



wwPDB NMR Structure Validation Summary Report ⓘ

May 28, 2026 – 09:21 AM EDT

PDB ID : 8XIC / pdb_00008xic
BMRB ID : 36627
Title : Structure of Trioxacarcin A covalently bound to guanosine-2'-fluorinated d(AACCGGTT)2
Authors : Gao, R.Q.; Cao, C.; Tang, G.L.
Deposited on : 2023-12-19

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 ⓘ 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 : **FAILED**
Mogul : 2022.3.0, CSD as543be (2022)
Buster-report : wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
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) : 2.49

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 23%.

There are no overall percentile quality scores available for this entry.

The sequence quality summary graphics cannot be shown.

2 Ensemble composition and analysis ⓘ

This entry contains 20 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.

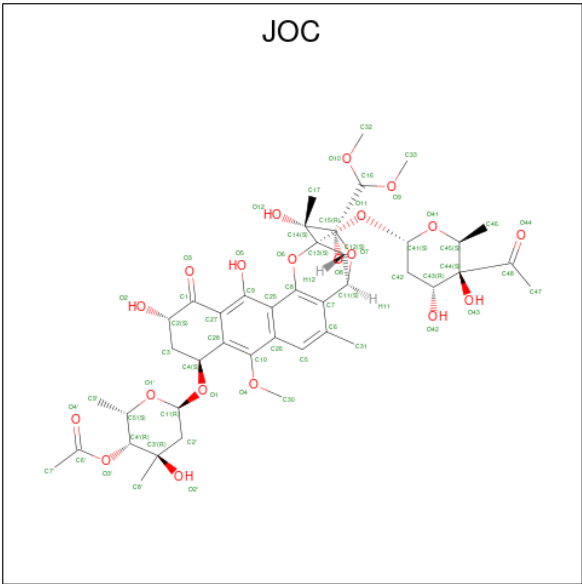
3 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 736 atoms, of which 286 are hydrogens and 0 are deuteriums.

- Molecule 1 is a DNA chain called DNA (5'-D(*AP*AP*CP*CP*2''F-GP*2''F-GP*TP*T)-3').

Mol	Chain	Residues	Atoms							Trace
1	A	8	Total	C	F	H	N	O	P	0
			253	78	2	90	30	46	7	
1	B	8	Total	C	F	H	N	O	P	0
			253	78	2	90	30	46	7	

- Molecule 2 is Trioxacarcin A, bound form (CCD ID: JOC) (formula: C₄₂H₅₄O₂₀) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms			
2	A	1	Total	C	H	O
			115	42	53	20
2	B	1	Total	C	H	O
			115	42	53	20

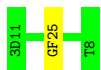
4 Residue-property plots [i](#)

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: DNA (5'-D(*AP*AP*CP*CP*2''F-GP*2''F-GP*TP*T)-3')

Chain A: 



- Molecule 1: DNA (5'-D(*AP*AP*CP*CP*2''F-GP*2''F-GP*TP*T)-3')

Chain B: 



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.

- Molecule 1: DNA (5'-D(*AP*AP*CP*CP*2''F-GP*2''F-GP*TP*T)-3')

Chain A: 



- Molecule 1: DNA (5'-D(*AP*AP*CP*CP*2''F-GP*2''F-GP*TP*T)-3')

Chain B: 



5 Refinement protocol and experimental data overview

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

Of the 100 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	structure calculation	
X-PLOR NIH	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	90
Number of shifts mapped to atoms	76
Number of unparsed shifts	0
Number of shifts with mapping errors	14
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	23%

6 Model quality [i](#)

6.1 Standard geometry [i](#)

MolProbity failed to run properly - this section will have to be empty.

6.2 Too-close contacts [i](#)

MolProbity failed to run properly - this section will have to be empty.

6.3 Torsion angles [i](#)

6.3.1 Protein backbone [i](#)

MolProbity failed to run properly - this section will have to be empty.

6.3.2 Protein sidechains [i](#)

MolProbity failed to run properly - this section will have to be empty.

6.3.3 RNA [i](#)

MolProbity failed to run properly - this section will have to be empty.

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

MolProbity failed to run properly - this section will have to be empty.

6.5 Carbohydrates [i](#)

MolProbity failed to run properly - this section will have to be empty.

6.6 Ligand geometry [i](#)

MolProbity failed to run properly - this section will have to be empty.

6.7 Other polymers [i](#)

MolProbity failed to run properly - this section will have to be empty.

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 23% for the well-defined parts and 23% for the entire structure.

7.1 Chemical shift list 1

File name: `working_cs.cif`

Chemical shift list name: *starch_output*

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	90
Number of shifts mapped to atoms	76
Number of unparsed shifts	0
Number of shifts with mapping errors	14
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 14) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	6	GF2	H3M3	1.066	.	.
1	A	6	GF2	H3O	4.155	.	.
1	A	6	GF2	H4M3	2.198	.	.
1	A	6	GF2	H63	1.17	.	.
1	A	6	GF2	HBL	4.315	.	.
1	A	6	GF2	HBN	2.798	.	.
1	A	6	GF2	HBP	5.153	.	.
1	A	6	GF2	HBR	6.78	.	.
1	A	6	GF2	HBS	12.506	.	.
1	A	6	GF2	HCD	4.971	.	.
1	A	6	GF2	HCK	5.096	.	.
1	A	6	GF2	HCQ	3.764	.	.
1	A	6	GF2	HCS	2.18	.	.
1	A	6	GF2	HCT	3.685	.	.

7.1.2 Chemical shift referencing [i](#)

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

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 23%, i.e. 44 atoms were assigned a chemical shift out of a possible 192. 0 out of 0 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹ H	¹³ C	¹⁵ N
Sugar	32/120 (27%)	32/70 (46%)	0/50 (0%)	0/0 (—%)
Base	12/72 (17%)	12/42 (29%)	0/20 (0%)	0/10 (0%)
Overall	44/192 (23%)	44/112 (39%)	0/70 (0%)	0/10 (0%)

7.1.4 Statistically unusual chemical shifts [i](#)

There are no statistically unusual chemical shifts.

7.1.5 Random Coil Index (RCI) plots [i](#)

No *random coil index*(RCI) plot could be generated from the current chemical shift list. RCI is only applicable to proteins

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	438
Intra-residue ($ i-j =0$)	228
Sequential ($ i-j =1$)	176
Medium range ($ i-j >1$ and $ i-j <5$)	6
Long range ($ i-j \geq 5$)	0
Inter-chain	28
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	56
Number of restraints per residue	24.3
Number of long range restraints per residue ¹	0.0

¹Long 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)	9.7	0.2
0.2-0.5 (Medium)	17.3	0.38
>0.5 (Large)	65.0	7.29

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 ⓘ

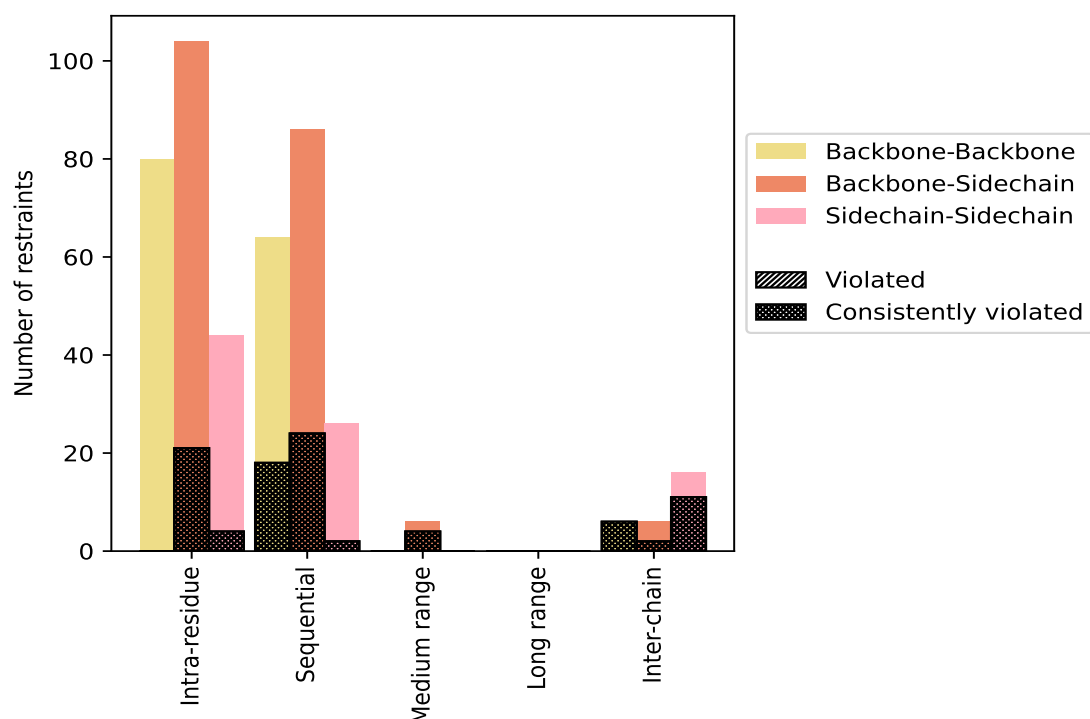
9.1 Summary of distance violations ⓘ

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	% ¹	Violated ³			Consistently Violated ⁴		
			Count	% ²	% ¹	Count	% ²	% ¹
Intra-residue (i-j =0)	228	52.1	25	11.0	5.7	25	11.0	5.7
Backbone-Backbone	80	18.3	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	104	23.7	21	20.2	4.8	21	20.2	4.8
Sidechain-Sidechain	44	10.0	4	9.1	0.9	4	9.1	0.9
Sequential (i-j =1)	176	40.2	44	25.0	10.0	44	25.0	10.0
Backbone-Backbone	64	14.6	18	28.1	4.1	18	28.1	4.1
Backbone-Sidechain	86	19.6	24	27.9	5.5	24	27.9	5.5
Sidechain-Sidechain	26	5.9	2	7.7	0.5	2	7.7	0.5
Medium range (i-j >1 & i-j <5)	6	1.4	4	66.7	0.9	4	66.7	0.9
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	6	1.4	4	66.7	0.9	4	66.7	0.9
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Long range (i-j ≥5)	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
Inter-chain	28	6.4	19	67.9	4.3	19	67.9	4.3
Backbone-Backbone	6	1.4	6	100.0	1.4	6	100.0	1.4
Backbone-Sidechain	6	1.4	2	33.3	0.5	2	33.3	0.5
Sidechain-Sidechain	16	3.7	11	68.8	2.5	11	68.8	2.5
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	438	100.0	92	21.0	21.0	92	21.0	21.0
Backbone-Backbone	150	34.2	24	16.0	5.5	24	16.0	5.5
Backbone-Sidechain	202	46.1	51	25.2	11.6	51	25.2	11.6
Sidechain-Sidechain	86	19.6	17	19.8	3.9	17	19.8	3.9

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

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



Violated and consistently violated restraints are shown using different hatch patterns in their respective categories. The hydrogen bonds and 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						Mean (Å)	Max (Å)	SD ⁶ (Å)	Median (Å)
	IR ¹	SQ ²	MR ³	LR ⁴	IC ⁵	Total				
1	25	44	4	0	19	92	2.24	7.28	1.94	1.94
2	25	44	4	0	19	92	2.24	7.28	1.94	1.94
3	25	44	4	0	19	92	2.24	7.28	1.94	1.94
4	25	44	4	0	19	92	2.24	7.27	1.94	1.94
5	25	44	4	0	19	92	2.24	7.28	1.94	1.94
6	25	44	4	0	19	92	2.24	7.28	1.94	1.94
7	25	44	4	0	19	92	2.24	7.28	1.94	1.94
8	25	44	4	0	19	92	2.24	7.28	1.94	1.95
9	25	44	4	0	19	92	2.24	7.28	1.94	1.94
10	25	44	4	0	19	92	2.24	7.28	1.94	1.94

Continued on next page...

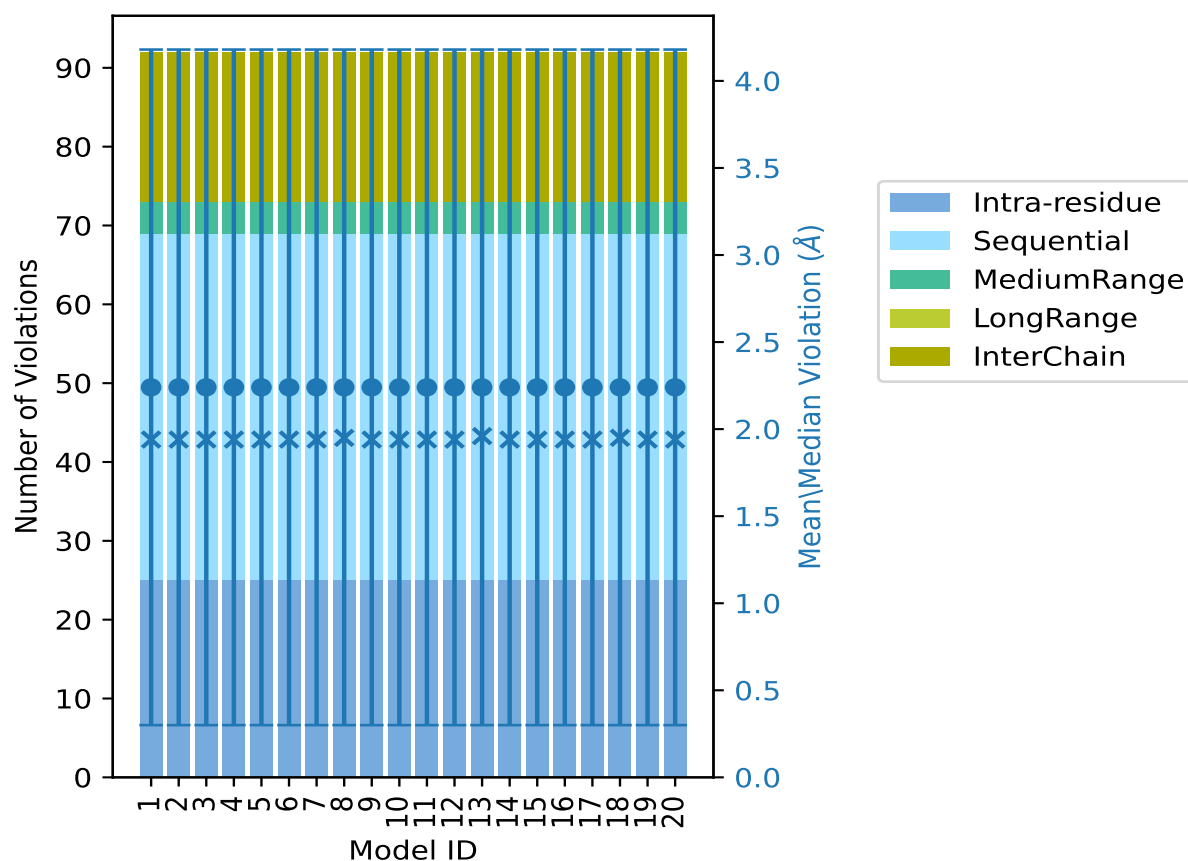
Continued from previous page...

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	25	44	4	0	19	92	2.24	7.28	1.94	1.94
12	25	44	4	0	19	92	2.24	7.28	1.94	1.94
13	25	44	4	0	19	92	2.24	7.29	1.94	1.96
14	25	44	4	0	19	92	2.24	7.29	1.94	1.94
15	25	44	4	0	19	92	2.24	7.28	1.94	1.94
16	25	44	4	0	19	92	2.24	7.28	1.94	1.94
17	25	44	4	0	19	92	2.24	7.28	1.94	1.94
18	25	44	4	0	19	92	2.24	7.28	1.94	1.95
19	25	44	4	0	19	92	2.24	7.29	1.94	1.94
20	25	44	4	0	19	92	2.24	7.28	1.94	1.94

¹Intra-residue restraints, ²Sequential restraints, ³Medium range restraints, ⁴Long range restraints,

⁵Inter-chain restraints, ⁶Standard 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

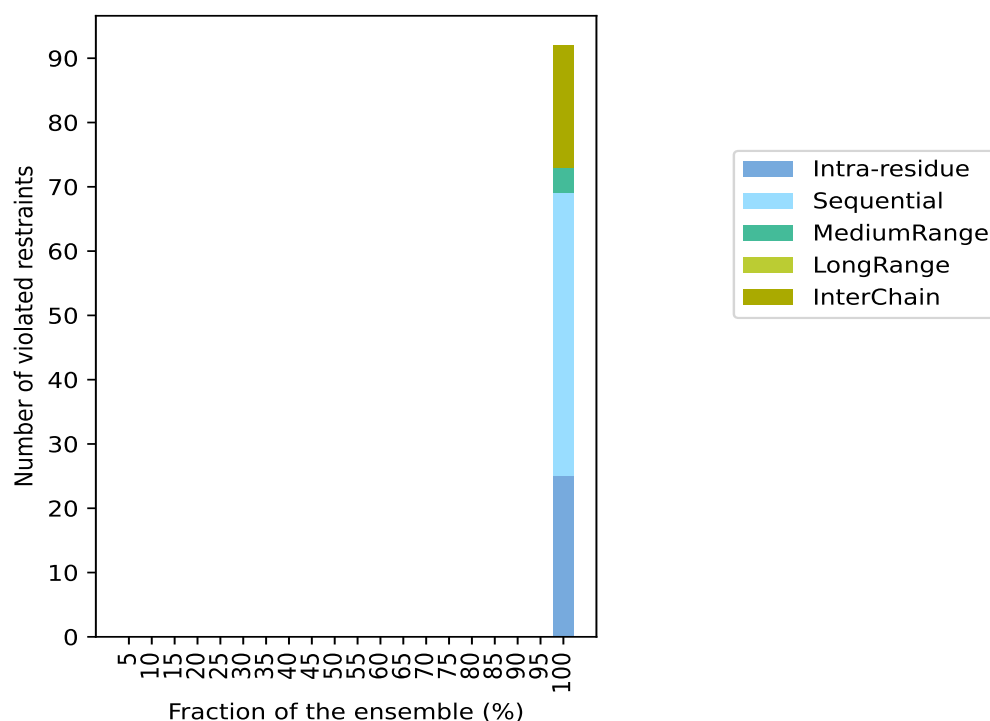
Violation analysis may find that some restraints are violated in few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of the ensemble. In total, 346(IR:203, SQ:132, MR:2, LR:0, IC:9) restraints are not 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	%
0	0	0	0	0	0	1	5.0
0	0	0	0	0	0	2	10.0
0	0	0	0	0	0	3	15.0
0	0	0	0	0	0	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	0	0	0	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	0	0	0	15	75.0
0	0	0	0	0	0	16	80.0
0	0	0	0	0	0	17	85.0
0	0	0	0	0	0	18	90.0
0	0	0	0	0	0	19	95.0
25	44	4	0	19	92	20	100.0

¹Intra-residue restraints, ²Sequential restraints, ³Medium range restraints, ⁴Long range restraints,

⁵Inter-chain restraints, ⁶ Number of models with violations

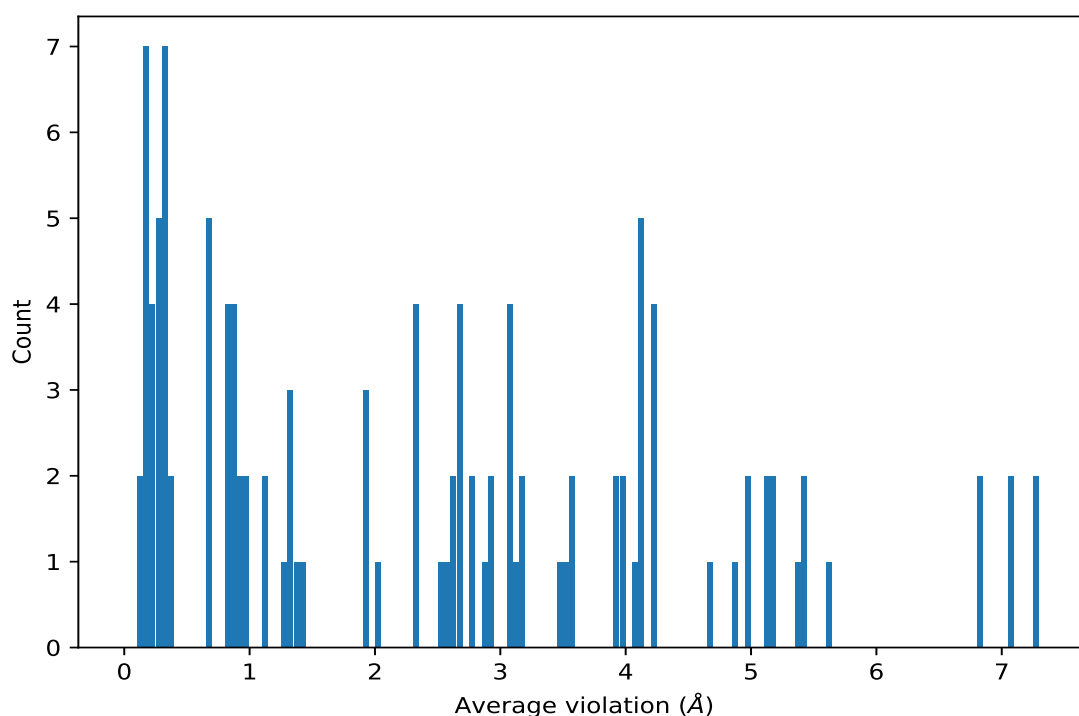
9.3.1 Bar graph : Distance violation statistics for the ensemble [i](#)



9.4 Most violated distance restraints in the ensemble [i](#)

9.4.1 Histogram : Distribution of mean distance violations [i](#)

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



9.4.2 Table: Most violated distance restraints [i](#)

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

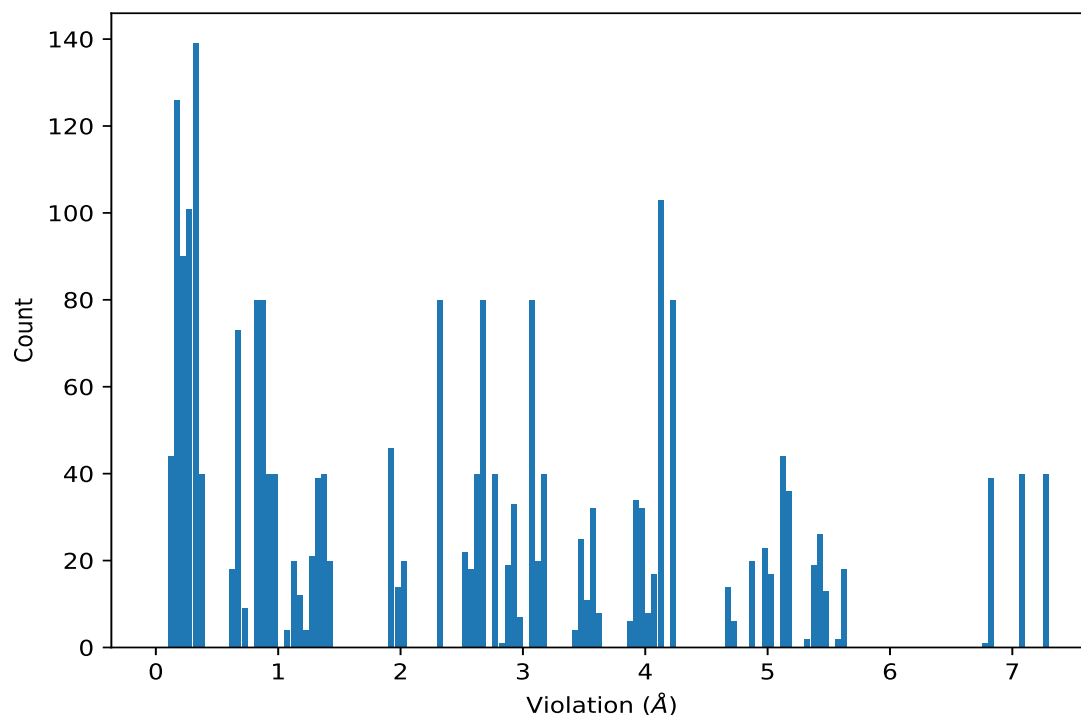
Key	Atom-1	Atom-2	Models ¹	Mean (Å)	SD ¹ (Å)	Median (Å)
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5'	20	7.28	0.0	7.28
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5''	20	7.28	0.0	7.28
(1,422)	1:14:B:GF2:H8	1:16:B:DT:H5'	20	7.07	0.01	7.07
(1,422)	1:14:B:GF2:H8	1:16:B:DT:H5''	20	7.07	0.01	7.07
(1,317)	1:16:B:DT:H2'	1:15:B:DT:H6	20	6.81	0.01	6.81
(1,315)	1:8:A:DT:H2'	1:7:A:DT:H6	20	6.81	0.01	6.81
(1,375)	1:6:A:GF2:H8	1:10:B:DA:H2	20	5.6	0.01	5.6
(1,376)	1:14:B:GF2:H8	1:2:A:DA:H2	20	5.44	0.01	5.44
(1,377)	1:6:A:GF2:H8	1:10:B:DA:H8	20	5.42	0.02	5.42
(1,378)	1:14:B:GF2:H8	1:2:A:DA:H8	20	5.36	0.01	5.36
(1,311)	1:8:A:DT:H2'	1:7:A:DT:H1'	20	5.15	0.01	5.15
(1,311)	1:8:A:DT:H2''	1:7:A:DT:H1'	20	5.15	0.01	5.15
(1,312)	1:16:B:DT:H2'	1:15:B:DT:H1'	20	5.14	0.01	5.14

¹Number of violated models, ²Standard 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)
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5'	13	7.29
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5''	13	7.29
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5'	14	7.29
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5''	14	7.29
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5'	19	7.29
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5''	19	7.29
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5'	1	7.28
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5''	1	7.28
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5'	2	7.28
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5''	2	7.28
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5'	3	7.28

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5''	3	7.28
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5'	5	7.28
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5''	5	7.28
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5'	6	7.28
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5''	6	7.28
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5'	7	7.28
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5''	7	7.28
(1,421)	1:6:A:GF2:H8	1:8:A:DT:H5'	8	7.28

10 Dihedral-angle violation analysis ⓘ

No dihedral-angle restraints found