wwPDB NMR Structure Validation Summary Report

May 28, 2020 – 11:16 pm BST

PDB ID	:	2LV6
Title	:	The complex between Ca-Calmodulin and skeletal muscle myosin light chain kinase from combination of NMR and aqueous and contrast-matched SAXS data
Authors	:	Grishaev, A.V.; Anthis, N.J.; Clore, G.M.
Deposited on	:	2012-06-29

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:

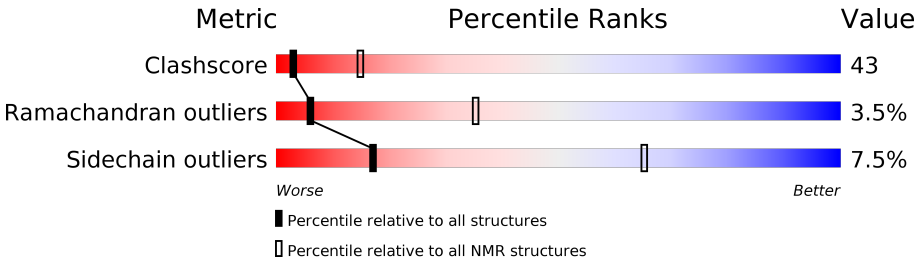
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.11
Ideal geometry (proteins)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

1 Overall quality at a glance (i)

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

The overall completeness of chemical shifts assignment is 14%.

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	148	76% 18% 6%
2	B	26	12% 54% 35%

2 Ensemble composition and analysis (i)

This entry contains 1 models. Identification of well-defined residues and clustering analysis are not possible.

3 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2700 atoms, of which 1323 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Calmodulin.

Mol	Chain	Residues	Atoms						Trace
1	A	148	Total	C	H	N	O	S	0
			2257	713	1094	187	254	9	

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	99	PHE	TYR	SEE REMARK 999	UNP P62158
A	143	THR	GLN	SEE REMARK 999	UNP P62158

- Molecule 2 is a protein called Myosin light chain kinase 2, skeletal/cardiac muscle.

Mol	Chain	Residues	Atoms					Trace
2	B	26	Total	C	H	N	O	0
			439	134	229	43	33	

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

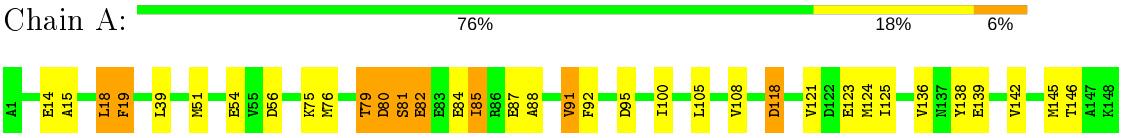
Mol	Chain	Residues	Atoms	
3	A	4	Total	Ca
			4	4

4 Residue-property plots (i)

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.

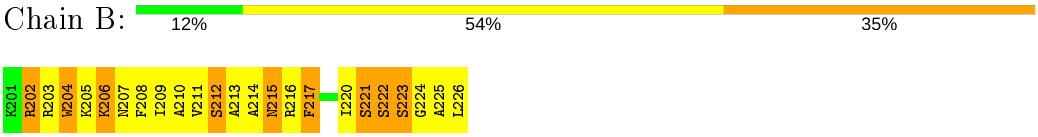
- Molecule 1: Calmodulin

Chain A:



- Molecule 2: Myosin light chain kinase 2, skeletal/cardiac muscle

Chain B:



5 Refinement protocol and experimental data overview

The models were refined using the following method: *rigid-body optimization, simulated annealing.*

Of the 1 calculated structures, 1 were deposited, based on the following criterion: *structures with the least restraint violations.*

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

Software name	Classification	Version
Custom	structure solution	
CNS	refinement	1.0
Custom	refinement	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 6 of this report.

Chemical shift file(s)	input_cs.cif
Number of chemical shift lists	1
Total number of shifts	288
Number of shifts mapped to atoms	288
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	14%

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

COVALENT-GEOMETRY INFOmissingINFO

5.1 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	1163	1094	1095	99
2	B	210	229	228	95
All	All	1377	1323	1323	117

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 43.

5 of 117 clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)
1:A:18:LEU:HB3	2:B:210:ALA:CB	1.13	1.74
1:A:145:MET:SD	2:B:205:LYS:HA	1.04	1.93
1:A:18:LEU:HB3	2:B:210:ALA:HB2	0.94	1.35
1:A:145:MET:CE	2:B:208:PHE:CB	0.88	2.51
1:A:105:LEU:HD22	2:B:204:TRP:CH2	0.86	2.06

5.2 Torsion angles (i)

5.2.1 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	146/148 (99%)	141 (97%)	2 (1%)	3 (2%)	10 50
2	B	24/26 (92%)	17 (71%)	4 (17%)	3 (12%)	1 6
All	All	170/174 (98%)	158 (93%)	6 (4%)	6 (4%)	6 35

5 of 6 Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
2	B	223	SER
1	A	82	GLU
1	A	80	ASP
1	A	79	THR
2	B	221	SER

5.2.2 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	126/126 (100%)	121 (96%)	5 (4%)	35 83
2	B	21/21 (100%)	15 (71%)	6 (29%)	2 18
All	All	147/147 (100%)	136 (93%)	11 (7%)	17 65

5 of 11 residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	A	81	SER
1	A	91	VAL
2	B	204	TRP
2	B	217	PHE
1	A	76	MET

5.2.3 RNA (i)

There are no RNA molecules in this entry.

5.3 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.4 Carbohydrates (i)

There are no carbohydrates in this entry.

5.5 Ligand geometry (i)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

5.6 Other polymers (i)

There are no such molecules in this entry.

5.7 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 14% for the well-defined parts and 14% for the entire structure.

6.1 Chemical shift list 1

File name: input_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

6.1.1 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	288
Number of shifts mapped to atoms	288
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

6.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	0	—	None (insufficient data)
$^{13}C_{\beta}$	0	—	None (insufficient data)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	144	0.12 ± 0.32	None needed (< 0.5 ppm)

6.1.3 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 14%, i.e. 288 atoms were assigned a chemical shift out of a possible 2111. 0 out of 18 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	288/866 (33%)	144/346 (42%)	0/348 (0%)	144/172 (84%)
Sidechain	0/1118 (0%)	0/649 (0%)	0/416 (0%)	0/53 (0%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	0/127 (0%)	0/69 (0%)	0/55 (0%)	0/3 (0%)
Overall	288/2111 (14%)	144/1064 (14%)	0/819 (0%)	144/228 (63%)

6.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

6.1.5 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: