# wwPDB NMR Structure Validation Summary Report (i)

Apr 26, 2016 – 03:47 PM BST

| | | |
|---|---|---|
| PDB ID | : | 1I16 |
| Title | : | STRUCTURE OF INTERLEUKIN 16: IMPLICATIONS FOR FUNCTION, NMR, 20 STRUCTURES |
| Authors | : | Muehlhahn, P.; Zweckstetter, M.; Georgescu, J.; Ciosto, C.; Renner, C.; Lanzendoerfer, M.; Lang, K.; Ambrosius, D.; Baier, M.; Kurth, R.; Holak, T.A. |
| Deposited on | : | 1998-05-20 |

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
http://wwpdb.org/validation/2016/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

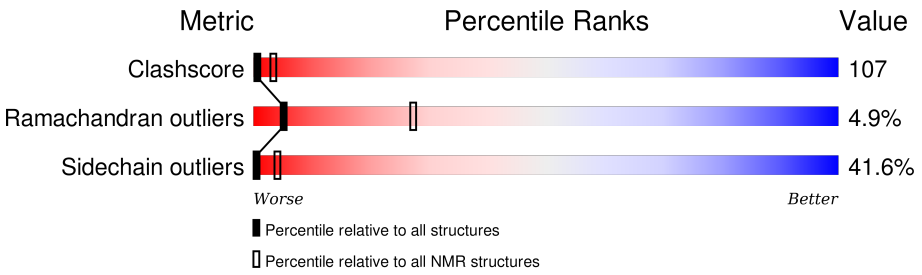
| | | |
|---|---|---|
| Cyrange | : | Kirchner and Güntert (2011) |
| NmrClust | : | Kelley et al. (1996) |
| MolProbity | : | 4.02b-467 |
| Mogul | : | unknown |
| Percentile statistics | : | 20151230.v01 (using entries in the PDB archive December 30th 2015) |
| RCI | : | v_1n_11_5_13_A (Berjanski et al., 2005) |
| PANAV | : | Wang et al. (2010) |
| ShiftChecker | : | rb-20027457 |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | rb-20027457 |

# 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 | 114402 | 11133 |
| Ramachandran outliers | 111179 | 9975 |
| Sidechain outliers | 111093 | 9958 |

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 | 130 | |

# 2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 8 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:29-A:37, A:41-A:47, A:59-A:65, A:74-A:116 (66) | 0.30 | 8 |

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

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

The models can be grouped into 3 clusters and 6 single-model clusters were found.

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

# 3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1869 atoms, of which 933 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called INTERLEUKIN 16.

| Mol | Chain | Residues | Atoms | | | | | | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | H | N | O | S | |
| 1 | A | 130 | 1869 | 573 | 933 | 160 | 199 | 4 | 0 |

# 4 Residue-property plots (i)

## 4.1 Average score per residue in the NMR ensemble

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: INTERLEUKIN 16

Chain A: 6% 27% 16% • 49%

M1 P2 D3 L4 N5 S6 S7 T8 D9 S10 A11 A12 S13 A14 S15 A16 A17 S18 D19 V20 S21 V22 E23 S24 T25 A26 E27 A28 T29 V30 C31 T32 V33 T34 L35 E36 K37 M38 S39 A40 G41 L42 G43 F44 S45 L46 E47 G48 G49 K50 G51 S52 L53 H54 G55 D56 K57 P58 L59 T60

I61 N62 R63 I64 F65 K66 G67 A68 A69 S70 E71 Q72 S73 E74 T75 V76 Q77 D80 E81 I82 L83 Q84 L85 G86 G87 T88 A89 M90 Q91 G92 L93 T94 R95 F96 E97 A98 W99 N100 I101 I102 K103 A104 L105 P106 V110 T111 I112 V113 I114 R115 R116 K117 S118 L119 Q120 S121 K122 E123

T124 T125 A126 A127 G128 D129 S130

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

The representative model is number 8. Colouring as in section 4.1 above.

- Molecule 1: INTERLEUKIN 16

Chain A: 9% 22% 18% • 49%

M1 P2 D3 L4 N5 S6 S7 T8 D9 S10 A11 A12 S13 A14 S15 A16 A17 S18 D19 V20 S21 V22 E23 S24 T25 A26 E27 A28 C31 T32 V33 T34 L35 E36 K37 M38 S39 A40 G41 L42 G43 F44 S45 L46 E47 G48 G49 K50 G51 S52 L53 H54 G55 D56 K57 P58 L59 T60 I61

N62 R63 I64 F65 K66 G67 A68 A69 S70 E71 Q72 S73 E74 T75 V76 Q77 P78 G79 D80 E81 I82 L83 Q84 L85 G86 G87 T88 A89 M90 Q91 G92 L93 T94 R95 F96 E97 I101 I102 K103 A104 L105 P106 D107 V110 T111 I112 V113 I114 R115 R116 K117 S118 L119 Q120 S121 K122 E123 T124

T125 A126 A127 G128 D129 S130

# 5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *SIMULATED ANNEALING*.

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *LEAST RESTRAINT VIOLATION, LEAST TOTAL ENERGY*.

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

| Software name | Classification | Version |
|---|---|---|
| X-PLOR | refinement | 3.1 |
| CCNMR | structure solution | |

No chemical shift data was provided. No validations of the models with respect to experimental NMR restraints is performed at this time.

# 6 Model quality (i)

## 6.1 Standard geometry (i)

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

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

| Mol | Chain | Chirality | Planarity |
|---|---|---|---|
| 1 | A | 0.0±0.0 | 3.9±0.4 |
| All | All | 0 | 77 |

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

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

| Mol | Chain | Res | Type | Group | Models (Total) |
|---|---|---|---|---|---|
| 1 | A | 116 | ARG | Sidechain | 20 |
| 1 | A | 115 | ARG | Sidechain | 19 |
| 1 | A | 95 | ARG | Sidechain | 19 |
| 1 | A | 63 | ARG | Sidechain | 19 |

## 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 | 508 | 531 | 531 | 111±9 |
| All | All | 10160 | 10620 | 10620 | 2221 |

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 107.

5 of 746 unique clashes are listed below, sorted by their clash magnitude.

| Atom-1 | Atom-2 | Clash(Å) | Distance(Å) | Models | |
|---|---|---|---|---|---|
| | | | | Worst | Total |
| 1:A:88:THR:HG21 | 1:A:102:ILE:HD11 | 1.12 | 1.22 | 11 | 18 |
| 1:A:88:THR:HG21 | 1:A:102:ILE:CD1 | 1.03 | 1.83 | 6 | 20 |
| 1:A:90:MET:CE | 1:A:98:ALA:HB1 | 0.97 | 1.90 | 4 | 2 |
| 1:A:42:LEU:HD12 | 1:A:44:PHE:CE2 | 0.96 | 1.95 | 20 | 7 |
| 1:A:99:TRP:CZ3 | 1:A:102:ILE:HG21 | 0.96 | 1.94 | 1 | 8 |

## 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 | 66/130 (51%) | 50±2 (76±4%) | 13±3 (20±4%) | 3±1 (5±2%) | 5 27 |
| All | All | 1320/2600 (51%) | 997 (76%) | 258 (20%) | 65 (5%) | 5 27 |

5 of 8 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 | 36 | GLU | 16 |
| 1 | A | 42 | LEU | 13 |
| 1 | A | 41 | GLY | 11 |
| 1 | A | 91 | GLN | 9 |
| 1 | A | 74 | GLU | 7 |

### 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 | 56/102 (55%) | 33±2 (58±4%) | 23±2 (42±4%) | 0 4 |
| All | All | 1120/2040 (55%) | 654 (58%) | 466 (42%) | 0 4 |

5 of 48 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 | 110 | VAL | 20 |
| 1 | A | 33 | VAL | 20 |
| 1 | A | 34 | THR | 20 |
| 1 | A | 42 | LEU | 20 |
| 1 | A | 94 | THR | 19 |

#### 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 carbohydrates 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)

No chemical shift data were provided