Full wwPDB NMR Structure Validation Report (i)

Jun 15, 2024 – 07:14 AM EDT

PDB ID	:	2LWL
BMRB ID	:	18634
Title	:	Structural Basis for the Interaction of Human β-Defensin 6 and Its Putative Chemokine Receptor CCR2 and Breast Cancer Microvesicles
Authors	:	de Paula, V.S.; Gomes, N.S.F.; Lima, L.G.; Miyamoto, C.A.; Monteiro, R.Q.; Almeida, F.C.L.; Valente, A.
Deposited on	:	2012-08-02

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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

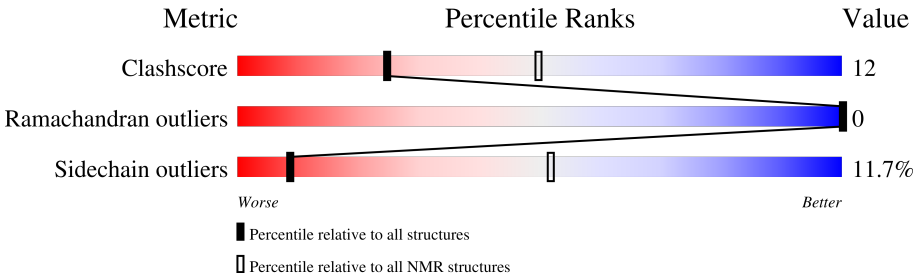
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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	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 <=5%

Mol	Chain	Length	Quality of chain
1	A	49	49% 12% • 27% 8%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 5 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:6-A:37 (32)	0.53	5

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 702 atoms, of which 356 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Beta-defensin 106.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	45	702	212	356	61	67	6	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	EXPRESSION TAG	UNP Q8N104
A	-2	SER	-	EXPRESSION TAG	UNP Q8N104
A	-1	HIS	-	EXPRESSION TAG	UNP Q8N104
A	0	MET	-	EXPRESSION TAG	UNP Q8N104
A	40	SER	CYS	ENGINEERED MUTATION	UNP Q8N104

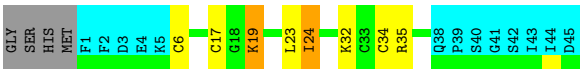
4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide 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: Beta-defensin 106

Chain A: 49% 12% • 27% 8%



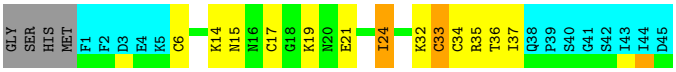
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Beta-defensin 106

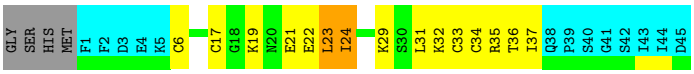
Chain A: 39% 22% • 27% 8%



4.2.2 Score per residue for model 2

• Molecule 1: Beta-defensin 106

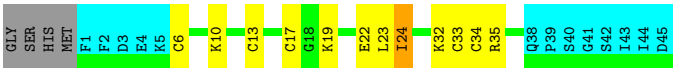
Chain A: 35% 27% • 27% 8%



4.2.3 Score per residue for model 3

- Molecule 1: Beta-defensin 106

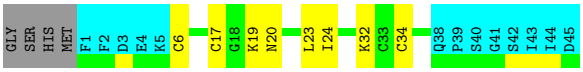
Chain A: 41% 22% • 27% 8%



4.2.4 Score per residue for model 4

- Molecule 1: Beta-defensin 106

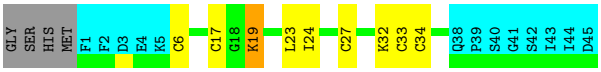
Chain A: 49% 16% 27% 8%



4.2.5 Score per residue for model 5 (medoid)

- Molecule 1: Beta-defensin 106

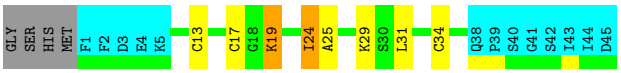
Chain A: 47% 16% • 27% 8%



4.2.6 Score per residue for model 6

- Molecule 1: Beta-defensin 106

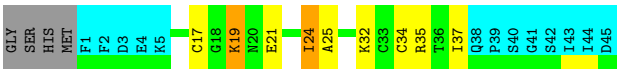
Chain A: 49% 12% • 27% 8%



4.2.7 Score per residue for model 7

- Molecule 1: Beta-defensin 106

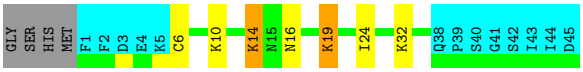
Chain A: 47% 14% • 27% 8%



4.2.8 Score per residue for model 8

- Molecule 1: Beta-defensin 106

Chain A: 51% 10% • 27% 8%



4.2.9 Score per residue for model 9

- Molecule 1: Beta-defensin 106

Chain A: 47% 14% • 27% 8%

4.2.10 Score per residue for model 10

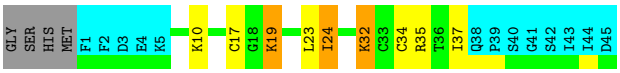
- Molecule 1: Beta-defensin 106

Chain A: 43% 18% • 27% 8%

4.2.11 Score per residue for model 11

- Molecule 1: Beta-defensin 106

Chain A: 47% 12% 6% 27% 8%



4.2.12 Score per residue for model 12

- Molecule 1: Beta-defensin 106

Chain A: 47% 16% • 27% 8%

4.2.13 Score per residue for model 13

- Molecule 1: Beta-defensin 106

Chain A: 41% 18% 6% 27% 8%

GLY SER HIS MET F1 F2 D3 E4 K5 C6 N15 N16 C17 G18 K19 N20 E21 E22 L23 I24 A25 K32 C33 C34 R35 T36 I37 Q38 P39 S40 G41 S42 I43 I44 D45

4.2.14 Score per residue for model 14

- Molecule 1: Beta-defensin 106

Chain A: 45% 16% • 27% 8%

GLY SER HIS MET F1 F2 D3 E4 K5 C6 N7 K8 C17 G18 K19 L23 I24 A25 K32 C33 C34 I37 Q38 P39 S40 G41 S42 I43 I44 D45

4.2.15 Score per residue for model 15

- Molecule 1: Beta-defensin 106

Chain A: 45% 18% • 27% 8%

GLY SER HIS MET F1 F2 D3 E4 K5 N16 C17 G18 K19 L23 I24 A25 K29 S30 L31 C34 I37 Q38 P39 S40 G41 S42 I43 I44 D45

4.2.16 Score per residue for model 16

- Molecule 1: Beta-defensin 106

Chain A: 33% 27% 6% 27% 8%

GLY SER HIS MET F1 F2 D3 E4 K5 C6 T12 C13 K14 N15 N16 C17 G18 K19 N20 E21 E22 L23 I24 A25 L26 C27 K32 C33 C34 R35 Q38 P39 S40 G41 S42 I43 I44 D45

4.2.17 Score per residue for model 17

- Molecule 1: Beta-defensin 106

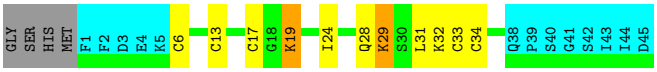
Chain A: 39% 24% • 27% 8%

GLY SER HIS MET F1 F2 D3 E4 K5 C6 N7 K8 C17 G18 K19 N20 E21 E22 L23 I24 K32 C33 C34 R35 T36 I37 Q38 P39 S40 G41 S42 I43 I44 D45

4.2.18 Score per residue for model 18

- Molecule 1: Beta-defensin 106

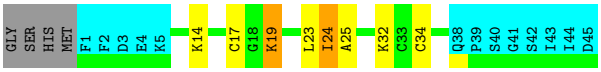
Chain A: 43% 18% • 27% 8%



4.2.19 Score per residue for model 19

- Molecule 1: Beta-defensin 106

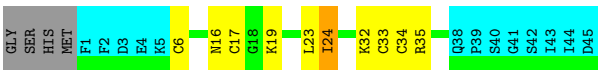
Chain A: 49% 12% • 27% 8%



4.2.20 Score per residue for model 20

- Molecule 1: Beta-defensin 106

Chain A: 45% 18% • 27% 8%



5 Refinement protocol and experimental data overview (i)

Of the 200 calculated structures, 20 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**
ARIA/CNS version	refinement	1.2

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)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	514
Number of shifts mapped to atoms	514
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%

6 Model quality (i)

6.1 Standard geometry (i)

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

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	241	261	261	6±2
All	All	4820	5220	5220	117

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 12.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:23:LEU:HD11	1:A:32:LYS:HB3	0.75	1.57	4	7
1:A:19:LYS:HD2	1:A:19:LYS:O	0.65	1.91	10	10
1:A:24:ILE:HG21	1:A:35:ARG:CD	0.62	2.24	13	8
1:A:14:LYS:C	1:A:14:LYS:HE3	0.59	2.17	8	1
1:A:23:LEU:HD11	1:A:32:LYS:CG	0.59	2.27	17	2
1:A:13:CYS:HA	1:A:32:LYS:O	0.58	1.98	3	2
1:A:14:LYS:HE3	1:A:14:LYS:O	0.56	2.00	8	1
1:A:17:CYS:SG	1:A:34:CYS:SG	0.56	3.03	10	19
1:A:23:LEU:HD21	1:A:32:LYS:CG	0.55	2.31	13	1
1:A:24:ILE:HG21	1:A:35:ARG:CG	0.54	2.32	9	3
1:A:10:LYS:O	1:A:35:ARG:HA	0.54	2.01	3	1
1:A:15:ASN:O	1:A:32:LYS:HE2	0.54	2.03	13	1
1:A:23:LEU:HD11	1:A:32:LYS:CB	0.53	2.32	10	7
1:A:24:ILE:HG21	1:A:35:ARG:HG3	0.53	1.80	9	2

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:14:LYS:HD2	1:A:16:ASN:O	0.52	2.05	8	1
1:A:23:LEU:HD11	1:A:32:LYS:HG3	0.50	1.83	11	2
1:A:22:GLU:O	1:A:34:CYS:HA	0.50	2.07	2	2
1:A:24:ILE:HG21	1:A:35:ARG:HD3	0.49	1.84	10	3
1:A:24:ILE:HG13	1:A:25:ALA:N	0.49	2.23	13	8
1:A:13:CYS:HB3	1:A:27:CYS:SG	0.49	2.48	16	1
1:A:6:CYS:CB	1:A:33:CYS:HG	0.48	2.21	10	3
1:A:6:CYS:SG	1:A:24:ILE:HD11	0.48	2.48	3	1
1:A:6:CYS:SG	1:A:33:CYS:HB3	0.48	2.48	20	3
1:A:21:GLU:HG2	1:A:36:THR:HA	0.48	1.86	2	5
1:A:28:GLN:O	1:A:29:LYS:HG2	0.47	2.10	18	1
1:A:29:LYS:HB2	1:A:31:LEU:HD23	0.46	1.87	6	3
1:A:21:GLU:HA	1:A:35:ARG:O	0.45	2.12	1	3
1:A:23:LEU:CD1	1:A:32:LYS:HB3	0.45	2.37	4	1
1:A:13:CYS:O	1:A:31:LEU:HD12	0.44	2.12	6	1
1:A:23:LEU:HD21	1:A:32:LYS:HG2	0.44	1.86	13	1
1:A:20:ASN:O	1:A:37:ILE:HG12	0.44	2.13	17	1
1:A:27:CYS:SG	1:A:33:CYS:SG	0.44	3.16	5	1
1:A:6:CYS:HG	1:A:33:CYS:CB	0.43	2.26	1	2
1:A:14:LYS:HD2	1:A:16:ASN:C	0.43	2.34	8	1
1:A:6:CYS:HB2	1:A:33:CYS:HG	0.42	1.73	2	1
1:A:23:LEU:HG	1:A:32:LYS:HD2	0.42	1.90	2	1
1:A:23:LEU:HD11	1:A:32:LYS:HB2	0.42	1.90	12	1
1:A:19:LYS:HG3	1:A:19:LYS:O	0.42	2.14	17	1
1:A:24:ILE:HG21	1:A:35:ARG:HG2	0.41	1.93	13	1
1:A:6:CYS:HG	1:A:33:CYS:HB3	0.40	1.77	1	1
1:A:17:CYS:SG	1:A:23:LEU:HD13	0.40	2.57	2	1
1:A:19:LYS:HD2	1:A:19:LYS:N	0.40	2.32	16	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	32/49 (65%)	31±1 (97±2%)	1±1 (3±2%)	0±0 (0±0%)	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	640/980 (65%)	619 (97%)	21 (3%)	0 (0%)	100 100

There are no Ramachandran outliers.

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	29/44 (66%)	26±1 (88±5%)	3±1 (12±5%)	9 52
All	All	580/880 (66%)	512 (88%)	68 (12%)	9 52

All 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	19	LYS	18
1	A	24	ILE	18
1	A	32	LYS	8
1	A	14	LYS	5
1	A	23	LEU	3
1	A	15	ASN	2
1	A	10	LYS	2
1	A	16	ASN	2
1	A	13	CYS	2
1	A	33	CYS	1
1	A	20	ASN	1
1	A	28	GLN	1
1	A	36	THR	1
1	A	8	LYS	1
1	A	12	THR	1
1	A	29	LYS	1
1	A	31	LEU	1

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 monosaccharides 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: working_cs.cif

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	514
Number of shifts mapped to atoms	514
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

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	45	0.10 ± 0.35	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	42	0.21 ± 0.26	None needed (< 0.5 ppm)
$^{13}C'$	43	0.23 ± 0.16	None needed (< 0.5 ppm)
^{15}N	44	0.68 ± 0.32	Should be applied

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. 380 atoms were assigned a chemical shift out of a possible 425. 0 out of 4 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	162/162 (100%)	66/66 (100%)	64/64 (100%)	32/32 (100%)
Sidechain	218/263 (83%)	144/168 (86%)	71/81 (88%)	3/14 (21%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Overall	380/425 (89%)	210/234 (90%)	135/145 (93%)	35/46 (76%)

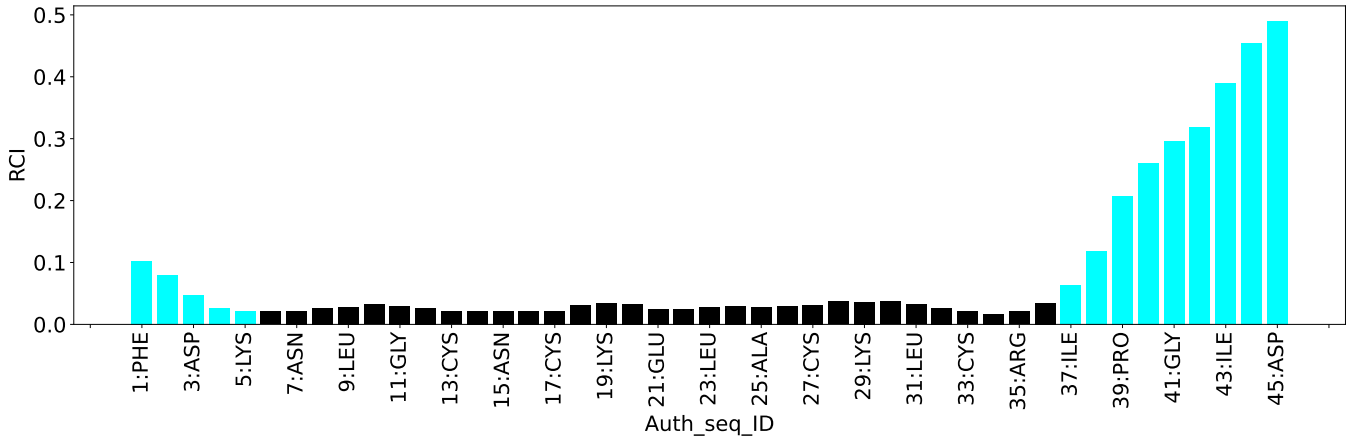
The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 87%, i.e. 514 atoms were assigned a chemical shift out of a possible 594. 0 out of 4 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	224/226 (99%)	92/92 (100%)	88/90 (98%)	44/44 (100%)
Sidechain	289/348 (83%)	191/222 (86%)	95/110 (86%)	3/16 (19%)
Aromatic	1/20 (5%)	1/10 (10%)	0/10 (0%)	0/0 (—%)
Overall	514/594 (87%)	284/324 (88%)	183/210 (87%)	47/60 (78%)

7.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

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.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A: