WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB NMR Structure Validation Report (i)

Jun 4, 2023 – 03:11 AM EDT

PDB ID	:	2L0D
BMRB ID	:	17039
Title	:	Solution NMR Structure of putative cell surface protein MA_4588 (272-376 domain) from Methanosarcina acetivorans, Northeast Structural Genomics Consortium Target MvR254A
Authors	:	Cort, J.R.; Lee, D.; Ciccosanti, C.; Janjua, H.; Acton, T.B.; Xiao, R.; Everett, J.K.; Montelione, G.T.; Ramelot, T.A.; Kennedy, M.A.; Northeast Structural Genomics Consortium (NESG)
Deposited on	:	2010-06-30

This is a Full wwPDB NMR Structure Validation 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 97%.

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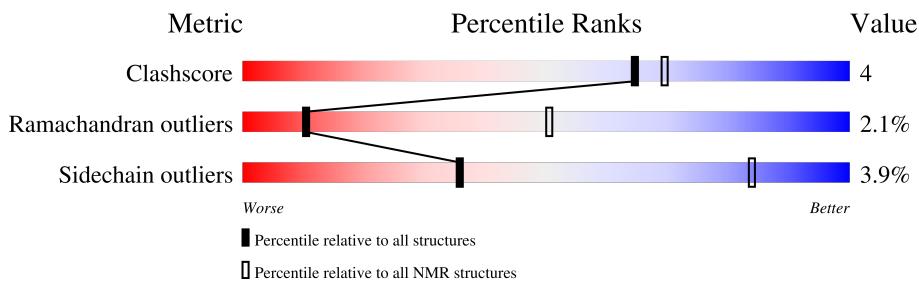
Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

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	114	87% 6% 7%

2 Ensemble composition and analysis ⓘ

This entry contains 20 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: *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:2-A:107 (106)	0.64	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 2 clusters. No single-model clusters were found.

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

3 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 1688 atoms, of which 824 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Cell surface protein.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	114	1688	529	824	146	187	2	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP Q8THC9
A	107	LEU	-	expression tag	UNP Q8THC9
A	108	GLU	-	expression tag	UNP Q8THC9
A	109	HIS	-	expression tag	UNP Q8THC9
A	110	HIS	-	expression tag	UNP Q8THC9
A	111	HIS	-	expression tag	UNP Q8THC9
A	112	HIS	-	expression tag	UNP Q8THC9
A	113	HIS	-	expression tag	UNP Q8THC9
A	114	HIS	-	expression tag	UNP Q8THC9

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.
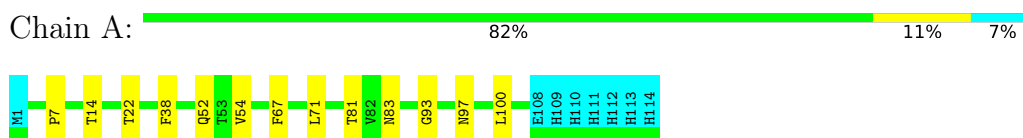
• Molecule 1: Cell surface protein

Chain A:

87%   6%   7%

M1 T14 L71 T81 D85 G93 L100 G105 E108 H109 H110 H111 H112 H113 H114
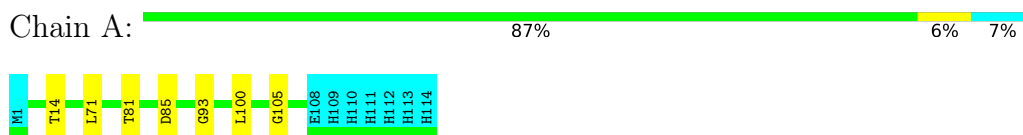
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Cell surface protein

Chain A:

82%   11%   7%

M1 P7 T14 T22 F38 Q52 T53 V54 F67 L71 T81 V82 N83 G93 N97 L100 E108 H109 H110 H111 H112 H113 H114

4.2.2 Score per residue for model 2

• Molecule 1: Cell surface protein

Chain A:

75%   18%   •   7%

M1 I2 L5 V6 P7 P12 V13 T14 V15 V16 P17 N18 M23 I27 Q52 L57 L71 T81 G93 N94 E95 S96 N97 N98 T99 L100 G105 E108 H109 H110 H111 H112 H113 H114

4.2.3 Score per residue for model 3

• Molecule 1: Cell surface protein

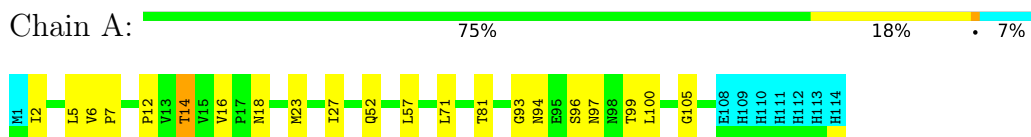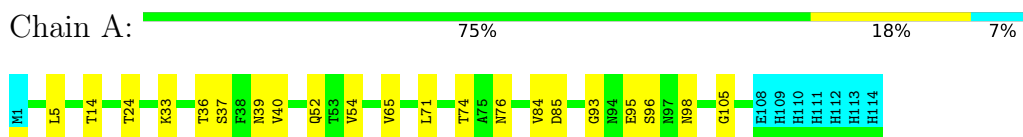### 4.2.4 Score per residue for model 4

• Molecule 1: Cell surface protein

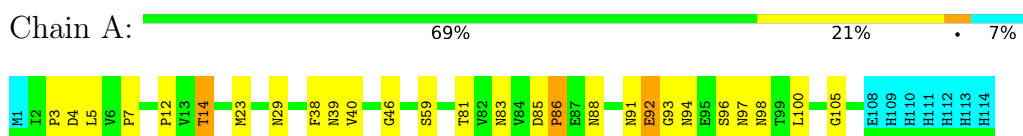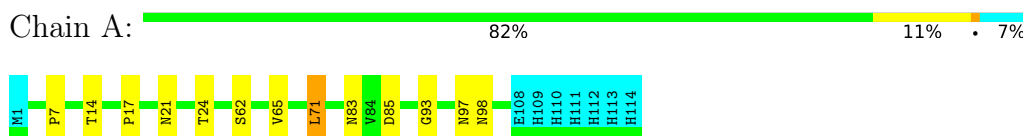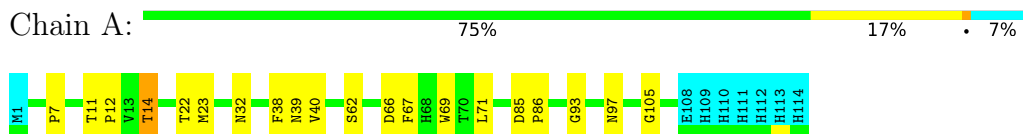### 4.2.5 Score per residue for model 5

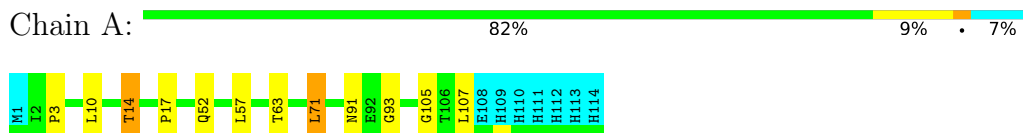• Molecule 1: Cell surface protein

4.2.6 Score per residue for model 6
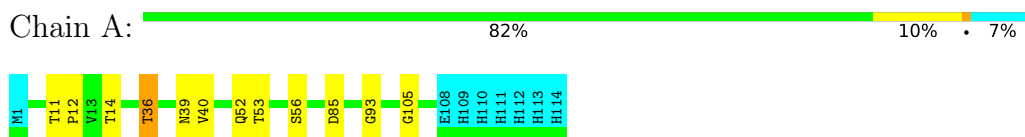
• Molecule 1: Cell surface protein

4.2.7 Score per residue for model 7 (medoid)

• Molecule 1: Cell surface protein

4.2.8 Score per residue for model 8

• Molecule 1: Cell surface protein

Chain A:

82% · 10% · 7%

M1 T11 P12 V13 T14 T36 N39 V40 Q52 T53 S56 D85 G93 G105 E108 H109 H110 H111 H112 H113 H114

4.2.9 Score per residue for model 9

• Molecule 1: Cell surface protein

Chain A:

78% 14% · 7%

M1 I2 P3 P7 T14 T24 F38 N39 V40 N61 V65 L71 Y78 T81 D85 N91 E92 G93 S96 T99 L100 E108 H109 H110 H111 H112 H113 H114

4.2.10 Score per residue for model 10

• Molecule 1: Cell surface protein

Chain A:

75% 18% 7%

M1 P7 T14 T22 M23 K33 T36 S37 F38 N39 V40 Q52 T55 S62 D66 F67 L71 T81 D85 G93 S96 N97 L100 E108 H109 H110 H111 H112 H114

4.2.11 Score per residue for model 11

• Molecule 1: Cell surface protein

Chain A:

77% 14% · 7%

M1 P7 V16 P17 N21 T22 M23 V44 D45 G46 I47 D66 L71 T74 D85 G93 N94 E95 S96 N97 N98 T99 L107 E108 H109 H110 H111 H112 H113 H114

4.2.12 Score per residue for model 12

• Molecule 1: Cell surface protein

Chain A:

79% 13% · 7%

M1 T14 V20 N29 K33 D34 E60 N61 S62 T70 T74 T81 D85 E92 G93 T99 L100 G105 E108 H109 H110 H111 H112 H113 H114

4.2.13 Score per residue for model 13

• Molecule 1: Cell surface protein

Chain A:

82% 11% 7%

M1 T14 P17 V20 Q52 L57 T63 T70 L71 T81 G93 N94 E95 L100 E108 H109 H110 H111 H112 H113 H114

4.2.14 Score per residue for model 14

• Molecule 1: Cell surface protein

Chain A:

78% 14% 7%

M1 L5 T11 P12 P17 N18 N21 T36 S37 V54 L71 D72 G73 N76 S77 Y78 E92 G93 N98 T106 L107 E108 H109 H110 H111 H112 H113 H114

4.2.15 Score per residue for model 15

• Molecule 1: Cell surface protein

Chain A:

77% 14% 7%

M1 I2 P3 P7 T14 V15 V16 P17 N29 S37 V54 L57 W69 T70 L71 T81 D85 E92 G93 L100 G105 T106 L107 E108 H109 H110 H111 H112 H113 H114

4.2.16 Score per residue for model 16

• Molecule 1: Cell surface protein

Chain A:

79% 13% 7%

M1 T14 V15 V16 P17 N21 N32 L57 T63 L71 N76 T81 D85 G93 L100 G105 T106 L107 E108 H109 H110 H111 H112 H113 H114

4.2.17 Score per residue for model 17

• Molecule 1: Cell surface protein

Chain A:

79% 14% 7%

M1 T14 V15 V16 P17 N18 T24 N32 V40 E60 V65 L71 T81 D85 G93 T99 L100 G105 T106 L107 E108 H109 H110 H111 H112 H113 H114
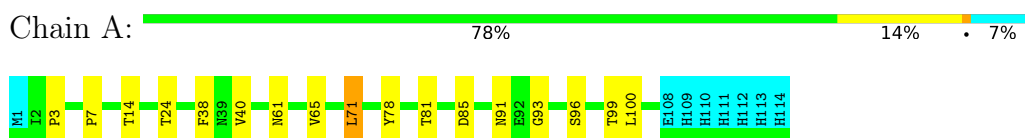
4.2.18 Score per residue for model 18

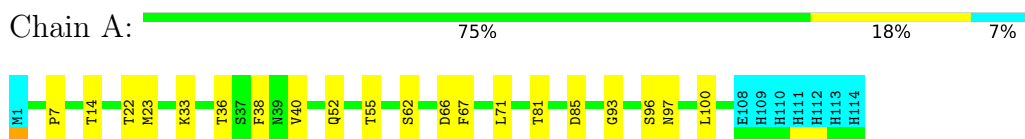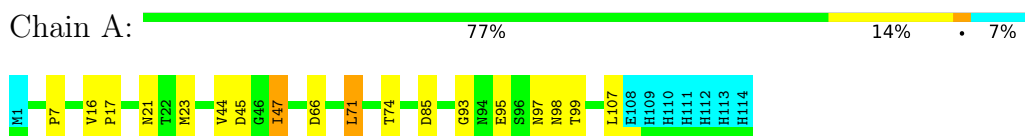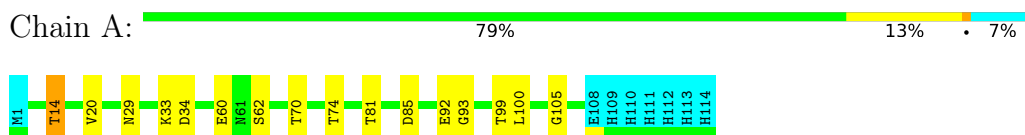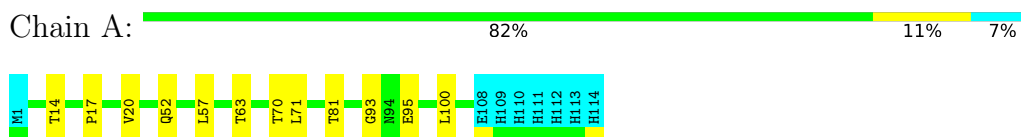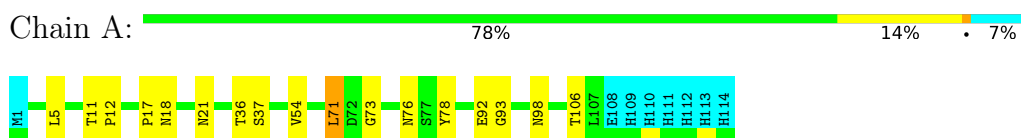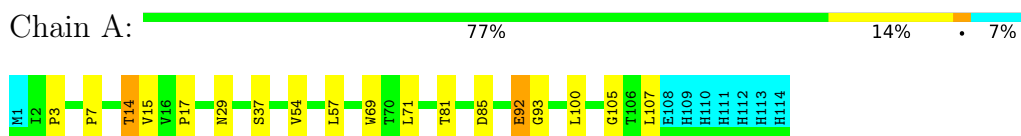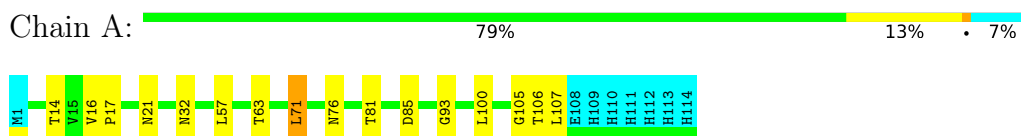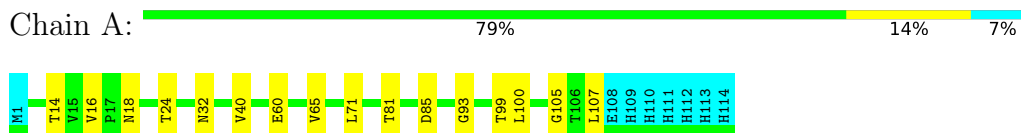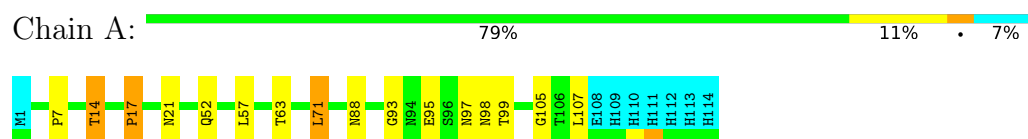• Molecule 1: Cell surface protein

Chain A:

4.2.19 Score per residue for model 19

• Molecule 1: Cell surface protein

Chain A:
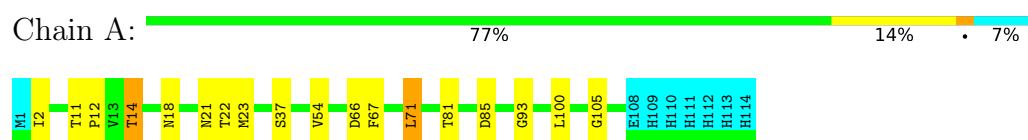
4.2.20 Score per residue for model 20

• Molecule 1: Cell surface protein

Chain A:
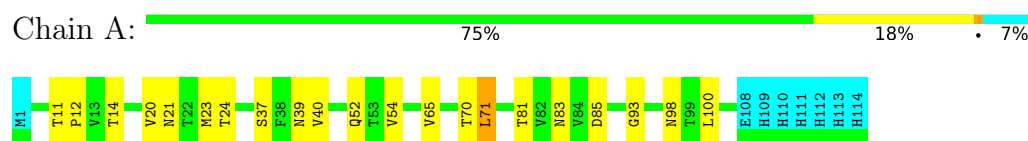
5 Refinement protocol and experimental data overview ⓘ

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

Of the 40 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy*.

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

Software name	Classification	Version
CNS	refinement	
CNS	structure solution	
CNS	geometry optimization	
CYANA	refinement	3.0
CYANA	geometry optimization	3.0
CYANA	structure solution	3.0
TALOS+	geometry optimization	
PALES	geometry optimization	
REDCAT	geometry optimization	
AutoStructure	refinement	2.1
Xplor-NIH	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	1313
Number of shifts mapped to atoms	1313
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	97%

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	786	759	758	6±2
All	All	15720	15180	15160	130

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 4.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Models Total
1:A:20:VAL:HG22	1:A:70:THR:HB	0.81	1.51	13	2
1:A:85:ASP:OD2	1:A:88:ASN:HA	0.65	1.92	4	1
1:A:5:LEU:HG	1:A:92:GLU:OE1	0.63	1.94	14	1
1:A:7:PRO:HD3	1:A:97:ASN:O	0.58	1.98	5	1
1:A:21:ASN:OD1	1:A:71:LEU:HD23	0.56	2.01	16	3
1:A:12:PRO:HD2	1:A:23:MET:SD	0.56	2.41	2	4
1:A:14:THR:HG22	1:A:107:LEU:HA	0.56	1.77	15	2
1:A:21:ASN:ND2	1:A:71:LEU:HD23	0.55	2.16	20	3
1:A:92:GLU:HG3	1:A:98:ASN:HD21	0.54	1.62	14	1
1:A:20:VAL:HA	1:A:70:THR:HA	0.54	1.78	12	1
1:A:23:MET:O	1:A:66:ASP:HA	0.53	2.03	10	3
1:A:14:THR:HA	1:A:105:GLY:O	0.53	2.03	3	12
1:A:3:PRO:O	1:A:92:GLU:HA	0.53	2.04	15	1
1:A:57:LEU:HD13	1:A:63:THR:HG23	0.52	1.80	13	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:22:THR:HA	1:A:67:PHE:O	0.52	2.05	19	4
1:A:24:THR:HA	1:A:65:VAL:O	0.52	2.04	20	5
1:A:95:GLU:HA	1:A:98:ASN:OD1	0.52	2.05	11	2
1:A:57:LEU:HD13	1:A:63:THR:OG1	0.52	2.04	7	3
1:A:94:ASN:HB2	1:A:97:ASN:HB3	0.51	1.82	2	1
1:A:69:TRP:CZ2	1:A:71:LEU:HD13	0.51	2.41	6	1
1:A:16:VAL:HG12	1:A:18:ASN:H	0.50	1.67	2	1
1:A:29:ASN:ND2	1:A:59:SER:HA	0.50	2.22	4	1
1:A:37:SER:HA	1:A:54:VAL:O	0.50	2.07	20	5
1:A:3:PRO:HD2	1:A:91:ASN:O	0.49	2.08	7	3
1:A:36:THR:O	1:A:56:SER:HA	0.48	2.07	8	1
1:A:17:PRO:HA	1:A:71:LEU:CB	0.48	2.38	15	4
1:A:38:PHE:CE2	1:A:40:VAL:HG23	0.48	2.44	6	4
1:A:39:ASN:O	1:A:86:PRO:HD3	0.48	2.08	4	2
1:A:16:VAL:HG22	1:A:107:LEU:HD22	0.48	1.85	11	3
1:A:11:THR:OG1	1:A:12:PRO:HA	0.48	2.09	14	4
1:A:4:ASP:HA	1:A:92:GLU:OE1	0.48	2.09	4	1
1:A:81:THR:HA	1:A:100:LEU:O	0.48	2.09	12	12
1:A:22:THR:CG2	1:A:66:ASP:HB3	0.47	2.40	19	1
1:A:71:LEU:HD11	1:A:78:TYR:CZ	0.47	2.45	9	1
1:A:7:PRO:HD2	1:A:97:ASN:O	0.47	2.09	10	5
1:A:73:GLY:HA3	1:A:78:TYR:OH	0.46	2.10	14	1
1:A:5:LEU:HD22	1:A:57:LEU:HD23	0.46	1.87	2	1
1:A:76:ASN:O	1:A:105:GLY:HA2	0.45	2.10	16	1
1:A:83:ASN:HA	1:A:98:ASN:O	0.45	2.12	20	3
1:A:5:LEU:O	1:A:97:ASN:HB3	0.45	2.11	4	1
1:A:44:VAL:O	1:A:47:ILE:HG23	0.45	2.11	11	1
1:A:20:VAL:HG22	1:A:70:THR:HG23	0.45	1.89	20	1
1:A:7:PRO:HB2	1:A:100:LEU:HB2	0.44	1.88	4	3
1:A:39:ASN:HA	1:A:53:THR:HA	0.44	1.88	8	1
1:A:17:PRO:HA	1:A:71:LEU:HG	0.44	1.89	14	2
1:A:6:VAL:O	1:A:27:ILE:HA	0.44	2.13	2	1
1:A:5:LEU:O	1:A:98:ASN:HA	0.43	2.13	3	1
1:A:15:VAL:HG21	1:A:69:TRP:CZ3	0.42	2.49	15	1
1:A:17:PRO:HA	1:A:71:LEU:HB3	0.42	1.90	11	2
1:A:40:VAL:HG22	1:A:84:VAL:HA	0.42	1.90	3	1
1:A:39:ASN:O	1:A:85:ASP:HA	0.41	2.15	20	2
1:A:38:PHE:CE1	1:A:54:VAL:HB	0.41	2.51	1	1
1:A:76:ASN:HA	1:A:106:THR:OG1	0.41	2.15	16	1
1:A:33:LYS:HG2	1:A:34:ASP:N	0.41	2.30	12	1
1:A:29:ASN:HB2	1:A:57:LEU:HG	0.41	1.93	15	1

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:10:LEU:C	1:A:10:LEU:HD12	0.41	2.36	7	1
1:A:29:ASN:O	1:A:60:GLU:HA	0.40	2.17	12	1
1:A:76:ASN:HA	1:A:106:THR:HG23	0.40	1.94	14	1

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	106/114 (93%)	96±3 (90±3%)	8±2 (8±2%)	2±1 (2±1%)	10	50
All	All	2120/2280 (93%)	1911 (90%)	165 (8%)	44 (2%)	10	50

All 10 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	93	GLY	20
1	A	85	ASP	11
1	A	96	SER	4
1	A	18	ASN	3
1	A	7	PRO	1
1	A	95	GLU	1
1	A	46	GLY	1
1	A	86	PRO	1
1	A	36	THR	1
1	A	17	PRO	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	97/105 (92%)	93±1 (96±2%)	4±1 (4±2%)	36	84
All	All	1940/2100 (92%)	1864 (96%)	76 (4%)	36	84

All 24 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	14	THR	14
1	A	71	LEU	14
1	A	52	GLN	9
1	A	99	THR	6
1	A	62	SER	4
1	A	36	THR	3
1	A	74	THR	3
1	A	92	GLU	3
1	A	40	VAL	3
1	A	2	ILE	2
1	A	32	ASN	2
1	A	83	ASN	1
1	A	96	SER	1
1	A	76	ASN	1
1	A	94	ASN	1
1	A	61	ASN	1
1	A	55	THR	1
1	A	21	ASN	1
1	A	45	ASP	1
1	A	47	ILE	1
1	A	95	GLU	1
1	A	60	GLU	1
1	A	88	ASN	1
1	A	11	THR	1

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 ⓘ

There are no ligands in this entry.

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 97% for the well-defined parts and 91% 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	1313
Number of shifts mapped to atoms	1313
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	2

7.1.2 Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_\alpha$	109	0.12 ± 0.16	None needed (< 0.5 ppm)
$^{13}C_\beta$	104	0.07 ± 0.12	None needed (< 0.5 ppm)
$^{13}C'$	101	0.45 ± 0.14	None needed (< 0.5 ppm)
^{15}N	102	-0.06 ± 0.54	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 97%, i.e. 1281 atoms were assigned a chemical shift out of a possible 1325. 0 out of 26 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	516/525 (98%)	211/212 (100%)	205/212 (97%)	100/101 (99%)
Sidechain	737/751 (98%)	492/492 (100%)	229/242 (95%)	16/17 (94%)

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	28/49 (57%)	14/24 (58%)	13/22 (59%)	1/3 (33%)
Overall	1281/1325 (97%)	717/728 (98%)	447/476 (94%)	117/121 (97%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 91%, i.e. 1312 atoms were assigned a chemical shift out of a possible 1448. 0 out of 26 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	528/565 (93%)	216/228 (95%)	210/228 (92%)	102/109 (94%)
Sidechain	756/786 (96%)	505/515 (98%)	235/254 (93%)	16/17 (94%)
Aromatic	28/97 (29%)	14/48 (29%)	13/34 (38%)	1/15 (7%)
Overall	1312/1448 (91%)	735/791 (93%)	458/516 (89%)	119/141 (84%)

### 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	42	LEU	HB2	-0.85	$-0.07 - 3.30$	-7.3
1	A	54	VAL	HB	0.16	$0.43 - 3.54$	-5.9

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:
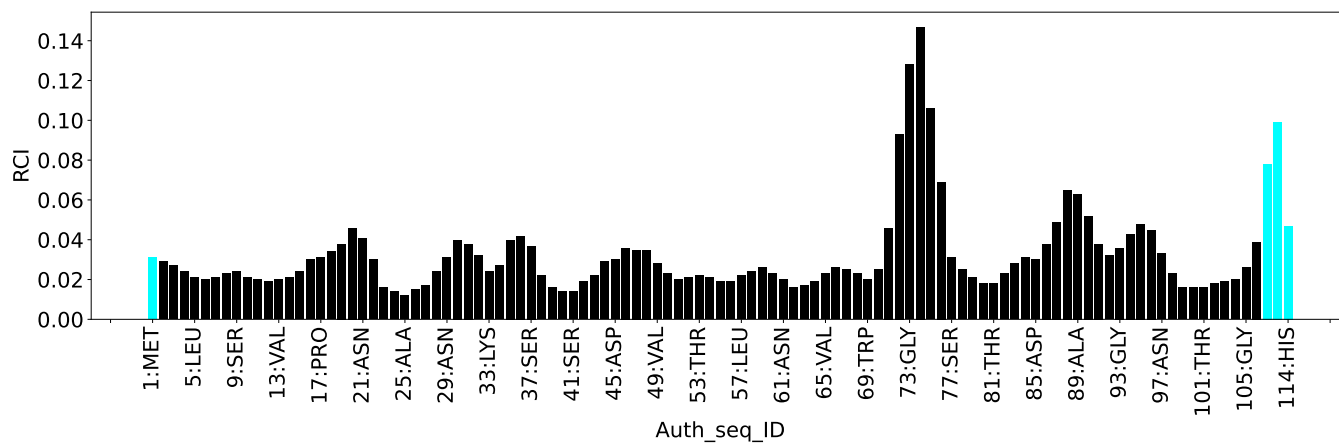
8 NMR restraints analysis ⓘ

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	3288
Intra-residue ($\|i-j\|=0$)	420
Sequential ($\|i-j\|=1$)	750
Medium range ($\|i-j\|>1$ and $\|i-j\|<5$)	342
Long range ($\|i-j\|\geq 5$)	1672
Inter-chain	0
Hydrogen bond restraints	104
Disulfide bond restraints	0
Total dihedral-angle restraints	111
Number of unmapped restraints	0
Number of restraints per residue	29.8
Number of long range restraints per residue[1]	15.5

[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)	1.3	0.15
0.2-0.5 (Medium)	None	None
>0.5 (Large)	None	None

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

8.2.2 Average number of dihedral-angle violations per model ⓘ

Dihedral-angle violations less than 1° are not included in the calculation.There are no dihedral-angle violations
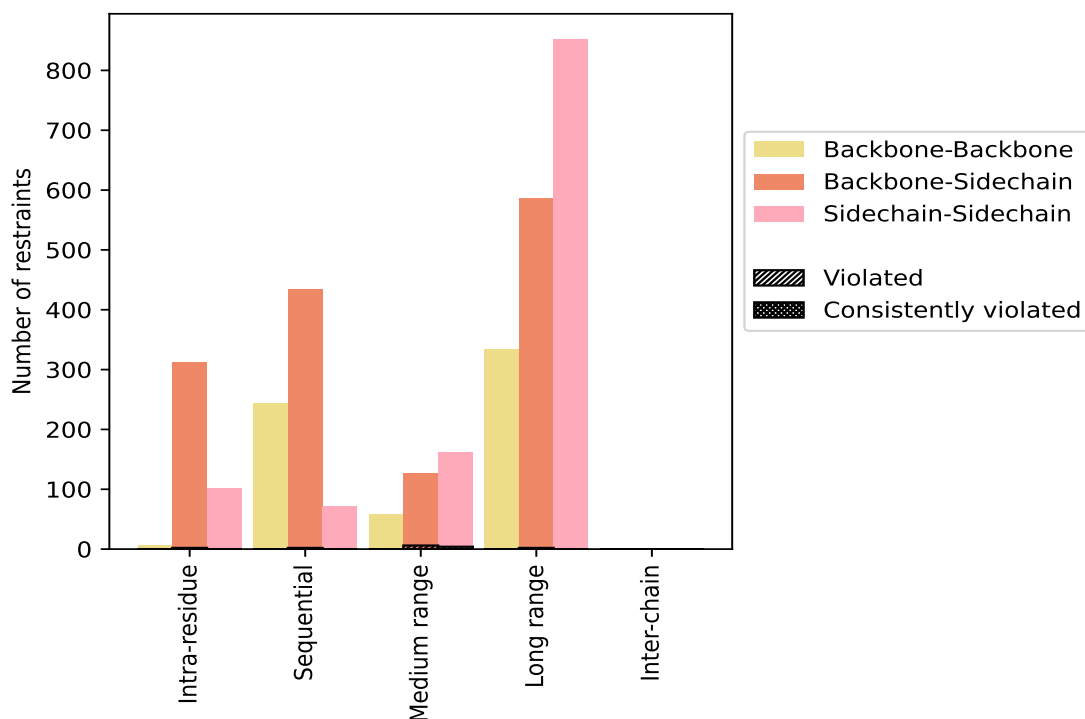
9　Distance violation analysis ⓘ

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)	420	12.8	2	0.5	0.1	0	0.0	0.0
Backbone-Backbone	6	0.2	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	312	9.5	2	0.6	0.1	0	0.0	0.0
Sidechain-Sidechain	102	3.1	0	0.0	0.0	0	0.0	0.0
Sequential (|i-j|=1)	750	22.8	2	0.3	0.1	0	0.0	0.0
Backbone-Backbone	244	7.4	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	434	13.2	2	0.5	0.1	0	0.0	0.0
Sidechain-Sidechain	72	2.2	0	0.0	0.0	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	342	10.4	10	2.9	0.3	0	0.0	0.0
Backbone-Backbone	54	1.6	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	126	3.8	6	4.8	0.2	0	0.0	0.0
Sidechain-Sidechain	162	4.9	4	2.5	0.1	0	0.0	0.0
Long range (|i-j|≥5)	1672	50.9	2	0.1	0.1	0	0.0	0.0
Backbone-Backbone	234	7.1	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	586	17.8	2	0.3	0.1	0	0.0	0.0
Sidechain-Sidechain	852	25.9	0	0.0	0.0	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	104	3.2	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	3288	100.0	16	0.5	0.5	0	0.0	0.0
Backbone-Backbone	642	19.5	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	1458	44.3	12	0.8	0.4	0	0.0	0.0
Sidechain-Sidechain	1188	36.1	4	0.3	0.1	0	0.0	0.0

[1] percentage calculated with respect to the total number of distance restraints, [2] percentage calculated with respect to the number of restraints in a particular restraint category, [3] violated in at least one model, [4] violated in all the models

9.1.1   Bar chart : Distribution of distance restraints and violations ⓘ

Violated and consistently violated restraints are shown using different hatch patterns in their respective categories. The hydrogen bonds and disulfied bonds are counted in their appropriate category on the x-axis

9.2   Distance violation statistics for each model ⓘ

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

Model ID	Number of violations						Mean (Å)	Max (Å)	SD^6 (Å)	Median (Å)
	IR^1	SQ^2	MR^3	LR^4	IC^5	Total				
1	0	2	0	0	0	2	0.11	0.11	0.0	0.11
2	0	0	0	0	0	0	0.0	0.0	0.0	0.0
3	0	0	0	0	0	0	0.0	0.0	0.0	0.0
4	0	0	2	2	0	4	0.12	0.12	0.0	0.12
5	0	0	0	0	0	0	0.0	0.0	0.0	0.0
6	0	0	2	0	0	2	0.11	0.11	0.0	0.11
7	0	0	0	0	0	0	0.0	0.0	0.0	0.0
8	2	0	0	0	0	2	0.14	0.14	0.0	0.14
9	0	0	0	0	0	0	0.0	0.0	0.0	0.0
10	0	0	0	0	0	0	0.0	0.0	0.0	0.0
11	2	0	2	0	0	4	0.12	0.12	0.0	0.12

Continued on next page...

Continued from previous page...

| Model ID | Number of violations | | | | | | Mean (Å) | Max (Å) | SD[6] (Å) | Median (Å) |
	IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total				
12	0	0	0	0	0	0	0.0	0.0	0.0	0.0
13	0	0	2	0	0	2	0.12	0.12	0.0	0.12
14	2	0	0	0	0	2	0.11	0.11	0.0	0.11
15	0	0	0	0	0	0	0.0	0.0	0.0	0.0
16	0	0	0	0	0	0	0.0	0.0	0.0	0.0
17	2	0	0	0	0	2	0.15	0.15	0.0	0.15
18	0	0	2	0	0	2	0.14	0.14	0.0	0.14
19	0	0	0	0	0	0	0.0	0.0	0.0	0.0
20	0	0	4	0	0	4	0.12	0.12	0.0	0.12

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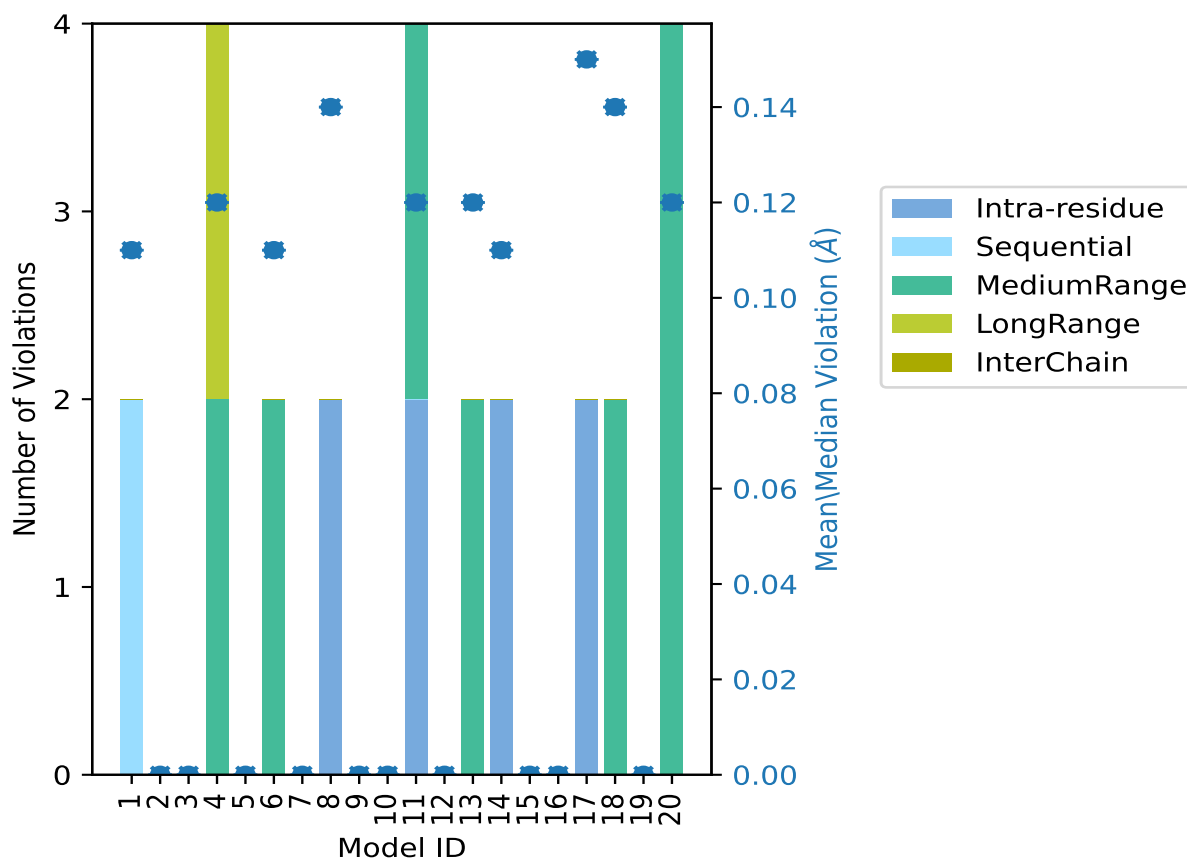
The mean(dot),median(x) and the standard deviation are shown in blue with respect to the y axis on the right

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, 3168(IR:418, SQ:748, MR:332, LR:1670, 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]	%
0	2	8	2	0	12	1	5.0
0	0	0	0	0	0	2	10.0
0	0	2	0	0	2	3	15.0
2	0	0	0	0	2	4	20.0
0	0	0	0	0	0	5	25.0
0	0	0	0	0	0	6	30.0
0	0	0	0	0	0	7	35.0
0	0	0	0	0	0	8	40.0
0	0	0	0	0	0	9	45.0
0	0	0	0	0	0	10	50.0
0	0	0	0	0	0	11	55.0
0	0	0	0	0	0	12	60.0
0	0	0	0	0	0	13	65.0
0	0	0	0	0	0	14	70.0
0	0	0	0	0	0	15	75.0
0	0	0	0	0	0	16	80.0
0	0	0	0	0	0	17	85.0
0	0	0	0	0	0	18	90.0
0	0	0	0	0	0	19	95.0
0	0	0	0	0	0	20	100.0

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

9.3.1 Bar graph : Distance violation statistics for the ensemble ⓘ
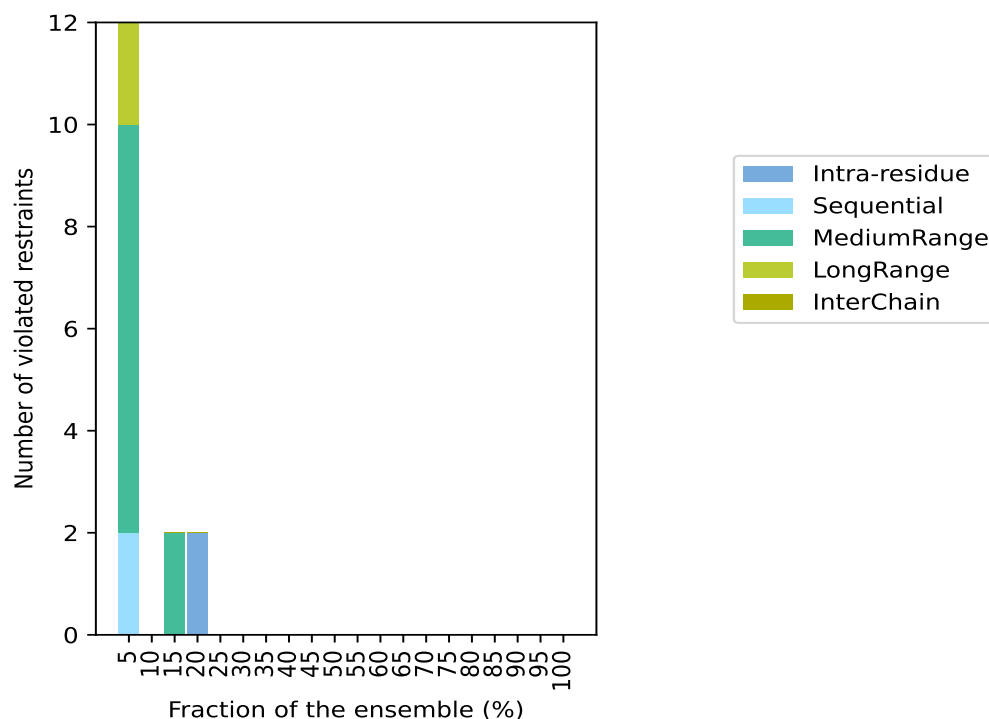
9.4 Most violated distance restraints in the ensemble ⓘ

9.4.1 Histogram : Distribution of mean distance violations ⓘ

The following histogram shows the distribution of the average value of the violation. The average is calculated for each restraint that is violated in more than one model over all the violated models in the ensemble
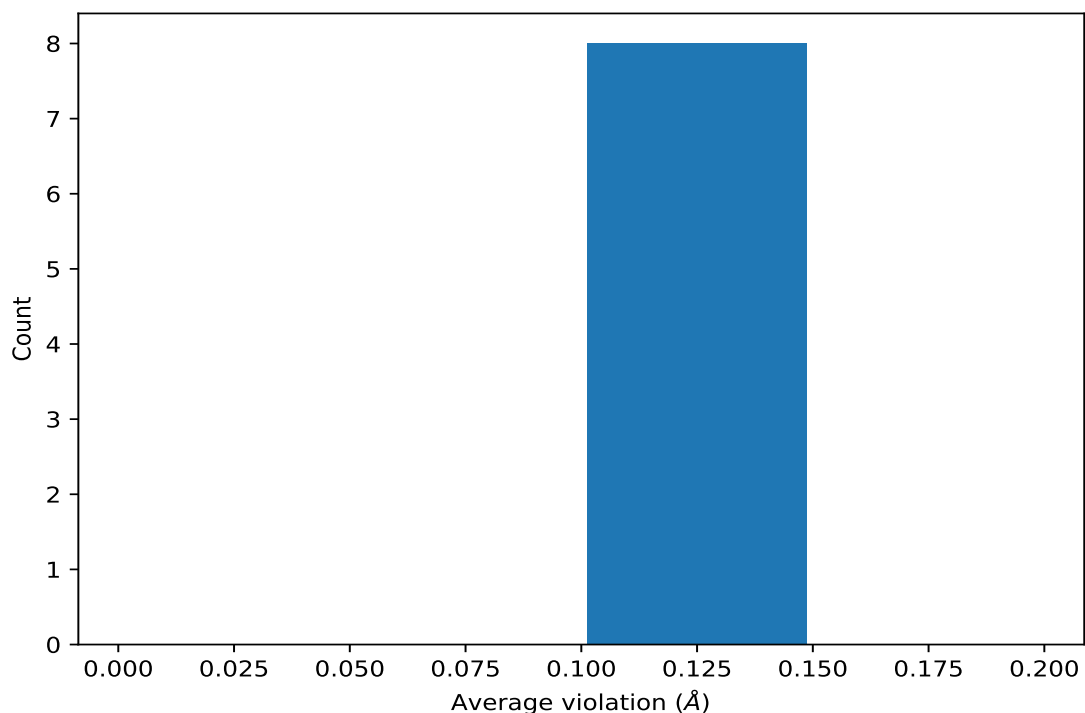
9.4.2   Table: Most violated distance restraints ⓘ

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

Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,1291)	1:A:14:THR:H	1:A:14:THR:HB	4	0.13	0.02	0.13
(1,2883)	1:A:14:THR:H	1:A:14:THR:HB	4	0.13	0.02	0.13
(1,537)	1:A:4:ASP:H	1:A:6:VAL:HG21	3	0.12	0.0	0.12
(1,537)	1:A:4:ASP:H	1:A:6:VAL:HG22	3	0.12	0.0	0.12
(1,537)	1:A:4:ASP:H	1:A:6:VAL:HG23	3	0.12	0.0	0.12
(1,2129)	1:A:4:ASP:H	1:A:6:VAL:HG21	3	0.12	0.0	0.12
(1,2129)	1:A:4:ASP:H	1:A:6:VAL:HG22	3	0.12	0.0	0.12
(1,2129)	1:A:4:ASP:H	1:A:6:VAL:HG23	3	0.12	0.0	0.12

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

9.5 All violated distance restraints (i)

9.5.1 Histogram : Distribution of distance violations (i)

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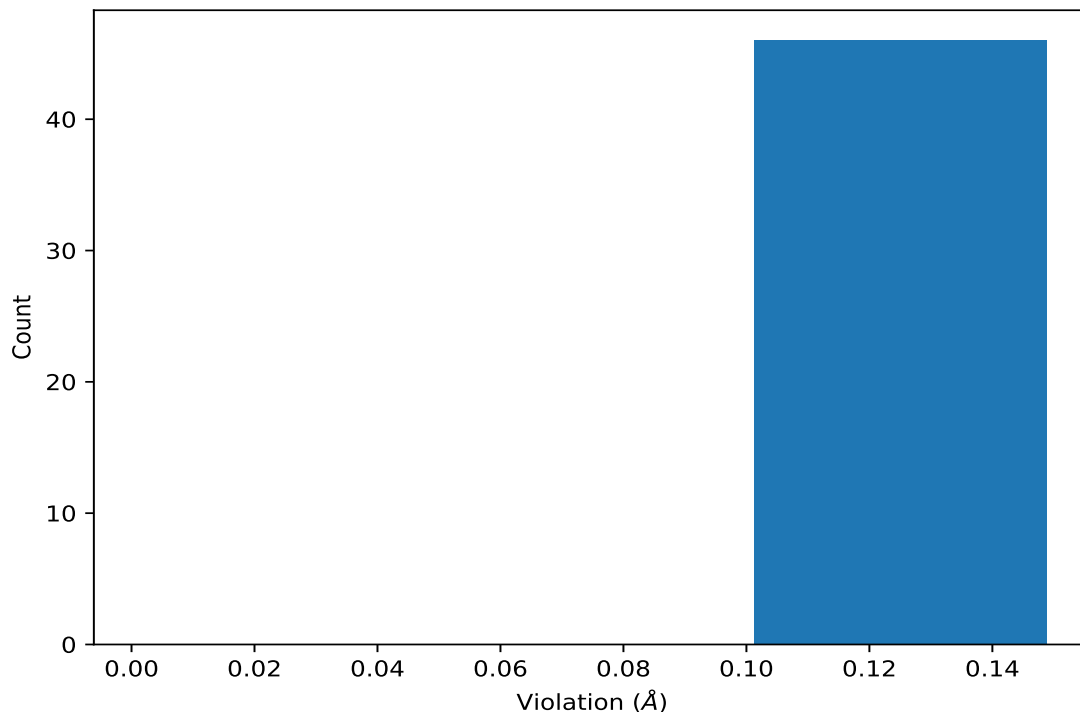
9.5.2 Table : All distance violations (i)

The following table lists the absolute value of the violation for each restraint in the ensemble sorted by its 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)
(1,2883)	1:A:14:THR:H	1:A:14:THR:HB	17	0.15
(1,1291)	1:A:14:THR:H	1:A:14:THR:HB	17	0.15
(1,916)	1:A:16:VAL:HB	1:A:18:ASN:HD21	18	0.14
(1,916)	1:A:16:VAL:HB	1:A:18:ASN:HD22	18	0.14
(1,2883)	1:A:14:THR:H	1:A:14:THR:HB	8	0.14
(1,2508)	1:A:16:VAL:HB	1:A:18:ASN:HD21	18	0.14
(1,2508)	1:A:16:VAL:HB	1:A:18:ASN:HD22	18	0.14
(1,1291)	1:A:14:THR:H	1:A:14:THR:HB	8	0.14
(1,611)	1:A:44:VAL:H	1:A:47:ILE:HB	11	0.12
(1,547)	1:A:8:VAL:H	1:A:100:LEU:HB2	4	0.12

Continued on next page...

Continued from previous page...

| Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,547)	1:A:8:VAL:H	1:A:100:LEU:HB3	4	0.12
(1,537)	1:A:4:ASP:H	1:A:6:VAL:HG21	13	0.12
(1,537)	1:A:4:ASP:H	1:A:6:VAL:HG22	13	0.12
(1,537)	1:A:4:ASP:H	1:A:6:VAL:HG23	13	0.12
(1,537)	1:A:4:ASP:H	1:A:6:VAL:HG21	20	0.12
(1,537)	1:A:4:ASP:H	1:A:6:VAL:HG22	20	0.12
(1,537)	1:A:4:ASP:H	1:A:6:VAL:HG23	20	0.12
(1,299)	1:A:11:THR:HG21	1:A:13:VAL:H	20	0.12
(1,299)	1:A:11:THR:HG22	1:A:13:VAL:H	20	0.12
(1,299)	1:A:11:THR:HG23	1:A:13:VAL:H	20	0.12
(1,2883)	1:A:14:THR:H	1:A:14:THR:HB	11	0.12
(1,2203)	1:A:44:VAL:H	1:A:47:ILE:HB	11	0.12
(1,2139)	1:A:8:VAL:H	1:A:100:LEU:HB2	4	0.12
(1,2139)	1:A:8:VAL:H	1:A:100:LEU:HB3	4	0.12
(1,2129)	1:A:4:ASP:H	1:A:6:VAL:HG21	13	0.12
(1,2129)	1:A:4:ASP:H	1:A:6:VAL:HG22	13	0.12
(1,2129)	1:A:4:ASP:H	1:A:6:VAL:HG23	13	0.12
(1,2129)	1:A:4:ASP:H	1:A:6:VAL:HG21	20	0.12
(1,2129)	1:A:4:ASP:H	1:A:6:VAL:HG22	20	0.12
(1,2129)	1:A:4:ASP:H	1:A:6:VAL:HG23	20	0.12
(1,1891)	1:A:11:THR:HG21	1:A:13:VAL:H	20	0.12
(1,1891)	1:A:11:THR:HG22	1:A:13:VAL:H	20	0.12
(1,1891)	1:A:11:THR:HG23	1:A:13:VAL:H	20	0.12
(1,1291)	1:A:14:THR:H	1:A:14:THR:HB	11	0.12
(1,537)	1:A:4:ASP:H	1:A:6:VAL:HG21	4	0.11
(1,537)	1:A:4:ASP:H	1:A:6:VAL:HG22	4	0.11
(1,537)	1:A:4:ASP:H	1:A:6:VAL:HG23	4	0.11
(1,2883)	1:A:14:THR:H	1:A:14:THR:HB	14	0.11
(1,2624)	1:A:38:PHE:HZ	1:A:40:VAL:HB	6	0.11
(1,2129)	1:A:4:ASP:H	1:A:6:VAL:HG21	4	0.11
(1,2129)	1:A:4:ASP:H	1:A:6:VAL:HG22	4	0.11
(1,2129)	1:A:4:ASP:H	1:A:6:VAL:HG23	4	0.11
(1,1751)	1:A:47:ILE:HB	1:A:48:VAL:H	1	0.11
(1,159)	1:A:47:ILE:HB	1:A:48:VAL:H	1	0.11
(1,1291)	1:A:14:THR:H	1:A:14:THR:HB	14	0.11
(1,1032)	1:A:38:PHE:HZ	1:A:40:VAL:HB	6	0.11

10　Dihedral-angle violation analysis (i)

10.1　Summary of dihedral-angle violations (i)

The following table provides the summary of dihedral-angle violations in different dihedral-angle types. Violations less than 1° are not included in the calculation.

Angle type	Count	$\%^1$	Violated[3]			Consistently Violated[4]		
			Count	$\%^2$	$\%^1$	Count	$\%^2$	$\%^1$
PHI	55	49.5	0	0.0	0.0	0	0.0	0.0
PSI	56	50.5	0	0.0	0.0	0	0.0	0.0
Total	111	100.0	0	0.0	0.0	0	0.0	0.0

[1] percentage calculated with respect to total number of dihedral-angle restraints, [2] percentage calculated with respect to number of restraints in a particular dihedral-angle type, [3] violated in at least one model, [4] violated in all the models

10.1.1　Bar chart : Distribution of dihedral-angles and violations (i)
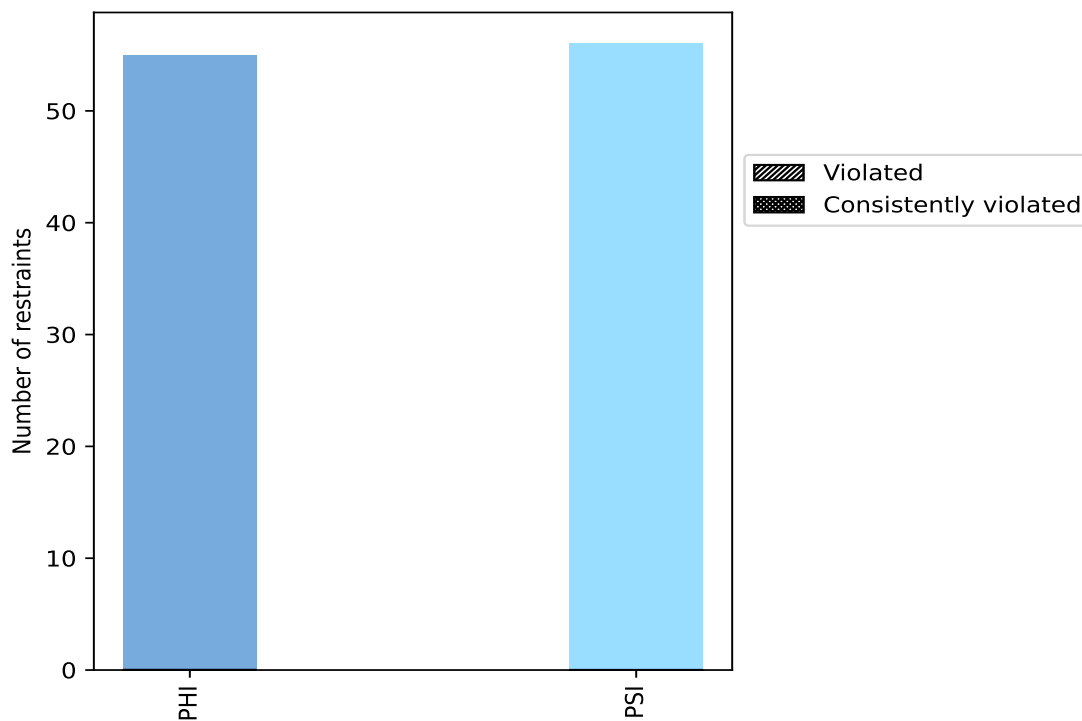
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

10.2 Dihedral-angle violation statistics for each model ⓘ

No violations found

10.3 Dihedral-angle violation statistics for the ensemble ⓘ

No violations found

10.4 Most violated dihedral-angle restraints in the ensemble ⓘ

No violations found

10.5 All violated dihedral-angle restraints ⓘ

No violations found