



# wwPDB NMR Structure Validation Summary Report ⓘ

Apr 6, 2026 – 11:06 AM UTC

PDB ID : 9ILL / pdb\_00009ill  
BMRB ID : 51661  
Title : monomeric SarA-E89Q in complex with DNA  
Authors : Xia, B.; Fu, D.H.  
Deposited on : 2024-06-30

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 : 4-5-2 with Phenix2.0  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
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) : 2.49

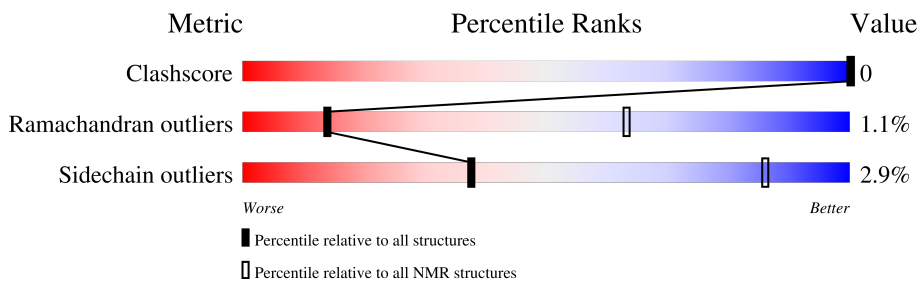
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 67%.

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	229148	14424
Ramachandran outliers	224038	12848
Sidechain outliers	223484	12823

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	114	
2	B	18	
2	C	18	

## 2 Ensemble composition and analysis i

This entry contains 20 models. Model 8 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:4-A:97 (94)	0.42	8

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

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

The models can be grouped into 4 clusters and 7 single-model clusters were found.

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

### 3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3111 atoms, of which 1407 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Transcriptional regulator SarA.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	114	1970	618	998	174	179	1	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP Q7A732
A	71	GLN	GLU	engineered mutation	UNP Q7A732
A	107	LEU	-	expression tag	UNP Q7A732
A	108	GLU	-	expression tag	UNP Q7A732
A	109	HIS	-	expression tag	UNP Q7A732
A	110	HIS	-	expression tag	UNP Q7A732
A	111	HIS	-	expression tag	UNP Q7A732
A	112	HIS	-	expression tag	UNP Q7A732
A	113	HIS	-	expression tag	UNP Q7A732
A	114	HIS	-	expression tag	UNP Q7A732

- Molecule 2 is a DNA chain called DNA.

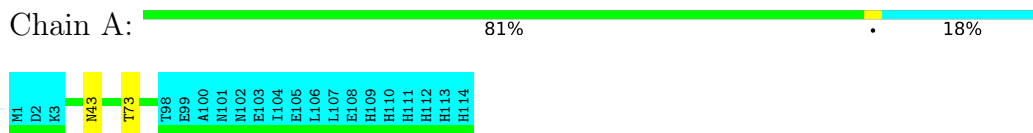
Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
2	B	18	570	176	204	67	106	17	0
2	C	18	571	176	205	67	106	17	0

## 4 Residue-property plots

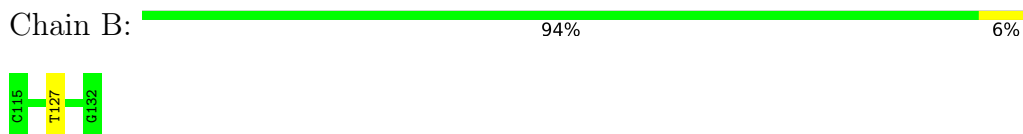
### 4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

- Molecule 1: Transcriptional regulator SarA



- Molecule 2: DNA



- Molecule 2: DNA

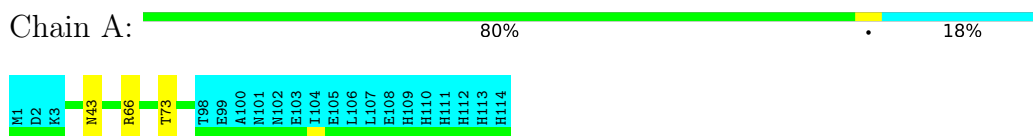


There are no outlier residues in this chain.

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

The representative model is number 8. Colouring as in section 4.1 above.

- Molecule 1: Transcriptional regulator SarA



- Molecule 2: DNA





- Molecule 2: DNA

Chain C:  100%

There are no outlier residues in this chain.

## 5 Refinement protocol and experimental data overview

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

Of the 100 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
Amber	structure calculation	
DYANA	structure calculation	
Amber	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	1635
Number of shifts mapped to atoms	1626
Number of unparsed shifts	0
Number of shifts with mapping errors	9
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	67%

## 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	0.86±0.01	0±0/811 ( 0.0± 0.0%)	1.40±0.02	1±1/1086 ( 0.1± 0.1%)
2	B	0.77±0.00	0±0/410 ( 0.0± 0.0%)	1.24±0.01	0±0/631 ( 0.0± 0.0%)
2	C	0.76±0.00	0±0/410 ( 0.0± 0.0%)	1.24±0.01	0±0/631 ( 0.0± 0.0%)
All	All	0.81	0/32620 ( 0.0%)	1.31	13/46960 ( 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	0.6±0.7
2	B	0.0±0.0	1.6±0.6
2	C	0.0±0.0	0.6±1.0
All	All	0	54

There are no bond-length outliers.

All 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	96	ARG	NE-CZ-NH2	6.03	124.63	119.20	5	1
1	A	66	ARG	NE-CZ-NH2	5.68	124.31	119.20	5	2
1	A	74	VAL	N-CA-C	5.59	115.99	108.17	2	7
1	A	83	ARG	NE-CZ-NH2	5.18	123.87	119.20	19	2
1	A	48	GLN	OE1-CD-NE2	-5.13	117.47	122.60	16	1

There are no chirality outliers.

5 of 13 unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
2	B	127	DT	Sidechain	16
1	A	72	ARG	Sidechain	7
2	B	116	DG	Sidechain	7
2	B	129	DC	Sidechain	4
2	C	138	DA	Sidechain	4

## 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	799	841	839	0±0
2	C	366	205	205	0±0
All	All	30620	25000	24980	1

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 0.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:36:LYS:HE3	2:C:139:DT:H3'	0.41	1.92	1	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	94/114 (82%)	92±1 (98±1%)	1±1 (1±1%)	1±0 (1±0%)	14	63
All	All	1880/2280 (82%)	1838 (98%)	22 (1%)	20 (1%)	14	63

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
1	A	43	ASN	20

### 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	91/110 (83%)	88±1 (97±2%)	3±1 (3±2%)	38 86
All	All	1820/2200 (83%)	1767 (97%)	53 (3%)	38 86

5 of 16 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	73	THR	20
1	A	70	ASP	7
1	A	64	LYS	4
1	A	14	ILE	4
1	A	72	ARG	3

### 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 [i](#)

The completeness of assignment taking into account all chemical shift lists is 67% for the well-defined parts and 66% for the entire structure.

### 7.1 Chemical shift list 1

File name: working\_cs.cif

Chemical shift list name: *starch\_output*

#### 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	1635
Number of shifts mapped to atoms	1626
Number of unparsed shifts	0
Number of shifts with mapping errors	9
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	1

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 9) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	C	151	DCZ	H1'	5.774	.	1
1	C	151	DCZ	H2'1	1.934	.	.
1	C	151	DCZ	H2'2	2.384	.	.
1	C	151	DCZ	H3'	4.683	.	1
1	C	151	DCZ	H4'	4.057	.	1
1	C	151	DCZ	H5	5.908	.	1
1	C	151	DCZ	H5'1	3.727	.	.
1	C	151	DCZ	H5'2	3.701	.	.
1	C	151	DCZ	H6	7.618	.	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
$^{13}\text{C}_\alpha$	107	$-0.51 \pm 0.07$	Should be checked
$^{13}\text{C}_\beta$	106	$-0.07 \pm 0.07$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}'$	0	—	None (insufficient data)
$^{15}\text{N}$	104	$0.31 \pm 0.36$	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 67%, i.e. 1434 atoms were assigned a chemical shift out of a possible 2134. 0 out of 17 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	364/468 (78%)	181/187 (97%)	92/188 (49%)	91/93 (98%)
Sidechain	707/857 (82%)	473/550 (86%)	222/266 (83%)	12/41 (29%)
Aromatic	66/101 (65%)	38/48 (79%)	28/49 (57%)	0/4 (0%)
Sugar	239/432 (55%)	239/252 (95%)	0/180 (0%)	0/0 (—%)
Base	58/276 (21%)	58/168 (35%)	0/64 (0%)	0/44 (0%)
Overall	1434/2134 (67%)	989/1205 (82%)	342/747 (46%)	103/182 (57%)

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

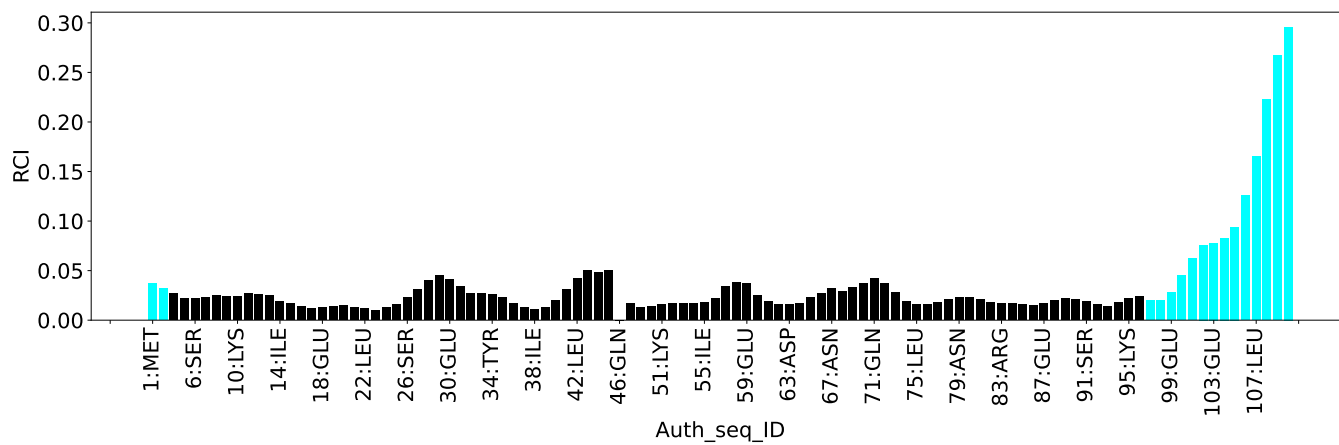
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	30	GLU	HB3	0.62	0.95 – 3.05	-6.5

### 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	5524
Intra-residue ( $ i-j =0$ )	2346
Sequential ( $ i-j =1$ )	1535
Medium range ( $ i-j >1$ and $ i-j <5$ )	468
Long range ( $ i-j \geq 5$ )	648
Inter-chain	297
Hydrogen bond restraints	230
Disulfide bond restraints	0
Total dihedral-angle restraints	756
Number of unmapped restraints	0
Number of restraints per residue	41.9
Number of long range restraints per residue <sup>1</sup>	4.5

<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)	3.9	0.19
0.2-0.5 (Medium)	1.9	0.5
>0.5 (Large)	123.9	55.84

### 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)	1.7	9.38
10.0-20.0 (Medium)	0.6	19.35
>20.0 (Large)	13.6	165.98

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

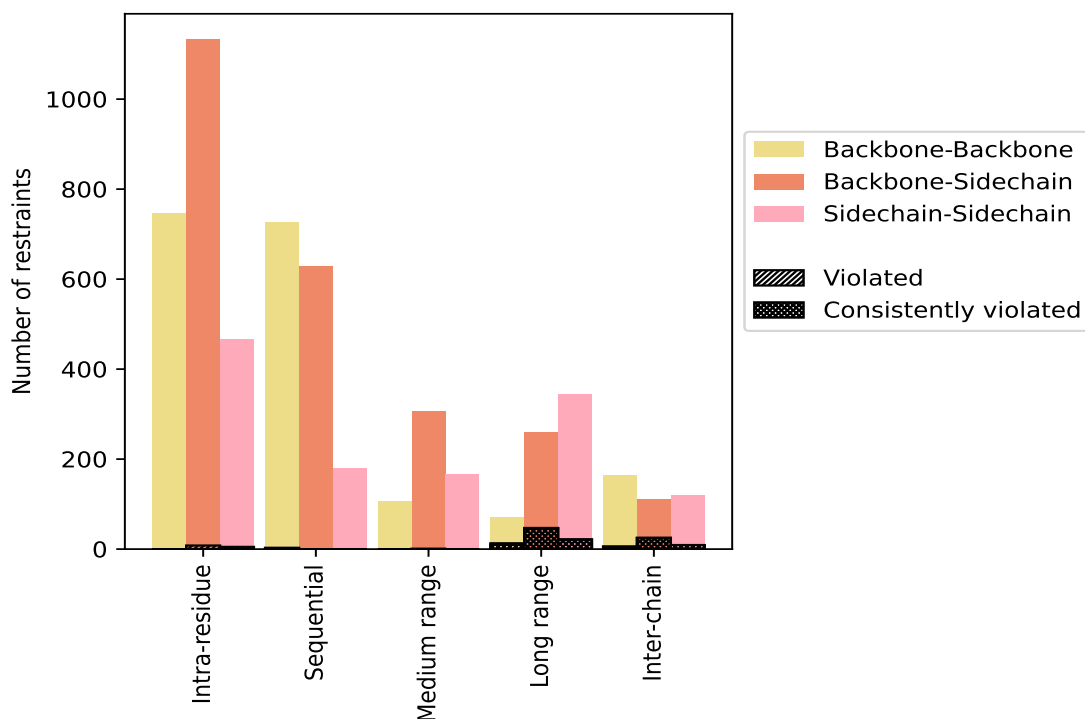
### 9.1 Summary of distance violations [i](#)

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>2346</b>	<b>42.5</b>	<b>13</b>	<b>0.6</b>	<b>0.2</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
Backbone-Backbone	746	13.5	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	1133	20.5	8	0.7	0.1	0	0.0	0.0
Sidechain-Sidechain	467	8.5	5	1.1	0.1	0	0.0	0.0
<b>Sequential (<math> i-j =1</math>)</b>	<b>1535</b>	<b>27.8</b>	<b>3</b>	<b>0.2</b>	<b>0.1</b>	<b>2</b>	<b>0.1</b>	<b>0.0</b>
Backbone-Backbone	727	13.2	3	0.4	0.1	2	0.3	0.0
Backbone-Sidechain	629	11.4	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	179	3.2	0	0.0	0.0	0	0.0	0.0
<b>Medium range (<math> i-j &gt;1</math> &amp; <math> i-j &lt;5</math>)</b>	<b>468</b>	<b>8.5</b>	<b>1</b>	<b>0.2</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
Backbone-Backbone	106	1.9	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	203	3.7	1	0.5	0.0	0	0.0	0.0
Sidechain-Sidechain	159	2.9	0	0.0	0.0	0	0.0	0.0
<b>Long range (<math> i-j \geq 5</math>)</b>	<b>648</b>	<b>11.7</b>	<b>73</b>	<b>11.3</b>	<b>1.3</b>	<b>67</b>	<b>10.3</b>	<b>1.2</b>
Backbone-Backbone	71	1.3	13	18.3	0.2	10	14.1	0.2
Backbone-Sidechain	237	4.3	38	16.0	0.7	37	15.6	0.7
Sidechain-Sidechain	340	6.2	22	6.5	0.4	20	5.9	0.4
<b>Inter-chain</b>	<b>297</b>	<b>5.4</b>	<b>29</b>	<b>9.8</b>	<b>0.5</b>	<b>28</b>	<b>9.4</b>	<b>0.5</b>
Backbone-Backbone	164	3.0	6	3.7	0.1	5	3.0	0.1
Backbone-Sidechain	103	1.9	19	18.4	0.3	19	18.4	0.3
Sidechain-Sidechain	30	0.5	4	13.3	0.1	4	13.3	0.1
<b>Hydrogen bond</b>	<b>230</b>	<b>4.2</b>	<b>20</b>	<b>8.7</b>	<b>0.4</b>	<b>20</b>	<b>8.7</b>	<b>0.4</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>5524</b>	<b>100.0</b>	<b>139</b>	<b>2.5</b>	<b>2.5</b>	<b>117</b>	<b>2.1</b>	<b>2.1</b>
Backbone-Backbone	1814	32.8	22	1.2	0.4	17	0.9	0.3
Backbone-Sidechain	2437	44.1	81	3.3	1.5	71	2.9	1.3
Sidechain-Sidechain	1273	23.0	36	2.8	0.7	29	2.3	0.5

<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	9	3	3	78	39	132	14.31	49.99	11.13	12.93
2	8	4	3	78	39	132	13.79	48.5	10.7	12.62
3	3	4	3	78	39	127	15.21	49.66	11.51	13.36
4	9	4	2	78	39	132	14.39	51.25	11.57	12.6
5	8	4	2	79	39	132	14.02	52.66	11.27	13.03
6	0	3	3	76	39	121	16.14	55.84	12.51	13.77
7	10	4	0	79	39	132	13.74	54.23	11.2	13.2
8	8	4	2	78	39	131	13.77	51.38	11.1	13.02
9	8	4	3	79	40	134	13.56	51.09	11.05	12.26
10	10	3	3	78	39	133	14.41	51.16	11.69	12.78

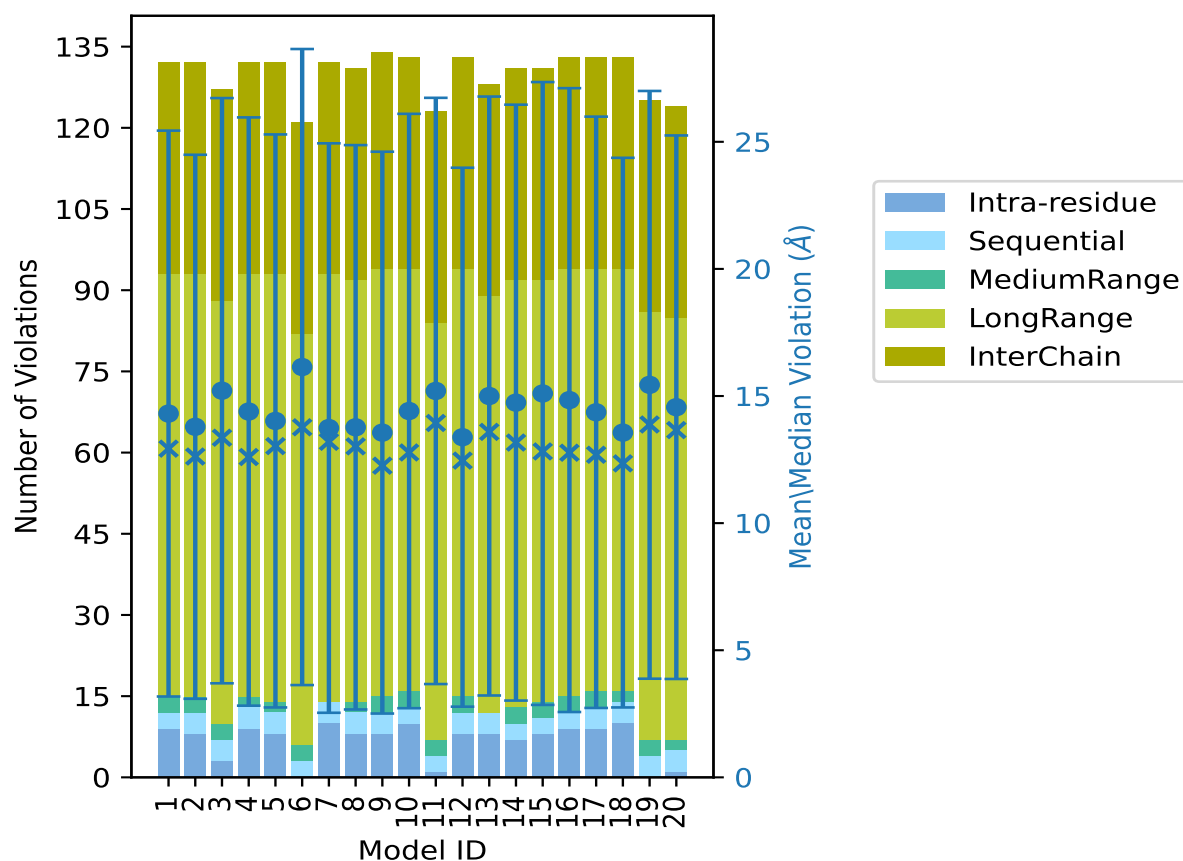
*Continued on next page...*

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	1	3	3	77	39	123	15.2	50.19	11.53	13.94
12	8	4	3	79	39	133	13.38	49.19	10.6	12.46
13	8	4	0	77	39	128	15.0	51.46	11.78	13.59
14	7	3	3	79	39	131	14.74	52.95	11.72	13.17
15	8	3	3	78	39	131	15.1	51.28	12.25	12.82
16	9	3	3	79	39	133	14.84	54.01	12.27	12.77
17	9	4	3	78	39	133	14.36	52.01	11.63	12.69
18	10	4	2	78	39	133	13.56	50.19	10.81	12.35
19	0	4	3	79	39	125	15.44	50.84	11.56	13.88
20	1	4	2	78	39	124	14.56	50.68	10.69	13.66

<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

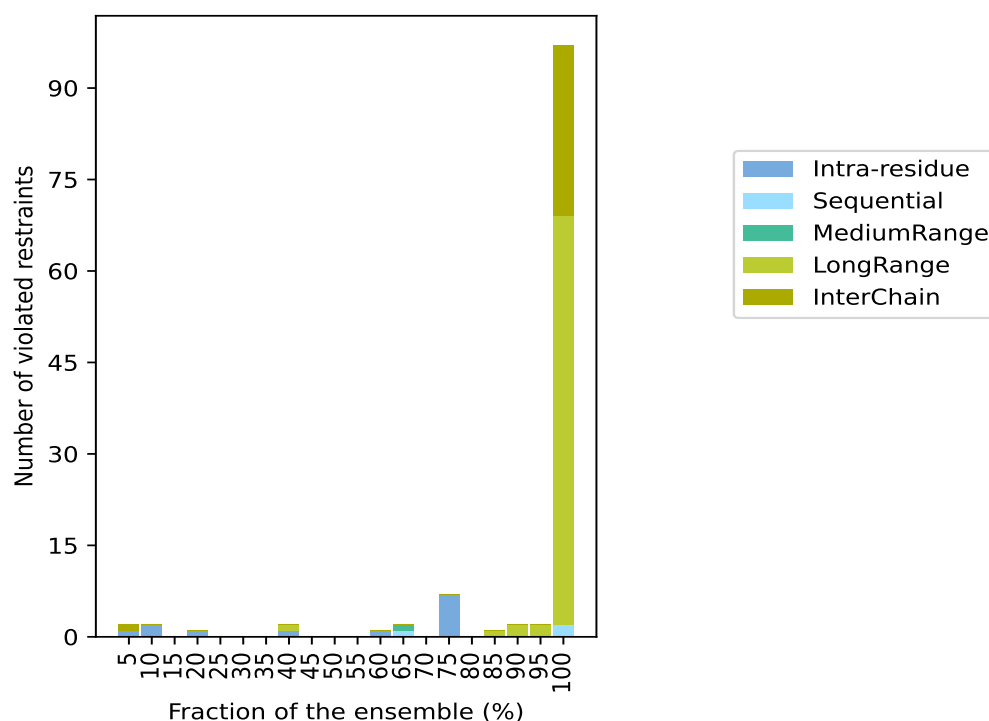
Violation analysis may find that some restraints are violated in few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of the ensemble. In total, 5175(IR:2333, SQ:1532, MR:467, LR:575, IC:268) restraints 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>	%
1	0	0	0	1	2	1	5.0
2	0	0	0	0	2	2	10.0
0	0	0	0	0	0	3	15.0
1	0	0	0	0	1	4	20.0
0	0	0	0	0	0	5	25.0
0	0	0	0	0	0	6	30.0
0	0	0	0	0	0	7	35.0
1	0	0	1	0	2	8	40.0
0	0	0	0	0	0	9	45.0
0	0	0	0	0	0	10	50.0
0	0	0	0	0	0	11	55.0
1	0	0	0	0	1	12	60.0
0	1	1	0	0	2	13	65.0
0	0	0	0	0	0	14	70.0
7	0	0	0	0	7	15	75.0
0	0	0	0	0	0	16	80.0
0	0	0	1	0	1	17	85.0
0	0	0	2	0	2	18	90.0
0	0	0	2	0	2	19	95.0
0	2	0	67	28	97	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

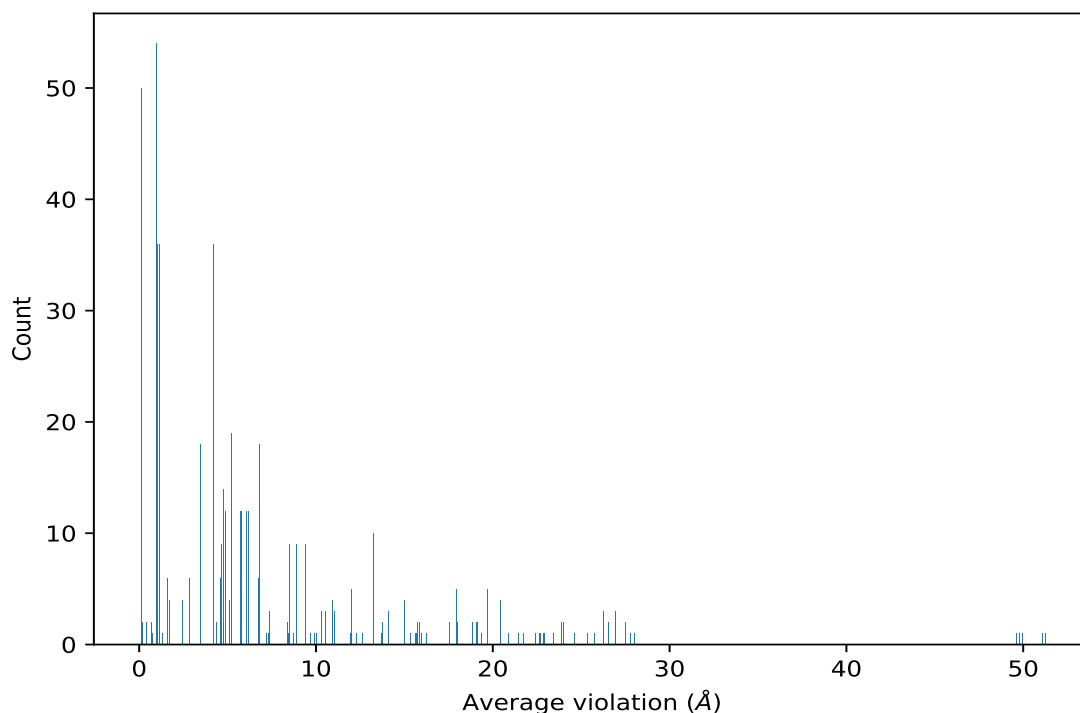
### 9.3.1 Bar graph : Distance violation statistics for the ensemble [\(i\)](#)



## 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.

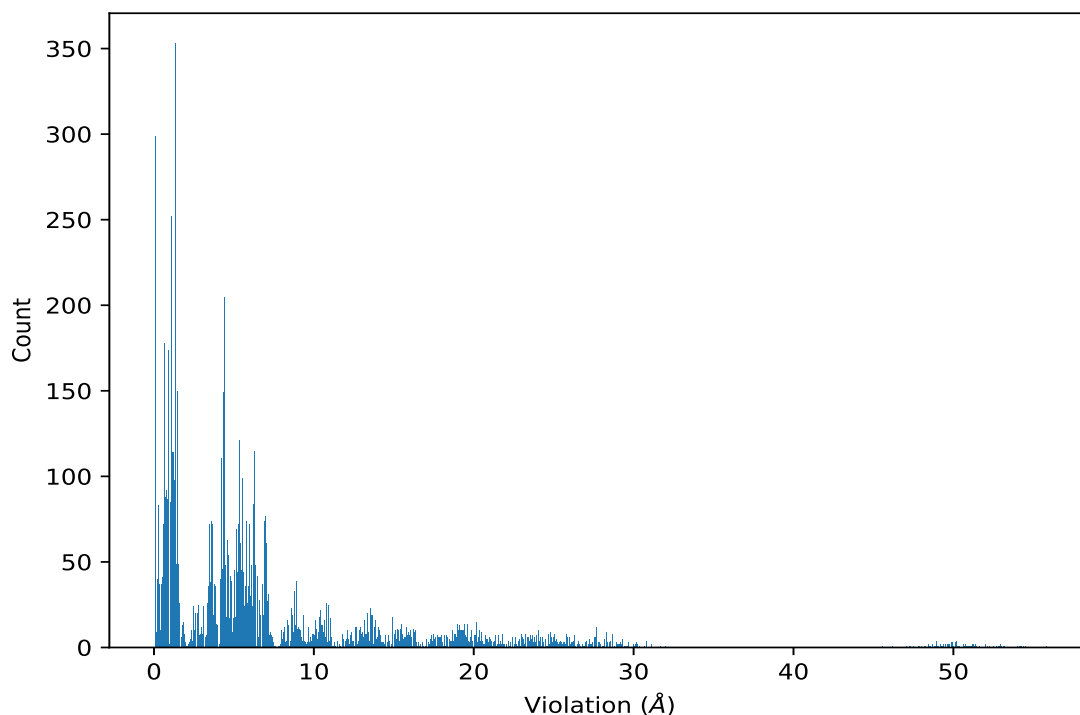
Key	Atom-1	Atom-2	Models <sup>1</sup>	Mean (Å)	SD <sup>1</sup> (Å)	Median (Å)
(2,40)	2:132:B:DG:H1	1:1:A:MET:N	20	51.25	1.93	51.2
(2,41)	2:132:B:DG:H21	1:1:A:MET:O	20	51.05	1.48	50.76
(2,42)	2:132:B:DG:N1	1:1:A:MET:N	20	49.95	1.97	49.88
(2,43)	2:132:B:DG:N2	1:1:A:MET:O	20	49.76	1.5	49.46
(2,44)	1:1:A:MET:N	2:132:B:DG:O6	20	49.62	2.08	49.46
(1,1250)	2:116:B:DG:H5 <sup>7</sup>	1:1:A:MET:H2	20	28.03	2.09	28.58
(1,1388)	2:116:B:DG:H5 <sup>7</sup>	1:1:A:MET:H2	20	27.76	2.09	28.32
(1,1460)	1:1:A:MET:H2	2:116:B:DG:H8	20	27.51	1.95	28.24
(1,1798)	2:116:B:DG:H8	1:1:A:MET:H2	20	27.51	1.95	28.24
(1,1269)	2:116:B:DG:H5 <sup>7</sup>	1:1:A:MET:H2	20	26.91	2.09	27.46

<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)
(2,40)	2:132:B:DG:H1	1:1:A:MET:N	6	55.84
(2,42)	2:132:B:DG:N1	1:1:A:MET:N	6	54.55
(2,44)	1:1:A:MET:N	2:132:B:DG:O6	6	54.4
(2,41)	2:132:B:DG:H21	1:1:A:MET:O	6	54.27
(2,40)	2:132:B:DG:H1	1:1:A:MET:N	7	54.23
(2,41)	2:132:B:DG:H21	1:1:A:MET:O	7	54.12
(2,40)	2:132:B:DG:H1	1:1:A:MET:N	16	54.01
(2,41)	2:132:B:DG:H21	1:1:A:MET:O	16	53.22
(2,44)	1:1:A:MET:N	2:132:B:DG:O6	7	53.14
(2,42)	2:132:B:DG:N1	1:1:A:MET:N	7	53.04

## 10 Dihedral-angle violation analysis

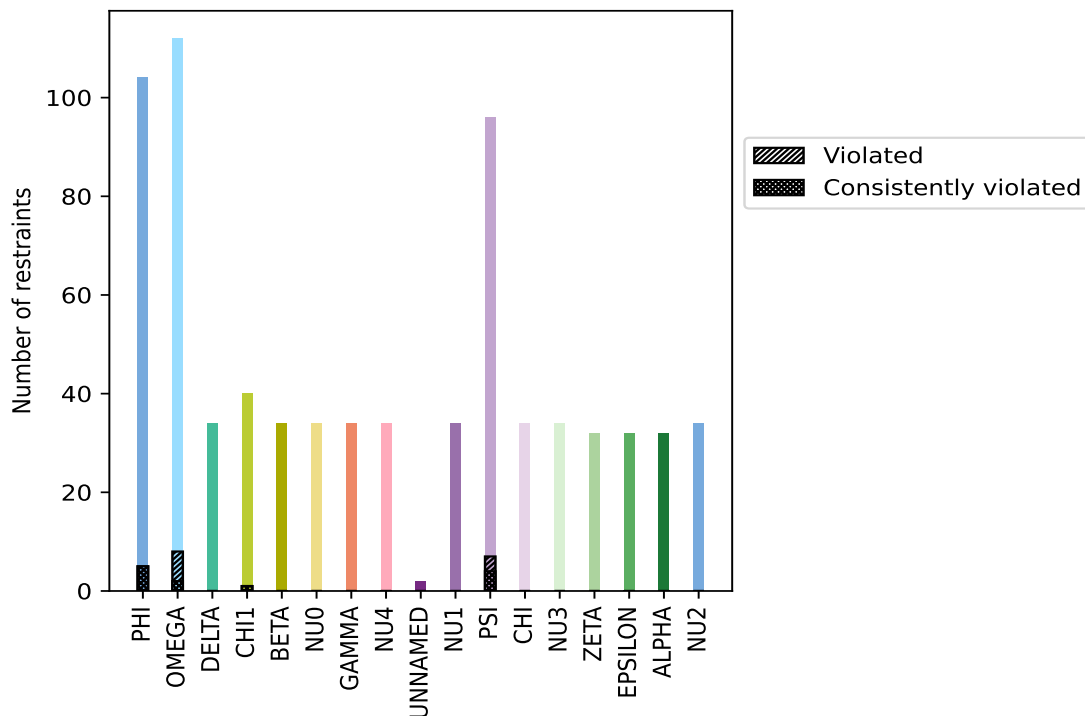
### 10.1 Summary of dihedral-angle violations

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	104	13.8	5	4.8	0.7	5	4.8	0.7
OMEGA	112	14.8	8	7.1	1.1	2	1.8	0.3
DELTA	34	4.5	0	0.0	0.0	0	0.0	0.0
CHI1	40	5.3	1	2.5	0.1	1	2.5	0.1
BETA	34	4.5	0	0.0	0.0	0	0.0	0.0
NU0	34	4.5	0	0.0	0.0	0	0.0	0.0
GAMMA	34	4.5	0	0.0	0.0	0	0.0	0.0
NU4	34	4.5	0	0.0	0.0	0	0.0	0.0
UNNAMED	2	0.3	0	0.0	0.0	0	0.0	0.0
NU1	34	4.5	0	0.0	0.0	0	0.0	0.0
PSI	96	12.7	7	7.3	0.9	4	4.2	0.5
CHI	34	4.5	0	0.0	0.0	0	0.0	0.0
NU3	34	4.5	0	0.0	0.0	0	0.0	0.0
ZETA	32	4.2	0	0.0	0.0	0	0.0	0.0
EPSILON	32	4.2	0	0.0	0.0	0	0.0	0.0
ALPHA	32	4.2	0	0.0	0.0	0	0.0	0.0
NU2	34	4.5	0	0.0	0.0	0	0.0	0.0
Total	756	100.0	21	2.8	2.8	12	1.6	1.6

<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](#)



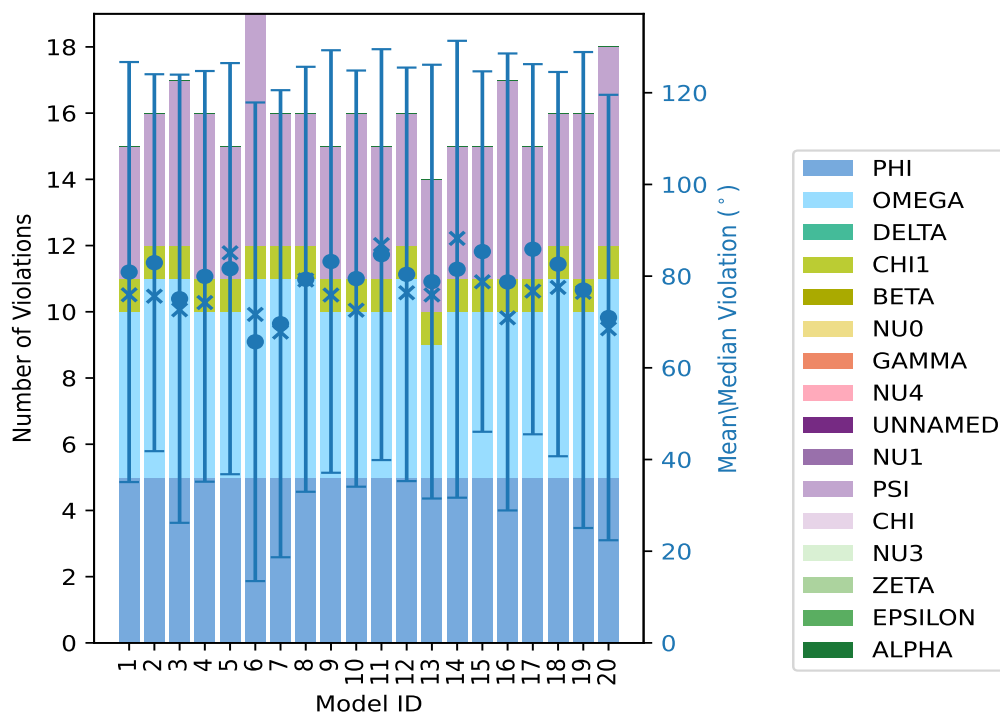
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

## 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															
	PHI	OMEGA	DELTA	CHI1	BETA	NU0	GAMMA	NU4	UNNAMED	NU1	PSI	CHI	NU3	ZETA	EPSILON	ALPHA
1	5	5	0	1	0	0	0	0	0	0	4	0	0	0	0	0
2	5	6	0	1	0	0	0	0	0	0	4	0	0	0	0	0
3	5	6	0	1	0	0	0	0	0	0	5	0	0	0	0	0
4	5	5	0	1	0	0	0	0	0	0	5	0	0	0	0	0
5	5	5	0	1	0	0	0	0	0	0	4	0	0	0	0	0
6	5	6	0	1	0	0	0	0	0	0	7	0	0	0	0	0
7	5	6	0	1	0	0	0	0	0	0	4	0	0	0	0	0
8	5	6	0	1	0	0	0	0	0	0	4	0	0	0	0	0
9	5	5	0	1	0	0	0	0	0	0	4	0	0	0	0	0
10	5	5	0	1	0	0	0	0	0	0	5	0	0	0	0	0
11	5	5	0	1	0	0	0	0	0	0	4	0	0	0	0	0
12	5	6	0	1	0	0	0	0	0	0	4	0	0	0	0	0
13	5	4	0	1	0	0	0	0	0	0	4	0	0	0	0	0
14	5	5	0	1	0	0	0	0	0	0	4	0	0	0	0	0
15	5	5	0	1	0	0	0	0	0	0	4	0	0	0	0	0
16	5	5	0	1	0	0	0	0	0	0	6	0	0	0	0	0
17	5	5	0	1	0	0	0	0	0	0	4	0	0	0	0	0
18	5	6	0	1	0	0	0	0	0	0	4	0	0	0	0	0
19	5	5	0	1	0	0	0	0	0	0	5	0	0	0	0	0
20	5	6	0	1	0	0	0	0	0	0	6	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.

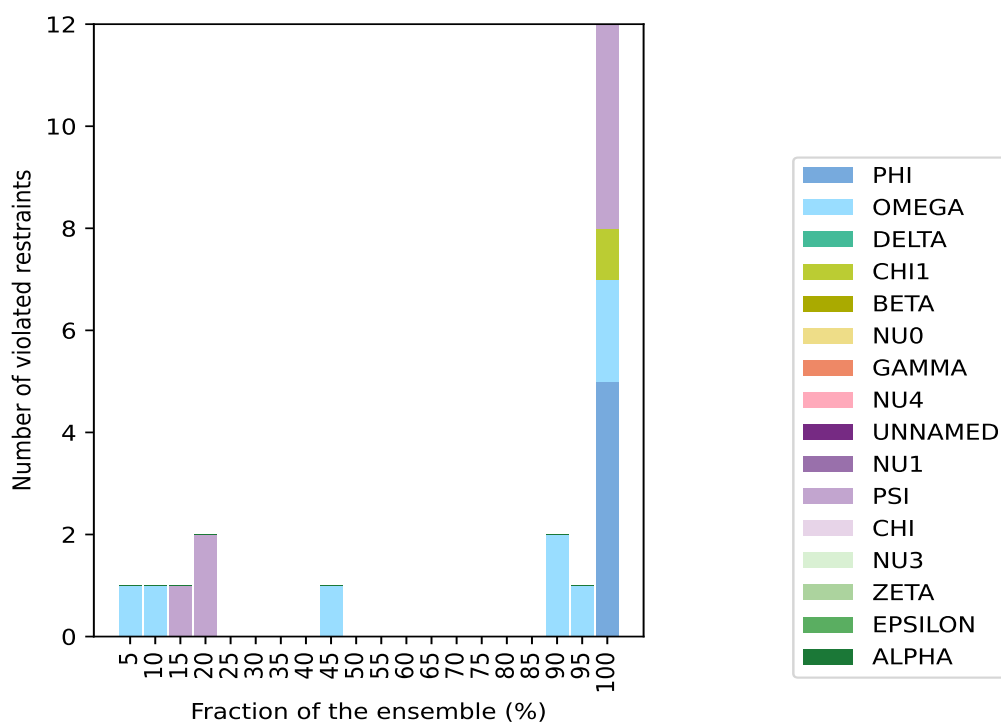
	Number of violated restraints													
	PHI	OMEGA	DELTA	CHI1	BETA	NU0	GAMMA	NU4	UNNAMED	NU1	PSI	CHI	NU3	ZETA
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	0	0	2	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Continued from previous page...

PHI	OMEGA	DELTA	CHI1	BETA	NU0	GAMMA	NU4	UNNAMED	NU1	Number of violated restraints			
										PSI	CHI	NU3	ZETA
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	2	0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	0	0
5	2	0	1	0	0	0	0	0	0	4	0	0	0

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

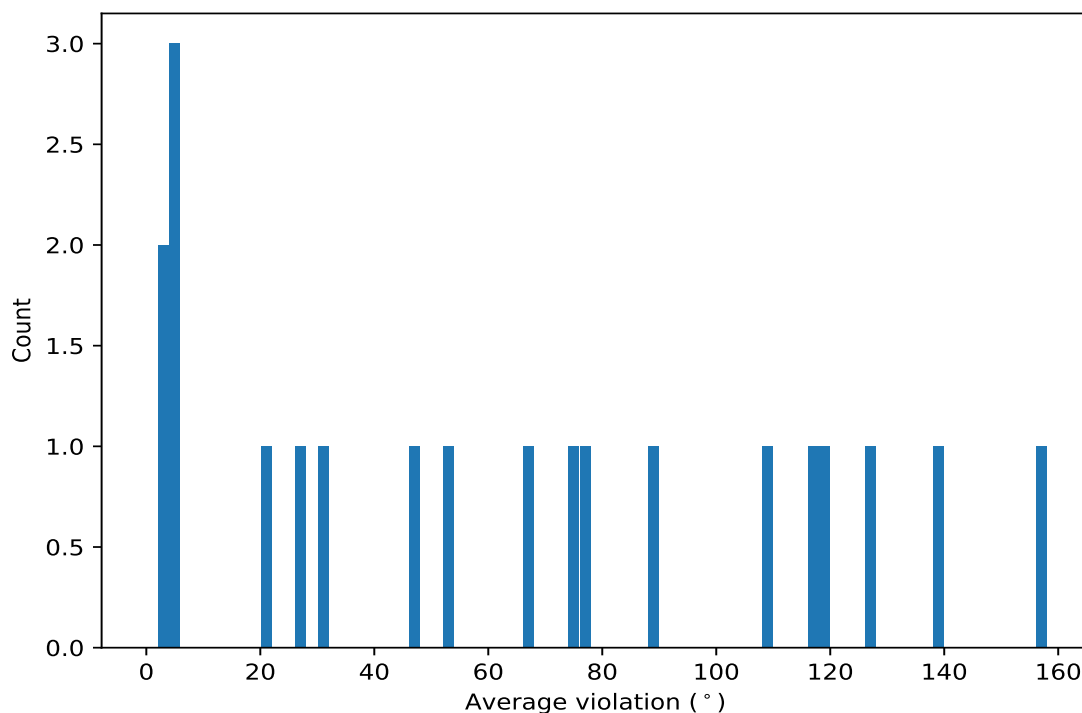
### 10.3.1 Bar graph : Dihedral-angle Violation statistics for the ensemble [i](#)



## 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.

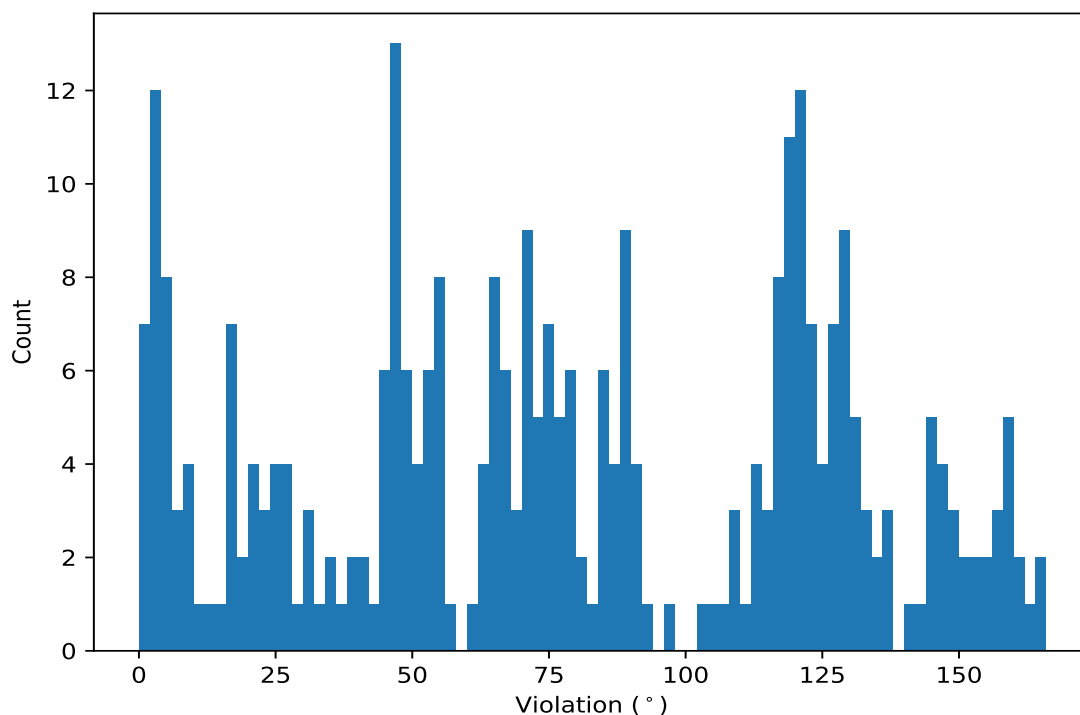
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models <sup>1</sup>	Mean	SD <sup>2</sup>	Median
(1,668)	1:23:A:THR:CA	1:23:A:THR:N	1:4:A:LEU:C	1:4:A:LEU:CA	20	157.27	4.76	158.07
(1,596)	1:87:A:GLU:C	1:106:A:LEU:N	1:106:A:LEU:CA	1:106:A:LEU:C	20	139.87	6.76	142.61
(1,699)	1:36:A:LYS:CA	1:36:A:LYS:N	1:53:A:VAL:C	1:53:A:VAL:CA	20	127.91	3.42	128.44
(1,444)	1:4:A:LEU:C	1:23:A:THR:N	1:23:A:THR:CA	1:23:A:THR:C	20	119.36	1.88	118.98
(1,595)	1:87:A:GLU:N	1:87:A:GLU:CA	1:87:A:GLU:C	1:106:A:LEU:N	20	117.2	6.07	117.0
(1,598)	1:89:A:LEU:C	1:108:A:GLU:N	1:108:A:GLU:CA	1:108:A:GLU:C	20	108.78	33.99	119.99
(1,502)	1:36:A:LYS:C	1:55:A:ILE:N	1:55:A:ILE:CA	1:55:A:ILE:C	20	88.25	2.01	88.56
(1,613)	1:4:A:LEU:N	1:4:A:LEU:CA	1:4:A:LEU:CB	1:4:A:LEU:CG	20	76.55	6.03	77.06
(1,443)	1:4:A:LEU:N	1:4:A:LEU:CA	1:4:A:LEU:C	1:23:A:THR:N	20	75.36	5.21	73.1
(1,442)	1:21:A:VAL:C	1:4:A:LEU:N	1:4:A:LEU:CA	1:4:A:LEU:C	20	66.48	3.03	65.72

<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 (°)
(1,668)	1:23:A:THR:CA	1:23:A:THR:N	1:4:A:LEU:C	1:4:A:LEU:CA	20	165.98
(1,668)	1:23:A:THR:CA	1:23:A:THR:N	1:4:A:LEU:C	1:4:A:LEU:CA	5	164.62
(1,668)	1:23:A:THR:CA	1:23:A:THR:N	1:4:A:LEU:C	1:4:A:LEU:CA	12	163.27
(1,668)	1:23:A:THR:CA	1:23:A:THR:N	1:4:A:LEU:C	1:4:A:LEU:CA	19	160.63
(1,668)	1:23:A:THR:CA	1:23:A:THR:N	1:4:A:LEU:C	1:4:A:LEU:CA	3	160.37
(1,668)	1:23:A:THR:CA	1:23:A:THR:N	1:4:A:LEU:C	1:4:A:LEU:CA	9	159.59
(1,668)	1:23:A:THR:CA	1:23:A:THR:N	1:4:A:LEU:C	1:4:A:LEU:CA	4	159.44
(1,668)	1:23:A:THR:CA	1:23:A:THR:N	1:4:A:LEU:C	1:4:A:LEU:CA	18	159.06
(1,668)	1:23:A:THR:CA	1:23:A:THR:N	1:4:A:LEU:C	1:4:A:LEU:CA	17	158.92
(1,668)	1:23:A:THR:CA	1:23:A:THR:N	1:4:A:LEU:C	1:4:A:LEU:CA	13	158.89