wwPDB NMR Structure Validation Summary Report

Feb 20, 2022 – 12:35 AM EST

PDB ID	:	1T4L
Title	:	Solution structure of double-stranded RNA binding domain of S. cerevisiae RNase III (Rnt1p) in complex with the 5' terminal RNA hairpin of snR47 precursor
Authors	:	Wu, H.; Henras, A.; Chanfreau, G.; Feigon, J.
Deposited on	:	2004-04-29

This is a wwPDB NMR Structure Validation Summary 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 ⓘ symbol.

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

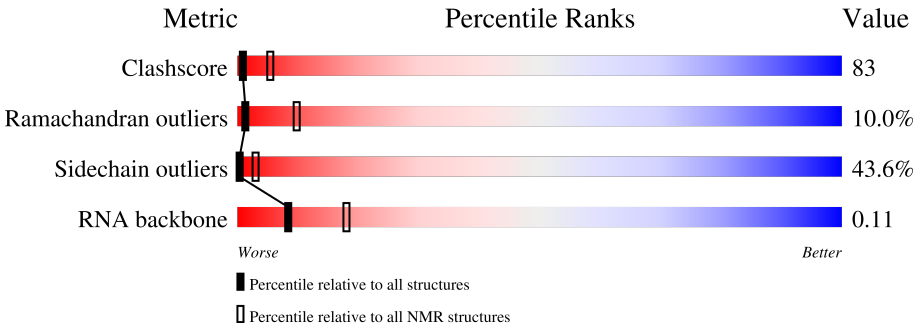
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.26
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.26

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)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428
RNA backbone	4643	676

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	32	44% 44% 12%
2	B	90	9% 50% 28% · 12%

2 Ensemble composition and analysis

This entry contains 15 models. Model 5 is the overall representative, medoid model (most similar to other models).

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	B:367-B:445 (79)	0.44	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 3 clusters and 1 single-model cluster was found.

Cluster number	Models
1	1, 3, 5, 6, 12
2	2, 7, 10, 11, 15
3	4, 8, 9, 13
Single-model clusters	14

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2437 atoms, of which 1067 are hydrogens and 0 are deuteriums.

- Molecule 1 is a RNA chain called 5' terminal hairpin of snR47 precursor.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
1	A	32	1028	306	346	124	221	31	0

- Molecule 2 is a protein called Ribonuclease III.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	90	1409	429	721	129	127	3	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	364	GLY	-	cloning artifact	UNP Q02555
B	365	SER	-	cloning artifact	UNP Q02555

4 Residue-property plots

4.1 Average score per residue in the NMR ensemble

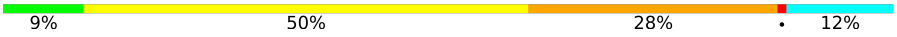
These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for 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: 5' terminal hairpin of snR47 precursor

Chain A:

• Molecule 2: Ribonuclease III

Chain B:



4.2 Residue scores for the representative (medoid) model from the NMR ensemble

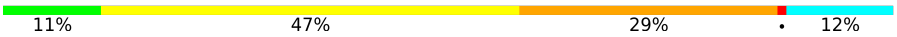
The representative model is number 5. Colouring as in section 4.1 above.

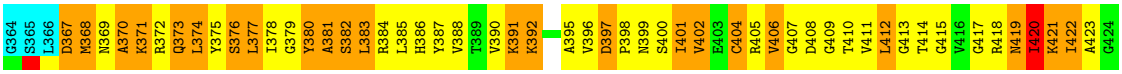
• Molecule 1: 5' terminal hairpin of snR47 precursor

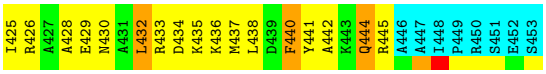
Chain A:

• Molecule 2: Ribonuclease III

Chain B:






I425
R426
A427
A428
E429
N430
A431
L432
R433
D434
K435
K436
M437
L438
D439
F440
Y441
A442
K443
Q444
R445
A446
A447
I448
P449
R450
S451
E452
S453

5 Refinement protocol and experimental data overview

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

Of the 100 calculated structures, 15 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
X-PLOR	refinement	NIH

No chemical shift data was provided.

6 Model quality

6.1 Standard geometry

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	1.13±0.01	0±0/763 (0.0± 0.0%)	2.00±0.01	44±1/1188 (3.7± 0.1%)
2	B	0.25±0.01	0±0/620 (0.0± 0.0%)	0.42±0.01	0±0/834 (0.0± 0.0%)
All	All	0.86	0/20745 (0.0%)	1.55	666/30330 (2.2%)

There are no bond-length outliers.

5 of 45 unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	26	G	N7-C8-N9	9.43	117.81	113.10	6	15
1	A	2	G	N7-C8-N9	9.40	117.80	113.10	14	15
1	A	16	G	N7-C8-N9	9.38	117.79	113.10	5	15
1	A	25	G	N7-C8-N9	9.37	117.78	113.10	7	15
1	A	23	G	N7-C8-N9	9.30	117.75	113.10	7	15

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	A	682	346	346	91±10
2	B	612	643	642	103±10
All	All	19410	14835	14820	2831

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 83.

5 of 871 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:B:390:VAL:HG23	2:B:401:ILE:HG23	1.04	1.29	6	14
2:B:406:VAL:HG13	2:B:412:LEU:HD21	1.02	1.20	15	3
2:B:380:TYR:CE2	2:B:383:LEU:HD12	1.00	1.91	9	2
1:A:30:C:O2'	1:A:31:U:H5'	0.95	1.61	11	15
2:B:377:LEU:O	2:B:378:ILE:HD13	0.93	1.62	8	12

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	
2	B	79/90 (88%)	56±3 (70±3%)	16±2 (20±3%)	8±2 (10±3%)	1	10
All	All	1185/1350 (88%)	834 (70%)	233 (20%)	118 (10%)	1	10

5 of 23 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
2	B	420	ILE	15
2	B	391	LYS	11
2	B	397	ASP	10
2	B	381	ALA	10
2	B	370	ALA	9

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	B	64/72 (89%)	36±4 (56±6%)	28±4 (44±6%)	0 3
All	All	960/1080 (89%)	541 (56%)	419 (44%)	0 3

5 of 53 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)
2	B	373	GLN	15
2	B	388	VAL	15
2	B	421	LYS	15
2	B	422	ILE	15
2	B	432	LEU	15

6.3.3 RNA

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers	Suiteness
1	A	31/32 (97%)	5±1 (16±4%)	2±1 (6±3%)	0.11±0.02
All	All	466/480 (97%)	76 (16%)	30 (6%)	0.11

The overall RNA backbone suiteness is 0.11.

5 of 15 unique RNA backbone outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
1	A	17	A	15
1	A	18	A	14
1	A	15	A	13
1	A	16	G	10
1	A	21	A	5

5 of 8 unique RNA pucker outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
1	A	16	G	11
1	A	17	A	8
1	A	28	A	3
1	A	25	G	2
1	A	3	G	2

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)

No chemical shift data were provided