



wwPDB NMR Structure Validation Summary Report ⓘ

Jul 2, 2020 – 12:14 AM CDT

PDB ID : 2JR2
Title : Solution NMR structure of homodimer CPS_2611 from *Colwellia psychrerythraea*. Northeast Structural Genomics Consortium target CsR4.
Authors : Ramelot, T.A.; Cort, J.R.; Wang, H.; Nwosu, C.; Cunningham, K.; Owens, L.; Ma, L.-C.; Xiao, R.; Liu, J.; Baran, M.C.; Swapna, G.; Acton, T.B.; Rost, B.; Montelione, G.T.; Kennedy, M.A.; Northeast Structural Genomics Consortium (NESG)
Deposited on : 2007-06-19

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 ⓘ symbol.

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

Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.6.dev1
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.6.dev1

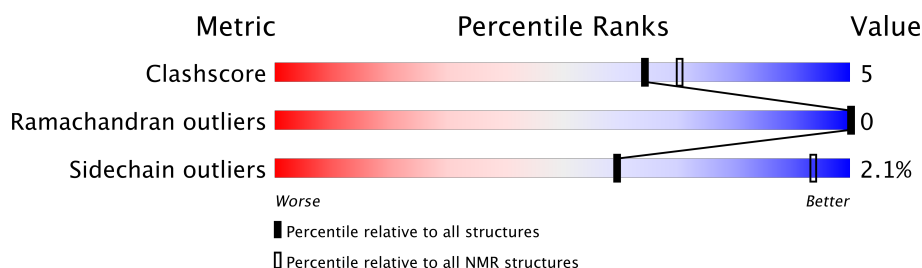
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 66%.

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	136327	12091
Ramachandran outliers	132723	10835
Sidechain outliers	132532	10811

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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	76	
1	B	76	

2 Ensemble composition and analysis

This entry contains 20 models. Model 1 is the overall representative, medoid model (most similar to other models).

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:8-A:69, B:8-B:68 (123)	0.35	1

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 4 single-model clusters were found.

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

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 2460 atoms, of which 1254 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called UPF0352 protein CPS_2611.

Mol	Chain	Residues	Atoms						Trace
1	A	76	Total	C	H	N	O	S	0
			1230	381	627	107	112	3	
1	B	76	Total	C	H	N	O	S	0
			1230	381	627	107	112	3	

There are 16 discrepancies between the modelled and reference sequences:

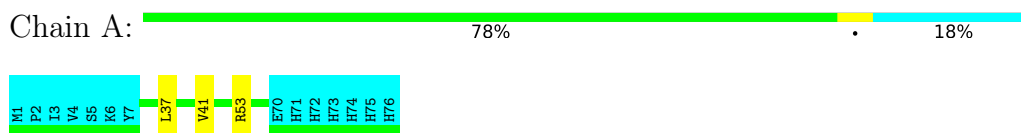
Chain	Residue	Modelled	Actual	Comment	Reference
A	69	LEU	-	CLONING ARTIFACT	UNP Q481E4
A	70	GLU	-	CLONING ARTIFACT	UNP Q481E4
A	71	HIS	-	CLONING ARTIFACT	UNP Q481E4
A	72	HIS	-	CLONING ARTIFACT	UNP Q481E4
A	73	HIS	-	CLONING ARTIFACT	UNP Q481E4
A	74	HIS	-	CLONING ARTIFACT	UNP Q481E4
A	75	HIS	-	CLONING ARTIFACT	UNP Q481E4
A	76	HIS	-	CLONING ARTIFACT	UNP Q481E4
B	69	LEU	-	CLONING ARTIFACT	UNP Q481E4
B	70	GLU	-	CLONING ARTIFACT	UNP Q481E4
B	71	HIS	-	CLONING ARTIFACT	UNP Q481E4
B	72	HIS	-	CLONING ARTIFACT	UNP Q481E4
B	73	HIS	-	CLONING ARTIFACT	UNP Q481E4
B	74	HIS	-	CLONING ARTIFACT	UNP Q481E4
B	75	HIS	-	CLONING ARTIFACT	UNP Q481E4
B	76	HIS	-	CLONING ARTIFACT	UNP Q481E4

4 Residue-property plots [i](#)

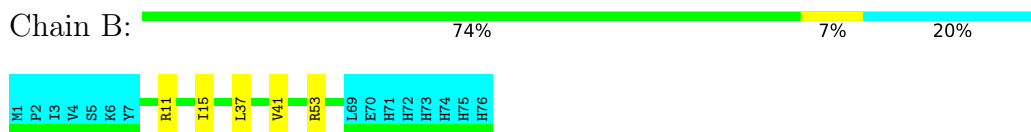
4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA 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: UPF0352 protein CPS_2611



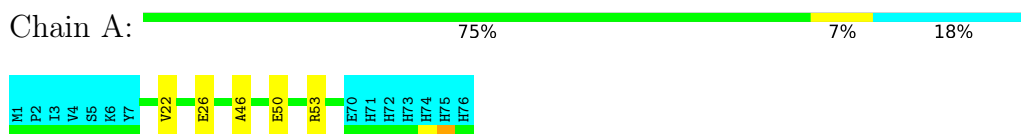
- Molecule 1: UPF0352 protein CPS_2611



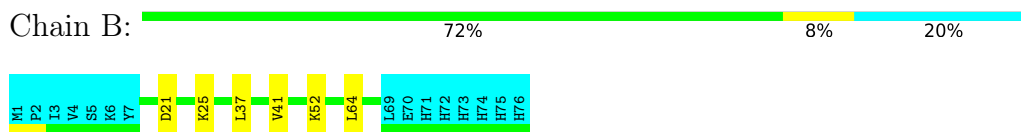
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 1. Colouring as in section 4.1 above.

- Molecule 1: UPF0352 protein CPS_2611



- Molecule 1: UPF0352 protein CPS_2611



5 Refinement protocol and experimental data overview

The models were refined using the following method: *simulated annealing, CNS water refinement*.

Of the 20 calculated structures, 20 were deposited, based on the following criterion: *all calculated structures submitted*.

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

Software name	Classification	Version
X-PLOR NIH	structure solution	2.15.0
CNS	refinement	1.1

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)	2jr2_nmr.cif
Number of chemical shift lists	1
Total number of shifts	1929
Number of shifts mapped to atoms	1183
Number of unparsed shifts	298
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	448
Assignment completeness (well-defined parts)	66%

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	476	508	508	5±2
1	B	468	497	497	5±2
All	All	18880	20100	20100	180

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 5.

5 of 76 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:B:22:VAL:HA	1:B:25:LYS:HE2	0.84	1.45	3	1
1:A:61:THR:HB	1:B:61:THR:HB	0.80	1.54	3	6
1:B:66:GLN:HA	1:B:66:GLN:HE21	0.74	1.42	6	1
1:A:30:PRO:HB3	1:B:48:VAL:HG22	0.67	1.66	10	1
1:A:48:VAL:O	1:A:53:ARG:HD3	0.61	1.96	18	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	62/76 (82%)	61±1 (99±1%)	1±1 (1±1%)	0±0 (0±0%)	100	100
1	B	61/76 (80%)	60±1 (99±1%)	1±1 (1±1%)	0±0 (0±0%)	100	100
All	All	2460/3040 (81%)	2432 (99%)	28 (1%)	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	56/70 (80%)	55±1 (98±1%)	1±1 (2±1%)	59	93
1	B	55/70 (79%)	54±1 (98±2%)	1±1 (2±2%)	59	93
All	All	2220/2800 (79%)	2174 (98%)	46 (2%)	59	93

5 of 20 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	18	ASP	5
1	A	52	LYS	5
1	B	52	LYS	5
1	A	29	THR	4
1	B	29	THR	4

6.3.3 RNA [i](#)

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.6 Ligand geometry [i](#)

There are no ligands in this entry.

6.7 Other polymers [i](#)

There are no such molecules in this entry.

6.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 66% for the well-defined parts and 61% for the entire structure.

7.1 Chemical shift list 1

File name: 2jr2_nmr.cif

Chemical shift list name: *nef_chemical_shift_list_2jr2.mr*

7.1.1 Bookkeeping

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

Total number of shifts	1929
Number of shifts mapped to atoms	1183
Number of unparsed shifts	298
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	448
Number of shift outliers (ShiftChecker)	0

The following errors were found when reading this chemical shift list.

- Chemical shift has been reported more than once. First 5 (of 298) occurrences are reported below.

Shift ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
5	A	2	PRO	HD%	3.370	0.020	1
17	A	3	ILE	HD1%	0.830	0.020	1
18	A	3	ILE	HD1%	0.830	0.020	1
22	A	3	ILE	HG2%	0.870	0.020	1
23	A	3	ILE	HG2%	0.870	0.020	1

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atoms found in structure. First 5 (of 448) occurrences are reported below.

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	11	ARG	HB%	1.95	0.02	1
B	44	ILE	HG1x	1.01	0.02	2

Continued on next page...

Continued from previous page...

Chain	Res	Type	Atom	Shift Data		
				Value	Uncertainty	Ambiguity
A	30	PRO	HGy	2.12	0.02	2
B	43	ASN	HBy	2.81	0.02	2
A	16	ILE	HG1y	1.88	0.02	2

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
$^{13}\text{C}_\alpha$	148	-0.40 ± 0.11	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	146	0.40 ± 0.09	None needed (< 0.5 ppm)
$^{13}\text{C}'$	142	-0.29 ± 0.12	None needed (< 0.5 ppm)
^{15}N	142	0.21 ± 0.16	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 66%, i.e. 992 atoms were assigned a chemical shift out of a possible 1498. 35 out of 35 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	601/607 (99%)	240/242 (99%)	242/246 (98%)	119/119 (100%)
Sidechain	383/873 (44%)	57/499 (11%)	308/338 (91%)	18/36 (50%)
Aromatic	8/18 (44%)	4/10 (40%)	4/8 (50%)	0/0 (—%)
Overall	992/1498 (66%)	301/751 (40%)	554/592 (94%)	137/155 (88%)

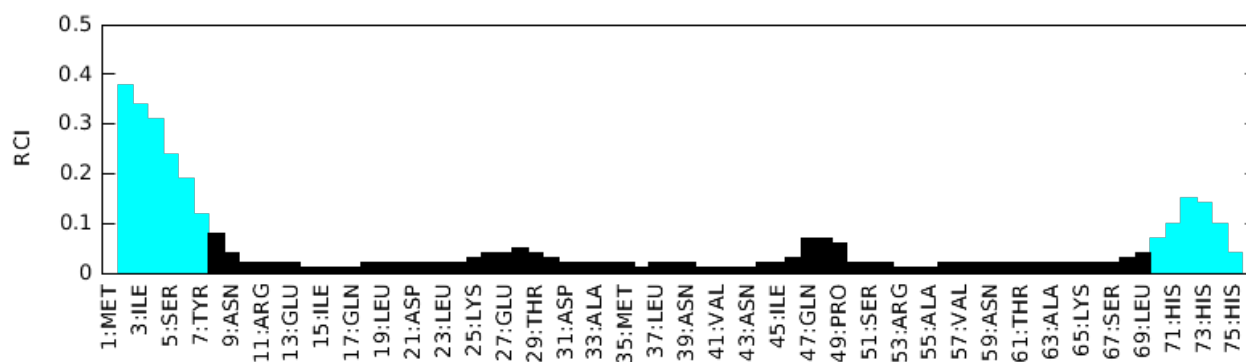
7.1.4 Statistically unusual chemical shifts [i](#)

There are no statistically unusual chemical shifts.

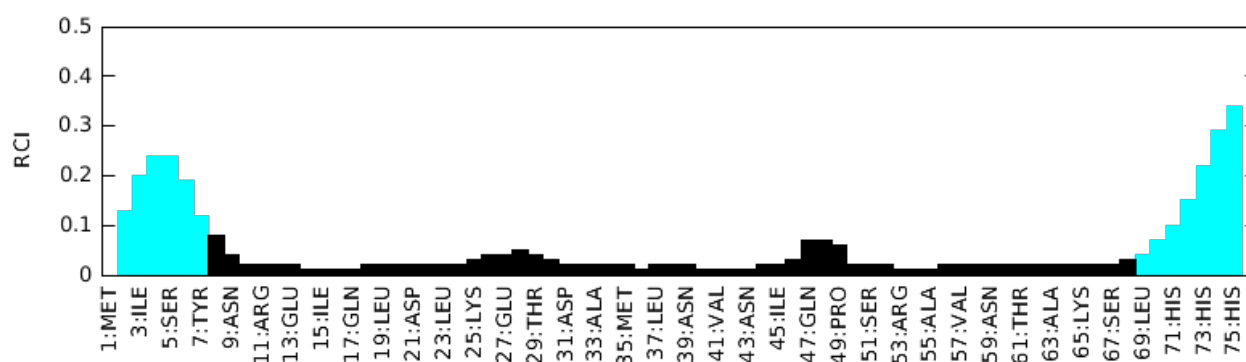
7.1.5 Random Coil Index (RCI) plots [i](#)

The images below report *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.

Random coil index (RCI) for chain A:



Random coil index (RCI) for chain B:



8 NMR restraints analysis

8.1 Conformationally restricting restraints

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

Description	Value
Total distance restraints	1506
Intra-residue ($ i-j =0$)	0
Sequential ($ i-j =1$)	216
Medium range ($ i-j >1$ and $ i-j <5$)	828
Long range ($ i-j \geq 5$)	96
Inter-chain	366
Total dihedral-angle restraints	212
Total hydrogen bond restraints	248
Total disulfide bond restraints	0
Number of unmapped restraints	0
Number of restraints per residue	11.1
Number of long range restraints per residue	0.7

8.2 Residual restraint violations

This section provides the overview of the restraint violations analysis. The violations are binned as small, medium and large violations based on its absolute value. Average number of violations per model is calculated by dividing the total number of violations in each bin by the size of the ensemble.

8.2.1 Average number of distance violations per model

Distance violations less than 0.1 Å are not included in the calculation. There are no distance violations

8.2.2 Average number of dihedral-angle violations per model

Dihedral-angle violations less than 1° are not included in the calculation. There are no dihedral-angle violations

9 Distance violation analysis

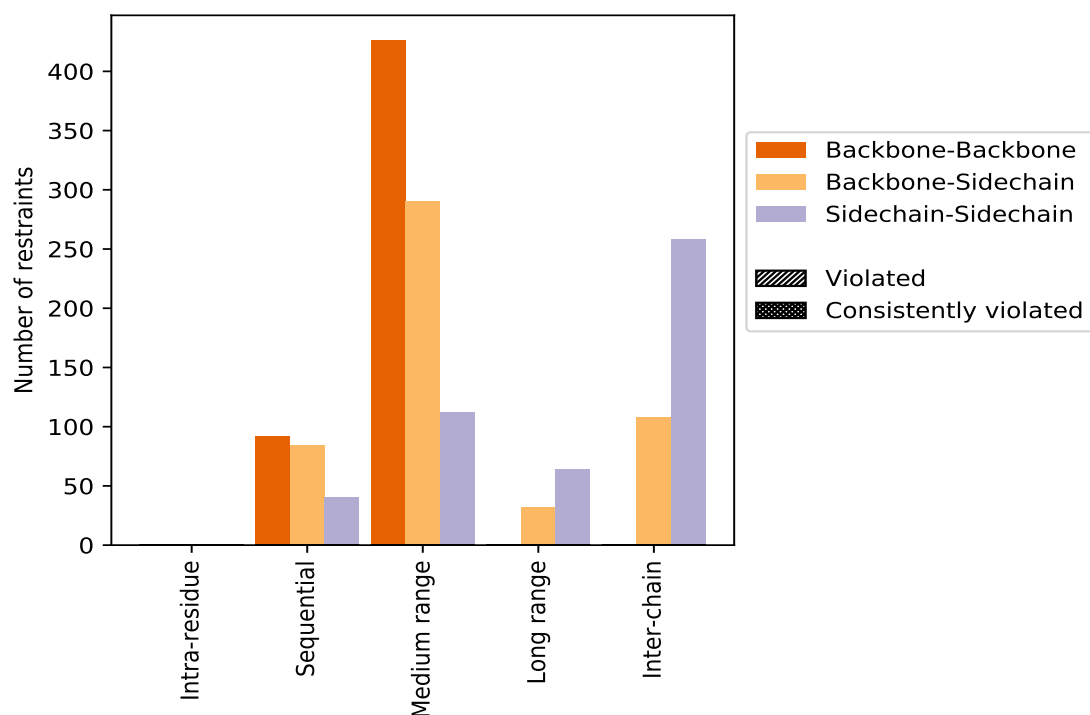
9.1 Summary of distance violations

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	% ¹	Violated ³			Consistently Violated ⁴		
			Count	% ²	% ¹	Count	% ²	% ¹
Intra-residue ($ i-j =0$)	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sequential ($ i-j =1$)	216	14.3	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	92	6.1	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	84	5.6	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	40	2.7	0	0.0	0.0	0	0.0	0.0
Medium range ($ i-j >1$ & $ i-j <5$)	828	55.0	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	426	28.3	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	290	19.3	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	112	7.4	0	0.0	0.0	0	0.0	0.0
Long range ($ i-j \geq 5$)	96	6.4	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	32	2.1	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	64	4.2	0	0.0	0.0	0	0.0	0.0
Inter-chain	366	24.3	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	108	7.2	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	258	17.1	0	0.0	0.0	0	0.0	0.0
Total	1506	100.0	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	518	34.4	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	514	34.1	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	474	31.5	0	0.0	0.0	0	0.0	0.0

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

9.2 Distance violation statistics for each model

No violations found

9.3 Distance violation statistics for the ensemble

No violations found

9.4 Most violated distance restraints in the ensemble

No violations found

9.5 All distance violations

No violations found

10 Dihedral-angle violation analysis

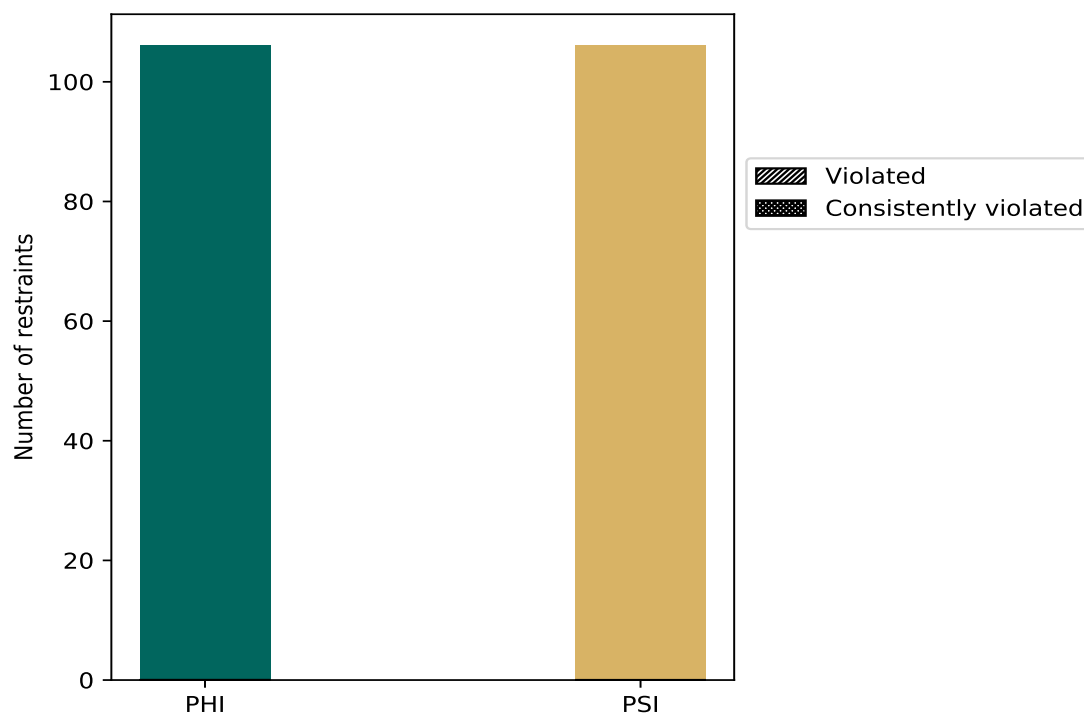
10.1 Summary of dihedral-angle violations

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	106	50.0	0	0.0	0.0	0	0.0	0.0
PSI	106	50.0	0	0.0	0.0	0	0.0	0.0
Total	212	100.0	0	0.0	0.0	0	0.0	0.0

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



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

10.2 Dihedral-angle violation statistics in each model

No violations found

10.3 Violation statistics in the ensemble

No violations found

10.4 Most violated dihedral-angle restraints

No violations found

10.5 All violated dihedral-angle restraints

No violations found