Full wwPDB NMR Structure Validation Report

Jun 4, 2023 – 10:49 AM EDT

PDB ID	:	2RSF
BMRB ID	:	11472
Title	:	Complex structure of WWE in RNF146 with ATP
Authors	:	He, F.; Muto, Y.; Inoue, M.; Kigawa, T.; Shirouzu, M.; Terada, T.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on	:	2012-01-31

This is a Full wwPDB NMR Structure Validation 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 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:

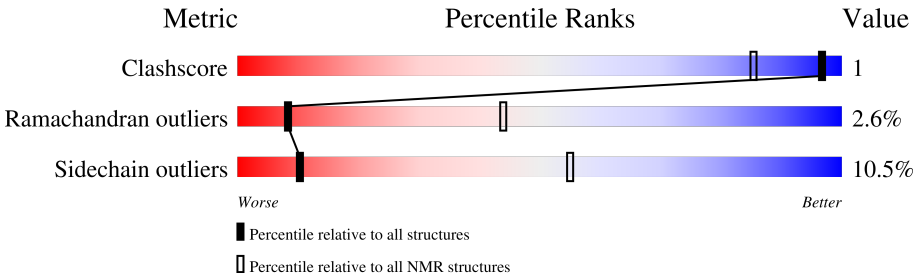
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints 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 89%.

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	110	60% 12% • 27%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 3 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:14-A:22 (9)	0.12	16
2	A:29-A:99 (71)	0.49	3

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

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

The models can be grouped into 3 clusters. No single-model clusters were found.

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

3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 1789 atoms, of which 872 are hydrogens and 0 are deuteriums.

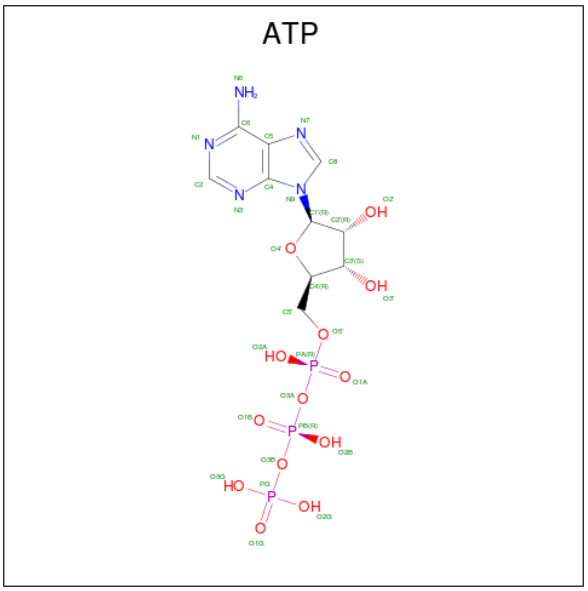
- Molecule 1 is a protein called E3 ubiquitin-protein ligase RNF146.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	110	1746	553	860	161	170	2	0

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	PRO	-	expression tag	UNP Q9CZW6
A	2	SER	-	expression tag	UNP Q9CZW6
A	3	SER	-	expression tag	UNP Q9CZW6
A	4	GLY	-	expression tag	UNP Q9CZW6
A	5	SER	-	expression tag	UNP Q9CZW6
A	6	SER	-	expression tag	UNP Q9CZW6
A	7	GLY	-	expression tag	UNP Q9CZW6
A	105	SER	-	expression tag	UNP Q9CZW6
A	106	GLY	-	expression tag	UNP Q9CZW6
A	107	PRO	-	expression tag	UNP Q9CZW6
A	108	SER	-	expression tag	UNP Q9CZW6
A	109	SER	-	expression tag	UNP Q9CZW6
A	110	GLY	-	expression tag	UNP Q9CZW6

- Molecule 2 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$).



Mol	Chain	Residues	Atoms					
			Total	C	H	N	O	P
2	A	1	43	10	12	5	13	3

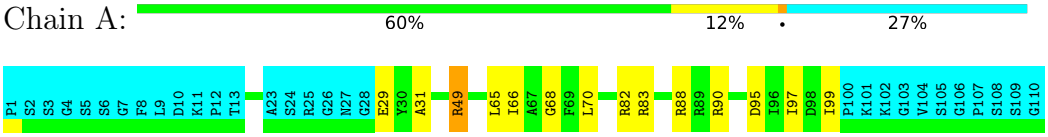
4 Residue-property plots

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: E3 ubiquitin-protein ligase RNF146

Chain A:



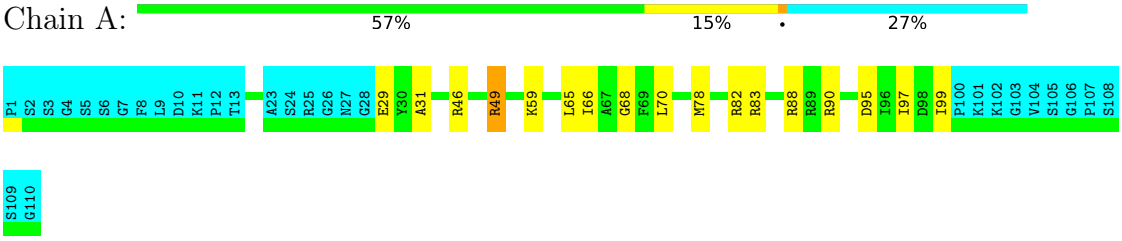
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: E3 ubiquitin-protein ligase RNF146

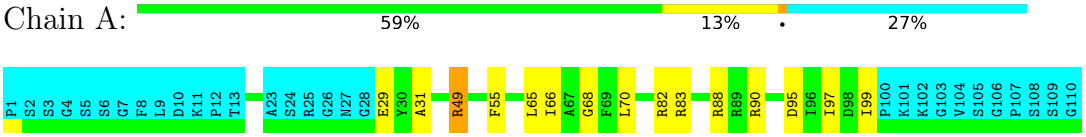
Chain A:



4.2.2 Score per residue for model 2

• Molecule 1: E3 ubiquitin-protein ligase RNF146

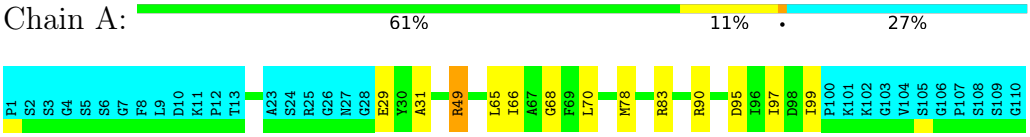
Chain A:



4.2.3 Score per residue for model 3 (medoid)

• Molecule 1: E3 ubiquitin-protein ligase RNF146

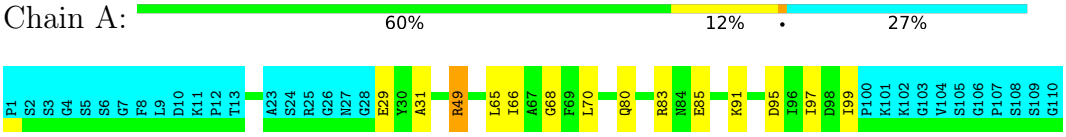
Chain A:



4.2.4 Score per residue for model 4

• Molecule 1: E3 ubiquitin-protein ligase RNF146

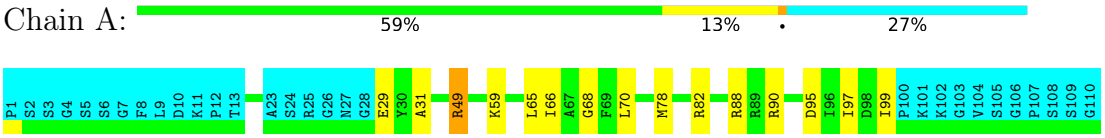
Chain A:



4.2.5 Score per residue for model 5

• Molecule 1: E3 ubiquitin-protein ligase RNF146

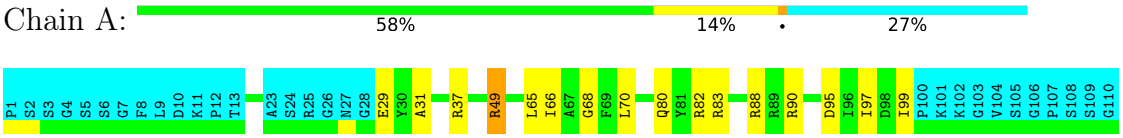
Chain A:



4.2.6 Score per residue for model 6

• Molecule 1: E3 ubiquitin-protein ligase RNF146

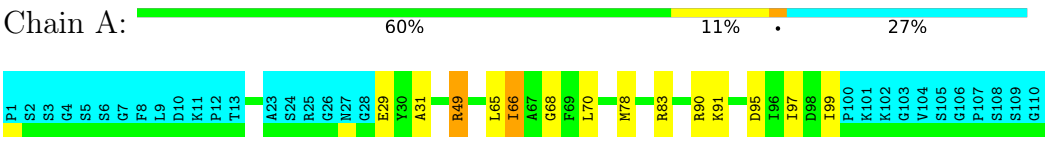
Chain A:



4.2.7 Score per residue for model 7

• Molecule 1: E3 ubiquitin-protein ligase RNF146

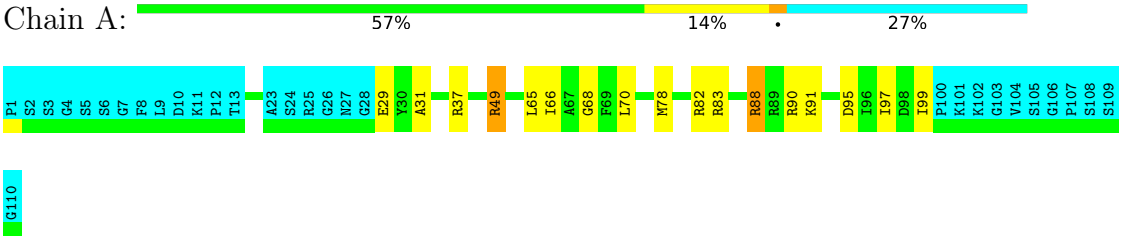
Chain A:



4.2.8 Score per residue for model 8

• Molecule 1: E3 ubiquitin-protein ligase RNF146

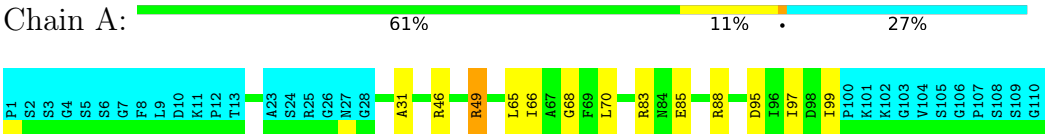
Chain A:



4.2.9 Score per residue for model 9

• Molecule 1: E3 ubiquitin-protein ligase RNF146

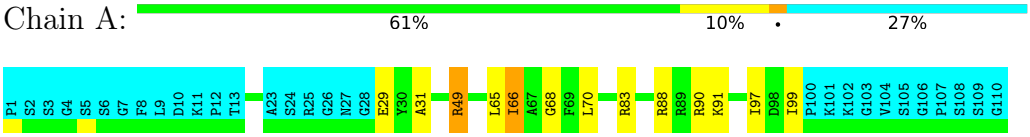
Chain A:



4.2.10 Score per residue for model 10

• Molecule 1: E3 ubiquitin-protein ligase RNF146

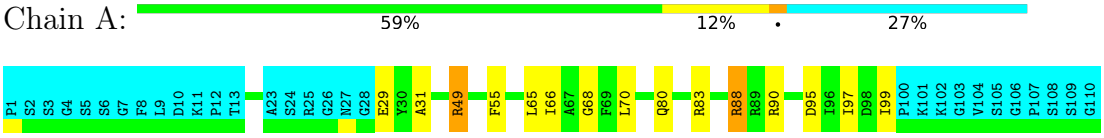
Chain A:



4.2.11 Score per residue for model 11

• Molecule 1: E3 ubiquitin-protein ligase RNF146

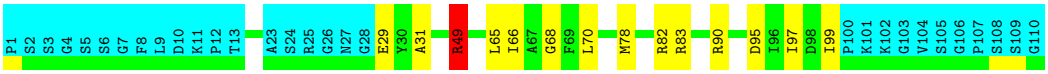
Chain A:



4.2.12 Score per residue for model 12

• Molecule 1: E3 ubiquitin-protein ligase RNF146

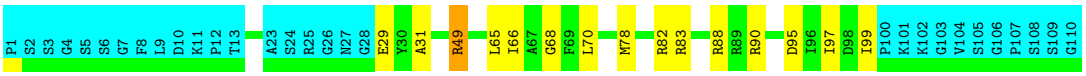
Chain A:



4.2.13 Score per residue for model 13

• Molecule 1: E3 ubiquitin-protein ligase RNF146

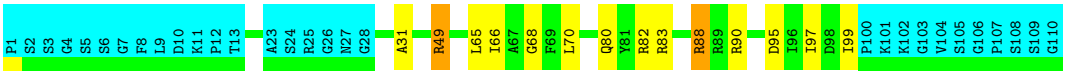
Chain A:



4.2.14 Score per residue for model 14

• Molecule 1: E3 ubiquitin-protein ligase RNF146

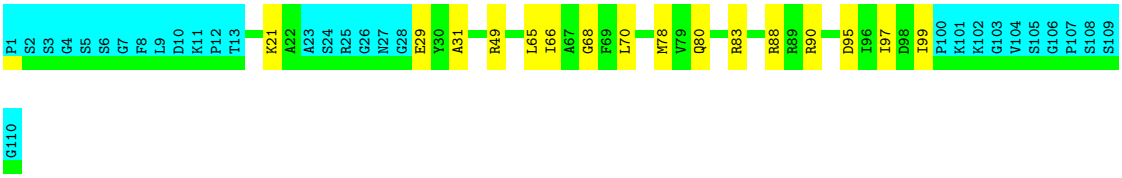
Chain A:



4.2.15 Score per residue for model 15

• Molecule 1: E3 ubiquitin-protein ligase RNF146

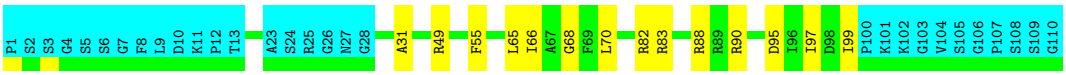
Chain A:



4.2.16 Score per residue for model 16

• Molecule 1: E3 ubiquitin-protein ligase RNF146

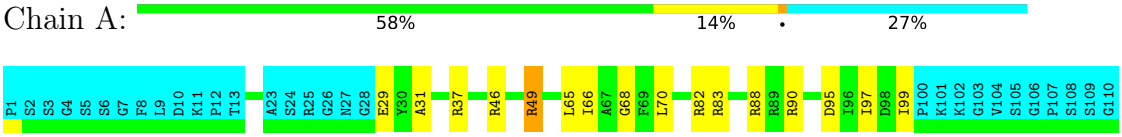
Chain A:



4.2.17 Score per residue for model 17

• Molecule 1: E3 ubiquitin-protein ligase RNF146

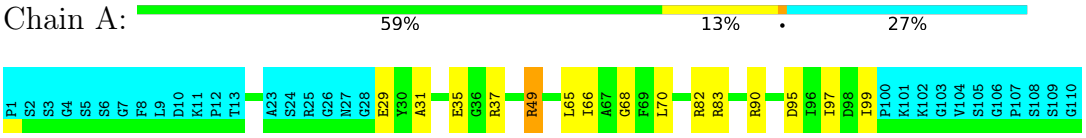
Chain A:



4.2.18 Score per residue for model 18

• Molecule 1: E3 ubiquitin-protein ligase RNF146

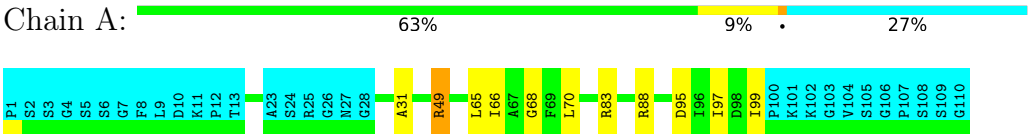
Chain A:



4.2.19 Score per residue for model 19

• Molecule 1: E3 ubiquitin-protein ligase RNF146

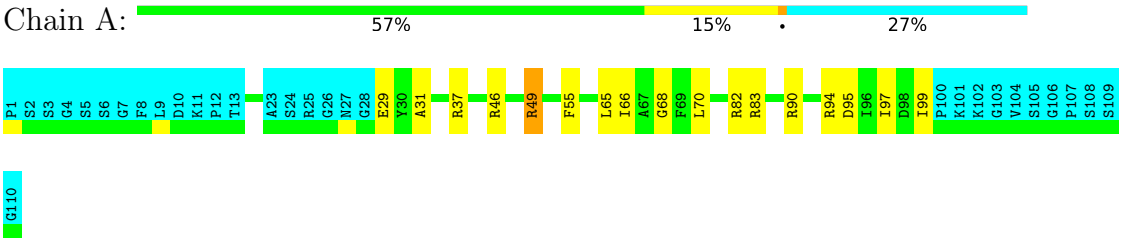
Chain A:



4.2.20 Score per residue for model 20

• Molecule 1: E3 ubiquitin-protein ligase RNF146

Chain A:



5 Refinement protocol and experimental data overview

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

Of the 100 calculated structures, 20 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
CYANA	structure solution	
Amber	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	1270
Number of shifts mapped to atoms	1270
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	89%

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: ATP

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	0.74±0.00	0±0/705 (0.0± 0.0%)	1.08±0.04	5±2/948 (0.6± 0.2%)
All	All	0.74	0/14100 (0.0%)	1.08	106/18960 (0.6%)

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	A	0.0±0.0	0.1±0.4
All	All	0	3

There are no bond-length outliers.

