



# Full wwPDB/EMDatabank EM Map/Model Validation Report ⓘ

Dec 9, 2019 – 02:29 AM EST

PDB ID : 6I2N  
EMDB ID: : EMD-0333  
Title : Helical RNA-bound Hantaan virus nucleocapsid  
Authors : Arragain, B.; Reguera, J.; Desfosses, A.; Gutsche, I.; Schoehn, G.; Malet, H.  
Deposited on : 2018-11-01  
Resolution : 3.30 Å(reported)  
Based on PDB ID : 5FSG

This is a Full wwPDB/EMDatabank EM Map/Model Validation Report  
for a publicly released PDB/EMDB 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/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

MolProbity : 4.02b-467  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.4

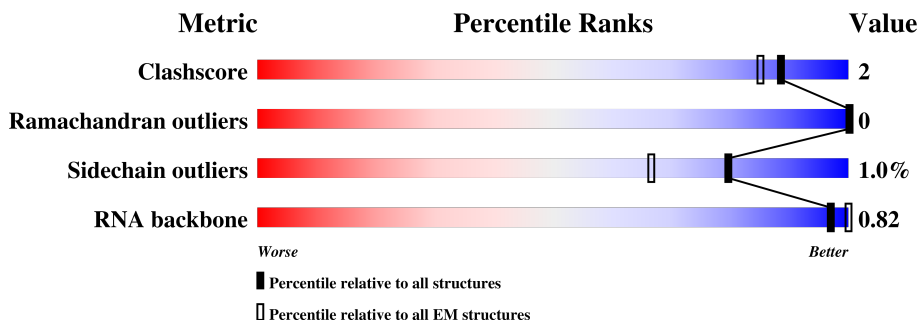
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.30 Å.

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) | EM structures<br>(#Entries) |
|-----------------------|-----------------------------|-----------------------------|
| Clashscore            | 136327                      | 1886                        |
| Ramachandran outliers | 132723                      | 1663                        |
| Sidechain outliers    | 132532                      | 1531                        |
| RNA backbone          | 3747                        | 458                         |

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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   | U     | 3      | 100%             |
| 2   | D     | 430    | 73% 5% 22%       |

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2710 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a RNA chain called RNA (5'-R(P\*UP\*UP\*U)-3').

| Mol | Chain | Residues | Atoms |    |   |    |   | AltConf | Trace |
|-----|-------|----------|-------|----|---|----|---|---------|-------|
| 1   | U     | 3        | Total | C  | N | O  | P | 0       | 0     |
|     |       |          | 60    | 27 | 6 | 24 | 3 |         |       |

- Molecule 2 is a protein called Nucleoprotein.

| Mol | Chain | Residues | Atoms |      |     |     |    | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|----|---------|-------|
| 2   | D     | 336      | Total | C    | N   | O   | S  | 0       | 0     |
|     |       |          | 2650  | 1674 | 461 | 498 | 17 |         |       |

There is a discrepancy between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment        | Reference  |
|-------|---------|----------|--------|----------------|------------|
| D     | 0       | GLY      | -      | expression tag | UNP P05133 |

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-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 outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

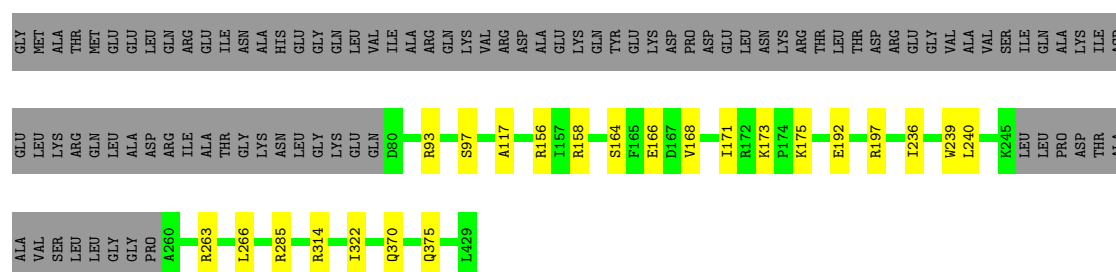
- Molecule 1: RNA (5'-R(P\*UP\*UP\*U)-3')

Chain U:  100%

There are no outlier residues recorded for this chain.

- Molecule 2: Nucleoprotein

Chain D:  73% 5% 22%



## 4 Experimental information

| Property                           | Value  | Source    |
|------------------------------------|--|-----------|
| Reconstruction method              | HELICAL  | Depositor |
| Imposed symmetry                   | HELICAL, twist=-99.95°, rise=18.87 Å, axial sym=C1 | Depositor |
| Number of segments used            | 105665   | Depositor |
| Resolution determination method    | FSC 0.143 CUT-OFF                                  | Depositor |
| CTF correction method              | PHASE FLIPPING AND AMPLITUDE CORRECTION            | Depositor |
| Microscope                         | FEI TITAN KRIOS                                    | Depositor |
| Voltage (kV)                       | 300  | Depositor |
| Electron dose ( $e^-/\text{Å}^2$ ) | 40   | Depositor |
| Minimum defocus (nm)               | 800  | Depositor |
| Maximum defocus (nm)               | 3500   | Depositor |
| Magnification                      | 46860  | Depositor |
| Image detector                     | GATAN K2 SUMMIT (4k x 4k)                          | Depositor |

## 5 Model quality [i](#)

### 5.1 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 root-mean-square of all Z scores of the bond lengths (or angles).

| Mol | Chain | Bond lengths |             | Bond angles |             |
|-----|-------|--------------|-------------|-------------|-------------|
|     |       | RMSZ         | $\# Z  > 2$ | RMSZ        | $\# Z  > 2$ |
| 1   | U     | 0.30         | 0/65        | 1.17        | 0/98        |
| 2   | D     | 0.35         | 0/2699      | 0.57        | 0/3645      |
| All | All   | 0.35         | 0/2764      | 0.59        | 0/3743      |

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 the 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 within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 1   | U     | 60    | 0        | 31       | 0       | 0            |
| 2   | D     | 2650  | 0        | 2678     | 11      | 0            |
| All | All   | 2710  | 0        | 2709     | 11      | 0            |

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 (11) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

| Atom-1           | Atom-2           | Interatomic distance (Å) | Clash overlap (Å) |
|------------------|------------------|--------------------------|-------------------|
| 2:D:285:ARG:HH12 | 2:D:322:ILE:HD11 | 1.65                     | 0.60              |
| 2:D:117:ALA:HB2  | 2:D:197:ARG:HE   | 1.66                     | 0.59              |

*Continued on next page...*

Continued from previous page...

| Atom-1           | Atom-2           | Interatomic distance (Å) | Clash overlap (Å) |
|------------------|------------------|--------------------------|-------------------|
| 2:D:93:ARG:HH12  | 2:D:97:SER:HA    | 1.67                     | 0.57              |
| 2:D:164:SER:HB2  | 2:D:175:LYS:HB3  | 1.89                     | 0.55              |
| 2:D:166:GLU:HB2  | 2:D:173:LYS:HB3  | 1.89                     | 0.54              |
| 2:D:370:GLN:HE22 | 2:D:375:GLN:HA   | 1.72                     | 0.54              |
| 2:D:239:TRP:HZ3  | 2:D:266:LEU:HD11 | 1.79                     | 0.47              |
| 2:D:168:VAL:HG23 | 2:D:171:ILE:HB   | 1.97                     | 0.46              |
| 2:D:156:ARG:HH22 | 2:D:192:GLU:HG3  | 1.82                     | 0.44              |
| 2:D:236:ILE:HG13 | 2:D:240:LEU:HD23 | 2.02                     | 0.42              |
| 2:D:239:TRP:HE3  | 2:D:240:LEU:HD22 | 1.85                     | 0.42              |

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 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 EM 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 |     |
|-----|-------|---------------|-----------|---------|----------|-------------|-----|
| 2   | D     | 332/430 (77%) | 312 (94%) | 20 (6%) | 0        | 100         | 100 |

There are no Ramachandran outliers to report.

### 5.3.2 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 EM 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 |    |
|-----|-------|---------------|-----------|----------|-------------|----|
| 2   | D     | 292/369 (79%) | 289 (99%) | 3 (1%)   | 78          | 87 |

All (3) residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 2   | D     | 158 | ARG  |
| 2   | D     | 263 | ARG  |
| 2   | D     | 314 | ARG  |

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA [i](#)

| Mol | Chain | Analysed  | Backbone Outliers | Pucker Outliers |
|-----|-------|-----------|-------------------|-----------------|
| 1   | U     | 2/3 (66%) | 0                 | 0               |

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.