B	1003	Total	C	N	O	S		0	0	0
			7773	4957	1266	1529	21				

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP Q6CMF0
A	-1	ASN	-	expression tag	UNP Q6CMF0
A	0	ALA	-	expression tag	UNP Q6CMF0
B	-2	SER	-	expression tag	UNP Q6CMF0
B	-1	ASN	-	expression tag	UNP Q6CMF0
B	0	ALA	-	expression tag	UNP Q6CMF0

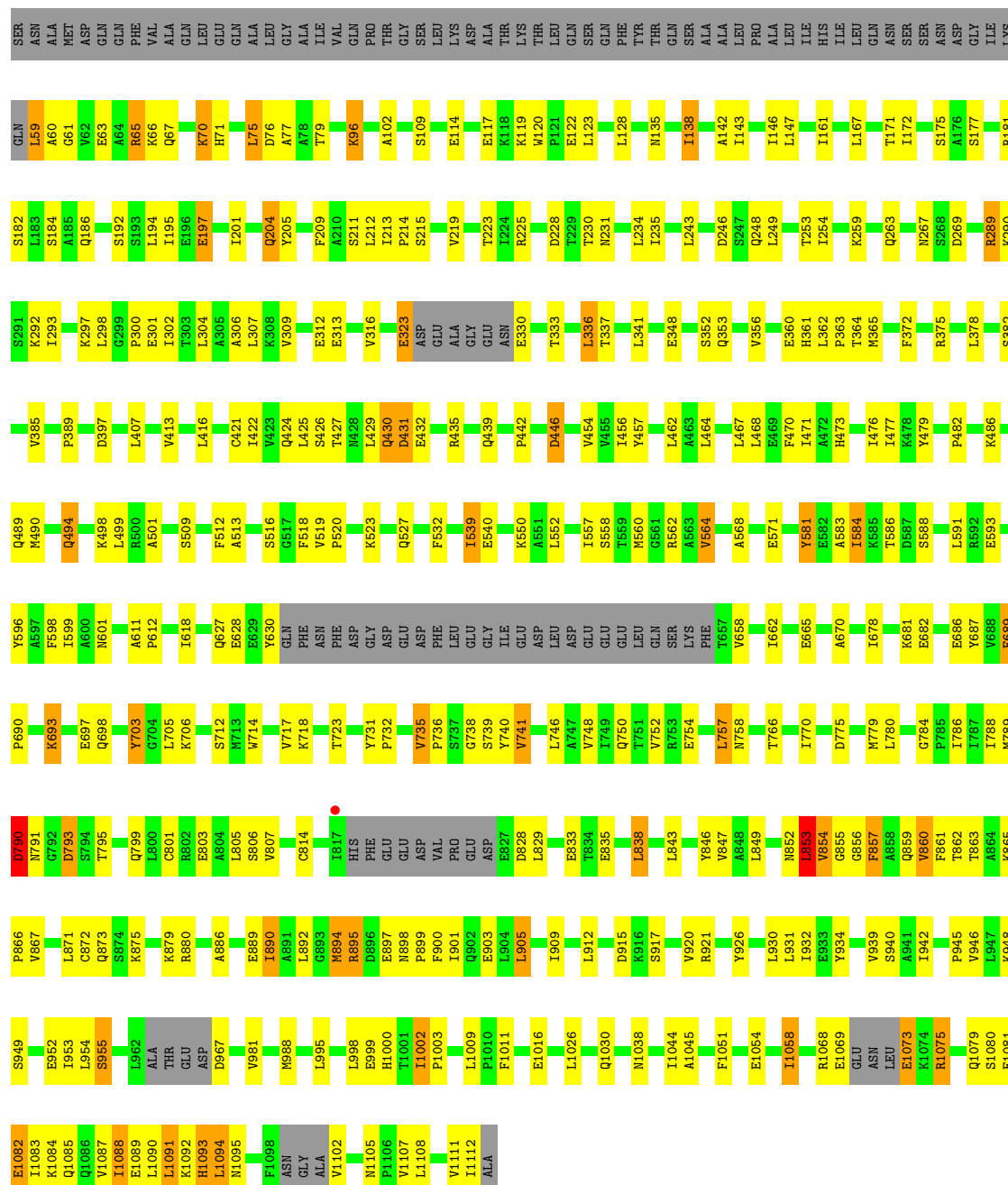
- Molecule 2 is a protein called Histone H4 1-34.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	7	Total	C	N	O	0	0	0
			54	31	16	7			
2	D	6	Total	C	N	O	0	0	0
			50	29	15	6			



