wwPDB NMR Structure Validation Summary Report (i)

Jun 3, 2023 – 10:09 PM EDT

PDB ID	:	7LOH
BMRB ID	:	30855
Title	:	Structure of the HIV-1 gp41 transmembrane domain and cytoplasmic tail
Authors	:	Piai, A.; Fu, Q.; Sharp, A.K.; Bighi, B.; Brown, A.M.; Chou, J.J.
Deposited on	:	2021-02-10

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:

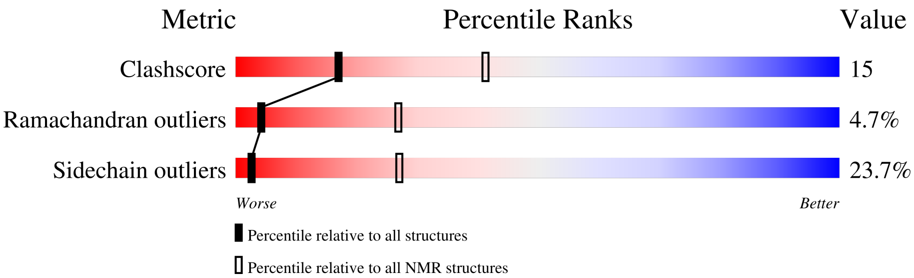
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal 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 8%.

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	180	39% 34% 6% 8% 12%
1	B	180	47% 29% 5% 7% 12%
1	C	180	47% 32% • 7% 12%

2 Ensemble composition and analysis (i)

This entry contains 15 models. Model 5 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:692-A:716, A:739-A:856, B:689-B:716, B:739-B:856, C:690-C:716, C:739-C:856 (434)	1.42	5

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

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

The models can be grouped into 4 clusters. No single-model clusters were found.

Cluster number	Models
1	3, 4, 5, 6, 11, 15
2	1, 2, 10, 13
3	9, 12, 14
4	7, 8

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 8049 atoms, of which 4131 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Transmembrane protein gp41.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
1	A	158	2683	844	1377	245	217	0
1	B	158	2683	844	1377	245	217	0
1	C	158	2683	844	1377	245	217	0

There are 24 discrepancies between the modelled and reference sequences:

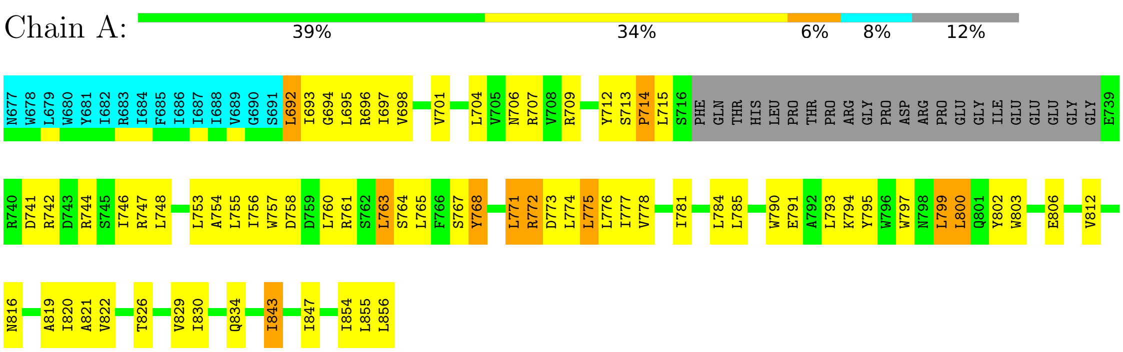
Chain	Residue	Modelled	Actual	Comment	Reference
A	683	ARG	LYS	conflict	UNP A0A386YSI0
A	687	ILE	MET	conflict	UNP A0A386YSI0
A	691	SER	GLY	conflict	UNP A0A386YSI0
A	693	ILE	VAL	conflict	UNP A0A386YSI0
A	704	LEU	ILE	conflict	UNP A0A386YSI0
A	764	SER	CYS	engineered mutation	UNP A0A386YSI0
A	823	GLY	ALA	engineered mutation	UNP A0A386YSI0
A	837	SER	CYS	engineered mutation	UNP A0A386YSI0
B	683	ARG	LYS	conflict	UNP A0A386YSI0
B	687	ILE	MET	conflict	UNP A0A386YSI0
B	691	SER	GLY	conflict	UNP A0A386YSI0
B	693	ILE	VAL	conflict	UNP A0A386YSI0
B	704	LEU	ILE	conflict	UNP A0A386YSI0
B	764	SER	CYS	engineered mutation	UNP A0A386YSI0
B	823	GLY	ALA	engineered mutation	UNP A0A386YSI0
B	837	SER	CYS	engineered mutation	UNP A0A386YSI0
C	683	ARG	LYS	conflict	UNP A0A386YSI0
C	687	ILE	MET	conflict	UNP A0A386YSI0
C	691	SER	GLY	conflict	UNP A0A386YSI0
C	693	ILE	VAL	conflict	UNP A0A386YSI0
C	704	LEU	ILE	conflict	UNP A0A386YSI0
C	764	SER	CYS	engineered mutation	UNP A0A386YSI0
C	823	GLY	ALA	engineered mutation	UNP A0A386YSI0
C	837	SER	CYS	engineered mutation	UNP A0A386YSI0

4 Residue-property plots (i)

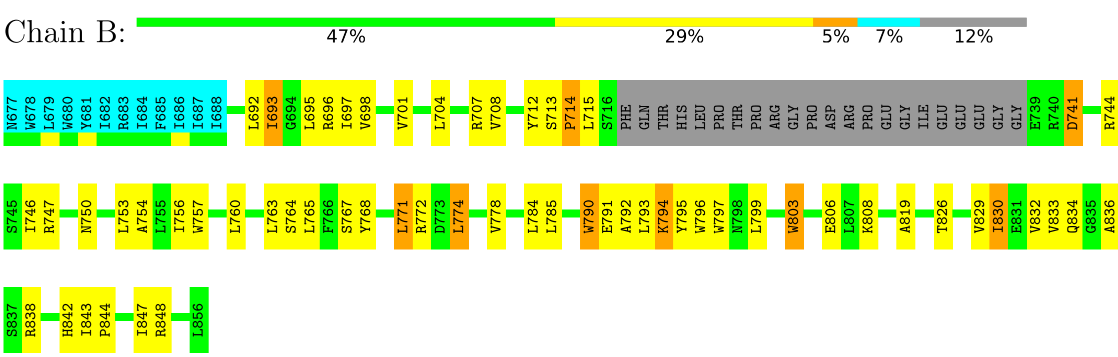
4.1 Average score per residue in the NMR ensemble

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.

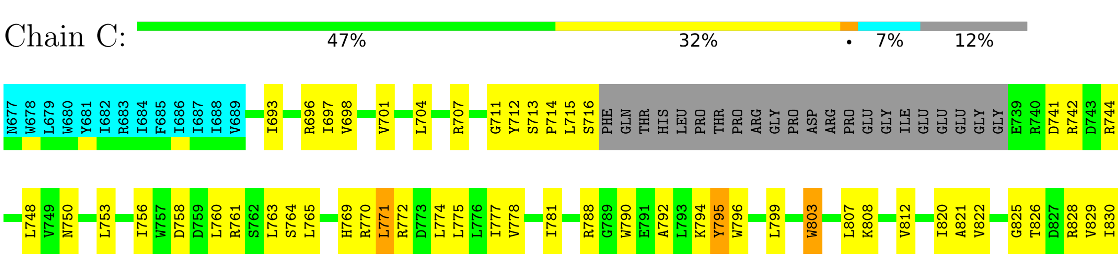
- Molecule 1: Transmembrane protein gp41

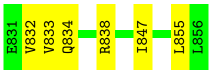


- Molecule 1: Transmembrane protein gp41



- Molecule 1: Transmembrane protein gp41

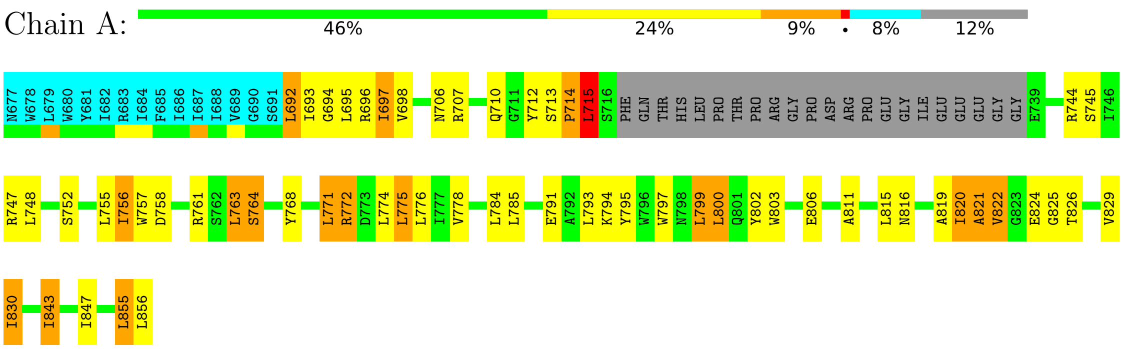




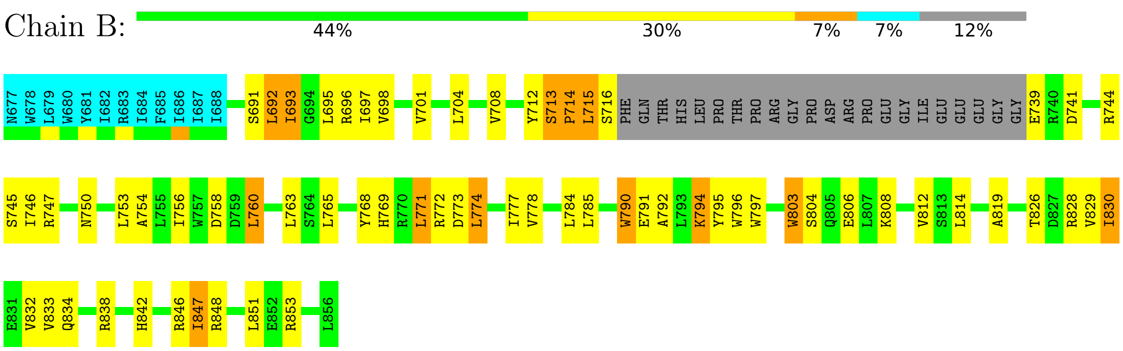
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

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

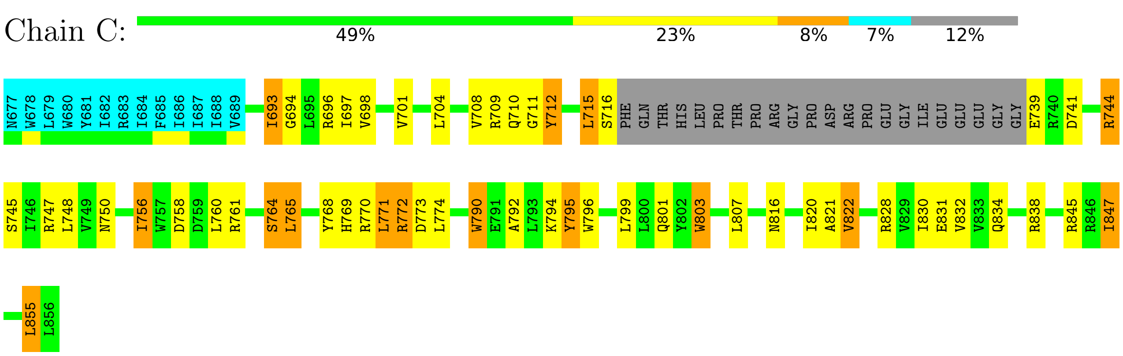
- Molecule 1: Transmembrane protein gp41



- Molecule 1: Transmembrane protein gp41



- Molecule 1: Transmembrane protein gp41



5 Refinement protocol and experimental data overview (i)

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

Of the 150 calculated structures, 15 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
X-PLOR NIH	refinement	
X-PLOR NIH	structure calculation	

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	596
Number of shifts mapped to atoms	561
Number of unparsed shifts	0
Number of shifts with mapping errors	35
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	8%

6 Model quality (i)

6.1 Standard geometry (i)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the (average) root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	0.81±0.01	0±0/1189 (0.0± 0.0%)	0.96±0.02	0±1/1606 (0.0± 0.1%)
1	B	0.79±0.01	0±0/1206 (0.0± 0.0%)	0.96±0.01	0±0/1629 (0.0± 0.0%)
1	C	0.80±0.01	0±0/1199 (0.0± 0.0%)	0.97±0.01	0±0/1619 (0.0± 0.0%)
All	All	0.80	0/53910 (0.0%)	0.96	10/72810 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	Chirality	Planarity
1	B	0.0±0.0	0.1±0.3
1	C	0.0±0.0	0.1±0.2
All	All	0	3

There are no bond-length outliers.

5 of 10 unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	795	TYR	CB-CG-CD1	-6.37	117.18	121.00	13	1
1	C	742	ARG	NE-CZ-NH2	5.65	123.12	120.30	14	1
1	A	712	TYR	CB-CG-CD1	-5.58	117.65	121.00	9	1
1	B	768	TYR	CB-CG-CD2	-5.30	117.82	121.00	13	1
1	C	795	TYR	CB-CG-CD1	-5.27	117.84	121.00	1	1

There are no chirality outliers.

All unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	C	761	ARG	Sidechain	1
1	B	761	ARG	Sidechain	1
1	B	707	ARG	Sidechain	1

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	1171	1235	1232	48±8
1	B	1188	1252	1249	42±8
1	C	1181	1243	1240	42±5
All	All	53100	55950	55815	1674

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 15.

5 of 776 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:754:ALA:HB1	1:B:715:LEU:HD12	0.89	1.41	1	1
1:A:692:LEU:HD12	1:A:693:ILE:N	0.82	1.90	12	4
1:A:692:LEU:HD23	1:A:696:ARG:CZ	0.80	2.06	13	1
1:B:771:LEU:HD22	1:B:772:ARG:N	0.78	1.93	7	5
1:A:756:ILE:HD13	1:C:795:TYR:CZ	0.78	2.13	13	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	140/180 (78%)	128±2 (91±1%)	7±2 (5±1%)	6±2 (4±1%)	5	30
1	B	143/180 (79%)	129±2 (90±1%)	7±1 (5±1%)	7±2 (5±1%)	4	26

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	142/180 (79%)	129±2 (91±2%)	6±2 (4±1%)	7±1 (5±1%)	4	26
All	All	6375/8100 (79%)	5787 (91%)	291 (5%)	297 (5%)	4	27

5 of 53 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	B	741	ASP	15
1	C	821	ALA	15
1	B	714	PRO	14
1	B	844	PRO	14
1	C	714	PRO	12

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	125/157 (80%)	92±3 (73±2%)	33±3 (27±2%)	2	22
1	B	127/157 (81%)	99±4 (78±3%)	28±4 (22±3%)	3	30
1	C	126/157 (80%)	98±3 (78±2%)	28±3 (22±2%)	3	29
All	All	5670/7065 (80%)	4325 (76%)	1345 (24%)	2	27

5 of 288 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	794	LYS	15
1	A	843	ILE	15
1	B	693	ILE	15
1	B	771	LEU	15
1	B	794	LYS	15

6.3.3 RNA ⓘ

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 8% for the well-defined parts and 8% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *HIV1_gp41_TMD-CT_chemical_shifts.tab*

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	596
Number of shifts mapped to atoms	561
Number of unparsed shifts	0
Number of shifts with mapping errors	35
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

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 35) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	717	PHE	H	8.05	0.01	1
1	A	717	PHE	C	175.43	0.05	1
1	A	717	PHE	CA	57.49	0.05	1
1	A	717	PHE	N	120.86	0.03	1
1	A	718	GLN	H	8.13	0.01	1
1	A	718	GLN	C	175.94	0.05	1
1	A	718	GLN	CA	55.56	0.05	1
1	A	718	GLN	N	120.21	0.03	1
1	A	719	THR	H	8.03	0.01	1
1	A	719	THR	CA	61.81	0.05	1
1	A	719	THR	N	114.2	0.03	1
1	A	720	HIS	C	174.75	0.05	1
1	A	720	HIS	CA	55.7	0.05	1
1	A	721	LEU	H	8.02	0.01	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	721	LEU	CA	52.47	0.05	1
1	A	721	LEU	N	123.97	0.03	1
1	A	722	PRO	C	176.74	0.05	1
1	A	722	PRO	CA	62.54	0.05	1
1	A	723	THR	H	8.15	0.01	1
1	A	723	THR	CA	59.45	0.05	1
1	A	723	THR	N	116.61	0.03	1
1	A	724	PRO	C	177.1	0.05	1
1	A	724	PRO	CA	62.91	0.05	1
1	A	725	ARG	H	8.45	0.01	1
1	A	725	ARG	C	177.14	0.05	1
1	A	725	ARG	CA	55.87	0.05	1
1	A	725	ARG	N	121.18	0.03	1
1	A	737	GLY	H	8.48	0.01	1
1	A	737	GLY	C	174.84	0.05	1
1	A	737	GLY	CA	45.06	0.05	1
1	A	737	GLY	N	110.12	0.03	1
1	A	738	GLY	H	8.32	0.01	1
1	A	738	GLY	C	174.61	0.05	1
1	A	738	GLY	CA	44.88	0.05	1
1	A	738	GLY	N	108.9	0.03	1

7.1.2 Chemical shift referencing ⓘ

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	157	-0.02 ± 0.12	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	0	—	None (insufficient data)
$^{13}C'$	145	-0.29 ± 0.11	None needed (< 0.5 ppm)
^{15}N	147	0.80 ± 0.18	Should be applied

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 8%, i.e. 508 atoms were assigned a chemical shift out of a possible 6514. 0 out of 115 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	508/2187 (23%)	126/891 (14%)	256/868 (29%)	126/428 (29%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Sidechain	0/3931 (0%)	0/2570 (0%)	0/1145 (0%)	0/216 (0%)
Aromatic	0/396 (0%)	0/192 (0%)	0/177 (0%)	0/27 (0%)
Overall	508/6514 (8%)	126/3653 (3%)	256/2190 (12%)	126/671 (19%)

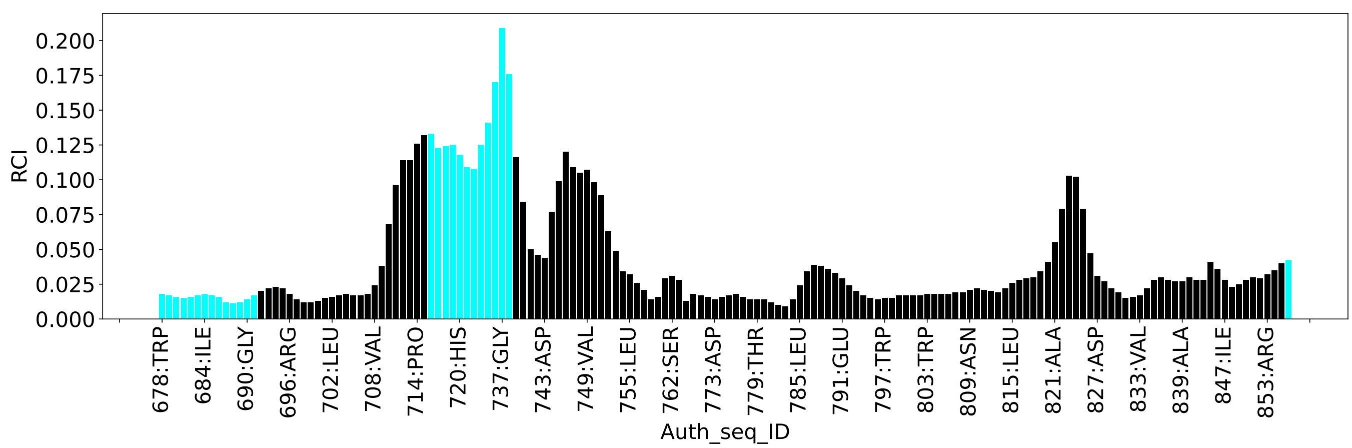
7.1.4 Statistically unusual chemical shifts (i)

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	1479
Intra-residue (|i-j|=0)	168
Sequential (|i-j|=1)	402
Medium range (|i-j|>1 and |i-j|<5)	483
Long range (|i-j|≥5)	168
Inter-chain	0
Hydrogen bond restraints	258
Disulfide bond restraints	0
Total dihedral-angle restraints	278
Number of unmapped restraints	38
Number of restraints per residue	3.3
Number of long range restraints per residue[1]	0.3

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

8.2 Residual restraint violations (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)	97.4	0.2
0.2-0.5 (Medium)	133.6	0.5
>0.5 (Large)	150.6	52.15

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

Dihedral-angle violations less than 1° are not included in the calculation.

Bins (°)	Average number of violations per model	Max (°)
1.0-10.0 (Small)	49.5	7.1
10.0-20.0 (Medium)	None	None
>20.0 (Large)	None	None

9 Distance violation analysis (i)

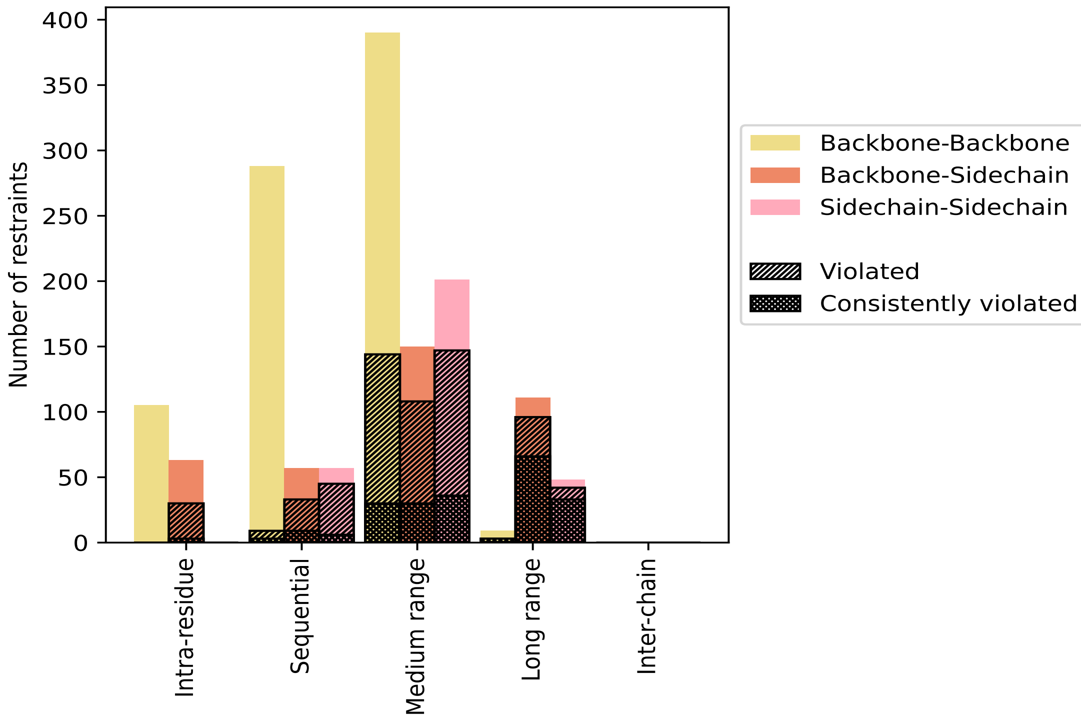
9.1 Summary of distance violations (i)

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)	168	11.4	30	17.9	2.0	3	1.8	0.2
Backbone-Backbone	105	7.1	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	63	4.3	30	47.6	2.0	3	4.8	0.2
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sequential (|i-j|=1)	402	27.2	87	21.6	5.9	18	4.5	1.2
Backbone-Backbone	288	19.5	9	3.1	0.6	3	1.0	0.2
Backbone-Sidechain	57	3.9	33	57.9	2.2	9	15.8	0.6
Sidechain-Sidechain	57	3.9	45	78.9	3.0	6	10.5	0.4
Medium range (|i-j|>1 & |i-j|<5)	483	32.7	327	67.7	22.1	87	18.0	5.9
Backbone-Backbone	132	8.9	72	54.5	4.9	21	15.9	1.4
Backbone-Sidechain	150	10.1	108	72.0	7.3	30	20.0	2.0
Sidechain-Sidechain	201	13.6	147	73.1	9.9	36	17.9	2.4
Long range (|i-j|≥5)	168	11.4	141	83.9	9.5	102	60.7	6.9
Backbone-Backbone	9	0.6	3	33.3	0.2	3	33.3	0.2
Backbone-Sidechain	111	7.5	96	86.5	6.5	66	59.5	4.5
Sidechain-Sidechain	48	3.2	42	87.5	2.8	33	68.8	2.2
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	258	17.4	72	27.9	4.9	9	3.5	0.6
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	1479	100.0	657	44.4	44.4	219	14.8	14.8
Backbone-Backbone	792	53.5	156	19.7	10.5	36	4.5	2.4
Backbone-Sidechain	381	25.8	267	70.1	18.1	108	28.3	7.3
Sidechain-Sidechain	306	20.7	234	76.5	15.8	75	24.5	5.1

[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	12	45	222	126	0	405	6.06	51.47	11.7	0.31
2	15	51	210	117	0	393	6.2	51.14	11.7	0.38
3	9	51	219	111	0	390	6.28	49.94	11.75	0.34
4	12	45	201	120	0	378	6.36	50.23	11.73	0.32
5	9	45	201	111	0	366	6.59	50.88	11.94	0.38
6	9	45	228	120	0	402	6.01	51.02	11.53	0.33
7	18	39	222	117	0	396	6.14	50.65	11.65	0.36
8	9	39	219	123	0	390	6.32	51.09	11.86	0.35
9	12	45	210	114	0	381	6.28	51.27	11.61	0.34
10	15	39	201	117	0	372	6.61	50.99	12.08	0.38
11	12	45	207	117	0	381	6.36	50.43	11.74	0.32

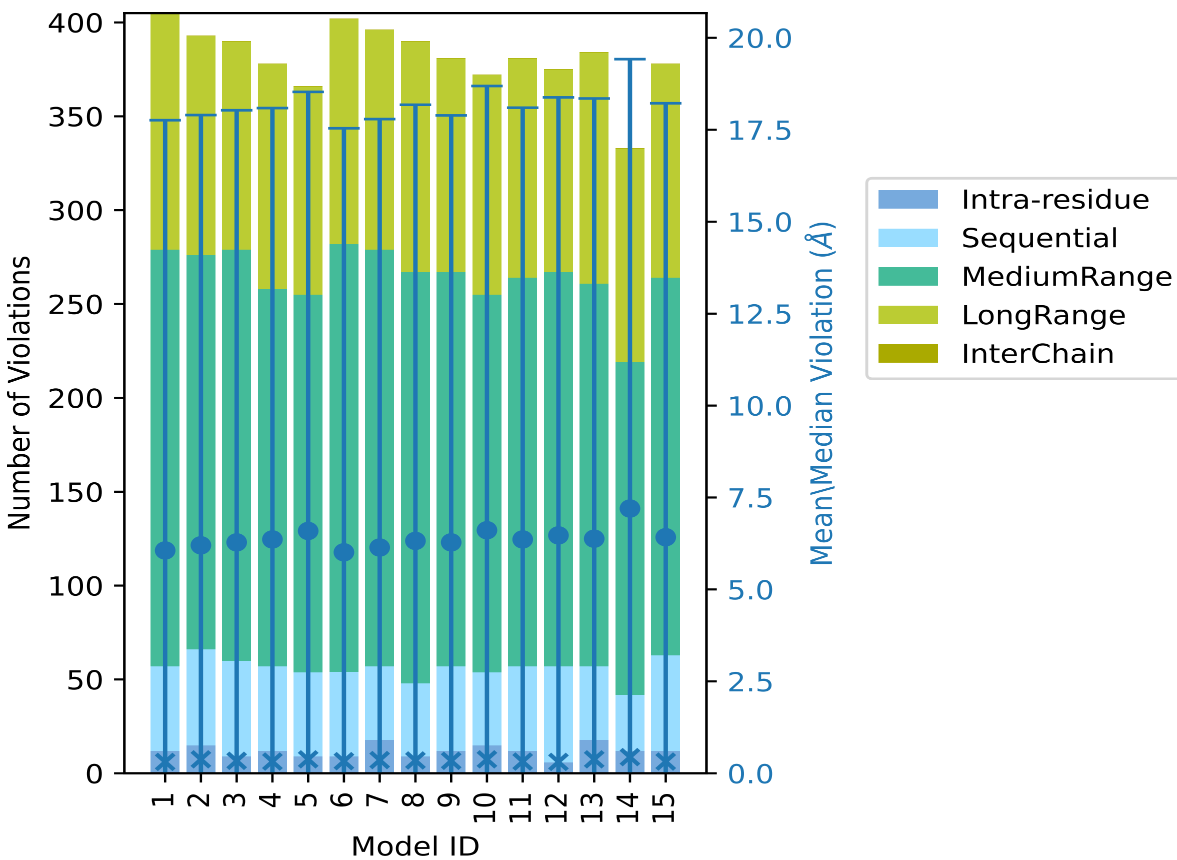
Continued on next page...

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	6	51	210	108	0	375	6.47	52.15	11.91	0.31
13	18	39	204	123	0	384	6.38	51.87	11.97	0.38
14	12	30	177	114	0	333	7.2	50.85	12.22	0.44
15	12	51	201	114	0	378	6.42	50.33	11.8	0.32

[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

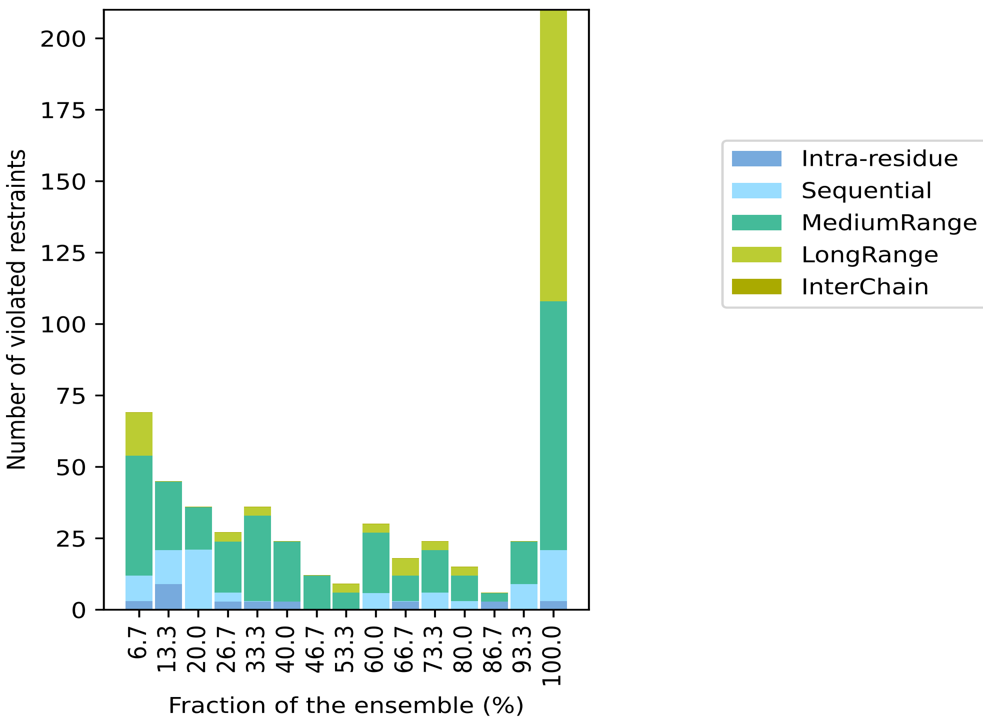
9.3 Distance violation statistics for the ensemble (i)

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, 636(IR:138, SQ:315, MR:156, LR:27, 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]	%
3	9	42	15	0	69	1	6.7
9	12	24	0	0	45	2	13.3
0	21	15	0	0	36	3	20.0
3	3	18	3	0	27	4	26.7
3	0	30	3	0	36	5	33.3
3	0	21	0	0	24	6	40.0
0	0	12	0	0	12	7	46.7
0	0	6	3	0	9	8	53.3
0	6	21	3	0	30	9	60.0
3	0	9	6	0	18	10	66.7
0	6	15	3	0	24	11	73.3
0	3	9	3	0	15	12	80.0
3	0	3	0	0	6	13	86.7
0	9	15	0	0	24	14	93.3
3	18	87	102	0	210	15	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

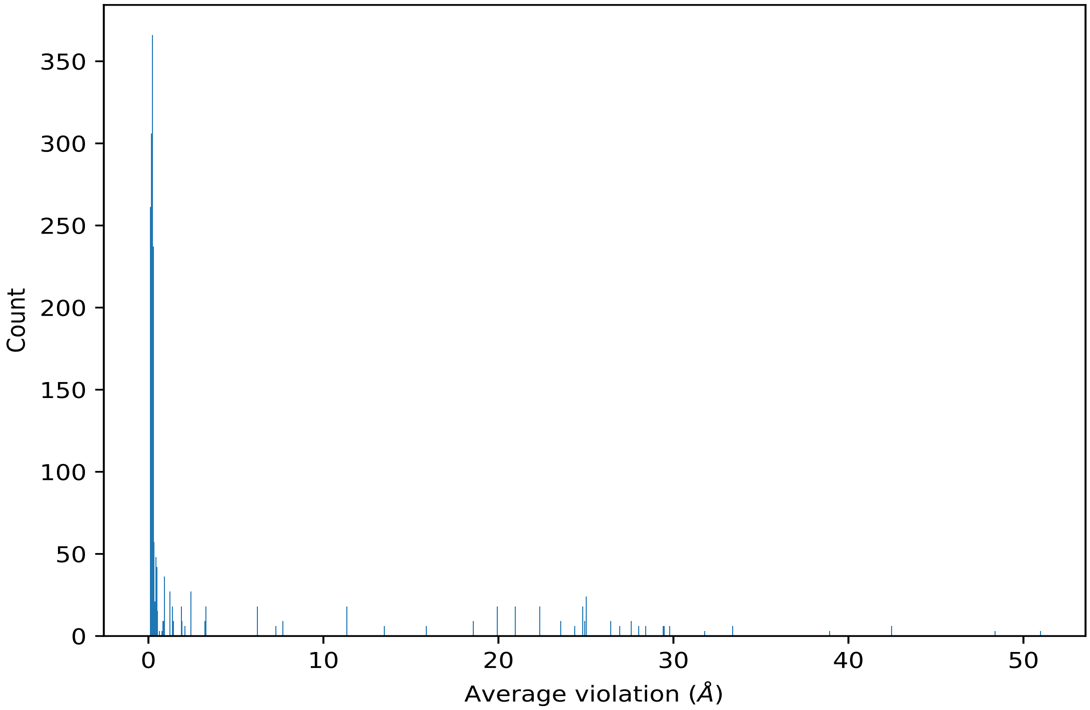
9.3.1 Bar graph : Distance violation statistics for the ensemble (i)



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.

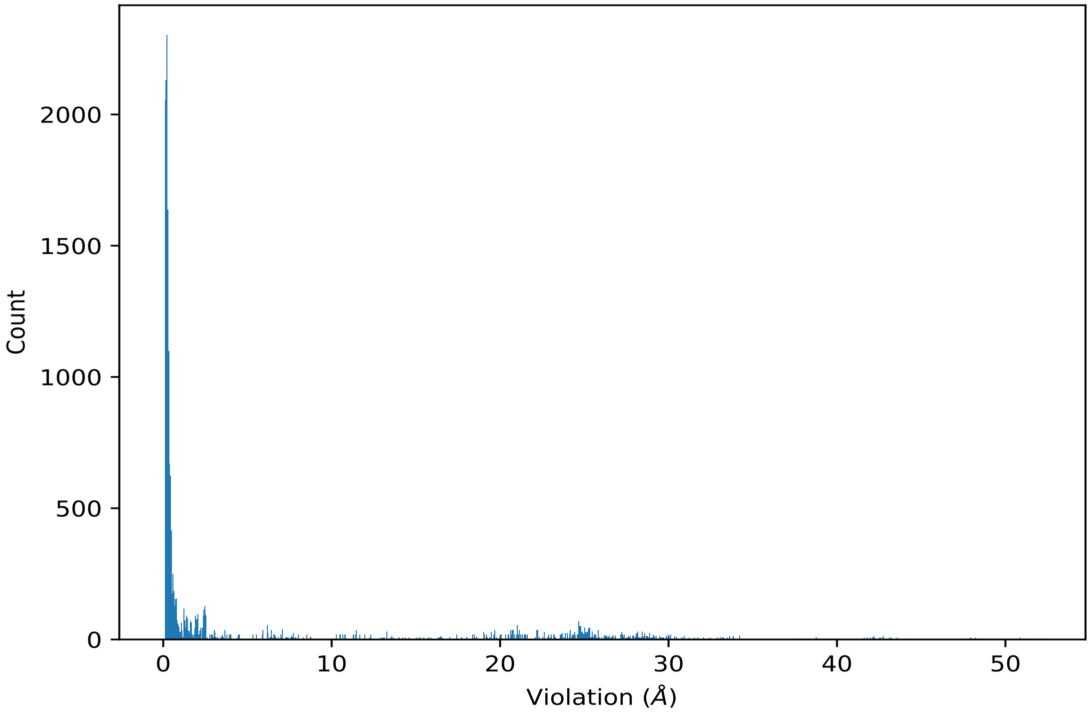
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,166)	1:C:806:GLU:H	1:C:772:ARG:HE	15	50.95	0.58	50.99
(1,167)	1:C:806:GLU:H	1:C:772:ARG:HE	15	50.95	0.58	50.99
(1,168)	1:C:806:GLU:H	1:C:772:ARG:HE	15	50.95	0.58	50.99
(1,163)	1:C:806:GLU:H	1:C:771:LEU:HG	15	48.4	0.58	48.22
(1,164)	1:C:806:GLU:H	1:C:771:LEU:HG	15	48.4	0.58	48.22
(1,165)	1:C:806:GLU:H	1:C:771:LEU:HG	15	48.4	0.58	48.22
(1,124)	1:C:771:LEU:H	1:C:806:GLU:HB2	15	42.49	0.53	42.5
(1,124)	1:C:771:LEU:H	1:C:806:GLU:HB3	15	42.49	0.53	42.5
(1,125)	1:C:771:LEU:H	1:C:806:GLU:HB2	15	42.49	0.53	42.5
(1,125)	1:C:771:LEU:H	1:C:806:GLU:HB3	15	42.49	0.53	42.5
(1,126)	1:C:771:LEU:H	1:C:806:GLU:HB2	15	42.49	0.53	42.5
(1,126)	1:C:771:LEU:H	1:C:806:GLU:HB3	15	42.49	0.53	42.5
(1,160)	1:C:803:TRP:HE1	1:C:764:SER:HG	15	38.92	0.64	38.8

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

9.5 All violated distance restraints

9.5.1 Histogram : Distribution of distance violations

The following histogram shows the distribution of the absolute value of the violation for all violated restraints in the ensemble.



9.5.2 Table : All distance violations

The following table provides the 10 worst performing restraints, sorted by the violation value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint. Rows with same key represent combinatorial or ambiguous restraints and are counted as a single restraint.

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,168)	1:C:806:GLU:H	1:C:772:ARG:HE	12	52.15
(1,167)	1:C:806:GLU:H	1:C:772:ARG:HE	12	52.15
(1,166)	1:C:806:GLU:H	1:C:772:ARG:HE	12	52.15
(1,168)	1:C:806:GLU:H	1:C:772:ARG:HE	13	51.87
(1,167)	1:C:806:GLU:H	1:C:772:ARG:HE	13	51.87
(1,166)	1:C:806:GLU:H	1:C:772:ARG:HE	13	51.87
(1,168)	1:C:806:GLU:H	1:C:772:ARG:HE	1	51.47
(1,167)	1:C:806:GLU:H	1:C:772:ARG:HE	1	51.47
(1,166)	1:C:806:GLU:H	1:C:772:ARG:HE	1	51.47
(1,168)	1:C:806:GLU:H	1:C:772:ARG:HE	9	51.27

10 Dihedral-angle violation analysis (i)

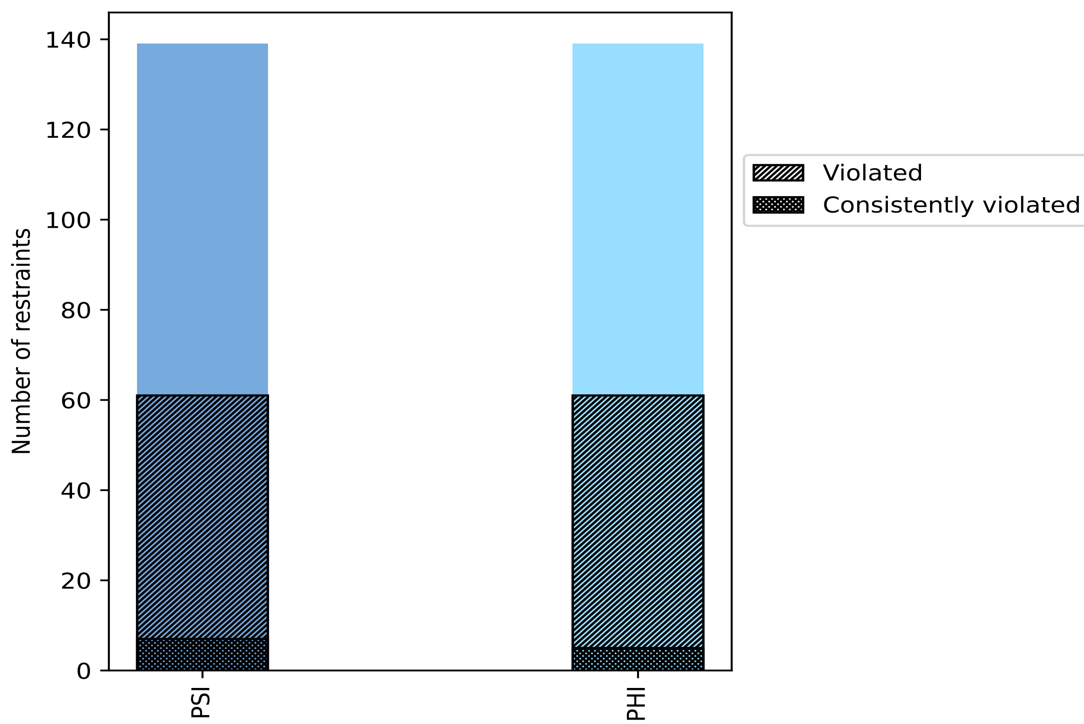
10.1 Summary of dihedral-angle violations (i)

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]
PSI	139	50.0	61	43.9	21.9	7	5.0	2.5
PHI	139	50.0	61	43.9	21.9	5	3.6	1.8
Total	278	100.0	122	43.9	43.9	12	4.3	4.3

[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)



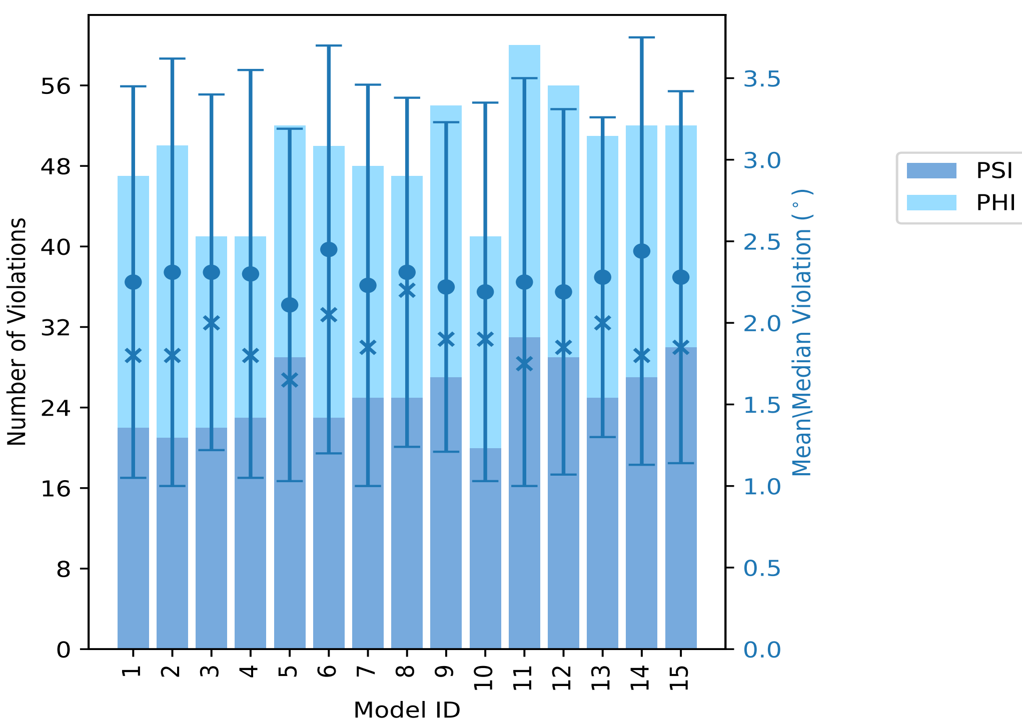
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

10.2 Dihedral-angle violation statistics for each model (i)

The following table provides the dihedral-angle violation statistics for each model in the ensemble. Violations less than 1° are not included in the statistics.

Model ID	Number of violations			Mean (°)	Max (°)	SD (°)	Median (°)
	PSI	PHI	Total				
1	22	25	47	2.25	5.8	1.2	1.8
2	21	29	50	2.31	6.5	1.31	1.8
3	22	19	41	2.31	5.2	1.09	2.0
4	23	18	41	2.3	6.2	1.25	1.8
5	29	23	52	2.11	5.5	1.08	1.65
6	23	27	50	2.45	6.3	1.25	2.05
7	25	23	48	2.23	7.1	1.23	1.85
8	25	22	47	2.31	5.6	1.07	2.2
9	27	27	54	2.22	4.8	1.01	1.9
10	20	21	41	2.19	5.3	1.16	1.9
11	31	29	60	2.25	6.9	1.25	1.75
12	29	27	56	2.19	6.4	1.12	1.85
13	25	26	51	2.28	5.1	0.98	2.0
14	27	25	52	2.44	6.1	1.31	1.8
15	30	22	52	2.28	5.7	1.14	1.85

10.2.1 Bar graph : Dihedral 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

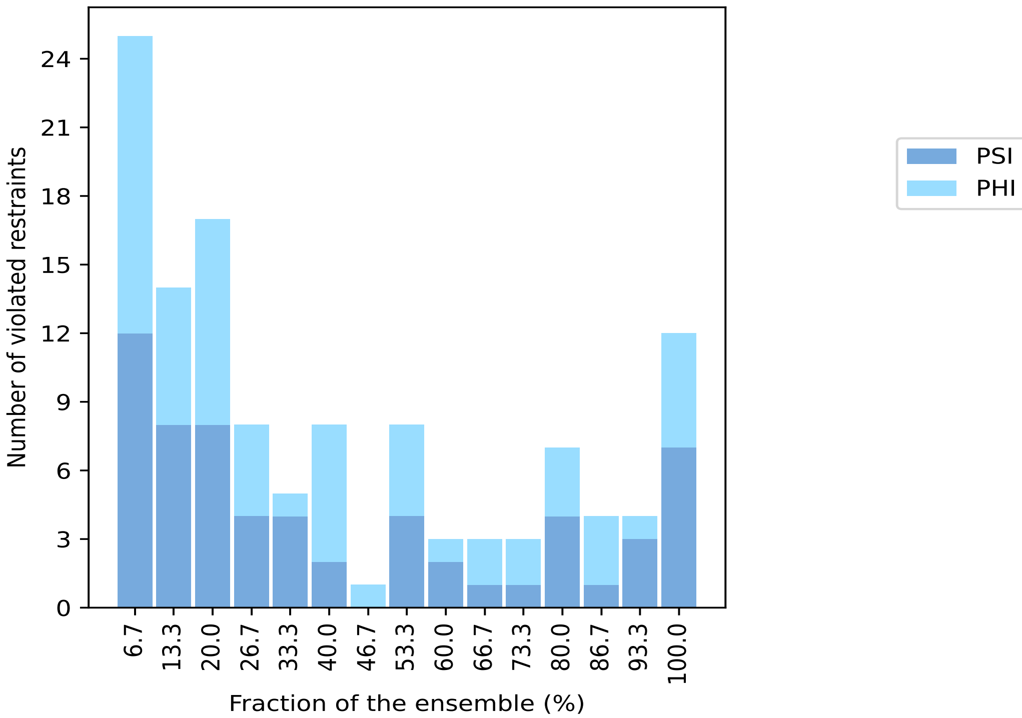
10.3 Dihedral-angle violation statistics for the ensemble (i)

Violation analysis may find that some restraints are violated in very 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 ensemble.

Number of violated restraints			Fraction of the ensemble	
PSI	PHI	Total	Count[1]	%
12	13	25	1	6.7
8	6	14	2	13.3
8	9	17	3	20.0
4	4	8	4	26.7
4	1	5	5	33.3
2	6	8	6	40.0
0	1	1	7	46.7
4	4	8	8	53.3
2	1	3	9	60.0
1	2	3	10	66.7
1	2	3	11	73.3
4	3	7	12	80.0
1	3	4	13	86.7
3	1	4	14	93.3
7	5	12	15	100.0

[1] Number of models with violations

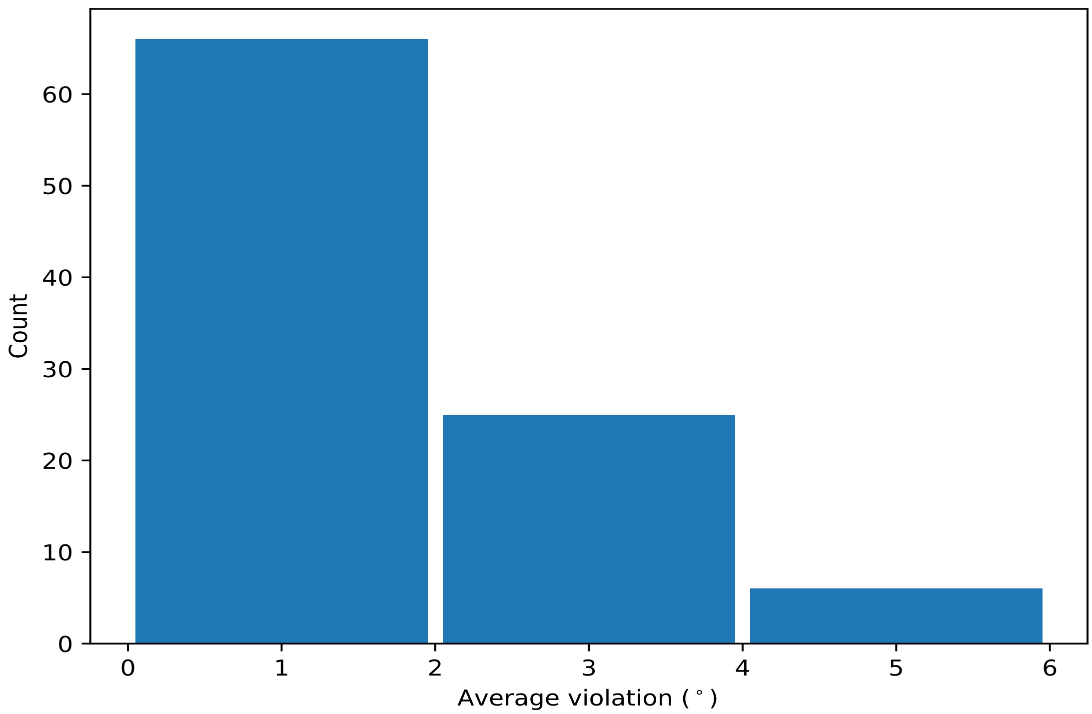
10.3.1 Bar graph : Dihedral-angle Violation statistics for the ensemble (i)



10.4 Most violated dihedral-angle restraints in the ensemble (i)

10.4.1 Histogram : Distribution of mean dihedral-angle 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



10.4.2 Table: Most violated dihedral-angle 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.

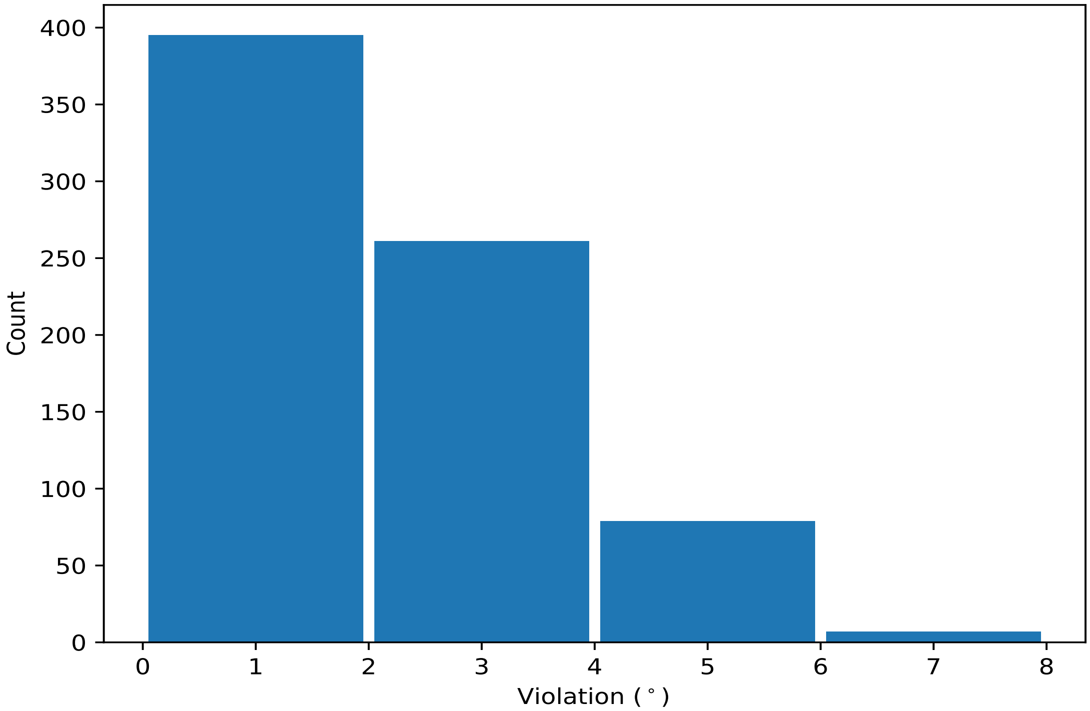
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models[1]	Mean	SD[2]	Median
(1,125)	1:C:763:LEU:C	1:C:764:SER:N	1:C:764:SER:CA	1:C:764:SER:C	15	4.73	1.46	4.9
(1,128)	1:C:770:ARG:N	1:C:770:ARG:CA	1:C:770:ARG:C	1:C:771:LEU:N	15	4.69	0.86	4.8
(1,168)	1:C:790:TRP:N	1:C:790:TRP:CA	1:C:790:TRP:C	1:C:791:GLU:N	15	4.49	0.99	4.3
(1,129)	1:C:770:ARG:C	1:C:771:LEU:N	1:C:771:LEU:CA	1:C:771:LEU:C	15	4.34	0.87	4.1
(1,37)	1:C:696:ARG:C	1:C:697:ILE:N	1:C:697:ILE:CA	1:C:697:ILE:C	15	4.08	0.69	4.1
(1,36)	1:C:696:ARG:N	1:C:696:ARG:CA	1:C:696:ARG:C	1:C:697:ILE:N	15	4.04	0.59	4.0
(1,42)	1:C:699:PHE:N	1:C:699:PHE:CA	1:C:699:PHE:C	1:C:700:ALA:N	15	3.59	0.95	3.6
(1,106)	1:C:754:ALA:N	1:C:754:ALA:CA	1:C:754:ALA:C	1:C:755:LEU:N	15	3.23	0.92	3.1
(1,43)	1:C:699:PHE:C	1:C:700:ALA:N	1:C:700:ALA:CA	1:C:700:ALA:C	15	2.97	0.83	3.1
(1,99)	1:C:750:ASN:C	1:C:751:GLY:N	1:C:751:GLY:CA	1:C:751:GLY:C	15	2.85	0.63	2.7

[1] Number of violated models, [2]Standard deviation, All angle values are in degree (°)

10.5 All violated dihedral-angle restraints (i)

10.5.1 Histogram : Distribution of violations (i)

The following histogram shows the distribution of the absolute value of the violation for all violated restraints in the ensemble.



10.5.2 Table: All violated dihedral-angle restraints (i)

The following table provides the list of violationsfor 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.

Key	Atom-1	Atom-2	Atom-3	Atom-4	Model ID	Violation (°)
(1,168)	1:C:790:TRP:N	1:C:790:TRP:CA	1:C:790:TRP:C	1:C:791:GLU:N	7	7.1
(1,125)	1:C:763:LEU:C	1:C:764:SER:N	1:C:764:SER:CA	1:C:764:SER:C	11	6.9
(1,129)	1:C:770:ARG:C	1:C:771:LEU:N	1:C:771:LEU:CA	1:C:771:LEU:C	2	6.5
(1,125)	1:C:763:LEU:C	1:C:764:SER:N	1:C:764:SER:CA	1:C:764:SER:C	12	6.4
(1,125)	1:C:763:LEU:C	1:C:764:SER:N	1:C:764:SER:CA	1:C:764:SER:C	6	6.3
(1,125)	1:C:763:LEU:C	1:C:764:SER:N	1:C:764:SER:CA	1:C:764:SER:C	4	6.2
(1,128)	1:C:770:ARG:N	1:C:770:ARG:CA	1:C:770:ARG:C	1:C:771:LEU:N	14	6.1
(1,128)	1:C:770:ARG:N	1:C:770:ARG:CA	1:C:770:ARG:C	1:C:771:LEU:N	2	5.8
(1,106)	1:C:754:ALA:N	1:C:754:ALA:CA	1:C:754:ALA:C	1:C:755:LEU:N	1	5.8
(1,125)	1:C:763:LEU:C	1:C:764:SER:N	1:C:764:SER:CA	1:C:764:SER:C	15	5.7