wwPDB NMR Structure Validation Summary Report (i)

Jun 3, 2023 – 05:06 PM EDT

PDB ID	:	2MOL
BMRB ID	:	19946
Title	:	3D NMR structure of the cytoplasmic rhodanese domain of the full-length inner membrane protein YgaP from Escherichia coli
Authors	:	Eichmann, C.; Tzitzilonis, C.; Bordignon, E.; Maslennikov, I.; Choe, S.; Riek, R.
Deposited on	:	2014-04-27

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:

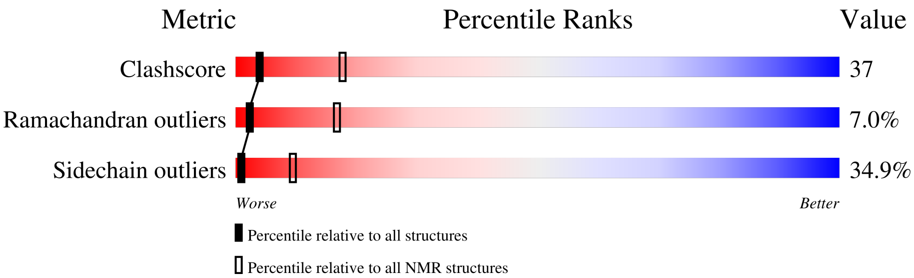
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal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

1 Overall quality at a glance (i)

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

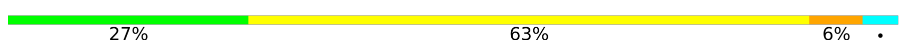
The overall completeness of chemical shifts assignment is 78%.

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	108	

2 Ensemble composition and analysis (i)

This entry contains 10 models. Model 2 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *fewest violations*.

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:105 (104)	1.29	2

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 1 clusters and 7 single-model clusters were found.

Cluster number	**Models**
1	2, 5, 6
Single-model clusters	1; 3; 4; 7; 8; 9; 10

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1528 atoms, of which 704 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Inner membrane protein YgaP.

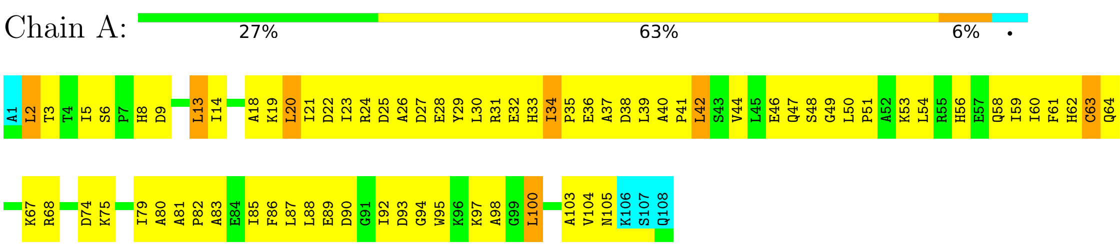
Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	108	1528	519	704	147	157	1	0

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.

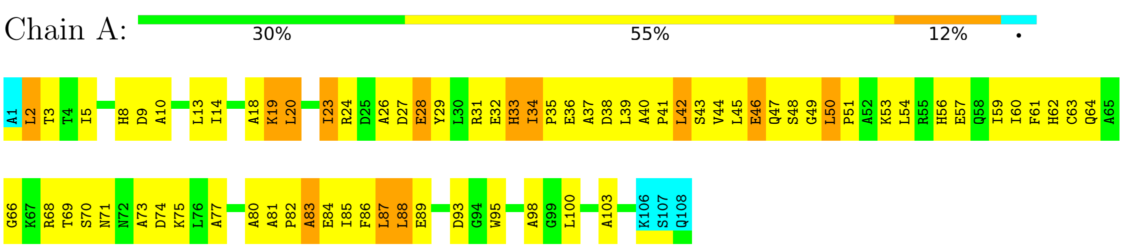
- Molecule 1: Inner membrane protein YgaP



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

The representative model is number 2. Colouring as in section 4.1 above.

- Molecule 1: Inner membrane protein YgaP



5 Refinement protocol and experimental data overview (i)

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

Of the 100 calculated structures, 10 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	1113
Number of shifts mapped to atoms	1004
Number of unparsed shifts	0
Number of shifts with mapping errors	109
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	78%

6 Model quality (i)

6.1 Standard geometry (i)

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 (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	795	681	804	60±11
All	All	7950	6810	8040	596

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 37.

5 of 358 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:98:ALA:HB3	1:A:100:LEU:HD11	1.07	1.17	6	3
1:A:98:ALA:HB3	1:A:100:LEU:CD1	0.98	1.87	10	3
1:A:59:ILE:HD11	1:A:61:PHE:CD2	0.97	1.94	10	1
1:A:34:ILE:HG22	1:A:37:ALA:HB2	0.94	1.39	9	2
1:A:3:THR:HG21	1:A:86:PHE:CE2	0.87	2.04	6	1

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	104/108 (96%)	72±3 (69±3%)	25±3 (24±3%)	7±2 (7±2%)	2	17
All	All	1040/1080 (96%)	721 (69%)	246 (24%)	73 (7%)	2	17

5 of 29 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	36	GLU	7
1	A	23	ILE	6
1	A	90	ASP	5
1	A	64	GLN	5
1	A	63	CYS	3

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	82/85 (96%)	53±2 (65±2%)	29±2 (35±2%)	1	9
All	All	820/850 (96%)	534 (65%)	286 (35%)	1	9

5 of 61 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	19	LYS	10
1	A	27	ASP	10
1	A	47	GLN	10
1	A	68	ARG	10
1	A	8	HIS	9

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 (i)

The completeness of assignment taking into account all chemical shift lists is 78% for the well-defined parts and 76% 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 (i)

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	1113
Number of shifts mapped to atoms	1004
Number of unparsed shifts	0
Number of shifts with mapping errors	109
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	19

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 109) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	2	LEU	HB2	1.553	.	.
1	A	5	ILE	HG12	1.586	.	.
1	A	7	PRO	HB2	2.294	.	.
1	A	7	PRO	HG2	2.108	.	.
1	A	7	PRO	HD2	3.896	.	.
1	A	8	HIS	HB2	2.976	.	.
1	A	9	ASP	HB2	2.829	.	.
1	A	11	GLN	HB2	2.054	.	.
1	A	11	GLN	HG2	2.296	.	.
1	A	12	GLU	HB2	2.089	.	.
1	A	12	GLU	HG2	2.265	.	.
1	A	13	LEU	HB2	1.511	.	.
1	A	14	ILE	HG12	1.59	.	.
1	A	16	ARG	HB2	2.144	.	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	16	ARG	HG2	1.673	.	.
1	A	16	ARG	HD2	2.64	.	.
1	A	19	LYS	HB2	1.578	.	.
1	A	19	LYS	HG2	1.262	.	.
1	A	19	LYS	HD2	1.531	.	.
1	A	19	LYS	HE2	2.763	.	.
1	A	20	LEU	HB2	1.681	.	.
1	A	21	ILE	HG12	1.225	.	.
1	A	22	ASP	HB2	2.678	.	.
1	A	23	ILE	HG12	1.062	.	.
1	A	24	ARG	HB2	2.212	.	.
1	A	24	ARG	HG2	1.892	.	.
1	A	25	ASP	HB2	2.866	.	.
1	A	27	ASP	HB2	2.903	.	.
1	A	28	GLU	HB2	2.168	.	.
1	A	28	GLU	HG2	2.332	.	.
1	A	29	TYR	HB2	2.941	.	.
1	A	30	LEU	HB2	1.718	.	.
1	A	31	ARG	HB2	1.871	.	.
1	A	31	ARG	HG2	1.765	.	.
1	A	31	ARG	HD2	3.243	.	.
1	A	32	GLU	HB2	1.995	.	.
1	A	32	GLU	HG2	1.857	.	.
1	A	33	HIS	HB2	2.877	.	.
1	A	34	ILE	HG12	0.931	.	.
1	A	35	PRO	HB2	2.113	.	.
1	A	35	PRO	HG2	2.005	.	.
1	A	35	PRO	HD2	3.997	.	.
1	A	36	GLU	HB2	1.876	.	.
1	A	36	GLU	HG2	2.262	.	.
1	A	38	ASP	HB2	2.388	.	.
1	A	39	LEU	HB2	1.116	.	.
1	A	41	PRO	HB2	2.393	.	.
1	A	41	PRO	HG2	2.2	.	.
1	A	41	PRO	HD2	3.437	.	.
1	A	42	LEU	HB2	1.805	.	.
1	A	43	SER	HB2	3.879	.	.
1	A	45	LEU	HB2	1.899	.	.
1	A	46	GLU	HB2	2.009	.	.
1	A	46	GLU	HG2	2.333	.	.
1	A	47	GLN	HB2	2.075	.	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	47	GLN	HG2	2.389	.	.
1	A	48	SER	HB2	3.836	.	.
1	A	50	LEU	HB2	1.66	.	.
1	A	51	PRO	HB2	2.324	.	.
1	A	51	PRO	HG2	1.994	.	.
1	A	53	LYS	HB2	1.814	.	.
1	A	53	LYS	HG2	1.403	.	.
1	A	53	LYS	HD2	1.666	.	.
1	A	53	LYS	HE2	2.913	.	.
1	A	54	LEU	HB2	1.702	.	.
1	A	55	ARG	HB2	1.718	.	.
1	A	55	ARG	HG2	1.625	.	.
1	A	55	ARG	HD2	3.13	.	.
1	A	56	HIS	HB2	3.001	.	.
1	A	57	GLU	HB2	2.167	.	.
1	A	57	GLU	HG2	2.485	.	.
1	A	58	GLN	HB2	2.16	.	.
1	A	58	GLN	HG2	2.04	.	.
1	A	59	ILE	HG12	1.359	.	.
1	A	60	ILE	HG12	1.473	.	.
1	A	61	PHE	HB2	2.789	.	.
1	A	62	HIS	HB2	1.162	.	.
1	A	63	CYS	HB2	3.672	.	.
1	A	67	LYS	HB2	2.136	.	.
1	A	70	SER	HB2	3.929	.	.
1	A	71	ASN	HB2	2.815	.	.
1	A	72	ASN	HB2	2.607	.	.
1	A	74	ASP	HB2	2.596	.	.
1	A	75	LYS	HB2	1.77	.	.
1	A	75	LYS	HG2	1.426	.	.
1	A	75	LYS	HD2	1.583	.	.
1	A	75	LYS	HE2	2.938	.	.
1	A	76	LEU	HB2	1.526	.	.
1	A	79	ILE	HG12	1.763	.	.
1	A	82	PRO	HB2	2.522	.	.
1	A	82	PRO	HG2	1.878	.	.
1	A	82	PRO	HD2	3.493	.	.
1	A	84	GLU	HB2	2.011	.	.
1	A	84	GLU	HG2	2.204	.	.
1	A	85	ILE	HG12	1.469	.	.
1	A	86	PHE	HB2	2.793	.	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	87	LEU	HB2	1.831	.	.
1	A	88	LEU	HB2	2.048	.	.
1	A	89	GLU	HB2	1.896	.	.
1	A	90	ASP	HB2	2.938	.	.
1	A	92	ILE	HG12	1.236	.	.
1	A	93	ASP	HB2	3.044	.	.
1	A	95	TRP	HB2	2.884	.	.
1	A	96	LYS	HB2	1.851	.	.
1	A	97	LYS	HB2	1.886	.	.
1	A	100	LEU	HB2	0.232	.	.
1	A	101	PRO	HB2	2.252	.	.
1	A	101	PRO	HD2	3.675	.	.
1	A	105	ASN	HB2	2.86	.	.

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	95	0.81 ± 0.09	Should be applied
$^{13}C_{\beta}$	90	0.70 ± 0.13	Should be applied
$^{13}C'$	0	—	None (insufficient data)
^{15}N	100	-0.39 ± 0.58	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 78%, i.e. 1095 atoms were assigned a chemical shift out of a possible 1404. 0 out of 16 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	395/514 (77%)	204/208 (98%)	94/208 (45%)	97/98 (99%)
Sidechain	668/821 (81%)	464/537 (86%)	195/256 (76%)	9/28 (32%)
Aromatic	32/69 (46%)	18/36 (50%)	12/28 (43%)	2/5 (40%)
Overall	1095/1404 (78%)	686/781 (88%)	301/492 (61%)	108/131 (82%)

7.1.4 Statistically unusual chemical shifts (i)

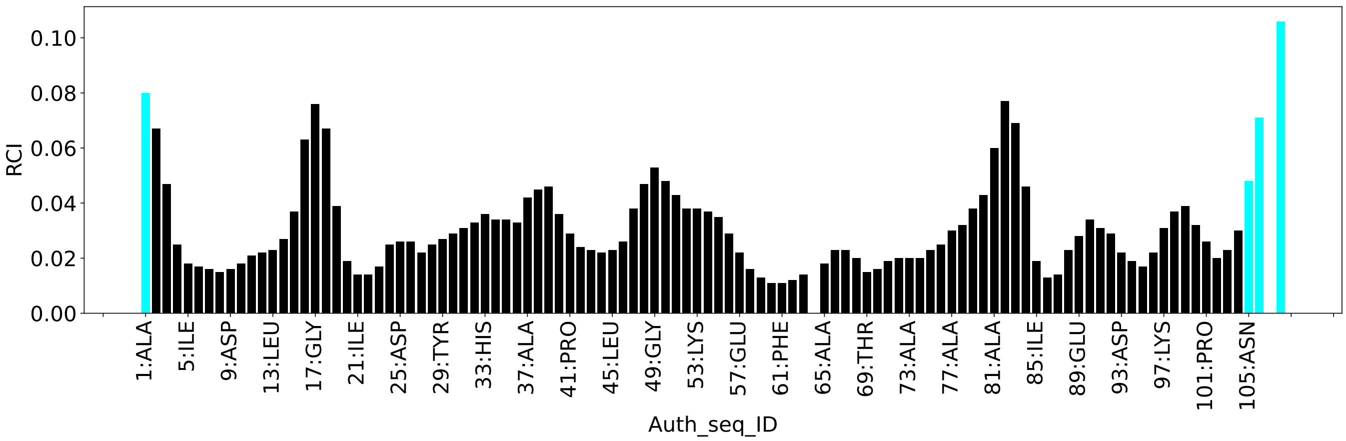
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	95	TRP	CH2	-6.33	116.19 – 131.43	-85.4
1	A	95	TRP	CZ3	-8.70	113.48 – 129.28	-82.3
1	A	95	TRP	CE3	-10.87	111.58 – 129.41	-73.7
1	A	86	PHE	CD2	51.72	125.53 – 137.61	-66.1
1	A	62	HIS	CD2	26.09	103.95 – 136.66	-28.8
1	A	24	ARG	NE	113.46	76.53 – 92.65	17.9
1	A	62	HIS	HD2	0.50	4.65 – 9.35	-13.8
1	A	24	ARG	NH2	107.99	57.68 – 87.89	11.7
1	A	95	TRP	NE1	106.17	118.53 – 139.98	-10.8
1	A	34	ILE	HD11	-0.97	-0.72 – 2.09	-5.9
1	A	34	ILE	HD12	-0.97	-0.72 – 2.09	-5.9
1	A	34	ILE	HD13	-0.97	-0.72 – 2.09	-5.9
1	A	100	LEU	HD21	-0.83	-0.65 – 2.13	-5.6
1	A	100	LEU	HD22	-0.83	-0.65 – 2.13	-5.6
1	A	100	LEU	HD23	-0.83	-0.65 – 2.13	-5.6
1	A	62	HIS	HB2	1.16	1.36 – 4.85	-5.6
1	A	39	LEU	HB3	-0.40	-0.26 – 3.31	-5.4
1	A	100	LEU	HB3	-0.40	-0.26 – 3.31	-5.4
1	A	64	GLN	NE2	120.71	103.38 – 120.35	5.2

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:


RCI
0.00
0.02
0.04
0.06
0.08
0.10
1:ALA
5:ILE
9:ASP
13:LEU
17:GLY
21:ILE
25:ASP
29:TYR
33:HIS
37:ALA
41:PRO
45:LEU
49:GLY
53:LYS
57:GLU
61:PHE
65:ALA
69:THR
73:ALA
77:ALA
81:ALA
85:ILE
89:GLU
93:ASP
97:LYS
101:PRO
105:ASN
Auth_seq_ID

8 NMR restraints analysis (i)

8.1 Conformationally restricting restraints (i)

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	1201
Intra-residue (|i-j|=0)	384
Sequential (|i-j|=1)	353
Medium range (|i-j|>1 and |i-j|<5)	203
Long range (|i-j|≥5)	261
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	482
Number of restraints per residue	11.1
Number of long range restraints per residue[1]	2.4

[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 (i)

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 (i)

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)	21.9	0.2
0.2-0.5 (Medium)	49.2	0.5
>0.5 (Large)	22.1	4.55

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)

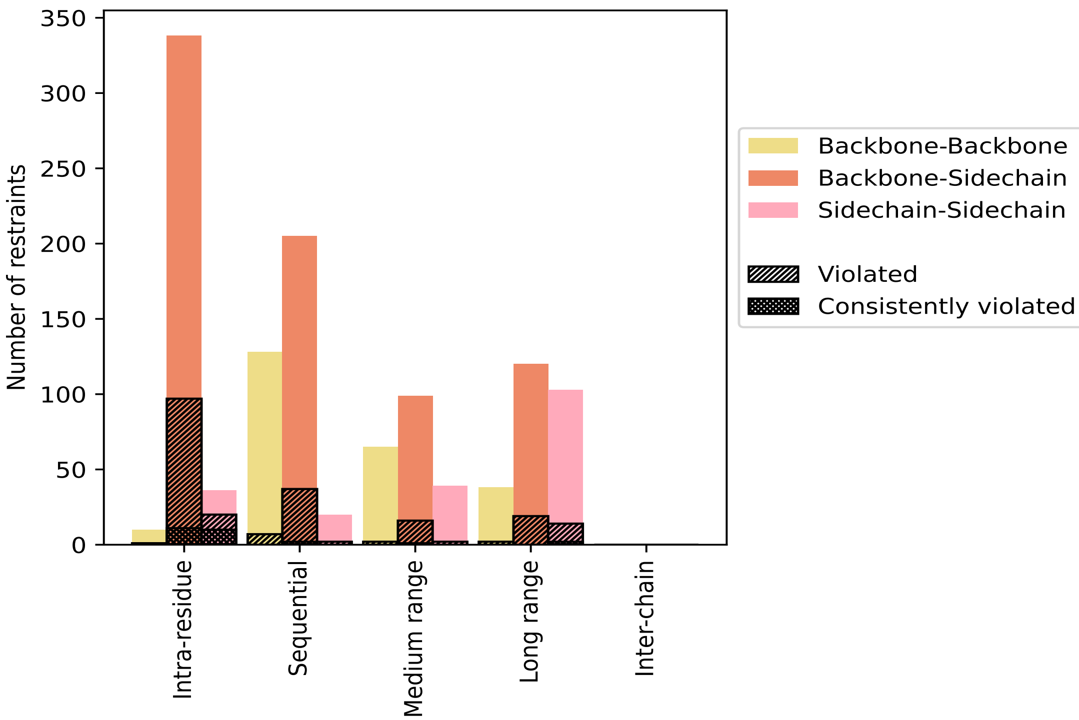
9.1 Summary of distance violations (i)

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)	384	32.0	118	30.7	9.8	21	5.5	1.7
Backbone-Backbone	10	0.8	1	10.0	0.1	0	0.0	0.0
Backbone-Sidechain	338	28.1	97	28.7	8.1	11	3.3	0.9
Sidechain-Sidechain	36	3.0	20	55.6	1.7	10	27.8	0.8
Sequential (|i-j|=1)	353	29.4	46	13.0	3.8	2	0.6	0.2
Backbone-Backbone	128	10.7	7	5.5	0.6	0	0.0	0.0
Backbone-Sidechain	205	17.1	37	18.0	3.1	2	1.0	0.2
Sidechain-Sidechain	20	1.7	2	10.0	0.2	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	203	16.9	20	9.9	1.7	1	0.5	0.1
Backbone-Backbone	65	5.4	2	3.1	0.2	0	0.0	0.0
Backbone-Sidechain	99	8.2	16	16.2	1.3	1	1.0	0.1
Sidechain-Sidechain	39	3.2	2	5.1	0.2	0	0.0	0.0
Long range (|i-j|≥5)	261	21.7	35	13.4	2.9	2	0.8	0.2
Backbone-Backbone	38	3.2	2	5.3	0.2	0	0.0	0.0
Backbone-Sidechain	120	10.0	19	15.8	1.6	0	0.0	0.0
Sidechain-Sidechain	103	8.6	14	13.6	1.2	2	1.9	0.2
Inter-chain	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Hydrogen bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	1201	100.0	219	18.2	18.2	26	2.2	2.2
Backbone-Backbone	241	20.1	12	5.0	1.0	0	0.0	0.0
Backbone-Sidechain	762	63.4	169	22.2	14.1	14	1.8	1.2
Sidechain-Sidechain	198	16.5	38	19.2	3.2	12	6.1	1.0

[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 (i)



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 (i)

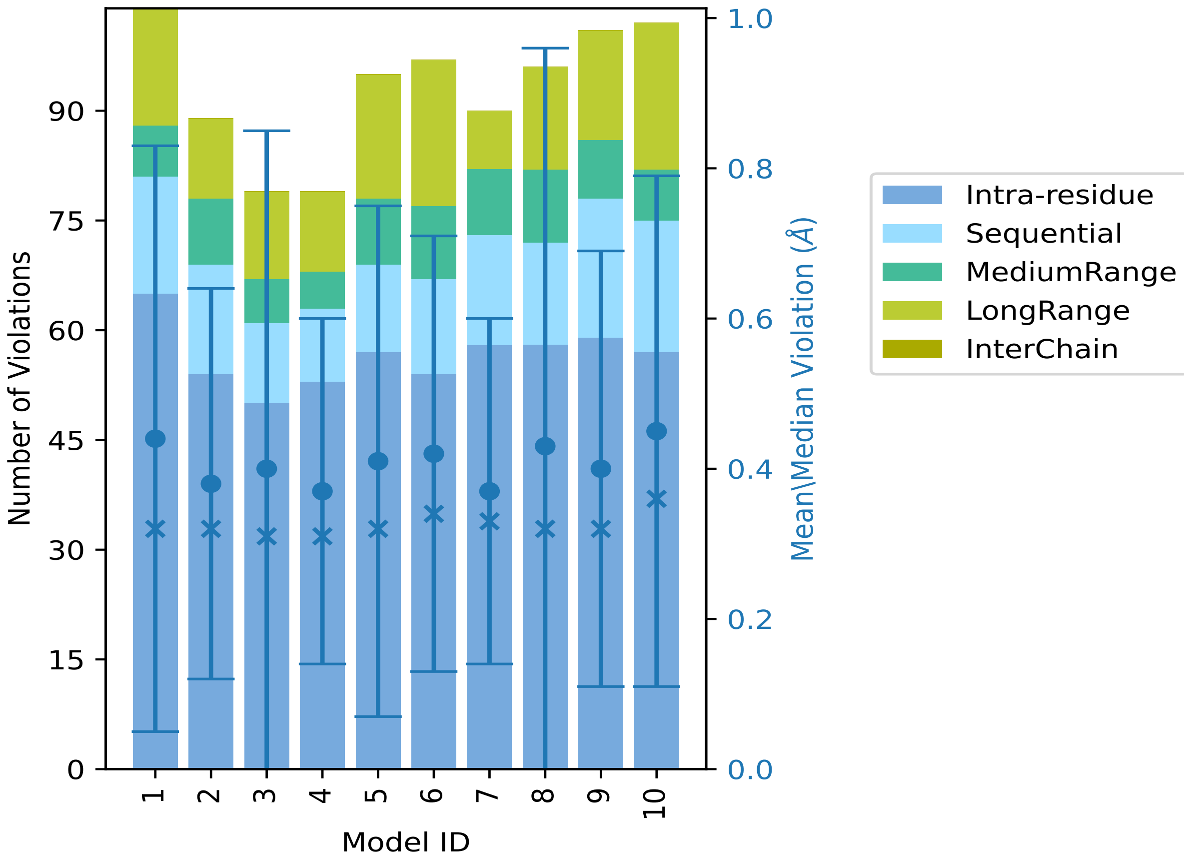
The following table provides the distance violation statistics for each model in the ensemble. Violations less than 0.1 Å are not included in the statistics.

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	65	16	7	16	0	104	0.44	3.11	0.39	0.32
2	54	15	9	11	0	89	0.38	1.64	0.26	0.32
3	50	11	6	12	0	79	0.4	3.34	0.45	0.31
4	53	10	5	11	0	79	0.37	1.4	0.23	0.31
5	57	12	9	17	0	95	0.41	2.68	0.34	0.32
6	54	13	10	20	0	97	0.42	1.55	0.29	0.34
7	58	15	9	8	0	90	0.37	1.65	0.23	0.33
8	58	14	10	14	0	96	0.43	4.55	0.53	0.32
9	59	19	8	15	0	101	0.4	1.46	0.29	0.32
10	57	18	7	20	0	102	0.45	2.24	0.34	0.36

[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, 982(IR:266, SQ:307, MR:183, LR:226, 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]	%
29	20	3	9	0	61	1	10.0
13	7	4	6	0	30	2	20.0
13	5	4	4	0	26	3	30.0
8	3	3	3	0	17	4	40.0

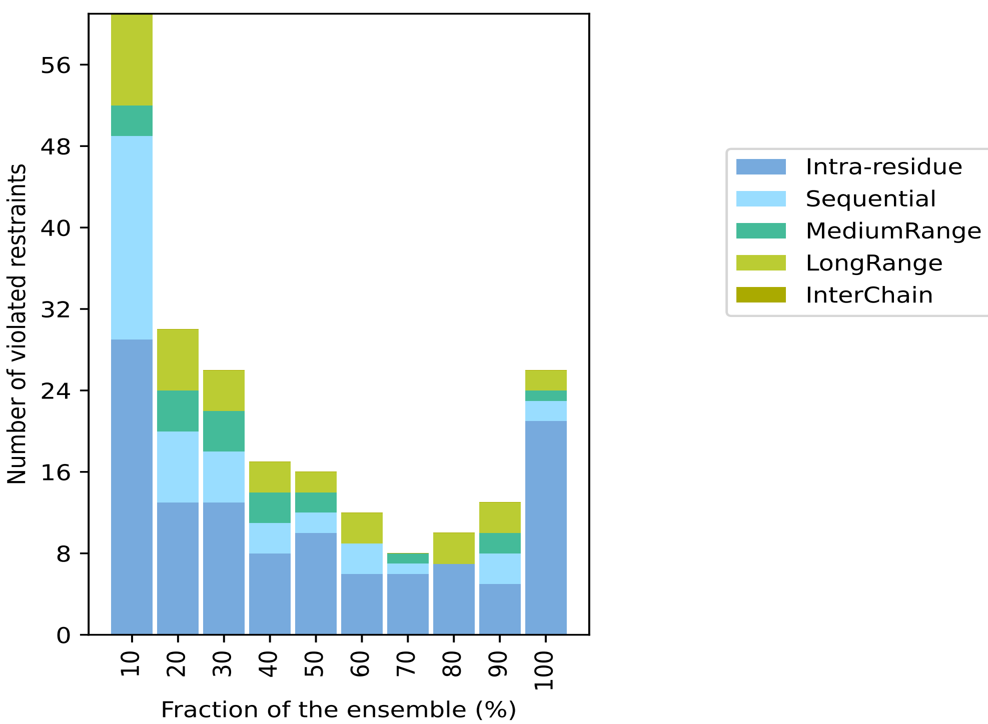
Continued on next page...

Continued from previous page...

Number of violated restraints						Fraction of the ensemble	
IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total	Count[6]	%
10	2	2	2	0	16	5	50.0
6	3	0	3	0	12	6	60.0
6	1	1	0	0	8	7	70.0
7	0	0	3	0	10	8	80.0
5	3	2	3	0	13	9	90.0
21	2	1	2	0	26	10	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

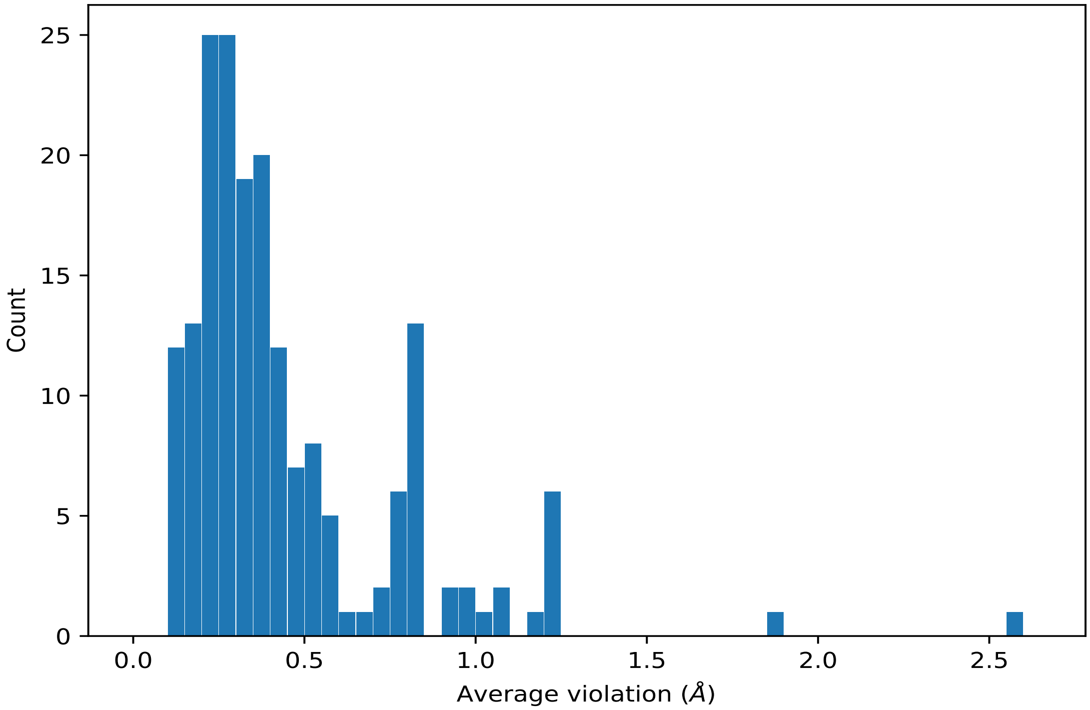
9.3.1 Bar graph : Distance violation statistics for the ensemble (i)



9.4 Most violated distance restraints in the ensemble (i)

9.4.1 Histogram : Distribution of mean distance violations (i)

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



9.4.2 Table: Most violated distance restraints (i)

The following table provides the mean and the standard deviation of the 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.

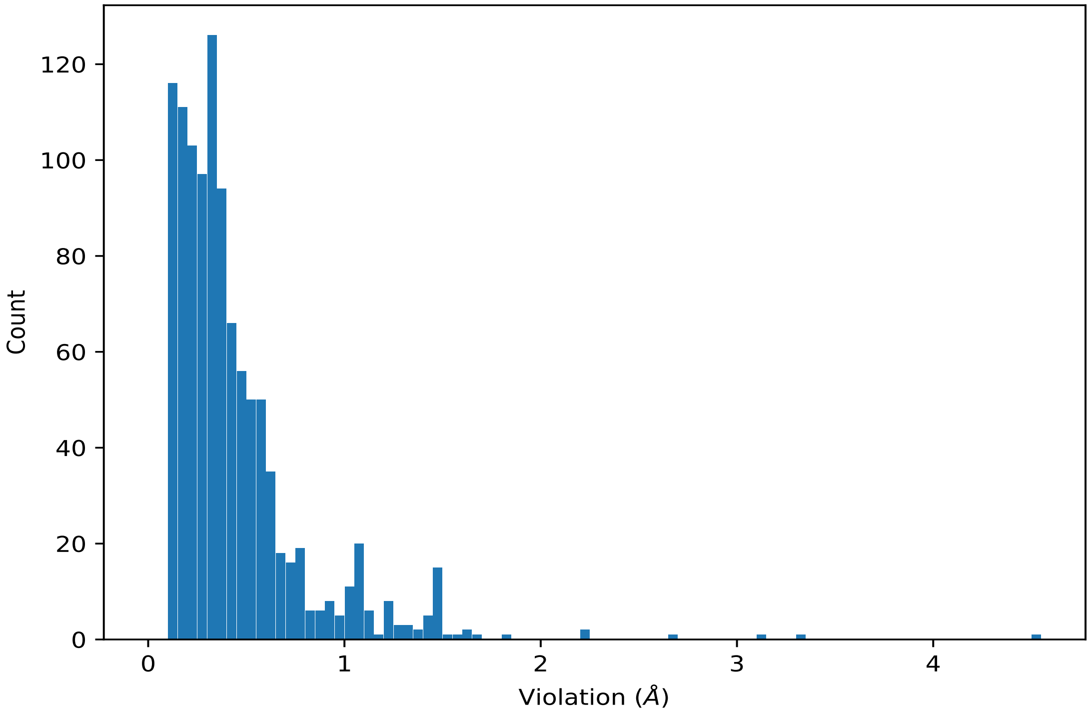
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,894)	1:A:95:TRP:HZ2	1:A:101:PRO:HD3	10	1.09	0.41	1.31
(1,284)	1:A:44:VAL:H	1:A:45:LEU:HB3	10	0.72	0.06	0.7
(1,796)	1:A:11:GLN:HB3	1:A:11:GLN:HG3	10	0.57	0.02	0.58
(1,872)	1:A:55:ARG:HB3	1:A:55:ARG:HG3	10	0.56	0.03	0.57
(1,836)	1:A:32:GLU:HB3	1:A:32:GLU:HG3	10	0.55	0.13	0.55
(1,287)	1:A:46:GLU:H	1:A:46:GLU:HB3	10	0.52	0.04	0.52
(1,802)	1:A:12:GLU:HB3	1:A:12:GLU:HG3	10	0.51	0.01	0.51
(1,832)	1:A:31:ARG:HB3	1:A:31:ARG:HG3	10	0.48	0.14	0.52
(1,860)	1:A:47:GLN:HB3	1:A:47:GLN:HG3	10	0.46	0.04	0.46
(1,841)	1:A:36:GLU:HA	1:A:36:GLU:HB3	10	0.46	0.07	0.49

[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)
(2,147)	1:A:24:ARG:HH22	1:A:33:HIS:HB3	8	4.55
(2,147)	1:A:24:ARG:HH22	1:A:33:HIS:HB3	3	3.34
(1,594)	1:A:24:ARG:HH22	1:A:29:TYR:HA	1	3.11
(1,594)	1:A:24:ARG:HH22	1:A:29:TYR:HA	5	2.68
(1,594)	1:A:24:ARG:HH22	1:A:29:TYR:HA	10	2.24
(1,594)	1:A:24:ARG:HH22	1:A:29:TYR:HA	8	2.23
(1,797)	1:A:11:GLN:HB3	1:A:95:TRP:HE3	3	1.84
(2,146)	1:A:24:ARG:HE	1:A:25:ASP:HB3	7	1.65
(1,822)	1:A:22:ASP:HA	1:A:24:ARG:HG3	2	1.64
(2,146)	1:A:24:ARG:HE	1:A:25:ASP:HB3	1	1.62

10 Dihedral-angle violation analysis (i)

No dihedral-angle restraints found