



wwPDB NMR Structure Validation Summary Report ⓘ

Mar 19, 2026 – 08:01 PM UTC

PDB ID : 2KDY / pdb_00002kdy
Title : NMR structure of LP2086-B01
Authors : Mascioni, A.; Bentley, B.E.; Camarda, R.; Dilts, D.A.; Fink, P.; Gusarova, V.; Hoiseth, S.; Jacob, J.; Lin, S.L.; Malakian, K.; McNeil, L.K.; Mininni, T.; Moy, F.; Murphy, E.; Novikova, E.; Sigethy, S.; Wen, Y.; Zlotnick, G.W.; Tsao, D.H.H.
Deposited on : 2009-01-20

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
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

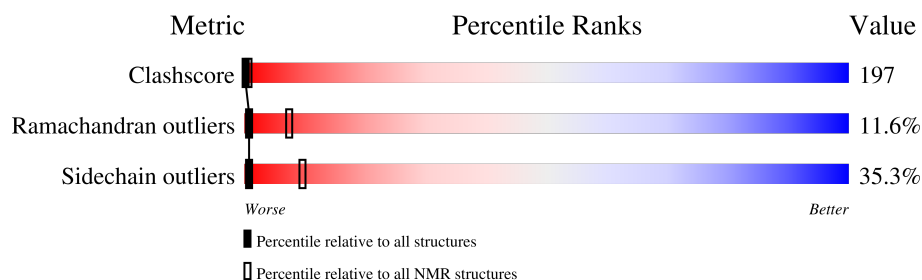
1 Overall quality at a glance

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	261	

2 Ensemble composition and analysis

This entry contains 17 models. Model 5 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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:19-A:122, A:129-A:190, A:195-A:261 (233)	1.73	5

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 2 clusters and 1 single-model cluster was found.

Cluster number	Models
1	2, 3, 4, 5, 8, 9, 12, 14, 15, 16, 17
2	1, 7, 10, 11, 13
Single-model clusters	6

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 3847 atoms, of which 1900 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Factor H binding protein variant B01_001.

Mol	Chain	Residues	Atoms						Trace
1	A	261	Total	C	H	N	O	S	0
			3847	1203	1900	340	401	3	

There are 2 discrepancies between the modelled and reference sequences:

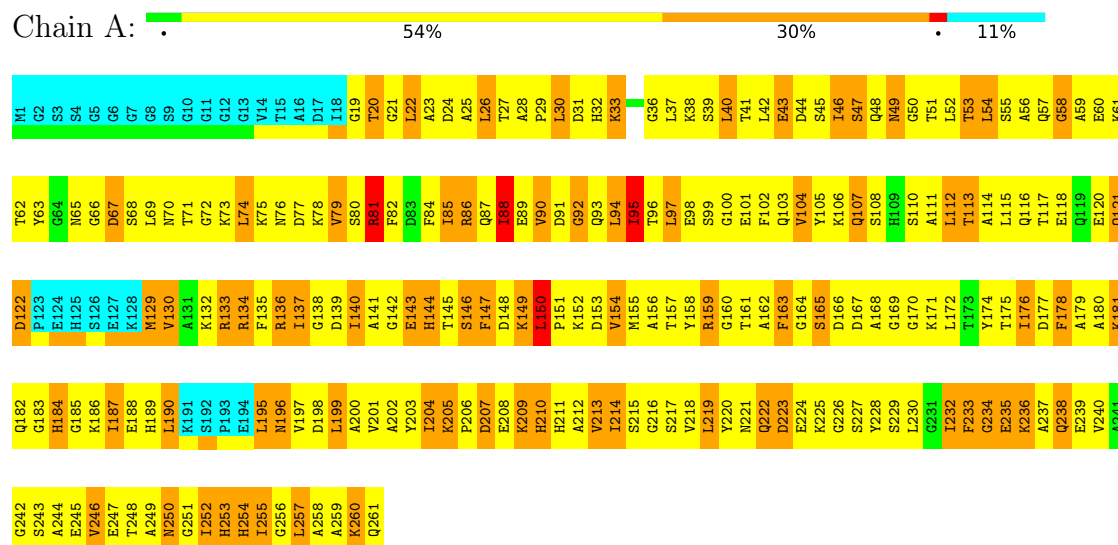
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	expression tag	UNP Q6VRY4
A	2	GLY	-	expression tag	UNP Q6VRY4

4 Residue-property plots

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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 outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

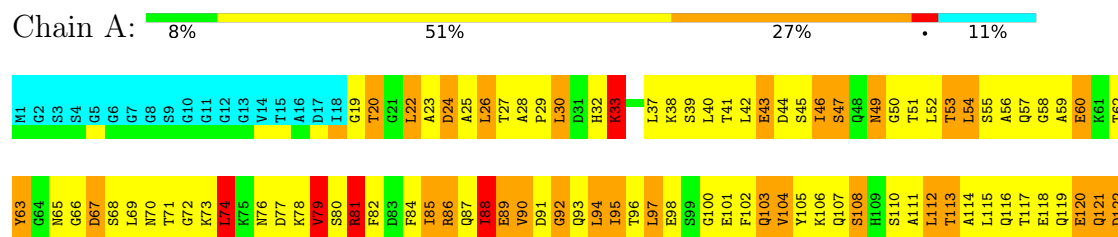
- Molecule 1: Factor H binding protein variant B01_001



4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 5. Colouring as in section 4.1 above.

- Molecule 1: Factor H binding protein variant B01_001



A244	H184	P123
E245	G185	E124
V246	K186	H125
E247	L187	S126
T248	E188	E127
A249	H189	K128
N250	L190	M129
G251	K191	V130
I252	S192	A131
H253	P193	K132
H254	E194	R133
I255	L195	R134
G256	M196	F135
L257	V197	R136
A258	D198	L137
A259	L199	G138
K260	A200	D139
Q261	V201	I140
	A202	A141
	Y203	G142
	I204	E143
	K205	H144
	P206	T145
	D207	S146
	E208	F147
	K209	D148
	H210	K149
	H211	L150
	A212	P151
	G213	K152
	I214	D153
	S215	V154
	G216	M155
	S217	A156
	V218	T157
	L219	Y158
	Y220	R159
	N221	G160
	Q222	T161
	D223	A162
	E224	F163
	K225	G164
	G226	S165
	S227	D166
	Y228	G169
	S229	L230
	L231	K171
	I232	L172
	G233	T173
	G234	Y174
	E235	T175
	K236	L176
	A237	D177
	Q238	F178
	E239	A179
	V240	A180
	A241	K181
	G242	O182
	S243	G183

5 Refinement protocol and experimental data overview ⓘ

The models were refined using the following method: *simulated annealing*.

Of the 100 calculated structures, 17 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
CNS	refinement	

No chemical shift data was provided.

6 Model quality i

6.1 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 (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	1.46±0.01	1±1/1795 (0.0± 0.0%)	1.65±0.01	8±2/2416 (0.3± 0.1%)
All	All	1.46	9/30515 (0.0%)	1.65	142/41072 (0.3%)

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	Planarity
1	A	0.0±0.0	6.0±0.0
All	All	0	102

All unique bond 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(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	207	ASP	N-CA	5.62	1.49	1.46	11	2
1	A	107	GLN	N-CA	5.58	1.49	1.46	11	3
1	A	64	GLY	N-CA	5.41	1.49	1.44	9	2
1	A	200	ALA	C-N	-5.21	1.30	1.33	10	1
1	A	164	GLY	N-CA	5.13	1.49	1.44	4	1

5 of 26 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	154	VAL	N-CA-CB	-7.40	102.63	112.10	17	14
1	A	90	VAL	N-CA-CB	-6.56	104.86	112.34	6	16
1	A	176	ILE	N-CA-CB	-6.49	104.23	112.60	13	11
1	A	214	ILE	N-CA-CB	-6.26	104.91	112.35	9	2
1	A	137	ILE	N-CA-CB	-6.22	103.99	110.95	16	8

There are no chirality outliers.

5 of 6 unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	81	ARG	Sidechain	17
1	A	86	ARG	Sidechain	17
1	A	133	ARG	Sidechain	17
1	A	134	ARG	Sidechain	17
1	A	136	ARG	Sidechain	17

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	1769	1736	1736	689±35
All	All	30073	29512	29512	11720

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 197.

5 of 5052 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:147:PHE:CD1	1:A:212:ALA:HB2	1.25	1.66	4	15
1:A:141:ALA:HB3	1:A:144:HIS:CD2	1.21	1.70	12	2
1:A:26:LEU:HD23	1:A:54:LEU:CD1	1.19	1.67	14	6
1:A:199:LEU:CD2	1:A:218:VAL:HG22	1.17	1.68	11	6
1:A:141:ALA:HB1	1:A:233:PHE:CD2	1.16	1.75	1	4

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	232/261 (89%)	155±5 (67±2%)	50±4 (21±2%)	27±4 (12±2%)	1 7
All	All	3944/4437 (89%)	2643 (67%)	845 (21%)	456 (12%)	1 7

5 of 77 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	58	GLY	17
1	A	95	ILE	17
1	A	178	PHE	17
1	A	92	GLY	15
1	A	67	ASP	15

6.3.2 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 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	185/203 (91%)	120±5 (65±3%)	65±5 (35±3%)	1 9
All	All	3145/3451 (91%)	2035 (65%)	1110 (35%)	1 9

5 of 163 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	20	THR	17
1	A	22	LEU	17
1	A	53	THR	17
1	A	187	ILE	17
1	A	255	ILE	17

6.3.3 RNA ⓘ

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

6.6 Ligand geometry [i](#)

There are no ligands in this entry.

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