# W O R L D W I D E
# wwPDB

## PROTEIN DATA BANK

# wwPDB X-ray Structure Validation Summary Report (i)

### Feb 27, 2014 – 01:24 PM GMT

|  |  |  |
|---|---|---|
| PDB ID | : | 2DPN |
| Title | : | Crystal Structure of the glycerol kinase from Thermus thermophilus HB8 |
| Authors | : | Asada, Y.; Sugahara, M.; Kunishima, N.; RIKEN Structural Genomics/Proteomics Initiative (RSGI) |
| Deposited on | : | 2006-05-12 |
| Resolution | : | 2.80 Å(reported) |

# 1 Overall quality at a glance ⓘ

The reported resolution of this entry is 2.80 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 |
|--------|------------------|-------|
| Rfree | | 0.266 |
| Clashscore | | 29 |
| Ramachandran outliers | | 2.6% |
| Sidechain outliers | | 7.3% |
| RSRZ outliers | | 2.2% |

Worse                                                                     Better

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
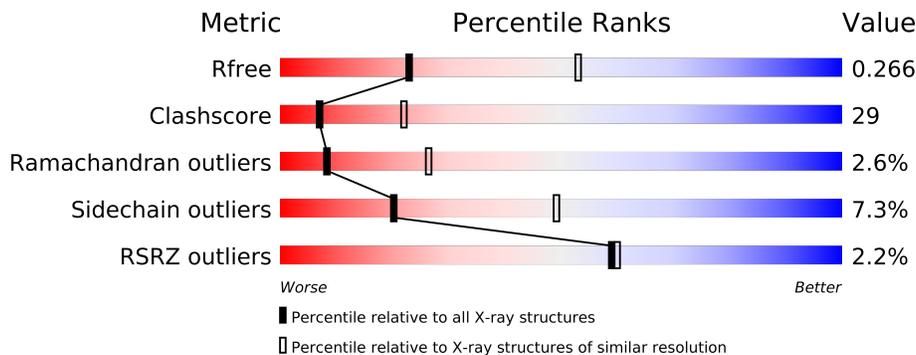
| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|--------|--------------------------|----------------------------------------------------|
| $R_{free}$ | 66092 | 1799 (2.80-2.80) |
| Clashscore | 79885 | 2295 (2.80-2.80) |
| Ramachandran outliers | 78287 | 2252 (2.80-2.80) |
| Sidechain outliers | 78261 | 2254 (2.80-2.80) |
| RSRZ outliers | 66119 | 1802 (2.80-2.80) |

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|------------------|
| 1 | A | 495 | |
| 1 | B | 495 | |

## 2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 7888 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Glycerol kinase.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---|---------|---------|-------|
| 1 | A | 492 | Total 3753 | C 2400 | N 664 | O 683 | S 6 | 4 | 0 | 0 |
| 1 | B | 492 | Total 3753 | C 2400 | N 664 | O 683 | S 6 | 4 | 0 | 0 |

- Molecule 2 is water.

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---------|---------|
| 2 | A | 243 | Total 243 | O 243 | 0 | 0 |
| 2 | B | 139 | Total 139 | O 139 | 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 errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). 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.

- Molecule 1: Glycerol kinase

Chain A:



- Molecule 1: Glycerol kinase

Chain B:

# 4 Data and refinement statistics (i)

| Property | Value | Source |
|---|---|---|
| Space group | P 21 21 21 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 67.99Å 90.51Å 182.99Å<br>90.00° 90.00° 90.00° | Depositor |
| Resolution (Å) | 29.77 – 2.80<br>29.77 – 2.80 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 99.1 (29.77-2.80)<br>99.2 (29.77-2.80) | Depositor<br>EDS |
| $R_{merge}$ | 0.10 | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 3.37 (at 2.80Å) | Xtriage |
| Refinement program | CNS 1.1 | Depositor |
| R, $R_{free}$ | 0.213 , 0.266<br>0.213 , 0.266 | Depositor<br>DCC |
| $R_{free}$ test set | 1425 reflections (5.04%) | DCC |
| Wilson B-factor (Å$^2$) | 42.5 | Xtriage |
| Anisotropy | 0.512 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.28 , 26.9 | EDS |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| L-test for twinning | $< |L| > = 0.48, < L^2 > = 0.31$ | Xtriage |
| Outliers | 0 of 28258 reflections | Xtriage |
| $F_o$,$F_c$ correlation | 0.92 | EDS |
| Total number of atoms | 7888 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 42.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.40% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

# 5   Model quality ⓘ

## 5.1   Standard geometry ⓘ

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 | 0.46 | 1/3837 (0.0%) | 0.72 | 0/5221 |
| 1 | B | 0.47 | 1/3837 (0.0%) | 0.72 | 0/5221 |
| All | All | 0.47 | 2/7674 (0.0%) | 0.72 | 0/10442 |

All (2) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|-------|-----|-------------|----------|
| 1 | A | 352 | TRP | NE1-CE2 | 8.70 | 1.48 | 1.37 |
| 1 | B | 352 | TRP | NE1-CE2 | 8.47 | 1.48 | 1.37 |

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## 5.2   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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 1 | A | 3753 | 0 | 3786 | 143 | 0 |
| 1 | B | 3753 | 0 | 3786 | 295 | 0 |
| 2 | A | 243 | 0 | 0 | 8 | 0 |
| 2 | B | 139 | 0 | 0 | 9 | 0 |
| All | All | 7888 | 0 | 7572 | 431 | 0 |

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including

hydrogens) of the entry. The overall clashscore for this entry is 29.

The worst 5 of 431 close contacts within the same asymmetric unit are listed below.

| Atom-1 | Atom-2 | Distance(Å) | Clash(Å) |
|---|---|---|---|
| 1:B:320:ARG:HG3 | 1:B:321:GLU:H | 1.18 | 1.07 |
| 1:B:47:PRO:HG3 | 1:B:85:LEU:HD21 | 1.35 | 1.06 |
| 1:B:186:ARG:HB3 | 1:B:186:ARG:HH11 | 1.26 | 1.00 |
| 1:A:86:LEU:HD12 | 1:A:142:LEU:HD13 | 1.43 | 0.98 |
| 1:A:17:ILE:HD12 | 1:A:28:VAL:HG22 | 1.42 | 0.96 |

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. 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 | 490/495 (99%) | 452 (92%) | 26 (5%) | 12 (2%) | 9 | 29 |
| 1 | B | 490/495 (99%) | 424 (86%) | 53 (11%) | 13 (3%) | 8 | 25 |
| All | All | 980/990 (99%) | 876 (89%) | 79 (8%) | 25 (3%) | 8 | 26 |

5 of 25 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 147 | PRO |
| 1 | A | 393 | GLU |
| 1 | A | 397 | VAL |
| 1 | B | 493 | ARG |
| 1 | A | 96 | HIS |

### 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 | 371/372 (100%) | 347 (94%) | 24 (6%) | 24 | 57 |
| 1 | B | 371/372 (100%) | 341 (92%) | 30 (8%) | 17 | 43 |
| All | All | 742/744 (100%) | 688 (93%) | 54 (7%) | 20 | 49 |

5 of 54 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | B | 9 | GLN |
| 1 | B | 145 | ASN |
| 1 | B | 392 | GLU |
| 1 | B | 53 | THR |
| 1 | B | 73 | LEU |

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | A | 411 | ASN |

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

There are no ligands in this entry.

## 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 ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, $95^{th}$ percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

| Mol | Chain | Analysed | <RSRZ> | #RSRZ>2 | | | OWAB(Å$^2$) | Q<0.9 |
|-----|-------|----------|--------|---------|------|------|-------------|-------|
| 1 | A | 492/495 (99%) | -0.54 | 2 (0%) | 90 | 91 | 13, 29, 49, 77 | 1 (0%) |
| 1 | B | 492/495 (99%) | 0.05 | 20 (4%) | 35 | 36 | 16, 57, 82, 88 | 1 (0%) |
| All | All | 984/990 (99%) | -0.25 | 22 (2%) | 59 | 60 | 13, 35, 78, 88 | 2 (0%) |

The worst 5 of 22 RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|-----|-------|-----|------|------|
| 1 | B | 76 | GLY | 3.4 |
| 1 | B | 83 | THR | 3.3 |
| 1 | B | 147 | PRO | 3.3 |
| 1 | B | 161 | GLY | 3.1 |
| 1 | B | 81 | ARG | 2.8 |

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

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

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.