wwPDB NMR Structure Validation Summary Report

Sep 15, 2026 – 12:34 PM JST

PDB ID	:	21HJ / pdb_000021hj
BMRB ID	:	36817
Title	:	Solution structures of BRD9 bromodomain in complex with histone H3 lactyl-lysine 18 (H3K18la) peptide
Authors	:	Wei, E.; Zeng, L.
Deposited on	:	2025-12-12

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 (i)) were used in the production of this report:

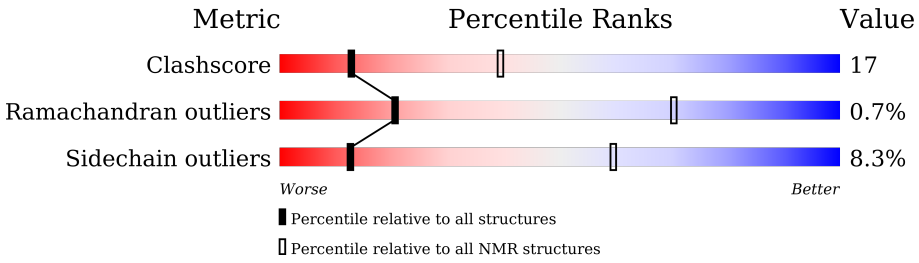
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2.1
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.50

1 Overall quality at a glance (i)

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

The overall completeness of chemical shifts assignment is 81%.

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 >=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 <=5%

Mol	Chain	Length	Quality of chain
1	A	117	57% 26% • 16%
2	B	15	100%

2 Ensemble composition and analysis

This entry contains 20 models. Model 9 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:7-A:104 (98)	0.56	9

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

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

The models can be grouped into 5 clusters and 2 single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2110 atoms, of which 1062 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Bromodomain-containing protein 9.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	117	1877	602	939	156	172	8	0

- Molecule 2 is a protein called Histone H3.3C.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
2	B	15	233	67	123	22	21	0

4 Residue-property plots

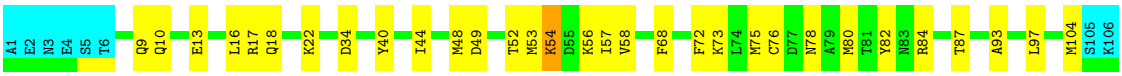
4.1 Average score per residue in the NMR ensemble

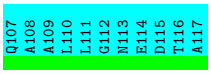
These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for 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: Bromodomain-containing protein 9

Chain A:

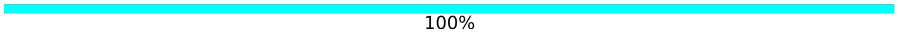


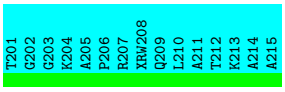




• Molecule 2: Histone H3.3C

Chain B:





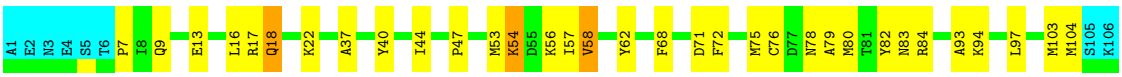
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

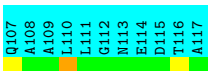
The representative model is number 9. Colouring as in section 4.1 above.

• Molecule 1: Bromodomain-containing protein 9

Chain A:

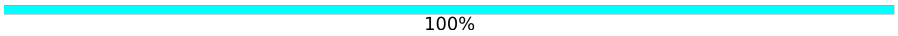


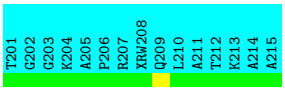




• Molecule 2: Histone H3.3C

Chain B:




T201
G202
G203
K204
A205
P206
R207
XRW208
Q209
L210
A211
T212
K213
A214
A215

5 Refinement protocol and experimental data overview

The models were refined using the following method: *torsion angle dynamics*.

Of the 200 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
ARIA	refinement	
CNS	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	1386
Number of shifts mapped to atoms	1386
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	81%

6 Model quality

6.1 Standard geometry

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

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

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	802	813	809	27±5
2	B	0	0	0	0±0
All	All	16040	16260	16180	536

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 17.

5 of 198 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:7:PRO:HB3	1:A:10:GLN:HG2	0.85	1.45	17	1
1:A:7:PRO:HA	1:A:10:GLN:HB3	0.85	1.48	7	2
1:A:44:ILE:HD11	1:A:82:TYR:HB2	0.81	1.50	5	14
1:A:84:ARG:HD2	1:A:84:ARG:H	0.80	1.34	9	11
1:A:69:LYS:HD3	1:A:104:MET:HE3	0.80	1.51	5	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	98/117 (84%)	93±2 (95±2%)	4±2 (4±2%)	1±1 (1±1%)	20 70
2	B	0	-	-	-	-
All	All	1960/2640 (74%)	1863 (95%)	84 (4%)	13 (1%)	20 70

All 4 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	38	PRO	8
1	A	7	PRO	3
1	A	60	ASN	1
1	A	85	PRO	1

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	86/100 (86%)	79±2 (92±2%)	7±2 (8±2%)	12 59
2	B	0	-	-	-
All	All	1720/2160 (80%)	1577 (92%)	143 (8%)	12 59

5 of 36 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	54	LYS	20
1	A	49	ASP	14
1	A	16	LEU	9
1	A	18	GLN	8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	9	GLN	7

6.3.3 RNA (i)

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains (i)

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
2	XRW	B	208	2	12,13,14	0.39±0.02	0±0 (0±0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of angles that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
2	XRW	B	208	2	10,15,17	0.54±0.07	0±0 (0±0%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	XRW	B	208	2	-	1±0,13,14,16	-

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

All unique torsion outliers are listed below.

Mol	Chain	Res	Type	Atoms	Models (Total)
2	B	208	XRW	C24-C22-NZ-CE	16

There are no ring outliers.

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 81% for the well-defined parts and 77% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

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	1386
Number of shifts mapped to atoms	1386
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	2

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	129	-0.50 ± 0.22	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	122	-0.01 ± 0.08	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	103	0.51 ± 0.44	None needed (imprecise)

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 81%, i.e. 1130 atoms were assigned a chemical shift out of a possible 1401. 10 out of 11 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	376/482 (78%)	189/194 (97%)	98/196 (50%)	89/92 (97%)
Sidechain	630/762 (83%)	427/497 (86%)	201/238 (84%)	2/27 (7%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	124/157 (79%)	62/76 (82%)	62/73 (85%)	0/8 (0%)
Overall	1130/1401 (81%)	678/767 (88%)	361/507 (71%)	91/127 (72%)

7.1.4 Statistically unusual chemical shifts (i)

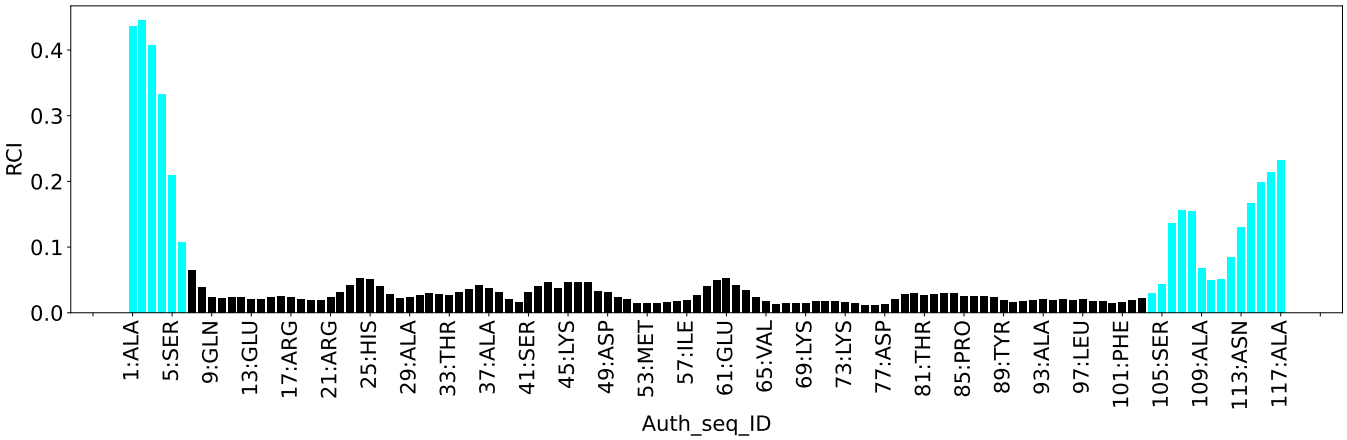
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	87	THR	HG1	5.45	0.08 – 2.19	20.4
1	A	85	PRO	HA	2.72	2.78 – 6.00	-5.2

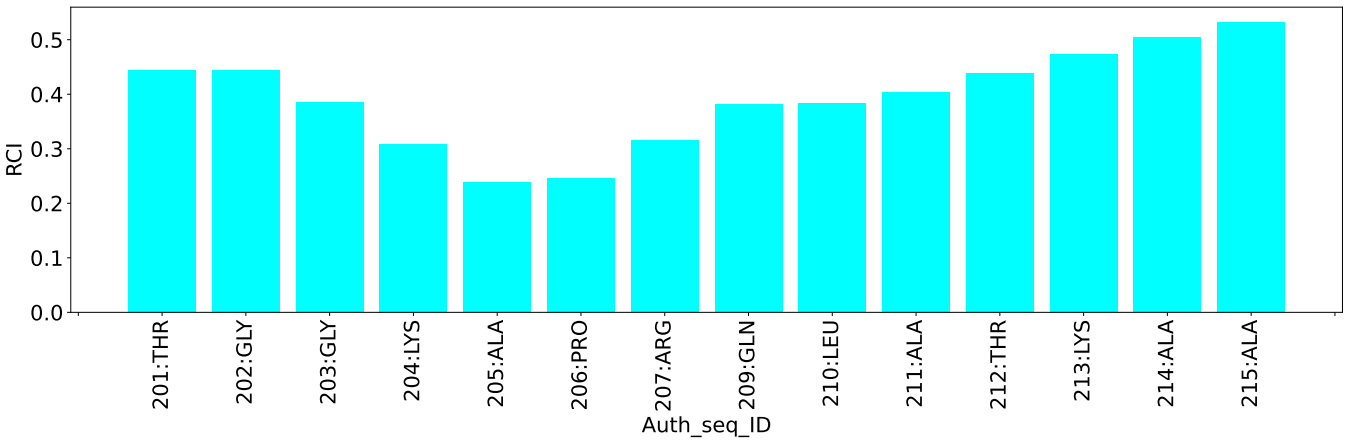
7.1.5 Random Coil Index (RCI) plots (i)

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:



Random coil index (RCI) for chain B:


RCI
0.0
0.1
0.2
0.3
0.4
0.5
201:THR
202:GLY
203:GLY
204:LYS
205:ALA
206:PRO
207:ARG
209:GLN
210:LEU
211:ALA
212:THR
213:LYS
214:ALA
215:ALA
Auth_seq_ID

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	1527
Intra-residue (|i-j|=0)	547
Sequential (|i-j|=1)	298
Medium range (|i-j|>1 and |i-j|<5)	250
Long range (|i-j|≥5)	291
Inter-chain	22
Hydrogen bond restraints	119
Disulfide bond restraints	0
Total dihedral-angle restraints	196
Number of unmapped restraints	0
Number of restraints per residue	13.1
Number of long range restraints per residue[1]	2.2

[1]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)	65.7	0.2
0.2-0.5 (Medium)	65.8	0.5
>0.5 (Large)	97.0	3.45

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)	3.9	5.49
10.0-20.0 (Medium)	None	None
>20.0 (Large)	None	None

9 Distance violation analysis (i)

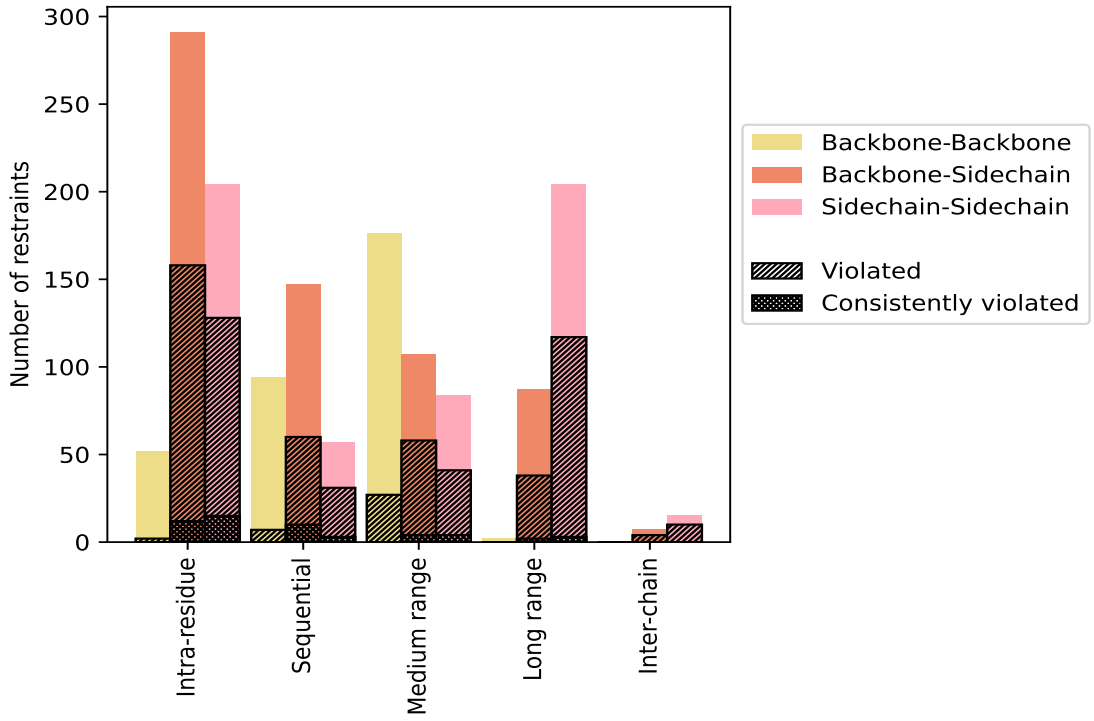
9.1 Summary of distance violations (i)

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.

Restraints type	Count	%[1]	Violated[3]			Consistently Violated[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
Intra-residue (|i-j|=0)	547	35.8	288	52.7	18.9	27	4.9	1.8
Backbone-Backbone	52	3.4	2	3.8	0.1	0	0.0	0.0
Backbone-Sidechain	291	19.1	158	54.3	10.3	12	4.1	0.8
Sidechain-Sidechain	204	13.4	128	62.7	8.4	15	7.4	1.0
Sequential (|i-j|=1)	298	19.5	98	32.9	6.4	13	4.4	0.9
Backbone-Backbone	94	6.2	7	7.4	0.5	0	0.0	0.0
Backbone-Sidechain	147	9.6	60	40.8	3.9	10	6.8	0.7
Sidechain-Sidechain	57	3.7	31	54.4	2.0	3	5.3	0.2
Medium range (|i-j|>1 & |i-j|<5)	250	16.4	106	42.4	6.9	8	3.2	0.5
Backbone-Backbone	59	3.9	7	11.9	0.5	0	0.0	0.0
Backbone-Sidechain	107	7.0	58	54.2	3.8	4	3.7	0.3
Sidechain-Sidechain	84	5.5	41	48.8	2.7	4	4.8	0.3
Long range (|i-j|≥5)	291	19.1	155	53.3	10.2	5	1.7	0.3
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	87	5.7	38	43.7	2.5	2	2.3	0.1
Sidechain-Sidechain	204	13.4	117	57.4	7.7	3	1.5	0.2
Inter-chain	22	1.4	14	63.6	0.9	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	7	0.5	4	57.1	0.3	0	0.0	0.0
Sidechain-Sidechain	15	1.0	10	66.7	0.7	0	0.0	0.0
Hydrogen bond	119	7.8	20	16.8	1.3	0	0.0	0.0
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	1527	100.0	681	44.6	44.6	53	3.5	3.5
Backbone-Backbone	324	21.2	36	11.1	2.4	0	0.0	0.0
Backbone-Sidechain	639	41.8	318	49.8	20.8	28	4.4	1.8
Sidechain-Sidechain	564	36.9	327	58.0	21.4	25	4.4	1.6

[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



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

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[6] (Å)	Median (Å)
	IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total				
1	106	35	33	36	2	212	0.54	1.75	0.42	0.42
2	112	41	36	46	3	238	0.53	2.18	0.45	0.35
3	125	44	32	50	0	251	0.62	3.09	0.54	0.43
4	115	33	26	40	4	218	0.57	2.02	0.44	0.43
5	110	42	33	31	1	217	0.48	1.82	0.42	0.29
6	115	41	32	40	2	230	0.48	1.75	0.39	0.35
7	123	43	37	38	0	241	0.54	1.81	0.44	0.38
8	125	41	42	45	3	256	0.56	2.13	0.44	0.41
9	130	38	31	48	1	248	0.54	2.12	0.44	0.4
10	111	39	31	40	2	223	0.56	1.77	0.45	0.41

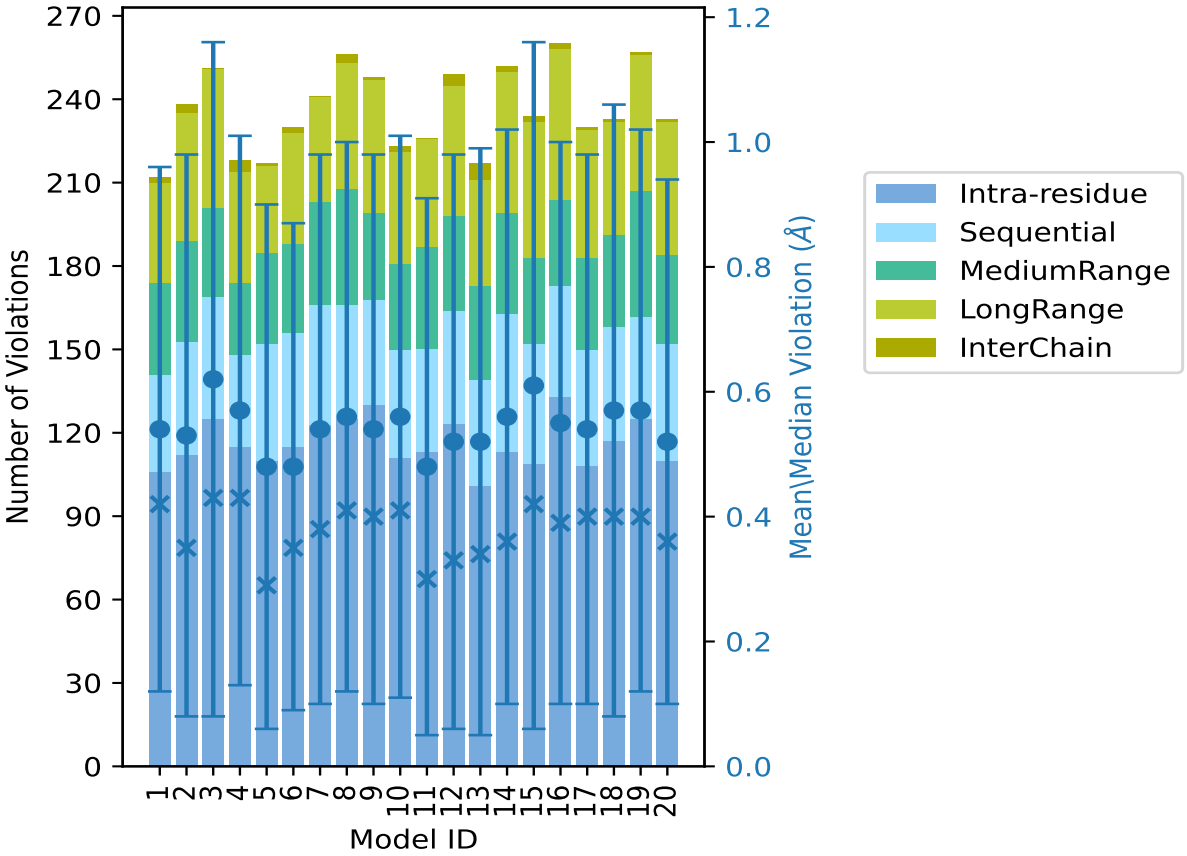
Continued on next page...

Continued from previous page...

Model ID	Number of violations						Mean (Å)	Max (Å)	SD[6] (Å)	Median (Å)
	IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total				
11	113	37	37	39	0	226	0.48	2.87	0.43	0.3
12	123	41	34	47	4	249	0.52	2.37	0.46	0.33
13	101	38	34	38	6	217	0.52	3.13	0.47	0.34
14	113	50	36	51	2	252	0.56	2.26	0.46	0.36
15	109	43	31	49	2	234	0.61	3.45	0.55	0.42
16	133	40	31	54	2	260	0.55	2.69	0.45	0.39
17	108	42	33	46	1	230	0.54	1.84	0.44	0.4
18	117	41	33	41	1	233	0.57	3.19	0.49	0.4
19	125	37	45	49	1	257	0.57	2.01	0.45	0.4
20	110	42	32	48	1	233	0.52	2.02	0.42	0.36

[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



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

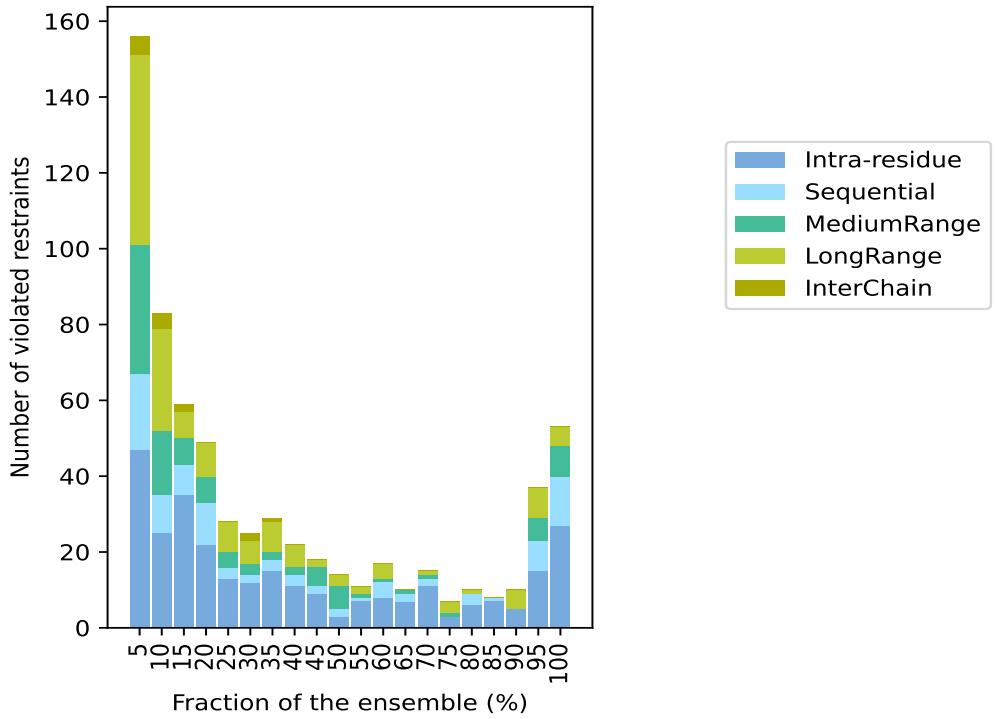
9.3 Distance violation statistics for the ensemble (i)

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, 747(IR:259, SQ:200, MR:144, LR:136, IC:8) restraints are not violated in the ensemble.

Number of violated restraints						Fraction of the ensemble	
IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total	Count[6]	%
47	20	34	50	5	156	1	5.0
25	10	17	27	4	83	2	10.0
35	8	7	7	2	59	3	15.0
22	11	7	9	0	49	4	20.0
13	3	4	8	0	28	5	25.0
12	2	3	6	2	25	6	30.0
15	3	2	8	1	29	7	35.0
11	3	2	6	0	22	8	40.0
9	2	5	2	0	18	9	45.0
3	2	6	3	0	14	10	50.0
7	1	1	2	0	11	11	55.0
8	4	1	4	0	17	12	60.0
7	2	1	0	0	10	13	65.0
11	2	1	1	0	15	14	70.0
3	0	1	3	0	7	15	75.0
6	3	0	1	0	10	16	80.0
7	1	0	0	0	8	17	85.0
5	0	0	5	0	10	18	90.0
15	8	6	8	0	37	19	95.0
27	13	8	5	0	53	20	100.0

[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

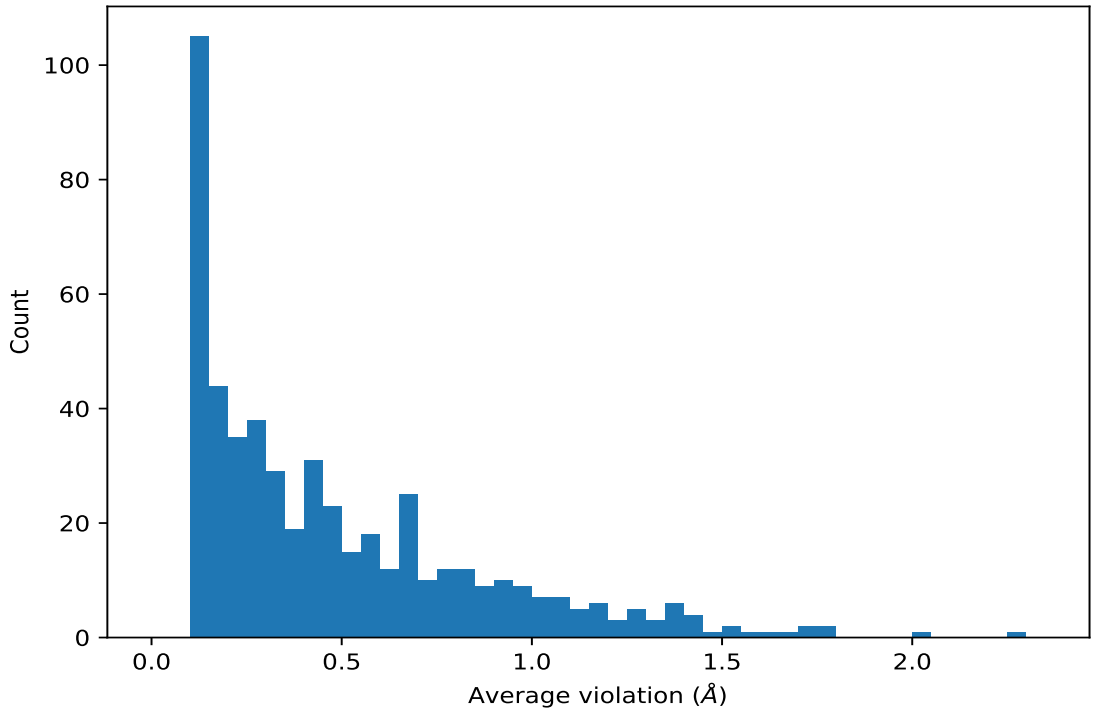
9.3.1 Bar graph : Distance violation statistics for the ensemble



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

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.

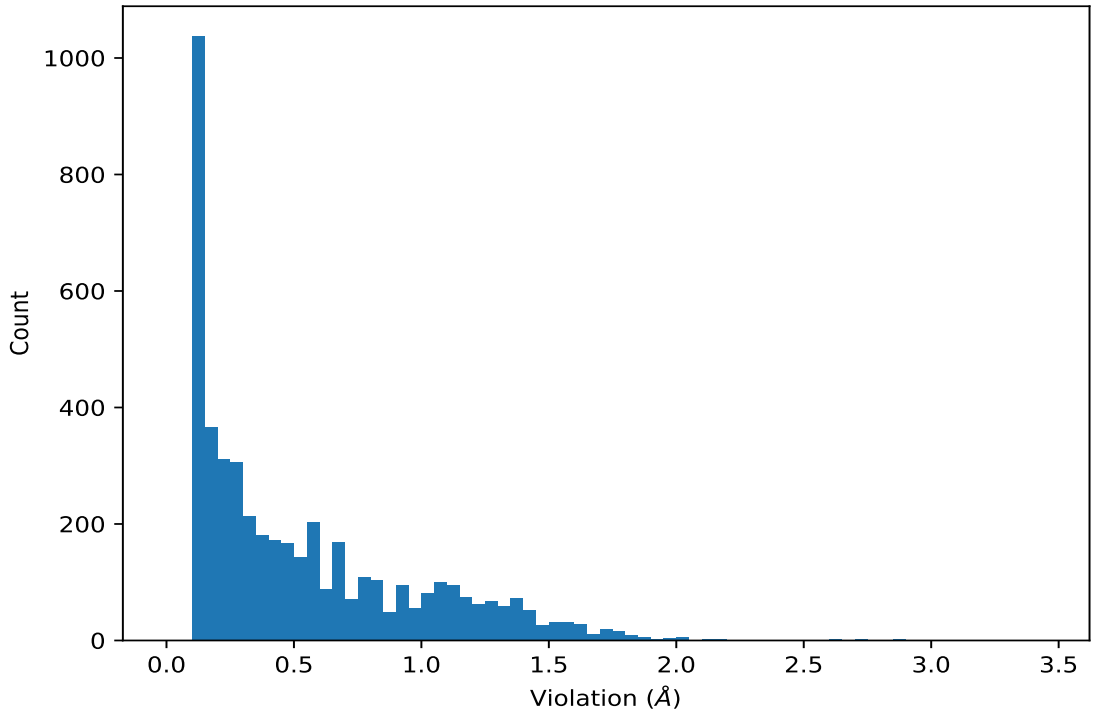
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,98)	1:100:A:GLY:HA3	1:72:A:PHE:HE1	20	1.75	0.06	1.75
(1,495)	1:54:A:LYS:HG2	1:58:A:VAL:HG11	20	1.61	0.35	1.52
(1,246)	1:87:A:THR:HB	1:88:A:VAL:HG21	20	1.54	0.33	1.62
(1,803)	1:28:A:PHE:HD1	1:23:A:ASP:HB2	20	1.38	0.2	1.38
(1,830)	1:97:A:LEU:HB2	1:94:A:LYS:HA	20	1.37	0.25	1.42
(1,614)	1:47:A:PRO:HB2	1:44:A:ILE:HG21	20	1.34	0.12	1.38
(1,1024)	1:88:A:VAL:H	1:88:A:VAL:HG21	20	1.29	0.26	1.35
(1,683)	1:94:A:LYS:HB2	1:91:A:LYS:HA	20	1.27	0.28	1.33
(1,364)	1:54:A:LYS:HB2	1:16:A:LEU:HD11	20	1.25	0.28	1.28
(1,1111)	1:46:A:HIS:H	1:45:A:LYS:HB3	20	1.22	0.07	1.21

[1]Number of violated models, [2]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)
(1,797)	1:68:A:PHE:HD1	1:19:A:LEU:HD11	15	3.45
(1,755)	1:11:A:LEU:HD21	1:8:A:ILE:HB	18	3.19
(1,878)	2:210:B:LEU:HD11	1:31:A:PRO:HA	13	3.13
(1,495)	1:54:A:LYS:HG2	1:58:A:VAL:HG11	3	3.09
(1,499)	1:54:A:LYS:HE3	1:58:A:VAL:HG11	3	2.89
(1,165)	1:27:A:PHE:HD2	1:89:A:TYR:HE2	11	2.87
(1,497)	1:54:A:LYS:HD3	1:58:A:VAL:HG11	3	2.74
(1,699)	1:110:A:LEU:HD11	1:116:A:THR:HB	15	2.71
(1,151)	1:40:A:TYR:HB3	1:82:A:TYR:HE2	16	2.69
(1,799)	1:53:A:MET:HB3	1:19:A:LEU:HD11	15	2.63

10 Dihedral-angle violation analysis

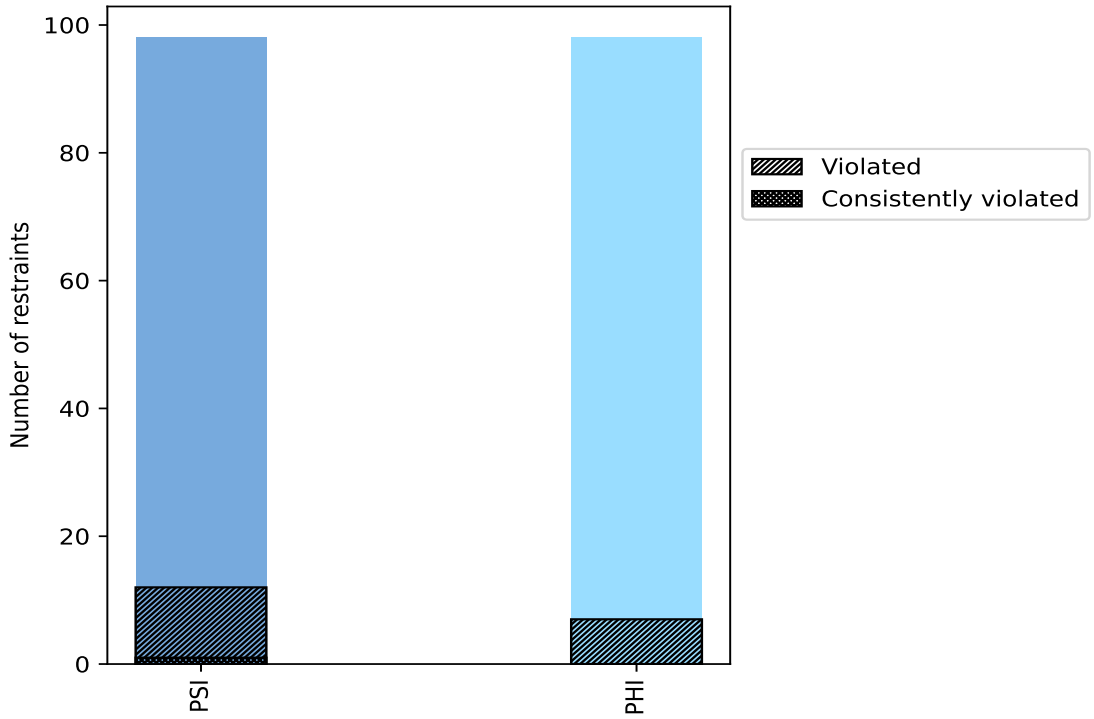
10.1 Summary of dihedral-angle violations

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	%[1]	Violated[3]			Consistently Violated[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
PSI	98	50.0	12	12.2	6.1	1	1.0	0.5
PHI	98	50.0	7	7.1	3.6	0	0.0	0.0
Total	196	100.0	19	9.7	9.7	1	0.5	0.5

[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



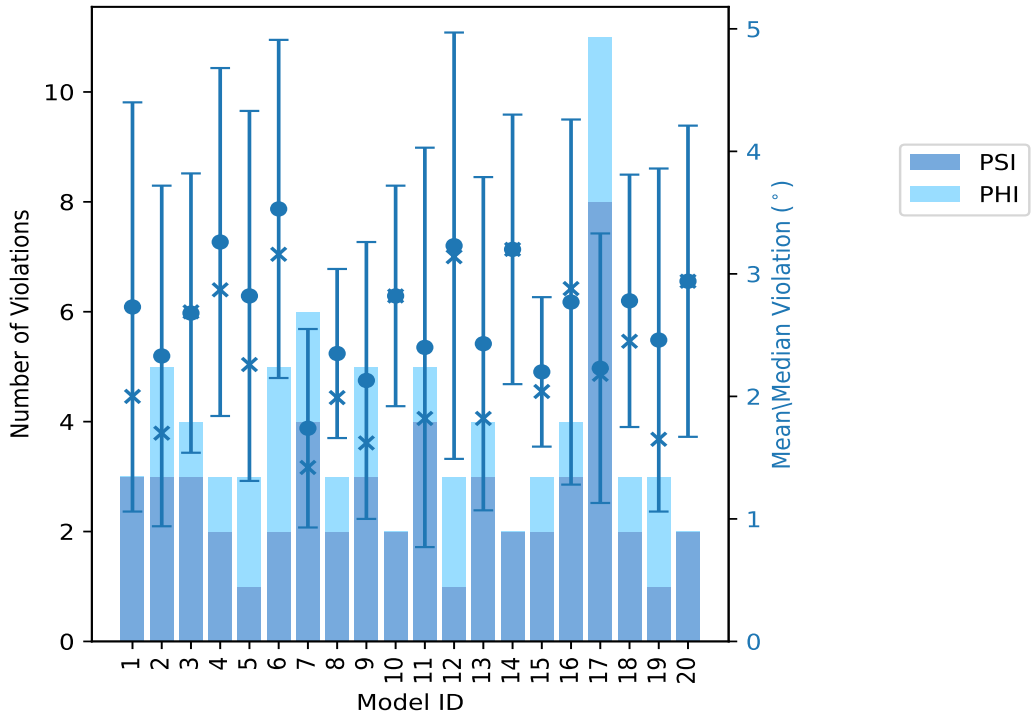
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

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

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

Model ID	Number of violations			Mean (°)	Max (°)	SD (°)	Median (°)
	PSI	PHI	Total				
1	3	0	3	2.73	5.04	1.67	2.0
2	3	2	5	2.33	5.05	1.39	1.7
3	3	1	4	2.68	4.08	1.14	2.69
4	2	1	3	3.26	5.17	1.42	2.87
5	1	2	3	2.82	4.89	1.51	2.26
6	2	3	5	3.53	5.17	1.38	3.16
7	4	2	6	1.74	3.43	0.81	1.42
8	2	1	3	2.35	3.31	0.69	1.99
9	3	2	5	2.13	4.16	1.13	1.62
10	2	0	2	2.82	3.71	0.9	2.82
11	4	1	5	2.4	5.49	1.63	1.82
12	1	2	3	3.23	5.4	1.74	3.14
13	3	1	4	2.43	4.74	1.36	1.82
14	2	0	2	3.2	4.29	1.1	3.2
15	2	1	3	2.2	3.02	0.61	2.04
16	3	1	4	2.77	4.28	1.49	2.88
17	8	3	11	2.23	4.75	1.1	2.18
18	2	1	3	2.78	4.17	1.03	2.45
19	1	2	3	2.46	4.44	1.4	1.65
20	2	0	2	2.94	4.21	1.27	2.94

10.2.1 Bar graph : Dihedral violation statistics for each model



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

10.3 Dihedral-angle violation statistics for the ensemble

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

Number of violated restraints			Fraction of the ensemble	
PSI	PHI	Total	Count[1]	%
4	2	6	1	5.0
2	1	3	2	10.0
3	0	3	3	15.0
0	1	1	4	20.0
0	1	1	5	25.0
0	1	1	6	30.0
1	1	2	7	35.0
0	0	0	8	40.0
1	0	1	9	45.0
0	0	0	10	50.0
0	0	0	11	55.0

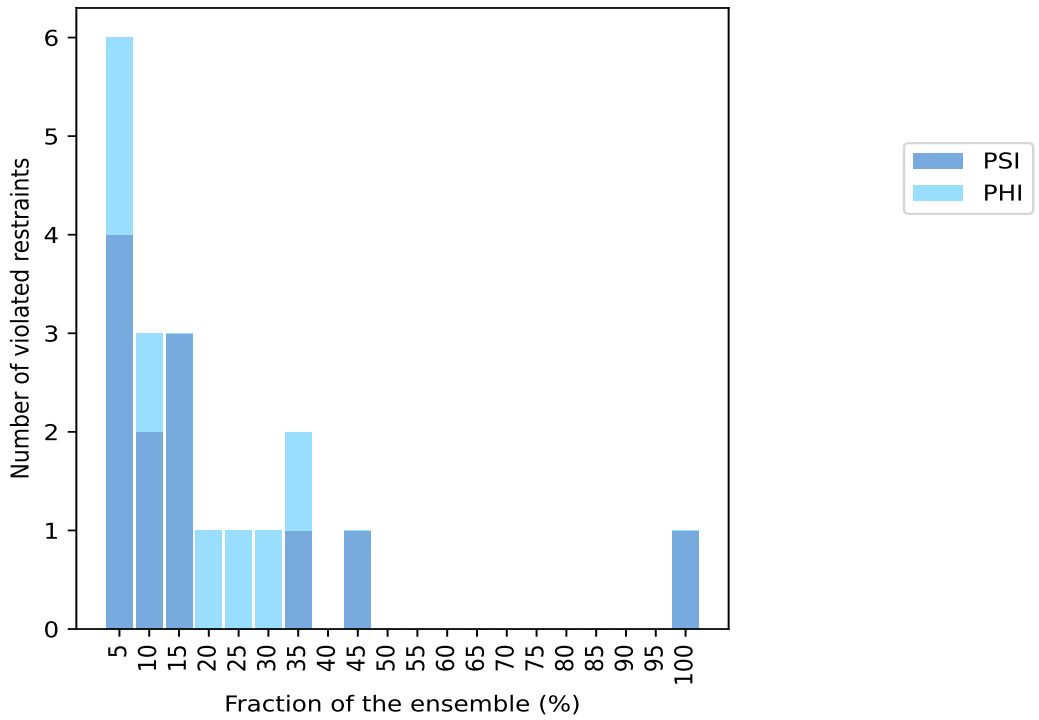
Continued on next page...

Continued from previous page...

Number of violated restraints			Fraction of the ensemble	
PSI	PHI	Total	Count[1]	%
0	0	0	12	60.0
0	0	0	13	65.0
0	0	0	14	70.0
0	0	0	15	75.0
0	0	0	16	80.0
0	0	0	17	85.0
0	0	0	18	90.0
0	0	0	19	95.0
1	0	1	20	100.0

[1] Number of models with violations

10.3.1 Bar graph : Dihedral-angle Violation statistics for the ensemble

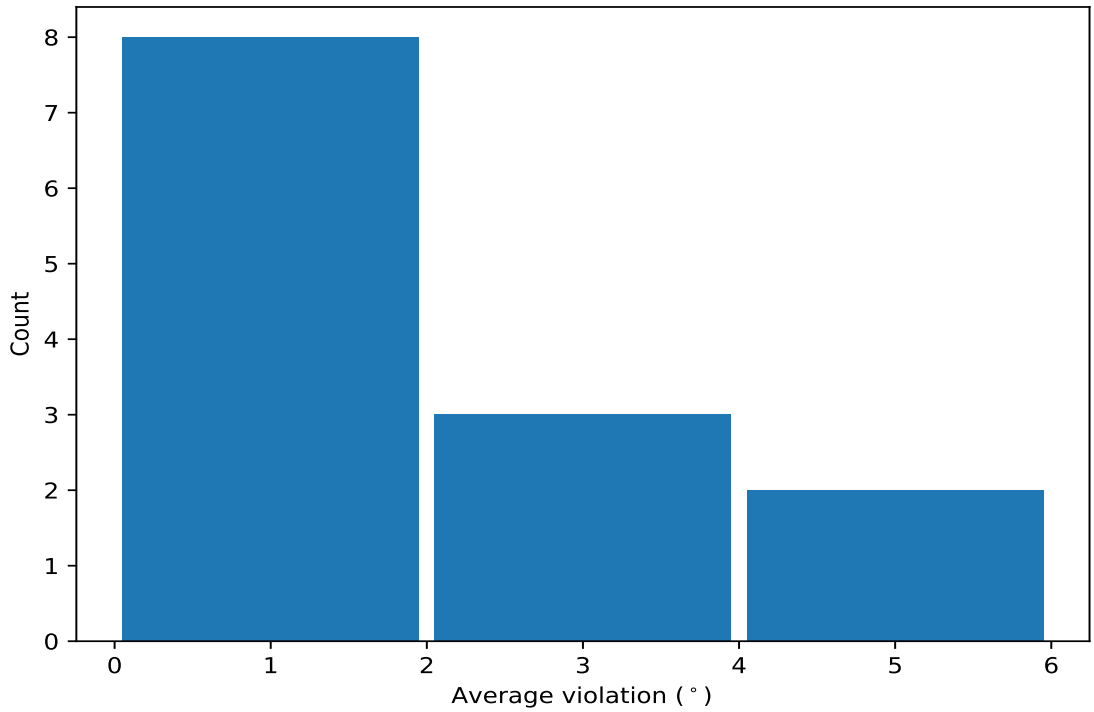


10.4 Most violated dihedral-angle restraints in the ensemble

10.4.1 Histogram : Distribution of mean dihedral-angle 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



10.4.2 Table: Most violated dihedral-angle restraints

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.

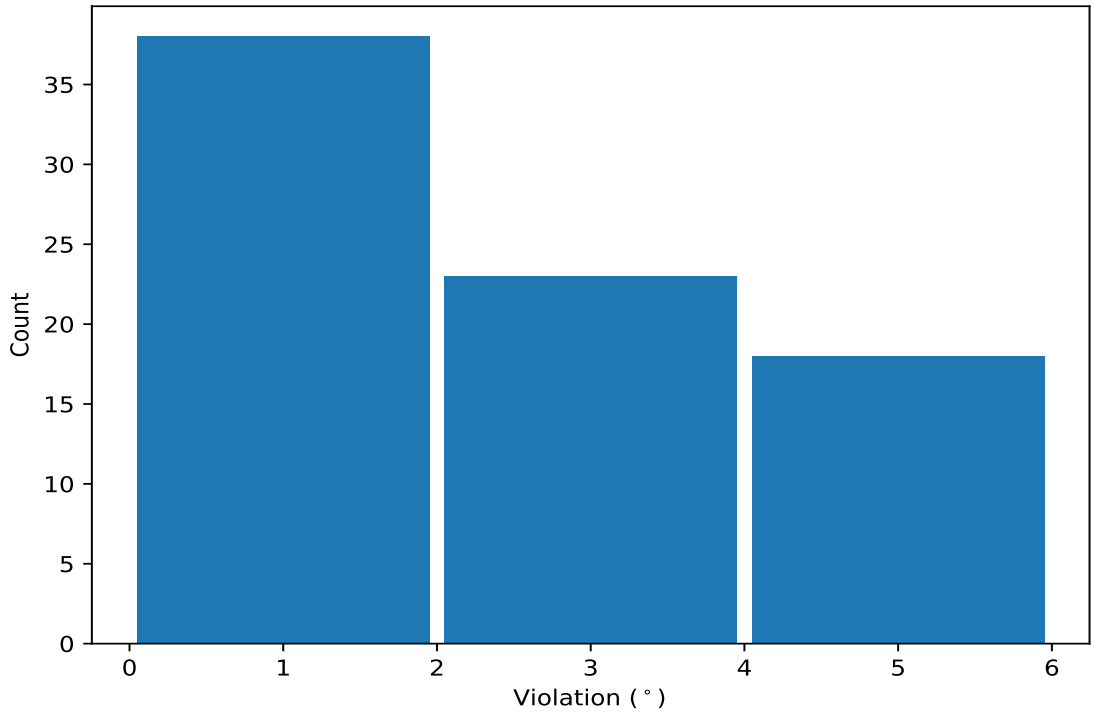
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models[1]	Mean	SD[2]	Median
(1,38)	1:26:A:GLY:N	1:26:A:GLY:CA	1:26:A:GLY:C	1:27:A:PHE:N	20	4.43	0.68	4.36
(1,74)	1:47:A:PRO:N	1:47:A:PRO:CA	1:47:A:PRO:C	1:48:A:MET:N	9	1.76	0.41	1.62
(1,2)	1:7:A:PRO:N	1:7:A:PRO:CA	1:7:A:PRO:C	1:8:A:ILE:N	7	1.82	0.83	1.32
(1,181)	1:103:A:MET:C	1:104:A:MET:N	1:104:A:MET:CA	1:104:A:MET:C	7	1.48	0.5	1.3
(1,75)	1:47:A:PRO:C	1:48:A:MET:N	1:48:A:MET:CA	1:48:A:MET:C	6	1.78	0.34	1.78
(1,77)	1:48:A:MET:C	1:49:A:ASP:N	1:49:A:ASP:CA	1:49:A:ASP:C	5	2.12	0.63	2.17
(1,37)	1:25:A:HIS:C	1:26:A:GLY:N	1:26:A:GLY:CA	1:26:A:GLY:C	4	1.54	0.23	1.64
(1,54)	1:34:A:ASP:N	1:34:A:ASP:CA	1:34:A:ASP:C	1:35:A:ALA:N	3	2.08	0.46	1.82
(1,182)	1:104:A:MET:N	1:104:A:MET:CA	1:104:A:MET:C	1:105:A:SER:N	3	1.91	0.6	1.99
(1,102)	1:61:A:GLU:N	1:61:A:GLU:CA	1:61:A:GLU:C	1:62:A:TYR:N	3	1.82	0.27	1.92

[1] Number of violated models, [2]Standard deviation, All angle values are in degree (°)

10.5 All violated dihedral-angle restraints

10.5.1 Histogram : Distribution of violations

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

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 (°)
(1,38)	1:26:A:GLY:N	1:26:A:GLY:CA	1:26:A:GLY:C	1:27:A:PHE:N	11	5.49
(1,38)	1:26:A:GLY:N	1:26:A:GLY:CA	1:26:A:GLY:C	1:27:A:PHE:N	12	5.4
(1,195)	1:110:A:LEU:C	1:111:A:LEU:N	1:111:A:LEU:CA	1:111:A:LEU:C	6	5.17
(1,38)	1:26:A:GLY:N	1:26:A:GLY:CA	1:26:A:GLY:C	1:27:A:PHE:N	4	5.17
(1,38)	1:26:A:GLY:N	1:26:A:GLY:CA	1:26:A:GLY:C	1:27:A:PHE:N	6	5.1
(1,38)	1:26:A:GLY:N	1:26:A:GLY:CA	1:26:A:GLY:C	1:27:A:PHE:N	2	5.05
(1,38)	1:26:A:GLY:N	1:26:A:GLY:CA	1:26:A:GLY:C	1:27:A:PHE:N	1	5.04
(1,38)	1:26:A:GLY:N	1:26:A:GLY:CA	1:26:A:GLY:C	1:27:A:PHE:N	5	4.89
(1,38)	1:26:A:GLY:N	1:26:A:GLY:CA	1:26:A:GLY:C	1:27:A:PHE:N	17	4.75
(1,38)	1:26:A:GLY:N	1:26:A:GLY:CA	1:26:A:GLY:C	1:27:A:PHE:N	13	4.74