wwPDB NMR Structure Validation Summary Report (i)

Jun 5, 2023 – 02:41 AM EDT

PDB ID	:	2LQW
BMRB ID	:	18333
Title	:	Solution structure of phosphorylated CRKL
Authors	:	Jankowski, W.; Saleh, T.; Kalodimos, C.
Deposited on	:	2012-03-16

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

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

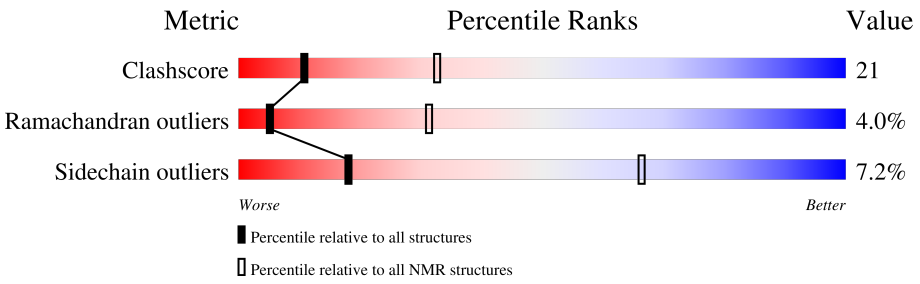
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (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 52%.

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	303	43% 25% • 31%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 10 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *closest to the average*.

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:4-A:104, A:125-A:180 (157)	0.80	10
2	A:238-A:247, A:254-A:295 (52)	0.86	1

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 4 clusters and 3 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 4705 atoms, of which 2317 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Crk-like protein.

Mol	Chain	Residues	Atoms							Trace
			Total	C	H	N	O	P	S	
1	A	303	4705	1504	2317	419	456	1	8	0

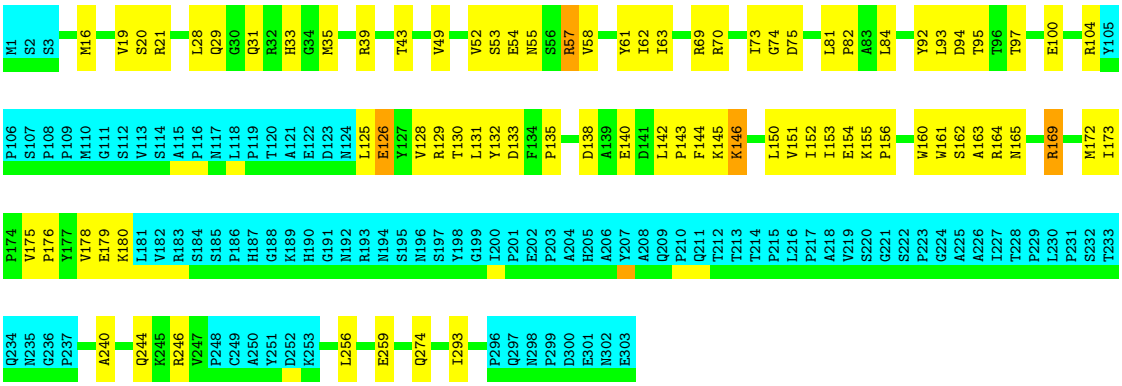
4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

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

- Molecule 1: Crk-like protein

Chain A: 43% 25% • 31%

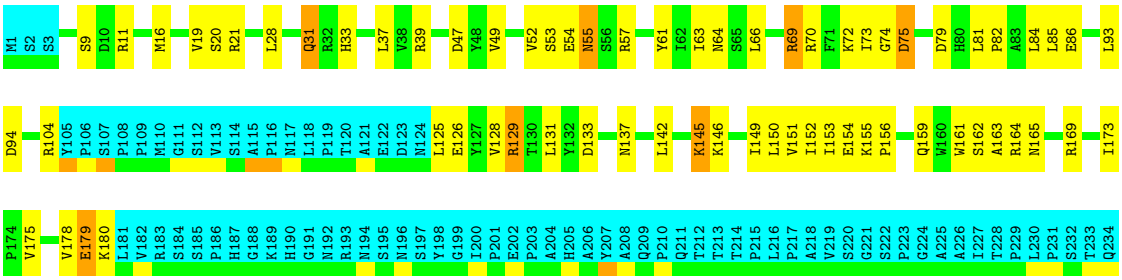


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 10. Colouring as in section 4.1 above.

- Molecule 1: Crk-like protein

Chain A: 44% 22% • 31%



N235
G236
P237
K241
R246
V247
P248
C249
A250
Y251
D252
K253
L256
R268
E276
H290
F294
D295
P296
Q297
N298
P299
D300
E301
N302
E303

5 Refinement protocol and experimental data overview (i)

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

Of the 400 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
CNS	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	1952
Number of shifts mapped to atoms	1952
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	52%

6 Model quality (i)

6.1 Standard geometry (i)

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

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	0.35±0.02	0±0/1745 (0.0± 0.0%)	0.49±0.04	0±1/2366 (0.0± 0.0%)
All	All	0.35	0/34900 (0.0%)	0.49	4/47320 (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	0.8±1.1
All	All	0	16

There are no bond-length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	57	ARG	N-CA-CB	-6.79	98.38	110.60	15	2
1	A	127	TYR	CB-CG-CD1	-5.65	117.61	121.00	9	1
1	A	68	ASN	N-CA-CB	5.09	119.76	110.60	15	1

There are no chirality outliers.

5 of 7 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	169	ARG	Sidechain	10
1	A	39	ARG	Sidechain	1
1	A	178	VAL	Peptide	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	56	SER	Peptide	1
1	A	57	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	1701	1681	1681	70±14
All	All	34020	33620	33620	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 21.

5 of 682 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:142:LEU:HD11	1:A:165:ASN:HB2	1.09	1.24	13	14
1:A:242:ALA:HB1	1:A:259:GLU:HA	1.00	1.33	12	3
1:A:152:ILE:HD13	1:A:163:ALA:HB2	0.98	1.36	20	20
1:A:238:VAL:H	1:A:266:VAL:HG22	0.96	1.21	7	1
1:A:238:VAL:N	1:A:266:VAL:HG22	0.95	1.75	7	1

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	209/303 (69%)	172±5 (82±2%)	29±5 (14±2%)	8±2 (4±1%)	5	31
All	All	4180/6060 (69%)	3431 (82%)	582 (14%)	167 (4%)	5	31

5 of 60 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	126	GLU	15
1	A	55	ASN	9
1	A	74	GLY	9
1	A	9	SER	8
1	A	156	PRO	8

6.3.2 Protein sidechains (i)

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	184/261 (70%)	171±3 (93±1%)	13±3 (7±1%)	18 66
All	All	3680/5220 (70%)	3416 (93%)	264 (7%)	18 66

5 of 90 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	146	LYS	19
1	A	172	MET	10
1	A	129	ARG	9
1	A	169	ARG	9
1	A	180	LYS	9

6.3.3 RNA (i)

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is

considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
1	PTR	A	207	1	15,16,17	0.88±0.02	1±0 (6±0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of angles that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
1	PTR	A	207	1	19,22,24	0.65±0.02	0±0 (0±0%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	PTR	A	207	1	-	0±0,10,11,13	0±0,1,1,1

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	207	PTR	P-OH	2.77	1.63	1.59	20	20

There are no bond-angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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 52% for the well-defined parts and 48% 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	1952
Number of shifts mapped to atoms	1952
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	16

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	270	0.13 ± 0.06	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	244	0.21 ± 0.20	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	234	0.18 ± 0.25	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 52%, i.e. 1524 atoms were assigned a chemical shift out of a possible 2953. 0 out of 33 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	714/1035 (69%)	333/420 (79%)	199/418 (48%)	182/197 (92%)
Sidechain	800/1653 (48%)	466/1069 (44%)	334/508 (66%)	0/76 (0%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	10/265 (4%)	10/131 (8%)	0/125 (0%)	0/9 (0%)
Overall	1524/2953 (52%)	809/1620 (50%)	533/1051 (51%)	182/282 (65%)

7.1.4 Statistically unusual chemical shifts (i)

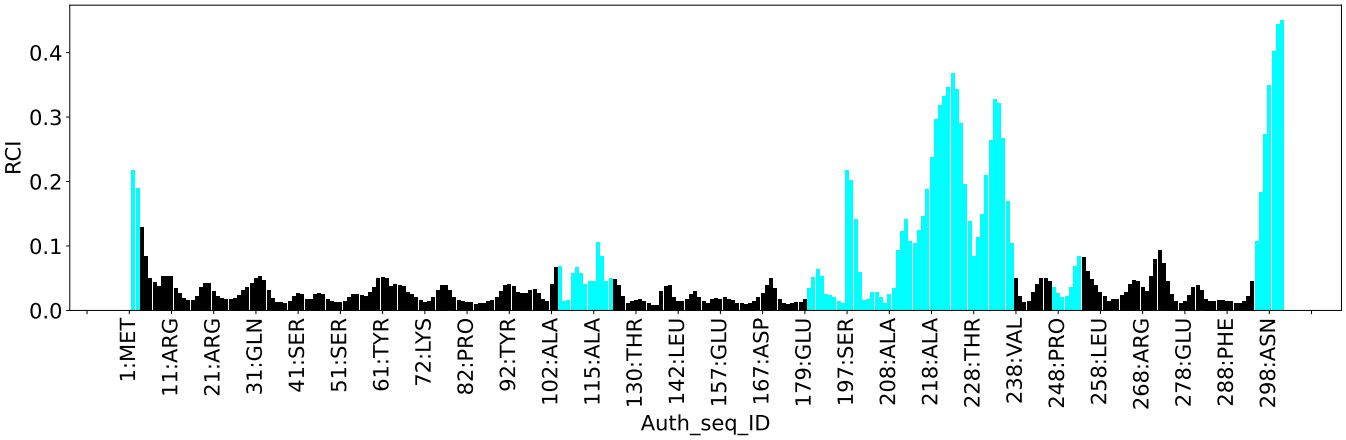
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	1	MET	HE1	10.20	-0.03 – 3.80	21.7
1	A	1	MET	HE2	10.20	-0.03 – 3.80	21.7
1	A	1	MET	HE3	10.20	-0.03 – 3.80	21.7
1	A	231	PRO	CD	42.81	45.11 – 55.58	-7.2
1	A	206	ALA	CA	40.20	43.52 – 62.81	-6.7
1	A	101	PRO	CA	53.31	55.85 – 70.84	-6.7
1	A	201	PRO	CA	53.89	55.85 – 70.84	-6.3
1	A	269	MET	HG3	0.14	0.54 – 4.26	-6.1
1	A	85	LEU	HD21	-0.74	-0.65 – 2.13	-5.3
1	A	85	LEU	HD22	-0.74	-0.65 – 2.13	-5.3
1	A	85	LEU	HD23	-0.74	-0.65 – 2.13	-5.3
1	A	84	LEU	HA	1.92	2.04 – 6.55	-5.3
1	A	21	ARG	HG2	0.24	0.26 – 2.87	-5.1
1	A	93	LEU	HD21	-0.67	-0.65 – 2.13	-5.1
1	A	93	LEU	HD22	-0.67	-0.65 – 2.13	-5.1
1	A	93	LEU	HD23	-0.67	-0.65 – 2.13	-5.1

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
0.0
0.1
0.2
0.3
0.4
1:MET
11:ARG
21:ARG
31:GLN
41:SER
51:SER
61:TYR
72:LYS
82:PRO
92:TYR
102:ALA
115:ALA
130:THR
142:LEU
157:GLU
167:ASP
179:GLU
197:SER
208:ALA
218:ALA
228:THR
238:VAL
248:PRO
258:LEU
268:ARG
278:GLU
288:PHE
298:ASN
Auth_seq_ID

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	3525
Intra-residue (|i-j|=0)	711
Sequential (|i-j|=1)	968
Medium range (|i-j|>1 and |i-j|<5)	427
Long range (|i-j|≥5)	1418
Inter-chain	0
Hydrogen bond restraints	1
Disulfide bond restraints	0
Total dihedral-angle restraints	176
Number of unmapped restraints	0
Number of restraints per residue	12.2
Number of long range restraints per residue[1]	4.7

[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)	13.7	0.2
0.2-0.5 (Medium)	12.4	0.5
>0.5 (Large)	0.6	0.74

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)	14.1	9.3
10.0-20.0 (Medium)	0.1	13.0
>20.0 (Large)	None	None

9 Distance violation analysis (i)

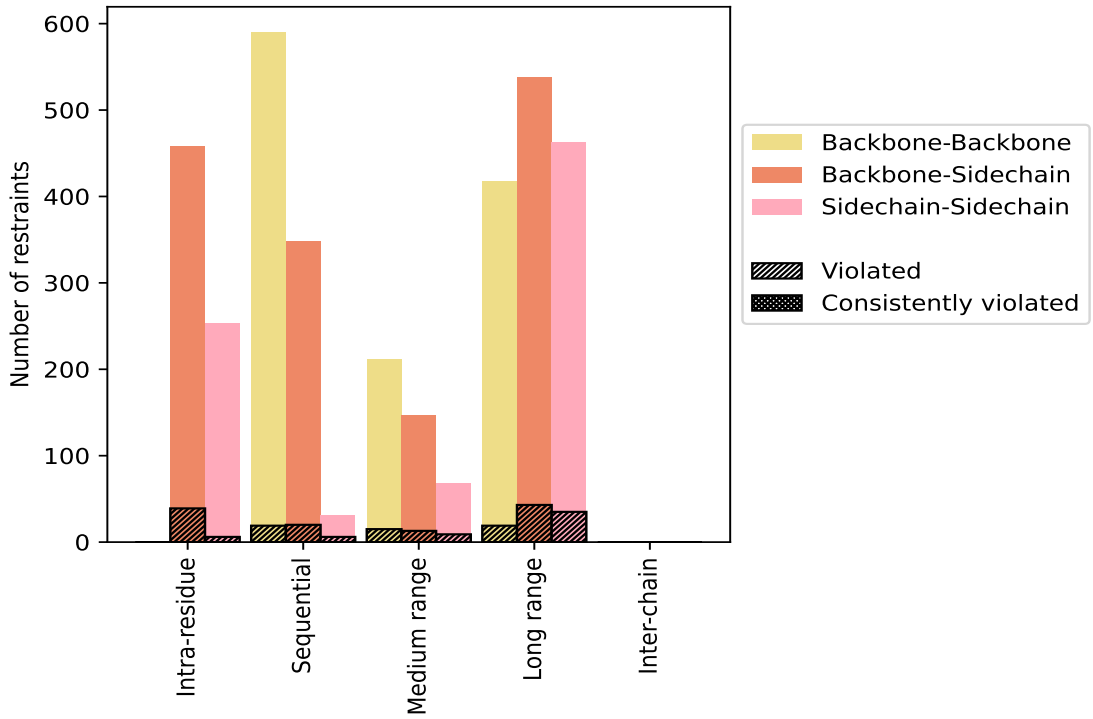
9.1 Summary of distance violations (i)

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)	711	20.2	45	6.3	1.3	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	458	13.0	39	8.5	1.1	0	0.0	0.0
Sidechain-Sidechain	253	7.2	6	2.4	0.2	0	0.0	0.0
Sequential (|i-j|=1)	968	27.5	45	4.6	1.3	0	0.0	0.0
Backbone-Backbone	589	16.7	19	3.2	0.5	0	0.0	0.0
Backbone-Sidechain	348	9.9	20	5.7	0.6	0	0.0	0.0
Sidechain-Sidechain	31	0.9	6	19.4	0.2	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	427	12.1	37	8.7	1.0	0	0.0	0.0
Backbone-Backbone	212	6.0	15	7.1	0.4	0	0.0	0.0
Backbone-Sidechain	147	4.2	13	8.8	0.4	0	0.0	0.0
Sidechain-Sidechain	68	1.9	9	13.2	0.3	0	0.0	0.0
Long range (|i-j|≥5)	1418	40.2	97	6.8	2.8	0	0.0	0.0
Backbone-Backbone	418	11.9	19	4.5	0.5	0	0.0	0.0
Backbone-Sidechain	538	15.3	43	8.0	1.2	0	0.0	0.0
Sidechain-Sidechain	462	13.1	35	7.6	1.0	0	0.0	0.0
Inter-chain	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Hydrogen bond	1	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	3525	100.0	224	6.4	6.4	0	0.0	0.0
Backbone-Backbone	1220	34.6	53	4.3	1.5	0	0.0	0.0
Backbone-Sidechain	1491	42.3	115	7.7	3.3	0	0.0	0.0
Sidechain-Sidechain	814	23.1	56	6.9	1.6	0	0.0	0.0

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

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



Violated and consistently violated restraints are shown using different hatch patterns in their respective categories. The hydrogen bonds and 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	9	2	3	6	0	20	0.22	0.42	0.1	0.18
2	10	5	5	11	0	31	0.23	0.54	0.11	0.23
3	6	9	3	7	0	25	0.27	0.73	0.15	0.22
4	7	7	6	10	0	30	0.19	0.48	0.1	0.16
5	13	3	3	10	0	29	0.23	0.46	0.12	0.17
6	8	7	8	12	0	35	0.25	0.51	0.12	0.21
7	8	7	7	10	0	32	0.21	0.4	0.1	0.16
8	8	7	5	13	0	33	0.27	0.57	0.12	0.27
9	13	7	3	19	0	42	0.26	0.5	0.11	0.24
10	3	7	2	7	0	19	0.26	0.49	0.13	0.2
11	11	11	10	12	0	44	0.23	0.53	0.12	0.18

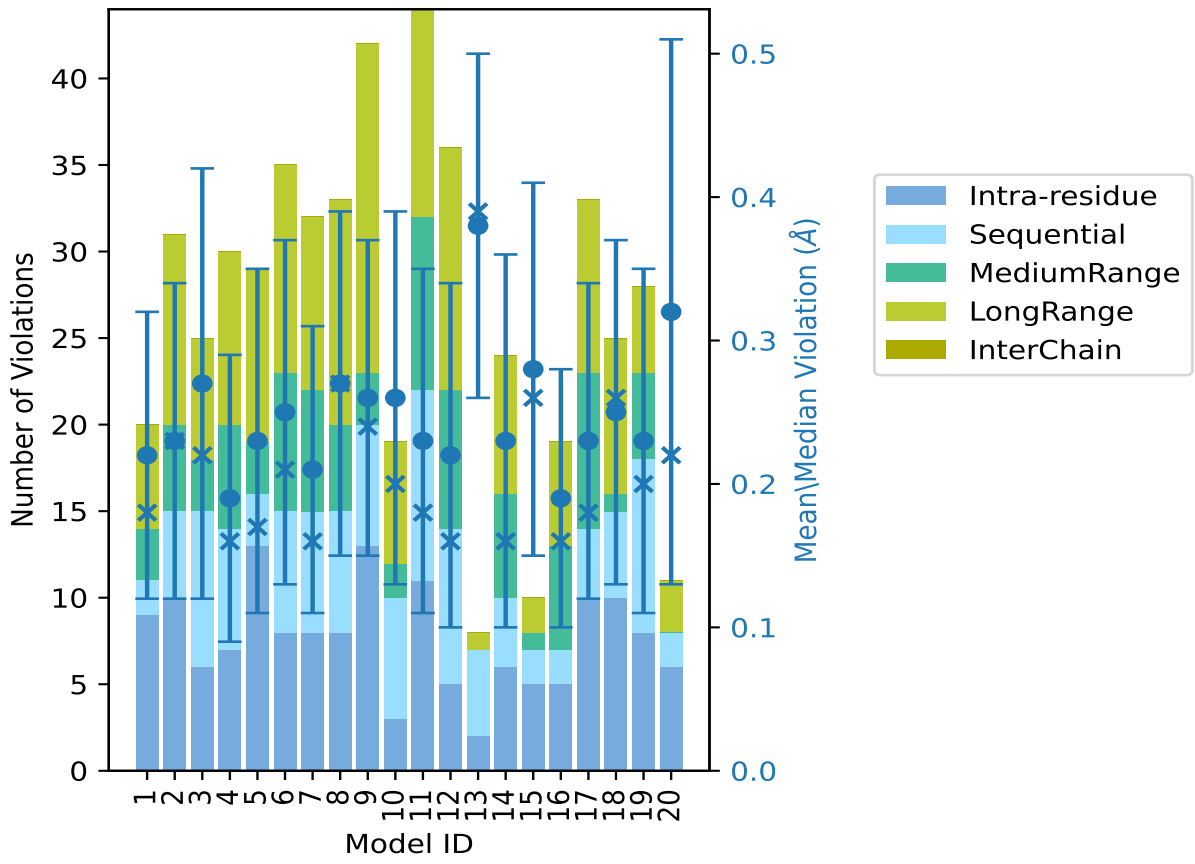
Continued on next page...

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	5	9	8	14	0	36	0.22	0.5	0.12	0.16
13	2	5	0	1	0	8	0.38	0.55	0.12	0.39
14	6	4	6	8	0	24	0.23	0.5	0.13	0.16
15	5	2	1	2	0	10	0.28	0.51	0.13	0.26
16	5	2	6	6	0	19	0.19	0.4	0.09	0.16
17	10	4	9	10	0	33	0.23	0.5	0.11	0.18
18	10	5	1	9	0	25	0.25	0.49	0.12	0.26
19	8	10	5	5	0	28	0.23	0.47	0.12	0.2
20	6	2	0	3	0	11	0.32	0.74	0.19	0.22

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

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



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

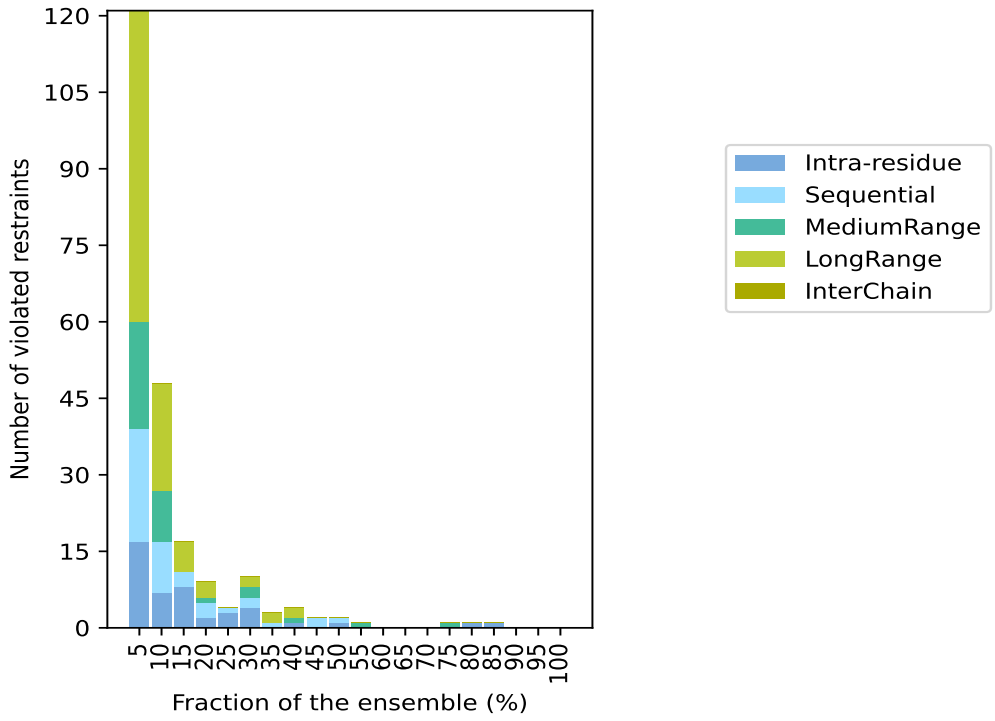
9.3 Distance violation statistics for the ensemble (i)

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, 3300(IR:666, SQ:923, MR:390, LR:1321, 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]	%
17	22	21	61	0	121	1	5.0
7	10	10	21	0	48	2	10.0
8	3	0	6	0	17	3	15.0
2	3	1	3	0	9	4	20.0
3	1	0	0	0	4	5	25.0
4	2	2	2	0	10	6	30.0
0	1	0	2	0	3	7	35.0
1	0	1	2	0	4	8	40.0
0	2	0	0	0	2	9	45.0
1	1	0	0	0	2	10	50.0
0	0	1	0	0	1	11	55.0
0	0	0	0	0	0	12	60.0
0	0	0	0	0	0	13	65.0
0	0	0	0	0	0	14	70.0
0	0	1	0	0	1	15	75.0
1	0	0	0	0	1	16	80.0
1	0	0	0	0	1	17	85.0
0	0	0	0	0	0	18	90.0
0	0	0	0	0	0	19	95.0
0	0	0	0	0	0	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

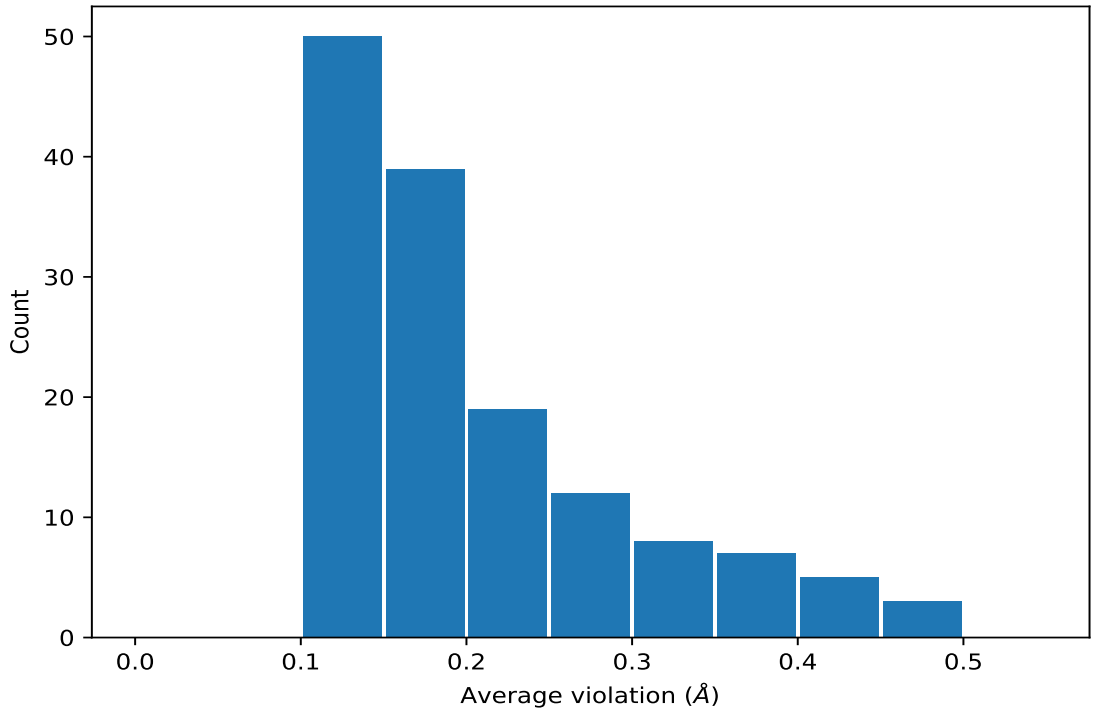
9.3.1 Bar graph : Distance violation statistics for the ensemble (i)



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.

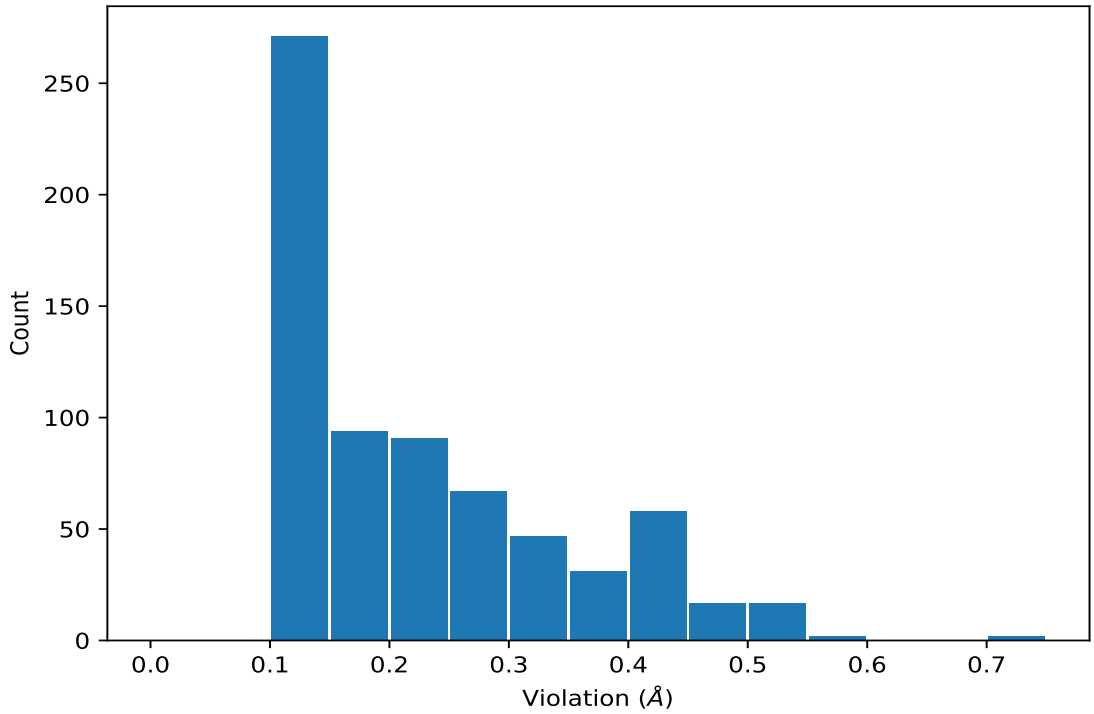
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,2606)	1:A:187:HIS:HB2	1:A:187:HIS:HD2	17	0.43	0.08	0.44
(1,1636)	1:A:132:TYR:HB2	1:A:132:TYR:HD1	16	0.42	0.02	0.43
(1,1200)	1:A:87:PHE:HZ	1:A:91:HIS:HE1	15	0.16	0.04	0.15
(1,249)	1:A:20:SER:H	1:A:23:GLU:H	11	0.21	0.04	0.21
(1,2631)	1:A:210:PRO:HA	1:A:211:GLN:H	10	0.26	0.07	0.28
(1,1631)	1:A:132:TYR:HA	1:A:132:TYR:HD1	10	0.21	0.05	0.24
(1,2638)	1:A:221:GLY:HA2	1:A:222:SER:H	9	0.36	0.14	0.4
(1,2633)	1:A:211:GLN:HB2	1:A:212:THR:HA	9	0.35	0.14	0.42
(1,2630)	1:A:207:PTR:HA	1:A:207:PTR:HD1	8	0.33	0.11	0.34
(1,458)	1:A:33:HIS:HE2	1:A:56:SER:HA	8	0.18	0.06	0.17

[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)
(1,1880)	1:A:144:PHE:HD1	1:A:145:LYS:H	20	0.74
(1,2365)	1:A:169:ARG:H	1:A:169:ARG:HB2	3	0.73
(1,3270)	1:A:270:ASN:HD21	1:A:275:TRP:HD1	8	0.57
(1,2633)	1:A:211:GLN:HB2	1:A:212:THR:HA	13	0.55
(1,2365)	1:A:169:ARG:H	1:A:169:ARG:HB2	2	0.54
(1,2094)	1:A:155:LYS:HG2	1:A:156:PRO:HD2	3	0.54
(1,1560)	1:A:130:THR:H	1:A:144:PHE:HD1	20	0.54
(1,2458)	1:A:174:PRO:HB2	1:A:177:TYR:HB2	11	0.53
(1,2458)	1:A:174:PRO:HB3	1:A:177:TYR:HB2	11	0.53
(1,2640)	1:A:223:PRO:HA	1:A:224:GLY:H	6	0.51
(1,2606)	1:A:187:HIS:HB2	1:A:187:HIS:HD2	3	0.51

10 Dihedral-angle violation analysis (i)

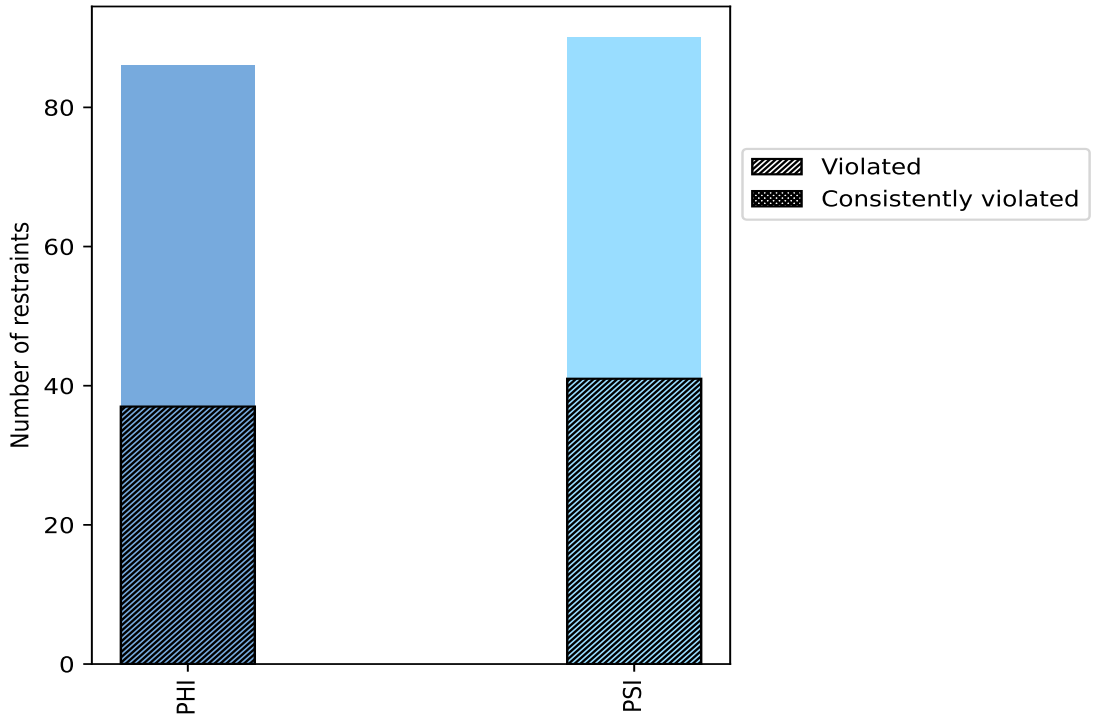
10.1 Summary of dihedral-angle violations (i)

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]
PHI	86	48.9	37	43.0	21.0	0	0.0	0.0
PSI	90	51.1	41	45.6	23.3	0	0.0	0.0
Total	176	100.0	78	44.3	44.3	0	0.0	0.0

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

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



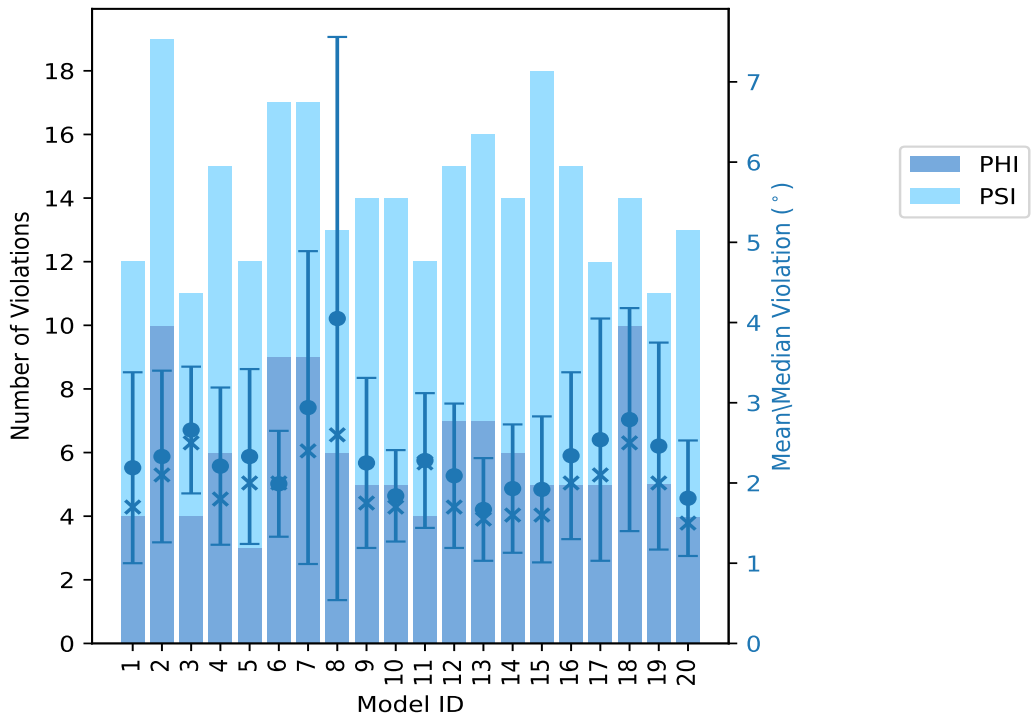
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

10.2 Dihedral-angle violation statistics for each model (i)

The following table provides the dihedral-angle violation statistics for each model in the ensemble. Violations less than 1° are not included in the statistics.

Model ID	Number of violations			Mean (°)	Max (°)	SD (°)	Median (°)
	PHI	PSI	Total				
1	4	8	12	2.19	5.0	1.19	1.7
2	10	9	19	2.33	5.4	1.07	2.1
3	4	7	11	2.66	3.7	0.79	2.5
4	6	9	15	2.21	4.0	0.98	1.8
5	3	9	12	2.33	4.6	1.09	2.0
6	9	8	17	1.99	3.4	0.66	2.0
7	9	8	17	2.94	9.3	1.95	2.4
8	6	7	13	4.05	13.0	3.51	2.6
9	5	9	14	2.25	4.2	1.06	1.75
10	5	9	14	1.84	3.5	0.57	1.7
11	4	8	12	2.28	4.2	0.84	2.25
12	7	8	15	2.09	4.1	0.9	1.7
13	7	9	16	1.67	3.9	0.64	1.55
14	6	8	14	1.93	4.1	0.8	1.6
15	5	13	18	1.92	4.6	0.91	1.6
16	5	10	15	2.34	5.2	1.04	2.0
17	5	7	12	2.54	6.3	1.51	2.1
18	10	4	14	2.79	6.2	1.39	2.5
19	5	6	11	2.46	5.1	1.29	2.0
20	4	9	13	1.81	3.6	0.72	1.5

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	
PHI	PSI	Total	Count[1]	%
21	13	34	1	5.0
2	7	9	2	10.0
3	8	11	3	15.0
2	1	3	4	20.0
2	2	4	5	25.0
2	2	4	6	30.0
2	2	4	7	35.0
0	1	1	8	40.0
1	1	2	9	45.0
0	0	0	10	50.0
0	1	1	11	55.0

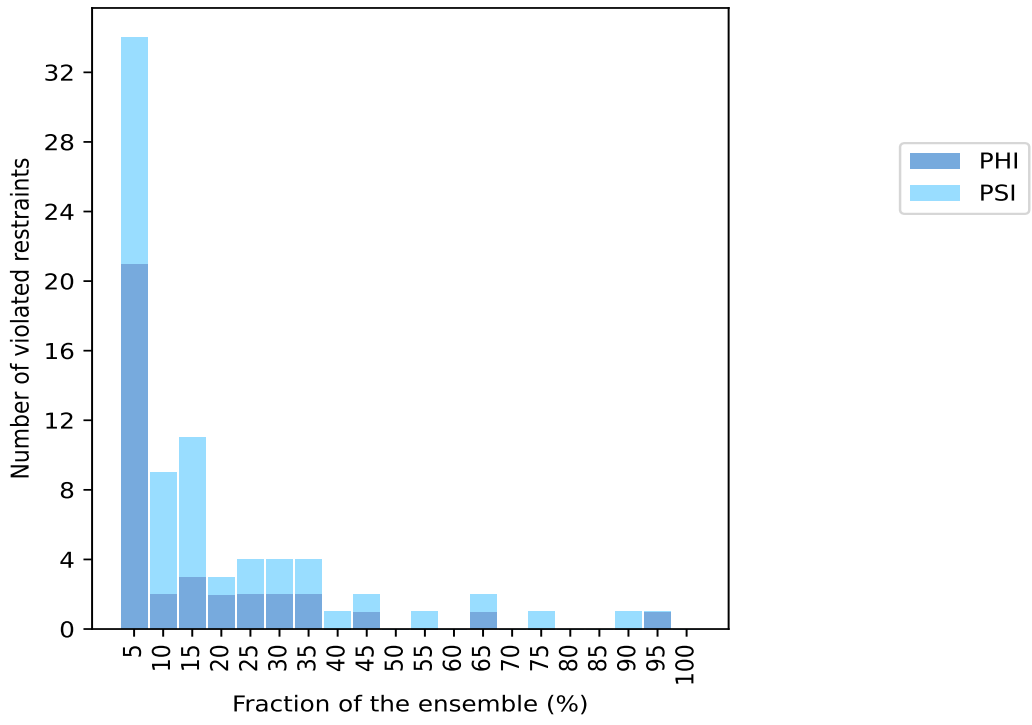
Continued on next page...

Continued from previous page...

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

[1] Number of models with violations

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

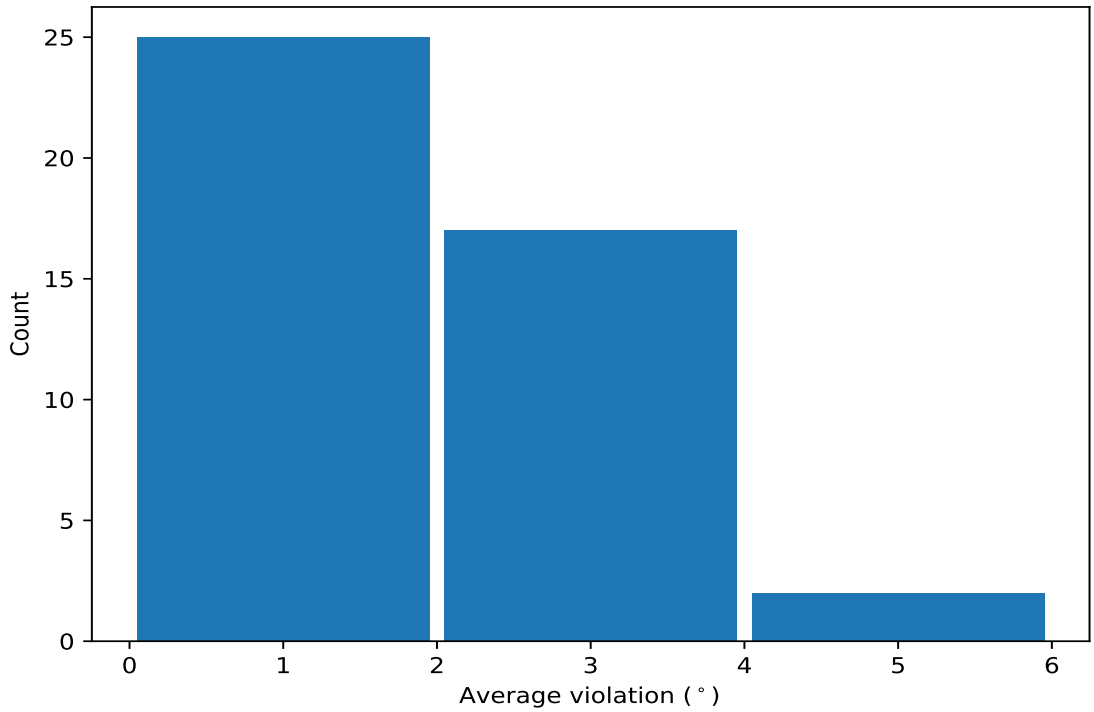


10.4 Most violated dihedral-angle restraints in the ensemble (i)

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.

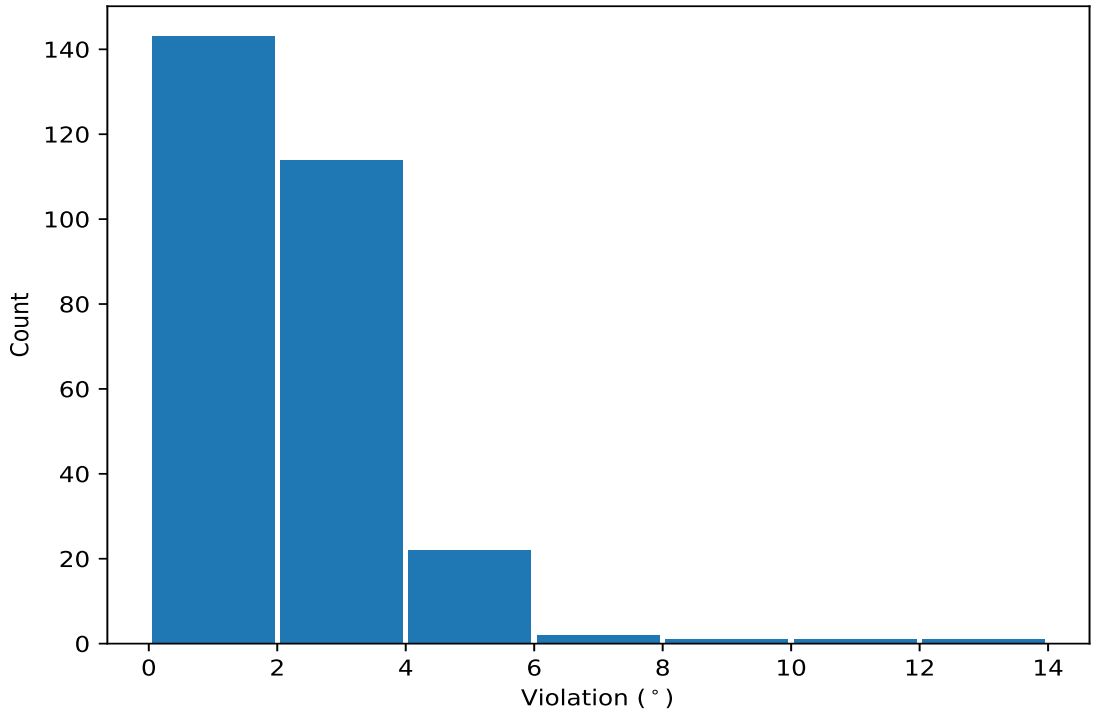
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models[1]	Mean	SD[2]	Median
(1,12)	1:A:16:MET:C	1:A:17:GLY:N	1:A:17:GLY:CA	1:A:17:GLY:C	19	2.44	0.44	2.5
(1,13)	1:A:17:GLY:N	1:A:17:GLY:CA	1:A:17:GLY:C	1:A:18:PRO:N	18	2.94	1.03	2.8
(1,108)	1:A:154:GLU:N	1:A:154:GLU:CA	1:A:154:GLU:C	1:A:155:LYS:N	15	3.51	1.31	3.9
(1,19)	1:A:23:GLU:N	1:A:23:GLU:CA	1:A:23:GLU:C	1:A:24:ALA:N	13	2.55	0.91	2.4
(1,45)	1:A:50:LEU:C	1:A:51:SER:N	1:A:51:SER:CA	1:A:51:SER:C	13	1.96	0.49	2.0
(1,54)	1:A:59:SER:N	1:A:59:SER:CA	1:A:59:SER:C	1:A:60:HIS:N	11	3.12	0.92	3.5
(1,62)	1:A:63:ILE:N	1:A:63:ILE:CA	1:A:63:ILE:C	1:A:64:ASN:N	9	4.03	1.65	3.6
(1,51)	1:A:56:SER:C	1:A:57:ARG:N	1:A:57:ARG:CA	1:A:57:ARG:C	9	1.73	0.31	1.7
(1,87)	1:A:87:PHE:N	1:A:87:PHE:CA	1:A:87:PHE:C	1:A:88:TYR:N	8	1.5	0.28	1.5
(1,106)	1:A:153:ILE:N	1:A:153:ILE:CA	1:A:153:ILE:C	1:A:154:GLU:N	7	2.09	0.93	1.8

[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 (°)
(1,156)	1:A:275:TRP:N	1:A:275:TRP:CA	1:A:275:TRP:C	1:A:276:GLU:N	8	13.0
(1,155)	1:A:274:GLN:C	1:A:275:TRP:N	1:A:275:TRP:CA	1:A:275:TRP:C	8	10.3
(1,133)	1:A:238:VAL:C	1:A:239:PHE:N	1:A:239:PHE:CA	1:A:239:PHE:C	7	9.3
(1,62)	1:A:63:ILE:N	1:A:63:ILE:CA	1:A:63:ILE:C	1:A:64:ASN:N	17	6.3
(1,62)	1:A:63:ILE:N	1:A:63:ILE:CA	1:A:63:ILE:C	1:A:64:ASN:N	18	6.2
(1,133)	1:A:238:VAL:C	1:A:239:PHE:N	1:A:239:PHE:CA	1:A:239:PHE:C	2	5.4
(1,108)	1:A:154:GLU:N	1:A:154:GLU:CA	1:A:154:GLU:C	1:A:155:LYS:N	8	5.2
(1,108)	1:A:154:GLU:N	1:A:154:GLU:CA	1:A:154:GLU:C	1:A:155:LYS:N	16	5.2
(1,62)	1:A:63:ILE:N	1:A:63:ILE:CA	1:A:63:ILE:C	1:A:64:ASN:N	19	5.1
(1,62)	1:A:63:ILE:N	1:A:63:ILE:CA	1:A:63:ILE:C	1:A:64:ASN:N	1	5.0