Full wwPDB NMR Structure Validation Report (i)

Jun 6, 2023 – 05:18 AM EDT

PDB ID	:	2MK4
BMRB ID	:	19751
Title	:	Solution structure of ORF2
Authors	:	Miyakawa, T.; Kobayashi, H.; Tashiro, M.; Yamanaka, H.; Tanokura, M.
Deposited on	:	2014-01-24

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 (1)) were used in the production of this report:

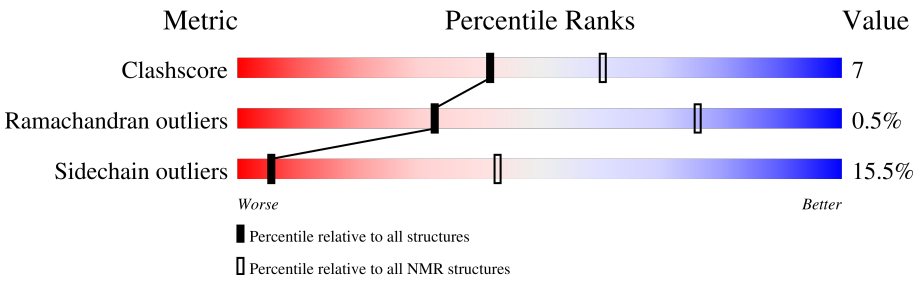
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal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

1 Overall quality at a glance (i)

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

The overall completeness of chemical shifts assignment is 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	133	46% 12% 40% .

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 1 is the overall representative, medoid model (most similar to other models).

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:6-A:26 (21)	2.36	17
2	A:63-A:88, A:93-A:122 (56)	0.24	1

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 and 4 single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1944 atoms, of which 982 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Open reading frame 2.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	130	1944	590	982	172	198	2	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: Open reading frame 2

Chain A: 46% 12% 40% •

GLY PRO VAL Q1 E2 S3 V4 T5 G27 S28 K29 K30 R31 V32 A33 R34 S35 L36 D37 S38 K39 V40 P41 Q42 Q43 T44 L45 R46 R47 G48 D49 V50 L51 M52 Q53 G54 A55 A56 S57 P58 E59 L60 T61 V62 T65 L66 L67 V68 D71 A79 L84 S89 S90 G91

G92 I93 A94 L95 L96 L104 A108 V118 Q119 I120 E121 L122 S123 G124 A125 E126 Q127 Q128 P129 K130

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 (medoid)

• Molecule 1: Open reading frame 2

Chain A: 45% 13% 40% •

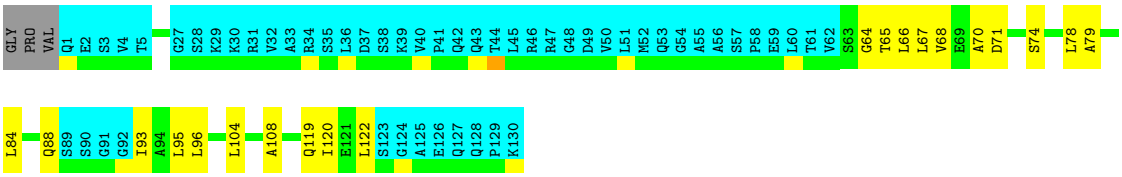
GLY PRO VAL Q1 E2 S3 V4 T5 K9 G27 S28 K29 K30 R31 V32 A33 R34 S35 L36 D37 S38 K39 V40 P41 Q42 Q43 T44 L45 R46 R47 G48 D49 V50 L51 M52 Q53 G54 A55 A56 S57 P58 E59 L60 T61 V62 T65 L66 L67 V68 D71 A79 T80 R81 L84

S89 S90 G91 G92 I93 A94 L95 L96 L104 A108 V118 Q119 I120 S123 G124 A125 E126 Q127 Q128 P129 K130

4.2.2 Score per residue for model 2

• Molecule 1: Open reading frame 2

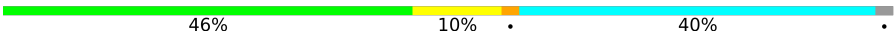
Chain A: 43% 15% 40% •

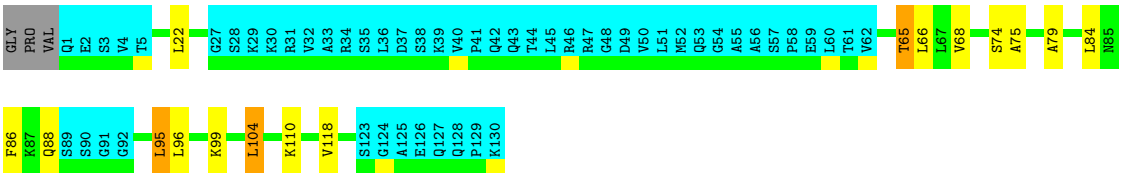


4.2.3 Score per residue for model 3

- Molecule 1: Open reading frame 2

Chain A:

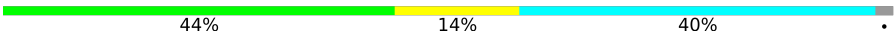


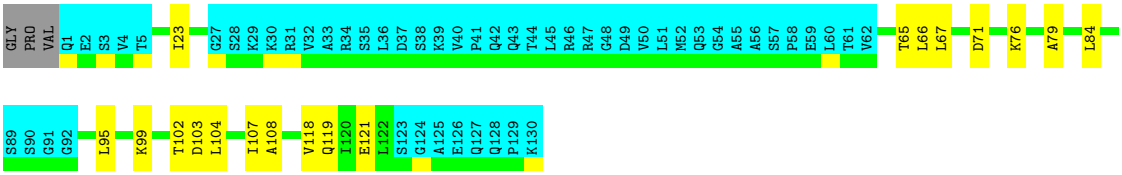


4.2.4 Score per residue for model 4

- Molecule 1: Open reading frame 2

Chain A:

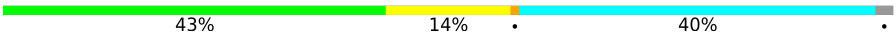


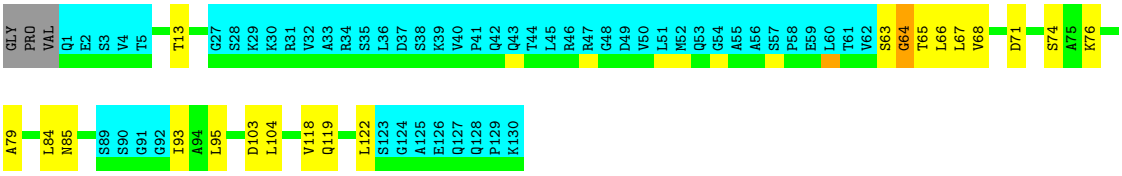


4.2.5 Score per residue for model 5

- Molecule 1: Open reading frame 2

Chain A:





4.2.6 Score per residue for model 6

- Molecule 1: Open reading frame 2

Chain A: 36% 21% • 40% •

GLY PRO VAL Q1 E2 S3 V4 T5 I23 G27 S28 K29 K30 R31 V32 A33 R34 S35 L36 D37 S38 K39 V40 P41 Q42 Q43 T44 L45 R46 R47 G48 D49 V50 L51 M52 Q53 G54 A55 A56 S57 P58 E59 L60 T61 V62 T65 L66 L67 V68 E69 A70 D71 S74 A75 K76 A77

L78 A79 L84 Q88 S89 S90 G91 G92 I93 A94 L95 T102 D103 L104 N105 A106 I107 A108 T109 K110 L111 V116 N117 V118 Q119 I120 E121 L122 S123 G124 A125 E126 Q127 Q128 P129 K130

4.2.7 Score per residue for model 7

- Molecule 1: Open reading frame 2

Chain A: 41% 16% • 40% •

GLY PRO VAL Q1 E2 S3 V4 T5 M6 D7 G8 K9 L22 I23 G27 S28 K29 K30 R31 V32 A33 R34 S35 L36 D37 S38 K39 V40 P41 Q42 Q43 T44 L45 R46 R47 G48 D49 V50 L51 M52 Q53 G54 A55 A56 S57 P58 E59 L60 T61 V62 T65 L66 L67 V68 E69 A70 D71

S74 L78 A79 L84 S89 S90 G91 G92 I93 A94 L95 L96 L104 V118 Q119 I120 E121 L122 S123 G124 A125 E126 Q127 Q128 P129 K130

4.2.8 Score per residue for model 8

- Molecule 1: Open reading frame 2

Chain A: 41% 15% • 40% •

GLY PRO VAL Q1 E2 S3 V4 T5 G8 K9 L22 I23 G27 S28 K29 K30 R31 V32 A33 R34 S35 L36 D37 S38 K39 V40 P41 Q42 Q43 T44 L45 R46 R47 G48 D49 V50 L51 M52 Q53 G54 A55 A56 S57 P58 E59 L60 T61 V62 S63 G64 T65 L66 L67 V68 D71

K76 A77 L78 A79 L84 N85 F86 K87 Q88 S89 S90 G91 G92 I93 A94 L95 L104 A108 V118 L122 S123 G124 A125 E126 Q127 Q128 P129 K130

4.2.9 Score per residue for model 9

- Molecule 1: Open reading frame 2

Chain A: 44% 14% • 40% •

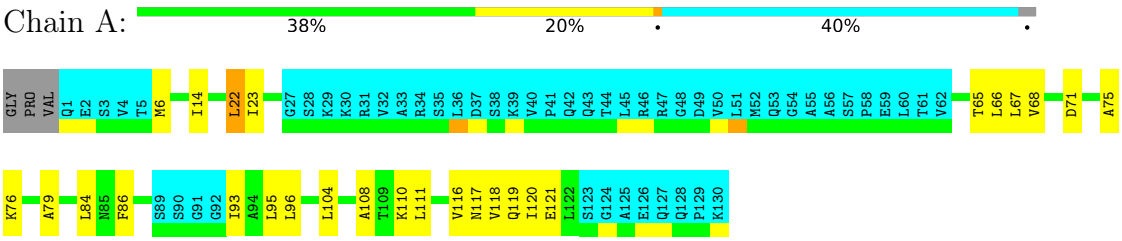
GLY PRO VAL Q1 E2 S3 V4 T5 E15 I23 G27 S28 K29 K30 R31 V32 A33 R34 S35 L36 D37 S38 K39 V40 P41 Q42 Q43 T44 L45 R46 R47 G48 D49 V50 L51 M52 Q53 G54 A55 A56 S57 P58 E59 L60 T61 V62 T65 L66 L67 V68 S74 A79 L84

Q88 S89 S90 G91 G92 I93 A94 L95 L104 N105 A108 V118 Q119 I120 E121 L122 S123 G124 A125 E126 Q127 Q128 P129 K130

4.2.10 Score per residue for model 10

- Molecule 1: Open reading frame 2

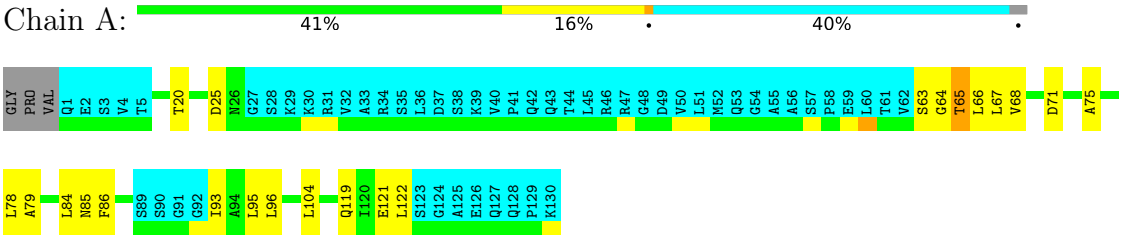
Chain A:



4.2.11 Score per residue for model 11

- Molecule 1: Open reading frame 2

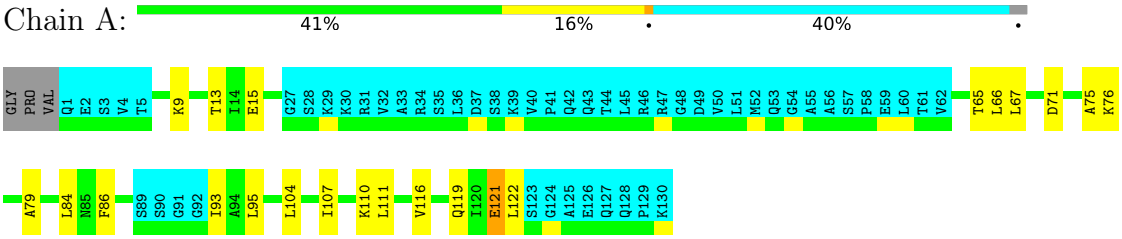
Chain A:



4.2.12 Score per residue for model 12

- Molecule 1: Open reading frame 2

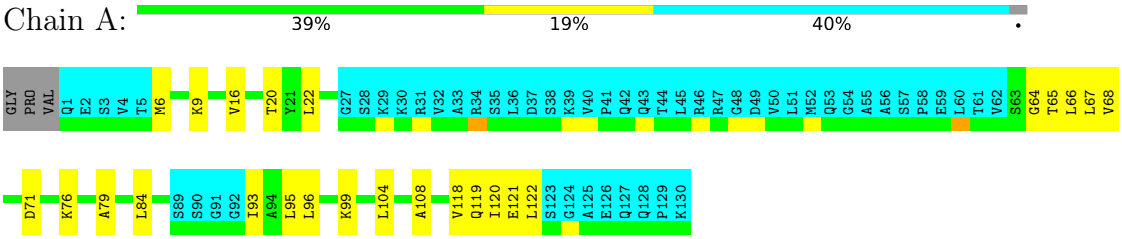
Chain A:



4.2.13 Score per residue for model 13

- Molecule 1: Open reading frame 2

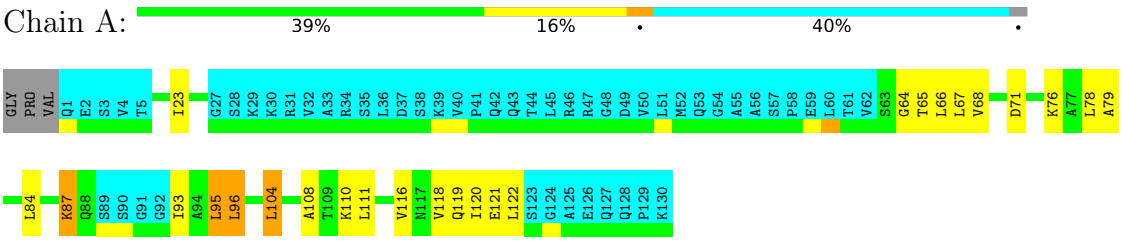
Chain A:



4.2.14 Score per residue for model 14

- Molecule 1: Open reading frame 2

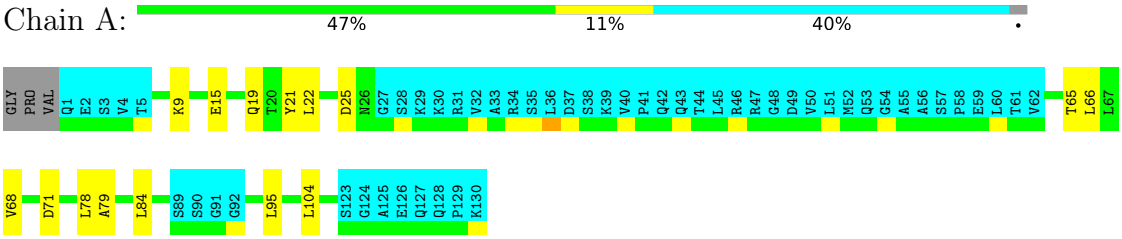
Chain A:



4.2.15 Score per residue for model 15

- Molecule 1: Open reading frame 2

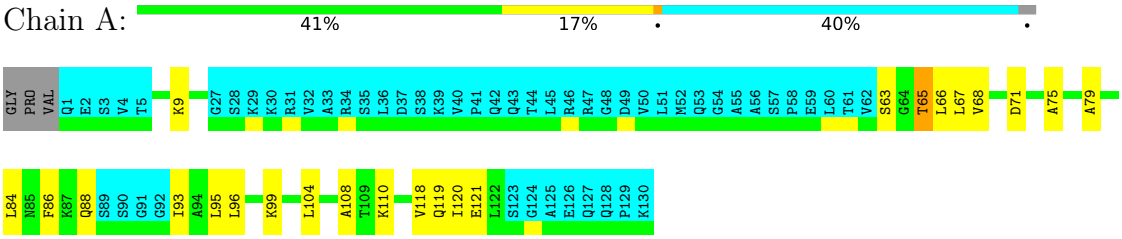
Chain A:



4.2.16 Score per residue for model 16

- Molecule 1: Open reading frame 2

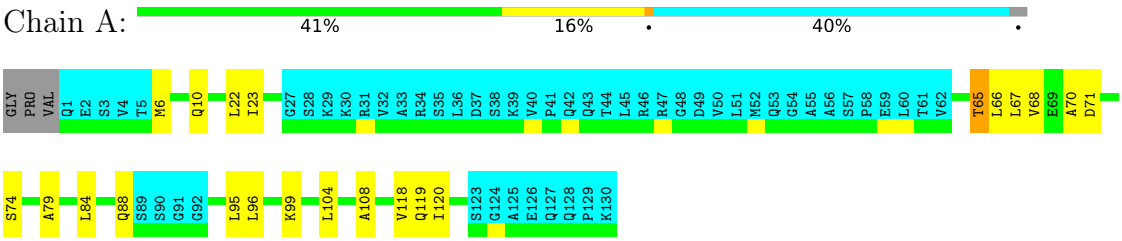
Chain A:



4.2.17 Score per residue for model 17

- Molecule 1: Open reading frame 2

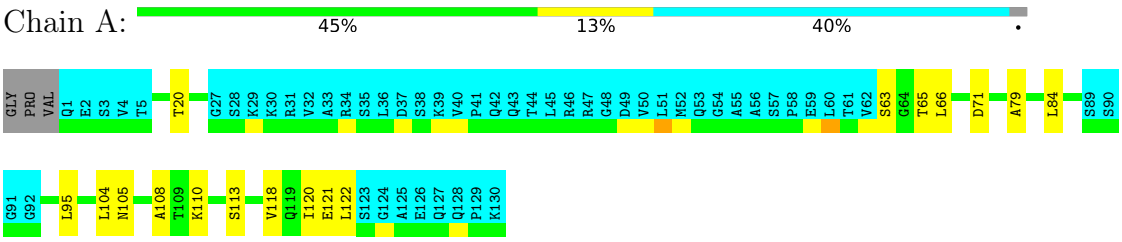
Chain A:



4.2.18 Score per residue for model 18

- Molecule 1: Open reading frame 2

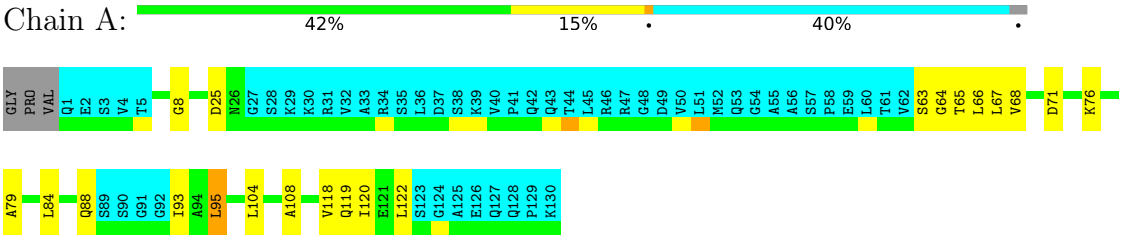
Chain A:



4.2.19 Score per residue for model 19

- Molecule 1: Open reading frame 2

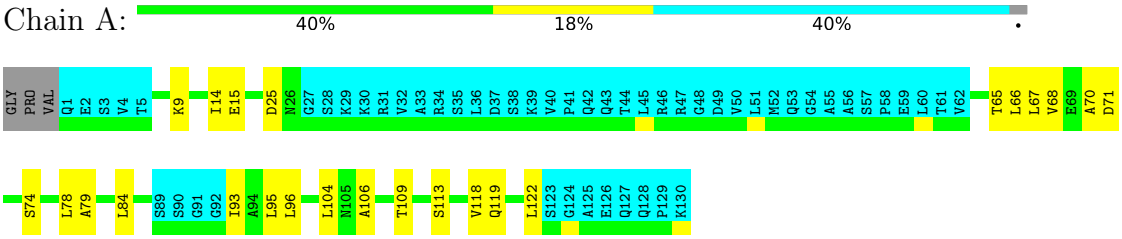
Chain A:



4.2.20 Score per residue for model 20

- Molecule 1: Open reading frame 2

Chain A:



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *DGSA-distance geometry simulated annealing.*

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

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

Software name	Classification	Version
CYANA	structure solution	
CYANA	geometry optimization	
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	1329
Number of shifts mapped to atoms	1329
Number of unparsed shifts	0
Number of shifts with mapping errors	0
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	576	586	586	8±3
All	All	11520	11720	11720	169

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 7.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:104:LEU:HD12	1:A:120:ILE:HG23	0.67	1.64	14	1
1:A:70:ALA:HB1	1:A:74:SER:HB2	0.61	1.73	17	4
1:A:108:ALA:HB2	1:A:120:ILE:HD11	0.61	1.71	1	8
1:A:16:VAL:HG22	1:A:20:THR:HB	0.60	1.72	13	1
1:A:67:LEU:HD13	1:A:93:ILE:HD12	0.57	1.77	1	6
1:A:108:ALA:CB	1:A:120:ILE:HD11	0.56	2.30	16	6
1:A:70:ALA:HB1	1:A:74:SER:CB	0.55	2.31	17	3
1:A:121:GLU:O	1:A:122:LEU:HD22	0.55	2.01	12	5
1:A:87:LYS:HE3	1:A:95:LEU:HD12	0.55	1.78	14	1
1:A:84:LEU:HD12	1:A:96:LEU:HD12	0.54	1.79	14	8
1:A:67:LEU:HD13	1:A:93:ILE:CD1	0.54	2.32	16	6
1:A:64:GLY:HA3	1:A:122:LEU:HD13	0.54	1.77	11	3
1:A:108:ALA:HB1	1:A:118:VAL:CG2	0.54	2.31	9	6
1:A:105:ASN:OD1	1:A:120:ILE:HG21	0.53	2.03	6	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:87:LYS:HD3	1:A:95:LEU:HD12	0.53	1.81	8	1
1:A:102:THR:CG2	1:A:107:ILE:HD11	0.52	2.34	4	2
1:A:108:ALA:HB1	1:A:118:VAL:HG21	0.52	1.82	9	3
1:A:65:THR:HG22	1:A:96:LEU:O	0.52	2.05	17	4
1:A:68:VAL:HA	1:A:118:VAL:HG12	0.52	1.82	6	14
1:A:67:LEU:HD12	1:A:119:GLN:HB3	0.51	1.82	14	16
1:A:84:LEU:HD21	1:A:107:ILE:HD13	0.51	1.81	12	1
1:A:68:VAL:HG21	1:A:78:LEU:CD1	0.51	2.36	11	8
1:A:67:LEU:HB3	1:A:93:ILE:HD13	0.51	1.83	12	4
1:A:88:GLN:HG3	1:A:95:LEU:HD21	0.51	1.81	6	3
1:A:84:LEU:HD21	1:A:107:ILE:CD1	0.51	2.36	12	1
1:A:22:LEU:O	1:A:23:ILE:HD13	0.50	2.05	17	1
1:A:79:ALA:HB1	1:A:84:LEU:O	0.49	2.08	5	20
1:A:93:ILE:HD12	1:A:93:ILE:O	0.48	2.08	12	3
1:A:23:ILE:O	1:A:23:ILE:HG23	0.47	2.09	14	6
1:A:88:GLN:CG	1:A:95:LEU:HD21	0.46	2.39	6	2
1:A:87:LYS:CE	1:A:95:LEU:HD12	0.46	2.40	14	1
1:A:22:LEU:HD13	1:A:23:ILE:N	0.46	2.25	10	2
1:A:75:ALA:HB1	1:A:86:PHE:HB2	0.45	1.87	12	4
1:A:111:LEU:O	1:A:116:VAL:HG12	0.45	2.12	6	1
1:A:111:LEU:C	1:A:116:VAL:HG12	0.45	2.32	12	4
1:A:21:TYR:C	1:A:22:LEU:HD22	0.44	2.33	15	1
1:A:22:LEU:HD12	1:A:22:LEU:O	0.44	2.13	8	1
1:A:67:LEU:HD13	1:A:93:ILE:HD13	0.44	1.89	10	1
1:A:67:LEU:CD1	1:A:93:ILE:HD12	0.43	2.43	5	1
1:A:75:ALA:HB1	1:A:86:PHE:CD1	0.42	2.50	11	1
1:A:106:ALA:O	1:A:109:THR:HG22	0.42	2.15	20	1
1:A:67:LEU:HD11	1:A:121:GLU:OE2	0.41	2.14	11	1
1:A:22:LEU:HD13	1:A:23:ILE:H	0.41	1.75	7	1
1:A:21:TYR:O	1:A:22:LEU:HD22	0.41	2.15	15	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	77/133 (58%)	70±2 (91±2%)	7±2 (9±3%)	0±1 (1±1%)	32 76
All	All	1540/2660 (58%)	1401 (91%)	131 (9%)	8 (1%)	32 76

All 2 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	64	GLY	5
1	A	8	GLY	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	62/107 (58%)	52±2 (85±3%)	10±2 (15±3%)	5 43
All	All	1240/2140 (58%)	1048 (85%)	192 (15%)	5 43

All 33 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	65	THR	20
1	A	66	LEU	20
1	A	95	LEU	20
1	A	104	LEU	20
1	A	71	ASP	18
1	A	76	LYS	9
1	A	9	LYS	8
1	A	110	LYS	7
1	A	121	GLU	7
1	A	122	LEU	6
1	A	22	LEU	5
1	A	99	LYS	5
1	A	63	SER	5
1	A	15	GLU	4
1	A	25	ASP	4
1	A	88	GLN	3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	74	SER	3
1	A	103	ASP	3
1	A	85	ASN	3
1	A	6	MET	3
1	A	13	THR	2
1	A	105	ASN	2
1	A	14	ILE	2
1	A	20	THR	2
1	A	93	ILE	2
1	A	113	SER	2
1	A	81	ARG	1
1	A	7	ASP	1
1	A	117	ASN	1
1	A	87	LYS	1
1	A	96	LEU	1
1	A	19	GLN	1
1	A	10	GLN	1

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 79% 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	1329
Number of shifts mapped to atoms	1329
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	122	-0.26 ± 0.12	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	110	0.12 ± 0.08	None needed (< 0.5 ppm)
$^{13}C'$	69	-0.08 ± 0.08	None needed (< 0.5 ppm)
^{15}N	119	-0.63 ± 0.27	Should be applied

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. 789 atoms were assigned a chemical shift out of a possible 1007. 0 out of 14 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	340/387 (88%)	149/158 (94%)	120/154 (78%)	71/75 (95%)
Sidechain	449/585 (77%)	318/382 (83%)	131/185 (71%)	0/18 (0%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	0/35 (0%)	0/17 (0%)	0/17 (0%)	0/1 (0%)
Overall	789/1007 (78%)	467/557 (84%)	251/356 (71%)	71/94 (76%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 79%, i.e. 1329 atoms were assigned a chemical shift out of a possible 1682. 0 out of 23 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	565/652 (87%)	255/267 (96%)	191/260 (73%)	119/125 (95%)
Sidechain	764/995 (77%)	539/646 (83%)	225/309 (73%)	0/40 (0%)
Aromatic	0/35 (0%)	0/17 (0%)	0/17 (0%)	0/1 (0%)
Overall	1329/1682 (79%)	794/930 (85%)	416/586 (71%)	119/166 (72%)

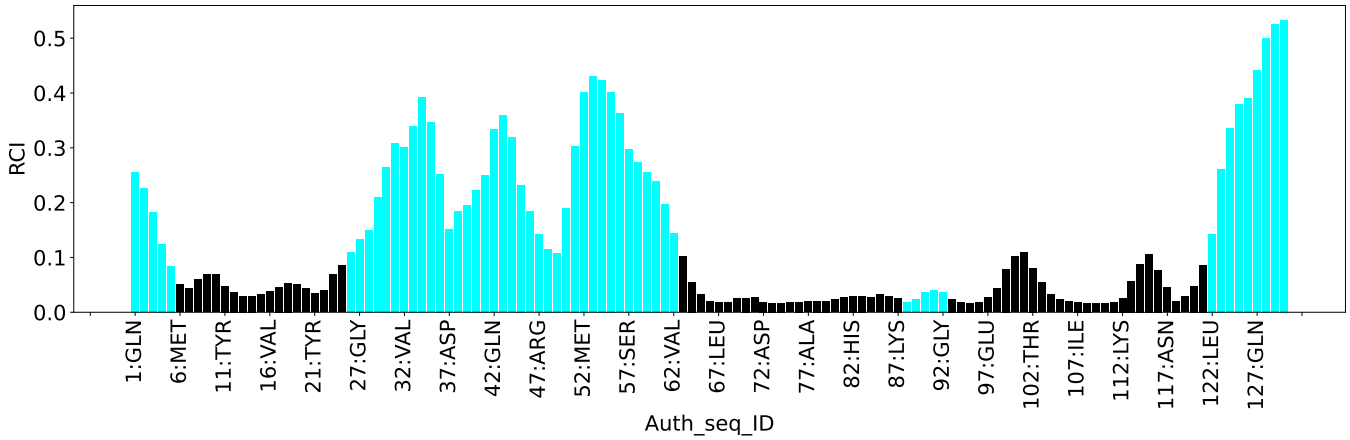
7.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

7.1.5 Random Coil Index (RCI) plots (i)

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

Random coil index (RCI) for chain A:



8 NMR restraints analysis (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	1253
Intra-residue (|i-j|=0)	240
Sequential (|i-j|=1)	345
Medium range (|i-j|>1 and |i-j|<5)	250
Long range (|i-j|≥5)	418
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	0
Number of restraints per residue	9.4
Number of long range restraints per residue[1]	3.1

[1]Long range hydrogen bonds and disulfide bonds are counted as long range restraints while calculating the number of long range restraints per residue

8.2 Residual restraint violations (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)	0.2	0.17
0.2-0.5 (Medium)	None	None
>0.5 (Large)	None	None

8.2.2 Average number of dihedral-angle violations per model (i)

Dihedral-angle violations less than 1° are not included in the calculation.There are no dihedral-angle violations

9 Distance violation analysis (i)

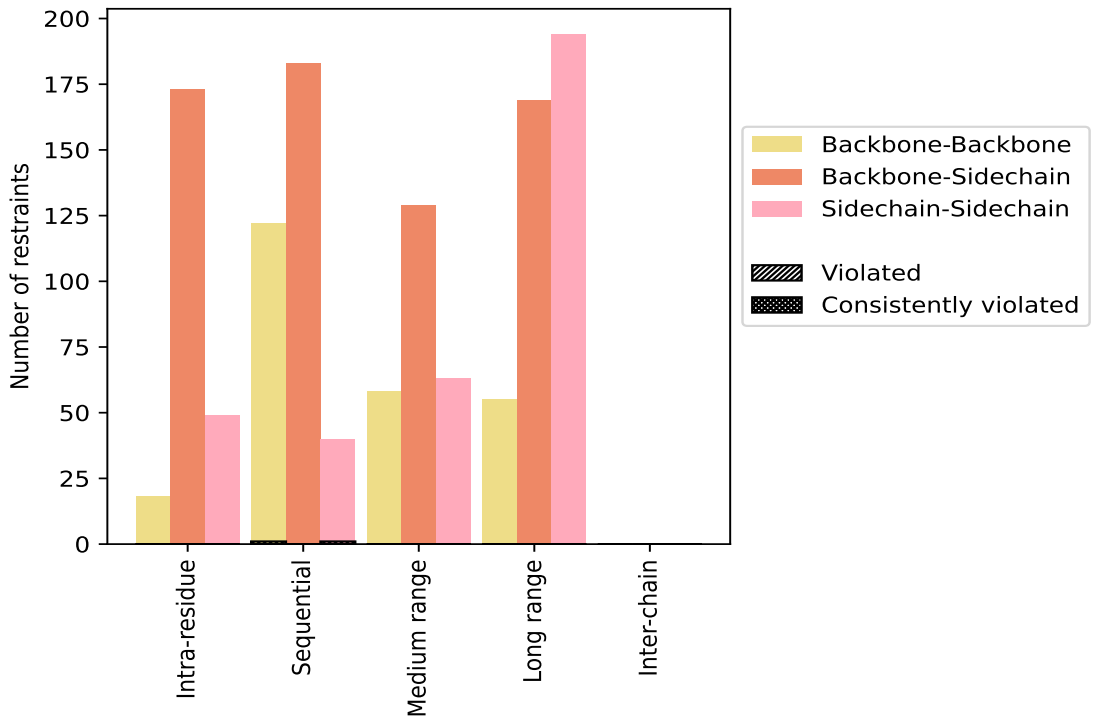
9.1 Summary of distance violations (i)

The following table shows the summary of distance violations in different restraint categories based on the sequence separation of the atoms involved. Each category is further sub-divided into three sub-categories based on the atoms involved. Violations less than 0.1 Å are not included in the statistics.

Restraints type	Count	%[1]	Violated[3]			Consistently Violated[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
Intra-residue (|i-j|=0)	240	19.2	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	18	1.4	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	173	13.8	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	49	3.9	0	0.0	0.0	0	0.0	0.0
Sequential (|i-j|=1)	345	27.5	2	0.6	0.2	0	0.0	0.0
Backbone-Backbone	122	9.7	1	0.8	0.1	0	0.0	0.0
Backbone-Sidechain	183	14.6	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	40	3.2	1	2.5	0.1	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	250	20.0	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	58	4.6	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	129	10.3	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	63	5.0	0	0.0	0.0	0	0.0	0.0
Long range (|i-j|≥5)	418	33.4	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	55	4.4	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	169	13.5	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	194	15.5	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	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	1253	100.0	2	0.2	0.2	0	0.0	0.0
Backbone-Backbone	253	20.2	1	0.4	0.1	0	0.0	0.0
Backbone-Sidechain	654	52.2	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	346	27.6	1	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 (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	0	0	0	0	0	0	0.0	0.0	0.0	0.0
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	0	0	0	0	0.0	0.0	0.0	0.0
5	0	0	0	0	0	0	0.0	0.0	0.0	0.0
6	0	0	0	0	0	0	0.0	0.0	0.0	0.0
7	0	0	0	0	0	0	0.0	0.0	0.0	0.0
8	0	0	0	0	0	0	0.0	0.0	0.0	0.0
9	0	1	0	0	0	1	0.16	0.16	0.0	0.16
10	0	0	0	0	0	0	0.0	0.0	0.0	0.0
11	0	0	0	0	0	0	0.0	0.0	0.0	0.0

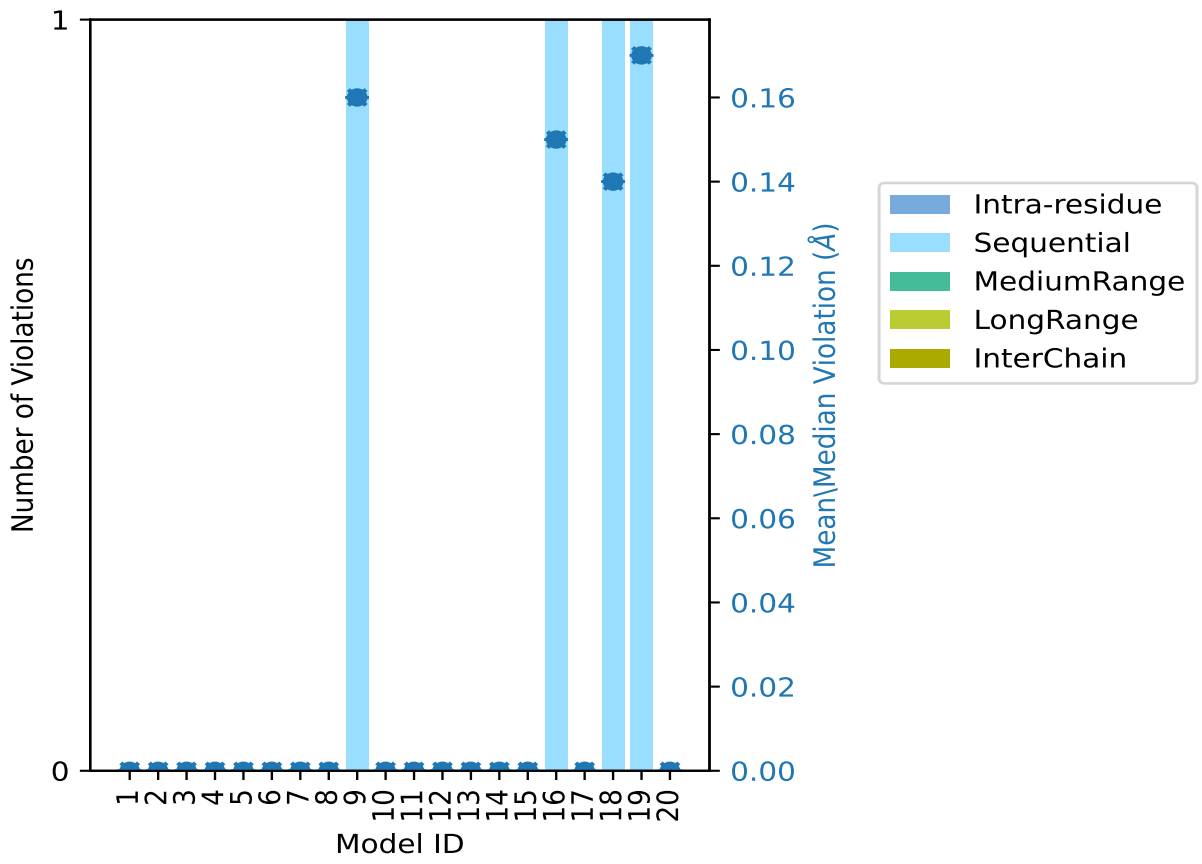
Continued on next page...

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	0	0	0	0	0.0	0.0	0.0	0.0
14	0	0	0	0	0	0	0.0	0.0	0.0	0.0
15	0	0	0	0	0	0	0.0	0.0	0.0	0.0
16	0	1	0	0	0	1	0.15	0.15	0.0	0.15
17	0	0	0	0	0	0	0.0	0.0	0.0	0.0
18	0	1	0	0	0	1	0.14	0.14	0.0	0.14
19	0	1	0	0	0	1	0.17	0.17	0.0	0.17
20	0	0	0	0	0	0	0.0	0.0	0.0	0.0

[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 (i)



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

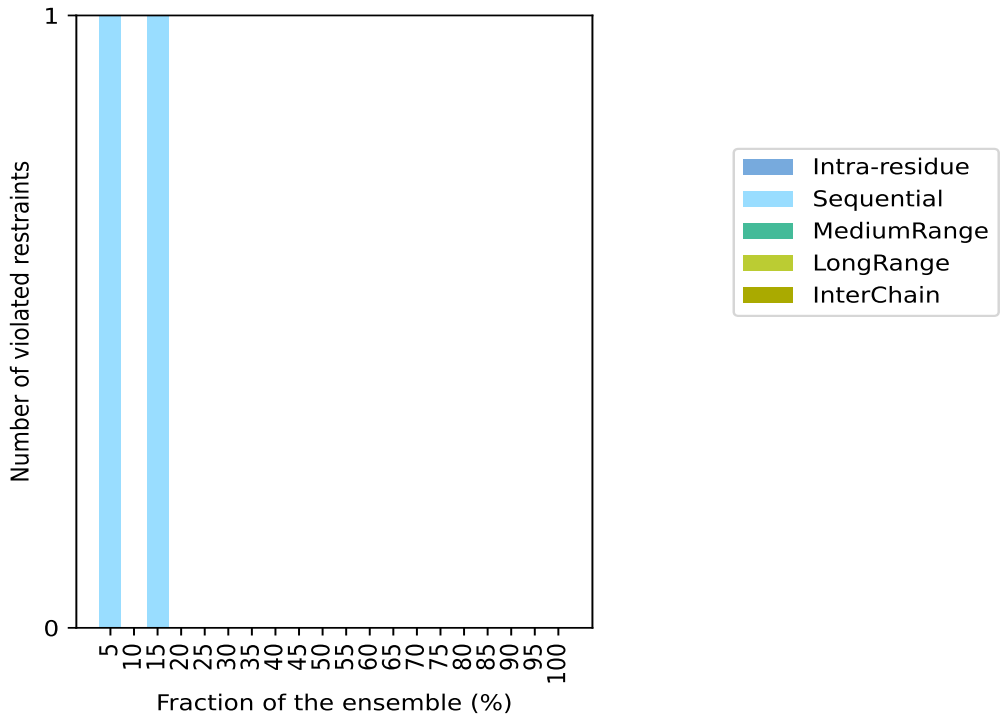
9.3 Distance violation statistics for the ensemble (i)

Violation analysis may find that some restraints are violated in few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of the ensemble. In total, 1251(IR:240, SQ:343, MR:250, LR:418, 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	1	0	0	0	1	1	5.0
0	0	0	0	0	0	2	10.0
0	1	0	0	0	1	3	15.0
0	0	0	0	0	0	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

