wwPDB NMR Structure Validation Summary Report

Aug 14, 2023 – 04:11 PM EDT

PDB ID	:	8G55
BMRB ID	:	31075
Title	:	Temperature-dependent structures of tau aggregates
Authors	:	El Mammeri, N.; Duan, P.; Dregni, A.J.; Hong, M.
Deposited on	:	2023-02-11

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:

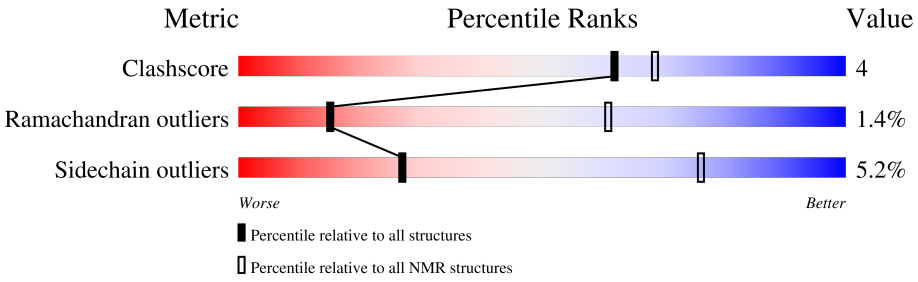
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLID-STATE NMR

The overall completeness of chemical shifts assignment is 3%.

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	202	24% 7% 69%
1	B	202	24% 7% 69%
1	C	202	24% 7% 69%
1	D	202	24% 7% 69%
1	E	202	24% 7% 69%
1	F	202	24% 7% 69%
1	G	202	24% 7% 69%
1	H	202	24% 7% 69%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	I	202	24% 7% 69%
1	J	202	24% 7% 69%

2 Ensemble composition and analysis

This entry contains 10 models. Model 4 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *medoid*.

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:274-A:322, B:274-B:322, C:274-C:322, D:274-D:322, E:274-E:322, F:274-F:322, G:274-G:322, H:274-H:322, I:274-I:322, J:274-J:322 (490)	3.48	4

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, 4, 5, 6, 7, 9
2	2, 3, 8, 10

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 9530 atoms, of which 4910 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Microtubule-associated protein tau.

Mol	Chain	Residues	Atoms						Trace
1	A	63	Total	C	H	N	O	S	0
			953	287	491	85	88	2	
1	B	63	Total	C	H	N	O	S	0
			953	287	491	85	88	2	
1	C	63	Total	C	H	N	O	S	0
			953	287	491	85	88	2	
1	D	63	Total	C	H	N	O	S	0
			953	287	491	85	88	2	
1	E	63	Total	C	H	N	O	S	0
			953	287	491	85	88	2	
1	F	63	Total	C	H	N	O	S	0
			953	287	491	85	88	2	
1	G	63	Total	C	H	N	O	S	0
			953	287	491	85	88	2	
1	H	63	Total	C	H	N	O	S	0
			953	287	491	85	88	2	
1	I	63	Total	C	H	N	O	S	0
			953	287	491	85	88	2	
1	J	63	Total	C	H	N	O	S	0
			953	287	491	85	88	2	

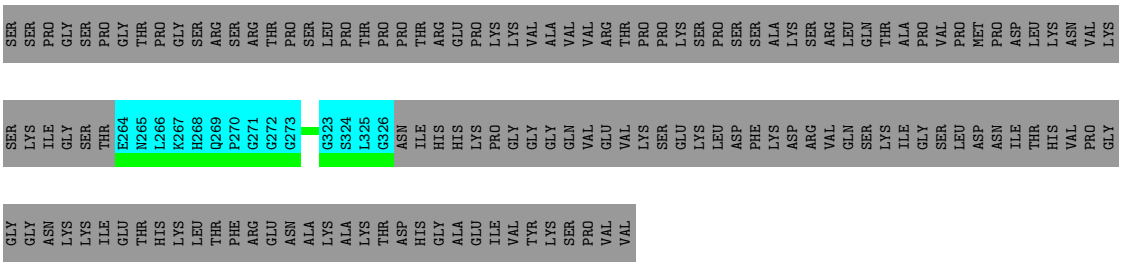
4 Residue-property plots

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.

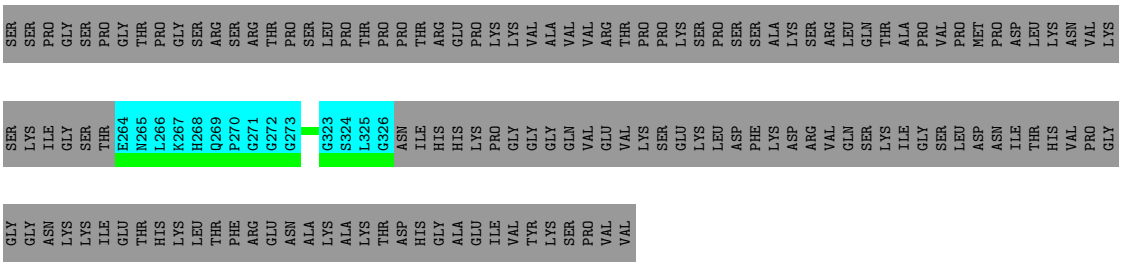
• Molecule 1: Microtubule-associated protein tau

Chain A:



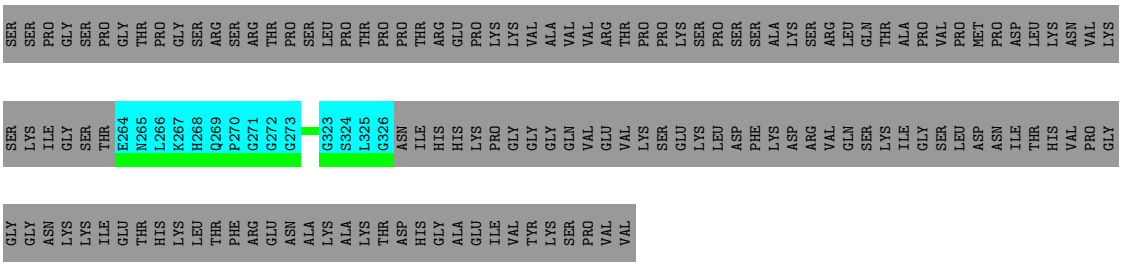
• Molecule 1: Microtubule-associated protein tau

Chain B:



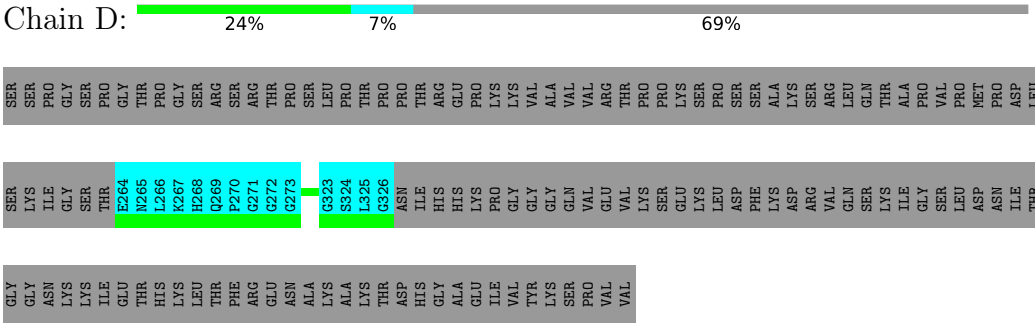
• Molecule 1: Microtubule-associated protein tau

Chain C:



