W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB NMR Structure Validation Summary Report ⓘ

Jun 22, 2024 – 05:25 PM EDT

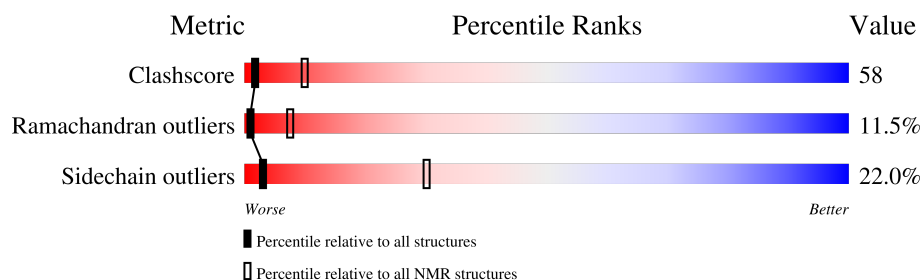|		
PDB ID	:	6BA6
BMRB ID	:	30353
Title	:	Solution structure of Rap1b/talin complex
Authors	:	Zhu, L.; Yang, J.; Qin, J.
Deposited on	:	2017-10-12

1   Overall quality at a glance ⓘ

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

The overall completeness of chemical shifts assignment is 73%.

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	Percentile Ranks	Value
Clashscore		58
Ramachandran outliers		11.5%
Sidechain outliers		22.0%

Worse   *Better*

▮ Percentile relative to all structures
⧫ Percentile relative to all NMR structures

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	93	16% 59% 15% ·· 8%
2	B	168	15% 68% 15% ··

2 Ensemble composition and analysis ⓘ

This entry contains 20 models. Model 2 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:2-A:86, B:1-B:167 (252)	0.73	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 4 clusters and 6 single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4045 atoms, of which 2022 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Talin-1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	86	1384	437	696	117	127	7	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	GLY	-	expression tag	UNP P26039
A	-5	ALA	-	expression tag	UNP P26039
A	-4	MET	-	expression tag	UNP P26039
A	-3	ASP	-	expression tag	UNP P26039
A	-2	PRO	-	expression tag	UNP P26039
A	-1	GLU	-	expression tag	UNP P26039
A	0	PHE	-	expression tag	UNP P26039

- Molecule 2 is a protein called Ras-related protein Rap-1b.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	167	2661	836	1326	228	263	8	0

There are 2 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
B	0	HIS	-	expression tag	UNP P61224
B	12	VAL	GLY	engineered mutation	UNP P61224

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: Talin-1

Chain A:
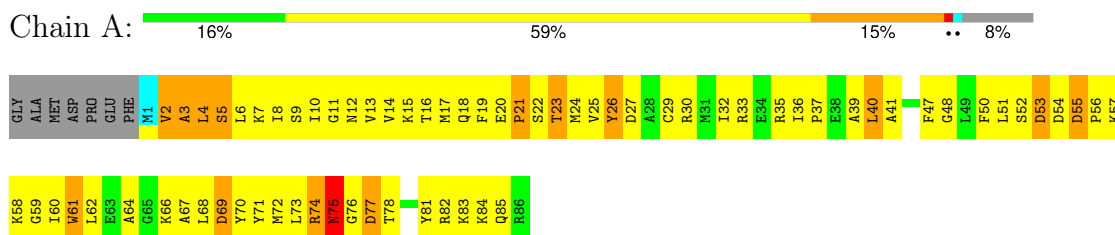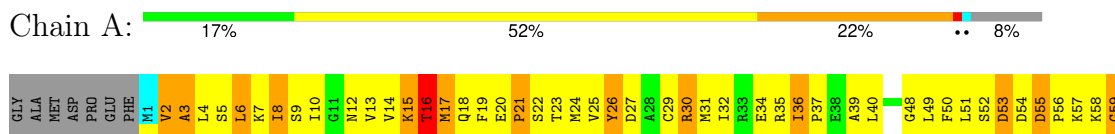

- Molecule 2: Ras-related protein Rap-1b

Chain B:
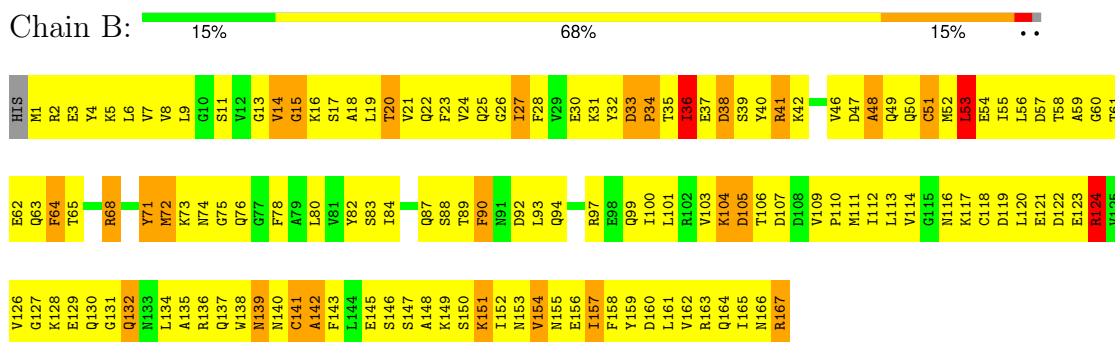
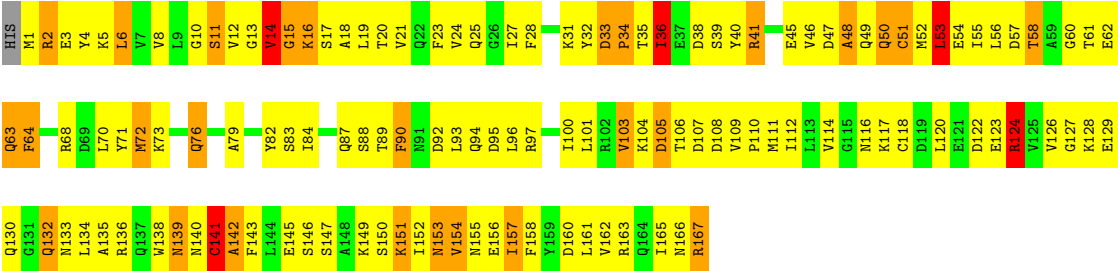## 4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 2. Colouring as in section 4.1 above.

- Molecule 1: Talin-1

Chain A:

I60 W61 L62 E63 A64 A67 L68 D69 Y70 Y71 M72 L73 R74 N75 G76 D77 T78 M79 E80 Y81 R82 K83 K84 Q85 R86

• Molecule 2: Ras-related protein Rap-1b

Chain B:

23%　　　　　　　　58%　　　　　　　　16%　　• •

HIS M1 R2 E3 Y4 K5 L6 V7 V8 L9 G10 S11 V12 G13 V14 G15 K16 S17 A18 L19 T20 V21 Q22 F23 V24 Q25 G26 I27 F28 K31 Y32 D33 P34 T35 I36 E37 D38 S39 Y40 R41 E45 V46 D47 A48 Q49 Q50 C51 M52 L53 E54 I55 L56 D57 T58 A59 G60 T61 E62

Q63 F64 R68 D69 L70 Y71 M72 K73 Q76 A79 Y82 S83 I84 Q87 S88 T89 F90 N91 D92 L93 Q94 D95 L96 R97 I100 L101 R102 V103 K104 D105 T106 D107 D108 V109 P110 M111 I112 L113 V114 G115 N116 K117 C118 D119 L120 E121 D122 E123 R124 V125 V126 G127 K128 E129

Q130 G131 Q132 N133 L134 A135 R136 Q137 W138 N139 N140 C141 A142 F143 L44 E145 S146 S147 A148 K149 S150 K151 I152 N153 V154 N155 E156 I157 F158 Y159 D160 L161 V162 R163 Q164 I165 N166 R167

5   Refinement protocol and experimental data overview   ⓘ

The models were refined using the following method: *simulated annealing.*

Of the 50 calculated structures, 20 were deposited, based on the following criterion: *20 structures for lowest energy.*

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

Software name	Classification	Version
X-PLOR NIH	refinement	
X-PLOR NIH	structure calculation	

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	3
Total number of shifts	4180
Number of shifts mapped to atoms	4180
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	73%

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	1.10±0.01	0±0/693 (0.0± 0.0%)	0.98±0.01	0±0/930 (0.0± 0.0%)
2	B	1.12±0.01	0±0/1351 (0.0± 0.0%)	1.00±0.01	1±1/1819 (0.0± 0.0%)
All	All	1.12	0/40880 (0.0%)	1.00	15/54980 (0.0%)

There are no bond-length outliers.

5 of 7 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
2	B	64	PHE	CB-CG-CD1	6.67	125.47	120.80	18	1
2	B	64	PHE	CB-CG-CD2	-6.01	116.60	120.80	18	1
2	B	71	TYR	CB-CG-CD1	5.75	124.45	121.00	11	2
2	B	4	TYR	CB-CG-CD2	-5.70	117.58	121.00	5	1
2	B	90	PHE	CB-CG-CD2	-5.52	116.93	120.80	15	2

There are no chirality outliers.

There are no planarity outliers.

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	680	687	687	83±8
2	B	1335	1326	1328	164±13
All	All	40300	40260	40300	4653

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 58.

5 of 1440 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Models Total
2:B:161:LEU:HD23	2:B:161:LEU:H	0.95	1.16	6	9
2:B:20:THR:O	2:B:22:GLN:N	0.91	2.04	16	3
1:A:3:ALA:HB1	1:A:19:PHE:O	0.86	1.71	17	15
1:A:35:ARG:NE	1:A:36:ILE:H	0.85	1.70	13	1
2:B:64:PHE:O	2:B:68:ARG:N	0.84	2.10	18	17

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	84/93 (90%)	51±3 (61±3%)	17±3 (20±4%)	16±2 (19±3%)	0	2
2	B	165/168 (98%)	129±3 (78±2%)	24±3 (14±2%)	12±1 (7±1%)	2	15
All	All	4980/5220 (95%)	3602 (72%)	804 (16%)	574 (12%)	1	7

5 of 60 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	21	PRO	20
1	A	53	ASP	20
1	A	56	PRO	20
1	A	69	ASP	20
1	A	75	ASN	20

6.3.2　Protein sidechains　ⓘ

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	73/79 (92%)	60±3 (82±4%)	13±3 (18±4%)	4	37
2	B	148/149 (99%)	113±4 (76±3%)	35±4 (24±3%)	2	27
All	All	4420/4560 (97%)	3447 (78%)	973 (22%)	3	30

5 of 146 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	26	TYR	20
1	A	61	TRP	20
1	A	83	LYS	20
1	A	84	LYS	20
2	B	36	ILE	20

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 monosaccharides 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 (i)

The completeness of assignment taking into account all chemical shift lists is 73% for the well-defined parts and 73% 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	1695
Number of shifts mapped to atoms	1695
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	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
$^{13}C_\alpha$	164	-0.10 ± 0.14	None needed (< 0.5 ppm)
$^{13}C_\beta$	153	0.56 ± 0.15	Should be checked
$^{13}C'$	156	0.32 ± 0.15	None needed (< 0.5 ppm)
^{15}N	159	0.36 ± 0.33	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 48%, i.e. 1688 atoms were assigned a chemical shift out of a possible 3504. 0 out of 44 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	808/1262 (64%)	331/513 (65%)	318/504 (63%)	159/245 (65%)
Sidechain	880/2028 (43%)	573/1310 (44%)	307/628 (49%)	0/90 (0%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	0/214 (0%)	0/102 (0%)	0/110 (0%)	0/2 (0%)
Overall	1688/3504 (48%)	904/1925 (47%)	625/1242 (50%)	159/337 (47%)

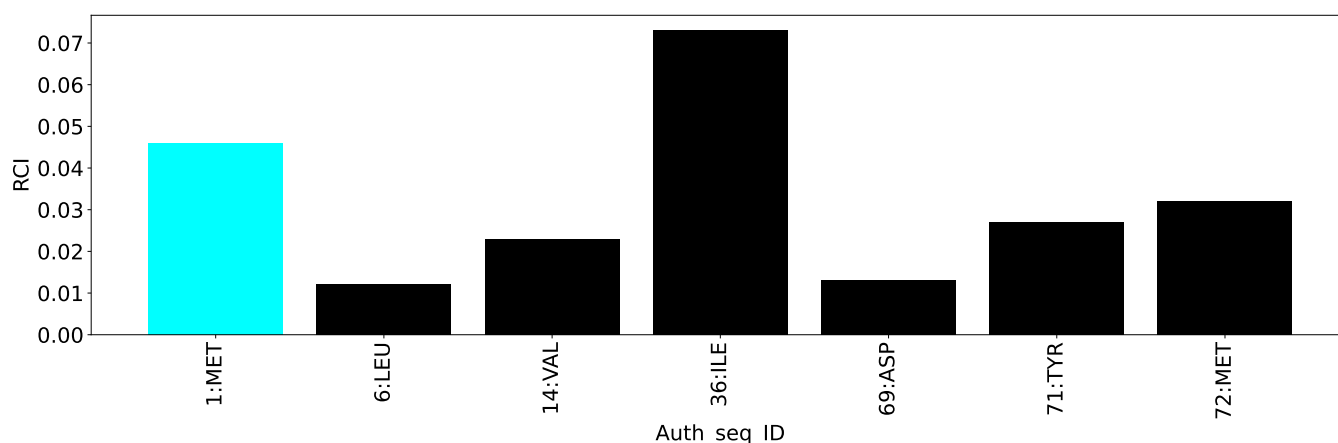### 7.1.4   Statistically unusual chemical shifts ⓘ

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.

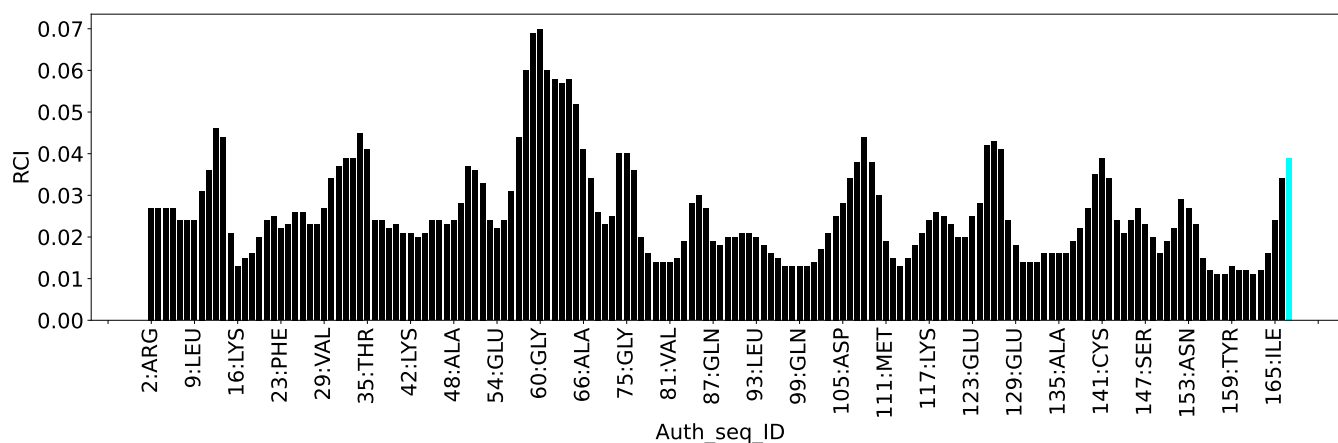| List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	B	134	LEU	HG	-0.28	$-0.13 - 3.16$	-5.5

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:

Random coil index (RCI) for chain B:

7.2 Chemical shift list 2

File name: working_cs.cif

Chemical shift list name: *talin-F0-bound-nc-noesy.str_ assigned_ chem_ shift_ list_ 1*

7.2.1 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	884
Number of shifts mapped to atoms	884
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	2

7.2.2 Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	84	0.18 ± 0.12	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	79	0.43 ± 0.15	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	73	-0.14 ± 0.28	None needed (< 0.5 ppm)

7.2.3   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 25%, i.e. 876 atoms were assigned a chemical shift out of a possible 3504. 0 out of 44 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	**^{1}H**	**^{13}C**	**^{15}N**
Backbone	319/1262 (25%)	163/513 (32%)	83/504 (16%)	73/245 (30%)
Sidechain	555/2028 (27%)	367/1310 (28%)	183/628 (29%)	5/90 (6%)
Aromatic	2/214 (1%)	1/102 (1%)	0/110 (0%)	1/2 (50%)
Overall	876/3504 (25%)	531/1925 (28%)	266/1242 (21%)	79/337 (23%)

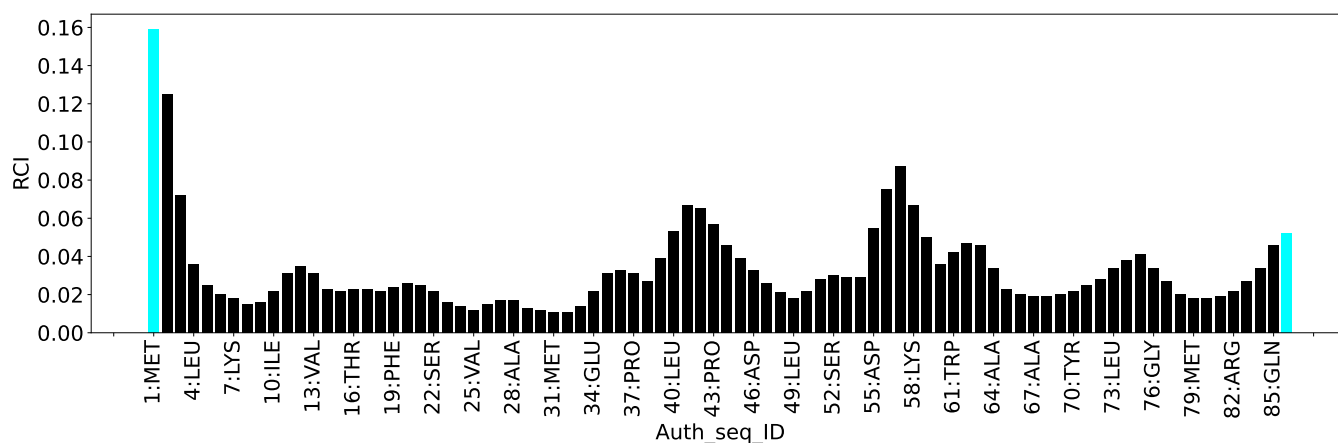### 7.2.4   Statistically unusual chemical shifts ⓘ

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
2	A	65	GLY	N	130.83	$91.59 - 127.52$	5.9
2	A	66	LYS	HD3	0.44	$0.54 - 2.65$	-5.5

7.2.5   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:

7.3   Chemical shift list 3

File name: working_cs.cif

Chemical shift list name: *rap1b-bound-nc-noesy.str_ assigned_ chem_ shift_ list_ 1*

7.3.1   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	1601
Number of shifts mapped to atoms	1601
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	1

7.3.2   Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	166	0.04 ± 0.09	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	156	0.50 ± 0.20	Should be checked
$^{13}C'$	0	—	None (insufficient data)
^{15}N	159	0.31 ± 0.40	None needed (< 0.5 ppm)

WORLDWIDE
wwPDB
PROTEIN DATA BANK

7.3.3   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 46%, i.e. 1595 atoms were assigned a chemical shift out of a possible 3504. 0 out of 44 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	**^{1}H**	**^{13}C**	**^{15}N**
Backbone	658/1262 (52%)	334/513 (65%)	165/504 (33%)	159/245 (65%)
Sidechain	935/2028 (46%)	609/1310 (46%)	325/628 (52%)	1/90 (1%)
Aromatic	2/214 (1%)	1/102 (1%)	0/110 (0%)	1/2 (50%)
Overall	1595/3504 (46%)	944/1925 (49%)	490/1242 (39%)	161/337 (48%)

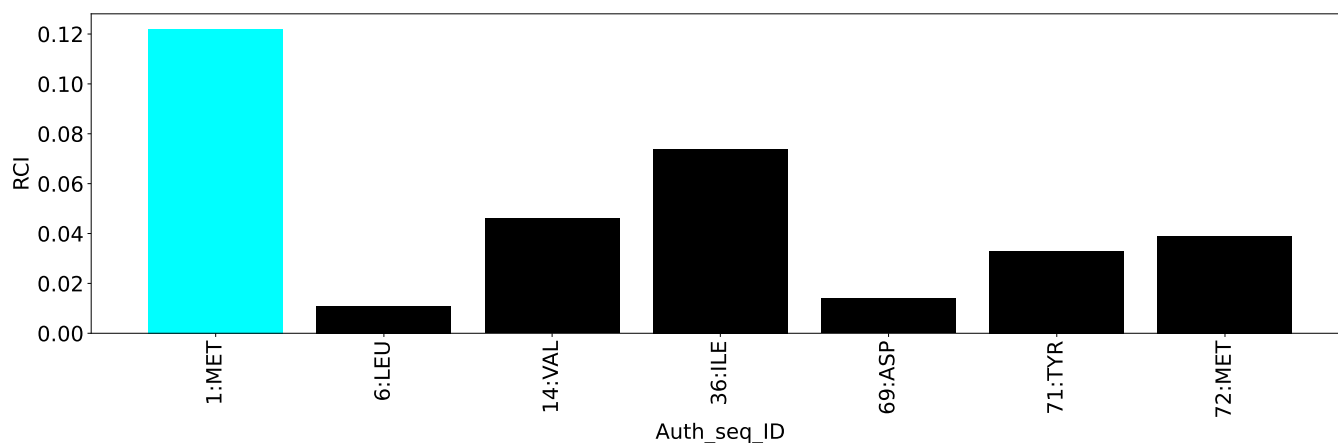### 7.3.4   Statistically unusual chemical shifts ⓘ

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
3	B	134	LEU	HG	-0.47	$-0.13 - 3.16$	-6.0

7.3.5   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:

Random coil index (RCI) for chain B: