# WORLDWIDE wwPDB PROTEIN DATA BANK

# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 2, 2021 – 08:55 PM EDT

|  |  |  |
|---|---|---|
| PDB ID | : | 1IV3 |
| Title | : | Structure of 2C-Methyl-D-erythritol-2,4-cyclodiphosphate Synthase (bound form MG atoms) |
| Authors | : | Kishida, H.; Wada, T.; Unzai, S.; Kuzuyama, T.; Terada, T.; Sirouzu, M.; Yokoyama, S.; Tame, J.R.H.; Park, S.-Y.; RIKEN Structural Genomics/Proteomics Initiative (RSGI) |
| Deposited on | : | 2002-03-11 |
| Resolution | : | 1.52 Å(reported) |

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

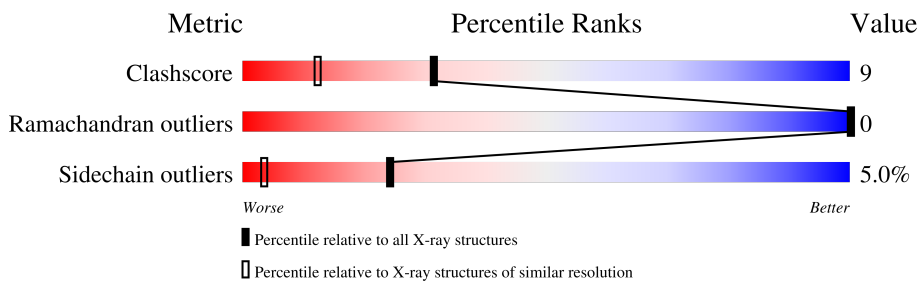|  |  |  |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Xtriage (Phenix) | : | NOT EXECUTED |
| EDS | : | NOT EXECUTED |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.23.2 |

# 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 1.52 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 | | 9 |
| Ramachandran outliers | | 0 |
| Sidechain outliers | | 5.0% |

*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 | 4249 (1.54-1.50) |
| Ramachandran outliers | 138981 | 4148 (1.54-1.50) |
| Sidechain outliers | 138945 | 4146 (1.54-1.50) |

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 of 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 | 152 | 80%    18%    ·· |
| 1 | B | 152 | 80%    16%    ·· |
| 1 | C | 152 | 74%    20%    · ·· |
| 1 | D | 152 | 76%    20%    ·· |
| 1 | E | 152 | 74%    18%    5% ·· |
| 1 | F | 152 | 80%    16%    ·· |

# 2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 7475 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 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A | 150 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1152 | 726 | 213 | 209 | 4 | | | |
| 1 | B | 150 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1152 | 726 | 213 | 209 | 4 | | | |
| 1 | C | 150 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1152 | 726 | 213 | 209 | 4 | | | |
| 1 | D | 150 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1152 | 726 | 213 | 209 | 4 | | | |
| 1 | E | 150 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1152 | 726 | 213 | 209 | 4 | | | |
| 1 | F | 150 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1152 | 726 | 213 | 209 | 4 | | | |

There are 18 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| A | 41 | MET | LEU | engineered mutation | UNP Q8RQP5 |
| A | 81 | MET | LEU | engineered mutation | UNP Q8RQP5 |
| A | 120 | MET | LEU | engineered mutation | UNP Q8RQP5 |
| B | 241 | MET | LEU | engineered mutation | UNP Q8RQP5 |
| B | 281 | MET | LEU | engineered mutation | UNP Q8RQP5 |
| B | 320 | MET | LEU | engineered mutation | UNP Q8RQP5 |
| C | 441 | MET | LEU | engineered mutation | UNP Q8RQP5 |
| C | 481 | MET | LEU | engineered mutation | UNP Q8RQP5 |
| C | 520 | MET | LEU | engineered mutation | UNP Q8RQP5 |
| D | 1041 | MET | LEU | engineered mutation | UNP Q8RQP5 |
| D | 1081 | MET | LEU | engineered mutation | UNP Q8RQP5 |
| D | 1120 | MET | LEU | engineered mutation | UNP Q8RQP5 |
| E | 1241 | MET | LEU | engineered mutation | UNP Q8RQP5 |
| E | 1281 | MET | LEU | engineered mutation | UNP Q8RQP5 |
| E | 1320 | MET | LEU | engineered mutation | UNP Q8RQP5 |
| F | 1441 | MET | LEU | engineered mutation | UNP Q8RQP5 |
| F | 1481 | MET | LEU | engineered mutation | UNP Q8RQP5 |

*Continued from previous page...*

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------|-----------|
| F | 1520 | MET | LEU | engineered mutation | UNP Q8RQP5 |

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---------|---------|
| 2 | A | 1 | Total 1 | Mg 1 | 0 | 0 |
| 2 | B | 1 | Total 1 | Mg 1 | 0 | 0 |
| 2 | C | 1 | Total 1 | Mg 1 | 0 | 0 |
| 2 | D | 1 | Total 1 | Mg 1 | 0 | 0 |
| 2 | E | 1 | Total 1 | Mg 1 | 0 | 0 |
| 2 | F | 1 | Total 1 | Mg 1 | 0 | 0 |

- Molecule 3 is water.

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---------|---------|
| 3 | A | 112 | Total 112 | O 112 | 0 | 0 |
| 3 | B | 94 | Total 94 | O 94 | 0 | 0 |
| 3 | C | 90 | Total 90 | O 90 | 0 | 0 |
| 3 | D | 87 | Total 87 | O 87 | 0 | 0 |
| 3 | E | 90 | Total 90 | O 90 | 0 | 0 |
| 3 | F | 84 | Total 84 | O 84 | 0 | 0 |

# 3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase

Chain A:



• Molecule 1: 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase

Chain B:



• Molecule 1: 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase

Chain C:



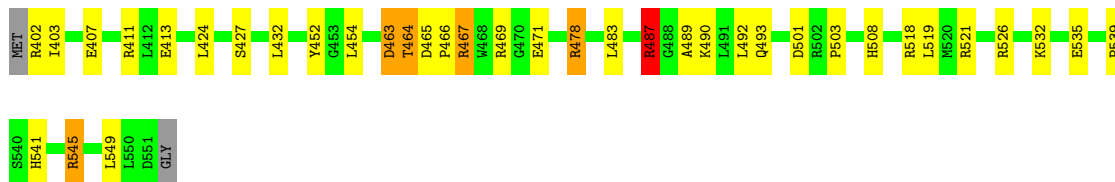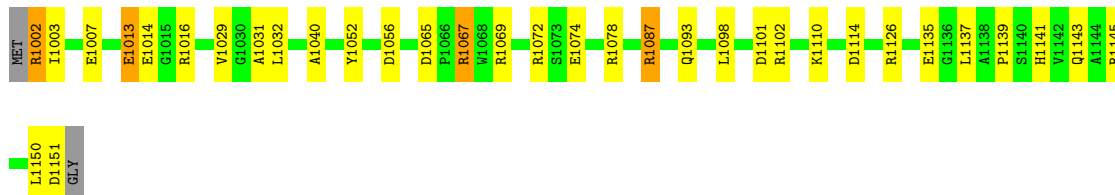• Molecule 1: 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase

Chain D:



• Molecule 1: 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase

Chain E:

MET R1202 I1203 E1207 D1208 S1209 H1210 R1211 L1212 E1213 E1214 G1215 R1216 L1224 V1229 G1230 A1231 L1232 M1241 D1265 P1266 R1267 W1268 R1269 R1272 S1273 E1274 R1278 R1282 E1285 G1288 A1289 K1290 Q1293 R1302 P1303 H1308 S1315 L1319 M1320 R1321 R1326 E1335

H1341 V1342 Q1343 A1344 R1345 L1350 D1351 GLY

• Molecule 1: 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase

Chain F:

80%　　　16%

MET R1402 E1407 D1408 S1409 H1410 R1411 L1412 E1413 R1416 A1431 L1432 A1433 H1434 G1437 D1463 E1474 R1478 K1490 Q1493 D1501 R1518 R1526 T1530 E1535 P1539 S1540 H1541 R1545 A1546 V1547 L1550 D1551 GLY

# 4 Data and refinement statistics ⓘ

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

| Property | Value | Source |
|----------|-------|--------|
| Space group | P 41 21 2 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 106.18Å    106.18Å    148.81Å<br>90.00°    90.00°    90.00° | Depositor |
| Resolution (Å) | 20.00 – 1.52 | Depositor |
| % Data completeness<br>(in resolution range) | 93.5 (20.00-1.52) | Depositor |
| $R_{merge}$ | 0.05 | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| Refinement program | SHELXL-97 | Depositor |
| R, $R_{free}$ | 0.202 , 0.257 | Depositor |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| Total number of atoms | 7475 | wwPDB-VP |
| Average B, all atoms (Å²) | 20.0 | wwPDB-VP |

# 5 Model quality ⓘ

## 5.1 Standard geometry ⓘ

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

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     | 0.39 | 0/1173  | 1.15 | 7/1587 (0.4%)  |
| 1   | B     | 0.39 | 0/1173  | 1.12 | 6/1587 (0.4%)  |
| 1   | C     | 0.41 | 0/1173  | 1.25 | 8/1587 (0.5%)  |
| 1   | D     | 0.39 | 0/1173  | 1.17 | 10/1587 (0.6%) |
| 1   | E     | 0.40 | 0/1173  | 1.24 | 13/1587 (0.8%) |
| 1   | F     | 0.38 | 0/1173  | 1.14 | 4/1587 (0.3%)  |
| All | All   | 0.39 | 0/7038  | 1.18 | 48/9522 (0.5%) |

There are no bond length outliers.

The worst 5 of 48 bond angle outliers are listed below:

| Mol | Chain | Res  | Type | Atoms      | Z      | Observed($^o$) | Ideal($^o$) |
|-----|-------|------|------|------------|--------|----------------|-------------|
| 1   | D     | 1072 | ARG  | NE-CZ-NH2  | -13.71 | 113.45         | 120.30      |
| 1   | C     | 487  | ARG  | CD-NE-CZ   | 13.30  | 142.23         | 123.60      |
| 1   | E     | 1272 | ARG  | NE-CZ-NH2  | 9.34   | 124.97         | 120.30      |
| 1   | C     | 518  | ARG  | NE-CZ-NH2  | -9.04  | 115.78         | 120.30      |
| 1   | C     | 487  | ARG  | NE-CZ-NH1  | 8.99   | 124.80         | 120.30      |

There are no chirality outliers.

There are no planarity outliers.

## 5.2 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 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 | 1152 | 0 | 1182 | 24 | 0 |
| 1 | B | 1152 | 0 | 1182 | 18 | 0 |
| 1 | C | 1152 | 0 | 1182 | 28 | 0 |
| 1 | D | 1152 | 0 | 1182 | 30 | 0 |
| 1 | E | 1152 | 0 | 1182 | 31 | 0 |
| 1 | F | 1152 | 0 | 1180 | 21 | 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 |
| 2 | E | 1 | 0 | 0 | 0 | 0 |
| 2 | F | 1 | 0 | 0 | 0 | 0 |
| 3 | A | 112 | 0 | 0 | 1 | 0 |
| 3 | B | 94 | 0 | 0 | 1 | 0 |
| 3 | C | 90 | 0 | 0 | 1 | 0 |
| 3 | D | 87 | 0 | 0 | 1 | 0 |
| 3 | E | 90 | 0 | 0 | 4 | 0 |
| 3 | F | 84 | 0 | 0 | 1 | 0 |
| All | All | 7475 | 0 | 7090 | 126 | 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 9.

The worst 5 of 126 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:E:1216:ARG:NH2 | 1:E:1232:LEU:HD12 | 1.63 | 1.13 |
| 1:E:1216:ARG:NH2 | 1:E:1232:LEU:CD1 | 2.28 | 0.97 |
| 1:E:1216:ARG:HH21 | 1:E:1232:LEU:HD12 | 1.28 | 0.90 |
| 1:C:483:LEU:HD22 | 1:C:487:ARG:HH22 | 1.39 | 0.86 |
| 1:D:1065:ASP:OD1 | 1:D:1067:ARG:HG3 | 1.84 | 0.78 |

There are no symmetry-related clashes.

## 5.3   Torsion angles ⓘ

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

WORLDWIDE
PDB
PROTEIN DATA BANK

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 | 148/152 (97%) | 145 (98%) | 3 (2%) | 0 | 100 | 100 |
| 1 | B | 148/152 (97%) | 147 (99%) | 1 (1%) | 0 | 100 | 100 |
| 1 | C | 148/152 (97%) | 143 (97%) | 5 (3%) | 0 | 100 | 100 |
| 1 | D | 148/152 (97%) | 145 (98%) | 3 (2%) | 0 | 100 | 100 |
| 1 | E | 148/152 (97%) | 146 (99%) | 2 (1%) | 0 | 100 | 100 |
| 1 | F | 148/152 (97%) | 146 (99%) | 2 (1%) | 0 | 100 | 100 |
| All | All | 888/912 (97%) | 872 (98%) | 16 (2%) | 0 | 100 | 100 |

There are no Ramachandran outliers to report.

## 5.3.2 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 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 | 121/122 (99%) | 115 (95%) | 6 (5%) | 24 | 3 |
| 1 | B | 121/122 (99%) | 114 (94%) | 7 (6%) | 20 | 2 |
| 1 | C | 121/122 (99%) | 113 (93%) | 8 (7%) | 16 | 1 |
| 1 | D | 121/122 (99%) | 119 (98%) | 2 (2%) | 60 | 32 |
| 1 | E | 121/122 (99%) | 114 (94%) | 7 (6%) | 20 | 2 |
| 1 | F | 121/122 (99%) | 115 (95%) | 6 (5%) | 24 | 3 |
| All | All | 726/732 (99%) | 690 (95%) | 36 (5%) | 24 | 3 |

5 of 36 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | E | 1272 | ARG |
| 1 | F | 1490 | LYS |
| 1 | E | 1290 | LYS |
| 1 | F | 1432 | LEU |
| 1 | B | 292 | LEU |

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

| Mol | Chain | Res | Type |
|-----|-------|------|------|
| 1 | E | 1341 | HIS |
| 1 | F | 1541 | HIS |
| 1 | F | 1543 | GLN |
| 1 | F | 1493 | GLN |
| 1 | C | 543 | GLN |

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

Of 6 ligands modelled in this entry, 6 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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

## 5.8 Polymer linkage issues (i)

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.