



Full wwPDB NMR Structure Validation Report ⓘ

May 29, 2020 – 12:07 am BST

PDB ID : 2KMX
Title : Solution structure of the nucleotide binding domain of the human Menkes protein in the ATP-bound 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

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
Mogul : 1.8.5 (274361), CSD as541be (2020)
buster-report : 1.1.7 (2018)
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

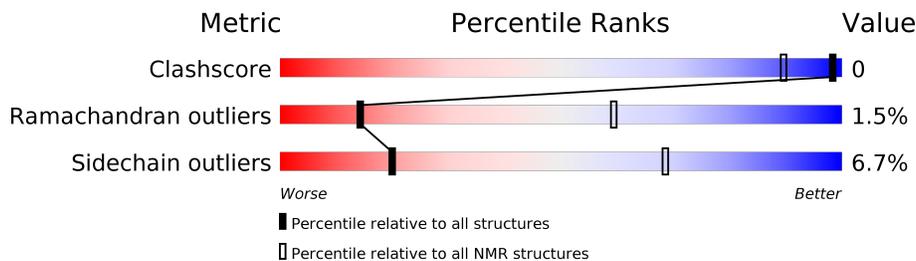
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 83%.

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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	185	

2 Ensemble composition and analysis i

This entry contains 20 models. Model 15 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *fewest violations*.

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:1053-A:1125, A:1177-A:1229 (126)	0.49	15

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 3 single-model clusters were found.

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

3 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 2863 atoms, of which 1417 are hydrogens and 0 are deuteriums.

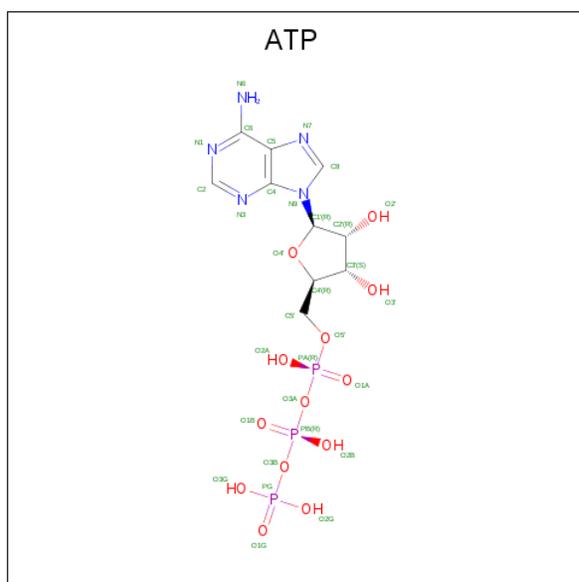
- Molecule 1 is a protein called Copper-transporting ATPase 1.

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

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

- Molecule 2 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$).



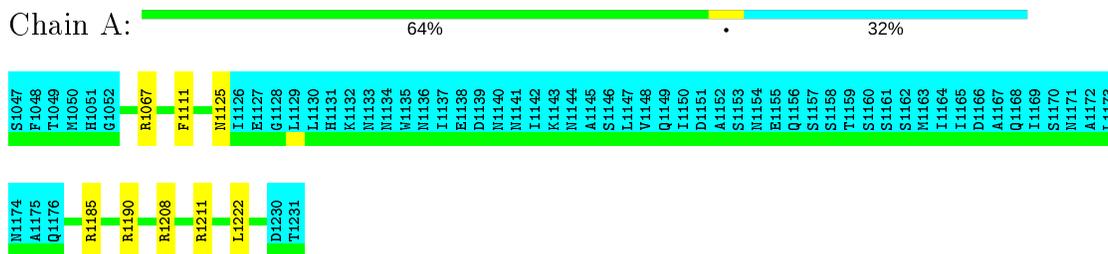
Mol	Chain	Residues	Atoms					
			Total	C	H	N	O	P
2	A	1	43	10	12	5	13	3

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

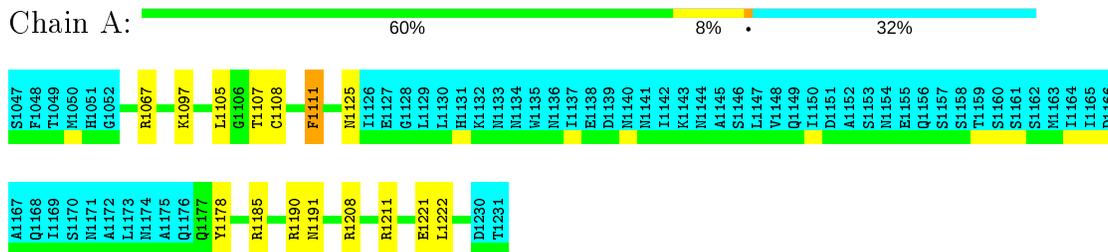


4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

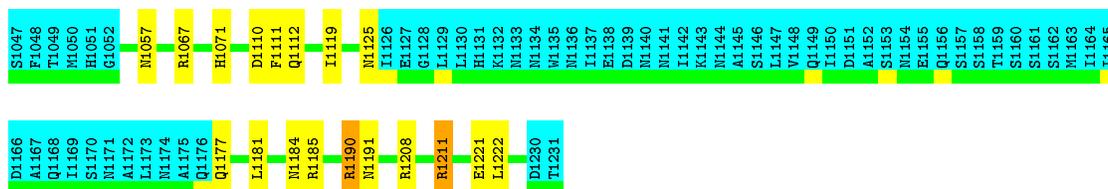
- Molecule 1: Copper-transporting ATPase 1



4.2.2 Score per residue for model 2

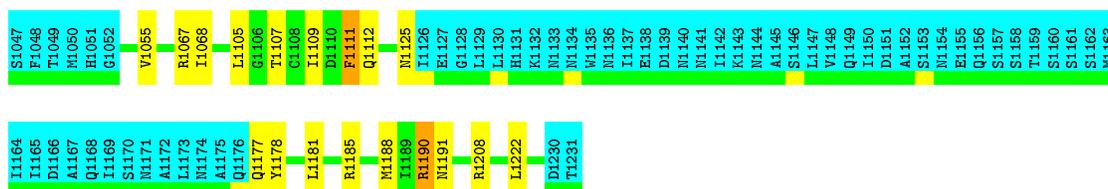
- Molecule 1: Copper-transporting ATPase 1





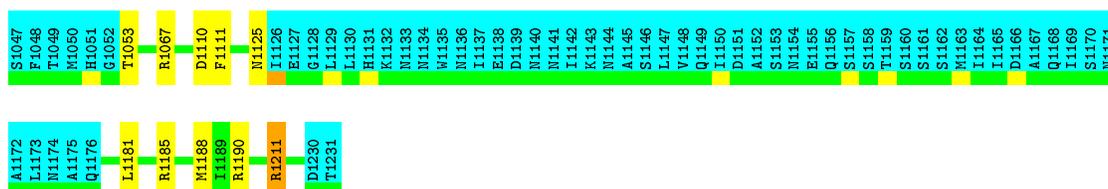
4.2.3 Score per residue for model 3

- Molecule 1: Copper-transporting ATPase 1



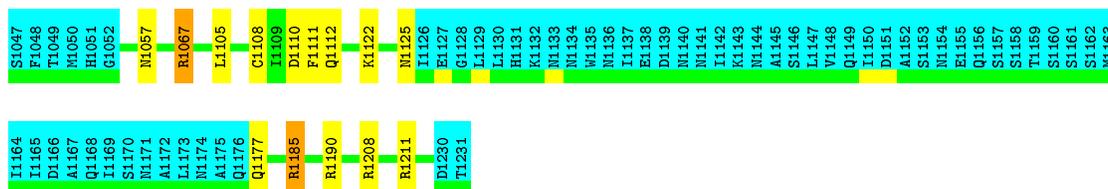
4.2.4 Score per residue for model 4

- Molecule 1: Copper-transporting ATPase 1



4.2.5 Score per residue for model 5

- Molecule 1: Copper-transporting ATPase 1





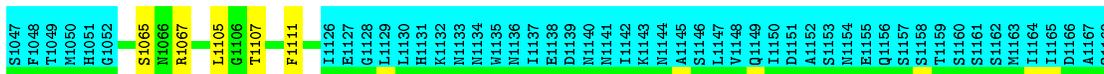
4.2.10 Score per residue for model 10

- Molecule 1: Copper-transporting ATPase 1



4.2.11 Score per residue for model 11

- Molecule 1: Copper-transporting ATPase 1



4.2.12 Score per residue for model 12

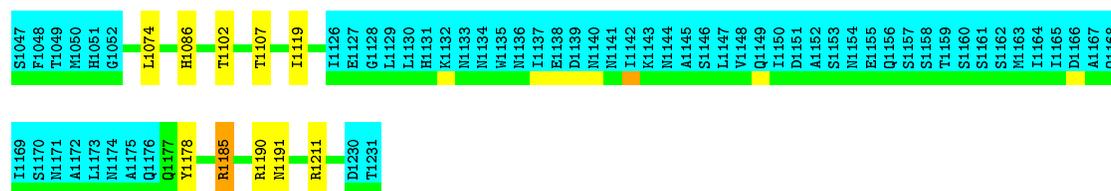
- Molecule 1: Copper-transporting ATPase 1



4.2.13 Score per residue for model 13

- Molecule 1: Copper-transporting ATPase 1

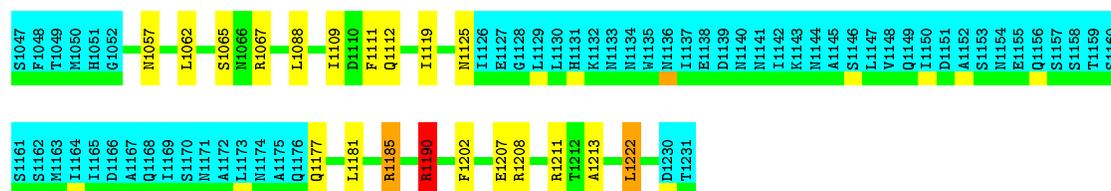
Chain A:  63% 5% 32%



4.2.14 Score per residue for model 14

- Molecule 1: Copper-transporting ATPase 1

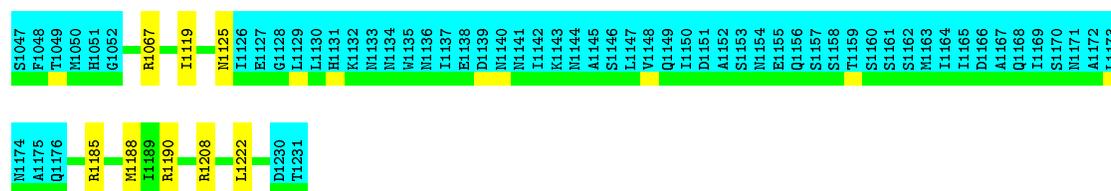
Chain A:  57% 9% 32%



4.2.15 Score per residue for model 15 (medoid)

- Molecule 1: Copper-transporting ATPase 1

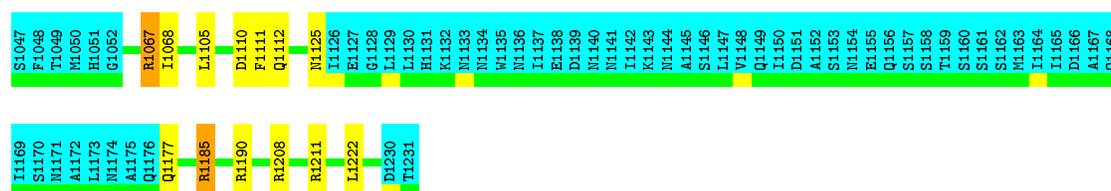
Chain A:  64% 32%



4.2.16 Score per residue for model 16

- Molecule 1: Copper-transporting ATPase 1

Chain A:  61% 6% 32%



L1173	R1185	D1230
M1174	R1190	T1231
A1175	R1208	
Q1176	L1222	

5 Refinement protocol and experimental data overview (i)

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	2054
Number of shifts mapped to atoms	2046
Number of unparsed shifts	0
Number of shifts with mapping errors	8
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	83%

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

COVALENT-GEOMETRY INFOmissingINFO

5.1 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	966	979	979	0±1
All	All	19940	19820	19820	6

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:1107:THR:HG22	1:A:1108:CYS:H	0.56	1.60	7	2
1:A:1062:LEU:HD23	1:A:1222:LEU:HD22	0.51	1.82	6	3
1:A:1207:GLU:HG2	1:A:1213:ALA:HB3	0.45	1.89	14	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	126/185 (68%)	115±2 (91±1%)	9±2 (7±2%)	2±1 (2±1%)	14	59
All	All	2520/3700 (68%)	2294 (91%)	187 (7%)	39 (2%)	14	59

All 10 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	1125	ASN	15
1	A	1111	PHE	6
1	A	1177	GLN	5
1	A	1190	ARG	3
1	A	1103	GLU	2
1	A	1107	THR	2
1	A	1067	ARG	2
1	A	1065	SER	2
1	A	1068	ILE	1
1	A	1185	ARG	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	110/162 (68%)	103±3 (93±3%)	7±3 (7±3%)	20	68
All	All	2200/3240 (68%)	2053 (93%)	147 (7%)	20	68

All 40 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	1222	LEU	16
1	A	1111	PHE	9
1	A	1112	GLN	8
1	A	1181	LEU	8
1	A	1110	ASP	7
1	A	1057	ASN	7
1	A	1191	ASN	7
1	A	1119	ILE	7
1	A	1105	LEU	6
1	A	1190	ARG	5
1	A	1188	MET	5
1	A	1185	ARG	5
1	A	1109	ILE	4
1	A	1184	ASN	4
1	A	1178	TYR	4
1	A	1177	GLN	4
1	A	1086	HIS	4
1	A	1097	LYS	3
1	A	1067	ARG	3
1	A	1068	ILE	3
1	A	1202	PHE	2
1	A	1088	LEU	2
1	A	1221	GLU	2
1	A	1211	ARG	2
1	A	1107	THR	2
1	A	1124	THR	2
1	A	1053	THR	2
1	A	1074	LEU	2
1	A	1085	GLU	1
1	A	1055	VAL	1
1	A	1070	HIS	1
1	A	1215	LEU	1
1	A	1219	ASP	1
1	A	1122	LYS	1
1	A	1108	CYS	1
1	A	1083	ASN	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	1209	LYS	1
1	A	1071	HIS	1
1	A	1198	ASP	1
1	A	1102	THR	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.

LIGAND-GEOMETRY INFOmissingINFO

5.5 Other polymers [i](#)

There are no such molecules in this entry.

5.6 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Chemical shift validation [i](#)

The completeness of assignment taking into account all chemical shift lists is 83% for the well-defined parts and 82% 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 [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	2054
Number of shifts mapped to atoms	2046
Number of unparsed shifts	0
Number of shifts with mapping errors	8
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.

- Chain not found in structure. All 8 occurrences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
UNMAPPED	1	ATP	H5'1	3.475	0.03	2
UNMAPPED	1	ATP	H1'	5.93	0.03	1
UNMAPPED	1	ATP	H3'	3.091	0.03	1
UNMAPPED	1	ATP	H5'2	3.586	0.03	2
UNMAPPED	1	ATP	H4'	4.082	0.03	1
UNMAPPED	1	ATP	H8	7.175	0.03	1
UNMAPPED	1	ATP	H2	8.189	0.03	1
UNMAPPED	1	ATP	H2'	3.543	0.03	1

6.1.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	175	-0.69 \pm 0.11	Should be applied
$^{13}\text{C}_\beta$	164	-0.16 \pm 0.11	None needed (< 0.5 ppm)
$^{13}\text{C}'$	157	0.03 \pm 0.09	None needed (< 0.5 ppm)
^{15}N	172	-0.26 \pm 0.25	None needed (< 0.5 ppm)

6.1.3 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 83%, i.e. 1250 atoms were assigned a chemical shift out of a possible 1500. 17 out of 24 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	597/624 (96%)	245/249 (98%)	231/252 (92%)	121/123 (98%)
Sidechain	633/802 (79%)	384/461 (83%)	239/306 (78%)	10/35 (29%)
Aromatic	20/74 (27%)	20/40 (50%)	0/29 (0%)	0/5 (0%)
Overall	1250/1500 (83%)	649/750 (87%)	470/587 (80%)	131/163 (80%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 82%, i.e. 1782 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	^1H	^{13}C	^{15}N
Backbone	851/919 (93%)	347/367 (95%)	332/370 (90%)	172/182 (95%)
Sidechain	903/1150 (79%)	552/662 (83%)	330/438 (75%)	21/50 (42%)
Aromatic	28/109 (26%)	28/59 (47%)	0/42 (0%)	0/8 (0%)
Overall	1782/2178 (82%)	927/1088 (85%)	662/850 (78%)	193/240 (80%)

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:

