wwPDB NMR Structure Validation Summary Report

Jun 6, 2023 – 06:14 pm BST

PDB ID	:	3ZG4
BMRB ID	:	18900
Title	:	NMR structure of the catalytic domain from E. faecium L,D- transpeptidase
Authors	:	Lecoq, L.; Dubee, V.; Triboulet, S.; Bougault, C.; Hugonnet, J.E.; Arthur, M.; Simorre, J.P.
Deposited on	:	2012-12-14

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:

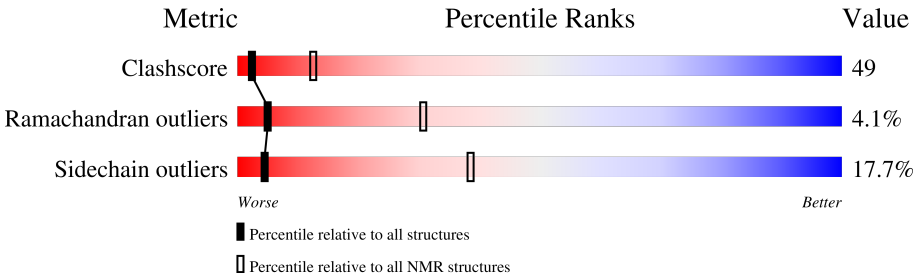
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal 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 96%.

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	129	28% 49% 11% • 11%

2 Ensemble composition and analysis

This entry contains 20 models. Model 17 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative.

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:343-A:394, A:404-A:466 (115)	0.18	17

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1983 atoms, of which 956 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called ERFK/YBIS/YCFS/YNHG.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	129	1983	657	956	161	203	6	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	338	GLY	-	expression tag	UNP Q3Y185
A	339	HIS	-	expression tag	UNP Q3Y185
A	340	MET	-	expression tag	UNP Q3Y185

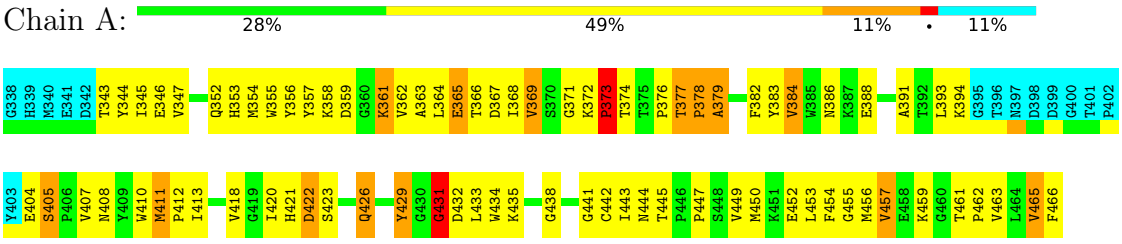
4 Residue-property plots

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: ERFK/YBIS/YCFS/YNHG

Chain A:

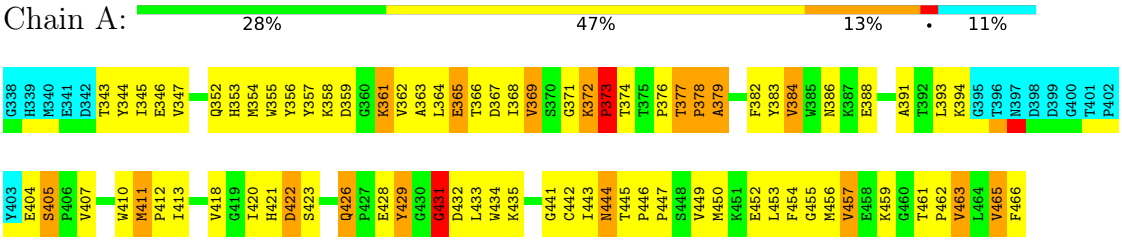


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 17. Colouring as in section 4.1 above.

- Molecule 1: ERFK/YBIS/YCFS/YNHG

Chain A:



5 Refinement protocol and experimental data overview

The models were refined using the following method: *ARIA2.3*.

Of the 1700 calculated structures, 20 were deposited, based on the following criterion: *TOTAL ENERGY*.

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

Software name	Classification	Version
CNS	refinement	
UNIO	structure solution	10
TALOS	structure solution	
CcpNmr Analysis	structure solution	ANALYSIS

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	1597
Number of shifts mapped to atoms	1597
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	96%

6 Model quality

6.1 Standard geometry

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	0.90±0.03	1±1/955 (0.1± 0.1%)	0.82±0.01	0±0/1305 (0.0± 0.0%)
All	All	0.90	11/19100 (0.1%)	0.82	0/26100 (0.0%)

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

Mol	Chain	Chirality	Planarity
1	A	0.0±0.0	0.9±0.5
All	All	0	19

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	409	TYR	CE2-CZ	5.99	1.46	1.38	15	3
1	A	409	TYR	CE1-CZ	-5.88	1.30	1.38	15	3
1	A	357	TYR	CE1-CZ	-5.55	1.31	1.38	20	3
1	A	357	TYR	CE2-CZ	5.15	1.45	1.38	20	1
1	A	382	PHE	CE1-CZ	5.02	1.46	1.37	4	1

There are no bond-angle outliers.

There are no chirality outliers.

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

Mol	Chain	Res	Type	Group	Models (Total)
1	A	431	GLY	Peptide	16
1	A	356	TYR	Sidechain	3

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	923	874	872	88±4
All	All	18460	17480	17440	1767

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 49.

5 of 201 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:352:GLN:OE1	1:A:369:VAL:HA	0.79	1.76	6	12
1:A:411:MET:SD	1:A:457:VAL:HG22	0.73	2.22	6	1
1:A:352:GLN:NE2	1:A:368:ILE:HD12	0.71	2.00	1	1
1:A:373:PRO:CD	1:A:435:LYS:HA	0.71	2.15	13	20
1:A:407:VAL:HG21	1:A:421:HIS:HB2	0.71	1.60	5	17

6.3 Torsion angles

6.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the backbone conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	114/129 (88%)	99±1 (87±1%)	10±1 (9±1%)	5±1 (4±1%)	5	31
All	All	2280/2580 (88%)	1984 (87%)	203 (9%)	93 (4%)	5	31

All 5 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	379	ALA	20
1	A	431	GLY	20
1	A	373	PRO	18

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	378	PRO	18
1	A	432	ASP	17

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	101/112 (90%)	83±2 (82±2%)	18±2 (18±2%)	4 38
All	All	2020/2240 (90%)	1663 (82%)	357 (18%)	4 38

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	369	VAL	20
1	A	373	PRO	20
1	A	377	THR	20
1	A	384	VAL	20
1	A	420	ILE	20

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

The completeness of assignment taking into account all chemical shift lists is 96% for the well-defined parts and 95% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chemical_shifts_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	1597
Number of shifts mapped to atoms	1597
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	12

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	128	-0.07 ± 0.11	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	115	-0.02 ± 0.24	None needed (< 0.5 ppm)
$^{13}C'$	116	0.26 ± 0.14	None needed (< 0.5 ppm)
^{15}N	116	0.38 ± 0.38	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 96%, i.e. 1463 atoms were assigned a chemical shift out of a possible 1530. 0 out of 17 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	557/568 (98%)	231/232 (100%)	220/230 (96%)	106/106 (100%)
Sidechain	740/783 (95%)	499/510 (98%)	235/256 (92%)	6/17 (35%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	166/179 (93%)	82/87 (94%)	73/81 (90%)	11/11 (100%)
Overall	1463/1530 (96%)	812/829 (98%)	528/567 (93%)	123/134 (92%)

7.1.4 Statistically unusual chemical shifts (i)

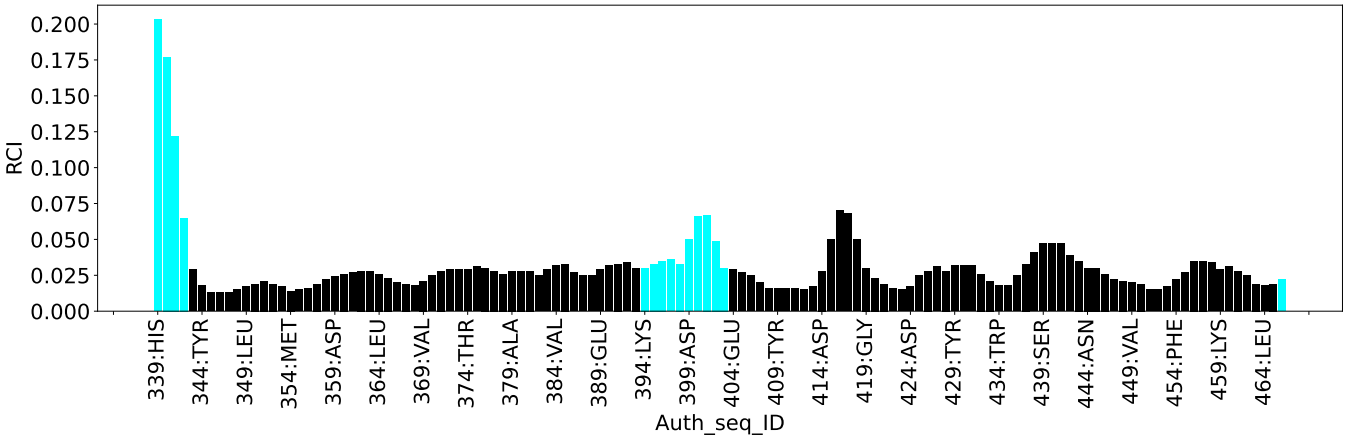
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	366	THR	HG1	5.51	0.08 – 2.19	20.7
1	A	412	PRO	HD3	1.00	1.76 – 5.48	-7.0
1	A	446	PRO	HG2	-0.10	0.41 – 3.45	-6.7
1	A	435	LYS	HG3	-0.33	0.04 – 2.67	-6.4
1	A	435	LYS	HG2	-0.18	0.13 – 2.61	-6.2
1	A	423	SER	HB3	2.20	2.49 – 5.20	-6.1
1	A	423	SER	HB2	2.43	2.61 – 5.13	-5.7
1	A	353	HIS	HD2	4.38	4.65 – 9.35	-5.6
1	A	412	PRO	HG2	0.24	0.41 – 3.45	-5.6
1	A	393	LEU	HD11	-0.69	-0.61 – 2.12	-5.3
1	A	393	LEU	HD12	-0.69	-0.61 – 2.12	-5.3
1	A	393	LEU	HD13	-0.69	-0.61 – 2.12	-5.3

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:


RCI
0.200
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
339:HIS
344:TYR
349:LEU
354:MET
359:ASP
364:LEU
369:VAL
374:THR
379:ALA
384:VAL
389:GLU
394:LYS
399:ASP
404:GLU
409:TYR
414:ASP
419:GLY
424:ASP
429:TYR
434:TRP
439:SER
444:ASN
449:VAL
454:PHE
459:LYS
464:LEU
Auth_seq_ID

8 NMR restraints analysis

8.1 Conformationally restricting restraints

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

Description	Value
Total distance restraints	5008
Intra-residue (|i-j|=0)	2735
Sequential (|i-j|=1)	1534
Medium range (|i-j|>1 and |i-j|<5)	209
Long range (|i-j|≥5)	530
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	220
Number of unmapped restraints	0
Number of restraints per residue	40.5
Number of long range restraints per residue[1]	4.1

[1]Long range hydrogen bonds and disulfide bonds are counted as long range restraints while calculating the number of long range restraints per residue

8.2 Residual restraint violations

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

8.2.1 Average number of distance violations per model

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

Bins (Å)	Average number of violations per model	Max (Å)
0.1-0.2 (Small)	304.6	0.2
0.2-0.5 (Medium)	788.1	0.5
>0.5 (Large)	1206.5	23.83

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)	34.1	10.0
10.0-20.0 (Medium)	3.0	16.6
>20.0 (Large)	None	None

9 Distance violation analysis (i)

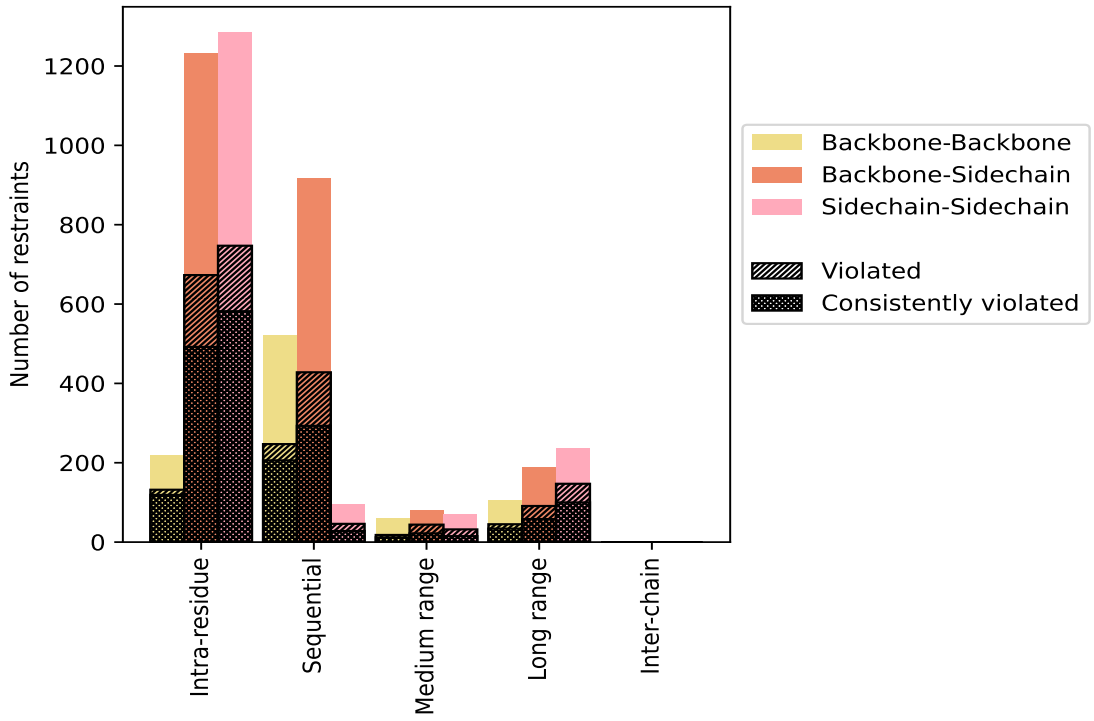
9.1 Summary of distance violations (i)

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)	2735	54.6	1552	56.7	31.0	1193	43.6	23.8
Backbone-Backbone	218	4.4	132	60.6	2.6	119	54.6	2.4
Backbone-Sidechain	1232	24.6	673	54.6	13.4	492	39.9	9.8
Sidechain-Sidechain	1285	25.7	747	58.1	14.9	582	45.3	11.6
Sequential (|i-j|=1)	1534	30.6	721	47.0	14.4	527	34.4	10.5
Backbone-Backbone	521	10.4	247	47.4	4.9	206	39.5	4.1
Backbone-Sidechain	917	18.3	428	46.7	8.5	293	32.0	5.9
Sidechain-Sidechain	96	1.9	46	47.9	0.9	28	29.2	0.6
Medium range (|i-j|>1 & |i-j|<5)	209	4.2	94	45.0	1.9	49	23.4	1.0
Backbone-Backbone	59	1.2	18	30.5	0.4	12	20.3	0.2
Backbone-Sidechain	81	1.6	44	54.3	0.9	22	27.2	0.4
Sidechain-Sidechain	69	1.4	32	46.4	0.6	15	21.7	0.3
Long range (|i-j|≥5)	530	10.6	283	53.4	5.7	192	36.2	3.8
Backbone-Backbone	105	2.1	45	42.9	0.9	34	32.4	0.7
Backbone-Sidechain	189	3.8	91	48.1	1.8	58	30.7	1.2
Sidechain-Sidechain	236	4.7	147	62.3	2.9	100	42.4	2.0
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	5008	100.0	2650	52.9	52.9	1961	39.2	39.2
Backbone-Backbone	903	18.0	442	48.9	8.8	371	41.1	7.4
Backbone-Sidechain	2419	48.3	1236	51.1	24.7	865	35.8	17.3
Sidechain-Sidechain	1686	33.7	972	57.7	19.4	725	43.0	14.5

[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	1360	619	76	235	0	2290	0.72	23.71	1.06	0.53
2	1375	617	73	242	0	2307	0.71	23.71	1.05	0.53
3	1374	613	72	236	0	2295	0.72	23.53	1.05	0.53
4	1360	628	76	234	0	2298	0.72	23.68	1.06	0.53
5	1386	616	78	239	0	2319	0.72	23.46	1.05	0.53
6	1360	623	75	237	0	2295	0.72	23.68	1.05	0.54
7	1365	610	73	233	0	2281	0.72	23.67	1.06	0.54
8	1380	609	75	236	0	2300	0.72	23.64	1.06	0.53
9	1367	618	74	234	0	2293	0.71	23.77	1.06	0.53
10	1368	623	74	241	0	2306	0.71	23.73	1.06	0.53
11	1370	611	71	238	0	2290	0.72	23.64	1.06	0.53

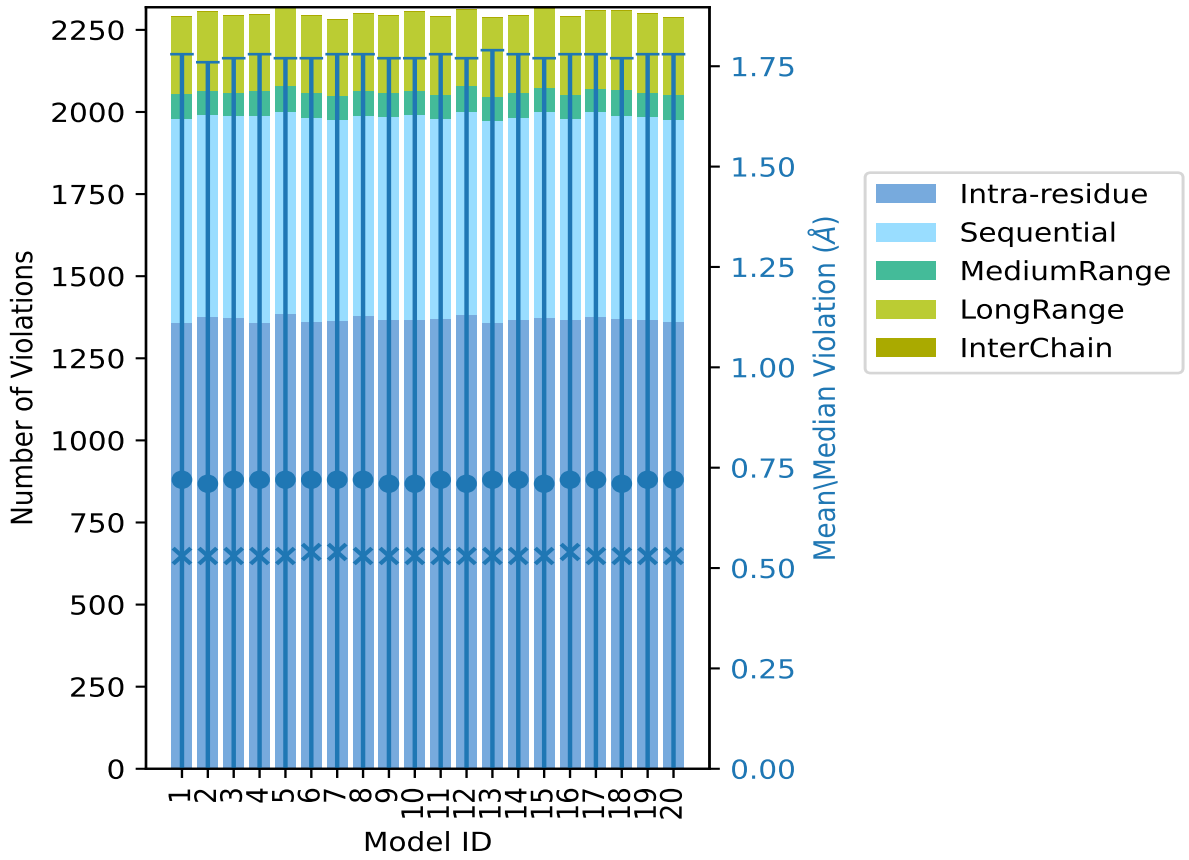
Continued on next page...

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	1381	620	77	234	0	2312	0.71	23.7	1.06	0.53
13	1360	613	74	241	0	2288	0.72	23.83	1.07	0.53
14	1368	615	76	235	0	2294	0.72	23.69	1.06	0.53
15	1374	625	75	242	0	2316	0.71	23.79	1.06	0.53
16	1368	610	75	238	0	2291	0.72	23.64	1.06	0.54
17	1376	624	71	238	0	2309	0.72	23.65	1.06	0.53
18	1369	621	77	243	0	2310	0.71	23.67	1.06	0.53
19	1366	620	74	241	0	2301	0.72	23.82	1.06	0.53
20	1362	614	76	236	0	2288	0.72	23.75	1.06	0.53

[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

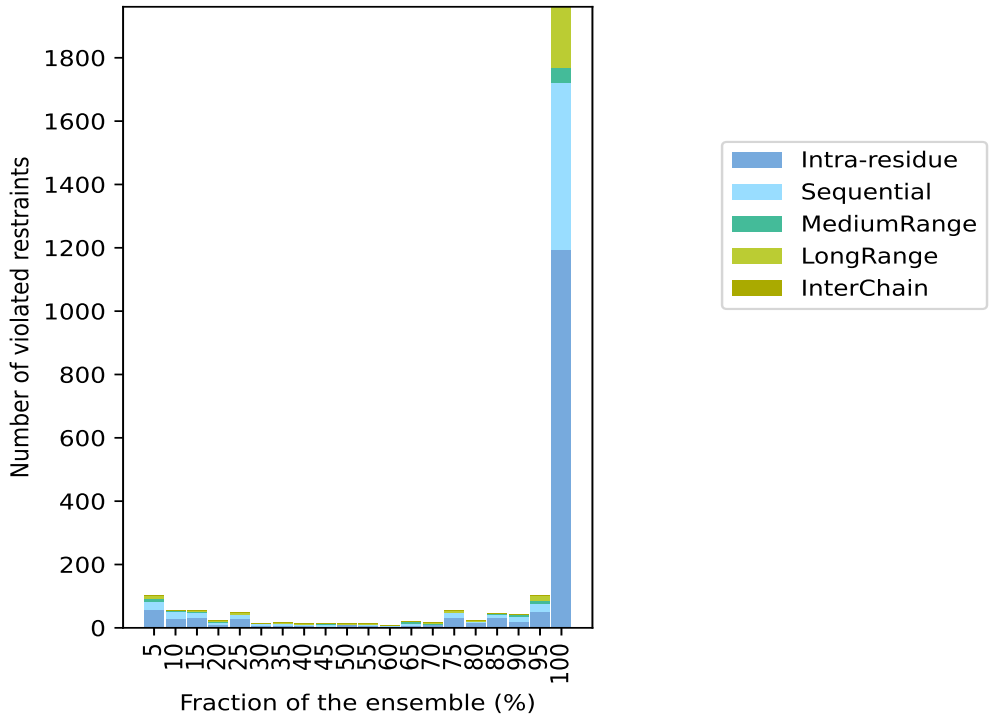
9.3 Distance violation statistics for the ensemble (i)

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, 2358(IR:1183, SQ:813, MR:115, LR:247, 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]	%
57	24	9	12	0	102	1	5.0
29	21	2	3	0	55	2	10.0
30	18	2	6	0	56	3	15.0
11	7	0	6	0	24	4	20.0
27	15	1	7	0	50	5	25.0
5	8	0	1	0	14	6	30.0
7	5	1	4	0	17	7	35.0
6	5	0	2	0	13	8	40.0
3	6	3	3	0	15	9	45.0
9	2	0	2	0	13	10	50.0
7	1	2	5	0	15	11	55.0
4	2	1	1	0	8	12	60.0
7	7	4	2	0	20	13	65.0
9	3	0	5	0	17	14	70.0
31	16	0	8	0	55	15	75.0
16	2	1	4	0	23	16	80.0
31	11	2	2	0	46	17	85.0
19	17	5	2	0	43	18	90.0
51	24	12	16	0	103	19	95.0
1193	527	49	192	0	1961	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

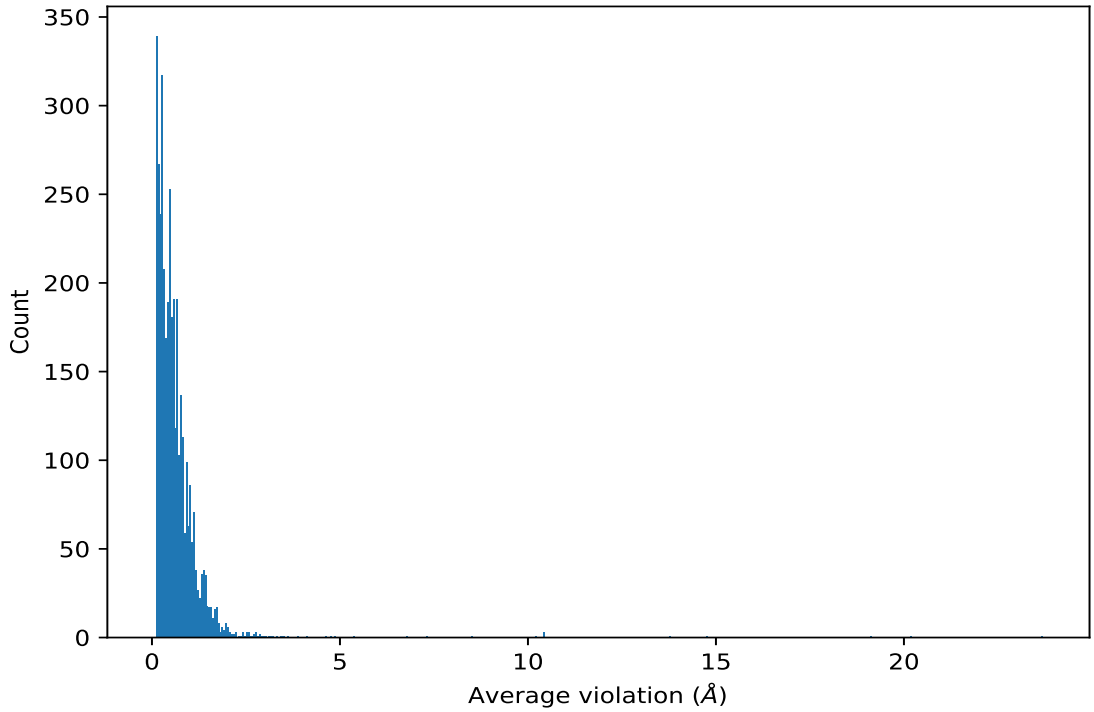
9.3.1 Bar graph : Distance violation statistics for the ensemble



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 (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.

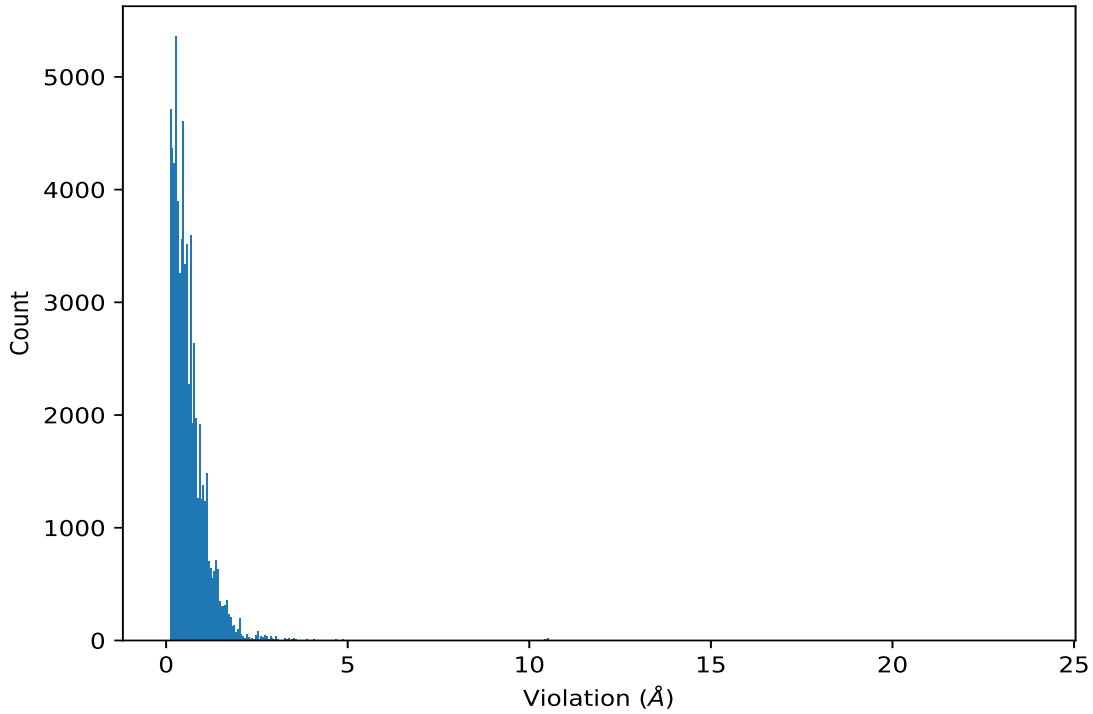
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,4925)	1:A:391:ALA:H	1:A:361:LYS:HA	20	23.69	0.09	23.68
(1,4064)	1:A:388:GLU:H	1:A:361:LYS:HA	20	20.17	0.07	20.18
(1,4147)	1:A:410:TRP:H	1:A:361:LYS:HA	20	19.12	0.09	19.12
(1,4371)	1:A:425:TRP:HE1	1:A:356:TYR:HA	20	14.77	0.12	14.82
(1,4424)	1:A:423:SER:H	1:A:458:GLU:HG2	20	13.78	0.26	13.85
(1,4978)	1:A:461:THR:H	1:A:362:VAL:HG21	20	10.43	0.19	10.47
(1,4978)	1:A:461:THR:H	1:A:362:VAL:HG22	20	10.43	0.19	10.47
(1,4978)	1:A:461:THR:H	1:A:362:VAL:HG23	20	10.43	0.19	10.47
(1,3937)	1:A:460:GLY:H	1:A:443:ILE:HG12	20	10.21	0.11	10.22
(1,4936)	1:A:425:TRP:HE1	1:A:419:GLY:HA2	20	8.51	0.14	8.57
(1,4835)	1:A:425:TRP:HE1	1:A:397:ASN:HD21	20	7.35	0.36	7.36
(1,4216)	1:A:360:GLY:H	1:A:464:LEU:HB3	20	6.76	0.16	6.78

[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)
(1,4925)	1:A:391:ALA:H	1:A:361:LYS:HA	13	23.83
(1,4925)	1:A:391:ALA:H	1:A:361:LYS:HA	19	23.82
(1,4925)	1:A:391:ALA:H	1:A:361:LYS:HA	15	23.79
(1,4925)	1:A:391:ALA:H	1:A:361:LYS:HA	9	23.77
(1,4925)	1:A:391:ALA:H	1:A:361:LYS:HA	20	23.75
(1,4925)	1:A:391:ALA:H	1:A:361:LYS:HA	10	23.73
(1,4925)	1:A:391:ALA:H	1:A:361:LYS:HA	1	23.71
(1,4925)	1:A:391:ALA:H	1:A:361:LYS:HA	2	23.71
(1,4925)	1:A:391:ALA:H	1:A:361:LYS:HA	12	23.7
(1,4925)	1:A:391:ALA:H	1:A:361:LYS:HA	14	23.69

10 Dihedral-angle violation analysis

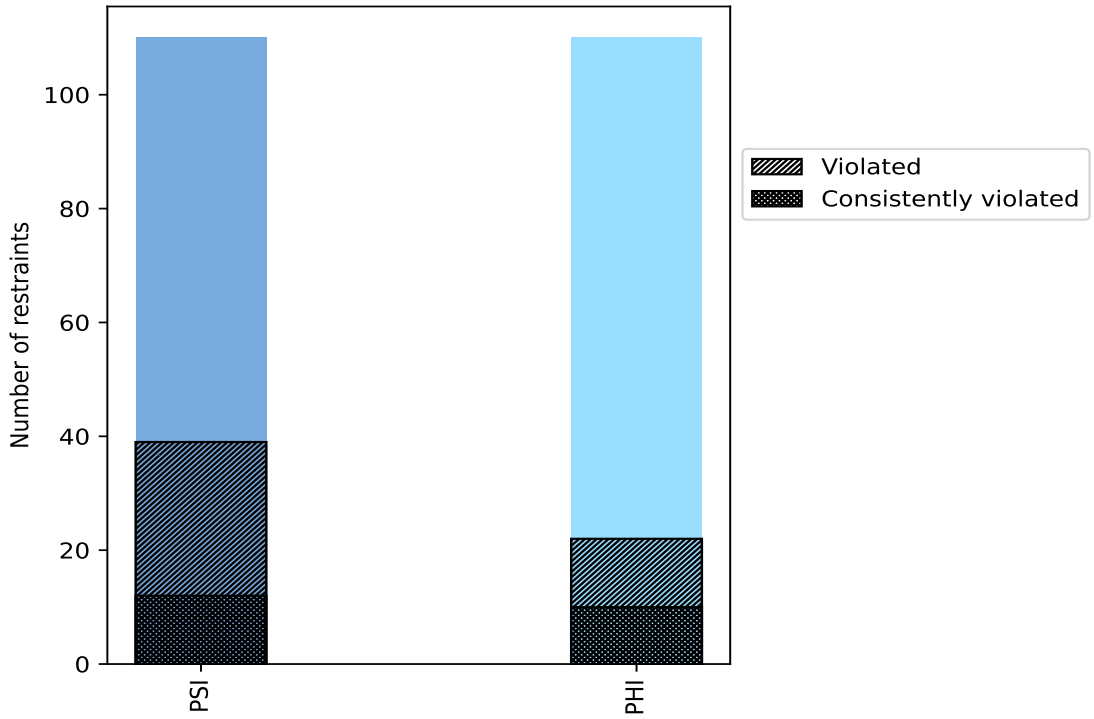
10.1 Summary of dihedral-angle violations

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

Angle type	Count	%[1]	Violated[3]			Consistently Violated[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
PSI	110	50.0	39	35.5	17.7	12	10.9	5.5
PHI	110	50.0	22	20.0	10.0	10	9.1	4.5
Total	220	100.0	61	27.7	27.7	22	10.0	10.0

[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



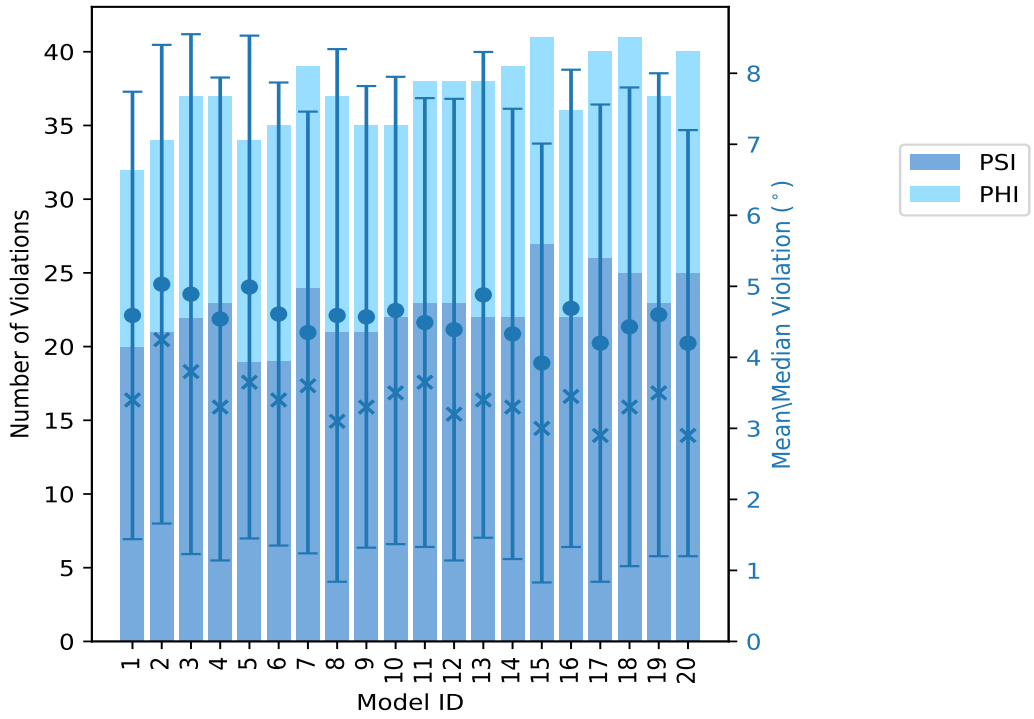
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

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	20	12	32	4.59	13.6	3.15	3.4
2	21	13	34	5.03	14.8	3.37	4.25
3	22	15	37	4.89	15.6	3.66	3.8
4	23	14	37	4.54	14.2	3.4	3.3
5	19	15	34	4.99	15.0	3.54	3.65
6	19	16	35	4.61	15.9	3.26	3.4
7	24	15	39	4.35	13.8	3.11	3.6
8	21	16	37	4.59	15.1	3.75	3.1
9	21	14	35	4.57	14.4	3.25	3.3
10	22	13	35	4.66	14.9	3.29	3.5
11	23	15	38	4.49	16.6	3.16	3.65
12	23	15	38	4.39	14.9	3.25	3.2
13	22	16	38	4.88	15.5	3.42	3.4
14	22	17	39	4.33	13.5	3.17	3.3
15	27	14	41	3.92	15.1	3.09	3.0
16	22	14	36	4.69	13.8	3.36	3.45
17	26	14	40	4.2	15.7	3.36	2.9
18	25	16	41	4.43	14.1	3.37	3.3
19	23	14	37	4.6	15.0	3.4	3.5
20	25	15	40	4.2	13.4	3.0	2.9

10.2.1 Bar graph : Dihedral violation statistics for each model



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

10.3 Dihedral-angle violation statistics for the ensemble

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

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

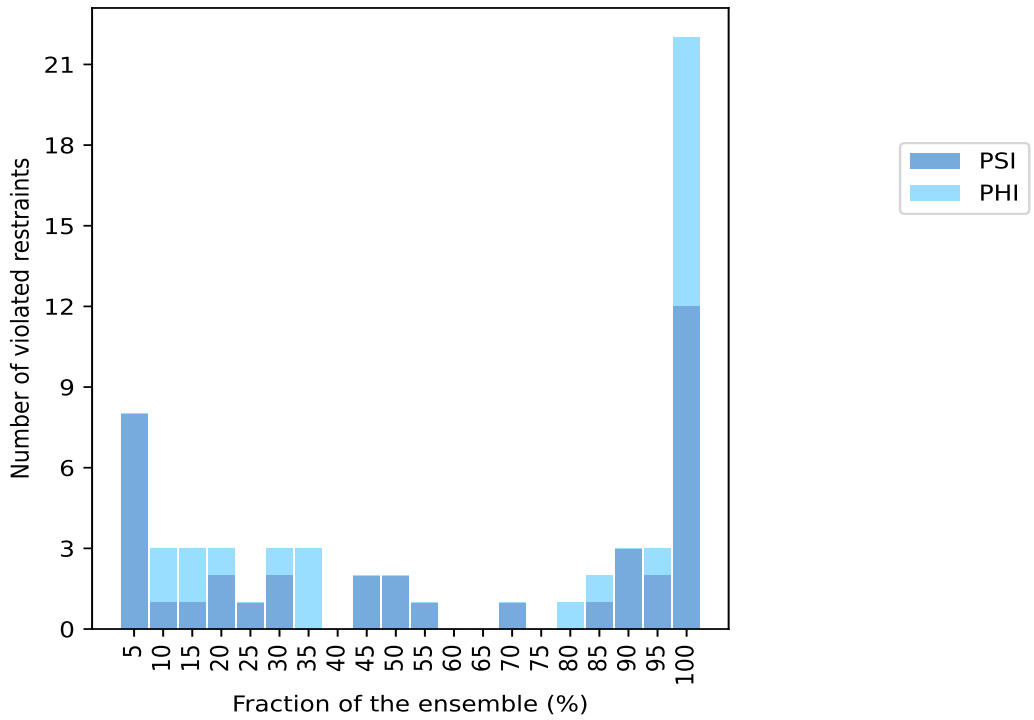
Continued on next page...

Continued from previous page...

Number of violated restraints			Fraction of the ensemble	
PSI	PHI	Total	Count[1]	%
0	0	0	12	60.0
0	0	0	13	65.0
1	0	1	14	70.0
0	0	0	15	75.0
0	1	1	16	80.0
1	1	2	17	85.0
3	0	3	18	90.0
2	1	3	19	95.0
12	10	22	20	100.0

[1] Number of models with violations

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

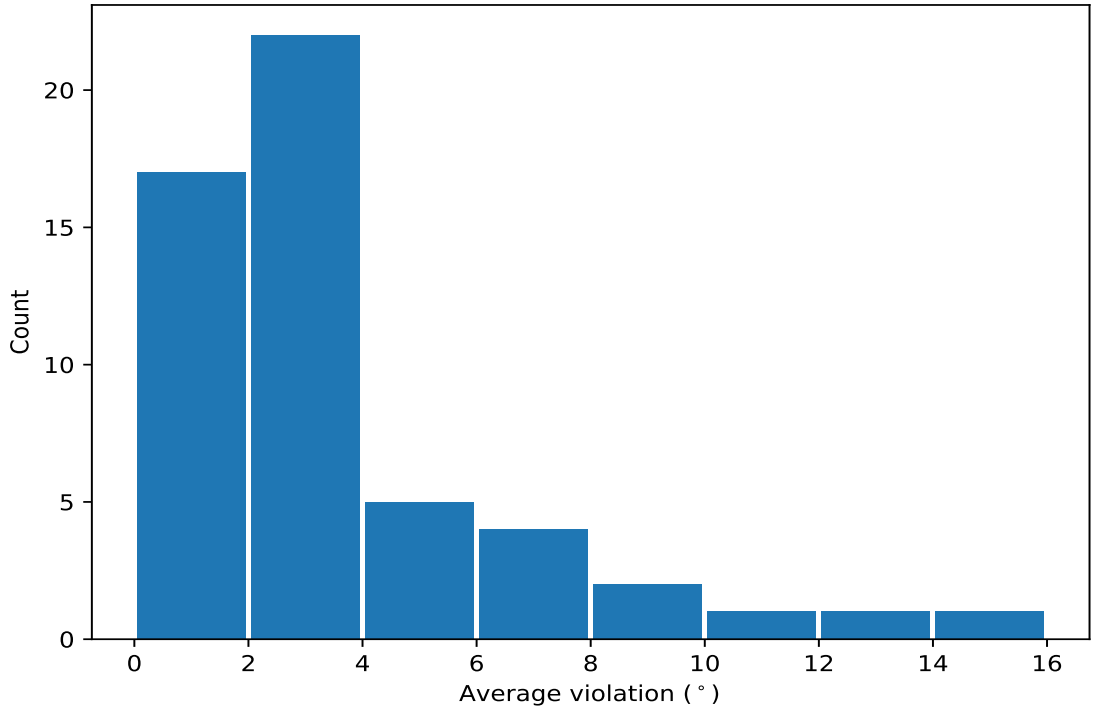


10.4 Most violated dihedral-angle restraints in the ensemble

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.

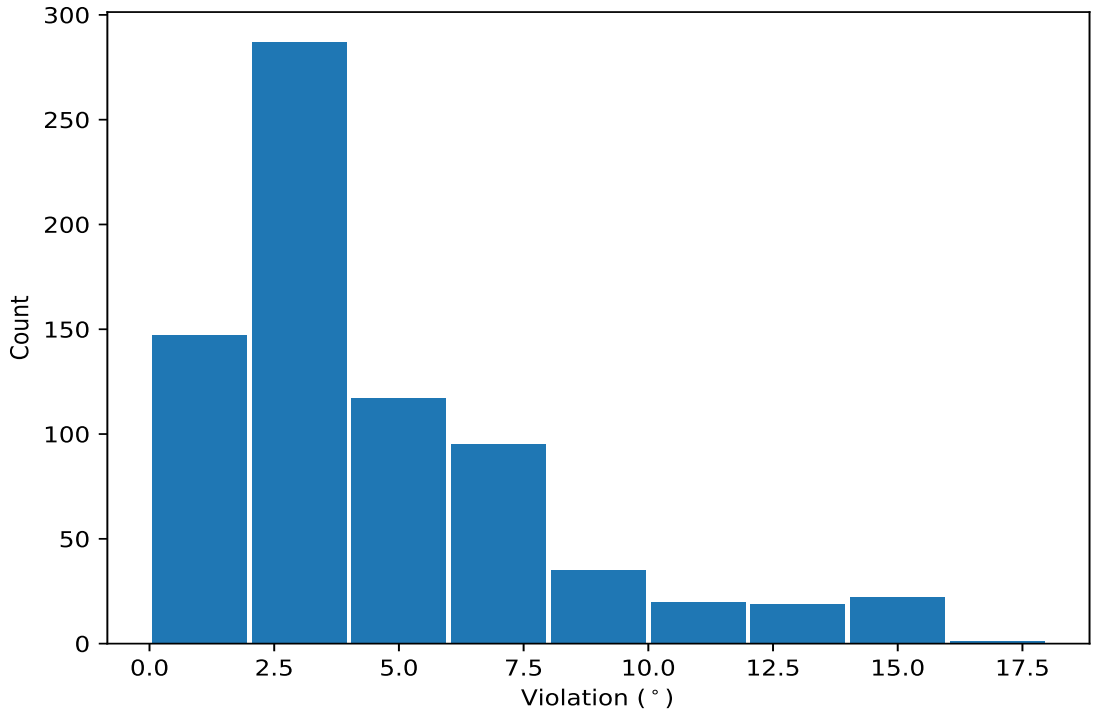
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models[1]	Mean	SD[2]	Median
(1,58)	1:A:371:GLY:N	1:A:371:GLY:CA	1:A:371:GLY:C	1:A:372:LYS:N	20	14.57	0.89	14.35
(1,2)	1:A:342:ASP:N	1:A:342:ASP:CA	1:A:342:ASP:C	1:A:343:THR:N	20	13.41	1.07	13.1
(1,157)	1:A:431:GLY:C	1:A:432:ASP:N	1:A:432:ASP:CA	1:A:432:ASP:C	20	11.34	1.59	11.0
(1,142)	1:A:423:SER:N	1:A:423:SER:CA	1:A:423:SER:C	1:A:424:ASP:N	20	8.74	1.67	8.65
(1,158)	1:A:432:ASP:N	1:A:432:ASP:CA	1:A:432:ASP:C	1:A:433:LEU:N	20	8.13	0.39	8.25
(1,153)	1:A:429:TYR:C	1:A:430:GLY:N	1:A:430:GLY:CA	1:A:430:GLY:C	20	7.42	0.53	7.35
(1,49)	1:A:365:GLU:C	1:A:366:THR:N	1:A:366:THR:CA	1:A:366:THR:C	20	7.24	0.7	7.3
(1,159)	1:A:432:ASP:C	1:A:433:LEU:N	1:A:433:LEU:CA	1:A:433:LEU:C	20	7.14	0.41	7.2
(1,143)	1:A:423:SER:C	1:A:424:ASP:N	1:A:424:ASP:CA	1:A:424:ASP:C	20	5.44	0.91	5.6
(1,206)	1:A:458:GLU:N	1:A:458:GLU:CA	1:A:458:GLU:C	1:A:459:LYS:N	20	4.93	0.85	4.6

[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 (°)
(1,58)	1:A:371:GLY:N	1:A:371:GLY:CA	1:A:371:GLY:C	1:A:372:LYS:N	11	16.6
(1,58)	1:A:371:GLY:N	1:A:371:GLY:CA	1:A:371:GLY:C	1:A:372:LYS:N	6	15.9
(1,58)	1:A:371:GLY:N	1:A:371:GLY:CA	1:A:371:GLY:C	1:A:372:LYS:N	17	15.7
(1,58)	1:A:371:GLY:N	1:A:371:GLY:CA	1:A:371:GLY:C	1:A:372:LYS:N	3	15.6
(1,2)	1:A:342:ASP:N	1:A:342:ASP:CA	1:A:342:ASP:C	1:A:343:THR:N	13	15.5
(1,58)	1:A:371:GLY:N	1:A:371:GLY:CA	1:A:371:GLY:C	1:A:372:LYS:N	13	15.1
(1,58)	1:A:371:GLY:N	1:A:371:GLY:CA	1:A:371:GLY:C	1:A:372:LYS:N	15	15.1
(1,2)	1:A:342:ASP:N	1:A:342:ASP:CA	1:A:342:ASP:C	1:A:343:THR:N	8	15.1
(1,58)	1:A:371:GLY:N	1:A:371:GLY:CA	1:A:371:GLY:C	1:A:372:LYS:N	5	15.0
(1,2)	1:A:342:ASP:N	1:A:342:ASP:CA	1:A:342:ASP:C	1:A:343:THR:N	19	15.0