wwPDB NMR Structure Validation Summary Report

Aug 26, 2024 – 10:26 AM EDT

PDB ID	:	8VRC
BMRB ID	:	31140
Title	:	Tetrahymena thermophila MLP1 RRM domain
Authors	:	Donaldson, L.W.
Deposited on	:	2024-01-21

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 (1)) were used in the production of this report:

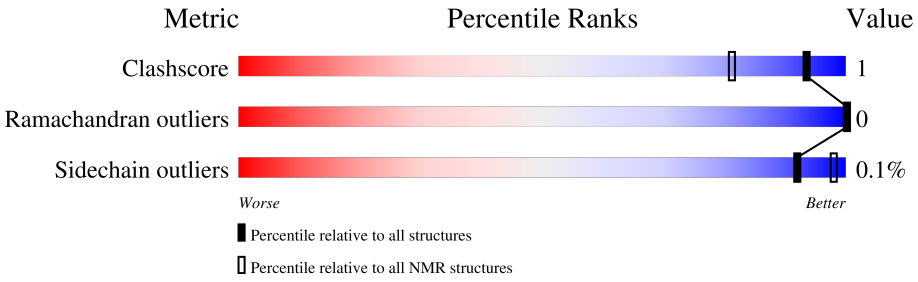
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.38.3

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 89%.

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



Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	210492	14027
Ramachandran outliers	207382	12486
Sidechain outliers	206894	12463

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

Mol	Chain	Length	Quality of chain
1	A	100	72% • 26%

2 Ensemble composition and analysis

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

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

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:35-A:108 (74)	0.29	20

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 8 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1579 atoms, of which 778 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called LA motif RNA-binding domain protein.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	100	1579	508	778	139	151	3	0

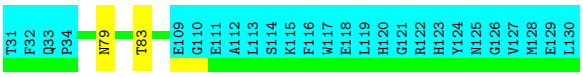
4 Residue-property plots

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: LA motif RNA-binding domain protein

Chain A:

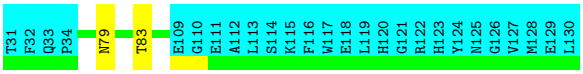


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 20. Colouring as in section 4.1 above.

- Molecule 1: LA motif RNA-binding domain protein

Chain A:



5 Refinement protocol and experimental data overview

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

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy.*

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

Software name	Classification	Version
CYANA	structure calculation	
Rosetta	refinement	

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

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	1440
Number of shifts mapped to atoms	1200
Number of unparsed shifts	0
Number of shifts with mapping errors	240
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	89%

6 Model quality

6.1 Standard geometry

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

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	584	579	579	1±1
All	All	11680	11580	11580	27

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:79:ASN:O	1:A:83:THR:N	0.49	2.46	14	20
1:A:66:TYR:C	1:A:66:TYR:CD2	0.48	2.87	17	2
1:A:79:ASN:OD1	1:A:80:LYS:N	0.44	2.50	3	2
1:A:57:LEU:O	1:A:61:ASN:N	0.41	2.53	15	3

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	74/100 (74%)	74±0 (100±0%)	0±0 (0±0%)	0±0 (0±0%)	100	100
All	All	1480/2000 (74%)	1480 (100%)	0 (0%)	0 (0%)	100	100

There are no Ramachandran outliers.

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	64/86 (74%)	64±0 (100±0%)	0±0 (0±0%)	92	98
All	All	1280/1720 (74%)	1279 (100%)	1 (0%)	92	98

All 1 unique residues with a non-rotameric sidechain are listed below.

Mol	Chain	Res	Type	Models (Total)
1	A	73	GLU	1

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 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 89% for the well-defined parts and 88% 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

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	1440
Number of shifts mapped to atoms	1200
Number of unparsed shifts	0
Number of shifts with mapping errors	240
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	5

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 240) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	5	ALA	H	8.296	0.024	1
1	A	5	ALA	HB1	1.4	0.000	1
1	A	5	ALA	HB2	1.4	0.000	1
1	A	5	ALA	HB3	1.4	0.000	1
1	A	5	ALA	CA	52.882	0.126	1
1	A	5	ALA	CB	19.415	0.123	1
1	A	5	ALA	N	124.95	0.000	1
1	A	6	SER	H	8.226	0.011	1
1	A	6	SER	CA	58.19	0.041	1
1	A	6	SER	CB	63.787	0.041	1
1	A	6	SER	N	114.605	0.000	1
1	A	7	MET	H	8.266	0.000	1
1	A	7	MET	HE1	2.005	0.000	1
1	A	7	MET	HE2	2.005	0.000	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	7	MET	HE3	2.005	0.000	1
1	A	7	MET	CA	55.31	0.000	1
1	A	7	MET	CB	33.329	0.081	1
1	A	7	MET	CE	16.951	0.000	1
1	A	7	MET	N	122.033	0.000	1
1	A	8	ASP	H	8.548	0.003	1
1	A	8	ASP	CA	52.553	0.000	1
1	A	8	ASP	CB	41.44	0.000	1
1	A	8	ASP	N	124.298	0.000	1
1	A	9	PRO	CA	64.31	0.000	1
1	A	9	PRO	CB	32.309	0.000	1
1	A	10	ALA	H	8.43	0.003	1
1	A	10	ALA	HB1	1.45	0.000	1
1	A	10	ALA	HB2	1.45	0.000	1
1	A	10	ALA	HB3	1.45	0.000	1
1	A	10	ALA	CA	54.144	0.051	1
1	A	10	ALA	CB	18.618	0.184	1
1	A	10	ALA	N	122.455	0.000	1
1	A	11	GLU	H	8.013	0.007	1
1	A	11	GLU	CA	57.94	0.034	1
1	A	11	GLU	CB	29.72	0.000	1
1	A	11	GLU	N	119.955	0.000	1
1	A	12	ALA	H	8.187	0.003	1
1	A	12	ALA	HB1	1.51	0.000	1
1	A	12	ALA	HB2	1.51	0.000	1
1	A	12	ALA	HB3	1.51	0.000	1
1	A	12	ALA	CA	54.506	0.088	1
1	A	12	ALA	CB	18.652	0.120	1
1	A	12	ALA	N	123.193	0.000	1
1	A	13	ALA	H	8.065	0.003	1
1	A	13	ALA	HB1	1.447	0.033	1
1	A	13	ALA	HB2	1.447	0.033	1
1	A	13	ALA	HB3	1.447	0.033	1
1	A	13	ALA	CA	54.418	0.000	1
1	A	13	ALA	CB	18.557	0.154	1
1	A	13	ALA	N	121.245	0.000	1
1	A	14	GLU	H	8.078	0.006	1
1	A	14	GLU	CA	59.06	0.090	1
1	A	14	GLU	CB	32.288	0.000	1
1	A	14	GLU	N	120.755	0.000	1
1	A	15	LYS	H	8.204	0.012	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	15	LYS	CA	58.788	0.000	1
1	A	15	LYS	N	119.555	0.000	1
1	A	19	GLU	CA	54.737	0.000	1
1	A	19	GLU	CB	29.633	0.000	1
1	A	20	ALA	H	8.138	0.002	1
1	A	20	ALA	CA	54.323	0.000	1
1	A	20	ALA	CB	18.322	0.089	1
1	A	20	ALA	N	122.994	0.000	1
1	A	21	GLU	H	8.209	0.005	1
1	A	21	GLU	CA	58.34	0.000	1
1	A	21	GLU	N	119.122	0.000	1
1	A	24	GLU	CA	58.407	0.000	1
1	A	25	LEU	H	8.072	0.009	1
1	A	25	LEU	HA	4.221	0.000	1
1	A	25	LEU	HB2	1.741	0.009	2
1	A	25	LEU	HB3	1.823	0.000	2
1	A	25	LEU	HD11	0.906	0.004	2
1	A	25	LEU	HD12	0.906	0.004	2
1	A	25	LEU	HD13	0.906	0.004	2
1	A	25	LEU	HD21	0.877	0.008	2
1	A	25	LEU	HD22	0.877	0.008	2
1	A	25	LEU	HD23	0.877	0.008	2
1	A	25	LEU	C	178.349	0.000	1
1	A	25	LEU	CA	57.436	0.035	1
1	A	25	LEU	CB	42.179	0.041	1
1	A	25	LEU	CD1	24.81	0.024	2
1	A	25	LEU	CD2	24.028	0.000	2
1	A	25	LEU	CG	26.938	0.000	1
1	A	25	LEU	N	122.813	0.000	1
1	A	26	ILE	H	8.287	0.011	1
1	A	26	ILE	HA	3.9	0.005	1
1	A	26	ILE	HB	1.916	0.004	1
1	A	26	ILE	HD11	0.896	0.006	1
1	A	26	ILE	HD12	0.896	0.006	1
1	A	26	ILE	HD13	0.896	0.006	1
1	A	26	ILE	HG12	1.701	0.007	1
1	A	26	ILE	HG21	0.97	0.038	1
1	A	26	ILE	HG22	0.97	0.038	1
1	A	26	ILE	HG23	0.97	0.038	1
1	A	26	ILE	C	177.96	0.000	1
1	A	26	ILE	CA	64.399	0.063	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	26	ILE	CB	38.25	0.015	1
1	A	26	ILE	CD1	13.642	0.029	1
1	A	26	ILE	CG1	29.068	0.000	1
1	A	26	ILE	CG2	17.436	0.025	1
1	A	26	ILE	N	119.857	0.000	1
1	A	27	ASN	H	8.149	0.011	1
1	A	27	ASN	HA	4.555	0.012	1
1	A	27	ASN	HB2	2.834	0.005	1
1	A	27	ASN	HD21	7.77	0.016	1
1	A	27	ASN	HD22	7.062	0.030	1
1	A	27	ASN	C	176.764	0.016	1
1	A	27	ASN	CA	56.165	0.089	1
1	A	27	ASN	CB	39.056	0.000	1
1	A	27	ASN	N	118.966	0.065	1
1	A	27	ASN	ND2	113.045	0.000	1
1	A	28	PHE	H	8.417	0.009	1
1	A	28	PHE	HA	4.192	0.024	1
1	A	28	PHE	HB2	3.2	0.007	2
1	A	28	PHE	HB3	3.249	0.006	2
1	A	28	PHE	HD1	6.734	0.004	1
1	A	28	PHE	HD2	6.734	0.004	1
1	A	28	PHE	HE1	7.17	0.003	1
1	A	28	PHE	HE2	7.17	0.003	1
1	A	28	PHE	C	176.655	0.020	1
1	A	28	PHE	CA	60.612	0.027	1
1	A	28	PHE	CB	38.872	0.036	1
1	A	28	PHE	CD1	133.323	0.000	1
1	A	28	PHE	CD2	133.323	0.000	1
1	A	28	PHE	CE1	130.886	0.000	1
1	A	28	PHE	CE2	130.886	0.000	1
1	A	28	PHE	N	122.235	0.000	1
1	A	29	TYR	H	7.981	0.013	1
1	A	29	TYR	HA	4.235	0.026	1
1	A	29	TYR	HB2	2.903	0.014	2
1	A	29	TYR	HB3	3.07	0.010	2
1	A	29	TYR	HD1	7.112	0.028	1
1	A	29	TYR	HD2	7.112	0.028	1
1	A	29	TYR	HE1	6.688	0.008	1
1	A	29	TYR	HE2	6.688	0.008	1
1	A	29	TYR	C	177.733	0.000	1
1	A	29	TYR	CA	61.133	0.000	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	29	TYR	CB	38.648	0.000	1
1	A	29	TYR	CD1	132.779	0.038	1
1	A	29	TYR	CD2	132.779	0.038	1
1	A	29	TYR	CE1	118.2	0.033	1
1	A	29	TYR	CE2	118.2	0.033	1
1	A	29	TYR	N	115.54	0.024	1
1	A	30	GLU	H	8.581	0.018	1
1	A	30	GLU	HA	4.745	0.010	1
1	A	30	GLU	HB2	2.224	0.003	2
1	A	30	GLU	HB3	2.156	0.021	2
1	A	30	GLU	HG2	2.424	0.000	2
1	A	30	GLU	HG3	2.348	0.006	2
1	A	30	GLU	C	177.884	0.000	1
1	A	30	GLU	CA	58.939	0.016	1
1	A	30	GLU	CB	30.356	0.000	1
1	A	30	GLU	CG	36.45	0.000	1
1	A	30	GLU	N	120.286	0.000	1
1	A	131	LYS	H	7.959	0.005	1
1	A	131	LYS	HA	3.951	0.000	1
1	A	131	LYS	HB2	1.618	0.000	2
1	A	131	LYS	HB3	1.315	0.000	2
1	A	131	LYS	HG2	1.758	0.000	1
1	A	131	LYS	C	178.537	0.015	1
1	A	131	LYS	CA	58.026	0.155	1
1	A	131	LYS	CB	30.697	0.060	1
1	A	131	LYS	CD	28.044	0.053	1
1	A	131	LYS	CG	24.927	0.000	1
1	A	131	LYS	N	119.383	0.000	1
1	A	132	LYS	H	8.25	0.000	1
1	A	132	LYS	HA	3.767	0.004	1
1	A	132	LYS	C	178.112	0.046	1
1	A	132	LYS	CA	59.002	0.336	1
1	A	132	LYS	CB	32.431	0.085	1
1	A	132	LYS	CD	29.677	0.000	1
1	A	132	LYS	CE	42.241	0.000	1
1	A	132	LYS	CG	25.693	0.000	1
1	A	132	LYS	N	118.011	0.000	1
1	A	133	LYS	H	7.775	0.006	1
1	A	133	LYS	HA	4.121	0.007	1
1	A	133	LYS	HB2	1.882	0.010	2
1	A	133	LYS	HB3	1.893	0.000	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	133	LYS	C	179.877	0.000	1
1	A	133	LYS	CA	58.378	0.003	1
1	A	133	LYS	CB	32.616	0.000	1
1	A	133	LYS	CD	29.524	0.000	1
1	A	133	LYS	CE	45.261	0.000	1
1	A	133	LYS	CG	25.234	0.000	1
1	A	133	LYS	N	119.089	0.000	1
1	A	134	GLU	H	8.045	0.006	1
1	A	134	GLU	HA	4.013	0.030	1
1	A	134	GLU	HB2	2.101	0.020	2
1	A	134	GLU	HB3	2.069	0.000	2
1	A	134	GLU	HG2	2.395	0.001	2
1	A	134	GLU	HG3	2.394	0.000	2
1	A	134	GLU	C	178.393	0.000	1
1	A	134	GLU	CA	58.381	0.000	1
1	A	134	GLU	CB	30.129	0.000	1
1	A	134	GLU	CG	36.566	0.005	1
1	A	134	GLU	N	119.306	0.106	1
1	A	135	VAL	H	7.741	0.009	1
1	A	135	VAL	HA	4.059	0.018	1
1	A	135	VAL	HB	2.216	0.022	1
1	A	135	VAL	HG11	1.113	0.000	2
1	A	135	VAL	HG12	1.113	0.000	2
1	A	135	VAL	HG13	1.113	0.000	2
1	A	135	VAL	HG21	1.058	0.011	2
1	A	135	VAL	HG22	1.058	0.011	2
1	A	135	VAL	HG23	1.058	0.011	2
1	A	135	VAL	C	176.407	0.019	1
1	A	135	VAL	CA	63.951	0.001	1
1	A	135	VAL	CB	31.964	0.032	1
1	A	135	VAL	CG1	21.414	0.000	2
1	A	135	VAL	CG2	21.992	0.000	2
1	A	135	VAL	N	116.37	0.000	1
1	A	136	ASN	H	7.889	0.012	1
1	A	136	ASN	HA	4.708	0.013	1
1	A	136	ASN	HB2	2.835	0.010	2
1	A	136	ASN	HB3	2.757	0.008	2
1	A	136	ASN	C	175.183	0.048	1
1	A	136	ASN	CA	54.101	0.057	1
1	A	136	ASN	CB	39.48	0.000	1
1	A	136	ASN	N	118.517	0.000	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	137	GLN	H	7.924	0.005	1
1	A	137	GLN	HA	4.393	0.008	1
1	A	137	GLN	HB2	2.177	0.012	2
1	A	137	GLN	HB3	2.053	0.007	2
1	A	137	GLN	HG2	2.415	0.002	1
1	A	137	GLN	C	175.347	0.000	1
1	A	137	GLN	CA	56.191	0.001	1
1	A	137	GLN	CB	29.532	0.000	1
1	A	137	GLN	CG	34.044	0.000	1
1	A	137	GLN	N	119.742	0.000	1
1	A	138	THR	H	7.797	0.006	1
1	A	138	THR	HA	4.146	0.003	1
1	A	138	THR	HB	4.244	0.009	1
1	A	138	THR	HG21	1.196	0.007	1
1	A	138	THR	HG22	1.196	0.007	1
1	A	138	THR	HG23	1.196	0.007	1
1	A	138	THR	C	179.202	0.000	1
1	A	138	THR	CA	63.547	0.088	1
1	A	138	THR	CB	70.834	0.019	1
1	A	138	THR	CG2	22.126	0.006	1
1	A	138	THR	N	120.428	0.000	1

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	127	-0.25 ± 0.29	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	115	-0.08 ± 0.16	None needed (< 0.5 ppm)
$^{13}C'$	113	-0.08 ± 0.23	None needed (< 0.5 ppm)
^{15}N	124	0.16 ± 0.37	None needed (< 0.5 ppm)

7.1.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 89%, i.e. 896 atoms were assigned a chemical shift out of a possible 1009. 0 out of 12 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	370/373 (99%)	150/152 (99%)	147/148 (99%)	73/73 (100%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Sidechain	472/561 (84%)	319/362 (88%)	144/176 (82%)	9/23 (39%)
Aromatic	54/75 (72%)	29/38 (76%)	24/34 (71%)	1/3 (33%)
Overall	896/1009 (89%)	498/552 (90%)	315/358 (88%)	83/99 (84%)

7.1.4 Statistically unusual chemical shifts (i)

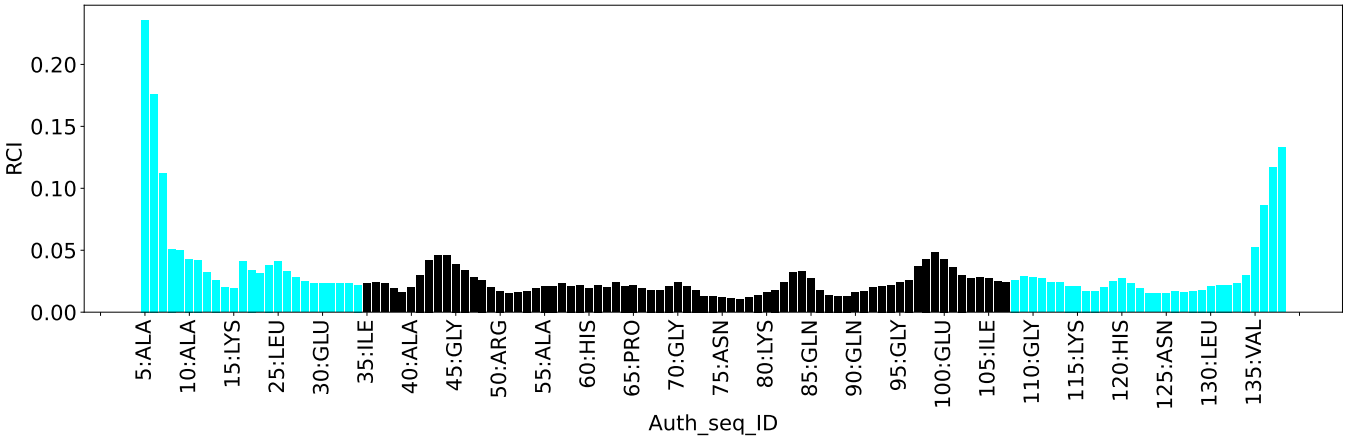
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	37	PHE	HB3	0.80	1.03 – 4.85	-5.6
1	A	47	ALA	HB1	0.02	0.14 – 2.58	-5.5
1	A	47	ALA	HB2	0.02	0.14 – 2.58	-5.5
1	A	47	ALA	HB3	0.02	0.14 – 2.58	-5.5
1	A	75	ASN	CB	47.70	30.50 – 46.89	5.5

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:



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	4510
Intra-residue (|i-j|=0)	1854
Sequential (|i-j|=1)	1063
Medium range (|i-j|>1 and |i-j|<5)	576
Long range (|i-j|≥5)	864
Inter-chain	0
Hydrogen bond restraints	153
Disulfide bond restraints	0
Total dihedral-angle restraints	376
Number of unmapped restraints	319
Number of restraints per residue	35.4
Number of long range restraints per residue[1]	6.6

[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

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)	45.9	0.2
0.2-0.5 (Medium)	81.2	0.5
>0.5 (Large)	125.3	3.97

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)	10.1	9.72
10.0-20.0 (Medium)	6.6	19.91
>20.0 (Large)	174.4	179.56

9 Distance violation analysis (i)

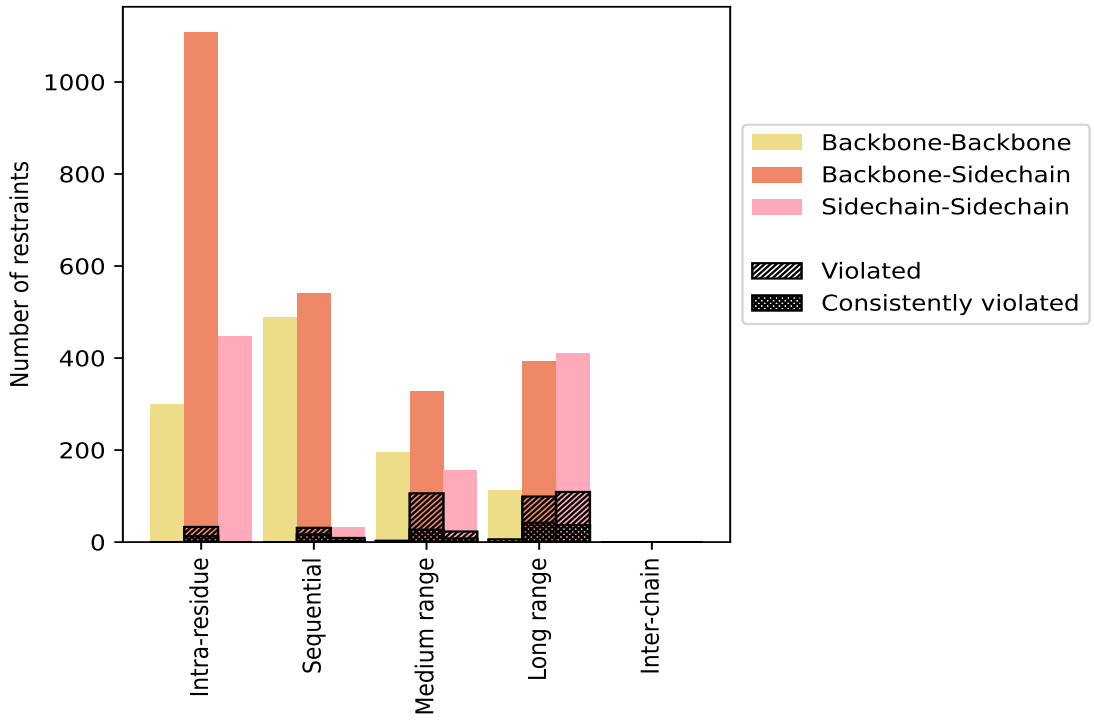
9.1 Summary of distance violations (i)

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)	1854	41.1	33	1.8	0.7	13	0.7	0.3
Backbone-Backbone	299	6.6	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	1108	24.6	33	3.0	0.7	13	1.2	0.3
Sidechain-Sidechain	447	9.9	0	0.0	0.0	0	0.0	0.0
Sequential (|i-j|=1)	1063	23.6	40	3.8	0.9	20	1.9	0.4
Backbone-Backbone	489	10.8	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	541	12.0	31	5.7	0.7	17	3.1	0.4
Sidechain-Sidechain	33	0.7	9	27.3	0.2	3	9.1	0.1
Medium range (|i-j|>1 & |i-j|<5)	576	12.8	86	14.9	1.9	28	4.9	0.6
Backbone-Backbone	194	4.3	3	1.5	0.1	1	0.5	0.0
Backbone-Sidechain	226	5.0	60	26.5	1.3	19	8.4	0.4
Sidechain-Sidechain	156	3.5	23	14.7	0.5	8	5.1	0.2
Long range (|i-j|≥5)	864	19.2	197	22.8	4.4	73	8.4	1.6
Backbone-Backbone	112	2.5	6	5.4	0.1	0	0.0	0.0
Backbone-Sidechain	341	7.6	82	24.0	1.8	36	10.6	0.8
Sidechain-Sidechain	411	9.1	109	26.5	2.4	37	9.0	0.8
Inter-chain	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Hydrogen bond	153	3.4	63	41.2	1.4	14	9.2	0.3
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	4510	100.0	419	9.3	9.3	148	3.3	3.3
Backbone-Backbone	1094	24.3	9	0.8	0.2	1	0.1	0.0
Backbone-Sidechain	2369	52.5	269	11.4	6.0	99	4.2	2.2
Sidechain-Sidechain	1047	23.2	141	13.5	3.1	48	4.6	1.1

[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



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

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	19	36	83	147	0	285	0.69	3.66	0.65	0.49
2	17	30	79	131	0	257	0.73	3.37	0.66	0.49
3	16	30	70	122	0	238	0.72	3.83	0.65	0.5
4	18	29	75	139	0	261	0.71	3.54	0.64	0.47
5	17	33	71	135	0	256	0.71	3.68	0.68	0.5
6	18	31	79	135	0	263	0.74	3.73	0.65	0.54
7	15	28	75	128	0	246	0.64	3.51	0.57	0.45
8	17	29	79	124	0	249	0.7	3.76	0.63	0.48
9	16	32	69	131	0	248	0.75	3.86	0.68	0.54
10	15	30	76	137	0	258	0.68	3.66	0.63	0.48

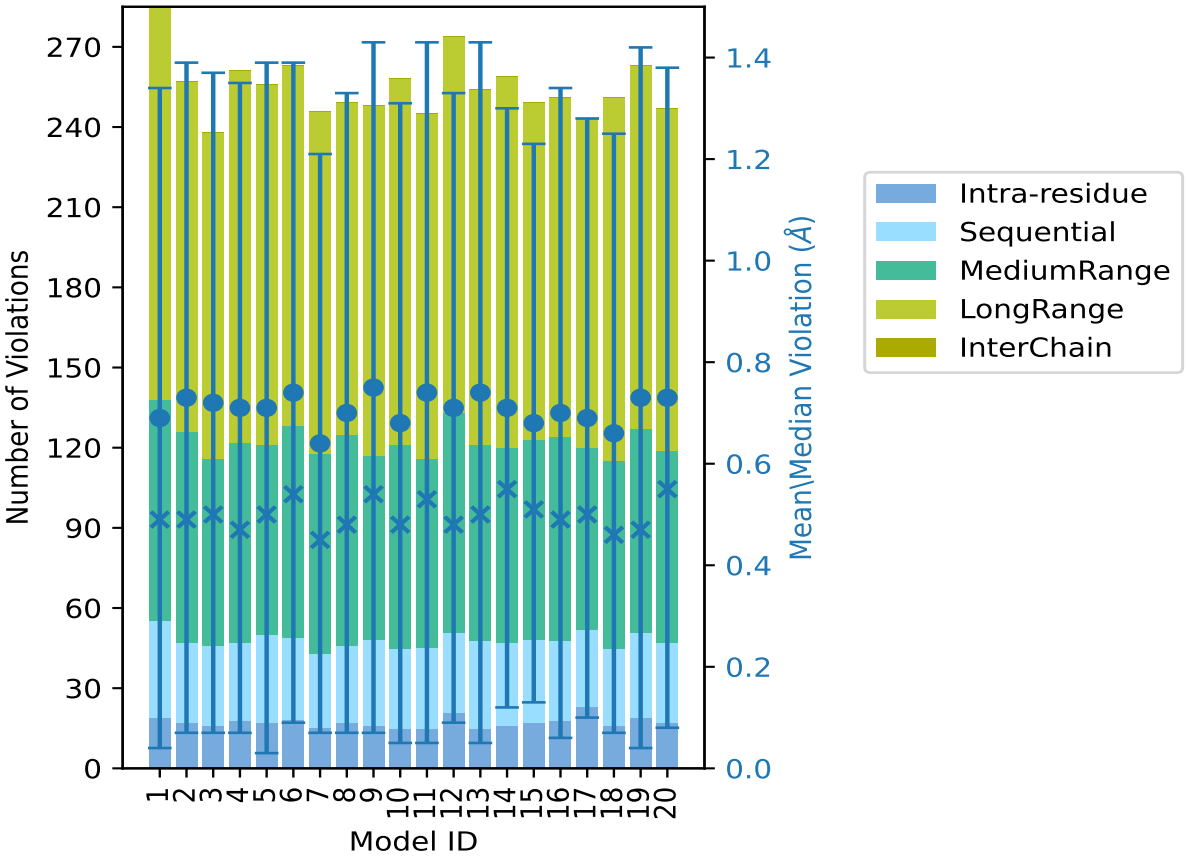
Continued on next page...

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	15	30	71	129	0	245	0.74	3.96	0.69	0.53
12	21	30	82	141	0	274	0.71	3.28	0.62	0.48
13	15	33	73	133	0	254	0.74	3.97	0.69	0.5
14	16	31	73	139	0	259	0.71	3.33	0.59	0.55
15	17	31	75	126	0	249	0.68	3.27	0.55	0.51
16	18	30	76	127	0	251	0.7	3.6	0.64	0.49
17	23	29	68	123	0	243	0.69	3.34	0.59	0.5
18	16	29	70	136	0	251	0.66	3.62	0.59	0.46
19	19	32	76	136	0	263	0.73	3.59	0.69	0.47
20	17	30	72	128	0	247	0.73	3.82	0.65	0.55

[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

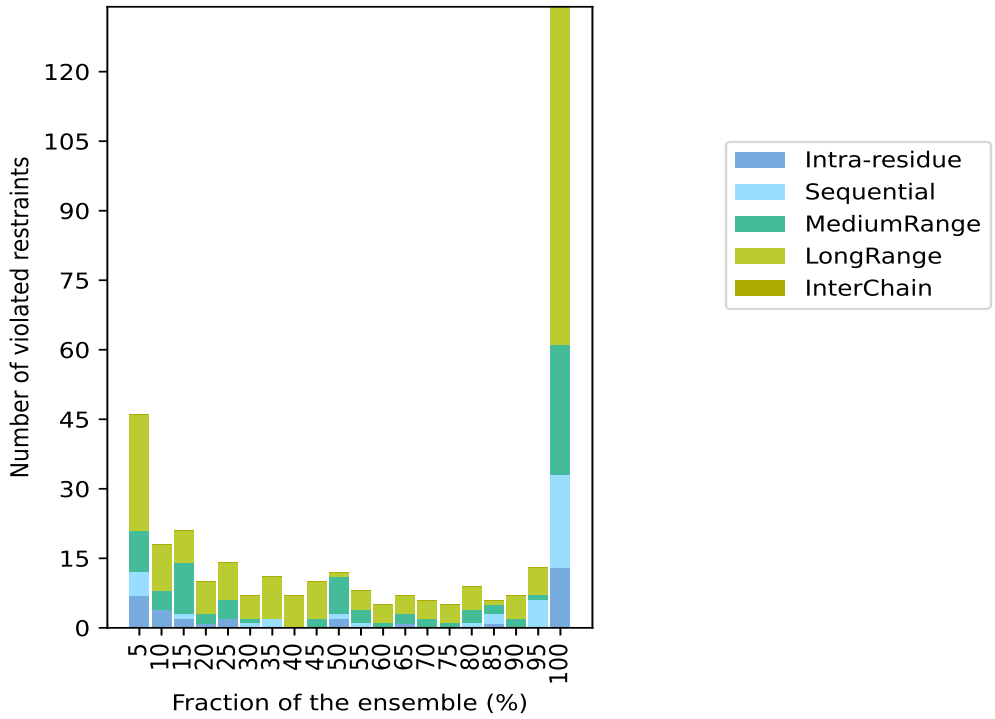
9.3 Distance violation statistics for the ensemble (i)

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, 4001(IR:1821, SQ:1023, MR:490, LR:667, IC:0) restraints are not violated in the ensemble.

Number of violated restraints						Fraction of the ensemble	
IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total	Count[6]	%
7	5	9	25	0	46	1	5.0
4	0	4	10	0	18	2	10.0
2	1	11	7	0	21	3	15.0
1	0	2	7	0	10	4	20.0
2	0	4	8	0	14	5	25.0
0	1	1	5	0	7	6	30.0
0	2	0	9	0	11	7	35.0
0	0	0	7	0	7	8	40.0
0	0	2	8	0	10	9	45.0
2	1	8	1	0	12	10	50.0
0	1	3	4	0	8	11	55.0
0	0	1	4	0	5	12	60.0
1	0	2	4	0	7	13	65.0
0	0	2	4	0	6	14	70.0
0	0	1	4	0	5	15	75.0
0	1	3	5	0	9	16	80.0
1	2	2	1	0	6	17	85.0
0	0	2	5	0	7	18	90.0
0	6	1	6	0	13	19	95.0
13	20	28	73	0	134	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

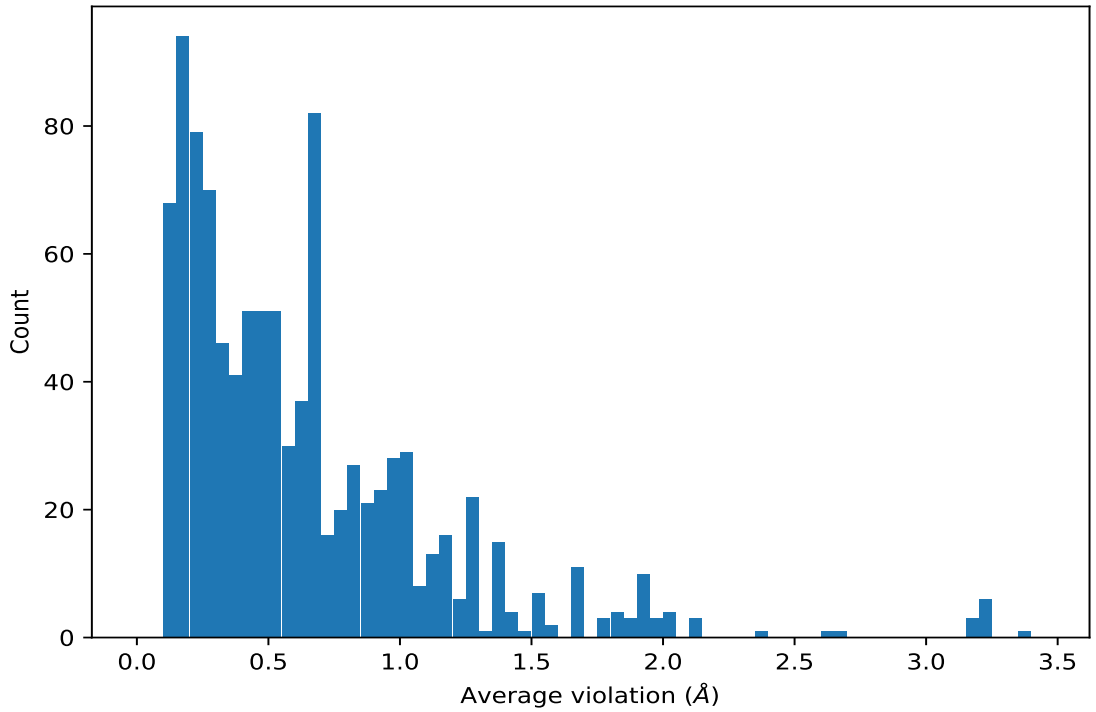
9.3.1 Bar graph : Distance violation statistics for the ensemble



9.4 Most violated distance restraints in the ensemble

9.4.1 Histogram : Distribution of mean distance 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



9.4.2 Table: Most violated distance restraints

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

Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(4,1794)	1:97:A:GLN:HG2	1:102:A:LYS:HG2	20	3.38	0.15	3.32
(4,2163)	1:62:A:VAL:HG11	1:56:A:LEU:H	20	3.16	0.03	3.17
(4,2163)	1:62:A:VAL:HG12	1:56:A:LEU:H	20	3.16	0.03	3.17
(4,2163)	1:62:A:VAL:HG13	1:56:A:LEU:H	20	3.16	0.03	3.17
(4,630)	1:51:A:ASN:H	1:54:A:GLU:HG3	20	2.69	0.03	2.7
(3,63)	1:79:A:ASN:H	1:82:A:LYS:O	20	2.65	0.09	2.64
(4,1795)	1:97:A:GLN:HG2	1:102:A:LYS:HE2	20	2.1	0.17	2.04
(4,1795)	1:97:A:GLN:HG2	1:102:A:LYS:HE3	20	2.1	0.17	2.04
(3,64)	1:79:A:ASN:N	1:82:A:LYS:O	20	2.05	0.08	2.05
(4,402)	1:116:A:PHE:H	1:119:A:LEU:HD11	20	2.05	0.06	2.05
(4,402)	1:116:A:PHE:H	1:119:A:LEU:HD12	20	2.05	0.06	2.05
(4,402)	1:116:A:PHE:H	1:119:A:LEU:HD13	20	2.05	0.06	2.05
(4,1211)	1:46:A:VAL:HG11	1:98:A:PHE:HA	20	1.98	0.44	1.81
(4,1211)	1:46:A:VAL:HG12	1:98:A:PHE:HA	20	1.98	0.44	1.81
(4,1211)	1:46:A:VAL:HG13	1:98:A:PHE:HA	20	1.98	0.44	1.81
(4,1218)	1:62:A:VAL:HG11	1:56:A:LEU:HA	20	1.94	0.03	1.95

Continued on next page...

Continued from previous page...

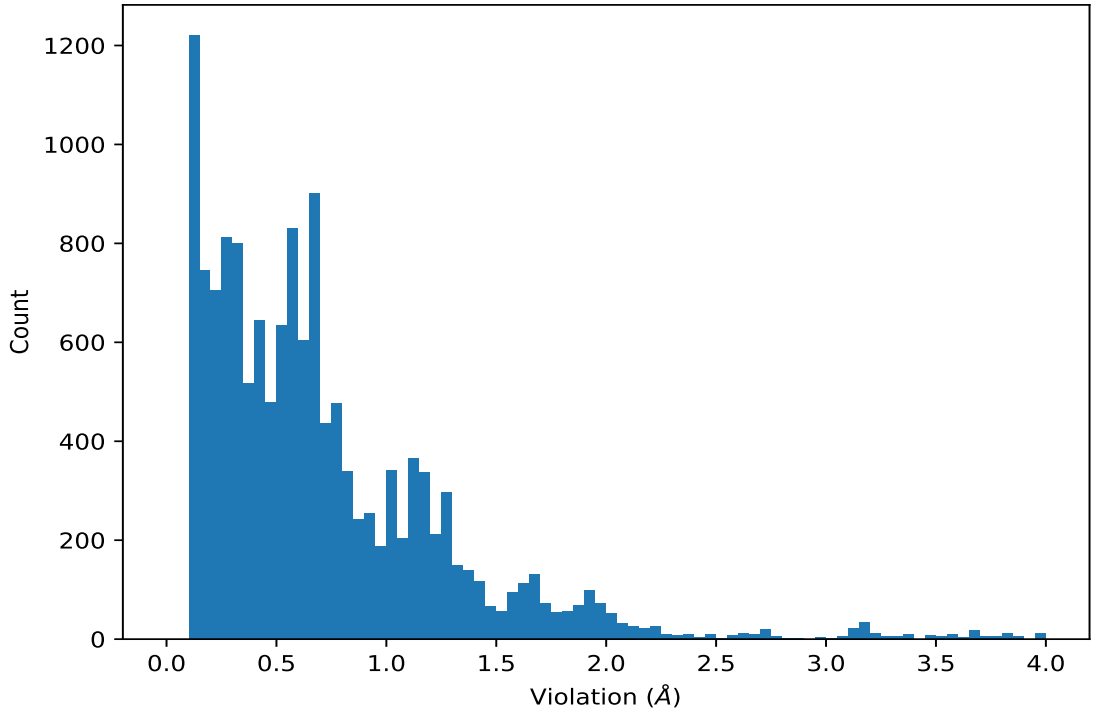
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(4,1218)	1:62:A:VAL:HG12	1:56:A:LEU:HA	20	1.94	0.03	1.95
(4,1218)	1:62:A:VAL:HG13	1:56:A:LEU:HA	20	1.94	0.03	1.95
(4,1433)	1:50:A:ARG:HD2	1:52:A:ILE:H	20	1.94	0.24	1.8

[1]Number of violated models, [2]Standard deviation

9.5 All violated distance restraints

9.5.1 Histogram : Distribution of distance violations

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

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

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(4,2133)	1:123:A:HIS:HD2	1:127:A:VAL:HG11	13	3.97
(4,2133)	1:123:A:HIS:HD2	1:127:A:VAL:HG12	13	3.97

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(4,2133)	1:123:A:HIS:HD2	1:127:A:VAL:HG13	13	3.97
(4,1283)	1:123:A:HIS:HD2	1:127:A:VAL:HG11	13	3.97
(4,1283)	1:123:A:HIS:HD2	1:127:A:VAL:HG12	13	3.97
(4,1283)	1:123:A:HIS:HD2	1:127:A:VAL:HG13	13	3.97
(4,2133)	1:123:A:HIS:HD2	1:127:A:VAL:HG11	11	3.96
(4,2133)	1:123:A:HIS:HD2	1:127:A:VAL:HG12	11	3.96
(4,2133)	1:123:A:HIS:HD2	1:127:A:VAL:HG13	11	3.96
(4,1283)	1:123:A:HIS:HD2	1:127:A:VAL:HG11	11	3.96
(4,1283)	1:123:A:HIS:HD2	1:127:A:VAL:HG12	11	3.96
(4,1283)	1:123:A:HIS:HD2	1:127:A:VAL:HG13	11	3.96
(4,2133)	1:123:A:HIS:HD2	1:127:A:VAL:HG11	9	3.86
(4,2133)	1:123:A:HIS:HD2	1:127:A:VAL:HG12	9	3.86
(4,2133)	1:123:A:HIS:HD2	1:127:A:VAL:HG13	9	3.86
(4,1283)	1:123:A:HIS:HD2	1:127:A:VAL:HG11	9	3.86
(4,1283)	1:123:A:HIS:HD2	1:127:A:VAL:HG12	9	3.86
(4,1283)	1:123:A:HIS:HD2	1:127:A:VAL:HG13	9	3.86
(4,2133)	1:123:A:HIS:HD2	1:127:A:VAL:HG11	3	3.83
(4,2133)	1:123:A:HIS:HD2	1:127:A:VAL:HG12	3	3.83
(4,2133)	1:123:A:HIS:HD2	1:127:A:VAL:HG13	3	3.83
(4,1283)	1:123:A:HIS:HD2	1:127:A:VAL:HG11	3	3.83
(4,1283)	1:123:A:HIS:HD2	1:127:A:VAL:HG12	3	3.83
(4,1283)	1:123:A:HIS:HD2	1:127:A:VAL:HG13	3	3.83
(4,2133)	1:123:A:HIS:HD2	1:127:A:VAL:HG11	20	3.82
(4,2133)	1:123:A:HIS:HD2	1:127:A:VAL:HG12	20	3.82
(4,2133)	1:123:A:HIS:HD2	1:127:A:VAL:HG13	20	3.82
(4,1283)	1:123:A:HIS:HD2	1:127:A:VAL:HG11	20	3.82

10 Dihedral-angle violation analysis

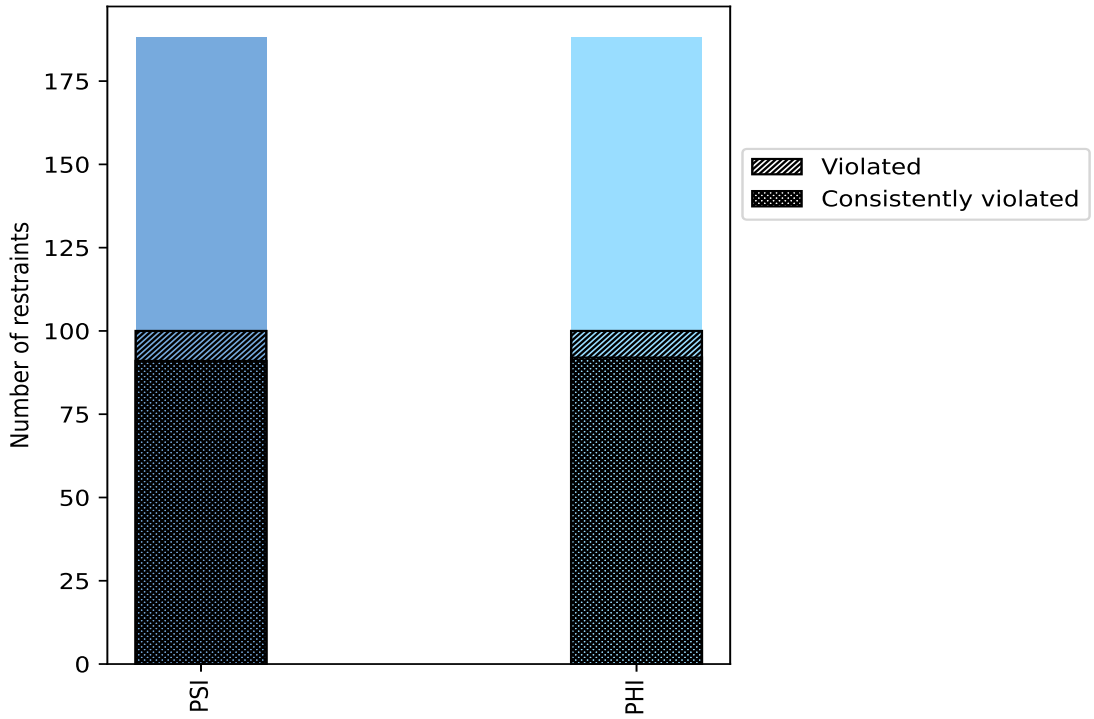
10.1 Summary of dihedral-angle violations

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	188	50.0	100	53.2	26.6	91	48.4	24.2
PHI	188	50.0	100	53.2	26.6	92	48.9	24.5
Total	376	100.0	200	53.2	53.2	183	48.7	48.7

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

10.1.1 Bar chart : Distribution of dihedral-angles and violations



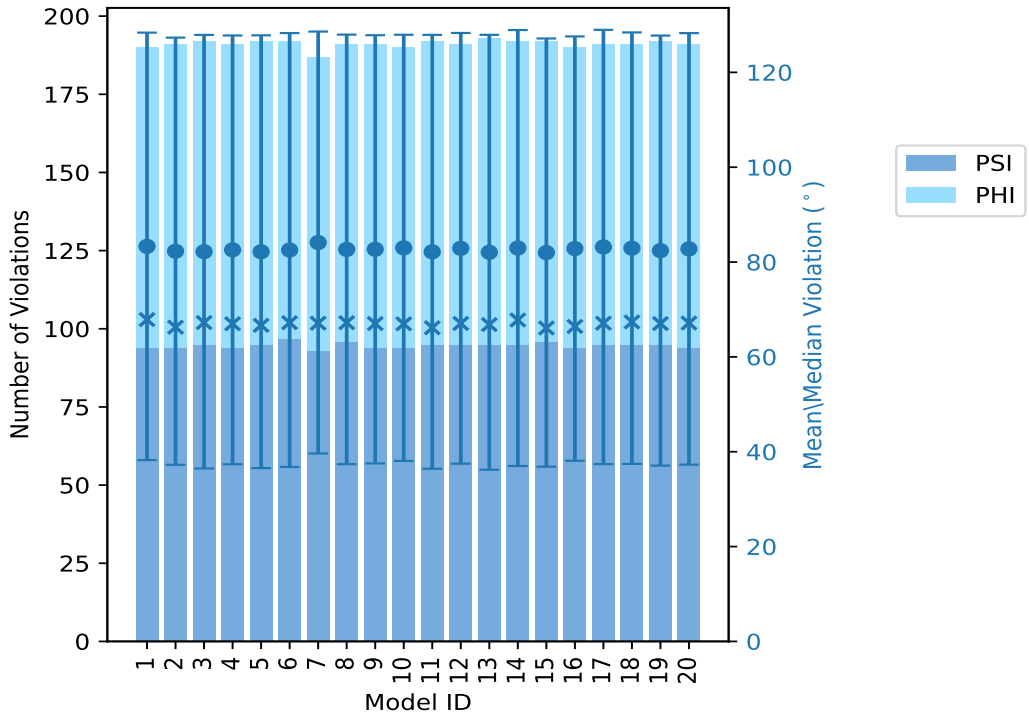
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

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	94	96	190	83.31	178.18	45.07	67.87
2	94	97	191	82.29	173.42	45.04	66.27
3	95	97	192	82.19	178.34	45.71	67.24
4	94	97	191	82.57	173.19	45.21	66.99
5	95	97	192	82.18	177.96	45.63	66.66
6	97	95	192	82.54	179.43	45.75	67.22
7	93	94	187	84.12	177.78	44.49	67.11
8	96	95	191	82.67	178.2	45.29	67.22
9	94	97	191	82.69	174.52	45.15	67.03
10	94	96	190	83.0	176.95	44.93	66.98
11	95	97	192	82.16	176.72	45.75	66.15
12	95	96	191	82.9	178.56	45.41	67.06
13	95	98	193	82.06	177.43	45.86	66.82
14	95	97	192	82.96	179.56	45.97	67.78
15	96	96	192	82.0	178.6	45.15	66.09
16	94	96	190	82.84	177.55	44.74	66.42
17	95	96	191	83.18	178.88	45.79	67.11
18	95	96	191	82.93	178.74	45.49	67.37
19	95	97	192	82.42	176.81	45.34	67.03
20	94	97	191	82.78	178.6	45.5	67.14

10.2.1 Bar graph : Dihedral 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

10.3 Dihedral-angle violation statistics for the ensemble

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

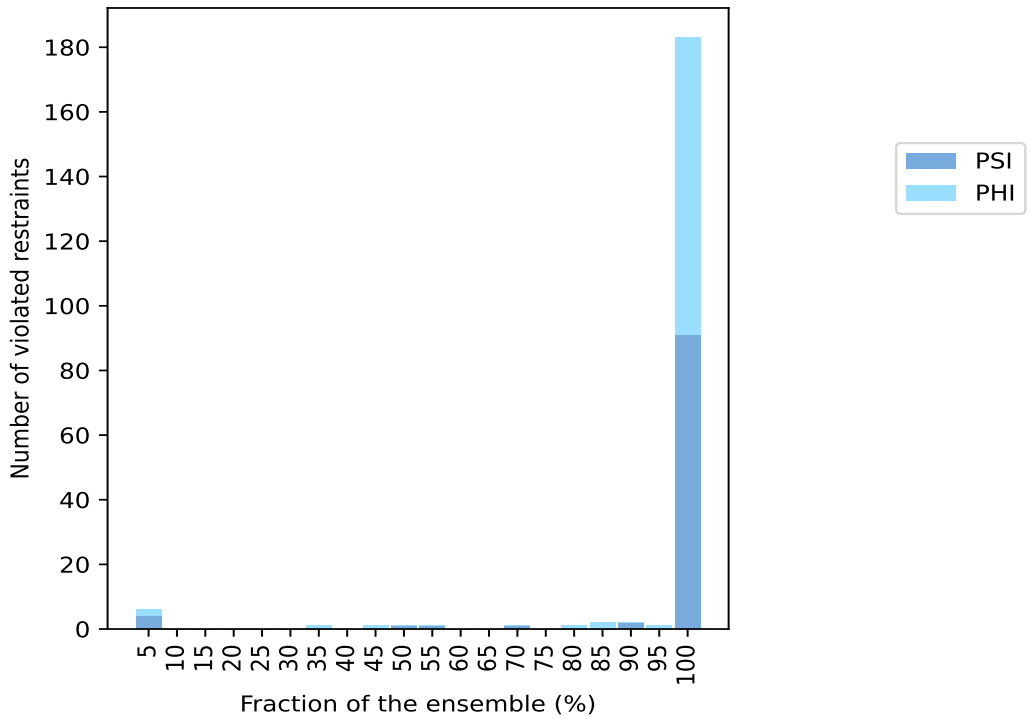
Continued on next page...

Continued from previous page...

Number of violated restraints			Fraction of the ensemble	
PSI	PHI	Total	Count[1]	%
0	0	0	12	60.0
0	0	0	13	65.0
1	0	1	14	70.0
0	0	0	15	75.0
0	1	1	16	80.0
0	2	2	17	85.0
2	0	2	18	90.0
0	1	1	19	95.0
91	92	183	20	100.0

[1] Number of models with violations

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

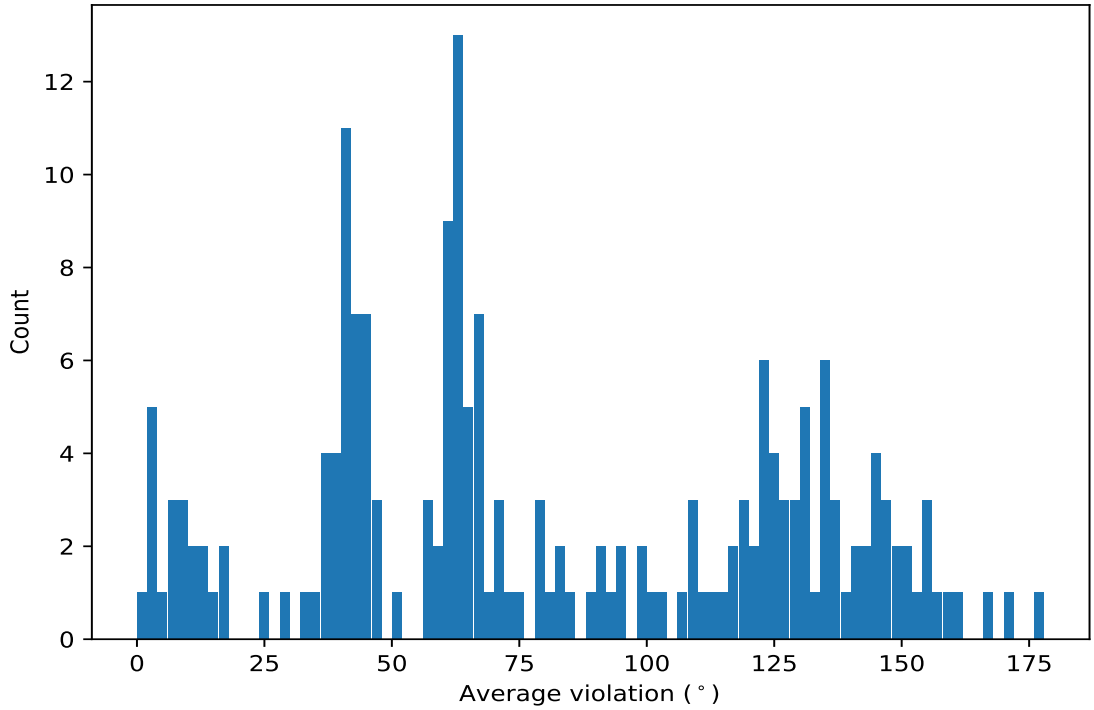


10.4 Most violated dihedral-angle restraints in the ensemble

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

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.

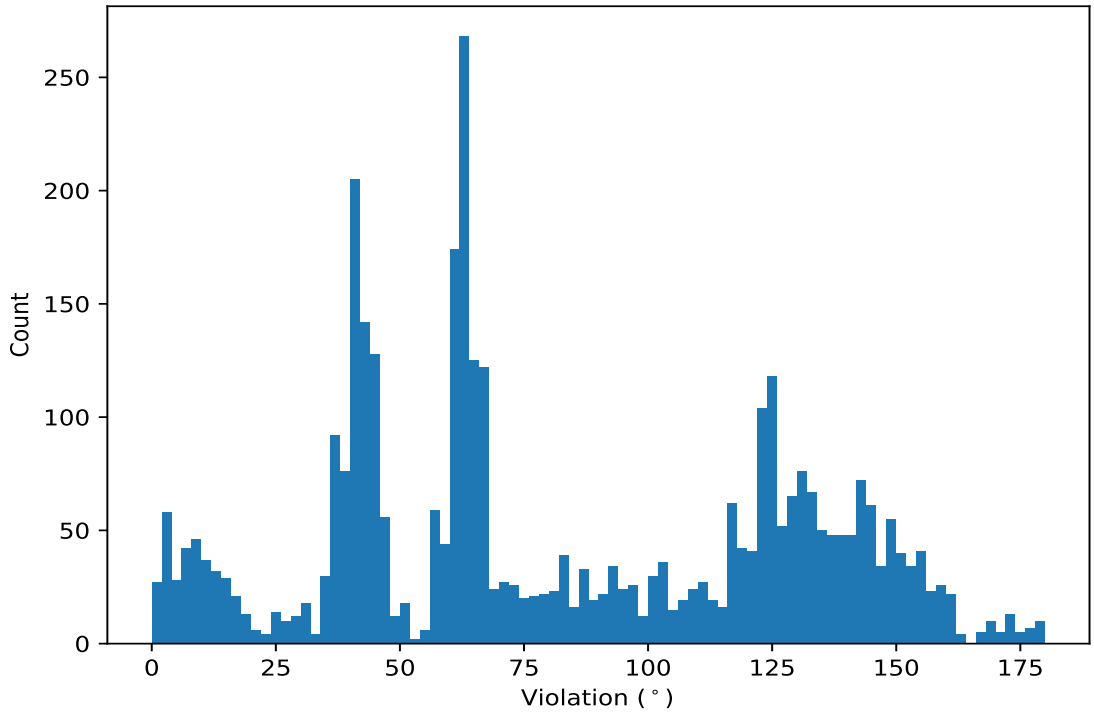
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models[1]	Mean	SD[2]	Median
(2,88)	1:74:A:GLY:N	1:74:A:GLY:CA	1:74:A:GLY:C	1:75:A:ASN:N	20	177.47	1.77	178.07
(2,152)	1:109:A:GLU:N	1:109:A:GLU:CA	1:109:A:GLU:C	1:110:A:GLY:N	20	170.89	5.91	173.3
(2,40)	1:48:A:ASN:N	1:48:A:ASN:CA	1:48:A:ASN:C	1:49:A:TRP:N	20	166.84	5.24	168.46
(2,92)	1:76:A:PHE:N	1:76:A:PHE:CA	1:76:A:PHE:C	1:77:A:ALA:N	20	160.42	1.36	160.61
(2,87)	1:73:A:GLU:C	1:74:A:GLY:N	1:74:A:GLY:CA	1:74:A:GLY:C	20	159.75	2.77	160.04
(2,86)	1:73:A:GLU:N	1:73:A:GLU:CA	1:73:A:GLU:C	1:74:A:GLY:N	20	156.2	3.99	157.41
(2,26)	1:41:A:CYS:N	1:41:A:CYS:CA	1:41:A:CYS:C	1:42:A:GLU:N	20	155.2	3.79	155.26
(2,90)	1:75:A:ASN:N	1:75:A:ASN:CA	1:75:A:ASN:C	1:76:A:PHE:N	20	155.12	0.71	155.17
(2,38)	1:47:A:ALA:N	1:47:A:ALA:CA	1:47:A:ALA:C	1:48:A:ASN:N	20	154.4	3.17	154.69
(2,12)	1:34:A:PRO:N	1:34:A:PRO:CA	1:34:A:PRO:C	1:35:A:ILE:N	20	153.22	1.54	153.34

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

10.5 All violated dihedral-angle restraints

10.5.1 Histogram : Distribution of violations

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

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 (°)
(2,88)	1:74:A:GLY:N	1:74:A:GLY:CA	1:74:A:GLY:C	1:75:A:ASN:N	14	179.56
(2,88)	1:74:A:GLY:N	1:74:A:GLY:CA	1:74:A:GLY:C	1:75:A:ASN:N	6	179.43
(2,88)	1:74:A:GLY:N	1:74:A:GLY:CA	1:74:A:GLY:C	1:75:A:ASN:N	17	178.88
(2,88)	1:74:A:GLY:N	1:74:A:GLY:CA	1:74:A:GLY:C	1:75:A:ASN:N	18	178.74
(2,88)	1:74:A:GLY:N	1:74:A:GLY:CA	1:74:A:GLY:C	1:75:A:ASN:N	15	178.6
(2,88)	1:74:A:GLY:N	1:74:A:GLY:CA	1:74:A:GLY:C	1:75:A:ASN:N	20	178.6
(2,88)	1:74:A:GLY:N	1:74:A:GLY:CA	1:74:A:GLY:C	1:75:A:ASN:N	12	178.56
(2,88)	1:74:A:GLY:N	1:74:A:GLY:CA	1:74:A:GLY:C	1:75:A:ASN:N	3	178.34
(2,88)	1:74:A:GLY:N	1:74:A:GLY:CA	1:74:A:GLY:C	1:75:A:ASN:N	8	178.2
(2,88)	1:74:A:GLY:N	1:74:A:GLY:CA	1:74:A:GLY:C	1:75:A:ASN:N	1	178.18