wwPDB NMR Structure Validation Summary Report (i)

Sep 28, 2025 – 02:32 PM EDT

PDB ID	:	9ONB / pdb_00009onb
BMRB ID	:	31248
Title	:	Immature HIV-1 CACTD-SP1 lattice with Maturation inhibitor PF-46396 (S) and Inositol hexakisphosphate (IP6)
Authors	:	Zadorozhnyi, R.; Quinn, C.M.; Zadrozny, K.K.; Ablan, S.D.; Kennedy, B.J.; Yap, G.P.A.; Sanner, D.; Kraml, C.; Freed, E.O.; Ganser-Pornillos, B.K.; Pornillos, O.; Gronenborn, A.M.; Polenova, T.
Deposited on	:	2025-05-14

The following versions of software and data (see references (1)) were used in the production of this report:

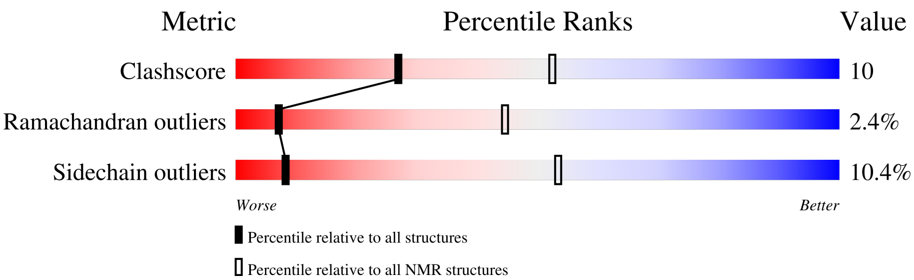
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.46

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 6%.

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	210492	14027
Ramachandran outliers	207382	12486
Sidechain outliers	206894	12463

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	G	102	73% 23% •• •
1	H	102	66% 26% • • •
1	I	102	57% 26% 5% 12%
1	J	102	69% 22% •• 7%
1	K	102	69% 21% 5% 6%
1	L	102	68% 25% 5% •

2 Ensemble composition and analysis (i)

This entry contains 10 models. Model 2 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	G:146-G:244, H:147-H:245, I:147-I:148, I:153-I:205, I:210-I:244, J:147-J:205, J:210-J:245, K:144-K:147, K:154-K:245, L:147-L:245 (578)	0.62	2

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

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

The models can be grouped into 2 clusters and 2 single-model clusters were found.

Cluster number	Models
1	1, 3, 4, 8
2	2, 6, 9, 10
Single-model clusters	5; 7

3 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9410 atoms, of which 4703 are hydrogens and 0 are deuteriums.

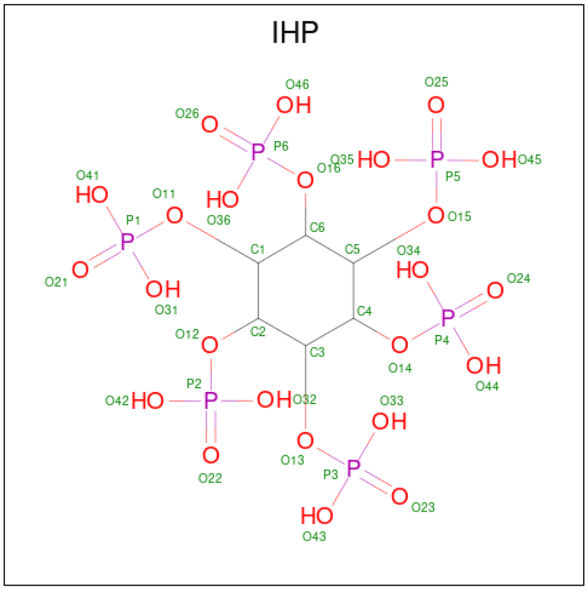
- Molecule 1 is a protein called Capsid protein p24.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	G	102	1551	479	778	136	151	7	0
1	H	102	1551	479	778	136	151	7	0
1	I	102	1551	479	778	136	151	7	0
1	J	102	1551	479	778	136	151	7	0
1	K	102	1551	479	778	136	151	7	0
1	L	102	1551	479	778	136	151	7	0

There are 18 discrepancies between the modelled and reference sequences:

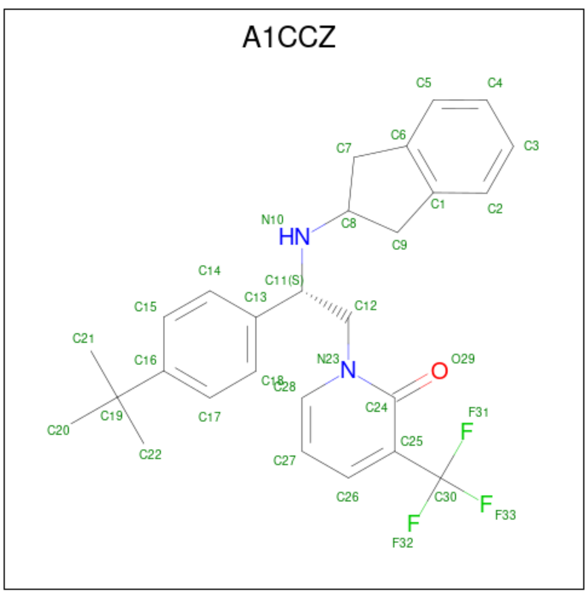
Chain	Residue	Modelled	Actual	Comment	Reference
G	144	GLY	-	expression tag	UNP P12497
G	145	GLY	-	expression tag	UNP P12497
G	241	THR	PRO	conflict	UNP P12497
H	144	GLY	-	expression tag	UNP P12497
H	145	GLY	-	expression tag	UNP P12497
H	241	THR	PRO	conflict	UNP P12497
I	144	GLY	-	expression tag	UNP P12497
I	145	GLY	-	expression tag	UNP P12497
I	241	THR	PRO	conflict	UNP P12497
J	144	GLY	-	expression tag	UNP P12497
J	145	GLY	-	expression tag	UNP P12497
J	241	THR	PRO	conflict	UNP P12497
K	144	GLY	-	expression tag	UNP P12497
K	145	GLY	-	expression tag	UNP P12497
K	241	THR	PRO	conflict	UNP P12497
L	144	GLY	-	expression tag	UNP P12497
L	145	GLY	-	expression tag	UNP P12497
L	241	THR	PRO	conflict	UNP P12497

- Molecule 2 is INOSITOL HEXAKISPHOSPHATE (CCD ID: IHP) (formula: $C_6H_{18}O_{24}P_6$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				
			Total	C	H	O	P
2	H	1	42	6	6	24	6

- Molecule 3 is 1-{(2S)-2-(4-tert-butylphenyl)-2-[(2,3-dihydro-1H-inden-2-yl)amino]ethyl}-3-(trifluoromethyl)pyridin-2(1H)-one (CCD ID: A1CCZ) (formula: $C_{27}H_{29}F_3N_2O$) (labeled as "Ligand of Interest" by depositor).



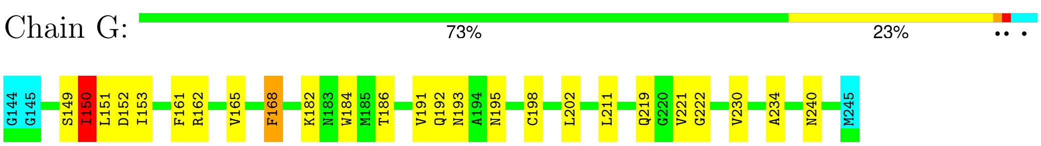
Mol	Chain	Residues	Atoms					
			Total	C	F	H	N	O
3	H	1	62	27	3	29	2	1

4 Residue-property plots (i)

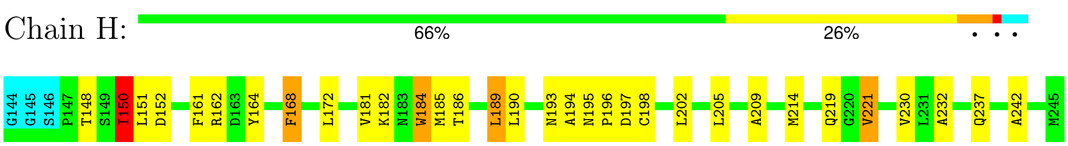
4.1 Average score per residue in the NMR ensemble

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.

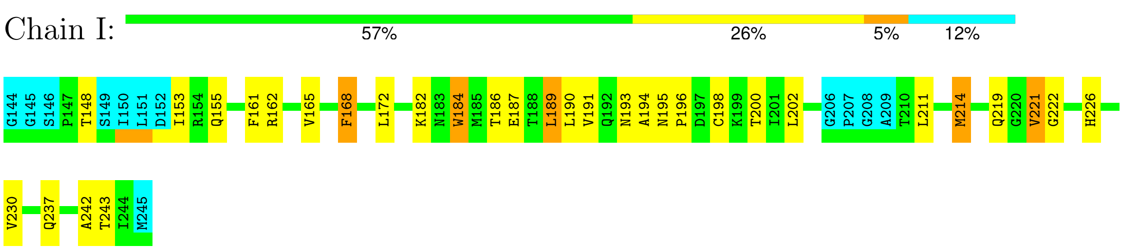
- Molecule 1: Capsid protein p24



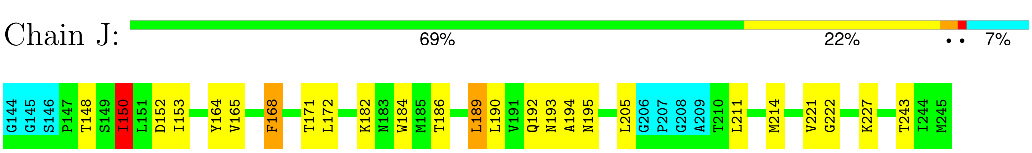
- Molecule 1: Capsid protein p24



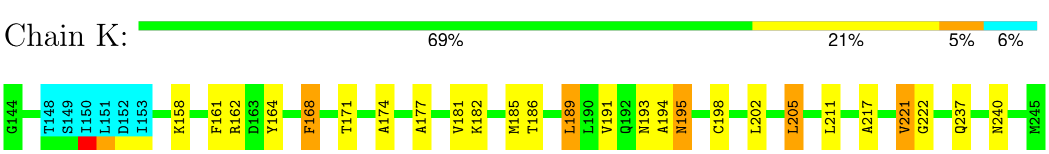
- Molecule 1: Capsid protein p24



- Molecule 1: Capsid protein p24

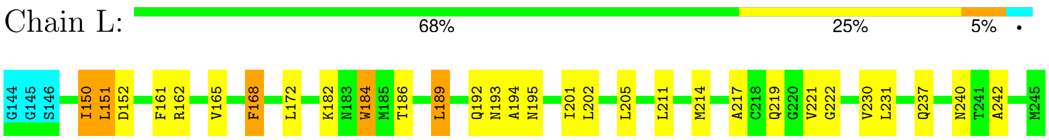


- Molecule 1: Capsid protein p24



- Molecule 1: Capsid protein p24

Chain L:

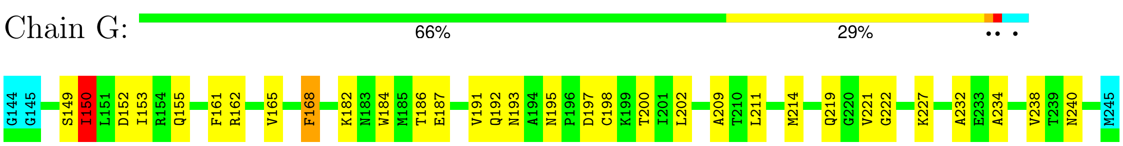


4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 2. Colouring as in section 4.1 above.

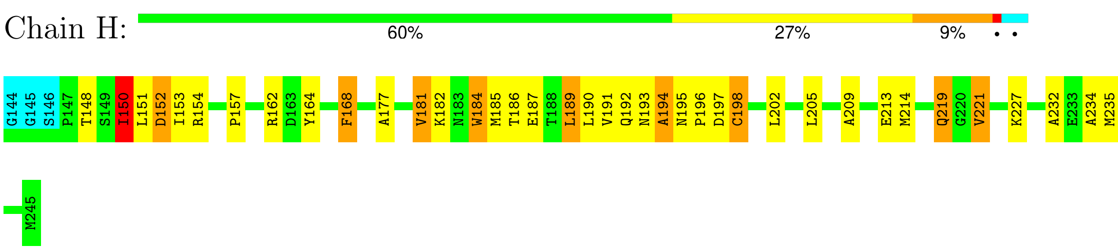
- Molecule 1: Capsid protein p24

Chain G:



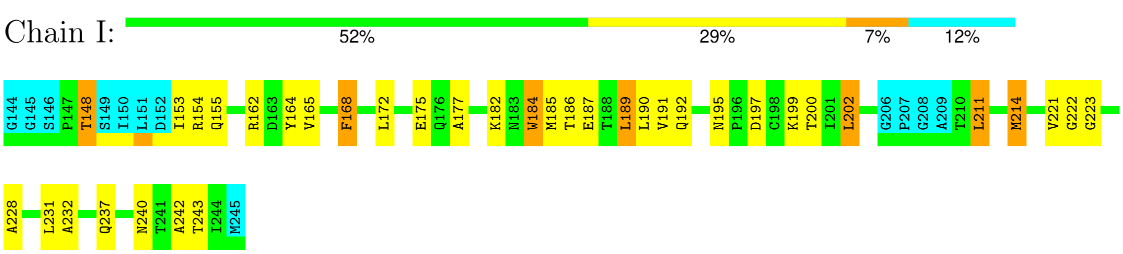
- Molecule 1: Capsid protein p24

Chain H:



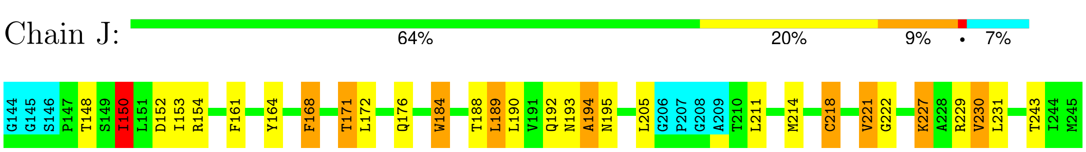
- Molecule 1: Capsid protein p24

Chain I:



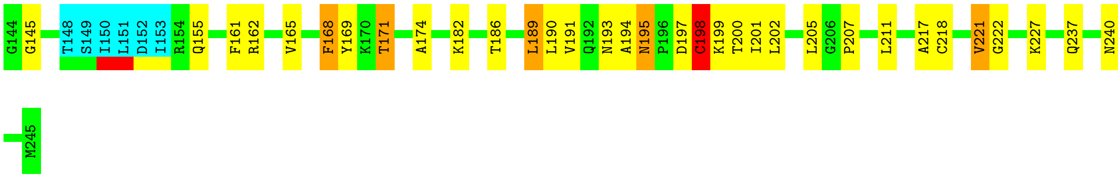
- Molecule 1: Capsid protein p24

Chain J:



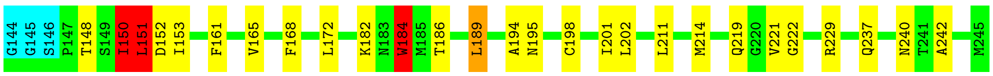
- Molecule 1: Capsid protein p24

Chain K:



• Molecule 1: Capsid protein p24

Chain L:



5 Refinement protocol and experimental data overview (i)

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

Of the 100 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	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	514
Number of shifts mapped to atoms	514
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	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.

6 Model quality (i)

6.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: IHP, A1CCZ

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	G	1.39±0.02	6±1/769 (0.7± 0.2%)	1.50±0.02	10±2/1043 (1.0± 0.1%)
1	H	1.33±0.03	4±2/771 (0.6± 0.3%)	1.54±0.03	10±2/1041 (1.0± 0.2%)
1	I	1.27±0.05	2±2/712 (0.3± 0.2%)	1.56±0.03	11±2/964 (1.1± 0.2%)
1	J	1.17±0.02	1±1/750 (0.2± 0.1%)	1.48±0.04	11±2/1012 (1.0± 0.2%)
1	K	1.12±0.04	1±1/740 (0.2± 0.1%)	1.42±0.04	7±1/998 (0.7± 0.1%)
1	L	1.25±0.04	4±2/771 (0.5± 0.2%)	1.56±0.05	15±4/1041 (1.4± 0.4%)
All	All	1.26	185/45130 (0.4%)	1.51	635/60990 (1.0%)

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	I	0.0±0.0	0.1±0.3
All	All	0	1

5 of 64 unique bond 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(Å)	Ideal(Å)	Models	
								Worst	Total
1	L	221	VAL	C-N	13.95	1.44	1.33	7	10
1	I	221	VAL	C-N	9.88	1.42	1.32	10	4
1	L	152	ASP	C-N	9.85	1.43	1.33	9	6
1	J	221	VAL	C-N	9.01	1.41	1.32	9	5
1	H	152	ASP	C-N	8.13	1.41	1.33	9	2

5 of 169 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(°)	Ideal(°)	Models	
								Worst	Total
1	L	184	TRP	N-CA-C	-10.72	99.56	111.14	5	10
1	L	151	LEU	N-CA-C	-10.49	98.72	112.68	7	6
1	I	222	GLY	CA-C-O	-10.25	115.07	122.45	1	5
1	H	151	LEU	N-CA-C	-9.94	99.46	112.68	2	3
1	H	152	ASP	N-CA-CB	-8.69	96.22	111.42	2	1

There are no chirality outliers.

All unique planar outliers are listed below.

Mol	Chain	Res	Type	Group	Models (Total)
1	I	162	ARG	Sidechain	1

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	G	756	761	761	12±3
1	H	759	765	765	21±4
1	I	700	707	707	18±6
1	J	739	747	747	13±4
1	K	728	729	726	13±3
1	L	759	765	765	14±2
3	H	33	29	0	0±1
All	All	45100	45090	44770	868

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 10.

5 of 366 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:L:150:ILE:HD11	1:L:171:THR:HG23	0.87	1.44	6	4
1:G:234:ALA:HB1	1:H:232:ALA:HB2	0.85	1.47	1	7
1:H:234:ALA:HB1	1:I:232:ALA:HB2	0.80	1.53	9	2
1:I:153:ILE:HG22	1:I:189:LEU:HD21	0.79	1.53	9	3
1:H:152:ASP:HA	1:H:189:LEU:HD11	0.79	1.54	2	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	G	99/102 (97%)	91±1 (92±1%)	6±1 (6±1%)	2±0 (2±0%)	8	50
1	H	98/102 (96%)	87±1 (89±1%)	8±1 (8±1%)	3±1 (3±1%)	5	34
1	I	90/102 (88%)	80±2 (89±2%)	8±2 (9±3%)	2±0 (2±1%)	9	52
1	J	94/102 (92%)	82±2 (88±2%)	9±2 (9±3%)	3±1 (3±1%)	6	39
1	K	94/102 (92%)	85±2 (90±2%)	8±2 (8±2%)	2±0 (2±1%)	10	55
1	L	98/102 (96%)	86±1 (88±1%)	9±1 (9±1%)	2±1 (3±1%)	6	42
All	All	5730/6120 (94%)	5117 (89%)	477 (8%)	136 (2%)	7	44

5 of 24 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	G	150	ILE	10
1	G	195	ASN	10
1	H	150	ILE	10
1	H	195	ASN	10
1	I	195	ASN	10

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	G	82/83 (99%)	74±1 (91±1%)	8±1 (9±1%)	10	57
1	H	82/83 (99%)	72±2 (88±2%)	10±2 (12±2%)	7	50
1	I	76/83 (92%)	68±1 (90±2%)	8±1 (10±2%)	9	54
1	J	81/83 (98%)	72±2 (89±2%)	9±2 (11±2%)	7	51

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	K	77/83 (93%)	68±2 (89±3%)	8±2 (11±3%)	7	52
1	L	82/83 (99%)	74±2 (91±2%)	8±2 (9±2%)	10	57
All	All	4800/4980 (96%)	4299 (90%)	501 (10%)	8	53

5 of 129 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	G	150	ILE	10
1	G	168	PHE	10
1	G	184	TRP	10
1	G	193	ASN	10
1	G	211	LEU	10

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 oligosaccharides in this entry.

6.6 Ligand geometry (i)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
2	IHP	H	301	-	36,36,36	1.67±0.00	6±1 (15±1%)
3	A1CCZ	H	302	-	36,36,36	2.03±0.00	13±0 (36±0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of angles that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
2	IHP	H	301	-	60,60,60	1.30±0.00	7±0 (11±0%)
3	A1CCZ	H	302	-	44,54,54	2.20±0.00	12±0 (27±0%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	IHP	H	301	-	-	0±0,30,54,54	0±0,1,1,1
3	A1CCZ	H	302	-	-	0±0,24,32,32	0±0,4,4,4

5 of 20 unique bond 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(Å)	Ideal(Å)	Models	
								Worst	Total
2	H	301	IHP	C6-C5	5.43	1.63	1.52	3	10
3	H	302	A1CCZ	C12-N23	5.05	1.53	1.47	9	10
3	H	302	A1CCZ	C8-N10	3.82	1.54	1.48	3	10
3	H	302	A1CCZ	C7-C8	3.73	1.47	1.53	1	10
3	H	302	A1CCZ	C18-C17	3.53	1.33	1.38	9	10

5 of 19 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
3	H	302	A1CCZ	C11-C12-N23	6.40	121.30	111.78	3	10

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
3	H	302	A1CCZ	C7-C6-C1	6.16	106.36	110.50	10	10
3	H	302	A1CCZ	C1-C9-C8	4.94	98.02	102.89	2	10
3	H	302	A1CCZ	C7-C6-C5	4.06	136.77	129.17	4	10
2	H	301	IHP	C5-C4-C3	3.82	102.05	110.43	9	10

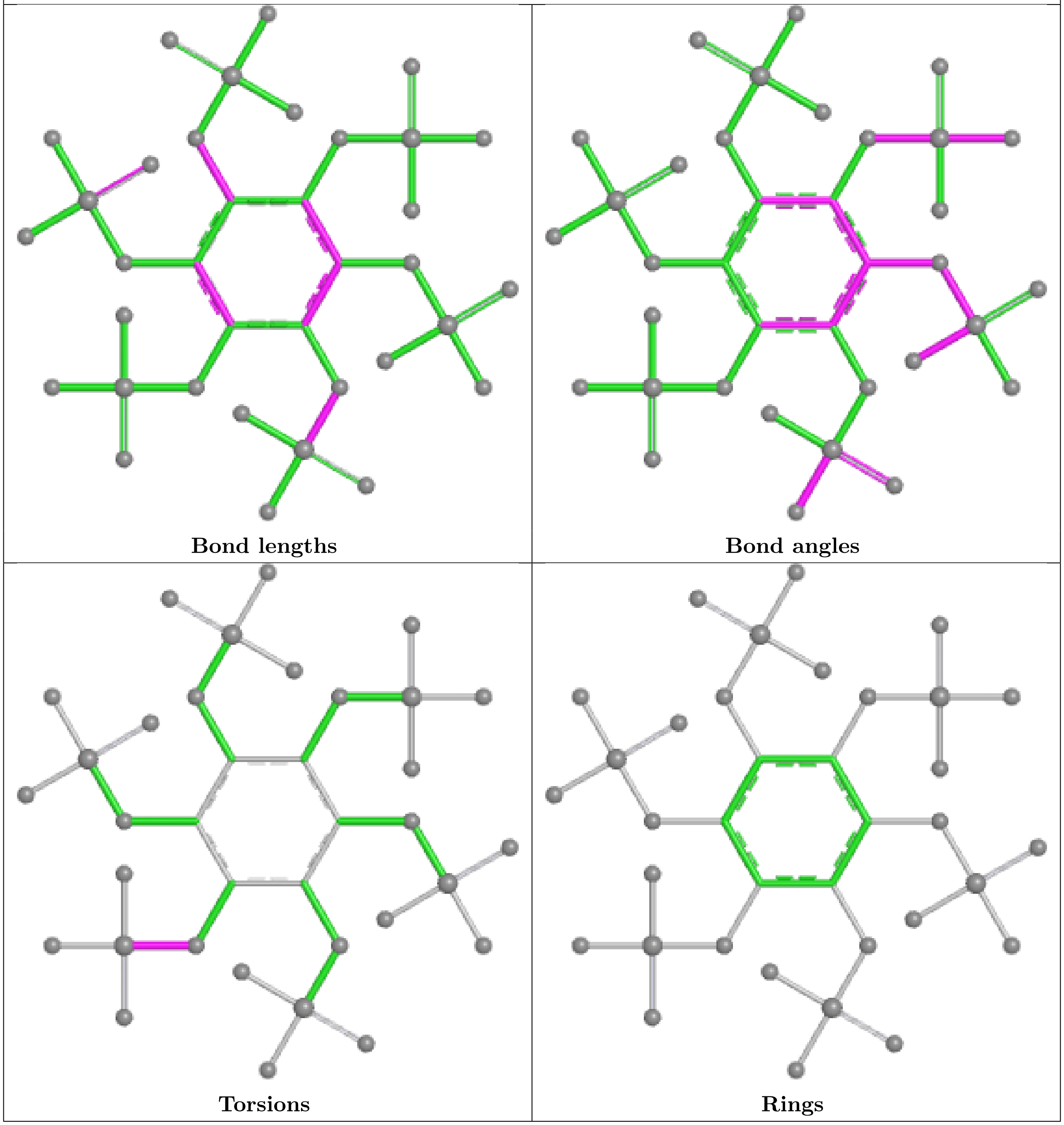
There are no chirality outliers.

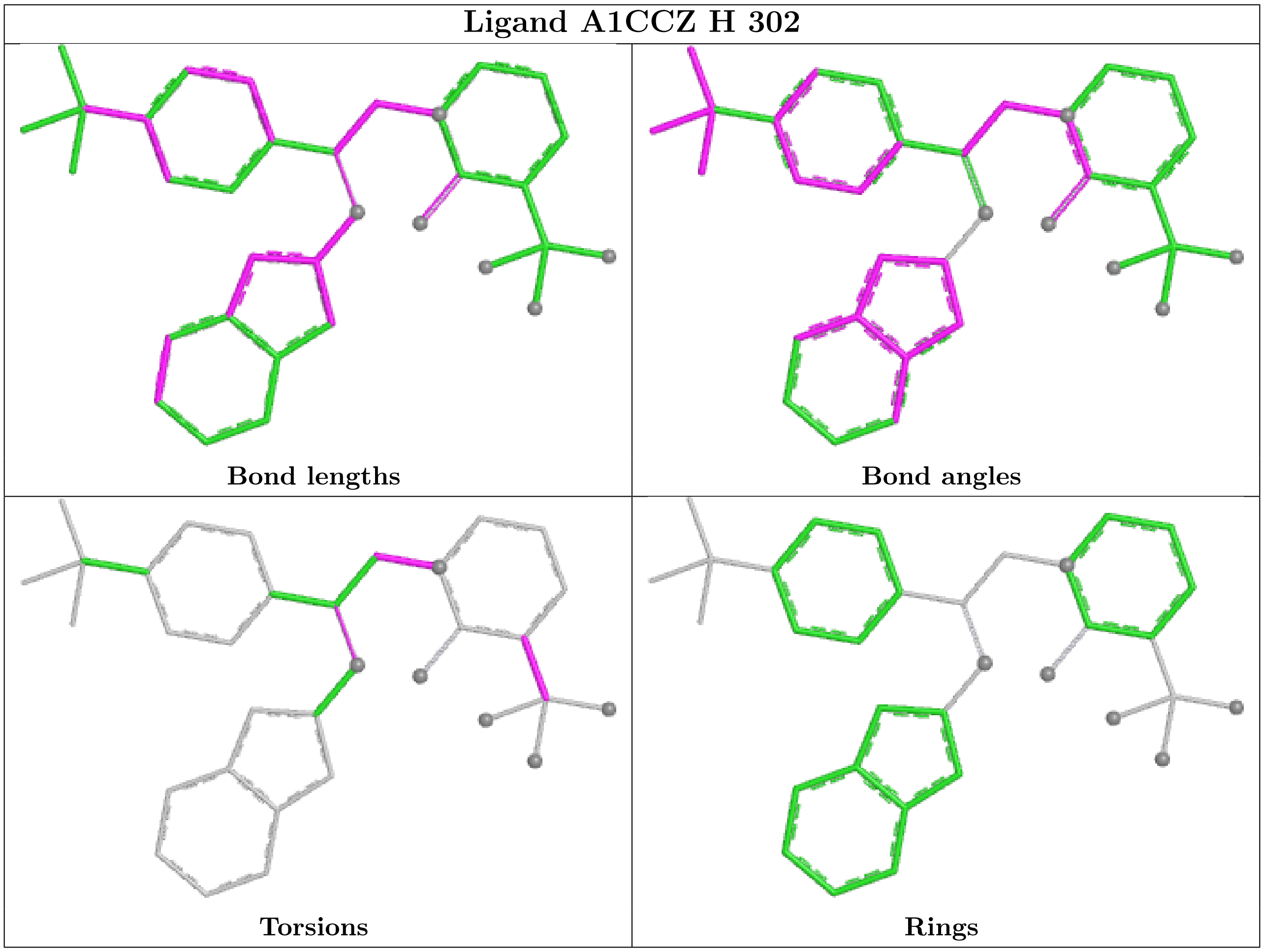
There are no torsion outliers.

There are no ring outliers.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand IHP H 301





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 6% for the well-defined parts and 6% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *D_1000295827_cs_P1.str.V1*

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	514
Number of shifts mapped to atoms	514
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	2

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	96	-0.78 ± 0.21	Should be checked
$^{13}C_{\beta}$	87	0.07 ± 0.17	None needed (< 0.5 ppm)
$^{13}C'$	92	-0.75 ± 0.16	Should be applied
^{15}N	93	-0.46 ± 0.78	None needed (< 0.5 ppm)

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 6%, i.e. 494 atoms were assigned a chemical shift out of a possible 7723. 0 out of 82 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	272/2862 (10%)	2/1162 (0%)	184/1156 (16%)	86/544 (16%)
Sidechain	203/4519 (4%)	6/2939 (0%)	196/1394 (14%)	1/186 (1%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	19/342 (6%)	1/168 (1%)	18/162 (11%)	0/12 (0%)
Overall	494/7723 (6%)	9/4269 (0%)	398/2712 (15%)	87/742 (12%)

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)

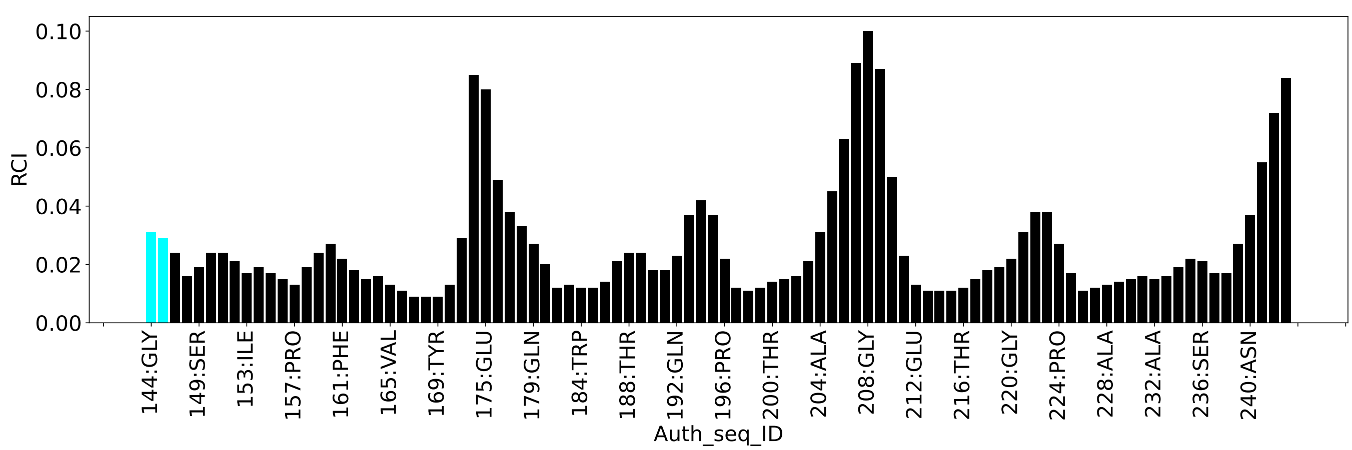
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	G	229	ARG	NE	141.70	76.53 – 92.65	35.4
1	G	158	LYS	H	4.70	5.24 – 11.12	-5.9

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 G:



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	2427
Intra-residue (|i-j|=0)	834
Sequential (|i-j|=1)	284
Medium range (|i-j|>1 and |i-j|<5)	488
Long range (|i-j|≥5)	414
Inter-chain	407
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	1116
Number of unmapped restraints	0
Number of restraints per residue	5.8
Number of long range restraints per residue[1]	0.7

[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)	22.2	0.2
0.2-0.5 (Medium)	42.0	0.5
>0.5 (Large)	20.4	1.49

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)	188.3	10.0
10.0-20.0 (Medium)	19.8	19.86
>20.0 (Large)	2.2	28.71

9 Distance violation analysis (i)

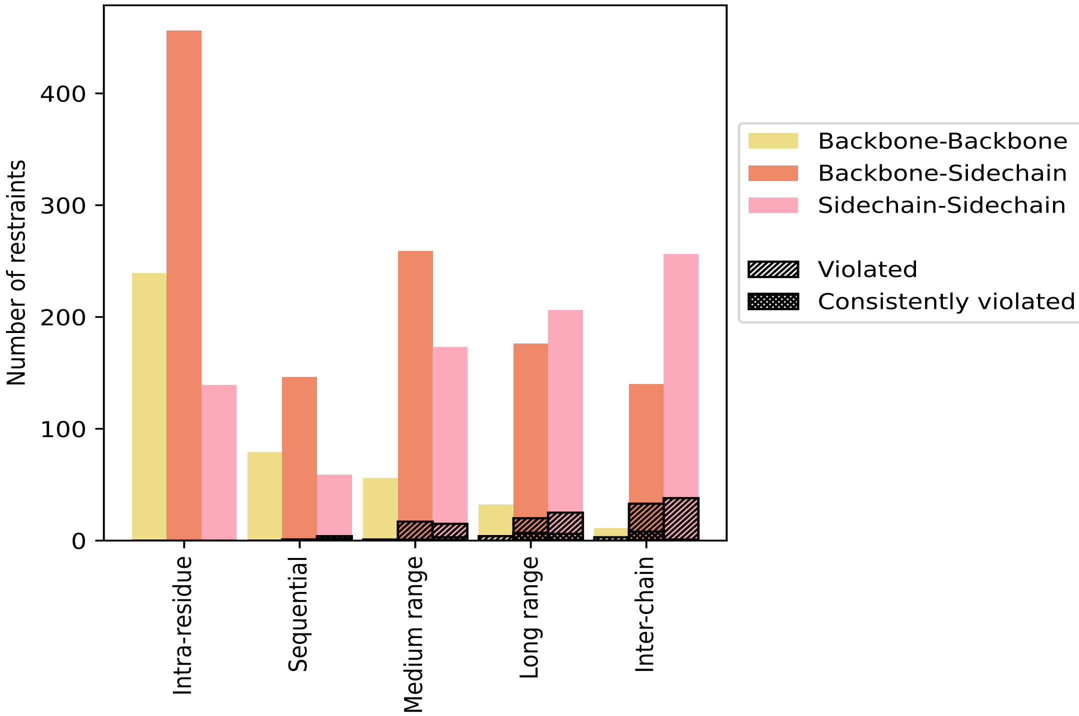
9.1 Summary of distance violations (i)

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)	834	34.4	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	239	9.8	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	456	18.8	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	139	5.7	0	0.0	0.0	0	0.0	0.0
Sequential (|i-j|=1)	284	11.7	5	1.8	0.2	3	1.1	0.1
Backbone-Backbone	79	3.3	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	146	6.0	1	0.7	0.0	1	0.7	0.0
Sidechain-Sidechain	59	2.4	4	6.8	0.2	2	3.4	0.1
Medium range (|i-j|>1 & |i-j|<5)	488	20.1	33	6.8	1.4	4	0.8	0.2
Backbone-Backbone	56	2.3	1	1.8	0.0	0	0.0	0.0
Backbone-Sidechain	259	10.7	17	6.6	0.7	1	0.4	0.0
Sidechain-Sidechain	173	7.1	15	8.7	0.6	3	1.7	0.1
Long range (|i-j|≥5)	414	17.1	49	11.8	2.0	13	3.1	0.5
Backbone-Backbone	32	1.3	4	12.5	0.2	0	0.0	0.0
Backbone-Sidechain	176	7.3	20	11.4	0.8	7	4.0	0.3
Sidechain-Sidechain	206	8.5	25	12.1	1.0	6	2.9	0.2
Inter-chain	407	16.8	74	18.2	3.0	9	2.2	0.4
Backbone-Backbone	11	0.5	3	27.3	0.1	0	0.0	0.0
Backbone-Sidechain	140	5.8	33	23.6	1.4	8	5.7	0.3
Sidechain-Sidechain	256	10.5	38	14.8	1.6	1	0.4	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	2427	100.0	161	6.6	6.6	29	1.2	1.2
Backbone-Backbone	417	17.2	8	1.9	0.3	0	0.0	0.0
Backbone-Sidechain	1177	48.5	71	6.0	2.9	17	1.4	0.7
Sidechain-Sidechain	833	34.3	82	9.8	3.4	12	1.4	0.5

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

9.1.1 Bar chart : Distribution of distance restraints and violations (i)



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

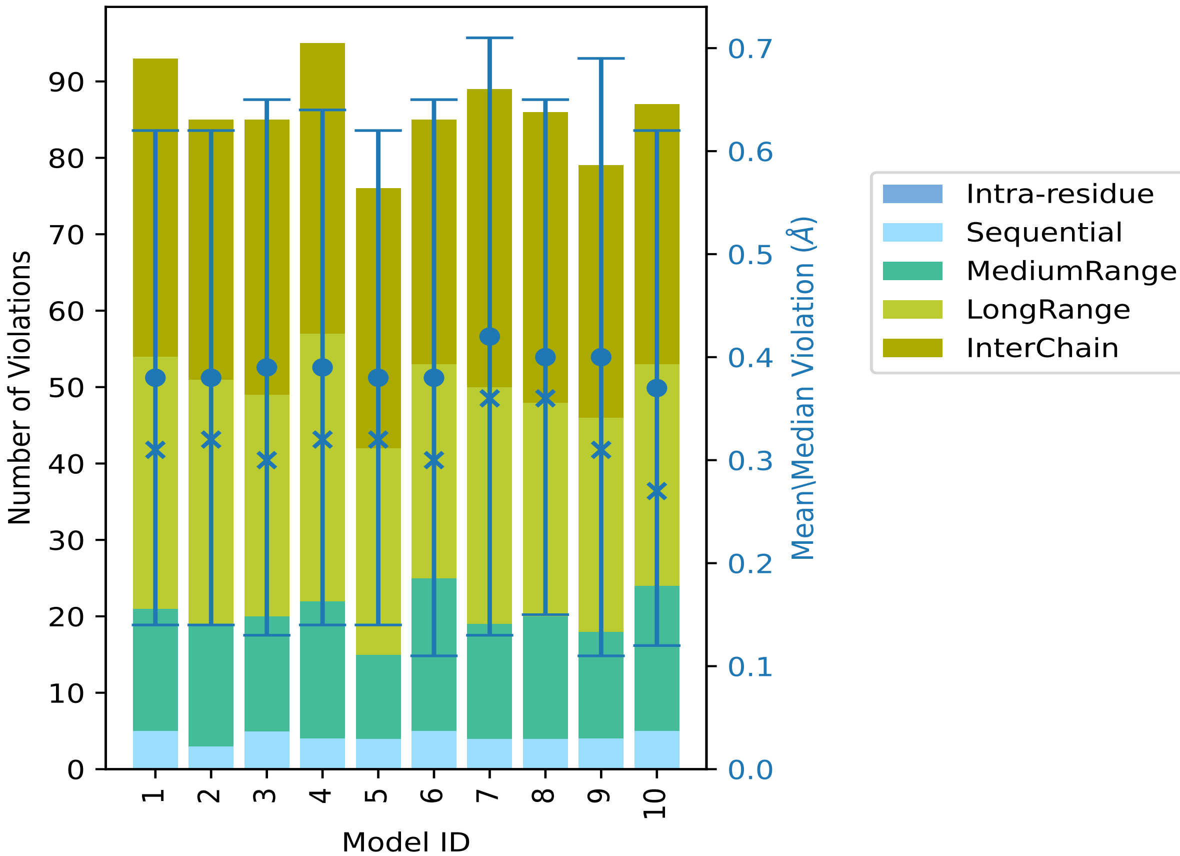
9.2 Distance violation statistics for each model (i)

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

Model ID	Number of violations						Mean (Å)	Max (Å)	SD[6] (Å)	Median (Å)
	IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total				
1	0	5	16	33	39	93	0.38	1.19	0.24	0.31
2	0	3	16	32	34	85	0.38	1.25	0.24	0.32
3	0	5	15	29	36	85	0.39	1.24	0.26	0.3
4	0	4	18	35	38	95	0.39	1.24	0.25	0.32
5	0	4	11	27	34	76	0.38	1.25	0.24	0.32
6	0	5	20	28	32	85	0.38	1.38	0.27	0.3
7	0	4	15	31	39	89	0.42	1.32	0.29	0.36
8	0	4	16	28	38	86	0.4	1.33	0.25	0.36
9	0	4	14	28	33	79	0.4	1.49	0.29	0.31
10	0	5	19	29	34	87	0.37	1.2	0.25	0.27

[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, 2266(IR:834, SQ:279, MR:455, LR:365, IC:333) 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	7	5	17	29	1	10.0
0	0	3	3	9	15	2	20.0
0	0	4	6	6	16	3	30.0

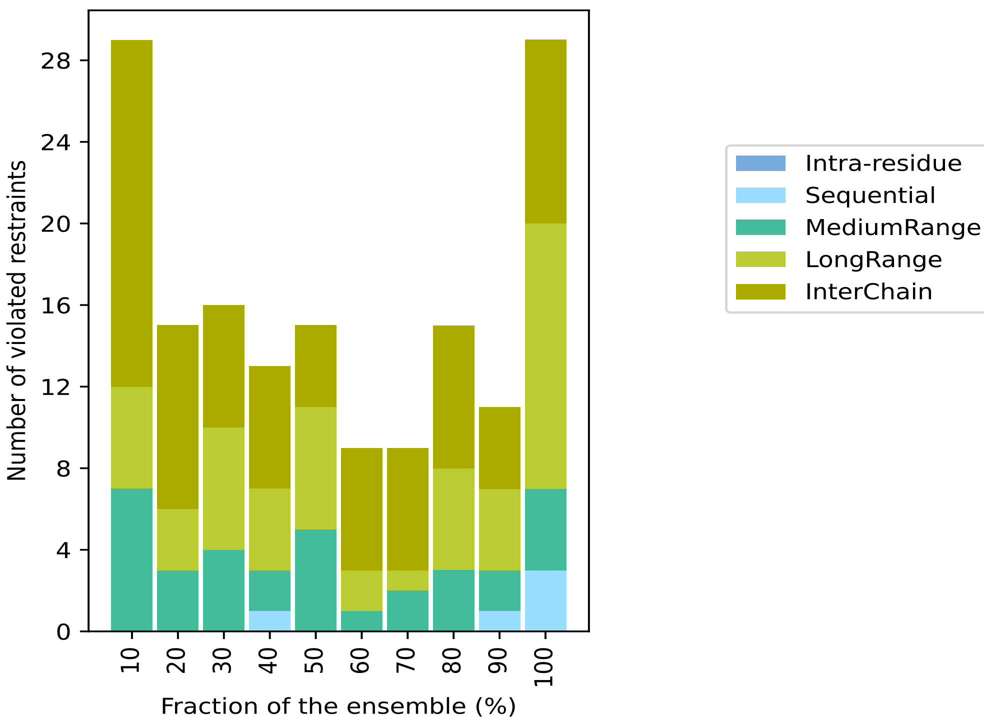
Continued on next page...

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	1	2	4	6	13	4	40.0
0	0	5	6	4	15	5	50.0
0	0	1	2	6	9	6	60.0
0	0	2	1	6	9	7	70.0
0	0	3	5	7	15	8	80.0
0	1	2	4	4	11	9	90.0
0	3	4	13	9	29	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

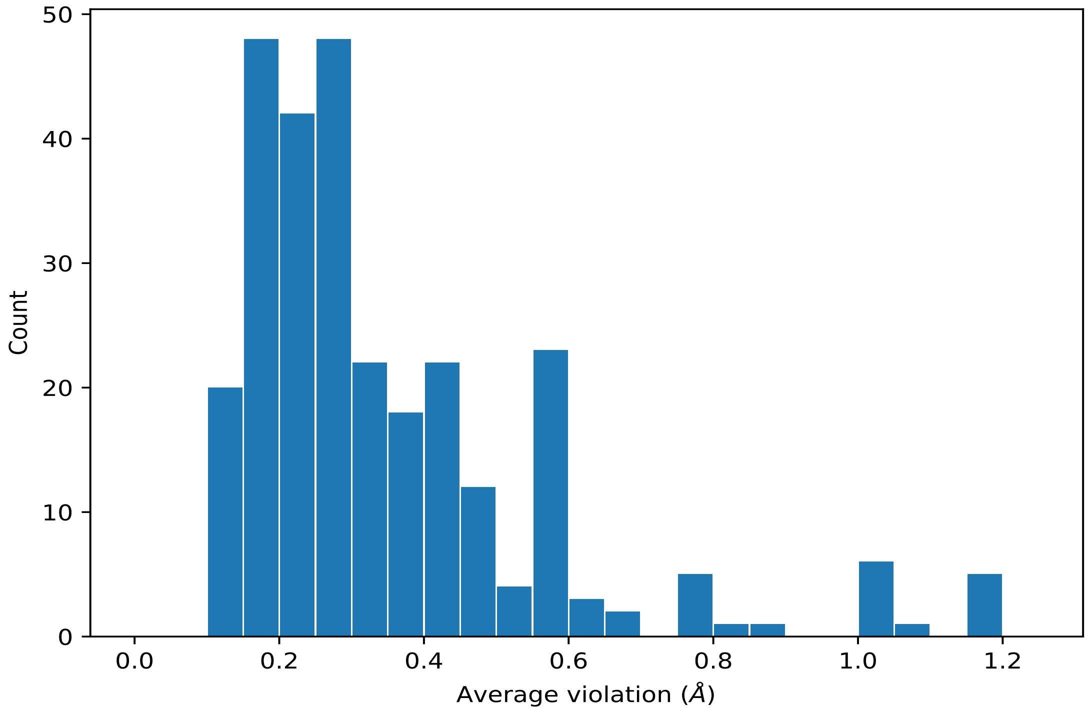
9.3.1 Bar graph : Distance violation statistics for the ensemble (i)



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)
(1,1096)	1:171:H:THR:CG2	1:152:H:ASP:CB	10	1.16	0.23	1.22
(1,1096)	1:171:J:THR:CG2	1:152:J:ASP:CB	10	1.16	0.23	1.22
(1,1096)	1:171:L:THR:CG2	1:152:L:ASP:CB	10	1.16	0.23	1.22
(1,1096)	1:171:K:THR:CG2	1:152:K:ASP:CB	10	1.16	0.23	1.22
(1,1096)	1:171:I:THR:CG2	1:152:I:ASP:CB	10	1.16	0.23	1.22
(1,2172)	2:301:H:IHP:P1	1:158:H:LYS:HA	10	1.05	0.09	1.03
(1,833)	1:221:J:VAL:CG2	1:230:J:VAL:CG2	10	1.03	0.24	1.07
(1,833)	1:221:I:VAL:CG2	1:230:I:VAL:CG1	10	1.03	0.24	1.07
(1,833)	1:221:G:VAL:CG2	1:230:G:VAL:CG2	10	1.03	0.24	1.07
(1,833)	1:221:I:VAL:CG2	1:230:I:VAL:CG2	10	1.03	0.24	1.07
(1,833)	1:221:H:VAL:CG2	1:230:H:VAL:CG1	10	1.03	0.24	1.07
(1,833)	1:221:G:VAL:CG2	1:230:G:VAL:CG1	10	1.03	0.24	1.07
(1,2118)	2:301:H:IHP:P4	1:158:J:LYS:HA	10	0.89	0.24	0.96
(1,2207)	2:301:H:IHP:O35	3:302:H:A1CCZ:F32	10	0.85	0.15	0.86
(1,2174)	2:301:H:IHP:P2	1:158:L:LYS:HA	10	0.78	0.36	0.74

Continued on next page...

Continued from previous page...

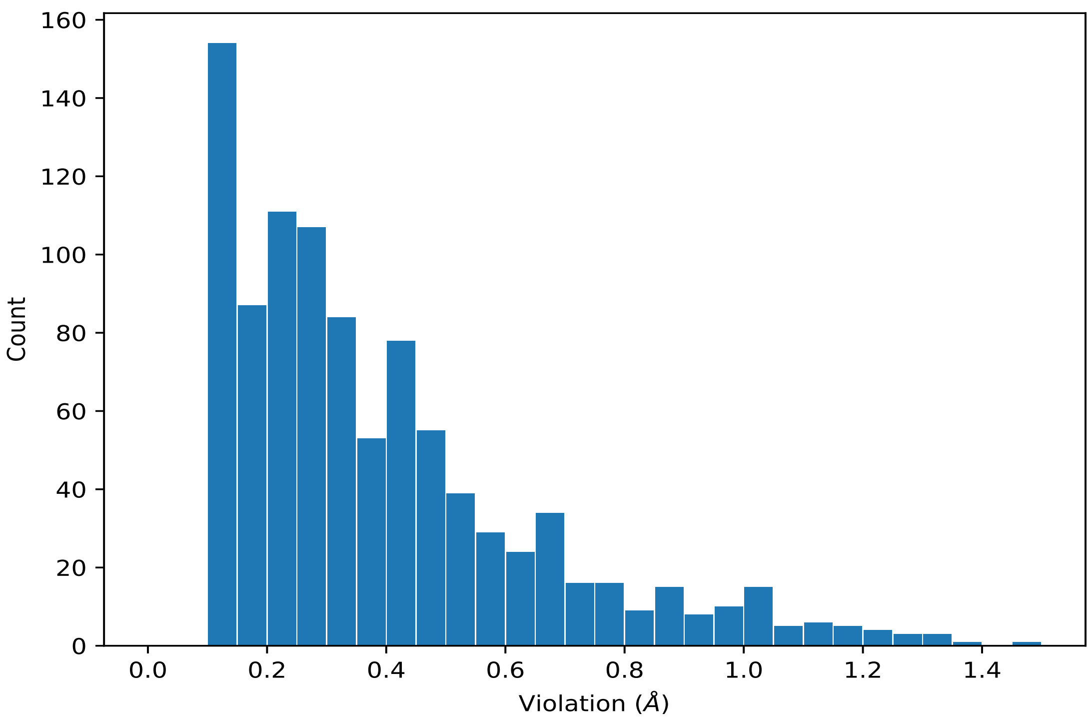
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,2115)	2:301:H:IHP:P4	1:158:G:LYS:HA	10	0.68	0.21	0.68
(1,29)	1:219:L:GLN:CG	1:155:K:GLN:CA	10	0.63	0.2	0.58
(1,2262)	3:302:H:A1CCZ:H7A	1:232:H:ALA:CA	10	0.59	0.21	0.68
(1,2262)	3:302:H:A1CCZ:H7B	1:232:H:ALA:CA	10	0.59	0.21	0.68
(1,412)	1:202:L:LEU:CG	1:194:L:ALA:CA	10	0.57	0.2	0.55

[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)
(1,1096)	1:171:I:THR:CG2	1:152:I:ASP:CB	9	1.49
(1,2174)	2:301:H:IHP:P2	1:158:L:LYS:HA	6	1.38
(1,1096)	1:171:J:THR:CG2	1:152:J:ASP:CB	8	1.33
(1,833)	1:221:G:VAL:CG2	1:230:G:VAL:CG1	9	1.33
(1,1096)	1:171:I:THR:CG2	1:152:I:ASP:CB	7	1.32
(1,2174)	2:301:H:IHP:P2	1:158:L:LYS:HA	7	1.28
(1,2172)	2:301:H:IHP:P1	1:158:H:LYS:HA	5	1.25
(1,1096)	1:171:J:THR:CG2	1:152:J:ASP:CB	2	1.25
(1,1096)	1:171:L:THR:CG2	1:152:L:ASP:CB	3	1.24
(1,833)	1:221:G:VAL:CG2	1:230:G:VAL:CG2	4	1.24

10 Dihedral-angle violation analysis (i)

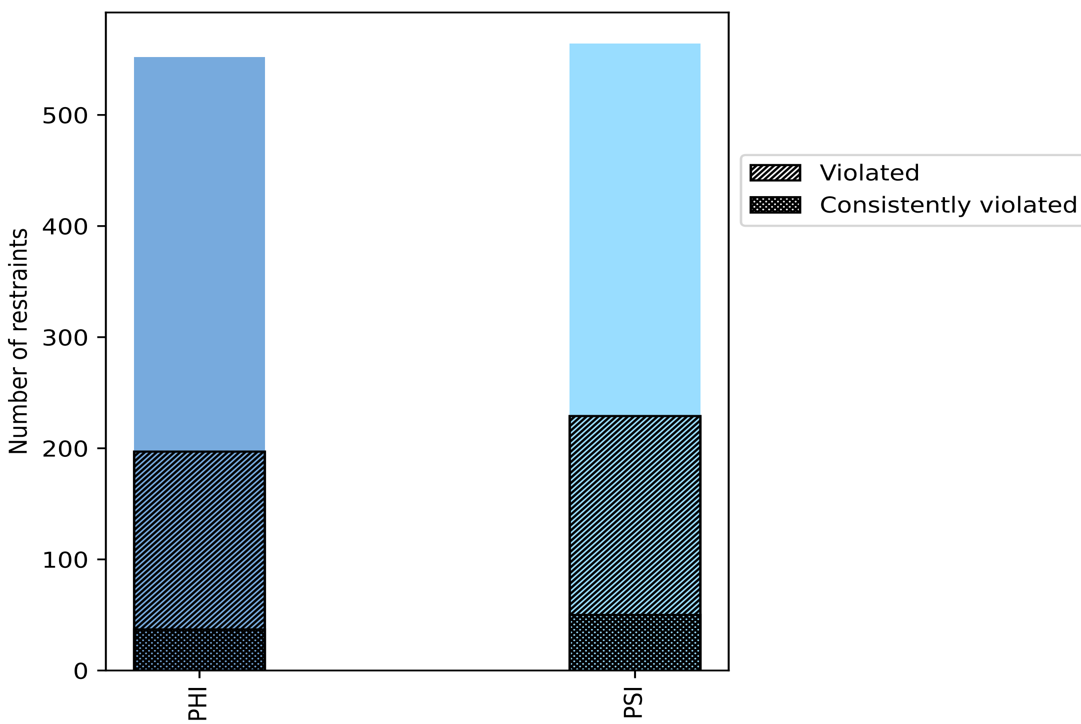
10.1 Summary of dihedral-angle violations (i)

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

Angle type	Count	%[1]	Violated[3]			Consistently Violated[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
PHI	552	49.5	197	35.7	17.7	37	6.7	3.3
PSI	564	50.5	229	40.6	20.5	50	8.9	4.5
Total	1116	100.0	426	38.2	38.2	87	7.8	7.8

[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)



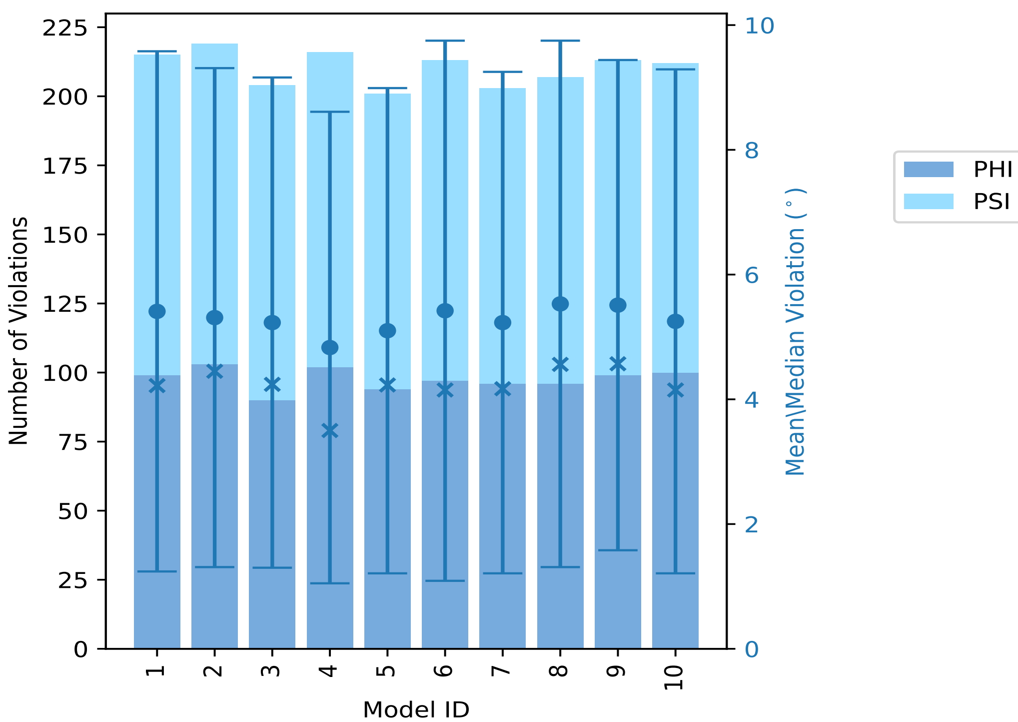
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

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 (°)
	PHI	PSI	Total				
1	99	116	215	5.41	28.06	4.17	4.22
2	103	116	219	5.31	21.78	4.0	4.45
3	90	114	204	5.23	28.32	3.93	4.24
4	102	114	216	4.83	27.69	3.78	3.5
5	94	107	201	5.1	21.91	3.89	4.23
6	97	116	213	5.42	26.76	4.33	4.15
7	96	107	203	5.23	27.99	4.02	4.17
8	96	111	207	5.53	28.43	4.22	4.56
9	99	114	213	5.51	22.44	3.93	4.57
10	100	112	212	5.25	28.71	4.04	4.15

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

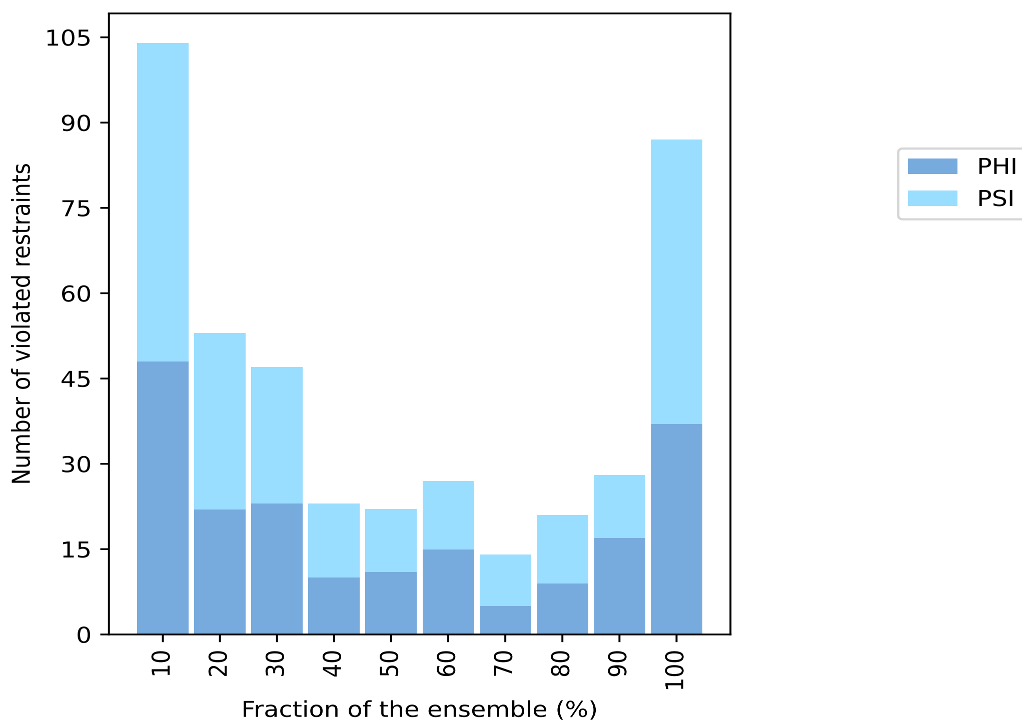
10.3 Dihedral-angle violation statistics for the ensemble (i)

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	
PHI	PSI	Total	Count[1]	%
48	56	104	1	10.0
22	31	53	2	20.0
23	24	47	3	30.0
10	13	23	4	40.0
11	11	22	5	50.0
15	12	27	6	60.0
5	9	14	7	70.0
9	12	21	8	80.0
17	11	28	9	90.0
37	50	87	10	100.0

[1] Number of models with violations

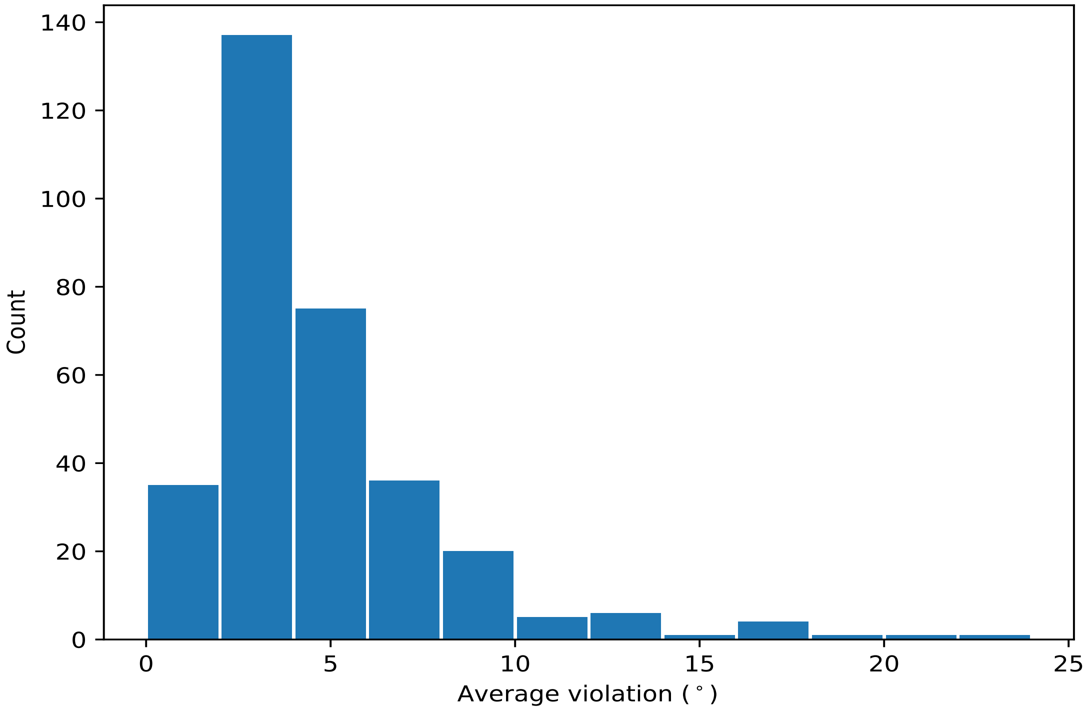
10.3.1 Bar graph : Dihedral-angle Violation statistics for the ensemble (i)



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.

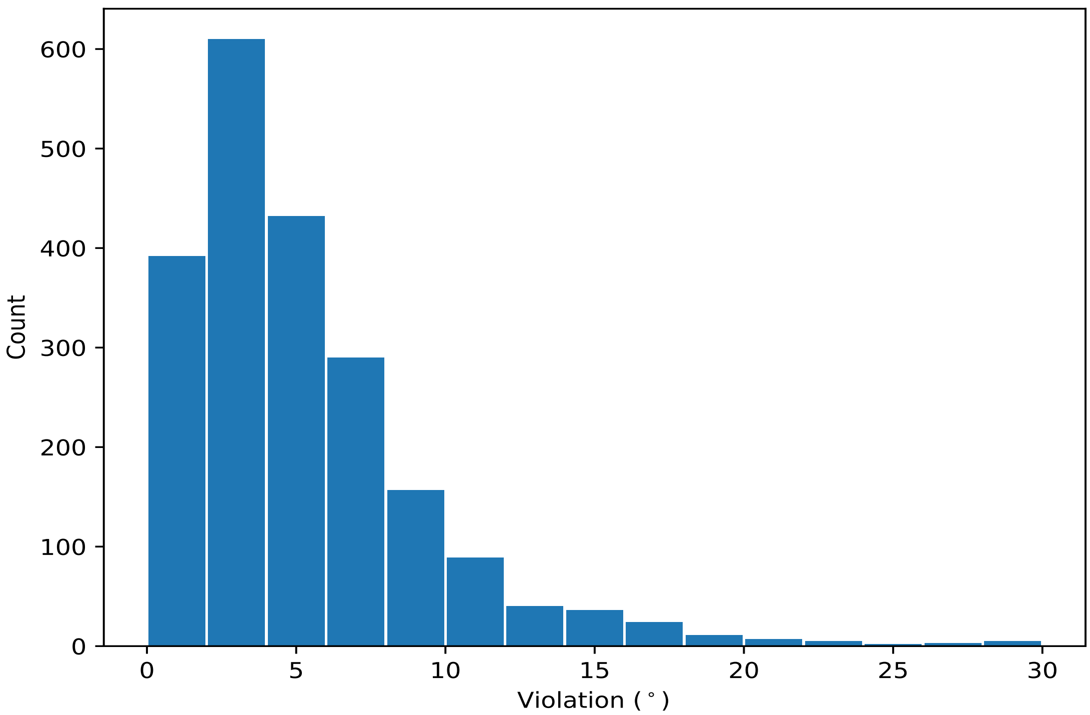
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models[1]	Mean	SD[2]	Median
(1,526)	1:195:J:ASN:N	1:195:J:ASN:CA	1:195:J:ASN:C	1:196:J:PRO:N	10	23.43	4.67	24.34
(1,525)	1:195:I:ASN:N	1:195:I:ASN:CA	1:195:I:ASN:C	1:196:I:PRO:N	10	20.64	5.45	21.29
(1,528)	1:195:L:ASN:N	1:195:L:ASN:CA	1:195:L:ASN:C	1:196:L:PRO:N	10	17.69	2.63	16.67
(1,523)	1:195:G:ASN:N	1:195:G:ASN:CA	1:195:G:ASN:C	1:196:G:PRO:N	10	17.37	0.85	17.56
(1,524)	1:195:H:ASN:N	1:195:H:ASN:CA	1:195:H:ASN:C	1:196:H:PRO:N	10	13.18	3.13	13.54
(1,43)	1:149:G:SER:N	1:149:G:SER:CA	1:149:G:SER:C	1:150:G:ILE:N	10	12.83	1.55	13.0
(1,890)	1:228:H:ALA:C	1:229:H:ARG:N	1:229:H:ARG:CA	1:229:H:ARG:C	10	12.4	1.7	13.06
(1,527)	1:195:K:ASN:N	1:195:K:ASN:CA	1:195:K:ASN:C	1:196:K:PRO:N	10	12.14	4.05	10.38
(1,448)	1:187:J:GLU:C	1:188:J:THR:N	1:188:J:THR:CA	1:188:J:THR:C	10	10.9	3.49	10.0
(1,1045)	1:242:G:ALA:C	1:243:G:THR:N	1:243:G:THR:CA	1:243:G:THR:C	10	10.55	0.43	10.4

[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 violations for 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 (°)
(1,525)	1:195:I:ASN:N	1:195:I:ASN:CA	1:195:I:ASN:C	1:196:I:PRO:N	10	28.71
(1,526)	1:195:J:ASN:N	1:195:J:ASN:CA	1:195:J:ASN:C	1:196:J:PRO:N	8	28.43
(1,54)	1:150:L:ILE:C	1:151:L:LEU:N	1:151:L:LEU:CA	1:151:L:LEU:C	8	28.36
(1,526)	1:195:J:ASN:N	1:195:J:ASN:CA	1:195:J:ASN:C	1:196:J:PRO:N	3	28.32
(1,525)	1:195:I:ASN:N	1:195:I:ASN:CA	1:195:I:ASN:C	1:196:I:PRO:N	1	28.06
(1,526)	1:195:J:ASN:N	1:195:J:ASN:CA	1:195:J:ASN:C	1:196:J:PRO:N	7	27.99
(1,526)	1:195:J:ASN:N	1:195:J:ASN:CA	1:195:J:ASN:C	1:196:J:PRO:N	4	27.69
(1,526)	1:195:J:ASN:N	1:195:J:ASN:CA	1:195:J:ASN:C	1:196:J:PRO:N	6	26.76
(1,51)	1:150:I:ILE:C	1:151:I:LEU:N	1:151:I:LEU:CA	1:151:I:LEU:C	6	25.79
(1,65)	1:152:K:ASP:N	1:152:K:ASP:CA	1:152:K:ASP:C	1:153:K:ILE:N	10	25.44