wwPDB NMR Structure Validation Summary Report

Jun 6, 2023 – 04:19 AM EDT

PDB ID	:	2MMG
BMRB ID	:	19854
Title	:	Structural Characterization of the Mengovirus Leader Protein Bound to Ran GTPase by Nuclear Magnetic Resonance
Authors	:	Bacot-Davis, V.R.; Palmenberg, A.C.; Cornilescu, C.C.; Markley, J.L.
Deposited on	:	2014-03-15

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
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 (1)) were used in the production of this report:

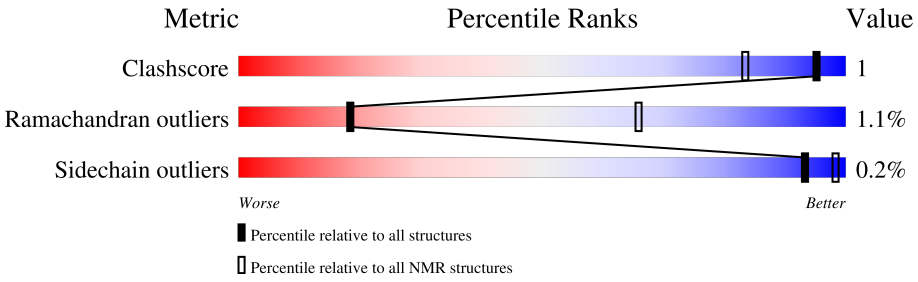
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.33

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 85%.

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	216	82% · 14%

2 Ensemble composition and analysis

This entry contains 10 models. Model 6 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:9-A:175 (167)	0.22	6
2	A:182-A:190 (9)	0.97	9
3	A:199-A:207 (9)	0.15	2

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 2 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 3449 atoms, of which 1725 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called GTP-binding nuclear protein Ran.

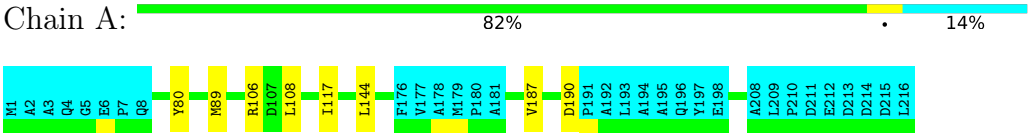
Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	216	3449	1109	1725	295	313	7	0

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.

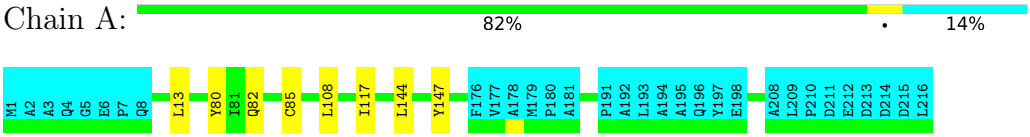
• Molecule 1: GTP-binding nuclear protein Ran



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

The representative model is number 6. Colouring as in section 4.1 above.

• Molecule 1: GTP-binding nuclear protein Ran



5 Refinement protocol and experimental data overview

The models were refined using the following method: *torsion angle dynamics*.

Of the 50 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
CYANA	structure solution	
TALOS	geometry optimization	
SPARTA+	geometry optimization	
CYANA	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)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	2631
Number of shifts mapped to atoms	2625
Number of unparsed shifts	0
Number of shifts with mapping errors	6
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	85%

6 Model quality

6.1 Standard geometry

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	0.78±0.01	0±0/1532 (0.0± 0.0%)	1.09±0.02	1±2/2076 (0.0± 0.1%)
All	All	0.78	0/15320 (0.0%)	1.09	10/20760 (0.0%)

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	0.9±0.9
All	All	0	9

There are no bond-length outliers.

5 of 10 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	13	LEU	CB-CA-C	-6.43	97.99	110.20	5	1
1	A	60	LYS	N-CA-CB	5.82	121.07	110.60	8	1
1	A	190	ASP	N-CA-CB	5.57	120.62	110.60	1	1
1	A	158	GLU	CB-CA-C	-5.54	99.33	110.40	5	1
1	A	52	PHE	N-CA-CB	5.16	119.89	110.60	5	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	89	MET	Mainchain	5
1	A	189	MET	Mainchain	3
1	A	80	TYR	Sidechain	1

6.2 Too-close contacts

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	1495	1516	1516	3±1
All	All	14950	15160	15160	29

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 1.

5 of 14 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:13:LEU:HD22	1:A:85:CYS:SG	0.74	2.23	5	1
1:A:29:ARG:HH21	1:A:151:ALA:HA	0.55	1.61	1	2
1:A:117:ILE:HB	1:A:144:LEU:HD22	0.51	1.82	5	9
1:A:80:TYR:CE1	1:A:108:LEU:HG	0.48	2.43	3	7
1:A:104:TRP:O	1:A:108:LEU:HD13	0.47	2.09	5	1

6.3 Torsion angles

6.3.1 Protein backbone

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	185/216 (86%)	181±1 (98±1%)	2±1 (1±0%)	2±1 (1±1%)	18	66
All	All	1850/2160 (86%)	1809 (98%)	21 (1%)	20 (1%)	18	66

All 4 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	187	VAL	7
1	A	106	ARG	6
1	A	190	ASP	6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	189	MET	1

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	163/185 (88%)	163±0 (100±0%)	0±0 (0±0%)	93	98
All	All	1630/1850 (88%)	1627 (100%)	3 (0%)	93	98

All 2 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	77	ASP	2
1	A	82	GLN	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

The completeness of assignment taking into account all chemical shift lists is 85% for the well-defined parts and 85% 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

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	2631
Number of shifts mapped to atoms	2625
Number of unparsed shifts	0
Number of shifts with mapping errors	6
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	92

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

- No matching atom found in the structure. First 5 (of 6) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	30	HIS	HE2	11.376	0.020	1
1	A	48	HIS	HE2	11.298	0.020	1
1	A	53	HIS	HE2	11.083	0.020	1
1	A	105	HIS	HE2	11.112	0.020	1
1	A	139	HIS	HE2	12.494	0.020	1
1	A	199	HIS	HE2	12.445	0.020	1

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	212	-1.31 ± 0.41	Should be checked
$^{13}C_{\beta}$	196	-0.41 ± 0.22	None needed (< 0.5 ppm)

Continued on next page...

Continued from previous page...

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C'$	211	-0.48 ± 0.20	None needed (< 0.5 ppm)
^{15}N	215	0.64 ± 0.55	None needed (imprecise)

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 85%, i.e. 2236 atoms were assigned a chemical shift out of a possible 2630. 0 out of 36 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	740/922 (80%)	202/375 (54%)	362/370 (98%)	176/177 (99%)
Sidechain	1249/1461 (85%)	880/948 (93%)	308/452 (68%)	61/61 (100%)
Aromatic	247/247 (100%)	120/120 (100%)	112/112 (100%)	15/15 (100%)
Overall	2236/2630 (85%)	1202/1443 (83%)	782/934 (84%)	252/253 (100%)

7.1.4 Statistically unusual chemical shifts

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.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	114	ASN	ND2	1114.29	101.55 – 123.95	447.1
1	A	104	TRP	CE3	174.81	111.58 – 129.41	30.5
1	A	32	THR	HG1	6.41	0.08 – 2.19	25.0
1	A	66	THR	HG1	6.01	0.08 – 2.19	23.1
1	A	207	THR	HG1	5.99	0.08 – 2.19	23.0
1	A	206	THR	HG1	5.96	0.08 – 2.19	22.8
1	A	24	THR	HG1	5.91	0.08 – 2.19	22.6
1	A	54	THR	HG1	5.59	0.08 – 2.19	21.1
1	A	93	THR	HG1	5.56	0.08 – 2.19	21.0
1	A	97	THR	HG1	5.48	0.08 – 2.19	20.6
1	A	25	THR	HG1	5.39	0.08 – 2.19	20.1
1	A	127	LYS	HB3	6.92	0.46 – 3.04	20.0
1	A	42	THR	HG1	5.31	0.08 – 2.19	19.8
1	A	127	LYS	HD2	5.61	0.58 – 2.64	19.4
1	A	127	LYS	HD3	5.66	0.54 – 2.65	19.3
1	A	84	GLN	HB3	6.96	0.71 – 3.33	18.9
1	A	21	THR	HG1	5.08	0.08 – 2.19	18.7
1	A	127	LYS	HG2	5.43	0.13 – 2.61	16.4

Continued on next page...

Continued from previous page...

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	127	LYS	HG3	5.58	0.04 – 2.67	16.1
1	A	140	ARG	NE	108.39	76.53 – 92.65	14.8
1	A	127	LYS	HE3	5.80	1.92 – 3.89	14.7
1	A	127	LYS	HE2	5.73	1.95 – 3.88	14.6
1	A	127	LYS	HB2	5.20	0.58 – 2.97	14.3
1	A	104	TRP	CD2	155.17	118.40 – 137.65	14.1
1	A	134	LYS	CB	54.58	24.03 – 41.47	12.5
1	A	82	GLN	NE2	132.38	103.38 – 120.35	12.1
1	A	183	ALA	HB1	4.16	0.14 – 2.58	11.5
1	A	183	ALA	HB2	4.16	0.14 – 2.58	11.5
1	A	183	ALA	HB3	4.16	0.14 – 2.58	11.5
1	A	36	GLU	CB	48.14	21.56 – 38.37	10.8
1	A	85	CYS	HB3	7.56	0.69 – 5.10	10.6
1	A	139	HIS	HB3	6.88	1.18 – 4.91	10.3
1	A	85	CYS	HB2	7.31	0.81 – 5.11	10.1
1	A	28	LYS	NZ	59.20	19.79 – 46.09	10.0
1	A	84	GLN	NE2	128.39	103.38 – 120.35	9.7
1	A	205	GLN	NE2	128.39	103.38 – 120.35	9.7
1	A	139	HIS	HB2	6.46	1.36 – 4.85	9.6
1	A	175	GLU	CA	76.25	47.03 – 67.62	9.2
1	A	45	VAL	CB	17.10	23.86 – 41.50	-8.8
1	A	68	GLY	C	189.53	164.92 – 182.89	8.7
1	A	151	ALA	C	160.74	167.61 – 188.05	-8.4
1	A	60	LYS	CA	74.29	46.18 – 67.77	8.0
1	A	84	GLN	HB2	4.01	0.80 – 3.29	7.9
1	A	123	LYS	NZ	53.49	19.79 – 46.09	7.8
1	A	23	LYS	NZ	53.30	19.79 – 46.09	7.7
1	A	99	LYS	NZ	53.30	19.79 – 46.09	7.7
1	A	127	LYS	NZ	53.29	19.79 – 46.09	7.7
1	A	37	LYS	NZ	53.20	19.79 – 46.09	7.7
1	A	32	THR	CG2	29.98	16.06 – 27.03	7.7
1	A	82	GLN	CD	170.41	173.59 – 185.85	-7.6
1	A	186	GLU	CD	198.61	171.45 – 193.13	7.5
1	A	38	LYS	NZ	52.40	19.79 – 46.09	7.4
1	A	143	ASN	ND2	129.29	101.55 – 123.95	7.4
1	A	158	GLU	CA	72.15	47.03 – 67.62	7.2
1	A	44	GLY	CA	54.58	38.93 – 51.79	7.2
1	A	99	LYS	CE	47.85	37.57 – 46.21	6.9
1	A	29	ARG	CA	72.15	45.44 – 68.13	6.8
1	A	33	GLY	C	161.91	164.92 – 182.89	-6.7
1	A	23	LYS	CE	47.55	37.57 – 46.21	6.6
1	A	60	LYS	NZ	49.99	19.79 – 46.09	6.5

Continued on next page...

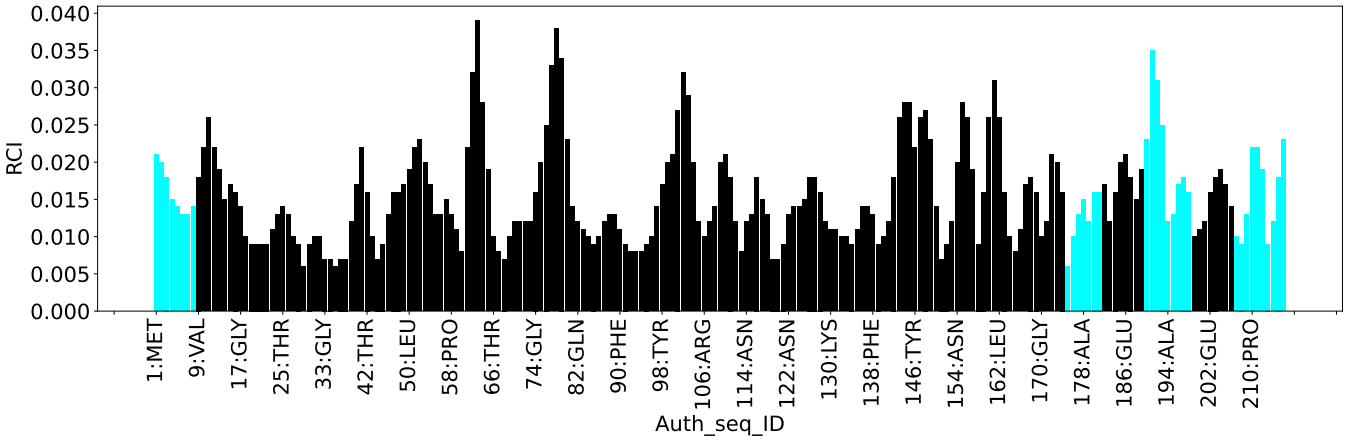
Continued from previous page...

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	62	ASN	ND2	98.29	101.55 – 123.95	-6.5
1	A	12	LYS	NZ	49.89	19.79 – 46.09	6.5
1	A	170	GLY	C	162.50	164.92 – 182.89	-6.3
1	A	132	LYS	NZ	49.59	19.79 – 46.09	6.3
1	A	94	SER	C	163.97	166.15 – 183.14	-6.3
1	A	134	LYS	NZ	49.30	19.79 – 46.09	6.2
1	A	141	LYS	NZ	49.24	19.79 – 46.09	6.2
1	A	89	MET	CE	27.93	8.39 – 25.85	6.2
1	A	142	LYS	NZ	49.22	19.79 – 46.09	6.2
1	A	152	LYS	NZ	49.22	19.79 – 46.09	6.2
1	A	71	LYS	NZ	49.00	19.79 – 46.09	6.1
1	A	119	LEU	CA	42.87	45.17 – 66.21	-6.1
1	A	22	GLY	C	184.81	164.92 – 182.89	6.1
1	A	99	LYS	HE2	4.08	1.95 – 3.88	6.1
1	A	124	VAL	CG1	29.68	14.71 – 28.29	6.0
1	A	84	GLN	HG3	3.94	0.91 – 3.68	5.9
1	A	165	ALA	C	165.72	167.61 – 188.05	-5.9
1	A	127	LYS	CE	46.97	37.57 – 46.21	5.9
1	A	110	ARG	NE	94.02	76.53 – 92.65	5.8
1	A	104	TRP	CD1	137.30	117.34 – 135.80	5.8
1	A	113	GLU	CA	69.22	47.03 – 67.62	5.8
1	A	124	VAL	CB	42.57	23.86 – 41.50	5.6
1	A	154	ASN	ND2	125.21	101.55 – 123.95	5.6
1	A	213	ASP	CB	49.61	32.98 – 48.76	5.5
1	A	99	LYS	CD	34.95	23.50 – 34.42	5.5
1	A	99	LYS	HE3	3.98	1.92 – 3.89	5.5
1	A	29	ARG	NE	93.29	76.53 – 92.65	5.4
1	A	76	ARG	NE	93.29	76.53 – 92.65	5.4
1	A	194	ALA	C	166.90	167.61 – 188.05	-5.3
1	A	167	LYS	NZ	46.77	19.79 – 46.09	5.3
1	A	84	GLN	HG2	3.66	1.01 – 3.62	5.2
1	A	140	ARG	CB	39.64	21.74 – 39.52	5.1

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:



8 NMR restraints analysis (i)

No restraints data found

9 Distance violation analysis (i)

No distance restraints data found

10 Dihedral-angle violation analysis (i)

Dihedral angle analysis failed due to data error in the dihedral angle restraints, possibly missing target value