

# wwPDB NMR Structure Validation Summary Report (i)

Dec 25, 2024 – 06:42 PM EST

PDB ID	:	7SXC
BMRB ID	:	30967
Title	:	cTnC-TnI chimera complexed with calcium
Authors	:	Poppe, L.; Hartman, J.J.; Romero, A.; Reagan, J.D.
Deposited on	:	2021-11-22

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

We welcome your comments at 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 (i) 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 (1)) 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

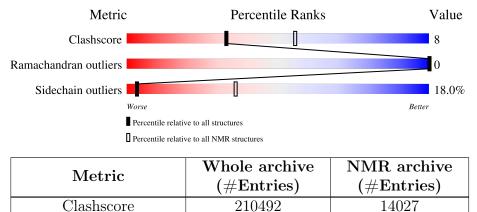
Ramachandran outliers

#### Overall quality at a glance (i) 1

The following experimental techniques were used to determine the structure: SOLUTION NMR

The overall completeness of chemical shifts assignment is 95%.

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.



207382

	Sidechain outliers	206894	12463	
		1	l	1
Tł	ne table below summari	ses the geometric issue	ues observed across t	the polymeric chains and their
fit	to the experimental d	ata. The red, orange	e, yellow and green s	segments indicate the fraction
of	residues that contain o	outliers for $>=3, 2, 1$	and 0 types of geom	netric quality criteria. A cyan
seg	gment indicates the frac	tion of residues that a	re not part of the wel	ll-defined cores, and a grey seg-
me	ent represents the fracti	on of residues that are	e not modelled. The	numeric value for each fraction
is	indicated below the cor	responding segment,	with a dot represent	ing fractions $<=5\%$

12486

Mol	Chain	Length	Quality of	of chain		
1	٨	190				
	A	129	53%	14%	•	32%



# 2 Ensemble composition and analysis (i)

This entry contains 10 models. Model 10 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *target function*.

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:3-A:85, A:98-A:102 (88)	0.23	10		

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 1 clusters and 6 single-model clusters were found.

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



# 3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 1961 atoms, of which 973 are hydrogens and 0 are deuteriums.

• Molecule 1 is a protein called Troponin C, slow skeletal and cardiac muscles, Troponin I, cardiac muscle chimera.

Mol	Chain	Residues		Atoms				Trace	
1	٨	190	Total	С	Η	Ν	0	$\mathbf{S}$	0
	А	129	1960	609	973	165	203	10	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP P63316
А	92	ALA	-	linker	UNP P63316
А	93	GLY	-	linker	UNP P63316
А	94	ALA	-	linker	UNP P63316
А	95	GLY	-	linker	UNP P63316

• Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	
2	А	1	Total Ca	
			1 1	

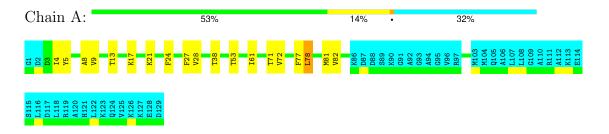


# 4 Residue-property plots (i)

# 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.

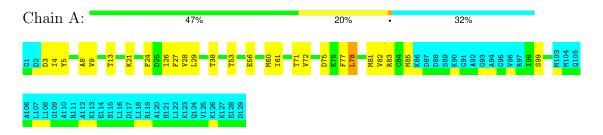
• Molecule 1: Troponin C, slow skeletal and cardiac muscles, Troponin I, cardiac muscle chimera



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

The representative model is number 10. Colouring as in section 4.1 above.

• Molecule 1: Troponin C, slow skeletal and cardiac muscles, Troponin I, cardiac muscle chimera





# 5 Refinement protocol and experimental data overview (i)

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

Of the 100 calculated structures, 10 were deposited, based on the following criterion: target function.

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

Software name	Classification	Version
CYANA	refinement	3.98.13
CYANA	structure calculation	3.98.13

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	1570
Number of shifts mapped to atoms	1570
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	95%



# 6 Model quality (i)

# 6.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: CA

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

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 6.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	А	682	655	654	$11 \pm 3$
All	All	6830	6550	6540	110

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 8.

5 of 19 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
Atom-1	Atom-2	Clash(A)	Distance(A)	Worst	Total
1:A:9:VAL:HG23	1:A:78:LEU:HD12	0.74	1.57	6	9
1:A:9:VAL:CG2	1:A:78:LEU:HD12	0.61	2.25	6	7
1:A:17:LYS:HD3	1:A:78:LEU:HD11	0.59	1.74	4	1
1:A:17:LYS:CD	1:A:78:LEU:HD21	0.55	2.31	6	2
1:A:9:VAL:HB	1:A:82:VAL:HG21	0.54	1.80	1	10



# 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	А	88/129~(68%)	$84\pm1$ (95 $\pm1\%$ )	$4\pm1~(5\pm1\%)$	0±0 (0±0%)	100 100
All	All	880/1290 (68%)	836 (95%)	44 (5%)	0 (0%)	100 100

There are no Ramachandran outliers.

#### 6.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent side chain 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 side chain conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	А	75/105~(71%)	$62\pm2$ (82 $\pm3\%$ )	$14\pm2~(18\pm3\%)$	3 36
All	All	750/1050~(71%)	615 (82%)	135 (18%)	3 36

5 of 30 unique residues with a non-rotameric side chain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	А	13	THR	10
1	А	53	THR	10
1	А	71	THR	10
1	А	77	PHE	10
1	А	78	LEU	10

#### 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)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

## 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 95% for the well-defined parts and 93% 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	1570
Number of shifts mapped to atoms	1570
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	4

#### 7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	${\rm Correction}\pm{\rm precision},ppm$	Suggested action
$^{13}C_{\alpha}$	128	$-0.16 \pm 0.03$	None needed ( $< 0.5$ ppm)
$^{13}C_{\beta}$	117	$0.22 \pm 0.05$	None needed ( $< 0.5$ ppm)
$^{13}C'$	125	$-0.31 \pm 0.09$	None needed ( $< 0.5$ ppm)
$^{15}N$	125	$-0.35 \pm 0.31$	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 95%, i.e. 1089 atoms were assigned a chemical shift out of a possible 1152. 0 out of 13 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^{1}\mathbf{H}$	$^{13}\mathrm{C}$	$^{15}\mathbf{N}$
Backbone	440/442~(100%)	180/180~(100%)	174/176~(99%)	86/86~(100%)
Sidechain	605/651~(93%)	413/421 (98%)	184/213~(86%)	8/17~(47%)

Continued on next page...



Continucu	from previous page.	••		
	Total	$^{1}\mathrm{H}$	$^{13}\mathrm{C}$	$^{15}\mathbf{N}$
Aromatic	44/59~(75%)	22/29~(76%)	22/30~(73%)	0/0 (%)
Overall	1089/1152~(95%)	615/630~(98%)	380/419~(91%)	94/103~(91%)

Continued from previous page...

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

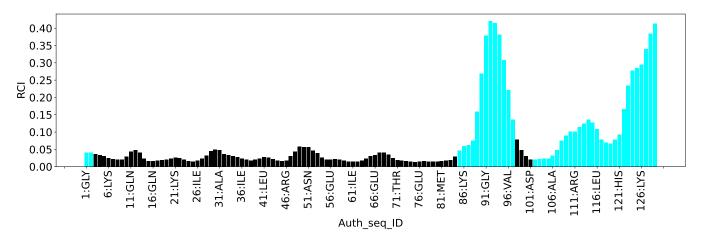
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	А	46	ARG	NE	110.80	76.53 - 92.65	16.3
1	А	83	ARG	NE	109.50	76.53 - 92.65	15.4
1	А	97	ARG	CZ	110.80	141.81 - 177.93	-13.6
1	А	21	LYS	HD3	0.24	0.54 - 2.65	-6.4

#### 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 (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	2242
Intra-residue ( i-j =0)	510
Sequential ( i-j =1)	642
Medium range ( $ i-j >1$ and $ i-j <5$ )	634
Long range $( i-j  \ge 5)$	453
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	204
Number of unmapped restraints	0
Number of restraints per residue	18.8
Number of long range restraints per residue <sup>1</sup>	3.5

<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 (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)	1.9	0.15
0.2-0.5 (Medium)	None	None
>0.5 (Large)	None	None



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

Dihedral-angle violations less than  $1^\circ$  are not included in the calculation.

Bins ( $^{\circ}$ )	Average number of violations per model	Max ( $^{\circ}$ )
1.0-10.0 (Small)	6.2	2.86
10.0-20.0 (Medium)	None	None
>20.0 (Large)	None	None



# 9 Distance violation analysis (i)

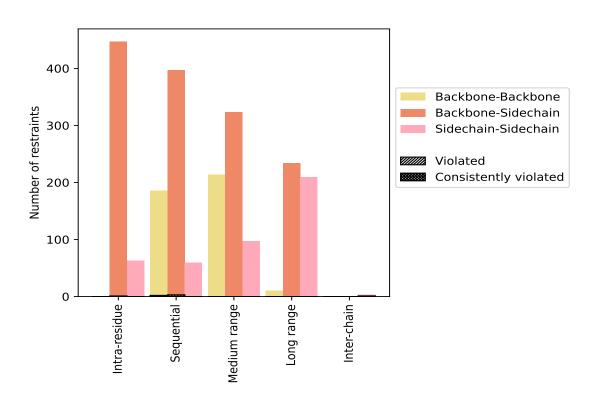
# 9.1 Summary of distance violations (i)

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.

Destusints torus	Count	$\%^1$	Vio	lated	3	Consis	tently	$\vee$ Violated <sup>4</sup>
Restraints type	Count	701	Count	$\%^2$	$\%^1$	Count	$\%^2$	$\%^1$
Intra-residue ( i-j =0)	510	22.7	1	0.2	0.0	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	447	19.9	1	0.2	0.0	0	0.0	0.0
Sidechain-Sidechain	63	2.8	0	0.0	0.0	0	0.0	0.0
Sequential ( i-j =1)	642	28.6	5	0.8	0.2	0	0.0	0.0
Backbone-Backbone	186	8.3	2	1.1	0.1	0	0.0	0.0
Backbone-Sidechain	397	17.7	3	0.8	0.1	0	0.0	0.0
Sidechain-Sidechain	59	2.6	0	0.0	0.0	0	0.0	0.0
Medium range ( $ i-j  > 1 \&  i-j  < 5$ )	634	28.3	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	214	9.5	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	323	14.4	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	97	4.3	0	0.0	0.0	0	0.0	0.0
Long range $( i-j  \ge 5)$	453	20.2	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	10	0.4	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	234	10.4	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	209	9.3	0	0.0	0.0	0	0.0	0.0
Inter-chain	0	0.0	0	0.0	0.0	0	0.0	0.0
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
Hydrogen bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	2242	100.0	7	0.3	0.3	1	0.0	0.0
Backbone-Backbone	410	18.3	2	0.5	0.1	0	0.0	0.0
Backbone-Sidechain	1401	62.5	4	0.3	0.2	0	0.0	0.0
Sidechain-Sidechain	431	19.2	1	0.2	0.0	1	0.2	0.0

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





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

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

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

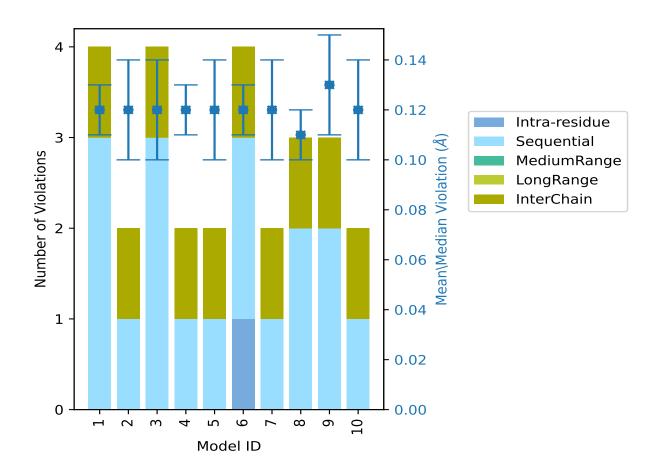
The following table provides the distance violation statistics for each model in the ensemble. Violations less than 0.1 Å

are not included in the statistics.

Model ID		Nun	nber o	f viola	ations	5	Mean (Å)	Max (Å)	$SD^{6}$ (Å)	Madian (Å)
Model ID	$IR^1$	$SQ^2$	$MR^3$	$LR^4$	$  IC^5  $	Total	Mean (A)	Max (A)	$SD^{\circ}(A)$	Median (Å)
1	0	3	0	0	1	4	0.12	0.13	0.01	0.12
2	0	1	0	0	1	2	0.12	0.13	0.02	0.12
3	0	3	0	0	1	4	0.12	0.14	0.02	0.12
4	0	1	0	0	1	2	0.12	0.13	0.01	0.12
5	0	1	0	0	1	2	0.12	0.13	0.02	0.12
6	1	2	0	0	1	4	0.12	0.13	0.01	0.12
7	0	1	0	0	1	2	0.12	0.13	0.02	0.12
8	0	2	0	0	1	3	0.11	0.13	0.01	0.11
9	0	2	0	0	1	3	0.13	0.15	0.02	0.13
10	0	1	0	0	1	2	0.12	0.13	0.02	0.12



 $^1$ Intra-residue restraints,  $^2$ Sequential restraints,  $^3$ Medium range restraints,  $^4$ Long range restraints,  $^5$ Inter-chain restraints,  $^6$ 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, 2233(IR:509, SQ:637, MR:634, LR:453, IC:0) restraints are not violated in the ensemble.

Nu	Number of violated restraints								
$\mathrm{IR}^{1}$	$SQ^2$	$MR^3$	LR <sup>4</sup>	$IC^5$	Total	$\operatorname{Count}^6$	%		
1	2	0	0	0	3	1	10.0		
0	0	0	0	0	0	2	20.0		
0	2	0	0	0	2	3	30.0		

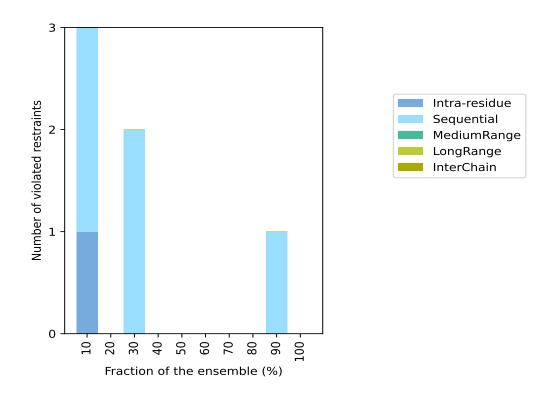
Continued on next page...



	Continueu from pretious page											
Nu	ımber	of vio	lated	restra	aints	Fraction of the ensemble						
$\mathrm{IR}^{1}$	$SQ^2$	$MR^3$	$LR^4$	IC <sup>5</sup>	Total	$\operatorname{Count}^6$	%					
0	0	0	0	0	0	4	40.0					
0	0	0	0	0	0	5	50.0					
0	0	0	0	0	0	6	60.0					
0	0	0	0	0	0	7	70.0					
0	0	0	0	0	0	8	80.0					
0	1	0	0	0	1	9	90.0					
0	0	0	0	0	0	10	100.0					

Continued from previous page..

 $^{1}$ Intra-residue restraints,  $^{2}$ Sequential restraints,  $^{3}$ Medium range restraints,  $^{4}$ Long range restraints,  $^{5}$ Inter-chain restraints,  $^{6}$  Number of models with violations



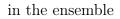
#### 9.3.1 Bar graph : Distance violation statistics for the ensemble (i)

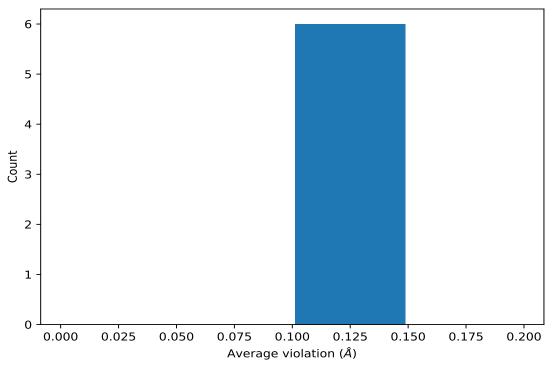
### 9.4 Most violated distance restraints in the ensemble (i)

#### 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







#### 9.4.2 Table: Most violated distance 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. Rows with same key represent combinatorial or ambiguous restraints and are counted as a single restraint.

Key	Atom-1	Atom-2	$\mathbf{Models}^1$	Mean (Å)	$SD^1$ (Å)	Median (Å)
(3,1)	1:65:A:ASP:OD1	2:201:A:CA:CA	10	0.13	0.0	0.13
(1,1057)	1:9:A:VAL:HB	1:10:A:GLU:H	9	0.1	0.0	0.1
(1,2084)	1:104:A:MET:HG2	1:105:A:GLN:H	3	0.13	0.01	0.13
(1,2084)	1:104:A:MET:HG3	1:105:A:GLN:H	3	0.13	0.01	0.13
(1,1605)	1:6:A:LYS:HB2	1:7:A:ALA:H	3	0.11	0.0	0.11
(1,1605)	1:6:A:LYS:HB3	1:7:A:ALA:H	3	0.11	0.0	0.11

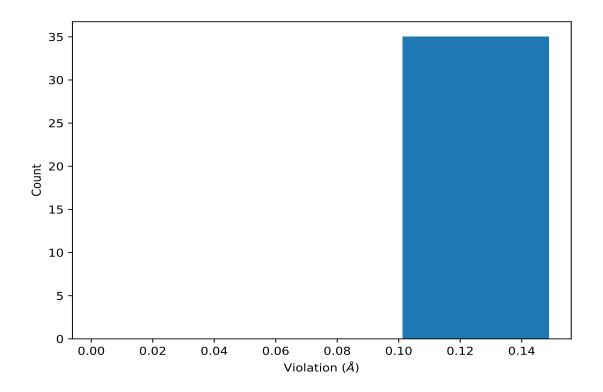
<sup>1</sup>Number of violated models, <sup>2</sup>Standard deviation

### 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)
(1,733)	1:92:A:ALA:H	1:93:A:GLY:H	9	0.15
(1,2084)	1:104:A:MET:HG2	1:105:A:GLN:H	3	0.14
(1,2084)	1:104:A:MET:HG3	1:105:A:GLN:H	3	0.14
(3,1)	1:65:A:ASP:OD1	2:201:A:CA:CA	1	0.13
(3,1)	1:65:A:ASP:OD1	2:201:A:CA:CA	2	0.13
(3,1)	1:65:A:ASP:OD1	2:201:A:CA:CA	3	0.13
(3,1)	1:65:A:ASP:OD1	2:201:A:CA:CA	4	0.13
(3,1)	1:65:A:ASP:OD1	2:201:A:CA:CA	5	0.13
(3,1)	1:65:A:ASP:OD1	2:201:A:CA:CA	6	0.13
(3,1)	1:65:A:ASP:OD1	2:201:A:CA:CA	7	0.13
(3,1)	1:65:A:ASP:OD1	2:201:A:CA:CA	8	0.13



# 10 Dihedral-angle violation analysis (i)

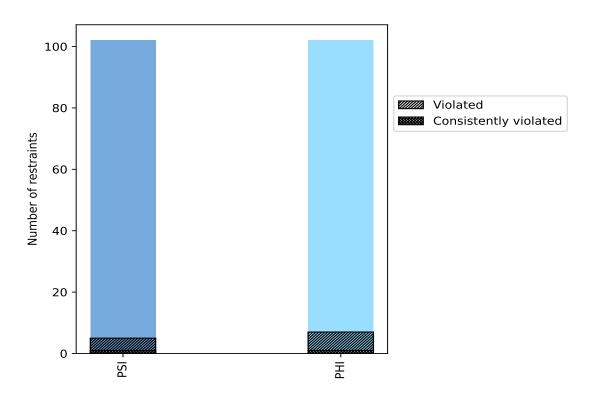
# 10.1 Summary of dihedral-angle violations (i)

The following table provides the summary of dihedral-angle violations in different dihedral-angle types. Violations less than  $1^{\circ}$  are not included in the calculation.

	Count	$\%^1$	$Violated^3$			Consistently Violated <sup>4</sup>		
Angle type	Count	/0	Count	$\%^2$	$\%^1$	Count	$\%^2$	$\%^1$
PSI	102	50.0	5	4.9	2.5	1	1.0	0.5
PHI	102	50.0	7	6.9	3.4	1	1.0	0.5
Total	204	100.0	12	5.9	5.9	2	1.0	1.0

 $^1$  percentage calculated with respect to total number of dihedral-angle restraints,  $^2$  percentage calculated with respect to number of restraints in a particular dihedral-angle type,  $^3$  violated in at least one model,  $^4$  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

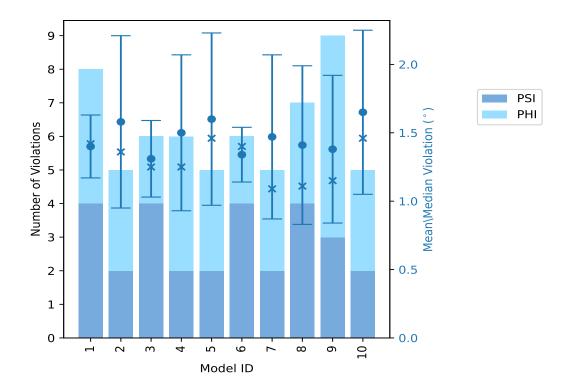


# 10.2 Dihedral-angle violation statistics for each model (i)

The following table provides the dihedral-angle violation statistics for each model in the ensemble. Violations less than  $1^{\circ}$  are not included in the statistics.

Model ID	Nun	nber c	of violations	Mean (°)	Max (°)	SD (°)	Median (°)
Model ID	PSI	PHI	Total	Mean ()	Max ()		Median ()
1	4	4	8	1.4	1.75	0.23	1.42
2	2	3	5	1.58	2.8	0.63	1.36
3	4	2	6	1.31	1.72	0.28	1.25
4	2	4	6	1.5	2.75	0.57	1.25
5	2	3	5	1.6	2.81	0.63	1.46
6	4	2	6	1.34	1.61	0.2	1.4
7	2	3	5	1.47	2.61	0.6	1.09
8	4	3	7	1.41	2.78	0.58	1.11
9	3	6	9	1.38	2.86	0.54	1.15
10	2	3	5	1.65	2.81	0.6	1.46

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  ${\bf y}$  axis on the right



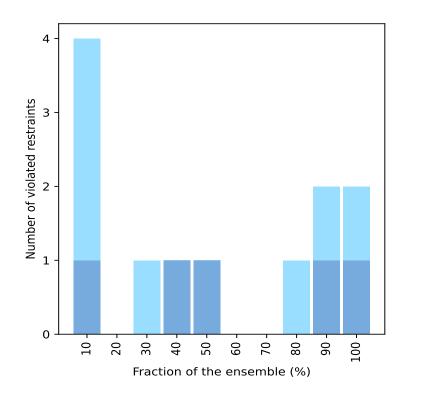
## 10.3 Dihedral-angle violation statistics for the ensemble (i)

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.

Nun	nber o	f violated restraints	Fractio	n of the ensemble
PSI	PHI	Total	$\operatorname{Count}^1$	%
1	3	4	1	10.0
0	0	0	2	20.0
0	1	1	3	30.0
1	0	1	4	40.0
1	0	1	5	50.0
0	0	0	6	60.0
0	0	0	7	70.0
0	1	1	8	80.0
1	1	2	9	90.0
1	1	2	10	100.0

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

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



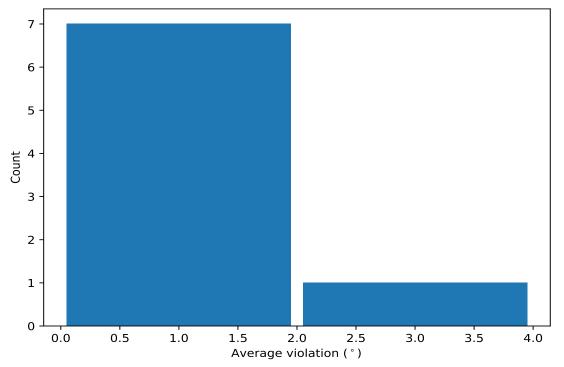




# 10.4 Most violated dihedral-angle restraints in the ensemble (i)

#### 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	$\mathbf{Models}^1$	Mean	$\mathbf{SD}^2$	Median
(1,91)	1:48:A:LEU:C	1:49:A:GLY:N	1:49:A:GLY:CA	1:49:A:GLY:C	10	1.46	0.12	1.52
(1,124)	1:67:A:ASP:N	1:67:A:ASP:CA	1:67:A:ASP:C	1:68:A:GLY:N	10	1.11	0.05	1.1
(1,169)	1:100:A:ALA:C	1:101:A:ASP:N	1:101:A:ASP:CA	1:101:A:ASP:C	9	1.37	0.19	1.36
(1,128)	1:70:A:GLY:N	1:70:A:GLY:CA	1:70:A:GLY:C	1:71:A:THR:N	9	1.08	0.03	1.08
(1,171)	1:101:A:ASP:C	1:102:A:ALA:N	1:102:A:ALA:CA	1:102:A:ALA:C	8	2.58	0.52	2.79
(1,72)	1:39:A:LYS:N	1:39:A:LYS:CA	1:39:A:LYS:C	1:40:A:GLU:N	5	1.24	0.14	1.32
(1,176)	1:104:A:MET:N	1:104:A:MET:CA	1:104:A:MET:C	1:105:A:GLN:N	4	1.4	0.2	1.47
(1,79)	1:42:A:GLY:C	1:43:A:LYS:N	1:43:A:LYS:CA	1:43:A:LYS:C	3	1.61	0.11	1.61

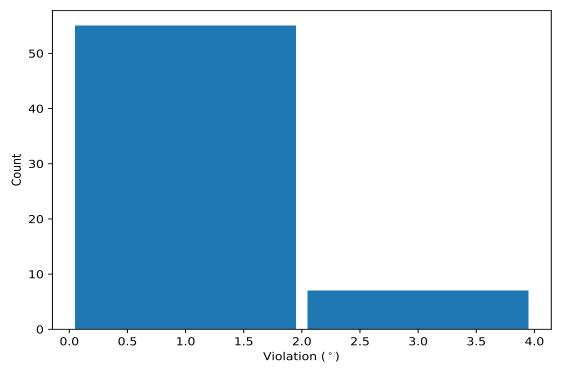
<sup>1</sup> Number of violated models, <sup>2</sup>Standard deviation, All angle values are in degree (°)



### 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 ( $^{\circ}$ )
(1,171)	1:101:A:ASP:C	1:102:A:ALA:N	1:102:A:ALA:CA	1:102:A:ALA:C	9	2.86
(1,171)	1:101:A:ASP:C	1:102:A:ALA:N	1:102:A:ALA:CA	1:102:A:ALA:C	5	2.81
(1,171)	1:101:A:ASP:C	1:102:A:ALA:N	1:102:A:ALA:CA	1:102:A:ALA:C	10	2.81
(1,171)	1:101:A:ASP:C	1:102:A:ALA:N	1:102:A:ALA:CA	1:102:A:ALA:C	2	2.8
(1,171)	1:101:A:ASP:C	1:102:A:ALA:N	1:102:A:ALA:CA	1:102:A:ALA:C	8	2.78
(1,171)	1:101:A:ASP:C	1:102:A:ALA:N	1:102:A:ALA:CA	1:102:A:ALA:C	4	2.75
(1,171)	1:101:A:ASP:C	1:102:A:ALA:N	1:102:A:ALA:CA	1:102:A:ALA:C	7	2.61
(1,79)	1:42:A:GLY:C	1:43:A:LYS:N	1:43:A:LYS:CA	1:43:A:LYS:C	1	1.75
(1,169)	1:100:A:ALA:C	1:101:A:ASP:N	1:101:A:ASP:CA	1:101:A:ASP:C	3	1.72
(1,79)	1:42:A:GLY:C	1:43:A:LYS:N	1:43:A:LYS:CA	1:43:A:LYS:C	6	1.61

