# Full wwPDB NMR Structure Validation Report (i)

Jun 12, 2024 – 05:47 AM EDT

| | | |
|---|---|---|
| PDB ID | : | 2KMX |
| BMRB ID | : | 16441 |
| Title | : | Solution structure of the nucleotide binding domain of the human Menkes protein in the ATP-bound form |
| Authors | : | Banci, L.; Bertini, I.; Cantini, F.; Inagaki, S.; Migliardi, M.; Rosato, A. |
| Deposited on | : | 2009-08-05 |

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:

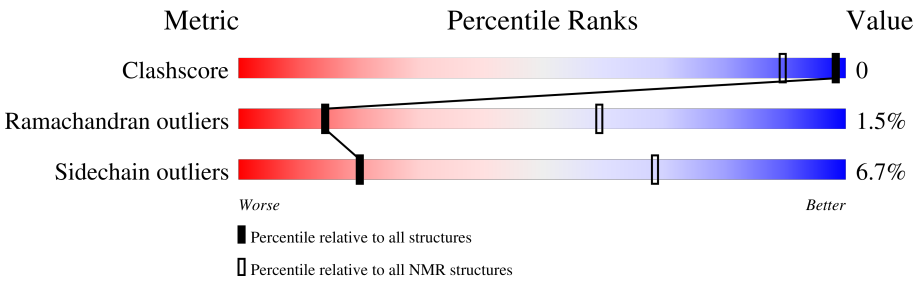
| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Mogul | : | 2022.3.0, CSD as543be (2022) |
| buster-report | : | 1.1.7 (2018) |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| wwPDB-RCI | : | v_1n_11_5_13_A (Berjanski et al., 2005) |
| PANAV | : | Wang et al. (2010) |
| wwPDB-ShiftChecker | : | v1.2 |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.36.2 |

# 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 86%.

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 | 185 | 64% • 32% |

# 2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 15 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *fewest violations*.

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:1053-A:1125, A:1177-A:1229 (126) | 0.84 | 15 |

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 3 single-model clusters were found.

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

# 3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2863 atoms, of which 1417 are hydrogens and 0 are deuteriums.

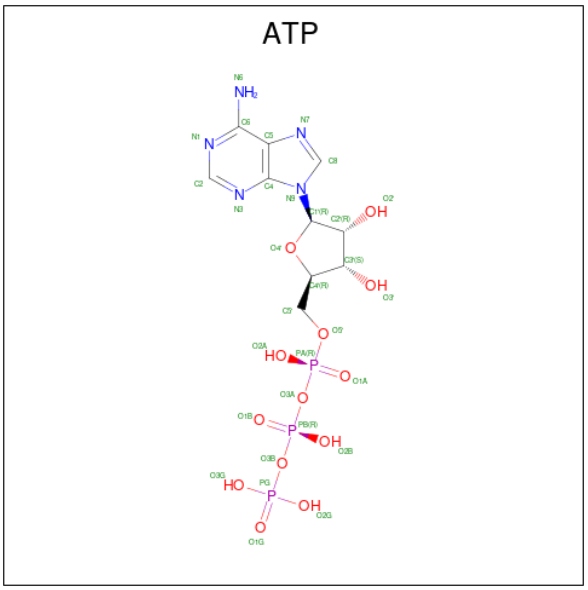
- Molecule 1 is a protein called Copper-transporting ATPase 1.

| Mol | Chain | Residues | Atoms | | | | | | Trace |
|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | H | N | O | S | |
| 1 | A | 185 | 2820 | 872 | 1405 | 249 | 285 | 9 | 0 |

There are 4 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| A | 1047 | SER | - | expression tag | UNP Q04656 |
| A | 1048 | PHE | - | expression tag | UNP Q04656 |
| A | 1049 | THR | - | expression tag | UNP Q04656 |
| A | 1050 | MET | - | expression tag | UNP Q04656 |

- Molecule 2 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$).




| Mol | Chain | Residues | Atoms | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Total | C | H | N | O | P |
| 2 | A | 1 | 43 | 10 | 12 | 5 | 13 | 3 |

# 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: Copper-transporting ATPase 1

Chain A: 64% • 32%

S1047 F1048 T1049 M1050 H1051 G1052 R1067 F1111 N1125 I1126 E1127 G1128 L1129 L1130 H1131 K1132 N1133 N1134 W1135 N1136 I1137 E1138 D1139 N1140 N1141 I1142 K1143 N1144 A1145 S1146 L1147 V1148 Q1149 I1150 D1151 A1152 S1153 N1154 E1155 Q1156 S1157 S1158 T1159 S1160 S1161 S1162 M1163 I1164 I1165 D1166 A1167 Q1168 I1169 S1170 N1171 A1172 L1173

N1174 A1175 Q1176 R1185 R1190 R1208 R1211 L1222 D1230 T1231

## 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: Copper-transporting ATPase 1

Chain A: 60% 8% • 32%

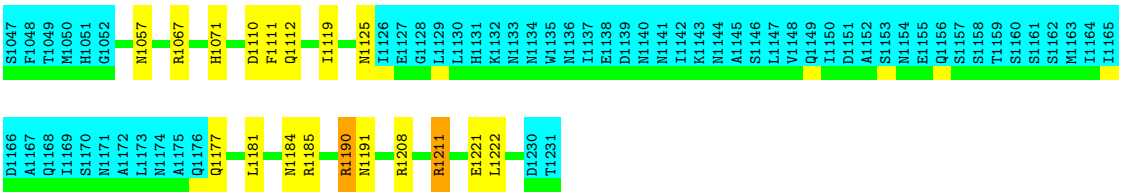
S1047 F1048 T1049 M1050 H1051 G1052 R1067 K1097 L1105 G1106 T1107 C1108 F1111 N1125 I1126 E1127 G1128 L1129 L1130 H1131 K1132 N1133 N1134 W1135 N1136 I1137 E1138 D1139 N1140 N1141 I1142 K1143 N1144 A1145 S1146 L1147 V1148 Q1149 I1150 D1151 A1152 S1153 N1154 E1155 Q1156 S1157 S1158 T1159 S1160 S1161 S1162 M1163 I1164 I1165 D1166

A1167 Q1168 I1169 S1170 N1171 A1172 L1173 N1174 A1175 Q1176 Q1177 Y1178 R1185 R1190 N1191 R1208 R1211 E1221 L1222 D1230 T1231

### 4.2.2 Score per residue for model 2

• Molecule 1: Copper-transporting ATPase 1

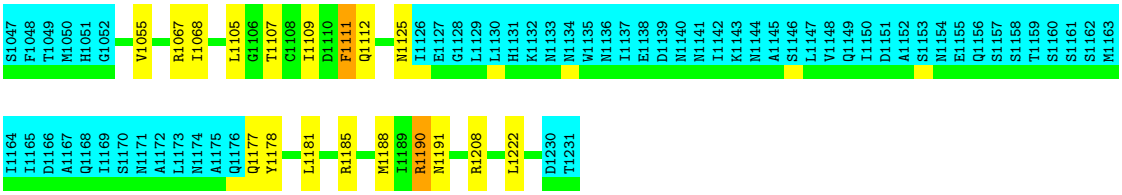
Chain A: 58% 9% • 32%




### 4.2.3 Score per residue for model 3

- Molecule 1: Copper-transporting ATPase 1

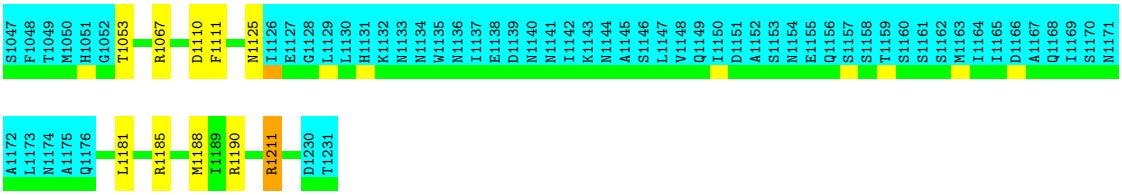
Chain A:




### 4.2.4 Score per residue for model 4

- Molecule 1: Copper-transporting ATPase 1

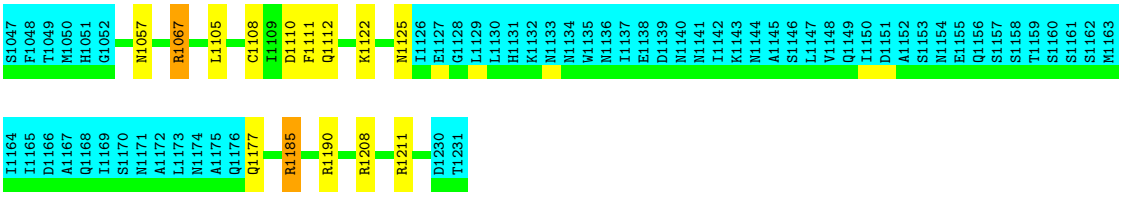
Chain A:




### 4.2.5 Score per residue for model 5

- Molecule 1: Copper-transporting ATPase 1

Chain A:



### 4.2.6 Score per residue for model 6

- Molecule 1: Copper-transporting ATPase 1

Chain A: 62% 6% • 32%

S1047 F1048 T1049 M1050 H1051 G1052 N1057 L1062 R1067 D1110 F1111 I1126 E1127 G1128 L1129 L1130 H1131 K1132 N1133 N1134 W1135 N1136 I1137 E1138 D1139 N1140 N1141 I1142 K1143 N1144 A1145 S1146 L1147 V1148 Q1149 I1150 D1151 A1152 S1153 N1154 E1155 Q1156 S1157 S1158 T1159 S1160 S1161 S1162 M1163 I1164 I1165 D1166 A1167 Q1168 I1169

S1170 N1171 A1172 L1173 N1174 A1175 Q1176 L1181 N1184 R1185 R1190 R1208 R1211 L1222 D1230 T1231

### 4.2.7 Score per residue for model 7

- Molecule 1: Copper-transporting ATPase 1

Chain A: 59% 8% • 32%

S1047 F1048 T1049 M1050 H1051 G1052 N1057 L1088 E1103 T1107 C1108 Q1112 N1125 I1126 E1127 G1128 L1129 L1130 H1131 K1132 N1133 N1134 W1135 N1136 I1137 E1138 D1139 N1140 N1141 I1142 K1143 N1144 A1145 S1146 L1147 V1148 Q1149 I1150 D1151 A1152 S1153 N1154 E1155 Q1156 S1157 S1158 T1159 S1160 S1161 S1162 M1163 I1164 I1165 D1166

A1167 Q1168 I1169 S1170 N1171 A1172 L1173 N1174 A1175 Q1176 Q1177 R1185 M1188 I1189 R1190 N1191 R1208 K1209 G1210 R1211 L1222 D1230 T1231

### 4.2.8 Score per residue for model 8

- Molecule 1: Copper-transporting ATPase 1

Chain A: 63% • • 32%

S1047 F1048 T1049 M1050 H1051 G1052 L1062 I1068 I1109 D1110 F1111 N1125 I1126 E1127 G1128 L1129 L1130 H1131 K1132 N1133 N1134 W1135 N1136 I1137 E1138 D1139 N1140 N1141 I1142 K1143 N1144 A1145 S1146 L1147 V1148 Q1149 I1150 D1151 A1152 S1153 N1154 E1155 Q1156 S1157 S1158 T1159 S1160 S1161 S1162 M1163 I1164 I1165 D1166 A1167 Q1168 I1169

S1170 N1171 A1172 L1173 N1174 A1175 Q1176 Q1177 R1190 F1202 L1222 D1230 T1231

### 4.2.9 Score per residue for model 9

- Molecule 1: Copper-transporting ATPase 1

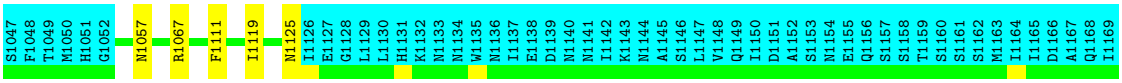
Chain A: 63% 5% 32%

S1047 F1048 T1049 M1050 H1051 G1052 T1053 R1067 N1083 D1110 N1125 I1126 E1127 G1128 L1129 L1130 H1131 K1132 N1133 N1134 W1135 N1136 I1137 E1138 D1139 N1140 N1141 I1142 K1143 N1144 A1145 S1146 L1147 V1148 Q1149 I1150 D1151 A1152 S1153 N1154 E1155 Q1156 S1157 S1158 T1159 S1160 S1161 S1162 M1163 I1164 I1165 D1166 A1167 Q1168 I1169 S1170

### 4.2.10 Score per residue for model 10

- Molecule 1: Copper-transporting ATPase 1

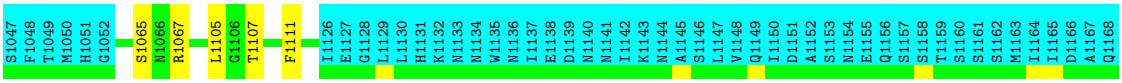
Chain A:



### 4.2.11 Score per residue for model 11

- Molecule 1: Copper-transporting ATPase 1

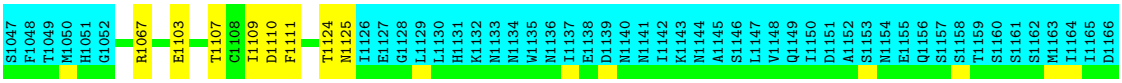
Chain A:

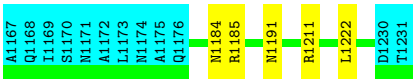


### 4.2.12 Score per residue for model 12

- Molecule 1: Copper-transporting ATPase 1

Chain A:




### 4.2.13 Score per residue for model 13

- Molecule 1: Copper-transporting ATPase 1

Chain A: 63% 5% • 32%

S1047 F1048 T1049 M1050 H1051 G1052 L1074 H1086 T1102 T1107 I1119 I1126 E1127 G1128 L1129 L1130 H1131 K1132 N1133 N1134 W1135 N1136 I1137 E1138 D1139 N1140 N1141 I1142 K1143 N1144 A1145 S1146 L1147 V1148 Q1149 I1150 D1151 A1152 S1153 N1154 E1155 Q1156 S1157 S1158 T1159 S1160 S1161 S1162 M1163 I1164 I1165 D1166 A1167 Q1168

I1169 S1170 N1171 A1172 L1173 N1174 A1175 Q1176 Q1177 Y1178 R1185 R1190 N1191 R1211 D1230 T1231

### 4.2.14 Score per residue for model 14

- Molecule 1: Copper-transporting ATPase 1

Chain A: 57% 9% •• 32%

S1047 F1048 T1049 M1050 H1051 G1052 N1057 L1062 S1065 N1066 R1067 L1088 I1109 D1110 F1111 Q1112 I1119 N1125 I1126 E1127 G1128 L1129 L1130 H1131 K1132 N1133 N1134 W1135 N1136 I1137 E1138 D1139 N1140 N1141 I1142 K1143 N1144 A1145 S1146 L1147 V1148 Q1149 I1150 D1151 A1152 S1153 N1154 E1155 Q1156 S1157 S1158 S1160

S1161 S1162 M1163 I1164 I1165 D1166 A1167 Q1168 I1169 S1170 N1171 A1172 L1173 N1174 A1175 Q1176 Q1177 L1181 R1185 R1190 F1202 E1207 R1208 R1211 T1212 A1213 L1222 D1230 T1231

### 4.2.15 Score per residue for model 15 (medoid)

- Molecule 1: Copper-transporting ATPase 1

Chain A: 64% • 32%

S1047 F1048 T1049 M1050 H1051 G1052 R1067 I1119 N1125 I1126 E1127 G1128 L1129 L1130 H1131 K1132 N1133 N1134 W1135 N1136 I1137 E1138 D1139 N1140 N1141 I1142 K1143 N1144 A1145 S1146 L1147 V1148 Q1149 I1150 D1151 A1152 S1153 N1154 E1155 Q1156 S1157 S1158 T1159 S1160 S1161 S1162 M1163 I1164 I1165 D1166 A1167 Q1168 I1169 S1170 N1171 A1172 L1173

N1174 A1175 Q1176 R1185 M1188 I1189 R1190 R1208 L1222 D1230 T1231

### 4.2.16 Score per residue for model 16

- Molecule 1: Copper-transporting ATPase 1

Chain A: 61% 6% • 32%

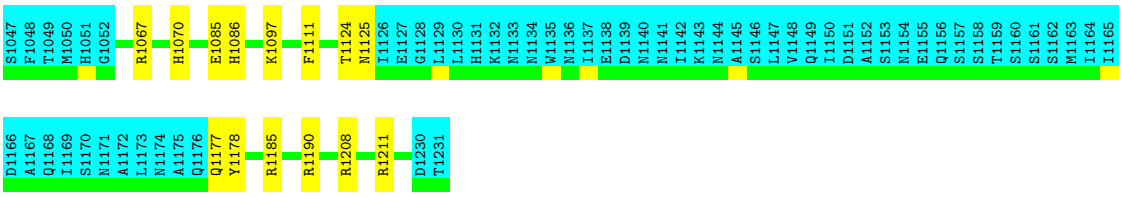
S1047 F1048 T1049 M1050 H1051 G1052 R1067 I1068 L1105 D1110 F1111 Q1112 N1125 I1126 E1127 G1128 L1129 L1130 H1131 K1132 N1133 N1134 W1135 N1136 I1137 E1138 D1139 N1140 N1141 I1142 K1143 N1144 A1145 S1146 L1147 V1148 Q1149 I1150 D1151 A1152 S1153 N1154 E1155 Q1156 S1157 S1158 T1159 S1160 S1161 S1162 M1163 I1164 I1165 D1166 A1167 Q1168

I1169 S1170 N1171 A1172 L1173 N1174 A1175 Q1176 Q1177 R1185 R1190 R1208 R1211 L1222 D1230 T1231

### 4.2.17 Score per residue for model 17

- Molecule 1: Copper-transporting ATPase 1

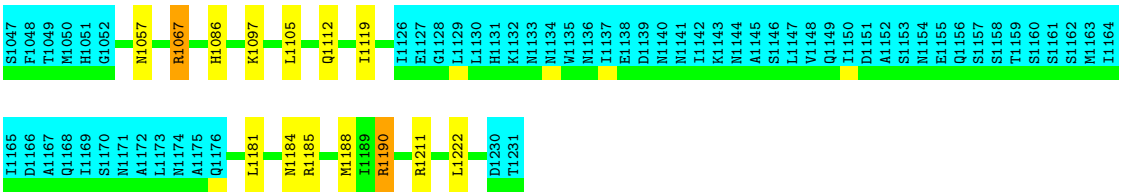
Chain A: 61% 8% 32%




### 4.2.18 Score per residue for model 18

- Molecule 1: Copper-transporting ATPase 1

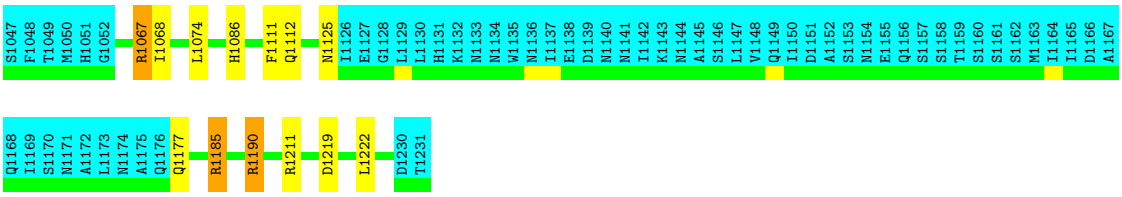
Chain A: 61% 6% • 32%




### 4.2.19 Score per residue for model 19

- Molecule 1: Copper-transporting ATPase 1

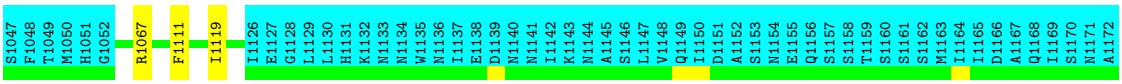
Chain A: 61% 5% • 32%




### 4.2.20 Score per residue for model 20

- Molecule 1: Copper-transporting ATPase 1

Chain A: 64% • 32%



L1173
N1174
A1175
Q1176
R1185
R1190
R1208
L1222
D1230
T1231

# 5 Refinement protocol and experimental data overview (i)

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

Of the 500 calculated structures, 20 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 solution | |
| Amber | refinement | 10 |

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 | 2054 |
| Number of shifts mapped to atoms | 2046 |
| Number of unparsed shifts | 0 |
| Number of shifts with mapping errors | 8 |
| Number of shifts with mapping warnings | 0 |
| Assignment completeness (well-defined parts) | 86% |

# 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: ATP

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the (average) root-mean-square of all Z scores of the bond lengths (or angles).

| Mol | Chain | Bond lengths | | Bond angles | |
|---|---|---|---|---|---|
| | | RMSZ | #Z>5 | RMSZ | #Z>5 |
| 1 | A | 0.65±0.01 | 0±0/979 ( 0.0± 0.0%) | 1.13±0.03 | 4±1/1330 ( 0.3± 0.1%) |
| All | All | 0.65 | 0/19580 ( 0.0%) | 1.13 | 86/26600 ( 0.3%) |

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 | 0.1±0.3 |
| All | All | 0 | 2 |

There are no bond-length outliers.

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) | Models | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Worst | Total |
| 1 | A | 1211 | ARG | NE-CZ-NH1 | 10.14 | 125.37 | 120.30 | 14 | 15 |
| 1 | A | 1067 | ARG | NE-CZ-NH1 | 9.93 | 125.26 | 120.30 | 2 | 16 |
| 1 | A | 1185 | ARG | NE-CZ-NH1 | 9.63 | 125.12 | 120.30 | 7 | 18 |
| 1 | A | 1208 | ARG | NE-CZ-NH1 | 9.54 | 125.07 | 120.30 | 20 | 13 |
| 1 | A | 1190 | ARG | NE-CZ-NH1 | 9.04 | 124.82 | 120.30 | 10 | 17 |
| 1 | A | 1190 | ARG | NE-CZ-NH2 | -6.47 | 117.06 | 120.30 | 1 | 4 |
| 1 | A | 1067 | ARG | NE-CZ-NH2 | -5.57 | 117.52 | 120.30 | 11 | 1 |
| 1 | A | 1178 | TYR | CB-CG-CD2 | -5.19 | 117.89 | 121.00 | 13 | 1 |
| 1 | A | 1111 | PHE | CB-CG-CD2 | -5.16 | 117.19 | 120.80 | 1 | 1 |

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 | 1111 | PHE | Sidechain | 1 |
| 1 | A | 1185 | ARG | Sidechain | 1 |

## 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 | 966 | 979 | 979 | 0±1 |
| All | All | 19940 | 19820 | 19820 | 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 0.

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

| Atom-1 | Atom-2 | Clash(Å) | Distance(Å) | Models | |
|---|---|---|---|---|---|
| | | | | Worst | Total |
| 1:A:1107:THR:HG22 | 1:A:1108:CYS:H | 0.56 | 1.60 | 7 | 2 |
| 1:A:1062:LEU:HD23 | 1:A:1222:LEU:HD22 | 0.51 | 1.82 | 6 | 3 |
| 1:A:1207:GLU:HG2 | 1:A:1213:ALA:HB3 | 0.45 | 1.89 | 14 | 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 | 126/185 (68%) | 115±2 (91±1%) | 9±2 (7±2%) | 2±1 (2±1%) | 14 | 59 |
| All | All | 2520/3700 (68%) | 2294 (91%) | 187 (7%) | 39 (2%) | 14 | 59 |

All 10 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 | 1125 | ASN | 15 |
| 1 | A | 1111 | PHE | 6 |
| 1 | A | 1177 | GLN | 5 |
| 1 | A | 1190 | ARG | 3 |
| 1 | A | 1067 | ARG | 2 |
| 1 | A | 1103 | GLU | 2 |
| 1 | A | 1065 | SER | 2 |
| 1 | A | 1107 | THR | 2 |
| 1 | A | 1185 | ARG | 1 |
| 1 | A | 1068 | ILE | 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 | 110/162 (68%) | 103±3 (93±3%) | 7±3 (7±3%) | 20 68 |
| All | All | 2200/3240 (68%) | 2053 (93%) | 147 (7%) | 20 68 |

All 40 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 | 1222 | LEU | 16 |
| 1 | A | 1111 | PHE | 9 |
| 1 | A | 1112 | GLN | 8 |
| 1 | A | 1181 | LEU | 8 |
| 1 | A | 1191 | ASN | 7 |
| 1 | A | 1057 | ASN | 7 |
| 1 | A | 1110 | ASP | 7 |
| 1 | A | 1119 | ILE | 7 |
| 1 | A | 1105 | LEU | 6 |
| 1 | A | 1188 | MET | 5 |
| 1 | A | 1190 | ARG | 5 |
| 1 | A | 1185 | ARG | 5 |
| 1 | A | 1178 | TYR | 4 |
| 1 | A | 1177 | GLN | 4 |
| 1 | A | 1184 | ASN | 4 |
| 1 | A | 1109 | ILE | 4 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | Models (Total) |
|---|---|---|---|---|
| 1 | A | 1086 | HIS | 4 |
| 1 | A | 1097 | LYS | 3 |
| 1 | A | 1067 | ARG | 3 |
| 1 | A | 1068 | ILE | 3 |
| 1 | A | 1221 | GLU | 2 |
| 1 | A | 1211 | ARG | 2 |
| 1 | A | 1107 | THR | 2 |
| 1 | A | 1053 | THR | 2 |
| 1 | A | 1088 | LEU | 2 |
| 1 | A | 1202 | PHE | 2 |
| 1 | A | 1124 | THR | 2 |
| 1 | A | 1074 | LEU | 2 |
| 1 | A | 1071 | HIS | 1 |
| 1 | A | 1055 | VAL | 1 |
| 1 | A | 1108 | CYS | 1 |
| 1 | A | 1122 | LYS | 1 |
| 1 | A | 1209 | LYS | 1 |
| 1 | A | 1083 | ASN | 1 |
| 1 | A | 1198 | ASP | 1 |
| 1 | A | 1215 | LEU | 1 |
| 1 | A | 1102 | THR | 1 |
| 1 | A | 1070 | HIS | 1 |
| 1 | A | 1085 | GLU | 1 |
| 1 | A | 1219 | ASP | 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 monosaccharides in this entry.

## 6.6 Ligand geometry (i)

1 ligand 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 | ATP | A | 1232 | - | 28,33,33 | 2.12±0.19 | 5±0 (16±1%) |

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 | ATP | A | 1232 | - | 34,52,52 | 1.53±0.31 | 5±1 (14±3%) |

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 | ATP | A | 1232 | - | - | 0±0,18,38,38 | 0±0,3,3,3 |

All unique bond outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) | Models | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Worst | Total |
| 2 | A | 1232 | ATP | PB-O3B | 7.39 | 1.51 | 1.59 | 19 | 20 |
| 2 | A | 1232 | ATP | PA-O3A | 6.80 | 1.52 | 1.59 | 2 | 20 |
| 2 | A | 1232 | ATP | PB-O3A | 6.55 | 1.52 | 1.59 | 19 | 20 |
| 2 | A | 1232 | ATP | PG-O3G | 2.39 | 1.45 | 1.54 | 9 | 18 |
| 2 | A | 1232 | ATP | PG-O2G | 2.31 | 1.46 | 1.54 | 18 | 17 |

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) | Models | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Worst | Total |
| 2 | A | 1232 | ATP | O4'-C1'-N9 | 8.49 | 120.00 | 108.75 | 13 | 17 |
| 2 | A | 1232 | ATP | C4'-O4'-C1' | 6.37 | 104.09 | 109.92 | 5 | 14 |
| 2 | A | 1232 | ATP | C5-C6-N6 | 5.61 | 128.86 | 120.31 | 13 | 20 |
| 2 | A | 1232 | ATP | N6-C6-N1 | 3.89 | 110.01 | 118.33 | 13 | 19 |
| 2 | A | 1232 | ATP | C1'-N9-C4 | 3.14 | 121.12 | 126.64 | 13 | 5 |
| 2 | A | 1232 | ATP | O3'-C3'-C2' | 2.66 | 103.28 | 111.82 | 17 | 5 |
| 2 | A | 1232 | ATP | C6-C5-C4 | 2.60 | 112.85 | 117.90 | 17 | 18 |
| 2 | A | 1232 | ATP | O3G-PG-O2G | 2.07 | 115.58 | 107.80 | 9 | 4 |

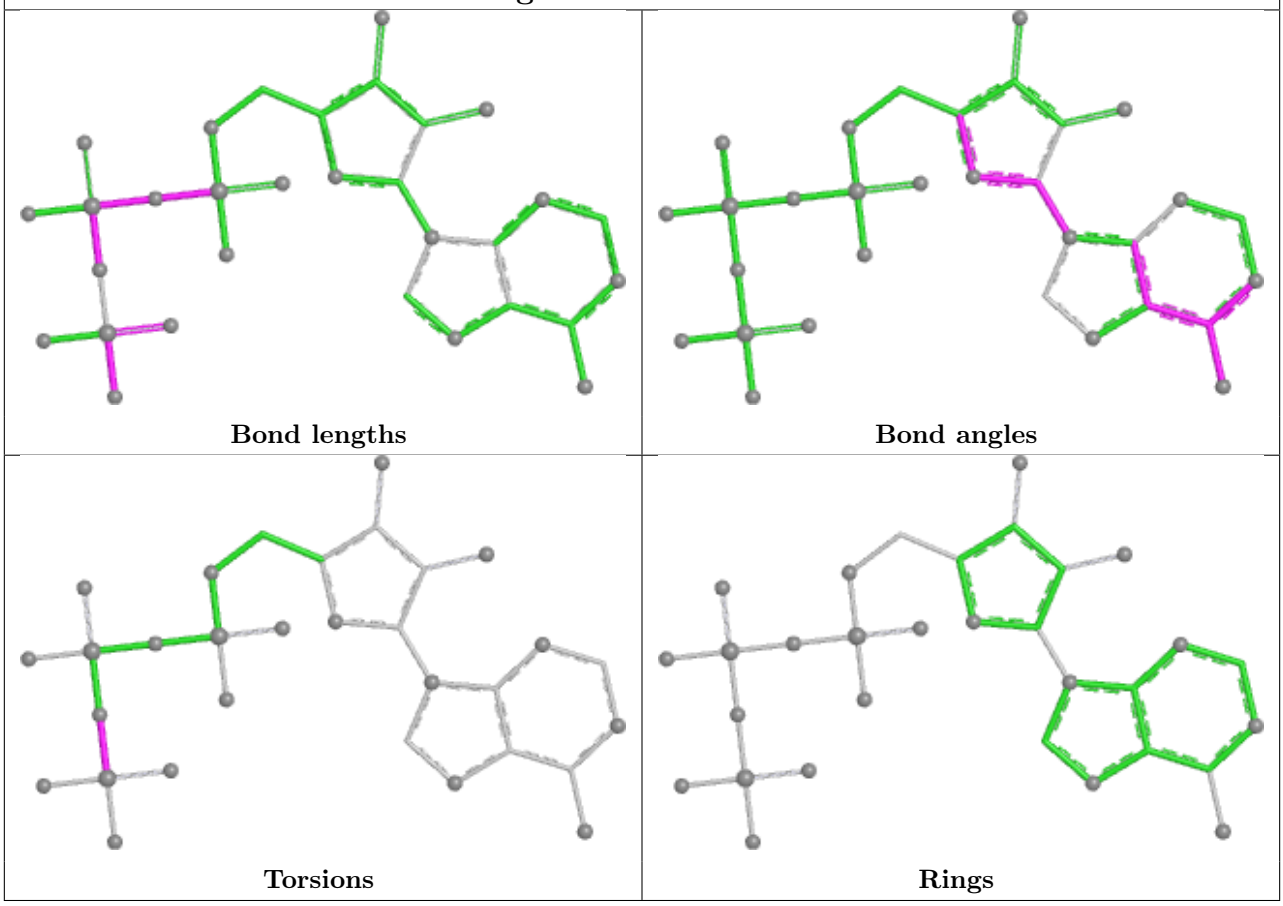
There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

**Ligand ATP A 1232**




## 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 86% for the well-defined parts and 84% for the entire structure.

## 7.1 Chemical shift list 1

File name: working_cs.cif

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 | 2054 |
| Number of shifts mapped to atoms | 2046 |
| Number of unparsed shifts | 0 |
| Number of shifts with mapping errors | 8 |
| Number of shifts with mapping warnings | 0 |
| Number of shift outliers (ShiftChecker) | 1 |

The following assigned chemical shifts were not mapped to the molecules present in the coordinate file.

- No matching atom found in the structure. All 8 occurrences are reported below.

| List ID | Chain | Res | Type | Atom | Shift Data | | |
|---|---|---|---|---|---|---|---|
| | | | | | Value | Uncertainty | Ambiguity |
| 1 | UNMAPPED | 1 | ATP | H5'1 | 3.475 | 0.030 | 2 |
| 1 | UNMAPPED | 1 | ATP | H5'2 | 3.586 | 0.030 | 2 |
| 1 | UNMAPPED | 1 | ATP | H4' | 4.082 | 0.030 | 1 |
| 1 | UNMAPPED | 1 | ATP | H3' | 3.091 | 0.030 | 1 |
| 1 | UNMAPPED | 1 | ATP | H2' | 3.543 | 0.030 | 1 |
| 1 | UNMAPPED | 1 | ATP | H1' | 5.93 | 0.030 | 1 |
| 1 | UNMAPPED | 1 | ATP | H2 | 8.189 | 0.030 | 1 |
| 1 | UNMAPPED | 1 | ATP | H8 | 7.175 | 0.030 | 1 |

### 7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

| Nucleus | # values | Correction ± precision, *ppm* | Suggested action |
|---|---|---|---|
| $^{13}C_{\alpha}$ | 175 | -0.69 ± 0.10 | Should be checked |
| $^{13}C_{\beta}$ | 164 | -0.16 ± 0.09 | None needed (< 0.5 ppm) |
| $^{13}C'$ | 157 | 0.03 ± 0.09 | None needed (< 0.5 ppm) |
| $^{15}N$ | 172 | -0.27 ± 0.25 | 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 86%, i.e. 1453 atoms were assigned a chemical shift out of a possible 1693. 0 out of 24 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

| | Total | $^{1}H$ | $^{13}C$ | $^{15}N$ |
|---|---|---|---|---|
| Backbone | 606/633 (96%) | 254/258 (98%) | 231/252 (92%) | 121/123 (98%) |
| Sidechain | 827/982 (84%) | 567/641 (88%) | 250/306 (82%) | 10/35 (29%) |
| Aromatic | 20/78 (26%) | 20/40 (50%) | 0/33 (0%) | 0/5 (0%) |
| Overall | 1453/1693 (86%) | 841/939 (90%) | 481/591 (81%) | 131/163 (80%) |

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 84%, i.e. 2045 atoms were assigned a chemical shift out of a possible 2442. 0 out of 29 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

| | Total | $^{1}H$ | $^{13}C$ | $^{15}N$ |
|---|---|---|---|---|
| Backbone | 862/930 (93%) | 358/378 (95%) | 332/370 (90%) | 172/182 (95%) |
| Sidechain | 1155/1398 (83%) | 791/910 (87%) | 343/438 (78%) | 21/50 (42%) |
| Aromatic | 28/114 (25%) | 28/59 (47%) | 0/47 (0%) | 0/8 (0%) |
| Overall | 2045/2442 (84%) | 1177/1347 (87%) | 675/855 (79%) | 193/240 (80%) |

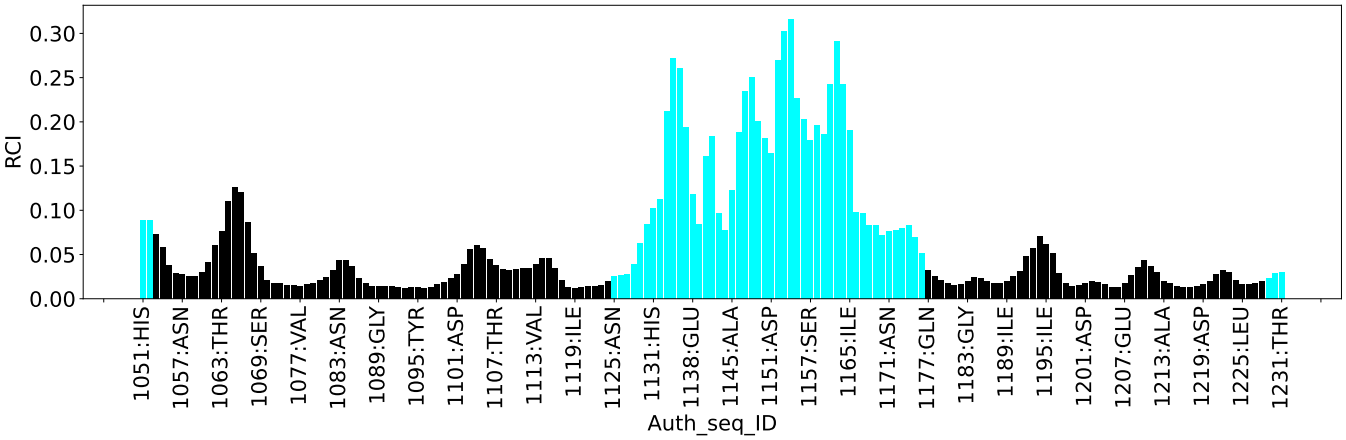
### 7.1.4 Statistically unusual chemical shifts (i)

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.

| List Id | Chain | Res | Type | Atom | Shift, *ppm* | Expected range, *ppm* | Z-score |
|---|---|---|---|---|---|---|---|
| 1 | A | 1079 | THR | HG1 | 5.44 | 0.08 – 2.19 | 20.4 |

### 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: