



# Full wwPDB NMR Structure Validation Report ⓘ

May 28, 2020 – 10:41 pm BST

PDB ID : 2KMV  
Title : Solution structure of the nucleotide binding domain of the human Menkes protein in the ATP-free form  
Authors : Banci, L.; Bertini, I.; Cantini, F.; Inagaki, S.; Migliardi, M.; Rosato, A.  
Deposited on : 2009-08-05

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

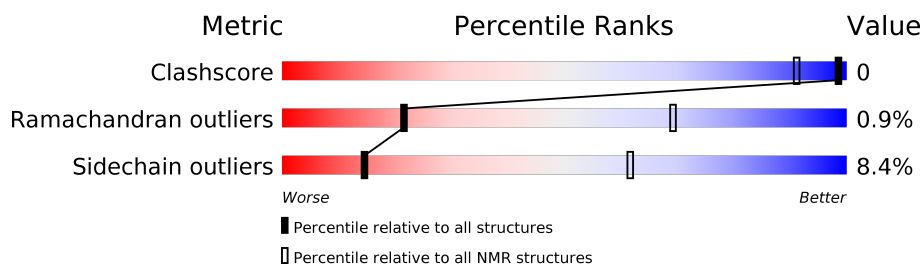
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 82%.

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	185	

## 2 Ensemble composition and analysis

This entry contains 20 models. Model 1 is the overall representative, medoid model (most similar to other models).

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:1055-A:1125, A:1177-A:1229 (124)	0.62	1

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 3 single-model clusters were found.

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 2820 atoms, of which 1405 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Copper-transporting ATPase 1.

Mol	Chain	Residues	Atoms						Trace
1	A	185	Total	C	H	N	O	S	0
			2820	872	1405	249	285	9	

There are 4 discrepancies between the modelled and reference sequences:

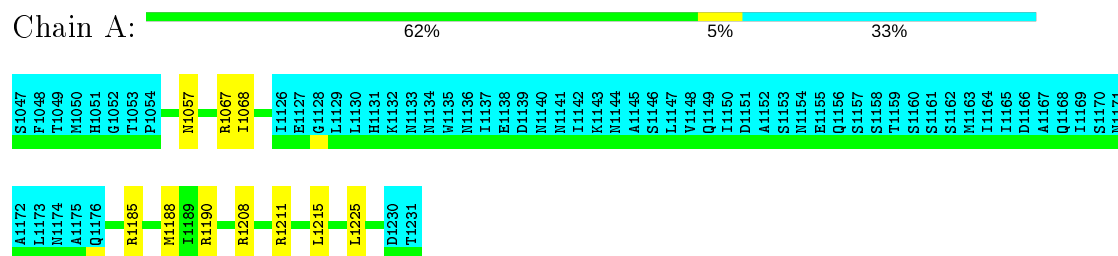
Chain	Residue	Modelled	Actual	Comment	Reference
A	1047	SER	-	expression tag	UNP Q04656
A	1048	PHE	-	expression tag	UNP Q04656
A	1049	THR	-	expression tag	UNP Q04656
A	1050	MET	-	expression tag	UNP Q04656

## 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: Copper-transporting ATPase 1

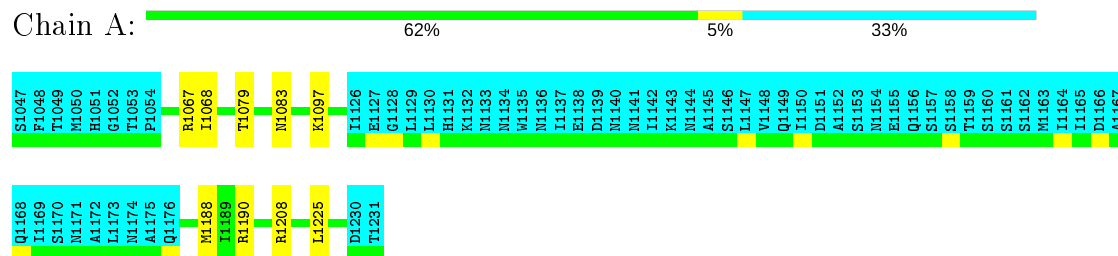


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

#### 4.2.1 Score per residue for model 1 (medoid)

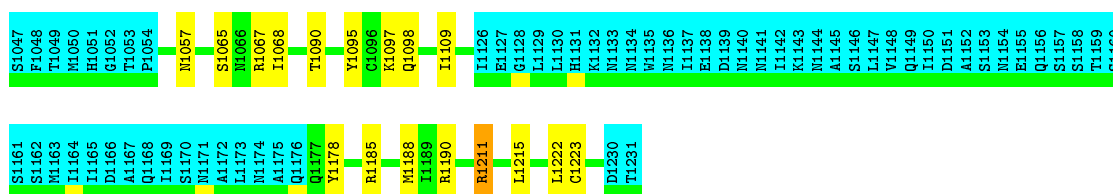
- Molecule 1: Copper-transporting ATPase 1



#### 4.2.2 Score per residue for model 2

- Molecule 1: Copper-transporting ATPase 1

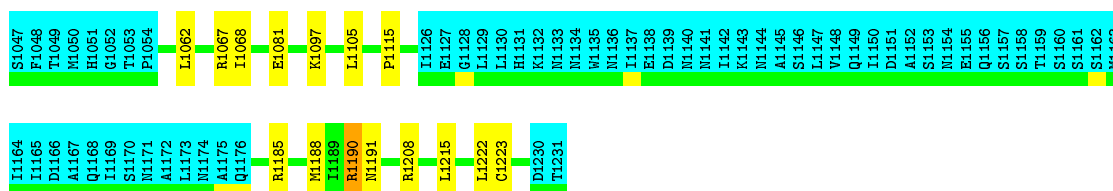




### 4.2.3 Score per residue for model 3

- Molecule 1: Copper-transporting ATPase 1

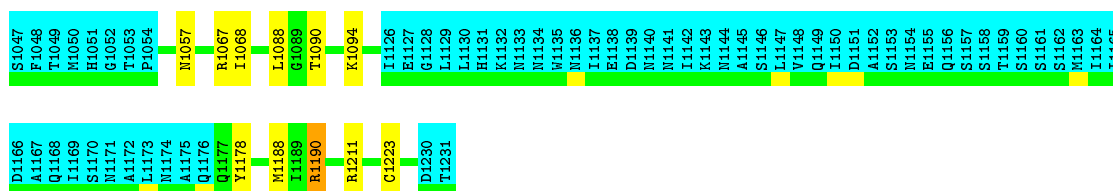
Chain A: 59% 8% 33%



### 4.2.4 Score per residue for model 4

- Molecule 1: Copper-transporting ATPase 1

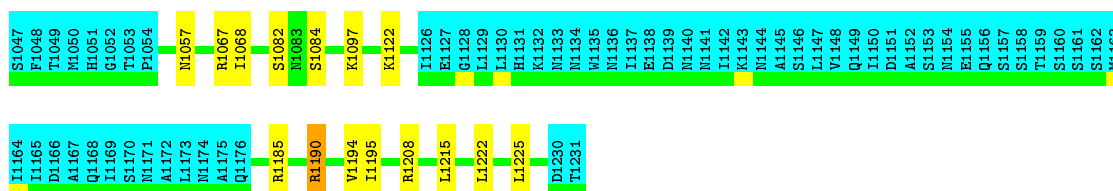
Chain A: 61% 5% 33%



### 4.2.5 Score per residue for model 5

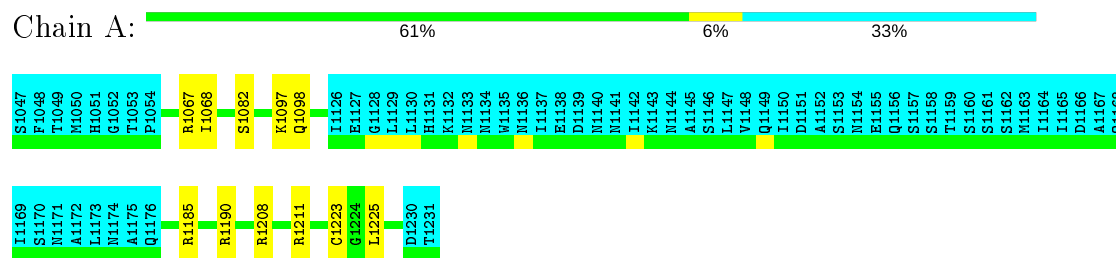
- Molecule 1: Copper-transporting ATPase 1

Chain A: 59% 8% 33%

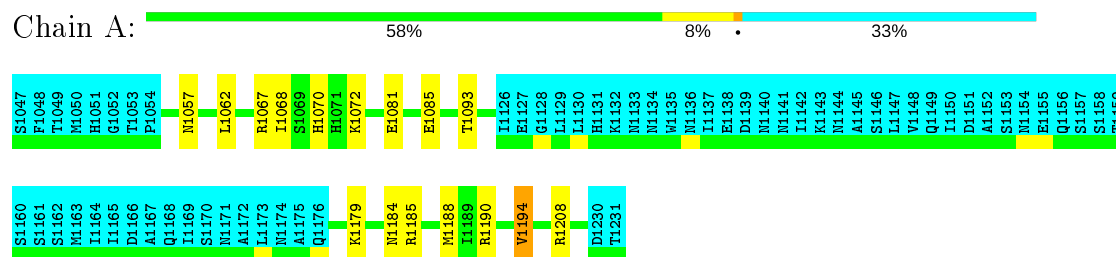


### 4.2.6 Score per residue for model 6

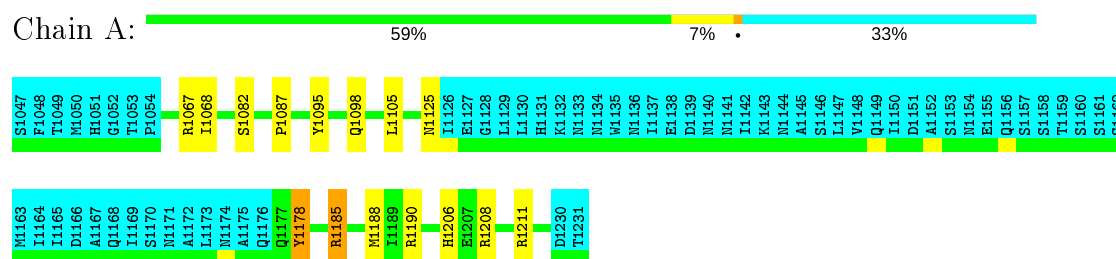
- Molecule 1: Copper-transporting ATPase 1



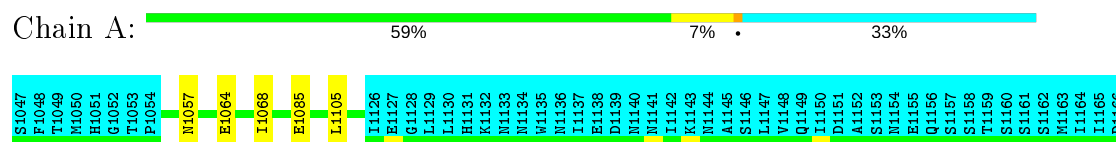
- Molecule 1: Copper-transporting ATPase 1



- Molecule 1: Copper-transporting ATPase 1



- Molecule 1: Copper-transporting ATPase 1

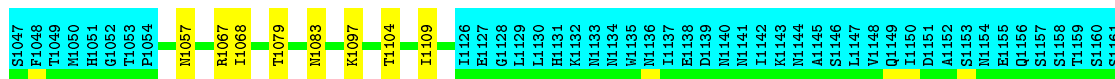




#### 4.2.10 Score per residue for model 10

- Molecule 1: Copper-transporting ATPase 1

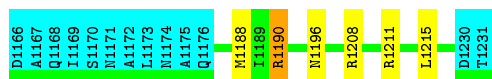
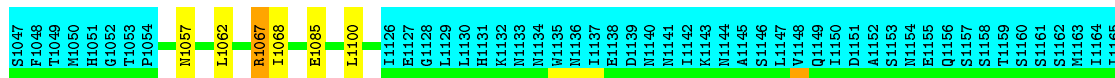
Chain A: 58% 8% 33%



