# W O R L D W I D E
# wwPDB
## P R O T E I N   D A T A   B A N K

# Full wwPDB NMR Structure Validation Report ⓘ

Feb 26, 2022 – 05:43 PM EST

|  |  |  |
|---|---|---|
| PDB ID | : | 2BCB |
| Title | : | HIGH-RESOLUTION SOLUTION STRUCTURE OF CALCIUM-LOADED CALBINDIN D9K |
| Authors | : | Kordel, J.; Skelton, N.J.; Akke, M.; Chazin, W.J. |
| Deposited on | : | 1993-08-18 |

# 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 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.
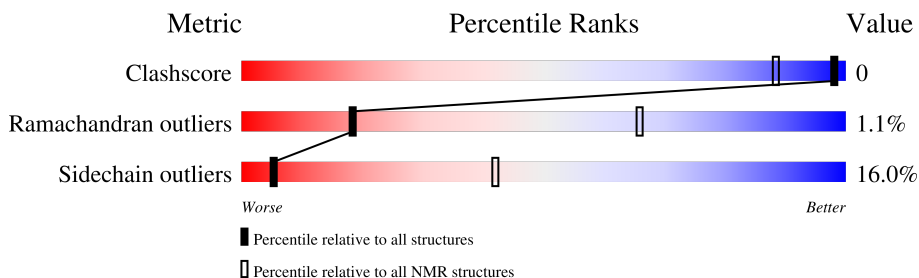


| Metric | Whole archive (#Entries) | NMR archive (#Entries) |
|---|---|---|
| Clashscore | 158937 | 12864 |
| Ramachandran outliers | 154571 | 11451 |
| Sidechain outliers | 154315 | 11428 |

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 | 75 | 76% 12% 12% |

# 2 Ensemble composition and analysis ⓘ

This entry contains 32 models. Model 24 is the overall representative, medoid model (most similar to other models).

The following residues are included in the computation of the global validation metrics.

| Well-defined (core) protein residues | | | |
|---|---|---|---|
| Well-defined core | Residue range (total) | Backbone RMSD (Å) | Medoid model |
| 1 | A:3-A:37, A:45-A:75 (66) | 0.48 | 24 |

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 5 clusters and 3 single-model clusters were found.

| Cluster number | Models |
|---|---|
| 1 | 1, 2, 3, 5, 6, 8, 10, 13, 14, 18, 21, 22, 23, 24, 25, 27, 30 |
| 2 | 7, 16, 19, 26 |
| 3 | 12, 15, 29 |
| 4 | 17, 20, 32 |
| 5 | 4, 9 |
| Single-model clusters | 11; 28; 31 |

# 3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1194 atoms, of which 598 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called CALBINDIN D9K.

| Mol | Chain | Residues | Atoms | | | | | Trace |
|-----|-------|----------|-------|-----|-----|----|-----|-------|
| 1 | A | 75 | Total | C | H | N | O | 0 |
| | | | 1194 | 381 | 598 | 91 | 124 | |

There is a discrepancy between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|----------|------------|
| A | 43 | GLY | PRO | conflict | UNP P02633 |

# 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.
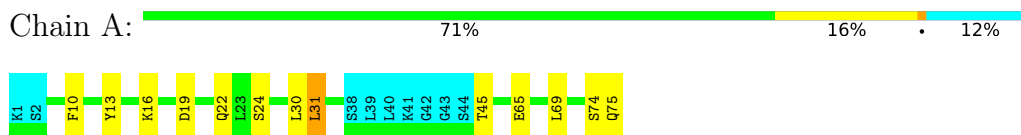
• Molecule 1: CALBINDIN D9K
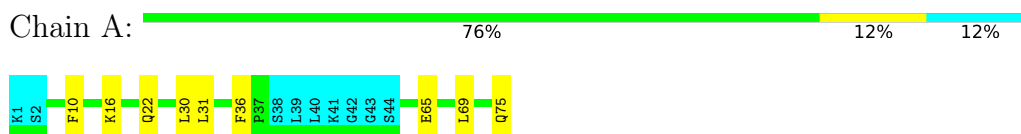


## 4.2  Scores per residue for each member of the ensemble
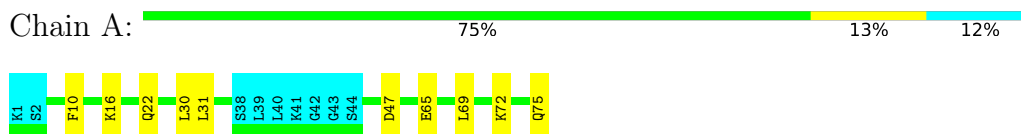
Colouring as in section 4.1 above.

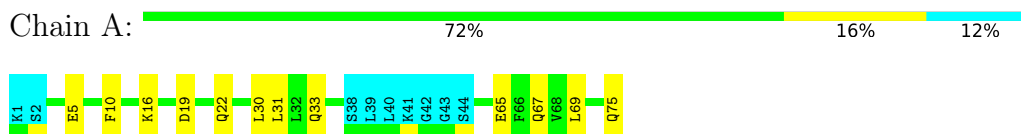### 4.2.1  Score per residue for model 1

• Molecule 1: CALBINDIN D9K



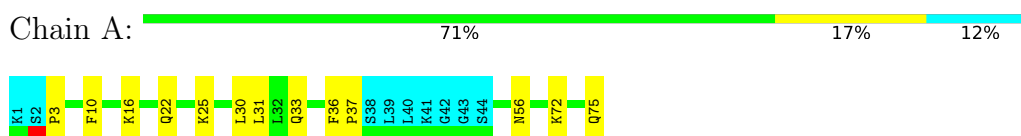### 4.2.2  Score per residue for model 2

• Molecule 1: CALBINDIN D9K

### 4.2.3 Score per residue for model 3

• Molecule 1: CALBINDIN D9K

Chain A:

### 4.2.4 Score per residue for model 4

• Molecule 1: CALBINDIN D9K

Chain A:

### 4.2.5 Score per residue for model 5

• Molecule 1: CALBINDIN D9K

Chain A:

### 4.2.6 Score per residue for model 6

• Molecule 1: CALBINDIN D9K

Chain A:

### 4.2.7 Score per residue for model 7

• Molecule 1: CALBINDIN D9K

Chain A:

#### 4.2.8 Score per residue for model 8
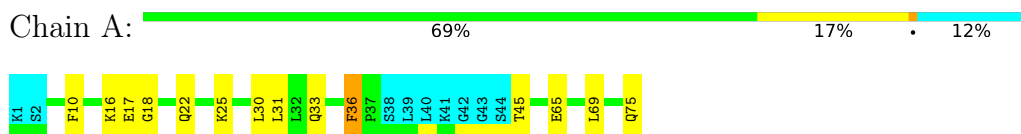
• Molecule 1: CALBINDIN D9K

Chain A:



#### 4.2.9 Score per residue for model 9

• Molecule 1: CALBINDIN D9K

Chain A:



#### 4.2.10 Score per residue for model 10

• Molecule 1: CALBINDIN D9K

Chain A:
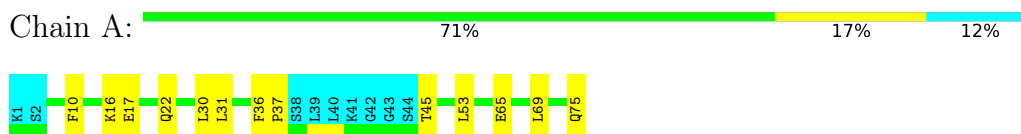


#### 4.2.11 Score per residue for model 11

• Molecule 1: CALBINDIN D9K

Chain A:
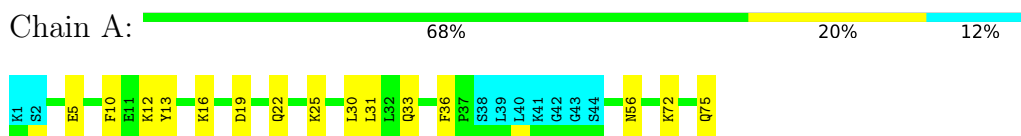


#### 4.2.12 Score per residue for model 12

• Molecule 1: CALBINDIN D9K

Chain A:
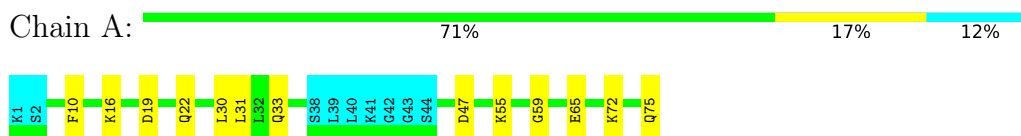
#### 4.2.13 Score per residue for model 13

• Molecule 1: CALBINDIN D9K

Chain A:



#### 4.2.14 Score per residue for model 14

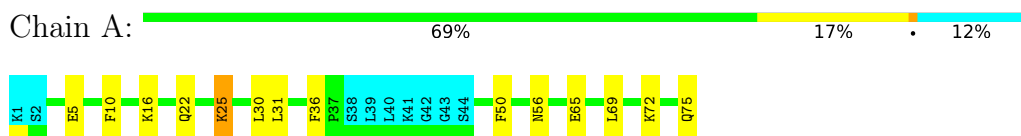• Molecule 1: CALBINDIN D9K

Chain A:



#### 4.2.15 Score per residue for model 15

• Molecule 1: CALBINDIN D9K
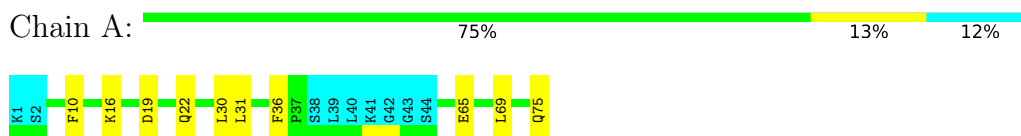
Chain A:



#### 4.2.16 Score per residue for model 16
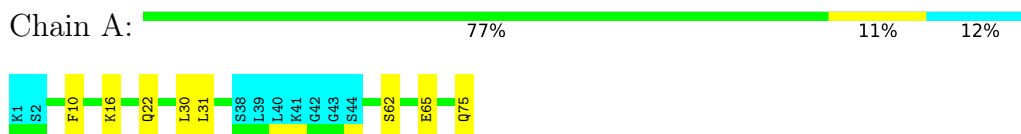
• Molecule 1: CALBINDIN D9K

Chain A:



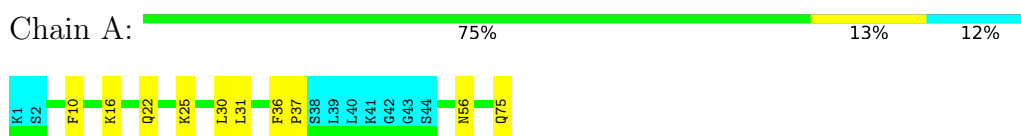#### 4.2.17 Score per residue for model 17

• Molecule 1: CALBINDIN D9K

Chain A:

### 4.2.18 Score per residue for model 18

• Molecule 1: CALBINDIN D9K
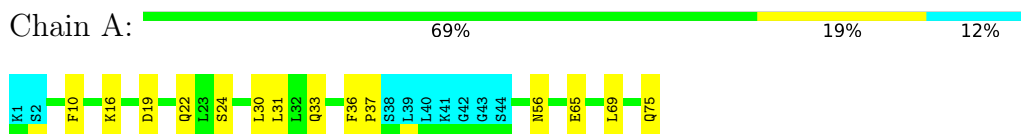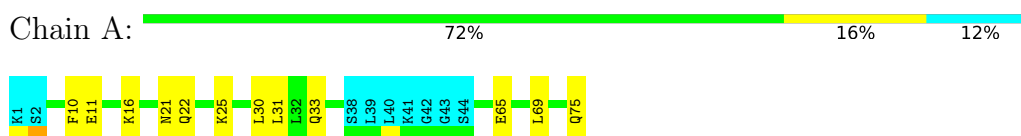
Chain A:                                    73%                      15%          12%

K1 S2 F10 K16 Q22 L30 L31 P37 S38 L39 L40 K41 G42 G43 S44 E65 L69 K72 I73 S74 Q75
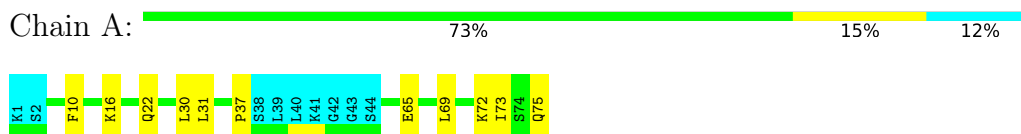
### 4.2.19 Score per residue for model 19

• Molecule 1: CALBINDIN D9K

Chain A:                                    71%                      17%          12%

K1 S2 F10 Y13 K16 D19 Q22 L30 L31 T34 E35 F36 P37 S38 L39 L40 K41 G42 G43 S44 N56 E65 K72 Q75

### 4.2.20 Score per residue for model 20

• Molecule 1: CALBINDIN D9K

Chain A:                                    72%                      15%    •     12%

K1 S2 F10 K16 D19 Q22 L30 L31 F36 P37 S38 L39 L40 K41 G42 G43 S44 N56 G57 D58 E65 K72 Q75

### 4.2.21 Score per residue for model 21

• Molecule 1: CALBINDIN D9K

Chain A:                                    71%                      16%    •     12%

K1 S2 E5 F10 K16 Q22 L23 S24 L30 L31 F36 P37 S38 L39 L40 K41 G42 G43 S44 E64 E65 L69 K72 Q75

### 4.2.22 Score per residue for model 22

• Molecule 1: CALBINDIN D9K

Chain A:                                    68%                      20%          12%

K1 S2 F10 E11 K12 Y13 K16 E17 Q22 K25 L30 L31 L32 Q33 S38 L39 L40 K41 G42 G43 S44 L49 F50 E65 L69 Q75

#### 4.2.23 Score per residue for model 23

• Molecule 1: CALBINDIN D9K

Chain A:

73% 15% 12%

K1 S2 F10 K16 Q22 K25 L30 L31 L32 Q33 S38 L39 L40 K41 G42 G43 S44 E60 E65 K72 Q75
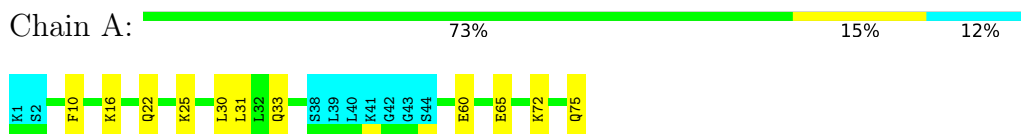
#### 4.2.24 Score per residue for model 24 (medoid)

• Molecule 1: CALBINDIN D9K

Chain A:

76% 12% 12%

K1 S2 F10 K16 Q22 L30 L31 S38 L39 L40 K41 G42 G43 S44 T45 E65 L69 Q75
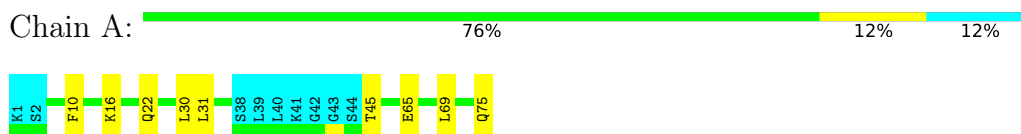
#### 4.2.25 Score per residue for model 25

• Molecule 1: CALBINDIN D9K

Chain A:

69% 19% 12%

K1 S2 E5 F10 K16 Q22 L23 S24 L30 L31 S38 L39 L40 K41 G42 G43 S44 T45 L46 D47 E64 E65 L69 K72 Q75
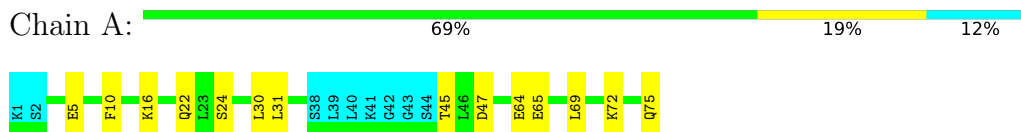
#### 4.2.26 Score per residue for model 26

• Molecule 1: CALBINDIN D9K

Chain A:

72% 15% 12%

K1 S2 F10 Y13 K16 D19 Q22 K25 L30 L31 F36 P37 S38 L39 L40 K41 G42 G43 S44 E65 L69 Q75
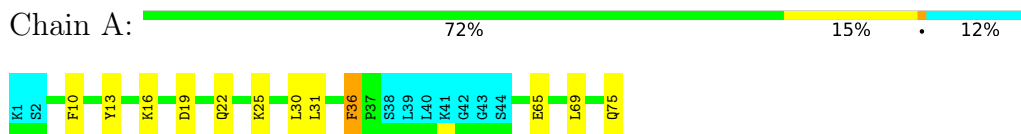
#### 4.2.27 Score per residue for model 27

• Molecule 1: CALBINDIN D9K

Chain A:

73% 15% 12%

K1 S2 F10 Y13 K16 Q22 L30 L31 F36 P37 S38 L39 L40 K41 G42 G43 S44 T45 E65 L69 Q75
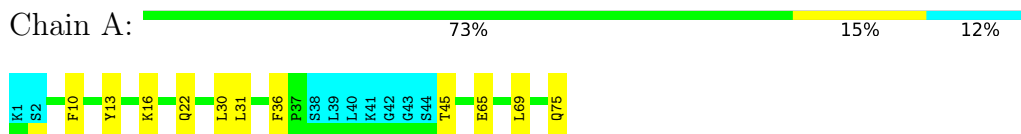
#### 4.2.28 Score per residue for model 28

• Molecule 1: CALBINDIN D9K

Chain A:
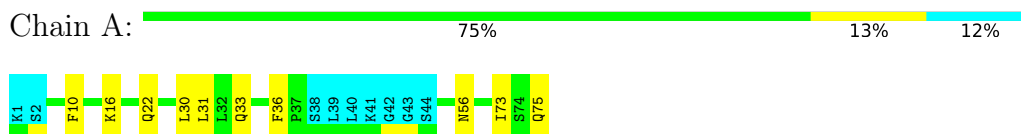


#### 4.2.29 Score per residue for model 29

• Molecule 1: CALBINDIN D9K

Chain A:
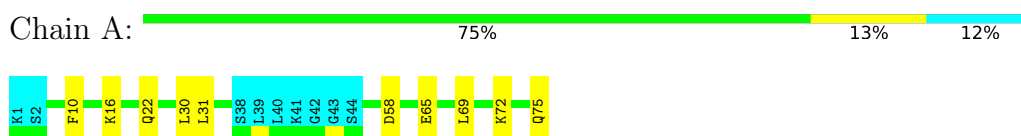


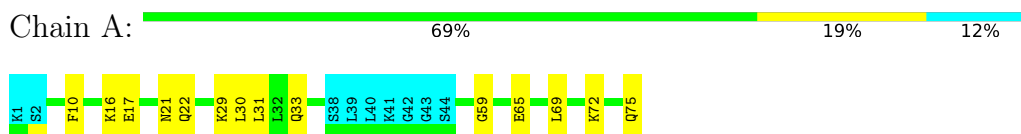#### 4.2.30 Score per residue for model 30

• Molecule 1: CALBINDIN D9K

Chain A:



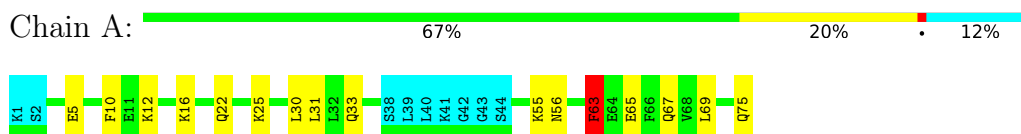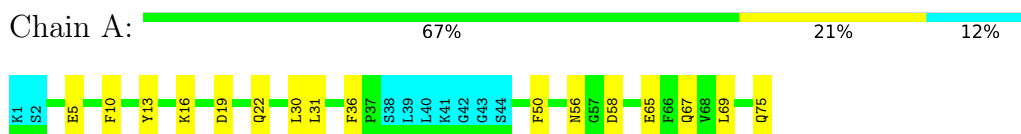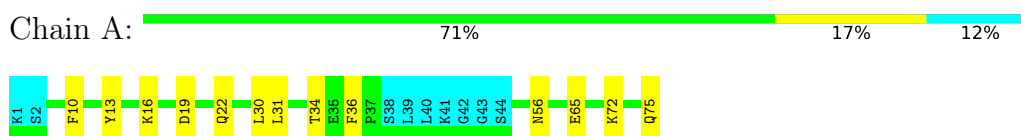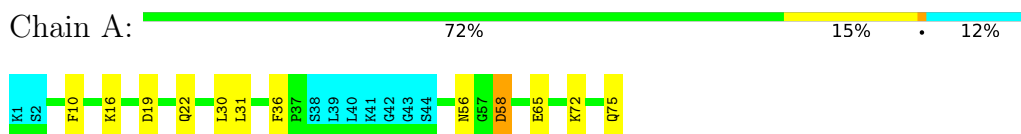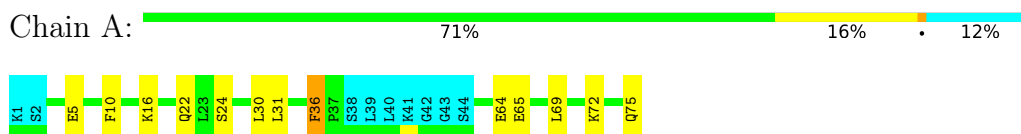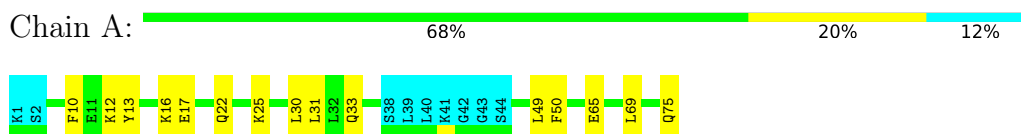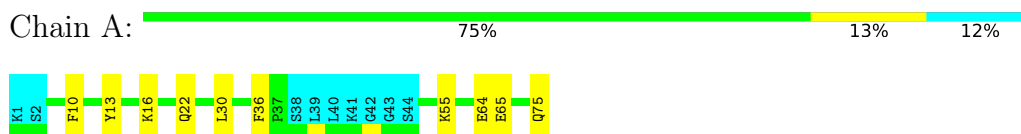#### 4.2.31 Score per residue for model 31

• Molecule 1: CALBINDIN D9K

Chain A:
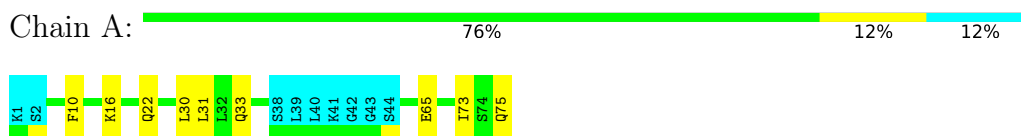


#### 4.2.32 Score per residue for model 32

• Molecule 1: CALBINDIN D9K

Chain A:
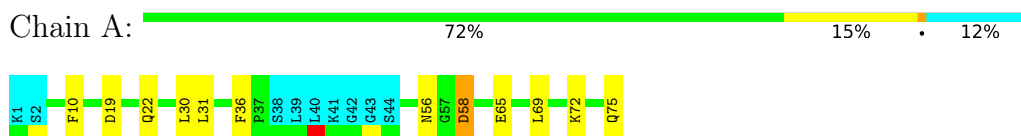
# 5 Refinement protocol and experimental data overview ⓘ

Of the ? calculated structures, 32 were deposited, based on the following criterion: ?.

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

| Software name | Classification | Version |
|---------------|----------------|---------|
| DISGEO | refinement | |
| Amber | refinement | |

No chemical shift data was provided.

# 6 Model quality (i)

## 6.1 Standard geometry (i)

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 | 0.56±0.00 | 0±0/544 ( 0.0± 0.0%) | 0.93±0.03 | 0±0/728 ( 0.0± 0.0%) |
| All | All | 0.56 | 0/17408 ( 0.0%) | 0.93 | 3/23296 ( 0.0%) |

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 | 1.4±0.7 |
| All | All | 0 | 45 |

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 | 63 | PHE | CB-CG-CD2 | -7.30 | 115.69 | 120.80 | 16 | 1 |
| 1 | A | 31 | LEU | CB-CG-CD1 | 6.03 | 121.24 | 111.00 | 1 | 1 |
| 1 | A | 36 | PHE | CB-CG-CD2 | -5.09 | 117.24 | 120.80 | 19 | 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 | 10 | PHE | Sidechain | 32 |
| 1 | A | 13 | TYR | Sidechain | 7 |
| 1 | A | 36 | PHE | Sidechain | 3 |
| 1 | A | 50 | PHE | Sidechain | 2 |
| 1 | A | 63 | PHE | Sidechain | 1 |

## 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 | 536 | 527 | 527 | 0±0 |
| All | All | 17152 | 16864 | 16863 | 9 |

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:13:TYR:CD1 | 1:A:31:LEU:HD12 | 0.56 | 2.35 | 1 | 1 |
| 1:A:22:GLN:HB3 | 1:A:60:GLU:CB | 0.45 | 2.42 | 23 | 1 |
| 1:A:25:LYS:HB2 | 1:A:50:PHE:CG | 0.44 | 2.48 | 6 | 1 |
| 1:A:72:LYS:HA | 1:A:75:GLN:HB2 | 0.43 | 1.89 | 19 | 3 |
| 1:A:36:PHE:N | 1:A:37:PRO:CD | 0.42 | 2.82 | 13 | 3 |

## 6.3 Torsion angles ⓘ

### 6.3.1 Protein backbone ⓘ

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 | 65/75 (87%) | 59±2 (90±4%) | 6±2 (9±3%) | 1±1 (1±1%) | 18 | 66 |
| All | All | 2080/2400 (87%) | 1878 (90%) | 179 (9%) | 23 (1%) | 18 | 66 |

All 7 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 | 19 | ASP | 11 |
| 1 | A | 58 | ASP | 3 |
| 1 | A | 64 | GLU | 3 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | Models (Total) |
|-----|-------|-----|------|----------------|
| 1 | A | 55 | LYS | 2 |
| 1 | A | 59 | GLY | 2 |
| 1 | A | 18 | GLY | 1 |
| 1 | A | 37 | PRO | 1 |

### 6.3.2 Protein sidechains ⓘ

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 | 60/67 (90%) | 50±2 (84±3%) | 10±2 (16±3%) | 5 | 42 |
| All | All | 1920/2144 (90%) | 1613 (84%) | 307 (16%) | 5 | 42 |

All 31 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 | 30 | LEU | 32 |
| 1 | A | 16 | LYS | 31 |
| 1 | A | 22 | GLN | 31 |
| 1 | A | 31 | LEU | 30 |
| 1 | A | 75 | GLN | 29 |
| 1 | A | 65 | GLU | 28 |
| 1 | A | 69 | LEU | 21 |
| 1 | A | 33 | GLN | 13 |
| 1 | A | 36 | PHE | 13 |
| 1 | A | 72 | LYS | 11 |
| 1 | A | 56 | ASN | 11 |
| 1 | A | 25 | LYS | 10 |
| 1 | A | 5 | GLU | 7 |
| 1 | A | 45 | THR | 6 |
| 1 | A | 24 | SER | 4 |
| 1 | A | 17 | GLU | 4 |
| 1 | A | 47 | ASP | 3 |
| 1 | A | 67 | GLN | 3 |
| 1 | A | 12 | LYS | 3 |
| 1 | A | 73 | ILE | 3 |
| 1 | A | 58 | ASP | 3 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | Models (Total) |
|-----|-------|-----|------|----------------|
| 1 | A | 21 | ASN | 2 |
| 1 | A | 74 | SER | 1 |
| 1 | A | 62 | SER | 1 |
| 1 | A | 53 | LEU | 1 |
| 1 | A | 11 | GLU | 1 |
| 1 | A | 29 | LYS | 1 |
| 1 | A | 63 | PHE | 1 |
| 1 | A | 34 | THR | 1 |
| 1 | A | 49 | LEU | 1 |
| 1 | A | 55 | LYS | 1 |

### 6.3.3   RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 6.5   Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 6.6   Ligand geometry ⓘ

There are no ligands in this entry.

## 6.7   Other polymers ⓘ

There are no such molecules in this entry.

## 6.8   Polymer linkage issues ⓘ

There are no chain breaks in this entry.

# 7 Chemical shift validation ⓘ

No chemical shift data were provided