# W O R L D W I D E
# wwPDB
## PROTEIN DATA BANK

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

Oct 3, 2023 – 04:26 AM EDT

| | | |
|---:|:--:|:---|
| PDB ID | : | 6O41 |
| Title | : | Crystal structure of the unbound PGZL1 germline Fab fragment (PGZL1_gVmDmJ) |
| Authors | : | Irimia, A.; Wilson, I.A. |
| Deposited on | : | 2019-02-27 |
| Resolution | : | 2.46 Å(reported) |

———————————————————

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

| | | |
|---:|:--:|:---|
| MolProbity | : | FAILED |
| Mogul | : | 1.8.5 (274361), CSD as541be (2020) |
| Xtriage (Phenix) | : | 1.13 |
| EDS | : | FAILED |
| 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) | : | 2.35.1 |

# 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 2.46 Å.

There are no overall percentile quality scores available for this entry.

MolProbity and EDS failed to run properly - the sequence quality summary graphics cannot be shown.

# 2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 11744 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 germline PGZL1_gVmDmJ light chain.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---|---------|---------|-------|
| 1 | L | 213 | Total 1668 | C 1037 | N 291 | O 335 | S 5 | 0 | 4 | 0 |
| 1 | A | 213 | Total 1635 | C 1019 | N 279 | O 332 | S 5 | 0 | 1 | 0 |
| 1 | C | 213 | Total 1635 | C 1019 | N 279 | O 332 | S 5 | 0 | 1 | 0 |

- Molecule 2 is a protein called germline PGZL1_gVmDmJ heavy chain.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---|---------|---------|-------|
| 2 | H | 223 | Total 1652 | C 1044 | N 275 | O 324 | S 9 | 0 | 4 | 0 |
| 2 | B | 222 | Total 1643 | C 1039 | N 274 | O 321 | S 9 | 0 | 3 | 0 |
| 2 | D | 222 | Total 1639 | C 1036 | N 273 | O 321 | S 9 | 0 | 3 | 0 |

- Molecule 3 is a protein called Immunoglobulin G-binding protein G (DOMAIN III).

| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---------|---------|-------|
| 3 | M | 60 | Total 465 | C 292 | N 76 | O 97 | 0 | 0 | 0 |
| 3 | N | 60 | Total 473 | C 296 | N 78 | O 99 | 0 | 1 | 0 |
| 3 | O | 59 | Total 460 | C 289 | N 75 | O 96 | 0 | 0 | 0 |

There are 18 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------|-----------|
| M | 1 | MET | - | initiating methionine | UNP P19909 |
| M | 62 | HIS | - | expression tag | UNP P19909 |

*Continued on next page...*

*Continued from previous page...*

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------|-----------|
| M | 63 | HIS | - | expression tag | UNP P19909 |
| M | 64 | HIS | - | expression tag | UNP P19909 |
| M | 65 | HIS | - | expression tag | UNP P19909 |
| M | 66 | HIS | - | expression tag | UNP P19909 |
| N | 1 | MET | - | initiating methionine | UNP P19909 |
| N | 62 | HIS | - | expression tag | UNP P19909 |
| N | 63 | HIS | - | expression tag | UNP P19909 |
| N | 64 | HIS | - | expression tag | UNP P19909 |
| N | 65 | HIS | - | expression tag | UNP P19909 |
| N | 66 | HIS | - | expression tag | UNP P19909 |
| O | 1 | MET | - | initiating methionine | UNP P19909 |
| O | 62 | HIS | - | expression tag | UNP P19909 |
| O | 63 | HIS | - | expression tag | UNP P19909 |
| O | 64 | HIS | - | expression tag | UNP P19909 |
| O | 65 | HIS | - | expression tag | UNP P19909 |
| O | 66 | HIS | - | expression tag | UNP P19909 |

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).
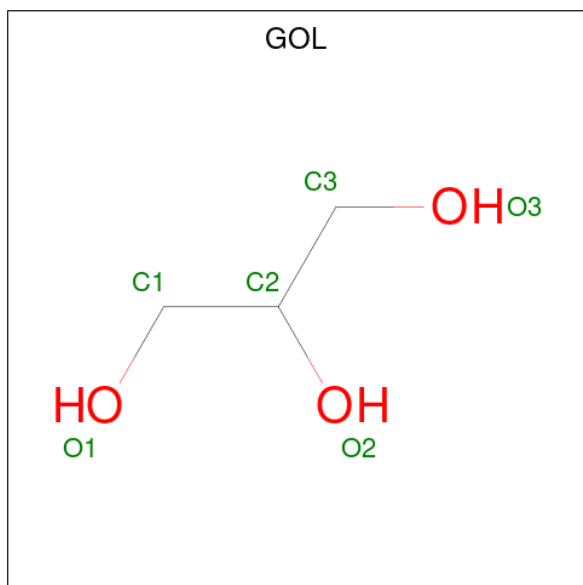


| Mol | Chain | Residues | Atoms | | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---|---------|---------|
| 4 | L | 1 | Total<br>6 | C<br>3 | O<br>3 | 0 | 0 |
| 4 | L | 1 | Total<br>6 | C<br>3 | O<br>3 | 0 | 0 |
| 4 | L | 1 | Total<br>6 | C<br>3 | O<br>3 | 0 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Residues | Atoms | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|
| 4 | L | 1 | Total 6 | C 3 | O 3 | 0 | 0 |
| 4 | L | 1 | Total 6 | C 3 | O 3 | 0 | 0 |
| 4 | H | 1 | Total 6 | C 3 | O 3 | 0 | 0 |
| 4 | H | 1 | Total 6 | C 3 | O 3 | 0 | 0 |
| 4 | A | 1 | Total 6 | C 3 | O 3 | 0 | 0 |
| 4 | B | 1 | Total 6 | C 3 | O 3 | 0 | 0 |
| 4 | C | 1 | Total 6 | C 3 | O 3 | 0 | 0 |
| 4 | C | 1 | Total 6 | C 3 | O 3 | 0 | 0 |
| 4 | D | 1 | Total 6 | C 3 | O 3 | 0 | 0 |
| 4 | O | 1 | Total 6 | C 3 | O 3 | 0 | 0 |

- Molecule 5 is water.

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| 5 | L | 79 | Total 79 | O 79 | 0 | 0 |
| 5 | H | 58 | Total 58 | O 58 | 0 | 0 |
| 5 | A | 59 | Total 59 | O 59 | 0 | 0 |
| 5 | B | 47 | Total 47 | O 47 | 0 | 0 |
| 5 | C | 72 | Total 72 | O 72 | 0 | 0 |
| 5 | D | 42 | Total 42 | O 42 | 0 | 0 |
| 5 | M | 16 | Total 16 | O 16 | 0 | 0 |
| 5 | N | 13 | Total 13 | O 13 | 0 | 0 |
| 5 | O | 10 | Total 10 | O 10 | 0 | 0 |

MolProbity and EDS failed to run properly - this section is therefore empty.

# 3 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

| Property | Value | Source |
|----------|-------|--------|
| Space group | P 1 21 1 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 78.84Å   107.49Å   131.96Å<br>90.00°   94.61°   90.00° | Depositor |
| Resolution (Å) | 39.67   –   2.46 | Depositor |
| % Data completeness<br>(in resolution range) | 97.8 (39.67-2.46) | Depositor |
| $R_{merge}$ | (Not available) | Depositor |
| $R_{sym}$ | 0.10 | Depositor |
| $<I/\sigma(I)>$ [1] | 1.68 (at 2.48Å) | Xtriage |
| Refinement program | PHENIX (1.12_2829: ???) | Depositor |
| R, $R_{free}$ | 0.169   ,   0.219 | Depositor |
| Wilson B-factor (Å$^2$) | 50.0 | Xtriage |
| Anisotropy | 0.211 | Xtriage |
| L-test for twinning[2] | $<|L|> = 0.51$, $<L^2> = 0.34$ | Xtriage |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| Total number of atoms | 11744 | 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 3.12% 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.

# 4 Model quality ⓘ

## 4.1 Standard geometry ⓘ

MolProbity failed to run properly - this section is therefore empty.

## 4.2 Too-close contacts ⓘ

MolProbity failed to run properly - this section is therefore empty.

## 4.3 Torsion angles ⓘ

### 4.3.1 Protein backbone ⓘ

MolProbity failed to run properly - this section is therefore empty.

### 4.3.2 Protein sidechains ⓘ

MolProbity failed to run properly - this section is therefore empty.

### 4.3.3 RNA ⓘ

MolProbity failed to run properly - this section is therefore empty.

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

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

## 4.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 4.6 Ligand geometry ⓘ

13 ligands are 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 | #$|Z| > 2$ |
| 4 | GOL | L | 304 | - | 5,5,5 | 0.90 | 0 | 5,5,5 | 1.02 | 0 |
| 4 | GOL | L | 303 | - | 5,5,5 | 0.92 | 0 | 5,5,5 | 0.95 | 0 |
| 4 | GOL | C | 302 | - | 5,5,5 | 0.97 | 0 | 5,5,5 | 1.01 | 0 |
| 4 | GOL | D | 301 | - | 5,5,5 | 0.82 | 0 | 5,5,5 | 1.20 | 1 (20%) |
| 4 | GOL | A | 301 | - | 5,5,5 | 0.72 | 0 | 5,5,5 | 1.20 | 0 |
| 4 | GOL | H | 302 | - | 5,5,5 | 1.03 | 0 | 5,5,5 | 0.98 | 0 |
| 4 | GOL | L | 301 | - | 5,5,5 | 0.92 | 0 | 5,5,5 | 0.98 | 0 |
| 4 | GOL | L | 302 | - | 5,5,5 | 0.97 | 0 | 5,5,5 | 0.93 | 0 |
| 4 | GOL | B | 301 | - | 5,5,5 | 1.00 | 0 | 5,5,5 | 0.89 | 0 |
| 4 | GOL | L | 305 | - | 5,5,5 | 1.01 | 0 | 5,5,5 | 0.85 | 0 |
| 4 | GOL | O | 101 | - | 5,5,5 | 0.88 | 0 | 5,5,5 | 0.84 | 0 |
| 4 | GOL | H | 301 | - | 5,5,5 | 1.05 | 0 | 5,5,5 | 0.95 | 0 |
| 4 | GOL | C | 301 | - | 5,5,5 | 0.80 | 0 | 5,5,5 | 1.20 | 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 | GOL | L | 304 | - | - | 0/4/4/4 | - |
| 4 | GOL | L | 303 | - | - | 4/4/4/4 | - |
| 4 | GOL | C | 302 | - | - | 2/4/4/4 | - |
| 4 | GOL | D | 301 | - | - | 2/4/4/4 | - |
| 4 | GOL | A | 301 | - | - | 4/4/4/4 | - |
| 4 | GOL | H | 302 | - | - | 2/4/4/4 | - |
| 4 | GOL | L | 301 | - | - | 1/4/4/4 | - |
| 4 | GOL | L | 302 | - | - | 2/4/4/4 | - |
| 4 | GOL | B | 301 | - | - | 0/4/4/4 | - |
| 4 | GOL | L | 305 | - | - | 2/4/4/4 | - |
| 4 | GOL | O | 101 | - | - | 1/4/4/4 | - |
| 4 | GOL | H | 301 | - | - | 2/4/4/4 | - |
| 4 | GOL | C | 301 | - | - | 0/4/4/4 | - |

There are no bond length outliers.

All (1) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|-----|-------|-----|------|----------|-------|----------------|-------------|
| 4 | D | 301 | GOL | C3-C2-C1 | -2.21 | 103.11 | 111.70 |

There are no chirality outliers.

5 of 22 torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|-----|-------|-----|------|-------------|
| 4 | L | 303 | GOL | O1-C1-C2-O2 |
| 4 | L | 303 | GOL | O1-C1-C2-C3 |
| 4 | H | 301 | GOL | C1-C2-C3-O3 |
| 4 | H | 302 | GOL | O1-C1-C2-C3 |
| 4 | A | 301 | GOL | O1-C1-C2-C3 |

There are no ring outliers.

No monomer is involved in short contacts.

## 4.7 Other polymers ⓘ

There are no such residues in this entry.

## 4.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

# 5 Fit of model and data ⓘ

## 5.1 Protein, DNA and RNA chains ⓘ

EDS failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

## 5.3 Carbohydrates ⓘ

EDS failed to run properly - this section is therefore empty.

## 5.4 Ligands ⓘ

EDS failed to run properly - this section is therefore empty.

## 5.5 Other polymers ⓘ

EDS failed to run properly - this section is therefore empty.