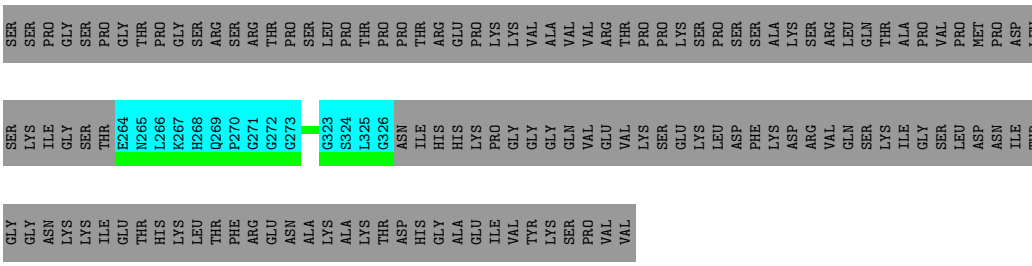
• Molecule 1: Microtubule-associated protein tau

Chain D:



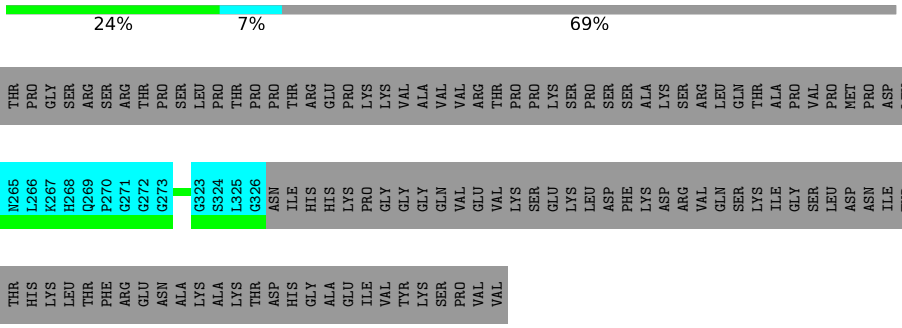
• Molecule 1: Microtubule-associated protein tau

Chain E:



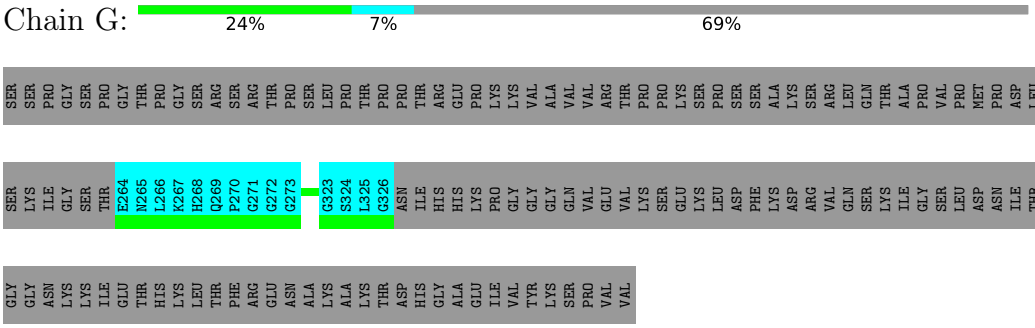
• Molecule 1: Microtubule-associated protein tau

Chain F:



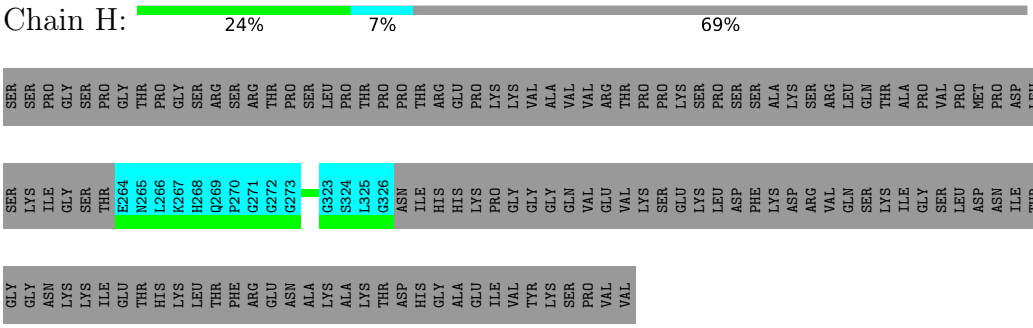
• Molecule 1: Microtubule-associated protein tau

Chain G:



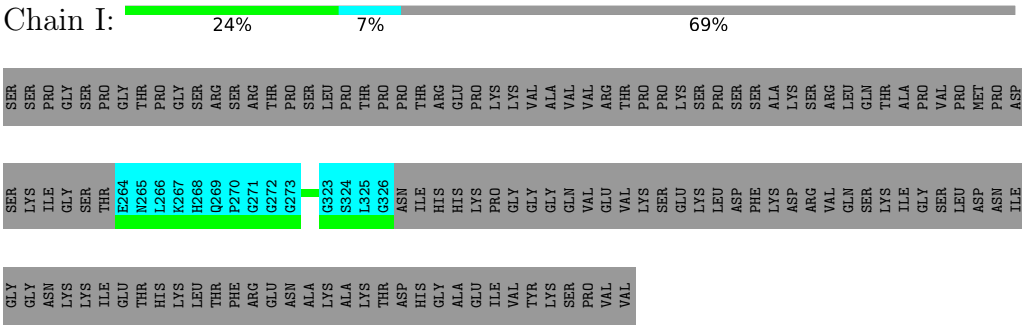
• Molecule 1: Microtubule-associated protein tau

Chain H:



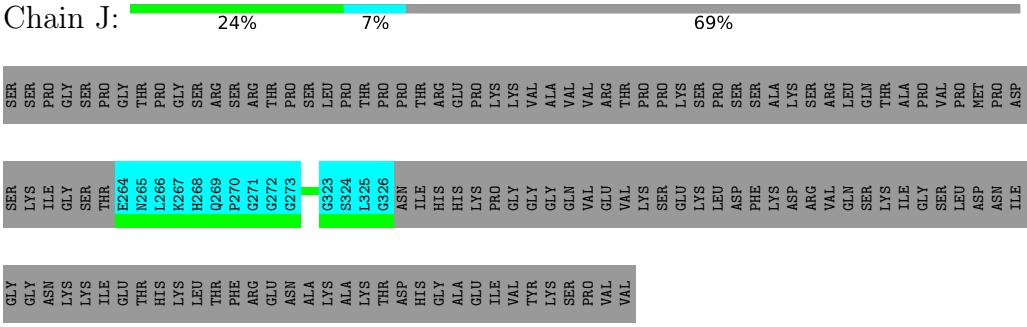
• Molecule 1: Microtubule-associated protein tau

Chain I:



• Molecule 1: Microtubule-associated protein tau

Chain J:



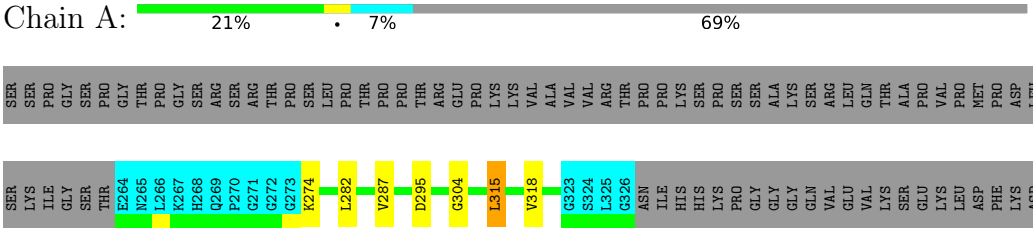
• Molecule 1: Microtubule-associated protein tau

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

The representative model is number 4. Colouring as in section 4.1 above.

