wwPDB NMR Structure Validation Summary Report

May 28, 2020 – 10:56 pm BST

PDB ID	:	2L6F
Title	:	NMR Solution structure of FAT domain of FAK complexed with LD2 and LD4 motifs of PAXILLIN
Authors	:	Bertolucci, C.M.; Guibao, C.; Zhang, C.; Zheng, J.
Deposited on	:	2010-11-19

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 following versions of software and data (see references (1)) were used in the production of this report:

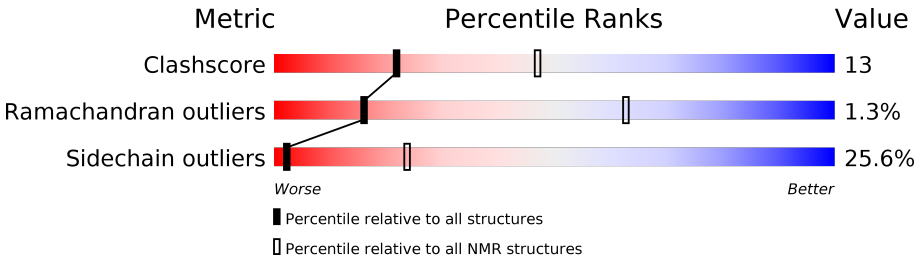
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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

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	215	42% 29% • 9% 18%

2 Ensemble composition and analysis

This entry contains 20 models. Model 6 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:10-A:137, A:160-A:174, A:205-A:218 (157)	0.39	6

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 2 clusters and 2 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2778 atoms, of which 1411 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Focal adhesion kinase 1, linker1, Paxillin, linker2, Paxillin.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	176	2778	854	1411	235	269	9	0

There are 41 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	4	MET	-	INITIATING METHIONINE	UNP Q00944
A	143	GLY	-	LINKER	UNP Q00944
A	144	GLY	-	LINKER	UNP Q00944
A	145	SER	-	LINKER	UNP Q00944
A	146	GLY	-	LINKER	UNP Q00944
A	147	GLY	-	LINKER	UNP Q00944
A	148	SER	-	LINKER	UNP Q00944
A	149	GLY	-	LINKER	UNP Q00944
A	150	SER	-	LINKER	UNP Q00944
A	151	GLY	-	LINKER	UNP Q00944
A	152	GLY	-	LINKER	UNP Q00944
A	153	SER	-	LINKER	UNP Q00944
A	154	GLY	-	LINKER	UNP Q00944
A	155	GLY	-	LINKER	UNP Q00944
A	156	SER	-	LINKER	UNP Q00944
A	157	GLY	-	LINKER	UNP Q00944
A	180	GLY	-	LINKER	UNP P49024
A	181	SER	-	LINKER	UNP P49024
A	182	GLY	-	LINKER	UNP P49024
A	183	SER	-	LINKER	UNP P49024
A	184	GLY	-	LINKER	UNP P49024
A	185	SER	-	LINKER	UNP P49024
A	186	GLY	-	LINKER	UNP P49024
A	187	GLY	-	LINKER	UNP P49024
A	188	SER	-	LINKER	UNP P49024
A	189	GLY	-	LINKER	UNP P49024
A	190	GLY	-	LINKER	UNP P49024
A	191	SER	-	LINKER	UNP P49024
A	192	GLY	-	LINKER	UNP P49024
A	193	GLY	-	LINKER	UNP P49024
A	194	SER	-	LINKER	UNP P49024

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	195	GLY	-	LINKER	UNP P49024
A	196	GLY	-	LINKER	UNP P49024
A	197	SER	-	LINKER	UNP P49024
A	198	GLY	-	LINKER	UNP P49024
A	199	GLY	-	LINKER	UNP P49024
A	200	SER	-	LINKER	UNP P49024
A	201	GLY	-	LINKER	UNP P49024
A	202	GLY	-	LINKER	UNP P49024
A	203	SER	-	LINKER	UNP P49024
A	204	SER	-	LINKER	UNP P49024

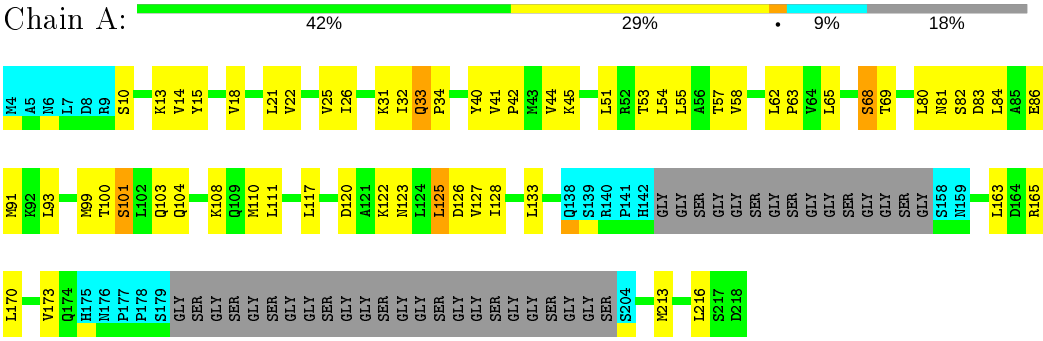
4 Residue-property plots (i)

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: Focal adhesion kinase 1, linker1, Paxillin, linker2, Paxillin

Chain A:

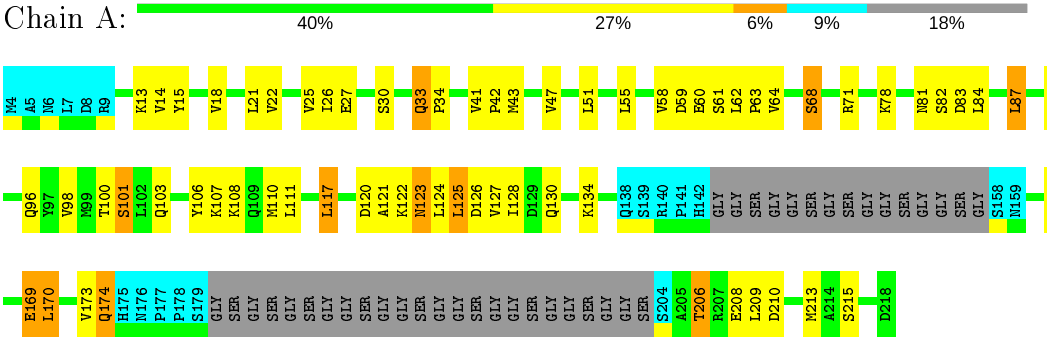


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 6. Colouring as in section 4.1 above.

- Molecule 1: Focal adhesion kinase 1, linker1, Paxillin, linker2, Paxillin

Chain A:



5 Refinement protocol and experimental data overview

The models were refined using the following method: *SIMULATED ANNEALING, TORSION ANGLE DYNAMICS*.

Of the 500 calculated structures, 20 were deposited, based on the following criterion: *target function*.

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

Software name	Classification	Version
CYANA	refinement	
CYANA	structure solution	
SPARKY	structure solution	3
NMRPIPE	structure solution	
MOLMOL	structure solution	

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	A	1218	1275	1275	32±5
All	All	24360	25500	25500	645

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 13.

5 of 246 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:54:LEU:HD21	1:A:80:LEU:HD13	0.94	1.39	20	2
1:A:25:VAL:HG21	1:A:118:ALA:HB2	0.85	1.49	5	1
1:A:51:LEU:HD13	1:A:84:LEU:HD13	0.81	1.52	12	2
1:A:119:VAL:HG22	1:A:167:LEU:HD21	0.81	1.51	18	1
1:A:51:LEU:HD11	1:A:117:LEU:HD22	0.80	1.52	7	1

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
1	A	156/215 (73%)	146±2 (93±1%)	8±2 (5±1%)	2±0 (1±0%)	16 63
All	All	3120/4300 (73%)	2914 (93%)	166 (5%)	40 (1%)	16 63

All 2 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	101	SER	20
1	A	68	SER	20

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	138/170 (81%)	103±5 (74±4%)	35±5 (26±4%)	2 24
All	All	2760/3400 (81%)	2053 (74%)	707 (26%)	2 24

5 of 96 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	125	LEU	20
1	A	126	ASP	20
1	A	122	LYS	19
1	A	87	LEU	17
1	A	82	SER	15

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

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