wwPDB NMR Structure Validation Summary Report

Jun 3, 2023 – 09:12 AM EDT

PDB ID	:	2HAJ
BMRB ID	:	6284
Title	:	Solution structure of the helicase-binding domain of Escherichia coli primase
Authors	:	Su, X.C.; Loscha, K.V.; Dixon, N.E.; Otting, G.
Deposited on	:	2006-06-13

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 (1)) were used in the production of this report:

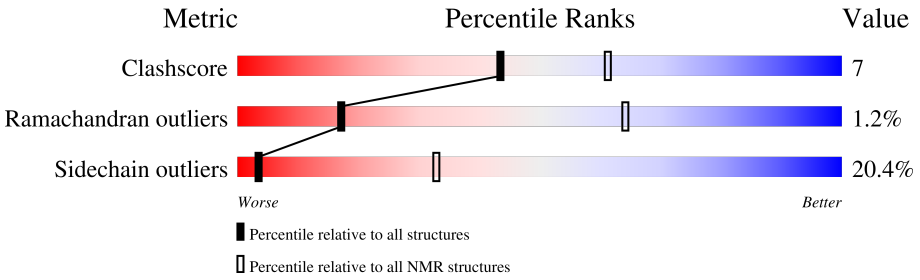
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal 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 86%.

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	148	64% 22% 5% 9%

2 Ensemble composition and analysis

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:451-A:578 (128)	1.01	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 7 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2167 atoms, of which 1086 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called DNA primase.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	135	2167	677	1086	189	211	4	0

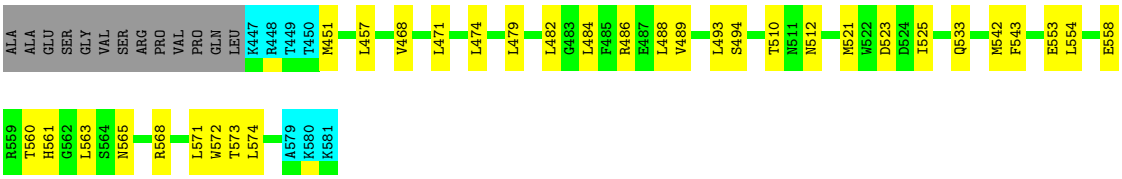
4 Residue-property plots

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 primase

Chain A:

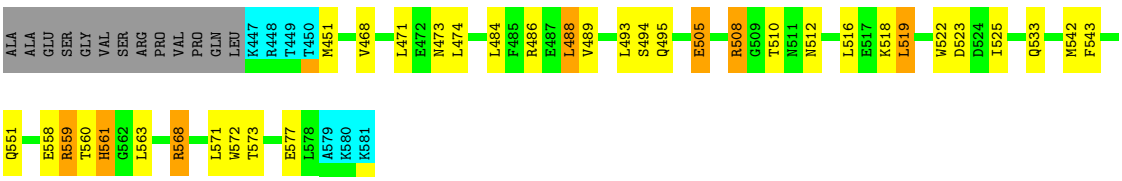


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

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

• Molecule 1: DNA primase

Chain A:



5 Refinement protocol and experimental data overview

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

Of the 200 calculated structures, 20 were deposited, based on the following criterion: *structures with the least restraint violations*.

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

Software name	Classification	Version
CYANA	structure solution	2.0
CYANA	refinement	2.0

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	2
Total number of shifts	2335
Number of shifts mapped to atoms	2242
Number of unparsed shifts	0
Number of shifts with mapping errors	93
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	86%

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	0.62±0.04	0±0/1040 (0.0± 0.0%)	1.03±0.07	1±1/1413 (0.1± 0.1%)
All	All	0.62	0/20800 (0.0%)	1.03	16/28260 (0.1%)

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

Mol	Chain	Chirality	Planarity
1	A	0.0±0.0	0.3±0.5
All	All	0	6

There are no bond-length outliers.

5 of 12 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	486	ARG	NE-CZ-NH2	-7.86	116.37	120.30	6	4
1	A	508	ARG	NE-CZ-NH2	-7.84	116.38	120.30	1	2
1	A	468	VAL	CG1-CB-CG2	-6.71	100.17	110.90	3	1
1	A	519	LEU	CB-CG-CD1	6.18	121.50	111.00	1	1
1	A	457	LEU	CB-CG-CD2	6.00	121.19	111.00	3	1

There are no chirality outliers.

5 of 6 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	559	ARG	Sidechain	1
1	A	538	SER	Mainchain	1
1	A	540	ASN	Mainchain	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	573	THR	Mainchain	1
1	A	507	TYR	Sidechain	1

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	1024	1018	1021	14±13
All	All	20480	20360	20420	286

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 7.

