



wwPDB NMR Structure Validation Summary Report ⓘ

Apr 26, 2016 – 03:47 PM BST

PDB ID : 1LUI
Title : NMR Structures of Itk SH2 domain, Pro287cis isoform, ensemble of 20 low energy structures
Authors : Mallis, R.J.; Brazin, K.N.; Fulton, B.F.; Andreotti, A.M.
Deposited on : 2002-05-22

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

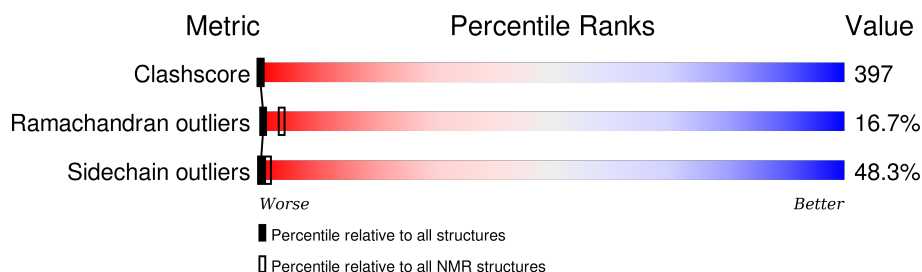
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 81%.

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

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	110	

2 Ensemble composition and analysis

This entry contains 20 models. Model 11 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:6-A:53, A:60-A:70, A:74-A:111 (97)	0.35	11

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

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1764 atoms, of which 877 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Tyrosine-protein kinase ITK/TSK.

Mol	Chain	Residues	Atoms						Trace
1	A	110	Total	C	H	N	O	S	1
			1764	567	877	151	166	3	

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	111	GLY	-	CLONING ARTIFACT	UNP Q03526

5 Refinement protocol and experimental data overview

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

Of the 200 calculated structures, 20 were deposited, based on the following criterion: *structures with favorable non-bond energy*.

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

Software name	Classification	Version
CNS	refinement	1.0

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)	BMRB entry 5461
Number of chemical shift lists	2
Total number of shifts	2441
Number of shifts mapped to atoms	2315
Number of unparsed shifts	0
Number of shifts with mapping errors	126
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	81%

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

6 Model quality

6.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ACE, NH2

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	800	798	797	634±21
All	All	16000	15960	15940	12672

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 397.

5 of 2659 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:6:LEU:HD12	1:A:12:TYR:CD1	1.27	1.62	10	20
1:A:34:PHE:CG	1:A:105:LEU:HD12	1.24	1.64	17	17
1:A:94:HIS:CB	1:A:105:LEU:HD22	1.24	1.60	6	17
1:A:94:HIS:CD2	1:A:101:LEU:HD22	1.24	1.66	3	11
1:A:11:TRP:CG	1:A:34:PHE:CE2	1.23	2.26	15	20

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	97/110 (88%)	67±2 (69±2%)	14±2 (15±3%)	16±2 (17±2%)	0	4
All	All	1940/2200 (88%)	1333 (69%)	283 (15%)	324 (17%)	0	4

5 of 31 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	52	LYS	20
1	A	41	THR	20
1	A	42	PRO	20
1	A	80	GLU	19
1	A	104	ARG	19

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	86/97 (89%)	45±3 (52±3%)	42±3 (48±3%)	0	1
All	All	1720/1940 (89%)	890 (52%)	830 (48%)	0	1

5 of 68 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	92	GLN	20
1	A	76	TYR	20
1	A	37	ARG	20
1	A	16	ILE	20
1	A	66	ILE	20

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.

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 81% for the well-defined parts and 81% for the entire structure.

7.1 Chemical shift list 1

File name: BMRB entry 5461

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	1213
Number of shifts mapped to atoms	1150
Number of unparsed shifts	0
Number of shifts with mapping errors	63
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	5

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

- Residue not found in structure. First 5 (of 63) occurrences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	116	HIS	HD2	6.982	-1.0	2
A	115	ILE	HG23	0.824	-1.0	4
A	115	ILE	HD12	0.821	-1.0	1
A	116	HIS	CD2	116.712	-1.0	1
A	114	GLY	HA2	4.221	-1.0	2

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	113	-0.34 ± 0.29	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	106	-0.23 ± 0.26	None needed (< 0.5 ppm)
$^{13}\text{C}'$	0	—	—

Continued on next page...

Continued from previous page...

Nucleus	# values	Correction \pm precision, ppm	Suggested action
^{15}N	107	-0.17 \pm 0.41	None needed (< 0.5 ppm)

7.1.3 Completeness of resonance assignments [i](#)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 80%, i.e. 997 atoms were assigned a chemical shift out of a possible 1250. 0 out of 15 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	379/479 (79%)	189/191 (99%)	96/194 (49%)	94/94 (100%)
Sidechain	517/630 (82%)	324/371 (87%)	186/227 (82%)	7/32 (22%)
Aromatic	101/141 (72%)	71/73 (97%)	30/63 (48%)	0/5 (0%)
Overall	997/1250 (80%)	584/635 (92%)	312/484 (64%)	101/131 (77%)

7.1.4 Statistically unusual chemical shifts [i](#)

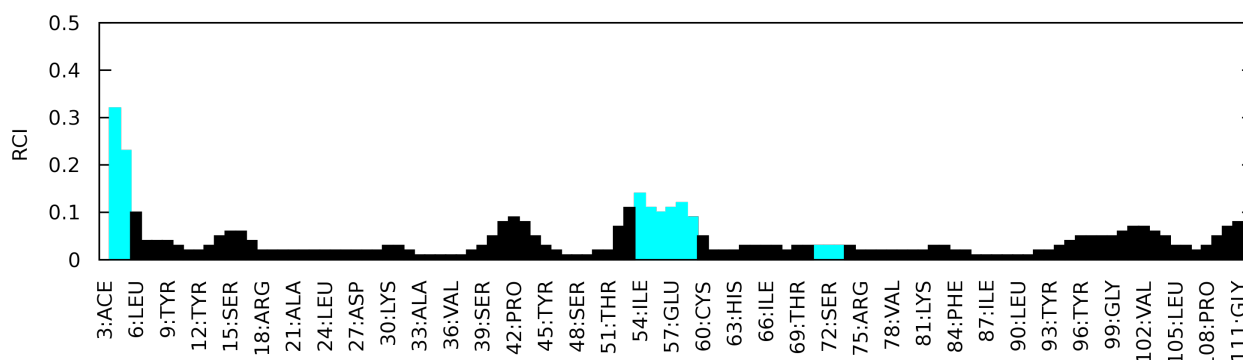
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	75	ARG	NE	121.48	92.63 – 76.73	23.1
1	A	37	ARG	NE	117.93	92.63 – 76.73	20.9
1	A	105	LEU	HB3	-1.01	3.34 – -0.26	-7.1
1	A	9	TYR	HB3	0.65	4.75 – 0.95	-5.8
1	A	67	LYS	HB3	0.29	3.10 – 0.40	-5.4

7.1.5 Random Coil Index (RCI) plots [i](#)

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.

Random coil index (RCI) for chain A:



7.2 Chemical shift list 2

File name: BMRB entry 5461

Chemical shift list name: *assigned_chem_shift_list_2*

7.2.1 Bookkeeping [i](#)

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	1228
Number of shifts mapped to atoms	1165
Number of unparsed shifts	0
Number of shifts with mapping errors	63
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	6

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

- Residue not found in structure. First 5 (of 63) occurrences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	116	HIS	HD2	6.982	-1.0	2
A	115	ILE	HG23	0.824	-1.0	4
A	115	ILE	HD12	0.821	-1.0	1
A	116	HIS	CD2	116.712	-1.0	1
A	114	GLY	HA2	4.221	-1.0	2

7.2.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	114	-0.31 ± 0.13	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	107	-0.21 ± 0.13	None needed (< 0.5 ppm)
$^{13}\text{C}'$	0	—	—
^{15}N	107	-0.34 ± 0.42	None needed (< 0.5 ppm)

7.2.3 Completeness of resonance assignments [i](#)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 81%, i.e. 1009 atoms were assigned a chemical shift out of a possible 1250. 0 out of 15 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	380/479 (79%)	189/191 (99%)	97/194 (50%)	94/94 (100%)
Sidechain	528/630 (84%)	328/371 (88%)	193/227 (85%)	7/32 (22%)
Aromatic	101/141 (72%)	71/73 (97%)	30/63 (48%)	0/5 (0%)
Overall	1009/1250 (81%)	588/635 (93%)	320/484 (66%)	101/131 (77%)

7.2.4 Statistically unusual chemical shifts [i](#)

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.

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	75	ARG	NE	121.48	92.63 – 76.73	23.1
1	A	37	ARG	NE	117.93	92.63 – 76.73	20.9
1	A	105	LEU	HB3	-1.01	3.34 – -0.26	-7.1
1	A	26	LEU	HG	3.64	3.16 – -0.14	6.5
1	A	9	TYR	HB3	0.65	4.75 – 0.95	-5.8
1	A	67	LYS	HB3	0.29	3.10 – 0.40	-5.4

7.2.5 Random Coil Index (RCI) plots [i](#)

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.

Random coil index (RCI) for chain A:

