



# wwPDB NMR Structure Validation Summary Report ⓘ

Dec 25, 2024 – 06:04 PM EST

PDB ID : 7LOI  
BMRB ID : 30856  
Title : Model of the HIV-1 gp41 membrane-proximal external region, transmembrane domain and cytoplasmic tail  
Authors : Piai, A.; Fu, Q.; Sharp, A.K.; Bighi, B.; Brown, A.M.; Chou, J.J.  
Deposited on : 2021-02-10

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 : 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) : 2.40

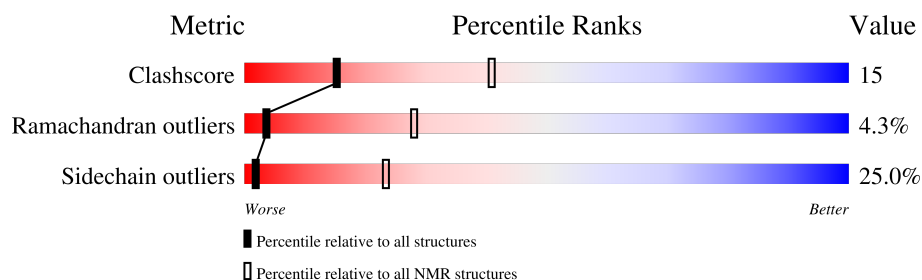
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 8%.

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  $\geq 3$ , 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	197	
1	B	197	
1	C	197	

## 2 Ensemble composition and analysis

This entry contains 15 models. Model 4 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:660-A:716, A:739-A:856, B:660-B:716, B:739-B:856, C:660-C:716, C:739-C:856 (525)	1.58	4

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 3 clusters and 1 single-model cluster was found.

Cluster number	Models
1	2, 3, 4, 6, 7, 9, 10, 11, 13
2	1, 5, 12
3	8, 15
Single-model clusters	14

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 8940 atoms, of which 4563 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Transmembrane protein gp41.

Mol	Chain	Residues	Atoms					Trace
1	A	175	Total	C	H	N	O	0
			2980	949	1521	267	243	
1	B	175	Total	C	H	N	O	0
			2980	949	1521	267	243	
1	C	175	Total	C	H	N	O	0
			2980	949	1521	267	243	

There are 27 discrepancies between the modelled and reference sequences:

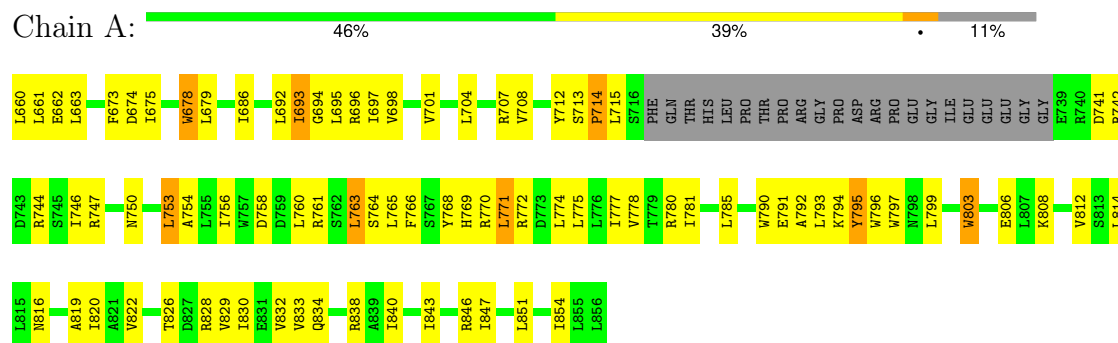
Chain	Residue	Modelled	Actual	Comment	Reference
A	674	ASP	ASN	conflict	UNP A0A386YSI0
A	683	ARG	LYS	conflict	UNP A0A386YSI0
A	687	ILE	MET	conflict	UNP A0A386YSI0
A	691	SER	GLY	conflict	UNP A0A386YSI0
A	693	ILE	VAL	conflict	UNP A0A386YSI0
A	704	LEU	ILE	conflict	UNP A0A386YSI0
A	764	SER	CYS	engineered mutation	UNP A0A386YSI0
A	823	GLY	ALA	engineered mutation	UNP A0A386YSI0
A	837	SER	CYS	engineered mutation	UNP A0A386YSI0
B	674	ASP	ASN	conflict	UNP A0A386YSI0
B	683	ARG	LYS	conflict	UNP A0A386YSI0
B	687	ILE	MET	conflict	UNP A0A386YSI0
B	691	SER	GLY	conflict	UNP A0A386YSI0
B	693	ILE	VAL	conflict	UNP A0A386YSI0
B	704	LEU	ILE	conflict	UNP A0A386YSI0
B	764	SER	CYS	engineered mutation	UNP A0A386YSI0
B	823	GLY	ALA	engineered mutation	UNP A0A386YSI0
B	837	SER	CYS	engineered mutation	UNP A0A386YSI0
C	674	ASP	ASN	conflict	UNP A0A386YSI0
C	683	ARG	LYS	conflict	UNP A0A386YSI0
C	687	ILE	MET	conflict	UNP A0A386YSI0
C	691	SER	GLY	conflict	UNP A0A386YSI0
C	693	ILE	VAL	conflict	UNP A0A386YSI0
C	704	LEU	ILE	conflict	UNP A0A386YSI0
C	764	SER	CYS	engineered mutation	UNP A0A386YSI0
C	823	GLY	ALA	engineered mutation	UNP A0A386YSI0
C	837	SER	CYS	engineered mutation	UNP A0A386YSI0

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

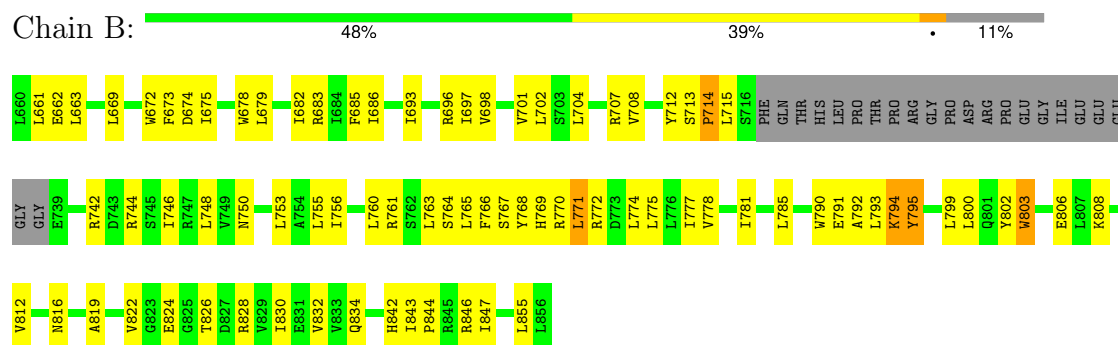
### 4.1 Average score per residue in the NMR ensemble

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: Transmembrane protein gp41



- Molecule 1: Transmembrane protein gp41



- Molecule 1: Transmembrane protein gp41



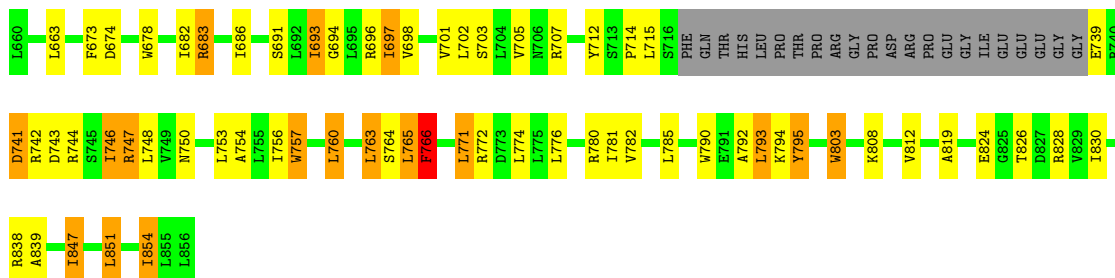


## 4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 4. Colouring as in section 4.1 above.

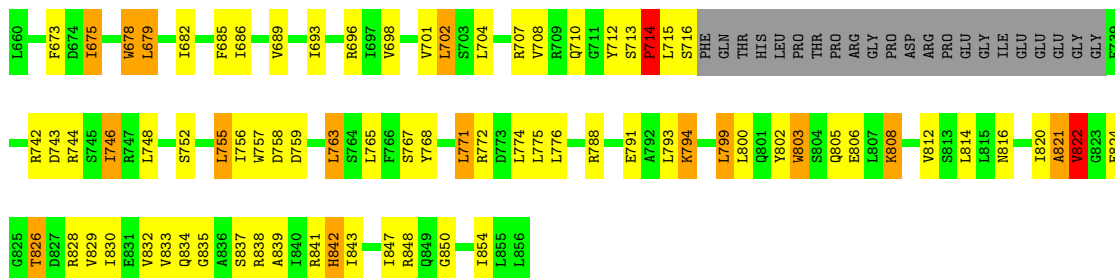
- Molecule 1: Transmembrane protein gp41

Chain A: 56% 24% 9% 11%



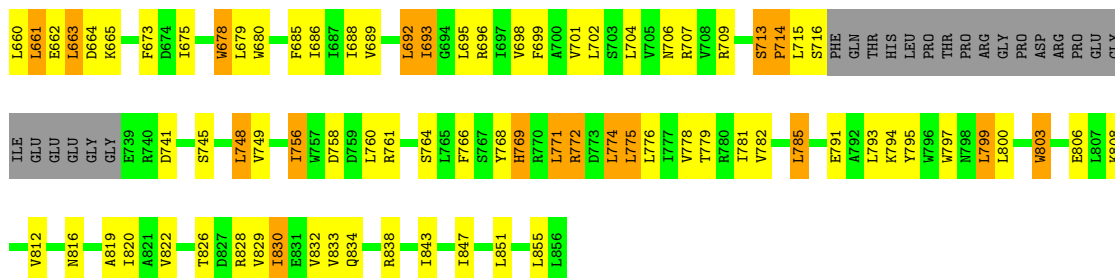
- Molecule 1: Transmembrane protein gp41

Chain B: 49% 31% 8% 11%



- Molecule 1: Transmembrane protein gp41

Chain C: 48% 31% 9% 11%



## 5 Refinement protocol and experimental data overview

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

Of the 150 calculated structures, 15 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 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	657
Number of shifts mapped to atoms	622
Number of unparsed shifts	0
Number of shifts with mapping errors	35
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	8%

## 6 Model quality

### 6.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the (average) root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	0.79±0.01	0±0/1490 ( 0.0± 0.0%)	0.97±0.01	0±1/2021 ( 0.0± 0.0%)
1	B	0.79±0.01	0±0/1490 ( 0.0± 0.0%)	0.96±0.01	0±1/2021 ( 0.0± 0.0%)
1	C	0.79±0.01	0±0/1490 ( 0.0± 0.0%)	0.96±0.01	0±0/2021 ( 0.0± 0.0%)
All	All	0.79	0/67050 ( 0.0%)	0.97	12/90945 ( 0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	Chirality	Planarity
1	B	0.0±0.0	0.1±0.2
All	All	0	1

There are no bond-length outliers.

5 of 10 unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	A	795	TYR	CB-CG-CD1	-7.20	116.68	121.00	4	1
1	B	795	TYR	CB-CG-CD1	-6.27	117.24	121.00	6	1
1	A	707	ARG	NE-CZ-NH2	5.66	123.13	120.30	1	2
1	A	685	PHE	CB-CG-CD2	-5.65	116.85	120.80	15	1
1	C	790	TRP	CD1-NE1-CE2	5.44	113.89	109.00	7	1

There are no chirality outliers.

All unique planar outliers are listed below.

Mol	Chain	Res	Type	Group	Models (Total)
1	B	770	ARG	Sidechain	1



## 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	1459	1521	1515	55±5
1	B	1459	1521	1515	52±8
1	C	1459	1521	1515	48±6
All	All	65655	68445	68175	1955

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 15.

5 of 1178 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:746:ILE:HG23	1:C:781:ILE:HD12	1.00	1.27	11	1
1:C:819:ALA:HB3	1:C:826:THR:HG21	0.97	1.35	14	1
1:B:819:ALA:HB1	1:B:826:THR:HG23	0.90	1.41	1	1
1:A:686:ILE:HD11	1:C:686:ILE:HG21	0.90	1.43	4	4
1:B:669:LEU:HD13	1:B:676:THR:HG21	0.89	1.43	8	1

## 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	171/197 (87%)	154±2 (90±1%)	10±2 (6±1%)	7±2 (4±1%)	4	31
1	B	171/197 (87%)	152±2 (89±1%)	10±2 (6±1%)	8±2 (5±1%)	3	24
1	C	171/197 (87%)	154±2 (90±1%)	10±2 (6±1%)	7±3 (4±2%)	4	29
All	All	7695/8865 (87%)	6905 (90%)	459 (6%)	331 (4%)	3	28

5 of 68 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	714	PRO	15
1	C	714	PRO	15
1	B	714	PRO	14
1	C	712	TYR	11
1	A	741	ASP	9

### 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	155/173 (90%)	116±6 (75±4%)	39±6 (25±4%)	2	23
1	B	155/173 (90%)	117±6 (76±4%)	38±6 (24±4%)	2	24
1	C	155/173 (90%)	115±5 (74±3%)	40±5 (26±3%)	2	22
All	All	6975/7785 (90%)	5231 (75%)	1744 (25%)	2	23

5 of 409 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	803	TRP	15
1	B	803	TRP	15
1	C	771	LEU	15
1	C	794	LYS	15
1	C	803	TRP	15

### 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 oligosaccharides in this entry.

## 6.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 6.7 Other polymers [i](#)

There are no such molecules in this entry.

## 6.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 8% for the well-defined parts and 8% for the entire structure.

### 7.1 Chemical shift list 1

File name: working\_cs.cif

Chemical shift list name: *HIV1\_gp41\_MPER-TMD-CT\_chemical\_shifts.tab*

#### 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	657
Number of shifts mapped to atoms	622
Number of unparsed shifts	0
Number of shifts with mapping errors	35
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 35) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	717	PHE	H	8.05	0.01	1
1	A	717	PHE	C	175.43	0.05	1
1	A	717	PHE	CA	57.49	0.05	1
1	A	717	PHE	N	120.86	0.03	1
1	A	718	GLN	H	8.13	0.01	1
1	A	718	GLN	C	175.94	0.05	1
1	A	718	GLN	CA	55.56	0.05	1
1	A	718	GLN	N	120.21	0.03	1
1	A	719	THR	H	8.03	0.01	1
1	A	719	THR	CA	61.81	0.05	1
1	A	719	THR	N	114.2	0.03	1
1	A	720	HIS	C	174.75	0.05	1
1	A	720	HIS	CA	55.7	0.05	1
1	A	721	LEU	H	8.02	0.01	1

*Continued on next page...*

*Continued from previous page...*

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	721	LEU	CA	52.47	0.05	1
1	A	721	LEU	N	123.97	0.03	1
1	A	722	PRO	C	176.74	0.05	1
1	A	722	PRO	CA	62.54	0.05	1
1	A	723	THR	H	8.15	0.01	1
1	A	723	THR	CA	59.45	0.05	1
1	A	723	THR	N	116.61	0.03	1
1	A	724	PRO	C	177.1	0.05	1
1	A	724	PRO	CA	62.91	0.05	1
1	A	725	ARG	H	8.45	0.01	1
1	A	725	ARG	C	177.14	0.05	1
1	A	725	ARG	CA	55.87	0.05	1
1	A	725	ARG	N	121.18	0.03	1
1	A	737	GLY	H	8.48	0.01	1
1	A	737	GLY	C	174.84	0.05	1
1	A	737	GLY	CA	45.06	0.05	1
1	A	737	GLY	N	110.12	0.03	1
1	A	738	GLY	H	8.32	0.01	1
1	A	738	GLY	C	174.61	0.05	1
1	A	738	GLY	CA	44.88	0.05	1
1	A	738	GLY	N	108.9	0.03	1

### 7.1.2 Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	171	$-0.04 \pm 0.13$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}_\beta$	0	—	None (insufficient data)
$^{13}\text{C}'$	158	$-0.28 \pm 0.08$	None needed ( $< 0.5$ ppm)
$^{15}\text{N}$	164	$0.90 \pm 0.26$	Should be applied

### 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 8%, i.e. 622 atoms were assigned a chemical shift out of a possible 7980. 0 out of 132 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

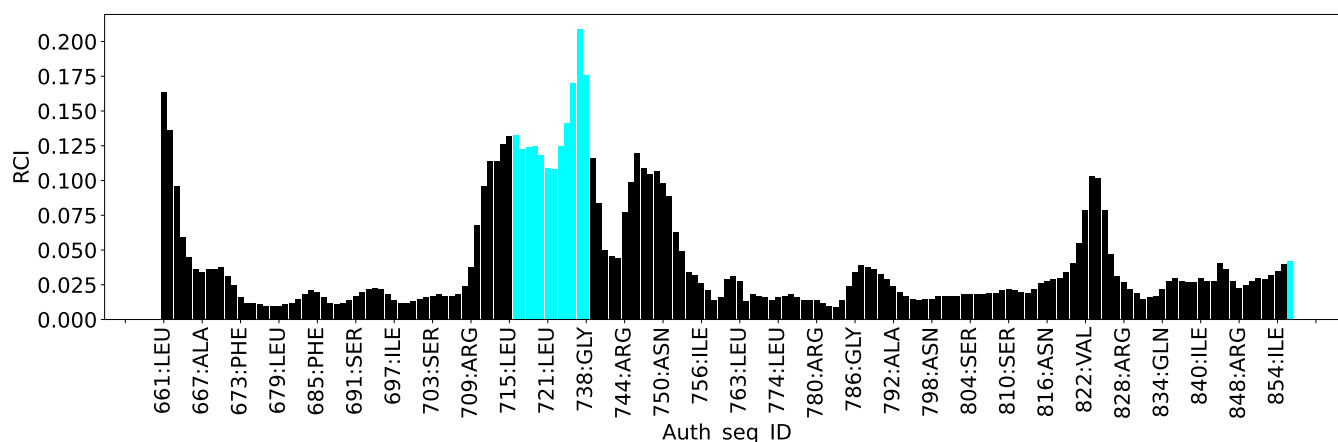
	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	622/2643 (24%)	156/1074 (15%)	310/1050 (30%)	156/519 (30%)

*Continued on next page...*

	<b>Total</b>	<b><sup>1</sup>H</b>	<b><sup>13</sup>C</b>	<b><sup>15</sup>N</b>
Sidechain	0/4674 (0%)	0/3063 (0%)	0/1377 (0%)	0/234 (0%)
Aromatic	0/663 (0%)	0/324 (0%)	0/297 (0%)	0/42 (0%)
Overall	622/7980 (8%)	156/4461 (3%)	310/2724 (11%)	156/795 (20%)

There are no statistically unusual chemical shifts.

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	1746
Intra-residue ( $ i-j =0$ )	191
Sequential ( $ i-j =1$ )	522
Medium range ( $ i-j >1$ and $ i-j <5$ )	517
Long range ( $ i-j \geq 5$ )	131
Inter-chain	112
Hydrogen bond restraints	273
Disulfide bond restraints	0
Total dihedral-angle restraints	298
Number of unmapped restraints	38
Number of restraints per residue	3.5
Number of long range restraints per residue <sup>1</sup>	0.2

<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)	125.3	0.2
0.2-0.5 (Medium)	116.4	0.5
>0.5 (Large)	144.9	56.95

### 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)	44.5	7.23
10.0-20.0 (Medium)	None	None
>20.0 (Large)	None	None



## 9 Distance violation analysis ⓘ

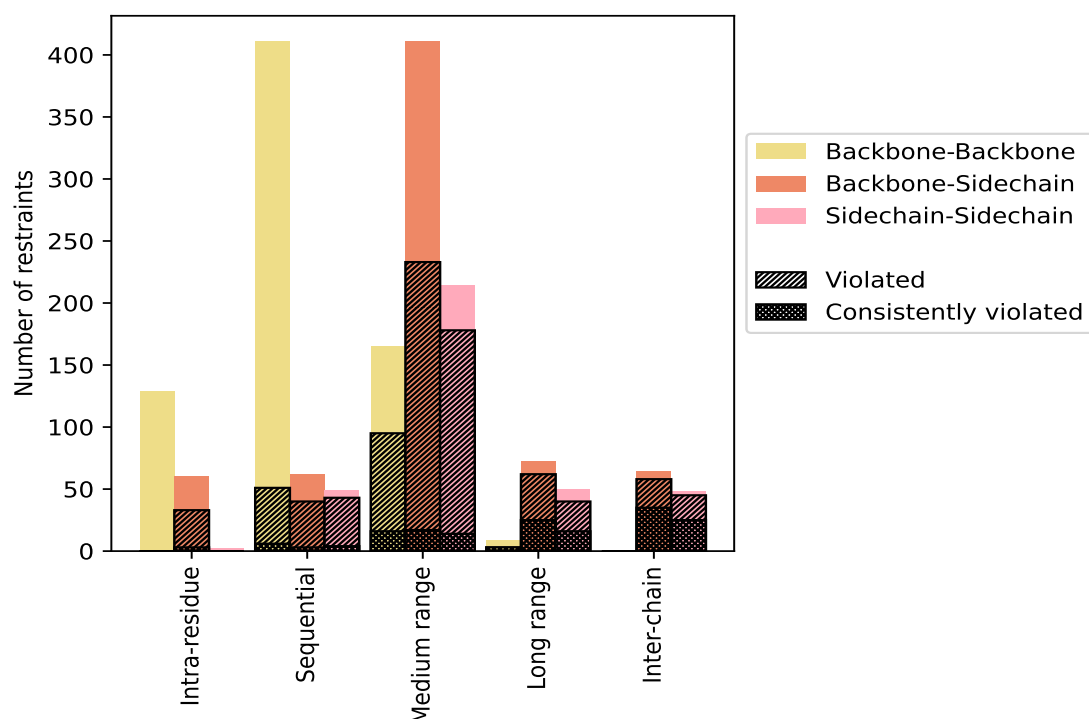
### 9.1 Summary of distance violations ⓘ

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>191</b>	<b>10.9</b>	<b>33</b>	<b>17.3</b>	<b>1.9</b>	<b>3</b>	<b>1.6</b>	<b>0.2</b>
Backbone-Backbone	129	7.4	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	60	3.4	33	55.0	1.9	3	5.0	0.2
Sidechain-Sidechain	2	0.1	0	0.0	0.0	0	0.0	0.0
<b>Sequential (<math> i-j =1</math>)</b>	<b>522</b>	<b>29.9</b>	<b>134</b>	<b>25.7</b>	<b>7.7</b>	<b>13</b>	<b>2.5</b>	<b>0.7</b>
Backbone-Backbone	411	23.5	51	12.4	2.9	6	1.5	0.3
Backbone-Sidechain	62	3.6	40	64.5	2.3	3	4.8	0.2
Sidechain-Sidechain	49	2.8	43	87.8	2.5	4	8.2	0.2
<b>Medium range (<math> i-j &gt;1</math> &amp; <math> i-j &lt;5</math>)</b>	<b>517</b>	<b>29.6</b>	<b>374</b>	<b>72.3</b>	<b>21.4</b>	<b>34</b>	<b>6.6</b>	<b>1.9</b>
Backbone-Backbone	165	9.5	95	57.6	5.4	16	9.7	0.9
Backbone-Sidechain	153	8.8	116	75.8	6.6	11	7.2	0.6
Sidechain-Sidechain	199	11.4	163	81.9	9.3	7	3.5	0.4
<b>Long range (<math> i-j \geq 5</math>)</b>	<b>131</b>	<b>7.5</b>	<b>105</b>	<b>80.2</b>	<b>6.0</b>	<b>44</b>	<b>33.6</b>	<b>2.5</b>
Backbone-Backbone	9	0.5	3	33.3	0.2	3	33.3	0.2
Backbone-Sidechain	72	4.1	62	86.1	3.6	25	34.7	1.4
Sidechain-Sidechain	50	2.9	40	80.0	2.3	16	32.0	0.9
<b>Inter-chain</b>	<b>112</b>	<b>6.4</b>	<b>103</b>	<b>92.0</b>	<b>5.9</b>	<b>60</b>	<b>53.6</b>	<b>3.4</b>
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	64	3.7	58	90.6	3.3	35	54.7	2.0
Sidechain-Sidechain	48	2.7	45	93.8	2.6	25	52.1	1.4
<b>Hydrogen bond</b>	<b>273</b>	<b>15.6</b>	<b>132</b>	<b>48.4</b>	<b>7.6</b>	<b>13</b>	<b>4.8</b>	<b>0.7</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>1746</b>	<b>100.0</b>	<b>881</b>	<b>50.5</b>	<b>50.5</b>	<b>167</b>	<b>9.6</b>	<b>9.6</b>
Backbone-Backbone	714	40.9	149	20.9	8.5	25	3.5	1.4
Backbone-Sidechain	669	38.3	426	63.7	24.4	83	12.4	4.8
Sidechain-Sidechain	363	20.8	306	84.3	17.5	59	16.3	3.4

<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 disulfied 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	9	54	187	62	76	388	4.82	56.95	10.19	0.32
2	11	42	166	68	74	361	5.16	54.31	10.43	0.36
3	10	49	203	62	76	400	4.67	55.51	10.03	0.33
4	9	38	183	62	81	373	4.92	56.86	10.31	0.32
5	15	49	168	68	78	378	4.83	55.88	10.17	0.31
6	17	62	196	63	77	415	4.44	55.56	9.81	0.32
7	10	49	204	64	75	402	4.57	54.51	9.91	0.33
8	12	45	209	63	80	409	4.57	56.35	9.94	0.3
9	11	46	178	68	79	382	4.89	53.78	10.17	0.32
10	14	46	190	68	86	404	4.58	55.73	9.92	0.3

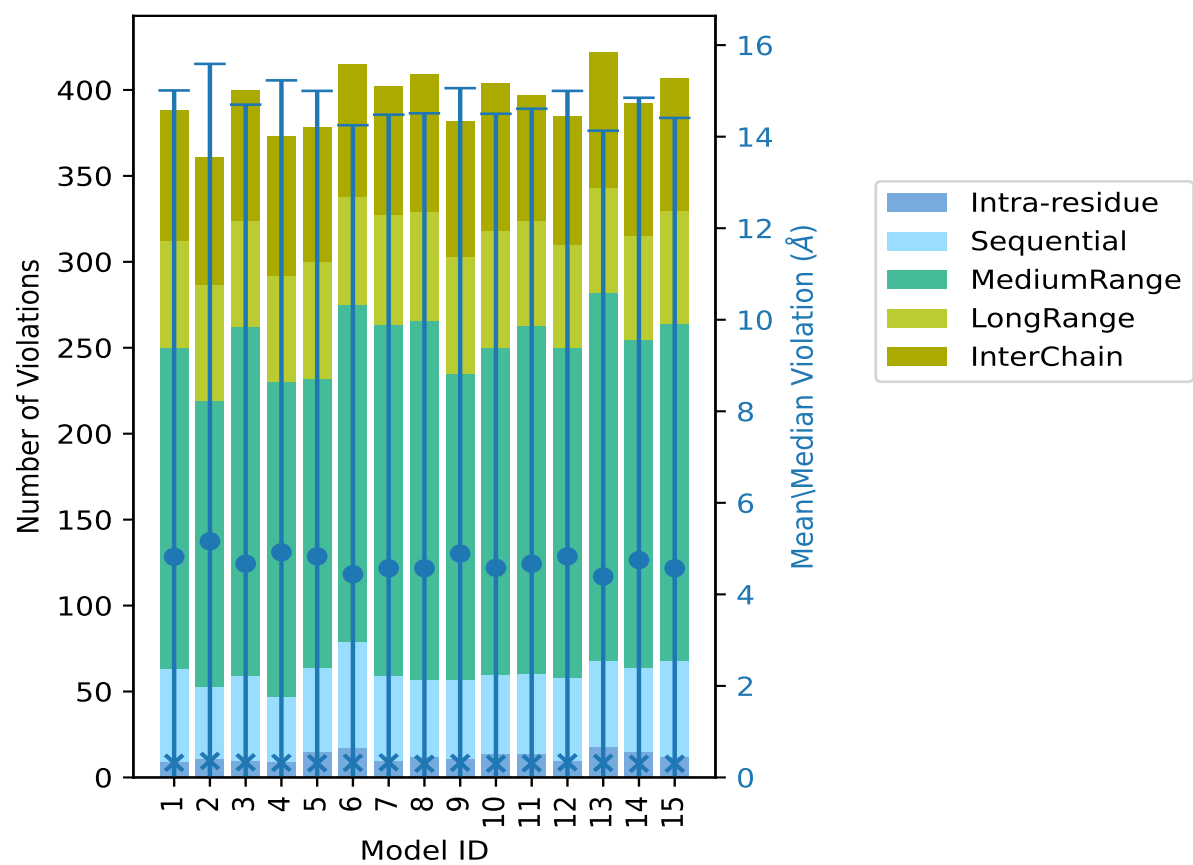
*Continued on next page...*

Continued from previous page...

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				
11	14	46	203	61	73	397	4.67	54.48	9.94	0.31
12	10	48	192	60	75	385	4.83	55.54	10.17	0.33
13	18	50	214	61	79	422	4.39	55.28	9.74	0.32
14	15	49	191	60	77	392	4.75	54.55	10.1	0.3
15	12	56	196	66	77	407	4.57	56.0	9.84	0.3

<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 ⓘ



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 ⓘ

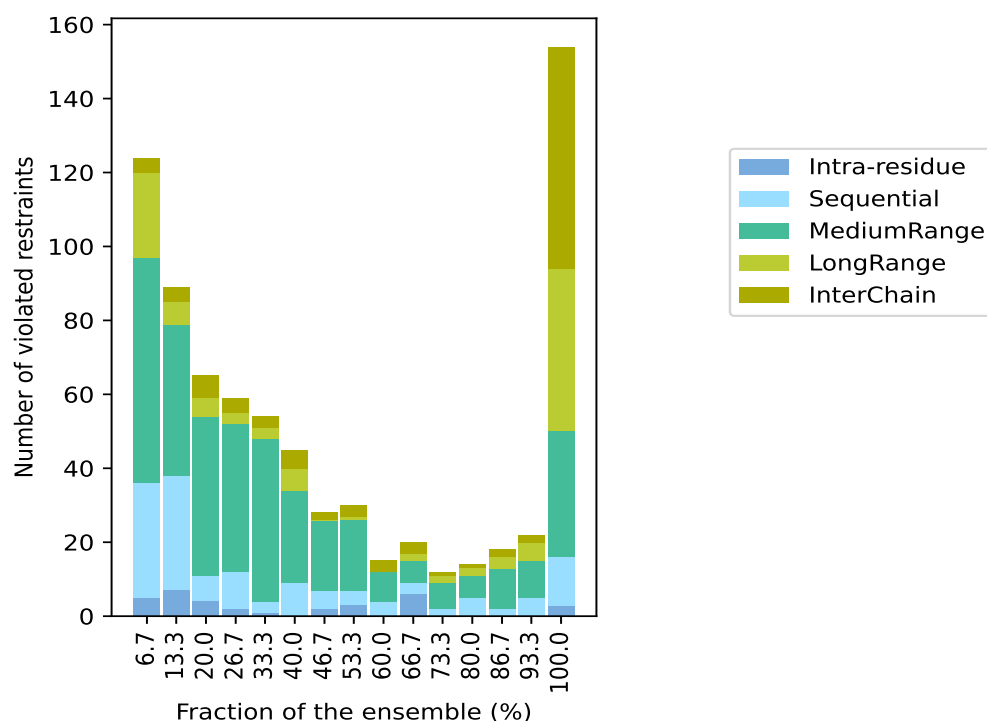
Violation analysis may find that some restraints are violated in few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of the ensemble. In total, 724(IR:158, SQ:388, MR:143, LR:26, IC:9) restraints are not 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>	%
5	31	61	23	4	124	1	6.7
7	31	41	6	4	89	2	13.3
4	7	43	5	6	65	3	20.0
2	10	40	3	4	59	4	26.7
1	3	44	3	3	54	5	33.3
0	9	25	6	5	45	6	40.0
2	5	19	0	2	28	7	46.7
3	4	19	1	3	30	8	53.3
0	4	8	0	3	15	9	60.0
6	3	6	2	3	20	10	66.7
0	2	7	2	1	12	11	73.3
0	5	6	2	1	14	12	80.0
0	2	11	3	2	18	13	86.7
0	5	10	5	2	22	14	93.3
3	13	34	44	60	154	15	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

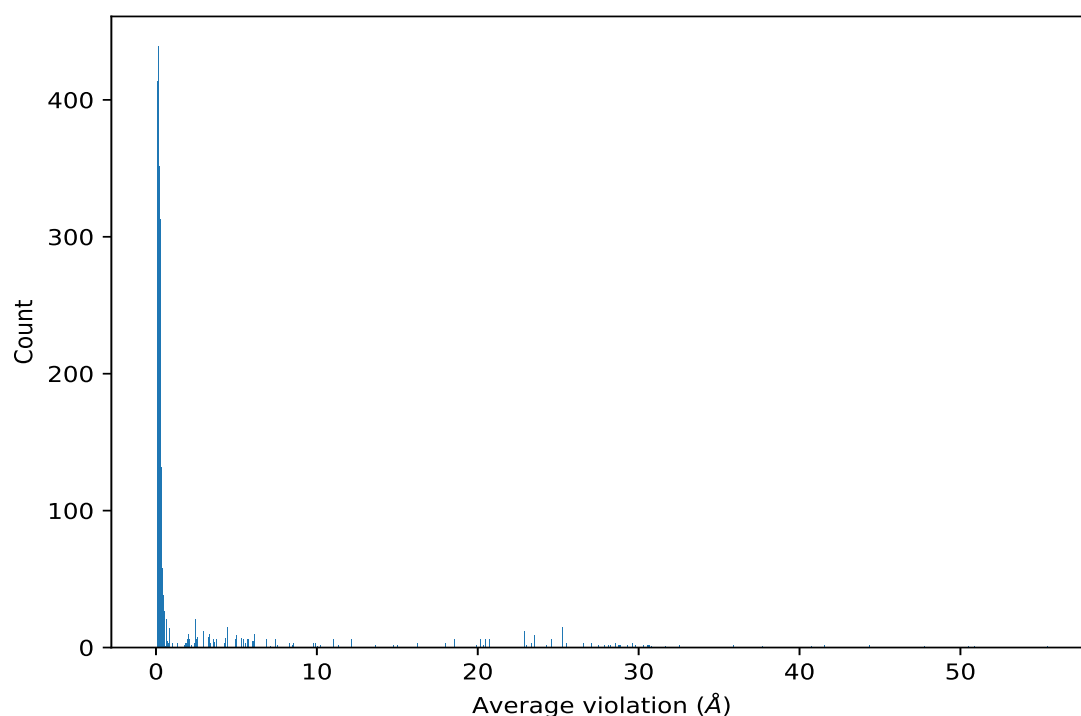
### 9.3.1 Bar graph : Distance violation statistics for the ensemble [i](#)



## 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.

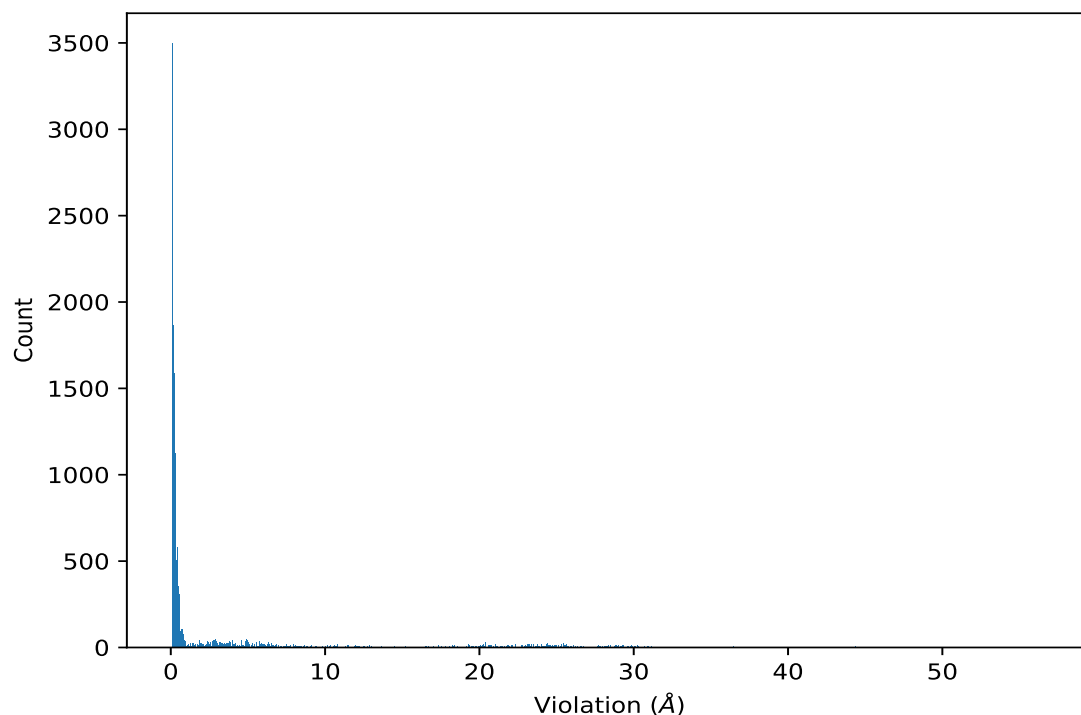
Key	Atom-1	Atom-2	Models <sup>1</sup>	Mean (Å)	SD <sup>1</sup> (Å)	Median (Å)
(2,181)	1:806:B:GLU:H	1:772:A:ARG:HE	15	55.42	0.91	55.54
(2,178)	1:806:B:GLU:H	1:771:A:LEU:HG	15	50.87	0.91	50.41
(2,182)	1:806:A:GLU:H	1:772:A:ARG:HE	15	50.52	0.88	50.68
(2,179)	1:806:A:GLU:H	1:771:A:LEU:HG	15	47.76	0.56	47.6
(2,140)	1:771:A:LEU:H	1:806:B:GLU:HB2	15	44.31	0.6	44.38
(2,140)	1:771:A:LEU:H	1:806:B:GLU:HB3	15	44.31	0.6	44.38
(2,139)	1:771:A:LEU:H	1:806:A:GLU:HB2	15	41.56	0.56	41.55
(2,139)	1:771:A:LEU:H	1:806:A:GLU:HB3	15	41.56	0.56	41.55
(2,175)	1:803:B:TRP:HE1	1:764:A:SER:HG	15	40.79	1.03	40.8
(2,176)	1:803:A:TRP:HE1	1:764:A:SER:HG	15	37.7	0.93	37.76
(2,172)	1:803:B:TRP:HE1	1:764:A:SER:HB2	15	35.91	0.64	35.98
(2,172)	1:803:B:TRP:HE1	1:764:A:SER:HB3	15	35.91	0.64	35.98
(2,173)	1:803:A:TRP:HE1	1:764:A:SER:HB2	15	32.52	0.8	32.43

<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)
(2,181)	1:806:B:GLU:H	1:772:A:ARG:HE	1	56.95
(2,181)	1:806:B:GLU:H	1:772:A:ARG:HE	4	56.86
(2,181)	1:806:B:GLU:H	1:772:A:ARG:HE	8	56.35
(2,181)	1:806:B:GLU:H	1:772:A:ARG:HE	15	56.0
(2,181)	1:806:B:GLU:H	1:772:A:ARG:HE	5	55.88
(2,181)	1:806:B:GLU:H	1:772:A:ARG:HE	10	55.73
(2,181)	1:806:B:GLU:H	1:772:A:ARG:HE	6	55.56
(2,181)	1:806:B:GLU:H	1:772:A:ARG:HE	12	55.54
(2,181)	1:806:B:GLU:H	1:772:A:ARG:HE	3	55.51
(2,181)	1:806:B:GLU:H	1:772:A:ARG:HE	13	55.28

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

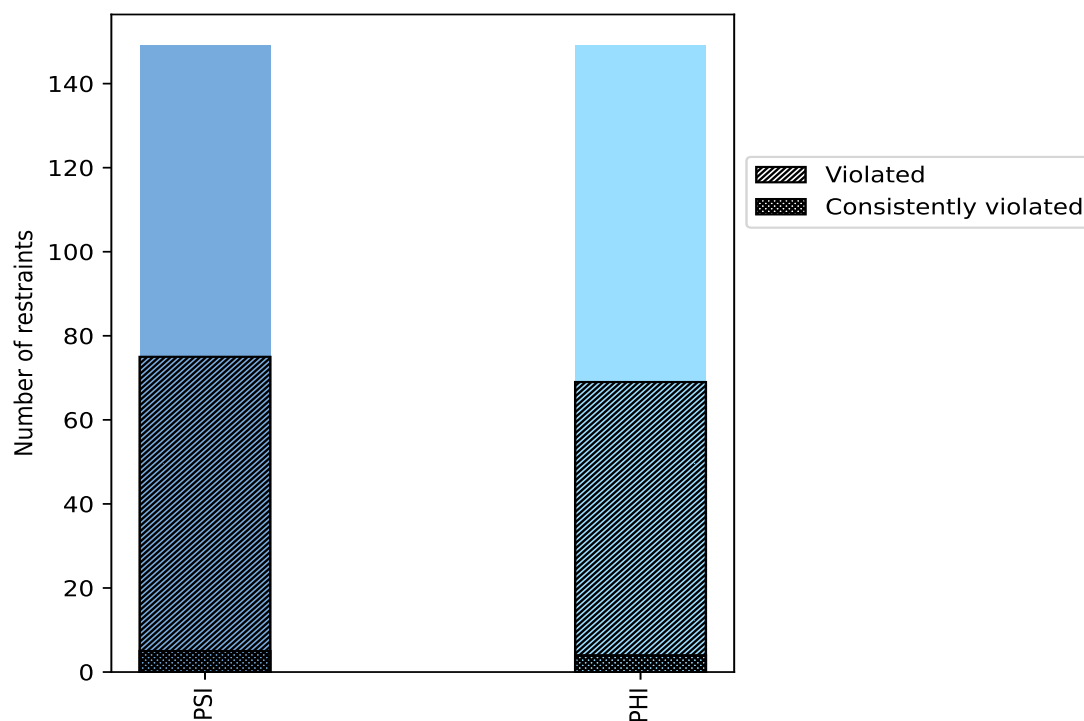
### 10.1 Summary of dihedral-angle violations [i](#)

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>
PSI	149	50.0	75	50.3	25.2	5	3.4	1.7
PHI	149	50.0	69	46.3	23.2	4	2.7	1.3
Total	298	100.0	144	48.3	48.3	9	3.0	3.0

<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

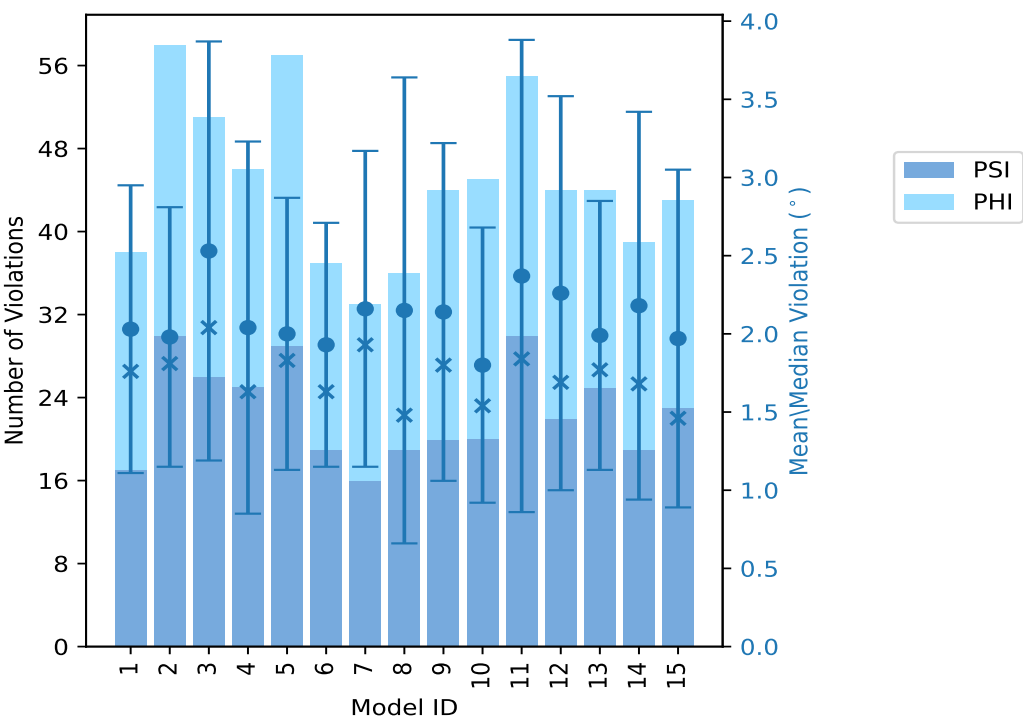


10.2 Dihedral-angle violation statistics for each model ⓘ

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 (°)
	PSI	PHI	Total				
1	17	21	38	2.03	4.62	0.92	1.76
2	30	28	58	1.98	3.88	0.83	1.81
3	26	25	51	2.53	5.3	1.34	2.04
4	25	21	46	2.04	6.43	1.19	1.63
5	29	28	57	2.0	4.74	0.87	1.83
6	19	18	37	1.93	4.07	0.78	1.63
7	16	17	33	2.16	4.77	1.01	1.93
8	19	17	36	2.15	7.23	1.49	1.48
9	20	24	44	2.14	4.93	1.08	1.8
10	20	25	45	1.8	5.42	0.88	1.54
11	30	25	55	2.37	7.09	1.51	1.84
12	22	22	44	2.26	5.32	1.26	1.69
13	25	19	44	1.99	4.64	0.86	1.77
14	19	20	39	2.18	5.41	1.24	1.68
15	23	20	43	1.97	4.97	1.08	1.46

10.2.1 Bar graph : Dihedral violation statistics for each model ⓘ



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

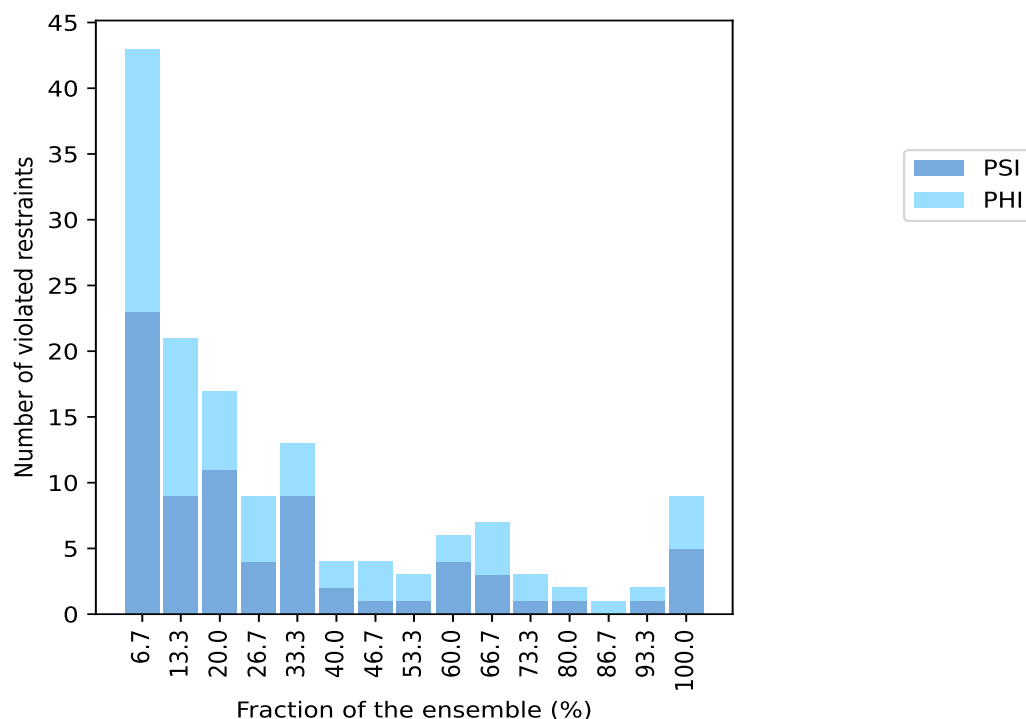
### 10.3 Dihedral-angle violation statistics for the ensemble [i](#)

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	
PSI	PHI	Total	Count <sup>1</sup>	%
23	20	43	1	6.7
9	12	21	2	13.3
11	6	17	3	20.0
4	5	9	4	26.7
9	4	13	5	33.3
2	2	4	6	40.0
1	3	4	7	46.7
1	2	3	8	53.3
4	2	6	9	60.0
3	4	7	10	66.7
1	2	3	11	73.3
1	1	2	12	80.0
0	1	1	13	86.7
1	1	2	14	93.3
5	4	9	15	100.0

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

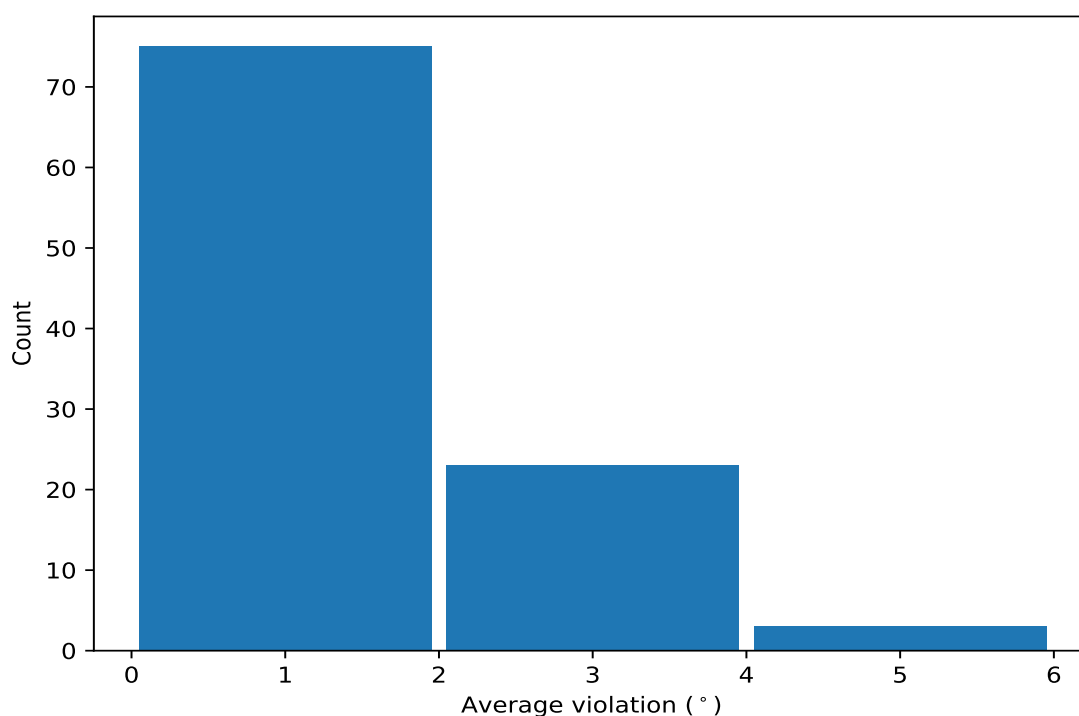
### 10.3.1 Bar graph : Dihedral-angle Violation statistics for the ensemble [i](#)



## 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 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.

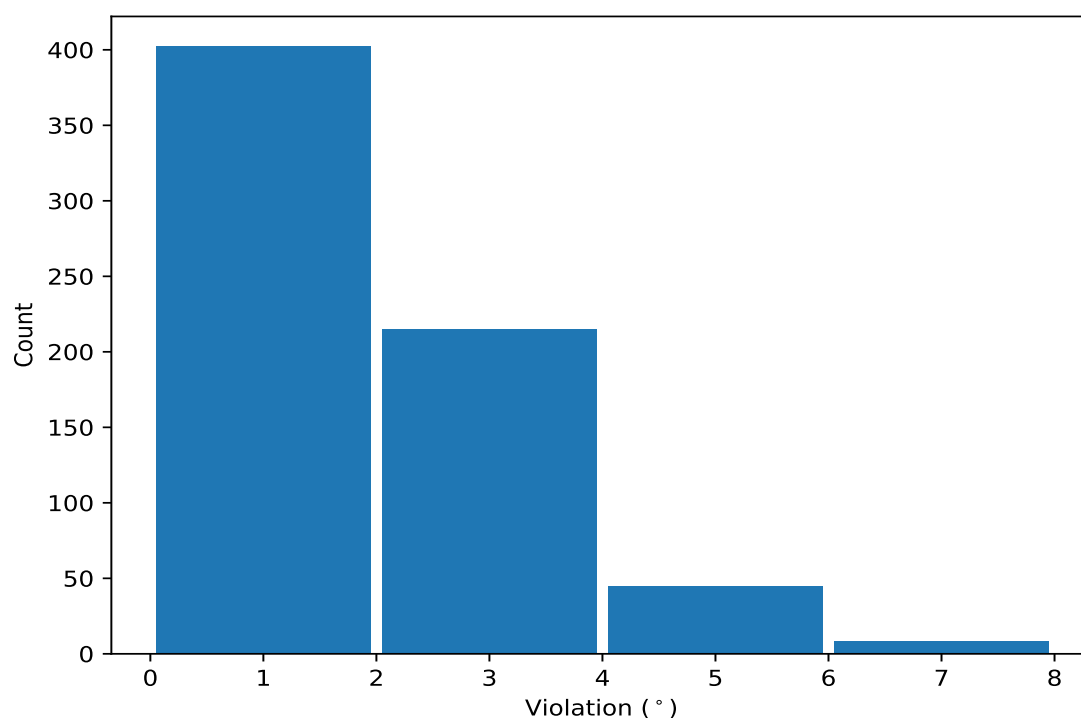
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models <sup>1</sup>	Mean	SD <sup>2</sup>	Median
(1,148)	1:770:B:ARG:N	1:770:B:ARG:CA	1:770:B:ARG:C	1:771:B:LEU:N	15	4.51	1.33	4.64
(1,149)	1:770:B:ARG:C	1:771:B:LEU:N	1:771:B:LEU:CA	1:771:B:LEU:C	15	4.41	1.43	4.69
(1,188)	1:790:B:TRP:N	1:790:B:TRP:CA	1:790:B:TRP:C	1:791:B:GLU:N	15	4.09	0.66	4.07
(1,62)	1:699:B:PHE:N	1:699:B:PHE:CA	1:699:B:PHE:C	1:700:B:ALA:N	15	3.88	0.8	3.83
(1,126)	1:754:B:ALA:N	1:754:B:ALA:CA	1:754:B:ALA:C	1:755:B:LEU:N	15	3.46	1.28	3.29
(1,127)	1:754:B:ALA:C	1:755:B:LEU:N	1:755:B:LEU:CA	1:755:B:LEU:C	15	3.27	1.31	3.13
(1,63)	1:699:B:PHE:C	1:700:B:ALA:N	1:700:B:ALA:CA	1:700:B:ALA:C	15	3.04	0.68	2.97
(1,119)	1:750:B:ASN:C	1:751:B:GLY:N	1:751:B:GLY:CA	1:751:B:GLY:C	15	2.57	0.62	2.75
(1,118)	1:750:B:ASN:N	1:750:B:ASN:CA	1:750:B:ASN:C	1:751:B:GLY:N	15	2.19	0.82	2.1
(1,56)	1:696:B:ARG:N	1:696:B:ARG:CA	1:696:B:ARG:C	1:697:B:ILE:N	14	2.3	0.88	2.14

<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 (°)
(1,148)	1:770:B:ARG:N	1:770:B:ARG:CA	1:770:B:ARG:C	1:771:B:LEU:N	8	7.23
(1,149)	1:770:B:ARG:C	1:771:B:LEU:N	1:771:B:LEU:CA	1:771:B:LEU:C	8	7.2
(1,133)	1:757:B:TRP:C	1:758:B:ASP:N	1:758:B:ASP:CA	1:758:B:ASP:C	11	7.09
(1,148)	1:770:B:ARG:N	1:770:B:ARG:CA	1:770:B:ARG:C	1:771:B:LEU:N	11	6.96
(1,149)	1:770:B:ARG:C	1:771:B:LEU:N	1:771:B:LEU:CA	1:771:B:LEU:C	11	6.75
(1,132)	1:757:B:TRP:N	1:757:B:TRP:CA	1:757:B:TRP:C	1:758:B:ASP:N	11	6.66
(1,126)	1:754:B:ALA:N	1:754:B:ALA:CA	1:754:B:ALA:C	1:755:B:LEU:N	4	6.43
(1,127)	1:754:B:ALA:C	1:755:B:LEU:N	1:755:B:LEU:CA	1:755:B:LEU:C	4	6.1
(1,188)	1:790:B:TRP:N	1:790:B:TRP:CA	1:790:B:TRP:C	1:791:B:GLU:N	10	5.42
(1,148)	1:770:B:ARG:N	1:770:B:ARG:CA	1:770:B:ARG:C	1:771:B:LEU:N	14	5.41