wwPDB NMR Structure Validation Summary Report

Jun 3, 2023 – 07:33 PM EDT

PDB ID	:	2N2E
BMRB ID	:	25194
Title	:	NMR solution structure of the C-terminal domain of NisI, a lipoprotein from Lactococcus lactis which confers immunity against nisin
Authors	:	Hacker, C.; Christ, N.A.; Korn, S.; Duchardt-Ferner, E.; Hellmich, U.A.; Duesterhus, S.; Koetter, P.; Entian, K.; Woehnert, J.
Deposited on	:	2015-05-08

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:

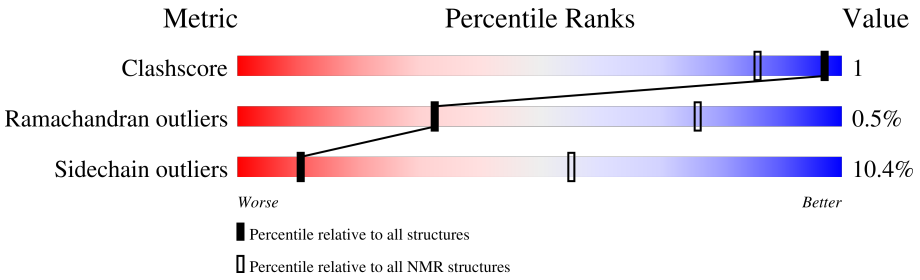
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal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

1 Overall quality at a glance

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

The overall completeness of chemical shifts assignment is 89%.

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



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

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

Mol	Chain	Length	Quality of chain
1	A	129	67% 5% 12% 16%

2 Ensemble composition and analysis

This entry contains 20 models. Model 2 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *minimized average structure*.

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:120-A:176, A:190-A:224 (92)	0.40	2

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 4 clusters and 5 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1673 atoms, of which 814 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Nisin immunity protein.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	108	1673	530	814	146	182	1	0

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: Nisin immunity protein

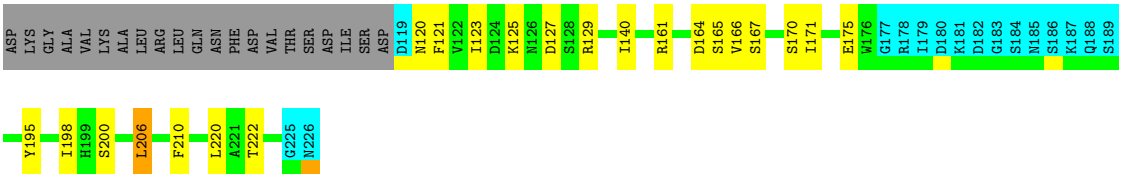
Chain A:

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

The representative model is number 2. Colouring as in section 4.1 above.

- Molecule 1: Nisin immunity protein

Chain A:



5 Refinement protocol and experimental data overview

The models were refined using the following method: *torsion angle dynamics, water refinement.*

Of the 100 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
TopSpin	structure solution	
TopSpin	structure solution	
TopSpin	structure solution	
TopSpin	structure solution	
CYANA	refinement	
OPALp	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	1511
Number of shifts mapped to atoms	1267
Number of unparsed shifts	0
Number of shifts with mapping errors	244
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	89%

6 Model quality

6.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the (average) root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	0.59±0.09	0±0/757 (0.0± 0.0%)	1.06±0.17	1±1/1024 (0.1± 0.1%)
All	All	0.59	0/15140 (0.0%)	1.07	19/20480 (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	1.3±1.1
All	All	0	26

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(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	195	TYR	CB-CG-CD2	-6.70	116.98	121.00	7	4
1	A	173	ARG	NE-CZ-NH2	-6.40	117.10	120.30	16	2
1	A	195	TYR	C-N-CA	6.22	135.35	122.30	9	6
1	A	161	ARG	NE-CZ-NH2	-5.80	117.40	120.30	17	1
1	A	167	SER	N-CA-CB	-5.79	101.81	110.50	17	1

There are no chirality outliers.

5 of 8 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	161	ARG	Sidechain	5
1	A	202	ARG	Sidechain	4
1	A	190	ARG	Sidechain	4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	138	TYR	Sidechain	4
1	A	121	PHE	Sidechain	4

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	743	708	708	1±2
All	All	14860	14160	14160	18

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 1.

5 of 13 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:123:ILE:HG21	1:A:195:TYR:OH	0.62	1.94	2	1
1:A:129:ARG:O	1:A:140:ILE:HD12	0.61	1.94	2	1
1:A:121:PHE:CD2	1:A:198:ILE:HD13	0.53	2.37	2	1
1:A:210:PHE:O	1:A:220:LEU:HD22	0.53	2.02	2	1
1:A:164:ASP:CG	1:A:166:VAL:HG22	0.49	2.27	2	2

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	92/129 (71%)	86±1 (93±1%)	6±1 (6±2%)	0±1 (0±1%)	32	76
All	All	1840/2580 (71%)	1717 (93%)	114 (6%)	9 (0%)	32	76

All 3 unique Ramachandran outliers are listed below. They are sorted by the frequency of occur-

rence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	158	ALA	7
1	A	151	GLY	1
1	A	141	SER	1

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	83/114 (73%)	74±2 (90±2%)	9±2 (10±2%)	10	55
All	All	1660/2280 (73%)	1488 (90%)	172 (10%)	10	55

5 of 35 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	120	ASN	20
1	A	206	LEU	18
1	A	171	ILE	15
1	A	127	ASP	13
1	A	161	ARG	11

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

There are no ligands in this entry.

6.7 Other polymers

There are no such molecules in this entry.

6.8 Polymer linkage issues

There are no chain breaks in this entry.

7 Chemical shift validation

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

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

7.1.1 Bookkeeping

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

Total number of shifts	1511
Number of shifts mapped to atoms	1267
Number of unparsed shifts	0
Number of shifts with mapping errors	244
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	7

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 244) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	98	ASP	HA	4.692	0.05	1
1	A	98	ASP	HB2	2.72	0.05	2
1	A	98	ASP	HB3	2.656	0.05	2
1	A	98	ASP	C	176.394	0.20	1
1	A	98	ASP	CA	54.37	0.20	1
1	A	98	ASP	CB	41.171	0.20	1
1	A	99	LYS	H	8.499	0.05	1
1	A	99	LYS	HA	4.265	0.05	1
1	A	99	LYS	HB2	1.881	0.05	2
1	A	99	LYS	HB3	1.787	0.05	2
1	A	99	LYS	HG2	1.442	0.05	1
1	A	99	LYS	HG3	1.442	0.05	1
1	A	99	LYS	HD2	1.676	0.05	1
1	A	99	LYS	HD3	1.676	0.05	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	99	LYS	HE2	2.99	0.05	1
1	A	99	LYS	HE3	2.99	0.05	1
1	A	99	LYS	C	177.405	0.20	1
1	A	99	LYS	CA	56.976	0.20	1
1	A	99	LYS	CB	32.693	0.20	1
1	A	99	LYS	CG	24.746	0.20	1
1	A	99	LYS	CD	29.01	0.20	1
1	A	99	LYS	CE	42.025	0.20	1
1	A	99	LYS	N	122.39	0.30	1
1	A	100	GLY	H	8.495	0.05	1
1	A	100	GLY	HA2	3.923	0.05	1
1	A	100	GLY	HA3	3.923	0.05	1
1	A	100	GLY	C	173.815	0.20	1
1	A	100	GLY	CA	45.486	0.20	1
1	A	100	GLY	N	109.651	0.30	1
1	A	101	ALA	H	8.055	0.05	1
1	A	101	ALA	HA	4.319	0.05	1
1	A	101	ALA	HB1	1.394	0.05	1
1	A	101	ALA	HB2	1.394	0.05	1
1	A	101	ALA	HB3	1.394	0.05	1
1	A	101	ALA	C	178.104	0.20	1
1	A	101	ALA	CA	52.763	0.20	1
1	A	101	ALA	CB	19.263	0.20	1
1	A	101	ALA	N	123.889	0.30	1
1	A	102	VAL	H	8.051	0.05	1
1	A	102	VAL	HA	4.016	0.05	1
1	A	102	VAL	HB	2.072	0.05	1
1	A	102	VAL	HG11	0.936	0.05	1
1	A	102	VAL	HG12	0.936	0.05	1
1	A	102	VAL	HG13	0.936	0.05	1
1	A	102	VAL	HG21	0.936	0.05	1
1	A	102	VAL	HG22	0.936	0.05	1
1	A	102	VAL	HG23	0.936	0.05	1
1	A	102	VAL	C	176.625	0.20	1
1	A	102	VAL	CA	62.896	0.20	1
1	A	102	VAL	CB	32.6	0.20	1
1	A	102	VAL	CG1	21.172	0.20	1
1	A	102	VAL	N	119.472	0.30	1
1	A	103	LYS	H	8.313	0.05	1
1	A	103	LYS	HA	4.254	0.05	1
1	A	103	LYS	HB2	1.866	0.05	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	103	LYS	HB3	1.785	0.05	2
1	A	103	LYS	HG2	1.437	0.05	1
1	A	103	LYS	HG3	1.437	0.05	1
1	A	103	LYS	HD2	1.677	0.05	1
1	A	103	LYS	HD3	1.677	0.05	1
1	A	103	LYS	HE2	2.991	0.05	1
1	A	103	LYS	HE3	2.991	0.05	1
1	A	103	LYS	C	176.515	0.20	1
1	A	103	LYS	CA	56.931	0.20	1
1	A	103	LYS	CB	32.685	0.20	1
1	A	103	LYS	CG	24.675	0.20	1
1	A	103	LYS	CD	29.098	0.20	1
1	A	103	LYS	CE	42.025	0.20	1
1	A	103	LYS	N	124.993	0.30	1
1	A	104	ALA	H	8.204	0.05	1
1	A	104	ALA	HA	4.265	0.05	1
1	A	104	ALA	HB1	1.385	0.05	1
1	A	104	ALA	HB2	1.385	0.05	1
1	A	104	ALA	HB3	1.385	0.05	1
1	A	104	ALA	C	177.816	0.20	1
1	A	104	ALA	CA	52.716	0.20	1
1	A	104	ALA	CB	19.12	0.20	1
1	A	104	ALA	N	124.495	0.30	1
1	A	105	LEU	H	8.101	0.05	1
1	A	105	LEU	HA	4.272	0.05	1
1	A	105	LEU	HB2	1.658	0.05	2
1	A	105	LEU	HB3	1.581	0.05	2
1	A	105	LEU	HD11	0.922	0.05	2
1	A	105	LEU	HD12	0.922	0.05	2
1	A	105	LEU	HD13	0.922	0.05	2
1	A	105	LEU	HD21	0.873	0.05	2
1	A	105	LEU	HD22	0.873	0.05	2
1	A	105	LEU	HD23	0.873	0.05	2
1	A	105	LEU	C	177.521	0.20	1
1	A	105	LEU	CA	55.458	0.20	1
1	A	105	LEU	CB	42.317	0.20	1
1	A	105	LEU	CD1	24.999	0.20	1
1	A	105	LEU	CD2	23.424	0.20	1
1	A	105	LEU	N	121.309	0.30	1
1	A	106	ARG	H	8.232	0.05	1
1	A	106	ARG	HA	4.319	0.05	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	106	ARG	HB2	1.819	0.05	1
1	A	106	ARG	HB3	1.819	0.05	1
1	A	106	ARG	HG2	1.587	0.05	1
1	A	106	ARG	HG3	1.632	0.05	1
1	A	106	ARG	HD2	3.186	0.05	1
1	A	106	ARG	HD3	3.186	0.05	1
1	A	106	ARG	CA	56.278	0.20	1
1	A	106	ARG	CB	30.639	0.20	1
1	A	106	ARG	CG	27.147	0.20	1
1	A	106	ARG	CD	43.419	0.20	1
1	A	106	ARG	N	121.55	0.30	1
1	A	107	LEU	H	8.197	0.05	1
1	A	107	LEU	HA	4.321	0.05	1
1	A	107	LEU	HB2	1.659	0.05	2
1	A	107	LEU	HB3	1.591	0.05	2
1	A	107	LEU	HD11	0.873	0.05	2
1	A	107	LEU	HD12	0.873	0.05	2
1	A	107	LEU	HD13	0.873	0.05	2
1	A	107	LEU	HD21	0.918	0.05	2
1	A	107	LEU	HD22	0.918	0.05	2
1	A	107	LEU	HD23	0.918	0.05	2
1	A	107	LEU	C	177.078	0.20	1
1	A	107	LEU	CA	55.279	0.20	1
1	A	107	LEU	CB	42.319	0.20	1
1	A	107	LEU	CD1	23.903	0.20	1
1	A	107	LEU	CD2	24.534	0.20	1
1	A	107	LEU	N	122.855	0.30	1
1	A	108	GLN	H	8.232	0.05	1
1	A	108	GLN	HA	4.25	0.05	1
1	A	108	GLN	HB2	1.977	0.05	2
1	A	108	GLN	HB3	1.911	0.05	2
1	A	108	GLN	HG2	2.268	0.05	1
1	A	108	GLN	HG3	2.268	0.05	1
1	A	108	GLN	HE21	7.137	0.05	1
1	A	108	GLN	HE22	6.738	0.05	1
1	A	108	GLN	C	175.187	0.20	1
1	A	108	GLN	CA	55.802	0.20	1
1	A	108	GLN	CB	29.548	0.20	1
1	A	108	GLN	CG	33.686	0.20	1
1	A	108	GLN	N	120.212	0.30	1
1	A	108	GLN	NE2	113.527	0.30	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	109	ASN	H	8.293	0.05	1
1	A	109	ASN	HA	4.643	0.05	1
1	A	109	ASN	HB2	2.764	0.05	2
1	A	109	ASN	HB3	2.683	0.05	2
1	A	109	ASN	HD21	7.566	0.05	1
1	A	109	ASN	HD22	6.862	0.05	1
1	A	109	ASN	C	174.744	0.20	1
1	A	109	ASN	CA	53.118	0.20	1
1	A	109	ASN	CB	38.856	0.20	1
1	A	109	ASN	N	119.229	0.30	1
1	A	109	ASN	ND2	112.768	0.30	1
1	A	110	PHE	H	8.126	0.05	1
1	A	110	PHE	HA	4.613	0.05	1
1	A	110	PHE	HB2	3.191	0.05	2
1	A	110	PHE	HB3	3.002	0.05	2
1	A	110	PHE	HD1	7.247	0.05	3
1	A	110	PHE	HD2	7.247	0.05	3
1	A	110	PHE	C	175.22	0.20	1
1	A	110	PHE	CA	57.87	0.20	1
1	A	110	PHE	CB	39.607	0.20	1
1	A	110	PHE	CD1	131.885	0.20	3
1	A	110	PHE	CD2	131.885	0.20	3
1	A	110	PHE	N	120.374	0.30	1
1	A	111	ASP	H	8.336	0.05	1
1	A	111	ASP	HA	4.636	0.05	1
1	A	111	ASP	HB2	2.708	0.05	2
1	A	111	ASP	HB3	2.59	0.05	2
1	A	111	ASP	C	176.153	0.20	1
1	A	111	ASP	CA	54.401	0.20	1
1	A	111	ASP	CB	41.166	0.20	1
1	A	111	ASP	N	121.757	0.30	1
1	A	112	VAL	H	8.071	0.05	1
1	A	112	VAL	HA	4.199	0.05	1
1	A	112	VAL	HB	2.175	0.05	1
1	A	112	VAL	HG11	0.957	0.05	1
1	A	112	VAL	HG12	0.957	0.05	1
1	A	112	VAL	HG13	0.957	0.05	1
1	A	112	VAL	HG21	0.957	0.05	1
1	A	112	VAL	HG22	0.957	0.05	1
1	A	112	VAL	HG23	0.957	0.05	1
1	A	112	VAL	C	176.448	0.20	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	112	VAL	CA	62.368	0.20	1
1	A	112	VAL	CB	32.655	0.20	1
1	A	112	VAL	CG1	20.263	0.20	1
1	A	112	VAL	CG2	20.263	0.20	1
1	A	112	VAL	N	119.947	0.30	1
1	A	113	THR	H	8.285	0.05	1
1	A	113	THR	HA	4.394	0.05	1
1	A	113	THR	HB	4.261	0.05	1
1	A	113	THR	HG21	1.221	0.05	1
1	A	113	THR	HG22	1.221	0.05	1
1	A	113	THR	HG23	1.221	0.05	1
1	A	113	THR	C	174.623	0.20	1
1	A	113	THR	CA	61.961	0.20	1
1	A	113	THR	CB	69.745	0.20	1
1	A	113	THR	CG2	21.61	0.20	1
1	A	113	THR	N	116.901	0.30	1
1	A	114	SER	H	8.213	0.05	1
1	A	114	SER	HA	4.48	0.05	1
1	A	114	SER	HB2	3.819	0.05	1
1	A	114	SER	HB3	3.819	0.05	1
1	A	114	SER	C	174.02	0.20	1
1	A	114	SER	CA	58.114	0.20	1
1	A	114	SER	CB	64.094	0.20	1
1	A	114	SER	N	117.913	0.30	1
1	A	115	ASP	H	8.384	0.05	1
1	A	115	ASP	HA	4.637	0.05	1
1	A	115	ASP	HB2	2.712	0.05	2
1	A	115	ASP	HB3	2.59	0.05	2
1	A	115	ASP	C	175.965	0.20	1
1	A	115	ASP	CA	54.319	0.20	1
1	A	115	ASP	CB	41.116	0.20	1
1	A	115	ASP	N	122.728	0.30	1
1	A	116	ILE	H	8.04	0.05	1
1	A	116	ILE	HA	4.213	0.05	1
1	A	116	ILE	HB	1.874	0.05	1
1	A	116	ILE	HG12	1.447	0.05	2
1	A	116	ILE	HG13	1.17	0.05	2
1	A	116	ILE	HG21	0.893	0.05	1
1	A	116	ILE	HG22	0.893	0.05	1
1	A	116	ILE	HG23	0.893	0.05	1
1	A	116	ILE	HD11	0.839	0.05	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	116	ILE	HD12	0.839	0.05	1
1	A	116	ILE	HD13	0.839	0.05	1
1	A	116	ILE	C	176.113	0.20	1
1	A	116	ILE	CA	61.008	0.20	1
1	A	116	ILE	CB	38.904	0.20	1
1	A	116	ILE	CG1	27.162	0.20	1
1	A	116	ILE	CG2	17.545	0.20	1
1	A	116	ILE	CD1	13.005	0.20	1
1	A	116	ILE	N	120.615	0.30	1
1	A	117	SER	H	8.373	0.05	1
1	A	117	SER	HA	4.471	0.05	1
1	A	117	SER	HB2	3.869	0.05	2
1	A	117	SER	HB3	3.811	0.05	2
1	A	117	SER	C	174.447	0.20	1
1	A	117	SER	CA	58.096	0.20	1
1	A	117	SER	CB	63.956	0.20	1
1	A	117	SER	N	120.205	0.30	1
1	A	118	ASP	H	8.39	0.05	1
1	A	118	ASP	HA	4.575	0.05	1
1	A	118	ASP	HB2	2.666	0.05	2
1	A	118	ASP	HB3	2.575	0.05	2
1	A	118	ASP	C	175.522	0.20	1
1	A	118	ASP	CA	54.84	0.20	1
1	A	118	ASP	CB	40.904	0.20	1
1	A	118	ASP	N	123.18	0.30	1

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	129	-0.16 ± 0.13	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	120	-0.13 ± 0.19	None needed (< 0.5 ppm)
$^{13}C'$	128	0.02 ± 0.14	None needed (< 0.5 ppm)
^{15}N	127	-1.18 ± 0.42	Should be applied

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 89%, i.e. 1102 atoms were assigned a chemical shift out of a possible 1243. 0 out of 12 assigned methyl groups (LEU and VAL) were assigned

stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	463/463 (100%)	188/188 (100%)	184/184 (100%)	91/91 (100%)
Sidechain	553/673 (82%)	376/432 (87%)	172/215 (80%)	5/26 (19%)
Aromatic	86/107 (80%)	43/52 (83%)	41/52 (79%)	2/3 (67%)
Overall	1102/1243 (89%)	607/672 (90%)	397/451 (88%)	98/120 (82%)

7.1.4 Statistically unusual chemical shifts (i)

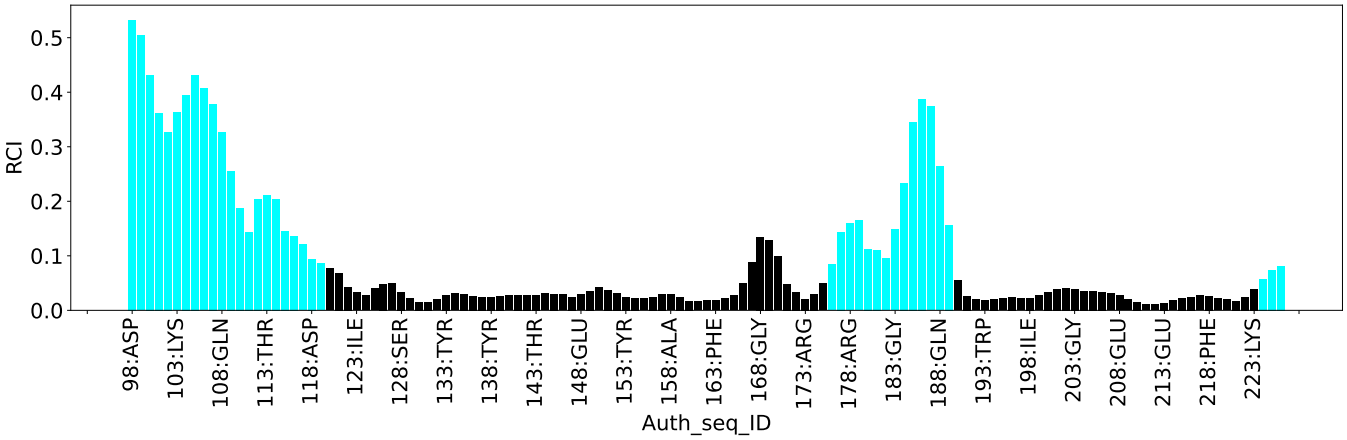
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	129	ARG	NE	117.99	76.53 – 92.65	20.7
1	A	213	GLU	HG2	0.76	1.24 – 3.30	-7.3
1	A	191	THR	HB	2.04	2.57 – 5.77	-6.6
1	A	214	ILE	HG21	-0.85	-0.56 – 2.11	-6.1
1	A	214	ILE	HG22	-0.85	-0.56 – 2.11	-6.1
1	A	214	ILE	HG23	-0.85	-0.56 – 2.11	-6.1
1	A	173	ARG	HA	2.06	2.06 – 6.51	-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:


RCI
0.0
0.1
0.2
0.3
0.4
0.5
98:ASP
103:LYS
108:GLN
113:THR
118:ASP
123:ILE
128:SER
133:TYR
138:TYR
143:THR
148:GLU
153:TYR
158:ALA
163:PHE
168:GLY
173:ARG
178:ARG
183:GLY
188:GLN
193:TRP
198:ILE
203:GLY
208:GLU
213:GLU
218:PHE
223:LYS
Auth_seq_ID

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	2238
Intra-residue (|i-j|=0)	416
Sequential (|i-j|=1)	631
Medium range (|i-j|>1 and |i-j|<5)	309
Long range (|i-j|≥5)	882
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	0
Number of restraints per residue	17.3
Number of long range restraints per residue[1]	6.8

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

8.2 Residual restraint violations

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

8.2.1 Average number of distance violations per model

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

Bins (Å)	Average number of violations per model	Max (Å)
0.1-0.2 (Small)	3.5	0.2
0.2-0.5 (Medium)	2.3	0.5
>0.5 (Large)	4.5	1.61

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

Dihedral-angle violations less than 1° are not included in the calculation.There are no dihedral-angle violations

9 Distance violation analysis (i)

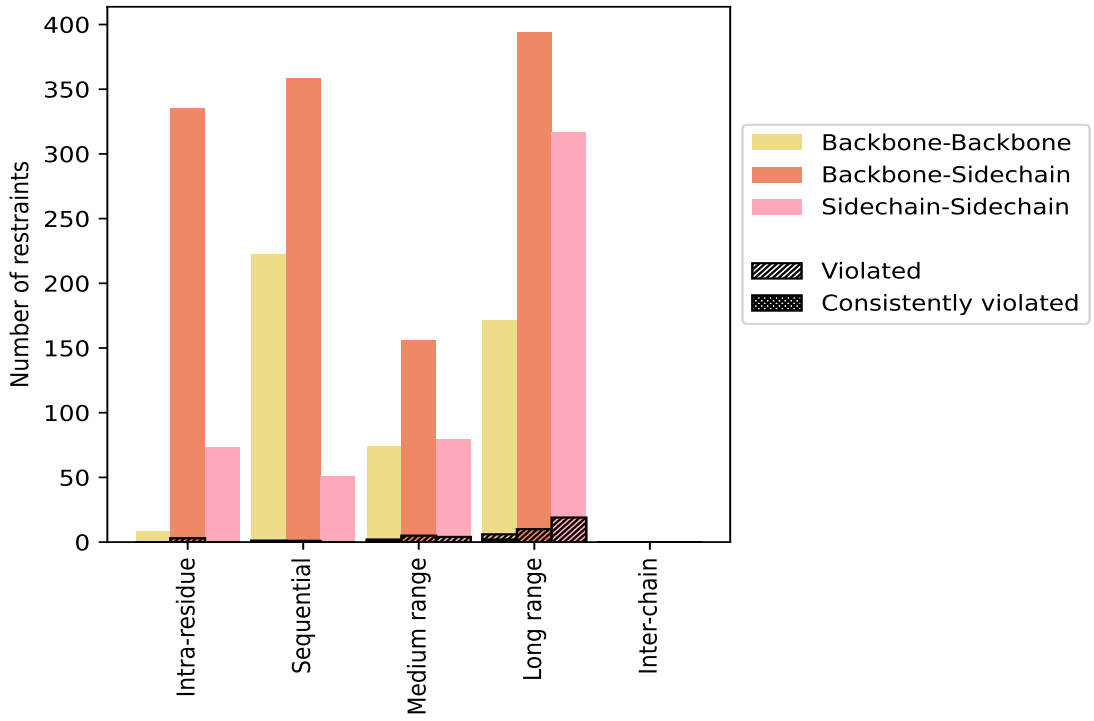
9.1 Summary of distance violations (i)

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.

Restraints type	Count	%[1]	Violated[3]			Consistently Violated[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
Intra-residue (|i-j|=0)	416	18.6	3	0.7	0.1	0	0.0	0.0
Backbone-Backbone	8	0.4	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	335	15.0	3	0.9	0.1	0	0.0	0.0
Sidechain-Sidechain	73	3.3	0	0.0	0.0	0	0.0	0.0
Sequential (|i-j|=1)	631	28.2	2	0.3	0.1	1	0.2	0.0
Backbone-Backbone	222	9.9	1	0.5	0.0	1	0.5	0.0
Backbone-Sidechain	358	16.0	1	0.3	0.0	0	0.0	0.0
Sidechain-Sidechain	51	2.3	0	0.0	0.0	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	309	13.8	11	3.6	0.5	1	0.3	0.0
Backbone-Backbone	74	3.3	2	2.7	0.1	1	1.4	0.0
Backbone-Sidechain	156	7.0	5	3.2	0.2	0	0.0	0.0
Sidechain-Sidechain	79	3.5	4	5.1	0.2	0	0.0	0.0
Long range (|i-j|≥5)	882	39.4	35	4.0	1.6	2	0.2	0.1
Backbone-Backbone	171	7.6	6	3.5	0.3	2	1.2	0.1
Backbone-Sidechain	394	17.6	10	2.5	0.4	0	0.0	0.0
Sidechain-Sidechain	317	14.2	19	6.0	0.8	0	0.0	0.0
Inter-chain	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Hydrogen bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	2238	100.0	51	2.3	2.3	4	0.2	0.2
Backbone-Backbone	475	21.2	9	1.9	0.4	4	0.8	0.2
Backbone-Sidechain	1243	55.5	19	1.5	0.8	0	0.0	0.0
Sidechain-Sidechain	520	23.2	23	4.4	1.0	0	0.0	0.0

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

9.1.1 Bar chart : Distribution of distance restraints and violations



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

9.2 Distance violation statistics for each model

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

Model ID	Number of violations						Mean (Å)	Max (Å)	SD[6] (Å)	Median (Å)
	IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total				
1	1	1	2	5	0	9	0.48	0.91	0.31	0.39
2	0	1	2	8	0	11	0.45	1.28	0.38	0.32
3	1	1	3	7	0	12	0.5	1.04	0.35	0.38
4	0	1	2	6	0	9	0.51	0.92	0.31	0.49
5	0	2	2	6	0	10	0.57	1.12	0.39	0.53
6	0	1	3	7	0	11	0.54	1.08	0.35	0.49
7	0	1	2	7	0	10	0.52	1.19	0.34	0.46
8	1	1	2	5	0	9	0.67	1.11	0.34	0.54
9	0	1	2	5	0	8	0.7	1.61	0.47	0.66
10	0	1	3	5	0	9	0.63	1.49	0.43	0.47
11	0	1	3	4	0	8	0.63	1.17	0.4	0.61

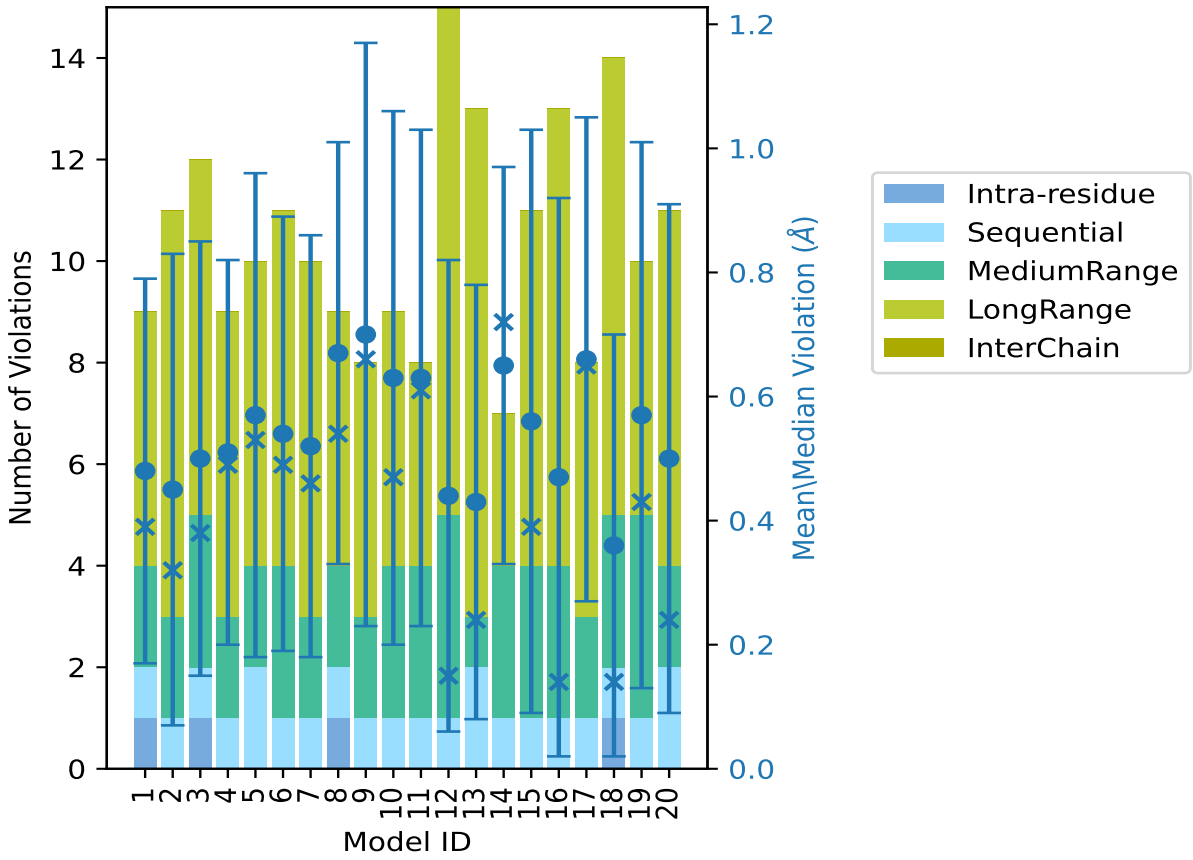
Continued on next page...

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	0	1	4	10	0	15	0.44	1.11	0.38	0.15
13	0	2	1	10	0	13	0.43	1.11	0.35	0.24
14	0	1	3	3	0	7	0.65	1.04	0.32	0.72
15	0	1	3	7	0	11	0.56	1.55	0.47	0.39
16	0	1	3	9	0	13	0.47	1.47	0.45	0.14
17	0	1	2	5	0	8	0.66	1.34	0.39	0.65
18	1	1	3	9	0	14	0.36	0.92	0.34	0.14
19	0	1	4	5	0	10	0.57	1.56	0.44	0.43
20	0	2	2	7	0	11	0.5	1.36	0.41	0.24

[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

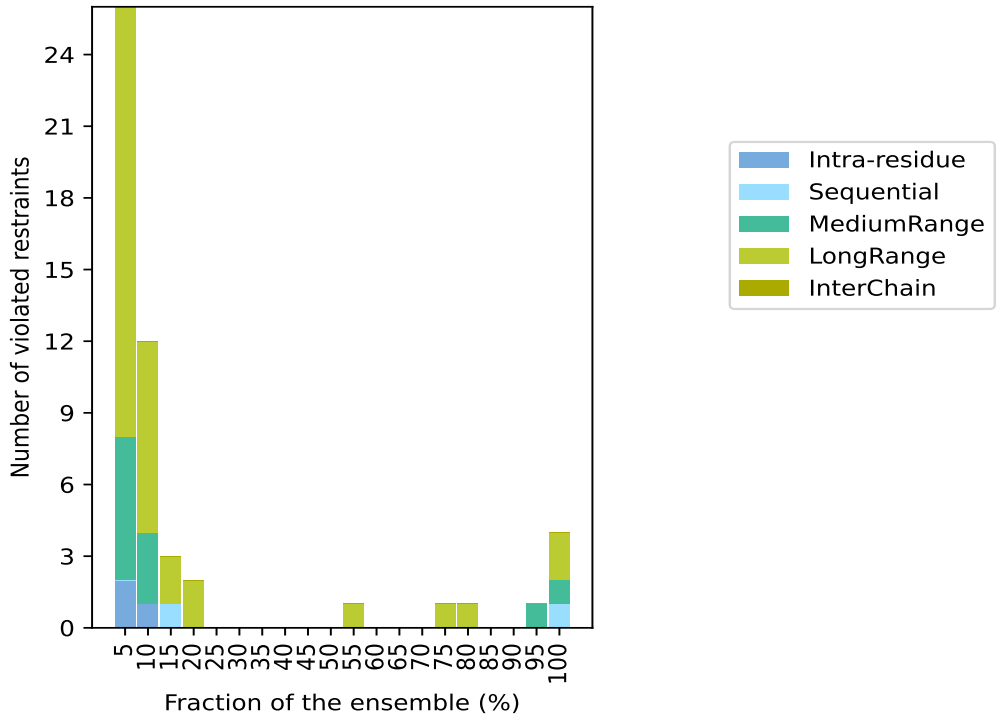
9.3 Distance violation statistics for the ensemble (i)

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, 2187(IR:413, SQ:629, MR:298, LR:847, 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]	%
2	0	6	18	0	26	1	5.0
1	0	3	8	0	12	2	10.0
0	1	0	2	0	3	3	15.0
0	0	0	2	0	2	4	20.0
0	0	0	0	0	0	5	25.0
0	0	0	0	0	0	6	30.0
0	0	0	0	0	0	7	35.0
0	0	0	0	0	0	8	40.0
0	0	0	0	0	0	9	45.0
0	0	0	0	0	0	10	50.0
0	0	0	1	0	1	11	55.0
0	0	0	0	0	0	12	60.0
0	0	0	0	0	0	13	65.0
0	0	0	0	0	0	14	70.0
0	0	0	1	0	1	15	75.0
0	0	0	1	0	1	16	80.0
0	0	0	0	0	0	17	85.0
0	0	0	0	0	0	18	90.0
0	0	1	0	0	1	19	95.0
0	1	1	2	0	4	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

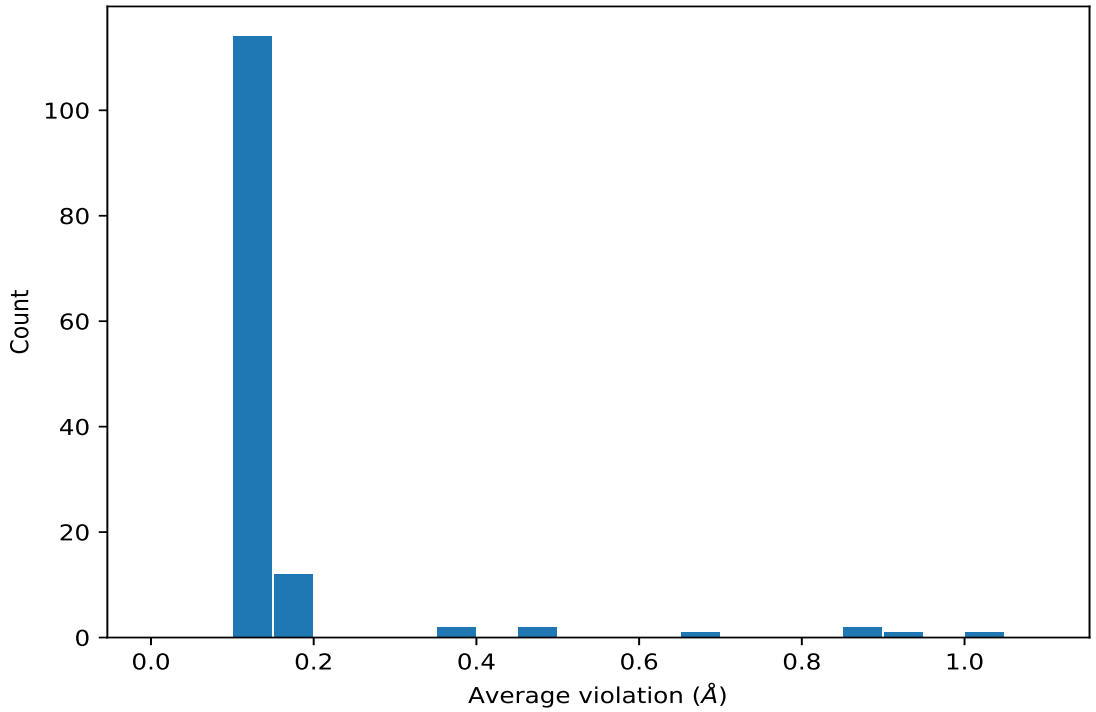
9.3.1 Bar graph : Distance violation statistics for the ensemble



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 (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.

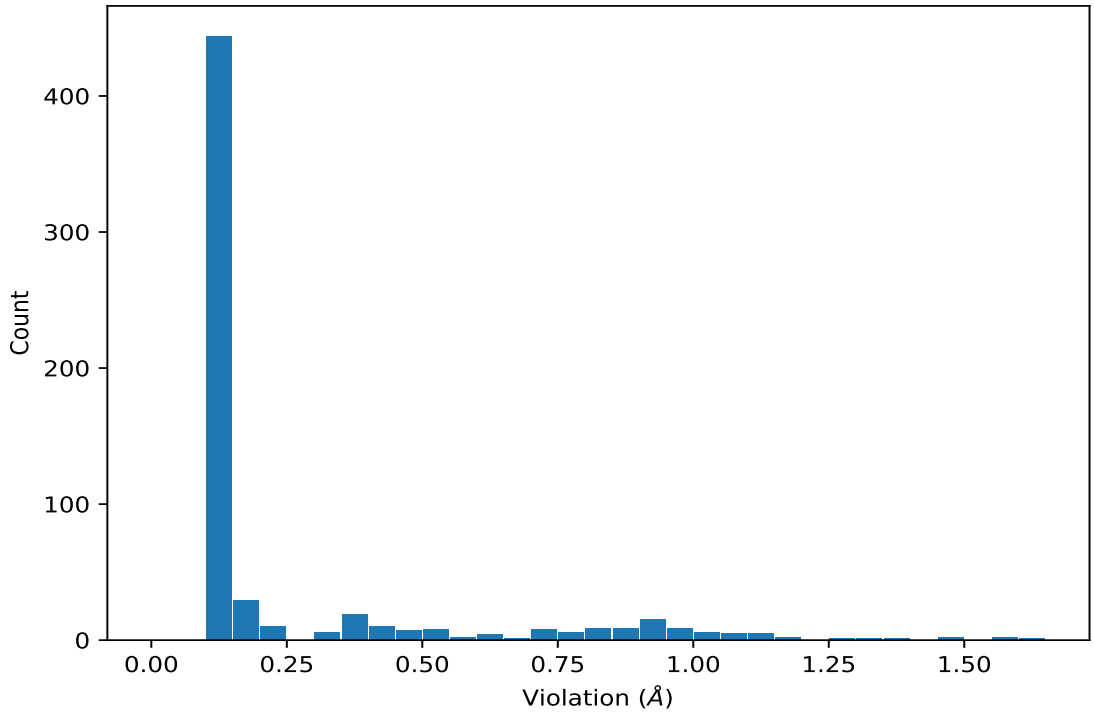
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,903)	1:A:151:GLY:HA2	1:A:200:SER:H	20	1.01	0.13	1.03
(1,709)	1:A:166:VAL:H	1:A:168:GLY:HA3	20	0.88	0.1	0.91
(1,24)	1:A:196:GLY:HA2	1:A:213:GLU:H	20	0.86	0.07	0.84
(1,372)	1:A:151:GLY:HA2	1:A:152:GLU:H	20	0.37	0.04	0.38
(1,1735)	1:A:165:SER:HA	1:A:168:GLY:HA3	19	0.93	0.45	0.75
(1,1519)	1:A:157:LEU:HG	1:A:196:GLY:HA3	16	0.65	0.38	0.48
(1,1736)	1:A:163:PHE:HB2	1:A:168:GLY:HA3	15	0.39	0.17	0.37
(1,1365)	1:A:151:GLY:HA3	1:A:202:ARG:HD2	11	0.49	0.19	0.49
(1,1365)	1:A:151:GLY:HA3	1:A:202:ARG:HD3	11	0.49	0.19	0.49
(1,1614)	1:A:121:PHE:HD1	1:A:198:ILE:HG21	4	0.12	0.01	0.12
(1,1614)	1:A:121:PHE:HD1	1:A:198:ILE:HG22	4	0.12	0.01	0.12
(1,1614)	1:A:121:PHE:HD1	1:A:198:ILE:HG23	4	0.12	0.01	0.12
(1,1614)	1:A:121:PHE:HD2	1:A:198:ILE:HG21	4	0.12	0.01	0.12
(1,1614)	1:A:121:PHE:HD2	1:A:198:ILE:HG22	4	0.12	0.01	0.12
(1,1614)	1:A:121:PHE:HD2	1:A:198:ILE:HG23	4	0.12	0.01	0.12
(1,2198)	1:A:211:ALA:H	1:A:220:LEU:HD11	4	0.11	0.0	0.11

[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)
(1,1735)	1:A:165:SER:HA	1:A:168:GLY:HA3	9	1.61
(1,1735)	1:A:165:SER:HA	1:A:168:GLY:HA3	19	1.56
(1,1735)	1:A:165:SER:HA	1:A:168:GLY:HA3	15	1.55
(1,1735)	1:A:165:SER:HA	1:A:168:GLY:HA3	10	1.49
(1,1735)	1:A:165:SER:HA	1:A:168:GLY:HA3	16	1.47
(1,1735)	1:A:165:SER:HA	1:A:168:GLY:HA3	20	1.36
(1,1735)	1:A:165:SER:HA	1:A:168:GLY:HA3	17	1.34
(1,903)	1:A:151:GLY:HA2	1:A:200:SER:H	2	1.28

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,903)	1:A:151:GLY:HA2	1:A:200:SER:H	7	1.19
(1,1519)	1:A:157:LEU:HG	1:A:196:GLY:HA3	11	1.17

10 Dihedral-angle violation analysis (i)

Dihedral angle analysis failed due to data error in the dihedral angle restraints, possibly missing target value