Full wwPDB NMR Structure Validation Report (i)

Jun 4, 2023 – 03:53 PM EDT

PDB ID	:	2LNX
BMRB ID	:	18183
Title	:	Solution structure of Vav2 SH2 domain
Authors	:	Wu, B.; Zhang, J.; Wu, J.; Shi, Y.
Deposited on	:	2012-01-05

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 (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:

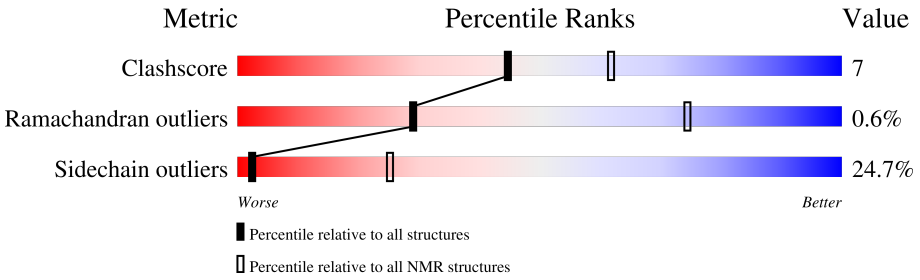
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal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

1 Overall quality at a glance

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

The overall completeness of chemical shifts assignment is 86%.

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	122	56% 22% 15% 7%

2 Ensemble composition and analysis

This entry contains 20 models. Model 19 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:667-A:699, A:706-A:767 (95)	0.73	19

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

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

The models can be grouped into 2 clusters. No single-model clusters were found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1887 atoms, of which 935 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Guanine nucleotide exchange factor VAV2.

Mol	Chain	Residues	Atoms						Trace
1	A	113	Total	C	H	N	O	S	0
			1887	607	935	163	180	2	

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	650	MET	-	expression tag	UNP P52735
A	651	GLY	-	expression tag	UNP P52735
A	652	HIS	-	expression tag	UNP P52735
A	653	HIS	-	expression tag	UNP P52735
A	654	HIS	-	expression tag	UNP P52735
A	655	HIS	-	expression tag	UNP P52735
A	656	HIS	-	expression tag	UNP P52735
A	657	HIS	-	expression tag	UNP P52735
A	658	MET	-	expression tag	UNP P52735

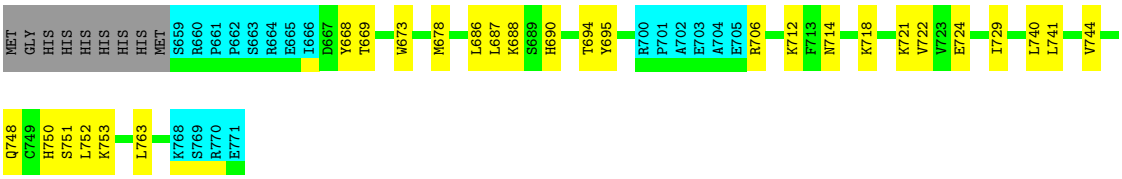
4 Residue-property plots (i)

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: Guanine nucleotide exchange factor VAV2

Chain A:



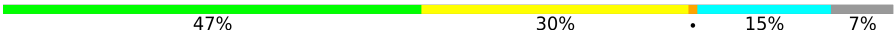
4.2 Scores per residue for each member of the ensemble

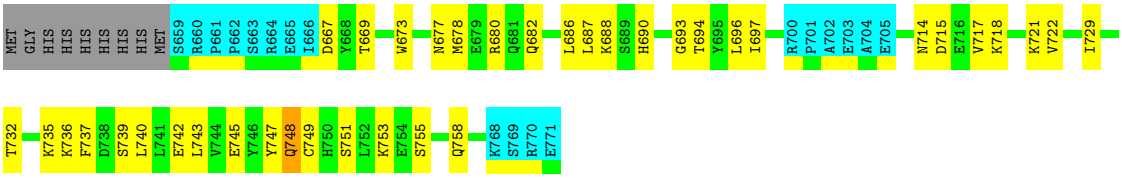
Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

• Molecule 1: Guanine nucleotide exchange factor VAV2

Chain A:

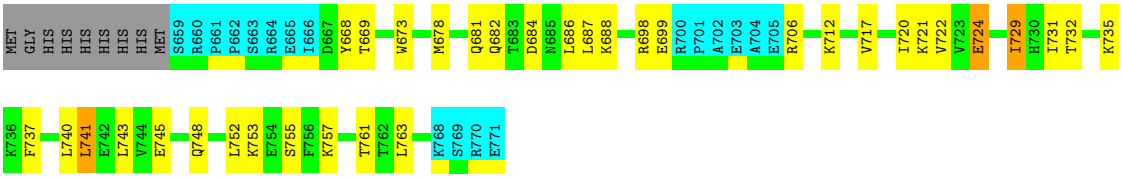




4.2.2 Score per residue for model 2

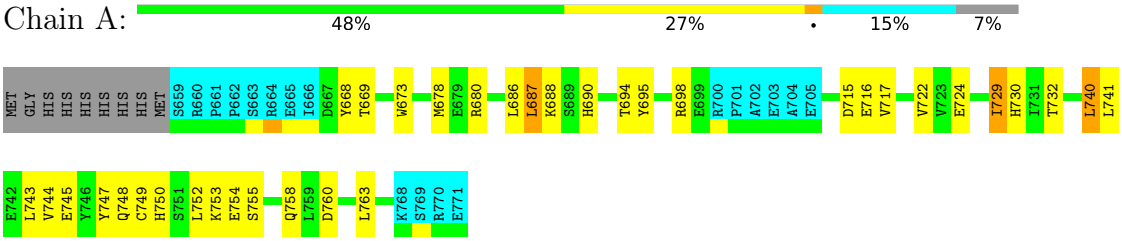
• Molecule 1: Guanine nucleotide exchange factor VAV2

Chain A:



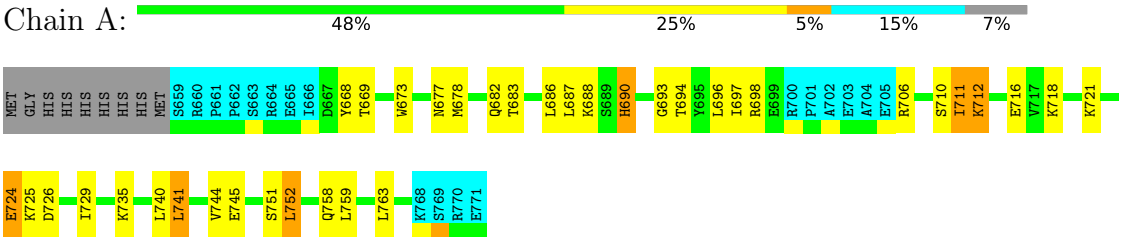
4.2.3 Score per residue for model 3

• Molecule 1: Guanine nucleotide exchange factor VAV2



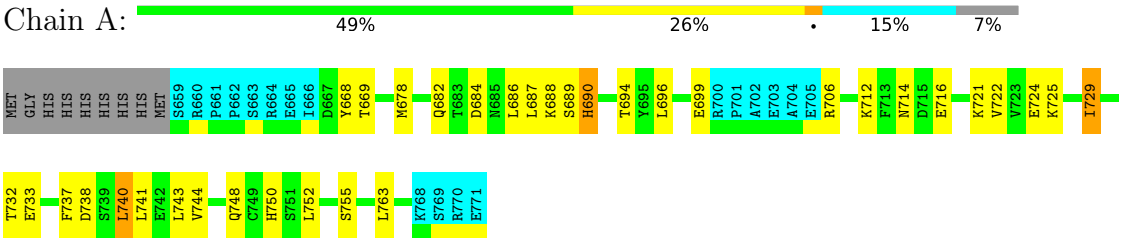
4.2.4 Score per residue for model 4

• Molecule 1: Guanine nucleotide exchange factor VAV2



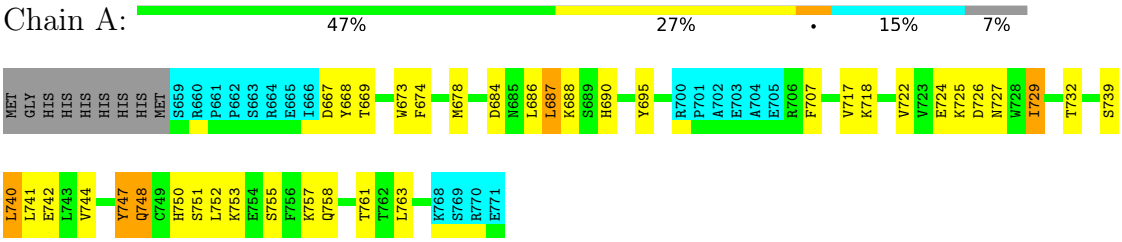
4.2.5 Score per residue for model 5

• Molecule 1: Guanine nucleotide exchange factor VAV2



4.2.6 Score per residue for model 6

• Molecule 1: Guanine nucleotide exchange factor VAV2

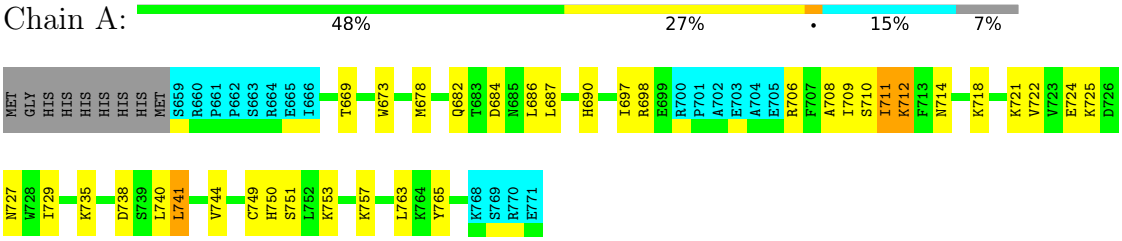


4.2.7 Score per residue for model 7

• Molecule 1: Guanine nucleotide exchange factor VAV2

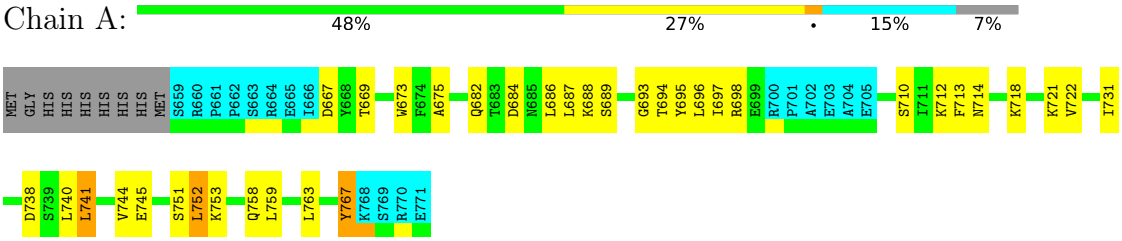
4.2.8 Score per residue for model 8

• Molecule 1: Guanine nucleotide exchange factor VAV2



4.2.9 Score per residue for model 9

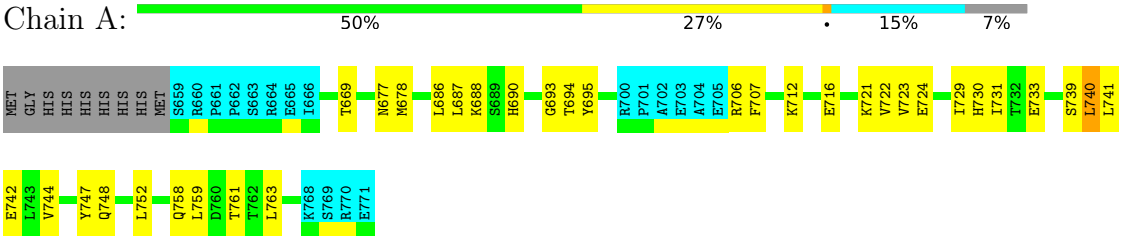
• Molecule 1: Guanine nucleotide exchange factor VAV2



4.2.10 Score per residue for model 10

• Molecule 1: Guanine nucleotide exchange factor VAV2

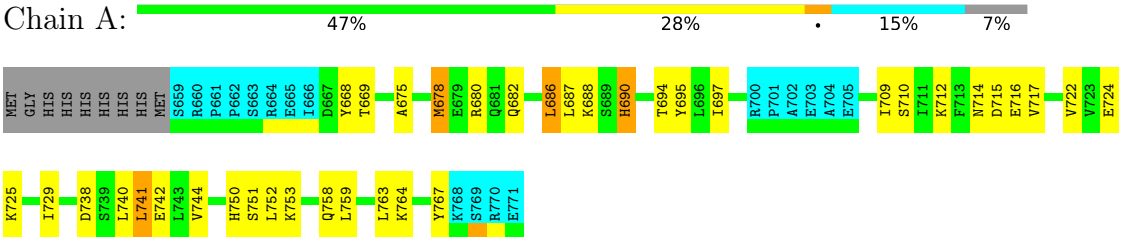
Chain A:



4.2.11 Score per residue for model 11

• Molecule 1: Guanine nucleotide exchange factor VAV2

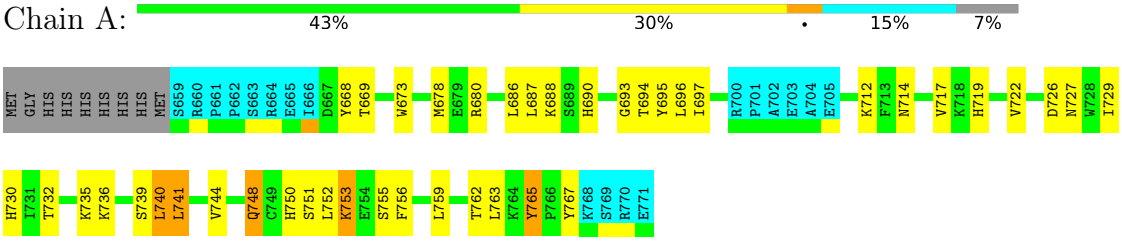
Chain A:



4.2.12 Score per residue for model 12

• Molecule 1: Guanine nucleotide exchange factor VAV2

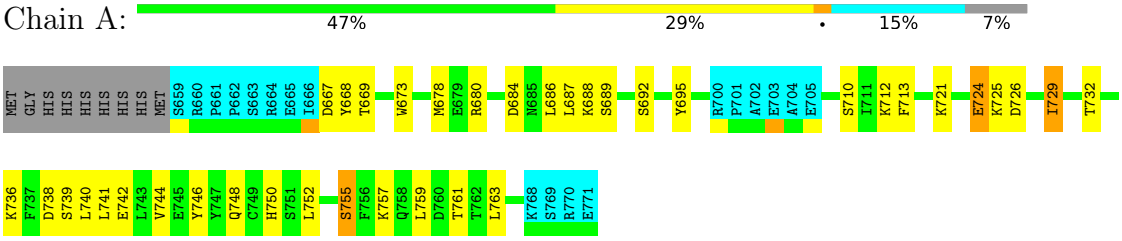
Chain A:



4.2.13 Score per residue for model 13

• Molecule 1: Guanine nucleotide exchange factor VAV2

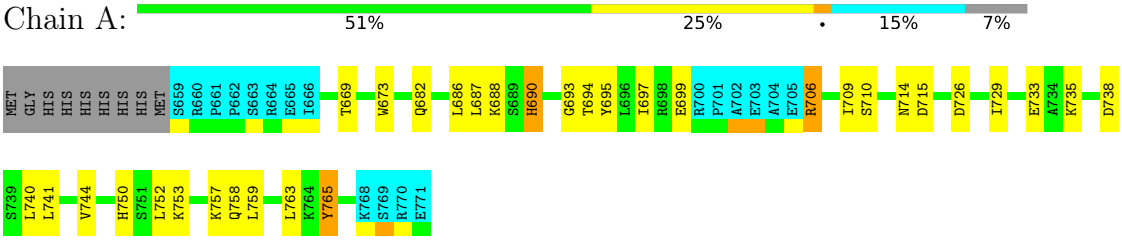
Chain A:



4.2.14 Score per residue for model 14

• Molecule 1: Guanine nucleotide exchange factor VAV2

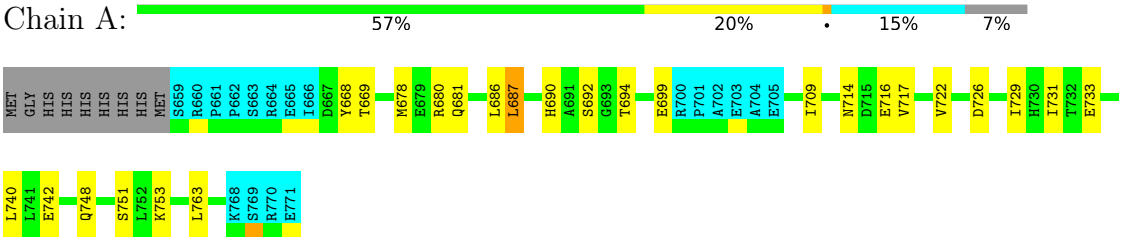
Chain A:



4.2.15 Score per residue for model 15

• Molecule 1: Guanine nucleotide exchange factor VAV2

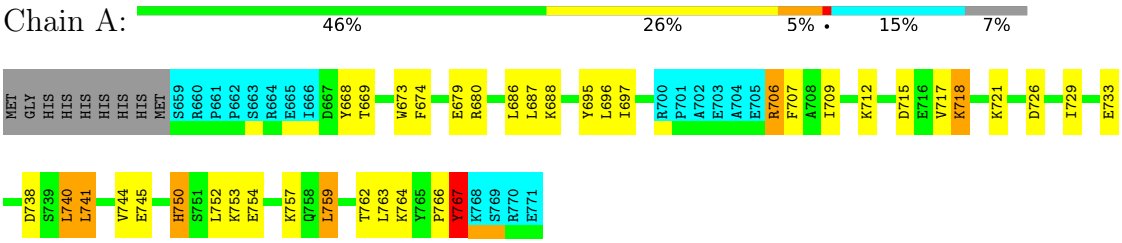
Chain A:



4.2.16 Score per residue for model 16

• Molecule 1: Guanine nucleotide exchange factor VAV2

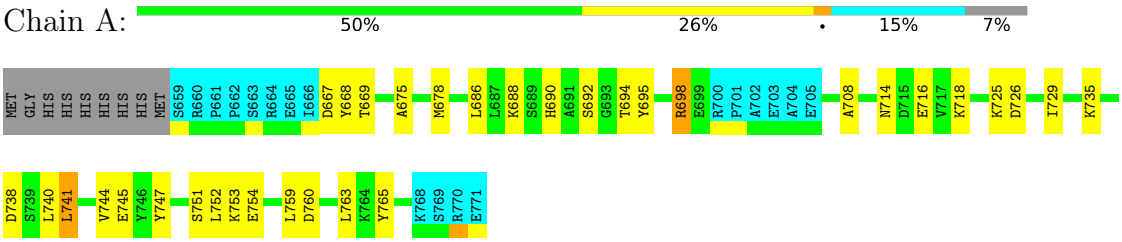
Chain A:



4.2.17 Score per residue for model 17

• Molecule 1: Guanine nucleotide exchange factor VAV2

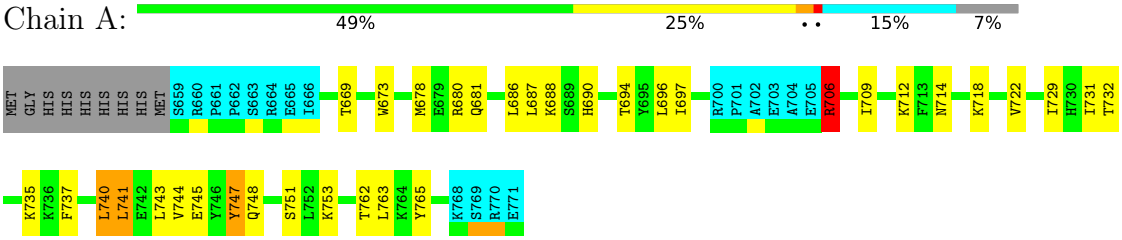
Chain A:



4.2.18 Score per residue for model 18

• Molecule 1: Guanine nucleotide exchange factor VAV2

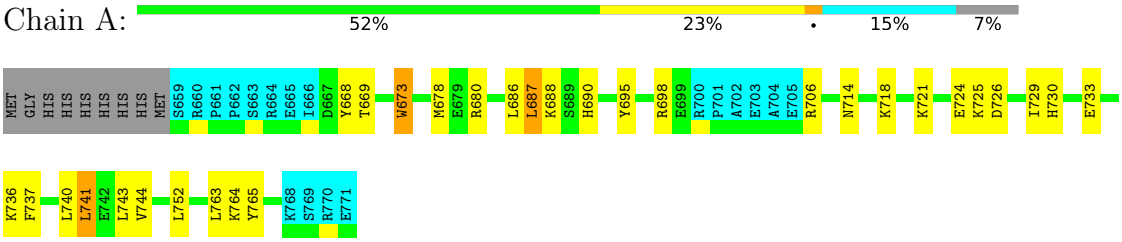
Chain A:



4.2.19 Score per residue for model 19 (medoid)

• Molecule 1: Guanine nucleotide exchange factor VAV2

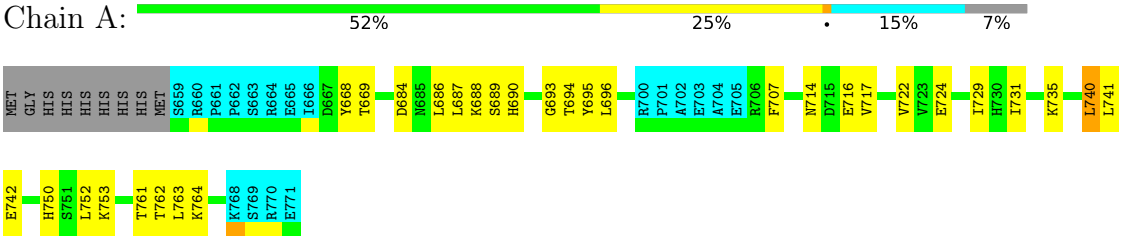
Chain A:



4.2.20 Score per residue for model 20

• Molecule 1: Guanine nucleotide exchange factor VAV2

Chain A:



5 Refinement protocol and experimental data overview

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

Of the 200 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
CNS	structure solution	
CNS	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	1368
Number of shifts mapped to atoms	1368
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	86%

6 Model quality

6.1 Standard geometry

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

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	805	789	785	11±3
All	All	16100	15780	15700	215

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 7.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:722:VAL:HG12	1:A:729:ILE:HD11	0.75	1.57	3	8
1:A:673:TRP:CE3	1:A:697:ILE:HD13	0.75	2.16	9	6
1:A:724:GLU:HB2	1:A:729:ILE:HD12	0.68	1.63	13	5
1:A:748:GLN:HA	1:A:763:LEU:HD13	0.68	1.65	7	1
1:A:732:THR:HG21	1:A:755:SER:HB2	0.67	1.67	2	2
1:A:748:GLN:CG	1:A:763:LEU:HD22	0.67	2.20	6	2
1:A:748:GLN:HG2	1:A:763:LEU:HD22	0.66	1.66	6	3
1:A:748:GLN:HG3	1:A:763:LEU:HD12	0.65	1.68	15	1
1:A:750:HIS:O	1:A:762:THR:HG22	0.65	1.92	7	3
1:A:687:LEU:HD12	1:A:712:LYS:HB2	0.64	1.68	18	1
1:A:675:ALA:HB1	1:A:678:MET:CE	0.63	2.24	17	1
1:A:697:ILE:CD1	1:A:709:ILE:HG23	0.63	2.24	11	2
1:A:732:THR:HG21	1:A:755:SER:OG	0.62	1.94	1	5
1:A:675:ALA:HB1	1:A:678:MET:SD	0.62	2.35	11	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:667:ASP:HB3	1:A:670:ALA:HB3	0.62	1.70	7	1
1:A:697:ILE:HD13	1:A:709:ILE:HG23	0.61	1.73	11	2
1:A:722:VAL:HG22	1:A:731:ILE:HG22	0.61	1.73	18	4
1:A:740:LEU:O	1:A:744:VAL:HG23	0.61	1.96	18	10
1:A:752:LEU:HD11	1:A:759:LEU:HD23	0.60	1.71	10	3
1:A:687:LEU:HD22	1:A:712:LYS:HB2	0.60	1.71	12	4
1:A:748:GLN:CG	1:A:763:LEU:HD12	0.59	2.28	15	2
1:A:690:HIS:CD2	1:A:694:THR:HG21	0.58	2.32	12	6
1:A:737:PHE:CD1	1:A:743:LEU:HD12	0.58	2.34	18	3
1:A:687:LEU:HD12	1:A:717:VAL:HG13	0.58	1.75	1	3
1:A:695:TYR:CG	1:A:763:LEU:HD22	0.57	2.34	3	4
1:A:695:TYR:CD2	1:A:763:LEU:HD23	0.57	2.34	14	2
1:A:695:TYR:CE1	1:A:763:LEU:HD23	0.56	2.35	6	1
1:A:697:ILE:HG13	1:A:709:ILE:HG23	0.56	1.77	14	3
1:A:695:TYR:CE2	1:A:763:LEU:HD23	0.56	2.36	20	4
1:A:696:LEU:HD22	1:A:767:TYR:CE1	0.55	2.37	16	1
1:A:748:GLN:CB	1:A:763:LEU:HD13	0.55	2.32	18	1
1:A:709:ILE:HD12	1:A:731:ILE:HG21	0.55	1.77	15	1
1:A:687:LEU:HD23	1:A:712:LYS:HB2	0.55	1.78	13	2
1:A:697:ILE:HD12	1:A:709:ILE:HG12	0.55	1.79	11	2
1:A:748:GLN:HG2	1:A:763:LEU:HD12	0.54	1.78	5	2
1:A:752:LEU:HD23	1:A:761:THR:O	0.53	2.03	10	2
1:A:673:TRP:CZ3	1:A:744:VAL:HG11	0.53	2.38	4	1
1:A:722:VAL:CG1	1:A:729:ILE:HD11	0.53	2.33	6	5
1:A:707:PHE:CD2	1:A:740:LEU:HD11	0.53	2.39	20	1
1:A:687:LEU:HD13	1:A:717:VAL:HG13	0.53	1.81	3	3
1:A:675:ALA:HB2	1:A:767:TYR:OH	0.51	2.06	9	1
1:A:722:VAL:HG22	1:A:731:ILE:CG2	0.50	2.36	20	3
1:A:744:VAL:HG13	1:A:763:LEU:CD1	0.50	2.35	16	2
1:A:722:VAL:HA	1:A:731:ILE:HG22	0.49	1.82	10	1
1:A:690:HIS:HD2	1:A:694:THR:HG21	0.49	1.67	12	1
1:A:687:LEU:HD23	1:A:690:HIS:CD2	0.49	2.42	15	3
1:A:748:GLN:CB	1:A:763:LEU:HD12	0.49	2.37	3	1
1:A:687:LEU:HD11	1:A:696:LEU:HD23	0.49	1.84	4	2
1:A:673:TRP:CE3	1:A:744:VAL:HG11	0.49	2.42	19	1
1:A:748:GLN:HB3	1:A:763:LEU:HD12	0.49	1.85	3	1
1:A:751:SER:HA	1:A:762:THR:HG22	0.48	1.85	18	1
1:A:690:HIS:CG	1:A:694:THR:HG21	0.48	2.43	10	2
1:A:687:LEU:HD21	1:A:696:LEU:HD23	0.48	1.85	1	1
1:A:744:VAL:CG1	1:A:763:LEU:HD21	0.48	2.39	13	1
1:A:687:LEU:HD13	1:A:690:HIS:CB	0.48	2.39	18	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:695:TYR:CB	1:A:763:LEU:HD22	0.48	2.39	16	1
1:A:751:SER:N	1:A:762:THR:HG22	0.48	2.23	12	1
1:A:695:TYR:CE1	1:A:763:LEU:HB3	0.47	2.45	12	2
1:A:687:LEU:HD11	1:A:696:LEU:HD21	0.47	1.86	5	1
1:A:694:THR:HA	1:A:765:TYR:O	0.47	2.10	12	3
1:A:718:LYS:HE3	1:A:759:LEU:HD13	0.47	1.86	17	1
1:A:684:ASP:HA	1:A:717:VAL:HG21	0.47	1.86	20	1
1:A:752:LEU:CD2	1:A:759:LEU:HD23	0.47	2.39	11	1
1:A:711:ILE:HD13	1:A:712:LYS:N	0.47	2.25	8	2
1:A:752:LEU:HD12	1:A:761:THR:O	0.47	2.10	13	2
1:A:687:LEU:HD22	1:A:694:THR:HB	0.47	1.85	15	3
1:A:696:LEU:HD12	1:A:697:ILE:N	0.47	2.25	4	1
1:A:744:VAL:HG13	1:A:763:LEU:CD2	0.46	2.40	14	1
1:A:673:TRP:CZ3	1:A:744:VAL:CG1	0.46	2.98	4	1
1:A:673:TRP:CH2	1:A:744:VAL:HG21	0.46	2.45	8	1
1:A:687:LEU:HD11	1:A:694:THR:CG2	0.46	2.40	18	1
1:A:752:LEU:HD13	1:A:761:THR:O	0.46	2.11	20	1
1:A:722:VAL:HG12	1:A:729:ILE:CD1	0.46	2.41	8	2
1:A:698:ARG:HG3	1:A:708:ALA:HB3	0.45	1.88	8	2
1:A:711:ILE:HD13	1:A:712:LYS:H	0.45	1.72	4	2
1:A:696:LEU:HD22	1:A:767:TYR:CZ	0.45	2.47	12	1
1:A:696:LEU:HB2	1:A:767:TYR:CE2	0.44	2.46	9	1
1:A:687:LEU:HD12	1:A:717:VAL:CG1	0.44	2.42	1	1
1:A:695:TYR:CD2	1:A:763:LEU:HD22	0.44	2.47	3	1
1:A:747:TYR:CZ	1:A:763:LEU:HD11	0.44	2.48	6	1
1:A:750:HIS:C	1:A:762:THR:HG22	0.44	2.33	12	1
1:A:737:PHE:CD1	1:A:743:LEU:HA	0.44	2.48	2	1
1:A:763:LEU:HD12	1:A:763:LEU:N	0.44	2.27	14	1
1:A:673:TRP:HE3	1:A:697:ILE:HD13	0.44	1.73	12	1
1:A:744:VAL:HG13	1:A:763:LEU:HD22	0.44	1.89	17	1
1:A:741:LEU:HD12	1:A:741:LEU:C	0.43	2.33	7	10
1:A:722:VAL:HG12	1:A:729:ILE:CG1	0.43	2.43	8	1
1:A:737:PHE:CD1	1:A:743:LEU:HG	0.43	2.48	1	1
1:A:747:TYR:CD1	1:A:752:LEU:HD23	0.43	2.48	17	1
1:A:763:LEU:HD22	1:A:763:LEU:N	0.43	2.28	4	2
1:A:717:VAL:HG12	1:A:719:HIS:CE1	0.43	2.49	12	1
1:A:687:LEU:HD22	1:A:712:LYS:CB	0.42	2.42	12	1
1:A:669:THR:HG22	1:A:674:PHE:CE2	0.42	2.49	16	1
1:A:687:LEU:CD1	1:A:717:VAL:HG13	0.42	2.45	2	1
1:A:696:LEU:CB	1:A:767:TYR:CE2	0.42	3.02	9	1
1:A:709:ILE:HD12	1:A:731:ILE:CG2	0.42	2.44	15	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:687:LEU:HD12	1:A:717:VAL:CG2	0.42	2.45	15	1
1:A:687:LEU:HG	1:A:696:LEU:HD21	0.42	1.91	12	1
1:A:731:ILE:HG13	1:A:732:THR:HG22	0.42	1.90	18	1
1:A:745:GLU:HA	1:A:748:GLN:HG2	0.41	1.91	1	1
1:A:707:PHE:CD2	1:A:740:LEU:HD13	0.41	2.50	6	1
1:A:673:TRP:HB3	1:A:695:TYR:CE2	0.41	2.50	14	1
1:A:720:ILE:HG21	1:A:731:ILE:HD12	0.41	1.93	2	1
1:A:752:LEU:HD12	1:A:759:LEU:CD2	0.41	2.46	16	1
1:A:695:TYR:CG	1:A:763:LEU:HD23	0.41	2.51	11	1
1:A:753:LYS:HA	1:A:756:PHE:O	0.41	2.16	12	1
1:A:744:VAL:HA	1:A:747:TYR:CE1	0.41	2.51	18	1
1:A:687:LEU:HD22	1:A:694:THR:CG2	0.41	2.46	20	1
1:A:751:SER:CA	1:A:762:THR:HG22	0.41	2.45	18	1
1:A:729:ILE:HG23	1:A:743:LEU:HD23	0.40	1.93	3	1
1:A:687:LEU:HD21	1:A:694:THR:HG22	0.40	1.93	9	1
1:A:718:LYS:NZ	1:A:759:LEU:HD11	0.40	2.31	16	1
1:A:752:LEU:HD22	1:A:759:LEU:HD23	0.40	1.92	4	1
1:A:686:LEU:HD12	1:A:767:TYR:HB2	0.40	1.93	11	1

6.3 Torsion angles (i)

6.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the backbone conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	95/122 (78%)	87±1 (92±2%)	7±1 (7±1%)	1±1 (1±1%)	29	74
All	All	1900/2440 (78%)	1748 (92%)	140 (7%)	12 (1%)	29	74

All 4 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	693	GLY	7
1	A	767	TYR	2
1	A	706	ARG	2
1	A	673	TRP	1

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	88/112 (79%)	66±4 (75±4%)	22±4 (25±4%)	2 25
All	All	1760/2240 (79%)	1326 (75%)	434 (25%)	2 25

All 64 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	686	LEU	20
1	A	669	THR	19
1	A	688	LYS	18
1	A	740	LEU	17
1	A	741	LEU	16
1	A	678	MET	14
1	A	753	LYS	14
1	A	714	ASN	13
1	A	721	LYS	11
1	A	729	ILE	11
1	A	680	ARG	10
1	A	718	LYS	9
1	A	735	LYS	9
1	A	751	SER	9
1	A	706	ARG	9
1	A	724	GLU	9
1	A	690	HIS	9
1	A	726	ASP	9
1	A	682	GLN	8
1	A	758	GLN	8
1	A	716	GLU	8
1	A	750	HIS	8
1	A	725	LYS	8
1	A	738	ASP	8
1	A	742	GLU	7
1	A	684	ASP	7
1	A	745	GLU	7
1	A	752	LEU	7
1	A	667	ASP	6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	715	ASP	6
1	A	747	TYR	6
1	A	698	ARG	6
1	A	757	LYS	6
1	A	710	SER	6
1	A	733	GLU	6
1	A	739	SER	5
1	A	748	GLN	5
1	A	673	TRP	5
1	A	687	LEU	5
1	A	730	HIS	5
1	A	712	LYS	5
1	A	765	TYR	5
1	A	736	LYS	4
1	A	681	GLN	4
1	A	699	GLU	4
1	A	689	SER	4
1	A	764	LYS	4
1	A	677	ASN	3
1	A	749	CYS	3
1	A	754	GLU	3
1	A	760	ASP	3
1	A	727	ASN	3
1	A	759	LEU	3
1	A	692	SER	3
1	A	711	ILE	2
1	A	696	LEU	2
1	A	713	PHE	2
1	A	707	PHE	2
1	A	683	THR	1
1	A	674	PHE	1
1	A	746	TYR	1
1	A	755	SER	1
1	A	679	GLU	1
1	A	767	TYR	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

There are no ligands in this entry.

6.7 Other polymers

There are no such molecules in this entry.

6.8 Polymer linkage issues

There are no chain breaks in this entry.

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 86% 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	1368
Number of shifts mapped to atoms	1368
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

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	105	-0.25 ± 0.07	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	103	0.22 ± 0.16	None needed (< 0.5 ppm)
$^{13}C'$	107	0.43 ± 0.13	None needed (< 0.5 ppm)
^{15}N	106	0.19 ± 0.18	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 86%, i.e. 1182 atoms were assigned a chemical shift out of a possible 1368. 0 out of 13 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	466/473 (99%)	189/190 (99%)	185/190 (97%)	92/93 (99%)
Sidechain	651/726 (90%)	451/468 (96%)	191/231 (83%)	9/27 (33%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	65/169 (38%)	64/81 (79%)	0/78 (0%)	1/10 (10%)
Overall	1182/1368 (86%)	704/739 (95%)	376/499 (75%)	102/130 (78%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 84%, i.e. 1367 atoms were assigned a chemical shift out of a possible 1622. 0 out of 13 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	539/557 (97%)	221/223 (99%)	212/226 (94%)	106/108 (98%)
Sidechain	763/896 (85%)	535/575 (93%)	219/281 (78%)	9/40 (22%)
Aromatic	65/169 (38%)	64/81 (79%)	0/78 (0%)	1/10 (10%)
Overall	1367/1622 (84%)	820/879 (93%)	431/585 (74%)	116/158 (73%)

7.1.4 Statistically unusual chemical shifts (i)

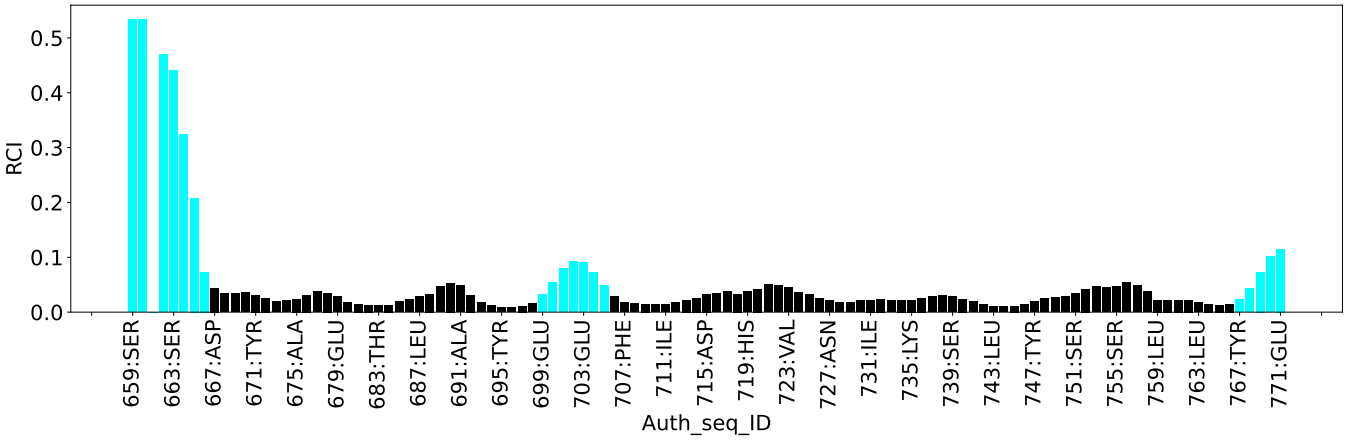
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	743	LEU	HA	1.96	2.04 – 6.55	-5.2

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
659:SER
663:SER
667:ASP
671:TYR
675:ALA
679:GLU
683:THR
687:LEU
691:ALA
695:TYR
699:GLU
703:GLU
707:PHE
711:ILE
715:ASP
719:HIS
723:VAL
727:ASN
731:ILE
735:LYS
739:SER
743:LEU
747:TYR
751:SER
755:SER
759:LEU
763:LEU
767:TYR
771:GLU
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	1644
Intra-residue (|i-j|=0)	496
Sequential (|i-j|=1)	439
Medium range (|i-j|>1 and |i-j|<5)	238
Long range (|i-j|≥5)	421
Inter-chain	0
Hydrogen bond restraints	50
Disulfide bond restraints	0
Total dihedral-angle restraints	102
Number of unmapped restraints	0
Number of restraints per residue	14.3
Number of long range restraints per residue[1]	3.7

[1]Long range hydrogen bonds and disulfide bonds are counted as long range restraints while calculating the number of long range restraints per residue

8.2 Residual restraint violations

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)	0.6	0.17
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)

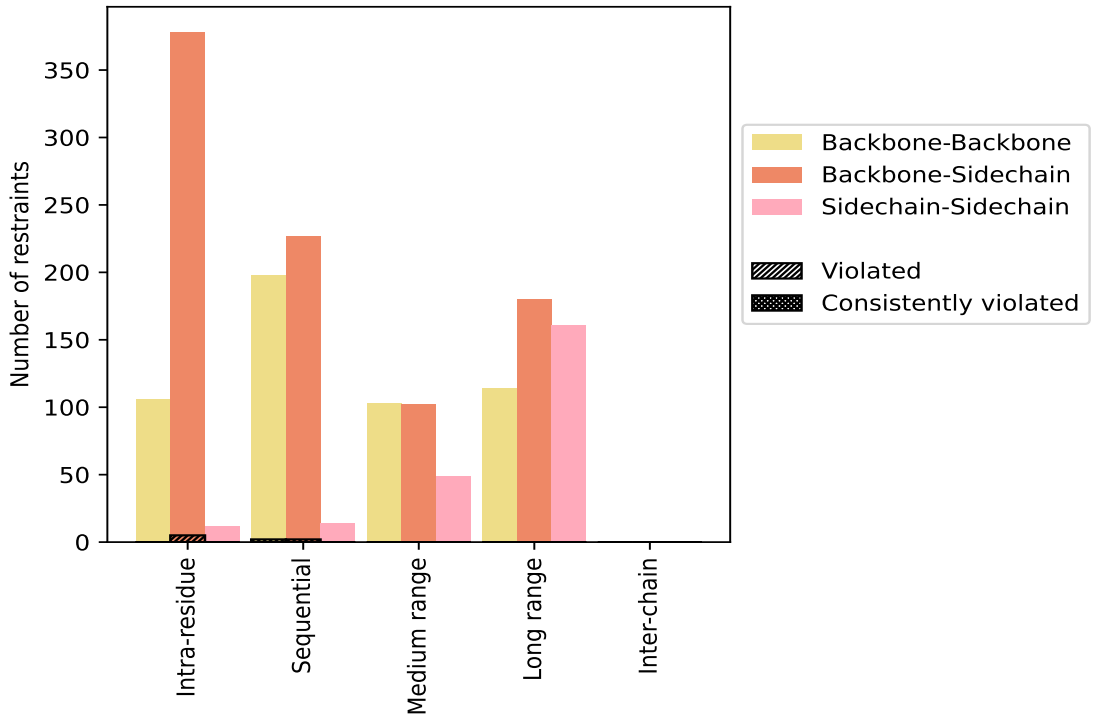
9.1 Summary of distance violations (i)

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)	496	30.2	5	1.0	0.3	0	0.0	0.0
Backbone-Backbone	106	6.4	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	378	23.0	5	1.3	0.3	0	0.0	0.0
Sidechain-Sidechain	12	0.7	0	0.0	0.0	0	0.0	0.0
Sequential (|i-j|=1)	439	26.7	4	0.9	0.2	0	0.0	0.0
Backbone-Backbone	198	12.0	2	1.0	0.1	0	0.0	0.0
Backbone-Sidechain	227	13.8	2	0.9	0.1	0	0.0	0.0
Sidechain-Sidechain	14	0.9	0	0.0	0.0	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	238	14.5	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	87	5.3	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	102	6.2	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	49	3.0	0	0.0	0.0	0	0.0	0.0
Long range (|i-j|≥5)	421	25.6	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	80	4.9	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	180	10.9	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	161	9.8	0	0.0	0.0	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	50	3.0	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	1644	100.0	9	0.5	0.5	0	0.0	0.0
Backbone-Backbone	521	31.7	2	0.4	0.1	0	0.0	0.0
Backbone-Sidechain	887	54.0	7	0.8	0.4	0	0.0	0.0
Sidechain-Sidechain	236	14.4	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. 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	2	0	0	0	2	0.14	0.16	0.02	0.14
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	2	0	0	0	0	2	0.16	0.16	0.0	0.16
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	2	1	0	0	0	3	0.14	0.15	0.01	0.15
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

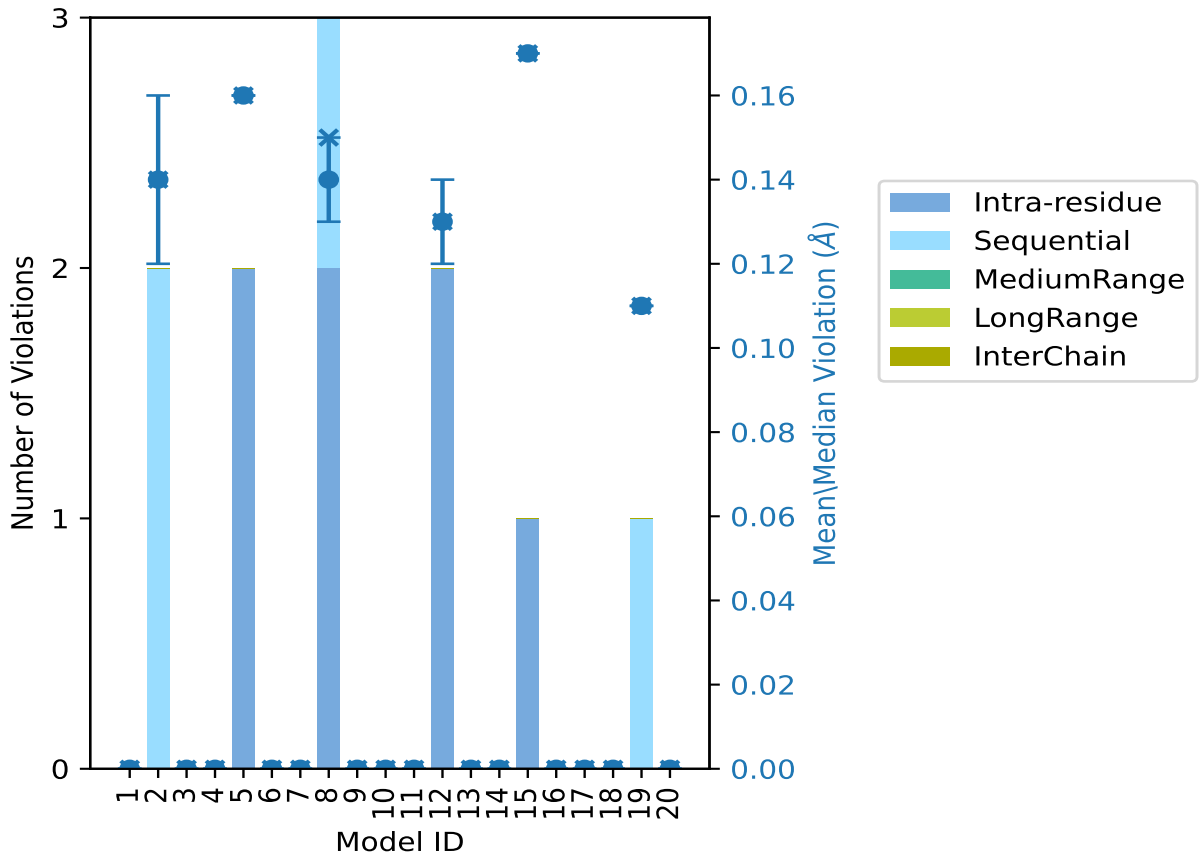
Continued on next page...

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	2	0	0	0	0	2	0.13	0.14	0.01	0.13
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	1	0	0	0	0	1	0.17	0.17	0.0	0.17
16	0	0	0	0	0	0	0.0	0.0	0.0	0.0
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	1	0	0	0	1	0.11	0.11	0.0	0.11
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

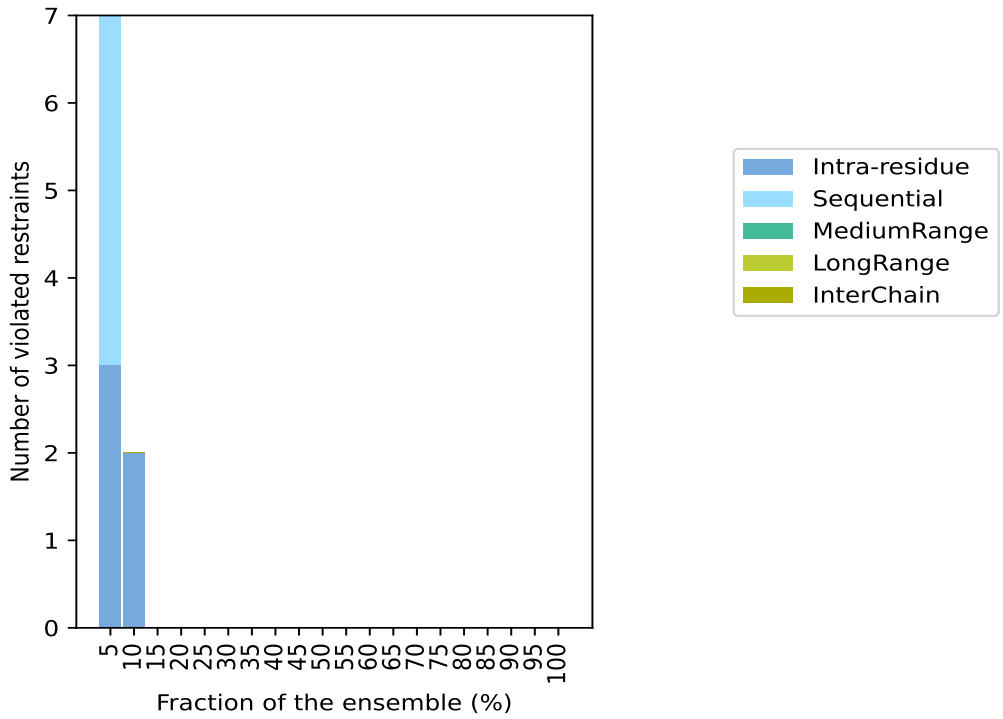
9.3 Distance violation statistics for the ensemble (i)

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, 1585(IR:491, SQ:435, MR:238, LR:421, 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]	%
3	4	0	0	0	7	1	5.0
2	0	0	0	0	2	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

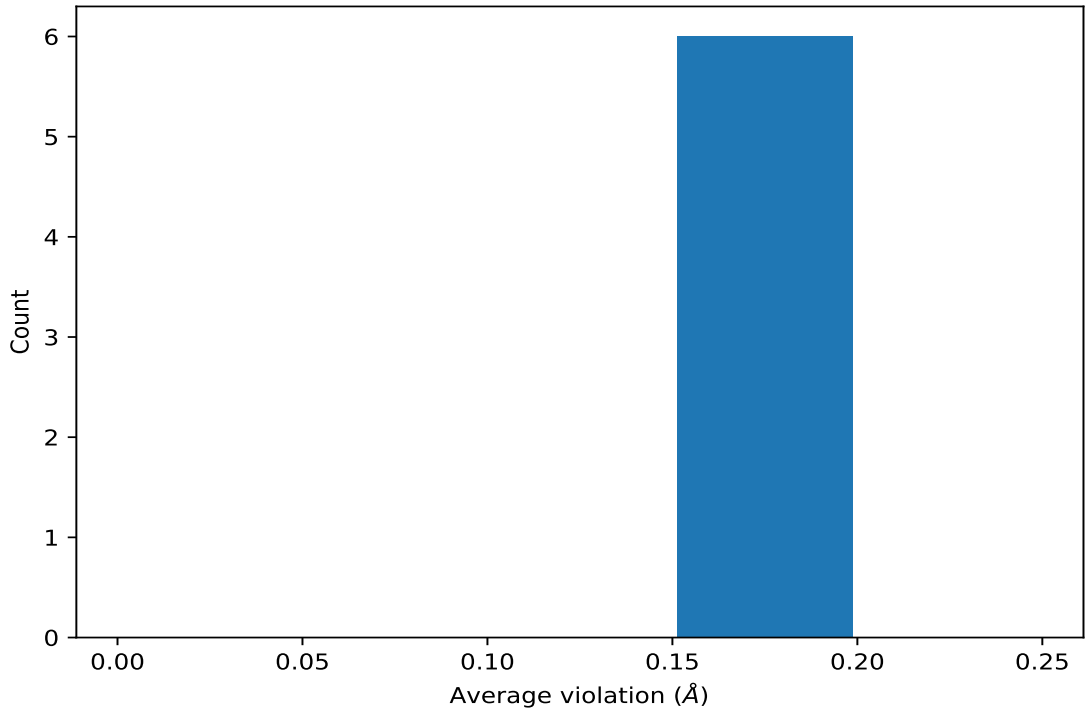
9.3.1 Bar graph : Distance violation statistics for the ensemble



9.4 Most violated distance restraints in the ensemble

9.4.1 Histogram : Distribution of mean distance violations

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

The following table provides the mean and the standard deviation of the violation for each restraint sorted by number of violated models and the mean 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.

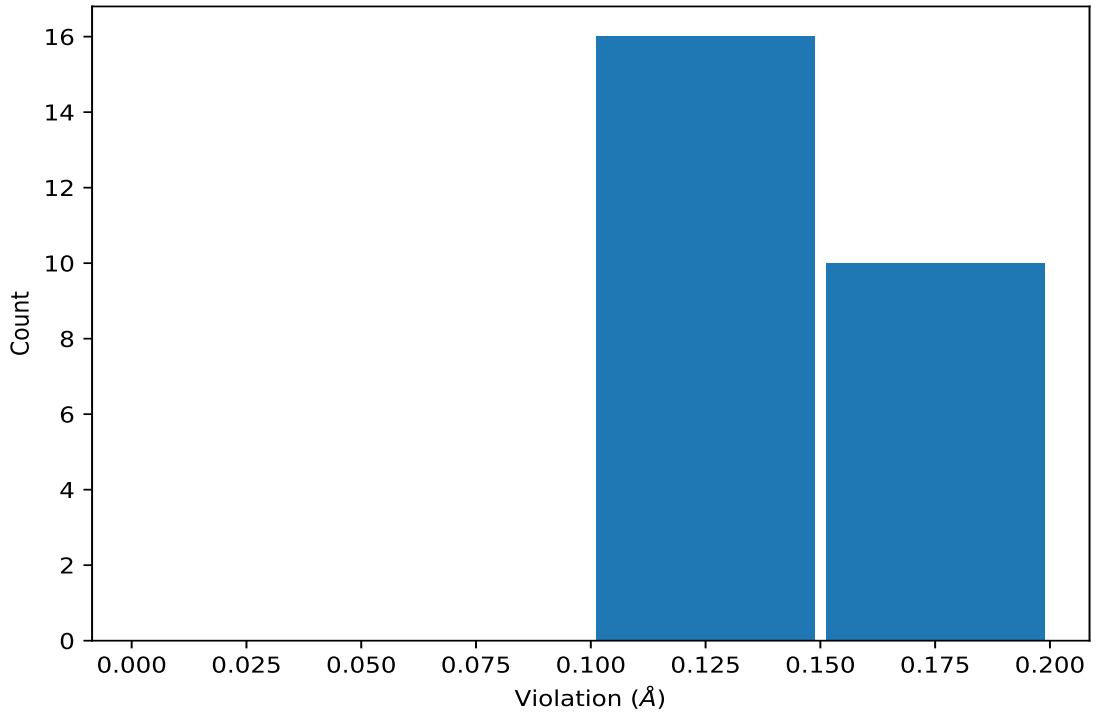
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,341)	1:A:686:LEU:HD11	1:A:686:LEU:HA	2	0.16	0.01	0.16
(1,341)	1:A:686:LEU:HD12	1:A:686:LEU:HA	2	0.16	0.01	0.16
(1,341)	1:A:686:LEU:HD13	1:A:686:LEU:HA	2	0.16	0.01	0.16
(1,342)	1:A:686:LEU:HD21	1:A:686:LEU:HA	2	0.16	0.01	0.16
(1,342)	1:A:686:LEU:HD22	1:A:686:LEU:HA	2	0.16	0.01	0.16
(1,342)	1:A:686:LEU:HD23	1:A:686:LEU:HA	2	0.16	0.01	0.16

[1]Number of violated models, [2]Standard deviation

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)
(1,761)	1:A:717:VAL:HG11	1:A:717:VAL:HA	15	0.17
(1,761)	1:A:717:VAL:HG12	1:A:717:VAL:HA	15	0.17
(1,761)	1:A:717:VAL:HG13	1:A:717:VAL:HA	15	0.17
(1,342)	1:A:686:LEU:HD21	1:A:686:LEU:HA	5	0.16
(1,342)	1:A:686:LEU:HD22	1:A:686:LEU:HA	5	0.16
(1,342)	1:A:686:LEU:HD23	1:A:686:LEU:HA	5	0.16
(1,341)	1:A:686:LEU:HD11	1:A:686:LEU:HA	5	0.16
(1,341)	1:A:686:LEU:HD12	1:A:686:LEU:HA	5	0.16
(1,341)	1:A:686:LEU:HD13	1:A:686:LEU:HA	5	0.16
(1,21)	1:A:666:ILE:H	1:A:665:GLU:H	2	0.16
(1,342)	1:A:686:LEU:HD21	1:A:686:LEU:HA	8	0.15
(1,342)	1:A:686:LEU:HD22	1:A:686:LEU:HA	8	0.15
(1,342)	1:A:686:LEU:HD23	1:A:686:LEU:HA	8	0.15
(1,341)	1:A:686:LEU:HD11	1:A:686:LEU:HA	8	0.15
(1,341)	1:A:686:LEU:HD12	1:A:686:LEU:HA	8	0.15
(1,341)	1:A:686:LEU:HD13	1:A:686:LEU:HA	8	0.15

Continued on next page...

Continued from previous page...

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,32)	1:A:666:ILE:HG21	1:A:666:ILE:HA	12	0.14
(1,32)	1:A:666:ILE:HG22	1:A:666:ILE:HA	12	0.14
(1,32)	1:A:666:ILE:HG23	1:A:666:ILE:HA	12	0.14
(1,28)	1:A:666:ILE:H	1:A:666:ILE:HG12	12	0.12
(1,28)	1:A:666:ILE:H	1:A:666:ILE:HG13	12	0.12
(1,2)	1:A:660:ARG:H	1:A:659:SER:HB2	8	0.12
(1,2)	1:A:660:ARG:H	1:A:659:SER:HB3	8	0.12
(1,385)	1:A:689:SER:H	1:A:688:LYS:HG2	2	0.11
(1,385)	1:A:689:SER:H	1:A:688:LYS:HG3	2	0.11
(1,1577)	1:A:769:SER:H	1:A:770:ARG:H	19	0.11

10 Dihedral-angle violation analysis

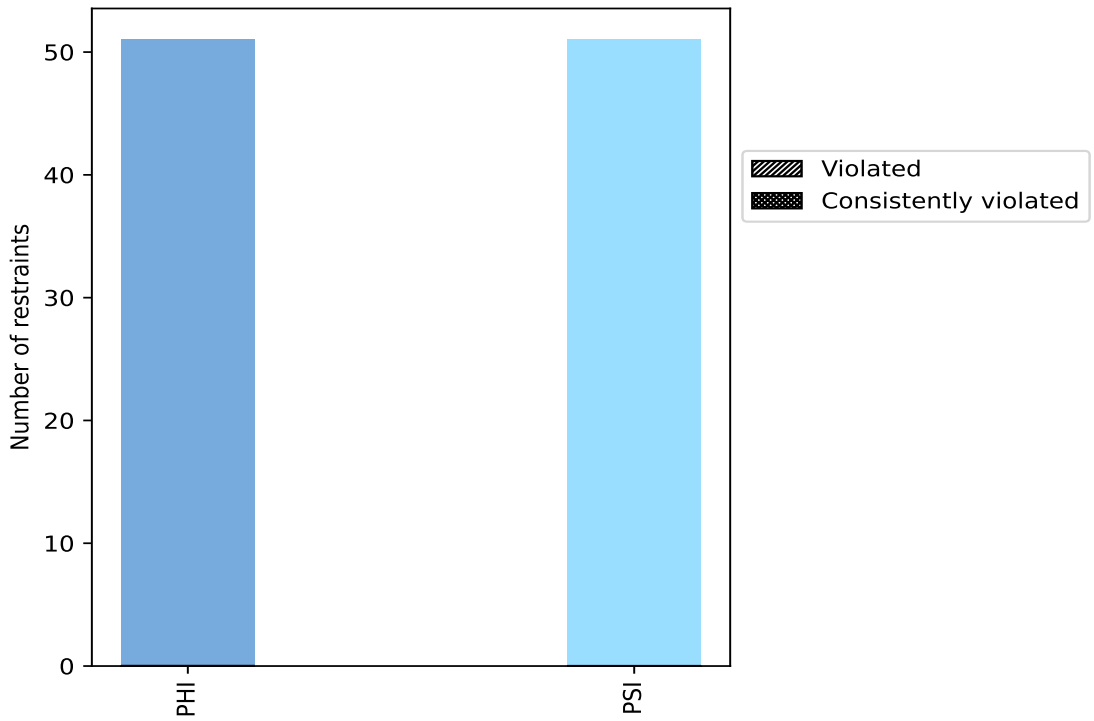
10.1 Summary of dihedral-angle violations

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	51	50.0	0	0.0	0.0	0	0.0	0.0
PSI	51	50.0	0	0.0	0.0	0	0.0	0.0
Total	102	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 for each model (i)

No violations found

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

No violations found

10.4 Most violated dihedral-angle restraints in the ensemble (i)

No violations found

10.5 All violated dihedral-angle restraints (i)

No violations found