wwPDB NMR Structure Validation Summary Report (i)

Jun 3, 2023 – 05:07 AM EDT

PDB ID	:	2NPB
BMRB ID	:	7324
Title	:	NMR solution structure of mouse SelW
Authors	:	Aachmann, F.L.; Fomenko, D.E.; Soragni, A.; Gladyshev, V.N.; Dikiy, A.
Deposited on	:	2006-10-27

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

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

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

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

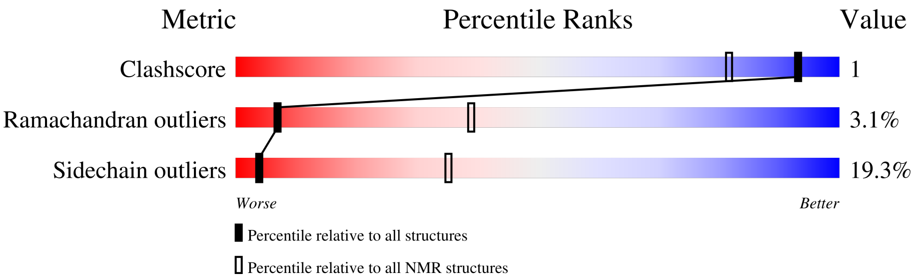
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal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

1 Overall quality at a glance (i)

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

The overall completeness of chemical shifts assignment is 90%.

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	96	

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 15 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *fewest violations*.

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

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:2-A:39, A:47-A:88 (80)	0.66	15

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

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

The models can be grouped into 3 clusters and 1 single-model cluster was found.

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

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1369 atoms, of which 695 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Selenoprotein W.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	88	1369	432	695	115	122	5	0

There are 10 discrepancies between the modelled and reference sequences:

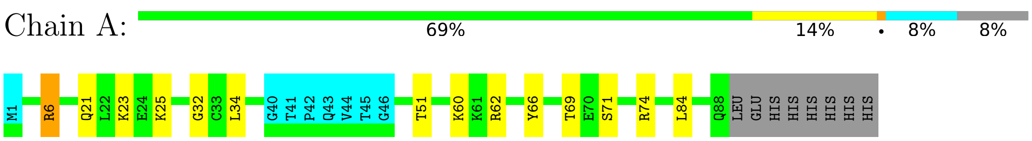
Chain	Residue	Modelled	Actual	Comment	Reference
A	10	SER	CYS	engineered mutation	UNP P63300
A	13	CYS	SEC	engineered mutation	UNP P63300
A	89	LEU	-	expression tag	UNP P63300
A	90	GLU	-	expression tag	UNP P63300
A	91	HIS	-	expression tag	UNP P63300
A	92	HIS	-	expression tag	UNP P63300
A	93	HIS	-	expression tag	UNP P63300
A	94	HIS	-	expression tag	UNP P63300
A	95	HIS	-	expression tag	UNP P63300
A	96	HIS	-	expression tag	UNP P63300

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.

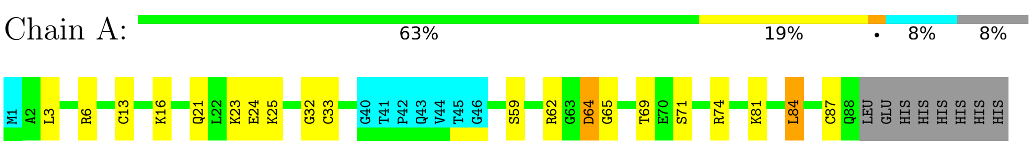
- Molecule 1: Selenoprotein W



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

The representative model is number 15. Colouring as in section 4.1 above.

- Molecule 1: Selenoprotein W



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *distance geometry, simulated annealing, energy minimisation, torsion angle dynamics*.

Of the 96 calculated structures, 20 were deposited, based on the following criterion: *target function*.

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

Software name	Classification	Version
CYANA	structure solution	2.1
Amber	refinement	9

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	1094
Number of shifts mapped to atoms	1020
Number of unparsed shifts	0
Number of shifts with mapping errors	74
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	90%

6 Model quality (i)

6.1 Standard geometry (i)

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	0.71±0.01	0±0/632 (0.0± 0.0%)	1.17±0.03	2±1/847 (0.3± 0.1%)
All	All	0.71	0/12640 (0.0%)	1.17	46/16940 (0.3%)

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	1.8±1.3
All	All	0	36

There are no bond-length outliers.

5 of 9 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	6	ARG	NE-CZ-NH1	11.27	125.93	120.30	19	13
1	A	74	ARG	NE-CZ-NH1	8.46	124.53	120.30	5	10
1	A	68	ASP	CB-CG-OD1	7.41	124.97	118.30	2	2
1	A	6	ARG	NE-CZ-NH2	6.98	123.79	120.30	6	6
1	A	62	ARG	NE-CZ-NH1	6.30	123.45	120.30	10	5

There are no chirality outliers.

5 of 11 unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	6	ARG	Sidechain	9
1	A	66	TYR	Peptide	7
1	A	62	ARG	Sidechain	5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	34	LEU	Peptide	4
1	A	58	HIS	Sidechain	4

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	621	640	640	2±1
All	All	12420	12800	12800	32

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.

5 of 14 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:33:CYS:CB	1:A:84:LEU:HD21	0.53	2.33	14	8
1:A:33:CYS:HB3	1:A:84:LEU:HD21	0.53	1.81	14	8
1:A:60:LYS:HE2	1:A:66:TYR:N	0.47	2.25	12	1
1:A:60:LYS:HE3	1:A:65:GLY:H	0.43	1.73	18	1
1:A:6:ARG:HH11	1:A:37:CYS:HB2	0.43	1.74	18	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	79/96 (82%)	69±3 (87±3%)	8±2 (10±3%)	2±1 (3±1%)	7	39
All	All	1580/1920 (82%)	1377 (87%)	154 (10%)	49 (3%)	7	39

5 of 12 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	32	GLY	13
1	A	66	TYR	9
1	A	14	GLY	8
1	A	13	CYS	5
1	A	67	VAL	3

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	65/79 (82%)	52±2 (81±3%)	13±2 (19±3%)	4	35
All	All	1300/1580 (82%)	1049 (81%)	251 (19%)	4	35

5 of 46 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	25	LYS	20
1	A	34	LEU	14
1	A	69	THR	14
1	A	62	ARG	13
1	A	71	SER	13

6.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

6.5 Carbohydrates (i)

There are no monosaccharides in this entry.

6.6 Ligand geometry (i)

There are no ligands in this entry.

6.7 Other polymers (i)

There are no such molecules in this entry.

6.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

7 Chemical shift validation (i)

The completeness of assignment taking into account all chemical shift lists is 90% for the well-defined parts and 86% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

7.1.1 Bookkeeping (i)

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

Total number of shifts	1094
Number of shifts mapped to atoms	1020
Number of unparsed shifts	0
Number of shifts with mapping errors	74
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	6

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

- No matching atom found in the structure. First 5 (of 74) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	89	LEU	H	7.849	0.02	1
1	A	89	LEU	HA	4.159	0.02	1
1	A	89	LEU	HB2	1.674	0.02	1
1	A	89	LEU	HB3	1.552	0.02	1
1	A	89	LEU	HG	1.617	0.02	1
1	A	89	LEU	HD11	0.855	0.02	1
1	A	89	LEU	HD12	0.855	0.02	1
1	A	89	LEU	HD13	0.855	0.02	1
1	A	89	LEU	HD21	0.818	0.02	1
1	A	89	LEU	HD22	0.818	0.02	1
1	A	89	LEU	HD23	0.818	0.02	1
1	A	89	LEU	C	178.074	0.2	1
1	A	89	LEU	CA	56.07	0.2	1
1	A	89	LEU	CB	42.133	0.2	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	89	LEU	CG	26.706	0.2	1
1	A	89	LEU	CD1	23.975	0.2	1
1	A	89	LEU	CD2	23.461	0.2	1
1	A	89	LEU	N	121.915	0.2	1
1	A	90	GLU	H	8.023	0.02	1
1	A	90	GLU	HA	4.084	0.02	1
1	A	90	GLU	HB2	1.905	0.02	1
1	A	90	GLU	HB3	1.89	0.02	1
1	A	90	GLU	HG2	2.136	0.02	1
1	A	90	GLU	HG3	2.274	0.02	1
1	A	90	GLU	C	176.698	0.2	1
1	A	90	GLU	CA	56.91	0.2	1
1	A	90	GLU	CB	29.766	0.2	1
1	A	90	GLU	CG	35.999	0.2	1
1	A	90	GLU	N	120.41	0.2	1
1	A	91	HIS	H	8.205	0.02	1
1	A	91	HIS	HA	4.578	0.02	1
1	A	91	HIS	HB2	3.163	0.02	1
1	A	91	HIS	HB3	3.068	0.02	1
1	A	91	HIS	HD1	7.19	0.02	1
1	A	91	HIS	C	174.427	0.2	1
1	A	91	HIS	CA	55.582	0.2	1
1	A	91	HIS	CB	29.128	0.2	1
1	A	91	HIS	N	119.116	0.2	1
1	A	92	HIS	H	8.368	0.02	1
1	A	92	HIS	HA	4.557	0.02	1
1	A	92	HIS	C	174.324	0.2	1
1	A	92	HIS	CA	55.545	0.2	1
1	A	92	HIS	CB	29.423	0.2	1
1	A	92	HIS	N	119.683	0.2	1
1	A	93	HIS	H	8.678	0.02	1
1	A	93	HIS	C	174.186	0.2	1
1	A	93	HIS	CA	55.555	0.2	1
1	A	93	HIS	CB	29.224	0.2	1
1	A	93	HIS	N	120.971	0.2	1
1	A	94	HIS	H	8.54	0.02	1
1	A	94	HIS	HA	4.602	0.02	1
1	A	94	HIS	HB2	3.162	0.02	1
1	A	94	HIS	HB3	3.067	0.02	1
1	A	94	HIS	C	174.307	0.2	1
1	A	94	HIS	CA	55.441	0.2	1

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	94	HIS	CB	29.322	0.2	1
1	A	94	HIS	N	120.461	0.2	1
1	A	95	HIS	H	8.668	0.02	1
1	A	95	HIS	HA	4.581	0.02	1
1	A	95	HIS	HB2	3.173	0.02	1
1	A	95	HIS	HB3	3.065	0.02	1
1	A	95	HIS	HD2	7.19	0.02	1
1	A	95	HIS	C	173.567	0.2	1
1	A	95	HIS	CA	55.569	0.2	1
1	A	95	HIS	CB	29.368	0.2	1
1	A	95	HIS	N	120.968	0.2	1
1	A	96	HIS	H	8.291	0.02	1
1	A	96	HIS	HA	4.403	0.02	1
1	A	96	HIS	HB2	3.181	0.02	1
1	A	96	HIS	HB3	3.073	0.02	1
1	A	96	HIS	C	178.951	0.2	1
1	A	96	HIS	CA	56.958	0.2	1
1	A	96	HIS	CB	29.506	0.2	1
1	A	96	HIS	N	126.046	0.2	1

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	93	-0.36 ± 0.34	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	84	0.49 ± 0.14	None needed (< 0.5 ppm)
$^{13}C'$	86	-0.38 ± 0.17	None needed (< 0.5 ppm)
^{15}N	86	-1.09 ± 0.45	Should be applied

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 90%, i.e. 978 atoms were assigned a chemical shift out of a possible 1090. 0 out of 16 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	394/403 (98%)	163/165 (99%)	155/160 (97%)	76/78 (97%)
Sidechain	545/597 (91%)	375/389 (96%)	162/186 (87%)	8/22 (36%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	39/90 (43%)	39/44 (89%)	0/44 (0%)	0/2 (0%)
Overall	978/1090 (90%)	577/598 (96%)	317/390 (81%)	84/102 (82%)

7.1.4 Statistically unusual chemical shifts (i)

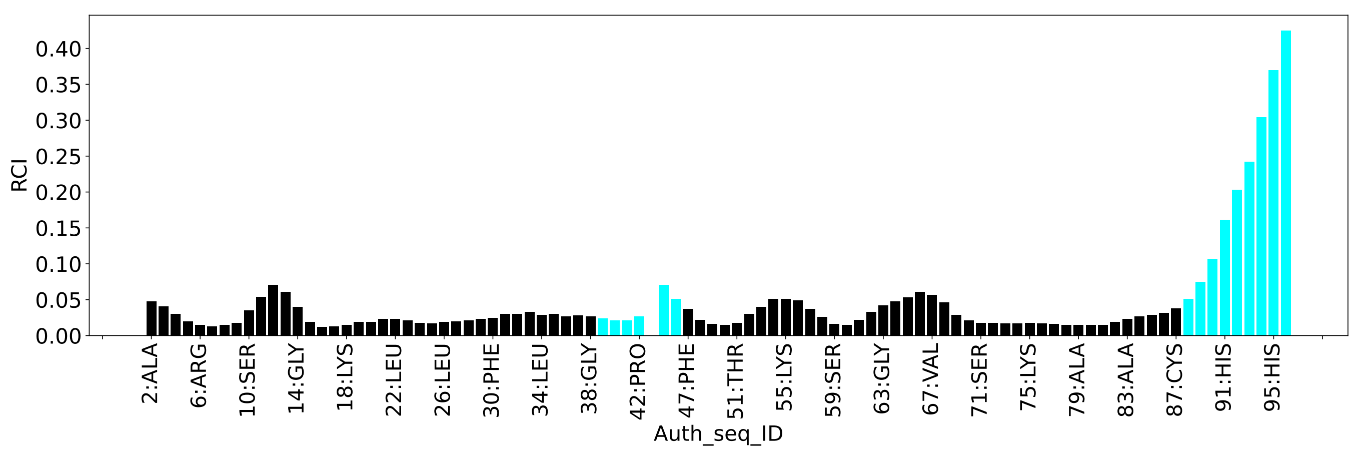
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	62	ARG	NE	122.43	76.53 – 92.65	23.5
1	A	74	ARG	NE	121.92	76.53 – 92.65	23.2
1	A	6	ARG	NE	121.72	76.53 – 92.65	23.0
1	A	6	ARG	NH2	109.27	57.68 – 87.89	12.1
1	A	6	ARG	NH1	108.69	49.05 – 99.42	6.8
1	A	64	ASP	HB3	1.29	1.32 – 4.00	-5.1

7.1.5 Random Coil Index (RCI) plots (i)

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

Random coil index (RCI) for chain A:



8 NMR restraints analysis (i)

8.1 Conformationally restricting restraints (i)

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

Description	Value
Total distance restraints	878
Intra-residue (|i-j|=0)	289
Sequential (|i-j|=1)	341
Medium range (|i-j|>1 and |i-j|<5)	152
Long range (|i-j|≥5)	96
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	11
Number of restraints per residue	9.1
Number of long range restraints per residue[1]	1.0

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

8.2 Residual restraint violations (i)

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

8.2.1 Average number of distance violations per model (i)

Distance violations less than 0.1 Å are not included in the calculation.

Bins (Å)	Average number of violations per model	Max (Å)
0.1-0.2 (Small)	36.8	0.2
0.2-0.5 (Medium)	13.6	0.49
>0.5 (Large)	3.5	1.24

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)

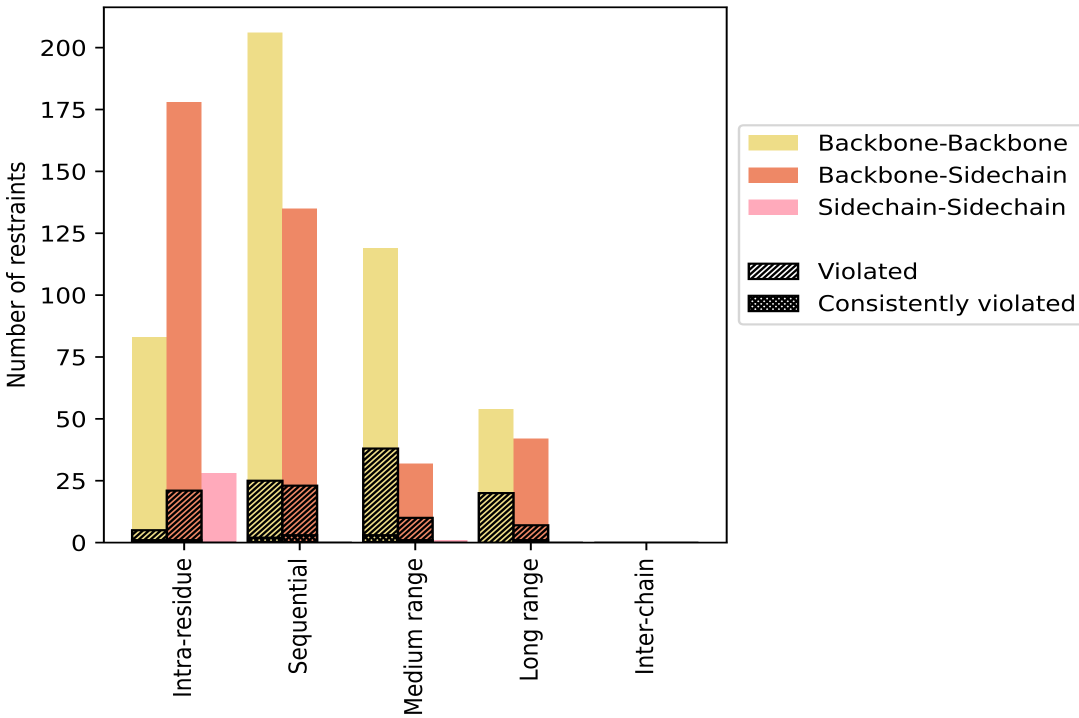
9.1 Summary of distance violations (i)

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)	289	32.9	26	9.0	3.0	2	0.7	0.2
Backbone-Backbone	83	9.5	5	6.0	0.6	1	1.2	0.1
Backbone-Sidechain	178	20.3	21	11.8	2.4	1	0.6	0.1
Sidechain-Sidechain	28	3.2	0	0.0	0.0	0	0.0	0.0
Sequential (|i-j|=1)	341	38.8	48	14.1	5.5	5	1.5	0.6
Backbone-Backbone	206	23.5	25	12.1	2.8	2	1.0	0.2
Backbone-Sidechain	135	15.4	23	17.0	2.6	3	2.2	0.3
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	152	17.3	48	31.6	5.5	4	2.6	0.5
Backbone-Backbone	119	13.6	38	31.9	4.3	3	2.5	0.3
Backbone-Sidechain	32	3.6	10	31.2	1.1	1	3.1	0.1
Sidechain-Sidechain	1	0.1	0	0.0	0.0	0	0.0	0.0
Long range (|i-j|≥5)	96	10.9	27	28.1	3.1	1	1.0	0.1
Backbone-Backbone	54	6.2	20	37.0	2.3	0	0.0	0.0
Backbone-Sidechain	42	4.8	7	16.7	0.8	1	2.4	0.1
Sidechain-Sidechain	0	0.0	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	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	878	100.0	149	17.0	17.0	12	1.4	1.4
Backbone-Backbone	462	52.6	88	19.0	10.0	6	1.3	0.7
Backbone-Sidechain	387	44.1	61	15.8	6.9	6	1.6	0.7
Sidechain-Sidechain	29	3.3	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 (i)



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 (i)

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	8	22	15	9	0	54	0.2	0.95	0.16	0.15
2	8	22	19	13	0	62	0.21	1.1	0.16	0.15
3	10	21	15	9	0	55	0.23	1.09	0.2	0.17
4	8	22	16	10	0	56	0.21	1.24	0.18	0.14
5	7	21	12	10	0	50	0.21	1.16	0.18	0.16
6	5	22	18	5	0	50	0.22	1.03	0.17	0.15
7	8	20	18	8	0	54	0.23	1.24	0.19	0.17
8	10	21	19	8	0	58	0.22	1.05	0.17	0.16
9	8	20	19	9	0	56	0.23	1.06	0.18	0.16
10	6	19	15	9	0	49	0.23	1.14	0.19	0.17
11	7	19	15	7	0	48	0.23	1.07	0.2	0.16

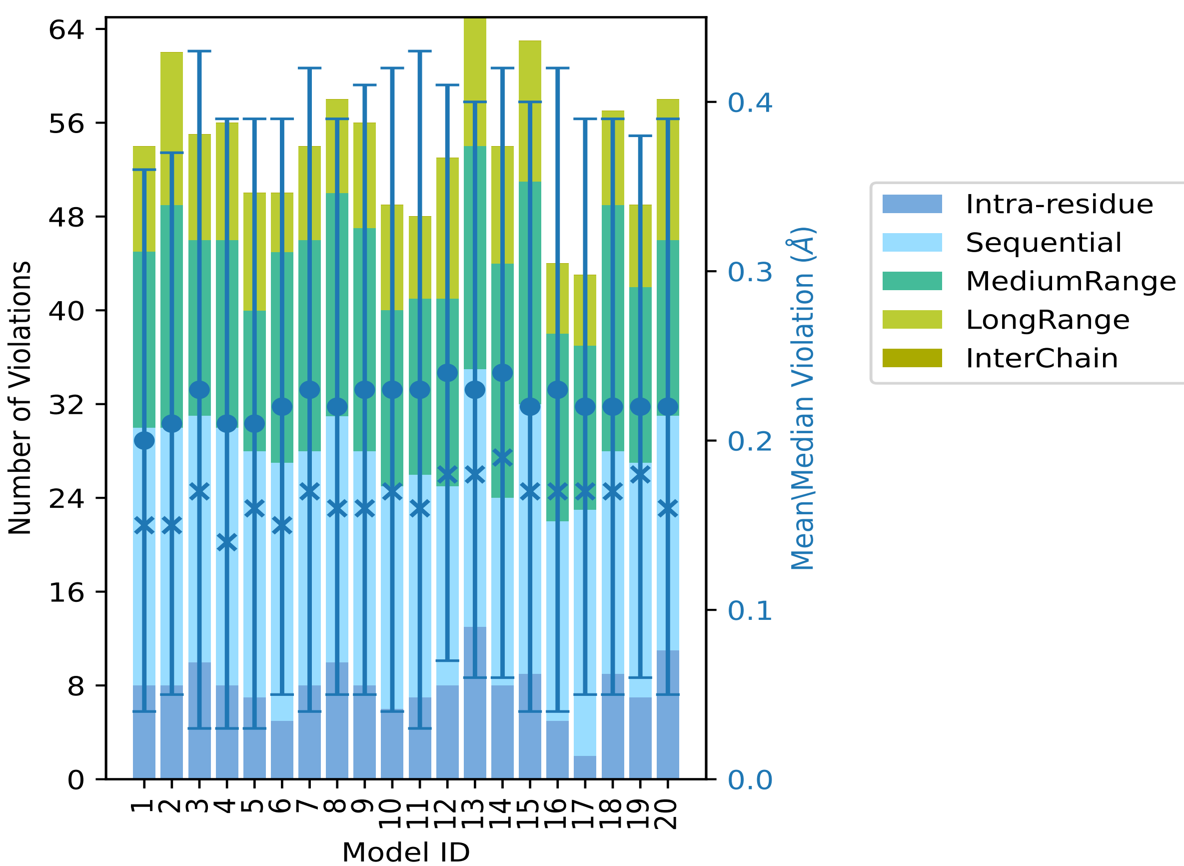
Continued on next page...

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	8	17	16	12	0	53	0.24	0.96	0.17	0.18
13	13	22	19	11	0	65	0.23	1.05	0.17	0.18
14	8	16	20	10	0	54	0.24	1.17	0.18	0.19
15	9	23	19	12	0	63	0.22	1.2	0.18	0.17
16	5	17	16	6	0	44	0.23	1.1	0.19	0.17
17	2	21	14	6	0	43	0.22	0.91	0.17	0.17
18	9	19	21	8	0	57	0.22	1.09	0.17	0.17
19	7	20	15	7	0	49	0.22	0.99	0.16	0.18
20	11	20	15	12	0	58	0.22	1.03	0.17	0.16

[1]Intra-residue restraints, [2]Sequential restraints, [3]Medium range restraints, [4]Long range restraints, [5]Inter-chain restraints, [6]Standard deviation

9.2.1 Bar graph : Distance Violation statistics for each model (i)



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

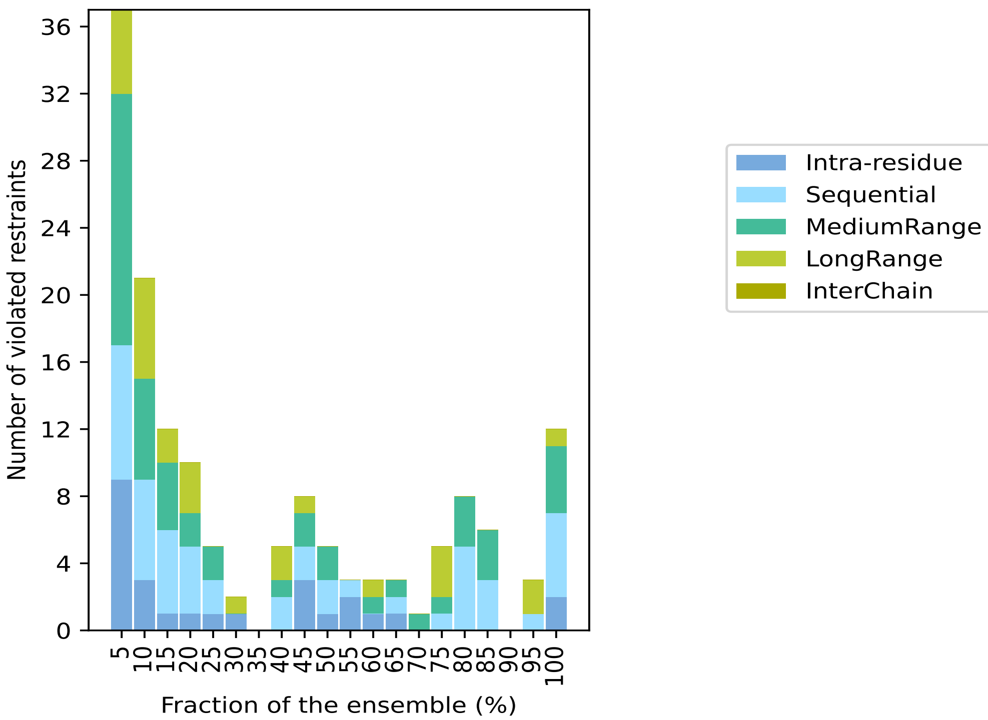
9.3 Distance violation statistics for the ensemble (i)

Violation analysis may find that some restraints are violated in few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of the ensemble. In total, 729(IR:263, SQ:293, MR:104, LR:69, 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]	%
9	8	15	5	0	37	1	5.0
3	6	6	6	0	21	2	10.0
1	5	4	2	0	12	3	15.0
1	4	2	3	0	10	4	20.0
1	2	2	0	0	5	5	25.0
1	0	0	1	0	2	6	30.0
0	0	0	0	0	0	7	35.0
0	2	1	2	0	5	8	40.0
3	2	2	1	0	8	9	45.0
1	2	2	0	0	5	10	50.0
2	1	0	0	0	3	11	55.0
1	0	1	1	0	3	12	60.0
1	1	1	0	0	3	13	65.0
0	0	1	0	0	1	14	70.0
0	1	1	3	0	5	15	75.0
0	5	3	0	0	8	16	80.0
0	3	3	0	0	6	17	85.0
0	0	0	0	0	0	18	90.0
0	1	0	2	0	3	19	95.0
2	5	4	1	0	12	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

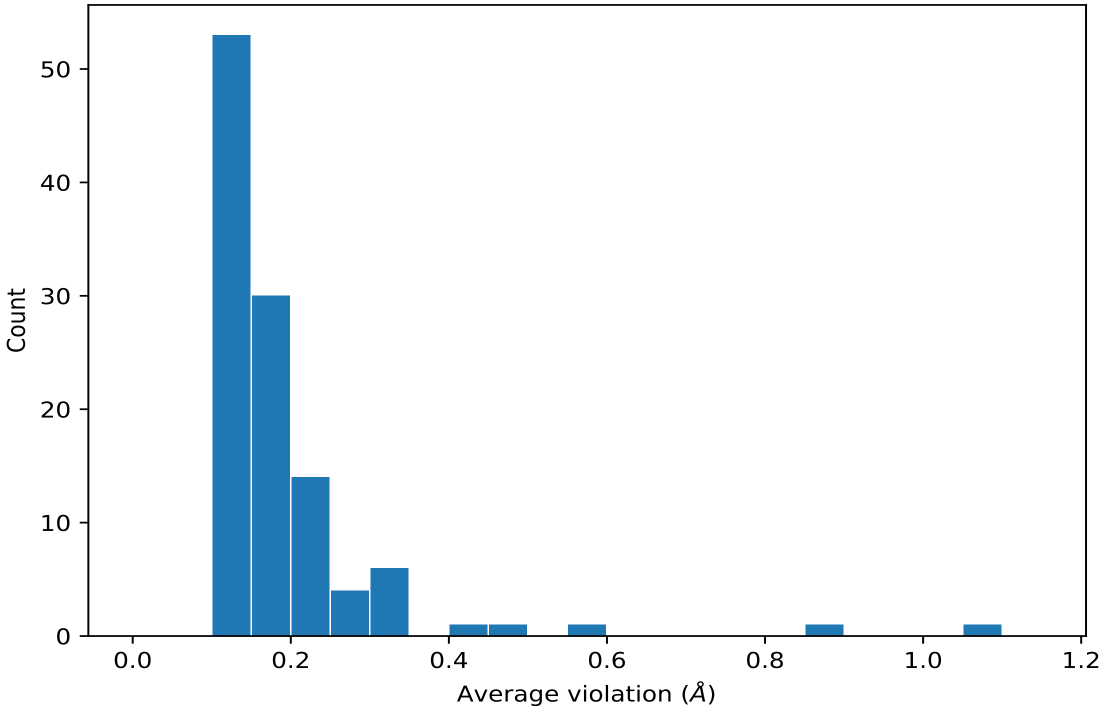
9.3.1 Bar graph : Distance violation statistics for the ensemble (i)



9.4 Most violated distance restraints in the ensemble (i)

9.4.1 Histogram : Distribution of mean distance violations (i)

The following histogram shows the distribution of the average value of the violation. The average is calculated for each restraint that is violated in more than one model over all the violated models in the ensemble



9.4.2 Table: Most violated distance restraints (i)

The following table provides the mean and the standard deviation of the violations for the 10 worst performing restraints, sorted by number of violated models and the mean violation value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint. Rows with same key represent combinatorial or ambiguous restraints and are counted as a single restraint.

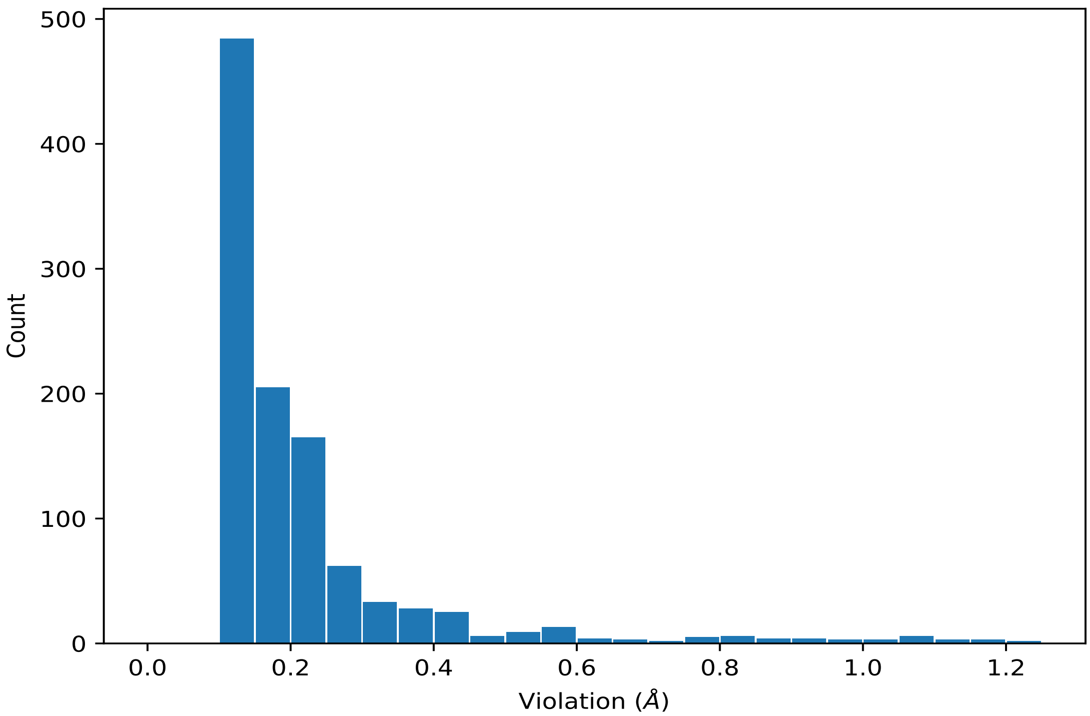
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,745)	1:A:3:LEU:HB2	1:A:53:ALA:H	20	1.08	0.1	1.08
(1,108)	1:A:22:LEU:HB2	1:A:23:LYS:H	20	0.85	0.08	0.85
(1,258)	1:A:54:GLY:H	1:A:54:GLY:HA3	20	0.4	0.03	0.4
(1,118)	1:A:20:LEU:HB2	1:A:24:GLU:H	20	0.33	0.1	0.34
(1,456)	1:A:79:ALA:HA	1:A:81:LYS:H	20	0.24	0.03	0.24
(1,187)	1:A:36:ILE:H	1:A:36:ILE:HB	20	0.22	0.03	0.23
(1,14)	1:A:5:VAL:HB	1:A:6:ARG:H	20	0.2	0.02	0.2
(1,474)	1:A:83:ALA:HA	1:A:84:LEU:H	20	0.2	0.02	0.2
(1,487)	1:A:86:GLN:HA	1:A:87:CYS:H	20	0.2	0.02	0.2
(1,481)	1:A:84:LEU:HB2	1:A:85:ALA:H	20	0.2	0.04	0.19

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

9.5 All violated distance restraints (i)

9.5.1 Histogram : Distribution of distance violations (i)

The following histogram shows the distribution of the absolute value of the violation for all violated restraints in the ensemble.



9.5.2 Table : All distance violations (i)

The following table provides the 10 worst performing restraints, sorted by the violation value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint. Rows with same key represent combinatorial or ambiguous restraints and are counted as a single restraint.

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,745)	1:A:3:LEU:HB2	1:A:53:ALA:H	4	1.24
(1,745)	1:A:3:LEU:HB2	1:A:53:ALA:H	7	1.24
(1,745)	1:A:3:LEU:HB2	1:A:53:ALA:H	15	1.2
(1,745)	1:A:3:LEU:HB2	1:A:53:ALA:H	14	1.17
(1,745)	1:A:3:LEU:HB2	1:A:53:ALA:H	5	1.16
(1,745)	1:A:3:LEU:HB2	1:A:53:ALA:H	10	1.14
(1,745)	1:A:3:LEU:HB2	1:A:53:ALA:H	2	1.1
(1,745)	1:A:3:LEU:HB2	1:A:53:ALA:H	16	1.1
(1,745)	1:A:3:LEU:HB2	1:A:53:ALA:H	3	1.09
(1,745)	1:A:3:LEU:HB2	1:A:53:ALA:H	18	1.09

10 Dihedral-angle violation analysis (i)

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