wwPDB NMR Structure Validation Summary Report

Apr 26, 2016 – 10:56 PM BST

PDB ID	:	2K6Q
Title	:	LC3 p62 complex structure
Authors	:	Noda, N.; Kumeta, H.; Nakatogawa, H.; Satoo, K.; Adachi, W.; Ishii, J.; Fujioka, Y.; Ohsumi, Y.; Inagaki, F.
Deposited on	:	2008-07-17

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 (i) symbol.

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

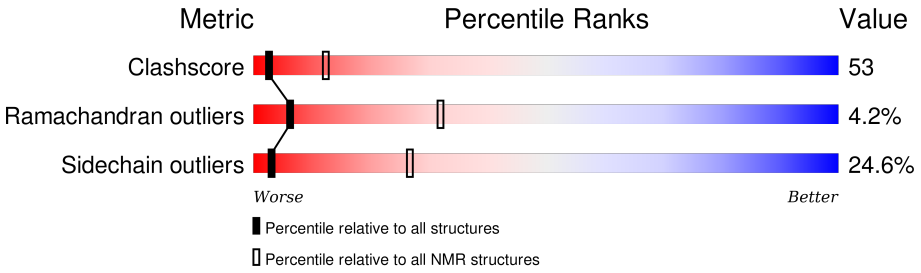
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal 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 (i)

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

The overall completeness of chemical shifts assignment is 92%.

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	121	17% 61% 7% • 12%
2	B	17	12% 41% 12% 35%

2 Ensemble composition and analysis

This entry contains 20 models. Model 4 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:13-A:118, B:337-B:347 (117)	0.28	4

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 15 single-model clusters were found.

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

3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2258 atoms, of which 1127 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Microtubule-associated proteins 1A/1B light chain 3B.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	121	2014	636	1013	176	185	4	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	EXPRESSION TAG	UNP Q62625

- Molecule 2 is a protein called p62_peptide from Sequestosome-1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	17	244	77	114	21	31	1	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	331	MET	-	EXPRESSION TAG	UNP O08623

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

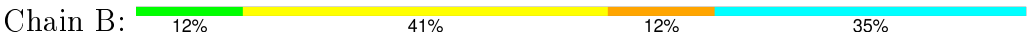
These plots are provided for all protein, RNA and DNA chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain 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.

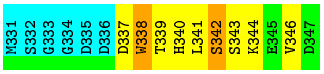
- Molecule 1: Microtubule-associated proteins 1A/1B light chain 3B

Chain A:

- Molecule 2: p62_peptide from Sequestosome-1

Chain B:



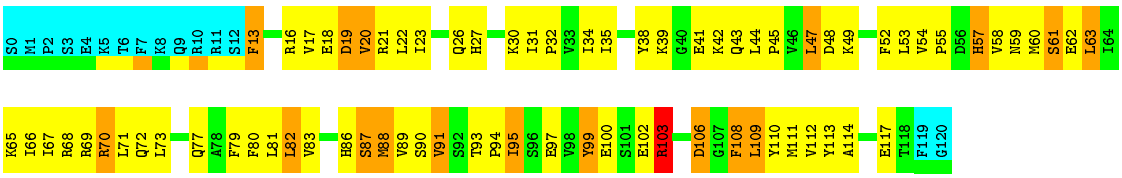


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 4. Colouring as in section 4.1 above.

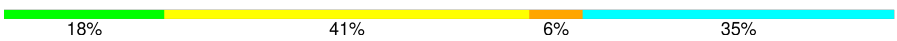
- Molecule 1: Microtubule-associated proteins 1A/1B light chain 3B

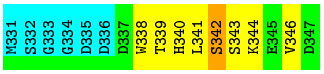
Chain A:



- Molecule 2: p62_peptide from Sequestosome-1

Chain B:




M331
S332
G333
G334
D335
D336
D337
W338
T339
H340
L341
S342
S343
K344
E345
V346
D347

5 Refinement protocol and experimental data overview

The models were refined using the following method: *simulated annealing, torsion angle dynamics.*

Of the 100 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
SPARKY	refinement	3.110
CYANA	structure solution	2.1
CYANA	refinement	2.1

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 15877
Number of chemical shift lists	1
Total number of shifts	1776
Number of shifts mapped to atoms	1776
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	92%

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	877	890	888	99±9
2	B	92	83	81	21±4
All	All	19380	19460	19380	2047

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 53.

5 of 563 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:47:LEU:HD22	1:A:71:LEU:HD21	1.03	1.28	16	8
1:A:35:ILE:HD13	1:A:67:ILE:HD11	0.99	1.34	18	10
1:A:63:LEU:HD22	1:A:63:LEU:O	0.95	1.60	13	1
1:A:13:PHE:O	1:A:17:VAL:HG13	0.94	1.63	16	5
1:A:66:ILE:HG21	2:B:341:LEU:HD13	0.90	1.39	20	8

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	106/121 (88%)	86±2 (81±2%)	16±2 (15±2%)	4±1 (4±1%)	7 34
2	B	10/17 (59%)	8±1 (78±8%)	1±1 (14±10%)	1±1 (8±6%)	2 15
All	All	2320/2760 (84%)	1872 (81%)	350 (15%)	98 (4%)	6 32

5 of 17 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	103	ARG	20
1	A	43	GLN	19
1	A	39	LYS	14
2	B	338	TRP	12
1	A	91	VAL	11

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	100/114 (88%)	78±2 (78±2%)	22±2 (22±2%)	4 31
2	B	11/15 (73%)	6±1 (52±9%)	5±1 (48±9%)	0 1
All	All	2220/2580 (86%)	1673 (75%)	547 (25%)	3 27

5 of 76 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	88	MET	20
1	A	59	ASN	20
1	A	95	ILE	20
1	A	47	LEU	20
1	A	49	LYS	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 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

The completeness of assignment taking into account all chemical shift lists is 92% for the well-defined parts and 91% for the entire structure.

7.1 Chemical shift list 1

File name: BMRB entry 15877

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	1776
Number of shifts mapped to atoms	1776
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	76

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	138	-0.25 ± 0.13	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	131	0.06 ± 0.09	None needed (< 0.5 ppm)
$^{13}C'$	128	-0.09 ± 0.12	None needed (< 0.5 ppm)
^{15}N	130	0.60 ± 0.40	None needed (imprecise)

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 92%, i.e. 1406 atoms were assigned a chemical shift out of a possible 1534. 0 out of 23 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	569/575 (99%)	229/229 (100%)	228/234 (97%)	112/112 (100%)
Sidechain	741/842 (88%)	450/493 (91%)	277/308 (90%)	14/41 (34%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	96/117 (82%)	52/63 (83%)	43/49 (88%)	1/5 (20%)
Overall	1406/1534 (92%)	731/785 (93%)	548/591 (93%)	127/158 (80%)

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
2	B	344	LYS	CE	72.06	46.00 – 37.80	36.8
2	B	339	THR	CG2	51.11	27.15 – 15.95	26.4
2	B	341	LEU	CG	56.68	32.55 – 21.05	26.0
2	B	344	LYS	CG	54.67	30.67 – 19.17	25.9
2	B	338	TRP	CZ3	161.54	129.20 – 113.60	25.7
2	B	344	LYS	CD	58.97	34.86 – 23.06	25.4
2	B	345	GLU	CG	66.84	42.24 – 29.94	25.0
2	B	331	MET	CG	61.86	38.33 – 25.73	23.7
2	B	334	GLY	CA	75.67	51.81 – 38.91	23.5
1	A	69	ARG	NE	121.89	92.63 – 76.73	23.4
2	B	333	GLY	CA	75.47	51.81 – 38.91	23.3
1	A	24	ARG	NE	119.16	92.63 – 76.73	21.7
1	A	37	ARG	NE	119.08	92.63 – 76.73	21.6
1	A	21	ARG	NE	118.97	92.63 – 76.73	21.6
2	B	342	SER	CB	95.11	71.24 – 56.34	21.0
2	B	346	VAL	CG1	49.98	28.40 – 14.60	20.6
2	B	332	SER	CB	93.93	71.24 – 56.34	20.2
2	B	338	TRP	CE3	155.09	129.06 – 111.96	20.2
2	B	338	TRP	CH2	160.82	133.06 – 114.56	20.0
2	B	343	SER	CB	93.47	71.24 – 56.34	19.9
2	B	338	TRP	CZ2	144.03	121.76 – 106.66	19.7
2	B	346	VAL	CG2	51.66	29.20 – 13.40	19.2
2	B	347	ASP	CB	72.28	49.06 – 32.66	19.2
2	B	337	ASP	CB	71.93	49.06 – 32.66	18.9
2	B	336	ASP	CB	71.55	49.06 – 32.66	18.7
2	B	341	LEU	CD1	54.91	32.77 – 16.57	18.7
2	B	335	ASP	CB	71.19	49.06 – 32.66	18.5
2	B	339	THR	CB	99.44	78.10 – 61.30	17.7
2	B	345	GLU	CB	60.18	38.65 – 21.35	17.4
2	B	344	LYS	CB	62.83	41.68 – 23.88	16.9
2	B	346	VAL	CB	63.11	41.76 – 23.66	16.8
2	B	341	LEU	CD2	52.53	32.60 – 15.60	16.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
2	B	341	LEU	CB	73.60	51.69 – 32.89	16.7
2	B	338	TRP	CD1	158.33	136.18 – 116.78	16.4
2	B	338	TRP	NE1	160.13	139.19 – 119.59	15.7
2	B	338	TRP	CB	61.22	40.02 – 19.92	15.5
2	B	347	ASP	CA	86.07	64.88 – 44.48	15.4
2	B	335	ASP	CA	85.07	64.88 – 44.48	14.9
2	B	340	HIS	CB	61.37	40.69 – 19.69	14.8
2	B	336	ASP	CA	84.71	64.88 – 44.48	14.7
2	B	343	SER	CA	89.55	69.25 – 48.25	14.7
2	B	337	ASP	CA	83.67	64.88 – 44.48	14.2
2	B	345	GLU	CA	87.15	67.86 – 46.86	14.2
2	B	332	SER	CA	88.44	69.25 – 48.25	14.1
2	B	344	LYS	CA	86.72	67.97 – 45.97	13.5
2	B	342	SER	CA	86.84	69.25 – 48.25	13.4
2	B	331	MET	CB	62.92	44.20 – 21.80	13.4
2	B	341	LEU	CA	83.68	66.36 – 44.96	13.1
2	B	340	HIS	CA	87.02	68.24 – 44.74	13.0
2	B	331	MET	CA	85.33	67.38 – 44.88	13.0
2	B	340	HIS	CE1	166.28	149.70 – 125.30	11.8
2	B	339	THR	CA	92.66	75.37 – 49.07	11.6
2	B	338	TRP	CA	87.02	70.54 – 44.84	11.4
2	B	346	VAL	CA	91.71	76.93 – 48.03	10.1
2	B	340	HIS	N	159.13	139.70 – 99.40	9.8
2	B	347	ASP	N	158.09	140.24 – 101.14	9.6
1	A	51	LYS	HE2	1.16	3.87 – 1.97	-9.3
2	B	343	SER	N	148.57	134.24 – 98.34	9.0
1	A	51	LYS	HE3	1.23	3.86 – 1.96	-8.8
2	B	332	SER	N	147.26	134.24 – 98.34	8.6
2	B	345	GLU	N	150.68	138.24 – 103.14	8.5
2	B	344	LYS	N	151.91	140.01 – 102.11	8.1
2	B	333	GLY	N	141.08	129.07 – 90.27	8.1
2	B	340	HIS	CD2	147.42	137.40 – 103.40	7.9
2	B	341	LEU	N	153.30	141.70 – 102.00	7.9
2	B	342	SER	N	144.02	134.24 – 98.34	7.7
2	B	334	GLY	N	138.89	129.07 – 90.27	7.5
2	B	335	ASP	N	149.64	140.24 – 101.14	7.4
2	B	337	ASP	N	149.56	140.24 – 101.14	7.4
2	B	336	ASP	N	148.88	140.24 – 101.14	7.2
2	B	339	THR	N	149.50	139.35 – 91.55	7.1
2	B	338	TRP	N	150.17	142.48 – 100.78	6.8
1	A	34	ILE	HD12	-1.02	2.13 – -0.77	-5.9
1	A	34	ILE	HD13	-1.02	2.13 – -0.77	-5.9

Continued on next page...

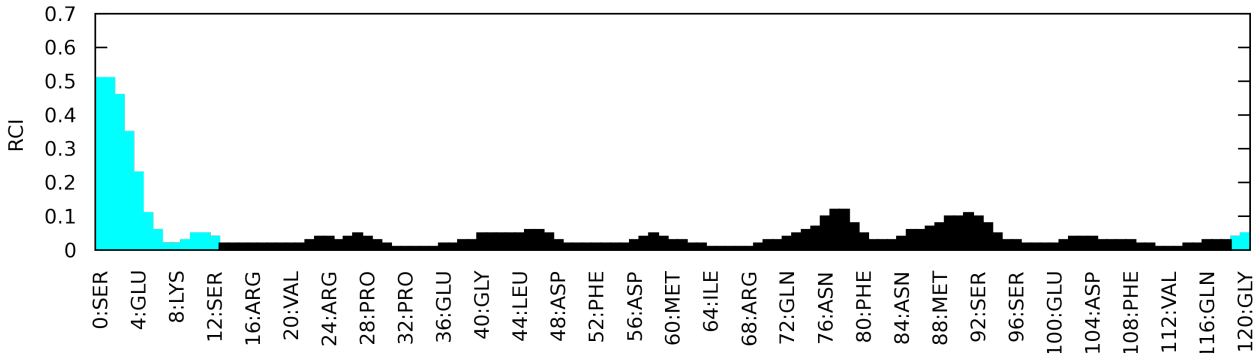
Continued from previous page...

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	34	ILE	HD11	-1.02	2.13 – -0.77	-5.9
2	B	346	VAL	N	146.72	144.09 – 98.19	5.6

7.1.5 Random Coil Index (RCI) plots

The images below report *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:



Random coil index (RCI) for chain B: