



# wwPDB NMR Structure Validation Summary Report ⓘ

Feb 14, 2017 – 07:28 PM EST

PDB ID : 2JPP  
Title : Structural basis of RsmA/CsrA RNA recognition: Structure of RsmE bound to the Shine-Dalgarno sequence of hcnA mRNA  
Authors : Schubert, M.; Lapouge, K.; Duss, O.; Oberstrass, F.C.; Jelesarov, I.; Haas, D.; Allain, F.H.-T.  
Deposited on : 2007-05-21

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

<http://wwpdb.org/validation/2016/NMRValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

Cyrange	:	NOT EXECUTED
NmrClust	:	NOT EXECUTED
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	NOT EXECUTED
PANAV	:	NOT EXECUTED
ShiftChecker	:	NOT EXECUTED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20028442

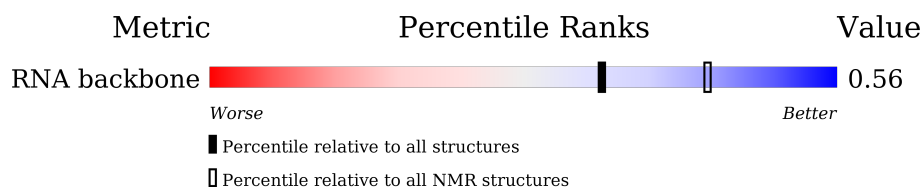
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment was not calculated.

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)
RNA backbone	3027	600

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	C	20	100%
1	D	20	100%
2	A	70	76% 24%
2	B	70	76% 24%

## 2 Ensemble composition and analysis ⓘ

This entry contains 10 models. The atoms present in the NMR models are not consistent. Some calculations may have failed as a result. All residues are included in the validation scores.

Cyrange was unable to find well-defined residues.

Error message: Cyrange did not run

NmrClust was unable to cluster the ensemble.

Error message: NmrClust did not run

### 3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2978 atoms, of which 1308 are hydrogens and 0 are deuteriums.

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

Mol	Chain	Residues	Atoms						Trace
1	C	20	Total	C	H	N	O	P	0
			645	191	219	79	137	19	
1	D	20	Total	C	H	N	O	P	0
			645	191	219	79	137	19	

- Molecule 2 is a protein called Translational repressor.

Mol	Chain	Residues	Atoms						Trace
2	A	53	Total	C	H	N	O	S	0
			844	256	435	75	77	1	
2	B	53	Total	C	H	N	O	S	0
			844	256	435	75	77	1	

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	65	HIS	-	EXPRESSION TAG	UNP Q5MXB2
A	66	HIS	-	EXPRESSION TAG	UNP Q5MXB2
A	67	HIS	-	EXPRESSION TAG	UNP Q5MXB2
A	68	HIS	-	EXPRESSION TAG	UNP Q5MXB2
A	69	HIS	-	EXPRESSION TAG	UNP Q5MXB2
A	70	HIS	-	EXPRESSION TAG	UNP Q5MXB2
B	65	HIS	-	EXPRESSION TAG	UNP Q5MXB2
B	66	HIS	-	EXPRESSION TAG	UNP Q5MXB2
B	67	HIS	-	EXPRESSION TAG	UNP Q5MXB2
B	68	HIS	-	EXPRESSION TAG	UNP Q5MXB2
B	69	HIS	-	EXPRESSION TAG	UNP Q5MXB2
B	70	HIS	-	EXPRESSION TAG	UNP Q5MXB2

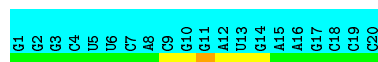
## 4 Residue-property plots

### 4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA 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\*GP\*GP\*CP\*UP\*UP\*CP\*AP\*CP\*GP\*GP\*AP\*UP\*GP\*AP\*AP\*GP\*CP\*CP\*C)-3')

Chain C:  100%





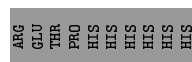
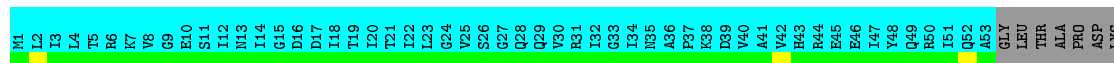
- Molecule 1: RNA (5'-R(\*GP\*GP\*GP\*CP\*UP\*UP\*CP\*AP\*CP\*GP\*GP\*AP\*UP\*GP\*AP\*AP\*GP\*CP\*CP\*C)-3')

Chain D:  100%





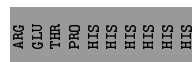
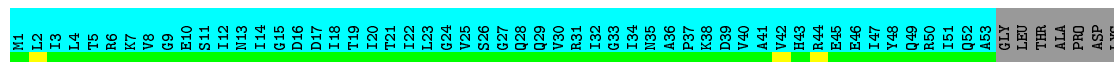
- Molecule 2: Translational repressor

Chain A:  76%  24%



- Molecule 2: Translational repressor

Chain B:  76%  24%

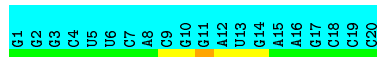


## 4.2 Residue scores for the representative (author defined) model from the NMR ensemble

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

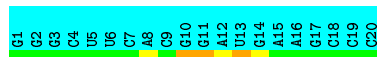
- Molecule 1: RNA (5'-R(\*GP\*GP\*GP\*CP\*UP\*UP\*CP\*AP\*CP\*GP\*GP\*AP\*UP\*GP\*AP\*AP\*GP\*CP\*CP\*C)-3')

Chain C:  100%





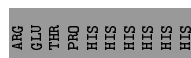
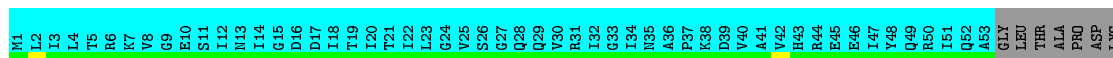
- Molecule 1: RNA (5'-R(\*GP\*GP\*GP\*CP\*UP\*UP\*CP\*AP\*CP\*GP\*GP\*AP\*UP\*GP\*AP\*AP\*GP\*CP\*CP\*C)-3')

Chain D:  100%





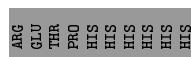
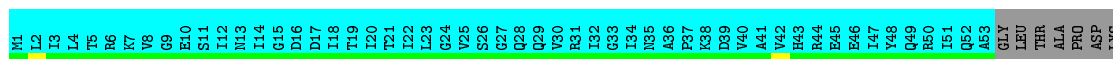
- Molecule 2: Translational repressor

Chain A:  76%  24%



- Molecule 2: Translational repressor

Chain B:  76%  24%



## 5 Refinement protocol and experimental data overview ⓘ

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

Of the 200 calculated structures, 10 were deposited, based on the following criterion: *structures with the least restraint violations*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
AMBER	refinement	7.0
DYANA	geometry optimization	3.02

No chemical shift data was provided. No validations of the models with respect to experimental NMR restraints is performed at this time.

## 6 Model quality ⓘ

### 6.1 Standard geometry ⓘ

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 ⓘ

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	D	0	0	0	0±0
1	C	0	0	0	0±0
2	A	0	0	0	0±0
2	B	0	0	0	0±0
All	All	0	0	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 -.

There are no clashes.

### 6.3 Torsion angles ⓘ

#### 6.3.1 Protein backbone ⓘ

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
2	A	0	-	-	-	-
2	B	0	-	-	-	-
All	All	0	-	-	-	-



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
2	A	0	-	-	-
2	B	0	-	-	-
All	All	0	-	-	-

There are no protein residues with a non-rotameric sidechain to report.

### 6.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers	Suiteness
1	C	19/20 (95%)	4±0 (22±2%)	1±1 (4±3%)	0.56±0.03
1	D	19/20 (95%)	4±0 (22±2%)	2±1 (8±3%)	0.57±0.04
All	All	380/400 (95%)	84 (22%)	22 (6%)	0.56

The overall RNA backbone suiteness is 0.56.

5 of 12 unique RNA backbone outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
1	D	14	G	10
1	C	14	G	10
1	C	11	G	10
1	D	10	G	10
1	C	10	G	10

All unique RNA pucker outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
1	D	11	G	5
1	D	10	G	5
1	D	13	U	5
1	C	13	U	4
1	C	11	G	3

## 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 carbohydrates 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

No chemical shift data were provided