wwPDB NMR Structure Validation Summary Report (i)

Oct 2, 2023 – 10:12 am BST

PDB ID	:	7PVC
BMRB ID	:	25839
Title	:	The structure of Kbp.K from E. coli with potassium bound.
Authors	:	Smith, B.O.
Deposited on	:	2021-10-01

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
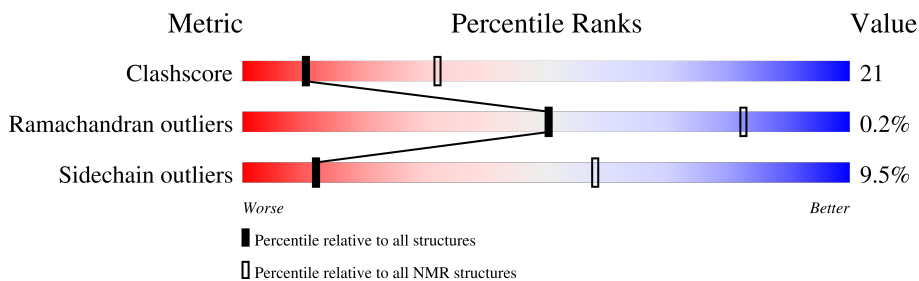http://www.wwpdb.org/validation/2017/FAQs#types.

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 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.

Metric	Percentile Ranks	Value
Clashscore		21
Ramachandran outliers		0.2%
Sidechain outliers		9.5%

Worse   Better

▮ Percentile relative to all structures
▯ Percentile relative to all NMR structures

Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=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	157	50% 35% · 13%

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: *closest to the average*.

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:15, A:25-A:148 (137)	0.51	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 3 clusters and 2 single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2413 atoms, of which 1204 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Potassium binding protein Kbp.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	157	2412	757	1204	211	238	2	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	150	LEU	-	expression tag	UNP P0ADE6
A	151	GLU	-	expression tag	UNP P0ADE6
A	152	HIS	-	expression tag	UNP P0ADE6
A	153	HIS	-	expression tag	UNP P0ADE6
A	154	HIS	-	expression tag	UNP P0ADE6
A	155	HIS	-	expression tag	UNP P0ADE6
A	156	HIS	-	expression tag	UNP P0ADE6
A	157	HIS	-	expression tag	UNP P0ADE6

- Molecule 2 is POTASSIUM ION (three-letter code: K) (formula: K) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	
			Total	K
2	A	1	1	1

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.

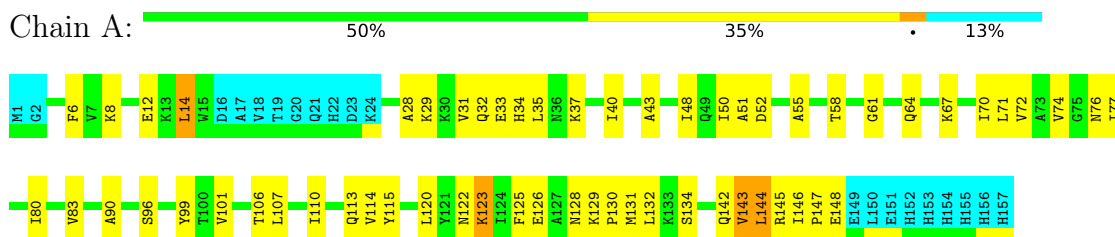• Molecule 1: Potassium binding protein Kbp
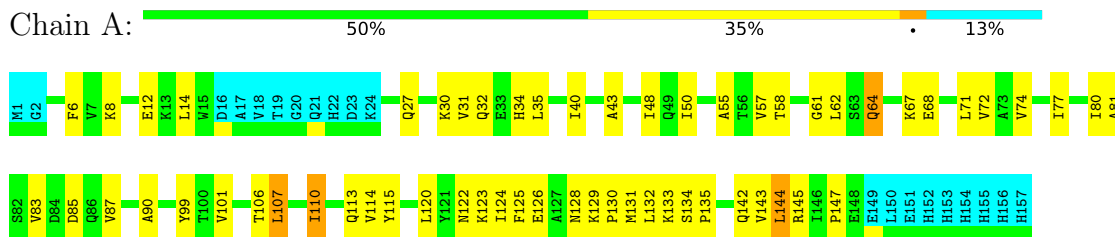
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

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

• Molecule 1: Potassium binding protein Kbp

5 Refinement protocol and experimental data overview ⓘ

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

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *structures with the least restraint violations.*

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

Software name	Classification	Version
ARIA	structure calculation	2.3
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	1879
Number of shifts mapped to atoms	1879
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 ⓘ

6.1 Standard geometry ⓘ

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

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the (average) root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	0.48±0.05	1±1/1059 (0.1± 0.1%)	0.51±0.02	0±0/1434 (0.0± 0.0%)
All	All	0.48	11/21180 (0.1%)	0.51	0/28680 (0.0%)

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

Mol	Chain	Chirality	Planarity
1	A	0.0±0.0	0.1±0.2
All	All	0	1

5 of 6 unique bond outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	115	TYR	CE1-CZ	-8.93	1.26	1.38	13	4
1	A	115	TYR	CE2-CZ	7.98	1.49	1.38	13	3
1	A	125	PHE	CE1-CZ	6.90	1.50	1.37	19	1
1	A	125	PHE	CE2-CZ	-5.89	1.26	1.37	19	1
1	A	6	PHE	CE1-CZ	5.32	1.47	1.37	5	1

There are no bond-angle outliers.

There are no chirality outliers.

All unique planar outliers are listed below.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	121	TYR	Sidechain	1

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	1042	1058	1057	45±6
All	All	20860	21160	21140	891

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 21.

5 of 261 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Total
1:A:7:VAL:HG11	1:A:10:ALA:HB3	0.94	1.37	8	5
1:A:35:LEU:HB3	1:A:43:ALA:HB1	0.87	1.47	19	10
1:A:32:GLN:NE2	1:A:48:ILE:HB	0.80	1.90	1	5
1:A:107:LEU:HD23	1:A:135:PRO:HA	0.79	1.55	16	5
1:A:125:PHE:CE1	1:A:129:LYS:HD3	0.78	2.14	2	2

6.3   Torsion angles ⓘ

6.3.1   Protein backbone ⓘ

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	137/157 (87%)	127±2 (93±2%)	10±2 (7±2%)	0±0 (0±0%)	50	82
All	All	2740/3140 (87%)	2538 (93%)	196 (7%)	6 (0%)	50	82

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	140	PRO	3
1	A	103	SER	1
1	A	85	ASP	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	53	GLY	1

6.3.2 Protein sidechains ⓘ

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	113/130 (87%)	102±2 (91±2%)	11±2 (9±2%)	12	58
All	All	2260/2600 (87%)	2046 (91%)	214 (9%)	12	58

5 of 43 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	144	LEU	17
1	A	64	GLN	16
1	A	107	LEU	14
1	A	58	THR	13
1	A	12	GLU	12

6.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.6 Ligand geometry ⓘ

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 ⓘ

The completeness of assignment taking into account all chemical shift lists is 95% for the well-defined parts and 91% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *nef_chemical_shift_list_K*

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	1879
Number of shifts mapped to atoms	1879
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 ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_\alpha$	153	-0.25 ± 0.18	None needed (< 0.5 ppm)
$^{13}C_\beta$	141	0.03 ± 0.11	None needed (< 0.5 ppm)
$^{13}C'$	153	-0.18 ± 0.14	None needed (< 0.5 ppm)
^{15}N	141	0.29 ± 0.18	None needed (< 0.5 ppm)

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 95%, i.e. 1715 atoms were assigned a chemical shift out of a possible 1806. 0 out of 20 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	676/683 (99%)	274/278 (99%)	274/274 (100%)	128/131 (98%)
Sidechain	959/1027 (93%)	656/666 (98%)	286/326 (88%)	17/35 (49%)

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	80/96 (83%)	40/46 (87%)	39/47 (83%)	1/3 (33%)
Overall	1715/1806 (95%)	970/990 (98%)	599/647 (93%)	146/169 (86%)

### 7.1.4   Statistically unusual chemical shifts   ⓘ

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	75	GLY	HA3	0.74	$2.08 - 5.71$	-8.7
1	A	27	GLN	HG3	0.24	$0.91 - 3.68$	-7.4
1	A	142	GLN	HB3	0.24	$0.71 - 3.33$	-6.8
1	A	142	GLN	NE2	103.33	$103.38 - 120.35$	-5.0

7.1.5   Random Coil Index (RCI) plots   ⓘ

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

Random coil index (RCI) for chain A:
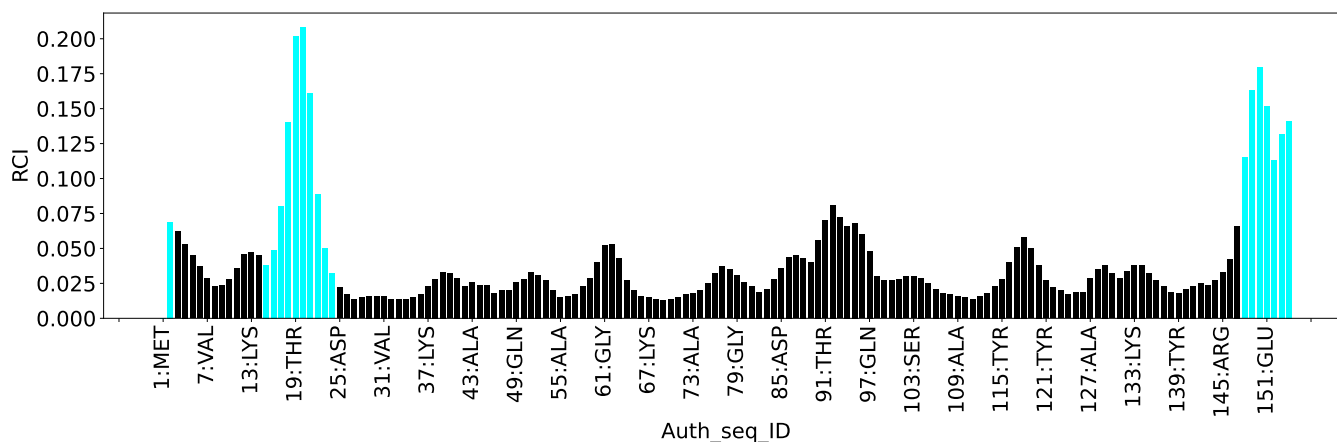
8 NMR restraints analysis ⓘ

8.1 Conformationally restricting restraints ⓘ

The following table provides the summary of experimentally observed NMR restraints in different categories. Restraints are classified into different categories based on the sequence separation of the atoms involved.

Description	Value
Total distance restraints	5562
Intra-residue ($\vert$i-j$\vert$=0)	1786
Sequential ($\vert$i-j$\vert$=1)	1086
Medium range ($\vert$i-j$\vert$>1 and $\vert$i-j$\vert$<5)	929
Long range ($\vert$i-j$\vert \geq$5)	1756
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	162
Number of unmapped restraints	0
Number of restraints per residue	36.2
Number of long range restraints per residue[1]	11.1

[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)	152.4	0.2
0.2-0.5 (Medium)	156.3	0.5
>0.5 (Large)	121.2	3.14

8.2.2 Average number of dihedral-angle violations per model ⓘ

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)	5.0	9.96
10.0-20.0 (Medium)	2.9	19.94
>20.0 (Large)	7.2	137.44

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.

Restraints type	Count	%[1]	Violated[3]			Consistently Violated[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
Intra-residue (|i-j|=0)	1786	32.1	262	14.7	4.7	34	1.9	0.6
Backbone-Backbone	209	3.8	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	1020	18.3	225	22.1	4.0	23	2.3	0.4
Sidechain-Sidechain	557	10.0	37	6.6	0.7	11	2.0	0.2
Sequential (|i-j|=1)	1086	19.5	205	18.9	3.7	18	1.7	0.3
Backbone-Backbone	436	7.8	17	3.9	0.3	2	0.5	0.0
Backbone-Sidechain	567	10.2	157	27.7	2.8	13	2.3	0.2
Sidechain-Sidechain	83	1.5	31	37.3	0.6	3	3.6	0.1
Medium range (|i-j|>1 & |i-j|<5)	929	16.7	194	20.9	3.5	7	0.8	0.1
Backbone-Backbone	261	4.7	7	2.7	0.1	0	0.0	0.0
Backbone-Sidechain	471	8.5	98	20.8	1.8	3	0.6	0.1
Sidechain-Sidechain	197	3.5	89	45.2	1.6	4	2.0	0.1
Long range (|i-j|≥5)	1756	31.6	478	27.2	8.6	27	1.5	0.5
Backbone-Backbone	223	4.0	19	8.5	0.3	2	0.9	0.0
Backbone-Sidechain	970	17.4	242	24.9	4.4	10	1.0	0.2
Sidechain-Sidechain	563	10.1	217	38.5	3.9	15	2.7	0.3
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	5562	100.0	1139	20.5	20.5	86	1.5	1.5
Backbone-Backbone	1129	20.3	43	3.8	0.8	4	0.4	0.1
Backbone-Sidechain	3028	54.4	722	23.8	13.0	49	1.6	0.9
Sidechain-Sidechain	1405	25.3	374	26.6	6.7	33	2.3	0.6

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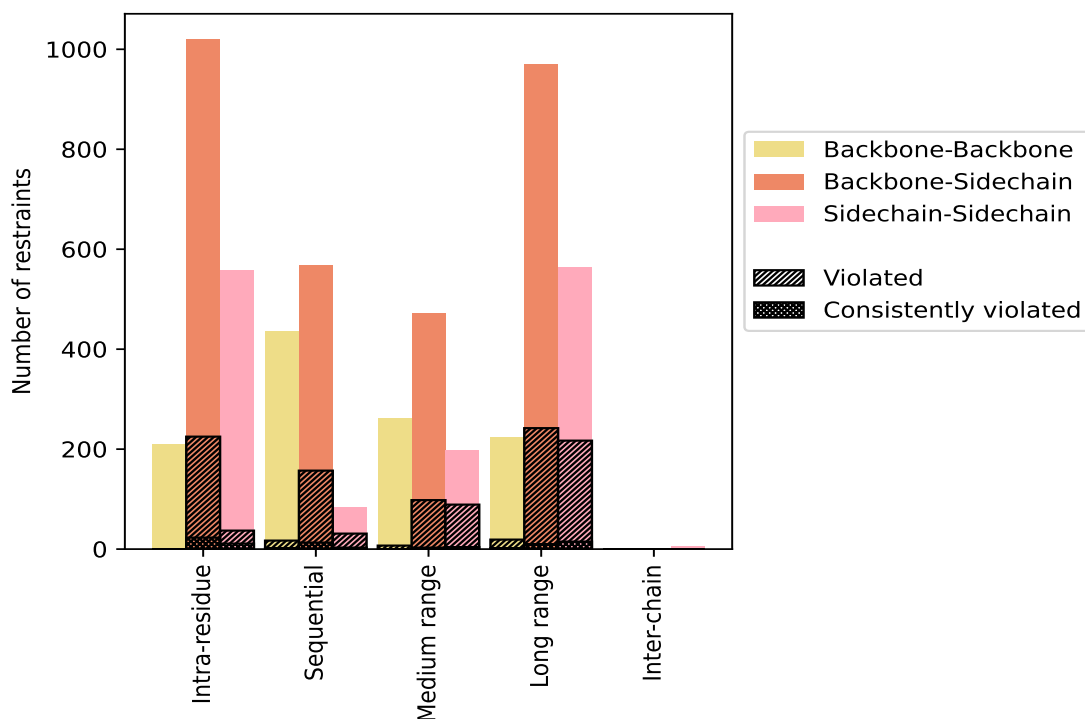
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories. The hydrogen bonds and disulfied bonds are counted in their appropriate category on the x-axis

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	100	79	57	168	0	404	0.42	2.4	0.4	0.25
2	116	92	65	195	0	468	0.47	2.79	0.46	0.28
3	113	79	68	184	0	444	0.43	2.41	0.43	0.26
4	103	92	58	175	0	428	0.46	2.74	0.46	0.27
5	111	91	78	197	0	477	0.5	2.8	0.49	0.29
6	109	90	65	221	0	485	0.54	3.14	0.52	0.3
7	96	84	66	179	0	425	0.45	2.78	0.45	0.28
8	98	75	64	180	0	417	0.41	2.37	0.41	0.24
9	109	82	62	177	0	430	0.43	2.93	0.42	0.26
10	107	90	73	186	0	456	0.46	2.89	0.45	0.27
11	108	79	60	180	0	427	0.46	2.55	0.46	0.25

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				
12	100	80	53	196	0	429	0.45	2.79	0.44	0.27
13	115	90	72	193	0	470	0.44	2.59	0.41	0.27
14	98	78	53	174	0	403	0.44	2.58	0.44	0.26
15	112	87	64	172	0	435	0.39	2.5	0.36	0.25
16	108	83	61	174	0	426	0.4	2.02	0.36	0.25
17	110	84	70	172	0	436	0.39	2.43	0.36	0.24
18	112	90	66	207	0	475	0.49	2.59	0.46	0.3
19	99	77	69	171	0	416	0.42	2.44	0.4	0.26
20	105	80	72	196	0	453	0.45	2.82	0.47	0.25

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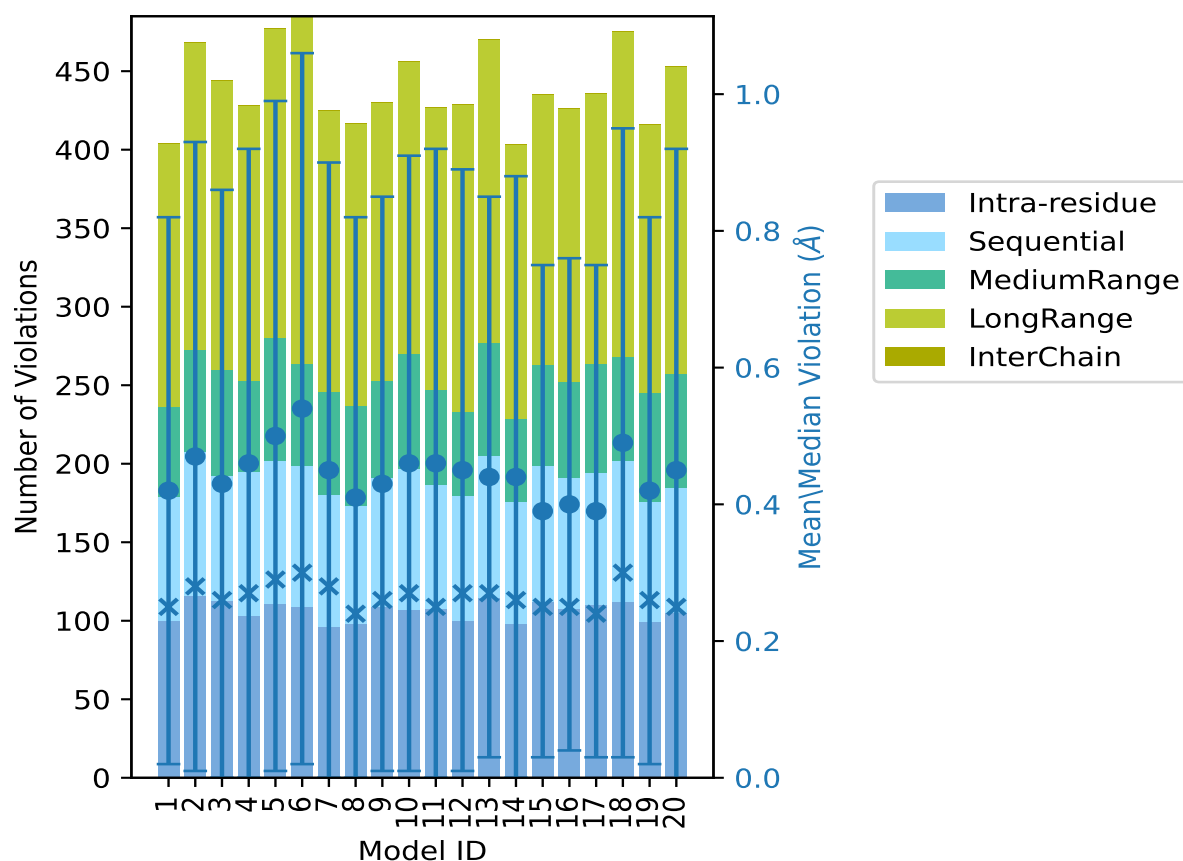
The mean(dot),median(x) and the standard deviation are shown in blue with respect to the y axis on the right

## 9.3 Distance violation statistics for the ensemble ⓘ

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, 4418(IR:1524, SQ:881, MR:735, LR:1278, IC:0) 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]	%
58	40	29	67	0	194	1	5.0
22	20	34	59	0	135	2	10.0
17	14	18	38	0	87	3	15.0
15	16	16	24	0	71	4	20.0
10	11	10	30	0	61	5	25.0
17	6	14	29	0	66	6	30.0
7	4	7	21	0	39	7	35.0
12	10	7	19	0	48	8	40.0
10	11	10	17	0	48	9	45.0
7	5	2	19	0	33	10	50.0
7	4	5	21	0	37	11	55.0
7	4	4	22	0	37	12	60.0
7	4	3	19	0	33	13	65.0
8	3	7	20	0	38	14	70.0
4	6	7	9	0	26	15	75.0
3	6	2	11	0	22	16	80.0
6	6	5	10	0	27	17	85.0
4	9	3	11	0	27	18	90.0
7	8	4	5	0	24	19	95.0
34	18	7	27	0	86	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 ⓘ
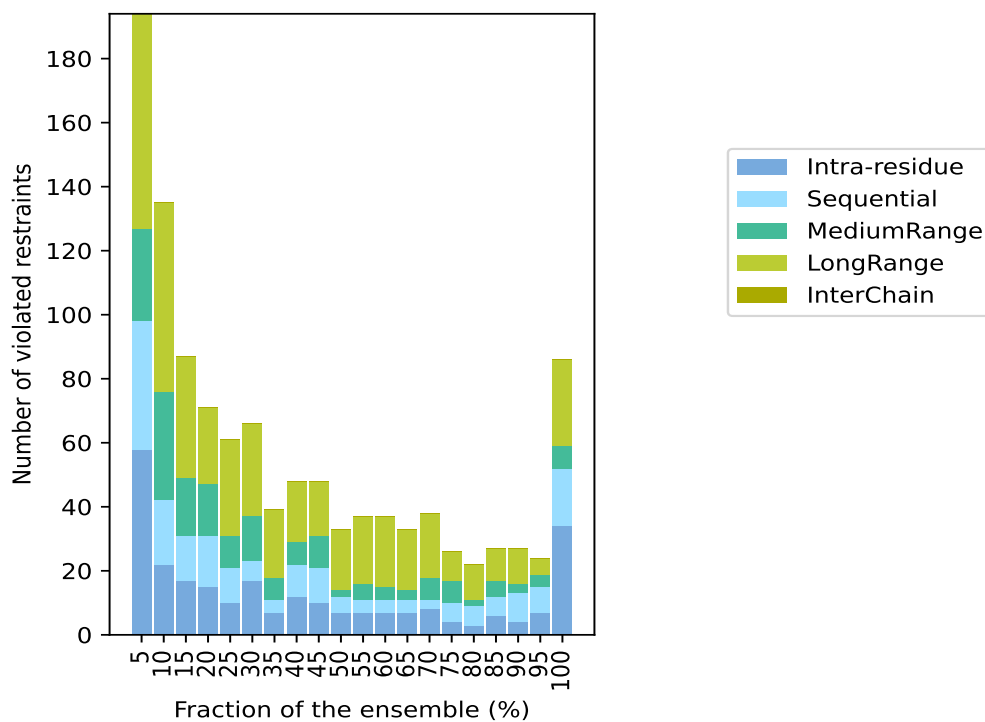
9.4 Most violated distance restraints in the ensemble ⓘ

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
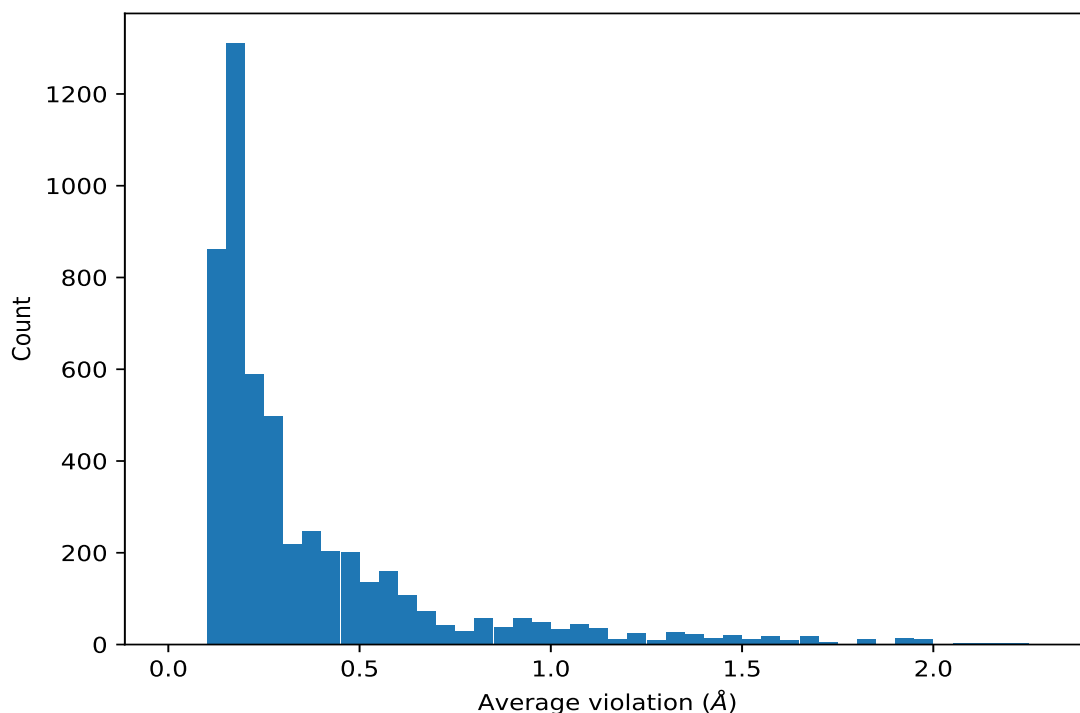
9.4.2 Table: Most violated distance restraints ⓘ

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)
(3,2623)	1:61:A:GLY:HA3	1:90:A:ALA:HA	20	1.79	0.05	1.8
(3,741)	1:98:A:PHE:HB2	1:143:A:VAL:HG21	20	1.7	0.9	2.33
(3,741)	1:98:A:PHE:HB2	1:143:A:VAL:HG22	20	1.7	0.9	2.33
(3,741)	1:98:A:PHE:HB2	1:143:A:VAL:HG23	20	1.7	0.9	2.33
(3,2303)	1:99:A:TYR:HB2	1:144:A:LEU:HD21	20	1.39	0.47	1.25
(3,2303)	1:99:A:TYR:HB2	1:144:A:LEU:HD22	20	1.39	0.47	1.25
(3,2303)	1:99:A:TYR:HB2	1:144:A:LEU:HD23	20	1.39	0.47	1.25
(3,2303)	1:135:A:PRO:HD2	1:144:A:LEU:HD21	20	1.39	0.47	1.25
(3,2303)	1:135:A:PRO:HD2	1:144:A:LEU:HD22	20	1.39	0.47	1.25
(3,2303)	1:135:A:PRO:HD2	1:144:A:LEU:HD23	20	1.39	0.47	1.25
(3,1291)	1:10:A:ALA:HB1	1:101:A:VAL:HG11	20	1.32	0.4	1.54
(3,1291)	1:10:A:ALA:HB1	1:101:A:VAL:HG12	20	1.32	0.4	1.54
(3,1291)	1:10:A:ALA:HB1	1:101:A:VAL:HG13	20	1.32	0.4	1.54
(3,1291)	1:10:A:ALA:HB2	1:101:A:VAL:HG11	20	1.32	0.4	1.54
(3,1291)	1:10:A:ALA:HB2	1:101:A:VAL:HG12	20	1.32	0.4	1.54
(3,1291)	1:10:A:ALA:HB2	1:101:A:VAL:HG13	20	1.32	0.4	1.54

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(3,1291)	1:10:A:ALA:HB3	1:101:A:VAL:HG11	20	1.32	0.4	1.54
(3,1291)	1:10:A:ALA:HB3	1:101:A:VAL:HG12	20	1.32	0.4	1.54
(3,1291)	1:10:A:ALA:HB3	1:101:A:VAL:HG13	20	1.32	0.4	1.54
(3,1291)	1:101:A:VAL:HG11	1:102:A:LYS:HG3	20	1.32	0.4	1.54
(3,1291)	1:101:A:VAL:HG12	1:102:A:LYS:HG3	20	1.32	0.4	1.54
(3,1291)	1:101:A:VAL:HG13	1:102:A:LYS:HG3	20	1.32	0.4	1.54
(3,1291)	1:101:A:VAL:HG11	1:110:A:ILE:HB	20	1.32	0.4	1.54
(3,1291)	1:101:A:VAL:HG12	1:110:A:ILE:HB	20	1.32	0.4	1.54
(3,1291)	1:101:A:VAL:HG13	1:110:A:ILE:HB	20	1.32	0.4	1.54
(3,1291)	1:101:A:VAL:HG11	1:132:A:LEU:HG	20	1.32	0.4	1.54
(3,1291)	1:101:A:VAL:HG12	1:132:A:LEU:HG	20	1.32	0.4	1.54
(3,1291)	1:101:A:VAL:HG13	1:132:A:LEU:HG	20	1.32	0.4	1.54
(3,1291)	1:101:A:VAL:HG11	1:144:A:LEU:HB3	20	1.32	0.4	1.54
(3,1291)	1:101:A:VAL:HG12	1:144:A:LEU:HB3	20	1.32	0.4	1.54
(3,1291)	1:101:A:VAL:HG13	1:144:A:LEU:HB3	20	1.32	0.4	1.54
(3,1107)	1:128:A:ASN:HD21	1:144:A:LEU:HB2	20	1.28	0.33	1.43
(3,692)	1:107:A:LEU:HB2	1:111:A:SER:HB3	20	1.24	0.47	1.48
(3,692)	1:133:A:LYS:HB3	1:134:A:SER:HB3	20	1.24	0.47	1.48
(3,692)	1:134:A:SER:HB3	1:137:A:LYS:HB3	20	1.24	0.47	1.48
(3,1753)	1:14:A:LEU:HD21	1:32:A:GLN:HB3	20	1.1	0.36	0.94
(3,1753)	1:14:A:LEU:HD22	1:32:A:GLN:HB3	20	1.1	0.36	0.94
(3,1753)	1:14:A:LEU:HD23	1:32:A:GLN:HB3	20	1.1	0.36	0.94
(3,1753)	1:14:A:LEU:HD21	1:32:A:GLN:HB2	20	1.1	0.36	0.94
(3,1753)	1:14:A:LEU:HD22	1:32:A:GLN:HB2	20	1.1	0.36	0.94
(3,1753)	1:14:A:LEU:HD23	1:32:A:GLN:HB2	20	1.1	0.36	0.94
(3,1753)	1:14:A:LEU:HD21	1:33:A:GLU:HB3	20	1.1	0.36	0.94
(3,1753)	1:14:A:LEU:HD22	1:33:A:GLU:HB3	20	1.1	0.36	0.94
(3,1753)	1:14:A:LEU:HD23	1:33:A:GLU:HB3	20	1.1	0.36	0.94
(3,1753)	1:14:A:LEU:HD21	1:131:A:MET:HE1	20	1.1	0.36	0.94
(3,1753)	1:14:A:LEU:HD21	1:131:A:MET:HE2	20	1.1	0.36	0.94
(3,1753)	1:14:A:LEU:HD21	1:131:A:MET:HE3	20	1.1	0.36	0.94
(3,1753)	1:14:A:LEU:HD22	1:131:A:MET:HE1	20	1.1	0.36	0.94
(3,1753)	1:14:A:LEU:HD22	1:131:A:MET:HE2	20	1.1	0.36	0.94
(3,1753)	1:14:A:LEU:HD22	1:131:A:MET:HE3	20	1.1	0.36	0.94
(3,1753)	1:14:A:LEU:HD23	1:131:A:MET:HE1	20	1.1	0.36	0.94
(3,1753)	1:14:A:LEU:HD23	1:131:A:MET:HE2	20	1.1	0.36	0.94
(3,1753)	1:14:A:LEU:HD23	1:131:A:MET:HE3	20	1.1	0.36	0.94
(3,1287)	1:101:A:VAL:HG21	1:105:A:ASP:HB3	20	1.08	0.37	1.02
(3,1287)	1:101:A:VAL:HG22	1:105:A:ASP:HB3	20	1.08	0.37	1.02
(3,1287)	1:101:A:VAL:HG23	1:105:A:ASP:HB3	20	1.08	0.37	1.02
(3,1401)	1:29:A:LYS:HG3	1:33:A:GLU:HB2	20	0.79	0.19	0.88
(3,1401)	1:30:A:LYS:HG2	1:33:A:GLU:HB2	20	0.79	0.19	0.88

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(3,1401)	1:33:A:GLU:HB2	1:37:A:LYS:HG2	20	0.79	0.19	0.88
(3,1401)	1:112:A:LYS:HD3	1:113:A:GLN:HB3	20	0.79	0.19	0.88
(3,1401)	1:113:A:GLN:HB3	1:114:A:VAL:HB	20	0.79	0.19	0.88
(3,1401)	1:113:A:GLN:HB2	1:114:A:VAL:HB	20	0.79	0.19	0.88
(3,2280)	1:39:A:GLY:H	1:41:A:PRO:HD2	20	0.74	0.03	0.74

[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.
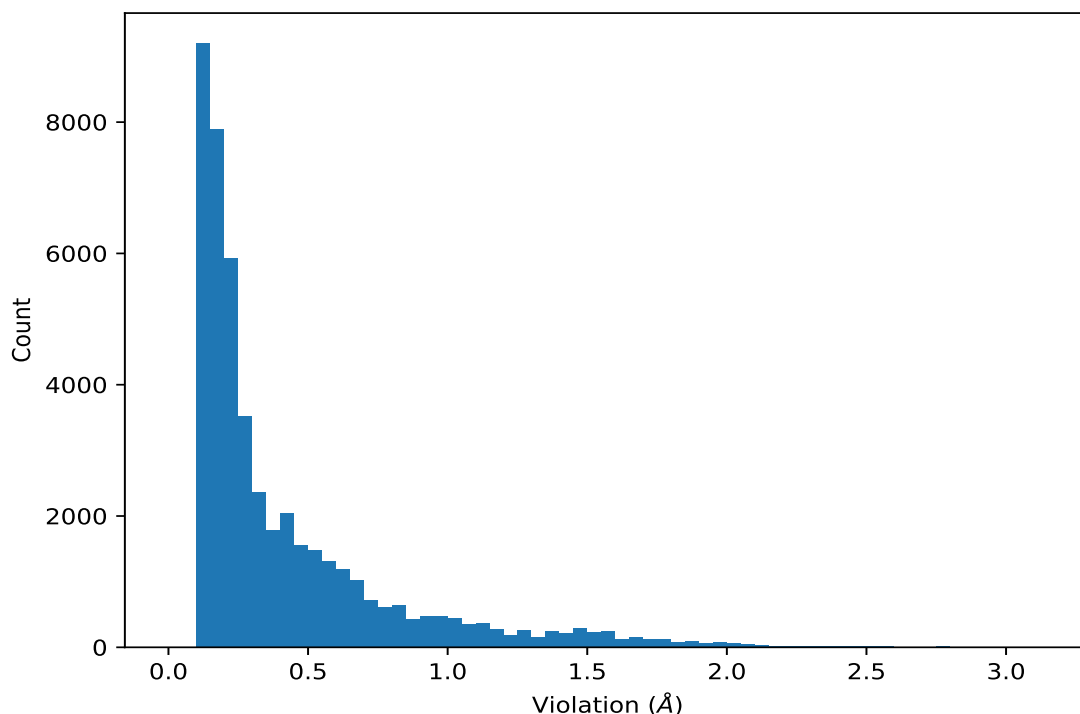
9.5.2 Table : All distance violations (i)

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)
(3,1270)	1:14:A:LEU:HD11	1:31:A:VAL:HG21	6	3.14
(3,1270)	1:14:A:LEU:HD11	1:31:A:VAL:HG22	6	3.14
(3,1270)	1:14:A:LEU:HD11	1:31:A:VAL:HG23	6	3.14
(3,1270)	1:14:A:LEU:HD12	1:31:A:VAL:HG21	6	3.14
(3,1270)	1:14:A:LEU:HD12	1:31:A:VAL:HG22	6	3.14
(3,1270)	1:14:A:LEU:HD12	1:31:A:VAL:HG23	6	3.14
(3,1270)	1:14:A:LEU:HD13	1:31:A:VAL:HG21	6	3.14
(3,1270)	1:14:A:LEU:HD13	1:31:A:VAL:HG22	6	3.14
(3,1270)	1:14:A:LEU:HD13	1:31:A:VAL:HG23	6	3.14
(3,1270)	1:14:A:LEU:HD11	1:31:A:VAL:HG21	9	2.93
(3,1270)	1:14:A:LEU:HD11	1:31:A:VAL:HG22	9	2.93
(3,1270)	1:14:A:LEU:HD11	1:31:A:VAL:HG23	9	2.93
(3,1270)	1:14:A:LEU:HD12	1:31:A:VAL:HG21	9	2.93
(3,1270)	1:14:A:LEU:HD12	1:31:A:VAL:HG22	9	2.93
(3,1270)	1:14:A:LEU:HD12	1:31:A:VAL:HG23	9	2.93
(3,1270)	1:14:A:LEU:HD13	1:31:A:VAL:HG21	9	2.93
(3,1270)	1:14:A:LEU:HD13	1:31:A:VAL:HG22	9	2.93
(3,1270)	1:14:A:LEU:HD13	1:31:A:VAL:HG23	9	2.93
(3,1270)	1:14:A:LEU:HD11	1:31:A:VAL:HG21	10	2.89
(3,1270)	1:14:A:LEU:HD11	1:31:A:VAL:HG22	10	2.89
(3,1270)	1:14:A:LEU:HD11	1:31:A:VAL:HG23	10	2.89
(3,1270)	1:14:A:LEU:HD12	1:31:A:VAL:HG21	10	2.89
(3,1270)	1:14:A:LEU:HD12	1:31:A:VAL:HG22	10	2.89
(3,1270)	1:14:A:LEU:HD12	1:31:A:VAL:HG23	10	2.89
(3,1270)	1:14:A:LEU:HD13	1:31:A:VAL:HG21	10	2.89
(3,1270)	1:14:A:LEU:HD13	1:31:A:VAL:HG22	10	2.89
(3,1270)	1:14:A:LEU:HD13	1:31:A:VAL:HG23	10	2.89
(3,1759)	1:7:A:VAL:HG11	1:9:A:ASP:HB2	20	2.82
(3,1759)	1:7:A:VAL:HG12	1:9:A:ASP:HB2	20	2.82
(3,1759)	1:7:A:VAL:HG13	1:9:A:ASP:HB2	20	2.82
(3,2438)	1:115:A:TYR:HA	1:120:A:LEU:HD21	5	2.8
(3,2438)	1:115:A:TYR:HA	1:120:A:LEU:HD22	5	2.8
(3,2438)	1:115:A:TYR:HA	1:120:A:LEU:HD23	5	2.8
(3,2438)	1:115:A:TYR:HA	1:120:A:LEU:HD21	6	2.79
(3,2438)	1:115:A:TYR:HA	1:120:A:LEU:HD22	6	2.79
(3,2438)	1:115:A:TYR:HA	1:120:A:LEU:HD23	6	2.79
(3,2438)	1:115:A:TYR:HA	1:120:A:LEU:HD21	12	2.79
(3,2438)	1:115:A:TYR:HA	1:120:A:LEU:HD22	12	2.79
(3,2438)	1:115:A:TYR:HA	1:120:A:LEU:HD23	12	2.79
(3,1759)	1:7:A:VAL:HG11	1:9:A:ASP:HB2	2	2.79
(3,1759)	1:7:A:VAL:HG12	1:9:A:ASP:HB2	2	2.79
(3,1759)	1:7:A:VAL:HG13	1:9:A:ASP:HB2	2	2.79
(3,2438)	1:115:A:TYR:HA	1:120:A:LEU:HD21	7	2.78

Continued on next page...

Continued from previous page...

| Key	Atom-1	Atom-2	Model ID	Violation (Å)
(3,2438)	1:115:A:TYR:HA	1:120:A:LEU:HD22	7	2.78
(3,2438)	1:115:A:TYR:HA	1:120:A:LEU:HD23	7	2.78
(3,2438)	1:115:A:TYR:HA	1:120:A:LEU:HD21	4	2.74

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	$\%^2$	$\%^1$
PHI	81	50.0	17	21.0	10.5	2	2.5	1.2
PSI	81	50.0	19	23.5	11.7	5	6.2	3.1
Total	162	100.0	36	22.2	22.2	7	4.3	4.3

[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 ⓘ
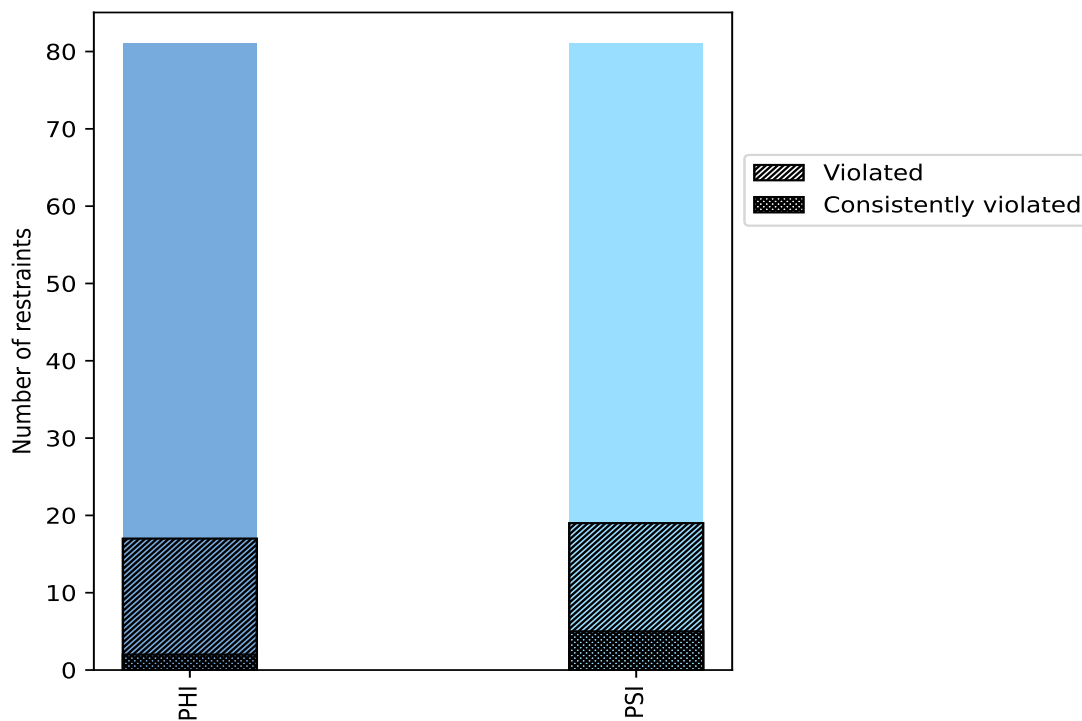

Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

10.2 Dihedral-angle violation statistics for each model ⓘ

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 (°)
	PHI	PSI	Total				
1	7	10	17	36.3	119.72	41.15	18.86
2	6	8	14	42.08	132.49	39.28	28.1
3	8	7	15	41.66	116.02	40.57	24.66
4	8	7	15	38.76	126.11	42.9	14.13
5	7	9	16	40.82	124.39	40.64	21.97
6	8	8	16	36.57	122.99	38.56	22.49
7	6	6	12	49.48	137.13	47.51	20.09
8	12	7	19	30.98	113.79	37.04	8.44
9	6	6	12	46.48	128.19	45.07	25.48
10	7	5	12	40.0	112.04	41.81	18.14
11	5	7	12	45.6	112.76	41.44	20.47
12	8	8	16	39.69	118.26	42.34	15.46
13	7	7	14	43.78	135.31	44.28	20.24
14	9	8	17	41.84	132.18	41.41	26.29
15	9	7	16	42.42	126.99	43.43	17.81
16	10	7	17	36.0	130.3	41.84	12.88
17	8	8	16	37.97	137.44	42.45	18.34
18	8	7	15	42.6	120.52	42.04	19.89
19	8	6	14	39.57	116.24	38.73	19.0
20	7	8	15	38.71	126.62	42.32	15.56

10.2.1 Bar graph : Dihedral violation statistics for each model ⓘ
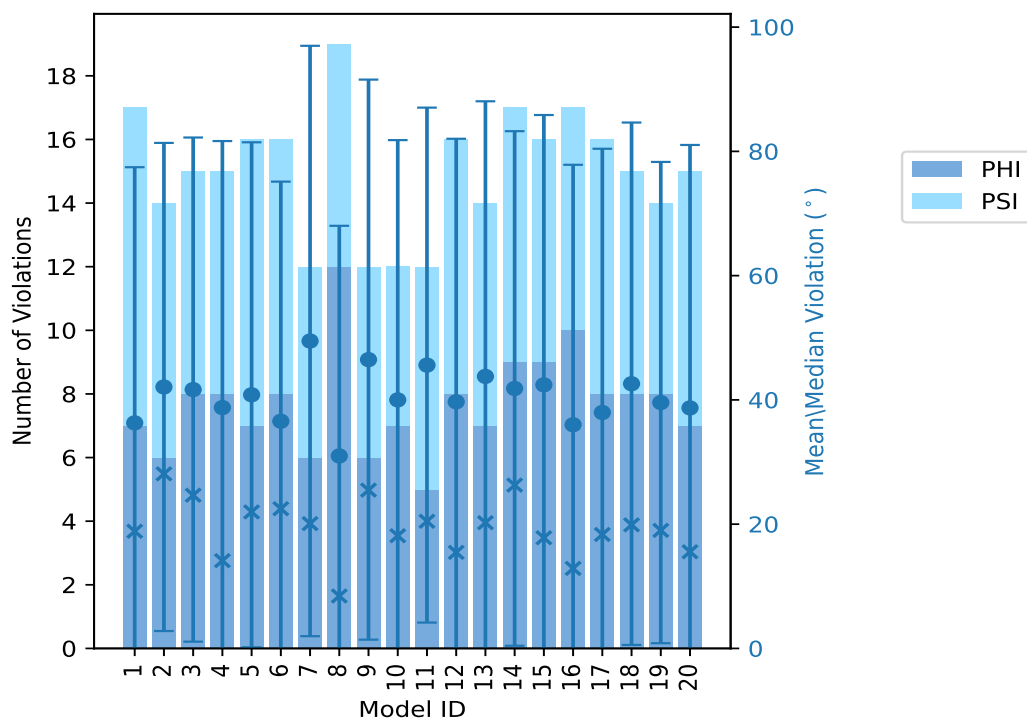
The mean(dot),median(x) and the standard deviation are shown in blue with respect to the y axis on the right

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	
PHI	PSI	Total	Count[1]	%
4	8	12	1	5.0
0	1	1	2	10.0
1	1	2	3	15.0
0	0	0	4	20.0
3	3	6	5	25.0
0	0	0	6	30.0
0	0	0	7	35.0
0	0	0	8	40.0
2	0	2	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	
PHI	PSI	Total	Count[1]	%
1	0	1	12	60.0
1	0	1	13	65.0
0	0	0	14	70.0
1	0	1	15	75.0
0	0	0	16	80.0
2	0	2	17	85.0
0	1	1	18	90.0
0	0	0	19	95.0
2	5	7	20	100.0

[1] Number of models with violations

### 10.3.1   Bar graph : Dihedral-angle Violation statistics for the ensemble ⓘ
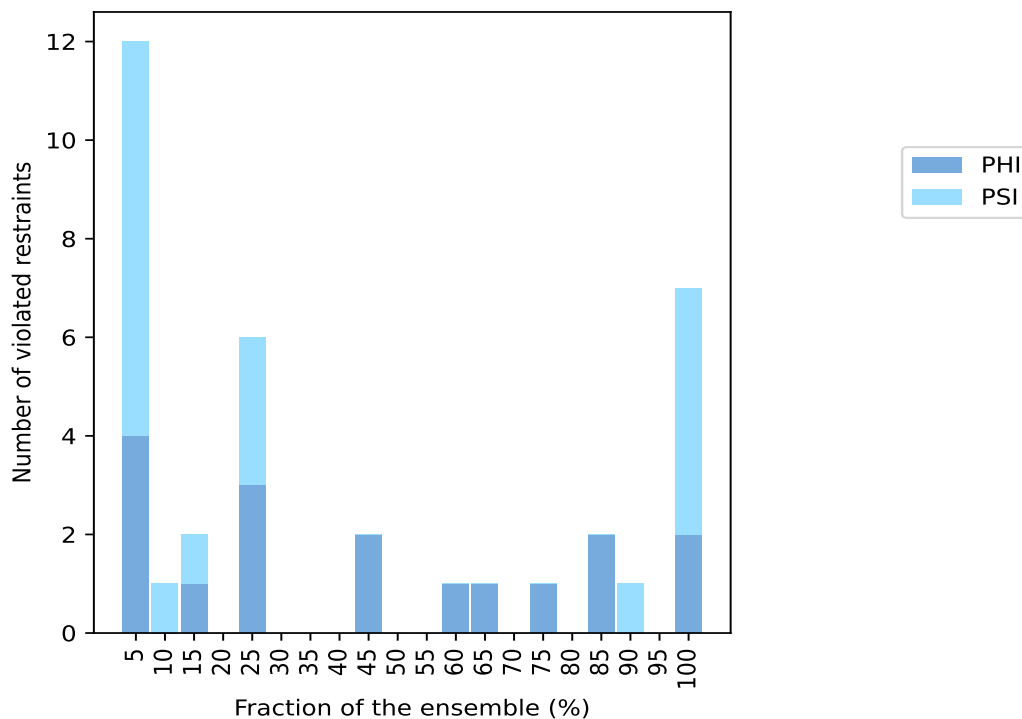
10.4   Most violated dihedral-angle restraints in the ensemble ⓘ

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
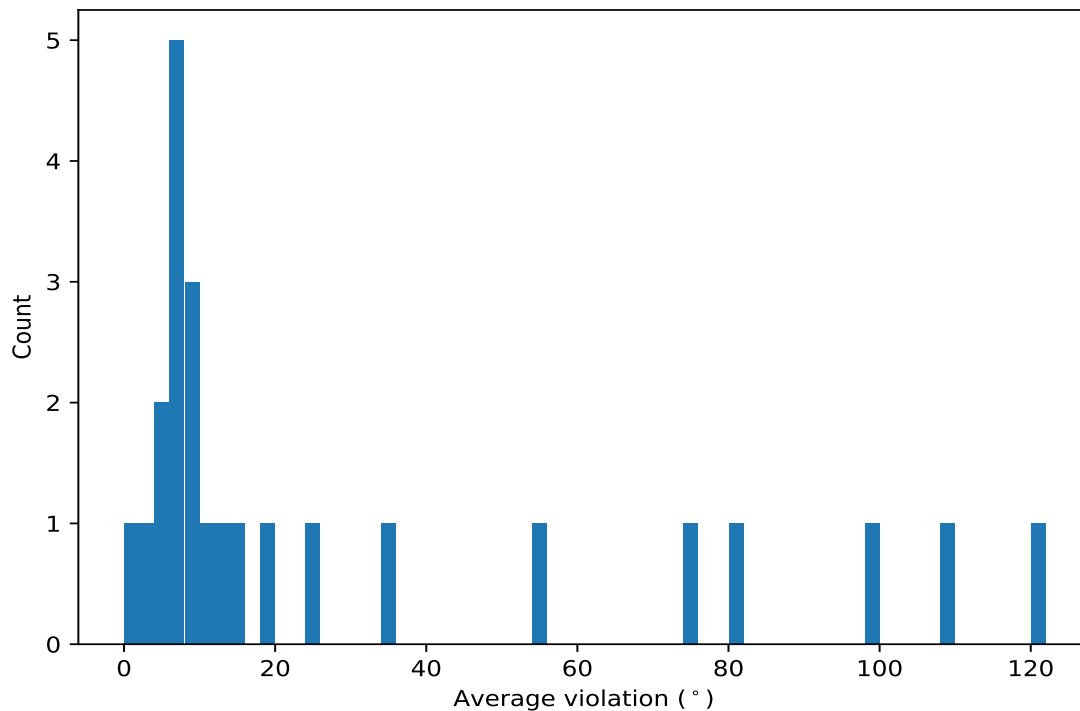
10.4.2 Table: Most violated dihedral-angle restraints ⓘ

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
(1,92)	1:81:A:ALA:N	1:81:A:ALA:CA	1:81:A:ALA:C	1:82:A:SER:N	20	108.9	10.83	112.28
(1,39)	1:41:A:PRO:C	1:42:A:ASP:N	1:42:A:ASP:CA	1:42:A:ASP:C	20	99.77	4.0	99.98
(1,102)	1:86:A:GLN:N	1:86:A:GLN:CA	1:86:A:GLN:C	1:87:A:VAL:N	20	81.31	7.36	79.94
(1,56)	1:51:A:ALA:N	1:51:A:ALA:CA	1:51:A:ALA:C	1:52:A:ASP:N	20	35.09	31.34	24.9
(1,142)	1:120:A:LEU:N	1:120:A:LEU:CA	1:120:A:LEU:C	1:121:A:TYR:N	20	24.1	2.24	23.7
(1,40)	1:42:A:ASP:N	1:42:A:ASP:CA	1:42:A:ASP:C	1:43:A:ALA:N	20	19.04	4.53	18.83
(1,119)	1:100:A:THR:C	1:101:A:VAL:N	1:101:A:VAL:CA	1:101:A:VAL:C	20	7.28	3.2	7.98
(1,10)	1:23:A:ASP:N	1:23:A:ASP:CA	1:23:A:ASP:C	1:24:A:LYS:N	18	120.75	14.68	125.04
(1,9)	1:22:A:HIS:C	1:23:A:ASP:N	1:23:A:ASP:CA	1:23:A:ASP:C	17	55.78	18.68	56.07
(1,67)	1:58:A:THR:C	1:59:A:GLY:N	1:59:A:GLY:CA	1:59:A:GLY:C	17	12.02	5.6	15.48

[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.
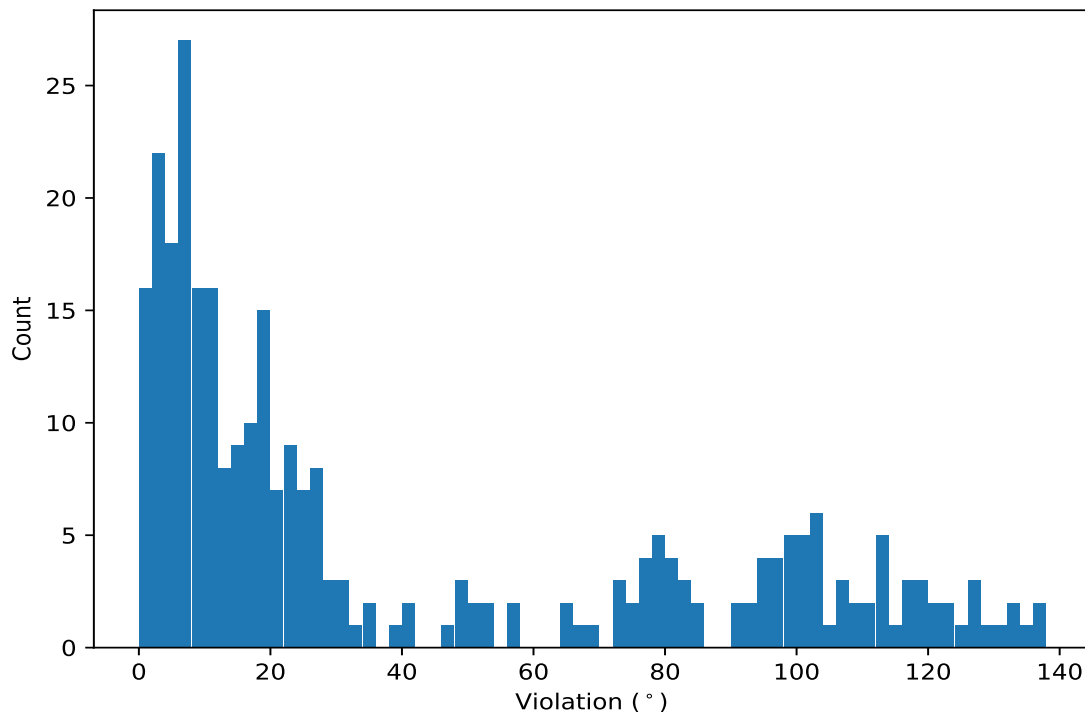
10.5.2   Table: All violated dihedral-angle restraints ⓘ

The following table provides the list of violationsfor 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 (°)
(1,10)	1:23:A:ASP:N	1:23:A:ASP:CA	1:23:A:ASP:C	1:24:A:LYS:N	17	137.44
(1,10)	1:23:A:ASP:N	1:23:A:ASP:CA	1:23:A:ASP:C	1:24:A:LYS:N	7	137.13
(1,10)	1:23:A:ASP:N	1:23:A:ASP:CA	1:23:A:ASP:C	1:24:A:LYS:N	13	135.31
(1,10)	1:23:A:ASP:N	1:23:A:ASP:CA	1:23:A:ASP:C	1:24:A:LYS:N	2	132.49
(1,10)	1:23:A:ASP:N	1:23:A:ASP:CA	1:23:A:ASP:C	1:24:A:LYS:N	14	132.18
(1,10)	1:23:A:ASP:N	1:23:A:ASP:CA	1:23:A:ASP:C	1:24:A:LYS:N	16	130.3
(1,10)	1:23:A:ASP:N	1:23:A:ASP:CA	1:23:A:ASP:C	1:24:A:LYS:N	9	128.19
(1,56)	1:51:A:ALA:N	1:51:A:ALA:CA	1:51:A:ALA:C	1:52:A:ASP:N	15	126.99
(1,10)	1:23:A:ASP:N	1:23:A:ASP:CA	1:23:A:ASP:C	1:24:A:LYS:N	20	126.62
(1,10)	1:23:A:ASP:N	1:23:A:ASP:CA	1:23:A:ASP:C	1:24:A:LYS:N	4	126.11