wwPDB NMR Structure Validation Summary Report

Jun 4, 2023 – 09:43 AM EDT

PDB ID	:	2KWV
BMRB ID	:	16885
Title	:	Solution Structure of UBM1 of murine Polymerase iota in Complex with Ubiquitin
Authors	:	Burschowsky, D.; Rudolf, F.; Rabut, G.; Herrmann, T.; Peter, M.; Wider, G.
Deposited on	:	2010-04-20

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:

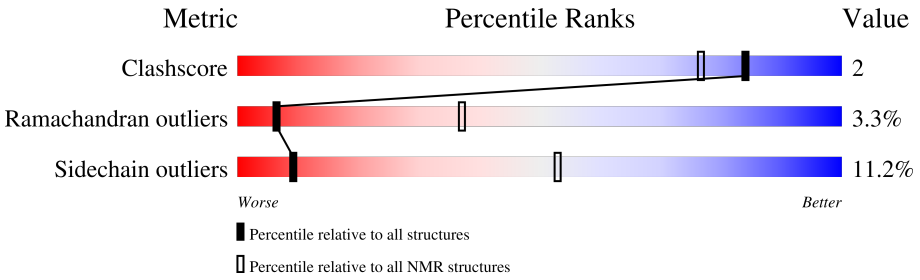
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal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

1 Overall quality at a glance

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

The overall completeness of chemical shifts assignment is 87%.

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	48	58% 15% . 25%
2	B	76	89% 8% .

2 Ensemble composition and analysis

This entry contains 20 models. Model 15 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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:491-A:526, B:1-B:74 (110)	0.80	15

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 1804 atoms, of which 919 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called DNA polymerase iota.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
1	A	36	573	179	290	47	57	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	485	GLY	-	expression tag	UNP Q6R3M4
A	486	SER	-	expression tag	UNP Q6R3M4
A	518	SER	TYR	variant	UNP Q6R3M4

- Molecule 2 is a protein called Ubiquitin.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	76	1231	378	629	105	118	1	0

4 Residue-property plots

4.1 Average score per residue in the NMR ensemble

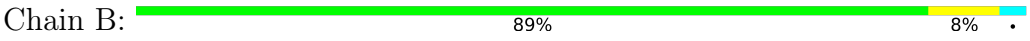
These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for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

Chain A:

• Molecule 2: Ubiquitin

Chain B:



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

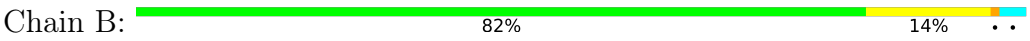
The representative model is number 15. Colouring as in section 4.1 above.

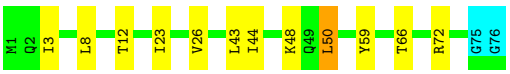
• Molecule 1: DNA polymerase iota

Chain A:

• Molecule 2: Ubiquitin

Chain B:





5 Refinement protocol and experimental data overview

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

Of the 200 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	9
MOLMOL	geometry optimization	2k.2
UNIO'08	structure solution	1.0.4
DYANA	structure solution	

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	1463
Number of shifts mapped to atoms	1365
Number of unparsed shifts	0
Number of shifts with mapping errors	98
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	87%

6 Model quality

6.1 Standard geometry

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	0.43±0.01	0±0/286 (0.0± 0.0%)	0.80±0.02	0±0/385 (0.0± 0.0%)
2	B	0.43±0.00	0±0/600 (0.0± 0.0%)	0.76±0.02	0±0/809 (0.0± 0.0%)
All	All	0.43	0/17720 (0.0%)	0.78	2/23880 (0.0%)

There are no bond-length outliers.

All unique angle outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
2	B	59	TYR	CB-CG-CD2	-5.37	117.78	121.00	8	2

There are no chirality outliers.

There are no planarity outliers.

6.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	283	290	289	2±1
2	B	593	623	623	2±2
All	All	17520	18260	18240	83

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:493:LEU:HD13	1:A:493:LEU:H	0.64	1.53	1	1
2:B:4:PHE:HB3	2:B:12:THR:HG21	0.63	1.69	1	3
1:A:513:GLN:O	1:A:517:LEU:HD22	0.59	1.97	12	2
1:A:491:LEU:N	1:A:491:LEU:HD22	0.59	2.12	7	1
2:B:23:ILE:HA	2:B:26:VAL:CG2	0.57	2.29	8	3

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	34/48 (71%)	25±1 (74±4%)	7±2 (20±5%)	2±1 (6±3%)	2 18
2	B	73/76 (96%)	63±1 (87±2%)	9±1 (12±2%)	1±1 (2±1%)	12 54
All	All	2140/2480 (86%)	1766 (83%)	304 (14%)	70 (3%)	6 37

5 of 8 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	493	LEU	17
2	B	66	THR	16
1	A	499	GLY	10
1	A	525	LEU	8
2	B	74	ARG	6

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	32/41 (78%)	25±1 (80±4%)	7±1 (20±4%)	3 33
2	B	68/68 (100%)	63±2 (93±2%)	5±2 (7±2%)	19 68

