



# wwPDB NMR Structure Validation Summary Report ⓘ

Jun 12, 2024 – 02:53 PM EDT

PDB ID : 2L2N  
BMRB ID : 17143  
Title : Backbone <sup>1</sup>H, <sup>13</sup>C, and <sup>15</sup>N Chemical Shift Assignments for the first dsRBD of protein HYL1  
Authors : Rasia, R.M.; Mateos, J.L.; Bologna, N.G.; Burdisso, P.; Imbert, L.; Palatnik, J.F.; Boisbouvier, J.  
Deposited on : 2010-08-23

This is a wwPDB NMR 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/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 : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
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) : 2.36.2

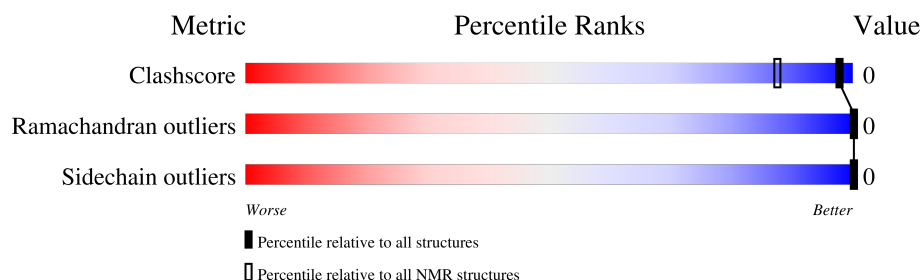
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 37%.

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	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

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  $\geq 3$ , 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	103	

## 2 Ensemble composition and analysis

This entry contains 10 models. Model 1 is the overall representative, medoid model (most similar to other models).

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:20-A:86 (67)	1.04	1

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 3 single-model clusters were found.

Cluster number	Models
1	1, 2, 3, 6, 8
2	7, 9
Single-model clusters	4; 5; 10

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1143 atoms, of which 576 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Hyponastic leave 1.

Mol	Chain	Residues	Atoms					Trace
1	A	71	Total	C	H	N	O	0
			1143	364	576	98	105	

There are 3 discrepancies between the modelled and reference sequences:

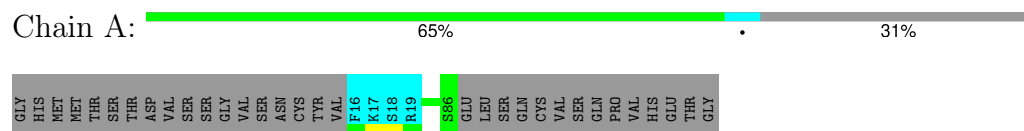
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP O04492
A	-1	HIS	-	expression tag	UNP O04492
A	0	MET	-	expression tag	UNP O04492

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

### 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.

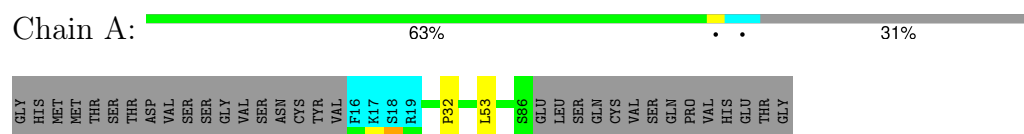
- Molecule 1: Hyponastic leave 1



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

The representative model is number 1. Colouring as in section 4.1 above.

- Molecule 1: Hyponastic leave 1



## 5 Refinement protocol and experimental data overview

The models were refined using the following method: *rosetta full atom relaxation*.

Of the 1000 calculated structures, 10 were deposited, based on the following criterion: *target function*.

The authors did not provide any information on software used for structure solution, optimization or refinement.

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	500
Number of shifts mapped to atoms	363
Number of unparsed shifts	0
Number of shifts with mapping errors	137
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	37%

## 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	0.75±0.01	0±0/541 ( 0.0± 0.1%)	0.50±0.01	0±0/730 ( 0.0± 0.0%)
All	All	0.75	1/5410 ( 0.0%)	0.50	0/7300 ( 0.0%)

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	62	PRO	CG-CD	5.95	1.70	1.50	5	1

There are no bond-angle outliers.

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	530	536	536	0±0
All	All	5300	5360	5360	4

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 0.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:32:PRO:HG3	1:A:53:LEU:HD13	0.53	1.80	1	1

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:20:LEU:HB3	1:A:32:PRO:HG2	0.48	1.85	5	1
1:A:20:LEU:O	1:A:20:LEU:HD23	0.46	2.11	10	1
1:A:44:LYS:HG2	1:A:44:LYS:O	0.41	2.16	3	1

## 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	66/103 (64%)	65±0 (99±1%)	1±0 (1±1%)	0±0 (0±0%)	100	100
All	All	660/1030 (64%)	652 (99%)	8 (1%)	0 (0%)	100	100

There are no Ramachandran outliers.

### 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	57/90 (63%)	57±0 (100±0%)	0±0 (0±0%)	100	100
All	All	570/900 (63%)	570 (100%)	0 (0%)	100	100

There are no protein residues with a non-rotameric sidechain to report.

### 6.3.3 RNA [i](#)

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 monosaccharides 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

The completeness of assignment taking into account all chemical shift lists is 37% for the well-defined parts and 37% for the entire structure.

### 7.1 Chemical shift list 1

File name: working\_cs.cif

Chemical shift list name: *assigned\_chem\_shift\_list\_1*

#### 7.1.1 Bookkeeping

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	500
Number of shifts mapped to atoms	363
Number of unparsed shifts	0
Number of shifts with mapping errors	137
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. First 5 (of 137) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	3	SER	HA	4.41	0.030	1
1	A	3	SER	C	175.02	0.300	1
1	A	3	SER	CA	58.59	0.300	1
1	A	3	SER	CB	64.026	0.300	1
1	A	4	THR	H	8.16	0.030	1
1	A	4	THR	HA	4.21	0.030	1
1	A	4	THR	C	174.39	0.300	1
1	A	4	THR	CA	62.151	0.300	1
1	A	4	THR	CB	69.719	0.300	1
1	A	4	THR	N	115.16	0.300	1
1	A	5	ASP	H	8.234	0.030	1
1	A	5	ASP	HA	4.54	0.030	1
1	A	5	ASP	C	176.58	0.300	1
1	A	5	ASP	CA	54.542	0.300	1

*Continued on next page...*

*Continued from previous page...*

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	5	ASP	CB	41.236	0.300	1
1	A	5	ASP	N	122.75	0.300	1
1	A	6	VAL	H	8.072	0.030	1
1	A	6	VAL	HA	4.06	0.030	1
1	A	6	VAL	C	176.65	0.300	1
1	A	6	VAL	CA	62.527	0.300	1
1	A	6	VAL	CB	32.394	0.300	1
1	A	6	VAL	N	119.83	0.300	1
1	A	7	SER	H	8.371	0.030	1
1	A	7	SER	HA	4.33	0.030	1
1	A	7	SER	C	175.09	0.300	1
1	A	7	SER	CA	58.898	0.300	1
1	A	7	SER	CB	63.849	0.300	1
1	A	7	SER	N	118.64	0.300	1
1	A	8	SER	H	8.324	0.030	1
1	A	8	SER	C	175.11	0.300	1
1	A	8	SER	N	110.59	0.300	1
1	A	9	GLY	C	174.31	0.300	1
1	A	9	GLY	CA	45.465	0.300	1
1	A	10	VAL	H	7.886	0.030	1
1	A	10	VAL	HA	4.05	0.030	1
1	A	10	VAL	C	176.46	0.300	1
1	A	10	VAL	CA	62.483	0.300	1
1	A	10	VAL	CB	32.506	0.300	1
1	A	10	VAL	N	118.69	0.300	1
1	A	11	SER	H	8.312	0.030	1
1	A	11	SER	C	174.51	0.300	1
1	A	11	SER	CA	58.418	0.300	1
1	A	11	SER	CB	63.79	0.300	1
1	A	11	SER	N	118.46	0.300	1
1	A	13	CYS	HA	4.28	0.030	1
1	A	13	CYS	C	174.45	0.300	1
1	A	13	CYS	CA	58.764	0.300	1
1	A	13	CYS	CB	27.622	0.300	1
1	A	14	TYR	H	8.131	0.030	1
1	A	14	TYR	HA	4.52	0.030	1
1	A	14	TYR	C	176.78	0.300	1
1	A	14	TYR	CA	56.692	0.300	1
1	A	14	TYR	CB	37.716	0.300	1
1	A	14	TYR	N	121.16	0.300	1
1	A	15	VAL	H	7.387	0.030	1

*Continued on next page...*

*Continued from previous page...*

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	15	VAL	HA	4.27	0.030	1
1	A	15	VAL	C	176.42	0.300	1
1	A	15	VAL	CA	61.599	0.300	1
1	A	15	VAL	CB	32.453	0.300	1
1	A	15	VAL	N	113.53	0.300	1
1	A	87	GLU	HA	4.16	0.030	1
1	A	87	GLU	C	177.18	0.300	1
1	A	87	GLU	CA	57.316	0.300	1
1	A	87	GLU	CB	29.793	0.300	1
1	A	88	LEU	H	8.058	0.030	1
1	A	88	LEU	HA	4.25	0.030	1
1	A	88	LEU	C	177.91	0.300	1
1	A	88	LEU	CA	55.725	0.300	1
1	A	88	LEU	CB	42.155	0.300	1
1	A	88	LEU	N	121.56	0.300	1
1	A	89	SER	H	8.122	0.030	1
1	A	89	SER	HA	4.32	0.030	1
1	A	89	SER	C	174.8	0.300	1
1	A	89	SER	CA	58.91	0.300	1
1	A	89	SER	CB	63.709	0.300	1
1	A	89	SER	N	115.48	0.300	1
1	A	90	GLN	H	8.242	0.030	1
1	A	90	GLN	HA	4.26	0.030	1
1	A	90	GLN	C	175.96	0.300	1
1	A	90	GLN	CA	55.963	0.300	1
1	A	90	GLN	CB	29.286	0.300	1
1	A	90	GLN	N	121.21	0.300	1
1	A	91	CYS	H	8.256	0.030	1
1	A	91	CYS	HA	4.39	0.030	1
1	A	91	CYS	C	174.72	0.300	1
1	A	91	CYS	CA	58.741	0.300	1
1	A	91	CYS	CB	27.983	0.300	1
1	A	91	CYS	N	120.04	0.300	1
1	A	92	VAL	H	8.185	0.030	1
1	A	92	VAL	HA	4.1	0.030	1
1	A	92	VAL	C	176.11	0.300	1
1	A	92	VAL	CA	62.37	0.300	1
1	A	92	VAL	CB	32.747	0.300	1
1	A	92	VAL	N	122.03	0.300	1
1	A	93	SER	H	8.317	0.030	1
1	A	93	SER	HA	4.34	0.030	1

*Continued on next page...*

*Continued from previous page...*

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	93	SER	C	174.02	0.300	1
1	A	93	SER	CA	58.26	0.300	1
1	A	93	SER	CB	63.856	0.300	1
1	A	93	SER	N	119.35	0.300	1
1	A	94	GLN	H	8.297	0.030	1
1	A	94	GLN	C	173.82	0.300	1
1	A	94	GLN	CA	53.664	0.300	1
1	A	94	GLN	CB	29.116	0.300	1
1	A	94	GLN	N	123.04	0.300	1
1	A	95	PRO	HA	3.87	0.030	1
1	A	95	PRO	C	176.79	0.300	1
1	A	95	PRO	CA	63.138	0.300	1
1	A	95	PRO	CB	32.007	0.300	1
1	A	96	VAL	H	8.202	0.030	1
1	A	96	VAL	HA	3.94	0.030	1
1	A	96	VAL	C	175.99	0.300	1
1	A	96	VAL	CA	62.415	0.300	1
1	A	96	VAL	CB	32.64	0.300	1
1	A	96	VAL	N	120.27	0.300	1
1	A	97	HIS	H	8.352	0.030	1
1	A	97	HIS	HA	4.59	0.030	1
1	A	97	HIS	C	175.09	0.300	1
1	A	97	HIS	CA	55.761	0.300	1
1	A	97	HIS	CB	30.585	0.300	1
1	A	97	HIS	N	122.91	0.300	1
1	A	98	GLU	H	8.446	0.030	1
1	A	98	GLU	HA	4.27	0.030	1
1	A	98	GLU	C	176.57	0.300	1
1	A	98	GLU	CA	56.706	0.300	1
1	A	98	GLU	CB	30.386	0.300	1
1	A	98	GLU	N	122.98	0.300	1
1	A	99	THR	H	8.194	0.030	1
1	A	99	THR	HA	4.23	0.030	1
1	A	99	THR	C	174.36	0.300	1
1	A	99	THR	CA	61.843	0.300	1
1	A	99	THR	CB	70.065	0.300	1
1	A	99	THR	N	114.42	0.300	1
1	A	100	GLY	H	8.031	0.030	1
1	A	100	GLY	C	179.05	0.300	1
1	A	100	GLY	CA	46.279	0.300	1
1	A	100	GLY	N	117.31	0.300	1

### 7.1.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	90	$-0.21 \pm 0.15$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}_\beta$	84	$0.03 \pm 0.19$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}'$	90	$-0.16 \pm 0.19$	None needed ( $< 0.5$ ppm)
$^{15}\text{N}$	80	$0.58 \pm 0.54$	None needed (imprecise)

### 7.1.3 Completeness of resonance assignments [i](#)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 37%, i.e. 345 atoms were assigned a chemical shift out of a possible 923. 0 out of 12 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	287/330 (87%)	109/133 (82%)	122/134 (91%)	56/63 (89%)
Sidechain	58/520 (11%)	0/338 (0%)	58/161 (36%)	0/21 (0%)
Aromatic	0/73 (0%)	0/35 (0%)	0/37 (0%)	0/1 (0%)
Overall	345/923 (37%)	109/506 (22%)	180/332 (54%)	56/85 (66%)

### 7.1.4 Statistically unusual chemical shifts [i](#)

There are no statistically unusual chemical shifts.

### 7.1.5 Random Coil Index (RCI) plots [i](#)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition. If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:

