

wwPDB NMR Structure Validation Summary Report (i)

Jun 3, 2023 – 09:38 AM EDT

PDB ID	:	5W77
BMRB ID	:	27144
Title	:	Solution structure of the MYC G-quadruplex bound to small molecule DC-34
Authors	:	Chen, X.; Walters, K.J.
Deposited on	:	2017-06-19

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 (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:

buster-report Percentile statistics	: :	4.02b-467 1.8.5 (274361), CSD as541be (2020) 1.1.7 (2018) 20191225.v01 (using entries in the PDB archive December 25th 2019) v 1n 11 5 13 A (Berjanski et al., 2005)
		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

Clashscore

1 Overall quality at a glance (i)

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

The overall completeness of chemical shifts assignment is 43%.

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.

Metri	c	Percent	Percentile Ranks				
Clashscore	e 📃			0			
	Worse			Better			
	Perce	ntile relative to all structures					
	Perce	ntile relative to all NMR structures					
Мани		Whole archive	NMR archive				
Metr	ic	(# Entries)	(# Entries)				

158937

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%

12864

Mol	Chain	Length	Quality of chain
1	А	22	100%



2 Ensemble composition and analysis (i)

This entry contains 15 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)

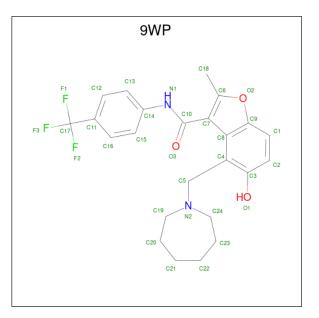
There are 3 unique types of molecules in this entry. The entry contains 832 atoms, of which 299 are hydrogens and 0 are deuteriums.

Mol	Chain	Residues		Trace					
1	Δ	22	Total	С	Η	Ν	0	Р	0
1 A		716	220	249	95	131	21	0	

• Molecule 2 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms
2	А	2	Total K 2 2

• Molecule 3 is 4-[(azepan-1-yl)methyl]-5-hydroxy-2-methyl-N-[4-(trifluoromethyl)phenyl]-1-b enzofuran-3-carboxamide (three-letter code: 9WP) (formula: $C_{24}H_{25}F_3N_2O_3$).



Mol	Chain	Residues		A	ton	ns		
9	3 A	1	Total	С	F	Η	Ν	0
0			57	24	3	25	2	3
9	А	1	Total	С	F	Η	Ν	0
3		1	57	24	3	25	2	3



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(*TP*GP*AP*GP*GP*GP*GP*TP*GP*GP*GP*TP*AP*GP*GP*GP*TP *GP*GP*TP*AP*A)-3')

Chain A:

100%

There are no outlier residues in this chain.

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(*TP*GP*AP*GP*GP*GP*GP*TP*GP*GP*GP*TP*AP*GP*GP*GP*TP *GP*GP*GP*TP*AP*A)-3')

Chain A:

100%

There are no outlier residues in this chain.



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *simulated annealing*.

Of the 100 calculated structures, 15 were deposited, based on the following criterion: *structures with the lowest energy*.

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

Software name	Classification	Version
X-PLOR NIH	refinement	2.45

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	253
Number of shifts mapped to atoms	253
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	43%



6 Model quality (i)

6.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: K, $9\mathrm{WP}$

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

6.2 Too-close contacts (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
All	All	7995	4485	3735	-

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

There are no clashes.

6.3 Torsion angles (i)

6.3.1 Protein backbone (i)

There are no protein molecules in this entry.

6.3.2 Protein sidechains (i)

There are no protein molecules in this entry.

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)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 for Mogul analysis.

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.

Mal	Mol Type		Dec	Tiple		Bond len	gths
	туре	Unam	nes		Counts	RMSZ	#Z>2
3	9WP	А	104	-	$31,\!35,\!35$	$1.52{\pm}0.02$	5±0 (16±0%)
3	9WP	А	103	-	31,35,35	$1.49{\pm}0.01$	5±0 (16±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	Dec Link		Bond angles			
	Type	Chain	nes	LIIIK	Counts	RMSZ	#Z>2	
3	9WP	А	104	-	40,51,51	$1.87{\pm}0.02$	8±0 (21±1%)	
3	9WP	А	103	-	40,51,51	$1.82{\pm}0.01$	8±0 (19±1%)	

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	Chirals Torsions	
3	9WP	А	103	-	-	$0\pm0,\!14,\!27,\!27$	$0\pm0,4,4,4$
3	9WP	А	104	-	-	$0\pm0,\!14,\!27,\!27$	$0\pm0,4,4,4$

5 of 10 unique bond outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Iol Chain R		Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
10101	Ullalli	\mathbf{Res}	Type	Atoms		Observeu(A)	Iueai(A)	Worst	Total
3	А	104	9WP	C7-C10	4.57	1.44	1.51	1	15
3	А	103	9WP	C7-C10	4.43	1.44	1.51	8	15
3	А	104	9WP	C5-C4	4.12	1.54	1.51	9	15
3	A	103	9WP	C5-C4	3.94	1.54	1.51	5	15
3	А	104	9WP	C3-C4	3.22	1.41	1.38	11	15

5 of 17 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	$Observed(^{o})$	$Ideal(^{o})$	Moo	dels
	Ullaili	nes	туре	Atoms				Worst	Total
3	А	104	9WP	C7-C8-C4	4.73	138.26	132.03	12	15
3	А	103	9WP	C7-C8-C4	4.67	138.19	132.03	5	15
3	А	104	9WP	C7-C8-C9	4.56	104.13	110.52	15	15
3	А	103	9WP	C7-C8-C9	4.49	104.23	110.52	10	15
3	А	104	9WP	C6-C7-C8	4.45	104.93	108.57	1	15

There are no chirality outliers.

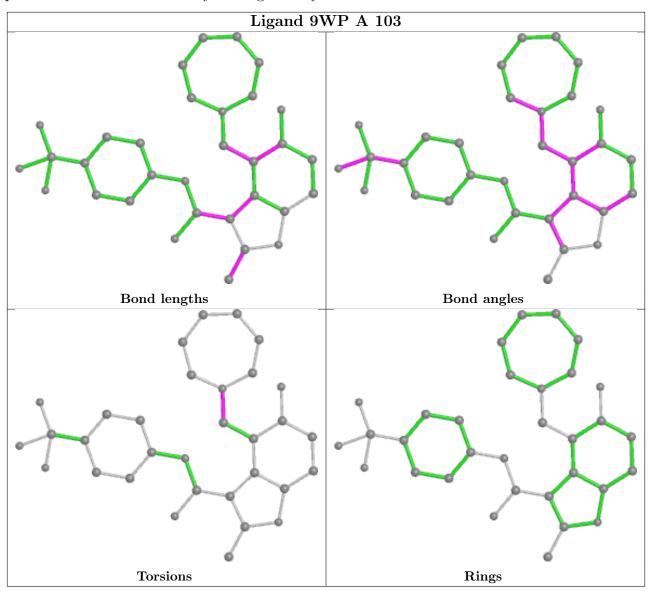
There are no torsion outliers.

All unique ring outliers are listed below.

Mol	Chain	Res	Type	Atoms	Models (Total)
3	А	104	9WP	C19-C20-C21-C22-C23-C24-N2	2

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient

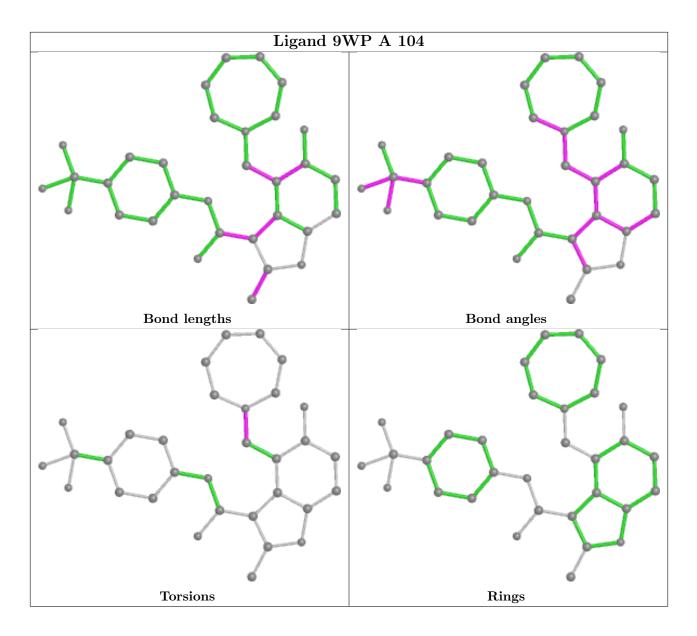




equivalents in the CSD to analyse the geometry.







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 43% for the well-defined parts and 43% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: Assignment_STAR_3.1.txt

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	253
Number of shifts mapped to atoms	253
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

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 43%, i.e. 192 atoms were assigned a chemical shift out of a possible 443. 0 out of 0 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^{1}\mathrm{H}$	$^{13}\mathrm{C}$	$^{15}\mathbf{N}$
Sugar	154/264~(58%)	154/154~(100%)	0/110~(0%)	$0/0 \ (\%)$
Base	38/179~(21%)	38/113~(34%)	0/31~(0%)	0/35~(0%)
Overall	192/443~(43%)	192/267~(72%)	0/141~(0%)	0/35~(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 (i)

8.1 Conformationally restricting restraints (i)

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	1023
Intra-residue (i-j =0)	610
Sequential (i-j =1)	254
Medium range ($ i-j >1$ and $ i-j <5$)	52
Long range $(i-j \ge 5)$	107
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	12
Number of unmapped restraints	0
Number of restraints per residue	47.0
Number of long range restraints per residue ¹	4.9

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 (i)

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 (i)

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)	55.8	0.2
0.2-0.5 (Medium)	17.8	0.5
>0.5 (Large)	2.1	1.64



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

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

Bins $(^{\circ})$	Average number of violations per model	Max ($^{\circ}$)
1.0-10.0 (Small)	0.1	3.3
10.0-20.0 (Medium)	None	None
>20.0 (Large)	None	None



9 Distance violation analysis (i)

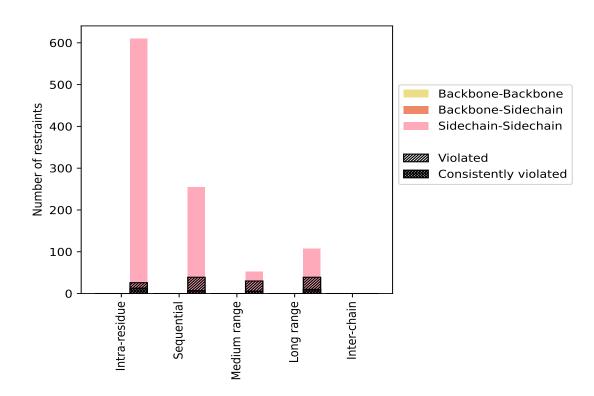
9.1 Summary of distance violations (i)

The following table shows the summary of distance violations in different restraint categories based on the sequence separation of the atoms involved. Each category is further sub-divided into three sub-categories based on the atoms involved. Violations less than 0.1 Å are not included in the statistics.

Destroints type	Count	$\%^1$	Vi	olated	3	Consis	tently	Violated ⁴
Restraints type	Count	701	Count	$\%^2$	$\%^1$	Count	$\%^2$	$\%^1$
Intra-residue (i-j =0)	610	59.6	26	4.3	2.5	13	2.1	1.3
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	610	59.6	26	4.3	2.5	13	2.1	1.3
Sequential (i-j =1)	254	24.8	39	15.4	3.8	7	2.8	0.7
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	254	24.8	39	15.4	3.8	7	2.8	0.7
Medium range ($ i-j > 1 \& i-j < 5$)	52	5.1	30	57.7	2.9	6	11.5	0.6
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	52	5.1	30	57.7	2.9	6	11.5	0.6
Long range $(i-j \ge 5)$	107	10.5	39	36.4	3.8	10	9.3	1.0
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	107	10.5	39	36.4	3.8	10	9.3	1.0
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	1023	100.0	134	13.1	13.1	36	3.5	3.5
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	1023	100.0	134	13.1	13.1	36	3.5	3.5

 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 disulfied 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.

Madal ID		Nun	nber o	f viola	ations	5	Maan (Å)	Max (Å)	SD^6 (Å)	Madian (Å)
Model ID	IR^1	SQ^2	MR^3	LR^4	$ IC^5 $	Total	Mean (Å)	Max (A)	$SD^{*}(A)$	Median (Å)
1	19	18	19	20	0	76	0.21	1.6	0.19	0.16
2	18	18	18	23	0	77	0.21	1.58	0.18	0.17
3	17	17	19	24	0	77	0.2	1.57	0.18	0.16
4	17	19	16	24	0	76	0.22	1.56	0.17	0.17
5	18	16	16	22	0	72	0.21	1.64	0.19	0.16
6	17	17	18	24	0	76	0.2	1.58	0.18	0.16
7	17	22	17	22	0	78	0.23	1.52	0.18	0.17
8	15	19	18	21	0	73	0.22	1.58	0.19	0.17
9	16	22	18	25	0	81	0.23	1.56	0.2	0.17
10	20	17	19	19	0	75	0.21	1.58	0.18	0.17
11	20	14	17	23	0	74	0.21	1.6	0.19	0.16

Continued on next page...

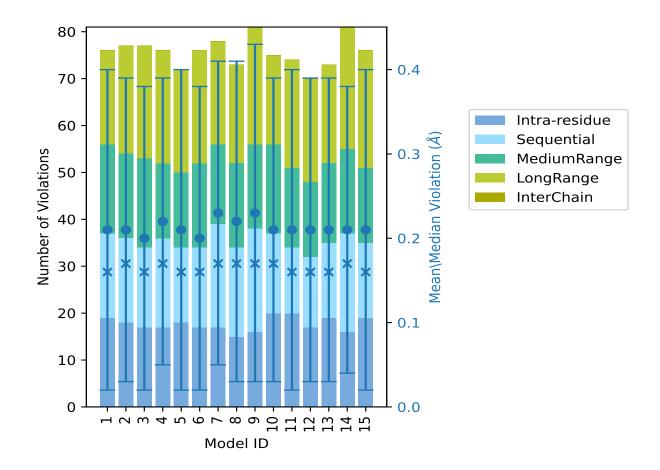


MadalID		Nur	nber o	f viola	ations	5	Mean (Å)	Max (Å)	SD^6 (Å)	Median (Å)	
Model ID	IR^{1}	SQ^2	MR^3	LR^4	IC^5	Total	Mean (A)	Max (A)	$SD^*(A)$	Median (A)	
12	17	15	16	22	0	70	0.21	1.55	0.18	0.16	
13	19	16	17	21	0	73	0.21	1.57	0.18	0.16	
14	16	21	18	26	0	81	0.21	1.51	0.17	0.17	
15	19	16	16	25	0	76	0.21	1.55	0.19	0.16	

Continued from previous page...

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 (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 (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

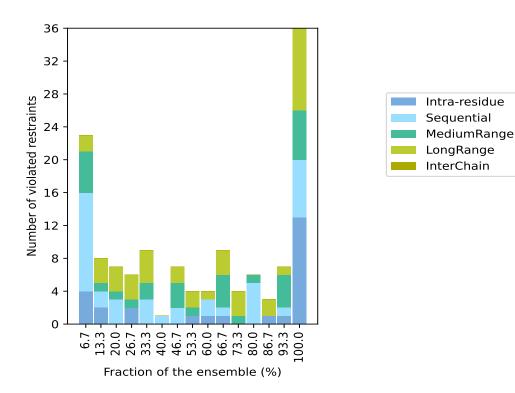


Nu	mber	of vio	lated	restra	aints	Fraction of the ensemble				
IR^1	SQ^2	MR^3	LR ⁴	IC ⁵	Total	Count^6	%			
4	12	5	2	0	23	1	6.7			
2	2	1	3	0	8	2	13.3			
0	3	1	3	0	7	3	20.0			
2	0	1	3	0	6	4	26.7			
0	3	2	4	0	9	5	33.3			
0	1	0	0	0	1	6	40.0			
0	2	3	2	0	7	7	46.7			
1	0	1	2	0	4	8	53.3			
1	2	0	1	0	4	9	60.0			
1	1	4	3	0	9	10	66.7			
0	0	1	3	0	4	11	73.3			
0	5	1	0	0	6	12	80.0			
1	0	0	2	0	3	13	86.7			
1	1	4	1	0	7	14	93.3			
13	7	6	10	0	36	15	100.0			

for a given fraction of the ensemble. In total, 889(IR:584, SQ:215, MR:22, LR:68, IC:0) restraints are not violated in the ensemble.

 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





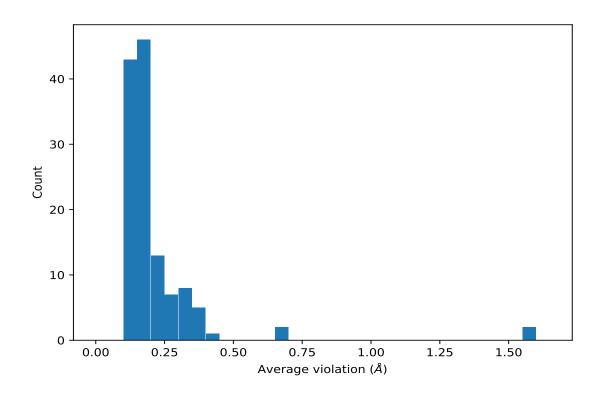
9.3.1 Bar graph : Distance violation statistics for the ensemble (i)

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)
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5'	15	1.57	0.03	1.57
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5"	15	1.57	0.03	1.57
(3,858)	1:A:4:DT:H5'	1:A:6:DA:H2	15	0.65	0.1	0.67
(3,858)	1:A:4:DT:H5"	1:A:6:DA:H2	15	0.65	0.1	0.67
(3,822)	1:A:5:DG:H2'	1:A:5:DG:H4'	15	0.36	0.0	0.36
(3,339)	1:A:8:DG:H1'	1:A:8:DG:H8	15	0.35	0.01	0.35
(3,312)	1:A:17:DG:H1'	1:A:17:DG:H8	15	0.32	0.05	0.35
(3,870)	1:A:23:DT:H5'	1:A:23:DT:H6	15	0.31	0.04	0.31
(3,870)	1:A:23:DT:H5"	1:A:23:DT:H6	15	0.31	0.04	0.31
(1,50)	1:A:16:DG:OP1	1:A:12:DG:H22	15	0.31	0.05	0.33
(3,838)	1:A:5:DG:H3'	1:A:4:DT:H2'	15	0.29	0.01	0.29
(3,275)	1:A:20:DG:H1'	1:A:20:DG:H8	15	0.29	0.06	0.32
(3,199)	1:A:12:DG:H1'	1:A:12:DG:H8	15	0.25	0.08	0.2

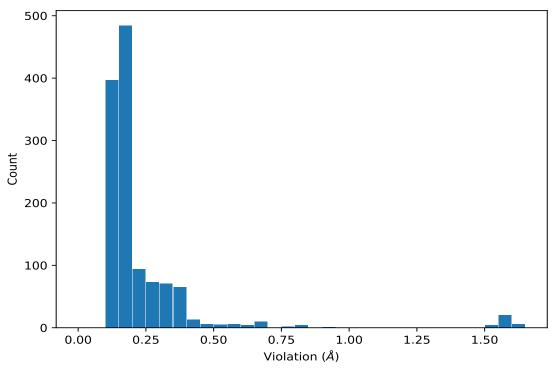
 $^1\mathrm{Number}$ of violated models, $^2\mathrm{Standard}$ deviation



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)
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5'	5	1.64
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5"	5	1.64
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5'	1	1.6
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5"	1	1.6
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5'	11	1.6
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5"	11	1.6
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5'	2	1.58
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5"	2	1.58
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5'	6	1.58
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5"	6	1.58
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5'	8	1.58

Continued on next page...



Key	Atom-1	Atom-2	Model ID	Violation (Å)
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5"	8	1.58
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5'	10	1.58
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5"	10	1.58
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5'	3	1.57
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5"	3	1.57
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5'	13	1.57
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5"	13	1.57
(3,419)	1:A:22:DG:H8	1:A:23:DT:H5'	4	1.56

Continued from previous page...



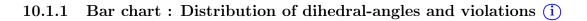
10 Dihedral-angle violation analysis (i)

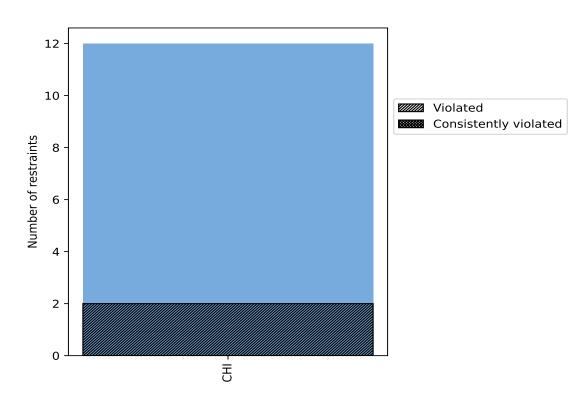
10.1 Summary of dihedral-angle violations (i)

The following table provides the summary of dihedral-angle violations in different dihedral-angle types. Violations less than 1° are not included in the calculation.

	Count	071					Consistently Violated ⁴		
Angle type	Count	$\%^1$	Count	$\%^2$	$\%^1$	Count	$\%^2$	$\%^1$	
CHI	12	100.0	2	16.7	16.7	0	0.0	0.0	
Total	12	100.0	2	16.7	16.7	0	0.0	0.0	

 1 percentage calculated with respect to total number of dihedral-angle restraints, 2 percentage calculated with respect to number of restraints in a particular dihedral-angle type, 3 violated in at least one model, 4 violated in all the models





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

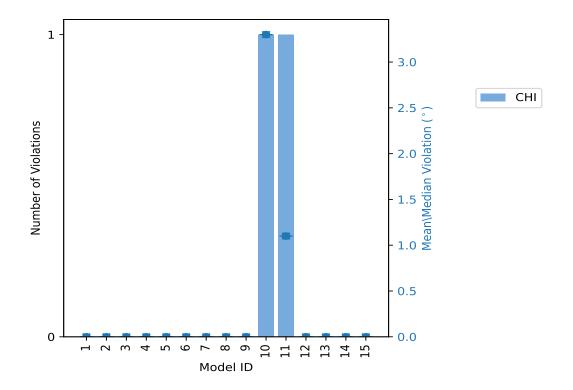


10.2 Dihedral-angle violation statistics for each model (i)

The following table provides the dihedral-angle violation statistics for each model in the ensemble. Violations less than 1° are not included in the statistics.

Model ID	Number of violations		Mean (°)	Max (°)	SD (°)	Median (°)
Model ID	CHI	Total	Mean ()	Max ()	SD ()	Median ()
1	0	0	0.0	0.0	0.0	0.0
2	0	0	0.0	0.0	0.0	0.0
3	0	0	0.0	0.0	0.0	0.0
4	0	0	0.0	0.0	0.0	0.0
5	0	0	0.0	0.0	0.0	0.0
6	0	0	0.0	0.0	0.0	0.0
7	0	0	0.0	0.0	0.0	0.0
8	0	0	0.0	0.0	0.0	0.0
9	0	0	0.0	0.0	0.0	0.0
10	1	1	3.3	3.3	0.0	3.3
11	1	1	1.1	1.1	0.0	1.1
12	0	0	0.0	0.0	0.0	0.0
13	0	0	0.0	0.0	0.0	0.0
14	0	0	0.0	0.0	0.0	0.0
15	0	0	0.0	0.0	0.0	0.0

10.2.1	Bar graph :	Dihedral	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

10.3 Dihedral-angle violation statistics for the ensemble (i)

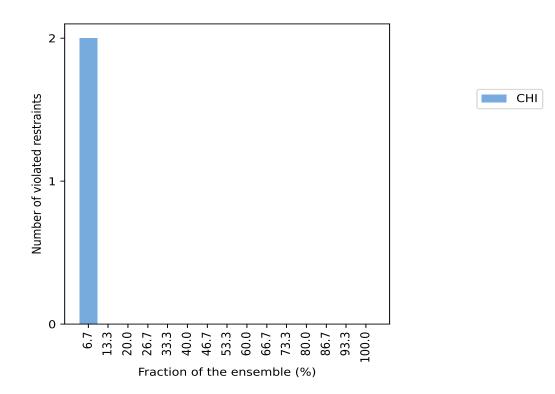
Violation analysis may find that some restraints are violated in very few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of ensemble.

Num	ber of violated restraints	Fractio	n of the ensemble
CHI	Total	Count^1	%
2	2	1	6.7
0	0	2	13.3
0	0	3	20.0
0	0	4	26.7
0	0	5	33.3
0	0	6	40.0
0	0	7	46.7
0	0	8	53.3
0	0	9	60.0
0	0	10	66.7
0	0	11	73.3
0	0	12	80.0
0	0	13	86.7
0	0	14	93.3
0	0	15	100.0

¹ Number of models with violations



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



10.4 Most violated dihedral-angle restraints in the ensemble (i)

No violations found

10.5 All violated dihedral-angle restraints (i)

10.5.1 Histogram : Distribution of violations (i)

The following histogram shows the distribution of the absolute value of the violation for all violated restraints in the ensemble.

Data insufficient to plot histogram

10.5.2 Table: All violated dihedral-angle restraints (i)

The following table lists the absolute value of the violation for each restraint in the ensemble sorted by its value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint.

Key	Atom-1	Atom-2	Atom-3	Atom-4	Model ID	Violation ($^{\circ}$)
(1,6)	1:A:13:DG:O4'	1:A:13:DG:C1'	1:A:13:DG:N9	1:A:13:DG:C4	10	3.3
(1,9)	1:A:18:DG:O4'	1:A:18:DG:C1'	1:A:18:DG:N9	1:A:18:DG:C4	11	1.1

