



wwPDB NMR Structure Validation Summary Report ⓘ

Dec 25, 2024 – 03:11 AM EST

PDB ID : 4BS2
BMRB ID : 19290
Title : NMR structure of human TDP-43 tandem RRM_s in complex with UG-rich RNA
Authors : Lukavsky, P.J.; Daujotyte, D.; Tollervey, J.R.; Ule, J.; Stuani, C.; Buratti, E.; Baralle, F.E.; Damberger, F.F.; Allain, F.H.T.
Deposited on : 2013-06-06

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

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

MolProbity : 4.02b-467
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
wwPDB-RCI : v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV : Wang et al. (2010)
wwPDB-ShiftChecker : v1.2
BMRB Restraints Analysis : v1.2
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.40

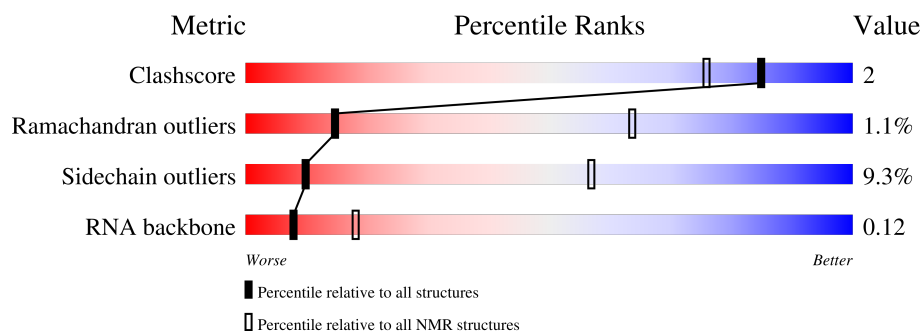
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 92%.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	210492	14027
Ramachandran outliers	207382	12486
Sidechain outliers	206894	12463
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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	174	 79% 10% 11%
2	B	12	 17% 42% 33% 8%

2 Ensemble composition and analysis

This entry contains 20 models. Model 4 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	Residue range (total)	Backbone RMSD (Å)	Medoid model
1	A:101-A:138, A:145-A:261 (155)	0.44	4

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 7 single-model clusters were found.

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

3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3143 atoms, of which 1485 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called TAR DNA-BINDING PROTEIN 43.

Mol	Chain	Residues	Atoms						Trace
1	A	174	Total	C	H	N	O	S	0
			2774	885	1373	241	265	10	

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	96	GLY	-	expression tag	UNP Q13148
A	97	SER	-	expression tag	UNP Q13148
A	98	HIS	-	expression tag	UNP Q13148
A	99	MET	-	expression tag	UNP Q13148
A	100	ALA	-	expression tag	UNP Q13148
A	101	SER	-	expression tag	UNP Q13148

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

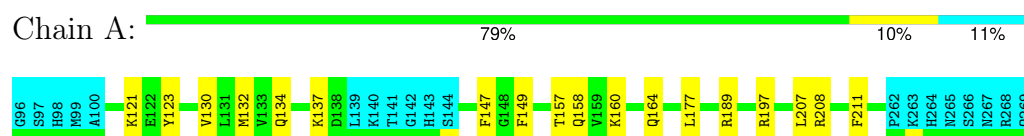
Mol	Chain	Residues	Atoms						Trace
2	B	12	Total	C	H	N	O	P	0
			369	116	112	48	82	11	

4 Residue-property plots

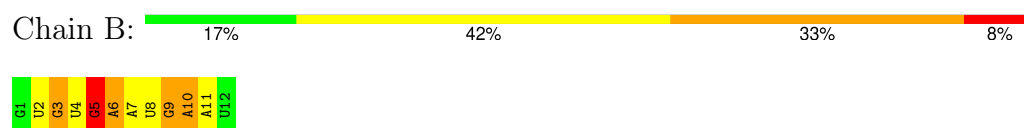
4.1 Average score per residue in the NMR ensemble

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

- Molecule 1: TAR DNA-BINDING PROTEIN 43



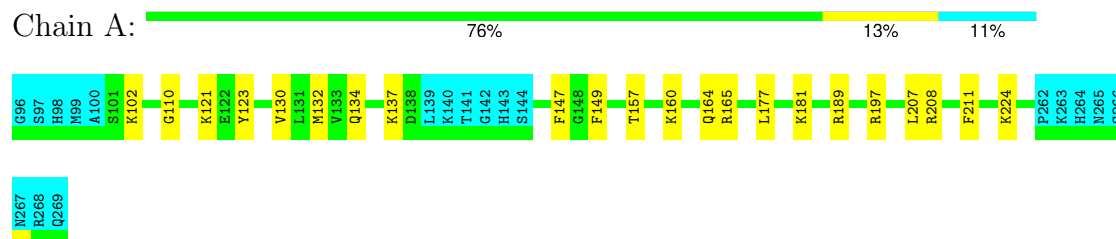
- Molecule 2: 5'-R(*GP*UP*GP*UP*GP*AP*AP*UP*GP*AP*AP*UP)-3'



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

The representative model is number 4. Colouring as in section 4.1 above.

- Molecule 1: TAR DNA-BINDING PROTEIN 43



- Molecule 2: 5'-R(*GP*UP*GP*UP*GP*AP*AP*UP*GP*AP*AP*UP)-3'



5 Refinement protocol and experimental data overview

The models were refined using the following method: *SIMULATED ANNEALING COMBINED WITH MOLECULAR DYNAMICS IN TORSION ANGLE SPACE*.

Of the 50 calculated structures, 20 were deposited, based on the following criterion: *LOWEST ENERGY AND RESTRAINT VIOLATION*.

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

Software name	Classification	Version
Amber	refinement	
CYANA	structure solution	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	2
Total number of shifts	2382
Number of shifts mapped to atoms	2375
Number of unparsed shifts	0
Number of shifts with mapping errors	7
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	92%

6 Model quality i

6.1 Standard geometry i

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	0.50±0.00	0±0/1283 (0.0± 0.0%)	0.74±0.01	0±0/1728 (0.0± 0.0%)
2	B	1.05±0.01	0±0/288 (0.0± 0.0%)	1.58±0.05	5±2/448 (1.0± 0.4%)
All	All	0.64	0/31420 (0.0%)	0.97	92/43520 (0.2%)

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

Mol	Chain	Chirality	Planarity
2	B	0.0±0.0	0.2±0.4
All	All	0	5

There are no bond-length outliers.

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	5	G	O4'-C1'-N9	7.52	114.22	108.20	11	10
2	B	5	G	C3'-C2'-C1'	-7.04	95.87	101.50	1	5
2	B	3	G	C1'-O4'-C4'	-6.95	104.34	109.90	3	12
2	B	3	G	O4'-C1'-N9	6.68	113.54	108.20	18	4
2	B	10	A	O4'-C1'-N9	6.52	113.41	108.20	8	4

There are no chirality outliers.

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

Mol	Chain	Res	Type	Group	Models (Total)
2	B	6	A	Sidechain	3
2	B	9	G	Sidechain	2

6.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	1256	1232	1231	5±2
2	B	257	112	130	3±1
All	All	30260	26880	27220	120

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 2.

5 of 43 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:228:ALA:HB2	2:B:6:A:C4	0.64	2.27	12	2
1:A:247:ASP:C	1:A:248:LEU:HD12	0.57	2.19	3	1
2:B:6:A:H1'	2:B:7:A:C8	0.56	2.36	12	5
1:A:149:PHE:CZ	2:B:5:G:C4	0.53	2.97	20	15
1:A:207:LEU:HD22	1:A:211:PHE:CE1	0.48	2.42	19	6

6.3 Torsion angles

6.3.1 Protein backbone

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	155/174 (89%)	138±1 (89±1%)	15±2 (10±1%)	2±1 (1±0%)	15	64
All	All	3100/3480 (89%)	2761 (89%)	304 (10%)	35 (1%)	15	64

All 5 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	130	VAL	20
1	A	228	ALA	5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	247	ASP	5
1	A	190	SER	3
1	A	215	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	141/157 (90%)	128±2 (91±1%)	13±2 (9±1%)	10	57
All	All	2820/3140 (90%)	2558 (91%)	262 (9%)	10	57

5 of 23 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	121	LYS	20
1	A	123	TYR	20
1	A	134	GLN	20
1	A	189	ARG	20
1	A	208	ARG	20

6.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers	Suiteness
2	B	11/12 (92%)	8±1 (72±12%)	3±1 (27±11%)	0.12±0.07
All	All	221/240 (92%)	159 (72%)	60 (27%)	0.11

The overall RNA backbone suiteness is 0.12.

5 of 11 unique RNA backbone outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
2	B	3	G	20
2	B	10	A	19
2	B	11	A	19
2	B	6	A	18

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
2	B	5	G	16

5 of 8 unique RNA pucker outliers are listed below:

Mol	Chain	Res	Type	Models (Total)
2	B	10	A	19
2	B	3	G	14
2	B	2	U	13
2	B	6	A	5
2	B	4	U	4

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 oligosaccharides in this entry.

6.6 Ligand geometry [i](#)

There are no ligands in this entry.

6.7 Other polymers [i](#)

There are no such molecules in this entry.

6.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 92% for the well-defined parts and 90% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *YP597CN-aug12.1_withRNA2_protein2.str.csh*

7.1.1 Bookkeeping

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	2197
Number of shifts mapped to atoms	2197
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	4

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	172	0.07 ± 0.11	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	162	0.03 ± 0.06	None needed (< 0.5 ppm)
$^{13}\text{C}'$	158	0.25 ± 0.10	None needed (< 0.5 ppm)
^{15}N	159	0.25 ± 0.17	None needed (< 0.5 ppm)

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 85%, i.e. 2002 atoms were assigned a chemical shift out of a possible 2368. 0 out of 23 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	758/774 (98%)	311/314 (99%)	300/310 (97%)	147/150 (98%)
Sidechain	1058/1178 (90%)	716/759 (94%)	324/372 (87%)	18/47 (38%)

Continued on next page...

Continued from previous page...

	Total	¹ H	¹³ C	¹⁵ N
Aromatic	186/196 (95%)	93/97 (96%)	91/94 (97%)	2/5 (40%)
Sugar	0/132 (0%)	0/72 (0%)	0/60 (0%)	0/0 (—%)
Base	0/88 (0%)	0/56 (0%)	0/20 (0%)	0/12 (0%)
Overall	2002/2368 (85%)	1120/1298 (86%)	715/856 (84%)	167/214 (78%)

7.1.4 Statistically unusual chemical shifts ⓘ

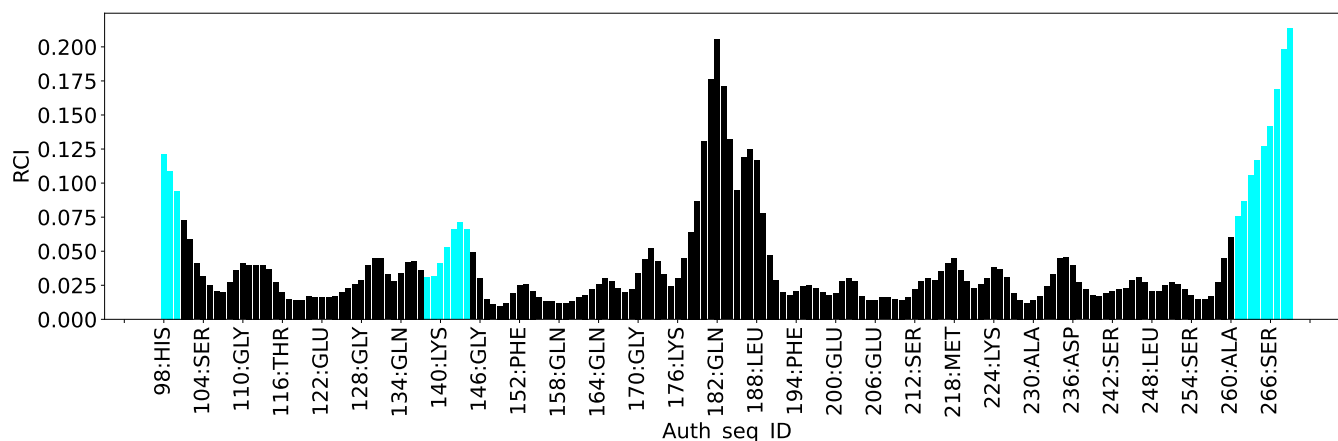
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	250	ILE	HG13	-1.28	-0.82 – 3.23	-6.1
1	A	167	MET	HB3	0.05	0.33 – 3.66	-5.8
1	A	168	ILE	HG13	-0.86	-0.82 – 3.23	-5.1
1	A	200	GLU	HB3	0.93	0.95 – 3.05	-5.1

7.1.5 Random Coil Index (RCI) plots ⓘ

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition. If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:



7.2 Chemical shift list 2

File name: working_cs.cif

Chemical shift list name: *YP597CN-aug12.1_withRNA2_RNA2.str.csh*

7.2.1 Bookkeeping

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	185
Number of shifts mapped to atoms	178
Number of unparsed shifts	0
Number of shifts with mapping errors	7
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	3

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 7) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
2	B	5	G	H1	12.173	0.020	1
2	B	5	G	H22	5.568	0.020	1
2	B	6	A	H2	7.805	0.020	1
2	B	7	A	H2	8.156	0.020	1
2	B	9	G	H1	10.046	0.020	1
2	B	10	A	H2	7.892	0.020	1
2	B	11	A	H2	8.104	0.020	1

7.2.2 Chemical shift referencing

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

7.2.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 7%, i.e. 166 atoms were assigned a chemical shift out of a possible 2368. 0 out of 23 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹ H	¹³ C	¹⁵ N
Backbone	0/774 (0%)	0/314 (0%)	0/310 (0%)	0/150 (0%)
Sidechain	0/1178 (0%)	0/759 (0%)	0/372 (0%)	0/47 (0%)
Aromatic	0/196 (0%)	0/97 (0%)	0/94 (0%)	0/5 (0%)
Sugar	131/132 (99%)	71/72 (99%)	60/60 (100%)	0/0 (—%)
Base	35/88 (40%)	19/56 (34%)	16/20 (80%)	0/12 (0%)
Overall	166/2368 (7%)	90/1298 (7%)	76/856 (9%)	0/214 (0%)

7.2.4 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
2	B	4	U	H4'	3.09	3.44 – 5.28	-6.9
2	B	4	U	H5''	3.01	3.05 – 5.24	-5.2
2	B	9	G	H5'	5.42	3.12 – 5.42	5.0

7.2.5 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

8.1 Conformationally restricting restraints

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

Description	Value
Total distance restraints	5131
Intra-residue ($ i-j =0$)	1060
Sequential ($ i-j =1$)	1256
Medium range ($ i-j >1$ and $ i-j <5$)	791
Long range ($ i-j \geq 5$)	1698
Inter-chain	214
Hydrogen bond restraints	112
Disulfide bond restraints	0
Total dihedral-angle restraints	686
Number of unmapped restraints	0
Number of restraints per residue	31.3
Number of long range restraints per residue ¹	9.4

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

8.2 Residual restraint violations

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

8.2.1 Average number of distance violations per model

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

Bins (Å)	Average number of violations per model	Max (Å)
0.1-0.2 (Small)	25.4	0.2
0.2-0.5 (Medium)	4.2	0.31
>0.5 (Large)	1.0	1.67

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)	0.7	3.31
10.0-20.0 (Medium)	None	None
>20.0 (Large)	169.0	109.61

9 Distance violation analysis ⓘ

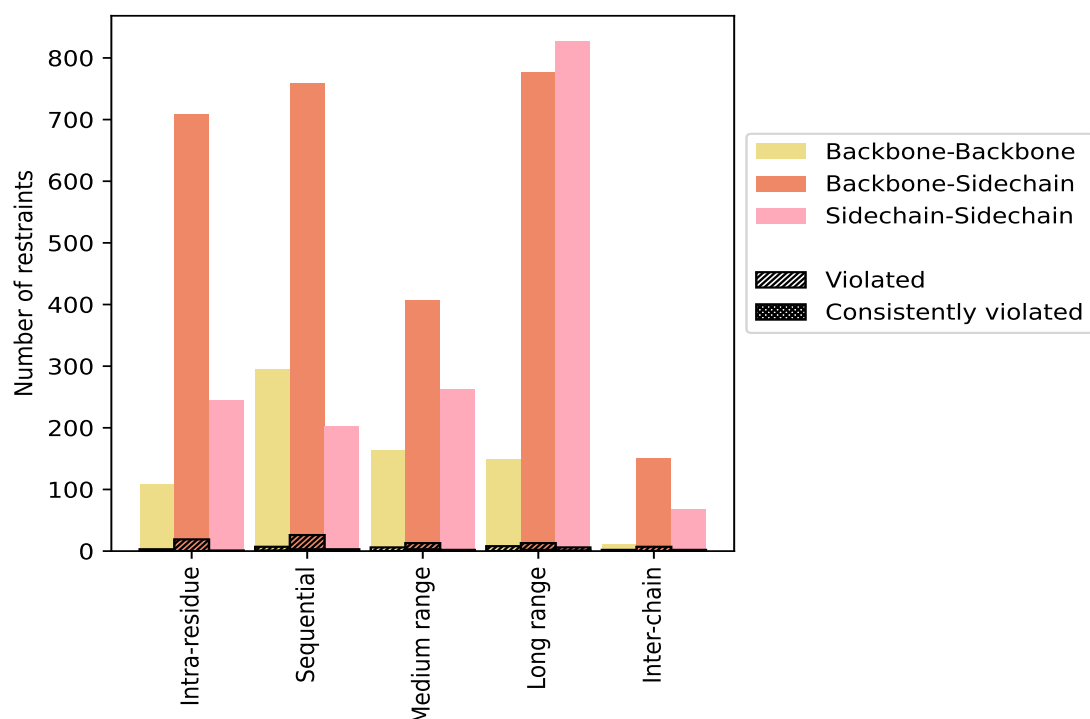
9.1 Summary of distance violations ⓘ

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

Restrains type	Count	% ¹	Violated ³			Consistently Violated ⁴		
			Count	% ²	% ¹	Count	% ²	% ¹
Intra-residue (i-j =0)	1060	20.7	23	2.2	0.4	0	0.0	0.0
Backbone-Backbone	108	2.1	3	2.8	0.1	0	0.0	0.0
Backbone-Sidechain	708	13.8	19	2.7	0.4	0	0.0	0.0
Sidechain-Sidechain	244	4.8	1	0.4	0.0	0	0.0	0.0
Sequential (i-j =1)	1256	24.5	36	2.9	0.7	4	0.3	0.1
Backbone-Backbone	295	5.7	7	2.4	0.1	1	0.3	0.0
Backbone-Sidechain	759	14.8	26	3.4	0.5	3	0.4	0.1
Sidechain-Sidechain	202	3.9	3	1.5	0.1	0	0.0	0.0
Medium range (i-j >1 & i-j <5)	791	15.4	19	2.4	0.4	1	0.1	0.0
Backbone-Backbone	163	3.2	6	3.7	0.1	0	0.0	0.0
Backbone-Sidechain	365	7.1	11	3.0	0.2	1	0.3	0.0
Sidechain-Sidechain	263	5.1	2	0.8	0.0	0	0.0	0.0
Long range (i-j ≥5)	1698	33.1	21	1.2	0.4	3	0.2	0.1
Backbone-Backbone	148	2.9	8	5.4	0.2	0	0.0	0.0
Backbone-Sidechain	723	14.1	7	1.0	0.1	2	0.3	0.0
Sidechain-Sidechain	827	16.1	6	0.7	0.1	1	0.1	0.0
Inter-chain	214	4.2	9	4.2	0.2	0	0.0	0.0
Backbone-Backbone	11	0.2	2	18.2	0.0	0	0.0	0.0
Backbone-Sidechain	147	2.9	6	4.1	0.1	0	0.0	0.0
Sidechain-Sidechain	56	1.1	1	1.8	0.0	0	0.0	0.0
Hydrogen bond	112	2.2	10	8.9	0.2	3	2.7	0.1
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	5131	100.0	118	2.3	2.3	11	0.2	0.2
Backbone-Backbone	725	14.1	26	3.6	0.5	1	0.1	0.0
Backbone-Sidechain	2802	54.6	78	2.8	1.5	8	0.3	0.2
Sidechain-Sidechain	1604	31.3	14	0.9	0.3	2	0.1	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 disulfid 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	7	8	4	9	3	31	0.19	1.21	0.19	0.14
2	8	9	4	7	5	33	0.2	1.67	0.26	0.15
3	3	8	5	8	7	31	0.19	1.36	0.22	0.13
4	1	9	7	8	7	32	0.18	1.18	0.18	0.15
5	5	13	6	8	7	39	0.18	1.54	0.22	0.13
6	6	11	7	9	4	37	0.18	1.67	0.25	0.13
7	6	12	3	9	2	32	0.2	1.65	0.27	0.14
8	7	10	7	9	5	38	0.18	1.53	0.23	0.12
9	8	7	5	8	7	35	0.17	1.21	0.18	0.13
10	5	9	8	6	3	31	0.19	1.18	0.19	0.16

Continued on next page...

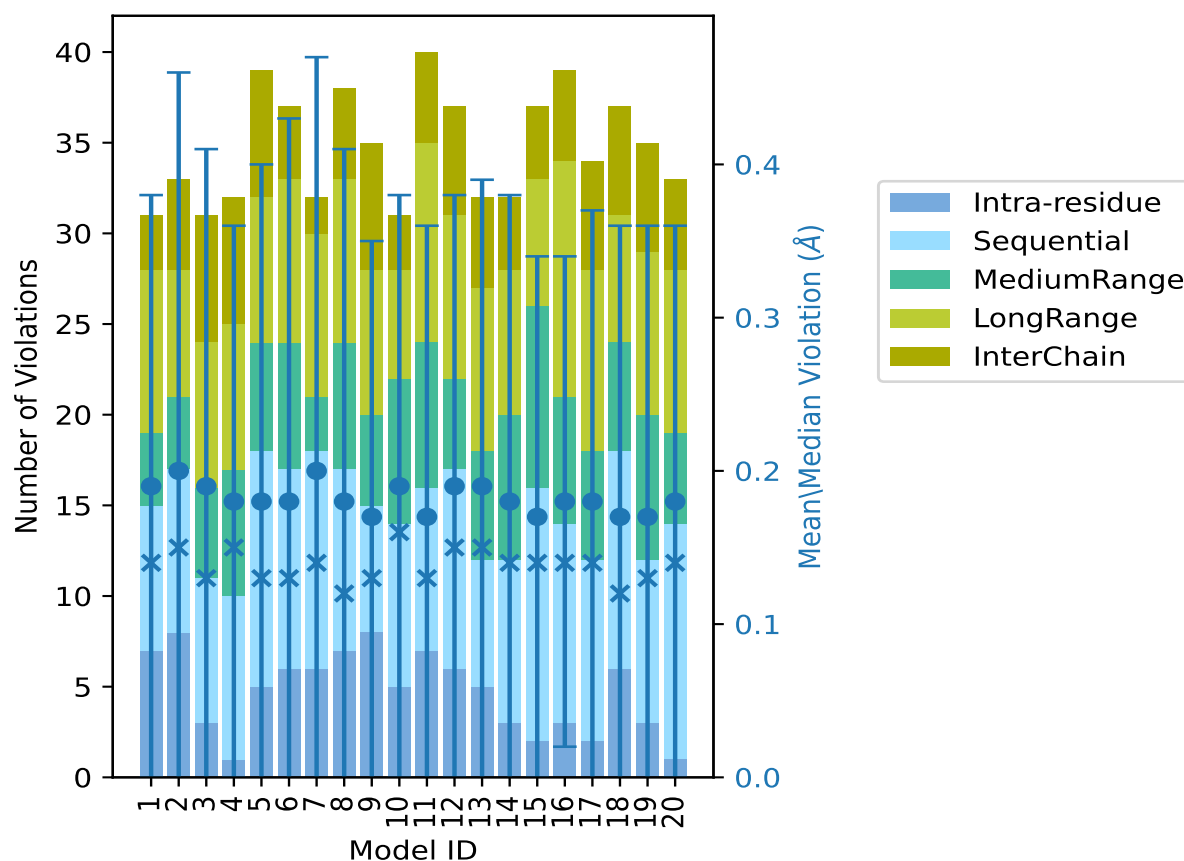
Continued from previous page...

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	7	9	8	11	5	40	0.17	1.31	0.19	0.13
12	6	11	5	9	6	37	0.19	1.3	0.19	0.15
13	5	7	6	9	5	32	0.19	1.28	0.2	0.15
14	3	9	8	8	4	32	0.18	1.29	0.2	0.14
15	2	14	10	7	4	37	0.17	1.19	0.17	0.14
16	3	11	7	13	5	39	0.18	1.15	0.16	0.14
17	2	10	6	10	6	34	0.18	1.25	0.19	0.14
18	6	12	6	7	6	37	0.17	1.29	0.19	0.12
19	3	9	8	9	6	35	0.17	1.24	0.19	0.13
20	1	13	5	9	5	33	0.18	1.2	0.18	0.14

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

⁵Inter-chain restraints, ⁶Standard deviation

9.2.1 Bar graph : Distance Violation statistics for each model ⓘ



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

9.3 Distance violation statistics for the ensemble

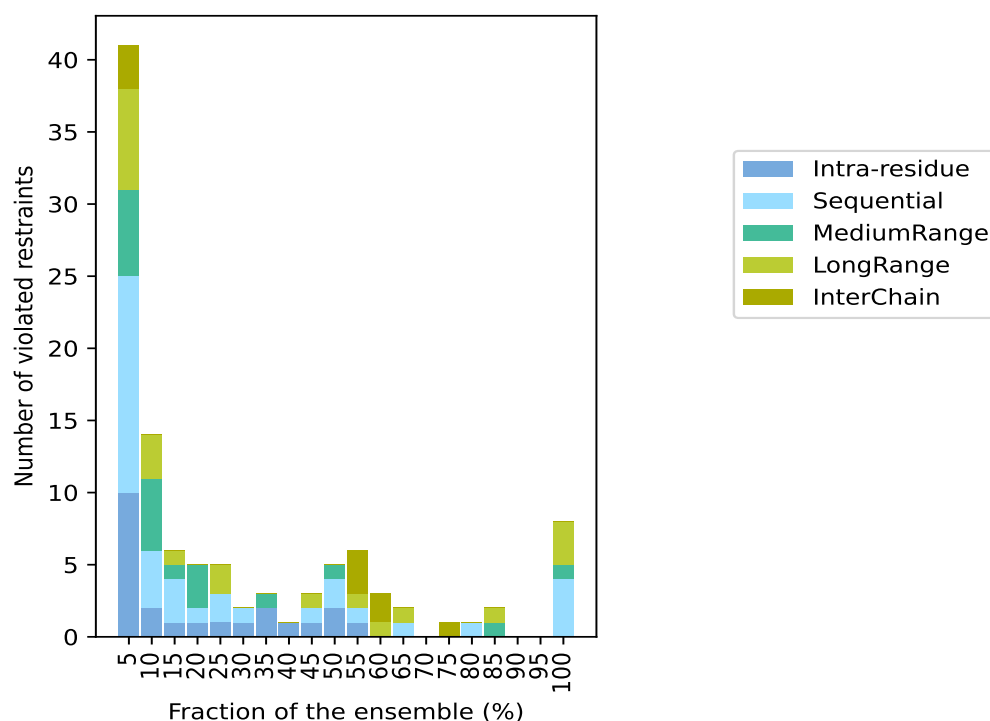
Violation analysis may find that some restraints are violated in few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of the ensemble. In total, 4911(IR:1037, SQ:1220, MR:772, LR:1677, IC:205) restraints 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	%
10	15	6	7	3	41	1	5.0
2	4	5	3	0	14	2	10.0
1	3	1	1	0	6	3	15.0
1	1	3	0	0	5	4	20.0
1	2	0	2	0	5	5	25.0
1	1	0	0	0	2	6	30.0
2	0	1	0	0	3	7	35.0
1	0	0	0	0	1	8	40.0
1	1	0	1	0	3	9	45.0
2	2	1	0	0	5	10	50.0
1	1	0	1	3	6	11	55.0
0	0	0	1	2	3	12	60.0
0	1	0	1	0	2	13	65.0
0	0	0	0	0	0	14	70.0
0	0	0	0	1	1	15	75.0
0	1	0	0	0	1	16	80.0
0	0	1	1	0	2	17	85.0
0	0	0	0	0	0	18	90.0
0	0	0	0	0	0	19	95.0
0	4	1	3	0	8	20	100.0

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

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

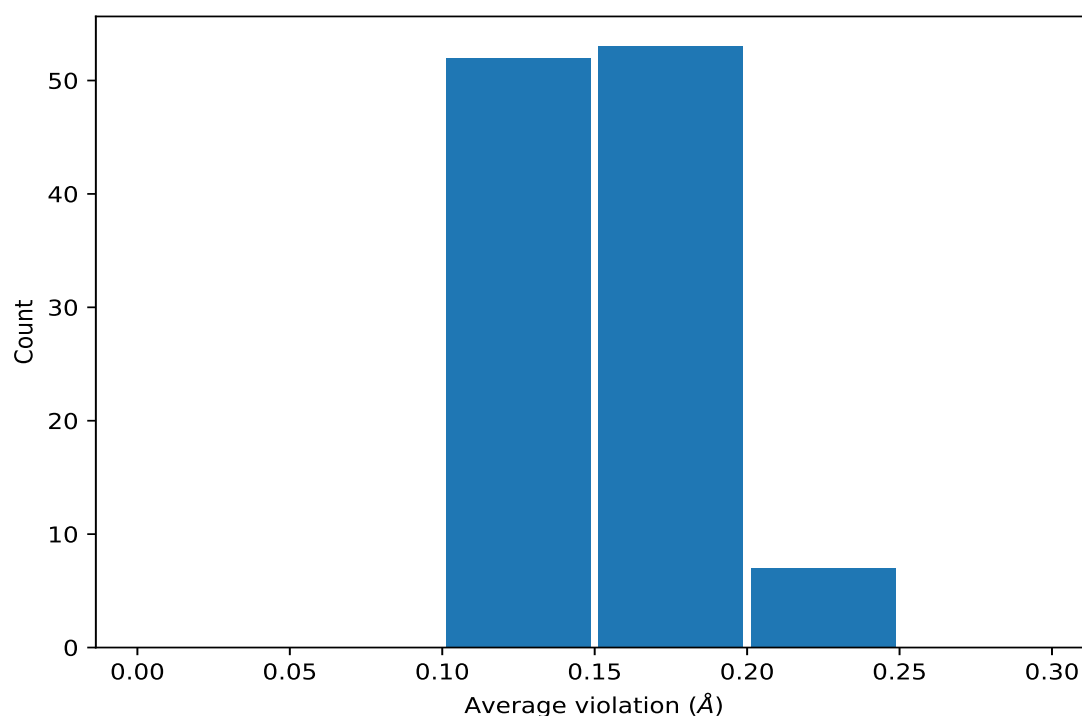
9.3.1 Bar graph : Distance violation statistics for the ensemble [i](#)



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

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

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



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

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

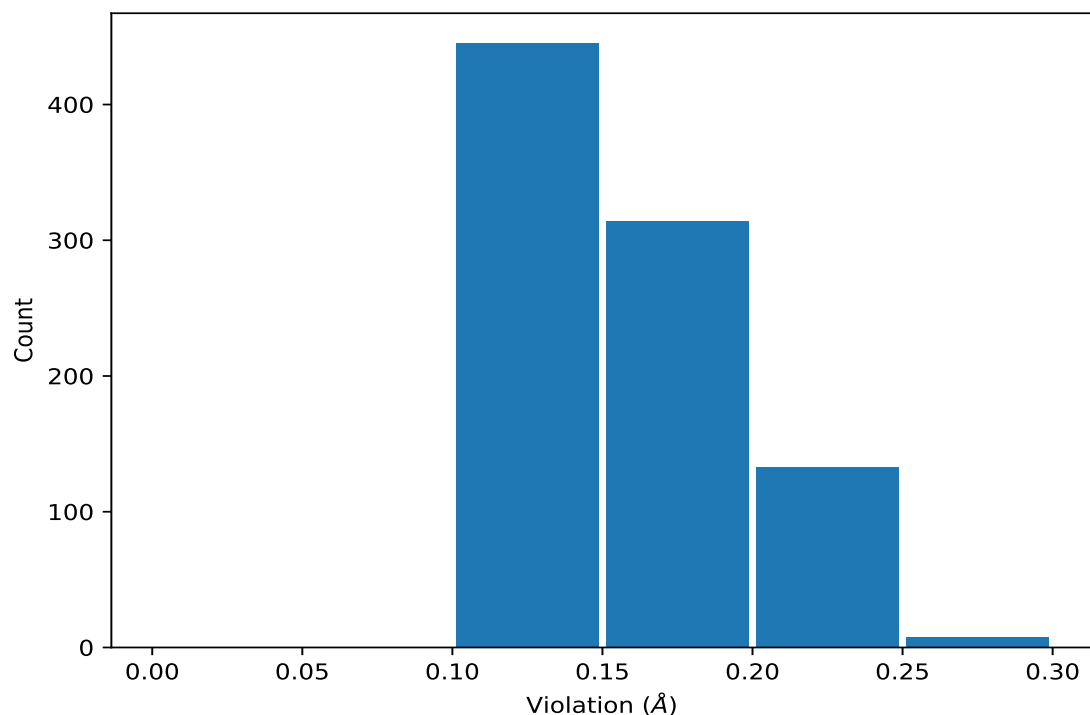
Key	Atom-1	Atom-2	Models ¹	Mean (Å)	SD ¹ (Å)	Median (Å)
(3,2)	1:121:A:LYS:H	1:117:A:GLU:O	20	0.23	0.02	0.23
(1,1877)	1:119:A:ASP:H	1:122:A:GLU:HG2	20	0.21	0.01	0.21
(1,2077)	1:140:A:LYS:HB2	1:141:A:THR:H	20	0.2	0.02	0.19
(1,813)	1:153:A:THR:HB	1:154:A:GLU:H	20	0.17	0.01	0.17
(3,12)	1:121:A:LYS:N	1:117:A:GLU:O	20	0.17	0.02	0.17
(1,2041)	1:208:A:ARG:HD2	1:217:A:VAL:HB	20	0.17	0.02	0.17
(1,791)	1:202:A:MET:H	1:253:A:ILE:HG13	20	0.16	0.01	0.16
(1,222)	1:135:A:VAL:H	1:136:A:LYS:H	20	0.15	0.01	0.16
(1,2144)	1:154:A:GLU:HB2	1:155:A:TYR:HA	20	0.13	0.01	0.13
(1,4740)	1:166:A:HIS:HE1	1:161:A:VAL:H	20	0.12	0.02	0.12

¹Number of violated models, ²Standard deviation

9.5 All violated distance restraints [i](#)

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

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



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

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

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(3,33)	1:147:A:PHE:H	1:136:A:LYS:O	6	0.27
(3,33)	1:147:A:PHE:H	1:136:A:LYS:O	12	0.26
(1,674)	1:207:A:LEU:H	1:207:A:LEU:HG	7	0.26
(3,2)	1:121:A:LYS:H	1:117:A:GLU:O	1	0.25
(3,2)	1:121:A:LYS:H	1:117:A:GLU:O	2	0.25
(3,2)	1:121:A:LYS:H	1:117:A:GLU:O	7	0.25
(3,2)	1:121:A:LYS:H	1:117:A:GLU:O	18	0.25
(1,4878)	2:5:B:G:H8	2:5:B:G:H4'	13	0.25
(3,33)	1:147:A:PHE:H	1:136:A:LYS:O	8	0.24
(3,2)	1:121:A:LYS:H	1:117:A:GLU:O	9	0.24

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

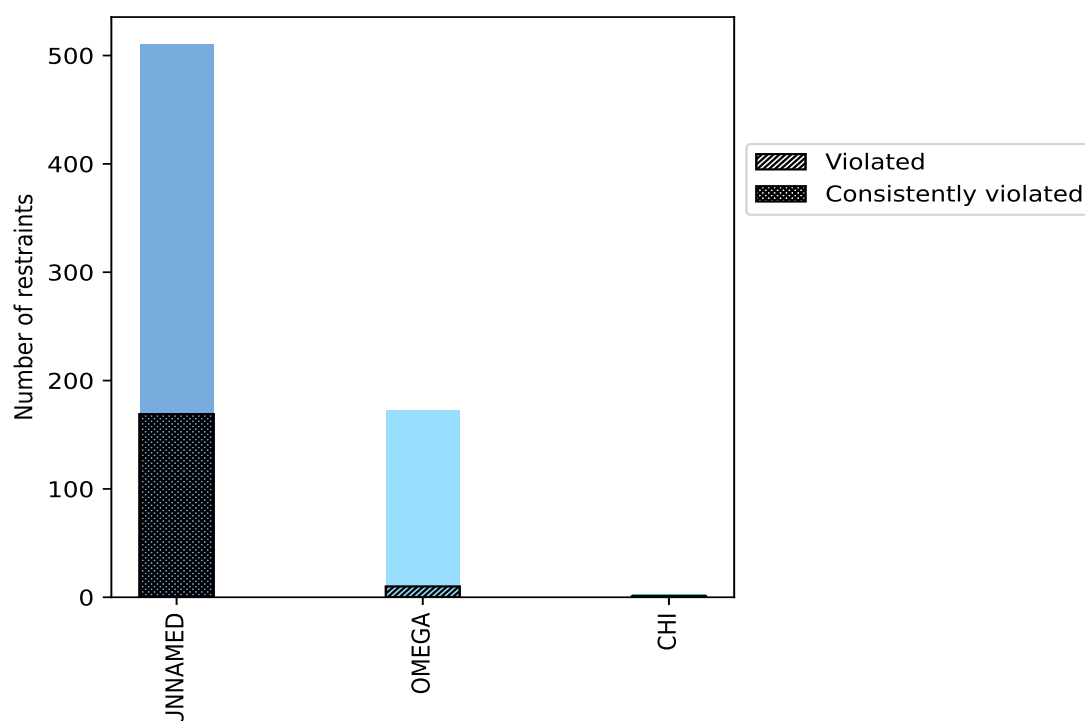
10.1 Summary of dihedral-angle violations [i](#)

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
UNNAMED	510	74.3	169	33.1	24.6	169	33.1	24.6
OMEGA	173	25.2	10	5.8	1.5	0	0.0	0.0
CHI	3	0.4	1	33.3	0.1	0	0.0	0.0
Total	686	100.0	180	26.2	26.2	169	24.6	24.6

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

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



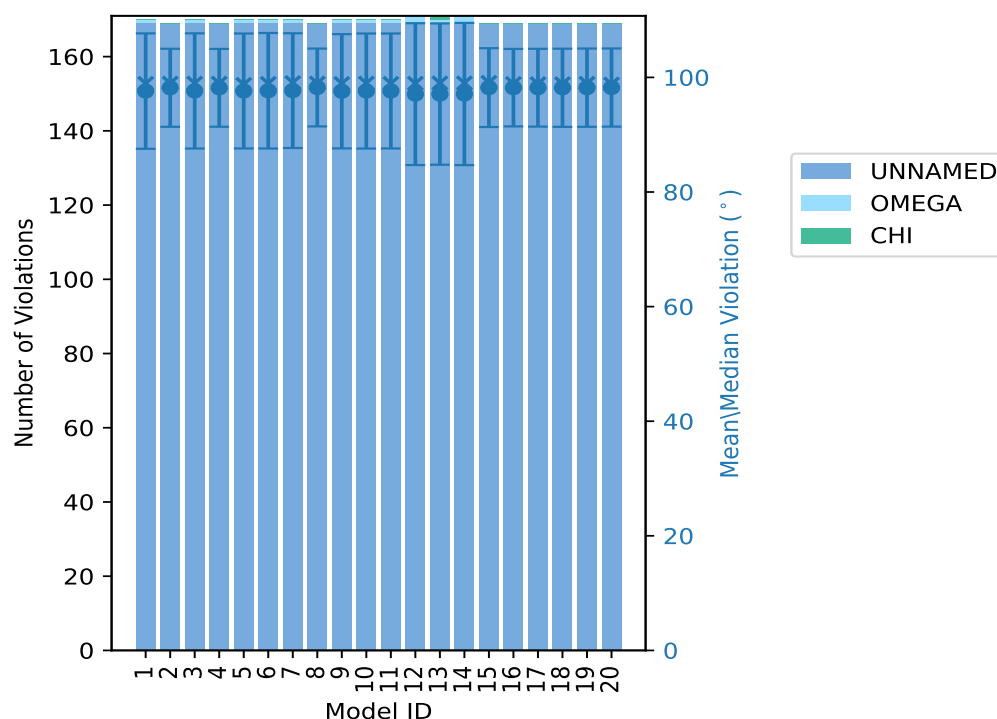
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

10.2 Dihedral-angle violation statistics for each model

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

Model ID	Number of violations				Mean (°)	Max (°)	SD (°)	Median (°)
	UNNAMED	OMEGA	CHI	Total				
1	169	1	0	170	97.61	108.35	10.07	98.94
2	169	0	0	169	98.19	109.31	6.81	98.99
3	169	1	0	170	97.64	108.46	10.05	98.99
4	169	0	0	169	98.18	107.78	6.79	98.98
5	169	1	0	170	97.63	108.08	10.03	98.77
6	169	1	0	170	97.67	109.55	10.08	98.94
7	169	1	0	170	97.69	109.46	10.01	99.08
8	169	0	0	169	98.24	107.82	6.79	99.0
9	169	1	0	170	97.58	108.47	9.97	98.93
10	169	1	0	170	97.62	108.63	10.04	98.99
11	169	1	0	170	97.63	107.67	10.03	98.94
12	169	2	0	171	97.08	109.47	12.38	98.96
13	169	1	1	171	97.09	108.26	12.33	99.02
14	169	2	0	171	97.1	108.56	12.41	99.02
15	169	0	0	169	98.22	109.61	6.88	99.12
16	169	0	0	169	98.21	108.23	6.76	98.93
17	169	0	0	169	98.2	108.03	6.79	98.89
18	169	0	0	169	98.19	108.19	6.82	98.89
19	169	0	0	169	98.21	109.31	6.83	98.95
20	169	0	0	169	98.23	108.39	6.82	98.83

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

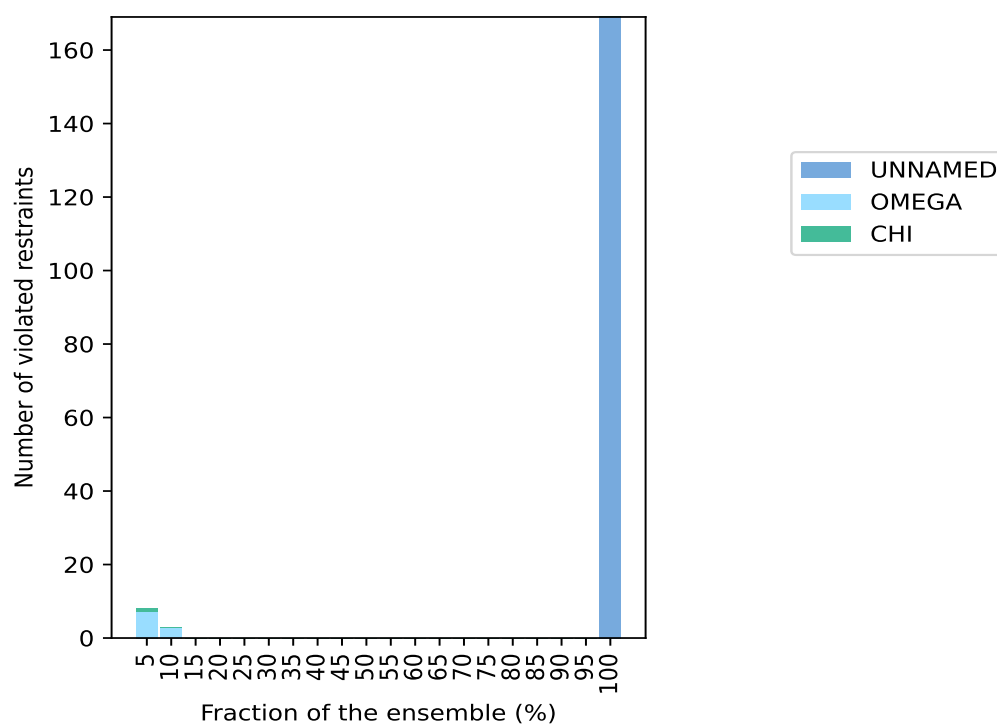
Continued on next page...

Continued from previous page...

Number of violated restraints				Fraction of the ensemble	
UNNAMED	OMEGA	CHI	Total	Count ¹	%
0	0	0	0	12	60.0
0	0	0	0	13	65.0
0	0	0	0	14	70.0
0	0	0	0	15	75.0
0	0	0	0	16	80.0
0	0	0	0	17	85.0
0	0	0	0	18	90.0
0	0	0	0	19	95.0
169	0	0	169	20	100.0

¹ Number of models with violations

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

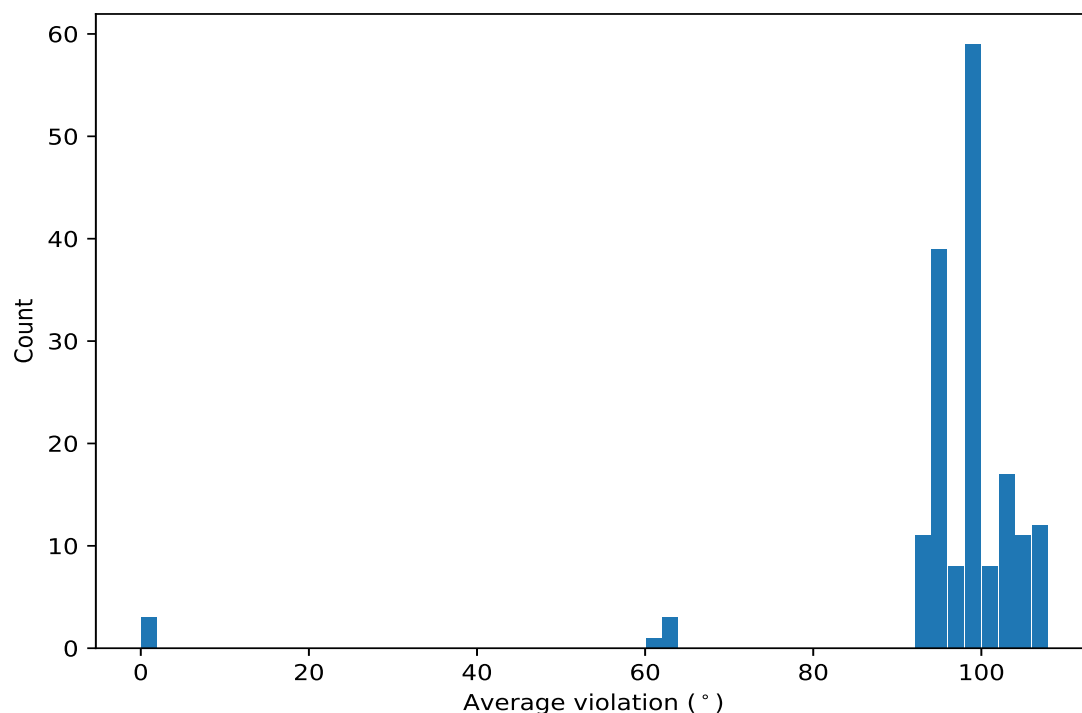


10.4 Most violated dihedral-angle restraints in the ensemble ⓘ

10.4.1 Histogram : Distribution of mean dihedral-angle violations ⓘ

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

in the ensemble



10.4.2 Table: Most violated dihedral-angle restraints [i](#)

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

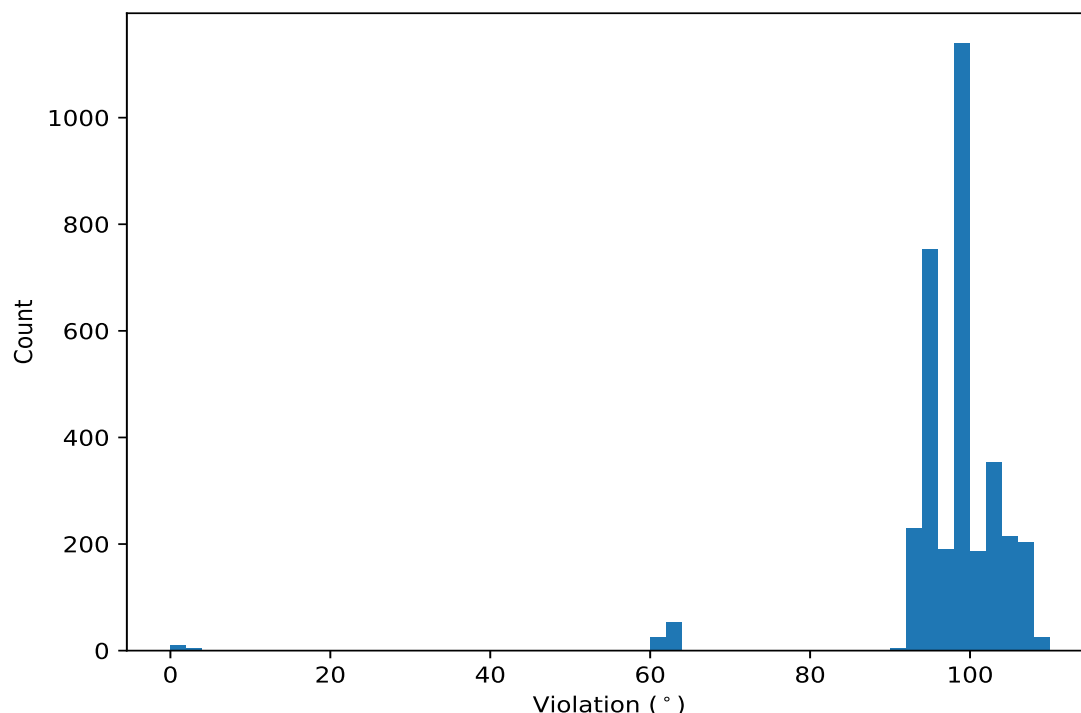
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models ¹	Mean	SD ²	Median
(1,506)	2:11:B:A:O3'	2:11:B:A:C2'	2:11:B:A:C4'	2:11:B:A:H3'	20	107.78	0.6	107.84
(1,471)	2:4:B:U:O3'	2:4:B:U:C2'	2:4:B:U:C4'	2:4:B:U:H3'	20	107.76	0.92	107.58
(1,511)	2:12:B:U:O3'	2:12:B:U:C2'	2:12:B:U:C4'	2:12:B:U:H3'	20	107.51	0.32	107.48
(1,476)	2:5:B:G:O3'	2:5:B:G:C2'	2:5:B:G:C4'	2:5:B:G:H3'	20	107.26	0.83	107.24
(1,486)	2:7:B:A:O3'	2:7:B:A:C2'	2:7:B:A:C4'	2:7:B:A:H3'	20	107.19	0.38	107.24
(1,496)	2:9:B:G:O3'	2:9:B:G:C2'	2:9:B:G:C4'	2:9:B:G:H3'	20	107.13	0.49	107.13
(1,481)	2:6:B:A:O3'	2:6:B:A:C2'	2:6:B:A:C4'	2:6:B:A:H3'	20	107.09	0.24	107.14
(1,456)	2:1:B:G:O3'	2:1:B:G:C2'	2:1:B:G:C4'	2:1:B:G:H3'	20	106.96	0.4	106.87
(1,501)	2:10:B:A:O3'	2:10:B:A:C2'	2:10:B:A:C4'	2:10:B:A:H3'	20	106.94	0.9	106.76
(1,461)	2:2:B:U:O3'	2:2:B:U:C2'	2:2:B:U:C4'	2:2:B:U:H3'	20	106.81	0.41	106.77

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

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

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

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



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

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

Key	Atom-1	Atom-2	Atom-3	Atom-4	Model ID	Violation (°)
(1,501)	2:10:B:A:O3'	2:10:B:A:C2'	2:10:B:A:C4'	2:10:B:A:H3'	15	109.61
(1,471)	2:4:B:U:O3'	2:4:B:U:C2'	2:4:B:U:C4'	2:4:B:U:H3'	6	109.55
(1,471)	2:4:B:U:O3'	2:4:B:U:C2'	2:4:B:U:C4'	2:4:B:U:H3'	12	109.47
(1,471)	2:4:B:U:O3'	2:4:B:U:C2'	2:4:B:U:C4'	2:4:B:U:H3'	7	109.46
(1,501)	2:10:B:A:O3'	2:10:B:A:C2'	2:10:B:A:C4'	2:10:B:A:H3'	19	109.31
(1,471)	2:4:B:U:O3'	2:4:B:U:C2'	2:4:B:U:C4'	2:4:B:U:H3'	2	109.31
(1,456)	2:1:B:G:O3'	2:1:B:G:C2'	2:1:B:G:C4'	2:1:B:G:H3'	10	108.63
(1,506)	2:11:B:A:O3'	2:11:B:A:C2'	2:11:B:A:C4'	2:11:B:A:H3'	12	108.61
(1,496)	2:9:B:G:O3'	2:9:B:G:C2'	2:9:B:G:C4'	2:9:B:G:H3'	14	108.56
(1,506)	2:11:B:A:O3'	2:11:B:A:C2'	2:11:B:A:C4'	2:11:B:A:H3'	19	108.49