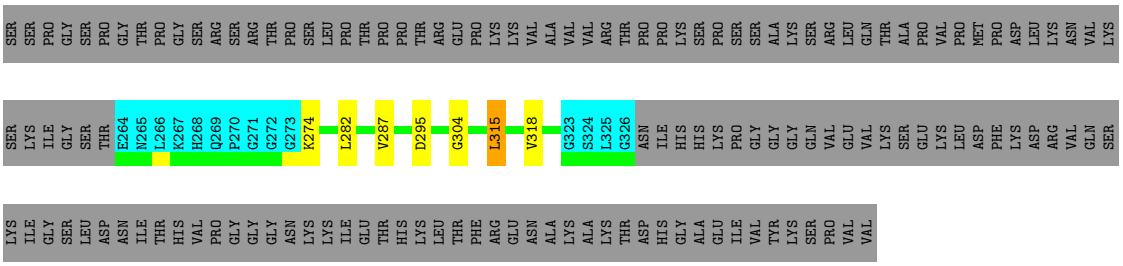
• Molecule 1: Microtubule-associated protein tau

Chain A:



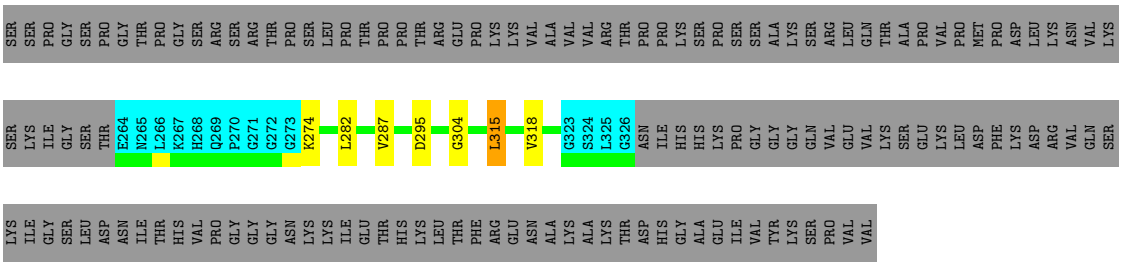
• Molecule 1: Microtubule-associated protein tau

Chain B:



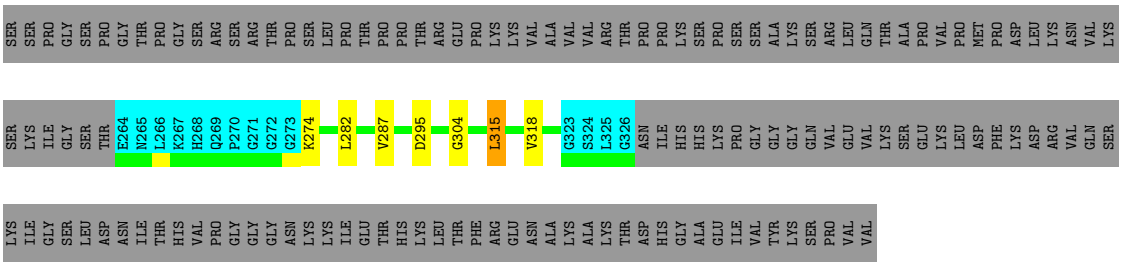
• Molecule 1: Microtubule-associated protein tau

Chain C:



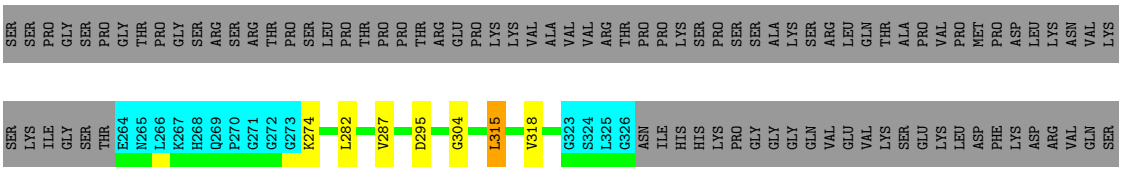
• Molecule 1: Microtubule-associated protein tau

Chain D:



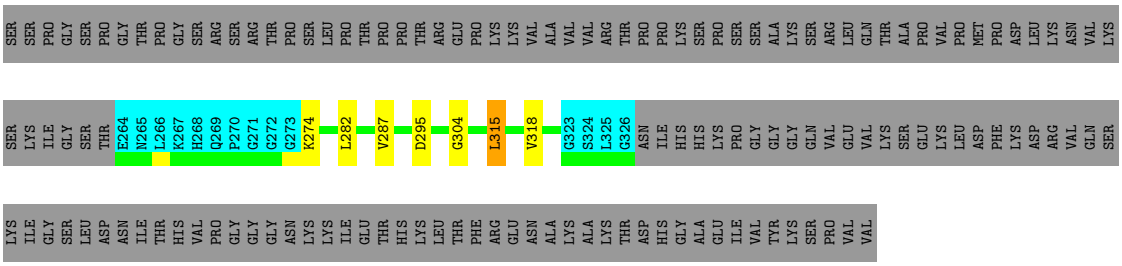
• Molecule 1: Microtubule-associated protein tau

Chain E:



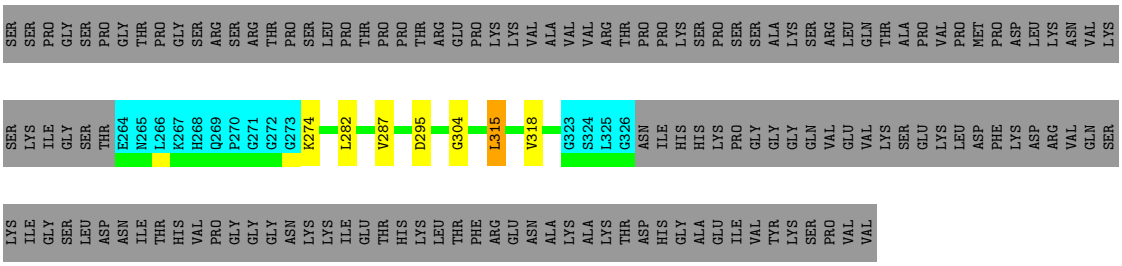
• Molecule 1: Microtubule-associated protein tau

Chain F:



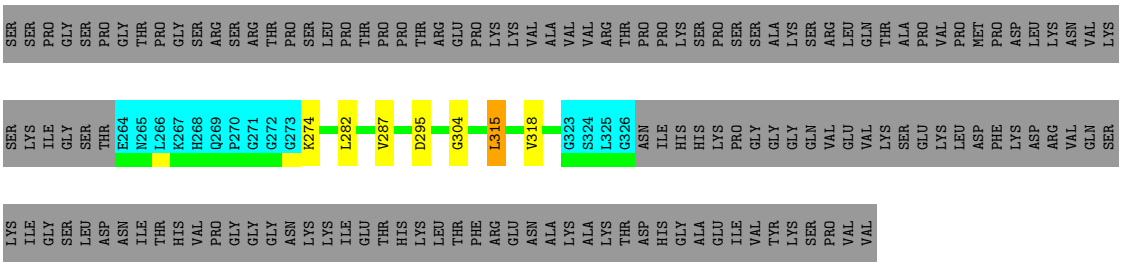
• Molecule 1: Microtubule-associated protein tau

Chain G:



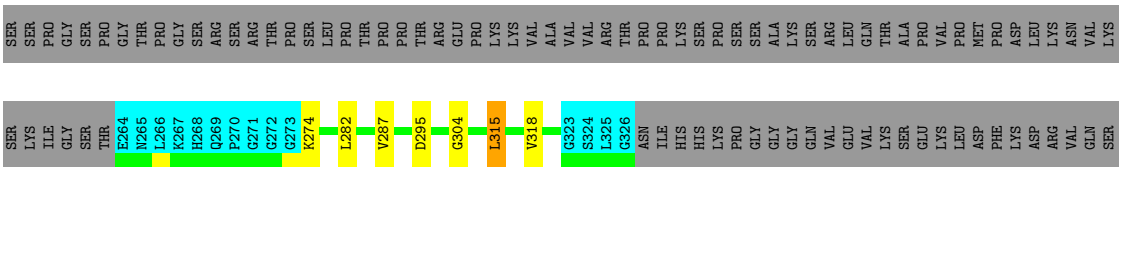
• Molecule 1: Microtubule-associated protein tau

Chain H:



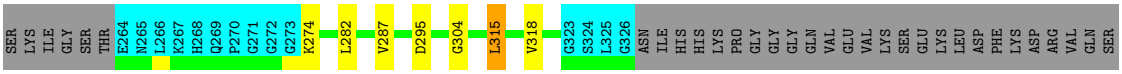
• Molecule 1: Microtubule-associated protein tau

Chain I:



• Molecule 1: Microtubule-associated protein tau

Chain J:



5 Refinement protocol and experimental data overview

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

Of the 1000 calculated structures, 10 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	329
Number of shifts mapped to atoms	227
Number of unparsed shifts	0
Number of shifts with mapping errors	102
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	3%

Note: This is a solid-state NMR structure, where hydrogen atoms are typically not assigned a chemical shift value, which may lead to lower completeness of assignment measure.

6 Model quality

6.1 Standard geometry

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

6.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	368	401	400	4±2
1	B	368	401	400	4±2
1	C	368	401	400	6±3
1	D	368	401	400	6±3
1	E	368	401	400	6±3
1	F	368	401	400	6±3
1	G	368	401	400	6±3
1	H	368	401	400	6±3
1	I	368	401	400	4±2
1	J	368	401	400	4±2
All	All	36800	40100	40000	343

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.

5 of 239 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:G:300:VAL:HG12	1:I:300:VAL:HG22	0.76	1.58	6	4
1:A:300:VAL:HG12	1:C:300:VAL:HG22	0.76	1.58	6	4
1:H:300:VAL:HG12	1:J:300:VAL:HG22	0.76	1.58	6	4
1:E:300:VAL:HG12	1:G:300:VAL:HG22	0.76	1.58	6	4
1:B:300:VAL:HG12	1:D:300:VAL:HG22	0.76	1.58	6	4

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	49/202 (24%)	45±1 (92±2%)	3±1 (6±2%)	1±1 (1±1%)	15	61
1	B	49/202 (24%)	45±1 (92±2%)	3±1 (6±2%)	1±1 (1±1%)	15	61
1	C	49/202 (24%)	45±1 (92±2%)	3±1 (6±2%)	1±1 (1±1%)	15	61
1	D	49/202 (24%)	45±1 (92±2%)	3±1 (6±2%)	1±1 (1±1%)	15	61
1	E	49/202 (24%)	45±1 (92±2%)	3±1 (6±2%)	1±1 (1±1%)	15	61
1	F	49/202 (24%)	45±1 (92±2%)	3±1 (6±2%)	1±1 (1±1%)	15	61
1	G	49/202 (24%)	45±1 (92±2%)	3±1 (6±2%)	1±1 (1±1%)	15	61
1	H	49/202 (24%)	45±1 (92±2%)	3±1 (6±2%)	1±1 (1±1%)	15	61
1	I	49/202 (24%)	45±1 (92±2%)	3±1 (6±2%)	1±1 (1±1%)	15	61
1	J	49/202 (24%)	45±1 (92±2%)	3±1 (6±2%)	1±1 (1±1%)	15	61
All	All	4900/20200 (24%)	4530 (92%)	300 (6%)	70 (1%)	15	61

5 of 50 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	274	LYS	2
1	B	274	LYS	2
1	C	274	LYS	2
1	D	274	LYS	2
1	E	274	LYS	2

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	45/175 (26%)	43±1 (95±1%)	2±1 (5±1%)	26	75
1	B	45/175 (26%)	43±1 (95±1%)	2±1 (5±1%)	26	75
1	C	45/175 (26%)	43±1 (95±1%)	2±1 (5±1%)	26	75
1	D	45/175 (26%)	43±1 (95±2%)	2±1 (5±2%)	27	77
1	E	45/175 (26%)	43±1 (95±2%)	2±1 (5±2%)	27	77
1	F	45/175 (26%)	43±1 (95±2%)	2±1 (5±2%)	27	77
1	G	45/175 (26%)	43±1 (95±2%)	2±1 (5±2%)	27	77
1	H	45/175 (26%)	43±1 (95±2%)	2±1 (5±2%)	27	77
1	I	45/175 (26%)	43±1 (95±2%)	2±1 (5±2%)	27	77
1	J	45/175 (26%)	43±1 (95±2%)	2±1 (5±2%)	27	77
All	All	4500/17500 (26%)	4267 (95%)	233 (5%)	27	76

5 of 133 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	288	GLN	4
1	B	288	GLN	4
1	C	288	GLN	4
1	D	288	GLN	4
1	E	288	GLN	4

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

The completeness of assignment taking into account all chemical shift lists is 3% for the well-defined parts and 3% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *starch_ output*

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	329
Number of shifts mapped to atoms	227
Number of unparsed shifts	0
Number of shifts with mapping errors	102
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 102) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	329	HIS	C	171.312	.	.
1	A	329	HIS	CA	53.017	.	.
1	A	329	HIS	CB	33.177	.	.
1	A	329	HIS	CD2	114.657	.	.
1	A	329	HIS	CE1	134.415	.	.
1	A	329	HIS	N	126.468	.	.
1	A	330	HIS	C	170.615	.	.
1	A	330	HIS	CA	54.85	.	.
1	A	330	HIS	CB	29.771	.	.
1	A	330	HIS	CD2	114.139	.	.
1	A	330	HIS	CE1	132.203	.	.
1	A	330	HIS	N	126.386	.	.
1	A	331	LYS	C	173.087	.	.
1	A	331	LYS	CA	52.602	.	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	331	LYS	CB	31.039	.	.
1	A	331	LYS	CD	28.661	.	.
1	A	331	LYS	CG	23.905	.	.
1	A	331	LYS	N	126.423	.	.
1	A	355	GLY	C	172.8	.	.
1	A	355	GLY	CA	46.259	.	.
1	A	355	GLY	N	119.144	.	.
1	A	356	SER	C	169.267	.	.
1	A	356	SER	CA	56.825	.	.
1	A	356	SER	CB	66.356	.	.
1	A	356	SER	N	111.365	.	.
1	A	361	THR	C	171.314	.	.
1	A	361	THR	CA	54.681	.	.
1	A	361	THR	CB	64.461	.	.
1	A	361	THR	N	117.933	.	.
1	A	362	HIS	C	172.025	.	.
1	A	362	HIS	CA	53.107	.	.
1	A	362	HIS	CB	33.183	.	.
1	A	362	HIS	CD2	114.554	.	.
1	A	362	HIS	CE1	135.843	.	.
1	A	362	HIS	CG	134.97	.	.
1	A	362	HIS	N	123.797	.	.
1	A	367	GLY	C	172.227	.	.
1	A	367	GLY	CA	45.757	.	.
1	A	367	GLY	N	106.945	.	.
1	A	368	ASN	CA	51.188	.	.
1	A	368	ASN	CB	40.099	.	.
1	A	368	ASN	N	117.029	.	.
1	A	372	GLU	C	174.186	.	.
1	A	372	GLU	CA	51.814	.	.
1	A	372	GLU	CB	29.264	.	.
1	A	372	GLU	CD	181.675	.	.
1	A	372	GLU	CG	35.812	.	.
1	A	372	GLU	N	123.093	.	.
1	A	373	THR	C	174.257	.	.
1	A	373	THR	CA	59.601	.	.
1	A	373	THR	CB	66.186	.	.
1	A	373	THR	CG2	21.498	.	.
1	A	373	THR	N	113.907	.	.
1	A	374	HIS	C	171.584	.	.
1	A	374	HIS	CA	53.443	.	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	374	HIS	CB	34.63	.	.
1	A	374	HIS	CD2	113.796	.	.
1	A	374	HIS	CE1	134.765	.	.
1	A	374	HIS	N	125.721	.	.
1	A	375	LYS	CA	52.269	.	.
1	A	375	LYS	CB	29.148	.	.
1	A	375	LYS	CD	25.333	.	.
1	A	375	LYS	CE	44.057	.	.
1	A	375	LYS	CG	21.855	.	.
1	A	375	LYS	N	126.043	.	.
1	A	382	ALA	C	173.006	.	.
1	A	382	ALA	CA	48.654	.	.
1	A	382	ALA	CB	20.786	.	.
1	A	382	ALA	N	128.749	.	.
1	A	383	LYS	C	172.364	.	.
1	A	383	LYS	CA	53.08	.	.
1	A	383	LYS	CB	32.179	.	.
1	A	383	LYS	N	118.358	.	.
1	A	384	ALA	C	173.198	.	.
1	A	384	ALA	CA	48.714	.	.
1	A	384	ALA	CB	20.673	.	.
1	A	384	ALA	N	127.074	.	.
1	A	385	LYS	C	174.91	.	.
1	A	385	LYS	CA	53.63	.	.
1	A	385	LYS	CB	31.227	.	.
1	A	385	LYS	N	119.634	.	.
1	A	386	THR	C	172.605	.	.
1	A	386	THR	CA	57.703	.	.
1	A	386	THR	CB	67.055	.	.
1	A	386	THR	CG2	19.744	.	.
1	A	386	THR	N	108.507	.	.
1	A	388	HIS	CA	50.481	.	.
1	A	388	HIS	CB	29.717	.	.
1	A	388	HIS	CE1	135.106	.	.
1	A	388	HIS	CG	127.2	.	.
1	A	388	HIS	N	128.229	.	.
1	A	389	GLY	C	170.969	.	.
1	A	389	GLY	CA	44.943	.	.
1	A	389	GLY	N	108.718	.	.
1	A	390	ALA	C	174.544	.	.
1	A	390	ALA	CA	49.537	.	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	390	ALA	CB	18.01	.	.
1	A	390	ALA	N	121.808	.	.
1	A	391	GLU	CA	53.637	.	.
1	A	391	GLU	CB	31.744	.	.
1	A	391	GLU	CG	34.681	.	.
1	A	391	GLU	N	121.829	.	.

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	68	2.08 ± 0.34	Should be checked
$^{13}C_{\beta}$	56	0.91 ± 0.33	Should be checked
$^{13}C'$	59	2.30 ± 0.35	Should be applied
^{15}N	67	-1.99 ± 0.43	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 3%, i.e. 220 atoms were assigned a chemical shift out of a possible 6570. 0 out of 100 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	126/2450 (5%)	0/1000 (0%)	85/980 (9%)	41/470 (9%)
Sidechain	92/3950 (2%)	0/2560 (0%)	89/1240 (7%)	3/150 (2%)
Aromatic	2/170 (1%)	0/80 (0%)	2/70 (3%)	0/20 (0%)
Overall	220/6570 (3%)	0/3640 (0%)	176/2290 (8%)	44/640 (7%)

Note: This is a solid-state NMR structure, where hydrogen atoms are typically not assigned a chemical shift value, which may lead to lower completeness of assignment measure.

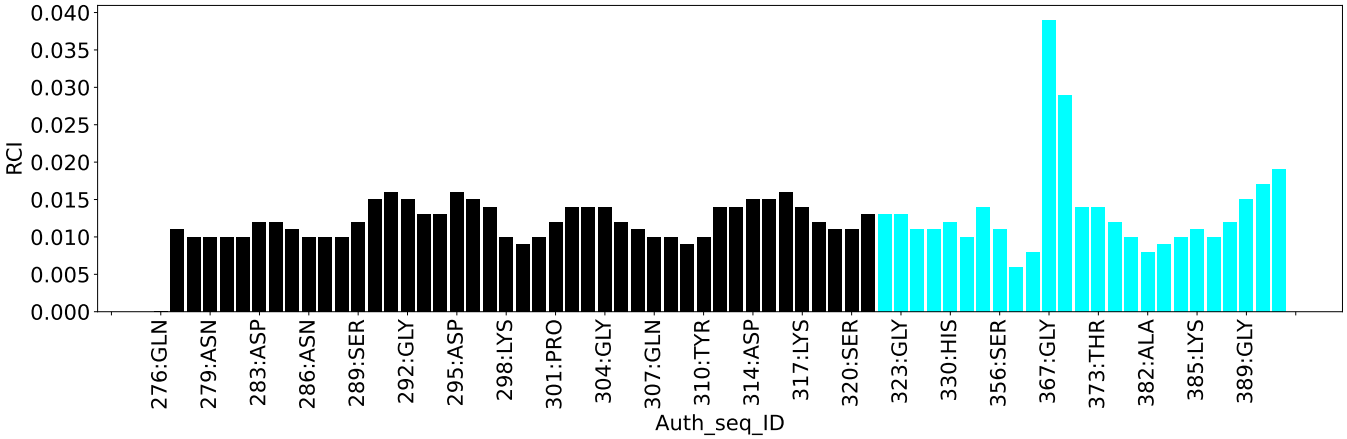
7.1.4 Statistically unusual chemical shifts

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

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	46
Intra-residue (|i-j|=0)	0
Sequential (|i-j|=1)	0
Medium range (|i-j|>1 and |i-j|<5)	0
Long range (|i-j|≥5)	46
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	0.0
Number of long range restraints per residue[1]	0.0

[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)	1.0	0.2
0.2-0.5 (Medium)	1.4	0.46
>0.5 (Large)	11.6	4.34

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)

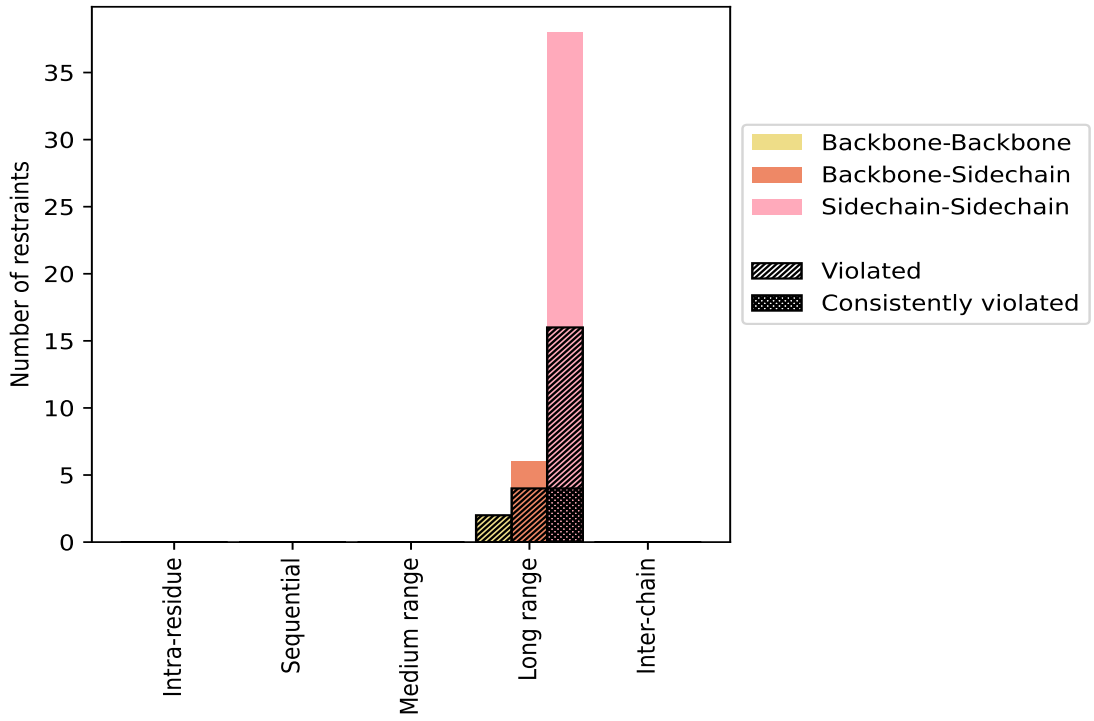
9.1 Summary of distance violations (i)

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)	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
Sequential (|i-j|=1)	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
Medium range (|i-j|>1 & |i-j|<5)	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
Long range (|i-j|≥5)	46	100.0	22	47.8	47.8	4	8.7	8.7
Backbone-Backbone	2	4.3	2	100.0	4.3	0	0.0	0.0
Backbone-Sidechain	6	13.0	4	66.7	8.7	0	0.0	0.0
Sidechain-Sidechain	38	82.6	16	42.1	34.8	4	10.5	8.7
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	46	100.0	22	47.8	47.8	4	8.7	8.7
Backbone-Backbone	2	4.3	2	100.0	4.3	0	0.0	0.0
Backbone-Sidechain	6	13.0	4	66.7	8.7	0	0.0	0.0
Sidechain-Sidechain	38	82.6	16	42.1	34.8	4	10.5	8.7

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

9.1.1 Bar chart : Distribution of distance restraints and violations



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

9.2 Distance violation statistics for each model

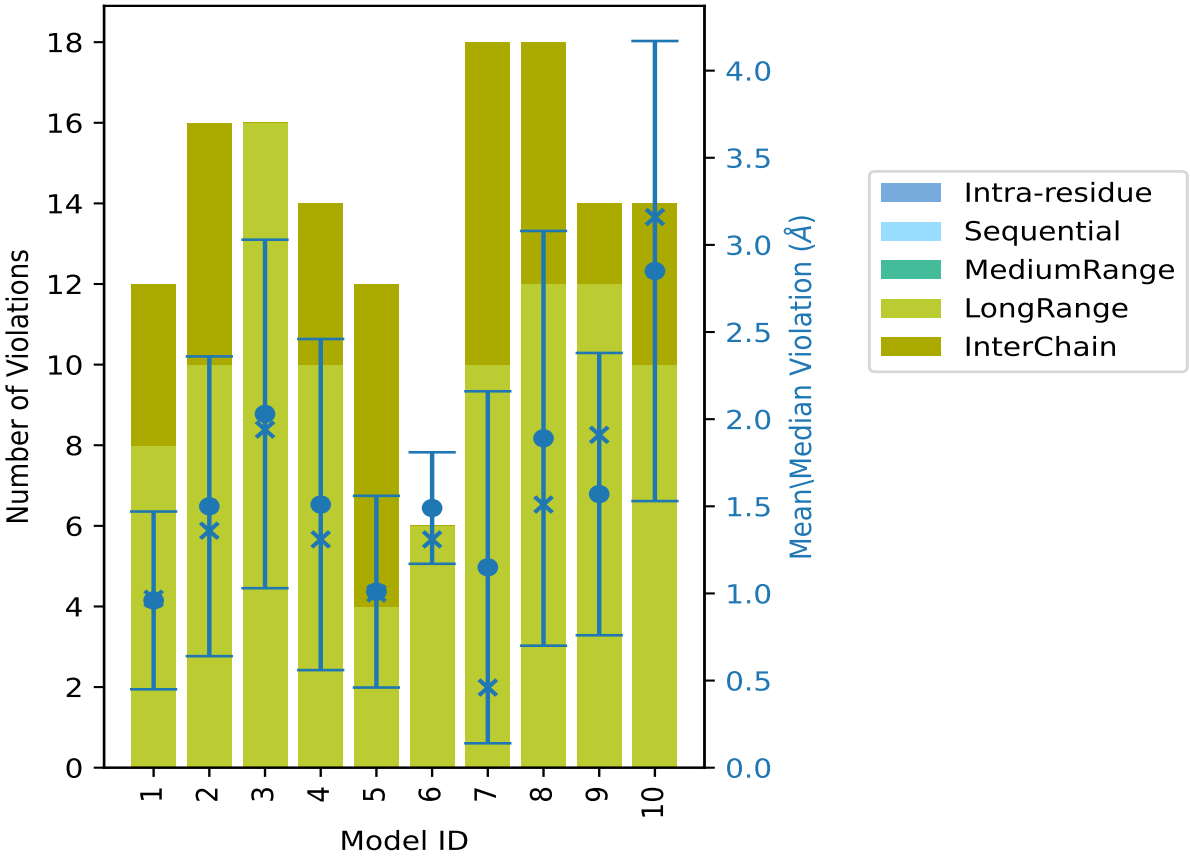
The following table provides the distance violation statistics for each model in the ensemble. Violations less than 0.1 Å are not included in the statistics.

Model ID	Number of violations						Mean (Å)	Max (Å)	SD[6] (Å)	Median (Å)
	IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total				
1	0	0	0	8	4	12	0.96	1.86	0.51	0.97
2	0	0	0	10	6	16	1.5	3.17	0.86	1.36
3	0	0	0	16	0	16	2.03	3.58	1.0	1.94
4	0	0	0	10	4	14	1.51	3.2	0.95	1.31
5	0	0	0	4	8	12	1.01	2.03	0.55	1.0
6	0	0	0	6	0	6	1.49	1.94	0.32	1.31
7	0	0	0	10	8	18	1.15	3.11	1.01	0.46
8	0	0	0	12	6	18	1.89	4.09	1.19	1.51
9	0	0	0	12	2	14	1.57	2.93	0.81	1.91
10	0	0	0	10	4	14	2.85	4.34	1.32	3.16

[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, 24(IR:0, SQ:0, MR:0, LR:24, 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	0	0	2	0	2	1	10.0
0	0	0	2	0	2	2	20.0
0	0	0	0	0	0	3	30.0
0	0	0	0	0	0	4	40.0

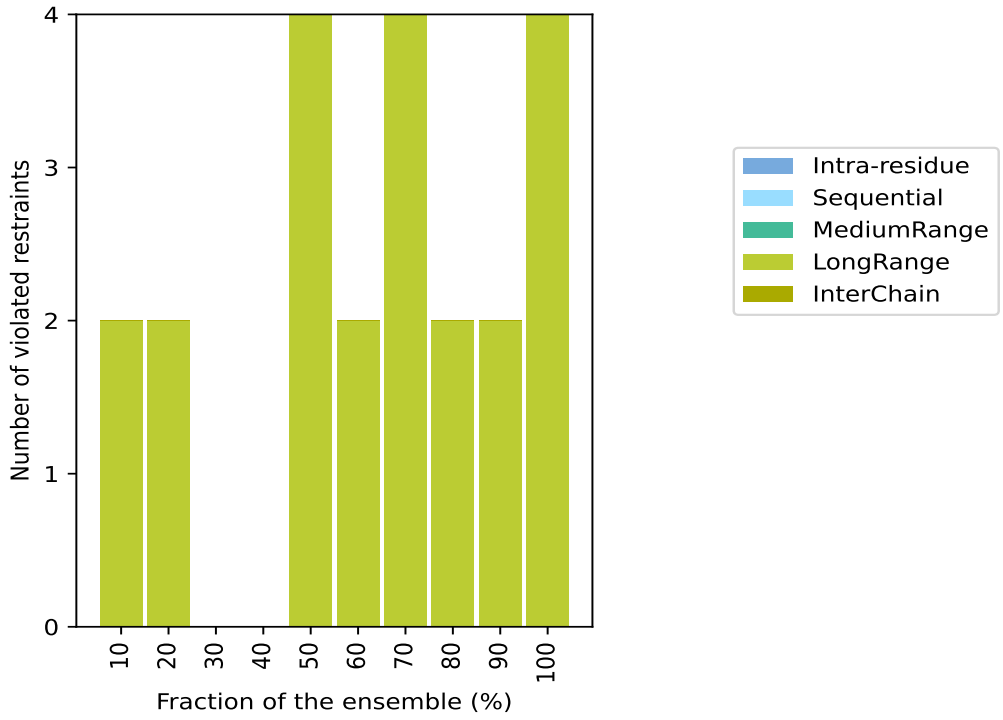
Continued on next page...

Continued from previous page...

Number of violated restraints						Fraction of the ensemble	
IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total	Count[6]	%
0	0	0	4	0	4	5	50.0
0	0	0	2	0	2	6	60.0
0	0	0	4	0	4	7	70.0
0	0	0	2	0	2	8	80.0
0	0	0	2	0	2	9	90.0
0	0	0	4	0	4	10	100.0

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

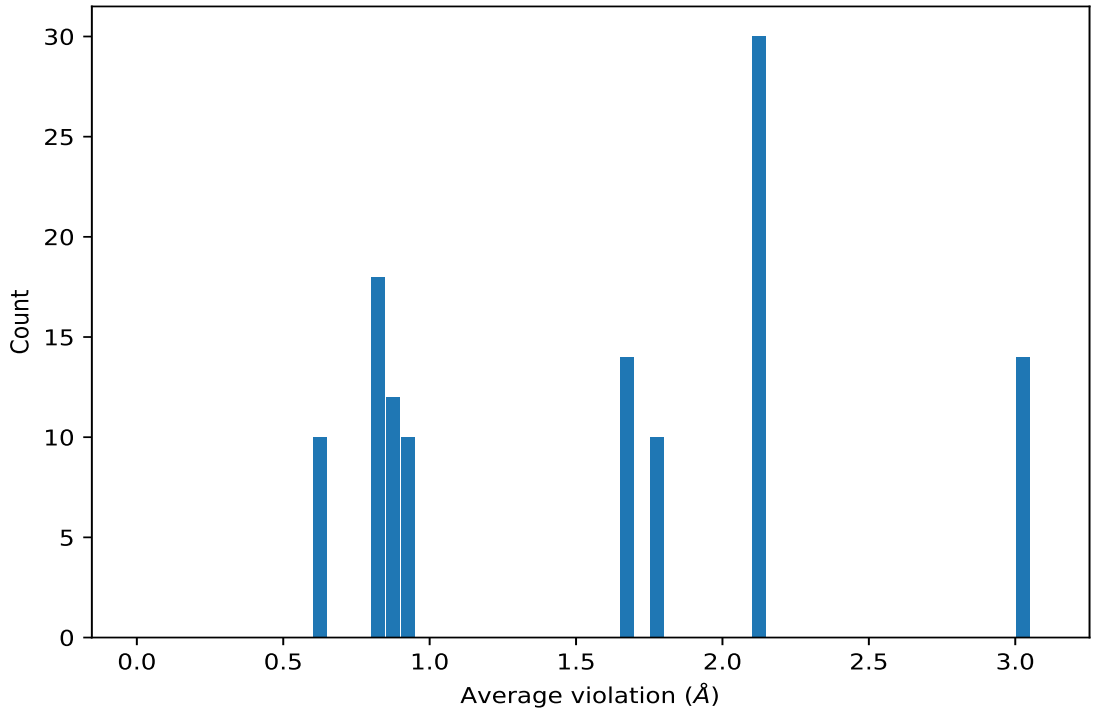
9.3.1 Bar graph : Distance violation statistics for the ensemble



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 (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)
(1,17)	1:297:I:ILE:CB	1:308:I:ILE:CD1	10	3.02	0.82	3.14
(1,17)	1:297:J:ILE:CB	1:308:J:ILE:CD1	10	3.02	0.82	3.14
(1,17)	1:297:G:ILE:CB	1:308:G:ILE:CD1	10	3.02	0.82	3.14
(1,17)	1:297:G:ILE:CB	1:308:I:ILE:CD1	10	3.02	0.82	3.14
(1,17)	1:297:E:ILE:CB	1:308:E:ILE:CD1	10	3.02	0.82	3.14
(1,17)	1:297:D:ILE:CB	1:308:D:ILE:CD1	10	3.02	0.82	3.14
(1,17)	1:297:H:ILE:CB	1:308:H:ILE:CD1	10	3.02	0.82	3.14
(2,17)	1:297:I:ILE:CB	1:308:I:ILE:CD1	10	3.02	0.82	3.14
(2,17)	1:297:J:ILE:CB	1:308:J:ILE:CD1	10	3.02	0.82	3.14
(2,17)	1:297:G:ILE:CB	1:308:G:ILE:CD1	10	3.02	0.82	3.14
(2,17)	1:297:G:ILE:CB	1:308:I:ILE:CD1	10	3.02	0.82	3.14
(2,17)	1:297:E:ILE:CB	1:308:E:ILE:CD1	10	3.02	0.82	3.14
(2,17)	1:297:D:ILE:CB	1:308:D:ILE:CD1	10	3.02	0.82	3.14
(2,17)	1:297:H:ILE:CB	1:308:H:ILE:CD1	10	3.02	0.82	3.14
(1,16)	1:297:J:ILE:CG1	1:308:J:ILE:CD1	10	2.12	0.75	2.16
(1,16)	1:297:J:ILE:CG2	1:308:J:ILE:CD1	10	2.12	0.75	2.16

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,16)	1:297:C:ILE:CG1	1:308:C:ILE:CD1	10	2.12	0.75	2.16
(1,16)	1:297:A:ILE:CG1	1:308:A:ILE:CD1	10	2.12	0.75	2.16
(1,16)	1:297:E:ILE:CG2	1:308:E:ILE:CD1	10	2.12	0.75	2.16
(1,16)	1:297:D:ILE:CG1	1:308:D:ILE:CD1	10	2.12	0.75	2.16
(1,16)	1:297:C:ILE:CG2	1:308:C:ILE:CD1	10	2.12	0.75	2.16
(1,16)	1:297:H:ILE:CG2	1:308:F:ILE:CD1	10	2.12	0.75	2.16
(1,16)	1:297:I:ILE:CG2	1:308:I:ILE:CD1	10	2.12	0.75	2.16
(2,16)	1:297:J:ILE:CG1	1:308:J:ILE:CD1	10	2.12	0.75	2.16
(2,16)	1:297:J:ILE:CG2	1:308:J:ILE:CD1	10	2.12	0.75	2.16
(2,16)	1:297:C:ILE:CG1	1:308:C:ILE:CD1	10	2.12	0.75	2.16
(2,16)	1:297:A:ILE:CG1	1:308:A:ILE:CD1	10	2.12	0.75	2.16
(2,16)	1:297:E:ILE:CG2	1:308:E:ILE:CD1	10	2.12	0.75	2.16
(2,16)	1:297:D:ILE:CG1	1:308:D:ILE:CD1	10	2.12	0.75	2.16
(2,16)	1:297:C:ILE:CG2	1:308:C:ILE:CD1	10	2.12	0.75	2.16
(2,16)	1:297:H:ILE:CG2	1:308:F:ILE:CD1	10	2.12	0.75	2.16
(2,16)	1:297:I:ILE:CG2	1:308:I:ILE:CD1	10	2.12	0.75	2.16
(1,2)	1:310:B:TYR:CB	1:297:B:ILE:CD1	9	0.82	0.39	0.78
(1,2)	1:310:A:TYR:CB	1:297:A:ILE:CD1	9	0.82	0.39	0.78
(1,2)	1:310:J:TYR:CB	1:297:J:ILE:CD1	9	0.82	0.39	0.78
(1,2)	1:310:I:TYR:CB	1:297:G:ILE:CD1	9	0.82	0.39	0.78
(1,2)	1:310:F:TYR:CB	1:297:F:ILE:CD1	9	0.82	0.39	0.78
(1,2)	1:310:D:TYR:CB	1:297:D:ILE:CD1	9	0.82	0.39	0.78
(1,2)	1:310:C:TYR:CB	1:297:C:ILE:CD1	9	0.82	0.39	0.78
(2,2)	1:310:B:TYR:CB	1:297:B:ILE:CD1	9	0.82	0.39	0.78
(2,2)	1:310:A:TYR:CB	1:297:A:ILE:CD1	9	0.82	0.39	0.78
(2,2)	1:310:J:TYR:CB	1:297:J:ILE:CD1	9	0.82	0.39	0.78
(2,2)	1:310:I:TYR:CB	1:297:G:ILE:CD1	9	0.82	0.39	0.78
(2,2)	1:310:F:TYR:CB	1:297:F:ILE:CD1	9	0.82	0.39	0.78
(2,2)	1:310:D:TYR:CB	1:297:D:ILE:CD1	9	0.82	0.39	0.78
(2,2)	1:310:C:TYR:CB	1:297:C:ILE:CD1	9	0.82	0.39	0.78
(1,5)	1:297:J:ILE:CD1	1:308:J:ILE:CD1	8	1.7	0.6	1.7
(1,5)	1:297:F:ILE:CD1	1:308:F:ILE:CD1	8	1.7	0.6	1.7
(1,5)	1:297:A:ILE:CD1	1:308:A:ILE:CD1	8	1.7	0.6	1.7
(1,5)	1:297:G:ILE:CD1	1:308:I:ILE:CD1	8	1.7	0.6	1.7
(1,5)	1:297:C:ILE:CD1	1:308:C:ILE:CD1	8	1.7	0.6	1.7
(1,5)	1:297:I:ILE:CG2	1:308:I:ILE:CD1	8	1.7	0.6	1.7
(1,5)	1:297:H:ILE:CD1	1:308:H:ILE:CD1	8	1.7	0.6	1.7
(2,5)	1:297:J:ILE:CD1	1:308:J:ILE:CD1	8	1.7	0.6	1.7
(2,5)	1:297:F:ILE:CD1	1:308:F:ILE:CD1	8	1.7	0.6	1.7
(2,5)	1:297:A:ILE:CD1	1:308:A:ILE:CD1	8	1.7	0.6	1.7
(2,5)	1:297:G:ILE:CD1	1:308:I:ILE:CD1	8	1.7	0.6	1.7
(2,5)	1:297:C:ILE:CD1	1:308:C:ILE:CD1	8	1.7	0.6	1.7

Continued on next page...

Continued from previous page...

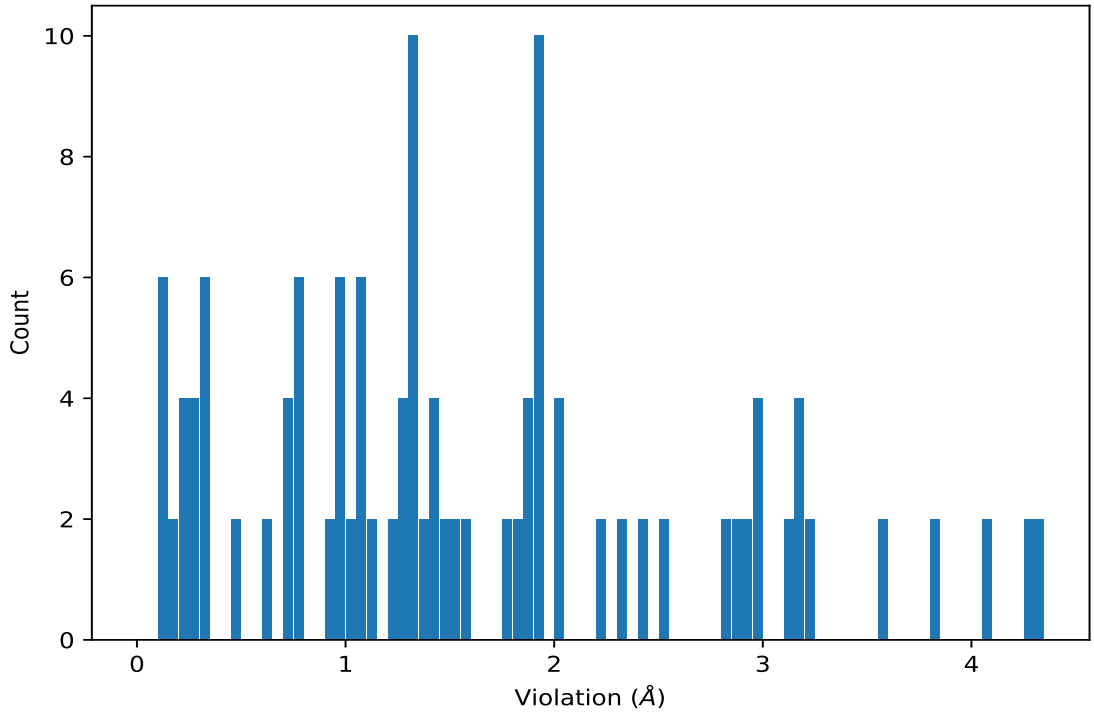
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(2,5)	1:297:I:ILE:CG2	1:308:I:ILE:CD1	8	1.7	0.6	1.7
(2,5)	1:297:H:ILE:CD1	1:308:H:ILE:CD1	8	1.7	0.6	1.7
(1,19)	1:283:C:ASP:CB	1:316:A:SER:CB	7	1.79	1.05	1.34
(1,19)	1:283:A:ASP:CB	1:316:A:SER:CB	7	1.79	1.05	1.34
(1,19)	1:283:D:ASP:CB	1:316:D:SER:CB	7	1.79	1.05	1.34
(1,19)	1:283:D:ASP:CB	1:316:B:SER:CB	7	1.79	1.05	1.34
(1,19)	1:283:I:ASP:CB	1:316:I:SER:CB	7	1.79	1.05	1.34
(2,19)	1:283:C:ASP:CB	1:316:A:SER:CB	7	1.79	1.05	1.34

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

9.5 All violated distance restraints (i)

9.5.1 Histogram : Distribution of distance violations (i)

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



9.5.2 Table : All distance violations (i)

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

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(2,17)	1:297:H:ILE:CB	1:308:H:ILE:CD1	10	4.34
(1,17)	1:297:H:ILE:CB	1:308:H:ILE:CD1	10	4.34
(2,18)	1:282:C:LEU:CG	1:316:A:SER:CB	10	4.26
(1,18)	1:282:C:LEU:CG	1:316:A:SER:CB	10	4.26
(2,17)	1:297:E:ILE:CB	1:308:E:ILE:CD1	8	4.09
(1,17)	1:297:E:ILE:CB	1:308:E:ILE:CD1	8	4.09
(2,19)	1:283:C:ASP:CB	1:316:A:SER:CB	10	3.84
(1,19)	1:283:C:ASP:CB	1:316:A:SER:CB	10	3.84
(2,17)	1:297:G:ILE:CB	1:308:G:ILE:CD1	3	3.58
(1,17)	1:297:G:ILE:CB	1:308:G:ILE:CD1	3	3.58

10 Dihedral-angle violation analysis (i)

No dihedral-angle restraints found