



wwPDB NMR Structure Validation Summary Report ⓘ

Apr 27, 2016 – 12:11 AM BST

PDB ID : 2KZ7
Title : Solution structure of the CARMIL CAH3a/b domain bound to capping protein (CP)
Authors : Zwolak, A.; Tjandra, N.
Deposited on : 2010-06-12

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
<http://wwpdb.org/validation/2016/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 : unknown
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI : v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV : Wang et al. (2010)
ShiftChecker : rb-20027457
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027457

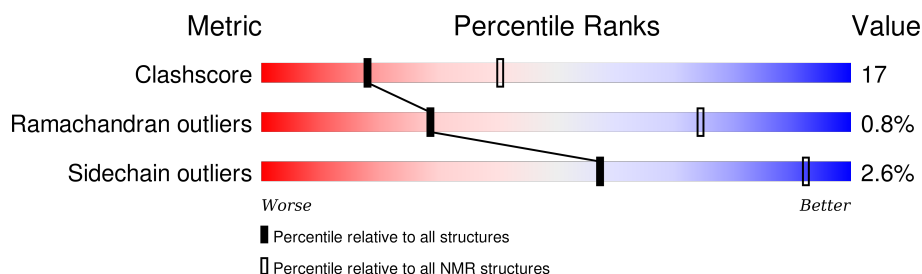
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 28%.

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	114402	11133
Ramachandran outliers	111179	9975
Sidechain outliers	111093	9958

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	286	 60% 33% • •
2	B	277	 76% 14% • 6% •
3	C	85	 96% •

2 Ensemble composition and analysis ⓘ

This entry contains 10 models. Model 7 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:8-A:281, B:302-B:554 (527)	0.05	7

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 4 clusters. No single-model clusters were found.

Cluster number	Models
1	1, 4, 7, 10
2	3, 6
3	5, 9
4	2, 8

3 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 9973 atoms, of which 4953 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called F-actin-capping protein subunit alpha-1.

Mol	Chain	Residues	Atoms						Trace
1	A	275	Total	C	H	N	O	S	0
			4413	1413	2175	392	428	5	

- Molecule 2 is a protein called F-actin-capping protein subunit beta isoforms 1 and 2.

Mol	Chain	Residues	Atoms						Trace
2	B	270	Total	C	H	N	O	S	0
			4273	1334	2133	374	422	10	

- Molecule 3 is a protein called Leucine-rich repeat-containing protein 16A.

Mol	Chain	Residues	Atoms						Trace
3	C	82	Total	C	H	N	O	S	0
			1287	393	645	121	124	4	

There are 10 discrepancies between the modelled and reference sequences:

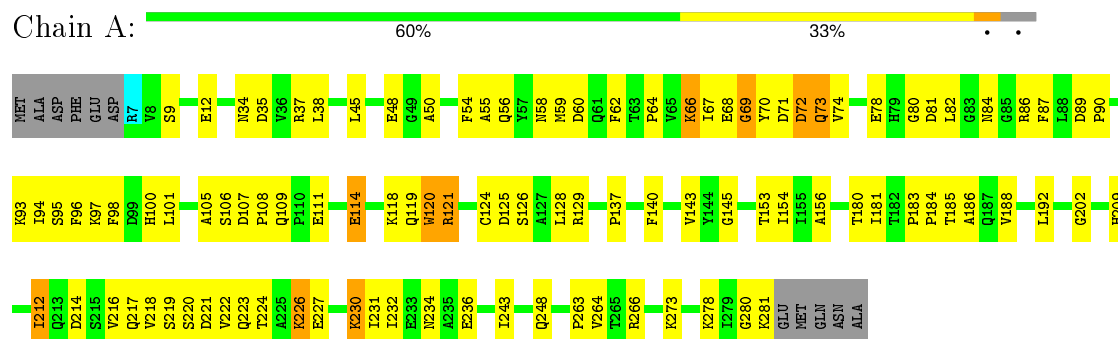
Chain	Residue	Modelled	Actual	Comment	Reference
C	598	GLY	-	EXPRESSION TAG	UNP Q6EDY6
C	599	ALA	-	EXPRESSION TAG	UNP Q6EDY6
C	600	MET	-	EXPRESSION TAG	UNP Q6EDY6
C	601	GLY	-	EXPRESSION TAG	UNP Q6EDY6
C	602	SER	-	EXPRESSION TAG	UNP Q6EDY6
C	603	TRP	-	EXPRESSION TAG	UNP Q6EDY6
C	604	GLY	-	EXPRESSION TAG	UNP Q6EDY6
C	605	CYS	-	EXPRESSION TAG	UNP Q6EDY6
C	681	GLY	-	EXPRESSION TAG	UNP Q6EDY6
C	682	CYS	-	EXPRESSION TAG	UNP Q6EDY6

4 Residue-property plots [i](#)

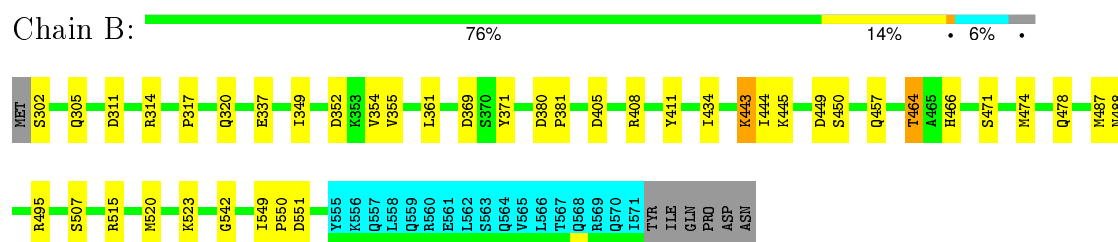
4.1 Average score per residue in the NMR ensemble

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.

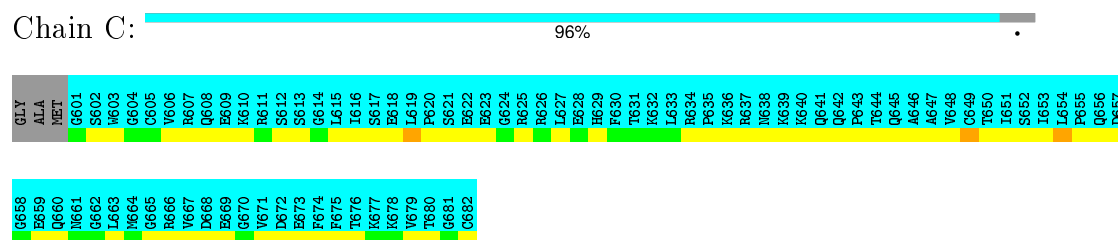
- Molecule 1: F-actin-capping protein subunit alpha-1



- Molecule 2: F-actin-capping protein subunit beta isoforms 1 and 2



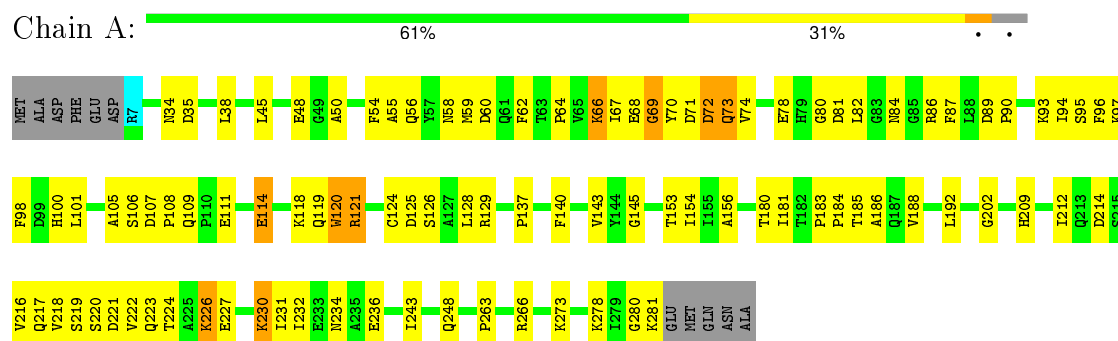
- Molecule 3: Leucine-rich repeat-containing protein 16A



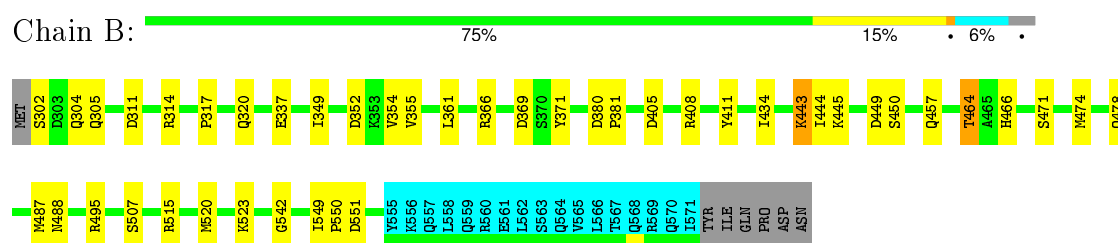
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 7. Colouring as in section 4.1 above.

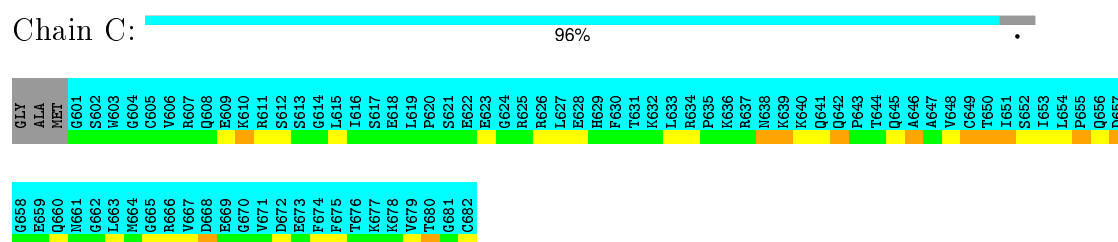
- Molecule 1: F-actin-capping protein subunit alpha-1



- Molecule 2: F-actin-capping protein subunit beta isoforms 1 and 2



- Molecule 3: Leucine-rich repeat-containing protein 16A



5 Refinement protocol and experimental data overview

The models were refined using the following method: *molecular dynamics, simulated annealing, torsion angle dynamics*.

Of the 200 calculated structures, 10 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
X-PLOR NIH	structure solution	2.23
X-PLOR NIH	refinement	2.23

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)	BMRB entry 17016
Number of chemical shift lists	1
Total number of shifts	2452
Number of shifts mapped to atoms	2272
Number of unparsed shifts	0
Number of shifts with mapping errors	180
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	28%

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

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	0.36±0.00	0±0/2279 (0.0±0.0%)	0.68±0.00	1±0/3088 (0.0±0.0%)
2	B	0.37±0.00	0±0/2026 (0.0±0.0%)	0.64±0.00	1±0/2738 (0.0±0.0%)
All	All	0.36	0/43050 (0.0%)	0.66	20/58260 (0.0%)

There are no bond-length outliers.

All 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
2	B	464	THR	N-CA-C	-7.10	91.82	111.00	8	10
1	A	72	ASP	N-CA-C	5.82	126.72	111.00	1	10

There are no chirality outliers.

There are no planarity outliers.

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	2227	2162	2154	120±3
2	B	1991	1974	1969	31±1
3	C	0	0	0	0±0
All	All	42180	41360	41230	1450

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 17.

5 of 179 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:67:ILE:HG22	1:A:72:ASP:CG	0.89	1.88	9	10
1:A:202:GLY:O	2:B:314:ARG:HD2	0.82	1.75	5	10
2:B:443:LYS:HD3	2:B:443:LYS:N	0.78	1.94	4	7
1:A:82:LEU:HD12	1:A:86:ARG:HB2	0.78	1.56	4	10
2:B:443:LYS:N	2:B:443:LYS:HD3	0.77	1.94	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	273/286 (95%)	248±1 (91±0%)	22±1 (8±0%)	3±0 (1±0%)	21	68
2	B	252/277 (91%)	244±0 (97±0%)	7±0 (3±0%)	1±0 (0±0%)	43	81
3	C	0	-	-	-	-	-
All	All	5250/6480 (81%)	4918 (94%)	289 (6%)	43 (1%)	29	74

5 of 6 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
2	B	550	PRO	10
1	A	73	GLN	10
1	A	69	GLY	10
1	A	68	GLU	10
1	A	8	VAL	2

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	242/252 (96%)	232±0 (96±0%)	10±0 (4±0%)	41	84
2	B	224/248 (90%)	222±0 (99±0%)	2±0 (1±0%)	85	97
3	C	0	-	-	-	-
All	All	4660/5720 (81%)	4539 (97%)	121 (3%)	57	92

5 of 13 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	120	TRP	10
1	A	84	ASN	10
1	A	121	ARG	10
2	B	464	THR	10
1	A	278	LYS	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 carbohydrates 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

The completeness of assignment taking into account all chemical shift lists is 28% for the well-defined parts and 28% for the entire structure.

7.1 Chemical shift list 1

File name: BMRB entry 17016

Chemical shift list name: *assigned_chem_shift_list_1*

7.1.1 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	2452
Number of shifts mapped to atoms	2272
Number of unparsed shifts	0
Number of shifts with mapping errors	180
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	16

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

- Residue not found in structure. First 5 (of 180) occurrences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
B	257	ALA	H	7.91	-1.0	1
B	250	ALA	CB	18.38	-1.0	1
A	220	ASN	CA	50.56	-1.0	1
A	85	SER	CB	63.93	-1.0	1
A	113	VAL	CA	60.21	-1.0	1

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	523	0.15 ± 0.08	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	460	0.69 ± 0.10	Should be applied
$^{13}\text{C}'$	521	-0.03 ± 0.06	None needed (< 0.5 ppm)

Continued on next page...

Continued from previous page...

Nucleus	# values	Correction \pm precision, ppm	Suggested action
^{15}N	474	-0.42 \pm 0.12	None needed (< 0.5 ppm)

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 28%, i.e. 1835 atoms were assigned a chemical shift out of a possible 6580. 0 out of 76 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	1491/2593 (58%)	353/1033 (34%)	785/1054 (74%)	353/506 (70%)
Sidechain	344/3469 (10%)	0/2032 (0%)	344/1264 (27%)	0/173 (0%)
Aromatic	0/518 (0%)	0/268 (0%)	0/217 (0%)	0/33 (0%)
Overall	1835/6580 (28%)	353/3333 (11%)	1129/2535 (45%)	353/712 (50%)

7.1.4 Statistically unusual chemical shifts [i](#)

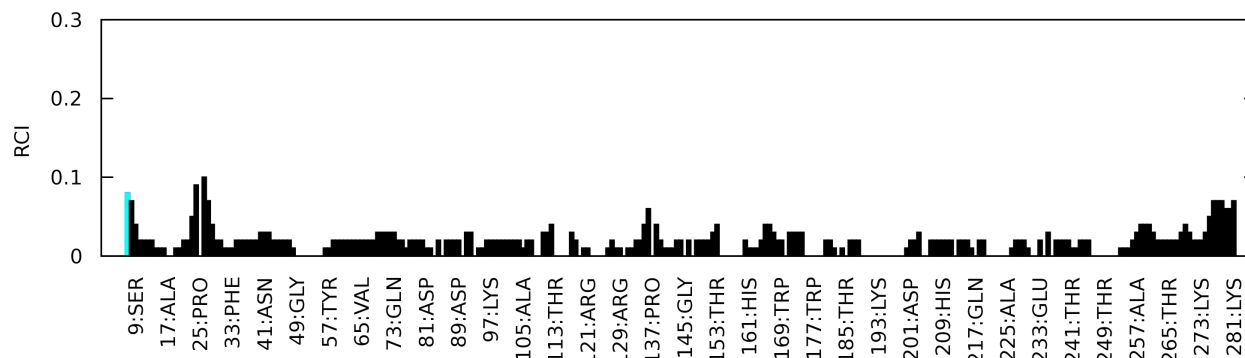
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
2	B	257	ALA	CB	73.88	28.03 – 9.93	30.3
2	B	387	ALA	CB	73.78	28.03 – 9.93	30.3
2	B	265	ALA	CB	73.78	28.03 – 9.93	30.3
1	A	130	ALA	CB	73.70	28.03 – 9.93	30.2
1	A	53	ALA	CB	73.65	28.03 – 9.93	30.2
1	A	18	ALA	CB	73.62	28.03 – 9.93	30.2
1	A	17	ALA	CB	73.54	28.03 – 9.93	30.1
1	A	51	ALA	CB	73.47	28.03 – 9.93	30.1
2	B	391	ALA	CB	73.40	28.03 – 9.93	30.1
1	A	50	ALA	CB	73.10	28.03 – 9.93	29.9
1	A	55	ALA	CB	72.78	28.03 – 9.93	29.7
1	A	162	GLN	N	19.34	138.01 – 101.71	-27.7
1	A	52	HIS	CB	73.65	40.69 – 19.69	20.7
1	A	83	GLY	CA	67.77	51.81 – 38.91	17.4
1	A	139	GLY	N	131.06	129.07 – 90.27	5.5
1	A	80	GLY	N	129.34	129.07 – 90.27	5.1

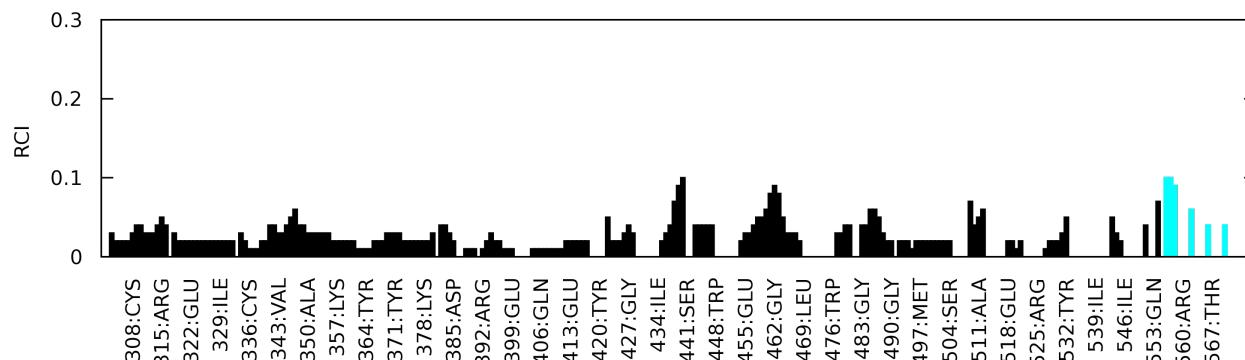
7.1.5 Random Coil Index (RCI) plots [i](#)

The images below report *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:



Random coil index (RCI) for chain B:



Random coil index (RCI) for chain C:

