



# wwPDB NMR Structure Validation Summary Report ⓘ

Jun 6, 2023 – 05:42 AM EDT

PDB ID : 2MMB  
BMRB ID : 19851  
Title : NMR structure of the protein YP\_001712342.1 from *Acinetobacter baumannii*  
Authors : Proudfoot, A.; Serrano, P.; Geralt, M.; Wuthrich, K.; Joint Center for Structural Genomics (JCSG)  
Deposited on : 2014-03-14

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

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 ⓘ 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 ⓘ](#)) 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

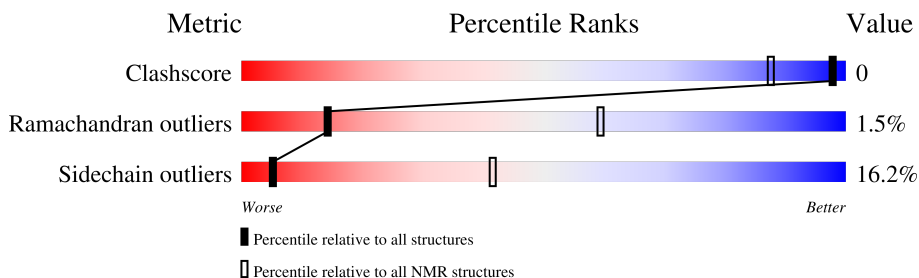
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 86%.

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  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	107	 72% 9% 19%

## 2 Ensemble composition and analysis i

This entry contains 20 models. Model 7 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:21-A:107 (87)	0.56	7

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

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

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

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

### 3 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 1559 atoms, of which 704 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Uncharacterized protein.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	107	1559	539	704	154	159	3	0

There is a discrepancy between the modelled and reference sequences:

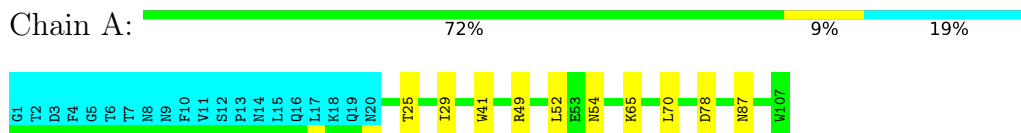
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP B0V7M7

## 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.

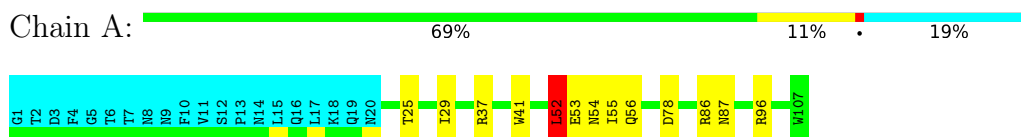
- Molecule 1: Uncharacterized protein



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

The representative model is number 7. Colouring as in section 4.1 above.

- Molecule 1: Uncharacterized protein



## 5 Refinement protocol and experimental data overview

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

Of the 80 calculated structures, 20 were deposited, based on the following criterion: *target function*.

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

Software name	Classification	Version
CYANA	refinement	
TopSpin	refinement	
UNIO	structure solution	
OPALp	refinement	
CARA	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	1231
Number of shifts mapped to atoms	1088
Number of unparsed shifts	0
Number of shifts with mapping errors	143
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	86%

## 6 Model quality i

### 6.1 Standard geometry i

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	0.58±0.01	0±0/717 ( 0.0± 0.0%)	1.04±0.03	1±1/971 ( 0.1± 0.1%)
All	All	0.58	0/14340 ( 0.0%)	1.04	20/19420 ( 0.1%)

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

Mol	Chain	Chirality	Planarity
1	A	0.0±0.0	1.0±0.9
All	All	0	20

There are no bond-length outliers.

5 of 10 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(°)	Ideal(°)	Models	
								Worst	Total
1	A	70	LEU	CB-CG-CD2	6.81	122.58	111.00	3	5
1	A	98	ARG	NE-CZ-NH2	-6.62	116.99	120.30	5	1
1	A	49	ARG	NE-CZ-NH2	-6.36	117.12	120.30	19	3
1	A	52	LEU	CB-CG-CD2	6.21	121.57	111.00	19	4
1	A	96	ARG	NE-CZ-NH1	5.97	123.28	120.30	3	1

There are no chirality outliers.

5 of 9 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	86	ARG	Sidechain	4
1	A	51	ARG	Sidechain	3
1	A	80	TYR	Sidechain	3

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	98	ARG	Sidechain	2
1	A	37	ARG	Sidechain	2

## 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	701	580	706	1±1
All	All	14020	11600	14120	12

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

5 of 10 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:81:GLU:HG2	1:A:97:ALA:HB1	0.62	1.69	12	3
1:A:81:GLU:CG	1:A:97:ALA:HB1	0.55	2.30	12	1
1:A:80:TYR:HB2	1:A:97:ALA:HB2	0.47	1.86	12	1
1:A:52:LEU:HD13	1:A:53:GLU:HG2	0.46	1.88	7	1
1:A:81:GLU:HG3	1:A:97:ALA:HB1	0.44	1.90	15	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.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	86/107 (80%)	73±2 (85±2%)	12±2 (14±3%)	1±1 (2±1%)	14	59
All	All	1720/2140 (80%)	1458 (85%)	236 (14%)	26 (2%)	14	59

5 of 12 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	61	SER	5
1	A	35	GLY	4
1	A	43	THR	4
1	A	41	TRP	3
1	A	27	LYS	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	73/91 (80%)	61±2 (84±3%)	12±2 (16±3%)	5 41
All	All	1460/1820 (80%)	1223 (84%)	237 (16%)	5 41

5 of 42 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	29	ILE	20
1	A	41	TRP	20
1	A	87	ASN	17
1	A	65	LYS	16
1	A	25	THR	14

### 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 86% for the well-defined parts and 83% for the entire structure.

### 7.1 Chemical shift list 1

File name: working\_cs.cif

Chemical shift list name: *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	1231
Number of shifts mapped to atoms	1088
Number of unparsed shifts	0
Number of shifts with mapping errors	143
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	9

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 143) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	3	ASP	HB2	2.455	0.000	.
1	A	4	PHE	HB2	3.118	0.000	.
1	A	8	ASN	HB2	2.653	0.000	.
1	A	9	ASN	HB2	2.552	0.000	.
1	A	10	PHE	HB2	2.945	0.000	.
1	A	12	SER	HB2	3.775	0.000	.
1	A	13	PRO	HB2	2.199	0.000	.
1	A	13	PRO	HG2	1.94	0.000	.
1	A	13	PRO	HD2	3.708	0.000	.
1	A	14	ASN	HB2	2.64	0.000	.
1	A	15	LEU	HB2	1.505	0.000	.
1	A	16	GLN	HB2	1.89	0.000	.
1	A	17	LEU	HB2	1.547	0.000	.
1	A	18	LYS	HB2	1.717	0.000	.

*Continued on next page...*

*Continued from previous page...*

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	18	LYS	HG2	1.344	0.000	.
1	A	19	GLN	HB2	1.866	0.000	.
1	A	19	GLN	HG2	2.248	0.000	.
1	A	20	ASN	HB2	2.719	0.000	.
1	A	22	LEU	HB2	1.404	0.000	.
1	A	23	PRO	HB2	-0.185	0.000	.
1	A	23	PRO	HG2	-0.112	0.000	.
1	A	23	PRO	HD2	2.555	0.000	.
1	A	24	PRO	HB2	2.222	0.000	.
1	A	24	PRO	HG2	1.836	0.000	.
1	A	24	PRO	HD2	3.477	0.000	.
1	A	26	PRO	HB2	2.118	0.000	.
1	A	26	PRO	HG2	2.022	0.000	.
1	A	26	PRO	HD2	3.379	0.000	.
1	A	27	LYS	HB2	1.816	0.000	.
1	A	27	LYS	HG2	1.355	0.000	.
1	A	27	LYS	HD2	1.612	0.000	.
1	A	27	LYS	HE2	2.918	0.000	.
1	A	28	ASN	HB2	2.935	0.000	.
1	A	29	ILE	HG12	1.319	0.000	.
1	A	30	PRO	HB2	1.941	0.000	.
1	A	30	PRO	HG2	1.996	0.000	.
1	A	30	PRO	HD2	3.494	0.000	.
1	A	31	LEU	HB2	1.78	0.000	.
1	A	32	PRO	HB2	2.319	0.000	.
1	A	32	PRO	HG2	2.074	0.000	.
1	A	32	PRO	HD2	3.849	0.000	.
1	A	34	PHE	HB2	2.828	0.000	.
1	A	36	GLN	HB2	1.994	0.000	.
1	A	36	GLN	HG2	2.314	0.000	.
1	A	37	ARG	HB2	1.811	0.000	.
1	A	37	ARG	HG2	1.442	0.000	.
1	A	37	ARG	HD2	3.053	0.000	.
1	A	38	ILE	HG12	1.38	0.000	.
1	A	39	ILE	HG12	-0.223	0.000	.
1	A	41	TRP	HB2	2.641	0.000	.
1	A	46	GLU	HB2	1.86	0.000	.
1	A	46	GLU	HG2	2.253	0.000	.
1	A	49	ARG	HB2	1.352	0.000	.
1	A	49	ARG	HG2	1.022	0.000	.
1	A	49	ARG	HD2	2.548	0.000	.

*Continued on next page...*

*Continued from previous page...*

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	50	GLN	HB2	1.967	0.000	.
1	A	50	GLN	HG2	2.392	0.000	.
1	A	51	ARG	HB2	1.96	0.000	.
1	A	51	ARG	HG2	1.111	0.000	.
1	A	51	ARG	HD2	3.582	0.000	.
1	A	52	LEU	HB2	2.243	0.000	.
1	A	53	GLU	HB2	1.918	0.000	.
1	A	53	GLU	HG2	2.192	0.000	.
1	A	54	ASN	HB2	2.666	0.000	.
1	A	55	ILE	HG12	1.523	0.000	.
1	A	56	GLN	HB2	2.431	0.000	.
1	A	56	GLN	HG2	2.45	0.000	.
1	A	57	PRO	HB2	2.462	0.000	.
1	A	57	PRO	HG2	2.262	0.000	.
1	A	57	PRO	HD2	3.916	0.000	.
1	A	59	ASP	HB2	2.536	0.000	.
1	A	61	SER	HB2	3.856	0.000	.
1	A	62	MET	HB2	2.193	0.000	.
1	A	62	MET	HG2	2.585	0.000	.
1	A	63	ILE	HG12	0.869	0.000	.
1	A	64	LYS	HB2	1.64	0.000	.
1	A	64	LYS	HG2	0.544	0.000	.
1	A	64	LYS	HD2	1.054	0.000	.
1	A	65	LYS	HB2	1.936	0.000	.
1	A	65	LYS	HG2	1.549	0.000	.
1	A	65	LYS	HD2	1.596	0.000	.
1	A	65	LYS	HE2	2.871	0.000	.
1	A	66	GLN	HB2	2.401	0.000	.
1	A	66	GLN	HG2	2.395	0.000	.
1	A	70	LEU	HB2	1.309	0.000	.
1	A	71	GLU	HB2	1.776	0.000	.
1	A	71	GLU	HG2	2.241	0.000	.
1	A	72	MET	HB2	2.687	0.000	.
1	A	72	MET	HG2	1.954	0.000	.
1	A	73	ILE	HG12	1.083	0.000	.
1	A	76	TRP	HB2	3.66	0.000	.
1	A	77	GLN	HB2	2.172	0.000	.
1	A	77	GLN	HG2	2.168	0.000	.
1	A	78	ASP	HB2	2.739	0.000	.
1	A	79	PHE	HB2	2.988	0.000	.
1	A	80	TYR	HB2	3.645	0.000	.

*Continued on next page...*

*Continued from previous page...*

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	81	GLU	HB2	2.232	0.000	.
1	A	81	GLU	HG2	2.596	0.000	.
1	A	82	GLN	HB2	1.928	0.000	.
1	A	82	GLN	HG2	2.264	0.000	.
1	A	83	GLU	HB2	1.892	0.000	.
1	A	83	GLU	HG2	1.743	0.000	.
1	A	84	GLN	HB2	2.026	0.000	.
1	A	84	GLN	HG2	2.392	0.000	.
1	A	85	GLN	HB2	2.062	0.000	.
1	A	85	GLN	HG2	2.395	0.000	.
1	A	86	ARG	HB2	1.745	0.000	.
1	A	86	ARG	HG2	1.522	0.000	.
1	A	86	ARG	HD2	3.053	0.000	.
1	A	87	ASN	HB2	2.836	0.000	.
1	A	88	GLU	HB2	1.916	0.000	.
1	A	88	GLU	HG2	2.18	0.000	.
1	A	89	ASN	HB2	2.66	0.000	.
1	A	90	ASN	HB2	2.963	0.000	.
1	A	91	PRO	HB2	2.314	0.000	.
1	A	91	PRO	HG2	2.003	0.000	.
1	A	91	PRO	HD2	3.935	0.000	.
1	A	94	LYS	HB2	1.534	0.000	.
1	A	94	LYS	HG2	1.142	0.000	.
1	A	94	LYS	HD2	1.361	0.000	.
1	A	94	LYS	HE2	2.536	0.000	.
1	A	95	TYR	HB2	2.684	0.000	.
1	A	96	ARG	HB2	1.232	0.000	.
1	A	96	ARG	HG2	0.641	0.000	.
1	A	96	ARG	HD2	1.227	0.000	.
1	A	98	ARG	HB2	1.926	0.000	.
1	A	98	ARG	HG2	1.691	0.000	.
1	A	98	ARG	HD2	2.805	0.000	.
1	A	99	LEU	HB2	2.272	0.000	.
1	A	100	MET	HB2	2.326	0.000	.
1	A	100	MET	HG2	3.441	0.000	.
1	A	101	LYS	HB2	1.71	0.000	.
1	A	101	LYS	HG2	1.386	0.000	.
1	A	101	LYS	HD2	1.577	0.000	.
1	A	101	LYS	HE2	2.96	0.000	.
1	A	102	LYS	HB2	2.066	0.000	.
1	A	102	LYS	HG2	1.298	0.000	.

*Continued on next page...*

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	102	LYS	HD2	1.733	0.000	.
1	A	102	LYS	HE2	2.903	0.000	.
1	A	103	ILE	HG12	0.715	0.000	.
1	A	105	ASP	HB2	3.01	0.000	.
1	A	106	LEU	HB2	1.602	0.000	.
1	A	107	TRP	HB2	3.124	0.000	.

### 7.1.2 Chemical shift referencing [i](#)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	105	2.40 $\pm$ 0.09	Should be checked
$^{13}\text{C}_\beta$	98	2.80 $\pm$ 0.05	Should be checked
$^{13}\text{C}'$	0	—	None (insufficient data)
$^{15}\text{N}$	97	0.18 $\pm$ 0.28	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 86%, i.e. 1054 atoms were assigned a chemical shift out of a possible 1230. 0 out of 8 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	340/427 (80%)	173/173 (100%)	87/174 (50%)	80/80 (100%)
Sidechain	644/729 (88%)	438/470 (93%)	193/222 (87%)	13/37 (35%)
Aromatic	70/74 (95%)	35/36 (97%)	32/35 (91%)	3/3 (100%)
Overall	1054/1230 (86%)	646/679 (95%)	312/431 (72%)	96/120 (80%)

### 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	80	TYR	CG	36.18	112.42 – 146.96	-27.1
1	A	96	ARG	HD3	0.42	1.81 – 4.39	-10.4
1	A	96	ARG	HD2	1.23	1.97 – 4.26	-8.3

Continued on next page...

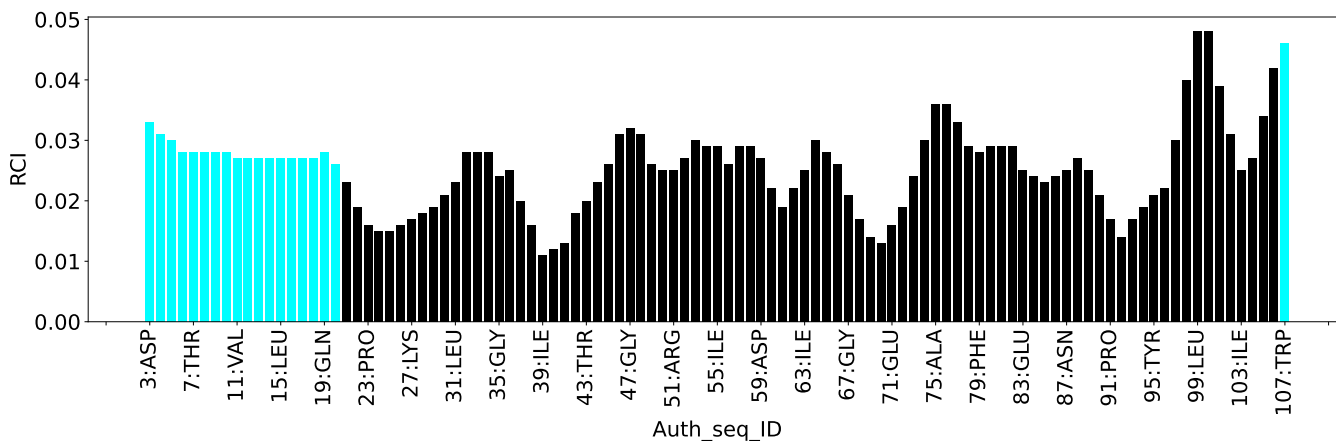
Continued from previous page...

List Id	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	35	GLY	HA2	1.34	2.15 – 5.77	-7.2
1	A	23	PRO	HG2	-0.11	0.41 – 3.45	-6.7
1	A	23	PRO	HB2	-0.18	0.37 – 3.78	-6.6
1	A	23	PRO	HD3	1.27	1.76 – 5.48	-6.3
1	A	96	ARG	HB3	0.24	0.43 – 3.11	-5.7
1	A	23	PRO	CB	25.82	26.06 – 37.61	-5.2

### 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:





## 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	2161
Intra-residue ( $ i-j =0$ )	516
Sequential ( $ i-j =1$ )	630
Medium range ( $ i-j >1$ and $ i-j <5$ )	534
Long range ( $ i-j \geq 5$ )	481
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	711
Number of restraints per residue	20.2
Number of long range restraints per residue <sup>1</sup>	4.5

<sup>1</sup>Long range hydrogen bonds and disulfide bonds are counted as long range restraints while calculating the number of long range restraints per residue

### 8.2 Residual restraint violations

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

#### 8.2.1 Average number of distance violations per model

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

Bins (Å)	Average number of violations per model	Max (Å)
0.1-0.2 (Small)	26.6	0.2
0.2-0.5 (Medium)	59.4	0.5
>0.5 (Large)	101.3	3.61

### 8.2.2 Average number of dihedral-angle violations per model

Dihedral-angle violations less than  $1^\circ$  are not included in the calculation. There are no dihedral-angle violations

## 9 Distance violation analysis [\(i\)](#)

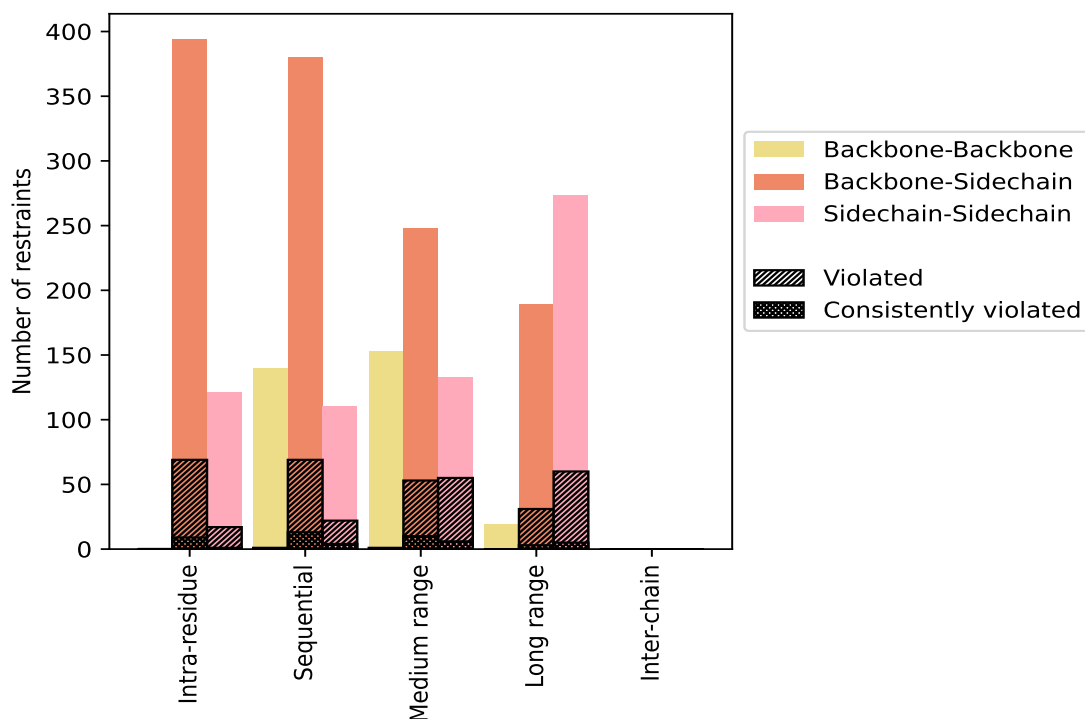
### 9.1 Summary of distance violations [\(i\)](#)

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.

Restrains type	Count	% <sup>1</sup>	Violated <sup>3</sup>			Consistently Violated <sup>4</sup>		
			Count	% <sup>2</sup>	% <sup>1</sup>	Count	% <sup>2</sup>	% <sup>1</sup>
<b>Intra-residue (<math> i-j =0</math>)</b>	<b>516</b>	<b>23.9</b>	<b>86</b>	<b>16.7</b>	<b>4.0</b>	<b>10</b>	<b>1.9</b>	<b>0.5</b>
Backbone-Backbone	1	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	394	18.2	69	17.5	3.2	9	2.3	0.4
Sidechain-Sidechain	121	5.6	17	14.0	0.8	1	0.8	0.0
<b>Sequential (<math> i-j =1</math>)</b>	<b>630</b>	<b>29.2</b>	<b>92</b>	<b>14.6</b>	<b>4.3</b>	<b>17</b>	<b>2.7</b>	<b>0.8</b>
Backbone-Backbone	140	6.5	1	0.7	0.0	0	0.0	0.0
Backbone-Sidechain	380	17.6	69	18.2	3.2	13	3.4	0.6
Sidechain-Sidechain	110	5.1	22	20.0	1.0	4	3.6	0.2
<b>Medium range (<math> i-j &gt;1</math> &amp; <math> i-j &lt;5</math>)</b>	<b>534</b>	<b>24.7</b>	<b>109</b>	<b>20.4</b>	<b>5.0</b>	<b>16</b>	<b>3.0</b>	<b>0.7</b>
Backbone-Backbone	153	7.1	1	0.7	0.0	0	0.0	0.0
Backbone-Sidechain	248	11.5	53	21.4	2.5	10	4.0	0.5
Sidechain-Sidechain	133	6.2	55	41.4	2.5	6	4.5	0.3
<b>Long range (<math> i-j \geq 5</math>)</b>	<b>481</b>	<b>22.3</b>	<b>91</b>	<b>18.9</b>	<b>4.2</b>	<b>8</b>	<b>1.7</b>	<b>0.4</b>
Backbone-Backbone	19	0.9	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	189	8.7	31	16.4	1.4	3	1.6	0.1
Sidechain-Sidechain	273	12.6	60	22.0	2.8	5	1.8	0.2
<b>Inter-chain</b>	<b>0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
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
<b>Hydrogen bond</b>	<b>0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
<b>Disulfide bond</b>	<b>0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
<b>Total</b>	<b>2161</b>	<b>100.0</b>	<b>378</b>	<b>17.5</b>	<b>17.5</b>	<b>51</b>	<b>2.4</b>	<b>2.4</b>
Backbone-Backbone	313	14.5	2	0.6	0.1	0	0.0	0.0
Backbone-Sidechain	1211	56.0	222	18.3	10.3	35	2.9	1.6
Sidechain-Sidechain	637	29.5	154	24.2	7.1	16	2.5	0.7

<sup>1</sup> percentage calculated with respect to the total number of distance restraints, <sup>2</sup> percentage calculated with respect to the number of restraints in a particular restraint category, <sup>3</sup> violated in at least one model, <sup>4</sup> 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 disulfid 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					Total	Mean (Å)	Max (Å)	SD <sup>6</sup> (Å)	Median (Å)
	IR <sup>1</sup>	SQ <sup>2</sup>	MR <sup>3</sup>	LR <sup>4</sup>	IC <sup>5</sup>					
1	40	55	60	50	0	205	0.68	3.16	0.56	0.48
2	41	45	50	40	0	176	0.73	2.75	0.51	0.57
3	41	45	51	48	0	185	0.73	2.9	0.56	0.56
4	42	50	54	54	0	200	0.73	2.88	0.52	0.59
5	45	51	51	44	0	191	0.69	2.54	0.5	0.58
6	39	48	53	38	0	178	0.73	3.36	0.56	0.62
7	38	55	56	32	0	181	0.72	2.66	0.52	0.57
8	37	43	55	57	0	192	0.8	2.49	0.55	0.64
9	35	40	43	38	0	156	0.73	2.72	0.56	0.56
10	41	45	53	48	0	187	0.71	2.82	0.57	0.56
11	45	48	48	45	0	186	0.68	2.61	0.5	0.55

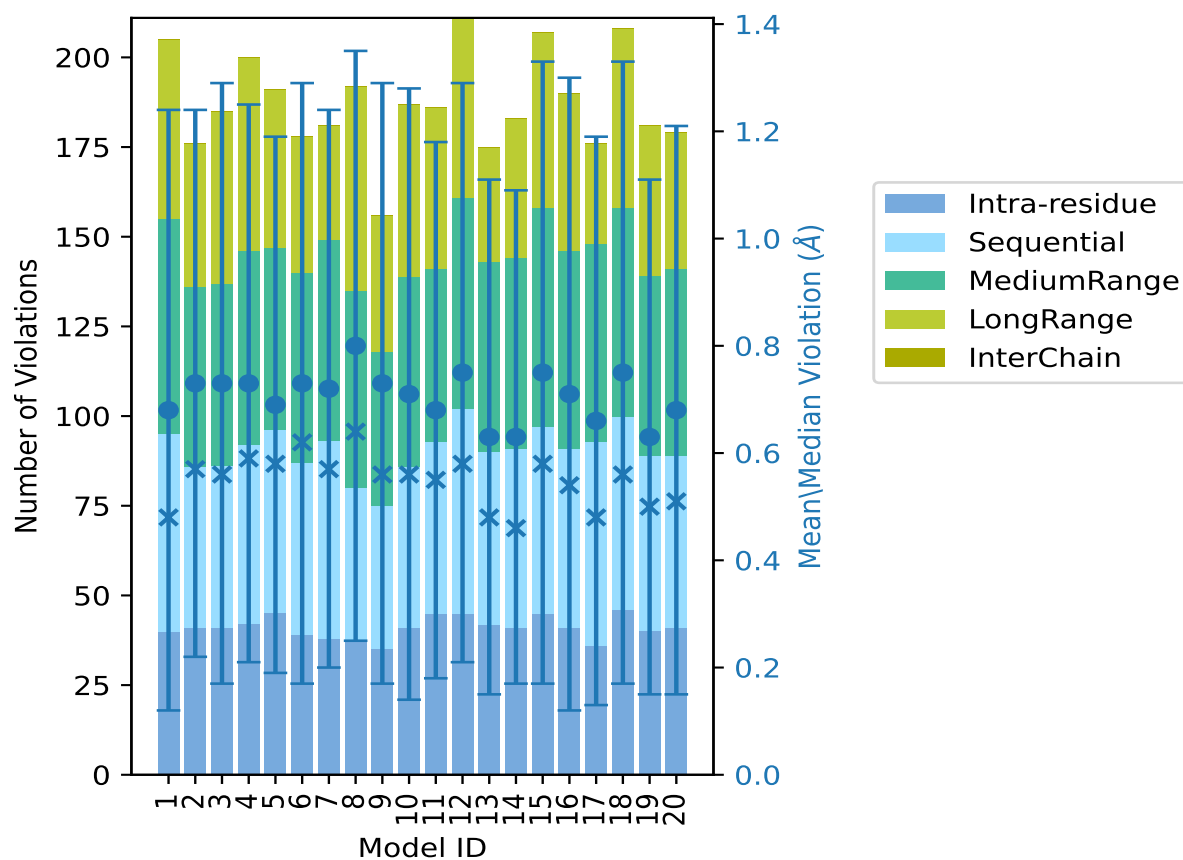
*Continued on next page...*

Continued from previous page...

Model ID	Number of violations					Total	Mean (Å)	Max (Å)	SD <sup>6</sup> (Å)	Median (Å)
	IR <sup>1</sup>	SQ <sup>2</sup>	MR <sup>3</sup>	LR <sup>4</sup>	IC <sup>5</sup>					
12	45	57	59	50	0	211	0.75	2.72	0.54	0.58
13	42	48	53	32	0	175	0.63	2.28	0.48	0.48
14	41	50	53	39	0	183	0.63	1.97	0.46	0.46
15	45	52	61	49	0	207	0.75	3.49	0.58	0.58
16	41	50	55	44	0	190	0.71	3.58	0.59	0.54
17	36	57	55	28	0	176	0.66	3.0	0.53	0.48
18	46	54	58	50	0	208	0.75	3.61	0.58	0.56
19	40	49	50	42	0	181	0.63	2.86	0.48	0.5
20	41	48	52	38	0	179	0.68	2.57	0.53	0.51

<sup>1</sup>Intra-residue restraints, <sup>2</sup>Sequential restraints, <sup>3</sup>Medium range restraints, <sup>4</sup>Long range restraints, <sup>5</sup>Inter-chain restraints, <sup>6</sup>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

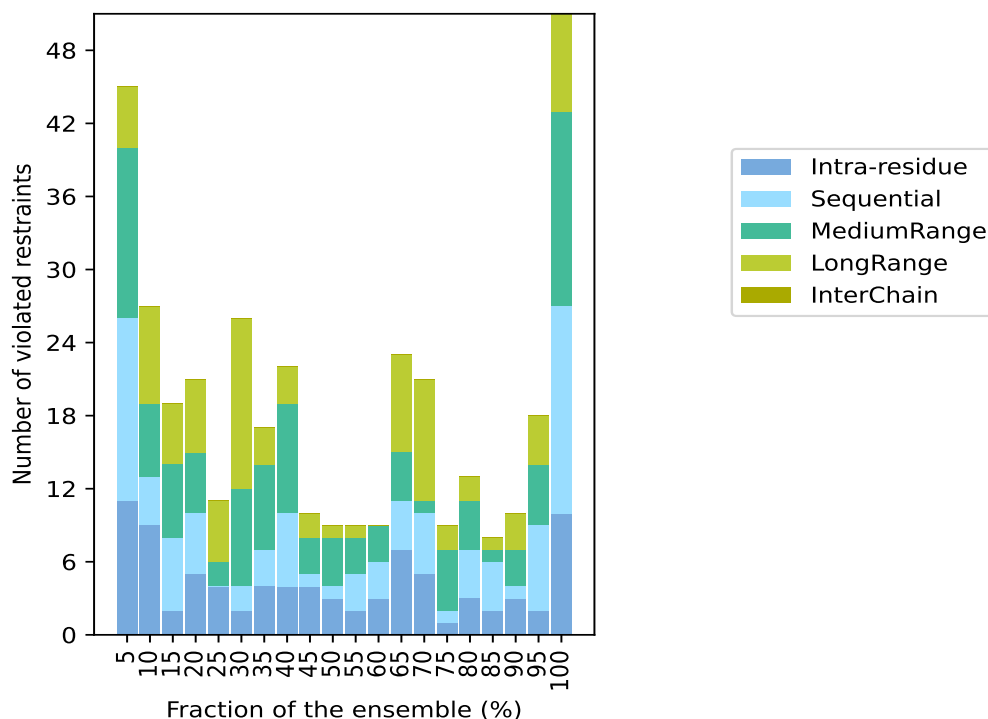
### 9.3 Distance violation statistics for the ensemble

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, 1783(IR:430, SQ:538, MR:425, LR:390, IC:0) restraints are not violated in the ensemble.

Number of violated restraints						Fraction of the ensemble	
IR <sup>1</sup>	SQ <sup>2</sup>	MR <sup>3</sup>	LR <sup>4</sup>	IC <sup>5</sup>	Total	Count <sup>6</sup>	%
11	15	14	5	0	45	1	5.0
9	4	6	8	0	27	2	10.0
2	6	6	5	0	19	3	15.0
5	5	5	6	0	21	4	20.0
4	0	2	5	0	11	5	25.0
2	2	8	14	0	26	6	30.0
4	3	7	3	0	17	7	35.0
4	6	9	3	0	22	8	40.0
4	1	3	2	0	10	9	45.0
3	1	4	1	0	9	10	50.0
2	3	3	1	0	9	11	55.0
3	3	3	0	0	9	12	60.0
7	4	4	8	0	23	13	65.0
5	5	1	10	0	21	14	70.0
1	1	5	2	0	9	15	75.0
3	4	4	2	0	13	16	80.0
2	4	1	1	0	8	17	85.0
3	1	3	3	0	10	18	90.0
2	7	5	4	0	18	19	95.0
10	17	16	8	0	51	20	100.0

<sup>1</sup>Intra-residue restraints, <sup>2</sup>Sequential restraints, <sup>3</sup>Medium range restraints, <sup>4</sup>Long range restraints, <sup>5</sup>Inter-chain restraints, <sup>6</sup> Number of models with violations

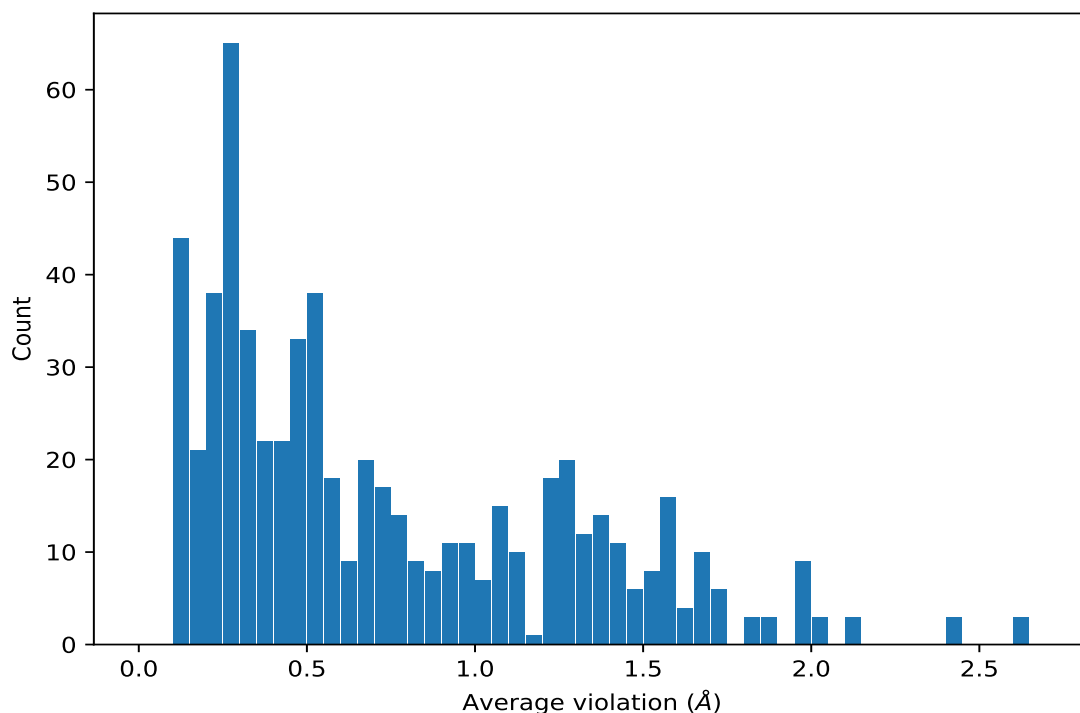
### 9.3.1 Bar graph : Distance violation statistics for the ensemble [\(i\)](#)



## 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 <sup>1</sup>	Mean (Å)	SD <sup>1</sup> (Å)	Median (Å)
(1,1344)	1:A:103:ILE:H	1:A:106:LEU:HD21	20	1.71	0.26	1.8
(1,1344)	1:A:103:ILE:H	1:A:106:LEU:HD22	20	1.71	0.26	1.8
(1,1344)	1:A:103:ILE:H	1:A:106:LEU:HD23	20	1.71	0.26	1.8
(1,986)	1:A:30:PRO:HB3	1:A:32:PRO:HD3	20	1.68	0.04	1.66
(1,1358)	1:A:103:ILE:HA	1:A:106:LEU:HD21	20	1.65	0.25	1.67
(1,1358)	1:A:103:ILE:HA	1:A:106:LEU:HD22	20	1.65	0.25	1.67
(1,1358)	1:A:103:ILE:HA	1:A:106:LEU:HD23	20	1.65	0.25	1.67
(1,683)	1:A:35:GLY:H	1:A:37:ARG:HB3	20	1.61	0.28	1.74
(1,832)	1:A:74:THR:HA	1:A:78:ASP:HB3	20	1.54	0.23	1.58
(1,803)	1:A:15:LEU:HB3	1:A:17:LEU:HD21	20	1.49	1.01	0.9
(1,803)	1:A:15:LEU:HB3	1:A:17:LEU:HD22	20	1.49	1.01	0.9
(1,803)	1:A:15:LEU:HB3	1:A:17:LEU:HD23	20	1.49	1.01	0.9
(1,1229)	1:A:81:GLU:HA	1:A:84:GLN:HB3	20	1.43	0.64	1.75
(1,871)	1:A:29:ILE:HG21	1:A:34:PHE:HB3	20	1.42	0.08	1.41
(1,871)	1:A:29:ILE:HG22	1:A:34:PHE:HB3	20	1.42	0.08	1.41
(1,871)	1:A:29:ILE:HG23	1:A:34:PHE:HB3	20	1.42	0.08	1.41

*Continued on next page...*



Continued from previous page...

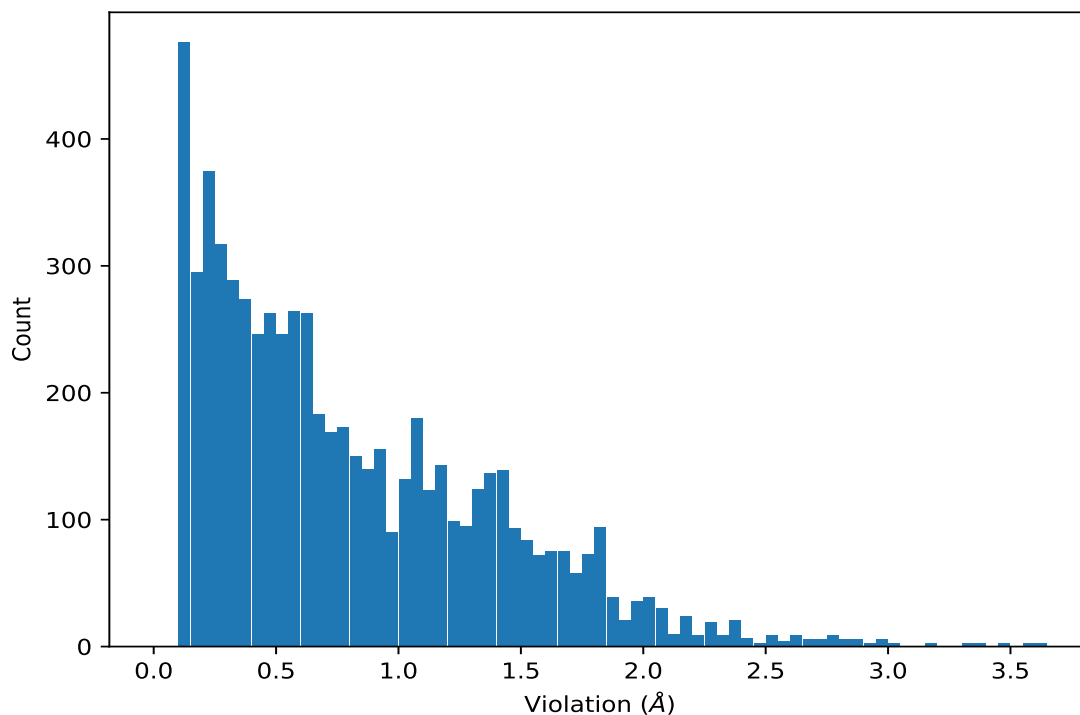
Key	Atom-1	Atom-2	Models <sup>1</sup>	Mean (Å)	SD <sup>1</sup> (Å)	Median (Å)
(1,1865)	1:A:33:ALA:HA	1:A:37:ARG:HB3	20	1.35	0.31	1.4
(1,1629)	1:A:92:THR:HG21	1:A:95:TYR:HB3	20	1.32	0.12	1.37

<sup>1</sup>Number of violated models, <sup>2</sup>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)
(1,1402)	1:A:14:ASN:HB3	1:A:17:LEU:HD21	18	3.61
(1,1402)	1:A:14:ASN:HB3	1:A:17:LEU:HD22	18	3.61
(1,1402)	1:A:14:ASN:HB3	1:A:17:LEU:HD23	18	3.61

Continued on next page...

*Continued from previous page...*

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,1402)	1:A:14:ASN:HB3	1:A:17:LEU:HD21	16	3.58
(1,1402)	1:A:14:ASN:HB3	1:A:17:LEU:HD22	16	3.58
(1,1402)	1:A:14:ASN:HB3	1:A:17:LEU:HD23	16	3.58
(1,585)	1:A:19:GLN:HE22	1:A:21:VAL:HG21	15	3.49
(1,585)	1:A:19:GLN:HE22	1:A:21:VAL:HG22	15	3.49
(1,585)	1:A:19:GLN:HE22	1:A:21:VAL:HG23	15	3.49
(1,1443)	1:A:52:LEU:HD21	1:A:102:LYS:HE3	6	3.36
(1,1443)	1:A:52:LEU:HD22	1:A:102:LYS:HE3	6	3.36
(1,1443)	1:A:52:LEU:HD23	1:A:102:LYS:HE3	6	3.36
(1,1443)	1:A:52:LEU:HD21	1:A:102:LYS:HE3	16	3.3
(1,1443)	1:A:52:LEU:HD22	1:A:102:LYS:HE3	16	3.3
(1,1443)	1:A:52:LEU:HD23	1:A:102:LYS:HE3	16	3.3
(1,803)	1:A:15:LEU:HB3	1:A:17:LEU:HD21	1	3.16
(1,803)	1:A:15:LEU:HB3	1:A:17:LEU:HD22	1	3.16
(1,803)	1:A:15:LEU:HB3	1:A:17:LEU:HD23	1	3.16
(1,983)	1:A:19:GLN:HG3	1:A:21:VAL:HG21	17	3.0
(1,983)	1:A:19:GLN:HG3	1:A:21:VAL:HG22	17	3.0
(1,983)	1:A:19:GLN:HG3	1:A:21:VAL:HG23	17	3.0
(1,1402)	1:A:14:ASN:HB3	1:A:17:LEU:HD21	6	2.98
(1,1402)	1:A:14:ASN:HB3	1:A:17:LEU:HD22	6	2.98
(1,1402)	1:A:14:ASN:HB3	1:A:17:LEU:HD23	6	2.98
(1,1443)	1:A:52:LEU:HD21	1:A:102:LYS:HE3	1	2.97
(1,1443)	1:A:52:LEU:HD22	1:A:102:LYS:HE3	1	2.97
(1,1443)	1:A:52:LEU:HD23	1:A:102:LYS:HE3	1	2.97
(1,1443)	1:A:52:LEU:HD21	1:A:102:LYS:HE3	3	2.9

## 10 Dihedral-angle violation analysis

No dihedral-angle restraints found