All 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
1	A	90	ARG	NE-CZ-NH1	9.39	125.00	120.30	6	17
1	A	88	ARG	NE-CZ-NH1	8.43	124.52	120.30	5	14
1	A	90	ARG	NE-CZ-NH2	-8.03	116.28	120.30	10	8
1	A	37	ARG	NE-CZ-NH1	7.96	124.28	120.30	18	4
1	A	82	ARG	NE-CZ-NH1	7.90	124.25	120.30	5	12
1	A	88	ARG	NE-CZ-NH2	-7.03	116.79	120.30	9	3
1	A	94	ARG	NE-CZ-NH1	6.79	123.69	120.30	20	1
1	A	83	ARG	NE-CZ-NH2	-6.74	116.93	120.30	20	4
1	A	83	ARG	NE-CZ-NH1	6.37	123.49	120.30	4	19
1	A	49	ARG	NE-CZ-NH1	6.16	123.38	120.30	13	16
1	A	46	ARG	NE-CZ-NH2	-6.12	117.24	120.30	9	3
1	A	82	ARG	NE-CZ-NH2	-5.79	117.41	120.30	2	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	46	ARG	NE-CZ-NH1	5.66	123.13	120.30	20	3
1	A	37	ARG	NE-CZ-NH2	-5.65	117.48	120.30	17	1

There are no chirality outliers.

All unique planar outliers are listed below.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	49	ARG	Sidechain	3

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	689	668	668	1±0
All	All	14400	13600	13600	20

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:31:ALA:HB2	1:A:97:ILE:HD11	0.53	1.79	3	20

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.

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	80/110 (73%)	76±1 (94±1%)	2±1 (3±1%)	2±0 (3±0%)	8 44
All	All	1600/2200 (73%)	1511 (94%)	47 (3%)	42 (3%)	8 44

All 3 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	68	GLY	20
1	A	99	ILE	20
1	A	66	ILE	2

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	70/92 (76%)	63±1 (90±2%)	7±1 (10±2%)	10 55
All	All	1400/1840 (76%)	1253 (90%)	147 (10%)	10 55

All 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	49	ARG	20
1	A	65	LEU	20
1	A	66	ILE	20
1	A	70	LEU	20
1	A	95	ASP	19
1	A	29	GLU	16
1	A	78	MET	8
1	A	80	GLN	5
1	A	55	PHE	4
1	A	91	LYS	4
1	A	88	ARG	3
1	A	59	LYS	2
1	A	85	GLU	2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	46	ARG	1
1	A	37	ARG	1
1	A	21	LYS	1
1	A	35	GLU	1

6.3.3 RNA

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains

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

6.5 Carbohydrates

There are no monosaccharides in this entry.

6.6 Ligand geometry

1 ligand is 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	ATP	A	201	-	26,33,33	0.92±0.02	0±0 (0±1%)

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	ATP	A	201	-	31,52,52	1.13±0.09	2±1 (7±2%)

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	ATP	A	201	-	-	0±0,18,38,38	0±0,3,3,3

All unique bond outliers are listed below.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
2	A	201	ATP	C8-N7	2.14	1.30	1.34	16	5

All 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
2	A	201	ATP	C5-C6-N6	3.61	125.84	120.35	11	20
2	A	201	ATP	PB-O3B-PG	3.57	145.06	132.83	20	5
2	A	201	ATP	N6-C6-N1	3.00	112.35	118.57	11	17
2	A	201	ATP	PA-O3A-PB	2.61	123.87	132.83	6	2
2	A	201	ATP	O4'-C1'-C2'	2.23	103.67	106.93	2	1

There are no chirality outliers.

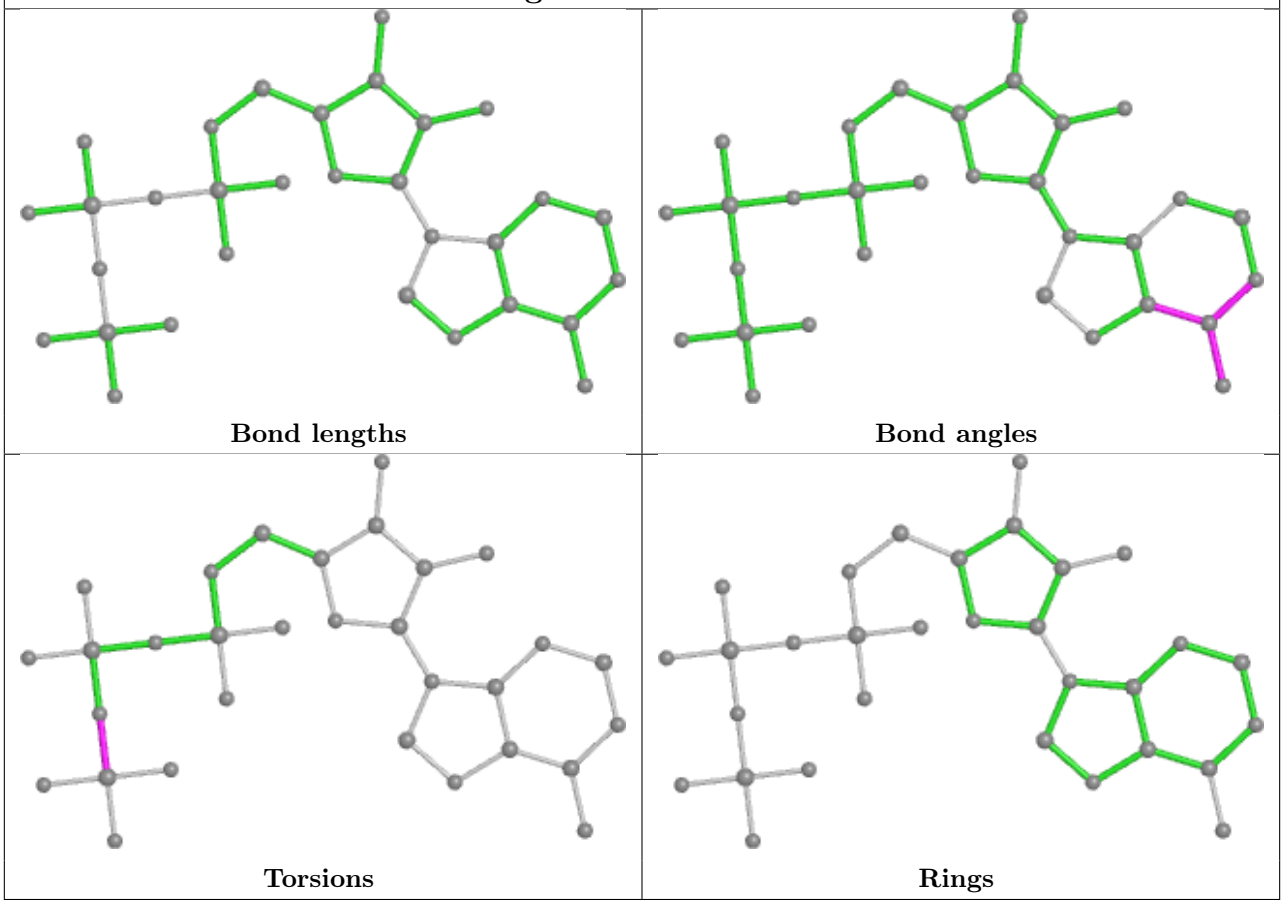
There are no torsion outliers.

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 ATP A 201



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 89% for the well-defined parts and 84% 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

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	1270
Number of shifts mapped to atoms	1270
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	9

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	98	-0.05 ± 0.05	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	89	-0.03 ± 0.17	None needed (< 0.5 ppm)
$^{13}C'$	97	0.09 ± 0.14	None needed (< 0.5 ppm)
^{15}N	95	-0.32 ± 0.41	None needed (< 0.5 ppm)

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 89%, i.e. 1054 atoms were assigned a chemical shift out of a possible 1180. 0 out of 9 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	403/403 (100%)	164/164 (100%)	160/160 (100%)	79/79 (100%)
Sidechain	551/660 (83%)	374/420 (89%)	171/201 (85%)	6/39 (15%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	100/117 (85%)	50/56 (89%)	47/57 (82%)	3/4 (75%)
Overall	1054/1180 (89%)	588/640 (92%)	378/418 (90%)	88/122 (72%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 84%, i.e. 1262 atoms were assigned a chemical shift out of a possible 1503. 0 out of 11 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	492/552 (89%)	202/227 (89%)	195/220 (89%)	95/105 (90%)
Sidechain	662/824 (80%)	449/526 (85%)	206/252 (82%)	7/46 (15%)
Aromatic	108/127 (85%)	54/61 (89%)	51/62 (82%)	3/4 (75%)
Overall	1262/1503 (84%)	705/814 (87%)	452/534 (85%)	105/155 (68%)

7.1.4 Statistically unusual chemical shifts (i)

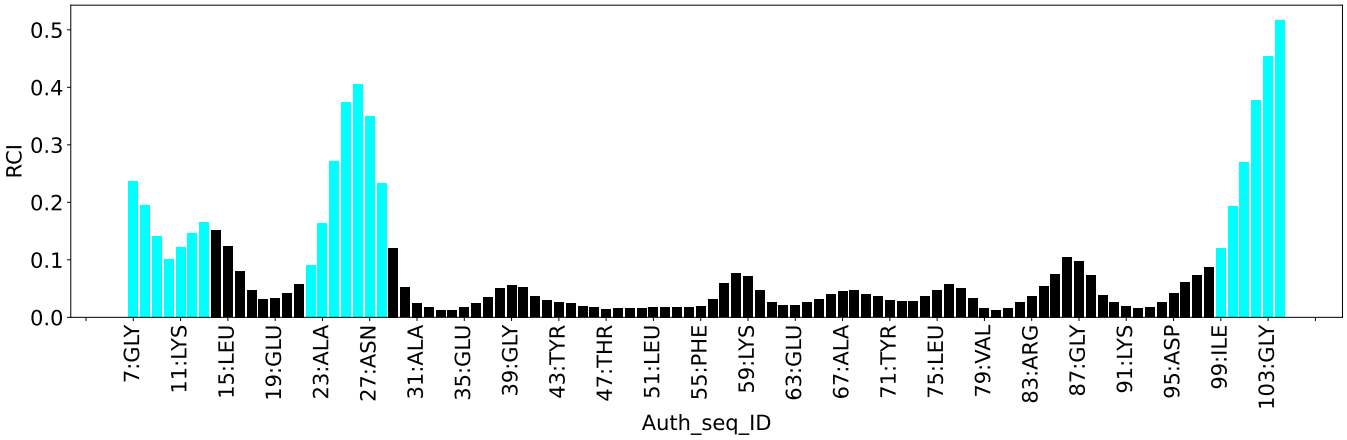
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	93	LYS	HE3	0.57	1.92 – 3.89	-11.9
1	A	94	ARG	HB3	-1.06	0.43 – 3.11	-10.6
1	A	93	LYS	HE2	1.28	1.95 – 3.88	-8.4
1	A	51	LEU	HB3	-1.29	-0.26 – 3.31	-7.9
1	A	93	LYS	HD2	0.08	0.58 – 2.64	-7.4
1	A	94	ARG	HG2	-0.32	0.26 – 2.87	-7.2
1	A	94	ARG	HD2	1.64	1.97 – 4.26	-6.4
1	A	94	ARG	HB2	0.42	0.52 – 3.08	-5.4
1	A	94	ARG	HD3	1.74	1.81 – 4.39	-5.3

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:


RCI
0.0
0.1
0.2
0.3
0.4
0.5
7:GLY
11:LYS
15:LEU
19:GLU
23:ALA
27:ASN
31:ALA
35:GLU
39:GLY
43:TYR
47:THR
51:LEU
55:PHE
59:LYS
63:GLU
67:ALA
71:TYR
75:LEU
79:VAL
83:ARG
87:GLY
91:LYS
95:ASP
99:ILE
103:GLY
Auth_seq_ID

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	1527
Intra-residue (|i-j|=0)	345
Sequential (|i-j|=1)	456
Medium range (|i-j|>1 and |i-j|<5)	228
Long range (|i-j|≥5)	498
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	13.9
Number of long range restraints per residue[1]	4.5

[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

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.

Bins (Å)	Average number of violations per model	Max (Å)
0.1-0.2 (Small)	0.1	0.11
0.2-0.5 (Medium)	None	None
>0.5 (Large)	None	None

8.2.2 Average number of dihedral-angle violations per model (i)

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

9 Distance violation analysis (i)

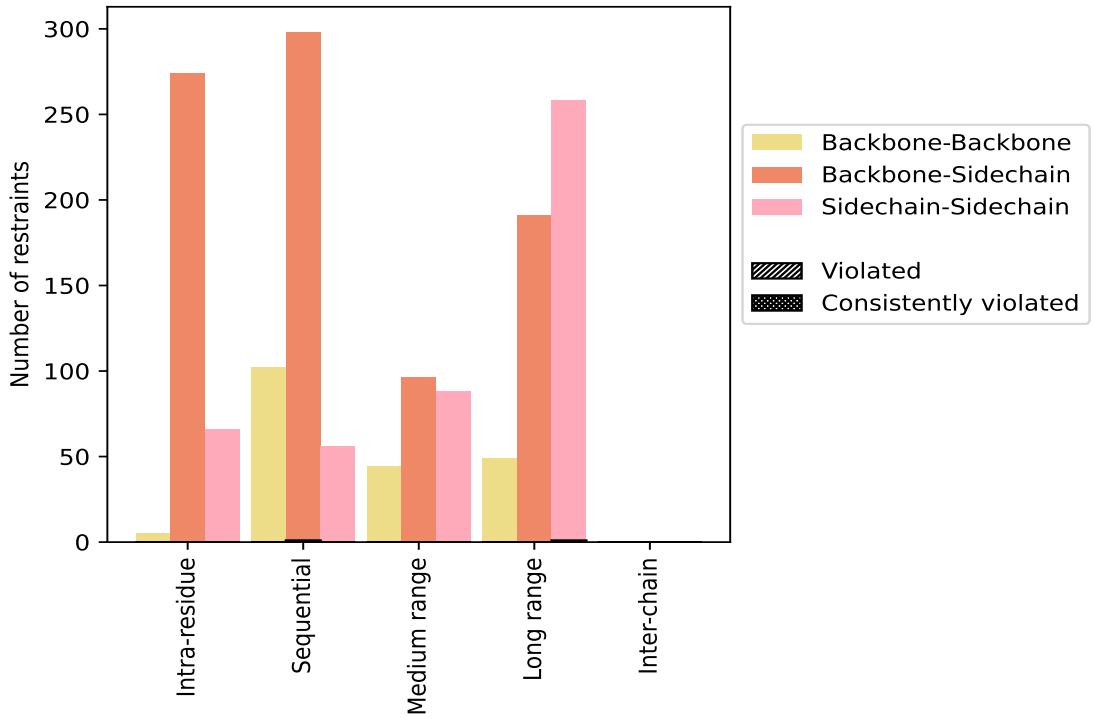
9.1 Summary of distance violations (i)

The following table shows the summary of distance violations in different restraint categories based on the sequence separation of the atoms involved. Each category is further sub-divided into three sub-categories based on the atoms involved. Violations less than 0.1 Å are not included in the statistics.

Restraints type	Count	%[1]	Violated[3]			Consistently Violated[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
Intra-residue (|i-j|=0)	345	22.6	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	5	0.3	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	274	17.9	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	66	4.3	0	0.0	0.0	0	0.0	0.0
Sequential (|i-j|=1)	456	29.9	1	0.2	0.1	0	0.0	0.0
Backbone-Backbone	102	6.7	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	298	19.5	1	0.3	0.1	0	0.0	0.0
Sidechain-Sidechain	56	3.7	0	0.0	0.0	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	228	14.9	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	44	2.9	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	96	6.3	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	88	5.8	0	0.0	0.0	0	0.0	0.0
Long range (|i-j|≥5)	498	32.6	1	0.2	0.1	0	0.0	0.0
Backbone-Backbone	49	3.2	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	191	12.5	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	258	16.9	1	0.4	0.1	0	0.0	0.0
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	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	1527	100.0	2	0.1	0.1	0	0.0	0.0
Backbone-Backbone	200	13.1	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	859	56.3	1	0.1	0.1	0	0.0	0.0
Sidechain-Sidechain	468	30.6	1	0.2	0.1	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. The hydrogen bonds and disulfied bonds are counted in their appropriate category on the x-axis

9.2 Distance violation statistics for each model

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

Model ID	Number of violations						Mean (Å)	Max (Å)	SD[6] (Å)	Median (Å)
	IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total				
1	0	0	0	0	0	0	0.0	0.0	0.0	0.0
2	0	0	0	0	0	0	0.0	0.0	0.0	0.0
3	0	0	0	0	0	0	0.0	0.0	0.0	0.0
4	0	0	0	0	0	0	0.0	0.0	0.0	0.0
5	0	0	0	0	0	0	0.0	0.0	0.0	0.0
6	0	0	0	0	0	0	0.0	0.0	0.0	0.0
7	0	0	0	0	0	0	0.0	0.0	0.0	0.0
8	0	0	0	0	0	0	0.0	0.0	0.0	0.0
9	0	0	0	0	0	0	0.0	0.0	0.0	0.0
10	0	0	0	0	0	0	0.0	0.0	0.0	0.0
11	0	0	0	0	0	0	0.0	0.0	0.0	0.0

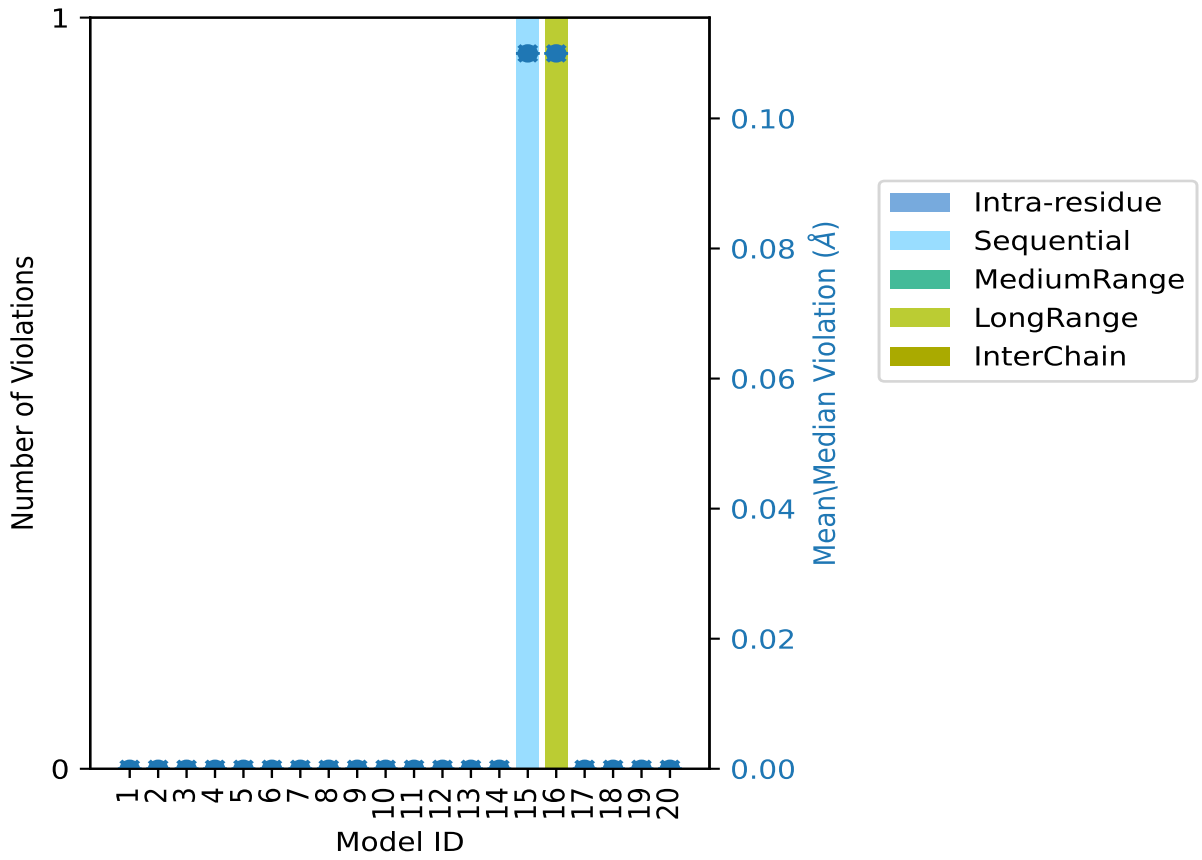
Continued on next page...

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	0	0	0	0	0	0	0.0	0.0	0.0	0.0
13	0	0	0	0	0	0	0.0	0.0	0.0	0.0
14	0	0	0	0	0	0	0.0	0.0	0.0	0.0
15	0	1	0	0	0	1	0.11	0.11	0.0	0.11
16	0	0	0	1	0	1	0.11	0.11	0.0	0.11
17	0	0	0	0	0	0	0.0	0.0	0.0	0.0
18	0	0	0	0	0	0	0.0	0.0	0.0	0.0
19	0	0	0	0	0	0	0.0	0.0	0.0	0.0
20	0	0	0	0	0	0	0.0	0.0	0.0	0.0

[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



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

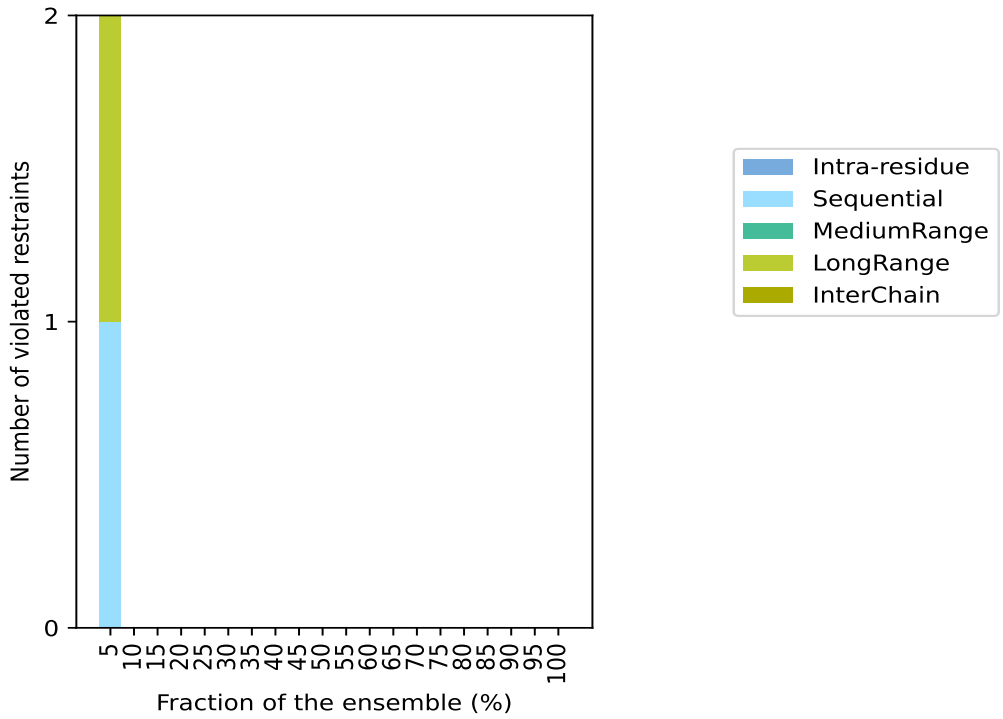
9.3 Distance violation statistics for the ensemble (i)

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, 1525(IR:345, SQ:455, MR:228, LR:497, 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]	%
0	1	0	1	0	2	1	5.0
0	0	0	0	0	0	2	10.0
0	0	0	0	0	0	3	15.0
0	0	0	0	0	0	4	20.0
0	0	0	0	0	0	5	25.0
0	0	0	0	0	0	6	30.0
0	0	0	0	0	0	7	35.0
0	0	0	0	0	0	8	40.0
0	0	0	0	0	0	9	45.0
0	0	0	0	0	0	10	50.0
0	0	0	0	0	0	11	55.0
0	0	0	0	0	0	12	60.0
0	0	0	0	0	0	13	65.0
0	0	0	0	0	0	14	70.0
0	0	0	0	0	0	15	75.0
0	0	0	0	0	0	16	80.0
0	0	0	0	0	0	17	85.0
0	0	0	0	0	0	18	90.0
0	0	0	0	0	0	19	95.0
0	0	0	0	0	0	20	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



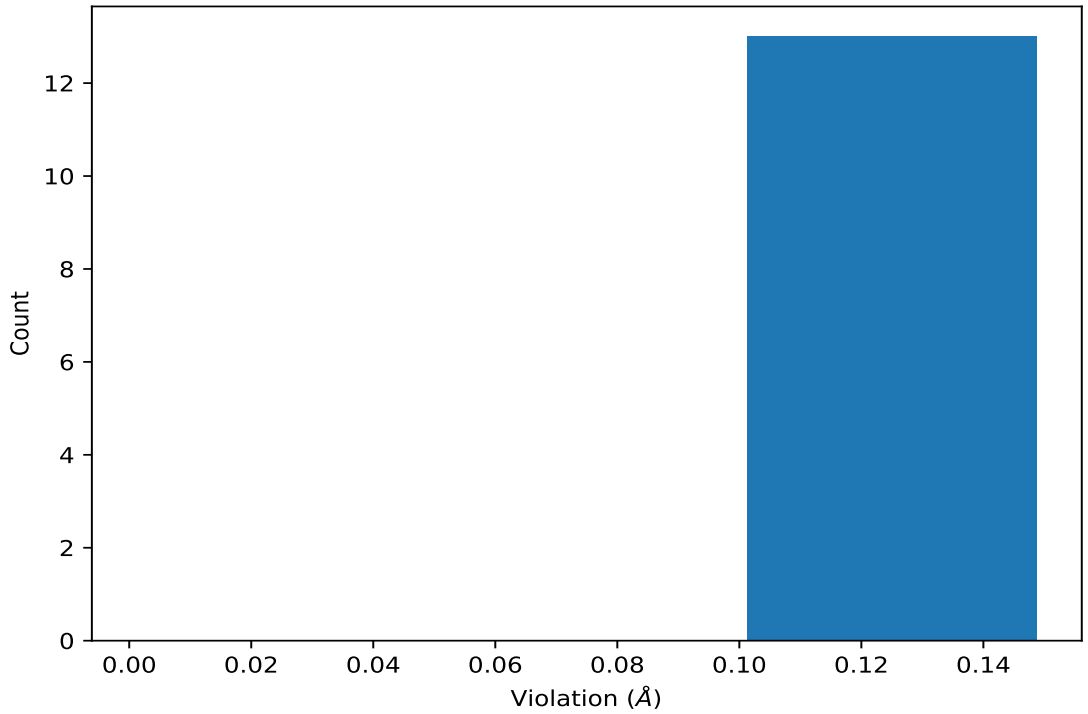
9.4 Most violated distance restraints in the ensemble

No violations found

9.5 All violated distance restraints

9.5.1 Histogram : Distribution of distance violations

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 lists the absolute value of the violation for each restraint in the ensemble sorted by its 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)
(1,519)	1:A:38:ASN:HB3	1:A:39:GLY:H	15	0.11
(1,381)	1:A:32:TRP:HH2	1:A:75:LEU:HD11	16	0.11
(1,381)	1:A:32:TRP:HH2	1:A:75:LEU:HD12	16	0.11
(1,381)	1:A:32:TRP:HH2	1:A:75:LEU:HD13	16	0.11
(1,381)	1:A:32:TRP:HH2	1:A:75:LEU:HD21	16	0.11
(1,381)	1:A:32:TRP:HH2	1:A:75:LEU:HD22	16	0.11
(1,381)	1:A:32:TRP:HH2	1:A:75:LEU:HD23	16	0.11
(1,381)	1:A:32:TRP:HH2	1:A:75:LEU:HD11	16	0.11
(1,381)	1:A:32:TRP:HH2	1:A:75:LEU:HD12	16	0.11
(1,381)	1:A:32:TRP:HH2	1:A:75:LEU:HD13	16	0.11
(1,381)	1:A:32:TRP:HH2	1:A:75:LEU:HD21	16	0.11
(1,381)	1:A:32:TRP:HH2	1:A:75:LEU:HD22	16	0.11
(1,381)	1:A:32:TRP:HH2	1:A:75:LEU:HD23	16	0.11

10 Dihedral-angle violation analysis (i)

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