



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 4, 2022 – 02:08 PM EDT

PDB ID : 7MSN
Title : SunS glycosin S-glycosyltransferase
Authors : Garg, N.; Nair, S.K.
Deposited on : 2021-05-11
Resolution : 3.00 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Xtrriage (Phenix) : 1.13
EDS : **FAILED**
buster-report : 1.1.7 (2018)
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) : 2.27

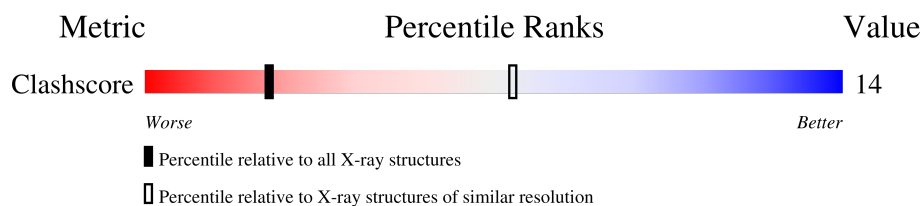
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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 | 2416 (3.00-3.00) |

2 Entry composition [i](#)

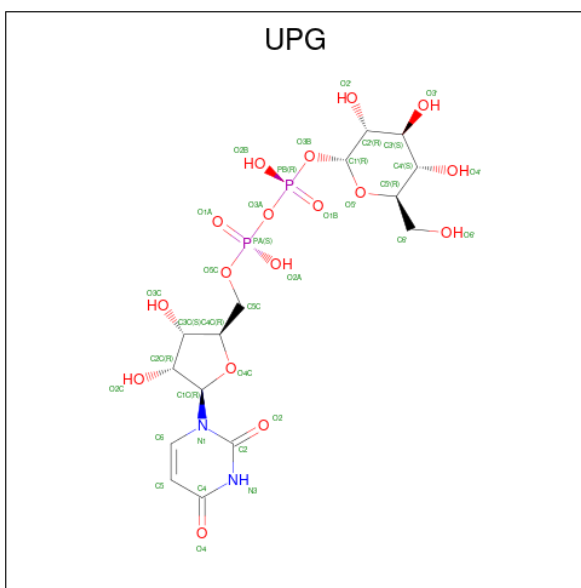
There are 2 unique types of molecules in this entry. The entry contains 6995 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 SPbeta prophage-derived glycosyltransferase SunS.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|---------------|-----------|----------|----------|---------|---------|---------|-------|
| | | | Total | C | N | O | S | | | |
| 1 | A | 417 | Total 3462 | C 2218 | N 570 | O 659 | S 15 | 0 | 0 | 0 |
| 1 | B | 416 | Total 3461 | C 2217 | N 569 | O 660 | S 15 | 0 | 0 | 0 |

- Molecule 2 is URIDINE-5'-DIPHOSPHATE-GLUCOSE (three-letter code: UPG) (formula: $C_{15}H_{24}N_2O_{17}P_2$) (labeled as "Ligand of Interest" by depositor).



| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf |
|-----|-------|----------|-------------|---------|--------|---------|--------|---------|---------|
| | | | Total | C | N | O | P | | |
| 2 | A | 1 | Total 36 | C 15 | N 2 | O 17 | P 2 | 0 | 0 |
| 2 | B | 1 | Total 36 | C 15 | N 2 | O 17 | P 2 | 0 | 0 |

SEQUENCE-PLOTS INFOmissingINFO

3 Data and refinement statistics

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

| Property | Value | Source |
|--|---|-----------|
| Space group | P 43 21 2 | Depositor |
| Cell constants a, b, c, α , β , γ | 108.73Å 108.73Å 210.70Å 90.00° 90.00° 90.00° | Depositor |
| Resolution (Å) | 47.38 – 3.00 | Depositor |
| % Data completeness (in resolution range) | 98.8 (47.38-3.00) | Depositor |
| R_{merge} | (Not available) | Depositor |
| R_{sym} | (Not available) | Depositor |
| $\langle I/\sigma(I) \rangle$ ¹ | 2.19 (at 3.01Å) | Xtrriage |
| Refinement program | PHENIX 1.18.2_3874 | Depositor |
| R, R_{free} | 0.204 , 0.280 | Depositor |
| Wilson B-factor (Å ²) | 70.6 | Xtrriage |
| Anisotropy | 0.188 | Xtrriage |
| L-test for twinning ² | $\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$ | Xtrriage |
| Estimated twinning fraction | No twinning to report. | Xtrriage |
| Total number of atoms | 6995 | wwPDB-VP |
| Average B, all atoms (Å ²) | 64.0 | wwPDB-VP |

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.20% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

4 Model quality [i](#)

4.1 Standard geometry [i](#)

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

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 | 0.53 | 0/3528 | 0.77 | 4/4760 (0.1%) |
| 1 | B | 0.56 | 0/3527 | 0.82 | 6/4758 (0.1%) |
| All | All | 0.54 | 0/7055 | 0.79 | 10/9518 (0.1%) |

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 | B | 0 | 5 |

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(°) | Ideal(°) |
|-----|-------|-----|------|-----------|-------|-------------|----------|
| 1 | B | 26 | LEU | CB-CG-CD2 | -7.31 | 98.57 | 111.00 |
| 1 | B | 294 | LEU | CA-CB-CG | 7.13 | 131.69 | 115.30 |
| 1 | A | 281 | TYR | CB-CG-CD1 | -6.84 | 116.89 | 121.00 |
| 1 | B | 281 | TYR | CB-CG-CD1 | -5.67 | 117.60 | 121.00 |
| 1 | A | 281 | TYR | CB-CG-CD2 | 5.45 | 124.27 | 121.00 |

There are no chirality outliers.

All (5) planarity outliers are listed below:

| Mol | Chain | Res | Type | Group |
|-----|-------|-----|------|---------|
| 1 | B | 202 | ASN | Peptide |
| 1 | B | 205 | HIS | Peptide |
| 1 | B | 206 | SER | Peptide |

Continued on next page...

Continued from previous page...

| Mol | Chain | Res | Type | Group |
|-----|-------|-----|------|---------|
| 1 | B | 27 | VAL | Peptide |
| 1 | B | 62 | VAL | Peptide |

4.2 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 | 3462 | 0 | 3450 | 100 | 0 |
| 1 | B | 3461 | 0 | 3452 | 101 | 0 |
| 2 | A | 36 | 0 | 21 | 2 | 0 |
| 2 | B | 36 | 0 | 21 | 5 | 0 |
| All | All | 6995 | 0 | 6944 | 190 | 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 14.

The worst 5 of 190 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) |
|-----------------|-----------------|--------------------------|-------------------|
| 2:B:500:UPG:C4C | 2:B:500:UPG:O4C | 1.66 | 1.15 |
| 2:A:500:UPG:O4C | 2:A:500:UPG:C4C | 1.66 | 1.14 |
| 1:A:28:ASN:ND2 | 1:A:210:ASN:O | 2.08 | 0.86 |
| 1:A:121:ASN:HA | 1:A:191:LYS:HE2 | 1.67 | 0.77 |
| 1:A:165:VAL:HB | 1:A:183:MET:HB3 | 1.68 | 0.76 |

There are no symmetry-related clashes.

4.3 Torsion angles [i](#)

4.3.1 Protein backbone [i](#)

There are no protein backbone outliers to report in this entry.

4.3.2 Protein sidechains [i](#)

There are no protein residues with a non-rotameric sidechain to report in this entry.

4.3.3 RNA [i](#)

There are no RNA molecules in this entry.

4.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

4.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

4.6 Ligand geometry [i](#)

2 ligands are modelled in this entry.

There are no bond length outliers.

There are no bond angle outliers.

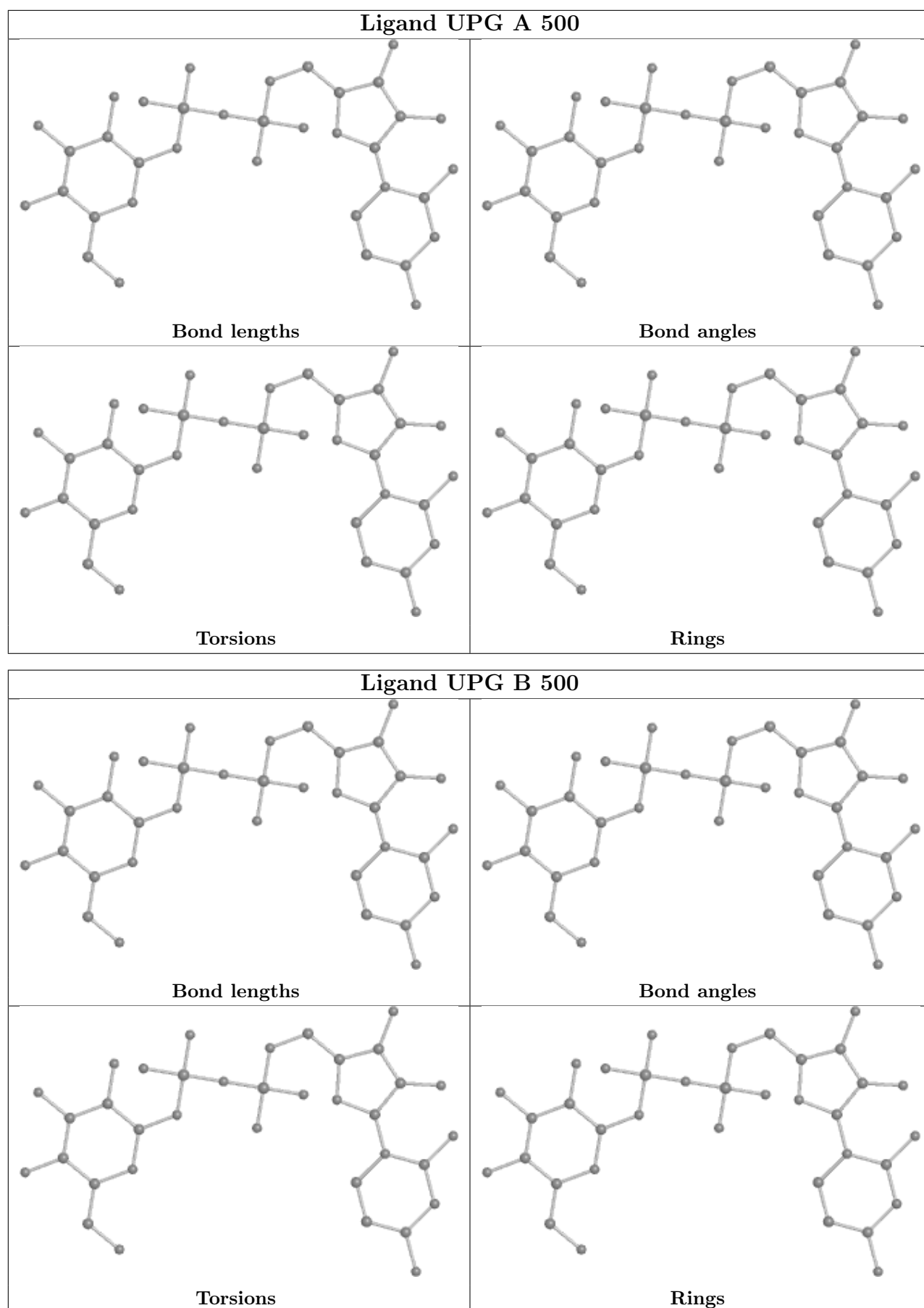
There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



4.7 Other polymers [i](#)

There are no such residues in this entry.

4.8 Polymer linkage issues [i](#)

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.