



# wwPDB NMR Structure Validation Summary Report ⓘ

Jun 12, 2024 – 04:28 AM EDT

PDB ID : 1Q80  
BMRB ID : 4129  
Title : Solution structure and dynamics of Nereis sarcoplasmic calcium binding protein  
Authors : Rabah, G.; Popescu, R.; Cox, J.A.; Engelborghs, Y.; Craescu, C.T.  
Deposited on : 2003-08-20

This is a wwPDB NMR Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) 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)  
wwPDB-RCI : v\_1n\_11\_5\_13\_A (Berjanski et al., 2005)  
PANAV : Wang et al. (2010)  
wwPDB-ShiftChecker : v1.2  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36.2

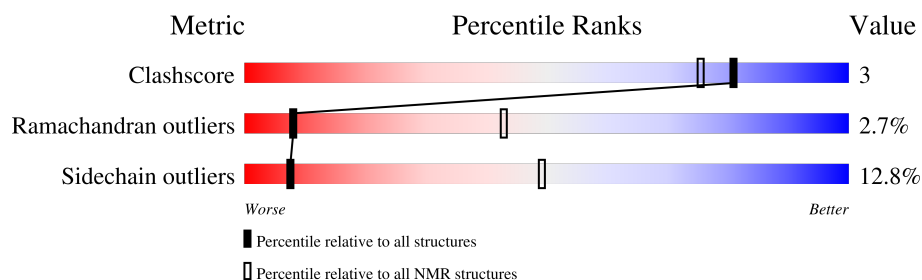
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*


The overall completeness of chemical shifts assignment is 52%.

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  $\geq 3$ , 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	174	

## 2 Ensemble composition and analysis

This entry contains 17 models. Model 9 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *closest to the average*.

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:1-A:174 (174)	1.51	9

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

### 3 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 2678 atoms, of which 1310 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Sarcoplasmic calcium-binding protein.

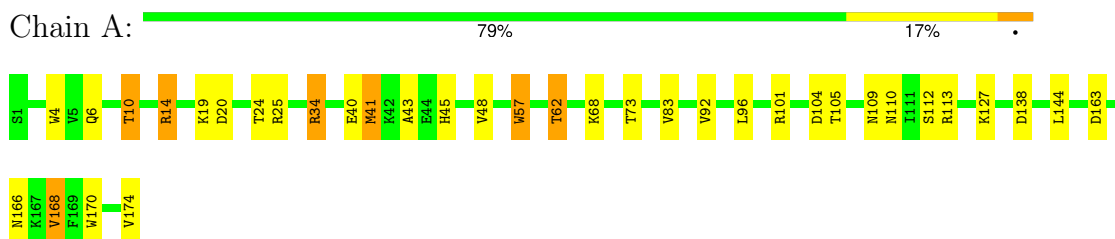
Mol	Chain	Residues	Atoms						Trace
1	A	174	Total	C	H	N	O	S	0
			2678	868	1310	218	272	10	

## 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.

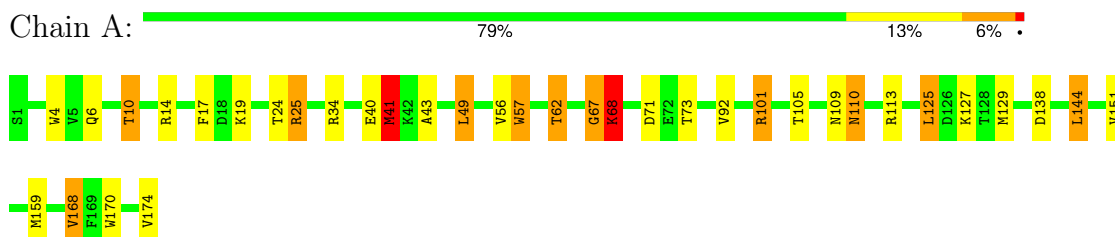
- Molecule 1: Sarcoplasmic calcium-binding protein



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

The representative model is number 9. Colouring as in section 4.1 above.

- Molecule 1: Sarcoplasmic calcium-binding protein



## 5 Refinement protocol and experimental data overview

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

Of the 17 calculated structures, 17 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
DGII	structure solution	2
Discover	structure solution	3
Discover	refinement	3

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)	working_cs.cif
Number of chemical shift lists	3
Total number of shifts	1210
Number of shifts mapped to atoms	1195
Number of unparsed shifts	0
Number of shifts with mapping errors	15
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	52%

## 6 Model quality i

### 6.1 Standard geometry i

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	1.28±0.01	1±0/1396 ( 0.1± 0.0%)	1.92±0.29	24±5/1880 ( 1.3± 0.3%)
All	All	1.28	17/23732 ( 0.1%)	1.94	409/31960 ( 1.3%)

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	5.1±2.5
All	All	0	86

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	174	VAL	C-OXT	7.58	1.37	1.23	17	17

5 of 127 unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	43	ALA	O-C-N	-38.28	61.45	122.70	15	6
1	A	41	MET	O-C-N	-38.05	61.83	122.70	15	6
1	A	64	VAL	O-C-N	-37.76	62.28	122.70	14	1
1	A	68	LYS	O-C-N	-36.27	61.55	123.20	10	3
1	A	165	THR	O-C-N	-36.04	65.03	122.70	13	1

There are no chirality outliers.

5 of 34 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	168	VAL	Mainchain	8
1	A	25	ARG	Sidechain	7
1	A	34	ARG	Sidechain	6
1	A	43	ALA	Mainchain	6
1	A	41	MET	Mainchain	6

## 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	1368	1310	1310	7±2
All	All	23256	22270	22253	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 3.

5 of 39 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:136:ALA:HB2	1:A:165:THR:OG1	0.89	1.68	5	3
1:A:6:GLN:O	1:A:10:THR:OG1	0.81	1.97	14	11
1:A:41:MET:CB	1:A:105:THR:OG1	0.73	2.36	15	8
1:A:57:TRP:O	1:A:62:THR:OG1	0.68	2.11	7	17
1:A:41:MET:SD	1:A:105:THR:OG1	0.68	2.50	10	3

## 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	172/174 (99%)	154±2 (90±1%)	13±3 (8±2%)	5±1 (3±1%)	8	43
All	All	2924/2958 (99%)	2624 (90%)	222 (8%)	78 (3%)	8	43

5 of 25 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	43	ALA	12
1	A	138	ASP	12
1	A	104	ASP	6
1	A	163	ASP	6
1	A	164	SER	6

### 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	148/148 (100%)	129±4 (87±3%)	19±4 (13±3%)	7	49
All	All	2516/2516 (100%)	2195 (87%)	321 (13%)	7	49

5 of 81 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	73	THR	16
1	A	20	ASP	15
1	A	10	THR	14
1	A	19	LYS	14
1	A	24	THR	14

### 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](#)

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 i

The completeness of assignment taking into account all chemical shift lists is 52% for the well-defined parts and 52% for the entire structure.

### 7.1 Chemical shift list 1

File name: working\_cs.cif

Chemical shift list name: *assigned\_chem\_shift\_list\_1*

#### 7.1.1 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	1195
Number of shifts mapped to atoms	1195
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	5

#### 7.1.2 Chemical shift referencing i

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	0	—	None (insufficient data)
$^{13}\text{C}_\beta$	0	—	None (insufficient data)
$^{13}\text{C}'$	0	—	None (insufficient data)
$^{15}\text{N}$	165	$2.68 \pm 0.22$	Should be applied

#### 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 52%, i.e. 1195 atoms were assigned a chemical shift out of a possible 2303. 0 out of 23 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	511/873 (59%)	346/356 (97%)	0/348 (0%)	165/169 (98%)
Sidechain	582/1219 (48%)	571/789 (72%)	0/391 (0%)	11/39 (28%)

*Continued on next page...*

Continued from previous page...

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Aromatic	102/211 (48%)	99/105 (94%)	0/102 (0%)	3/4 (75%)
Overall	1195/2303 (52%)	1016/1250 (81%)	0/841 (0%)	179/212 (84%)

#### 7.1.4 Statistically unusual chemical shifts [i](#)

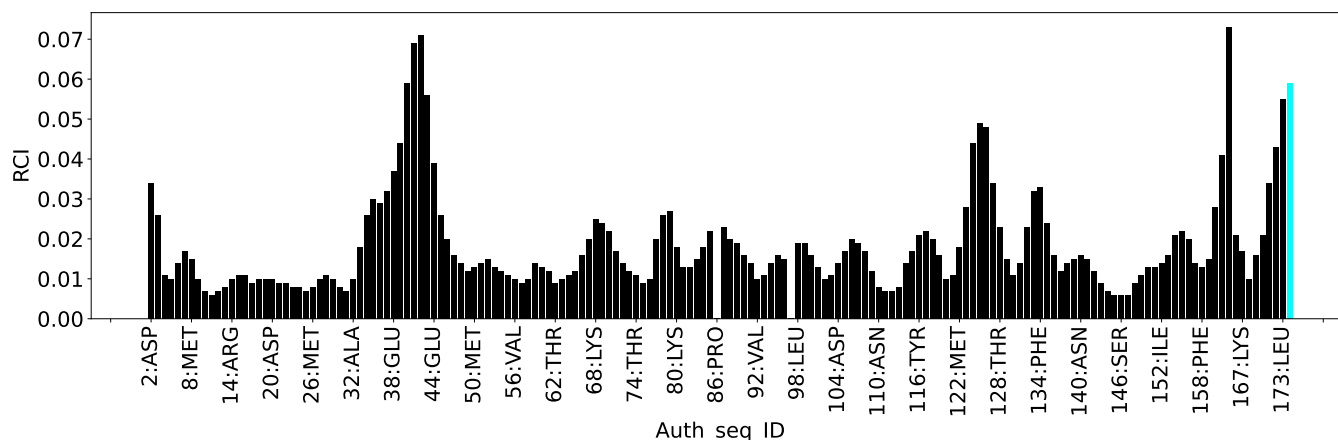
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	25	ARG	NE	115.80	76.53 – 92.65	19.4
1	A	14	ARG	NE	115.50	76.53 – 92.65	19.2
1	A	113	ARG	HG2	-0.06	0.26 – 2.87	-6.2
1	A	25	ARG	HB3	0.22	0.43 – 3.11	-5.8
1	A	154	GLY	HA3	1.81	2.08 – 5.71	-5.7

#### 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. If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:



## 7.2 Chemical shift list 2

File name: working\_cs.cif

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	4
Number of shifts mapped to atoms	0
Number of unparsed shifts	0
Number of shifts with mapping errors	4
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

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

- No matching atom found in the structure. All 4 occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
2	UNMAPPED	1	SAC	H3	1.84	0.01	1
2	UNMAPPED	1	SAC	H	9.54	0.01	1
2	UNMAPPED	1	SAC	HA	4.4	0.01	1
2	UNMAPPED	1	SAC	HB2	4.33	0.01	2

### 7.2.2 Chemical shift referencing [i](#)

No chemical shift referencing corrections were calculated (not enough data).

### 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 0%, i.e. 0 atoms were assigned a chemical shift out of a possible 2303. 0 out of 23 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Backbone	0/873 (0%)	0/356 (0%)	0/348 (0%)	0/169 (0%)
Sidechain	0/1219 (0%)	0/789 (0%)	0/391 (0%)	0/39 (0%)

*Continued on next page...*

*Continued from previous page...*

	<b>Total</b>	<b><sup>1</sup>H</b>	<b><sup>13</sup>C</b>	<b><sup>15</sup>N</b>
Aromatic	0/211 (0%)	0/105 (0%)	0/102 (0%)	0/4 (0%)
Overall	0/2303 (0%)	0/1250 (0%)	0/841 (0%)	0/212 (0%)

#### 7.2.4 Statistically unusual chemical shifts [i](#)

There are no statistically unusual chemical shifts.

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

No *random coil index*(RCI) plot could be generated from the current chemical shift list. RCI is only applicable to proteins

### 7.3 Chemical shift list 3

File name: working\_cs.cif

Chemical shift list name: *assigned\_chem\_shift\_list\_3*

#### 7.3.1 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	11
Number of shifts mapped to atoms	0
Number of unparsed shifts	0
Number of shifts with mapping errors	11
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

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

- No matching atom found in the structure. First 5 (of 11) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
3	UNMAPPED	5	VAL	H	9.08	0.01	1
3	UNMAPPED	7	LYS	H	8.81	0.01	1
3	UNMAPPED	8	MET	H	8.24	0.01	1
3	UNMAPPED	9	LYS	H	8.95	0.01	1

*Continued on next page...*

*Continued from previous page...*

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
3	UNMAPPED	80	LYS	H	8.73	0.01	1
3	UNMAPPED	84	LYS	H	6.88	0.01	1
3	UNMAPPED	114	ASP	H	8.34	0.01	1
3	UNMAPPED	159	MET	H	8.83	0.01	1
3	UNMAPPED	160	ASN	H	8.35	0.01	1
3	UNMAPPED	161	ASP	H	9.27	0.01	1
3	UNMAPPED	162	GLY	H	8.83	0.01	1

### 7.3.2 Chemical shift referencing [i](#)

No chemical shift referencing corrections were calculated (not enough data).

### 7.3.3 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 0%, i.e. 0 atoms were assigned a chemical shift out of a possible 2303. 0 out of 23 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Backbone	0/873 (0%)	0/356 (0%)	0/348 (0%)	0/169 (0%)
Sidechain	0/1219 (0%)	0/789 (0%)	0/391 (0%)	0/39 (0%)
Aromatic	0/211 (0%)	0/105 (0%)	0/102 (0%)	0/4 (0%)
Overall	0/2303 (0%)	0/1250 (0%)	0/841 (0%)	0/212 (0%)

### 7.3.4 Statistically unusual chemical shifts [i](#)

There are no statistically unusual chemical shifts.

### 7.3.5 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. If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain UNMAPPED:

