# W O R L D W I D E

# wwPDB

## PROTEIN DATA BANK

# Full wwPDB NMR Structure Validation Report ⓘ

May 28, 2020 – 09:06 pm BST

| | | |
|---|---|---|
| PDB ID | : | 1X9L |
| Title | : | Solution structure of CuI-DR1885 from Deinococcus Radiodurans |
| Authors | : | Banci, L.; Bertini, I.; Ciofi-Baffoni, S.; Katsari, E.; Katsaros, N.; Kubicek, K.; Structural Proteomics in Europe (SPINE) |
| Deposited on | : | 2004-08-23 |

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 ⓘ symbol.

---

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

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

# 1 Overall quality at a glance (i)

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

The overall completeness of chemical shifts assignment was not calculated.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.
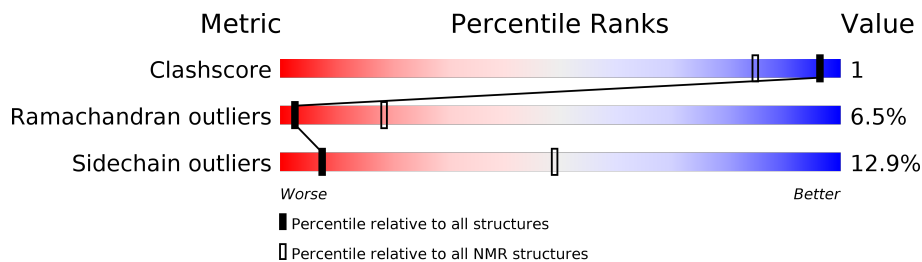


| Metric | Whole archive (#Entries) | NMR archive (#Entries) |
|---|---|---|
| Clashscore | 158937 | 12864 |
| Ramachandran outliers | 154571 | 11451 |
| Sidechain outliers | 154315 | 11428 |

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 | 149 | 69% 9% 21% |

# 2 Ensemble composition and analysis (i)

This entry contains 20 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: *minimized average structure*.

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:28-A:144 (117) | 0.52 | 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 5 single-model clusters were found.

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

# 3 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2216 atoms, of which 1137 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called CuI-DR1885.

| Mol | Chain | Residues | Atoms | | | | | | Trace |
|-----|-------|----------|-------|-----|------|-----|-----|-----|-------|
| 1 | A | 149 | Total | C | H | N | O | S | 0 |
| | | | 2215 | 673 | 1137 | 195 | 200 | 10 | |

There are 5 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------|-----------|
| A | 1 | MET | - | INITIATING METHIONINE | UNP Q9RT80 |
| A | 146 | ILE | - | CLONING ARTIFACT | UNP Q9RT80 |
| A | 147 | GLU | - | CLONING ARTIFACT | UNP Q9RT80 |
| A | 148 | GLY | - | CLONING ARTIFACT | UNP Q9RT80 |
| A | 149 | ARG | - | CLONING ARTIFACT | UNP Q9RT80 |

- Molecule 2 is COPPER (I) ION (three-letter code: CU1) (formula: Cu).

| Mol | Chain | Residues | Atoms | |
|-----|-------|----------|-------|-----|
| 2 | A | 1 | Total | Cu |
| | | | 1 | 1 |

# 4 Residue-property plots ⓘ

## 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.
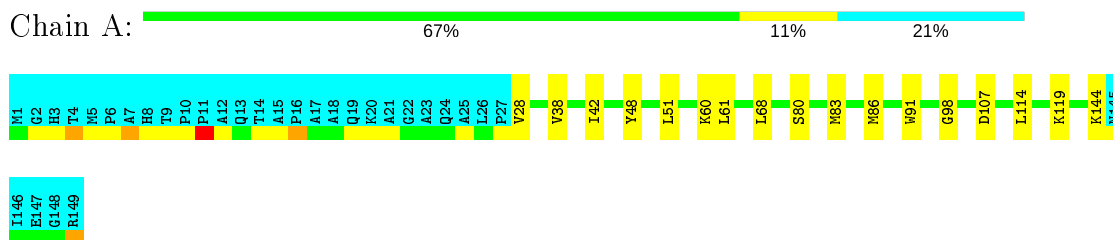
- Molecule 1: CuI-DR1885
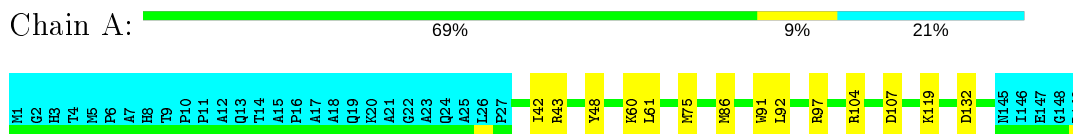


## 4.2 Scores per residue for each member of the ensemble
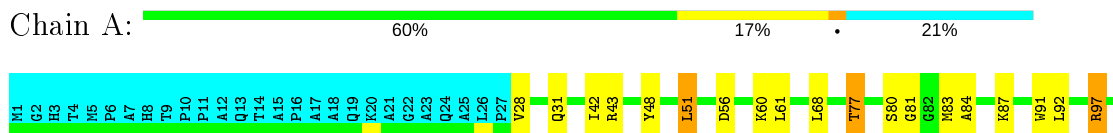
Colouring as in section 4.1 above.

### 4.2.1 Score per residue for model 1

- Molecule 1: CuI-DR1885

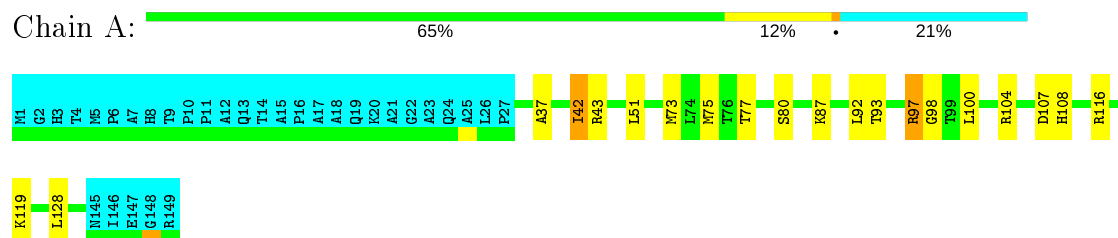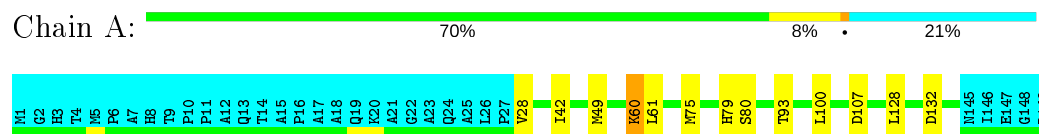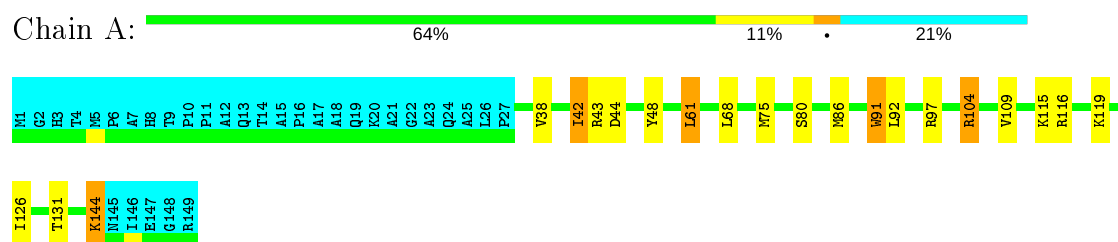

### 4.2.2 Score per residue for model 2

- Molecule 1: CuI-DR1885

L100 D105 G106 D107 L118 E122 D132 G133 R134 T135 L136 K143 K144 N145 I146 E147 G148 R149

### 4.2.3 Score per residue for model 3

• Molecule 1: CuI-DR1885

Chain A: 65% 12% • 21%

M1 G2 H3 T4 M5 P6 A7 H8 T9 P10 P11 A12 Q13 T14 A15 P16 A17 A18 Q19 K20 A21 G22 A23 Q24 A25 L26 P27 A37 I42 R43 L51 M73 L74 M75 T76 T77 S80 K87 L92 T93 R97 G98 T99 L100 R104 D107 H108 R116

K119 L128 N145 I146 E147 G148 R149

### 4.2.4 Score per residue for model 4

• Molecule 1: CuI-DR1885

Chain A: 70% 8% • 21%

M1 G2 H3 T4 M5 P6 A7 H8 T9 P10 P11 A12 Q13 T14 A15 P16 A17 A18 Q19 K20 A21 G22 A23 Q24 A25 L26 P27 V28 I42 M49 K60 L61 M75 H79 S80 T93 L100 D107 L128 D132 N145 I146 E147 G148 R149

### 4.2.5 Score per residue for model 5

• Molecule 1: CuI-DR1885

Chain A: 64% 11% • 21%

M1 G2 H3 T4 M5 P6 A7 H8 T9 P10 P11 A12 Q13 T14 A15 P16 A17 A18 Q19 K20 A21 G22 A23 Q24 A25 L26 P27 V38 I42 R43 D44 Y48 L61 L68 M75 S80 M86 W91 L92 R97 R104 V109 K115 R116 K119

I126 T131 K144 N145 I146 E147 G148 R149

### 4.2.6 Score per residue for model 6

• Molecule 1: CuI-DR1885

Chain A: 67% 11% • 21%

M1 G2 H3 T4 M5 P6 A7 H8 T9 P10 P11 A12 Q13 T14 A15 P16 A17 A18 Q19 K20 A21 G22 A23 Q24 A25 L26 P27 A37 I42 L61 V62 M75 T76 M83 A84 K87 M88 V89 R97 D107 H108 K115 R116 P117 L118 E122 N145 I146 E147

G148
R149

## 4.2.7 Score per residue for model 7

• Molecule 1: CuI-DR1885

Chain A:	65%	11%	• •	21%

M1 G2 H3 T4 M5 P6 A7 H8 T9 P10 P11 A12 Q13 T14 A15 P16 A17 A18 Q19 K20 A21 G22 A23 Q24 A25 L26 P27 V38 I42 R43 D44 Y48 M49 T50 L51 D56 L61 M75 M86 R97 L100 R104 D107 L114 L118 K119

L128 T131 N145 I146 E147 G148 R149

## 4.2.8 Score per residue for model 8

• Molecule 1: CuI-DR1885

Chain A:	60%	17%	•	21%

M1 G2 H3 T4 M5 P6 A7 H8 T9 P10 P11 A12 Q13 T14 A15 P16 A17 A18 Q19 K20 A21 G22 A23 Q24 A25 L26 P27 V30 A37 I42 R43 Y48 K54 Q57 K60 L61 V62 S80 M86 W91 L92 R97 L100 D105 G106 D107

K115 R116 P117 L118 K119 V120 D132 G133 R134 T135 L136 K143 K144 N145 I146 E147 G148 R149

## 4.2.9 Score per residue for model 9 (medoid)

• Molecule 1: CuI-DR1885

Chain A:	66%	11%	•	21%

M1 G2 H3 T4 M5 P6 A7 H8 T9 P10 P11 A12 Q13 T14 A15 P16 A17 A18 Q19 K20 A21 G22 A23 Q24 A25 L26 P27 I42 R43 D44 S55 K60 L61 S71 G81 M86 W91 L92 R97 G98 Q103 D107 L118 E122 D132 G133 R134

N145 I146 E147 G148 R149

## 4.2.10 Score per residue for model 10

• Molecule 1: CuI-DR1885

Chain A:	64%	12%	•	21%

### 4.2.11 Score per residue for model 11

- Molecule 1: CuI-DR1885



### 4.2.12 Score per residue for model 12

- Molecule 1: CuI-DR1885



### 4.2.13 Score per residue for model 13

- Molecule 1: CuI-DR1885



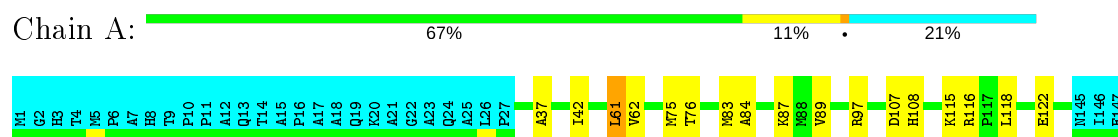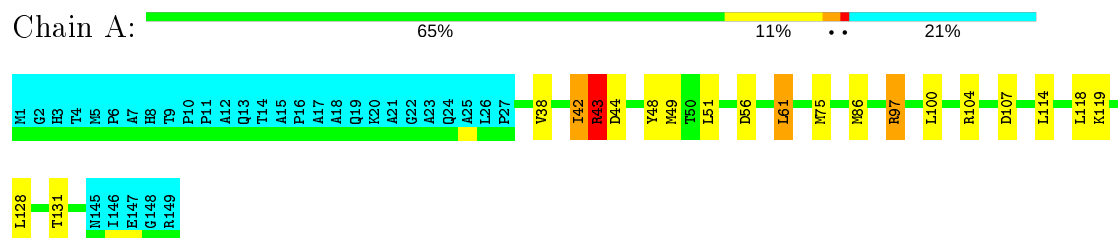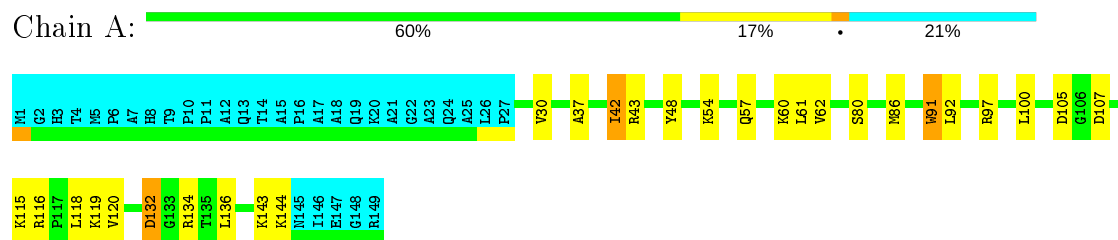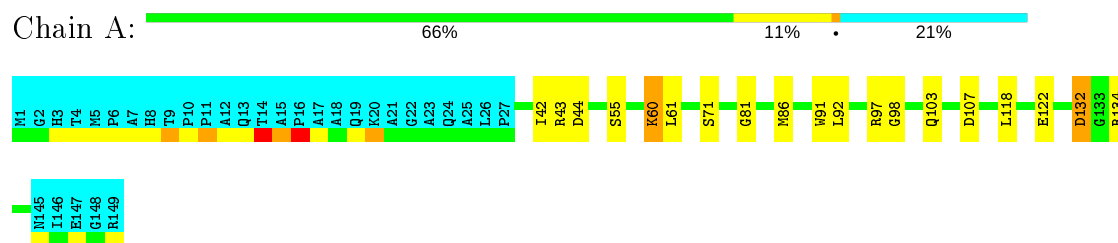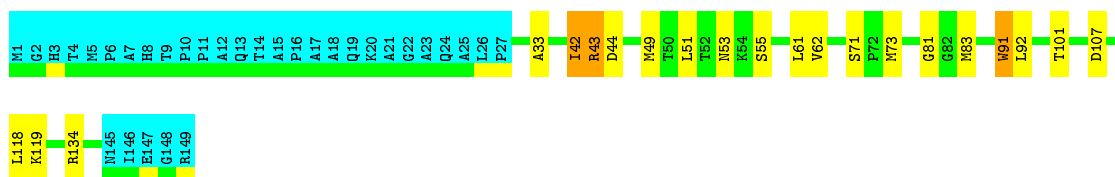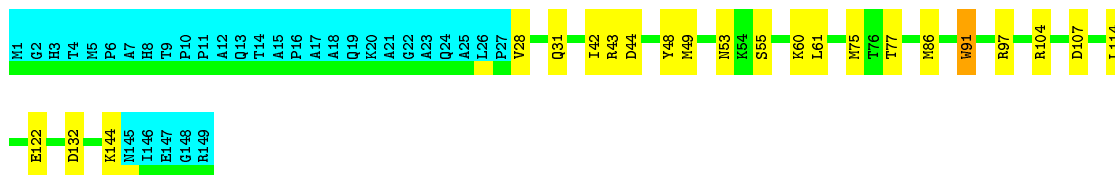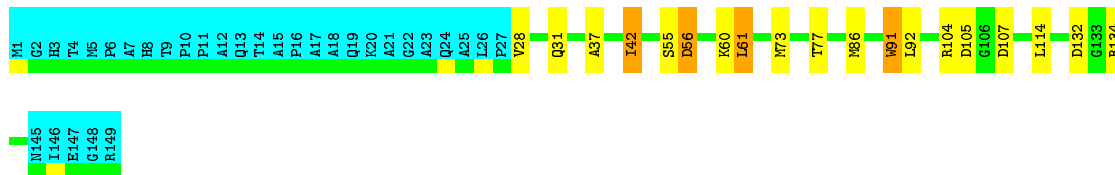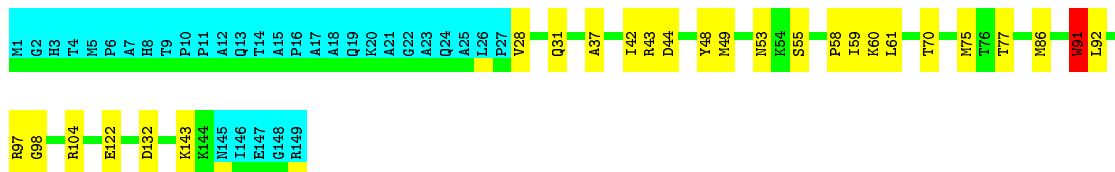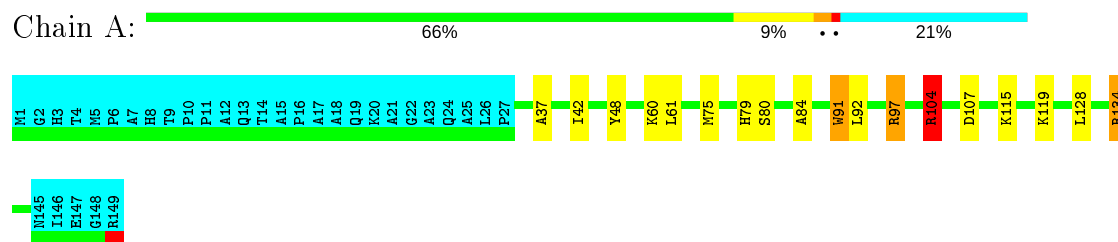### 4.2.14 Score per residue for model 14

- Molecule 1: CuI-DR1885

Chain A: 66% 9% 21%

M1 G2 H3 T4 M5 P6 A7 H8 T9 P10 P11 A12 Q13 T14 A15 P16 A17 A18 Q19 K20 A21 G22 A23 Q24 A25 L26 P27 A37 I42 Y48 K60 L61 M75 H79 S80 A84 W91 L92 R97 R104 D107 K115 K119 L128 R134

N145 I146 E147 G148 R149

## 4.2.15 Score per residue for model 15

• Molecule 1: CuI-DR1885

Chain A: 64% 13% 21%

M1 G2 H3 T4 M5 P6 A7 H8 T9 P10 P11 A12 Q13 T14 A15 P16 A17 A18 Q19 K20 A21 G22 A23 Q24 A25 L26 P27 V38 I42 R43 Y48 L51 T52 N53 K54 S55 M75 A84 G85 M86 L92 A96 R97 L100 D105 G106 K119 E122

T127 L128 K144 N145 I146 E147 G148 R149

## 4.2.16 Score per residue for model 16

• Molecule 1: CuI-DR1885

Chain A: 63% 14% 21%

M1 G2 H3 T4 M5 P6 A7 H8 T9 P10 P11 A12 Q13 T14 A15 P16 A17 A18 Q19 K20 A21 G22 A23 Q24 A25 L26 P27 V30 A37 I42 Y48 K60 L61 V62 L68 T76 A84 G85 W91 L92 R97 Q103 R116 P117 L118 K119 V120 G121

E122 L128 K129 D132 L136 K144 N145 I146 E147 G148 R149

## 4.2.17 Score per residue for model 17

• Molecule 1: CuI-DR1885

Chain A: 63% 13% 21%

M1 G2 H3 T4 M5 P6 A7 H8 T9 P10 P11 A12 Q13 T14 A15 P16 A17 A18 Q19 K20 A21 G22 A23 Q24 A25 L26 P27 V28 I42 M49 S55 D56 K60 H79 S80 M83 A84 G85 M86 W91 A96 R97 G106 D107 K115 K119 E122

L128 D132 L136 K144 N145 I146 E147 G148 R149
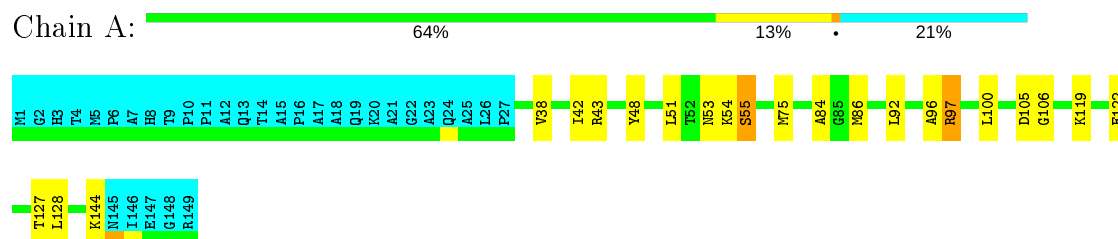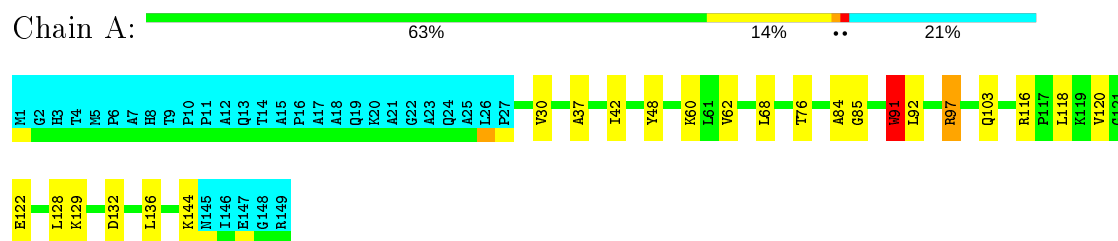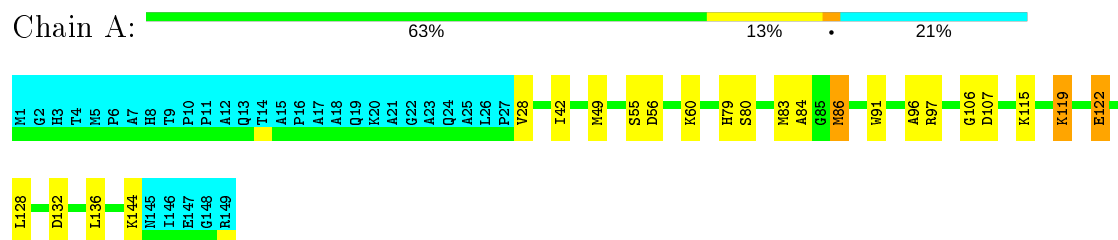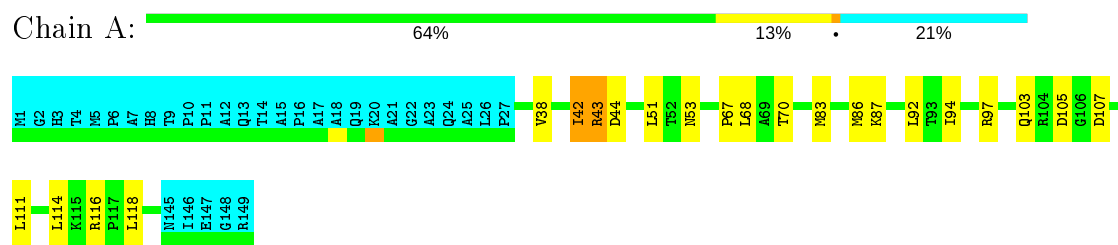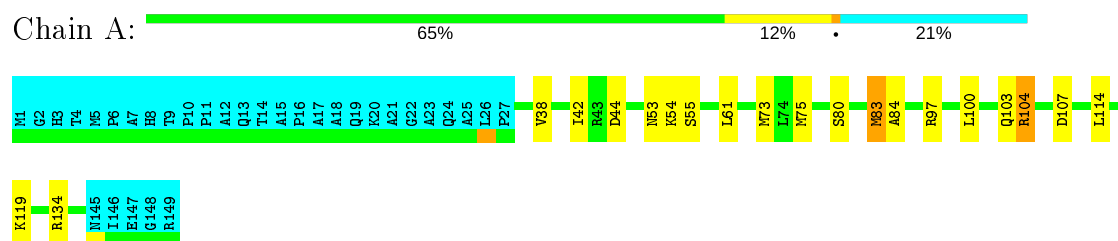
#### 4.2.18 Score per residue for model 18
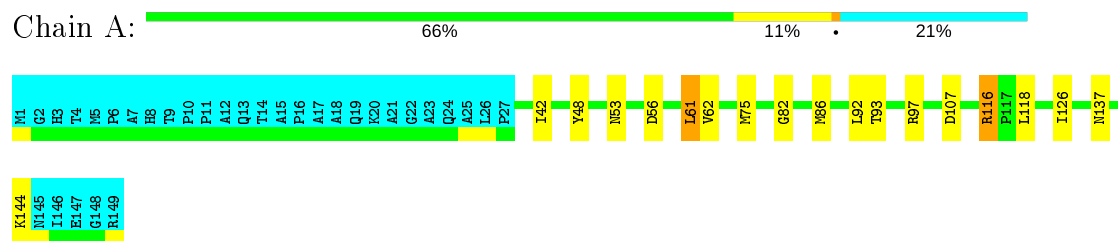
- Molecule 1: CuI-DR1885



#### 4.2.19 Score per residue for model 19

- Molecule 1: CuI-DR1885



#### 4.2.20 Score per residue for model 20

- Molecule 1: CuI-DR1885

# 5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *simulated annealing, torsion angle dynamics*.

Of the 200 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 |
|---|---|---|
| DYANA | refinement | 1.5 |

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

COVALENT-GEOMETRY INFOmissingINFO

## 5.1 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 | 853 | 914 | 914 | 1±1 |
| All | All | 17080 | 18280 | 18280 | 27 |

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 1.

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:53:ASN:HD21 | 1:A:94:ILE:HD12 | 0.58 | 1.57 | 18 | 1 |
| 1:A:42:ILE:HG22 | 1:A:43:ARG:H | 0.46 | 1.69 | 8 | 5 |
| 1:A:62:VAL:HB | 1:A:91:TRP:CD1 | 0.45 | 2.46 | 16 | 1 |
| 1:A:61:LEU:HD13 | 1:A:61:LEU:C | 0.45 | 2.32 | 6 | 3 |
| 1:A:61:LEU:C | 1:A:61:LEU:HD13 | 0.44 | 2.33 | 12 | 3 |
| 1:A:42:ILE:HG22 | 1:A:43:ARG:N | 0.43 | 2.28 | 3 | 2 |
| 1:A:60:LYS:NZ | 1:A:132:ASP:OD1 | 0.43 | 2.51 | 4 | 1 |
| 1:A:28:VAL:HG22 | 1:A:51:LEU:HG | 0.43 | 1.91 | 2 | 1 |
| 1:A:51:LEU:HD23 | 1:A:94:ILE:HD13 | 0.42 | 1.92 | 18 | 1 |
| 1:A:91:TRP:C | 1:A:92:LEU:HD22 | 0.42 | 2.35 | 13 | 1 |

*Continued on next page...*

*Continued from previous page...*

| Atom-1 | Atom-2 | Clash(Å) | Distance(Å) | Models | |
|---|---|---|---|---|---|
| | | | | Worst | Total |
| 1:A:58:PRO:C | 1:A:59:ILE:HD12 | 0.42 | 2.36 | 13 | 1 |
| 1:A:81:GLY:HA3 | 1:A:100:LEU:HD13 | 0.41 | 1.91 | 2 | 1 |
| 1:A:60:LYS:NZ | 1:A:132:ASP:OD2 | 0.41 | 2.53 | 17 | 2 |
| 1:A:79:HIS:CD2 | 1:A:84:ALA:CB | 0.41 | 3.04 | 17 | 1 |
| 1:A:28:VAL:O | 1:A:28:VAL:HG13 | 0.40 | 2.16 | 2 | 1 |
| 1:A:119:LYS:NZ | 1:A:122:GLU:OE2 | 0.40 | 2.54 | 17 | 1 |
| 1:A:56:ASP:OD2 | 1:A:97:ARG:NH1 | 0.40 | 2.55 | 20 | 1 |

## 5.2 Torsion angles (i)

### 5.2.1 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 | 117/149 (79%) | 81±4 (69±3%) | 28±3 (24±3%) | 8±1 (6±1%) | 2 | 18 |
| All | All | 2340/2980 (79%) | 1626 (69%) | 563 (24%) | 151 (6%) | 2 | 18 |

All 35 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 | 42 | ILE | 20 |
| 1 | A | 107 | ASP | 15 |
| 1 | A | 91 | TRP | 11 |
| 1 | A | 80 | SER | 9 |
| 1 | A | 55 | SER | 8 |
| 1 | A | 116 | ARG | 7 |
| 1 | A | 97 | ARG | 7 |
| 1 | A | 38 | VAL | 6 |
| 1 | A | 84 | ALA | 6 |
| 1 | A | 28 | VAL | 6 |
| 1 | A | 37 | ALA | 5 |
| 1 | A | 98 | GLY | 4 |
| 1 | A | 62 | VAL | 4 |
| 1 | A | 86 | MET | 4 |
| 1 | A | 43 | ARG | 3 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | Models (Total) |
|-----|-------|-----|------|----------------|
| 1 | A | 105 | ASP | 3 |
| 1 | A | 132 | ASP | 3 |
| 1 | A | 144 | LYS | 3 |
| 1 | A | 120 | VAL | 2 |
| 1 | A | 79 | HIS | 2 |
| 1 | A | 106 | GLY | 2 |
| 1 | A | 56 | ASP | 2 |
| 1 | A | 131 | THR | 2 |
| 1 | A | 83 | MET | 2 |
| 1 | A | 114 | LEU | 2 |
| 1 | A | 81 | GLY | 2 |
| 1 | A | 96 | ALA | 2 |
| 1 | A | 30 | VAL | 2 |
| 1 | A | 118 | LEU | 1 |
| 1 | A | 68 | LEU | 1 |
| 1 | A | 67 | PRO | 1 |
| 1 | A | 108 | HIS | 1 |
| 1 | A | 82 | GLY | 1 |
| 1 | A | 85 | GLY | 1 |
| 1 | A | 33 | ALA | 1 |

## 5.2.2 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 | 94/115 (82%) | 82±2 (87±3%) | 12±2 (13±3%) | 7 | 49 |
| All | All | 1880/2300 (82%) | 1638 (87%) | 242 (13%) | 7 | 49 |

All 51 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 | 61 | LEU | 14 |
| 1 | A | 92 | LEU | 12 |
| 1 | A | 75 | MET | 11 |
| 1 | A | 119 | LYS | 10 |
| 1 | A | 97 | ARG | 10 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | Models (Total) |
|-----|-------|-----|------|----------------|
| 1 | A | 60 | LYS | 10 |
| 1 | A | 91 | TRP | 9 |
| 1 | A | 86 | MET | 9 |
| 1 | A | 118 | LEU | 8 |
| 1 | A | 44 | ASP | 8 |
| 1 | A | 122 | GLU | 8 |
| 1 | A | 104 | ARG | 8 |
| 1 | A | 128 | LEU | 7 |
| 1 | A | 53 | ASN | 6 |
| 1 | A | 43 | ARG | 6 |
| 1 | A | 49 | MET | 6 |
| 1 | A | 100 | LEU | 6 |
| 1 | A | 51 | LEU | 6 |
| 1 | A | 83 | MET | 5 |
| 1 | A | 77 | THR | 5 |
| 1 | A | 132 | ASP | 5 |
| 1 | A | 115 | LYS | 5 |
| 1 | A | 144 | LYS | 5 |
| 1 | A | 73 | MET | 4 |
| 1 | A | 134 | ARG | 4 |
| 1 | A | 136 | LEU | 4 |
| 1 | A | 68 | LEU | 4 |
| 1 | A | 103 | GLN | 4 |
| 1 | A | 114 | LEU | 4 |
| 1 | A | 87 | LYS | 4 |
| 1 | A | 31 | GLN | 4 |
| 1 | A | 56 | ASP | 3 |
| 1 | A | 93 | THR | 3 |
| 1 | A | 126 | ILE | 2 |
| 1 | A | 71 | SER | 2 |
| 1 | A | 105 | ASP | 2 |
| 1 | A | 54 | LYS | 2 |
| 1 | A | 70 | THR | 2 |
| 1 | A | 76 | THR | 2 |
| 1 | A | 143 | LYS | 2 |
| 1 | A | 107 | ASP | 1 |
| 1 | A | 101 | THR | 1 |
| 1 | A | 57 | GLN | 1 |
| 1 | A | 116 | ARG | 1 |
| 1 | A | 127 | THR | 1 |
| 1 | A | 109 | VAL | 1 |
| 1 | A | 129 | LYS | 1 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | Models (Total) |
|-----|-------|-----|------|----------------|
| 1 | A | 111 | LEU | 1 |
| 1 | A | 137 | ASN | 1 |
| 1 | A | 42 | ILE | 1 |
| 1 | A | 89 | VAL | 1 |

#### 5.2.3　RNA (i)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 5.4　Carbohydrates (i)

There are no carbohydrates in this entry.

### 5.5　Ligand geometry (i)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

### 5.6　Other polymers (i)

There are no such molecules in this entry.

### 5.7　Polymer linkage issues (i)

There are no chain breaks in this entry.

# 6 Chemical shift validation (i)

No chemical shift data were provided