



# wwPDB NMR Structure Validation Summary Report ⓘ

Mar 5, 2026 – 04:29 AM UTC

PDB ID : 5VH7 / pdb\_00005vh7  
BMRB ID : 30282  
Title : Structure and dynamics of RNA repeat expansions that cause Huntington's Disease and myotonic dystrophy type 1  
Authors : Chen, J.L.; VanEtten, D.M.; Fountain, M.A.; Yildirim, I.; Disney, M.D.  
Deposited on : 2017-04-12

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

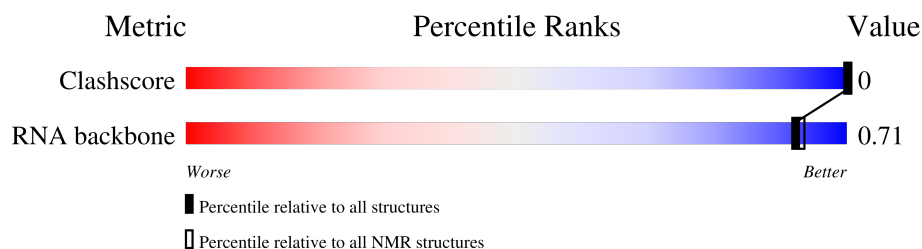
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

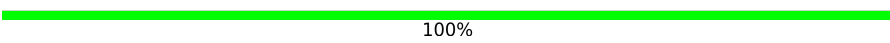
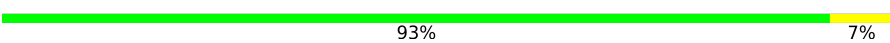
The overall completeness of chemical shifts assignment is 15%.

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<br>(#Entries) | NMR archive<br>(#Entries) |
|--------------|-----------------------------|---------------------------|
| Clashscore   | 229148                      | 14424                     |
| RNA backbone | 8273                        | 777                       |

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     | 15     |  100%   |
| 1   | B     | 15     |  93% 7% |

## 2 Ensemble composition and analysis ⓘ

This entry contains 20 models. This entry does not contain polypeptide chains, therefore identification of well-defined residues and clustering analysis are not possible. All residues are included in the validation scores.

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

There is only 1 type of molecule in this entry. The entry contains 972 atoms, of which 332 are hydrogens and 0 are deuteriums.

- Molecule 1 is a RNA chain called RNA (5'-R(\*GP\*AP\*CP\*CP\*AP\*GP\*CP\*AP\*G)-3').

| Mol | Chain | Residues | Atoms |     |     |    |     |    | Trace |
|-----|-------|----------|-------|-----|-----|----|-----|----|-------|
| 1   | A     | 15       | Total | C   | H   | N  | O   | P  | 0     |
|     |       |          | 486   | 144 | 166 | 62 | 100 | 14 |       |
| 1   | B     | 15       | Total | C   | H   | N  | O   | P  | 0     |
|     |       |          | 486   | 144 | 166 | 62 | 100 | 14 |       |

## 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: RNA (5'-R(\*GP\*AP\*CP\*CP\*AP\*GP\*CP\*AP\*G)-3')

Chain A:  100%

There are no outlier residues in this chain.

- Molecule 1: RNA (5'-R(\*GP\*AP\*CP\*CP\*AP\*GP\*CP\*AP\*G)-3')

Chain B:  93% 7%



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

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

- Molecule 1: RNA (5'-R(\*GP\*AP\*CP\*CP\*AP\*GP\*CP\*AP\*G)-3')

Chain A:  100%

There are no outlier residues in this chain.

- Molecule 1: RNA (5'-R(\*GP\*AP\*CP\*CP\*AP\*GP\*CP\*AP\*G)-3')

Chain B:  93% 7%



## 5 Refinement protocol and experimental data overview

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

Of the 40 calculated structures, 20 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 |
|---------------|-----------------------|---------|
| Amber         | refinement            | 16      |
| Amber         | structure calculation |         |

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               | 2              |
| Total number of shifts                       | 108            |
| Number of shifts mapped to atoms             | 108            |
| Number of unparsed shifts                    | 0              |
| Number of shifts with mapping errors         | 0              |
| Number of shifts with mapping warnings       | 0              |
| Assignment completeness (well-defined parts) | 15%            |

## 6 Model quality [i](#)

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

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

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   | 0.6±0.6   |
| 1   | B     | 0.0±0.0   | 0.8±1.0   |
| All | All   | 0         | 27        |

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

5 of 10 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   | B     | 21  | G    | Sidechain | 7              |
| 1   | A     | 9   | G    | Sidechain | 6              |
| 1   | A     | 13  | G    | Sidechain | 3              |
| 1   | B     | 19  | C    | Sidechain | 3              |
| 1   | B     | 24  | G    | Sidechain | 2              |

### 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 |
|-----|-------|-------|----------|----------|---------|
| All | All   | 12800 | 6640     | 6640     | -       |

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 -.

There are no clashes.

## 6.3 Torsion angles [i](#)

### 6.3.1 Protein backbone [i](#)

There are no protein molecules in this entry.

### 6.3.2 Protein sidechains [i](#)

There are no protein molecules in this entry.

### 6.3.3 RNA [i](#)

| Mol | Chain | Analysed      | Backbone Outliers | Pucker Outliers | Suiteness |
|-----|-------|---------------|-------------------|-----------------|-----------|
| 1   | A     | 14/15 (93%)   | 0±0 (0±0%)        | 0±0 (0±0%)      | 0.67±0.00 |
| 1   | B     | 14/15 (93%)   | 1±0 (7±0%)        | 0±0 (0±0%)      | 0.74±0.00 |
| All | All   | 560/600 (93%) | 20 (4%)           | 0 (0%)          | 0.70      |

The overall RNA backbone suiteness is 0.71.

All unique RNA backbone outliers are listed below:

| Mol | Chain | Res | Type | Models (Total) |
|-----|-------|-----|------|----------------|
| 1   | B     | 23  | A    | 20             |

There are no RNA pucker outliers to report.

## 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 ⓘ

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 15% for the well-defined parts and 15% for the entire structure.

### 7.1 Chemical shift list 1

File name: working\_cs.cif

Chemical shift list name: *CAG\_NMRSTAR3.1*

#### 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                  | 69 |
| Number of shifts mapped to atoms        | 69 |
| 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  |

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

No chemical shift referencing corrections were calculated (not enough 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 12%, i.e. 68 atoms were assigned a chemical shift out of a possible 576. 0 out of 0 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

|         | Total        | <sup>1</sup> H | <sup>13</sup> C | <sup>15</sup> N |
|---------|--------------|----------------|-----------------|-----------------|
| Sugar   | 44/330 (13%) | 44/180 (24%)   | 0/150 (0%)      | 0/0 (—%)        |
| Base    | 24/246 (10%) | 24/156 (15%)   | 0/50 (0%)       | 0/40 (0%)       |
| Overall | 68/576 (12%) | 68/336 (20%)   | 0/200 (0%)      | 0/40 (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](#)

No *random coil index*(RCI) plot could be generated from the current chemical shift list. RCI is only applicable to proteins

## 7.2 Chemical shift list 2

File name: working\_cs.cif

Chemical shift list name: CAG\_95H2O\_NMRSTAR3.1

### 7.2.1 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                  | 39 |
| Number of shifts mapped to atoms        | 39 |
| 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  |

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

No chemical shift referencing corrections were calculated (not enough data).

### 7.2.3 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 7%, i.e. 38 atoms were assigned a chemical shift out of a possible 576. 0 out of 0 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

|         | Total        | <sup>1</sup> H | <sup>13</sup> C | <sup>15</sup> N |
|---------|--------------|----------------|-----------------|-----------------|
| Sugar   | 9/330 (3%)   | 9/180 (5%)     | 0/150 (0%)      | 0/0 (—%)        |
| Base    | 29/246 (12%) | 29/156 (19%)   | 0/50 (0%)       | 0/40 (0%)       |
| Overall | 38/576 (7%)  | 38/336 (11%)   | 0/200 (0%)      | 0/40 (0%)       |

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

There are no statistically unusual chemical shifts.

### 7.2.5 Random Coil Index (RCI) plots

No *random coil index*(RCI) plot could be generated from the current chemical shift list. RCI is only applicable to proteins