# WORLDWIDE PDB PROTEIN DATA BANK

# Full wwPDB NMR Structure Validation Report (i)

Aug 8, 2023 – 01:21 pm BST

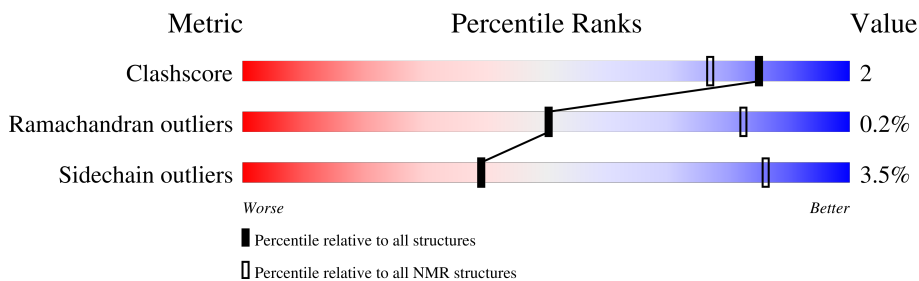|   |   |   |
|---|---|---|
| PDB ID | : | 8ORZ |
| BMRB ID | : | 34807 |
| Title | : | Solution NMR structure of Notch1 G1740-1743 TMD |
| Authors | : | Guschtschin-Schmidt, N.; Muhle-Goll, C. |
| Deposited on | : | 2023-04-17 |

# 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 74%.

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 | Percentile Ranks | Value |
|---|---|---|
| Clashscore | | 2 |
| Ramachandran outliers | | 0.2% |
| Sidechain outliers | | 3.5% |

*Worse*                                        *Better*

▮ Percentile relative to all structures
▯ Percentile relative to all NMR structures

| 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 >=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 | 63%     7%     30% |

# 2 Ensemble composition and analysis ⓘ

This entry contains 40 models. Model 9 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:1740 (8) | 0.39 | 18 |
| 2 | A:1747-A:1759 (13) | 0.77 | 9 |

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 6 clusters. No single-model clusters were found.

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

# 3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 494 atoms, of which 261 are hydrogens and 0 are deuteriums.

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

| Mol | Chain | Residues | Atoms | | | | | | Trace |
|-----|-------|----------|-------|---|---|---|---|---|-------|
| | | | Total | C | H | N | O | S | |
| 1 | A | 30 | 494 | 160 | 261 | 38 | 33 | 2 | 0 |

There are 10 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 | 1740 | GLY | ALA | engineered mutation | UNP P46531 |
| A | 1741 | GLY | ALA | engineered mutation | UNP P46531 |
| A | 1742 | GLY | ALA | engineered mutation | UNP P46531 |
| A | 1743 | GLY | ALA | engineered mutation | UNP P46531 |
| A | 1758 | LYS | - | expression tag | UNP P46531 |
| A | 1759 | LYS | - | expression tag | UNP P46531 |
| A | 1760 | LYS | - | expression tag | 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.

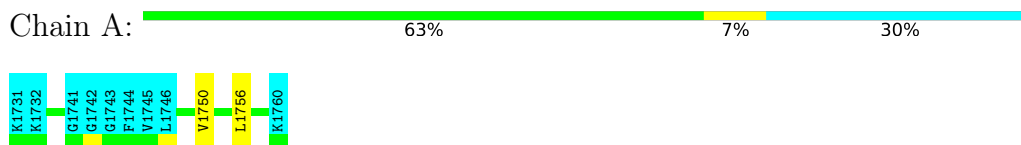• Molecule 1: Notch 1 extracellular truncation
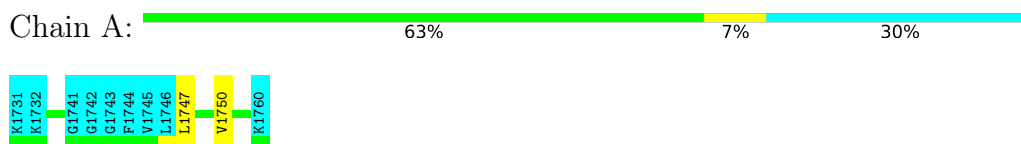


## 4.2 Scores per residue for each member of the ensemble
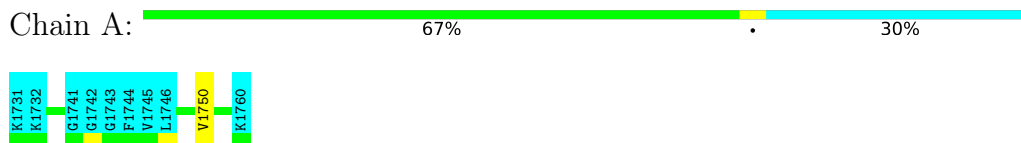
Colouring as in section 4.1 above.

### 4.2.1 Score per residue for model 1

• Molecule 1: Notch 1 extracellular truncation



### 4.2.2 Score per residue for model 2

• Molecule 1: Notch 1 extracellular truncation

### 4.2.3 Score per residue for model 3

• Molecule 1: Notch 1 extracellular truncation

Chain A:

### 4.2.4 Score per residue for model 4

• Molecule 1: Notch 1 extracellular truncation

Chain A:

### 4.2.5 Score per residue for model 5

• Molecule 1: Notch 1 extracellular truncation

Chain A:

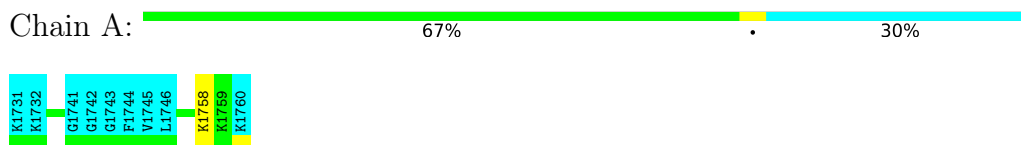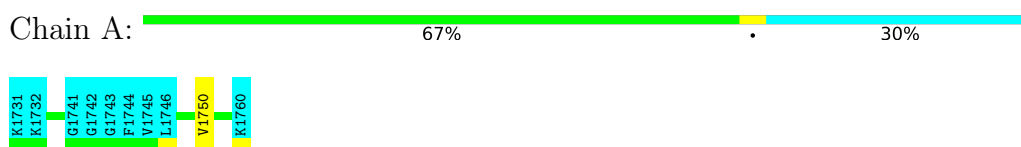### 4.2.6 Score per residue for model 6

• Molecule 1: Notch 1 extracellular truncation

Chain A:

### 4.2.7 Score per residue for model 7

• Molecule 1: Notch 1 extracellular truncation

Chain A:

#### 4.2.8 Score per residue for model 8

• Molecule 1: Notch 1 extracellular truncation

Chain A:

70%          30%

K1731 K1732 G1741 G1742 G1743 F1744 V1745 L1746 K1760

#### 4.2.9 Score per residue for model 9 (medoid)

• Molecule 1: Notch 1 extracellular truncation

Chain A:

63%          7%          30%

K1731 K1732 G1741 G1742 G1743 F1744 V1745 L1746 L1747 V1750 K1760

#### 4.2.10 Score per residue for model 10

• Molecule 1: Notch 1 extracellular truncation

Chain A:

63%          7%          30%

K1731 K1732 G1741 G1742 G1743 F1744 V1745 L1746 L1747 V1750 K1760

#### 4.2.11 Score per residue for model 11

• Molecule 1: Notch 1 extracellular truncation

Chain A:

60%          10%          30%

K1731 K1732 G1741 G1742 G1743 F1744 V1745 L1746 L1747 V1750 K1758 K1759 K1760

#### 4.2.12 Score per residue for model 12

• Molecule 1: Notch 1 extracellular truncation

Chain A:

63%          7%          30%

K1731 K1732 G1741 G1742 G1743 F1744 V1745 L1746 L1747 V1750 K1760

### 4.2.13  Score per residue for model 13

• Molecule 1: Notch 1 extracellular truncation
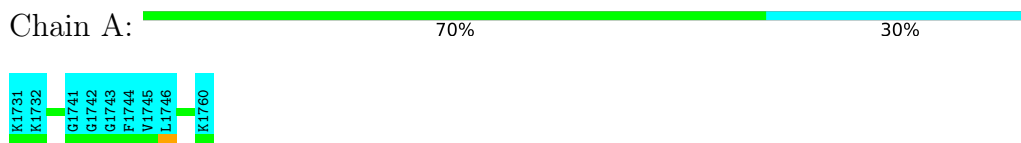
Chain A:



