# WORLDWIDE PDB
## PROTEIN DATA BANK

# wwPDB X-ray Structure Validation Summary Report  ⓘ

Jun 12, 2024 – 02:39 AM EDT

| | | |
|---|---|---|
| PDB ID | : | 1YI8 |
| Title | : | Crystal structure of tryptophanyl trRNA synthetase II from Deinococcus radiodurans in complex with L-Trp |
| Authors | : | Buddha, M.R.; Crane, B.R. |
| Deposited on | : | 2005-01-11 |
| Resolution | : | 2.10 Å(reported) |

---

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

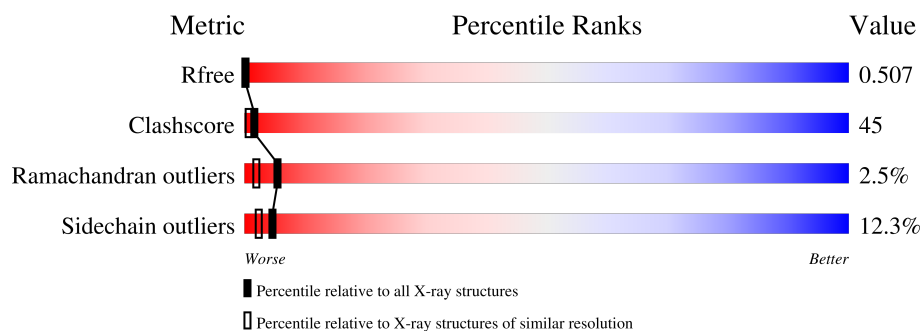| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Mogul | : | 2022.3.0, CSD as543be (2022) |
| Xtriage (Phenix) | : | 1.20.1 |
| EDS | : | 2.36.2 |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Refmac | : | 5.8.0158 |
| CCP4 | : | 7.0.044 (Gargrove) |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.36.2 |

# 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 2.10 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 | | 0.507 |
| Clashscore | | 45 |
| Ramachandran outliers | | 2.5% |
| Sidechain outliers | | 12.3% |

*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)) |
|--------|--------------------------|------------------------------------------------------|
| $R_{free}$ | 130704 | 5197 (2.10-2.10) |
| Clashscore | 141614 | 5710 (2.10-2.10) |
| Ramachandran outliers | 138981 | 5647 (2.10-2.10) |
| Sidechain outliers | 138945 | 5648 (2.10-2.10) |

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%

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|------------------|
| 1 | A | 351 | 32% 51% 11% 6% |
| 1 | B | 351 | 45% 41% 7% 6% |
| 1 | C | 351 | 48% 39% 6% · 6% |

# 2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 8661 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 tryptophanyl-tRNA synthetase.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|---|---------|---------|-------|
| 1 | B | 331 | Total | C | N | O | S | 2 | 0 | 0 |
| | | | 2544 | 1599 | 468 | 471 | 6 | | | |
| 1 | A | 331 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 2511 | 1581 | 457 | 467 | 6 | | | |
| 1 | C | 331 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 2532 | 1593 | 462 | 471 | 6 | | | |

- Molecule 2 is TRYPTOPHAN (three-letter code: TRP) (formula: $C_{11}H_{12}N_2O_2$).
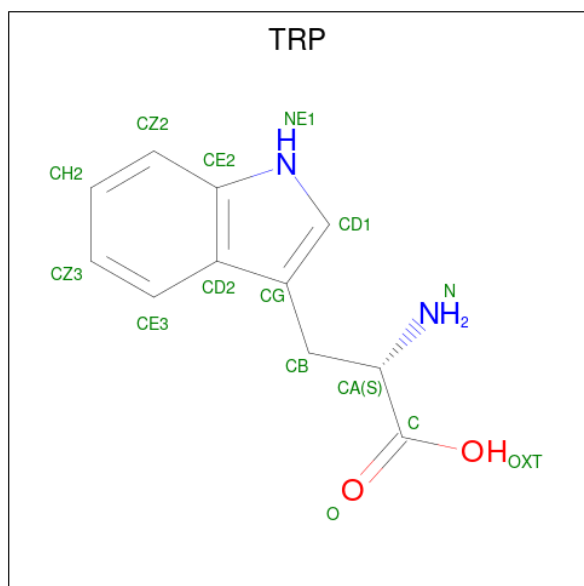


| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|-----|-------|----------|-------|----|---|---|---------|---------|
| 2 | C | 1 | Total | C | N | O | 0 | 0 |
| | | | 15 | 11 | 2 | 2 | | |

- Molecule 3 is water.

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| 3 | B | 308 | Total 308 | O 308 | 0 | 0 |
| 3 | A | 375 | Total 375 | O 375 | 0 | 0 |
| 3 | C | 376 | Total 376 | O 376 | 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 secon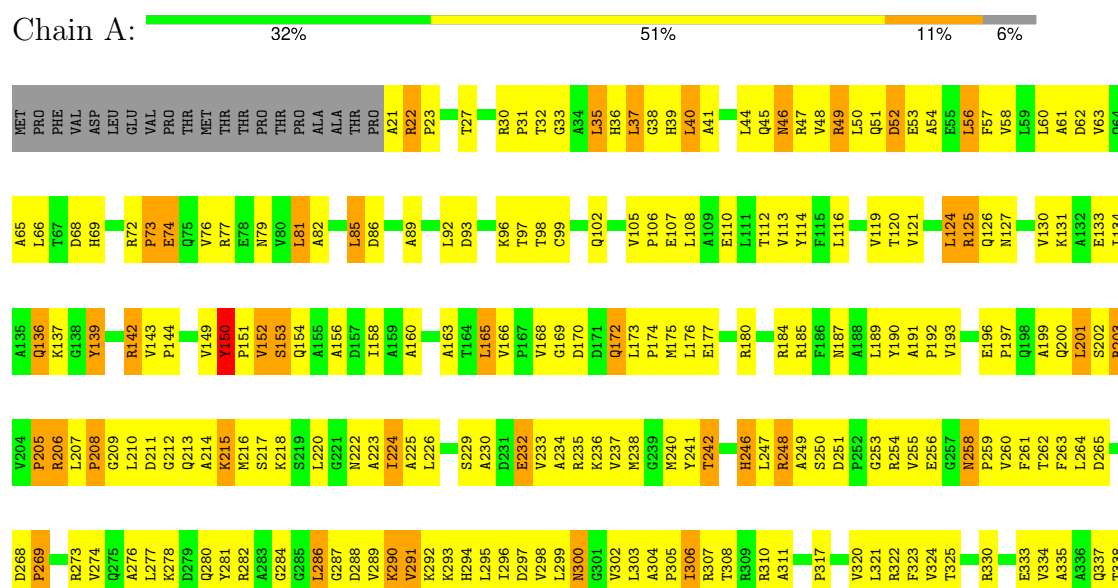d graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: tryptophanyl-tRNA synthetase



- Molecule 1: tryptophanyl-tRNA synthetase

L339 G340 Q341 R344 L348 H351

• Molecule 1: tryptophanyl-tRNA synthetase

Chain C:

48%　　　　　　　　　　　39%　　　　　6% · 6%

MET PRO PHE VAL ASP LEU GLU VAL PRO THR MET THR THR PRO THR PRO ALA ALA THR PRO A21 R22 P23 R24 V25 L26 T27 R30 P31 T32 H36 L40 A41 G42 S43 L44 Q45 N46 R47 V48 R49 L50 Q51 D52 E55 L56 F57 V58 L59 L60 A61 D62 V63 Q64 A65 L66

T67 D68 H69 F70 D71 E74 Q75 V76 R77 E78 N79 V80 L81 A82 V83 A84 L85 L92 D93 P94 Q95 K96 T97 T98 C99 V100 V101 Q102 S103 L108 T112 L116 N117 T120 H123 N127 P128 T129 V130 E133 Q136 K137 E141 A145 Y150 P151 V152

S153 Q154 A155 A156 D157 I158 L165 V166 P167 V168 G169 D170 D171 Q172 M175 L176 R180 V183 R184 R185 F186 N187 A188 L189 P192 V193 L194 A195 E196 P197 Q200 L201 S202 R203 V204 P205 L210 D211 M216 S217 K218 A223 I224 A230 D231 E232 V233 A234 R235

M238 G239 M240 Y241 T242 D243 P244 G245 H246 L247 R248 D251 R254 V255 E256 G257 N258 F267 D268 P269 D270 P271 A272 R273 A276 L277 K278 D279 Q280 A283 L286 K290 V291 K292 K293 H294 L295 I296 D297 V298 L299 N300 A304 P305 I306 R307 R310 A311 E314 R315

V320 L321 R322 F323 V324 T325 E326 R330 G331 R332 L339 R343 R344 R347 H351

# 4 Data and refinement statistics ⓘ

| Property | Value | Source |
|---|---|---|
| Space group | C 1 2 1 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 211.80Å   58.50Å   88.90Å<br>90.00°   100.72°   90.00° | Depositor |
| Resolution (Å) | 50.00  –  2.10<br>30.02  –  2.12 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 88.7 (50.00-2.10)<br>89.4 (30.02-2.12) | Depositor<br>EDS |
| $R_{merge}$ | 0.08 | Depositor |
| $R_{sym}$ | 0.06 | Depositor |
| $< I/\sigma(I) >$ [1] | 1.45 (at 2.12Å) | Xtriage |
| Refinement program | CNS | Depositor |
| R, $R_{free}$ | 0.260  ,  0.280<br>0.497  ,  0.507 | Depositor<br>DCC |
| $R_{free}$ test set | 5520 reflections (10.14%) | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 35.9 | Xtriage |
| Anisotropy | 0.313 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.26 , 39.0 | EDS |
| L-test for twinning[2] | $< |L| > = 0.49, < L^2 > = 0.33$ | Xtriage |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| $F_o$,$F_c$ correlation | 0.64 | EDS |
| Total number of atoms | 8661 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 52.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.28% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

# 5 Model quality (i)

## 5.1 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 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 | 0.36 | 0/2558 | 0.68 | 0/3478 |
| 1 | B | 0.39 | 0/2591 | 0.68 | 0/3519 |
| 1 | C | 0.48 | 0/2579 | 0.73 | 3/3505 (0.1%) |
| All | All | 0.41 | 0/7728 | 0.70 | 3/10502 (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 outliers | #Planarity outliers |
|-----|-------|---------------------|---------------------|
| 1 | A | 0 | 2 |

There are no bond length outliers.

All (3) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|-----|-------|-----|------|-------|-----|------------|---------|
| 1 | C | 200 | GLN | N-CA-C | -5.49 | 96.19 | 111.00 |
| 1 | C | 202 | SER | N-CA-C | 5.07 | 124.69 | 111.00 |
| 1 | C | 169 | GLY | N-CA-C | -5.06 | 100.45 | 113.10 |

There are no chirality outliers.

All (2) planarity outliers are listed below:

| Mol | Chain | Res | Type | Group |
|-----|-------|-----|------|-------|
| 1 | A | 139 | TYR | Sidechain |
| 1 | A | 150 | TYR | Sidechain |

## 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 | 2511 | 0 | 2507 | 297 | 15 |
| 1 | B | 2544 | 0 | 2568 | 206 | 7 |
| 1 | C | 2532 | 0 | 2546 | 197 | 8 |
| 2 | C | 15 | 0 | 9 | 4 | 0 |
| 3 | A | 375 | 0 | 0 | 16 | 6 |
| 3 | B | 308 | 0 | 0 | 7 | 6 |
| 3 | C | 376 | 0 | 0 | 16 | 18 |
| All | All | 8661 | 0 | 7630 | 684 | 32 |

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 45.

The worst 5 of 684 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:168:VAL:HG23 | 1:A:172:GLN:HB2 | 1.20 | 1.15 |
| 1:A:226:LEU:HB3 | 1:A:306:ILE:HD11 | 1.34 | 1.09 |
| 1:A:207:LEU:HD23 | 1:A:216:MET:HE1 | 1.38 | 1.04 |
| 1:C:180:ARG:HG3 | 1:C:180:ARG:HH11 | 1.21 | 1.03 |
| 1:B:51:GLN:HE21 | 1:B:97:THR:HG22 | 1.18 | 1.02 |

The worst 5 of 32 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

| Atom-1 | Atom-2 | Interatomic distance (Å) | Clash overlap (Å) |
|--------|--------|--------------------------|-------------------|
| 1:A:235:ARG:CA | 3:C:2021:HOH:O[1_554] | 0.49 | 1.71 |
| 3:C:1200:HOH:O | 3:C:1789:HOH:O[4_546] | 0.74 | 1.46 |
| 1:C:315:ARG:NH1 | 3:C:1771:HOH:O[4_556] | 0.82 | 1.38 |
| 1:A:235:ARG:CZ | 3:B:1589:HOH:O[1_554] | 0.84 | 1.36 |
| 1:B:322:ARG:NH1 | 3:A:1468:HOH:O[1_546] | 0.88 | 1.32 |

## 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 | 329/351 (94%) | 278 (84%) | 37 (11%) | 14 (4%) | 2 | 0 |
| 1 | B | 329/351 (94%) | 294 (89%) | 28 (8%) | 7 (2%) | 7 | 3 |
| 1 | C | 329/351 (94%) | 312 (95%) | 13 (4%) | 4 (1%) | 13 | 8 |
| All | All | 987/1053 (94%) | 884 (90%) | 78 (8%) | 25 (2%) | 5 | 2 |

5 of 25 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | B | 22 | ARG |
| 1 | A | 61 | ALA |
| 1 | C | 22 | ARG |
| 1 | C | 202 | SER |
| 1 | C | 205 | PRO |

### 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 | 256/282 (91%) | 224 (88%) | 32 (12%) | 4 | 2 |
| 1 | B | 264/282 (94%) | 230 (87%) | 34 (13%) | 4 | 2 |
| 1 | C | 262/282 (93%) | 232 (88%) | 30 (12%) | 5 | 3 |
| All | All | 782/846 (92%) | 686 (88%) | 96 (12%) | 4 | 2 |

5 of 96 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | A | 232 | GLU |
| 1 | C | 97 | THR |
| 1 | A | 248 | ARG |
| 1 | A | 348 | LEU |
| 1 | C | 152 | VAL |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | C | 117 | ASN |
| 1 | C | 126 | GLN |
| 1 | C | 258 | ASN |
| 1 | A | 36 | HIS |
| 1 | B | 337 | 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 ⓘ

1 ligand is modelled in this entry.

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 | #$\|Z\|>2$ |
| 2 | TRP | C | 3000 | - | 14,16,16 | 1.31 | 2 (14%) | 13,22,22 | 0.82 | 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 |
|-----|------|-------|-----|------|---------|----------|-------|
| 2 | TRP | C | 3000 | - | - | 2/7/8/8 | 0/2/2/2 |

All (2) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|---------|-------|-------------|----------|
| 2 | C | 3000 | TRP | OXT-C | -2.94 | 1.21 | 1.30 |
| 2 | C | 3000 | TRP | CH2-CZ3 | 2.06 | 1.42 | 1.38 |

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|-----|-------|-----|------|-------------|
| 2 | C | 3000 | TRP | OXT-C-CA-CB |
| 2 | C | 3000 | TRP | O-C-CA-CB |

There are no ring outliers.

1 monomer is involved in 4 short contacts:

| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|-----|-------|-----|------|---------|--------------|
| 2 | C | 3000 | TRP | 4 | 0 |

### 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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

## 6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

## 6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

## 6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.