wwPDB NMR Structure Validation Summary Report (i)

Apr 20, 2024 – 08:28 AM EDT

PDB ID	:	2NBW
BMRB ID	:	25825
Title	:	Solution structure of the Rpn1 T1 site with the Rad23 UBL domain
Authors	:	Chen, X.; Walters, K.J.
Deposited on	:	2016-03-14

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:

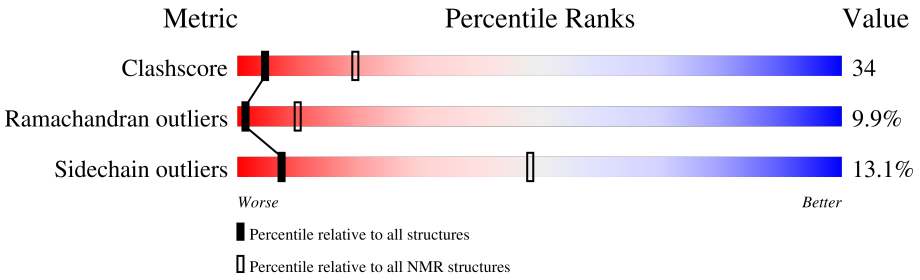
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 39%.

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	131	47% 36% • 15%
2	B	78	33% 41% 22% •

2 Ensemble composition and analysis (i)

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: *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:497-A:607, B:1-B:75 (186)	1.16	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 and 1 single-model cluster was found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3169 atoms, of which 1608 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called 26S proteasome regulatory subunit RPN1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	131	1921	606	967	150	192	6	0

- Molecule 2 is a protein called UV excision repair protein RAD23.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	78	1248	383	641	97	123	4	0

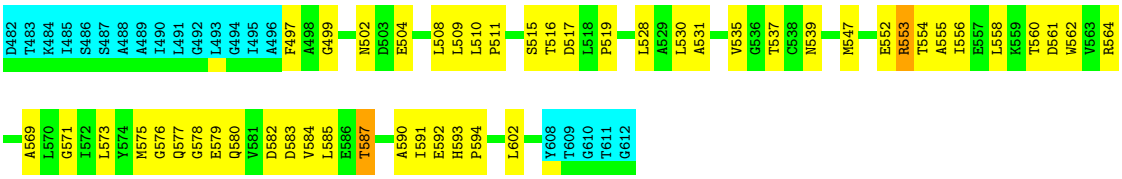
4 Residue-property plots (i)

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.

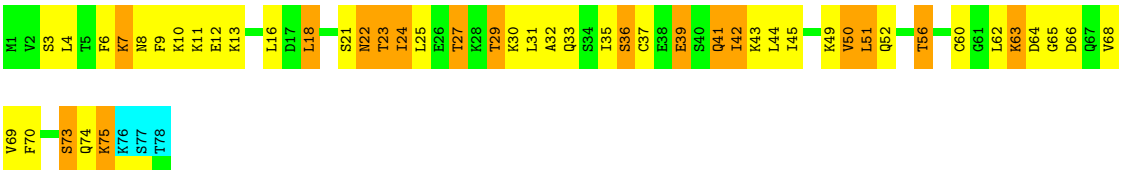
• Molecule 1: 26S proteasome regulatory subunit RPN1

Chain A: 47% 36% • 15%



• Molecule 2: UV excision repair protein RAD23

Chain B: 33% 41% 22% •



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

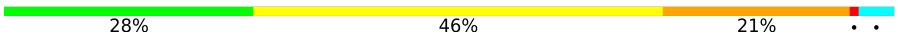
The representative model is number 3. Colouring as in section 4.1 above.

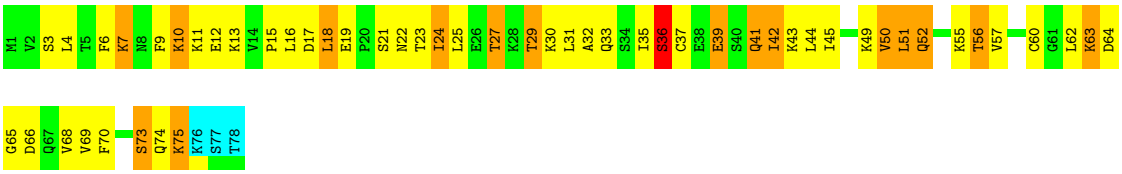
• Molecule 1: 26S proteasome regulatory subunit RPN1

Chain A: 50% 30% 5% 15%

• Molecule 2: UV excision repair protein RAD23

Chain B:





5 Refinement protocol and experimental data overview (i)

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

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
X-PLOR NIH	refinement	
X-PLOR NIH	structure solution	

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	1177
Number of shifts mapped to atoms	1003
Number of unparsed shifts	0
Number of shifts with mapping errors	174
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	39%

6 Model quality (i)

6.1 Standard geometry (i)

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

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	1.6±0.5
All	All	0	16

There are no bond-length outliers.

There are no bond-angle outliers.

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	564	ARG	Sidechain	10
1	A	553	ARG	Sidechain	6

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	820	826	826	38±5
2	B	584	616	616	61±6
All	All	14040	14420	14420	981

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 34.

5 of 371 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:B:32:ALA:HB1	2:B:37:CYS:O	1.00	1.56	8	10
2:B:4:LEU:HD23	2:B:66:ASP:H	0.91	1.25	1	10
2:B:24:ILE:O	2:B:28:LYS:HG3	0.87	1.69	8	1
2:B:45:ILE:HG23	2:B:50:VAL:N	0.82	1.89	2	10
2:B:25:LEU:C	2:B:25:LEU:HD13	0.80	1.96	4	6

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	111/131 (85%)	64±5 (58±4%)	34±3 (30±3%)	13±3 (12±2%)	1 7
2	B	74/78 (95%)	59±1 (80±2%)	10±1 (14±2%)	5±1 (7±1%)	2 17
All	All	1850/2090 (89%)	1230 (66%)	437 (24%)	183 (10%)	1 10

5 of 57 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
2	B	50	VAL	10
2	B	31	LEU	9
2	B	36	SER	9
1	A	591	ILE	8
2	B	51	LEU	8

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	89/102 (87%)	84±1 (94±2%)	5±1 (6±2%)	22 71
2	B	71/74 (96%)	56±2 (78±3%)	16±2 (22±3%)	3 30

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1600/1760 (91%)	1391 (87%)	209 (13%)	7 48

5 of 56 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)
2	B	7	LYS	10
2	B	24	ILE	10
2	B	27	THR	10
2	B	29	THR	10
2	B	41	GLN	10

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 39% for the well-defined parts and 37% 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	1177
Number of shifts mapped to atoms	1003
Number of unparsed shifts	0
Number of shifts with mapping errors	174
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

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 174) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	B	79	ASP	H	8.28	0	1
1	B	79	ASP	HA	4.878	0	1
1	B	79	ASP	HB2	2.774	0	2
1	B	79	ASP	HB3	2.574	0	2
1	B	79	ASP	C	175.013	0	1
1	B	79	ASP	CA	52.087	0	1
1	B	79	ASP	CB	41.205	0	1
1	B	79	ASP	N	124.382	0	1
1	B	80	PRO	HA	4.423	0	1
1	B	80	PRO	HB2	2.3	0	2
1	B	80	PRO	HB3	1.966	0	2
1	B	80	PRO	HG2	2.015	0	2
1	B	80	PRO	HG3	2.015	0	2
1	B	80	PRO	HD2	3.889	0	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	B	80	PRO	HD3	3.835	0	2
1	B	80	PRO	C	177.067	0	1
1	B	80	PRO	CA	63.802	0	1
1	B	80	PRO	CB	32.021	0	1
1	B	80	PRO	CG	26.959	0	1
1	B	80	PRO	CD	50.583	0	1
1	B	81	ASN	H	8.52	0	1
1	B	81	ASN	HA	4.736	0	1
1	B	81	ASN	HB2	2.863	0	2
1	B	81	ASN	HB3	2.789	0	2
1	B	81	ASN	HD21	6.95	0	2
1	B	81	ASN	HD22	7.756	0	2
1	B	81	ASN	C	175.519	0	1
1	B	81	ASN	CA	53.396	0	1
1	B	81	ASN	CB	38.771	0	1
1	B	81	ASN	N	117.478	0	1
1	B	81	ASN	ND2	113.877	0	1
1	B	82	SER	H	8.023	0	1
1	B	82	SER	HA	4.402	0	1
1	B	82	SER	HB2	3.875	0	2
1	B	82	SER	HB3	3.875	0	2
1	B	82	SER	C	174.653	0	1
1	B	82	SER	CA	58.739	0	1
1	B	82	SER	CB	63.802	0	1
1	B	82	SER	N	115.818	0	1
1	B	83	SER	H	8.331	0	1
1	B	83	SER	HA	4.491	0	1
1	B	83	SER	HB2	3.894	0	2
1	B	83	SER	HB3	3.894	0	2
1	B	83	SER	C	174.662	0	1
1	B	83	SER	CA	58.458	0	1
1	B	83	SER	CB	63.802	0	1
1	B	83	SER	N	117.494	0	1
1	B	84	SER	H	8.283	0	1
1	B	84	SER	C	174.886	0	1
1	B	84	SER	CA	58.461	0	1
1	B	84	SER	CB	63.523	0	1
1	B	84	SER	N	117.712	0	1
1	B	85	VAL	H	8.034	0	1
1	B	85	VAL	HA	4.062	0	1
1	B	85	VAL	HB	2.098	0	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	B	85	VAL	HG11	0.931	0	2
1	B	85	VAL	HG12	0.931	0	2
1	B	85	VAL	HG13	0.931	0	2
1	B	85	VAL	HG21	0.938	0	2
1	B	85	VAL	HG22	0.938	0	2
1	B	85	VAL	HG23	0.938	0	2
1	B	85	VAL	C	176.142	0	1
1	B	85	VAL	CA	62.889	0	1
1	B	85	VAL	CB	32.518	0	1
1	B	85	VAL	CG1	20.885	0	2
1	B	85	VAL	CG2	20.483	0	2
1	B	85	VAL	N	121.117	0	1
1	B	86	ASP	H	8.269	0	1
1	B	86	ASP	HA	4.565	0	1
1	B	86	ASP	HB2	2.713	0	2
1	B	86	ASP	HB3	2.612	0	2
1	B	86	ASP	C	176.766	0	1
1	B	86	ASP	CA	54.546	0	1
1	B	86	ASP	CB	40.752	0	1
1	B	86	ASP	N	123.029	0	1
1	B	87	LYS	H	8.172	0	1
1	B	87	LYS	HA	4.196	0	1
1	B	87	LYS	HB2	1.858	0	2
1	B	87	LYS	HB3	1.773	0	2
1	B	87	LYS	HG2	1.476	0	2
1	B	87	LYS	HG3	1.417	0	2
1	B	87	LYS	HD2	1.664	0	2
1	B	87	LYS	HD3	1.664	0	2
1	B	87	LYS	HE2	2.986	0	2
1	B	87	LYS	HE3	2.986	0	2
1	B	87	LYS	C	177.35	0	1
1	B	87	LYS	CA	57.052	0	1
1	B	87	LYS	CB	32.583	0	1
1	B	87	LYS	CG	24.709	0	1
1	B	87	LYS	CD	28.927	0	1
1	B	87	LYS	CE	41.865	0	1
1	B	87	LYS	N	121.906	0	1
1	B	88	LEU	H	8.101	0	1
1	B	88	LEU	HA	4.261	0	1
1	B	88	LEU	HB2	1.713	0	2
1	B	88	LEU	HB3	1.592	0	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	B	88	LEU	HG	1.62	0	1
1	B	88	LEU	HD11	0.909	0	2
1	B	88	LEU	HD12	0.909	0	2
1	B	88	LEU	HD13	0.909	0	2
1	B	88	LEU	HD21	0.851	0	2
1	B	88	LEU	HD22	0.851	0	2
1	B	88	LEU	HD23	0.851	0	2
1	B	88	LEU	C	177.798	0	1
1	B	88	LEU	CA	55.646	0	1
1	B	88	LEU	CB	41.583	0	1
1	B	88	LEU	CG	26.959	0	1
1	B	88	LEU	CD1	24.709	0	2
1	B	88	LEU	CD2	23.302	0	2
1	B	88	LEU	N	121.846	0	1
1	B	89	ALA	H	8.001	0	1
1	B	89	ALA	HA	4.189	0	1
1	B	89	ALA	HB1	1.395	0	2
1	B	89	ALA	HB2	1.395	0	2
1	B	89	ALA	HB3	1.395	0	2
1	B	89	ALA	C	178.382	0	1
1	B	89	ALA	CA	53.389	0	1
1	B	89	ALA	CB	18.521	0	1
1	B	89	ALA	N	123.573	0	1
1	B	90	ALA	H	8.023	0	1
1	B	90	ALA	HA	4.232	0	1
1	B	90	ALA	HB1	1.395	0	2
1	B	90	ALA	HB2	1.395	0	2
1	B	90	ALA	HB3	1.395	0	2
1	B	90	ALA	C	178.119	0	1
1	B	90	ALA	CA	52.833	0	1
1	B	90	ALA	CB	19.084	0	1
1	B	90	ALA	N	121.852	0	1
1	B	91	ALA	H	7.95	0	1
1	B	91	ALA	HA	4.232	0	1
1	B	91	ALA	HB1	1.394	0	2
1	B	91	ALA	HB2	1.394	0	2
1	B	91	ALA	HB3	1.394	0	2
1	B	91	ALA	C	178.1	0	1
1	B	91	ALA	CA	52.552	0	1
1	B	91	ALA	CB	18.802	0	1
1	B	91	ALA	N	121.781	0	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	B	92	LEU	H	7.928	0	1
1	B	92	LEU	HA	4.24	0	1
1	B	92	LEU	HB2	1.644	0	2
1	B	92	LEU	HB3	1.519	0	2
1	B	92	LEU	HG	1.641	0	1
1	B	92	LEU	HD11	0.889	0	2
1	B	92	LEU	HD12	0.889	0	2
1	B	92	LEU	HD13	0.889	0	2
1	B	92	LEU	HD21	0.835	0	2
1	B	92	LEU	HD22	0.835	0	2
1	B	92	LEU	HD23	0.835	0	2
1	B	92	LEU	C	177.584	0	1
1	B	92	LEU	CA	55.109	0	1
1	B	92	LEU	CB	42.159	0	1
1	B	92	LEU	CG	26.676	0	1
1	B	92	LEU	CD1	24.987	0	2
1	B	92	LEU	CD2	23.298	0	2
1	B	92	LEU	N	120.154	0	1
1	B	93	GLU	H	8.061	0	1
1	B	93	GLU	HA	4.16	0	1
1	B	93	GLU	HB2	1.897	0	2
1	B	93	GLU	HB3	1.897	0	2
1	B	93	GLU	HG2	2.205	0	2
1	B	93	GLU	HG3	2.121	0	2
1	B	93	GLU	C	176.308	0	1
1	B	93	GLU	CA	56.771	0	1
1	B	93	GLU	CB	30.052	0	1
1	B	93	GLU	CG	35.958	0	1
1	B	93	GLU	N	120.382	0	1
1	B	94	HIS	H	8.18	0	1
1	B	94	HIS	HA	4.565	0	1
1	B	94	HIS	HB2	3.057	0	2
1	B	94	HIS	HB3	2.98	0	2
1	B	94	HIS	C	174.955	0	1
1	B	94	HIS	CA	55.927	0	1
1	B	94	HIS	CB	30.333	0	1
1	B	94	HIS	N	119.338	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
$^{13}C_{\alpha}$	94	-0.01 ± 0.06	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	91	0.12 ± 0.15	None needed (< 0.5 ppm)
$^{13}C'$	94	-0.05 ± 0.23	None needed (< 0.5 ppm)
^{15}N	90	0.42 ± 0.40	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 39%, i.e. 967 atoms were assigned a chemical shift out of a possible 2460. 0 out of 38 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	372/931 (40%)	150/378 (40%)	150/372 (40%)	72/181 (40%)
Sidechain	580/1415 (41%)	394/932 (42%)	179/452 (40%)	7/31 (23%)
Aromatic	15/114 (13%)	15/57 (26%)	0/54 (0%)	0/3 (0%)
Overall	967/2460 (39%)	559/1367 (41%)	329/878 (37%)	79/215 (37%)

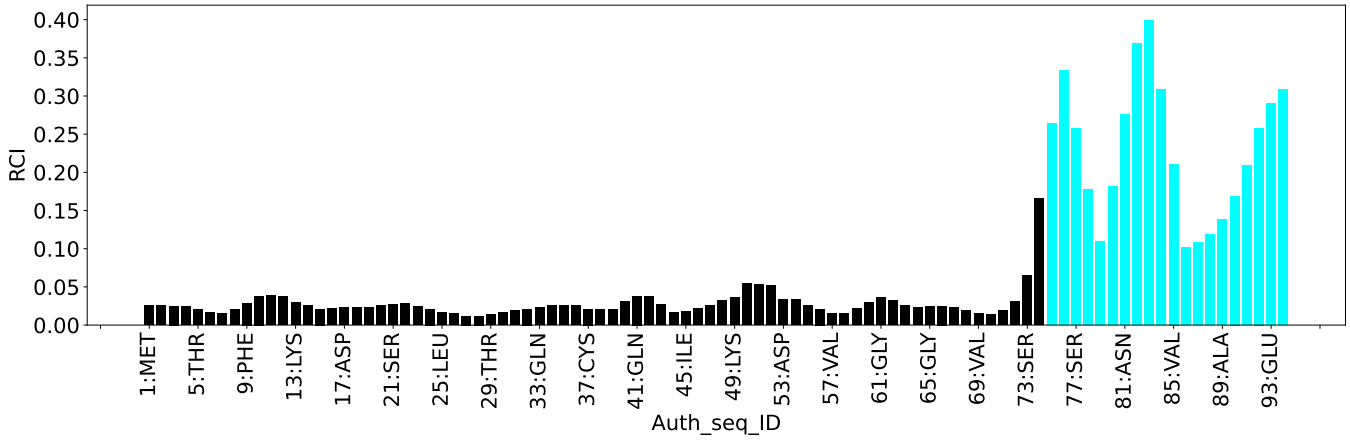
7.1.4 Statistically unusual chemical shifts (i)

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 B:


RCI
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
1:MET
5:THR
9:PHE
13:LYS
17:ASP
21:SER
25:LEU
29:THR
33:GLN
37:CYS
41:GLN
45:ILE
49:LYS
53:ASP
57:VAL
61:GLY
65:GLY
69:VAL
73:SER
77:SER
81:ASN
85:VAL
89:ALA
93:GLU
Auth_seq_ID