



# Full wwPDB NMR Structure Validation Report ⓘ

Jun 3, 2023 – 05:24 PM EDT

PDB ID : 2MPQ  
BMRB ID : 19998  
Title : Solution structure of the sodium channel toxin Hd1a  
Authors : Klint, J.K.; Mobli, M.; King, G.F.  
Deposited on : 2014-06-01

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

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 ⓘ 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 ⓘ](#)) 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

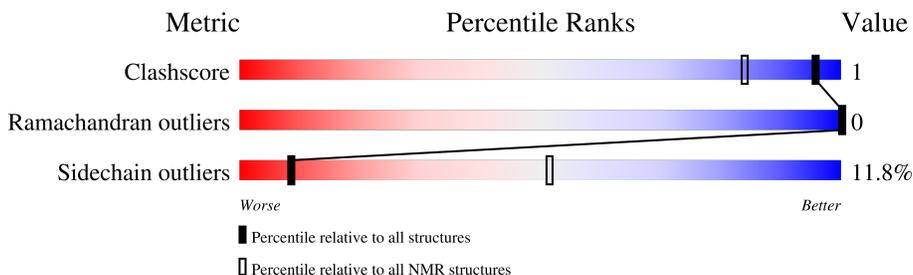
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*SOLUTION NMR*

The overall completeness of chemical shifts assignment is 93%.

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  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	36	92% 6% •

## 2 Ensemble composition and analysis i

This entry contains 20 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: *best mobprobability score*.

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:2-A:36 (35)	0.10	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 5 clusters and 3 single-model clusters were found.

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

### 3 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 516 atoms, of which 250 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Hd1a.

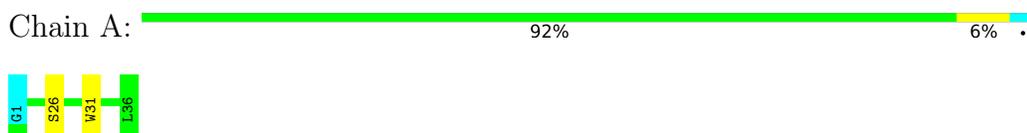
Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	36	516	162	250	47	51	6	0

## 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: Hd1a

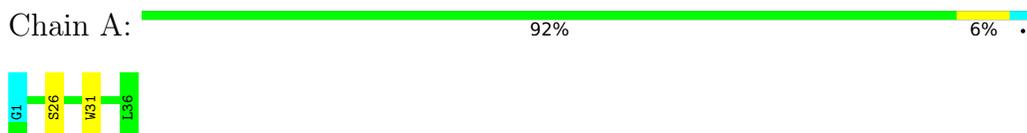


### 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

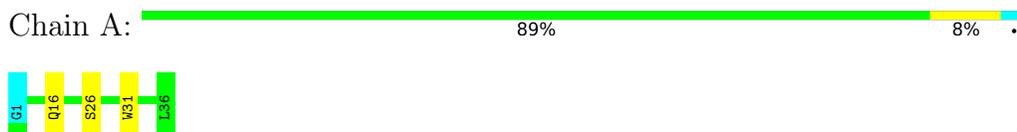
#### 4.2.1 Score per residue for model 1

- Molecule 1: Hd1a



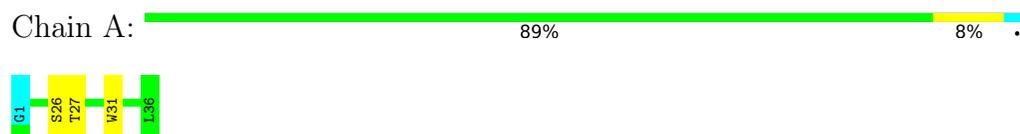
#### 4.2.2 Score per residue for model 2 (medoid)

