wwPDB NMR Structure Validation Summary Report

Jun 4, 2023 – 02:36 AM EDT

PDB ID	:	2LB9
BMRB ID	:	15717
Title	:	Refined solution structure of a cyanobacterial phytochrome gaf domain in the red light-absorbing ground state (corrected pyrrole ring planarity)
Authors	:	Cornilescu, C.C.; Cornilescu, G.; Ulijasz, A.T.; Vierstra, R.D.; Markley, J.L.
Deposited on	:	2011-03-23

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:

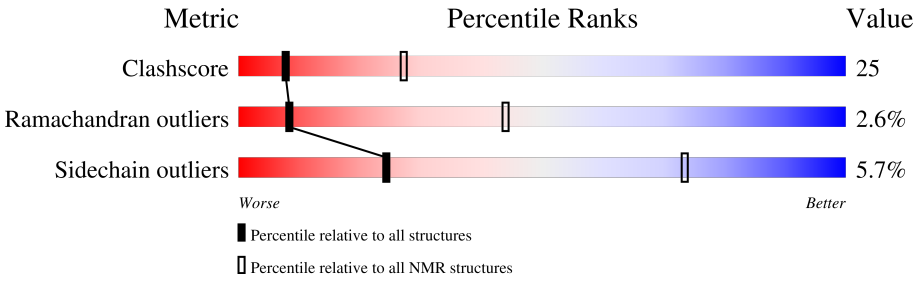
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.33

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 64%.

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	208	49% 23% • 10% 17%

2 Ensemble composition and analysis

This entry contains 20 models. Model 18 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:31-A:112, A:133-A:202 (152)	0.53	18

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

3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2776 atoms, of which 1390 are hydrogens and 0 are deuteriums.

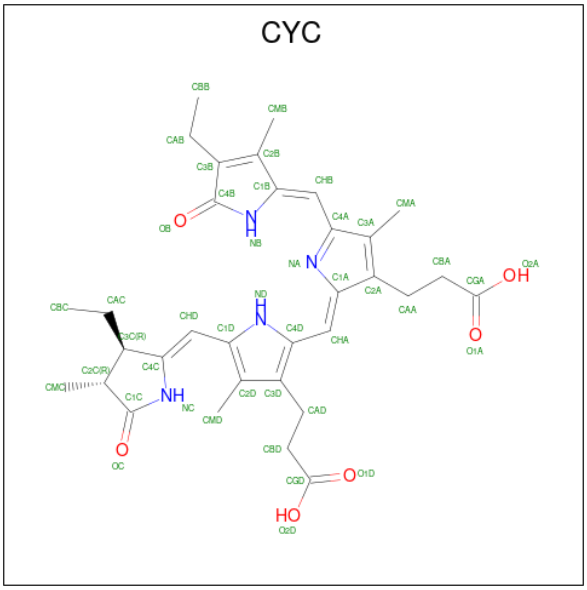
- Molecule 1 is a protein called Sensor histidine kinase.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	172	2695	850	1352	243	247	3	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	201	SER	-	expression tag	UNP Q2JIZ5
A	202	LEU	-	expression tag	UNP Q2JIZ5
A	203	HIS	-	expression tag	UNP Q2JIZ5
A	204	HIS	-	expression tag	UNP Q2JIZ5
A	205	HIS	-	expression tag	UNP Q2JIZ5
A	206	HIS	-	expression tag	UNP Q2JIZ5
A	207	HIS	-	expression tag	UNP Q2JIZ5
A	208	HIS	-	expression tag	UNP Q2JIZ5

- Molecule 2 is PHYCOCYANOBILIN (three-letter code: CYC) (formula: $C_{33}H_{40}N_4O_6$).



Mol	Chain	Residues	Atoms				
			Total	C	H	N	O
2	A	1	81	33	38	4	6

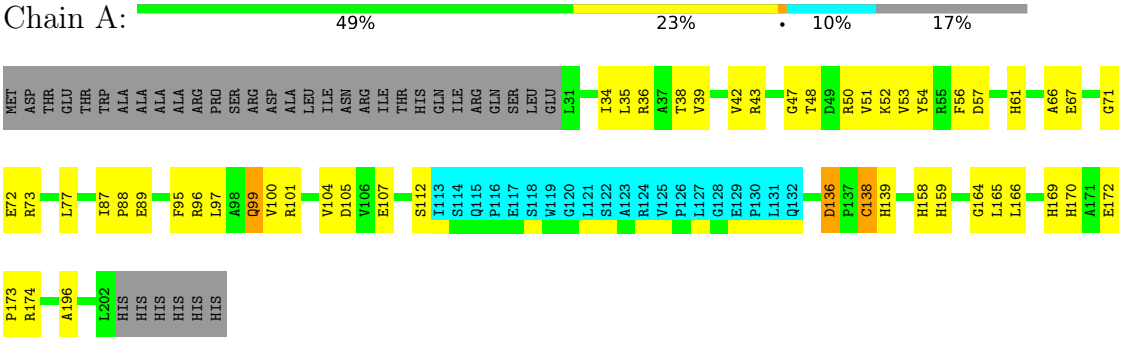
4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

• Molecule 1: Sensor histidine kinase

Chain A:

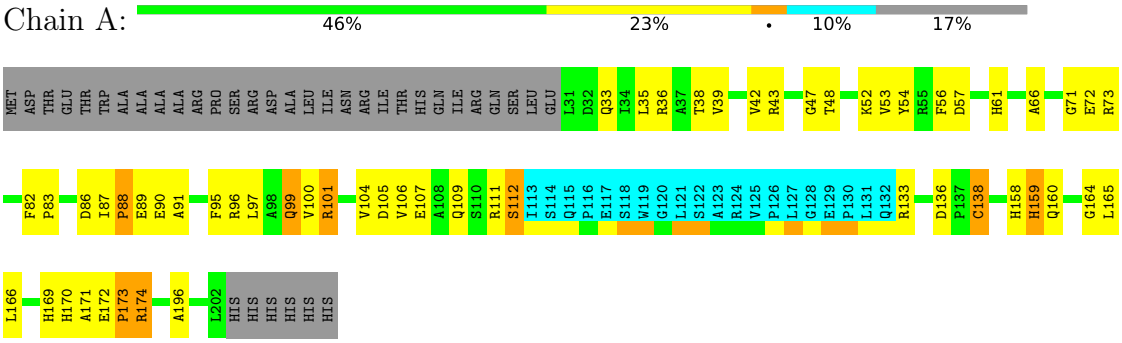


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 18. Colouring as in section 4.1 above.

• Molecule 1: Sensor histidine kinase

Chain A:



5 Refinement protocol and experimental data overview

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

Of the 100 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
X-PLOR NIH	geometry optimization	
X-PLOR NIH	structure solution	
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	1490
Number of shifts mapped to atoms	1462
Number of unparsed shifts	0
Number of shifts with mapping errors	28
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	64%

6 Model quality

6.1 Standard geometry

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

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	1191	1201	1200	62±7
2	A	43	38	37	6±2
All	All	24680	24780	24740	1256

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 25.

5 of 448 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:139:HIS:N	1:A:139:HIS:ND1	0.91	2.15	10	1
1:A:138:CYS:SG	2:A:301:CYC:HHD	0.88	2.08	15	8
1:A:43:ARG:O	1:A:47:GLY:N	0.78	2.16	2	20
1:A:34:ILE:O	1:A:38:THR:HG23	0.77	1.79	6	16
1:A:99:GLN:N	1:A:99:GLN:NE2	0.76	2.34	14	13

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	150/208 (72%)	135±2 (90±1%)	11±2 (8±1%)	4±1 (3±1%)	8	44
All	All	3000/4160 (72%)	2696 (90%)	225 (8%)	79 (3%)	8	44

5 of 7 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	88	PRO	19
1	A	72	GLU	18
1	A	89	GLU	13
1	A	159	HIS	10
1	A	173	PRO	8

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	129/177 (73%)	122±2 (94±1%)	7±2 (6±1%)	24	73
All	All	2580/3540 (73%)	2432 (94%)	148 (6%)	24	73

5 of 33 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	138	CYS	20
1	A	112	SER	14
1	A	99	GLN	13
1	A	136	ASP	12
1	A	101	ARG	11

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)

1 ligand 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	CYC	A	301	1	42,46,46	1.88±0.01	7±0 (16±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	CYC	A	301	1	50,67,67	1.89±0.01	14±0 (28±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	CYC	A	301	1	-	1±0,25,74,74	0±0,4,4,4

5 of 7 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
2	A	301	CYC	CHA-C1A	6.76	1.40	1.35	12	20
2	A	301	CYC	C4B-C3B	4.91	1.38	1.48	3	20
2	A	301	CYC	C1A-C2A	3.66	1.39	1.45	15	20
2	A	301	CYC	C1B-C2B	3.09	1.39	1.45	11	20
2	A	301	CYC	C1A-NA	2.48	1.33	1.38	4	20

5 of 15 unique angle 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(^{o})	Ideal(^{o})	Models	
								Worst	Total
2	A	301	CYC	CHD-C4C-NC	5.48	131.73	125.20	20	20
2	A	301	CYC	C4D-CHA-C1A	4.47	123.47	128.81	9	20
2	A	301	CYC	CHB-C1B-C2B	3.63	119.75	126.95	6	20
2	A	301	CYC	CMB-C2B-C1B	3.54	128.58	124.17	10	20
2	A	301	CYC	C1A-NA-C4A	3.24	112.62	106.51	8	20

There are no chirality outliers.

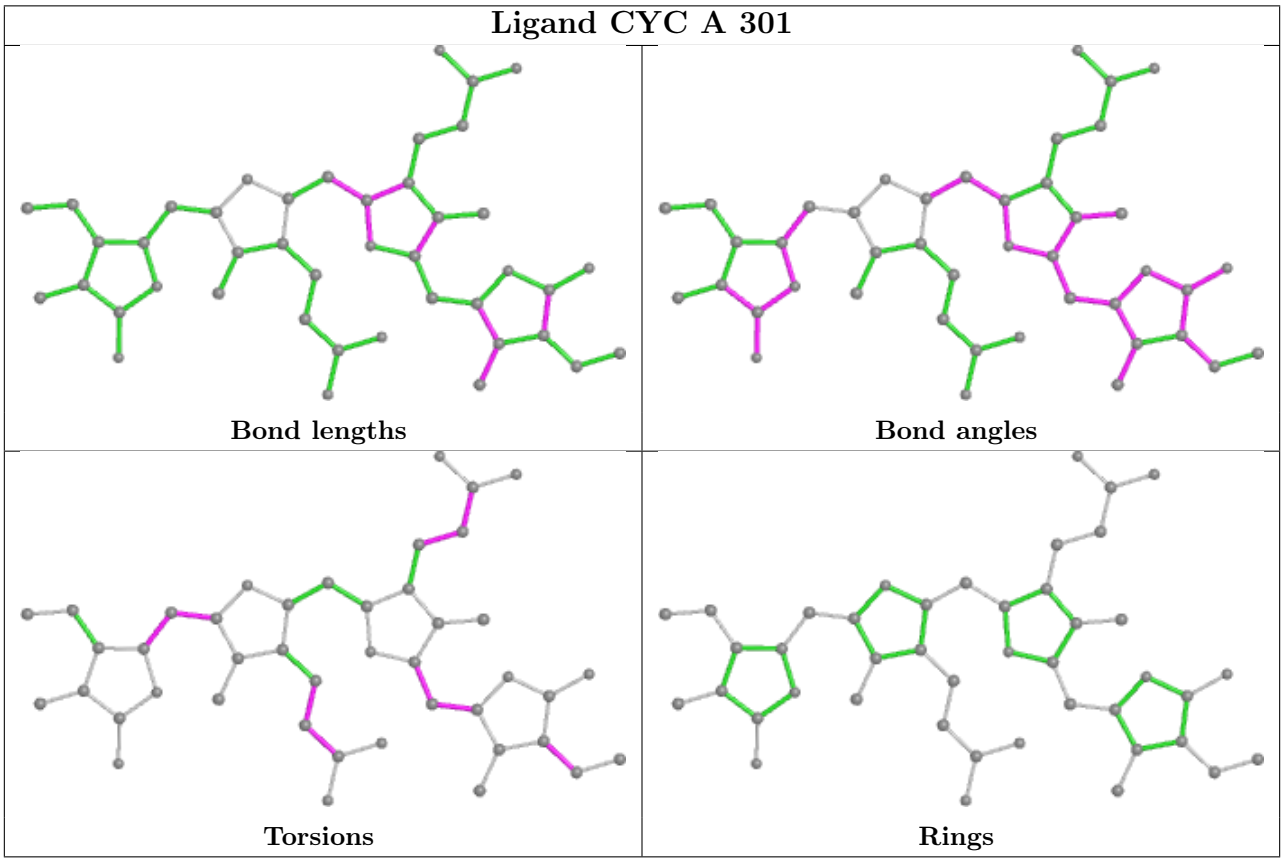
All unique torsion outliers are listed below.

Mol	Chain	Res	Type	Atoms	Models (Total)
2	A	301	CYC	NC-C4C-CHD-C1D	9

There are no ring outliers.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier.

The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



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 64% for the well-defined parts and 61% 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	1490
Number of shifts mapped to atoms	1462
Number of unparsed shifts	0
Number of shifts with mapping errors	28
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	2

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 28) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	12	PRO	C	176.214	.	.
1	A	12	PRO	CA	63.477	.	.
1	A	12	PRO	CB	32.013	.	.
1	A	13	SER	H	7.993	.	.
1	A	13	SER	C	177.801	.	.
1	A	13	SER	CA	64.627	.	.
1	A	13	SER	CB	64.76	.	.
1	A	13	SER	N	121.843	.	.
1	A	14	ARG	H	7.522	.	.
1	A	14	ARG	N	123.632	.	.
1	A	17	LEU	HA	3.816	.	.
1	A	17	LEU	HB2	1.648	.	.
1	A	17	LEU	C	177.942	.	.
1	A	17	LEU	CA	58.148	.	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	17	LEU	CB	41.845	.	.
1	A	17	LEU	CD1	25.386	.	.
1	A	18	ILE	H	7.91	.	.
1	A	18	ILE	N	118.063	.	.
1	A	26	ARG	C	178.781	.	.
1	A	26	ARG	CA	60.067	.	.
1	A	26	ARG	CB	29.282	.	.
1	A	27	GLN	H	8.003	.	.
1	A	27	GLN	C	176.793	.	.
1	A	27	GLN	CA	58.619	.	.
1	A	27	GLN	CB	30.212	.	.
1	A	27	GLN	N	118.92	.	.
1	A	28	SER	H	8.175	.	.
1	A	28	SER	N	118.066	.	.

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	159	-0.17 ± 0.10	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	148	-0.09 ± 0.11	None needed (< 0.5 ppm)
$^{13}C'$	144	-0.14 ± 0.16	None needed (< 0.5 ppm)
^{15}N	142	0.17 ± 0.26	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 64%, i.e. 1368 atoms were assigned a chemical shift out of a possible 2122. 0 out of 39 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	656/751 (87%)	254/304 (84%)	273/304 (90%)	129/143 (90%)
Sidechain	658/1243 (53%)	417/815 (51%)	237/378 (63%)	4/50 (8%)
Aromatic	54/128 (42%)	32/66 (48%)	21/54 (39%)	1/8 (12%)
Overall	1368/2122 (64%)	703/1185 (59%)	531/736 (72%)	134/201 (67%)

7.1.4 Statistically unusual chemical shifts (i)

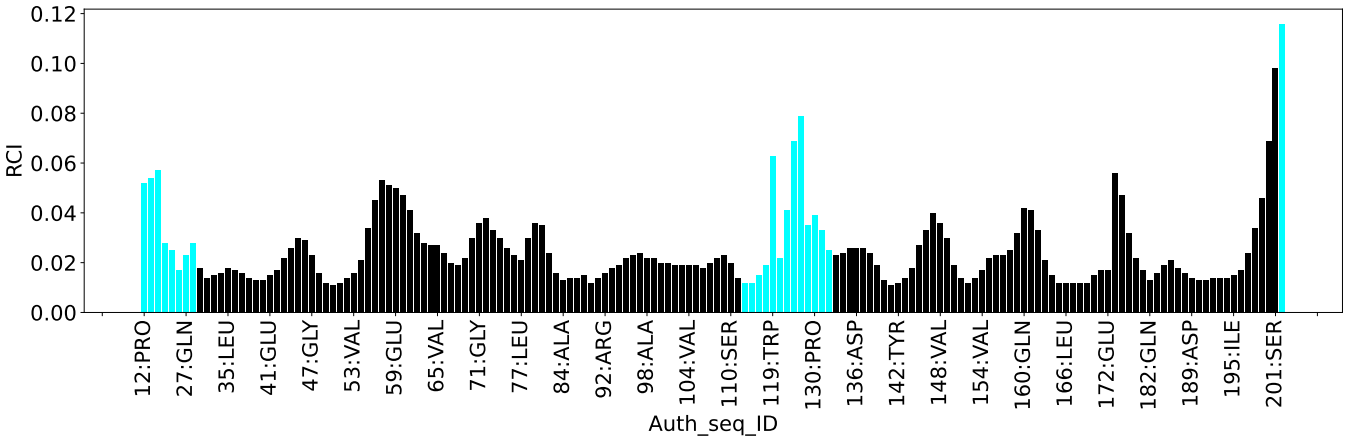
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	85	GLY	N	134.16	91.59 – 127.52	6.8
1	A	77	LEU	H	11.57	5.09 – 11.34	5.4

7.1.5 Random Coil Index (RCI) plots (i)

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

Random coil index (RCI) for chain A:



8 NMR restraints analysis

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	2492
Intra-residue (|i-j|=0)	865
Sequential (|i-j|=1)	688
Medium range (|i-j|>1 and |i-j|<5)	364
Long range (|i-j|≥5)	417
Inter-chain	0
Hydrogen bond restraints	158
Disulfide bond restraints	0
Total dihedral-angle restraints	504
Number of unmapped restraints	26
Number of restraints per residue	14.4
Number of long range restraints per residue[1]	2.3

[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)	92.0	0.2
0.2-0.5 (Medium)	103.5	0.5
>0.5 (Large)	2.3	0.78

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)	42.0	10.0
10.0-20.0 (Medium)	5.3	19.8
>20.0 (Large)	1.2	150.0

9 Distance violation analysis (i)

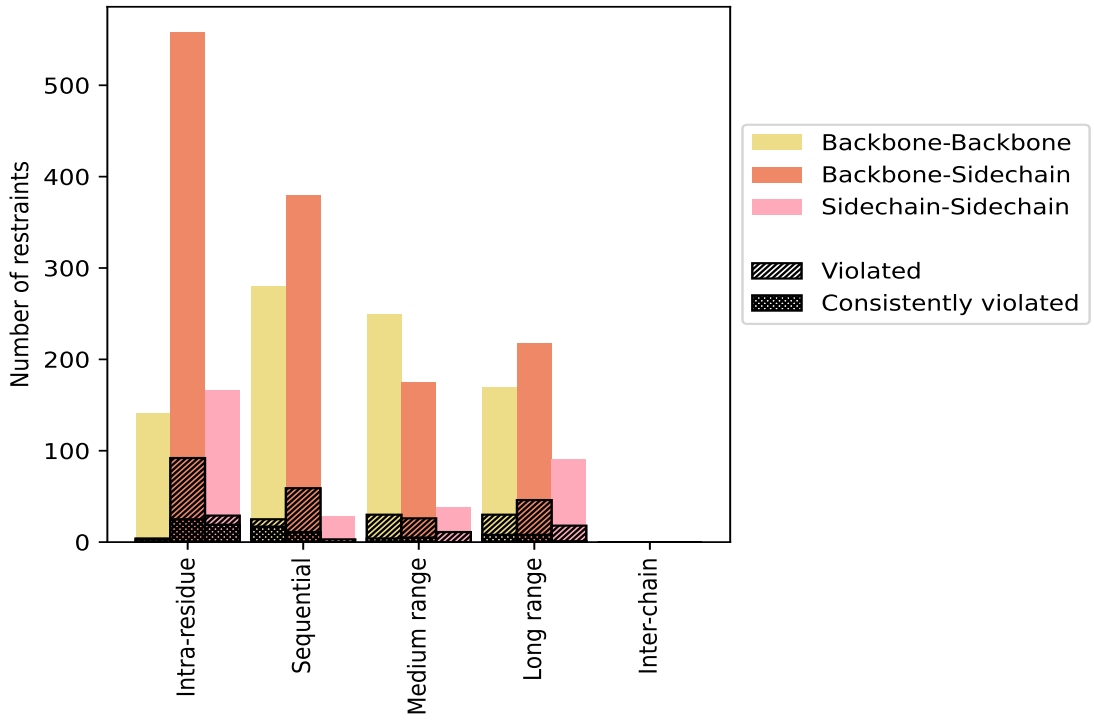
9.1 Summary of distance violations (i)

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)	865	34.7	125	14.5	5.0	47	5.4	1.9
Backbone-Backbone	141	5.7	4	2.8	0.2	3	2.1	0.1
Backbone-Sidechain	558	22.4	92	16.5	3.7	25	4.5	1.0
Sidechain-Sidechain	166	6.7	29	17.5	1.2	19	11.4	0.8
Sequential (|i-j|=1)	688	27.6	87	12.6	3.5	28	4.1	1.1
Backbone-Backbone	280	11.2	25	8.9	1.0	17	6.1	0.7
Backbone-Sidechain	380	15.2	59	15.5	2.4	11	2.9	0.4
Sidechain-Sidechain	28	1.1	3	10.7	0.1	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	364	14.6	44	12.1	1.8	7	1.9	0.3
Backbone-Backbone	151	6.1	7	4.6	0.3	2	1.3	0.1
Backbone-Sidechain	175	7.0	26	14.9	1.0	5	2.9	0.2
Sidechain-Sidechain	38	1.5	11	28.9	0.4	0	0.0	0.0
Long range (|i-j|≥5)	417	16.7	78	18.7	3.1	13	3.1	0.5
Backbone-Backbone	109	4.4	14	12.8	0.6	4	3.7	0.2
Backbone-Sidechain	217	8.7	46	21.2	1.8	8	3.7	0.3
Sidechain-Sidechain	91	3.7	18	19.8	0.7	1	1.1	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	158	6.3	39	24.7	1.6	6	3.8	0.2
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	2492	100.0	373	15.0	15.0	101	4.1	4.1
Backbone-Backbone	839	33.7	89	10.6	3.6	32	3.8	1.3
Backbone-Sidechain	1330	53.4	223	16.8	8.9	49	3.7	2.0
Sidechain-Sidechain	323	13.0	61	18.9	2.4	20	6.2	0.8

[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	77	53	33	46	0	209	0.24	0.63	0.1	0.22
2	75	56	28	43	0	202	0.24	0.64	0.1	0.22
3	80	47	33	44	0	204	0.23	0.53	0.09	0.2
4	83	56	26	46	0	211	0.23	0.6	0.1	0.21
5	75	44	32	47	0	198	0.24	0.6	0.1	0.22
6	77	50	27	42	0	196	0.24	0.78	0.1	0.22
7	80	49	31	45	0	205	0.23	0.65	0.1	0.21
8	78	51	33	36	0	198	0.23	0.5	0.1	0.21
9	76	46	25	42	0	189	0.24	0.59	0.1	0.21
10	81	46	33	40	0	200	0.24	0.72	0.11	0.21
11	77	54	26	42	0	199	0.23	0.69	0.1	0.21

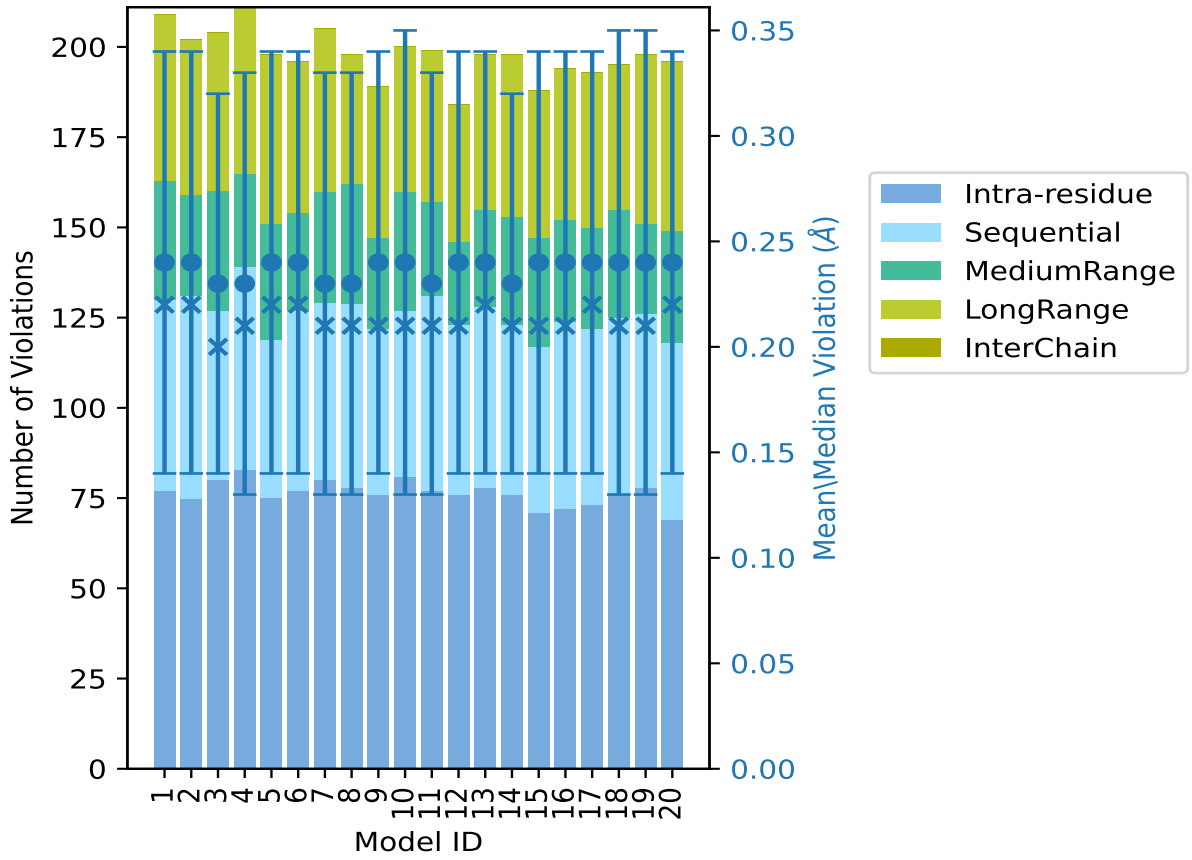
Continued on next page...

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	76	47	23	38	0	184	0.24	0.55	0.1	0.21
13	78	50	27	43	0	198	0.24	0.57	0.1	0.22
14	76	47	30	45	0	198	0.23	0.54	0.09	0.21
15	71	46	30	41	0	188	0.24	0.58	0.1	0.21
16	72	52	28	42	0	194	0.24	0.6	0.1	0.21
17	73	49	28	43	0	193	0.24	0.56	0.1	0.22
18	76	49	30	40	0	195	0.24	0.53	0.11	0.21
19	78	48	25	47	0	198	0.24	0.63	0.11	0.21
20	69	49	31	47	0	196	0.24	0.62	0.1	0.22

[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

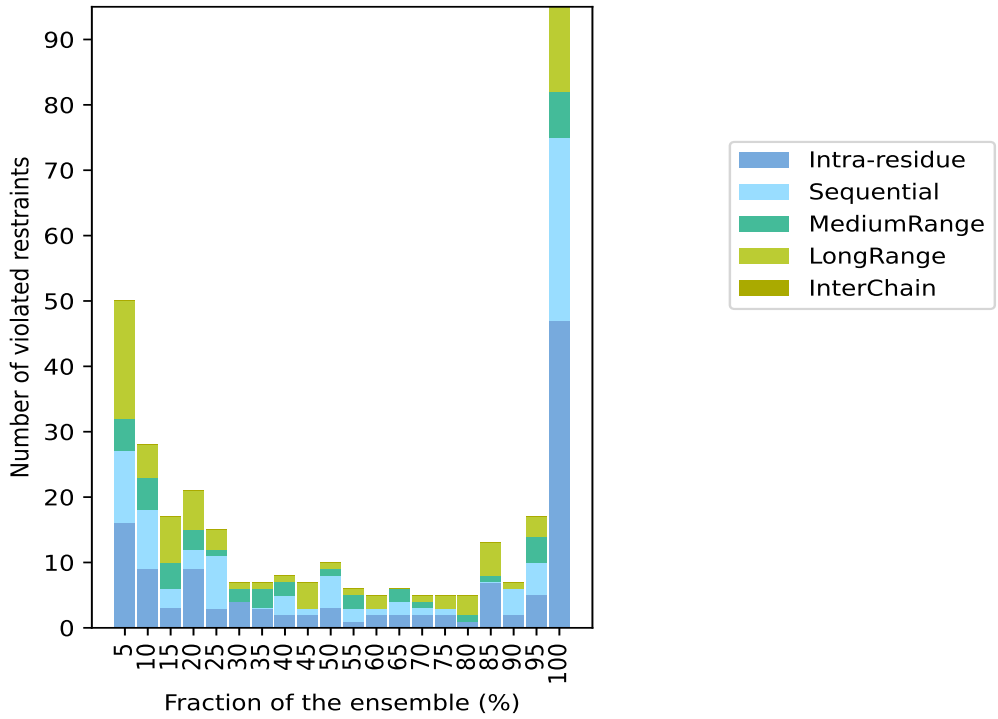
9.3 Distance violation statistics for the ensemble (i)

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, 2000(IR:740, SQ:601, MR:320, LR:339, 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]	%
16	11	5	18	0	50	1	5.0
9	9	5	5	0	28	2	10.0
3	3	4	7	0	17	3	15.0
9	3	3	6	0	21	4	20.0
3	8	1	3	0	15	5	25.0
4	0	2	1	0	7	6	30.0
3	0	3	1	0	7	7	35.0
2	3	2	1	0	8	8	40.0
2	1	0	4	0	7	9	45.0
3	5	1	1	0	10	10	50.0
1	2	2	1	0	6	11	55.0
2	1	0	2	0	5	12	60.0
2	2	2	0	0	6	13	65.0
2	1	1	1	0	5	14	70.0
2	1	0	2	0	5	15	75.0
1	0	1	3	0	5	16	80.0
7	0	1	5	0	13	17	85.0
2	4	0	1	0	7	18	90.0
5	5	4	3	0	17	19	95.0
47	28	7	13	0	95	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

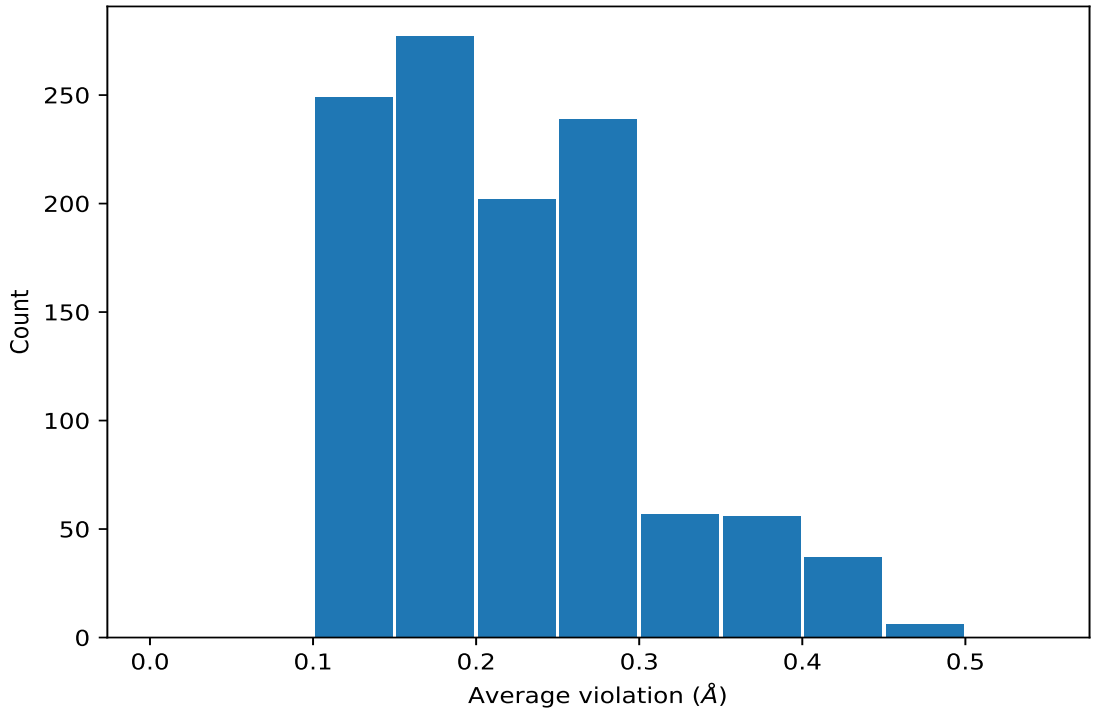
9.3.1 Bar graph : Distance violation statistics for the ensemble



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)
(1,2087)	1:A:157:MET:H	1:A:195:ILE:HG12	20	0.49	0.15	0.52
(1,2087)	1:A:157:MET:H	1:A:195:ILE:HG13	20	0.49	0.15	0.52
(1,310)	1:A:78:LEU:HB2	1:A:78:LEU:H	20	0.44	0.07	0.45
(1,310)	1:A:78:LEU:HB3	1:A:78:LEU:H	20	0.44	0.07	0.45
(1,1719)	1:A:104:VAL:H	1:A:151:SER:H	20	0.43	0.05	0.42
(1,666)	1:A:135:VAL:HA	1:A:136:ASP:H	20	0.42	0.04	0.43
(1,1541)	1:A:80:LEU:HG	1:A:81:THR:H	20	0.42	0.03	0.42
(1,711)	1:A:143:LEU:HA	1:A:143:LEU:HD11	20	0.42	0.04	0.42
(1,711)	1:A:143:LEU:HA	1:A:143:LEU:HD12	20	0.42	0.04	0.42
(1,711)	1:A:143:LEU:HA	1:A:143:LEU:HD13	20	0.42	0.04	0.42
(1,711)	1:A:143:LEU:HA	1:A:143:LEU:HD21	20	0.42	0.04	0.42
(1,711)	1:A:143:LEU:HA	1:A:143:LEU:HD22	20	0.42	0.04	0.42
(1,711)	1:A:143:LEU:HA	1:A:143:LEU:HD23	20	0.42	0.04	0.42
(1,894)	1:A:164:GLY:HA2	1:A:165:LEU:H	20	0.41	0.0	0.41
(1,894)	1:A:164:GLY:HA3	1:A:165:LEU:H	20	0.41	0.0	0.41
(1,1293)	1:A:51:VAL:H	1:A:168:SER:HA	20	0.4	0.1	0.39

Continued on next page...

Continued from previous page...

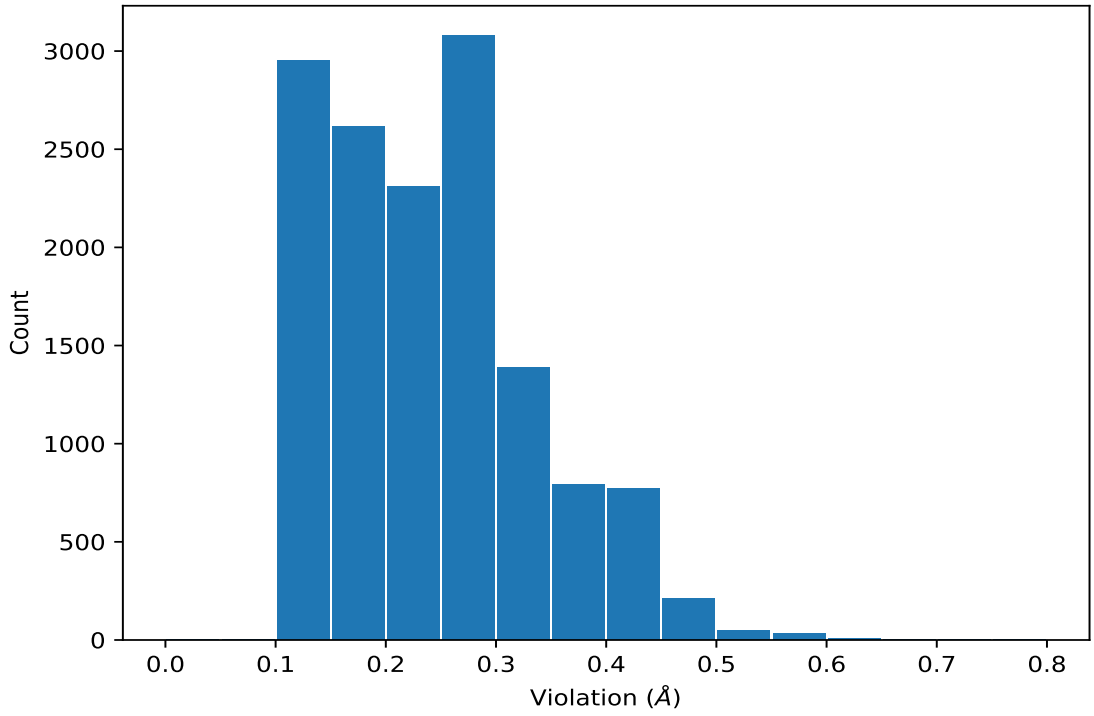
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,902)	1:A:166:LEU:HB2	1:A:166:LEU:HD11	20	0.4	0.01	0.4
(1,902)	1:A:166:LEU:HB2	1:A:166:LEU:HD12	20	0.4	0.01	0.4
(1,902)	1:A:166:LEU:HB2	1:A:166:LEU:HD13	20	0.4	0.01	0.4
(1,902)	1:A:166:LEU:HB2	1:A:166:LEU:HD21	20	0.4	0.01	0.4
(1,902)	1:A:166:LEU:HB2	1:A:166:LEU:HD22	20	0.4	0.01	0.4
(1,902)	1:A:166:LEU:HB2	1:A:166:LEU:HD23	20	0.4	0.01	0.4
(1,902)	1:A:166:LEU:HB3	1:A:166:LEU:HD11	20	0.4	0.01	0.4
(1,902)	1:A:166:LEU:HB3	1:A:166:LEU:HD12	20	0.4	0.01	0.4
(1,902)	1:A:166:LEU:HB3	1:A:166:LEU:HD13	20	0.4	0.01	0.4
(1,902)	1:A:166:LEU:HB3	1:A:166:LEU:HD21	20	0.4	0.01	0.4
(1,902)	1:A:166:LEU:HB3	1:A:166:LEU:HD22	20	0.4	0.01	0.4
(1,902)	1:A:166:LEU:HB3	1:A:166:LEU:HD23	20	0.4	0.01	0.4
(1,189)	1:A:52:LYS:HA	1:A:52:LYS:HB2	20	0.4	0.04	0.42

[1]Number of violated models, [2]Standard deviation

9.5 All violated distance restraints

9.5.1 Histogram : Distribution of distance violations

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)
(1,1497)	1:A:73:ARG:HG2	1:A:74:LEU:H	6	0.78
(1,1497)	1:A:73:ARG:HG3	1:A:74:LEU:H	6	0.78
(1,2087)	1:A:157:MET:H	1:A:195:ILE:HG12	10	0.72
(1,2087)	1:A:157:MET:H	1:A:195:ILE:HG13	10	0.72
(1,2087)	1:A:157:MET:H	1:A:195:ILE:HG12	11	0.69
(1,2087)	1:A:157:MET:H	1:A:195:ILE:HG13	11	0.69
(1,2087)	1:A:157:MET:H	1:A:195:ILE:HG12	7	0.65
(1,2087)	1:A:157:MET:H	1:A:195:ILE:HG13	7	0.65
(1,482)	1:A:101:ARG:HA	1:A:101:ARG:HB2	2	0.64
(1,482)	1:A:101:ARG:HA	1:A:101:ARG:HB3	2	0.64
(1,637)	1:A:130:PRO:HB2	1:A:132:GLN:H	1	0.63
(1,637)	1:A:130:PRO:HB3	1:A:132:GLN:H	1	0.63
(1,192)	1:A:52:LYS:HB2	1:A:67:GLU:HA	19	0.63
(1,192)	1:A:52:LYS:HB3	1:A:67:GLU:HA	19	0.63
(1,1270)	1:A:49:ASP:H	1:A:169:HIS:HB2	20	0.62
(1,1270)	1:A:49:ASP:H	1:A:169:HIS:HB3	20	0.62
(1,310)	1:A:78:LEU:HB2	1:A:78:LEU:H	16	0.6
(1,310)	1:A:78:LEU:HB3	1:A:78:LEU:H	16	0.6
(1,2087)	1:A:157:MET:H	1:A:195:ILE:HG12	4	0.6

10 Dihedral-angle violation analysis

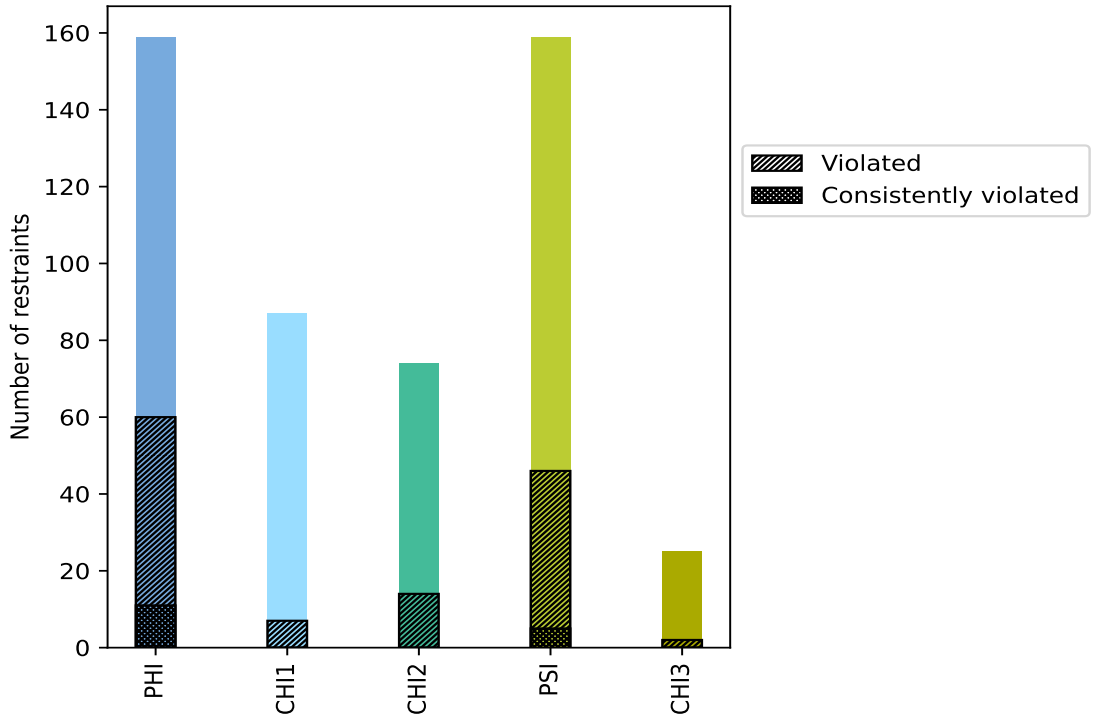
10.1 Summary of dihedral-angle violations

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	159	31.5	60	37.7	11.9	11	6.9	2.2
CHI1	87	17.3	7	8.0	1.4	0	0.0	0.0
CHI2	74	14.7	14	18.9	2.8	0	0.0	0.0
PSI	159	31.5	46	28.9	9.1	5	3.1	1.0
CHI3	25	5.0	2	8.0	0.4	0	0.0	0.0
Total	504	100.0	129	25.6	25.6	16	3.2	3.2

[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



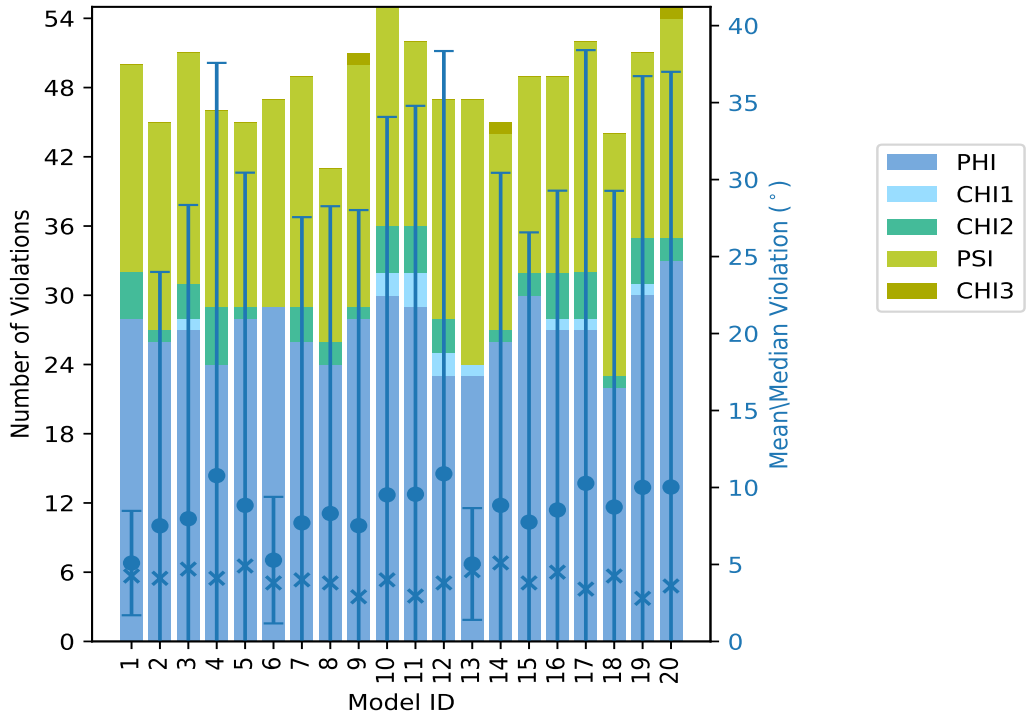
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

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 (°)
	PHI	CHI1	CHI2	PSI	CHI3	Total				
1	28	0	4	18	0	50	5.09	17.6	3.39	4.25
2	26	0	1	18	0	45	7.51	114.1	16.49	4.1
3	27	1	3	20	0	51	7.97	149.7	20.38	4.7
4	24	0	5	17	0	46	10.77	149.8	26.81	4.1
5	28	0	1	16	0	45	8.84	149.9	21.61	4.9
6	29	0	0	18	0	47	5.28	17.7	4.11	3.8
7	26	0	3	20	0	49	7.7	143.0	19.86	4.0
8	24	0	2	15	0	41	8.31	132.1	19.96	3.8
9	28	0	1	21	1	51	7.52	149.9	20.5	2.9
10	30	2	4	19	0	55	9.52	150.0	24.55	4.0
11	29	3	4	16	0	52	9.56	149.7	25.23	2.95
12	23	2	3	19	0	47	10.89	149.9	27.46	3.8
13	23	1	0	23	0	47	5.03	15.8	3.63	4.6
14	26	0	1	17	1	45	8.84	149.8	21.6	5.1
15	30	0	2	17	0	49	7.75	135.5	18.82	3.8
16	27	1	4	17	0	49	8.54	149.6	20.74	4.5
17	27	1	4	20	0	52	10.27	149.7	28.14	3.4
18	22	0	1	21	0	44	8.73	140.6	20.54	4.25
19	30	1	4	16	0	51	10.01	149.9	26.71	2.8
20	33	0	2	19	1	55	10.03	149.9	26.97	3.6

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	CHI1	CHI2	PSI	CHI3	Total	Count[1]	%
9	6	4	9	1	29	1	5.0
7	0	5	5	1	18	2	10.0
2	0	0	5	0	7	3	15.0
5	0	1	5	0	11	4	20.0
6	0	1	2	0	9	5	25.0
3	1	0	0	0	4	6	30.0
2	0	1	0	0	3	7	35.0
3	0	1	1	0	5	8	40.0
0	0	0	2	0	2	9	45.0
1	0	0	1	0	2	10	50.0
1	0	1	1	0	3	11	55.0

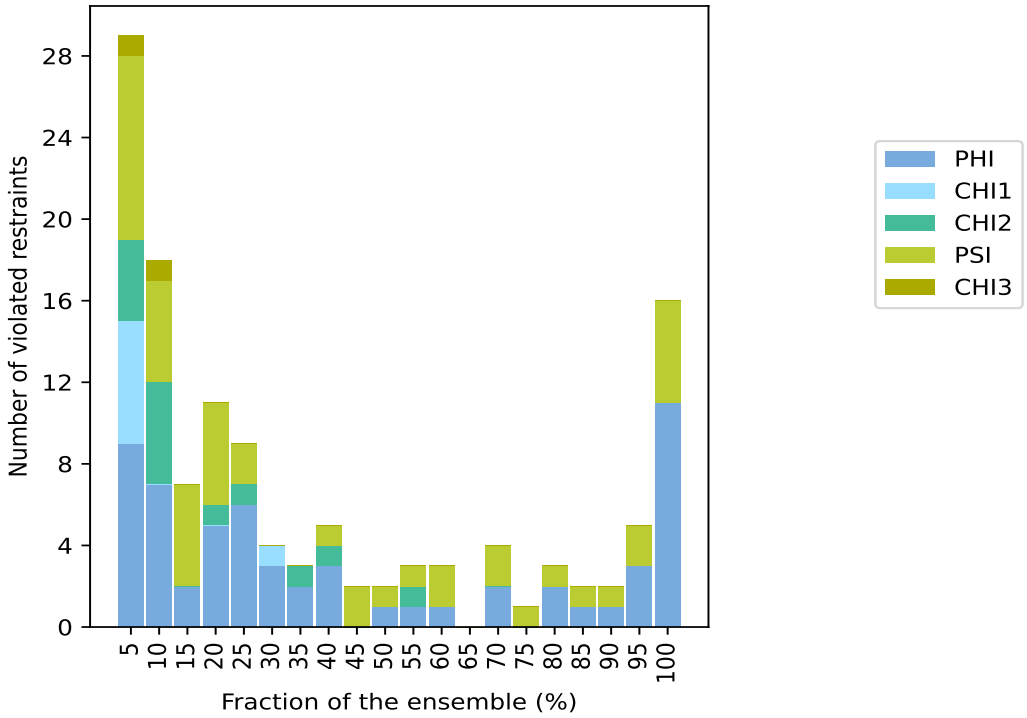
Continued on next page...

Continued from previous page...

Number of violated restraints						Fraction of the ensemble	
PHI	CHI1	CHI2	PSI	CHI3	Total	Count[1]	%
1	0	0	2	0	3	12	60.0
0	0	0	0	0	0	13	65.0
2	0	0	2	0	4	14	70.0
0	0	0	1	0	1	15	75.0
2	0	0	1	0	3	16	80.0
1	0	0	1	0	2	17	85.0
1	0	0	1	0	2	18	90.0
3	0	0	2	0	5	19	95.0
11	0	0	5	0	16	20	100.0

[1] Number of models with violations

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

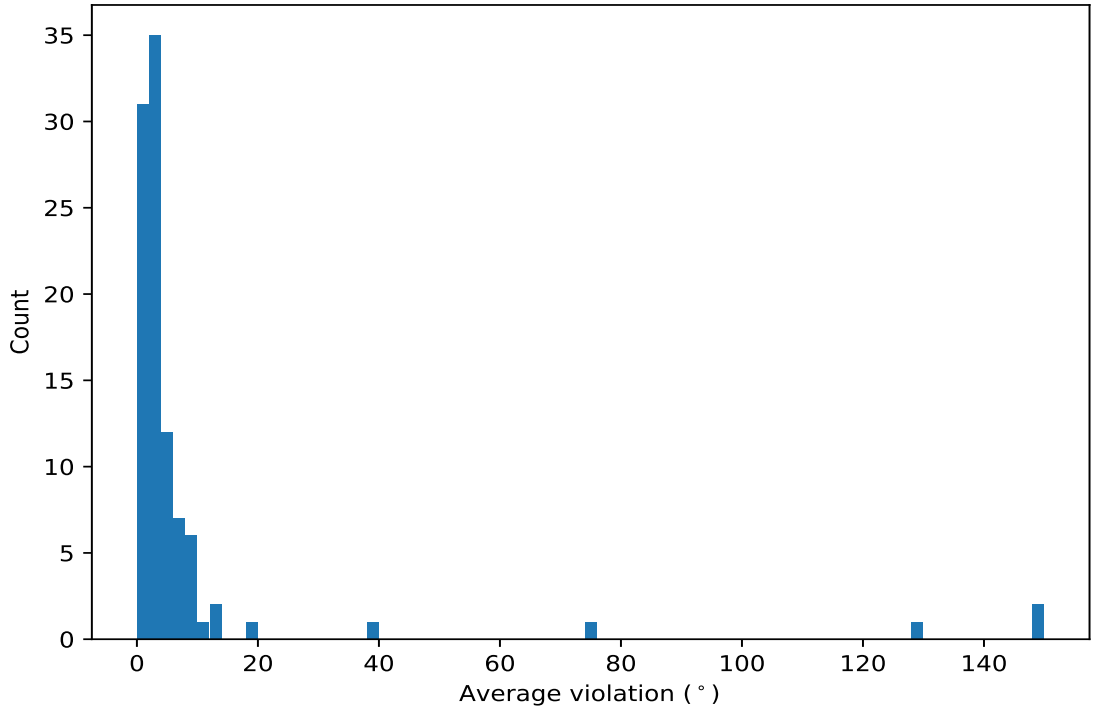


10.4 Most violated dihedral-angle restraints in the ensemble

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.

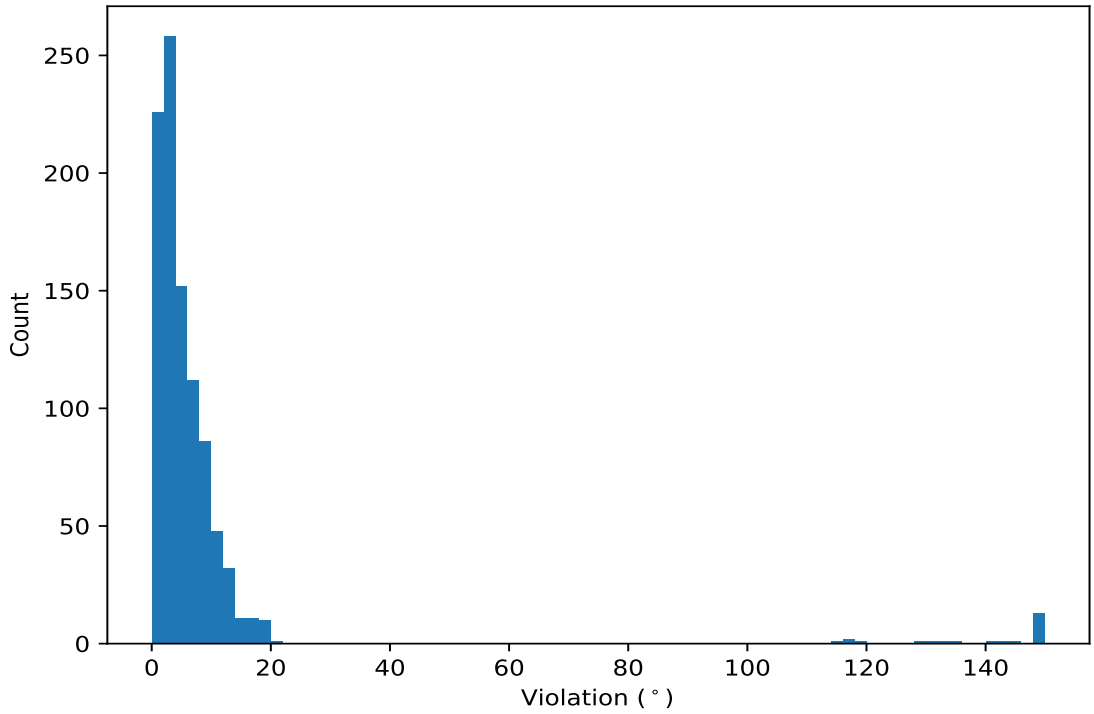
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models[1]	Mean	SD[2]	Median
(1,236)	1:A:104:VAL:N	1:A:104:VAL:CA	1:A:104:VAL:C	1:A:105:ASP:N	20	18.12	1.23	18.05
(1,140)	1:A:182:GLN:C	1:A:183:VAL:N	1:A:183:VAL:CA	1:A:183:VAL:C	20	13.14	1.35	13.3
(1,77)	1:A:103:ILE:C	1:A:104:VAL:N	1:A:104:VAL:CA	1:A:104:VAL:C	20	12.76	0.9	12.95
(1,78)	1:A:104:VAL:C	1:A:105:ASP:N	1:A:105:ASP:CA	1:A:105:ASP:C	20	11.62	1.12	11.35
(1,122)	1:A:159:HIS:C	1:A:160:GLN:N	1:A:160:GLN:CA	1:A:160:GLN:C	20	9.34	1.03	9.45
(1,106)	1:A:141:HIS:C	1:A:142:TYR:N	1:A:142:TYR:CA	1:A:142:TYR:C	20	8.56	1.16	8.8
(1,81)	1:A:107:GLU:C	1:A:108:ALA:N	1:A:108:ALA:CA	1:A:108:ALA:C	20	8.1	1.83	8.4
(1,241)	1:A:109:GLN:N	1:A:109:GLN:CA	1:A:109:GLN:C	1:A:110:SER:N	20	7.88	1.61	7.5
(1,142)	1:A:184:VAL:C	1:A:185:GLN:N	1:A:185:GLN:CA	1:A:185:GLN:C	20	7.22	1.94	7.3
(1,230)	1:A:96:ARG:N	1:A:96:ARG:CA	1:A:96:ARG:C	1:A:97:LEU:N	20	7.12	2.14	7.45

[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 (°)
(1,375)	1:A:176:TYR:CA	1:A:176:TYR:CB	1:A:176:TYR:CG	1:A:176:TYR:CD1	10	150.0
(1,375)	1:A:176:TYR:CA	1:A:176:TYR:CB	1:A:176:TYR:CG	1:A:176:TYR:CD1	5	149.9
(1,375)	1:A:176:TYR:CA	1:A:176:TYR:CB	1:A:176:TYR:CG	1:A:176:TYR:CD1	12	149.9
(1,375)	1:A:176:TYR:CA	1:A:176:TYR:CB	1:A:176:TYR:CG	1:A:176:TYR:CD1	19	149.9
(1,375)	1:A:176:TYR:CA	1:A:176:TYR:CB	1:A:176:TYR:CG	1:A:176:TYR:CD1	20	149.9
(1,365)	1:A:56:PHE:CA	1:A:56:PHE:CB	1:A:56:PHE:CG	1:A:56:PHE:CD1	9	149.9
(1,373)	1:A:142:TYR:CA	1:A:142:TYR:CB	1:A:142:TYR:CG	1:A:142:TYR:CD1	14	149.8
(1,365)	1:A:56:PHE:CA	1:A:56:PHE:CB	1:A:56:PHE:CG	1:A:56:PHE:CD1	4	149.8
(1,371)	1:A:54:TYR:CA	1:A:54:TYR:CB	1:A:54:TYR:CG	1:A:54:TYR:CD1	11	149.7
(1,364)	1:A:45:PHE:CA	1:A:45:PHE:CB	1:A:45:PHE:CG	1:A:45:PHE:CD1	3	149.7