wwPDB NMR Structure Validation Summary Report

Jun 3, 2023 – 08:34 PM EDT

PDB ID	:	2N6A
BMRB ID	:	25757
Title	:	NMR structure of a human calmodulin/connexin-36 peptide hybrid
Authors	:	Donaldson, L.
Deposited on	:	2015-08-14

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:

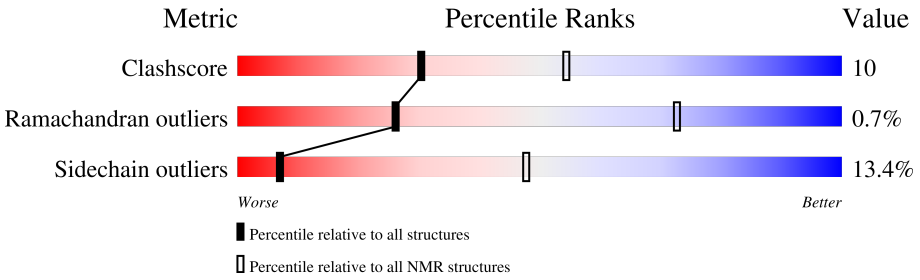
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal 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 72%.

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	173	71% 21% •• 6%

2 Ensemble composition and analysis

This entry contains 18 models. Model 5 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:8-A:168 (161)	0.87	5

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 3 clusters and 4 single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2235 atoms, of which 975 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called human calmodulin/connexin-36 peptide hybrid.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	162	2231	772	975	211	264	9	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	HIS	-	expression tag	UNP P62158
A	0	HIS	-	expression tag	UNP P62158
A	1	HIS	-	expression tag	UNP P62158
A	2	HIS	-	expression tag	UNP P62158
A	3	HIS	-	expression tag	UNP P62158
A	4	HIS	-	expression tag	UNP P62158
A	147	GLY	-	linker	UNP P62158
A	148	ALA	-	linker	UNP P62158
A	149	SER	-	linker	UNP P62158
A	150	THR	-	linker	UNP P62158
A	151	ALA	-	linker	UNP P62158
A	152	ALA	-	linker	UNP P62158
A	153	GLY	-	linker	UNP P62158
A	154	SER	-	linker	UNP P62158

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

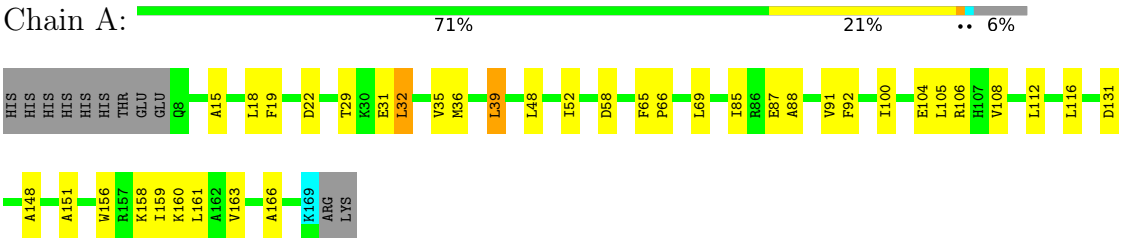
Mol	Chain	Residues	Atoms	
			Total	Ca
2	A	4	4	4

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

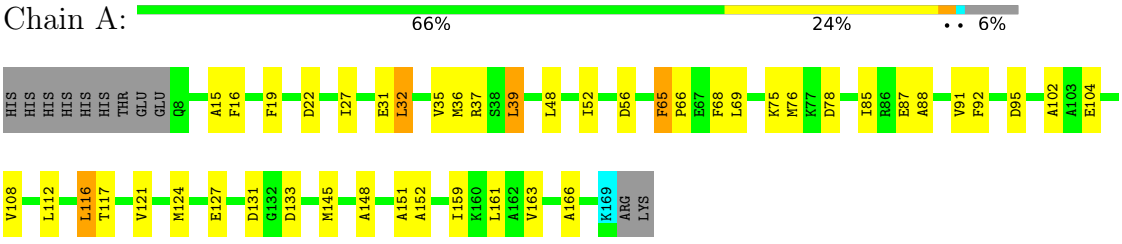
• Molecule 1: human calmodulin/connexin-36 peptide hybrid



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

The representative model is number 5. Colouring as in section 4.1 above.

• Molecule 1: human calmodulin/connexin-36 peptide hybrid



5 Refinement protocol and experimental data overview

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

Of the 400 calculated structures, 18 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	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	1575
Number of shifts mapped to atoms	1380
Number of unparsed shifts	0
Number of shifts with mapping errors	195
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	72%

6 Model quality

6.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CA

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	1247	966	1188	24±3
All	All	22518	17388	21384	434

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 10.

5 of 120 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:15:ALA:HB2	1:A:161:LEU:HD22	0.98	1.35	2	12
1:A:15:ALA:HB2	1:A:161:LEU:HD23	0.98	1.31	3	3
1:A:100:ILE:HG21	1:A:105:LEU:HD13	0.79	1.51	6	11
1:A:18:LEU:HD11	1:A:162:ALA:HB2	0.75	1.57	11	6
1:A:148:ALA:HB1	1:A:151:ALA:HB3	0.75	1.59	15	11

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	160/173 (92%)	146±3 (91±2%)	13±3 (8±2%)	1±0 (1±0%)	26 73
All	All	2880/3114 (92%)	2623 (91%)	238 (8%)	19 (1%)	26 73

All 3 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	85	ILE	17
1	A	114	GLU	1
1	A	78	ASP	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	131/143 (92%)	113±4 (87±3%)	18±4 (13±3%)	7 48
All	All	2358/2574 (92%)	2041 (87%)	317 (13%)	7 48

5 of 71 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	32	LEU	18
1	A	112	LEU	16
1	A	58	ASP	15
1	A	19	PHE	14
1	A	131	ASP	14

6.3.3 RNA

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates

There are no monosaccharides in this entry.

6.6 Ligand geometry

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

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 72% for the well-defined parts and 72% 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	1575
Number of shifts mapped to atoms	1380
Number of unparsed shifts	0
Number of shifts with mapping errors	195
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	1

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	4	HIS	HA	4.671	0.000	1
1	A	4	HIS	HB2	2.508	0.000	2
1	A	4	HIS	HB3	1.98	0.000	2
1	A	4	HIS	C	176.173	0.000	1
1	A	4	HIS	CA	55.635	0.000	1
1	A	4	HIS	CB	35.158	0.000	1
1	A	5	THR	H	8.585	0.003	1
1	A	5	THR	HA	4.779	0.000	1
1	A	5	THR	HB	4.43	0.000	1
1	A	5	THR	HG21	1.307	0.000	1
1	A	5	THR	HG22	1.307	0.000	1
1	A	5	THR	HG23	1.307	0.000	1
1	A	5	THR	C	175.352	0.000	1
1	A	5	THR	CA	60.9	0.000	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	5	THR	CB	71.022	0.059	1
1	A	5	THR	N	112.91	0.000	1
1	A	6	GLU	H	9.071	0.004	1
1	A	6	GLU	HA	3.95	0.000	1
1	A	6	GLU	HB2	2.028	0.000	1
1	A	6	GLU	HG2	2.352	0.000	2
1	A	6	GLU	HG3	2.34	0.000	2
1	A	6	GLU	C	179.499	0.014	1
1	A	6	GLU	CA	59.964	0.000	1
1	A	6	GLU	CB	29.19	0.000	1
1	A	6	GLU	N	120.47	0.000	1
1	A	7	GLU	H	8.78	0.003	1
1	A	7	GLU	HA	4.058	0.000	1
1	A	7	GLU	HB2	1.98	0.000	2
1	A	7	GLU	HB3	1.932	0.000	2
1	A	7	GLU	HG2	2.304	0.000	2
1	A	7	GLU	HG3	2.328	0.000	2
1	A	7	GLU	C	179.075	0.000	1
1	A	7	GLU	CA	60.023	0.059	1
1	A	7	GLU	CB	29.307	0.000	1
1	A	7	GLU	N	119.5	0.000	1
1	A	8	GLN	HB2	1.644	0.000	2
1	A	8	GLN	HG2	1.92	0.000	2
1	A	9	ILE	HG12	1.115	0.001	2
1	A	11	GLU	HB2	1.98	0.000	2
1	A	11	GLU	HG2	2.304	0.000	2
1	A	12	PHE	HB2	3.416	0.032	1
1	A	13	LYS	HB2	1.879	0.017	1
1	A	13	LYS	HG2	1.173	0.014	1
1	A	13	LYS	HD2	0.954	0.012	1
1	A	13	LYS	HE2	2.472	0.000	1
1	A	14	GLU	HB2	2.184	0.000	1
1	A	14	GLU	HG2	2.436	0.000	1
1	A	16	PHE	HB2	3.188	0.018	2
1	A	18	LEU	HB2	1.843	0.016	2
1	A	22	ASP	HB2	3.037	0.000	2
1	A	24	ASP	HB2	3.009	0.004	2
1	A	25	GLY	HA3	3.688	0.021	2
1	A	30	LYS	HB2	1.794	0.003	1
1	A	30	LYS	HG2	1.436	0.010	1
1	A	30	LYS	HE2	2.952	0.000	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	31	GLU	HB2	2.62	0.013	1
1	A	31	GLU	HG2	2.36	0.019	1
1	A	32	LEU	HB2	1.839	0.012	2
1	A	33	GLY	HA3	4.125	0.016	2
1	A	38	SER	HB2	3.768	0.011	1
1	A	39	LEU	HB2	1.897	0.021	2
1	A	40	GLY	HA3	3.729	0.011	2
1	A	43	PRO	HB2	2.173	0.000	1
1	A	43	PRO	HG2	1.958	0.000	1
1	A	43	PRO	HD2	3.645	0.000	1
1	A	45	GLU	HB2	2.002	0.017	1
1	A	45	GLU	HG2	2.303	0.009	1
1	A	47	GLU	HB2	1.872	0.000	1
1	A	47	GLU	HG2	2.334	0.000	1
1	A	48	LEU	HB2	2.079	0.013	2
1	A	49	GLN	HB2	2.152	0.000	1
1	A	49	GLN	HG2	2.431	0.000	1
1	A	50	ASP	HB2	2.774	0.012	2
1	A	52	ILE	HG12	0.985	0.015	2
1	A	53	ASN	HB2	2.99	0.000	2
1	A	54	GLU	HB2	2.216	0.000	2
1	A	54	GLU	HG2	2.443	0.010	1
1	A	56	ASP	HB2	2.732	0.000	2
1	A	58	ASP	HB2	3.024	0.020	2
1	A	60	ASN	HB2	3.309	0.019	2
1	A	61	GLY	HA3	3.442	0.016	2
1	A	63	ILE	HG12	1.582	0.015	1
1	A	64	ASP	HB2	3.065	0.011	2
1	A	65	PHE	HB2	2.206	0.000	2
1	A	66	PRO	HB2	2.549	0.000	1
1	A	66	PRO	HG2	2.055	0.000	1
1	A	66	PRO	HD2	2.915	0.000	1
1	A	68	PHE	HB2	3.178	0.024	1
1	A	69	LEU	HB2	1.508	0.020	1
1	A	71	MET	HB2	2.457	0.007	2
1	A	72	MET	HB2	1.24	0.007	1
1	A	72	MET	HG2	1.056	0.013	1
1	A	75	LYS	HB2	1.915	0.000	1
1	A	75	LYS	HG2	1.69	0.000	1
1	A	75	LYS	HE2	3.216	0.000	1
1	A	78	ASP	HB2	2.754	0.000	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	80	ASP	HB2	2.754	0.000	2
1	A	82	GLU	HB2	2.152	0.000	1
1	A	82	GLU	HG2	2.356	0.000	1
1	A	83	GLU	HB2	2.034	0.000	1
1	A	83	GLU	HG2	2.313	0.000	1
1	A	84	GLU	HB2	2.098	0.000	1
1	A	84	GLU	HG2	2.334	0.000	1
1	A	85	ILE	HG12	1.775	0.020	2
1	A	86	ARG	HB2	1.884	0.012	2
1	A	86	ARG	HG2	2.049	0.003	1
1	A	86	ARG	HD2	2.99	0.000	1
1	A	87	GLU	HB2	2.055	0.021	1
1	A	87	GLU	HG2	2.351	0.005	1
1	A	89	PHE	HB2	3.199	0.014	2
1	A	90	ARG	HB2	1.59	0.024	1
1	A	90	ARG	HG2	1.909	0.004	1
1	A	90	ARG	HD2	3.183	0.000	1
1	A	92	PHE	HB2	2.754	0.000	2
1	A	93	ASP	HB2	1.321	0.014	1
1	A	97	ASN	HB2	3.406	0.003	2
1	A	98	GLY	HA3	3.408	0.017	2
1	A	99	TYR	HB2	1.941	0.000	2
1	A	101	SER	HB2	4.046	0.025	2
1	A	104	GLU	HB2	2.484	0.027	2
1	A	105	LEU	HB2	1.769	0.008	2
1	A	106	ARG	HB2	1.577	0.000	1
1	A	106	ARG	HG2	1.907	0.012	1
1	A	106	ARG	HD2	3.163	0.011	1
1	A	107	HIS	HB2	3.288	0.000	2
1	A	107	HIS	HB3	3.174	0.000	2
1	A	111	ASN	HB2	2.922	0.005	2
1	A	112	LEU	HB2	1.767	0.009	2
1	A	113	GLY	HA3	3.644	0.013	2
1	A	114	GLU	HB2	1.563	0.014	1
1	A	114	GLU	HG2	1.872	0.000	1
1	A	115	LYS	HB2	1.691	0.000	2
1	A	115	LYS	HG2	1.317	0.007	2
1	A	115	LYS	HE2	2.898	0.000	1
1	A	116	LEU	HB2	1.569	0.019	2
1	A	120	GLU	HB2	1.891	0.000	2
1	A	120	GLU	HG2	2.347	0.000	2

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	122	ASP	HB2	2.794	0.000	2
1	A	123	GLU	HB2	2.138	0.000	1
1	A	123	GLU	HG2	2.356	0.000	1
1	A	125	ILE	HG12	1.465	0.002	2
1	A	127	GLU	HB2	2.329	0.015	2
1	A	127	GLU	HG2	2.583	0.008	2
1	A	129	ASP	HB2	2.56	0.015	1
1	A	130	ILE	HG12	1.181	0.019	2
1	A	131	ASP	HB2	3.02	0.022	2
1	A	132	GLY	HA3	3.793	0.013	2
1	A	133	ASP	HB2	2.879	0.026	2
1	A	134	GLY	HA3	3.391	0.016	2
1	A	135	GLN	HB2	1.701	0.010	1
1	A	135	GLN	HG2	1.942	0.004	1
1	A	137	ASN	HB2	3.293	0.008	1
1	A	138	TYR	HB2	2.322	0.017	2
1	A	139	GLU	HB2	1.996	0.017	1
1	A	139	GLU	HG2	2.28	0.000	1
1	A	140	GLU	HB2	2.376	0.000	2
1	A	140	GLU	HG2	2.821	0.021	2
1	A	141	PHE	HB2	3.367	0.016	2
1	A	143	GLN	HB2	2.332	0.015	2
1	A	144	MET	HB2	1.686	0.012	2
1	A	144	MET	HG2	1.109	0.021	2
1	A	145	MET	HG2	1.52	0.000	1
1	A	153	GLY	HA3	4.02	0.000	2
1	A	154	SER	HB2	3.859	0.000	1
1	A	155	GLY	HA3	3.722	0.013	2
1	A	156	TRP	HB2	3.084	0.010	2
1	A	157	ARG	HB2	1.938	0.000	1
1	A	157	ARG	HG2	1.691	0.000	1
1	A	157	ARG	HD2	3.203	0.000	1
1	A	158	LYS	HB2	1.907	0.006	1
1	A	158	LYS	HG2	1.682	0.000	1
1	A	158	LYS	HD2	1.387	0.000	1
1	A	158	LYS	HE2	2.813	0.000	1
1	A	159	ILE	HG12	0.783	0.016	2
1	A	160	LYS	HB2	1.971	0.006	1
1	A	160	LYS	HG2	1.682	0.000	1
1	A	161	LEU	HB2	1.995	0.000	1
1	A	164	ARG	HB2	1.942	0.022	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	164	ARG	HG2	1.796	0.000	1
1	A	165	GLY	HA3	3.671	0.026	2
1	A	167	GLN	HB2	2.09	0.000	2
1	A	167	GLN	HG2	2.402	0.021	1
1	A	169	LYS	HB2	1.406	0.000	2
1	A	169	LYS	HE2	2.927	0.000	1
1	A	170	ARG	H	8.509	0.004	1
1	A	170	ARG	HG2	1.853	0.000	1
1	A	170	ARG	C	175.469	0.007	1
1	A	170	ARG	CA	56.117	0.000	1
1	A	170	ARG	CB	30.711	0.000	1
1	A	170	ARG	N	124.713	0.000	1
1	A	171	LYS	H	8.079	0.003	1
1	A	171	LYS	C	181.31	0.000	1
1	A	171	LYS	CA	57.91	0.000	1
1	A	171	LYS	CB	33.543	0.000	1
1	A	171	LYS	N	128.205	0.000	1

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	158	-0.33 ± 0.08	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	144	0.16 ± 0.11	None needed (< 0.5 ppm)
$^{13}C'$	155	-0.49 ± 0.07	None needed (< 0.5 ppm)
^{15}N	151	0.37 ± 0.26	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 72%, i.e. 1517 atoms were assigned a chemical shift out of a possible 2100. 0 out of 17 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	738/816 (90%)	294/335 (88%)	299/322 (93%)	145/159 (91%)
Sidechain	711/1168 (61%)	471/750 (63%)	240/373 (64%)	0/45 (0%)
Aromatic	68/116 (59%)	34/58 (59%)	33/57 (58%)	1/1 (100%)
Overall	1517/2100 (72%)	799/1143 (70%)	572/752 (76%)	146/205 (71%)

7.1.4 Statistically unusual chemical shifts (i)

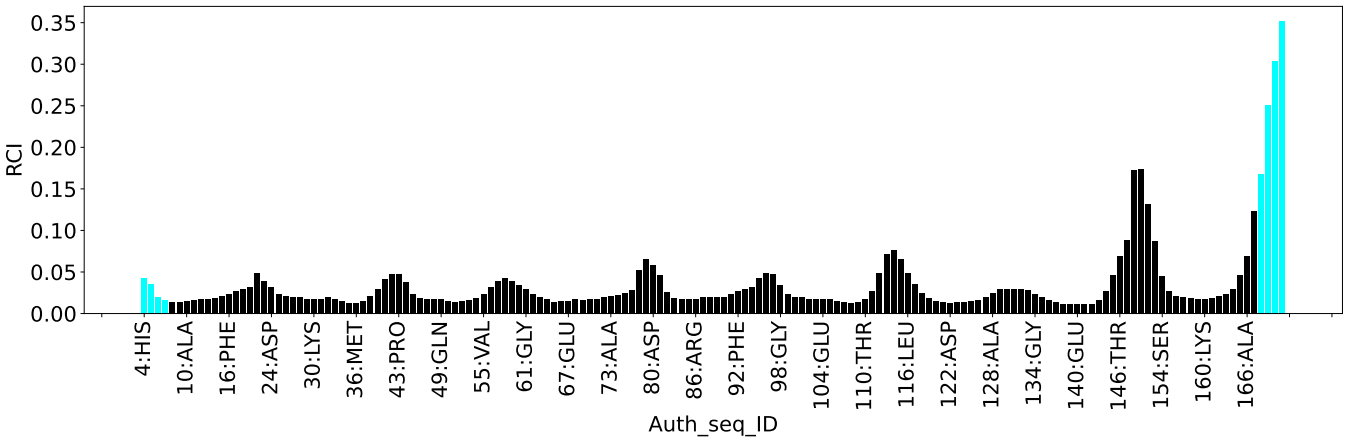
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	93	ASP	HB2	1.32	1.41 – 4.01	-5.3

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	1147
Intra-residue (|i-j|=0)	396
Sequential (|i-j|=1)	291
Medium range (|i-j|>1 and |i-j|<5)	115
Long range (|i-j|≥5)	213
Inter-chain	0
Hydrogen bond restraints	132
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	212
Number of restraints per residue	6.6
Number of long range restraints per residue[1]	1.3

[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)	47.8	0.2
0.2-0.5 (Medium)	39.0	0.5
>0.5 (Large)	26.6	3.82

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)

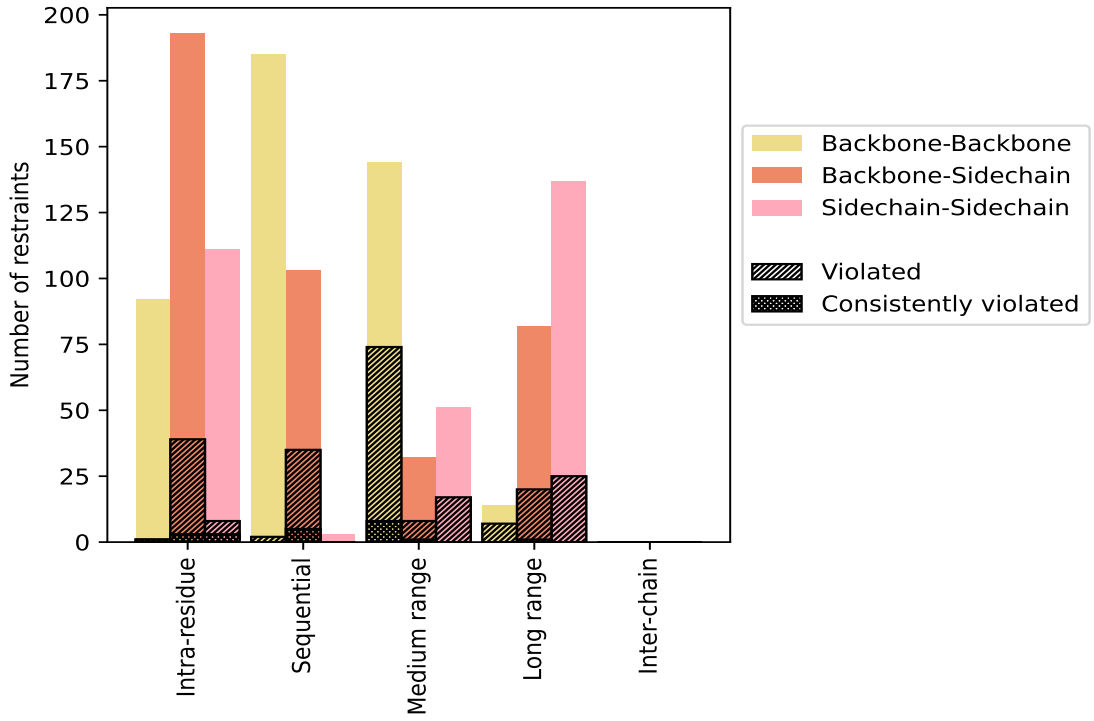
9.1 Summary of distance violations (i)

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)	396	34.5	48	12.1	4.2	7	1.8	0.6
Backbone-Backbone	92	8.0	1	1.1	0.1	1	1.1	0.1
Backbone-Sidechain	193	16.8	39	20.2	3.4	3	1.6	0.3
Sidechain-Sidechain	111	9.7	8	7.2	0.7	3	2.7	0.3
Sequential (|i-j|=1)	291	25.4	37	12.7	3.2	5	1.7	0.4
Backbone-Backbone	185	16.1	2	1.1	0.2	0	0.0	0.0
Backbone-Sidechain	103	9.0	35	34.0	3.1	5	4.9	0.4
Sidechain-Sidechain	3	0.3	0	0.0	0.0	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	115	10.0	28	24.3	2.4	1	0.9	0.1
Backbone-Backbone	36	3.1	5	13.9	0.4	0	0.0	0.0
Backbone-Sidechain	28	2.4	6	21.4	0.5	1	3.6	0.1
Sidechain-Sidechain	51	4.4	17	33.3	1.5	0	0.0	0.0
Long range (|i-j|≥5)	213	18.6	40	18.8	3.5	1	0.5	0.1
Backbone-Backbone	2	0.2	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	74	6.5	15	20.3	1.3	1	1.4	0.1
Sidechain-Sidechain	137	11.9	25	18.2	2.2	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	132	11.5	83	62.9	7.2	8	6.1	0.7
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	1147	100.0	236	20.6	20.6	22	1.9	1.9
Backbone-Backbone	435	37.9	84	19.3	7.3	9	2.1	0.8
Backbone-Sidechain	410	35.7	102	24.9	8.9	10	2.4	0.9
Sidechain-Sidechain	302	26.3	50	16.6	4.4	3	1.0	0.3

[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	23	15	51	22	0	111	0.47	3.46	0.56	0.27
2	25	19	50	17	0	111	0.37	1.6	0.33	0.25
3	24	21	48	21	0	114	0.37	1.74	0.31	0.25
4	22	20	49	23	0	114	0.44	3.43	0.53	0.26
5	27	20	53	19	0	119	0.4	2.09	0.4	0.24
6	20	15	50	26	0	111	0.48	3.44	0.58	0.25
7	20	19	52	22	0	113	0.4	2.15	0.39	0.25
8	18	17	51	25	0	111	0.44	2.87	0.49	0.25
9	22	15	46	24	0	107	0.39	3.55	0.48	0.21
10	21	22	49	25	0	117	0.38	1.67	0.36	0.22
11	19	21	54	22	0	116	0.43	3.82	0.56	0.24

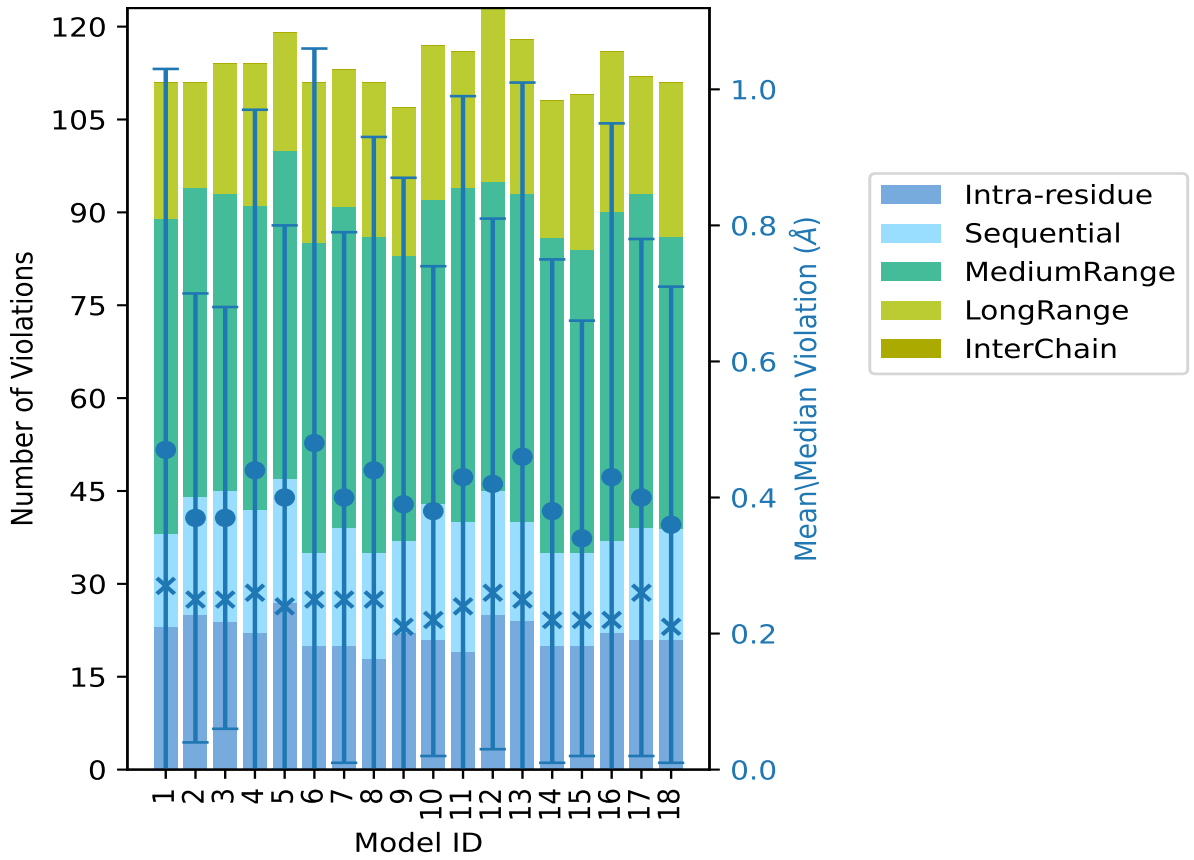
Continued on next page...

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	25	20	50	28	0	123	0.42	1.94	0.39	0.26
13	24	16	53	25	0	118	0.46	3.43	0.55	0.25
14	20	15	51	22	0	108	0.38	1.68	0.37	0.22
15	20	15	49	25	0	109	0.34	1.78	0.32	0.22
16	22	15	53	26	0	116	0.43	3.39	0.52	0.22
17	21	18	54	19	0	112	0.4	2.08	0.38	0.26
18	21	18	47	25	0	111	0.36	1.74	0.35	0.21

[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

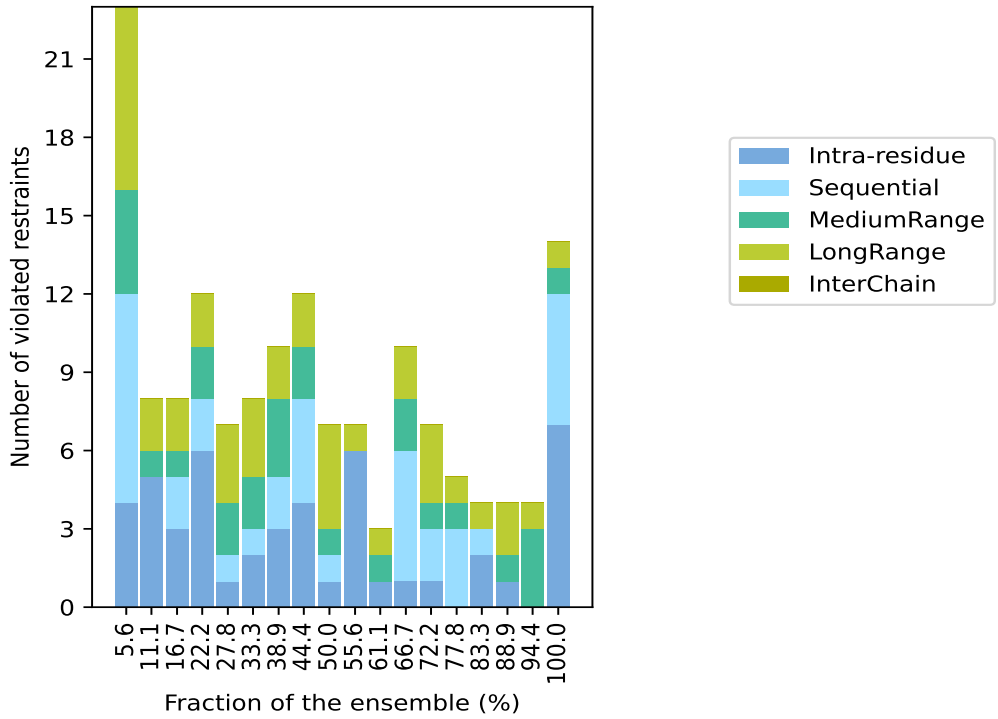
9.3 Distance violation statistics for the ensemble (i)

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, 862(IR:348, SQ:254, MR:87, LR:173, 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]	%
4	8	4	7	0	23	1	5.6
5	0	1	2	0	8	2	11.1
3	2	1	2	0	8	3	16.7
6	2	2	2	0	12	4	22.2
1	1	2	3	0	7	5	27.8
2	1	2	3	0	8	6	33.3
3	2	3	2	0	10	7	38.9
4	4	2	2	0	12	8	44.4
1	1	1	4	0	7	9	50.0
6	0	0	1	0	7	10	55.6
1	0	1	1	0	3	11	61.1
1	5	2	2	0	10	12	66.7
1	2	1	3	0	7	13	72.2
0	3	1	1	0	5	14	77.8
2	1	0	1	0	4	15	83.3
1	0	1	2	0	4	16	88.9
0	0	3	1	0	4	17	94.4
7	5	1	1	0	14	18	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

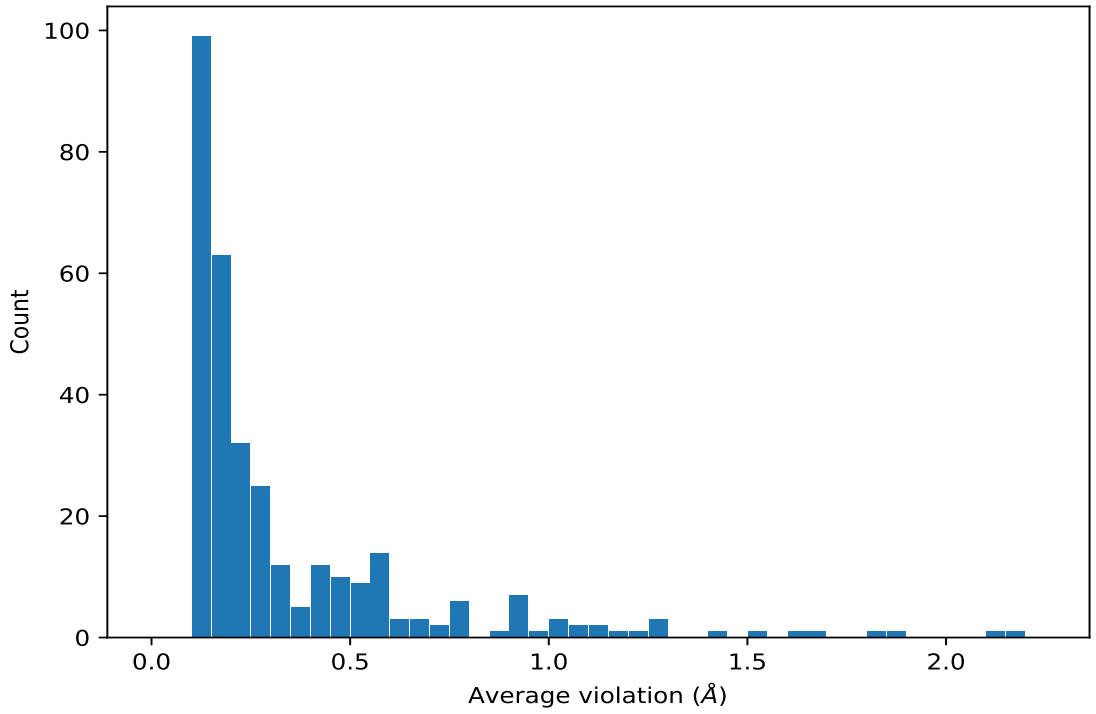
9.3.1 Bar graph : Distance violation statistics for the ensemble



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.

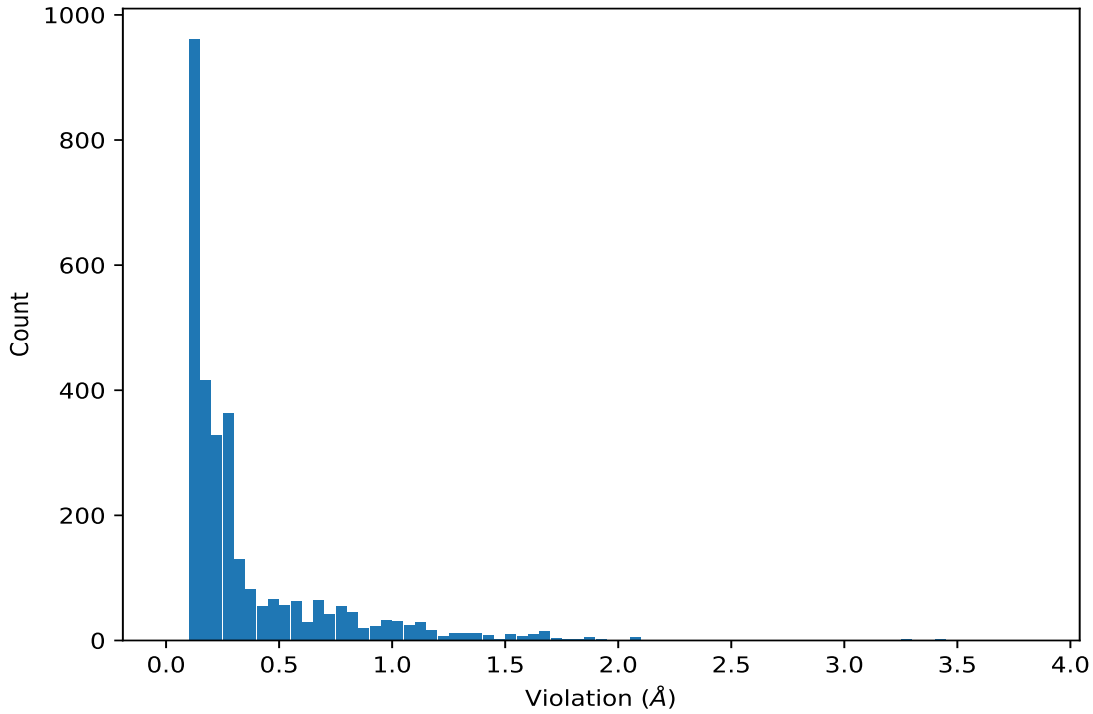
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,122)	1:A:23:GLY:H	1:A:22:ASP:HB3	18	0.77	0.07	0.78
(1,220)	1:A:62:THR:H	1:A:60:ASN:HB3	18	0.74	0.17	0.78
(1,436)	1:A:160:LYS:H	1:A:159:ILE:HG13	18	0.69	0.04	0.7
(1,780)	1:A:124:MET:HE1	1:A:155:GLY:HA2	18	0.58	0.17	0.56
(1,780)	1:A:124:MET:HE2	1:A:155:GLY:HA2	18	0.58	0.17	0.56
(1,780)	1:A:124:MET:HE3	1:A:155:GLY:HA2	18	0.58	0.17	0.56
(1,25)	1:A:156:TRP:H	1:A:156:TRP:HB3	18	0.49	0.0	0.49
(1,32)	1:A:141:PHE:H	1:A:140:GLU:HB3	18	0.39	0.06	0.37
(1,368)	1:A:135:GLN:H	1:A:135:GLN:HB3	18	0.38	0.15	0.31
(1,368)	1:A:135:GLN:H	1:A:135:GLN:HB3	18	0.38	0.15	0.31
(1,503)	1:A:52:ILE:H	1:A:52:ILE:HG13	18	0.34	0.05	0.35
(2,98)	1:A:153:GLY:O	1:A:157:ARG:N	18	0.3	0.03	0.3
(1,387)	1:A:132:GLY:H	1:A:131:ASP:HB3	18	0.3	0.03	0.3

[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)
(4,1)	1:A:20:ASP:OD1	1:A:22:ASP:OD2	11	3.82
(4,3)	1:A:22:ASP:OD2	1:A:24:ASP:OD1	9	3.55
(4,24)	1:A:140:GLU:OE2	1:A:131:ASP:OD2	1	3.46
(4,24)	1:A:140:GLU:OE2	1:A:131:ASP:OD2	6	3.44
(4,24)	1:A:140:GLU:OE2	1:A:131:ASP:OD2	4	3.43
(4,24)	1:A:140:GLU:OE2	1:A:131:ASP:OD2	13	3.43
(4,24)	1:A:140:GLU:OE2	1:A:131:ASP:OD2	16	3.39
(4,21)	1:A:131:ASP:OD2	1:A:133:ASP:OD1	1	3.32
(4,21)	1:A:131:ASP:OD2	1:A:133:ASP:OD1	11	3.28
(4,21)	1:A:131:ASP:OD2	1:A:133:ASP:OD1	13	3.26

10 Dihedral-angle violation analysis (i)

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