



wwPDB NMR Structure Validation Summary Report ⓘ

Apr 26, 2016 – 04:57 PM BST

PDB ID : 1FJE
Title : SOLUTION STRUCTURE OF NUCLEOLIN RBD12 IN COMPLEX WITH
SNRE RNA
Authors : Allain, F.H.T.; Bouvet, P.; Dieckmann, T.; Feigon, J.
Deposited on : 2000-08-08

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
<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 : Kirchner and Güntert (2011)
NmrClust : Kelley et al. (1996)
MolProbity : 4.02b-467
Mogul : unknown
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI : v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV : Wang et al. (2010)
ShiftChecker : rb-20027457
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027457

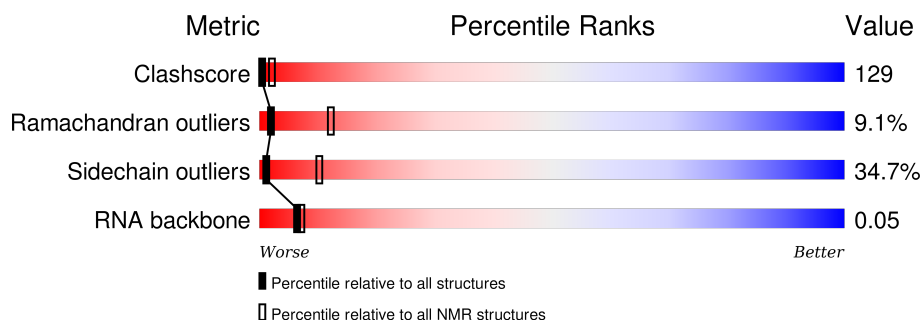
1 Overall quality at a glance

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	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958
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 ≥ 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	22	<div> <div>5%</div> <div>41%</div> <div>55%</div> </div>
2	B	175	<div> <div>7%</div> <div>58%</div> <div>24%</div> <div>• 10%</div> </div>

2 Ensemble composition and analysis

This entry contains 19 models. Model 10 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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:10-B:130, B:136-B:172 (158)	0.84	10

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 4 clusters and 3 single-model clusters were found.

Cluster number	Models
1	1, 11, 12, 14, 15
2	3, 6, 10, 19
3	5, 7, 8, 18
4	2, 13, 16
Single-model clusters	4; 9; 17

3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3444 atoms, of which 1611 are hydrogens and 0 are deuteriums.

- Molecule 1 is a RNA chain called SNRE RNA.

Mol	Chain	Residues	Atoms						Trace
1	A	22	Total	C	H	N	O	P	0
			712	211	242	90	148	21	

- Molecule 2 is a protein called NUCLEOLIN RBD12.

Mol	Chain	Residues	Atoms						Trace
2	B	175	Total	C	H	N	O	S	0
			2732	856	1369	229	277	1	

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	GLY	LYS	CONFLICT	UNP P08199
B	2	SER	LYS	CONFLICT	UNP P08199
B	3	HIS	GLN	CONFLICT	UNP P08199
B	4	MET	LYS	CONFLICT	UNP P08199
B	37	LEU	PRO	CONFLICT	UNP P08199

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.

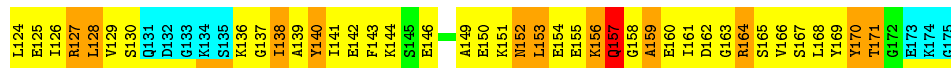
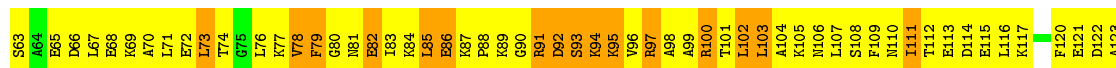
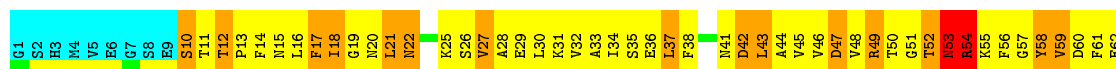
- Molecule 1: SNRE RNA

Chain A: 



- Molecule 2: NUCLEOLIN RBD12

Chain B: 



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

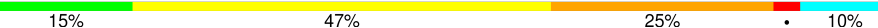
The representative model is number 10. Colouring as in section 4.1 above.

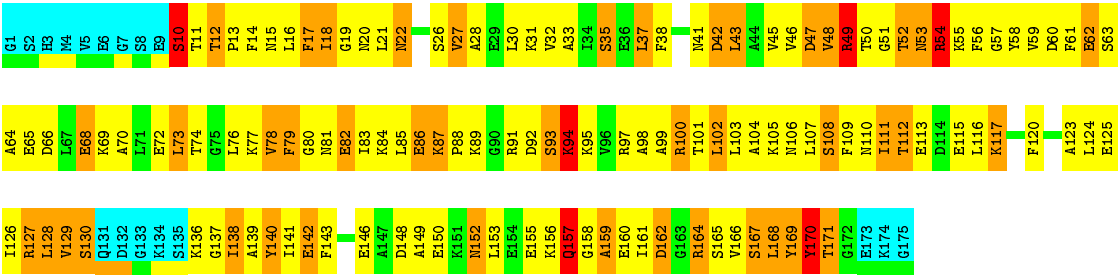
- Molecule 1: SNRE RNA

Chain A: 



- Molecule 2: NUCLEOLIN RBD12

Chain B: 



5 Refinement protocol and experimental data overview ⓘ

The models were refined using the following method: *simulated annealing using XPLOR 3.841*.

Of the 40 calculated structures, 19 were deposited, based on the following criterion: *all calculated structures submitted*.

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

Software name	Classification	Version
X-PLOR	refinement	3.841
X-PLOR	structure solution	3.841

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 (i)

6.1 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 (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	1.21±0.01	0±1/526 (0.1±0.1%)	2.10±0.01	32±2/819 (3.9±0.2%)
2	B	1.05±0.01	0±0/1257 (0.0±0.0%)	1.25±0.01	0±0/1690 (0.0±0.0%)
All	All	1.10	9/33877 (0.0%)	1.58	614/47671 (1.3%)

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	0.1±0.2
2	B	0.0±0.0	6.8±0.4
All	All	0	131

All unique bond 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(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	11	C	N1-C6	-5.76	1.33	1.37	14	1
1	A	10	C	N1-C6	-5.50	1.33	1.37	16	4
1	A	12	C	N1-C6	-5.41	1.33	1.37	18	1
1	A	13	G	N9-C8	-5.28	1.34	1.37	17	3

5 of 48 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(°)	Ideal(°)	Models	
								Worst	Total
1	A	13	G	N7-C8-N9	10.33	118.26	113.10	14	19
1	A	5	G	N7-C8-N9	9.73	117.96	113.10	11	19
1	A	20	G	N7-C8-N9	9.51	117.86	113.10	4	19
1	A	1	G	N7-C8-N9	9.39	117.80	113.10	12	19

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	16	G	N7-C8-N9	9.36	117.78	113.10	15	19

There are no chirality outliers.

5 of 8 unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
2	B	164	ARG	Sidechain	19
2	B	49	ARG	Sidechain	19
2	B	97	ARG	Sidechain	19
2	B	127	ARG	Sidechain	19
2	B	100	ARG	Sidechain	19

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	470	242	242	172±15
2	B	1241	1259	1259	297±19
All	All	32509	28519	28519	7866

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 129.

5 of 2534 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:B:102:LEU:HD21	2:B:149:ALA:HB1	1.14	1.18	3	6
2:B:21:LEU:HD22	2:B:30:LEU:HD22	1.12	1.19	18	1
2:B:102:LEU:HD23	2:B:149:ALA:HB1	1.12	1.18	14	5
2:B:73:LEU:HD12	2:B:76:LEU:HD11	1.11	1.14	12	5
1:A:3:C:O2'	1:A:4:C:H5'	1.08	1.48	8	11

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	158/175 (90%)	119±3 (75±2%)	25±2 (16±2%)	14±2 (9±1%)	2	12
All	All	3002/3325 (90%)	2252 (75%)	476 (16%)	274 (9%)	2	12

5 of 38 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	157	GLN	19
2	B	52	THR	19
2	B	42	ASP	19
2	B	79	PHE	18
2	B	170	TYR	17

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	134/147 (91%)	87±6 (65±4%)	47±6 (35±4%)	1	10
All	All	2546/2793 (91%)	1662 (65%)	884 (35%)	1	10

5 of 111 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	17	PHE	19
2	B	140	TYR	19
2	B	138	ILE	19
2	B	65	GLU	19
2	B	152	ASN	18

6.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers	Suiteness
1	A	21/22 (95%)	18±1 (85±6%)	8±1 (38±6%)	0.05±0.04
All	All	401/418 (96%)	339 (85%)	152 (38%)	0.05

The overall RNA backbone suiteness is 0.05.

5 of 21 unique RNA backbone outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
1	A	12	C	19
1	A	13	G	19
1	A	14	A	19
1	A	15	A	19
1	A	5	G	19

5 of 18 unique RNA pucker outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
1	A	12	C	19
1	A	13	G	19
1	A	4	C	19
1	A	15	A	18
1	A	9	U	17

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 carbohydrates 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

No chemical shift data were provided