wwPDB NMR Structure Validation Summary Report

Mar 26, 2026 – 06:16 AM UTC

PDB ID	:	2LXL / pdb_00002lxl
BMRB ID	:	18681
Title	:	Lip5(mit)2
Authors	:	Skalicky, J.J.; Sundquist, W.I.
Deposited on	:	2012-08-29

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

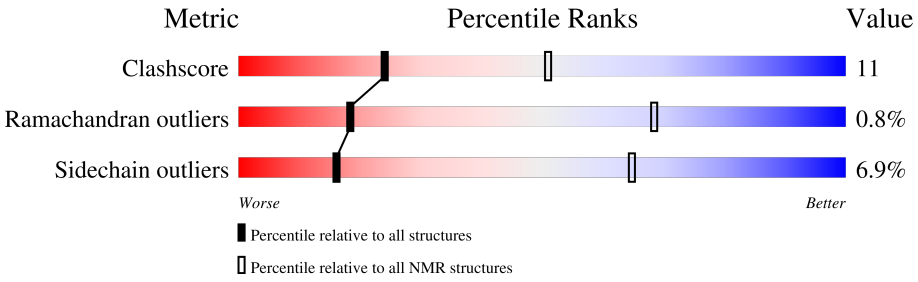
MolProbity	:	4-5-2 with Phenix2.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

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 73%.

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	229148	14424
Ramachandran outliers	224038	12848
Sidechain outliers	223484	12823

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	183	71% 16% • 11%

2 Ensemble composition and analysis

This entry contains 10 models. Model 1 is the overall representative, medoid model (most similar to other models).

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

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:4-A:163 (160)	0.73	1

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 2 clusters and 2 single-model clusters were found.

Cluster number	Models
1	1, 4, 7, 8, 9
2	3, 5, 10
Single-model clusters	2; 6

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2642 atoms, of which 1323 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Vacuolar protein sorting-associated protein VTA1 homolog.

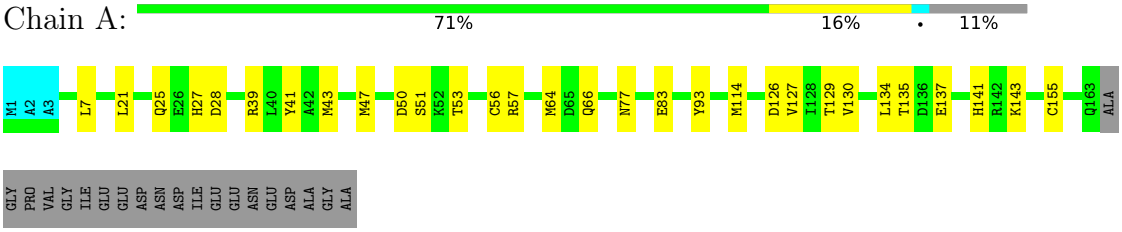
Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	163	2642	839	1323	232	238	10	0

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

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

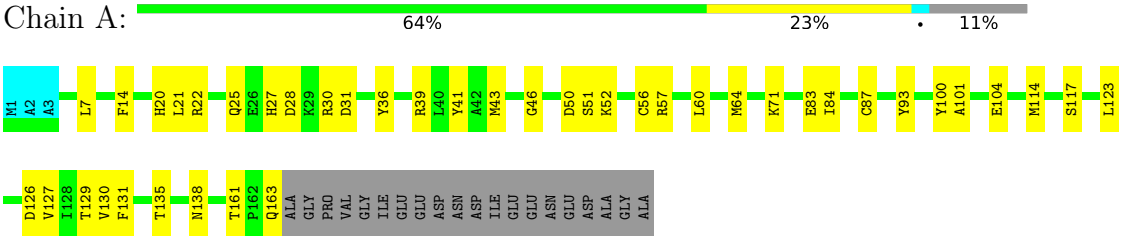
• Molecule 1: Vacuolar protein sorting-associated protein VTA1 homolog



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

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

• Molecule 1: Vacuolar protein sorting-associated protein VTA1 homolog



5 Refinement protocol and experimental data overview

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

Of the 400 calculated structures, 10 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
XPLOR-NIH	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	1867
Number of shifts mapped to atoms	1682
Number of unparsed shifts	0
Number of shifts with mapping errors	185
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	73%

6 Model quality

6.1 Standard geometry

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	0.66±0.01	0±0/1330 (0.0± 0.0%)	1.02±0.01	0±0/1793 (0.0± 0.0%)
All	All	0.66	0/13300 (0.0%)	1.02	1/17930 (0.0%)

There are no bond-length outliers.

All unique angle outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	141	HIS	CA-CB-CG	-5.15	108.65	113.80	2	1

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	1301	1305	1305	28±6
All	All	13010	13050	13050	281

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 11.

5 of 196 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:129:THR:HG22	1:A:134:LEU:HD21	0.86	1.46	6	2

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:7:LEU:N	1:A:7:LEU:HD22	0.79	1.93	1	4
1:A:14:PHE:CD2	1:A:17:ILE:HD11	0.66	2.26	7	1
1:A:127:VAL:O	1:A:130:VAL:HG22	0.66	1.91	2	10
1:A:14:PHE:CZ	1:A:56:CYS:SG	0.64	2.91	7	1

6.3 Torsion angles (i)

6.3.1 Protein backbone (i)

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	159/183 (87%)	152±2 (96±1%)	6±2 (3±1%)	1±1 (1±0%)	18	68
All	All	1590/1830 (87%)	1523 (96%)	55 (3%)	12 (1%)	18	68

All 4 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	51	SER	9
1	A	56	CYS	1
1	A	49	ILE	1
1	A	54	PRO	1

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	140/155 (90%)	130±3 (93±2%)	10±3 (7±2%)	16	65
All	All	1400/1550 (90%)	1303 (93%)	97 (7%)	16	65

5 of 44 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	21	LEU	7
1	A	53	THR	6
1	A	134	LEU	6
1	A	93	TYR	5
1	A	64	MET	5

6.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

6.5 Carbohydrates (i)

There are no 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 73% for the well-defined parts and 73% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

7.1.1 Bookkeeping

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	1867
Number of shifts mapped to atoms	1682
Number of unparsed shifts	0
Number of shifts with mapping errors	185
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 185) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	164	ALA	H	8.35	0.002	1
1	A	164	ALA	HA	4.428	0.003	1
1	A	164	ALA	HB1	1.451	0.001	1
1	A	164	ALA	HB2	1.451	0.001	1
1	A	164	ALA	HB3	1.451	0.001	1
1	A	164	ALA	CA	52.423	0.103	1
1	A	164	ALA	CB	20.313	0.062	1
1	A	164	ALA	N	126.533	0.017	1
1	A	165	GLY	H	8.999	0.004	1
1	A	165	GLY	HA2	3.86	0.003	.
1	A	165	GLY	HA3	4.189	0.003	.
1	A	165	GLY	CA	44.637	0.051	1
1	A	165	GLY	N	110.327	0.012	1
1	A	166	PRO	HA	4.538	0.001	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	166	PRO	HB2	1.89	0.000	.
1	A	166	PRO	HB3	1.947	0.000	.
1	A	166	PRO	HD2	3.457	0.002	.
1	A	166	PRO	HD3	3.621	0.003	.
1	A	166	PRO	CA	63.233	0.008	1
1	A	166	PRO	CB	34.367	0.007	1
1	A	166	PRO	CD	49.285	0.031	1
1	A	167	VAL	H	8.21	0.002	1
1	A	167	VAL	HA	3.836	0.002	1
1	A	167	VAL	HB	1.769	0.002	1
1	A	167	VAL	HG11	0.713	0.004	.
1	A	167	VAL	HG12	0.713	0.004	.
1	A	167	VAL	HG13	0.713	0.004	.
1	A	167	VAL	HG21	0.721	0.005	.
1	A	167	VAL	HG22	0.721	0.005	.
1	A	167	VAL	HG23	0.721	0.005	.
1	A	167	VAL	CA	61.891	0.010	1
1	A	167	VAL	CB	33.033	0.092	1
1	A	167	VAL	CG1	20.784	0.034	.
1	A	167	VAL	CG2	21.186	0.000	.
1	A	167	VAL	N	121.37	0.011	1
1	A	168	GLY	H	8.104	0.005	1
1	A	168	GLY	HA2	3.735	0.006	.
1	A	168	GLY	HA3	3.735	0.006	.
1	A	168	GLY	CA	45.118	0.061	1
1	A	168	GLY	N	111.732	0.012	1
1	A	169	ILE	H	7.433	0.005	1
1	A	169	ILE	HA	3.767	0.003	1
1	A	169	ILE	HB	1.582	0.004	1
1	A	169	ILE	HD11	0.655	0.001	1
1	A	169	ILE	HD12	0.655	0.001	1
1	A	169	ILE	HD13	0.655	0.001	1
1	A	169	ILE	HG12	0.774	0.004	.
1	A	169	ILE	HG13	1.077	0.002	.
1	A	169	ILE	HG21	0.634	0.010	1
1	A	169	ILE	HG22	0.634	0.010	1
1	A	169	ILE	HG23	0.634	0.010	1
1	A	169	ILE	CA	60.973	0.000	1
1	A	169	ILE	CB	38.512	0.005	1
1	A	169	ILE	CD1	13.252	0.016	1
1	A	169	ILE	CG1	26.712	0.058	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	169	ILE	CG2	17.346	0.105	1
1	A	169	ILE	N	118.646	0.006	1
1	A	170	GLU	H	8.165	0.003	1
1	A	170	GLU	HA	3.797	0.004	1
1	A	170	GLU	HB2	1.692	0.003	.
1	A	170	GLU	HB3	1.79	0.005	.
1	A	170	GLU	HG2	2.0	0.001	.
1	A	170	GLU	HG3	2.0	0.001	.
1	A	170	GLU	CA	56.631	0.032	1
1	A	170	GLU	CB	29.84	0.111	1
1	A	170	GLU	CG	35.971	0.008	1
1	A	170	GLU	N	123.982	0.017	1
1	A	171	GLU	H	8.019	0.004	1
1	A	171	GLU	HA	4.038	0.006	1
1	A	171	GLU	HB2	1.776	0.005	.
1	A	171	GLU	HB3	1.924	0.003	.
1	A	171	GLU	HG2	2.088	0.005	.
1	A	171	GLU	HG3	2.088	0.005	.
1	A	171	GLU	CA	56.326	0.123	1
1	A	171	GLU	CB	30.343	0.134	1
1	A	171	GLU	CG	36.256	0.000	1
1	A	171	GLU	N	120.341	0.038	1
1	A	172	ASP	H	8.137	0.003	1
1	A	172	ASP	HA	4.554	0.001	1
1	A	172	ASP	HB2	2.517	0.002	.
1	A	172	ASP	HB3	2.653	0.005	.
1	A	172	ASP	CA	54.123	0.032	1
1	A	172	ASP	CB	40.951	0.004	1
1	A	172	ASP	N	120.462	0.008	1
1	A	173	ASN	H	8.14	0.002	1
1	A	173	ASN	HA	4.684	0.006	1
1	A	173	ASN	HB2	2.634	0.007	.
1	A	173	ASN	HB3	2.717	0.005	.
1	A	173	ASN	HD21	6.812	0.004	.
1	A	173	ASN	HD22	7.477	0.004	.
1	A	173	ASN	CA	53.117	0.094	1
1	A	173	ASN	CB	39.62	0.132	1
1	A	173	ASN	N	119.022	0.024	1
1	A	173	ASN	ND2	113.312	0.011	1
1	A	174	ASP	H	8.392	0.003	1
1	A	174	ASP	HA	4.625	0.002	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	174	ASP	HB2	2.554	0.004	.
1	A	174	ASP	HB3	2.654	0.002	.
1	A	174	ASP	CA	54.055	0.034	1
1	A	174	ASP	CB	41.136	0.014	1
1	A	174	ASP	N	120.952	0.007	1
1	A	175	ILE	H	7.981	0.003	1
1	A	175	ILE	HA	4.011	0.001	1
1	A	175	ILE	HB	1.635	0.001	1
1	A	175	ILE	HD11	0.64	0.010	1
1	A	175	ILE	HD12	0.64	0.010	1
1	A	175	ILE	HD13	0.64	0.010	1
1	A	175	ILE	HG12	0.94	0.003	.
1	A	175	ILE	HG13	1.283	0.002	.
1	A	175	ILE	HG21	0.683	0.003	1
1	A	175	ILE	HG22	0.683	0.003	1
1	A	175	ILE	HG23	0.683	0.003	1
1	A	175	ILE	CA	60.852	0.012	1
1	A	175	ILE	CB	38.79	0.026	1
1	A	175	ILE	CD1	13.159	0.023	1
1	A	175	ILE	CG1	27.099	0.048	1
1	A	175	ILE	CG2	17.395	0.077	1
1	A	175	ILE	N	120.253	0.035	1
1	A	176	GLU	H	8.31	0.002	1
1	A	176	GLU	HA	4.226	0.006	1
1	A	176	GLU	HB2	1.867	0.000	.
1	A	176	GLU	HB3	1.971	0.002	.
1	A	176	GLU	HG2	2.195	0.003	.
1	A	176	GLU	HG3	2.195	0.003	.
1	A	176	GLU	CA	56.462	0.096	1
1	A	176	GLU	CB	30.41	0.000	1
1	A	176	GLU	CG	36.225	0.000	1
1	A	176	GLU	N	124.792	0.021	1
1	A	177	GLU	H	8.411	0.006	1
1	A	177	GLU	HA	4.219	0.003	1
1	A	177	GLU	HB2	1.869	0.002	.
1	A	177	GLU	HB3	1.973	0.008	.
1	A	177	GLU	HG2	2.193	0.004	.
1	A	177	GLU	HG3	2.193	0.004	.
1	A	177	GLU	CA	56.586	0.002	1
1	A	177	GLU	CB	30.21	0.004	1
1	A	177	GLU	CG	35.982	0.000	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	177	GLU	N	122.343	0.028	1
1	A	178	ASN	H	8.459	0.003	1
1	A	178	ASN	HA	4.665	0.002	1
1	A	178	ASN	HB2	2.67	0.006	.
1	A	178	ASN	HB3	2.731	0.006	.
1	A	178	ASN	HD21	6.866	0.002	.
1	A	178	ASN	HD22	7.579	0.003	.
1	A	178	ASN	CA	53.126	0.000	1
1	A	178	ASN	CB	39.275	0.101	1
1	A	178	ASN	N	120.046	0.047	1
1	A	178	ASN	ND2	112.783	0.010	1
1	A	179	GLU	H	8.502	0.001	1
1	A	179	GLU	HA	4.195	0.003	1
1	A	179	GLU	HB2	1.88	0.006	.
1	A	179	GLU	HB3	1.989	0.005	.
1	A	179	GLU	HG2	2.197	0.001	.
1	A	179	GLU	HG3	2.197	0.001	.
1	A	179	GLU	CA	56.977	0.000	1
1	A	179	GLU	CB	30.271	0.004	1
1	A	179	GLU	CG	36.138	0.000	1
1	A	179	GLU	N	121.956	0.005	1
1	A	180	ASP	H	8.304	0.002	1
1	A	180	ASP	HA	4.475	0.002	1
1	A	180	ASP	HB2	2.52	0.002	.
1	A	180	ASP	HB3	2.639	0.002	.
1	A	180	ASP	CB	41.0	0.135	1
1	A	180	ASP	N	120.953	0.026	1
1	A	181	ALA	H	8.006	0.002	1
1	A	181	ALA	HA	3.97	0.003	1
1	A	181	ALA	HB1	1.179	0.003	1
1	A	181	ALA	HB2	1.179	0.003	1
1	A	181	ALA	HB3	1.179	0.003	1
1	A	181	ALA	CA	52.742	0.119	1
1	A	181	ALA	CB	18.844	0.094	1
1	A	181	ALA	N	124.298	0.017	1
1	A	182	GLY	H	8.112	0.004	1
1	A	182	GLY	HA2	3.562	0.001	.
1	A	182	GLY	HA3	3.724	0.003	.
1	A	182	GLY	CA	44.977	0.105	1
1	A	182	GLY	N	107.168	0.009	1
1	A	183	ALA	H	7.465	0.001	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	183	ALA	HA	3.998	0.001	1
1	A	183	ALA	HB1	1.241	0.002	1
1	A	183	ALA	HB2	1.241	0.002	1
1	A	183	ALA	HB3	1.241	0.002	1
1	A	183	ALA	CA	53.582	0.005	1
1	A	183	ALA	CB	19.911	0.008	1
1	A	183	ALA	N	128.407	0.005	1

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	171	-0.56 ± 0.09	Should be checked
$^{13}C_{\beta}$	161	0.33 ± 0.09	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	168	0.19 ± 0.20	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 73%, i.e. 1654 atoms were assigned a chemical shift out of a possible 2260. 0 out of 23 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	615/792 (78%)	319/319 (100%)	149/320 (47%)	147/153 (96%)
Sidechain	927/1275 (73%)	638/826 (77%)	275/395 (70%)	14/54 (26%)
Aromatic	112/193 (58%)	71/96 (74%)	40/89 (45%)	1/8 (12%)
Overall	1654/2260 (73%)	1028/1241 (83%)	464/804 (58%)	162/215 (75%)

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	101	ALA	HB1	0.13	0.14 – 2.58	-5.0
1	A	101	ALA	HB2	0.13	0.14 – 2.58	-5.0

Continued on next page...

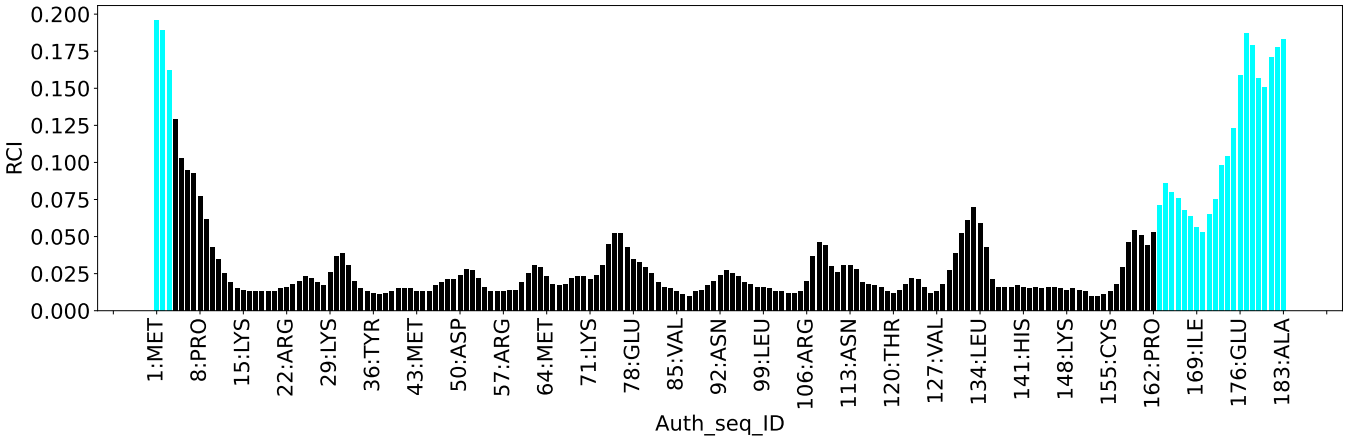
Continued from previous page...

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	101	ALA	HB3	0.13	0.14 – 2.58	-5.0

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	3360
Intra-residue (|i-j|=0)	775
Sequential (|i-j|=1)	960
Medium range (|i-j|>1 and |i-j|<5)	787
Long range (|i-j|≥5)	838
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	354
Number of unmapped restraints	290
Number of restraints per residue	20.3
Number of long range restraints per residue[1]	4.6

[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)	22.1	0.2
0.2-0.5 (Medium)	10.8	0.49
>0.5 (Large)	0.3	0.58

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)	7.4	5.44
10.0-20.0 (Medium)	None	None
>20.0 (Large)	None	None

9 Distance violation analysis (i)

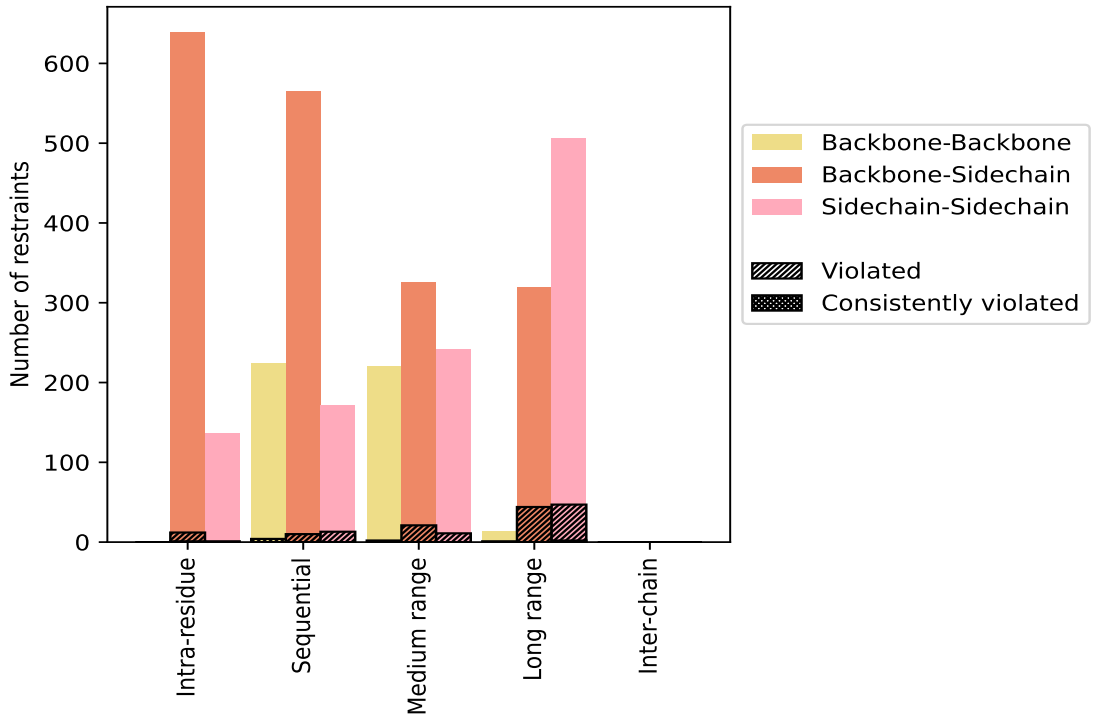
9.1 Summary of distance violations (i)

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

Restraints type	Count	%[1]	Violated[3]			Consistently Violated[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
Intra-residue (|i-j|=0)	775	23.1	13	1.7	0.4	1	0.1	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	639	19.0	12	1.9	0.4	1	0.2	0.0
Sidechain-Sidechain	136	4.0	1	0.7	0.0	0	0.0	0.0
Sequential (|i-j|=1)	960	28.6	27	2.8	0.8	0	0.0	0.0
Backbone-Backbone	224	6.7	4	1.8	0.1	0	0.0	0.0
Backbone-Sidechain	565	16.8	10	1.8	0.3	0	0.0	0.0
Sidechain-Sidechain	171	5.1	13	7.6	0.4	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	787	23.4	34	4.3	1.0	0	0.0	0.0
Backbone-Backbone	220	6.5	2	0.9	0.1	0	0.0	0.0
Backbone-Sidechain	325	9.7	21	6.5	0.6	0	0.0	0.0
Sidechain-Sidechain	242	7.2	11	4.5	0.3	0	0.0	0.0
Long range (|i-j|≥5)	838	24.9	92	11.0	2.7	2	0.2	0.1
Backbone-Backbone	13	0.4	1	7.7	0.0	0	0.0	0.0
Backbone-Sidechain	319	9.5	44	13.8	1.3	0	0.0	0.0
Sidechain-Sidechain	506	15.1	47	9.3	1.4	2	0.4	0.1
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	0	0.0	0	0.0	0.0	0	0.0	0.0
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	3360	100.0	166	4.9	4.9	3	0.1	0.1
Backbone-Backbone	457	13.6	7	1.5	0.2	0	0.0	0.0
Backbone-Sidechain	1848	55.0	87	4.7	2.6	1	0.1	0.0
Sidechain-Sidechain	1055	31.4	72	6.8	2.1	2	0.2	0.1

[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

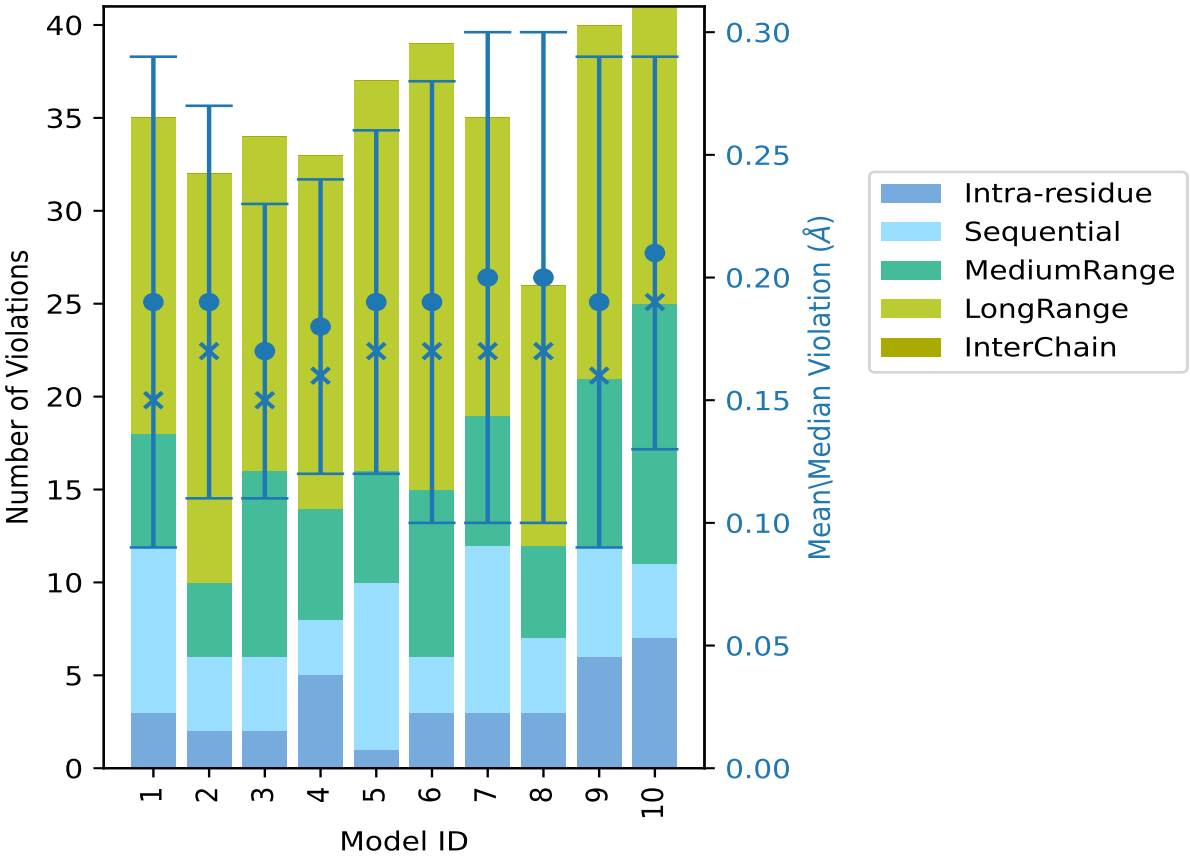
9.2 Distance violation statistics for each model

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	3	9	6	17	0	35	0.19	0.44	0.1	0.15
2	2	4	4	22	0	32	0.19	0.49	0.08	0.17
3	2	4	10	18	0	34	0.17	0.35	0.06	0.15
4	5	3	6	19	0	33	0.18	0.34	0.06	0.16
5	1	9	6	21	0	37	0.19	0.4	0.07	0.17
6	3	3	9	24	0	39	0.19	0.56	0.09	0.17
7	3	9	7	16	0	35	0.2	0.53	0.1	0.17
8	3	4	5	14	0	26	0.2	0.47	0.1	0.17
9	6	6	9	19	0	40	0.19	0.58	0.1	0.16
10	7	4	14	16	0	41	0.21	0.49	0.08	0.19

[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

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, 3194(IR:762, SQ:933, MR:753, LR:746, 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]	%
6	14	19	50	0	89	1	10.0
2	6	5	24	0	37	2	20.0
1	4	3	7	0	15	3	30.0

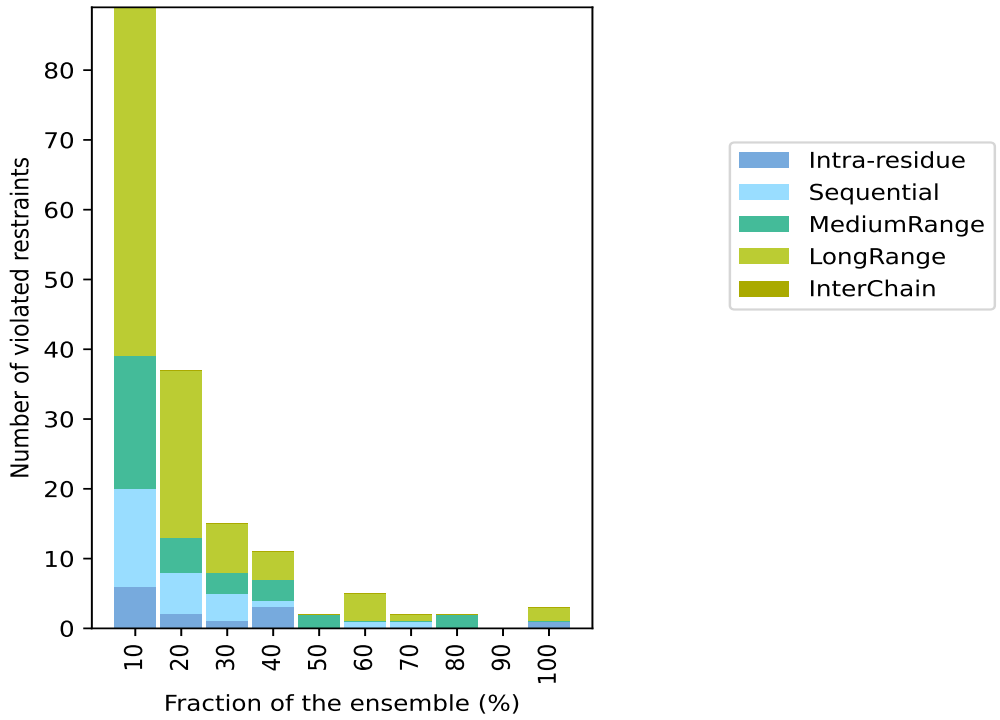
Continued on next page...

Continued from previous page...

Number of violated restraints						Fraction of the ensemble	
IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total	Count[6]	%
3	1	3	4	0	11	4	40.0
0	0	2	0	0	2	5	50.0
0	1	0	4	0	5	6	60.0
0	1	0	1	0	2	7	70.0
0	0	2	0	0	2	8	80.0
0	0	0	0	0	0	9	90.0
1	0	0	2	0	3	10	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

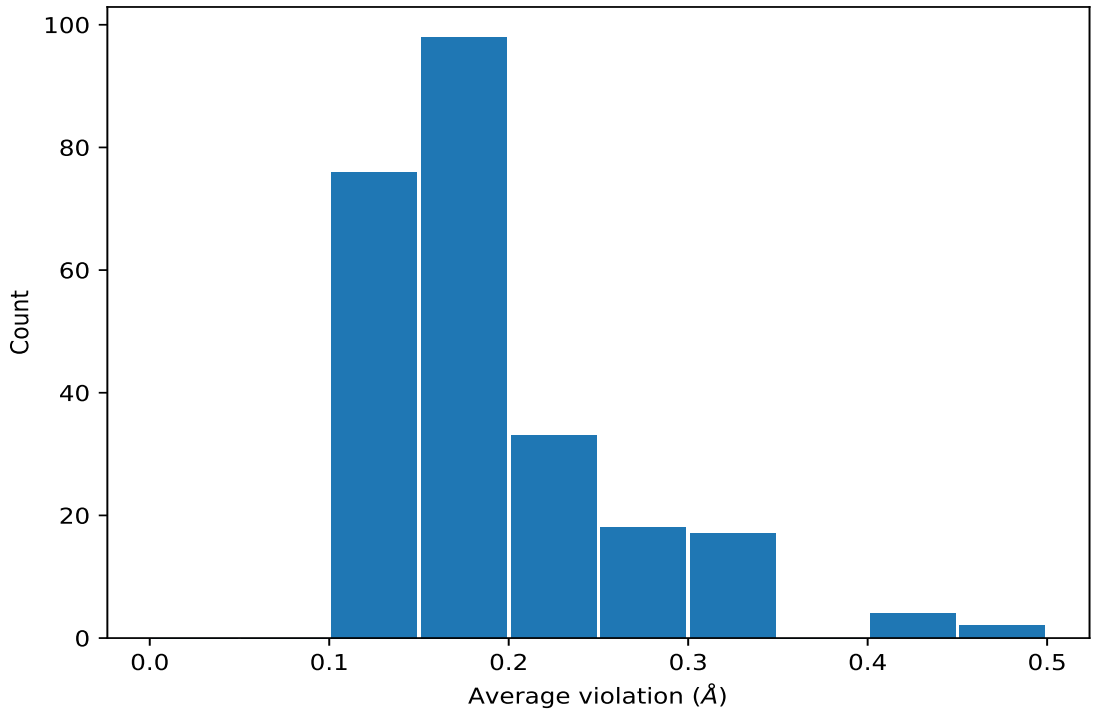


9.4 Most violated distance restraints in the ensemble

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)
(1,3083)	1:51:A:SER:HB2	1:60:A:LEU:HD11	10	0.33	0.12	0.31
(1,3083)	1:51:A:SER:HB2	1:60:A:LEU:HD12	10	0.33	0.12	0.31
(1,3083)	1:51:A:SER:HB2	1:60:A:LEU:HD13	10	0.33	0.12	0.31
(1,3083)	1:51:A:SER:HB2	1:60:A:LEU:HD21	10	0.33	0.12	0.31
(1,3083)	1:51:A:SER:HB2	1:60:A:LEU:HD22	10	0.33	0.12	0.31
(1,3083)	1:51:A:SER:HB2	1:60:A:LEU:HD23	10	0.33	0.12	0.31
(1,3083)	1:51:A:SER:HB3	1:60:A:LEU:HD11	10	0.33	0.12	0.31
(1,3083)	1:51:A:SER:HB3	1:60:A:LEU:HD12	10	0.33	0.12	0.31
(1,3083)	1:51:A:SER:HB3	1:60:A:LEU:HD13	10	0.33	0.12	0.31
(1,3083)	1:51:A:SER:HB3	1:60:A:LEU:HD21	10	0.33	0.12	0.31
(1,3083)	1:51:A:SER:HB3	1:60:A:LEU:HD22	10	0.33	0.12	0.31
(1,3083)	1:51:A:SER:HB3	1:60:A:LEU:HD23	10	0.33	0.12	0.31
(1,971)	1:18:A:GLN:H	1:18:A:GLN:HG2	10	0.29	0.03	0.3
(1,245)	1:90:A:LEU:HD11	1:124:A:LEU:HB3	10	0.2	0.06	0.19
(1,245)	1:90:A:LEU:HD12	1:124:A:LEU:HB3	10	0.2	0.06	0.19

Continued on next page...

Continued from previous page...

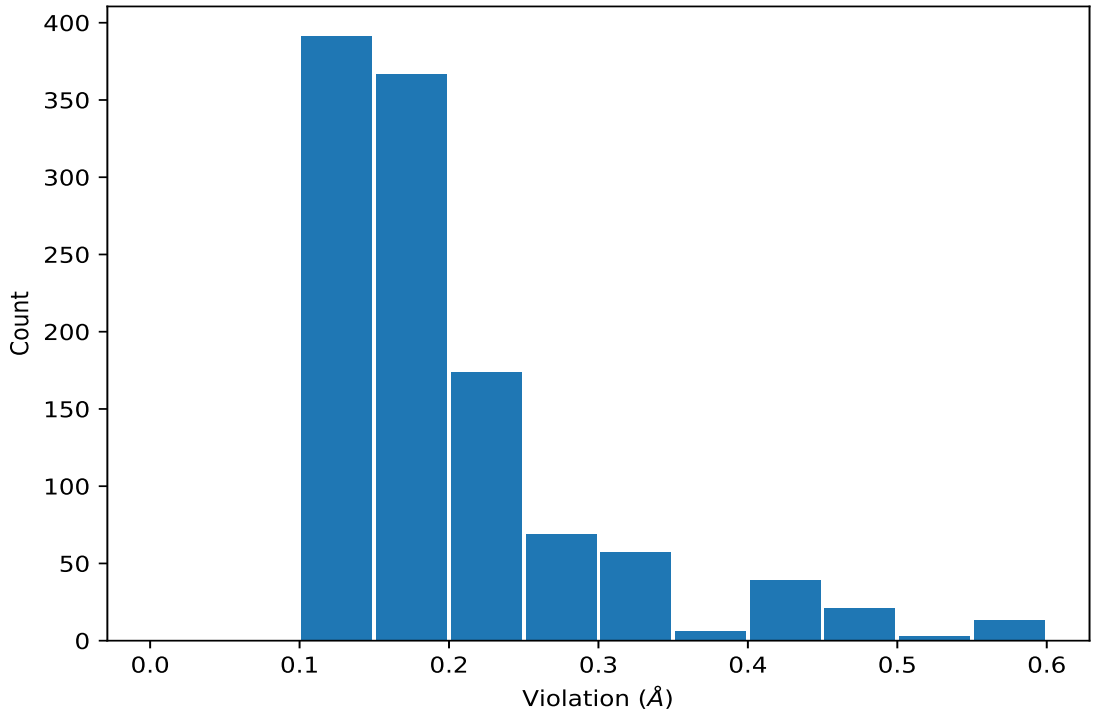
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,245)	1:90:A:LEU:HD13	1:124:A:LEU:HB3	10	0.2	0.06	0.19
(1,1261)	1:49:A:ILE:HG21	1:51:A:SER:H	8	0.26	0.1	0.27
(1,1261)	1:49:A:ILE:HG22	1:51:A:SER:H	8	0.26	0.1	0.27
(1,1261)	1:49:A:ILE:HG23	1:51:A:SER:H	8	0.26	0.1	0.27
(1,2410)	1:84:A:ILE:HD11	1:87:A:CYS:H	8	0.18	0.05	0.16
(1,2410)	1:84:A:ILE:HD12	1:87:A:CYS:H	8	0.18	0.05	0.16
(1,2410)	1:84:A:ILE:HD13	1:87:A:CYS:H	8	0.18	0.05	0.16
(1,1499)	1:70:A:LEU:HB2	1:80:A:ILE:HD11	7	0.19	0.04	0.21
(1,1499)	1:70:A:LEU:HB2	1:80:A:ILE:HD12	7	0.19	0.04	0.21
(1,1499)	1:70:A:LEU:HB2	1:80:A:ILE:HD13	7	0.19	0.04	0.21
(1,2511)	1:76:A:ASP:H	1:77:A:ASN:H	7	0.13	0.02	0.14
(1,2808)	1:102:A:ASP:H	1:144:A:TYR:HE1	6	0.19	0.05	0.2
(1,2808)	1:102:A:ASP:H	1:144:A:TYR:HE2	6	0.19	0.05	0.2
(1,3081)	1:51:A:SER:H	1:52:A:LYS:HG2	6	0.17	0.04	0.18
(1,3081)	1:51:A:SER:H	1:52:A:LYS:HG3	6	0.17	0.04	0.18
(1,2827)	1:36:A:TYR:H	1:70:A:LEU:HG	6	0.16	0.04	0.18

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

9.5 All violated distance restraints (i)

9.5.1 Histogram : Distribution of distance violations (i)

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



9.5.2 Table : All distance violations (i)

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

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,92)	1:147:A:TRP:HE3	1:148:A:LYS:HB2	9	0.58
(1,3083)	1:51:A:SER:HB2	1:60:A:LEU:HD11	6	0.56
(1,3083)	1:51:A:SER:HB2	1:60:A:LEU:HD12	6	0.56
(1,3083)	1:51:A:SER:HB2	1:60:A:LEU:HD13	6	0.56
(1,3083)	1:51:A:SER:HB2	1:60:A:LEU:HD21	6	0.56
(1,3083)	1:51:A:SER:HB2	1:60:A:LEU:HD22	6	0.56
(1,3083)	1:51:A:SER:HB2	1:60:A:LEU:HD23	6	0.56
(1,3083)	1:51:A:SER:HB3	1:60:A:LEU:HD11	6	0.56
(1,3083)	1:51:A:SER:HB3	1:60:A:LEU:HD12	6	0.56
(1,3083)	1:51:A:SER:HB3	1:60:A:LEU:HD13	6	0.56
(1,3083)	1:51:A:SER:HB3	1:60:A:LEU:HD21	6	0.56
(1,3083)	1:51:A:SER:HB3	1:60:A:LEU:HD22	6	0.56
(1,3083)	1:51:A:SER:HB3	1:60:A:LEU:HD23	6	0.56
(1,1218)	1:44:A:GLN:H	1:64:A:MET:HE1	7	0.53
(1,1218)	1:44:A:GLN:H	1:64:A:MET:HE2	7	0.53
(1,1218)	1:44:A:GLN:H	1:64:A:MET:HE3	7	0.53
(1,1859)	1:141:A:HIS:HA	1:141:A:HIS:HD2	2	0.49

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,1242)	1:44:A:GLN:HA	1:47:A:MET:HG3	10	0.49
(1,3083)	1:51:A:SER:HB2	1:60:A:LEU:HD11	8	0.47
(1,3083)	1:51:A:SER:HB2	1:60:A:LEU:HD12	8	0.47
(1,3083)	1:51:A:SER:HB2	1:60:A:LEU:HD13	8	0.47
(1,3083)	1:51:A:SER:HB2	1:60:A:LEU:HD21	8	0.47
(1,3083)	1:51:A:SER:HB2	1:60:A:LEU:HD22	8	0.47
(1,3083)	1:51:A:SER:HB2	1:60:A:LEU:HD23	8	0.47
(1,3083)	1:51:A:SER:HB3	1:60:A:LEU:HD11	8	0.47
(1,3083)	1:51:A:SER:HB3	1:60:A:LEU:HD12	8	0.47
(1,3083)	1:51:A:SER:HB3	1:60:A:LEU:HD13	8	0.47
(1,3083)	1:51:A:SER:HB3	1:60:A:LEU:HD21	8	0.47
(1,3083)	1:51:A:SER:HB3	1:60:A:LEU:HD22	8	0.47
(1,3083)	1:51:A:SER:HB3	1:60:A:LEU:HD23	8	0.47
(1,498)	1:140:A:LYS:HG3	1:140:A:LYS:HE2	7	0.45
(1,498)	1:140:A:LYS:HG3	1:140:A:LYS:HE3	7	0.45
(1,498)	1:140:A:LYS:HG3	1:140:A:LYS:HE2	8	0.45
(1,498)	1:140:A:LYS:HG3	1:140:A:LYS:HE3	8	0.45
(1,406)	1:134:A:LEU:HD11	1:140:A:LYS:H	8	0.45
(1,406)	1:134:A:LEU:HD12	1:140:A:LYS:H	8	0.45
(1,406)	1:134:A:LEU:HD13	1:140:A:LYS:H	8	0.45
(1,3083)	1:51:A:SER:HB2	1:60:A:LEU:HD11	1	0.44

10 Dihedral-angle violation analysis

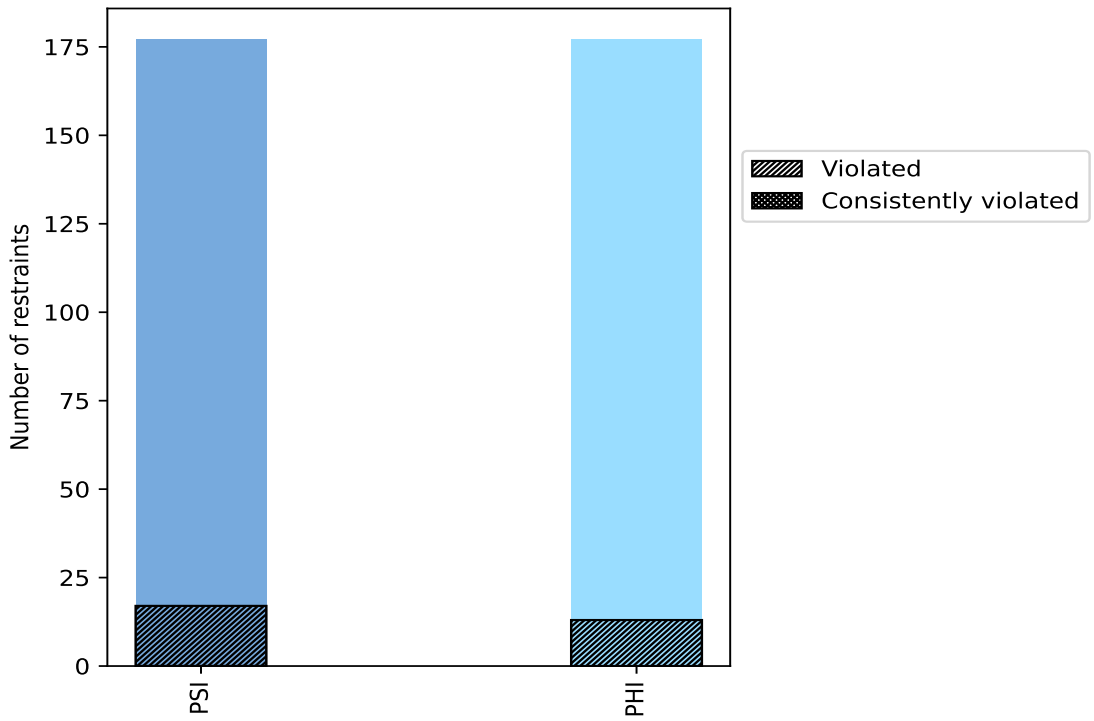
10.1 Summary of dihedral-angle violations

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	177	50.0	17	9.6	4.8	0	0.0	0.0
PHI	177	50.0	13	7.3	3.7	0	0.0	0.0
Total	354	100.0	30	8.5	8.5	0	0.0	0.0

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

10.1.1 Bar chart : Distribution of dihedral-angles and violations



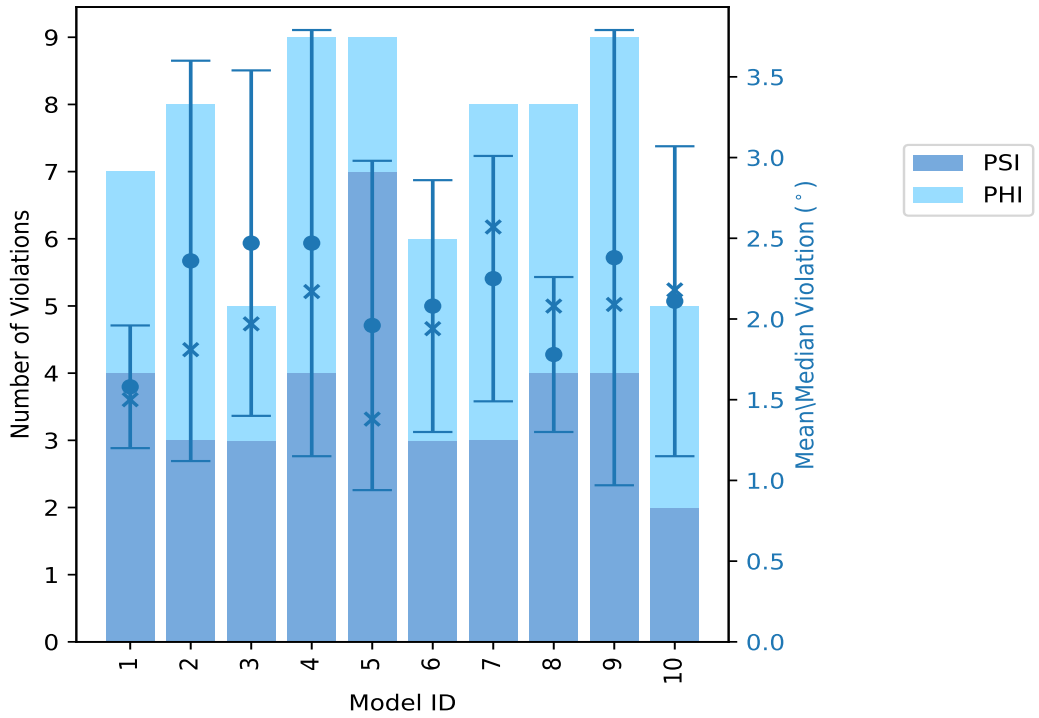
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

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	4	3	7	1.58	2.35	0.38	1.5
2	3	5	8	2.36	4.71	1.24	1.81
3	3	2	5	2.47	4.47	1.07	1.97
4	4	5	9	2.47	4.84	1.32	2.17
5	7	2	9	1.96	3.6	1.02	1.38
6	3	3	6	2.08	3.63	0.78	1.94
7	3	5	8	2.25	3.26	0.76	2.57
8	4	4	8	1.78	2.24	0.48	2.08
9	4	5	9	2.38	5.44	1.41	2.09
10	2	3	5	2.11	3.56	0.96	2.18

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

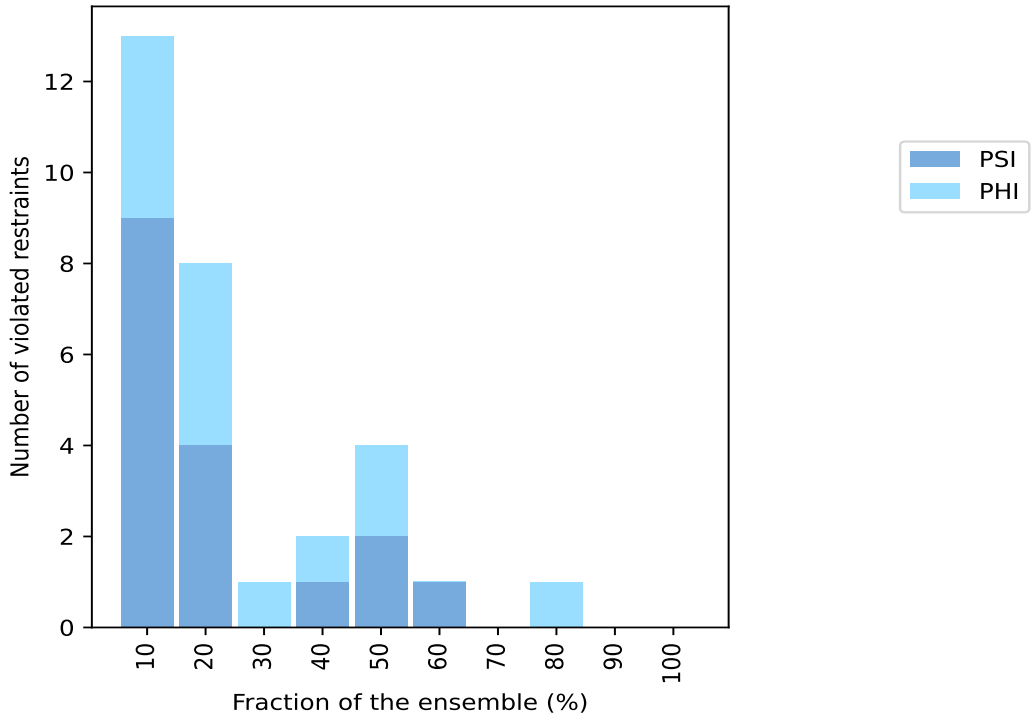
10.3 Dihedral-angle violation statistics for the ensemble

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]	%
9	4	13	1	10.0
4	4	8	2	20.0
0	1	1	3	30.0
1	1	2	4	40.0
2	2	4	5	50.0
1	0	1	6	60.0
0	0	0	7	70.0
0	1	1	8	80.0
0	0	0	9	90.0
0	0	0	10	100.0

[1] Number of models with violations

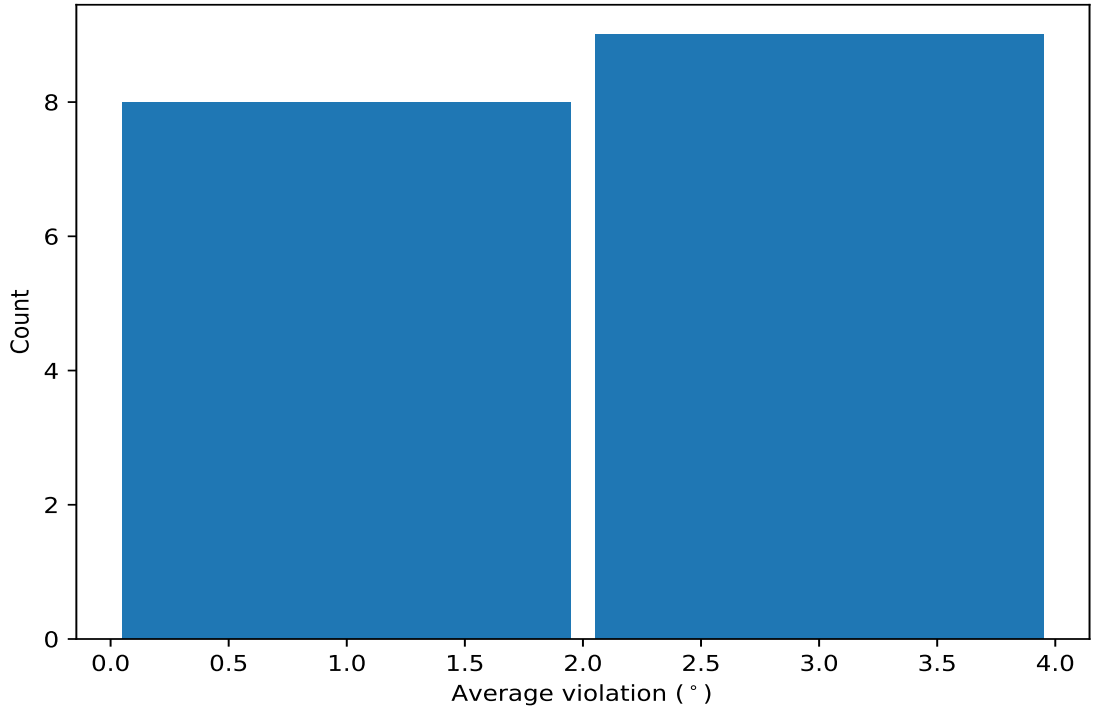
10.3.1 Bar graph : Dihedral-angle Violation statistics for the ensemble



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.

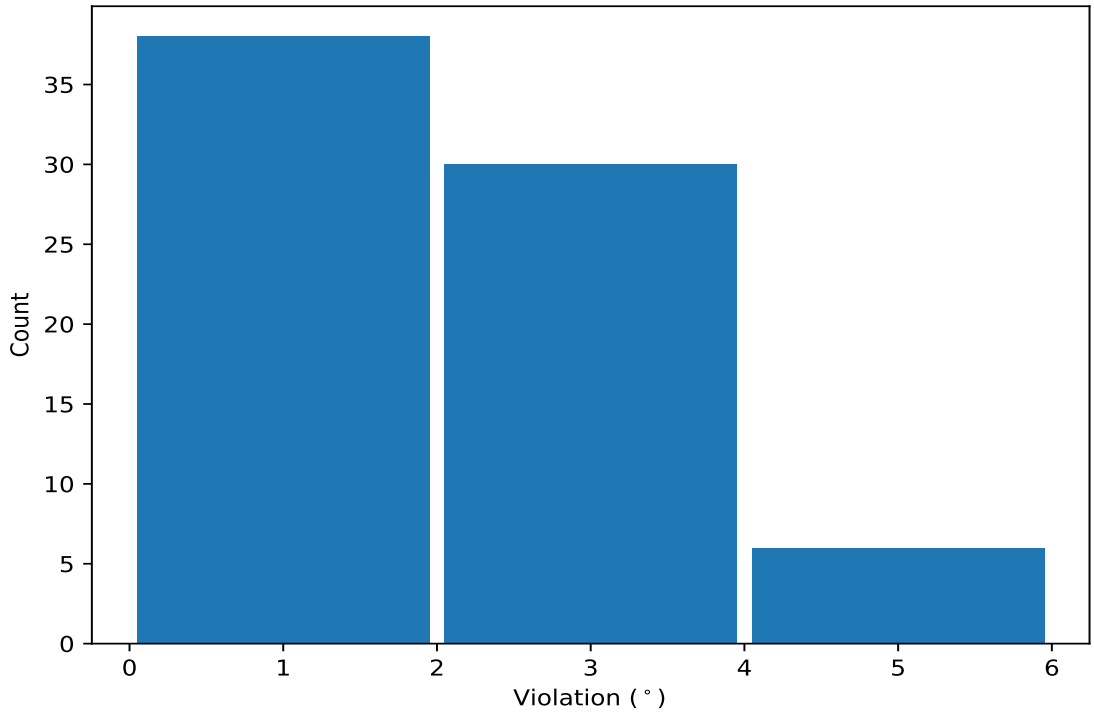
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models[1]	Mean	SD[2]	Median
(1,251)	1:129:A:THR:C	1:130:A:VAL:N	1:130:A:VAL:CA	1:130:A:VAL:C	8	2.45	0.71	2.34
(1,150)	1:79:A:ALA:N	1:79:A:ALA:CA	1:79:A:ALA:C	1:80:A:ILE:N	6	3.38	1.01	3.34
(1,106)	1:56:A:CYS:N	1:56:A:CYS:CA	1:56:A:CYS:C	1:57:A:ARG:N	5	2.45	0.8	2.24
(1,205)	1:106:A:ARG:C	1:107:A:ALA:N	1:107:A:ALA:CA	1:107:A:ALA:C	5	2.25	0.91	2.21
(1,247)	1:127:A:VAL:C	1:128:A:ILE:N	1:128:A:ILE:CA	1:128:A:ILE:C	5	1.87	0.79	1.4
(1,52)	1:27:A:HIS:N	1:27:A:HIS:CA	1:27:A:HIS:C	1:28:A:ASP:N	5	1.55	0.41	1.38
(1,213)	1:110:A:PHE:C	1:111:A:HIS:N	1:111:A:HIS:CA	1:111:A:HIS:C	4	3.48	1.68	3.48
(1,212)	1:110:A:PHE:N	1:110:A:PHE:CA	1:110:A:PHE:C	1:111:A:HIS:N	4	2.43	1.11	2.2
(1,103)	1:54:A:PRO:C	1:55:A:GLU:N	1:55:A:GLU:CA	1:55:A:GLU:C	3	1.5	0.29	1.32
(1,104)	1:55:A:GLU:N	1:55:A:GLU:CA	1:55:A:GLU:C	1:56:A:CYS:N	2	3.46	0.92	3.46

[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 (°)
(1,213)	1:110:A:PHE:C	1:111:A:HIS:N	1:111:A:HIS:CA	1:111:A:HIS:C	9	5.44
(1,213)	1:110:A:PHE:C	1:111:A:HIS:N	1:111:A:HIS:CA	1:111:A:HIS:C	4	4.84
(1,150)	1:79:A:ALA:N	1:79:A:ALA:CA	1:79:A:ALA:C	1:80:A:ILE:N	2	4.71
(1,150)	1:79:A:ALA:N	1:79:A:ALA:CA	1:79:A:ALA:C	1:80:A:ILE:N	3	4.47
(1,104)	1:55:A:GLU:N	1:55:A:GLU:CA	1:55:A:GLU:C	1:56:A:CYS:N	4	4.37
(1,212)	1:110:A:PHE:N	1:110:A:PHE:CA	1:110:A:PHE:C	1:111:A:HIS:N	9	4.19
(1,140)	1:73:A:GLN:N	1:73:A:GLN:CA	1:73:A:GLN:C	1:74:A:LEU:N	6	3.63
(1,251)	1:129:A:THR:C	1:130:A:VAL:N	1:130:A:VAL:CA	1:130:A:VAL:C	5	3.6
(1,205)	1:106:A:ARG:C	1:107:A:ALA:N	1:107:A:ALA:CA	1:107:A:ALA:C	2	3.58
(1,150)	1:79:A:ALA:N	1:79:A:ALA:CA	1:79:A:ALA:C	1:80:A:ILE:N	10	3.56