



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 4, 2026 – 10:35 PM UTC

PDB ID : 2G99 / pdb_00002g99
Title : Structural basis for the specific recognition of methylated histone H3 lysine 4
by the WD-40 protein WDR5
Authors : Chai, J.; Han, Z.; Wang, H.; Shen, Y.
Deposited on : 2006-03-06
Resolution : 1.90 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
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

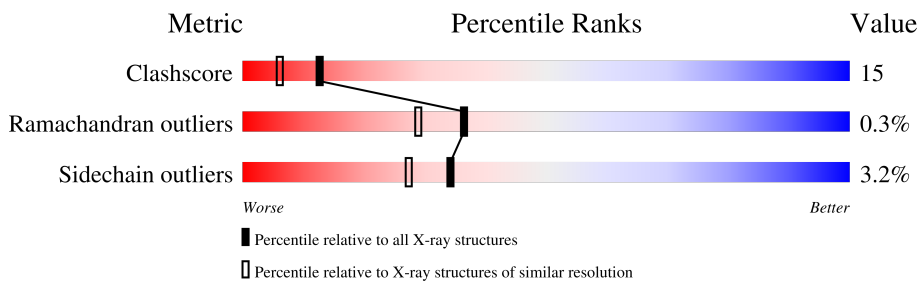
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.90 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	190562	8410 (1.90-1.90)
Ramachandran outliers	187476	8333 (1.90-1.90)
Sidechain outliers	187428	8333 (1.90-1.90)

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 $\leq 5\%$.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	308	 71% 25% ...
1	B	308	 75% 21% ...
2	C	10	 30% 10% 10% 50%
2	D	10	 10% 30% 10% 50%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4800 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 WD-repeat protein 5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	304	Total	C	N	O	S	0	0	0
			2357	1505	393	450	9			
1	B	304	Total	C	N	O	S	0	0	0
			2357	1505	393	450	9			

- Molecule 2 is a protein called Histone H3.

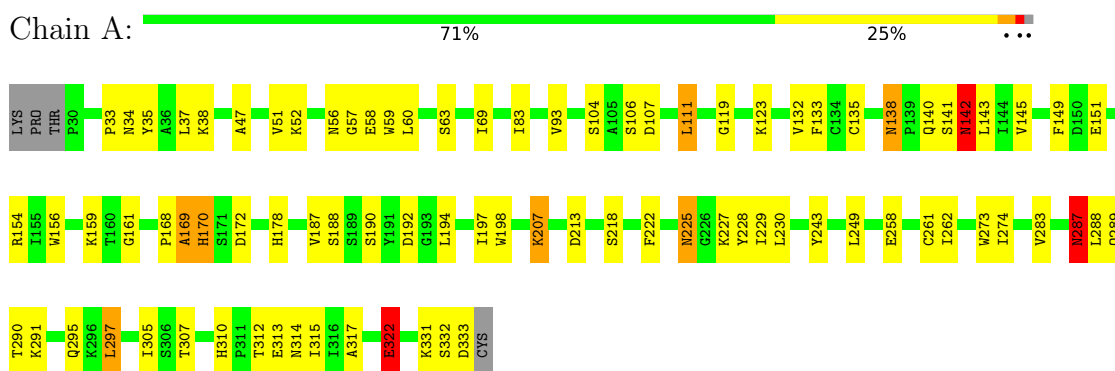
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	5	Total	C	N	O	0	0	0
			43	26	10	7			
2	D	5	Total	C	N	O	0	0	0
			43	26	10	7			

3 Residue-property plots

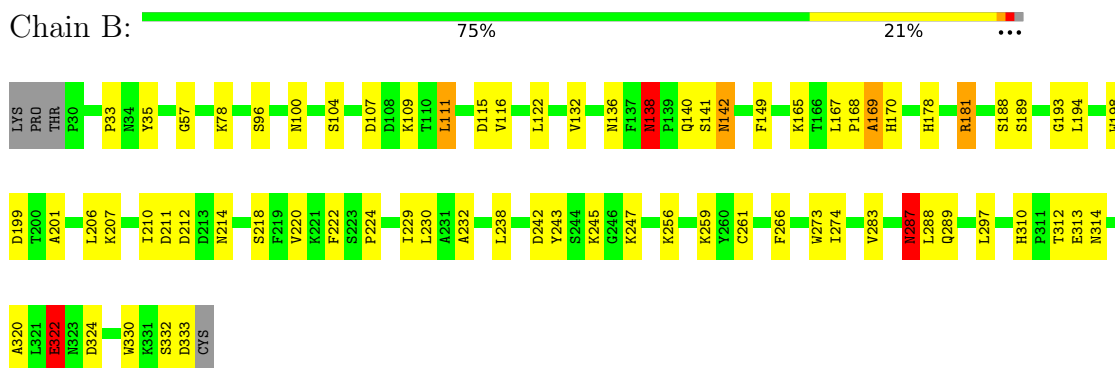
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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. 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.

Note EDS was not executed.

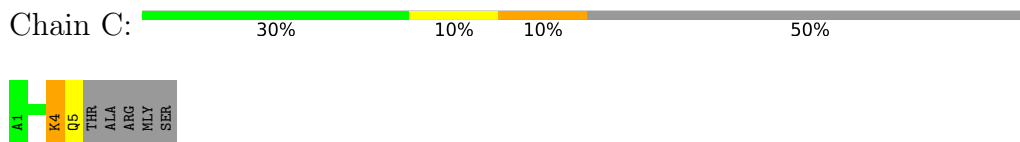
- Molecule 1: WD-repeat protein 5



- Molecule 1: WD-repeat protein 5



- Molecule 2: Histone H3



- Molecule 2: Histone H3



A1	R2	T3	K4	Q5	THR	ALA	ARG	MLY	SER
----	----	----	----	----	-----	-----	-----	-----	-----

4 Data and refinement statistics

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

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	134.19Å 46.89Å 112.31Å 90.00° 117.40° 90.00°	Depositor
Resolution (Å)	20.00 – 1.90	Depositor
% Data completeness (in resolution range)	(Not available) (20.00-1.90)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.212 , 0.243	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4800	wwPDB-VP
Average B, all atoms (Å ²)	21.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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

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	0.39	0/2414	0.97	10/3275 (0.3%)
1	B	0.38	0/2414	0.98	13/3275 (0.4%)
2	C	0.28	0/31	0.66	0/40
2	D	0.31	0/31	1.44	0/40
All	All	0.38	0/4890	0.98	23/6630 (0.3%)

There are no bond length outliers.

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	169	ALA	N-CA-C	8.64	120.32	111.07
1	B	169	ALA	N-CA-C	7.89	119.51	111.07
1	B	181	ARG	N-CA-C	7.14	121.58	113.01
1	B	168	PRO	N-CA-C	-6.86	103.23	112.48
1	B	111	LEU	N-CA-C	-6.50	99.33	109.72

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	2357	0	2342	85	0
1	B	2357	0	2342	57	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	43	0	50	2	0
2	D	43	0	50	5	0
All	All	4800	0	4784	146	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 15.

The worst 5 of 146 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:225:ASN:HD22	1:A:227:LYS:HG2	1.28	0.96
1:A:142:ASN:HD22	1:A:142:ASN:H	1.20	0.88
1:A:207:LYS:HZ3	1:A:243:TYR:HD1	1.18	0.88
1:B:115:ASP:HB2	1:B:122:LEU:HD11	1.58	0.86
1:A:187:VAL:HG22	1:A:197:ILE:HG13	1.59	0.84

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	302/308 (98%)	289 (96%)	12 (4%)	1 (0%)	36	29
1	B	302/308 (98%)	291 (96%)	10 (3%)	1 (0%)	36	29
2	C	2/10 (20%)	2 (100%)	0	0	100	100
2	D	2/10 (20%)	1 (50%)	1 (50%)	0	100	100
All	All	608/636 (96%)	583 (96%)	23 (4%)	2 (0%)	36	29

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	322	GLU
1	B	322	GLU

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	266/270 (98%)	255 (96%)	11 (4%)	27	19
1	B	266/270 (98%)	260 (98%)	6 (2%)	44	40
2	C	3/6 (50%)	3 (100%)	0	100	100
2	D	3/6 (50%)	3 (100%)	0	100	100
All	All	538/552 (98%)	521 (97%)	17 (3%)	34	27

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	194	LEU
1	B	287	ASN
1	A	287	ASN
1	A	289	GLN
1	A	297	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	178	HIS
1	B	310	HIS
1	A	287	ASN
1	A	289	GLN
1	A	310	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues 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	MLY	C	4	2	9,10,11	0.75	0	6,11,13	1.65	2 (33%)
2	MLY	D	4	2	9,10,11	0.86	0	6,11,13	2.43	4 (66%)

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	MLY	C	4	2	-	4/8/9/11	-
2	MLY	D	4	2	-	2/8/9/11	-

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	4	MLY	CD-CE-NZ	3.44	122.61	113.71
2	D	4	MLY	CH1-NZ-CE	3.08	122.97	110.75
2	D	4	MLY	CH2-NZ-CE	2.79	121.82	110.75
2	C	4	MLY	CH2-NZ-CE	2.43	120.38	110.75
2	D	4	MLY	CH2-NZ-CH1	-2.37	103.65	109.72

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	4	MLY	N-CA-CB-CG
2	C	4	MLY	C-CA-CB-CG
2	D	4	MLY	C-CA-CB-CG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	C	4	MLY	CA-CB-CG-CD
2	C	4	MLY	CG-CD-CE-NZ

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	4	MLY	1	0
2	D	4	MLY	4	0

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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.