wwPDB NMR Structure Validation Summary Report (i)

Jun 5, 2023 – 02:18 PM EDT

PDB ID	:	2MC1
BMRB ID	:	19423
Title	:	Solution structure of the Vav1 SH2 domain complexed with a Syk-derived singly phosphorylated peptide
Authors	:	Chen, C.; Piraner, D.; Gorenstein, N.M.; Geahlen, R.L.; Post, C.B.
Deposited on	:	2013-08-13

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:

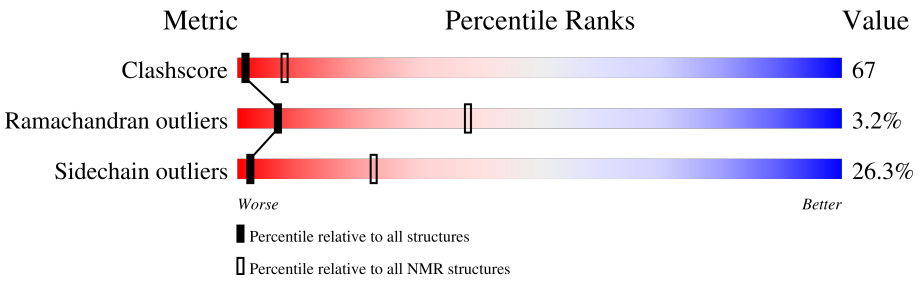
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (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 73%.

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	107	14% 55% 19% 12%
2	B	13	100%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 9 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:666-A:696, A:705-A:767 (94)	0.39	9

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 4 clusters and 4 single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 1937 atoms, of which 958 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Proto-oncogene vav.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	107	1741	555	873	149	160	4	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	661	GLY	-	expression tag	UNP P15498
A	662	HIS	-	expression tag	UNP P15498
A	663	MET	-	expression tag	UNP P15498

- Molecule 2 is a protein called Tyrosine-protein kinase SYK.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	P	
2	B	13	196	66	85	13	31	1	0

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: Proto-oncogene vav

Chain A: 14% 55% 19% 12%

G661 H662 M663 Q664 D665 H669 L670 W671 Y672 A673 G674 P675 M676 E677 R678 A679 E682 S683 I684 L685 A686 N687 R688 S689 D690 G691 T692 F693 L694 V695 R696 Q697 R698 V699 K700 D701 A702 A703 E704 F705 A706 I707 S708 I709 K710 Y711 N712 V713 E714 V715 K716 H717 I718 K719 I720 M721 T722 A723

E724 G725 L726 Y727 R728 I729 T730 E731 K732 K733 A734 F735 R736 G737 L738 T739 E740 L741 V742 E743 F744 Y745 Q746 Q747 N748 S749 L750 K751 D752 C753 F754 K755 S756 L757 D758 T759 T760 L761 Q762 F763 P764 F765 K766 E767

• Molecule 2: Tyrosine-protein kinase SYK

Chain B: 100%

D338 T339 E340 V341 Y342 E343 S344 P345 Y346 A347 D348 P349 E350

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

The representative model is number 9. Colouring as in section 4.1 above.

• Molecule 1: Proto-oncogene vav

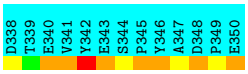
Chain A: 15% 54% 19% 12%

G661 H662 M663 Q664 D665 H669 L670 W671 Y672 A673 G674 P675 M676 E677 R678 A679 G680 A681 E682 S683 I684 N687 R688 S689 D690 G691 T692 F693 L694 V695 R696 Q697 R698 V699 K700 D701 A702 A703 E704 F705 A706 I707 S708 I709 K710 Y711 N712 V713 E714 V715 K716 H717 I718 K719 I720 M721 T722 A723

E724 G725 L726 Y727 R728 I729 T730 E731 K732 K733 A734 F735 R736 G737 L738 T739 E740 L741 V742 E743 F744 Y745 Q746 Q747 N748 S749 L750 K751 D752 C753 F754 K755 S756 L757 D758 T759 T760 L761 Q762 F763 P764 F765 K766 E767

• Molecule 2: Tyrosine-protein kinase SYK

Chain B: 100%


D338
T339
E340
V341
Y342
E343
S344
P345
Y346
A347
D348
P349
E350

5 Refinement protocol and experimental data overview (i)

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

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy*.

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

Software name	Classification	Version
X-PLOR NIH	structure solution	
X-PLOR NIH	refinement	
CYANA	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	2
Total number of shifts	1088
Number of shifts mapped to atoms	1088
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	73%

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: PTR

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
1	A	766	776	774	103±7
2	B	0	0	0	0±0
All	All	15320	15520	15480	2065

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 67.

5 of 394 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:745:TYR:CZ	1:A:753:CYS:SG	1.03	2.51	11	4
1:A:745:TYR:CE2	1:A:753:CYS:SG	0.91	2.64	17	4
1:A:729:ILE:HD11	1:A:741:LEU:HD21	0.80	1.52	7	18
1:A:728:ARG:NE	1:A:733:LYS:O	0.77	2.17	18	4
1:A:693:PHE:CD2	1:A:761:LEU:HD12	0.76	2.14	13	20

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	93/107 (87%)	76±2 (82±2%)	14±2 (15±2%)	3±1 (3±1%)	7 38
2	B	0	-	-	-	-
All	All	1860/2400 (78%)	1524 (82%)	276 (15%)	60 (3%)	7 38

5 of 12 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	713	VAL	20
1	A	723	ALA	15
1	A	757	LEU	7
1	A	674	GLY	6
1	A	688	ARG	3

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	82/92 (89%)	60±2 (74±3%)	22±2 (26±3%)	2 22
2	B	0	-	-	-
All	All	1640/2060 (80%)	1209 (74%)	431 (26%)	2 22

5 of 41 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	729	ILE	20
1	A	757	LEU	20
1	A	763	PHE	20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	766	LYS	20
1	A	687	ASN	18

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	PTR	B	346	2	15,16,17	1.12±0.04	1±0 (6±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	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
2	PTR	B	346	2	19,22,24	0.85±0.04	0±0 (0±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	Torsions	Rings
2	PTR	B	346	2	-	0±0,10,11,13	0±0,1,1,1

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
2	B	346	PTR	P-OH	3.71	1.65	1.59	9	20

All unique angle outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
2	B	346	PTR	CG-CB-CA	2.14	109.77	114.10	1	1

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)

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 73% for the well-defined parts and 65% 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	1015
Number of shifts mapped to atoms	1015
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	4

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	95	-0.02 ± 0.15	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	92	-0.04 ± 0.12	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	92	0.67 ± 0.31	Should be applied

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 73%, i.e. 975 atoms were assigned a chemical shift out of a possible 1341. 0 out of 15 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	365/471 (77%)	185/191 (97%)	91/188 (48%)	89/92 (97%)
Sidechain	588/736 (80%)	388/479 (81%)	200/228 (88%)	0/29 (0%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	22/134 (16%)	11/65 (17%)	11/64 (17%)	0/5 (0%)
Overall	975/1341 (73%)	584/735 (79%)	302/480 (63%)	89/126 (71%)

7.1.4 Statistically unusual chemical shifts (i)

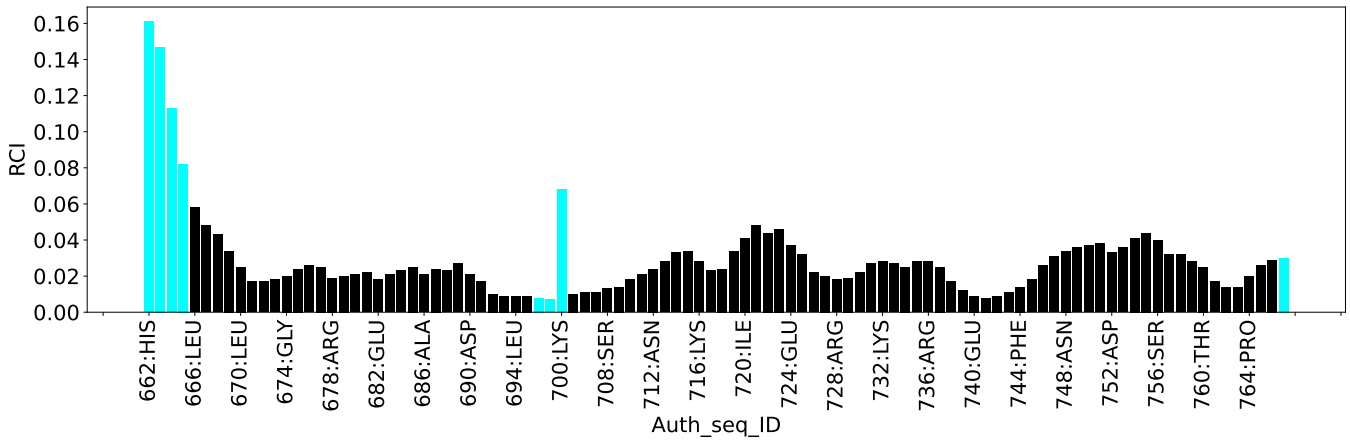
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	669	HIS	HB3	0.78	1.18 – 4.91	-6.1
1	A	766	LYS	HG3	-0.13	0.04 – 2.67	-5.7
1	A	741	LEU	HA	2.01	2.04 – 6.55	-5.1
1	A	674	GLY	H	5.21	5.23 – 11.42	-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:



7.2 Chemical shift list 2

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list_2*

7.2.1 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	73
Number of shifts mapped to atoms	73
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

7.2.2 Chemical shift referencing (i)

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

7.2.3 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 0%, i.e. 0 atoms were assigned a chemical shift out of a possible 1341. 0 out of 15 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}**H**	^{13}**C**	^{15}**N**
Backbone	0/471 (0%)	0/191 (0%)	0/188 (0%)	0/92 (0%)
Sidechain	0/736 (0%)	0/479 (0%)	0/228 (0%)	0/29 (0%)
Aromatic	0/134 (0%)	0/65 (0%)	0/64 (0%)	0/5 (0%)
Overall	0/1341 (0%)	0/735 (0%)	0/480 (0%)	0/126 (0%)

7.2.4 Statistically unusual chemical shifts (i)

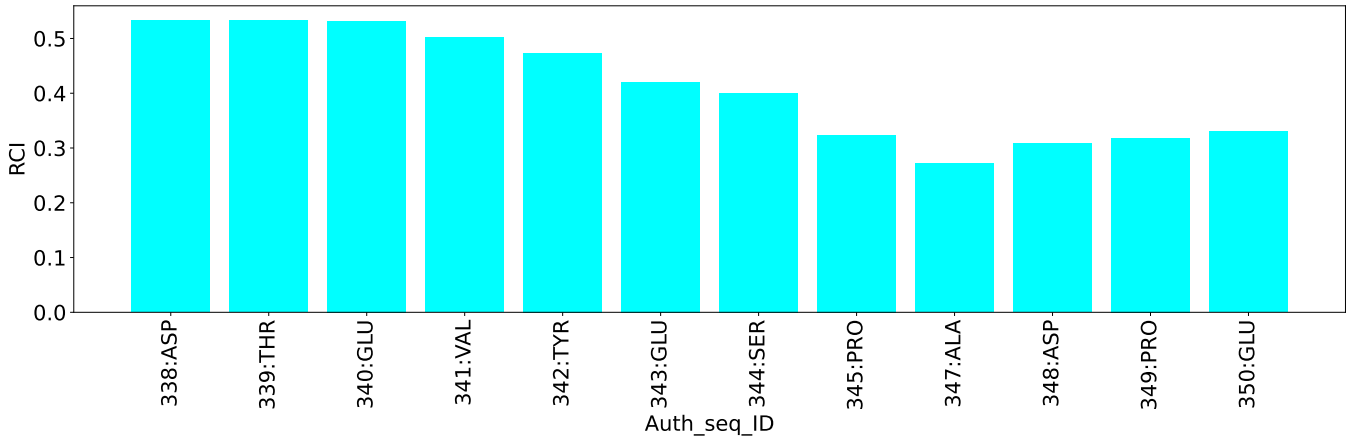
There are no statistically unusual chemical shifts.

7.2.5 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 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	2319
Intra-residue (|i-j|=0)	452
Sequential (|i-j|=1)	603
Medium range (|i-j|>1 and |i-j|<5)	471
Long range (|i-j|≥5)	767
Inter-chain	26
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	0
Number of restraints per residue	19.3
Number of long range restraints per residue[1]	6.4

[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)	28.0	0.2
0.2-0.5 (Medium)	6.0	0.3
>0.5 (Large)	None	None

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)

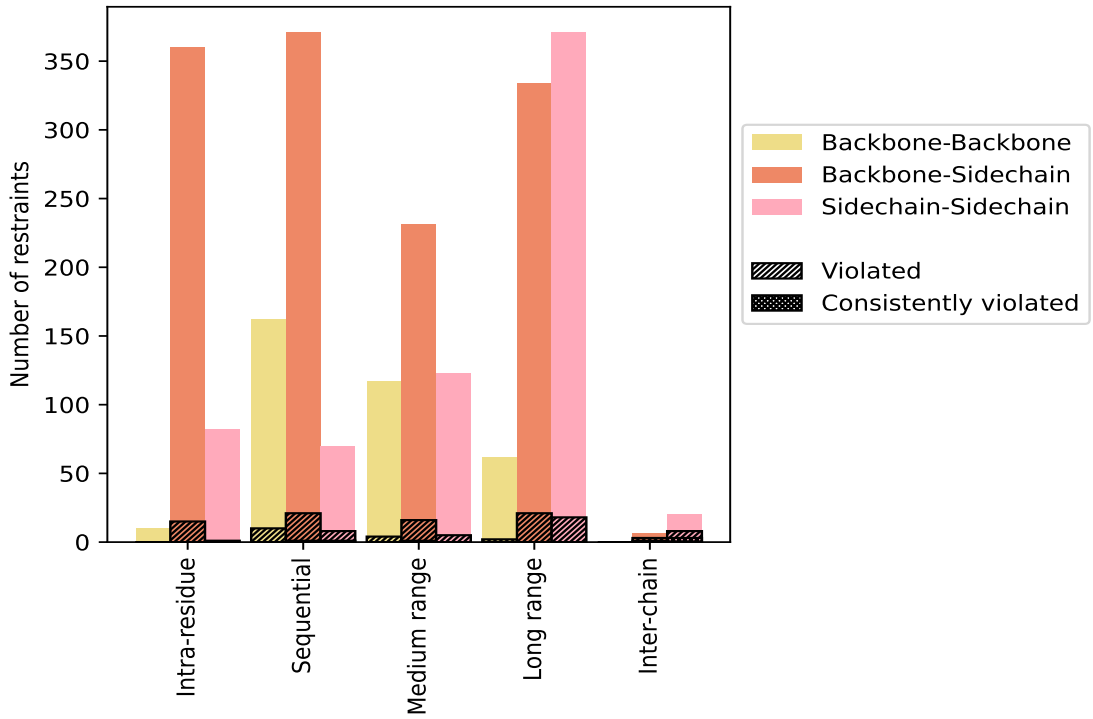
9.1 Summary of distance violations (i)

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)	452	19.5	16	3.5	0.7	0	0.0	0.0
Backbone-Backbone	10	0.4	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	360	15.5	15	4.2	0.6	0	0.0	0.0
Sidechain-Sidechain	82	3.5	1	1.2	0.0	0	0.0	0.0
Sequential (|i-j|=1)	603	26.0	39	6.5	1.7	2	0.3	0.1
Backbone-Backbone	162	7.0	10	6.2	0.4	0	0.0	0.0
Backbone-Sidechain	371	16.0	21	5.7	0.9	1	0.3	0.0
Sidechain-Sidechain	70	3.0	8	11.4	0.3	1	1.4	0.0
Medium range (|i-j|>1 & |i-j|<5)	471	20.3	25	5.3	1.1	1	0.2	0.0
Backbone-Backbone	117	5.0	4	3.4	0.2	0	0.0	0.0
Backbone-Sidechain	231	10.0	16	6.9	0.7	1	0.4	0.0
Sidechain-Sidechain	123	5.3	5	4.1	0.2	0	0.0	0.0
Long range (|i-j|≥5)	767	33.1	41	5.3	1.8	0	0.0	0.0
Backbone-Backbone	62	2.7	2	3.2	0.1	0	0.0	0.0
Backbone-Sidechain	334	14.4	21	6.3	0.9	0	0.0	0.0
Sidechain-Sidechain	371	16.0	18	4.9	0.8	0	0.0	0.0
Inter-chain	26	1.1	11	42.3	0.5	4	15.4	0.2
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	6	0.3	3	50.0	0.1	1	16.7	0.0
Sidechain-Sidechain	20	0.9	8	40.0	0.3	3	15.0	0.1
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	2319	100.0	132	5.7	5.7	7	0.3	0.3
Backbone-Backbone	351	15.1	16	4.6	0.7	0	0.0	0.0
Backbone-Sidechain	1302	56.1	76	5.8	3.3	3	0.2	0.1
Sidechain-Sidechain	666	28.7	40	6.0	1.7	4	0.6	0.2

[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	10	7	7	7	33	0.16	0.28	0.04	0.15
2	3	11	8	4	8	34	0.17	0.27	0.05	0.15
3	3	11	7	7	5	33	0.17	0.28	0.05	0.15
4	3	9	5	10	5	32	0.16	0.28	0.05	0.16
5	5	11	7	6	6	35	0.15	0.25	0.04	0.14
6	4	11	9	8	4	36	0.16	0.27	0.05	0.15
7	3	11	6	10	5	35	0.15	0.26	0.04	0.15
8	4	10	4	10	5	33	0.16	0.26	0.05	0.15
9	2	12	7	5	6	32	0.16	0.29	0.05	0.16
10	2	9	5	5	4	25	0.17	0.29	0.05	0.15
11	7	9	11	10	5	42	0.15	0.27	0.03	0.14

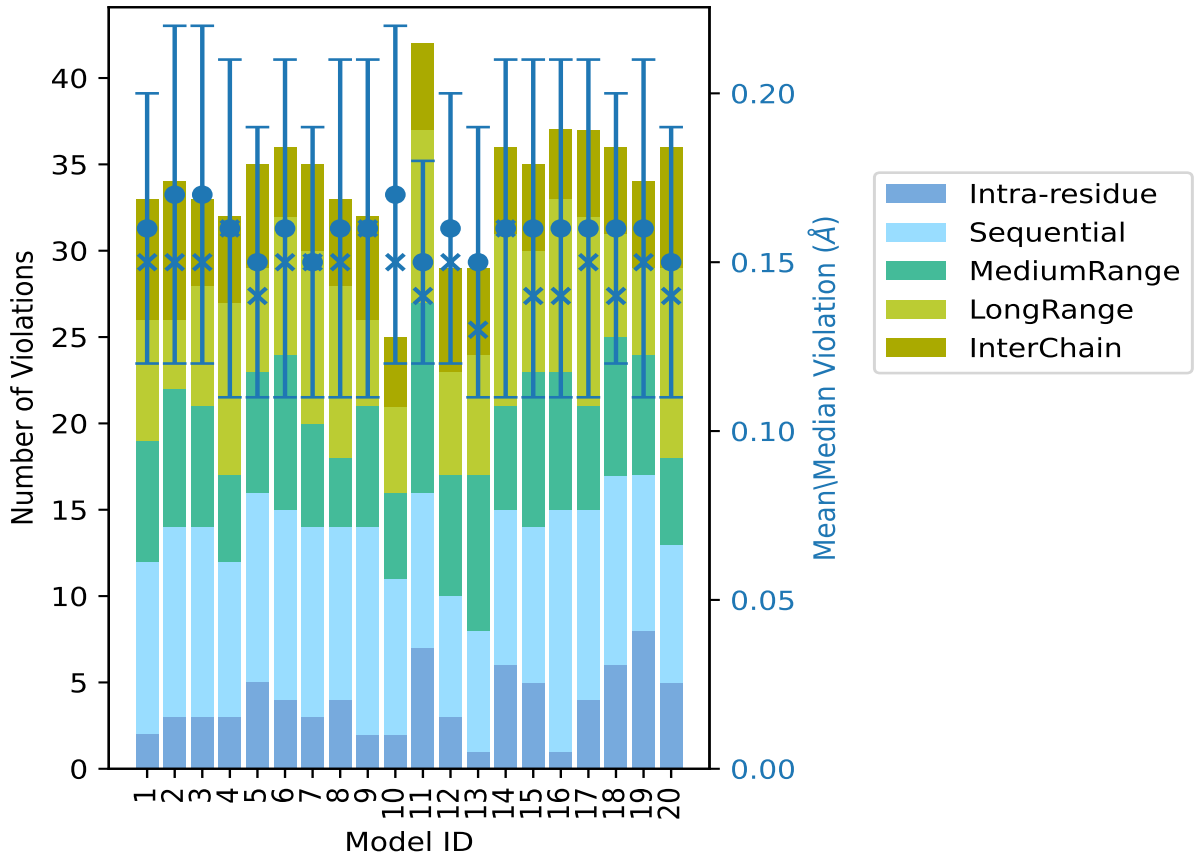
Continued on next page...

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	3	7	7	6	6	29	0.16	0.25	0.04	0.15
13	1	7	9	7	5	29	0.15	0.3	0.04	0.13
14	6	9	6	10	5	36	0.16	0.3	0.05	0.16
15	5	9	9	7	5	35	0.16	0.27	0.05	0.14
16	1	14	8	10	4	37	0.16	0.28	0.05	0.14
17	4	11	6	11	5	37	0.16	0.28	0.05	0.15
18	6	11	8	6	5	36	0.16	0.29	0.04	0.14
19	8	9	7	5	5	34	0.16	0.28	0.05	0.15
20	5	8	5	11	7	36	0.15	0.25	0.04	0.14

[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

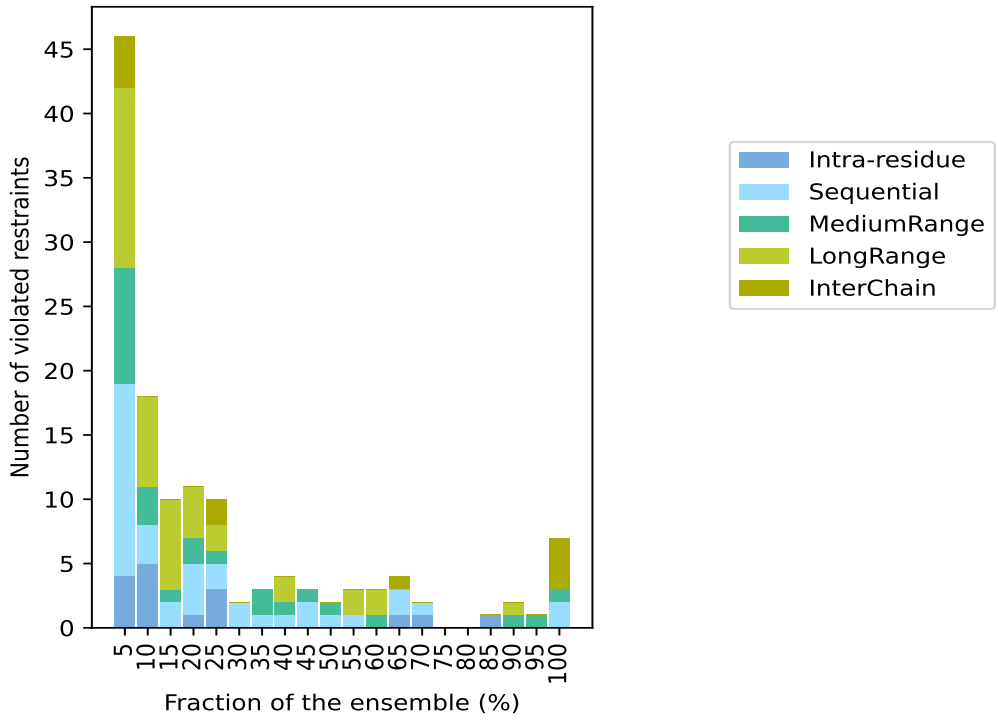
9.3 Distance violation statistics for the ensemble (i)

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, 2187(IR:436, SQ:564, MR:446, LR:726, IC:15) 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]	%
4	15	9	14	4	46	1	5.0
5	3	3	7	0	18	2	10.0
0	2	1	7	0	10	3	15.0
1	4	2	4	0	11	4	20.0
3	2	1	2	2	10	5	25.0
0	2	0	0	0	2	6	30.0
0	1	2	0	0	3	7	35.0
0	1	1	2	0	4	8	40.0
0	2	1	0	0	3	9	45.0
0	1	1	0	0	2	10	50.0
0	1	0	2	0	3	11	55.0
0	0	1	2	0	3	12	60.0
1	2	0	0	1	4	13	65.0
1	1	0	0	0	2	14	70.0
0	0	0	0	0	0	15	75.0
0	0	0	0	0	0	16	80.0
1	0	0	0	0	1	17	85.0
0	0	1	1	0	2	18	90.0
0	0	1	0	0	1	19	95.0
0	2	1	0	4	7	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

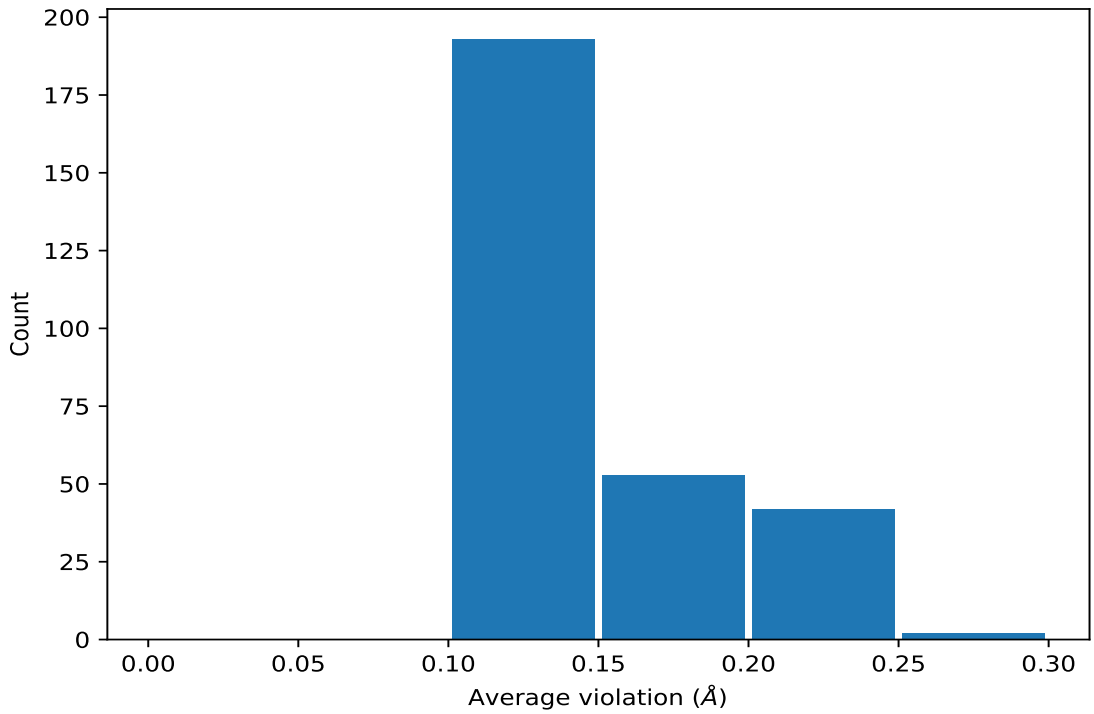
9.3.1 Bar graph : Distance violation statistics for the ensemble (i)



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)
(1,2307)	1:A:757:LEU:HD11	2:B:345:PRO:HB2	20	0.24	0.04	0.24
(1,2307)	1:A:757:LEU:HD11	2:B:345:PRO:HB3	20	0.24	0.04	0.24
(1,2307)	1:A:757:LEU:HD12	2:B:345:PRO:HB2	20	0.24	0.04	0.24
(1,2307)	1:A:757:LEU:HD12	2:B:345:PRO:HB3	20	0.24	0.04	0.24
(1,2307)	1:A:757:LEU:HD13	2:B:345:PRO:HB2	20	0.24	0.04	0.24
(1,2307)	1:A:757:LEU:HD13	2:B:345:PRO:HB3	20	0.24	0.04	0.24
(1,2307)	1:A:757:LEU:HD21	2:B:345:PRO:HB2	20	0.24	0.04	0.24
(1,2307)	1:A:757:LEU:HD21	2:B:345:PRO:HB3	20	0.24	0.04	0.24
(1,2307)	1:A:757:LEU:HD22	2:B:345:PRO:HB2	20	0.24	0.04	0.24
(1,2307)	1:A:757:LEU:HD22	2:B:345:PRO:HB3	20	0.24	0.04	0.24
(1,2307)	1:A:757:LEU:HD23	2:B:345:PRO:HB2	20	0.24	0.04	0.24
(1,2307)	1:A:757:LEU:HD23	2:B:345:PRO:HB3	20	0.24	0.04	0.24
(1,2309)	1:A:757:LEU:HD11	2:B:345:PRO:HB2	20	0.24	0.04	0.24
(1,2309)	1:A:757:LEU:HD11	2:B:345:PRO:HB3	20	0.24	0.04	0.24
(1,2309)	1:A:757:LEU:HD12	2:B:345:PRO:HB2	20	0.24	0.04	0.24
(1,2309)	1:A:757:LEU:HD12	2:B:345:PRO:HB3	20	0.24	0.04	0.24

Continued on next page...

Continued from previous page...

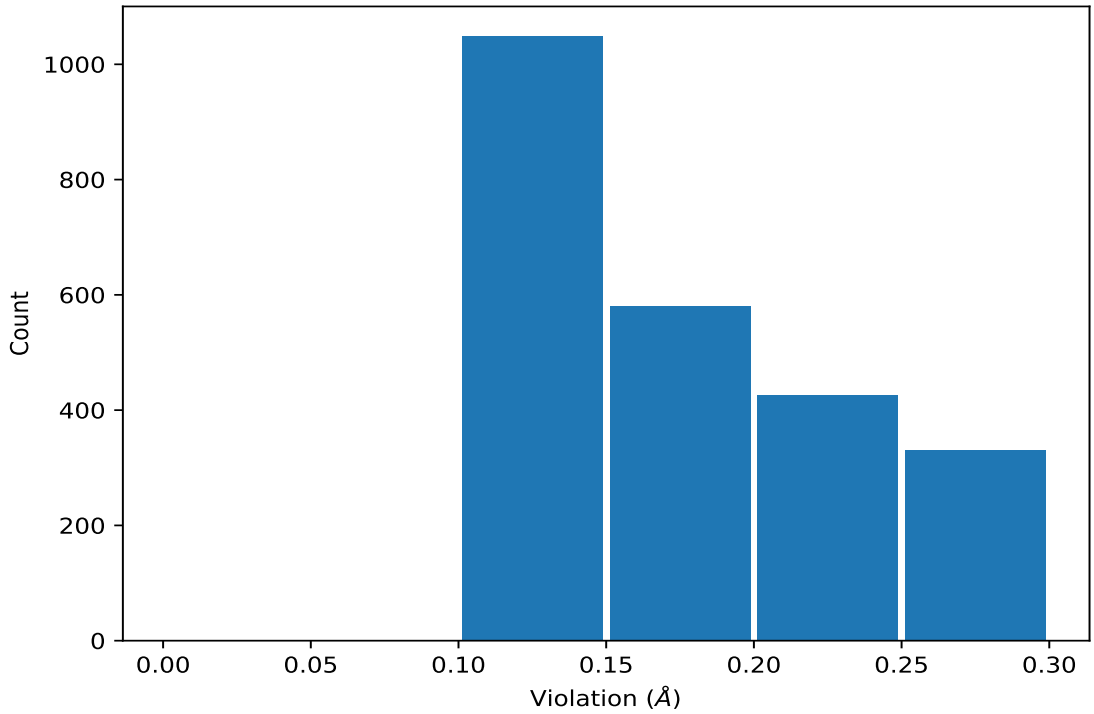
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,2309)	1:A:757:LEU:HD13	2:B:345:PRO:HB2	20	0.24	0.04	0.24
(1,2309)	1:A:757:LEU:HD13	2:B:345:PRO:HB3	20	0.24	0.04	0.24
(1,2309)	1:A:757:LEU:HD21	2:B:345:PRO:HB2	20	0.24	0.04	0.24
(1,2309)	1:A:757:LEU:HD21	2:B:345:PRO:HB3	20	0.24	0.04	0.24
(1,2309)	1:A:757:LEU:HD22	2:B:345:PRO:HB2	20	0.24	0.04	0.24
(1,2309)	1:A:757:LEU:HD22	2:B:345:PRO:HB3	20	0.24	0.04	0.24
(1,2309)	1:A:757:LEU:HD23	2:B:345:PRO:HB2	20	0.24	0.04	0.24
(1,2309)	1:A:757:LEU:HD23	2:B:345:PRO:HB3	20	0.24	0.04	0.24
(1,1628)	1:A:757:LEU:HD21	1:A:759:THR:H	20	0.23	0.03	0.23
(1,1628)	1:A:757:LEU:HD22	1:A:759:THR:H	20	0.23	0.03	0.23
(1,1628)	1:A:757:LEU:HD23	1:A:759:THR:H	20	0.23	0.03	0.23
(1,2312)	1:A:718:ILE:HB	2:B:346:PTR:HE1	20	0.19	0.04	0.2
(1,2312)	1:A:718:ILE:HB	2:B:346:PTR:HE2	20	0.19	0.04	0.2
(1,2201)	1:A:741:LEU:H	1:A:742:VAL:HG11	20	0.19	0.02	0.19
(1,2201)	1:A:741:LEU:H	1:A:742:VAL:HG12	20	0.19	0.02	0.19
(1,2201)	1:A:741:LEU:H	1:A:742:VAL:HG13	20	0.19	0.02	0.19
(1,2201)	1:A:741:LEU:H	1:A:742:VAL:HG21	20	0.19	0.02	0.19
(1,2201)	1:A:741:LEU:H	1:A:742:VAL:HG22	20	0.19	0.02	0.19
(1,2201)	1:A:741:LEU:H	1:A:742:VAL:HG23	20	0.19	0.02	0.19
(1,2314)	1:A:731:GLU:HB2	2:B:347:ALA:HA	20	0.18	0.04	0.18
(1,2314)	1:A:731:GLU:HB3	2:B:347:ALA:HA	20	0.18	0.04	0.18
(1,2208)	1:A:742:VAL:HG11	1:A:743:GLU:HG3	20	0.15	0.03	0.14
(1,2208)	1:A:742:VAL:HG12	1:A:743:GLU:HG3	20	0.15	0.03	0.14
(1,2208)	1:A:742:VAL:HG13	1:A:743:GLU:HG3	20	0.15	0.03	0.14
(1,2208)	1:A:742:VAL:HG21	1:A:743:GLU:HG3	20	0.15	0.03	0.14
(1,2208)	1:A:742:VAL:HG22	1:A:743:GLU:HG3	20	0.15	0.03	0.14
(1,2208)	1:A:742:VAL:HG23	1:A:743:GLU:HG3	20	0.15	0.03	0.14
(1,1552)	1:A:750:LEU:HD11	1:A:752:ASP:H	19	0.16	0.03	0.17
(1,1552)	1:A:750:LEU:HD12	1:A:752:ASP:H	19	0.16	0.03	0.17
(1,1552)	1:A:750:LEU:HD13	1:A:752:ASP:H	19	0.16	0.03	0.17
(1,1751)	1:A:682:GLU:HG3	1:A:717:HIS:HD2	18	0.16	0.02	0.16
(1,1508)	1:A:749:SER:H	1:A:751:LYS:HD2	18	0.15	0.02	0.15

[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)
(1,2318)	1:A:695:VAL:HG11	2:B:342:TYR:HD1	14	0.3
(1,2318)	1:A:695:VAL:HG11	2:B:342:TYR:HD2	14	0.3
(1,2318)	1:A:695:VAL:HG12	2:B:342:TYR:HD1	14	0.3
(1,2318)	1:A:695:VAL:HG12	2:B:342:TYR:HD2	14	0.3
(1,2318)	1:A:695:VAL:HG13	2:B:342:TYR:HD1	14	0.3
(1,2318)	1:A:695:VAL:HG13	2:B:342:TYR:HD2	14	0.3
(1,2318)	1:A:695:VAL:HG21	2:B:342:TYR:HD1	14	0.3
(1,2318)	1:A:695:VAL:HG21	2:B:342:TYR:HD2	14	0.3
(1,2318)	1:A:695:VAL:HG22	2:B:342:TYR:HD1	14	0.3
(1,2318)	1:A:695:VAL:HG22	2:B:342:TYR:HD2	14	0.3
(1,2318)	1:A:695:VAL:HG23	2:B:342:TYR:HD1	14	0.3
(1,2318)	1:A:695:VAL:HG23	2:B:342:TYR:HD2	14	0.3
(1,2180)	1:A:736:ARG:HG2	1:A:737:GLY:H	13	0.3
(1,2180)	1:A:736:ARG:HG3	1:A:737:GLY:H	13	0.3
(1,319)	1:A:753:CYS:HB2	1:A:754:PHE:HA	9	0.29
(1,2309)	1:A:757:LEU:HD11	2:B:345:PRO:HB2	10	0.29
(1,2309)	1:A:757:LEU:HD11	2:B:345:PRO:HB3	10	0.29

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,2309)	1:A:757:LEU:HD12	2:B:345:PRO:HB2	10	0.29
(1,2309)	1:A:757:LEU:HD12	2:B:345:PRO:HB3	10	0.29
(1,2309)	1:A:757:LEU:HD13	2:B:345:PRO:HB2	10	0.29
(1,2309)	1:A:757:LEU:HD13	2:B:345:PRO:HB3	10	0.29
(1,2309)	1:A:757:LEU:HD21	2:B:345:PRO:HB2	10	0.29
(1,2309)	1:A:757:LEU:HD21	2:B:345:PRO:HB3	10	0.29
(1,2309)	1:A:757:LEU:HD22	2:B:345:PRO:HB2	10	0.29
(1,2309)	1:A:757:LEU:HD22	2:B:345:PRO:HB3	10	0.29
(1,2309)	1:A:757:LEU:HD23	2:B:345:PRO:HB2	10	0.29
(1,2309)	1:A:757:LEU:HD23	2:B:345:PRO:HB3	10	0.29
(1,2307)	1:A:757:LEU:HD11	2:B:345:PRO:HB2	10	0.29
(1,2307)	1:A:757:LEU:HD11	2:B:345:PRO:HB3	10	0.29
(1,2307)	1:A:757:LEU:HD12	2:B:345:PRO:HB2	10	0.29
(1,2307)	1:A:757:LEU:HD12	2:B:345:PRO:HB3	10	0.29
(1,2307)	1:A:757:LEU:HD13	2:B:345:PRO:HB2	10	0.29
(1,2307)	1:A:757:LEU:HD13	2:B:345:PRO:HB3	10	0.29
(1,2307)	1:A:757:LEU:HD21	2:B:345:PRO:HB2	10	0.29
(1,2307)	1:A:757:LEU:HD21	2:B:345:PRO:HB3	10	0.29
(1,2307)	1:A:757:LEU:HD22	2:B:345:PRO:HB2	10	0.29
(1,2307)	1:A:757:LEU:HD22	2:B:345:PRO:HB3	10	0.29
(1,2307)	1:A:757:LEU:HD23	2:B:345:PRO:HB2	10	0.29
(1,2307)	1:A:757:LEU:HD23	2:B:345:PRO:HB3	10	0.29
(1,2158)	1:A:730:THR:HA	1:A:731:GLU:HG2	18	0.29
(1,2158)	1:A:730:THR:HA	1:A:731:GLU:HG3	18	0.29
(1,670)	1:A:732:LYS:HA	1:A:732:LYS:HE2	19	0.28
(1,670)	1:A:732:LYS:HA	1:A:732:LYS:HE3	19	0.28
(1,2318)	1:A:695:VAL:HG11	2:B:342:TYR:HD1	19	0.28
(1,2318)	1:A:695:VAL:HG11	2:B:342:TYR:HD2	19	0.28
(1,2318)	1:A:695:VAL:HG12	2:B:342:TYR:HD1	19	0.28
(1,2318)	1:A:695:VAL:HG12	2:B:342:TYR:HD2	19	0.28
(1,2318)	1:A:695:VAL:HG13	2:B:342:TYR:HD1	19	0.28
(1,2318)	1:A:695:VAL:HG13	2:B:342:TYR:HD2	19	0.28
(1,2318)	1:A:695:VAL:HG21	2:B:342:TYR:HD1	19	0.28
(1,2318)	1:A:695:VAL:HG21	2:B:342:TYR:HD2	19	0.28
(1,2318)	1:A:695:VAL:HG22	2:B:342:TYR:HD1	19	0.28
(1,2318)	1:A:695:VAL:HG22	2:B:342:TYR:HD2	19	0.28
(1,2318)	1:A:695:VAL:HG23	2:B:342:TYR:HD1	19	0.28
(1,2318)	1:A:695:VAL:HG23	2:B:342:TYR:HD2	19	0.28
(1,2314)	1:A:731:GLU:HB2	2:B:347:ALA:HA	1	0.28
(1,2314)	1:A:731:GLU:HB3	2:B:347:ALA:HA	1	0.28
(1,2309)	1:A:757:LEU:HD11	2:B:345:PRO:HB2	3	0.28

10 Dihedral-angle violation analysis (i)

No dihedral-angle restraints found