



# wwPDB NMR Structure Validation Summary Report ⓘ

Mar 5, 2026 – 09:29 PM UTC

PDB ID : 2MPC / pdb\_00002mpc  
BMRB ID : 19875  
Title : Solution structure of the pyrin domain of human Pyrin  
Authors : Soh, S.L.; Smith, S.J.; Hill, J.M.  
Deposited on : 2014-05-15

This is a wwPDB NMR Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](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 : 4-5-2 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

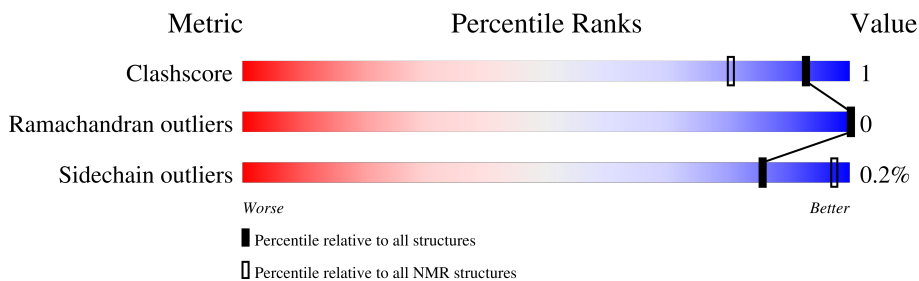
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 40%.

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  $\geq 3$ , 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	100	 88% 10%

## 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:4-A:93 (90)	0.98	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 2 single-model clusters were found.

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

### 3 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 1508 atoms, of which 766 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Pysin.

Mol	Chain	Residues	Atoms					Trace	
			Total	C	H	N	O		S
1	A	90	1508	474	766	132	135	1	0

There are 8 discrepancies between the modelled and reference sequences:

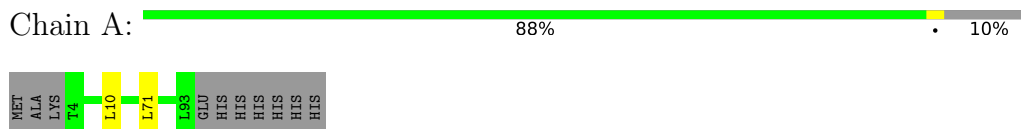
Chain	Residue	Modelled	Actual	Comment	Reference
A	93	LEU	-	expression tag	UNP O15553
A	94	GLU	-	expression tag	UNP O15553
A	95	HIS	-	expression tag	UNP O15553
A	96	HIS	-	expression tag	UNP O15553
A	97	HIS	-	expression tag	UNP O15553
A	98	HIS	-	expression tag	UNP O15553
A	99	HIS	-	expression tag	UNP O15553
A	100	HIS	-	expression tag	UNP O15553

## 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.

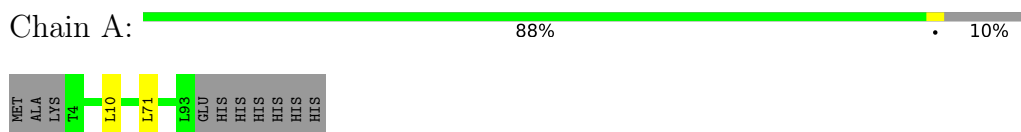
- Molecule 1: Pysin



### 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: Pysin



## 5 Refinement protocol and experimental data overview

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

Of the 5000 calculated structures, 10 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
CS-ROSETTA	refinement	3.4

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	584
Number of shifts mapped to atoms	527
Number of unparsed shifts	0
Number of shifts with mapping errors	57
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	40%

## 6 Model quality [i](#)

### 6.1 Standard geometry [i](#)

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

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	742	766	765	2±1
All	All	7420	7660	7650	21

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 1.

5 of 10 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:10:LEU:HD23	1:A:10:LEU:C	0.64	2.18	6	6
1:A:74:LEU:HD12	1:A:74:LEU:C	0.59	2.22	2	1
1:A:71:LEU:O	1:A:71:LEU:HD23	0.56	1.99	4	1
1:A:71:LEU:HD23	1:A:71:LEU:C	0.56	2.26	7	6
1:A:10:LEU:C	1:A:10:LEU:CD2	0.52	2.83	6	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	88/100 (88%)	87±1 (99±1%)	1±1 (1±1%)	0±0 (0±0%)	100	100
All	All	880/1000 (88%)	869 (99%)	11 (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	82/91 (90%)	82±0 (100±0%)	0±0 (0±0%)	85	97
All	All	820/910 (90%)	818 (100%)	2 (0%)	85	97

All 2 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	56	LEU	1
1	A	69	LEU	1

### 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 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 i

The completeness of assignment taking into account all chemical shift lists is 40% for the well-defined parts and 40% 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 i

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	584
Number of shifts mapped to atoms	527
Number of unparsed shifts	0
Number of shifts with mapping errors	57
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 57) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	1	MET	C	175.653	.	1
1	A	1	MET	CA	54.513	.	1
1	A	1	MET	CB	33.382	.	1
1	A	1	MET	N	128.59	.	1
1	A	2	ALA	H	8.418	.	1
1	A	2	ALA	HA	4.11	.	1
1	A	2	ALA	C	177.445	.	1
1	A	2	ALA	CA	52.866	.	1
1	A	2	ALA	CB	19.074	.	1
1	A	3	LYS	H	8.408	.	1
1	A	3	LYS	HA	4.56	.	1
1	A	3	LYS	C	176.614	.	1
1	A	3	LYS	CA	56.238	.	1
1	A	3	LYS	CB	33.96	.	1

*Continued on next page...*

*Continued from previous page...*

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	3	LYS	N	120.839	.	1
1	A	94	GLU	H	7.817	.	1
1	A	94	GLU	HA	4.141	.	1
1	A	94	GLU	C	176.59	.	1
1	A	94	GLU	CA	56.992	.	1
1	A	94	GLU	CB	29.465	.	1
1	A	94	GLU	N	119.448	.	1
1	A	95	HIS	H	8.271	.	1
1	A	95	HIS	HA	4.633	.	1
1	A	95	HIS	C	174.489	.	1
1	A	95	HIS	CA	55.474	.	1
1	A	95	HIS	CB	28.814	.	1
1	A	95	HIS	N	118.24	.	1
1	A	96	HIS	H	8.436	.	1
1	A	96	HIS	HA	4.636	.	1
1	A	96	HIS	C	174.245	.	1
1	A	96	HIS	CA	55.638	.	1
1	A	96	HIS	CB	28.979	.	1
1	A	96	HIS	N	118.815	.	1
1	A	97	HIS	H	8.591	.	1
1	A	97	HIS	HA	4.677	.	1
1	A	97	HIS	C	174.218	.	1
1	A	97	HIS	CA	55.545	.	1
1	A	97	HIS	CB	29.126	.	1
1	A	97	HIS	N	119.516	.	1
1	A	98	HIS	H	8.69	.	1
1	A	98	HIS	HA	4.669	.	1
1	A	98	HIS	C	173.992	.	1
1	A	98	HIS	CA	55.696	.	1
1	A	98	HIS	CB	29.366	.	1
1	A	98	HIS	N	120.384	.	1
1	A	99	HIS	H	8.62	.	1
1	A	99	HIS	HA	4.628	.	1
1	A	99	HIS	C	173.538	.	1
1	A	99	HIS	CA	55.711	.	1
1	A	99	HIS	CB	29.393	.	1
1	A	99	HIS	N	120.617	.	1
1	A	100	HIS	H	8.392	.	1
1	A	100	HIS	HA	4.448	.	1
1	A	100	HIS	C	178.848	.	1
1	A	100	HIS	CA	57.465	.	1

*Continued on next page...*

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	100	HIS	CB	29.817	.	1
1	A	100	HIS	N	125.55	.	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
$^{13}\text{C}_\alpha$	100	$-0.76 \pm 0.14$	Should be checked
$^{13}\text{C}_\beta$	99	$0.40 \pm 0.13$	None needed (< 0.5 ppm)
$^{13}\text{C}'$	98	$-0.40 \pm 0.14$	None needed (< 0.5 ppm)
$^{15}\text{N}$	93	$0.61 \pm 0.17$	Should be applied

### 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 40%, i.e. 527 atoms were assigned a chemical shift out of a possible 1328. 0 out of 20 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	438/443 (99%)	176/177 (99%)	178/180 (99%)	84/86 (98%)
Sidechain	89/798 (11%)	0/519 (0%)	89/243 (37%)	0/36 (0%)
Aromatic	0/87 (0%)	0/43 (0%)	0/41 (0%)	0/3 (0%)
Overall	527/1328 (40%)	176/739 (24%)	267/464 (58%)	84/125 (67%)

### 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:

