WORLDWIDE PDB PROTEIN DATA BANK

wwPDB NMR Structure Validation Summary Report (i)

Jun 6, 2023 – 05:54 pm BST

PDB ID	:	6TI6
BMRB ID	:	34455
Title	:	Mixing Abeta(1-40) and Abeta(1-42) peptides generates unique amyloid fibrils
Authors	:	Cerofolini, L.; Ravera, E.; Bologna, S.; Wiglenda, T.; Boddrich, A.; Purfurst, B.; Benilova, A.; Korsak, M.; Gallo, G.; Rizzo, D.; Gonnelli, L.; Fragai, M.; De Strooper, B.; Wanker, E.E.; Luchinat, C.
Deposited on	:	2019-11-21

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:

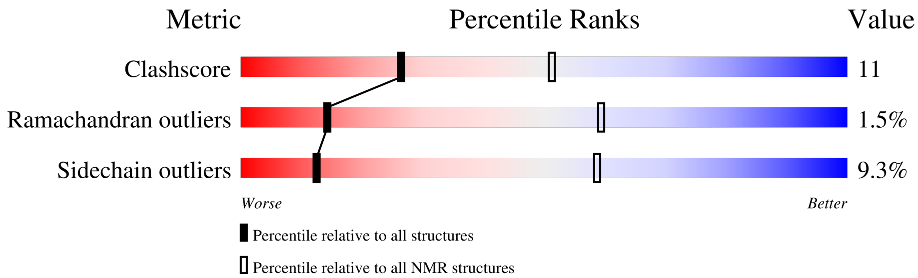
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal 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:
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	40	62% 10% • 25%
1	C	40	62% 12% 25%
1	E	40	50% 22% • 25%
1	G	40	52% 22% 25%
1	I	40	58% 12% 5% 25%
1	K	40	62% 10% • 25%
1	M	40	50% 22% • 25%
1	O	40	62% 12% 25%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	42	52% 14% • 7% 24%
2	D	42	57% 17% • 24%
2	F	42	60% 14% • 24%
2	H	42	55% 12% 5% 5% 24%
2	J	42	50% 14% 5% 7% 24%
2	L	42	57% 17% • 24%
2	N	42	57% 17% • 24%
2	P	42	60% 17% 24%

2 Ensemble composition and analysis

This entry contains 4 models. Model 3 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:12-A:40, B:14-B:42, C:11-C:40, D:11-D:42, E:11-E:40, F:11-F:42, G:11-G:40, H:11-H:40, I:13-I:40, J:14-J:42, K:11-K:40, L:11-L:42, M:11-M:40, N:11-N:42, O:11-O:40, P:11-P:42 (485)	0.32	3

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, 2
2	3, 4

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 7322 atoms, of which 3680 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Amyloid-beta precursor protein.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	30	443	143	222	38	39	1	0
1	C	30	443	143	222	38	39	1	0
1	E	30	443	143	222	38	39	1	0
1	G	30	443	143	222	38	39	1	0
1	I	30	443	143	222	38	39	1	0
1	K	30	443	143	222	38	39	1	0
1	M	30	443	143	222	38	39	1	0
1	O	30	443	143	222	38	39	1	0

- Molecule 2 is a protein called Amyloid-beta precursor protein.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	32	472	152	238	40	41	1	0
2	D	32	472	152	238	40	41	1	0
2	F	32	472	152	238	40	41	1	0
2	H	32	473	152	238	40	42	1	0
2	J	32	472	152	238	40	41	1	0
2	L	32	472	152	238	40	41	1	0
2	N	32	472	152	238	40	41	1	0
2	P	32	473	152	238	40	42	1	0

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: Amyloid-beta precursor protein

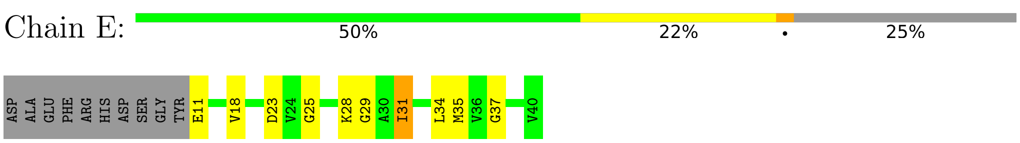
Chain A:

- Molecule 1: Amyloid-beta precursor protein

Chain C:

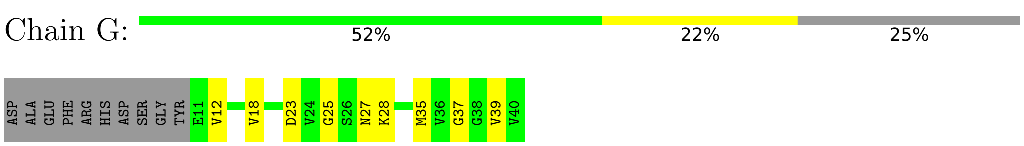
- Molecule 1: Amyloid-beta precursor protein

Chain E:



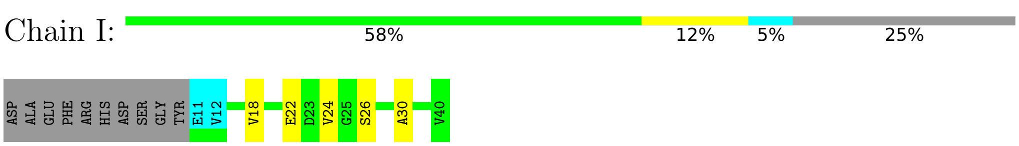
- Molecule 1: Amyloid-beta precursor protein

Chain G:



- Molecule 1: Amyloid-beta precursor protein

Chain I:

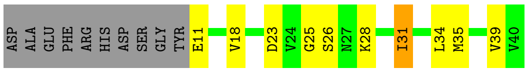


- Molecule 1: Amyloid-beta precursor protein

Chain K:

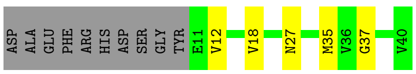
• Molecule 1: Amyloid-beta precursor protein

Chain M:



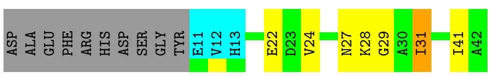
• Molecule 1: Amyloid-beta precursor protein

Chain O:



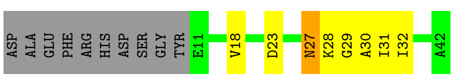
• Molecule 2: Amyloid-beta precursor protein

Chain B:



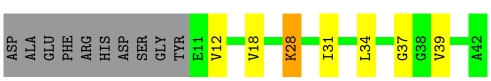
• Molecule 2: Amyloid-beta precursor protein

Chain D:



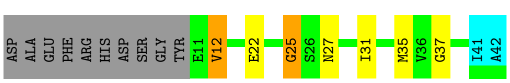
• Molecule 2: Amyloid-beta precursor protein

Chain F:



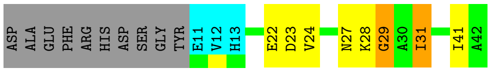
• Molecule 2: Amyloid-beta precursor protein

Chain H:



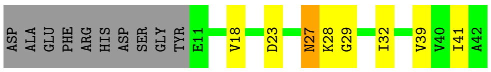
• Molecule 2: Amyloid-beta precursor protein

Chain J:



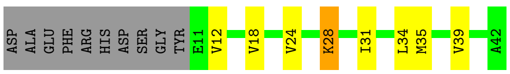
• Molecule 2: Amyloid-beta precursor protein

Chain L:



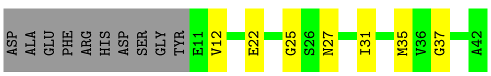
• Molecule 2: Amyloid-beta precursor protein

Chain N:



• Molecule 2: Amyloid-beta precursor protein

Chain P:

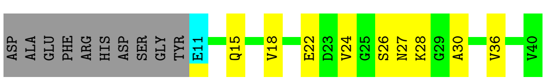


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 3. Colouring as in section 4.1 above.

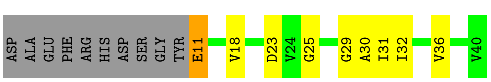
• Molecule 1: Amyloid-beta precursor protein

Chain A:



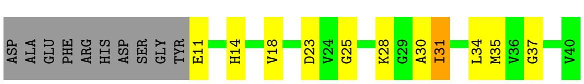
• Molecule 1: Amyloid-beta precursor protein

Chain C:



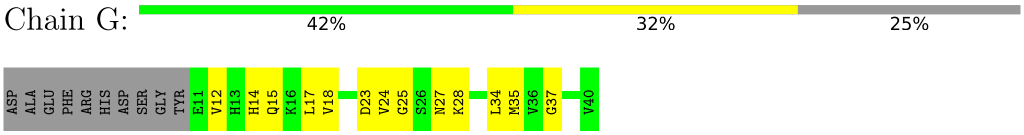
• Molecule 1: Amyloid-beta precursor protein

Chain E:



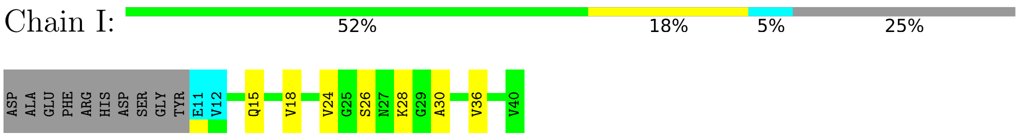
- Molecule 1: Amyloid-beta precursor protein

Chain G:



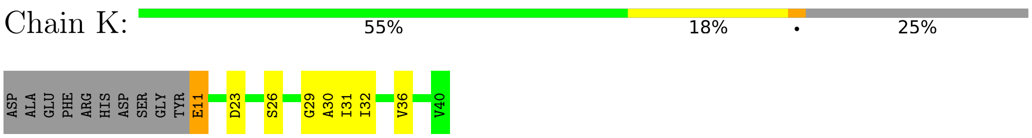
- Molecule 1: Amyloid-beta precursor protein

Chain I:



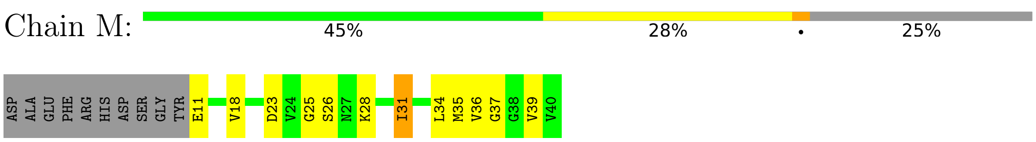
- Molecule 1: Amyloid-beta precursor protein

Chain K:



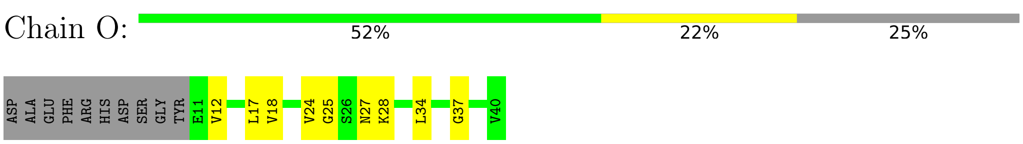
- Molecule 1: Amyloid-beta precursor protein

Chain M:



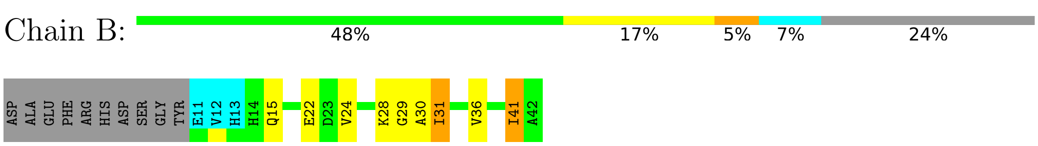
- Molecule 1: Amyloid-beta precursor protein

Chain O:



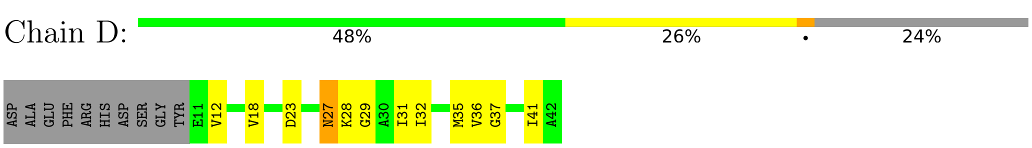
- Molecule 2: Amyloid-beta precursor protein

Chain B:



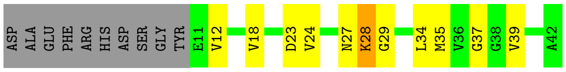
- Molecule 2: Amyloid-beta precursor protein

Chain D:

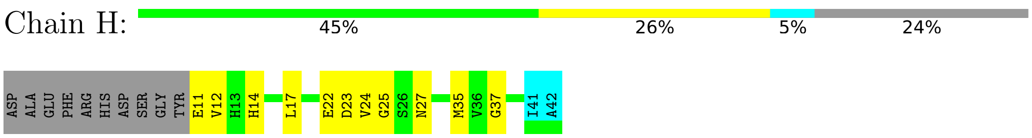


- Molecule 2: Amyloid-beta precursor protein

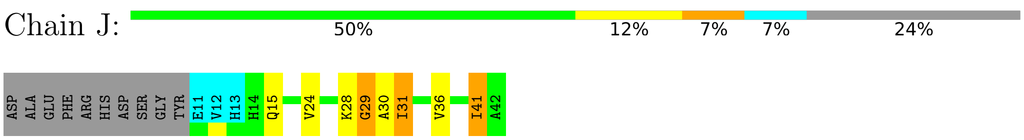
Chain F:



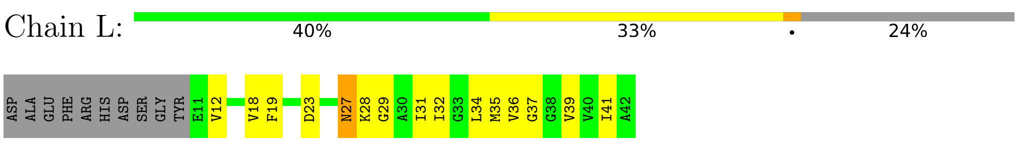
- Molecule 2: Amyloid-beta precursor protein



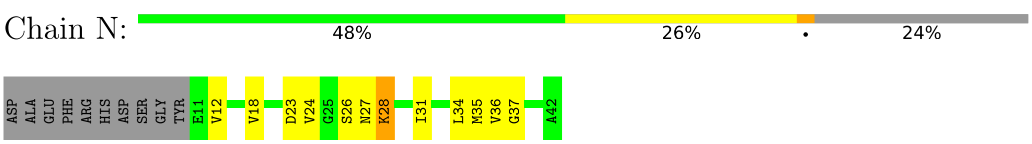
- Molecule 2: Amyloid-beta precursor protein



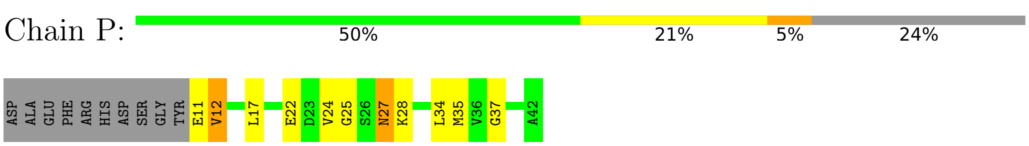
- Molecule 2: Amyloid-beta precursor protein



- Molecule 2: Amyloid-beta precursor protein



- Molecule 2: Amyloid-beta precursor protein



5 Refinement protocol and experimental data overview (i)

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

Of the 200 calculated structures, 4 were deposited, based on the following criterion: *Top 4 Structures From Haddock*.

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

Software name	Classification	Version
HADDOCK	structure calculation	2.2
MODELLER	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	233
Number of shifts mapped to atoms	178
Number of unparsed shifts	0
Number of shifts with mapping errors	55
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	212	217	219	6±1
1	C	221	222	224	8±3
1	E	221	222	224	11±3
1	G	221	222	224	11±1
1	I	205	208	210	7±1
1	K	221	222	224	7±3
1	M	221	222	224	11±1
1	O	221	222	224	10±3
2	B	208	218	219	12±2
2	D	234	238	240	12±2
2	F	234	238	240	13±0
2	H	221	222	224	7±1
2	J	208	218	219	11±1
2	L	234	238	240	11±4
2	N	234	238	240	12±2
2	P	235	238	240	6±2
All	All	14204	14420	14540	319

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 11.

5 of 201 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:F:12:VAL:HB	1:G:12:VAL:HG12	0.78	1.53	3	4
2:H:37:GLY:HA3	2:P:35:MET:HG2	0.77	1.56	3	3
1:G:35:MET:HG2	1:O:37:GLY:HA3	0.77	1.56	3	2
2:N:12:VAL:HB	1:O:12:VAL:HG12	0.75	1.58	4	4
2:L:28:LYS:HD3	1:M:28:LYS:HA	0.75	1.57	3	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	28/40 (70%)	21±2 (75±7%)	7±2 (25±7%)	0±0 (0±0%)	100	100
1	C	28/40 (70%)	23±0 (81±2%)	5±0 (17±2%)	0±0 (2±2%)	12	54
1	E	28/40 (70%)	22±1 (79±5%)	5±1 (17±5%)	1±0 (4±0%)	6	34
1	G	28/40 (70%)	25±1 (89±3%)	2±1 (9±4%)	0±0 (2±2%)	12	54
1	I	27/40 (68%)	22±1 (80±3%)	6±1 (20±3%)	0±0 (0±0%)	100	100
1	K	28/40 (70%)	22±0 (79±2%)	5±1 (17±4%)	1±1 (4±3%)	6	34
1	M	28/40 (70%)	23±1 (81±4%)	4±1 (16±4%)	1±0 (3±2%)	8	43
1	O	28/40 (70%)	25±1 (88±3%)	3±1 (11±3%)	0±0 (1±2%)	21	69
2	B	28/42 (67%)	21±1 (76±3%)	6±1 (23±3%)	0±0 (1±2%)	21	69
2	D	30/42 (71%)	27±1 (90±2%)	3±1 (10±2%)	0±0 (0±0%)	100	100
2	F	30/42 (71%)	27±1 (90±4%)	3±1 (10±4%)	0±0 (0±0%)	100	100
2	H	29/42 (69%)	24±0 (82±1%)	4±0 (15±1%)	1±0 (3±0%)	6	36
2	J	28/42 (67%)	22±0 (77±2%)	6±1 (21±3%)	0±0 (2±2%)	12	54
2	L	30/42 (71%)	26±1 (88±4%)	4±1 (12±4%)	0±0 (0±0%)	100	100
2	N	30/42 (71%)	27±0 (91±1%)	3±0 (9±1%)	0±0 (0±0%)	100	100
2	P	30/42 (71%)	24±1 (79±4%)	5±1 (18±4%)	1±0 (3±0%)	6	37
All	All	1832/2624 (70%)	1520 (83%)	285 (16%)	27 (1%)	14	59

5 of 12 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	E	25	GLY	4
2	H	25	GLY	4
1	K	26	SER	3
1	M	25	GLY	3
2	P	25	GLY	3

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	22/31 (71%)	20±1 (91±3%)	2±1 (9±3%)	13	59
1	C	23/31 (74%)	22±0 (95±2%)	1±0 (5±2%)	26	75
1	E	23/31 (74%)	20±1 (86±4%)	3±1 (14±4%)	6	46
1	G	23/31 (74%)	22±0 (96±0%)	1±0 (4±0%)	33	81
1	I	21/31 (68%)	19±1 (92±4%)	2±1 (8±4%)	15	62
1	K	23/31 (74%)	21±1 (90±4%)	2±1 (10±4%)	11	57
1	M	23/31 (74%)	19±1 (84±4%)	4±1 (16±4%)	5	41
1	O	23/31 (74%)	23±0 (99±2%)	0±0 (1±2%)	74	96
2	B	21/32 (66%)	20±1 (94±4%)	1±1 (6±4%)	23	72
2	D	24/32 (75%)	21±1 (89±5%)	3±1 (11±5%)	9	52
2	F	24/32 (75%)	22±1 (94±4%)	2±1 (6±4%)	21	70
2	H	23/32 (72%)	19±0 (84±2%)	4±0 (16±2%)	5	41
2	J	21/32 (66%)	20±1 (95±5%)	1±1 (5±5%)	29	78
2	L	24/32 (75%)	22±0 (91±2%)	2±0 (9±2%)	12	58
2	N	24/32 (75%)	22±0 (94±2%)	2±0 (6±2%)	21	70
2	P	24/32 (75%)	20±2 (81±7%)	4±2 (19±7%)	4	36
All	All	1464/2016 (73%)	1328 (91%)	136 (9%)	12	59

5 of 68 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)
2	D	23	ASP	4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
2	D	27	ASN	4
1	E	23	ASP	4
2	F	28	LYS	4
2	H	12	VAL	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 (i)

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 (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	233
Number of shifts mapped to atoms	178
Number of unparsed shifts	0
Number of shifts with mapping errors	55
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 55) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	1	ASP	C	172.758	0.3	1
1	A	1	ASP	CA	52.028	0.3	1
1	A	1	ASP	CB	41.529	0.3	1
1	A	1	ASP	CG	179.54	0.3	1
1	A	1	ASP	N	130.337	0.3	1
1	A	2	ALA	C	174.804	0.3	1
1	A	2	ALA	CA	49.036	0.3	1
1	A	2	ALA	CB	22.229	0.3	1
1	A	2	ALA	N	125.125	0.3	1
1	A	3	GLU	C	172.758	0.3	1
1	A	3	GLU	CA	53.639	0.3	1
1	A	3	GLU	CB	32.019	0.3	1
1	A	3	GLU	CG	34.637	0.3	1
1	A	3	GLU	CD	180.565	0.3	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	3	GLU	N	124.357	0.3	1
1	A	4	PHE	C	173.99	0.3	1
1	A	4	PHE	CA	57.386	0.3	1
1	A	4	PHE	CB	40.375	0.3	1
1	A	4	PHE	CG	137.523	0.3	1
1	A	4	PHE	CD1	130.492	0.3	1
1	A	4	PHE	CE1	129.683	0.3	1
1	A	4	PHE	N	127.61	0.3	1
1	A	5	ARG	C	181.505	0.3	1
1	A	5	ARG	CA	53.94	0.3	1
1	A	5	ARG	CB	35.649	0.3	1
1	A	5	ARG	CG	25.921	0.3	1
1	A	5	ARG	CD	43.008	0.3	1
1	A	5	ARG	CZ	158.633	0.3	1
1	A	5	ARG	N	127.421	0.3	1
1	A	6	HIS	C	174.291	0.3	1
1	A	6	HIS	CA	53.453	0.3	1
1	A	6	HIS	CB	30.139	0.3	1
1	A	6	HIS	CG	131.642	0.3	1
1	A	6	HIS	CD2	116.388	0.3	1
1	A	6	HIS	CE1	134.609	0.3	1
1	A	6	HIS	N	127.638	0.3	1
1	A	7	ASP	C	175.923	0.3	1
1	A	7	ASP	CA	54.629	0.3	1
1	A	7	ASP	CB	33.967	0.3	1
1	A	7	ASP	CG	178.973	0.3	1
1	A	7	ASP	N	124.043	0.3	1
1	A	8	SER	C	172.307	0.3	1
1	A	8	SER	CA	55.449	0.3	1
1	A	8	SER	CB	63.447	0.3	1
1	A	8	SER	N	111.244	0.3	1
1	A	9	GLY	C	172.482	0.3	1
1	A	9	GLY	CA	42.174	0.3	1
1	A	9	GLY	N	114.288	0.3	1
1	A	10	TYR	C	173.368	0.3	1
1	A	10	TYR	CA	56.258	0.3	1
1	A	10	TYR	CB	34.772	0.3	1
1	A	10	TYR	CG	127.497	0.3	1
1	A	10	TYR	CD1	132.205	0.3	1
1	A	10	TYR	CE1	116.659	0.3	1
1	A	10	TYR	N	123.519	0.3	1

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	42	1.53 ± 0.40	Should be checked
$^{13}C_{\beta}$	36	0.35 ± 0.68	None needed (< 0.5 ppm)
$^{13}C'$	42	1.63 ± 0.32	Should be applied
^{15}N	42	-2.02 ± 0.79	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 3%, i.e. 168 atoms were assigned a chemical shift out of a possible 6356. 0 out of 125 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	93/2505 (4%)	0/1050 (0%)	62/970 (6%)	31/485 (6%)
Sidechain	66/3351 (2%)	0/2235 (0%)	66/1052 (6%)	0/64 (0%)
Aromatic	9/500 (2%)	0/280 (0%)	9/220 (4%)	0/0 (—%)
Overall	168/6356 (3%)	0/3565 (0%)	137/2242 (6%)	31/549 (6%)

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.

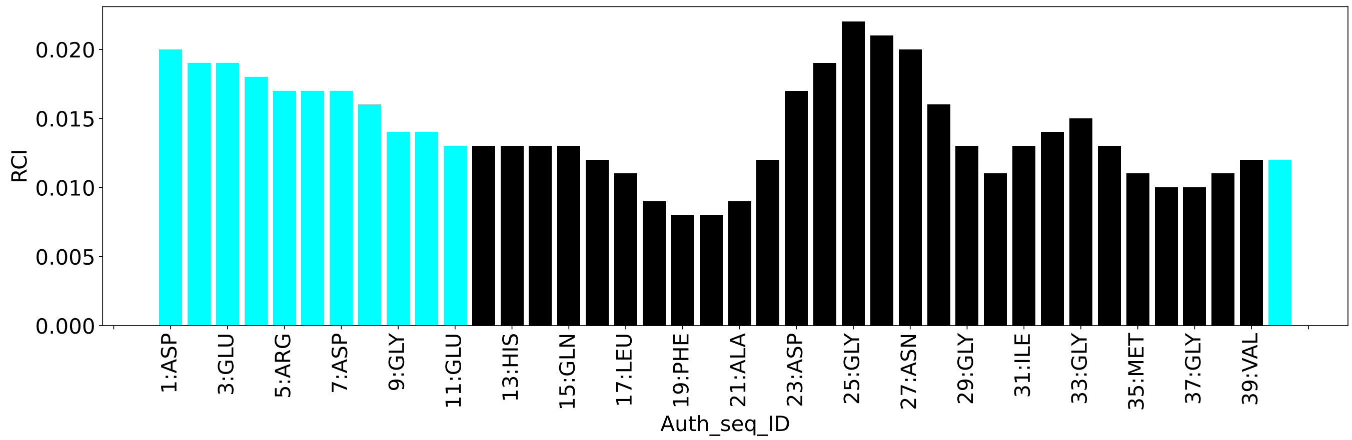
7.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

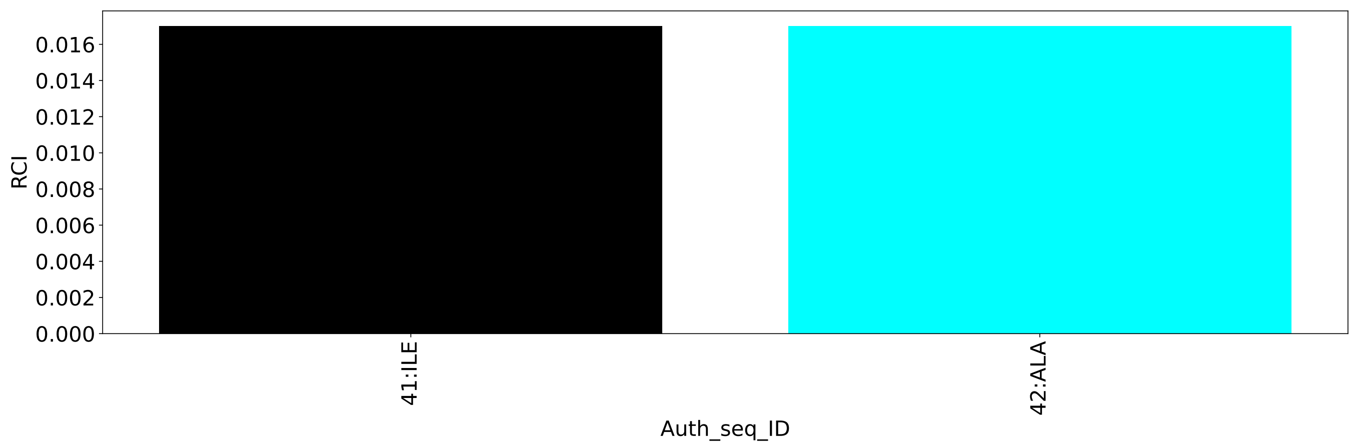
7.1.5 Random Coil Index (RCI) plots (i)

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:



Random coil index (RCI) for chain B:



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	27
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)	2
Inter-chain	25
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 (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)	None	None
0.2-0.5 (Medium)	None	None
>0.5 (Large)	26.0	24.18

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)

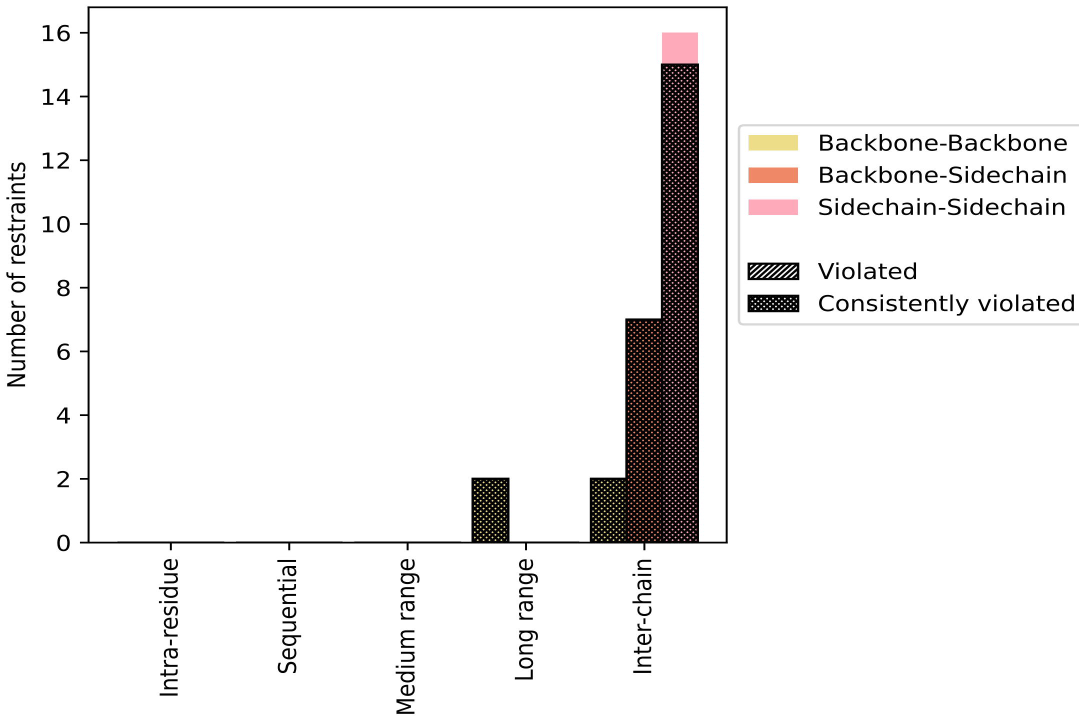
9.1 Summary of distance violations (i)

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)	2	7.4	2	100.0	7.4	2	100.0	7.4
Backbone-Backbone	2	7.4	2	100.0	7.4	2	100.0	7.4
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
Inter-chain	25	92.6	24	96.0	88.9	24	96.0	88.9
Backbone-Backbone	2	7.4	2	100.0	7.4	2	100.0	7.4
Backbone-Sidechain	7	25.9	7	100.0	25.9	7	100.0	25.9
Sidechain-Sidechain	16	59.3	15	93.8	55.6	15	93.8	55.6
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	27	100.0	26	96.3	96.3	26	96.3	96.3
Backbone-Backbone	4	14.8	4	100.0	14.8	4	100.0	14.8
Backbone-Sidechain	7	25.9	7	100.0	25.9	7	100.0	25.9
Sidechain-Sidechain	16	59.3	15	93.8	55.6	15	93.8	55.6

[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

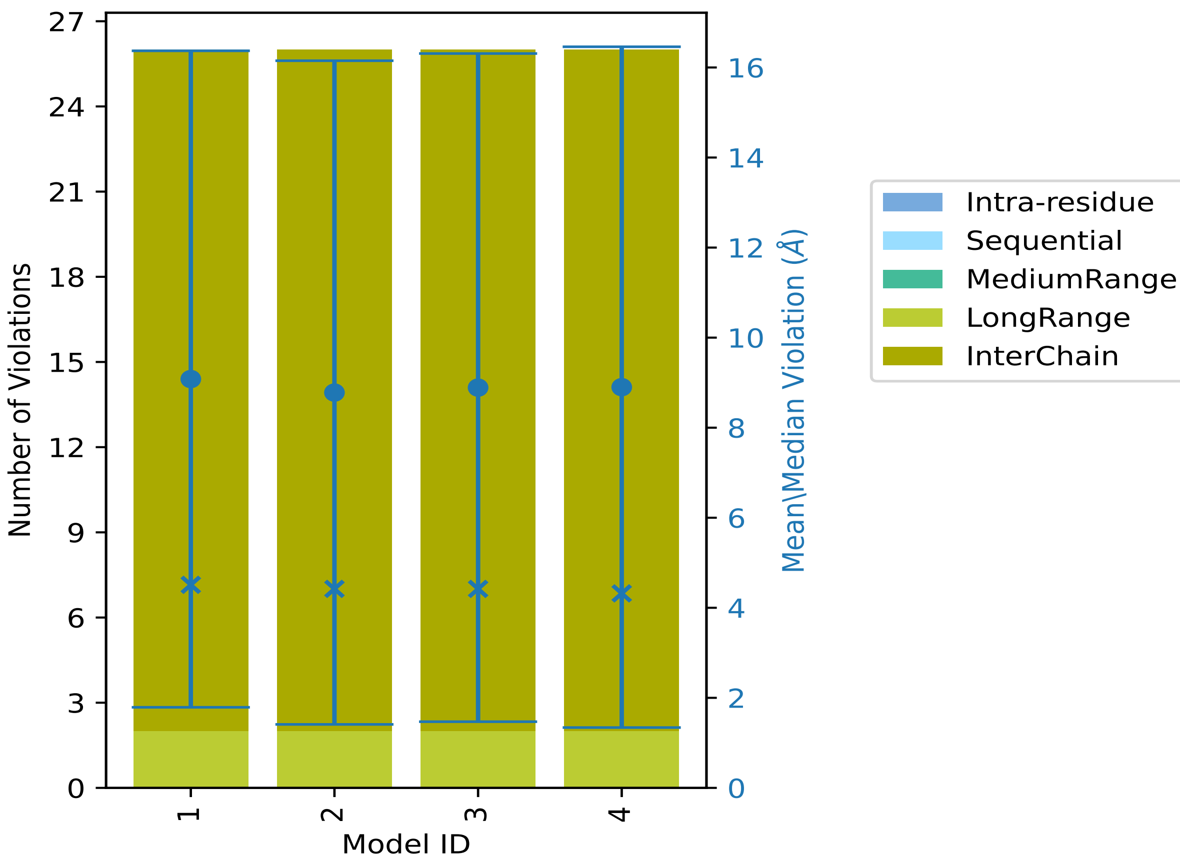
9.2 Distance violation statistics for each model ⓘ

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	2	24	26	9.08	24.16	7.29	4.51
2	0	0	0	2	24	26	8.78	23.71	7.37	4.42
3	0	0	0	2	24	26	8.89	24.1	7.42	4.42
4	0	0	0	2	24	26	8.9	24.18	7.56	4.32

[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

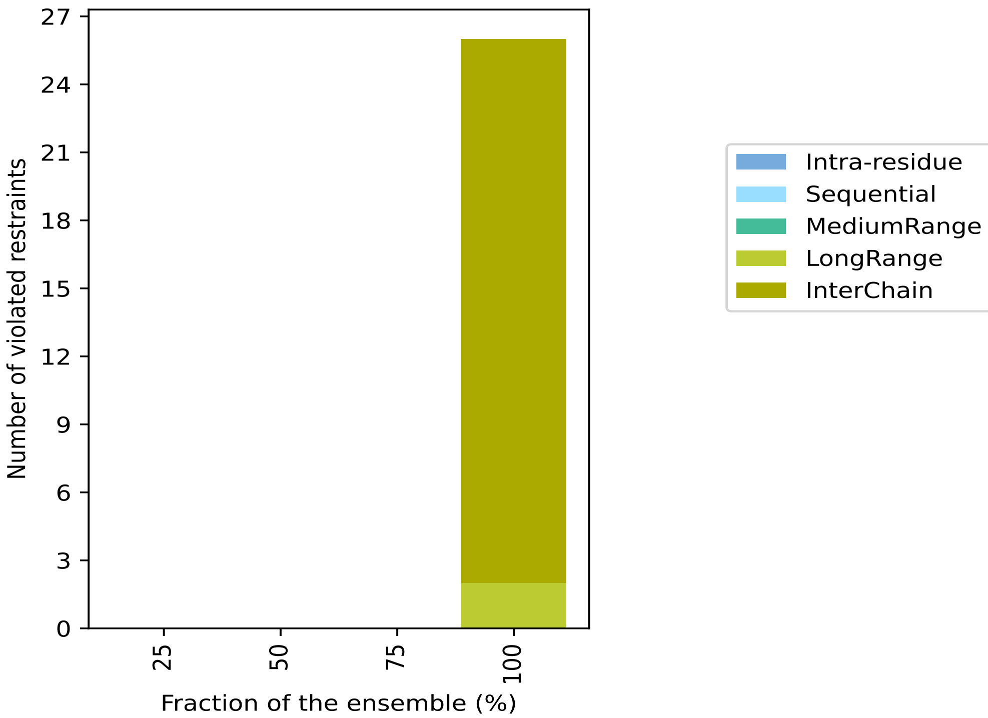
9.3 Distance violation statistics for the ensemble (i)

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, 1(IR:0, SQ:0, MR:0, LR:0, IC:1) 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	0	0	0	1	25.0
0	0	0	0	0	0	2	50.0
0	0	0	0	0	0	3	75.0
0	0	0	2	24	26	4	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

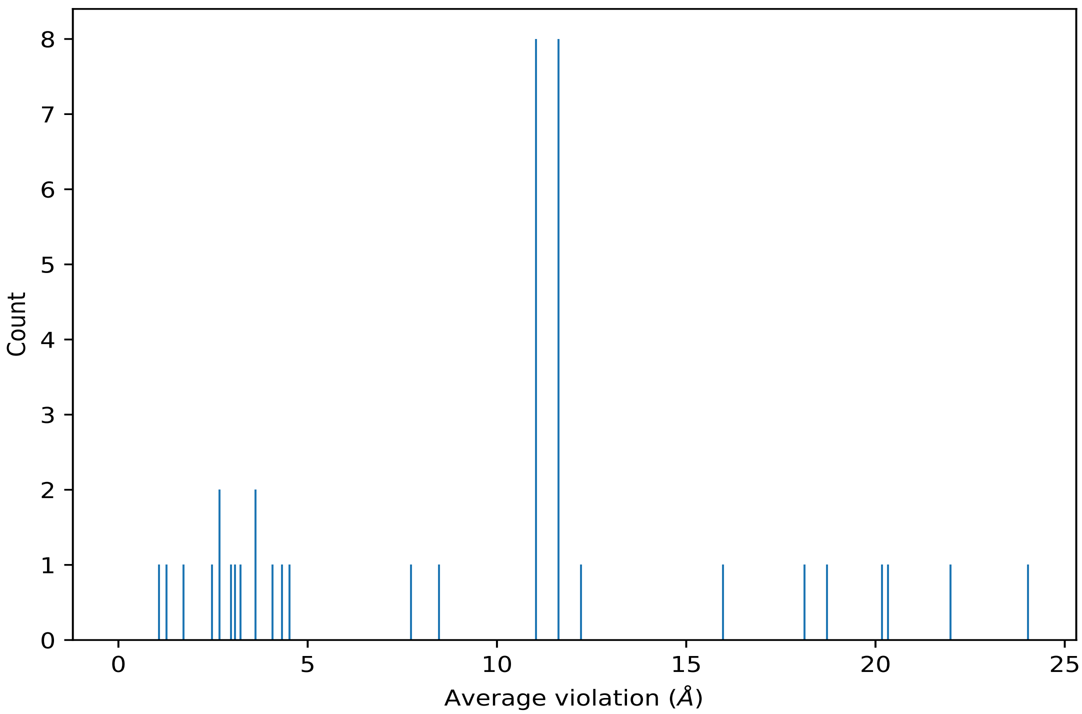
9.3.1 Bar graph : Distance violation statistics for the ensemble (i)



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)
(2,6)	1:A:31:ILE:CG2	2:B:40:VAL:CG1	4	24.04	0.19	24.13
(2,2)	1:A:31:ILE:CD1	2:B:39:VAL:CG2	4	21.96	0.8	21.96
(2,1)	1:A:31:ILE:CB	2:B:39:VAL:CG1	4	20.34	1.04	20.39
(2,3)	1:A:31:ILE:CG2	2:B:39:VAL:C	4	20.16	0.12	20.13
(2,5)	1:A:31:ILE:CG2	2:B:39:VAL:CB	4	18.7	0.13	18.71
(2,7)	1:A:31:ILE:CG2	2:B:39:VAL:CG2	4	18.14	0.78	18.12
(2,4)	1:A:31:ILE:CG2	2:B:38:GLY:CA	4	15.97	0.15	15.94
(3,4)	2:B:12:VAL:C	2:B:42:ALA:C	4	12.21	0.57	12.02
(3,2)	1:A:12:VAL:C	2:B:42:ALA:C	4	11.62	0.31	11.56
(3,2)	1:A:12:VAL:C	2:D:42:ALA:C	4	11.62	0.31	11.56
(3,2)	1:A:12:VAL:C	2:F:42:ALA:C	4	11.62	0.31	11.56
(3,2)	1:A:12:VAL:C	2:H:42:ALA:C	4	11.62	0.31	11.56
(3,2)	1:A:12:VAL:C	2:J:42:ALA:C	4	11.62	0.31	11.56
(3,2)	1:A:12:VAL:C	2:L:42:ALA:C	4	11.62	0.31	11.56
(3,2)	1:A:12:VAL:C	2:N:42:ALA:C	4	11.62	0.31	11.56
(3,2)	1:A:12:VAL:C	2:P:42:ALA:C	4	11.62	0.31	11.56

Continued on next page...

Continued from previous page...

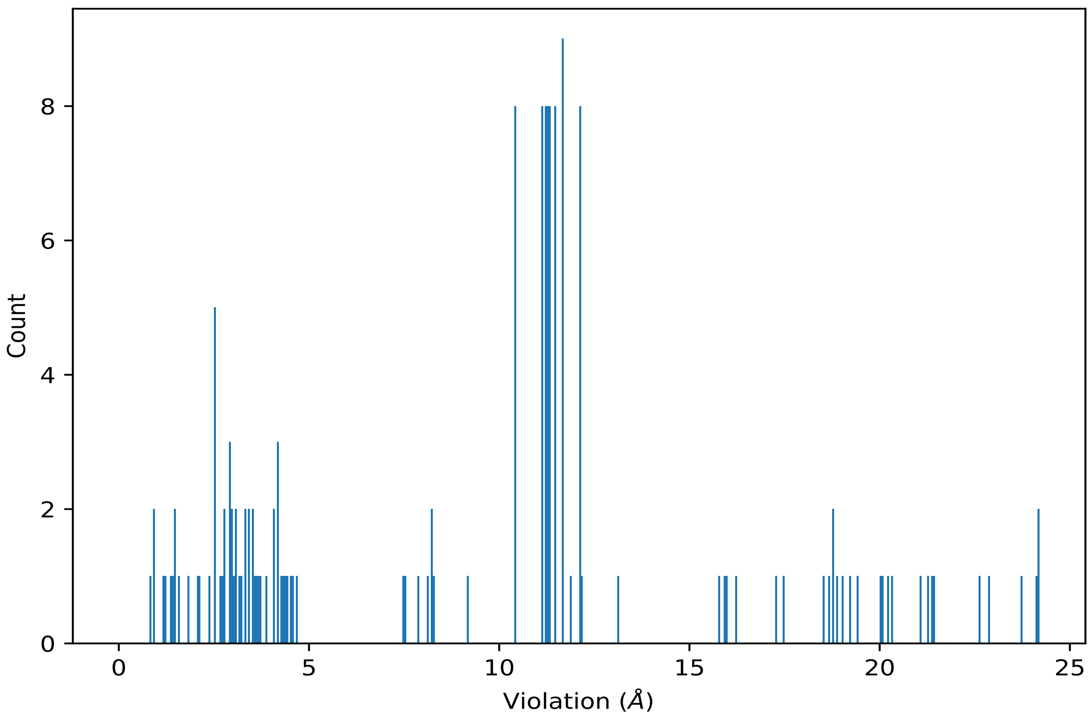
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(3,1)	1:A:11:GLU:C	2:B:41:ILE:C	4	11.02	0.35	11.16

[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 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)
(2,6)	1:A:31:ILE:CG2	2:B:40:VAL:CG1	4	24.18
(2,6)	1:A:31:ILE:CG2	2:B:40:VAL:CG1	1	24.16
(2,6)	1:A:31:ILE:CG2	2:B:40:VAL:CG1	3	24.1
(2,6)	1:A:31:ILE:CG2	2:B:40:VAL:CG1	2	23.71

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(2,2)	1:A:31:ILE:CD1	2:B:39:VAL:CG2	4	22.89
(2,2)	1:A:31:ILE:CD1	2:B:39:VAL:CG2	1	22.63
(2,1)	1:A:31:ILE:CB	2:B:39:VAL:CG1	2	21.4
(2,1)	1:A:31:ILE:CB	2:B:39:VAL:CG1	3	21.36
(2,2)	1:A:31:ILE:CD1	2:B:39:VAL:CG2	3	21.28
(2,2)	1:A:31:ILE:CD1	2:B:39:VAL:CG2	2	21.06

10 Dihedral-angle violation analysis (i)

No dihedral-angle restraints found