5 of 151 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:471:LEU:CD1	1:A:474:LEU:HD12	0.91	1.96	13	1
1:A:488:LEU:HD11	1:A:516:LEU:HD21	0.89	1.44	13	8
1:A:455:ILE:HD13	1:A:484:LEU:HD23	0.88	1.45	19	1
1:A:554:LEU:HD11	1:A:571:LEU:HD21	0.86	1.46	20	1
1:A:457:LEU:HD21	1:A:525:ILE:HG21	0.81	1.50	2	6

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	128/148 (86%)	117±2 (91±1%)	9±2 (7±2%)	2±1 (1±1%)	17	64
All	All	2560/2960 (86%)	2341 (91%)	189 (7%)	30 (1%)	17	64

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	475	ASP	8
1	A	526	ALA	5
1	A	451	MET	3
1	A	563	LEU	3
1	A	525	ILE	2

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	115/131 (88%)	92±5 (80±5%)	23±5 (20±5%)	3	33
All	All	2300/2620 (88%)	1831 (80%)	469 (20%)	3	33

5 of 74 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	572	TRP	20
1	A	493	LEU	19
1	A	523	ASP	16
1	A	542	MET	16
1	A	521	MET	15

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

There are no ligands in this entry.

6.7 Other polymers

There are no such molecules in this entry.

6.8 Polymer linkage issues

There are no chain breaks in this entry.

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 86% for the well-defined parts and 84% 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	915
Number of shifts mapped to atoms	842
Number of unparsed shifts	0
Number of shifts with mapping errors	73
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

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 73) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	434	ALA	CA	51.2	0.2	1
1	A	434	ALA	HA	4.0	0.02	1
1	A	434	ALA	CB	19.3	0.2	1
1	A	434	ALA	C	173.4	0.2	1
1	A	435	ALA	N	123.4	0.2	1
1	A	435	ALA	H	8.55	0.02	1
1	A	435	ALA	CA	52.5	0.2	1
1	A	435	ALA	HA	4.24	0.02	1
1	A	435	ALA	CB	18.8	0.2	1
1	A	435	ALA	C	177.5	0.2	1
1	A	436	GLU	N	120.8	0.2	1
1	A	436	GLU	H	8.52	0.02	1
1	A	436	GLU	CA	56.5	0.2	1
1	A	436	GLU	HA	4.22	0.02	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	436	GLU	CB	30.1	0.2	1
1	A	436	GLU	C	176.5	0.2	1
1	A	437	SER	N	116.9	0.2	1
1	A	437	SER	H	8.35	0.02	1
1	A	437	SER	CA	58.4	0.2	1
1	A	437	SER	HA	4.35	0.02	1
1	A	437	SER	CB	63.8	0.2	1
1	A	437	SER	C	174.9	0.2	1
1	A	438	GLY	N	110.9	0.2	1
1	A	438	GLY	H	8.39	0.02	1
1	A	438	GLY	CA	45.2	0.2	1
1	A	438	GLY	HA2	3.9	0.02	1
1	A	438	GLY	HA3	3.9	0.02	1
1	A	438	GLY	C	173.9	0.2	1
1	A	439	VAL	N	118.9	0.2	1
1	A	439	VAL	H	7.86	0.02	1
1	A	439	VAL	CA	62.0	0.2	1
1	A	439	VAL	HA	4.08	0.02	1
1	A	439	VAL	CB	32.9	0.2	1
1	A	439	VAL	C	176.1	0.2	1
1	A	440	SER	N	119.9	0.2	1
1	A	440	SER	H	8.33	0.02	1
1	A	440	SER	CA	58.1	0.2	1
1	A	440	SER	HA	4.36	0.02	1
1	A	440	SER	CB	63.7	0.2	1
1	A	440	SER	C	173.8	0.2	1
1	A	441	ARG	N	124.1	0.2	1
1	A	441	ARG	H	8.24	0.02	1
1	A	441	ARG	CA	53.7	0.2	1
1	A	441	ARG	HA	4.55	0.02	1
1	A	441	ARG	C	174.1	0.2	1
1	A	442	PRO	CA	62.8	0.2	1
1	A	442	PRO	HA	4.35	0.02	1
1	A	442	PRO	CB	32.1	0.2	1
1	A	442	PRO	C	176.5	0.2	1
1	A	443	VAL	N	121.9	0.2	1
1	A	443	VAL	H	8.22	0.02	1
1	A	443	VAL	CA	59.9	0.2	1
1	A	443	VAL	HA	4.25	0.02	1
1	A	443	VAL	C	174.6	0.2	1
1	A	444	PRO	CA	63.1	0.2	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	444	PRO	HA	4.27	0.02	1
1	A	444	PRO	CB	32.2	0.2	1
1	A	444	PRO	C	176.6	0.2	1
1	A	445	GLN	N	120.9	0.2	1
1	A	445	GLN	H	8.38	0.02	1
1	A	445	GLN	CA	55.6	0.2	1
1	A	445	GLN	HA	4.19	0.02	1
1	A	445	GLN	CB	29.7	0.2	1
1	A	445	GLN	NE2	112.7	0.2	1
1	A	445	GLN	HE21	7.48	0.02	2
1	A	445	GLN	HE22	6.84	0.02	2
1	A	445	GLN	C	176.0	0.2	1
1	A	446	LEU	N	123.6	0.2	1
1	A	446	LEU	H	8.18	0.02	1
1	A	446	LEU	CA	54.9	0.2	1
1	A	446	LEU	HA	4.2	0.02	1
1	A	446	LEU	CB	42.3	0.2	1
1	A	446	LEU	C	177.0	0.2	1

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	147	-0.62 ± 0.12	Should be checked
$^{13}C_{\beta}$	132	0.12 ± 0.08	None needed (< 0.5 ppm)
$^{13}C'$	146	-0.39 ± 0.10	None needed (< 0.5 ppm)
^{15}N	140	0.12 ± 0.27	None needed (< 0.5 ppm)

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 45%, i.e. 801 atoms were assigned a chemical shift out of a possible 1795. 0 out of 30 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	633/637 (99%)	257/258 (100%)	253/256 (99%)	123/123 (100%)
Sidechain	164/1074 (15%)	32/697 (5%)	116/337 (34%)	16/40 (40%)
Aromatic	4/84 (5%)	2/43 (5%)	0/36 (0%)	2/5 (40%)
Overall	801/1795 (45%)	291/998 (29%)	369/629 (59%)	141/168 (84%)

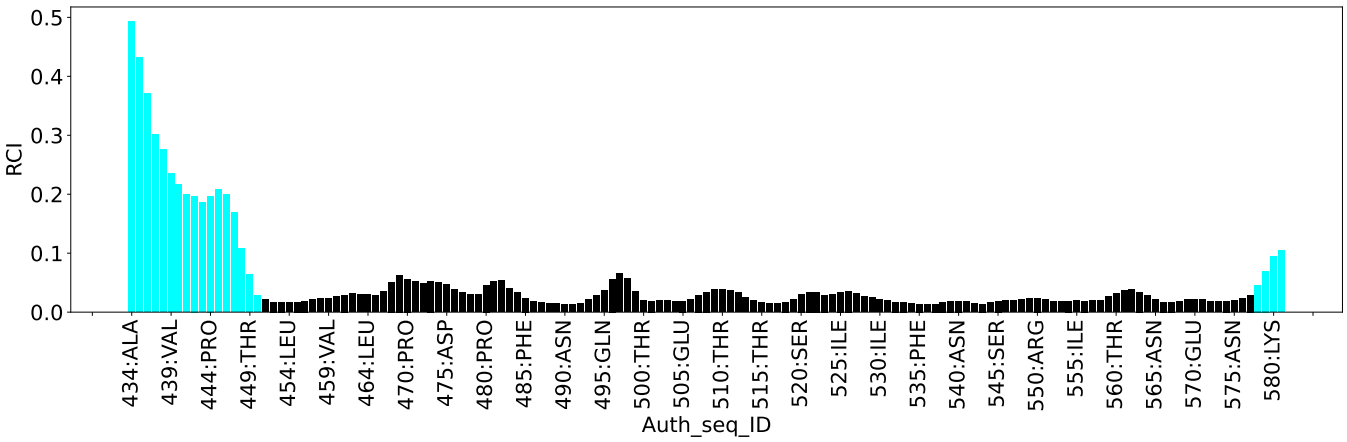
7.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

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	1420
Number of shifts mapped to atoms	1400
Number of unparsed shifts	0
Number of shifts with mapping errors	20
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 20) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
2	A	435	ALA	N	123.6	.	1
2	A	435	ALA	H	8.66	.	1
2	A	436	GLU	N	120.9	.	1
2	A	436	GLU	H	8.59	.	1
2	A	437	SER	N	116.9	.	1
2	A	437	SER	H	8.42	.	1
2	A	438	GLY	N	111.0	.	1
2	A	438	GLY	H	8.47	.	1
2	A	439	VAL	N	118.9	.	1
2	A	439	VAL	H	7.93	.	1
2	A	440	SER	N	119.9	.	1
2	A	440	SER	H	8.4	.	1
2	A	441	ARG	N	124.2	.	1
2	A	441	ARG	H	8.3	.	1
2	A	443	VAL	N	122.0	.	1
2	A	443	VAL	H	8.27	.	1
2	A	445	GLN	N	120.8	.	1
2	A	445	GLN	H	8.44	.	1
2	A	446	LEU	N	123.6	.	1
2	A	446	LEU	H	8.24	.	1

7.2.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	111	-0.65 ± 0.12	Should be checked
$^{13}C_{\beta}$	113	0.21 ± 0.06	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	140	0.12 ± 0.28	None needed (< 0.5 ppm)

7.2.3 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 76%, i.e. 1366 atoms were assigned a chemical shift out of a possible 1795. 0 out of 30 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	490/637 (77%)	258/258 (100%)	109/256 (43%)	123/123 (100%)
Sidechain	843/1074 (78%)	623/697 (89%)	219/337 (65%)	1/40 (2%)
Aromatic	33/84 (39%)	33/43 (77%)	0/36 (0%)	0/5 (0%)
Overall	1366/1795 (76%)	914/998 (92%)	328/629 (52%)	124/168 (74%)

7.2.4 Statistically unusual chemical shifts (i)

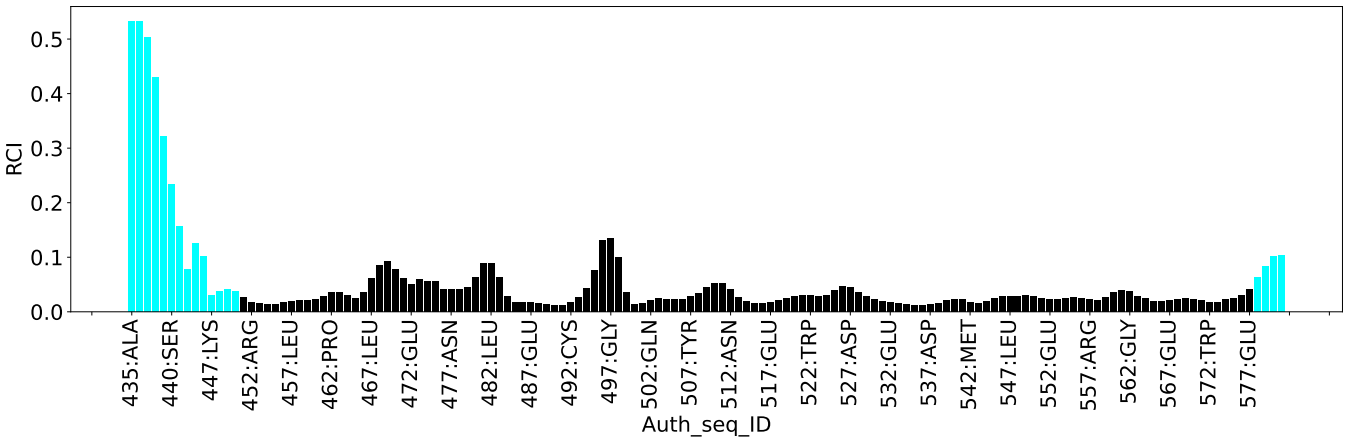
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
2	A	510	THR	HG1	5.99	0.08 – 2.19	23.0
2	A	499	THR	HG1	5.62	0.08 – 2.19	21.2

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 A:



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	1604
Intra-residue (|i-j|=0)	288
Sequential (|i-j|=1)	364
Medium range (|i-j|>1 and |i-j|<5)	481
Long range (|i-j|≥5)	471
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	0
Number of restraints per residue	10.8
Number of long range restraints per residue[1]	3.2

[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

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)	0.6	0.16
0.2-0.5 (Medium)	0.1	0.35
>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)

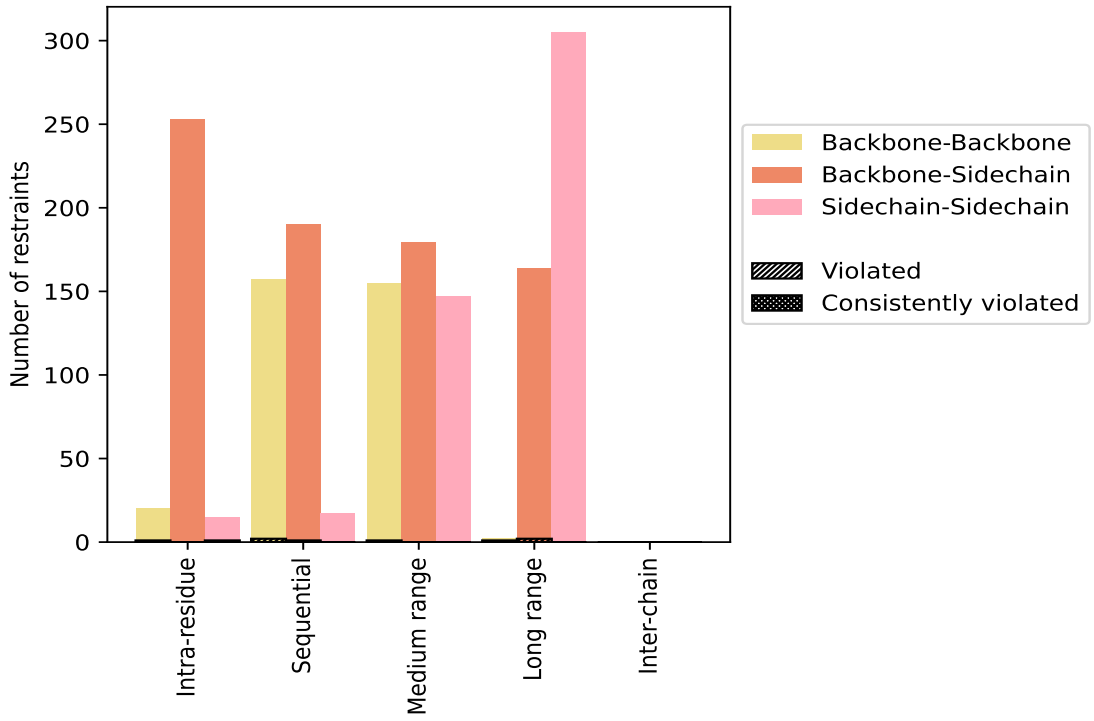
9.1 Summary of distance violations (i)

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)	288	18.0	2	0.7	0.1	0	0.0	0.0
Backbone-Backbone	20	1.2	1	5.0	0.1	0	0.0	0.0
Backbone-Sidechain	253	15.8	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	15	0.9	1	6.7	0.1	0	0.0	0.0
Sequential (|i-j|=1)	364	22.7	3	0.8	0.2	0	0.0	0.0
Backbone-Backbone	157	9.8	2	1.3	0.1	0	0.0	0.0
Backbone-Sidechain	190	11.8	1	0.5	0.1	0	0.0	0.0
Sidechain-Sidechain	17	1.1	0	0.0	0.0	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	481	30.0	1	0.2	0.1	0	0.0	0.0
Backbone-Backbone	155	9.7	1	0.6	0.1	0	0.0	0.0
Backbone-Sidechain	179	11.2	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	147	9.2	0	0.0	0.0	0	0.0	0.0
Long range (|i-j|≥5)	471	29.4	3	0.6	0.2	0	0.0	0.0
Backbone-Backbone	2	0.1	1	50.0	0.1	0	0.0	0.0
Backbone-Sidechain	164	10.2	2	1.2	0.1	0	0.0	0.0
Sidechain-Sidechain	305	19.0	0	0.0	0.0	0	0.0	0.0
Inter-chain	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Hydrogen bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	1604	100.0	9	0.6	0.6	0	0.0	0.0
Backbone-Backbone	334	20.8	5	1.5	0.3	0	0.0	0.0
Backbone-Sidechain	786	49.0	3	0.4	0.2	0	0.0	0.0
Sidechain-Sidechain	484	30.2	1	0.2	0.1	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



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	0	0	0	0	0	0	0.0	0.0	0.0	0.0
2	0	0	0	0	0	0	0.0	0.0	0.0	0.0
3	0	0	0	1	0	1	0.11	0.11	0.0	0.11
4	0	0	0	0	0	0	0.0	0.0	0.0	0.0
5	1	0	0	1	0	2	0.15	0.16	0.01	0.15
6	0	0	0	0	0	0	0.0	0.0	0.0	0.0
7	0	0	0	0	0	0	0.0	0.0	0.0	0.0
8	0	0	0	0	0	0	0.0	0.0	0.0	0.0
9	0	1	0	0	0	1	0.13	0.13	0.0	0.13
10	0	0	0	0	0	0	0.0	0.0	0.0	0.0
11	0	0	0	0	0	0	0.0	0.0	0.0	0.0

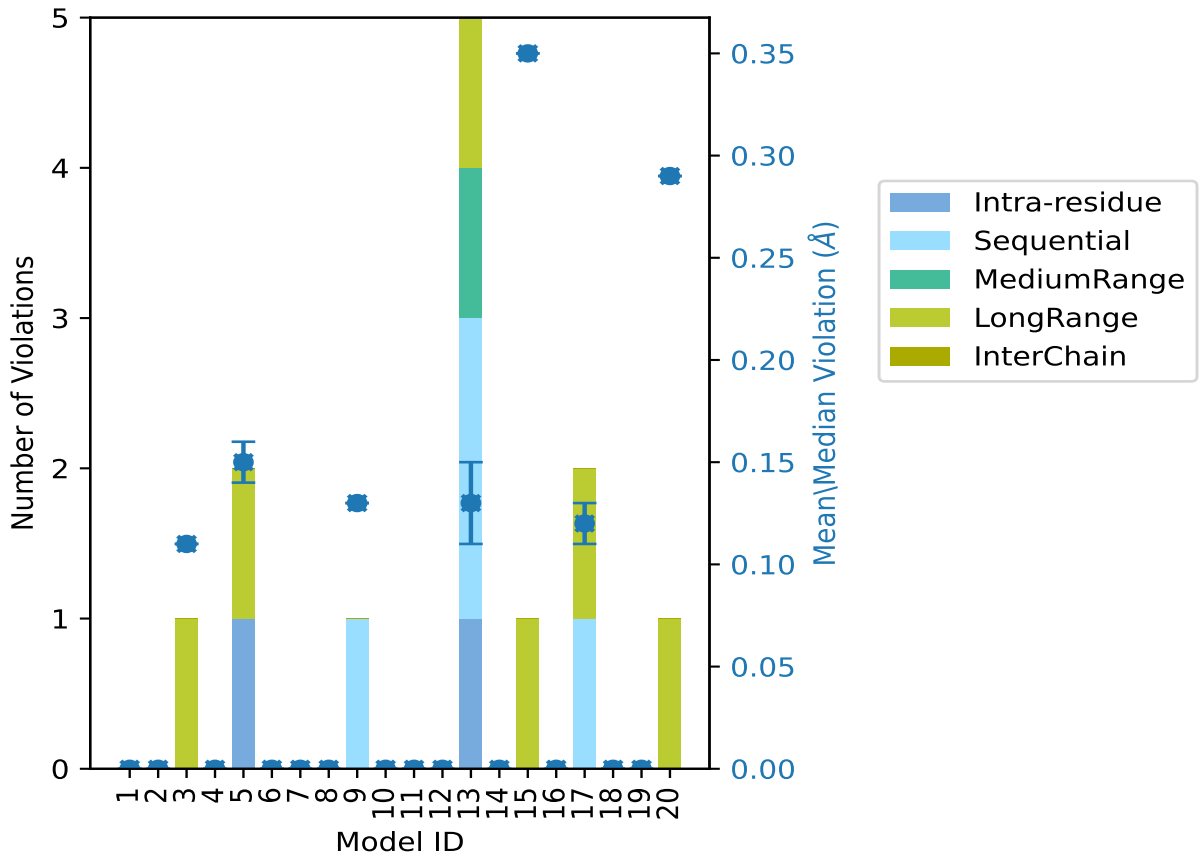
Continued on next page...

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	0	0	0	0	0	0	0.0	0.0	0.0	0.0
13	1	2	1	1	0	5	0.13	0.16	0.02	0.13
14	0	0	0	0	0	0	0.0	0.0	0.0	0.0
15	0	0	0	1	0	1	0.35	0.35	0.0	0.35
16	0	0	0	0	0	0	0.0	0.0	0.0	0.0
17	0	1	0	1	0	2	0.12	0.13	0.01	0.12
18	0	0	0	0	0	0	0.0	0.0	0.0	0.0
19	0	0	0	0	0	0	0.0	0.0	0.0	0.0
20	0	0	0	1	0	1	0.29	0.29	0.0	0.29

[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

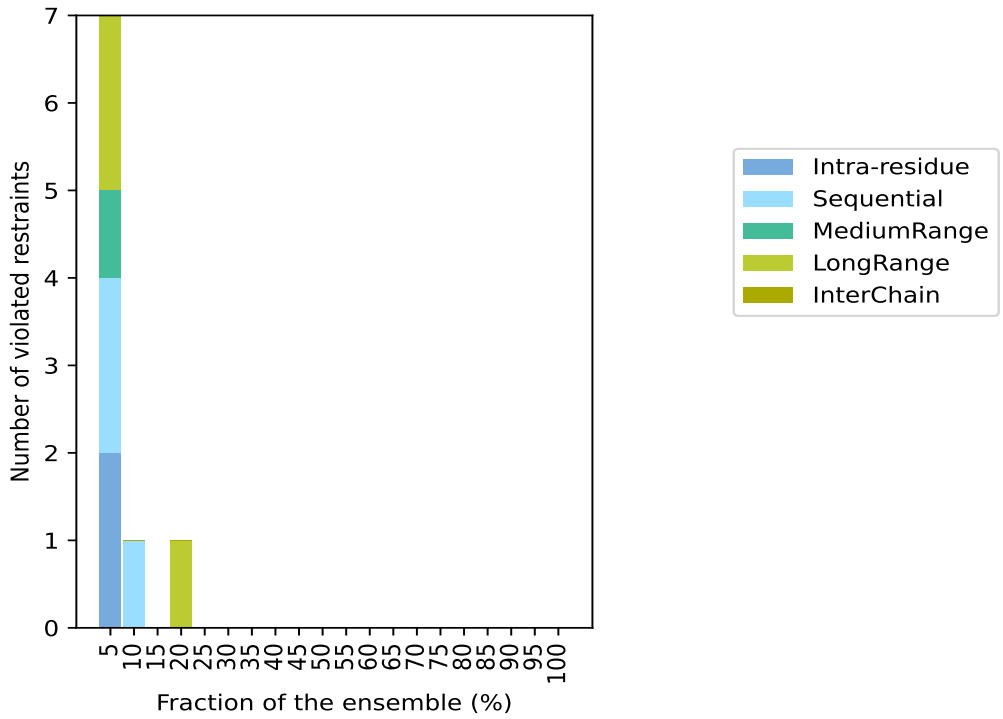
9.3 Distance violation statistics for the ensemble (i)

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, 1595(IR:286, SQ:361, MR:480, LR:468, IC:0) 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]	%
2	2	1	2	0	7	1	5.0
0	1	0	0	0	1	2	10.0
0	0	0	0	0	0	3	15.0
0	0	0	1	0	1	4	20.0
0	0	0	0	0	0	5	25.0
0	0	0	0	0	0	6	30.0
0	0	0	0	0	0	7	35.0
0	0	0	0	0	0	8	40.0
0	0	0	0	0	0	9	45.0
0	0	0	0	0	0	10	50.0
0	0	0	0	0	0	11	55.0
0	0	0	0	0	0	12	60.0
0	0	0	0	0	0	13	65.0
0	0	0	0	0	0	14	70.0
0	0	0	0	0	0	15	75.0
0	0	0	0	0	0	16	80.0
0	0	0	0	0	0	17	85.0
0	0	0	0	0	0	18	90.0
0	0	0	0	0	0	19	95.0
0	0	0	0	0	0	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

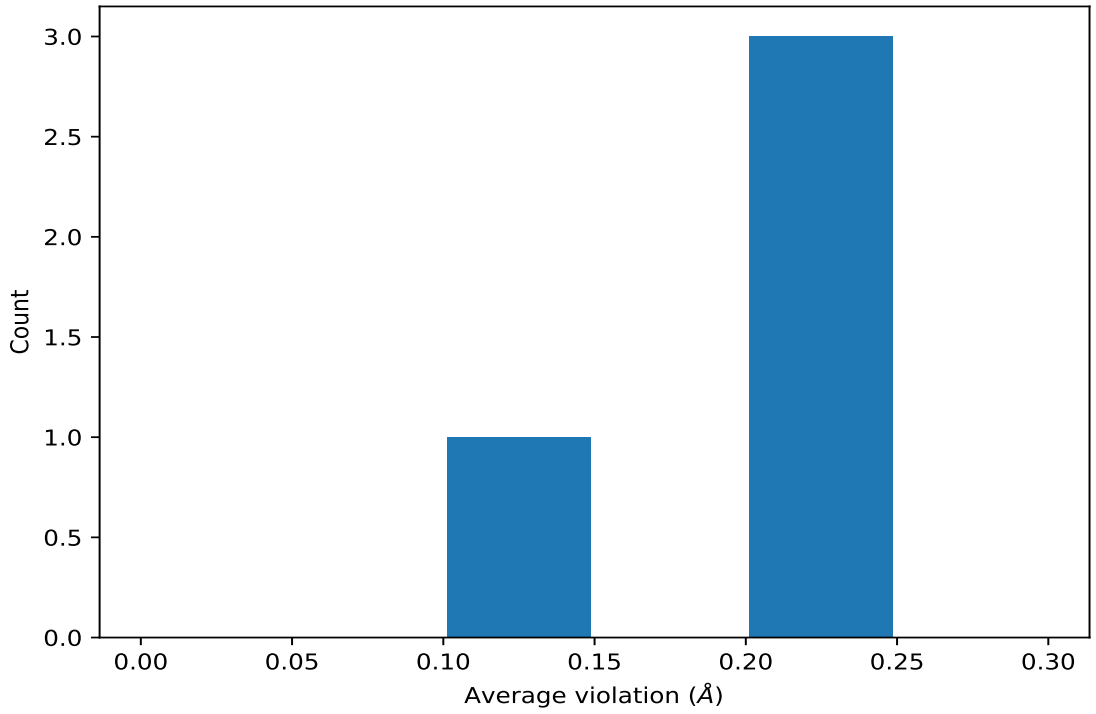
9.3.1 Bar graph : Distance violation statistics for the ensemble



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 violation for each restraint sorted by number of violated models and the mean 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.

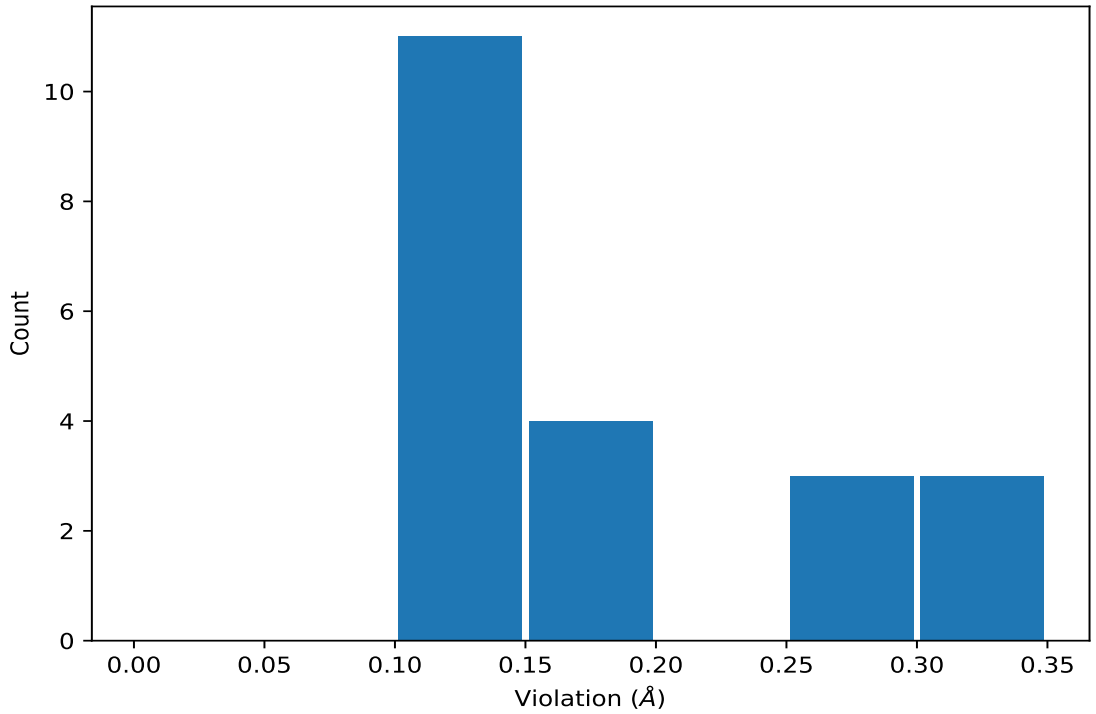
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,105)	1:A:456:GLY:HA2	1:A:525:ILE:HD11	4	0.23	0.1	0.22
(1,105)	1:A:456:GLY:HA2	1:A:525:ILE:HD12	4	0.23	0.1	0.22
(1,105)	1:A:456:GLY:HA2	1:A:525:ILE:HD13	4	0.23	0.1	0.22
(1,101)	1:A:456:GLY:HA2	1:A:457:LEU:H	2	0.12	0.01	0.12

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

9.5 All violated distance restraints

9.5.1 Histogram : Distribution of distance violations

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)
(1,105)	1:A:456:GLY:HA2	1:A:525:ILE:HD11	15	0.35
(1,105)	1:A:456:GLY:HA2	1:A:525:ILE:HD12	15	0.35
(1,105)	1:A:456:GLY:HA2	1:A:525:ILE:HD13	15	0.35
(1,105)	1:A:456:GLY:HA2	1:A:525:ILE:HD11	20	0.29
(1,105)	1:A:456:GLY:HA2	1:A:525:ILE:HD12	20	0.29
(1,105)	1:A:456:GLY:HA2	1:A:525:ILE:HD13	20	0.29
(1,308)	1:A:476:GLU:HB3	1:A:477:ASN:H	13	0.16
(1,105)	1:A:456:GLY:HA2	1:A:525:ILE:HD11	5	0.16
(1,105)	1:A:456:GLY:HA2	1:A:525:ILE:HD12	5	0.16
(1,105)	1:A:456:GLY:HA2	1:A:525:ILE:HD13	5	0.16
(1,41)	1:A:453:ILE:HB	1:A:453:ILE:HG12	5	0.14
(1,965)	1:A:558:GLU:HA	1:A:562:GLY:H	13	0.13
(1,964)	1:A:558:GLU:H	1:A:563:LEU:HG	13	0.13
(1,293)	1:A:475:ASP:HA	1:A:476:GLU:H	17	0.13
(1,101)	1:A:456:GLY:HA2	1:A:457:LEU:H	9	0.13
(1,101)	1:A:456:GLY:HA2	1:A:457:LEU:H	13	0.12

10 Dihedral-angle violation analysis (i)

No dihedral-angle restraints found