wwPDB NMR Structure Validation Summary Report (i)

Jun 3, 2023 – 04:06 PM EDT

PDB ID	:	6X4X
BMRB ID	:	30755
Title	:	B24Y DKP insulin
Authors	:	Weiss, M.A.; Yang, Y.
Deposited on	:	2020-05-24

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:

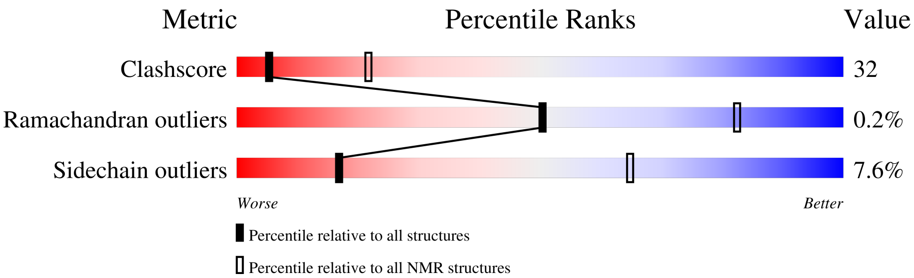
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal 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 81%.

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	21	29% 62% 5% 5%
2	B	30	43% 43% 13%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 11 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *fewest violations*.

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:2-A:21, B:22-B:47 (46)	0.20	11

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. No single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 783 atoms, of which 379 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Insulin chain A.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	21	312	99	149	25	35	4	0

- Molecule 2 is a protein called Insulin.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	30	471	156	230	38	45	2	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	31	ASP	HIS	variant	UNP P01308
B	45	TYR	PHE	variant	UNP P01308
B	49	LYS	PRO	variant	UNP P01308
B	50	PRO	LYS	variant	UNP P01308

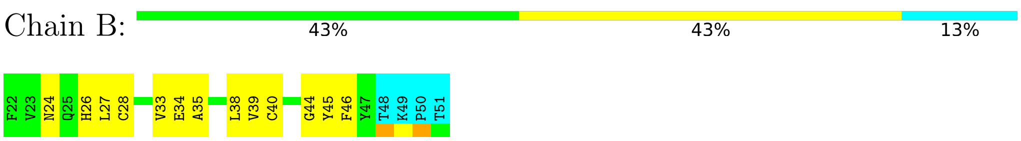
4 Residue-property plots (i)

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: Insulin chain A

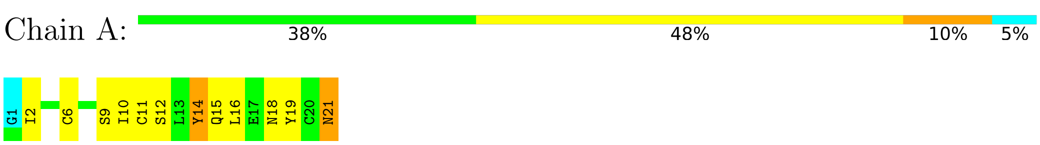
- Molecule 2: Insulin



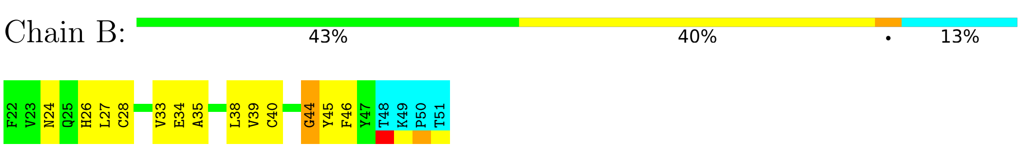
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 11. Colouring as in section 4.1 above.

- Molecule 1: Insulin chain A



- Molecule 2: Insulin



5 Refinement protocol and experimental data overview (i)

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

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *structures with the least restraint violations*.

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

Software name	Classification	Version
X-PLOR NIH	refinement	
X-PLOR NIH	structure calculation	

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	525
Number of shifts mapped to atoms	525
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	81%

6 Model quality (i)

6.1 Standard geometry (i)

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

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

6.2 Too-close contacts (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	159	144	144	14±2
2	B	210	196	192	13±2
All	All	7380	6800	6720	445

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 32.

5 of 69 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:10:ILE:HD11	2:B:26:HIS:CE1	0.89	2.03	5	7
1:A:2:ILE:HD12	1:A:19:TYR:CD2	0.80	2.11	19	20
1:A:16:LEU:HD23	2:B:39:VAL:HG21	0.68	1.64	14	20
1:A:6:CYS:N	1:A:11:CYS:SG	0.61	2.74	4	11
1:A:16:LEU:HD21	2:B:35:ALA:HB1	0.61	1.71	1	16

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	19/21 (90%)	16±1 (85±7%)	3±1 (15±7%)	0±0 (0±0%)	100	100
2	B	25/30 (83%)	22±1 (89±4%)	3±1 (11±4%)	0±0 (0±1%)	38	78
All	All	880/1020 (86%)	767 (87%)	111 (13%)	2 (0%)	50	82

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
2	B	44	GLY	2

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	20/20 (100%)	18±1 (92±3%)	2±1 (8±3%)	16	63
2	B	22/26 (85%)	20±1 (93±3%)	2±1 (7±3%)	18	66
All	All	840/920 (91%)	776 (92%)	64 (8%)	17	65

5 of 10 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	14	TYR	20
2	B	24	ASN	20
1	A	11	CYS	6
1	A	21	ASN	5
2	B	46	PHE	5

6.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

6.5 Carbohydrates (i)

There are no monosaccharides in this entry.

6.6 Ligand geometry (i)

There are no ligands in this entry.

6.7 Other polymers (i)

There are no such molecules in this entry.

6.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

7 Chemical shift validation (i)

The completeness of assignment taking into account all chemical shift lists is 81% for the well-defined parts and 78% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *starch_ output*

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	525
Number of shifts mapped to atoms	525
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

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	42	0.28 ± 0.22	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	38	0.19 ± 0.15	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	35	0.86 ± 0.74	None needed (imprecise)

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 81%, i.e. 493 atoms were assigned a chemical shift out of a possible 612. 0 out of 10 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	164/233 (70%)	88/95 (93%)	41/92 (45%)	35/46 (76%)
Sidechain	283/306 (92%)	196/200 (98%)	80/97 (82%)	7/9 (78%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	46/73 (63%)	29/34 (85%)	17/37 (46%)	0/2 (0%)
Overall	493/612 (81%)	313/329 (95%)	138/226 (61%)	42/57 (74%)

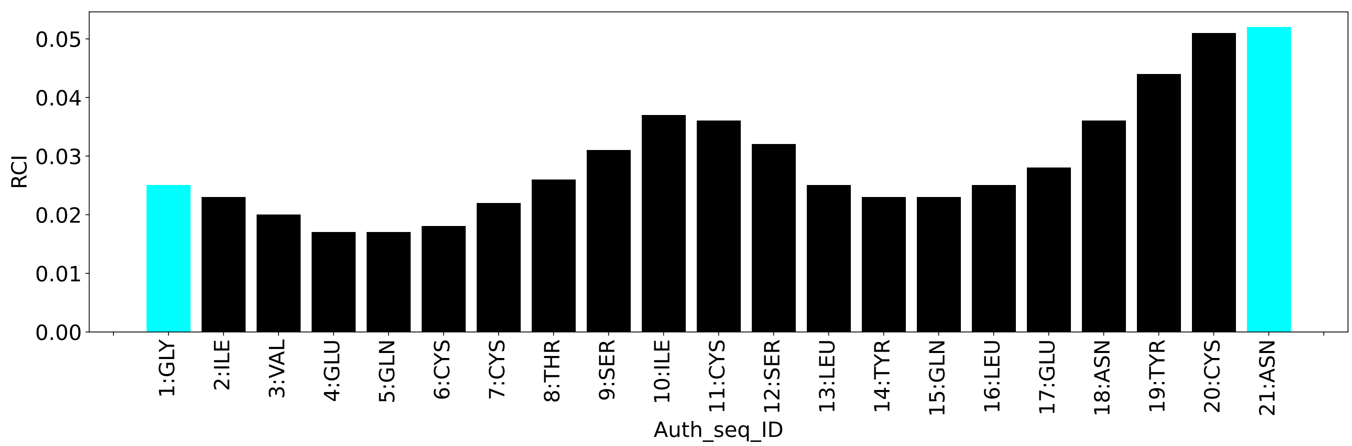
7.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

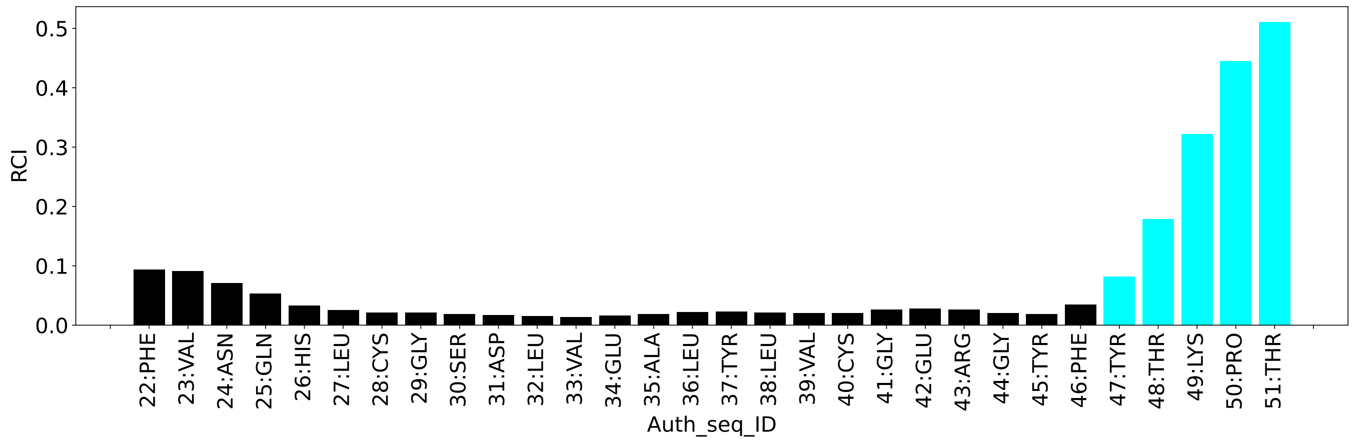
7.1.5 Random Coil Index (RCI) plots (i)

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:



Random coil index (RCI) for chain B:


RCI
0.0
0.1
0.2
0.3
0.4
0.5
22:PHE
23:VAL
24:ASN
25:GLN
26:HIS
27:LEU
28:CYS
29:GLY
30:SER
31:ASP
32:LEU
33:VAL
34:GLU
35:ALA
36:LEU
37:TYR
38:LEU
39:VAL
40:CYS
41:GLY
42:GLU
43:ARG
44:GLY
45:TYR
46:PHE
47:TYR
48:THR
49:LYS
50:PRO
51:THR
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	925
Intra-residue (|i-j|=0)	338
Sequential (|i-j|=1)	229
Medium range (|i-j|>1 and |i-j|<5)	147
Long range (|i-j|≥5)	73
Inter-chain	122
Hydrogen bond restraints	16
Disulfide bond restraints	0
Total dihedral-angle restraints	81
Number of unmapped restraints	0
Number of restraints per residue	19.7
Number of long range restraints per residue[1]	1.4

[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)	30.4	0.2
0.2-0.5 (Medium)	7.0	0.5
>0.5 (Large)	1.0	0.75

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)	5.0	3.4
10.0-20.0 (Medium)	None	None
>20.0 (Large)	None	None

9 Distance violation analysis (i)

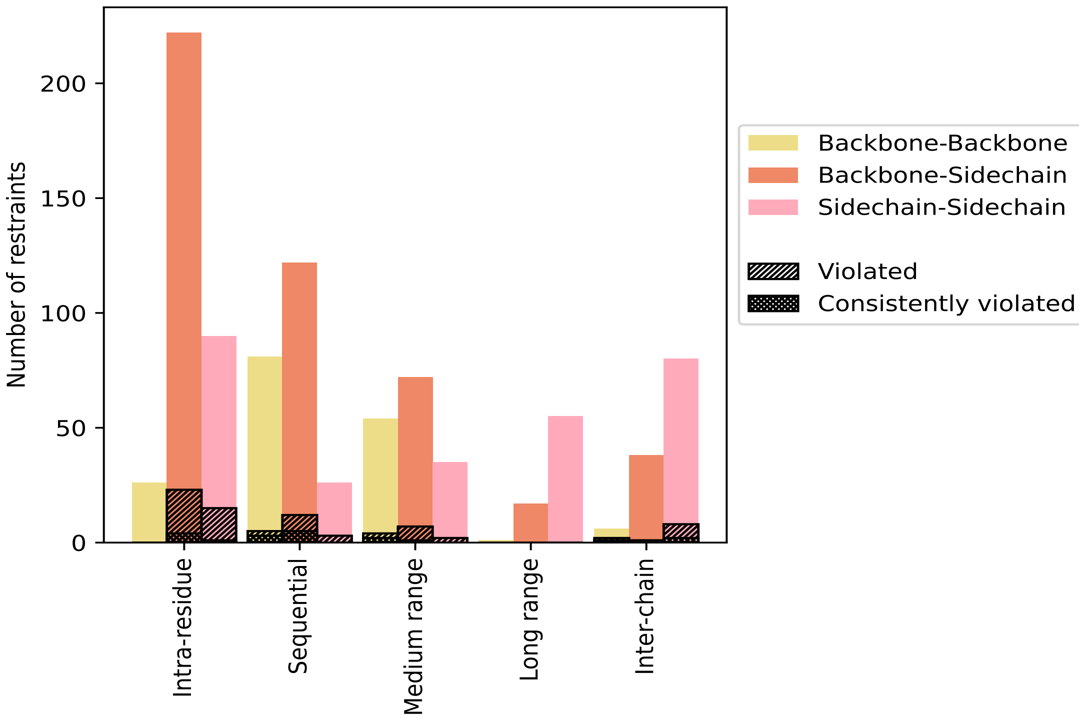
9.1 Summary of distance violations (i)

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)	338	36.5	38	11.2	4.1	5	1.5	0.5
Backbone-Backbone	26	2.8	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	222	24.0	23	10.4	2.5	4	1.8	0.4
Sidechain-Sidechain	90	9.7	15	16.7	1.6	1	1.1	0.1
Sequential (|i-j|=1)	229	24.8	20	8.7	2.2	8	3.5	0.9
Backbone-Backbone	81	8.8	5	6.2	0.5	3	3.7	0.3
Backbone-Sidechain	122	13.2	12	9.8	1.3	5	4.1	0.5
Sidechain-Sidechain	26	2.8	3	11.5	0.3	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	147	15.9	12	8.2	1.3	2	1.4	0.2
Backbone-Backbone	40	4.3	3	7.5	0.3	1	2.5	0.1
Backbone-Sidechain	72	7.8	7	9.7	0.8	1	1.4	0.1
Sidechain-Sidechain	35	3.8	2	5.7	0.2	0	0.0	0.0
Long range (|i-j|≥5)	73	7.9	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	1	0.1	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	17	1.8	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	55	5.9	0	0.0	0.0	0	0.0	0.0
Inter-chain	122	13.2	10	8.2	1.1	3	2.5	0.3
Backbone-Backbone	4	0.4	1	25.0	0.1	1	25.0	0.1
Backbone-Sidechain	38	4.1	1	2.6	0.1	0	0.0	0.0
Sidechain-Sidechain	80	8.6	8	10.0	0.9	2	2.5	0.2
Hydrogen bond	16	1.7	2	12.5	0.2	1	6.2	0.1
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	925	100.0	82	8.9	8.9	19	2.1	2.1
Backbone-Backbone	168	18.2	11	6.5	1.2	6	3.6	0.6
Backbone-Sidechain	471	50.9	43	9.1	4.6	10	2.1	1.1
Sidechain-Sidechain	286	30.9	28	9.8	3.0	3	1.0	0.3

[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	12	11	7	0	6	36	0.18	0.7	0.1	0.16
2	14	12	7	0	5	38	0.18	0.63	0.09	0.15
3	15	13	8	0	6	42	0.17	0.65	0.09	0.16
4	16	11	9	0	6	42	0.18	0.64	0.08	0.16
5	16	12	6	0	8	42	0.18	0.63	0.08	0.17
6	18	10	6	0	7	41	0.18	0.63	0.08	0.16
7	16	11	8	0	7	42	0.17	0.67	0.09	0.16
8	13	14	7	0	7	41	0.17	0.64	0.08	0.16
9	11	13	5	0	4	33	0.19	0.63	0.09	0.18
10	14	11	6	0	6	37	0.17	0.63	0.08	0.16
11	11	12	6	0	4	33	0.17	0.64	0.09	0.15

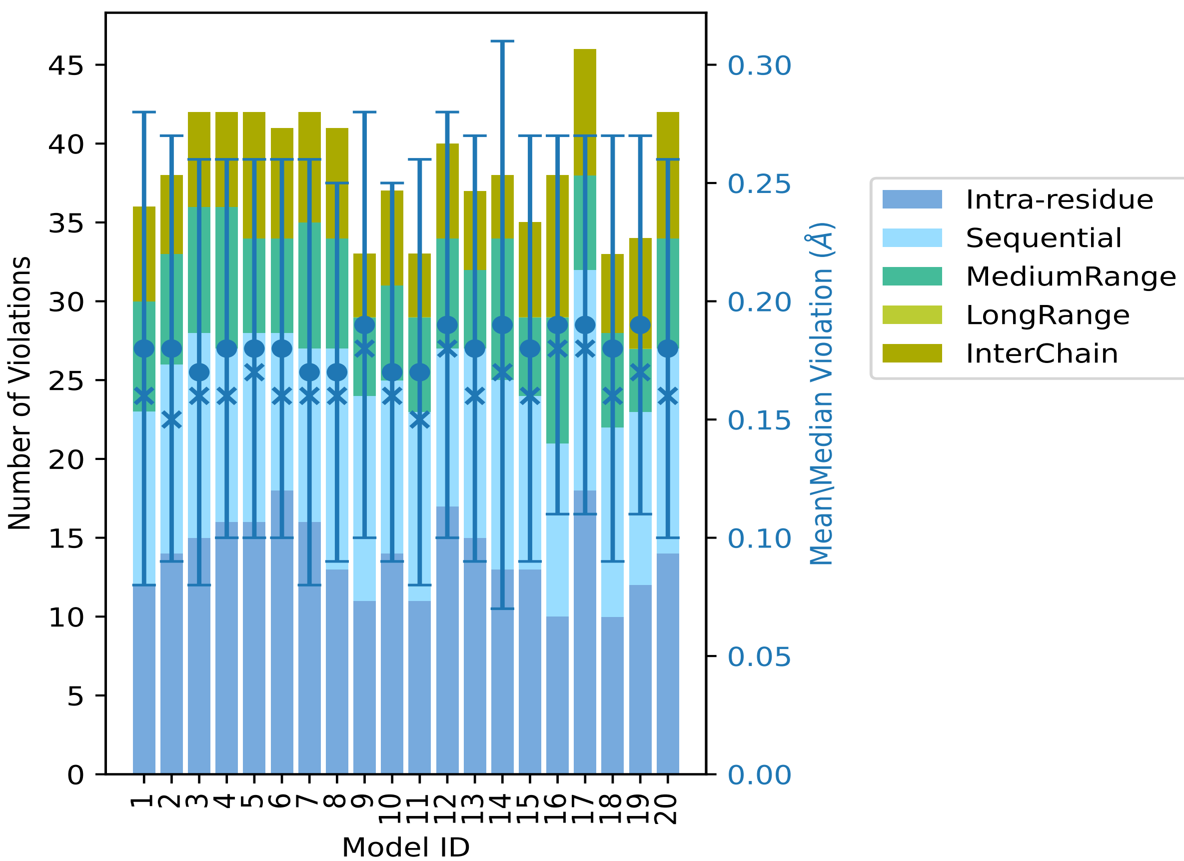
Continued on next page...

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	17	10	7	0	6	40	0.19	0.66	0.09	0.18
13	15	12	5	0	5	37	0.18	0.64	0.09	0.16
14	13	12	9	0	4	38	0.19	0.75	0.12	0.17
15	13	11	5	0	6	35	0.18	0.65	0.09	0.16
16	10	11	8	0	9	38	0.19	0.61	0.08	0.18
17	18	14	6	0	8	46	0.19	0.65	0.08	0.18
18	10	12	6	0	5	33	0.18	0.62	0.09	0.16
19	12	11	4	0	7	34	0.19	0.58	0.08	0.17
20	14	13	7	0	8	42	0.18	0.65	0.08	0.16

[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

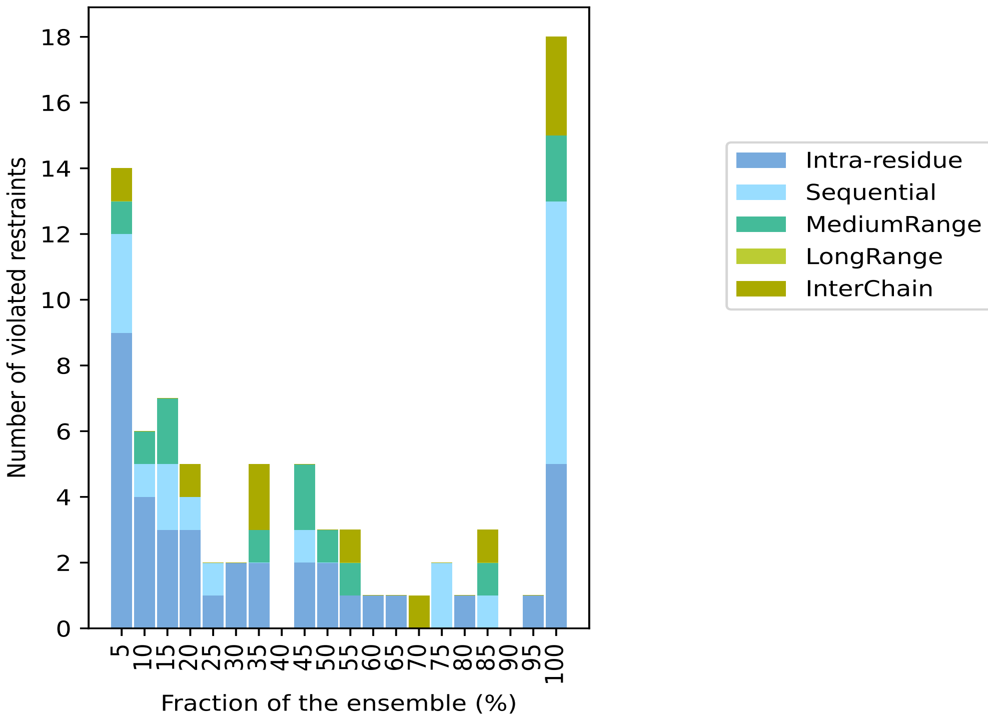
9.3 Distance violation statistics for the ensemble (i)

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, 829(IR:300, SQ:209, MR:135, LR:73, IC:112) 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]	%
9	3	1	0	1	14	1	5.0
4	1	1	0	0	6	2	10.0
3	2	2	0	0	7	3	15.0
3	1	0	0	1	5	4	20.0
1	1	0	0	0	2	5	25.0
2	0	0	0	0	2	6	30.0
2	0	1	0	2	5	7	35.0
0	0	0	0	0	0	8	40.0
2	1	2	0	0	5	9	45.0
2	0	1	0	0	3	10	50.0
1	0	1	0	1	3	11	55.0
1	0	0	0	0	1	12	60.0
1	0	0	0	0	1	13	65.0
0	0	0	0	1	1	14	70.0
0	2	0	0	0	2	15	75.0
1	0	0	0	0	1	16	80.0
0	1	1	0	1	3	17	85.0
0	0	0	0	0	0	18	90.0
1	0	0	0	0	1	19	95.0
5	8	2	0	3	18	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

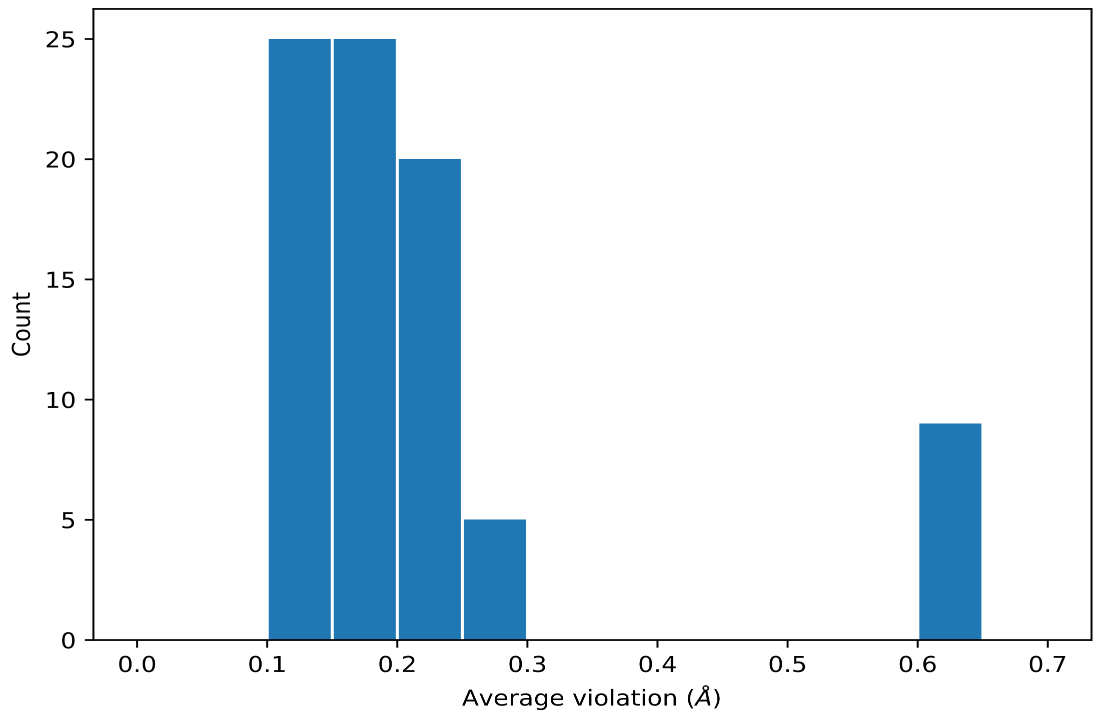
9.3.1 Bar graph : Distance violation statistics for the ensemble (i)



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)
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD11	20	0.64	0.03	0.64
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD12	20	0.64	0.03	0.64
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD13	20	0.64	0.03	0.64
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD11	20	0.64	0.03	0.64
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD12	20	0.64	0.03	0.64
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD13	20	0.64	0.03	0.64
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD11	20	0.64	0.03	0.64
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD12	20	0.64	0.03	0.64
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD13	20	0.64	0.03	0.64
(1,274)	1:A:14:TYR:HD2	1:A:15:GLN:HA	20	0.22	0.01	0.22
(1,825)	2:B:43:ARG:HA	2:B:44:GLY:H	20	0.2	0.03	0.2
(1,847)	2:B:45:TYR:HB2	2:B:45:TYR:HD2	20	0.2	0.0	0.2
(1,127)	1:A:7:CYS:HB2	1:A:8:THR:H	20	0.2	0.03	0.19
(1,366)	1:A:17:GLU:HG3	1:A:18:ASN:H	20	0.19	0.01	0.19
(1,268)	1:A:14:TYR:HD1	1:A:14:TYR:H	20	0.19	0.01	0.19
(1,142)	1:A:8:THR:H	1:A:9:SER:HB2	20	0.19	0.03	0.18

Continued on next page...

Continued from previous page...

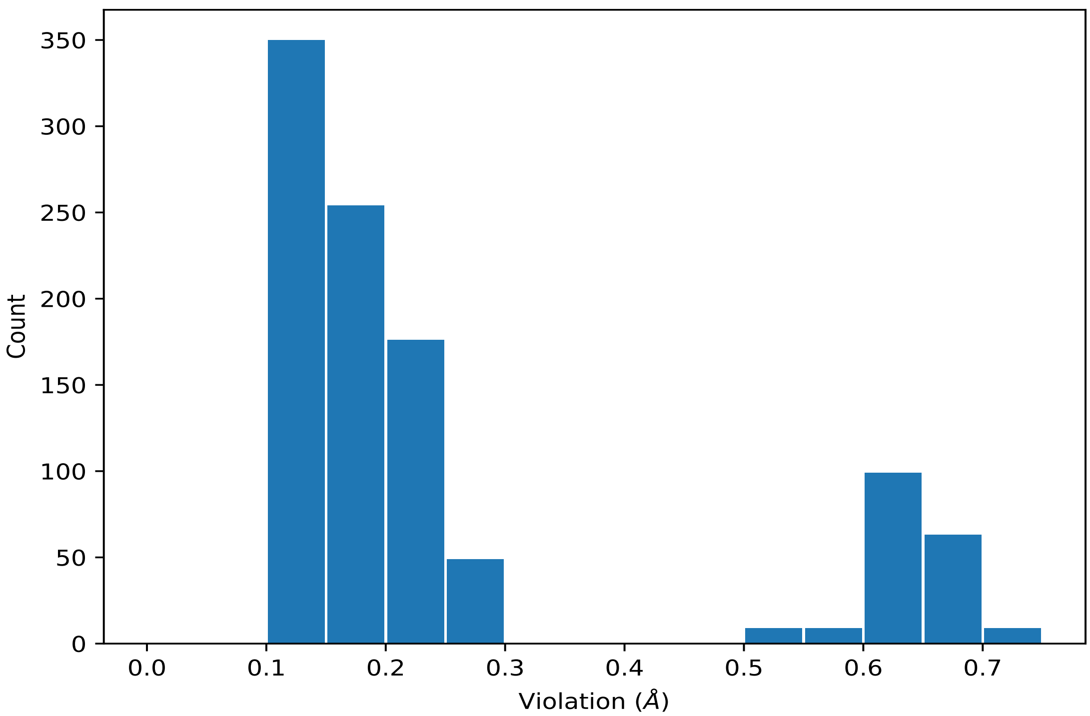
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,530)	2:B:27:LEU:HB2	2:B:28:CYS:H	20	0.19	0.01	0.19
(1,435)	1:A:20:CYS:HB2	2:B:45:TYR:HB2	20	0.18	0.02	0.18

[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)
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD11	14	0.75
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD12	14	0.75
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD13	14	0.75

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD11	14	0.75
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD12	14	0.75
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD13	14	0.75
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD11	14	0.75
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD12	14	0.75
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD13	14	0.75
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD11	1	0.7
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD12	1	0.7
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD13	1	0.7
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD11	1	0.7
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD12	1	0.7
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD13	1	0.7
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD11	1	0.7
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD12	1	0.7
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD13	1	0.7
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD11	7	0.67
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD12	7	0.67
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD13	7	0.67
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD11	7	0.67
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD12	7	0.67
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD13	7	0.67
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD11	7	0.67
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD12	7	0.67
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD13	7	0.67
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD11	12	0.66
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD12	12	0.66
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD13	12	0.66
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD11	12	0.66
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD12	12	0.66
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD13	12	0.66
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD11	12	0.66
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD12	12	0.66
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD13	12	0.66
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD11	3	0.65
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD12	3	0.65
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD13	3	0.65
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD11	3	0.65
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD12	3	0.65
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD13	3	0.65
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD11	3	0.65
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD12	3	0.65
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD13	3	0.65

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD11	15	0.65
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD12	15	0.65
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD13	15	0.65
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD11	15	0.65
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD12	15	0.65
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD13	15	0.65
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD11	15	0.65
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD12	15	0.65
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD13	15	0.65
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD11	17	0.65
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD12	17	0.65
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD13	17	0.65
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD11	17	0.65
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD12	17	0.65
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD13	17	0.65
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD11	17	0.65
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD12	17	0.65
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD13	17	0.65
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD11	20	0.65
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD12	20	0.65
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD13	20	0.65
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD11	20	0.65
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD12	20	0.65
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD13	20	0.65
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD11	20	0.65
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD12	20	0.65
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD13	20	0.65
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD11	4	0.64
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD12	4	0.64
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD13	4	0.64
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD11	4	0.64
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD12	4	0.64
(1,47)	1:A:2:ILE:HD12	2:B:36:LEU:HD13	4	0.64
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD11	4	0.64
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD12	4	0.64
(1,47)	1:A:2:ILE:HD13	2:B:36:LEU:HD13	4	0.64
(1,47)	1:A:2:ILE:HD11	2:B:36:LEU:HD11	8	0.64

10 Dihedral-angle violation analysis (i)

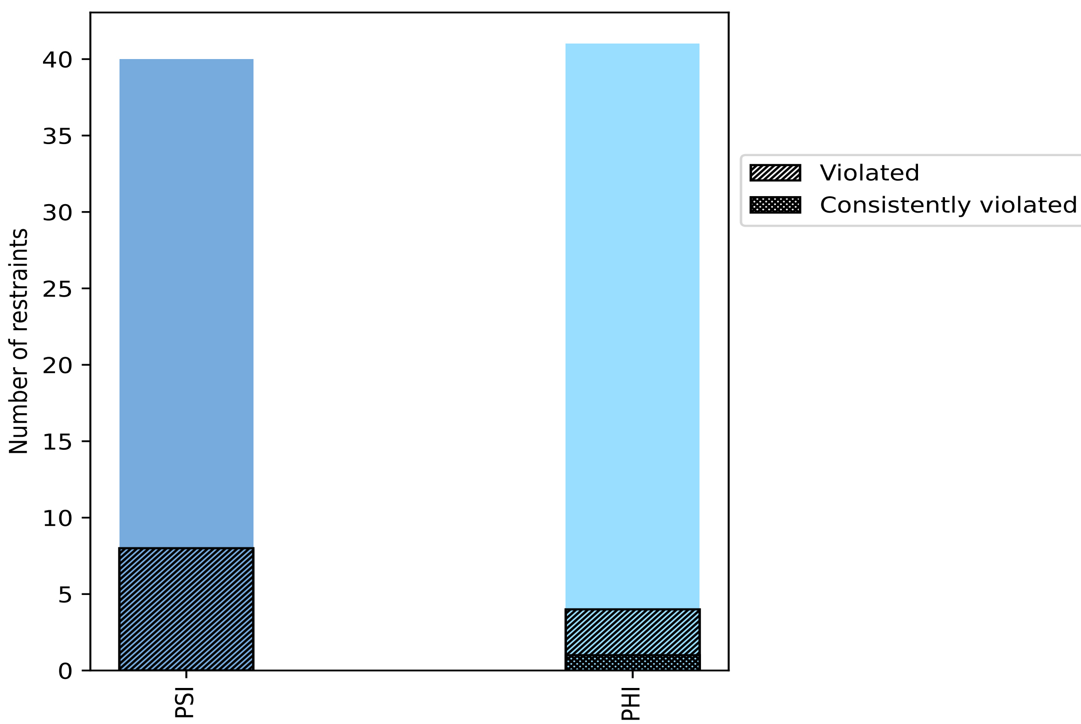
10.1 Summary of dihedral-angle violations (i)

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	40	49.4	8	20.0	9.9	0	0.0	0.0
PHI	41	50.6	4	9.8	4.9	1	2.4	1.2
Total	81	100.0	12	14.8	14.8	1	1.2	1.2

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

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



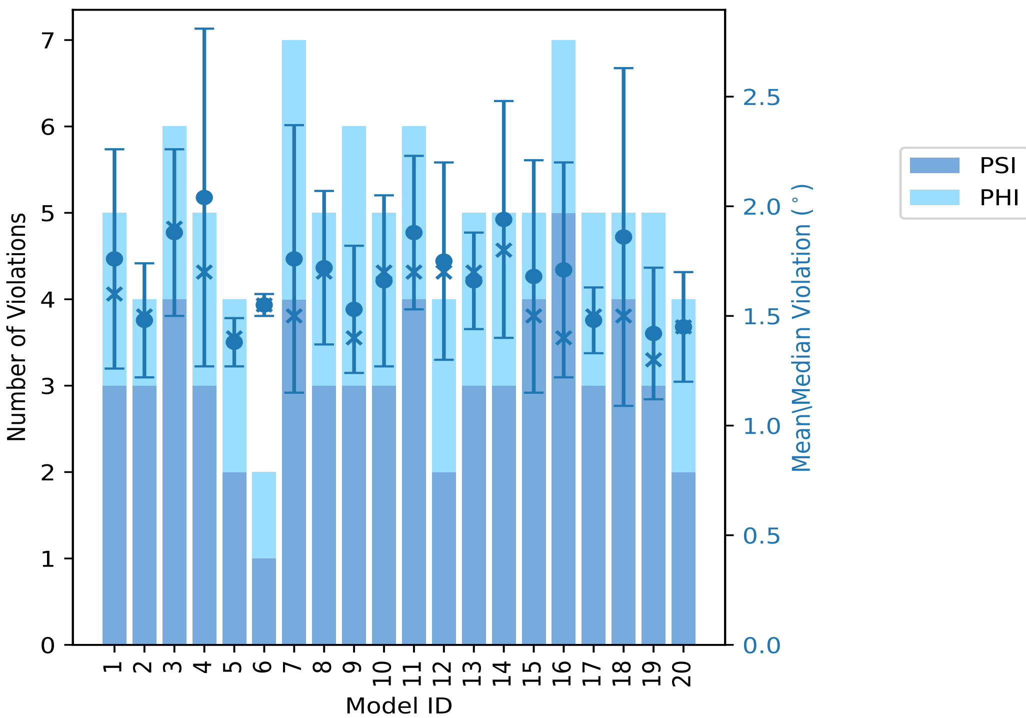
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

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

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

Model ID	Number of violations			Mean (°)	Max (°)	SD (°)	Median (°)
	PSI	PHI	Total				
1	3	2	5	1.76	2.6	0.5	1.6
2	3	1	4	1.48	1.8	0.26	1.5
3	4	2	6	1.88	2.6	0.38	1.9
4	3	2	5	2.04	3.2	0.77	1.7
5	2	2	4	1.38	1.5	0.11	1.4
6	1	1	2	1.55	1.6	0.05	1.55
7	4	3	7	1.76	2.7	0.61	1.5
8	3	2	5	1.72	2.2	0.35	1.7
9	3	3	6	1.53	2.1	0.29	1.4
10	3	2	5	1.66	2.3	0.39	1.7
11	4	2	6	1.88	2.5	0.35	1.7
12	2	2	4	1.75	2.4	0.45	1.7
13	3	2	5	1.66	2.0	0.22	1.7
14	3	2	5	1.94	2.7	0.54	1.8
15	4	1	5	1.68	2.4	0.53	1.5
16	5	2	7	1.71	2.5	0.49	1.4
17	3	2	5	1.48	1.6	0.15	1.5
18	4	1	5	1.86	3.4	0.77	1.5
19	3	2	5	1.42	2.0	0.3	1.3
20	2	2	4	1.45	1.7	0.25	1.45

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]	%
1	0	1	1	5.0
2	1	3	2	10.0
1	1	2	3	15.0
0	0	0	4	20.0
0	0	0	5	25.0
0	0	0	6	30.0
0	0	0	7	35.0
1	0	1	8	40.0
1	0	1	9	45.0
0	0	0	10	50.0
0	0	0	11	55.0

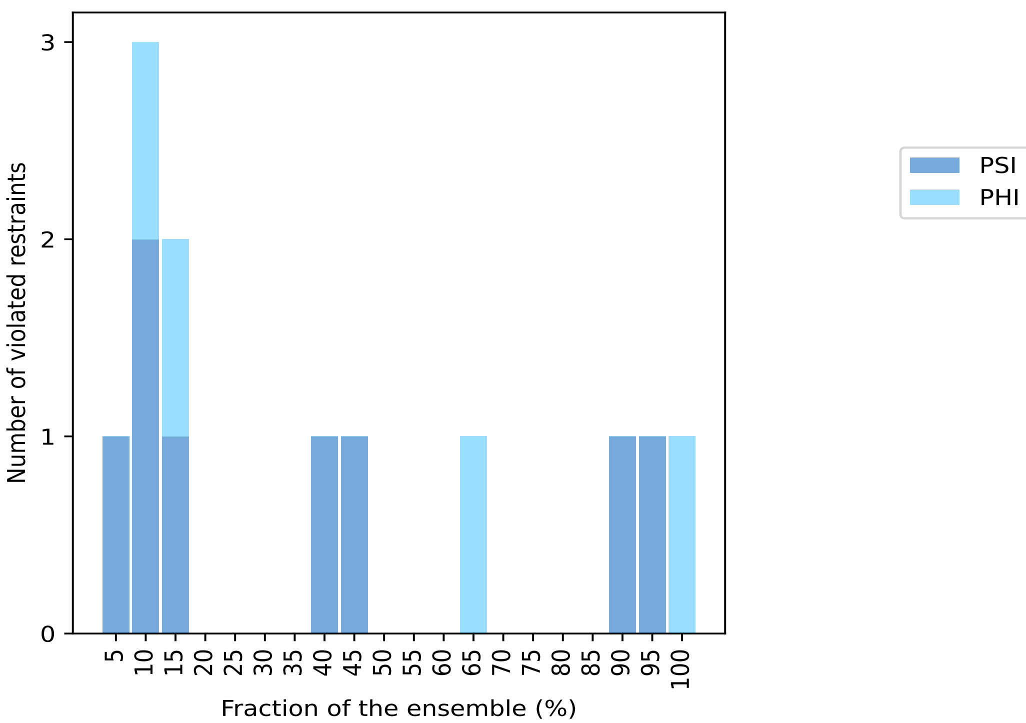
Continued on next page...

Continued from previous page...

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

[1] Number of models with violations

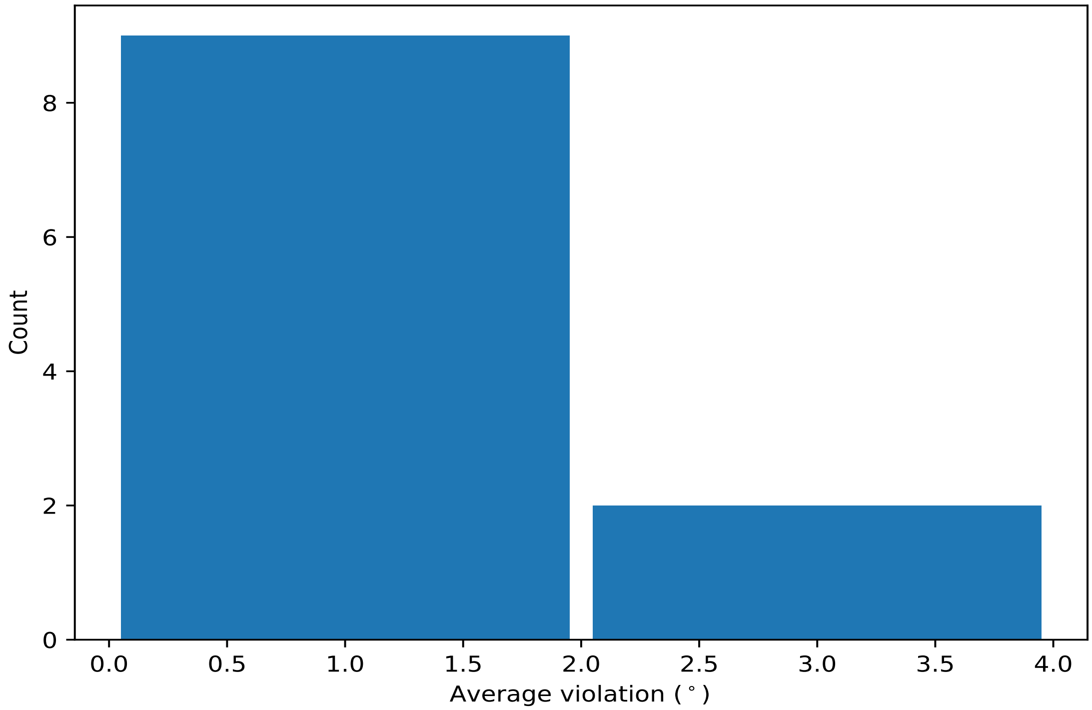
10.3.1 Bar graph : Dihedral-angle Violation statistics for the ensemble (i)



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.

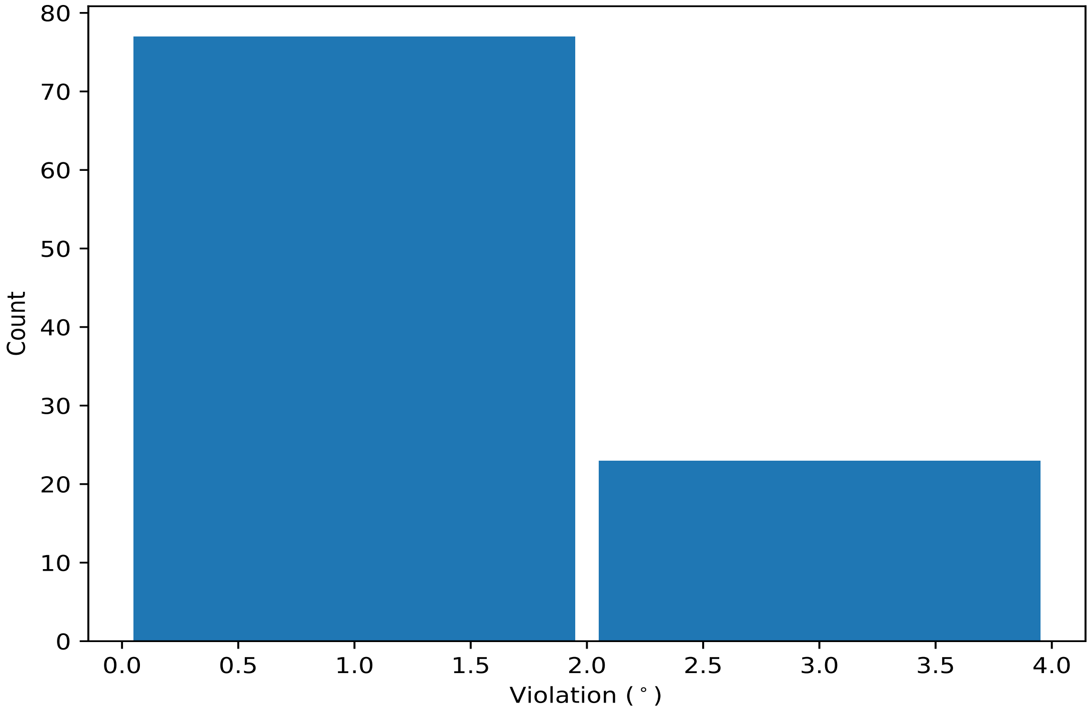
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models[1]	Mean	SD[2]	Median
(1,63)	2:B:37:TYR:C	2:B:38:LEU:N	2:B:38:LEU:CA	2:B:38:LEU:C	20	1.6	0.18	1.6
(1,40)	2:B:26:HIS:N	2:B:26:HIS:CA	2:B:26:HIS:C	2:B:27:LEU:N	19	2.19	0.53	2.2
(1,58)	2:B:35:ALA:N	2:B:35:ALA:CA	2:B:35:ALA:C	2:B:36:LEU:N	18	1.45	0.21	1.4
(1,71)	2:B:41:GLY:C	2:B:42:GLU:N	2:B:42:GLU:CA	2:B:42:GLU:C	13	1.45	0.23	1.4
(1,81)	2:B:47:TYR:N	2:B:47:TYR:CA	2:B:47:TYR:C	2:B:48:THR:N	9	2.36	0.46	2.4
(1,68)	2:B:40:CYS:N	2:B:40:CYS:CA	2:B:40:CYS:C	2:B:41:GLY:N	8	1.39	0.21	1.3
(1,80)	2:B:46:PHE:C	2:B:47:TYR:N	2:B:47:TYR:CA	2:B:47:TYR:C	3	1.53	0.21	1.5
(1,34)	1:A:20:CYS:N	1:A:20:CYS:CA	1:A:20:CYS:C	1:A:21:ASN:N	3	1.17	0.05	1.2
(1,7)	1:A:4:GLU:C	1:A:5:GLN:N	1:A:5:GLN:CA	1:A:5:GLN:C	2	1.65	0.45	1.65
(1,10)	1:A:6:CYS:N	1:A:6:CYS:CA	1:A:6:CYS:C	1:A:7:CYS:N	2	1.6	0.3	1.6

[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 (°)
(1,40)	2:B:26:HIS:N	2:B:26:HIS:CA	2:B:26:HIS:C	2:B:27:LEU:N	18	3.4
(1,81)	2:B:47:TYR:N	2:B:47:TYR:CA	2:B:47:TYR:C	2:B:48:THR:N	4	3.2
(1,81)	2:B:47:TYR:N	2:B:47:TYR:CA	2:B:47:TYR:C	2:B:48:THR:N	7	2.7
(1,40)	2:B:26:HIS:N	2:B:26:HIS:CA	2:B:26:HIS:C	2:B:27:LEU:N	4	2.7
(1,40)	2:B:26:HIS:N	2:B:26:HIS:CA	2:B:26:HIS:C	2:B:27:LEU:N	7	2.7
(1,40)	2:B:26:HIS:N	2:B:26:HIS:CA	2:B:26:HIS:C	2:B:27:LEU:N	14	2.7
(1,81)	2:B:47:TYR:N	2:B:47:TYR:CA	2:B:47:TYR:C	2:B:48:THR:N	3	2.6
(1,40)	2:B:26:HIS:N	2:B:26:HIS:CA	2:B:26:HIS:C	2:B:27:LEU:N	1	2.6
(1,40)	2:B:26:HIS:N	2:B:26:HIS:CA	2:B:26:HIS:C	2:B:27:LEU:N	11	2.5
(1,40)	2:B:26:HIS:N	2:B:26:HIS:CA	2:B:26:HIS:C	2:B:27:LEU:N	16	2.5