- Molecule 1: Hd1a



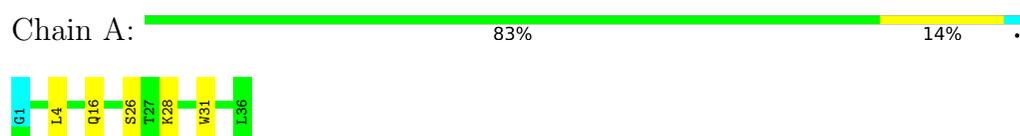
### 4.2.3 Score per residue for model 3

- Molecule 1: Hd1a



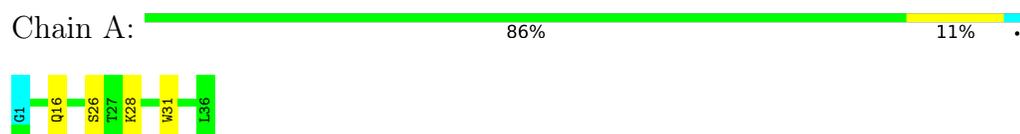
### 4.2.4 Score per residue for model 4

- Molecule 1: Hd1a



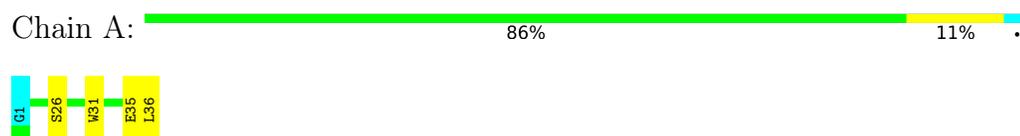
### 4.2.5 Score per residue for model 5

- Molecule 1: Hd1a



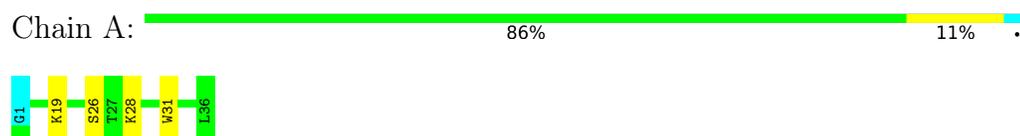
### 4.2.6 Score per residue for model 6

- Molecule 1: Hd1a



### 4.2.7 Score per residue for model 7

- Molecule 1: Hd1a



#### 4.2.8 Score per residue for model 8

- Molecule 1: Hd1a

Chain A:  86% 11%



#### 4.2.9 Score per residue for model 9

- Molecule 1: Hd1a

Chain A:  83% 14%



#### 4.2.10 Score per residue for model 10

- Molecule 1: Hd1a

Chain A:  83% 14%



#### 4.2.11 Score per residue for model 11

- Molecule 1: Hd1a

Chain A:  81% 17%



#### 4.2.12 Score per residue for model 12

- Molecule 1: Hd1a

Chain A:  86% 11%



#### 4.2.13 Score per residue for model 13

- Molecule 1: Hd1a

Chain A:  86% 11% .



#### 4.2.14 Score per residue for model 14

- Molecule 1: Hd1a

Chain A:  83% 14% .



#### 4.2.15 Score per residue for model 15

- Molecule 1: Hd1a

Chain A:  86% 11% .



#### 4.2.16 Score per residue for model 16

- Molecule 1: Hd1a

Chain A:  75% 19% . .



#### 4.2.17 Score per residue for model 17

- Molecule 1: Hd1a

Chain A:  78% 17% . .



#### 4.2.18 Score per residue for model 18

- Molecule 1: Hd1a

Chain A:  86% 11%



#### 4.2.19 Score per residue for model 19

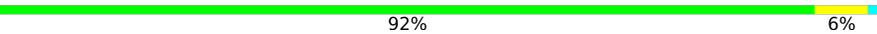
- Molecule 1: Hd1a

Chain A:  86% 11%



#### 4.2.20 Score per residue for model 20

- Molecule 1: Hd1a

Chain A:  92% 6%



## 5 Refinement protocol and experimental data overview

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

Of the 200 calculated structures, 20 were deposited, based on the following criterion: *Best MobProbability score*.

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

Software name	Classification	Version
TALOS	geometry optimization	+
CYANA	structure solution	3.0
CYANA	refinement	

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	398
Number of shifts mapped to atoms	398
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	93%

## 6 Model quality [i](#)

### 6.1 Standard geometry [i](#)

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 [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	262	247	247	0±1
All	All	5240	4940	4940	9

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 1.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:12:PRO:CB	1:A:27:THR:HG23	0.54	2.33	16	2
1:A:12:PRO:HB2	1:A:27:THR:HG23	0.47	1.86	16	1
1:A:35:GLU:C	1:A:36:LEU:HD12	0.47	2.30	11	1
1:A:18:CYS:HB2	1:A:23:LEU:HD12	0.41	1.92	16	2
1:A:24:ALA:HB3	1:A:35:GLU:HG3	0.40	1.93	10	1
1:A:4:LEU:HD12	1:A:16:GLN:O	0.40	2.17	4	2

### 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	34/36 (94%)	29±0 (86±1%)	5±0 (14±1%)	0±0 (0±0%)	100	100
All	All	680/720 (94%)	585 (86%)	95 (14%)	0 (0%)	100	100

There are no Ramachandran outliers.

### 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	31/31 (100%)	27±1 (88±3%)	4±1 (12±3%)	8	51
All	All	620/620 (100%)	547 (88%)	73 (12%)	8	51

All 14 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	26	SER	20
1	A	31	TRP	20
1	A	28	LYS	9
1	A	27	THR	4
1	A	35	GLU	4
1	A	13	SER	3
1	A	16	GLN	2
1	A	19	LYS	2
1	A	8	LYS	2
1	A	14	ASN	2
1	A	22	SER	2
1	A	36	LEU	1
1	A	9	SER	1
1	A	21	SER	1

### 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 93% for the well-defined parts and 93% 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	398
Number of shifts mapped to atoms	398
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

#### 7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction $\pm$ precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	35	$0.33 \pm 0.32$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}_\beta$	33	$0.28 \pm 0.44$	None needed ( $< 0.5$ ppm)
$^{13}\text{C}'$	31	$2.87 \pm 0.37$	Should be applied
$^{15}\text{N}$	34	$-0.16 \pm 0.54$	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 93%, i.e. 394 atoms were assigned a chemical shift out of a possible 423. 0 out of 3 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	$^1\text{H}$	$^{13}\text{C}$	$^{15}\text{N}$
Backbone	169/175 (97%)	71/71 (100%)	64/70 (91%)	34/34 (100%)
Sidechain	195/210 (93%)	132/135 (98%)	60/67 (90%)	3/8 (38%)

*Continued on next page...*

Continued from previous page...

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Aromatic	30/38 (79%)	17/19 (89%)	12/17 (71%)	1/2 (50%)
Overall	394/423 (93%)	220/225 (98%)	136/154 (88%)	38/44 (86%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 93%, i.e. 398 atoms were assigned a chemical shift out of a possible 429. 0 out of 3 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	<sup>1</sup> H	<sup>13</sup> C	<sup>15</sup> N
Backbone	173/181 (96%)	73/74 (99%)	66/72 (92%)	34/35 (97%)
Sidechain	195/210 (93%)	132/135 (98%)	60/67 (90%)	3/8 (38%)
Aromatic	30/38 (79%)	17/19 (89%)	12/17 (71%)	1/2 (50%)
Overall	398/429 (93%)	222/228 (97%)	138/156 (88%)	38/45 (84%)

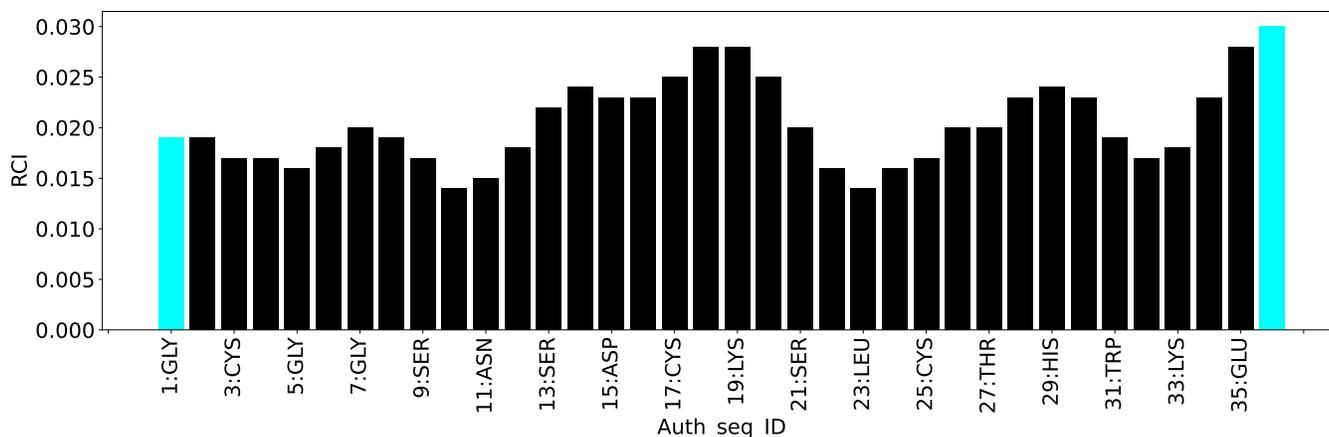
#### 7.1.4 Statistically unusual chemical shifts [i](#)

There are no statistically unusual chemical shifts.

#### 7.1.5 Random Coil Index (RCI) plots [i](#)

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

Random coil index (RCI) for chain A:



## 8 NMR restraints analysis [i](#)

### 8.1 Conformationally restricting restraints [i](#)

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

Description	Value
Total distance restraints	601
Intra-residue ( $ i-j =0$ )	131
Sequential ( $ i-j =1$ )	154
Medium range ( $ i-j >1$ and $ i-j <5$ )	109
Long range ( $ i-j \geq 5$ )	207
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	0
Number of restraints per residue	16.7
Number of long range restraints per residue <sup>1</sup>	5.8

<sup>1</sup>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. There are no distance violations

#### 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](#)

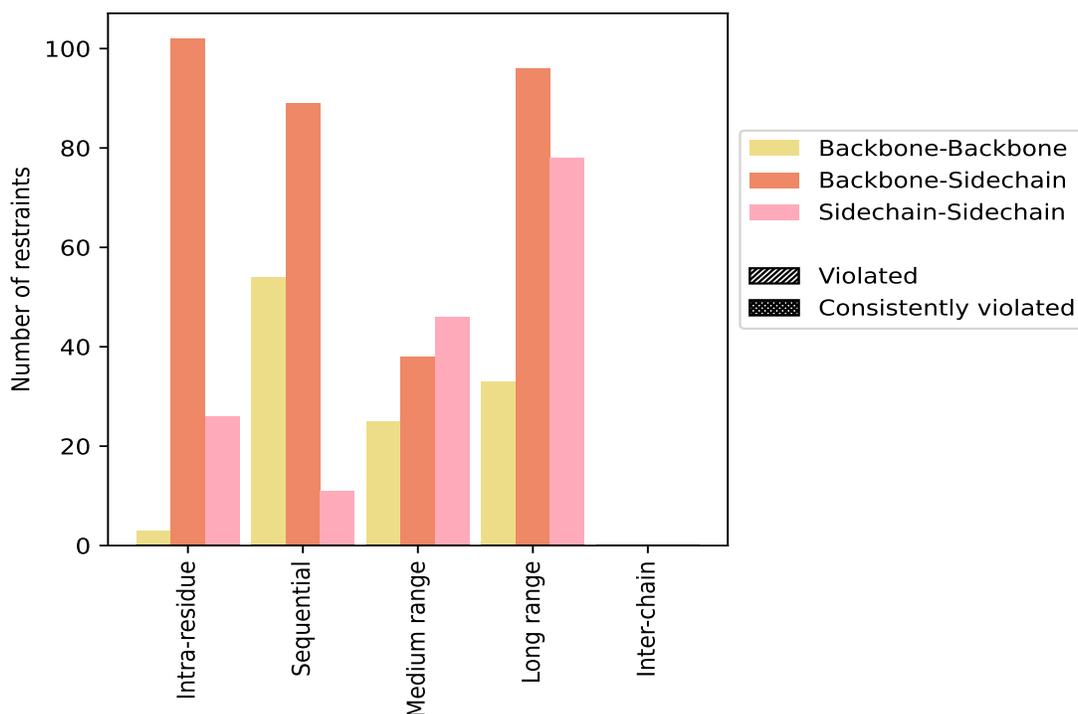
### 9.1 Summary of distance violations [i](#)

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.

Restrains type	Count	% <sup>1</sup>	Violated <sup>3</sup>			Consistently Violated <sup>4</sup>		
			Count	% <sup>2</sup>	% <sup>1</sup>	Count	% <sup>2</sup>	% <sup>1</sup>
<b>Intra-residue ( i-j =0)</b>	<b>131</b>	<b>21.8</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
Backbone-Backbone	3	0.5	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	102	17.0	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	26	4.3	0	0.0	0.0	0	0.0	0.0
<b>Sequential ( i-j =1)</b>	<b>154</b>	<b>25.6</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
Backbone-Backbone	54	9.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	89	14.8	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	11	1.8	0	0.0	0.0	0	0.0	0.0
<b>Medium range ( i-j &gt;1 &amp;  i-j &lt;5)</b>	<b>109</b>	<b>18.1</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
Backbone-Backbone	25	4.2	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	38	6.3	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	46	7.7	0	0.0	0.0	0	0.0	0.0
<b>Long range ( i-j ≥5)</b>	<b>207</b>	<b>34.4</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
Backbone-Backbone	33	5.5	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	96	16.0	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	78	13.0	0	0.0	0.0	0	0.0	0.0
<b>Inter-chain</b>	<b>0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
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
<b>Hydrogen bond</b>	<b>0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
<b>Disulfide bond</b>	<b>0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
<b>Total</b>	<b>601</b>	<b>100.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>	<b>0</b>	<b>0.0</b>	<b>0.0</b>
Backbone-Backbone	115	19.1	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	325	54.1	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	161	26.8	0	0.0	0.0	0	0.0	0.0

<sup>1</sup> percentage calculated with respect to the total number of distance restraints, <sup>2</sup> percentage calculated with respect to the number of restraints in a particular restraint category, <sup>3</sup> violated in at least one model, <sup>4</sup> 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 disulfid bonds are counted in their appropriate category on the x-axis

## 9.2 Distance violation statistics for each model [i](#)

No violations found

## 9.3 Distance violation statistics for the ensemble [i](#)

No violations found

## 9.4 Most violated distance restraints in the ensemble [i](#)

No violations found

## 9.5 All violated distance restraints [i](#)

No violations found

## 10 Dihedral-angle violation analysis

Dihedral angle analysis failed due to data error in the dihedral angle restraints, possibly missing target value