wwPDB NMR Structure Validation Summary Report (i)

Jun 5, 2023 – 08:22 PM EDT

PDB ID	:	2M68
BMRB ID	:	19117
Title	:	NMR solution structure ensemble of 3-4D mutant domain 11 IGF2R in complex with IGF2 (domain 11 structure only)
Authors	:	Strickland, M.; Williams, C.; Richards, E.; Minnall, L.; Crump, M.P.; Frago, S.; Hughes, J.; Garner, L.; Hoppe, H.; Rezgui, D.; Zaccheo, O.J.; Prince, S.N.; Hassan, A.B.; Whittaker, S.
Deposited on	:	2013-03-27

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:

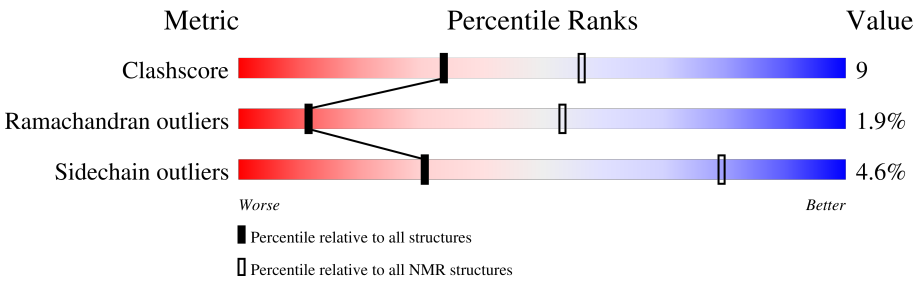
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal 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 90%.

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	142	73% 17% • 9%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 20 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *fewest violations*.

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:1516-A:1598, A:1604-A:1649 (129)	0.55	20

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

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

The models can be grouped into 3 clusters and 5 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2129 atoms, of which 1051 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Insulin-like growth factor 2 receptor variant.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	142	2129	676	1051	187	204	11	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1543	ALA	SER	engineered mutation	UNP Q59EZ3
A	1544	LYS	GLU	engineered mutation	UNP Q59EZ3
A	1545	GLY	LYS	engineered mutation	UNP Q59EZ3
A	1546	TRP	GLY	engineered mutation	UNP Q59EZ3
A	1547	GLY	LEU	engineered mutation	UNP Q59EZ3
A	1648	LEU	GLN	cloning artifact	UNP Q59EZ3

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

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

• Molecule 1: Insulin-like growth factor 2 receptor variant

Chain A: 73% 17% • 9%

M1508 K1509 S1510 N1511 E1512 H1513 D1514 D1515 C1516 Q1517 N1520 F1527 T1539 A1540 V1548 C1553 N1556 C1559 P1560 V1563 R1571 L1581 L1588 K1593 P1599 S1600 K1601 S1602 G1603 L1604 S1605 Y1606 K1607 F1612 P1624 L1629 D1630 K1631 L1636 F1637 F1638 E1647

L1648 A1649

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

The representative model is number 20. Colouring as in section 4.1 above.

• Molecule 1: Insulin-like growth factor 2 receptor variant

Chain A: 68% 22% • 9%

M1508 K1509 S1510 N1511 E1512 H1513 D1514 D1515 C1516 Q1517 N1520 L1526 F1527 S1533 A1540 V1548 Y1549 I1552 N1556 C1559 V1563 C1566 V1574 N1578 L1588 K1593 P1599 S1600 K1601 S1602 G1603 L1604 S1605 Y1606 K1607 S1608 F1612 A1618 T1621 P1624 M1625

L1626 L1629 D1630 K1631 C1634 T1635 L1636 F1637 F1638 E1647 L1648 A1649

5 Refinement protocol and experimental data overview (i)

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

Of the 200 calculated structures, 20 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**
ARIA	refinement	2.2
ARIA	structure solution	2.2
CING	refinement	iCing
TALOS	geometry optimization	TALOS+

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	1640
Number of shifts mapped to atoms	1640
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	90%

6 Model quality (i)

6.1 Standard geometry (i)

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

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	0.1±0.4
All	All	0	2

There are no bond-length outliers.

There are no bond-angle outliers.

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	1571	ARG	Sidechain	1
1	A	1580	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	980	961	959	17±4
All	All	19600	19220	19180	349

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 9.

5 of 84 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:1559:CYS:HB2	1:A:1563:VAL:HG11	0.87	1.45	14	19
1:A:1606:TYR:HA	1:A:1634:CYS:HB3	0.79	1.52	20	6
1:A:1517:GLN:HB2	1:A:1526:LEU:HD21	0.74	1.57	14	8
1:A:1588:LEU:HB2	1:A:1612:PHE:HB2	0.71	1.62	19	20
1:A:1616:PRO:HA	1:A:1643:PRO:HB3	0.69	1.64	4	3

6.3 Torsion angles (i)

6.3.1 Protein backbone (i)

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	128/142 (90%)	112±2 (88±2%)	14±2 (11±2%)	2±1 (2±1%)	11	53
All	All	2560/2840 (90%)	2240 (88%)	272 (11%)	48 (2%)	11	53

5 of 6 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	1571	ARG	18
1	A	1604	LEU	14
1	A	1553	CYS	8
1	A	1534	GLY	3
1	A	1533	SER	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	108/120 (90%)	103±1 (95±1%)	5±1 (5±1%)	31	79
All	All	2160/2400 (90%)	2060 (95%)	100 (5%)	31	79

5 of 20 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	1629	LEU	20
1	A	1631	LYS	20
1	A	1593	LYS	19
1	A	1647	GLU	11
1	A	1578	ASN	6

6.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

6.5 Carbohydrates (i)

There are no monosaccharides in this entry.

6.6 Ligand geometry (i)

There are no ligands in this entry.

6.7 Other polymers (i)

There are no such molecules in this entry.

6.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

7 Chemical shift validation (i)

The completeness of assignment taking into account all chemical shift lists is 90% 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 (i)

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

Total number of shifts	1640
Number of shifts mapped to atoms	1640
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	12

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	141	-0.52 ± 0.14	Should be applied
$^{13}C_{\beta}$	126	-0.63 ± 0.16	Should be applied
$^{13}C'$	130	-0.26 ± 0.12	None needed (< 0.5 ppm)
^{15}N	132	-0.43 ± 0.49	None needed (< 0.5 ppm)

7.1.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 90%, i.e. 1508 atoms were assigned a chemical shift out of a possible 1680. 0 out of 21 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	627/641 (98%)	261/262 (100%)	245/258 (95%)	121/121 (100%)
Sidechain	800/894 (89%)	544/586 (93%)	241/274 (88%)	15/34 (44%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	81/145 (56%)	55/70 (79%)	24/69 (35%)	2/6 (33%)
Overall	1508/1680 (90%)	860/918 (94%)	510/601 (85%)	138/161 (86%)

7.1.4 Statistically unusual chemical shifts (i)

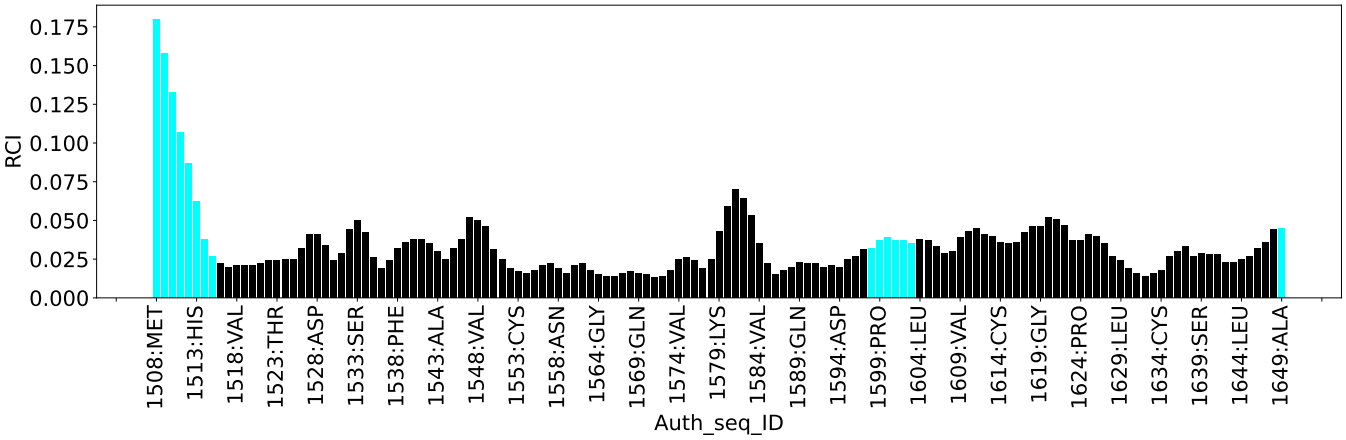
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	1642	THR	HG1	6.54	0.08 – 2.19	25.6
1	A	1572	ILE	HB	-0.31	0.35 – 3.22	-7.3
1	A	1630	ASP	HB3	4.32	1.32 – 4.00	6.2
1	A	1578	ASN	HB2	0.92	1.27 – 4.34	-6.1
1	A	1635	THR	HG21	-0.02	0.08 – 2.19	-5.5
1	A	1635	THR	HG22	-0.02	0.08 – 2.19	-5.5
1	A	1635	THR	HG23	-0.02	0.08 – 2.19	-5.5
1	A	1529	LEU	HD11	-0.67	-0.61 – 2.12	-5.2
1	A	1529	LEU	HD12	-0.67	-0.61 – 2.12	-5.2
1	A	1529	LEU	HD13	-0.67	-0.61 – 2.12	-5.2
1	A	1636	LEU	CG	32.30	21.37 – 32.19	5.1
1	A	1558	ASN	HB2	1.24	1.27 – 4.34	-5.1

7.1.5 Random Coil Index (RCI) plots (i)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:


RCI
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
1508:MET
1513:HIS
1518:VAL
1523:THR
1528:ASP
1533:SER
1538:PHE
1543:ALA
1548:VAL
1553:CYS
1558:ASN
1564:GLY
1569:GLN
1574:VAL
1579:LYS
1584:VAL
1589:GLN
1594:ASP
1599:PRO
1604:LEU
1609:VAL
1614:CYS
1619:GLY
1624:PRO
1629:LEU
1634:CYS
1639:SER
1644:LEU
1649:ALA
Auth_seq_ID

8 NMR restraints analysis (i)

8.1 Conformationally restricting restraints (i)

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

Description	Value
Total distance restraints	3228
Intra-residue (|i-j|=0)	1242
Sequential (|i-j|=1)	791
Medium range (|i-j|>1 and |i-j|<5)	327
Long range (|i-j|≥5)	816
Inter-chain	0
Hydrogen bond restraints	52
Disulfide bond restraints	0
Total dihedral-angle restraints	249
Number of unmapped restraints	0
Number of restraints per residue	24.5
Number of long range restraints per residue[1]	6.1

[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)	21.6	0.2
0.2-0.5 (Medium)	49.6	0.5
>0.5 (Large)	59.2	3.14

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)	15.9	3.5
10.0-20.0 (Medium)	None	None
>20.0 (Large)	None	None

9 Distance violation analysis (i)

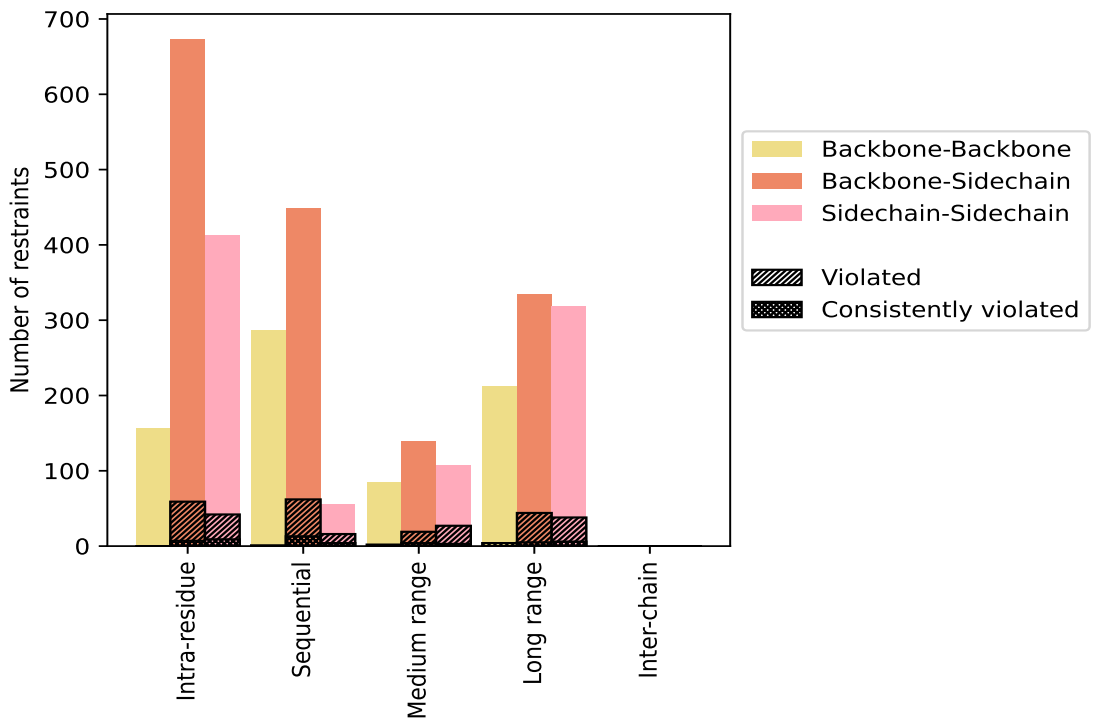
9.1 Summary of distance violations (i)

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

Restraints type	Count	%[1]	Violated[3]			Consistently Violated[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
Intra-residue (|i-j|=0)	1242	38.5	101	8.1	3.1	16	1.3	0.5
Backbone-Backbone	156	4.8	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	673	20.8	59	8.8	1.8	7	1.0	0.2
Sidechain-Sidechain	413	12.8	42	10.2	1.3	9	2.2	0.3
Sequential (|i-j|=1)	791	24.5	79	10.0	2.4	17	2.1	0.5
Backbone-Backbone	286	8.9	1	0.3	0.0	0	0.0	0.0
Backbone-Sidechain	449	13.9	62	13.8	1.9	13	2.9	0.4
Sidechain-Sidechain	56	1.7	16	28.6	0.5	4	7.1	0.1
Medium range (|i-j|>1 & |i-j|<5)	327	10.1	48	14.7	1.5	7	2.1	0.2
Backbone-Backbone	81	2.5	2	2.5	0.1	0	0.0	0.0
Backbone-Sidechain	139	4.3	19	13.7	0.6	4	2.9	0.1
Sidechain-Sidechain	107	3.3	27	25.2	0.8	3	2.8	0.1
Long range (|i-j|≥5)	816	25.3	86	10.5	2.7	11	1.3	0.3
Backbone-Backbone	164	5.1	4	2.4	0.1	0	0.0	0.0
Backbone-Sidechain	334	10.3	44	13.2	1.4	5	1.5	0.2
Sidechain-Sidechain	318	9.9	38	11.9	1.2	6	1.9	0.2
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	52	1.6	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	3228	100.0	314	9.7	9.7	51	1.6	1.6
Backbone-Backbone	739	22.9	7	0.9	0.2	0	0.0	0.0
Backbone-Sidechain	1595	49.4	184	11.5	5.7	29	1.8	0.9
Sidechain-Sidechain	894	27.7	123	13.8	3.8	22	2.5	0.7

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

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



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

9.2 Distance violation statistics for each model (i)

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

Model ID	Number of violations						Mean (Å)	Max (Å)	SD[6] (Å)	Median (Å)
	IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total				
1	45	33	20	33	0	131	0.56	1.63	0.37	0.45
2	47	33	20	29	0	129	0.57	3.14	0.43	0.45
3	44	32	18	37	0	131	0.58	1.93	0.4	0.49
4	38	30	18	35	0	121	0.6	2.0	0.38	0.5
5	41	34	18	39	0	132	0.57	2.45	0.42	0.44
6	39	31	13	42	0	125	0.61	2.22	0.42	0.5
7	44	28	17	36	0	125	0.57	1.72	0.39	0.46
8	42	35	19	38	0	134	0.57	2.33	0.41	0.43
9	43	35	18	33	0	129	0.6	2.56	0.46	0.44
10	43	30	21	30	0	124	0.6	2.41	0.45	0.47
11	41	33	20	37	0	131	0.55	1.73	0.38	0.45

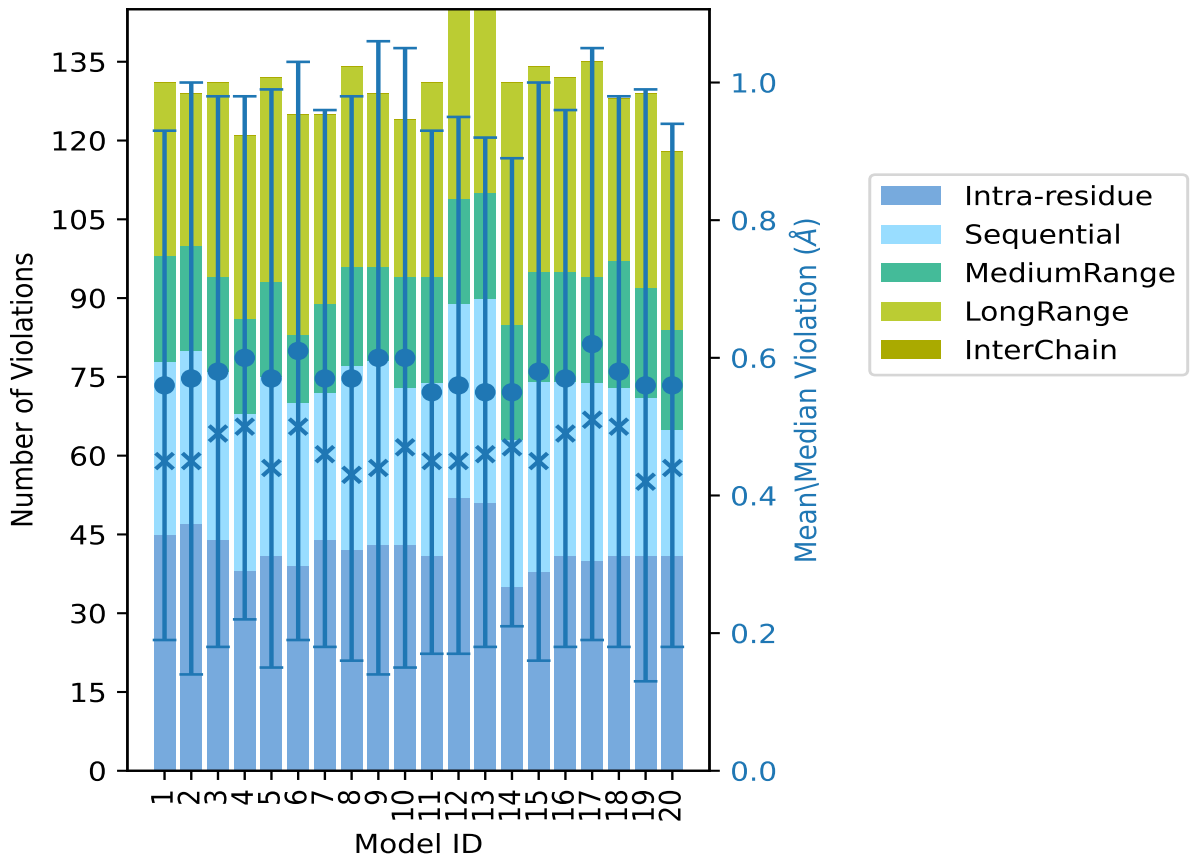
Continued on next page...

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	52	37	20	36	0	145	0.56	1.7	0.39	0.45
13	51	39	20	35	0	145	0.55	2.11	0.37	0.46
14	35	28	22	46	0	131	0.55	1.55	0.34	0.47
15	38	36	21	39	0	134	0.58	2.16	0.42	0.45
16	41	33	21	37	0	132	0.57	1.76	0.39	0.49
17	40	34	20	41	0	135	0.62	2.08	0.43	0.51
18	41	32	24	31	0	128	0.58	2.12	0.4	0.5
19	41	30	21	37	0	129	0.56	2.29	0.43	0.42
20	41	24	19	34	0	118	0.56	1.57	0.38	0.44

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

9.2.1 Bar graph : Distance Violation statistics for each model (i)



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

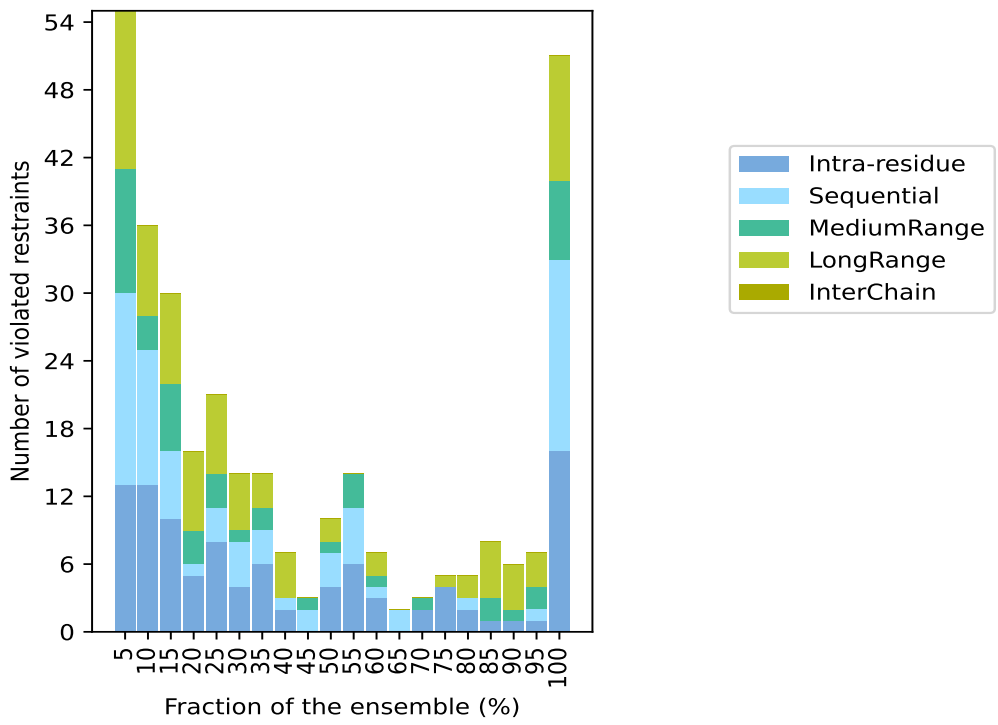
9.3 Distance violation statistics for the ensemble (i)

Violation analysis may find that some restraints are violated in few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of the ensemble. In total, 2862(IR:1141, SQ:712, MR:279, LR:730, 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]	%
13	17	11	14	0	55	1	5.0
13	12	3	8	0	36	2	10.0
10	6	6	8	0	30	3	15.0
5	1	3	7	0	16	4	20.0
8	3	3	7	0	21	5	25.0
4	4	1	5	0	14	6	30.0
6	3	2	3	0	14	7	35.0
2	1	0	4	0	7	8	40.0
0	2	1	0	0	3	9	45.0
4	3	1	2	0	10	10	50.0
6	5	3	0	0	14	11	55.0
3	1	1	2	0	7	12	60.0
0	2	0	0	0	2	13	65.0
2	0	1	0	0	3	14	70.0
4	0	0	1	0	5	15	75.0
2	1	0	2	0	5	16	80.0
1	0	2	5	0	8	17	85.0
1	0	1	4	0	6	18	90.0
1	1	2	3	0	7	19	95.0
16	17	7	11	0	51	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

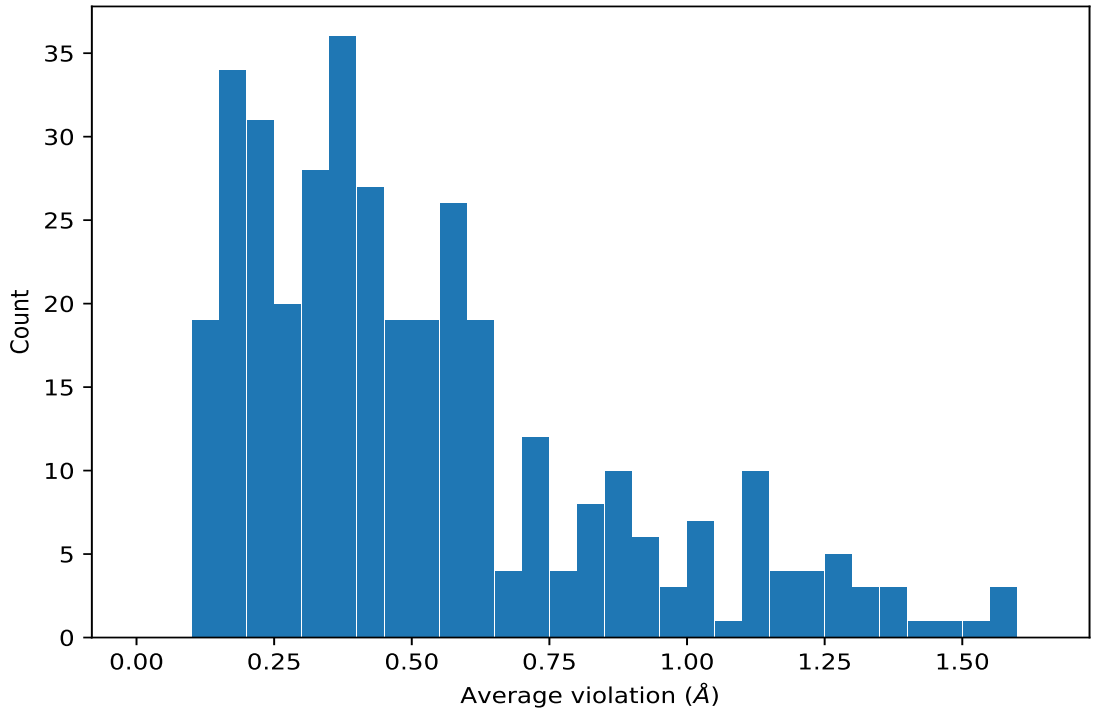
9.3.1 Bar graph : Distance violation statistics for the ensemble (i)



9.4 Most violated distance restraints in the ensemble (i)

9.4.1 Histogram : Distribution of mean distance violations (i)

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



9.4.2 Table: Most violated distance restraints (i)

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

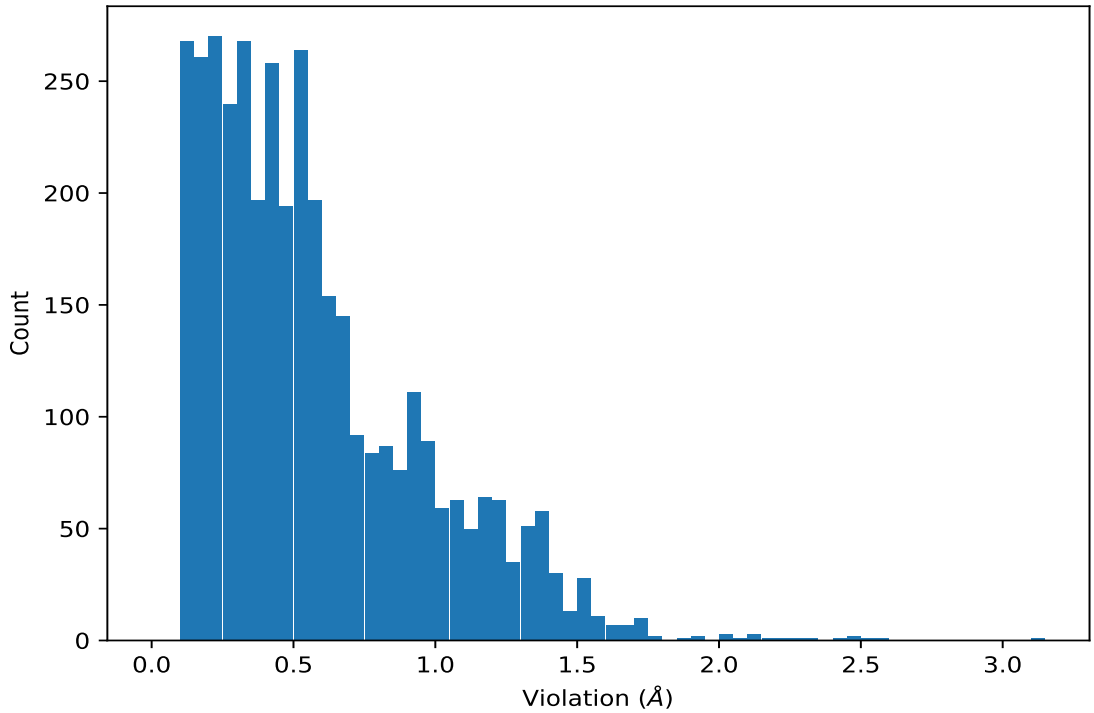
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(2,1074)	1:A:1597:PRO:HB2	1:A:1605:SER:HB2	20	1.56	0.41	1.53
(2,2239)	1:A:1619:GLY:H	1:A:1622:ASN:HD22	20	1.54	0.14	1.58
(2,826)	1:A:1582:ARG:HD2	1:A:1589:GLN:HB2	20	1.44	1.01	1.94
(2,2160)	1:A:1522:SER:H	1:A:1521:PRO:HG3	20	1.37	0.04	1.39
(2,1765)	1:A:1622:ASN:H	1:A:1622:ASN:HD22	20	1.35	0.02	1.35
(2,2236)	1:A:1619:GLY:H	1:A:1622:ASN:HB3	20	1.33	0.11	1.34
(2,125)	1:A:1522:SER:HB3	1:A:1521:PRO:HG3	20	1.27	0.33	1.31
(2,1494)	1:A:1589:GLN:HB3	1:A:1587:VAL:HG22	20	1.27	0.17	1.33
(2,1494)	1:A:1589:GLN:HB3	1:A:1587:VAL:HG21	20	1.27	0.17	1.33
(2,1494)	1:A:1589:GLN:HB3	1:A:1587:VAL:HG23	20	1.27	0.17	1.33
(2,1252)	1:A:1618:ALA:HB2	1:A:1643:PRO:HD3	20	1.16	0.18	1.21
(2,1252)	1:A:1618:ALA:HB1	1:A:1643:PRO:HD3	20	1.16	0.18	1.21
(2,1252)	1:A:1618:ALA:HB3	1:A:1643:PRO:HD3	20	1.16	0.18	1.21
(2,1641)	1:A:1630:ASP:H	1:A:1637:PHE:HB3	20	1.15	0.12	1.13

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

9.5 All violated distance restraints (i)

9.5.1 Histogram : Distribution of distance violations (i)

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



9.5.2 Table : All distance violations (i)

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

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(2,826)	1:A:1582:ARG:HD2	1:A:1589:GLN:HB2	2	3.14
(2,826)	1:A:1582:ARG:HD2	1:A:1589:GLN:HB2	9	2.56
(2,1074)	1:A:1597:PRO:HB2	1:A:1605:SER:HB2	9	2.52
(2,1074)	1:A:1597:PRO:HB2	1:A:1605:SER:HB2	2	2.47
(2,826)	1:A:1582:ARG:HD2	1:A:1589:GLN:HB2	5	2.45
(2,826)	1:A:1582:ARG:HD2	1:A:1589:GLN:HB2	10	2.41
(2,826)	1:A:1582:ARG:HD2	1:A:1589:GLN:HB2	8	2.33
(2,826)	1:A:1582:ARG:HD2	1:A:1589:GLN:HB2	19	2.29
(2,826)	1:A:1582:ARG:HD2	1:A:1589:GLN:HB2	6	2.22
(2,1394)	1:A:1615:ARG:HD3	1:A:1617:GLU:HB3	15	2.16

10 Dihedral-angle violation analysis (i)

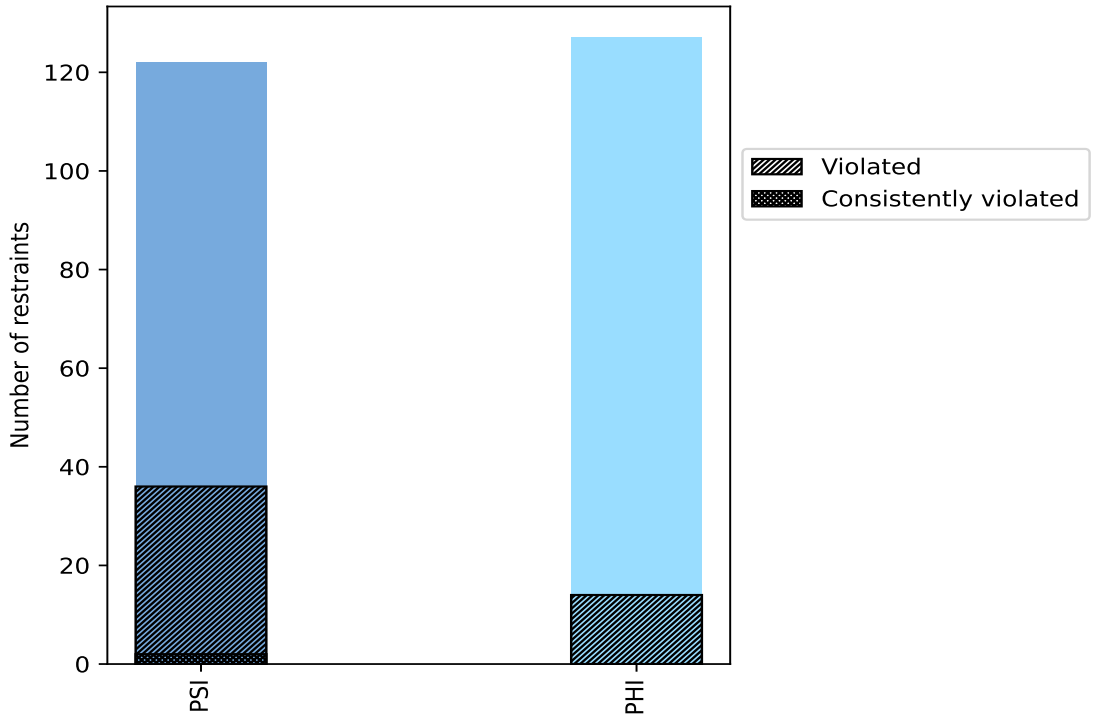
10.1 Summary of dihedral-angle violations (i)

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

Angle type	Count	%[1]	Violated[3]			Consistently Violated[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
PSI	122	49.0	36	29.5	14.5	2	1.6	0.8
PHI	127	51.0	14	11.0	5.6	0	0.0	0.0
Total	249	100.0	50	20.1	20.1	2	0.8	0.8

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

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



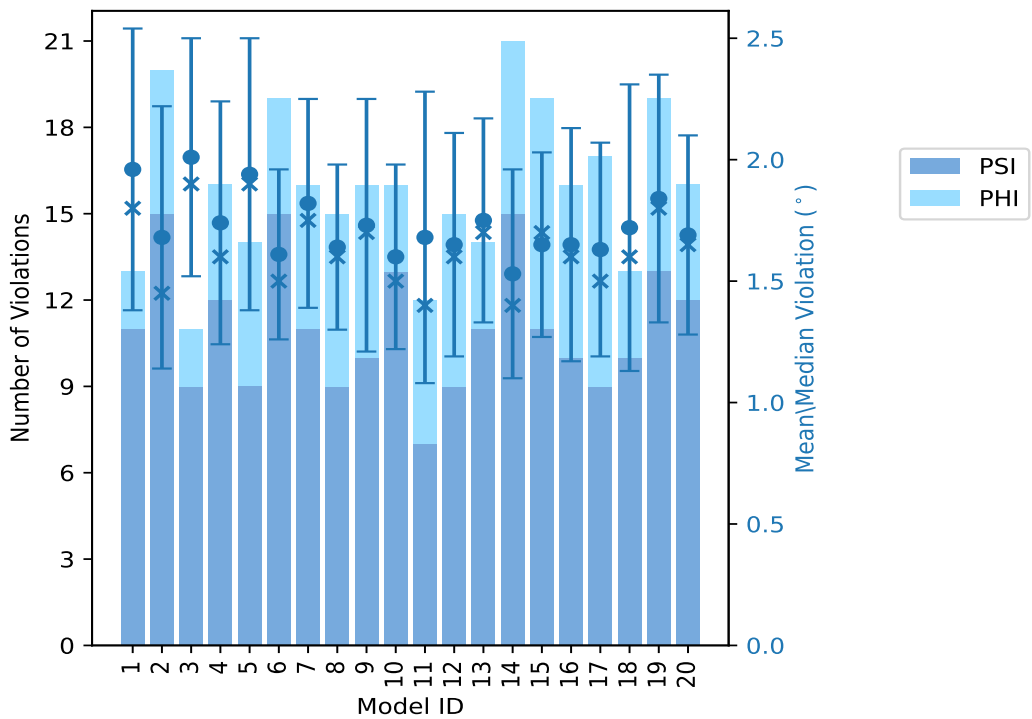
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

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

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

Model ID	Number of violations			Mean (°)	Max (°)	SD (°)	Median (°)
	PSI	PHI	Total				
1	11	2	13	1.96	3.3	0.58	1.8
2	15	5	20	1.68	2.8	0.54	1.45
3	9	2	11	2.01	3.2	0.49	1.9
4	12	4	16	1.74	2.6	0.5	1.6
5	9	5	14	1.94	3.5	0.56	1.9
6	15	4	19	1.61	2.6	0.35	1.5
7	11	5	16	1.82	3.0	0.43	1.75
8	9	6	15	1.64	2.3	0.34	1.6
9	10	6	16	1.73	2.9	0.52	1.7
10	13	3	16	1.6	2.4	0.38	1.5
11	7	5	12	1.68	3.0	0.6	1.4
12	9	6	15	1.65	2.8	0.46	1.6
13	11	3	14	1.75	2.7	0.42	1.7
14	15	6	21	1.53	2.6	0.43	1.4
15	11	8	19	1.65	2.3	0.38	1.7
16	10	6	16	1.65	3.1	0.48	1.6
17	9	8	17	1.63	2.6	0.44	1.5
18	10	3	13	1.72	2.9	0.59	1.6
19	13	6	19	1.84	3.4	0.51	1.8
20	12	4	16	1.69	2.7	0.41	1.65

10.2.1 Bar graph : Dihedral violation statistics for each model (i)



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

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

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

Number of violated restraints			Fraction of the ensemble	
PSI	PHI	Total	Count[1]	%
10	1	11	1	5.0
6	1	7	2	10.0
0	3	3	3	15.0
2	1	3	4	20.0
5	1	6	5	25.0
2	1	3	6	30.0
0	1	1	7	35.0
1	0	1	8	40.0
1	1	2	9	45.0
0	0	0	10	50.0
1	2	3	11	55.0

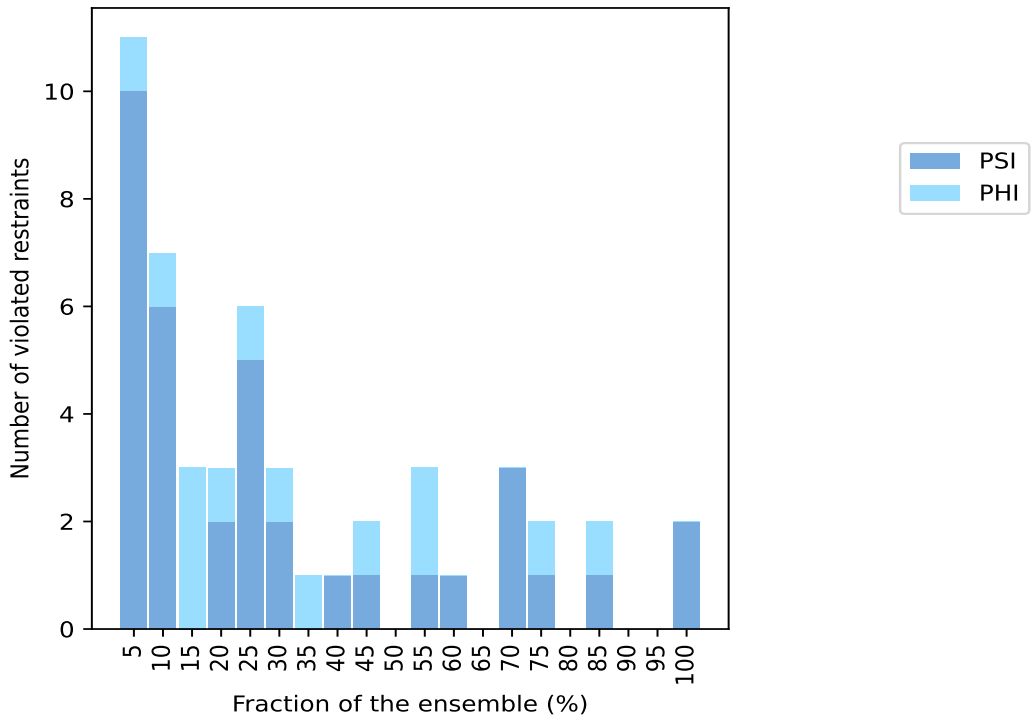
Continued on next page...

Continued from previous page...

Number of violated restraints			Fraction of the ensemble	
PSI	PHI	Total	Count[1]	%
1	0	1	12	60.0
0	0	0	13	65.0
3	0	3	14	70.0
1	1	2	15	75.0
0	0	0	16	80.0
1	1	2	17	85.0
0	0	0	18	90.0
0	0	0	19	95.0
2	0	2	20	100.0

[1] Number of models with violations

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

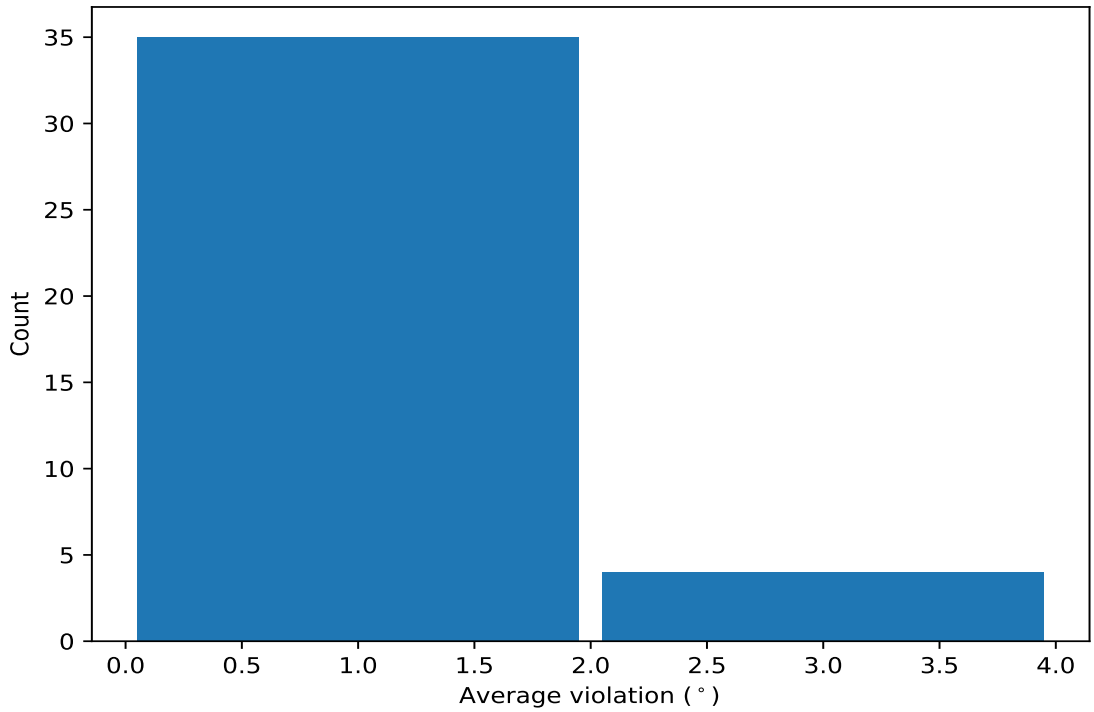


10.4 Most violated dihedral-angle restraints in the ensemble (i)

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.

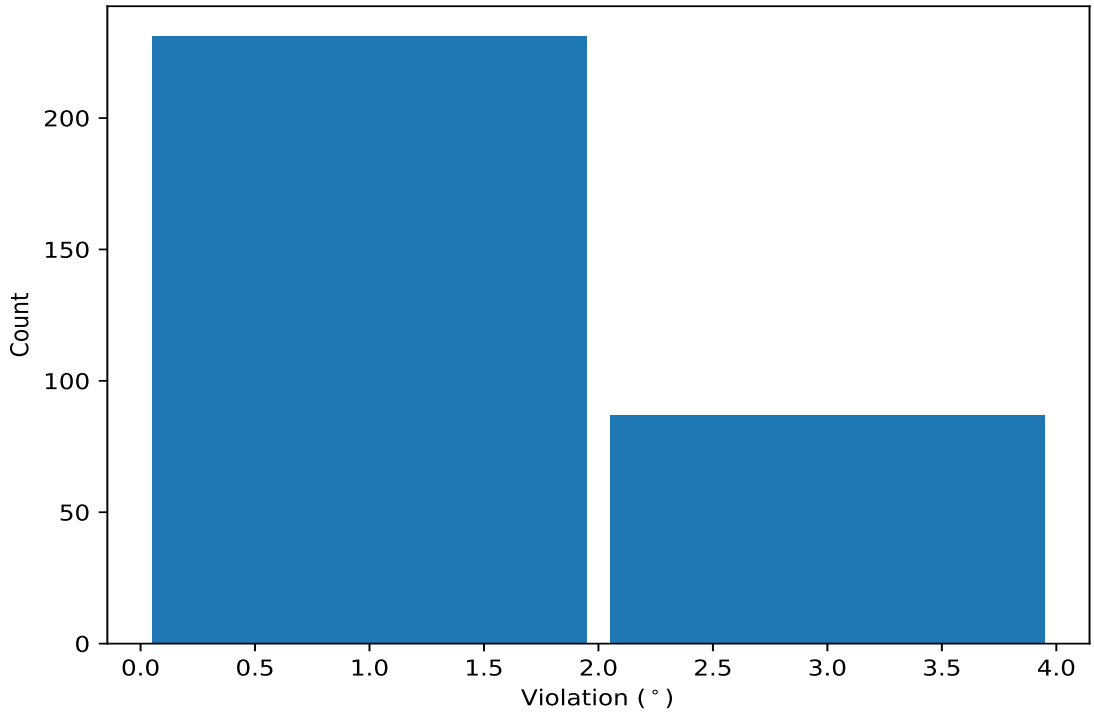
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models[1]	Mean	SD[2]	Medi
(1,151)	1:A:1597:PRO:N	1:A:1597:PRO:CA	1:A:1597:PRO:C	1:A:1598:CYS:N	20	2.14	0.42	2.0
(1,169)	1:A:1606:TYR:N	1:A:1606:TYR:CA	1:A:1606:TYR:C	1:A:1607:LYS:N	20	2.04	0.68	1.8
(1,35)	1:A:1528:ASP:C	1:A:1529:LEU:N	1:A:1529:LEU:CA	1:A:1529:LEU:C	17	1.93	0.42	1.8
(1,79)	1:A:1554:GLY:N	1:A:1554:GLY:CA	1:A:1554:GLY:C	1:A:1555:GLU:N	17	1.72	0.34	1.8
(1,44)	1:A:1533:SER:N	1:A:1533:SER:CA	1:A:1533:SER:C	1:A:1534:GLY:N	15	1.93	0.6	1.8
(1,83)	1:A:1556:ASN:C	1:A:1557:GLU:N	1:A:1557:GLU:CA	1:A:1557:GLU:C	15	1.75	0.35	1.7
(1,106)	1:A:1570:THR:N	1:A:1570:THR:CA	1:A:1570:THR:C	1:A:1571:ARG:N	14	1.89	0.3	1.9
(1,193)	1:A:1618:ALA:N	1:A:1618:ALA:CA	1:A:1618:ALA:C	1:A:1619:GLY:N	14	1.87	0.33	1.9
(1,147)	1:A:1595:GLY:N	1:A:1595:GLY:CA	1:A:1595:GLY:C	1:A:1596:SER:N	14	1.69	0.53	1.6
(1,247)	1:A:1647:GLU:N	1:A:1647:GLU:CA	1:A:1647:GLU:C	1:A:1648:LEU:N	12	1.66	0.35	1.6

[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 (i)

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

Key	Atom-1	Atom-2	Atom-3	Atom-4	Model ID	Violation (°)
(1,169)	1:A:1606:TYR:N	1:A:1606:TYR:CA	1:A:1606:TYR:C	1:A:1607:LYS:N	5	3.5
(1,44)	1:A:1533:SER:N	1:A:1533:SER:CA	1:A:1533:SER:C	1:A:1534:GLY:N	19	3.4
(1,169)	1:A:1606:TYR:N	1:A:1606:TYR:CA	1:A:1606:TYR:C	1:A:1607:LYS:N	1	3.3
(1,147)	1:A:1595:GLY:N	1:A:1595:GLY:CA	1:A:1595:GLY:C	1:A:1596:SER:N	3	3.2
(1,169)	1:A:1606:TYR:N	1:A:1606:TYR:CA	1:A:1606:TYR:C	1:A:1607:LYS:N	16	3.1
(1,44)	1:A:1533:SER:N	1:A:1533:SER:CA	1:A:1533:SER:C	1:A:1534:GLY:N	1	3.0
(1,151)	1:A:1597:PRO:N	1:A:1597:PRO:CA	1:A:1597:PRO:C	1:A:1598:CYS:N	11	3.0
(1,101)	1:A:1567:PHE:N	1:A:1567:PHE:CA	1:A:1567:PHE:C	1:A:1568:GLY:N	7	3.0
(1,151)	1:A:1597:PRO:N	1:A:1597:PRO:CA	1:A:1597:PRO:C	1:A:1598:CYS:N	9	2.9
(1,151)	1:A:1597:PRO:N	1:A:1597:PRO:CA	1:A:1597:PRO:C	1:A:1598:CYS:N	18	2.9