wwPDB NMR Structure Validation Summary Report

Jun 24, 2024 – 07:58 AM EDT

PDB ID	:	6U4N
BMRB ID	:	30659
Title	:	Solution structure of paxillin LIM4 in complex with kindlin-2 F0
Authors	:	Zhu, L.; Qin, J.
Deposited on	:	2019-08-26

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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

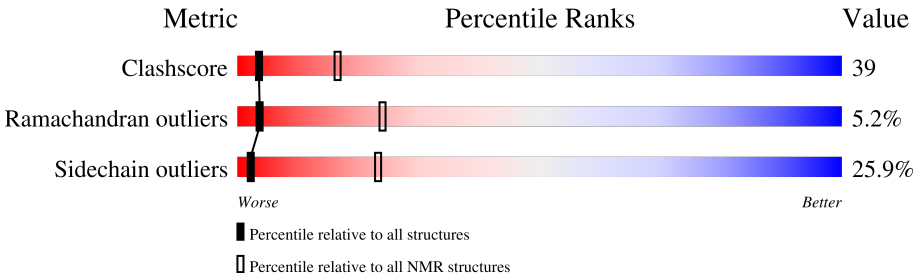
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

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

The overall completeness of chemical shifts assignment is 73%.

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

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	112	23% 38% 8% 14% 16%
2	B	72	26% 40% 15% 8% 10%

2 Ensemble composition and analysis

This entry contains 20 models. Model 12 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	A:16-A:93, B:547-B:605 (137)	0.88	12

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 1 single-model cluster was found.

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

3 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2573 atoms, of which 1279 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Fermitin family homolog 2.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	94	1532	494	765	132	137	4	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	GLY	-	expression tag	UNP Q96AC1
A	-5	ALA	-	expression tag	UNP Q96AC1
A	-4	MET	-	expression tag	UNP Q96AC1
A	-3	ASP	-	expression tag	UNP Q96AC1
A	-2	PRO	-	expression tag	UNP Q96AC1
A	-1	GLU	-	expression tag	UNP Q96AC1
A	0	PHE	-	expression tag	UNP Q96AC1

- Molecule 2 is a protein called Paxillin.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	65	1039	333	514	95	88	9	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	534	GLY	-	expression tag	UNP P49023
B	535	ALA	-	expression tag	UNP P49023
B	536	MET	-	expression tag	UNP P49023
B	537	ASP	-	expression tag	UNP P49023
B	538	PRO	-	expression tag	UNP P49023
B	539	GLU	-	expression tag	UNP P49023
B	540	PHE	-	expression tag	UNP P49023

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	
			Total	Zn
3	B	2	2	2

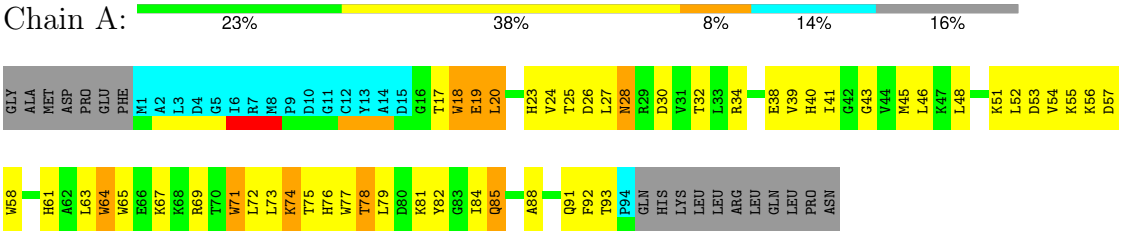
4 Residue-property plots

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.

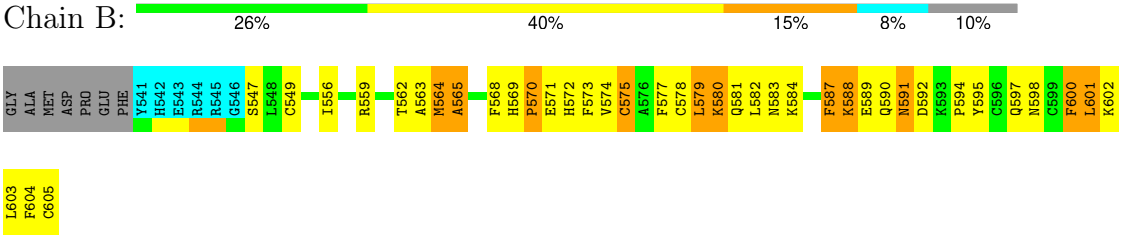
• Molecule 1: Fermitin family homolog 2

Chain A:



• Molecule 2: Paxillin

Chain B:

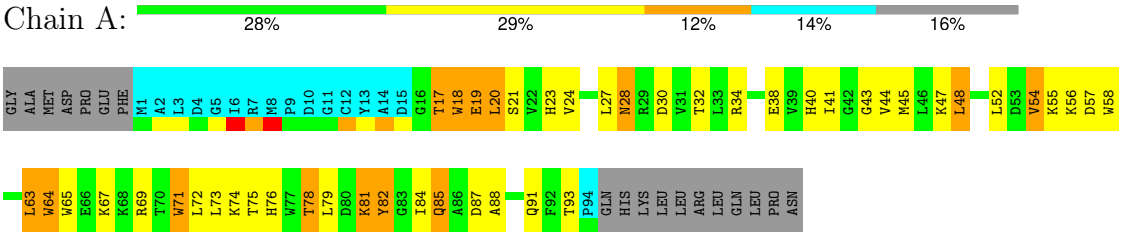


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 12. Colouring as in section 4.1 above.

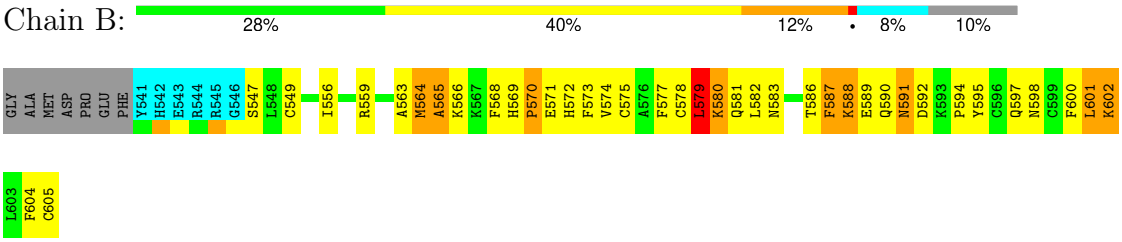
• Molecule 1: Fermitin family homolog 2

Chain A:



• Molecule 2: Paxillin

Chain B:


28%
40%
12%
8%
10%
GLY ALA MET ASP PRO GLU PHE Y541 H542 E543 R544 R545 G546 S547 L548 C549 I556 R559 A563 M564 A565 K566 K567 F568 H569 P570 E571 H572 F573 V574 C575 A576 F577 C578 L579 K580 Q581 L582 N583 T586 F587 K588 E589 Q590 N591 D592 K593 P594 Y595 C596 Q597 N598 C599 F600 L601 K602
L603 F604 C605

5 Refinement protocol and experimental data overview

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

Of the 100 calculated structures, 20 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 NIH	structure calculation	
X-PLOR NIH	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)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	1649
Number of shifts mapped to atoms	1599
Number of unparsed shifts	0
Number of shifts with mapping errors	50
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	73%

6 Model quality

6.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	A	650	656	656	47±6
2	B	468	463	465	43±5
All	All	22400	22380	22420	1770

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 39.

5 of 378 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:B:579:LEU:HD22	2:B:579:LEU:H	0.88	1.25	9	4
2:B:579:LEU:HD13	2:B:579:LEU:H	0.85	1.31	7	4
2:B:571:GLU:O	2:B:574:VAL:HG13	0.85	1.72	3	20
2:B:579:LEU:HD22	2:B:579:LEU:N	0.80	1.92	3	5
1:A:89:LYS:O	1:A:90:LEU:HD23	0.78	1.77	15	1

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
1	A	78/112 (70%)	61±2 (78±3%)	14±3 (18±3%)	3±1 (4±2%)	4 29
2	B	58/72 (81%)	40±1 (70±2%)	14±1 (24±2%)	4±0 (6±1%)	3 19
All	All	2720/3680 (74%)	2028 (75%)	550 (20%)	142 (5%)	4 24

5 of 18 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	72	LEU	20
2	B	564	MET	20
2	B	565	ALA	20
2	B	570	PRO	17
2	B	579	LEU	16

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
1	A	70/98 (71%)	51±3 (72±4%)	19±3 (28±4%)	2 20
2	B	53/63 (84%)	41±2 (77±4%)	12±2 (23±4%)	3 27
All	All	2460/3220 (76%)	1824 (74%)	636 (26%)	2 23

5 of 71 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)
1	A	18	TRP	20
1	A	19	GLU	20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	20	LEU	20
1	A	28	ASN	20
1	A	32	THR	20

6.3.3 RNA

There are no RNA molecules in this entry.

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 monosaccharides in this entry.

6.6 Ligand geometry

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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

The completeness of assignment taking into account all chemical shift lists is 73% for the well-defined parts and 72% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

7.1.1 Bookkeeping

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	1649
Number of shifts mapped to atoms	1599
Number of unparsed shifts	0
Number of shifts with mapping errors	50
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	13

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. First 5 (of 50) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	B	535	ALA	HA	4.112	0.03	1
1	B	535	ALA	HB1	1.265	0.03	1
1	B	535	ALA	HB2	1.265	0.03	1
1	B	535	ALA	HB3	1.265	0.03	1
1	B	535	ALA	CA	52.788	0.50	1
1	B	535	ALA	CB	18.966	0.50	1
1	B	536	MET	H	8.349	0.03	1
1	B	536	MET	HA	4.37	0.03	1
1	B	536	MET	HB2	1.95	0.03	2
1	B	536	MET	HB3	1.845	0.03	2
1	B	536	MET	HG2	2.427	0.03	2
1	B	536	MET	HG3	2.355	0.03	2
1	B	536	MET	CA	54.718	0.50	1
1	B	536	MET	CB	32.15	0.50	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	B	536	MET	CG	32.27	0.50	1
1	B	536	MET	N	118.129	0.25	1
1	B	537	ASP	H	8.056	0.03	1
1	B	537	ASP	HA	4.718	0.03	1
1	B	537	ASP	HB2	2.634	0.03	2
1	B	537	ASP	HB3	2.461	0.03	2
1	B	537	ASP	CA	52.519	0.50	1
1	B	537	ASP	CB	40.591	0.50	1
1	B	537	ASP	N	123.081	0.25	1
1	B	538	PRO	HA	4.227	0.03	1
1	B	538	PRO	HB2	2.142	0.03	2
1	B	538	PRO	HG2	1.898	0.03	2
1	B	538	PRO	HG3	1.735	0.03	2
1	B	538	PRO	HD2	3.755	0.03	2
1	B	538	PRO	HD3	3.624	0.03	2
1	B	538	PRO	CA	63.553	0.50	1
1	B	538	PRO	CB	31.831	0.50	1
1	B	538	PRO	CG	27.156	0.50	1
1	B	538	PRO	CD	50.515	0.50	1
1	B	539	GLU	H	8.415	0.03	1
1	B	539	GLU	HA	3.991	0.03	1
1	B	539	GLU	HB2	1.932	0.03	2
1	B	539	GLU	HB3	1.746	0.03	2
1	B	539	GLU	HG2	2.048	0.03	2
1	B	539	GLU	HG3	1.943	0.03	2
1	B	539	GLU	CA	56.892	0.50	1
1	B	539	GLU	CB	29.582	0.50	1
1	B	539	GLU	CG	36.09	0.50	1
1	B	539	GLU	N	119.47	0.25	1
1	B	540	PHE	H	7.893	0.03	1
1	B	540	PHE	HA	4.443	0.03	1
1	B	540	PHE	HB2	2.898	0.03	2
1	B	540	PHE	HB3	2.848	0.03	2
1	B	540	PHE	CA	57.948	0.50	1
1	B	540	PHE	CB	39.638	0.50	1
1	B	540	PHE	N	119.35	0.25	1

7.1.2 Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	165	0.14 ± 0.16	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	154	0.34 ± 0.08	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	135	0.32 ± 0.49	None needed (< 0.5 ppm)

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 73%, i.e. 1410 atoms were assigned a chemical shift out of a possible 1928. 0 out of 25 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	503/687 (73%)	255/279 (91%)	137/274 (50%)	111/134 (83%)
Sidechain	895/1041 (86%)	592/675 (88%)	293/327 (90%)	10/39 (26%)
Aromatic	12/200 (6%)	7/103 (7%)	0/87 (0%)	5/10 (50%)
Overall	1410/1928 (73%)	854/1057 (81%)	430/688 (62%)	126/183 (69%)

7.1.4 Statistically unusual chemical shifts

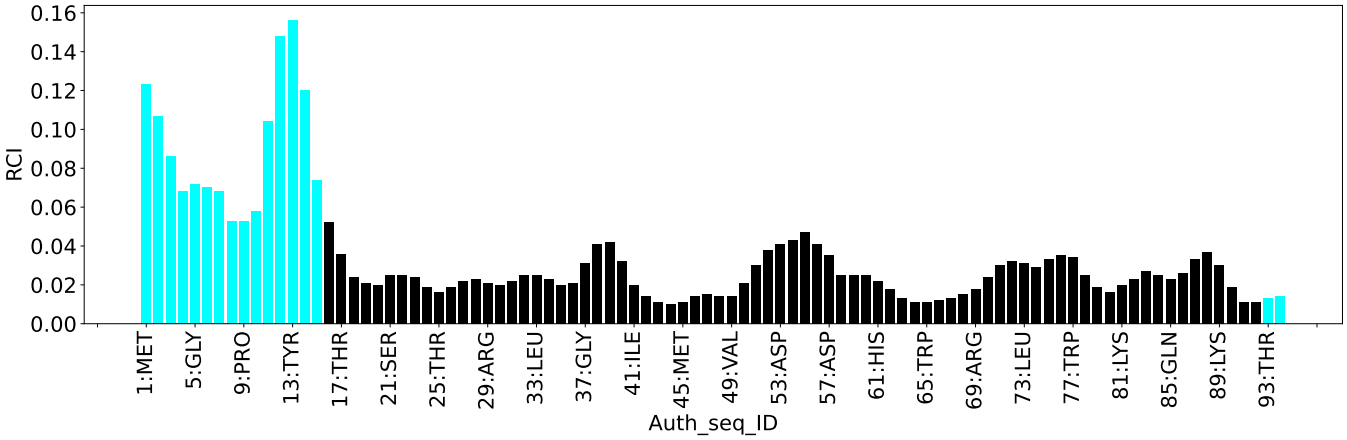
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	74	LYS	HD3	0.10	0.54 – 2.65	-7.1
1	A	68	LYS	HB3	0.12	0.46 – 3.04	-6.3
1	A	41	ILE	CG2	9.30	10.93 – 24.12	-6.2
1	A	35	VAL	HG21	-0.91	-0.58 – 2.19	-6.2
1	A	35	VAL	HG22	-0.91	-0.58 – 2.19	-6.2
1	A	35	VAL	HG23	-0.91	-0.58 – 2.19	-6.2
1	A	61	HIS	HB3	0.83	1.18 – 4.91	-6.0
1	A	68	LYS	HG2	-0.10	0.13 – 2.61	-5.9
1	A	81	LYS	HD3	0.37	0.54 – 2.65	-5.8
1	A	68	LYS	HG3	-0.15	0.04 – 2.67	-5.7
1	A	74	LYS	HB3	0.29	0.46 – 3.04	-5.7
1	B	603	LEU	HB3	-0.48	-0.26 – 3.31	-5.6
1	A	74	LYS	HD2	0.57	0.58 – 2.64	-5.0

7.1.5 Random Coil Index (RCI) plots

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:



Random coil index (RCI) for chain B: