wwPDB NMR Structure Validation Summary Report (i)

Apr 26, 2016 – 01:42 PM BST

PDB ID	:	1ADN
Title	:	SOLUTION STRUCTURE OF THE DNA METHYLPHOSPHOTRIESTER REPAIR DOMAIN OF ESCHERICHIA COLI ADA
Authors	:	Myers, L.C.; Verdine, G.L.; Wagner, G.
Deposited on	:	1993-09-30

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:

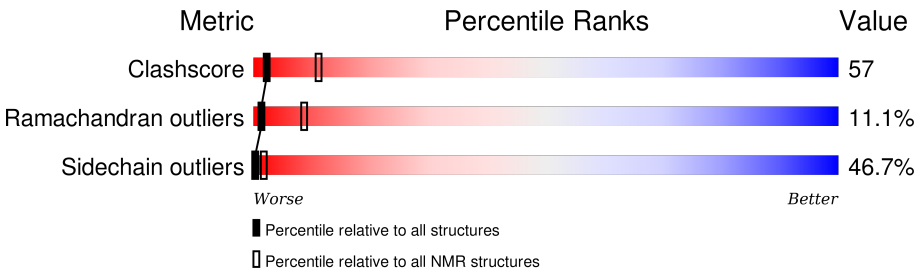
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	rb-20027457
Ideal 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 87%.

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	92	11% 39% 11% 39%

2 Ensemble composition and analysis (i)

This entry contains 14 models. Model 11 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:8-A:17, A:27-A:72 (56)	0.43	11

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 2 single-model clusters were found.

Cluster number	Models
1	3, 5, 6, 7, 8, 9, 10, 11, 12, 13
2	1, 2
Single-model clusters	4; 14

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 732 atoms, of which 0 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called N-ADA 10.

Mol	Chain	Residues	Atoms					Trace
			Total	C	N	O	S	
1	A	92	731	447	148	129	7	0

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	
			Total	Zn
2	A	1	1	1

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: N-ADA 10

Chain A: 11% 39% 11% 39%

M1 K2 K3 A4 T5 C6 L7 T8 D9 D10 Q11 R12 W13 Q14 S15 V16 L17 A18 R19 D20 P21 N22 A23 D24 G25 E26 F27 V28 F29 A30 V31 R32 T33 T34 G35 I36 F37 C38 R39 P40 S41 C42 R43 A44 R45 H46 A47 L48 R49 E50 N51 V52 S53 F54 Y55 A56 N57 A58 S59 E60

A61 L62 A63 A64 G65 F66 R67 P68 C69 K70 R71 G72 Q73 P74 D75 K76 A77 N78 P79 R80 Q81 H82 R83 L84 D85 K86 I87 T88 H89 A90 C91 R92

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

The representative model is number 11. Colouring as in section 4.1 above.

- Molecule 1: N-ADA 10

Chain A: 14% 34% 13% 39%

M1 K2 K3 A4 T5 C6 L7 T8 D9 D10 Q11 R12 W13 Q14 S15 V16 L17 A18 R19 D20 P21 N22 A23 D24 G25 E26 F27 V28 F29 A30 V31 R32 T33 T34 G35 I36 F37 C38 R39 P40 S41 C42 R43 A44 R45 H46 A47 L48 R49 E50 N51 V52 S53 F54 Y55 A56 N57 A58 S59 E60

A61 L62 A63 A64 G65 F66 R67 P68 C69 K70 R71 G72 Q73 P74 D75 K76 A77 N78 P79 R80 Q81 H82 R83 L84 D85 K86 I87 T88 H89 A90 C91 R92

5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *?*.

Of the ? calculated structures, 14 were deposited, based on the following criterion: *?*.

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

Software name	Classification	Version
DGII	refinement	

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 6053, BMRB entry 6054
Number of chemical shift lists	2
Total number of shifts	3396
Number of shifts mapped to atoms	2026
Number of unparsed shifts	0
Number of shifts with mapping errors	1370
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	87%

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)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	2.58±0.01	60±0/458 (13.1±0.0%)	1.71±0.03	22±1/619 (3.6±0.2%)
All	All	2.58	840/6412 (13.1%)	1.71	310/8666 (3.6%)

5 of 64 unique bond 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(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	60	GLU	CD-OE1	10.23	1.36	1.25	6	7
1	A	60	GLU	CD-OE2	10.04	1.36	1.25	13	7
1	A	50	GLU	CD-OE2	10.04	1.36	1.25	13	8
1	A	50	GLU	CD-OE1	10.01	1.36	1.25	11	6
1	A	64	ALA	C-N	8.88	1.49	1.33	6	14

5 of 27 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
1	A	32	ARG	NE-CZ-NH1	8.38	124.49	120.30	11	14
1	A	12	ARG	NE-CZ-NH1	8.35	124.47	120.30	4	14
1	A	45	ARG	NE-CZ-NH1	8.34	124.47	120.30	14	14
1	A	43	ARG	NE-CZ-NH1	8.32	124.46	120.30	14	14
1	A	39	ARG	NE-CZ-NH1	8.29	124.45	120.30	7	14

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	447	0	433	51±8
All	All	6272	0	6062	709

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 57.

5 of 274 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:31:VAL:HG21	1:A:47:ALA:HB1	1.13	1.16	10	7
1:A:31:VAL:HG21	1:A:47:ALA:HB3	1.09	1.16	5	5
1:A:48:LEU:HD22	1:A:50:GLU:HG3	1.06	1.07	13	2
1:A:28:VAL:HG23	1:A:55:TYR:HD1	1.05	1.05	1	4
1:A:37:PHE:HB3	1:A:66:PHE:CD2	1.02	1.90	7	14

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	56/92 (61%)	42±2 (76±4%)	7±2 (13±4%)	6±2 (11±3%)	1 9
All	All	784/1288 (61%)	593 (76%)	104 (13%)	87 (11%)	1 9

5 of 14 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	72	CYS	14
1	A	40	PRO	14
1	A	45	ARG	11

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	41	SER	11
1	A	46	HIS	7

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	46/76 (61%)	25±3 (53±7%)	22±3 (47±7%)	0 2
All	All	644/1064 (61%)	343 (53%)	301 (47%)	0 2

5 of 43 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	52	VAL	14
1	A	48	LEU	14
1	A	66	PHE	14
1	A	40	PRO	13
1	A	62	LEU	12

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)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

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 87% for the well-defined parts and 79% for the entire structure.

7.1 Chemical shift list 1

File name: BMRB entry 6053

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	1679
Number of shifts mapped to atoms	1011
Number of unparsed shifts	0
Number of shifts with mapping errors	668
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	12

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

- Residue not found in structure. First 5 (of 668) occurences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	127	THR	CB	68.31	0.05	1
A	121	LYS	N	122.19	0.05	1
A	110	ALA	N	123.33	0.05	1
A	123	THR	C	175.52	0.05	1
A	127	THR	HG22	1.48	0.01	1

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	144	-0.36 ± 0.24	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	137	-0.03 ± 0.13	None needed (< 0.5 ppm)
$^{13}C'$	130	-0.35 ± 0.22	None needed (< 0.5 ppm)

Continued on next page...

Continued from previous page...

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
^{15}N	130	0.27 ± 0.20	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 87%, i.e. 619 atoms were assigned a chemical shift out of a possible 714. 7 out of 7 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	274/276 (99%)	110/110 (100%)	110/112 (98%)	54/54 (100%)
Sidechain	286/367 (78%)	178/218 (82%)	101/120 (84%)	7/29 (24%)
Aromatic	59/71 (83%)	34/39 (87%)	24/31 (77%)	1/1 (100%)
Overall	619/714 (87%)	322/367 (88%)	235/263 (89%)	62/84 (74%)

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.

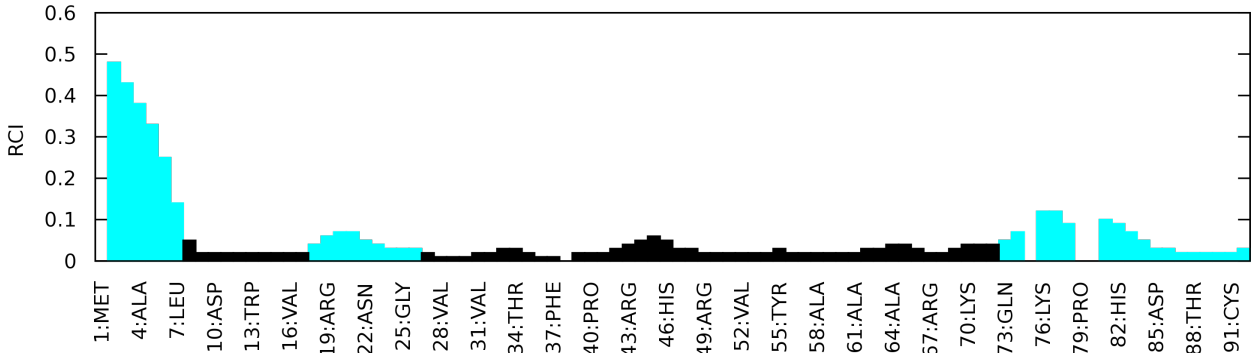
Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	49	ARG	NE	111.97	92.63 – 76.73	17.2
1	A	39	ARG	NE	111.93	92.63 – 76.73	17.1
1	A	12	ARG	NE	109.99	92.63 – 76.73	15.9
1	A	12	ARG	HB3	-0.18	3.17 – 0.37	-7.0
1	A	87	ILE	CG2	9.19	24.63 – 10.43	-5.9
1	A	49	ARG	HD3	1.65	4.36 – 1.86	-5.8
1	A	61	ALA	HB1	-0.05	2.61 – 0.11	-5.6
1	A	61	ALA	HB2	-0.05	2.61 – 0.11	-5.6
1	A	61	ALA	HB3	-0.05	2.61 – 0.11	-5.6
1	A	12	ARG	HD3	1.73	4.36 – 1.86	-5.5
1	A	49	ARG	HD2	1.96	4.27 – 1.97	-5.0
1	A	12	ARG	HG3	0.10	3.00 – 0.10	-5.0

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:



7.2 Chemical shift list 2

File name: BMRB entry 6054

Chemical shift list name: *assigned_chem_shift_list_1*

7.2.1 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	1717
Number of shifts mapped to atoms	1015
Number of unparsed shifts	0
Number of shifts with mapping errors	702
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	18

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

- Residue not found in structure. First 5 (of 702) occurences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	127	THR	CB	68.29	0.05	1
A	121	LYS	N	122.2	0.05	1
A	110	ALA	N	123.4	0.05	1
A	123	THR	C	175.52	0.05	1
A	127	THR	HG22	1.48	0.01	1

7.2.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	145	-0.37 ± 0.14	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	140	0.04 ± 0.16	None needed (< 0.5 ppm)
$^{13}C'$	133	-0.26 ± 0.13	None needed (< 0.5 ppm)
^{15}N	131	0.68 ± 0.43	None needed (imprecise)

7.2.3 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 86%, i.e. 613 atoms were assigned a chemical shift out of a possible 714. 7 out of 7 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	269/276 (97%)	108/110 (98%)	108/112 (96%)	53/54 (98%)
Sidechain	283/367 (77%)	176/218 (81%)	100/120 (83%)	7/29 (24%)
Aromatic	61/71 (86%)	36/39 (92%)	24/31 (77%)	1/1 (100%)
Overall	613/714 (86%)	320/367 (87%)	232/263 (88%)	61/84 (73%)

7.2.4 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.

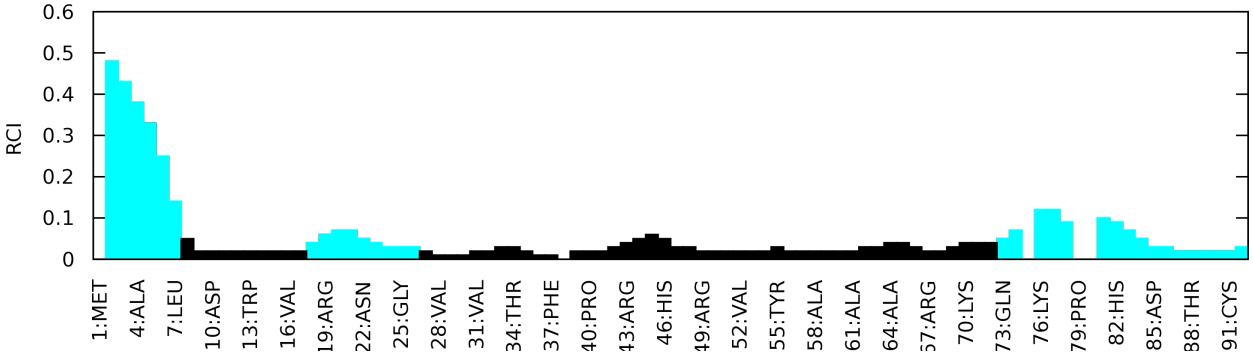
Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	49	ARG	NE	111.74	92.63 – 76.73	17.0
1	A	39	ARG	NE	111.43	92.63 – 76.73	16.8
1	A	12	ARG	NE	109.82	92.63 – 76.73	15.8
1	A	12	ARG	HB3	-0.11	3.17 – 0.37	-6.7
1	A	32	ARG	HG2	3.35	2.92 – 0.22	6.6
1	A	32	ARG	HG3	3.35	3.00 – 0.10	6.2
1	A	49	ARG	HD3	1.58	4.36 – 1.86	-6.1
1	A	87	ILE	CG2	9.21	24.63 – 10.43	-5.9
1	A	32	ARG	HD2	1.78	4.27 – 1.97	-5.8
1	A	65	GLY	N	131.24	129.07 – 90.27	5.6
1	A	12	ARG	HD3	1.74	4.36 – 1.86	-5.5
1	A	32	ARG	HD3	1.78	4.36 – 1.86	-5.3
1	A	61	ALA	HB1	0.03	2.61 – 0.11	-5.3
1	A	61	ALA	HB2	0.03	2.61 – 0.11	-5.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	61	ALA	HB3	0.03	2.61 – 0.11	-5.3
1	A	49	ARG	HD2	1.92	4.27 – 1.97	-5.2
1	A	68	PRO	HA	2.72	6.05 – 2.75	-5.1
1	A	12	ARG	HG3	0.10	3.00 – 0.10	-5.0

7.2.5 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: