# Full wwPDB NMR Structure Validation Report (i)

Apr 26, 2016 – 08:37 PM BST

| | | |
|---|---|---|
| PDB ID | : | 2H7E |
| Title | : | Solution structure of the talin F3 domain in complex with a chimeric beta3 integrin-PIP kinase peptide- minimized average structure |
| Authors | : | Wegener, K.L. |
| Deposited on | : | 2006-06-02 |

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
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:

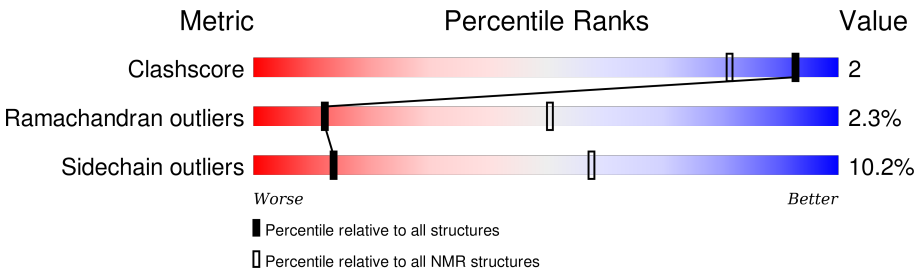
| | | |
|---|---|---|
| Cyrange | : | Kirchner and Güntert (2011) |
| NmrClust | : | Kelley et al. (1996) |
| MolProbity | : | 4.02b-467 |
| Mogul | : | 1.7.1 (RC1), CSD as537be (2016) |
| Percentile statistics | : | 20151230.v01 (using entries in the PDB archive December 30th 2015) |
| RCI | : | v_1n_11_5_13_A (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 is 69%.

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 | 101 | 89% 11% |
| 2 | B | 34 | 76% 18% 6% |

# 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 are 2 unique types of molecules in this entry. The entry contains 2255 atoms, of which 1145 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Talin-1.

| Mol | Chain | Residues | Atoms | | | | | | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | H | N | O | S | |
| 1 | A | 101 | 1650 | 522 | 842 | 133 | 151 | 2 | 0 |

There are 5 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| A | 305 | PRO | - | CLONING ARTIFACT | UNP P54939 |
| A | 306 | LEU | - | CLONING ARTIFACT | UNP P54939 |
| A | 307 | GLY | - | CLONING ARTIFACT | UNP P54939 |
| A | 308 | SER | - | CLONING ARTIFACT | UNP P54939 |
| A | 336 | SER | CYS | ENGINEERED | UNP P54939 |

- Molecule 2 is a protein called Chimera of 24-mer peptide from Integrin beta-3 and 10-mer peptide from Phosphatidylinositol-4-phosphate 5-kinase type-1 gamma.

| Mol | Chain | Residues | Atoms | | | | | | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | H | N | O | P | |
| 2 | B | 34 | 605 | 193 | 303 | 55 | 53 | 1 | 0 |

There is a discrepancy between the modelled and reference sequences:

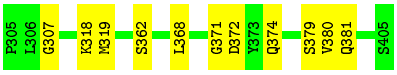
| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| B | 741 | PTR | TYR | MODIFIED RESIDUE | UNP O70161 |

# 4 Residue-property plots (i)

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: Talin-1

Chain A:




- Molecule 2: Chimera of 24-mer peptide from Integrin beta-3 and 10-mer peptide from Phosphatidylinositol-4-phosphate 5-kinase type-1 gamma

Chain B:

# 5 Refinement protocol and experimental data overview (i)

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

Of the ? calculated structures, 1 were deposited, based on the following criterion: *?*.

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

| Software name | Classification | Version |
|---|---|---|
| CNS | structure solution | 1.1 |
| CNS | refinement | 1.1 |

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) | BMRB entry 7150 |
|---|---|
| Number of chemical shift lists | 1 |
| Total number of shifts | 1380 |
| Number of shifts mapped to atoms | 1380 |
| Number of unparsed shifts | 0 |
| Number of shifts with mapping errors | 0 |
| Number of shifts with mapping warnings | 0 |
| Assignment completeness (well-defined parts) | 69% |

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)

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

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 | 808 | 842 | 842 | 1 |
| 2 | B | 302 | 303 | 297 | 3 |
| All | All | 1110 | 1145 | 1139 | 4 |

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 2.

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

| Atom-1 | Atom-2 | Clash(Å) | Distance(Å) |
|---|---|---|---|
| 2:B:730:PHE:CD1 | 2:B:731:GLU:N | 0.61 | 2.69 |
| 1:A:380:VAL:HG23 | 1:A:381:GLN:N | 0.48 | 2.24 |
| 2:B:719:ILE:HG21 | 2:B:723:ASP:CB | 0.45 | 2.42 |
| 2:B:719:ILE:HG21 | 2:B:723:ASP:HB3 | 0.43 | 1.89 |

## 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 | 99/101 (98%) | 93 (94%) | 4 (4%) | 2 (2%) | 14 | 55 |
| 2 | B | 31/34 (91%) | 27 (87%) | 3 (10%) | 1 (3%) | 8 | 40 |
| All | All | 130/135 (96%) | 120 (92%) | 7 (5%) | 3 (2%) | 12 | 51 |

All 3 Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

| Mol | Chain | Res | Type |
|---|---|---|---|
| 2 | B | 722 | HIS |
| 1 | A | 307 | GLY |
| 1 | A | 371 | GLY |

### 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 | 89/89 (100%) | 82 (92%) | 7 (8%) | 20 | 65 |
| 2 | B | 29/29 (100%) | 24 (83%) | 5 (17%) | 6 | 42 |
| All | All | 118/118 (100%) | 106 (90%) | 12 (10%) | 14 | 58 |

All 12 residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

| Mol | Chain | Res | Type |
|---|---|---|---|
| 2 | B | 749 | ARG |
| 1 | A | 319 | MET |
| 1 | A | 362 | SER |
| 1 | A | 372 | ASP |
| 1 | A | 379 | SER |
| 2 | B | 716 | LYS |
| 1 | A | 374 | GLN |
| 1 | A | 368 | LEU |
| 1 | A | 318 | LYS |
| 2 | B | 730 | PHE |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type |
|---|---|---|---|
| 2 | B | 722 | HIS |
| 2 | B | 734 | ARG |

### 6.3.3 RNA (i)

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue 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 | PTR | B | 741 | 2 | 13,16,17 | 0.66 | 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 |
| 2 | PTR | B | 741 | 2 | 19,22,24 | 0.89 | 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 |
|---|---|---|---|---|---|---|---|
| 2 | PTR | B | 741 | 2 | - | 0,9,11,13 | 0,1,1,1 |

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 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)

The completeness of assignment taking into account all chemical shift lists is 69% for the well-defined parts and 69% for the entire structure.

## 7.1 Chemical shift list 1

File name: BMRB entry 7150

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 | 1380 |
| Number of shifts mapped to atoms | 1380 |
| Number of unparsed shifts | 0 |
| Number of shifts with mapping errors | 0 |
| Number of shifts with mapping warnings | 0 |
| Number of shift outliers (ShiftChecker) | 5 |

### 7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

| Nucleus | # values | Correction ± precision, *ppm* | Suggested action |
|---|---|---|---|
| $^{13}C_{\alpha}$ | 100 | -0.26 ± 0.18 | None needed (< 0.5 ppm) |
| $^{13}C_{\beta}$ | 87 | -0.45 ± 0.16 | None needed (< 0.5 ppm) |
| $^{13}C'$ | 96 | -0.05 ± 0.11 | None needed (< 0.5 ppm) |
| $^{15}N$ | 97 | 0.05 ± 0.75 | None needed (< 0.5 ppm) |

### 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 69%, i.e. 1207 atoms were assigned a chemical shift out of a possible 1753. 0 out of 20 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

| | Total | $^{1}H$ | $^{13}C$ | $^{15}N$ |
|---|---|---|---|---|
| Backbone | 555/662 (84%) | 262/264 (99%) | 196/268 (73%) | 97/130 (75%) |
| Sidechain | 583/945 (62%) | 487/557 (87%) | 87/342 (25%) | 9/46 (20%) |

*Continued on next page...*

*Continued from previous page...*

| | Total | $^{1}$H | $^{13}$C | $^{15}$N |
|---|---|---|---|---|
| Aromatic | 69/146 (47%) | 67/76 (88%) | 0/63 (0%) | 2/7 (29%) |
| Overall | 1207/1753 (69%) | 816/897 (91%) | 283/673 (42%) | 108/183 (59%) |

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 69%, i.e. 1207 atoms were assigned a chemical shift out of a possible 1753. 0 out of 20 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

| | Total | $^{1}$H | $^{13}$C | $^{15}$N |
|---|---|---|---|---|
| Backbone | 555/662 (84%) | 262/264 (99%) | 196/268 (73%) | 97/130 (75%) |
| Sidechain | 583/945 (62%) | 487/557 (87%) | 87/342 (25%) | 9/46 (20%) |
| Aromatic | 69/146 (47%) | 67/76 (88%) | 0/63 (0%) | 2/7 (29%) |
| Overall | 1207/1753 (69%) | 816/897 (91%) | 283/673 (42%) | 108/183 (59%) |

### 7.1.4 Statistically unusual chemical shifts (i)

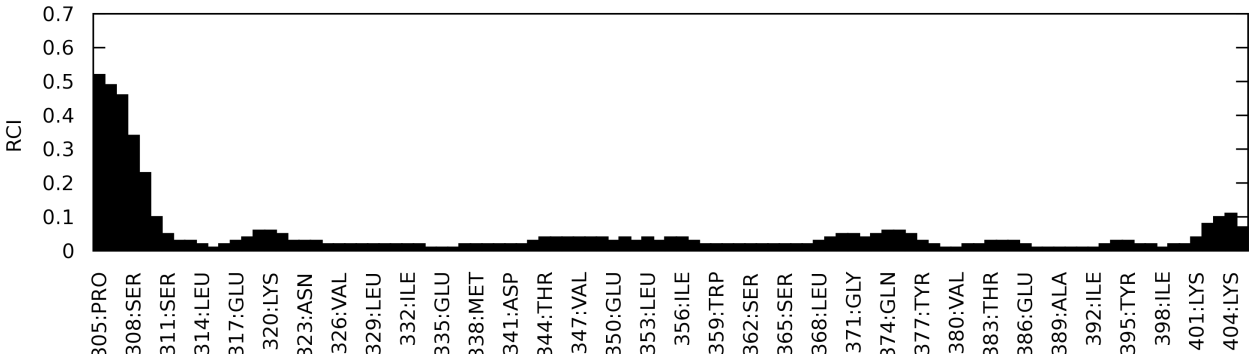
The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

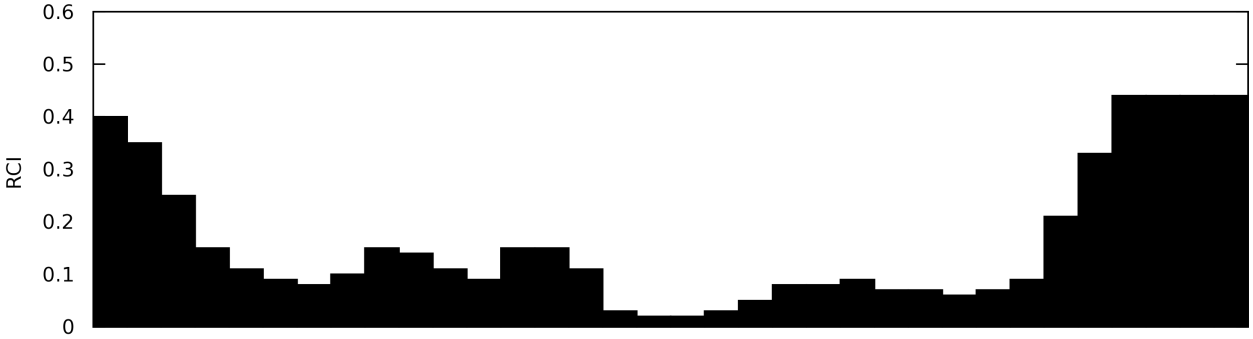
| Mol | Chain | Res | Type | Atom | Shift, *ppm* | Expected range, *ppm* | Z-score |
|---|---|---|---|---|---|---|---|
| 1 | A | 388 | ILE | HG12 | -1.29 | 3.27 – -0.73 | -6.4 |
| 1 | A | 360 | ALA | HB3 | -0.12 | 2.61 – 0.11 | -5.9 |
| 1 | A | 360 | ALA | HB2 | -0.12 | 2.61 – 0.11 | -5.9 |
| 1 | A | 360 | ALA | HB1 | -0.12 | 2.61 – 0.11 | -5.9 |
| 1 | A | 349 | GLN | HB2 | 0.64 | 3.30 – 0.80 | -5.6 |

### 7.1.5 Random Coil Index (RCI) plots (i)

The images below report *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.

Random coil index (RCI) for chain A:




Random coil index (RCI) for chain B: