



# wwPDB NMR Structure Validation Summary Report ⓘ

Jun 6, 2023 – 02:21 pm BST

PDB ID : 5NOC  
BMRB ID : 34122  
Title : Solution NMR Structure of the C-terminal domain of ParB (Spo0J)  
Authors : Higman, V.A.; Fisher, G.L.M.; Dillingham, M.S.; Crump, M.P.  
Deposited on : 2017-04-11

This is a wwPDB NMR Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 ⓘ 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 ⓘ](#)) 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

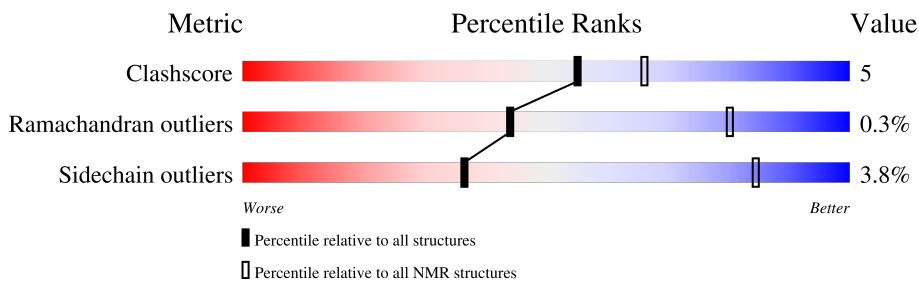
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 45%.

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  $\geq 3$ , 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	69	68% 9% . 22%
1	B	69	70% 9% 22%

## 2 Ensemble composition and analysis i

This entry contains 20 models. Model 13 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:230-A:282, B:229-B:282 (107)	0.65	13

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1838 atoms, of which 928 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Stage 0 sporulation protein J.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
1	A	54	919	287	464	78	90	0
1	B	54	919	287	464	78	90	0

There are 6 discrepancies between the modelled and reference sequences:

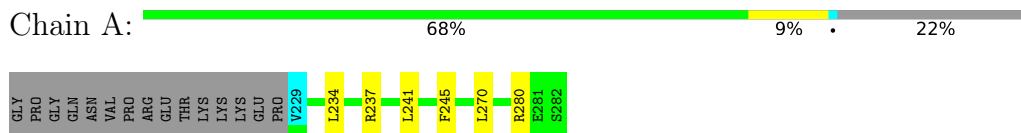
Chain	Residue	Modelled	Actual	Comment	Reference
A	214	GLY	-	expression tag	UNP P26497
A	215	PRO	-	expression tag	UNP P26497
A	216	GLY	-	expression tag	UNP P26497
B	214	GLY	-	expression tag	UNP P26497
B	215	PRO	-	expression tag	UNP P26497
B	216	GLY	-	expression tag	UNP P26497

## 4 Residue-property plots [i](#)

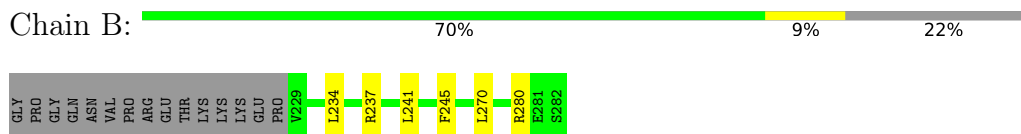
### 4.1 Average score per residue in the NMR ensemble

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: Stage 0 sporulation protein J



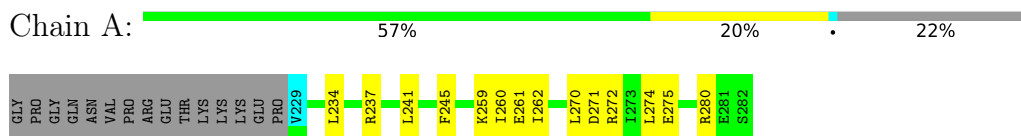
- Molecule 1: Stage 0 sporulation protein J



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

The representative model is number 13. Colouring as in section 4.1 above.

- Molecule 1: Stage 0 sporulation protein J



- Molecule 1: Stage 0 sporulation protein J



## 5 Refinement protocol and experimental data overview

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

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	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	1
Total number of shifts	911
Number of shifts mapped to atoms	731
Number of unparsed shifts	0
Number of shifts with mapping errors	180
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	45%

## 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	448	455	455	6±3
1	B	455	464	463	6±3
All	All	18060	18380	18360	192

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:234:LEU:HD11	1:B:270:LEU:HD21	0.63	1.69	5	9
1:A:270:LEU:HD21	1:B:234:LEU:HG	0.63	1.71	8	3
1:A:237:ARG:O	1:A:241:LEU:HG	0.61	1.95	17	18
1:B:237:ARG:O	1:B:241:LEU:HG	0.61	1.95	17	18
1:A:270:LEU:HD21	1:B:234:LEU:HD11	0.61	1.70	2	9

### 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	52/69 (75%)	48±1 (93±2%)	3±1 (7±2%)	0±0 (0±1%)	44	80
1	B	52/69 (75%)	48±1 (93±2%)	3±1 (7±2%)	0±0 (0±1%)	44	80
All	All	2080/2760 (75%)	1937 (93%)	137 (7%)	6 (0%)	44	80

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	255	LYS	2
1	B	255	LYS	2
1	A	279	GLU	1
1	B	279	GLU	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	50/64 (78%)	48±1 (96±2%)	2±1 (4±2%)	36	84
1	B	51/64 (80%)	49±1 (96±2%)	2±1 (4±2%)	37	85
All	All	2020/2560 (79%)	1944 (96%)	76 (4%)	36	84

5 of 14 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	280	ARG	15
1	B	280	ARG	15
1	A	245	PHE	14
1	B	245	PHE	14
1	A	272	ARG	4

### 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 45% for the well-defined parts and 46% for the entire structure.

### 7.1 Chemical shift list 1

File name: working\_cs.cif

Chemical shift list name: *assigned\_chem\_shift\_list*

#### 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	911
Number of shifts mapped to atoms	731
Number of unparsed shifts	0
Number of shifts with mapping errors	180
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 180) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	215	PRO	CG	27.092	0.055	1
1	A	215	PRO	CD	49.673	0.036	1
1	A	215	PRO	CA	63.529	0.120	1
1	A	215	PRO	C	177.449	0.000	1
1	A	215	PRO	CB	32.36	0.147	1
1	A	215	PRO	HA	4.47	0.007	1
1	A	215	PRO	HB2	1.984	0.000	2
1	A	215	PRO	HB3	2.301	0.002	2
1	A	215	PRO	HG2	2.029	0.000	1
1	A	215	PRO	HG3	2.029	0.000	1
1	A	215	PRO	HD2	3.57	0.000	2
1	A	215	PRO	HD3	3.595	0.000	2
1	A	216	GLY	C	174.074	0.000	1
1	A	216	GLY	CA	45.349	0.012	1

*Continued on next page...*

*Continued from previous page...*

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	216	GLY	H	8.567	0.001	1
1	A	216	GLY	HA2	3.954	0.000	1
1	A	216	GLY	HA3	3.954	0.000	1
1	A	216	GLY	N	109.785	0.008	1
1	A	217	GLN	CA	55.885	0.059	1
1	A	217	GLN	C	175.705	0.000	1
1	A	217	GLN	CB	29.742	0.106	1
1	A	217	GLN	CG	33.947	0.135	1
1	A	217	GLN	CD	180.587	0.015	1
1	A	217	GLN	H	8.115	0.001	1
1	A	217	GLN	HE21	6.818	0.001	1
1	A	217	GLN	HE22	7.482	0.001	1
1	A	217	GLN	HA	4.358	0.003	1
1	A	217	GLN	HG2	2.328	0.003	2
1	A	217	GLN	HG3	2.328	0.003	2
1	A	217	GLN	HB2	1.985	0.002	2
1	A	217	GLN	HB3	2.091	0.002	2
1	A	217	GLN	N	119.371	0.005	1
1	A	217	GLN	NE2	111.992	0.215	1
1	A	218	ASN	C	174.656	0.000	1
1	A	218	ASN	CB	38.882	0.027	1
1	A	218	ASN	CA	53.355	0.074	1
1	A	218	ASN	CG	177.134	0.007	1
1	A	218	ASN	HD21	6.871	0.002	1
1	A	218	ASN	HD22	7.527	0.022	1
1	A	218	ASN	H	8.498	0.001	1
1	A	218	ASN	HA	4.707	0.004	1
1	A	218	ASN	HB2	2.726	0.006	2
1	A	218	ASN	HB3	2.796	0.007	2
1	A	218	ASN	ND2	112.955	0.145	1
1	A	218	ASN	N	120.188	0.010	1
1	A	219	VAL	CB	32.61	0.059	1
1	A	219	VAL	CA	59.87	0.139	1
1	A	219	VAL	CG1	20.391	0.000	1
1	A	219	VAL	CG2	20.391	0.000	1
1	A	219	VAL	H	7.987	0.001	1
1	A	219	VAL	HA	4.41	0.003	1
1	A	219	VAL	HB	2.065	0.001	1
1	A	219	VAL	HG11	0.911	0.011	2
1	A	219	VAL	HG12	0.911	0.011	2
1	A	219	VAL	HG13	0.911	0.011	2

*Continued on next page...*

*Continued from previous page...*

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	219	VAL	HG21	0.942	0.012	2
1	A	219	VAL	HG22	0.942	0.012	2
1	A	219	VAL	HG23	0.942	0.012	2
1	A	219	VAL	N	121.592	0.016	1
1	A	220	PRO	CG	24.858	0.000	1
1	A	220	PRO	CD	50.203	0.000	1
1	A	220	PRO	C	176.051	0.000	1
1	A	220	PRO	CA	62.78	0.015	1
1	A	220	PRO	CB	34.487	0.007	1
1	A	220	PRO	HA	4.712	0.000	1
1	A	220	PRO	HD2	3.545	0.000	1
1	A	220	PRO	HD3	3.545	0.000	1
1	A	220	PRO	HG2	1.859	0.000	2
1	A	220	PRO	HB2	2.106	0.000	2
1	A	220	PRO	HB3	2.379	0.000	2
1	A	220	PRO	HG3	1.928	0.000	2
1	A	221	ARG	CB	30.93	0.056	1
1	A	221	ARG	CA	56.468	0.145	1
1	A	221	ARG	CG	27.137	0.000	1
1	A	221	ARG	CD	43.434	0.000	1
1	A	221	ARG	C	176.357	0.000	1
1	A	221	ARG	H	8.536	0.002	1
1	A	221	ARG	HD2	3.209	0.000	1
1	A	221	ARG	HD3	3.209	0.000	1
1	A	221	ARG	HA	4.303	0.002	1
1	A	221	ARG	HB2	1.802	0.000	1
1	A	221	ARG	HB3	1.802	0.000	1
1	A	221	ARG	HG2	1.673	0.000	1
1	A	221	ARG	HG3	1.673	0.000	1
1	A	221	ARG	N	121.447	0.022	1
1	A	222	GLU	CB	30.556	0.064	1
1	A	222	GLU	CG	36.259	0.000	1
1	A	222	GLU	C	176.536	0.000	1
1	A	222	GLU	CA	56.497	0.055	1
1	A	222	GLU	H	8.444	0.001	1
1	A	222	GLU	HB2	1.95	0.000	2
1	A	222	GLU	HB3	2.037	0.000	2
1	A	222	GLU	HG2	2.242	0.000	1
1	A	222	GLU	HG3	2.242	0.000	1
1	A	222	GLU	HA	4.364	0.000	1
1	A	222	GLU	N	122.061	0.005	1

*Continued on next page...*

*Continued from previous page...*

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	223	THR	C	174.443	0.000	1
1	A	223	THR	CB	69.854	0.062	1
1	A	223	THR	CA	61.995	0.055	1
1	A	223	THR	CG2	21.818	0.000	1
1	A	223	THR	H	8.232	0.001	1
1	A	223	THR	HB	4.19	0.000	1
1	A	223	THR	HG23	1.193	0.000	1
1	A	223	THR	HG22	1.193	0.000	1
1	A	223	THR	HG21	1.193	0.000	1
1	A	223	THR	HA	4.316	0.000	1
1	A	223	THR	N	116.334	0.007	1
1	A	224	LYS	CE	42.246	0.000	1
1	A	224	LYS	CG	24.662	0.000	1
1	A	224	LYS	CD	28.974	0.000	1
1	A	224	LYS	C	176.303	0.000	1
1	A	224	LYS	CA	56.29	0.055	1
1	A	224	LYS	CB	33.124	0.043	1
1	A	224	LYS	H	8.284	0.001	1
1	A	224	LYS	HA	4.331	0.000	1
1	A	224	LYS	HG2	1.426	0.000	1
1	A	224	LYS	HG3	1.426	0.000	1
1	A	224	LYS	HE2	2.997	0.000	1
1	A	224	LYS	HE3	2.997	0.000	1
1	A	224	LYS	HB2	1.748	0.000	2
1	A	224	LYS	HB3	1.815	0.000	2
1	A	224	LYS	HD2	1.685	0.000	1
1	A	224	LYS	HD3	1.685	0.000	1
1	A	224	LYS	N	124.115	0.021	1
1	A	225	LYS	CB	33.148	0.000	1
1	A	225	LYS	CG	24.709	0.000	1
1	A	225	LYS	CD	28.967	0.000	1
1	A	225	LYS	CE	42.265	0.000	1
1	A	225	LYS	C	176.276	0.000	1
1	A	225	LYS	CA	56.297	0.057	1
1	A	225	LYS	HE3	2.999	0.000	1
1	A	225	LYS	HG2	1.427	0.000	1
1	A	225	LYS	HG3	1.427	0.000	1
1	A	225	LYS	HD2	1.674	0.000	1
1	A	225	LYS	H	8.3	0.001	1
1	A	225	LYS	HD3	1.674	0.000	1
1	A	225	LYS	HB2	1.738	0.000	2

*Continued on next page...*

*Continued from previous page...*

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	225	LYS	HB3	1.806	0.000	2
1	A	225	LYS	HA	4.292	0.000	1
1	A	225	LYS	HE2	2.999	0.000	1
1	A	225	LYS	N	123.31	0.013	1
1	A	226	LYS	CB	33.183	0.015	1
1	A	226	LYS	CG	24.708	0.000	1
1	A	226	LYS	CD	29.029	0.000	1
1	A	226	LYS	CE	42.264	0.000	1
1	A	226	LYS	C	176.19	0.000	1
1	A	226	LYS	CA	56.184	0.072	1
1	A	226	LYS	H	8.353	0.005	1
1	A	226	LYS	HA	4.311	0.000	1
1	A	226	LYS	HE2	2.999	0.000	1
1	A	226	LYS	HE3	2.999	0.000	1
1	A	226	LYS	HG2	1.427	0.000	1
1	A	226	LYS	HG3	1.427	0.000	1
1	A	226	LYS	HB2	1.723	0.000	2
1	A	226	LYS	HD2	1.723	0.000	1
1	A	226	LYS	HD3	1.723	0.000	1
1	A	226	LYS	HB3	1.796	0.000	2
1	A	226	LYS	N	123.509	0.033	1
1	A	227	GLU	CB	29.949	0.000	1
1	A	227	GLU	CA	54.447	0.065	1
1	A	227	GLU	CG	36.128	0.000	1
1	A	227	GLU	H	8.368	0.002	1
1	A	227	GLU	HB2	1.879	0.000	2
1	A	227	GLU	HB3	2.032	0.000	2
1	A	227	GLU	HG2	2.287	0.000	2
1	A	227	GLU	HG3	2.283	0.000	2
1	A	227	GLU	HA	4.582	0.000	1
1	A	227	GLU	N	124.084	0.018	1
1	A	228	PRO	CA	62.868	0.007	1
1	A	228	PRO	CB	32.169	0.009	1
1	A	228	PRO	C	176.903	0.000	1
1	A	228	PRO	CG	27.493	0.000	1
1	A	228	PRO	CD	50.694	0.000	1
1	A	228	PRO	HA	4.453	0.000	1
1	A	228	PRO	HD2	3.697	0.000	2
1	A	228	PRO	HD3	3.813	0.000	2
1	A	228	PRO	HG2	2.034	0.000	1
1	A	228	PRO	HG3	2.034	0.000	1

*Continued on next page...*

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	228	PRO	HB2	1.858	0.000	2
1	A	228	PRO	HB3	2.266	0.000	2

### 7.1.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	68	$-0.56 \pm 0.13$	Should be checked
$^{13}\text{C}_\beta$	65	$0.09 \pm 0.09$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}'$	64	$-0.11 \pm 0.13$	None needed ( $< 0.5$ ppm)
$^{15}\text{N}$	64	$-0.04 \pm 0.26$	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 45%, i.e. 716 atoms were assigned a chemical shift out of a possible 1575. 0 out of 17 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	262/539 (49%)	106/218 (49%)	104/214 (49%)	52/107 (49%)
Sidechain	416/940 (44%)	278/597 (47%)	129/295 (44%)	9/48 (19%)
Aromatic	38/96 (40%)	19/46 (41%)	19/50 (38%)	0/0 (—%)
Overall	716/1575 (45%)	403/861 (47%)	252/559 (45%)	61/155 (39%)

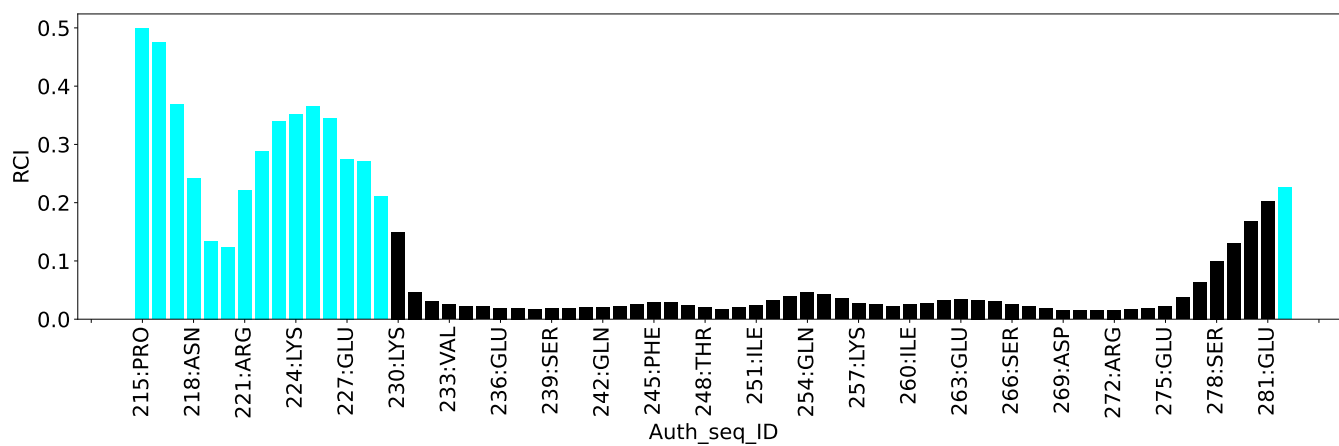
### 7.1.4 Statistically unusual chemical shifts [i](#)

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:





## 8 NMR restraints analysis

### 8.1 Conformationally restricting restraints

The following table provides the summary of experimentally observed NMR restraints in different categories. Restraints are classified into different categories based on the sequence separation of the atoms involved.

Description	Value
Total distance restraints	4044
Intra-residue ( $ i-j =0$ )	1080
Sequential ( $ i-j =1$ )	790
Medium range ( $ i-j >1$ and $ i-j <5$ )	962
Long range ( $ i-j \geq 5$ )	724
Inter-chain	438
Hydrogen bond restraints	50
Disulfide bond restraints	0
Total dihedral-angle restraints	160
Number of unmapped restraints	0
Number of restraints per residue	30.5
Number of long range restraints per residue <sup>1</sup>	5.3

<sup>1</sup>Long range hydrogen bonds and disulfide bonds are counted as long range restraints while calculating the number of long range restraints per residue

### 8.2 Residual restraint violations

This section provides the overview of the restraint violations analysis. The violations are binned as small, medium and large violations based on its absolute value. Average number of violations per model is calculated by dividing the total number of violations in each bin by the size of the ensemble.

#### 8.2.1 Average number of distance violations per model

Distance violations less than 0.1 Å are not included in the calculation.

Bins (Å)	Average number of violations per model	Max (Å)
0.1-0.2 (Small)	30.2	0.2
0.2-0.5 (Medium)	45.6	0.5
>0.5 (Large)	52.7	2.46

### 8.2.2 Average number of dihedral-angle violations per model [i](#)

Dihedral-angle violations less than 1° are not included in the calculation.

Bins (°)	Average number of violations per model	Max (°)
1.0-10.0 (Small)	1.4	2.8
10.0-20.0 (Medium)	None	None
>20.0 (Large)	None	None

## 9 Distance violation analysis [i](#)

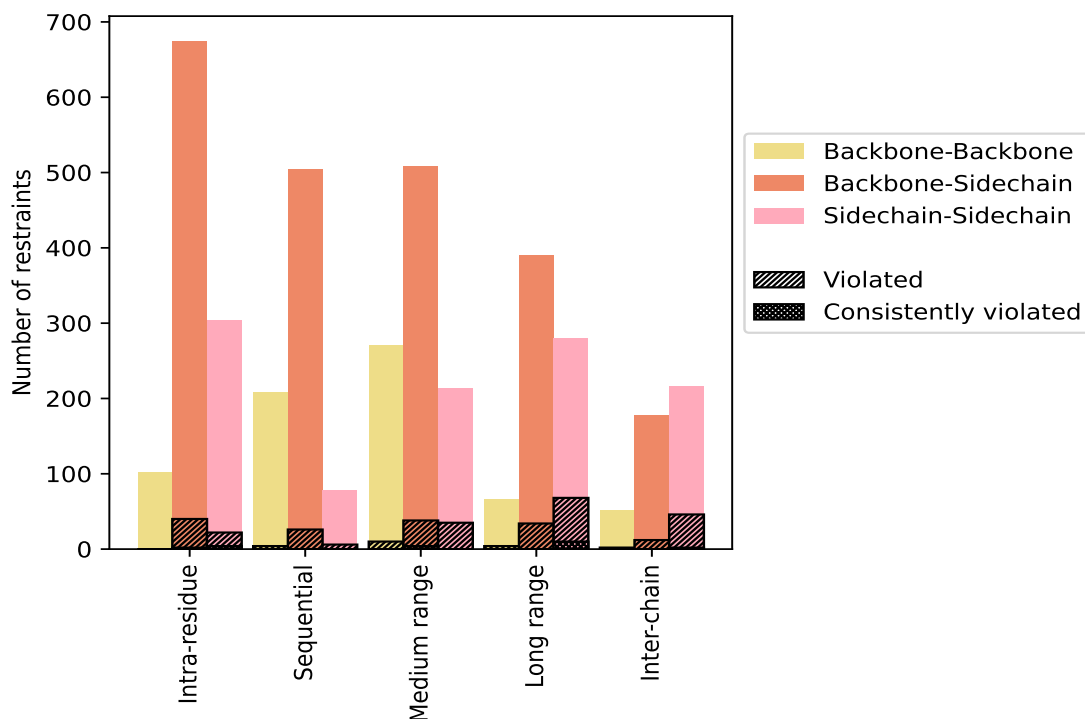
### 9.1 Summary of distance violations [i](#)

The following table shows the summary of distance violations in different restraint categories based on the sequence separation of the atoms involved. Each category is further sub-divided into three sub-categories based on the atoms involved. Violations less than 0.1 Å are not included in the statistics.

Restrains type	Count	% <sup>1</sup>	Violated <sup>3</sup>			Consistently Violated <sup>4</sup>		
			Count	% <sup>2</sup>	% <sup>1</sup>	Count	% <sup>2</sup>	% <sup>1</sup>
<b>Intra-residue (<math> i-j =0</math>)</b>	<b>1080</b>	<b>26.7</b>	<b>62</b>	<b>5.7</b>	<b>1.5</b>	<b>6</b>	<b>0.6</b>	<b>0.1</b>
Backbone-Backbone	102	2.5	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	674	16.7	40	5.9	1.0	2	0.3	0.0
Sidechain-Sidechain	304	7.5	22	7.2	0.5	4	1.3	0.1
<b>Sequential (<math> i-j =1</math>)</b>	<b>790</b>	<b>19.5</b>	<b>36</b>	<b>4.6</b>	<b>0.9</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
Backbone-Backbone	208	5.1	4	1.9	0.1	0	0.0	0.0
Backbone-Sidechain	504	12.5	26	5.2	0.6	0	0.0	0.0
Sidechain-Sidechain	78	1.9	6	7.7	0.1	0	0.0	0.0
<b>Medium range (<math> i-j &gt;1</math> &amp; <math> i-j &lt;5</math>)</b>	<b>962</b>	<b>23.8</b>	<b>75</b>	<b>7.8</b>	<b>1.9</b>	<b>4</b>	<b>0.4</b>	<b>0.1</b>
Backbone-Backbone	240	5.9	2	0.8	0.0	0	0.0	0.0
Backbone-Sidechain	508	12.6	38	7.5	0.9	4	0.8	0.1
Sidechain-Sidechain	214	5.3	35	16.4	0.9	0	0.0	0.0
<b>Long range (<math> i-j \geq 5</math>)</b>	<b>724</b>	<b>17.9</b>	<b>106</b>	<b>14.6</b>	<b>2.6</b>	<b>10</b>	<b>1.4</b>	<b>0.2</b>
Backbone-Backbone	54	1.3	4	7.4	0.1	0	0.0	0.0
Backbone-Sidechain	390	9.6	34	8.7	0.8	0	0.0	0.0
Sidechain-Sidechain	280	6.9	68	24.3	1.7	10	3.6	0.2
<b>Inter-chain</b>	<b>438</b>	<b>10.8</b>	<b>60</b>	<b>13.7</b>	<b>1.5</b>	<b>2</b>	<b>0.5</b>	<b>0.0</b>
Backbone-Backbone	44	1.1	2	4.5	0.0	0	0.0	0.0
Backbone-Sidechain	178	4.4	12	6.7	0.3	0	0.0	0.0
Sidechain-Sidechain	216	5.3	46	21.3	1.1	2	0.9	0.0
<b>Hydrogen bond</b>	<b>50</b>	<b>1.2</b>	<b>8</b>	<b>16.0</b>	<b>0.2</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
<b>Disulfide bond</b>	<b>0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
<b>Total</b>	<b>4044</b>	<b>100.0</b>	<b>347</b>	<b>8.6</b>	<b>8.6</b>	<b>22</b>	<b>0.5</b>	<b>0.5</b>
Backbone-Backbone	698	17.3	20	2.9	0.5	0	0.0	0.0
Backbone-Sidechain	2254	55.7	150	6.7	3.7	6	0.3	0.1
Sidechain-Sidechain	1092	27.0	177	16.2	4.4	16	1.5	0.4

<sup>1</sup> percentage calculated with respect to the total number of distance restraints, <sup>2</sup> percentage calculated with respect to the number of restraints in a particular restraint category, <sup>3</sup> violated in at least one model, <sup>4</sup> violated in all the models

### 9.1.1 Bar chart : Distribution of distance restraints and violations [i](#)



Violated and consistently violated restraints are shown using different hatch patterns in their respective categories. The hydrogen bonds and disulfid bonds are counted in their appropriate category on the x-axis

## 9.2 Distance violation statistics for each model [i](#)

The following table provides the distance violation statistics for each model in the ensemble. Violations less than 0.1 Å are not included in the statistics.

Model ID	Number of violations						Mean (Å)	Max (Å)	SD <sup>6</sup> (Å)	Median (Å)
	IR <sup>1</sup>	SQ <sup>2</sup>	MR <sup>3</sup>	LR <sup>4</sup>	IC <sup>5</sup>	Total				
1	22	10	27	46	22	127	0.62	2.46	0.51	0.47
2	26	12	20	30	24	112	0.62	2.42	0.58	0.46
3	28	8	36	46	23	141	0.53	2.29	0.42	0.4
4	28	8	24	31	20	111	0.6	2.24	0.48	0.51
5	28	14	23	40	23	128	0.5	2.22	0.4	0.38
6	26	14	36	32	26	134	0.55	2.27	0.46	0.42
7	30	14	28	36	21	129	0.59	2.41	0.54	0.38
8	26	12	24	56	32	150	0.58	2.21	0.45	0.43
9	27	12	22	42	26	129	0.56	2.21	0.43	0.44
10	23	14	30	38	26	131	0.5	1.98	0.36	0.45
11	32	14	40	42	22	150	0.56	2.09	0.5	0.36

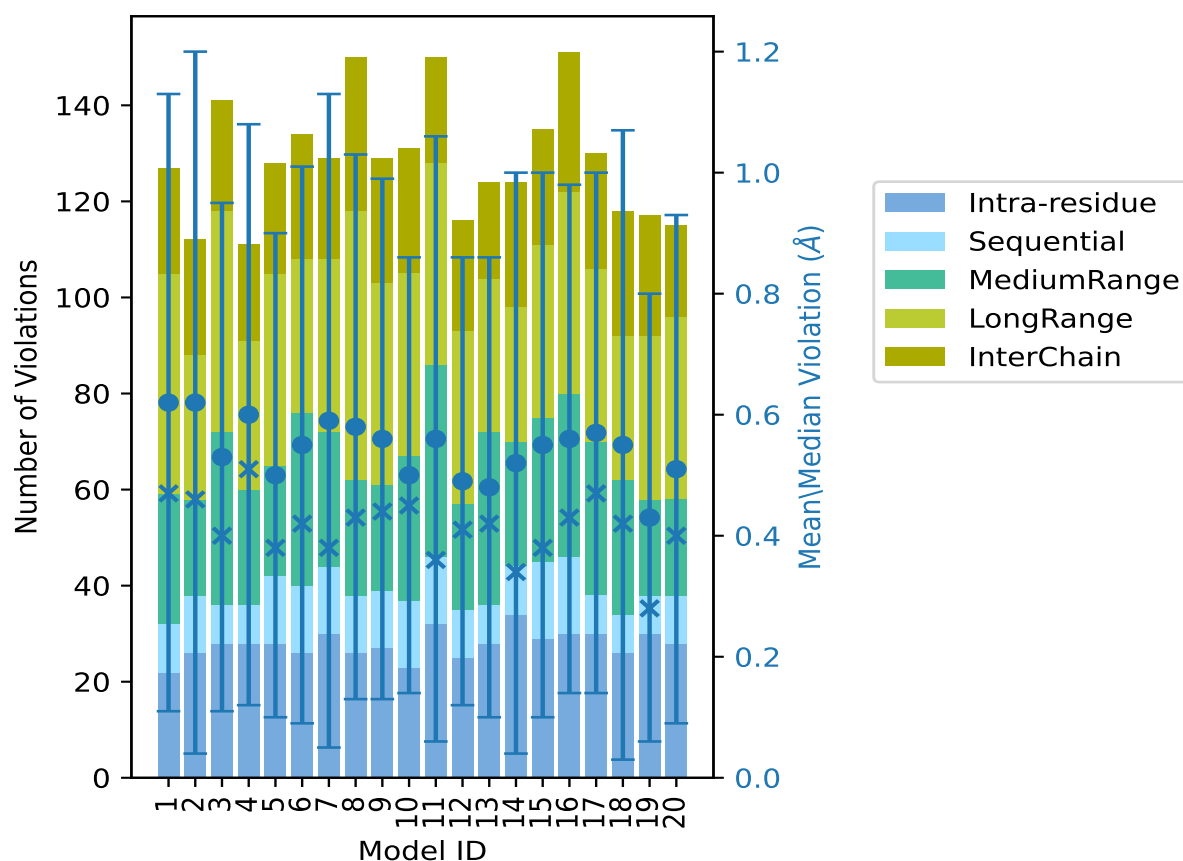
*Continued on next page...*

Continued from previous page...

Model ID	Number of violations					Total	Mean (Å)	Max (Å)	SD <sup>6</sup> (Å)	Median (Å)
	IR <sup>1</sup>	SQ <sup>2</sup>	MR <sup>3</sup>	LR <sup>4</sup>	IC <sup>5</sup>					
12	25	10	22	36	23	116	0.49	1.87	0.37	0.41
13	28	8	36	32	20	124	0.48	2.29	0.38	0.42
14	34	10	26	28	26	124	0.52	2.29	0.48	0.34
15	29	16	30	36	24	135	0.55	2.21	0.45	0.38
16	30	16	34	42	29	151	0.56	2.35	0.42	0.43
17	30	8	32	36	24	130	0.57	2.24	0.43	0.47
18	26	8	28	30	26	118	0.55	2.4	0.52	0.42
19	30	8	20	34	25	117	0.43	2.12	0.37	0.28
20	28	10	20	38	19	115	0.51	2.05	0.42	0.4

<sup>1</sup>Intra-residue restraints, <sup>2</sup>Sequential restraints, <sup>3</sup>Medium range restraints, <sup>4</sup>Long range restraints, <sup>5</sup>Inter-chain restraints, <sup>6</sup>Standard deviation

### 9.2.1 Bar graph : Distance Violation statistics for each model [\(i\)](#)



The mean(dot),median(x) and the standard deviation are shown in blue with respect to the y axis on the right

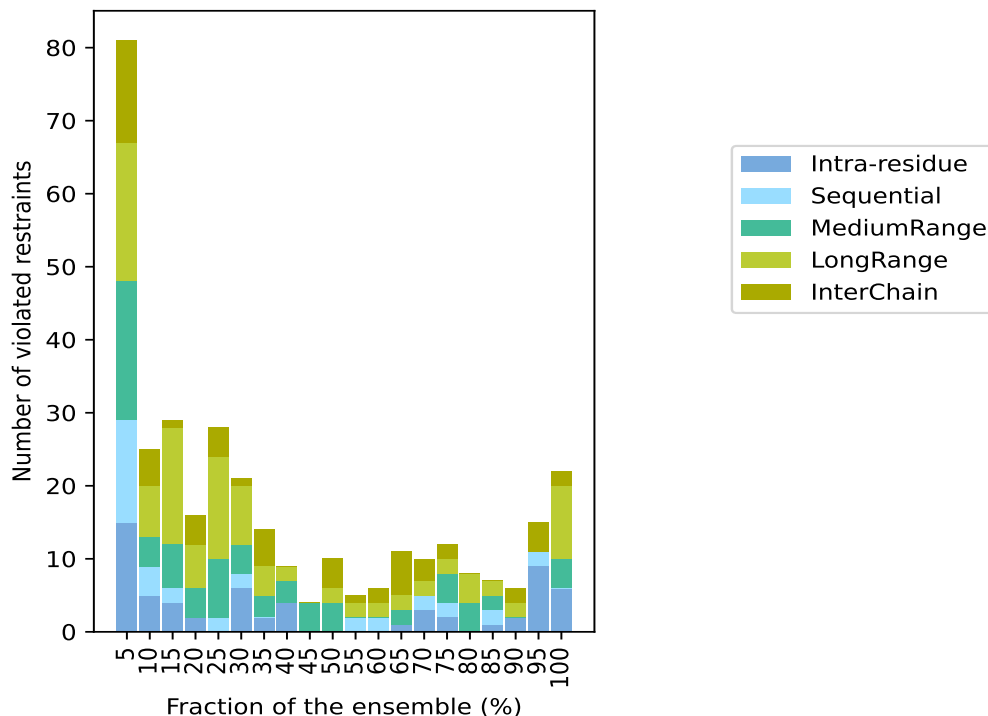
### 9.3 Distance violation statistics for the ensemble

Violation analysis may find that some restraints are violated in few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of the ensemble. In total, 3655(IR:1018, SQ:754, MR:887, LR:618, IC:378) restraints are not violated in the ensemble.

Number of violated restraints						Fraction of the ensemble	
IR <sup>1</sup>	SQ <sup>2</sup>	MR <sup>3</sup>	LR <sup>4</sup>	IC <sup>5</sup>	Total	Count <sup>6</sup>	%
15	14	19	19	14	81	1	5.0
5	4	4	7	5	25	2	10.0
4	2	6	16	1	29	3	15.0
2	0	4	6	4	16	4	20.0
0	2	8	14	4	28	5	25.0
6	2	4	8	1	21	6	30.0
2	0	3	4	5	14	7	35.0
4	0	3	2	0	9	8	40.0
0	0	4	0	0	4	9	45.0
0	0	4	2	4	10	10	50.0
0	2	0	2	1	5	11	55.0
0	2	0	2	2	6	12	60.0
1	0	2	2	6	11	13	65.0
3	2	0	2	3	10	14	70.0
2	2	4	2	2	12	15	75.0
0	0	4	4	0	8	16	80.0
1	2	2	2	0	7	17	85.0
2	0	0	2	2	6	18	90.0
9	2	0	0	4	15	19	95.0
6	0	4	10	2	22	20	100.0

<sup>1</sup>Intra-residue restraints, <sup>2</sup>Sequential restraints, <sup>3</sup>Medium range restraints, <sup>4</sup>Long range restraints, <sup>5</sup>Inter-chain restraints, <sup>6</sup> Number of models with violations

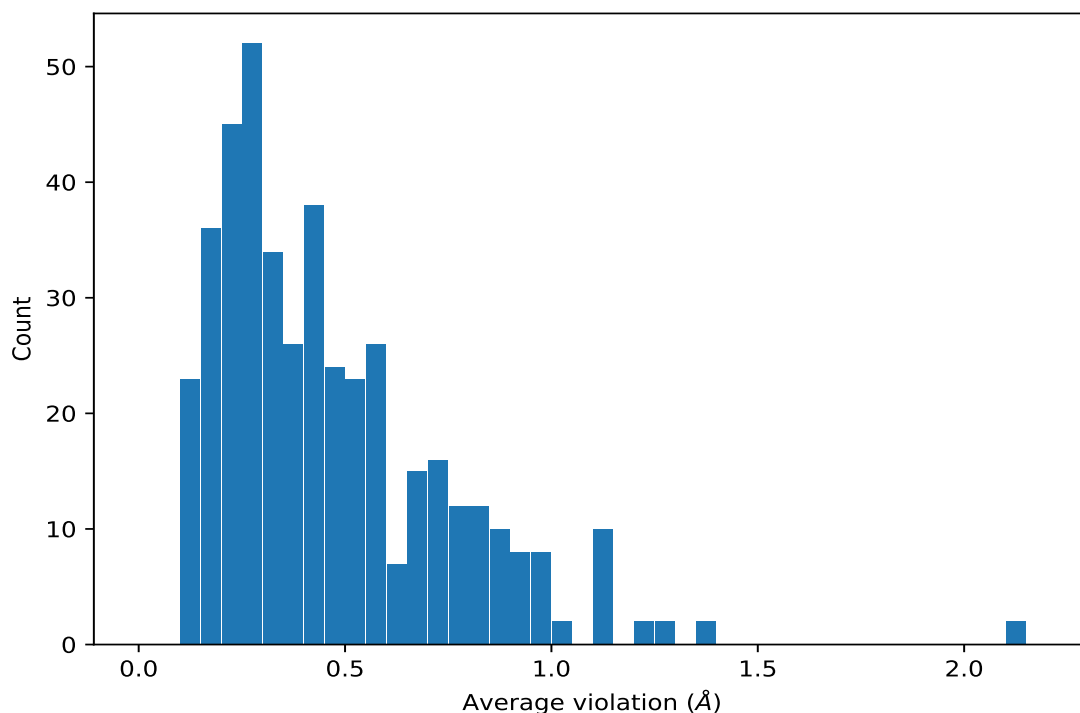
### 9.3.1 Bar graph : Distance violation statistics for the ensemble [i](#)



## 9.4 Most violated distance restraints in the ensemble [i](#)

### 9.4.1 Histogram : Distribution of mean distance violations [i](#)

The following histogram shows the distribution of the average value of the violation. The average is calculated for each restraint that is violated in more than one model over all the violated models in the ensemble



#### 9.4.2 Table: Most violated distance restraints [i](#)

The following table provides the mean and the standard deviation of the violations for the 10 worst performing restraints, sorted by number of violated models and the mean violation value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint. Rows with same key represent combinatorial or ambiguous restraints and are counted as a single restraint.

Key	Atom-1	Atom-2	Models <sup>1</sup>	Mean (Å)	SD <sup>1</sup> (Å)	Median (Å)
(1,1033)	1:A:273:ILE:HG21	1:A:262:ILE:HG12	20	2.13	0.15	2.17
(1,1034)	1:B:273:ILE:HG21	1:B:262:ILE:HG12	20	2.13	0.15	2.17
(1,310)	1:B:260:ILE:HD11	1:A:262:ILE:HG12	20	1.36	0.42	1.5
(1,309)	1:A:260:ILE:HD11	1:B:262:ILE:HG12	20	1.36	0.42	1.5
(1,886)	1:B:261:GLU:HG3	1:B:252:LYS:HG3	20	1.14	0.42	1.14
(1,885)	1:A:261:GLU:HG3	1:A:252:LYS:HG3	20	1.14	0.42	1.14
(1,985)	1:A:247:THR:HG21	1:A:272:ARG:HE	20	1.12	0.52	1.06
(1,986)	1:B:247:THR:HG21	1:B:272:ARG:HE	20	1.12	0.52	1.06
(1,377)	1:A:262:ILE:HG21	1:A:262:ILE:HG12	20	0.93	0.17	0.86
(1,378)	1:B:262:ILE:HG21	1:B:262:ILE:HG12	20	0.93	0.17	0.87

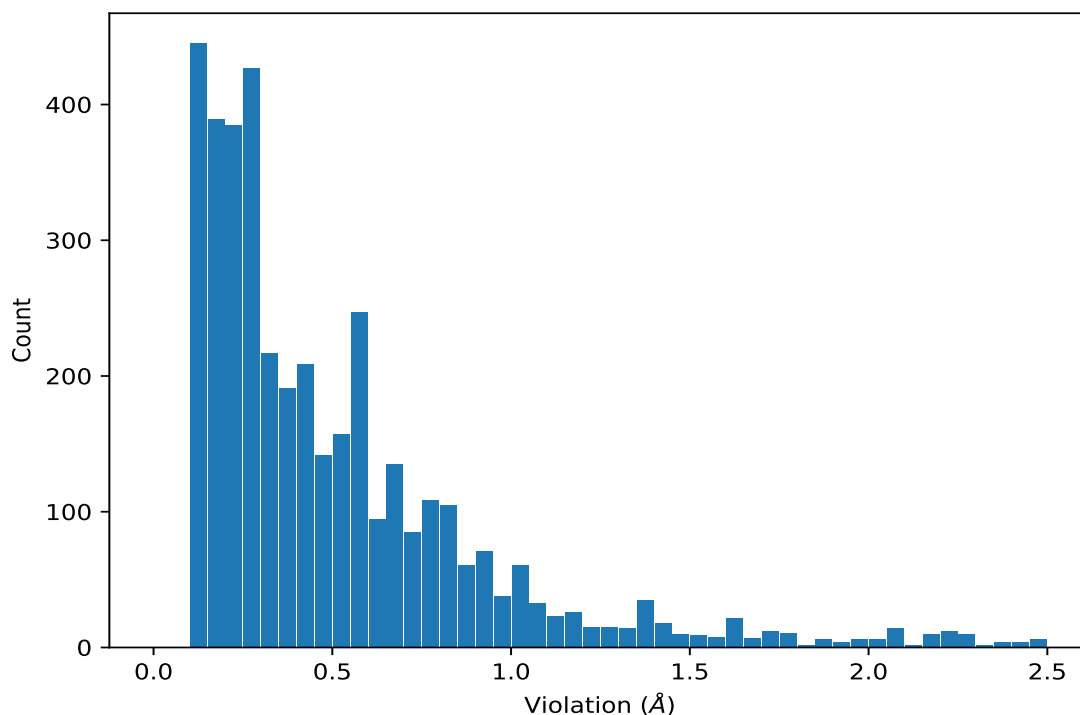
<sup>1</sup>Number of violated models, <sup>2</sup>Standard deviation



## 9.5 All violated distance restraints [i](#)

### 9.5.1 Histogram : Distribution of distance violations [i](#)

The following histogram shows the distribution of the absolute value of the violation for all violated restraints in the ensemble.



### 9.5.2 Table : All distance violations [i](#)

The following table provides the 10 worst performing restraints, sorted by the violation value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint. Rows with same key represent combinatorial or ambiguous restraints and are counted as a single restraint.

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,1272)	1:B:253:ARG:HD2	1:B:234:LEU:HD12	1	2.46
(1,1272)	1:B:234:LEU:HD11	1:B:253:ARG:HD2	1	2.46
(1,1272)	1:B:253:ARG:HD2	1:B:234:LEU:HD13	1	2.46
(1,1271)	1:A:253:ARG:HD2	1:A:234:LEU:HD12	1	2.46
(1,1271)	1:A:234:LEU:HD11	1:A:253:ARG:HD2	1	2.46
(1,1271)	1:A:253:ARG:HD2	1:A:234:LEU:HD13	1	2.46
(1,2682)	1:B:281:GLU:H	1:B:280:ARG:HD3	2	2.42
(1,2681)	1:A:281:GLU:H	1:A:280:ARG:HD3	2	2.42
(1,410)	1:B:269:ASP:HB2	1:B:272:ARG:HE	7	2.41
(1,409)	1:A:269:ASP:HB2	1:A:272:ARG:HE	7	2.41
(1,986)	1:B:247:THR:HG21	1:B:272:ARG:HE	18	2.4

*Continued on next page...*

*Continued from previous page...*

<b>Key</b>	<b>Atom-1</b>	<b>Atom-2</b>	<b>Model ID</b>	<b>Violation (Å)</b>
(1,985)	1:A:247:THR:HG21	1:A:272:ARG:HE	18	2.4
(1,1034)	1:B:273:ILE:HG21	1:B:262:ILE:HG12	16	2.35
(1,1033)	1:A:273:ILE:HG21	1:A:262:ILE:HG12	16	2.35

## 10 Dihedral-angle violation analysis [i](#)

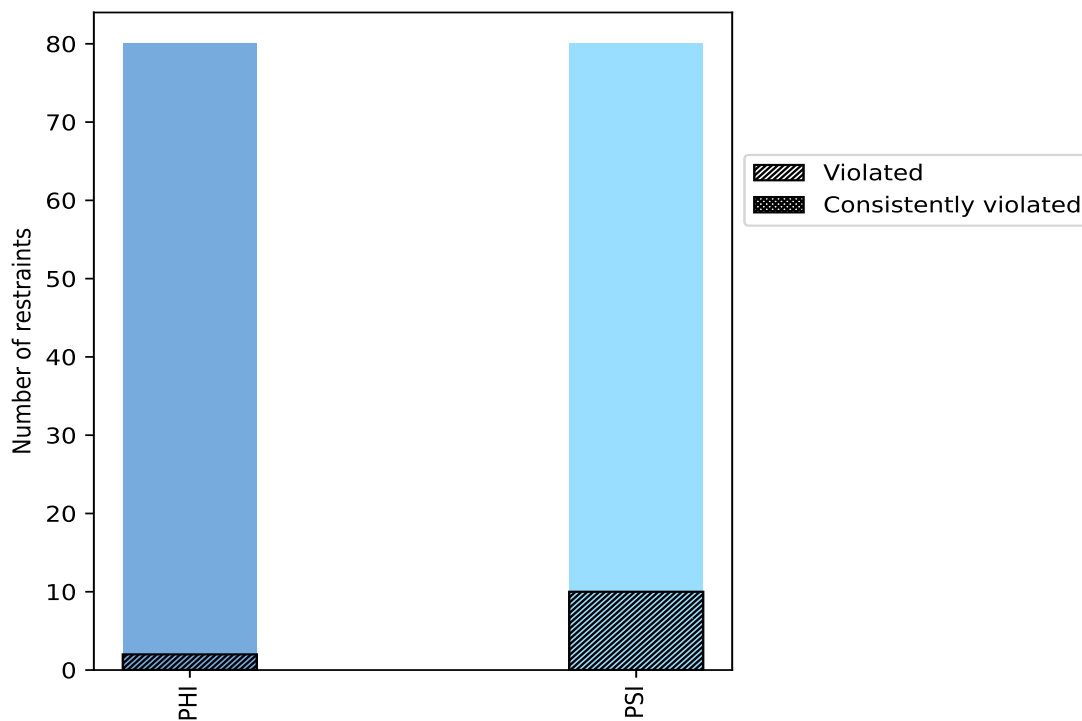
### 10.1 Summary of dihedral-angle violations [i](#)

The following table provides the summary of dihedral-angle violations in different dihedral-angle types. Violations less than 1° are not included in the calculation.

Angle type	Count	% <sup>1</sup>	Violated <sup>3</sup>			Consistently Violated <sup>4</sup>		
			Count	% <sup>2</sup>	% <sup>1</sup>	Count	% <sup>2</sup>	% <sup>1</sup>
PHI	80	50.0	2	2.5	1.2	0	0.0	0.0
PSI	80	50.0	10	12.5	6.2	0	0.0	0.0
Total	160	100.0	12	7.5	7.5	0	0.0	0.0

<sup>1</sup> percentage calculated with respect to total number of dihedral-angle restraints, <sup>2</sup> percentage calculated with respect to number of restraints in a particular dihedral-angle type, <sup>3</sup> violated in at least one model, <sup>4</sup> violated in all the models

#### 10.1.1 Bar chart : Distribution of dihedral-angles and violations [i](#)



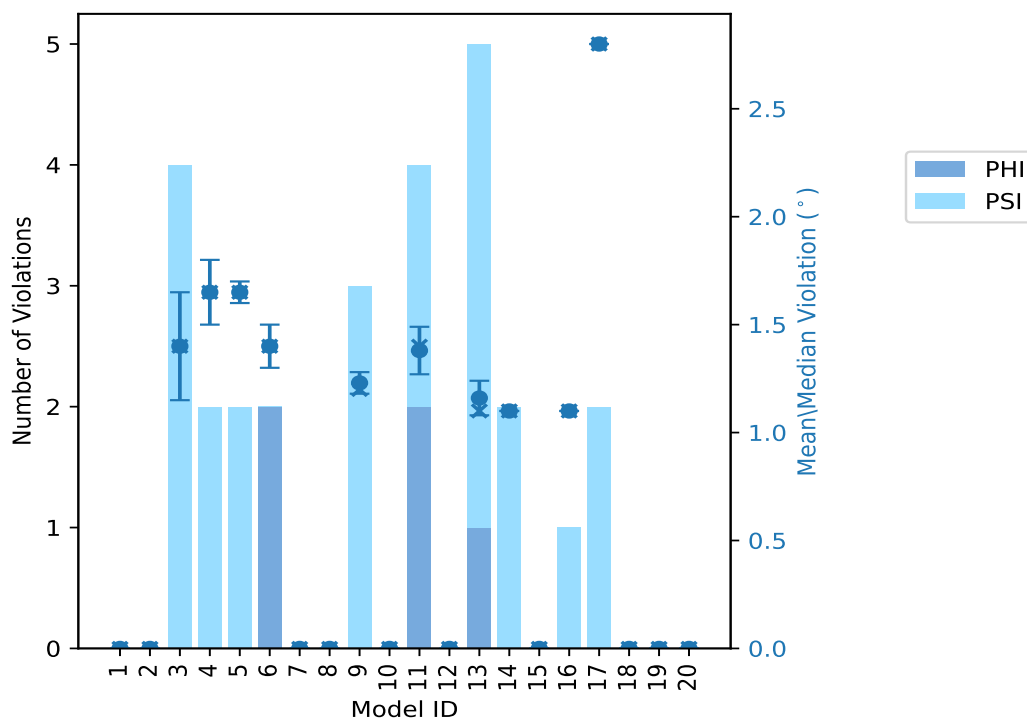
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

## 10.2 Dihedral-angle violation statistics for each model [i](#)

The following table provides the dihedral-angle violation statistics for each model in the ensemble. Violations less than 1° are not included in the statistics.

Model ID	Number of violations			Mean (°)	Max (°)	SD (°)	Median (°)
	PHI	PSI	Total				
1	0	0	0	0.0	0.0	0.0	0.0
2	0	0	0	0.0	0.0	0.0	0.0
3	0	4	4	1.4	1.7	0.25	1.4
4	0	2	2	1.65	1.8	0.15	1.65
5	0	2	2	1.65	1.7	0.05	1.65
6	2	0	2	1.4	1.5	0.1	1.4
7	0	0	0	0.0	0.0	0.0	0.0
8	0	0	0	0.0	0.0	0.0	0.0
9	0	3	3	1.23	1.3	0.05	1.2
10	0	0	0	0.0	0.0	0.0	0.0
11	2	2	4	1.38	1.5	0.11	1.4
12	0	0	0	0.0	0.0	0.0	0.0
13	1	4	5	1.16	1.3	0.08	1.1
14	0	2	2	1.1	1.1	0.0	1.1
15	0	0	0	0.0	0.0	0.0	0.0
16	0	1	1	1.1	1.1	0.0	1.1
17	0	2	2	2.8	2.8	0.0	2.8
18	0	0	0	0.0	0.0	0.0	0.0
19	0	0	0	0.0	0.0	0.0	0.0
20	0	0	0	0.0	0.0	0.0	0.0

### 10.2.1 Bar graph : Dihedral violation statistics for each model [i](#)



The mean(dot),median(x) and the standard deviation are shown in blue with respect to the y axis on the right

### 10.3 Dihedral-angle violation statistics for the ensemble [i](#)

Violation analysis may find that some restraints are violated in very few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of ensemble.

Number of violated restraints			Fraction of the ensemble	
PHI	PSI	Total	Count <sup>1</sup>	%
0	5	5	1	5.0
1	1	2	2	10.0
1	2	3	3	15.0
0	1	1	4	20.0
0	1	1	5	25.0
0	0	0	6	30.0
0	0	0	7	35.0
0	0	0	8	40.0
0	0	0	9	45.0
0	0	0	10	50.0
0	0	0	11	55.0

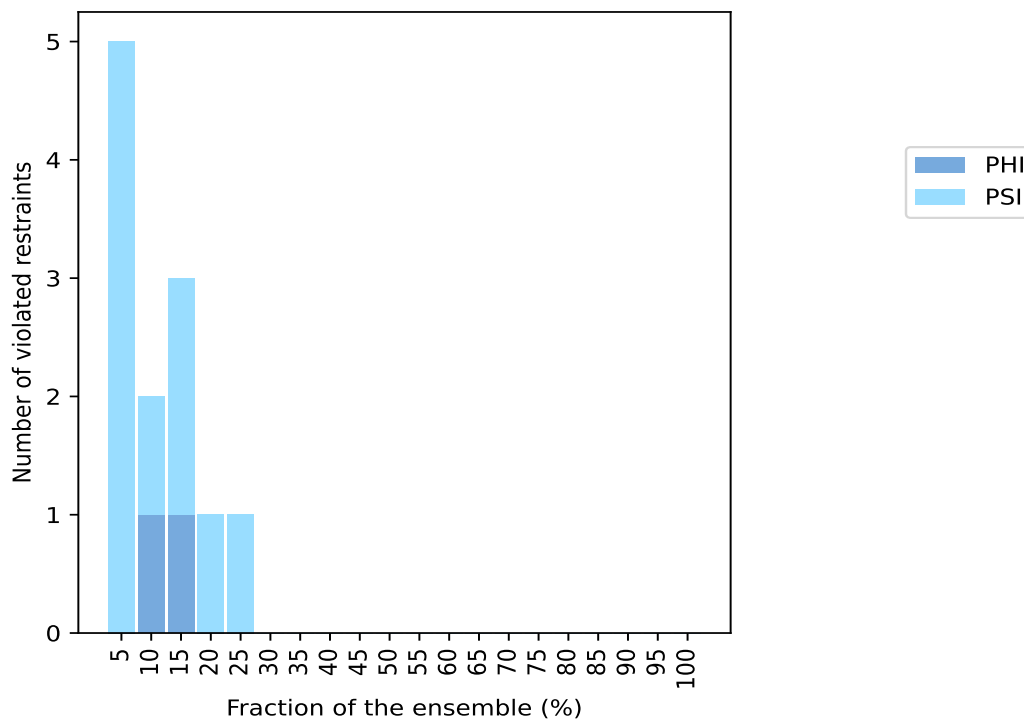
*Continued on next page...*

Continued from previous page...

Number of violated restraints			Fraction of the ensemble	
PHI	PSI	Total	Count <sup>1</sup>	%
0	0	0	12	60.0
0	0	0	13	65.0
0	0	0	14	70.0
0	0	0	15	75.0
0	0	0	16	80.0
0	0	0	17	85.0
0	0	0	18	90.0
0	0	0	19	95.0
0	0	0	20	100.0

<sup>1</sup> Number of models with violations

### 10.3.1 Bar graph : Dihedral-angle Violation statistics for the ensemble [i](#)

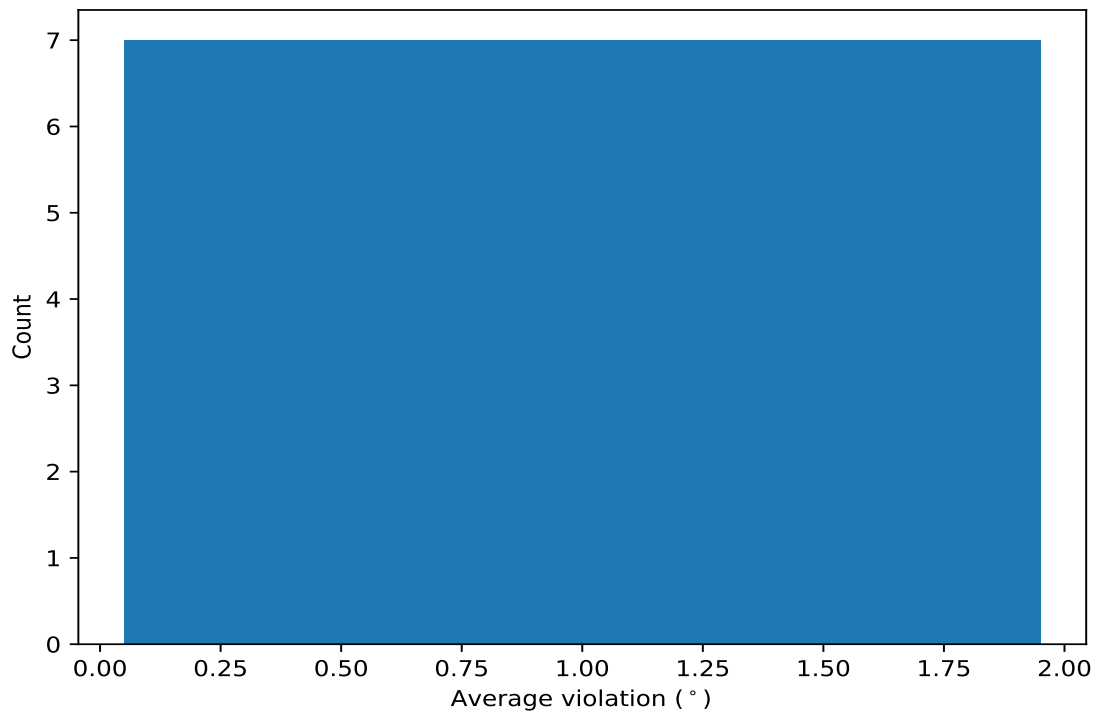


## 10.4 Most violated dihedral-angle restraints in the ensemble [i](#)

### 10.4.1 Histogram : Distribution of mean dihedral-angle violations [i](#)

The following histogram shows the distribution of the average value of the violation. The average is calculated for each restraint that is violated in more than one model over all the violated models

in the ensemble



#### 10.4.2 Table: Most violated dihedral-angle restraints [i](#)

The following table provides the mean and the standard deviation of the violation for each restraint sorted by number of violated models and the mean value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint.

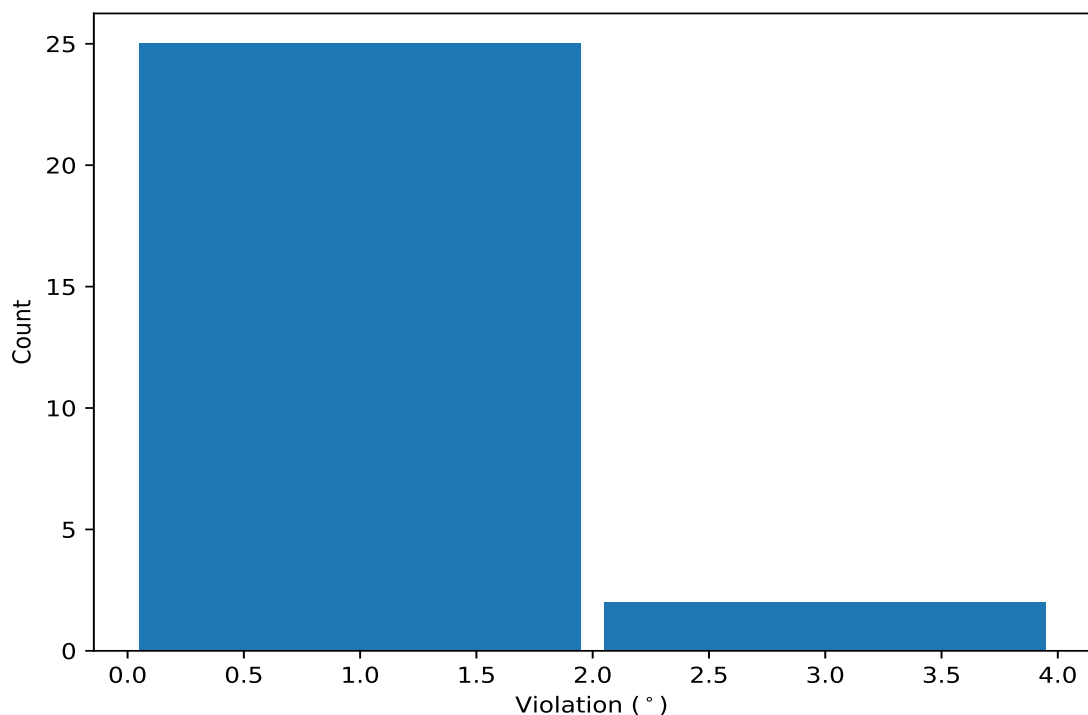
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models <sup>1</sup>	Mean	SD <sup>2</sup>	Median
(1,138)	1:B:264:PHE:N	1:B:264:PHE:CA	1:B:264:PHE:C	1:B:265:PHE:N	5	1.26	0.21	1.1
(1,58)	1:A:264:PHE:N	1:A:264:PHE:CA	1:A:264:PHE:C	1:A:265:PHE:N	4	1.3	0.23	1.2
(1,110)	1:B:245:PHE:N	1:B:245:PHE:CA	1:B:245:PHE:C	1:B:246:GLY:N	3	1.87	0.7	1.7
(1,30)	1:A:245:PHE:N	1:A:245:PHE:CA	1:A:245:PHE:C	1:A:246:GLY:N	3	1.83	0.71	1.6
(1,13)	1:A:236:GLU:C	1:A:237:ARG:N	1:A:237:ARG:CA	1:A:237:ARG:C	3	1.4	0.14	1.5
(1,114)	1:B:247:THR:N	1:B:247:THR:CA	1:B:247:THR:C	1:B:248:THR:N	2	1.55	0.25	1.55
(1,93)	1:B:236:GLU:C	1:B:237:ARG:N	1:B:237:ARG:CA	1:B:237:ARG:C	2	1.35	0.05	1.35

<sup>1</sup> Number of violated models, <sup>2</sup>Standard deviation, All angle values are in degree (°)

### 10.5 All violated dihedral-angle restraints [i](#)

#### 10.5.1 Histogram : Distribution of violations [i](#)

The following histogram shows the distribution of the absolute value of the violation for all violated restraints in the ensemble.



### 10.5.2 Table: All violated dihedral-angle restraints [i](#)

The following table provides the list of violations for the 10 worst performing restraints, sorted by the violation value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint.

Key	Atom-1	Atom-2	Atom-3	Atom-4	Model ID	Violation (°)
(1,30)	1:A:245:PHE:N	1:A:245:PHE:CA	1:A:245:PHE:C	1:A:246:GLY:N	17	2.8
(1,110)	1:B:245:PHE:N	1:B:245:PHE:CA	1:B:245:PHE:C	1:B:246:GLY:N	17	2.8
(1,114)	1:B:247:THR:N	1:B:247:THR:CA	1:B:247:THR:C	1:B:248:THR:N	4	1.8
(1,58)	1:A:264:PHE:N	1:A:264:PHE:CA	1:A:264:PHE:C	1:A:265:PHE:N	5	1.7
(1,110)	1:B:245:PHE:N	1:B:245:PHE:CA	1:B:245:PHE:C	1:B:246:GLY:N	3	1.7
(1,30)	1:A:245:PHE:N	1:A:245:PHE:CA	1:A:245:PHE:C	1:A:246:GLY:N	3	1.6
(1,138)	1:B:264:PHE:N	1:B:264:PHE:CA	1:B:264:PHE:C	1:B:265:PHE:N	5	1.6
(1,34)	1:A:247:THR:N	1:A:247:THR:CA	1:A:247:THR:C	1:A:248:THR:N	4	1.5
(1,13)	1:A:236:GLU:C	1:A:237:ARG:N	1:A:237:ARG:CA	1:A:237:ARG:C	6	1.5
(1,13)	1:A:236:GLU:C	1:A:237:ARG:N	1:A:237:ARG:CA	1:A:237:ARG:C	11	1.5