



wwPDB NMR Structure Validation Summary Report ⓘ

Jun 4, 2023 – 09:40 AM EDT

PDB ID : 2NDO
BMRB ID : 19664
Title : Structure of EcDsbA-sulfonamide1 complex
Authors : Williams, M.L.; Doak, B.C.; Vazirani, M.; Ilyichova, O.; Wang, G.; Bermel, W.; Simpson, J.S.; Chalmers, D.K.; King, G.F.; Mobli, M.; Scanlon, M.J.
Deposited on : 2016-08-22

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

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/NMRValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI : v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV : Wang et al. (2010)
wwPDB-ShiftChecker : v1.2
BMRB Restraints Analysis : v1.2
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.33

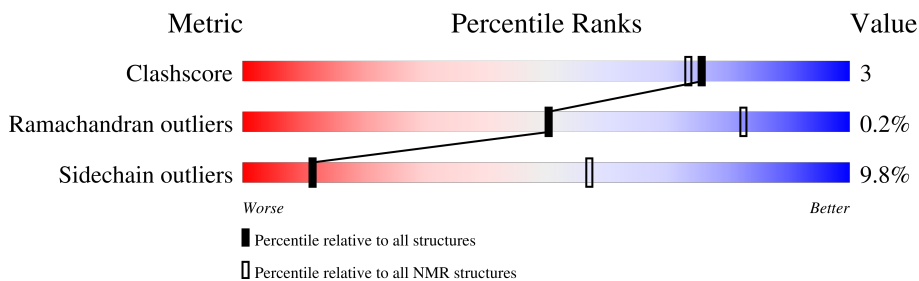
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

SOLUTION NMR

The overall completeness of chemical shifts assignment is 17%.

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	189	 88% 10% .

2 Ensemble composition and analysis

This entry contains 10 models. Model 1 is the overall representative, medoid model (most similar to other models).

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

Well-defined (core) protein residues			
Well-defined core	Residue range (total)	Backbone RMSD (Å)	Medoid model
1	A:3-A:187 (185)	0.50	1

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

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

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

Cluster number	Models
1	3, 6, 7, 8, 9, 10
2	1, 2, 4, 5

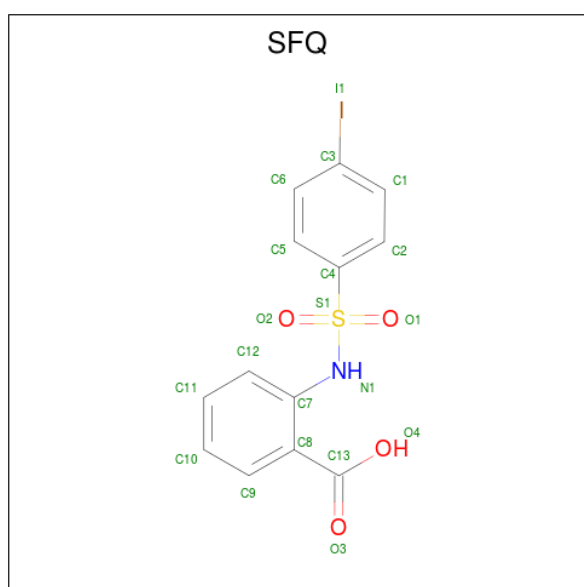
3 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 2984 atoms, of which 1476 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Thiol:disulfide interchange protein DsbA.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	189	2955	952	1467	244	284	8	0

- Molecule 2 is 2-[[4-iodophenyl]sulfonyl]amino}benzoic acid (three-letter code: SFQ) (formula: C₁₃H₁₀INO₄S).



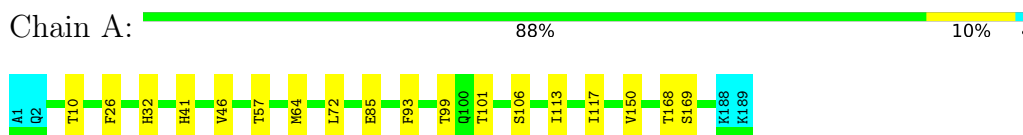
Mol	Chain	Residues	Atoms						
			Total	C	H	I	N	O	S
2	A	1	29	13	9	1	1	4	1

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.

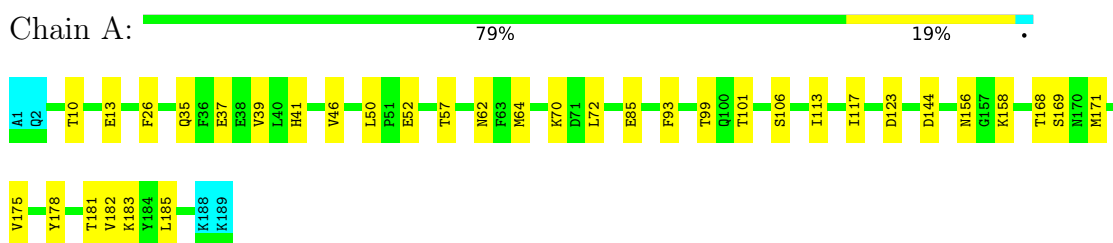
- Molecule 1: Thiol:disulfide interchange protein DsbA



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

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

- Molecule 1: Thiol:disulfide interchange protein DsbA



5 Refinement protocol and experimental data overview

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

Of the 200 calculated structures, 10 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
HADDOCK	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	454
Number of shifts mapped to atoms	454
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	17%

6 Model quality [i](#)

6.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SFQ

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

6.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	1455	1426	1423	8±3
2	A	20	9	0	2±1
All	All	14750	14350	14230	91

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 3.

5 of 47 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:35:GLN:O	1:A:39:VAL:HB	0.61	1.96	6	5
1:A:59:TYR:CD1	1:A:138:GLN:HG2	0.58	2.33	4	1
2:A:201:SFQ:O4	2:A:201:SFQ:N1	0.57	2.37	7	10
1:A:113:ILE:HA	1:A:117:ILE:O	0.56	2.00	3	7
1:A:46:VAL:O	1:A:50:LEU:HG	0.56	2.01	1	5

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	185/189 (98%)	177±3 (96±1%)	7±3 (4±2%)	0±0 (0±0%)	50	82
All	All	1850/1890 (98%)	1773 (96%)	73 (4%)	4 (0%)	50	82

All 1 unique Ramachandran outliers are listed below.

Mol	Chain	Res	Type	Models (Total)
1	A	150	VAL	4

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	156/159 (98%)	141±2 (90±1%)	15±2 (10±1%)	11	57
All	All	1560/1590 (98%)	1407 (90%)	153 (10%)	11	57

5 of 42 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	101	THR	10
1	A	168	THR	10
1	A	10	THR	9
1	A	57	THR	9
1	A	99	THR	9

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](#)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
2	SFQ	A	201	-	21,21,21	3.18±0.02	4±0 (19±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
2	SFQ	A	201	-	30,30,30	1.87±0.02	2±0 (6±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
2	SFQ	A	201	-	-	0±0,15,15,15	0±0,2,2,2

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
2	A	201	SFQ	O2-S1	8.94	1.53	1.43	6	10
2	A	201	SFQ	O1-S1	8.81	1.53	1.43	2	10
2	A	201	SFQ	C7-N1	5.66	1.33	1.42	7	10
2	A	201	SFQ	C8-C13	4.88	1.39	1.49	8	10

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

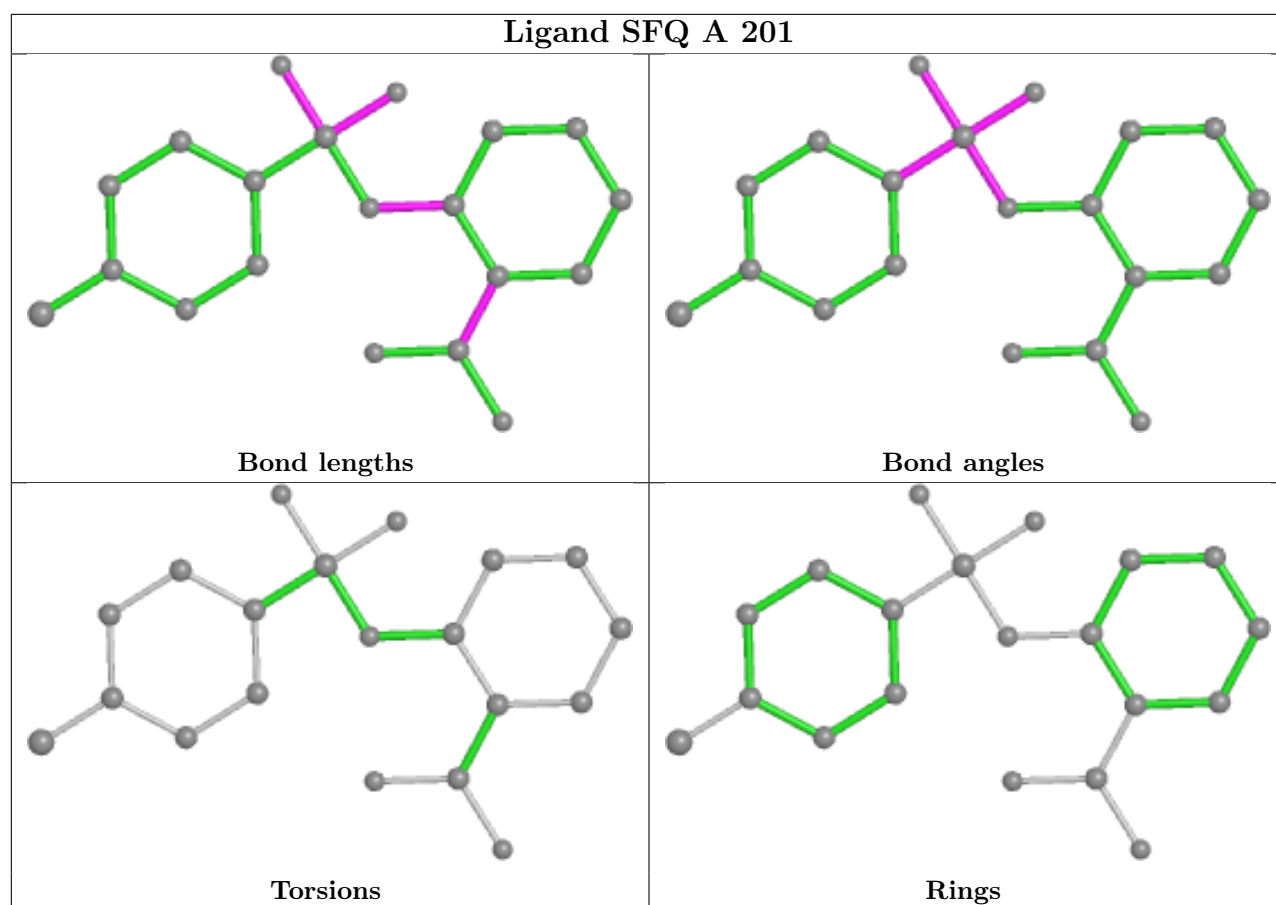
Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)	Models	
								Worst	Total
2	A	201	SFQ	O2-S1-O1	8.93	108.57	119.55	9	10
2	A	201	SFQ	C4-S1-N1	2.74	110.28	106.83	4	8

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 than 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.



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 17% for the well-defined parts and 17% 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	454
Number of shifts mapped to atoms	454
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	3

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

- Chemical shift has been reported more than once. First 5 (of 0) occurrences are reported below.

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	185	LEU	HD11	-1.191	0.000	.
1	A	185	LEU	HD12	-1.191	0.000	.
1	A	185	LEU	HD13	-1.191	0.000	.
1	A	161	LEU	HD11	0.296	0.000	.
1	A	161	LEU	HD12	0.296	0.000	.
1	A	161	LEU	HD13	0.296	0.000	.
1	A	23	LEU	HD11	0.314	0.000	.
1	A	23	LEU	HD12	0.314	0.000	.
1	A	23	LEU	HD13	0.314	0.000	.
1	A	92	LEU	HD11	0.413	0.000	.
1	A	92	LEU	HD12	0.413	0.000	.
1	A	92	LEU	HD13	0.413	0.000	.
1	A	50	LEU	HD11	0.546	0.000	.
1	A	50	LEU	HD12	0.546	0.000	.

Continued on next page...

Continued from previous page...

List ID	Chain	Res	Type	Atom	Shift Data		
					Value	Uncertainty	Ambiguity
1	A	50	LEU	HD13	0.546	0.000	.
1	A	72	LEU	HD11	0.568	0.000	.
1	A	72	LEU	HD12	0.568	0.000	.
1	A	72	LEU	HD13	0.568	0.000	.
1	A	40	LEU	HD11	0.589	0.000	.
1	A	40	LEU	HD12	0.589	0.000	.
1	A	40	LEU	HD13	0.589	0.000	.
1	A	82	LEU	HD11	0.663	0.000	.
1	A	82	LEU	HD12	0.663	0.000	.
1	A	82	LEU	HD13	0.663	0.000	.
1	A	134	LEU	HD11	0.749	0.000	.
1	A	134	LEU	HD12	0.749	0.000	.
1	A	134	LEU	HD13	0.749	0.000	.
1	A	68	LEU	HD11	0.771	0.000	.
1	A	68	LEU	HD12	0.771	0.000	.
1	A	68	LEU	HD13	0.771	0.000	.
1	A	12	LEU	HD11	0.847	0.000	.
1	A	12	LEU	HD12	0.847	0.000	.
1	A	12	LEU	HD13	0.847	0.000	.

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 17%, i.e. 418 atoms were assigned a chemical shift out of a possible 2488. 0 out of 33 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	¹ H	¹³ C	¹⁵ N
Backbone	0/924 (0%)	0/376 (0%)	0/370 (0%)	0/178 (0%)
Sidechain	418/1334 (31%)	312/867 (36%)	106/423 (25%)	0/44 (0%)
Aromatic	0/230 (0%)	0/111 (0%)	0/111 (0%)	0/8 (0%)
Overall	418/2488 (17%)	312/1354 (23%)	106/904 (12%)	0/230 (0%)

7.1.4 Statistically unusual chemical shifts [i](#)

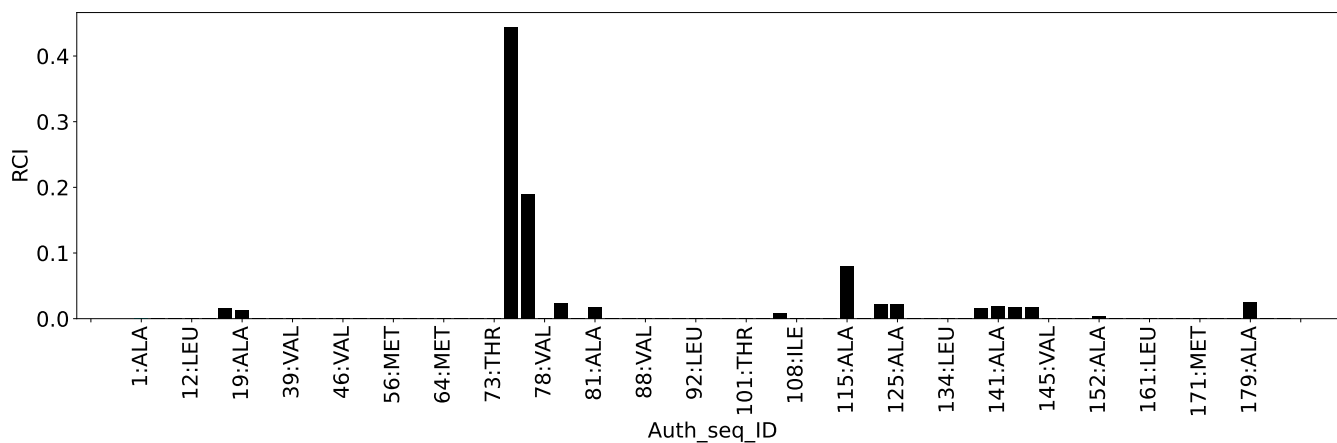
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

List Id	Chain	Res	Type	Atom	Shift, ppm	Expected range, ppm	Z-score
1	A	185	LEU	HD21	-1.19	-0.65 – 2.13	-6.9
1	A	185	LEU	HD22	-1.19	-0.65 – 2.13	-6.9
1	A	185	LEU	HD23	-1.19	-0.65 – 2.13	-6.9

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

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	19
Intra-residue ($ i-j =0$)	0
Sequential ($ i-j =1$)	0
Medium range ($ i-j >1$ and $ i-j <5$)	0
Long range ($ i-j \geq 5$)	19
Inter-chain	0
Hydrogen bond restraints	0
Disulfide bond restraints	0
Total dihedral-angle restraints	0
Number of unmapped restraints	0
Number of restraints per residue	0.1
Number of long range restraints per residue ¹	0.1

¹Long 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. There are no distance violations

8.2.2 Average number of dihedral-angle violations per model

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

9 Distance violation analysis

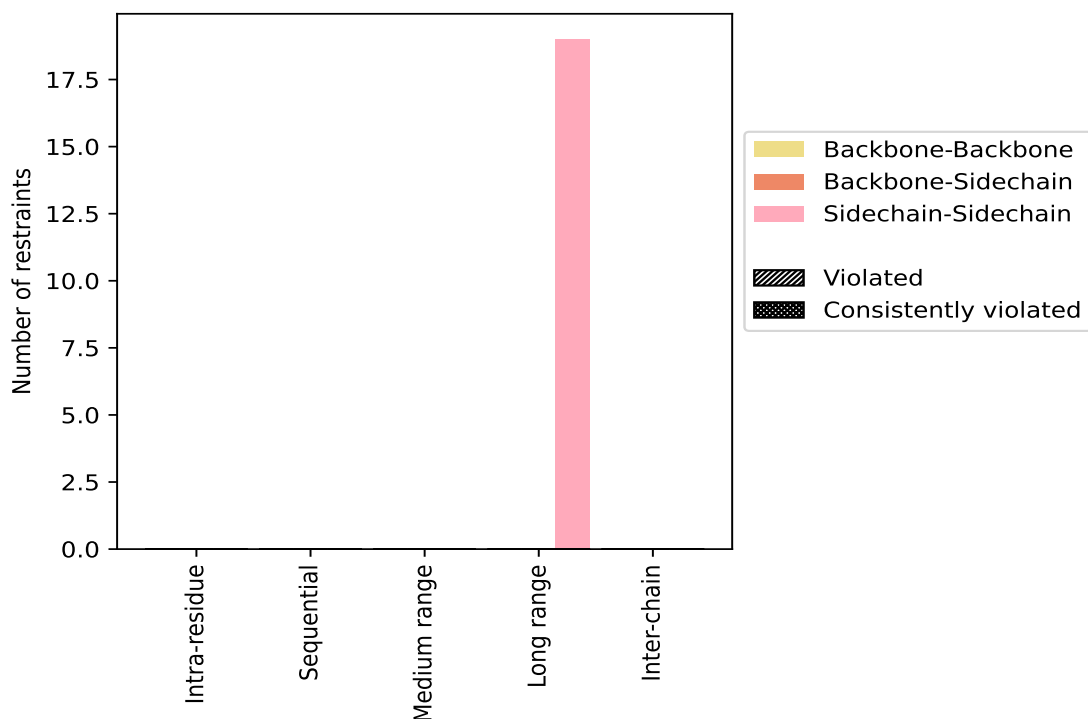
9.1 Summary of distance violations

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

Restrains type	Count	% ¹	Violated ³			Consistently Violated ⁴		
			Count	% ²	% ¹	Count	% ²	% ¹
Intra-residue (i-j =0)	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	0	0.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	0	0.0	0	0.0	0.0	0	0.0	0.0
Sequential (i-j =1)	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
Medium range (i-j >1 & 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
Long range (i-j ≥5)	19	100.0	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	19	100.0	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	19	100.0	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	19	100.0	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 disulfid bonds are counted in their appropriate category on the x-axis

9.2 Distance violation statistics for each model [i](#)

No violations found

9.3 Distance violation statistics for the ensemble [i](#)

No violations found

9.4 Most violated distance restraints in the ensemble [i](#)

No violations found

9.5 All violated distance restraints [i](#)

No violations found

10 Dihedral-angle violation analysis

No dihedral-angle restraints found