



wwPDB NMR Structure Validation Summary Report ⓘ

Feb 21, 2024 – 05:29 PM EST

PDB ID : 8V9U
BMRB ID : 31134
Title : Solution NMR structure of human DNMT1 N-terminal alpha-helical domain
Authors : Hu, Q.; Botuyan, M.V.; Mer, G.
Deposited on : 2023-12-09

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 ⓘ symbol.

The types of validation reports are described at

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

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

Cyrange : Kirchner and Güntert (2011)
NmrClust : Kelley et al. (1996)
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
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) : 2.36

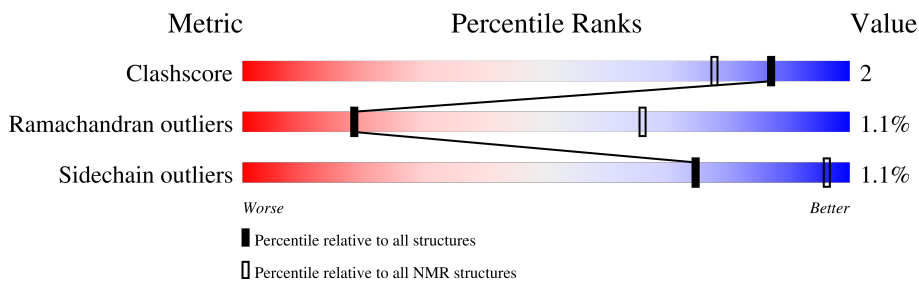
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 90%.

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	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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	81	

2 Ensemble composition and analysis i

This entry contains 30 models. Model 22 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:21-A:90 (70)	0.57	22

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 2 clusters and 1 single-model cluster was found.

Cluster number	Models
1	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 21, 22, 24, 25, 26, 28, 29
2	19, 27, 30
Single-model clusters	23

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1353 atoms, of which 690 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called DNA (cytosine-5)-methyltransferase 1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	81	1353	413	690	115	132	3	0

There are 3 discrepancies between the modelled and reference sequences:

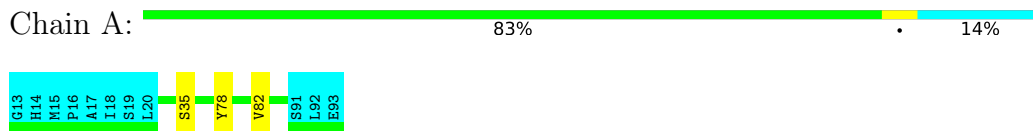
Chain	Residue	Modelled	Actual	Comment	Reference
A	13	GLY	-	expression tag	UNP P26358
A	14	HIS	-	expression tag	UNP P26358
A	15	MET	-	expression tag	UNP P26358

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.

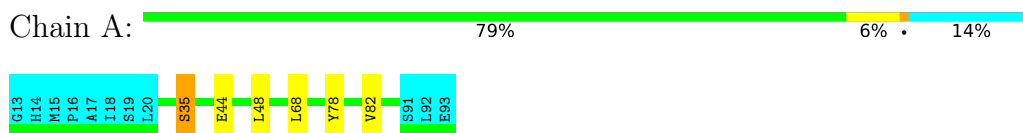
- Molecule 1: DNA (cytosine-5)-methyltransferase 1



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

The representative model is number 22. Colouring as in section 4.1 above.

- Molecule 1: DNA (cytosine-5)-methyltransferase 1



5 Refinement protocol and experimental data overview

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

Of the 200 calculated structures, 30 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
TALOS+	geometry optimization	
X-PLOR NIH	structure calculation	
X-PLOR NIH	refinement	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	1220
Number of shifts mapped to atoms	1030
Number of unparsed shifts	0
Number of shifts with mapping errors	190
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	90%

6 Model quality [i](#)

6.1 Standard geometry [i](#)

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	A	583	608	608	2±1
All	All	17490	18240	18240	71

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 2.

5 of 18 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:78:TYR:O	1:A:82:VAL:HG23	0.59	1.97	26	29
1:A:68:LEU:HD11	1:A:78:TYR:CD2	0.55	2.37	22	3
1:A:31:LEU:HD13	1:A:44:GLU:OE1	0.52	2.04	2	2
1:A:60:GLN:OE1	1:A:85:LEU:HD11	0.51	2.07	30	4
1:A:68:LEU:HD11	1:A:78:TYR:CG	0.51	2.41	28	5

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	70/81 (86%)	68±1 (98±1%)	1±1 (1±1%)	1±0 (1±1%)	18	66
All	All	2100/2430 (86%)	2051 (98%)	25 (1%)	24 (1%)	18	66

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
1	A	35	SER	24

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	68/77 (88%)	67±1 (99±1%)	1±1 (1±1%)	74	96
All	All	2040/2310 (88%)	2018 (99%)	22 (1%)	74	96

5 of 8 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	35	SER	13
1	A	89	ASP	2
1	A	60	GLN	2
1	A	40	GLU	1
1	A	62	CYS	1

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 monosaccharides in this entry.

6.6 Ligand geometry [i](#)

There are no ligands in this entry.

6.7 Other polymers [i](#)

There are no such molecules in this entry.

6.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

7 Chemical shift validation i

The completeness of assignment taking into account all chemical shift lists is 90% for the well-defined parts and 88% 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	1220
Number of shifts mapped to atoms	1030
Number of unparsed shifts	0
Number of shifts with mapping errors	190
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

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

- Chemical shift has been reported more than once. All 4 occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	120	VAL	HG12	0.883	.	1
1	A	120	VAL	HG13	0.883	.	1
1	A	120	VAL	HG22	0.883	.	1
1	A	120	VAL	HG23	0.883	.	1

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. First 5 (of 190) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	94	ASN	H	8.407	.	1
1	A	94	ASN	HB2	2.768	.	1
1	A	94	ASN	HB3	2.768	.	1
1	A	94	ASN	HD22	6.891	.	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	94	ASN	C	175.823	.	1
1	A	94	ASN	CA	53.482	.	1
1	A	94	ASN	CB	38.908	.	1
1	A	94	ASN	N	119.353	.	1
1	A	94	ASN	ND2	112.633	.	1
1	A	95	GLY	H	8.379	.	1
1	A	95	GLY	C	174.195	.	1
1	A	95	GLY	CA	45.666	.	1
1	A	95	GLY	N	109.043	.	1
1	A	96	ALA	H	8.084	.	1
1	A	96	ALA	CA	52.73	.	1
1	A	96	ALA	CB	19.142	.	1
1	A	96	ALA	N	123.157	.	1
1	A	97	HIS	H	8.146	.	1
1	A	97	HIS	CA	56.309	.	1
1	A	97	HIS	CB	30.651	.	1
1	A	97	HIS	N	117.761	.	1
1	A	98	ALA	H	8.025	.	1
1	A	98	ALA	C	177.325	.	1
1	A	98	ALA	CA	52.509	.	1
1	A	98	ALA	CB	19.306	.	1
1	A	98	ALA	N	124.282	.	1
1	A	99	TYR	H	8.087	.	1
1	A	99	TYR	HA	4.498	.	1
1	A	99	TYR	HB2	3.008	.	2
1	A	99	TYR	HB3	2.938	.	2
1	A	99	TYR	HD1	7.059	.	1
1	A	99	TYR	HD2	7.059	.	1
1	A	99	TYR	HE1	6.783	.	1
1	A	99	TYR	HE2	6.783	.	1
1	A	99	TYR	C	175.437	.	1
1	A	99	TYR	CA	57.809	.	1
1	A	99	TYR	CB	38.632	.	1
1	A	99	TYR	CD1	133.044	.	1
1	A	99	TYR	CD2	133.044	.	1
1	A	99	TYR	CE1	118.016	.	1
1	A	99	TYR	CE2	118.016	.	1
1	A	99	TYR	N	119.03	.	1
1	A	100	ASN	H	8.255	.	1
1	A	100	ASN	C	174.91	.	1
1	A	100	ASN	CA	53.161	.	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	100	ASN	CB	38.769	.	1
1	A	100	ASN	N	120.445	.	1
1	A	101	ARG	H	8.182	.	1
1	A	101	ARG	C	176.358	.	1
1	A	101	ARG	CA	56.666	.	1
1	A	101	ARG	CB	30.803	.	1
1	A	101	ARG	N	121.353	.	1
1	A	102	GLU	H	8.153	.	1
1	A	102	GLU	C	174.762	.	1
1	A	102	GLU	CA	56.192	.	1
1	A	102	GLU	N	121.132	.	1
1	A	103	VAL	H	8.159	.	1
1	A	103	VAL	C	176.099	.	1
1	A	103	VAL	CA	62.608	.	1
1	A	103	VAL	CB	32.624	.	1
1	A	103	VAL	N	121.282	.	1
1	A	104	ASN	H	8.48	.	1
1	A	104	ASN	C	175.686	.	1
1	A	104	ASN	CA	53.674	.	1
1	A	104	ASN	CB	38.817	.	1
1	A	104	ASN	N	121.506	.	1
1	A	105	GLY	H	8.329	.	1
1	A	105	GLY	C	173.899	.	1
1	A	105	GLY	CA	45.588	.	1
1	A	105	GLY	N	108.918	.	1
1	A	106	ARG	H	8.071	.	1
1	A	106	ARG	C	176.077	.	1
1	A	106	ARG	CA	56.016	.	1
1	A	106	ARG	CB	30.794	.	1
1	A	106	ARG	N	120.236	.	1
1	A	107	LEU	H	8.28	.	1
1	A	107	LEU	C	177.449	.	1
1	A	107	LEU	CA	55.284	.	1
1	A	107	LEU	CB	42.232	.	1
1	A	107	LEU	N	123.163	.	1
1	A	108	GLU	H	8.466	.	1
1	A	108	GLU	C	176.269	.	1
1	A	108	GLU	CA	56.7	.	1
1	A	108	GLU	CB	30.222	.	1
1	A	108	GLU	N	121.794	.	1
1	A	109	ASN	H	8.23	.	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	109	ASN	C	177.012	.	1
1	A	109	ASN	CA	54.468	.	1
1	A	109	ASN	CB	41.369	.	1
1	A	109	ASN	N	121.131	.	1
1	A	110	GLY	H	8.369	.	1
1	A	110	GLY	C	174.215	.	1
1	A	110	GLY	CA	45.716	.	1
1	A	110	GLY	N	109.351	.	1
1	A	111	ASN	H	8.307	.	1
1	A	111	ASN	C	175.607	.	1
1	A	111	ASN	CA	53.647	.	1
1	A	111	ASN	CB	38.728	.	1
1	A	111	ASN	N	118.592	.	1
1	A	112	GLN	H	8.371	.	1
1	A	112	GLN	C	175.997	.	1
1	A	112	GLN	CA	56.236	.	1
1	A	112	GLN	CB	29.181	.	1
1	A	112	GLN	N	120.552	.	1
1	A	113	ALA	H	8.281	.	1
1	A	113	ALA	C	177.869	.	1
1	A	113	ALA	CA	52.771	.	1
1	A	113	ALA	CB	19.119	.	1
1	A	113	ALA	N	124.576	.	1
1	A	114	ARG	H	8.161	.	1
1	A	114	ARG	CA	56.337	.	1
1	A	114	ARG	CB	30.69	.	1
1	A	114	ARG	N	119.817	.	1
1	A	115	SER	H	8.292	.	1
1	A	115	SER	CA	58.76	.	1
1	A	115	SER	CB	63.578	.	1
1	A	115	SER	N	116.816	.	1
1	A	116	GLU	H	8.452	.	1
1	A	116	GLU	CA	56.884	.	1
1	A	116	GLU	CB	30.228	.	1
1	A	116	GLU	N	122.907	.	1
1	A	117	ALA	H	8.191	.	1
1	A	117	ALA	N	124.414	.	1
1	A	118	ARG	C	176.275	.	1
1	A	118	ARG	CB	30.716	.	1
1	A	119	ARG	H	8.362	.	1
1	A	119	ARG	C	176.753	.	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	119	ARG	CA	56.762	.	1
1	A	119	ARG	CB	30.297	.	1
1	A	119	ARG	N	121.442	.	1
1	A	120	VAL	H	8.122	.	1
1	A	120	VAL	HA	4.033	.	1
1	A	120	VAL	HB	2.021	.	1
1	A	120	VAL	HG11	0.883	.	1
1	A	120	VAL	HG12	0.883	.	1
1	A	120	VAL	HG13	0.883	.	1
1	A	120	VAL	HG21	0.883	.	1
1	A	120	VAL	HG22	0.883	.	1
1	A	120	VAL	HG23	0.883	.	1
1	A	120	VAL	C	176.652	.	1
1	A	120	VAL	CA	62.58	.	1
1	A	120	VAL	CB	32.752	.	1
1	A	120	VAL	N	120.691	.	1
1	A	121	GLY	H	8.498	.	1
1	A	121	GLY	HA2	4.072	.	1
1	A	121	GLY	HA3	3.934	.	1
1	A	121	GLY	C	174.154	.	1
1	A	121	GLY	CA	45.243	.	1
1	A	121	GLY	N	112.464	.	1
1	A	122	MET	H	8.146	.	1
1	A	122	MET	C	176.153	.	1
1	A	122	MET	CA	55.495	.	1
1	A	122	MET	CB	33.108	.	1
1	A	122	MET	N	119.787	.	1
1	A	123	ALA	H	8.366	.	1
1	A	123	ALA	C	177.402	.	1
1	A	123	ALA	CA	52.654	.	1
1	A	123	ALA	CB	19.235	.	1
1	A	123	ALA	N	125.011	.	1
1	A	124	ASP	H	8.254	.	1
1	A	124	ASP	C	176.27	.	1
1	A	124	ASP	CA	54.136	.	1
1	A	124	ASP	CB	41.329	.	1
1	A	124	ASP	N	119.648	.	1
1	A	125	ALA	H	8.253	.	1
1	A	125	ALA	C	177.647	.	1
1	A	125	ALA	CA	52.915	.	1
1	A	125	ALA	CB	19.154	.	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	125	ALA	N	124.412	.	1
1	A	126	ASN	H	8.381	.	1
1	A	126	ASN	C	174.9	.	1
1	A	126	ASN	N	116.719	.	1
1	A	127	SER	H	8.005	.	1
1	A	127	SER	N	117.024	.	1
1	A	131	PRO	C	176.773	.	1
1	A	132	LEU	H	8.378	.	1
1	A	132	LEU	C	177.436	.	1
1	A	132	LEU	CA	55.203	.	1
1	A	132	LEU	CB	42.386	.	1
1	A	132	LEU	N	122.535	.	1
1	A	133	SER	H	8.247	.	1
1	A	133	SER	C	173.339	.	1
1	A	133	SER	CA	58.219	.	1
1	A	133	SER	CB	63.946	.	1
1	A	133	SER	N	117.093	.	1
1	A	134	LYS	H	7.957	.	1
1	A	134	LYS	C	181.206	.	1
1	A	134	LYS	CA	57.871	.	1
1	A	134	LYS	CB	33.679	.	1
1	A	134	LYS	N	128.09	.	1

7.1.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	111	-0.48 ± 0.10	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	106	0.03 ± 0.08	None needed (< 0.5 ppm)
$^{13}\text{C}'$	101	-0.37 ± 0.12	None needed (< 0.5 ppm)
^{15}N	111	0.47 ± 0.30	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 90%, i.e. 926 atoms were assigned a chemical shift out of a possible 1030. 0 out of 18 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹ H	¹³ C	¹⁵ N
Backbone	341/349 (98%)	139/140 (99%)	134/140 (96%)	68/69 (99%)
Sidechain	568/655 (87%)	385/417 (92%)	178/208 (86%)	5/30 (17%)
Aromatic	17/26 (65%)	9/13 (69%)	8/12 (67%)	0/1 (0%)
Overall	926/1030 (90%)	533/570 (94%)	320/360 (89%)	73/100 (73%)

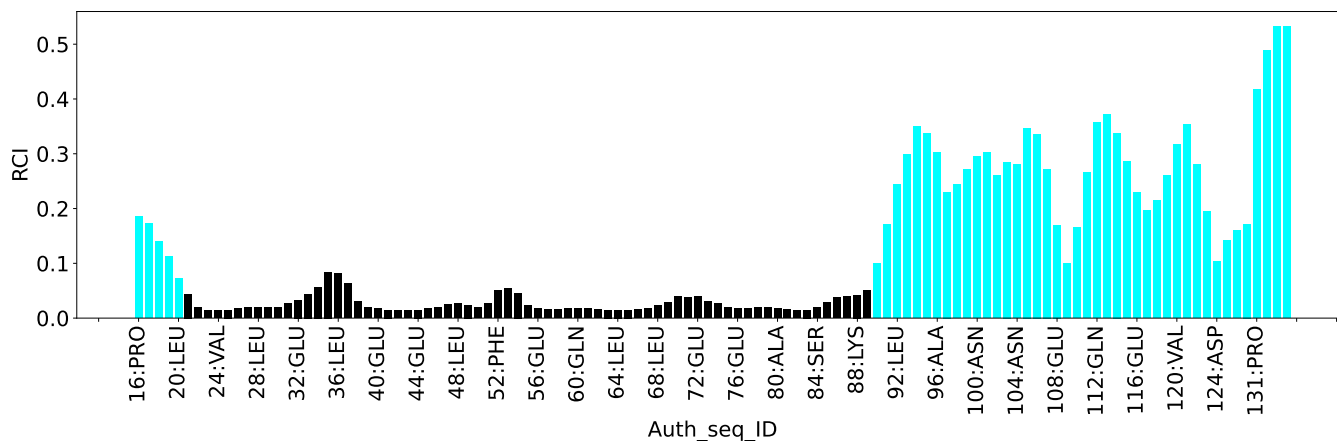
7.1.4 Statistically unusual chemical shifts [i](#)

There are no statistically unusual chemical shifts.

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

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

Random coil index (RCI) for chain A:



8 NMR restraints analysis

8.1 Conformationally restricting restraints

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

Description	Value
Total distance restraints	937
Intra-residue ($ i-j =0$)	243
Sequential ($ i-j =1$)	288
Medium range ($ i-j >1$ and $ i-j <5$)	306
Long range ($ i-j \geq 5$)	100
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	132
Number of unmapped restraints	0
Number of restraints per residue	13.2
Number of long range restraints per residue ¹	1.2

¹Long 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)	19.9	0.2
0.2-0.5 (Medium)	16.0	0.49
>0.5 (Large)	3.9	1.31

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

Dihedral-angle violations less than 1° are not included in the calculation.

Bins (°)	Average number of violations per model	Max (°)
1.0-10.0 (Small)	2.2	4.61
10.0-20.0 (Medium)	None	None
>20.0 (Large)	None	None

9 Distance violation analysis [i](#)

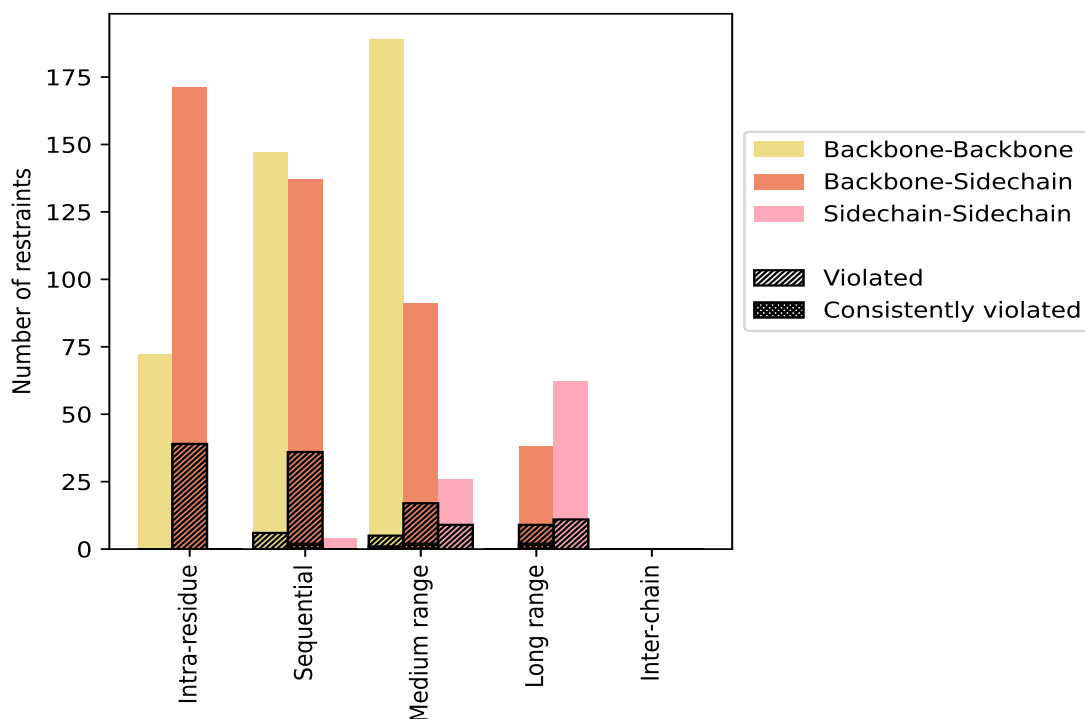
9.1 Summary of distance violations [i](#)

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.

Restrains type	Count	% ¹	Violated ³			Consistently Violated ⁴		
			Count	% ²	% ¹	Count	% ²	% ¹
Intra-residue ($i-j =0$)	243	25.9	39	16.0	4.2	0	0.0	0.0
Backbone-Backbone	72	7.7	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	171	18.2	39	22.8	4.2	0	0.0	0.0
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sequential ($i-j =1$)	288	30.7	42	14.6	4.5	2	0.7	0.2
Backbone-Backbone	147	15.7	6	4.1	0.6	0	0.0	0.0
Backbone-Sidechain	137	14.6	36	26.3	3.8	2	1.5	0.2
Sidechain-Sidechain	4	0.4	0	0.0	0.0	0	0.0	0.0
Medium range ($i-j >1$ & $i-j <5$)	306	32.7	31	10.1	3.3	3	1.0	0.3
Backbone-Backbone	189	20.2	5	2.6	0.5	1	0.5	0.1
Backbone-Sidechain	91	9.7	17	18.7	1.8	2	2.2	0.2
Sidechain-Sidechain	26	2.8	9	34.6	1.0	0	0.0	0.0
Long range ($i-j \geq 5$)	100	10.7	20	20.0	2.1	2	2.0	0.2
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	38	4.1	9	23.7	1.0	2	5.3	0.2
Sidechain-Sidechain	62	6.6	11	17.7	1.2	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	937	100.0	132	14.1	14.1	7	0.7	0.7
Backbone-Backbone	408	43.5	11	2.7	1.2	1	0.2	0.1
Backbone-Sidechain	437	46.6	101	23.1	10.8	6	1.4	0.6
Sidechain-Sidechain	92	9.8	20	21.7	2.1	0	0.0	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 disulfid bonds are counted in their appropriate category on the x-axis

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

The following table provides the distance violation statistics for each model in the ensemble. Violations less than 0.1 Å are not included in the statistics.

Model ID	Number of violations						Mean (Å)	Max (Å)	SD ⁶ (Å)	Median (Å)
	IR ¹	SQ ²	MR ³	LR ⁴	IC ⁵	Total				
1	13	13	7	8	0	41	0.28	1.15	0.22	0.18
2	7	17	8	9	0	41	0.28	0.91	0.2	0.21
3	7	12	10	7	0	36	0.26	1.14	0.22	0.2
4	11	17	13	9	0	50	0.24	0.81	0.15	0.2
5	13	18	6	6	0	43	0.25	0.96	0.2	0.16
6	6	16	8	5	0	35	0.24	0.9	0.17	0.15
7	8	16	11	9	0	44	0.27	1.16	0.25	0.16
8	8	19	7	9	0	43	0.26	1.2	0.2	0.21
9	14	12	6	9	0	41	0.31	1.12	0.24	0.23
10	6	14	7	6	0	33	0.29	1.21	0.25	0.22
11	12	18	11	6	0	47	0.26	1.0	0.18	0.19

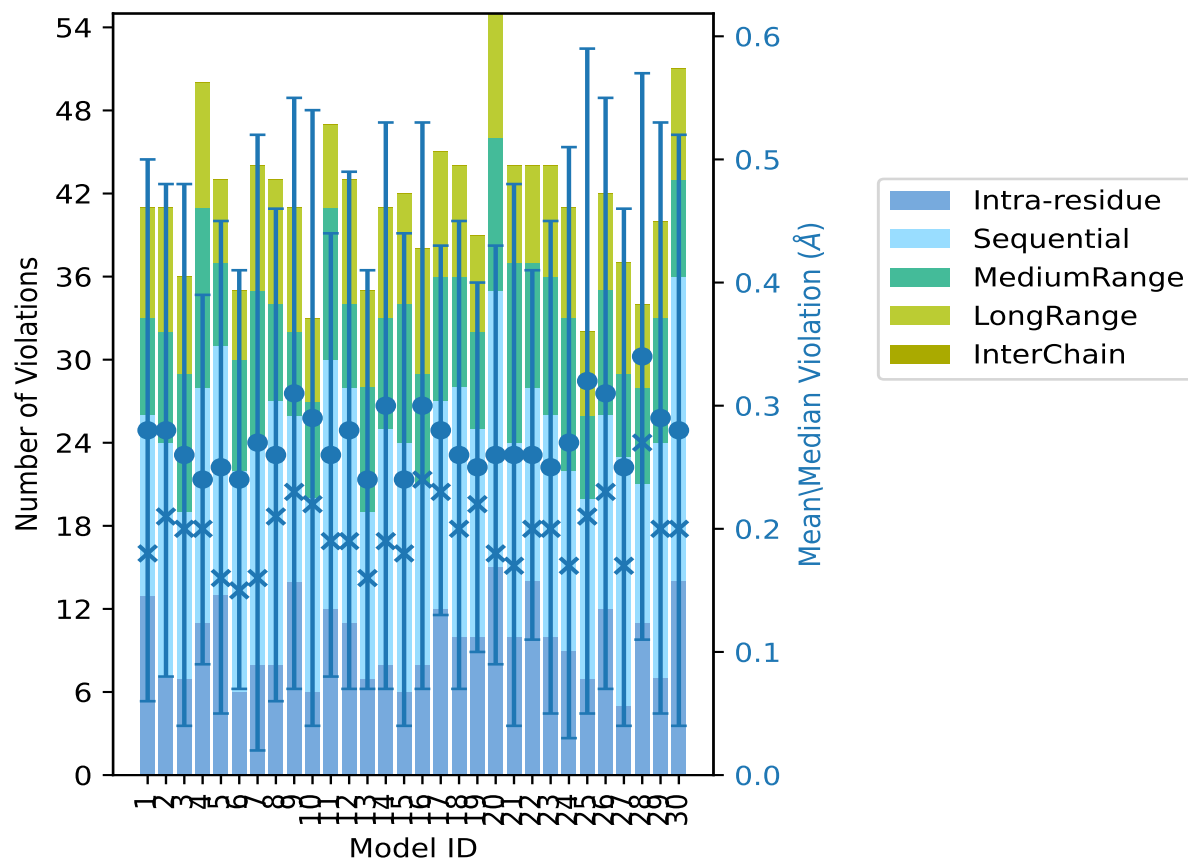
Continued on next page...

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	11	17	6	9	0	43	0.28	1.11	0.21	0.19
13	7	12	9	7	0	35	0.24	0.88	0.17	0.16
14	8	17	8	8	0	41	0.3	1.2	0.23	0.19
15	6	18	10	8	0	42	0.24	0.98	0.2	0.18
16	8	13	8	9	0	38	0.3	1.1	0.23	0.24
17	12	15	9	9	0	45	0.28	0.72	0.15	0.23
18	10	18	8	8	0	44	0.26	1.05	0.19	0.2
19	10	15	7	7	0	39	0.25	0.89	0.15	0.22
20	15	20	11	9	0	55	0.26	0.89	0.17	0.18
21	10	14	13	7	0	44	0.26	1.29	0.22	0.17
22	14	14	9	7	0	44	0.26	0.82	0.15	0.2
23	10	16	10	8	0	44	0.25	1.16	0.2	0.2
24	9	13	11	8	0	41	0.27	1.03	0.24	0.17
25	7	13	6	6	0	32	0.32	1.19	0.27	0.21
26	12	14	9	7	0	42	0.31	1.31	0.24	0.23
27	5	18	6	8	0	37	0.25	1.11	0.21	0.17
28	11	10	7	6	0	34	0.34	0.96	0.23	0.27
29	7	17	9	7	0	40	0.29	1.18	0.24	0.2
30	14	22	7	8	0	51	0.28	1.28	0.24	0.2

¹Intra-residue restraints, ²Sequential restraints, ³Medium range restraints, ⁴Long range restraints,
⁵Inter-chain restraints, ⁶Standard 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, 805(IR:204, SQ:246, MR:275, LR:80, 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	%
7	5	7	4	0	23	1	3.3
3	3	5	4	0	15	2	6.7
3	6	5	0	0	14	3	10.0
1	3	2	1	0	7	4	13.3
4	2	0	0	0	6	5	16.7
3	0	0	0	0	3	6	20.0

Continued on next page...

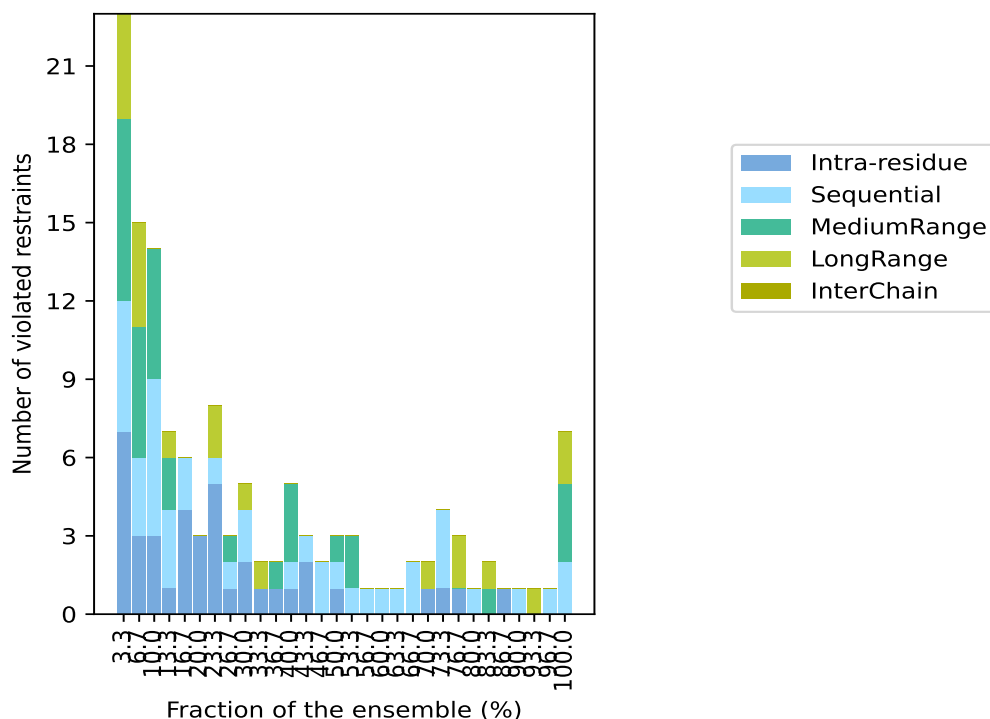
Continued from previous page...

Number of violated restraints						Fraction of the ensemble	
IR ¹	SQ ²	MR ³	LR ⁴	IC ⁵	Total	Count ⁶	%
5	1	0	2	0	8	7	23.3
1	1	1	0	0	3	8	26.7
2	2	0	1	0	5	9	30.0
1	0	0	1	0	2	10	33.3
1	0	1	0	0	2	11	36.7
1	1	3	0	0	5	12	40.0
2	1	0	0	0	3	13	43.3
0	2	0	0	0	2	14	46.7
1	1	1	0	0	3	15	50.0
0	1	2	0	0	3	16	53.3
0	1	0	0	0	1	17	56.7
0	1	0	0	0	1	18	60.0
0	1	0	0	0	1	19	63.3
0	2	0	0	0	2	20	66.7
1	0	0	1	0	2	21	70.0
1	3	0	0	0	4	22	73.3
1	0	0	2	0	3	23	76.7
0	1	0	0	0	1	24	80.0
0	0	1	1	0	2	25	83.3
1	0	0	0	0	1	26	86.7
0	1	0	0	0	1	27	90.0
0	0	0	1	0	1	28	93.3
0	1	0	0	0	1	29	96.7
0	2	3	2	0	7	30	100.0

¹Intra-residue restraints, ²Sequential restraints, ³Medium range restraints, ⁴Long range restraints,

⁵Inter-chain restraints, ⁶ Number of models with violations

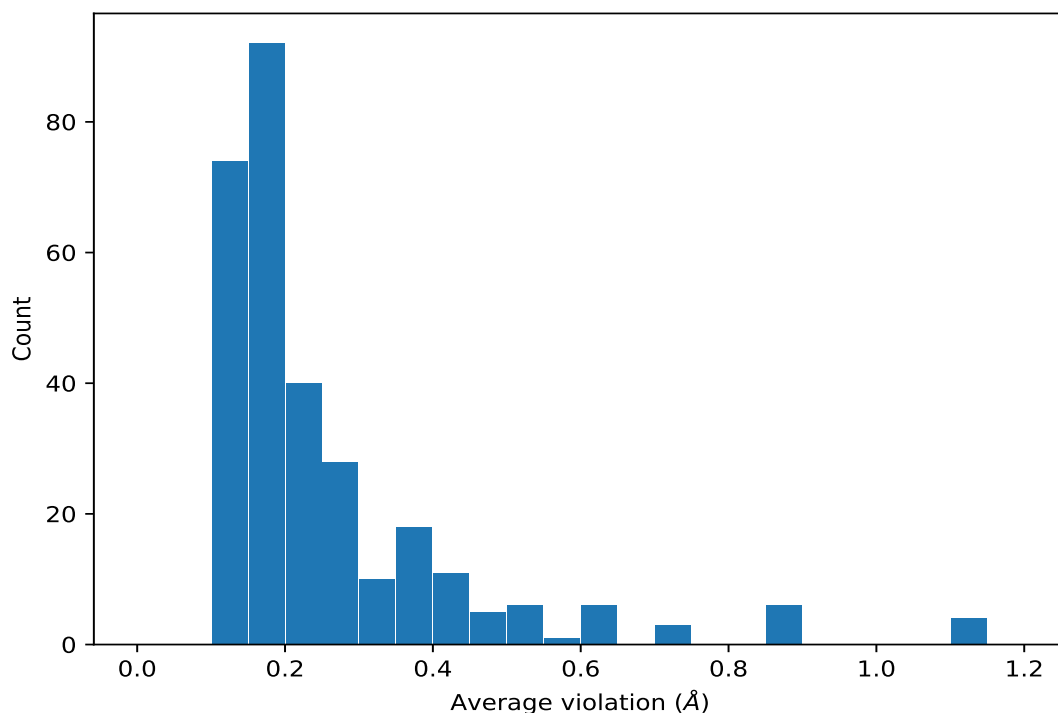
9.3.1 Bar graph : Distance violation statistics for the ensemble [i](#)



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 in the ensemble



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

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

Key	Atom-1	Atom-2	Models ¹	Mean (Å)	SD ¹ (Å)	Median (Å)
(1,921)	1:38:A:GLU:HA	1:79:A:LEU:HD11	30	0.86	0.1	0.87
(1,921)	1:38:A:GLU:HA	1:79:A:LEU:HD12	30	0.86	0.1	0.87
(1,921)	1:38:A:GLU:HA	1:79:A:LEU:HD13	30	0.86	0.1	0.87
(1,921)	1:38:A:GLU:HA	1:79:A:LEU:HD21	30	0.86	0.1	0.87
(1,921)	1:38:A:GLU:HA	1:79:A:LEU:HD22	30	0.86	0.1	0.87
(1,921)	1:38:A:GLU:HA	1:79:A:LEU:HD23	30	0.86	0.1	0.87
(1,375)	1:48:A:LEU:H	1:28:A:LEU:HD11	30	0.63	0.09	0.66
(1,375)	1:48:A:LEU:H	1:28:A:LEU:HD12	30	0.63	0.09	0.66
(1,375)	1:48:A:LEU:H	1:28:A:LEU:HD13	30	0.63	0.09	0.66
(1,375)	1:48:A:LEU:H	1:28:A:LEU:HD21	30	0.63	0.09	0.66
(1,375)	1:48:A:LEU:H	1:28:A:LEU:HD22	30	0.63	0.09	0.66
(1,375)	1:48:A:LEU:H	1:28:A:LEU:HD23	30	0.63	0.09	0.66
(1,673)	1:72:A:GLU:H	1:71:A:GLU:HB2	30	0.28	0.03	0.28
(1,673)	1:72:A:GLU:H	1:71:A:GLU:HB3	30	0.28	0.03	0.28
(1,647)	1:69:A:ARG:H	1:68:A:LEU:HG	30	0.27	0.03	0.28
(1,801)	1:82:A:VAL:HG11	1:79:A:LEU:HA	30	0.25	0.03	0.24

Continued on next page...

Continued from previous page...

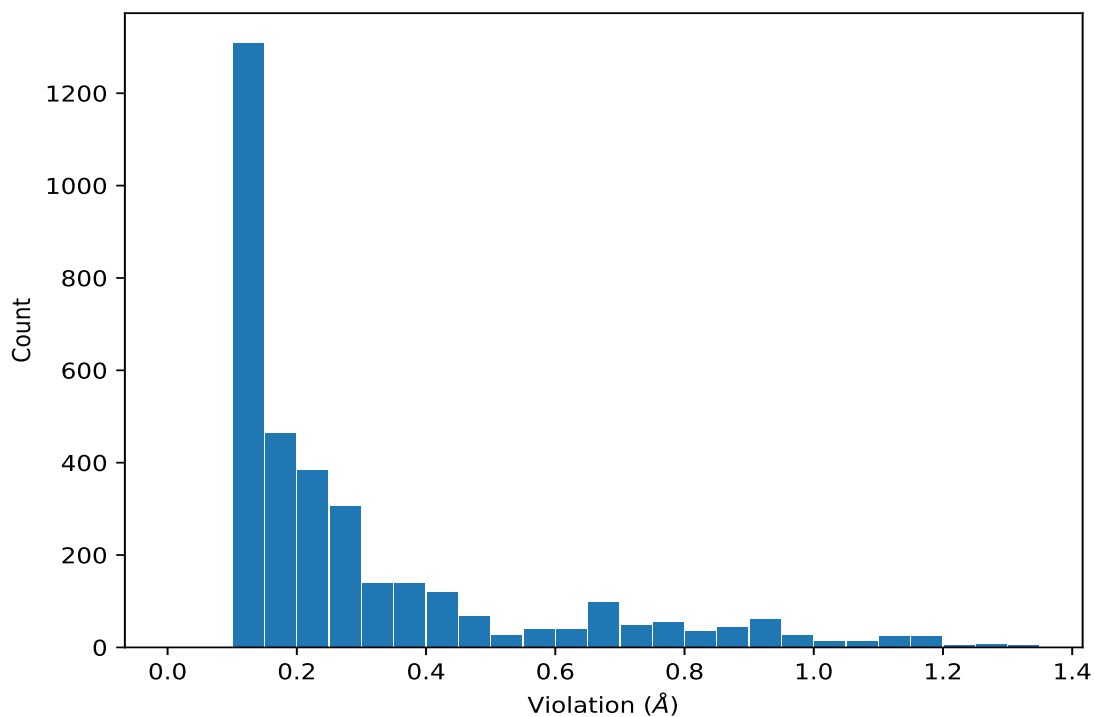
Key	Atom-1	Atom-2	Models ¹	Mean (Å)	SD ¹ (Å)	Median (Å)
(1,801)	1:82:A:VAL:HG12	1:79:A:LEU:HA	30	0.25	0.03	0.24
(1,801)	1:82:A:VAL:HG13	1:79:A:LEU:HA	30	0.25	0.03	0.24
(1,495)	1:60:A:GLN:H	1:57:A:ILE:HD11	30	0.25	0.04	0.25
(1,495)	1:60:A:GLN:H	1:57:A:ILE:HD12	30	0.25	0.04	0.25
(1,495)	1:60:A:GLN:H	1:57:A:ILE:HD13	30	0.25	0.04	0.25
(1,704)	1:74:A:SER:H	1:76:A:GLU:H	30	0.16	0.03	0.15
(1,231)	1:36:A:LEU:H	1:37:A:THR:H	29	0.36	0.05	0.37
(1,929)	1:47:A:ASN:HD21	1:58:A:LYS:HD2	28	0.19	0.06	0.18
(1,929)	1:47:A:ASN:HD21	1:58:A:LYS:HD3	28	0.19	0.06	0.18
(1,929)	1:47:A:ASN:HD22	1:58:A:LYS:HD2	28	0.19	0.06	0.18
(1,929)	1:47:A:ASN:HD22	1:58:A:LYS:HD3	28	0.19	0.06	0.18
(1,531)	1:62:A:CYS:H	1:61:A:LEU:HD11	27	0.15	0.03	0.14

¹Number of violated models, ²Standard 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)
(1,49)	1:20:A:LEU:HB2	1:25:A:ARG:HB2	26	1.31
(1,49)	1:20:A:LEU:HB2	1:25:A:ARG:HB3	26	1.31
(1,49)	1:20:A:LEU:HB3	1:25:A:ARG:HB2	26	1.31
(1,49)	1:20:A:LEU:HB3	1:25:A:ARG:HB3	26	1.31
(1,49)	1:20:A:LEU:HB2	1:25:A:ARG:HB2	21	1.29
(1,49)	1:20:A:LEU:HB2	1:25:A:ARG:HB3	21	1.29
(1,49)	1:20:A:LEU:HB3	1:25:A:ARG:HB2	21	1.29
(1,49)	1:20:A:LEU:HB3	1:25:A:ARG:HB3	21	1.29
(1,49)	1:20:A:LEU:HB2	1:25:A:ARG:HB2	30	1.28
(1,49)	1:20:A:LEU:HB2	1:25:A:ARG:HB3	30	1.28
(1,49)	1:20:A:LEU:HB3	1:25:A:ARG:HB2	30	1.28
(1,49)	1:20:A:LEU:HB3	1:25:A:ARG:HB3	30	1.28
(1,49)	1:20:A:LEU:HB2	1:25:A:ARG:HB2	10	1.21
(1,49)	1:20:A:LEU:HB2	1:25:A:ARG:HB3	10	1.21
(1,49)	1:20:A:LEU:HB3	1:25:A:ARG:HB2	10	1.21
(1,49)	1:20:A:LEU:HB3	1:25:A:ARG:HB3	10	1.21
(1,49)	1:20:A:LEU:HB2	1:25:A:ARG:HB2	8	1.2
(1,49)	1:20:A:LEU:HB2	1:25:A:ARG:HB3	8	1.2
(1,49)	1:20:A:LEU:HB3	1:25:A:ARG:HB2	8	1.2
(1,49)	1:20:A:LEU:HB3	1:25:A:ARG:HB3	8	1.2
(1,49)	1:20:A:LEU:HB2	1:25:A:ARG:HB2	14	1.2
(1,49)	1:20:A:LEU:HB2	1:25:A:ARG:HB3	14	1.2
(1,49)	1:20:A:LEU:HB3	1:25:A:ARG:HB2	14	1.2
(1,49)	1:20:A:LEU:HB3	1:25:A:ARG:HB3	14	1.2
(1,49)	1:20:A:LEU:HB2	1:25:A:ARG:HB2	25	1.19
(1,49)	1:20:A:LEU:HB2	1:25:A:ARG:HB3	25	1.19
(1,49)	1:20:A:LEU:HB3	1:25:A:ARG:HB2	25	1.19
(1,49)	1:20:A:LEU:HB3	1:25:A:ARG:HB3	25	1.19
(1,49)	1:20:A:LEU:HB2	1:25:A:ARG:HB2	29	1.18
(1,49)	1:20:A:LEU:HB2	1:25:A:ARG:HB3	29	1.18
(1,49)	1:20:A:LEU:HB3	1:25:A:ARG:HB2	29	1.18
(1,49)	1:20:A:LEU:HB3	1:25:A:ARG:HB3	29	1.18
(1,49)	1:20:A:LEU:HB2	1:25:A:ARG:HB2	7	1.16
(1,49)	1:20:A:LEU:HB2	1:25:A:ARG:HB3	7	1.16
(1,49)	1:20:A:LEU:HB3	1:25:A:ARG:HB2	7	1.16
(1,49)	1:20:A:LEU:HB3	1:25:A:ARG:HB3	7	1.16
(1,49)	1:20:A:LEU:HB2	1:25:A:ARG:HB2	23	1.16

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

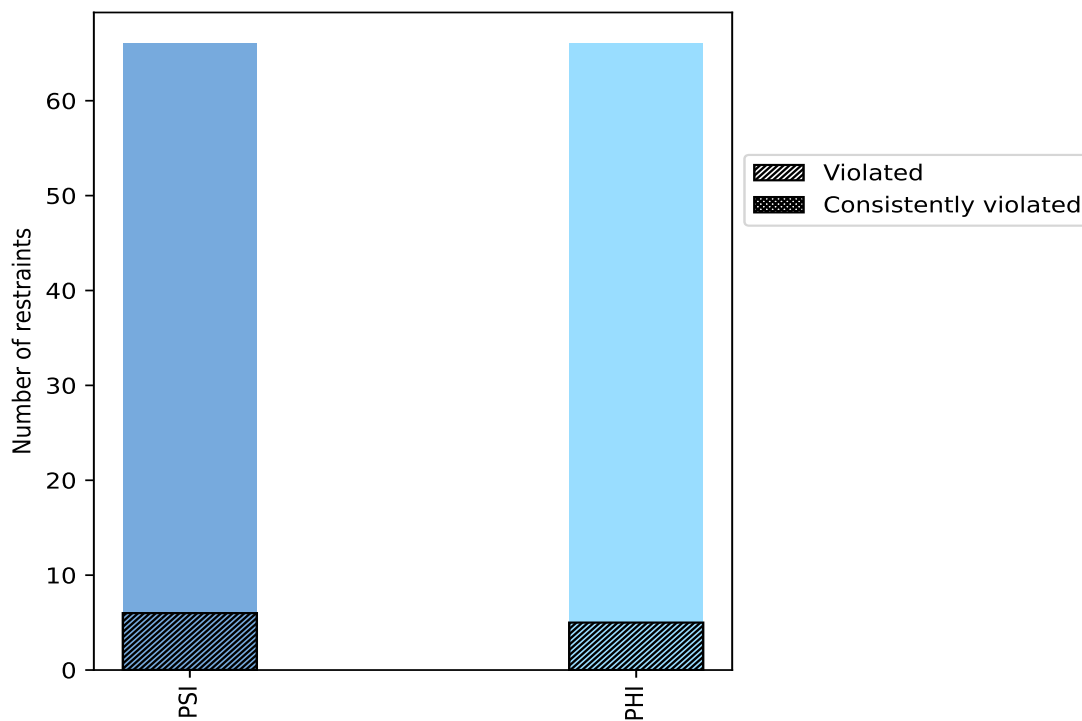
10.1 Summary of dihedral-angle violations [i](#)

The following table provides the summary of dihedral-angle violations in different dihedral-angle types. Violations less than 1° are not included in the calculation.

Angle type	Count	% ¹	Violated ³			Consistently Violated ⁴		
			Count	% ²	% ¹	Count	% ²	% ¹
PSI	66	50.0	6	9.1	4.5	0	0.0	0.0
PHI	66	50.0	5	7.6	3.8	0	0.0	0.0
Total	132	100.0	11	8.3	8.3	0	0.0	0.0

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](#)



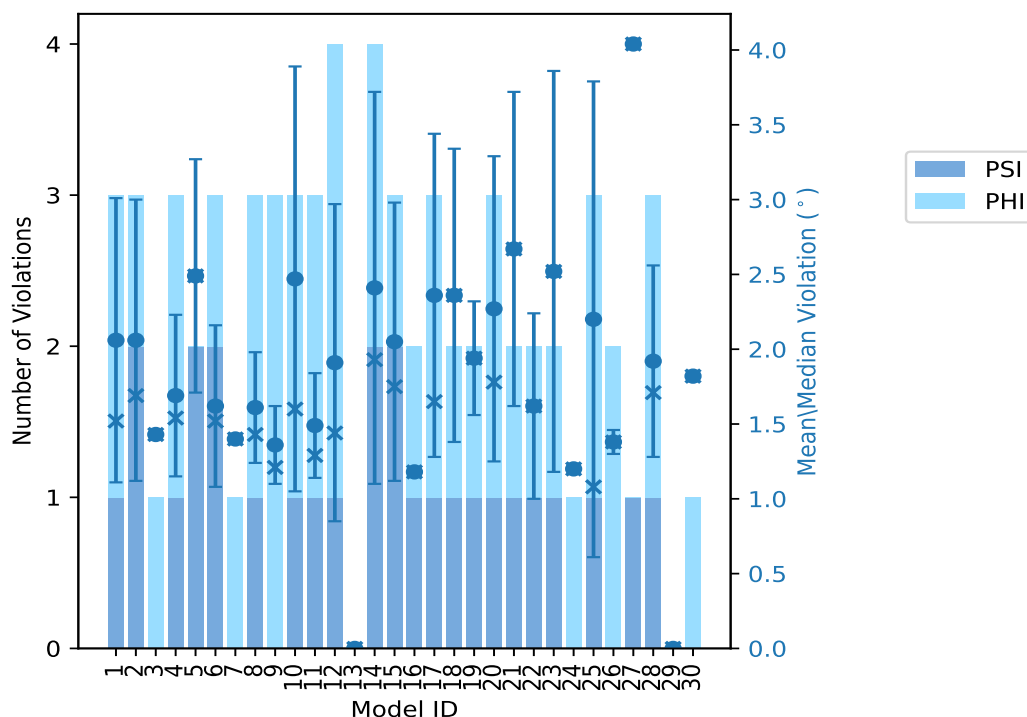
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

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° are not included in the statistics.

Model ID	Number of violations			Mean (°)	Max (°)	SD (°)	Median (°)
	PSI	PHI	Total				
1	1	2	3	2.06	3.39	0.95	1.52
2	2	1	3	2.06	3.35	0.94	1.69
3	0	1	1	1.43	1.43	0.0	1.43
4	1	2	3	1.69	2.41	0.54	1.54
5	2	0	2	2.49	3.27	0.78	2.49
6	2	1	3	1.62	2.33	0.54	1.52
7	0	1	1	1.4	1.4	0.0	1.4
8	1	2	3	1.61	2.13	0.37	1.43
9	0	3	3	1.36	1.72	0.26	1.21
10	1	2	3	2.47	4.47	1.42	1.6
11	1	2	3	1.49	1.98	0.35	1.29
12	1	3	4	1.91	3.72	1.06	1.44
13	0	0	0	0.0	0.0	0.0	0.0
14	2	2	4	2.41	4.61	1.31	1.93
15	2	1	3	2.05	3.3	0.93	1.75
16	1	1	2	1.18	1.21	0.03	1.18
17	1	2	3	2.36	3.88	1.08	1.65
18	1	1	2	2.36	3.33	0.98	2.36
19	1	1	2	1.94	2.33	0.38	1.94
20	1	2	3	2.27	3.7	1.02	1.78
21	1	1	2	2.67	3.72	1.05	2.67
22	1	1	2	1.62	2.23	0.62	1.62
23	1	1	2	2.52	3.85	1.34	2.52
24	0	1	1	1.2	1.2	0.0	1.2
25	1	2	3	2.2	4.44	1.59	1.08
26	0	2	2	1.38	1.46	0.08	1.38
27	1	0	1	4.04	4.04	0.0	4.04
28	1	2	3	1.92	2.79	0.64	1.71
29	0	0	0	0.0	0.0	0.0	0.0
30	0	1	1	1.82	1.82	0.0	1.82

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



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

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

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

Number of violated restraints			Fraction of the ensemble	
PSI	PHI	Total	Count ¹	%
3	2	5	1	3.3
0	0	0	2	6.7
2	0	2	3	10.0
0	1	1	4	13.3
0	0	0	5	16.7
0	0	0	6	20.0
0	0	0	7	23.3
0	0	0	8	26.7
0	0	0	9	30.0
0	0	0	10	33.3
0	1	1	11	36.7

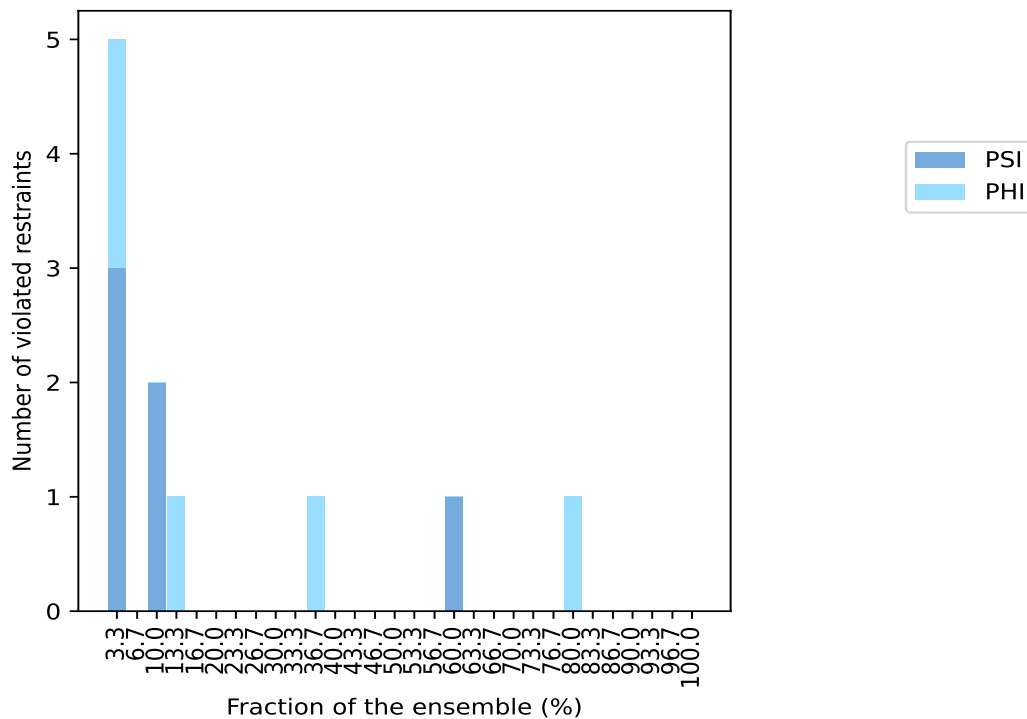
Continued on next page...

Continued from previous page...

Number of violated restraints			Fraction of the ensemble	
PSI	PHI	Total	Count ¹	%
0	0	0	12	40.0
0	0	0	13	43.3
0	0	0	14	46.7
0	0	0	15	50.0
0	0	0	16	53.3
0	0	0	17	56.7
1	0	1	18	60.0
0	0	0	19	63.3
0	0	0	20	66.7
0	0	0	21	70.0
0	0	0	22	73.3
0	0	0	23	76.7
0	1	1	24	80.0
0	0	0	25	83.3
0	0	0	26	86.7
0	0	0	27	90.0
0	0	0	28	93.3
0	0	0	29	96.7
0	0	0	30	100.0

¹ Number of models with violations

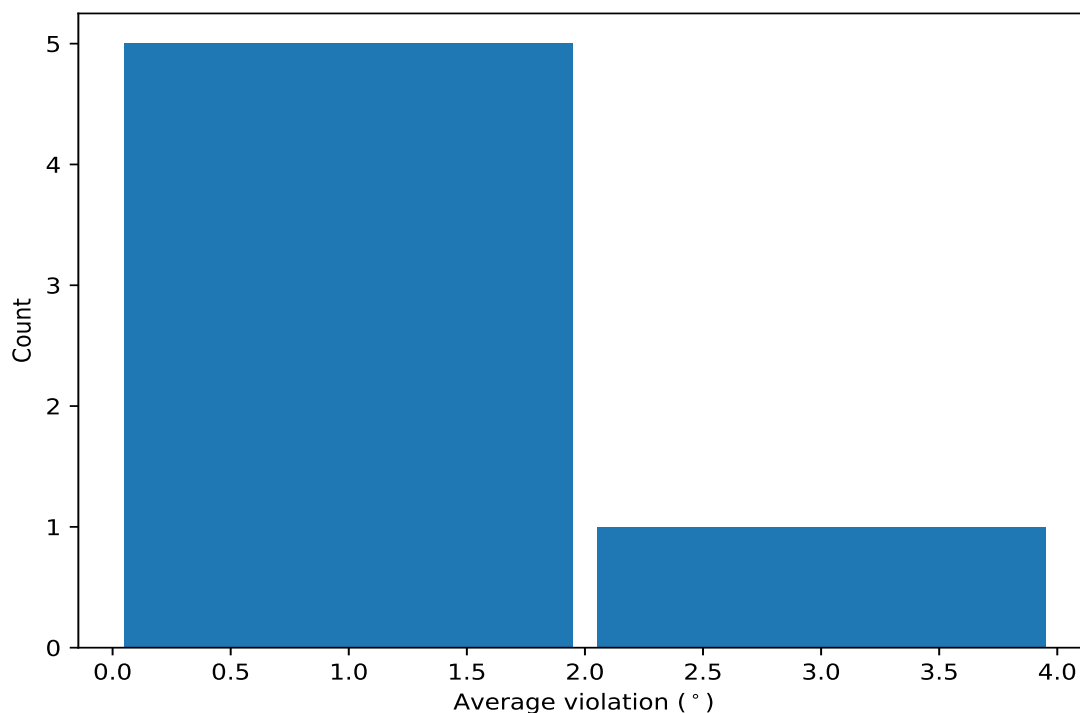
10.3.1 Bar graph : Dihedral-angle Violation statistics for the ensemble [i](#)



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

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

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



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

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

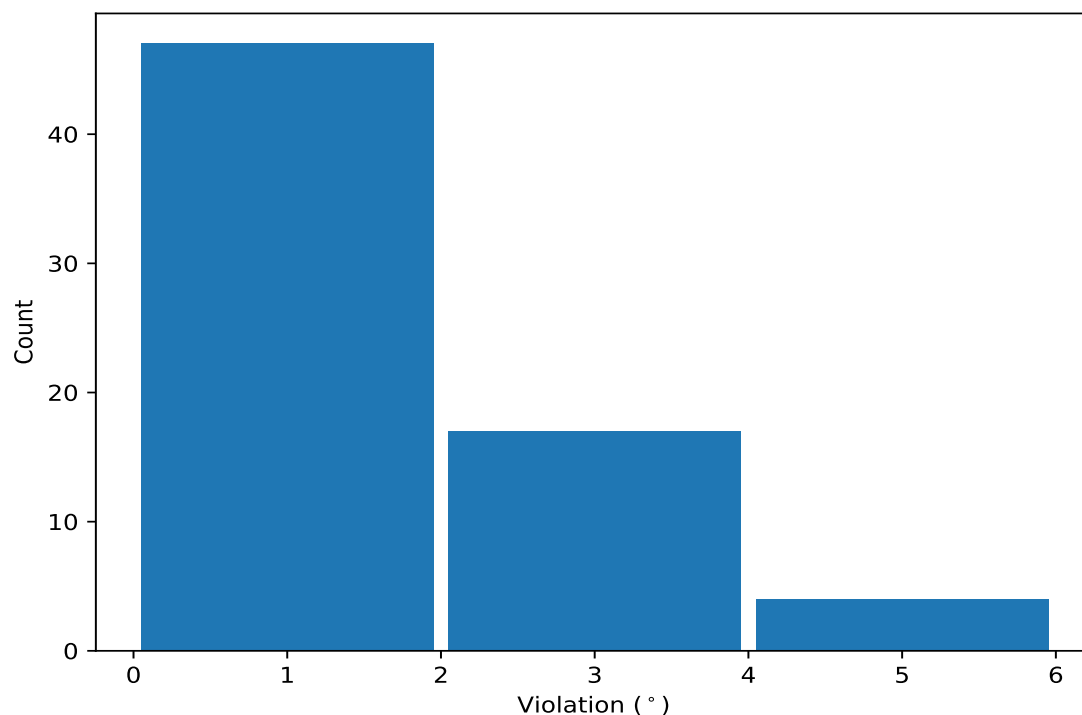
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models ¹	Mean	SD ²	Median
(1,97)	1:72:A:GLU:C	1:73:A:LEU:N	1:73:A:LEU:CA	1:73:A:LEU:C	24	1.56	0.29	1.53
(1,132)	1:90:A:LEU:N	1:90:A:LEU:CA	1:90:A:LEU:C	1:91:A:SER:N	18	3.43	0.82	3.54
(1,25)	1:33:A:ARG:C	1:34:A:ASP:N	1:34:A:ASP:CA	1:34:A:ASP:C	11	1.33	0.18	1.34
(1,55)	1:51:A:GLU:C	1:52:A:PHE:N	1:52:A:PHE:CA	1:52:A:PHE:C	4	1.13	0.11	1.12
(1,104)	1:76:A:GLU:N	1:76:A:GLU:CA	1:76:A:GLU:C	1:77:A:GLY:N	3	1.63	0.38	1.56
(1,130)	1:89:A:ASP:N	1:89:A:ASP:CA	1:89:A:ASP:C	1:90:A:LEU:N	3	1.31	0.28	1.13

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

10.5 All violated dihedral-angle restraints [i](#)

10.5.1 Histogram : Distribution of violations [i](#)

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



10.5.2 Table: All violated dihedral-angle restraints [i](#)

The following table provides the list of violations for the 10 worst performing restraints, sorted by the violation value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint.

Key	Atom-1	Atom-2	Atom-3	Atom-4	Model ID	Violation (°)
(1,132)	1:90:A:LEU:N	1:90:A:LEU:CA	1:90:A:LEU:C	1:91:A:SER:N	14	4.61
(1,132)	1:90:A:LEU:N	1:90:A:LEU:CA	1:90:A:LEU:C	1:91:A:SER:N	10	4.47
(1,132)	1:90:A:LEU:N	1:90:A:LEU:CA	1:90:A:LEU:C	1:91:A:SER:N	25	4.44
(1,132)	1:90:A:LEU:N	1:90:A:LEU:CA	1:90:A:LEU:C	1:91:A:SER:N	27	4.04
(1,132)	1:90:A:LEU:N	1:90:A:LEU:CA	1:90:A:LEU:C	1:91:A:SER:N	17	3.88
(1,132)	1:90:A:LEU:N	1:90:A:LEU:CA	1:90:A:LEU:C	1:91:A:SER:N	23	3.85
(1,132)	1:90:A:LEU:N	1:90:A:LEU:CA	1:90:A:LEU:C	1:91:A:SER:N	12	3.72
(1,132)	1:90:A:LEU:N	1:90:A:LEU:CA	1:90:A:LEU:C	1:91:A:SER:N	21	3.72
(1,132)	1:90:A:LEU:N	1:90:A:LEU:CA	1:90:A:LEU:C	1:91:A:SER:N	20	3.7
(1,132)	1:90:A:LEU:N	1:90:A:LEU:CA	1:90:A:LEU:C	1:91:A:SER:N	1	3.39