



# Full wwPDB NMR Structure Validation Report ⓘ

May 29, 2020 – 12:30 am BST

PDB ID : 2N8Y  
Title : Holo form of Calmodulin-Like Domain of Human Non-Muscle alpha-actinin 1  
Authors : Drmota Prebil, S.; Slapsak, U.; de Almeida Ribeiro, E.; Pavsic, M.; Ilc, G.; Zielinska, K.; Hartl, M.; Backman, L.; Plavec, J.; Lenarcic, B.; Djinovic-Carugo, K.  
Deposited on : 2015-10-28

This is a Full wwPDB NMR Structure Validation 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 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  
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

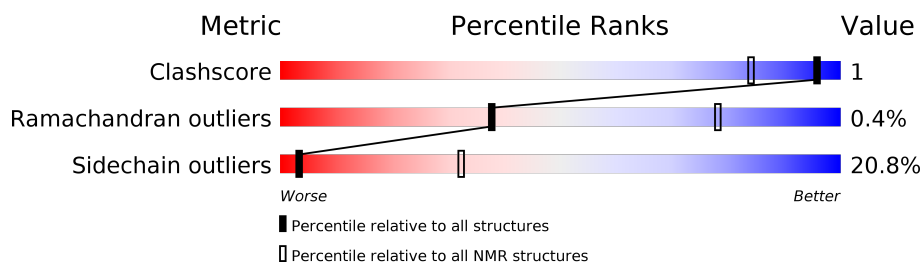
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 84%.

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	153	<div>74%14%12%</div>

## 2 Ensemble composition and analysis

This entry contains 20 models. Model 6 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:742-A:867, A:875-A:882 (134)	0.43	6

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 2 single-model clusters were found.

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

### 3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2305 atoms, of which 1114 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Alpha-actinin-1.

Mol	Chain	Residues	Atoms						Trace
1	A	153	Total	C	H	N	O	S	0
			2304	742	1114	195	244	9	

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	740	GLY	-	EXPRESSION TAG	UNP P12814
A	741	SER	-	EXPRESSION TAG	UNP P12814
A	742	SER	-	EXPRESSION TAG	UNP P12814

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

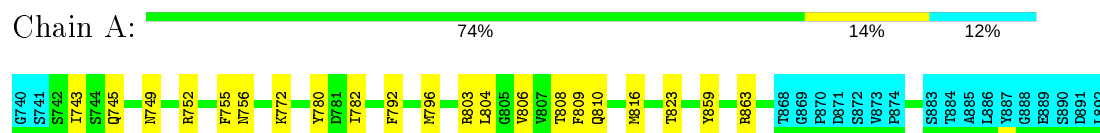
Mol	Chain	Residues	Atoms	
2	A	1	Total	Ca
			1	1

## 4 Residue-property plots

### 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: Alpha-actinin-1

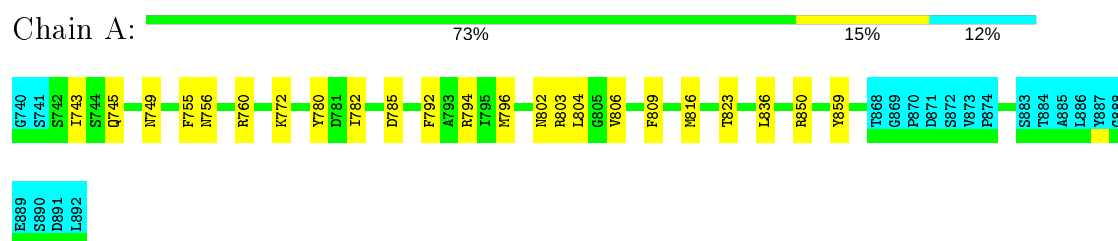


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

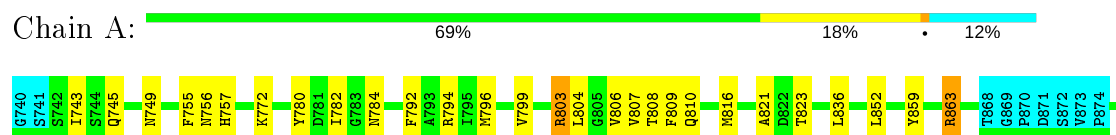
#### 4.2.1 Score per residue for model 1

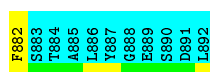
- Molecule 1: Alpha-actinin-1



#### 4.2.2 Score per residue for model 2

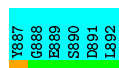
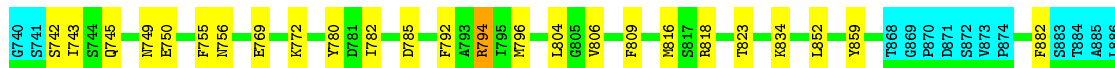
- Molecule 1: Alpha-actinin-1





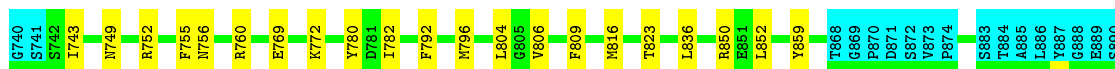
### 4.2.3 Score per residue for model 3

- Molecule 1: Alpha-actinin-1



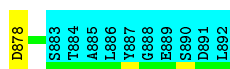
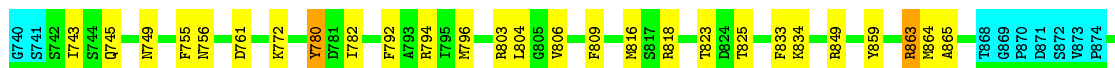
### 4.2.4 Score per residue for model 4

- Molecule 1: Alpha-actinin-1



### 4.2.5 Score per residue for model 5

- Molecule 1: Alpha-actinin-1



### 4.2.6 Score per residue for model 6 (medoid)

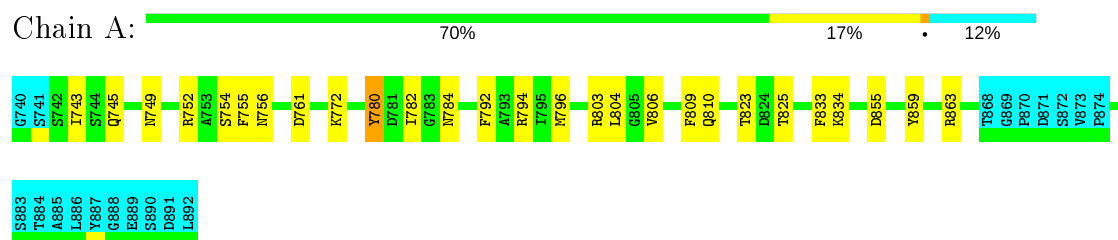
- Molecule 1: Alpha-actinin-1





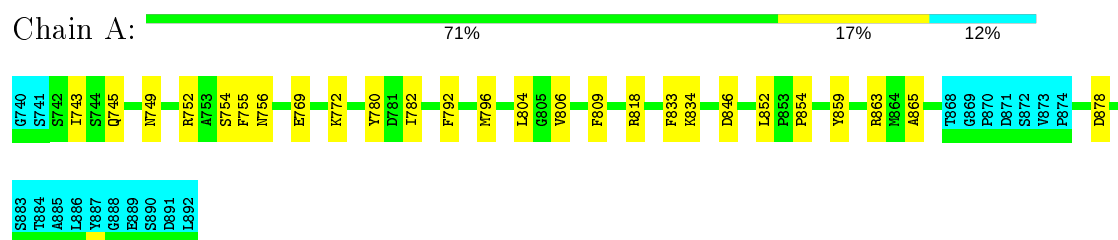
#### 4.2.7 Score per residue for model 7

- Molecule 1: Alpha-actinin-1



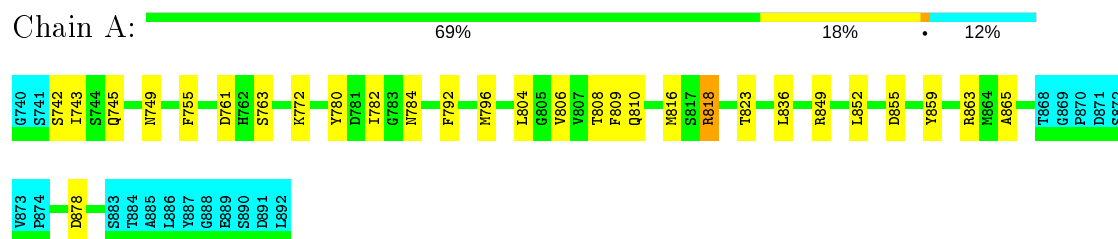
#### 4.2.8 Score per residue for model 8

- Molecule 1: Alpha-actinin-1



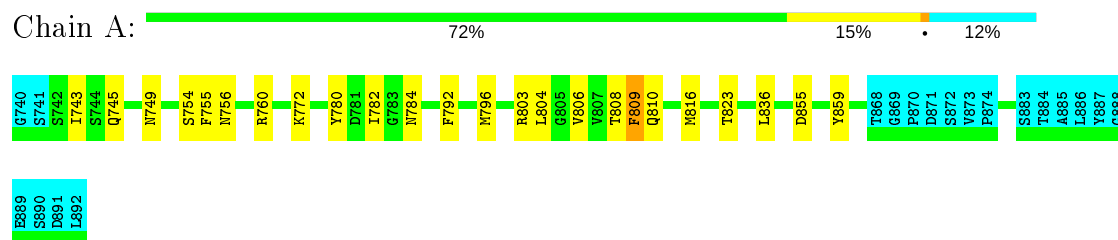
#### 4.2.9 Score per residue for model 9

- Molecule 1: Alpha-actinin-1



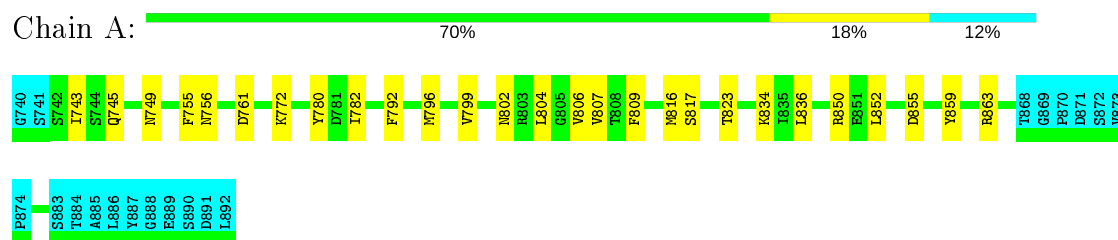
#### 4.2.10 Score per residue for model 10

- Molecule 1: Alpha-actinin-1



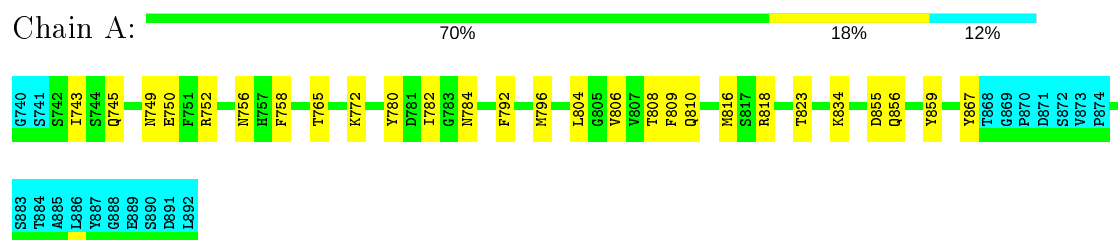
#### 4.2.11 Score per residue for model 11

- Molecule 1: Alpha-actinin-1



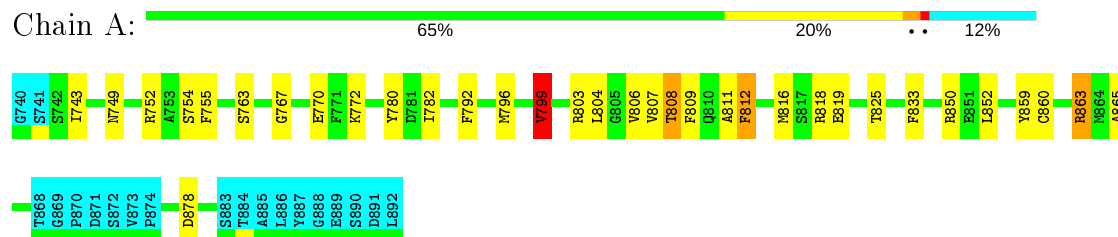
#### 4.2.12 Score per residue for model 12

- Molecule 1: Alpha-actinin-1



#### 4.2.13 Score per residue for model 13

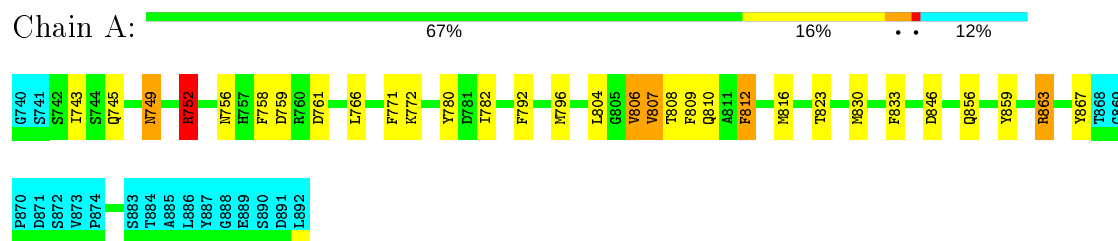
- Molecule 1: Alpha-actinin-1





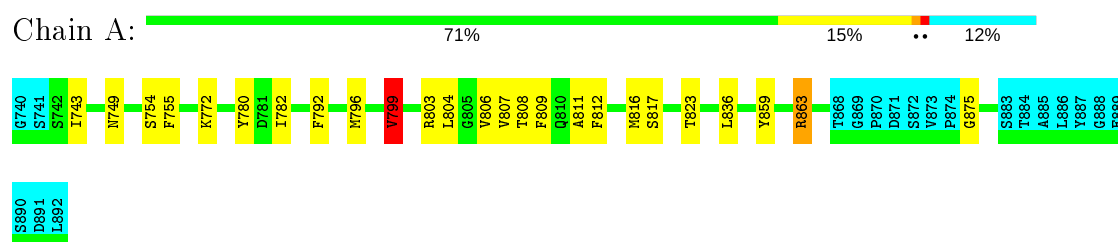
## 4.2.14 Score per residue for model 14

- Molecule 1: Alpha-actinin-1



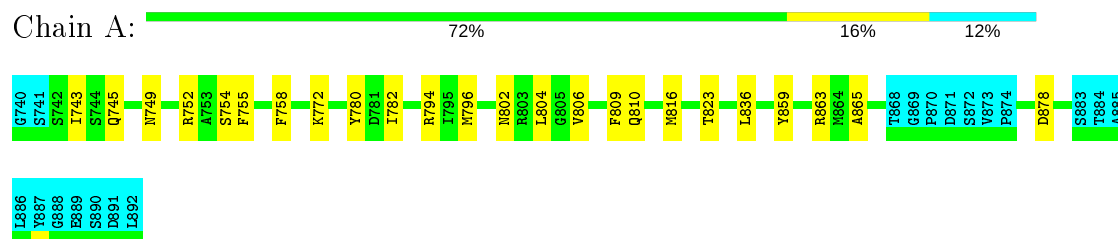
## 4.2.15 Score per residue for model 15

- Molecule 1: Alpha-actinin-1



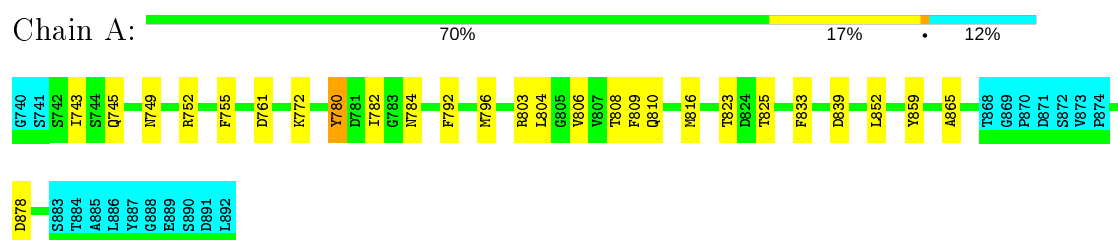
## 4.2.16 Score per residue for model 16

- Molecule 1: Alpha-actinin-1



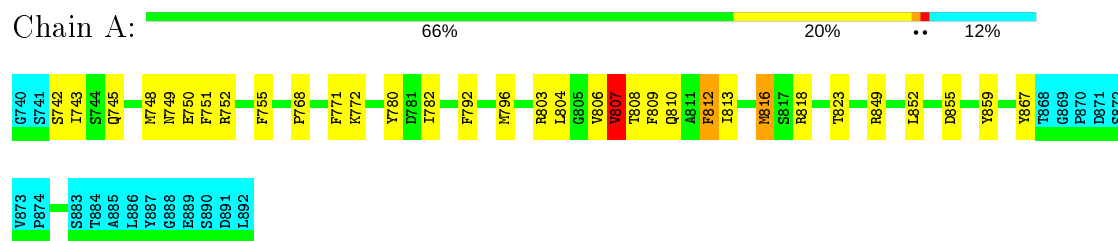
## 4.2.17 Score per residue for model 17

- Molecule 1: Alpha-actinin-1



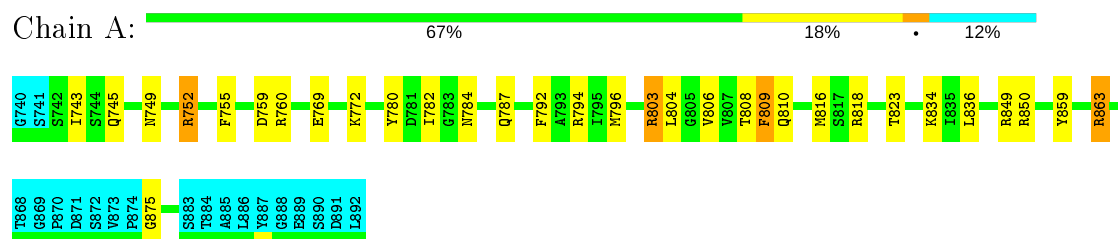
### 4.2.18 Score per residue for model 18

- Molecule 1: Alpha-actinin-1



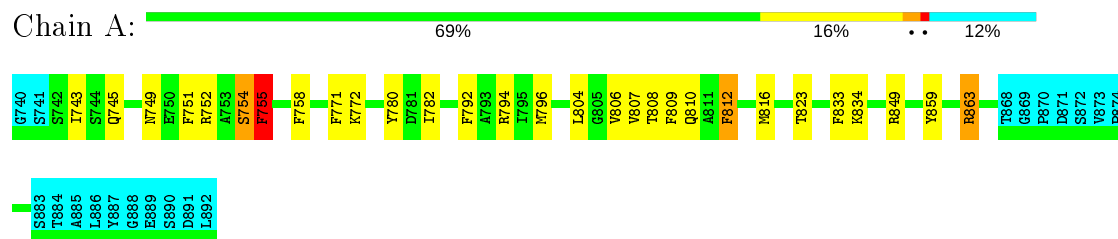
### 4.2.19 Score per residue for model 19

- Molecule 1: Alpha-actinin-1



### 4.2.20 Score per residue for model 20

- Molecule 1: Alpha-actinin-1



## 5 Refinement protocol and experimental data overview

The models were refined using the following method: *simulated annealing, water refinement*.

Of the 20 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	structure solution	3.0
YASARA	refinement	
CYANA	refinement	3.0

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	1690
Number of shifts mapped to atoms	1690
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	84%

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	1060	996	995	2±2
All	All	21220	19920	19900	39

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 1.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:799:VAL:HA	1:A:811:ALA:HB1	0.91	1.42	15	2
1:A:771:PHE:HB2	1:A:807:VAL:HG11	0.70	1.62	14	1
1:A:807:VAL:CG2	1:A:812:PHE:CZ	0.65	2.80	14	1
1:A:799:VAL:CA	1:A:811:ALA:HB1	0.64	2.20	15	2
1:A:752:ARG:O	1:A:755:PHE:HB3	0.57	1.99	20	1
1:A:771:PHE:CD2	1:A:807:VAL:HG11	0.55	2.36	18	1
1:A:771:PHE:CB	1:A:807:VAL:HG11	0.54	2.33	14	1
1:A:766:LEU:HD23	1:A:812:PHE:HZ	0.52	1.64	14	1
1:A:771:PHE:HD2	1:A:812:PHE:CE1	0.51	2.22	18	1
1:A:749:ASN:HA	1:A:752:ARG:HD2	0.50	1.83	14	1
1:A:799:VAL:HG13	1:A:807:VAL:HG22	0.48	1.84	15	2
1:A:766:LEU:O	1:A:806:VAL:HG12	0.48	2.09	14	1
1:A:771:PHE:CG	1:A:807:VAL:HG11	0.47	2.44	14	1
1:A:865:ALA:HB3	1:A:878:ASP:HB3	0.47	1.86	5	7
1:A:799:VAL:HG22	1:A:811:ALA:C	0.46	2.31	13	2
1:A:799:VAL:HG23	1:A:807:VAL:HG22	0.43	1.89	11	2
1:A:780:TYR:CD1	1:A:780:TYR:N	0.42	2.87	17	1
1:A:748:MET:HE2	1:A:813:ILE:HG21	0.42	1.90	18	1
1:A:771:PHE:CD2	1:A:812:PHE:CE1	0.42	3.06	18	1
1:A:808:THR:O	1:A:811:ALA:HB3	0.42	2.15	13	1
1:A:751:PHE:CD2	1:A:816:MET:HG3	0.42	2.50	18	1
1:A:852:LEU:C	1:A:854:PRO:HD2	0.41	2.36	8	1
1:A:771:PHE:CG	1:A:807:VAL:HG21	0.41	2.51	20	1
1:A:751:PHE:CE2	1:A:812:PHE:HB3	0.41	2.51	20	1
1:A:767:GLY:H	1:A:770:GLU:CD	0.41	2.19	13	1
1:A:771:PHE:HB2	1:A:807:VAL:CG1	0.41	2.42	14	1
1:A:780:TYR:N	1:A:780:TYR:CD1	0.41	2.87	5	2

## 5.2 Torsion angles ⓘ

### 5.2.1 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	134/153 (88%)	123±1 (92±1%)	10±2 (8±1%)	1±1 (0±1%)	38	78
All	All	2680/3060 (88%)	2465 (92%)	204 (8%)	11 (0%)	38	78

All 6 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	742	SER	3
1	A	759	ASP	2
1	A	875	GLY	2
1	A	812	PHE	2
1	A	821	ALA	1
1	A	768	PRO	1

### 5.2.2 Protein sidechains ⓘ

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	113/128 (88%)	90±2 (79±2%)	23±2 (21±2%)	3	32
All	All	2260/2560 (88%)	1791 (79%)	469 (21%)	3	32

All 57 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	749	ASN	20
1	A	780	TYR	20
1	A	782	ILE	20
1	A	772	LYS	20
1	A	804	LEU	20
1	A	859	TYR	20
1	A	743	ILE	20
1	A	806	VAL	20
1	A	796	MET	20
1	A	809	PHE	19
1	A	792	PHE	19
1	A	823	THR	18
1	A	816	MET	18
1	A	755	PHE	18
1	A	745	GLN	17
1	A	808	THR	12
1	A	863	ARG	12
1	A	810	GLN	12

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	756	ASN	11
1	A	752	ARG	11
1	A	836	LEU	10
1	A	834	LYS	9
1	A	852	LEU	9
1	A	784	ASN	8
1	A	754	SER	8
1	A	761	ASP	7
1	A	833	PHE	7
1	A	855	ASP	6
1	A	825	THR	4
1	A	802	ASN	4
1	A	769	GLU	4
1	A	758	PHE	4
1	A	750	GLU	3
1	A	803	ARG	3
1	A	760	ARG	3
1	A	867	TYR	3
1	A	846	ASP	2
1	A	785	ASP	2
1	A	818	ARG	2
1	A	856	GLN	2
1	A	817	SER	2
1	A	799	VAL	2
1	A	850	ARG	2
1	A	882	PHE	2
1	A	763	SER	2
1	A	787	GLN	1
1	A	864	MET	1
1	A	849	ARG	1
1	A	765	THR	1
1	A	757	HIS	1
1	A	839	ASP	1
1	A	742	SER	1
1	A	847	GLU	1
1	A	819	GLU	1
1	A	860	CYS	1
1	A	794	ARG	1
1	A	830	MET	1

### 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 1 ligands modelled in this entry, 1 is 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 84% for the well-defined parts and 83% 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	1690
Number of shifts mapped to atoms	1690
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 $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	151	$-0.10 \pm 0.08$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}_\beta$	141	$0.34 \pm 0.10$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}'$	0	—	None (insufficient data)
$^{15}\text{N}$	140	$0.26 \pm 0.35$	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 84%, i.e. 1360 atoms were assigned a chemical shift out of a possible 1623. 12 out of 12 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	522/658 (79%)	261/262 (100%)	134/268 (50%)	127/128 (99%)
Sidechain	700/821 (85%)	437/483 (90%)	252/300 (84%)	11/38 (29%)

*Continued on next page...*



Continued from previous page...

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Aromatic	138/144 (96%)	74/78 (95%)	64/64 (100%)	0/2 (0%)
Overall	1360/1623 (84%)	772/823 (94%)	450/632 (71%)	138/168 (82%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 83%, i.e. 1500 atoms were assigned a chemical shift out of a possible 1804. 15 out of 15 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Backbone	582/749 (78%)	291/298 (98%)	151/306 (49%)	140/145 (97%)
Sidechain	772/903 (85%)	481/531 (91%)	280/334 (84%)	11/38 (29%)
Aromatic	146/152 (96%)	78/82 (95%)	68/68 (100%)	0/2 (0%)
Overall	1500/1804 (83%)	850/911 (93%)	499/708 (70%)	151/185 (82%)

#### 6.1.4 Statistically unusual chemical shifts ⓘ

There are no statistically unusual chemical shifts.

#### 6.1.5 Random Coil Index (RCI) plots ⓘ

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:

