wwPDB NMR Structure Validation Summary Report

Jun 3, 2023 – 08:15 PM EDT

PDB ID	:	2MZ8
BMRB ID	:	25476
Title	:	Solution NMR structure of Salmonella Typhimurium transcriptional regulator protein Crl
Authors	:	Cavaliere, P.; Levi-Acobas, F.; Monteil, V.; Bellalou, J.; Mayer, C.; Norel, F.; Sizun, C.
Deposited on	:	2015-02-07

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:

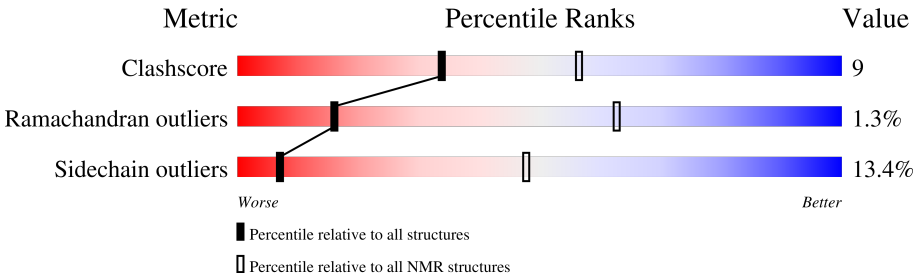
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal 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 77%.

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	154	64% 21% . 14%

2 Ensemble composition and analysis

This entry contains 20 models. Model 3 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:1-A:133 (133)	1.43	3

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 2 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2195 atoms, of which 1081 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Sigma factor-binding protein Crl.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	133	2195	719	1081	185	204	6	0

There are 21 discrepancies between the modelled and reference sequences:

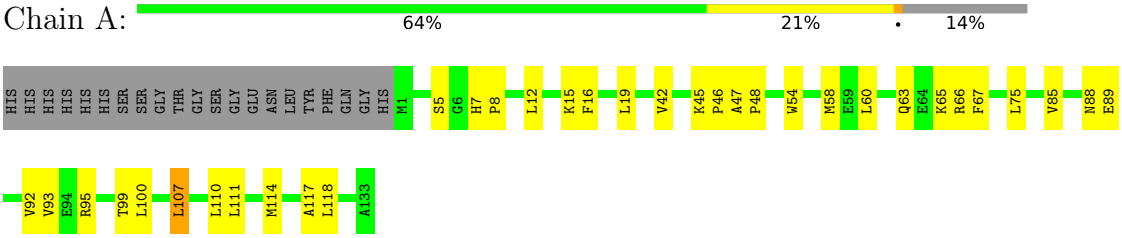
Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	HIS	-	expression tag	UNP Q7CR52
A	-19	HIS	-	expression tag	UNP Q7CR52
A	-18	HIS	-	expression tag	UNP Q7CR52
A	-17	HIS	-	expression tag	UNP Q7CR52
A	-16	HIS	-	expression tag	UNP Q7CR52
A	-15	HIS	-	expression tag	UNP Q7CR52
A	-14	SER	-	expression tag	UNP Q7CR52
A	-13	SER	-	expression tag	UNP Q7CR52
A	-12	GLY	-	expression tag	UNP Q7CR52
A	-11	THR	-	expression tag	UNP Q7CR52
A	-10	GLY	-	expression tag	UNP Q7CR52
A	-9	SER	-	expression tag	UNP Q7CR52
A	-8	GLY	-	expression tag	UNP Q7CR52
A	-7	GLU	-	expression tag	UNP Q7CR52
A	-6	ASN	-	expression tag	UNP Q7CR52
A	-5	LEU	-	expression tag	UNP Q7CR52
A	-4	TYR	-	expression tag	UNP Q7CR52
A	-3	PHE	-	expression tag	UNP Q7CR52
A	-2	GLN	-	expression tag	UNP Q7CR52
A	-1	GLY	-	expression tag	UNP Q7CR52
A	0	HIS	-	expression tag	UNP Q7CR52

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.

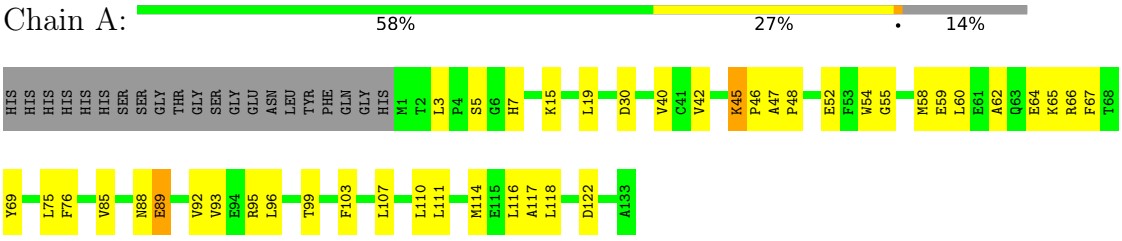
• Molecule 1: Sigma factor-binding protein Crl



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

The representative model is number 3. Colouring as in section 4.1 above.

• Molecule 1: Sigma factor-binding protein Crl



5 Refinement protocol and experimental data overview

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

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
CYANA	refinement	

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	1626
Number of shifts mapped to atoms	1466
Number of unparsed shifts	0
Number of shifts with mapping errors	160
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	77%

6 Model quality

6.1 Standard geometry

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

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	1114	1081	1083	20±3
All	All	22280	21620	21660	396

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 9.

5 of 130 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:19:LEU:HD11	1:A:110:LEU:HD22	0.89	1.42	3	20
1:A:75:LEU:HD13	1:A:85:VAL:HG11	0.79	1.55	19	3
1:A:75:LEU:HD13	1:A:85:VAL:HG21	0.77	1.55	6	5
1:A:111:LEU:HD22	1:A:118:LEU:HD12	0.71	1.62	6	1
1:A:12:LEU:HD12	1:A:16:PHE:CZ	0.70	2.20	11	8

6.3 Torsion angles

6.3.1 Protein backbone

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	131/154 (85%)	120±3 (92±2%)	9±3 (7±2%)	2±1 (1±1%)	16 63
All	All	2620/3080 (85%)	2408 (92%)	179 (7%)	33 (1%)	16 63

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	89	GLU	16
1	A	88	ASN	9
1	A	43	ASN	2
1	A	44	VAL	2
1	A	29	GLU	1

6.3.2 Protein sidechains (i)

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	121/138 (88%)	105±2 (87±2%)	16±2 (13±2%)	7 48
All	All	2420/2760 (88%)	2096 (87%)	324 (13%)	7 48

5 of 72 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	93	VAL	20
1	A	107	LEU	20
1	A	7	HIS	19
1	A	65	LYS	18
1	A	5	SER	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 77% for the well-defined parts and 77% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list*

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	1626
Number of shifts mapped to atoms	1466
Number of unparsed shifts	0
Number of shifts with mapping errors	160
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	3

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 160) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	-16	HIS	HA	4.59	0.02	1
1	A	-16	HIS	HB2	3.03	0.02	2
1	A	-16	HIS	HB3	3.08	0.02	2
1	A	-16	HIS	HD2	6.91	0.02	1
1	A	-16	HIS	CA	56.5	0.25	1
1	A	-16	HIS	CB	31.05	0.25	1
1	A	-15	HIS	HA	4.54	0.02	1
1	A	-15	HIS	HB2	2.94	0.02	2
1	A	-15	HIS	HB3	2.97	0.02	2
1	A	-15	HIS	HD2	6.82	0.02	1
1	A	-15	HIS	HE1	7.65	0.02	1
1	A	-15	HIS	CA	56.37	0.25	1
1	A	-15	HIS	CB	30.98	0.25	1
1	A	-15	HIS	CD2	119.28	0.25	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	-15	HIS	CE1	138.45	0.25	1
1	A	-14	SER	HA	4.44	0.02	1
1	A	-14	SER	HB2	3.79	0.02	2
1	A	-14	SER	HB3	3.82	0.02	2
1	A	-14	SER	CA	58.26	0.25	1
1	A	-14	SER	CB	63.45	0.25	1
1	A	-13	SER	H	8.57	0.02	1
1	A	-13	SER	HA	4.52	0.02	1
1	A	-13	SER	HB2	3.87	0.02	2
1	A	-13	SER	HB3	3.84	0.02	2
1	A	-13	SER	CA	58.06	0.25	1
1	A	-13	SER	CB	63.69	0.25	1
1	A	-13	SER	N	118.69	0.15	1
1	A	-12	GLY	H	8.53	0.02	1
1	A	-12	GLY	HA2	4.08	0.02	1
1	A	-12	GLY	HA3	4.08	0.02	1
1	A	-12	GLY	CA	45.26	0.25	1
1	A	-12	GLY	N	111.39	0.15	1
1	A	-11	THR	H	8.23	0.02	1
1	A	-11	THR	HA	4.38	0.02	1
1	A	-11	THR	HB	4.29	0.02	1
1	A	-11	THR	HG21	1.2	0.02	1
1	A	-11	THR	HG22	1.2	0.02	1
1	A	-11	THR	HG23	1.2	0.02	1
1	A	-11	THR	CA	61.78	0.25	1
1	A	-11	THR	CB	69.61	0.25	1
1	A	-11	THR	CG2	21.39	0.25	1
1	A	-11	THR	N	113.67	0.15	1
1	A	-10	GLY	H	8.59	0.02	1
1	A	-10	GLY	HA2	4.05	0.02	1
1	A	-10	GLY	HA3	4.05	0.02	1
1	A	-10	GLY	CA	45.21	0.25	1
1	A	-10	GLY	N	112.16	0.15	1
1	A	-9	SER	H	8.38	0.02	1
1	A	-9	SER	HA	4.49	0.02	1
1	A	-9	SER	HB2	3.89	0.02	2
1	A	-9	SER	HB3	3.92	0.02	2
1	A	-9	SER	CA	58.38	0.25	1
1	A	-9	SER	CB	63.62	0.25	1
1	A	-9	SER	N	116.39	0.15	1
1	A	-8	GLY	H	8.59	0.02	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	-8	GLY	HA2	3.98	0.02	2
1	A	-8	GLY	HA3	4.0	0.02	2
1	A	-8	GLY	C	174.2	0.25	1
1	A	-8	GLY	CA	45.34	0.25	1
1	A	-8	GLY	N	111.7	0.15	1
1	A	-7	GLU	H	8.33	0.02	1
1	A	-7	GLU	HA	4.26	0.02	1
1	A	-7	GLU	HB2	2.04	0.02	2
1	A	-7	GLU	HB3	1.92	0.02	2
1	A	-7	GLU	HG2	2.24	0.02	2
1	A	-7	GLU	HG3	2.22	0.02	2
1	A	-7	GLU	C	176.14	0.25	1
1	A	-7	GLU	CA	56.55	0.25	1
1	A	-7	GLU	CB	30.16	0.25	1
1	A	-7	GLU	CG	36.15	0.25	1
1	A	-7	GLU	N	121.11	0.15	1
1	A	-6	ASN	H	8.53	0.02	1
1	A	-6	ASN	HA	4.66	0.02	1
1	A	-6	ASN	HB2	2.79	0.02	2
1	A	-6	ASN	HB3	2.72	0.02	2
1	A	-6	ASN	HD21	6.86	0.02	1
1	A	-6	ASN	HD22	7.54	0.02	1
1	A	-6	ASN	C	174.93	0.25	1
1	A	-6	ASN	CA	53.21	0.25	1
1	A	-6	ASN	CB	38.55	0.25	1
1	A	-6	ASN	N	119.88	0.15	1
1	A	-6	ASN	ND2	113.06	0.15	1
1	A	-5	LEU	H	8.17	0.02	1
1	A	-5	LEU	HA	4.23	0.02	1
1	A	-5	LEU	HB2	1.47	0.02	2
1	A	-5	LEU	HB3	1.33	0.02	2
1	A	-5	LEU	HG	1.46	0.02	1
1	A	-5	LEU	HD11	0.79	0.02	2
1	A	-5	LEU	HD12	0.79	0.02	2
1	A	-5	LEU	HD13	0.79	0.02	2
1	A	-5	LEU	HD21	0.86	0.02	2
1	A	-5	LEU	HD22	0.86	0.02	2
1	A	-5	LEU	HD23	0.86	0.02	2
1	A	-5	LEU	C	176.97	0.25	1
1	A	-5	LEU	CA	55.33	0.25	1
1	A	-5	LEU	CB	42.12	0.25	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	-5	LEU	CG	26.73	0.25	1
1	A	-5	LEU	CD1	23.24	0.25	2
1	A	-5	LEU	CD2	24.78	0.25	2
1	A	-5	LEU	N	122.88	0.15	1
1	A	-4	TYR	H	8.12	0.02	1
1	A	-4	TYR	HA	4.56	0.02	1
1	A	-4	TYR	HB2	2.98	0.02	2
1	A	-4	TYR	HB3	2.84	0.02	2
1	A	-4	TYR	HD1	7.04	0.02	3
1	A	-4	TYR	HD2	7.04	0.02	3
1	A	-4	TYR	HE1	6.79	0.02	3
1	A	-4	TYR	HE2	6.79	0.02	3
1	A	-4	TYR	C	175.63	0.25	1
1	A	-4	TYR	CA	57.67	0.25	1
1	A	-4	TYR	CB	38.72	0.25	1
1	A	-4	TYR	CD1	133.18	0.25	3
1	A	-4	TYR	CD2	133.18	0.25	3
1	A	-4	TYR	CE1	118.01	0.25	3
1	A	-4	TYR	CE2	118.01	0.25	3
1	A	-4	TYR	N	120.35	0.15	1
1	A	-3	PHE	H	8.22	0.02	1
1	A	-3	PHE	HA	4.54	0.02	1
1	A	-3	PHE	HB2	3.06	0.02	2
1	A	-3	PHE	HB3	3.06	0.02	2
1	A	-3	PHE	HD1	7.24	0.02	3
1	A	-3	PHE	HD2	7.24	0.02	3
1	A	-3	PHE	HE1	7.32	0.02	3
1	A	-3	PHE	HE2	7.32	0.02	3
1	A	-3	PHE	HZ	7.26	0.02	1
1	A	-3	PHE	C	174.44	0.25	1
1	A	-3	PHE	CA	57.81	0.25	1
1	A	-3	PHE	CB	39.39	0.25	1
1	A	-3	PHE	CD1	131.68	0.25	3
1	A	-3	PHE	CD2	131.68	0.25	3
1	A	-3	PHE	CE1	131.51	0.25	3
1	A	-3	PHE	CE2	131.51	0.25	3
1	A	-3	PHE	CZ	129.94	0.25	1
1	A	-3	PHE	N	122.65	0.15	1
1	A	-2	GLN	H	8.34	0.02	1
1	A	-2	GLN	HA	4.18	0.02	1
1	A	-2	GLN	HB2	2.1	0.02	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	-2	GLN	HB3	1.88	0.02	2
1	A	-2	GLN	HG2	2.13	0.02	2
1	A	-2	GLN	HG3	2.25	0.02	2
1	A	-2	GLN	CA	55.89	0.25	1
1	A	-2	GLN	CB	28.81	0.25	1
1	A	-2	GLN	CG	33.6	0.25	1
1	A	-2	GLN	N	122.49	0.15	1
1	A	-1	GLY	H	7.91	0.02	1
1	A	-1	GLY	HA2	3.94	0.02	2
1	A	-1	GLY	HA3	3.95	0.02	2
1	A	-1	GLY	CA	45.3	0.25	1
1	A	-1	GLY	N	108.87	0.15	1
1	A	0	HIS	H	8.29	0.02	1
1	A	0	HIS	HA	4.51	0.02	1
1	A	0	HIS	HB2	2.91	0.02	2
1	A	0	HIS	HB3	3.0	0.02	2
1	A	0	HIS	HD2	6.81	0.02	1
1	A	0	HIS	HE1	7.73	0.02	1
1	A	0	HIS	CA	56.3	0.25	1
1	A	0	HIS	CB	31.04	0.25	1
1	A	0	HIS	CD2	119.4	0.25	1
1	A	0	HIS	CE1	138.57	0.25	1
1	A	0	HIS	N	119.0	0.15	1

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	146	-0.16 ± 0.22	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	126	0.38 ± 0.22	None needed (< 0.5 ppm)
$^{13}C'$	119	0.00 ± 0.19	None needed (< 0.5 ppm)
^{15}N	121	-0.37 ± 0.51	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 77%, i.e. 1466 atoms were assigned a chemical shift out of a possible 1904. 0 out of 22 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	588/655 (90%)	239/264 (91%)	242/266 (91%)	107/125 (86%)
Sidechain	787/1051 (75%)	528/676 (78%)	254/331 (77%)	5/44 (11%)
Aromatic	91/198 (46%)	68/98 (69%)	19/94 (20%)	4/6 (67%)
Overall	1466/1904 (77%)	835/1038 (80%)	515/691 (75%)	116/175 (66%)

7.1.4 Statistically unusual chemical shifts (i)

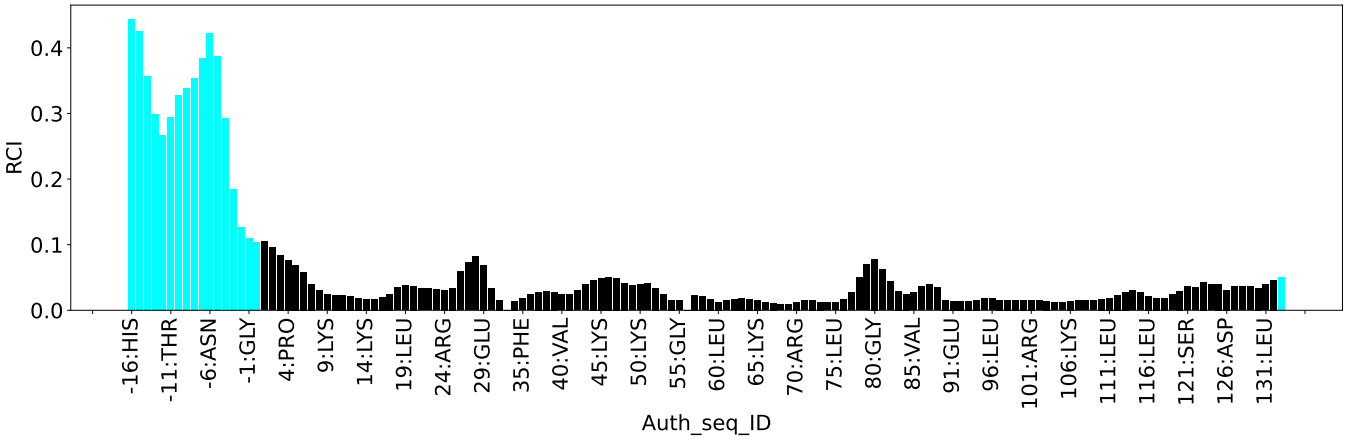
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	100	LEU	HD11	-0.97	-0.61 – 2.12	-6.3
1	A	100	LEU	HD12	-0.97	-0.61 – 2.12	-6.3
1	A	100	LEU	HD13	-0.97	-0.61 – 2.12	-6.3

7.1.5 Random Coil Index (RCI) plots (i)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:



8 NMR restraints analysis

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	1519
Intra-residue (|i-j|=0)	399
Sequential (|i-j|=1)	380
Medium range (|i-j|>1 and |i-j|<5)	236
Long range (|i-j|≥5)	504
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	9.9
Number of long range restraints per residue[1]	3.3

[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)	1.8	0.2
0.2-0.5 (Medium)	0.5	0.38
>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)

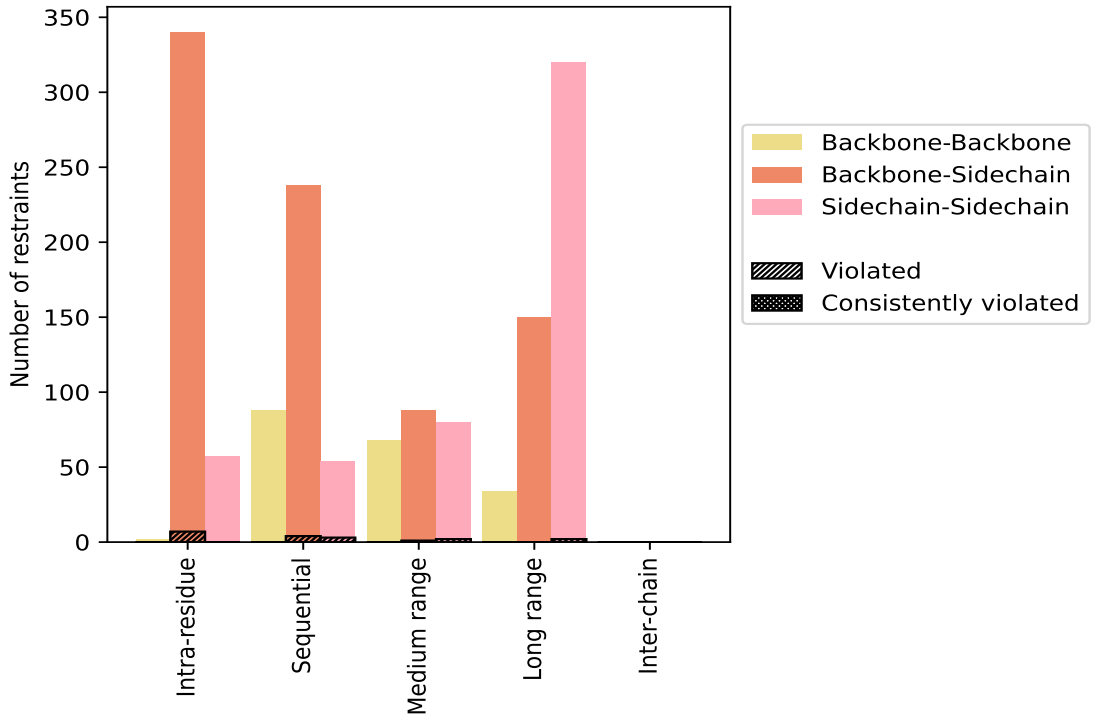
9.1 Summary of distance violations (i)

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)	399	26.3	7	1.8	0.5	0	0.0	0.0
Backbone-Backbone	2	0.1	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	340	22.4	7	2.1	0.5	0	0.0	0.0
Sidechain-Sidechain	57	3.8	0	0.0	0.0	0	0.0	0.0
Sequential (|i-j|=1)	380	25.0	7	1.8	0.5	0	0.0	0.0
Backbone-Backbone	88	5.8	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	238	15.7	4	1.7	0.3	0	0.0	0.0
Sidechain-Sidechain	54	3.6	3	5.6	0.2	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	236	15.5	3	1.3	0.2	0	0.0	0.0
Backbone-Backbone	68	4.5	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	88	5.8	1	1.1	0.1	0	0.0	0.0
Sidechain-Sidechain	80	5.3	2	2.5	0.1	0	0.0	0.0
Long range (|i-j|≥5)	504	33.2	2	0.4	0.1	0	0.0	0.0
Backbone-Backbone	34	2.2	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	150	9.9	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	320	21.1	2	0.6	0.1	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	1519	100.0	19	1.3	1.3	0	0.0	0.0
Backbone-Backbone	192	12.6	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	816	53.7	12	1.5	0.8	0	0.0	0.0
Sidechain-Sidechain	511	33.6	7	1.4	0.5	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	1	0	0	0	0	1	0.21	0.21	0.0	0.21
2	0	0	0	0	0	0	0.0	0.0	0.0	0.0
3	2	0	0	0	0	2	0.18	0.2	0.02	0.18
4	0	3	0	0	0	3	0.22	0.35	0.1	0.19
5	0	0	0	0	0	0	0.0	0.0	0.0	0.0
6	2	1	0	0	0	3	0.17	0.19	0.02	0.16
7	0	3	0	0	0	3	0.22	0.36	0.1	0.2
8	0	2	0	0	0	2	0.16	0.17	0.01	0.16
9	2	0	1	0	0	3	0.14	0.16	0.02	0.13
10	0	0	0	0	0	0	0.0	0.0	0.0	0.0
11	0	0	2	0	0	2	0.12	0.13	0.01	0.12

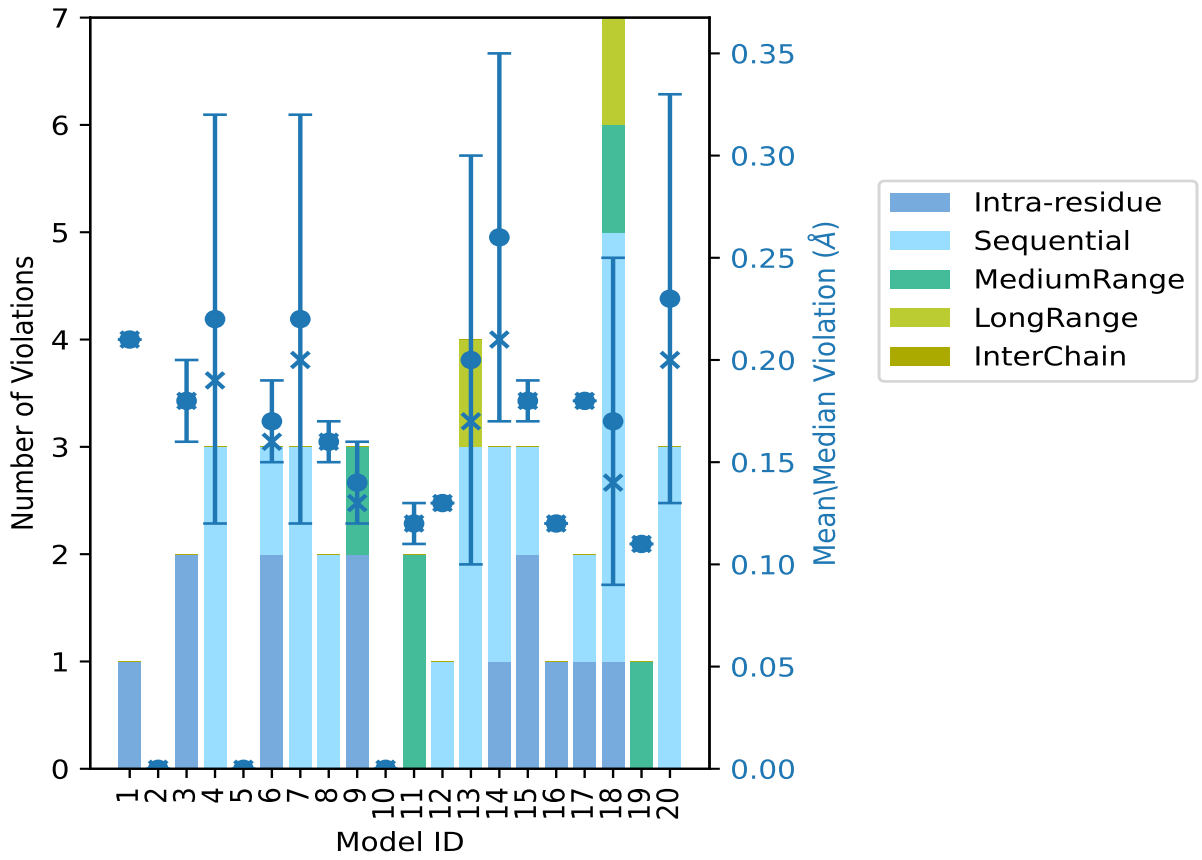
Continued on next page...

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	1	0	0	0	1	0.13	0.13	0.0	0.13
13	0	3	0	1	0	4	0.2	0.36	0.1	0.17
14	1	2	0	0	0	3	0.26	0.38	0.09	0.21
15	2	1	0	0	0	3	0.18	0.2	0.01	0.18
16	1	0	0	0	0	1	0.12	0.12	0.0	0.12
17	1	1	0	0	0	2	0.18	0.18	0.0	0.18
18	1	4	1	1	0	7	0.17	0.36	0.08	0.14
19	0	0	1	0	0	1	0.11	0.11	0.0	0.11
20	0	3	0	0	0	3	0.23	0.36	0.1	0.2

[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

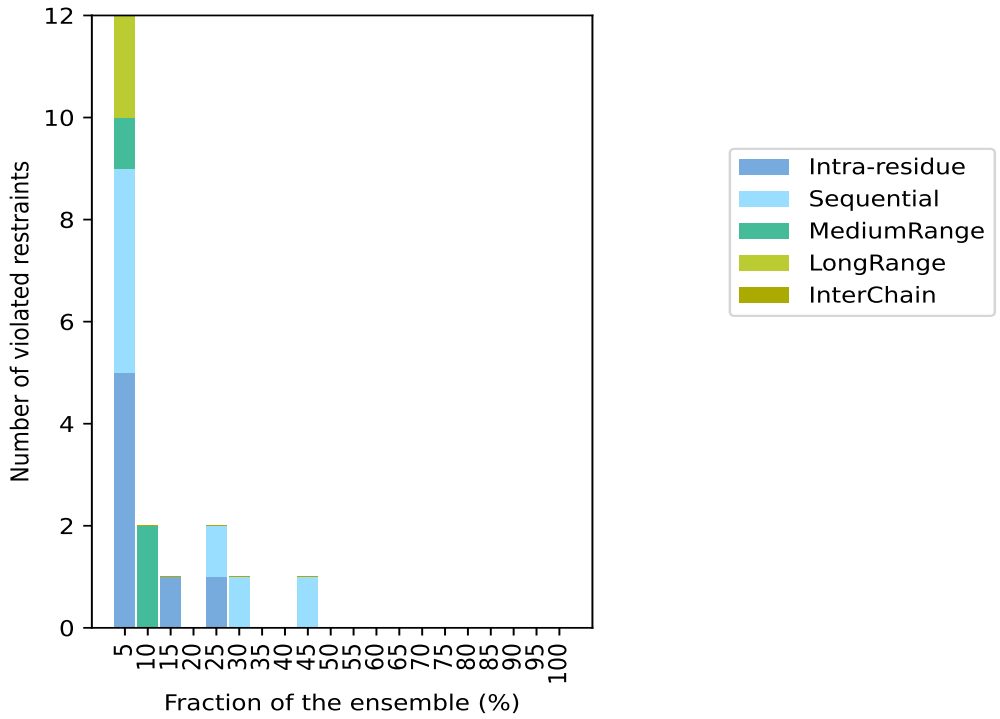
9.3 Distance violation statistics for the ensemble (i)

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, 1500(IR:392, SQ:373, MR:233, LR:502, 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]	%
5	4	1	2	0	12	1	5.0
0	0	2	0	0	2	2	10.0
1	0	0	0	0	1	3	15.0
0	0	0	0	0	0	4	20.0
1	1	0	0	0	2	5	25.0
0	1	0	0	0	1	6	30.0
0	0	0	0	0	0	7	35.0
0	0	0	0	0	0	8	40.0
0	1	0	0	0	1	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

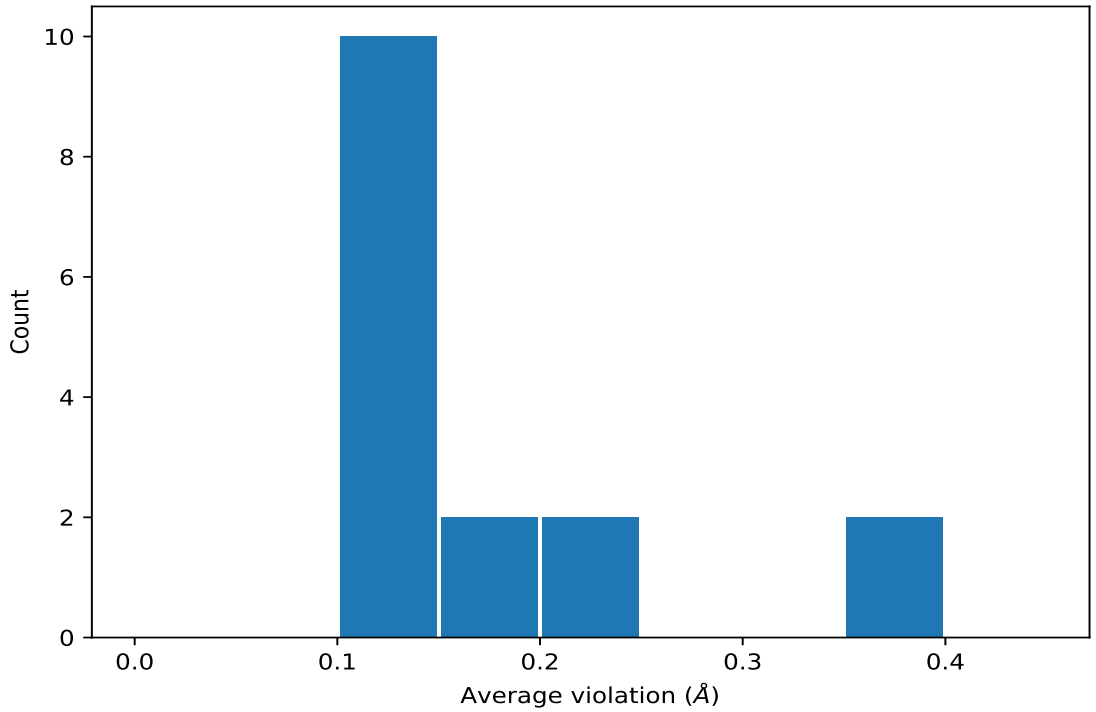
9.3.1 Bar graph : Distance violation statistics for the ensemble



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 (i)

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.

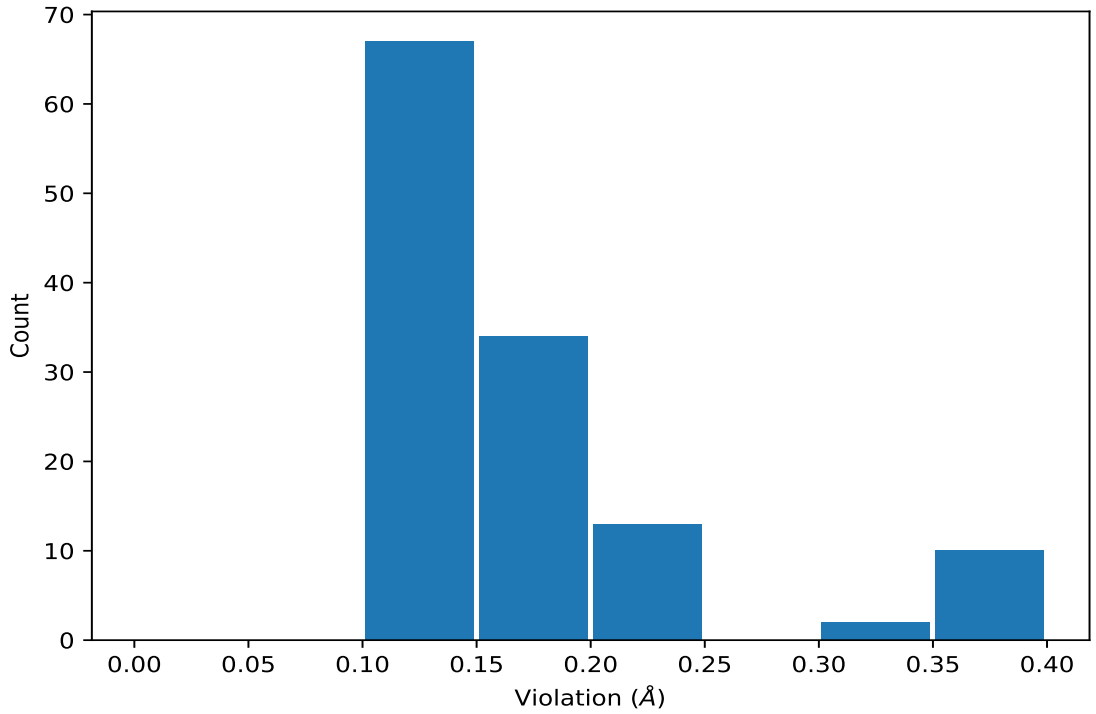
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,1239)	1:A:64:GLU:HB3	1:A:65:LYS:HG2	9	0.19	0.02	0.19
(1,1239)	1:A:64:GLU:HB3	1:A:65:LYS:HG3	9	0.19	0.02	0.19
(1,1246)	1:A:65:LYS:HG2	1:A:66:ARG:H	6	0.36	0.01	0.36
(1,1246)	1:A:65:LYS:HG3	1:A:66:ARG:H	6	0.36	0.01	0.36
(1,1226)	1:A:63:GLN:H	1:A:63:GLN:HG2	5	0.2	0.01	0.2
(1,1226)	1:A:63:GLN:H	1:A:63:GLN:HG3	5	0.2	0.01	0.2
(1,1247)	1:A:65:LYS:HD2	1:A:66:ARG:HA	5	0.11	0.0	0.11
(1,1247)	1:A:65:LYS:HD3	1:A:66:ARG:HA	5	0.11	0.0	0.11
(1,525)	1:A:79:GLU:HA	1:A:79:GLU:HG2	3	0.15	0.01	0.15
(1,1322)	1:A:79:GLU:HB2	1:A:81:ASN:HD21	2	0.14	0.01	0.14
(1,1322)	1:A:79:GLU:HB2	1:A:81:ASN:HD22	2	0.14	0.01	0.14
(1,1322)	1:A:79:GLU:HB3	1:A:81:ASN:HD21	2	0.14	0.01	0.14
(1,1322)	1:A:79:GLU:HB3	1:A:81:ASN:HD22	2	0.14	0.01	0.14
(1,910)	1:A:109:ASP:HA	1:A:112:ILE:HG21	2	0.12	0.01	0.12
(1,910)	1:A:109:ASP:HA	1:A:112:ILE:HG22	2	0.12	0.01	0.12
(1,910)	1:A:109:ASP:HA	1:A:112:ILE:HG23	2	0.12	0.01	0.12

[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)
(1,1246)	1:A:65:LYS:HG2	1:A:66:ARG:H	14	0.38
(1,1246)	1:A:65:LYS:HG3	1:A:66:ARG:H	14	0.38
(1,1246)	1:A:65:LYS:HG2	1:A:66:ARG:H	7	0.36
(1,1246)	1:A:65:LYS:HG3	1:A:66:ARG:H	7	0.36
(1,1246)	1:A:65:LYS:HG2	1:A:66:ARG:H	13	0.36
(1,1246)	1:A:65:LYS:HG3	1:A:66:ARG:H	13	0.36
(1,1246)	1:A:65:LYS:HG2	1:A:66:ARG:H	18	0.36
(1,1246)	1:A:65:LYS:HG3	1:A:66:ARG:H	18	0.36

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,1246)	1:A:65:LYS:HG2	1:A:66:ARG:H	20	0.36
(1,1246)	1:A:65:LYS:HG3	1:A:66:ARG:H	20	0.36
(1,1246)	1:A:65:LYS:HG2	1:A:66:ARG:H	4	0.35
(1,1246)	1:A:65:LYS:HG3	1:A:66:ARG:H	4	0.35
(1,1239)	1:A:64:GLU:HB3	1:A:65:LYS:HG2	13	0.22
(1,1239)	1:A:64:GLU:HB3	1:A:65:LYS:HG3	13	0.22
(1,802)	1:A:7:HIS:HA	1:A:7:HIS:HD2	1	0.21
(1,1226)	1:A:63:GLN:H	1:A:63:GLN:HG2	14	0.21
(1,1226)	1:A:63:GLN:H	1:A:63:GLN:HG3	14	0.21
(1,1239)	1:A:64:GLU:HB3	1:A:65:LYS:HG2	7	0.2

10 Dihedral-angle violation analysis (i)

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