wwPDB NMR Structure Validation Summary Report (i)

Dec 25, 2024 – 07:57 PM EST

PDB ID	:	8R8P
BMRB ID	:	34885
Title	:	Solution structure of SMN-CX bound to the RNA helix formed upon SMN2 exon7 5'-splice site recognition
Authors	:	Malard, F.; Campagne, S.
Deposited on	:	2023-11-29

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:

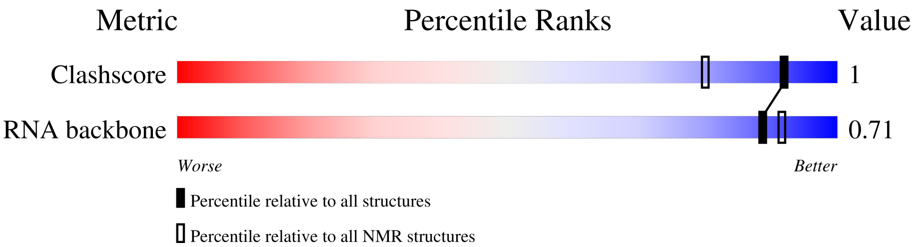
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints 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 (i)

The following experimental techniques were used to determine the structure:
SOLID-STATE NMR

The overall completeness of chemical shifts assignment is 75%.

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
RNA backbone	6643	756

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	11	18% 73% 9%
2	B	11	100%

2 Ensemble composition and analysis (i)

This entry contains 15 models. This entry does not contain polypeptide chains, therefore identification of well-defined residues and clustering analysis are not possible. All residues are included in the validation scores.

3 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 752 atoms, of which 258 are hydrogens and 0 are deuteriums.

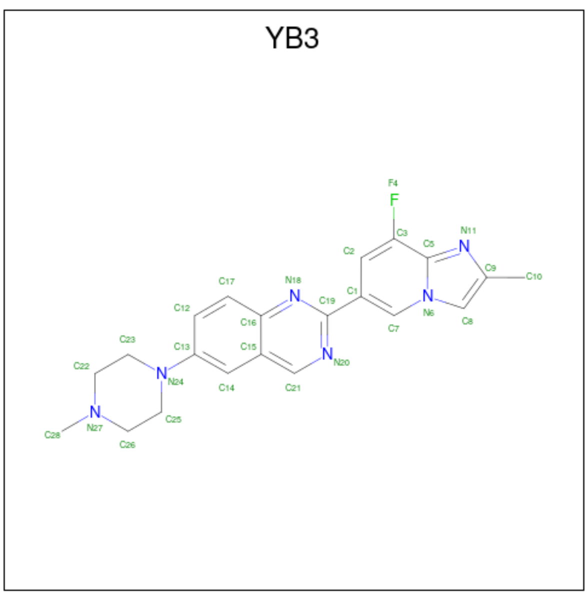
- Molecule 1 is a RNA chain called RNA (5'-R(*AP*UP*AP*CP*(PSU)P*(PSU)P*AP*CP*CP*UP*G)-3').

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
1	A	11	345	103	117	37	78	10	0

- Molecule 2 is a RNA chain called RNA (5'-R(P*GP*GP*AP*GP*UP*AP*AP*GP*UP*CP*U)-3').

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
2	B	11	356	106	118	44	77	11	0

- Molecule 3 is 2-(8-fluoranyl-2-methyl-imidazo[1,2-a]pyridin-6-yl)-6-(4-methylpiperazin-1-yl)quinazoline (three-letter code: YB3) (formula: $C_{21}H_{21}FN_6$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				
			Total	C	F	H	N
3	B	1	51	21	1	23	6

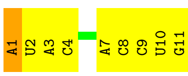
4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

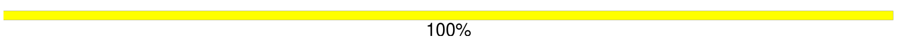
These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for 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: RNA (5'-R(*AP*UP*AP*CP*(PSU)P*(PSU)P*AP*CP*CP*UP*G)-3')

Chain A:

- Molecule 2: RNA (5'-R(P*GP*GP*AP*GP*UP*AP*AP*GP*UP*CP*U)-3')

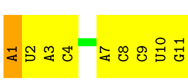
Chain B:


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

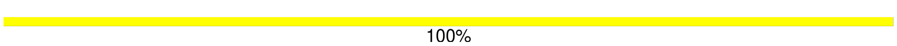
The representative model is number 1. Colouring as in section 4.1 above.

- Molecule 1: RNA (5'-R(*AP*UP*AP*CP*(PSU)P*(PSU)P*AP*CP*CP*UP*G)-3')

Chain A:

- Molecule 2: RNA (5'-R(P*GP*GP*AP*GP*UP*AP*AP*GP*UP*CP*U)-3')

Chain B:


5 Refinement protocol and experimental data overview (i)

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

Of the 50 calculated structures, 15 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
Amber	refinement	
CYANA	structure calculation	3.98

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	341
Number of shifts mapped to atoms	320
Number of unparsed shifts	0
Number of shifts with mapping errors	21
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	75%

Note: This is a solid-state NMR structure, where hydrogen atoms are typically not assigned a chemical shift value, which may lead to lower completeness of assignment measure.

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: YB3, PSU

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	1.53±0.01	3±0/207 (1.5± 0.1%)	1.83±0.02	8±0/320 (2.5± 0.0%)
2	B	1.50±0.01	3±0/266 (1.1± 0.0%)	1.90±0.02	11±0/413 (2.7± 0.0%)
All	All	1.51	91/7095 (1.3%)	1.87	285/10995 (2.6%)

5 of 7 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	B	14	A	C6-N1	-6.73	1.30	1.35	10	15
1	A	7	A	C6-N1	-6.61	1.30	1.35	14	15
2	B	17	A	C6-N1	-6.53	1.30	1.35	6	15
1	A	1	A	C6-N1	-6.43	1.31	1.35	4	15
1	A	3	A	C6-N1	-6.37	1.31	1.35	15	15

5 of 19 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(°)	Ideal(°)	Models	
								Worst	Total
2	B	12	G	OP1-P-OP2	-9.17	105.85	119.60	9	15
1	A	9	C	OP1-P-OP2	-7.72	108.02	119.60	15	15
2	B	16	U	OP1-P-OP2	-7.68	108.08	119.60	8	15
1	A	3	A	OP1-P-OP2	-7.56	108.26	119.60	7	15
2	B	19	G	OP1-P-OP2	-7.55	108.27	119.60	12	15

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	228	117	119	1±0
All	All	7410	3870	3570	15

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 1.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:1:A:OP2	1:A:1:A:OP1	0.42	2.38	10	15

6.3 Torsion angles (i)

6.3.1 Protein backbone (i)

There are no protein molecules in this entry.

6.3.2 Protein sidechains (i)

There are no protein molecules in this entry.

6.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers	Suiteness
1	A	10/11 (91%)	0±0 (0±0%)	0±0 (0±0%)	0.68±0.02
2	B	10/11 (91%)	0±0 (0±0%)	0±0 (0±0%)	0.73±0.01
All	All	300/330 (91%)	0 (0%)	0 (0%)	0.71

The overall RNA backbone suiteness is 0.71.

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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

2 non-standard protein/DNA/RNA residues are 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	PSU	A	5	2,1	18,21,22	0.75±0.02	0±0 (1±2%)
1	PSU	A	6	2,1	18,21,22	0.75±0.03	0±0 (1±2%)

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	PSU	A	5	2,1	21,30,33	0.78±0.02	0±0 (0±0%)
1	PSU	A	6	2,1	21,30,33	0.78±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	PSU	A	5	2,1	-	0±0,7,25,26	0±0,2,2,2
1	PSU	A	6	2,1	-	0±0,7,25,26	0±0,2,2,2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	6	PSU	C2-N1	2.12	1.39	1.36	10	5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	5	PSU	C2-N1	2.12	1.39	1.36	7	5

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 oligosaccharides in this entry.

6.6 Ligand geometry (i)

1 ligand 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
3	YB3	B	101	-	29,32,32	1.31±0.02	5±0 (17±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
3	YB3	B	101	-	39,47,47	1.64±0.04	7±1 (18±3%)

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
3	YB3	B	101	-	-	0±0,8,18,18	0±0,5,5,5

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
3	B	101	YB3	C22-N27	2.68	1.52	1.46	12	15
3	B	101	YB3	C26-N27	2.67	1.52	1.46	9	15
3	B	101	YB3	C2-C1	2.66	1.44	1.39	12	15
3	B	101	YB3	C28-N27	2.32	1.52	1.46	2	15
3	B	101	YB3	C19-N18	2.31	1.39	1.33	6	15

5 of 11 unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

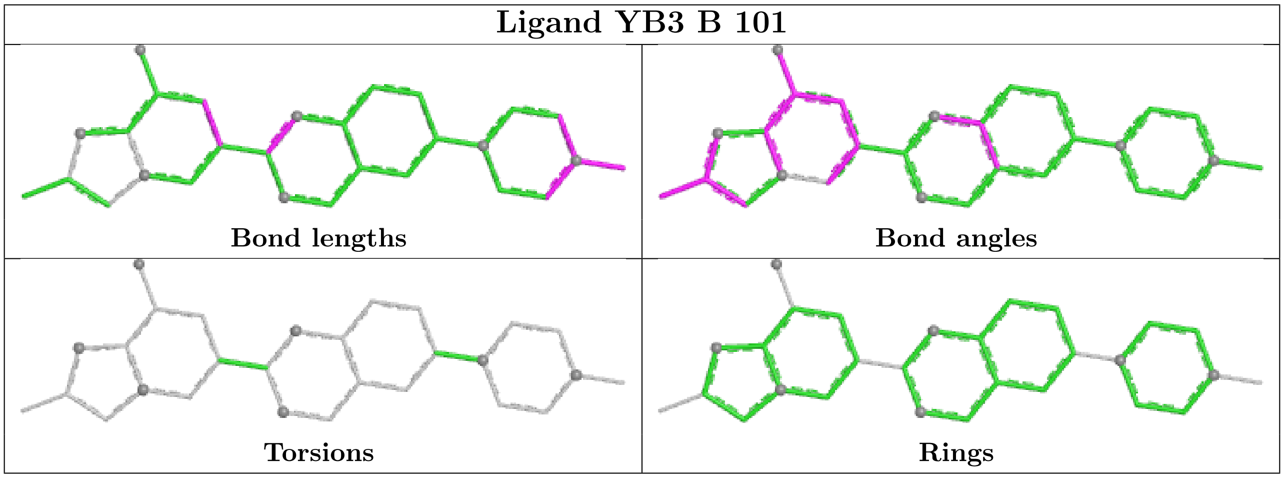
Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
3	B	101	YB3	F4-C3-C5	5.11	119.49	117.37	13	15
3	B	101	YB3	C7-C1-C2	4.58	123.88	117.76	9	15
3	B	101	YB3	C15-C16-N18	2.45	118.88	122.25	5	15
3	B	101	YB3	C10-C9-C8	2.42	122.72	129.07	10	15
3	B	101	YB3	C23-C22-N27	2.32	114.58	110.86	4	11

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



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 75% for the well-defined parts and 75% 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	341
Number of shifts mapped to atoms	320
Number of unparsed shifts	0
Number of shifts with mapping errors	21
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. First 5 (of 21) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	B	101	SCX	H10A	2.145	0.020	2
1	B	101	SCX	H10B	2.145	0.020	2
1	B	101	SCX	H10C	2.145	0.020	2
1	B	101	SCX	H121	7.449	0.020	1
1	B	101	SCX	H14	7.678	0.020	1
1	B	101	SCX	H17	7.212	0.020	1
1	B	101	SCX	H2	7.181	0.020	1
1	B	101	SCX	H21	6.61	0.020	1
1	B	101	SCX	H22A	3.553	0.020	2
1	B	101	SCX	H22B	3.553	0.020	2
1	B	101	SCX	H23A	3.672	0.020	2
1	B	101	SCX	H23B	3.672	0.020	2
1	B	101	SCX	H25A	3.731	0.020	2
1	B	101	SCX	H25B	3.731	0.020	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	B	101	SCX	H26A	3.116	0.020	2
1	B	101	SCX	H26B	3.116	0.020	2
1	B	101	SCX	H28A	2.887	0.020	2
1	B	101	SCX	H28B	2.887	0.020	2
1	B	101	SCX	H28C	2.887	0.020	2
1	B	101	SCX	H7	8.591	0.020	1
1	B	101	SCX	H8	7.356	0.020	1

7.1.2 Chemical shift referencing (i)

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

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 75%, i.e. 282 atoms were assigned a chemical shift out of a possible 375. 0 out of 0 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Sugar	220/220 (100%)	120/120 (100%)	100/100 (100%)	0/0 (—%)
Base	62/155 (40%)	32/95 (34%)	30/35 (86%)	0/25 (0%)
Overall	282/375 (75%)	152/215 (71%)	130/135 (96%)	0/25 (0%)

Note: This is a solid-state NMR structure, where hydrogen atoms are typically not assigned a chemical shift value, which may lead to lower completeness of assignment measure.

7.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

7.1.5 Random Coil Index (RCI) plots (i)

No *random coil index*(RCI) plot could be generated from the current chemical shift list. RCI is only applicable to proteins

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	419
Intra-residue (|i-j|=0)	248
Sequential (|i-j|=1)	87
Medium range (|i-j|>1 and |i-j|<5)	0
Long range (|i-j|≥5)	0
Inter-chain	12
Hydrogen bond restraints	72
Disulfide bond restraints	0
Total dihedral-angle restraints	165
Number of unmapped restraints	7
Number of restraints per residue	25.4
Number of long range restraints per residue[1]	0.0

[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)	16.7	0.2
0.2-0.5 (Medium)	26.9	0.5
>0.5 (Large)	15.1	5.78

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)	37.1	9.92
10.0-20.0 (Medium)	1.9	14.32
>20.0 (Large)	None	None

9 Distance violation analysis (i)

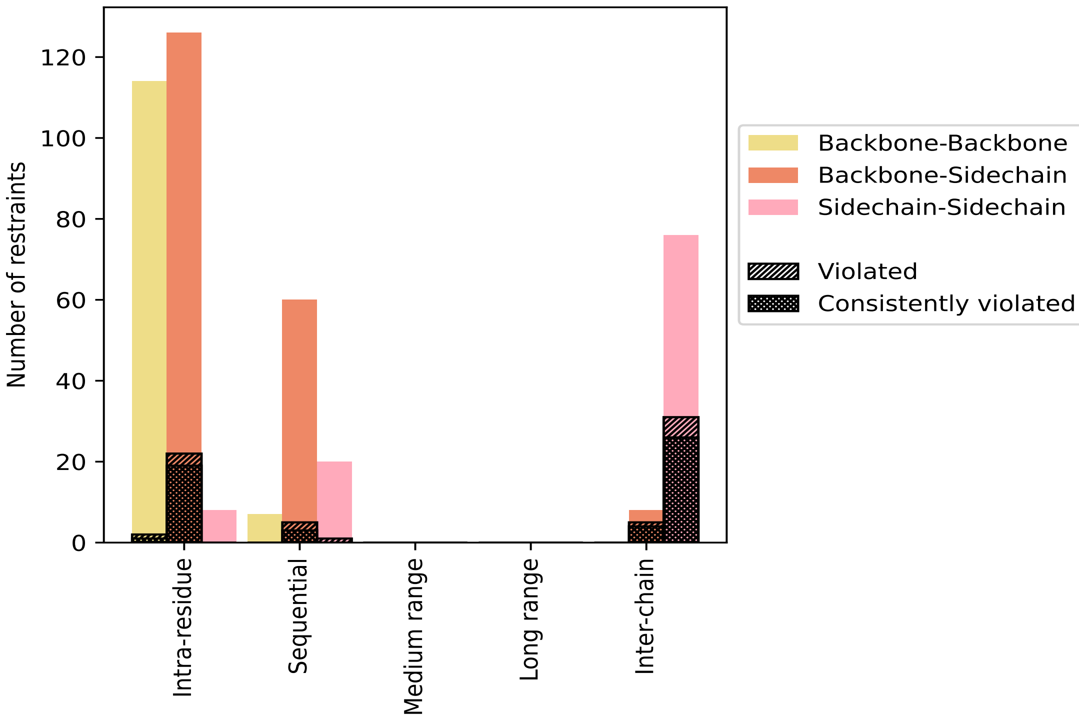
9.1 Summary of distance violations (i)

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)	248	59.2	24	9.7	5.7	20	8.1	4.8
Backbone-Backbone	114	27.2	2	1.8	0.5	1	0.9	0.2
Backbone-Sidechain	126	30.1	22	17.5	5.3	19	15.1	4.5
Sidechain-Sidechain	8	1.9	0	0.0	0.0	0	0.0	0.0
Sequential (|i-j|=1)	87	20.8	6	6.9	1.4	3	3.4	0.7
Backbone-Backbone	7	1.7	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	60	14.3	5	8.3	1.2	3	5.0	0.7
Sidechain-Sidechain	20	4.8	1	5.0	0.2	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	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
Long range (|i-j|≥5)	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
Inter-chain	12	2.9	1	8.3	0.2	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	4	1.0	1	25.0	0.2	0	0.0	0.0
Sidechain-Sidechain	8	1.9	0	0.0	0.0	0	0.0	0.0
Hydrogen bond	72	17.2	35	48.6	8.4	30	41.7	7.2
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	419	100.0	66	15.8	15.8	53	12.6	12.6
Backbone-Backbone	121	28.9	2	1.7	0.5	1	0.8	0.2
Backbone-Sidechain	194	46.3	32	16.5	7.6	26	13.4	6.2
Sidechain-Sidechain	104	24.8	32	30.8	7.6	26	25.0	6.2

[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	21	4	0	0	34	59	1.05	5.73	1.61	0.24
2	22	3	0	0	33	58	1.07	5.73	1.62	0.26
3	22	3	0	0	35	60	1.05	5.74	1.6	0.27
4	22	4	0	0	33	59	1.05	5.72	1.61	0.26
5	23	4	0	0	34	61	1.02	5.71	1.59	0.24
6	21	3	0	0	34	58	1.07	5.75	1.62	0.26
7	22	4	0	0	33	59	1.06	5.75	1.61	0.25
8	22	3	0	0	32	57	1.09	5.75	1.63	0.28
9	22	4	0	0	34	60	1.04	5.78	1.61	0.25
10	22	3	0	0	34	59	1.06	5.72	1.61	0.27

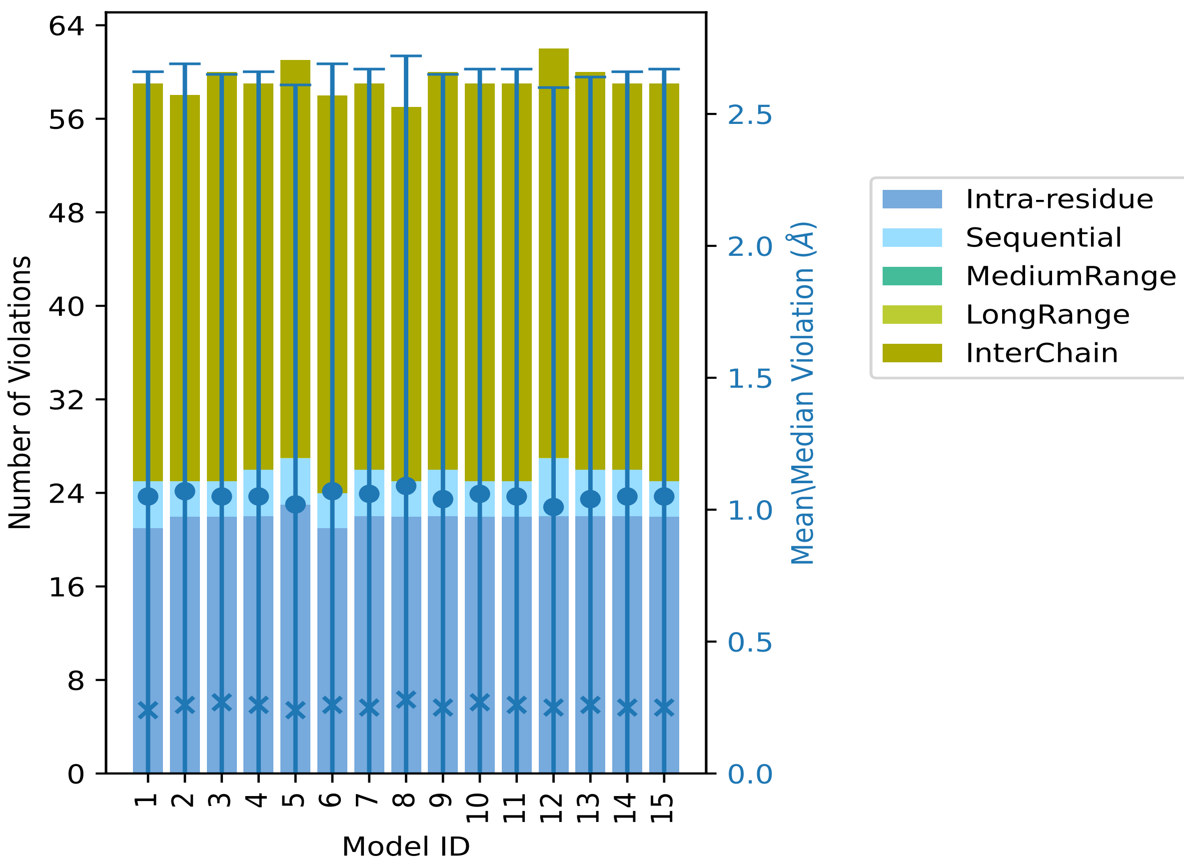
Continued on next page...

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	22	3	0	0	34	59	1.05	5.76	1.62	0.26
12	22	5	0	0	35	62	1.01	5.77	1.59	0.25
13	22	4	0	0	34	60	1.04	5.75	1.6	0.26
14	22	4	0	0	33	59	1.05	5.75	1.61	0.25
15	22	3	0	0	34	59	1.05	5.77	1.62	0.25

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

9.2.1 Bar graph : Distance Violation statistics for each model



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

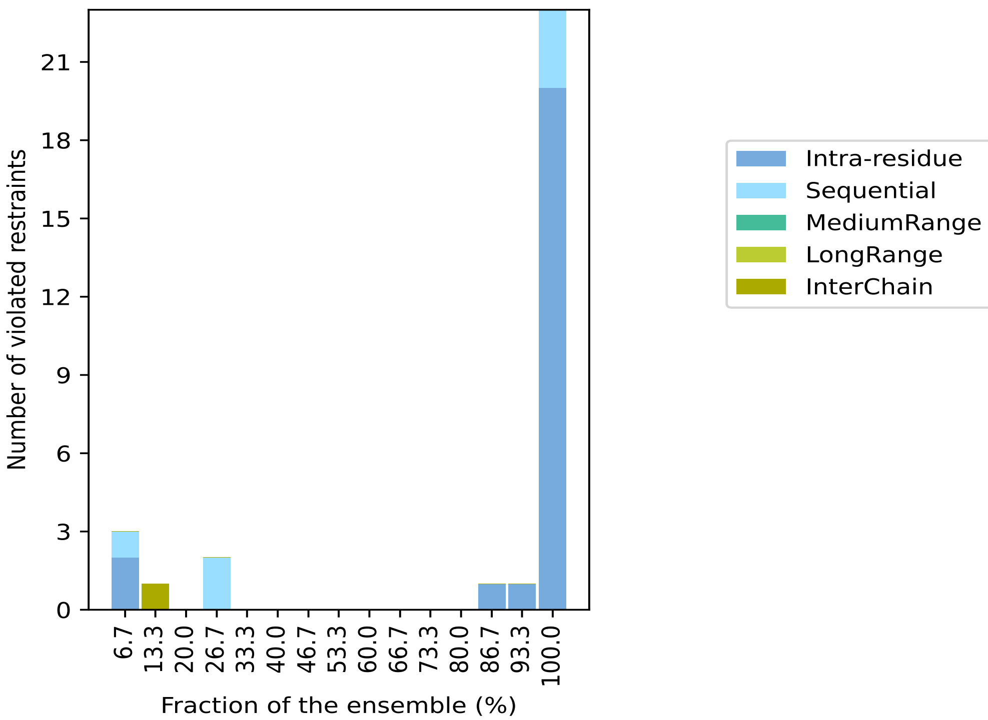
9.3 Distance violation statistics for the ensemble (i)

Violation analysis may find that some restraints are violated in few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of the ensemble. In total, 316(IR:224, SQ:81, MR:0, LR:0, IC:11) 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]	%
2	1	0	0	0	3	1	6.7
0	0	0	0	1	1	2	13.3
0	0	0	0	0	0	3	20.0
0	2	0	0	0	2	4	26.7
0	0	0	0	0	0	5	33.3
0	0	0	0	0	0	6	40.0
0	0	0	0	0	0	7	46.7
0	0	0	0	0	0	8	53.3
0	0	0	0	0	0	9	60.0
0	0	0	0	0	0	10	66.7
0	0	0	0	0	0	11	73.3
0	0	0	0	0	0	12	80.0
1	0	0	0	0	1	13	86.7
1	0	0	0	0	1	14	93.3
20	3	0	0	0	23	15	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

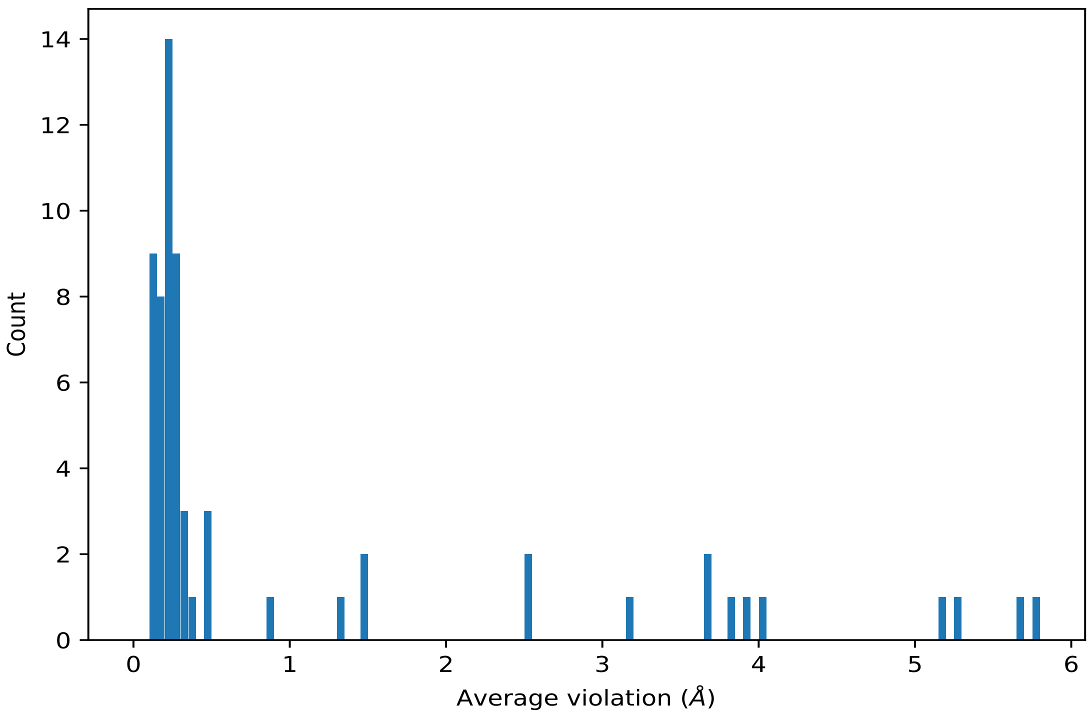
9.3.1 Bar graph : Distance violation statistics for the ensemble (i)



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.

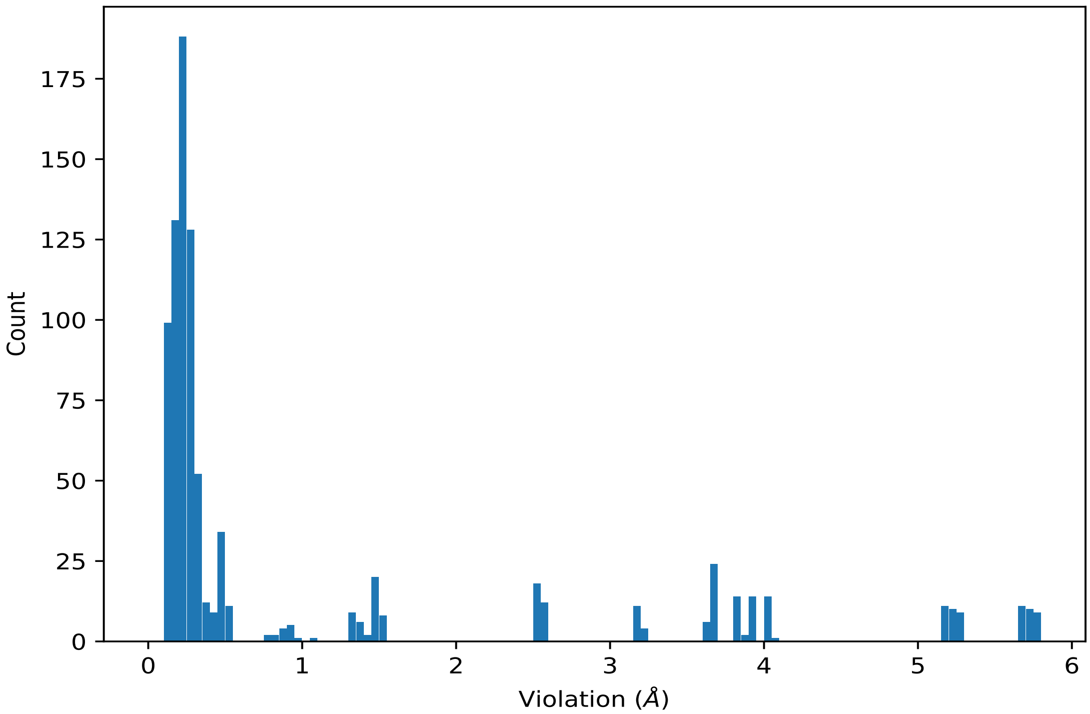
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(6,13)	1:5:A:PSU:H3'	2:18:B:A:N1	15	5.75	0.02	5.75
(6,17)	1:6:A:PSU:H3'	2:17:B:A:N1	15	5.68	0.02	5.69
(1,13)	1:5:A:PSU:H3'	2:18:B:A:N1	15	5.25	0.02	5.25
(1,17)	1:6:A:PSU:H3'	2:17:B:A:N1	15	5.18	0.02	5.19
(2,150)	1:9:A:C:H5	1:9:A:C:H4'	15	4.03	0.01	4.03
(2,245)	2:16:B:U:H5	2:16:B:U:H4'	15	3.91	0.01	3.91
(2,249)	2:16:B:U:H5	2:16:B:U:H4'	15	3.82	0.01	3.82
(2,322)	2:21:B:C:H5	2:21:B:C:H4'	15	3.65	0.01	3.65
(2,324)	2:21:B:C:H5	2:21:B:C:H4'	15	3.65	0.01	3.65
(2,25)	1:2:A:U:H4'	1:2:A:U:H5	15	3.19	0.01	3.19

[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)
(6,13)	1:5:A:PSU:H3'	2:18:B:A:N1	9	5.78
(6,13)	1:5:A:PSU:H3'	2:18:B:A:N1	12	5.77
(6,13)	1:5:A:PSU:H3'	2:18:B:A:N1	15	5.77
(6,13)	1:5:A:PSU:H3'	2:18:B:A:N1	11	5.76
(6,13)	1:5:A:PSU:H3'	2:18:B:A:N1	6	5.75
(6,13)	1:5:A:PSU:H3'	2:18:B:A:N1	7	5.75
(6,13)	1:5:A:PSU:H3'	2:18:B:A:N1	8	5.75
(6,13)	1:5:A:PSU:H3'	2:18:B:A:N1	13	5.75
(6,13)	1:5:A:PSU:H3'	2:18:B:A:N1	14	5.75
(6,17)	1:6:A:PSU:H3'	2:17:B:A:N1	3	5.74

10 Dihedral-angle violation analysis (i)

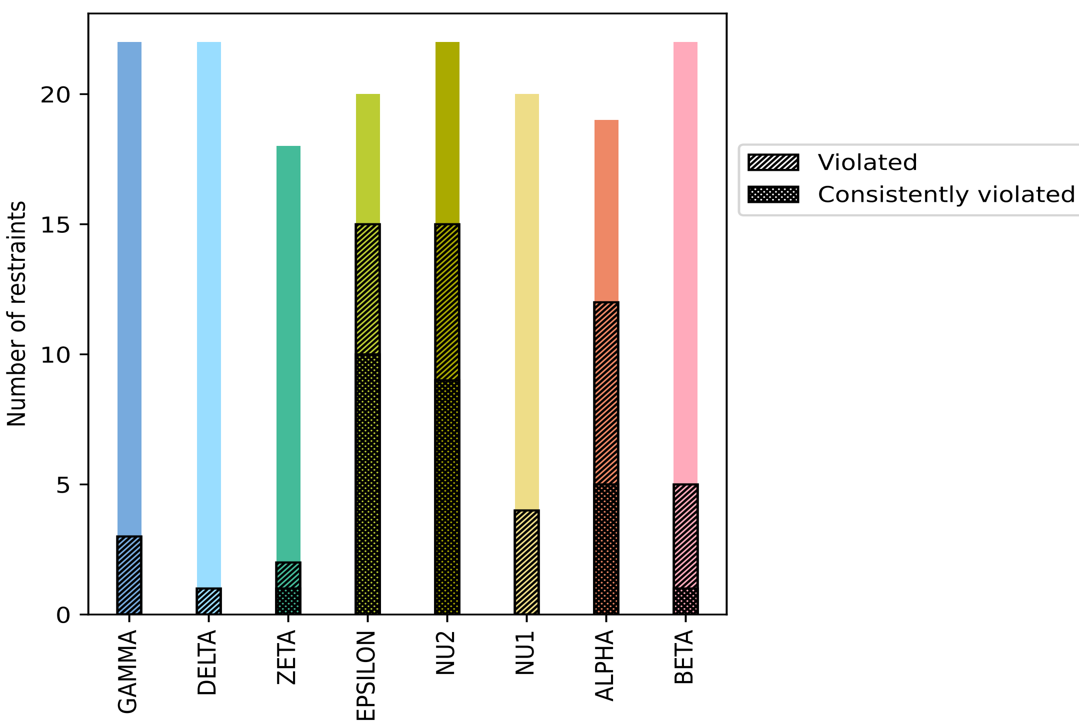
10.1 Summary of dihedral-angle violations (i)

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]
GAMMA	22	13.3	3	13.6	1.8	0	0.0	0.0
DELTA	22	13.3	1	4.5	0.6	0	0.0	0.0
ZETA	18	10.9	2	11.1	1.2	1	5.6	0.6
EPSILON	20	12.1	15	75.0	9.1	10	50.0	6.1
NU2	22	13.3	15	68.2	9.1	9	40.9	5.5
NU1	20	12.1	4	20.0	2.4	0	0.0	0.0
ALPHA	19	11.5	12	63.2	7.3	5	26.3	3.0
BETA	22	13.3	5	22.7	3.0	1	4.5	0.6
Total	165	100.0	57	34.5	34.5	26	15.8	15.8

[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)



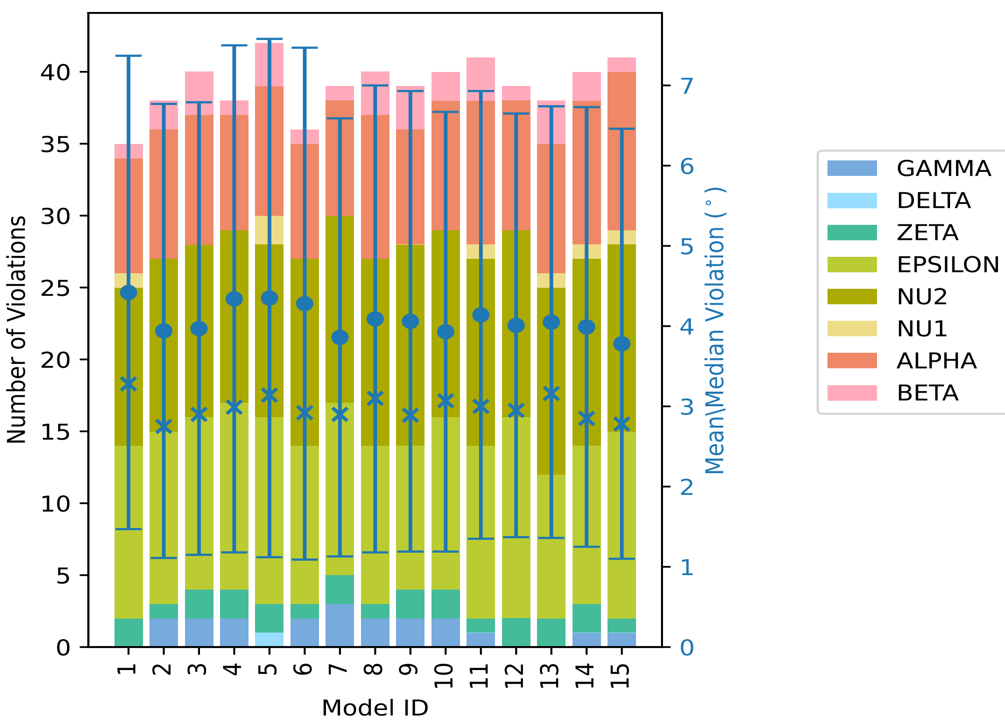
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

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 (°)
	GAMMA	DELTA	ZETA	EPSILON	NU2	NU1	ALPHA	BETA	Total				
1	0	0	2	12	11	1	8	1	35	4.42	11.03	2.95	3.28
2	2	0	1	12	12	0	9	2	38	3.94	11.0	2.83	2.75
3	2	0	2	12	12	0	9	3	40	3.97	11.33	2.82	2.9
4	2	0	2	13	12	0	8	1	38	4.34	13.69	3.16	2.99
5	0	1	2	13	12	2	9	3	42	4.35	14.32	3.23	3.14
6	2	0	1	11	13	0	8	1	36	4.28	12.38	3.19	2.92
7	3	0	2	12	13	0	8	1	39	3.86	10.48	2.73	2.9
8	2	0	1	11	13	0	10	3	40	4.09	11.86	2.91	3.1
9	2	0	2	10	14	0	8	3	39	4.06	11.32	2.87	2.89
10	2	0	2	12	13	0	9	2	40	3.93	11.23	2.74	3.07
11	1	0	1	12	13	1	10	3	41	4.14	10.3	2.79	3.0
12	0	0	2	14	13	0	9	1	39	4.01	10.77	2.64	2.95
13	0	0	2	10	13	1	9	3	38	4.05	11.43	2.69	3.16
14	1	0	2	11	13	1	10	2	40	3.99	10.45	2.74	2.85
15	1	0	1	13	13	1	11	1	41	3.78	10.33	2.68	2.78

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

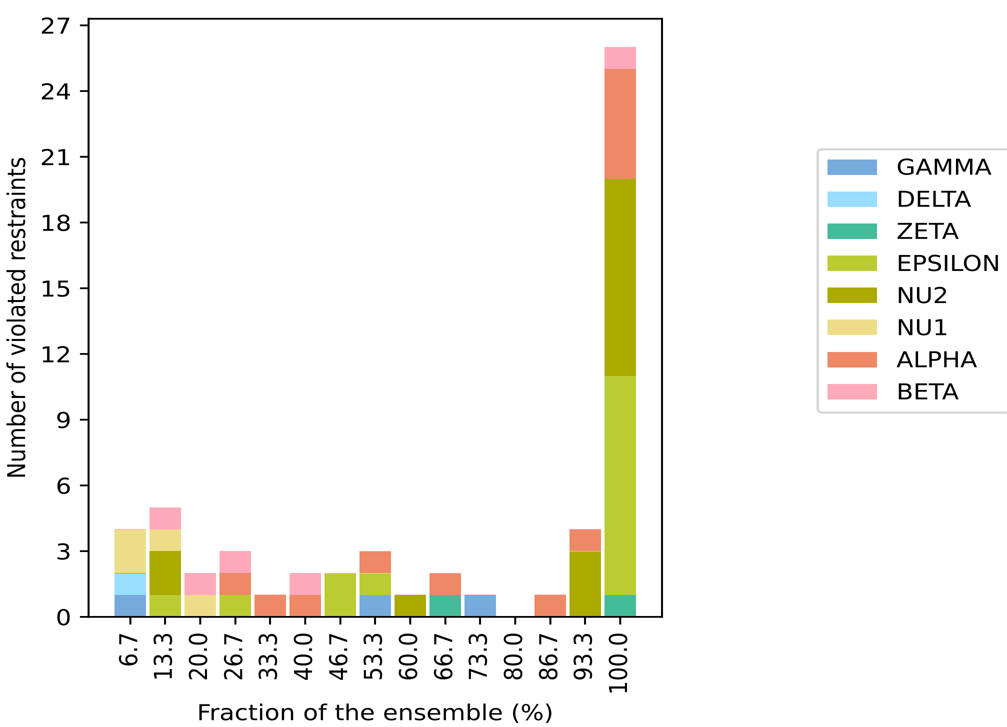
10.3 Dihedral-angle violation statistics for the ensemble (i)

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	
GAMMA	DELTA	ZETA	EPSILON	NU2	NU1	ALPHA	BETA	Total	Count[1]	%
1	1	0	0	0	2	0	0	4	1	6.7
0	0	0	1	2	1	0	1	5	2	13.3
0	0	0	0	0	1	0	1	2	3	20.0
0	0	0	1	0	0	1	1	3	4	26.7
0	0	0	0	0	0	1	0	1	5	33.3
0	0	0	0	0	0	1	1	2	6	40.0
0	0	0	2	0	0	0	0	2	7	46.7
1	0	0	1	0	0	1	0	3	8	53.3
0	0	0	0	1	0	0	0	1	9	60.0
0	0	1	0	0	0	1	0	2	10	66.7
1	0	0	0	0	0	0	0	1	11	73.3
0	0	0	0	0	0	0	0	0	12	80.0
0	0	0	0	0	0	1	0	1	13	86.7
0	0	0	0	3	0	1	0	4	14	93.3
0	0	1	10	9	0	5	1	26	15	100.0

[1] Number of models with violations

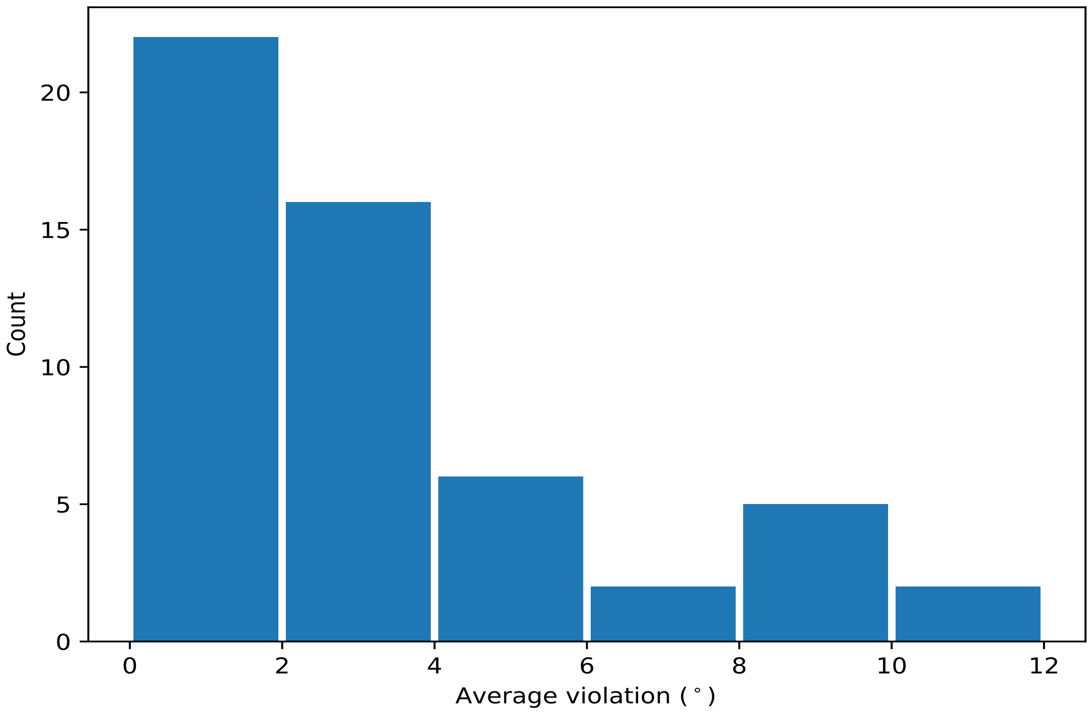
10.3.1 Bar graph : Dihedral-angle Violation statistics for the ensemble (i)



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.

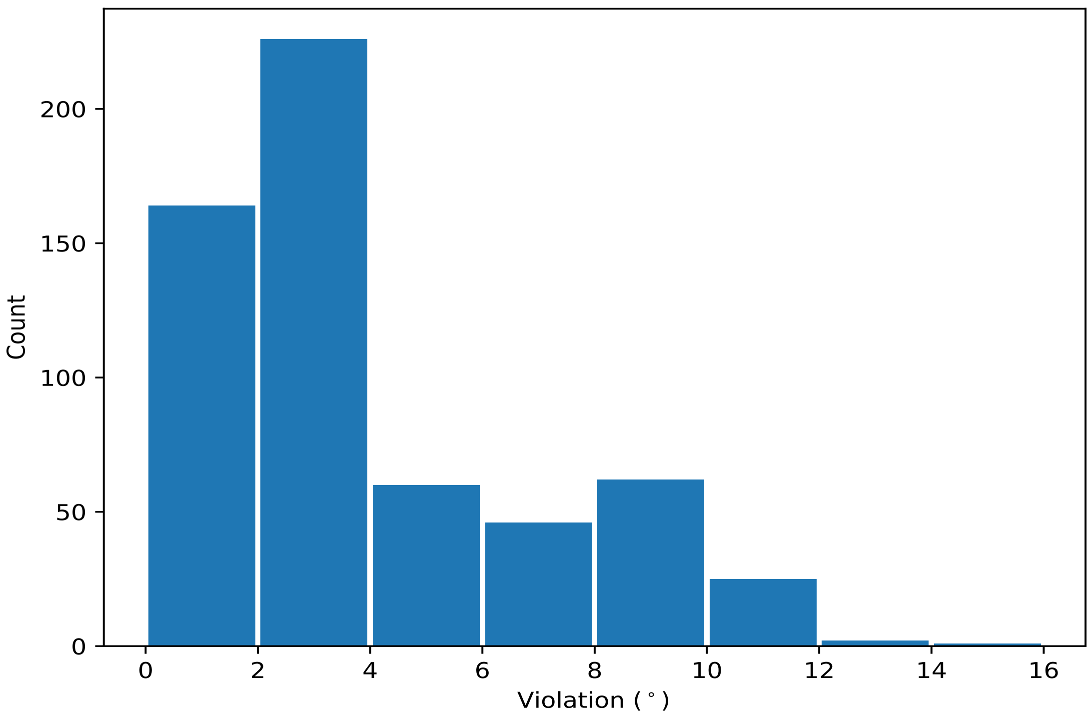
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models[1]	Mean	SD[2]	Median
(1,89)	2:12:B:G:O3'	2:13:B:G:P	2:13:B:G:O5'	2:13:B:G:C5'	15	10.85	1.31	11.0
(1,37)	1:5:A:PSU:C4'	1:5:A:PSU:C3'	1:5:A:PSU:O3'	1:6:A:PSU:P	15	10.35	0.83	10.43
(1,119)	2:16:B:U:C4'	2:16:B:U:C3'	2:16:B:U:O3'	2:17:B:A:P	15	8.75	0.98	8.56
(1,103)	2:14:B:A:C4'	2:14:B:A:C3'	2:14:B:A:O3'	2:15:B:G:P	15	8.65	0.85	8.51
(1,7)	1:1:A:A:O3'	1:2:A:U:P	1:2:A:U:O5'	1:2:A:U:C5'	15	8.49	0.62	8.36
(1,111)	2:15:B:G:C4'	2:15:B:G:C3'	2:15:B:G:O3'	2:16:B:U:P	15	8.48	0.62	8.31
(1,69)	1:9:A:C:O3'	1:10:A:U:P	1:10:A:U:O5'	1:10:A:U:C5'	15	8.45	1.04	8.59
(1,62)	1:8:A:C:O3'	1:9:A:C:P	1:9:A:C:O5'	1:9:A:C:C5'	15	7.1	1.02	6.96
(1,159)	2:21:B:C:C4'	2:21:B:C:C3'	2:21:B:C:O3'	2:22:B:U:P	15	5.86	0.96	5.74
(1,45)	1:6:A:PSU:C4'	1:6:A:PSU:C3'	1:6:A:PSU:O3'	1:7:A:A:P	15	5.8	1.61	6.01

[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 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 (°)
(1,78)	1:11:A:G:C5'	1:11:A:G:C4'	1:11:A:G:C3'	1:11:A:G:O3'	5	14.32
(1,89)	2:12:B:G:O3'	2:13:B:G:P	2:13:B:G:O5'	2:13:B:G:C5'	4	13.69
(1,89)	2:12:B:G:O3'	2:13:B:G:P	2:13:B:G:O5'	2:13:B:G:C5'	6	12.38
(1,89)	2:12:B:G:O3'	2:13:B:G:P	2:13:B:G:O5'	2:13:B:G:C5'	5	11.87
(1,89)	2:12:B:G:O3'	2:13:B:G:P	2:13:B:G:O5'	2:13:B:G:C5'	8	11.86
(1,37)	1:5:A:PSU:C4'	1:5:A:PSU:C3'	1:5:A:PSU:O3'	1:6:A:PSU:P	4	11.63
(1,37)	1:5:A:PSU:C4'	1:5:A:PSU:C3'	1:5:A:PSU:O3'	1:6:A:PSU:P	13	11.43
(1,37)	1:5:A:PSU:C4'	1:5:A:PSU:C3'	1:5:A:PSU:O3'	1:6:A:PSU:P	3	11.33
(1,89)	2:12:B:G:O3'	2:13:B:G:P	2:13:B:G:O5'	2:13:B:G:C5'	9	11.32
(1,89)	2:12:B:G:O3'	2:13:B:G:P	2:13:B:G:O5'	2:13:B:G:C5'	3	11.23