Full wwPDB NMR Structure Validation Report (i)

Jun 6, 2023 – 08:03 pm BST

PDB ID	:	6QAX
BMRB ID	:	34340
Title	:	P31-43
Authors	:	Calvanese, L.; D'Auria, G.; Falcigno, F.
Deposited on	:	2018-12-19

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:

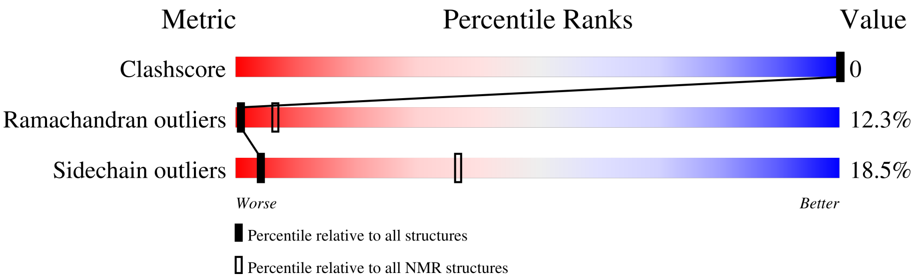
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal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

1 Overall quality at a glance (i)

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

The overall completeness of chemical shifts assignment is 46%.

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	13	85% 8% 8%

2 Ensemble composition and analysis (i)

This entry contains 40 models.

Cyrange was unable to find well-defined residues.

Error message: Only domains with < 8 residues could be identified.

NmrClust was unable to cluster the ensemble.

Error message: Wrapper check: not enough residues in core to run NmrClust

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 208 atoms, of which 100 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN -PRO-TYR.

Mol	Chain	Residues	Atoms					Trace
			Total	C	H	N	O	
1	A	13	208	71	100	18	19	0

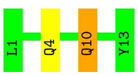
4 Residue-property plots (i)

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: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

Chain A:



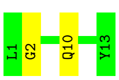
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

- Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

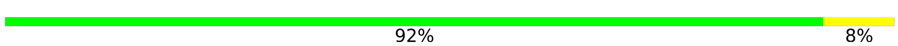
Chain A:

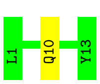


4.2.2 Score per residue for model 2

- Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

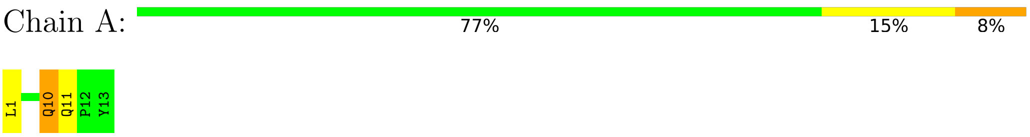
Chain A:





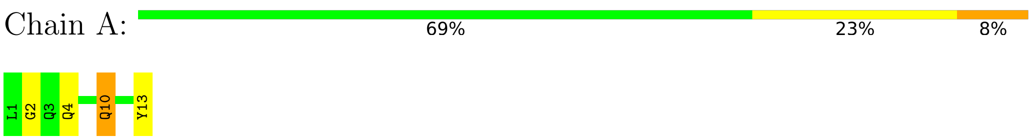
4.2.3 Score per residue for model 3

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR



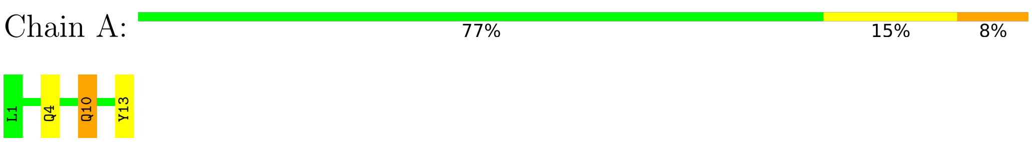
4.2.4 Score per residue for model 4

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR



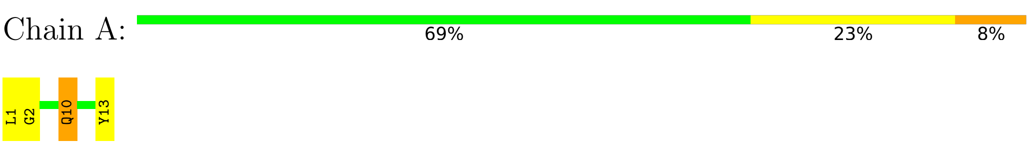
4.2.5 Score per residue for model 5

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR



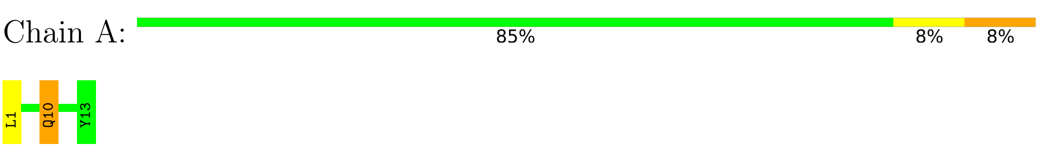
4.2.6 Score per residue for model 6

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR



4.2.7 Score per residue for model 7

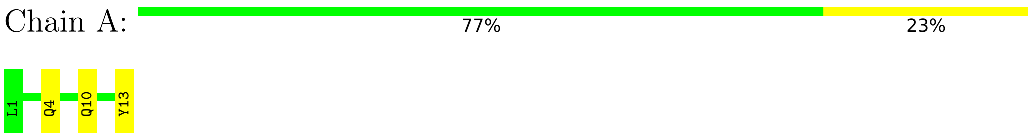
• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR



4.2.8 Score per residue for model 8

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

Chain A:



4.2.9 Score per residue for model 9

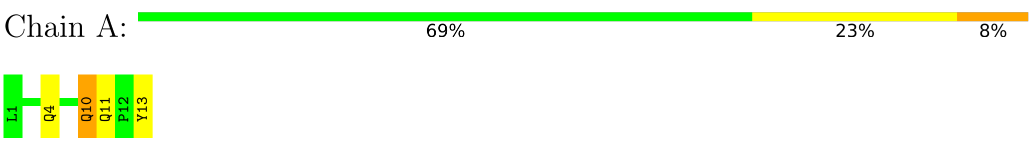
• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

Chain A:

4.2.10 Score per residue for model 10

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

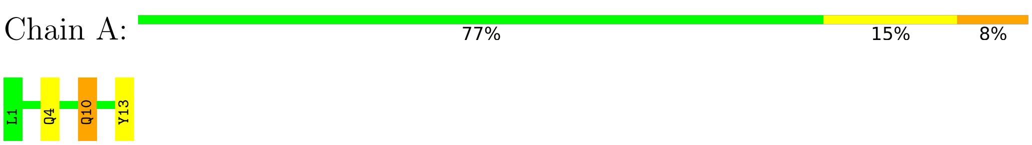
Chain A:



4.2.11 Score per residue for model 11

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

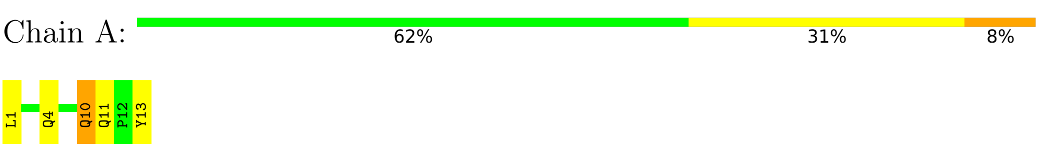
Chain A:



4.2.12 Score per residue for model 12

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

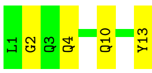
Chain A:



4.2.13 Score per residue for model 13

- Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

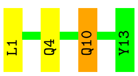
Chain A:

4.2.14 Score per residue for model 14

- Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

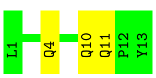
Chain A:

4.2.15 Score per residue for model 15

- Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

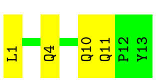
Chain A:

4.2.16 Score per residue for model 16

- Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

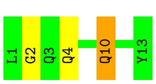
Chain A:

4.2.17 Score per residue for model 17

- Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

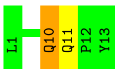
Chain A:

4.2.18 Score per residue for model 18

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

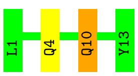
Chain A:



4.2.19 Score per residue for model 19

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

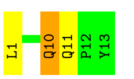
Chain A:



4.2.20 Score per residue for model 20

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

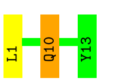
Chain A:



4.2.21 Score per residue for model 21

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

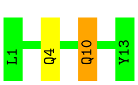
Chain A:



4.2.22 Score per residue for model 22

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

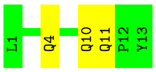
Chain A:



4.2.23 Score per residue for model 23

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

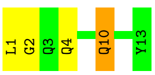
Chain A:



4.2.24 Score per residue for model 24

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

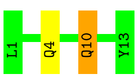
Chain A:



4.2.25 Score per residue for model 25

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

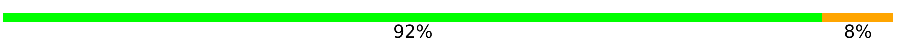
Chain A:

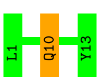


4.2.26 Score per residue for model 26

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

Chain A:

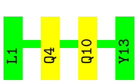




4.2.27 Score per residue for model 27

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

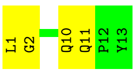
Chain A:



4.2.28 Score per residue for model 28

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

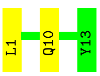
Chain A:



4.2.29 Score per residue for model 29

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

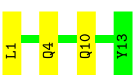
Chain A:



4.2.30 Score per residue for model 30

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

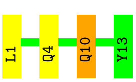
Chain A:



4.2.31 Score per residue for model 31

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

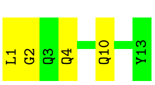
Chain A:



4.2.32 Score per residue for model 32

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

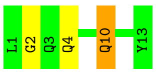
Chain A:



4.2.33 Score per residue for model 33

- Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

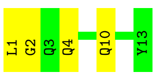
Chain A:



4.2.34 Score per residue for model 34

- Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

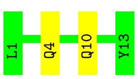
Chain A:



4.2.35 Score per residue for model 35

- Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

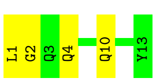
Chain A:



4.2.36 Score per residue for model 36

- Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

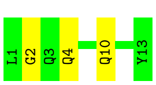
Chain A:



4.2.37 Score per residue for model 37

- Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

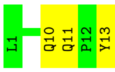
Chain A:



4.2.38 Score per residue for model 38

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

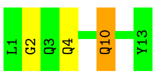
Chain A:



4.2.39 Score per residue for model 39

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

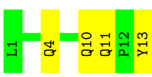
Chain A:



4.2.40 Score per residue for model 40

• Molecule 1: LEU-GLY-GLN-GLN-GLN-PRO-PHE-PRO-PRO-GLN-GLN-PRO-TYR

Chain A:



5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *na*.

Of the 100 calculated structures, 40 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 calculation	

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	83
Number of shifts mapped to atoms	83
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	46%

6 Model quality (i)

6.1 Standard geometry (i)

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
All	All	4320	4000	4080	-

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 -.

There are no clashes.

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	11/13 (85%)	6±1 (55±8%)	4±1 (32±8%)	1±0 (12±4%)	1	6
All	All	440/520 (85%)	243 (55%)	143 (32%)	54 (12%)	1	6

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)
1	A	10	GLN	40
1	A	2	GLY	14

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	12/12 (100%)	10±1 (81±8%)	2±1 (19±8%)	4	37
All	All	480/480 (100%)	391 (81%)	89 (19%)	4	37

All 5 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	4	GLN	28
1	A	10	GLN	22
1	A	1	LEU	17
1	A	11	GLN	11
1	A	13	TYR	11

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 46% for the well-defined parts and 46% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *1838.str*

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	83
Number of shifts mapped to atoms	83
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

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 46%, i.e. 83 atoms were assigned a chemical shift out of a possible 182. 0 out of 1 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	**^{1}H**	**^{13}C**	**^{15}N**
Backbone	22/58 (38%)	22/23 (96%)	0/26 (0%)	0/9 (0%)
Sidechain	55/105 (52%)	55/67 (82%)	0/33 (0%)	0/5 (0%)
Aromatic	6/19 (32%)	6/9 (67%)	0/10 (0%)	0/0 (—%)
Overall	83/182 (46%)	83/99 (84%)	0/69 (0%)	0/14 (0%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 46%, i.e. 83 atoms were assigned a chemical shift out of a possible 182. 0 out of 1 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	22/58 (38%)	22/23 (96%)	0/26 (0%)	0/9 (0%)
Sidechain	55/105 (52%)	55/67 (82%)	0/33 (0%)	0/5 (0%)
Aromatic	6/19 (32%)	6/9 (67%)	0/10 (0%)	0/0 (—%)
Overall	83/182 (46%)	83/99 (84%)	0/69 (0%)	0/14 (0%)

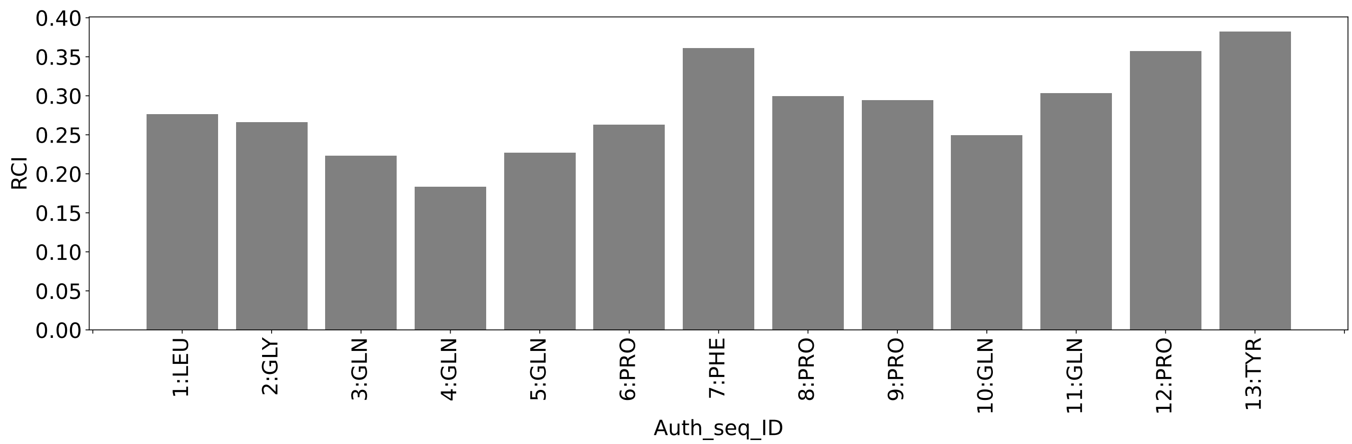
7.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

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	47
Intra-residue (|i-j|=0)	34
Sequential (|i-j|=1)	13
Medium range (|i-j|>1 and |i-j|<5)	0
Long range (|i-j|≥5)	0
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	3.6
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 (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.There are no distance violations

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)

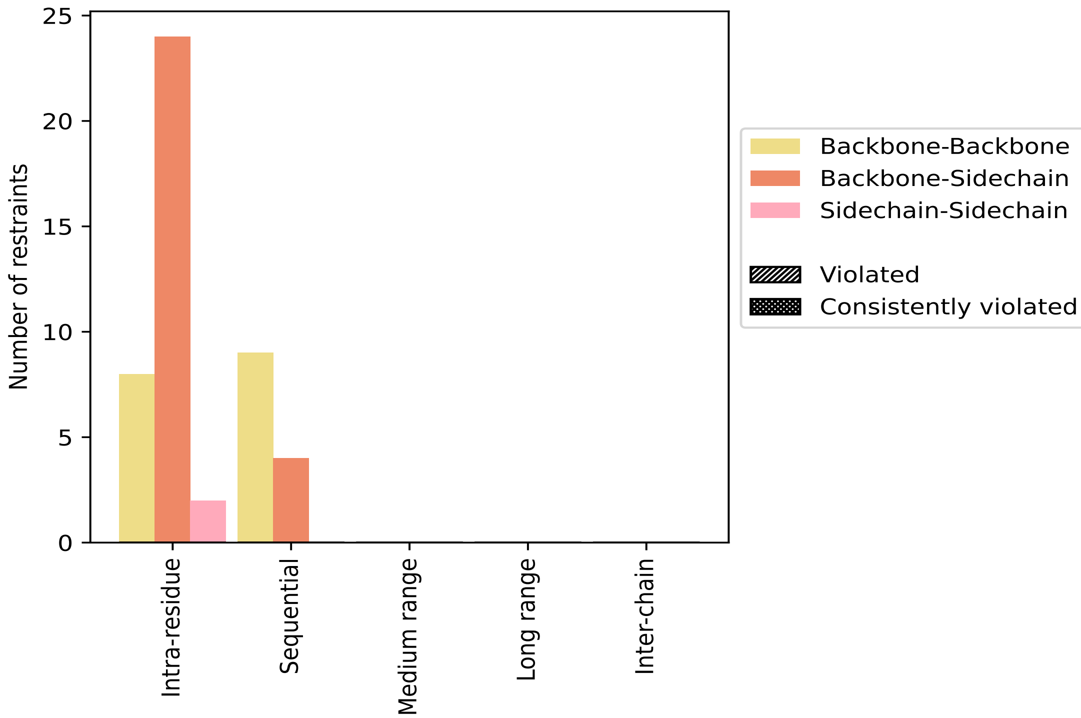
9.1 Summary of distance violations (i)

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)	34	72.3	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	8	17.0	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	24	51.1	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	2	4.3	0	0.0	0.0	0	0.0	0.0
Sequential (|i-j|=1)	13	27.7	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	9	19.1	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	4	8.5	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)	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	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	47	100.0	0	0.0	0.0	0	0.0	0.0
Backbone-Backbone	17	36.2	0	0.0	0.0	0	0.0	0.0
Backbone-Sidechain	28	59.6	0	0.0	0.0	0	0.0	0.0
Sidechain-Sidechain	2	4.3	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)

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 (i)

No dihedral-angle restraints found