



Full wwPDB NMR Structure Validation Report i

Feb 12, 2017 – 11:55 pm GMT

PDB ID : 2LCW
Title : solution structure of FUS/TLS RRM domain
Authors : Liu, X.; Ren, J.; Niu, C.; Gong, W.; Feng, W.
Deposited on : 2011-05-10

This is a Full wwPDB NMR Structure Validation 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 i symbol.

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

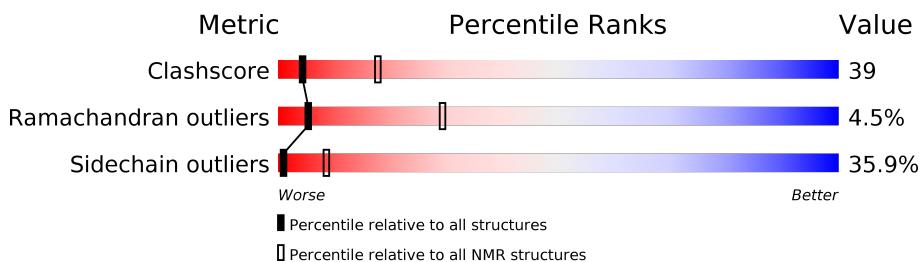
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	trunk28760
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
SOLUTION NMR

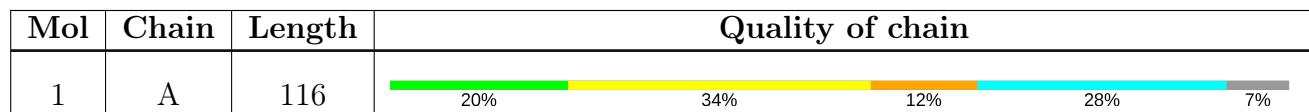
The overall completeness of chemical shifts assignment is 95%.

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	125131	11601
Ramachandran outliers	121729	10391
Sidechain outliers	121581	10367

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 $\leq 5\%$



2 Ensemble composition and analysis [\(i\)](#)

This entry contains 20 models. Model 17 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *closest to the average*.

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	A:284-A:325, A:335-A:368 (76)	0.33	17

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. No single-model clusters were found.

Cluster number	Models
1	5, 8, 12, 15, 16, 17, 18, 19
2	4, 6, 9, 11, 20
3	3, 7, 10, 14
4	1, 2, 13

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

There is only 1 type of molecule in this entry. The entry contains 1626 atoms, of which 799 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called RNA-binding protein FUS.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	108	1626	513	799	146	167	1	0

There are 8 discrepancies between the modelled and reference sequences:

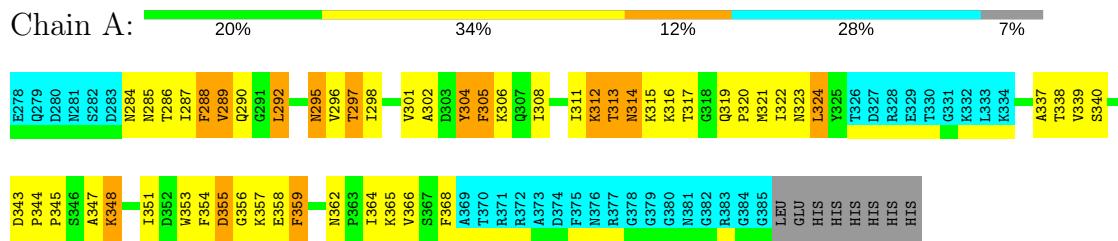
Chain	Residue	Modelled	Actual	Comment	Reference
A	386	LEU	-	EXPRESSION TAG	UNP P35637
A	387	GLU	-	EXPRESSION TAG	UNP P35637
A	388	HIS	-	EXPRESSION TAG	UNP P35637
A	389	HIS	-	EXPRESSION TAG	UNP P35637
A	390	HIS	-	EXPRESSION TAG	UNP P35637
A	391	HIS	-	EXPRESSION TAG	UNP P35637
A	392	HIS	-	EXPRESSION TAG	UNP P35637
A	393	HIS	-	EXPRESSION TAG	UNP P35637

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-binding protein FUS

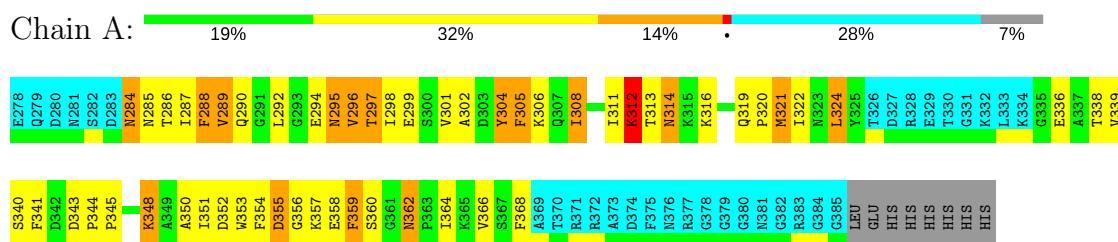


4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

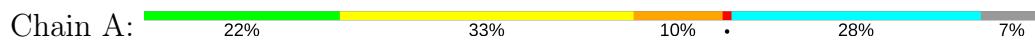
4.2.1 Score per residue for model 1

- Molecule 1: RNA-binding protein FUS



4.2.2 Score per residue for model 2

- Molecule 1: RNA-binding protein FUS



5 Refinement protocol and experimental data overview i

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

Of the 400 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy and the least restraint violations*.

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

Software name	Classification	Version
CNS	refinement	

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)	2lcw_cs.str
Number of chemical shift lists	1
Total number of shifts	1220
Number of shifts mapped to atoms	1220
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	95%

No validations of the models with respect to experimental NMR restraints is performed at this time.

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	336	GLU	6
1	A	307	GLN	5
1	A	358	GLU	5
1	A	294	GLU	5
1	A	299	GLU	5
1	A	317	THR	5
1	A	323	ASN	4
1	A	300	SER	4
1	A	354	PHE	3
1	A	340	SER	3
1	A	353	TRP	3
1	A	286	THR	1
1	A	341	PHE	1
1	A	338	THR	1
1	A	339	VAL	1
1	A	310	ILE	1

6.3.3 RNA [\(i\)](#)

There are no RNA molecules in this entry.

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.

