wwPDB NMR Structure Validation Summary Report (i)

Jun 3, 2023 – 01:34 PM EDT

PDB ID	:	6UJU
BMRB ID	:	30678
Title	:	Structure of the HIV-1 gp41 transmembrane domain and cytoplasmic tail (LLP2)
Authors	:	Piai, A.; Fu, Q.; Cai, Y.; Ghantous, F.; Xiao, T.; Shaik, M.M.; Peng, H.; Rits-Volloch, S.; Liu, Z.; Chen, W.; Seaman, M.S.; Chen, B.; Chou, J.J.
Deposited on	:	2019-10-03

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:

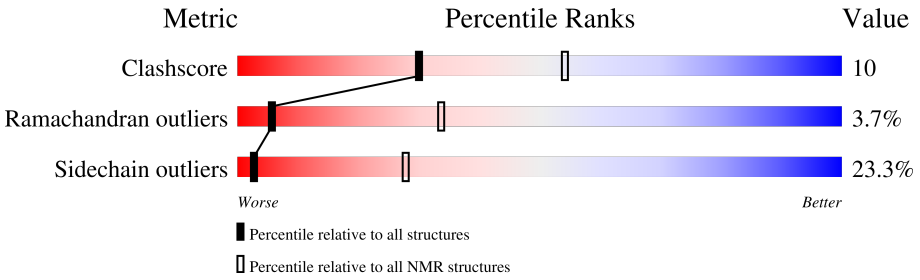
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.33

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 9%.

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	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

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	112	50% 29% • 20%
1	B	112	49% 30% • 20%
1	C	112	46% 34% • 20%

2 Ensemble composition and analysis (i)

This entry contains 15 models. Model 13 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:679-A:716, A:739-A:788, B:677-B:716, B:739-B:788, C:677-C:716, C:739-C:788 (268)	1.73	13

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 1 single-model cluster was found.

Cluster number	Models
1	1, 5, 6, 8, 11, 12, 15
2	2, 3, 4, 7, 13
3	9, 14
Single-model clusters	10

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 4686 atoms, of which 2427 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Envelope glycoprotein GP41.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
1	A	90	1562	490	809	141	122	0
1	B	90	1562	490	809	141	122	0
1	C	90	1562	490	809	141	122	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	683	ARG	LYS	conflict	UNP A0A386YSI0
A	687	ILE	MET	conflict	UNP A0A386YSI0
A	691	SER	GLY	conflict	UNP A0A386YSI0
A	693	ILE	VAL	conflict	UNP A0A386YSI0
A	704	LEU	ILE	conflict	UNP A0A386YSI0
A	764	SER	CYS	engineered mutation	UNP A0A386YSI0
B	683	ARG	LYS	conflict	UNP A0A386YSI0
B	687	ILE	MET	conflict	UNP A0A386YSI0
B	691	SER	GLY	conflict	UNP A0A386YSI0
B	693	ILE	VAL	conflict	UNP A0A386YSI0
B	704	LEU	ILE	conflict	UNP A0A386YSI0
B	764	SER	CYS	engineered mutation	UNP A0A386YSI0
C	683	ARG	LYS	conflict	UNP A0A386YSI0
C	687	ILE	MET	conflict	UNP A0A386YSI0
C	691	SER	GLY	conflict	UNP A0A386YSI0
C	693	ILE	VAL	conflict	UNP A0A386YSI0
C	704	LEU	ILE	conflict	UNP A0A386YSI0
C	764	SER	CYS	engineered mutation	UNP A0A386YSI0

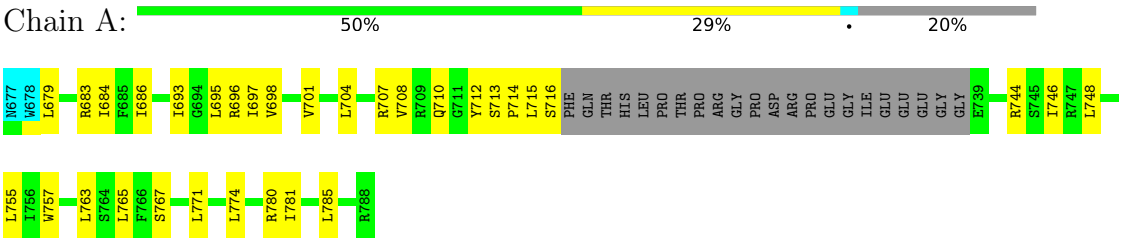
4 Residue-property plots (i)

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.

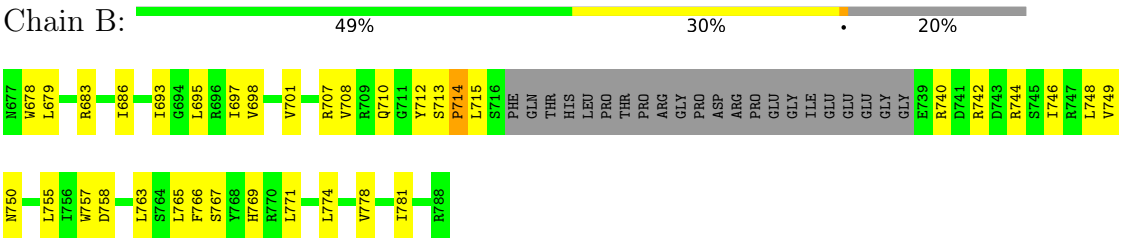
• Molecule 1: Envelope glycoprotein GP41

Chain A:



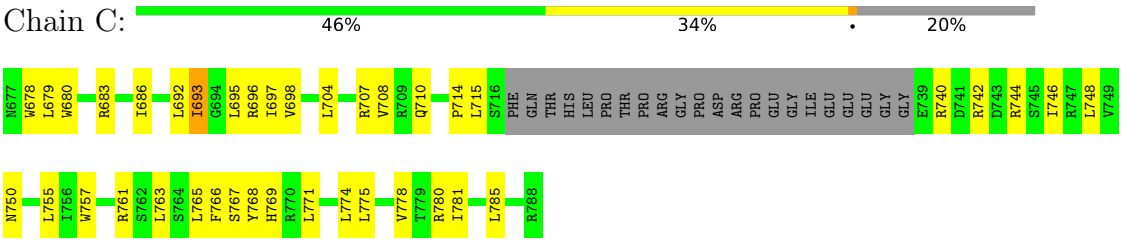
• Molecule 1: Envelope glycoprotein GP41

Chain B:



• Molecule 1: Envelope glycoprotein GP41

Chain C:

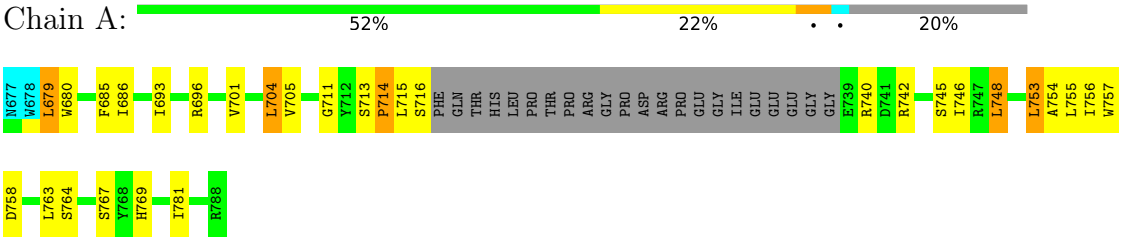


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 13. Colouring as in section 4.1 above.

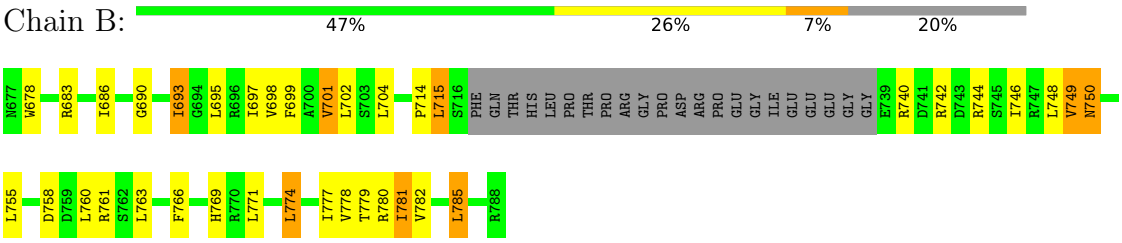
• Molecule 1: Envelope glycoprotein GP41

Chain A:



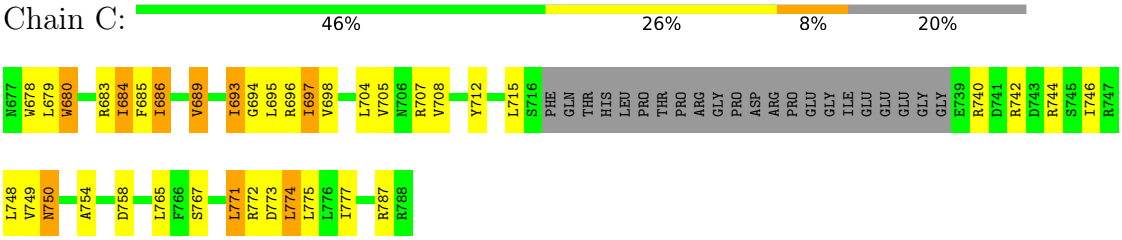
• Molecule 1: Envelope glycoprotein GP41

Chain B:



• Molecule 1: Envelope glycoprotein GP41

Chain C:



5 Refinement protocol and experimental data overview (i)

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

Of the 150 calculated structures, 15 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
X-PLOR NIH	refinement	
X-PLOR NIH	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	460
Number of shifts mapped to atoms	367
Number of unparsed shifts	0
Number of shifts with mapping errors	93
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	9%

6 Model quality (i)

6.1 Standard geometry (i)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the (average) root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	0.75±0.01	0±0/741 (0.0± 0.0%)	0.91±0.02	0±0/998 (0.0± 0.0%)
1	B	0.74±0.01	0±0/765 (0.0± 0.0%)	0.92±0.01	0±0/1032 (0.0± 0.0%)
1	C	0.75±0.01	0±0/765 (0.0± 0.0%)	0.92±0.02	0±0/1032 (0.0± 0.0%)
All	All	0.75	0/34065 (0.0%)	0.92	2/45930 (0.0%)

There are no bond-length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	C	742	ARG	NE-CZ-NH2	5.22	122.91	120.30	7	1
1	C	768	TYR	CB-CG-CD2	-5.06	117.96	121.00	4	1

There are no chirality outliers.

There are no planarity outliers.

6.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	731	791	789	19±5
1	B	753	809	804	19±6
1	C	753	809	804	20±5
All	All	33555	36135	35955	693

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 10.

5 of 471 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:781:ILE:HG23	1:B:746:ILE:HG23	1.06	1.23	12	2
1:A:754:ALA:HB2	1:B:715:LEU:HD22	1.02	1.21	4	1
1:A:746:ILE:HG23	1:C:781:ILE:HD11	0.84	1.49	15	1
1:B:692:LEU:HD13	1:B:693:ILE:N	0.82	1.90	1	1
1:B:746:ILE:O	1:B:749:VAL:HG22	0.81	1.76	3	2

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	85/112 (76%)	78±2 (91±2%)	5±1 (5±2%)	3±1 (3±2%)	7	39
1	B	86/112 (77%)	78±1 (91±1%)	4±1 (5±1%)	3±1 (4±1%)	5	31
1	C	86/112 (77%)	78±1 (91±1%)	4±1 (5±2%)	3±1 (4±2%)	5	31
All	All	3855/5040 (76%)	3516 (91%)	196 (5%)	143 (4%)	6	34

5 of 27 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	B	714	PRO	14
1	A	714	PRO	12
1	C	714	PRO	10
1	B	767	SER	8
1	C	765	LEU	8

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	81/101 (80%)	63±4 (78±5%)	18±4 (22±5%)	3	30
1	B	83/101 (82%)	65±3 (79±4%)	18±3 (21±4%)	3	31
1	C	83/101 (82%)	61±3 (74±3%)	22±3 (26±3%)	2	22
All	All	3705/4545 (82%)	2841 (77%)	864 (23%)	3	28

5 of 216 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	B	683	ARG	13
1	C	693	ILE	13
1	C	704	LEU	12
1	A	712	TYR	11
1	B	744	ARG	11

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 monosaccharides 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 (i)

The completeness of assignment taking into account all chemical shift lists is 9% for the well-defined parts and 9% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *HIV-1_gp41_TMD-CTLLP2_chemical_shifts.tab*

7.1.1 Bookkeeping (i)

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

Total number of shifts	460
Number of shifts mapped to atoms	367
Number of unparsed shifts	0
Number of shifts with mapping errors	93
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

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

- No matching atom found in the structure. First 5 (of 93) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	717	PHE	H	8.06	0.01	1
1	A	717	PHE	C	175.39	0.05	1
1	A	717	PHE	CA	57.46	0.05	1
1	A	717	PHE	CB	38.33	0.05	1
1	A	717	PHE	N	120.76	0.03	1
1	A	718	GLN	H	8.15	0.01	1
1	A	718	GLN	C	175.83	0.05	1
1	A	718	GLN	CA	55.52	0.05	1
1	A	718	GLN	CB	28.72	0.05	1
1	A	718	GLN	N	119.94	0.03	1
1	A	719	THR	H	8.04	0.01	1
1	A	719	THR	CA	61.73	0.05	1
1	A	719	THR	CB	68.95	0.05	1
1	A	719	THR	N	113.98	0.03	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	720	HIS	C	174.61	0.05	1
1	A	720	HIS	CA	55.59	0.05	1
1	A	720	HIS	CB	29.94	0.05	1
1	A	721	LEU	H	8.03	0.01	1
1	A	721	LEU	CA	52.32	0.05	1
1	A	721	LEU	CB	40.6	0.05	1
1	A	721	LEU	N	123.72	0.03	1
1	A	722	PRO	C	176.55	0.05	1
1	A	722	PRO	CA	62.43	0.05	1
1	A	722	PRO	CB	30.9	0.05	1
1	A	723	THR	H	8.17	0.01	1
1	A	723	THR	CA	59.3	0.05	1
1	A	723	THR	CB	68.95	0.05	1
1	A	723	THR	N	116.44	0.03	1
1	A	724	PRO	C	176.7	0.05	1
1	A	724	PRO	CA	62.79	0.05	1
1	A	724	PRO	CB	31.23	0.05	1
1	A	725	ARG	H	8.37	0.01	1
1	A	725	ARG	C	176.44	0.05	1
1	A	725	ARG	CA	55.51	0.05	1
1	A	725	ARG	CB	30.18	0.05	1
1	A	725	ARG	N	121.04	0.03	1
1	A	726	GLY	H	8.23	0.01	1
1	A	726	GLY	C	171.92	0.05	1
1	A	726	GLY	CA	44.04	0.05	1
1	A	726	GLY	N	110.02	0.03	1
1	A	727	PRO	C	176.72	0.05	1
1	A	727	PRO	CA	62.84	0.05	1
1	A	727	PRO	CB	31.14	0.05	1
1	A	728	ASP	H	8.38	0.01	1
1	A	728	ASP	C	175.65	0.05	1
1	A	728	ASP	CA	54.07	0.05	1
1	A	728	ASP	CB	40.05	0.05	1
1	A	728	ASP	N	119.29	0.03	1
1	A	729	ARG	H	8.02	0.01	1
1	A	729	ARG	C	174.0	0.05	1
1	A	729	ARG	CA	53.08	0.05	1
1	A	729	ARG	CB	29.34	0.05	1
1	A	729	ARG	N	120.86	0.03	1
1	A	730	PRO	C	176.84	0.05	1
1	A	730	PRO	CA	62.79	0.05	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	730	PRO	CB	31.09	0.05	1
1	A	731	GLU	H	8.57	0.01	1
1	A	731	GLU	C	177.03	0.05	1
1	A	731	GLU	CA	56.43	0.05	1
1	A	731	GLU	CB	29.29	0.05	1
1	A	731	GLU	N	120.32	0.03	1
1	A	732	GLY	H	8.36	0.01	1
1	A	732	GLY	C	173.94	0.05	1
1	A	732	GLY	CA	44.91	0.05	1
1	A	732	GLY	N	109.52	0.03	1
1	A	733	ILE	H	7.82	0.01	1
1	A	733	ILE	C	176.16	0.05	1
1	A	733	ILE	CA	60.51	0.05	1
1	A	733	ILE	CB	37.61	0.05	1
1	A	733	ILE	N	119.17	0.03	1
1	A	734	GLU	H	8.49	0.01	1
1	A	734	GLU	C	176.36	0.05	1
1	A	734	GLU	CA	56.11	0.05	1
1	A	734	GLU	CB	29.31	0.05	1
1	A	734	GLU	N	124.32	0.03	1
1	A	735	GLU	H	8.38	0.01	1
1	A	735	GLU	C	176.48	0.05	1
1	A	735	GLU	CA	56.11	0.05	1
1	A	735	GLU	CB	29.46	0.05	1
1	A	735	GLU	N	121.83	0.03	1
1	A	736	GLU	H	8.47	0.01	1
1	A	736	GLU	C	177.09	0.05	1
1	A	736	GLU	CA	56.32	0.05	1
1	A	736	GLU	CB	29.38	0.05	1
1	A	736	GLU	N	122.08	0.03	1
1	A	737	GLY	H	8.51	0.01	1
1	A	737	GLY	C	174.82	0.05	1
1	A	737	GLY	CA	45.04	0.05	1
1	A	737	GLY	N	109.96	0.03	1
1	A	738	GLY	H	8.27	0.01	1
1	A	738	GLY	C	174.5	0.05	1
1	A	738	GLY	CA	44.95	0.05	1
1	A	738	GLY	N	108.53	0.03	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
$^{13}C_{\alpha}$	106	-0.04 ± 0.15	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	55	1.27 ± 0.13	Should be checked
$^{13}C'$	101	-0.16 ± 0.08	None needed (< 0.5 ppm)
^{15}N	99	0.59 ± 0.15	Should be applied

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 9%, i.e. 365 atoms were assigned a chemical shift out of a possible 4192. 0 out of 78 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	328/1349 (24%)	82/548 (15%)	164/536 (31%)	82/265 (31%)
Sidechain	37/2552 (1%)	0/1677 (0%)	37/738 (5%)	0/137 (0%)
Aromatic	0/291 (0%)	0/141 (0%)	0/136 (0%)	0/14 (0%)
Overall	365/4192 (9%)	82/2366 (3%)	201/1410 (14%)	82/416 (20%)

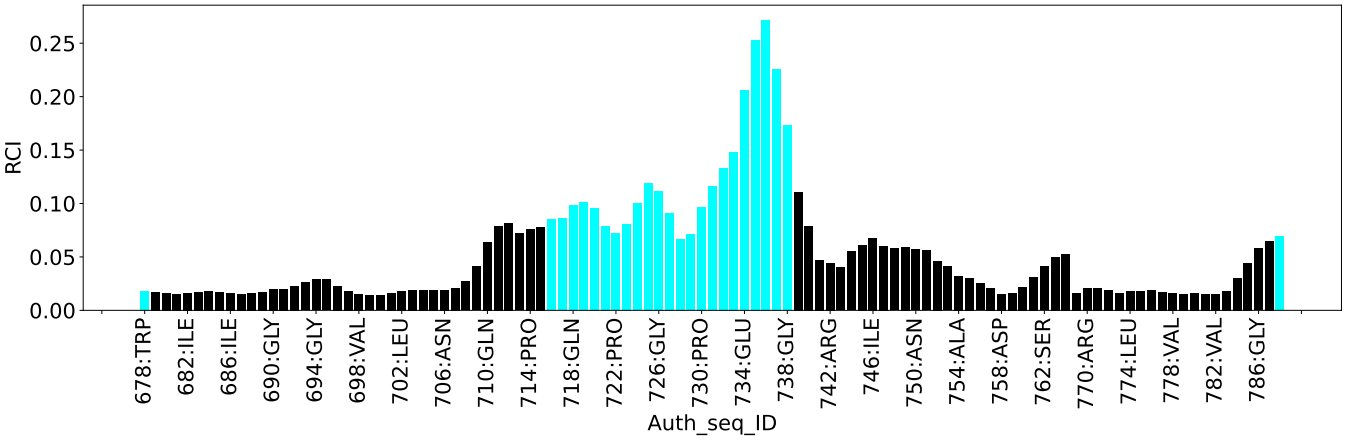
7.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

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:


RCI
0.00
0.05
0.10
0.15
0.20
0.25
678:TRP
682:ILE
686:ILE
690:GLY
694:GLY
698:VAL
702:LEU
706:ASN
710:GLN
714:PRO
718:GLN
722:PRO
726:GLY
730:PRO
734:GLU
738:GLY
742:ARG
746:ILE
750:ASN
754:ALA
758:ASP
762:SER
770:ARG
774:LEU
778:VAL
782:VAL
786:GLY
Auth_seq_ID

8 NMR restraints analysis (i)

8.1 Conformationally restricting restraints (i)

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

Description	Value
Total distance restraints	1173
Intra-residue (|i-j|=0)	159
Sequential (|i-j|=1)	372
Medium range (|i-j|>1 and |i-j|<5)	480
Long range (|i-j|≥5)	84
Inter-chain	0
Hydrogen bond restraints	78
Disulfide bond restraints	0
Total dihedral-angle restraints	160
Number of unmapped restraints	38
Number of restraints per residue	4.0
Number of long range restraints per residue[1]	0.2

[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 (i)

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

8.2.1 Average number of distance violations per model (i)

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

Bins (Å)	Average number of violations per model	Max (Å)
0.1-0.2 (Small)	83.8	0.2
0.2-0.5 (Medium)	77.2	0.5
>0.5 (Large)	129.4	28.73

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)	17.3	4.9
10.0-20.0 (Medium)	None	None
>20.0 (Large)	None	None

9 Distance violation analysis (i)

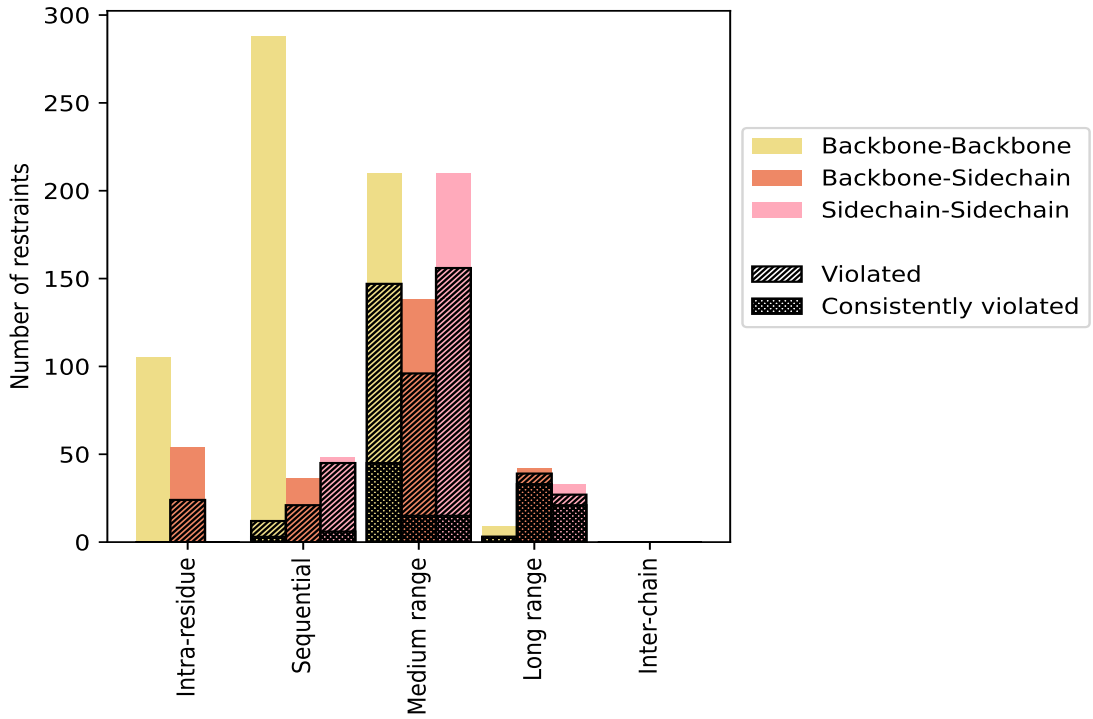
9.1 Summary of distance violations (i)

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)	159	13.6	24	15.1	2.0	0	0.0	0.0
Backbone-Backbone	105	9.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	54	4.6	24	44.4	2.0	0	0.0	0.0
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sequential (|i-j|=1)	372	31.7	78	21.0	6.6	9	2.4	0.8
Backbone-Backbone	288	24.6	12	4.2	1.0	3	1.0	0.3
Backbone-Sidechain	36	3.1	21	58.3	1.8	0	0.0	0.0
Sidechain-Sidechain	48	4.1	45	93.8	3.8	6	12.5	0.5
Medium range (|i-j|>1 & |i-j|<5)	480	40.9	339	70.6	28.9	39	8.1	3.3
Backbone-Backbone	132	11.3	87	65.9	7.4	9	6.8	0.8
Backbone-Sidechain	138	11.8	96	69.6	8.2	15	10.9	1.3
Sidechain-Sidechain	210	17.9	156	74.3	13.3	15	7.1	1.3
Long range (|i-j|≥5)	84	7.2	69	82.1	5.9	57	67.9	4.9
Backbone-Backbone	9	0.8	3	33.3	0.3	3	33.3	0.3
Backbone-Sidechain	42	3.6	39	92.9	3.3	33	78.6	2.8
Sidechain-Sidechain	33	2.8	27	81.8	2.3	21	63.6	1.8
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	78	6.6	60	76.9	5.1	36	46.2	3.1
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	1173	100.0	570	48.6	48.6	141	12.0	12.0
Backbone-Backbone	612	52.2	162	26.5	13.8	51	8.3	4.3
Backbone-Sidechain	270	23.0	180	66.7	15.3	48	17.8	4.1
Sidechain-Sidechain	291	24.8	228	78.4	19.4	42	14.4	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 (i)



Violated and consistently violated restraints are shown using different hatch patterns in their respective categories. The hydrogen bonds and disulfied bonds are counted in their appropriate category on the x-axis

9.2 Distance violation statistics for each model (i)

The following table provides the distance violation statistics for each model in the ensemble. Violations less than 0.1 Å are not included in the statistics.

Model ID	Number of violations						Mean (Å)	Max (Å)	SD[6] (Å)	Median (Å)
	IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total				
1	9	42	177	63	0	291	3.72	28.73	7.7	0.36
2	9	24	165	66	0	264	4.0	27.71	7.66	0.42
3	9	39	198	60	0	306	3.35	27.06	6.89	0.34
4	6	27	174	63	0	270	3.82	27.36	7.32	0.41
5	6	45	180	63	0	294	3.69	28.56	7.53	0.45
6	12	30	177	60	0	279	3.97	27.24	7.64	0.41
7	12	33	201	60	0	306	3.48	27.29	7.0	0.4
8	12	36	204	63	0	315	3.38	27.76	7.16	0.34
9	12	33	177	60	0	282	3.82	27.46	7.45	0.38
10	15	21	177	60	0	273	3.96	27.81	7.33	0.57
11	3	39	177	60	0	279	3.79	27.8	7.59	0.44

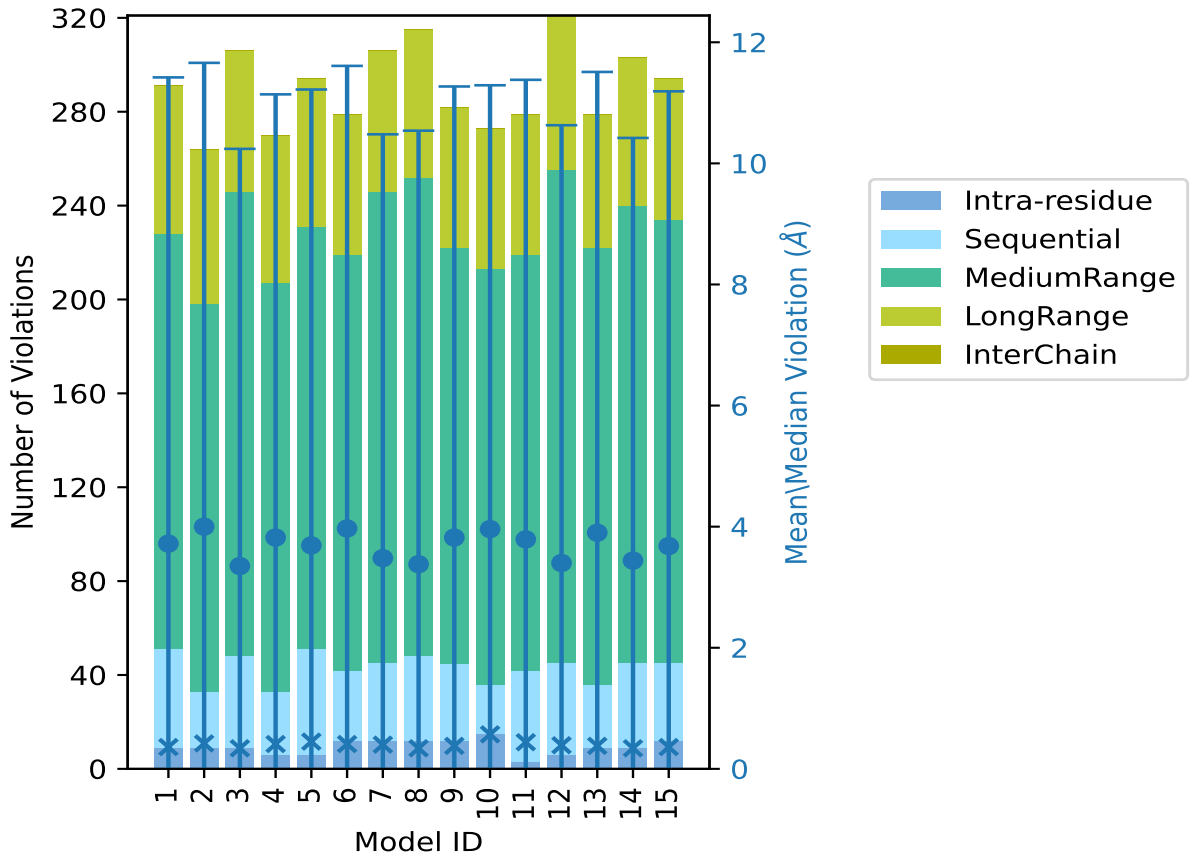
Continued on next page...

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				
12	6	39	210	66	0	321	3.4	28.31	7.23	0.39
13	9	27	186	57	0	279	3.9	27.62	7.61	0.38
14	9	36	195	63	0	303	3.44	27.37	6.98	0.34
15	12	33	189	60	0	294	3.68	28.16	7.51	0.36

[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 (i)



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)

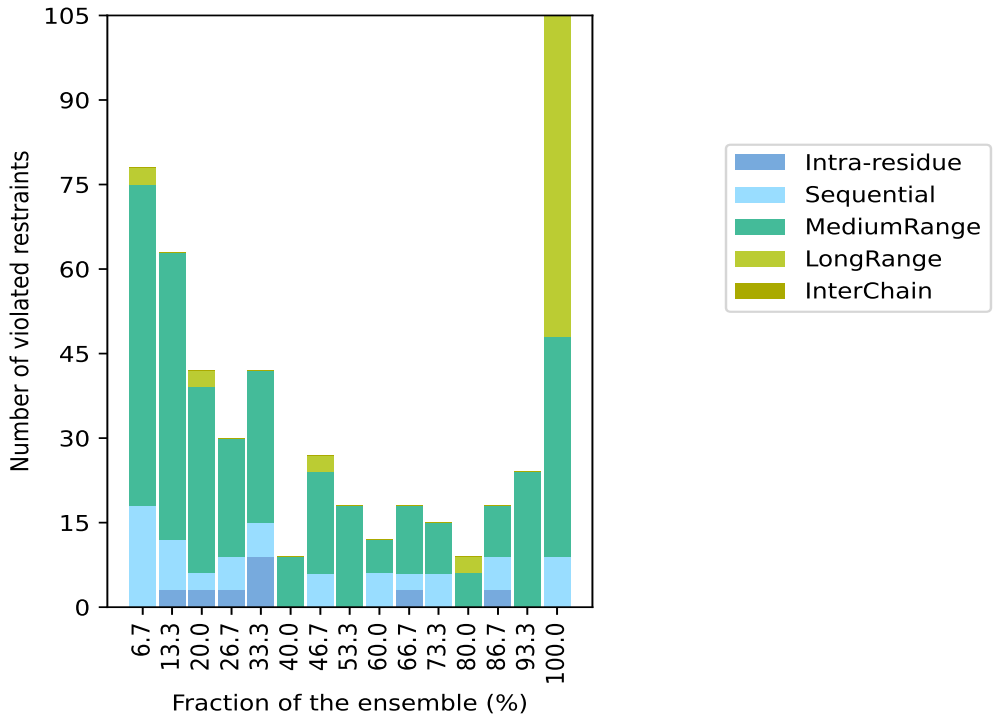
Violation analysis may find that some restraints are violated in few models and some are violated in most of models. The following table provides this information as number of violated restraints

for a given fraction of the ensemble. In total, 585(IR:135, SQ:294, MR:141, LR:15, 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]	%
0	18	57	3	0	78	1	6.7
3	9	51	0	0	63	2	13.3
3	3	33	3	0	42	3	20.0
3	6	21	0	0	30	4	26.7
9	6	27	0	0	42	5	33.3
0	0	9	0	0	9	6	40.0
0	6	18	3	0	27	7	46.7
0	0	18	0	0	18	8	53.3
0	6	6	0	0	12	9	60.0
3	3	12	0	0	18	10	66.7
0	6	9	0	0	15	11	73.3
0	0	6	3	0	9	12	80.0
3	6	9	0	0	18	13	86.7
0	0	24	0	0	24	14	93.3
0	9	39	57	0	105	15	100.0

[1]Intra-residue restraints, [2]Sequential restraints, [3]Medium range restraints, [4]Long range restraints, [5]Inter-chain restraints, [6] Number of models with violations

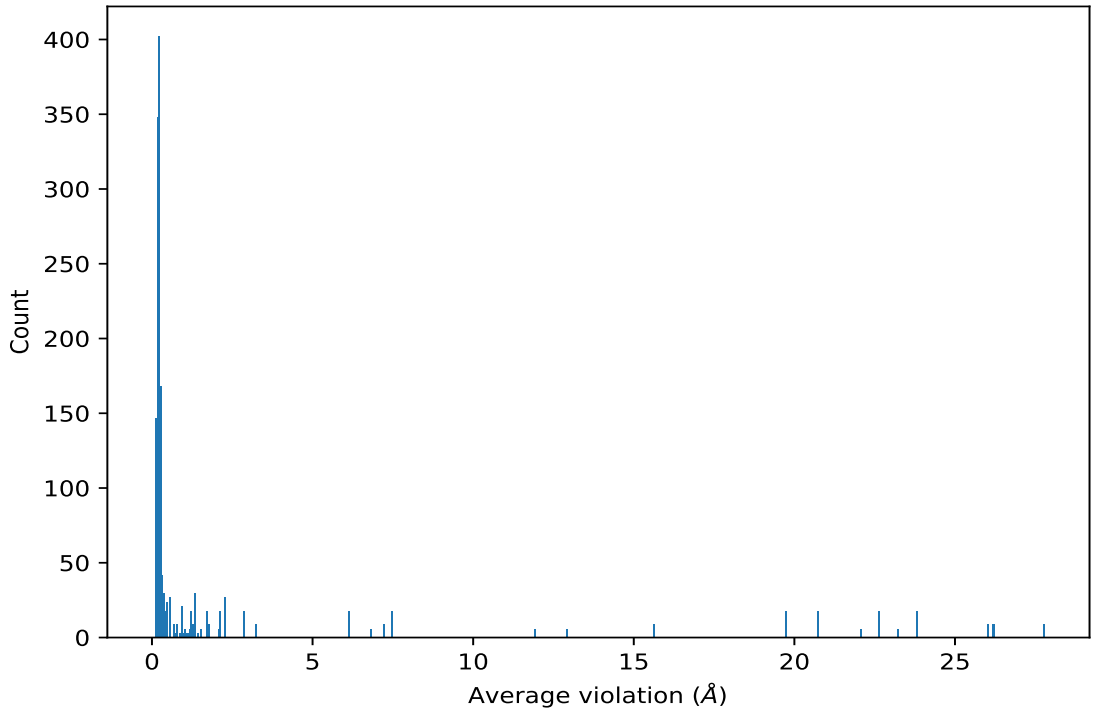
9.3.1 Bar graph : Distance violation statistics for the ensemble (i)



9.4 Most violated distance restraints in the ensemble (i)

9.4.1 Histogram : Distribution of mean distance violations (i)

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



9.4.2 Table: Most violated distance restraints (i)

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

Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(2,97)	1:C:781:ILE:H	1:C:746:ILE:HG21	15	27.75	0.48	27.71
(2,97)	1:C:781:ILE:H	1:C:746:ILE:HG22	15	27.75	0.48	27.71
(2,97)	1:C:781:ILE:H	1:C:746:ILE:HG23	15	27.75	0.48	27.71
(2,98)	1:C:781:ILE:H	1:C:746:ILE:HG21	15	27.75	0.48	27.71
(2,98)	1:C:781:ILE:H	1:C:746:ILE:HG22	15	27.75	0.48	27.71
(2,98)	1:C:781:ILE:H	1:C:746:ILE:HG23	15	27.75	0.48	27.71
(2,99)	1:C:781:ILE:H	1:C:746:ILE:HG21	15	27.75	0.48	27.71
(2,99)	1:C:781:ILE:H	1:C:746:ILE:HG22	15	27.75	0.48	27.71
(2,99)	1:C:781:ILE:H	1:C:746:ILE:HG23	15	27.75	0.48	27.71
(2,76)	1:C:749:VAL:HG11	1:C:782:VAL:H	15	26.22	0.66	26.06
(2,76)	1:C:749:VAL:HG12	1:C:782:VAL:H	15	26.22	0.66	26.06
(2,76)	1:C:749:VAL:HG13	1:C:782:VAL:H	15	26.22	0.66	26.06
(2,77)	1:C:749:VAL:HG11	1:C:782:VAL:H	15	26.22	0.66	26.06
(2,77)	1:C:749:VAL:HG12	1:C:782:VAL:H	15	26.22	0.66	26.06
(2,77)	1:C:749:VAL:HG13	1:C:782:VAL:H	15	26.22	0.66	26.06
(2,78)	1:C:749:VAL:HG11	1:C:782:VAL:H	15	26.22	0.66	26.06

Continued on next page...

Continued from previous page...

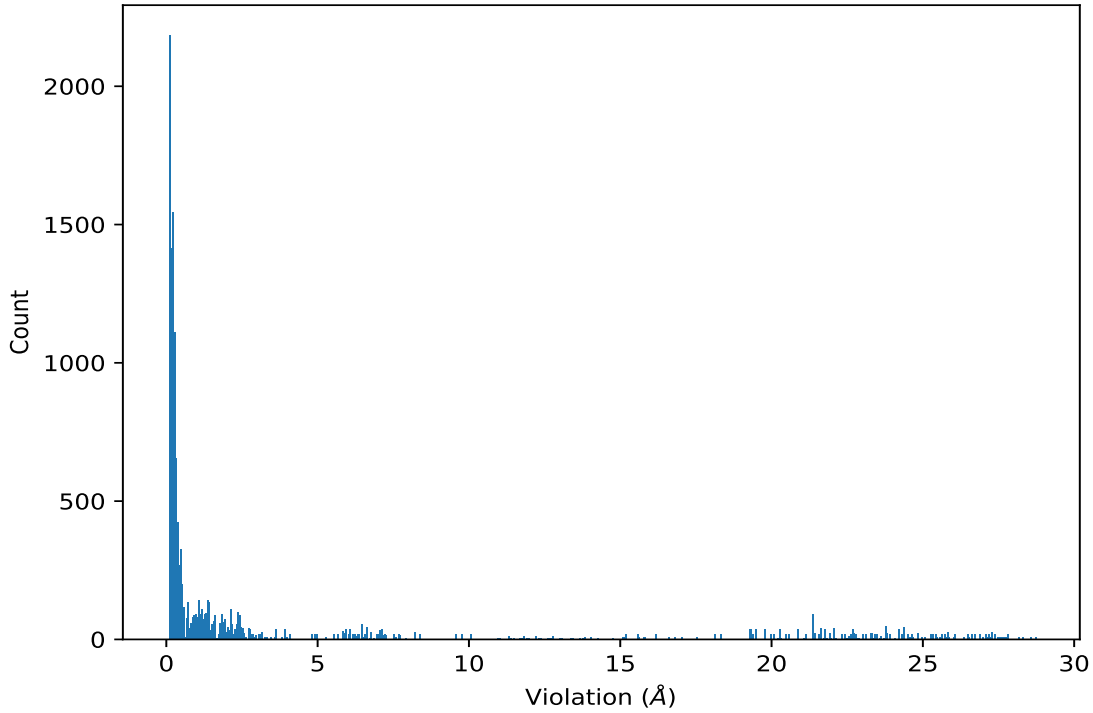
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(2,78)	1:C:749:VAL:HG12	1:C:782:VAL:H	15	26.22	0.66	26.06
(2,78)	1:C:749:VAL:HG13	1:C:782:VAL:H	15	26.22	0.66	26.06
(2,112)	1:C:785:LEU:HD21	1:C:749:VAL:HB	15	26.16	0.99	26.47
(2,112)	1:C:785:LEU:HD22	1:C:749:VAL:HB	15	26.16	0.99	26.47
(2,112)	1:C:785:LEU:HD23	1:C:749:VAL:HB	15	26.16	0.99	26.47
(2,113)	1:C:785:LEU:HD21	1:C:749:VAL:HB	15	26.16	0.99	26.47
(2,113)	1:C:785:LEU:HD22	1:C:749:VAL:HB	15	26.16	0.99	26.47
(2,113)	1:C:785:LEU:HD23	1:C:749:VAL:HB	15	26.16	0.99	26.47
(2,114)	1:C:785:LEU:HD21	1:C:749:VAL:HB	15	26.16	0.99	26.47
(2,114)	1:C:785:LEU:HD22	1:C:749:VAL:HB	15	26.16	0.99	26.47
(2,114)	1:C:785:LEU:HD23	1:C:749:VAL:HB	15	26.16	0.99	26.47
(2,100)	1:C:781:ILE:HD11	1:C:743:ASP:HA	15	26.0	0.85	25.81

[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)
(2,99)	1:C:781:ILE:H	1:C:746:ILE:HG21	1	28.73
(2,99)	1:C:781:ILE:H	1:C:746:ILE:HG22	1	28.73
(2,99)	1:C:781:ILE:H	1:C:746:ILE:HG23	1	28.73
(2,98)	1:C:781:ILE:H	1:C:746:ILE:HG21	1	28.73
(2,98)	1:C:781:ILE:H	1:C:746:ILE:HG22	1	28.73
(2,98)	1:C:781:ILE:H	1:C:746:ILE:HG23	1	28.73
(2,97)	1:C:781:ILE:H	1:C:746:ILE:HG21	1	28.73
(2,97)	1:C:781:ILE:H	1:C:746:ILE:HG22	1	28.73
(2,97)	1:C:781:ILE:H	1:C:746:ILE:HG23	1	28.73
(2,99)	1:C:781:ILE:H	1:C:746:ILE:HG21	5	28.56
(2,99)	1:C:781:ILE:H	1:C:746:ILE:HG22	5	28.56
(2,99)	1:C:781:ILE:H	1:C:746:ILE:HG23	5	28.56
(2,98)	1:C:781:ILE:H	1:C:746:ILE:HG21	5	28.56
(2,98)	1:C:781:ILE:H	1:C:746:ILE:HG22	5	28.56
(2,98)	1:C:781:ILE:H	1:C:746:ILE:HG23	5	28.56
(2,97)	1:C:781:ILE:H	1:C:746:ILE:HG21	5	28.56
(2,97)	1:C:781:ILE:H	1:C:746:ILE:HG22	5	28.56
(2,97)	1:C:781:ILE:H	1:C:746:ILE:HG23	5	28.56
(2,99)	1:C:781:ILE:H	1:C:746:ILE:HG21	12	28.31
(2,99)	1:C:781:ILE:H	1:C:746:ILE:HG22	12	28.31
(2,99)	1:C:781:ILE:H	1:C:746:ILE:HG23	12	28.31
(2,98)	1:C:781:ILE:H	1:C:746:ILE:HG21	12	28.31
(2,98)	1:C:781:ILE:H	1:C:746:ILE:HG22	12	28.31
(2,98)	1:C:781:ILE:H	1:C:746:ILE:HG23	12	28.31
(2,97)	1:C:781:ILE:H	1:C:746:ILE:HG21	12	28.31
(2,97)	1:C:781:ILE:H	1:C:746:ILE:HG22	12	28.31
(2,97)	1:C:781:ILE:H	1:C:746:ILE:HG23	12	28.31
(2,99)	1:C:781:ILE:H	1:C:746:ILE:HG21	15	28.16

10 Dihedral-angle violation analysis (i)

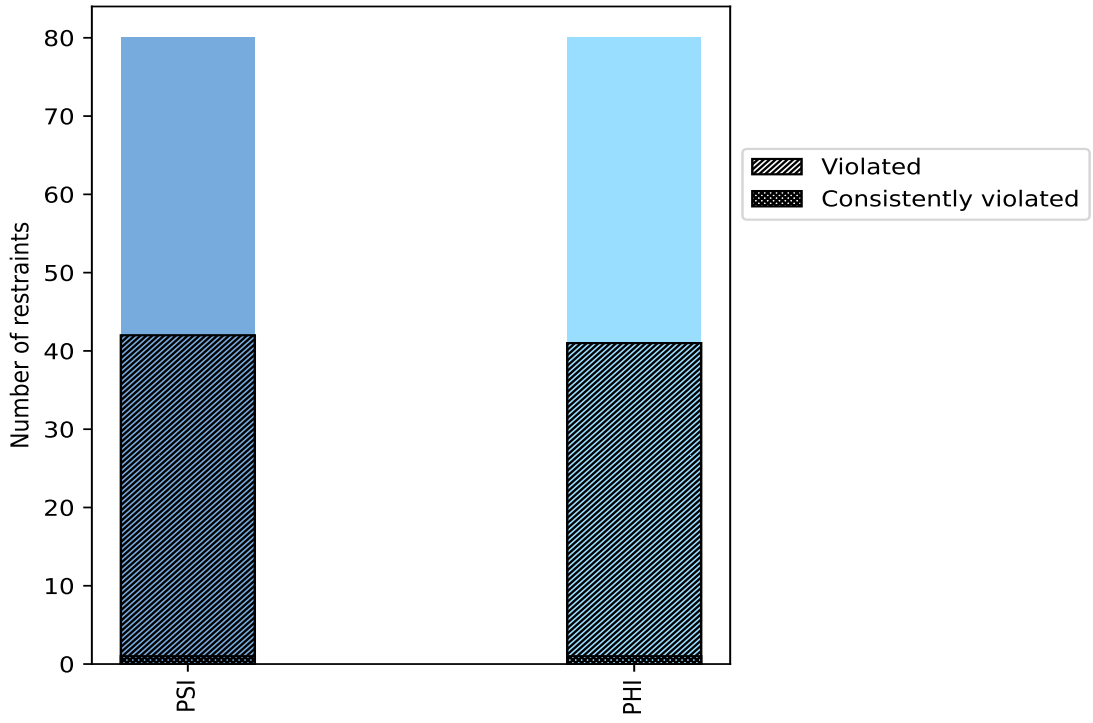
10.1 Summary of dihedral-angle violations (i)

The following table provides the summary of dihedral-angle violations in different dihedral-angle types. Violations less than 1° are not included in the calculation.

Angle type	Count	%[1]	Violated[3]			Consistently Violated[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
PSI	80	50.0	42	52.5	26.2	1	1.2	0.6
PHI	80	50.0	41	51.2	25.6	1	1.2	0.6
Total	160	100.0	83	51.9	51.9	2	1.2	1.2

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

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



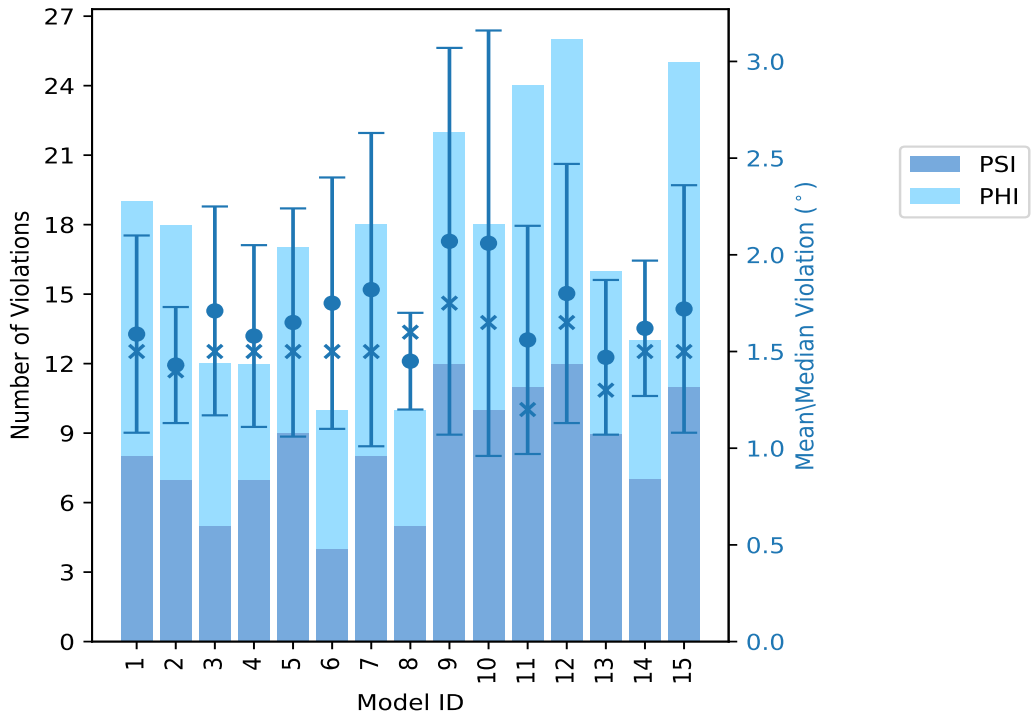
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

10.2 Dihedral-angle violation statistics for each model (i)

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

Model ID	Number of violations			Mean (°)	Max (°)	SD (°)	Median (°)
	PSI	PHI	Total				
1	8	11	19	1.59	2.9	0.51	1.5
2	7	11	18	1.43	2.1	0.3	1.4
3	5	7	12	1.71	2.8	0.54	1.5
4	7	5	12	1.58	2.7	0.47	1.5
5	9	8	17	1.65	3.2	0.59	1.5
6	4	6	10	1.75	3.0	0.65	1.5
7	8	10	18	1.82	3.8	0.81	1.5
8	5	5	10	1.45	1.8	0.25	1.6
9	12	10	22	2.07	4.9	1.0	1.75
10	10	8	18	2.06	4.8	1.1	1.65
11	11	13	24	1.56	3.0	0.59	1.2
12	12	14	26	1.8	3.4	0.67	1.65
13	9	7	16	1.47	2.6	0.4	1.3
14	7	6	13	1.62	2.3	0.35	1.5
15	11	14	25	1.72	3.4	0.64	1.5

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

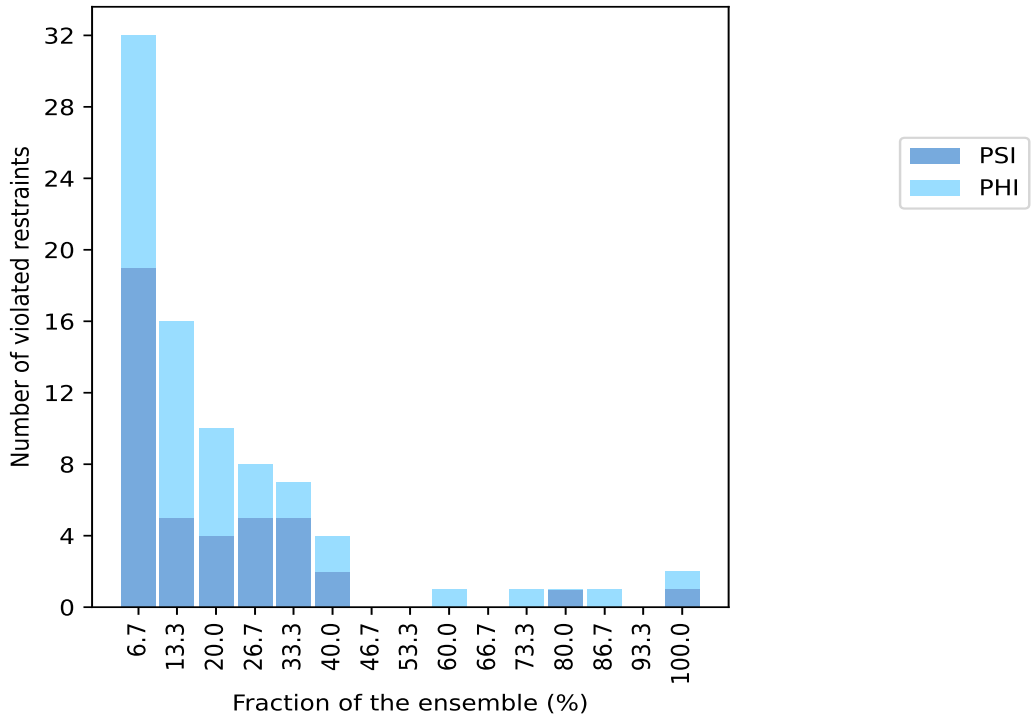
10.3 Dihedral-angle violation statistics for the ensemble (i)

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]	%
19	13	32	1	6.7
5	11	16	2	13.3
4	6	10	3	20.0
5	3	8	4	26.7
5	2	7	5	33.3
2	2	4	6	40.0
0	0	0	7	46.7
0	0	0	8	53.3
0	1	1	9	60.0
0	0	0	10	66.7
0	1	1	11	73.3
1	0	1	12	80.0
0	1	1	13	86.7
0	0	0	14	93.3
1	1	2	15	100.0

[1] Number of models with violations

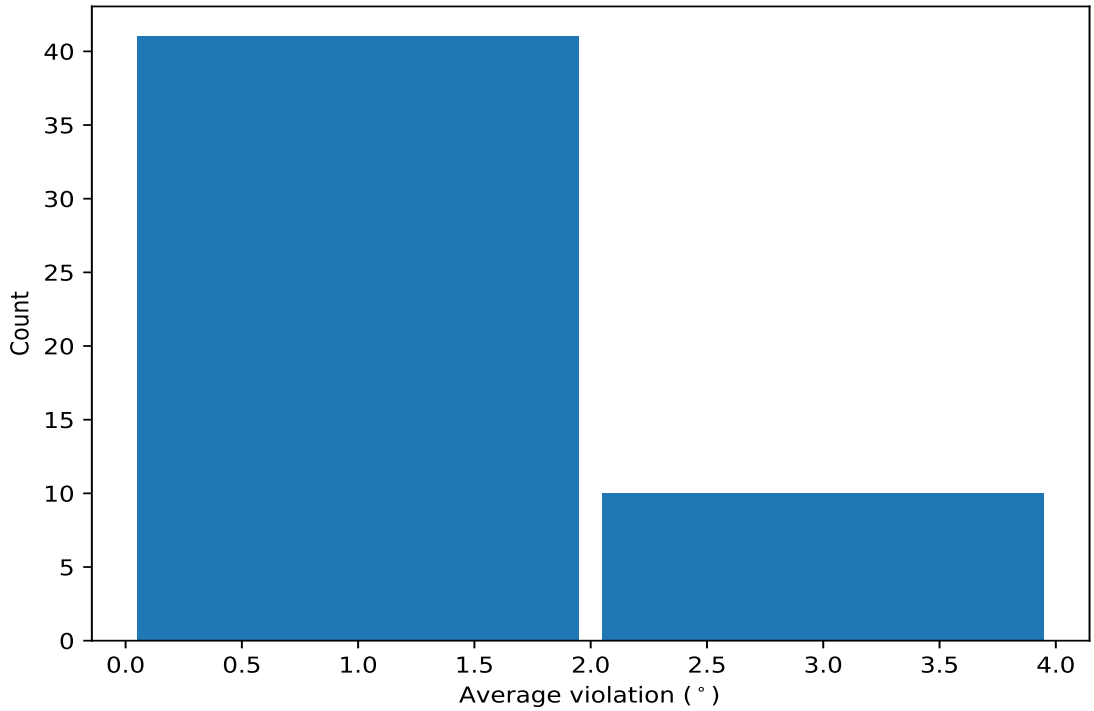
10.3.1 Bar graph : Dihedral-angle Violation statistics for the ensemble (i)



10.4 Most violated dihedral-angle restraints in the ensemble (i)

10.4.1 Histogram : Distribution of mean dihedral-angle violations (i)

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



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

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

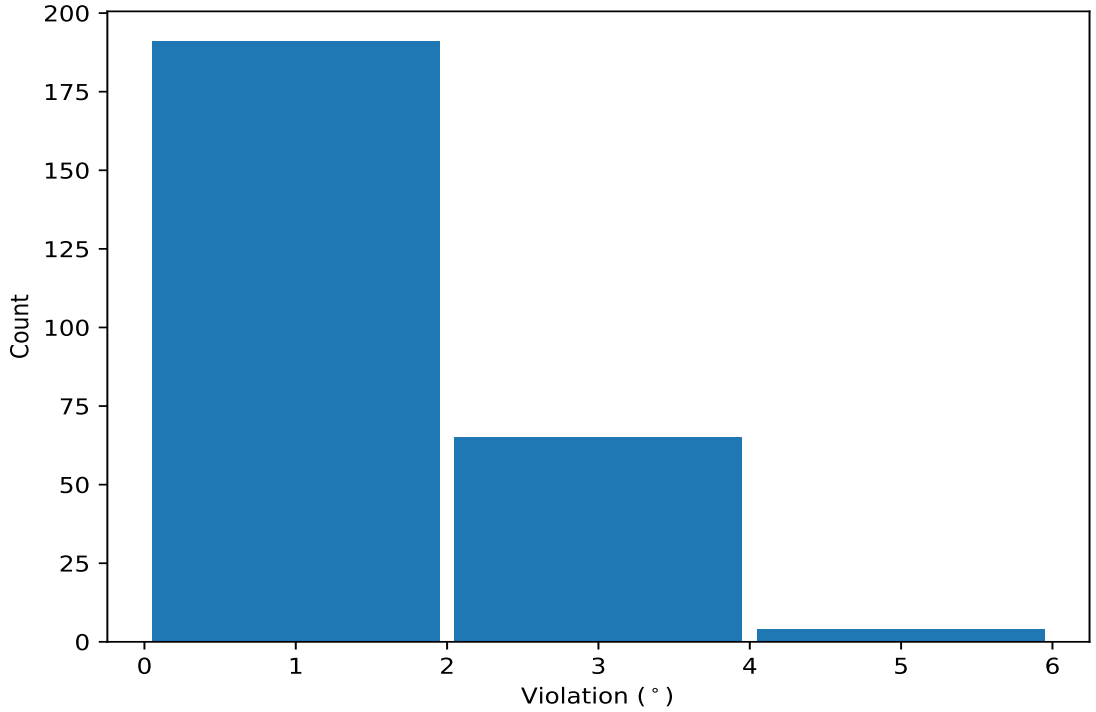
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models[1]	Mean	SD[2]	Median
(1,129)	1:C:770:ARG:C	1:C:771:LEU:N	1:C:771:LEU:CA	1:C:771:LEU:C	15	2.18	0.75	2.1
(1,128)	1:C:770:ARG:N	1:C:770:ARG:CA	1:C:770:ARG:C	1:C:771:LEU:N	15	1.95	0.81	1.9
(1,37)	1:C:696:ARG:C	1:C:697:ILE:N	1:C:697:ILE:CA	1:C:697:ILE:C	13	2.3	0.72	2.6
(1,36)	1:C:696:ARG:N	1:C:696:ARG:CA	1:C:696:ARG:C	1:C:697:ILE:N	12	2.55	0.64	2.8
(1,125)	1:C:763:LEU:C	1:C:764:SER:N	1:C:764:SER:CA	1:C:764:SER:C	11	1.45	0.44	1.3
(1,97)	1:C:749:VAL:C	1:C:750:ASN:N	1:C:750:ASN:CA	1:C:750:ASN:C	9	1.64	0.35	1.7
(1,64)	1:C:710:GLN:N	1:C:710:GLN:CA	1:C:710:GLN:C	1:C:711:GLY:N	6	2.35	1.24	1.85
(1,61)	1:C:708:VAL:C	1:C:709:ARG:N	1:C:709:ARG:CA	1:C:709:ARG:C	6	1.33	0.19	1.35
(1,106)	1:C:754:ALA:N	1:C:754:ALA:CA	1:C:754:ALA:C	1:C:755:LEU:N	6	1.3	0.15	1.35
(1,107)	1:C:754:ALA:C	1:C:755:LEU:N	1:C:755:LEU:CA	1:C:755:LEU:C	6	1.22	0.18	1.15

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

10.5 All violated dihedral-angle restraints (i)

10.5.1 Histogram : Distribution of violations (i)

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



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

The following table provides the list of violationsfor 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 (°)
(1,64)	1:C:710:GLN:N	1:C:710:GLN:CA	1:C:710:GLN:C	1:C:711:GLY:N	9	4.9
(1,129)	1:C:770:ARG:C	1:C:771:LEU:N	1:C:771:LEU:CA	1:C:771:LEU:C	10	4.8
(1,63)	1:C:709:ARG:C	1:C:710:GLN:N	1:C:710:GLN:CA	1:C:710:GLN:C	9	4.6
(1,128)	1:C:770:ARG:N	1:C:770:ARG:CA	1:C:770:ARG:C	1:C:771:LEU:N	10	4.6
(1,109)	1:C:755:LEU:C	1:C:756:ILE:N	1:C:756:ILE:CA	1:C:756:ILE:C	7	3.8
(1,108)	1:C:755:LEU:N	1:C:755:LEU:CA	1:C:755:LEU:C	1:C:756:ILE:N	7	3.7
(1,42)	1:C:699:PHE:N	1:C:699:PHE:CA	1:C:699:PHE:C	1:C:700:ALA:N	12	3.4
(1,36)	1:C:696:ARG:N	1:C:696:ARG:CA	1:C:696:ARG:C	1:C:697:ILE:N	15	3.4
(1,37)	1:C:696:ARG:C	1:C:697:ILE:N	1:C:697:ILE:CA	1:C:697:ILE:C	5	3.2
(1,37)	1:C:696:ARG:C	1:C:697:ILE:N	1:C:697:ILE:CA	1:C:697:ILE:C	15	3.2