



wwPDB NMR Structure Validation Summary Report ⓘ

Dec 25, 2024 – 02:53 PM EST

PDB ID : 6S3W
BMRB ID : 27397
Title : Solution NMR Structure of TolAIII Bound to a Peptide Derived from the N-terminus of TolB
Authors : Kleanthous, C.; Redfield, C.; Rajasekar, K.; Holmes, P.
Deposited on : 2019-06-26

This is a wwPDB NMR Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/NMRValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

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

MolProbity : 4.02b-467
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
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) : 2.40

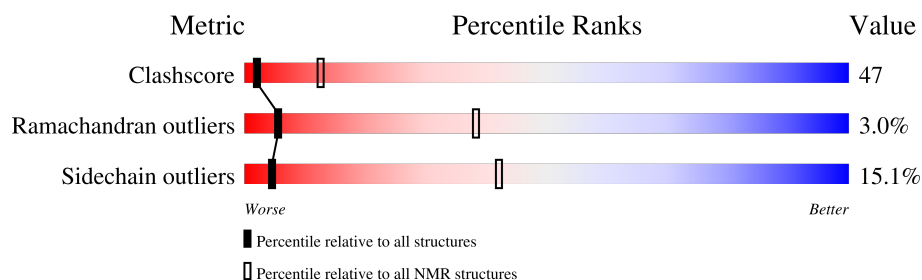
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 88%.

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	210492	14027
Ramachandran outliers	207382	12486
Sidechain outliers	206894	12463

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and a grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	B	13	
2	A	124	

2 Ensemble composition and analysis

This entry contains 20 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	B:3-B:7, A:254-A:342 (94)	0.30	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 2 clusters and 2 single-model clusters were found.

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

3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2110 atoms, of which 1056 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called TolBp.

Mol	Chain	Residues	Atoms					Trace
1	B	13	Total	C	H	N	O	0
			182	54	90	17	21	

- Molecule 2 is a protein called Cell envelope integrity/translocation protein TolA.

Mol	Chain	Residues	Atoms						Trace
2	A	124	Total	C	H	N	O	S	0
			1928	591	966	176	191	4	

There are 2 discrepancies between the modelled and reference sequences:

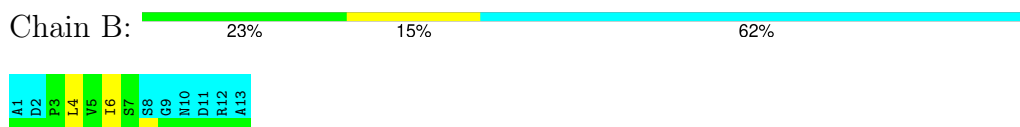
Chain	Residue	Modelled	Actual	Comment	Reference
A	224	HIS	-	expression tag	UNP A0A454LZ61
A	225	MET	-	expression tag	UNP A0A454LZ61

4 Residue-property plots

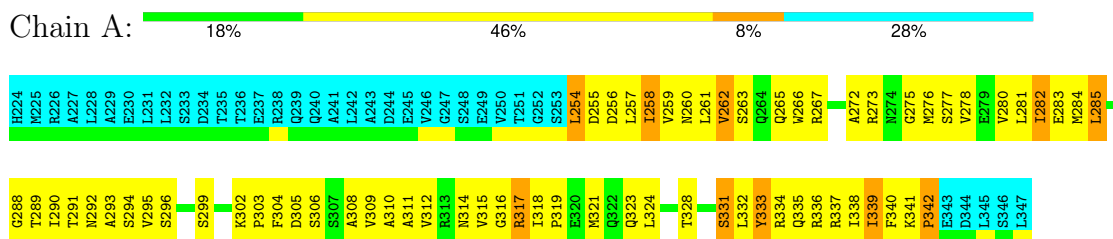
4.1 Average score per residue in the NMR ensemble

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: TolBp



- Molecule 2: Cell envelope integrity/translocation protein TolA



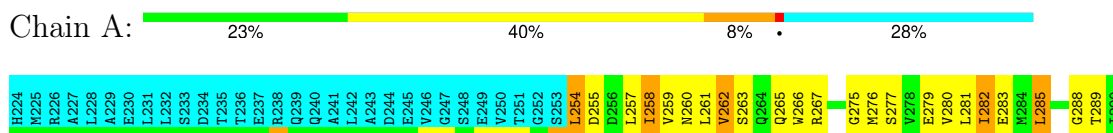
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

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

- Molecule 1: TolBp



- Molecule 2: Cell envelope integrity/translocation protein TolA



T291	N292	A293	S294	V295	S296	S299	K302	P303	F304	D305	S306	S307	A308	V309	A310	A311	V312	R313	N314	V315	G316	R317	T318	P319	E320	N321	L324	P325	R326	A327	T328	F329	D330	S331	L332	Y333	R334	Q335	R336	R337	I338	I339	F340	K341	P342	F343	D344	L345	S346	L347
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

5 Refinement protocol and experimental data overview

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

Of the 600 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
ARIA	structure calculation	2.3
CNS	structure calculation	1.2

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	2
Total number of shifts	1554
Number of shifts mapped to atoms	1554
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	88%

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	B	0.38±0.03	0±0/37 (0.0± 0.0%)	0.46±0.03	0±0/51 (0.0± 0.0%)
2	A	0.83±0.04	1±1/712 (0.1± 0.1%)	0.83±0.01	0±0/966 (0.0± 0.0%)
All	All	0.82	19/14980 (0.1%)	0.81	0/20340 (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
2	A	0.0±0.0	0.5±0.6
All	All	0	10

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
2	A	333	TYR	CE2-CZ	9.35	1.50	1.38	4	7
2	A	333	TYR	CE1-CZ	-8.73	1.27	1.38	6	7
2	A	329	PHE	CE1-CZ	6.24	1.49	1.37	8	5

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)
2	A	333	TYR	Sidechain	8
2	A	304	PHE	Sidechain	1
2	A	336	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	B	36	43	43	6±2
2	A	699	712	712	70±4
All	All	14700	15100	15100	1402

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 47.

5 of 249 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:A:272:ALA:HA	2:A:276:MET:SD	0.92	2.05	18	15
2:A:254:LEU:HB2	2:A:333:TYR:OH	0.81	1.75	17	9
2:A:258:ILE:HG13	2:A:318:ILE:HD13	0.79	1.55	13	5
2:A:281:LEU:HD11	2:A:335:GLN:HB3	0.79	1.55	12	20
2:A:336:ARG:NE	2:A:336:ARG:HA	0.78	1.91	1	3

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	B	5/13 (38%)	4±1 (90±13%)	0±1 (10±13%)	0±0 (0±0%)	100	100
2	A	89/124 (72%)	80±2 (90±2%)	6±1 (7±2%)	3±1 (3±1%)	5	36
All	All	1880/2740 (69%)	1687 (90%)	136 (7%)	57 (3%)	5	38

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)
2	A	342	PRO	19

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
2	A	275	GLY	16
2	A	316	GLY	10
2	A	319	PRO	8
2	A	254	LEU	2

6.3.2 Protein sidechains ⓘ

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	5/10 (50%)	5±0 (98±6%)	0±0 (2±6%)	50	92
2	A	79/108 (73%)	66±2 (84±3%)	13±2 (16±3%)	4	40
All	All	1680/2360 (71%)	1427 (85%)	253 (15%)	4	42

5 of 39 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)
2	A	258	ILE	20
2	A	294	SER	20
2	A	331	SER	20
2	A	339	ILE	20
2	A	262	VAL	18

6.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

6.5 Carbohydrates ⓘ

There are no oligosaccharides 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 88% for the well-defined parts and 84% 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	1428
Number of shifts mapped to atoms	1428
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 \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	115	0.01 ± 0.20	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	108	0.50 ± 0.13	Should be checked
$^{13}\text{C}'$	110	-0.20 ± 0.21	None needed (< 0.5 ppm)
^{15}N	115	-0.39 ± 0.32	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 84%, i.e. 1100 atoms were assigned a chemical shift out of a possible 1315. 0 out of 17 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	424/458 (93%)	174/184 (95%)	168/188 (89%)	82/86 (95%)
Sidechain	634/806 (79%)	431/525 (82%)	193/240 (80%)	10/41 (24%)

Continued on next page...

Continued from previous page...

	Total	¹ H	¹³ C	¹⁵ N
Aromatic	42/51 (82%)	21/25 (84%)	20/25 (80%)	1/1 (100%)
Overall	1100/1315 (84%)	626/734 (85%)	381/453 (84%)	93/128 (73%)

7.1.4 Statistically unusual chemical shifts [i](#)

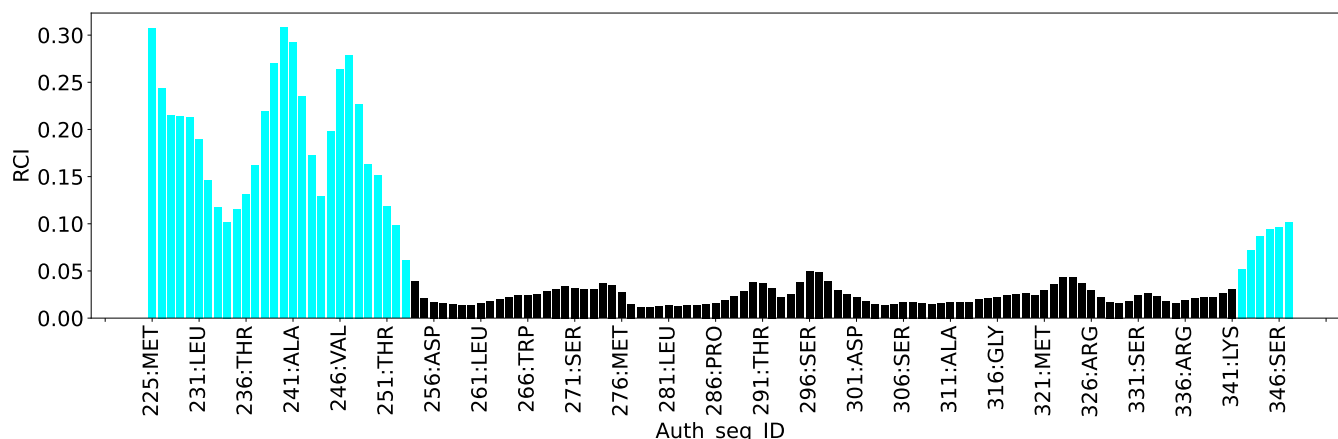
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	286	PRO	HA	2.36	2.78 – 6.00	-6.3
1	A	308	ALA	HB1	-0.18	0.14 – 2.58	-6.3
1	A	308	ALA	HB2	-0.18	0.14 – 2.58	-6.3
1	A	308	ALA	HB3	-0.18	0.14 – 2.58	-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:



7.2 Chemical shift list 2

File name: working_cs.cif

Chemical shift list name: *assignedchem_shift_list_2*

7.2.1 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	126
Number of shifts mapped to atoms	126
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

7.2.2 Chemical shift referencing [i](#)

No chemical shift referencing corrections were calculated (not enough data).

7.2.3 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 4%, i.e. 56 atoms were assigned a chemical shift out of a possible 1315. 0 out of 17 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹ H	¹³ C	¹⁵ N
Backbone	23/458 (5%)	9/184 (5%)	10/188 (5%)	4/86 (5%)
Sidechain	33/806 (4%)	26/525 (5%)	7/240 (3%)	0/41 (0%)
Aromatic	0/51 (0%)	0/25 (0%)	0/25 (0%)	0/1 (0%)
Overall	56/1315 (4%)	35/734 (5%)	17/453 (4%)	4/128 (3%)

7.2.4 Statistically unusual chemical shifts [i](#)

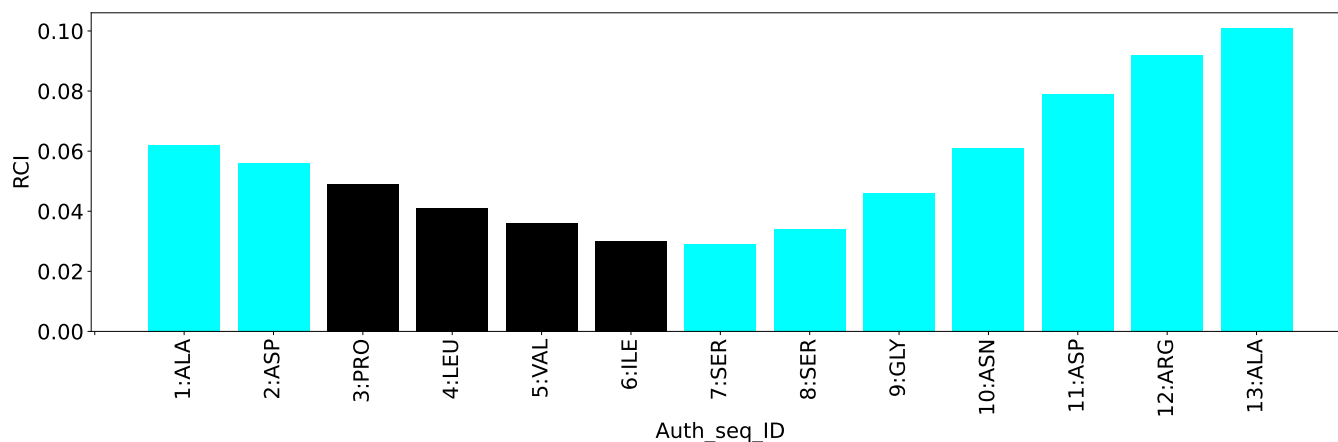
There are no statistically unusual chemical shifts.

7.2.5 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 B:



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	5115
Intra-residue ($ i-j =0$)	2371
Sequential ($ i-j =1$)	1147
Medium range ($ i-j >1$ and $ i-j <5$)	689
Long range ($ i-j \geq 5$)	790
Inter-chain	87
Hydrogen bond restraints	31
Disulfide bond restraints	0
Total dihedral-angle restraints	151
Number of unmapped restraints	0
Number of restraints per residue	38.4
Number of long range restraints per residue ¹	5.8

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

8.2 Residual restraint violations

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

8.2.1 Average number of distance violations per model

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

Bins (Å)	Average number of violations per model	Max (Å)
0.1-0.2 (Small)	291.2	0.2
0.2-0.5 (Medium)	482.9	0.5
>0.5 (Large)	282.3	3.64

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)	9.7	9.54
10.0-20.0 (Medium)	0.2	11.87
>20.0 (Large)	None	None

9 Distance violation analysis ⓘ

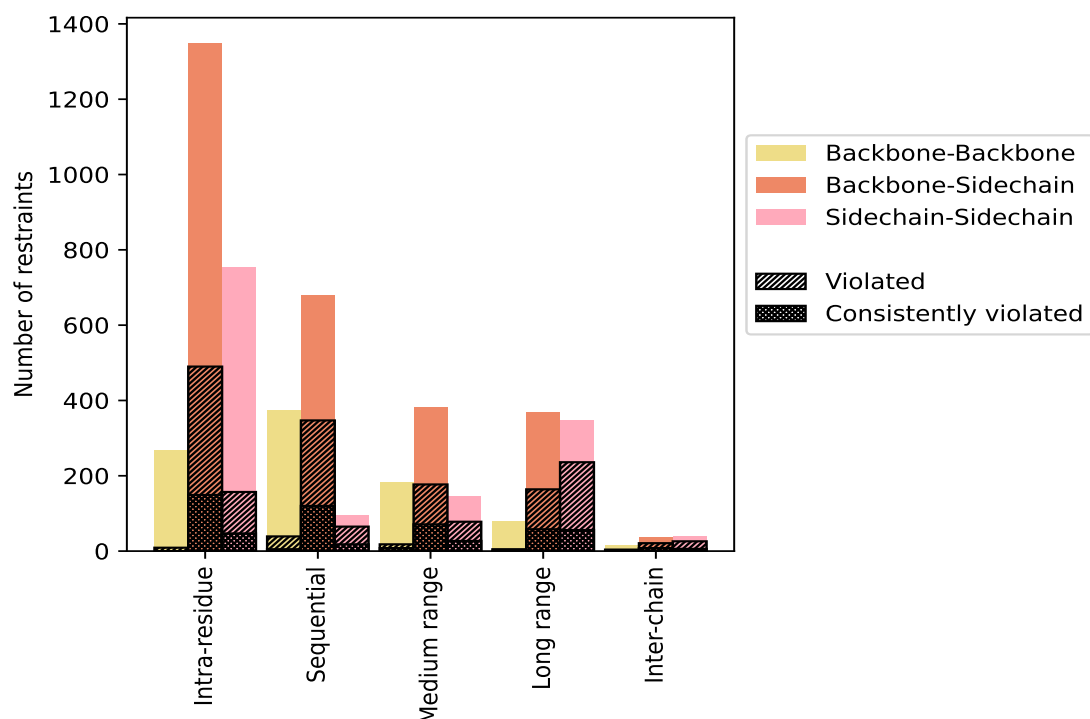
9.1 Summary of distance violations ⓘ

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

Restrains type	Count	% ¹	Violated ³			Consistently Violated ⁴		
			Count	% ²	% ¹	Count	% ²	% ¹
Intra-residue (i-j =0)	2371	46.4	656	27.7	12.8	196	8.3	3.8
Backbone-Backbone	268	5.2	9	3.4	0.2	0	0.0	0.0
Backbone-Sidechain	1349	26.4	490	36.3	9.6	149	11.0	2.9
Sidechain-Sidechain	754	14.7	157	20.8	3.1	47	6.2	0.9
Sequential (i-j =1)	1147	22.4	451	39.3	8.8	143	12.5	2.8
Backbone-Backbone	374	7.3	39	10.4	0.8	5	1.3	0.1
Backbone-Sidechain	679	13.3	347	51.1	6.8	120	17.7	2.3
Sidechain-Sidechain	94	1.8	65	69.1	1.3	18	19.1	0.4
Medium range (i-j >1 & i-j <5)	689	13.5	271	39.3	5.3	106	15.4	2.1
Backbone-Backbone	182	3.6	18	9.9	0.4	8	4.4	0.2
Backbone-Sidechain	361	7.1	175	48.5	3.4	71	19.7	1.4
Sidechain-Sidechain	146	2.9	78	53.4	1.5	27	18.5	0.5
Long range (i-j ≥5)	790	15.4	405	51.3	7.9	115	14.6	2.2
Backbone-Backbone	80	1.6	5	6.2	0.1	1	1.2	0.0
Backbone-Sidechain	363	7.1	164	45.2	3.2	58	16.0	1.1
Sidechain-Sidechain	347	6.8	236	68.0	4.6	56	16.1	1.1
Inter-chain	87	1.7	50	57.5	1.0	14	16.1	0.3
Backbone-Backbone	16	0.3	4	25.0	0.1	0	0.0	0.0
Backbone-Sidechain	33	0.6	20	60.6	0.4	8	24.2	0.2
Sidechain-Sidechain	38	0.7	26	68.4	0.5	6	15.8	0.1
Hydrogen bond	31	0.6	3	9.7	0.1	0	0.0	0.0
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	5115	100.0	1836	35.9	35.9	574	11.2	11.2
Backbone-Backbone	920	18.0	75	8.2	1.5	14	1.5	0.3
Backbone-Sidechain	2816	55.1	1199	42.6	23.4	406	14.4	7.9
Sidechain-Sidechain	1379	27.0	562	40.8	11.0	154	11.2	3.0

¹ percentage calculated with respect to the total number of distance restraints, ² percentage calculated with respect to the number of restraints in a particular restraint category, ³ violated in at least one model, ⁴ violated in all the models

9.1.1 Bar chart : Distribution of distance restraints and violations [i](#)



Violated and consistently violated restraints are shown using different hatch patterns in their respective categories. The hydrogen bonds and 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	363	250	173	224	27	1037	0.42	2.96	0.33	0.32
2	368	258	174	220	25	1045	0.41	2.9	0.34	0.3
3	379	260	178	236	24	1077	0.41	1.88	0.32	0.32
4	377	265	172	241	28	1083	0.43	3.6	0.41	0.31
5	385	272	181	258	30	1126	0.42	2.89	0.37	0.31
6	398	263	176	242	27	1106	0.43	3.64	0.4	0.31
7	385	262	175	234	27	1083	0.41	2.7	0.33	0.31
8	374	261	169	242	25	1071	0.42	2.74	0.33	0.32
9	376	274	175	224	31	1080	0.4	2.77	0.32	0.31
10	382	279	180	240	30	1111	0.43	3.38	0.39	0.31

Continued on next page...

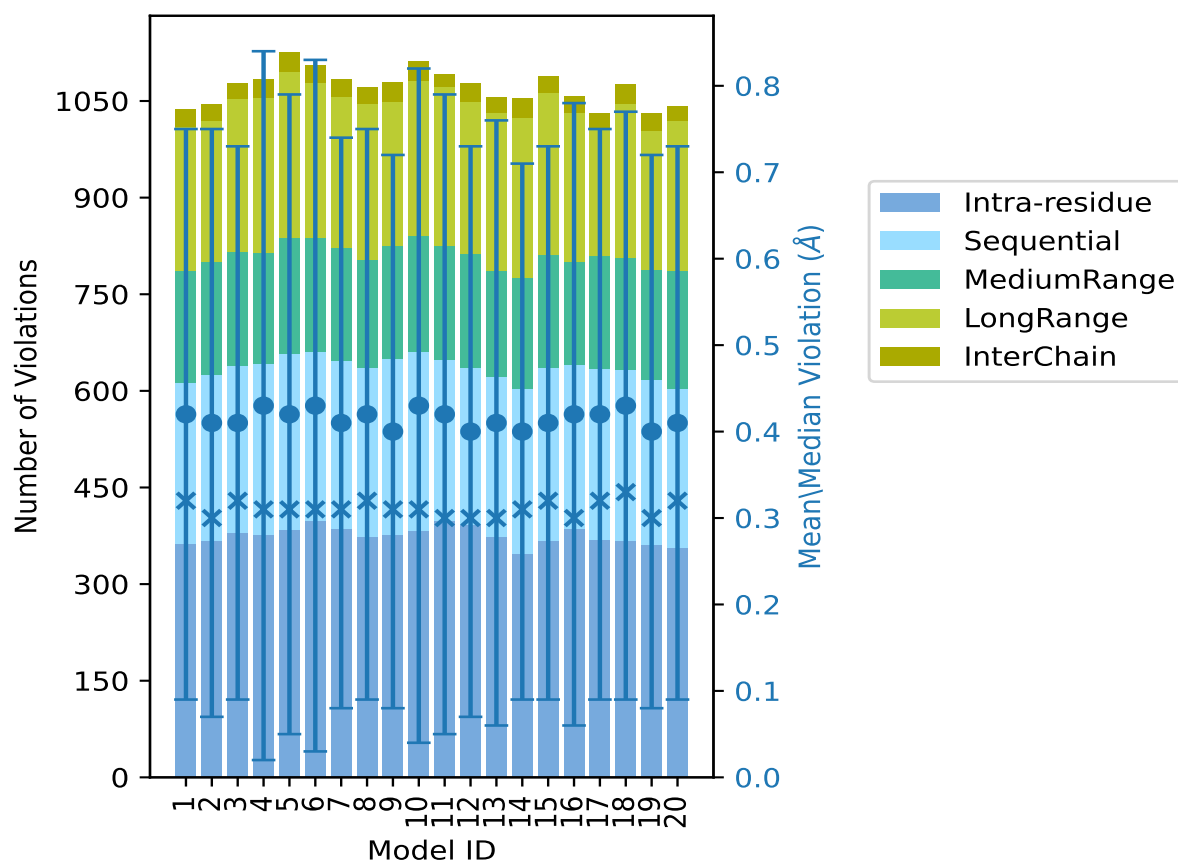
Continued from previous page...

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				
11	399	250	177	246	20	1092	0.42	3.0	0.37	0.3
12	392	244	177	236	28	1077	0.4	2.96	0.33	0.3
13	373	248	166	245	24	1056	0.41	3.0	0.35	0.3
14	347	256	173	248	30	1054	0.4	2.89	0.31	0.31
15	367	269	175	251	26	1088	0.41	1.93	0.32	0.32
16	386	254	161	231	25	1057	0.42	2.83	0.36	0.3
17	368	267	174	200	23	1032	0.42	1.78	0.33	0.32
18	367	265	175	239	30	1076	0.43	2.68	0.34	0.33
19	361	256	171	216	27	1031	0.4	1.8	0.32	0.3
20	356	247	183	234	22	1042	0.41	1.78	0.32	0.32

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

⁵Inter-chain restraints, ⁶Standard deviation

9.2.1 Bar graph : Distance Violation statistics for each model ⓘ



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

9.3 Distance violation statistics for the ensemble

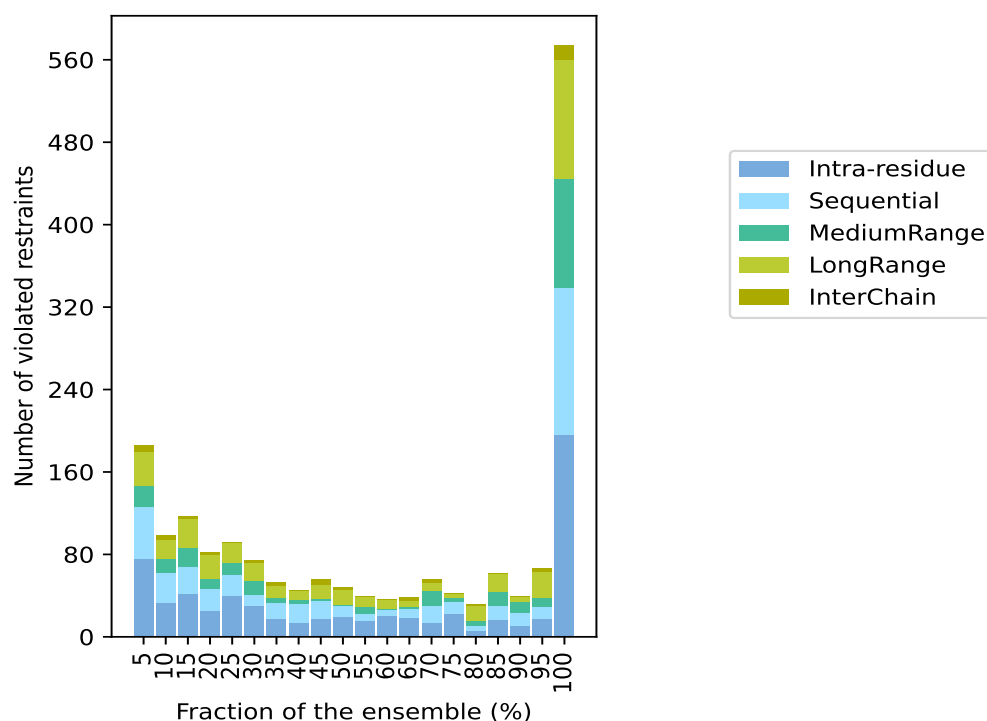
Violation analysis may find that some restraints are violated in few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of the ensemble. In total, 3251(IR:1715, SQ:696, MR:418, LR:385, IC:37) restraints are 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	%
76	50	21	33	6	186	1	5.0
33	29	14	18	4	98	2	10.0
42	26	18	29	2	117	3	15.0
25	22	9	24	2	82	4	20.0
40	20	12	20	0	92	5	25.0
30	11	13	18	2	74	6	30.0
18	15	5	12	3	53	7	35.0
14	18	4	9	0	45	8	40.0
18	17	2	14	5	56	9	45.0
20	10	1	15	2	48	10	50.0
16	7	6	10	0	39	11	55.0
21	6	0	9	0	36	12	60.0
19	8	2	6	3	38	13	65.0
14	16	15	8	3	56	14	70.0
22	12	4	4	0	42	15	75.0
6	5	5	14	1	31	16	80.0
17	13	14	17	0	61	17	85.0
11	12	11	5	0	39	18	90.0
18	11	9	25	3	66	19	95.0
196	143	106	115	14	574	20	100.0

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

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

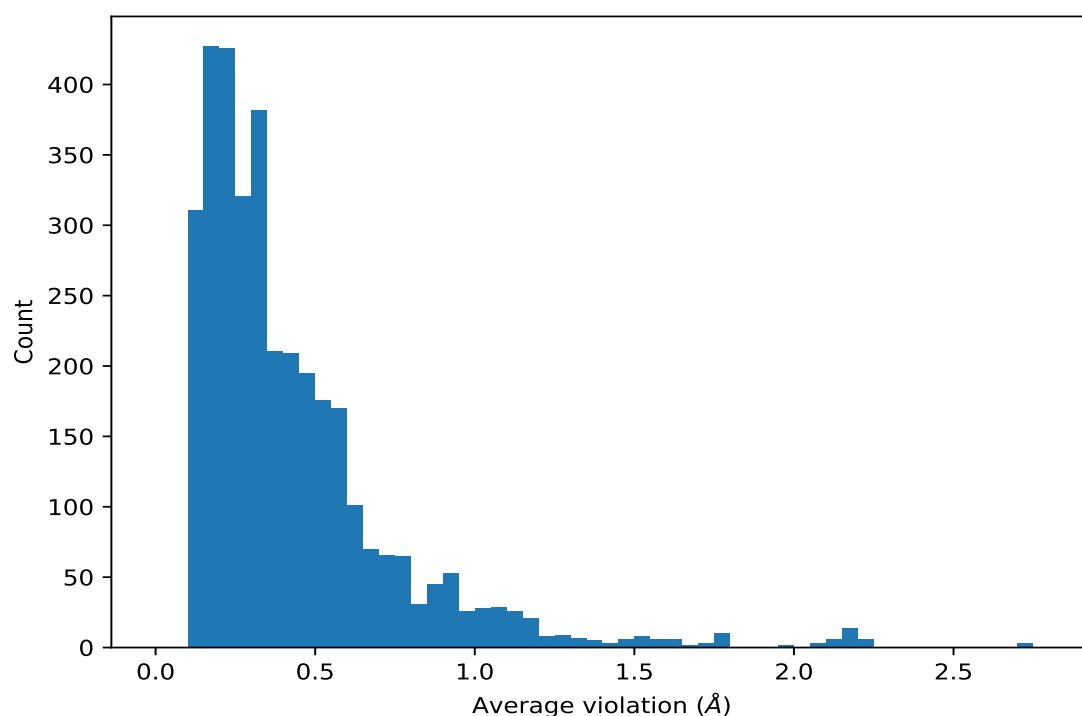
9.3.1 Bar graph : Distance violation statistics for the ensemble [i](#)



9.4 Most violated distance restraints in the ensemble [i](#)

9.4.1 Histogram : Distribution of mean distance violations [i](#)

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



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

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

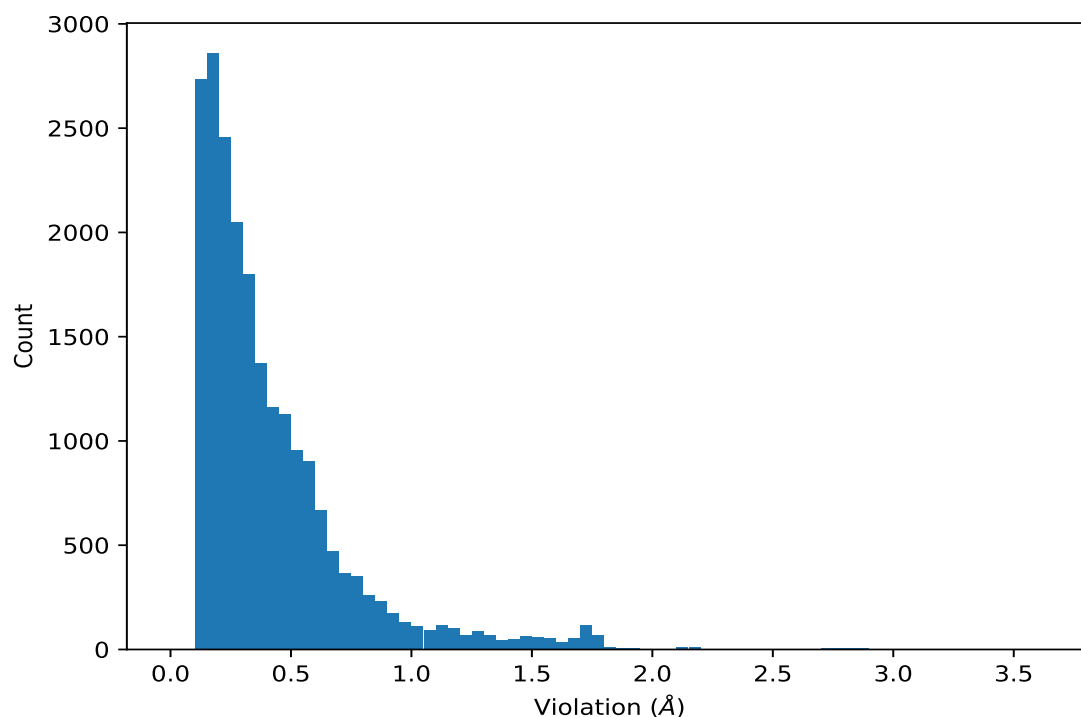
Key	Atom-1	Atom-2	Models ¹	Mean (Å)	SD ¹ (Å)	Median (Å)
(4,3319)	2:285:A:LEU:HD21	2:289:A:THR:HB	20	2.16	1.04	2.74
(4,3319)	2:285:A:LEU:HD23	2:289:A:THR:HB	20	2.16	1.04	2.74
(4,3319)	2:285:A:LEU:HD22	2:289:A:THR:HB	20	2.16	1.04	2.74
(4,431)	2:285:A:LEU:HD21	2:289:A:THR:HB	20	2.15	1.04	2.73
(4,431)	2:285:A:LEU:HD23	2:289:A:THR:HB	20	2.15	1.04	2.73
(4,431)	2:285:A:LEU:HD22	2:289:A:THR:HB	20	2.15	1.04	2.73
(4,3940)	2:264:A:GLN:H	2:265:A:GLN:HB3	20	1.75	0.02	1.75
(4,3288)	2:304:A:PHE:HA	2:303:A:PRO:HB3	20	1.74	0.01	1.74
(4,413)	2:304:A:PHE:HA	2:303:A:PRO:HB3	20	1.72	0.01	1.72
(4,904)	2:264:A:GLN:H	2:265:A:GLN:HB3	20	1.71	0.02	1.71
(4,4127)	2:275:A:GLY:HA2	2:342:A:PRO:HB3	20	1.69	0.12	1.73
(4,1054)	2:275:A:GLY:HA2	2:342:A:PRO:HB3	20	1.65	0.12	1.69
(4,2871)	2:254:A:LEU:HB3	2:251:A:THR:HA	20	1.61	0.14	1.65
(4,74)	2:254:A:LEU:HB3	2:251:A:THR:HA	20	1.59	0.15	1.63

¹Number of violated models, ²Standard deviation

9.5 All violated distance restraints [i](#)

9.5.1 Histogram : Distribution of distance violations [i](#)

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



9.5.2 Table : All distance violations [i](#)

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

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(4,3317)	2:295:A:VAL:HG12	2:278:A:VAL:HG11	6	3.64
(4,429)	2:295:A:VAL:HG12	2:278:A:VAL:HG11	6	3.62
(4,3317)	2:295:A:VAL:HG11	2:278:A:VAL:HG12	4	3.6
(4,429)	2:295:A:VAL:HG11	2:278:A:VAL:HG12	4	3.59
(4,3317)	2:295:A:VAL:HG11	2:278:A:VAL:HG13	10	3.38
(4,429)	2:295:A:VAL:HG11	2:278:A:VAL:HG13	10	3.37
(4,3319)	2:285:A:LEU:HD22	2:289:A:THR:HB	4	3.07
(4,431)	2:285:A:LEU:HD22	2:289:A:THR:HB	4	3.06
(4,2793)	2:340:A:PHE:HB3	2:278:A:VAL:HG21	4	3.05
(4,1)	2:340:A:PHE:HB3	2:278:A:VAL:HG21	4	3.03

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

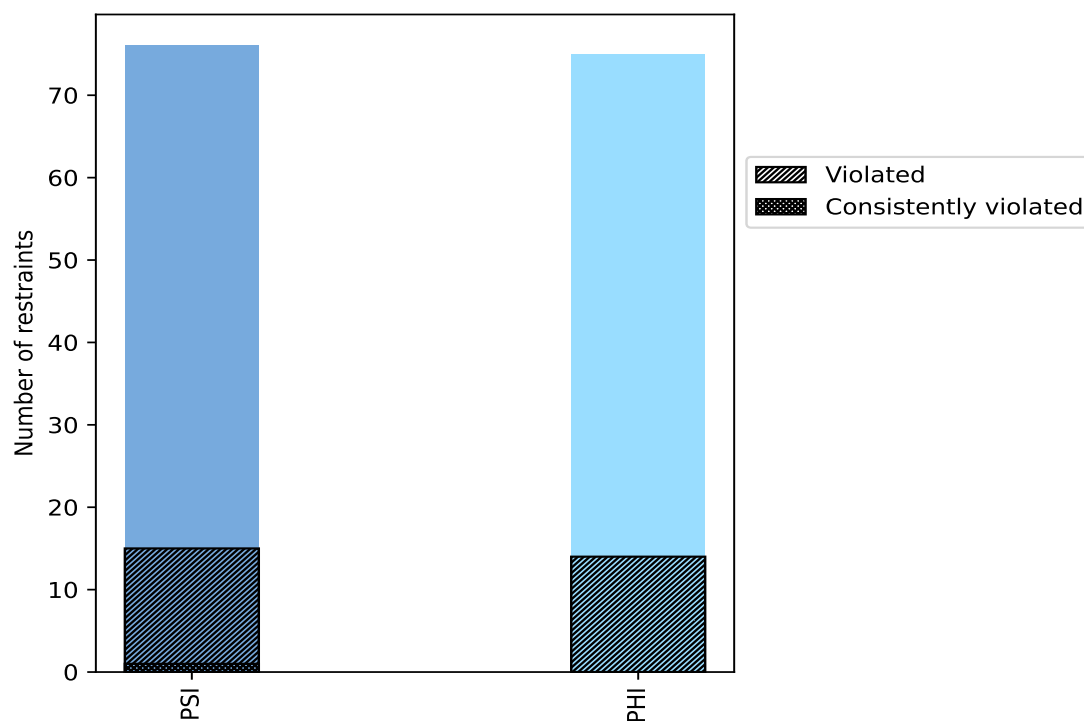
10.1 Summary of dihedral-angle violations [i](#)

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	76	50.3	15	19.7	9.9	1	1.3	0.7
PHI	75	49.7	14	18.7	9.3	0	0.0	0.0
Total	151	100.0	29	19.2	19.2	1	0.7	0.7

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](#)



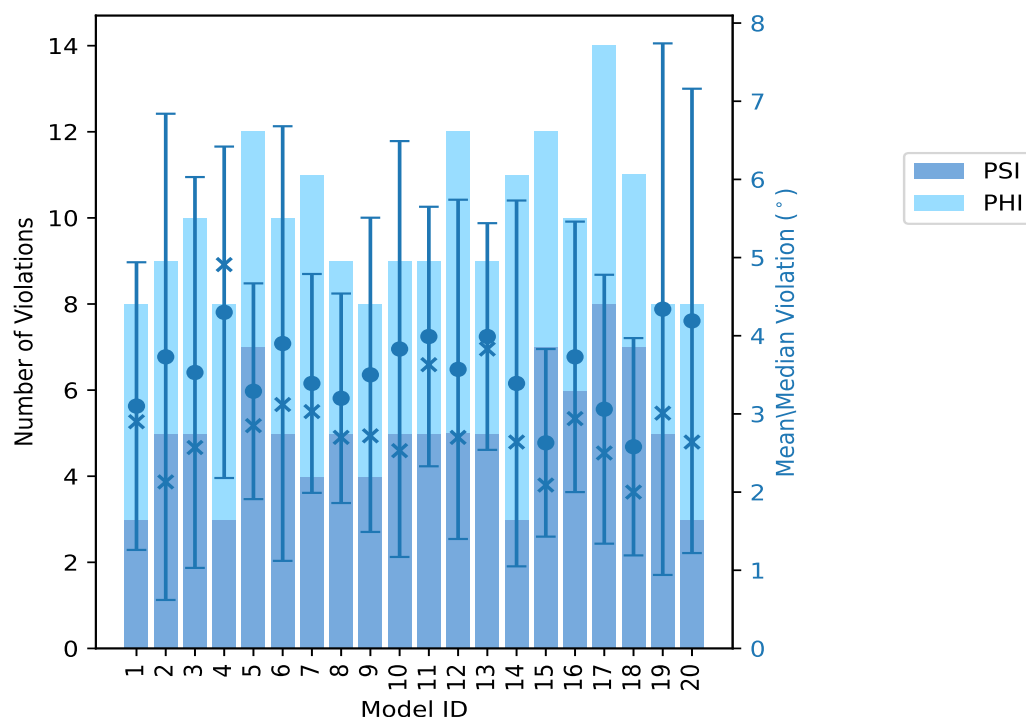
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

10.2 Dihedral-angle violation statistics for each model

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	3	5	8	3.1	6.2	1.84	2.9
2	5	4	9	3.73	10.75	3.11	2.13
3	5	5	10	3.53	10.07	2.5	2.57
4	3	5	8	4.3	7.52	2.12	4.91
5	7	5	12	3.29	6.11	1.38	2.85
6	5	5	10	3.9	11.04	2.78	3.12
7	4	7	11	3.39	5.7	1.4	3.03
8	5	4	9	3.2	6.16	1.34	2.7
9	4	4	8	3.5	6.54	2.01	2.72
10	5	4	9	3.83	9.54	2.66	2.53
11	5	4	9	3.99	6.98	1.66	3.63
12	5	7	12	3.57	7.75	2.17	2.7
13	5	4	9	3.99	6.69	1.45	3.83
14	3	8	11	3.39	9.01	2.34	2.64
15	7	5	12	2.63	5.3	1.2	2.09
16	6	4	10	3.73	7.18	1.73	2.94
17	8	6	14	3.06	7.11	1.72	2.5
18	7	4	11	2.58	5.14	1.39	2.0
19	5	3	8	4.34	11.87	3.4	3.01
20	3	5	8	4.19	9.35	2.97	2.64

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



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

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

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

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

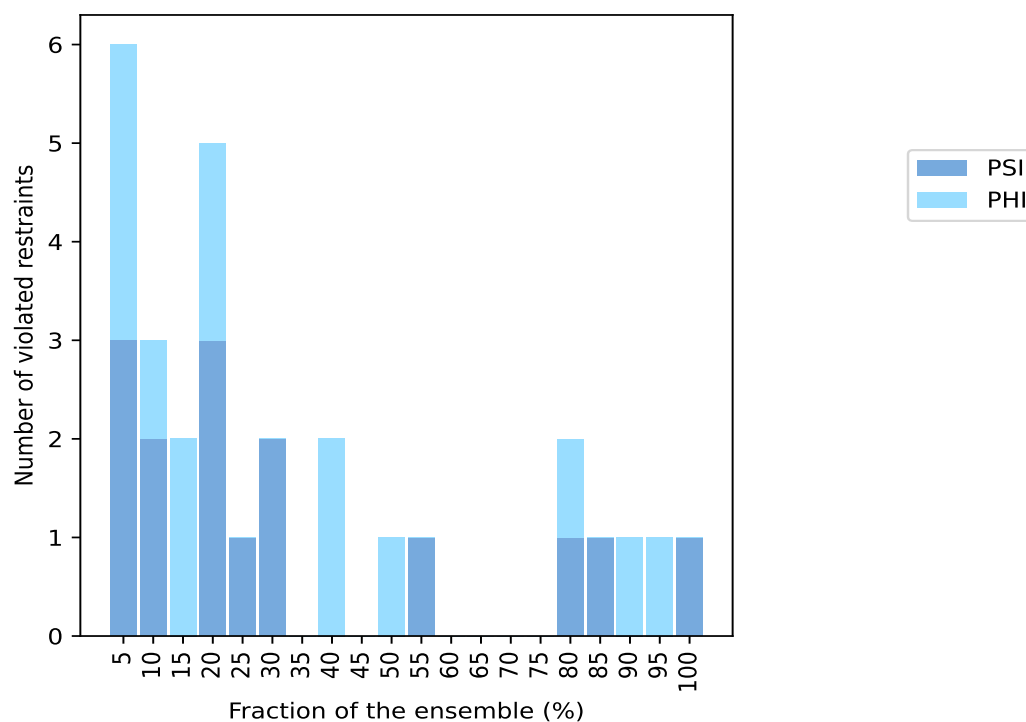
Continued on next page...

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
0	0	0	14	70.0
0	0	0	15	75.0
1	1	2	16	80.0
1	0	1	17	85.0
0	1	1	18	90.0
0	1	1	19	95.0
1	0	1	20	100.0

¹ Number of models with violations

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

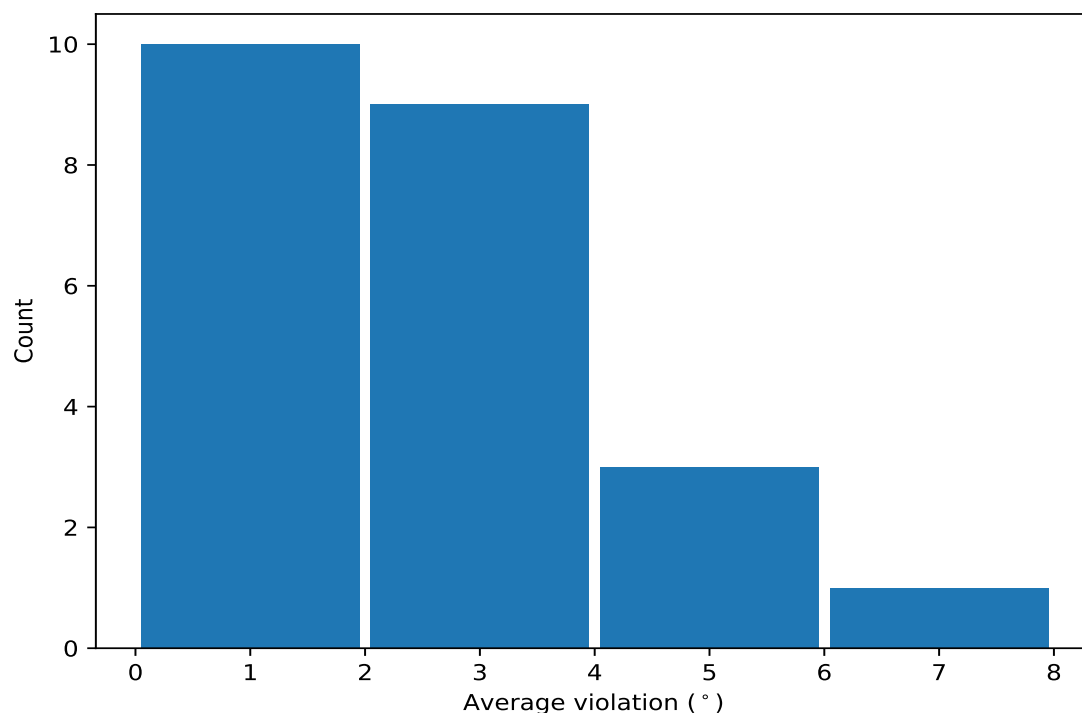


10.4 Most violated dihedral-angle restraints in the ensemble ⓘ

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 [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.

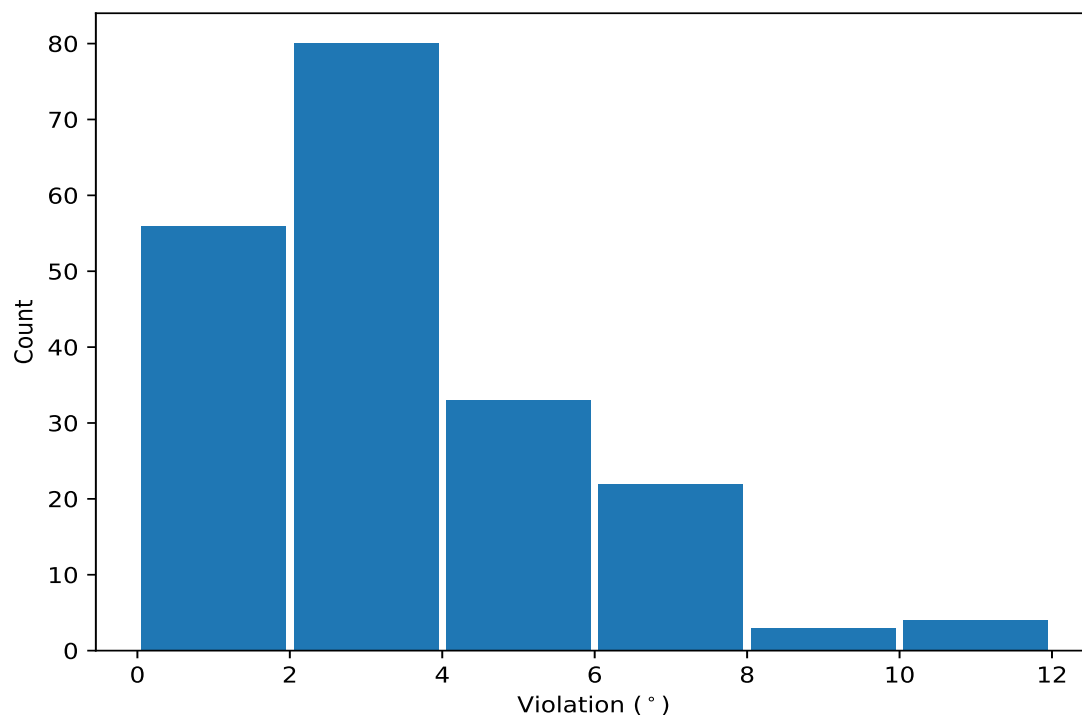
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models ¹	Mean	SD ²	Median
(1,64)	2:294:A:SER:N	2:294:A:SER:CA	2:294:A:SER:C	2:295:A:VAL:N	20	5.97	0.9	5.87
(1,25)	2:266:A:TRP:C	2:267:A:ARG:N	2:267:A:ARG:CA	2:267:A:ARG:C	19	4.13	1.28	3.98
(1,119)	2:331:A:SER:C	2:332:A:LEU:N	2:332:A:LEU:CA	2:332:A:LEU:C	18	3.1	0.81	2.99
(1,36)	2:273:A:ARG:N	2:273:A:ARG:CA	2:273:A:ARG:C	2:274:A:ASN:N	17	3.95	1.72	3.69
(1,125)	2:336:A:ARG:C	2:337:A:ARG:N	2:337:A:ARG:CA	2:337:A:ARG:C	16	7.97	2.42	7.49
(1,60)	2:292:A:ASN:N	2:292:A:ASN:CA	2:292:A:ASN:C	2:293:A:ALA:N	16	2.78	1.1	2.46
(1,24)	2:266:A:TRP:N	2:266:A:TRP:CA	2:266:A:TRP:C	2:267:A:ARG:N	11	2.18	0.65	1.98
(1,101)	2:317:A:ARG:C	2:318:A:ILE:N	2:318:A:ILE:CA	2:318:A:ILE:C	10	3.19	1.09	2.96
(1,17)	2:262:A:VAL:C	2:263:A:SER:N	2:263:A:SER:CA	2:263:A:SER:C	8	1.93	1.0	1.6
(1,99)	2:316:A:GLY:C	2:317:A:ARG:N	2:317:A:ARG:CA	2:317:A:ARG:C	8	1.67	0.4	1.48

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

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

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

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



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

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

Key	Atom-1	Atom-2	Atom-3	Atom-4	Model ID	Violation (°)
(1,125)	2:336:A:ARG:C	2:337:A:ARG:N	2:337:A:ARG:CA	2:337:A:ARG:C	19	11.87
(1,125)	2:336:A:ARG:C	2:337:A:ARG:N	2:337:A:ARG:CA	2:337:A:ARG:C	6	11.04
(1,125)	2:336:A:ARG:C	2:337:A:ARG:N	2:337:A:ARG:CA	2:337:A:ARG:C	2	10.75
(1,125)	2:336:A:ARG:C	2:337:A:ARG:N	2:337:A:ARG:CA	2:337:A:ARG:C	3	10.07
(1,125)	2:336:A:ARG:C	2:337:A:ARG:N	2:337:A:ARG:CA	2:337:A:ARG:C	10	9.54
(1,125)	2:336:A:ARG:C	2:337:A:ARG:N	2:337:A:ARG:CA	2:337:A:ARG:C	20	9.35
(1,125)	2:336:A:ARG:C	2:337:A:ARG:N	2:337:A:ARG:CA	2:337:A:ARG:C	14	9.01
(1,36)	2:273:A:ARG:N	2:273:A:ARG:CA	2:273:A:ARG:C	2:274:A:ASN:N	12	7.75
(1,64)	2:294:A:SER:N	2:294:A:SER:CA	2:294:A:SER:C	2:295:A:VAL:N	19	7.65
(1,64)	2:294:A:SER:N	2:294:A:SER:CA	2:294:A:SER:C	2:295:A:VAL:N	2	7.55