

# Full wwPDB NMR Structure Validation Report (i)

Aug 8, 2023 – 01:20 pm BST

PDB ID : 8OR5 BMRB ID : 34804

> Title : Solution NMR structure of Notch1 TMD Authors : Guschtschin-Schmidt, N.; Muhle-Goll, C.

Deposited on : 2023-04-13

This is a Full wwPDB NMR Structure Validation 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 (i) 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 (1)) 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)

 $\begin{array}{ccc} wwPDB\text{-}ShiftChecker &:& v1.2\\ BMRB \ Restraints \ Analysis &:& v1.2 \end{array}$ 

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

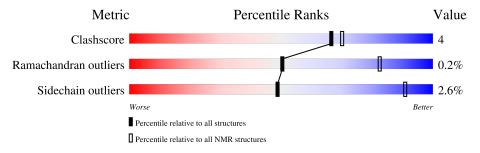
Validation Pipeline (wwPDB-VP) : 2.35

## 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:  $SOLUTION\ NMR$ 

The overall completeness of chemical shifts assignment is 69%.

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	NMR archive
Metric	$(\# \mathrm{Entries})$	$(\# \mathrm{Entries})$
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 >=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 <=5%

Mol	Chain	Length	Quality of chain		
1	A	30	87%	•	10%



## 2 Ensemble composition and analysis (i)

This entry contains 40 models. Model 23 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:1733-A:1747 (15)	0.23	23	
2	A:1748-A:1759 (12)	0.73	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 5 clusters. No single-model clusters were found.

Cluster number	Models
1	1, 3, 4, 5, 6, 9, 11, 12, 13, 14, 18, 19, 21, 24, 27,
1	29, 33, 34, 37, 40
2	7, 15, 26, 28, 30, 35, 38, 39
3	17, 22, 23, 31, 32, 36
4	2, 10, 25
5	8, 16, 20



## 3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 506 atoms, of which 269 are hydrogens and 0 are deuteriums.

• Molecule 1 is a protein called Notch 1 extracellular truncation.

Mol	Chain	Residues	Atoms			Trace			
1	Λ	20	Total	С	Н	N	О	S	0
1	A	30	506	164	269	38	33	2	U

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1731	LYS	-	expression tag	UNP P46531
A	1732	LYS	-	expression tag	UNP P46531
A	1733	LYS	-	expression tag	UNP P46531
A	1758	LYS	ARG	conflict	UNP P46531
A	1760	LYS	ARG	conflict	UNP P46531

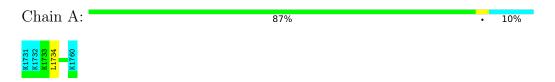


## 4 Residue-property plots (i)

### 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: Notch 1 extracellular truncation



#### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

#### 4.2.1 Score per residue for model 1

• Molecule 1: Notch 1 extracellular truncation

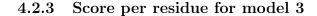
Chain A: 87% . 10%

#### 4.2.2 Score per residue for model 2

• Molecule 1: Notch 1 extracellular truncation

Chain A: 73% 17% 10%





• Molecule 1: Notch 1 extracellular truncation

Chain A: • 10%



#### 4.2.4 Score per residue for model 4

• Molecule 1: Notch 1 extracellular truncation

Chain A: 80% 10% 10%



#### 4.2.5 Score per residue for model 5

• Molecule 1: Notch 1 extracellular truncation

Chain A: 87% • 10%



#### 4.2.6 Score per residue for model 6

• Molecule 1: Notch 1 extracellular truncation

Chain A: 87% • 10%



#### 4.2.7 Score per residue for model 7

• Molecule 1: Notch 1 extracellular truncation

Chain A: 80% 7% • 10%





#### 4.2.8 Score per residue for model 8

• Molecule 1: Notch 1 extracellular truncation

Chain A: 83% 7% 10%



#### 4.2.9 Score per residue for model 9

• Molecule 1: Notch 1 extracellular truncation

Chain A: 83% . . 10%



#### 4.2.10 Score per residue for model 10

• Molecule 1: Notch 1 extracellular truncation

Chain A: 87% • 10%



#### 4.2.11 Score per residue for model 11

• Molecule 1: Notch 1 extracellular truncation

Chain A: 80% 7% · 10%



#### 4.2.12 Score per residue for model 12

• Molecule 1: Notch 1 extracellular truncation

Chain A: 80% 10% 10%

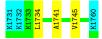




#### 4.2.13 Score per residue for model 13

• Molecule 1: Notch 1 extracellular truncation

Chain A: 80% 10% 10%



#### 4.2.14 Score per residue for model 14

• Molecule 1: Notch 1 extracellular truncation

Chain A: 83% . . 10%



#### 4.2.15 Score per residue for model 15

• Molecule 1: Notch 1 extracellular truncation

Chain A: 80% 10% 10%



#### 4.2.16 Score per residue for model 16

• Molecule 1: Notch 1 extracellular truncation

Chain A: 77% 13% 10%



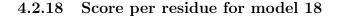
#### 4.2.17 Score per residue for model 17

• Molecule 1: Notch 1 extracellular truncation

Chain A: 83% 7% 10%







• Molecule 1: Notch 1 extracellular truncation

Chain A: 83% . . 10%



#### 4.2.19 Score per residue for model 19

• Molecule 1: Notch 1 extracellular truncation

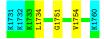
Chain A: 73% 17% 10%



#### 4.2.20 Score per residue for model 20

• Molecule 1: Notch 1 extracellular truncation

Chain A: 80% 10% 10%



#### 4.2.21 Score per residue for model 21

• Molecule 1: Notch 1 extracellular truncation

Chain A: 80% 7% · 10%



#### 4.2.22 Score per residue for model 22

• Molecule 1: Notch 1 extracellular truncation

Chain A: 87% • 10%





#### 4.2.23 Score per residue for model 23 (medoid)

• Molecule 1: Notch 1 extracellular truncation

Chain A: 73% 17% 10%



#### 4.2.24 Score per residue for model 24

• Molecule 1: Notch 1 extracellular truncation

Chain A: 80% 10% 10%



#### 4.2.25 Score per residue for model 25

• Molecule 1: Notch 1 extracellular truncation

Chain A: 80% 7% · 10%



#### 4.2.26 Score per residue for model 26

• Molecule 1: Notch 1 extracellular truncation

Chain A: 80% 7% · 10%



#### 4.2.27 Score per residue for model 27

• Molecule 1: Notch 1 extracellular truncation

Chain A: 87% • 10%





#### 4.2.28 Score per residue for model 28

• Molecule 1: Notch 1 extracellular truncation

Chain A: 83% . . 10%



#### 4.2.29 Score per residue for model 29

• Molecule 1: Notch 1 extracellular truncation

Chain A: 73% 17% 10%



#### 4.2.30 Score per residue for model 30

• Molecule 1: Notch 1 extracellular truncation

Chain A: 87% • 10%



#### 4.2.31 Score per residue for model 31

• Molecule 1: Notch 1 extracellular truncation

Chain A: 87% • 10%



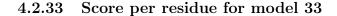
#### 4.2.32 Score per residue for model 32

• Molecule 1: Notch 1 extracellular truncation

Chain A: 83% 7% 10%







• Molecule 1: Notch 1 extracellular truncation

Chain A: • 10%



#### 4.2.34 Score per residue for model 34

• Molecule 1: Notch 1 extracellular truncation

Chain A: 80% 7% • 10%



#### 4.2.35 Score per residue for model 35

• Molecule 1: Notch 1 extracellular truncation

Chain A: 87% • 10%



#### 4.2.36 Score per residue for model 36

• Molecule 1: Notch 1 extracellular truncation

Chain A: 80% 7% · 10%



#### 4.2.37 Score per residue for model 37

• Molecule 1: Notch 1 extracellular truncation

Chain A: 83% 7% 10%





### 4.2.38 Score per residue for model 38

• Molecule 1: Notch 1 extracellular truncation

Chain A: 83% 7% 10%



#### 4.2.39 Score per residue for model 39

• Molecule 1: Notch 1 extracellular truncation

Chain A: 83% . . 10%



#### 4.2.40 Score per residue for model 40

• Molecule 1: Notch 1 extracellular truncation

Chain A: 87% • 10%





## 5 Refinement protocol and experimental data overview (i)

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

Of the 400 calculated structures, 40 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	1.2.1
ARIA	structure calculation	2.3.2

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	350
Number of shifts mapped to atoms	350
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	69%



## 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	209	228	227	2±1
All	All	8360	9120	9080	69

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 4.

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

Atom-1	Atom-2	Clash(Å)	$\operatorname{Distance}(\mathring{\mathrm{A}})$	Models	
Atom-1	Atom-2	Clash(A)	Distance(A)	Worst	Total
1:A:1734:LEU:O	1:A:1734:LEU:HD13	0.63	1.94	40	40
1:A:1741:ALA:O	1:A:1745:VAL:HG13	0.47	2.10	34	9
1:A:1733:LYS:HZ2	1:A:1737:MET:HG2	0.46	1.69	7	1
1:A:1754:VAL:HG13	1:A:1755:LEU:HD22	0.43	1.89	24	1
1:A:1734:LEU:O	1:A:1738:TYR:HB3	0.42	2.15	17	10
1:A:1736:PHE:O	1:A:1739:VAL:HG12	0.42	2.14	36	1
1:A:1751:GLY:O	1:A:1754:VAL:HG22	0.42	2.14	20	1
1:A:1734:LEU:HD22	1:A:1738:TYR:HB2	0.41	1.92	17	1
1:A:1744:PHE:O	1:A:1747:LEU:HD23	0.41	2.15	2	3
1:A:1733:LYS:HD2	1:A:1737:MET:HG3	0.41	1.92	19	1
1:A:1742:ALA:O	1:A:1746:LEU:HG	0.41	2.16	25	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 Outli		Outliers	Perce	ntiles	
1	A	27/30 (90%)	24±1 (89±3%)	3±1 (11±3%)	0±0 (0±1%)	50	82
All	All	1080/1200 (90%)	959 (89%)	119 (11%)	2 (0%)	50	82

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
1	A	1759	LYS	2

#### 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	Perce	ntiles
1	A	21/24 (88%)	20±1 (97±3%)	1±1 (3±3%)	49	91
All	All	840/960 (88%)	818 (97%)	22 (3%)	49	91

All 6 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	1734	LEU	13
1	A	1733	LYS	2
1	A	1756	LEU	2
1	A	1759	LYS	2
1	A	1737	MET	2
1	A	1736	PHE	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 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 (i)

The completeness of assignment taking into account all chemical shift lists is 69% for the well-defined parts and 69% for the entire structure.

### 7.1 Chemical shift list 1

File name: working cs.cif

Chemical shift list name: Notch1WT.str

### 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	350
Number of shifts mapped to atoms	350
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

The following errors were found when reading this chemical shift list.

• Chemical shift has been reported more than once. All 44 occurrences are reported below.

T:a4 ID	Chain	Chain Res	Type Atom	A 4	Shift Data		
List ID				Atom	Value	Uncertainty	Ambiguity
1	A	1734	LEU	HD12	0.953	•	
1	A	1734	LEU	HD13	0.953	•	
1	A	1734	LEU	HD22	0.913		
1	A	1734	LEU	HD23	0.913	•	
1	A	1739	VAL	HG12	0.960	•	
1	A	1739	VAL	HG13	0.960	•	•
1	A	1739	VAL	HG22	0.930	•	
1	A	1739	VAL	HG23	0.930	•	•
1	A	1740	ALA	HB2	1.316	•	
1	A	1740	ALA	HB3	1.316	•	
1	A	1741	ALA	HB2	1.475	•	
1	A	1741	ALA	HB3	1.475	•	
1	A	1742	ALA	HB2	1.422	•	
1	A	1742	ALA	HB3	1.422	•	
1	A	1743	ALA	HB2	1.456		

Continued on next page...



Continued from previous page...

				A 4		Shift Data		
List ID	Chain	Res	Type	Atom	Value	Uncertainty	Ambiguity	
1	A	1743	ALA	HB3	1.456			
1	A	1745	VAL	HG12	1.153	•	•	
1	A	1745	VAL	HG13	1.153	•	•	
1	A	1745	VAL	HG22	1.033	•	•	
1	A	1745	VAL	HG23	1.033	•	•	
1	A	1746	LEU	HD12	0.946	•		
1	A	1746	LEU	HD13	0.946	•	•	
1	A	1746	LEU	HD22	0.922	•	•	
1	A	1746	LEU	HD23	0.922	•		
1	A	1747	LEU	HD12	0.850	•	•	
1	A	1747	LEU	HD13	0.850	•		
1	A	1747	LEU	HD22	0.879	•	•	
1	A	1747	LEU	HD23	0.879	•	•	
1	A	1750	VAL	HG12	1.121			
1	A	1750	VAL	HG13	1.121	•	•	
1	A	1750	VAL	HG22	0.938	•	•	
1	A	1750	VAL	HG23	0.938	•	•	
1	A	1754	VAL	HG12	1.032	•	•	
1	A	1754	VAL	HG13	1.032	•	•	
1	A	1754	VAL	HG22	0.910	•		
1	A	1754	VAL	HG23	0.910	•	•	
1	A	1755	LEU	HD12	0.869	•		
1	A	1755	LEU	HD13	0.869	•	•	
1	A	1755	LEU	HD22	0.905	•	•	
1	A	1755	LEU	HD23	0.905	•		
1	A	1756	LEU	HD12	0.878			
1	A	1756	LEU	HD13	0.878	•	•	
1	A	1756	LEU	HD22	0.887	•	•	
1	A	1756	LEU	HD23	0.887	•	•	

## 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
$^{13}\mathrm{C}_{\alpha}$	30	$0.12 \pm 0.21$	None needed ( $< 0.5 \text{ ppm}$ )
$^{13}C_{\beta}$	28	$0.89 \pm 0.07$	Should be checked
<sup>13</sup> C′	0		None (insufficient data)
$^{15}N$	0	_	None (insufficient data)



#### 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 69%, i.e. 269 atoms were assigned a chemical shift out of a possible 388. 0 out of 9 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^{1}{ m H}$	$^{13}\mathbf{C}$	$^{15}{ m N}$
Backbone	83/137 (61%)	56/56~(100%)	27/54~(50%)	0/27~(0%)
Sidechain	167/194 (86%)	112/132 (85%)	55/59 (93%)	0/3 (0%)
Aromatic	19/57 (33%)	19/28~(68%)	0/27~(0%)	0/2~(0%)
Overall	$269/388 \ (69\%)$	187/216 (87%)	82/140 (59%)	0/32 (0%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 69%, i.e. 305 atoms were assigned a chemical shift out of a possible 442. 0 out of 9 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^{1}\mathrm{H}$	$^{13}\mathbf{C}$	$^{15}{ m N}$
Backbone	91/152~(60%)	61/62~(98%)	30/60 (50%)	0/30 (0%)
Sidechain	195/233~(84%)	128/156~(82%)	67/71 (94%)	0/6 (0%)
Aromatic	19/57 (33%)	19/28 (68%)	0/27 (0%)	0/2 (0%)
Overall	305/442~(69%)	208/246~(85%)	97/158 (61%)	0/38 (0%)

### 7.1.4 Statistically unusual chemical shifts (i)

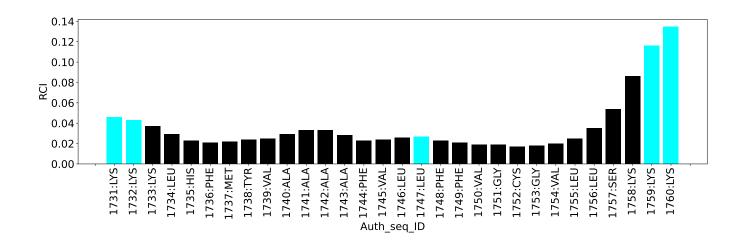
There are no statistically unusual chemical shifts.

## 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:







# 8 NMR restraints analysis (i)

No restraints data found



# 9 Distance violation analysis (i)

No distance restraints data found



# 10 Dihedral-angle violation analysis (i)

No dihedral-angle restraints found

