wwPDB NMR Structure Validation Summary Report (i)

Apr 30, 2024 – 08:47 PM EDT

PDB ID	:	8VSX
BMRB ID	:	51784
Title	:	NMR Structure of GCAP5 R22A
Authors	:	Cudia, D.L.; Ames, J.B.
Deposited on	:	2024-01-24

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
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)	:	2.36.2

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 68%.

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		103
Ramachandran outliers		5.9%
Sidechain outliers		8.6%

Worse   *Better*

▮ Percentile relative to all structures

▯ Percentile relative to all NMR structures
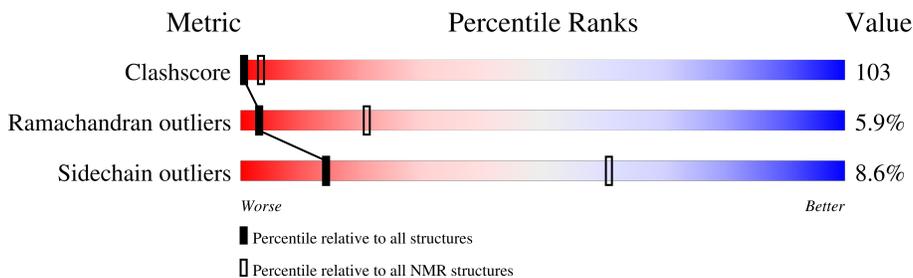
Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

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	198	19% 62% 6% 5% 8%

2 Ensemble composition and analysis ⓘ

This entry contains 10 models. Model 3 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *closest to the average.*

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:8-A:122, A:126-A:183 (173)	1.94	3

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

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

The models can be grouped into 2 clusters and 1 single-model cluster was found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2858 atoms, of which 1412 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Guanylyl cyclase-activating protein 1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	183	2858	926	1412	220	288	12	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MYR	-	expression tag	UNP Q5MAC8
A	22	ALA	ARG	engineered mutation	UNP Q5MAC8

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: Guanylyl cyclase-activating protein 1

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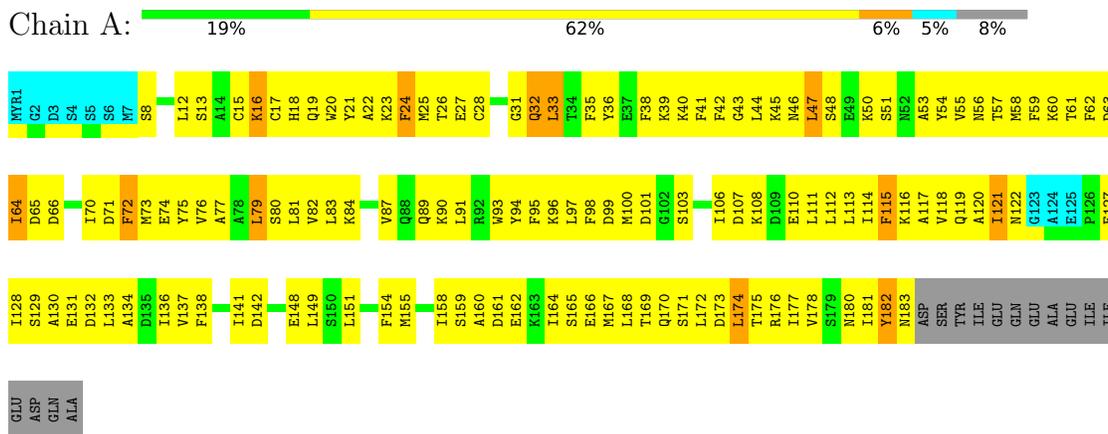
4.2   Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 3. Colouring as in section 4.1 above.

- Molecule 1: Guanylyl cyclase-activating protein 1

Chain A:
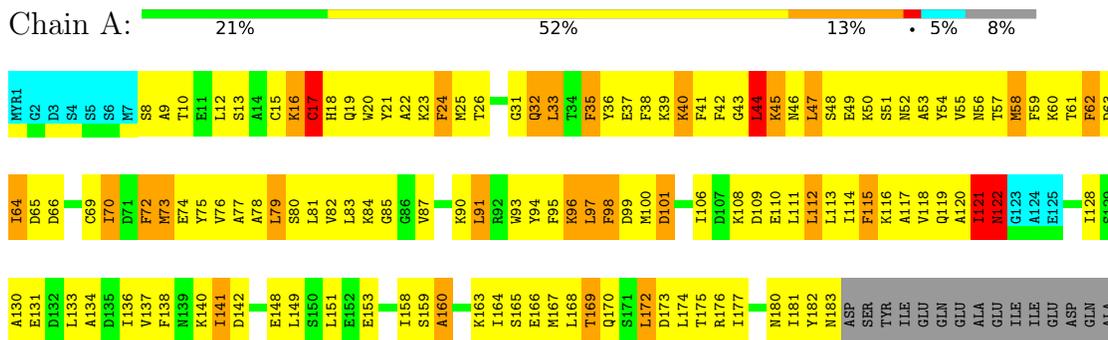
5 Refinement protocol and experimental data overview ⓘ

The models were refined using the following method: *DGSA-distance geometry simulated annealing.*

Of the 500 calculated structures, 10 were deposited, based on the following criterion: *structures with the least restraint violations.*

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

Software name	Classification	Version
X-PLOR NIH	structure calculation	
X-PLOR NIH	refinement	

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

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	1834
Number of shifts mapped to atoms	1665
Number of unparsed shifts	0
Number of shifts with mapping errors	169
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	68%

6 Model quality ⓘ

6.1 Standard geometry ⓘ

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

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

6.2 Too-close contacts ⓘ

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	1375	1340	1340	278±16
All	All	13750	13400	13400	2784

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 103.

5 of 2018 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Total
1:A:61:THR:HG23	1:A:97:LEU:HD23	1.10	1.24	5	2
1:A:97:LEU:HD13	1:A:98:PHE:N	1.09	1.63	5	1
1:A:121:ILE:HD13	1:A:122:ASN:N	1.07	1.64	2	8
1:A:118:VAL:O	1:A:121:ILE:HD12	1.05	1.47	7	9
1:A:115:PHE:CZ	1:A:130:ALA:HB1	1.04	1.87	9	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	172/198 (87%)	147±4 (85±2%)	15±2 (9±1%)	10±4 (6±2%)	3	21
All	All	1720/1980 (87%)	1470 (85%)	148 (9%)	102 (6%)	3	21

5 of 40 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	32	GLN	9
1	A	16	LYS	8
1	A	73	MET	6
1	A	18	HIS	5
1	A	101	ASP	5

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	152/171 (89%)	139±3 (91±2%)	13±3 (9±2%)	14	61
All	All	1520/1710 (89%)	1389 (91%)	131 (9%)	14	61

5 of 58 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	24	PHE	10
1	A	121	ILE	9
1	A	115	PHE	8
1	A	174	LEU	7
1	A	79	LEU	6

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 monosaccharides 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 68% for the well-defined parts and 69% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chemical_shifts_1*

7.1.1 Bookkeeping ⓘ

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

Total number of shifts	1834
Number of shifts mapped to atoms	1665
Number of unparsed shifts	0
Number of shifts with mapping errors	169
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 169) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	184	ASP	H	8.258	0.003	1
1	A	184	ASP	HA	4.549	0.001	1
1	A	184	ASP	HB2	2.65	0.002	1
1	A	184	ASP	HB3	2.65	0.002	1
1	A	184	ASP	C	176.198	0	1
1	A	184	ASP	CA	53.277	0.054	1
1	A	184	ASP	CB	39.776	0.038	1
1	A	184	ASP	N	121.156	0.048	1
1	A	185	SER	H	8.119	0.003	1
1	A	185	SER	HA	4.401	0.005	1
1	A	185	SER	HB2	3.806	0.007	1
1	A	185	SER	HB3	3.806	0.007	1
1	A	185	SER	C	173.966	0	1
1	A	185	SER	CA	57.21	0.1	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	185	SER	CB	62.573	0.036	1
1	A	185	SER	N	115.417	0.017	1
1	A	186	TYR	H	8.098	0.004	1
1	A	186	TYR	HA	4.531	0.002	1
1	A	186	TYR	HB2	2.984	0.006	1
1	A	186	TYR	HB3	2.984	0.006	1
1	A	186	TYR	HD1	7.084	0	1
1	A	186	TYR	HD2	7.084	0	1
1	A	186	TYR	HE1	6.805	0	1
1	A	186	TYR	HE2	6.805	0	1
1	A	186	TYR	C	175.191	0	1
1	A	186	TYR	CA	56.852	0.081	1
1	A	186	TYR	CB	37.434	0.034	1
1	A	186	TYR	N	122.631	0.014	1
1	A	187	ILE	H	7.8	0.004	1
1	A	187	ILE	HA	4.065	0.001	1
1	A	187	ILE	HB	1.726	0.005	1
1	A	187	ILE	HD11	0.816	0.002	1
1	A	187	ILE	HD12	0.816	0.002	1
1	A	187	ILE	HD13	0.816	0.002	1
1	A	187	ILE	HG12	1.393	0.004	1
1	A	187	ILE	HG13	1.09	0.003	1
1	A	187	ILE	HG21	0.828	0.008	1
1	A	187	ILE	HG22	0.828	0.008	1
1	A	187	ILE	HG23	0.828	0.008	1
1	A	187	ILE	C	175.451	0	1
1	A	187	ILE	CA	59.374	0.05	1
1	A	187	ILE	CB	37.836	0.086	1
1	A	187	ILE	CD1	11.613	0.039	1
1	A	187	ILE	CG1	25.987	0.094	1
1	A	187	ILE	CG2	16.163	0.055	1
1	A	187	ILE	N	124.043	0.026	1
1	A	188	GLU	H	8.335	0.001	1
1	A	188	GLU	HA	4.171	0	1
1	A	188	GLU	HB2	1.963	0.002	1
1	A	188	GLU	HB3	1.963	0.002	1
1	A	188	GLU	HG2	2.263	0	1
1	A	188	GLU	HG3	2.263	0	1
1	A	188	GLU	C	176.154	0	1
1	A	188	GLU	CA	55.325	0.039	1
1	A	188	GLU	CB	28.953	0.056	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	188	GLU	CG	35.026	0	1
1	A	188	GLU	N	125.584	0.014	1
1	A	189	GLN	H	8.323	0.001	1
1	A	189	GLN	HA	4.321	0.004	1
1	A	189	GLN	HB2	2.007	0.006	1
1	A	189	GLN	HB3	2.007	0.006	1
1	A	189	GLN	HG2	2.341	0	1
1	A	189	GLN	HG3	2.341	0	1
1	A	189	GLN	C	175.767	0	1
1	A	189	GLN	CA	54.38	0.05	1
1	A	189	GLN	CB	28.587	0.012	1
1	A	189	GLN	CG	32.49	0	1
1	A	189	GLN	N	122.19	0.026	1
1	A	190	GLU	H	8.489	0.001	1
1	A	190	GLU	HA	4.252	0.009	1
1	A	190	GLU	HB2	2.027	0.004	2
1	A	190	GLU	HB3	1.946	0.007	2
1	A	190	GLU	HG2	2.266	0	1
1	A	190	GLU	HG3	2.266	0	1
1	A	190	GLU	C	176.017	0	1
1	A	190	GLU	CA	55.254	0.101	1
1	A	190	GLU	CB	29.064	0.102	1
1	A	190	GLU	CG	35.047	0	1
1	A	190	GLU	N	123.118	0.015	1
1	A	191	ALA	H	8.288	0.002	1
1	A	191	ALA	HA	4.307	0.004	1
1	A	191	ALA	HB1	1.375	0.003	1
1	A	191	ALA	HB2	1.375	0.003	1
1	A	191	ALA	HB3	1.375	0.003	1
1	A	191	ALA	C	177.361	0	1
1	A	191	ALA	CA	51.047	0.063	1
1	A	191	ALA	CB	18.296	0.036	1
1	A	191	ALA	N	125.467	0.018	1
1	A	192	GLU	H	8.329	0.001	1
1	A	192	GLU	HA	4.26	0.003	1
1	A	192	GLU	C	176.099	0	1
1	A	192	GLU	CA	55.078	0.073	1
1	A	192	GLU	CB	29.253	0.038	1
1	A	192	GLU	CG	35.032	0	1
1	A	192	GLU	N	121.061	0.011	1
1	A	193	ILE	H	8.271	0.002	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	193	ILE	HA	4.154	0.002	1
1	A	193	ILE	HB	1.834	0.011	1
1	A	193	ILE	HD11	0.85	0.003	1
1	A	193	ILE	HD12	0.85	0.003	1
1	A	193	ILE	HD13	0.85	0.003	1
1	A	193	ILE	HG12	1.174	0.009	1
1	A	193	ILE	HG13	1.484	0.006	1
1	A	193	ILE	HG21	0.853	0.005	1
1	A	193	ILE	HG22	0.853	0.005	1
1	A	193	ILE	HG23	0.853	0.005	1
1	A	193	ILE	C	175.867	0	1
1	A	193	ILE	CA	59.543	0.077	1
1	A	193	ILE	CB	37.3	0.089	1
1	A	193	ILE	CD1	11.509	0.001	1
1	A	193	ILE	CG1	25.988	0.066	1
1	A	193	ILE	CG2	16.215	0.062	1
1	A	193	ILE	N	123.635	0.019	1
1	A	194	ILE	H	8.315	0.001	1
1	A	194	ILE	HA	4.202	0.004	1
1	A	194	ILE	HB	1.832	0.005	1
1	A	194	ILE	HD11	0.849	0.013	1
1	A	194	ILE	HD12	0.849	0.013	1
1	A	194	ILE	HD13	0.849	0.013	1
1	A	194	ILE	HG12	1.454	0.009	1
1	A	194	ILE	HG13	1.171	0.006	1
1	A	194	ILE	HG21	0.886	0.007	1
1	A	194	ILE	HG22	0.886	0.007	1
1	A	194	ILE	HG23	0.886	0.007	1
1	A	194	ILE	C	176.067	0	1
1	A	194	ILE	CA	59.376	0.074	1
1	A	194	ILE	CB	37.397	0.069	1
1	A	194	ILE	CD1	11.415	0.086	1
1	A	194	ILE	CG1	25.988	0.073	1
1	A	194	ILE	CG2	16.274	0.049	1
1	A	194	ILE	N	126.987	0.017	1
1	A	195	GLU	H	8.518	0.001	1
1	A	195	GLU	HA	4.307	0.002	1
1	A	195	GLU	HB2	2.021	0	2
1	A	195	GLU	HB3	1.929	0.023	2
1	A	195	GLU	HG2	2.225	0	1
1	A	195	GLU	HG3	2.225	0	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	195	GLU	C	175.952	0	1
1	A	195	GLU	CA	55.143	0.055	1
1	A	195	GLU	CB	29.344	0.032	1
1	A	195	GLU	CG	35.045	0	1
1	A	195	GLU	N	126.215	0.02	1
1	A	196	ASP	H	8.386	0.001	1
1	A	196	ASP	HA	4.571	0.007	1
1	A	196	ASP	HB2	2.642	0.007	1
1	A	196	ASP	HB3	2.642	0.007	1
1	A	196	ASP	C	175.969	0	1
1	A	196	ASP	CA	53.199	0.064	1
1	A	196	ASP	CB	39.902	0.029	1
1	A	196	ASP	N	122.145	0.014	1
1	A	197	GLN	H	8.286	0.002	1
1	A	197	GLN	HA	4.333	0.004	1
1	A	197	GLN	HB2	2.143	0.004	2
1	A	197	GLN	HB3	1.953	0.001	2
1	A	197	GLN	HG2	2.354	0	1
1	A	197	GLN	HG3	2.354	0	1
1	A	197	GLN	C	174.706	0	1
1	A	197	GLN	CA	54.369	0.046	1
1	A	197	GLN	CB	28.465	0.016	1
1	A	197	GLN	CG	32.54	0	1
1	A	197	GLN	N	121.186	0.032	1
1	A	198	ALA	H	7.994	0	1
1	A	198	ALA	HA	4.117	0.001	1
1	A	198	ALA	HB1	1.332	0.003	1
1	A	198	ALA	HB2	1.332	0.003	1
1	A	198	ALA	HB3	1.332	0.003	1
1	A	198	ALA	CA	52.672	0.004	1
1	A	198	ALA	CB	18.871	0.03	1
1	A	198	ALA	N	131.588	0.003	1

7.1.2   Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	196	0.82 ± 0.08	Should be checked
$^{13}C_{\beta}$	183	1.47 ± 0.04	Should be checked
$^{13}C'$	182	-0.28 ± 0.07	None needed (< 0.5 ppm)

Continued on next page...

Continued from previous page...

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
^{15}N	190	0.93 ± 0.24	Should be applied

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 68%, i.e. 1597 atoms were assigned a chemical shift out of a possible 2332. 0 out of 26 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	**^{1}H**	**^{13}C**	**^{15}N**
Backbone	801/871 (92%)	303/354 (86%)	332/346 (96%)	166/171 (97%)
Sidechain	736/1246 (59%)	474/809 (59%)	262/403 (65%)	0/34 (0%)
Aromatic	60/215 (28%)	59/105 (56%)	0/107 (0%)	1/3 (33%)
Overall	1597/2332 (68%)	836/1268 (66%)	594/856 (69%)	167/208 (80%)

7.1.4   Statistically unusual chemical shifts ⓘ

There are no statistically unusual chemical shifts.

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:
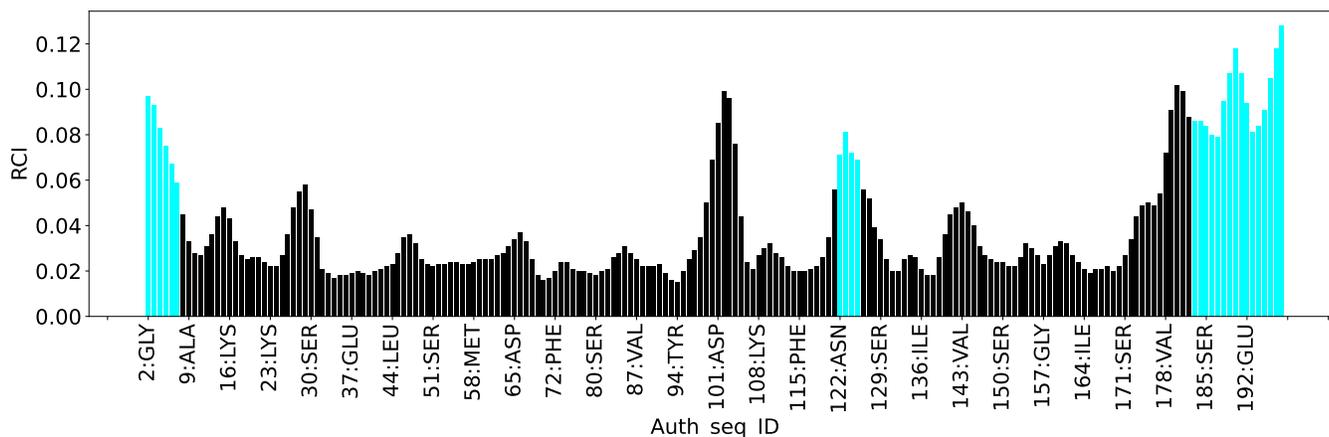
8 NMR restraints analysis ⓘ

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	738
Intra-residue ($\|i\text{-}j\|$=0)	2
Sequential ($\|i\text{-}j\|$=1)	18
Medium range ($\|i\text{-}j\|$>1 and $\|i\text{-}j\|$<5)	310
Long range ($\|i\text{-}j\|$≥5)	282
Inter-chain	0
Hydrogen bond restraints	126
Disulfide bond restraints	0
Total dihedral-angle restraints	230
Number of unmapped restraints	0
Number of restraints per residue	4.9
Number of long range restraints per residue[1]	1.4

[1]Long range hydrogen bonds and disulfide bonds are counted as long range restraints while calculating the number of long range restraints per residue

8.2 Residual restraint violations ⓘ

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

8.2.1 Average number of distance violations per model ⓘ

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

Bins (Å)	Average number of violations per model	Max (Å)
0.1-0.2 (Small)	26.7	0.2
0.2-0.5 (Medium)	14.3	0.48
>0.5 (Large)	None	None

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)	13.8	5.0
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)	2	0.3	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	1	0.1	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	1	0.1	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
Sequential (|i-j|=1)	18	2.4	2	11.1	0.3	1	5.6	0.1
Backbone-Backbone	13	1.8	1	7.7	0.1	0	0.0	0.0
Backbone-Sidechain	4	0.5	1	25.0	0.1	1	25.0	0.1
Sidechain-Sidechain	1	0.1	0	0.0	0.0	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	310	42.0	57	18.4	7.7	5	1.6	0.7
Backbone-Backbone	240	32.5	43	17.9	5.8	5	2.1	0.7
Backbone-Sidechain	65	8.8	11	16.9	1.5	0	0.0	0.0
Sidechain-Sidechain	5	0.7	3	60.0	0.4	0	0.0	0.0
Long range (|i-j|≥5)	282	38.2	55	19.5	7.5	4	1.4	0.5
Backbone-Backbone	28	3.8	9	32.1	1.2	0	0.0	0.0
Backbone-Sidechain	96	13.0	22	22.9	3.0	1	1.0	0.1
Sidechain-Sidechain	158	21.4	24	15.2	3.3	3	1.9	0.4
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	126	17.1	23	18.3	3.1	0	0.0	0.0
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	738	100.0	137	18.6	18.6	10	1.4	1.4
Backbone-Backbone	282	38.2	53	18.8	7.2	5	1.8	0.7
Backbone-Sidechain	292	39.6	57	19.5	7.7	2	0.7	0.3
Sidechain-Sidechain	164	22.2	27	16.5	3.7	3	1.8	0.4

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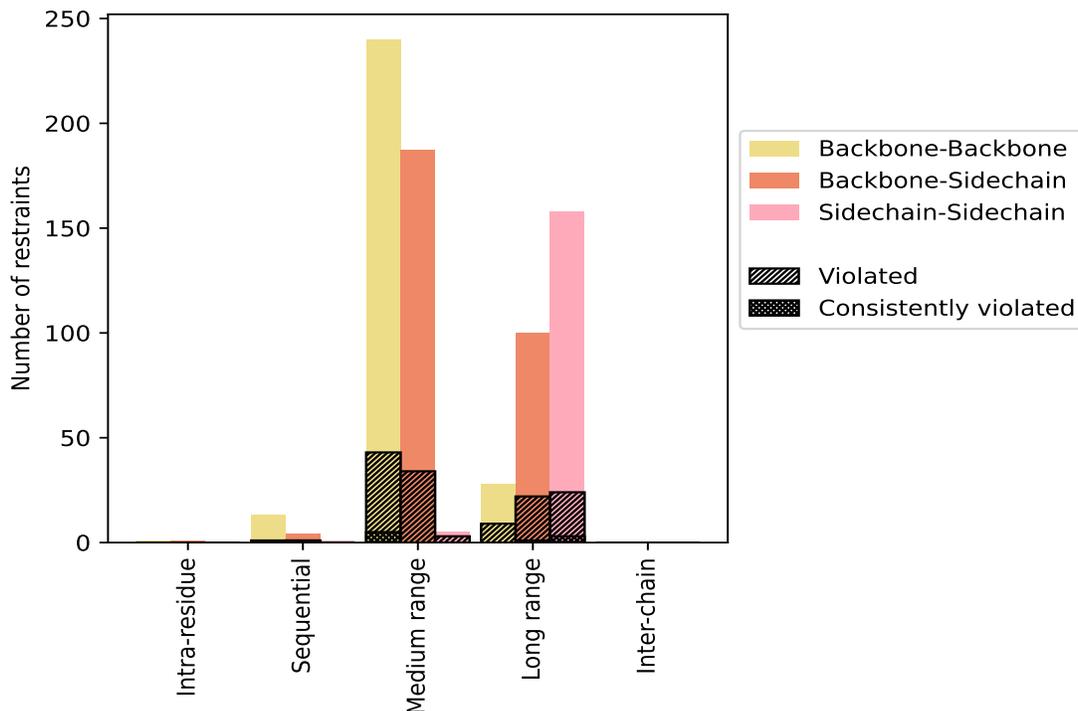
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories. The hydrogen bonds and disulfied bonds are counted in their appropriate category on the x-axis

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	1	27	19	0	47	0.19	0.43	0.07	0.15
2	0	1	19	16	0	36	0.18	0.42	0.07	0.16
3	0	2	23	13	0	38	0.18	0.4	0.06	0.17
4	0	1	21	20	0	42	0.18	0.41	0.06	0.16
5	0	1	20	15	0	36	0.19	0.41	0.08	0.18
6	0	1	21	18	0	40	0.2	0.42	0.09	0.15
7	0	2	34	17	0	53	0.2	0.48	0.09	0.17
8	0	1	27	20	0	48	0.2	0.48	0.09	0.18
9	0	1	26	21	0	48	0.18	0.35	0.06	0.16
10	0	1	24	19	0	44	0.18	0.46	0.08	0.15

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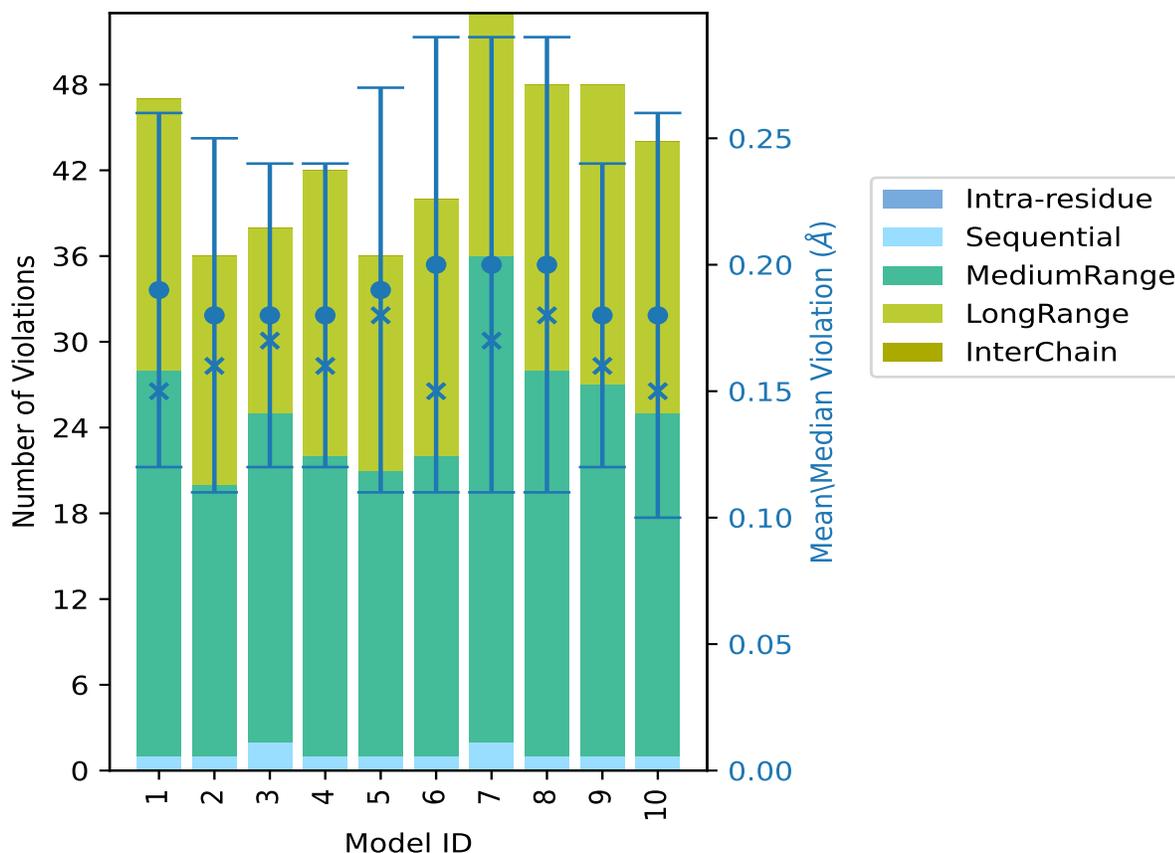
The mean(dot),median(x) and the standard deviation are shown in blue with respect to the y axis on the right

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, 498(IR:2, SQ:16, MR:253, LR:227, 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	0	17	23	0	40	1	10.0
0	1	16	10	0	27	2	20.0
0	0	7	4	0	11	3	30.0

Continued on next page...

WORLDWIDE
PDB
PROTEIN DATA BANK

Continued from previous page...

Number of violated restraints						Fraction of the ensemble	
IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total	Count[6]	%
0	0	3	5	0	8	4	40.0
0	0	4	2	0	6	5	50.0
0	0	1	0	0	1	6	60.0
0	0	1	4	0	5	7	70.0
0	0	1	2	0	3	8	80.0
0	0	2	1	0	3	9	90.0
0	1	5	4	0	10	10	100.0

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

9.3.1   Bar graph : Distance violation statistics for the ensemble (i)
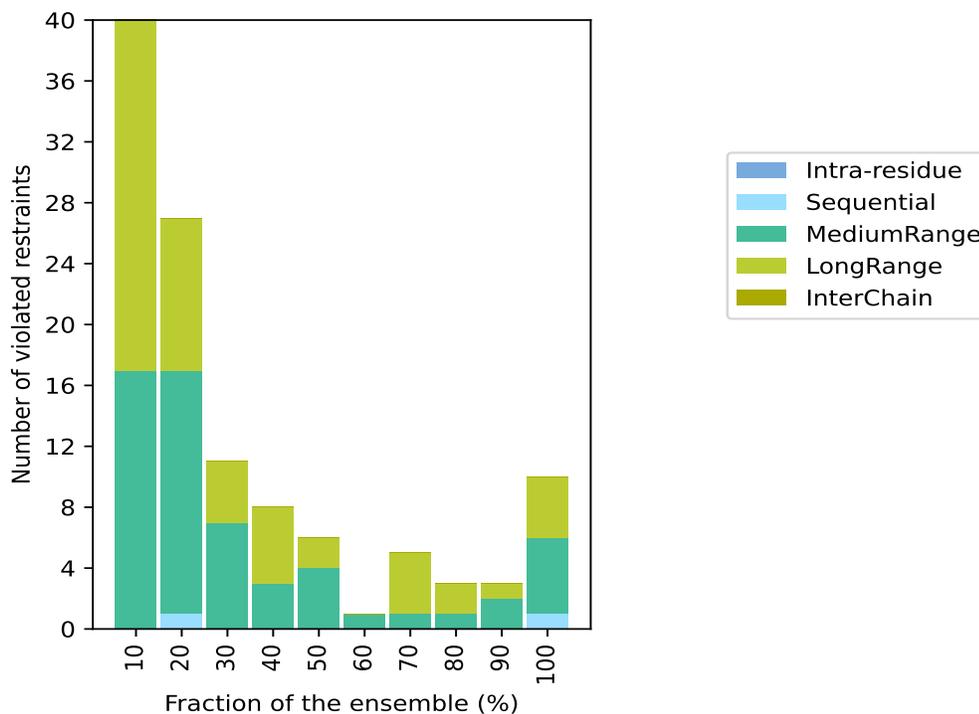
9.4   Most violated distance restraints in the ensemble (i)

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
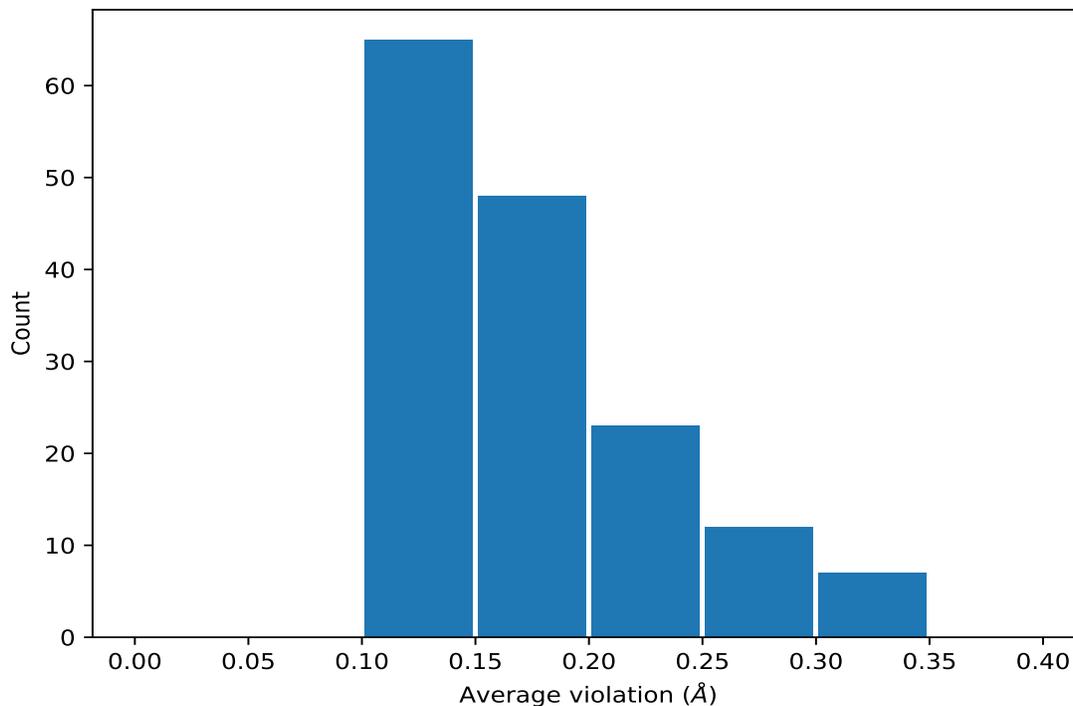
9.4.2   Table: Most violated distance restraints ⓘ

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

Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,600)	1:93:A:TRP:HB2	1:177:A:ILE:HG21	10	0.32	0.09	0.32
(1,600)	1:93:A:TRP:HB2	1:177:A:ILE:HG22	10	0.32	0.09	0.32
(1,600)	1:93:A:TRP:HB2	1:177:A:ILE:HG23	10	0.32	0.09	0.32
(1,600)	1:93:A:TRP:HB3	1:177:A:ILE:HG21	10	0.32	0.09	0.32
(1,600)	1:93:A:TRP:HB3	1:177:A:ILE:HG22	10	0.32	0.09	0.32
(1,600)	1:93:A:TRP:HB3	1:177:A:ILE:HG23	10	0.32	0.09	0.32
(1,468)	1:129:A:SER:HA	1:121:A:ILE:HD11	10	0.3	0.07	0.28
(1,468)	1:129:A:SER:HA	1:121:A:ILE:HD12	10	0.3	0.07	0.28
(1,468)	1:129:A:SER:HA	1:121:A:ILE:HD13	10	0.3	0.07	0.28
(1,599)	1:93:A:TRP:HB2	1:158:A:ILE:HG12	10	0.28	0.12	0.23
(1,599)	1:93:A:TRP:HB3	1:158:A:ILE:HG12	10	0.28	0.12	0.23
(1,124)	1:112:A:LEU:HA	1:115:A:PHE:H	10	0.25	0.06	0.24
(1,104)	1:55:A:VAL:HA	1:58:A:MET:H	10	0.24	0.04	0.23
(1,439)	1:15:A:CYS:H	1:16:A:LYS:HB3	10	0.24	0.08	0.24
(1,510)	1:35:A:PHE:HZ	1:55:A:VAL:HB	10	0.21	0.06	0.22

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,106)	1:57:A:THR:HA	1:60:A:LYS:H	10	0.19	0.05	0.19
(1,105)	1:56:A:ASN:HA	1:59:A:PHE:H	10	0.18	0.05	0.18
(1,90)	1:22:A:ALA:HA	1:25:A:MET:H	10	0.14	0.03	0.14

[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)
(2,79)	1:121:A:ILE:H	1:117:A:ALA:O	7	0.48
(1,599)	1:93:A:TRP:HB2	1:158:A:ILE:HG12	8	0.48

Continued on next page...

Continued from previous page...

| Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,599)	1:93:A:TRP:HB3	1:158:A:ILE:HG12	8	0.48
(1,600)	1:93:A:TRP:HB2	1:177:A:ILE:HG21	10	0.46
(1,600)	1:93:A:TRP:HB2	1:177:A:ILE:HG22	10	0.46
(1,600)	1:93:A:TRP:HB2	1:177:A:ILE:HG23	10	0.46
(1,600)	1:93:A:TRP:HB3	1:177:A:ILE:HG21	10	0.46
(1,600)	1:93:A:TRP:HB3	1:177:A:ILE:HG22	10	0.46
(1,600)	1:93:A:TRP:HB3	1:177:A:ILE:HG23	10	0.46
(1,468)	1:129:A:SER:HA	1:121:A:ILE:HD11	1	0.43
(1,468)	1:129:A:SER:HA	1:121:A:ILE:HD12	1	0.43
(1,468)	1:129:A:SER:HA	1:121:A:ILE:HD13	1	0.43
(1,600)	1:93:A:TRP:HB2	1:177:A:ILE:HG21	2	0.42
(1,600)	1:93:A:TRP:HB2	1:177:A:ILE:HG22	2	0.42
(1,600)	1:93:A:TRP:HB2	1:177:A:ILE:HG23	2	0.42
(1,600)	1:93:A:TRP:HB3	1:177:A:ILE:HG21	2	0.42
(1,600)	1:93:A:TRP:HB3	1:177:A:ILE:HG22	2	0.42
(1,600)	1:93:A:TRP:HB3	1:177:A:ILE:HG23	2	0.42
(1,290)	1:166:A:GLU:HA	1:169:A:THR:HB	6	0.42
(1,599)	1:93:A:TRP:HB2	1:158:A:ILE:HG12	4	0.41
(1,599)	1:93:A:TRP:HB3	1:158:A:ILE:HG12	4	0.41
(1,292)	1:174:A:LEU:HA	1:177:A:ILE:HB	5	0.41
(2,79)	1:121:A:ILE:H	1:117:A:ALA:O	3	0.4
(1,600)	1:93:A:TRP:HB2	1:177:A:ILE:HG21	8	0.4

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	116	50.4	31	26.7	13.5	1	0.9	0.4
PHI	114	49.6	17	14.9	7.4	0	0.0	0.0
Total	230	100.0	48	20.9	20.9	1	0.4	0.4

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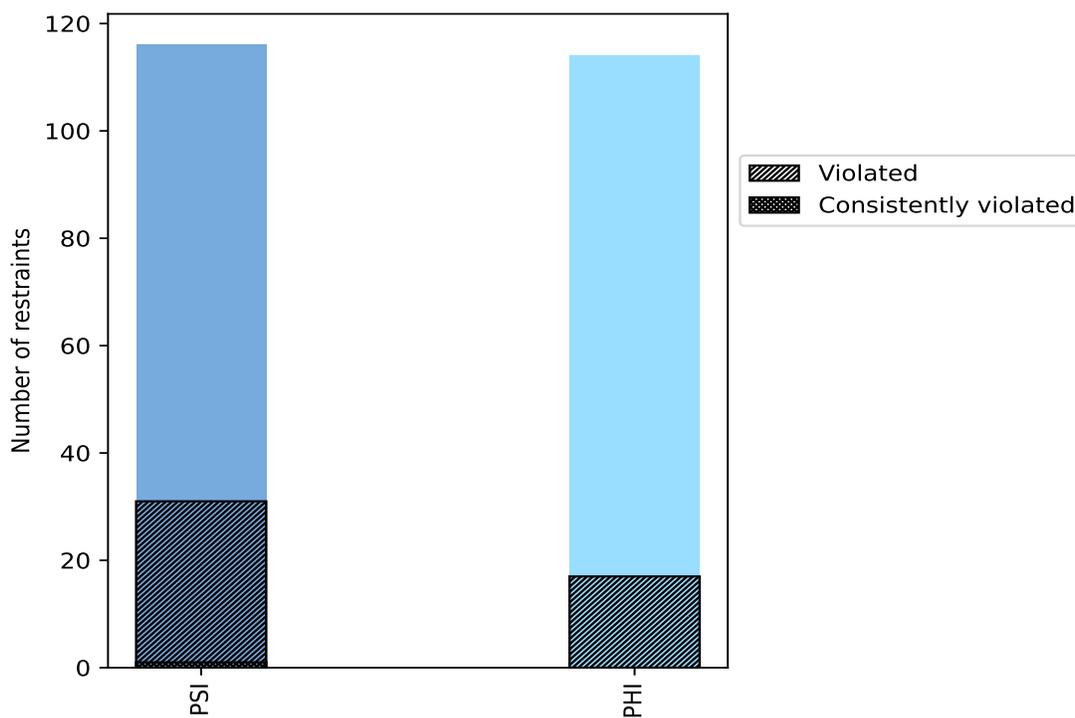
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

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	PHI	Total				
1	8	4	12	2.51	4.72	1.28	2.3
2	7	4	11	2.27	4.43	0.95	1.86
3	9	3	12	2.5	3.69	0.8	2.66
4	11	2	13	3.26	5.0	1.32	3.05
5	7	3	10	2.92	4.81	1.29	2.72
6	14	4	18	2.51	4.48	1.34	1.82
7	8	5	13	2.35	4.94	1.28	1.75
8	9	7	16	2.48	4.94	1.27	2.15
9	14	6	20	2.66	4.96	0.97	2.55
10	8	5	13	2.29	4.48	1.14	1.89

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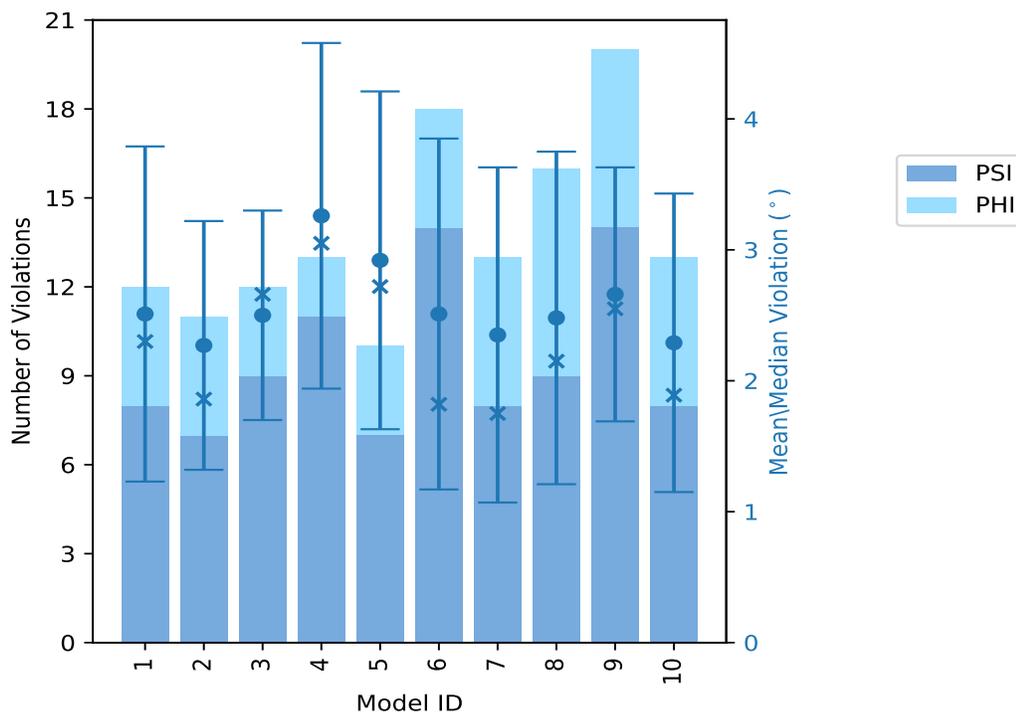
The mean(dot),median(x) and the standard deviation are shown in blue with respect to the y axis on the right

10.3 Dihedral-angle violation statistics for the ensemble (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]	%
11	6	17	1	10.0
3	7	10	2	20.0
8	1	9	3	30.0
3	1	4	4	40.0
2	0	2	5	50.0
0	0	0	6	60.0
2	1	3	7	70.0
1	0	1	8	80.0
0	1	1	9	90.0
1	0	1	10	100.0

[1] Number of models with violations

10.3.1 Bar graph : Dihedral-angle Violation statistics for the ensemble (i)
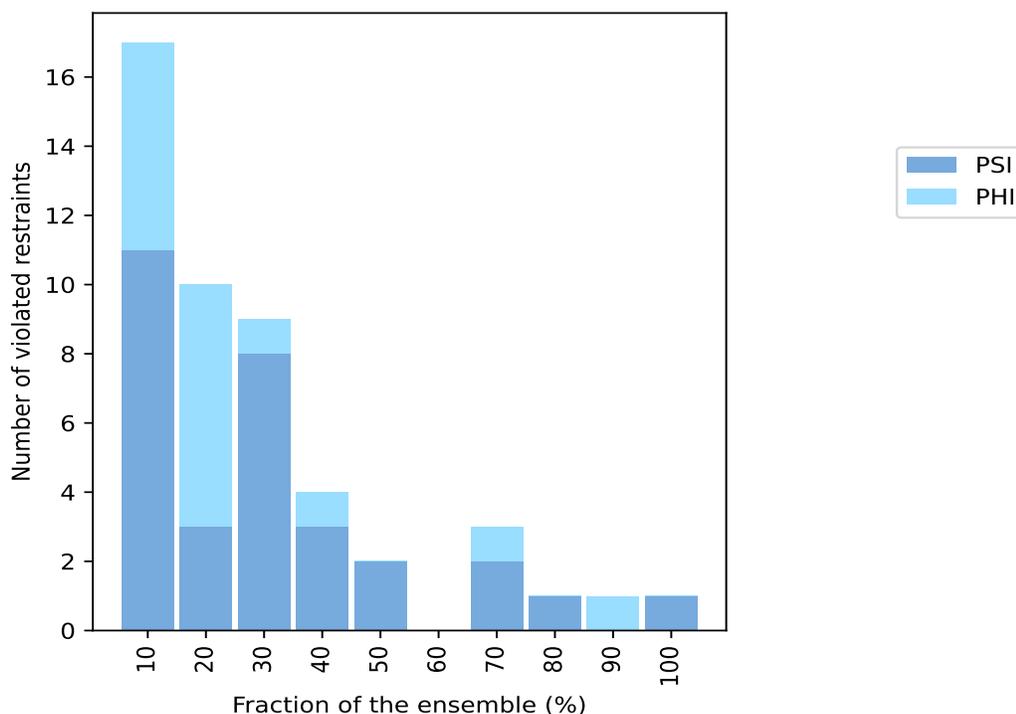
10.4 Most violated dihedral-angle restraints in the ensemble ⓘ

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
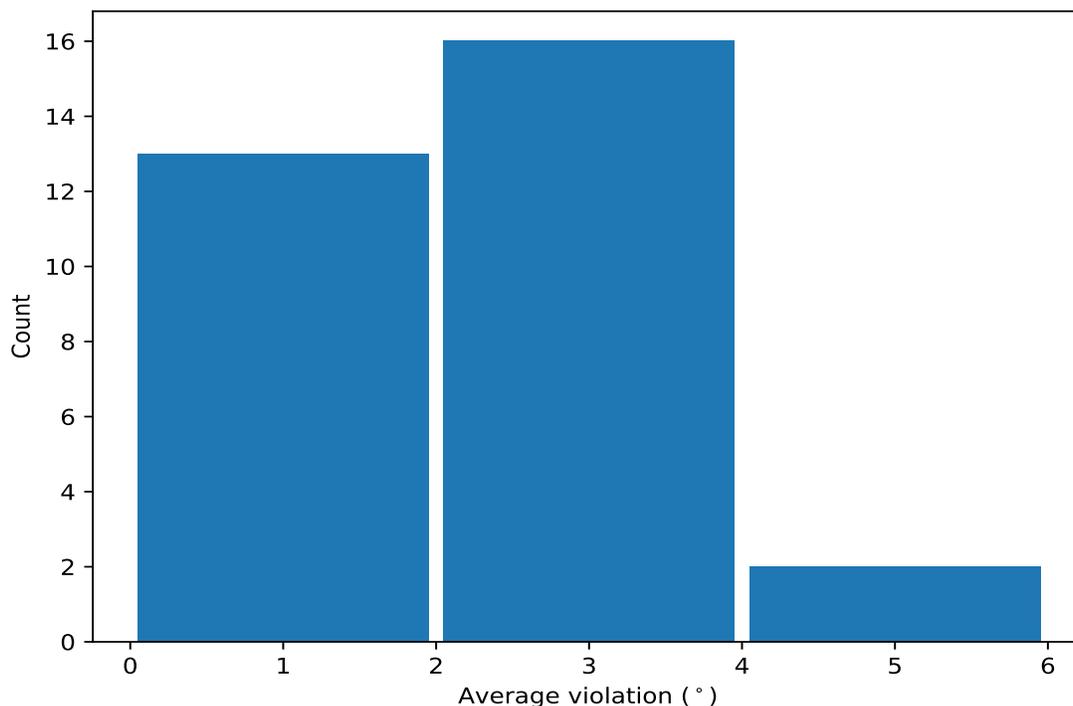
10.4.2 Table: Most violated dihedral-angle restraints ⓘ

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
(1,17)	1:20:A:TRP:N	1:20:A:TRP:CA	1:20:A:TRP:C	1:21:A:TYR:N	10	3.8	0.72	3.74
(1,146)	1:120:A:ALA:C	1:121:A:ILE:N	1:121:A:ILE:CA	1:121:A:ILE:C	9	4.0	0.59	4.06
(1,113)	1:94:A:TYR:N	1:94:A:TYR:CA	1:94:A:TYR:C	1:95:A:PHE:N	8	3.25	1.34	3.4
(1,112)	1:93:A:TRP:C	1:94:A:TYR:N	1:94:A:TYR:CA	1:94:A:TYR:C	7	4.29	0.38	4.39
(1,76)	1:62:A:PHE:N	1:62:A:PHE:CA	1:62:A:PHE:C	1:63:A:ASP:N	7	1.93	0.86	1.43
(1,29)	1:26:A:THR:N	1:26:A:THR:CA	1:26:A:THR:C	1:27:A:GLU:N	7	1.73	0.51	1.82
(1,60)	1:54:A:TYR:N	1:54:A:TYR:CA	1:54:A:TYR:C	1:55:A:VAL:N	5	3.52	1.36	3.91
(1,211)	1:159:A:SER:N	1:159:A:SER:CA	1:159:A:SER:C	1:160:A:ALA:N	5	3.32	1.35	4.09
(1,52)	1:50:A:LYS:N	1:50:A:LYS:CA	1:50:A:LYS:C	1:51:A:SER:N	4	2.87	1.02	2.54
(1,155)	1:128:A:ILE:N	1:128:A:ILE:CA	1:128:A:ILE:C	1:129:A:SER:N	4	2.29	0.95	2.31

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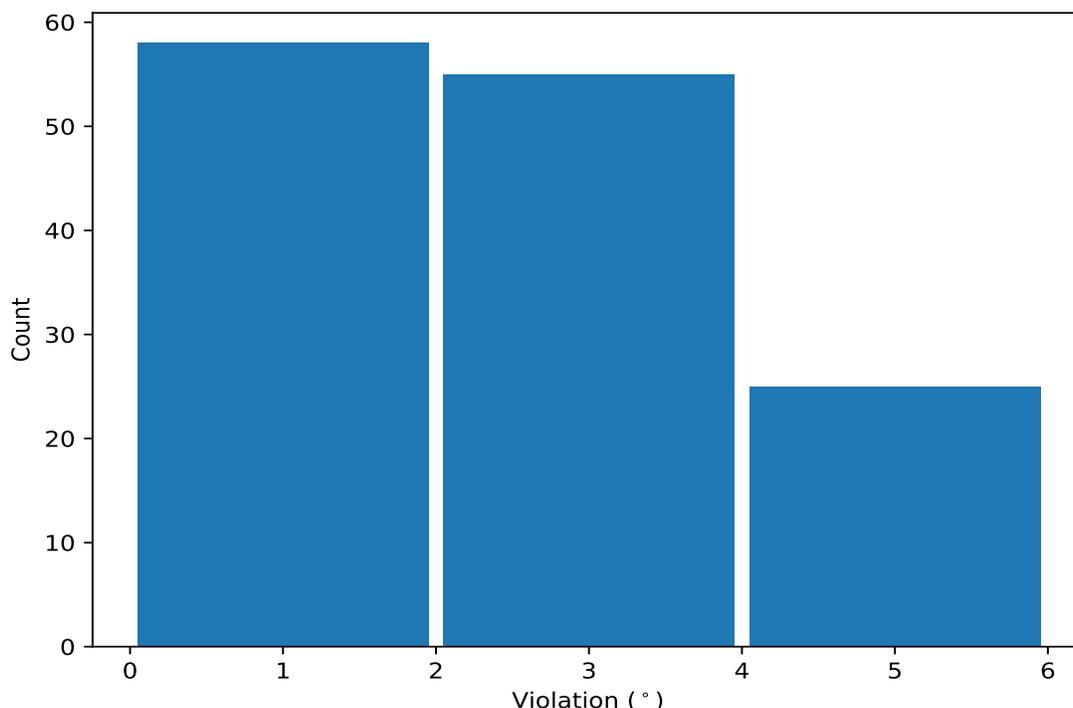
10.5.2 Table: All violated dihedral-angle restraints ⓘ

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 (°)
(1,60)	1:54:A:TYR:N	1:54:A:TYR:CA	1:54:A:TYR:C	1:55:A:VAL:N	4	5.0
(1,113)	1:94:A:TYR:N	1:94:A:TYR:CA	1:94:A:TYR:C	1:95:A:PHE:N	9	4.96
(1,112)	1:93:A:TRP:C	1:94:A:TYR:N	1:94:A:TYR:CA	1:94:A:TYR:C	8	4.94
(1,17)	1:20:A:TRP:N	1:20:A:TRP:CA	1:20:A:TRP:C	1:21:A:TYR:N	7	4.94
(1,17)	1:20:A:TRP:N	1:20:A:TRP:CA	1:20:A:TRP:C	1:21:A:TYR:N	5	4.81
(1,113)	1:94:A:TYR:N	1:94:A:TYR:CA	1:94:A:TYR:C	1:95:A:PHE:N	4	4.79
(1,111)	1:93:A:TRP:N	1:93:A:TRP:CA	1:93:A:TRP:C	1:94:A:TYR:N	4	4.74
(1,146)	1:120:A:ALA:C	1:121:A:ILE:N	1:121:A:ILE:CA	1:121:A:ILE:C	1	4.72
(1,211)	1:159:A:SER:N	1:159:A:SER:CA	1:159:A:SER:C	1:160:A:ALA:N	4	4.7
(1,146)	1:120:A:ALA:C	1:121:A:ILE:N	1:121:A:ILE:CA	1:121:A:ILE:C	7	4.65