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2000/2180 (92%)	1775 (89%)	225 (11%)	9 53

5 of 29 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	491	LEU	20
1	A	505	PHE	20
2	B	48	LYS	18
1	A	502	GLN	14
1	A	525	LEU	14

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

The completeness of assignment taking into account all chemical shift lists is 87% for the well-defined parts and 87% 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

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	1463
Number of shifts mapped to atoms	1365
Number of unparsed shifts	0
Number of shifts with mapping errors	98
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	2

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 98) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	485	GLY	HA2	3.915	0.020	2
1	A	485	GLY	HA3	3.931	0.020	2
1	A	485	GLY	CA	43.403	0.300	1
1	A	486	SER	H	8.809	0.020	1
1	A	486	SER	HA	4.532	0.020	1
1	A	486	SER	HB2	3.906	0.020	2
1	A	486	SER	HB3	3.906	0.020	2
1	A	486	SER	CA	58.447	0.300	1
1	A	486	SER	CB	63.801	0.300	1
1	A	486	SER	N	113.768	0.300	1
1	A	487	ASP	H	8.686	0.020	1
1	A	487	ASP	HA	4.699	0.020	1
1	A	487	ASP	HB2	2.699	0.020	2
1	A	487	ASP	HB3	2.759	0.020	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	487	ASP	CA	54.455	0.300	1
1	A	487	ASP	CB	41.052	0.300	1
1	A	487	ASP	N	120.187	0.300	1
1	A	488	THR	H	8.204	0.020	1
1	A	488	THR	HA	4.391	0.020	1
1	A	488	THR	HB	4.331	0.020	1
1	A	488	THR	HG21	1.21	0.020	1
1	A	488	THR	HG22	1.21	0.020	1
1	A	488	THR	HG23	1.21	0.020	1
1	A	488	THR	CA	61.81	0.300	1
1	A	488	THR	CB	69.578	0.300	1
1	A	488	THR	CG2	21.576	0.300	1
1	A	488	THR	N	111.886	0.300	1
1	A	489	SER	H	8.401	0.020	1
1	A	489	SER	HA	4.434	0.020	1
1	A	489	SER	HB2	3.877	0.020	2
1	A	489	SER	HB3	3.905	0.020	2
1	A	489	SER	CA	58.784	0.300	1
1	A	489	SER	CB	63.78	0.300	1
1	A	489	SER	N	115.891	0.300	1
1	A	490	ASP	H	8.386	0.020	1
1	A	490	ASP	HA	4.637	0.020	1
1	A	490	ASP	HB2	2.585	0.020	2
1	A	490	ASP	HB3	2.697	0.020	2
1	A	490	ASP	CA	54.224	0.300	1
1	A	490	ASP	CB	40.977	0.300	1
1	A	490	ASP	N	120.145	0.300	1
1	A	527	GLY	H	8.383	0.020	1
1	A	527	GLY	HA2	3.968	0.020	2
1	A	527	GLY	HA3	3.968	0.020	2
1	A	527	GLY	CA	45.278	0.300	1
1	A	527	GLY	N	107.468	0.300	1
1	A	528	LYS	H	8.32	0.020	1
1	A	528	LYS	HA	4.333	0.020	1
1	A	528	LYS	HB2	1.793	0.020	2
1	A	528	LYS	HB3	1.891	0.020	2
1	A	528	LYS	HD2	1.699	0.020	2
1	A	528	LYS	HD3	1.699	0.020	2
1	A	528	LYS	HE2	3.007	0.020	2
1	A	528	LYS	HE3	3.007	0.020	2
1	A	528	LYS	HG2	1.43	0.020	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	528	LYS	HG3	1.48	0.020	2
1	A	528	LYS	CA	56.5	0.300	1
1	A	528	LYS	CB	32.944	0.300	1
1	A	528	LYS	CD	29.089	0.300	1
1	A	528	LYS	CE	42.192	0.300	1
1	A	528	LYS	CG	24.73	0.300	1
1	A	528	LYS	N	118.8	0.300	1
1	A	529	GLY	H	8.586	0.020	1
1	A	529	GLY	HA2	3.992	0.020	2
1	A	529	GLY	HA3	3.992	0.020	2
1	A	529	GLY	CA	45.296	0.300	1
1	A	529	GLY	N	108.073	0.300	1
1	A	530	SER	H	8.264	0.020	1
1	A	530	SER	HA	4.478	0.020	1
1	A	530	SER	HB2	3.889	0.020	2
1	A	530	SER	HB3	3.889	0.020	2
1	A	530	SER	CA	58.379	0.300	1
1	A	530	SER	CB	63.864	0.300	1
1	A	530	SER	N	113.546	0.300	1
1	A	531	LEU	H	8.456	0.020	1
1	A	531	LEU	HA	4.448	0.020	1
1	A	531	LEU	HB2	1.678	0.020	2
1	A	531	LEU	HB3	1.678	0.020	2
1	A	531	LEU	HD11	0.882	0.020	2
1	A	531	LEU	HD12	0.882	0.020	2
1	A	531	LEU	HD13	0.882	0.020	2
1	A	531	LEU	HD21	0.938	0.020	2
1	A	531	LEU	HD22	0.938	0.020	2
1	A	531	LEU	HD23	0.938	0.020	2
1	A	531	LEU	HG	1.671	0.020	1
1	A	531	LEU	CA	55.255	0.300	1
1	A	531	LEU	CB	42.29	0.300	1
1	A	531	LEU	CD1	23.261	0.300	2
1	A	531	LEU	CD2	25.146	0.300	2
1	A	531	LEU	CG	26.921	0.300	1
1	A	531	LEU	N	122.135	0.300	1
1	A	532	SER	H	7.932	0.020	1
1	A	532	SER	HA	4.255	0.020	1
1	A	532	SER	HB2	3.844	0.020	2
1	A	532	SER	HB3	3.844	0.020	2
1	A	532	SER	CA	59.938	0.300	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	532	SER	CB	64.842	0.300	1
1	A	532	SER	N	119.382	0.300	1

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	124	-0.09 ± 0.10	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	113	0.16 ± 0.10	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	114	2.65 ± 0.57	Should be applied

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 87%, i.e. 1355 atoms were assigned a chemical shift out of a possible 1560. 0 out of 21 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	429/544 (79%)	218/220 (99%)	110/220 (50%)	101/104 (97%)
Sidechain	892/970 (92%)	605/626 (97%)	270/306 (88%)	17/38 (45%)
Aromatic	34/46 (74%)	21/23 (91%)	13/22 (59%)	0/1 (0%)
Overall	1355/1560 (87%)	844/869 (97%)	393/548 (72%)	118/143 (83%)

7.1.4 Statistically unusual chemical shifts

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.

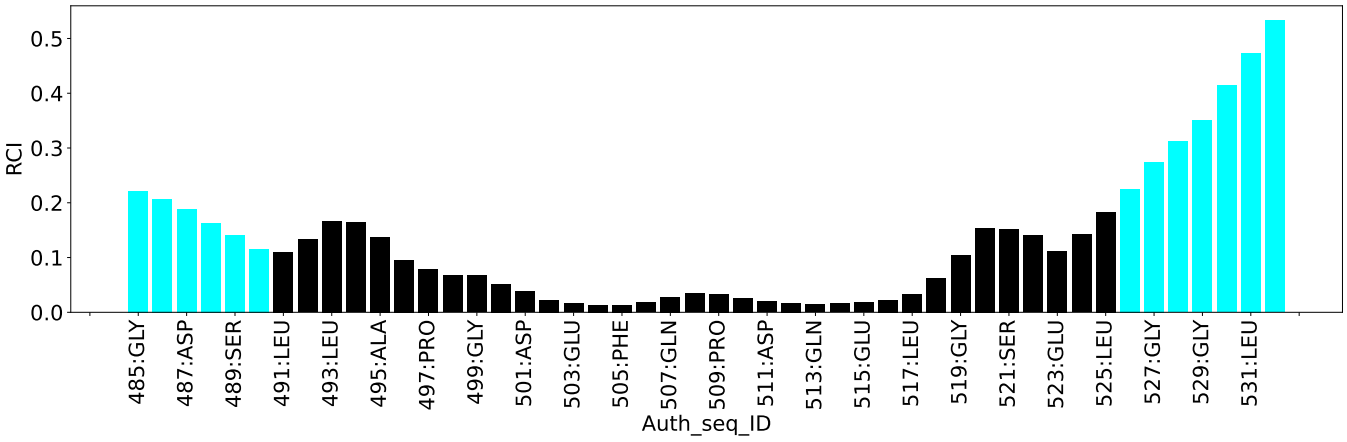
List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	513	GLN	HG2	0.70	1.01 – 3.62	-6.2
1	B	41	GLN	NE2	102.05	103.38 – 120.35	-5.8

7.1.5 Random Coil Index (RCI) plots

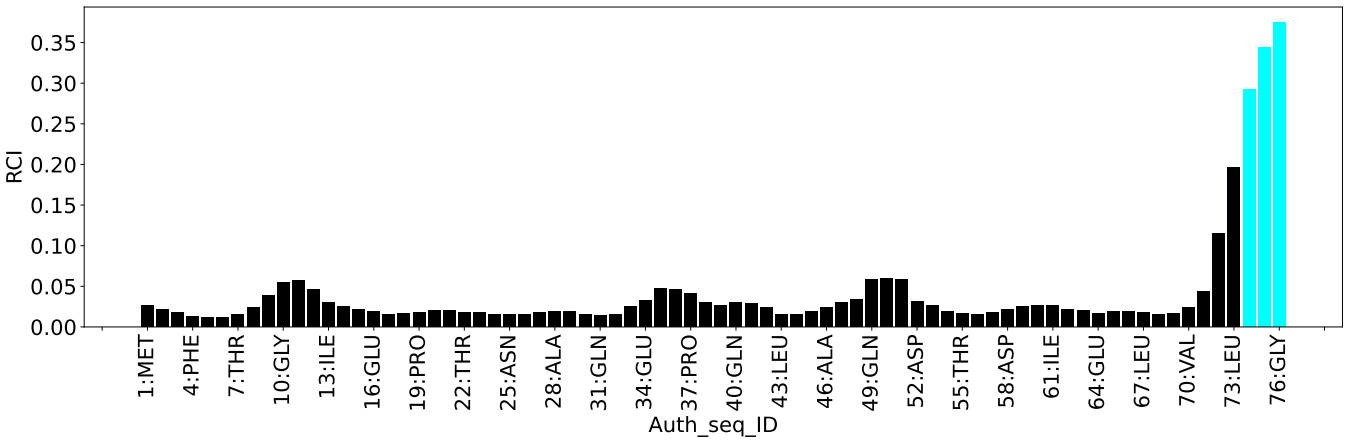
The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from

the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:



Random coil index (RCI) for chain B:



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	2764
Intra-residue (|i-j|=0)	626
Sequential (|i-j|=1)	784
Medium range (|i-j|>1 and |i-j|<5)	605
Long range (|i-j|≥5)	687
Inter-chain	62
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	118
Number of restraints per residue	22.3
Number of long range restraints per residue[1]	5.5

[1]Long range hydrogen bonds and disulfide bonds are counted as long range restraints while calculating the number of long range restraints per residue

8.2 Residual restraint violations

This section provides the overview of the restraint violations analysis. The violations are binned as small, medium and large violations based on its absolute value. Average number of violations per model is calculated by dividing the total number of violations in each bin by the size of the ensemble.

8.2.1 Average number of distance violations per model

Distance violations less than 0.1 Å are not included in the calculation.

Bins (Å)	Average number of violations per model	Max (Å)
0.1-0.2 (Small)	65.3	0.2
0.2-0.5 (Medium)	52.6	0.5
>0.5 (Large)	129.2	4.13

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)

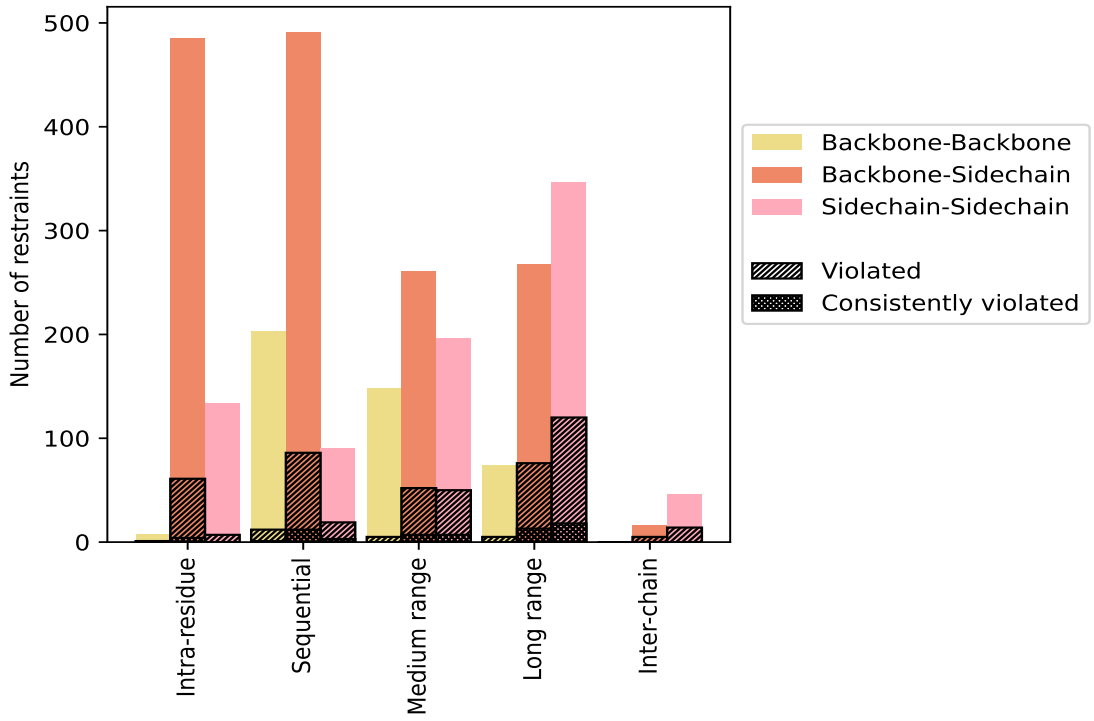
9.1 Summary of distance violations (i)

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)	626	22.6	69	11.0	2.5	4	0.6	0.1
Backbone-Backbone	7	0.3	1	14.3	0.0	0	0.0	0.0
Backbone-Sidechain	485	17.5	61	12.6	2.2	4	0.8	0.1
Sidechain-Sidechain	134	4.8	7	5.2	0.3	0	0.0	0.0
Sequential (|i-j|=1)	784	28.4	117	14.9	4.2	16	2.0	0.6
Backbone-Backbone	203	7.3	12	5.9	0.4	1	0.5	0.0
Backbone-Sidechain	491	17.8	86	17.5	3.1	12	2.4	0.4
Sidechain-Sidechain	90	3.3	19	21.1	0.7	3	3.3	0.1
Medium range (|i-j|>1 & |i-j|<5)	605	21.9	107	17.7	3.9	14	2.3	0.5
Backbone-Backbone	148	5.4	5	3.4	0.2	0	0.0	0.0
Backbone-Sidechain	261	9.4	52	19.9	1.9	7	2.7	0.3
Sidechain-Sidechain	196	7.1	50	25.5	1.8	7	3.6	0.3
Long range (|i-j|≥5)	687	24.9	201	29.3	7.3	31	4.5	1.1
Backbone-Backbone	74	2.7	5	6.8	0.2	0	0.0	0.0
Backbone-Sidechain	267	9.7	76	28.5	2.7	13	4.9	0.5
Sidechain-Sidechain	346	12.5	120	34.7	4.3	18	5.2	0.7
Inter-chain	62	2.2	19	30.6	0.7	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	16	0.6	5	31.2	0.2	0	0.0	0.0
Sidechain-Sidechain	46	1.7	14	30.4	0.5	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	2764	100.0	513	18.6	18.6	65	2.4	2.4
Backbone-Backbone	432	15.6	23	5.3	0.8	1	0.2	0.0
Backbone-Sidechain	1520	55.0	280	18.4	10.1	36	2.4	1.3
Sidechain-Sidechain	812	29.4	210	25.9	7.6	28	3.4	1.0

[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



Violated and consistently violated restraints are shown using different hatch patterns in their respective categories. The hydrogen bonds and disulfied bonds are counted in their appropriate category on the x-axis

9.2 Distance violation statistics for each model

The following table provides the distance violation statistics for each model in the ensemble. Violations less than 0.1 Å are not included in the statistics.

Model ID	Number of violations						Mean (Å)	Max (Å)	SD[6] (Å)	Median (Å)
	IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total				
1	27	54	50	111	5	247	0.81	4.11	0.77	0.56
2	26	56	52	112	4	250	0.8	4.06	0.75	0.54
3	28	57	52	97	3	237	0.67	2.68	0.6	0.45
4	26	61	59	101	6	253	0.77	4.09	0.75	0.46
5	25	64	50	108	6	253	0.82	4.1	0.75	0.58
6	24	52	45	106	11	238	0.83	4.0	0.77	0.55
7	27	47	51	100	9	234	0.72	2.7	0.6	0.54
8	32	56	51	89	5	233	0.68	2.45	0.56	0.48
9	29	62	57	110	6	264	0.82	4.11	0.75	0.62
10	27	55	54	78	10	224	0.69	2.54	0.57	0.48
11	31	54	55	108	8	256	0.74	2.61	0.61	0.57

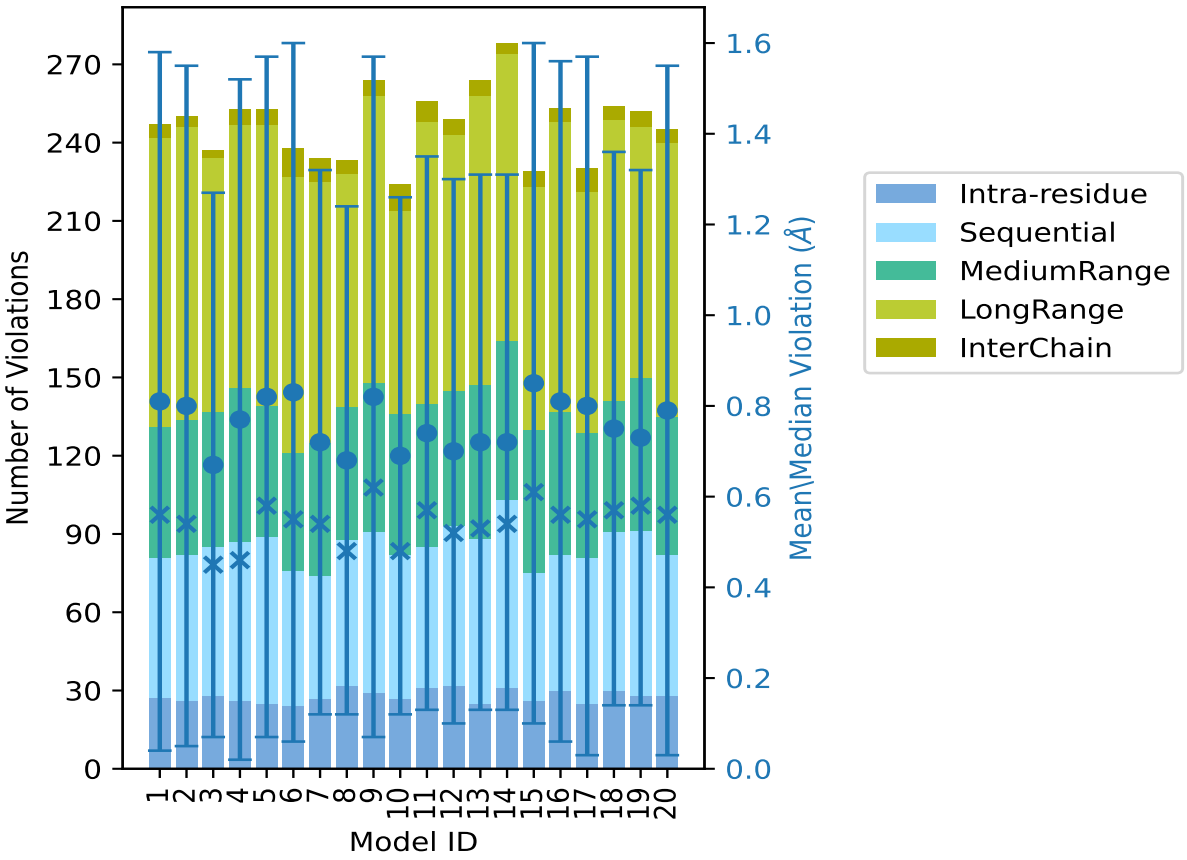
Continued on next page...

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	32	61	52	98	6	249	0.7	2.58	0.6	0.52
13	25	63	59	111	6	264	0.72	2.72	0.59	0.53
14	31	72	61	110	4	278	0.72	2.56	0.59	0.54
15	26	49	55	93	6	229	0.85	4.04	0.75	0.61
16	30	52	55	111	5	253	0.81	4.11	0.75	0.56
17	25	56	48	92	9	230	0.8	4.09	0.77	0.55
18	30	61	50	108	5	254	0.75	2.58	0.61	0.57
19	28	63	59	96	6	252	0.73	2.68	0.59	0.58
20	28	54	53	105	5	245	0.79	4.13	0.76	0.56

[1]Intra-residue restraints, [2]Sequential restraints, [3]Medium range restraints, [4]Long range restraints, [5]Inter-chain restraints, [6]Standard deviation

9.2.1 Bar graph : Distance Violation statistics for each model



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

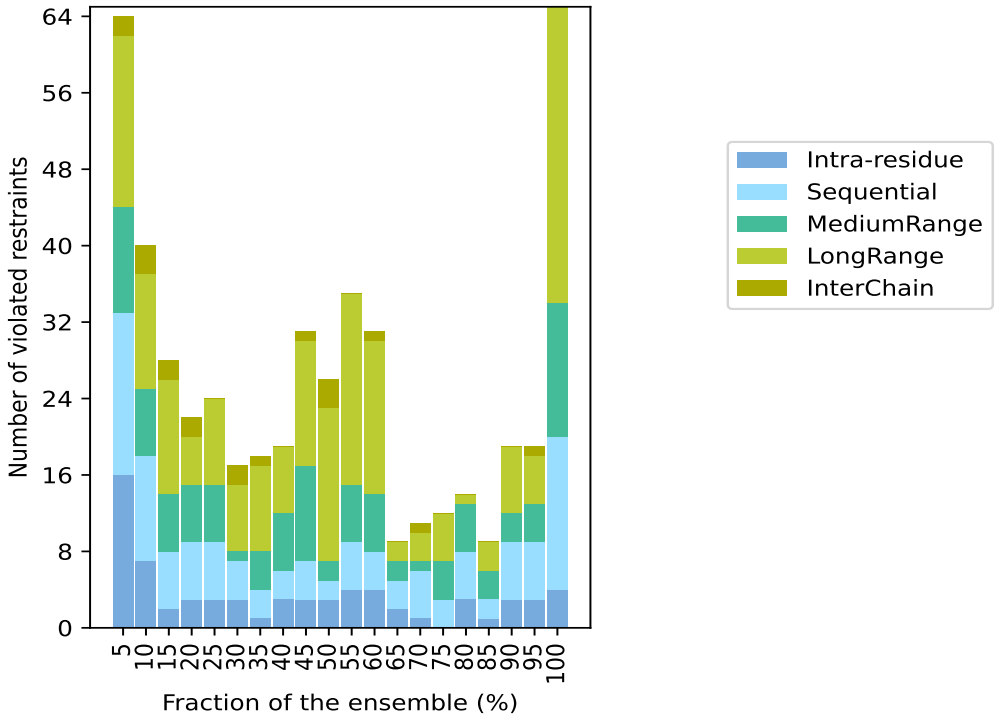
9.3 Distance violation statistics for the ensemble (i)

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, 2251(IR:557, SQ:667, MR:498, LR:486, IC:43) 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]	%
16	17	11	18	2	64	1	5.0
7	11	7	12	3	40	2	10.0
2	6	6	12	2	28	3	15.0
3	6	6	5	2	22	4	20.0
3	6	6	9	0	24	5	25.0
3	4	1	7	2	17	6	30.0
1	3	4	9	1	18	7	35.0
3	3	6	7	0	19	8	40.0
3	4	10	13	1	31	9	45.0
3	2	2	16	3	26	10	50.0
4	5	6	20	0	35	11	55.0
4	4	6	16	1	31	12	60.0
2	3	2	2	0	9	13	65.0
1	5	1	3	1	11	14	70.0
0	3	4	5	0	12	15	75.0
3	5	5	1	0	14	16	80.0
1	2	3	3	0	9	17	85.0
3	6	3	7	0	19	18	90.0
3	6	4	5	1	19	19	95.0
4	16	14	31	0	65	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

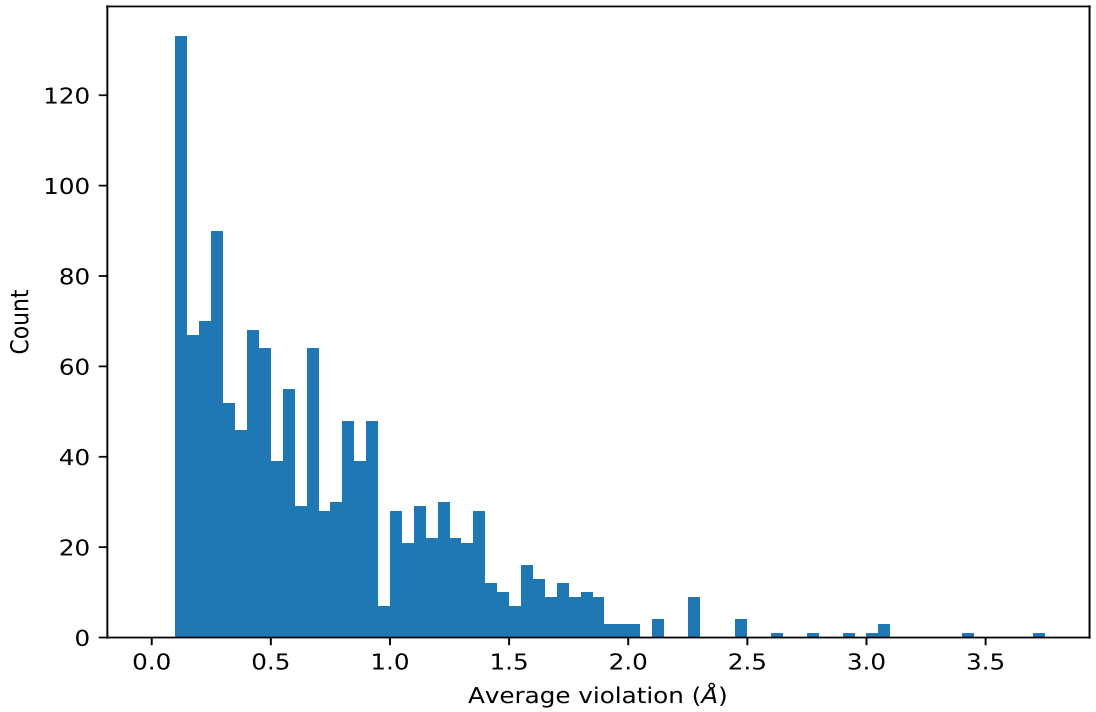
9.3.1 Bar graph : Distance violation statistics for the ensemble



9.4 Most violated distance restraints in the ensemble

9.4.1 Histogram : Distribution of mean distance violations

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



9.4.2 Table: Most violated distance restraints

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

Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,787)	1:A:500:VAL:HG21	1:A:506:LYS:H	20	2.49	0.07	2.49
(1,787)	1:A:500:VAL:HG22	1:A:506:LYS:H	20	2.49	0.07	2.49
(1,787)	1:A:500:VAL:HG23	1:A:506:LYS:H	20	2.49	0.07	2.49
(1,411)	1:A:500:VAL:HG21	1:A:505:PHE:H	20	2.14	0.1	2.13
(1,411)	1:A:500:VAL:HG22	1:A:505:PHE:H	20	2.14	0.1	2.13
(1,411)	1:A:500:VAL:HG23	1:A:505:PHE:H	20	2.14	0.1	2.13
(1,2076)	2:B:1:MET:HG3	2:B:17:VAL:HG11	20	2.03	0.62	2.46
(1,2076)	2:B:1:MET:HG3	2:B:17:VAL:HG12	20	2.03	0.62	2.46
(1,2076)	2:B:1:MET:HG3	2:B:17:VAL:HG13	20	2.03	0.62	2.46
(1,1064)	2:B:17:VAL:HG21	2:B:29:LYS:HE2	20	1.88	0.04	1.88
(1,1064)	2:B:17:VAL:HG22	2:B:29:LYS:HE2	20	1.88	0.04	1.88
(1,1064)	2:B:17:VAL:HG23	2:B:29:LYS:HE2	20	1.88	0.04	1.88
(1,1317)	2:B:17:VAL:HG21	2:B:29:LYS:HB2	20	1.86	0.22	1.8
(1,1317)	2:B:17:VAL:HG22	2:B:29:LYS:HB2	20	1.86	0.22	1.8
(1,1317)	2:B:17:VAL:HG23	2:B:29:LYS:HB2	20	1.86	0.22	1.8
(1,230)	1:A:500:VAL:HG21	1:A:502:GLN:HA	20	1.81	0.24	1.74

Continued on next page...

Continued from previous page...

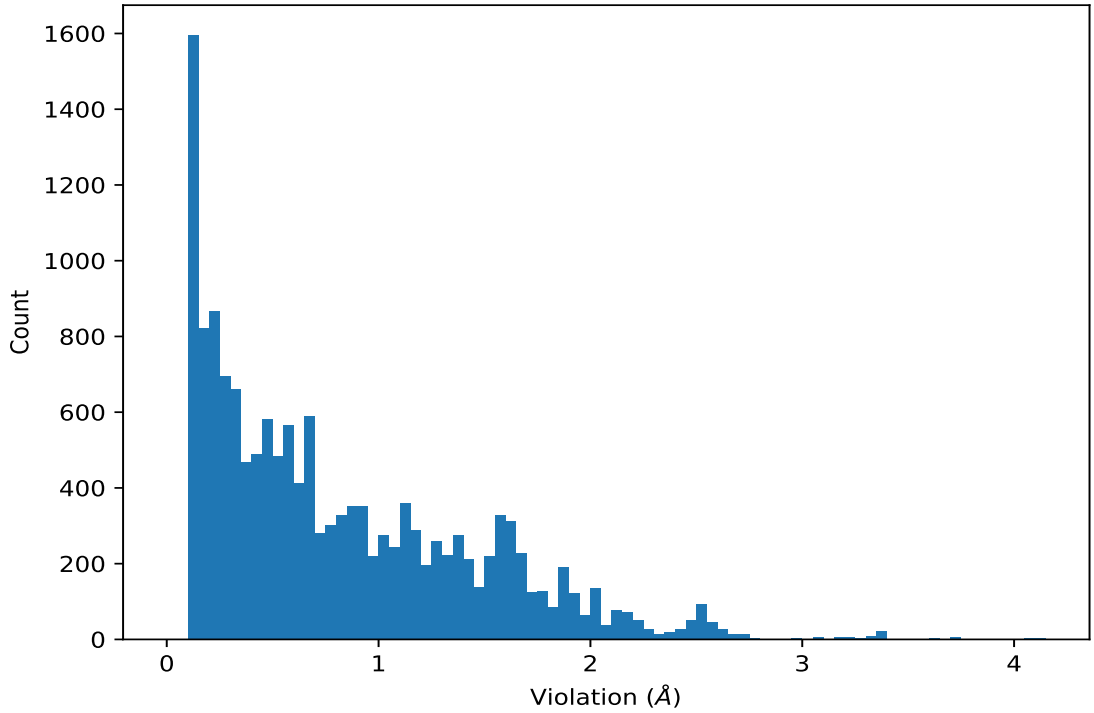
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,230)	1:A:500:VAL:HG22	1:A:502:GLN:HA	20	1.81	0.24	1.74
(1,230)	1:A:500:VAL:HG23	1:A:502:GLN:HA	20	1.81	0.24	1.74
(1,192)	1:A:494:GLN:HB3	1:A:517:LEU:HD11	20	1.76	0.48	1.93
(1,192)	1:A:494:GLN:HB3	1:A:517:LEU:HD12	20	1.76	0.48	1.93
(1,192)	1:A:494:GLN:HB3	1:A:517:LEU:HD13	20	1.76	0.48	1.93
(1,1325)	2:B:68:HIS:HB3	2:B:70:VAL:HG11	20	1.66	0.4	1.69
(1,1325)	2:B:68:HIS:HB3	2:B:70:VAL:HG12	20	1.66	0.4	1.69
(1,1325)	2:B:68:HIS:HB3	2:B:70:VAL:HG13	20	1.66	0.4	1.69
(1,1363)	2:B:17:VAL:HG11	2:B:19:PRO:HA	20	1.62	0.14	1.6
(1,1363)	2:B:17:VAL:HG12	2:B:19:PRO:HA	20	1.62	0.14	1.6
(1,1363)	2:B:17:VAL:HG13	2:B:19:PRO:HA	20	1.62	0.14	1.6
(1,2151)	2:B:1:MET:HE1	2:B:17:VAL:HG11	20	1.55	0.11	1.58

[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)
(1,1536)	2:B:23:ILE:HG12	2:B:59:TYR:HE1	20	4.13
(1,1536)	2:B:23:ILE:HG12	2:B:59:TYR:HE1	1	4.11
(1,1536)	2:B:23:ILE:HG12	2:B:59:TYR:HE1	9	4.11
(1,1536)	2:B:23:ILE:HG12	2:B:59:TYR:HE1	16	4.11
(1,1536)	2:B:23:ILE:HG12	2:B:59:TYR:HE1	5	4.1
(1,1536)	2:B:23:ILE:HG12	2:B:59:TYR:HE1	4	4.09
(1,1536)	2:B:23:ILE:HG12	2:B:59:TYR:HE1	17	4.09
(1,1536)	2:B:23:ILE:HG12	2:B:59:TYR:HE1	2	4.06
(1,1536)	2:B:23:ILE:HG12	2:B:59:TYR:HE1	15	4.04
(1,1536)	2:B:23:ILE:HG12	2:B:59:TYR:HE1	6	4.0

10 Dihedral-angle violation analysis (i)

No dihedral-angle restraints found