#### 4.2.11 Score per residue for model 11

- Molecule 1: Copper-transporting ATPase 1

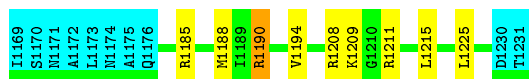
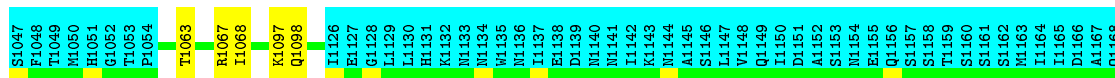
Chain A: 61% 5% 33%



#### 4.2.12 Score per residue for model 12

- Molecule 1: Copper-transporting ATPase 1

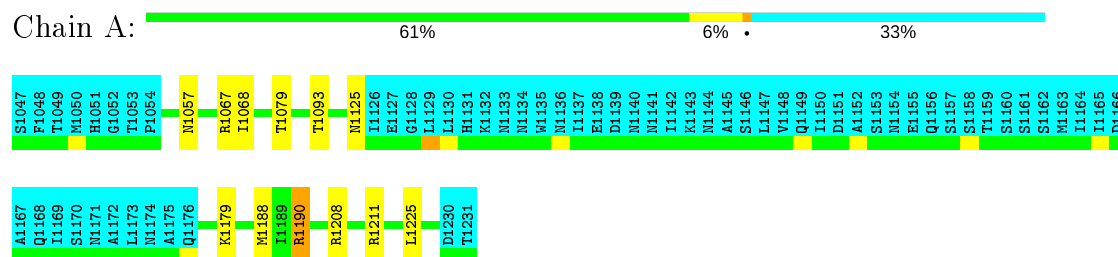
Chain A: 59% 7% 33%



#### 4.2.13 Score per residue for model 13

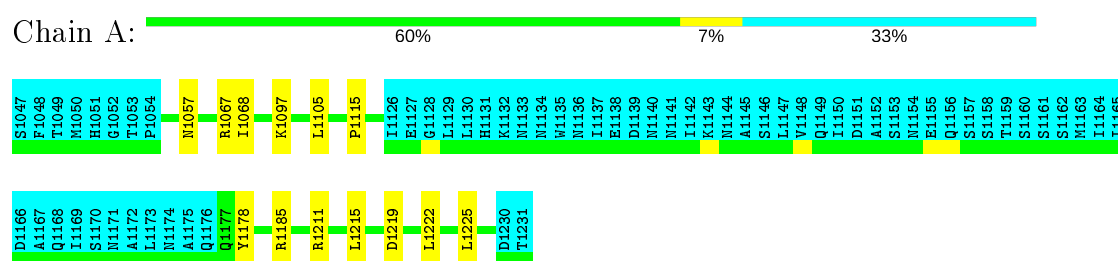
- Molecule 1: Copper-transporting ATPase 1





#### 4.2.14 Score per residue for model 14

- Molecule 1: Copper-transporting ATPase 1



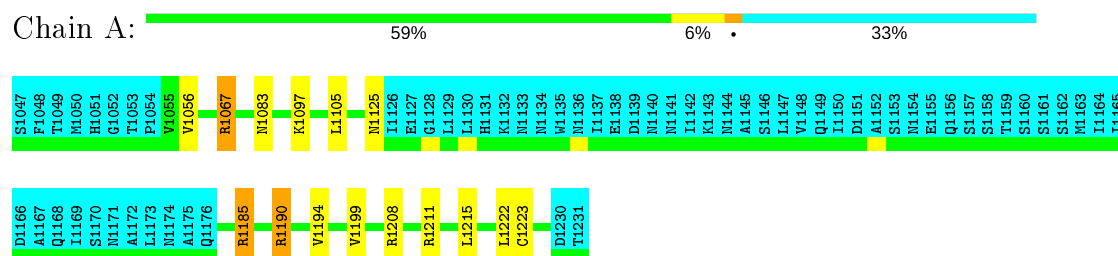
#### 4.2.15 Score per residue for model 15

- Molecule 1: Copper-transporting ATPase 1



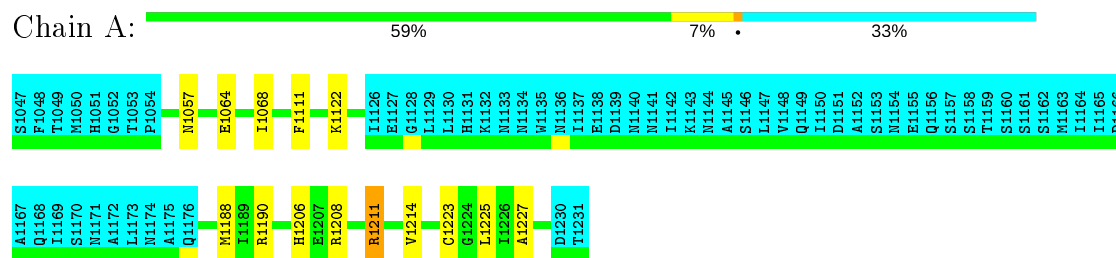
#### 4.2.16 Score per residue for model 16

- Molecule 1: Copper-transporting ATPase 1



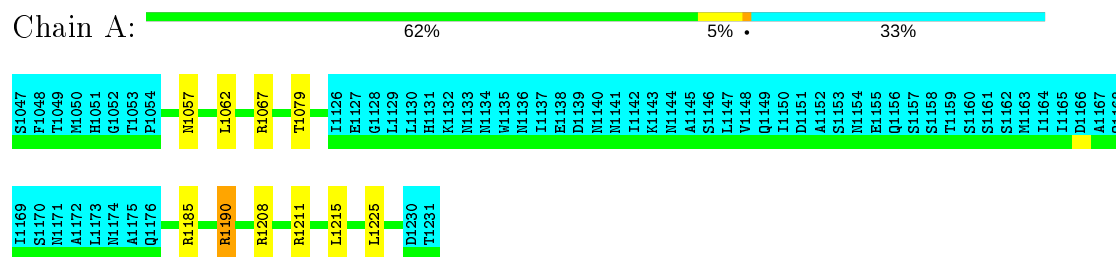
## 4.2.17 Score per residue for model 17

- Molecule 1: Copper-transporting ATPase 1



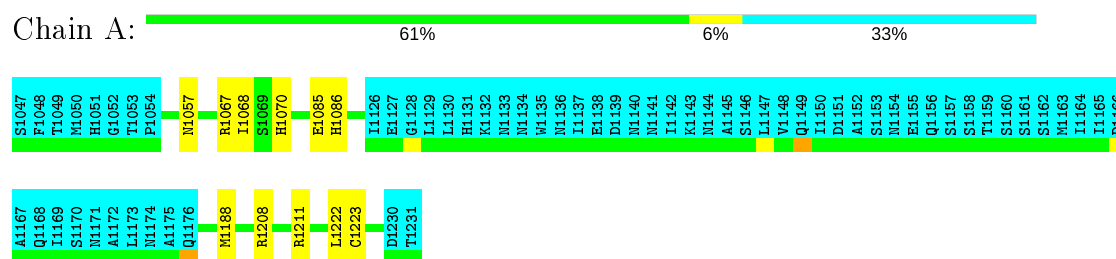
## 4.2.18 Score per residue for model 18

- Molecule 1: Copper-transporting ATPase 1



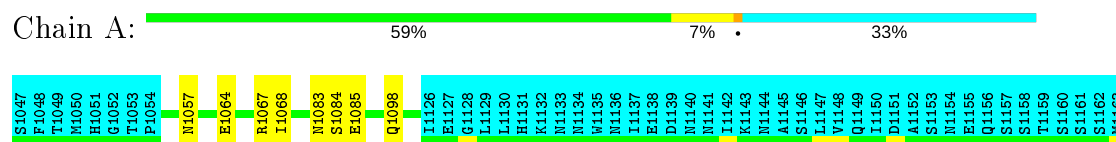
## 4.2.19 Score per residue for model 19

- Molecule 1: Copper-transporting ATPase 1



## 4.2.20 Score per residue for model 20

- Molecule 1: Copper-transporting ATPase 1



I1164	I1165	D1166	A1167	Q1168	I1169	S1170	R1171	A1172	L1173	N1174	A1175	Q1176	Q1177		R1185		R1190		R1211		L1215		L1222		L1225		D1230	T1231
-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	--	-------	-------

## 5 Refinement protocol and experimental data overview

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

Of the 500 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
CYANA	structure solution	
Amber	refinement	10

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	2021
Number of shifts mapped to atoms	2021
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	82%

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	952	965	965	0±0
All	All	19040	19300	19300	2

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 0.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:1206:HIS:CE1	1:A:1227:ALA:HB1	0.44	2.48	17	1
1:A:1062:LEU:HD23	1:A:1062:LEU:C	0.43	2.34	18	1

## 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	124/185 (67%)	116±1 (93±1%)	7±2 (6±2%)	1±1 (1±1%)	21	69
All	All	2480/3700 (67%)	2312 (93%)	146 (6%)	22 (1%)	21	69

All 11 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	1083	ASN	5
1	A	1064	GLU	3
1	A	1084	SER	3
1	A	1115	PRO	3
1	A	1125	ASN	2
1	A	1194	VAL	1
1	A	1087	PRO	1
1	A	1104	THR	1
1	A	1056	VAL	1
1	A	1065	SER	1
1	A	1085	GLU	1

### 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	108/162 (67%)	99±2 (92±2%)	9±2 (8±2%)	14	61
All	All	2160/3240 (67%)	1979 (92%)	181 (8%)	14	61

All 45 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	1068	ILE	18
1	A	1057	ASN	13
1	A	1190	ARG	12
1	A	1188	MET	12
1	A	1215	LEU	11
1	A	1225	LEU	11
1	A	1097	LYS	9
1	A	1222	LEU	9
1	A	1223	CYS	8
1	A	1105	LEU	5
1	A	1098	GLN	5
1	A	1085	GLU	5
1	A	1062	LEU	4
1	A	1067	ARG	4
1	A	1194	VAL	4
1	A	1185	ARG	3
1	A	1082	SER	3
1	A	1079	THR	3
1	A	1093	THR	3
1	A	1211	ARG	3
1	A	1209	LYS	3
1	A	1109	ILE	2
1	A	1184	ASN	2
1	A	1088	LEU	2
1	A	1179	LYS	2
1	A	1070	HIS	2
1	A	1122	LYS	2
1	A	1090	THR	2
1	A	1191	ASN	2
1	A	1081	GLU	2
1	A	1196	ASN	1
1	A	1208	ARG	1
1	A	1214	VAL	1
1	A	1125	ASN	1
1	A	1094	LYS	1
1	A	1100	LEU	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Models (Total)
1	A	1063	THR	1
1	A	1195	ILE	1
1	A	1219	ASP	1
1	A	1086	HIS	1
1	A	1177	GLN	1
1	A	1072	LYS	1
1	A	1186	GLU	1
1	A	1199	VAL	1
1	A	1206	HIS	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](#)

There are no ligands in this entry.

### 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 82% for the well-defined parts and 81% 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	2021
Number of shifts mapped to atoms	2021
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
$^{13}\text{C}_\alpha$	173	$-0.22 \pm 0.10$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}_\beta$	161	$0.10 \pm 0.06$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}'$	165	$0.04 \pm 0.09$	None needed ( $< 0.5$ ppm)
$^{15}\text{N}$	169	$-0.10 \pm 0.19$	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 82%, i.e. 1213 atoms were assigned a chemical shift out of a possible 1479. 18 out of 24 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	589/616 (96%)	238/246 (97%)	234/248 (94%)	117/122 (96%)
Sidechain	609/789 (77%)	375/453 (83%)	226/301 (75%)	8/35 (23%)

*Continued on next page...*



Continued from previous page...

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Aromatic	15/74 (20%)	15/40 (38%)	0/29 (0%)	0/5 (0%)
Overall	1213/1479 (82%)	628/739 (85%)	460/578 (80%)	125/162 (77%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 81%, i.e. 1761 atoms were assigned a chemical shift out of a possible 2178. 19 out of 29 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Backbone	850/919 (92%)	343/367 (93%)	338/370 (91%)	169/182 (93%)
Sidechain	888/1150 (77%)	549/662 (83%)	320/438 (73%)	19/50 (38%)
Aromatic	23/109 (21%)	23/59 (39%)	0/42 (0%)	0/8 (0%)
Overall	1761/2178 (81%)	915/1088 (84%)	658/850 (77%)	188/240 (78%)

#### 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:

