# W O R L D W I D E

# wwPDB

## PROTEIN DATA BANK

# Full wwPDB NMR Structure Validation Report (i)

Apr 27, 2016 – 02:55 AM BST

| | | |
|---|---|---|
| PDB ID | : | 2MCN |
| Title | : | Distinct ubiquitin binding modes exhibited by SH3 domains: molecular deter- minants and functional implications |
| Authors | : | Ortega-Roldan, J.; Salmon, L.; Azuaga, A.; Blackledge, M.; Van Nuland, N. |
| Deposited on | : | 2013-08-22 |

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:

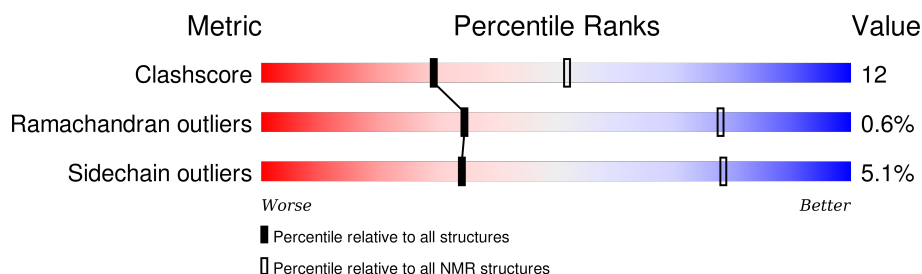| | | |
|---|---|---|
| Cyrange | : | Kirchner and Güntert (2011) |
| NmrClust | : | Kelley et al. (1996) |
| MolProbity | : | 4.02b-467 |
| Mogul | : | unknown |
| Percentile statistics | : | 20151230.v01 (using entries in the PDB archive December 30th 2015) |
| RCI | : | v_1n_11_5_13_A (Berjanski et 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 ⓘ

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

The overall completeness of chemical shifts assignment is 15%.

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 | Percentile Ranks | Value |
|--------|------------------|-------|
| Clashscore | | 12 |
| Ramachandran outliers | | 0.6% |
| Sidechain outliers | | 5.1% |

*Worse*                                                          *Better*

▌ Percentile relative to all structures

▯ Percentile relative to all NMR structures

| 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 | 61 | 56%    34%   • •   7% |
| 2 | B | 76 | 86%    7%    8% |

# 2 Ensemble composition and analysis (i)

This entry contains 10 models. Model 9 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

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:58, B:1-B:70 (126) | 0.30 | 9 |

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 and 1 single-model cluster was found.

| Cluster number | Models |
|---|---|
| 1 | 1, 2, 4, 7, 9, 10 |
| 2 | 3, 6, 8 |
| Single-model clusters | 5 |

# 3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2171 atoms, of which 1093 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called CD2-associated protein.

| Mol | Chain | Residues | Atoms | | | | | | Trace |
|-----|-------|----------|-------|---|---|---|---|---|-------|
| | | | Total | C | H | N | O | S | |
| 1 | A | 57 | 939 | 302 | 463 | 80 | 93 | 1 | 0 |

- Molecule 2 is a protein called Ubiquitin.

| Mol | Chain | Residues | Atoms | | | | | | Trace |
|-----|-------|----------|-------|---|---|---|---|---|-------|
| | | | Total | C | H | N | O | S | |
| 2 | B | 76 | 1232 | 378 | 630 | 105 | 118 | 1 | 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: CD2-associated protein
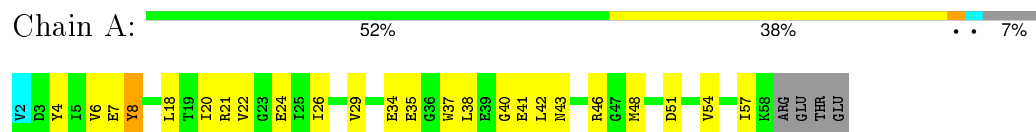
Chain A:



- Molecule 2: Ubiquitin

Chain B:
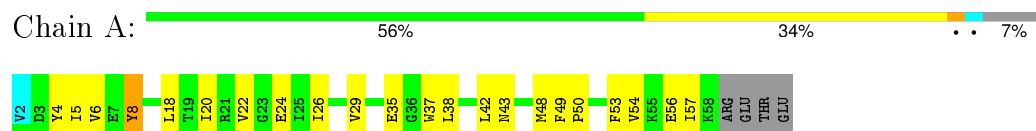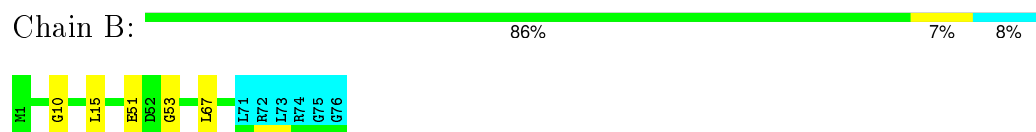


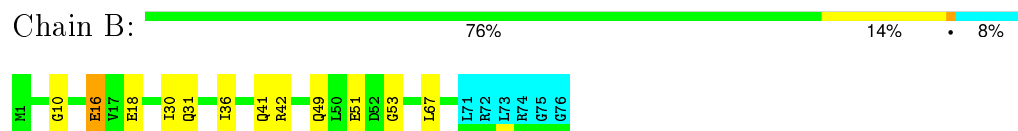## 4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

### 4.2.1 Score per residue for model 1

- Molecule 1: CD2-associated protein

Chain A:



- Molecule 2: Ubiquitin

Chain B:

#### 4.2.2 Score per residue for model 2

• Molecule 1: CD2-associated protein

Chain A:



• Molecule 2: Ubiquitin

Chain B:
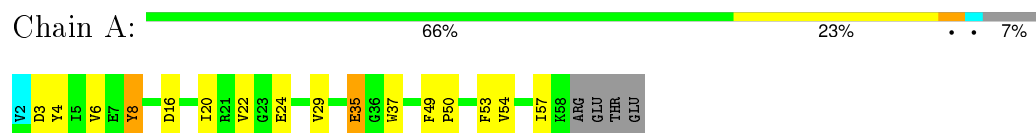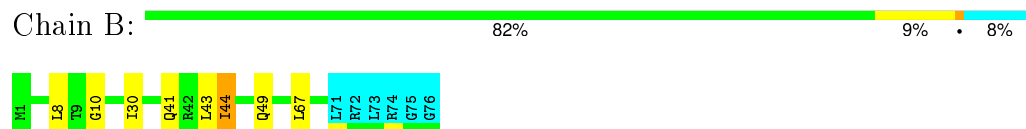


#### 4.2.3 Score per residue for model 3

• Molecule 1: CD2-associated protein

Chain A:



• Molecule 2: Ubiquitin

Chain B:
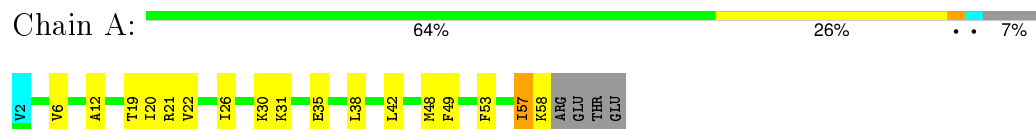


#### 4.2.4 Score per residue for model 4

• Molecule 1: CD2-associated protein

Chain A:



• Molecule 2: Ubiquitin

Chain B:
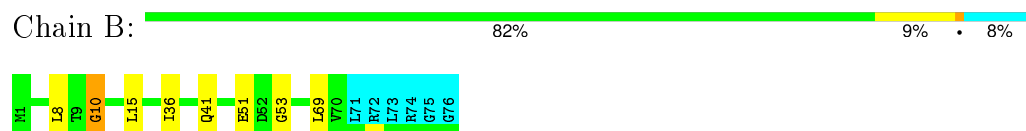
#### 4.2.5 Score per residue for model 5
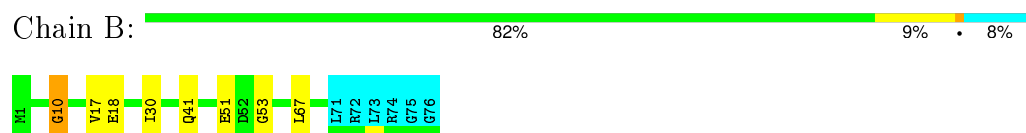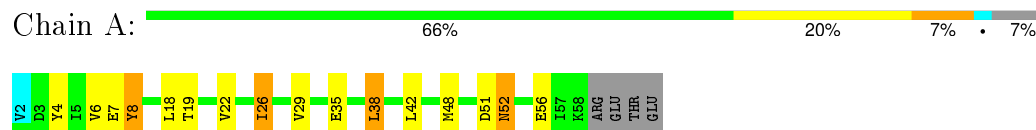
• Molecule 1: CD2-associated protein

Chain A:

62% 26% • • 7%

V2 D3 Y4 I5 V6 E7 Y8 D9 L18 T19 I20 R21 V22 I26 V29 L38 L42 N43 M48 F49 P50 F53 V54 K55 E56 I57 K58 ARG GLU THR GLU

• Molecule 2: Ubiquitin

Chain B:

80% 11% • 8%

M1 L15 E18 K29 I44 F45 E51 D52 G53 L67 H68 L71 R72 L73 R74 G75 G76
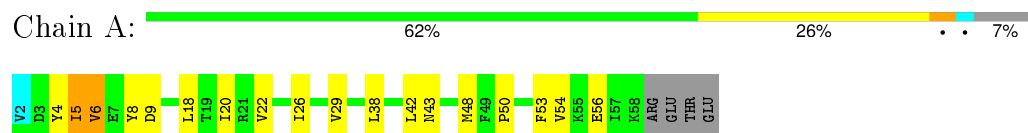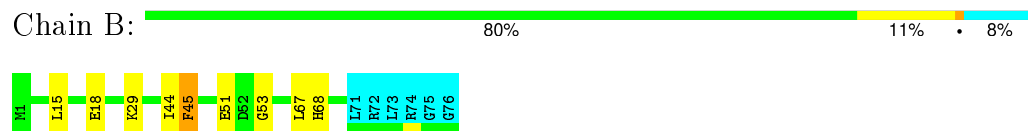
#### 4.2.6 Score per residue for model 6

• Molecule 1: CD2-associated protein

Chain A:

52% 36% • • 7%

V2 D3 Y4 I5 V6 E7 Y8 D9 Y10 E17 L18 T19 I20 R21 V22 G23 E24 V29 L38 E39 L42 N43 G44 R45 R46 G47 M48 F49 P50 F53 V54 K55 E56 I57 K58 ARG GLU THR GLU

• Molecule 2: Ubiquitin

Chain B:

79% 11% • 8%

M1 T7 L8 T9 G10 L43 I44 L50 E51 D52 G53 S65 T66 L67 L71 R72 L73 R74 G75 G76
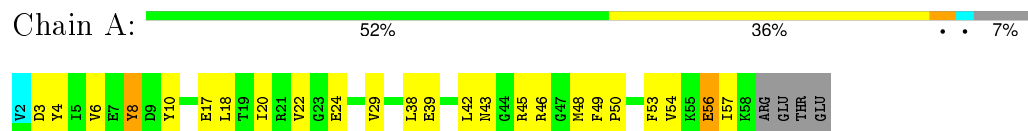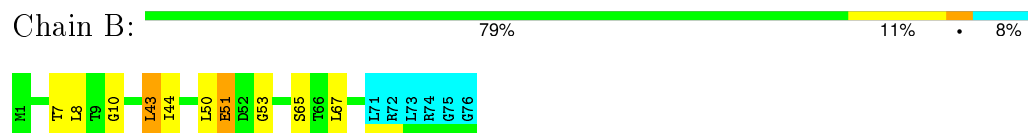
#### 4.2.7 Score per residue for model 7

• Molecule 1: CD2-associated protein

Chain A:

51% 38% • • 7%

V2 D3 Y4 I5 V6 E7 Y8 E17 L18 T19 I20 E24 I25 I26 R27 N28 V29 L32 Q33 E39 G40 E41 L42 R46 G47 M48 F49 P50 F53 V54 K55 E56 I57 K58 ARG GLU THR GLU

• Molecule 2: Ubiquitin

Chain B:

78% 14% 8%

M1 G10 I13 T14 L15 I30 Q31 I36 L50 E51 D52 G53 R54 T55 L56 S57 L71 R72 L73 R74 G75 G76
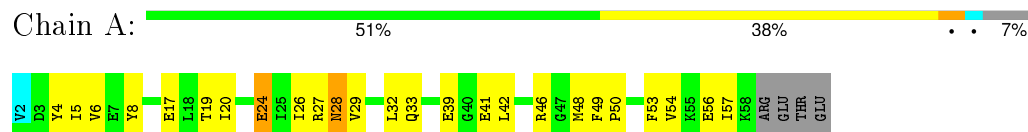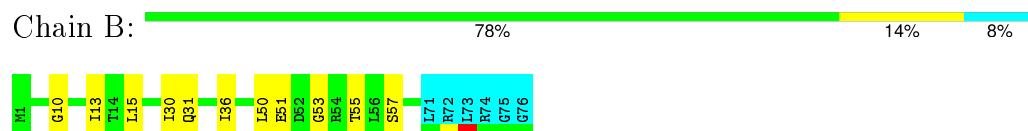
### 4.2.8 Score per residue for model 8

• Molecule 1: CD2-associated protein

Chain A:

59%    26%    7%    •    7%

V2 D3 Y4 I5 V6 E7 Y8 D15 L18 T19 I20 R21 V22 I25 I26 R27 N28 V29 Q33 E34 E35 G36 W37 L42 N43 F49 P50 F53 V54 K58 ARG GLU THR GLU

• Molecule 2: Ubiquitin

Chain B:

70%    21%    •    8%

M1 Q2 I3 G10 T14 L15 E16 V17 E18 I23 K29 K48 Q49 L50 E51 D52 G53 E64 H68 L69 V70 L71 R72 L73 R74 G75 G76
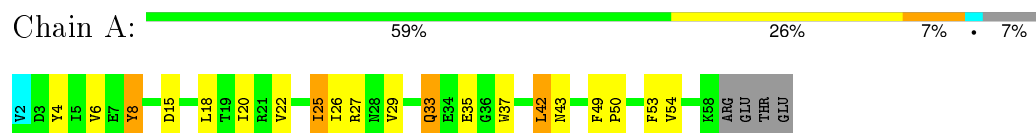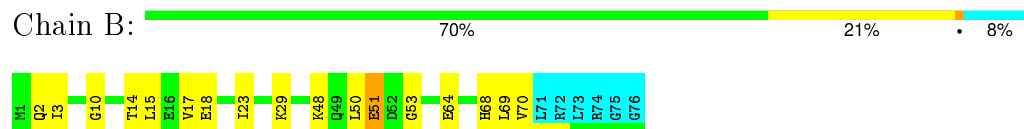
### 4.2.9 Score per residue for model 9 (medoid)

• Molecule 1: CD2-associated protein

Chain A:

66%    23%    •  •    7%

V2 D3 Y4 I5 V6 E7 Y8 L18 T19 I20 R21 V22 G23 E24 I25 I26 V29 E35 L38 M48 F49 P50 F53 V54 K58 ARG GLU THR GLU

• Molecule 2: Ubiquitin

Chain B:

80%    11%    •    8%

M1 Q2 I3 L8 T9 G10 L15 E16 V26 E51 D52 G53 K63 L71 R72 L73 R74 G75 G76
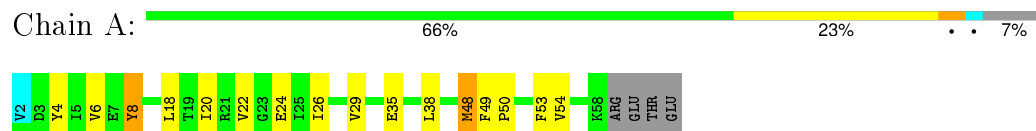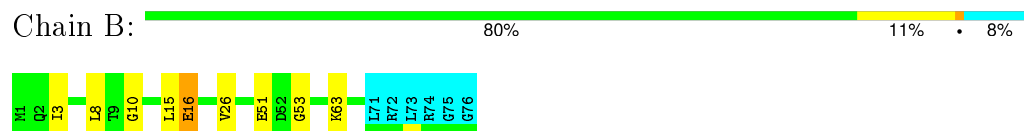
### 4.2.10 Score per residue for model 10

• Molecule 1: CD2-associated protein

Chain A:

57%    28%    7%    •    7%

V2 D3 Y4 I5 V6 E7 Y8 H14 D15 L18 T19 I20 R21 V22 N28 V29 Q33 W37 L38 L42 N43 M48 F49 P50 F53 V54 K58 ARG GLU THR GLU

• Molecule 2: Ubiquitin

Chain B:

84%    8%    8%

M1 Q2 I3 F4 Q31 L43 L67 V70 L71 R72 L73 R74 G75 G76
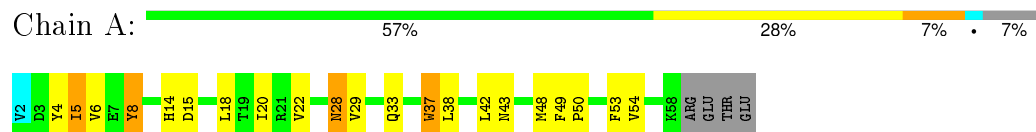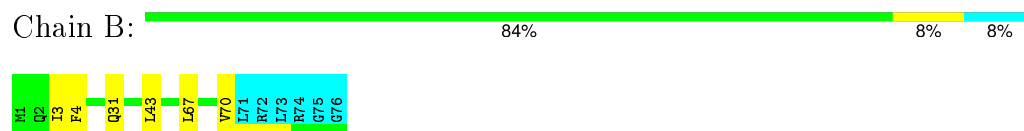
# 5 Refinement protocol and experimental data overview ⓘ

The models were refined using the following method: *simulated annealing, molecular dynamics, docking, simulated annealing, molecular dynamics.*

Of the 200 calculated structures, 10 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 |
|---|---|---|
| HADDOCK | refinement | |
| CNS/SCULPTOR | structure solution | |
| NMRVIEW | structure solution | |
| HADDOCK | structure solution | |
| CNS/SCULPTOR | refinement | |
| 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) | 2mcn_cs.str |
|---|---|
| Number of chemical shift lists | 2 |
| Total number of shifts | 264 |
| Number of shifts mapped to atoms | 264 |
| Number of unparsed shifts | 0 |
| Number of shifts with mapping errors | 0 |
| Number of shifts with mapping warnings | 0 |
| Assignment completeness (well-defined parts) | 15% |

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

# 6 Model quality ⓘ

## 6.1 Standard geometry ⓘ

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 | 469 | 452 | 451 | 18±3 |
| 2 | B | 555 | 576 | 575 | 8±3 |
| All | All | 10240 | 10280 | 10260 | 251 |

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 12.

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

| Atom-1 | Atom-2 | Clash(Å) | Distance(Å) | Models Worst | Models Total |
|--------|--------|----------|-------------|--------------|--------------|
| 1:A:6:VAL:HG12 | 1:A:54:VAL:HG12 | 0.79 | 1.53 | 7 | 2 |
| 1:A:6:VAL:HG22 | 1:A:54:VAL:HG12 | 0.75 | 1.59 | 10 | 3 |
| 1:A:6:VAL:HG21 | 1:A:20:ILE:HB | 0.70 | 1.63 | 1 | 4 |
| 1:A:38:LEU:HD12 | 1:A:54:VAL:HG11 | 0.69 | 1.64 | 1 | 1 |
| 1:A:6:VAL:CG1 | 1:A:54:VAL:HG12 | 0.66 | 2.21 | 7 | 2 |
| 1:A:26:ILE:N | 1:A:26:ILE:HD12 | 0.64 | 2.08 | 1 | 1 |
| 2:B:3:ILE:O | 2:B:3:ILE:HD12 | 0.64 | 1.93 | 9 | 1 |
| 1:A:6:VAL:HG22 | 1:A:24:GLU:O | 0.63 | 1.94 | 2 | 3 |
| 1:A:6:VAL:O | 1:A:6:VAL:HG23 | 0.62 | 1.93 | 7 | 1 |
| 2:B:43:LEU:HD12 | 2:B:50:LEU:HD22 | 0.62 | 1.69 | 6 | 1 |
| 1:A:20:ILE:HD12 | 1:A:20:ILE:O | 0.62 | 1.93 | 8 | 1 |
| 1:A:6:VAL:HG23 | 1:A:6:VAL:O | 0.61 | 1.95 | 2 | 2 |
| 1:A:8:TYR:CE2 | 1:A:53:PHE:CZ | 0.61 | 2.88 | 5 | 1 |

*Continued on next page...*

*Continued from previous page...*

| Atom-1 | Atom-2 | Clash(Å) | Distance(Å) | Models Worst | Total |
|---|---|---|---|---|---|
| 1:A:48:MET:N | 1:A:48:MET:SD | 0.61 | 2.73 | 9 | 2 |
| 2:B:44:ILE:HG22 | 2:B:45:PHE:N | 0.61 | 2.09 | 5 | 1 |
| 2:B:15:LEU:HD21 | 2:B:29:LYS:CB | 0.60 | 2.27 | 8 | 2 |
| 1:A:57:ILE:HD13 | 1:A:58:LYS:N | 0.60 | 2.12 | 3 | 1 |
| 1:A:6:VAL:HG11 | 1:A:20:ILE:HD12 | 0.60 | 1.73 | 3 | 3 |
| 1:A:12:ALA:H | 1:A:19:THR:HG22 | 0.59 | 1.56 | 3 | 1 |
| 1:A:6:VAL:CG2 | 1:A:54:VAL:HG12 | 0.58 | 2.28 | 10 | 2 |
| 1:A:42:LEU:HD23 | 1:A:43:ASN:N | 0.58 | 2.14 | 6 | 4 |
| 1:A:26:ILE:CD1 | 1:A:42:LEU:HD22 | 0.58 | 2.28 | 7 | 2 |
| 2:B:15:LEU:HD21 | 2:B:29:LYS:HB2 | 0.58 | 1.75 | 8 | 2 |
| 1:A:6:VAL:HG12 | 1:A:54:VAL:CG1 | 0.58 | 2.29 | 7 | 2 |
| 1:A:8:TYR:CE2 | 1:A:53:PHE:CE1 | 0.58 | 2.92 | 2 | 1 |
| 2:B:8:LEU:H | 2:B:8:LEU:HD12 | 0.57 | 1.58 | 9 | 1 |
| 2:B:44:ILE:HD13 | 2:B:44:ILE:C | 0.57 | 2.19 | 2 | 1 |
| 2:B:30:ILE:HD12 | 2:B:41:GLN:OE1 | 0.57 | 2.00 | 2 | 3 |
| 1:A:25:ILE:H | 1:A:25:ILE:HD13 | 0.56 | 1.58 | 8 | 1 |
| 1:A:49:PHE:CD1 | 1:A:54:VAL:HG11 | 0.56 | 2.36 | 9 | 5 |
| 1:A:6:VAL:CG1 | 1:A:20:ILE:HD13 | 0.56 | 2.31 | 8 | 1 |
| 1:A:4:TYR:HB2 | 1:A:29:VAL:HG21 | 0.56 | 1.77 | 6 | 6 |
| 1:A:20:ILE:HG21 | 1:A:26:ILE:HD11 | 0.55 | 1.78 | 5 | 1 |
| 1:A:8:TYR:CZ | 1:A:53:PHE:CE1 | 0.55 | 2.95 | 2 | 1 |
| 2:B:51:GLU:C | 2:B:53:GLY:N | 0.54 | 2.60 | 8 | 8 |
| 1:A:6:VAL:HG21 | 1:A:20:ILE:HD12 | 0.54 | 1.78 | 6 | 1 |
| 1:A:6:VAL:HG13 | 1:A:53:PHE:O | 0.54 | 2.01 | 10 | 3 |
| 1:A:19:THR:O | 1:A:42:LEU:HD23 | 0.54 | 2.02 | 3 | 2 |
| 1:A:50:PRO:CB | 2:B:68:HIS:CE1 | 0.53 | 2.90 | 5 | 1 |
| 1:A:50:PRO:HB2 | 1:A:53:PHE:CD1 | 0.53 | 2.38 | 6 | 4 |
| 1:A:6:VAL:HB | 1:A:54:VAL:HG12 | 0.53 | 1.79 | 5 | 1 |
| 1:A:6:VAL:HG12 | 1:A:20:ILE:HD13 | 0.52 | 1.81 | 8 | 1 |
| 1:A:26:ILE:O | 1:A:29:VAL:HG23 | 0.52 | 2.05 | 8 | 1 |
| 1:A:5:ILE:HD13 | 1:A:5:ILE:C | 0.52 | 2.25 | 5 | 2 |
| 1:A:3:ASP:O | 1:A:57:ILE:HG22 | 0.52 | 2.03 | 6 | 2 |
| 1:A:6:VAL:CG2 | 1:A:6:VAL:O | 0.52 | 2.57 | 7 | 1 |
| 1:A:8:TYR:O | 1:A:22:VAL:N | 0.52 | 2.43 | 4 | 6 |
| 1:A:35:GLU:CB | 2:B:10:GLY:HA2 | 0.51 | 2.35 | 1 | 6 |
| 1:A:53:PHE:CZ | 2:B:68:HIS:CE1 | 0.51 | 2.98 | 5 | 1 |
| 1:A:32:LEU:HD11 | 1:A:48:MET:SD | 0.51 | 2.45 | 7 | 1 |
| 2:B:23:ILE:HD12 | 2:B:51:GLU:O | 0.51 | 2.05 | 8 | 1 |
| 2:B:23:ILE:HD13 | 2:B:50:LEU:HB3 | 0.51 | 1.82 | 8 | 1 |
| 1:A:42:LEU:HD13 | 1:A:43:ASN:N | 0.51 | 2.21 | 5 | 1 |
| 1:A:27:ARG:O | 1:A:28:ASN:C | 0.51 | 2.48 | 7 | 1 |

*Continued on next page...*

*Continued from previous page...*

| Atom-1 | Atom-2 | Clash(Å) | Distance(Å) | Models Worst | Total |
|--------|--------|----------|-------------|--------------|-------|
| 2:B:43:LEU:CD1 | 2:B:50:LEU:HD22 | 0.50 | 2.37 | 6 | 1 |
| 1:A:8:TYR:CG | 1:A:53:PHE:CD2 | 0.49 | 2.99 | 6 | 1 |
| 2:B:8:LEU:HD23 | 2:B:8:LEU:N | 0.49 | 2.22 | 4 | 1 |
| 2:B:44:ILE:CG2 | 2:B:45:PHE:N | 0.49 | 2.75 | 5 | 1 |
| 1:A:51:ASP:HA | 1:A:54:VAL:HG12 | 0.49 | 1.85 | 1 | 1 |
| 1:A:25:ILE:H | 1:A:25:ILE:CD1 | 0.49 | 2.20 | 8 | 1 |
| 1:A:26:ILE:CD1 | 1:A:26:ILE:N | 0.49 | 2.75 | 1 | 1 |
| 2:B:30:ILE:HG22 | 2:B:36:ILE:HG22 | 0.48 | 1.84 | 7 | 1 |
| 1:A:6:VAL:HG22 | 1:A:6:VAL:O | 0.48 | 2.08 | 5 | 1 |
| 2:B:51:GLU:C | 2:B:53:GLY:H | 0.48 | 2.11 | 7 | 7 |
| 1:A:18:LEU:HD21 | 1:A:42:LEU:CB | 0.48 | 2.39 | 8 | 3 |
| 1:A:39:GLU:HB2 | 1:A:48:MET:SD | 0.48 | 2.49 | 6 | 1 |
| 1:A:18:LEU:HD13 | 1:A:19:THR:N | 0.48 | 2.24 | 4 | 1 |
| 2:B:67:LEU:HD12 | 2:B:67:LEU:N | 0.47 | 2.24 | 3 | 3 |
| 2:B:44:ILE:HG22 | 2:B:45:PHE:H | 0.47 | 1.65 | 5 | 1 |
| 1:A:4:TYR:CD1 | 1:A:29:VAL:HB | 0.47 | 2.44 | 10 | 4 |
| 2:B:43:LEU:HD11 | 2:B:67:LEU:HB3 | 0.47 | 1.86 | 6 | 1 |
| 1:A:4:TYR:CD1 | 1:A:56:GLU:HA | 0.47 | 2.45 | 4 | 1 |
| 1:A:20:ILE:HD13 | 1:A:49:PHE:HB2 | 0.47 | 1.85 | 3 | 1 |
| 2:B:44:ILE:N | 2:B:44:ILE:HD12 | 0.46 | 2.24 | 5 | 1 |
| 1:A:42:LEU:HD23 | 1:A:42:LEU:C | 0.46 | 2.30 | 1 | 1 |
| 2:B:67:LEU:N | 2:B:67:LEU:HD12 | 0.46 | 2.26 | 2 | 2 |
| 1:A:8:TYR:CD2 | 1:A:9:ASP:N | 0.46 | 2.84 | 5 | 1 |
| 2:B:43:LEU:N | 2:B:49:GLN:NE2 | 0.46 | 2.63 | 2 | 1 |
| 1:A:4:TYR:CD1 | 1:A:29:VAL:HG21 | 0.46 | 2.46 | 4 | 1 |
| 1:A:4:TYR:CE1 | 1:A:29:VAL:HG21 | 0.45 | 2.46 | 4 | 1 |
| 1:A:48:MET:SD | 1:A:48:MET:N | 0.45 | 2.88 | 3 | 1 |
| 1:A:31:LYS:HA | 1:A:38:LEU:HD23 | 0.45 | 1.89 | 3 | 1 |
| 1:A:49:PHE:CD1 | 1:A:54:VAL:HG21 | 0.45 | 2.46 | 6 | 1 |
| 1:A:4:TYR:CZ | 1:A:29:VAL:HG21 | 0.45 | 2.47 | 4 | 1 |
| 1:A:20:ILE:HD12 | 1:A:20:ILE:C | 0.45 | 2.32 | 8 | 1 |
| 1:A:6:VAL:HG22 | 1:A:54:VAL:CG1 | 0.45 | 2.38 | 10 | 1 |
| 2:B:44:ILE:H | 2:B:44:ILE:HD12 | 0.45 | 1.72 | 5 | 1 |
| 1:A:33:GLN:H | 1:A:33:GLN:NE2 | 0.45 | 2.10 | 8 | 1 |
| 1:A:42:LEU:C | 1:A:42:LEU:HD23 | 0.45 | 2.32 | 6 | 2 |
| 1:A:6:VAL:CG2 | 1:A:20:ILE:HD12 | 0.45 | 2.42 | 6 | 1 |
| 1:A:6:VAL:O | 1:A:6:VAL:CG2 | 0.45 | 2.63 | 2 | 1 |
| 1:A:4:TYR:CD2 | 1:A:56:GLU:HA | 0.44 | 2.47 | 6 | 1 |
| 2:B:18:GLU:CD | 2:B:18:GLU:N | 0.44 | 2.71 | 1 | 2 |
| 2:B:30:ILE:HG22 | 2:B:36:ILE:CG2 | 0.44 | 2.42 | 7 | 1 |
| 1:A:33:GLN:N | 1:A:33:GLN:NE2 | 0.44 | 2.65 | 7 | 1 |

*Continued on next page...*

*Continued from previous page...*

| Atom-1 | Atom-2 | Clash(Å) | Distance(Å) | Models | |
|---|---|---|---|---|---|
| | | | | Worst | Total |
| 2:B:43:LEU:HD12 | 2:B:43:LEU:N | 0.44 | 2.28 | 10 | 1 |
| 1:A:4:TYR:CE2 | 1:A:29:VAL:HG21 | 0.44 | 2.48 | 4 | 1 |
| 1:A:37:TRP:C | 1:A:38:LEU:HD12 | 0.44 | 2.32 | 10 | 1 |
| 2:B:51:GLU:O | 2:B:53:GLY:N | 0.44 | 2.50 | 8 | 1 |
| 1:A:52:ASN:HD22 | 1:A:52:ASN:N | 0.44 | 2.11 | 4 | 1 |
| 1:A:50:PRO:HG2 | 1:A:53:PHE:CD2 | 0.44 | 2.48 | 5 | 1 |
| 2:B:17:VAL:HG22 | 2:B:18:GLU:N | 0.44 | 2.27 | 8 | 2 |
| 2:B:8:LEU:HD23 | 2:B:8:LEU:H | 0.44 | 1.71 | 4 | 1 |
| 1:A:57:ILE:HG23 | 1:A:57:ILE:O | 0.44 | 2.13 | 1 | 1 |
| 2:B:50:LEU:HD12 | 2:B:50:LEU:N | 0.44 | 2.28 | 7 | 1 |
| 1:A:38:LEU:HB2 | 1:A:49:PHE:CE1 | 0.43 | 2.48 | 6 | 1 |
| 1:A:50:PRO:HG2 | 1:A:53:PHE:CE2 | 0.43 | 2.47 | 2 | 1 |
| 1:A:4:TYR:CD2 | 1:A:29:VAL:HG11 | 0.43 | 2.48 | 1 | 1 |
| 2:B:44:ILE:CG2 | 2:B:45:PHE:H | 0.43 | 2.26 | 5 | 1 |
| 1:A:5:ILE:HG23 | 1:A:57:ILE:HG21 | 0.43 | 1.88 | 7 | 1 |
| 2:B:68:HIS:CE1 | 2:B:70:VAL:CG2 | 0.43 | 3.01 | 8 | 1 |
| 1:A:24:GLU:HG2 | 1:A:42:LEU:HD11 | 0.43 | 1.90 | 6 | 1 |
| 1:A:35:GLU:CG | 2:B:10:GLY:HA2 | 0.43 | 2.44 | 2 | 1 |
| 1:A:26:ILE:CG2 | 1:A:40:GLY:HA3 | 0.43 | 2.42 | 1 | 1 |
| 1:A:8:TYR:N | 1:A:22:VAL:HG12 | 0.43 | 2.28 | 8 | 1 |
| 2:B:69:LEU:HD13 | 2:B:69:LEU:C | 0.43 | 2.32 | 4 | 1 |
| 2:B:31:GLN:HE21 | 2:B:36:ILE:C | 0.43 | 2.17 | 1 | 1 |
| 1:A:14:HIS:CG | 1:A:15:ASP:N | 0.43 | 2.87 | 10 | 1 |
| 1:A:4:TYR:CG | 1:A:29:VAL:HG21 | 0.43 | 2.48 | 4 | 1 |
| 2:B:15:LEU:N | 2:B:15:LEU:HD12 | 0.43 | 2.29 | 4 | 2 |
| 1:A:41:GLU:CD | 1:A:41:GLU:N | 0.43 | 2.72 | 7 | 1 |
| 1:A:7:GLU:O | 1:A:8:TYR:CD2 | 0.43 | 2.72 | 1 | 2 |
| 1:A:10:TYR:HB3 | 1:A:20:ILE:CG1 | 0.43 | 2.44 | 6 | 1 |
| 2:B:8:LEU:HD12 | 2:B:8:LEU:C | 0.43 | 2.34 | 2 | 1 |
| 1:A:26:ILE:HB | 1:A:29:VAL:CG2 | 0.43 | 2.44 | 8 | 1 |
| 1:A:4:TYR:CD2 | 1:A:29:VAL:HG21 | 0.43 | 2.48 | 4 | 1 |
| 2:B:16:GLU:CD | 2:B:16:GLU:H | 0.43 | 2.17 | 1 | 2 |
| 1:A:39:GLU:HG2 | 1:A:48:MET:SD | 0.43 | 2.54 | 7 | 1 |
| 2:B:55:THR:C | 2:B:57:SER:H | 0.43 | 2.17 | 7 | 1 |
| 1:A:37:TRP:CD1 | 1:A:48:MET:HB3 | 0.43 | 2.49 | 10 | 1 |
| 1:A:4:TYR:CZ | 1:A:56:GLU:HB2 | 0.43 | 2.49 | 7 | 1 |
| 1:A:28:ASN:HD22 | 1:A:28:ASN:N | 0.43 | 2.12 | 10 | 1 |
| 1:A:8:TYR:CD1 | 1:A:53:PHE:CE2 | 0.42 | 3.07 | 9 | 1 |
| 1:A:26:ILE:HD13 | 1:A:26:ILE:H | 0.42 | 1.74 | 4 | 1 |
| 1:A:38:LEU:O | 1:A:48:MET:HA | 0.42 | 2.14 | 4 | 3 |
| 2:B:36:ILE:O | 2:B:36:ILE:HG23 | 0.42 | 2.14 | 7 | 1 |

*Continued on next page...*

*Continued from previous page...*

| Atom-1 | Atom-2 | Clash(Å) | Distance(Å) | Models | |
|--------|--------|----------|-------------|--------|--------|
| | | | | Worst | Total |
| 2:B:15:LEU:HD23 | 2:B:15:LEU:H | 0.42 | 1.75 | 7 | 1 |
| 2:B:43:LEU:HD13 | 2:B:44:ILE:N | 0.42 | 2.28 | 6 | 1 |
| 2:B:7:THR:O | 2:B:8:LEU:C | 0.42 | 2.58 | 6 | 1 |
| 1:A:5:ILE:HG23 | 1:A:5:ILE:O | 0.42 | 2.14 | 10 | 1 |
| 1:A:50:PRO:HB2 | 1:A:53:PHE:CD2 | 0.42 | 2.49 | 7 | 1 |
| 1:A:8:TYR:CD1 | 1:A:53:PHE:CD2 | 0.42 | 3.07 | 9 | 1 |
| 1:A:6:VAL:HG23 | 1:A:21:ARG:O | 0.42 | 2.15 | 1 | 1 |
| 1:A:52:ASN:HD22 | 1:A:52:ASN:H | 0.42 | 1.57 | 4 | 1 |
| 1:A:6:VAL:CG2 | 1:A:20:ILE:HB | 0.42 | 2.42 | 1 | 2 |
| 1:A:38:LEU:HD12 | 1:A:54:VAL:HG21 | 0.42 | 1.91 | 5 | 1 |
| 1:A:37:TRP:CE3 | 1:A:48:MET:SD | 0.41 | 3.13 | 1 | 1 |
| 1:A:25:ILE:HD13 | 1:A:25:ILE:N | 0.41 | 2.29 | 8 | 1 |
| 1:A:6:VAL:HG21 | 1:A:20:ILE:CB | 0.41 | 2.45 | 6 | 1 |
| 1:A:6:VAL:O | 1:A:6:VAL:HG13 | 0.41 | 2.15 | 6 | 1 |
| 1:A:8:TYR:N | 1:A:22:VAL:HG22 | 0.41 | 2.30 | 9 | 1 |
| 1:A:29:VAL:HA | 1:A:39:GLU:O | 0.41 | 2.15 | 7 | 1 |
| 1:A:4:TYR:CE1 | 1:A:29:VAL:CG2 | 0.41 | 3.03 | 7 | 1 |
| 1:A:45:ARG:NE | 1:A:45:ARG:HA | 0.41 | 2.31 | 6 | 1 |
| 1:A:21:ARG:HG2 | 1:A:22:VAL:N | 0.41 | 2.30 | 3 | 1 |
| 2:B:3:ILE:O | 2:B:14:THR:HG23 | 0.41 | 2.16 | 8 | 1 |
| 1:A:5:ILE:HD13 | 1:A:6:VAL:N | 0.41 | 2.31 | 5 | 1 |
| 1:A:19:THR:O | 1:A:19:THR:HG23 | 0.41 | 2.16 | 7 | 1 |
| 1:A:50:PRO:HB3 | 2:B:68:HIS:CE1 | 0.41 | 2.50 | 5 | 1 |
| 2:B:69:LEU:C | 2:B:69:LEU:HD23 | 0.41 | 2.36 | 8 | 1 |
| 2:B:36:ILE:O | 2:B:41:GLN:NE2 | 0.41 | 2.53 | 4 | 1 |
| 2:B:42:ARG:HH22 | 2:B:49:GLN:HE21 | 0.41 | 1.58 | 1 | 1 |
| 1:A:14:HIS:CG | 1:A:15:ASP:H | 0.41 | 2.34 | 10 | 1 |
| 1:A:41:GLU:N | 1:A:41:GLU:CD | 0.40 | 2.74 | 1 | 1 |
| 1:A:4:TYR:CD1 | 1:A:29:VAL:HG22 | 0.40 | 2.51 | 7 | 1 |
| 1:A:4:TYR:CD1 | 1:A:29:VAL:CG2 | 0.40 | 3.04 | 7 | 1 |
| 1:A:26:ILE:N | 1:A:26:ILE:HD13 | 0.40 | 2.31 | 4 | 1 |
| 1:A:26:ILE:HD13 | 1:A:49:PHE:CE2 | 0.40 | 2.51 | 9 | 1 |
| 1:A:6:VAL:HG12 | 1:A:54:VAL:CB | 0.40 | 2.46 | 1 | 1 |
| 1:A:34:GLU:HB3 | 1:A:37:TRP:CD1 | 0.40 | 2.51 | 1 | 1 |
| 1:A:8:TYR:CE2 | 1:A:53:PHE:CE2 | 0.40 | 3.09 | 5 | 1 |

## 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 | 55/61 (90%) | 52±1 (95±2%) | 3±1 (5±2%) | 0±0 (0±1%) | 56 | 85 |
| 2 | B | 69/76 (91%) | 62±2 (90±3%) | 7±2 (9±3%) | 1±1 (1±1%) | 24 | 71 |
| All | All | 1240/1370 (91%) | 1142 (92%) | 90 (7%) | 8 (1%) | 34 | 78 |

All 5 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

| Mol | Chain | Res | Type | Models (Total) |
|------|-------|-----|------|----------------|
| 2 | B | 10 | GLY | 4 |
| 1 | A | 28 | ASN | 1 |
| 2 | B | 65 | SER | 1 |
| 2 | B | 13 | ILE | 1 |
| 2 | B | 45 | PHE | 1 |

### 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 | 50/55 (91%) | 46±1 (91±3%) | 4±1 (9±3%) | 18 | 63 |
| 2 | B | 64/68 (94%) | 63±1 (98±2%) | 2±1 (2±2%) | 61 | 93 |
| All | All | 1140/1230 (93%) | 1082 (95%) | 58 (5%) | 34 | 78 |

All 35 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 | 8 | TYR | 8 |
| 1 | A | 18 | LEU | 3 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | Models (Total) |
|-----|-------|-----|------|----------------|
| 1 | A | 46 | ARG | 3 |
| 1 | A | 37 | TRP | 3 |
| 1 | A | 24 | GLU | 2 |
| 1 | A | 6 | VAL | 2 |
| 2 | B | 16 | GLU | 2 |
| 1 | A | 33 | GLN | 2 |
| 1 | A | 5 | ILE | 2 |
| 2 | B | 51 | GLU | 2 |
| 2 | B | 31 | GLN | 2 |
| 1 | A | 56 | GLU | 2 |
| 1 | A | 17 | GLU | 2 |
| 1 | A | 42 | LEU | 2 |
| 1 | A | 57 | ILE | 1 |
| 1 | A | 26 | ILE | 1 |
| 2 | B | 26 | VAL | 1 |
| 1 | A | 16 | ASP | 1 |
| 1 | A | 35 | GLU | 1 |
| 2 | B | 43 | LEU | 1 |
| 1 | A | 52 | ASN | 1 |
| 1 | A | 51 | ASP | 1 |
| 2 | B | 63 | LYS | 1 |
| 1 | A | 38 | LEU | 1 |
| 2 | B | 2 | GLN | 1 |
| 2 | B | 3 | ILE | 1 |
| 1 | A | 27 | ARG | 1 |
| 2 | B | 44 | ILE | 1 |
| 1 | A | 48 | MET | 1 |
| 1 | A | 25 | ILE | 1 |
| 1 | A | 28 | ASN | 1 |
| 1 | A | 30 | LYS | 1 |
| 2 | B | 64 | GLU | 1 |
| 2 | B | 4 | PHE | 1 |
| 2 | B | 48 | LYS | 1 |

### 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)

The completeness of assignment taking into account all chemical shift lists is 15% for the well-defined parts and 15% for the entire structure.

## 7.1 Chemical shift list 1

File name: 2mcn_cs.str

Chemical shift list name: *CD2AP_SH3-A*

### 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 | 120 |
| Number of shifts mapped to atoms | 120 |
| 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)

The following table shows the suggested chemical shift referencing corrections.

| Nucleus | # values | Correction ± precision, *ppm* | Suggested action |
|---|---|---|---|
| $^{13}\mathrm{C}_\alpha$ | 0 | — | — |
| $^{13}\mathrm{C}_\beta$ | 0 | — | — |
| $^{13}\mathrm{C}'$ | 0 | — | — |
| $^{15}\mathrm{N}$ | 60 | $0.44 \pm 0.39$ | 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 7%, i.e. 110 atoms were assigned a chemical shift out of a possible 1609. 0 out of 20 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

| | Total | $^{1}\mathrm{H}$ | $^{13}\mathrm{C}$ | $^{15}\mathrm{N}$ |
|---|---|---|---|---|
| Backbone | 110/622 (18%) | 55/248 (22%) | 0/252 (0%) | 55/122 (45%) |
| Sidechain | 0/891 (0%) | 0/517 (0%) | 0/333 (0%) | 0/41 (0%) |

*Continued on next page...*

*Continued from previous page...*

|          | **Total**    | **$^1$H**   | **$^{13}$C** | **$^{15}$N** |
|----------|--------------|-------------|--------------|--------------|
| Aromatic | 0/96 (0%)    | 0/50 (0%)   | 0/41 (0%)    | 0/5 (0%)     |
| Overall  | 110/1609 (7%)| 55/815 (7%) | 0/626 (0%)   | 55/168 (33%) |

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 7%, i.e. 111 atoms were assigned a chemical shift out of a possible 1704. 0 out of 23 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

|           | **Total**     | **$^1$H**    | **$^{13}$C** | **$^{15}$N** |
|-----------|---------------|--------------|--------------|--------------|
| Backbone  | 111/657 (17%) | 55/262 (21%) | 0/266 (0%)   | 56/129 (43%) |
| Sidechain | 0/951 (0%)    | 0/552 (0%)   | 0/352 (0%)   | 0/47 (0%)    |
| Aromatic  | 0/96 (0%)     | 0/50 (0%)    | 0/41 (0%)    | 0/5 (0%)     |
| Overall   | 111/1704 (7%) | 55/864 (6%)  | 0/659 (0%)   | 56/181 (31%) |

### 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.

Random coil index (RCI) for chain A:
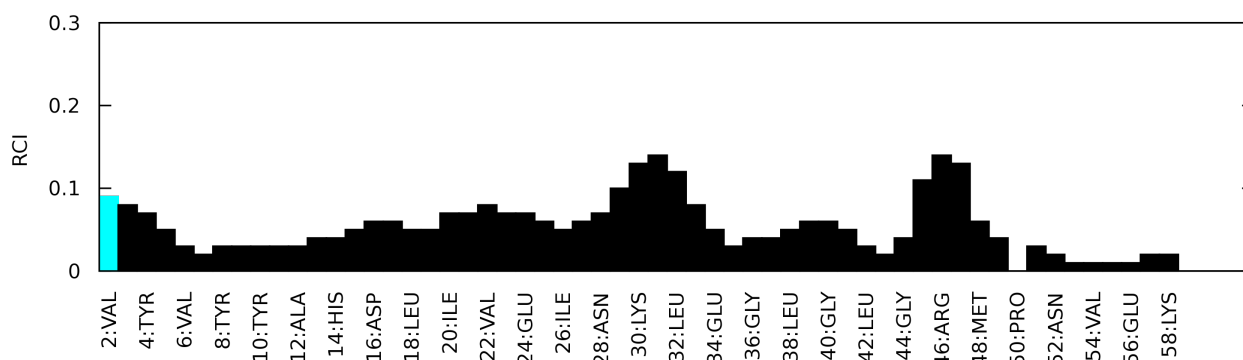


## 7.2 Chemical shift list 2

File name: 2mcn_cs.str

Chemical shift list name: *Ubiquitin-CS*

### 7.2.1 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 | 144 |
| Number of shifts mapped to atoms | 144 |
| 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 |

### 7.2.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

| Nucleus | # values | Correction ± precision, *ppm* | Suggested action |
|---|---|---|---|
| $^{13}C_{\alpha}$ | 0 | — | — |
| $^{13}C_{\beta}$ | 0 | — | — |
| $^{13}C'$ | 0 | — | — |
| $^{15}N$ | 72 | 0.81 ± 0.43 | None needed (imprecise) |

### 7.2.3 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 8%, i.e. 132 atoms were assigned a chemical shift out of a possible 1609. 0 out of 20 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

| | Total | $^1$H | $^{13}$C | $^{15}$N |
|---|---|---|---|---|
| Backbone | 132/622 (21%) | 66/248 (27%) | 0/252 (0%) | 66/122 (54%) |
| Sidechain | 0/891 (0%) | 0/517 (0%) | 0/333 (0%) | 0/41 (0%) |
| Aromatic | 0/96 (0%) | 0/50 (0%) | 0/41 (0%) | 0/5 (0%) |
| Overall | 132/1609 (8%) | 66/815 (8%) | 0/626 (0%) | 66/168 (39%) |

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 8%, i.e. 144 atoms were assigned a chemical shift out of a possible 1704. 0 out of 23 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

| | Total | $^1$H | $^{13}$C | $^{15}$N |
|---|---|---|---|---|
| Backbone | 144/657 (22%) | 72/262 (27%) | 0/266 (0%) | 72/129 (56%) |

*Continued from previous page...*

| | Total | $^1$H | $^{13}$C | $^{15}$N |
|---|---|---|---|---|
| Sidechain | 0/951 (0%) | 0/552 (0%) | 0/352 (0%) | 0/47 (0%) |
| Aromatic | 0/96 (0%) | 0/50 (0%) | 0/41 (0%) | 0/5 (0%) |
| Overall | 144/1704 (8%) | 72/864 (8%) | 0/659 (0%) | 72/181 (40%) |

### 7.2.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

### 7.2.5 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.

Random coil index (RCI) for chain B: