



# wwPDB NMR Structure Validation Summary Report ⓘ

Jun 3, 2023 – 01:39 PM EDT

PDB ID : 6UJV  
BMRB ID : 30503  
Title : Model of the HIV-1 gp41 membrane-proximal external region, transmembrane domain and cytoplasmic tail (LLP2)  
Authors : Piai, A.; Fu, Q.; Cai, Y.; Ghantous, F.; Xiao, T.; Shaik, M.M.; Peng, H.; Rits-Volloch, S.; Liu, Z.; Chen, W.; Seaman, M.S.; Chen, B.; Chou, J.J.  
Deposited on : 2019-10-03

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

We welcome your comments at [validation@mail.wwpdb.org](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

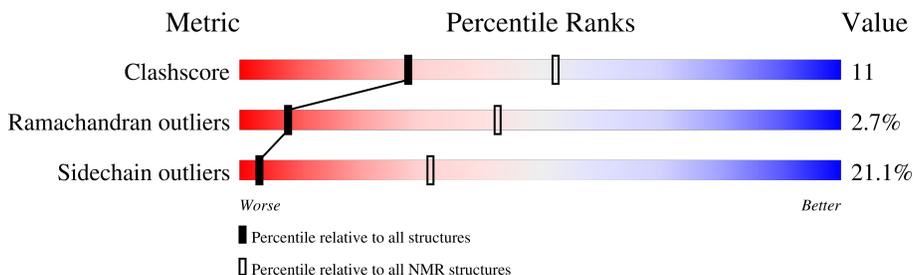
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 9%.

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	129	49% 30% 17%
1	B	129	51% 29% 17%
1	C	129	49% 29% 17%

## 2 Ensemble composition and analysis

This entry contains 15 models. Model 9 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:785, B:660-B:716, B:739-B:788, C:660-C:716, C:739-C:783 (313)	1.04	9

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 4 single-model clusters were found.

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 5577 atoms, of which 2859 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Envelope glycoprotein GP41.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
1	A	107	1859	595	953	163	148	0
1	B	107	1859	595	953	163	148	0
1	C	107	1859	595	953	163	148	0

There are 21 discrepancies between the modelled and reference sequences:

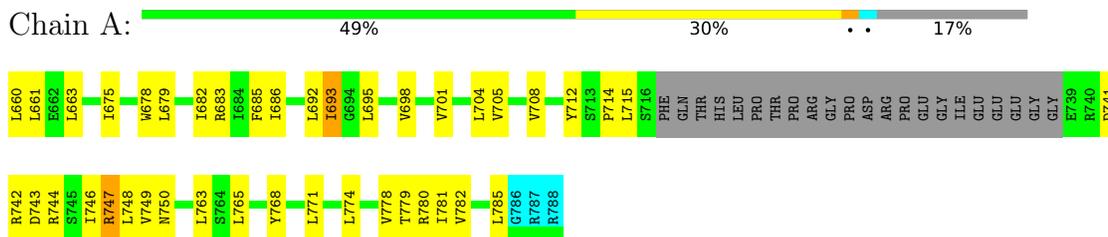
Chain	Residue	Modelled	Actual	Comment	Reference
A	683	ARG	LYS	conflict	UNP A0A060KRW4
A	684	ILE	LEU	conflict	UNP A0A060KRW4
A	687	ILE	MET	conflict	UNP A0A060KRW4
A	691	SER	GLY	conflict	UNP A0A060KRW4
A	704	LEU	ILE	conflict	UNP A0A060KRW4
A	723	THR	ILE	conflict	UNP A0A060KRW4
A	764	SER	CYS	engineered mutation	UNP A0A060KRW4
B	683	ARG	LYS	conflict	UNP A0A060KRW4
B	684	ILE	LEU	conflict	UNP A0A060KRW4
B	687	ILE	MET	conflict	UNP A0A060KRW4
B	691	SER	GLY	conflict	UNP A0A060KRW4
B	704	LEU	ILE	conflict	UNP A0A060KRW4
B	723	THR	ILE	conflict	UNP A0A060KRW4
B	764	SER	CYS	engineered mutation	UNP A0A060KRW4
C	683	ARG	LYS	conflict	UNP A0A060KRW4
C	684	ILE	LEU	conflict	UNP A0A060KRW4
C	687	ILE	MET	conflict	UNP A0A060KRW4
C	691	SER	GLY	conflict	UNP A0A060KRW4
C	704	LEU	ILE	conflict	UNP A0A060KRW4
C	723	THR	ILE	conflict	UNP A0A060KRW4
C	764	SER	CYS	engineered mutation	UNP A0A060KRW4

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

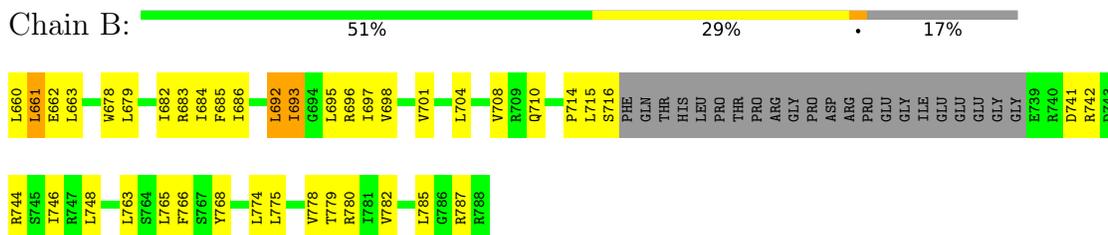
### 4.1 Average score per residue in the NMR ensemble

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.

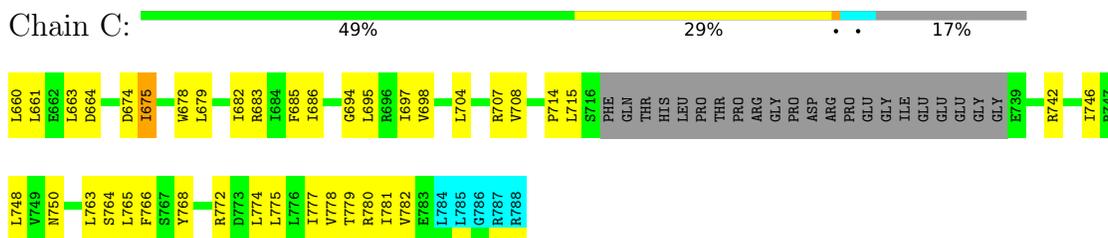
- Molecule 1: Envelope glycoprotein GP41



- Molecule 1: Envelope glycoprotein GP41



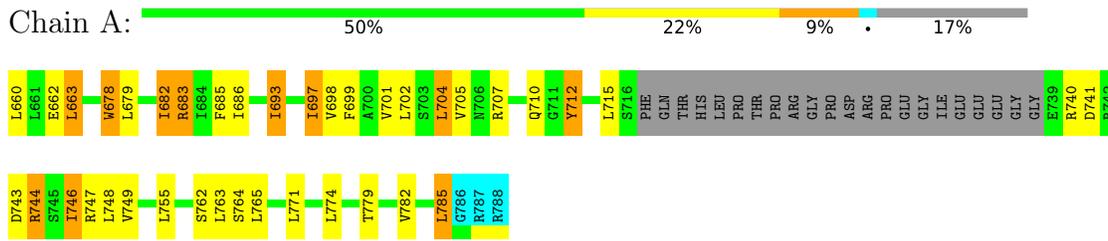
- Molecule 1: Envelope glycoprotein GP41



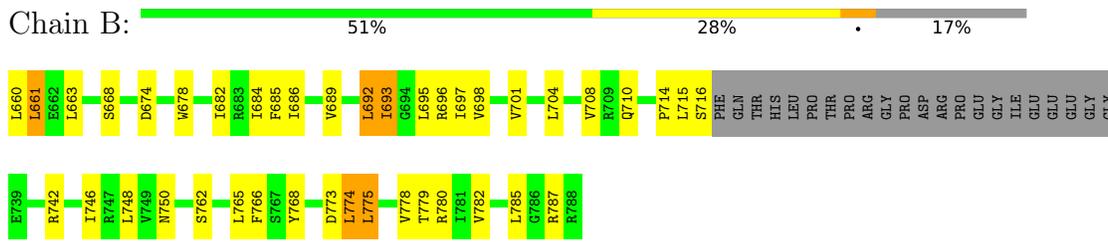
### 4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 9. Colouring as in section 4.1 above.

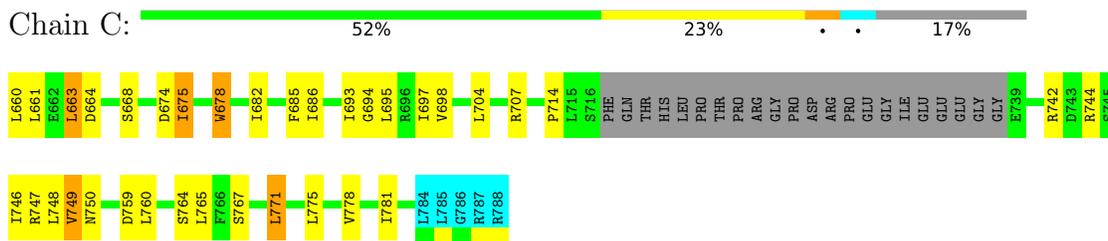
- Molecule 1: Envelope glycoprotein GP41



- Molecule 1: Envelope glycoprotein GP41



- Molecule 1: Envelope glycoprotein GP41



## 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	structure calculation	
X-PLOR NIH	refinement	

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	520
Number of shifts mapped to atoms	427
Number of unparsed shifts	0
Number of shifts with mapping errors	93
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	9%

## 6 Model quality i

### 6.1 Standard geometry i

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	0.76±0.01	0±0/899 ( 0.0± 0.0%)	0.94±0.01	0±0/1223 ( 0.0± 0.0%)
1	B	0.75±0.01	0±0/925 ( 0.0± 0.0%)	0.92±0.01	0±0/1253 ( 0.0± 0.0%)
1	C	0.74±0.01	0±0/883 ( 0.0± 0.0%)	0.92±0.01	0±0/1201 ( 0.0± 0.0%)
All	All	0.75	0/40605 ( 0.0%)	0.93	1/55155 ( 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	A	0.0±0.0	0.1±0.3
All	All	0	2

There are no bond-length outliers.

All unique angle outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
1	C	681	TYR	CB-CG-CD1	-5.21	117.87	121.00	5	1

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	707	ARG	Sidechain	1
1	A	747	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	879	924	919	24±5
1	B	906	953	948	25±4
1	C	863	902	897	23±5
All	All	39720	41685	41460	861

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 11.

5 of 434 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:781:ILE:HG23	1:B:746:ILE:HG23	0.86	1.47	8	1
1:A:686:ILE:HD12	1:C:686:ILE:HG21	0.81	1.52	9	12
1:A:705:VAL:HG22	1:A:748:LEU:HD11	0.80	1.53	13	1
1:B:781:ILE:HG23	1:C:746:ILE:HG23	0.79	1.52	4	2
1:A:686:ILE:HG21	1:B:686:ILE:HD12	0.78	1.55	15	12

## 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	101/129 (78%)	93±1 (92±1%)	6±1 (5±1%)	2±1 (2±1%)	9	48
1	B	103/129 (80%)	95±2 (92±2%)	5±2 (5±2%)	3±1 (3±1%)	8	42
1	C	99/129 (77%)	92±1 (93±1%)	4±1 (4±1%)	3±1 (3±1%)	7	40
All	All	4545/5805 (78%)	4207 (93%)	215 (5%)	123 (3%)	8	43

5 of 24 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	B	661	LEU	14
1	C	714	PRO	13
1	A	661	LEU	11
1	A	714	PRO	11
1	B	714	PRO	11

### 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	97/117 (83%)	76±3 (78±3%)	21±3 (22±3%)	<b>3</b> 30
1	B	99/117 (85%)	77±4 (78±4%)	22±4 (22±4%)	<b>3</b> 30
1	C	95/117 (81%)	77±3 (81±3%)	18±3 (19±3%)	<b>4</b> 35
All	All	4365/5265 (83%)	3444 (79%)	921 (21%)	<b>3</b> 32

5 of 215 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	C	685	PHE	15
1	A	704	LEU	14
1	B	685	PHE	14
1	B	692	LEU	14
1	A	712	TYR	14

### 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 9% for the well-defined parts and 9% for the entire structure.

### 7.1 Chemical shift list 1

File name: working\_cs.cif

Chemical shift list name: *HIV1\_gp41\_MPER-TMD-CTLLP2\_chemical\_shifts.tab*

#### 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	520
Number of shifts mapped to atoms	427
Number of unparsed shifts	0
Number of shifts with mapping errors	93
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 93) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	717	PHE	H	8.06	0.01	1
1	A	717	PHE	C	175.39	0.05	1
1	A	717	PHE	CA	57.46	0.05	1
1	A	717	PHE	CB	38.33	0.05	1
1	A	717	PHE	N	120.76	0.03	1
1	A	718	GLN	H	8.15	0.01	1
1	A	718	GLN	C	175.83	0.05	1
1	A	718	GLN	CA	55.52	0.05	1
1	A	718	GLN	CB	28.72	0.05	1
1	A	718	GLN	N	119.94	0.03	1
1	A	719	THR	H	8.04	0.01	1
1	A	719	THR	CA	61.73	0.05	1
1	A	719	THR	CB	68.95	0.05	1
1	A	719	THR	N	113.98	0.03	1

*Continued on next page...*

*Continued from previous page...*

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	720	HIS	C	174.61	0.05	1
1	A	720	HIS	CA	55.59	0.05	1
1	A	720	HIS	CB	29.94	0.05	1
1	A	721	LEU	H	8.03	0.01	1
1	A	721	LEU	CA	52.32	0.05	1
1	A	721	LEU	CB	40.6	0.05	1
1	A	721	LEU	N	123.72	0.03	1
1	A	722	PRO	C	176.55	0.05	1
1	A	722	PRO	CA	62.43	0.05	1
1	A	722	PRO	CB	30.9	0.05	1
1	A	723	THR	H	8.17	0.01	1
1	A	723	THR	CA	59.3	0.05	1
1	A	723	THR	CB	68.95	0.05	1
1	A	723	THR	N	116.44	0.03	1
1	A	724	PRO	C	176.7	0.05	1
1	A	724	PRO	CA	62.79	0.05	1
1	A	724	PRO	CB	31.23	0.05	1
1	A	725	ARG	H	8.37	0.01	1
1	A	725	ARG	C	176.44	0.05	1
1	A	725	ARG	CA	55.51	0.05	1
1	A	725	ARG	CB	30.18	0.05	1
1	A	725	ARG	N	121.04	0.03	1
1	A	726	GLY	H	8.23	0.01	1
1	A	726	GLY	C	171.92	0.05	1
1	A	726	GLY	CA	44.04	0.05	1
1	A	726	GLY	N	110.02	0.03	1
1	A	727	PRO	C	176.72	0.05	1
1	A	727	PRO	CA	62.84	0.05	1
1	A	727	PRO	CB	31.14	0.05	1
1	A	728	ASP	H	8.38	0.01	1
1	A	728	ASP	C	175.65	0.05	1
1	A	728	ASP	CA	54.07	0.05	1
1	A	728	ASP	CB	40.05	0.05	1
1	A	728	ASP	N	119.29	0.03	1
1	A	729	ARG	H	8.02	0.01	1
1	A	729	ARG	C	174.0	0.05	1
1	A	729	ARG	CA	53.08	0.05	1
1	A	729	ARG	CB	29.34	0.05	1
1	A	729	ARG	N	120.86	0.03	1
1	A	730	PRO	C	176.84	0.05	1
1	A	730	PRO	CA	62.79	0.05	1

*Continued on next page...*

*Continued from previous page...*

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	730	PRO	CB	31.09	0.05	1
1	A	731	GLU	H	8.57	0.01	1
1	A	731	GLU	C	177.03	0.05	1
1	A	731	GLU	CA	56.43	0.05	1
1	A	731	GLU	CB	29.29	0.05	1
1	A	731	GLU	N	120.32	0.03	1
1	A	732	GLY	H	8.36	0.01	1
1	A	732	GLY	C	173.94	0.05	1
1	A	732	GLY	CA	44.91	0.05	1
1	A	732	GLY	N	109.52	0.03	1
1	A	733	ILE	H	7.82	0.01	1
1	A	733	ILE	C	176.16	0.05	1
1	A	733	ILE	CA	60.51	0.05	1
1	A	733	ILE	CB	37.61	0.05	1
1	A	733	ILE	N	119.17	0.03	1
1	A	734	GLU	H	8.49	0.01	1
1	A	734	GLU	C	176.36	0.05	1
1	A	734	GLU	CA	56.11	0.05	1
1	A	734	GLU	CB	29.31	0.05	1
1	A	734	GLU	N	124.32	0.03	1
1	A	735	GLU	H	8.38	0.01	1
1	A	735	GLU	C	176.48	0.05	1
1	A	735	GLU	CA	56.11	0.05	1
1	A	735	GLU	CB	29.46	0.05	1
1	A	735	GLU	N	121.83	0.03	1
1	A	736	GLU	H	8.47	0.01	1
1	A	736	GLU	C	177.09	0.05	1
1	A	736	GLU	CA	56.32	0.05	1
1	A	736	GLU	CB	29.38	0.05	1
1	A	736	GLU	N	122.08	0.03	1
1	A	737	GLY	H	8.51	0.01	1
1	A	737	GLY	C	174.82	0.05	1
1	A	737	GLY	CA	45.04	0.05	1
1	A	737	GLY	N	109.96	0.03	1
1	A	738	GLY	H	8.27	0.01	1
1	A	738	GLY	C	174.5	0.05	1
1	A	738	GLY	CA	44.95	0.05	1
1	A	738	GLY	N	108.53	0.03	1

### 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
$^{13}\text{C}_\alpha$	120	0.00 $\pm$ 0.24	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	55	1.27 $\pm$ 0.09	Should be checked
$^{13}\text{C}'$	113	-0.13 $\pm$ 0.14	None needed (< 0.5 ppm)
$^{15}\text{N}$	116	0.72 $\pm$ 0.17	Should be applied

### 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 9%, i.e. 414 atoms were assigned a chemical shift out of a possible 4852. 0 out of 88 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	379/1572 (24%)	96/636 (15%)	187/626 (30%)	96/310 (31%)
Sidechain	35/2839 (1%)	0/1867 (0%)	35/840 (4%)	0/132 (0%)
Aromatic	0/441 (0%)	0/216 (0%)	0/201 (0%)	0/24 (0%)
Overall	414/4852 (9%)	96/2719 (4%)	222/1667 (13%)	96/466 (21%)

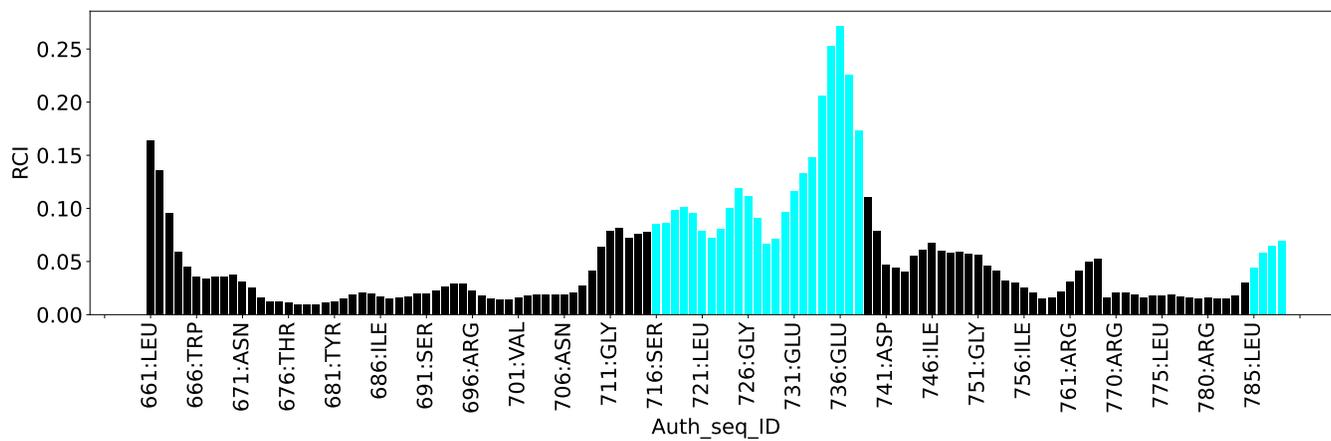
### 7.1.4 Statistically unusual chemical shifts [i](#)

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	1425
Intra-residue ( $ i-j =0$ )	192
Sequential ( $ i-j =1$ )	498
Medium range ( $ i-j >1$ and $ i-j <5$ )	540
Long range ( $ i-j \geq 5$ )	117
Inter-chain	0
Hydrogen bond restraints	78
Disulfide bond restraints	0
Total dihedral-angle restraints	180
Number of unmapped restraints	38
Number of restraints per residue	4.1
Number of long range restraints per residue <sup>1</sup>	0.3

<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)	74.6	0.2
0.2-0.5 (Medium)	82.0	0.5
>0.5 (Large)	132.4	29.07

### 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)	16.3	4.3
10.0-20.0 (Medium)	None	None
>20.0 (Large)	None	None

## 9 Distance violation analysis i

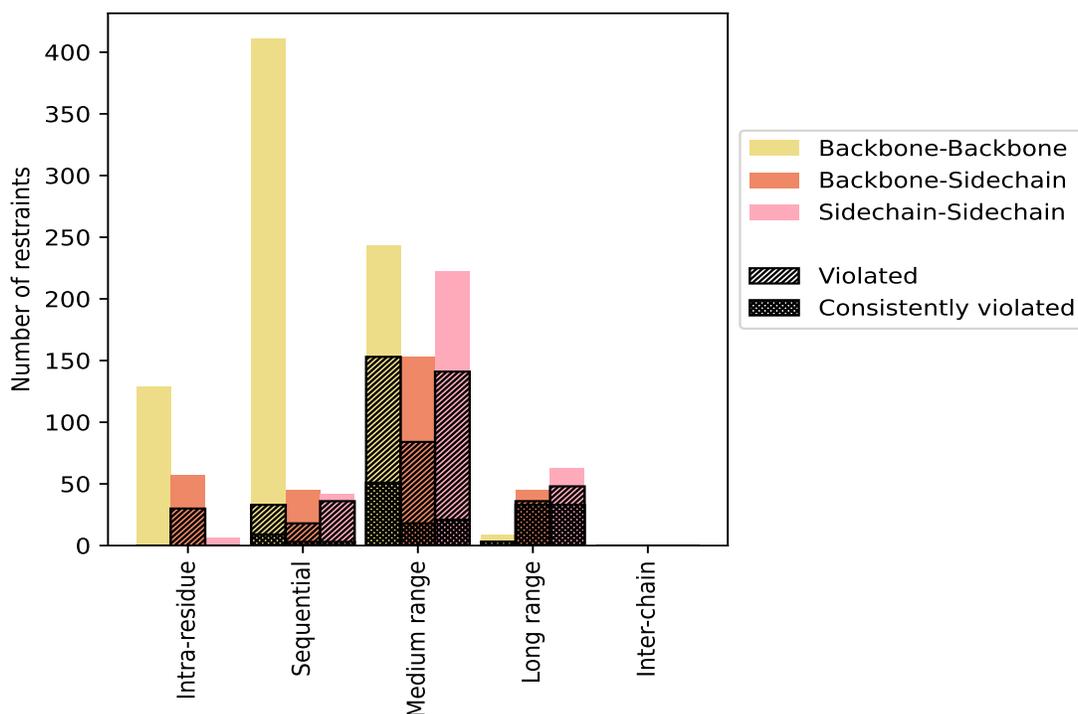
### 9.1 Summary of distance violations i

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>192</b>	<b>13.5</b>	<b>30</b>	<b>15.6</b>	<b>2.1</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
Backbone-Backbone	129	9.1	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	57	4.0	30	52.6	2.1	0	0.0	0.0
Sidechain-Sidechain	6	0.4	0	0.0	0.0	0	0.0	0.0
<b>Sequential (<math> i-j =1</math>)</b>	<b>498</b>	<b>34.9</b>	<b>87</b>	<b>17.5</b>	<b>6.1</b>	<b>15</b>	<b>3.0</b>	<b>1.1</b>
Backbone-Backbone	411	28.8	33	8.0	2.3	9	2.2	0.6
Backbone-Sidechain	45	3.2	18	40.0	1.3	3	6.7	0.2
Sidechain-Sidechain	42	2.9	36	85.7	2.5	3	7.1	0.2
<b>Medium range (<math> i-j &gt;1</math> &amp; <math> i-j &lt;5</math>)</b>	<b>540</b>	<b>37.9</b>	<b>321</b>	<b>59.4</b>	<b>22.5</b>	<b>57</b>	<b>10.6</b>	<b>4.0</b>
Backbone-Backbone	165	11.6	96	58.2	6.7	18	10.9	1.3
Backbone-Sidechain	153	10.7	84	54.9	5.9	18	11.8	1.3
Sidechain-Sidechain	222	15.6	141	63.5	9.9	21	9.5	1.5
<b>Long range (<math> i-j \geq 5</math>)</b>	<b>117</b>	<b>8.2</b>	<b>87</b>	<b>74.4</b>	<b>6.1</b>	<b>69</b>	<b>59.0</b>	<b>4.8</b>
Backbone-Backbone	9	0.6	3	33.3	0.2	3	33.3	0.2
Backbone-Sidechain	45	3.2	36	80.0	2.5	33	73.3	2.3
Sidechain-Sidechain	63	4.4	48	76.2	3.4	33	52.4	2.3
<b>Inter-chain</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>
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
<b>Hydrogen bond</b>	<b>78</b>	<b>5.5</b>	<b>57</b>	<b>73.1</b>	<b>4.0</b>	<b>33</b>	<b>42.3</b>	<b>2.3</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>1425</b>	<b>100.0</b>	<b>582</b>	<b>40.8</b>	<b>40.8</b>	<b>174</b>	<b>12.2</b>	<b>12.2</b>
Backbone-Backbone	792	55.6	189	23.9	13.3	63	8.0	4.4
Backbone-Sidechain	300	21.1	168	56.0	11.8	54	18.0	3.8
Sidechain-Sidechain	333	23.4	225	67.6	15.8	57	17.1	4.0

<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	6	27	147	72	0	252	4.45	27.75	7.88	0.5
2	15	27	165	75	0	282	4.03	28.91	7.72	0.39
3	15	30	171	75	0	291	3.96	28.36	7.49	0.43
4	12	30	159	78	0	279	4.08	28.66	7.59	0.44
5	12	30	165	75	0	282	4.05	27.57	7.54	0.4
6	15	30	171	72	0	288	3.99	28.52	7.48	0.48
7	9	30	171	72	0	282	4.07	29.07	7.74	0.4
8	9	54	189	72	0	324	3.58	27.54	7.17	0.38
9	6	36	177	81	0	300	3.76	28.3	7.35	0.35
10	12	30	165	69	0	276	4.23	28.9	7.82	0.5
11	15	27	165	78	0	285	4.09	27.96	7.56	0.46

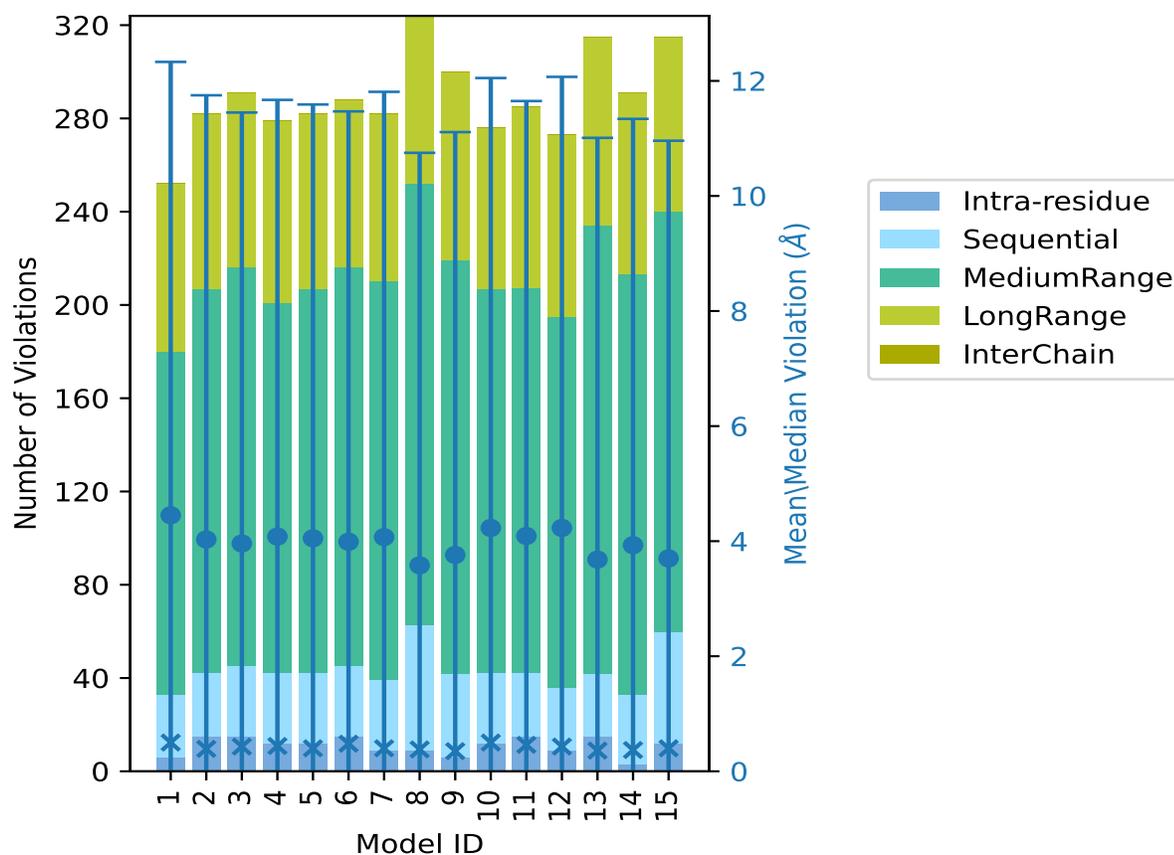
*Continued on next page...*

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	9	27	159	78	0	273	4.23	29.03	7.84	0.43
13	15	27	192	81	0	315	3.68	28.75	7.33	0.36
14	3	30	180	78	0	291	3.93	28.1	7.41	0.37
15	12	48	180	75	0	315	3.7	28.09	7.26	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 [i](#)

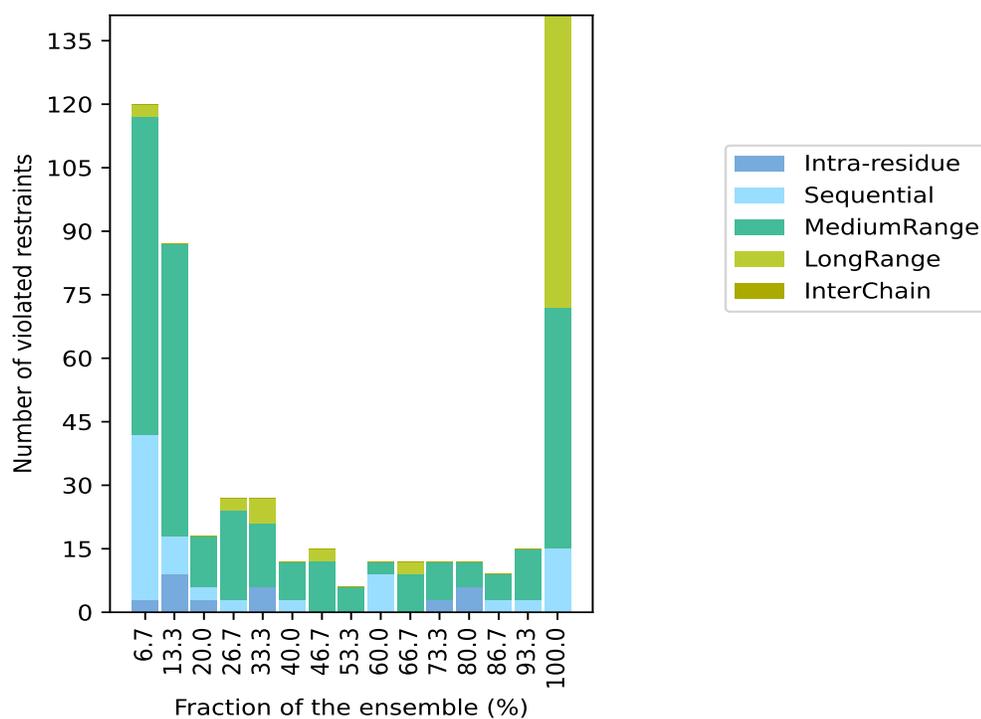
Violation analysis may find that some restraints are violated in few models and some are violated in most of models. The following table provides this information as number of violated restraints

for a given fraction of the ensemble. In total, 822(IR:162, SQ:411, MR:219, LR:30, IC:0) 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>	%
3	39	75	3	0	120	1	6.7
9	9	69	0	0	87	2	13.3
3	3	12	0	0	18	3	20.0
0	3	21	3	0	27	4	26.7
6	0	15	6	0	27	5	33.3
0	3	9	0	0	12	6	40.0
0	0	12	3	0	15	7	46.7
0	0	6	0	0	6	8	53.3
0	9	3	0	0	12	9	60.0
0	0	9	3	0	12	10	66.7
3	0	9	0	0	12	11	73.3
6	0	6	0	0	12	12	80.0
0	3	6	0	0	9	13	86.7
0	3	12	0	0	15	14	93.3
0	15	57	69	0	141	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

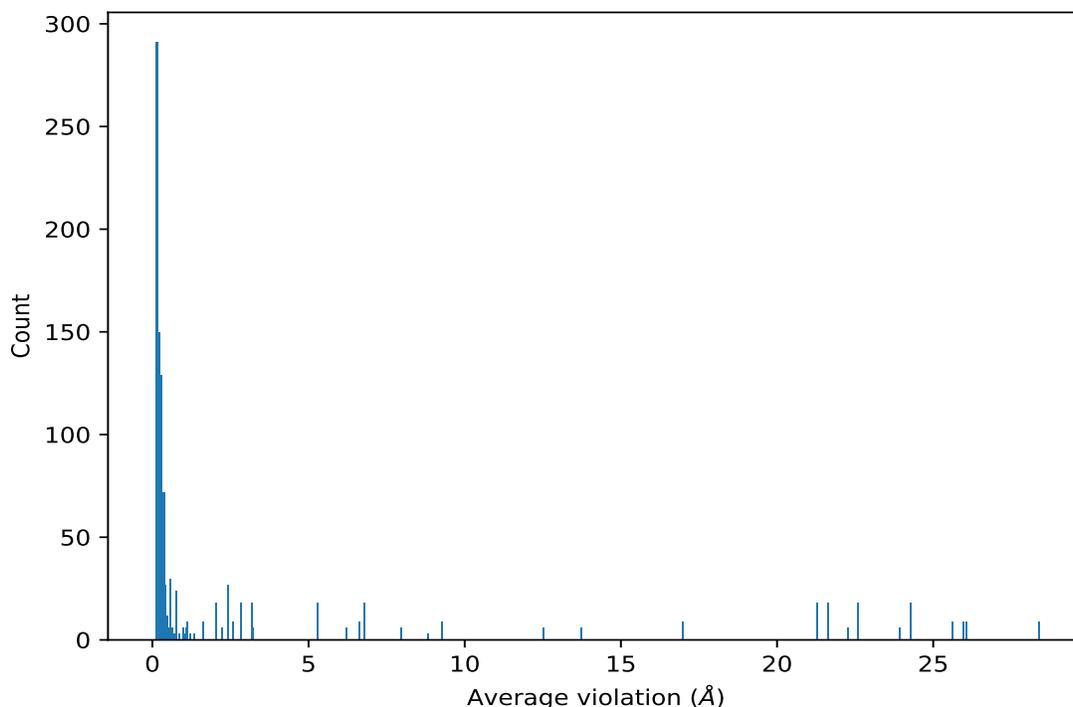
### 9.3.1 Bar graph : Distance violation statistics for the ensemble [i](#)



## 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)
(2,112)	1:A:781:ILE:H	1:A:746:ILE:HG21	15	28.37	0.5	28.36
(2,112)	1:A:781:ILE:H	1:A:746:ILE:HG22	15	28.37	0.5	28.36
(2,112)	1:A:781:ILE:H	1:A:746:ILE:HG23	15	28.37	0.5	28.36
(2,113)	1:A:781:ILE:H	1:A:746:ILE:HG21	15	28.37	0.5	28.36
(2,113)	1:A:781:ILE:H	1:A:746:ILE:HG22	15	28.37	0.5	28.36
(2,113)	1:A:781:ILE:H	1:A:746:ILE:HG23	15	28.37	0.5	28.36
(2,114)	1:A:781:ILE:H	1:A:746:ILE:HG21	15	28.37	0.5	28.36
(2,114)	1:A:781:ILE:H	1:A:746:ILE:HG22	15	28.37	0.5	28.36
(2,114)	1:A:781:ILE:H	1:A:746:ILE:HG23	15	28.37	0.5	28.36
(2,91)	1:A:749:VAL:HG11	1:A:782:VAL:H	15	26.05	0.59	25.96
(2,91)	1:A:749:VAL:HG12	1:A:782:VAL:H	15	26.05	0.59	25.96
(2,91)	1:A:749:VAL:HG13	1:A:782:VAL:H	15	26.05	0.59	25.96
(2,92)	1:A:749:VAL:HG11	1:A:782:VAL:H	15	26.05	0.59	25.96
(2,92)	1:A:749:VAL:HG12	1:A:782:VAL:H	15	26.05	0.59	25.96
(2,92)	1:A:749:VAL:HG13	1:A:782:VAL:H	15	26.05	0.59	25.96
(2,93)	1:A:749:VAL:HG11	1:A:782:VAL:H	15	26.05	0.59	25.96

*Continued on next page...*

Continued from previous page...

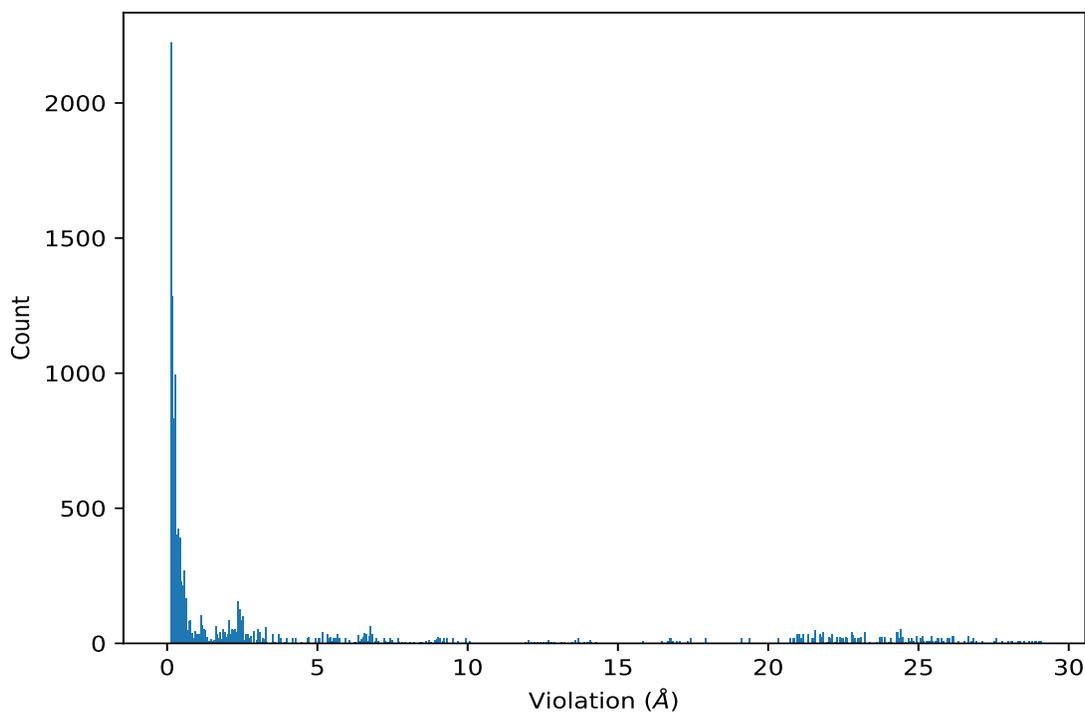
Key	Atom-1	Atom-2	Models <sup>1</sup>	Mean (Å)	SD <sup>1</sup> (Å)	Median (Å)
(2,93)	1:A:749:VAL:HG12	1:A:782:VAL:H	15	26.05	0.59	25.96
(2,93)	1:A:749:VAL:HG13	1:A:782:VAL:H	15	26.05	0.59	25.96
(2,115)	1:A:781:ILE:HD11	1:A:743:ASP:HA	15	25.97	0.69	25.75
(2,115)	1:A:781:ILE:HD12	1:A:743:ASP:HA	15	25.97	0.69	25.75
(2,115)	1:A:781:ILE:HD13	1:A:743:ASP:HA	15	25.97	0.69	25.75
(2,116)	1:A:781:ILE:HD11	1:A:743:ASP:HA	15	25.97	0.69	25.75
(2,116)	1:A:781:ILE:HD12	1:A:743:ASP:HA	15	25.97	0.69	25.75
(2,116)	1:A:781:ILE:HD13	1:A:743:ASP:HA	15	25.97	0.69	25.75
(2,117)	1:A:781:ILE:HD11	1:A:743:ASP:HA	15	25.97	0.69	25.75
(2,117)	1:A:781:ILE:HD12	1:A:743:ASP:HA	15	25.97	0.69	25.75
(2,117)	1:A:781:ILE:HD13	1:A:743:ASP:HA	15	25.97	0.69	25.75
(2,127)	1:A:785:LEU:HD21	1:A:749:VAL:HB	15	25.6	0.62	25.44

<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)
(2,114)	1:A:781:ILE:H	1:A:746:ILE:HG21	7	29.07
(2,114)	1:A:781:ILE:H	1:A:746:ILE:HG22	7	29.07
(2,114)	1:A:781:ILE:H	1:A:746:ILE:HG23	7	29.07
(2,113)	1:A:781:ILE:H	1:A:746:ILE:HG21	7	29.07
(2,113)	1:A:781:ILE:H	1:A:746:ILE:HG22	7	29.07
(2,113)	1:A:781:ILE:H	1:A:746:ILE:HG23	7	29.07
(2,112)	1:A:781:ILE:H	1:A:746:ILE:HG21	7	29.07
(2,112)	1:A:781:ILE:H	1:A:746:ILE:HG22	7	29.07
(2,112)	1:A:781:ILE:H	1:A:746:ILE:HG23	7	29.07
(2,114)	1:A:781:ILE:H	1:A:746:ILE:HG21	12	29.03
(2,114)	1:A:781:ILE:H	1:A:746:ILE:HG22	12	29.03
(2,114)	1:A:781:ILE:H	1:A:746:ILE:HG23	12	29.03
(2,113)	1:A:781:ILE:H	1:A:746:ILE:HG21	12	29.03
(2,113)	1:A:781:ILE:H	1:A:746:ILE:HG22	12	29.03
(2,113)	1:A:781:ILE:H	1:A:746:ILE:HG23	12	29.03
(2,112)	1:A:781:ILE:H	1:A:746:ILE:HG21	12	29.03
(2,112)	1:A:781:ILE:H	1:A:746:ILE:HG22	12	29.03
(2,112)	1:A:781:ILE:H	1:A:746:ILE:HG23	12	29.03
(2,114)	1:A:781:ILE:H	1:A:746:ILE:HG21	2	28.91
(2,114)	1:A:781:ILE:H	1:A:746:ILE:HG22	2	28.91
(2,114)	1:A:781:ILE:H	1:A:746:ILE:HG23	2	28.91
(2,113)	1:A:781:ILE:H	1:A:746:ILE:HG21	2	28.91
(2,113)	1:A:781:ILE:H	1:A:746:ILE:HG22	2	28.91
(2,113)	1:A:781:ILE:H	1:A:746:ILE:HG23	2	28.91
(2,112)	1:A:781:ILE:H	1:A:746:ILE:HG21	2	28.91
(2,112)	1:A:781:ILE:H	1:A:746:ILE:HG22	2	28.91
(2,112)	1:A:781:ILE:H	1:A:746:ILE:HG23	2	28.91
(2,114)	1:A:781:ILE:H	1:A:746:ILE:HG21	10	28.9

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

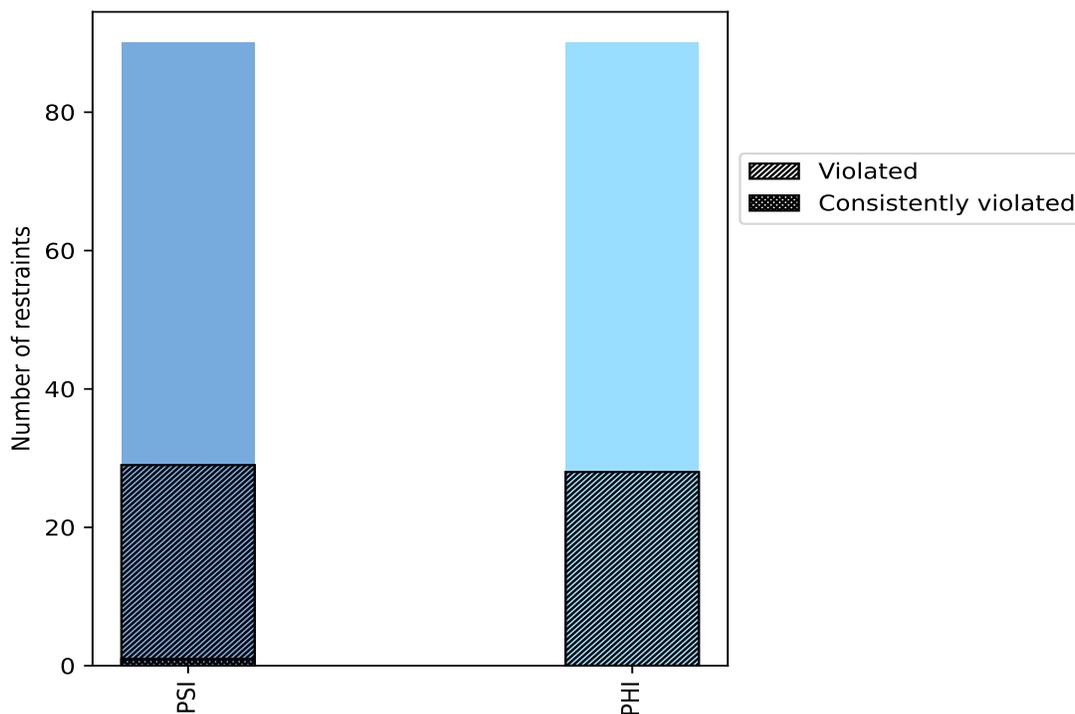
### 10.1 Summary of dihedral-angle violations [i](#)

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	90	50.0	29	32.2	16.1	1	1.1	0.6
PHI	90	50.0	28	31.1	15.6	0	0.0	0.0
Total	180	100.0	57	31.7	31.7	1	0.6	0.6

<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](#)



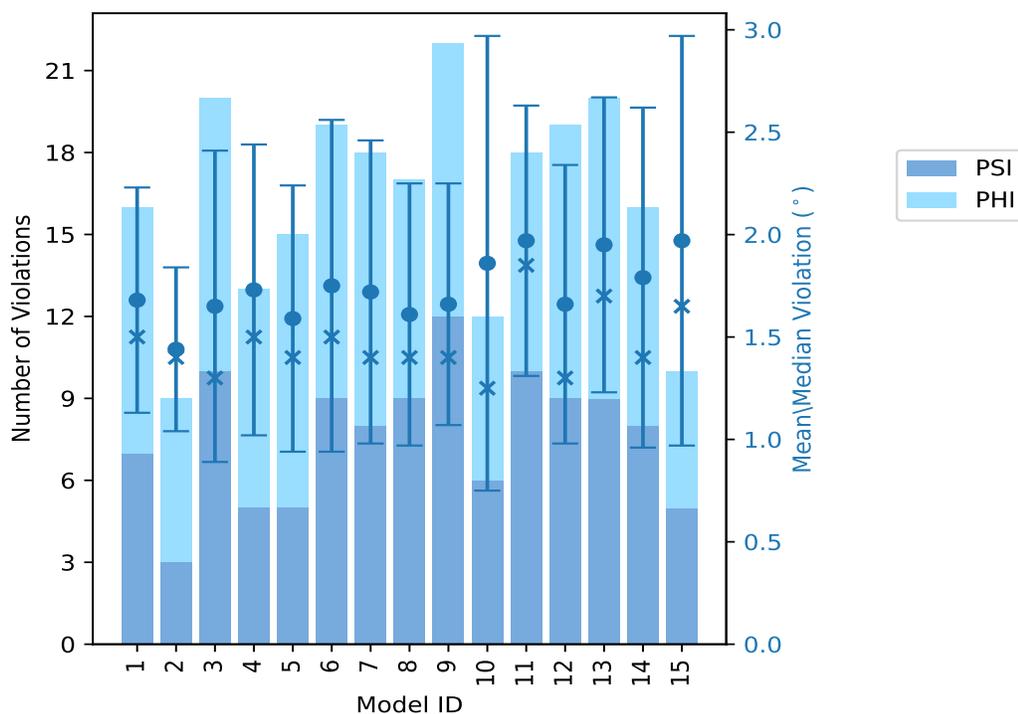
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

## 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 (°)
	PSI	PHI	Total				
1	7	9	16	1.68	2.8	0.55	1.5
2	3	6	9	1.44	2.5	0.4	1.4
3	10	10	20	1.65	3.6	0.76	1.3
4	5	8	13	1.73	3.3	0.71	1.5
5	5	10	15	1.59	3.4	0.65	1.4
6	9	10	19	1.75	3.8	0.81	1.5
7	8	10	18	1.72	3.7	0.74	1.4
8	9	8	17	1.61	3.3	0.64	1.4
9	12	10	22	1.66	3.0	0.59	1.4
10	6	6	12	1.86	4.3	1.11	1.25
11	10	8	18	1.97	3.4	0.66	1.85
12	9	10	19	1.66	3.3	0.68	1.3
13	9	11	20	1.95	3.6	0.72	1.7
14	8	8	16	1.79	3.7	0.83	1.4
15	5	5	10	1.97	3.9	1.0	1.65

### 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

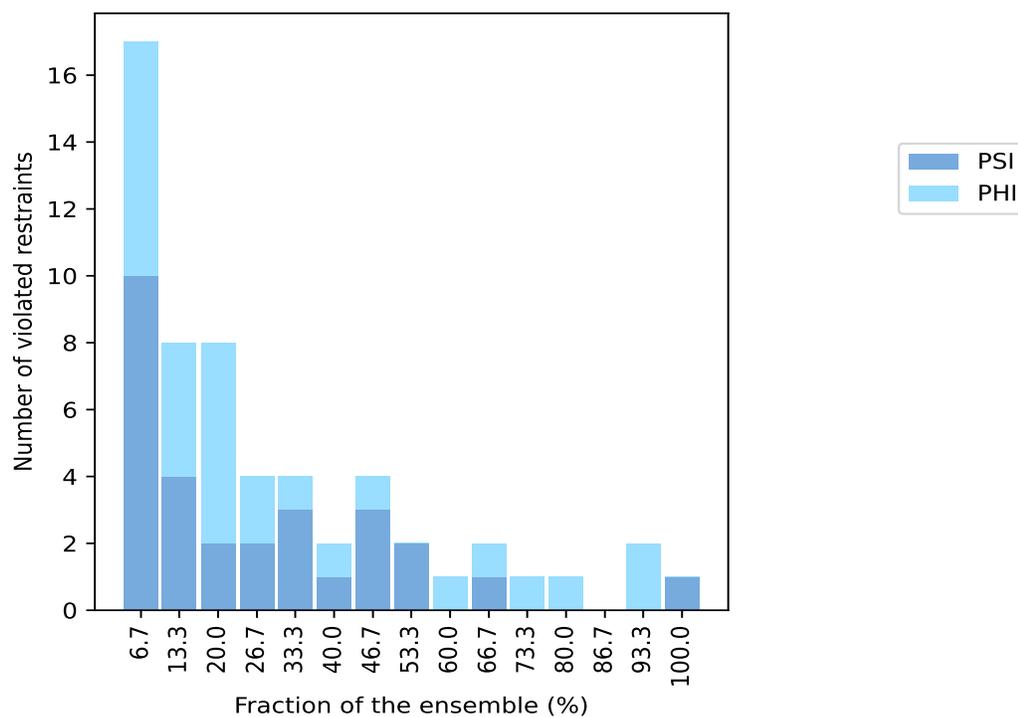
### 10.3 Dihedral-angle violation statistics for the ensemble

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>	%
10	7	17	1	6.7
4	4	8	2	13.3
2	6	8	3	20.0
2	2	4	4	26.7
3	1	4	5	33.3
1	1	2	6	40.0
3	1	4	7	46.7
2	0	2	8	53.3
0	1	1	9	60.0
1	1	2	10	66.7
0	1	1	11	73.3
0	1	1	12	80.0
0	0	0	13	86.7
0	2	2	14	93.3
1	0	1	15	100.0

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

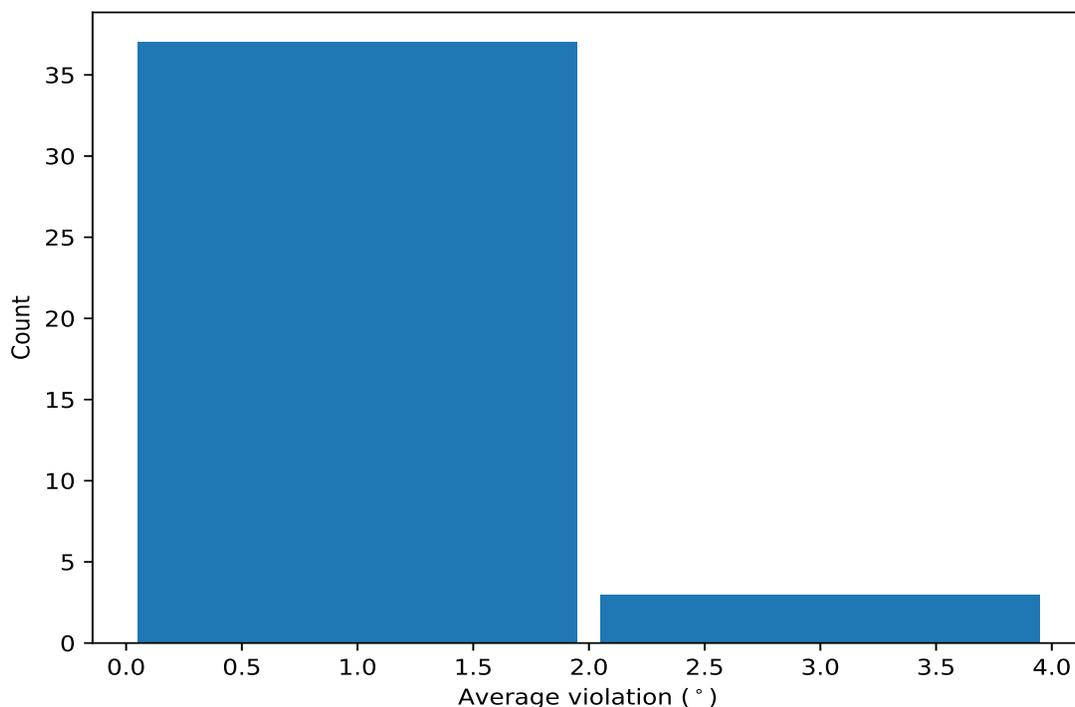
### 10.3.1 Bar graph : Dihedral-angle Violation statistics for the ensemble [i](#)



## 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.

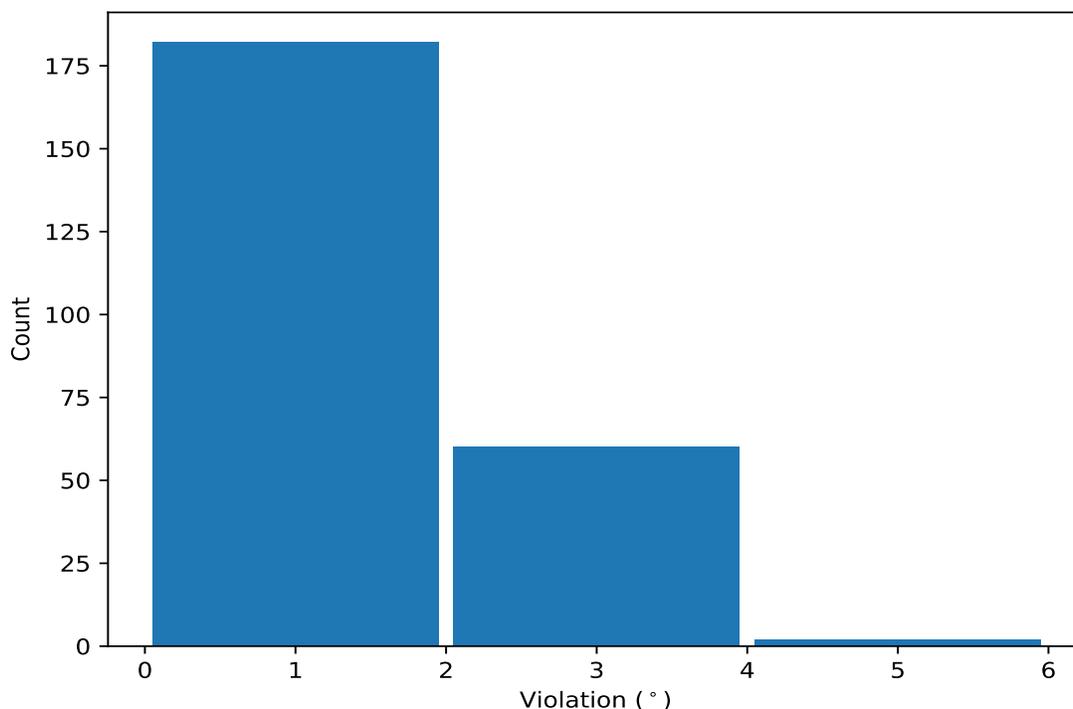
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models <sup>1</sup>	Mean	SD <sup>2</sup>	Median
(1,56)	1:C:696:ARG:N	1:C:696:ARG:CA	1:C:696:ARG:C	1:C:697:ILE:N	15	3.36	0.59	3.4
(1,57)	1:C:696:ARG:C	1:C:697:ILE:N	1:C:697:ILE:CA	1:C:697:ILE:C	14	3.34	0.45	3.3
(1,117)	1:C:749:VAL:C	1:C:750:ASN:N	1:C:750:ASN:CA	1:C:750:ASN:C	14	1.74	0.37	1.6
(1,83)	1:C:709:ARG:C	1:C:710:GLN:N	1:C:710:GLN:CA	1:C:710:GLN:C	12	1.84	0.6	1.65
(1,79)	1:C:707:ARG:C	1:C:708:VAL:N	1:C:708:VAL:CA	1:C:708:VAL:C	11	1.63	0.33	1.6
(1,171)	1:C:781:ILE:C	1:C:782:VAL:N	1:C:782:VAL:CA	1:C:782:VAL:C	10	1.95	0.34	1.9
(1,116)	1:C:749:VAL:N	1:C:749:VAL:CA	1:C:749:VAL:C	1:C:750:ASN:N	10	1.59	0.33	1.65
(1,55)	1:C:695:LEU:C	1:C:696:ARG:N	1:C:696:ARG:CA	1:C:696:ARG:C	9	1.31	0.17	1.3
(1,58)	1:C:697:ILE:N	1:C:697:ILE:CA	1:C:697:ILE:C	1:C:698:VAL:N	8	1.51	0.13	1.55
(1,172)	1:C:782:VAL:N	1:C:782:VAL:CA	1:C:782:VAL:C	1:C:783:GLU:N	8	1.34	0.36	1.2

<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 (°)
(1,57)	1:C:696:ARG:C	1:C:697:ILE:N	1:C:697:ILE:CA	1:C:697:ILE:C	10	4.3
(1,56)	1:C:696:ARG:N	1:C:696:ARG:CA	1:C:696:ARG:C	1:C:697:ILE:N	10	4.1
(1,57)	1:C:696:ARG:C	1:C:697:ILE:N	1:C:697:ILE:CA	1:C:697:ILE:C	15	3.9
(1,56)	1:C:696:ARG:N	1:C:696:ARG:CA	1:C:696:ARG:C	1:C:697:ILE:N	15	3.9
(1,56)	1:C:696:ARG:N	1:C:696:ARG:CA	1:C:696:ARG:C	1:C:697:ILE:N	6	3.8
(1,57)	1:C:696:ARG:C	1:C:697:ILE:N	1:C:697:ILE:CA	1:C:697:ILE:C	6	3.7
(1,56)	1:C:696:ARG:N	1:C:696:ARG:CA	1:C:696:ARG:C	1:C:697:ILE:N	7	3.7
(1,56)	1:C:696:ARG:N	1:C:696:ARG:CA	1:C:696:ARG:C	1:C:697:ILE:N	14	3.7
(1,57)	1:C:696:ARG:C	1:C:697:ILE:N	1:C:697:ILE:CA	1:C:697:ILE:C	13	3.6
(1,57)	1:C:696:ARG:C	1:C:697:ILE:N	1:C:697:ILE:CA	1:C:697:ILE:C	14	3.6