# WORLDWIDE PDB PROTEIN DATA BANK

# wwPDB X-ray Structure Validation Summary Report ⓘ

May 28, 2020 – 09:05 pm BST

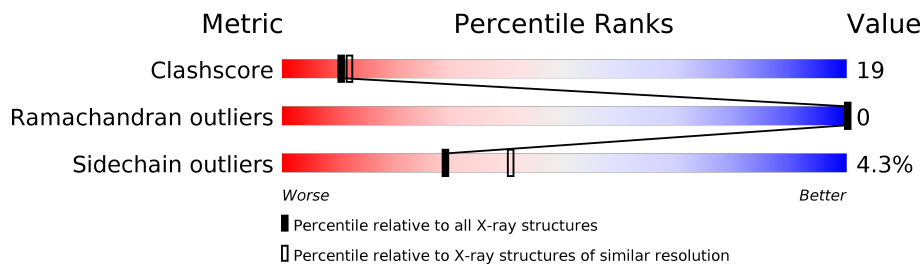| | | |
|---|---|---|
| PDB ID | : | 1YDN |
| Title | : | Crystal Structure of the HMG-CoA Lyase from Brucella melitensis, Northeast Structural Genomics Target LR35. |
| Authors | : | Forouhar, F.; Abashidze, M.; Hussain, M.; Vorobiev, S.M.; Xiao, R.; Ciano, M.; Acton, T.B.; Montelione, G.T.; Tong, L.; Hunt, J.F.; Northeast Structural Genomics Consortium (NESG) |
| Deposited on | : | 2004-12-24 |
| Resolution | : | 2.30 Å(reported) |

# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 Å.

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 | | 19 |
| Ramachandran outliers | | 0 |
| Sidechain outliers | | 4.3% |

*Worse*        *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|--------|--------------------------|-----------------------------------------------------|
| Clashscore | 141614 | 5643 (2.30-2.30) |
| Ramachandran outliers | 138981 | 5575 (2.30-2.30) |
| Sidechain outliers | 138945 | 5575 (2.30-2.30) |

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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%

Note EDS was not executed.

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|------------------|
| 1 | A | 295 | 62%     32% • • |
| 1 | B | 295 | 65%     30% • • |
| 1 | C | 295 | 66%     27% • • |
| 1 | D | 295 | 62%     32% • • |

# 2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 8972 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called HYDROXYMETHYLGLUTARYL-COA LYASE.

| Mol | Chain | Residues | Atoms | | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|----|---------|---------|-------|
| 1 | A | 283 | Total | C | N | O | S | Se | 0 | 0 | 0 |
| | | | 2093 | 1306 | 370 | 405 | 6 | 6 | | | |
| 1 | B | 283 | Total | C | N | O | S | Se | 0 | 0 | 0 |
| | | | 2093 | 1306 | 370 | 405 | 6 | 6 | | | |
| 1 | C | 283 | Total | C | N | O | S | Se | 0 | 0 | 0 |
| | | | 2093 | 1306 | 370 | 405 | 6 | 6 | | | |
| 1 | D | 283 | Total | C | N | O | S | Se | 0 | 0 | 0 |
| | | | 2093 | 1306 | 370 | 405 | 6 | 6 | | | |

There are 56 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------|-----------|
| A | 10 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| A | 64 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| A | 82 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| A | 189 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| A | 256 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| A | 260 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| A | 288 | LEU | - | CLONING ARTIFACT | UNP Q8YEF2 |
| A | 289 | GLU | - | CLONING ARTIFACT | UNP Q8YEF2 |
| A | 290 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| A | 291 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| A | 292 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| A | 293 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| A | 294 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| A | 295 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| B | 10 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| B | 64 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| B | 82 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| B | 189 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| B | 256 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| B | 260 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| B | 288 | LEU | - | CLONING ARTIFACT | UNP Q8YEF2 |

*Continued on next page...*

*Continued from previous page...*

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------|-----------|
| B | 289 | GLU | - | CLONING ARTIFACT | UNP Q8YEF2 |
| B | 290 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| B | 291 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| B | 292 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| B | 293 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| B | 294 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| B | 295 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| C | 10 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| C | 64 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| C | 82 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| C | 189 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| C | 256 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| C | 260 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| C | 288 | LEU | - | CLONING ARTIFACT | UNP Q8YEF2 |
| C | 289 | GLU | - | CLONING ARTIFACT | UNP Q8YEF2 |
| C | 290 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| C | 291 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| C | 292 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| C | 293 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| C | 294 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| C | 295 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| D | 10 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| D | 64 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| D | 82 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| D | 189 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| D | 256 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| D | 260 | MSE | MET | MODIFIED RESIDUE | UNP Q8YEF2 |
| D | 288 | LEU | - | CLONING ARTIFACT | UNP Q8YEF2 |
| D | 289 | GLU | - | CLONING ARTIFACT | UNP Q8YEF2 |
| D | 290 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| D | 291 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| D | 292 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| D | 293 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| D | 294 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |
| D | 295 | HIS | - | EXPRESSION TAG | UNP Q8YEF2 |

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---------|---------|
| 2 | B | 1 | Total 1 | Ca 1 | 0 | 0 |
| 2 | A | 1 | Total 1 | Ca 1 | 0 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---------|---------|
| 2 | D | 1 | Total<br>1 | Ca<br>1 | 0 | 0 |
| 2 | C | 1 | Total<br>1 | Ca<br>1 | 0 | 0 |

- Molecule 3 is water.

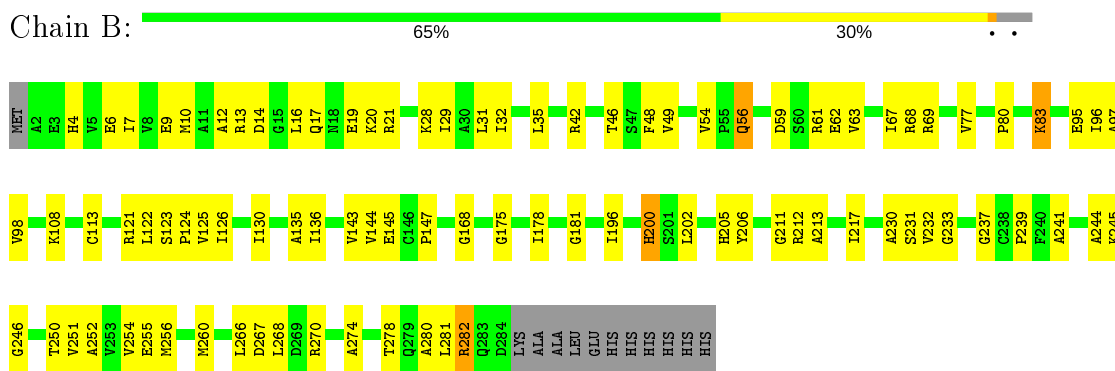| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---------|---------|
| 3 | A | 197 | Total<br>197 | O<br>197 | 0 | 0 |
| 3 | B | 131 | Total<br>131 | O<br>131 | 0 | 0 |
| 3 | C | 143 | Total<br>143 | O<br>143 | 0 | 0 |
| 3 | D | 125 | Total<br>125 | O<br>125 | 0 | 0 |

# 3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-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 outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.
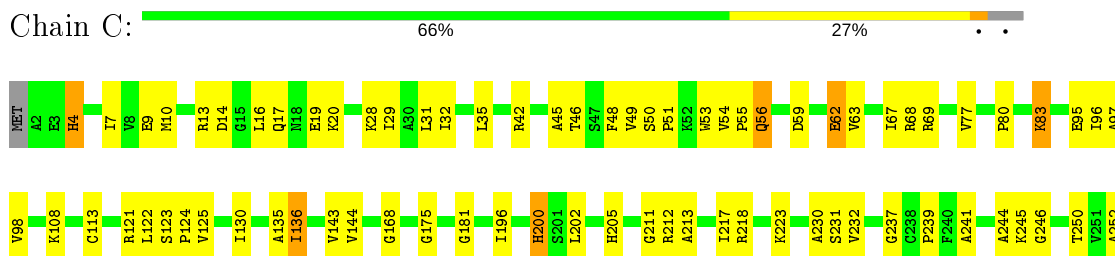
Note EDS was not executed.

- Molecule 1: HYDROXYMETHYLGLUTARYL-COA LYASE
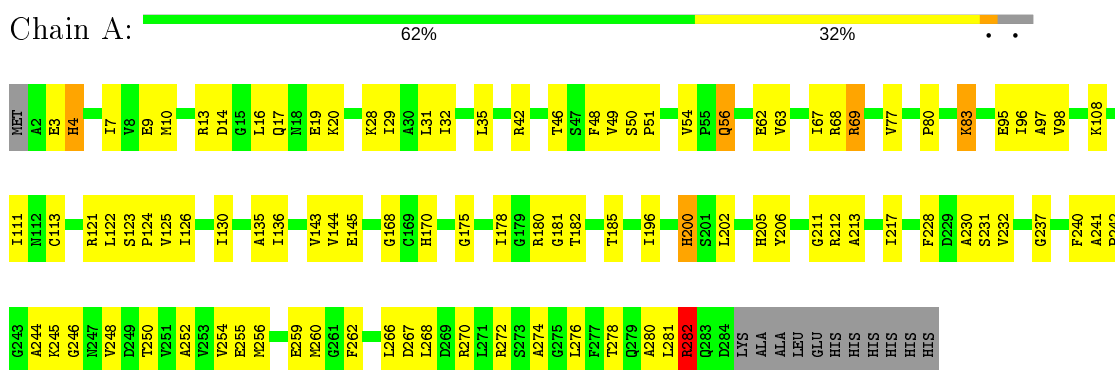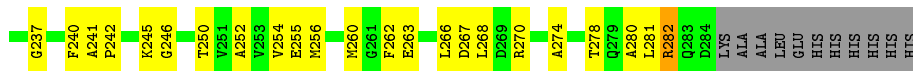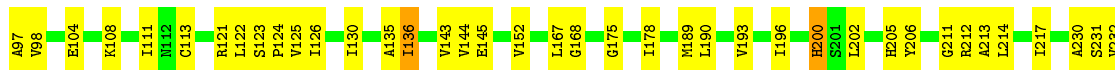


- Molecule 1: HYDROXYMETHYLGLUTARYL-COA LYASE



- Molecule 1: HYDROXYMETHYLGLUTARYL-COA LYASE

V253 V254 E255 M256 M260 L266 D267 L268 D269 R270 A274 T278 Q279 A280 L281 R282 Q283 D284 LYS ALA ALA LEU GLU HIS HIS HIS HIS HIS HIS

• Molecule 1: HYDROXYMETHYLGLUTARYL-COA LYASE

Chain D:

62%                          32%                       • •

MET A2 E3 H4 V5 E6 I7 V8 E9 M10 R13 D14 G15 L16 Q17 N18 E19 K28 I29 A30 L31 I32 L35 R42 T46 S47 F48 V49 S50 P51 V54 P55 Q56 D59 S60 R61 E62 V63 I67 R68 R69 R74 V77 P80 K83 E95 I96

A97 V98 E104 K108 I111 N112 C113 R121 L122 S123 P124 V125 I126 I130 A135 I136 V143 V144 E145 V152 L167 G168 G175 I178 M189 L190 V193 I196 H200 S201 L202 H205 Y206 G211 R212 A213 L214 I217 A230 S231 V232

G237 F240 A241 P242 K245 G246 T250 V251 A252 V253 V254 E255 M256 M260 G261 F262 E263 L266 D267 L268 D269 R270 A274 T278 Q279 A280 L281 R282 Q283 D284 LYS ALA ALA LEU GLU HIS HIS HIS HIS HIS HIS

# 4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

| Property | Value | Source |
|---|---|---|
| Space group | P 1 21 1 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 82.27Å   86.40Å   87.68Å<br>90.00°   118.70°   90.00° | Depositor |
| Resolution (Å) | 28.94  –  2.30 | Depositor |
| % Data completeness<br>(in resolution range) | 94.5 (28.94-2.30) | Depositor |
| $R_{merge}$ | 0.10 | Depositor |
| $R_{sym}$ | 0.10 | Depositor |
| Refinement program | CNS 1.1 | Depositor |
| R, $R_{free}$ | 0.271 , 0.304 | Depositor |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| Total number of atoms | 8972 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 30.0 | wwPDB-VP |

# 5 Model quality (i)

## 5.1 Standard geometry (i)

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

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 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 | 0.40 | 0/2122 | 1.00 | 6/2871 (0.2%) |
| 1 | B | 0.38 | 0/2122 | 0.61 | 4/2871 (0.1%) |
| 1 | C | 0.37 | 0/2122 | 0.62 | 4/2871 (0.1%) |
| 1 | D | 0.37 | 0/2122 | 0.62 | 3/2871 (0.1%) |
| All | All | 0.38 | 0/8488 | 0.73 | 17/11484 (0.1%) |

There are no bond length outliers.

The worst 5 of 17 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|-----|-------|-----|------|-------|---|-----------|---------|
| 1 | A | 69 | ARG | NE-CZ-NH1 | -23.89 | 108.36 | 120.30 |
| 1 | A | 69 | ARG | NE-CZ-NH2 | 21.20 | 130.90 | 120.30 |
| 1 | A | 282 | ARG | NE-CZ-NH2 | -19.27 | 110.66 | 120.30 |
| 1 | A | 282 | ARG | NE-CZ-NH1 | 19.09 | 129.84 | 120.30 |
| 1 | A | 69 | ARG | CD-NE-CZ | 9.99 | 137.59 | 123.60 |

There are no chirality outliers.

There are no planarity outliers.

## 5.2 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 the 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 within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 1 | A | 2093 | 0 | 2071 | 81 | 0 |
| 1 | B | 2093 | 0 | 2071 | 78 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 1 | C | 2093 | 0 | 2071 | 78 | 0 |
| 1 | D | 2093 | 0 | 2071 | 84 | 0 |
| 2 | A | 1 | 0 | 0 | 0 | 0 |
| 2 | B | 1 | 0 | 0 | 0 | 0 |
| 2 | C | 1 | 0 | 0 | 0 | 0 |
| 2 | D | 1 | 0 | 0 | 0 | 0 |
| 3 | A | 197 | 0 | 0 | 12 | 0 |
| 3 | B | 131 | 0 | 0 | 10 | 0 |
| 3 | C | 143 | 0 | 0 | 10 | 0 |
| 3 | D | 125 | 0 | 0 | 15 | 0 |
| All | All | 8972 | 0 | 8284 | 309 | 0 |

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 19.

The worst 5 of 309 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

| Atom-1 | Atom-2 | Interatomic distance (Å) | Clash overlap (Å) |
|--------|--------|--------------------------|-------------------|
| 1:A:7:ILE:HD11 | 1:A:230:ALA:HB3 | 1.59 | 0.85 |
| 1:B:7:ILE:HD11 | 1:B:230:ALA:HB3 | 1.59 | 0.85 |
| 1:C:7:ILE:HD11 | 1:C:230:ALA:HB3 | 1.60 | 0.84 |
| 1:A:259:GLU:HA | 3:A:674:HOH:O | 1.78 | 0.82 |
| 1:D:7:ILE:HD11 | 1:D:230:ALA:HB3 | 1.60 | 0.81 |

There are no symmetry-related clashes.

## 5.3  Torsion angles (i)

### 5.3.1  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 X-ray entries followed by that with respect to entries of similar resolution.

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 | 281/295 (95%) | 264 (94%) | 17 (6%) | 0 | 100 | 100 |
| 1 | B | 281/295 (95%) | 265 (94%) | 16 (6%) | 0 | 100 | 100 |

*Continued from previous page...*

| Mol | Chain | Analysed | Favoured | Allowed | Outliers | Percentiles | |
|-----|-------|----------|----------|---------|----------|-------------|---|
| 1 | C | 281/295 (95%) | 264 (94%) | 17 (6%) | 0 | 100 | 100 |
| 1 | D | 281/295 (95%) | 264 (94%) | 17 (6%) | 0 | 100 | 100 |
| All | All | 1124/1180 (95%) | 1057 (94%) | 67 (6%) | 0 | 100 | 100 |

There are no Ramachandran outliers to report.

### 5.3.2 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 X-ray entries followed by that with respect to entries of similar resolution.

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 | 217/221 (98%) | 206 (95%) | 11 (5%) | 24 | 33 |
| 1 | B | 217/221 (98%) | 209 (96%) | 8 (4%) | 34 | 48 |
| 1 | C | 217/221 (98%) | 208 (96%) | 9 (4%) | 30 | 43 |
| 1 | D | 217/221 (98%) | 208 (96%) | 9 (4%) | 30 | 43 |
| All | All | 868/884 (98%) | 831 (96%) | 37 (4%) | 29 | 40 |

5 of 37 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | B | 200 | HIS |
| 1 | C | 56 | GLN |
| 1 | D | 200 | HIS |
| 1 | B | 202 | LEU |
| 1 | B | 245 | LYS |

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | B | 283 | GLN |
| 1 | C | 56 | GLN |
| 1 | D | 110 | ASN |
| 1 | B | 200 | HIS |
| 1 | D | 112 | ASN |

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

# 6 Fit of model and data ⓘ

## 6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

## 6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

## 6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

## 6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.