



# wwPDB NMR Structure Validation Summary Report ⓘ

Dec 25, 2024 – 07:18 PM EST

PDB ID : 8HT7  
BMRB ID : 36531  
Title : The N-terminal region of Cdc6 specifically recognizes human DNA G-quadruplex  
Authors : Liu, C.; Zhu, G.; Geng, Y.; Xu, N.  
Deposited on : 2022-12-20

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

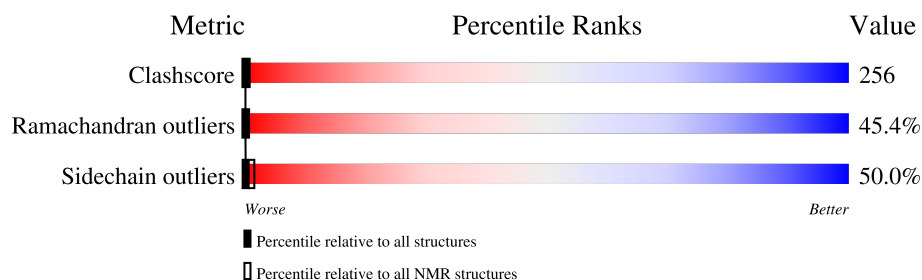
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 48%.

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	21	
2	B	15	

## 2 Ensemble composition and analysis ⓘ

This entry contains 10 models.

Cyrange was unable to find well-defined residues.

Error message: The number of core atoms ( 5) was below the domain threshold value ( 8).

The models can be grouped into 4 clusters and 1 single-model cluster was found.

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

### 3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 940 atoms, of which 372 are hydrogens and 0 are deuteriums.

- Molecule 1 is a DNA chain called DNA (5'-D(\*GP\*GP\*GP\*TP\*TP\*AP\*GP\*GP\*GP\*TP\*TP\*AP\*GP\*GP\*GP\*TP\*TP\*TP\*GP\*GP\*G)-3').

Mol	Chain	Residues	Atoms						Trace
1	A	21	Total	C	H	N	O	P	0
			683	210	240	84	129	20	

- Molecule 2 is a protein called GLN-ALA-GLN-ALA-THR-ILE-SER-PHE-PRO-LYS-ARG-LYS-LEU-SER-TRP.

Mol	Chain	Residues	Atoms					Trace
2	B	15	Total	C	H	N	O	0
			257	81	132	23	21	

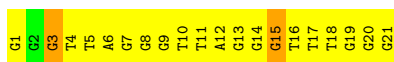
## 4 Residue-property plots [i](#)

### 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: DNA (5'-D(\*GP\*GP\*GP\*TP\*TP\*AP\*GP\*GP\*GP\*TP\*TP\*AP\*GP\*GP\*GP\*TP\*TP\*TP\*GP\*GP\*G)-3')

Chain A: 



- Molecule 2: GLN-ALA-GLN-ALA-THR-ILE-SER-PHE-PRO-LYS-ARG-LYS-LEU-SER-TRP

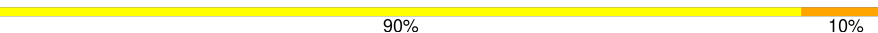
Chain B: 

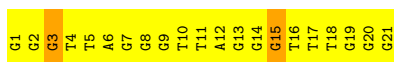


### 4.2 Residue scores for the representative (author defined) model from the NMR ensemble

The representative model is number 1. Colouring as in section 4.1 above.

- Molecule 1: DNA (5'-D(\*GP\*GP\*GP\*TP\*TP\*AP\*GP\*GP\*GP\*TP\*TP\*AP\*GP\*GP\*GP\*TP\*TP\*TP\*GP\*GP\*G)-3')

Chain A: 



- Molecule 2: GLN-ALA-GLN-ALA-THR-ILE-SER-PHE-PRO-LYS-ARG-LYS-LEU-SER-TRP

Chain B: 

G107
T111
I112
S113
F114
P115
K116
R117
K118
L119
S120
W121

## 5 Refinement protocol and experimental data overview

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

Of the 200 calculated structures, 10 were deposited, based on the following criterion: *structures with the lowest energy*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
X-PLOR NIH	refinement	
X-PLOR NIH	structure 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	345
Number of shifts mapped to atoms	345
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	48%

## 6 Model quality [i](#)

### 6.1 Standard geometry [i](#)

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

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

Mol	Chain	Chirality	Planarity
1	A	0.0±0.0	2.5±0.7
All	All	0	25

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

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

Mol	Chain	Res	Type	Group	Models (Total)
1	A	3	DG	Sidechain	10
1	A	15	DG	Sidechain	8
1	A	9	DG	Sidechain	4
1	A	6	DA	Sidechain	3

### 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	443	240	240	196±11
2	B	125	132	129	91±10
All	All	5680	3720	3690	2401

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 256.

5 of 951 unique clashes are listed below, sorted by their clash magnitude.



Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:16:DT:C5	1:A:18:DT:H72	1.27	1.61	1	2
1:A:16:DT:H72	1:A:18:DT:C4	1.12	1.80	5	6
1:A:16:DT:C4	1:A:18:DT:H72	1.11	1.81	1	1
1:A:16:DT:C5	1:A:18:DT:C7	1.10	2.33	1	3
1:A:4:DT:H73	1:A:6:DA:N7	1.10	1.62	2	8

## 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	
2	B	13/15 (87%)	3±1 (22±8%)	4±1 (32±11%)	6±1 (45±9%)	0	0
All	All	130/150 (87%)	29 (22%)	42 (32%)	59 (45%)	0	0

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

Mol	Chain	Res	Type	Models (Total)
2	B	114	PHE	7
2	B	115	PRO	7
2	B	120	SER	7
2	B	118	LYS	7
2	B	112	ILE	6

### 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	
2	B	13/13 (100%)	6±1 (50±8%)	6±1 (50±8%)	0	1
All	All	130/130 (100%)	65 (50%)	65 (50%)	0	1

5 of 10 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
2	B	116	LYS	10
2	B	118	LYS	10
2	B	119	LEU	10
2	B	114	PHE	9
2	B	111	THR	8

### 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 48% for the well-defined parts and 48% 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

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	345
Number of shifts mapped to atoms	345
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	1

The following errors were found when reading this chemical shift list.

- Chemical shift has been reported more than once. First 5 (of 0) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	B	108	ALA	HB2	1.300	.	.
1	B	108	ALA	HB3	1.300	.	.
1	B	110	ALA	HB2	1.315	.	.
1	B	110	ALA	HB3	1.315	.	.
1	B	111	THR	HG22	1.081	.	.
1	B	111	THR	HG23	1.081	.	.
1	B	112	ILE	HD12	0.713	.	.
1	B	112	ILE	HD13	0.713	.	.
1	B	112	ILE	HG22	0.647	.	.
1	B	112	ILE	HG23	0.647	.	.
1	B	119	LEU	HD12	0.744	.	.
1	B	119	LEU	HD13	0.744	.	.
1	B	119	LEU	HD22	0.687	.	.
1	B	119	LEU	HD23	0.687	.	.

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

No chemical shift referencing corrections were calculated (not enough data).

### 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 48%, i.e. 309 atoms were assigned a chemical shift out of a possible 638. 0 out of 1 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Backbone	57/73 (78%)	28/29 (97%)	15/30 (50%)	14/14 (100%)
Sidechain	106/125 (85%)	72/81 (89%)	34/37 (92%)	0/7 (0%)
Aromatic	7/22 (32%)	7/11 (64%)	0/10 (0%)	0/1 (0%)
Sugar	115/252 (46%)	115/147 (78%)	0/105 (0%)	0/0 (—%)
Base	24/166 (14%)	24/103 (23%)	0/30 (0%)	0/33 (0%)
Overall	309/638 (48%)	246/371 (66%)	49/212 (23%)	14/55 (25%)

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

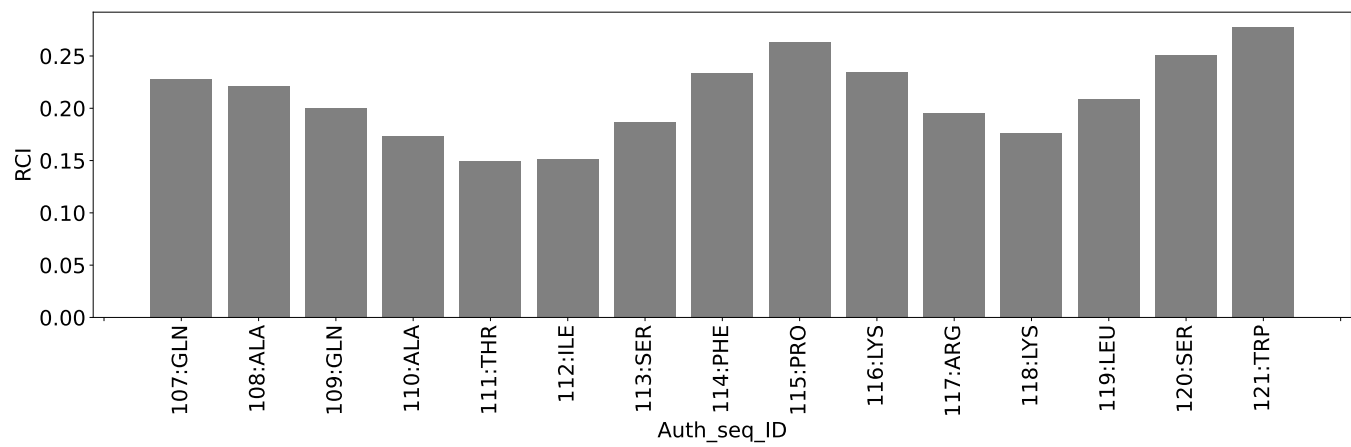
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	13	DG	H4'	3.25	3.37 – 5.35	-5.6

### 7.1.5 Random Coil Index (RCI) plots [i](#)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition. If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain B:



## 8 NMR restraints analysis [i](#)

### 8.1 Conformationally restricting restraints [i](#)

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	495
Intra-residue ( $ i-j =0$ )	190
Sequential ( $ i-j =1$ )	174
Medium range ( $ i-j >1$ and $ i-j <5$ )	35
Long range ( $ i-j \geq 5$ )	32
Inter-chain	12
Hydrogen bond restraints	52
Disulfide bond restraints	0
Total dihedral-angle restraints	58
Number of unmapped restraints	0
Number of restraints per residue	15.4
Number of long range restraints per residue <sup>1</sup>	2.0

<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 [i](#)

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 [i](#)

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)	6.0	0.2
0.2-0.5 (Medium)	4.9	0.5
>0.5 (Large)	24.3	6.17

### 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)	2.9	3.34
10.0-20.0 (Medium)	None	None
>20.0 (Large)	None	None

## 9 Distance violation analysis ⓘ

### 9.1 Summary of distance violations ⓘ

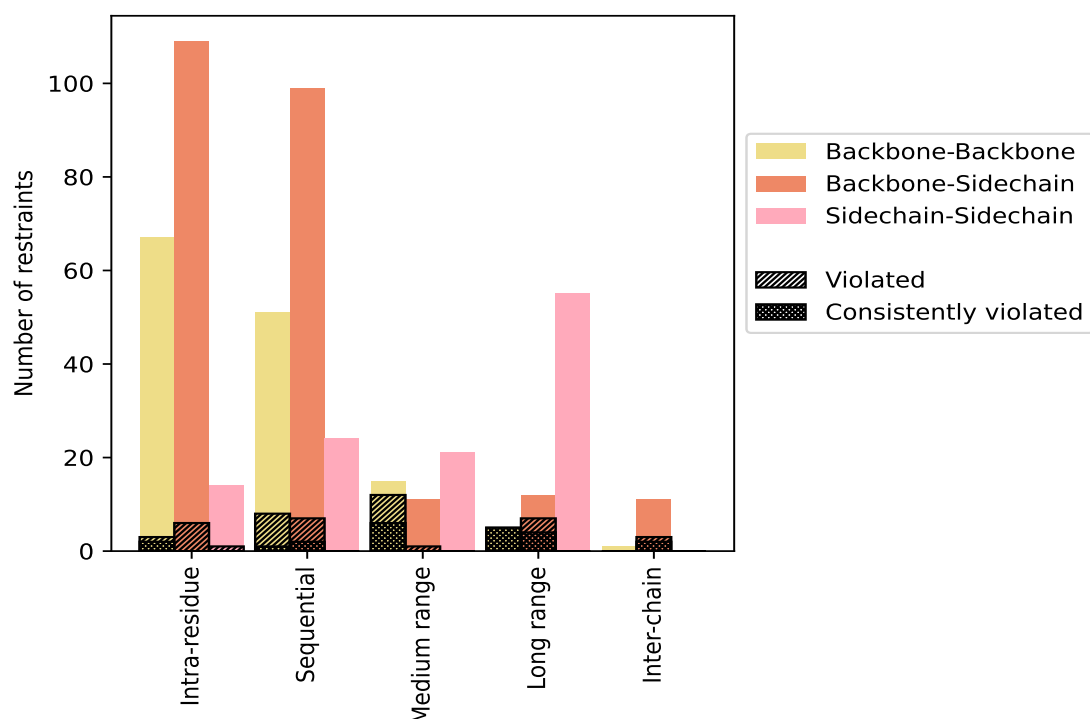
The following table shows the summary of distance violations in different restraint categories based on the sequence separation of the atoms involved. Each category is further sub-divided into three sub-categories based on the atoms involved. Violations less than 0.1 Å are not included in the statistics.

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>
<a href="#">Intra-residue ( i-j =0)</a>	<a href="#">190</a>	<a href="#">38.4</a>	<a href="#">10</a>	<a href="#">5.3</a>	<a href="#">2.0</a>	<a href="#">2</a>	<a href="#">1.1</a>	<a href="#">0.4</a>
Backbone-Backbone	67	13.5	3	4.5	0.6	2	3.0	0.4
Backbone-Sidechain	109	22.0	6	5.5	1.2	0	0.0	0.0
Sidechain-Sidechain	14	2.8	1	7.1	0.2	0	0.0	0.0
<a href="#">Sequential ( i-j =1)</a>	<a href="#">174</a>	<a href="#">35.2</a>	<a href="#">15</a>	<a href="#">8.6</a>	<a href="#">3.0</a>	<a href="#">3</a>	<a href="#">1.7</a>	<a href="#">0.6</a>
Backbone-Backbone	51	10.3	8	15.7	1.6	1	2.0	0.2
Backbone-Sidechain	99	20.0	7	7.1	1.4	2	2.0	0.4
Sidechain-Sidechain	24	4.8	0	0.0	0.0	0	0.0	0.0
<a href="#">Medium range ( i-j &gt;1 &amp;  i-j &lt;5)</a>	<a href="#">35</a>	<a href="#">7.1</a>	<a href="#">13</a>	<a href="#">37.1</a>	<a href="#">2.6</a>	<a href="#">6</a>	<a href="#">17.1</a>	<a href="#">1.2</a>
Backbone-Backbone	15	3.0	12	80.0	2.4	6	40.0	1.2
Backbone-Sidechain	11	2.2	1	9.1	0.2	0	0.0	0.0
Sidechain-Sidechain	9	1.8	0	0.0	0.0	0	0.0	0.0
<a href="#">Long range ( i-j ≥5)</a>	<a href="#">32</a>	<a href="#">6.5</a>	<a href="#">12</a>	<a href="#">37.5</a>	<a href="#">2.4</a>	<a href="#">9</a>	<a href="#">28.1</a>	<a href="#">1.8</a>
Backbone-Backbone	5	1.0	5	100.0	1.0	5	100.0	1.0
Backbone-Sidechain	12	2.4	7	58.3	1.4	4	33.3	0.8
Sidechain-Sidechain	15	3.0	0	0.0	0.0	0	0.0	0.0
<a href="#">Inter-chain</a>	<a href="#">12</a>	<a href="#">2.4</a>	<a href="#">3</a>	<a href="#">25.0</a>	<a href="#">0.6</a>	<a href="#">2</a>	<a href="#">16.7</a>	<a href="#">0.4</a>
Backbone-Backbone	1	0.2	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	11	2.2	3	27.3	0.6	2	18.2	0.4
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
<a href="#">Hydrogen bond</a>	<a href="#">52</a>	<a href="#">10.5</a>	<a href="#">0</a>	<a href="#">0.0</a>	<a href="#">0.0</a>	<a href="#">0</a>	<a href="#">0.0</a>	<a href="#">0.0</a>
<a href="#">Disulfide bond</a>	<a href="#">0</a>	<a href="#">0.0</a>	<a href="#">0</a>	<a href="#">0.0</a>	<a href="#">0.0</a>	<a href="#">0</a>	<a href="#">0.0</a>	<a href="#">0.0</a>
<a href="#">Total</a>	<a href="#">495</a>	<a href="#">100.0</a>	<a href="#">53</a>	<a href="#">10.7</a>	<a href="#">10.7</a>	<a href="#">22</a>	<a href="#">4.4</a>	<a href="#">4.4</a>
Backbone-Backbone	139	28.1	28	20.1	5.7	14	10.1	2.8
Backbone-Sidechain	242	48.9	24	9.9	4.8	8	3.3	1.6
Sidechain-Sidechain	114	23.0	1	0.9	0.2	0	0.0	0.0

<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

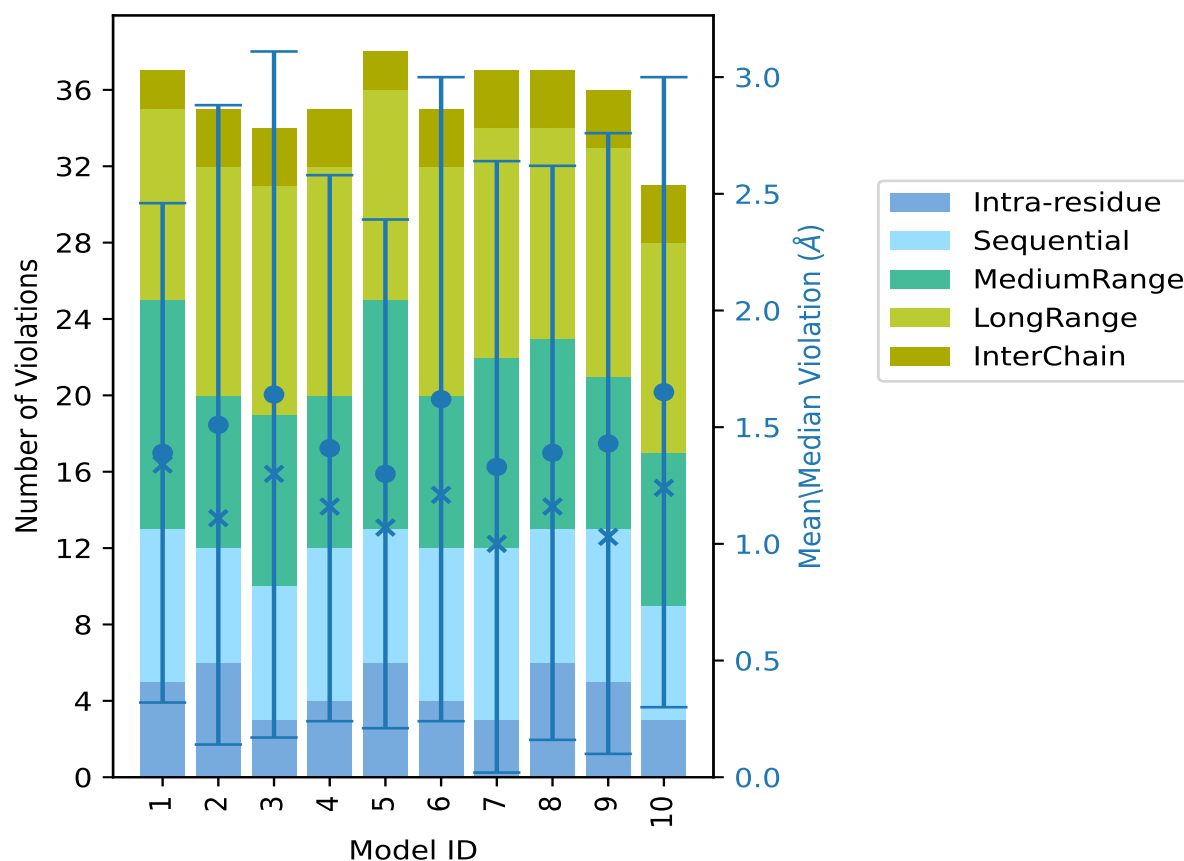
## 9.2 Distance violation statistics for each model [i](#)

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	5	8	12	10	2	37	1.39	4.34	1.07	1.34
2	6	6	8	12	3	35	1.51	5.72	1.37	1.11
3	3	7	9	12	3	34	1.64	6.17	1.47	1.3
4	4	8	8	12	3	35	1.41	4.72	1.17	1.16
5	6	7	12	11	2	38	1.3	4.6	1.09	1.07
6	4	8	8	12	3	35	1.62	5.9	1.38	1.21
7	3	9	10	12	3	37	1.33	5.8	1.31	1.0
8	6	7	10	11	3	37	1.39	5.29	1.23	1.16
9	5	8	8	12	3	36	1.43	5.48	1.33	1.03
10	3	6	8	11	3	31	1.65	5.77	1.35	1.24

<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, 390(IR:180, SQ:159, MR:22, LR:20, 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>	%
3	4	0	0	0	7	1	10.0
1	1	1	0	0	3	2	20.0
2	2	1	0	0	5	3	30.0

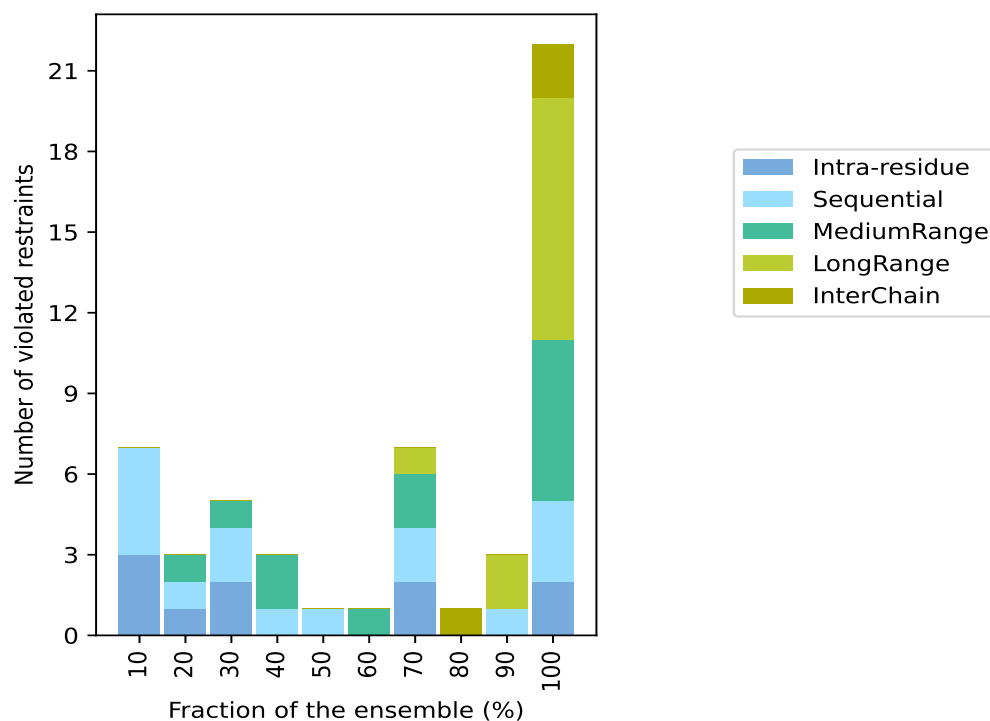
*Continued on next page...*

Continued from previous page...

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>	%
0	1	2	0	0	3	4	40.0
0	1	0	0	0	1	5	50.0
0	0	1	0	0	1	6	60.0
2	2	2	1	0	7	7	70.0
0	0	0	0	1	1	8	80.0
0	1	0	2	0	3	9	90.0
2	3	6	9	2	22	10	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 ⓘ

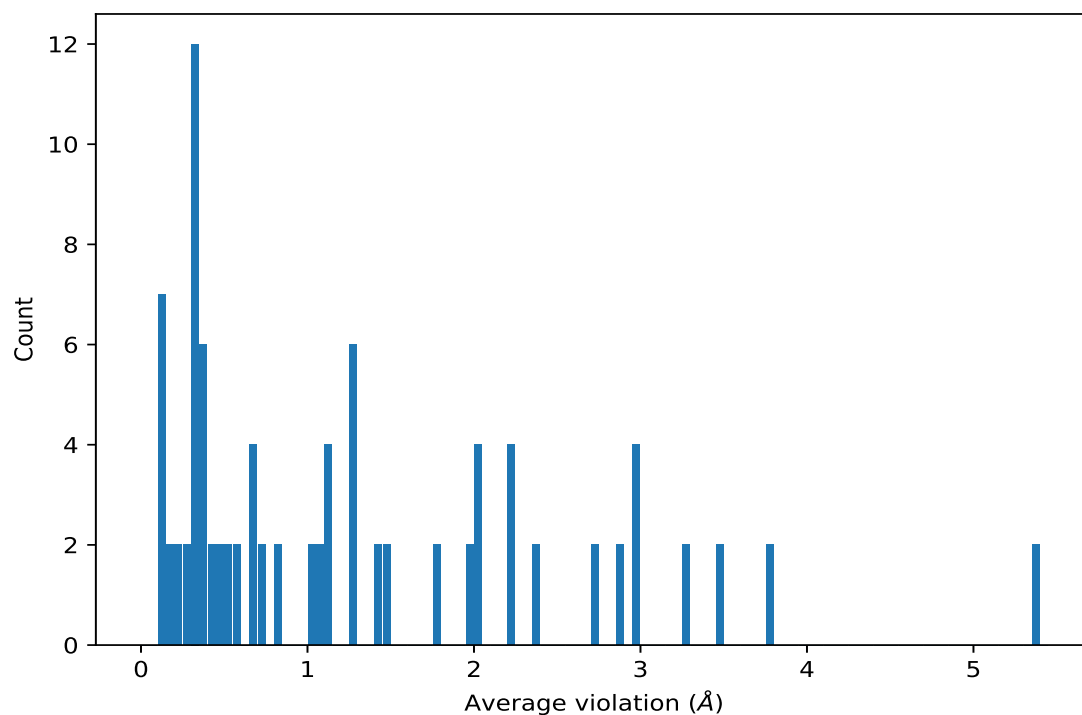


## 9.4 Most violated distance restraints in the ensemble ⓘ

### 9.4.1 Histogram : Distribution of mean distance violations ⓘ

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

in the ensemble



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

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

Key	Atom-1	Atom-2	Models <sup>1</sup>	Mean (Å)	SD <sup>1</sup> (Å)	Median (Å)
(1,66)	1:5:A:DT:H5'	1:18:A:DT:H1'	10	5.38	0.59	5.6
(1,66)	1:5:A:DT:H5''	1:18:A:DT:H1'	10	5.38	0.59	5.6
(1,39)	1:5:A:DT:H1'	1:17:A:DT:H5'	10	3.77	0.35	3.8
(1,39)	1:5:A:DT:H1'	1:17:A:DT:H5''	10	3.77	0.35	3.8
(1,67)	1:5:A:DT:H5'	1:18:A:DT:H5''	10	3.47	0.51	3.54
(1,67)	1:5:A:DT:H5''	1:18:A:DT:H5''	10	3.47	0.51	3.54
(1,64)	1:5:A:DT:H5'	1:17:A:DT:H2''	10	3.25	0.42	3.33
(1,64)	1:5:A:DT:H5''	1:17:A:DT:H2''	10	3.25	0.42	3.33
(1,442)	1:17:A:DT:H5'	2:120:B:SER:HB2	10	2.97	0.25	2.94
(1,442)	1:17:A:DT:H5'	2:120:B:SER:HB3	10	2.97	0.25	2.94
(1,442)	1:17:A:DT:H5''	2:120:B:SER:HB2	10	2.97	0.25	2.94
(1,442)	1:17:A:DT:H5''	2:120:B:SER:HB3	10	2.97	0.25	2.94
(1,166)	1:12:A:DA:H2	1:11:A:DT:H5'	10	2.85	0.45	3.0
(1,166)	1:12:A:DA:H2	1:11:A:DT:H5''	10	2.85	0.45	3.0
(1,252)	1:18:A:DT:H3	1:5:A:DT:H5'	10	2.73	0.53	2.88

*Continued on next page...*

Continued from previous page...

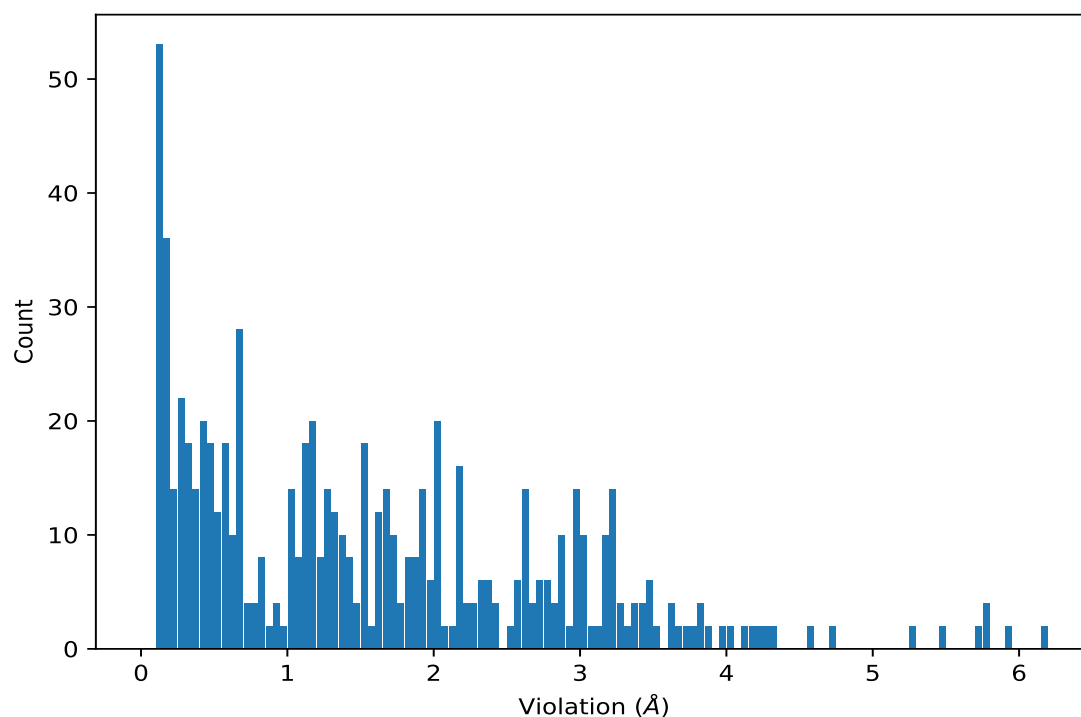
Key	Atom-1	Atom-2	Models <sup>1</sup>	Mean (Å)	SD <sup>1</sup> (Å)	Median (Å)
(1,252)	1:18:A:DT:H3	1:5:A:DT:H5''	10	2.73	0.53	2.88
(1,63)	1:5:A:DT:H5'	1:17:A:DT:H1'	10	2.35	0.47	2.37
(1,63)	1:5:A:DT:H5''	1:17:A:DT:H1'	10	2.35	0.47	2.37
(1,441)	1:17:A:DT:H5'	2:118:B:LYS:HE2	10	2.24	0.57	2.3
(1,441)	1:17:A:DT:H5'	2:118:B:LYS:HE3	10	2.24	0.57	2.3
(1,441)	1:17:A:DT:H5''	2:118:B:LYS:HE2	10	2.24	0.57	2.3
(1,441)	1:17:A:DT:H5''	2:118:B:LYS:HE3	10	2.24	0.57	2.3
(1,257)	1:18:A:DT:H5'	1:15:A:DG:H3'	10	2.04	0.21	1.97

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

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

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

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



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

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

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,66)	1:5:A:DT:H5'	1:18:A:DT:H1'	3	6.17
(1,66)	1:5:A:DT:H5''	1:18:A:DT:H1'	3	6.17
(1,66)	1:5:A:DT:H5'	1:18:A:DT:H1'	6	5.9
(1,66)	1:5:A:DT:H5''	1:18:A:DT:H1'	6	5.9
(1,66)	1:5:A:DT:H5'	1:18:A:DT:H1'	7	5.8
(1,66)	1:5:A:DT:H5''	1:18:A:DT:H1'	7	5.8
(1,66)	1:5:A:DT:H5'	1:18:A:DT:H1'	10	5.77
(1,66)	1:5:A:DT:H5''	1:18:A:DT:H1'	10	5.77
(1,66)	1:5:A:DT:H5'	1:18:A:DT:H1'	2	5.72
(1,66)	1:5:A:DT:H5''	1:18:A:DT:H1'	2	5.72
(1,66)	1:5:A:DT:H5'	1:18:A:DT:H1'	9	5.48
(1,66)	1:5:A:DT:H5''	1:18:A:DT:H1'	9	5.48
(1,66)	1:5:A:DT:H5'	1:18:A:DT:H1'	8	5.29
(1,66)	1:5:A:DT:H5''	1:18:A:DT:H1'	8	5.29
(1,66)	1:5:A:DT:H5'	1:18:A:DT:H1'	4	4.72
(1,66)	1:5:A:DT:H5''	1:18:A:DT:H1'	4	4.72
(1,66)	1:5:A:DT:H5'	1:18:A:DT:H1'	5	4.6
(1,66)	1:5:A:DT:H5''	1:18:A:DT:H1'	5	4.6
(1,66)	1:5:A:DT:H5'	1:18:A:DT:H1'	1	4.34

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

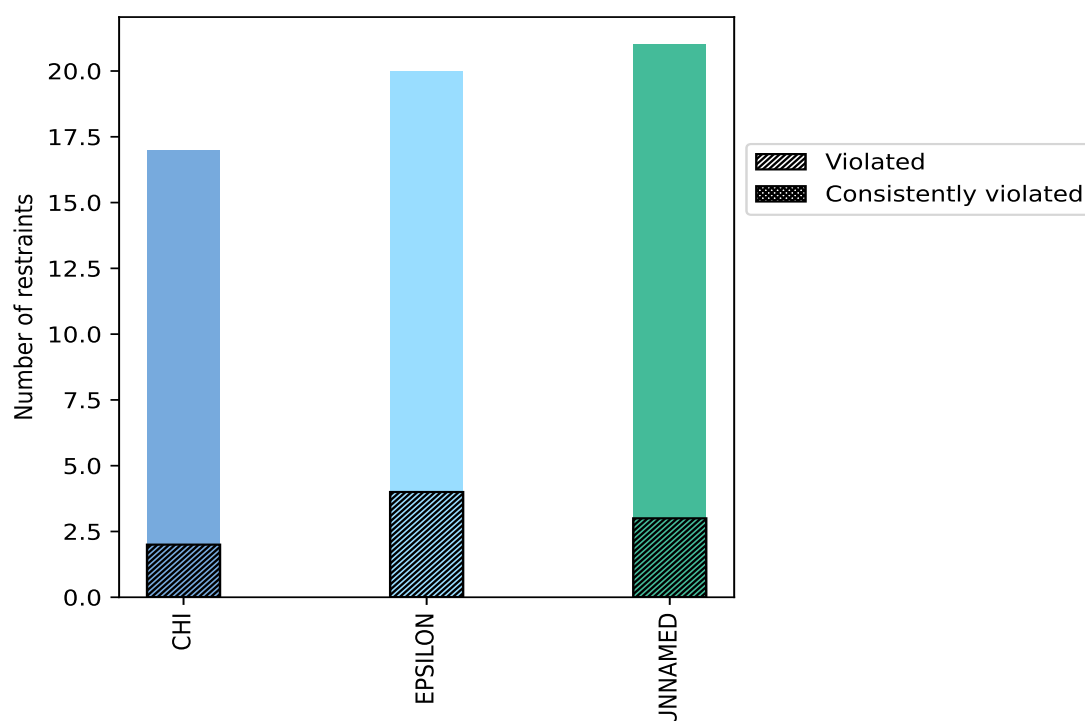
### 10.1 Summary of dihedral-angle violations [i](#)

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>
CHI	17	29.3	2	11.8	3.4	0	0.0	0.0
EPSILON	20	34.5	4	20.0	6.9	0	0.0	0.0
UNNAMED	21	36.2	3	14.3	5.2	0	0.0	0.0
Total	58	100.0	9	15.5	15.5	0	0.0	0.0

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

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



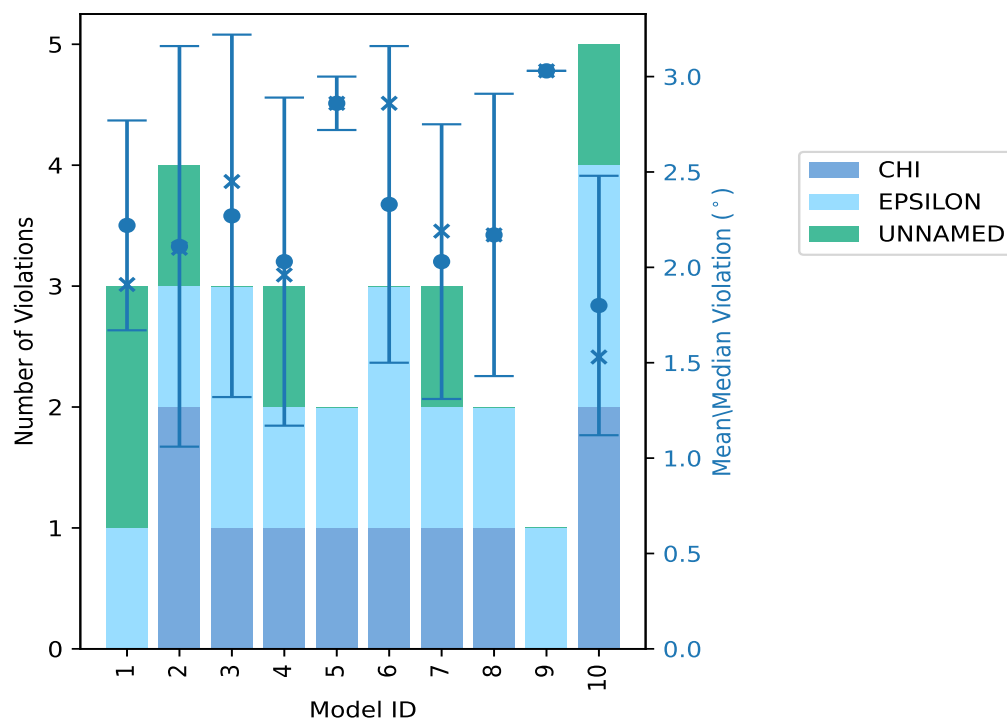
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

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

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

Model ID	Number of violations				Mean (°)	Max (°)	SD (°)	Median (°)
	CHI	EPSILON	UNNAMED	Total				
1	0	1	2	3	2.22	2.99	0.55	1.91
2	2	1	1	4	2.11	3.18	1.05	2.1
3	1	2	0	3	2.27	3.34	0.95	2.45
4	1	1	1	3	2.03	3.12	0.86	1.96
5	1	1	0	2	2.86	3.0	0.14	2.86
6	1	2	0	3	2.33	2.97	0.83	2.86
7	1	1	1	3	2.03	2.83	0.72	2.19
8	1	1	0	2	2.17	2.91	0.74	2.17
9	0	1	0	1	3.03	3.03	0.0	3.03
10	2	2	1	5	1.8	2.73	0.68	1.53

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



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



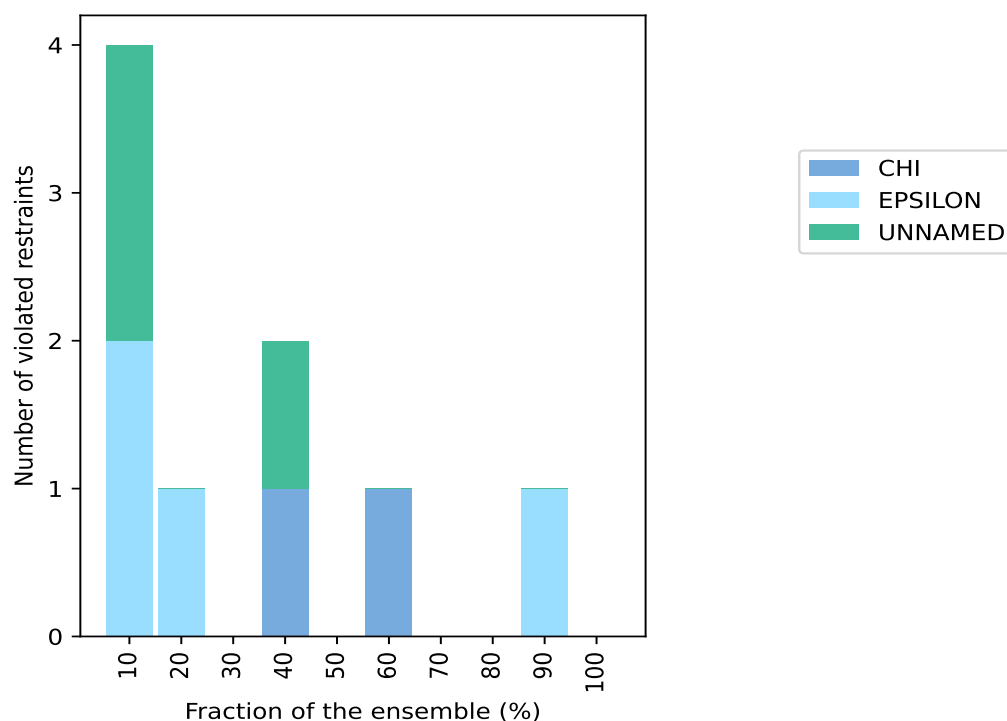
### 10.3 Dihedral-angle violation statistics for the ensemble [i](#)

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

Number of violated restraints				Fraction of the ensemble	
CHI	EPSILON	UNNAMED	Total	Count <sup>1</sup>	%
0	2	2	4	1	10.0
0	1	0	1	2	20.0
0	0	0	0	3	30.0
1	0	1	2	4	40.0
0	0	0	0	5	50.0
1	0	0	1	6	60.0
0	0	0	0	7	70.0
0	0	0	0	8	80.0
0	1	0	1	9	90.0
0	0	0	0	10	100.0

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

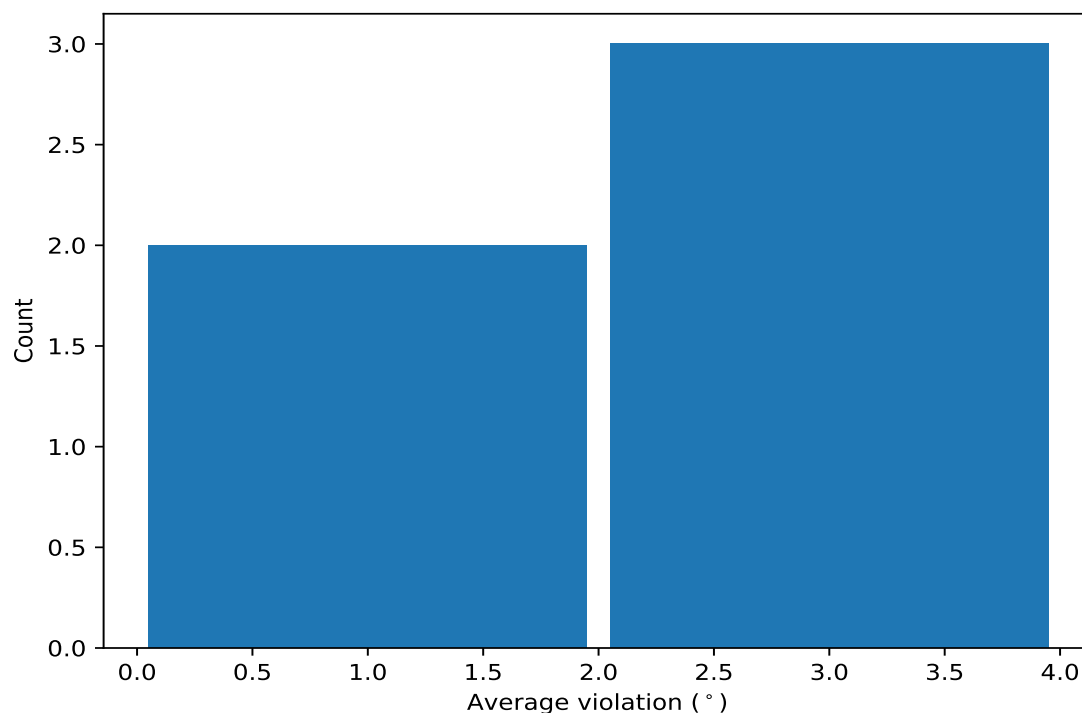
#### 10.3.1 Bar graph : Dihedral-angle Violation statistics for the ensemble [i](#)



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

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

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



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

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

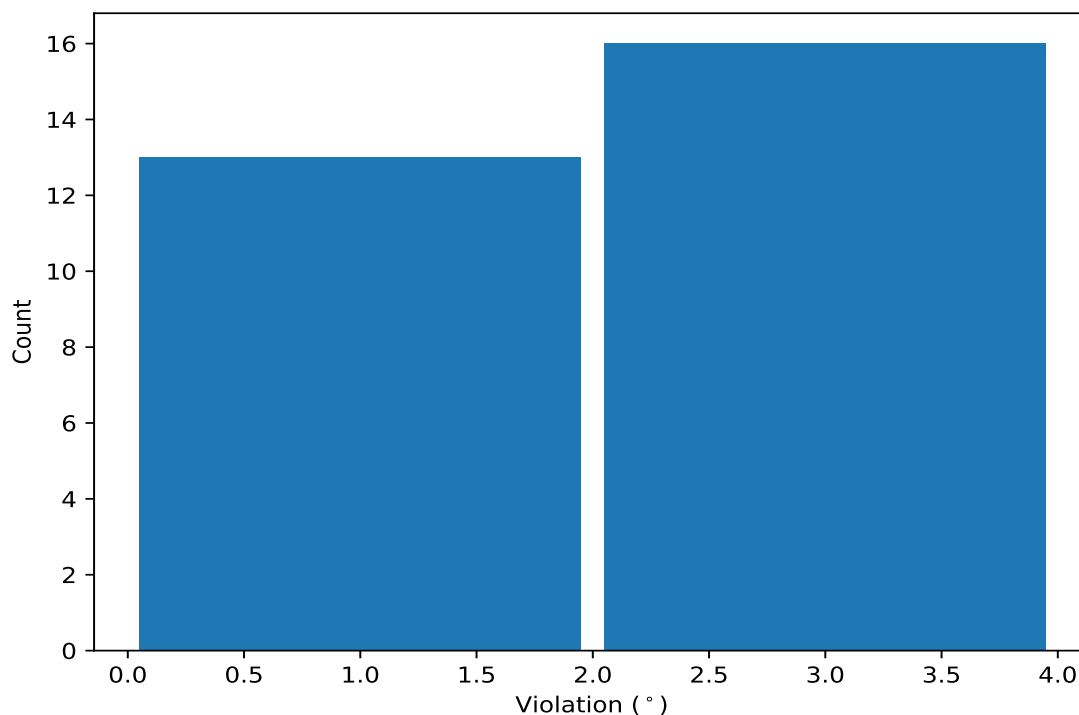
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models <sup>1</sup>	Mean	SD <sup>2</sup>	Median
(1,50)	1:12:A:DA:C4'	1:12:A:DA:C3'	1:12:A:DA:O3'	1:13:A:DG:P	9	3.02	0.17	3.0
(1,9)	1:15:A:DG:O4'	1:15:A:DG:C1'	1:15:A:DG:N9	1:15:A:DG:C4	6	2.17	0.71	2.34
(1,34)	1:16:A:DT:O4'	1:16:A:DT:C1'	1:16:A:DT:N1	1:16:A:DT:C2	4	2.56	0.35	2.46
(1,17)	1:8:A:DG:H1'	1:8:A:DG:C1'	1:8:A:DG:N9	1:8:A:DG:C4	4	1.24	0.31	1.08
(1,58)	1:20:A:DG:C4'	1:20:A:DG:C3'	1:20:A:DG:O3'	1:21:A:DG:P	2	1.1	0.07	1.1

<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 (°)
(1,50)	1:12:A:DA:C4'	1:12:A:DA:C3'	1:12:A:DA:O3'	1:13:A:DG:P	3	3.34
(1,50)	1:12:A:DA:C4'	1:12:A:DA:C3'	1:12:A:DA:O3'	1:13:A:DG:P	2	3.18
(1,34)	1:16:A:DT:O4'	1:16:A:DT:C1'	1:16:A:DT:N1	1:16:A:DT:C2	2	3.14
(1,50)	1:12:A:DA:C4'	1:12:A:DA:C3'	1:12:A:DA:O3'	1:13:A:DG:P	4	3.12
(1,50)	1:12:A:DA:C4'	1:12:A:DA:C3'	1:12:A:DA:O3'	1:13:A:DG:P	9	3.03
(1,50)	1:12:A:DA:C4'	1:12:A:DA:C3'	1:12:A:DA:O3'	1:13:A:DG:P	5	3.0
(1,50)	1:12:A:DA:C4'	1:12:A:DA:C3'	1:12:A:DA:O3'	1:13:A:DG:P	1	2.99
(1,50)	1:12:A:DA:C4'	1:12:A:DA:C3'	1:12:A:DA:O3'	1:13:A:DG:P	6	2.97
(1,9)	1:15:A:DG:O4'	1:15:A:DG:C1'	1:15:A:DG:N9	1:15:A:DG:C4	8	2.91
(1,9)	1:15:A:DG:O4'	1:15:A:DG:C1'	1:15:A:DG:N9	1:15:A:DG:C4	6	2.86