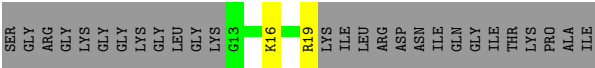
• Molecule 1: Kap123

Chain B: 58% 27% 10%



• Molecule 2: Histone H4 1-34

Chain C: 15% 6% 79%



● Molecule 2: Histone H4 1-34



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	78.79Å 87.70Å 101.62Å 79.59° 81.45° 71.70°	Depositor
Resolution (Å)	33.79 – 2.82 33.79 – 2.82	Depositor EDS
% Data completeness (in resolution range)	98.3 (33.79-2.82) 98.3 (33.79-2.82)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.79 (at 2.81Å)	Xtriage
Refinement program	PHENIX 1.11.1 _2575	Depositor
R, R_{free}	0.170 , 0.234 0.173 , 0.235	Depositor DCC
R_{free} test set	2000 reflections (3.29%)	wwPDB-VP
Wilson B-factor (Å ²)	75.8	Xtriage
Anisotropy	0.512	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 54.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	15787	wwPDB-VP
Average B, all atoms (Å ²)	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.38% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.46	0/7990	0.63	6/10843 (0.1%)
1	B	0.53	8/7889 (0.1%)	0.67	11/10705 (0.1%)
2	C	0.41	0/54	0.96	0/68
2	D	0.44	0/50	0.93	0/63
All	All	0.49	8/15983 (0.1%)	0.65	17/21679 (0.1%)

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	426	SER	C-O	-6.12	1.16	1.24
1	B	432	GLU	C-O	-5.74	1.17	1.24
1	B	427	THR	C-O	-5.68	1.17	1.24
1	B	892	LEU	C-O	-5.53	1.17	1.24
1	B	895	ARG	C-O	-5.39	1.19	1.24

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	430	GLN	N-CA-C	12.35	124.28	111.07
1	A	430	GLN	N-CA-C	11.43	123.43	110.97
1	B	581	TYR	CA-CB-CG	8.87	129.87	113.90
1	B	854	VAL	N-CA-CB	8.87	125.86	111.23
1	B	793	ASP	N-CA-C	7.71	122.09	110.14

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	7872	38	7970	286	0
1	B	7773	0	7854	322	0
2	C	54	0	56	9	0
2	D	50	0	53	3	0
All	All	15749	38	15933	607	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 19.

The worst 5 of 607 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:736:PRO:HD2	1:B:789:MET:CE	1.56	1.33
1:B:739:SER:N	1:B:791:ASN:OD1	1.61	1.32
1:B:738:GLY:HA2	1:B:791:ASN:OD1	1.29	1.24
1:B:738:GLY:HA2	1:B:791:ASN:CG	1.64	1.23
1:B:738:GLY:CA	1:B:791:ASN:OD1	1.86	1.21

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	1003/1116 (90%)	983 (98%)	19 (2%)	1 (0%)	48	75
1	B	989/1116 (89%)	968 (98%)	20 (2%)	1 (0%)	48	75
2	C	5/34 (15%)	5 (100%)	0	0	100	100
2	D	4/34 (12%)	4 (100%)	0	0	100	100
All	All	2001/2300 (87%)	1960 (98%)	39 (2%)	2 (0%)	48	75

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	854	VAL
1	B	854	VAL

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	873/958 (91%)	806 (92%)	67 (8%)	12	34
1	B	863/958 (90%)	794 (92%)	69 (8%)	11	32
2	C	4/23 (17%)	4 (100%)	0	100	100
2	D	4/23 (17%)	4 (100%)	0	100	100
All	All	1744/1962 (89%)	1608 (92%)	136 (8%)	11	34

5 of 136 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	862	THR
1	B	894	MET
1	B	1082	GLU
1	A	830	ASP
1	A	829	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	204	GLN
1	B	698	GLN
1	B	799	GLN
1	B	453	HIS
1	A	683	HIS

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, 95th 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 ²)	Q<0.9
1	A	1017/1116 (91%)	-0.74	1 (0%) 92 90	49, 77, 116, 142	0
1	B	1003/1116 (89%)	-0.70	1 (0%) 92 90	45, 82, 126, 152	0
2	C	7/34 (20%)	0.05	0 100 100	73, 98, 103, 104	0
2	D	6/34 (17%)	-0.15	0 100 100	64, 85, 88, 91	0
All	All	2033/2300 (88%)	-0.72	2 (0%) 92 90	45, 79, 122, 152	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	817	ILE	2.5
1	B	817	ILE	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 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.