wwPDB NMR Structure Validation Summary Report

Mar 25, 2025 – 04:04 PM EDT

PDB ID	:	9NM2
BMRB ID	:	31232
Title	:	Dimeric Structure of full-length CrgA, a Cell Division Protein from Mycobacterium tuberculosis, in Lipid Bilayers
Authors	:	Shin, Y.; Prasad, R.; Das, N.; Taylor, J.A.; Qin, H.; Hu, W.; Hu, Y.-Y.; Fu, R.; Zhang, R.; Zhou, H.-X.; Cross, T.A.
Deposited on	:	2025-03-03

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

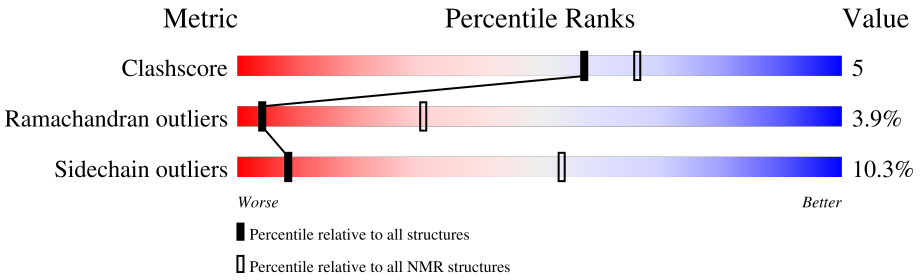
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41.4

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLID-STATE NMR

The overall completeness of chemical shifts assignment is 3%.

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	210492	14027
Ramachandran outliers	207382	12486
Sidechain outliers	206894	12463

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	102	48% 21% 7% 25%
1	B	102	50% 15% 9% • 25%

2 Ensemble composition and analysis

This entry contains 10 models. Model 3 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:17-A:93, B:17-B:93 (154)	0.25	3

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. No single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2478 atoms, of which 1266 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Cell division protein CrgA.

Mol	Chain	Residues	Atoms						Trace
1	A	77	Total	C	H	N	O	S	0
			1239	411	633	97	92	6	
1	B	77	Total	C	H	N	O	S	0
			1239	411	633	97	92	6	

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP P9WP57
A	94	LEU	-	expression tag	UNP P9WP57
A	95	GLU	-	expression tag	UNP P9WP57
A	96	HIS	-	expression tag	UNP P9WP57
A	97	HIS	-	expression tag	UNP P9WP57
A	98	HIS	-	expression tag	UNP P9WP57
A	99	HIS	-	expression tag	UNP P9WP57
A	100	HIS	-	expression tag	UNP P9WP57
A	101	HIS	-	expression tag	UNP P9WP57
B	0	MET	-	initiating methionine	UNP P9WP57
B	94	LEU	-	expression tag	UNP P9WP57
B	95	GLU	-	expression tag	UNP P9WP57
B	96	HIS	-	expression tag	UNP P9WP57
B	97	HIS	-	expression tag	UNP P9WP57
B	98	HIS	-	expression tag	UNP P9WP57
B	99	HIS	-	expression tag	UNP P9WP57
B	100	HIS	-	expression tag	UNP P9WP57
B	101	HIS	-	expression tag	UNP P9WP57

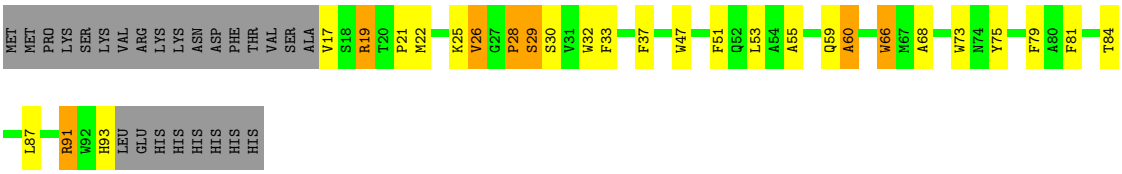
4 Residue-property plots

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.

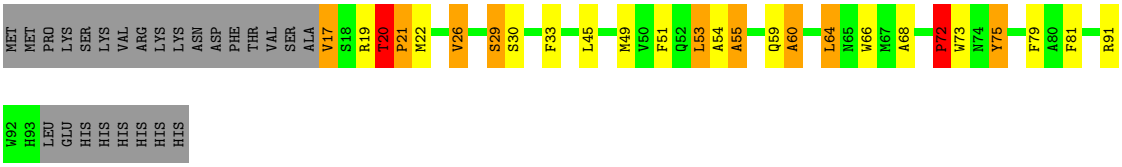
• Molecule 1: Cell division protein CrgA

Chain A:



• Molecule 1: Cell division protein CrgA

Chain B:

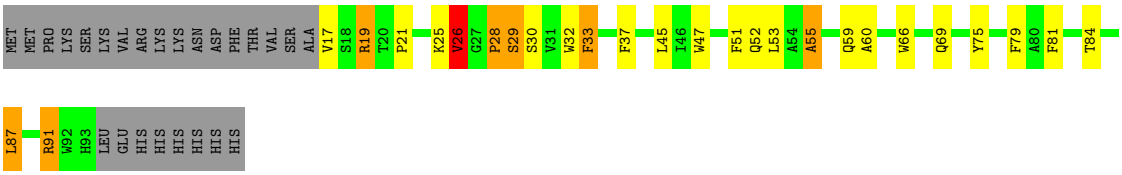


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 3. Colouring as in section 4.1 above.

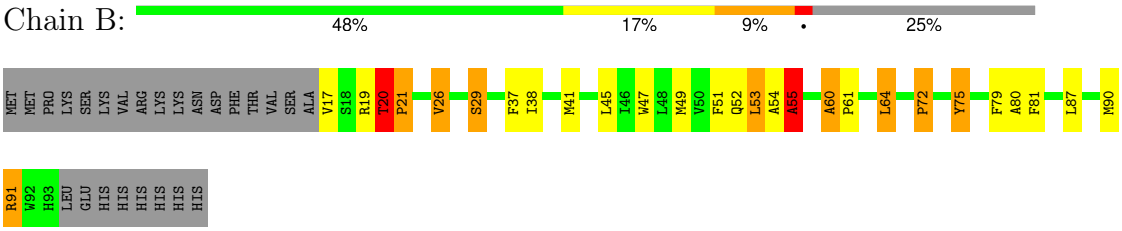
• Molecule 1: Cell division protein CrgA

Chain A:



• Molecule 1: Cell division protein CrgA

Chain B:


48%
17%
9%
25%
MET MET PRO LYS SER LYS VAL ARG LYS LYS ASN ASP PHE THR VAL SER ALA V17 S18 R19 T20 P21 V26 S29 F37 I38 M41 L45 I46 W47 L48 M49 V50 F51 Q52 L53 A54 A55 A60 P61 L64 P72 Y75 F79 A80 F81 L87 M90
R91 W92 H93 LEU GLU HIS HIS HIS HIS HIS HIS

5 Refinement protocol and experimental data overview

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

Of the 13 calculated structures, 10 were deposited, based on the following criterion: *structures with the least restraint violations*.

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

Software name	Classification	Version
NAMD	refinement	2.13
X-PLOR NIH	structure calculation	3.4

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	69
Number of shifts mapped to atoms	69
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	3%

Note: This is a solid-state NMR structure, where hydrogen atoms are typically not assigned a chemical shift value, which may lead to lower completeness of assignment measure.

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	1.62±0.04	3±2/626 (0.5± 0.2%)	2.19±0.10	25±3/854 (2.9± 0.4%)
1	B	1.67±0.07	5±2/626 (0.8± 0.3%)	2.32±0.08	26±4/854 (3.0± 0.4%)
All	All	1.65	76/12520 (0.6%)	2.25	505/17080 (3.0%)

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	6.9±1.4
1	B	1.0±0.0	6.6±1.0
All	All	10	135

5 of 55 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	B	20	THR	CA-CB	8.29	1.75	1.53	2	10
1	B	18	SER	CA-CB	8.04	1.65	1.52	7	2
1	A	51	PHE	CG-CD2	7.47	1.50	1.38	4	1
1	A	75	TYR	CE1-CZ	7.20	1.48	1.38	3	1
1	B	91	ARG	CD-NE	6.86	1.58	1.46	3	1

5 of 229 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	B	91	ARG	NE-CZ-NH1	24.44	132.52	120.30	9	7
1	A	91	ARG	NE-CZ-NH1	22.16	131.38	120.30	2	7
1	A	91	ARG	NE-CZ-NH2	-20.90	109.85	120.30	6	5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	B	20	THR	N-CA-C	18.31	160.44	111.00	8	10
1	A	19	ARG	NE-CZ-NH1	17.52	129.06	120.30	6	6

All unique chiral outliers are listed below.

Mol	Chain	Res	Type	Atoms	Models (Total)
1	B	20	THR	CA	10

5 of 27 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	19	ARG	Sidechain,Peptide,Mainchain	10
1	A	26	VAL	Peptide	10
1	B	20	THR	Peptide	10
1	B	29	SER	Peptide	10
1	B	72	PRO	Peptide	10

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	606	633	632	4±1
1	B	606	633	632	8±4
All	All	12120	12660	12640	114

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 5.

5 of 51 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:B:20:THR:CB	1:B:20:THR:CA	1.59	1.74	2	3
1:B:20:THR:CB	1:B:20:THR:C	1.16	2.13	10	3
1:B:20:THR:CB	1:B:20:THR:N	1.09	2.15	2	3
1:B:20:THR:N	1:B:20:THR:HB	0.96	1.74	9	3

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:B:20:THR:CA	1:B:20:THR:HB	0.89	1.97	9	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	75/102 (74%)	68±1 (90±1%)	5±1 (6±1%)	3±1 (4±1%)	4	31
1	B	75/102 (74%)	68±1 (91±1%)	4±1 (5±2%)	3±1 (4±1%)	4	30
All	All	1500/2040 (74%)	1360 (91%)	81 (5%)	59 (4%)	4	31

5 of 11 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	60	ALA	10
1	B	60	ALA	10
1	B	72	PRO	10
1	A	28	PRO	8
1	A	21	PRO	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	63/87 (72%)	58±2 (92±3%)	5±2 (8±3%)	12	60
1	B	63/87 (72%)	55±2 (88±3%)	8±2 (12±3%)	6	48
All	All	1260/1740 (72%)	1130 (90%)	130 (10%)	8	53

5 of 33 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	66	TRP	10
1	A	91	ARG	10
1	B	17	VAL	10
1	B	64	LEU	10
1	A	59	GLN	9

6.3.3 RNA

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates

There are no oligosaccharides in this entry.

6.6 Ligand geometry

There are no ligands in this entry.

6.7 Other polymers

There are no such molecules in this entry.

6.8 Polymer linkage issues

There are no chain breaks in this entry.

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 3% for the well-defined parts and 3% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chemical_shifts_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	69
Number of shifts mapped to atoms	69
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	65

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	0	—	None (insufficient data)
$^{13}C_{\beta}$	0	—	None (insufficient data)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	69	-7.01 ± 8.19	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 3%, i.e. 69 atoms were assigned a chemical shift out of a possible 2198. 0 out of 34 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	69/766 (9%)	0/312 (0%)	0/308 (0%)	69/146 (47%)
Sidechain	0/1180 (0%)	0/798 (0%)	0/356 (0%)	0/26 (0%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	0/252 (0%)	0/126 (0%)	0/114 (0%)	0/12 (0%)
Overall	69/2198 (3%)	0/1236 (0%)	0/778 (0%)	69/184 (38%)

Note: This is a solid-state NMR structure, where hydrogen atoms are typically not assigned a chemical shift value, which may lead to lower completeness of assignment measure.

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.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	35	SER	N	228.00	99.14 – 133.45	32.6
1	A	30	SER	N	225.00	99.14 – 133.45	31.7
1	A	39	GLY	N	220.00	91.59 – 127.52	30.7
1	A	80	ALA	N	229.00	106.13 – 140.55	30.7
1	A	49	MET	N	222.00	102.99 – 137.21	29.8
1	A	87	LEU	N	230.00	102.77 – 140.89	28.4
1	A	76	ALA	N	221.00	106.13 – 140.55	28.4
1	A	42	LEU	N	226.00	102.77 – 140.89	27.3
1	A	45	LEU	N	225.00	102.77 – 140.89	27.1
1	A	73	TRP	N	225.00	101.51 – 141.60	25.8
1	A	51	PHE	N	225.00	99.93 – 140.82	25.6
1	A	88	LEU	N	219.00	102.77 – 140.89	25.5
1	A	85	GLY	N	201.00	91.59 – 127.52	25.4
1	A	90	MET	N	207.00	102.99 – 137.21	25.4
1	A	46	ILE	N	227.00	100.55 – 142.30	25.3
1	A	48	LEU	N	216.00	102.77 – 140.89	24.7
1	A	41	MET	N	203.00	102.99 – 137.21	24.2
1	A	79	PHE	N	219.00	99.93 – 140.82	24.1
1	A	38	ILE	N	222.00	100.55 – 142.30	24.1
1	A	84	THR	N	227.00	91.89 – 138.78	23.8
1	A	77	ILE	N	220.00	100.55 – 142.30	23.6
1	A	83	ILE	N	219.00	100.55 – 142.30	23.4
1	A	74	ASN	N	209.00	99.66 – 138.23	23.4
1	A	44	GLY	N	193.00	91.59 – 127.52	23.2
1	A	31	VAL	N	221.00	99.23 – 142.92	22.9
1	A	32	TRP	N	213.00	101.51 – 141.60	22.8
1	A	92	TRP	N	213.00	101.51 – 141.60	22.8
1	A	78	ALA	N	201.00	106.13 – 140.55	22.6
1	A	36	LEU	N	207.00	102.77 – 140.89	22.3

Continued on next page...

Continued from previous page...

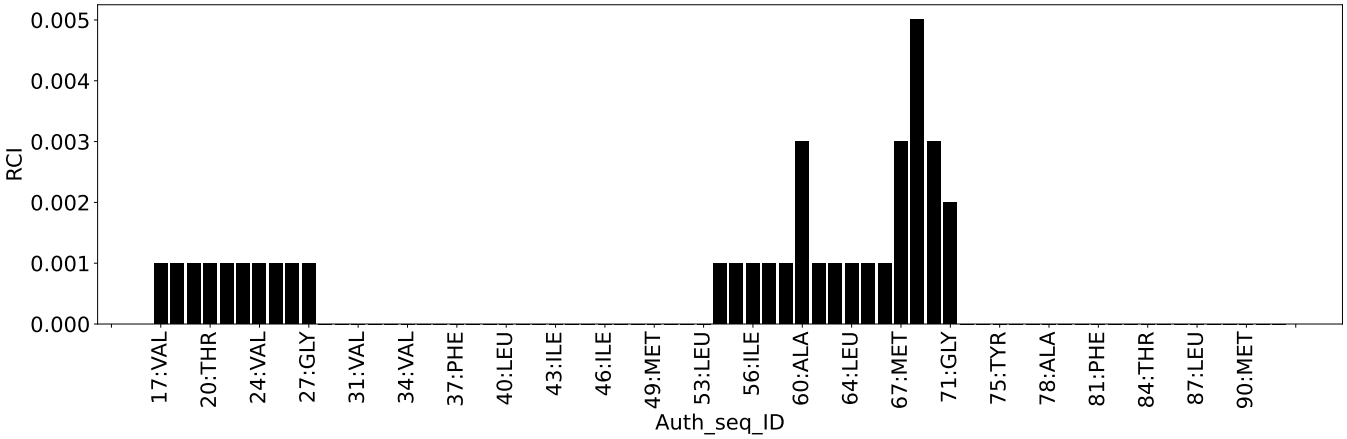
List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	37	PHE	N	209.00	99.93 – 140.82	21.7
1	A	81	PHE	N	209.00	99.93 – 140.82	21.7
1	A	82	MET	N	194.00	102.99 – 137.21	21.6
1	A	86	LEU	N	204.00	102.77 – 140.89	21.6
1	A	89	THR	N	216.00	91.89 – 138.78	21.5
1	A	43	ILE	N	209.00	100.55 – 142.30	21.0
1	A	34	VAL	N	211.00	99.23 – 142.92	20.6
1	A	33	PHE	N	204.00	99.93 – 140.82	20.4
1	A	40	LEU	N	198.00	102.77 – 140.89	20.0
1	A	29	SER	N	179.00	99.14 – 133.45	18.3
1	A	75	TYR	N	190.00	100.12 – 140.79	17.1
1	A	47	TRP	N	188.00	101.51 – 141.60	16.6
1	A	91	ARG	N	177.00	102.91 – 138.82	15.6
1	A	54	ALA	N	173.00	106.13 – 140.55	14.4
1	A	53	LEU	N	175.00	102.77 – 140.89	13.9
1	A	19	ARG	N	73.00	102.91 – 138.82	-13.3
1	A	63	ALA	N	78.00	106.13 – 140.55	-13.2
1	A	64	LEU	N	74.00	102.77 – 140.89	-12.6
1	A	55	ALA	N	166.00	106.13 – 140.55	12.4
1	A	17	VAL	N	175.00	99.23 – 142.92	12.3
1	A	23	LYS	N	77.00	102.74 – 139.42	-12.0
1	A	25	LYS	N	77.00	102.74 – 139.42	-12.0
1	A	18	SER	N	77.00	99.14 – 133.45	-11.4
1	A	58	SER	N	77.00	99.14 – 133.45	-11.4
1	A	50	VAL	N	171.00	99.23 – 142.92	11.4
1	A	62	THR	N	62.00	91.89 – 138.78	-11.4
1	A	24	VAL	N	72.00	99.23 – 142.92	-11.2
1	A	65	ASN	N	77.00	99.66 – 138.23	-10.9
1	A	26	VAL	N	74.00	99.23 – 142.92	-10.8
1	A	56	ILE	N	78.00	100.55 – 142.30	-10.4
1	A	57	GLY	N	74.00	91.59 – 127.52	-9.9
1	A	66	TRP	N	85.00	101.51 – 141.60	-9.1
1	A	22	MET	N	89.00	102.99 – 137.21	-9.1
1	A	20	THR	N	74.00	91.89 – 138.78	-8.8
1	A	27	GLY	N	78.00	91.59 – 127.52	-8.8
1	A	70	LEU	N	97.00	102.77 – 140.89	-6.5

7.1.5 Random Coil Index (RCI) plots

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)

No dihedral-angle restraints found