WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB NMR Structure Validation Summary Report (i)

Dec 25, 2024 – 06:31 PM EST

PDB ID	:	7OVC
BMRB ID	:	34638
Title	:	Structure of the human UFC1 protein in complex with the UBA5 C-terminal UFC1-binding motif.
Authors	:	Wesch, W.; Loehr, F.; Rogova, N.; Doetsch, V.; Rogov, V.V.
Deposited on	:	2021-06-14

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

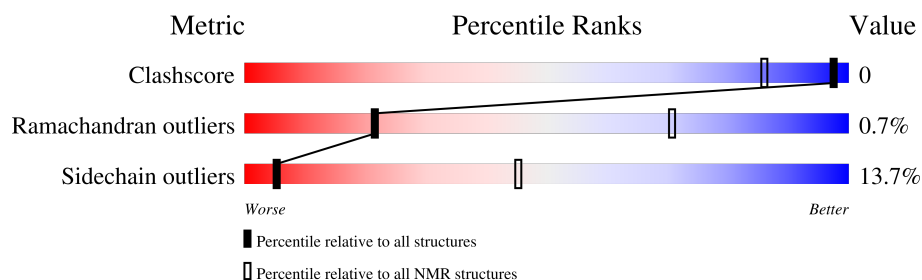|		
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
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 85%.

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	Percentile Ranks	Value
Clashscore		0
Ramachandran outliers		0.7%
Sidechain outliers		13.7%

Worse — *Better*

▮ Percentile relative to all structures

⟦⟧ Percentile relative to all NMR structures

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

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

Mol	Chain	Length	Quality of chain
1	A	167	81%　　　　　8%　10%
2	B	27	70%　　　　　11%　　19%

2 Ensemble composition and analysis ⓘ

This entry contains 20 models. Model 1 is the overall representative, medoid model (most similar to other models).

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

Well-defined (core) protein residues			
Well-defined core	Residue range (total)	Backbone RMSD (Å)	Medoid model
1	A:2-A:103, A:114-A:161, B:384-B:405 (172)	0.39	1

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

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

The models can be grouped into 4 clusters and 2 single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3147 atoms, of which 1567 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Ubiquitin-fold modifier-conjugating enzyme 1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	167	2739	884	1366	238	245	6	0

- Molecule 2 is a protein called Ubiquitin-like modifier-activating enzyme 5.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	27	408	127	201	31	45	4	0

There are 3 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
B	379	GLY	-	expression tag	UNP Q9GZZ9
B	380	MET	-	expression tag	UNP Q9GZZ9
B	405	TRP	-	expression tag	UNP Q9GZZ9

4 Residue-property plots ⓘ

4.1 Average score per residue in the NMR ensemble

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

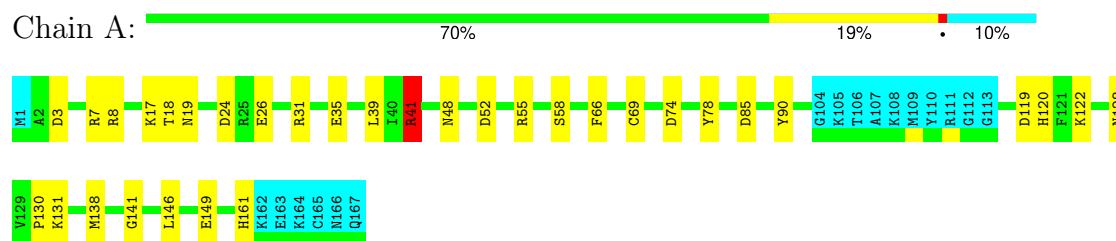• Molecule 1: Ubiquitin-fold modifier-conjugating enzyme 1
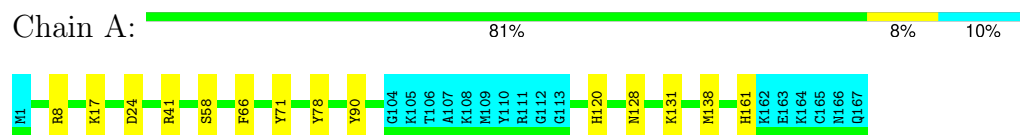
• Molecule 2: Ubiquitin-like modifier-activating enzyme 5
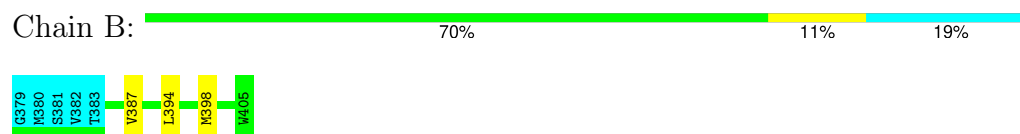
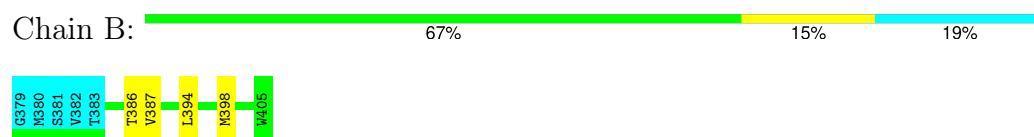
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

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

• Molecule 1: Ubiquitin-fold modifier-conjugating enzyme 1

• Molecule 2: Ubiquitin-like modifier-activating enzyme 5

5   Refinement protocol and experimental data overview ⓘ

The models were refined using the following method: *molecular dynamics.*

Of the 200 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy.*

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

Software name	Classification	Version
CYANA	structure calculation	3.98
OPALp	refinement	1.4

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	2268
Number of shifts mapped to atoms	2268
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	85%

6   Model quality ⓘ

6.1   Standard geometry ⓘ

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	0.62±0.01	0±0/1279 (0.0± 0.0%)	1.06±0.02	1±1/1743 (0.1± 0.1%)
2	B	0.51±0.01	0±0/176 (0.0± 0.0%)	1.06±0.04	0±0/234 (0.0± 0.1%)
All	All	0.60	0/29100 (0.0%)	1.06	30/39540 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	Chirality	Planarity
1	A	0.0±0.0	1.1±1.1
All	All	0	21

There are no bond-length outliers.

5 of 14 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	A	64	ARG	NE-CZ-NH1	8.98	124.79	120.30	13	1
1	A	64	ARG	NE-CZ-NH2	-7.80	116.40	120.30	13	1
1	A	64	ARG	CD-NE-CZ	6.46	132.64	123.60	13	1
1	A	31	ARG	CD-NE-CZ	6.39	132.55	123.60	5	7
1	A	31	ARG	NE-CZ-NH1	6.08	123.34	120.30	16	4

There are no chirality outliers.

5 of 10 unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	7	ARG	Sidechain	6
1	A	41	ARG	Sidechain	3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	55	ARG	Sidechain	3
1	A	8	ARG	Sidechain	2
1	A	42	TYR	Sidechain	2

6.2 Too-close contacts ⓘ

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	1242	1228	1228	1±1
2	B	175	168	168	1±1
All	All	28340	27920	27920	27

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 0.

5 of 13 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Total
1:A:17:LYS:HE2	2:B:386:THR:HG23	0.55	1.77	7	6
1:A:28:TRP:CZ2	1:A:32:LEU:HD13	0.51	2.39	6	1
1:A:94:ALA:HB2	1:A:125:TRP:CD2	0.50	2.41	13	2
1:A:99:VAL:CG1	1:A:115:ILE:HG21	0.48	2.39	16	1
1:A:94:ALA:HB2	1:A:125:TRP:CE2	0.47	2.45	13	2

6.3 Torsion angles ⓘ

6.3.1 Protein backbone ⓘ

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	150/167 (90%)	140±2 (93±1%)	10±2 (6±1%)	1±1 (1±1%)	24	72
2	B	21/27 (78%)	17±1 (82±6%)	3±1 (16±6%)	0±0 (1±2%)	12	59

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	3420/3880 (88%)	3138 (92%)	259 (8%)	23 (1%)	21 71

5 of 8 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	141	GLY	7
2	B	389	ASP	5
1	A	119	ASP	4
1	A	130	PRO	3
1	A	52	ASP	1

6.3.2   Protein sidechains ⓘ

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	133/146 (91%)	116±3 (87±3%)	18±3 (13±3%)	6 46
2	B	20/24 (83%)	17±1 (83±5%)	3±1 (17±5%)	4 38
All	All	3060/3400 (90%)	2642 (86%)	418 (14%)	5 45

5 of 66 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	66	PHE	20
1	A	90	TYR	20
1	A	128	ASN	20
1	A	131	LYS	20
2	B	394	LEU	20

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 ⓘ

There are no ligands in this entry.

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 85% for the well-defined parts and 83% 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	2268
Number of shifts mapped to atoms	2268
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	14

7.1.2 Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_\alpha$	190	-0.47 ± 0.09	None needed (< 0.5 ppm)
$^{13}C_\beta$	179	0.03 ± 0.12	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	175	0.44 ± 0.33	None needed (< 0.5 ppm)

7.1.3 Completeness of resonance assignments ⓘ

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 85%, i.e. 2079 atoms were assigned a chemical shift out of a possible 2445. 0 out of 29 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	671/848 (79%)	339/342 (99%)	172/344 (50%)	160/162 (99%)
Sidechain	1245/1380 (90%)	849/894 (95%)	384/435 (88%)	12/51 (24%)

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

|	Total	^{1}H	^{13}C	^{15}N
Aromatic	163/217 (75%)	91/108 (84%)	65/98 (66%)	7/11 (64%)
Overall	2079/2445 (85%)	1279/1344 (95%)	621/877 (71%)	179/224 (80%)

7.1.4 Statistically unusual chemical shifts ⓘ

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

| List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	58	SER	HB2	0.81	2.61 − 5.13	-12.1
1	A	149	GLU	HB2	0.28	1.00 − 3.05	-8.5
1	A	149	GLU	HG2	0.84	1.24 − 3.30	-6.9
1	A	65	TRP	HB2	1.03	1.51 − 4.87	-6.4
1	A	128	ASN	HB2	0.94	1.27 − 4.34	-6.1
1	A	22	PRO	HA	2.44	2.78 − 6.00	-6.0
1	A	46	ASN	HD21	4.58	4.94 − 9.72	-5.8
1	A	94	ALA	HB1	0.03	0.14 − 2.58	-5.5
1	A	94	ALA	HB2	0.03	0.14 − 2.58	-5.5
1	A	94	ALA	HB3	0.03	0.14 − 2.58	-5.5
1	A	54	PHE	HD2	5.40	5.52 − 8.61	-5.4
1	A	54	PHE	HD1	5.40	5.51 − 8.60	-5.4
1	A	115	ILE	HB	0.29	0.35 − 3.22	-5.2
1	A	55	ARG	HB2	0.50	0.52 − 3.08	-5.0

7.1.5 Random Coil Index (RCI) plots ⓘ

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:
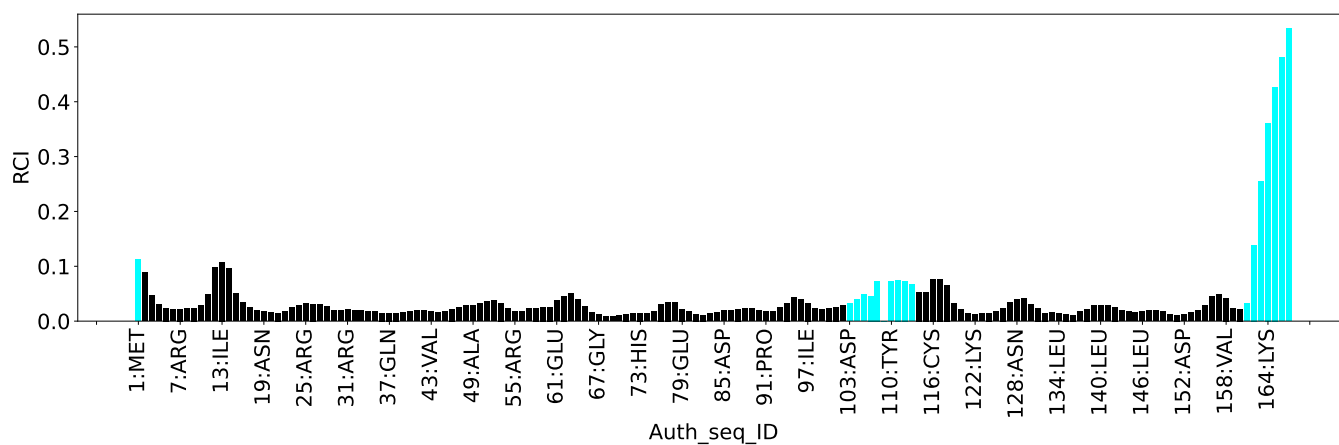
Random coil index (RCI) for chain B:
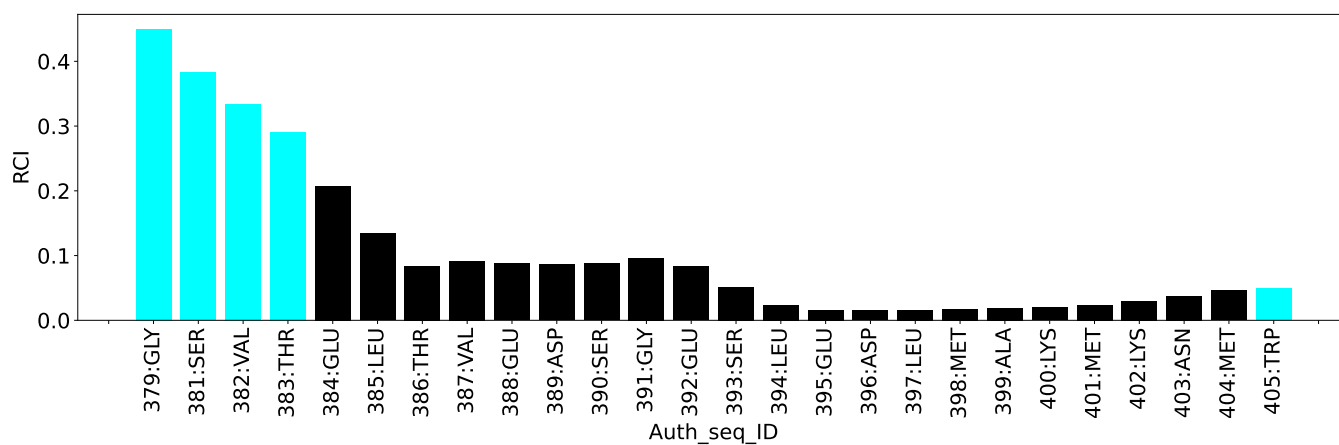
8   NMR restraints analysis (i)

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	3903
Intra-residue ($|$i-j$|$=0)	717
Sequential ($|$i-j$|$=1)	895
Medium range ($|$i-j$|$>1 and $|$i-j$|$<5)	841
Long range ($|$i-j$|$≥5)	931
Inter-chain	377
Hydrogen bond restraints	142
Disulfide bond restraints	0
Total dihedral-angle restraints	558
Number of unmapped restraints	0
Number of restraints per residue	23.0
Number of long range restraints per residue[1]	4.9

[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)	0.1	0.18
0.2-0.5 (Medium)	0.7	0.48
>0.5 (Large)	0.1	0.57

8.2.2 Average number of dihedral-angle violations per model ⓘ

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)	75.4	4.17
10.0-20.0 (Medium)	None	None
>20.0 (Large)	None	None

9 Distance violation analysis ⓘ

9.1 Summary of distance violations ⓘ

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)	717	18.4	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	5	0.1	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	515	13.2	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	197	5.0	0	0.0	0.0	0	0.0	0.0
Sequential (|i-j|=1)	895	22.9	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	185	4.7	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	562	14.4	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	148	3.8	0	0.0	0.0	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	841	21.5	1	0.1	0.0	0	0.0	0.0
Backbone-Backbone	248	6.4	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	335	8.6	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	258	6.6	1	0.4	0.0	0	0.0	0.0
Long range (|i-j|≥5)	931	23.9	3	0.3	0.1	0	0.0	0.0
Backbone-Backbone	52	1.3	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	358	9.2	1	0.3	0.0	0	0.0	0.0
Sidechain-Sidechain	521	13.3	2	0.4	0.1	0	0.0	0.0
Inter-chain	377	9.7	3	0.8	0.1	0	0.0	0.0
Backbone-Backbone	15	0.4	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	175	4.5	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	187	4.8	3	1.6	0.1	0	0.0	0.0
Hydrogen bond	142	3.6	0	0.0	0.0	0	0.0	0.0
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	3903	100.0	7	0.2	0.2	0	0.0	0.0
Backbone-Backbone	505	12.9	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	2087	53.5	1	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	1311	33.6	6	0.5	0.2	0	0.0	0.0

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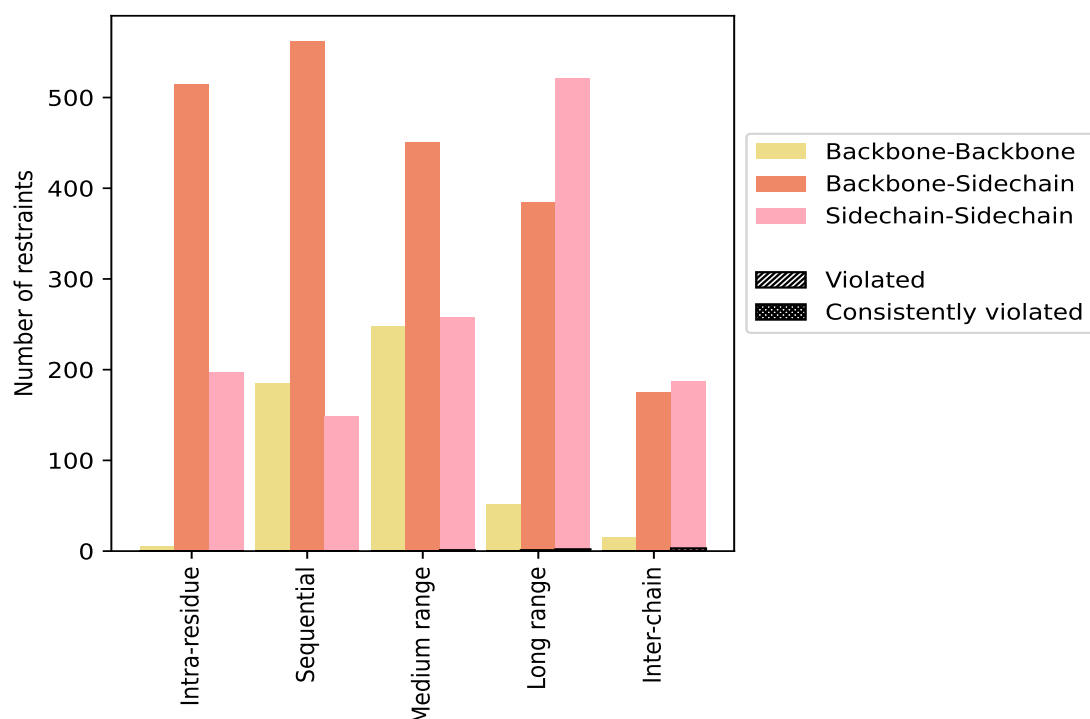
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories. The hydrogen bonds and disulfied bonds are counted in their appropriate category on the x-axis

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	0	0	0	0	2	2	0.21	0.32	0.11	0.21
2	0	0	1	0	2	3	0.18	0.35	0.12	0.1
3	0	0	0	0	0	0	0.0	0.0	0.0	0.0
4	0	0	0	0	1	1	0.48	0.48	0.0	0.48
5	0	0	0	0	1	1	0.42	0.42	0.0	0.42
6	0	0	0	0	1	1	0.18	0.18	0.0	0.18
7	0	0	0	0	1	1	0.25	0.25	0.0	0.25
8	0	0	0	0	1	1	0.57	0.57	0.0	0.57
9	0	0	0	1	0	1	0.1	0.1	0.0	0.1
10	0	0	0	0	1	1	0.28	0.28	0.0	0.28

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	0	0	0	0	1	1	0.57	0.57	0.0	0.57
12	0	0	0	0	1	1	0.32	0.32	0.0	0.32
13	0	0	0	1	0	1	0.1	0.1	0.0	0.1
14	0	0	0	0	1	1	0.23	0.23	0.0	0.23
15	0	0	0	1	1	2	0.29	0.46	0.17	0.29
16	0	0	0	1	1	2	0.22	0.35	0.12	0.22
17	0	0	0	0	1	1	0.43	0.43	0.0	0.43
18	0	0	0	0	0	0	0.0	0.0	0.0	0.0
19	0	0	0	0	1	1	0.24	0.24	0.0	0.24
20	0	0	0	0	1	1	0.3	0.3	0.0	0.3

[1]Intra-residue restraints, [2]Sequential restraints, [3]Medium range restraints, [4]Long range restraints, [5]Inter-chain restraints, [6]Standard deviation

9.2.1   Bar graph : Distance Violation statistics for each model ⓘ
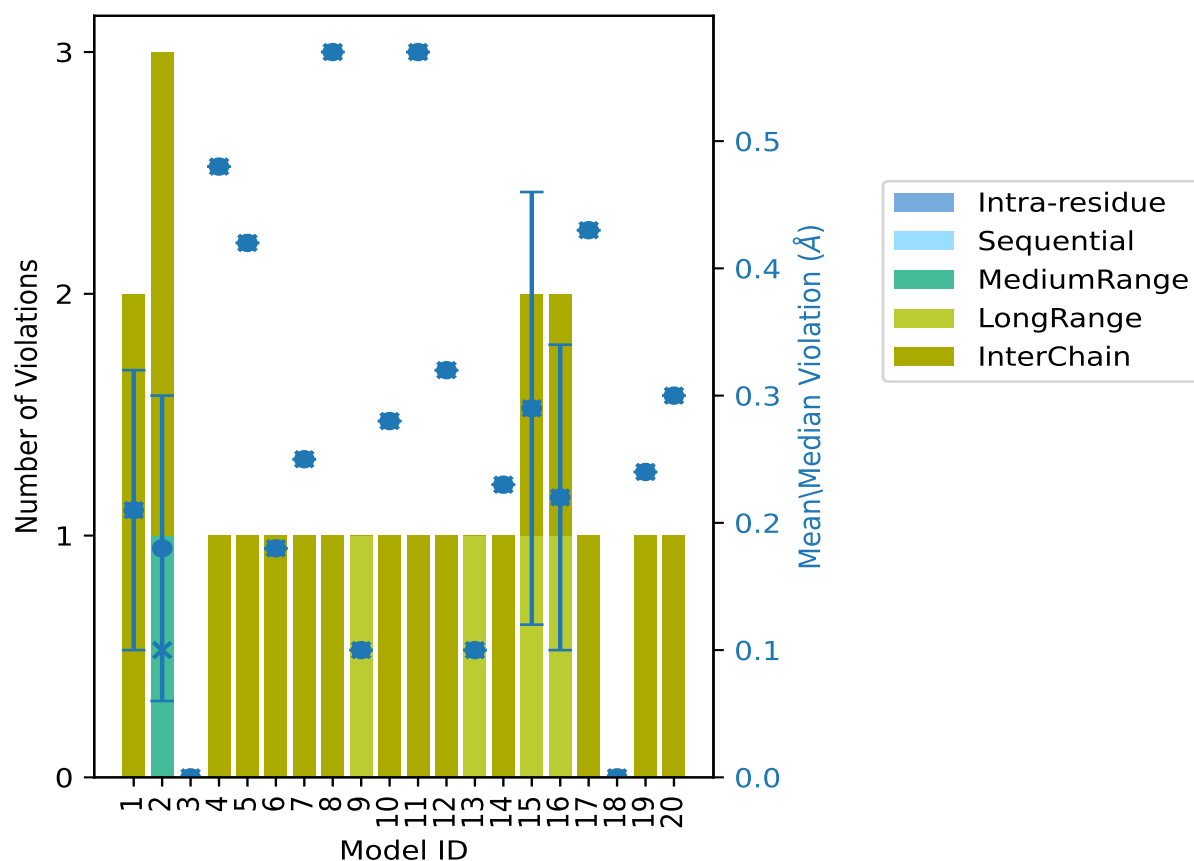
The mean(dot),median(x) and the standard deviation are shown in blue with respect to the y axis on the right

## 9.3 Distance violation statistics for the ensemble ⓘ

Violation analysis may find that some restraints are violated in few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of the ensemble. In total, 3754(IR:717, SQ:895, MR:840, LR:928, IC:374) 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	0	1	2	2	5	1	5.0
0	0	0	1	0	1	2	10.0
0	0	0	0	0	0	3	15.0
0	0	0	0	0	0	4	20.0
0	0	0	0	0	0	5	25.0
0	0	0	0	0	0	6	30.0
0	0	0	0	0	0	7	35.0
0	0	0	0	0	0	8	40.0
0	0	0	0	0	0	9	45.0
0	0	0	0	0	0	10	50.0
0	0	0	0	0	0	11	55.0
0	0	0	0	0	0	12	60.0
0	0	0	0	0	0	13	65.0
0	0	0	0	0	0	14	70.0
0	0	0	0	0	0	15	75.0
0	0	0	0	1	1	16	80.0
0	0	0	0	0	0	17	85.0
0	0	0	0	0	0	18	90.0
0	0	0	0	0	0	19	95.0
0	0	0	0	0	0	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 ⓘ
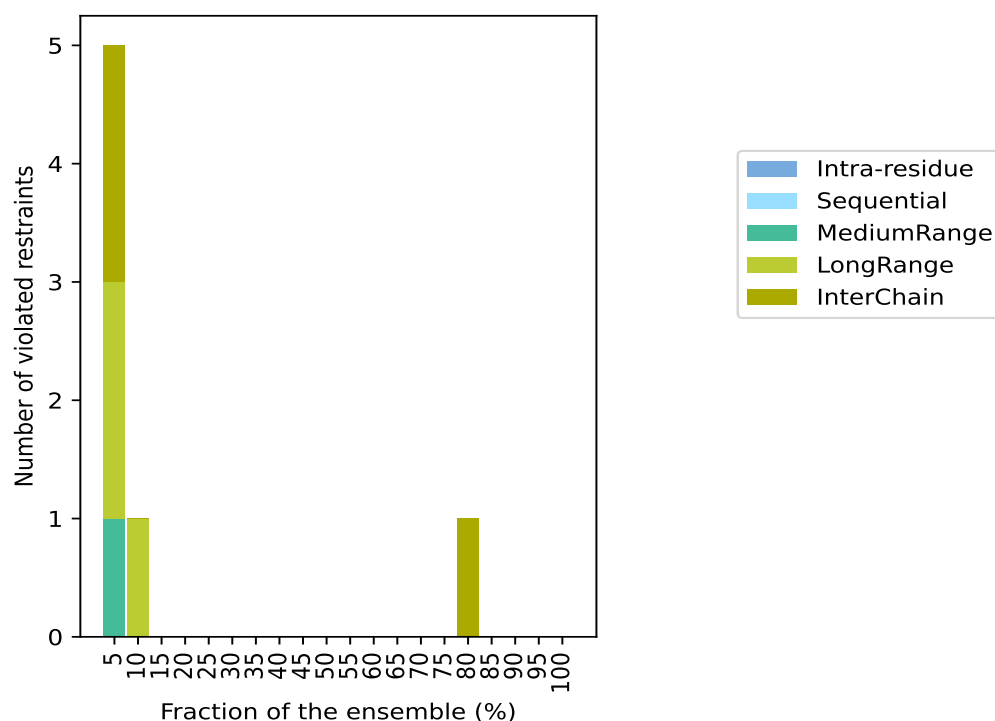
9.4 Most violated distance restraints in the ensemble ⓘ

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
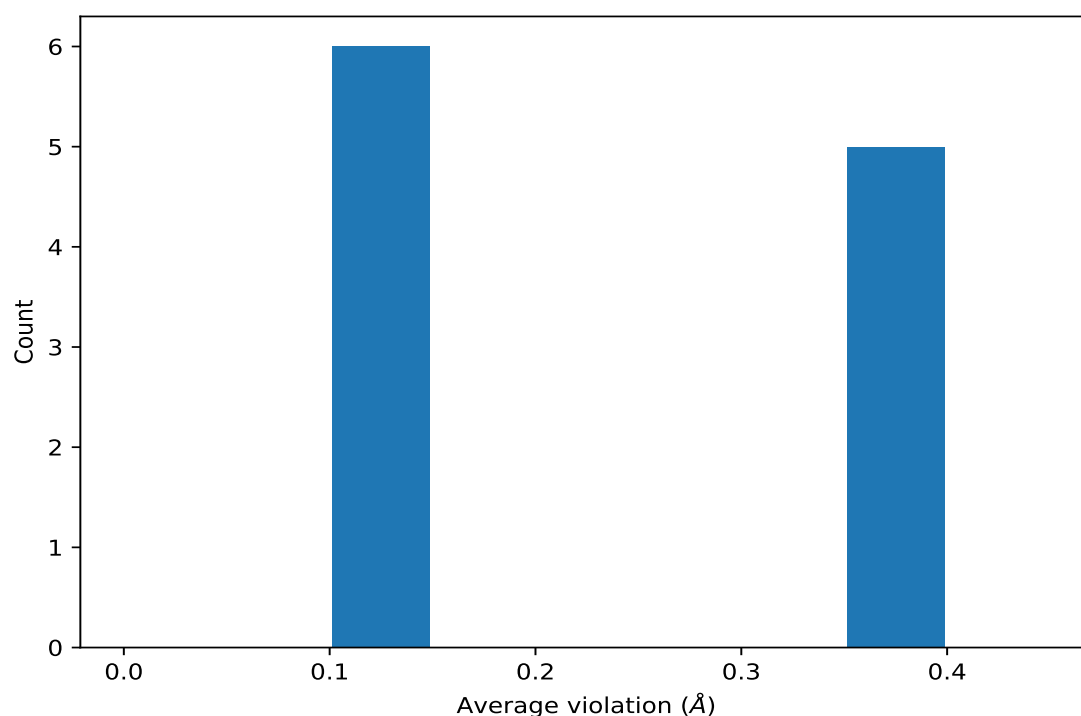
9.4.2   Table: Most violated distance restraints ⓘ

The following table provides the mean and the standard deviation of theviolation for each restraint sorted by number of violated models and the mean value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint. Rows with same key represent combinatorial or ambiguous restraints and are counted as a single restraint.
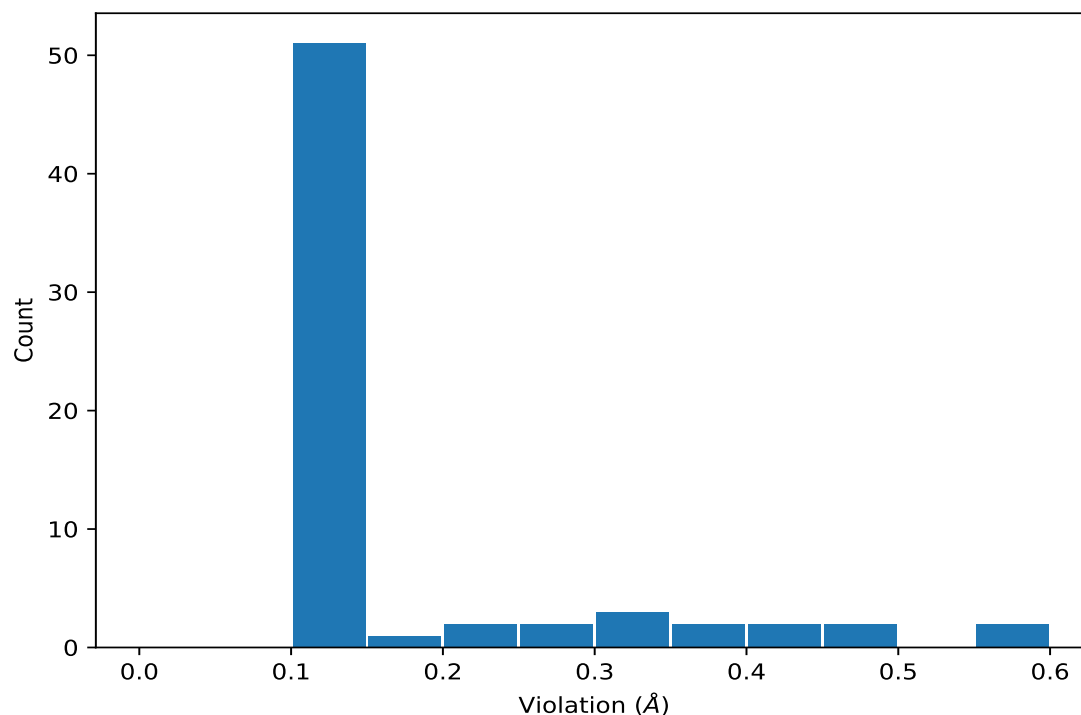
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,3717)	1:40:A:ILE:HG13	2:397:B:LEU:HD11	16	0.36	0.11	0.34
(1,3717)	1:40:A:ILE:HB	2:397:B:LEU:HD13	16	0.36	0.11	0.34
(1,3717)	1:40:A:ILE:HG13	2:397:B:LEU:HD13	16	0.36	0.11	0.34
(1,3717)	1:40:A:ILE:HG13	2:397:B:LEU:HD12	16	0.36	0.11	0.34
(1,3717)	1:40:A:ILE:HB	2:397:B:LEU:HD11	16	0.36	0.11	0.34
(1,2506)	1:117:A:LEU:HD11	1:142:A:LEU:H	2	0.1	0.0	0.1
(1,2506)	1:117:A:LEU:HD12	1:142:A:LEU:H	2	0.1	0.0	0.1
(1,2506)	1:117:A:LEU:HD13	1:142:A:LEU:H	2	0.1	0.0	0.1
(1,2506)	1:117:A:LEU:HD21	1:142:A:LEU:H	2	0.1	0.0	0.1
(1,2506)	1:117:A:LEU:HD22	1:142:A:LEU:H	2	0.1	0.0	0.1
(1,2506)	1:117:A:LEU:HD23	1:142:A:LEU:H	2	0.1	0.0	0.1

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

9.5 All violated distance restraints ⓘ

9.5.1 Histogram : Distribution of distance violations ⓘ

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

9.5.2 Table : All distance violations ⓘ

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

| Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,3717)	1:40:A:ILE:HG13	2:397:B:LEU:HD12	8	0.57
(1,3717)	1:40:A:ILE:HG13	2:397:B:LEU:HD13	11	0.57
(1,3717)	1:40:A:ILE:HB	2:397:B:LEU:HD13	4	0.48
(1,3717)	1:40:A:ILE:HB	2:397:B:LEU:HD11	15	0.46
(1,3717)	1:40:A:ILE:HG13	2:397:B:LEU:HD12	17	0.43
(1,3717)	1:40:A:ILE:HG13	2:397:B:LEU:HD11	5	0.42
(1,3717)	1:40:A:ILE:HG13	2:397:B:LEU:HD11	2	0.35
(1,3717)	1:40:A:ILE:HG13	2:397:B:LEU:HD13	16	0.35
(1,3717)	1:40:A:ILE:HG13	2:397:B:LEU:HD11	1	0.32
(1,3717)	1:40:A:ILE:HB	2:397:B:LEU:HD11	12	0.32

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	189	33.9	46	24.3	8.2	0	0.0	0.0
OMEGA	190	34.1	179	94.2	32.1	1	0.5	0.2
PHI	179	32.1	37	20.7	6.6	0	0.0	0.0
Total	558	100.0	262	47.0	47.0	1	0.2	0.2

[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 ⓘ
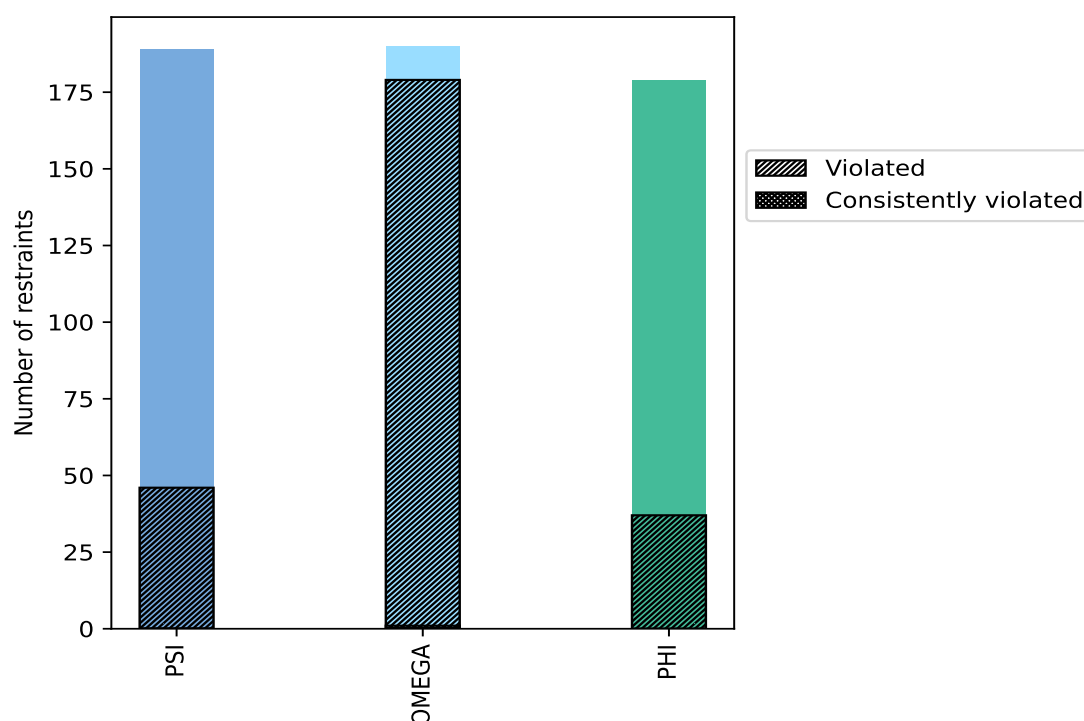
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories
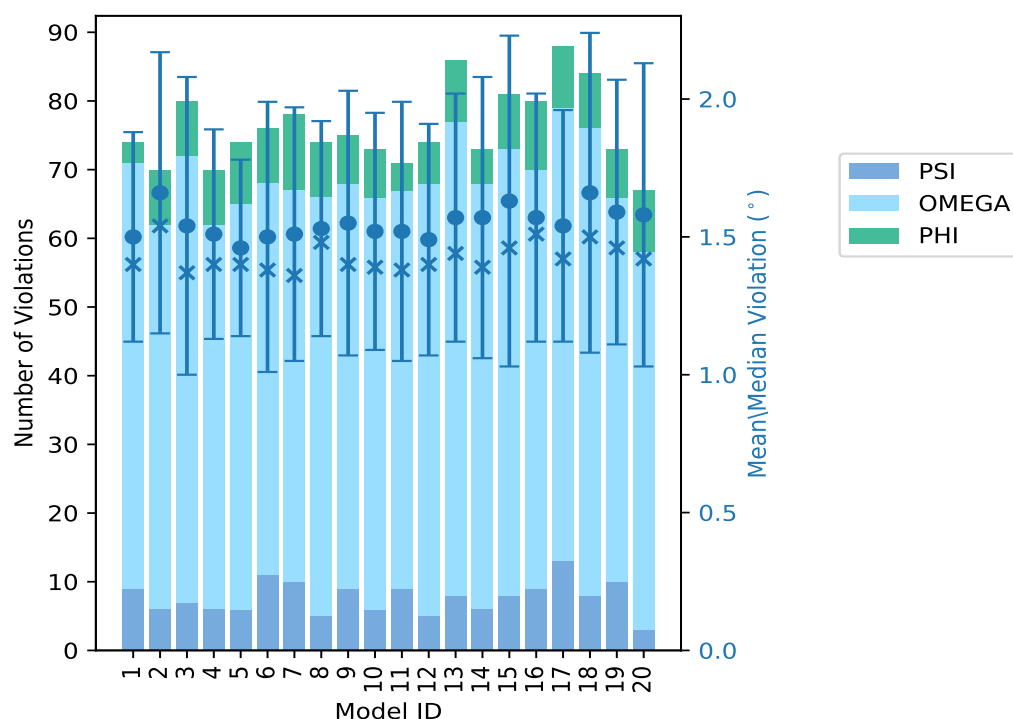
10.2   Dihedral-angle violation statistics for each model ⓘ

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	OMEGA	PHI	Total				
1	9	62	3	74	1.5	2.63	0.38	1.4
2	6	56	8	70	1.66	3.42	0.51	1.54
3	7	65	8	80	1.54	3.48	0.54	1.37
4	6	56	8	70	1.51	2.73	0.38	1.4
5	6	59	9	74	1.46	2.48	0.32	1.4
6	11	57	8	76	1.5	3.77	0.49	1.38
7	10	57	11	78	1.51	2.91	0.46	1.36
8	5	61	8	74	1.53	2.93	0.39	1.48
9	9	59	7	75	1.55	3.09	0.48	1.4
10	6	60	7	73	1.52	2.79	0.43	1.39
11	9	58	4	71	1.52	2.77	0.47	1.38
12	5	63	6	74	1.49	2.82	0.42	1.4
13	8	69	9	86	1.57	2.81	0.45	1.44
14	6	62	5	73	1.57	3.25	0.51	1.39
15	8	65	8	81	1.63	4.17	0.6	1.46
16	9	61	10	80	1.57	2.95	0.45	1.51
17	13	66	9	88	1.54	2.67	0.42	1.42
18	8	68	8	84	1.66	3.89	0.58	1.5
19	10	56	7	73	1.59	2.86	0.48	1.46
20	3	55	9	67	1.58	3.49	0.55	1.42

10.2.1   Bar graph : Dihedral violation statistics for each model ⓘ

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

10.3   Dihedral-angle violation statistics for the ensemble ⓘ

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

Number of violated restraints				Fraction of the ensemble	
PSI	OMEGA	PHI	Total	Count[1]	%
19	22	13	54	1	5.0
7	15	5	27	2	10.0
6	18	2	26	3	15.0
2	16	5	23	4	20.0
2	15	2	19	5	25.0
1	18	3	22	6	30.0
2	9	0	11	7	35.0
2	11	1	14	8	40.0
2	8	1	11	9	45.0
2	8	2	12	10	50.0
1	8	1	10	11	55.0

Continued on next page...

Continued from previous page...

Number of violated restraints				Fraction of the ensemble	
PSI	OMEGA	PHI	Total	Count[1]	%
0	3	1	4	12	60.0
0	5	0	5	13	65.0
0	5	0	5	14	70.0
0	5	1	6	15	75.0
0	4	0	4	16	80.0
0	4	0	4	17	85.0
0	3	0	3	18	90.0
0	1	0	1	19	95.0
0	1	0	1	20	100.0

[1] Number of models with violations

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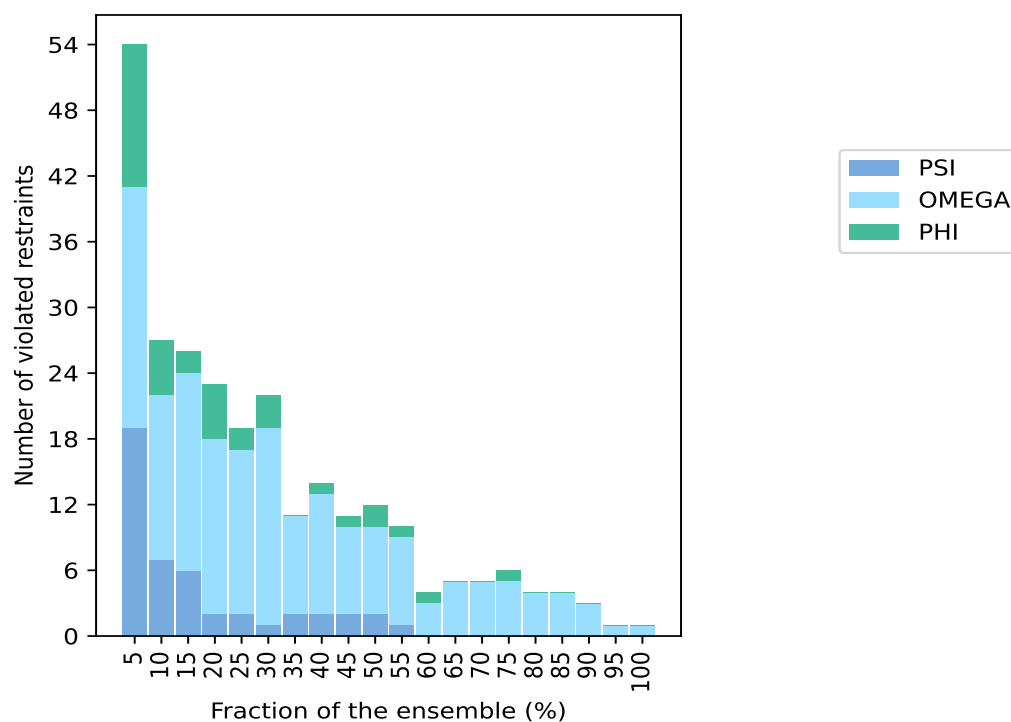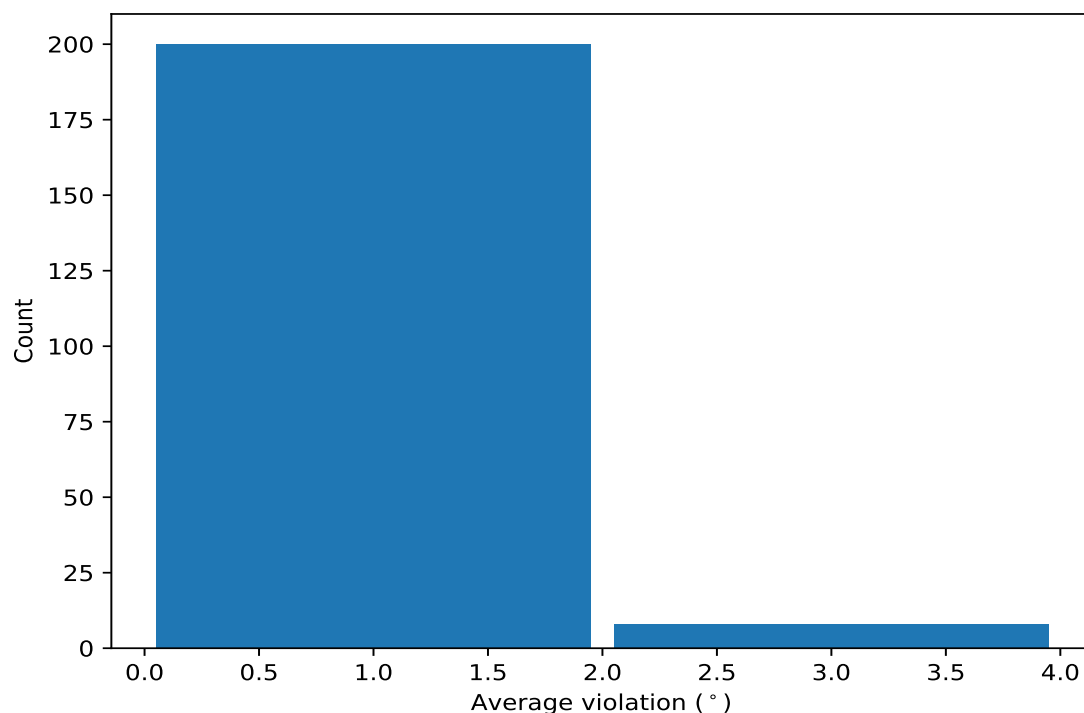

10.4   Most violated dihedral-angle restraints in the ensemble ⓘ

10.4.1   Histogram : Distribution of mean dihedral-angle violations ⓘ

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

in the ensemble

10.4.2   Table: Most violated dihedral-angle restraints ⓘ

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

Key	Atom-1	Atom-2	Atom-3	Atom-4	Models[1]	Mean	SD[2]	Median
(1,464)	1:97:A:ILE:CA	1:97:A:ILE:C	1:98:A:ALA:N	1:98:A:ALA:CA	20	2.0	0.46	2.02
(1,425)	1:58:A:SER:CA	1:58:A:SER:C	1:59:A:ASN:N	1:59:A:ASN:CA	19	1.84	0.71	1.61
(1,461)	1:94:A:ALA:CA	1:94:A:ALA:C	1:95:A:PRO:N	1:95:A:PRO:CA	18	1.91	0.49	1.83
(1,501)	1:134:A:LEU:CA	1:134:A:LEU:C	1:135:A:ALA:N	1:135:A:ALA:CA	18	1.87	0.37	1.82
(1,495)	1:128:A:ASN:CA	1:128:A:ASN:C	1:129:A:VAL:N	1:129:A:VAL:CA	18	1.82	0.45	1.84
(1,450)	1:83:A:GLU:CA	1:83:A:GLU:C	1:84:A:PHE:N	1:84:A:PHE:CA	17	2.1	0.61	2.08
(1,484)	1:117:A:LEU:CA	1:117:A:LEU:C	1:118:A:THR:N	1:118:A:THR:CA	17	1.73	0.39	1.77
(1,375)	1:8:A:ARG:CA	1:8:A:ARG:C	1:9:A:VAL:N	1:9:A:VAL:CA	17	1.5	0.31	1.38
(1,511)	1:144:A:PRO:CA	1:144:A:PRO:C	1:145:A:TRP:N	1:145:A:TRP:CA	17	1.45	0.25	1.44
(1,456)	1:89:A:THR:CA	1:89:A:THR:C	1:90:A:TYR:N	1:90:A:TYR:CA	16	1.9	0.57	1.78

[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.
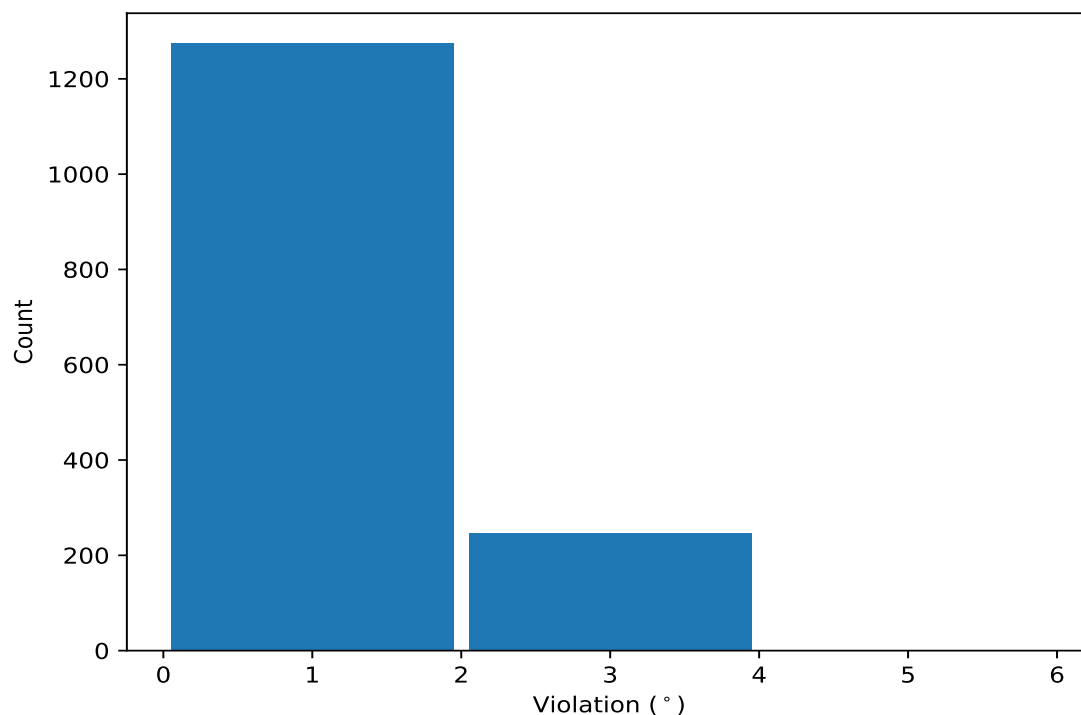
10.5.2   Table: All violated dihedral-angle restraints ⓘ

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 (°)
(1,381)	1:14:A:PRO:CA	1:14:A:PRO:C	1:15:A:VAL:N	1:15:A:VAL:CA	15	4.17
(1,515)	1:148:A:VAL:CA	1:148:A:VAL:C	1:149:A:GLU:N	1:149:A:GLU:CA	18	3.89
(1,220)	1:114:A:LYS:N	1:114:A:LYS:CA	1:114:A:LYS:C	1:115:A:ILE:N	18	3.85
(1,425)	1:58:A:SER:CA	1:58:A:SER:C	1:59:A:ASN:N	1:59:A:ASN:CA	6	3.77
(1,486)	1:119:A:ASP:CA	1:119:A:ASP:C	1:120:A:HIS:N	1:120:A:HIS:CA	15	3.76
(1,340)	2:390:B:SER:N	2:390:B:SER:CA	2:390:B:SER:C	2:391:B:GLY:N	6	3.5
(1,481)	1:114:A:LYS:CA	1:114:A:LYS:C	1:115:A:ILE:N	1:115:A:ILE:CA	20	3.49
(1,220)	1:114:A:LYS:N	1:114:A:LYS:CA	1:114:A:LYS:C	1:115:A:ILE:N	3	3.48
(1,381)	1:14:A:PRO:CA	1:14:A:PRO:C	1:15:A:VAL:N	1:15:A:VAL:CA	20	3.46
(1,450)	1:83:A:GLU:CA	1:83:A:GLU:C	1:84:A:PHE:N	1:84:A:PHE:CA	2	3.42

WORLDWIDE
wwPDB
PROTEIN DATA BANK