# Full wwPDB NMR Structure Validation Report (i)

May 29, 2020 – 01:04 am BST

| | | |
|---|---|---|
| PDB ID | : | 2MUZ |
| Title | : | ssNMR structure of a designed rocker protein |
| Authors | : | Wang, T.; Joh, N.; Wu, Y.; DeGrado, W.F.; Hong, M. |
| Deposited on | : | 2014-09-18 |

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 following versions of software and data (see references (1)) were used in the production of this report:

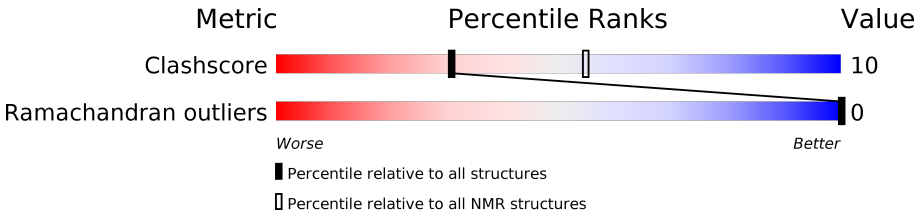
| | | |
|---|---|---|
| Cyrange | : | Kirchner and Güntert (2011) |
| NmrClust | : | Kelley et al. (1996) |
| MolProbity | : | 4.02b-467 |
| Mogul | : | 1.8.5 (274361), CSD as541be (2020) |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| RCI | : | v_1n_11_5_13_A (Berjanski et al., 2005) |
| PANAV | : | Wang et al. (2010) |
| ShiftChecker | : | 2.11 |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.11 |

# 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 was not calculated.

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 |

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 | 26 | 85% 15% |
| 1 | B | 26 | 100% |
| 1 | C | 26 | 85% 15% |
| 1 | D | 26 | 100% |

## 2 Ensemble composition and analysis (i)

This entry contains 1 models. Identification of well-defined residues and clustering analysis are not possible.

# 3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 464 atoms, of which 0 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called designed rocker protein.

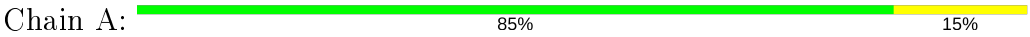
| Mol | Chain | Residues | Atoms | | | | | Trace |
|---|---|---|---|---|---|---|---|---|
| 1 | A | 26 | Total | C | F | N | O | 1 |
| | | | 116 | 63 | 2 | 26 | 25 | |
| 1 | B | 26 | Total | C | F | N | O | 1 |
| | | | 116 | 63 | 2 | 26 | 25 | |
| 1 | C | 26 | Total | C | F | N | O | 1 |
| | | | 116 | 63 | 2 | 26 | 25 | |
| 1 | D | 26 | Total | C | F | N | O | 1 |
| | | | 116 | 63 | 2 | 26 | 25 | |

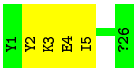
# 4 Residue-property plots

These plots are provided for all protein, RNA and DNA 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: designed rocker protein

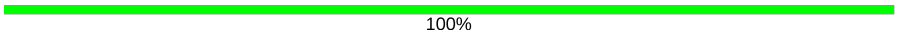
Chain A:




- Molecule 1: designed rocker protein

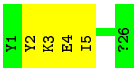
Chain B:




There are no outlier residues in this chain.

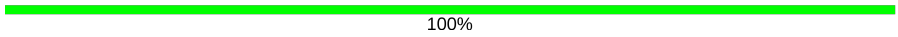
- Molecule 1: designed rocker protein

Chain C:


- Molecule 1: designed rocker protein

Chain D:




There are no outlier residues in this chain.

# 5 Refinement protocol and experimental data overview

The models were refined using the following method: *Torsion angles ajustments using Chimera and Pymol, CODEX data fitting by Matlab.*.

Of the 1 calculated structures, 1 were deposited, based on the following criterion: *structures with the least restraint violations.*

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

| Software name | Classification | Version |
|---|---|---|
| Matlab | 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) | input_cs.cif |
|---|---|
| Number of chemical shift lists | 1 |
| Total number of shifts | 1 |
| Number of shifts mapped to atoms | 1 |
| Number of unparsed shifts | 0 |
| Number of shifts with mapping errors | 0 |
| Number of shifts with mapping warnings | 0 |
| Assignment completeness (well-defined parts) | 0% |

No validations of the models with respect to experimental NMR restraints is performed at this time.

# 6 Model quality

## 6.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PFF, NH2

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

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## 6.2 Too-close contacts

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

| Mol | Chain | Non-H | H(model) | H(added) | Clashes |
|---|---|---|---|---|---|
| 1 | A | 116 | 0 | 38 | 3 |
| 1 | C | 116 | 0 | 38 | 3 |
| All | All | 464 | 0 | 152 | 6 |

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 10.

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

| Atom-1 | Atom-2 | Clash(Å) | Distance(Å) |
|---|---|---|---|
| 1:C:2:TYR:O | 1:C:5:ILE:N | 0.63 | 2.32 |
| 1:A:2:TYR:O | 1:A:5:ILE:N | 0.62 | 2.32 |
| 1:A:2:TYR:O | 1:A:3:LYS:C | 0.42 | 2.56 |
| 1:C:2:TYR:O | 1:C:3:LYS:C | 0.41 | 2.56 |
| 1:C:2:TYR:O | 1:C:4:GLU:N | 0.41 | 2.54 |
| 1:A:2:TYR:O | 1:A:4:GLU:N | 0.41 | 2.54 |

## 6.3 Torsion angles

### 6.3.1 Protein backbone

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 | 23/26 (88%) | 22 (96%) | 1 (4%) | 0 (0%) | 100 | 100 |
| 1 | B | 23/26 (88%) | 22 (96%) | 1 (4%) | 0 (0%) | 100 | 100 |
| 1 | C | 23/26 (88%) | 22 (96%) | 1 (4%) | 0 (0%) | 100 | 100 |
| 1 | D | 23/26 (88%) | 22 (96%) | 1 (4%) | 0 (0%) | 100 | 100 |
| All | All | 92/104 (88%) | 88 (96%) | 4 (4%) | 0 (0%) | 100 | 100 |

There are no Ramachandran outliers.

### 6.3.2 Protein sidechains

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 | 0 | - | - | - |
| 1 | B | 0 | - | - | - |
| 1 | C | 0 | - | - | - |
| 1 | D | 0 | - | - | - |
| All | All | 0 | - | - | - |

There are no protein residues with a non-rotameric sidechain to report.

### 6.3.3 RNA

There are no RNA molecules in this entry.

MODRES-GEOMETRY INFOmissingINFO

## 6.4 Carbohydrates

There are no carbohydrates in this entry.

## 6.5 Ligand geometry

There are no ligands in this entry.

## 6.6 Other polymers

There are no such molecules in this entry.

## 6.7 Polymer linkage issues

There are no chain breaks in this entry.

# 7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 0% for the well-defined parts and 0% for the entire structure.

## 7.1 Chemical shift list 1

File name: input_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

### 7.1.1 Bookkeeping

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

| | |
|---|---|
| Total number of shifts | 1 |
| Number of shifts mapped to atoms | 1 |
| 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

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

### 7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 0%, i.e. 0 atoms were assigned a chemical shift out of a possible 1224. 0 out of 20 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

| | Total | $^{1}$H | $^{13}$C | $^{15}$N |
|---|---|---|---|---|
| Backbone | 0/480 (0%) | 0/192 (0%) | 0/192 (0%) | 0/96 (0%) |
| Sidechain | 0/556 (0%) | 0/324 (0%) | 0/216 (0%) | 0/16 (0%) |
| Aromatic | 0/188 (0%) | 0/100 (0%) | 0/88 (0%) | 0/0 (—%) |
| Overall | 0/1224 (0%) | 0/616 (0%) | 0/496 (0%) | 0/112 (0%) |

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 0%, i.e. 0 atoms were assigned a chemical shift out of a possible 1224. 0 out of 20 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

| | Total | $^{1}$H | $^{13}$C | $^{15}$N |
|---|---|---|---|---|
| Backbone | 0/480 (0%) | 0/192 (0%) | 0/192 (0%) | 0/96 (0%) |
| Sidechain | 0/556 (0%) | 0/324 (0%) | 0/216 (0%) | 0/16 (0%) |
| Aromatic | 0/188 (0%) | 0/100 (0%) | 0/88 (0%) | 0/0 (—%) |
| Overall | 0/1224 (0%) | 0/616 (0%) | 0/496 (0%) | 0/112 (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)

No *random coil index* (RCI) plot could be generated from the current chemical shift list (assigned_chem_shift_list_1). RCI is only applicable to proteins.