wwPDB NMR Structure Validation Summary Report

Feb 18, 2025 – 09:02 am GMT

PDB ID	:	8S2A
BMRB ID	:	34901
Title	:	NMR structure of xantholysin A in micellar DPC solution
Authors	:	Kovacs, B.; Geudens, N.; Martins, J.C.
Deposited on	:	2024-02-16

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:

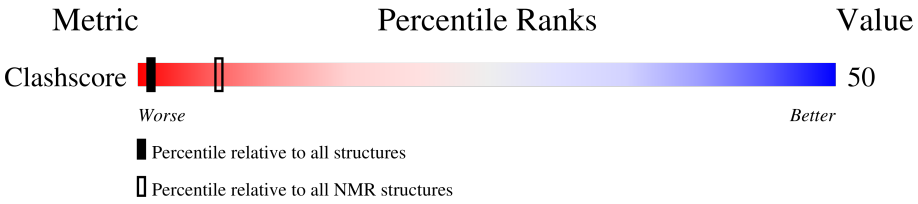
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41

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 84%.

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	210492	14027

Molprobity failed to run

2 Ensemble composition and analysis (i)

This entry contains 11 models. This entry does not contain polypeptide chains, therefore identification of well-defined residues and clustering analysis are not possible. All residues are included in the validation scores.

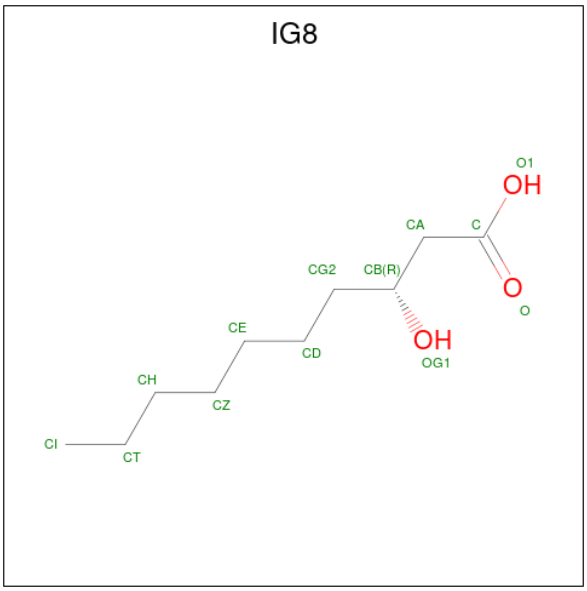
3 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 270 atoms, of which 145 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein (with D amino acids) called Xantholysin A.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
1	A	14	239	74	126	18	21	0

- Molecule 2 is (3 {R})-3-oxidanyldecanoic acid (three-letter code: IG8) (formula: $C_{10}H_{20}O_3$).



Mol	Chain	Residues	Atoms			
			Total	C	H	O
2	A	1	31	10	19	2

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.

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

The representative model is number 1. Colouring as in section 4.1 above.

5 Refinement protocol and experimental data overview

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

Of the 100 calculated structures, 11 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
Amber	refinement	
CNS	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	1
Total number of shifts	178
Number of shifts mapped to atoms	172
Number of unparsed shifts	0
Number of shifts with mapping errors	6
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	84%

6 Model quality

6.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: DLE, DGN, DGL, DVA, DSN, IG8

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	0.70±0.21	0±0/39 (0.0± 0.0%)	1.01±0.28	0±0/50 (0.2± 0.6%)
All	All	0.73	0/429 (0.0%)	1.05	1/550 (0.2%)

There are no bond-length outliers.

All unique angle outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	12	LEU	CB-CA-C	5.20	120.08	110.20	1	1

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	113	126	119	12±6
2	A	12	19	0	0±0
All	All	1375	1595	1307	134

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 50.

5 of 31 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:10:DLE:HD12	1:A:11:DGN:N	0.76	1.94	5	3
1:A:6:DLE:O	1:A:10:DLE:N	0.75	2.20	2	10
1:A:10:DLE:HD12	1:A:10:DLE:C	0.73	2.13	3	3
1:A:5:DVA:O	1:A:9:DVA:HG13	0.66	1.90	5	10
1:A:10:DLE:C	1:A:10:DLE:CD1	0.62	2.77	3	3

6.3 Torsion angles

6.3.1 Protein backbone

There are no protein molecules in this entry.

6.3.2 Protein sidechains

There are no protein molecules in this entry.

6.3.3 RNA

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains

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

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6.5 Carbohydrates

There are no oligosaccharides in this entry.

6.6 Ligand geometry (i)

1 ligand is 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	IG8	A	101	1	11,11,12	0.73±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	IG8	A	101	1	11,11,13	1.09±0.03	0±0 (0±2%)

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	IG8	A	101	1	-	0±0,10,10,11	-

There are no bond-length outliers.

All unique angle outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
2	A	101	IG8	CD-CG2-CB	2.44	107.83	114.85	1	1

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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 84% for the well-defined parts and 84% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: */home/hdgsgsg/Desktop/CCPNMR_old_project/XLA_water_DPC_175_mM_29*

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	178
Number of shifts mapped to atoms	172
Number of unparsed shifts	0
Number of shifts with mapping errors	6
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. First 5 (of 6) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	1	IG8	HA1	2.75	0.00	2
1	A	1	IG8	HA2	2.42	0.00	2
1	A	1	IG8	HB	3.96	0.00	1
1	A	1	IG8	CA	43.6	0.00	1
1	A	1	IG8	CB	68.3	0.00	1
1	A	2	LEU	H	9.19	0.00	1

7.1.2 Chemical shift referencing

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

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 84%, i.e. 73 atoms were assigned a chemical shift out of a possible 87. 0 out of 3 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	15/25 (60%)	10/10 (100%)	5/10 (50%)	0/5 (0%)
Sidechain	58/62 (94%)	40/42 (95%)	18/19 (95%)	0/1 (0%)
Overall	73/87 (84%)	50/52 (96%)	23/29 (79%)	0/6 (0%)

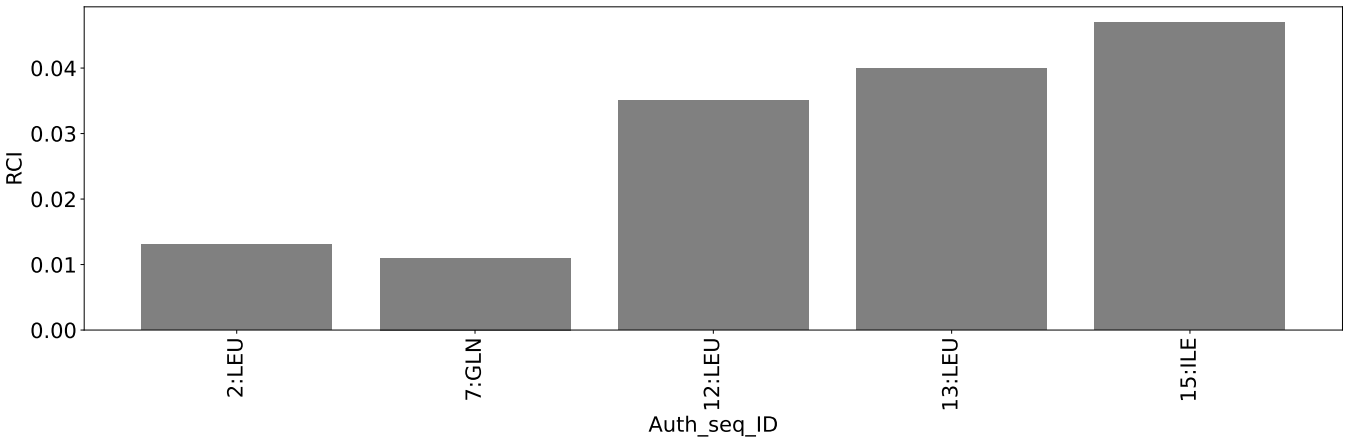
7.1.4 Statistically unusual chemical shifts

There are no statistically unusual chemical shifts.

7.1.5 Random Coil Index (RCI) plots

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

Random coil index (RCI) for chain A:



8 NMR restraints analysis

8.1 Conformationally restricting restraints

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

Description	Value
Total distance restraints	310
Intra-residue (|i-j|=0)	128
Sequential (|i-j|=1)	92
Medium range (|i-j|>1 and |i-j|<5)	87
Long range (|i-j|≥5)	3
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	22
Number of restraints per residue	20.7
Number of long range restraints per residue[1]	0.2

[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)	13.1	0.2
0.2-0.5 (Medium)	14.6	0.5
>0.5 (Large)	16.3	1.88

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)

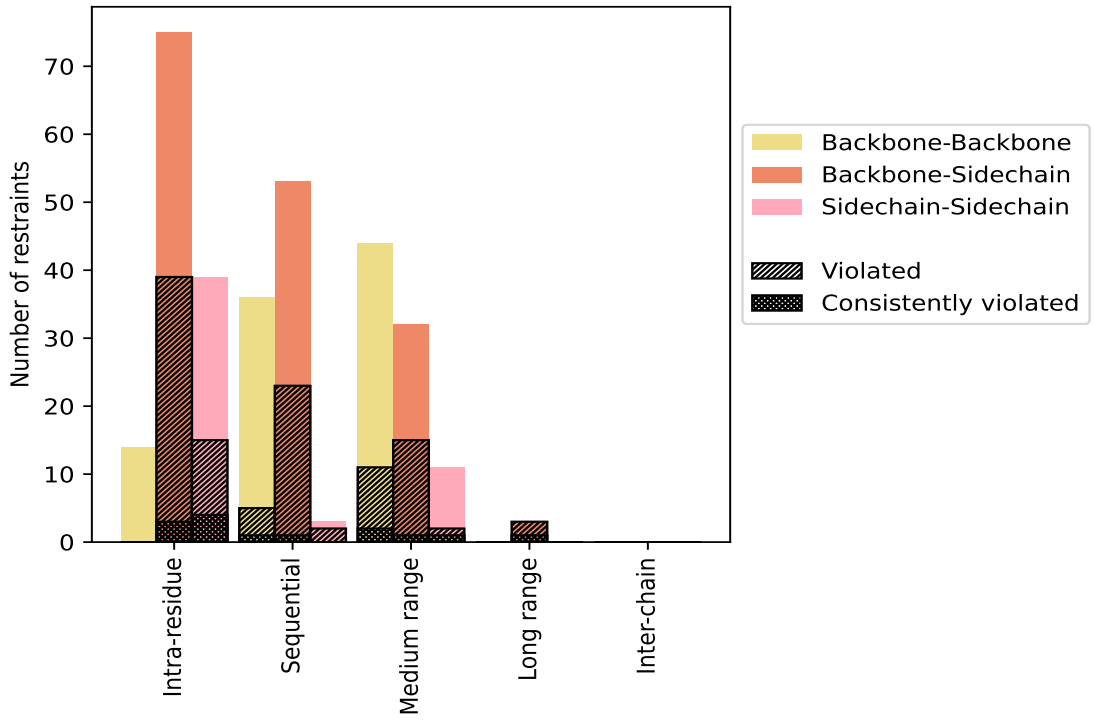
9.1 Summary of distance violations (i)

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)	128	41.3	54	42.2	17.4	7	5.5	2.3
Backbone-Backbone	14	4.5	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	75	24.2	39	52.0	12.6	3	4.0	1.0
Sidechain-Sidechain	39	12.6	15	38.5	4.8	4	10.3	1.3
Sequential (|i-j|=1)	92	29.7	30	32.6	9.7	2	2.2	0.6
Backbone-Backbone	36	11.6	5	13.9	1.6	1	2.8	0.3
Backbone-Sidechain	53	17.1	23	43.4	7.4	1	1.9	0.3
Sidechain-Sidechain	3	1.0	2	66.7	0.6	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	87	28.1	28	32.2	9.0	4	4.6	1.3
Backbone-Backbone	44	14.2	11	25.0	3.5	2	4.5	0.6
Backbone-Sidechain	32	10.3	15	46.9	4.8	1	3.1	0.3
Sidechain-Sidechain	11	3.5	2	18.2	0.6	1	9.1	0.3
Long range (|i-j|≥5)	3	1.0	3	100.0	1.0	1	33.3	0.3
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	3	1.0	3	100.0	1.0	1	33.3	0.3
Sidechain-Sidechain	0	0.0	0	0.0	0.0	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	310	100.0	115	37.1	37.1	14	4.5	4.5
Backbone-Backbone	94	30.3	16	17.0	5.2	3	3.2	1.0
Backbone-Sidechain	163	52.6	80	49.1	25.8	6	3.7	1.9
Sidechain-Sidechain	53	17.1	19	35.8	6.1	5	9.4	1.6

[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	31	17	19	3	0	70	0.51	1.58	0.35	0.4
2	18	9	11	1	0	39	0.5	1.76	0.49	0.26
3	21	10	10	1	0	42	0.49	1.88	0.42	0.31
4	23	8	8	1	0	40	0.49	1.66	0.44	0.26
5	18	7	7	1	0	33	0.4	1.43	0.32	0.26
6	18	9	11	1	0	39	0.45	1.42	0.37	0.29
7	25	17	13	2	0	57	0.5	1.8	0.45	0.31
8	25	11	10	1	0	47	0.45	1.6	0.39	0.27
9	17	9	11	1	0	38	0.5	1.77	0.48	0.26
10	19	9	12	1	0	41	0.51	1.63	0.44	0.28

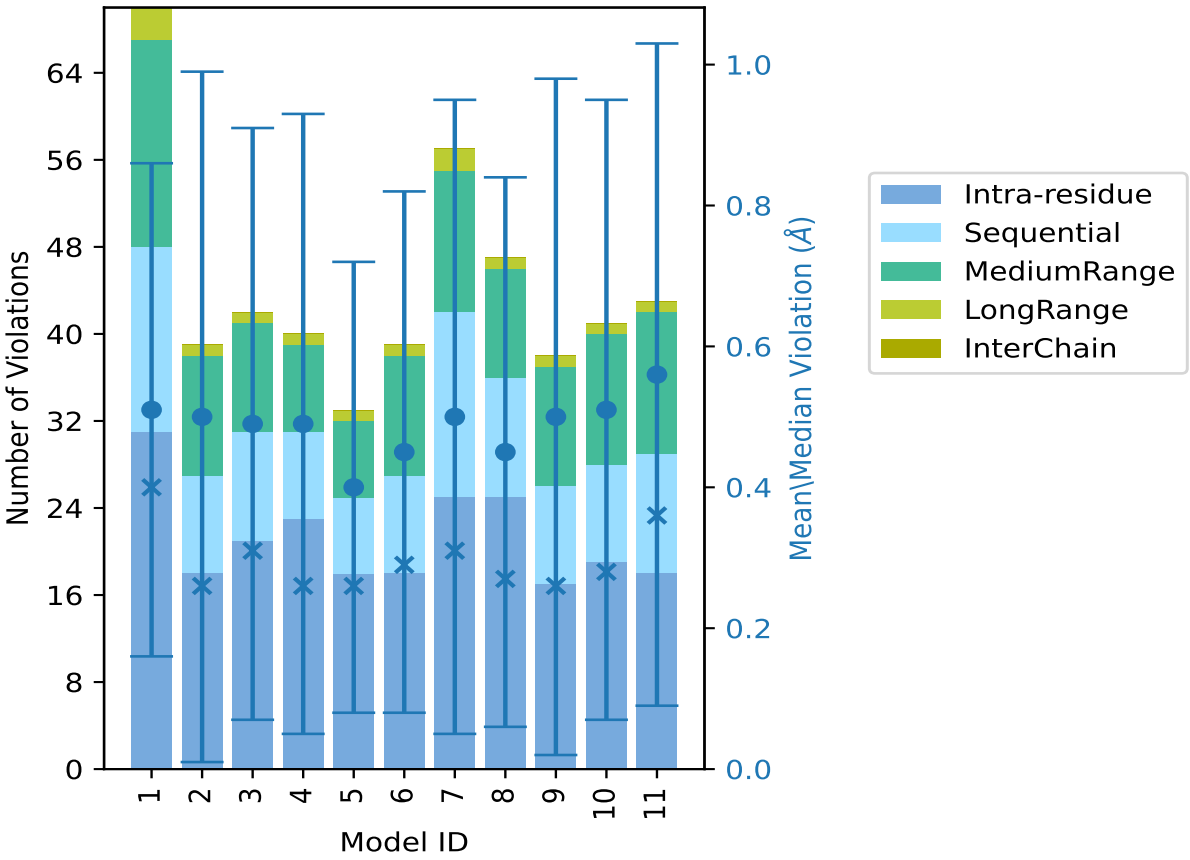
Continued on next page...

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	18	11	13	1	0	43	0.56	1.77	0.47	0.36

[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

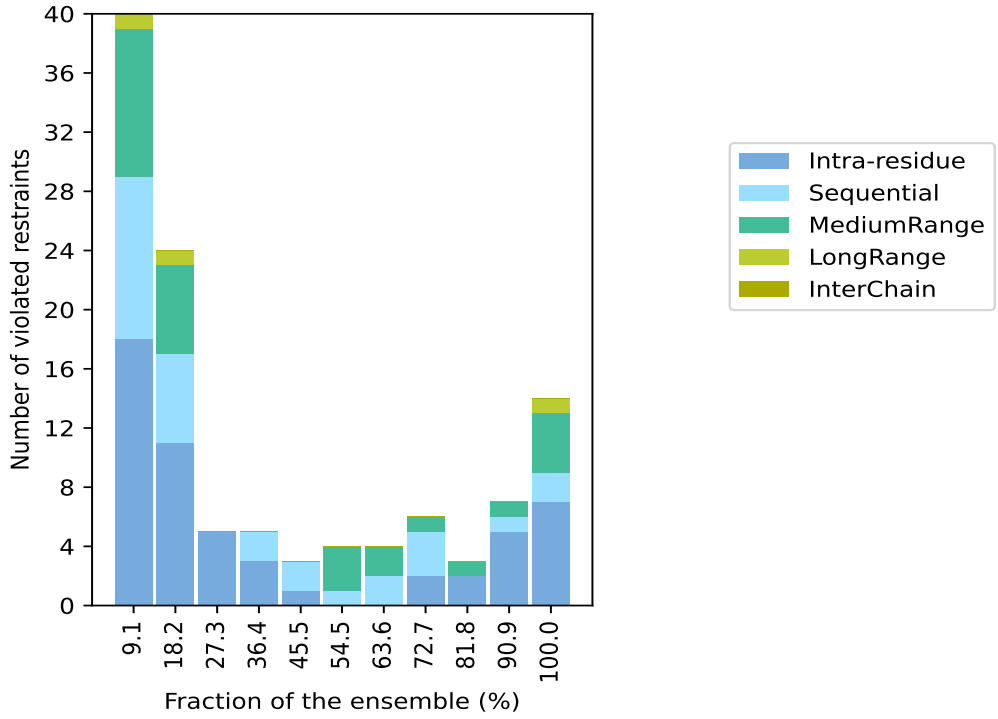
9.3 Distance violation statistics for the ensemble

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, 195(IR:74, SQ:62, MR:59, LR:0, 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]	%
18	11	10	1	0	40	1	9.1
11	6	6	1	0	24	2	18.2
5	0	0	0	0	5	3	27.3
3	2	0	0	0	5	4	36.4
1	2	0	0	0	3	5	45.5
0	1	3	0	0	4	6	54.5
0	2	2	0	0	4	7	63.6
2	3	1	0	0	6	8	72.7
2	0	1	0	0	3	9	81.8
5	1	1	0	0	7	10	90.9
7	2	4	1	0	14	11	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

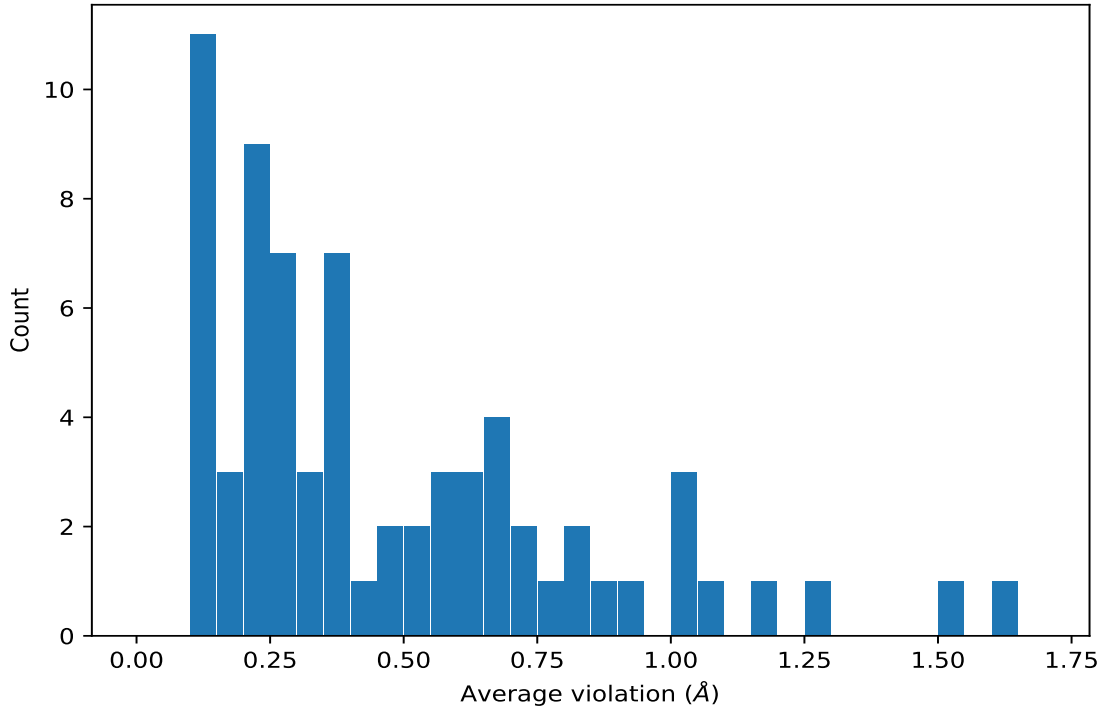
9.3.1 Bar graph : Distance violation statistics for the ensemble ⓘ



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)
(1,118)	1:8:A:DSN:H	1:13:A:LEU:HB3	11	1.5	0.13	1.55
(1,241)	1:2:A:LEU:HB3	1:5:A:DVA:HG11	11	1.03	0.51	1.35
(1,138)	1:13:A:LEU:HA	1:13:A:LEU:HB3	11	0.9	0.02	0.9
(1,253)	1:5:A:DVA:HA	1:6:A:DLE:HB2	11	0.76	0.26	0.65
(1,182)	1:9:A:DVA:HA	1:12:A:LEU:HB3	11	0.67	0.13	0.67
(1,209)	1:15:A:ILE:HB	1:15:A:ILE:HG12	11	0.53	0.07	0.55
(1,54)	1:13:A:LEU:H	1:13:A:LEU:HB2	11	0.37	0.01	0.37
(1,101)	1:4:A:DGN:H	1:8:A:DSN:H	11	0.31	0.19	0.25
(1,139)	1:12:A:LEU:HA	1:10:A:DLE:HA	11	0.27	0.15	0.25

Continued on next page...

Continued from previous page...

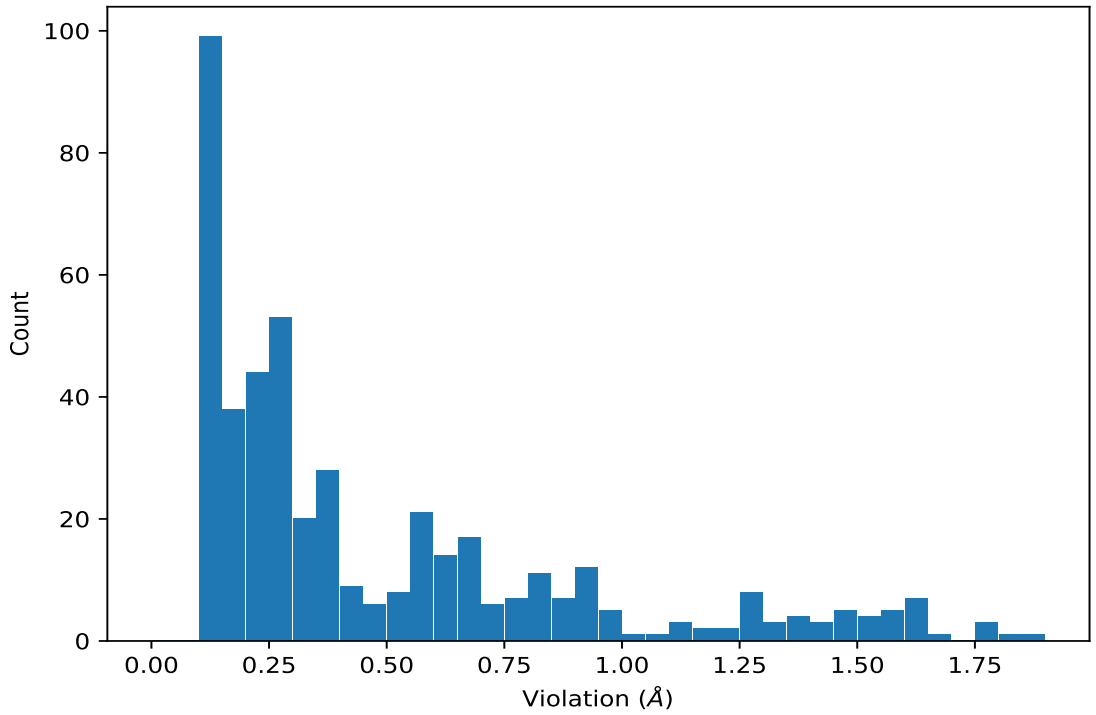
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,90)	1:7:A:GLN:HB3	1:7:A:GLN:H	11	0.23	0.03	0.25

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

9.5 All violated distance restraints

9.5.1 Histogram : Distribution of distance violations

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

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)
(1,299)	1:10:A:DLE:HA	1:10:A:DLE:HD11	3	1.88
(1,277)	1:9:A:DVA:H	1:10:A:DLE:HG	7	1.8
(1,277)	1:9:A:DVA:H	1:10:A:DLE:HG	9	1.77
(1,277)	1:9:A:DVA:H	1:10:A:DLE:HG	11	1.77

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,277)	1:9:A:DVA:H	1:10:A:DLE:HG	2	1.76
(1,277)	1:9:A:DVA:H	1:10:A:DLE:HG	4	1.66
(1,241)	1:2:A:LEU:HB3	1:5:A:DVA:HG11	10	1.63
(1,257)	1:6:A:DLE:HA	1:10:A:DLE:HG	2	1.62
(1,257)	1:6:A:DLE:HA	1:10:A:DLE:HG	9	1.62
(1,118)	1:8:A:DSN:H	1:13:A:LEU:HB3	7	1.62

10 Dihedral-angle violation analysis (i)

No dihedral-angle restraints found