



wwPDB NMR Structure Validation Summary Report ⓘ

Mar 4, 2026 – 08:22 PM UTC

PDB ID : 2JMJ / pdb_00002jnj
Title : NMR solution structure of the PHD domain from the yeast YNG1 protein in complex with H3(1-9)K4me3 peptide
Authors : Ilin, S.; Taverna, S.D.; Rogers, R.S.; Tanny, J.C.; Lavender, H.; Li, H.; Baker, L.; Boyle, J.; Blair, L.P.; Chait, B.T.; Patel, D.J.; Aitchison, J.D.; Tackett, A.J.; Allis, C.D.
Deposited on : 2006-11-15

This is a wwPDB NMR 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/NMRValidationReportHelp>

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)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

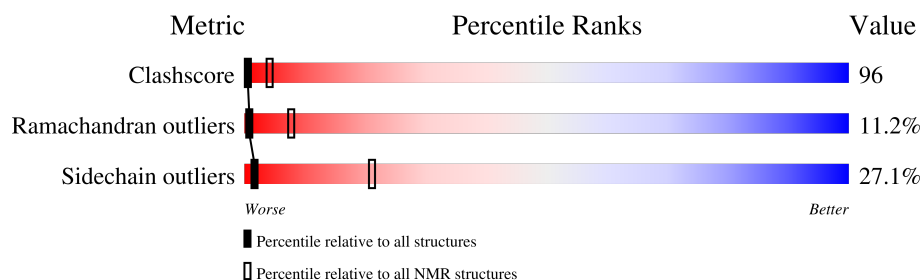
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment was not calculated.

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)	NMR archive (#Entries)
Clashscore	229148	14424
Ramachandran outliers	224038	12848
Sidechain outliers	223484	12823

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	90	
2	P	9	

2 Ensemble composition and analysis

This entry contains 20 models. Model 19 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *closest to the average*.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	Residue range (total)	Backbone RMSD (Å)	Medoid model
1	A:25-A:78, P:1-P:3 (57)	0.41	19

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 3 clusters. No single-model clusters were found.

Cluster number	Models
1	2, 3, 5, 6, 9, 11, 12, 13, 15, 16, 18, 19
2	1, 7, 8, 10, 17
3	4, 14, 20

3 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 1094 atoms, of which 536 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Protein YNG1.

Mol	Chain	Residues	Atoms						Trace
1	A	60	Total	C	H	N	O	S	0
			924	304	444	82	86	8	

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	cloning artifact	UNP Q08465
A	2	PRO	-	cloning artifact	UNP Q08465
A	3	LEU	-	cloning artifact	UNP Q08465
A	4	GLY	-	cloning artifact	UNP Q08465
A	5	SER	-	cloning artifact	UNP Q08465
A	6	HIS	-	cloning artifact	UNP Q08465
A	7	MET	-	cloning artifact	UNP Q08465
A	8	ALA	-	cloning artifact	UNP Q08465
A	9	SER	-	cloning artifact	UNP Q08465
A	10	GLU	-	cloning artifact	UNP Q08465
A	11	PHE	-	cloning artifact	UNP Q08465

- Molecule 2 is a protein called Histone H3.

Mol	Chain	Residues	Atoms					Trace
2	P	9	Total	C	H	N	O	0
			168	46	92	18	12	

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
P	4	M3L	LYS	modified residue	UNP P61830

- Molecule 3 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	
3	A	2	Total	Zn
			2	2

A1	R2	T3	K4	Q5	T6	A7	R8	K9
----	----	----	----	----	----	----	----	----

5 Refinement protocol and experimental data overview ⓘ

The models were refined using the following method: *torsion angle dynamics*.

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
X-PLOR NIH	structure solution	
X-PLOR NIH	refinement	

No chemical shift data was provided.

6 Model quality [i](#)

6.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: M3L, ZN

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 (average) 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	0.65±0.01	0±0/446 (0.0± 0.0%)	1.05±0.01	0±1/605 (0.1± 0.1%)
2	P	0.72±0.02	0±0/22 (0.0± 0.0%)	0.76±0.04	0±0/28 (0.0± 0.0%)
All	All	0.65	0/9360 (0.0%)	1.04	7/12660 (0.1%)

There are no bond-length outliers.

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	55	GLY	N-CA-C	-5.86	107.39	114.66	9	6
1	A	52	PHE	CA-CB-CG	-5.03	108.77	113.80	4	1

There are no chirality outliers.

There are no planarity outliers.

6.2 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 each 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 averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	430	392	398	83±7
2	P	23	27	27	4±2
3	A	2	0	0	0±1
All	All	9100	8380	8490	1682

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 96.

5 of 277 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:47:CYS:SG	1:A:69:CYS:SG	1.47	1.48	17	7
1:A:42:CYS:CB	1:A:69:CYS:SG	1.22	2.27	18	20
1:A:69:CYS:SG	1:A:73:CYS:CB	1.22	2.28	9	20
1:A:42:CYS:SG	1:A:69:CYS:CB	1.12	2.38	15	20
1:A:31:CYS:CB	1:A:56:CYS:SG	1.10	2.38	5	20

6.3 Torsion angles [i](#)

6.3.1 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 PDB entries followed by that with respect to all NMR entries. 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	54/90 (60%)	33±2 (62±3%)	14±2 (26±3%)	6±2 (12±4%)	1	7
2	P	2/9 (22%)	2±0 (88±22%)	0±0 (12±22%)	0±0 (0±0%)	100	100
All	All	1120/1980 (57%)	704 (63%)	290 (26%)	126 (11%)	1	8

5 of 21 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	47	CYS	20
1	A	57	VAL	20
1	A	69	CYS	20
1	A	37	GLY	12
1	A	61	GLN	7

6.3.2 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 PDB entries followed by that with respect to all NMR entries. 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	46/77 (60%)	34±1 (74±3%)	12±1 (26±3%)	2	23
2	P	2/6 (33%)	1±0 (48±11%)	1±0 (52±11%)	0	1
All	All	960/1660 (58%)	700 (73%)	260 (27%)	2	21

5 of 25 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	31	CYS	20
1	A	59	LEU	20
1	A	69	CYS	20
2	P	2	ARG	20
1	A	33	ASN	19

6.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds 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 is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
2	M3L	P	4	2	10,11,12	0.52±0.08	0±0 (0±0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of 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 angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
2	M3L	P	4	2	9,14,16	0.61±0.06	0±0 (0±0%)

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	M3L	P	4	2	-	0±0,9,10,12	-

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.

6.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

6.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

6.7 Other polymers [i](#)

There are no such molecules in this entry.

6.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

7 Chemical shift validation

No chemical shift data were provided