wwPDB NMR Structure Validation Summary Report (i)

Jun 6, 2023 – 07:06 pm BST

PDB ID	:	6QXZ
BMRB ID	:	34368
Title	:	Solution structure of the ASHH2 CW domain with the N-terminal histone H3 tail mimicking peptide monomethylated on lysine 4
Authors	:	Dobrovolska, O.; Madeleine, N.; Teigen, K.; Halskau, O.; Bril'kov, M.
Deposited on	:	2019-03-08

This is a wwPDB NMR Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

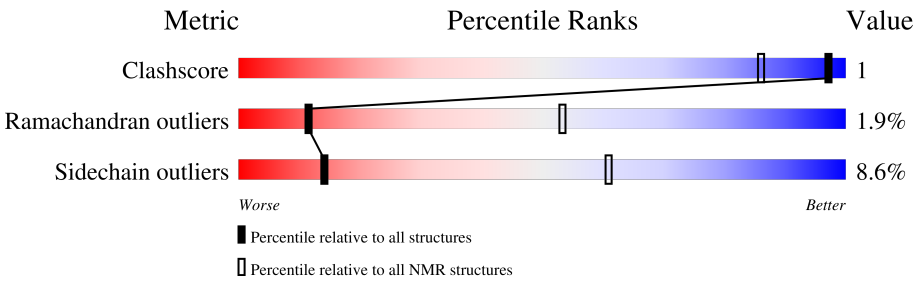
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.33

1 Overall quality at a glance (i)

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

The overall completeness of chemical shifts assignment is 80%.

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



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

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

Mol	Chain	Length	Quality of chain
1	A	79	63% 6% 30%
2	B	9	33% 67%

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 model**
1	A:863-A:885, A:890-A:921, B:3-B:3, B:5-B:6 (58)	0.48	1

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 2 clusters and 1 single-model cluster was found.

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

3 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 1348 atoms, of which 650 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Histone-lysine N-methyltransferase ASHH2.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	79	1185	370	566	112	131	6	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	850	GLY	-	expression tag	UNP Q2LAE1
A	851	SER	-	expression tag	UNP Q2LAE1
A	852	ARG	-	expression tag	UNP Q2LAE1
A	853	ARG	-	expression tag	UNP Q2LAE1
A	854	ALA	-	expression tag	UNP Q2LAE1
A	855	SER	-	expression tag	UNP Q2LAE1
A	856	VAL	-	expression tag	UNP Q2LAE1
A	857	GLY	-	expression tag	UNP Q2LAE1
A	858	SER	-	expression tag	UNP Q2LAE1
A	859	GLU	-	expression tag	UNP Q2LAE1
A	860	PHE	-	expression tag	UNP Q2LAE1

- Molecule 2 is a protein called ALA-ARG-THR-MLZ-GLN-THR-ALA-ARG-TYR.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
2	B	9	162	47	84	17	14	0

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	
			Total	Zn
3	A	1	1	1

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

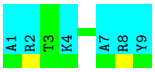
These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for 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: Histone-lysine N-methyltransferase ASHH2

Chain A: 63% 6% 30%

• Molecule 2: ALA-ARG-THR-MLZ-GLN-THR-ALA-ARG-TYR

Chain B: 33% 67%



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

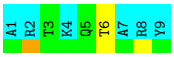
The representative model is number 1. Colouring as in section 4.1 above.

• Molecule 1: Histone-lysine N-methyltransferase ASHH2

Chain A: 59% 9% • 30%

• Molecule 2: ALA-ARG-THR-MLZ-GLN-THR-ALA-ARG-TYR

Chain B: 22% 11% 67%



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *distance geometry*.

Of the 20 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
Amber	refinement	
CYANA	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	880
Number of shifts mapped to atoms	879
Number of unparsed shifts	0
Number of shifts with mapping errors	1
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	80%

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: ZN, MLZ

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	0.74±0.01	0±0/454 (0.0± 0.0%)	1.15±0.02	2±1/611 (0.3± 0.1%)
2	B	0.61±0.06	0±0/22 (0.0± 0.0%)	1.05±0.15	0±0/29 (0.0± 0.0%)
All	All	0.74	0/9520 (0.0%)	1.14	37/12800 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	Chirality	Planarity
1	A	0.0±0.0	0.4±0.7
All	All	0	8

There are no bond-length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	867	ARG	NE-CZ-NH1	8.02	124.31	120.30	8	12
1	A	876	ARG	NE-CZ-NH1	7.54	124.07	120.30	3	13
1	A	890	ARG	NE-CZ-NH1	6.84	123.72	120.30	20	4
1	A	900	ARG	NE-CZ-NH1	6.72	123.66	120.30	15	2
1	A	875	ARG	NE-CZ-NH1	6.22	123.41	120.30	14	6

There are no chirality outliers.

All 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	867	ARG	Sidechain	4
1	A	876	ARG	Sidechain	4

6.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	445	420	420	1±1
2	B	23	22	22	0±0
All	All	9380	8840	8840	11

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:864:ALA:HB3	1:A:882:VAL:HG11	0.58	1.76	20	4
1:A:873:LYS:HE2	1:A:893:CYS:SG	0.53	2.43	18	4
1:A:899:LYS:CA	1:A:899:LYS:HE3	0.44	2.42	13	1
1:A:919:LEU:CD2	2:B:5:GLN:H	0.41	2.29	18	1
1:A:867:ARG:HB2	1:A:874:TRP:CZ3	0.40	2.50	4	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	55/79 (70%)	50±1 (90±2%)	4±1 (8±2%)	1±0 (2±0%)	12	54
2	B	3/9 (33%)	2±1 (68±20%)	1±1 (28±19%)	0±0 (3±10%)	6	37
All	All	1160/1760 (66%)	1031 (89%)	107 (9%)	22 (2%)	11	53

All 3 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	872	PHE	20
2	B	3	THR	1
2	B	6	THR	1

6.3.2 Protein sidechains (i)

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	50/68 (74%)	46±1 (92±3%)	4±1 (8±3%)	16 64
2	B	3/6 (50%)	2±1 (78±22%)	1±1 (22±22%)	3 30
All	All	1060/1480 (72%)	969 (91%)	91 (9%)	14 61

5 of 17 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	895	ASN	20
1	A	882	VAL	16
1	A	876	ARG	11
1	A	867	ARG	9
2	B	6	THR	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)

1 non-standard protein/DNA/RNA residue 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	MLZ	B	4	2	8,9,10	0.86±0.07	0±0 (0±2%)

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	MLZ	B	4	2	4,9,11	0.79±0.29	0±0 (2±7%)

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	MLZ	B	4	2	-	0±0,7,8,10	-

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
2	B	4	MLZ	CB-CA	2.13	1.56	1.53	11	1

All unique angle outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
2	B	4	MLZ	CM-NZ-CE	2.71	104.13	111.95	16	2

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6.5 Carbohydrates (i)

There are no monosaccharides in this entry.

6.6 Ligand geometry (i)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

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 80% for the well-defined parts and 77% 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 (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	880
Number of shifts mapped to atoms	879
Number of unparsed shifts	0
Number of shifts with mapping errors	1
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	15

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

- No matching atom found in the structure. All 1 occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	B	4	MLZ	HC	1.031	0.020	.

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	80	2.38 ± 0.16	Should be applied
$^{13}C_{\beta}$	76	2.74 ± 0.17	Should be applied
$^{13}C'$	0	—	None (insufficient data)
^{15}N	80	-0.93 ± 0.55	None needed (imprecise)

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 80%, i.e. 624 atoms were assigned a chemical shift out of a possible 783. 0 out of 4 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	224/290 (77%)	116/117 (99%)	54/116 (47%)	54/57 (95%)
Sidechain	352/437 (81%)	241/279 (86%)	103/134 (77%)	8/24 (33%)
Aromatic	48/56 (86%)	26/28 (93%)	19/25 (76%)	3/3 (100%)
Overall	624/783 (80%)	383/424 (90%)	176/275 (64%)	65/84 (77%)

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.

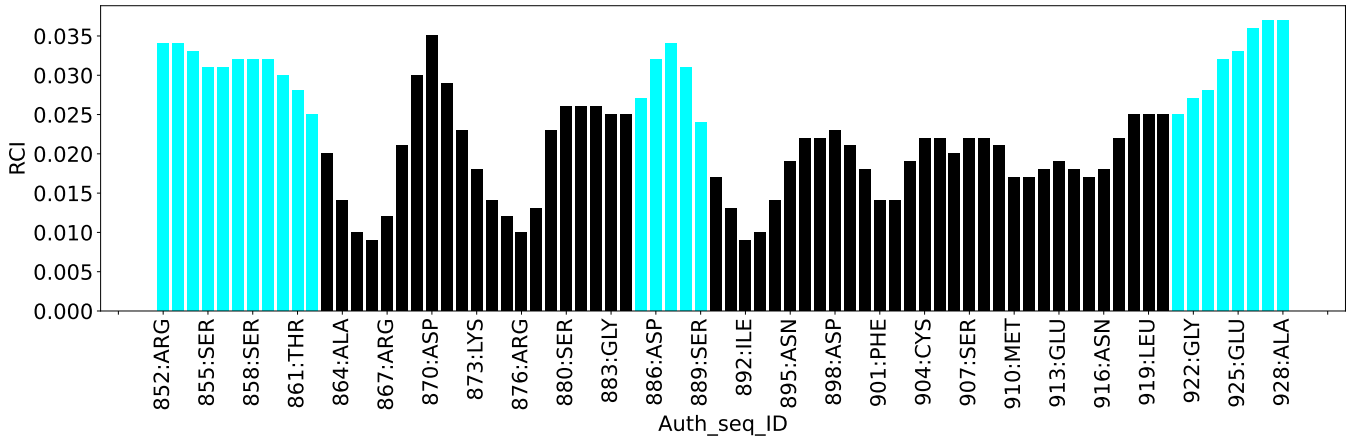
List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	876	ARG	NE	132.31	76.53 – 92.65	29.6
1	B	6	THR	HG1	5.77	0.08 – 2.19	22.0
1	A	875	ARG	NE	119.31	76.53 – 92.65	21.5
1	A	867	ARG	NE	118.83	76.53 – 92.65	21.2
1	B	3	THR	HG1	5.61	0.08 – 2.19	21.2
1	A	915	ILE	HG12	-1.35	-0.69 – 3.24	-6.7
1	A	915	ILE	HD11	-0.95	-0.72 – 2.09	-5.8
1	A	915	ILE	HD12	-0.95	-0.72 – 2.09	-5.8
1	A	915	ILE	HD13	-0.95	-0.72 – 2.09	-5.8
1	A	866	VAL	HG21	-0.78	-0.58 – 2.19	-5.7
1	A	866	VAL	HG22	-0.78	-0.58 – 2.19	-5.7
1	A	866	VAL	HG23	-0.78	-0.58 – 2.19	-5.7
1	A	876	ARG	HB2	0.33	0.52 – 3.08	-5.7
1	A	867	ARG	HG2	0.12	0.26 – 2.87	-5.5
1	A	866	VAL	CG2	13.57	13.71 – 28.88	-5.1

7.1.5 Random Coil Index (RCI) plots (i)

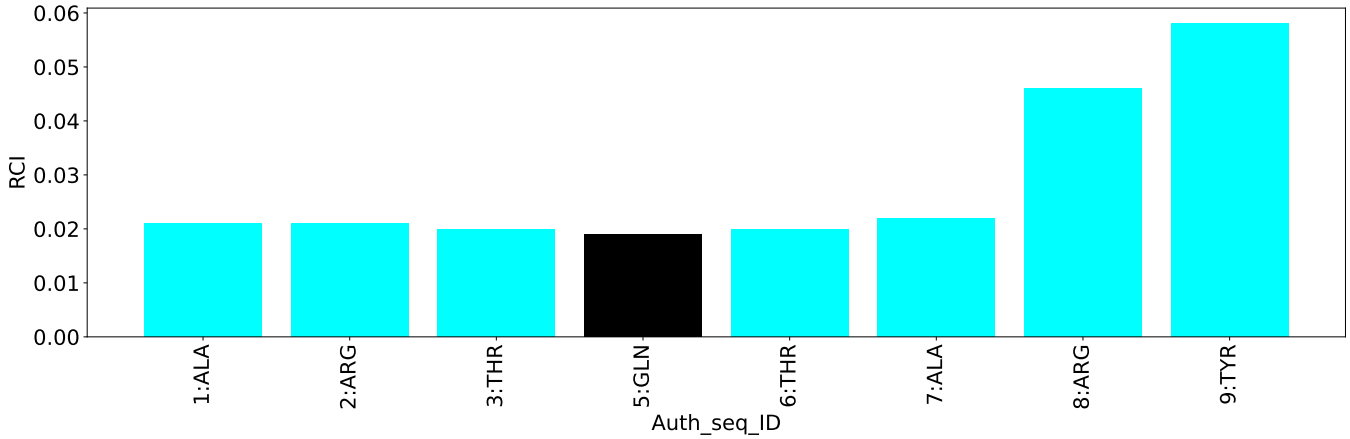
The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted 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:



Random coil index (RCI) for chain B:



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	947
Intra-residue (|i-j|=0)	233
Sequential (|i-j|=1)	302
Medium range (|i-j|>1 and |i-j|<5)	150
Long range (|i-j|≥5)	187
Inter-chain	71
Hydrogen bond restraints	0
Disulfide bond restraints	4
Total dihedral-angle restraints	0
Number of unmapped restraints	1
Number of restraints per residue	10.8
Number of long range restraints per residue[1]	2.2

[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 (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)	27.4	0.2
0.2-0.5 (Medium)	12.8	0.5
>0.5 (Large)	5.2	2.28

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)

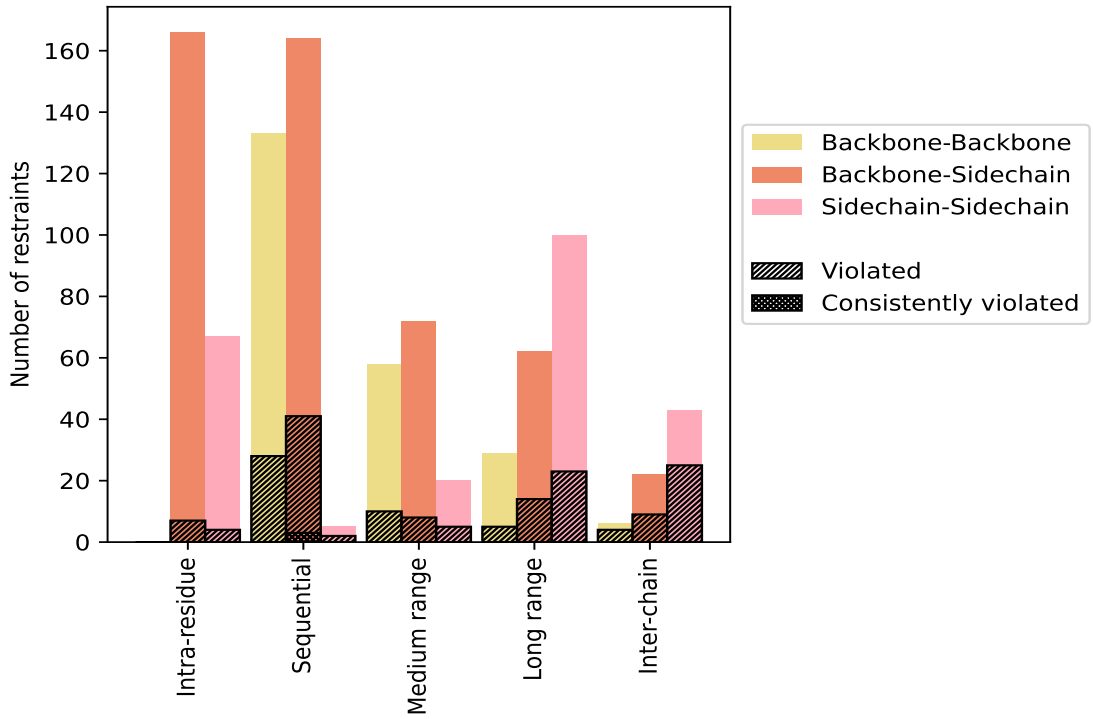
9.1 Summary of distance violations (i)

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)	233	24.6	11	4.7	1.2	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	166	17.5	7	4.2	0.7	0	0.0	0.0
Sidechain-Sidechain	67	7.1	4	6.0	0.4	0	0.0	0.0
Sequential (|i-j|=1)	302	31.9	71	23.5	7.5	3	1.0	0.3
Backbone-Backbone	133	14.0	28	21.1	3.0	0	0.0	0.0
Backbone-Sidechain	164	17.3	41	25.0	4.3	3	1.8	0.3
Sidechain-Sidechain	5	0.5	2	40.0	0.2	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	150	15.8	23	15.3	2.4	0	0.0	0.0
Backbone-Backbone	58	6.1	10	17.2	1.1	0	0.0	0.0
Backbone-Sidechain	72	7.6	8	11.1	0.8	0	0.0	0.0
Sidechain-Sidechain	20	2.1	5	25.0	0.5	0	0.0	0.0
Long range (|i-j|≥5)	187	19.7	39	20.9	4.1	0	0.0	0.0
Backbone-Backbone	29	3.1	5	17.2	0.5	0	0.0	0.0
Backbone-Sidechain	62	6.5	14	22.6	1.5	0	0.0	0.0
Sidechain-Sidechain	96	10.1	20	20.8	2.1	0	0.0	0.0
Inter-chain	71	7.5	38	53.5	4.0	0	0.0	0.0
Backbone-Backbone	6	0.6	4	66.7	0.4	0	0.0	0.0
Backbone-Sidechain	22	2.3	9	40.9	1.0	0	0.0	0.0
Sidechain-Sidechain	43	4.5	25	58.1	2.6	0	0.0	0.0
Hydrogen bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Disulfide bond	4	0.4	3	75.0	0.3	0	0.0	0.0
Total	947	100.0	185	19.5	19.5	3	0.3	0.3
Backbone-Backbone	226	23.9	47	20.8	5.0	0	0.0	0.0
Backbone-Sidechain	486	51.3	79	16.3	8.3	3	0.6	0.3
Sidechain-Sidechain	235	24.8	59	25.1	6.2	0	0.0	0.0

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

9.1.1 Bar chart : Distribution of distance restraints and violations (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	Number of violations						Mean (Å)	Max (Å)	SD[6] (Å)	Median (Å)
	IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total				
1	2	16	8	8	17	51	0.29	1.37	0.32	0.16
2	2	14	6	12	15	49	0.24	0.83	0.18	0.17
3	2	13	7	15	5	42	0.2	0.69	0.12	0.15
4	3	16	7	9	8	43	0.23	0.82	0.15	0.16
5	3	14	6	15	18	56	0.28	1.11	0.24	0.18
6	2	14	7	13	13	49	0.26	1.03	0.19	0.18
7	1	10	7	7	10	35	0.24	0.7	0.15	0.18
8	2	13	9	9	6	39	0.21	0.67	0.12	0.16
9	2	18	9	15	11	55	0.27	1.31	0.29	0.16
10	2	9	8	11	6	36	0.21	0.71	0.15	0.15
11	3	14	9	5	8	39	0.28	1.06	0.21	0.19

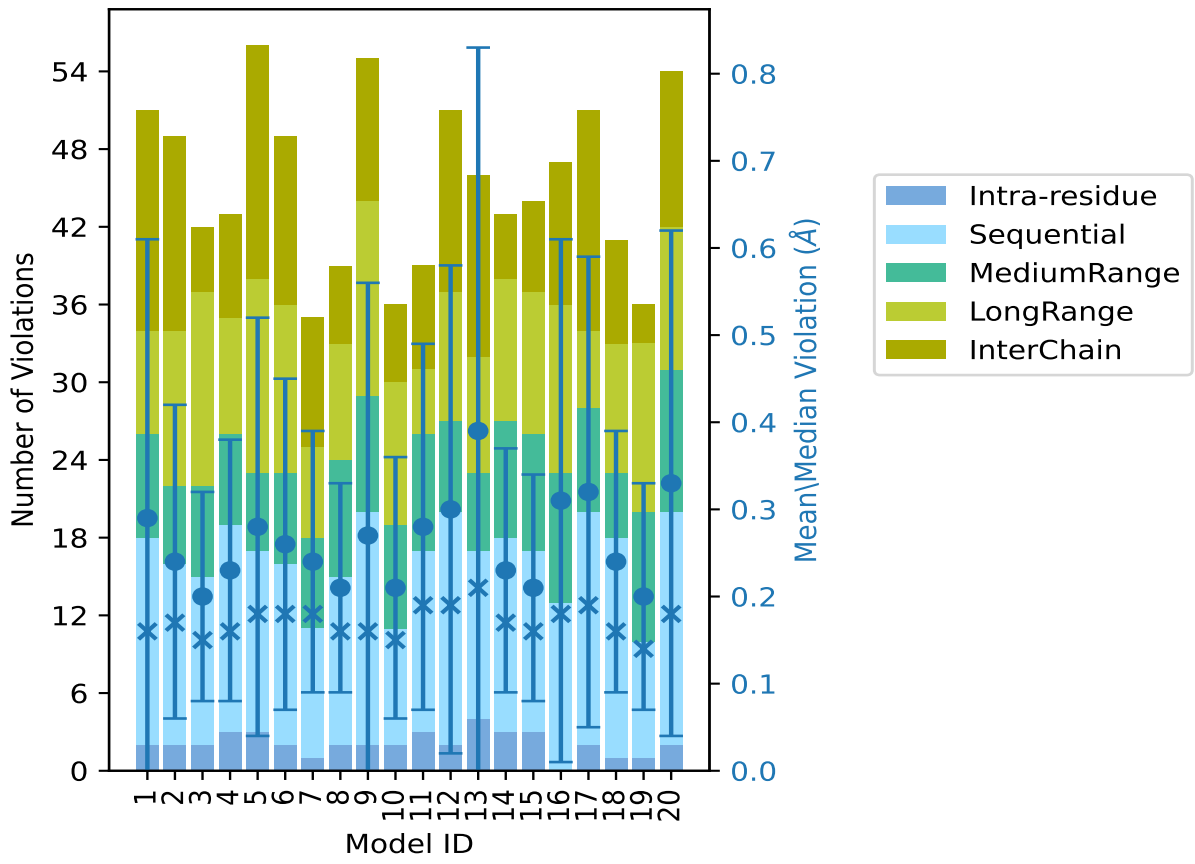
Continued on next page...

Continued from previous page...

Model ID	Number of violations						Mean (Å)	Max (Å)	SD[6] (Å)	Median (Å)
	IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total				
12	2	18	7	10	14	51	0.3	1.54	0.28	0.19
13	4	13	6	9	14	46	0.39	2.28	0.44	0.21
14	3	15	9	11	5	43	0.23	0.7	0.14	0.17
15	3	14	9	11	7	44	0.21	0.63	0.13	0.16
16	0	13	10	13	11	47	0.31	1.29	0.3	0.18
17	2	18	8	6	17	51	0.32	1.17	0.27	0.19
18	1	17	5	10	8	41	0.24	0.62	0.15	0.16
19	1	9	10	13	3	36	0.2	0.65	0.13	0.14
20	2	18	11	11	12	54	0.33	1.17	0.29	0.18

[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 (i)



The mean(dot),median(x) and the standard deviation are shown in blue with respect to the y axis on the right

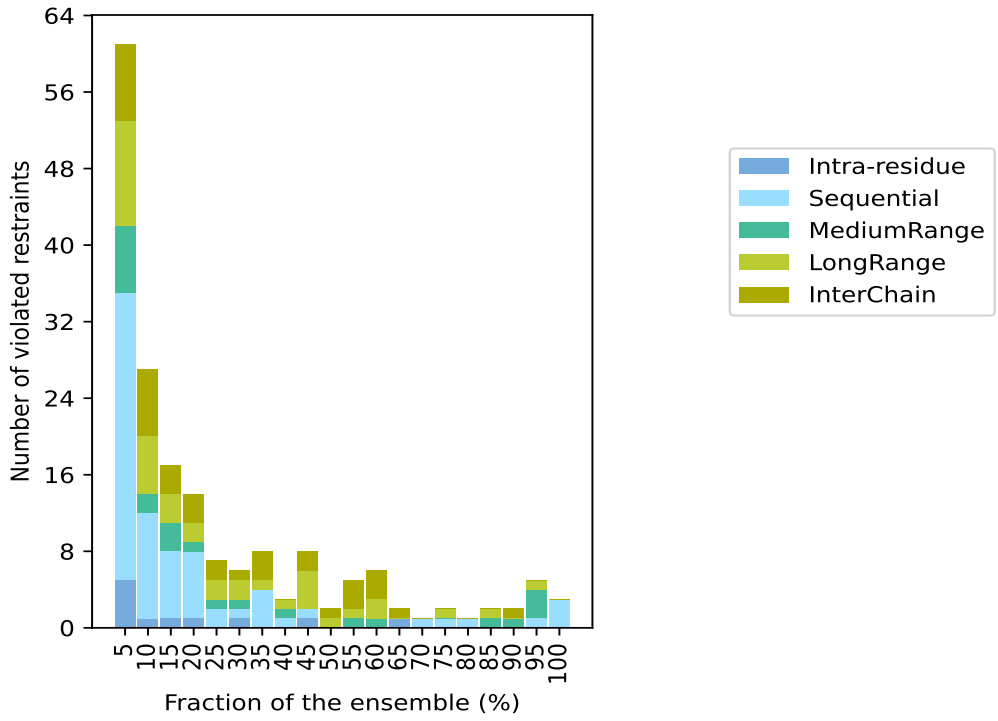
9.3 Distance violation statistics for the ensemble (i)

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, 761(IR:222, SQ:231, MR:127, LR:148, IC:33) 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]	%
5	30	7	11	8	61	1	5.0
1	11	2	6	7	27	2	10.0
1	7	3	3	3	17	3	15.0
1	7	1	2	3	14	4	20.0
0	2	1	2	2	7	5	25.0
1	1	1	2	1	6	6	30.0
0	4	0	1	3	8	7	35.0
0	1	1	1	0	3	8	40.0
1	1	0	4	2	8	9	45.0
0	0	0	1	1	2	10	50.0
0	0	1	1	3	5	11	55.0
0	0	1	2	3	6	12	60.0
1	0	0	0	1	2	13	65.0
0	1	0	0	0	1	14	70.0
0	1	0	1	0	2	15	75.0
0	1	0	0	0	1	16	80.0
0	0	1	1	0	2	17	85.0
0	0	1	0	1	2	18	90.0
0	1	3	1	0	5	19	95.0
0	3	0	0	0	3	20	100.0

[1]Intra-residue restraints, [2]Sequential restraints, [3]Medium range restraints, [4]Long range restraints, [5]Inter-chain restraints, [6] Number of models with violations

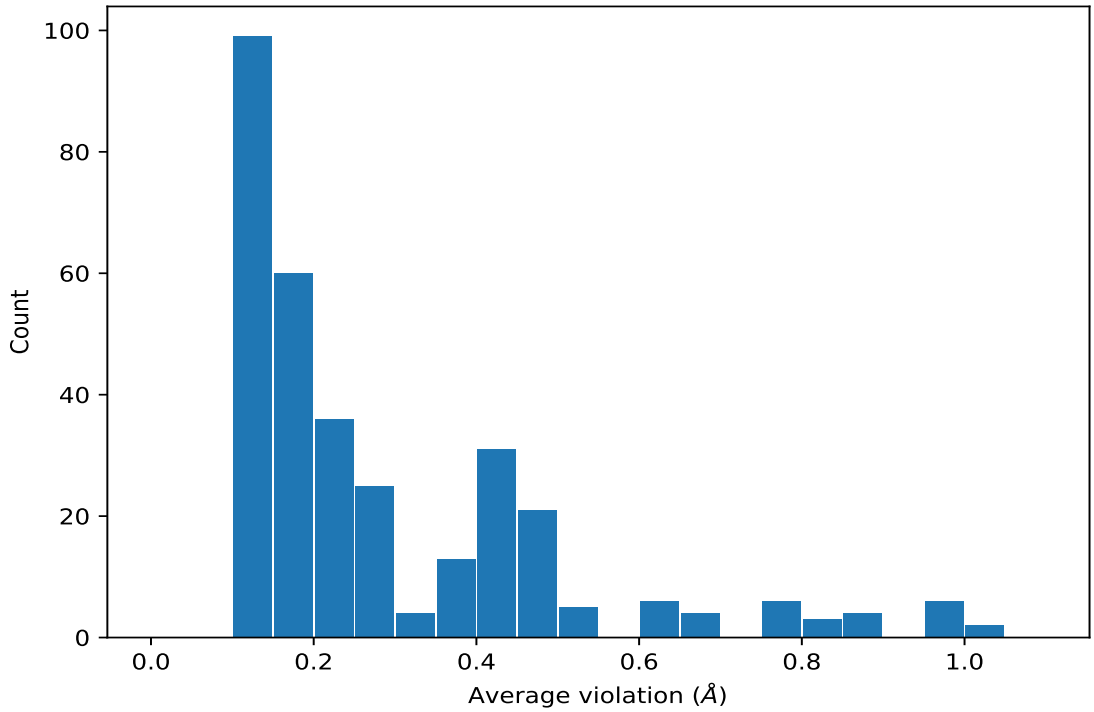
9.3.1 Bar graph : Distance violation statistics for the ensemble (i)



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)
(1,508)	1:A:892:ILE:HG12	1:A:893:CYS:H	20	0.65	0.02	0.66
(1,508)	1:A:892:ILE:HG13	1:A:893:CYS:H	20	0.65	0.02	0.66
(1,409)	1:A:882:VAL:HG11	1:A:883:GLY:HA2	20	0.3	0.08	0.31
(1,409)	1:A:882:VAL:HG11	1:A:883:GLY:HA3	20	0.3	0.08	0.31
(1,409)	1:A:882:VAL:HG12	1:A:883:GLY:HA2	20	0.3	0.08	0.31
(1,409)	1:A:882:VAL:HG12	1:A:883:GLY:HA3	20	0.3	0.08	0.31
(1,409)	1:A:882:VAL:HG13	1:A:883:GLY:HA2	20	0.3	0.08	0.31
(1,409)	1:A:882:VAL:HG13	1:A:883:GLY:HA3	20	0.3	0.08	0.31
(1,409)	1:A:882:VAL:HG21	1:A:883:GLY:HA2	20	0.3	0.08	0.31
(1,409)	1:A:882:VAL:HG21	1:A:883:GLY:HA3	20	0.3	0.08	0.31
(1,409)	1:A:882:VAL:HG22	1:A:883:GLY:HA2	20	0.3	0.08	0.31
(1,409)	1:A:882:VAL:HG22	1:A:883:GLY:HA3	20	0.3	0.08	0.31
(1,409)	1:A:882:VAL:HG23	1:A:883:GLY:HA2	20	0.3	0.08	0.31
(1,409)	1:A:882:VAL:HG23	1:A:883:GLY:HA3	20	0.3	0.08	0.31
(1,791)	1:A:915:ILE:HD11	1:A:916:ASN:H	20	0.2	0.03	0.2
(1,791)	1:A:915:ILE:HD12	1:A:916:ASN:H	20	0.2	0.03	0.2

Continued on next page...

Continued from previous page...

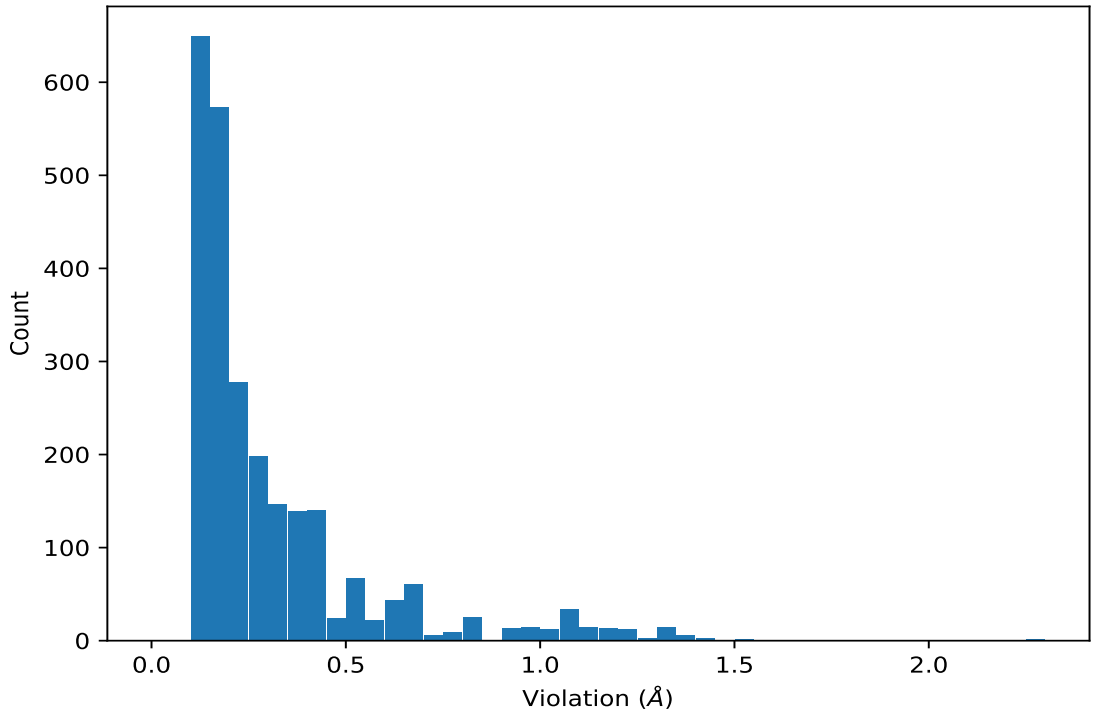
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,791)	1:A:915:ILE:HD13	1:A:916:ASN:H	20	0.2	0.03	0.2
(1,221)	1:A:868:CYS:H	1:A:874:TRP:HA	19	0.26	0.09	0.26
(1,520)	1:A:893:CYS:HA	1:A:895:ASN:H	19	0.24	0.06	0.24
(1,392)	1:A:881:VAL:HG11	1:A:884:SER:H	19	0.21	0.04	0.21
(1,392)	1:A:881:VAL:HG12	1:A:884:SER:H	19	0.21	0.04	0.21
(1,392)	1:A:881:VAL:HG13	1:A:884:SER:H	19	0.21	0.04	0.21
(1,392)	1:A:881:VAL:HG21	1:A:884:SER:H	19	0.21	0.04	0.21
(1,392)	1:A:881:VAL:HG22	1:A:884:SER:H	19	0.21	0.04	0.21
(1,392)	1:A:881:VAL:HG23	1:A:884:SER:H	19	0.21	0.04	0.21
(1,393)	1:A:881:VAL:HG11	1:A:885:ILE:H	19	0.18	0.05	0.17
(1,393)	1:A:881:VAL:HG12	1:A:885:ILE:H	19	0.18	0.05	0.17
(1,393)	1:A:881:VAL:HG13	1:A:885:ILE:H	19	0.18	0.05	0.17
(1,393)	1:A:881:VAL:HG21	1:A:885:ILE:H	19	0.18	0.05	0.17
(1,393)	1:A:881:VAL:HG22	1:A:885:ILE:H	19	0.18	0.05	0.17
(1,393)	1:A:881:VAL:HG23	1:A:885:ILE:H	19	0.18	0.05	0.17
(1,519)	1:A:893:CYS:HA	1:A:894:MET:H	19	0.12	0.01	0.12
(1,216)	1:A:867:ARG:HG2	2:B:4:MLZ:HCM1	18	0.4	0.22	0.36
(1,216)	1:A:867:ARG:HG2	2:B:4:MLZ:HCM2	18	0.4	0.22	0.36
(1,216)	1:A:867:ARG:HG2	2:B:4:MLZ:HCM3	18	0.4	0.22	0.36
(1,216)	1:A:867:ARG:HG3	2:B:4:MLZ:HCM1	18	0.4	0.22	0.36
(1,216)	1:A:867:ARG:HG3	2:B:4:MLZ:HCM2	18	0.4	0.22	0.36
(1,216)	1:A:867:ARG:HG3	2:B:4:MLZ:HCM3	18	0.4	0.22	0.36
(1,394)	1:A:881:VAL:HG11	1:A:885:ILE:HB	18	0.25	0.21	0.18

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

9.5 All violated distance restraints (i)

9.5.1 Histogram : Distribution of distance violations (i)

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



9.5.2 Table : All distance violations (i)

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

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,319)	1:A:874:TRP:HZ2	2:B:4:MLZ:HE2	13	2.28
(1,319)	1:A:874:TRP:HZ2	2:B:4:MLZ:HE3	13	2.28
(1,319)	1:A:874:TRP:HZ2	2:B:4:MLZ:HE2	12	1.54
(1,319)	1:A:874:TRP:HZ2	2:B:4:MLZ:HE3	12	1.54
(1,318)	1:A:874:TRP:HZ2	2:B:4:MLZ:HCM1	13	1.44
(1,318)	1:A:874:TRP:HZ2	2:B:4:MLZ:HCM2	13	1.44
(1,318)	1:A:874:TRP:HZ2	2:B:4:MLZ:HCM3	13	1.44
(1,795)	1:A:915:ILE:HD11	2:B:4:MLZ:HB2	1	1.37
(1,795)	1:A:915:ILE:HD11	2:B:4:MLZ:HB3	1	1.37
(1,795)	1:A:915:ILE:HD12	2:B:4:MLZ:HB2	1	1.37
(1,795)	1:A:915:ILE:HD12	2:B:4:MLZ:HB3	1	1.37
(1,795)	1:A:915:ILE:HD13	2:B:4:MLZ:HB2	1	1.37
(1,795)	1:A:915:ILE:HD13	2:B:4:MLZ:HB3	1	1.37
(1,795)	1:A:915:ILE:HD11	2:B:4:MLZ:HB2	13	1.32
(1,795)	1:A:915:ILE:HD11	2:B:4:MLZ:HB3	13	1.32
(1,795)	1:A:915:ILE:HD12	2:B:4:MLZ:HB2	13	1.32
(1,795)	1:A:915:ILE:HD12	2:B:4:MLZ:HB3	13	1.32

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,795)	1:A:915:ILE:HD13	2:B:4:MLZ:HB2	13	1.32
(1,795)	1:A:915:ILE:HD13	2:B:4:MLZ:HB3	13	1.32
(1,795)	1:A:915:ILE:HD11	2:B:4:MLZ:HB2	9	1.31
(1,795)	1:A:915:ILE:HD11	2:B:4:MLZ:HB3	9	1.31
(1,795)	1:A:915:ILE:HD12	2:B:4:MLZ:HB2	9	1.31
(1,795)	1:A:915:ILE:HD12	2:B:4:MLZ:HB3	9	1.31
(1,795)	1:A:915:ILE:HD13	2:B:4:MLZ:HB2	9	1.31
(1,795)	1:A:915:ILE:HD13	2:B:4:MLZ:HB3	9	1.31
(1,319)	1:A:874:TRP:HZ2	2:B:4:MLZ:HE2	1	1.3
(1,319)	1:A:874:TRP:HZ2	2:B:4:MLZ:HE3	1	1.3
(1,318)	1:A:874:TRP:HZ2	2:B:4:MLZ:HCM1	16	1.29
(1,318)	1:A:874:TRP:HZ2	2:B:4:MLZ:HCM2	16	1.29
(1,318)	1:A:874:TRP:HZ2	2:B:4:MLZ:HCM3	16	1.29
(1,801)	1:A:915:ILE:HG12	2:B:4:MLZ:HG2	1	1.22
(1,801)	1:A:915:ILE:HG12	2:B:4:MLZ:HG3	1	1.22
(1,801)	1:A:915:ILE:HG13	2:B:4:MLZ:HG2	1	1.22
(1,801)	1:A:915:ILE:HG13	2:B:4:MLZ:HG3	1	1.22
(1,801)	1:A:915:ILE:HG12	2:B:4:MLZ:HG2	16	1.22

10 Dihedral-angle violation analysis (i)

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