wwPDB NMR Structure Validation Summary Report

Aug 25, 2026 – 07:50 am BST

PDB ID	:	9GSA / pdb_00009gsa
BMRB ID	:	34959
Title	:	Lys9DabMC6*a 1-Delta
Authors	:	Maglio, O.; Lombardi, A.; Chino, M.; Pirro, F.
Deposited on	:	2024-09-13

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:

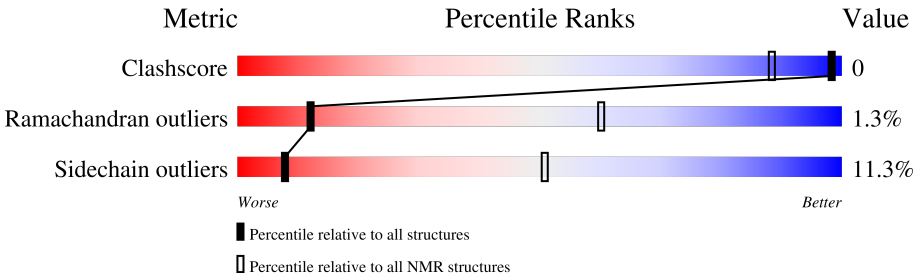
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	1.8.4, CSD as541be (2020)
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2.1
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.50

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 19%.

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	229148	14424
Ramachandran outliers	224038	12848
Sidechain outliers	223484	12823

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	12	42% 58%
2	B	16	88% 12%

2 Ensemble composition and analysis

This entry contains 20 models. Model 11 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:4-A:6, A:8-A:8, A:10-A:10, B:1-B:14 (19)	0.65	11

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 4 clusters. No single-model clusters were found.

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

3 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 485 atoms, of which 241 are hydrogens and 0 are deuteriums.

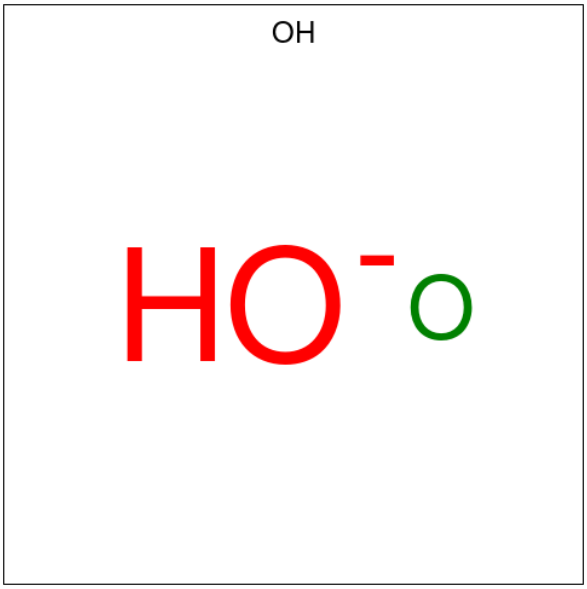
- Molecule 1 is a protein called ACE-ASP-GLU-AIB-GLN-LEU-SER-AIB-GLN-DAB-ARG-NH2.

Mol	Chain	Residues	Atoms					Trace
1	A	12	Total	C	H	N	O	1
			164	48	81	17	18	

- Molecule 2 is a protein called ACE-ASP-LEU-GLN-GLN-LEU-HIS-SER-GLN-LYS-ARG-LYS-ILE-THR-LEU-NH2.

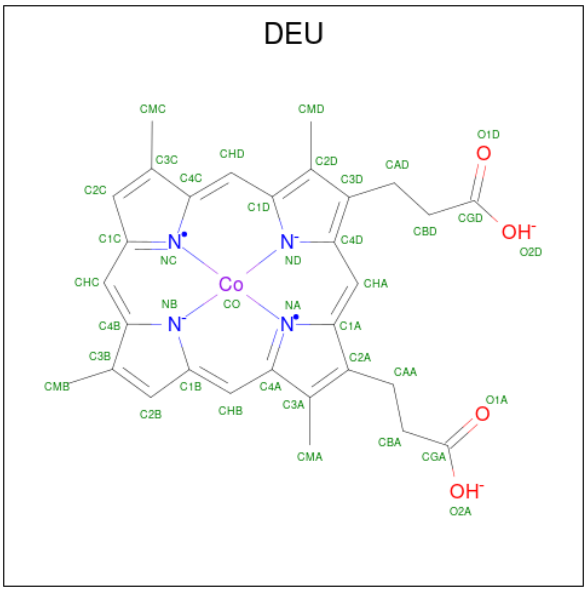
Mol	Chain	Residues	Atoms					Trace
2	B	16	Total	C	H	N	O	1
			256	76	133	25	22	

- Molecule 3 is HYDROXIDE ION (CCD ID: OH) (formula: HO).



Mol	Chain	Residues	Atoms			
3	A	1	Total	H	O	
			2	1	1	

- Molecule 4 is CO(III)-(DEUTEROPORPHYRIN IX) (CCD ID: DEU) (formula: $C_{30}H_{28}CoN_4O_4$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					
			Total	C	Co	H	N	O
4	A	1	63	30	1	26	4	2

4 Residue-property plots

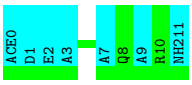
4.1 Average score per residue in the NMR ensemble

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: ACE-ASP-GLU-AIB-GLN-LEU-SER-AIB-GLN-DAB-ARG-NH2

Chain A:





• Molecule 2: ACE-ASP-LEU-GLN-GLN-LEU-HIS-SER-GLN-LYS-ARG-LYS-ILE-THR-LEU-NH2

Chain B:





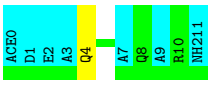
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 11. Colouring as in section 4.1 above.

• Molecule 1: ACE-ASP-GLU-AIB-GLN-LEU-SER-AIB-GLN-DAB-ARG-NH2

Chain A:

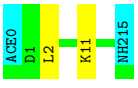




• Molecule 2: ACE-ASP-LEU-GLN-GLN-LEU-HIS-SER-GLN-LYS-ARG-LYS-ILE-THR-LEU-NH2

Chain B:





5 Refinement protocol and experimental data overview

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

Of the 400 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 calculation	2.1
Amber	refinement	10

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	167
Number of shifts mapped to atoms	165
Number of unparsed shifts	0
Number of shifts with mapping errors	2
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	19%

6 Model quality

6.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NH2, AIB, DEU, DAB, ACE, OH

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	0.72±0.03	0±0/40 (0.0± 0.0%)	1.74±0.12	1±1/48 (1.9± 2.1%)
2	B	0.81±0.02	0±0/119 (0.0± 0.0%)	1.63±0.06	0±1/157 (0.3± 0.6%)
All	All	0.79	0/3180 (0.0%)	1.66	28/4100 (0.7%)

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	4	GLN	CA-C-N	6.58	129.39	120.38	1	9
1	A	4	GLN	C-N-CA	6.58	129.39	120.38	1	9
2	B	1	ASP	CA-C-N	5.66	127.87	120.28	20	5
2	B	1	ASP	C-N-CA	5.66	127.87	120.28	20	5

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	43	45	45	0±0
3	A	1	1	0	0±0
4	A	37	26	26	0±0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes
All	All	4000	4000	3972	2

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 0.

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

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:10:ARG:HD3	4:A:102:DEU:HMD2	0.52	1.79	19	1
1:A:6:SER:OG	3:A:101:OH:O	0.42	2.37	13	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	5/12 (42%)	4±1 (85±19%)	1±1 (14±18%)	0±0 (1±4%)	15	65
2	B	14/16 (88%)	13±0 (95±3%)	0±0 (4±4%)	0±0 (1±3%)	11	58
All	All	380/560 (68%)	351 (92%)	24 (6%)	5 (1%)	12	60

All 2 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
2	B	1	ASP	4
1	A	10	ARG	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	5/7 (71%)	5±1 (94±11%)	0±1 (6±11%)	19 68
2	B	14/14 (100%)	12±1 (87±5%)	2±1 (13±5%)	6 46
All	All	380/420 (90%)	337 (89%)	43 (11%)	8 51

5 of 13 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)
2	B	11	LYS	10
2	B	8	GLN	9
2	B	2	LEU	8
2	B	9	LYS	3
2	B	4	GLN	2

6.3.3 RNA (i)

There are no RNA molecules in this entry.

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

3 non-standard protein/DNA/RNA residues are 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
1	DAB	A	9	1,4	5,6,7	0.69±0.10	0±0 (0±0%)
1	AIB	A	7	1	1,5,6	1.25±0.04	0±0 (0±0%)
1	AIB	A	3	1	1,5,6	1.15±0.05	0±0 (0±0%)

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
1	DAB	A	9	1,4	1,6,8	0.33±0.14	0±0 (0±0%)
1	AIB	A	7	1	2,7,9	0.80±0.17	0±0 (0±0%)
1	AIB	A	3	1	2,7,9	0.42±0.14	0±0 (0±0%)

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
1	AIB	A	7	1	-	0±0,2,3,6	-
1	DAB	A	9	1,4	-	0±0,4,5,7	-
1	AIB	A	3	1	-	0±0,2,3,6	-

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

6.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 1 is modelled with single atom - leaving 1 for Mogul analysis.

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
4	DEU	A	102	3,1,2	44,44,46	1.08±0.03	3±0 (6±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
4	DEU	A	102	3,1,2	55,72,76	0.95±0.03	2±1 (4±1%)

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
4	DEU	A	102	3,1,2	-	0±0,7,48,50	-

All unique bond 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(Å)	Ideal(Å)	Models	
								Worst	Total
4	A	102	DEU	CO-NA	4.33	2.09	1.96	14	20
4	A	102	DEU	O2D-CGD	3.59	1.23	1.42	17	20
4	A	102	DEU	CO-NB	2.33	2.04	1.96	3	16
4	A	102	DEU	CO-ND	2.10	2.03	1.96	4	1

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
4	A	102	DEU	C3D-C4D-ND	2.55	108.26	110.68	20	19
4	A	102	DEU	CHD-C4C-C3C	2.46	123.91	129.77	14	18
4	A	102	DEU	C1C-C2C-C3C	2.28	106.37	108.89	13	9
4	A	102	DEU	CMD-C2D-C3D	2.04	130.01	125.61	7	2

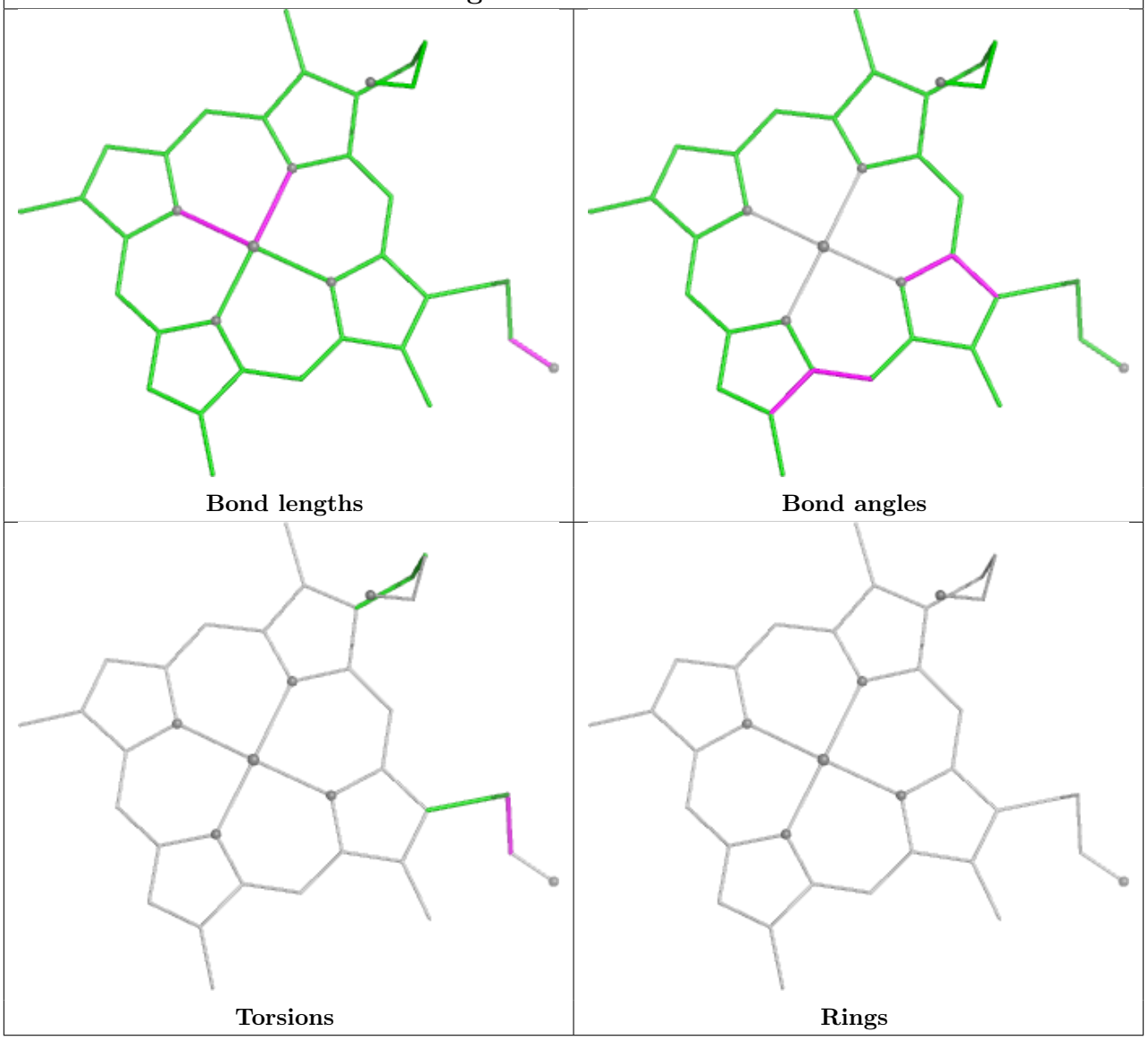
There are no chirality outliers.

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 DEU A 102



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 19% for the well-defined parts and 19% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *starch_ output*

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	167
Number of shifts mapped to atoms	165
Number of unparsed shifts	0
Number of shifts with mapping errors	2
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	3

- Can not parse the chemical shift value. First 5 (of 73) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	1	ASP	HB2	2.794	0.010	.
1	A	1	ASP	HB3	2.794	0.010	.
1	A	2	GLU	HB2	2.054	0.005	.
1	A	2	GLU	HB3	2.054	0.005	.
1	A	2	GLU	HG2	2.427	0.002	.
1	A	2	GLU	HG3	2.427	0.002	.
1	A	4	GLN	HB3	1.942	0.003	.
1	A	4	GLN	HE21	6.550	0.004	.
1	A	4	GLN	HE22	7.267	0.001	.
1	A	4	GLN	HG2	2.276	0.007	.
1	A	4	GLN	HG3	2.276	0.007	.
1	A	5	LEU	HB3	1.645	0.008	.
1	A	5	LEU	HD11	0.800	0.003	.
1	A	5	LEU	HD12	0.800	0.003	.
1	A	5	LEU	HD13	0.800	0.003	.
1	A	5	LEU	HD21	0.697	0.004	.
1	A	5	LEU	HD22	0.697	0.004	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	5	LEU	HD23	0.697	0.004	.
1	A	6	SER	HB3	3.577	0.002	.
1	A	8	GLN	HB3	1.575	0.006	.
1	A	8	GLN	HE21	6.055	0.003	.
1	A	8	GLN	HE22	6.976	0.002	.
1	A	8	GLN	HG3	2.032	0.005	.
1	A	10	ARG	HB3	1.828	0.006	.
1	A	10	ARG	HD3	3.042	0.006	.
1	A	10	ARG	HG3	1.448	0.011	.
1	B	1	ASP	HB2	2.715	0.002	.
1	B	1	ASP	HB3	2.715	0.002	.
1	B	2	LEU	HB3	1.423	0.005	.
1	B	2	LEU	HD11	0.976	0.004	.
1	B	2	LEU	HD12	0.976	0.004	.
1	B	2	LEU	HD13	0.976	0.004	.
1	B	2	LEU	HD21	0.682	0.003	.
1	B	2	LEU	HD22	0.682	0.003	.
1	B	2	LEU	HD23	0.682	0.003	.
1	B	3	GLN	HB3	1.902	0.008	.
1	B	3	GLN	HE21	6.494	0.003	.
1	B	3	GLN	HE22	7.180	0.003	.
1	B	3	GLN	HG3	2.045	0.004	.
1	B	4	GLN	HB3	1.903	0.005	.
1	B	4	GLN	HE22	7.267	0.001	.
1	B	4	GLN	HG3	2.279	0.009	.
1	B	5	LEU	HB3	1.056	0.009	.
1	B	5	LEU	HD11	0.586	0.006	.
1	B	5	LEU	HD12	0.586	0.006	.
1	B	5	LEU	HD13	0.586	0.006	.
1	B	5	LEU	HD21	0.008	0.004	.
1	B	5	LEU	HD22	0.008	0.004	.
1	B	5	LEU	HD23	0.008	0.004	.
1	B	6	HIS	HB3	1.481	0.007	.
1	B	7	SER	HB3	3.739	0.002	.
1	B	8	GLN	HB3	2.050	0.008	.
1	B	8	GLN	HE21	6.600	0.004	.
1	B	8	GLN	HE22	6.980	0.003	.
1	B	8	GLN	HG3	2.286	0.007	.
1	B	9	LYS	HB3	1.636	0.010	.
1	B	9	LYS	HD3	2.204	0.005	.
1	B	9	LYS	HE3	3.167	0.008	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	B	9	LYS	HG3	1.460	0.013	.
1	B	10	ARG	HB3	1.824	0.004	.
1	B	10	ARG	HD3	3.109	0.002	.
1	B	10	ARG	HG3	1.717	0.007	.
1	B	11	LYS	HB3	1.939	0.007	.
1	B	11	LYS	HD3	1.665	0.009	.
1	B	11	LYS	HG3	1.547	0.006	.
1	B	12	ILE	HG13	1.706	0.010	.
1	B	14	LEU	HB3	1.845	0.007	.
1	B	14	LEU	HD11	0.978	0.006	.
1	B	14	LEU	HD12	0.978	0.006	.
1	B	14	LEU	HD13	0.978	0.006	.
1	B	14	LEU	HD21	0.978	0.006	.
1	B	14	LEU	HD22	0.978	0.006	.
1	B	14	LEU	HD23	0.978	0.006	.

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

- No matching atom found in the structure. All 2 occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	B	9	LYS	HZ2	8.499	0.003	1
1	B	9	LYS	HZ3	8.499	0.003	1

7.1.2 Chemical shift referencing (i)

No chemical shift referencing corrections were calculated (not enough data).

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 19%, i.e. 58 atoms were assigned a chemical shift out of a possible 298. 0 out of 4 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	38/95 (40%)	38/38 (100%)	0/38 (0%)	0/19 (0%)
Sidechain	17/196 (9%)	17/125 (14%)	0/58 (0%)	0/13 (0%)
Aromatic	3/7 (43%)	3/4 (75%)	0/2 (0%)	0/1 (0%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Overall	58/298 (19%)	58/167 (35%)	0/98 (0%)	0/33 (0%)

7.1.4 Statistically unusual chemical shifts (i)

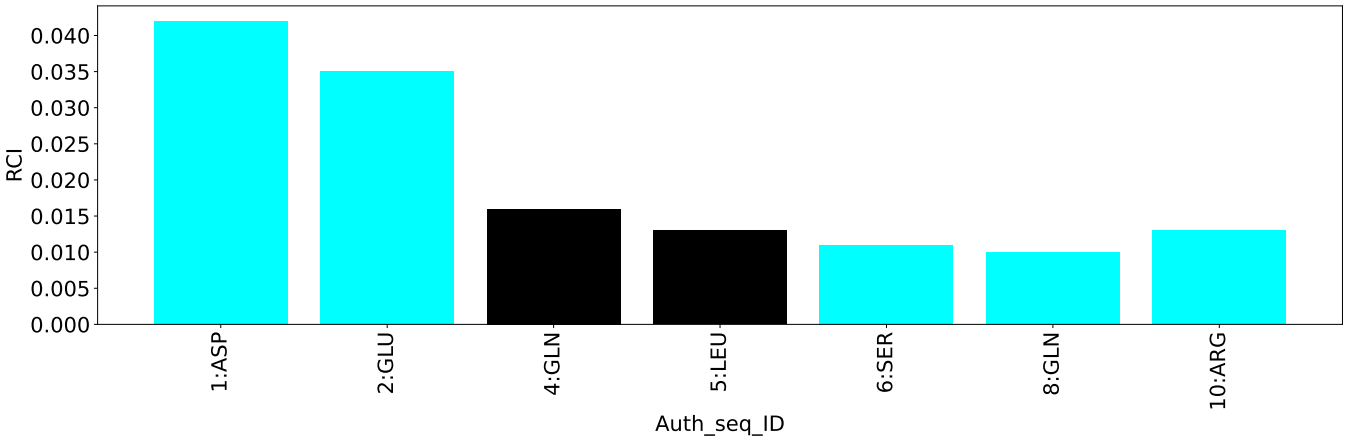
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	B	6	HIS	HD2	-0.46	4.65 – 9.35	-15.9
1	B	6	HIS	HE1	-0.20	5.13 – 10.76	-14.5
1	B	6	HIS	HA	2.30	2.49 – 6.71	-5.5

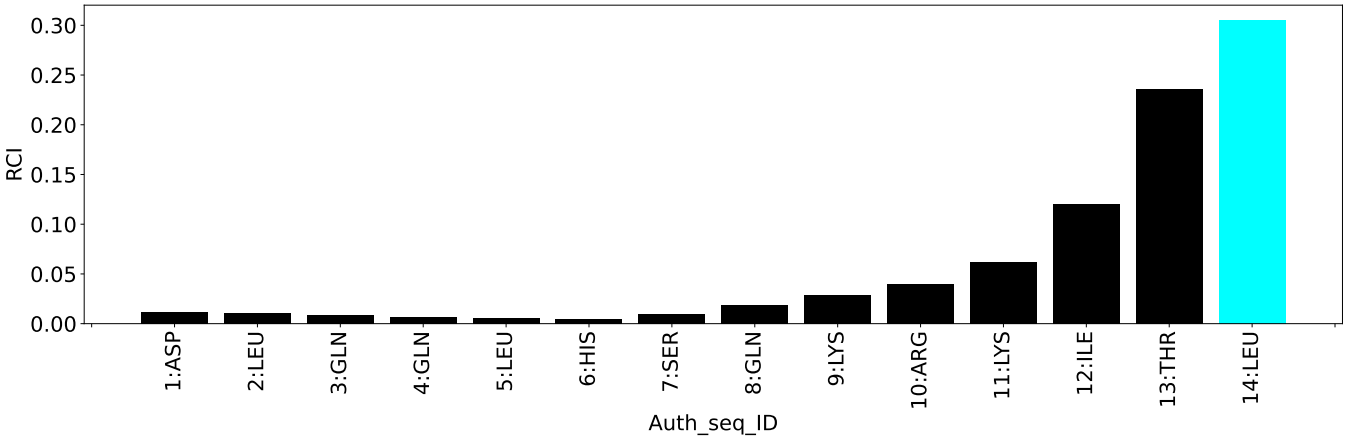
7.1.5 Random Coil Index (RCI) plots (i)

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:



Random coil index (RCI) for chain B:


RCI
0.30
0.25
0.20
0.15
0.10
0.05
0.00
1:ASP
2:LEU
3:GLN
4:GLN
5:LEU
6:HIS
7:SER
8:GLN
9:LYS
10:ARG
11:LYS
12:ILE
13:THR
14:LEU
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	239
Intra-residue (|i-j|=0)	82
Sequential (|i-j|=1)	75
Medium range (|i-j|>1 and |i-j|<5)	52
Long range (|i-j|≥5)	0
Inter-chain	30
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	0
Number of restraints per residue	8.0
Number of long range restraints per residue[1]	0.0

[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

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)	7.7	0.2
0.2-0.5 (Medium)	8.1	0.5
>0.5 (Large)	34.8	5.31

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)

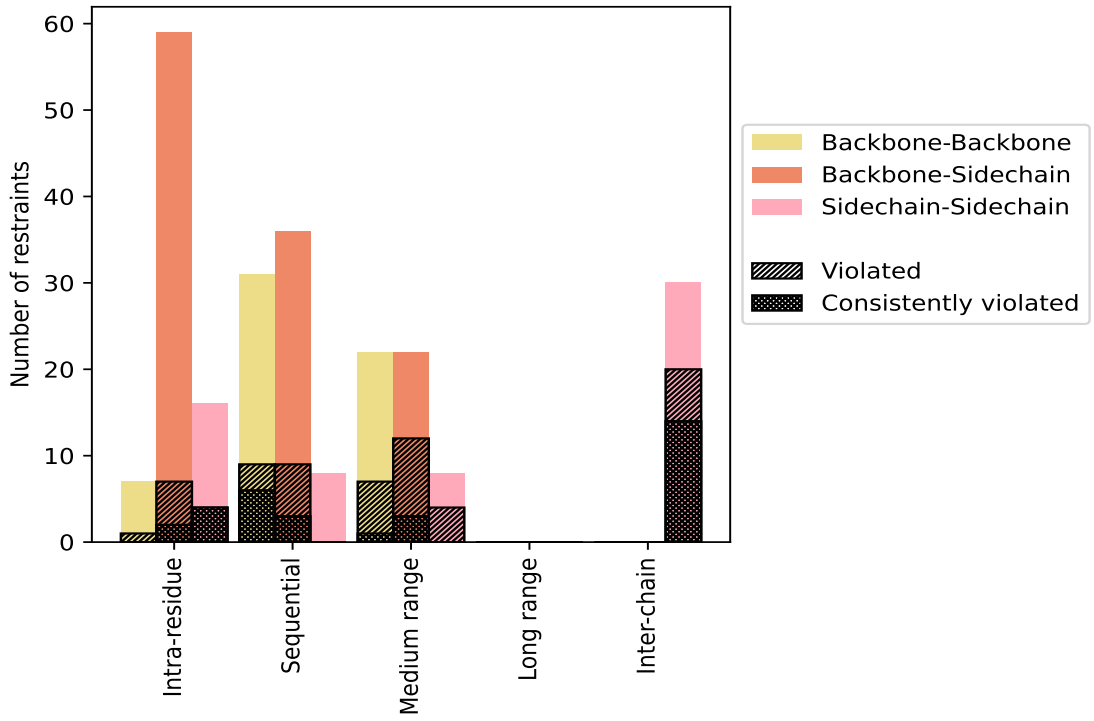
9.1 Summary of distance violations (i)

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)	82	34.3	12	14.6	5.0	6	7.3	2.5
Backbone-Backbone	7	2.9	1	14.3	0.4	0	0.0	0.0
Backbone-Sidechain	59	24.7	7	11.9	2.9	2	3.4	0.8
Sidechain-Sidechain	16	6.7	4	25.0	1.7	4	25.0	1.7
Sequential (|i-j|=1)	75	31.4	18	24.0	7.5	9	12.0	3.8
Backbone-Backbone	31	13.0	9	29.0	3.8	6	19.4	2.5
Backbone-Sidechain	36	15.1	9	25.0	3.8	3	8.3	1.3
Sidechain-Sidechain	8	3.3	0	0.0	0.0	0	0.0	0.0
Medium range (|i-j|>1 & |i-j|<5)	52	21.8	23	44.2	9.6	4	7.7	1.7
Backbone-Backbone	22	9.2	7	31.8	2.9	1	4.5	0.4
Backbone-Sidechain	22	9.2	12	54.5	5.0	3	13.6	1.3
Sidechain-Sidechain	8	3.3	4	50.0	1.7	0	0.0	0.0
Long range (|i-j|≥5)	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
Inter-chain	30	12.6	20	66.7	8.4	14	46.7	5.9
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	30	12.6	20	66.7	8.4	14	46.7	5.9
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	239	100.0	73	30.5	30.5	33	13.8	13.8
Backbone-Backbone	60	25.1	17	28.3	7.1	7	11.7	2.9
Backbone-Sidechain	117	49.0	28	23.9	11.7	8	6.8	3.3
Sidechain-Sidechain	62	25.9	28	45.2	11.7	18	29.0	7.5

[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	9	15	15	0	17	56	1.18	4.25	1.12	0.94
2	9	14	13	0	17	53	1.21	3.98	1.13	0.9
3	8	12	16	0	17	53	1.39	5.07	1.33	0.98
4	8	14	13	0	14	49	1.36	5.18	1.15	1.08
5	8	13	11	0	16	48	1.3	4.64	1.07	1.13
6	8	12	11	0	17	48	1.28	4.46	1.18	1.0
7	7	12	14	0	16	49	1.11	4.55	1.06	0.84
8	9	14	13	0	16	52	1.19	4.03	0.93	1.02
9	7	11	9	0	16	43	1.37	4.78	1.05	1.29
10	8	14	13	0	17	52	1.35	5.17	1.28	1.02

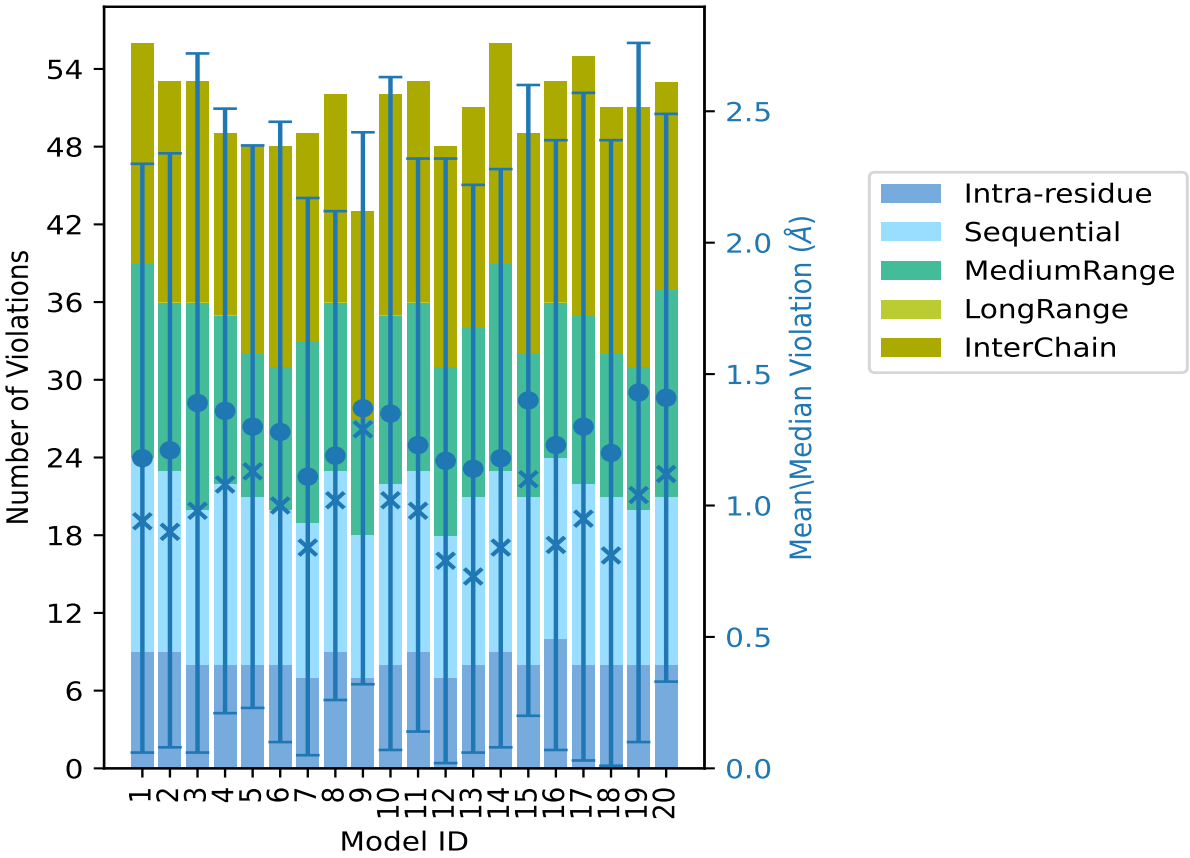
Continued on next page...

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				
11	9	14	13	0	17	53	1.23	4.07	1.09	0.98
12	7	11	13	0	17	48	1.17	4.22	1.15	0.79
13	8	13	13	0	17	51	1.14	3.98	1.08	0.73
14	9	14	16	0	17	56	1.18	4.22	1.1	0.84
15	8	13	11	0	17	49	1.4	4.1	1.2	1.1
16	10	14	12	0	17	53	1.23	5.16	1.16	0.85
17	8	14	13	0	20	55	1.3	5.18	1.27	0.95
18	8	13	11	0	19	51	1.2	4.37	1.19	0.81
19	8	12	11	0	20	51	1.43	5.31	1.33	1.04
20	8	13	16	0	16	53	1.41	4.68	1.08	1.12

[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

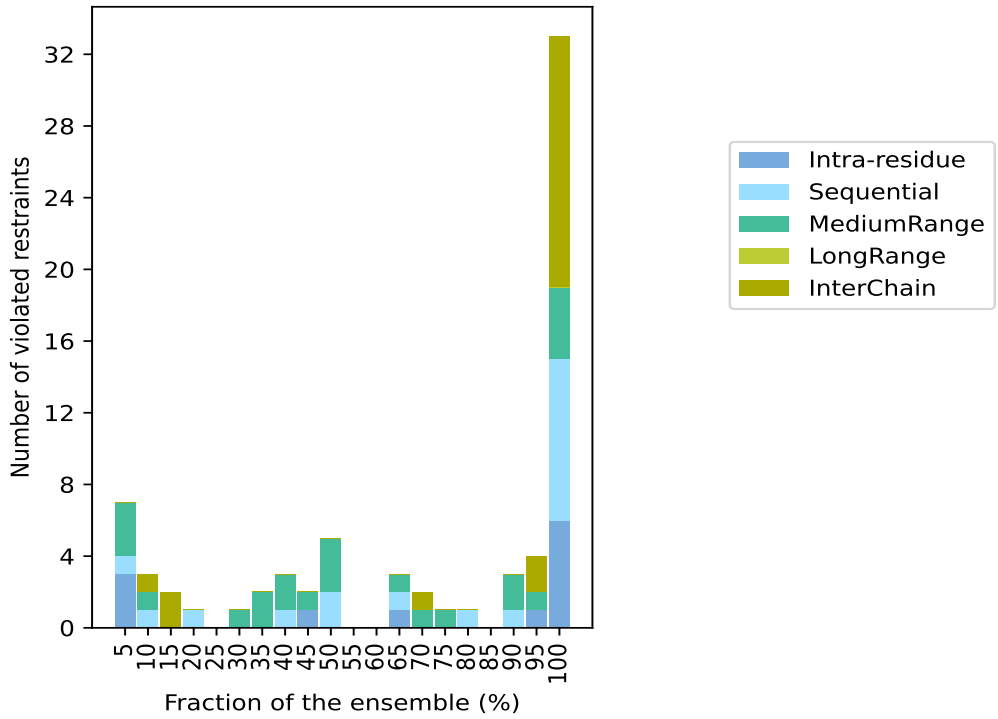
9.3 Distance violation statistics for the ensemble (i)

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, 166(IR:70, SQ:57, MR:29, LR:0, IC:10) 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	1	3	0	0	7	1	5.0
0	1	1	0	1	3	2	10.0
0	0	0	0	2	2	3	15.0
0	1	0	0	0	1	4	20.0
0	0	0	0	0	0	5	25.0
0	0	1	0	0	1	6	30.0
0	0	2	0	0	2	7	35.0
0	1	2	0	0	3	8	40.0
1	0	1	0	0	2	9	45.0
0	2	3	0	0	5	10	50.0
0	0	0	0	0	0	11	55.0
0	0	0	0	0	0	12	60.0
1	1	1	0	0	3	13	65.0
0	0	1	0	1	2	14	70.0
0	0	1	0	0	1	15	75.0
0	1	0	0	0	1	16	80.0
0	0	0	0	0	0	17	85.0
0	1	2	0	0	3	18	90.0
1	0	1	0	2	4	19	95.0
6	9	4	0	14	33	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

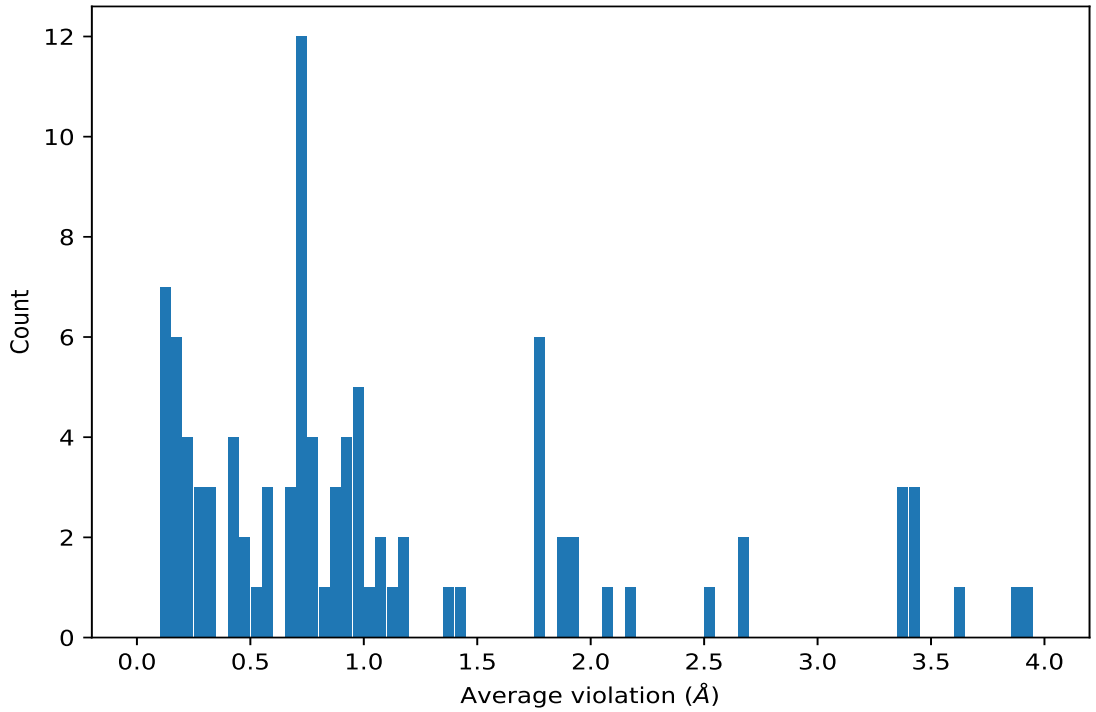
9.3.1 Bar graph : Distance violation statistics for the ensemble



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 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.

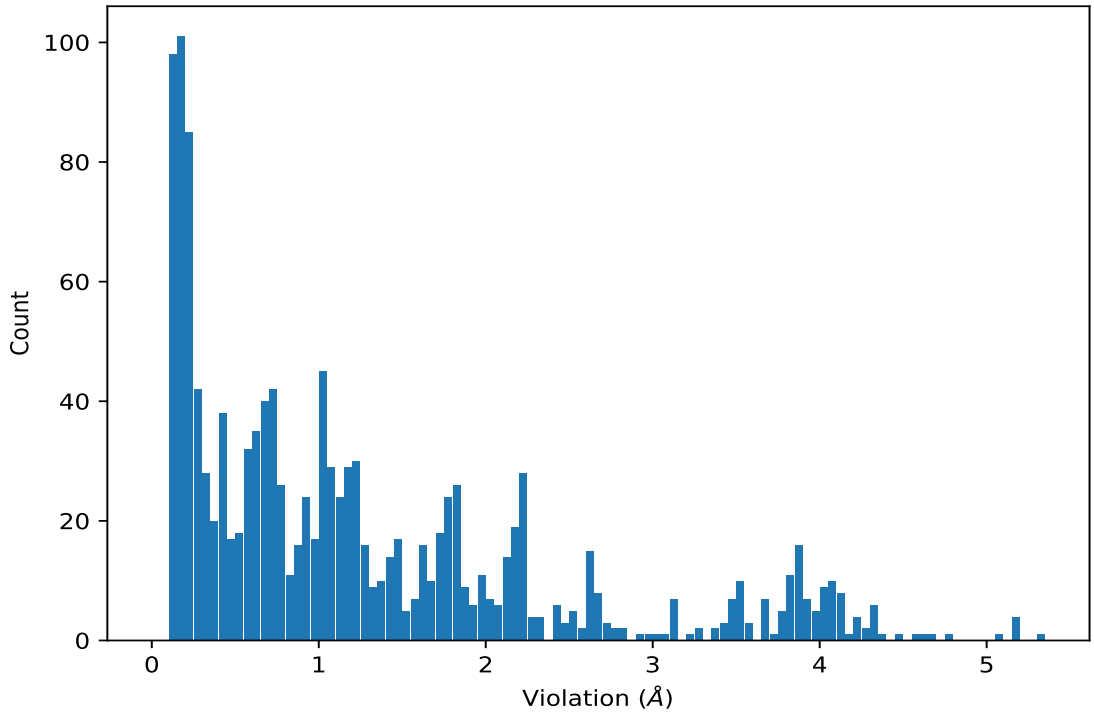
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,1)	1:9:A:DAB:HD1	4:102:A:DEU:HBA1	20	3.93	0.45	3.92
(1,158)	4:102:A:DEU:H2C	4:102:A:DEU:H2B	20	3.85	0.02	3.85
(1,107)	2:15:B:NH2:HN2	2:12:B:ILE:H	20	3.63	1.2	3.46
(1,152)	2:10:B:ARG:HD2	4:102:A:DEU:HHB	20	3.44	0.71	3.39
(1,203)	2:9:B:LYS:HD2	4:102:A:DEU:H2C	20	3.43	0.89	4.0
(1,203)	2:9:B:LYS:HD3	4:102:A:DEU:H2C	20	3.43	0.89	4.0
(2,8)	2:9:B:LYS:HD3	4:102:A:DEU:H2C	20	3.35	1.08	4.07
(2,8)	2:9:B:LYS:HD2	4:102:A:DEU:H2C	20	3.35	1.08	4.07
(2,8)	2:9:B:LYS:HD2	4:102:A:DEU:H2B	20	3.35	1.08	4.07
(1,22)	2:15:B:NH2:HN1	2:12:B:ILE:H	20	2.68	0.86	2.17
(1,22)	2:15:B:NH2:HN2	2:12:B:ILE:H	20	2.68	0.86	2.17
(1,105)	2:14:B:LEU:HA	2:13:B:THR:H	20	2.5	0.23	2.64
(1,141)	2:10:B:ARG:HD2	4:102:A:DEU:HHA	20	2.15	0.67	2.37
(1,43)	2:13:B:THR:HG1	4:102:A:DEU:HAD1	20	2.06	0.23	2.01

[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)
(1,107)	2:15:B:NH2:HN2	2:12:B:ILE:H	19	5.31
(1,107)	2:15:B:NH2:HN2	2:12:B:ILE:H	4	5.18
(1,107)	2:15:B:NH2:HN2	2:12:B:ILE:H	17	5.18
(1,107)	2:15:B:NH2:HN2	2:12:B:ILE:H	10	5.17
(1,107)	2:15:B:NH2:HN2	2:12:B:ILE:H	16	5.16
(1,107)	2:15:B:NH2:HN2	2:12:B:ILE:H	3	5.07
(1,1)	1:9:A:DAB:HD1	4:102:A:DEU:HBA1	9	4.78
(1,1)	1:9:A:DAB:HD1	4:102:A:DEU:HBA1	20	4.68
(1,152)	2:10:B:ARG:HD2	4:102:A:DEU:HHB	5	4.64
(1,1)	1:9:A:DAB:HD1	4:102:A:DEU:HBA1	7	4.55

10 Dihedral-angle violation analysis (i)

No dihedral-angle restraints found