wwPDB NMR Structure Validation Summary Report

Jun 4, 2023 – 10:26 PM EDT

PDB ID	:	2LSJ
BMRB ID	:	18433
Title	:	Solution structure of the mouse Rev1 CTD in complex with the Rev1-interacting Region (RIR)of Pol Kappa
Authors	:	Liu, J.; Wojtaszek, J.; Zhou, P.
Deposited on	:	2012-05-01

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

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

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

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

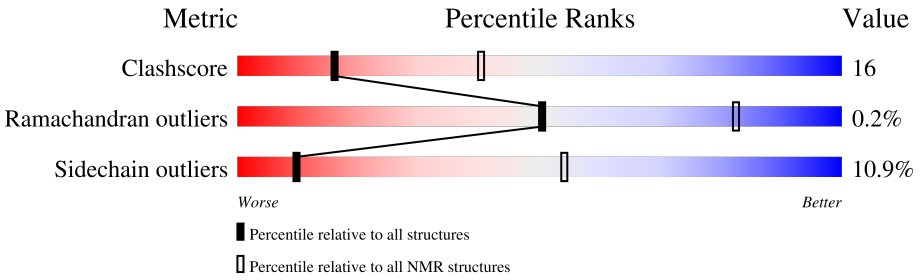
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal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

1 Overall quality at a glance

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

The overall completeness of chemical shifts assignment is 87%.

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



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

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

Mol	Chain	Length	Quality of chain
1	A	119	48% 28% • • 18%
2	B	25	20% 8% 28% 44%

2 Ensemble composition and analysis

This entry contains 20 models. Model 12 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:25-A:116, B:208-B:214 (99)	0.26	12

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. No single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 1822 atoms, of which 919 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called DNA repair protein REV1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	97	1575	502	794	121	153	5	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP Q920Q2
A	2	SER	-	expression tag	UNP Q920Q2
A	3	GLY	-	expression tag	UNP Q920Q2
A	4	GLY	-	expression tag	UNP Q920Q2

- Molecule 2 is a protein called DNA polymerase kappa.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
2	B	14	247	76	125	24	22	0

There are 2 discrepancies between the modelled and reference sequences:

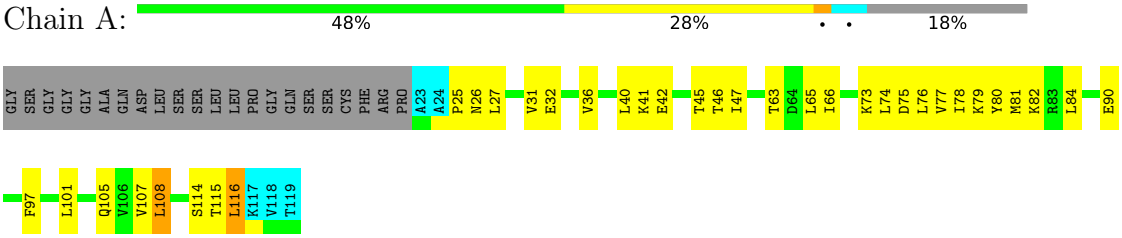
Chain	Residue	Modelled	Actual	Comment	Reference
B	201	SER	-	expression tag	UNP Q9QUG2
B	202	HIS	-	expression tag	UNP Q9QUG2

4 Residue-property plots

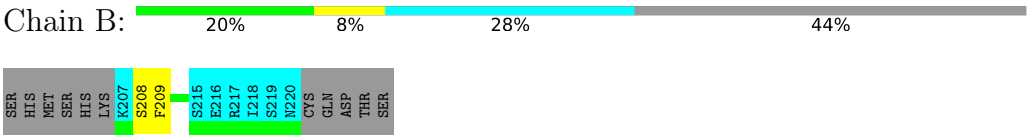
4.1 Average score per residue in the NMR ensemble

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 repair protein REV1



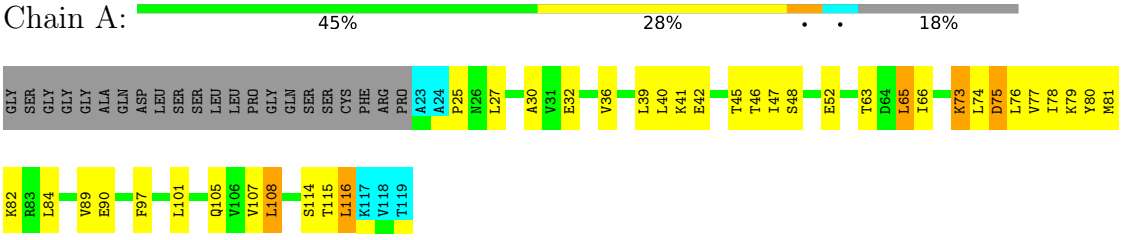
• Molecule 2: DNA polymerase kappa



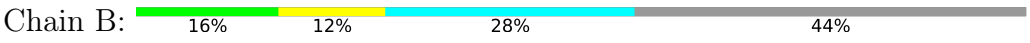
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

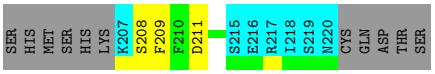
The representative model is number 12. Colouring as in section 4.1 above.

• Molecule 1: DNA repair protein REV1



• Molecule 2: DNA polymerase kappa




SER
HIS
MET
SER
HIS
LYS
K207
S208
F209
F210
D211
S215
E216
R217
I218
S219
N220
CYS
GLN
ASP
THR
SER

5 Refinement protocol and experimental data overview

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

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *target function.*

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

Software name	Classification	Version
CYANA	structure solution	
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	1552
Number of shifts mapped to atoms	1357
Number of unparsed shifts	0
Number of shifts with mapping errors	195
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	87%

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	748	755	755	26±3
2	B	65	66	66	1±1
All	All	16260	16420	16420	523

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 16.

5 of 70 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:78:ILE:HG21	1:A:116:LEU:HD22	0.78	1.54	2	16
1:A:97:PHE:CZ	1:A:101:LEU:HD11	0.71	2.21	7	20
1:A:42:GLU:O	1:A:46:THR:HG22	0.68	1.89	9	20
1:A:63:THR:HG22	1:A:107:VAL:HG11	0.67	1.66	9	15
1:A:28:ALA:HB3	1:A:39:LEU:HD13	0.67	1.67	9	2

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	92/119 (77%)	89±1 (97±1%)	3±1 (3±1%)	0±0 (0±0%)	50 82
2	B	7/25 (28%)	6±0 (89±6%)	1±0 (11±6%)	0±0 (0±0%)	100 100
All	All	1980/2880 (69%)	1906 (96%)	70 (4%)	4 (0%)	50 82

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
1	A	88	SER	4

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	87/106 (82%)	78±2 (90±2%)	9±2 (10±2%)	11 57
2	B	7/25 (28%)	5±1 (77±11%)	2±1 (23±11%)	3 28
All	All	1880/2620 (72%)	1675 (89%)	205 (11%)	10 54

5 of 31 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	79	LYS	20
1	A	82	LYS	20
1	A	108	LEU	20
1	A	116	LEU	20
2	B	208	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 87% for the well-defined parts and 86% 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	1552
Number of shifts mapped to atoms	1357
Number of unparsed shifts	0
Number of shifts with mapping errors	195
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 195) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	5	GLY	H	8.286	0.006	.
1	A	5	GLY	HA2	3.962	0.004	.
1	A	5	GLY	HA3	3.962	0.004	.
1	A	5	GLY	CA	45.182	0.016	.
1	A	5	GLY	N	108.925	0.041	.
1	A	6	ALA	H	8.189	0.006	.
1	A	6	ALA	HA	4.28	0.005	.
1	A	6	ALA	HB1	1.368	0.007	.
1	A	6	ALA	HB2	1.368	0.007	.
1	A	6	ALA	HB3	1.368	0.007	.
1	A	6	ALA	CA	52.52	0.141	.
1	A	6	ALA	CB	19.156	0.036	.
1	A	6	ALA	N	123.712	0.024	.
1	A	7	GLN	H	8.298	0.005	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	7	GLN	HA	4.297	0.005	.
1	A	7	GLN	HB2	1.964	0.000	.
1	A	7	GLN	HB3	2.097	0.000	.
1	A	7	GLN	HG2	2.368	0.006	.
1	A	7	GLN	HG3	2.368	0.006	.
1	A	7	GLN	CA	55.798	0.061	.
1	A	7	GLN	CB	29.678	0.022	.
1	A	7	GLN	CG	36.13	0.000	.
1	A	7	GLN	N	118.944	0.023	.
1	A	8	ASP	H	8.238	0.005	.
1	A	8	ASP	HA	4.598	0.004	.
1	A	8	ASP	HB2	2.605	0.007	.
1	A	8	ASP	HB3	2.736	0.002	.
1	A	8	ASP	CA	54.144	0.079	.
1	A	8	ASP	CB	41.233	0.024	.
1	A	8	ASP	N	121.339	0.007	.
1	A	9	LEU	H	8.268	0.005	.
1	A	9	LEU	HA	4.294	0.007	.
1	A	9	LEU	HB2	1.621	0.009	.
1	A	9	LEU	HB3	1.621	0.009	.
1	A	9	LEU	CA	55.65	0.058	.
1	A	9	LEU	CB	41.984	0.115	.
1	A	9	LEU	N	123.5	0.010	.
1	A	10	SER	H	8.301	0.005	.
1	A	10	SER	HA	4.318	0.008	.
1	A	10	SER	HB2	3.887	0.006	.
1	A	10	SER	HB3	3.887	0.006	.
1	A	10	SER	CA	59.535	0.114	.
1	A	10	SER	CB	63.309	0.001	.
1	A	10	SER	N	115.706	0.006	.
1	A	11	SER	H	8.019	0.005	.
1	A	11	SER	HA	4.413	0.007	.
1	A	11	SER	HB2	3.854	0.001	.
1	A	11	SER	HB3	3.854	0.001	.
1	A	11	SER	CA	58.523	0.063	.
1	A	11	SER	CB	63.577	0.000	.
1	A	11	SER	N	116.763	0.013	.
1	A	12	LEU	H	7.862	0.003	.
1	A	12	LEU	HA	4.316	0.010	.
1	A	12	LEU	HB2	1.561	0.007	.
1	A	12	LEU	HB3	1.606	0.005	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	12	LEU	HG	1.611	0.009	.
1	A	12	LEU	HD11	0.82	0.009	.
1	A	12	LEU	HD12	0.82	0.009	.
1	A	12	LEU	HD13	0.82	0.009	.
1	A	12	LEU	HD21	0.882	0.004	.
1	A	12	LEU	HD22	0.882	0.004	.
1	A	12	LEU	HD23	0.882	0.004	.
1	A	12	LEU	CA	55.265	0.120	.
1	A	12	LEU	CB	42.244	0.061	.
1	A	12	LEU	CG	26.891	0.132	.
1	A	12	LEU	CD1	23.34	0.068	.
1	A	12	LEU	CD2	24.972	0.108	.
1	A	12	LEU	N	122.926	0.018	.
1	A	13	LEU	H	7.98	0.003	.
1	A	13	LEU	HA	4.587	0.006	.
1	A	13	LEU	HB2	1.528	0.007	.
1	A	13	LEU	HB3	1.566	0.002	.
1	A	13	LEU	HG	1.642	0.005	.
1	A	13	LEU	HD11	0.874	0.006	.
1	A	13	LEU	HD12	0.874	0.006	.
1	A	13	LEU	HD13	0.874	0.006	.
1	A	13	LEU	HD21	0.898	0.009	.
1	A	13	LEU	HD22	0.898	0.009	.
1	A	13	LEU	HD23	0.898	0.009	.
1	A	13	LEU	CA	52.882	0.046	.
1	A	13	LEU	CB	41.508	0.068	.
1	A	13	LEU	CG	26.921	0.044	.
1	A	13	LEU	CD1	23.22	0.051	.
1	A	13	LEU	CD2	25.262	0.048	.
1	A	13	LEU	N	123.305	0.009	.
1	A	14	PRO	HA	4.366	0.006	.
1	A	14	PRO	HB2	1.9	0.004	.
1	A	14	PRO	HB3	2.265	0.004	.
1	A	14	PRO	HG2	1.988	0.003	.
1	A	14	PRO	HG3	2.046	0.001	.
1	A	14	PRO	HD2	3.61	0.011	.
1	A	14	PRO	HD3	3.8	0.006	.
1	A	14	PRO	CA	63.559	0.079	.
1	A	14	PRO	CB	31.877	0.066	.
1	A	14	PRO	CG	27.511	0.035	.
1	A	14	PRO	CD	50.531	0.028	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	15	GLY	H	8.437	0.006	.
1	A	15	GLY	HA2	3.906	0.006	.
1	A	15	GLY	HA3	3.906	0.006	.
1	A	15	GLY	CA	45.382	0.018	.
1	A	15	GLY	N	109.337	0.019	.
1	A	16	GLN	H	8.077	0.006	.
1	A	16	GLN	HA	4.371	0.010	.
1	A	16	GLN	HB2	2.006	0.015	.
1	A	16	GLN	HB3	2.122	0.007	.
1	A	16	GLN	HG2	2.326	0.006	.
1	A	16	GLN	HG3	2.326	0.006	.
1	A	16	GLN	CA	55.896	0.041	.
1	A	16	GLN	CB	29.474	0.019	.
1	A	16	GLN	CG	33.977	0.076	.
1	A	16	GLN	N	119.531	0.015	.
1	A	19	CYS	HA	4.442	0.003	.
1	A	19	CYS	HB2	2.805	0.003	.
1	A	19	CYS	HB3	2.805	0.003	.
1	A	19	CYS	CA	58.332	0.045	.
1	A	19	CYS	CB	27.88	0.014	.
1	A	20	PHE	H	8.169	0.005	.
1	A	20	PHE	HA	4.59	0.003	.
1	A	20	PHE	HB2	3.011	0.010	.
1	A	20	PHE	HB3	3.051	0.008	.
1	A	20	PHE	HD1	7.204	0.004	.
1	A	20	PHE	HD2	7.204	0.004	.
1	A	20	PHE	HE1	7.297	0.008	.
1	A	20	PHE	HE2	7.297	0.008	.
1	A	20	PHE	CA	57.643	0.077	.
1	A	20	PHE	CB	39.533	0.064	.
1	A	20	PHE	CD1	131.783	0.045	.
1	A	20	PHE	CE1	131.445	0.145	.
1	A	20	PHE	N	122.585	0.001	.
1	A	21	ARG	H	8.003	0.004	.
1	A	21	ARG	HA	4.544	0.005	.
1	A	21	ARG	HB2	1.624	0.007	.
1	A	21	ARG	HB3	1.745	0.006	.
1	A	21	ARG	HG2	1.555	0.005	.
1	A	21	ARG	HG3	1.555	0.005	.
1	A	21	ARG	HD2	3.153	0.004	.
1	A	21	ARG	HD3	3.153	0.004	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	21	ARG	CA	53.265	0.060	.
1	A	21	ARG	CB	30.684	0.032	.
1	A	21	ARG	CG	26.647	0.022	.
1	A	21	ARG	CD	43.163	0.033	.
1	A	21	ARG	N	124.42	0.019	.
1	A	22	PRO	HA	4.293	0.007	.
1	A	22	PRO	HB2	1.879	0.008	.
1	A	22	PRO	HB3	2.259	0.006	.
1	A	22	PRO	HG2	1.932	0.011	.
1	A	22	PRO	HG3	1.932	0.011	.
1	A	22	PRO	HD2	3.542	0.004	.
1	A	22	PRO	HD3	3.542	0.004	.
1	A	22	PRO	CA	63.011	0.144	.
1	A	22	PRO	CB	31.999	0.076	.
1	A	22	PRO	CG	27.142	0.058	.
1	A	22	PRO	CD	50.512	0.067	.
1	B	205	HIS	HA	4.634	0.005	.
1	B	205	HIS	HB2	3.084	0.005	.
1	B	205	HIS	HB3	3.145	0.005	.
1	B	205	HIS	HD2	7.026	0.013	.
1	B	205	HIS	CA	56.314	0.005	.
1	B	205	HIS	CB	30.814	0.015	.
1	B	205	HIS	CD2	119.637	0.140	.
1	B	221	CYS	H	8.14	0.005	.
1	B	221	CYS	HA	4.463	0.005	.
1	B	221	CYS	HB2	2.92	0.005	.
1	B	221	CYS	HB3	2.92	0.005	.
1	B	221	CYS	CA	58.651	0.129	.
1	B	221	CYS	CB	27.58	0.078	.
1	B	221	CYS	N	118.834	0.032	.
1	B	222	GLN	H	8.348	0.005	.
1	B	222	GLN	HA	4.293	0.000	.
1	B	222	GLN	HB2	1.656	0.000	.
1	B	222	GLN	HB3	1.656	0.000	.
1	B	222	GLN	CA	55.9	0.000	.
1	B	222	GLN	CB	29.134	0.000	.
1	B	222	GLN	N	122.118	0.018	.
1	B	223	ASP	H	8.308	0.005	.
1	B	223	ASP	HA	4.664	0.006	.
1	B	223	ASP	HB2	2.615	0.004	.
1	B	223	ASP	HB3	2.731	0.004	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	B	223	ASP	CA	54.554	0.083	.
1	B	223	ASP	CB	41.257	0.033	.
1	B	223	ASP	N	121.582	0.014	.
1	B	224	THR	H	8.056	0.006	.
1	B	224	THR	HA	4.4	0.004	.
1	B	224	THR	HB	4.314	0.004	.
1	B	224	THR	HG21	1.167	0.005	.
1	B	224	THR	HG22	1.167	0.005	.
1	B	224	THR	HG23	1.167	0.005	.
1	B	224	THR	CA	61.197	0.048	.
1	B	224	THR	CB	69.572	0.025	.
1	B	224	THR	CG2	21.269	0.069	.
1	B	224	THR	N	113.896	0.027	.
1	B	225	SER	H	7.957	0.004	.
1	B	225	SER	HA	4.227	0.002	.
1	B	225	SER	CA	60.168	0.062	.
1	B	225	SER	N	123.678	0.013	.

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	133	-0.60 ± 0.33	None needed (imprecise)
$^{13}C_{\beta}$	128	0.31 ± 0.07	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	121	0.33 ± 0.19	None needed (< 0.5 ppm)

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 87%, i.e. 1227 atoms were assigned a chemical shift out of a possible 1416. 0 out of 22 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	384/493 (78%)	193/198 (97%)	99/198 (50%)	92/97 (95%)
Sidechain	761/822 (93%)	519/534 (97%)	233/262 (89%)	9/26 (35%)
Aromatic	82/101 (81%)	49/49 (100%)	31/50 (62%)	2/2 (100%)
Overall	1227/1416 (87%)	761/781 (97%)	363/510 (71%)	103/125 (82%)

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.

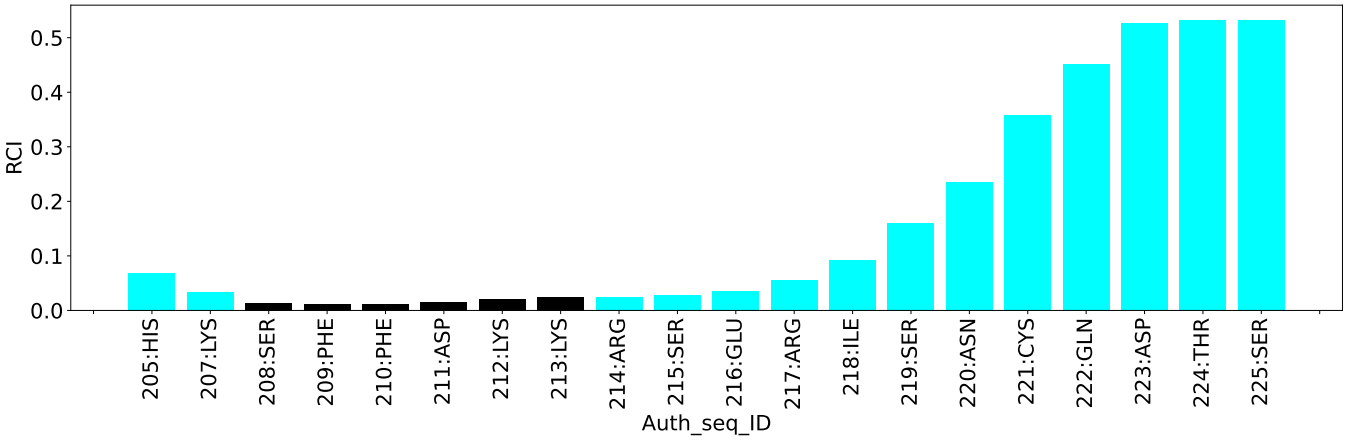
List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	93	TRP	HH2	4.37	5.24 – 8.73	-7.5
1	A	50	PRO	HG2	-0.19	0.41 – 3.45	-7.0
1	A	37	LYS	HG2	0.11	0.13 – 2.61	-5.1

7.1.5 Random Coil Index (RCI) plots (i)

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

Random coil index (RCI) for chain A:

Random coil index (RCI) for chain B:


RCI
0.0
0.1
0.2
0.3
0.4
0.5
205:HIS
207:LYS
208:SER
209:PHE
210:PHE
211:ASP
212:LYS
213:LYS
214:ARG
215:SER
216:GLU
217:ARG
218:ILE
219:SER
220:ASN
221:CYS
222:GLN
223:ASP
224:THR
225:SER
Auth_seq_ID

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	3375
Intra-residue (|i-j|=0)	662
Sequential (|i-j|=1)	709
Medium range (|i-j|>1 and |i-j|<5)	983
Long range (|i-j|≥5)	881
Inter-chain	140
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	85
Number of restraints per residue	23.4
Number of long range restraints per residue[1]	6.1

[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)	4.0	0.2
0.2-0.5 (Medium)	2.4	0.5
>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)

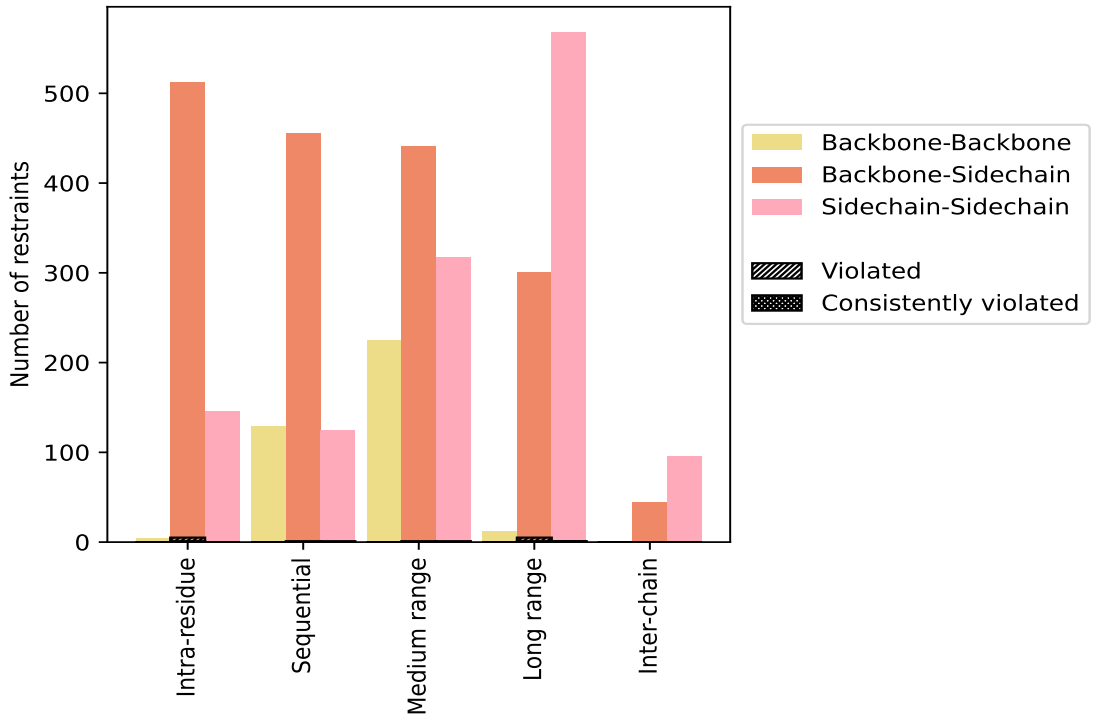
9.1 Summary of distance violations (i)

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)	662	19.6	5	0.8	0.1	1	0.2	0.0
Backbone-Backbone	4	0.1	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	512	15.2	5	1.0	0.1	1	0.2	0.0
Sidechain-Sidechain	146	4.3	0	0.0	0.0	0	0.0	0.0
Sequential (|i-j|=1)	709	21.0	2	0.3	0.1	0	0.0	0.0
Backbone-Backbone	129	3.8	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	455	13.5	1	0.2	0.0	0	0.0	0.0
Sidechain-Sidechain	125	3.7	1	0.8	0.0	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	983	29.1	2	0.2	0.1	0	0.0	0.0
Backbone-Backbone	225	6.7	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	441	13.1	1	0.2	0.0	0	0.0	0.0
Sidechain-Sidechain	317	9.4	1	0.3	0.0	0	0.0	0.0
Long range (|i-j|≥5)	881	26.1	6	0.7	0.2	2	0.2	0.1
Backbone-Backbone	12	0.4	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	301	8.9	5	1.7	0.1	1	0.3	0.0
Sidechain-Sidechain	568	16.8	1	0.2	0.0	1	0.2	0.0
Inter-chain	140	4.1	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	1	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	44	1.3	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	95	2.8	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	3375	100.0	15	0.4	0.4	3	0.1	0.1
Backbone-Backbone	371	11.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	1753	51.9	12	0.7	0.4	2	0.1	0.1
Sidechain-Sidechain	1251	37.1	3	0.2	0.1	1	0.1	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	3	0	4	0.29	0.48	0.16	0.29
2	1	0	0	5	0	6	0.24	0.49	0.16	0.15
3	3	0	0	2	0	5	0.3	0.48	0.12	0.25
4	1	1	1	5	0	8	0.22	0.5	0.14	0.14
5	1	1	1	4	0	7	0.24	0.5	0.14	0.17
6	2	0	0	5	0	7	0.23	0.5	0.15	0.15
7	1	0	0	5	0	6	0.25	0.48	0.14	0.17
8	1	1	2	4	0	8	0.21	0.5	0.15	0.12
9	1	0	0	5	0	6	0.24	0.49	0.15	0.17
10	1	1	0	3	0	5	0.28	0.48	0.16	0.17
11	1	1	1	4	0	7	0.24	0.48	0.15	0.18

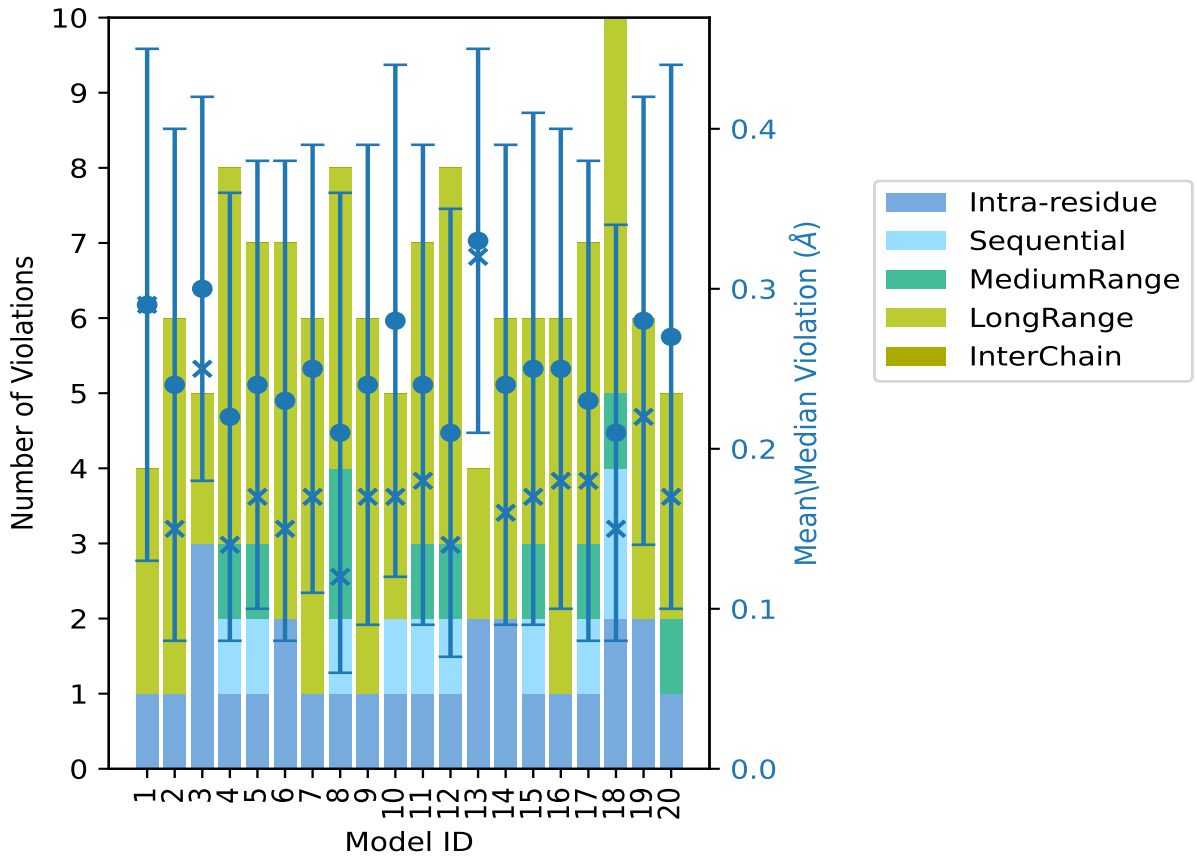
Continued on next page...

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	1	1	1	5	0	8	0.21	0.5	0.14	0.14
13	2	0	0	2	0	4	0.33	0.48	0.12	0.32
14	2	0	0	4	0	6	0.24	0.48	0.15	0.16
15	1	1	1	3	0	6	0.25	0.48	0.16	0.17
16	1	0	0	5	0	6	0.25	0.5	0.15	0.18
17	1	1	1	4	0	7	0.23	0.47	0.15	0.18
18	2	2	1	5	0	10	0.21	0.49	0.13	0.15
19	2	0	0	4	0	6	0.28	0.49	0.14	0.22
20	1	0	1	3	0	5	0.27	0.49	0.17	0.17

[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

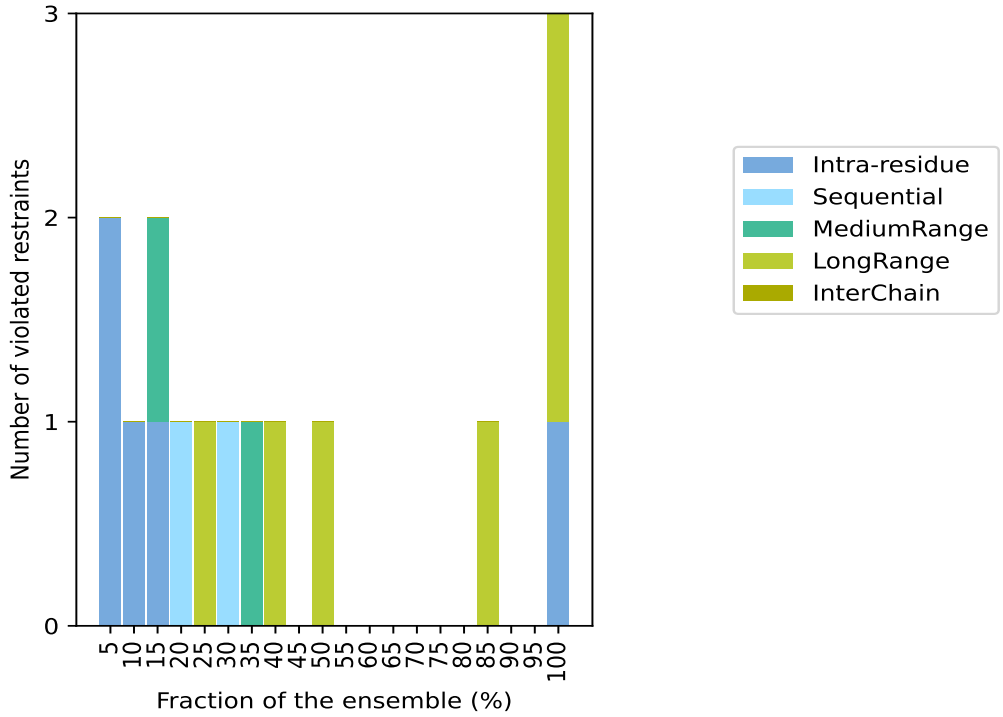
9.3 Distance violation statistics for the ensemble (i)

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, 3360(IR:657, SQ:707, MR:981, LR:875, IC:140) 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	0	0	0	0	2	1	5.0
1	0	0	0	0	1	2	10.0
1	0	1	0	0	2	3	15.0
0	1	0	0	0	1	4	20.0
0	0	0	1	0	1	5	25.0
0	1	0	0	0	1	6	30.0
0	0	1	0	0	1	7	35.0
0	0	0	1	0	1	8	40.0
0	0	0	0	0	0	9	45.0
0	0	0	1	0	1	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	1	0	1	17	85.0
0	0	0	0	0	0	18	90.0
0	0	0	0	0	0	19	95.0
1	0	0	2	0	3	20	100.0

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

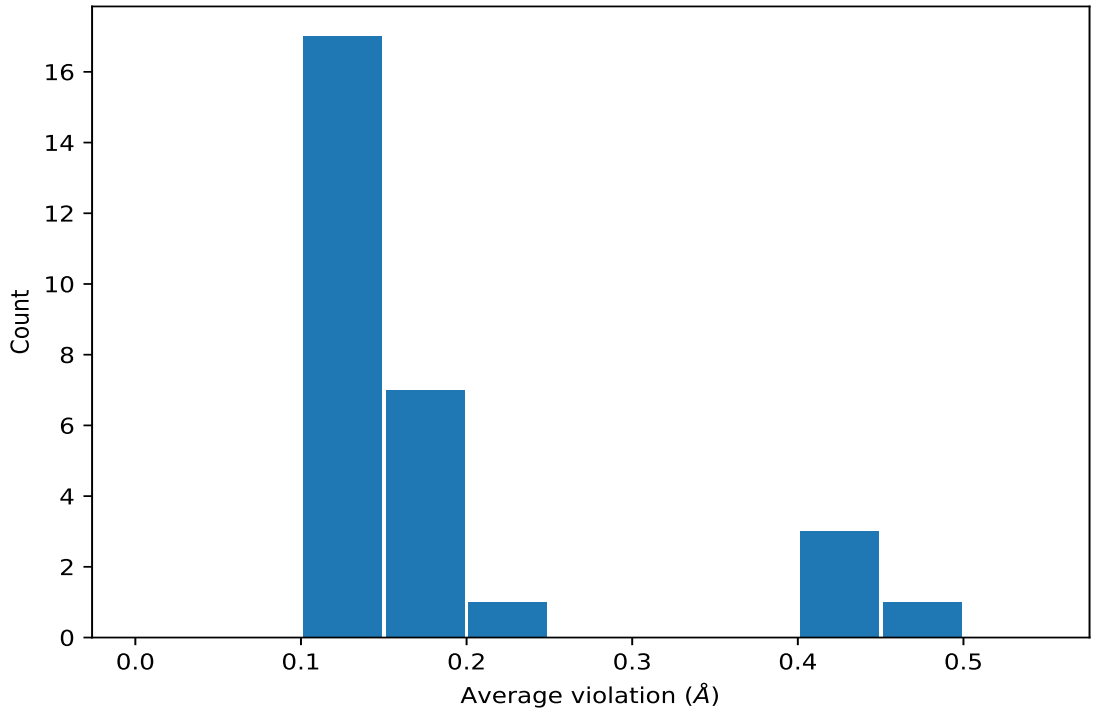
9.3.1 Bar graph : Distance violation statistics for the ensemble



9.4 Most violated distance restraints in the ensemble

9.4.1 Histogram : Distribution of mean distance violations

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



9.4.2 Table: Most violated distance restraints

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

Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,2847)	1:A:108:LEU:H	1:A:108:LEU:HG	20	0.49	0.01	0.49
(1,1623)	1:A:66:ILE:HD11	1:A:108:LEU:HG	20	0.42	0.02	0.41
(1,1623)	1:A:66:ILE:HD12	1:A:108:LEU:HG	20	0.42	0.02	0.41
(1,1623)	1:A:66:ILE:HD13	1:A:108:LEU:HG	20	0.42	0.02	0.41
(1,1535)	1:A:65:LEU:HB2	1:A:70:ASP:H	20	0.19	0.02	0.19
(1,2885)	1:A:108:LEU:HB2	1:A:115:THR:HA	17	0.12	0.01	0.11
(1,2885)	1:A:108:LEU:HB3	1:A:115:THR:HA	17	0.12	0.01	0.11
(1,1630)	1:A:66:ILE:HG12	1:A:108:LEU:H	10	0.12	0.01	0.12
(1,1551)	1:A:65:LEU:HD11	1:A:73:LYS:H	8	0.15	0.01	0.15
(1,1551)	1:A:65:LEU:HD12	1:A:73:LYS:H	8	0.15	0.01	0.15
(1,1551)	1:A:65:LEU:HD13	1:A:73:LYS:H	8	0.15	0.01	0.15
(1,3034)	1:A:117:LYS:HD2	1:A:119:THR:H	7	0.12	0.01	0.12
(1,3034)	1:A:117:LYS:HD3	1:A:119:THR:H	7	0.12	0.01	0.12
(1,3041)	1:A:118:VAL:HB	1:A:119:THR:H	6	0.14	0.03	0.14
(1,1636)	1:A:66:ILE:HA	1:A:108:LEU:HD11	5	0.11	0.0	0.11
(1,1636)	1:A:66:ILE:HA	1:A:108:LEU:HD12	5	0.11	0.0	0.11

Continued on next page...

Continued from previous page...

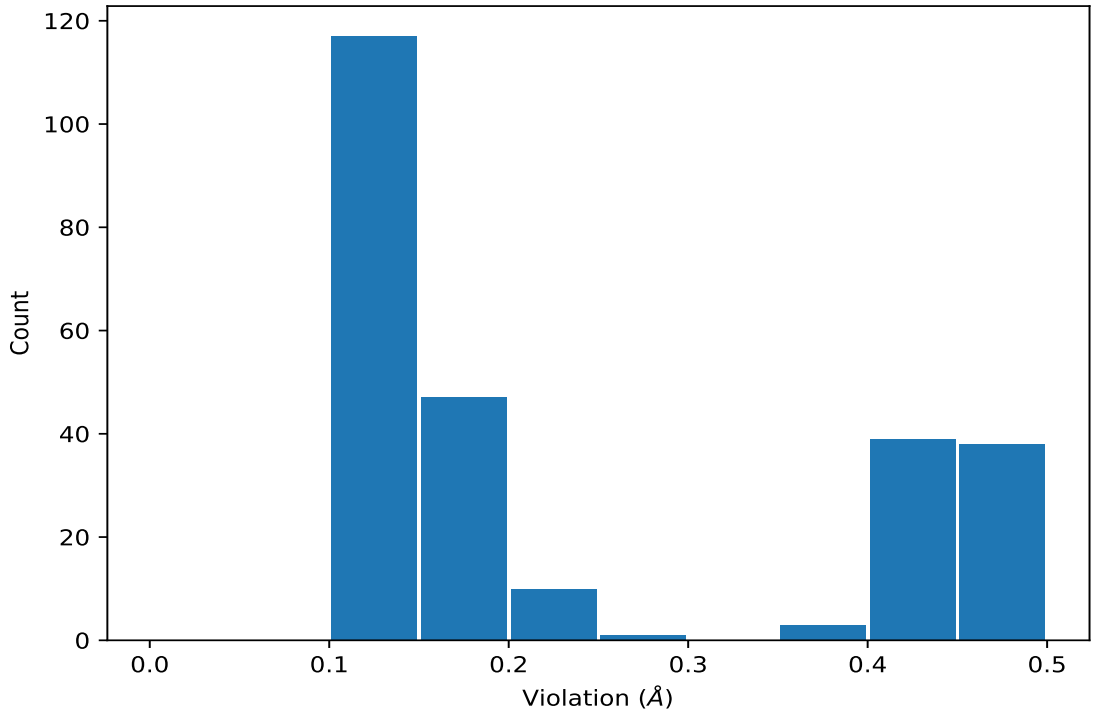
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,1636)	1:A:66:ILE:HA	1:A:108:LEU:HD13	5	0.11	0.0	0.11
(1,1636)	1:A:66:ILE:HA	1:A:108:LEU:HD21	5	0.11	0.0	0.11
(1,1636)	1:A:66:ILE:HA	1:A:108:LEU:HD22	5	0.11	0.0	0.11
(1,1636)	1:A:66:ILE:HA	1:A:108:LEU:HD23	5	0.11	0.0	0.11
(1,2382)	1:A:89:VAL:HG11	1:A:90:GLU:HG2	4	0.17	0.0	0.17

[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)
(1,2847)	1:A:108:LEU:H	1:A:108:LEU:HG	4	0.5
(1,2847)	1:A:108:LEU:H	1:A:108:LEU:HG	5	0.5
(1,2847)	1:A:108:LEU:H	1:A:108:LEU:HG	6	0.5
(1,2847)	1:A:108:LEU:H	1:A:108:LEU:HG	8	0.5
(1,2847)	1:A:108:LEU:H	1:A:108:LEU:HG	12	0.5
(1,2847)	1:A:108:LEU:H	1:A:108:LEU:HG	16	0.5
(1,2847)	1:A:108:LEU:H	1:A:108:LEU:HG	2	0.49
(1,2847)	1:A:108:LEU:H	1:A:108:LEU:HG	9	0.49
(1,2847)	1:A:108:LEU:H	1:A:108:LEU:HG	18	0.49
(1,2847)	1:A:108:LEU:H	1:A:108:LEU:HG	19	0.49

10 Dihedral-angle violation analysis (i)

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