wwPDB NMR Structure Validation Summary Report (i)

Apr 26, 2016 – 06:42 PM BST

PDB ID	:	2AAV
Title	:	Solution NMR structure of Filamin A domain 17
Authors	:	Nakamura, F.; Pudas, R.; Heikkinen, O.; Permi, P.; Kilpelainen, I.; Munday, A.D.; Hartwig, J.H.; Stossel, T.P.; Ylanne, J.
Deposited on	:	2005-07-14

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 (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

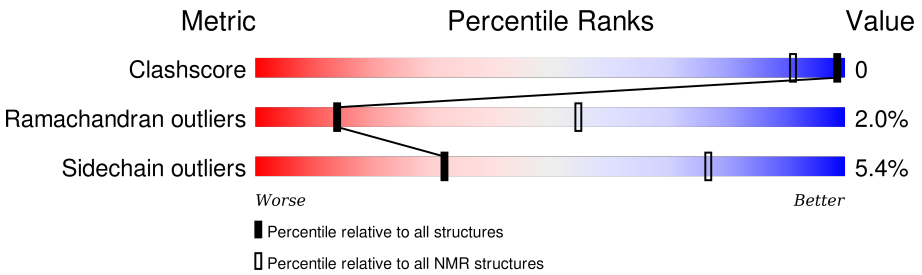
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal 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 (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment is 89%.

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	98	84% • 12%

2 Ensemble composition and analysis (i)

This entry contains 20 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: *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:1868-A:1894, A:1898-A:1956 (86)	0.61	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 3 clusters and 1 single-model cluster was found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1385 atoms, of which 679 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Filamin A.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	98	1385	438	679	118	146	4	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	GLY	-	CLONING ARTIFACT	UNP P21333
A	-3	ALA	-	CLONING ARTIFACT	UNP P21333
A	-2	MET	-	CLONING ARTIFACT	UNP P21333
A	-1	VAL	-	CLONING ARTIFACT	UNP P21333

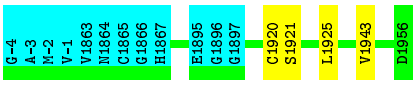
4 Residue-property plots (i)

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: Filamin A

Chain A:

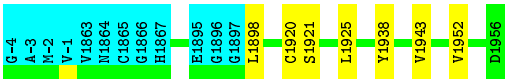


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

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

- Molecule 1: Filamin A

Chain A:



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *The structure calculations were made with the program CYANA 2.0 using the automated NOE assignment and structure calculation algorithm. The structure refinement of the 20 final CYANA structures was made with AMBER 8.0 program using the generalized Born continuum solvent model..*

Of the 300 calculated structures, 20 were deposited, based on the following criterion: *structures with favorable non-bond energy.*

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
CYANA	structure solution	2.0
AMBER	refinement	8.0

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 6730
Number of chemical shift lists	1
Total number of shifts	1065
Number of shifts mapped to atoms	1065
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	89%

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	0.75±0.02	0±0/644 (0.0±0.0%)	1.08±0.03	1±1/881 (0.1±0.1%)
All	All	0.75	0/12880 (0.0%)	1.08	17/17620 (0.1%)

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	0.1±0.3
All	All	0	2

There are no bond-length outliers.

5 of 6 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(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	1888	VAL	CA-CB-CG1	10.18	126.17	110.90	3	1
1	A	1943	VAL	CA-CB-CG1	6.84	121.17	110.90	12	4
1	A	1951	ARG	NE-CZ-NH1	6.54	123.57	120.30	11	7
1	A	1880	VAL	CA-CB-CG1	5.69	119.44	110.90	19	3
1	A	1938	TYR	CB-CG-CD1	-5.40	117.76	121.00	10	1

There are no chirality outliers.

All 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	1938	TYR	Sidechain	1
1	A	1932	TYR	Sidechain	1

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	630	609	609	0±1
All	All	12600	12180	12180	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.

5 of 6 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:1888:VAL:HG12	1:A:1920:CYS:HB3	0.74	1.57	3	1
1:A:1888:VAL:CG2	1:A:1920:CYS:H	0.58	2.11	19	1
1:A:1888:VAL:HG12	1:A:1920:CYS:CB	0.56	2.29	3	1
1:A:1900:LEU:HD11	1:A:1922:VAL:HG21	0.53	1.79	3	1
1:A:1888:VAL:HG21	1:A:1920:CYS:CB	0.42	2.44	19	1

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	85/98 (87%)	75±2 (89±3%)	8±2 (9±3%)	2±1 (2±1%)	14	55
All	All	1700/1960 (87%)	1506 (89%)	160 (9%)	34 (2%)	14	55

5 of 13 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	1921	SER	13
1	A	1907	LYS	3
1	A	1952	VAL	3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	1889	ASN	2
1	A	1891	LYS	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	71/78 (91%)	67±1 (95±2%)	4±1 (5±2%)	32 77
All	All	1420/1560 (91%)	1343 (95%)	77 (5%)	32 77

5 of 17 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	1943	VAL	20
1	A	1925	LEU	13
1	A	1920	CYS	11
1	A	1898	LEU	7
1	A	1882	LYS	5

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 (i)

The completeness of assignment taking into account all chemical shift lists is 89% for the well-defined parts and 87% for the entire structure.

7.1 Chemical shift list 1

File name: BMRB entry 6730

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	1065
Number of shifts mapped to atoms	1065
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	108

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	96	0.52 ± 0.09	Should be applied
$^{13}C_{\beta}$	84	0.80 ± 0.26	Should be applied
$^{13}C'$	79	0.00 ± 0.00	None needed (< 0.5 ppm)
^{15}N	88	-0.61 ± 0.33	None needed (imprecise)

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 89%, i.e. 835 atoms were assigned a chemical shift out of a possible 940. 14 out of 15 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	400/416 (96%)	165/165 (100%)	157/172 (91%)	78/79 (99%)
Sidechain	401/460 (87%)	240/267 (90%)	155/180 (86%)	6/13 (46%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	34/64 (53%)	18/34 (53%)	16/28 (57%)	0/2 (0%)
Overall	835/940 (89%)	423/466 (91%)	328/380 (86%)	84/94 (89%)

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
1	A	1932	TYR	CE1	40.44	124.14 – 111.74	-62.5
1	A	1871	TYR	CE1	41.43	124.14 – 111.74	-61.7
1	A	1886	PHE	CD1	56.58	137.63 – 125.43	-61.4
1	A	1948	PHE	CD1	57.05	137.63 – 125.43	-61.0
1	A	1938	TYR	CE1	42.30	124.14 – 111.74	-61.0
1	A	1924	TYR	CE1	42.30	124.14 – 111.74	-61.0
1	A	1938	TYR	CD1	56.54	139.11 – 126.41	-60.0
1	A	1932	TYR	CD1	57.91	139.11 – 126.41	-58.9
1	A	1871	TYR	CD1	58.34	139.11 – 126.41	-58.6
1	A	1924	TYR	CD1	58.34	139.11 – 126.41	-58.6
1	A	1886	PHE	CE1	53.57	137.92 – 123.42	-53.2
1	A	1948	PHE	CE1	54.32	137.92 – 123.42	-52.7
1	A	1948	PHE	CZ	54.13	137.04 – 121.44	-48.1
1	A	1886	PHE	CZ	54.14	137.04 – 121.44	-48.1
1	A	1942	HIS	CE1	60.42	149.70 – 125.30	-31.6
1	A	1877	HIS	CE1	61.17	149.70 – 125.30	-31.3
1	A	1867	HIS	CE1	61.48	149.70 – 125.30	-31.2
1	A	1929	PRO	C	130.31	184.42 – 169.02	-30.1
1	A	1883	PRO	C	132.09	184.42 – 169.02	-29.0
1	A	1905	PRO	C	132.24	184.42 – 169.02	-28.9
1	A	1947	PRO	C	132.45	184.42 – 169.02	-28.7
1	A	1944	PRO	C	133.28	184.42 – 169.02	-28.2
1	A	1926	PRO	C	133.43	184.42 – 169.02	-28.1
1	A	1873	PRO	C	134.63	184.42 – 169.02	-27.3
1	A	1899	SER	C	128.34	183.48 – 165.88	-26.3
1	A	1911	SER	C	128.59	183.48 – 165.88	-26.2
1	A	1869	THR	C	128.37	183.43 – 165.73	-26.1
1	A	1923	SER	C	128.78	183.48 – 165.88	-26.1
1	A	1949	THR	C	128.49	183.43 – 165.73	-26.0
1	A	1921	SER	C	129.05	183.48 – 165.88	-25.9
1	A	1913	THR	C	128.69	183.43 – 165.73	-25.9
1	A	1906	SER	C	129.20	183.48 – 165.88	-25.8

Continued on next page...

Continued from previous page...

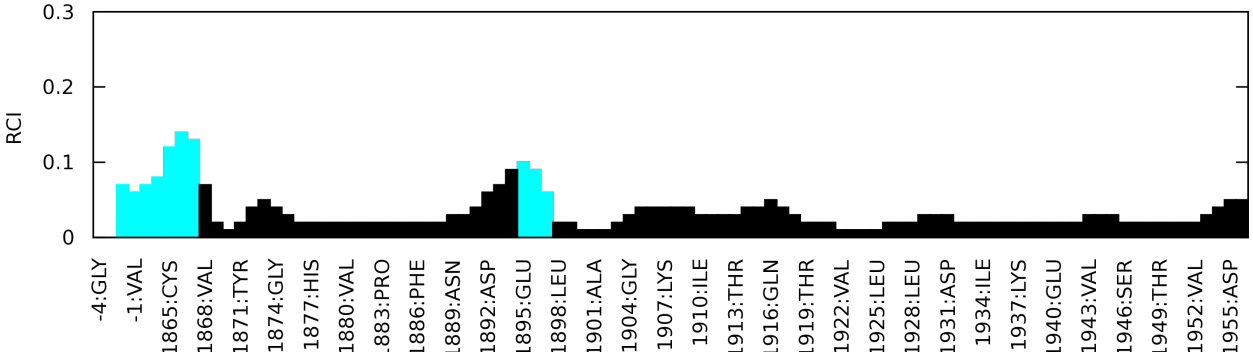
Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	1887	THR	C	128.94	183.43 – 165.73	-25.8
1	A	1955	ASP	C	130.92	185.29 – 167.59	-25.7
1	A	1933	SER	C	129.65	183.48 – 165.88	-25.6
1	A	1919	THR	C	129.79	183.43 – 165.73	-25.3
1	A	1892	ASP	C	131.69	185.29 – 167.59	-25.3
1	A	1878	GLY	C	126.48	183.33 – 164.53	-25.2
1	A	1931	ASP	C	131.84	185.29 – 167.59	-25.2
1	A	1881	ASN	C	129.58	184.41 – 166.21	-25.1
1	A	1953	THR	C	130.17	183.43 – 165.73	-25.1
1	A	1885	THR	C	130.38	183.43 – 165.73	-25.0
1	A	1930	GLY	C	127.65	183.33 – 164.53	-24.6
1	A	1890	THR	C	131.32	183.43 – 165.73	-24.4
1	A	1939	ASN	C	130.87	184.41 – 166.21	-24.4
1	A	1945	GLY	C	128.10	183.33 – 164.53	-24.4
1	A	1864	ASN	C	131.11	184.41 – 166.21	-24.3
1	A	1889	ASN	C	131.29	184.41 – 166.21	-24.2
1	A	1896	GLY	C	128.73	183.33 – 164.53	-24.0
1	A	1886	PHE	C	127.65	185.43 – 165.53	-24.0
1	A	1897	GLY	C	128.89	183.33 – 164.53	-24.0
1	A	1874	GLY	C	128.97	183.33 – 164.53	-23.9
1	A	1918	GLY	C	129.18	183.33 – 164.53	-23.8
1	A	1888	VAL	C	130.51	185.16 – 166.16	-23.8
1	A	1922	VAL	C	130.58	185.16 – 166.16	-23.7
1	A	1876	THR	C	132.59	183.43 – 165.73	-23.7
1	A	1954	GLY	C	129.37	183.33 – 164.53	-23.7
1	A	1866	GLY	C	129.43	183.33 – 164.53	-23.7
1	A	1879	VAL	C	130.75	185.16 – 166.16	-23.6
1	A	1880	VAL	C	131.08	185.16 – 166.16	-23.5
1	A	1877	HIS	C	128.46	185.27 – 165.27	-23.4
1	A	1927	VAL	C	131.22	185.16 – 166.16	-23.4
1	A	1894	GLY	C	130.00	183.33 – 164.53	-23.4
1	A	1940	GLU	C	130.81	186.83 – 167.03	-23.3
1	A	1937	LYS	C	130.56	186.58 – 166.78	-23.3
1	A	1934	ILE	C	130.51	185.63 – 166.13	-23.3
1	A	1868	VAL	C	131.53	185.16 – 166.16	-23.2
1	A	1924	TYR	C	128.99	185.42 – 165.42	-23.2
1	A	1903	GLU	C	131.22	186.83 – 167.03	-23.1
1	A	1863	VAL	C	131.87	185.16 – 166.16	-23.0
1	A	1942	HIS	CD2	42.45	137.40 – 103.40	-22.9
1	A	1902	ILE	C	131.24	185.63 – 166.13	-22.9
1	A	1871	TYR	C	129.68	185.42 – 165.42	-22.9
1	A	1910	ILE	C	131.52	185.63 – 166.13	-22.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	1907	LYS	C	131.91	186.58 – 166.78	-22.6
1	A	1867	HIS	C	130.18	185.27 – 165.27	-22.5
1	A	1935	LEU	C	131.72	187.06 – 166.96	-22.5
1	A	1932	TYR	C	130.60	185.42 – 165.42	-22.4
1	A	1895	GLU	C	132.59	186.83 – 167.03	-22.4
1	A	1941	GLN	C	132.27	186.20 – 166.50	-22.4
1	A	1867	HIS	CD2	44.34	137.40 – 103.40	-22.4
1	A	1909	GLU	C	132.77	186.83 – 167.03	-22.3
1	A	1952	VAL	C	133.30	185.16 – 166.16	-22.3
1	A	1875	LEU	C	132.28	187.06 – 166.96	-22.3
1	A	1877	HIS	CD2	44.87	137.40 – 103.40	-22.2
1	A	1898	LEU	C	132.60	187.06 – 166.96	-22.1
1	A	1901	ALA	C	130.41	188.57 – 166.97	-21.9
1	A	1912	CYS	C	129.77	185.15 – 164.55	-21.9
1	A	1891	LYS	C	133.38	186.58 – 166.78	-21.9
1	A	1951	ARG	C	131.68	186.68 – 166.18	-21.8
1	A	1948	PHE	C	132.24	185.43 – 165.53	-21.7
1	A	1920	CYS	C	130.17	185.15 – 164.55	-21.7
1	A	1870	ALA	C	131.43	188.57 – 166.97	-21.5
1	A	1865	CYS	C	130.84	185.15 – 164.55	-21.4
1	A	1950	ALA	C	131.97	188.57 – 166.97	-21.2
1	A	1938	TYR	C	133.10	185.42 – 165.42	-21.2
1	A	1942	HIS	C	133.60	185.27 – 165.27	-20.8
1	A	1884	ALA	C	133.01	188.57 – 166.97	-20.7
1	A	1893	ALA	C	134.11	188.57 – 166.97	-20.2
1	A	1951	ARG	NE	116.67	92.63 – 76.73	20.1
1	A	1951	ARG	CZ	128.85	176.59 – 143.59	-9.5
1	A	1907	LYS	HZ1	3.10	10.62 – 4.22	-6.7
1	A	1907	LYS	HZ2	3.10	10.62 – 4.22	-6.7
1	A	1907	LYS	HZ3	3.10	10.62 – 4.22	-6.7
1	A	1875	LEU	HD21	-1.09	2.14 – -0.66	-6.5
1	A	1875	LEU	HD23	-1.09	2.14 – -0.66	-6.5
1	A	1875	LEU	HD22	-1.09	2.14 – -0.66	-6.5
1	A	1922	VAL	HB	0.11	3.59 – 0.39	-5.9

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.

Random coil index (RCI) for chain A: