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

Oct 17, 2023 – 12:34 AM EDT

| | | |
|---|---|---|
| PDB ID | : | 1IT7 |
| Title | : | Crystal structure of archaeosine tRNA-guanine transglycosylase complexed with guanine |
| Authors | : | Ishitani, R.; Nureki, O.; Fukai, S.; Kijimoto, T.; Nameki, N.; Watanabe, M.; Kondo, H.; Sekine, M.; Okada, N.; Nishimura, S.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI) |
| Deposited on | : | 2002-01-11 |
| Resolution | : | 2.30 Å(reported) |

This is a wwPDB X-ray Structure Validation Summary 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/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

---

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

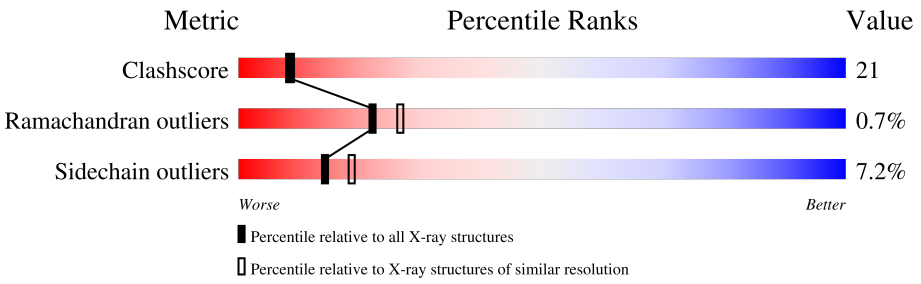
| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Mogul | : | 1.8.5 (274361), CSD as541be (2020) |
| Xtriage (Phenix) | : | NOT EXECUTED |
| EDS | : | NOT EXECUTED |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.36 |

# 1 Overall quality at a glance (i)

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 | 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 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 | 582 | 61% 33% 5% • |
| 1 | B | 582 | 59% 37% • • |

# 2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 9503 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 Archaeosine tRNA-guanine transglycosylase.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | | |
| 1 | A | 577 | 4652 | 2970 | 815 | 848 | 19 | 0 | 0 | 0 |
| 1 | B | 577 | 4652 | 2970 | 815 | 848 | 19 | 0 | 0 | 0 |

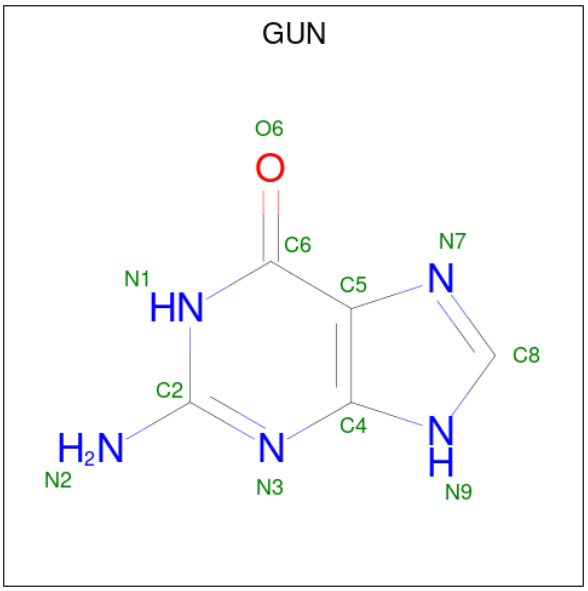
- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| | | | Total | Mg | | |
| 3 | A | 1 | 1 | 1 | 0 | 0 |
| 3 | B | 1 | 1 | 1 | 0 | 0 |

- Molecule 4 is GUANINE (three-letter code: GUN) (formula: $C_5H_5N_5O$).




| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | | |
| 4 | A | 1 | 11 | 5 | 5 | 1 | 0 | 0 |

- Molecule 5 is water.

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| | | | Total | O | | |
| 5 | A | 105 | 105 | 105 | 0 | 0 |
| | | | Total | O | | |
| 5 | B | 79 | 79 | 79 | 0 | 0 |

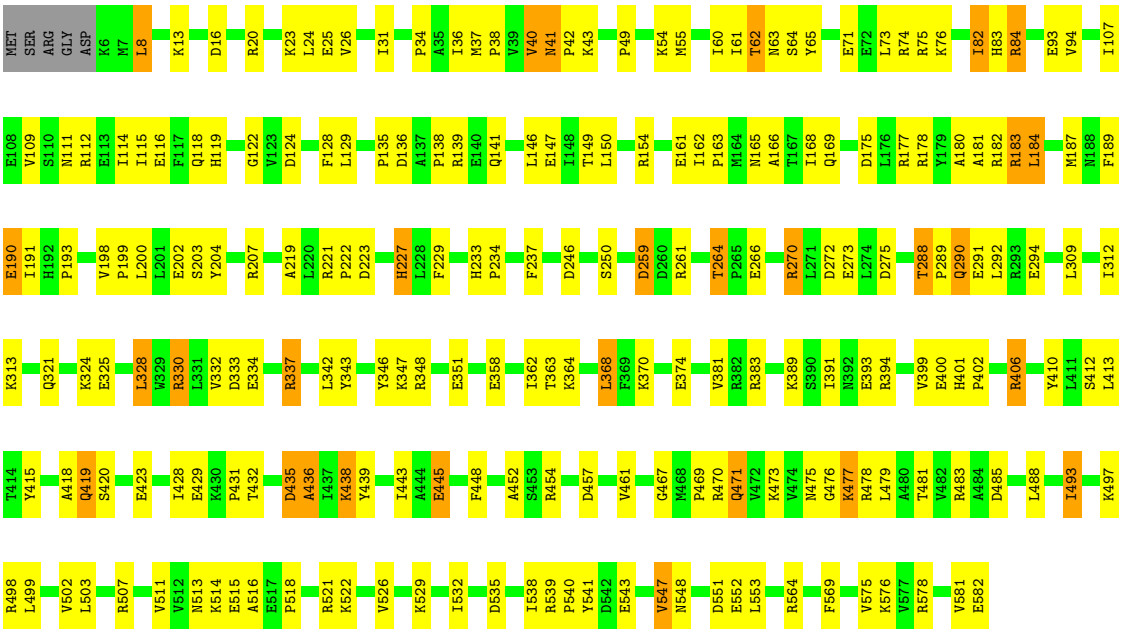
# 3 Residue-property plots (i)

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.

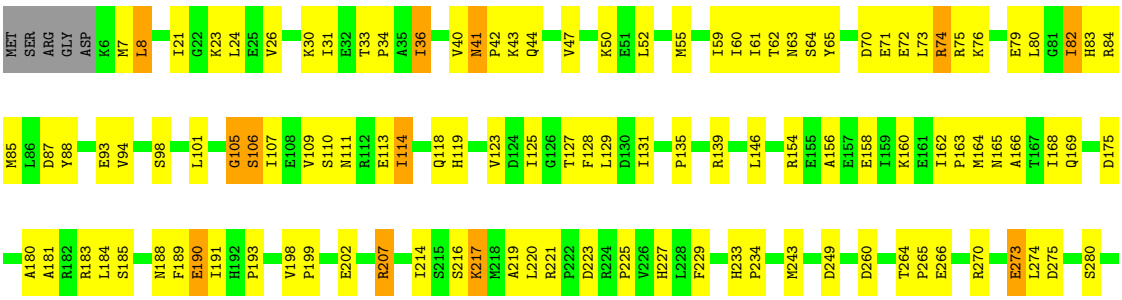
• Molecule 1: Archaeosine tRNA-guanine transglycosylase

Chain A: 61% 33% 5% •




• Molecule 1: Archaeosine tRNA-guanine transglycosylase

Chain B: 59% 37% • •



S285
T288
P289
Q290
E291
L292
R293
E294
M295
P296
K297
R300
T301
R302
L303
I316
Q321
K324
E325
L328
W329
R330
E334
R335
H339
P340
K341
L342
Y343
S344
K347
F355
F359
I362
K365
L368
F369
K370
N373
L376
R377
V381
E386
K389
S390
I391
R394
F395
G396
E397
R406
R409
Y410
L413
T414
Y415
P416
Q419
S420
I428
E429
K430
P431
E434
D435
A436
I437
K438
Y439
I443
F448
A452
A459
K460
V461
T466
G467
M468
P469
R470
K473
V474
N475
G476
K477
R478
R483
A484
D485
L488
I493
A496
K497
R498
L499
H500
R501
V502
L503
P504
R507
M508
R509
V510
V511
K514
F519
A520
R521
K522
G523
K524
D525
V526
K529
F530
V531
I532
D535
I538
Y541
V546
V547
E552
A555
T556
S562
E565
F569
R573
A574
V575
K576
V577
G580
V581
E582

# 4 Data and refinement statistics (i)

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

| Property | Value | Source |
|---|---|---|
| Space group | P 43 21 2 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 100.20Å 100.20Å 364.66Å<br>90.00° 90.00° 90.00° | Depositor |
| Resolution (Å) | 50.00 – 2.30 | Depositor |
| % Data completeness<br>(in resolution range) | 98.3 (50.00-2.30) | Depositor |
| $R_{merge}$ | 0.11 | Depositor |
| $R_{sym}$ | 0.11 | Depositor |
| Refinement program | CNS | Depositor |
| R, $R_{free}$ | 0.229 , 0.271 | Depositor |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| Total number of atoms | 9503 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 35.0 | 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: MG, GUN, ZN

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 | 0.38 | 0/4745 | 0.63 | 0/6391 |
| 1 | B | 0.36 | 0/4745 | 0.62 | 0/6391 |
| All | All | 0.37 | 0/9490 | 0.63 | 0/12782 |

There are no bond length outliers.

There are no bond angle outliers.

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 | 4652 | 0 | 4747 | 193 | 0 |
| 1 | B | 4652 | 0 | 4747 | 208 | 0 |
| 2 | A | 1 | 0 | 0 | 0 | 0 |
| 2 | B | 1 | 0 | 0 | 0 | 0 |
| 3 | A | 1 | 0 | 0 | 0 | 0 |
| 3 | B | 1 | 0 | 0 | 0 | 0 |
| 4 | A | 11 | 0 | 5 | 0 | 0 |
| 5 | A | 105 | 0 | 0 | 4 | 0 |
| 5 | B | 79 | 0 | 0 | 3 | 0 |
| All | All | 9503 | 0 | 9499 | 395 | 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 21.

The worst 5 of 395 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:84:ARG:HA | 1:A:84:ARG:HH11 | 1.16 | 1.04 |
| 1:B:264:THR:HG22 | 1:B:266:GLU:H | 1.24 | 1.03 |
| 1:B:514:LYS:H | 1:B:514:LYS:HD2 | 1.24 | 0.99 |
| 1:A:471:GLN:HG2 | 1:A:473:LYS:HE3 | 1.47 | 0.96 |
| 1:A:503:LEU:HD22 | 1:A:507:ARG:HH21 | 1.26 | 0.95 |

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 | 575/582 (99%) | 553 (96%) | 19 (3%) | 3 (0%) | 29 | 35 |
| 1 | B | 575/582 (99%) | 542 (94%) | 28 (5%) | 5 (1%) | 17 | 20 |
| All | All | 1150/1164 (99%) | 1095 (95%) | 47 (4%) | 8 (1%) | 22 | 26 |

5 of 8 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 475 | ASN |
| 1 | A | 477 | LYS |
| 1 | B | 82 | ILE |
| 1 | B | 106 | SER |
| 1 | A | 436 | ALA |

### 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 | 499/503 (99%) | 457 (92%) | 42 (8%) | 11 | 13 |
| 1 | B | 499/503 (99%) | 469 (94%) | 30 (6%) | 19 | 26 |
| All | All | 998/1006 (99%) | 926 (93%) | 72 (7%) | 14 | 18 |

5 of 72 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | B | 292 | LEU |
| 1 | B | 582 | GLU |
| 1 | B | 302 | ARG |
| 1 | B | 406 | ARG |
| 1 | A | 330 | ARG |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | B | 290 | GLN |
| 1 | B | 475 | ASN |
| 1 | A | 513 | ASN |
| 1 | B | 41 | ASN |
| 1 | B | 83 | HIS |

### 5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry (i)

Of 5 ligands modelled in this entry, 4 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

| Mol | Type | Chain | Res | Link | Bond lengths | | | Bond angles | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Counts | RMSZ | $\#\|Z\| > 2$ | Counts | RMSZ | $\#\|Z\| > 2$ |
| 4 | GUN | A | 602 | - | 7,12,12 | 1.37 | 1 (14%) | 8,17,17 | 0.80 | 0 |

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

| Mol | Type | Chain | Res | Link | Chirals | Torsions | Rings |
|---|---|---|---|---|---|---|---|
| 4 | GUN | A | 602 | - | - | - | 0/2/2/2 |

All (1) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 4 | A | 602 | GUN | C6-N1 | 2.76 | 1.42 | 1.37 |

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 (i)

There are no such residues in this entry.

## 5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

# 6 Fit of model and data (i)

## 6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

## 6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

## 6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

## 6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.