

# wwPDB NMR Structure Validation Summary Report (i)

Jun 4, 2023 – 03:56 AM EDT

PDB ID : 2L1N BMRB ID : 17090

Title: Solution NMR structure of the protein YP 399305.1

Authors: Mohanty, B.; Serrano, P.; Geralt, M.; Horst, R.; Wuthrich, K.; Joint Center

for Structural Genomics (JCSG)

Deposited on : 2010-07-30

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)

 $\begin{array}{ccc} wwPDB\text{-}ShiftChecker &: & v1.2 \\ BMRB \ Restraints \ Analysis &: & v1.2 \\ \end{array}$ 

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

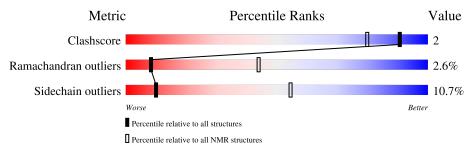
Validation Pipeline (wwPDB-VP) : 2.33

# 1 Overall quality at a glance (i)

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

The overall completeness of chemical shifts assignment is 91%.

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 $(\# \mathrm{Entries})$	$egin{array}{c} { m NMR \ archive} \ (\#{ m Entries}) \end{array}$	
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	120	79%	10%	11%



# 2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 1 is the overall representative, medoid model (most similar to other models).

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 mode						
1	A:2-A:89,	A:100-A:118	0.70	1		
	(107)					

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 3 clusters and 1 single-model cluster was found.

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



# 3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1941 atoms, of which 964 are hydrogens and 0 are deuteriums.

• Molecule 1 is a protein called Uncharacterized protein.

Mol	Chain	Residues	Atoms				Trace		
1	٨	120	Total	С	Н	N	О	S	0
1	A	120	1941	627	964	165	182	3	U

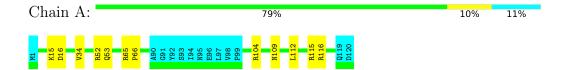


# 4 Residue-property plots (i)

#### 4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

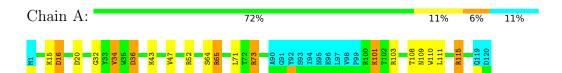
• Molecule 1: Uncharacterized protein



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

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

• Molecule 1: Uncharacterized protein





#### Refinement protocol and experimental data overview (i) 5



The models were refined using the following method: torsion angle dynamics.

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

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

Software name	Classification	Version
CYANA	structure solution	CYANA3.0
OPAL	refinement	
UNIO	structure solution	2.0.0

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	1529
Number of shifts mapped to atoms	1529
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	91%



# 6 Model quality (i)

### 6.1 Standard geometry (i)

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

Mol	l Chain B		Sond lengths	Bond angles	
WIOI	Chain	RMSZ	#Z>5	RMSZ	#Z>5
1	A	$0.61 \pm 0.01$	$0\pm0/903~(~0.0\pm~0.0\%)$	$1.08\pm0.03$	$1\pm1/1234~(~0.1\pm~0.1\%)$
All	All	0.61	0/18060 ( 0.0%)	1.08	20/24680 ( 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	$0.0\pm0.0$	$1.8 \pm 0.9$
All	All	0	36

There are no bond-length outliers.

5 of 16 unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	$f Res \ Type \ Atoms \ Z \ Observed (^o)$		$\operatorname{Ideal}({}^{o})$	Models				
MIOI	Chain	nes	туре	Atoms	L	Observed()	ideai()	Worst	Total
1	A	100	ARG	NE-CZ-NH2	-5.97	117.31	120.30	8	1
1	A	52	ARG	NE-CZ-NH2	-5.77	117.42	120.30	8	2
1	A	34	VAL	CA-CB-CG1	5.74	119.52	110.90	13	1
1	A	6	THR	CA-CB-CG2	-5.59	104.58	112.40	12	1
1	A	47	VAL	CA-CB-CG1	5.58	119.27	110.90	5	3

There are no chirality outliers.

5 of 11 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	55	TYR	Sidechain	8
1	A	65	ARG	Sidechain	7
1	A	116	ARG	Sidechain	4

Continued on next page...



Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	52	ARG	Sidechain	4
1	A	73	ARG	Sidechain	3

### 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	878	871	871	3±2
All	All	17560	17420	17420	57

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 2.

5 of 39 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	$Distance(\mathring{A})$	Models	
Atom-1	Atom-2	Clash(A)	Distance(A)	Worst	Total
1:A:79:TYR:HB3	1:A:82:LEU:HD21	0.74	1.57	18	2
1:A:79:TYR:CB	1:A:82:LEU:HD21	0.57	2.28	18	2
1:A:34:VAL:HG13	1:A:43:LYS:HE3	0.56	1.77	16	2
1:A:116:ARG:HD2	1:A:117:SER:N	0.56	2.15	20	1
1:A:82:LEU:HD23	1:A:85:GLU:CD	0.54	2.21	5	1

### 6.3 Torsion angles (i)

### 6.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the backbone conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	107/120 (89%)	96±3 (90±3%)	8±3 (8±2%)	3±1 (3±1%)	8 44
All	All	2140/2400 (89%)	1916 (90%)	169 (8%)	55 (3%)	8 44

5 of 6 unique Ramachandran outliers are listed below. They are sorted by the frequency of



occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	66	PRO	15
1	A	15	LYS	13
1	A	16	ASP	13
1	A	34	VAL	12
1	A	20	ASP	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	Perce	entiles
1	A	92/103 (89%)	82±3 (89±3%)	10±3 (11±3%)	10	54
All	All	1840/2060 (89%)	1643 (89%)	197 (11%)	10	54

5 of 46 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	109	ASN	11
1	A	53	GLN	11
1	A	64	SER	10
1	A	65	ARG	9
1	A	104	ARG	9

#### 6.3.3 RNA (i)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.5 Carbohydrates (i)

There are no monosaccharides in this entry.



### 6.6 Ligand geometry (i)

There are no ligands in this entry.

### 6.7 Other polymers (i)

There are no such molecules in this entry.

### 6.8 Polymer linkage issues (i)

There are no chain breaks in this entry.



# 7 Chemical shift validation (i)

The completeness of assignment taking into account all chemical shift lists is 91% for the well-defined parts and 91% for the entire structure.

#### 7.1 Chemical shift list 1

File name: working cs.cif

Chemical shift list name: assigned\_chem\_shift\_list\_1

#### 7.1.1 Bookkeeping (i)

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

Total number of shifts	1529
Number of shifts mapped to atoms	1529
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	14

#### 7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, $ppm$	Suggested action
$^{13}\mathrm{C}_{\alpha}$	120	$-0.28 \pm 0.10$	None needed ( $< 0.5 \text{ ppm}$ )
$^{13}C_{\beta}$	116	$0.13 \pm 0.12$	None needed ( $< 0.5 \text{ ppm}$ )
<sup>13</sup> C′	91	$0.37 \pm 0.12$	None needed ( $< 0.5 \text{ ppm}$ )
$^{15}N$	110	$0.07 \pm 0.25$	None needed ( $< 0.5 \text{ ppm}$ )

### 7.1.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 91%, i.e. 1379 atoms were assigned a chemical shift out of a possible 1519. 0 out of 17 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^{1}\mathrm{H}$	$^{13}\mathbf{C}$	$^{15}{ m N}$
Backbone	499/524~(95%)	210/210 (100%)	189/214 (88%)	100/100 (100%)
Sidechain	772/867 (89%)	526/563 (93%)	234/267 (88%)	$12/37 \ (32\%)$

Continued on next page...



Continued from previous page...

	Total	$^{1}\mathrm{H}$	$^{13}\mathbf{C}$	$^{15}{ m N}$
Aromatic	108/128 (84%)	61/62 (98%)	41/60 (68%)	6/6 (100%)
Overall	1379/1519 (91%)	797/835~(95%)	464/541 (86%)	118/143 (83%)

#### 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.

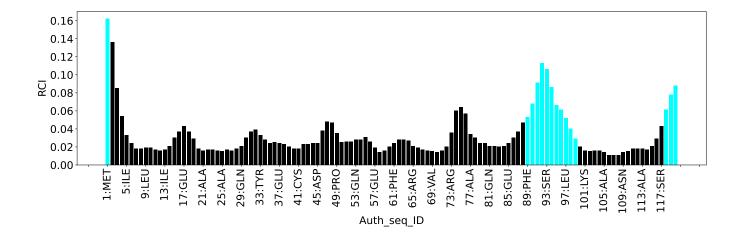
List Id	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	58	PRO	HB3	-1.75	0.25 - 3.76	-10.7
1	A	58	PRO	HG3	-1.45	0.33 - 3.48	-10.7
1	A	58	PRO	HB2	-1.50	0.37 - 3.78	-10.5
1	A	58	PRO	HG2	-1.22	0.41 - 3.45	-10.4
1	A	58	PRO	HD2	1.46	1.93 - 5.38	-6.3
1	A	58	PRO	HD3	1.54	1.76 - 5.48	-5.6
1	A	24	ASN	HB2	1.14	1.27 - 4.34	-5.4
1	A	71	LEU	HD11	-0.69	-0.61 - 2.12	-5.3
1	A	71	LEU	HD12	-0.69	-0.61 - 2.12	-5.3
1	A	71	LEU	HD13	-0.69	-0.61 - 2.12	-5.3
1	A	14	LEU	HD11	-0.67	-0.61 - 2.12	-5.2
1	A	14	LEU	HD12	-0.67	-0.61 - 2.12	-5.2
1	A	14	LEU	HD13	-0.67	-0.61 - 2.12	-5.2
1	A	59	PRO	HG2	0.40	0.41 - 3.45	-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:







### 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	2253
Intra-residue ( $ i-j =0$ )	577
Sequential ( $ i-j =1$ )	594
Medium range ( $ i-j >1$ and $ i-j <5$ )	556
Long range ( i-j ≥5)	526
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	18.8
Number of long range restraints per residue <sup>1</sup>	4.4

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

### 8.2 Residual restraint violations (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)	34.3	0.2
0.2-0.5 (Medium)	51.5	0.5
>0.5 (Large)	90.5	3.04



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

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



# 9 Distance violation analysis (i)

### 9.1 Summary of distance violations (i)

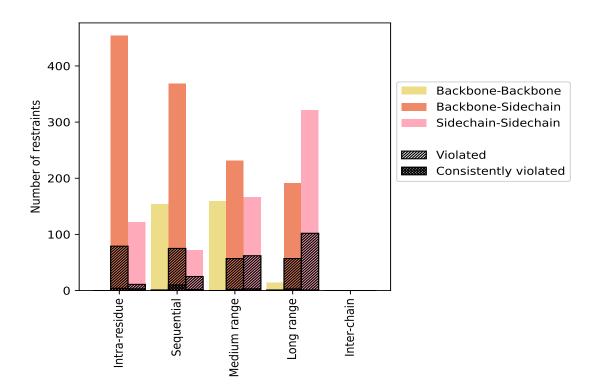
The following table shows the summary of distance violations in different restraint categories based on the sequence separation of the atoms involved. Each category is further sub-divided into three sub-categories based on the atoms involved. Violations less than 0.1 Å are not included in the statistics.

Doodnointe tour	Count	<b>%</b> <sup>1</sup>	Vi	${f Violated^3}$			Consistently Violated <sup>4</sup>		
Restraints type	Count	70	Count	$\%^2$	$\%^{1}$	Count	$\%^2$	$\%^1$	
Intra-residue ( i-j =0)	577	25.6	90	15.6	4.0	7	1.2	0.3	
Backbone-Backbone	1	0.0	0	0.0	0.0	0	0.0	0.0	
Backbone-Sidechain	454	20.2	79	17.4	3.5	4	0.9	0.2	
Sidechain-Sidechain	122	5.4	11	9.0	0.5	3	2.5	0.1	
Sequential ( i-j =1)	594	26.4	101	17.0	4.5	13	2.2	0.6	
Backbone-Backbone	154	6.8	1	0.6	0.0	1	0.6	0.0	
Backbone-Sidechain	368	16.3	75	20.4	3.3	10	2.7	0.4	
Sidechain-Sidechain	72	3.2	25	34.7	1.1	2	2.8	0.1	
Medium range ( $ i-j >1 \&  i-j <5$ )	556	24.7	119	21.4	5.3	5	0.9	0.2	
Backbone-Backbone	159	7.1	0	0.0	0.0	0	0.0	0.0	
Backbone-Sidechain	231	10.3	57	24.7	2.5	2	0.9	0.1	
Sidechain-Sidechain	166	7.4	62	37.3	2.8	3	1.8	0.1	
Long range ( $ i-j  \ge 5$ )	526	23.3	160	30.4	7.1	4	0.8	0.2	
Backbone-Backbone	14	0.6	1	7.1	0.0	0	0.0	0.0	
Backbone-Sidechain	191	8.5	57	29.8	2.5	3	1.6	0.1	
Sidechain-Sidechain	321	14.2	102	31.8	4.5	1	0.3	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	2253	100.0	470	20.9	20.9	29	1.3	1.3	
Backbone-Backbone	328	14.6	2	0.6	0.1	1	0.3	0.0	
Backbone-Sidechain	1244	55.2	268	21.5	11.9	19	1.5	0.8	
Sidechain-Sidechain	681	30.2	200	29.4	8.9	9	1.3	0.4	

 $<sup>^1</sup>$  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.

Model ID		Nun	nber o	f viola	ations	<u> </u>	Mean (Å)	Mar (Å)	$\mathbf{SD}^6$ (Å)	Median (Å)	
Model ID	$IR^1$	$SQ^2$	$MR^3$	$LR^4$	$IC^5$	Total	Mean (A)	Max (Å)	SD* (A)	Median (A)	
1	39	39	48	58	0	184	0.73	2.64	0.57	0.55	
2	38	39	42	43	0	162	0.56	1.95	0.44	0.44	
3	35	46	45	59	0	185	0.69	2.64	0.51	0.57	
4	35	39	49	68	0	191	0.74	2.62	0.6	0.57	
5	34	39	44	50	0	167	0.64	2.01	0.48	0.5	
6	33	39	32	33	0	137	0.55	2.12	0.43	0.43	
7	37	36	41	58	0	172	0.67	2.77	0.53	0.52	
8	38	39	55	68	0	200	0.77	2.63	0.6	0.64	
9	38	43	45	41	0	167	0.65	2.94	0.57	0.47	
10	37	45	58	66	0	206	0.83	2.94	0.63	0.66	
11	31	39	40	57	0	167	0.65	2.3	0.49	0.52	

Continued on next page...

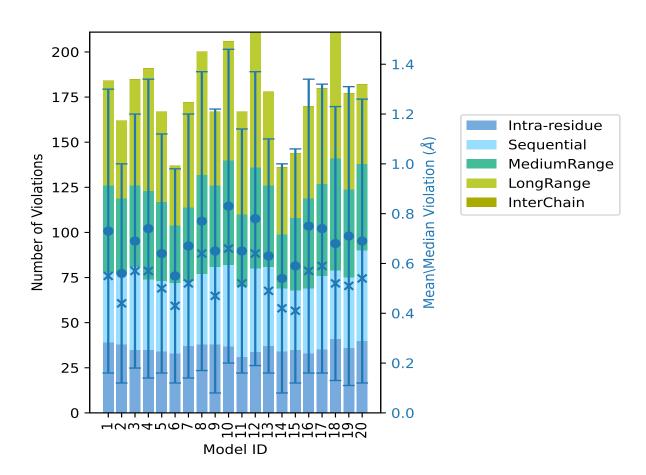


Continued from previous page...

Model ID		Nun	nber o	f viola	ations	3	Mean (Å)	Max (Å)	$SD^6$ (Å)	Median (Å)
Model 1D	$IR^1$	$SQ^2$	$MR^3$	$LR^4$	$IC^5$	Total	Mean (A)	Max (A)	$SD^*(A)$	Median (A)
12	34	46	56	75	0	211	0.78	2.76	0.59	0.64
13	37	44	45	52	0	178	0.63	2.36	0.47	0.49
14	34	35	30	37	0	136	0.54	2.76	0.46	0.42
15	35	33	40	36	0	144	0.59	2.23	0.47	0.41
16	33	36	50	51	0	170	0.75	2.82	0.59	0.57
17	35	41	51	53	0	180	0.74	3.0	0.58	0.59
18	41	38	62	70	0	211	0.68	2.8	0.55	0.52
19	36	39	49	53	0	177	0.71	2.69	0.6	0.51
20	40	50	48	44	0	182	0.69	3.04	0.57	0.54

 $<sup>^1</sup>$ 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 (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)

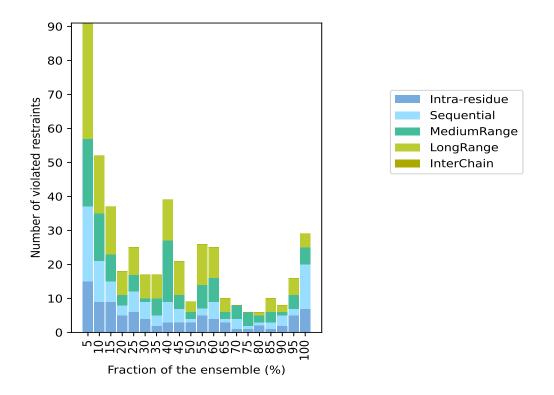
Violation analysis may find that some restraints are violated in few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of the ensemble. In total, 1783(IR:487, SQ:493, MR:437, LR:366, IC:0) restraints are not violated in the ensemble.

Nu	mber	of vio	lated	restra	aints	Fraction	n of the ensemble
$IR^1$	$SQ^2$	$MR^3$	$LR^4$	$IC^5$	Total	Count <sup>6</sup>	%
15	22	20	34	0	91	1	5.0
9	12	14	17	0	52	2	10.0
9	6	8	14	0	37	3	15.0
5	3	3	7	0	18	4	20.0
6	6	5	8	0	25	5	25.0
4	5	1	7	0	17	6	30.0
2	3	5	7	0	17	7	35.0
3	6	18	12	0	39	8	40.0
3	4	4	10	0	21	9	45.0
3	1	2	3	0	9	10	50.0
5	2	7	12	0	26	11	55.0
4	5	7	9	0	25	12	60.0
3	1	2	4	0	10	13	65.0
1	3	4	0	0	8	14	70.0
1	1	4	0	0	6	15	75.0
2	1	2	1	0	6	16	80.0
1	2	3	4	0	10	17	85.0
2	3	1	2	0	8	18	90.0
5	2	4	5	0	16	19	95.0
7	13	5	4	0	29	20	100.0

 $<sup>^1</sup>$ 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)

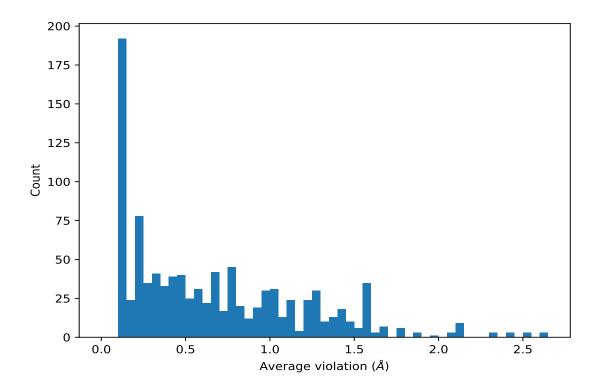


### 9.4 Most violated distance restraints in the ensemble (i)

### 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	$\mathbf{Models}^1$	Mean (Å)	$\mathrm{SD}^1$ (Å)	Median (Å)
(1,1432)	1:A:43:LYS:HG3	1:A:45:ASP:HB3	20	1.96	0.67	1.92
(1,875)	1:A:3:ILE:HB	1:A:29:GLN:HG2	20	1.55	0.18	1.6
(1,652)	1:A:18:LEU:HB2	1:A:24:ASN:H	20	1.36	0.26	1.48
(1,1177)	1:A:64:SER:HA	1:A:65:ARG:HB2	20	1.24	0.07	1.26
(1,1118)	1:A:28:TRP:HB2	1:A:29:GLN:HB2	20	1.2	0.12	1.19
(1,1178)	1:A:64:SER:HA	1:A:65:ARG:HG2	20	1.16	0.18	1.23
(1,655)	1:A:17:GLU:H	1:A:18:LEU:HB3	20	1.14	0.27	1.14
(1,606)	1:A:26:LEU:HB2	1:A:29:GLN:HE22	20	1.06	0.26	1.12
(1,1601)	1:A:35:TRP:HB3	1:A:42:TRP:HA	20	1.06	0.1	1.1
(1,502)	1:A:50:GLU:H	1:A:51:TRP:HB3	20	1.05	0.12	1.04

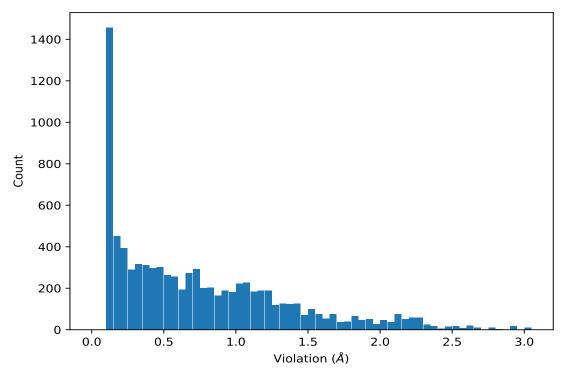
<sup>&</sup>lt;sup>1</sup>Number of violated models, <sup>2</sup>Standard deviation



### 9.5 All violated distance restraints (i)

#### 9.5.1 Histogram: Distribution of distance violations (i)

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



#### 9.5.2 Table : All distance violations (i)

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

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,1656)	1:A:34:VAL:HG21	1:A:47:VAL:HG11	20	3.04
(1,1656)	1:A:34:VAL:HG21	1:A:47:VAL:HG12	20	3.04
(1,1656)	1:A:34:VAL:HG21	1:A:47:VAL:HG13	20	3.04
(1,1656)	1:A:34:VAL:HG22	1:A:47:VAL:HG11	20	3.04
(1,1656)	1:A:34:VAL:HG22	1:A:47:VAL:HG12	20	3.04
(1,1656)	1:A:34:VAL:HG22	1:A:47:VAL:HG13	20	3.04
(1,1656)	1:A:34:VAL:HG23	1:A:47:VAL:HG11	20	3.04
(1,1656)	1:A:34:VAL:HG23	1:A:47:VAL:HG12	20	3.04
(1,1656)	1:A:34:VAL:HG23	1:A:47:VAL:HG13	20	3.04
(1,1432)	1:A:43:LYS:HG3	1:A:45:ASP:HB3	17	3.0
(1,1656)	1:A:34:VAL:HG21	1:A:47:VAL:HG11	9	2.94

Continued on next page...



Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,1656)	1:A:34:VAL:HG21	1:A:47:VAL:HG12	9	2.94
(1,1656)	1:A:34:VAL:HG21	1:A:47:VAL:HG13	9	2.94
(1,1656)	1:A:34:VAL:HG22	1:A:47:VAL:HG11	9	2.94
(1,1656)	1:A:34:VAL:HG22	1:A:47:VAL:HG12	9	2.94
(1,1656)	1:A:34:VAL:HG22	1:A:47:VAL:HG13	9	2.94
(1,1656)	1:A:34:VAL:HG23	1:A:47:VAL:HG11	9	2.94
(1,1656)	1:A:34:VAL:HG23	1:A:47:VAL:HG12	9	2.94
(1,1656)	1:A:34:VAL:HG23	1:A:47:VAL:HG13	9	2.94
(1,1656)	1:A:34:VAL:HG21	1:A:47:VAL:HG11	10	2.94
(1,1656)	1:A:34:VAL:HG21	1:A:47:VAL:HG12	10	2.94
(1,1656)	1:A:34:VAL:HG21	1:A:47:VAL:HG13	10	2.94
(1,1656)	1:A:34:VAL:HG22	1:A:47:VAL:HG11	10	2.94
(1,1656)	1:A:34:VAL:HG22	1:A:47:VAL:HG12	10	2.94
(1,1656)	1:A:34:VAL:HG22	1:A:47:VAL:HG13	10	2.94
(1,1656)	1:A:34:VAL:HG23	1:A:47:VAL:HG11	10	2.94
(1,1656)	1:A:34:VAL:HG23	1:A:47:VAL:HG12	10	2.94
(1,1656)	1:A:34:VAL:HG23	1:A:47:VAL:HG13	10	2.94
(1,222)	1:A:109:ASN:H	1:A:112:LEU:HD21	16	2.82
(1,222)	1:A:109:ASN:H	1:A:112:LEU:HD22	16	2.82
(1,222)	1:A:109:ASN:H	1:A:112:LEU:HD23	16	2.82
(1,1432)	1:A:43:LYS:HG3	1:A:45:ASP:HB3	18	2.8
(1,1432)	1:A:43:LYS:HG3	1:A:45:ASP:HB3	7	2.77
(1,1656)	1:A:34:VAL:HG21	1:A:47:VAL:HG11	12	2.76
(1,1656)	1:A:34:VAL:HG21	1:A:47:VAL:HG12	12	2.76
(1,1656)	1:A:34:VAL:HG21	1:A:47:VAL:HG13	12	2.76
(1,1656)	1:A:34:VAL:HG22	1:A:47:VAL:HG11	12	2.76
(1,1656)	1:A:34:VAL:HG22	1:A:47:VAL:HG12	12	2.76
(1,1656)	1:A:34:VAL:HG22	1:A:47:VAL:HG13	12	2.76
(1,1656)	1:A:34:VAL:HG23	1:A:47:VAL:HG11	12	2.76
(1,1656)	1:A:34:VAL:HG23	1:A:47:VAL:HG12	12	2.76
(1,1656)	1:A:34:VAL:HG23	1:A:47:VAL:HG13	12	2.76
(1,1432)	1:A:43:LYS:HG3	1:A:45:ASP:HB3	14	2.76
(1,1432)	1:A:43:LYS:HG3	1:A:45:ASP:HB3	16	2.74



# 10 Dihedral-angle violation analysis (i)

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