### 4.2.14  Score per residue for model 14

• Molecule 1: Notch 1 extracellular truncation

Chain A:



### 4.2.15  Score per residue for model 15
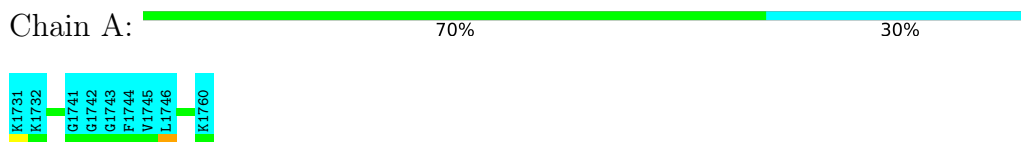
• Molecule 1: Notch 1 extracellular truncation

Chain A:



### 4.2.16  Score per residue for model 16
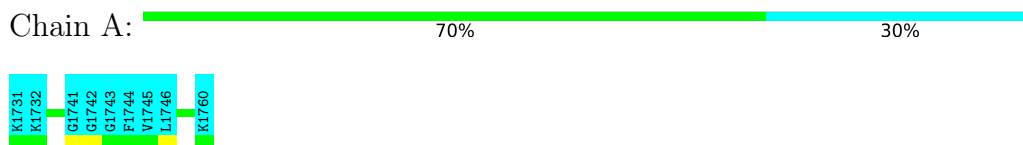
• Molecule 1: Notch 1 extracellular truncation

Chain A:



### 4.2.17  Score per residue for model 17

• Molecule 1: Notch 1 extracellular truncation

Chain A:

#### 4.2.18 Score per residue for model 18

• Molecule 1: Notch 1 extracellular truncation

Chain A:



#### 4.2.19 Score per residue for model 19

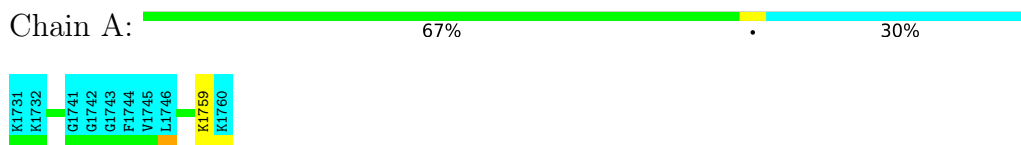• Molecule 1: Notch 1 extracellular truncation

Chain A:



#### 4.2.20 Score per residue for model 20

• Molecule 1: Notch 1 extracellular truncation

Chain A:



#### 4.2.21 Score per residue for model 21

• Molecule 1: Notch 1 extracellular truncation

Chain A:



#### 4.2.22 Score per residue for model 22

• Molecule 1: Notch 1 extracellular truncation

Chain A:

#### 4.2.23 Score per residue for model 23

• Molecule 1: Notch 1 extracellular truncation

Chain A:



#### 4.2.24 Score per residue for model 24

• Molecule 1: Notch 1 extracellular truncation

Chain A:



#### 4.2.25 Score per residue for model 25

• Molecule 1: Notch 1 extracellular truncation

Chain A:



#### 4.2.26 Score per residue for model 26

• Molecule 1: Notch 1 extracellular truncation

Chain A:



#### 4.2.27 Score per residue for model 27

• Molecule 1: Notch 1 extracellular truncation

Chain A:

#### 4.2.28 Score per residue for model 28

• Molecule 1: Notch 1 extracellular truncation

Chain A:

#### 4.2.29 Score per residue for model 29

• Molecule 1: Notch 1 extracellular truncation

Chain A:

#### 4.2.30 Score per residue for model 30

• Molecule 1: Notch 1 extracellular truncation

Chain A:

#### 4.2.31 Score per residue for model 31

• Molecule 1: Notch 1 extracellular truncation
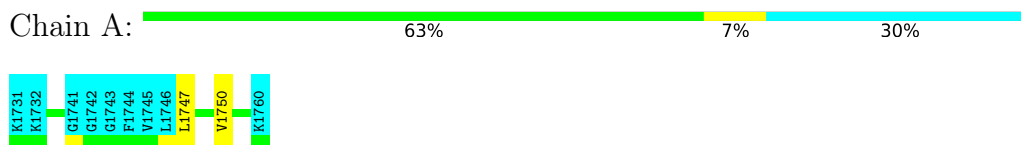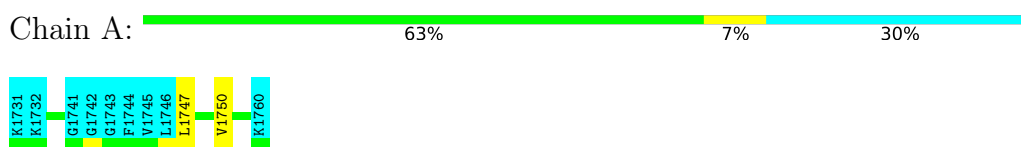
Chain A:

#### 4.2.32 Score per residue for model 32

• Molecule 1: Notch 1 extracellular truncation

Chain A:

### 4.2.33 Score per residue for model 33
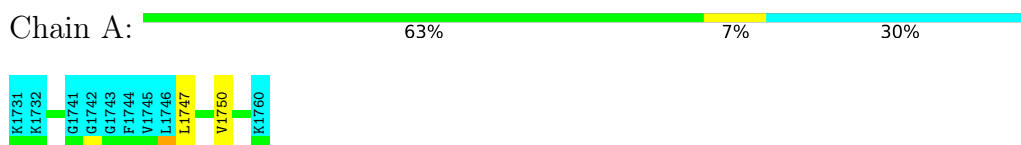
• Molecule 1: Notch 1 extracellular truncation

Chain A: 63% 7% 30%

K1731 K1732 G1741 G1742 G1743 F1744 V1745 L1746 L1747 V1750 K1760

### 4.2.34 Score per residue for model 34

• Molecule 1: Notch 1 extracellular truncation

Chain A: 63% 7% 30%

K1731 K1732 G1741 G1742 G1743 F1744 V1745 L1746 L1747 V1750 K1760

### 4.2.35 Score per residue for model 35

• Molecule 1: Notch 1 extracellular truncation

Chain A: 60% 10% 30%

K1731 K1732 G1741 G1742 G1743 F1744 V1745 L1746 L1747 V1750 L1756 K1760

### 4.2.36 Score per residue for model 36

• Molecule 1: Notch 1 extracellular truncation

Chain A: 67% • 30%

K1731 K1732 K1733 L1734 G1741 G1742 G1743 F1744 V1745 L1746 K1760

### 4.2.37 Score per residue for model 37

• Molecule 1: Notch 1 extracellular truncation

Chain A: 60% 10% 30%

K1731 K1732 G1741 G1742 G1743 F1744 V1745 L1746 L1747 V1750 K1758 K1759 K1760

#### 4.2.38 Score per residue for model 38

• Molecule 1: Notch 1 extracellular truncation

Chain A:



#### 4.2.39 Score per residue for model 39

• Molecule 1: Notch 1 extracellular truncation

Chain A:



#### 4.2.40 Score per residue for model 40
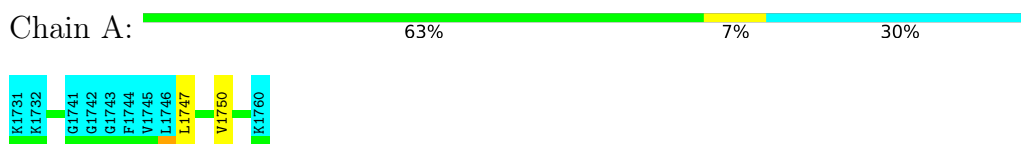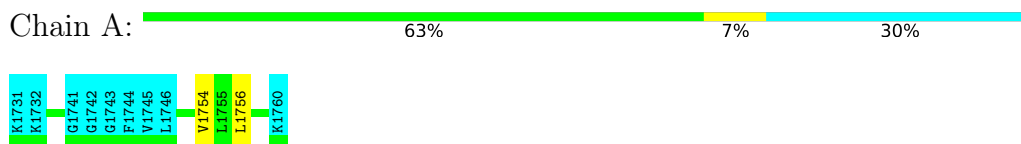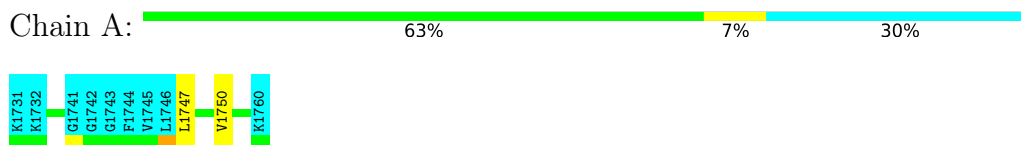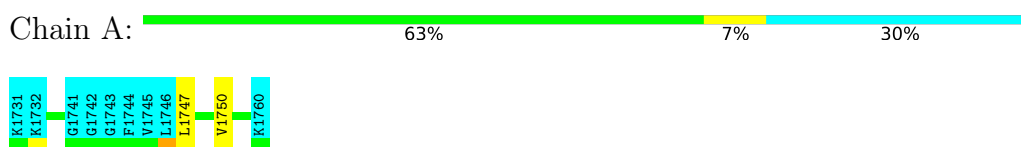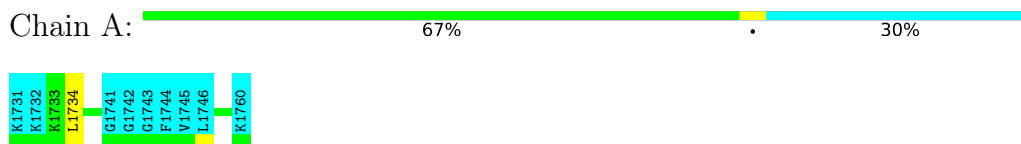
• Molecule 1: Notch 1 extracellular truncation

Chain A:
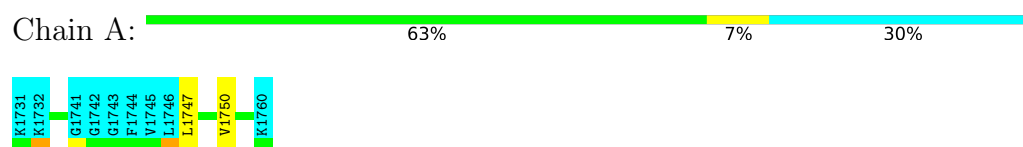
# 5 Refinement protocol and experimental data overview ⓘ

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 | 347 |
| Number of shifts mapped to atoms | 347 |
| Number of unparsed shifts | 0 |
| Number of shifts with mapping errors | 0 |
| Number of shifts with mapping warnings | 0 |
| Assignment completeness (well-defined parts) | 74% |

# 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     | 167   | 182      | 181      | 1±1     |
| All | All   | 6680  | 7280     | 7240     | 31      |

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 2.

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

| Atom-1 | Atom-2 | Clash(Å) | Distance(Å) | Models Worst | Total |
|--------|--------|----------|-------------|--------------|-------|
| 1:A:1747:LEU:O | 1:A:1750:VAL:HG22 | 0.56 | 2.00 | 24 | 25 |
| 1:A:1755:LEU:HD13 | 1:A:1756:LEU:N | 0.47 | 2.25 | 28 | 1 |
| 1:A:1754:VAL:HG23 | 1:A:1755:LEU:N | 0.46 | 2.25 | 28 | 1 |
| 1:A:1755:LEU:HD22 | 1:A:1755:LEU:C | 0.45 | 2.31 | 28 | 1 |
| 1:A:1734:LEU:C | 1:A:1734:LEU:HD22 | 0.43 | 2.34 | 20 | 1 |
| 1:A:1747:LEU:HA | 1:A:1750:VAL:HG12 | 0.41 | 1.91 | 38 | 1 |
| 1:A:1734:LEU:CD1 | 1:A:1734:LEU:H | 0.40 | 2.30 | 20 | 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 | 21/30 (70%) | 21±1 (98±3%) | 0±1 (2±3%) | 0±0 (0±1%) | 50 | 82 |
| All | All | 840/1200 (70%) | 821 (98%) | 17 (2%) | 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 ⓘ

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 | 18/24 (75%) | 17±1 (97±4%) | 1±1 (3±4%) | 39 | 86 |
| All | All | 720/960 (75%) | 695 (97%) | 25 (3%) | 39 | 86 |

All 8 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 | 1756 | LEU | 9 |
| 1 | A | 1734 | LEU | 5 |
| 1 | A | 1758 | LYS | 4 |
| 1 | A | 1737 | MET | 3 |
| 1 | A | 1759 | LYS | 1 |
| 1 | A | 1755 | LEU | 1 |
| 1 | A | 1754 | VAL | 1 |
| 1 | A | 1736 | PHE | 1 |

### 6.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 6.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 6.6 Ligand geometry ⓘ

There are no ligands in this entry.

## 6.7 Other polymers ⓘ

There are no such molecules in this entry.

## 6.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

# 7 Chemical shift validation ⓘ

The completeness of assignment taking into account all chemical shift lists is 74% for the well-defined parts and 73% for the entire structure.

## 7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *Notch1G.str*

### 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 | 347 |
| Number of shifts mapped to atoms | 347 |
| 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 30 occurrences are reported below.

| List ID | Chain | Res | Type | Atom | Shift Data | | |
|---|---|---|---|---|---|---|---|
| | | | | | Value | Uncertainty | Ambiguity |
| 1 | A | 1734 | LEU | HD12 | 0.916 | . | . |
| 1 | A | 1734 | LEU | HD13 | 0.916 | . | . |
| 1 | A | 1734 | LEU | HD22 | 0.958 | . | . |
| 1 | A | 1734 | LEU | HD23 | 0.958 | . | . |
| 1 | A | 1739 | VAL | HG12 | 0.916 | . | . |
| 1 | A | 1739 | VAL | HG13 | 0.916 | . | . |
| 1 | A | 1745 | VAL | HG12 | 1.033 | . | . |
| 1 | A | 1745 | VAL | HG13 | 1.033 | . | . |
| 1 | A | 1746 | LEU | HD12 | 0.933 | . | . |
| 1 | A | 1746 | LEU | HD13 | 0.933 | . | . |
| 1 | A | 1746 | LEU | HD22 | 0.990 | . | . |
| 1 | A | 1746 | LEU | HD23 | 0.990 | . | . |
| 1 | A | 1747 | LEU | HD12 | 0.921 | . | . |
| 1 | A | 1747 | LEU | HD13 | 0.921 | . | . |
| 1 | A | 1747 | LEU | HD22 | 0.885 | . | . |

*Continued on next page...*

*Continued from previous page...*

| List ID | Chain | Res | Type | Atom | Shift Data | | |
|---------|-------|-----|------|------|-------|-------------|-----------|
| | | | | | Value | Uncertainty | Ambiguity |
| 1 | A | 1747 | LEU | HD23 | 0.885 | . | . |
| 1 | A | 1750 | VAL | HG12 | 1.135 | . | . |
| 1 | A | 1750 | VAL | HG13 | 1.135 | . | . |
| 1 | A | 1750 | VAL | HG22 | 0.951 | . | . |
| 1 | A | 1750 | VAL | HG23 | 0.951 | . | . |
| 1 | A | 1754 | VAL | HG12 | 1.053 | . | . |
| 1 | A | 1754 | VAL | HG13 | 1.053 | . | . |
| 1 | A | 1754 | VAL | HG22 | 0.921 | . | . |
| 1 | A | 1754 | VAL | HG23 | 0.921 | . | . |
| 1 | A | 1755 | LEU | HD12 | 0.886 | . | . |
| 1 | A | 1755 | LEU | HD13 | 0.886 | . | . |
| 1 | A | 1755 | LEU | HD22 | 0.917 | . | . |
| 1 | A | 1755 | LEU | HD23 | 0.917 | . | . |
| 1 | A | 1756 | LEU | HD12 | 0.889 | . | . |
| 1 | A | 1756 | LEU | HD13 | 0.889 | . | . |

### 7.1.2   Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

| Nucleus | # values | Correction ± precision, *ppm* | Suggested action |
|---------|----------|-------------------------------|------------------|
| $^{13}C_\alpha$ | 30 | 0.02 ± 0.29 | None needed (< 0.5 ppm) |
| $^{13}C_\beta$ | 24 | — | None (insufficient data) |
| $^{13}C'$ | 0 | — | None (insufficient data) |
| $^{15}N$ | 30 | 0.87 ± 0.71 | None needed (imprecise) |

### 7.1.3   Completeness of resonance assignments ⓘ

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 74%, i.e. 228 atoms were assigned a chemical shift out of a possible 307. 0 out of 7 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

| | Total | $^1$H | $^{13}$C | $^{15}$N |
|---|-------|-------|----------|----------|
| Backbone | 87/108 (81%) | 45/45 (100%) | 21/42 (50%) | 21/21 (100%) |
| Sidechain | 123/152 (81%) | 80/102 (78%) | 43/47 (91%) | 0/3 (0%) |
| Aromatic | 18/47 (38%) | 13/23 (57%) | 5/22 (23%) | 0/2 (0%) |
| Overall | 228/307 (74%) | 138/170 (81%) | 69/111 (62%) | 21/26 (81%) |

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 73%, i.e. 316 atoms were assigned a chemical shift out of a possible

430. 0 out of 9 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

|  | Total | $^1$H | $^{13}$C | $^{15}$N |
|---|---|---|---|---|
| Backbone | 123/156 (79%) | 63/66 (95%) | 30/60 (50%) | 30/30 (100%) |
| Sidechain | 170/217 (78%) | 109/144 (76%) | 61/67 (91%) | 0/6 (0%) |
| Aromatic | 23/57 (40%) | 17/28 (61%) | 6/27 (22%) | 0/2 (0%) |
| Overall | 316/430 (73%) | 189/238 (79%) | 97/154 (63%) | 30/38 (79%) |

### 7.1.4 Statistically unusual chemical shifts ⓘ

There are no statistically unusual chemical shifts.

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

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:
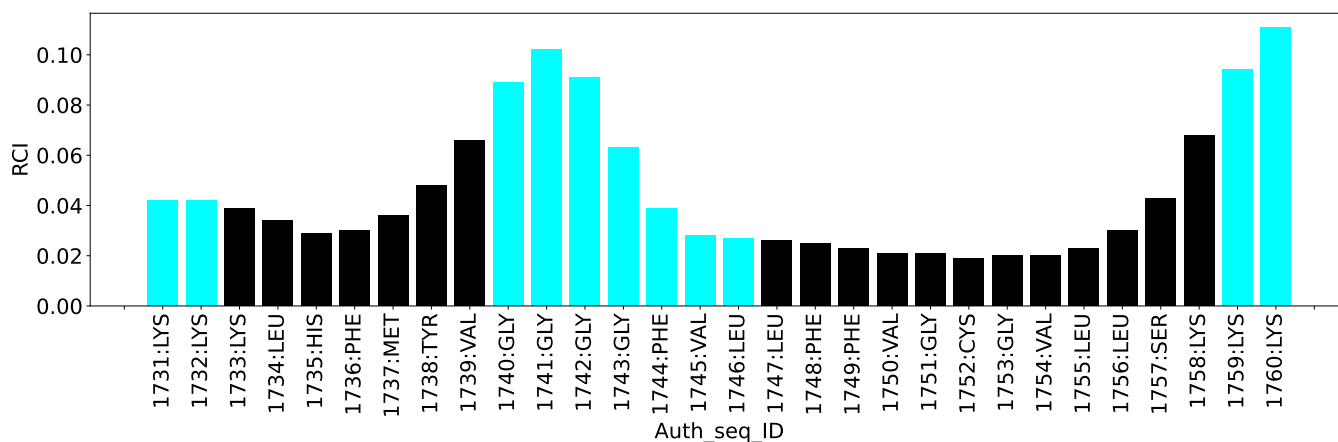
# 8 NMR restraints analysis (i)

No restraints data found

# 9 Distance violation analysis ⓘ

No distance restraints data found

# 10    Dihedral-angle violation analysis ⓘ

No dihedral-angle restraints found