wwPDB NMR Structure Validation Summary Report (i)

Jun 4, 2023 – 10:35 AM EDT

PDB ID	:	2LF8
BMRB ID	:	17742
Title	:	Intramolecular regulation of the ETS Domain within ETV6 sequence R335 to R458
Authors	:	Coyne III, H.; Green, S.M.; Graves, B.J.; Mcintosh, L.P.
Deposited on	:	2011-06-28

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:

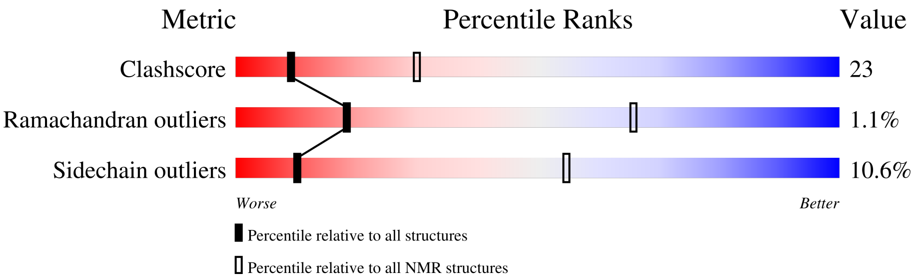
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal 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 89%.

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	128	42% 34% • 22%

2 Ensemble composition and analysis (i)

This entry contains 19 models. Model 3 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:336-A:382, A:386-A:438 (100)	0.38	3

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 3 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2173 atoms, of which 1076 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Transcription factor ETV6.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	128	2173	693	1076	200	199	5	0

There are 4 discrepancies between the modelled and reference sequences:

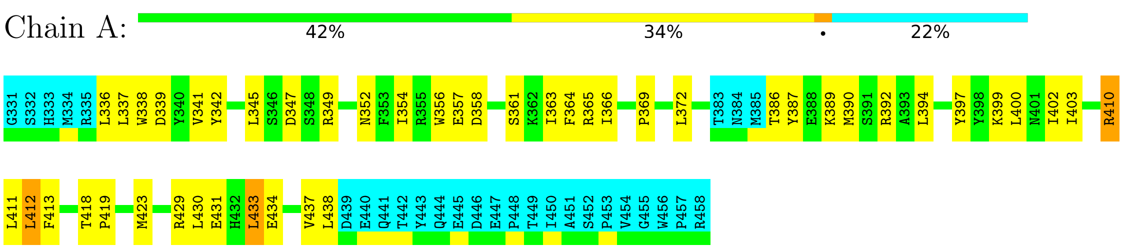
Chain	Residue	Modelled	Actual	Comment	Reference
A	331	GLY	-	expression tag	UNP P97360
A	332	SER	-	expression tag	UNP P97360
A	333	HIS	-	expression tag	UNP P97360
A	334	MET	-	expression tag	UNP P97360

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.

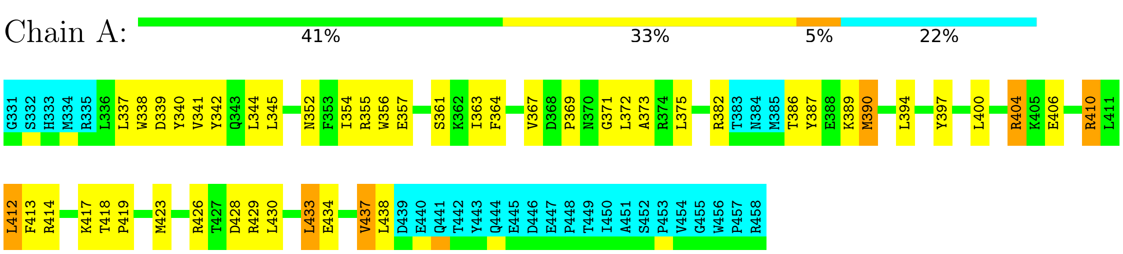
- Molecule 1: Transcription factor ETV6



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

The representative model is number 3. Colouring as in section 4.1 above.

- Molecule 1: Transcription factor ETV6



5 Refinement protocol and experimental data overview (i)

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

Of the 20 calculated structures, 19 were deposited, based on the following criterion: *structures with acceptable covalent geometry*.

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

Software name	Classification	Version
ARIA	structure solution	2.3
TALOS+	geometry optimization	
ARIA	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	1630
Number of shifts mapped to atoms	1630
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	89%

6 Model quality (i)

6.1 Standard geometry (i)

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

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

There are no bond-length outliers.

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	349	ARG	Sidechain	1

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	871	872	871	40±6
All	All	16549	16568	16549	759

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 23.

5 of 188 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:347:ASP:HB3	1:A:349:ARG:CD	0.83	2.04	13	2
1:A:358:ASP:HB2	1:A:363:ILE:HB	0.80	1.53	17	11

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:349:ARG:NE	1:A:350:TYR:H	0.78	1.76	13	2
1:A:434:GLU:O	1:A:438:LEU:HG	0.75	1.81	2	18
1:A:400:LEU:HD11	1:A:430:LEU:HD21	0.75	1.58	2	18

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	100/128 (78%)	94±1 (94±1%)	5±2 (5±2%)	1±1 (1±1%)	18	66
All	All	1900/2432 (78%)	1791 (94%)	88 (5%)	21 (1%)	18	66

All 2 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	410	ARG	17
1	A	425	GLY	4

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	94/119 (79%)	84±1 (89±1%)	10±1 (11±1%)	10	55
All	All	1786/2261 (79%)	1596 (89%)	190 (11%)	10	55

5 of 34 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	372	LEU	19
1	A	390	MET	19
1	A	410	ARG	19
1	A	412	LEU	19
1	A	433	LEU	19

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 89% for the well-defined parts and 86% 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 (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	1630
Number of shifts mapped to atoms	1630
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	4

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	122	-0.31 ± 0.09	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	117	0.23 ± 0.09	None needed (< 0.5 ppm)
$^{13}C'$	111	-0.22 ± 0.10	None needed (< 0.5 ppm)
^{15}N	115	0.41 ± 0.23	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 89%, i.e. 1367 atoms were assigned a chemical shift out of a possible 1530. 0 out of 16 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	483/498 (97%)	197/201 (98%)	192/200 (96%)	94/97 (97%)
Sidechain	744/878 (85%)	504/563 (90%)	227/264 (86%)	13/51 (25%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	140/154 (91%)	67/74 (91%)	64/71 (90%)	9/9 (100%)
Overall	1367/1530 (89%)	768/838 (92%)	483/535 (90%)	116/157 (74%)

7.1.4 Statistically unusual chemical shifts (i)

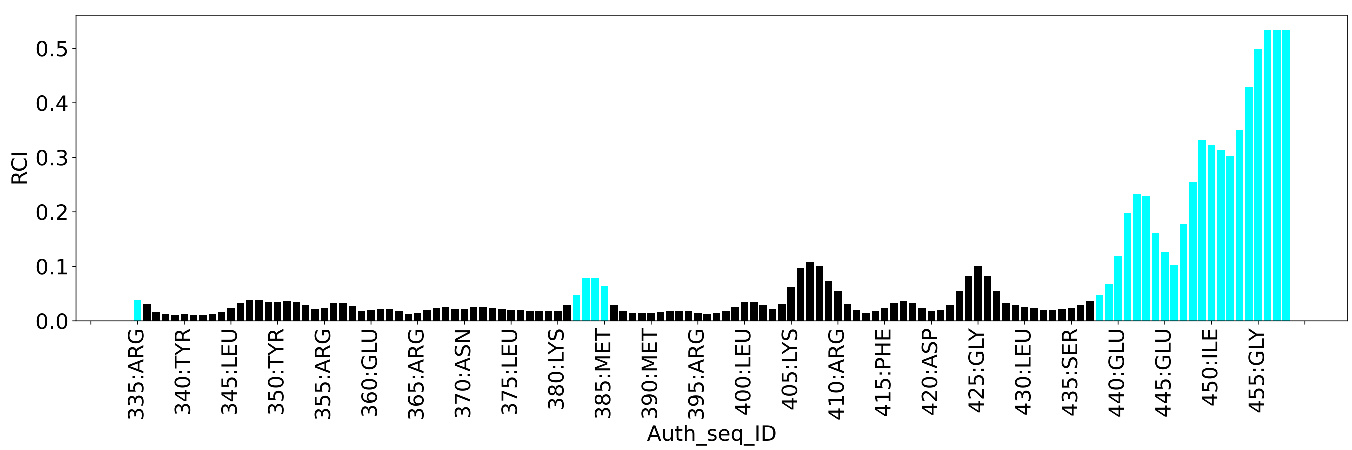
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	413	PHE	HB2	0.66	1.20 – 4.80	-6.5
1	A	345	LEU	HD11	-0.96	-0.61 – 2.12	-6.3
1	A	345	LEU	HD12	-0.96	-0.61 – 2.12	-6.3
1	A	345	LEU	HD13	-0.96	-0.61 – 2.12	-6.3

7.1.5 Random Coil Index (RCI) plots (i)

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

Random coil index (RCI) for chain A:



8 NMR restraints analysis (i)

8.1 Conformationally restricting restraints (i)

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

Description	Value
Total distance restraints	2992
Intra-residue (|i-j|=0)	1448
Sequential (|i-j|=1)	525
Medium range (|i-j|>1 and |i-j|<5)	388
Long range (|i-j|≥5)	631
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	190
Number of unmapped restraints	0
Number of restraints per residue	24.9
Number of long range restraints per residue[1]	4.9

[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 (i)

This section provides the overview of the restraint violations analysis. The violations are binned as small, medium and large violations based on its absolute value. Average number of violations per model is calculated by dividing the total number of violations in each bin by the size of the ensemble.

8.2.1 Average number of distance violations per model (i)

Distance violations less than 0.1 Å are not included in the calculation.

Bins (Å)	Average number of violations per model	Max (Å)
0.1-0.2 (Small)	93.6	0.2
0.2-0.5 (Medium)	87.7	0.5
>0.5 (Large)	152.1	7.03

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)	12.1	6.4
10.0-20.0 (Medium)	0.1	12.3
>20.0 (Large)	None	None

9 Distance violation analysis (i)

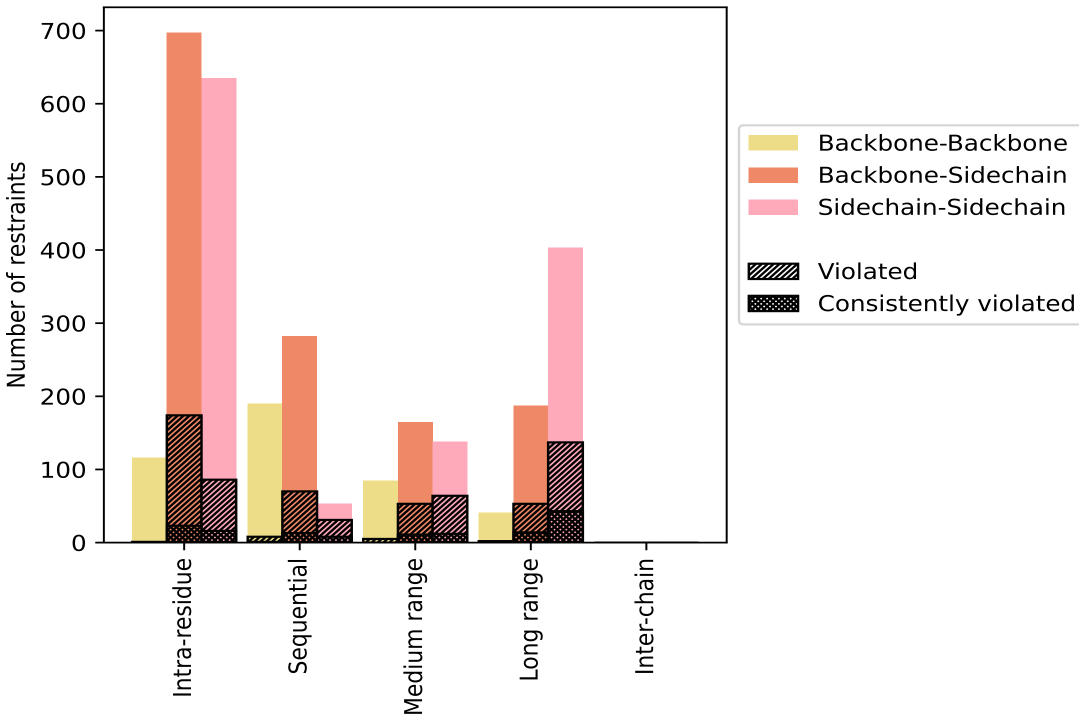
9.1 Summary of distance violations (i)

The following table shows the summary of distance violations in different restraint categories based on the sequence separation of the atoms involved. Each category is further sub-divided into three sub-categories based on the atoms involved. Violations less than 0.1 Å are not included in the statistics.

Restraints type	Count	%[1]	Violated[3]			Consistently Violated[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
Intra-residue (|i-j|=0)	1448	48.4	261	18.0	8.7	39	2.7	1.3
Backbone-Backbone	116	3.9	1	0.9	0.0	0	0.0	0.0
Backbone-Sidechain	697	23.3	174	25.0	5.8	23	3.3	0.8
Sidechain-Sidechain	635	21.2	86	13.5	2.9	16	2.5	0.5
Sequential (|i-j|=1)	525	17.5	109	20.8	3.6	22	4.2	0.7
Backbone-Backbone	190	6.4	8	4.2	0.3	1	0.5	0.0
Backbone-Sidechain	282	9.4	70	24.8	2.3	13	4.6	0.4
Sidechain-Sidechain	53	1.8	31	58.5	1.0	8	15.1	0.3
Medium range (|i-j|>1 & |i-j|<5)	388	13.0	122	31.4	4.1	23	5.9	0.8
Backbone-Backbone	85	2.8	5	5.9	0.2	0	0.0	0.0
Backbone-Sidechain	165	5.5	53	32.1	1.8	11	6.7	0.4
Sidechain-Sidechain	138	4.6	64	46.4	2.1	12	8.7	0.4
Long range (|i-j|≥5)	631	21.1	192	30.4	6.4	58	9.2	1.9
Backbone-Backbone	41	1.4	2	4.9	0.1	1	2.4	0.0
Backbone-Sidechain	187	6.2	53	28.3	1.8	14	7.5	0.5
Sidechain-Sidechain	403	13.5	137	34.0	4.6	43	10.7	1.4
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	2992	100.0	684	22.9	22.9	142	4.7	4.7
Backbone-Backbone	432	14.4	16	3.7	0.5	2	0.5	0.1
Backbone-Sidechain	1331	44.5	350	26.3	11.7	61	4.6	2.0
Sidechain-Sidechain	1229	41.1	318	25.9	10.6	79	6.4	2.6

[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 (i)



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

9.2 Distance violation statistics for each model (i)

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

Model ID	Number of violations						Mean (Å)	Max (Å)	SD[6] (Å)	Median (Å)
	IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total				
1	106	53	60	114	0	333	0.79	3.94	0.8	0.44
2	107	48	62	112	0	329	0.8	4.19	0.8	0.46
3	108	48	56	103	0	315	0.75	3.87	0.75	0.44
4	115	52	63	110	0	340	0.81	4.2	0.81	0.46
5	117	52	59	108	0	336	0.78	4.11	0.81	0.43
6	100	50	59	112	0	321	0.81	4.14	0.83	0.44
7	100	52	61	115	0	328	0.79	4.03	0.8	0.45
8	107	50	66	112	0	335	0.8	3.93	0.8	0.46
9	117	50	59	111	0	337	0.77	4.08	0.77	0.43
10	125	49	61	104	0	339	0.79	4.03	0.8	0.44
11	113	52	58	113	0	336	0.78	4.01	0.79	0.44

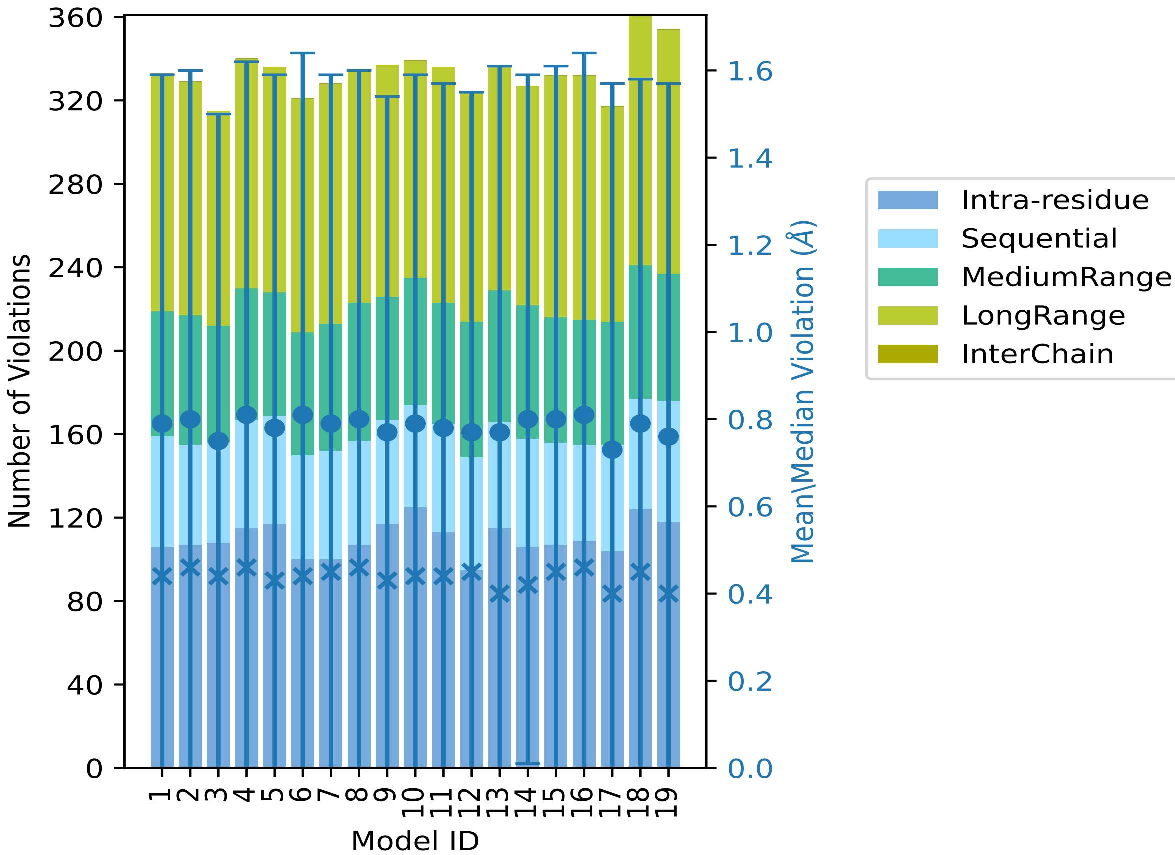
Continued on next page...

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	95	54	65	110	0	324	0.77	3.83	0.78	0.45
13	115	51	63	108	0	337	0.77	4.6	0.84	0.4
14	106	52	64	105	0	327	0.8	3.93	0.79	0.42
15	107	49	60	116	0	332	0.8	4.08	0.81	0.45
16	109	46	60	117	0	332	0.81	4.08	0.83	0.46
17	104	51	59	103	0	317	0.73	7.03	0.84	0.4
18	124	53	64	120	0	361	0.79	4.07	0.79	0.45
19	118	58	61	117	0	354	0.76	4.33	0.81	0.4

[1]Intra-residue restraints, [2]Sequential restraints, [3]Medium range restraints, [4]Long range restraints, [5]Inter-chain restraints, [6]Standard deviation

9.2.1 Bar graph : Distance Violation statistics for each model (i)



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

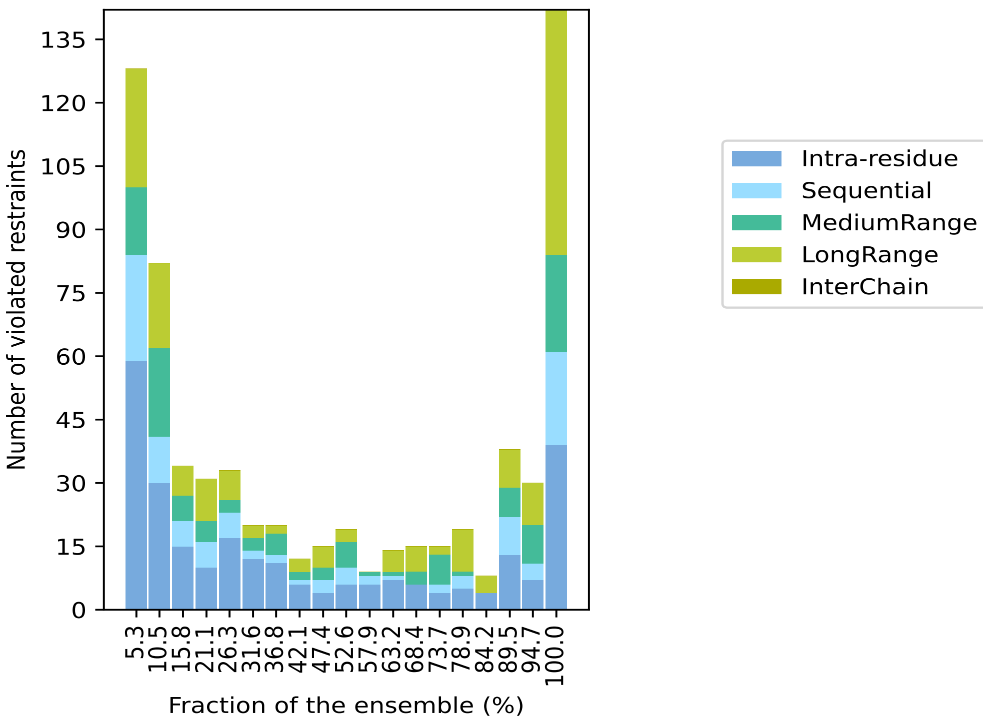
9.3 Distance violation statistics for the ensemble (i)

Violation analysis may find that some restraints are violated in few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of the ensemble. In total, 2308(IR:1187, SQ:416, MR:266, LR:439, 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]	%
59	25	16	28	0	128	1	5.3
30	11	21	20	0	82	2	10.5
15	6	6	7	0	34	3	15.8
10	6	5	10	0	31	4	21.1
17	6	3	7	0	33	5	26.3
12	2	3	3	0	20	6	31.6
11	2	5	2	0	20	7	36.8
6	1	2	3	0	12	8	42.1
4	3	3	5	0	15	9	47.4
6	4	6	3	0	19	10	52.6
6	2	1	0	0	9	11	57.9
7	1	1	5	0	14	12	63.2
6	0	3	6	0	15	13	68.4
4	2	7	2	0	15	14	73.7
5	3	1	10	0	19	15	78.9
4	0	0	4	0	8	16	84.2
13	9	7	9	0	38	17	89.5
7	4	9	10	0	30	18	94.7
39	22	23	58	0	142	19	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

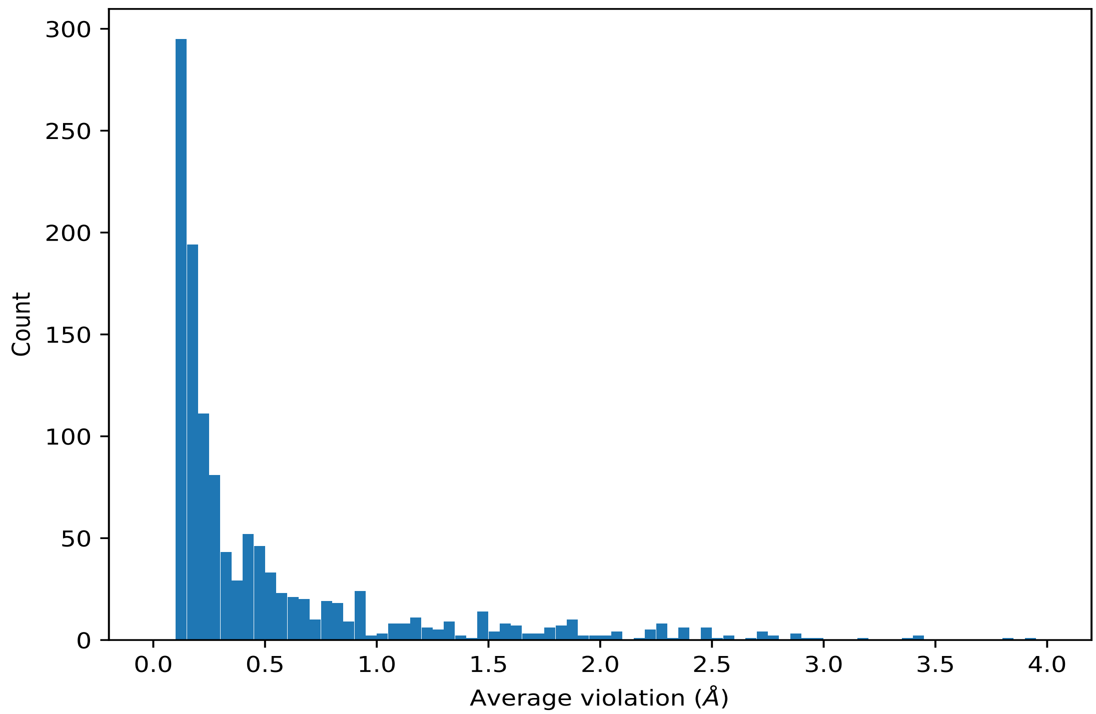
9.3.1 Bar graph : Distance violation statistics for the ensemble (i)



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.

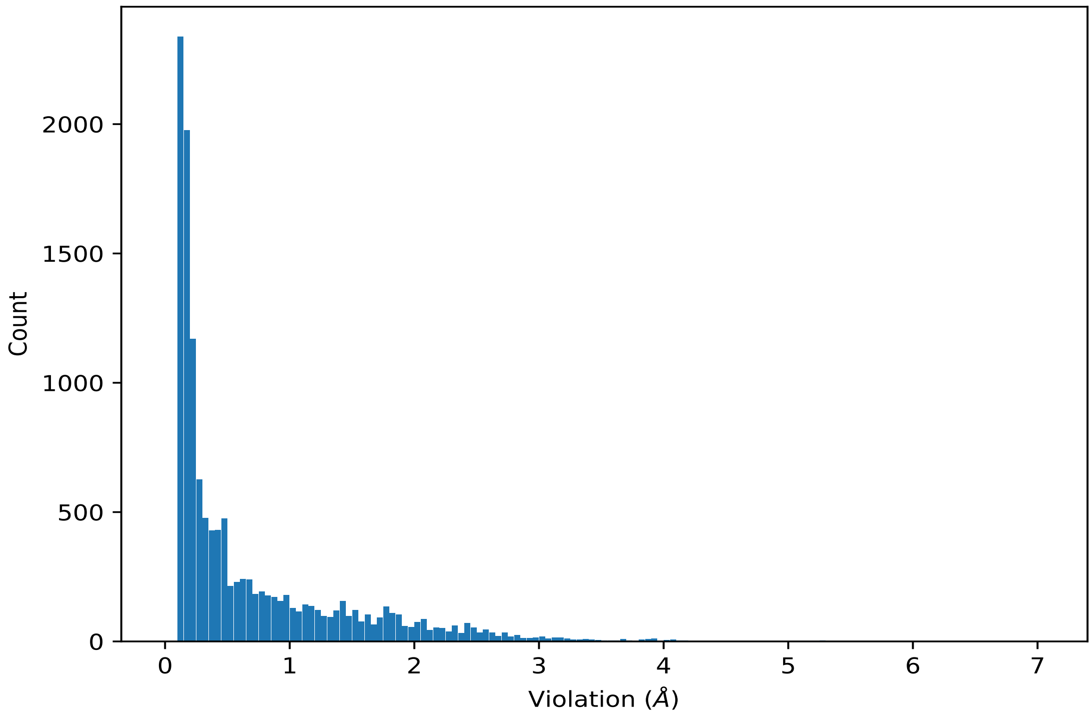
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,1069)	1:A:372:LEU:HG	1:A:387:TYR:HD1	19	3.91	0.08	3.92
(1,2761)	1:A:398:TYR:HD2	1:A:404:ARG:HB2	19	3.43	0.2	3.43
(1,2761)	1:A:398:TYR:HD2	1:A:404:ARG:HB3	19	3.43	0.2	3.43
(1,1387)	1:A:388:GLU:HA	1:A:387:TYR:HD2	19	3.16	0.1	3.16
(1,2098)	1:A:373:ALA:H	1:A:387:TYR:HD1	19	2.97	0.13	2.95
(1,1345)	1:A:398:TYR:HD1	1:A:395:ARG:HA	19	2.71	0.26	2.68
(1,1004)	1:A:403:ILE:HG22	1:A:398:TYR:HE2	19	2.7	0.33	2.8
(1,1004)	1:A:403:ILE:HG21	1:A:398:TYR:HE2	19	2.7	0.33	2.8
(1,1004)	1:A:403:ILE:HG23	1:A:398:TYR:HE2	19	2.7	0.33	2.8
(1,2209)	1:A:406:GLU:H	1:A:413:PHE:HD2	19	2.68	0.16	2.65
(1,636)	1:A:405:LYS:HA	1:A:413:PHE:HD2	19	2.57	0.15	2.52
(1,1809)	1:A:419:PRO:HG2	1:A:342:TYR:HD1	19	2.53	0.35	2.49
(1,1658)	1:A:395:ARG:HD2	1:A:398:TYR:HE1	19	2.49	0.43	2.37

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

9.5 All violated distance restraints (i)

9.5.1 Histogram : Distribution of distance violations (i)

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



9.5.2 Table : All distance violations (i)

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

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,2447)	1:A:385:MET:H	1:A:444:GLN:HE21	17	7.03
(1,888)	1:A:437:VAL:HG22	1:A:441:GLN:HE22	17	5.25
(1,888)	1:A:437:VAL:HG21	1:A:441:GLN:HE22	17	5.25
(1,888)	1:A:437:VAL:HG23	1:A:441:GLN:HE22	17	5.25
(1,1052)	1:A:347:ASP:HB2	1:A:350:TYR:HD1	13	4.6
(1,1057)	1:A:347:ASP:HB2	1:A:350:TYR:HE1	13	4.46
(1,1052)	1:A:347:ASP:HB2	1:A:350:TYR:HD1	19	4.33
(1,52)	1:A:443:TYR:HE2	1:A:456:TRP:HD1	4	4.2
(1,52)	1:A:443:TYR:HE2	1:A:456:TRP:HD1	2	4.19
(1,1057)	1:A:347:ASP:HB2	1:A:350:TYR:HE1	19	4.18
(1,646)	1:A:405:LYS:HG2	1:A:413:PHE:HE2	6	4.14

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,1069)	1:A:372:LEU:HG	1:A:387:TYR:HD1	5	4.11

10 Dihedral-angle violation analysis (i)

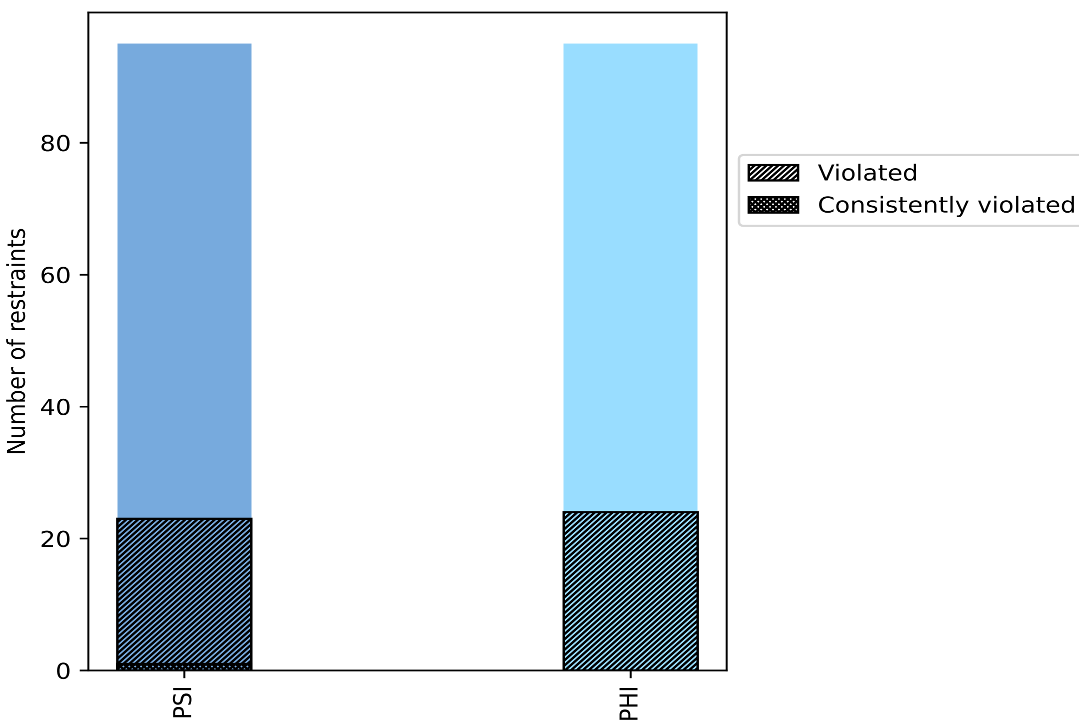
10.1 Summary of dihedral-angle violations (i)

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	95	50.0	23	24.2	12.1	1	1.1	0.5
PHI	95	50.0	24	25.3	12.6	0	0.0	0.0
Total	190	100.0	47	24.7	24.7	1	0.5	0.5

[1] percentage calculated with respect to total number of dihedral-angle restraints, [2] percentage calculated with respect to number of restraints in a particular dihedral-angle type, [3] violated in at least one model, [4] violated in all the models

10.1.1 Bar chart : Distribution of dihedral-angles and violations (i)



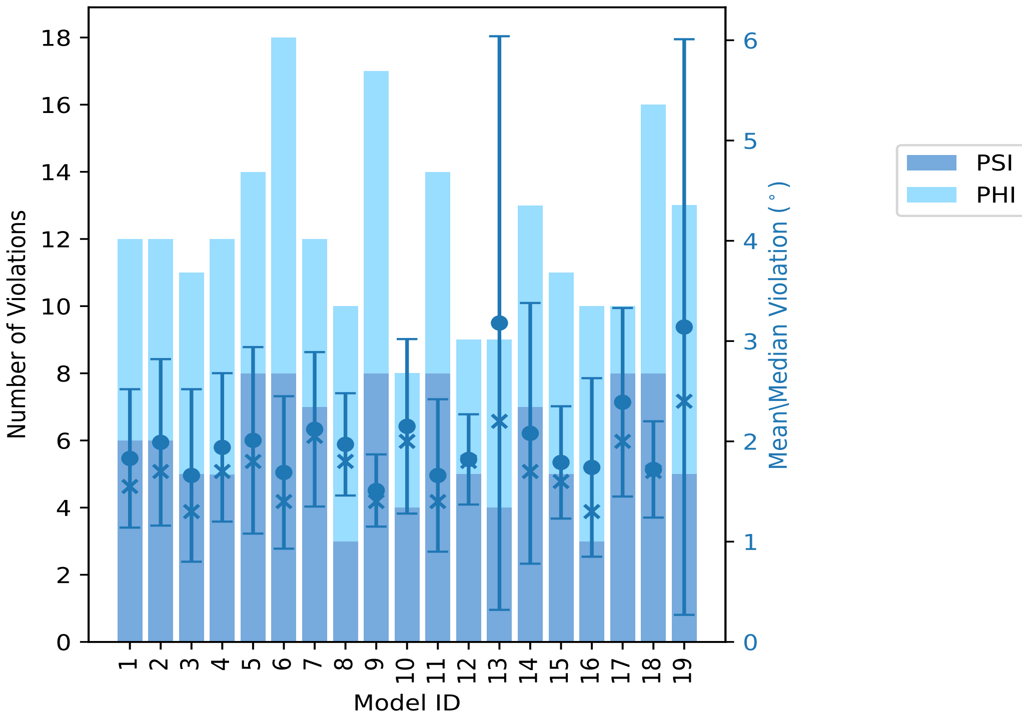
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

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	6	6	12	1.83	3.0	0.69	1.55
2	6	6	12	1.99	4.1	0.83	1.7
3	5	6	11	1.66	4.3	0.86	1.3
4	5	7	12	1.94	3.2	0.74	1.7
5	8	6	14	2.01	4.7	0.93	1.8
6	8	10	18	1.69	4.1	0.76	1.4
7	7	5	12	2.12	3.6	0.77	2.05
8	3	7	10	1.97	3.1	0.51	1.8
9	8	9	17	1.51	2.1	0.36	1.4
10	4	4	8	2.15	4.0	0.87	2.0
11	8	6	14	1.66	4.0	0.76	1.4
12	5	4	9	1.82	2.9	0.45	1.8
13	4	5	9	3.18	11.1	2.86	2.2
14	7	6	13	2.08	6.4	1.3	1.7
15	5	6	11	1.79	3.1	0.56	1.6
16	3	7	10	1.74	4.1	0.89	1.3
17	8	2	10	2.39	4.1	0.94	2.0
18	8	8	16	1.72	2.5	0.48	1.7
19	5	8	13	3.14	12.3	2.87	2.4

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]	%
5	8	13	1	5.3
6	4	10	2	10.5
2	1	3	3	15.8
0	1	1	4	21.1
4	3	7	5	26.3
1	1	2	6	31.6
1	0	1	7	36.8
1	1	2	8	42.1
0	1	1	9	47.4
0	0	0	10	52.6
0	0	0	11	57.9

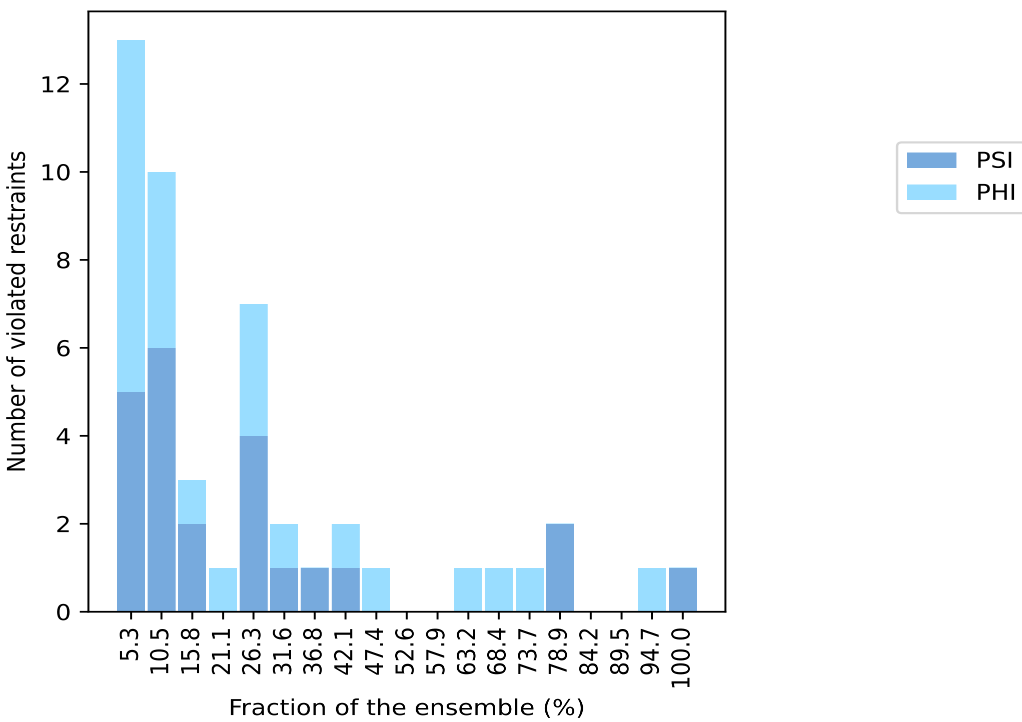
Continued on next page...

Continued from previous page...

Number of violated restraints			Fraction of the ensemble	
PSI	PHI	Total	Count[1]	%
0	1	1	12	63.2
0	1	1	13	68.4
0	1	1	14	73.7
2	0	2	15	78.9
0	0	0	16	84.2
0	0	0	17	89.5
0	1	1	18	94.7
1	0	1	19	100.0

[1] Number of models with violations

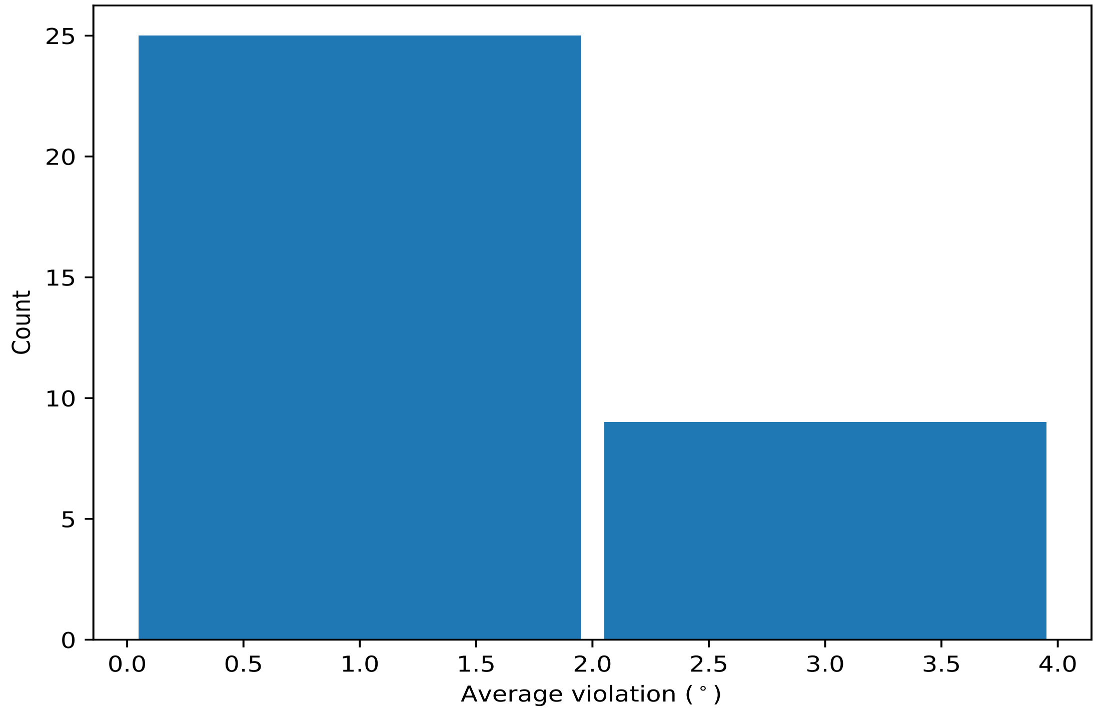
10.3.1 Bar graph : Dihedral-angle Violation statistics for the ensemble (i)



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 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.

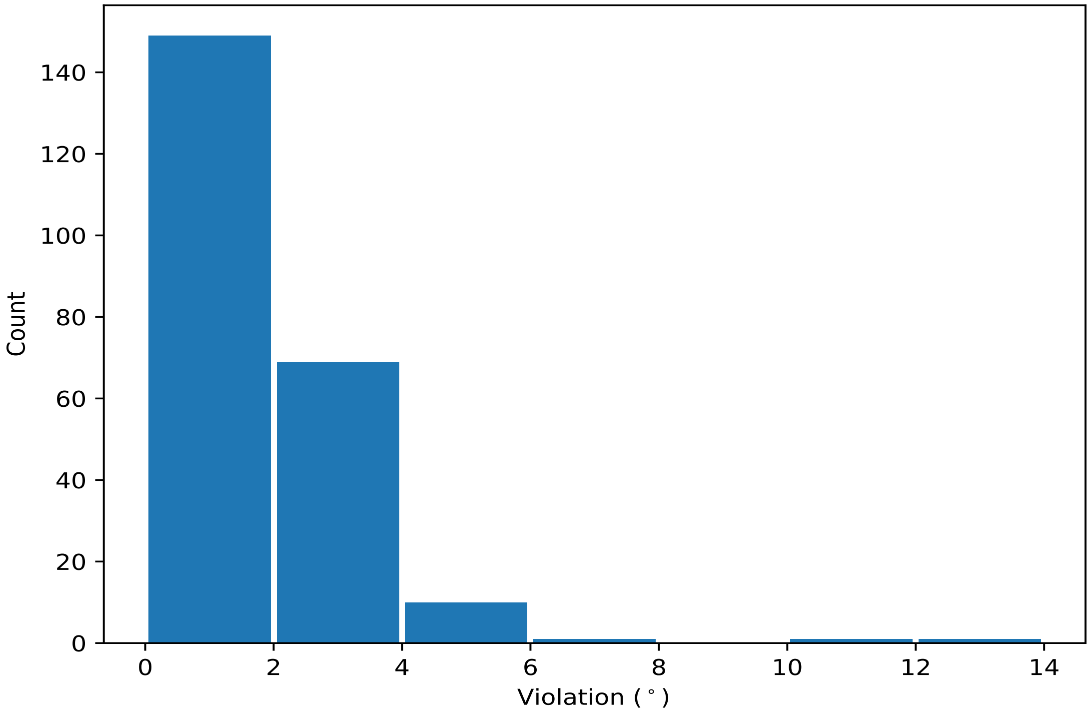
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models[1]	Mean	SD[2]	Median
(1,164)	1:A:411:LEU:N	1:A:411:LEU:CA	1:A:411:LEU:C	1:A:412:LEU:N	19	3.45	0.83	3.1
(1,17)	1:A:353:PHE:C	1:A:354:ILE:N	1:A:354:ILE:CA	1:A:354:ILE:C	18	1.8	0.41	1.75
(1,97)	1:A:338:TRP:N	1:A:338:TRP:CA	1:A:338:TRP:C	1:A:339:ASP:N	15	1.81	0.39	1.7
(1,145)	1:A:391:SER:N	1:A:391:SER:CA	1:A:391:SER:C	1:A:392:ARG:N	15	1.78	0.39	1.8
(1,90)	1:A:435:SER:C	1:A:436:GLN:N	1:A:436:GLN:CA	1:A:436:GLN:C	14	2.14	0.76	2.0
(1,59)	1:A:399:LYS:C	1:A:400:LEU:N	1:A:400:LEU:CA	1:A:400:LEU:C	13	1.93	0.57	1.8
(1,18)	1:A:354:ILE:C	1:A:355:ARG:N	1:A:355:ARG:CA	1:A:355:ARG:C	12	1.58	0.36	1.5
(1,48)	1:A:388:GLU:C	1:A:389:LYS:N	1:A:389:LYS:CA	1:A:389:LYS:C	9	1.87	0.66	1.6
(1,11)	1:A:347:ASP:C	1:A:348:SER:N	1:A:348:SER:CA	1:A:348:SER:C	8	3.95	4.49	1.4
(1,161)	1:A:407:PRO:N	1:A:407:PRO:CA	1:A:407:PRO:C	1:A:408:GLY:N	8	2.24	1.6	1.75

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

10.5 All violated dihedral-angle restraints (i)

10.5.1 Histogram : Distribution of violations (i)

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



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

The following table provides the list of 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 (°)
(1,11)	1:A:347:ASP:C	1:A:348:SER:N	1:A:348:SER:CA	1:A:348:SER:C	19	12.3
(1,11)	1:A:347:ASP:C	1:A:348:SER:N	1:A:348:SER:CA	1:A:348:SER:C	13	11.1
(1,161)	1:A:407:PRO:N	1:A:407:PRO:CA	1:A:407:PRO:C	1:A:408:GLY:N	14	6.4
(1,164)	1:A:411:LEU:N	1:A:411:LEU:CA	1:A:411:LEU:C	1:A:412:LEU:N	19	5.5
(1,81)	1:A:426:ARG:C	1:A:427:THR:N	1:A:427:THR:CA	1:A:427:THR:C	5	4.7
(1,164)	1:A:411:LEU:N	1:A:411:LEU:CA	1:A:411:LEU:C	1:A:412:LEU:N	3	4.3
(1,164)	1:A:411:LEU:N	1:A:411:LEU:CA	1:A:411:LEU:C	1:A:412:LEU:N	2	4.1
(1,164)	1:A:411:LEU:N	1:A:411:LEU:CA	1:A:411:LEU:C	1:A:412:LEU:N	6	4.1
(1,164)	1:A:411:LEU:N	1:A:411:LEU:CA	1:A:411:LEU:C	1:A:412:LEU:N	16	4.1
(1,108)	1:A:350:TYR:N	1:A:350:TYR:CA	1:A:350:TYR:C	1:A:351:GLU:N	17	4.1