wwPDB NMR Structure Validation Summary Report

Apr 29, 2024 – 05:30 PM EDT

PDB ID	:	8SG2
BMRB ID	:	31080
Title	:	BIVALENT INTERACTIONS OF PIN1 WITH THE C-TERMINAL TAIL OF PKC
Authors	:	Dixit, K.; Yang, Y.; Chen, X.R.; Igumenova, T.I.
Deposited on	:	2023-04-11

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 (i)) were used in the production of this report:

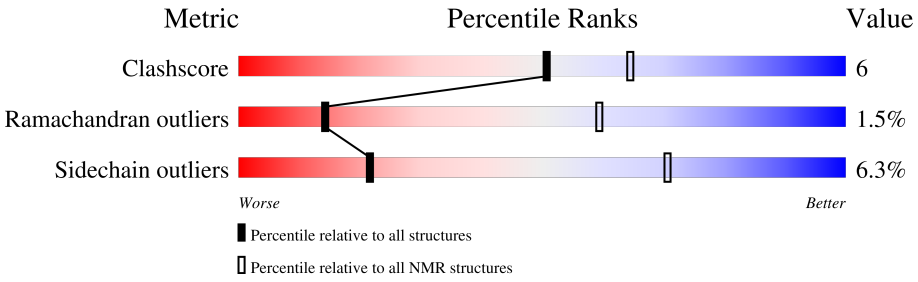
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (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 (i)

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

The overall completeness of chemical shifts assignment is 80%.

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	163	80% 8% 12%
2	B	25	40% 32% 28%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 4 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:7-A:37, A:52-A:163, B:639-B:640, B:642-B:646, B:650-B:659, B:661-B:661 (161)	1.07	4

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 1 clusters and 10 single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2902 atoms, of which 1420 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	163	2526	786	1244	239	251	6	0

- Molecule 2 is a protein called Protein kinase C beta type.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
2	B	25	376	122	176	30	46	2	1

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	638	ACE	-	acetylation	UNP P05771
B	662	NH2	-	amidation	UNP P05771

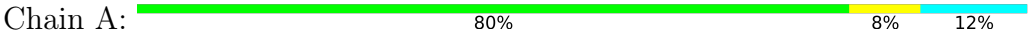
4 Residue-property plots

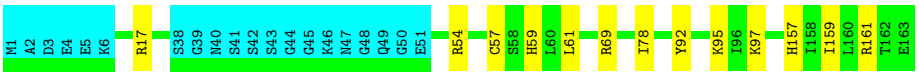
4.1 Average score per residue in the NMR ensemble

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: Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1

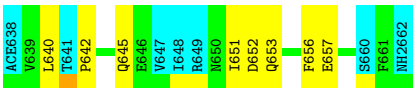
Chain A:





• Molecule 2: Protein kinase C beta type

Chain B:

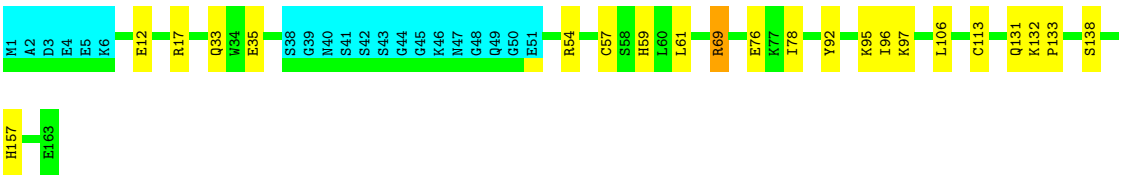


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 4. Colouring as in section 4.1 above.

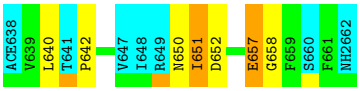
• Molecule 1: Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1

Chain A:



• Molecule 2: Protein kinase C beta type

Chain B:


ACE638
V639
L640
T641
P642
V647
I648
R649
N650
I651
D652
E657
G658
P659
S660
F661
NH2662

5 Refinement protocol and experimental data overview

The models were refined using the following method: *torsion angle dynamics, simulated annealing.*

Of the 50 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy.*

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

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

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	2
Total number of shifts	1960
Number of shifts mapped to atoms	1960
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	80%

6 Model quality

6.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NH2, SEP, TPO, ACE

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	1147	1125	1121	12±3
2	B	149	128	128	7±2
All	All	25920	25060	24980	298

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 6.

5 of 165 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:B:640:LEU:O	2:B:642:PRO:HD2	0.70	1.85	11	14
2:B:653:GLN:H	2:B:653:GLN:NE2	0.68	1.86	15	1
1:A:61:LEU:HG	1:A:113:CYS:SG	0.65	2.31	18	1
1:A:63:LYS:O	1:A:80:ARG:HD3	0.64	1.93	18	1
1:A:107:ALA:O	1:A:111:SER:HB2	0.62	1.95	14	4

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	142/163 (87%)	135±2 (95±1%)	6±2 (4±1%)	1±1 (1±0%)	32 76
2	B	18/25 (72%)	11±2 (62±10%)	5±1 (29±8%)	2±1 (9±5%)	1 12
All	All	3200/3760 (85%)	2927 (91%)	226 (7%)	47 (1%)	14 59

5 of 14 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
2	B	651	ILE	14
1	A	78	ILE	11
2	B	652	ASP	4
2	B	639	VAL	3
1	A	52	PRO	2

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	124/138 (90%)	117±2 (94±1%)	7±2 (6±1%)	24 73
2	B	17/20 (85%)	15±1 (89±7%)	2±1 (11±7%)	10 54
All	All	2820/3160 (89%)	2642 (94%)	178 (6%)	21 70

5 of 47 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	54	ARG	16
1	A	97	LYS	14

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	157	HIS	13
1	A	59	HIS	13
1	A	69	ARG	10

6.3.3 RNA (i)

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
2	TPO	B	641	2	8,10,11	1.57±0.08	1±0 (14±4%)
2	SEP	B	660	2	8,9,10	1.01±0.05	0±0 (0±0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of angles that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
2	TPO	B	641	2	10,14,16	1.65±0.03	4±1 (39±6%)
2	SEP	B	660	2	8,12,14	2.04±0.09	2±0 (25±0%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SEP	B	660	2	-	0±0,5,8,10	-
2	TPO	B	641	2	-	0±0,9,11,13	-

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
2	B	641	TPO	P-O1P	3.78	1.62	1.50	18	20
2	B	641	TPO	P-O3P	2.09	1.62	1.54	4	3

5 of 7 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
2	B	660	SEP	OG-P-O1P	4.08	117.92	106.47	7	20
2	B	660	SEP	OG-CB-CA	3.54	111.59	108.14	2	20
2	B	641	TPO	O-C-CA	2.60	117.98	124.78	13	20
2	B	641	TPO	O2P-P-OG1	2.59	117.60	105.99	19	20
2	B	641	TPO	O3P-P-O1P	2.45	101.08	110.68	16	20

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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

The completeness of assignment taking into account all chemical shift lists is 80% for the well-defined parts and 79% 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	1776
Number of shifts mapped to atoms	1776
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	10

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	161	-0.27 ± 0.18	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	143	-0.13 ± 0.17	None needed (< 0.5 ppm)
$^{13}C'$	160	-0.05 ± 0.14	None needed (< 0.5 ppm)
^{15}N	153	0.09 ± 0.20	None needed (< 0.5 ppm)

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 74%, i.e. 1623 atoms were assigned a chemical shift out of a possible 2206. 0 out of 15 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	701/798 (88%)	283/324 (87%)	283/322 (88%)	135/152 (89%)
Sidechain	817/1213 (67%)	534/778 (69%)	258/372 (69%)	25/63 (40%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	105/195 (54%)	59/96 (61%)	43/88 (49%)	3/11 (27%)
Overall	1623/2206 (74%)	876/1198 (73%)	584/782 (75%)	163/226 (72%)

7.1.4 Statistically unusual chemical shifts (i)

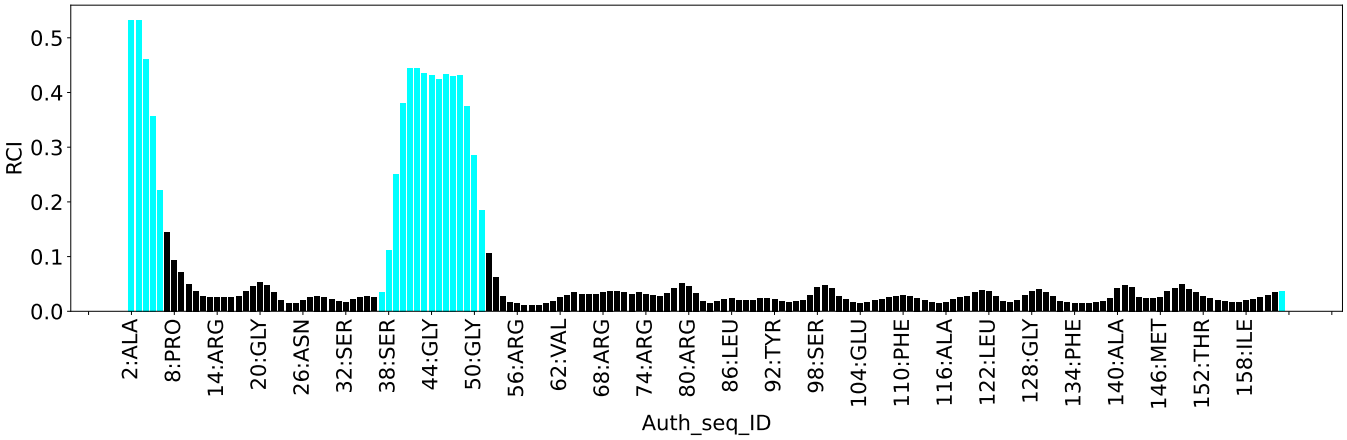
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	26	ASN	HB2	-0.66	1.27 – 4.34	-11.3
1	A	74	ARG	HG3	-0.55	0.15 – 2.94	-7.5
1	A	37	PRO	HG2	-0.03	0.41 – 3.45	-6.4
1	A	14	ARG	HB3	0.08	0.43 – 3.11	-6.3
1	A	26	ASN	HD22	4.24	4.69 – 9.61	-5.9
1	A	74	ARG	HG2	0.03	0.26 – 2.87	-5.9
1	A	109	GLN	HB2	0.77	0.80 – 3.29	-5.1
1	A	55	VAL	HG21	-0.58	-0.58 – 2.19	-5.0
1	A	55	VAL	HG22	-0.58	-0.58 – 2.19	-5.0
1	A	55	VAL	HG23	-0.58	-0.58 – 2.19	-5.0

7.1.5 Random Coil Index (RCI) plots (i)

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

Random coil index (RCI) for chain A:



7.2 Chemical shift list 2

File name: working_cs.cif

Chemical shift list name: *assigned_chemical_shifts_2*

7.2.1 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	184
Number of shifts mapped to atoms	184
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	1

7.2.2 Chemical shift referencing

No chemical shift referencing corrections were calculated (not enough data).

7.2.3 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 6%, i.e. 137 atoms were assigned a chemical shift out of a possible 2206. 0 out of 15 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	41/798 (5%)	33/324 (10%)	5/322 (2%)	3/152 (2%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Sidechain	76/1213 (6%)	67/778 (9%)	9/372 (2%)	0/63 (0%)
Aromatic	20/195 (10%)	10/96 (10%)	10/88 (11%)	0/11 (0%)
Overall	137/2206 (6%)	110/1198 (9%)	24/782 (3%)	3/226 (1%)

7.2.4 Statistically unusual chemical shifts (i)

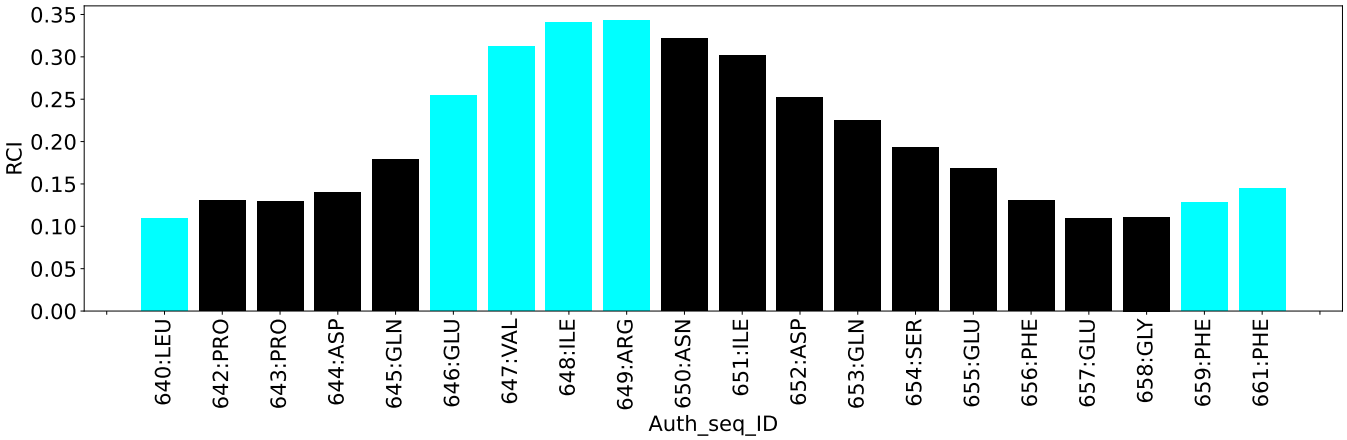
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
2	B	642	PRO	HG3	0.15	0.33 – 3.48	-5.6

7.2.5 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 B:



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	5666
Intra-residue (|i-j|=0)	2179
Sequential (|i-j|=1)	1277
Medium range (|i-j|>1 and |i-j|<5)	732
Long range (|i-j|≥5)	1284
Inter-chain	111
Hydrogen bond restraints	83
Disulfide bond restraints	0
Total dihedral-angle restraints	244
Number of unmapped restraints	0
Number of restraints per residue	31.4
Number of long range restraints per residue[1]	7.0

[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)	58.7	0.2
0.2-0.5 (Medium)	130.8	0.5
>0.5 (Large)	75.0	12.07

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

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

Bins (°)	Average number of violations per model	Max (°)
1.0-10.0 (Small)	10.6	5.69
10.0-20.0 (Medium)	None	None
>20.0 (Large)	None	None

9 Distance violation analysis (i)

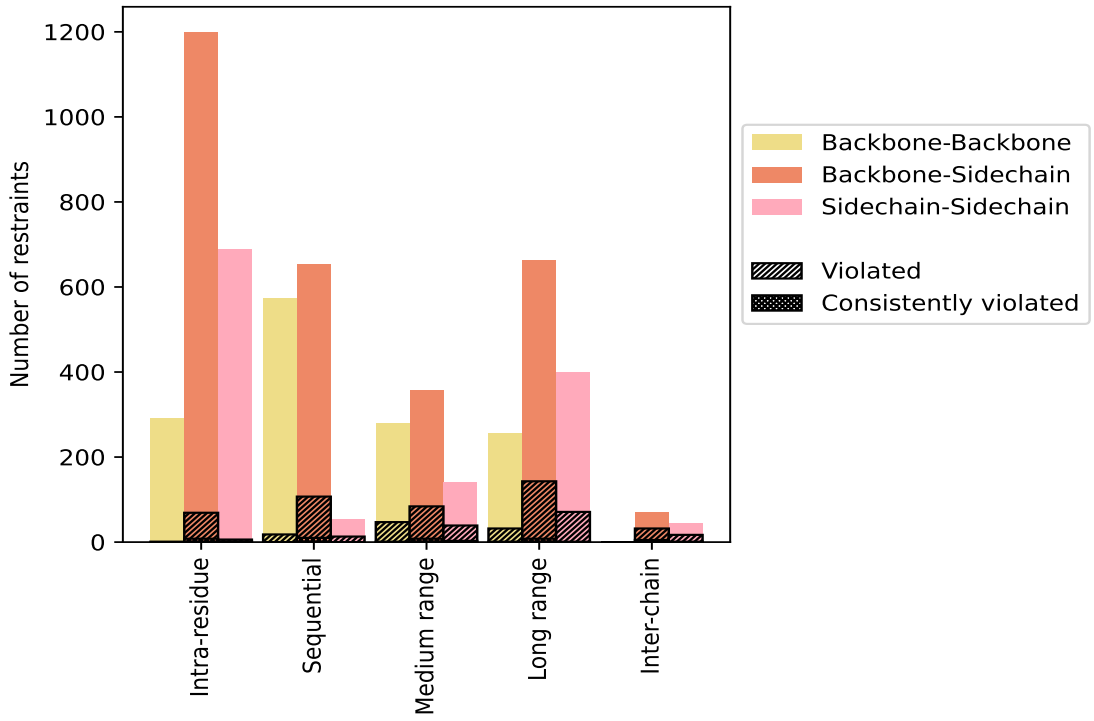
9.1 Summary of distance violations (i)

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)	2179	38.5	76	3.5	1.3	10	0.5	0.2
Backbone-Backbone	292	5.2	1	0.3	0.0	0	0.0	0.0
Backbone-Sidechain	1199	21.2	69	5.8	1.2	8	0.7	0.1
Sidechain-Sidechain	688	12.1	6	0.9	0.1	2	0.3	0.0
Sequential (|i-j|=1)	1277	22.5	138	10.8	2.4	10	0.8	0.2
Backbone-Backbone	572	10.1	18	3.1	0.3	0	0.0	0.0
Backbone-Sidechain	652	11.5	107	16.4	1.9	10	1.5	0.2
Sidechain-Sidechain	53	0.9	13	24.5	0.2	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	732	12.9	154	21.0	2.7	11	1.5	0.2
Backbone-Backbone	280	4.9	47	16.8	0.8	1	0.4	0.0
Backbone-Sidechain	311	5.5	68	21.9	1.2	7	2.3	0.1
Sidechain-Sidechain	141	2.5	39	27.7	0.7	3	2.1	0.1
Long range (|i-j|≥5)	1284	22.7	238	18.5	4.2	10	0.8	0.2
Backbone-Backbone	256	4.5	32	12.5	0.6	1	0.4	0.0
Backbone-Sidechain	629	11.1	135	21.5	2.4	8	1.3	0.1
Sidechain-Sidechain	399	7.0	71	17.8	1.3	1	0.3	0.0
Inter-chain	111	2.0	49	44.1	0.9	6	5.4	0.1
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	69	1.2	32	46.4	0.6	5	7.2	0.1
Sidechain-Sidechain	42	0.7	17	40.5	0.3	1	2.4	0.0
Hydrogen bond	83	1.5	24	28.9	0.4	1	1.2	0.0
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	5666	100.0	679	12.0	12.0	48	0.8	0.8
Backbone-Backbone	1400	24.7	98	7.0	1.7	2	0.1	0.0
Backbone-Sidechain	2940	51.9	435	14.8	7.7	39	1.3	0.7
Sidechain-Sidechain	1326	23.4	146	11.0	2.6	7	0.5	0.1

[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	37	52	69	93	24	275	0.47	10.81	0.92	0.37
2	34	52	64	85	23	258	0.46	9.46	0.78	0.38
3	30	56	60	103	26	275	0.47	11.07	0.93	0.37
4	34	49	65	89	24	261	0.47	11.49	0.98	0.37
5	32	61	68	94	25	280	0.46	10.54	0.89	0.37
6	32	50	70	95	28	275	0.47	9.95	0.89	0.4
7	31	59	68	85	20	263	0.46	10.86	0.94	0.36
8	31	58	57	93	24	263	0.47	11.25	0.94	0.36
9	30	53	55	92	27	257	0.49	11.6	0.97	0.37
10	32	47	62	95	26	262	0.47	8.4	0.74	0.4

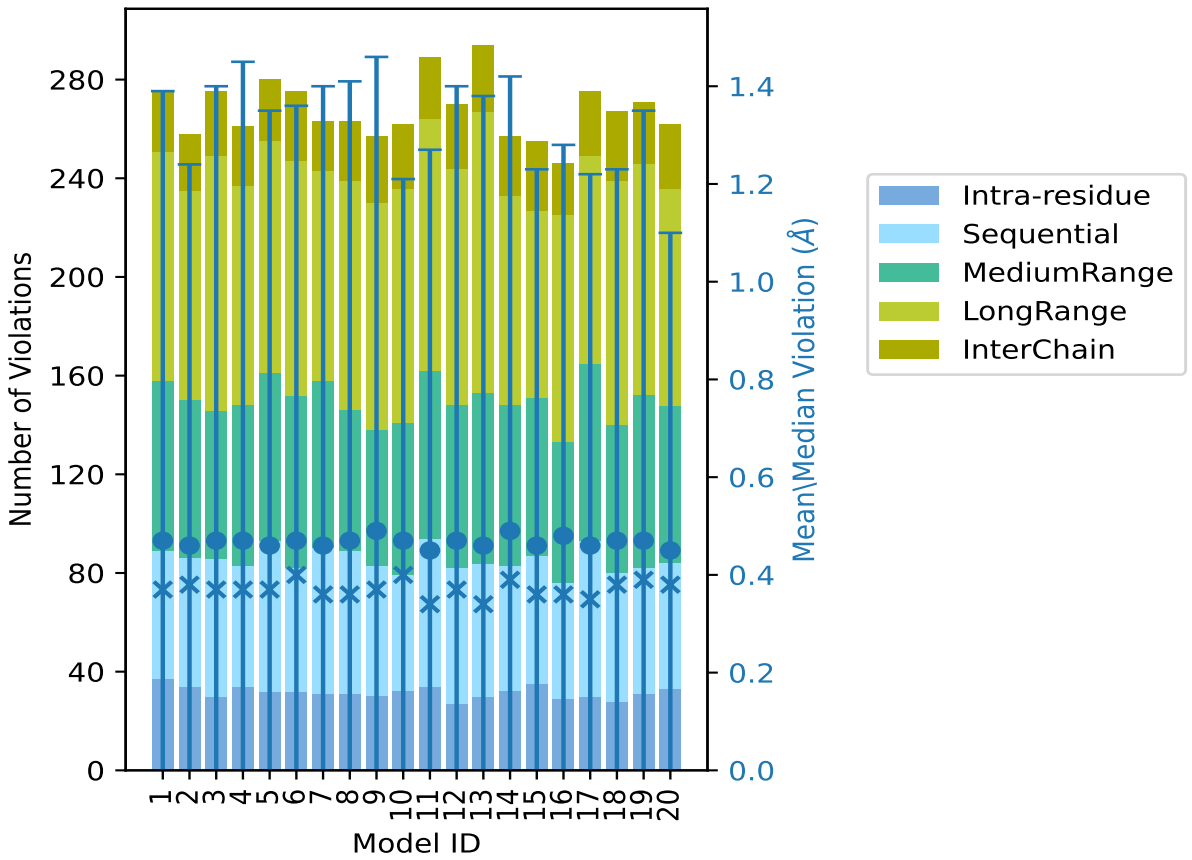
Continued on next page...

Continued from previous page...

Model ID	Number of violations						Mean (Å)	Max (Å)	SD[6] (Å)	Median (Å)
	IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total				
11	34	60	68	102	25	289	0.45	9.85	0.82	0.34
12	27	55	66	96	26	270	0.47	11.26	0.93	0.37
13	30	54	69	114	27	294	0.46	12.07	0.92	0.34
14	32	51	65	85	24	257	0.49	10.51	0.93	0.39
15	35	52	64	76	28	255	0.46	8.88	0.77	0.36
16	29	47	57	92	21	246	0.48	9.24	0.8	0.36
17	30	63	72	84	26	275	0.46	8.86	0.76	0.35
18	28	52	60	99	28	267	0.47	8.25	0.76	0.38
19	31	51	70	94	25	271	0.47	10.5	0.88	0.39
20	33	51	64	88	26	262	0.45	7.35	0.65	0.38

[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

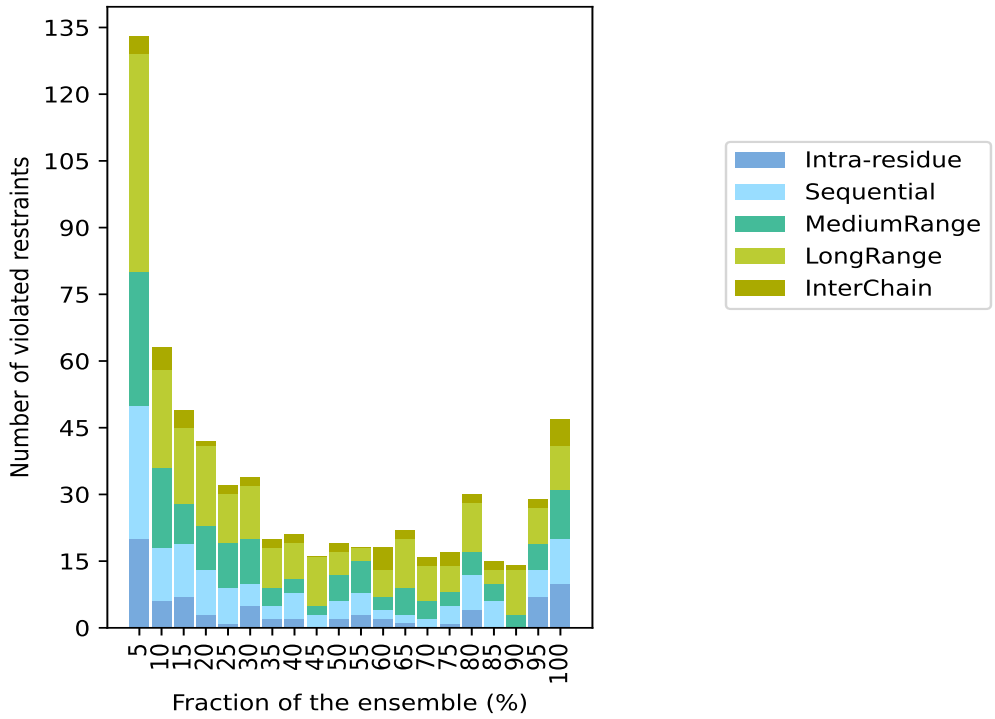
9.3 Distance violation statistics for the ensemble (i)

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, 4928(IR:2103, SQ:1139, MR:578, LR:1046, IC:62) 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]	%
20	30	30	49	4	133	1	5.0
6	12	18	22	5	63	2	10.0
7	12	9	17	4	49	3	15.0
3	10	10	18	1	42	4	20.0
1	8	10	11	2	32	5	25.0
5	5	10	12	2	34	6	30.0
2	3	4	9	2	20	7	35.0
2	6	3	8	2	21	8	40.0
0	3	2	11	0	16	9	45.0
2	4	6	5	2	19	10	50.0
3	5	7	3	0	18	11	55.0
2	2	3	6	5	18	12	60.0
1	2	6	11	2	22	13	65.0
0	2	4	8	2	16	14	70.0
1	4	3	6	3	17	15	75.0
4	8	5	11	2	30	16	80.0
0	6	4	3	2	15	17	85.0
0	0	3	10	1	14	18	90.0
7	6	6	8	2	29	19	95.0
10	10	11	10	6	47	20	100.0

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

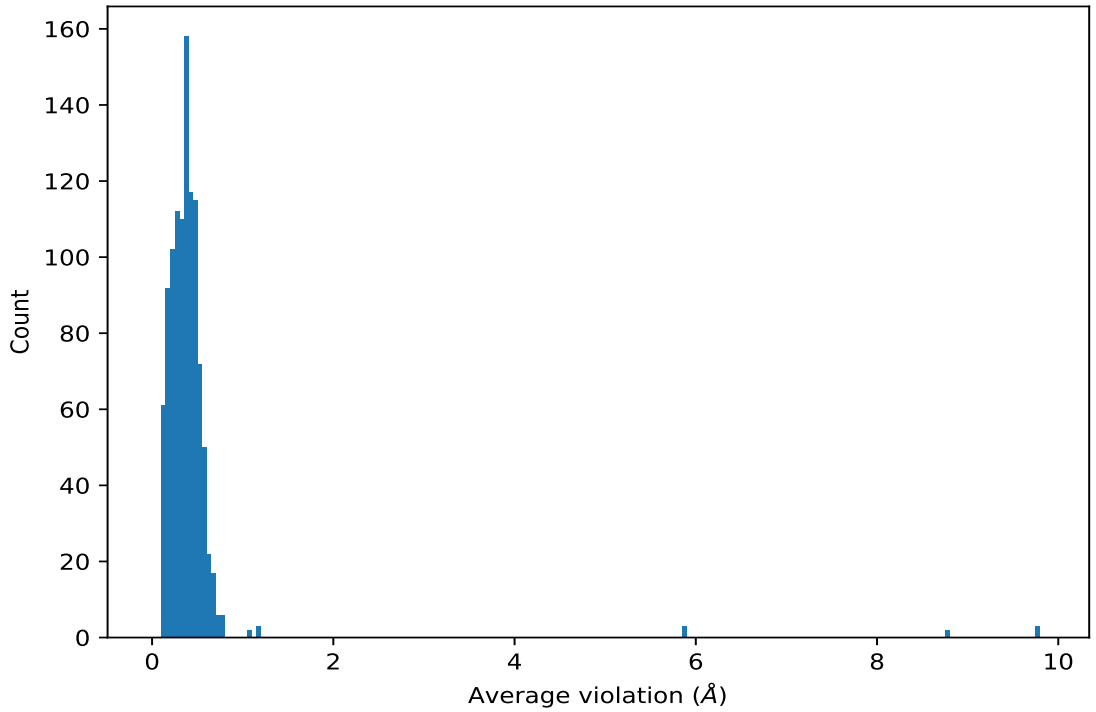
9.3.1 Bar graph : Distance violation statistics for the ensemble



9.4 Most violated distance restraints in the ensemble

9.4.1 Histogram : Distribution of mean distance violations

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



9.4.2 Table: Most violated distance restraints

The following table provides the mean and the standard deviation of 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)
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG21	20	9.76	1.66	10.5
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG22	20	9.76	1.66	10.5
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG23	20	9.76	1.66	10.5
(1,2595)	1:131:A:GLN:HG2	2:652:B:ASP:HA	20	8.75	0.86	9.09
(1,2595)	1:131:A:GLN:HG3	2:652:B:ASP:HA	20	8.75	0.86	9.09
(1,1319)	1:157:A:HIS:H	1:159:A:ILE:HD11	20	5.87	0.88	6.39
(1,1319)	1:157:A:HIS:H	1:159:A:ILE:HD12	20	5.87	0.88	6.39
(1,1319)	1:157:A:HIS:H	1:159:A:ILE:HD13	20	5.87	0.88	6.39
(2,85)	1:24:A:TYR:H	2:641:B:TPO:HG21	20	1.19	0.42	1.18
(2,85)	1:24:A:TYR:H	2:641:B:TPO:HG22	20	1.19	0.42	1.18
(2,85)	1:24:A:TYR:H	2:641:B:TPO:HG23	20	1.19	0.42	1.18
(2,84)	1:15:A:MET:H	2:641:B:TPO:HG21	20	0.75	0.08	0.75
(2,84)	1:15:A:MET:H	2:641:B:TPO:HG22	20	0.75	0.08	0.75
(2,84)	1:15:A:MET:H	2:641:B:TPO:HG23	20	0.75	0.08	0.75
(2,44)	1:131:A:GLN:HE21	1:153:A:ASP:H	20	0.71	0.11	0.72
(1,1227)	1:89:A:ILE:HG12	1:150:A:VAL:H	20	0.68	0.04	0.68

Continued on next page...

Continued from previous page...

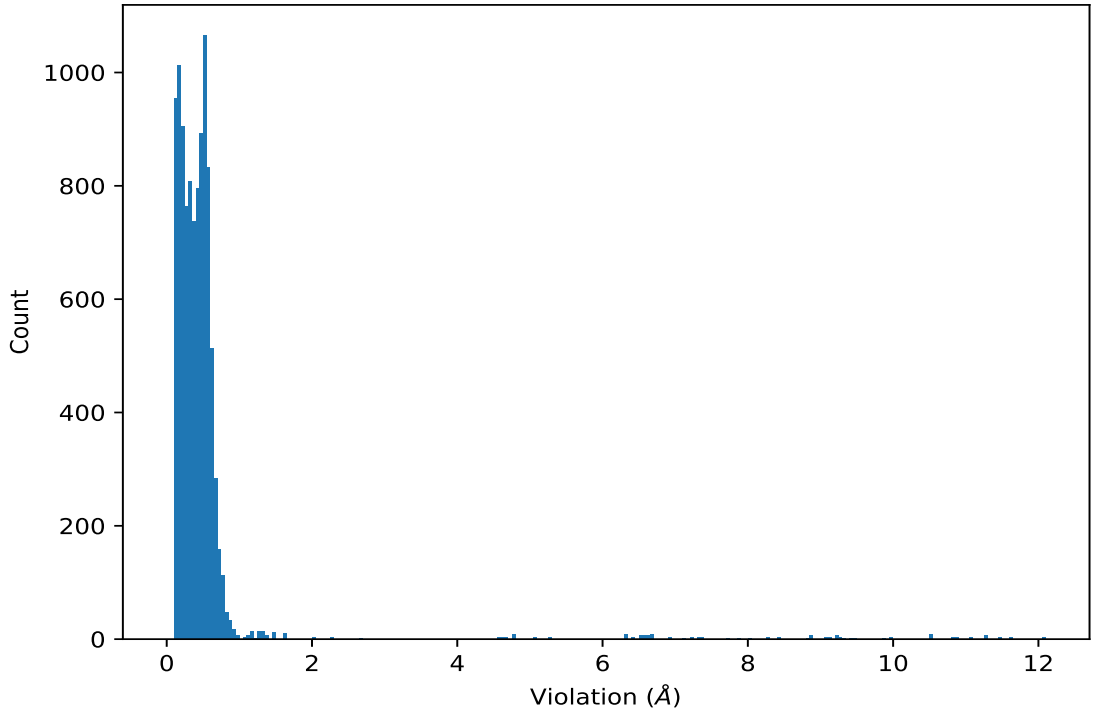
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,689)	1:86:A:LEU:HG	1:90:A:ASN:HD21	20	0.64	0.09	0.64
(2,45)	1:131:A:GLN:HE21	1:152:A:THR:HB	20	0.62	0.11	0.66
(2,103)	1:137:A:ALA:H	2:656:B:PHE:HB2	20	0.62	0.22	0.65

[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)
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG21	13	12.07
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG22	13	12.07

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG23	13	12.07
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG21	9	11.6
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG22	9	11.6
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG23	9	11.6
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG21	4	11.49
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG22	4	11.49
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG23	4	11.49
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG21	12	11.26
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG22	12	11.26
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG23	12	11.26
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG21	8	11.25
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG22	8	11.25
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG23	8	11.25
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG21	3	11.07
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG22	3	11.07
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG23	3	11.07
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG21	7	10.86
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG22	7	10.86
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG23	7	10.86
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG21	1	10.81
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG22	1	10.81
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG23	1	10.81
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG21	5	10.54
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG22	5	10.54
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG23	5	10.54
(1,2592)	1:129:A:GLN:HE21	2:651:B:ILE:HG21	14	10.51

10 Dihedral-angle violation analysis

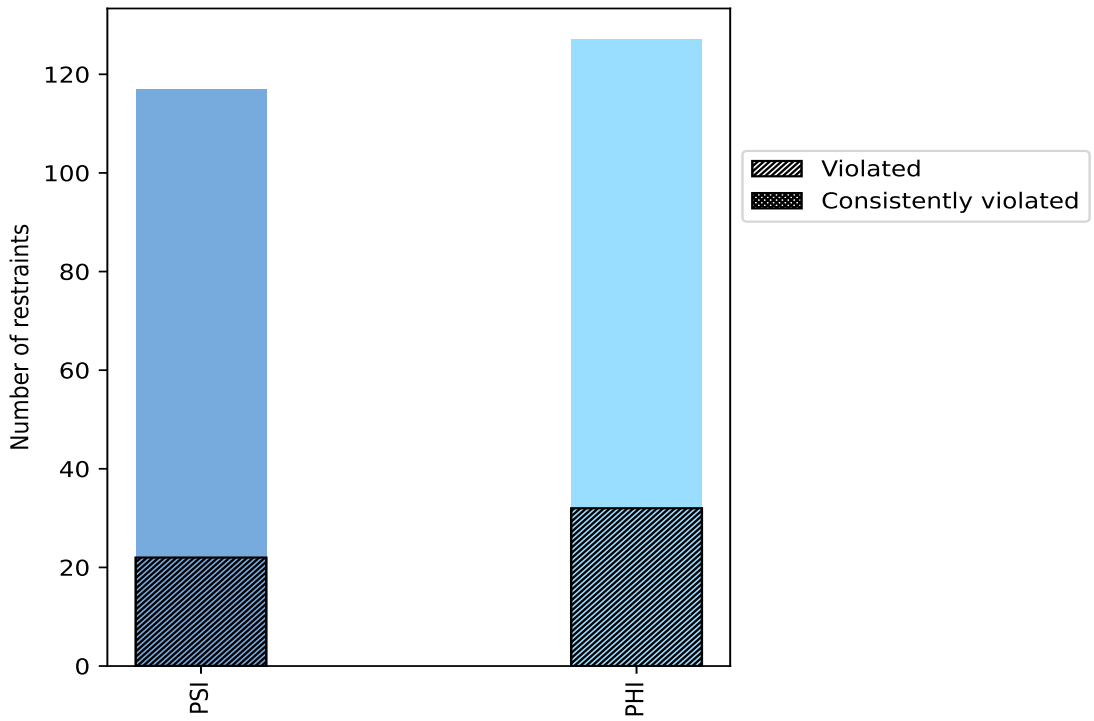
10.1 Summary of dihedral-angle violations

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	117	48.0	22	18.8	9.0	0	0.0	0.0
PHI	127	52.0	32	25.2	13.1	0	0.0	0.0
Total	244	100.0	54	22.1	22.1	0	0.0	0.0

[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



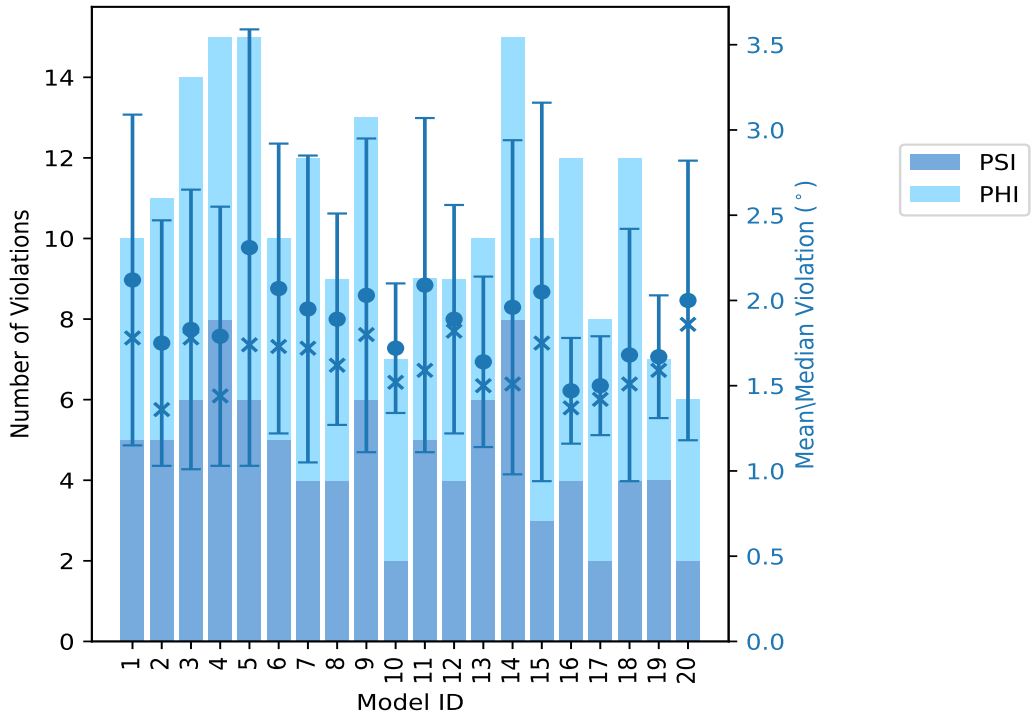
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

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 (°)
	PSI	PHI	Total				
1	5	5	10	2.12	4.52	0.97	1.78
2	5	6	11	1.75	3.45	0.72	1.36
3	6	8	14	1.83	4.56	0.82	1.78
4	8	7	15	1.79	3.34	0.76	1.44
5	6	9	15	2.31	5.69	1.28	1.74
6	5	5	10	2.07	4.2	0.85	1.73
7	4	8	12	1.95	4.06	0.9	1.72
8	4	5	9	1.89	3.18	0.62	1.62
9	6	7	13	2.03	4.38	0.92	1.8
10	2	5	7	1.72	2.3	0.38	1.52
11	5	4	9	2.09	3.26	0.98	1.59
12	4	5	9	1.89	3.44	0.67	1.82
13	6	4	10	1.64	2.44	0.5	1.5
14	8	7	15	1.96	4.09	0.98	1.51
15	3	7	10	2.05	4.89	1.11	1.75
16	4	8	12	1.47	2.1	0.31	1.37
17	2	6	8	1.5	2.2	0.29	1.42
18	4	8	12	1.68	3.69	0.74	1.51
19	4	3	7	1.67	2.3	0.36	1.59
20	2	4	6	2.0	3.59	0.82	1.86

10.2.1 Bar graph : Dihedral violation statistics for each model



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

10.3 Dihedral-angle violation statistics for the ensemble

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

Number of violated restraints			Fraction of the ensemble	
PSI	PHI	Total	Count[1]	%
12	11	23	1	5.0
1	7	8	2	10.0
0	5	5	3	15.0
3	2	5	4	20.0
1	1	2	5	25.0
0	1	1	6	30.0
1	0	1	7	35.0
0	2	2	8	40.0
0	0	0	9	45.0
1	0	1	10	50.0
0	0	0	11	55.0

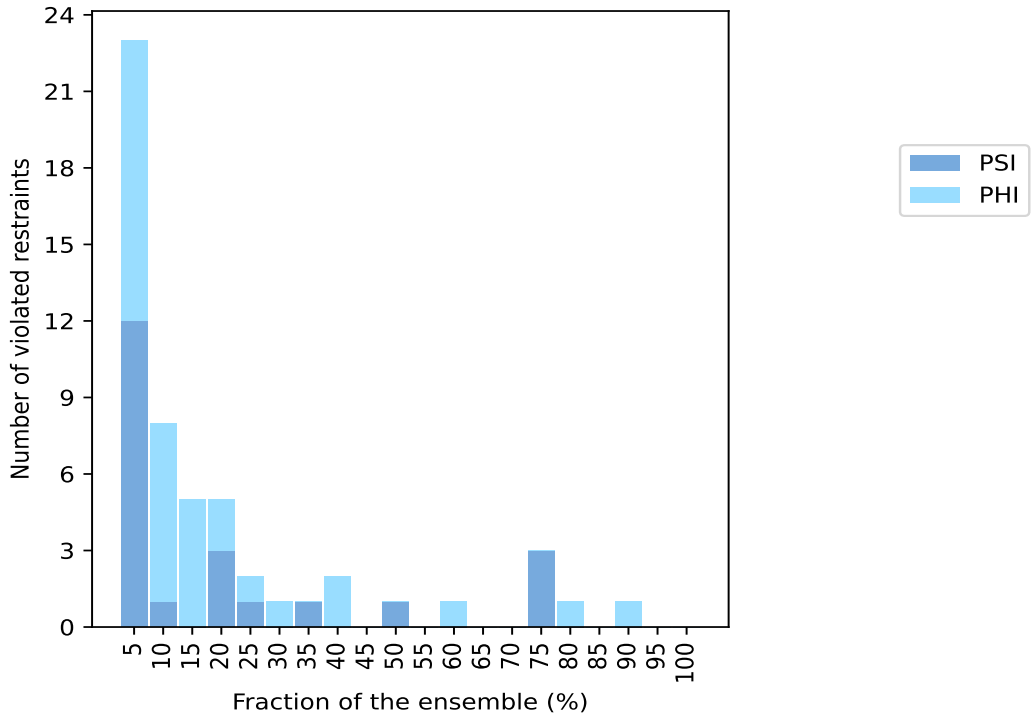
Continued on next page...

Continued from previous page...

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

[1] Number of models with violations

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

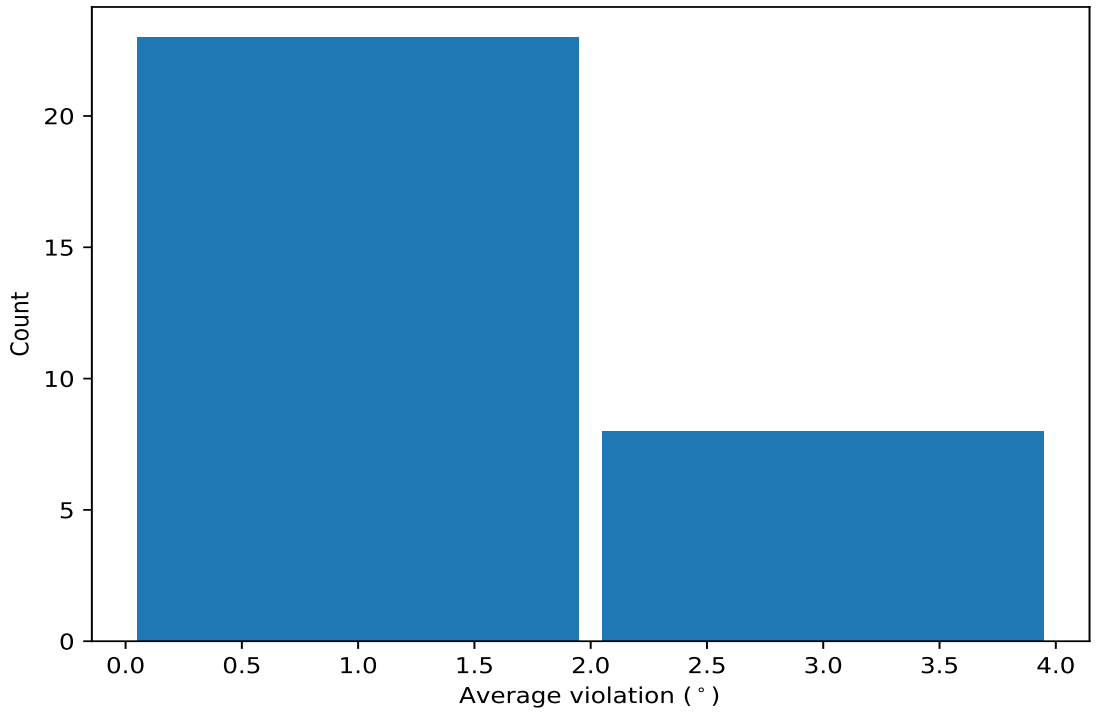


10.4 Most violated dihedral-angle restraints in the ensemble

10.4.1 Histogram : Distribution of mean dihedral-angle violations

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

in the ensemble



10.4.2 Table: Most violated dihedral-angle restraints

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

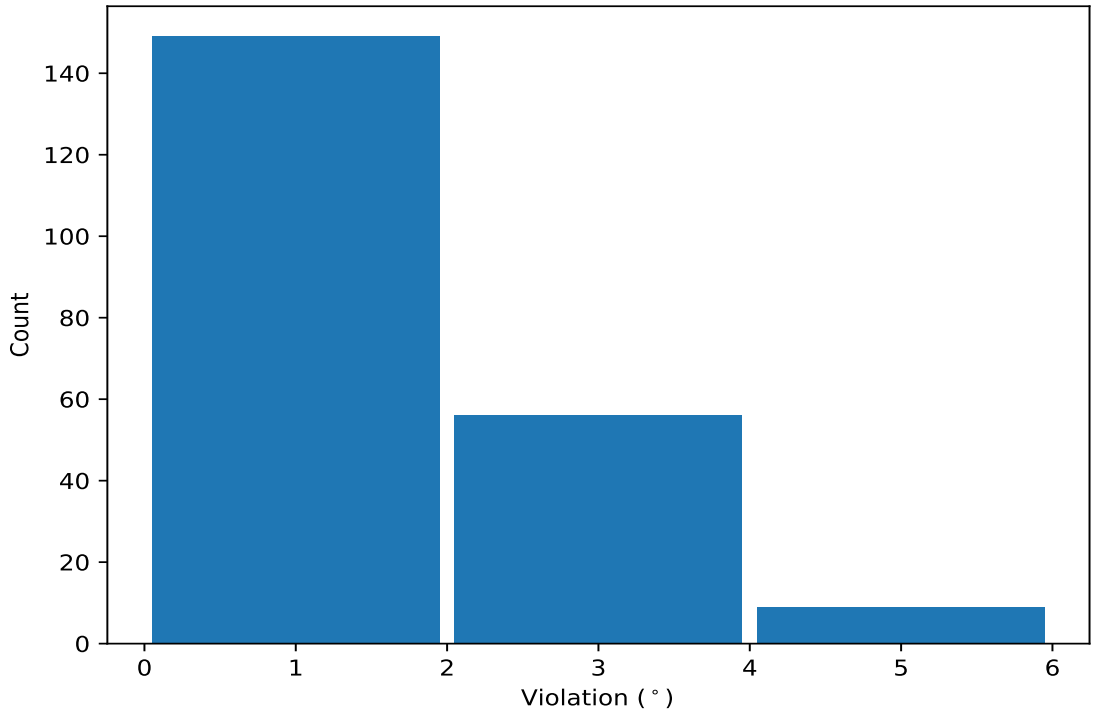
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models[1]	Mean	SD[2]	Median
(1,227)	1:151:A:PHE:C	1:152:A:THR:N	1:152:A:THR:CA	1:152:A:THR:C	18	2.09	0.6	1.92
(1,91)	1:62:A:VAL:C	1:63:A:LYS:N	1:63:A:LYS:CA	1:63:A:LYS:C	16	1.71	0.39	1.72
(1,107)	1:77:A:LYS:N	1:77:A:LYS:CA	1:77:A:LYS:C	1:78:A:ILE:N	15	2.55	1.32	1.99
(1,228)	1:152:A:THR:N	1:152:A:THR:CA	1:152:A:THR:C	1:153:A:ASP:N	15	2.49	1.04	2.13
(1,211)	1:139:A:PHE:N	1:139:A:PHE:CA	1:139:A:PHE:C	1:140:A:ALA:N	15	1.82	0.5	1.87
(1,99)	1:68:A:ARG:C	1:69:A:ARG:N	1:69:A:ARG:CA	1:69:A:ARG:C	12	3.36	1.11	3.39
(1,42)	1:30:A:ASN:N	1:30:A:ASN:CA	1:30:A:ASN:C	1:31:A:ALA:N	10	1.77	0.59	1.82
(1,220)	1:144:A:GLY:C	1:145:A:GLU:N	1:145:A:GLU:CA	1:145:A:GLU:C	8	1.68	0.44	1.86
(1,97)	1:66:A:GLN:C	1:67:A:SER:N	1:67:A:SER:CA	1:67:A:SER:C	8	1.64	0.43	1.54
(1,213)	1:140:A:ALA:N	1:140:A:ALA:CA	1:140:A:ALA:C	1:141:A:LEU:N	7	1.44	0.23	1.44

[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 (°)
(1,107)	1:77:A:LYS:N	1:77:A:LYS:CA	1:77:A:LYS:C	1:78:A:ILE:N	5	5.69
(1,107)	1:77:A:LYS:N	1:77:A:LYS:CA	1:77:A:LYS:C	1:78:A:ILE:N	15	4.89
(1,99)	1:68:A:ARG:C	1:69:A:ARG:N	1:69:A:ARG:CA	1:69:A:ARG:C	3	4.56
(1,228)	1:152:A:THR:N	1:152:A:THR:CA	1:152:A:THR:C	1:153:A:ASP:N	1	4.52
(1,99)	1:68:A:ARG:C	1:69:A:ARG:N	1:69:A:ARG:CA	1:69:A:ARG:C	5	4.5
(1,99)	1:68:A:ARG:C	1:69:A:ARG:N	1:69:A:ARG:CA	1:69:A:ARG:C	9	4.38
(1,99)	1:68:A:ARG:C	1:69:A:ARG:N	1:69:A:ARG:CA	1:69:A:ARG:C	6	4.2
(1,99)	1:68:A:ARG:C	1:69:A:ARG:N	1:69:A:ARG:CA	1:69:A:ARG:C	14	4.09
(1,228)	1:152:A:THR:N	1:152:A:THR:CA	1:152:A:THR:C	1:153:A:ASP:N	7	4.06
(1,228)	1:152:A:THR:N	1:152:A:THR:CA	1:152:A:THR:C	1:153:A:ASP:N	14	3.97