# wwPDB NMR Structure Validation Summary Report (i)

Jun 3, 2023 – 07:08 AM EDT

| | | |
|---|---|---|
| PDB ID | : | 1IIO |
| BMRB ID | : | 4996 |
| Title | : | NMR-Based Structure of the Conserved Protein MTH865 from the Archea Methanobacterium thermoautotrophicum |
| Authors | : | Lee, G.M.; Edwards, A.M.; Arrowsmith, C.H.; McIntosh, L.P. |
| Deposited on | : | 2001-04-23 |

This is a wwPDB NMR Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

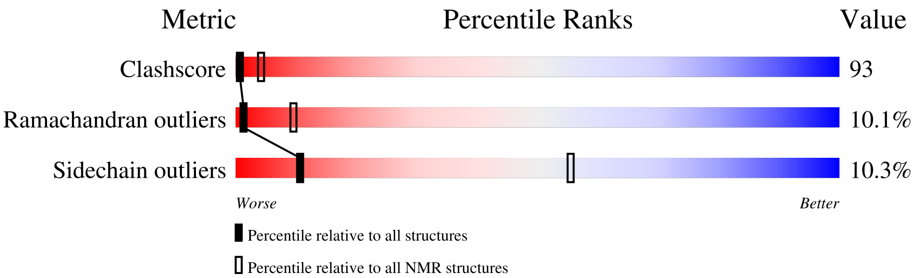
| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| wwPDB-RCI | : | v_1n_11_5_13_A (Berjanski et al., 2005) |
| PANAV | : | Wang et al. (2010) |
| wwPDB-ShiftChecker | : | v1.2 |
| BMRB Restraints Analysis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

# 1 Overall quality at a glance (i)

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

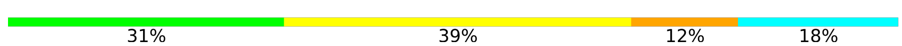
The overall completeness of chemical shifts assignment is 96%.

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 | 84 |  |

# 2 Ensemble composition and analysis (i)

This entry contains 25 models. Model 5 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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

| Well-defined (core) protein residues | | | |
|---|---|---|---|
| **Well-defined core** | **Residue range (total)** | **Backbone RMSD (Å)** | **Medoid model** |
| 1 | A:5-A:35, A:40-A:43, A:48-A:81 (69) | 0.32 | 5 |

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

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

The models can be grouped into 3 clusters and 7 single-model clusters were found.

| Cluster number | Models |
|---|---|
| 1 | 1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 13, 14 |
| 2 | 15, 17, 21, 22 |
| 3 | 18, 23 |
| Single-model clusters | 8; 12; 16; 19; 20; 24; 25 |

# 3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1204 atoms, of which 598 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called conserved hypothetical protein MTH865.

| Mol | Chain | Residues | Atoms | | | | | | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | H | N | O | S | |
| 1 | A | 84 | 1204 | 375 | 598 | 100 | 127 | 4 | 0 |

There are 3 discrepancies between the modelled and reference sequences:

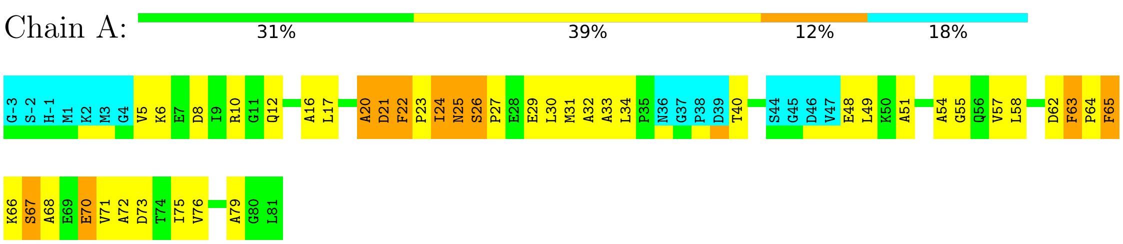
| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| A | -3 | GLY | - | expression tag | UNP Q04926 |
| A | -2 | SER | - | expression tag | UNP Q04926 |
| A | -1 | HIS | - | expression tag | UNP Q04926 |

# 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.

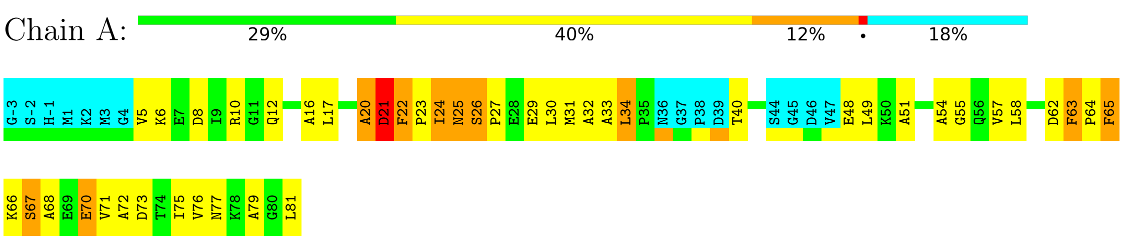
- Molecule 1: conserved hypothetical protein MTH865




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

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

- Molecule 1: conserved hypothetical protein MTH865



# 5 Refinement protocol and experimental data overview (i)

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

Of the 50 calculated structures, 25 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 |
|---|---|---|
| ARIA/CNS | structure solution | 1.0 |
| ARIA/CNS | refinement | 1.0 |

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 | 945 |
| Number of shifts mapped to atoms | 945 |
| Number of unparsed shifts | 0 |
| Number of shifts with mapping errors | 0 |
| Number of shifts with mapping warnings | 0 |
| Assignment completeness (well-defined parts) | 96% |

# 6 Model quality (i)

## 6.1 Standard geometry (i)

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

| Mol | Chain | Bond lengths | | Bond angles | |
|---|---|---|---|---|---|
| | | RMSZ | #Z>5 | RMSZ | #Z>5 |
| 1 | A | 0.34±0.06 | 0±0/511 ( 0.0± 0.1%) | 0.44±0.01 | 0±0/691 ( 0.0± 0.0%) |
| All | All | 0.35 | 6/12775 ( 0.0%) | 0.44 | 0/17275 ( 0.0%) |

All unique bond outliers are listed below.

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) | Models | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Worst | Total |
| 1 | A | 65 | PHE | CE2-CZ | 5.85 | 1.48 | 1.37 | 2 | 6 |

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## 6.2 Too-close contacts (i)

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

| Mol | Chain | Non-H | H(model) | H(added) | Clashes |
|---|---|---|---|---|---|
| 1 | A | 505 | 505 | 505 | 94±6 |
| All | All | 12625 | 12625 | 12625 | 2348 |

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 93.

5 of 285 unique clashes are listed below, sorted by their clash magnitude.

| Atom-1 | Atom-2 | Clash(Å) | Distance(Å) | Models | |
|---|---|---|---|---|---|
| | | | | Worst | Total |
| 1:A:24:ILE:HG21 | 1:A:71:VAL:HG21 | 1.12 | 1.19 | 1 | 25 |

*Continued on next page...*

*Continued from previous page...*

| Atom-1 | Atom-2 | Clash(Å) | Distance(Å) | Models | |
|---|---|---|---|---|---|
| | | | | Worst | Total |
| 1:A:17:LEU:HD22 | 1:A:24:ILE:HD12 | 0.96 | 1.35 | 9 | 25 |
| 1:A:24:ILE:HG23 | 1:A:30:LEU:HD22 | 0.94 | 1.39 | 16 | 25 |
| 1:A:24:ILE:HG21 | 1:A:71:VAL:CG2 | 0.91 | 1.96 | 2 | 25 |
| 1:A:17:LEU:HD22 | 1:A:24:ILE:CD1 | 0.90 | 1.96 | 20 | 25 |

## 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 | 68/84 (81%) | 51±2 (75±2%) | 10±2 (15±3%) | 7±1 (10±1%) | 1 | 9 |
| All | All | 1700/2100 (81%) | 1280 (75%) | 249 (15%) | 171 (10%) | 1 | 9 |

5 of 14 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 | 24 | ILE | 25 |
| 1 | A | 25 | ASN | 25 |
| 1 | A | 26 | SER | 25 |
| 1 | A | 21 | ASP | 24 |
| 1 | A | 20 | ALA | 23 |

### 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 | 53/64 (83%) | 48±1 (90±3%) | 5±1 (10±3%) | 11 | 55 |
| All | All | 1325/1600 (83%) | 1188 (90%) | 137 (10%) | 11 | 55 |

5 of 17 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 | 22 | PHE | 25 |
| 1 | A | 65 | PHE | 25 |
| 1 | A | 70 | GLU | 21 |
| 1 | A | 63 | PHE | 16 |
| 1 | A | 21 | ASP | 12 |

#### 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 96% for the well-defined parts and 91% for the entire structure.

## 7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

### 7.1.1 Bookkeeping (i)

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

| | |
|---|---|
| Total number of shifts | 945 |
| Number of shifts mapped to atoms | 945 |
| Number of unparsed shifts | 0 |
| Number of shifts with mapping errors | 0 |
| Number of shifts with mapping warnings | 0 |
| Number of shift outliers (ShiftChecker) | 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 |
|---|---|---|---|
| $^{13}C_{\alpha}$ | 80 | -0.33 ± 0.22 | None needed (< 0.5 ppm) |
| $^{13}C_{\beta}$ | 72 | -0.11 ± 0.10 | None needed (< 0.5 ppm) |
| $^{13}C'$ | 70 | -0.24 ± 0.18 | None needed (< 0.5 ppm) |
| $^{15}N$ | 75 | 0.69 ± 0.59 | None needed (imprecise) |

### 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 96%, i.e. 834 atoms were assigned a chemical shift out of a possible 873. 0 out of 10 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

| | Total | $^{1}H$ | $^{13}C$ | $^{15}N$ |
|---|---|---|---|---|
| Backbone | 333/342 (97%) | 139/139 (100%) | 129/138 (93%) | 65/65 (100%) |
| Sidechain | 471/501 (94%) | 324/328 (99%) | 142/161 (88%) | 5/12 (42%) |

*Continued on next page...*

*Continued from previous page...*

| | Total | $^{1}$H | $^{13}$C | $^{15}$N |
|---|---|---|---|---|
| Aromatic | 30/30 (100%) | 15/15 (100%) | 15/15 (100%) | 0/0 (—%) |
| Overall | 834/873 (96%) | 478/482 (99%) | 286/314 (91%) | 70/77 (91%) |

### 7.1.4 Statistically unusual chemical shifts

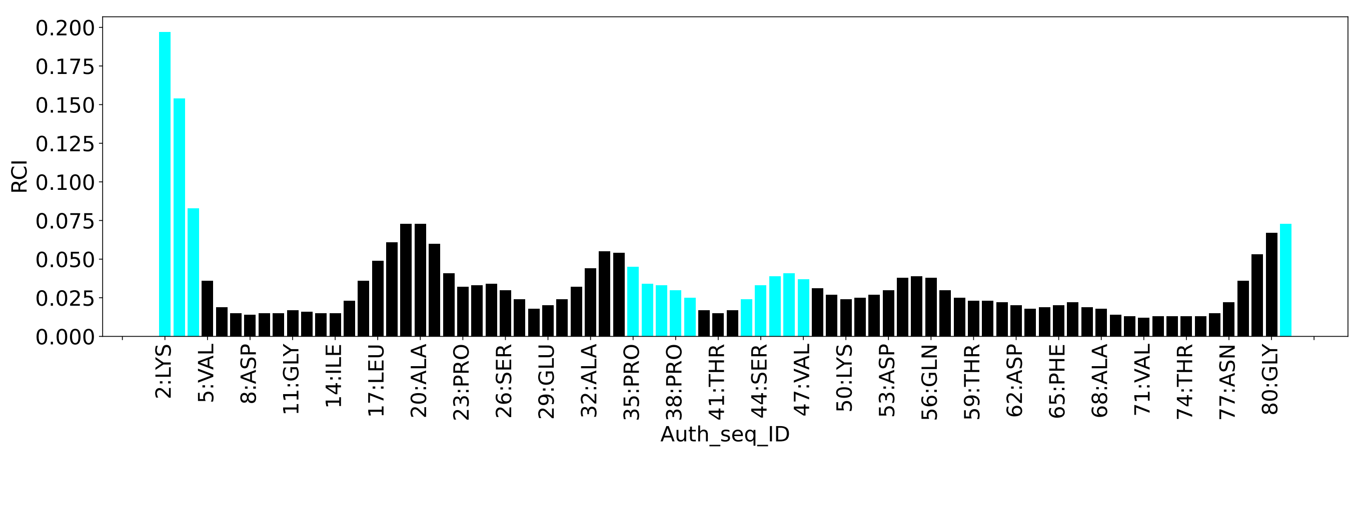
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

| List Id | Chain | Res | Type | Atom | Shift, *ppm* | Expected range, *ppm* | Z-score |
|---|---|---|---|---|---|---|---|
| 1 | A | 10 | ARG | NE | 113.83 | 76.53 – 92.65 | 18.1 |

### 7.1.5 Random Coil Index (RCI) plots

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

Random coil index (RCI) for chain A:



# 8 NMR restraints analysis (i)

## 8.1 Conformationally restricting restraints (i)

The following table provides the summary of experimentally observed NMR restraints in different categories. Restraints are classified into different categories based on the sequence separation of the atoms involved.

| Description | Value |
|---|---|
| Total distance restraints | 1915 |
| Intra-residue (\|i-j\|=0) | 379 |
| Sequential (\|i-j\|=1) | 267 |
| Medium range (\|i-j\|>1 and \|i-j\|<5) | 678 |
| Long range (\|i-j\|≥5) | 540 |
| Inter-chain | 0 |
| Hydrogen bond restraints | 51 |
| Disulfide bond restraints | 0 |
| Total dihedral-angle restraints | 102 |
| Number of unmapped restraints | 0 |
| Number of restraints per residue | 24.0 |
| Number of long range restraints per residue[1] | 6.4 |

[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) | 32.3 | 0.2 |
| 0.2-0.5 (Medium) | 62.7 | 0.5 |
| >0.5 (Large) | 659.0 | 18.4 |

### 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) | 10.9 | 9.9 |
| 10.0-20.0 (Medium) | 5.0 | 19.8 |
| >20.0 (Large) | 31.4 | 165.3 |

# 9 Distance violation analysis (i)

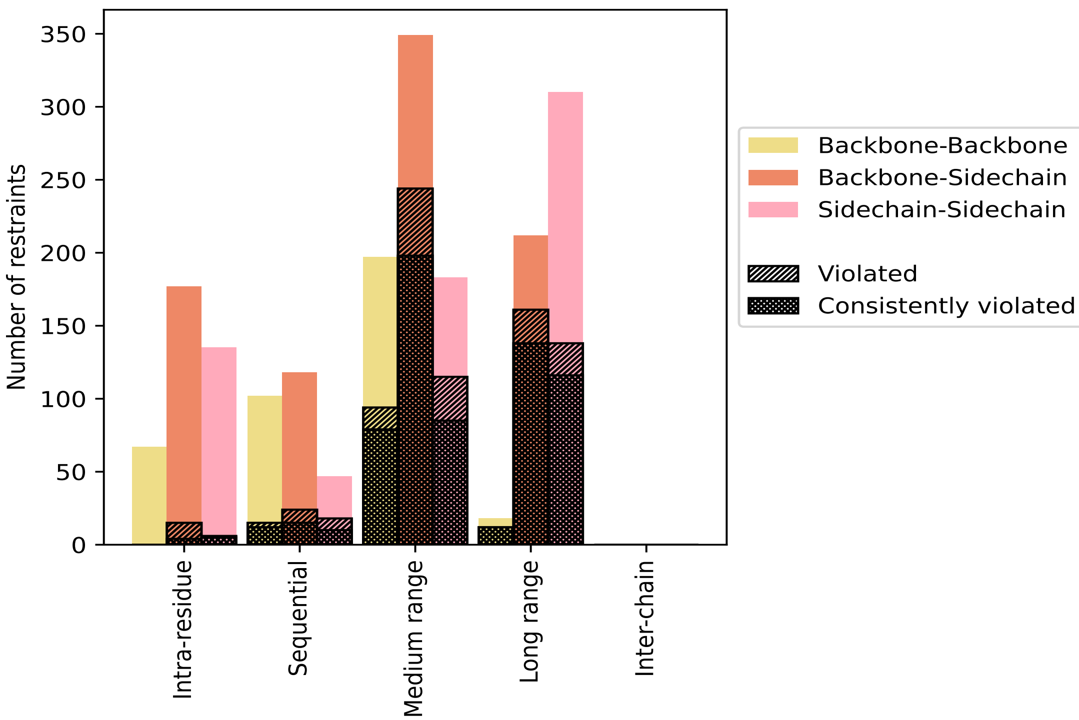
## 9.1 Summary of distance violations (i)

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) | 379 | 19.8 | 21 | 5.5 | 1.1 | 9 | 2.4 | 0.5 |
| Backbone-Backbone | 67 | 3.5 | 0 | 0.0 | 0.0 | 0 | 0.0 | 0.0 |
| Backbone-Sidechain | 177 | 9.2 | 15 | 8.5 | 0.8 | 4 | 2.3 | 0.2 |
| Sidechain-Sidechain | 135 | 7.0 | 6 | 4.4 | 0.3 | 5 | 3.7 | 0.3 |
| Sequential (\|i-j\|=1) | 267 | 13.9 | 57 | 21.3 | 3.0 | 37 | 13.9 | 1.9 |
| Backbone-Backbone | 102 | 5.3 | 15 | 14.7 | 0.8 | 12 | 11.8 | 0.6 |
| Backbone-Sidechain | 118 | 6.2 | 24 | 20.3 | 1.3 | 15 | 12.7 | 0.8 |
| Sidechain-Sidechain | 47 | 2.5 | 18 | 38.3 | 0.9 | 10 | 21.3 | 0.5 |
| Medium range (\|i-j\|>1 & \|i-j\|<5) | 678 | 35.4 | 431 | 63.6 | 22.5 | 343 | 50.6 | 17.9 |
| Backbone-Backbone | 146 | 7.6 | 72 | 49.3 | 3.8 | 60 | 41.1 | 3.1 |
| Backbone-Sidechain | 349 | 18.2 | 244 | 69.9 | 12.7 | 198 | 56.7 | 10.3 |
| Sidechain-Sidechain | 183 | 9.6 | 115 | 62.8 | 6.0 | 85 | 46.4 | 4.4 |
| Long range (\|i-j\|≥5) | 540 | 28.2 | 311 | 57.6 | 16.2 | 266 | 49.3 | 13.9 |
| Backbone-Backbone | 18 | 0.9 | 12 | 66.7 | 0.6 | 12 | 66.7 | 0.6 |
| Backbone-Sidechain | 212 | 11.1 | 161 | 75.9 | 8.4 | 138 | 65.1 | 7.2 |
| Sidechain-Sidechain | 310 | 16.2 | 138 | 44.5 | 7.2 | 116 | 37.4 | 6.1 |
| 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 | 51 | 2.7 | 22 | 43.1 | 1.1 | 19 | 37.3 | 1.0 |
| Disulfide bond | 0 | 0.0 | 0 | 0.0 | 0.0 | 0 | 0.0 | 0.0 |
| Total | 1915 | 100.0 | 842 | 44.0 | 44.0 | 674 | 35.2 | 35.2 |
| Backbone-Backbone | 384 | 20.1 | 121 | 31.5 | 6.3 | 103 | 26.8 | 5.4 |
| Backbone-Sidechain | 856 | 44.7 | 444 | 51.9 | 23.2 | 355 | 41.5 | 18.5 |
| Sidechain-Sidechain | 675 | 35.2 | 277 | 41.0 | 14.5 | 216 | 32.0 | 11.3 |

[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 | 13 | 49 | 398 | 289 | 0 | 749 | 2.8 | 14.92 | 2.31 | 2.29 |
| 2 | 13 | 46 | 410 | 284 | 0 | 753 | 2.8 | 17.45 | 2.36 | 2.17 |
| 3 | 12 | 47 | 402 | 286 | 0 | 747 | 2.81 | 15.57 | 2.35 | 2.27 |
| 4 | 14 | 49 | 399 | 290 | 0 | 752 | 2.82 | 18.4 | 2.42 | 2.27 |
| 5 | 12 | 45 | 404 | 286 | 0 | 747 | 2.79 | 16.7 | 2.37 | 2.14 |
| 6 | 14 | 49 | 406 | 288 | 0 | 757 | 2.81 | 14.87 | 2.33 | 2.27 |
| 7 | 15 | 48 | 399 | 286 | 0 | 748 | 2.82 | 17.49 | 2.36 | 2.28 |
| 8 | 13 | 52 | 397 | 287 | 0 | 749 | 2.83 | 15.02 | 2.38 | 2.25 |
| 9 | 13 | 48 | 402 | 288 | 0 | 751 | 2.77 | 16.08 | 2.34 | 2.22 |
| 10 | 16 | 47 | 406 | 293 | 0 | 762 | 2.76 | 16.16 | 2.36 | 2.15 |
| 11 | 13 | 43 | 404 | 293 | 0 | 753 | 2.79 | 16.76 | 2.39 | 2.14 |

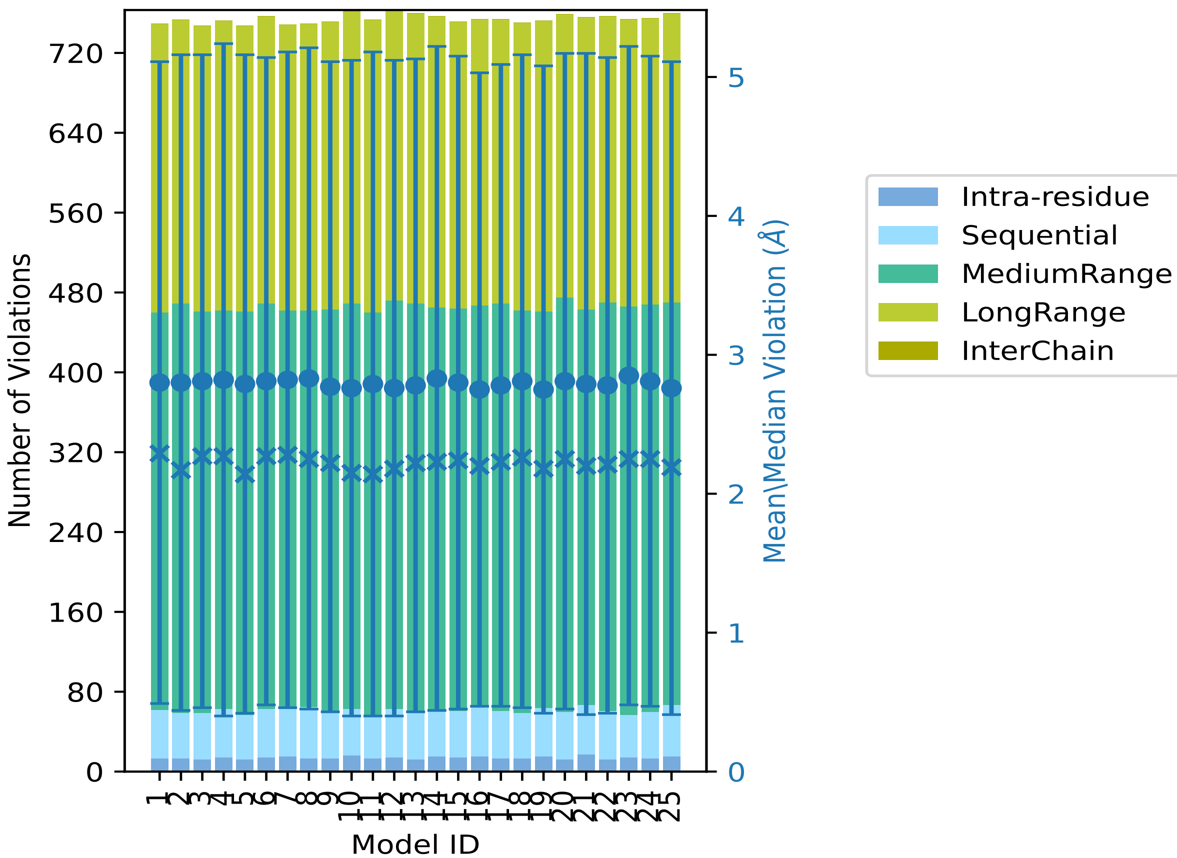
*Continued on next page...*

*Continued from previous page...*

| Model ID | Number of violations | | | | | | Mean (Å) | Max (Å) | SD[6] (Å) | Median (Å) |
|---|---|---|---|---|---|---|---|---|---|---|
| | IR[1] | SQ[2] | MR[3] | LR[4] | IC[5] | Total | | | | |
| 12 | 14 | 49 | 409 | 291 | 0 | 763 | 2.76 | 15.44 | 2.36 | 2.18 |
| 13 | 12 | 48 | 409 | 291 | 0 | 760 | 2.78 | 15.21 | 2.35 | 2.22 |
| 14 | 15 | 47 | 403 | 292 | 0 | 757 | 2.83 | 16.64 | 2.39 | 2.23 |
| 15 | 14 | 47 | 403 | 287 | 0 | 751 | 2.8 | 14.61 | 2.35 | 2.24 |
| 16 | 15 | 50 | 402 | 287 | 0 | 754 | 2.75 | 14.95 | 2.28 | 2.2 |
| 17 | 13 | 48 | 408 | 285 | 0 | 754 | 2.78 | 14.6 | 2.31 | 2.23 |
| 18 | 13 | 46 | 403 | 288 | 0 | 750 | 2.81 | 14.96 | 2.35 | 2.26 |
| 19 | 15 | 49 | 397 | 291 | 0 | 752 | 2.75 | 15.92 | 2.33 | 2.18 |
| 20 | 12 | 48 | 415 | 284 | 0 | 759 | 2.81 | 15.3 | 2.36 | 2.25 |
| 21 | 17 | 50 | 396 | 293 | 0 | 756 | 2.79 | 16.71 | 2.38 | 2.2 |
| 22 | 12 | 49 | 409 | 287 | 0 | 757 | 2.78 | 16.18 | 2.36 | 2.21 |
| 23 | 14 | 43 | 409 | 288 | 0 | 754 | 2.85 | 16.73 | 2.37 | 2.25 |
| 24 | 13 | 47 | 408 | 287 | 0 | 755 | 2.81 | 16.22 | 2.34 | 2.25 |
| 25 | 15 | 52 | 403 | 290 | 0 | 760 | 2.76 | 15.5 | 2.35 | 2.19 |

[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, 1044(IR:358, SQ:210, MR:247, LR:229, 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] | % |
| 1 | 1 | 10 | 6 | 0 | 18 | 1 | 4.0 |
| 1 | 1 | 3 | 4 | 0 | 9 | 2 | 8.0 |
| 0 | 0 | 4 | 1 | 0 | 5 | 3 | 12.0 |
| 1 | 0 | 5 | 2 | 0 | 8 | 4 | 16.0 |
| 1 | 0 | 4 | 1 | 0 | 6 | 5 | 20.0 |
| 1 | 1 | 3 | 0 | 0 | 5 | 6 | 24.0 |

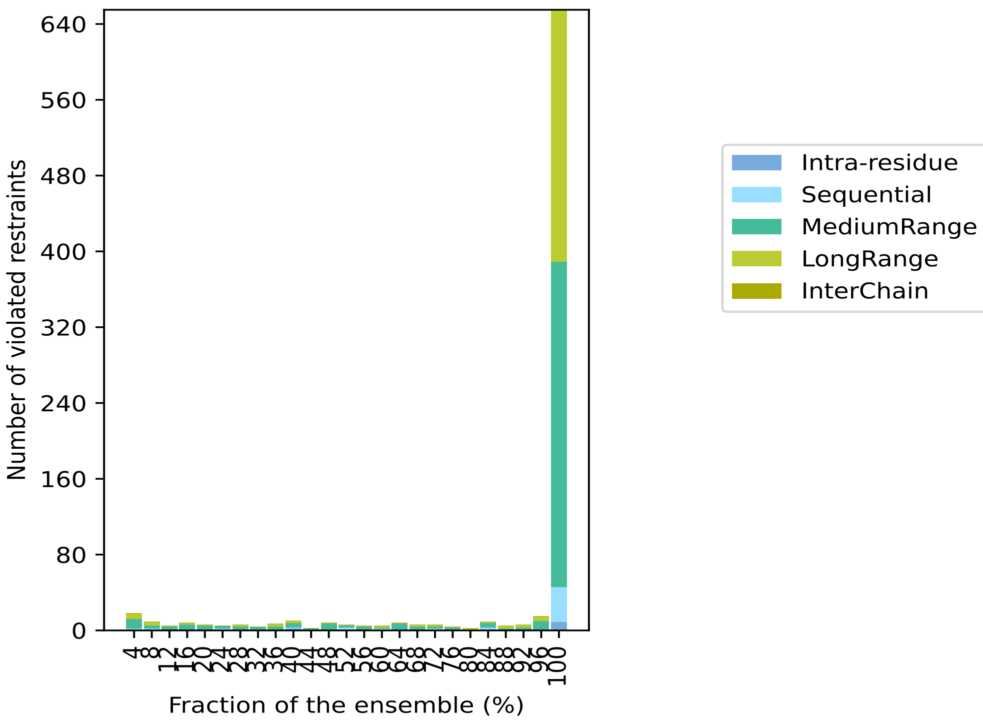
*Continued on next page...*

*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 | 3 | 2 | 0 | 6 | 7 | 28.0 |
| 0 | 1 | 3 | 0 | 0 | 4 | 8 | 32.0 |
| 0 | 1 | 3 | 3 | 0 | 7 | 9 | 36.0 |
| 2 | 1 | 5 | 2 | 0 | 10 | 10 | 40.0 |
| 0 | 1 | 1 | 0 | 0 | 2 | 11 | 44.0 |
| 1 | 0 | 6 | 1 | 0 | 8 | 12 | 48.0 |
| 1 | 2 | 2 | 1 | 0 | 6 | 13 | 52.0 |
| 0 | 1 | 3 | 1 | 0 | 5 | 14 | 56.0 |
| 1 | 1 | 1 | 2 | 0 | 5 | 15 | 60.0 |
| 0 | 1 | 6 | 1 | 0 | 8 | 16 | 64.0 |
| 0 | 1 | 3 | 2 | 0 | 6 | 17 | 68.0 |
| 1 | 1 | 2 | 2 | 0 | 6 | 18 | 72.0 |
| 0 | 0 | 3 | 1 | 0 | 4 | 19 | 76.0 |
| 0 | 0 | 1 | 1 | 0 | 2 | 20 | 80.0 |
| 1 | 2 | 5 | 1 | 0 | 9 | 21 | 84.0 |
| 0 | 1 | 1 | 3 | 0 | 5 | 22 | 88.0 |
| 0 | 1 | 2 | 3 | 0 | 6 | 23 | 92.0 |
| 0 | 1 | 9 | 5 | 0 | 15 | 24 | 96.0 |
| 9 | 37 | 343 | 266 | 0 | 655 | 25 | 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

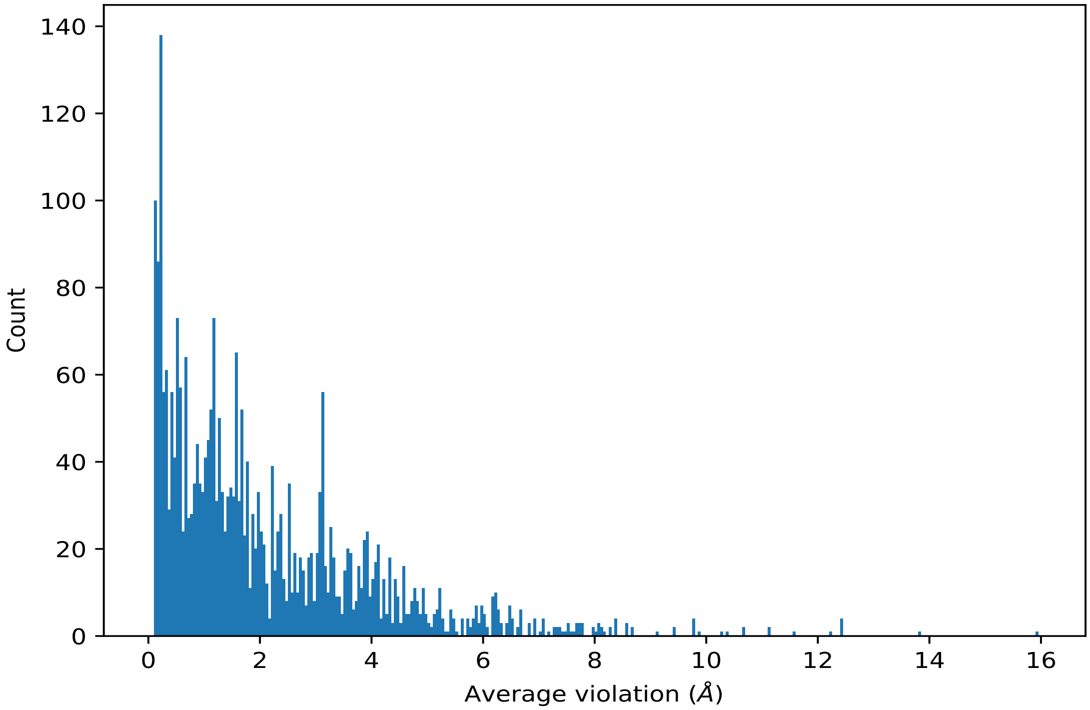
### 9.3.1 Bar graph : Distance violation statistics for the ensemble (i)




## 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.

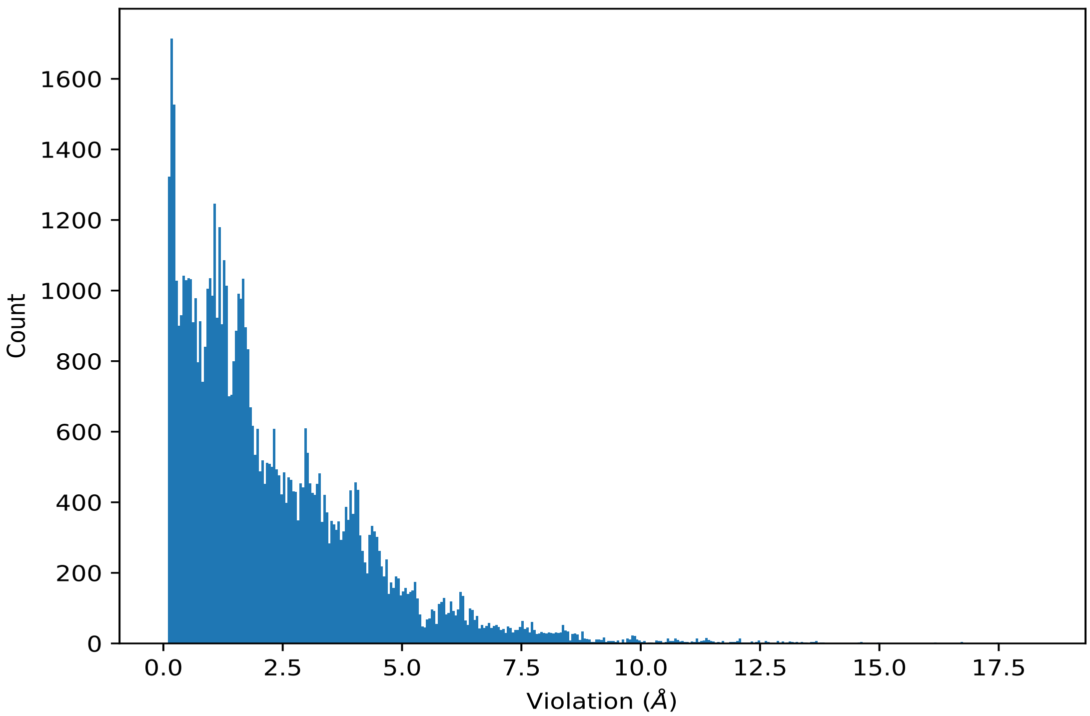
| Key | Atom-1 | Atom-2 | Models[1] | Mean (Å) | SD[1] (Å) | Median (Å) |
|---|---|---|---|---|---|---|
| (1,19) | 1:A:38:PRO:HA | 1:A:46:ASP:H | 25 | 15.94 | 0.98 | 15.92 |
| (1,800) | 1:A:39:ASP:HB3 | 1:A:6:LYS:HA | 25 | 13.85 | 0.48 | 13.78 |
| (1,1101) | 1:A:36:ASN:HB3 | 1:A:28:GLU:HG2 | 25 | 12.42 | 0.84 | 12.46 |
| (1,1101) | 1:A:36:ASN:HB3 | 1:A:28:GLU:HG3 | 25 | 12.42 | 0.84 | 12.46 |
| (1,1541) | 1:A:46:ASP:HB3 | 1:A:39:ASP:HB2 | 25 | 12.42 | 0.63 | 12.26 |
| (1,1541) | 1:A:46:ASP:HB2 | 1:A:39:ASP:HB2 | 25 | 12.42 | 0.63 | 12.26 |
| (1,21) | 1:A:39:ASP:HB2 | 1:A:46:ASP:H | 25 | 12.21 | 0.79 | 12.31 |
| (1,90) | 1:A:39:ASP:HB3 | 1:A:9:ILE:H | 25 | 11.59 | 0.52 | 11.45 |
| (1,913) | 1:A:40:THR:HA | 1:A:45:GLY:HA2 | 25 | 11.13 | 0.54 | 11.27 |
| (1,913) | 1:A:40:THR:HA | 1:A:45:GLY:HA3 | 25 | 11.13 | 0.54 | 11.27 |
| (1,285) | 1:A:21:ASP:HA | 1:A:62:ASP:H | 25 | 10.69 | 0.08 | 10.7 |
| (1,891) | 1:A:18:ALA:HB3 | 1:A:26:SER:HB2 | 25 | 10.67 | 0.25 | 10.65 |
| (1,18) | 1:A:40:THR:HA | 1:A:46:ASP:H | 25 | 10.36 | 0.91 | 9.93 |

[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) |
|---|---|---|---|---|
| (1,19) | 1:A:38:PRO:HA | 1:A:46:ASP:H | 4 | 18.4 |
| (1,19) | 1:A:38:PRO:HA | 1:A:46:ASP:H | 7 | 17.49 |
| (1,19) | 1:A:38:PRO:HA | 1:A:46:ASP:H | 2 | 17.45 |
| (1,19) | 1:A:38:PRO:HA | 1:A:46:ASP:H | 11 | 16.76 |
| (1,19) | 1:A:38:PRO:HA | 1:A:46:ASP:H | 23 | 16.73 |
| (1,19) | 1:A:38:PRO:HA | 1:A:46:ASP:H | 21 | 16.71 |
| (1,19) | 1:A:38:PRO:HA | 1:A:46:ASP:H | 5 | 16.7 |
| (1,19) | 1:A:38:PRO:HA | 1:A:46:ASP:H | 14 | 16.64 |
| (1,19) | 1:A:38:PRO:HA | 1:A:46:ASP:H | 24 | 16.22 |
| (1,19) | 1:A:38:PRO:HA | 1:A:46:ASP:H | 22 | 16.18 |

# 10 Dihedral-angle violation analysis (i)

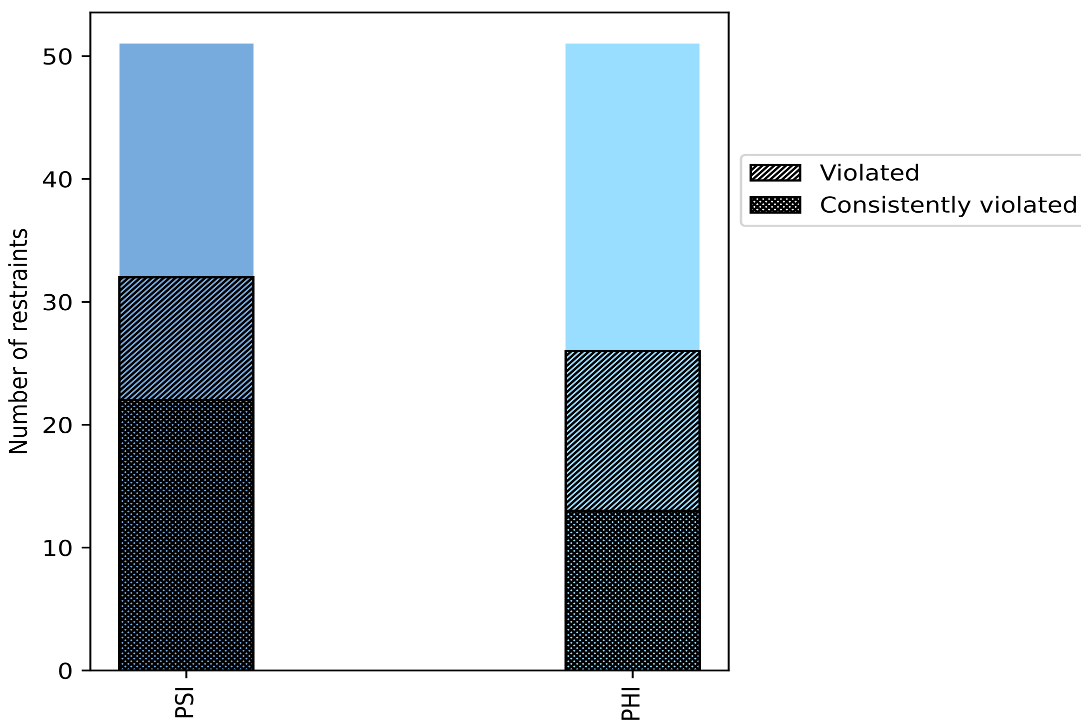
## 10.1 Summary of dihedral-angle violations (i)

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

| Angle type | Count | %[1] | Violated[3] | | | Consistently Violated[4] | | |
|---|---|---|---|---|---|---|---|---|
| | | | Count | %[2] | %[1] | Count | %[2] | %[1] |
| PSI | 51 | 50.0 | 32 | 62.7 | 31.4 | 22 | 43.1 | 21.6 |
| PHI | 51 | 50.0 | 26 | 51.0 | 25.5 | 13 | 25.5 | 12.7 |
| Total | 102 | 100.0 | 58 | 56.9 | 56.9 | 35 | 34.3 | 34.3 |

[1] percentage calculated with respect to total number of dihedral-angle restraints, [2] percentage calculated with respect to number of restraints in a particular dihedral-angle type, [3] violated in at least one model, [4] violated in all the models

### 10.1.1 Bar chart : Distribution of dihedral-angles and violations (i)



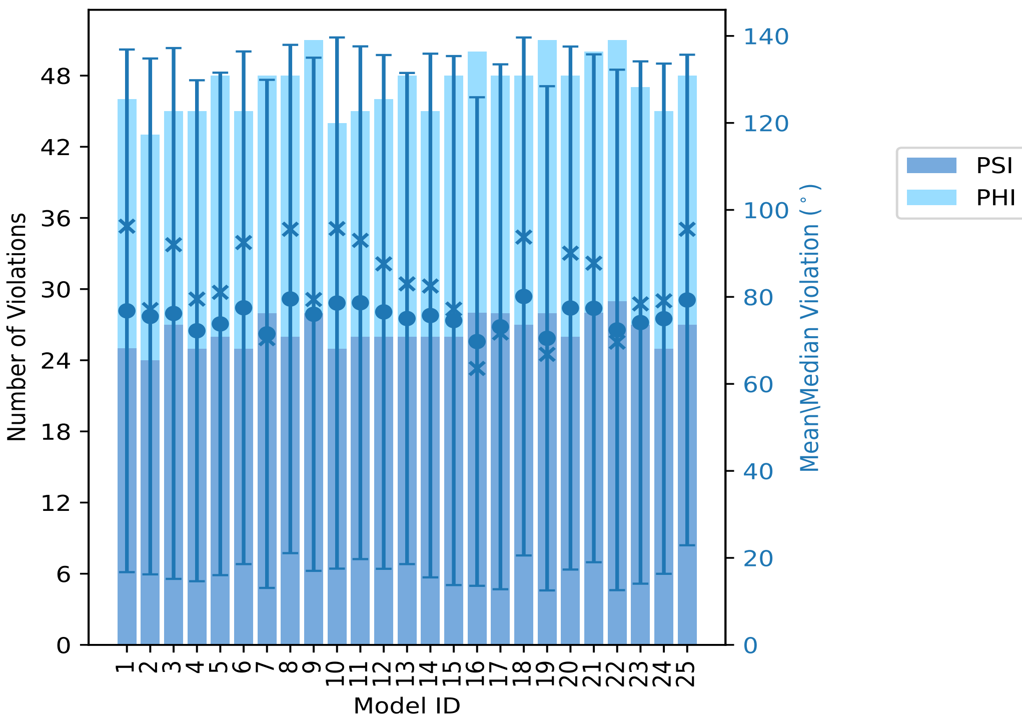

Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

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

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

| Model ID | Number of violations | | | Mean (°) | Max (°) | SD (°) | Median (°) |
|---|---|---|---|---|---|---|---|
| | PSI | PHI | Total | | | | |
| 1 | 25 | 21 | 46 | 76.8 | 155.1 | 60.08 | 96.25 |
| 2 | 24 | 19 | 43 | 75.51 | 155.2 | 59.3 | 77.1 |
| 3 | 27 | 18 | 45 | 76.19 | 164.7 | 61.02 | 92.0 |
| 4 | 25 | 20 | 45 | 72.22 | 155.9 | 57.59 | 79.5 |
| 5 | 26 | 22 | 48 | 73.79 | 158.1 | 57.76 | 81.05 |
| 6 | 25 | 20 | 45 | 77.5 | 156.1 | 58.93 | 92.5 |
| 7 | 28 | 20 | 48 | 71.5 | 158.1 | 58.41 | 70.4 |
| 8 | 26 | 22 | 48 | 79.52 | 159.2 | 58.43 | 95.55 |
| 9 | 28 | 23 | 51 | 76.01 | 164.1 | 58.99 | 79.4 |
| 10 | 25 | 19 | 44 | 78.59 | 157.7 | 61.06 | 95.7 |
| 11 | 26 | 19 | 45 | 78.66 | 163.8 | 58.94 | 93.0 |
| 12 | 26 | 20 | 46 | 76.55 | 161.6 | 59.06 | 87.6 |
| 13 | 26 | 22 | 48 | 75.03 | 155.4 | 56.46 | 83.0 |
| 14 | 26 | 19 | 45 | 75.72 | 161.4 | 60.2 | 82.5 |
| 15 | 26 | 22 | 48 | 74.56 | 162.1 | 60.81 | 77.2 |
| 16 | 28 | 22 | 50 | 69.74 | 157.1 | 56.16 | 63.55 |
| 17 | 28 | 20 | 48 | 73.13 | 158.6 | 60.34 | 71.7 |
| 18 | 27 | 21 | 48 | 80.1 | 165.3 | 59.54 | 93.75 |
| 19 | 28 | 23 | 51 | 70.48 | 159.0 | 57.96 | 66.8 |
| 20 | 26 | 22 | 48 | 77.44 | 164.0 | 60.12 | 90.05 |
| 21 | 28 | 22 | 50 | 77.38 | 165.2 | 58.38 | 87.75 |
| 22 | 29 | 22 | 51 | 72.41 | 157.7 | 59.83 | 69.6 |
| 23 | 27 | 20 | 47 | 74.11 | 156.2 | 60.04 | 78.4 |
| 24 | 25 | 20 | 45 | 74.99 | 156.0 | 58.65 | 79.1 |
| 25 | 27 | 21 | 48 | 79.29 | 156.4 | 56.39 | 95.55 |

### 10.2.1 Bar graph : Dihedral violation statistics for each model (i)




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

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

Violation analysis may find that some restraints are violated in very few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of ensemble.

| Number of violated restraints | | | Fraction of the ensemble | |
|---|---|---|---|---|
| PSI | PHI | Total | Count[1] | % |
| 1 | 1 | 2 | 1 | 4.0 |
| 1 | 0 | 1 | 2 | 8.0 |
| 2 | 0 | 2 | 3 | 12.0 |
| 0 | 0 | 0 | 4 | 16.0 |
| 0 | 0 | 0 | 5 | 20.0 |
| 0 | 1 | 1 | 6 | 24.0 |
| 0 | 0 | 0 | 7 | 28.0 |
| 0 | 1 | 1 | 8 | 32.0 |
| 0 | 0 | 0 | 9 | 36.0 |
| 0 | 0 | 0 | 10 | 40.0 |
| 1 | 1 | 2 | 11 | 44.0 |

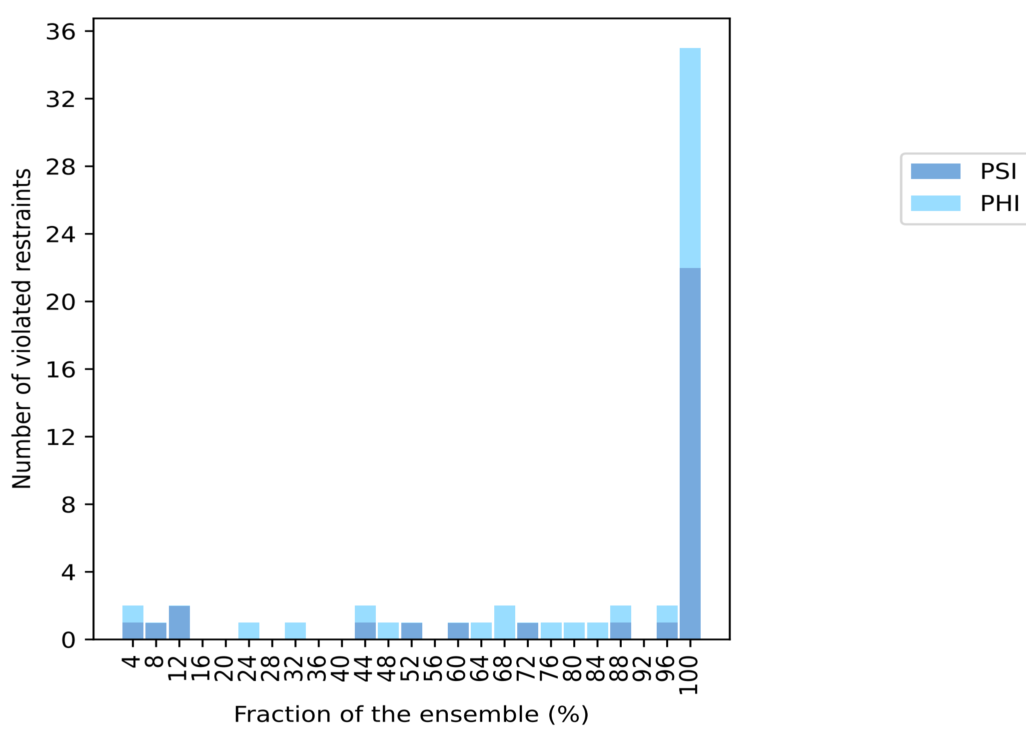
*Continued on next page...*

*Continued from previous page...*

| Number of violated restraints | | | Fraction of the ensemble | |
|---|---|---|---|---|
| PSI | PHI | Total | Count[1] | % |
| 0 | 1 | 1 | 12 | 48.0 |
| 1 | 0 | 1 | 13 | 52.0 |
| 0 | 0 | 0 | 14 | 56.0 |
| 1 | 0 | 1 | 15 | 60.0 |
| 0 | 1 | 1 | 16 | 64.0 |
| 0 | 2 | 2 | 17 | 68.0 |
| 1 | 0 | 1 | 18 | 72.0 |
| 0 | 1 | 1 | 19 | 76.0 |
| 0 | 1 | 1 | 20 | 80.0 |
| 0 | 1 | 1 | 21 | 84.0 |
| 1 | 1 | 2 | 22 | 88.0 |
| 0 | 0 | 0 | 23 | 92.0 |
| 1 | 1 | 2 | 24 | 96.0 |
| 22 | 13 | 35 | 25 | 100.0 |

[1] Number of models with violations

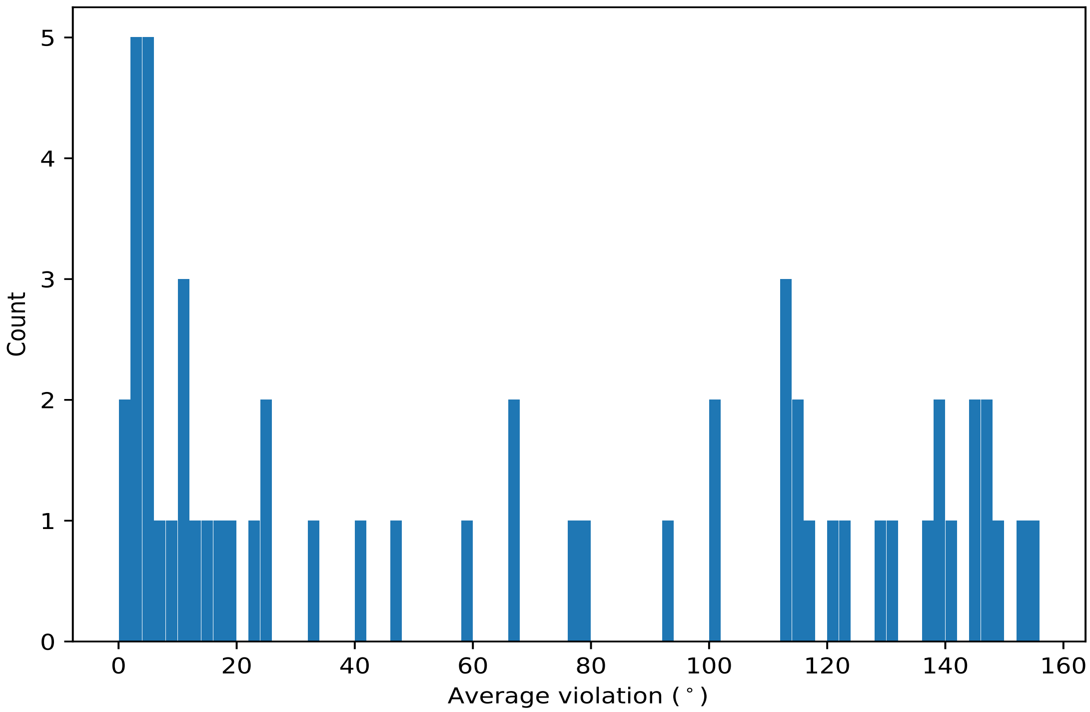
### 10.3.1 Bar graph : Dihedral-angle Violation statistics for the ensemble (i)



## 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.

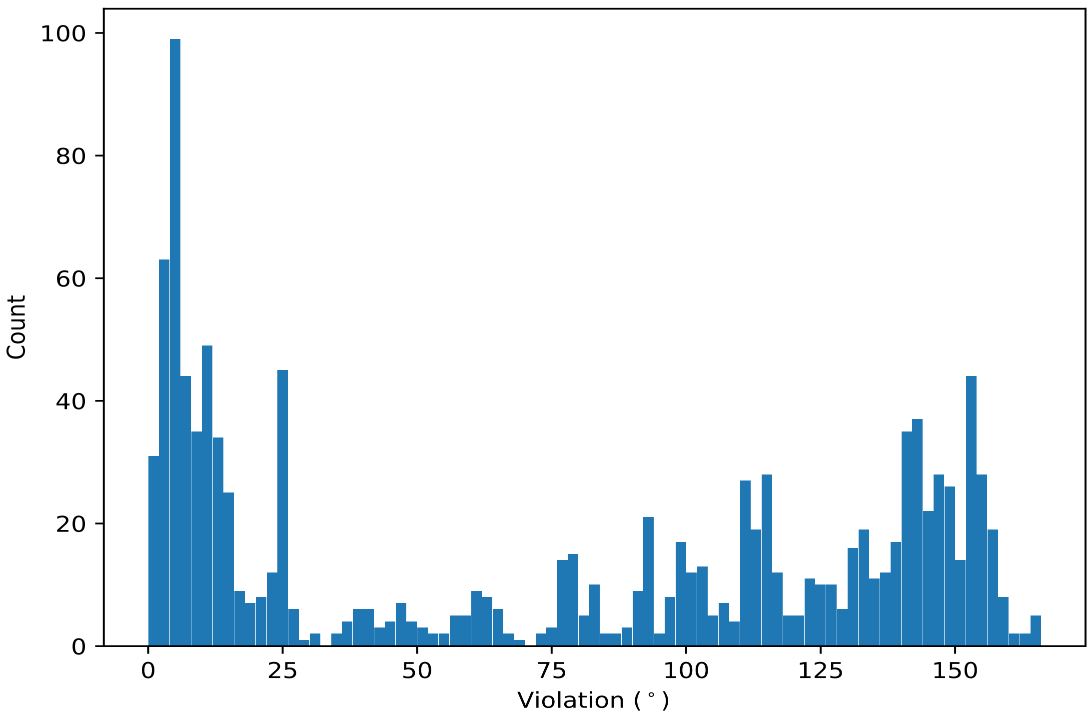
| Key | Atom-1 | Atom-2 | Atom-3 | Atom-4 | Models[1] | Mean | SD[2] | Median |
|---|---|---|---|---|---|---|---|---|
| (1,91) | 1:A:65:PHE:N | 1:A:65:PHE:CA | 1:A:65:PHE:C | 1:A:66:LYS:N | 25 | 155.63 | 1.24 | 155.6 |
| (1,14) | 1:A:23:PRO:C | 1:A:24:ILE:N | 1:A:24:ILE:CA | 1:A:24:ILE:C | 25 | 152.6 | 1.58 | 152.6 |
| (1,89) | 1:A:59:THR:N | 1:A:59:THR:CA | 1:A:59:THR:C | 1:A:60:ALA:N | 25 | 149.6 | 8.47 | 152.9 |
| (1,67) | 1:A:26:SER:N | 1:A:26:SER:CA | 1:A:26:SER:C | 1:A:27:PRO:N | 25 | 147.45 | 2.59 | 147.1 |
| (1,78) | 1:A:46:ASP:N | 1:A:46:ASP:CA | 1:A:46:ASP:C | 1:A:47:VAL:N | 25 | 147.23 | 21.7 | 155.1 |
| (1,22) | 1:A:36:ASN:C | 1:A:37:GLY:N | 1:A:37:GLY:CA | 1:A:37:GLY:C | 25 | 145.18 | 3.89 | 146.3 |
| (1,79) | 1:A:48:GLU:N | 1:A:48:GLU:CA | 1:A:48:GLU:C | 1:A:49:LEU:N | 25 | 144.97 | 3.09 | 145.5 |
| (1,93) | 1:A:67:SER:N | 1:A:67:SER:CA | 1:A:67:SER:C | 1:A:68:ALA:N | 25 | 141.85 | 0.98 | 141.8 |
| (1,80) | 1:A:49:LEU:N | 1:A:49:LEU:CA | 1:A:49:LEU:C | 1:A:50:LYS:N | 25 | 139.4 | 3.68 | 140.2 |
| (1,86) | 1:A:56:GLN:N | 1:A:56:GLN:CA | 1:A:56:GLN:C | 1:A:57:VAL:N | 25 | 138.86 | 3.84 | 138.4 |

[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 violationsfor the 10 worst performing restraints, sorted by the violation value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint.

| Key | Atom-1 | Atom-2 | Atom-3 | Atom-4 | Model ID | Violation (°) |
|---|---|---|---|---|---|---|
| (1,53) | 1:A:3:MET:N | 1:A:3:MET:CA | 1:A:3:MET:C | 1:A:4:GLY:N | 18 | 165.3 |
| (1,53) | 1:A:3:MET:N | 1:A:3:MET:CA | 1:A:3:MET:C | 1:A:4:GLY:N | 21 | 165.2 |
| (1,53) | 1:A:3:MET:N | 1:A:3:MET:CA | 1:A:3:MET:C | 1:A:4:GLY:N | 3 | 164.7 |
| (1,53) | 1:A:3:MET:N | 1:A:3:MET:CA | 1:A:3:MET:C | 1:A:4:GLY:N | 9 | 164.1 |
| (1,53) | 1:A:3:MET:N | 1:A:3:MET:CA | 1:A:3:MET:C | 1:A:4:GLY:N | 20 | 164.0 |
| (1,53) | 1:A:3:MET:N | 1:A:3:MET:CA | 1:A:3:MET:C | 1:A:4:GLY:N | 11 | 163.8 |
| (1,53) | 1:A:3:MET:N | 1:A:3:MET:CA | 1:A:3:MET:C | 1:A:4:GLY:N | 15 | 162.1 |
| (1,53) | 1:A:3:MET:N | 1:A:3:MET:CA | 1:A:3:MET:C | 1:A:4:GLY:N | 12 | 161.6 |
| (1,53) | 1:A:3:MET:N | 1:A:3:MET:CA | 1:A:3:MET:C | 1:A:4:GLY:N | 14 | 161.4 |
| (1,53) | 1:A:3:MET:N | 1:A:3:MET:CA | 1:A:3:MET:C | 1:A:4:GLY:N | 8 | 159.2 |