wwPDB NMR Structure Validation Summary Report

Jun 3, 2023 – 06:23 PM EDT

PDB ID	:	2MRJ
BMRB ID	:	25081
Title	:	Structure of Fyn protein SH2 bound
Authors	:	Huculeci, R.; Buts, L.; Lenaerts, T.; Van Nuland, N.
Deposited on	:	2014-07-09

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

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

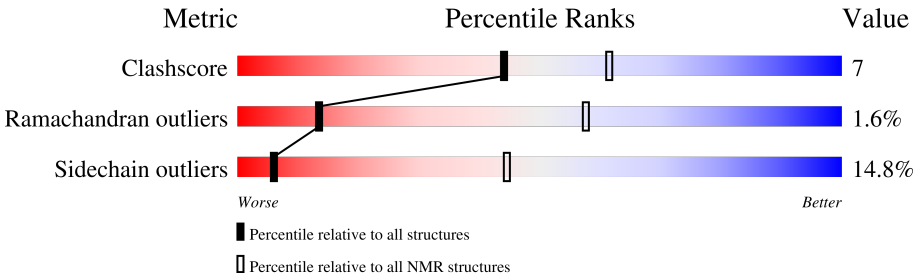
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal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

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment is 89%.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

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

Mol	Chain	Length	Quality of chain
1	A	112	60% 25% • • 11%

2 Ensemble composition and analysis

This entry contains 18 models. Model 17 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:149-A:244 (96)	0.71	17

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 5 clusters and 4 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1635 atoms, of which 812 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Tyrosine-protein kinase Fyn.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	100	1635	520	812	152	147	4	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	137	SER	-	expression tag	UNP P06241
A	138	SER	-	expression tag	UNP P06241
A	139	GLY	-	expression tag	UNP P06241
A	140	LEU	-	expression tag	UNP P06241
A	141	VAL	-	expression tag	UNP P06241
A	142	PRO	-	expression tag	UNP P06241
A	143	ARG	-	expression tag	UNP P06241
A	144	GLY	-	expression tag	UNP P06241
A	145	SER	-	expression tag	UNP P06241
A	146	HIS	-	expression tag	UNP P06241
A	147	MET	-	expression tag	UNP P06241

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: Tyrosine-protein kinase Fyn

Chain A:

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

The representative model is number 17. Colouring as in section 4.1 above.

- Molecule 1: Tyrosine-protein kinase Fyn

Chain A:

5 Refinement protocol and experimental data overview

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

Of the 100 calculated structures, 18 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
TALOS	geometry optimization	
CYANA	structure solution	
CNS	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	1386
Number of shifts mapped to atoms	1260
Number of unparsed shifts	0
Number of shifts with mapping errors	126
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	89%

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	0.92±0.07	0±0/808 (0.0± 0.1%)	0.76±0.15	0±1/1087 (0.0± 0.1%)
All	All	0.92	2/14544 (0.0%)	0.78	6/19566 (0.0%)

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	0.3±0.6
All	All	0	6

All unique bond 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(Å)	Ideal(Å)	Models	
								Worst	Total
1	A	200	VAL	C-N	15.10	1.68	1.34	18	1
1	A	199	HIS	C-N	13.77	1.65	1.34	18	1

5 of 6 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	200	VAL	CA-C-N	-22.17	68.43	117.20	18	1
1	A	199	HIS	O-C-N	-20.95	89.17	122.70	18	1
1	A	200	VAL	C-N-CA	-13.59	87.72	121.70	18	1
1	A	200	VAL	O-C-N	-12.88	102.10	122.70	18	1
1	A	199	HIS	CA-C-N	7.88	134.53	117.20	18	1

There are no chirality outliers.

All 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	234	ARG	Sidechain	2
1	A	176	ARG	Sidechain	1
1	A	206	ARG	Sidechain	1
1	A	156	ARG	Sidechain	1
1	A	161	ARG	Sidechain	1

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	790	780	779	11±4
All	All	14220	14040	14022	195

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 7.

5 of 92 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:200:VAL:C	1:A:201:LYS:N	1.45	1.68	18	1
1:A:200:VAL:CA	1:A:201:LYS:N	1.43	1.81	18	1
1:A:200:VAL:HA	1:A:201:LYS:N	1.30	1.33	18	1
1:A:200:VAL:C	1:A:201:LYS:CA	1.13	2.17	18	1
1:A:200:VAL:HA	1:A:201:LYS:H	0.90	1.14	18	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.

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	95/112 (85%)	86±2 (91±2%)	7±2 (8±2%)	2±1 (2±1%)	13 57
All	All	1710/2016 (85%)	1549 (91%)	133 (8%)	28 (2%)	13 57

All 3 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	155	GLY	18
1	A	198	ASP	7
1	A	199	HIS	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	82/96 (85%)	70±2 (85±3%)	12±2 (15±3%)	6 44
All	All	1476/1728 (85%)	1257 (85%)	219 (15%)	6 44

5 of 35 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	192	TRP	18
1	A	234	ARG	17
1	A	162	GLN	16
1	A	157	LYS	15
1	A	166	PHE	13

6.3.3 RNA (i)

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains (i)

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

6.5 Carbohydrates (i)

There are no monosaccharides in this entry.

6.6 Ligand geometry (i)

There are no ligands in this entry.

6.7 Other polymers (i)

There are no such molecules in this entry.

6.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	18-A	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
18	A	200:VAL	C	201:LYS	N	1.68
18	A	199:HIS	C	200:VAL	N	1.65

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 89% for the well-defined parts and 89% 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

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	1386
Number of shifts mapped to atoms	1260
Number of unparsed shifts	0
Number of shifts with mapping errors	126
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 126) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	137	SER	C	174.67	0.000	1
1	A	137	SER	CA	58.424	0.044	1
1	A	137	SER	CB	63.866	0.030	1
1	A	137	SER	HA	4.467	0.000	1
1	A	137	SER	HB2	3.851	0.004	1
1	A	137	SER	HB3	3.851	0.004	1
1	A	138	SER	C	174.992	0.000	1
1	A	138	SER	CA	58.731	0.061	1
1	A	138	SER	CB	63.862	0.017	1
1	A	138	SER	H	8.534	0.002	1
1	A	138	SER	HA	4.448	0.002	1
1	A	138	SER	HB2	3.894	0.001	1
1	A	138	SER	HB3	3.894	0.001	1
1	A	138	SER	N	118.235	0.022	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	139	GLY	C	173.845	0.000	1
1	A	139	GLY	CA	45.308	0.022	1
1	A	139	GLY	H	8.417	0.001	1
1	A	139	GLY	HA2	4.038	0.001	2
1	A	139	GLY	HA3	3.869	0.001	2
1	A	139	GLY	N	110.557	0.005	1
1	A	140	LEU	C	177.052	0.000	1
1	A	140	LEU	CA	55.149	0.025	1
1	A	140	LEU	CB	42.395	0.041	1
1	A	140	LEU	CD1	24.897	0.049	1
1	A	140	LEU	CD2	23.471	0.034	1
1	A	140	LEU	CG	26.975	0.070	1
1	A	140	LEU	H	8.074	0.001	1
1	A	140	LEU	HA	4.326	0.007	1
1	A	140	LEU	HB2	1.464	0.011	2
1	A	140	LEU	HB3	1.57	0.011	2
1	A	140	LEU	HD11	0.808	0.006	1
1	A	140	LEU	HD12	0.808	0.006	1
1	A	140	LEU	HD13	0.808	0.006	1
1	A	140	LEU	HD21	0.762	0.007	1
1	A	140	LEU	HD22	0.762	0.007	1
1	A	140	LEU	HD23	0.762	0.007	1
1	A	140	LEU	HG	1.509	0.002	1
1	A	140	LEU	N	121.573	0.008	1
1	A	141	VAL	CA	59.728	0.047	1
1	A	141	VAL	CB	32.903	0.065	1
1	A	141	VAL	CG1	21.0	0.076	1
1	A	141	VAL	CG2	20.447	0.033	1
1	A	141	VAL	H	8.02	0.002	1
1	A	141	VAL	HA	4.379	0.007	1
1	A	141	VAL	HB	2.0	0.006	1
1	A	141	VAL	HG11	0.879	0.003	1
1	A	141	VAL	HG12	0.879	0.003	1
1	A	141	VAL	HG13	0.879	0.003	1
1	A	141	VAL	HG21	0.874	0.001	1
1	A	141	VAL	HG22	0.874	0.001	1
1	A	141	VAL	HG23	0.874	0.001	1
1	A	141	VAL	N	122.512	0.002	1
1	A	142	PRO	C	176.786	0.000	1
1	A	142	PRO	CA	63.072	0.071	1
1	A	142	PRO	CB	31.93	0.004	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	142	PRO	CD	51.037	0.039	1
1	A	142	PRO	CG	27.476	0.039	1
1	A	142	PRO	HA	4.351	0.007	1
1	A	142	PRO	HB2	1.787	0.001	2
1	A	142	PRO	HB3	2.157	0.010	2
1	A	142	PRO	HD2	3.55	0.009	2
1	A	142	PRO	HD3	3.828	0.008	2
1	A	142	PRO	HG2	1.909	0.008	2
1	A	142	PRO	HG3	1.846	0.008	2
1	A	143	ARG	C	176.905	0.000	1
1	A	143	ARG	CA	56.505	0.038	1
1	A	143	ARG	CB	30.464	0.041	1
1	A	143	ARG	CD	43.246	0.056	1
1	A	143	ARG	CG	27.08	0.057	1
1	A	143	ARG	H	8.495	0.001	1
1	A	143	ARG	HA	4.218	0.006	1
1	A	143	ARG	HB2	1.772	0.005	2
1	A	143	ARG	HB3	1.831	0.004	2
1	A	143	ARG	HD2	3.126	0.001	1
1	A	143	ARG	HD3	3.126	0.001	1
1	A	143	ARG	HG2	1.643	0.005	2
1	A	143	ARG	HG3	1.58	0.002	2
1	A	143	ARG	N	121.547	0.004	1
1	A	144	GLY	C	173.933	0.000	1
1	A	144	GLY	CA	45.285	0.021	1
1	A	144	GLY	H	8.487	0.001	1
1	A	144	GLY	HA2	3.895	0.000	1
1	A	144	GLY	HA3	3.895	0.000	1
1	A	144	GLY	N	110.254	0.010	1
1	A	145	SER	C	174.295	0.000	1
1	A	145	SER	CA	58.585	0.016	1
1	A	145	SER	CB	63.744	0.043	1
1	A	145	SER	H	8.06	0.002	1
1	A	145	SER	HA	4.244	0.000	1
1	A	145	SER	HB2	3.679	0.000	1
1	A	145	SER	HB3	3.679	0.000	1
1	A	145	SER	N	115.029	0.006	1
1	A	146	HIS	C	174.687	0.000	1
1	A	146	HIS	CA	56.078	0.000	1
1	A	146	HIS	CB	29.461	0.000	1
1	A	146	HIS	H	8.379	0.003	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	146	HIS	HA	4.488	0.000	1
1	A	146	HIS	HB2	3.051	0.000	2
1	A	146	HIS	HB3	3.154	0.000	2
1	A	146	HIS	N	120.14	0.025	1
1	A	147	MET	C	177.111	0.000	1
1	A	147	MET	CA	54.312	0.016	1
1	A	147	MET	CB	31.956	0.039	1
1	A	147	MET	CE	17.126	0.078	1
1	A	147	MET	CG	31.916	0.012	1
1	A	147	MET	H	8.086	0.003	1
1	A	147	MET	HA	3.73	0.009	1
1	A	147	MET	HB2	1.105	0.008	2
1	A	147	MET	HB3	0.974	0.004	2
1	A	147	MET	HE1	1.829	0.003	1
1	A	147	MET	HE2	1.829	0.003	1
1	A	147	MET	HE3	1.829	0.003	1
1	A	147	MET	HG2	2.149	0.000	2
1	A	147	MET	HG3	1.981	0.000	2
1	A	147	MET	N	119.891	0.020	1
1	A	148	GLU	C	175.563	0.000	1
1	A	148	GLU	CA	59.298	0.035	1
1	A	148	GLU	CB	29.132	0.041	1
1	A	148	GLU	CG	36.418	0.024	1
1	A	148	GLU	H	9.08	0.002	1
1	A	148	GLU	HA	3.805	0.006	1
1	A	148	GLU	HB2	1.884	0.014	1
1	A	148	GLU	HB3	1.884	0.014	1
1	A	148	GLU	HG2	2.242	0.007	1
1	A	148	GLU	HG3	2.242	0.007	1
1	A	148	GLU	N	121.568	0.025	1

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	112	-0.36 ± 0.18	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	101	-0.04 ± 0.16	None needed (< 0.5 ppm)
$^{13}C'$	106	0.13 ± 0.21	None needed (< 0.5 ppm)
^{15}N	106	-0.65 ± 0.37	None needed (imprecise)

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 89%, i.e. 1214 atoms were assigned a chemical shift out of a possible 1365. 0 out of 14 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	479/487 (98%)	198/200 (99%)	188/192 (98%)	93/95 (98%)
Sidechain	621/739 (84%)	422/475 (89%)	192/223 (86%)	7/41 (17%)
Aromatic	114/139 (82%)	57/68 (84%)	55/66 (83%)	2/5 (40%)
Overall	1214/1365 (89%)	677/743 (91%)	435/481 (90%)	102/141 (72%)

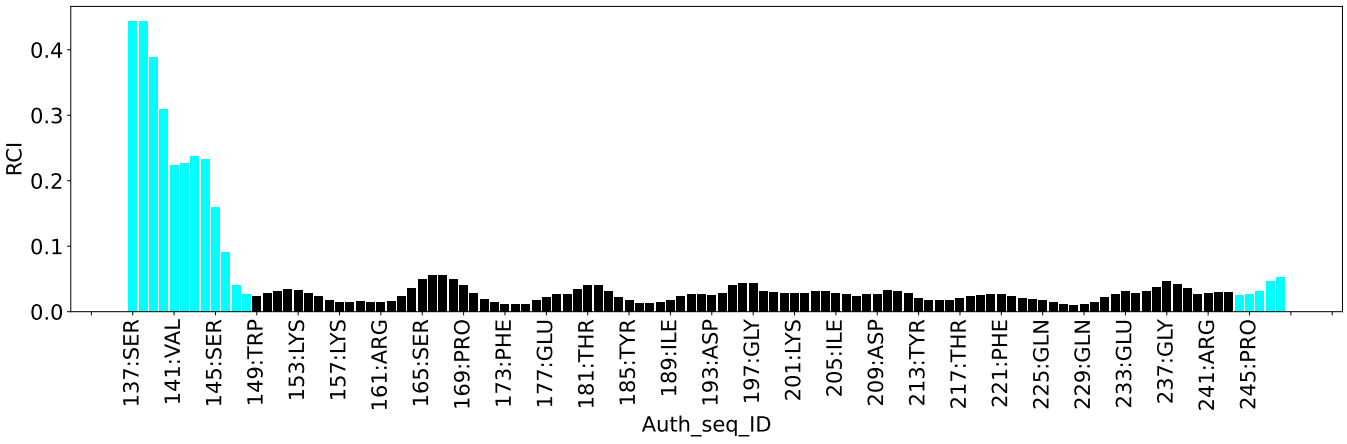
7.1.4 Statistically unusual chemical shifts

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	2395
Intra-residue (|i-j|=0)	617
Sequential (|i-j|=1)	768
Medium range (|i-j|>1 and |i-j|<5)	365
Long range (|i-j|≥5)	645
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	164
Number of unmapped restraints	162
Number of restraints per residue	22.8
Number of long range restraints per residue[1]	5.8

[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)	17.2	0.2
0.2-0.5 (Medium)	3.6	0.5
>0.5 (Large)	2.8	1.76

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)	14.8	9.6
10.0-20.0 (Medium)	0.2	15.9
>20.0 (Large)	0.1	26.3

9 Distance violation analysis (i)

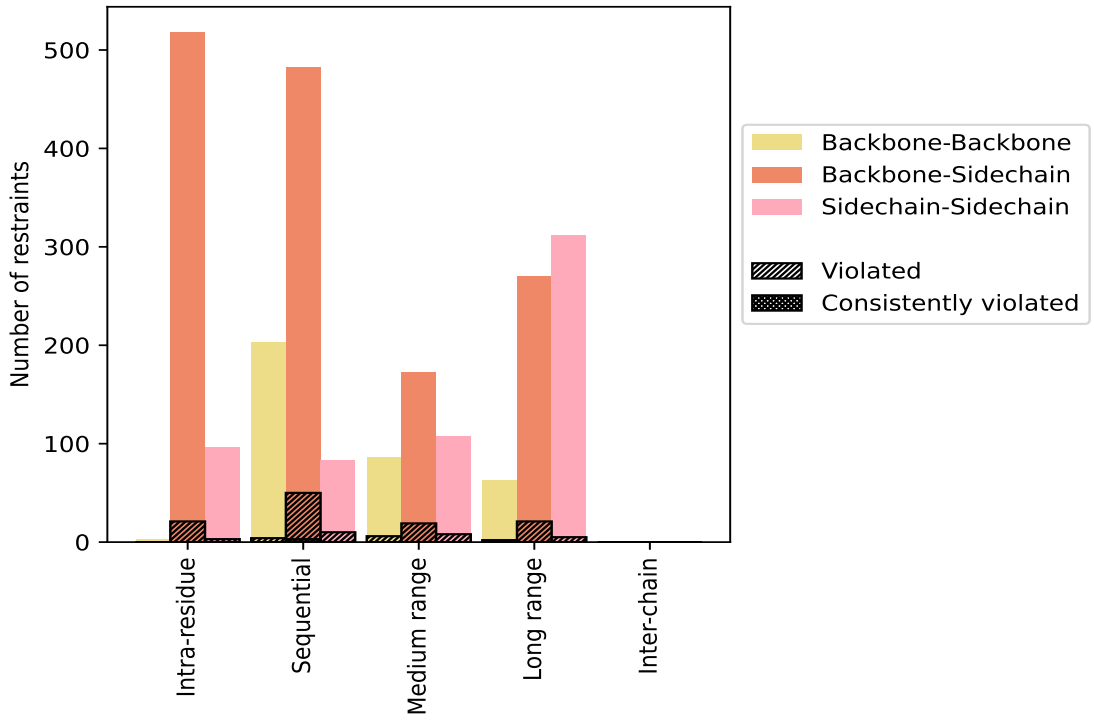
9.1 Summary of distance violations (i)

The following table shows the summary of distance violations in different restraint categories based on the sequence separation of the atoms involved. Each category is further sub-divided into three sub-categories based on the atoms involved. Violations less than 0.1 Å are not included in the statistics.

Restraints type	Count	%[1]	Violated[3]			Consistently Violated[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
Intra-residue (|i-j|=0)	617	25.8	24	3.9	1.0	0	0.0	0.0
Backbone-Backbone	3	0.1	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	518	21.6	21	4.1	0.9	0	0.0	0.0
Sidechain-Sidechain	96	4.0	3	3.1	0.1	0	0.0	0.0
Sequential (|i-j|=1)	768	32.1	64	8.3	2.7	3	0.4	0.1
Backbone-Backbone	203	8.5	4	2.0	0.2	0	0.0	0.0
Backbone-Sidechain	482	20.1	50	10.4	2.1	3	0.6	0.1
Sidechain-Sidechain	83	3.5	10	12.0	0.4	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	365	15.2	33	9.0	1.4	0	0.0	0.0
Backbone-Backbone	86	3.6	6	7.0	0.3	0	0.0	0.0
Backbone-Sidechain	172	7.2	19	11.0	0.8	0	0.0	0.0
Sidechain-Sidechain	107	4.5	8	7.5	0.3	0	0.0	0.0
Long range (|i-j|≥5)	645	26.9	28	4.3	1.2	0	0.0	0.0
Backbone-Backbone	63	2.6	2	3.2	0.1	0	0.0	0.0
Backbone-Sidechain	270	11.3	21	7.8	0.9	0	0.0	0.0
Sidechain-Sidechain	312	13.0	5	1.6	0.2	0	0.0	0.0
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	0	0.0	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	2395	100.0	149	6.2	6.2	3	0.1	0.1
Backbone-Backbone	355	14.8	12	3.4	0.5	0	0.0	0.0
Backbone-Sidechain	1442	60.2	111	7.7	4.6	3	0.2	0.1
Sidechain-Sidechain	598	25.0	26	4.3	1.1	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	4	15	3	0	0	22	0.24	1.31	0.28	0.13
2	4	12	2	4	0	22	0.23	1.29	0.26	0.13
3	2	14	6	0	0	22	0.23	1.32	0.27	0.14
4	5	17	3	2	0	27	0.22	1.1	0.21	0.14
5	3	8	2	3	0	16	0.28	1.24	0.3	0.14
6	2	13	2	3	0	20	0.26	1.28	0.27	0.16
7	4	15	3	2	0	24	0.21	1.26	0.24	0.14
8	6	13	4	4	0	27	0.24	1.28	0.24	0.16
9	4	13	4	0	0	21	0.24	1.23	0.26	0.14
10	2	13	3	3	0	21	0.23	1.28	0.26	0.13
11	2	9	1	1	0	13	0.29	1.3	0.32	0.16

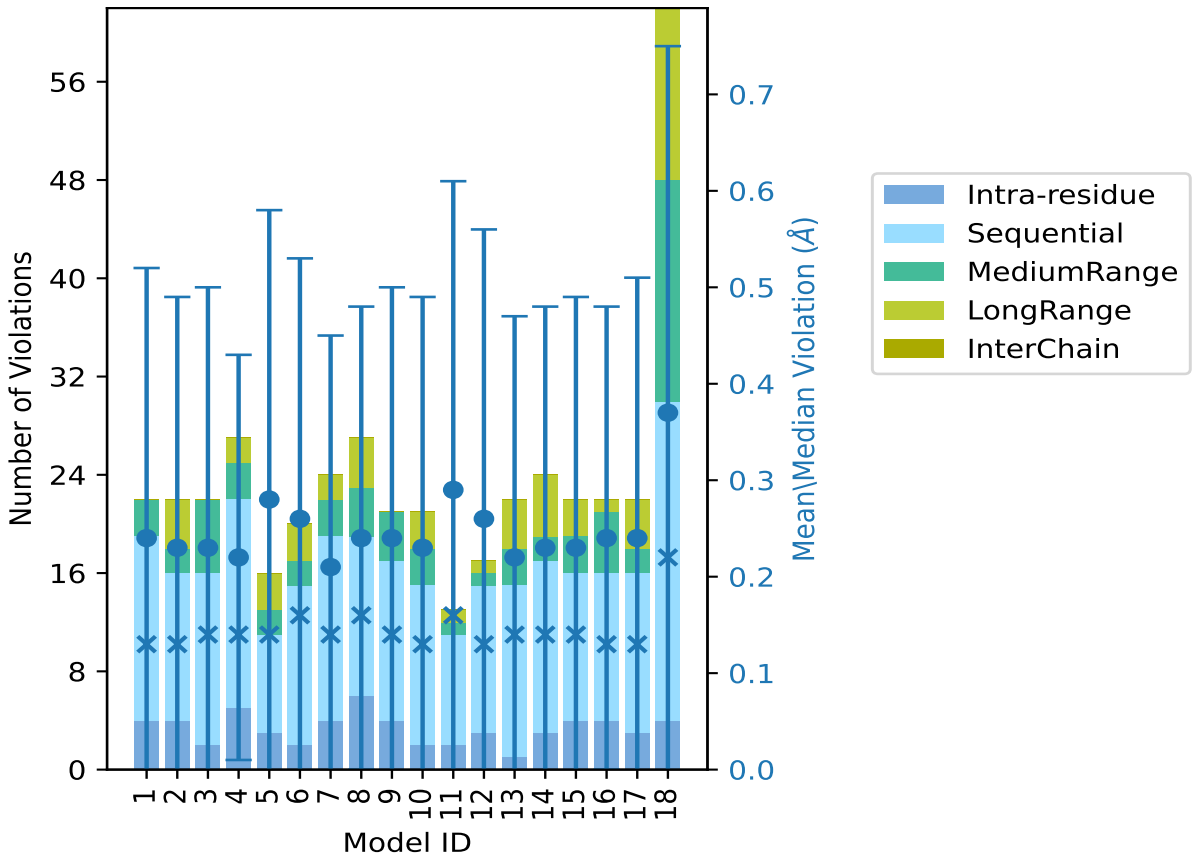
Continued on next page...

Continued from previous page...

Model ID	Number of violations						Mean (Å)	Max (Å)	SD[6] (Å)	Median (Å)
	IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total				
12	3	12	1	1	0	17	0.26	1.25	0.3	0.13
13	1	14	3	4	0	22	0.22	1.27	0.25	0.14
14	3	14	2	5	0	24	0.23	1.28	0.25	0.14
15	4	12	3	3	0	22	0.23	1.3	0.26	0.14
16	4	12	5	1	0	22	0.24	1.0	0.24	0.13
17	3	13	2	4	0	22	0.24	1.31	0.27	0.13
18	4	26	18	14	0	62	0.37	1.76	0.38	0.22

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

9.2.1 Bar graph : Distance Violation statistics for each model



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

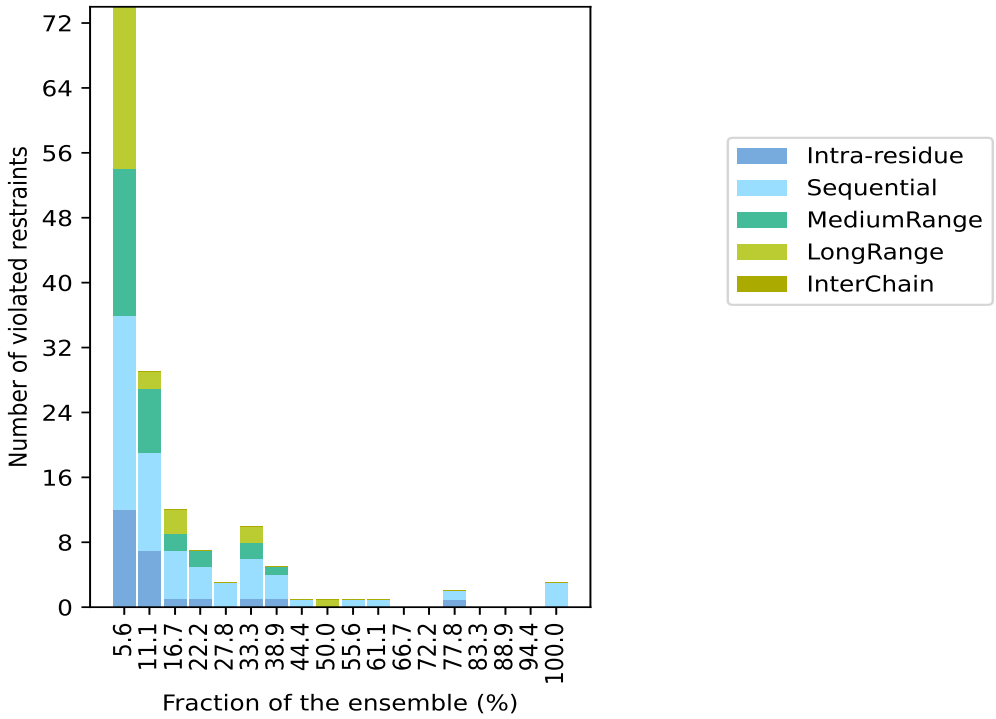
9.3 Distance violation statistics for the ensemble (i)

Violation analysis may find that some restraints are violated in few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of the ensemble. In total, 2246(IR:593, SQ:704, MR:332, LR:617, 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]	%
12	24	18	20	0	74	1	5.6
7	12	8	2	0	29	2	11.1
1	6	2	3	0	12	3	16.7
1	4	2	0	0	7	4	22.2
0	3	0	0	0	3	5	27.8
1	5	2	2	0	10	6	33.3
1	3	1	0	0	5	7	38.9
0	1	0	0	0	1	8	44.4
0	0	0	1	0	1	9	50.0
0	1	0	0	0	1	10	55.6
0	1	0	0	0	1	11	61.1
0	0	0	0	0	0	12	66.7
0	0	0	0	0	0	13	72.2
1	1	0	0	0	2	14	77.8
0	0	0	0	0	0	15	83.3
0	0	0	0	0	0	16	88.9
0	0	0	0	0	0	17	94.4
0	3	0	0	0	3	18	100.0

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

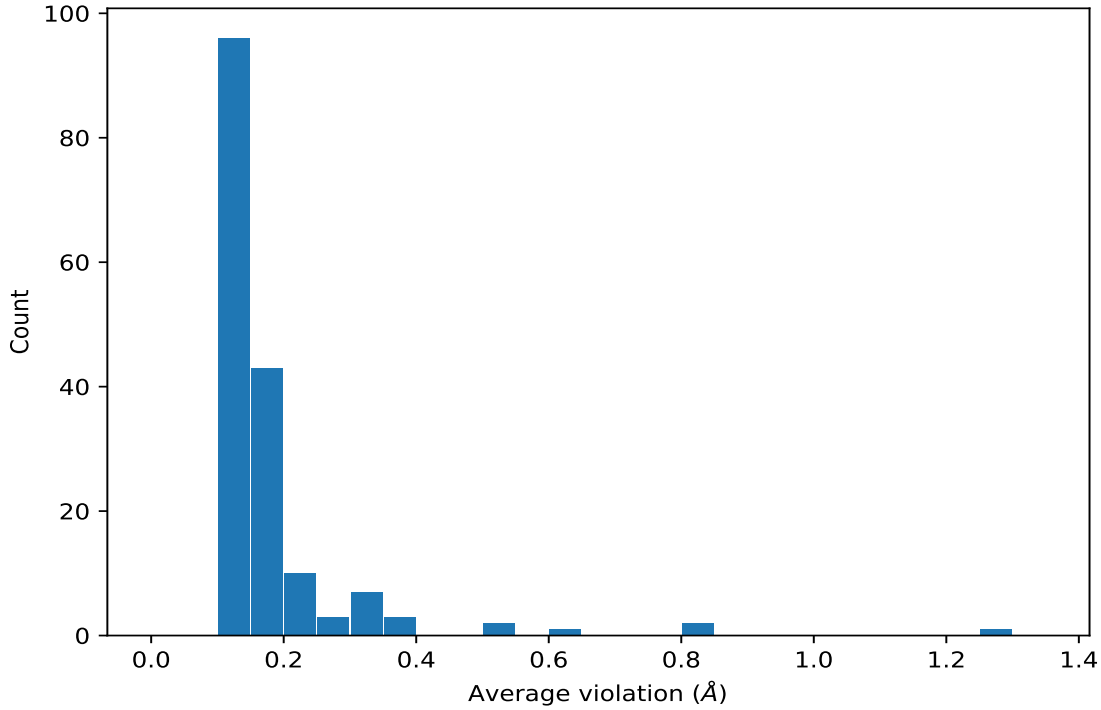
9.3.1 Bar graph : Distance violation statistics for the ensemble



9.4 Most violated distance restraints in the ensemble

9.4.1 Histogram : Distribution of mean distance violations

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



9.4.2 Table: Most violated distance restraints

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

Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,665)	1:A:245:PRO:HD3	1:A:246:CYS:H	18	1.26	0.08	1.28
(1,411)	1:A:245:PRO:HD2	1:A:246:CYS:H	18	0.64	0.08	0.63
(1,2384)	1:A:245:PRO:HG2	1:A:246:CYS:H	18	0.5	0.09	0.48
(1,2384)	1:A:245:PRO:HG3	1:A:246:CYS:H	18	0.5	0.09	0.48
(1,219)	1:A:234:ARG:HG2	1:A:235:ALA:H	14	0.13	0.02	0.13
(1,581)	1:A:226:GLN:HB3	1:A:226:GLN:HE21	14	0.12	0.01	0.12
(1,1613)	1:A:192:TRP:HZ3	1:A:193:ASP:HA	11	0.17	0.04	0.17
(1,457)	1:A:221:PHE:H	1:A:222:GLU:HG2	10	0.12	0.01	0.12
(1,354)	1:A:192:TRP:HE1	1:A:198:ASP:H	9	0.13	0.02	0.13
(1,2379)	1:A:244:VAL:HG11	1:A:245:PRO:HG2	8	0.18	0.03	0.18
(1,2379)	1:A:244:VAL:HG11	1:A:245:PRO:HG3	8	0.18	0.03	0.18
(1,2379)	1:A:244:VAL:HG12	1:A:245:PRO:HG2	8	0.18	0.03	0.18
(1,2379)	1:A:244:VAL:HG12	1:A:245:PRO:HG3	8	0.18	0.03	0.18
(1,2379)	1:A:244:VAL:HG13	1:A:245:PRO:HG2	8	0.18	0.03	0.18
(1,2379)	1:A:244:VAL:HG13	1:A:245:PRO:HG3	8	0.18	0.03	0.18
(1,2379)	1:A:244:VAL:HG21	1:A:245:PRO:HG2	8	0.18	0.03	0.18

Continued on next page...

Continued from previous page...

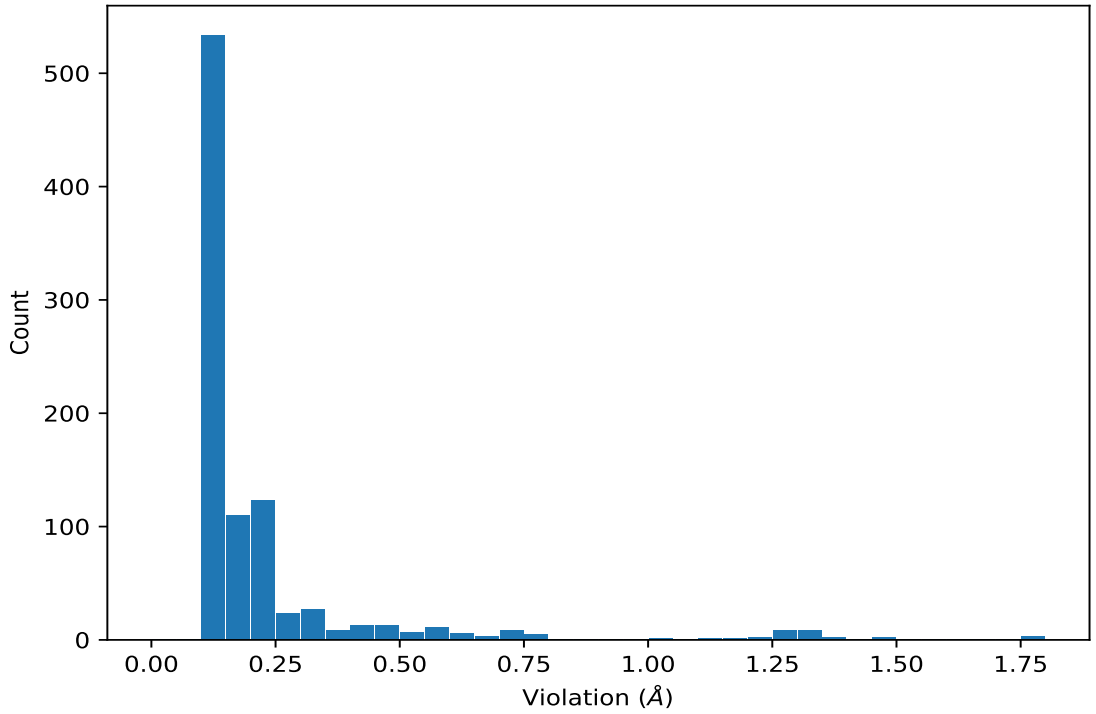
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,2379)	1:A:244:VAL:HG21	1:A:245:PRO:HG3	8	0.18	0.03	0.18
(1,2379)	1:A:244:VAL:HG22	1:A:245:PRO:HG2	8	0.18	0.03	0.18
(1,2379)	1:A:244:VAL:HG22	1:A:245:PRO:HG3	8	0.18	0.03	0.18
(1,2379)	1:A:244:VAL:HG23	1:A:245:PRO:HG2	8	0.18	0.03	0.18
(1,2379)	1:A:244:VAL:HG23	1:A:245:PRO:HG3	8	0.18	0.03	0.18
(1,128)	1:A:244:VAL:HB	1:A:245:PRO:HD2	7	0.24	0.03	0.24

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

9.5 All violated distance restraints (i)

9.5.1 Histogram : Distribution of distance violations (i)

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



9.5.2 Table : All distance violations (i)

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

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,1521)	1:A:164:LEU:HD11	1:A:199:HIS:HA	18	1.76
(1,1521)	1:A:164:LEU:HD12	1:A:199:HIS:HA	18	1.76
(1,1521)	1:A:164:LEU:HD13	1:A:199:HIS:HA	18	1.76
(1,1873)	1:A:156:ARG:HA	1:A:157:LYS:HE2	18	1.49
(1,1873)	1:A:156:ARG:HA	1:A:157:LYS:HE3	18	1.49
(1,1886)	1:A:157:LYS:HD2	1:A:158:ASP:H	18	1.35
(1,1886)	1:A:157:LYS:HD3	1:A:158:ASP:H	18	1.35
(1,1879)	1:A:157:LYS:H	1:A:157:LYS:HD2	18	1.34
(1,1879)	1:A:157:LYS:H	1:A:157:LYS:HD3	18	1.34
(1,665)	1:A:245:PRO:HD3	1:A:246:CYS:H	3	1.32
(1,665)	1:A:245:PRO:HD3	1:A:246:CYS:H	1	1.31
(1,665)	1:A:245:PRO:HD3	1:A:246:CYS:H	17	1.31
(1,665)	1:A:245:PRO:HD3	1:A:246:CYS:H	11	1.3
(1,665)	1:A:245:PRO:HD3	1:A:246:CYS:H	15	1.3
(1,665)	1:A:245:PRO:HD3	1:A:246:CYS:H	18	1.3

10 Dihedral-angle violation analysis

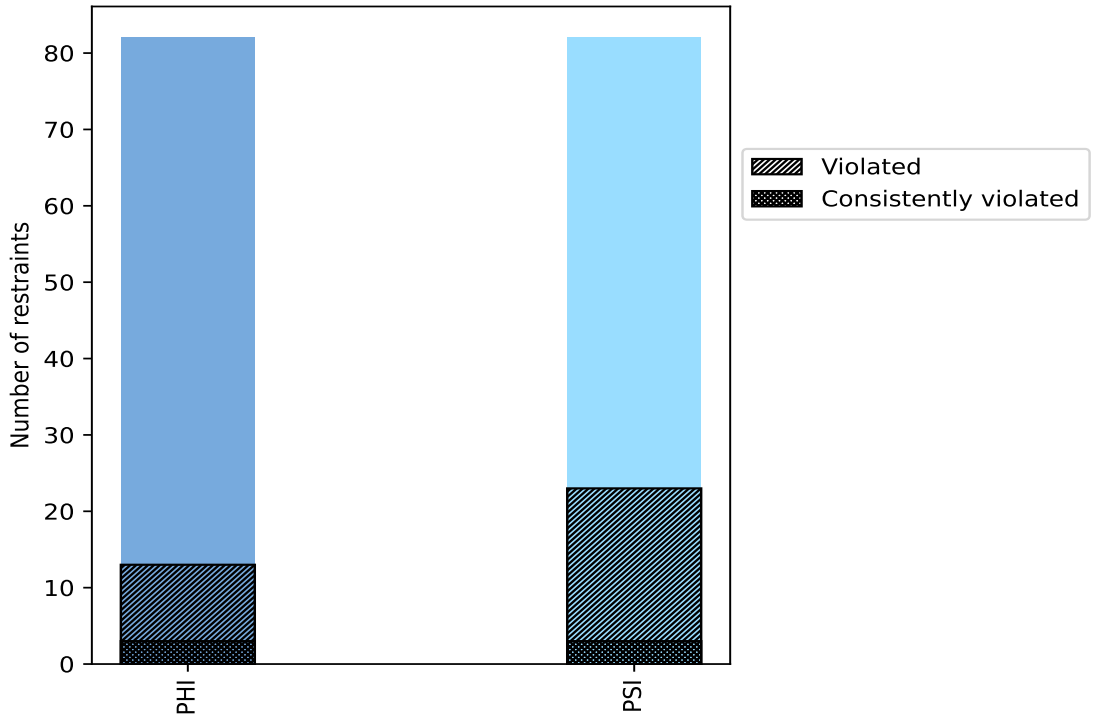
10.1 Summary of dihedral-angle violations

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]
PHI	82	50.0	13	15.9	7.9	3	3.7	1.8
PSI	82	50.0	23	28.0	14.0	3	3.7	1.8
Total	164	100.0	36	22.0	22.0	6	3.7	3.7

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

10.1.1 Bar chart : Distribution of dihedral-angles and violations



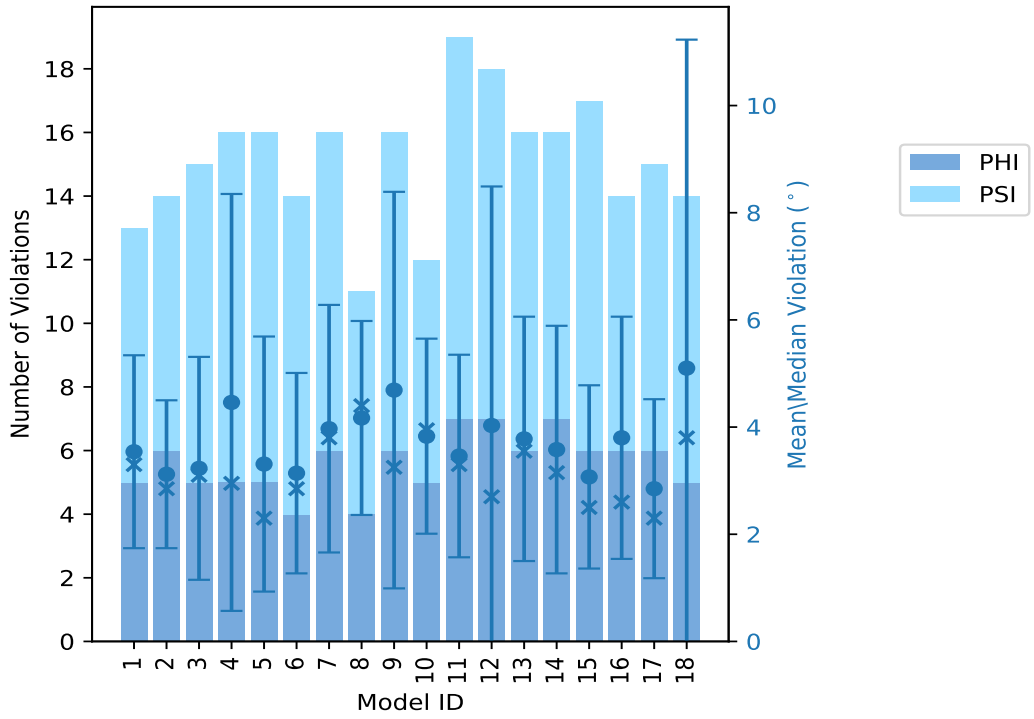
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

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

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

Model ID	Number of violations			Mean (°)	Max (°)	SD (°)	Median (°)
	PHI	PSI	Total				
1	5	8	13	3.54	6.5	1.8	3.3
2	6	8	14	3.12	6.2	1.38	2.85
3	5	10	15	3.23	7.9	2.08	3.1
4	5	11	16	4.46	15.9	3.89	2.95
5	5	11	16	3.31	8.5	2.38	2.3
6	4	10	14	3.14	7.8	1.87	2.85
7	6	10	16	3.97	8.9	2.31	3.8
8	4	7	11	4.17	6.5	1.81	4.4
9	6	10	16	4.69	15.2	3.7	3.25
10	5	7	12	3.83	6.7	1.82	3.95
11	7	12	19	3.46	6.7	1.89	3.3
12	7	11	18	4.03	20.9	4.46	2.7
13	6	10	16	3.78	7.5	2.28	3.55
14	7	9	16	3.58	9.5	2.31	3.15
15	6	11	17	3.07	6.9	1.71	2.5
16	6	8	14	3.8	8.0	2.26	2.6
17	6	9	15	2.85	6.1	1.67	2.3
18	5	9	14	5.1	26.3	6.13	3.8

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	
PHI	PSI	Total	Count[1]	%
0	6	6	1	5.6
4	1	5	2	11.1
2	2	4	3	16.7
0	3	3	4	22.2
2	0	2	5	27.8
1	1	2	6	33.3
0	0	0	7	38.9
0	1	1	8	44.4
0	1	1	9	50.0
0	1	1	10	55.6
0	1	1	11	61.1

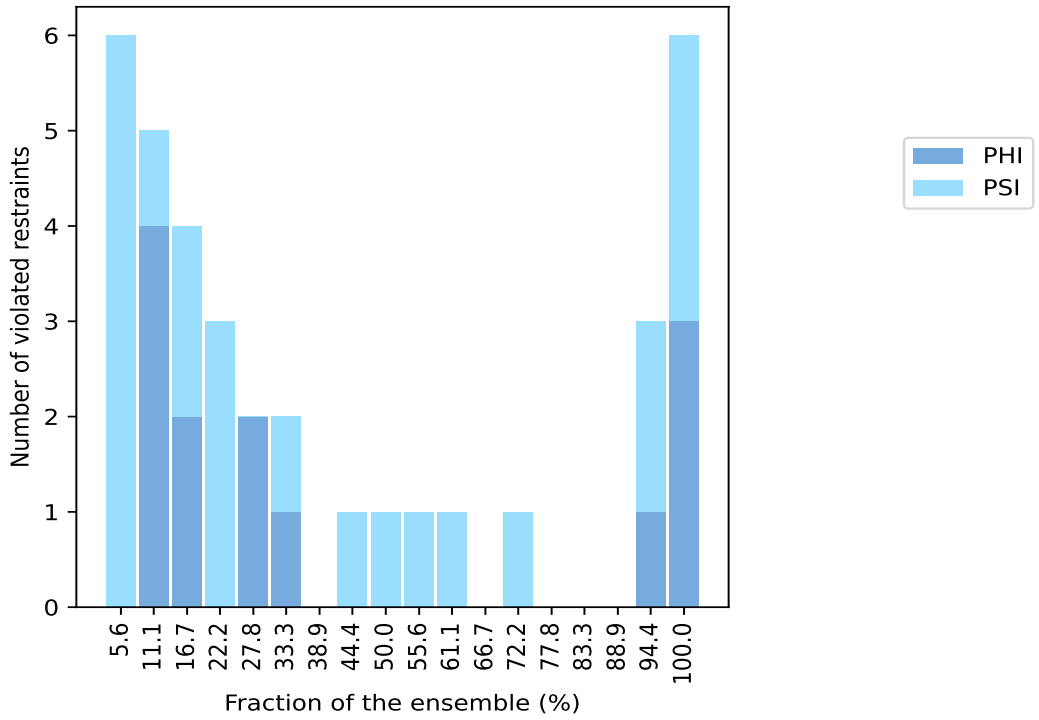
Continued on next page...

Continued from previous page...

Number of violated restraints			Fraction of the ensemble	
PHI	PSI	Total	Count[1]	%
0	0	0	12	66.7
0	1	1	13	72.2
0	0	0	14	77.8
0	0	0	15	83.3
0	0	0	16	88.9
1	2	3	17	94.4
3	3	6	18	100.0

[1] Number of models with violations

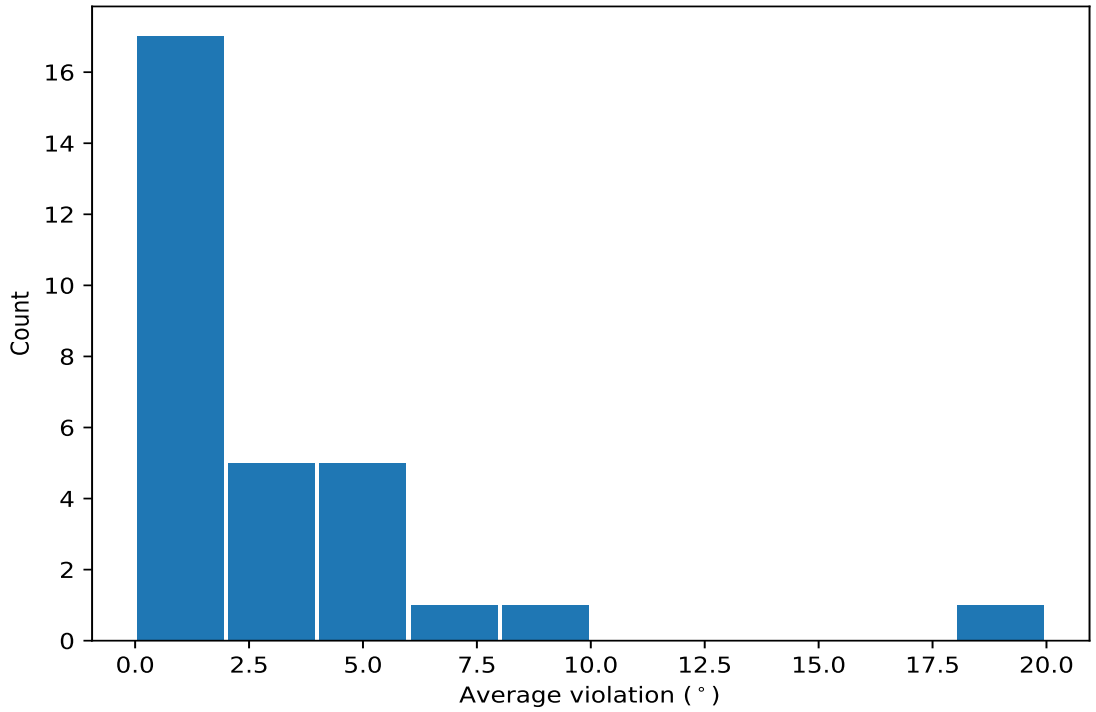
10.3.1 Bar graph : Dihedral-angle Violation statistics for the ensemble (i)



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

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

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



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

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

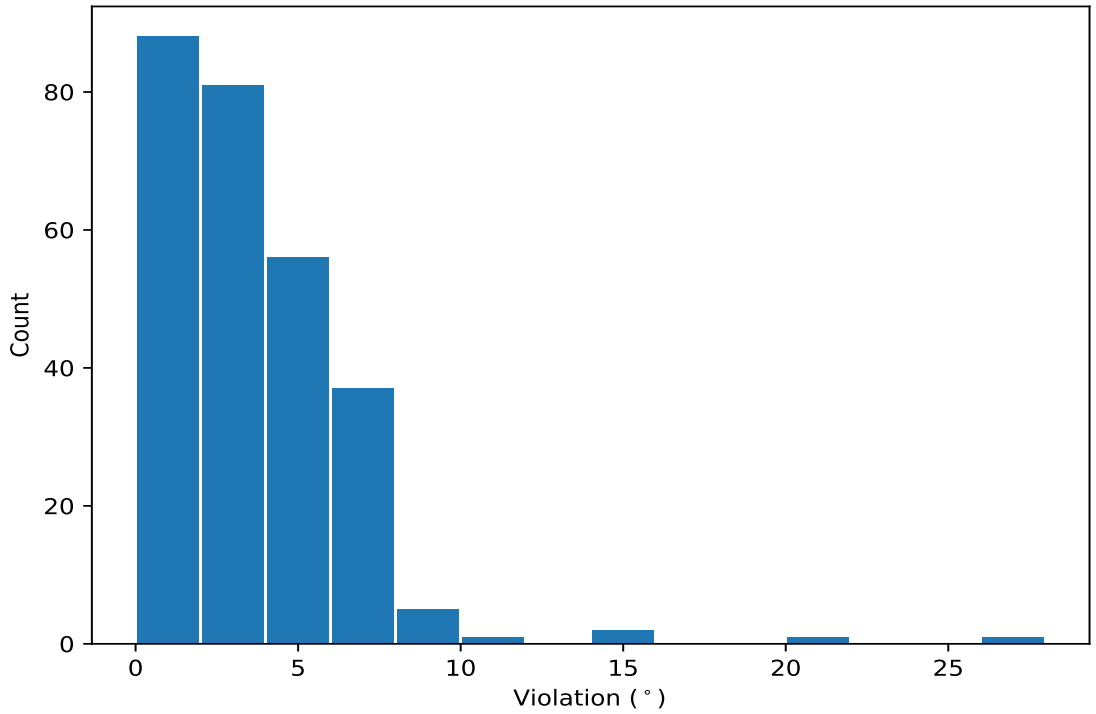
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models[1]	Mean	SD[2]	Median
(1,2)	1:A:149:TRP:N	1:A:149:TRP:CA	1:A:149:TRP:C	1:A:150:TYR:N	18	6.98	1.3	6.8
(1,3)	1:A:149:TRP:C	1:A:150:TYR:N	1:A:150:TYR:CA	1:A:150:TYR:C	18	5.99	0.92	6.05
(1,13)	1:A:156:ARG:C	1:A:157:LYS:N	1:A:157:LYS:CA	1:A:157:LYS:C	18	5.83	5.19	4.85
(1,78)	1:A:192:TRP:N	1:A:192:TRP:CA	1:A:192:TRP:C	1:A:193:ASP:N	18	5.39	1.25	5.1
(1,151)	1:A:236:ALA:C	1:A:237:GLY:N	1:A:237:GLY:CA	1:A:237:GLY:C	18	3.64	0.53	3.65
(1,150)	1:A:236:ALA:N	1:A:236:ALA:CA	1:A:236:ALA:C	1:A:237:GLY:N	18	3.2	0.7	3.3
(1,12)	1:A:156:ARG:N	1:A:156:ARG:CA	1:A:156:ARG:C	1:A:157:LYS:N	17	4.08	1.18	4.0
(1,11)	1:A:155:GLY:C	1:A:156:ARG:N	1:A:156:ARG:CA	1:A:156:ARG:C	17	2.79	0.71	2.5
(1,34)	1:A:169:PRO:N	1:A:169:PRO:CA	1:A:169:PRO:C	1:A:170:ARG:N	17	2.26	1.22	1.8
(1,160)	1:A:245:PRO:N	1:A:245:PRO:CA	1:A:245:PRO:C	1:A:246:CYS:N	13	5.76	2.32	5.6

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

10.5 All violated dihedral-angle restraints (i)

10.5.1 Histogram : Distribution of violations (i)

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



10.5.2 Table: All violated dihedral-angle restraints

The following table provides the list of violationsfor the 10 worst performing restraints, sorted by the violation value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint.

Key	Atom-1	Atom-2	Atom-3	Atom-4	Model ID	Violation (°)
(1,13)	1:A:156:ARG:C	1:A:157:LYS:N	1:A:157:LYS:CA	1:A:157:LYS:C	18	26.3
(1,161)	1:A:245:PRO:C	1:A:246:CYS:N	1:A:246:CYS:CA	1:A:246:CYS:C	12	20.9
(1,162)	1:A:246:CYS:N	1:A:246:CYS:CA	1:A:246:CYS:C	1:A:247:HIS:N	4	15.9
(1,161)	1:A:245:PRO:C	1:A:246:CYS:N	1:A:246:CYS:CA	1:A:246:CYS:C	9	15.2
(1,160)	1:A:245:PRO:N	1:A:245:PRO:CA	1:A:245:PRO:C	1:A:246:CYS:N	4	10.4
(1,2)	1:A:149:TRP:N	1:A:149:TRP:CA	1:A:149:TRP:C	1:A:150:TYR:N	9	9.6
(1,2)	1:A:149:TRP:N	1:A:149:TRP:CA	1:A:149:TRP:C	1:A:150:TYR:N	14	9.5
(1,160)	1:A:245:PRO:N	1:A:245:PRO:CA	1:A:245:PRO:C	1:A:246:CYS:N	7	8.9
(1,2)	1:A:149:TRP:N	1:A:149:TRP:CA	1:A:149:TRP:C	1:A:150:TYR:N	5	8.5
(1,78)	1:A:192:TRP:N	1:A:192:TRP:CA	1:A:192:TRP:C	1:A:193:ASP:N	16	8.0