



# wwPDB NMR Structure Validation Summary Report ⓘ

Dec 24, 2024 – 05:34 PM EST

PDB ID : 2LS8  
BMRB ID : 18415  
Title : Solution structure of human C-type lectin domain family 4 member D  
Authors : Harris, R.; Gaudette, J.; Bandaranayake, A.D.; Banu, R.; Bonanno, J.B.; Calarese, D.A.; Celikgil, A.; Chamala, S.; Chan, M.K.; Chaparro, R.; Evans, B.; Garforth, S.; Gizzi, A.; Hillerich, B.; Kar, A.; Lafleur, J.; Lim, S.; Love, J.; Matikainen, B.; Patel, H.; Seidel, R.D.; Smith, B.; Stead, M.; Girvin, M.E.; Almo, S.C.; New York Structural Genomics Research Consortium (NYSGRG)  
Deposited on : 2012-04-23

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

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

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

---

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

MolProbity : 4.02b-467  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
wwPDB-RCI : v\_1n\_11\_5\_13\_A (Berjanski et al., 2005)  
PANAV : Wang et al. (2010)  
wwPDB-ShiftChecker : v1.2  
BMRB Restraints Analysis : v1.2  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.40

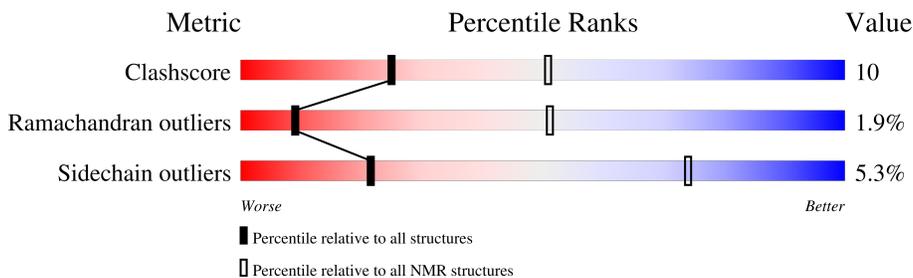
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 91%.

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and a grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	156	

## 2 Ensemble composition and analysis

This entry contains 20 models. Model 18 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:3-A:89, A:100-A:130 (118)	0.91	18
2	A:131-A:142 (12)	0.61	9

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. No single-model clusters were found.

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

### 3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 2273 atoms, of which 1096 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called C-type lectin domain family 4 member D.

Mol	Chain	Residues	Atoms					Trace	
			Total	C	H	N	O		S
1	A	142	2273	744	1096	207	217	9	0

There are 24 discrepancies between the modelled and reference sequences:

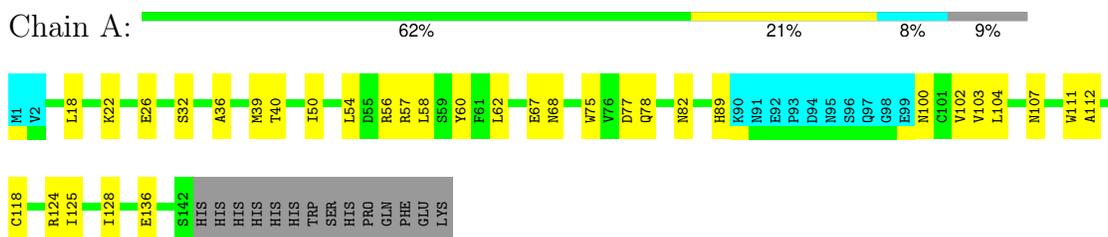
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	expression tag	UNP Q8WXI8
A	2	VAL	-	expression tag	UNP Q8WXI8
A	135	ALA	-	expression tag	UNP Q8WXI8
A	136	GLU	-	expression tag	UNP Q8WXI8
A	137	ASN	-	expression tag	UNP Q8WXI8
A	138	LEU	-	expression tag	UNP Q8WXI8
A	139	TYR	-	expression tag	UNP Q8WXI8
A	140	PHE	-	expression tag	UNP Q8WXI8
A	141	GLN	-	expression tag	UNP Q8WXI8
A	142	SER	-	expression tag	UNP Q8WXI8
A	143	HIS	-	expression tag	UNP Q8WXI8
A	144	HIS	-	expression tag	UNP Q8WXI8
A	145	HIS	-	expression tag	UNP Q8WXI8
A	146	HIS	-	expression tag	UNP Q8WXI8
A	147	HIS	-	expression tag	UNP Q8WXI8
A	148	HIS	-	expression tag	UNP Q8WXI8
A	149	TRP	-	expression tag	UNP Q8WXI8
A	150	SER	-	expression tag	UNP Q8WXI8
A	151	HIS	-	expression tag	UNP Q8WXI8
A	152	PRO	-	expression tag	UNP Q8WXI8
A	153	GLN	-	expression tag	UNP Q8WXI8
A	154	PHE	-	expression tag	UNP Q8WXI8
A	155	GLU	-	expression tag	UNP Q8WXI8
A	156	LYS	-	expression tag	UNP Q8WXI8

## 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.

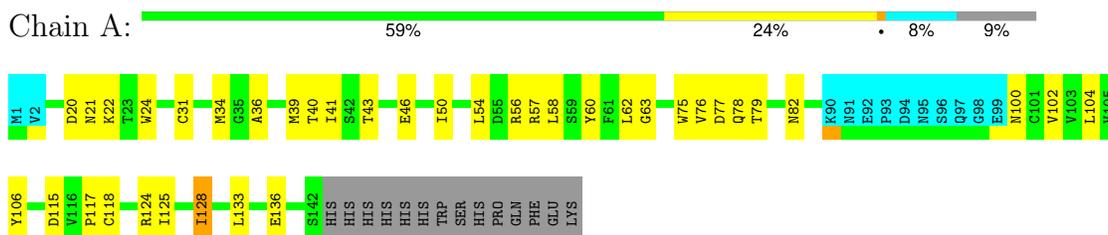
- Molecule 1: C-type lectin domain family 4 member D



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

The representative model is number 18. Colouring as in section 4.1 above.

- Molecule 1: C-type lectin domain family 4 member D



## 5 Refinement protocol and experimental data overview

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

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

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

Software name	Classification	Version
CNS	structure solution	1.21
CNS	refinement	1.21
ARIA	structure solution	2.3

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	1889
Number of shifts mapped to atoms	1766
Number of unparsed shifts	0
Number of shifts with mapping errors	123
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	91%

## 6 Model quality [i](#)

### 6.1 Standard geometry [i](#)

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 [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	1085	1012	1010	22±5
All	All	21700	20240	20200	431

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 238 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:22:LYS:HD3	1:A:26:GLU:HB3	0.85	1.48	3	2
1:A:67:GLU:HG2	1:A:70:LYS:HD2	0.79	1.51	9	1
1:A:39:MET:HB2	1:A:125:ILE:HB	0.77	1.54	1	8
1:A:45:ALA:HA	1:A:48:ASN:ND2	0.74	1.97	1	3
1:A:64:LEU:HB2	1:A:102:VAL:HG12	0.72	1.58	2	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	129/156 (83%)	114±3 (88±3%)	13±3 (10±2%)	2±1 (2±1%)	9	51
All	All	2580/3120 (83%)	2278 (88%)	253 (10%)	49 (2%)	9	51

5 of 19 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	11	GLN	9
1	A	76	VAL	4
1	A	69	ALA	4
1	A	5	ILE	3
1	A	57	ARG	3

### 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	117/142 (82%)	111±3 (95±3%)	6±3 (5±3%)	21	74
All	All	2340/2840 (82%)	2215 (95%)	125 (5%)	21	74

5 of 54 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	136	GLU	12
1	A	82	ASN	8
1	A	78	GLN	7
1	A	67	GLU	6
1	A	68	ASN	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 i

The completeness of assignment taking into account all chemical shift lists is 91% for the well-defined parts and 90% for the entire structure.

### 7.1 Chemical shift list 1

File name: working\_cs.cif

Chemical shift list name: *assigned\_chem\_shift\_list\_1*

#### 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	1889
Number of shifts mapped to atoms	1766
Number of unparsed shifts	0
Number of shifts with mapping errors	123
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	14

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 123) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	143	HIS	H	8.018	0.003	1
1	A	143	HIS	HA	4.496	0.001	1
1	A	143	HIS	HB2	3.033	0.001	2
1	A	143	HIS	HB3	2.734	0.007	2
1	A	143	HIS	HD2	6.99	0.000	1
1	A	143	HIS	C	174.397	0.000	1
1	A	143	HIS	CA	56.043	0.025	1
1	A	143	HIS	CB	28.949	0.084	1
1	A	143	HIS	N	118.375	0.030	1
1	A	148	HIS	HA	4.636	0.006	1
1	A	148	HIS	HB2	3.034	0.003	1
1	A	148	HIS	HB3	3.034	0.003	1
1	A	148	HIS	C	174.268	0.005	1
1	A	148	HIS	CA	55.709	0.041	1

*Continued on next page...*

*Continued from previous page...*

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	148	HIS	CB	29.358	0.006	1
1	A	149	TRP	H	8.141	0.001	1
1	A	149	TRP	HA	4.664	0.004	1
1	A	149	TRP	HB2	3.234	0.004	2
1	A	149	TRP	HB3	3.142	0.003	2
1	A	149	TRP	HD1	7.167	0.005	1
1	A	149	TRP	HE1	10.097	0.000	1
1	A	149	TRP	HE3	7.497	0.002	1
1	A	149	TRP	HZ2	7.32	0.006	1
1	A	149	TRP	HZ3	7.031	0.000	1
1	A	149	TRP	HH2	7.095	0.007	1
1	A	149	TRP	C	176.122	0.017	1
1	A	149	TRP	CA	57.129	0.024	1
1	A	149	TRP	CB	29.935	0.054	1
1	A	149	TRP	CD1	127.203	0.011	1
1	A	149	TRP	CE3	121.053	0.014	1
1	A	149	TRP	CZ2	114.65	0.001	1
1	A	149	TRP	CZ3	122.204	0.002	1
1	A	149	TRP	CH2	124.785	0.028	1
1	A	149	TRP	N	122.905	0.020	1
1	A	149	TRP	NE1	129.415	0.011	1
1	A	150	SER	H	8.231	0.001	1
1	A	150	SER	HA	4.331	0.004	1
1	A	150	SER	HB2	3.702	0.004	2
1	A	150	SER	HB3	3.662	0.000	2
1	A	150	SER	C	173.711	0.016	1
1	A	150	SER	CA	58.331	0.013	1
1	A	150	SER	CB	63.802	0.025	1
1	A	150	SER	N	117.58	0.056	1
1	A	151	HIS	H	8.167	0.002	1
1	A	151	HIS	HA	4.834	0.001	1
1	A	151	HIS	HB2	3.139	0.008	2
1	A	151	HIS	HB3	3.024	0.002	2
1	A	151	HIS	C	172.492	0.000	1
1	A	151	HIS	CA	53.656	0.006	1
1	A	151	HIS	CB	28.947	0.005	1
1	A	151	HIS	N	120.167	0.021	1
1	A	152	PRO	HA	4.313	0.003	1
1	A	152	PRO	HB2	1.765	0.002	2
1	A	152	PRO	HB3	2.2	0.003	2
1	A	152	PRO	HG2	1.945	0.003	1

*Continued on next page...*

*Continued from previous page...*

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	152	PRO	HG3	1.946	0.003	1
1	A	152	PRO	HD2	3.688	0.002	2
1	A	152	PRO	HD3	3.497	0.002	2
1	A	152	PRO	C	176.941	0.001	1
1	A	152	PRO	CA	63.605	0.005	1
1	A	152	PRO	CB	32.251	0.007	1
1	A	152	PRO	CG	27.521	0.008	1
1	A	152	PRO	CD	50.786	0.010	1
1	A	153	GLN	H	8.475	0.001	1
1	A	153	GLN	HA	4.189	0.002	1
1	A	153	GLN	HB2	1.884	0.008	2
1	A	153	GLN	HB3	1.855	0.010	2
1	A	153	GLN	HG2	2.168	0.001	2
1	A	153	GLN	HG3	2.1	0.002	2
1	A	153	GLN	HE21	6.797	0.000	1
1	A	153	GLN	HE22	7.41	0.002	1
1	A	153	GLN	C	175.723	0.010	1
1	A	153	GLN	CA	56.305	0.012	1
1	A	153	GLN	CB	29.431	0.012	1
1	A	153	GLN	CG	33.683	0.000	1
1	A	153	GLN	N	120.519	0.028	1
1	A	153	GLN	NE2	112.157	0.009	1
1	A	154	PHE	H	8.113	0.003	1
1	A	154	PHE	HA	4.677	0.004	1
1	A	154	PHE	HB2	3.192	0.004	2
1	A	154	PHE	HB3	2.998	0.008	2
1	A	154	PHE	HD1	7.225	0.004	3
1	A	154	PHE	HD2	7.225	0.004	3
1	A	154	PHE	HE1	7.293	0.002	3
1	A	154	PHE	HE2	7.293	0.002	3
1	A	154	PHE	HZ	7.237	0.000	1
1	A	154	PHE	C	175.595	0.027	1
1	A	154	PHE	CA	57.368	0.013	1
1	A	154	PHE	CB	39.663	0.013	1
1	A	154	PHE	CD1	131.909	0.086	3
1	A	154	PHE	CD2	131.909	0.086	3
1	A	154	PHE	CE1	131.732	0.048	3
1	A	154	PHE	CE2	131.732	0.048	3
1	A	154	PHE	CZ	129.993	0.000	1
1	A	154	PHE	N	119.769	0.023	1
1	A	155	GLU	H	8.213	0.003	1

*Continued on next page...*

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	155	GLU	HA	4.272	0.006	1
1	A	155	GLU	HB2	2.036	0.000	2
1	A	155	GLU	HB3	1.885	0.004	2
1	A	155	GLU	HG2	2.233	0.008	1
1	A	155	GLU	HG3	2.233	0.008	1
1	A	155	GLU	C	175.213	0.012	1
1	A	155	GLU	CA	56.535	0.007	1
1	A	155	GLU	CB	30.531	0.172	1
1	A	155	GLU	CG	34.256	0.003	1
1	A	155	GLU	N	122.522	0.019	1
1	A	156	LYS	H	7.881	0.002	1
1	A	156	LYS	HA	4.128	0.008	1
1	A	156	LYS	HB2	1.826	0.004	2
1	A	156	LYS	HB3	1.722	0.003	2
1	A	156	LYS	HG2	1.406	0.011	1
1	A	156	LYS	HG3	1.406	0.011	1
1	A	156	LYS	HD2	1.684	0.000	1
1	A	156	LYS	HD3	1.684	0.000	1
1	A	156	LYS	HE2	2.997	0.001	1
1	A	156	LYS	HE3	2.997	0.001	1
1	A	156	LYS	C	181.221	0.000	1
1	A	156	LYS	CA	57.796	0.014	1
1	A	156	LYS	CB	33.884	0.015	1
1	A	156	LYS	CG	24.833	0.000	1
1	A	156	LYS	CD	29.245	0.000	1
1	A	156	LYS	CE	42.072	0.000	1
1	A	156	LYS	N	126.928	0.025	1

### 7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	147	$-0.01 \pm 0.17$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}_\beta$	141	$-0.09 \pm 0.17$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}'$	142	$0.07 \pm 0.18$	None needed ( $< 0.5$ ppm)
$^{15}\text{N}$	131	$0.36 \pm 0.35$	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 91%, i.e. 1655 atoms were assigned a chemical shift out of a possible 1822. 0 out of 15 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Backbone	617/643 (96%)	250/259 (97%)	251/260 (97%)	116/124 (94%)
Sidechain	828/953 (87%)	562/612 (92%)	240/291 (82%)	26/50 (52%)
Aromatic	210/226 (93%)	105/111 (95%)	98/104 (94%)	7/11 (64%)
Overall	1655/1822 (91%)	917/982 (93%)	589/655 (90%)	149/185 (81%)

### 7.1.4 Statistically unusual chemical shifts [i](#)

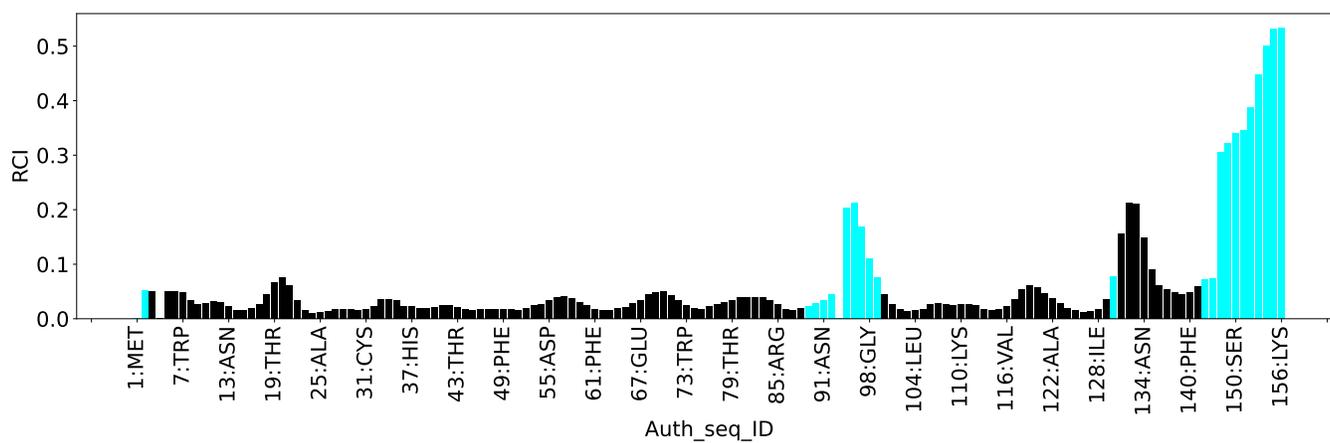
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	66	ASP	HB3	0.57	1.32 – 4.00	-7.8
1	A	106	TYR	HB2	0.30	1.09 – 4.72	-7.2
1	A	16	PHE	HB3	0.49	1.03 – 4.85	-6.4
1	A	113	TRP	HE3	4.97	5.27 – 9.37	-5.8
1	A	87	PHE	HB3	0.75	1.03 – 4.85	-5.7
1	A	47	GLN	HE21	4.84	5.02 – 9.43	-5.4
1	A	65	ARG	HD2	1.89	1.97 – 4.26	-5.3
1	A	11	GLN	HG3	0.84	0.91 – 3.68	-5.2
1	A	73	TRP	HB3	1.27	1.31 – 4.93	-5.1
1	A	64	LEU	HD21	-0.68	-0.65 – 2.13	-5.1
1	A	64	LEU	HD22	-0.68	-0.65 – 2.13	-5.1
1	A	64	LEU	HD23	-0.68	-0.65 – 2.13	-5.1
1	A	4	PRO	HG2	0.39	0.41 – 3.45	-5.0
1	A	65	ARG	HD3	1.80	1.81 – 4.39	-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	2332
Intra-residue ( $ i-j =0$ )	668
Sequential ( $ i-j =1$ )	559
Medium range ( $ i-j >1$ and $ i-j <5$ )	314
Long range ( $ i-j \geq 5$ )	716
Inter-chain	0
Hydrogen bond restraints	72
Disulfide bond restraints	3
Total dihedral-angle restraints	167
Number of unmapped restraints	1
Number of restraints per residue	16.0
Number of long range restraints per residue <sup>1</sup>	4.9

<sup>1</sup>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)	12.3	0.2
0.2-0.5 (Medium)	8.1	0.5
>0.5 (Large)	44.9	4.18

### 8.2.2 Average number of dihedral-angle violations per model [i](#)

Dihedral-angle violations less than 1° are not included in the calculation.

Bins (°)	Average number of violations per model	Max (°)
1.0-10.0 (Small)	10.9	7.97
10.0-20.0 (Medium)	None	None
>20.0 (Large)	None	None

## 9 Distance violation analysis [i](#)

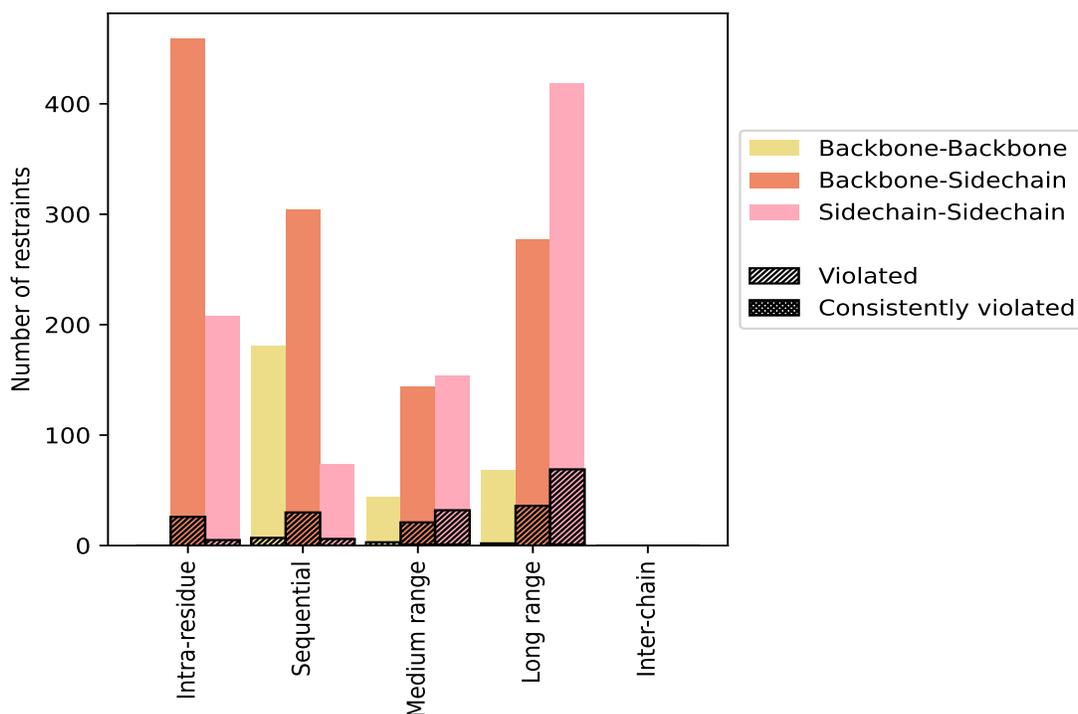
### 9.1 Summary of distance violations [i](#)

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.

Restrains type	Count	% <sup>1</sup>	Violated <sup>3</sup>			Consistently Violated <sup>4</sup>		
			Count	% <sup>2</sup>	% <sup>1</sup>	Count	% <sup>2</sup>	% <sup>1</sup>
<b>Intra-residue (<math> i-j =0</math>)</b>	<b>668</b>	<b>28.6</b>	<b>31</b>	<b>4.6</b>	<b>1.3</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
Backbone-Backbone	1	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	459	19.7	26	5.7	1.1	0	0.0	0.0
Sidechain-Sidechain	208	8.9	5	2.4	0.2	0	0.0	0.0
<b>Sequential (<math> i-j =1</math>)</b>	<b>559</b>	<b>24.0</b>	<b>43</b>	<b>7.7</b>	<b>1.8</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
Backbone-Backbone	181	7.8	7	3.9	0.3	0	0.0	0.0
Backbone-Sidechain	304	13.0	30	9.9	1.3	0	0.0	0.0
Sidechain-Sidechain	74	3.2	6	8.1	0.3	0	0.0	0.0
<b>Medium range (<math> i-j &gt;1</math> &amp; <math> i-j &lt;5</math>)</b>	<b>314</b>	<b>13.5</b>	<b>51</b>	<b>16.2</b>	<b>2.2</b>	<b>2</b>	<b>0.6</b>	<b>0.1</b>
Backbone-Backbone	44	1.9	3	6.8	0.1	0	0.0	0.0
Backbone-Sidechain	116	5.0	16	13.8	0.7	1	0.9	0.0
Sidechain-Sidechain	154	6.6	32	20.8	1.4	1	0.6	0.0
<b>Long range (<math> i-j \geq 5</math>)</b>	<b>716</b>	<b>30.7</b>	<b>106</b>	<b>14.8</b>	<b>4.5</b>	<b>1</b>	<b>0.1</b>	<b>0.0</b>
Backbone-Backbone	68	2.9	2	2.9	0.1	0	0.0	0.0
Backbone-Sidechain	233	10.0	35	15.0	1.5	0	0.0	0.0
Sidechain-Sidechain	415	17.8	69	16.6	3.0	1	0.2	0.0
<b>Inter-chain</b>	<b>0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
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
<b>Hydrogen bond</b>	<b>72</b>	<b>3.1</b>	<b>6</b>	<b>8.3</b>	<b>0.3</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
<b>Disulfide bond</b>	<b>3</b>	<b>0.1</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
<b>Total</b>	<b>2332</b>	<b>100.0</b>	<b>237</b>	<b>10.2</b>	<b>10.2</b>	<b>3</b>	<b>0.1</b>	<b>0.1</b>
Backbone-Backbone	294	12.6	12	4.1	0.5	0	0.0	0.0
Backbone-Sidechain	1184	50.8	113	9.5	4.8	1	0.1	0.0
Sidechain-Sidechain	854	36.6	112	13.1	4.8	2	0.2	0.1

<sup>1</sup> percentage calculated with respect to the total number of distance restraints, <sup>2</sup> percentage calculated with respect to the number of restraints in a particular restraint category, <sup>3</sup> violated in at least one model, <sup>4</sup> violated in all the models

### 9.1.1 Bar chart : Distribution of distance restraints and violations [i](#)



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

## 9.2 Distance violation statistics for each model [i](#)

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

Model ID	Number of violations						Mean (Å)	Max (Å)	SD <sup>6</sup> (Å)	Median (Å)
	IR <sup>1</sup>	SQ <sup>2</sup>	MR <sup>3</sup>	LR <sup>4</sup>	IC <sup>5</sup>	Total				
1	4	6	22	39	0	71	1.23	4.16	1.03	0.96
2	5	10	21	49	0	85	1.37	4.12	1.01	1.15
3	6	10	21	50	0	87	1.34	3.92	1.07	1.23
4	6	5	16	31	0	58	1.01	4.07	1.03	0.48
5	5	5	18	33	0	61	1.5	4.04	1.15	1.33
6	5	3	19	41	0	68	1.23	4.18	0.96	1.11
7	5	5	19	34	0	63	1.29	3.93	1.06	1.09
8	7	3	20	39	0	69	1.25	3.92	1.12	0.85
9	6	5	18	21	0	50	1.11	4.13	0.9	0.96
10	5	6	18	46	0	75	1.42	4.14	1.08	1.33

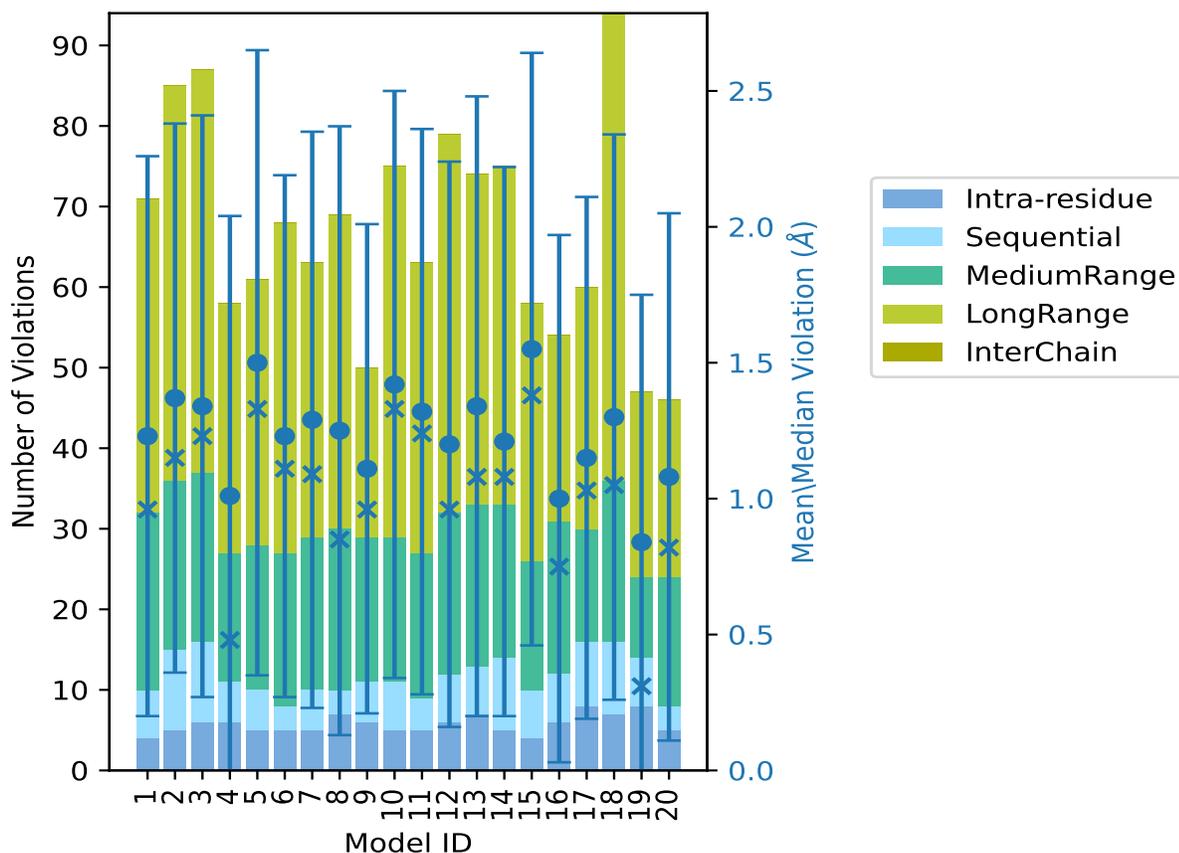
*Continued on next page...*

Continued from previous page...

Model ID	Number of violations					Total	Mean (Å)	Max (Å)	SD <sup>6</sup> (Å)	Median (Å)
	IR <sup>1</sup>	SQ <sup>2</sup>	MR <sup>3</sup>	LR <sup>4</sup>	IC <sup>5</sup>					
11	5	4	18	36	0	63	1.32	3.87	1.04	1.24
12	6	6	20	47	0	79	1.2	4.04	1.04	0.96
13	7	6	20	41	0	74	1.34	3.9	1.14	1.08
14	5	9	19	42	0	75	1.21	4.03	1.01	1.08
15	4	6	16	32	0	58	1.55	4.16	1.09	1.38
16	6	6	19	23	0	54	1.0	4.04	0.97	0.75
17	8	8	14	30	0	60	1.15	3.6	0.96	1.03
18	7	9	20	58	0	94	1.3	4.1	1.04	1.05
19	8	6	10	23	0	47	0.84	3.77	0.91	0.31
20	5	3	16	22	0	46	1.08	3.9	0.97	0.82

<sup>1</sup>Intra-residue restraints, <sup>2</sup>Sequential restraints, <sup>3</sup>Medium range restraints, <sup>4</sup>Long range restraints, <sup>5</sup>Inter-chain restraints, <sup>6</sup>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

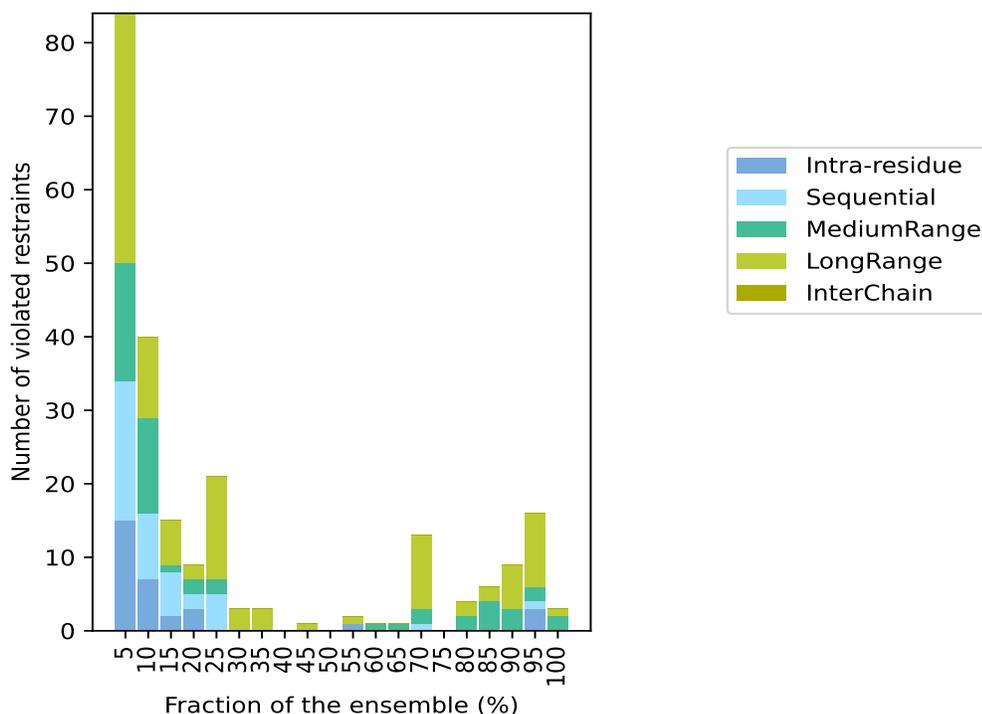
### 9.3 Distance violation statistics for the ensemble

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, 2026(IR:637, SQ:516, MR:263, LR:610, IC:0) restraints are not violated in the ensemble.

Number of violated restraints						Fraction of the ensemble	
IR <sup>1</sup>	SQ <sup>2</sup>	MR <sup>3</sup>	LR <sup>4</sup>	IC <sup>5</sup>	Total	Count <sup>6</sup>	%
15	19	16	34	0	84	1	5.0
7	9	13	11	0	40	2	10.0
2	6	1	6	0	15	3	15.0
3	2	2	2	0	9	4	20.0
0	5	2	14	0	21	5	25.0
0	0	0	3	0	3	6	30.0
0	0	0	3	0	3	7	35.0
0	0	0	0	0	0	8	40.0
0	0	0	1	0	1	9	45.0
0	0	0	0	0	0	10	50.0
1	0	0	1	0	2	11	55.0
0	0	1	0	0	1	12	60.0
0	0	1	0	0	1	13	65.0
0	1	2	10	0	13	14	70.0
0	0	0	0	0	0	15	75.0
0	0	2	2	0	4	16	80.0
0	0	4	2	0	6	17	85.0
0	0	3	6	0	9	18	90.0
3	1	2	10	0	16	19	95.0
0	0	2	1	0	3	20	100.0

<sup>1</sup>Intra-residue restraints, <sup>2</sup>Sequential restraints, <sup>3</sup>Medium range restraints, <sup>4</sup>Long range restraints, <sup>5</sup>Inter-chain restraints, <sup>6</sup> Number of models with violations

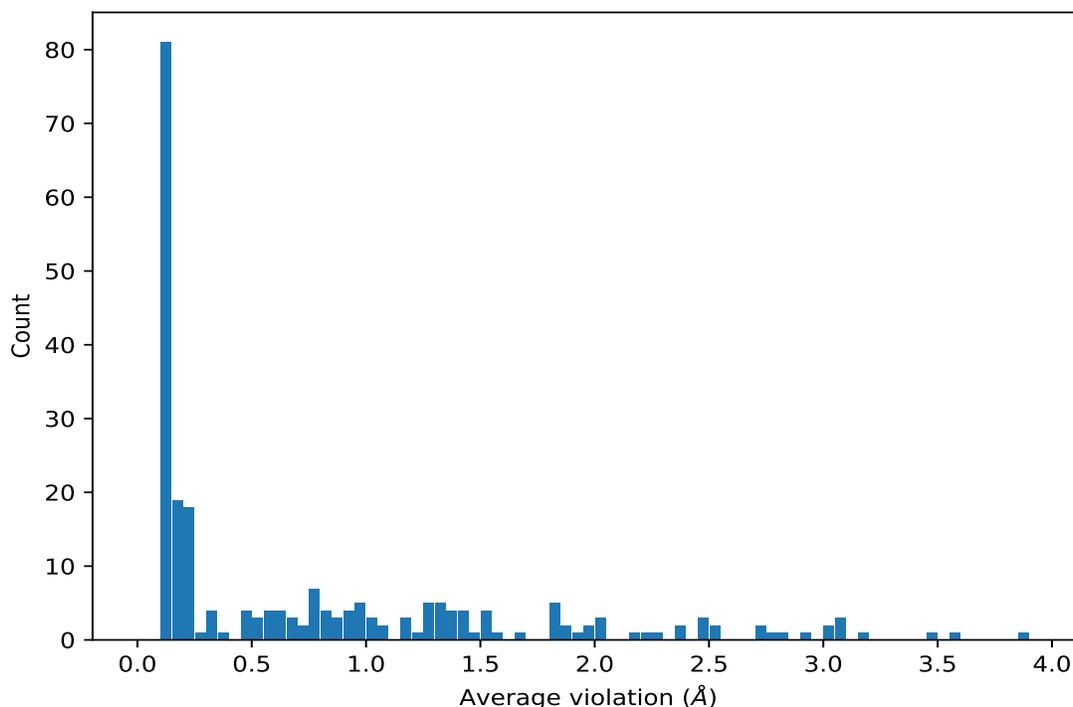
### 9.3.1 Bar graph : Distance violation statistics for the ensemble [i](#)



## 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.

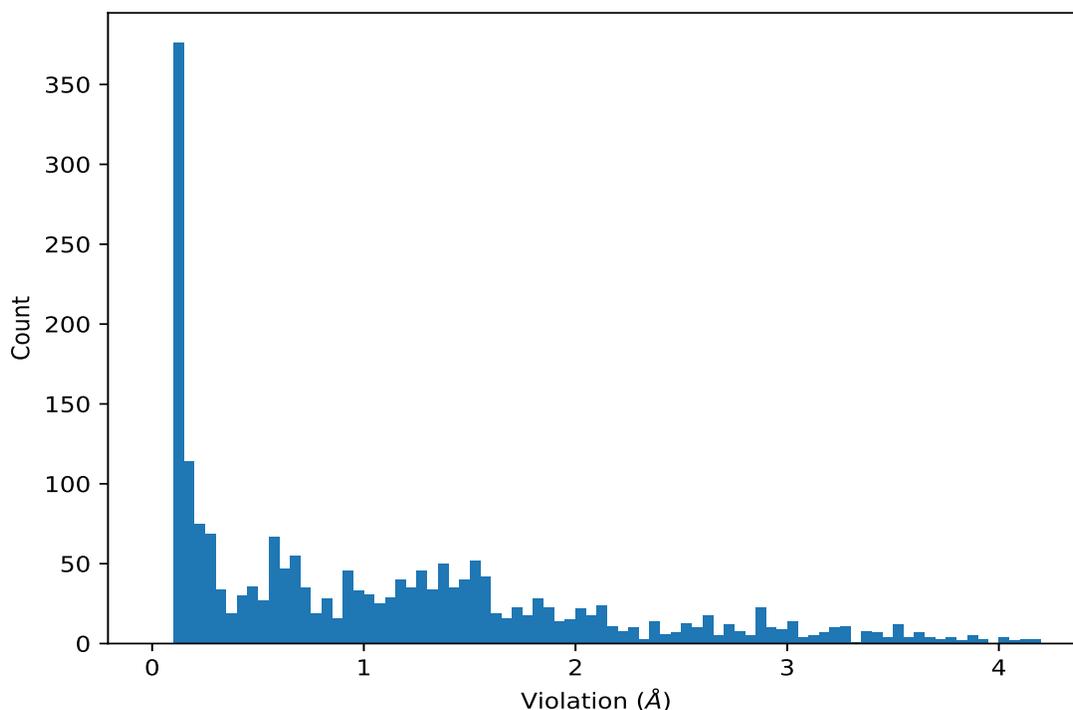
Key	Atom-1	Atom-2	Models <sup>1</sup>	Mean (Å)	SD <sup>1</sup> (Å)	Median (Å)
(1,1635)	1:10:A:PHE:HD2	1:49:A:PHE:HB3	20	1.85	0.61	1.84
(1,1532)	1:120:A:PHE:HE1	1:122:A:ALA:HA	20	1.07	0.32	1.04
(1,1527)	1:117:A:PRO:HB2	1:120:A:PHE:HD2	20	0.8	0.3	0.74
(1,1995)	1:59:A:SER:HB3	1:120:A:PHE:HE1	19	3.87	0.67	4.04
(1,2011)	1:60:A:TYR:HD1	1:124:A:ARG:HB2	19	3.58	0.26	3.64
(1,2017)	1:60:A:TYR:HE2	1:111:A:TRP:HZ2	19	3.17	0.29	3.21
(1,2016)	1:60:A:TYR:HE2	1:105:A:VAL:HA	19	2.7	0.32	2.71
(1,906)	1:59:A:SER:HA	1:60:A:TYR:HD2	19	2.0	0.06	1.99
(1,1694)	1:17:A:PRO:HB2	1:60:A:TYR:HE1	19	1.93	0.38	1.87
(1,1408)	1:58:A:LEU:HA	1:60:A:TYR:HE2	19	1.83	0.16	1.85

<sup>1</sup>Number of violated models, <sup>2</sup>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)
(1,1995)	1:59:A:SER:HB3	1:120:A:PHE:HE1	6	4.18
(1,1995)	1:59:A:SER:HB3	1:120:A:PHE:HE1	1	4.16
(1,1995)	1:59:A:SER:HB3	1:120:A:PHE:HE1	15	4.16
(1,1995)	1:59:A:SER:HB3	1:120:A:PHE:HE1	10	4.14
(1,1995)	1:59:A:SER:HB3	1:120:A:PHE:HE1	9	4.13
(1,1995)	1:59:A:SER:HB3	1:120:A:PHE:HE1	2	4.12
(1,1995)	1:59:A:SER:HB3	1:120:A:PHE:HE1	18	4.1
(1,1995)	1:59:A:SER:HB3	1:120:A:PHE:HE1	4	4.07
(1,1995)	1:59:A:SER:HB3	1:120:A:PHE:HE1	5	4.04
(1,1995)	1:59:A:SER:HB3	1:120:A:PHE:HE1	12	4.04

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

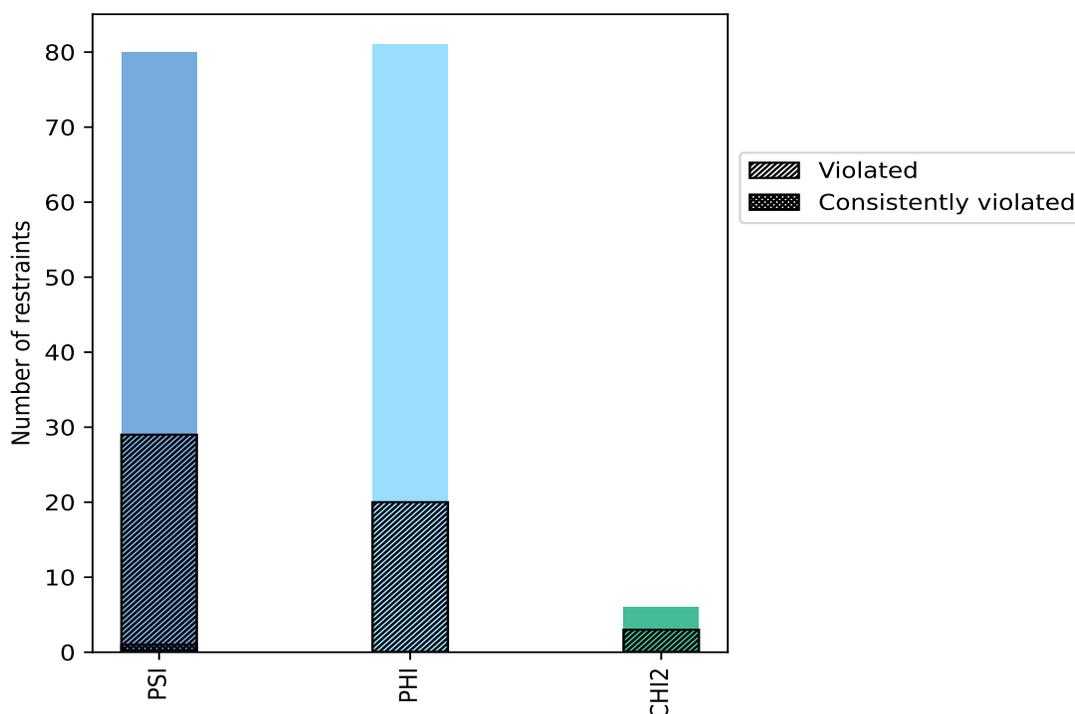
### 10.1 Summary of dihedral-angle violations [i](#)

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	% <sup>1</sup>	Violated <sup>3</sup>			Consistently Violated <sup>4</sup>		
			Count	% <sup>2</sup>	% <sup>1</sup>	Count	% <sup>2</sup>	% <sup>1</sup>
PSI	80	47.9	29	36.2	17.4	1	1.2	0.6
PHI	81	48.5	20	24.7	12.0	0	0.0	0.0
CHI2	6	3.6	3	50.0	1.8	0	0.0	0.0
Total	167	100.0	52	31.1	31.1	1	0.6	0.6

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

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



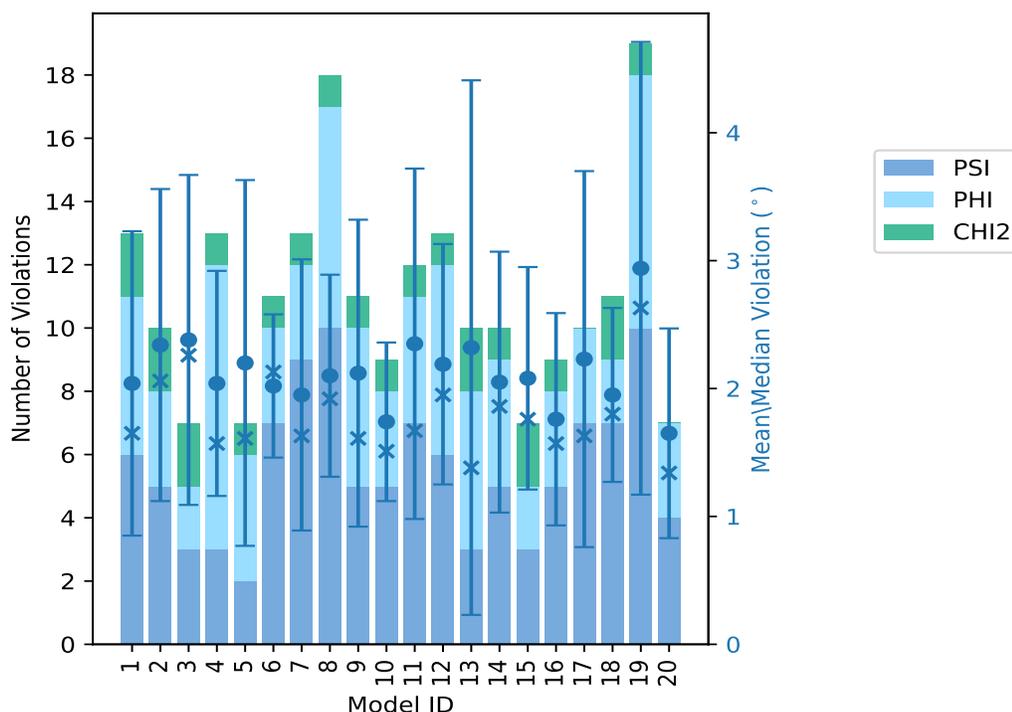
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

## 10.2 Dihedral-angle violation statistics for each model

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

Model ID	Number of violations				Mean (°)	Max (°)	SD (°)	Median (°)
	PSI	PHI	CHI2	Total				
1	6	5	2	13	2.04	5.59	1.19	1.65
2	5	3	2	10	2.34	5.76	1.22	2.06
3	3	2	2	7	2.38	5.0	1.29	2.26
4	3	9	1	13	2.04	4.42	0.88	1.57
5	2	4	1	7	2.2	5.55	1.43	1.61
6	7	3	1	11	2.02	2.78	0.56	2.13
7	9	3	1	13	1.95	4.23	1.06	1.63
8	10	7	1	18	2.1	3.81	0.79	1.92
9	5	5	1	11	2.12	5.1	1.2	1.61
10	5	3	1	9	1.74	2.94	0.62	1.51
11	7	4	1	12	2.35	5.12	1.37	1.67
12	6	6	1	13	2.19	4.5	0.94	1.95
13	3	5	2	10	2.32	7.97	2.09	1.38
14	5	4	1	10	2.05	4.91	1.02	1.86
15	3	2	2	7	2.08	3.61	0.87	1.76
16	5	3	1	9	1.76	3.9	0.83	1.57
17	7	3	0	10	2.23	6.33	1.47	1.63
18	7	2	2	11	1.95	3.36	0.68	1.8
19	10	8	1	19	2.94	7.02	1.77	2.63
20	4	3	0	7	1.65	3.54	0.82	1.34

### 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	CHI2	Total	Count <sup>1</sup>	%
11	6	1	18	1	5.0
4	4	0	8	2	10.0
2	1	0	3	3	15.0
3	1	0	4	4	20.0
1	1	0	2	5	25.0
2	1	1	4	6	30.0
3	3	0	6	7	35.0
1	1	0	2	8	40.0
1	0	0	1	9	45.0
0	0	0	0	10	50.0
0	1	0	1	11	55.0

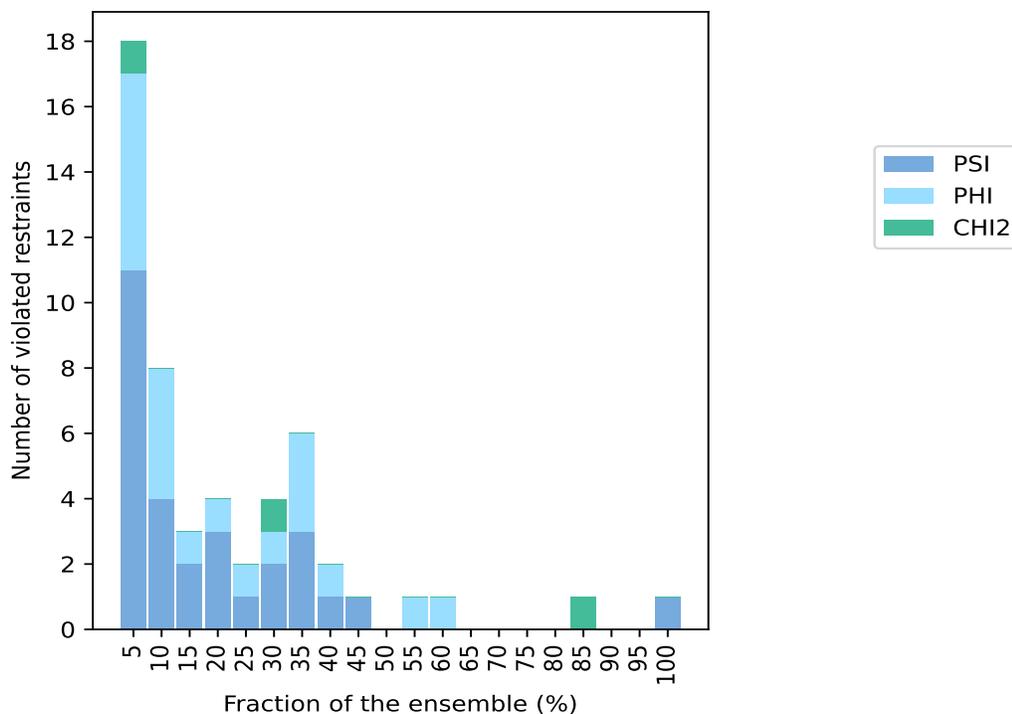
*Continued on next page...*

Continued from previous page...

Number of violated restraints				Fraction of the ensemble	
PSI	PHI	CHI2	Total	Count <sup>1</sup>	%
0	1	0	1	12	60.0
0	0	0	0	13	65.0
0	0	0	0	14	70.0
0	0	0	0	15	75.0
0	0	0	0	16	80.0
0	0	1	1	17	85.0
0	0	0	0	18	90.0
0	0	0	0	19	95.0
1	0	0	1	20	100.0

<sup>1</sup> Number of models with violations

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

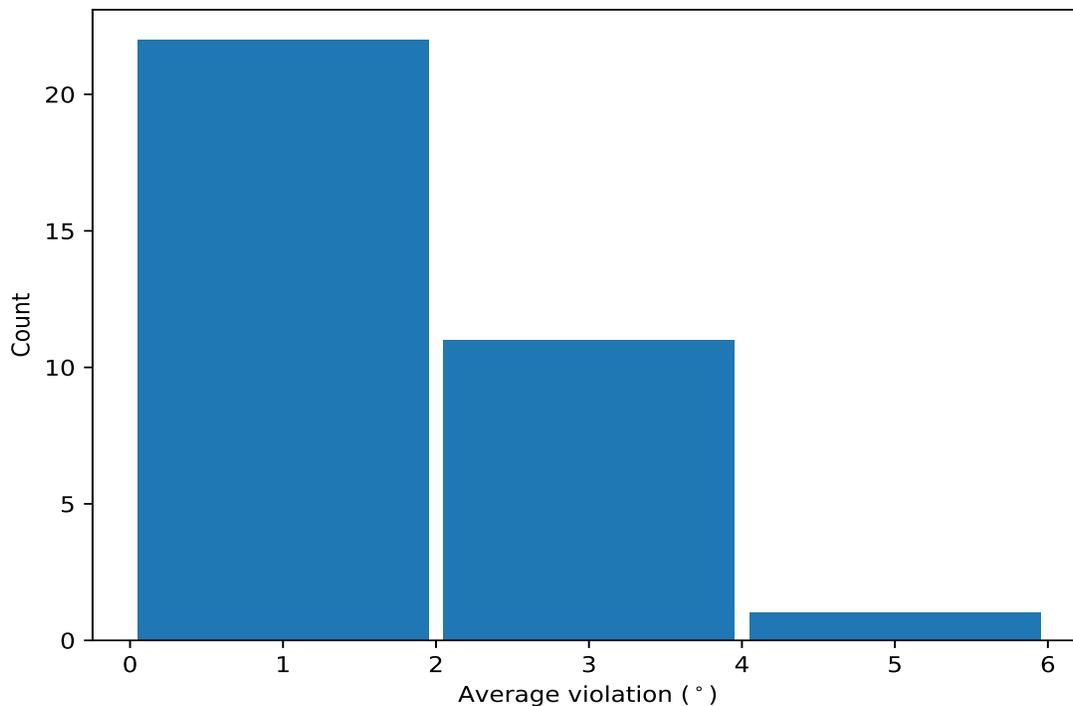


## 10.4 Most violated dihedral-angle restraints in the ensemble [i](#)

### 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.

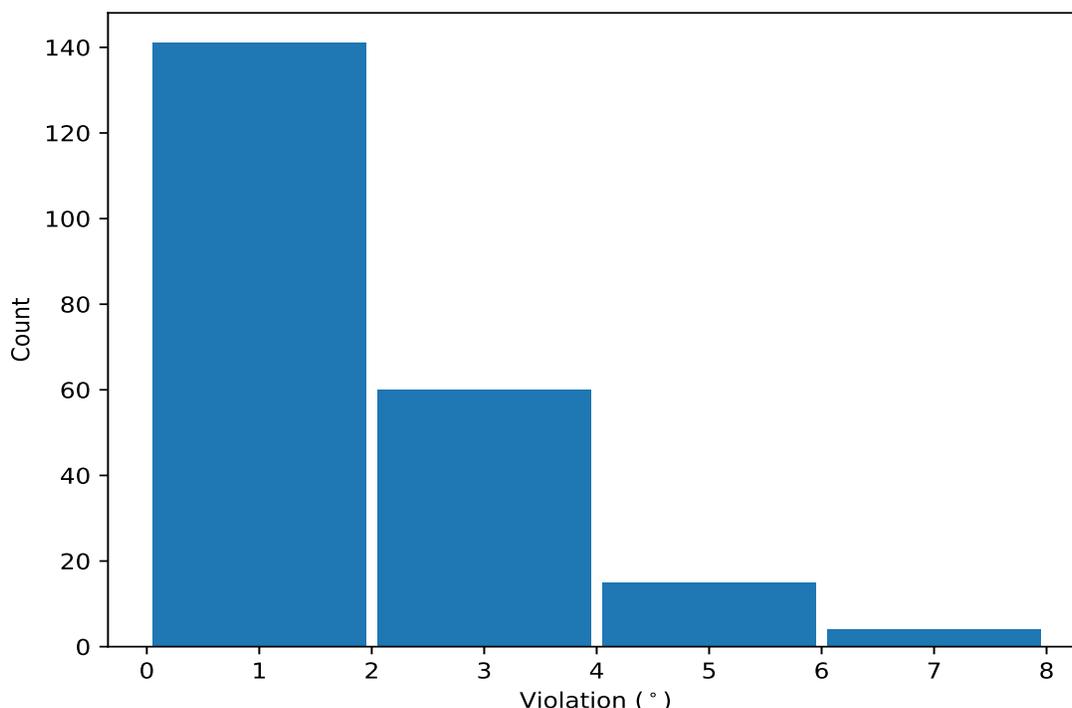
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models <sup>1</sup>	Mean	SD <sup>2</sup>	Median
(1,44)	1:31:A:CYS:N	1:31:A:CYS:CA	1:31:A:CYS:C	1:32:A:SER:N	20	4.32	1.39	4.28
(1,163)	1:44:A:GLU:CA	1:44:A:GLU:CB	1:44:A:GLU:CG	1:44:A:GLU:CD	17	2.53	1.12	2.4
(1,147)	1:134:A:ASN:C	1:135:A:ALA:N	1:135:A:ALA:CA	1:135:A:ALA:C	12	1.98	0.66	1.88
(1,160)	1:141:A:GLN:C	1:142:A:SER:N	1:142:A:SER:CA	1:142:A:SER:C	11	1.4	0.25	1.36
(1,20)	1:16:A:PHE:N	1:16:A:PHE:CA	1:16:A:PHE:C	1:17:A:PRO:N	9	2.0	0.82	1.84
(1,48)	1:37:A:HIS:N	1:37:A:HIS:CA	1:37:A:HIS:C	1:38:A:LEU:N	8	1.83	0.55	1.66
(1,7)	1:7:A:TRP:C	1:8:A:ARG:N	1:8:A:ARG:CA	1:8:A:ARG:C	8	1.6	0.51	1.44
(1,46)	1:36:A:ALA:N	1:36:A:ALA:CA	1:36:A:ALA:C	1:37:A:HIS:N	7	2.29	0.66	2.63
(1,38)	1:28:A:GLU:N	1:28:A:GLU:CA	1:28:A:GLU:C	1:29:A:ARG:N	7	2.01	0.93	1.61
(1,145)	1:133:A:LEU:C	1:134:A:ASN:N	1:134:A:ASN:CA	1:134:A:ASN:C	7	1.99	0.91	1.91

<sup>1</sup> Number of violated models, <sup>2</sup>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 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 (°)
(1,44)	1:31:A:CYS:N	1:31:A:CYS:CA	1:31:A:CYS:C	1:32:A:SER:N	13	7.97
(1,2)	1:2:A:VAL:N	1:2:A:VAL:CA	1:2:A:VAL:C	1:3:A:CYS:N	19	7.02
(1,3)	1:2:A:VAL:C	1:3:A:CYS:N	1:3:A:CYS:CA	1:3:A:CYS:C	19	6.46
(1,44)	1:31:A:CYS:N	1:31:A:CYS:CA	1:31:A:CYS:C	1:32:A:SER:N	17	6.33
(1,44)	1:31:A:CYS:N	1:31:A:CYS:CA	1:31:A:CYS:C	1:32:A:SER:N	2	5.76
(1,163)	1:44:A:GLU:CA	1:44:A:GLU:CB	1:44:A:GLU:CG	1:44:A:GLU:CD	1	5.59
(1,6)	1:4:A:PRO:N	1:4:A:PRO:CA	1:4:A:PRO:C	1:5:A:ILE:N	19	5.58
(1,44)	1:31:A:CYS:N	1:31:A:CYS:CA	1:31:A:CYS:C	1:32:A:SER:N	5	5.55
(1,164)	1:65:A:ARG:CA	1:65:A:ARG:CB	1:65:A:ARG:CG	1:65:A:ARG:CD	11	5.12
(1,44)	1:31:A:CYS:N	1:31:A:CYS:CA	1:31:A:CYS:C	1:32:A:SER:N	9	5.1