wwPDB NMR Structure Validation Summary Report

Dec 25, 2024 – 07:12 PM EST

PDB ID	:	8E1D
BMRB ID	:	31038
Title	:	NMR-derived ensemble of the TAZ2 domain of p300 bound to the microphth almia-associated transcription factor
Authors	:	Langelaan, D.N.; Branch, M.
Deposited on	:	2022-08-10

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:

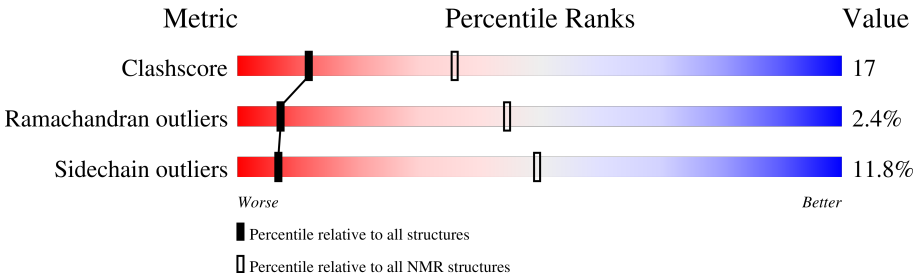
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal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

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 87%.

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	B	34	18% 9% 74%
2	A	92	63% 27% . 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, 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	B:126-B:134, A:1727-A:1811 (94)	0.91	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 3 clusters and 6 single-model clusters were found.

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

3 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 1922 atoms, of which 964 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Microphthalmia-associated transcription factor.

Mol	Chain	Residues	Atoms						Trace
1	B	34	Total	C	H	N	O	S	0
			498	154	242	39	59	4	

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	108	GLY	-	expression tag	UNP O75030

- Molecule 2 is a protein called Histone acetyltransferase p300.

Mol	Chain	Residues	Atoms						Trace
2	A	92	Total	C	H	N	O	S	0
			1421	425	722	142	122	10	

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1721	GLY	-	expression tag	UNP Q09472
A	1722	SER	-	expression tag	UNP Q09472
A	1738	ALA	CYS	engineered mutation	UNP Q09472
A	1746	ALA	CYS	engineered mutation	UNP Q09472
A	1789	ALA	CYS	engineered mutation	UNP Q09472
A	1790	ALA	CYS	engineered mutation	UNP Q09472

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	
3	A	3	Total	Zn
			3	3

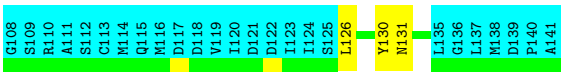
4 Residue-property plots (i)

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: Microphthalmia-associated transcription factor

Chain B:



• Molecule 2: Histone acetyltransferase p300

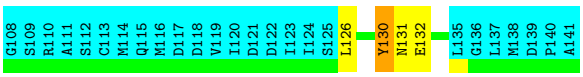
Chain A:

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.

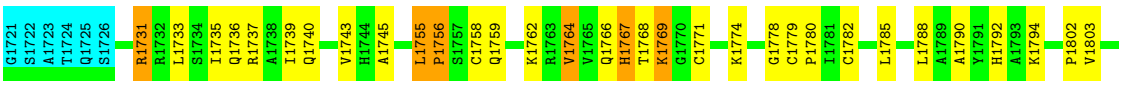
• Molecule 1: Microphthalmia-associated transcription factor

Chain B:



• Molecule 2: Histone acetyltransferase p300

Chain A:




L1807
K1812

5 Refinement protocol and experimental data overview

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

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

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

Software name	Classification	Version
ARIA	refinement	
CNS	refinement	
ARIA	structure calculation	

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

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	1557
Number of shifts mapped to atoms	1557
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	87%

6 Model quality

6.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	B	75	65	65	1±1
2	A	652	674	673	25±4
All	All	14600	14780	14760	506

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 17.

5 of 122 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:A:1782:CYS:HA	2:A:1785:LEU:HD12	0.80	1.54	19	11
2:A:1768:THR:HG21	2:A:1785:LEU:HD22	0.79	1.53	18	9
2:A:1768:THR:HA	2:A:1771:CYS:SG	0.78	2.19	3	11
2:A:1755:LEU:CD1	2:A:1758:CYS:HB3	0.78	2.09	17	8
2:A:1755:LEU:HD11	2:A:1758:CYS:HB3	0.77	1.56	8	8

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	B	9/34 (26%)	7±0 (74±5%)	1±1 (16±8%)	1±1 (10±7%)	1 9
2	A	85/92 (92%)	77±2 (91±2%)	6±2 (7±2%)	1±1 (2±1%)	10 55
All	All	1880/2520 (75%)	1681 (89%)	154 (8%)	45 (2%)	7 44

5 of 6 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
2	A	1802	PRO	18
1	B	131	ASN	14
2	A	1756	PRO	8
1	B	129	SER	2
1	B	130	TYR	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	B	9/30 (30%)	8±1 (94±7%)	0±1 (6±7%)	20 72
2	A	73/78 (94%)	64±2 (87±2%)	9±2 (13±2%)	6 48
All	All	1640/2160 (76%)	1447 (88%)	193 (12%)	7 49

5 of 31 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)
2	A	1755	LEU	20
2	A	1788	LEU	20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
2	A	1731	ARG	19
2	A	1803	VAL	19
2	A	1767	HIS	16

6.3.3 RNA

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains

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

6.5 Carbohydrates

There are no oligosaccharides in this entry.

6.6 Ligand geometry

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

6.7 Other polymers

There are no such molecules in this entry.

6.8 Polymer linkage issues

There are no chain breaks in this entry.

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 87% for the well-defined parts and 86% 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	1557
Number of shifts mapped to atoms	1557
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	1

The following errors were found when reading this chemical shift list.

- Chemical shift has been reported more than once. First 5 (of 0) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	1723	ALA	HB2	1.417	0.003	.
1	A	1723	ALA	HB3	1.417	0.003	.
1	A	1724	THR	HG22	1.186	0.003	.
1	A	1724	THR	HG23	1.186	0.003	.
1	A	1733	LEU	HD12	0.870	0.009	.
1	A	1733	LEU	HD13	0.870	0.009	.
1	A	1733	LEU	HD22	0.895	0.009	.
1	A	1733	LEU	HD23	0.895	0.009	.
1	A	1735	ILE	HG22	0.548	0.003	.
1	A	1735	ILE	HG23	0.548	0.003	.
1	A	1735	ILE	HD12	0.553	0.003	.
1	A	1735	ILE	HD13	0.553	0.003	.
1	A	1738	ALA	HB2	1.501	0.007	.
1	A	1738	ALA	HB3	1.501	0.007	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	1739	ILE	HG22	1.035	0.007	.
1	A	1739	ILE	HG23	1.035	0.007	.
1	A	1739	ILE	HD12	0.905	0.008	.
1	A	1739	ILE	HD13	0.905	0.008	.
1	A	1742	LEU	HD12	0.836	0.011	.
1	A	1742	LEU	HD13	0.836	0.011	.
1	A	1742	LEU	HD22	0.891	0.010	.
1	A	1742	LEU	HD23	0.891	0.010	.
1	A	1743	VAL	HG12	0.899	0.007	.
1	A	1743	VAL	HG13	0.899	0.007	.
1	A	1743	VAL	HG22	0.888	0.003	.
1	A	1743	VAL	HG23	0.888	0.003	.
1	A	1745	ALA	HB2	1.537	0.009	.
1	A	1745	ALA	HB3	1.537	0.009	.
1	A	1746	ALA	HB2	1.363	0.003	.
1	A	1746	ALA	HB3	1.363	0.003	.
1	A	1751	ALA	HB2	1.406	0.002	.
1	A	1751	ALA	HB3	1.406	0.002	.
1	A	1755	LEU	HD12	0.580	0.008	.
1	A	1755	LEU	HD13	0.580	0.008	.
1	A	1755	LEU	HD22	0.539	0.009	.
1	A	1755	LEU	HD23	0.539	0.009	.
1	A	1761	MET	HE2	2.037	0.006	.
1	A	1761	MET	HE3	2.037	0.006	.
1	A	1764	VAL	HG12	1.117	0.004	.
1	A	1764	VAL	HG13	1.117	0.004	.
1	A	1764	VAL	HG22	0.971	0.005	.
1	A	1764	VAL	HG23	0.971	0.005	.
1	A	1765	VAL	HG12	0.916	0.006	.
1	A	1765	VAL	HG13	0.916	0.006	.
1	A	1765	VAL	HG22	0.950	0.008	.
1	A	1765	VAL	HG23	0.950	0.008	.
1	A	1768	THR	HG22	1.629	0.003	.
1	A	1768	THR	HG23	1.629	0.003	.
1	A	1775	THR	HG22	1.169	0.002	.
1	A	1775	THR	HG23	1.169	0.002	.
1	A	1781	ILE	HG22	0.669	0.002	.
1	A	1781	ILE	HG23	0.669	0.002	.
1	A	1781	ILE	HD12	0.448	0.005	.
1	A	1781	ILE	HD13	0.448	0.005	.
1	A	1785	LEU	HD12	0.993	0.006	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	1785	LEU	HD13	0.993	0.006	.
1	A	1785	LEU	HD22	0.976	0.004	.
1	A	1785	LEU	HD23	0.976	0.004	.
1	A	1786	ILE	HG22	0.768	0.005	.
1	A	1786	ILE	HG23	0.768	0.005	.
1	A	1786	ILE	HD12	0.385	0.004	.
1	A	1786	ILE	HD13	0.385	0.004	.
1	A	1787	ALA	HB2	1.495	0.006	.
1	A	1787	ALA	HB3	1.495	0.006	.
1	A	1788	LEU	HD12	0.877	0.010	.
1	A	1788	LEU	HD13	0.877	0.010	.
1	A	1788	LEU	HD22	0.827	0.012	.
1	A	1788	LEU	HD23	0.827	0.012	.
1	A	1789	ALA	HB2	1.645	0.006	.
1	A	1789	ALA	HB3	1.645	0.006	.
1	A	1790	ALA	HB2	1.494	0.002	.
1	A	1790	ALA	HB3	1.494	0.002	.
1	A	1793	ALA	HB2	1.572	0.008	.
1	A	1793	ALA	HB3	1.572	0.008	.
1	A	1803	VAL	HG12	0.413	0.003	.
1	A	1803	VAL	HG13	0.413	0.003	.
1	A	1803	VAL	HG22	0.646	0.002	.
1	A	1803	VAL	HG23	0.646	0.002	.
1	A	1807	LEU	HD12	0.913	0.007	.
1	A	1807	LEU	HD13	0.913	0.007	.
1	A	1807	LEU	HD22	0.902	0.011	.
1	A	1807	LEU	HD23	0.902	0.011	.
1	A	1809	ILE	HG22	0.880	0.011	.
1	A	1809	ILE	HG23	0.880	0.011	.
1	A	1809	ILE	HD12	0.781	0.004	.
1	A	1809	ILE	HD13	0.781	0.004	.
1	B	111	ALA	HB2	1.379	0.004	.
1	B	111	ALA	HB3	1.379	0.004	.
1	B	119	VAL	HG12	0.899	0.006	.
1	B	119	VAL	HG13	0.899	0.006	.
1	B	119	VAL	HG22	0.870	0.005	.
1	B	119	VAL	HG23	0.870	0.005	.
1	B	120	ILE	HG22	0.849	0.007	.
1	B	120	ILE	HG23	0.849	0.007	.
1	B	120	ILE	HD12	0.796	0.013	.
1	B	120	ILE	HD13	0.796	0.013	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	B	123	ILE	HG22	0.834	0.018	.
1	B	123	ILE	HG23	0.834	0.018	.
1	B	123	ILE	HD12	0.743	0.002	.
1	B	123	ILE	HD13	0.743	0.002	.
1	B	124	ILE	HG22	0.842	0.010	.
1	B	124	ILE	HG23	0.842	0.010	.
1	B	124	ILE	HD12	0.806	0.015	.
1	B	124	ILE	HD13	0.806	0.015	.
1	B	126	LEU	HD12	0.817	0.008	.
1	B	126	LEU	HD13	0.817	0.008	.
1	B	126	LEU	HD22	0.775	0.015	.
1	B	126	LEU	HD23	0.775	0.015	.
1	B	134	ILE	HG22	0.816	0.018	.
1	B	134	ILE	HG23	0.816	0.018	.
1	B	134	ILE	HD12	0.789	0.014	.
1	B	134	ILE	HD13	0.789	0.014	.
1	B	135	LEU	HD12	0.810	0.006	.
1	B	135	LEU	HD13	0.810	0.006	.
1	B	135	LEU	HD22	0.868	0.002	.
1	B	135	LEU	HD23	0.868	0.002	.
1	B	137	LEU	HD12	0.810	0.007	.
1	B	137	LEU	HD13	0.810	0.007	.
1	B	137	LEU	HD22	0.864	0.008	.
1	B	137	LEU	HD23	0.864	0.008	.
1	B	141	ALA	HB2	1.293	0.002	.
1	B	141	ALA	HB3	1.293	0.002	.

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	126	-0.30 ± 0.15	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	119	0.12 ± 0.06	None needed (< 0.5 ppm)
$^{13}C'$	113	-0.22 ± 0.11	None needed (< 0.5 ppm)
^{15}N	115	-0.29 ± 0.27	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 87%, i.e. 1112 atoms were assigned a chemical

shift out of a possible 1277. 0 out of 11 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	451/464 (97%)	185/187 (99%)	177/188 (94%)	89/89 (100%)
Sidechain	631/756 (83%)	415/488 (85%)	202/227 (89%)	14/41 (34%)
Aromatic	30/57 (53%)	18/29 (62%)	12/23 (52%)	0/5 (0%)
Overall	1112/1277 (87%)	618/704 (88%)	391/438 (89%)	103/135 (76%)

7.1.4 Statistically unusual chemical shifts (i)

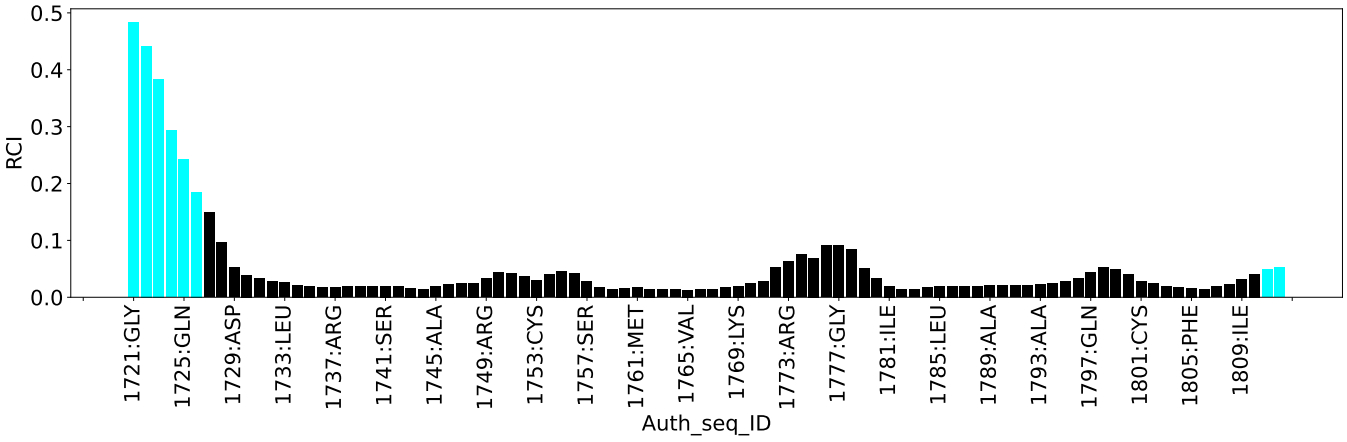
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	1768	THR	HG1	5.44	0.08 – 2.19	20.4

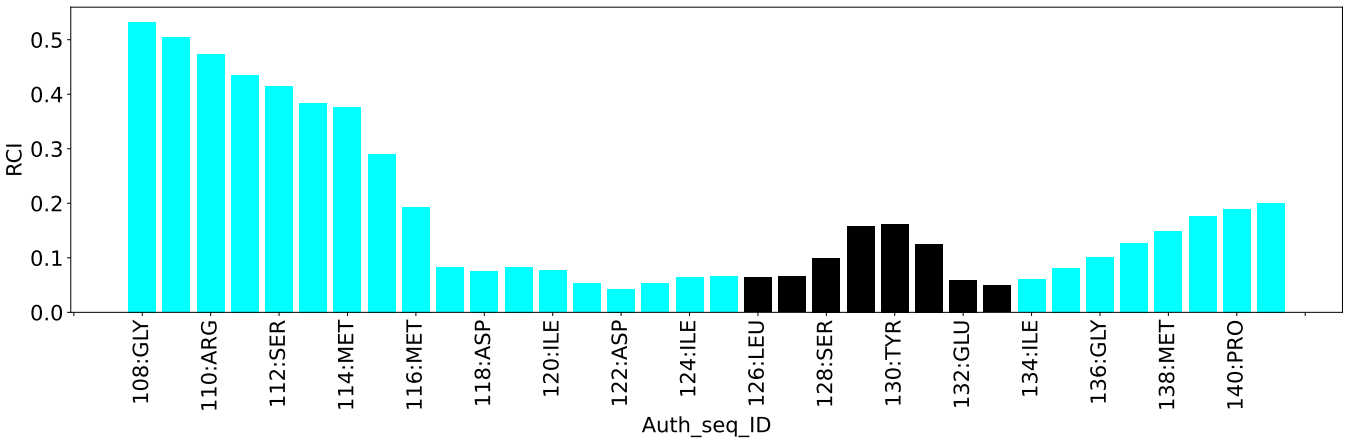
7.1.5 Random Coil Index (RCI) plots (i)

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

Random coil index (RCI) for chain A:



Random coil index (RCI) for chain B:


RCI
0.0
0.1
0.2
0.3
0.4
0.5
108:GLY
110:ARG
112:SER
114:MET
116:MET
118:ASP
120:ILE
122:ASP
124:ILE
126:LEU
128:SER
130:TYR
132:GLU
134:ILE
136:GLY
138:MET
140:PRO
Auth_seq_ID

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	1873
Intra-residue (|i-j|=0)	946
Sequential (|i-j|=1)	410
Medium range (|i-j|>1 and |i-j|<5)	257
Long range (|i-j|≥5)	155
Inter-chain	15
Hydrogen bond restraints	90
Disulfide bond restraints	0
Total dihedral-angle restraints	190
Number of unmapped restraints	0
Number of restraints per residue	16.0
Number of long range restraints per residue[1]	1.2

[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)	70.5	0.2
0.2-0.5 (Medium)	116.7	0.5
>0.5 (Large)	140.5	4.27

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)	4.6	5.05
10.0-20.0 (Medium)	None	None
>20.0 (Large)	None	None

9 Distance violation analysis (i)

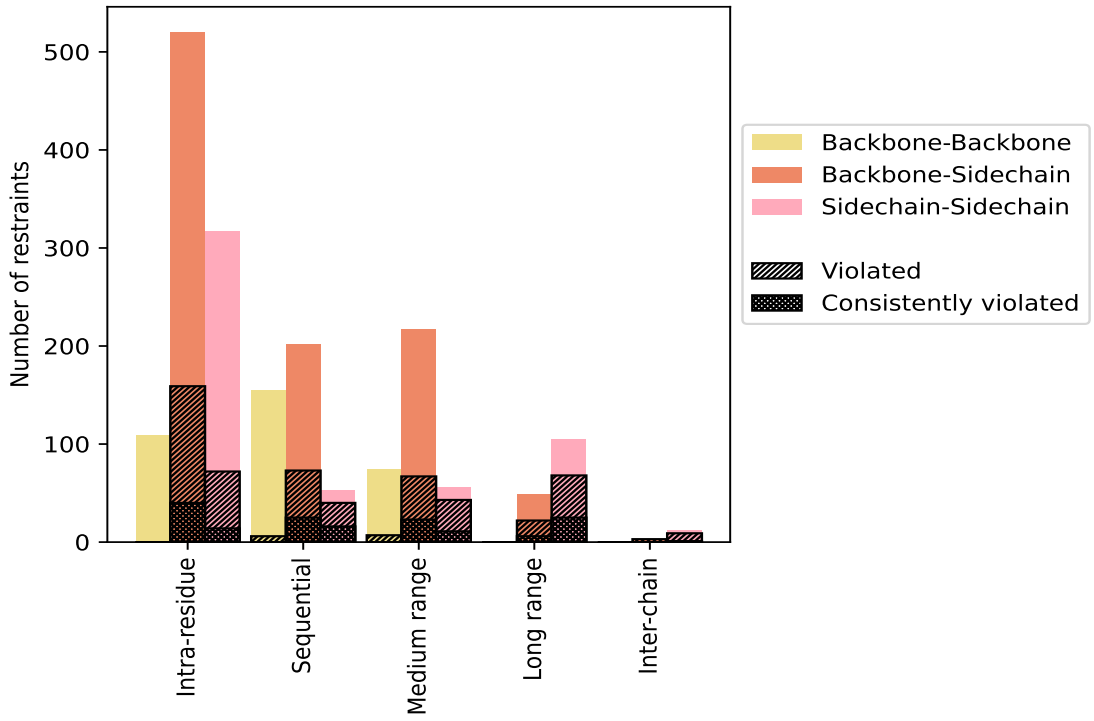
9.1 Summary of distance violations (i)

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)	946	50.5	231	24.4	12.3	54	5.7	2.9
Backbone-Backbone	109	5.8	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	520	27.8	159	30.6	8.5	40	7.7	2.1
Sidechain-Sidechain	317	16.9	72	22.7	3.8	14	4.4	0.7
Sequential (|i-j|=1)	410	21.9	119	29.0	6.4	41	10.0	2.2
Backbone-Backbone	155	8.3	6	3.9	0.3	0	0.0	0.0
Backbone-Sidechain	202	10.8	73	36.1	3.9	25	12.4	1.3
Sidechain-Sidechain	53	2.8	40	75.5	2.1	16	30.2	0.9
Medium range (|i-j|>1 & |i-j|<5)	257	13.7	114	44.4	6.1	34	13.2	1.8
Backbone-Backbone	74	4.0	7	9.5	0.4	0	0.0	0.0
Backbone-Sidechain	127	6.8	64	50.4	3.4	23	18.1	1.2
Sidechain-Sidechain	56	3.0	43	76.8	2.3	11	19.6	0.6
Long range (|i-j|≥5)	155	8.3	90	58.1	4.8	31	20.0	1.7
Backbone-Backbone	1	0.1	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	49	2.6	22	44.9	1.2	6	12.2	0.3
Sidechain-Sidechain	105	5.6	68	64.8	3.6	25	23.8	1.3
Inter-chain	15	0.8	12	80.0	0.6	1	6.7	0.1
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	3	0.2	3	100.0	0.2	0	0.0	0.0
Sidechain-Sidechain	12	0.6	9	75.0	0.5	1	8.3	0.1
Hydrogen bond	90	4.8	3	3.3	0.2	0	0.0	0.0
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	1873	100.0	569	30.4	30.4	161	8.6	8.6
Backbone-Backbone	339	18.1	13	3.8	0.7	0	0.0	0.0
Backbone-Sidechain	991	52.9	324	32.7	17.3	94	9.5	5.0
Sidechain-Sidechain	543	29.0	232	42.7	12.4	67	12.3	3.6

[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	130	72	67	55	3	327	0.55	2.58	0.45	0.43
2	132	74	61	59	4	330	0.59	3.51	0.51	0.46
3	127	73	75	57	6	338	0.57	2.77	0.48	0.45
4	128	76	69	55	4	332	0.6	3.59	0.52	0.46
5	126	76	62	64	3	331	0.59	4.27	0.55	0.44
6	126	70	70	54	4	324	0.56	3.52	0.49	0.42
7	130	73	66	59	6	334	0.58	2.44	0.47	0.44
8	134	82	63	53	5	337	0.57	2.75	0.47	0.43
9	133	77	71	53	4	338	0.6	3.33	0.51	0.46
10	134	73	56	60	4	327	0.6	3.41	0.52	0.45

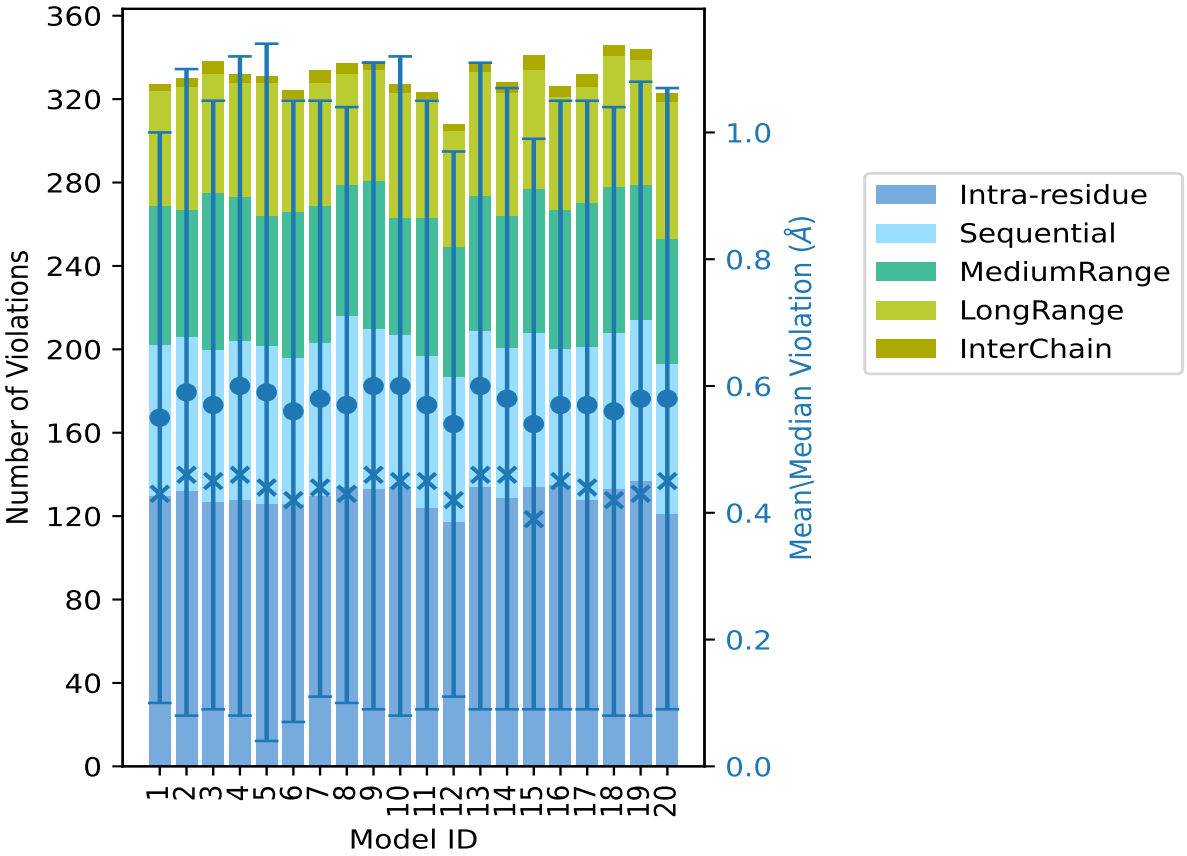
Continued on next page...

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	124	73	66	57	3	323	0.57	3.62	0.48	0.45
12	117	70	62	56	3	308	0.54	2.39	0.43	0.42
13	134	75	65	59	4	337	0.6	3.42	0.51	0.46
14	129	72	63	59	5	328	0.58	3.38	0.49	0.46
15	134	74	69	57	7	341	0.54	2.52	0.45	0.39
16	135	65	67	54	5	326	0.57	3.41	0.48	0.45
17	128	73	69	56	6	332	0.57	3.44	0.48	0.44
18	133	75	70	63	5	346	0.56	3.67	0.48	0.42
19	137	77	65	60	5	344	0.58	3.61	0.5	0.43
20	121	72	60	66	4	323	0.58	3.39	0.49	0.45

[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

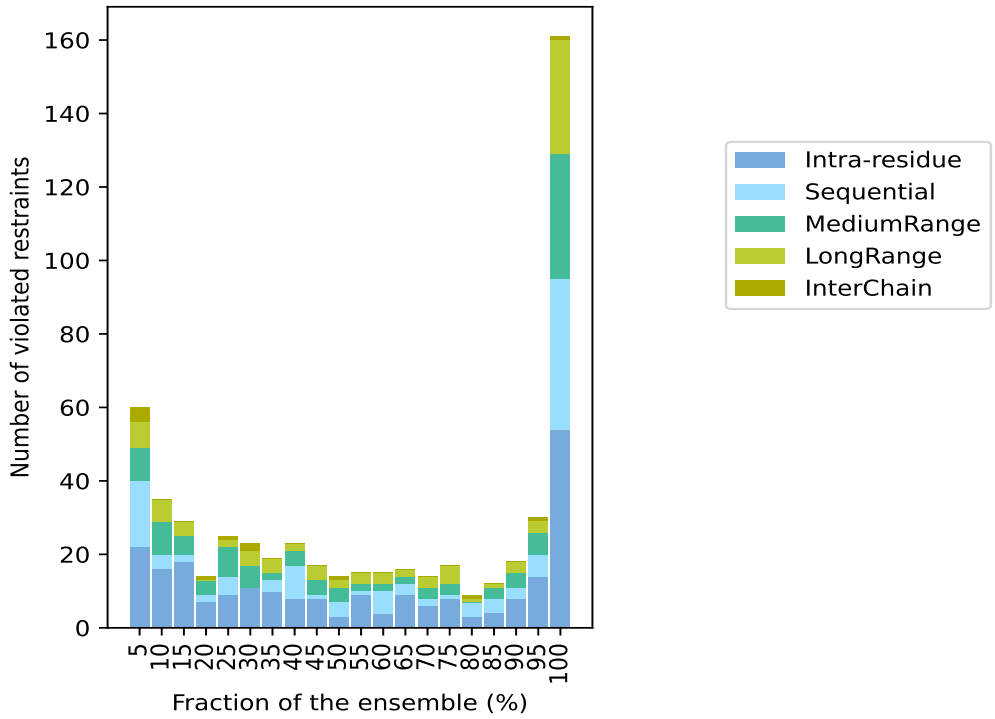
9.3 Distance violation statistics for the ensemble (i)

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, 1217(IR:715, SQ:291, MR:143, LR:65, IC:3) 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]	%
22	18	9	7	4	60	1	5.0
16	4	9	6	0	35	2	10.0
18	2	5	4	0	29	3	15.0
7	2	4	0	1	14	4	20.0
9	5	8	2	1	25	5	25.0
11	0	6	4	2	23	6	30.0
10	3	2	4	0	19	7	35.0
8	9	4	2	0	23	8	40.0
8	1	4	4	0	17	9	45.0
3	4	4	2	1	14	10	50.0
9	1	2	3	0	15	11	55.0
4	6	2	3	0	15	12	60.0
9	3	2	2	0	16	13	65.0
6	2	3	3	0	14	14	70.0
8	1	3	5	0	17	15	75.0
3	4	0	1	1	9	16	80.0
4	4	3	1	0	12	17	85.0
8	3	4	3	0	18	18	90.0
14	6	6	3	1	30	19	95.0
54	41	34	31	1	161	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

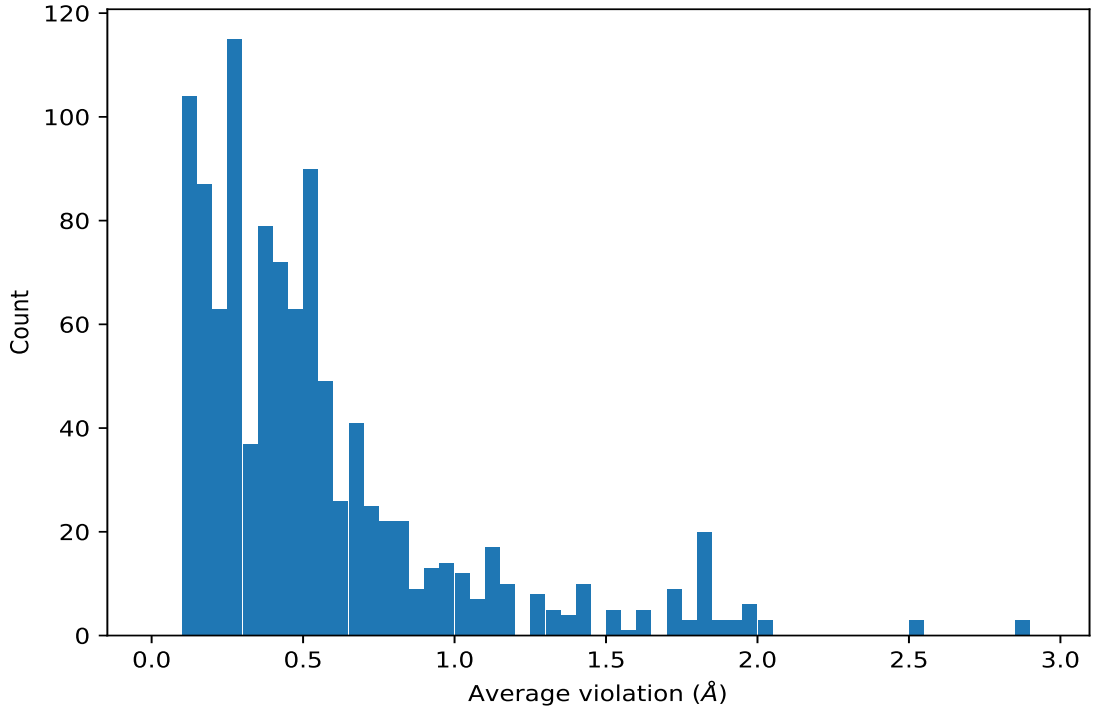
9.3.1 Bar graph : Distance violation statistics for the ensemble



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)
(2,665)	1:119:B:VAL:HG13	1:120:B:ILE:HG13	20	2.9	0.85	3.41
(2,665)	1:119:B:VAL:HG12	1:120:B:ILE:HG13	20	2.9	0.85	3.41
(2,665)	1:119:B:VAL:HG11	1:120:B:ILE:HG13	20	2.9	0.85	3.41
(2,1015)	2:1735:A:ILE:HB	2:1788:A:LEU:HD23	20	2.53	0.13	2.52
(2,1015)	2:1735:A:ILE:HB	2:1788:A:LEU:HD21	20	2.53	0.13	2.52
(2,1015)	2:1735:A:ILE:HB	2:1788:A:LEU:HD22	20	2.53	0.13	2.52
(2,1634)	2:1788:A:LEU:HD21	2:1791:A:TYR:HD2	20	1.96	0.06	1.97
(2,1634)	2:1788:A:LEU:HD22	2:1791:A:TYR:HD1	20	1.96	0.06	1.97
(2,1634)	2:1788:A:LEU:HD23	2:1791:A:TYR:HD2	20	1.96	0.06	1.97
(2,1634)	2:1788:A:LEU:HD23	2:1791:A:TYR:HD1	20	1.96	0.06	1.97
(2,1634)	2:1788:A:LEU:HD22	2:1791:A:TYR:HD2	20	1.96	0.06	1.97
(2,1634)	2:1788:A:LEU:HD21	2:1791:A:TYR:HD1	20	1.96	0.06	1.97
(2,561)	1:119:B:VAL:HG21	1:118:B:ASP:HB3	20	1.95	0.41	2.1
(2,561)	1:119:B:VAL:HG23	1:118:B:ASP:HB3	20	1.95	0.41	2.1
(2,561)	1:119:B:VAL:HG22	1:118:B:ASP:HB3	20	1.95	0.41	2.1
(2,1024)	2:1742:A:LEU:HA	2:1742:A:LEU:HD22	20	1.84	0.01	1.84

Continued on next page...

Continued from previous page...

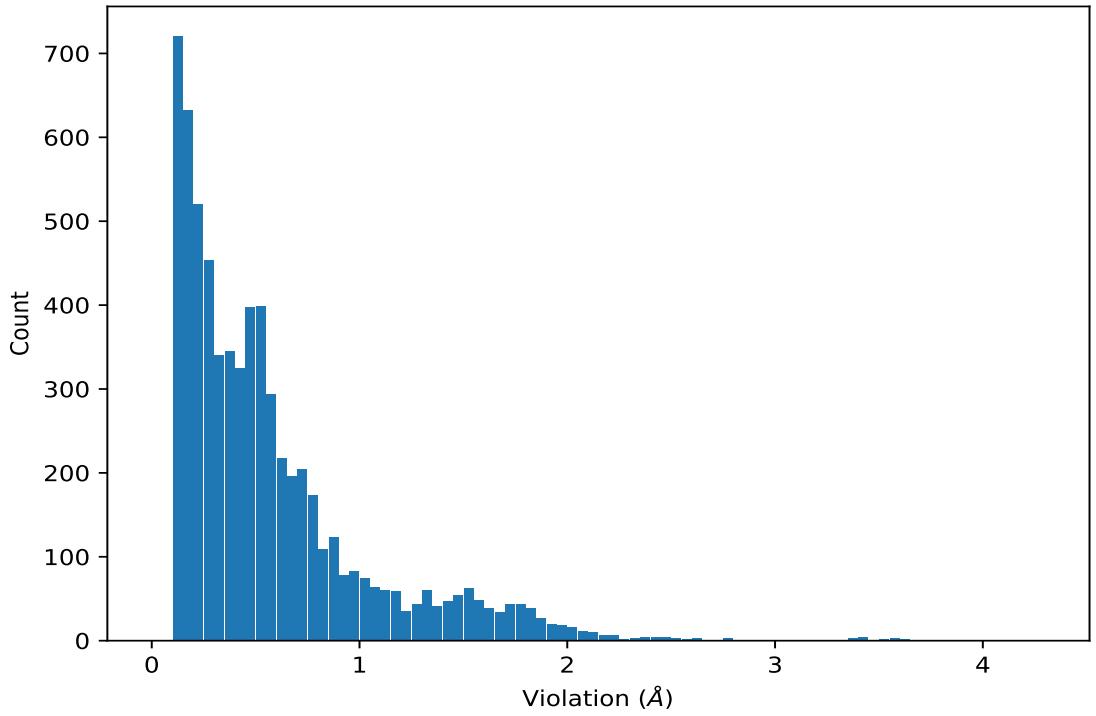
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(2,1024)	2:1742:A:LEU:HA	2:1742:A:LEU:HD21	20	1.84	0.01	1.84
(2,1024)	2:1742:A:LEU:HA	2:1742:A:LEU:HD23	20	1.84	0.01	1.84
(2,1000)	2:1738:A:ALA:HB3	2:1788:A:LEU:HD11	20	1.82	0.25	1.85
(2,1000)	2:1738:A:ALA:HB3	2:1788:A:LEU:HD13	20	1.82	0.25	1.85
(2,1000)	2:1738:A:ALA:HB1	2:1788:A:LEU:HD12	20	1.82	0.25	1.85
(2,1000)	2:1738:A:ALA:HB1	2:1788:A:LEU:HD13	20	1.82	0.25	1.85
(2,1000)	2:1738:A:ALA:HB2	2:1788:A:LEU:HD11	20	1.82	0.25	1.85
(2,1000)	2:1738:A:ALA:HB1	2:1788:A:LEU:HD11	20	1.82	0.25	1.85
(2,1000)	2:1738:A:ALA:HB2	2:1788:A:LEU:HD12	20	1.82	0.25	1.85
(2,1000)	2:1738:A:ALA:HB3	2:1788:A:LEU:HD12	20	1.82	0.25	1.85
(2,1404)	2:1788:A:LEU:HD11	2:1739:A:ILE:H	20	1.82	0.19	1.81
(2,1404)	2:1788:A:LEU:HD12	2:1739:A:ILE:H	20	1.82	0.19	1.81
(2,1404)	2:1788:A:LEU:HD13	2:1739:A:ILE:H	20	1.82	0.19	1.81
(2,975)	2:1788:A:LEU:HA	2:1788:A:LEU:HD22	20	1.81	0.04	1.82
(2,975)	2:1788:A:LEU:HA	2:1788:A:LEU:HD23	20	1.81	0.04	1.82
(2,975)	2:1788:A:LEU:HA	2:1788:A:LEU:HD21	20	1.81	0.04	1.82
(2,1500)	2:1781:A:ILE:HD11	2:1780:A:PRO:HD3	20	1.81	0.15	1.81
(2,1500)	2:1781:A:ILE:HD12	2:1780:A:PRO:HD3	20	1.81	0.15	1.81
(2,1500)	2:1781:A:ILE:HD13	2:1780:A:PRO:HD3	20	1.81	0.15	1.81
(2,1443)	2:1791:A:TYR:HB2	2:1739:A:ILE:HD11	20	1.75	0.04	1.75

[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)
(2,636)	1:135:B:LEU:HD21	1:132:B:GLU:HB2	5	4.27
(2,665)	1:119:B:VAL:HG11	1:120:B:ILE:HG13	18	3.67
(2,665)	1:119:B:VAL:HG13	1:120:B:ILE:HG13	11	3.62
(2,665)	1:119:B:VAL:HG13	1:120:B:ILE:HG13	19	3.61
(2,680)	1:138:B:MET:HG2	1:137:B:LEU:HD11	4	3.59
(2,665)	1:119:B:VAL:HG12	1:120:B:ILE:HG13	4	3.59
(2,637)	1:135:B:LEU:HD23	1:132:B:GLU:HG2	5	3.58
(2,665)	1:119:B:VAL:HG11	1:120:B:ILE:HG13	6	3.52
(2,665)	1:119:B:VAL:HG12	1:120:B:ILE:HG13	2	3.51
(2,665)	1:119:B:VAL:HG12	1:120:B:ILE:HG13	5	3.47

10 Dihedral-angle violation analysis

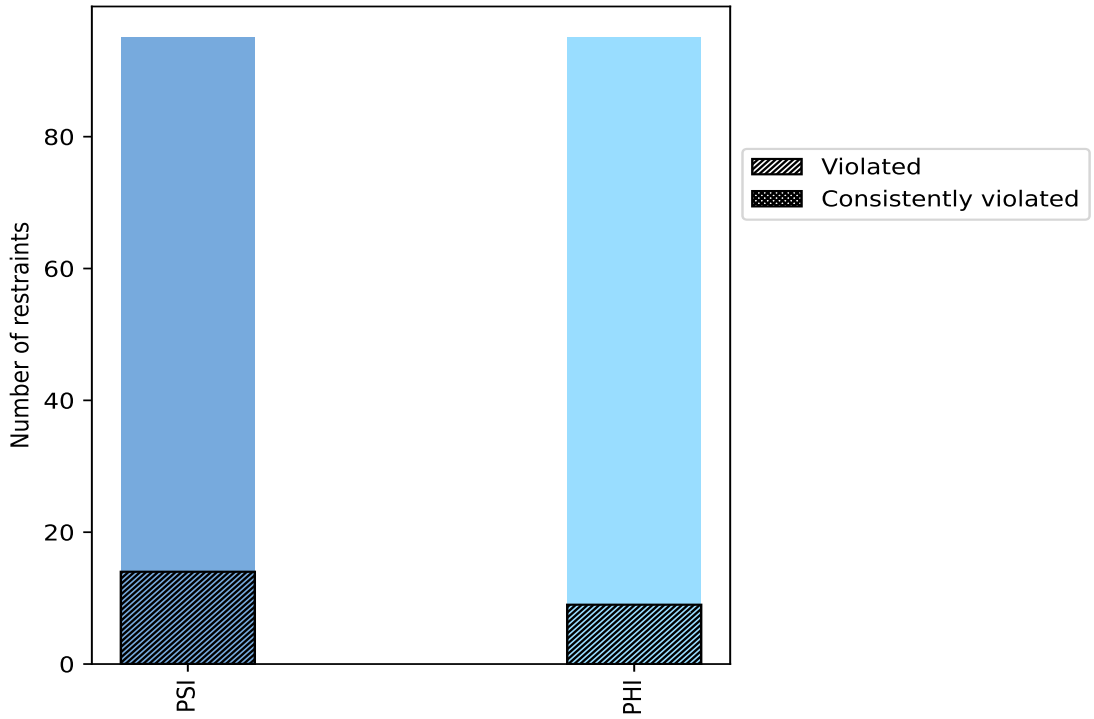
10.1 Summary of dihedral-angle violations

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	95	50.0	14	14.7	7.4	0	0.0	0.0
PHI	95	50.0	9	9.5	4.7	0	0.0	0.0
Total	190	100.0	23	12.1	12.1	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



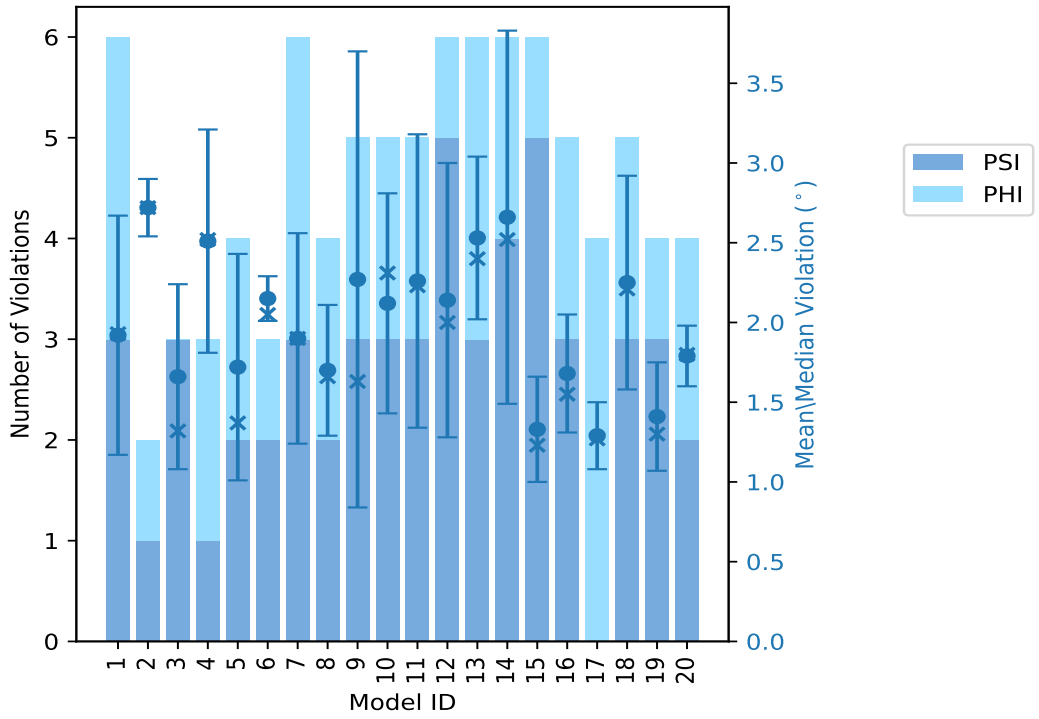
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

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	3	3	6	1.92	2.87	0.75	1.93
2	1	1	2	2.72	2.9	0.18	2.72
3	3	0	3	1.66	2.47	0.58	1.32
4	1	2	3	2.51	3.36	0.7	2.52
5	2	2	4	1.72	2.93	0.71	1.37
6	2	1	3	2.15	2.35	0.14	2.05
7	3	3	6	1.9	2.72	0.66	1.9
8	2	2	4	1.7	2.19	0.41	1.66
9	3	2	5	2.27	5.05	1.43	1.63
10	3	2	5	2.12	3.01	0.69	2.31
11	3	2	5	2.26	3.74	0.92	2.23
12	5	1	6	2.14	3.86	0.86	2.0
13	3	3	6	2.53	3.51	0.51	2.4
14	4	2	6	2.66	4.96	1.17	2.52
15	5	1	6	1.33	2.06	0.33	1.23
16	3	2	5	1.68	2.22	0.37	1.55
17	0	4	4	1.29	1.61	0.21	1.27
18	3	2	5	2.25	3.42	0.67	2.21
19	3	1	4	1.41	1.95	0.34	1.3
20	2	2	4	1.79	2.03	0.19	1.8

10.2.1 Bar graph : Dihedral violation statistics for each model (i)



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

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

Violation analysis may find that some restraints are violated in very few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of ensemble.

Number of violated restraints			Fraction of the ensemble	
PSI	PHI	Total	Count[1]	%
5	6	11	1	5.0
4	1	5	2	10.0
0	0	0	3	15.0
1	0	1	4	20.0
0	0	0	5	25.0
1	0	1	6	30.0
0	0	0	7	35.0
1	0	1	8	40.0
1	0	1	9	45.0
0	0	0	10	50.0
0	0	0	11	55.0

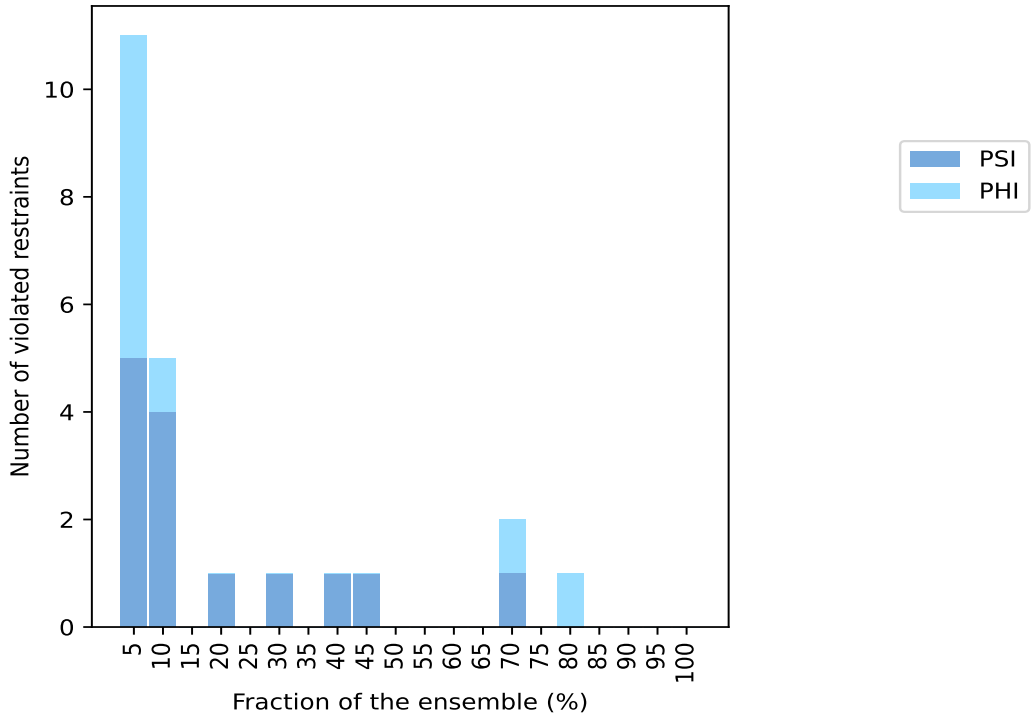
Continued on next page...

Continued from previous page...

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

[1] Number of models with violations

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

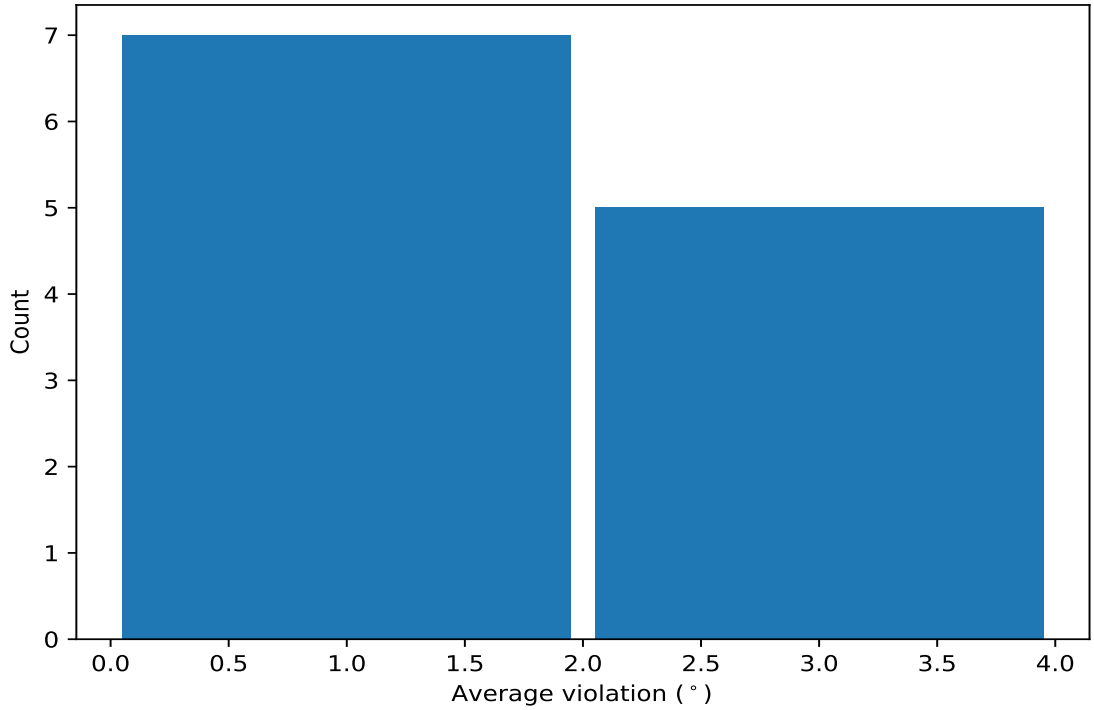


10.4 Most violated dihedral-angle restraints in the ensemble

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.

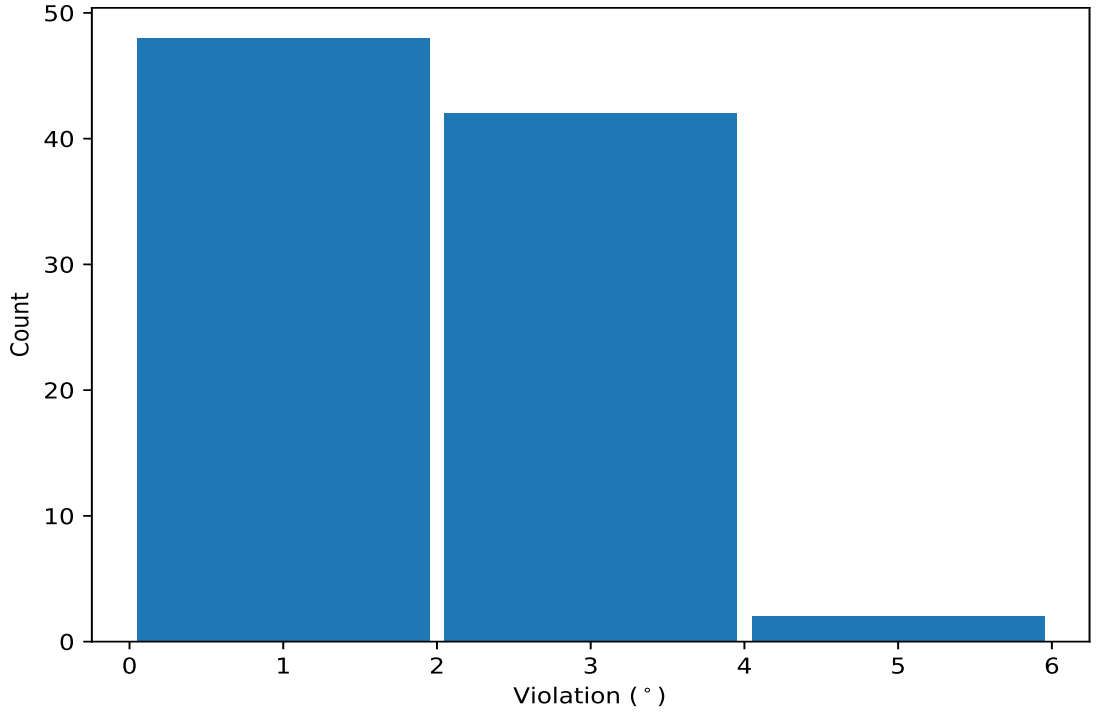
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models[1]	Mean	SD[2]	Medi
(1,57)	2:1756:A:PRO:C	2:1757:A:SER:N	2:1757:A:SER:CA	2:1757:A:SER:C	16	2.03	0.91	1.94
(1,178)	1:132:B:GLU:N	1:132:B:GLU:CA	1:132:B:GLU:C	1:133:B:GLU:N	14	2.56	1.03	2.42
(1,129)	2:1803:A:VAL:C	2:1804:A:PRO:N	2:1804:A:PRO:CA	2:1804:A:PRO:C	14	2.26	0.44	2.19
(1,132)	2:1805:A:PHE:N	2:1805:A:PHE:CA	2:1805:A:PHE:C	2:1806:A:CYS:N	9	1.61	0.58	1.29
(1,130)	2:1804:A:PRO:N	2:1804:A:PRO:CA	2:1804:A:PRO:C	2:1805:A:PHE:N	8	2.02	0.57	2.14
(1,92)	2:1779:A:CYS:N	2:1779:A:CYS:CA	2:1779:A:CYS:C	2:1780:A:PRO:N	6	1.99	0.91	1.62
(1,184)	1:135:B:LEU:N	1:135:B:LEU:CA	1:135:B:LEU:C	1:136:B:GLY:N	4	1.95	0.71	1.93
(1,83)	2:1773:A:ARG:C	2:1774:A:LYS:N	2:1774:A:LYS:CA	2:1774:A:LYS:C	2	2.19	0.51	2.19
(1,2)	2:1723:A:ALA:N	2:1723:A:ALA:CA	2:1723:A:ALA:C	2:1724:A:THR:N	2	1.92	0.02	1.92
(1,80)	2:1768:A:THR:N	2:1768:A:THR:CA	2:1768:A:THR:C	2:1769:A:LYS:N	2	1.66	0.55	1.66

[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 (°)
(1,57)	2:1756:A:PRO:C	2:1757:A:SER:N	2:1757:A:SER:CA	2:1757:A:SER:C	9	5.05
(1,178)	1:132:B:GLU:N	1:132:B:GLU:CA	1:132:B:GLU:C	1:133:B:GLU:N	14	4.96
(1,178)	1:132:B:GLU:N	1:132:B:GLU:CA	1:132:B:GLU:C	1:133:B:GLU:N	12	3.86
(1,92)	2:1779:A:CYS:N	2:1779:A:CYS:CA	2:1779:A:CYS:C	2:1780:A:PRO:N	11	3.74
(1,123)	2:1794:A:LYS:C	2:1795:A:HIS:N	2:1795:A:HIS:CA	2:1795:A:HIS:C	13	3.51
(1,178)	1:132:B:GLU:N	1:132:B:GLU:CA	1:132:B:GLU:C	1:133:B:GLU:N	18	3.42
(1,178)	1:132:B:GLU:N	1:132:B:GLU:CA	1:132:B:GLU:C	1:133:B:GLU:N	4	3.36
(1,57)	2:1756:A:PRO:C	2:1757:A:SER:N	2:1757:A:SER:CA	2:1757:A:SER:C	10	3.01
(1,129)	2:1803:A:VAL:C	2:1804:A:PRO:N	2:1804:A:PRO:CA	2:1804:A:PRO:C	5	2.93
(1,129)	2:1803:A:VAL:C	2:1804:A:PRO:N	2:1804:A:PRO:CA	2:1804:A:PRO:C	2	2.9