wwPDB NMR Structure Validation Summary Report (i)

Mar 6, 2022 – 12:11 PM EST

PDB ID	:	2KDU
Title	:	Structural basis of the Munc13-1/Ca2+-Calmodulin interaction: A novel 1-26 calmodulin binding motif with a bipartite binding mode
Authors	:	Rodriguez-Castaneda, F.A.; Maestre-Martinez, M.; Coudevylle, N.; Dimova, K.; Jahn, O.; Junge, H.; Becker, S.; Brose, N.; Carlomagno, T.; Griesinger, C.
Deposited on	:	2009-01-19

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 (i)) were used in the production of this report:

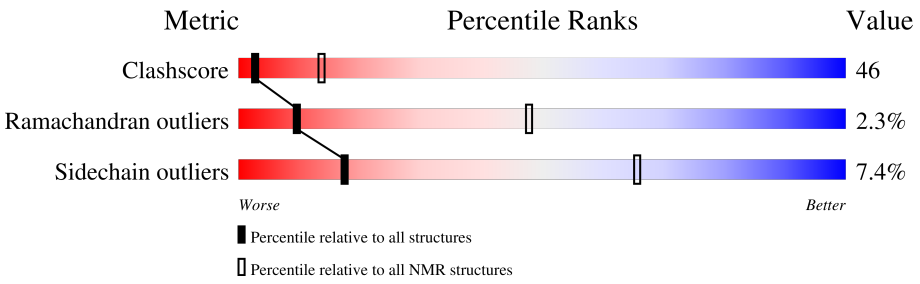
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.27
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.27

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 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	148	46% 41% • 9%
2	B	36	8% 33% 8% 50%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 3 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:4-A:74 (71)	0.55	20
2	A:84-A:146, B:459-B:476 (81)	0.43	3

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

3 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2856 atoms, of which 1393 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Calmodulin.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	148	2262	714	1096	188	255	9	0

- Molecule 2 is a protein called Protein unc-13 homolog A.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	36	590	183	297	58	50	2	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	457	GLY	-	expression tag	UNP Q62768

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

Mol	Chain	Residues	Atoms	
			Total	Ca
3	A	4	4	4

4 Residue-property plots (i)

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: Calmodulin

Chain A: 46% 41% • 9%

A1 D2 Q3 L4 Q8 I9 F12 F16 F19 I27 L32 V35 M36 R37 S38 L39 L48 M51 I52 N53 E54 V55 D56 A57 N60 G61 T62 I63 D64 F65 P66 L69 R74 K75 M76 K77 D78 T79 D80 S81 E82 E83 E84 I85 R86 E87 A88 F89

R90 V91 F92 D93 K94 D95 G98 Y99 I100 E104 L105 R106 H107 V108 M109 L112 G113 E114 K115 L116 V121 M124 I125 R126 E127 A128 D129 I130 V136 N137 Y138 E139 E140 F141 V142 Q143 M144 A147 K148

• Molecule 2: Protein unc-13 homolog A

Chain B: 8% 33% 8% 50%

G457 S458 R459 A460 K461 A462 N463 W464 L465 R466 A467 F468 N469 K470 V471 R472 M473 Q474 L475 Q476 E477 A478 R479 G480 E481 G482 E483 M484 S485 K486 S487 L488 W489 F490 K491 G492

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

The representative model is number 3. Colouring as in section 4.1 above.

• Molecule 1: Calmodulin

Chain A: 41% 45% • • 9%

A1 D2 Q3 L4 Q8 I9 F12 K13 F19 I27 T28 T29 L32 V35 M36 R37 S38 L39 G40 Q41 N42 P43 L48 M51 I52 V55 D56 A57 D58 G59 N60 I63 D64 F65 P66 E67 F68 L69 T70 M71 M72 A73 R74 K75 M76 K77 D78 T79 D80 S81

E82 E83 E84 I85 R86 E87 A88 F89 R90 V91 F92 D93 K94 D95 G96 N97 G98 Y99 I100 S101 A102 A103 E104 L105 R106 H107 V108 M109 L112 G113 E114 K115 L116 T117 E120 V121 M124 I125 R126 E127 A128 D129 I130 V136 N137 Y138 E139 E140 F141 V142 Q143 M144 A147 K148

• Molecule 2: Protein unc-13 homolog A

Chain B: 11% 31% 8% 50%

G457
S458
R459
A460
K461
A462
N463
W464
L465
R466
A467
F468
N469
K470
V471
R472
M473
Q474
L475
Q476
E477
A478
R479
G480
E481
G482
E483
M484
S485
K486
S487
L488
W489
F490
K491
G492

5 Refinement protocol and experimental data overview (i)

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
CYANA	geometry optimization	
CYANA	structure solution	
X-PLOR NIH	refinement	
CANDID	structure solution	

No chemical shift data was provided.

6 Model quality (i)

6.1 Standard geometry (i)

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

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the (average) root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	1.12±0.00	0±0/1069 (0.0± 0.0%)	1.21±0.01	0±0/1439 (0.0± 0.0%)
2	B	1.03±0.01	0±0/159 (0.0± 0.0%)	1.34±0.02	0±0/212 (0.0± 0.0%)
All	All	1.11	0/24560 (0.0%)	1.23	1/33020 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	Chirality	Planarity
1	A	0.0±0.0	6.0±0.0
2	B	0.0±0.0	3.0±0.0
All	All	0	180

There are no bond-length outliers.

All unique angle outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	51	MET	N-CA-CB	-5.21	101.22	110.60	14	1

There are no chirality outliers.

5 of 9 unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	37	ARG	Sidechain	20
1	A	74	ARG	Sidechain	20
1	A	86	ARG	Sidechain	20
1	A	90	ARG	Sidechain	20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	106	ARG	Sidechain	20

6.2 Too-close contacts (i)

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	1056	992	992	102±7
2	B	156	167	167	45±6
All	All	24320	23180	23180	2181

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 46.

5 of 694 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:88:ALA:HB2	2:B:475:LEU:HD11	1.00	1.25	11	7
1:A:55:VAL:HG11	1:A:63:ILE:HD13	1.00	1.33	16	1
1:A:4:LEU:HD23	1:A:73:ALA:HB2	0.98	1.36	8	1
1:A:88:ALA:HB2	2:B:475:LEU:HD21	0.97	1.33	13	12
1:A:128:ALA:HB2	2:B:464:TRP:CZ2	0.95	1.96	2	19

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	134/148 (91%)	120±3 (90±2%)	10±3 (8±2%)	3±1 (2±1%)	9 45
2	B	18/36 (50%)	16±1 (86±4%)	2±1 (12±4%)	0±0 (1±2%)	18 66
All	All	3040/3680 (83%)	2715 (89%)	255 (8%)	70 (2%)	9 48

5 of 19 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	129	ASP	12
1	A	4	LEU	10
1	A	95	ASP	9
1	A	114	GLU	7
1	A	41	GLN	5

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	114/126 (90%)	106±3 (93±2%)	8±3 (7±2%)	19 68
2	B	15/28 (54%)	13±1 (88±5%)	2±1 (12±5%)	8 51
All	All	2580/3080 (84%)	2388 (93%)	192 (7%)	17 65

5 of 46 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	112	LEU	19
2	B	474	GLN	16
1	A	35	VAL	13
1	A	144	MET	10
2	B	464	TRP	10

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 monosaccharides in this entry.

6.6 Ligand geometry (i)

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

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 (i)

No chemical shift data were provided