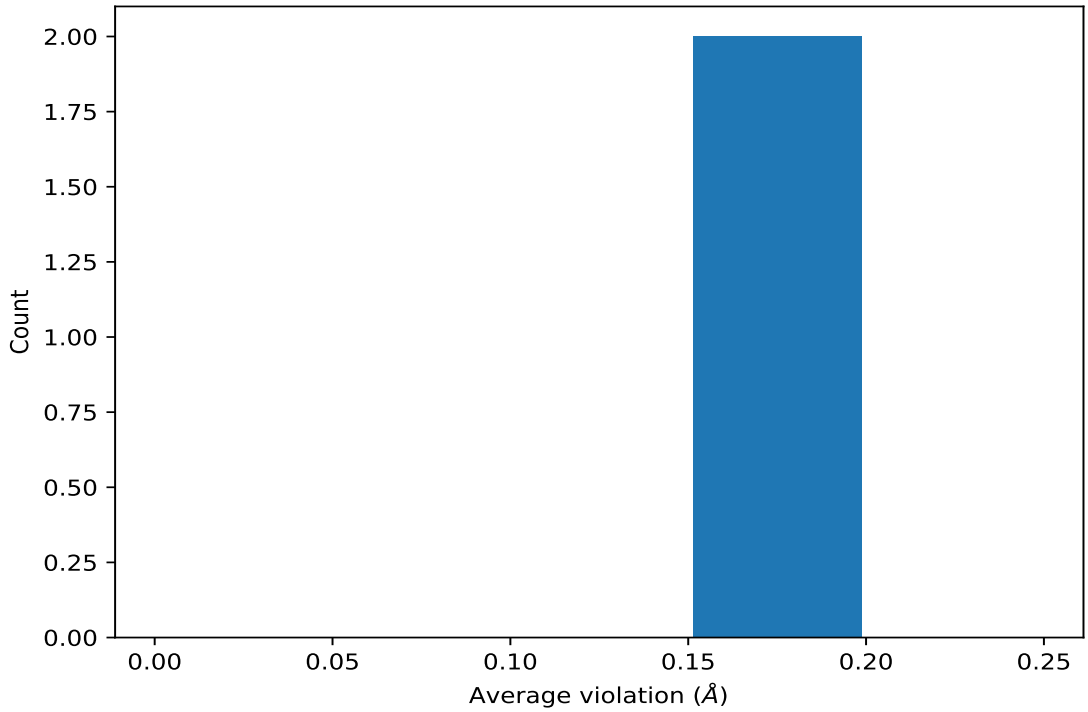
9.3.1 Bar graph : Distance violation statistics for the ensemble (i)



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 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.

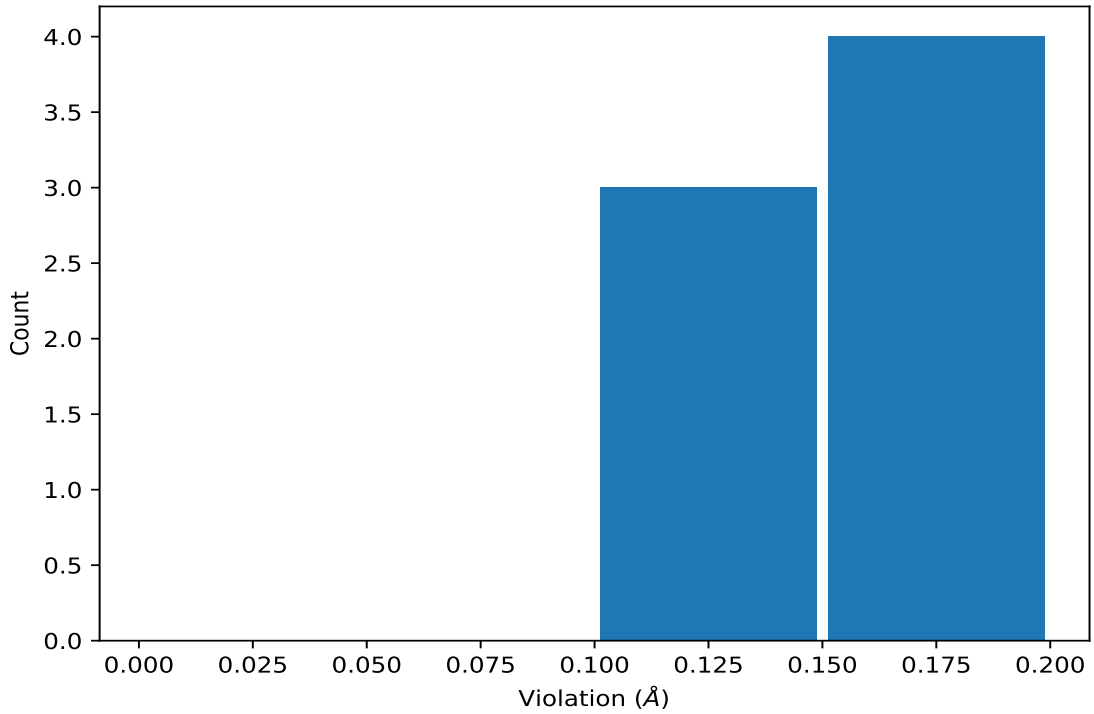
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,388)	1:A:39:LYS:HB2	1:A:40:VAL:HB	3	0.16	0.01	0.16
(1,388)	1:A:39:LYS:HB3	1:A:40:VAL:HB	3	0.16	0.01	0.16

[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)
(1,388)	1:A:39:LYS:HB2	1:A:40:VAL:HB	19	0.17
(1,388)	1:A:39:LYS:HB3	1:A:40:VAL:HB	19	0.17
(1,388)	1:A:39:LYS:HB2	1:A:40:VAL:HB	9	0.16
(1,388)	1:A:39:LYS:HB3	1:A:40:VAL:HB	9	0.16
(1,388)	1:A:39:LYS:HB2	1:A:40:VAL:HB	16	0.15
(1,388)	1:A:39:LYS:HB3	1:A:40:VAL:HB	16	0.15
(1,634)	1:A:61:THR:HA	1:A:62:VAL:H	18	0.14

10 Dihedral-angle violation analysis (i)

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