WORLDWIDE PDB

PROTEIN DATA BANK

wwPDB NMR Structure Validation Summary Report ⓘ

Jun 5, 2023 – 07:32 PM EDT

PDB ID	:	2M33
BMRB ID	:	18919
Title	:	Solution NMR structure of full-length oxidized microsomal rabbit cytochrome b5
Authors	:	Subramanian, V.; Ahuja, S.; Popovych, N.; Huang, R.; Le Clair, S.V.; Jahr, N.; Soong, R.; Xu, J.; Yamamoto, K.; Nanga, R.P.; Im, S.; Waskell, L.; Ramamoorthy, A.
Deposited on	:	2013-01-08

The following versions of software and data (see references ①) were used in the production of this report:

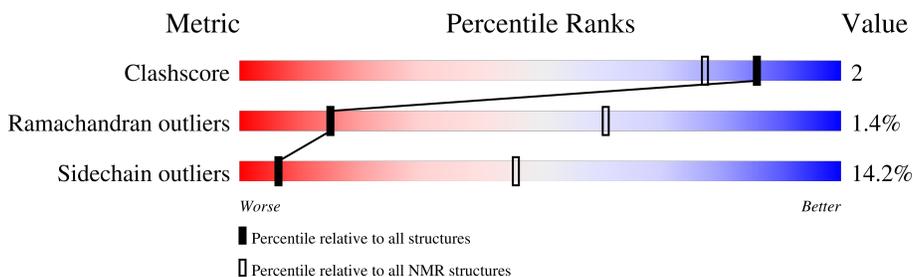|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints 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 61%.

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	Percentile Ranks	Value
Clashscore		2
Ramachandran outliers		1.4%
Sidechain outliers		14.2%

Worse — *Better*

■ Percentile relative to all structures
▯ Percentile relative to all NMR structures

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	134	69% 7% · 22%

2 Ensemble composition and analysis ⓘ

This entry contains 23 models. Model 7 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *minimized average structure*.

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:2-A:104 (103)	0.62	7

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 5 clusters. No single-model clusters were found.

Cluster number	Models
1	1, 5, 7, 13, 14, 20, 23
2	2, 3, 4, 12, 15, 19
3	6, 11, 16, 18, 21
4	8, 9, 10
5	17, 22

3   Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 1724 atoms, of which 845 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Cytochrome b5.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	104	1651	523	815	142	169	2	0

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).

HEM

Mol	Chain	Residues	Atoms					
			Total	C	Fe	H	N	O
2	A	1	73	34	1	30	4	4

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.
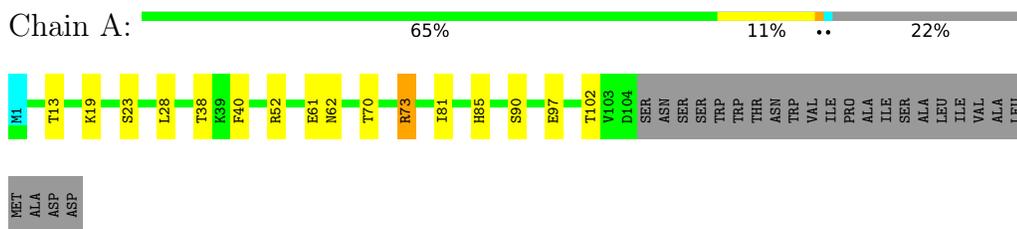
• Molecule 1: Cytochrome b5
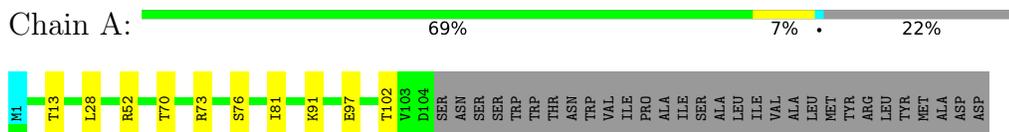
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 7. Colouring as in section 4.1 above.

• Molecule 1: Cytochrome b5

5 Refinement protocol and experimental data overview ⓘ

The models were refined using the following method: *distance geometry, torsion angle dynamics, simulated annealing.*

Of the 50 calculated structures, 23 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	2.1
HADDOCK	refinement	
CYANA	refinement	2.1

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	939
Number of shifts mapped to atoms	849
Number of unparsed shifts	0
Number of shifts with mapping errors	90
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	61%

6   Model quality ⓘ

6.1   Standard geometry ⓘ

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

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	828	804	799	4±1
2	A	43	30	30	0±1
All	All	20033	19182	19067	82

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 2.

5 of 33 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Total
1:A:17:ILE:HD13	1:A:29:ILE:HD11	0.60	1.71	6	6
1:A:60:THR:O	1:A:64:GLU:HG2	0.54	2.02	14	10
1:A:62:ASN:OD1	2:A:400:HEM:HAC	0.54	2.02	21	7
1:A:70:THR:O	1:A:73:ARG:HG3	0.51	2.05	18	9
1:A:49:GLU:OE1	1:A:52:ARG:HD2	0.50	2.06	16	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	102/134 (76%)	91±2 (90±2%)	9±2 (9±2%)	1±1 (1±1%)	15	61
All	All	2346/3082 (76%)	2100 (90%)	214 (9%)	32 (1%)	15	61

All 5 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	97	GLU	19
1	A	91	LYS	9
1	A	9	VAL	2
1	A	46	GLY	1
1	A	103	VAL	1

6.3.2 Protein sidechains ⓘ

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	92/119 (77%)	79±3 (86±3%)	13±3 (14±3%)	6	46
All	All	2116/2737 (77%)	1815 (86%)	301 (14%)	6	46

5 of 49 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	13	THR	23
1	A	73	ARG	23
1	A	81	ILE	20
1	A	52	ARG	20
1	A	102	THR	18

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 ⓘ

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

| Mol | Type | Chain | Res | Link | Bond lengths | | |
					Counts	RMSZ	#Z>2
2	HEM	A	400	1	41,50,50	1.61±0.01	5±0 (12±0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of angles that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

| Mol | Type | Chain | Res | Link | Bond angles | | |
					Counts	RMSZ	#Z>2
2	HEM	A	400	1	45,82,82	0.99±0.01	2±1 (4±1%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means

no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	HEM	A	400	1	-	0±0,12,54,54	-

All unique bond outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models Worst	Total
2	A	400	HEM	C3C-CAC	4.77	1.38	1.47	9	23
2	A	400	HEM	CBB-CAB	4.74	1.53	1.30	4	23
2	A	400	HEM	CBC-CAC	3.69	1.53	1.29	14	23
2	A	400	HEM	C3C-C2C	3.56	1.35	1.40	18	23
2	A	400	HEM	CAB-C3B	3.34	1.38	1.47	21	23

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models Worst	Total
2	A	400	HEM	C4B-CHC-C1C	2.22	125.49	122.56	16	17
2	A	400	HEM	C4C-CHD-C1D	2.20	125.46	122.56	22	17
2	A	400	HEM	CAD-CBD-CGD	2.17	108.94	113.60	22	12

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand HEM A 400

Bond lengths

Bond angles

Torsions

Rings
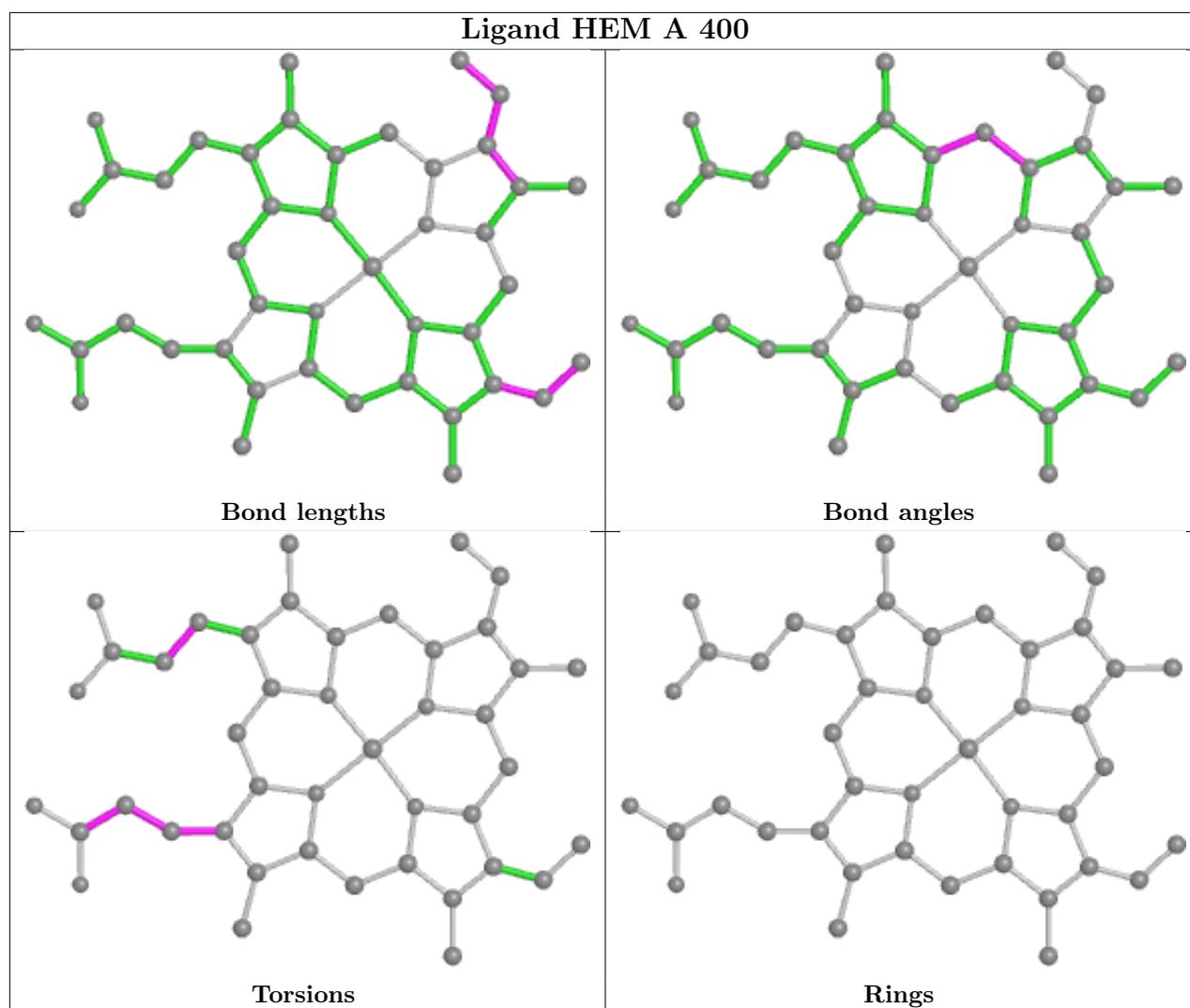
6.7 Other polymers ⓘ

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 61% for the well-defined parts and 60% 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	939
Number of shifts mapped to atoms	849
Number of unparsed shifts	0
Number of shifts with mapping errors	90
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	29

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 90) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	105	SER	HA	4.114	0.000	.
1	A	105	SER	HB2	4.004	0.000	.
1	A	105	SER	HB3	4.004	0.000	.
1	A	106	ASN	H	7.581	0.002	.
1	A	106	ASN	HA	4.264	0.000	.
1	A	106	ASN	HB2	2.977	0.000	.
1	A	106	ASN	HB3	2.782	0.000	.
1	A	106	ASN	N	127.088	0.028	.
1	A	111	THR	HG21	1.025	0.001	.
1	A	111	THR	HG22	1.025	0.001	.
1	A	111	THR	HG23	1.025	0.001	.
1	A	112	ASN	HB2	2.647	0.004	.
1	A	112	ASN	HB3	2.705	0.003	.
1	A	112	ASN	HD21	6.815	0.005	.

Continued on next page...

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	112	ASN	HD22	7.566	0.004	.
1	A	112	ASN	ND2	113.64	0.051	.
1	A	113	TRP	HA	4.266	0.000	.
1	A	113	TRP	HB2	2.855	0.000	.
1	A	113	TRP	HB3	2.855	0.000	.
1	A	114	VAL	HA	3.834	0.000	.
1	A	116	PRO	HB2	2.075	0.000	.
1	A	116	PRO	HB3	2.075	0.000	.
1	A	116	PRO	HG2	1.774	0.000	.
1	A	116	PRO	HG3	1.774	0.000	.
1	A	116	PRO	HD3	3.271	0.003	.
1	A	116	PRO	CD	47.675	0.111	.
1	A	120	ALA	HA	3.652	0.000	.
1	A	122	ILE	H	7.758	0.000	.
1	A	122	ILE	HA	4.474	0.690	.
1	A	122	ILE	HB	1.533	0.000	.
1	A	123	VAL	HA	4.207	0.000	.
1	A	123	VAL	HG11	-2.403	0.000	.
1	A	123	VAL	HG12	-2.403	0.000	.
1	A	123	VAL	HG13	-2.403	0.000	.
1	A	124	ALA	H	8.154	0.003	.
1	A	124	ALA	HB1	1.23	0.006	.
1	A	124	ALA	HB2	1.23	0.006	.
1	A	124	ALA	HB3	1.23	0.006	.
1	A	124	ALA	N	124.593	0.010	.
1	A	125	LEU	HA	4.888	0.003	.
1	A	126	MET	H	8.144	0.013	.
1	A	126	MET	HA	4.04	0.002	.
1	A	126	MET	HB2	1.673	0.001	.
1	A	126	MET	HB3	1.74	0.003	.
1	A	126	MET	HG2	1.966	0.001	.
1	A	126	MET	HG3	1.966	0.001	.
1	A	126	MET	N	123.137	0.018	.
1	A	127	TYR	H	8.186	0.003	.
1	A	127	TYR	HA	4.487	0.006	.
1	A	127	TYR	HB2	2.478	0.002	.
1	A	127	TYR	HB3	2.584	0.005	.
1	A	127	TYR	HD1	6.986	0.000	.
1	A	127	TYR	HD2	6.986	0.000	.
1	A	127	TYR	CA	52.638	0.019	.
1	A	127	TYR	CB	38.952	0.039	.

Continued on next page...

Continued from previous page...

| List ID | Chain | Res | Type | Atom | Shift Data | | Ambiguity |
					Value	Uncertainty	
1	A	127	TYR	N	119.91	0.041	.
1	A	128	ARG	H	8.139	0.004	.
1	A	128	ARG	HD2	2.984	0.014	.
1	A	128	ARG	HD3	2.984	0.014	.
1	A	128	ARG	N	121.788	0.027	.
1	A	129	LEU	HB2	1.362	0.000	.
1	A	129	LEU	HB3	1.866	0.000	.
1	A	129	LEU	HG	1.446	0.000	.
1	A	129	LEU	HD11	0.672	0.007	.
1	A	129	LEU	HD12	0.672	0.007	.
1	A	129	LEU	HD13	0.672	0.007	.
1	A	129	LEU	HD21	0.633	0.003	.
1	A	129	LEU	HD22	0.633	0.003	.
1	A	129	LEU	HD23	0.633	0.003	.
1	A	131	MET	HB3	1.776	0.002	.
1	A	132	ALA	HA	4.007	0.006	.
1	A	132	ALA	HB1	1.211	0.004	.
1	A	132	ALA	HB2	1.211	0.004	.
1	A	132	ALA	HB3	1.211	0.004	.
1	A	132	ALA	CA	52.332	0.030	.
1	A	132	ALA	CB	19.055	0.068	.
1	A	133	ASP	H	8.128	0.004	.
1	A	133	ASP	HA	4.439	0.007	.
1	A	133	ASP	HB2	2.334	0.008	.
1	A	133	ASP	HB3	2.537	0.004	.
1	A	133	ASP	CA	53.656	0.022	.
1	A	133	ASP	CB	41.39	0.033	.
1	A	133	ASP	N	120.504	0.016	.
1	A	134	ASP	H	7.753	0.003	.
1	A	134	ASP	HA	4.147	0.004	.
1	A	134	ASP	HB2	2.353	0.001	.
1	A	134	ASP	HB3	2.437	0.003	.
1	A	134	ASP	CA	55.487	0.016	.
1	A	134	ASP	CB	42.232	0.038	.
1	A	134	ASP	N	125.6	0.058	.

7.1.2   Chemical shift referencing   ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_\alpha$	74	0.08 ± 0.27	None needed (< 0.5 ppm)
$^{13}C_\beta$	46	0.11 ± 0.12	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	95	0.34 ± 0.46	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 61%, i.e. 849 atoms were assigned a chemical shift out of a possible 1395. 0 out of 15 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	356/515 (69%)	198/209 (95%)	70/206 (34%)	88/100 (88%)
Sidechain	462/757 (61%)	380/487 (78%)	78/247 (32%)	4/23 (17%)
Aromatic	31/123 (25%)	30/61 (49%)	0/49 (0%)	1/13 (8%)
Overall	849/1395 (61%)	608/757 (80%)	148/502 (29%)	93/136 (68%)

7.1.4   Statistically unusual chemical shifts ⓘ

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	123	VAL	HG11	-2.40	$-0.48 - 2.12$	-12.4
1	A	123	VAL	HG12	-2.40	$-0.48 - 2.12$	-12.4
1	A	123	VAL	HG13	-2.40	$-0.48 - 2.12$	-12.4
1	A	30	LEU	HD11	-2.33	$-0.61 - 2.12$	-11.3
1	A	30	LEU	HD12	-2.33	$-0.61 - 2.12$	-11.3
1	A	30	LEU	HD13	-2.33	$-0.61 - 2.12$	-11.3
1	A	44	HIS	HB3	7.26	$1.18 - 4.91$	11.3
1	A	81	ILE	HD11	-1.38	$-0.72 - 2.09$	-7.4
1	A	81	ILE	HD12	-1.38	$-0.72 - 2.09$	-7.4
1	A	81	ILE	HD13	-1.38	$-0.72 - 2.09$	-7.4
1	A	30	LEU	HD21	-1.08	$-0.65 - 2.13$	-6.5
1	A	30	LEU	HD22	-1.08	$-0.65 - 2.13$	-6.5
1	A	30	LEU	HD23	-1.08	$-0.65 - 2.13$	-6.5
1	A	59	ALA	HB1	-0.17	$0.14 - 2.58$	-6.2
1	A	59	ALA	HB2	-0.17	$0.14 - 2.58$	-6.2
1	A	59	ALA	HB3	-0.17	$0.14 - 2.58$	-6.2
1	A	63	PHE	HB3	0.56	$1.03 - 4.85$	-6.2

Continued on next page...

Continued from previous page...

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	28	LEU	HD21	-0.99	-0.65 − 2.13	-6.2
1	A	28	LEU	HD22	-0.99	-0.65 − 2.13	-6.2
1	A	28	LEU	HD23	-0.99	-0.65 − 2.13	-6.2
1	A	86	PRO	HD3	1.34	1.76 − 5.48	-6.1
1	A	69	SER	H	11.65	5.45 − 11.10	6.0
1	A	30	LEU	HG	-0.41	-0.13 − 3.16	-5.8
1	A	28	LEU	HD11	-0.81	-0.61 − 2.12	-5.8
1	A	28	LEU	HD12	-0.81	-0.61 − 2.12	-5.8
1	A	28	LEU	HD13	-0.81	-0.61 − 2.12	-5.8
1	A	29	ILE	CG1	18.21	19.24 − 36.26	-5.6
1	A	47	GLY	HA2	5.98	2.15 − 5.77	5.6
1	A	54	GLN	HB3	0.57	0.71 − 3.33	-5.5

7.1.5   Random Coil Index (RCI) plots   ⓘ

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:
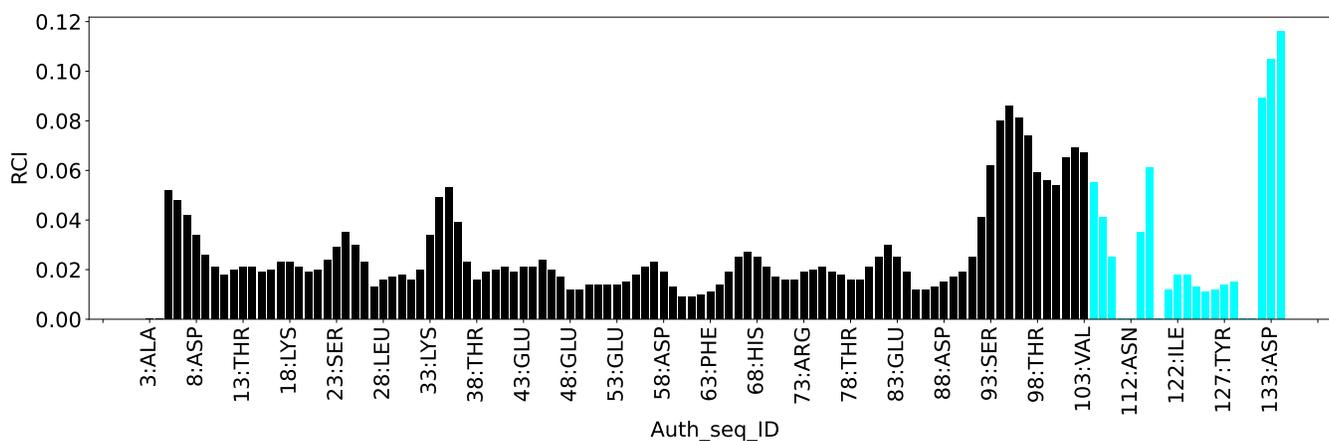
8 NMR restraints analysis ⓘ

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	1649
Intra-residue ($|i-j|$=0)	332
Sequential ($|i-j|$=1)	483
Medium range ($|i-j|$>1 and $|i-j|$<5)	340
Long range ($|i-j|$≥5)	455
Inter-chain	0
Hydrogen bond restraints	39
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	38
Number of restraints per residue	12.3
Number of long range restraints per residue[1]	3.5

[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)	59.2	0.2
0.2-0.5 (Medium)	106.7	0.5
>0.5 (Large)	128.3	5.42

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

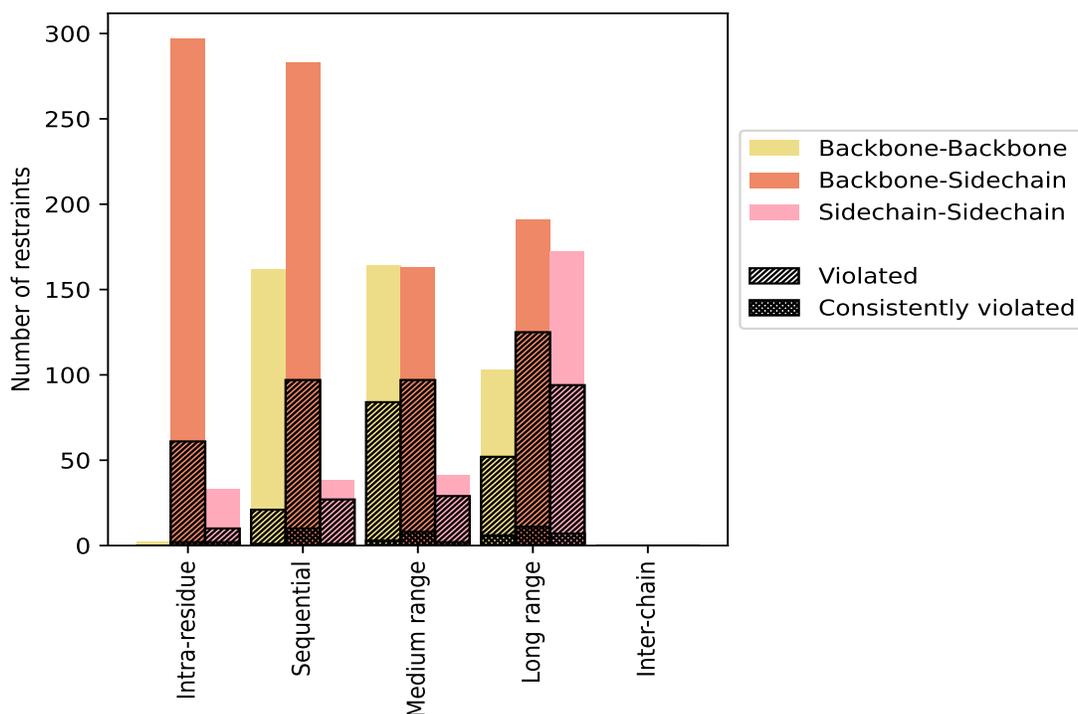# 9 Distance violation analysis ⓘ

9.1 Summary of distance violations ⓘ

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)	332	20.1	71	21.4	4.3	4	1.2	0.2
Backbone-Backbone	2	0.1	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	297	18.0	61	20.5	3.7	2	0.7	0.1
Sidechain-Sidechain	33	2.0	10	30.3	0.6	2	6.1	0.1
Sequential (|i-j|=1)	483	29.3	145	30.0	8.8	12	2.5	0.7
Backbone-Backbone	162	9.8	21	13.0	1.3	1	0.6	0.1
Backbone-Sidechain	283	17.2	97	34.3	5.9	10	3.5	0.6
Sidechain-Sidechain	38	2.3	27	71.1	1.6	1	2.6	0.1
Medium range (|i-j|>1 & |i-j|<5)	340	20.6	192	56.5	11.6	11	3.2	0.7
Backbone-Backbone	136	8.2	66	48.5	4.0	1	0.7	0.1
Backbone-Sidechain	163	9.9	97	59.5	5.9	8	4.9	0.5
Sidechain-Sidechain	41	2.5	29	70.7	1.8	2	4.9	0.1
Long range (|i-j|≥5)	455	27.6	269	59.1	16.3	24	5.3	1.5
Backbone-Backbone	92	5.6	50	54.3	3.0	6	6.5	0.4
Backbone-Sidechain	191	11.6	125	65.4	7.6	11	5.8	0.7
Sidechain-Sidechain	172	10.4	94	54.7	5.7	7	4.1	0.4
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	39	2.4	20	51.3	1.2	2	5.1	0.1
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	1649	100.0	697	42.3	42.3	53	3.2	3.2
Backbone-Backbone	431	26.1	157	36.4	9.5	10	2.3	0.6
Backbone-Sidechain	934	56.6	380	40.7	23.0	31	3.3	1.9
Sidechain-Sidechain	284	17.2	160	56.3	9.7	12	4.2	0.7

[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)	SD⁶ (Å)	Median (Å)
	IR¹	SQ²	MR³	LR⁴	IC⁵	Total				
1	29	53	66	125	0	273	0.6	3.32	0.5	0.44
2	28	60	86	117	0	291	0.61	3.42	0.58	0.4
3	34	59	101	133	0	327	0.6	3.73	0.53	0.42
4	30	71	92	124	0	317	0.58	3.02	0.48	0.4
5	34	58	72	122	0	286	0.62	3.52	0.51	0.42
6	31	57	81	129	0	298	0.61	3.37	0.49	0.44
7	24	52	66	116	0	258	0.56	2.64	0.47	0.39
8	30	58	73	146	0	307	0.62	2.83	0.48	0.44
9	24	63	75	140	0	302	0.69	5.42	0.69	0.49
10	29	65	63	123	0	280	0.55	2.18	0.41	0.42
11	28	59	68	130	0	285	0.59	2.48	0.45	0.47

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	62	80	131	0	298	0.67	3.42	0.59	0.45
13	32	66	88	150	0	336	0.61	3.87	0.54	0.42
14	28	52	76	136	0	292	0.6	2.18	0.43	0.47
15	31	56	80	126	0	293	0.6	4.05	0.53	0.41
16	29	59	88	118	0	294	0.65	3.39	0.6	0.42
17	27	58	80	114	0	279	0.55	3.46	0.45	0.41
18	29	58	71	115	0	273	0.6	4.53	0.56	0.43
19	26	53	83	121	0	283	0.57	2.32	0.42	0.46
20	32	57	69	130	0	288	0.62	4.61	0.57	0.43
21	27	60	71	125	0	283	0.68	4.39	0.63	0.44
22	31	60	86	144	0	321	0.6	3.02	0.52	0.41
23	32	56	80	134	0	302	0.66	3.51	0.58	0.44

[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 ⓘ
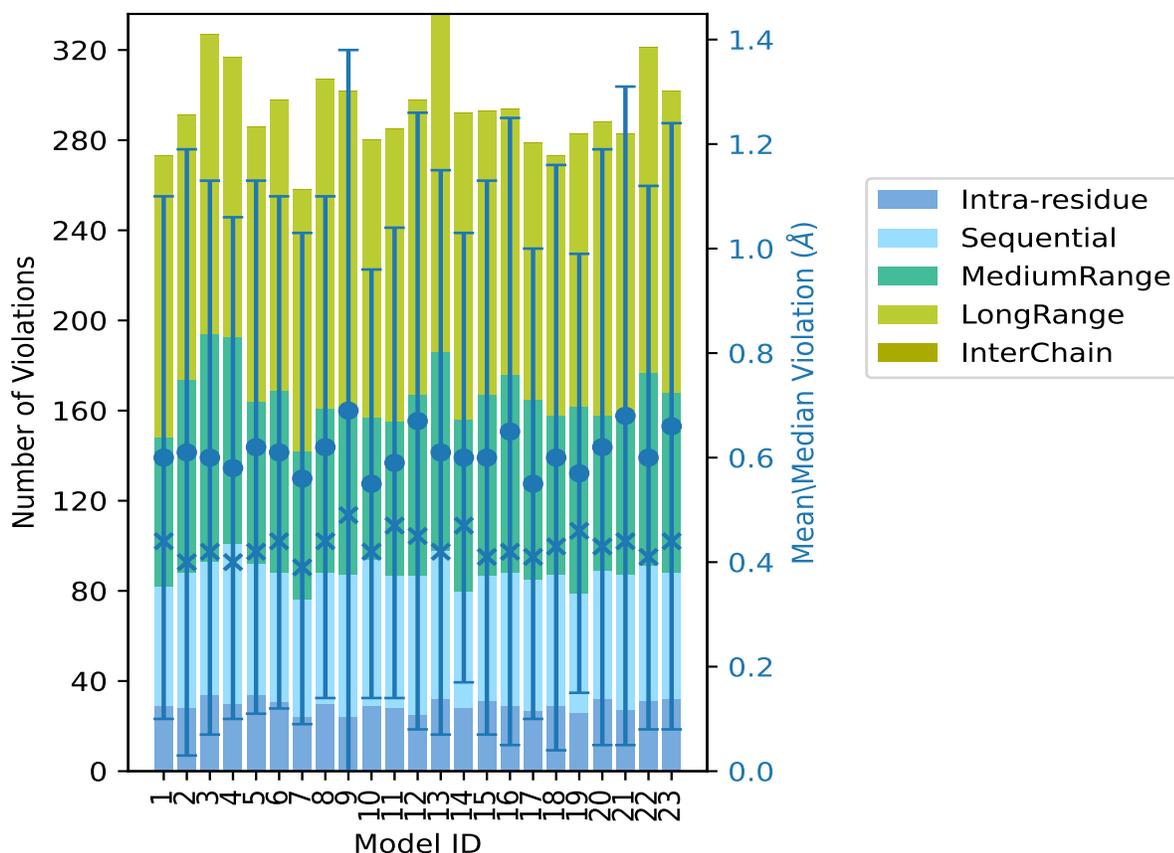
The mean(dot),median(x) and the standard deviation are shown in blue with respect to the y axis on the right

## 9.3 Distance violation statistics for the ensemble ⓘ

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, 933(IR:261, SQ:338, MR:148, LR:186, 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]	%
11	24	34	27	0	96	1	4.3
7	15	22	29	0	73	2	8.7
5	14	17	13	0	49	3	13.0
2	14	11	14	0	41	4	17.4
5	4	4	7	0	20	5	21.7
5	5	8	13	0	31	6	26.1
3	3	7	14	0	27	7	30.4
3	3	8	10	0	24	8	34.8
1	3	10	8	0	22	9	39.1
3	1	1	9	0	14	10	43.5
1	6	5	9	0	21	11	47.8
0	2	5	4	0	11	12	52.2
2	2	8	8	0	20	13	56.5
1	4	4	7	0	16	14	60.9
2	4	2	6	0	14	15	65.2
1	5	5	6	0	17	16	69.6
3	3	3	5	0	14	17	73.9
2	3	6	9	0	20	18	78.3
3	3	8	10	0	24	19	82.6
1	4	7	6	0	18	20	87.0
2	4	4	16	0	26	21	91.3
4	7	2	15	0	28	22	95.7
4	12	11	24	0	51	23	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 ⓘ
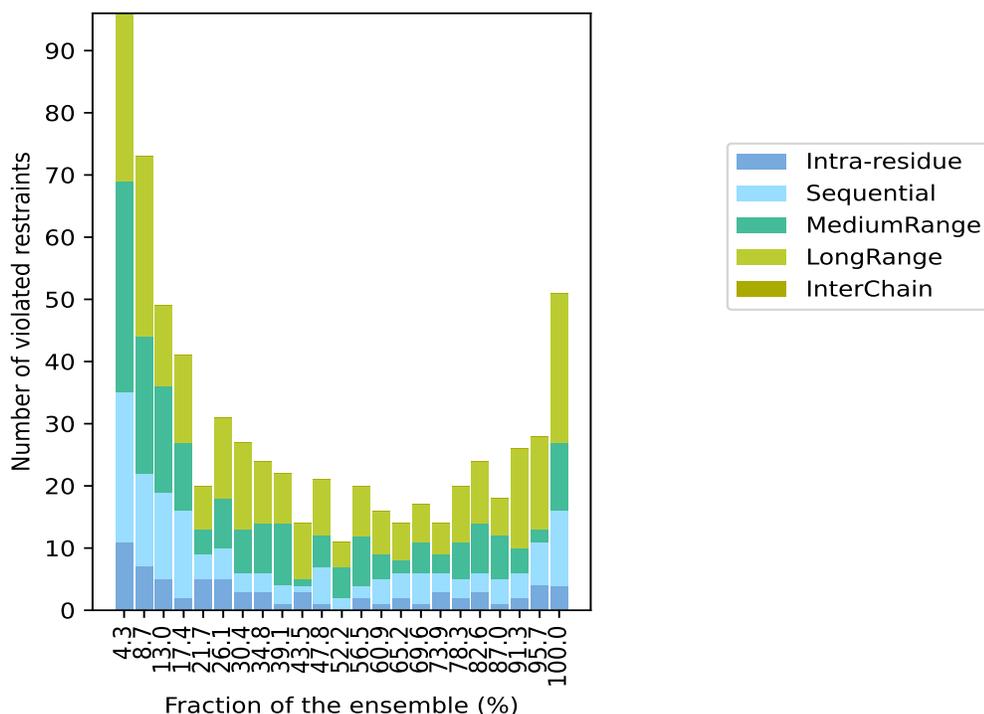
9.4 Most violated distance restraints in the ensemble ⓘ

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
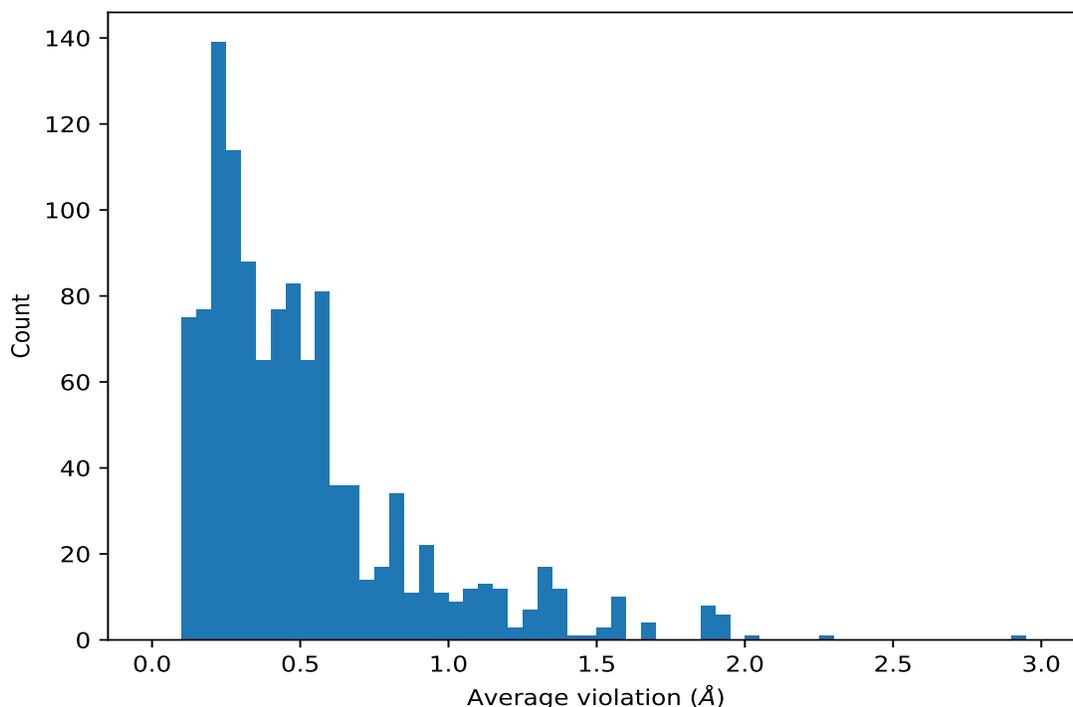
9.4.2　Table: Most violated distance restraints ⓘ

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)
(1,949)	1:A:54:GLN:HE22	1:A:62:ASN:HD22	23	2.91	1.37	3.32
(1,893)	1:A:35:TYR:H	1:A:81:ILE:HG23	23	2.3	0.32	2.34
(1,1351)	1:A:12:TYR:H	1:A:14:LEU:HD11	23	1.87	0.45	1.99
(1,1351)	1:A:12:TYR:H	1:A:14:LEU:HD12	23	1.87	0.45	1.99
(1,1351)	1:A:12:TYR:H	1:A:14:LEU:HD13	23	1.87	0.45	1.99
(1,1351)	1:A:12:TYR:H	1:A:14:LEU:HD21	23	1.87	0.45	1.99
(1,1351)	1:A:12:TYR:H	1:A:14:LEU:HD22	23	1.87	0.45	1.99
(1,1351)	1:A:12:TYR:H	1:A:14:LEU:HD23	23	1.87	0.45	1.99
(1,938)	1:A:14:LEU:HD21	1:A:88:ASP:H	23	1.56	0.33	1.51
(1,938)	1:A:14:LEU:HD22	1:A:88:ASP:H	23	1.56	0.33	1.51
(1,938)	1:A:14:LEU:HD23	1:A:88:ASP:H	23	1.56	0.33	1.51
(1,696)	1:A:81:ILE:HG23	1:A:82:GLY:H	23	1.48	0.23	1.52
(1,1243)	1:A:17:ILE:HG21	1:A:27:TRP:HH2	23	1.43	0.5	1.43
(1,948)	1:A:62:ASN:H	1:A:62:ASN:HD22	23	1.4	0.54	1.66
(1,1068)	1:A:20:HIS:H	1:A:21:ASN:HD21	23	1.35	0.47	1.39
(1,1246)	1:A:12:TYR:HB2	1:A:27:TRP:HH2	23	1.32	0.45	1.32

Continued from previous page...

Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,984)	1:A:36:ASP:H	1:A:81:ILE:HG22	23	1.32	0.26	1.31

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

9.5 All violated distance restraints ⓘ

9.5.1 Histogram : Distribution of distance violations ⓘ

The following histogram shows the distribution of the absolute value of the violation for all violated restraints in the ensemble.
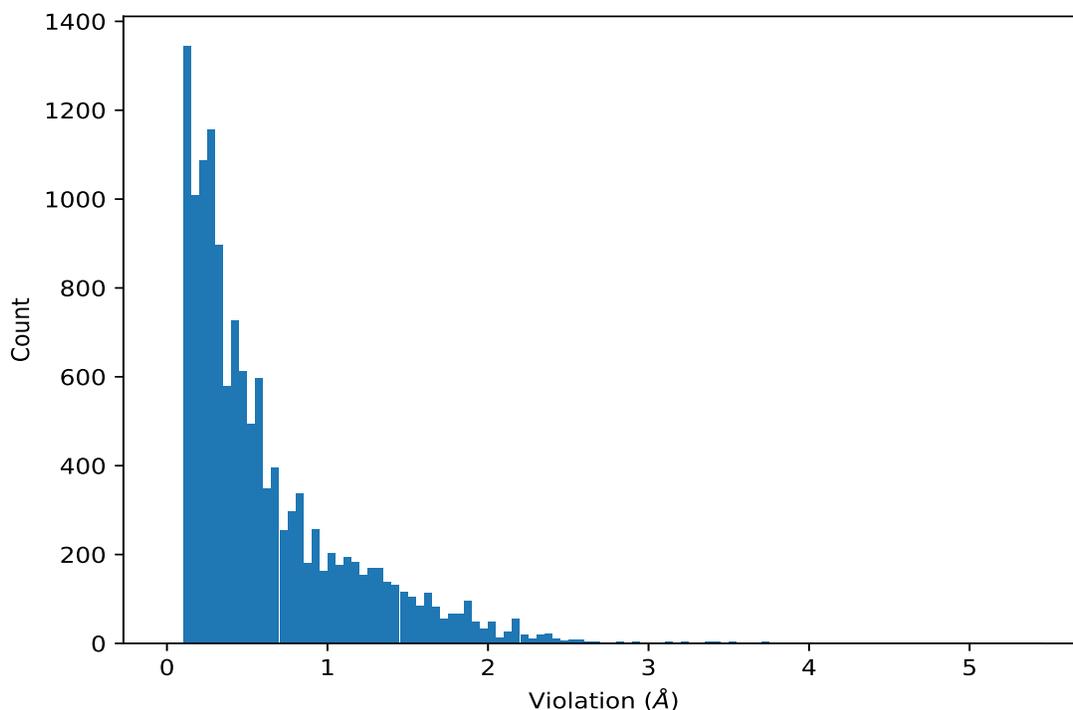
9.5.2 Table : All distance violations ⓘ

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)
(1,949)	1:A:54:GLN:HE22	1:A:62:ASN:HD22	9	5.42
(1,949)	1:A:54:GLN:HE22	1:A:62:ASN:HD22	20	4.61
(1,949)	1:A:54:GLN:HE22	1:A:62:ASN:HD22	18	4.53
(1,949)	1:A:54:GLN:HE22	1:A:62:ASN:HD22	21	4.39

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,449)	1:A:54:GLN:HE22	1:A:62:ASN:HD21	9	4.14
(1,441)	1:A:54:GLN:HE21	1:A:62:ASN:HD22	9	4.11
(1,949)	1:A:54:GLN:HE22	1:A:62:ASN:HD22	15	4.05
(1,441)	1:A:54:GLN:HE21	1:A:62:ASN:HD22	20	3.88
(1,449)	1:A:54:GLN:HE22	1:A:62:ASN:HD21	13	3.87
(1,949)	1:A:54:GLN:HE22	1:A:62:ASN:HD22	13	3.81

10 Dihedral-angle violation analysis ⓘ

No dihedral-angle restraints found