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

Jun 12, 2024 – 09:16 PM EDT

| | | |
|---|---|---|
| PDB ID | : | 3R2W |
| Title | : | Crystal Structure of UDP-glucose Pyrophosphorylase of Homo Sapiens |
| Authors | : | Zheng, X.; Yu, Q. |
| Deposited on | : | 2011-03-14 |
| Resolution | : | 3.60 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 (1)) were used in the production of this report:

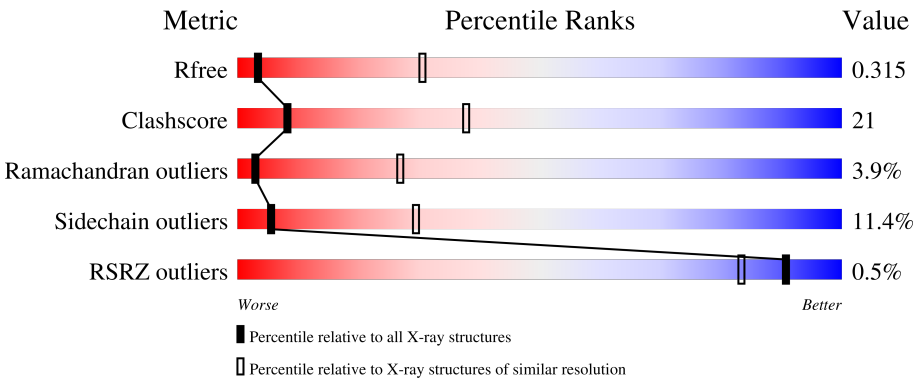
| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Xtriage (Phenix) | : | 1.20.1 |
| EDS | : | 2.36.2 |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Refmac | : | 5.8.0158 |
| CCP4 | : | 7.0.044 (Gargrove) |
| Ideal geometry (proteins) | : |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

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.60 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)) |
|---|---|---|
| $R_{free}$ | 130704 | 1257 (3.70-3.50) |
| Clashscore | 141614 | 1353 (3.70-3.50) |
| Ramachandran outliers | 138981 | 1307 (3.70-3.50) |
| Sidechain outliers | 138945 | 1307 (3.70-3.50) |
| RSRZ outliers | 127900 | 1161 (3.70-3.50) |

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% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 1 | A | 528 | 51% 32% 5% • 12% |
| 1 | B | 528 | 47% 34% 6% • 12% |
| 1 | C | 528 | % 55% 26% • 15% |
| 1 | D | 528 | 46% 36% 6% 11% |

# 2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 13938 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 UTP--glucose-1-phosphate uridylyltransferase.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | | |
| 1 | A | 465 | 3539 | 2253 | 604 | 672 | 10 | 0 | 0 | 0 |
| 1 | B | 467 | 3534 | 2251 | 605 | 667 | 11 | 0 | 0 | 0 |
| 1 | C | 450 | 3299 | 2088 | 557 | 646 | 8 | 0 | 0 | 0 |
| 1 | D | 468 | 3566 | 2275 | 608 | 672 | 11 | 0 | 0 | 0 |

There are 80 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| A | -30 | MET | - | expression tag | UNP Q16851 |
| A | -29 | GLY | - | expression tag | UNP Q16851 |
| A | -28 | SER | - | expression tag | UNP Q16851 |
| A | -27 | SER | - | expression tag | UNP Q16851 |
| A | -26 | HIS | - | expression tag | UNP Q16851 |
| A | -25 | HIS | - | expression tag | UNP Q16851 |
| A | -24 | HIS | - | expression tag | UNP Q16851 |
| A | -23 | HIS | - | expression tag | UNP Q16851 |
| A | -22 | HIS | - | expression tag | UNP Q16851 |
| A | -21 | HIS | - | expression tag | UNP Q16851 |
| A | -20 | SER | - | expression tag | UNP Q16851 |
| A | -19 | SER | - | expression tag | UNP Q16851 |
| A | -18 | GLY | - | expression tag | UNP Q16851 |
| A | -17 | LEU | - | expression tag | UNP Q16851 |
| A | -16 | VAL | - | expression tag | UNP Q16851 |
| A | -15 | PRO | - | expression tag | UNP Q16851 |
| A | -14 | ARG | - | expression tag | UNP Q16851 |
| A | -13 | GLY | - | expression tag | UNP Q16851 |
| A | -12 | SER | - | expression tag | UNP Q16851 |
| A | -11 | HIS | - | expression tag | UNP Q16851 |
| B | -30 | MET | - | expression tag | UNP Q16851 |

*Continued on next page...*

*Continued from previous page...*

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| B | -29 | GLY | - | expression tag | UNP Q16851 |
| B | -28 | SER | - | expression tag | UNP Q16851 |
| B | -27 | SER | - | expression tag | UNP Q16851 |
| B | -26 | HIS | - | expression tag | UNP Q16851 |
| B | -25 | HIS | - | expression tag | UNP Q16851 |
| B | -24 | HIS | - | expression tag | UNP Q16851 |
| B | -23 | HIS | - | expression tag | UNP Q16851 |
| B | -22 | HIS | - | expression tag | UNP Q16851 |
| B | -21 | HIS | - | expression tag | UNP Q16851 |
| B | -20 | SER | - | expression tag | UNP Q16851 |
| B | -19 | SER | - | expression tag | UNP Q16851 |
| B | -18 | GLY | - | expression tag | UNP Q16851 |
| B | -17 | LEU | - | expression tag | UNP Q16851 |
| B | -16 | VAL | - | expression tag | UNP Q16851 |
| B | -15 | PRO | - | expression tag | UNP Q16851 |
| B | -14 | ARG | - | expression tag | UNP Q16851 |
| B | -13 | GLY | - | expression tag | UNP Q16851 |
| B | -12 | SER | - | expression tag | UNP Q16851 |
| B | -11 | HIS | - | expression tag | UNP Q16851 |
| C | -30 | MET | - | expression tag | UNP Q16851 |
| C | -29 | GLY | - | expression tag | UNP Q16851 |
| C | -28 | SER | - | expression tag | UNP Q16851 |
| C | -27 | SER | - | expression tag | UNP Q16851 |
| C | -26 | HIS | - | expression tag | UNP Q16851 |
| C | -25 | HIS | - | expression tag | UNP Q16851 |
| C | -24 | HIS | - | expression tag | UNP Q16851 |
| C | -23 | HIS | - | expression tag | UNP Q16851 |
| C | -22 | HIS | - | expression tag | UNP Q16851 |
| C | -21 | HIS | - | expression tag | UNP Q16851 |
| C | -20 | SER | - | expression tag | UNP Q16851 |
| C | -19 | SER | - | expression tag | UNP Q16851 |
| C | -18 | GLY | - | expression tag | UNP Q16851 |
| C | -17 | LEU | - | expression tag | UNP Q16851 |
| C | -16 | VAL | - | expression tag | UNP Q16851 |
| C | -15 | PRO | - | expression tag | UNP Q16851 |
| C | -14 | ARG | - | expression tag | UNP Q16851 |
| C | -13 | GLY | - | expression tag | UNP Q16851 |
| C | -12 | SER | - | expression tag | UNP Q16851 |
| C | -11 | HIS | - | expression tag | UNP Q16851 |
| D | -30 | MET | - | expression tag | UNP Q16851 |
| D | -29 | GLY | - | expression tag | UNP Q16851 |
| D | -28 | SER | - | expression tag | UNP Q16851 |

*Continued on next page...*

*Continued from previous page...*

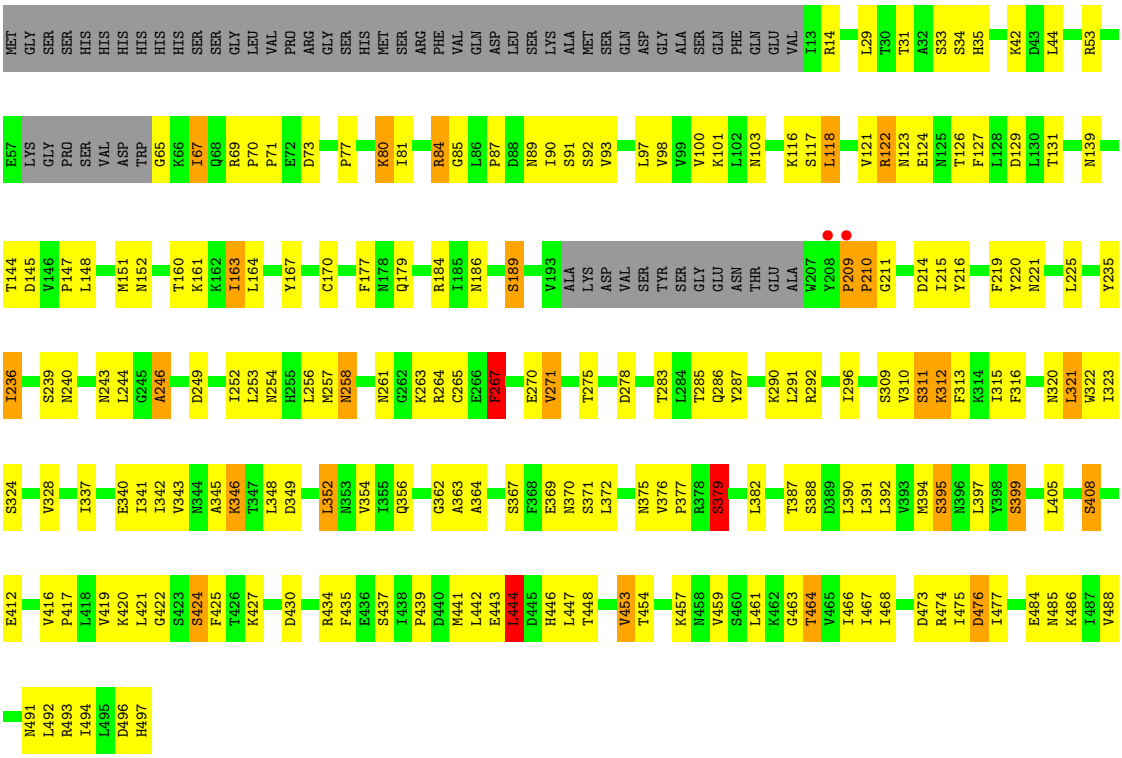
| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| D | -27 | SER | - | expression tag | UNP Q16851 |
| D | -26 | HIS | - | expression tag | UNP Q16851 |
| D | -25 | HIS | - | expression tag | UNP Q16851 |
| D | -24 | HIS | - | expression tag | UNP Q16851 |
| D | -23 | HIS | - | expression tag | UNP Q16851 |
| D | -22 | HIS | - | expression tag | UNP Q16851 |
| D | -21 | HIS | - | expression tag | UNP Q16851 |
| D | -20 | SER | - | expression tag | UNP Q16851 |
| D | -19 | SER | - | expression tag | UNP Q16851 |
| D | -18 | GLY | - | expression tag | UNP Q16851 |
| D | -17 | LEU | - | expression tag | UNP Q16851 |
| D | -16 | VAL | - | expression tag | UNP Q16851 |
| D | -15 | PRO | - | expression tag | UNP Q16851 |
| D | -14 | ARG | - | expression tag | UNP Q16851 |
| D | -13 | GLY | - | expression tag | UNP Q16851 |
| D | -12 | SER | - | expression tag | UNP Q16851 |
| D | -11 | HIS | - | expression tag | UNP Q16851 |

# 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 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.

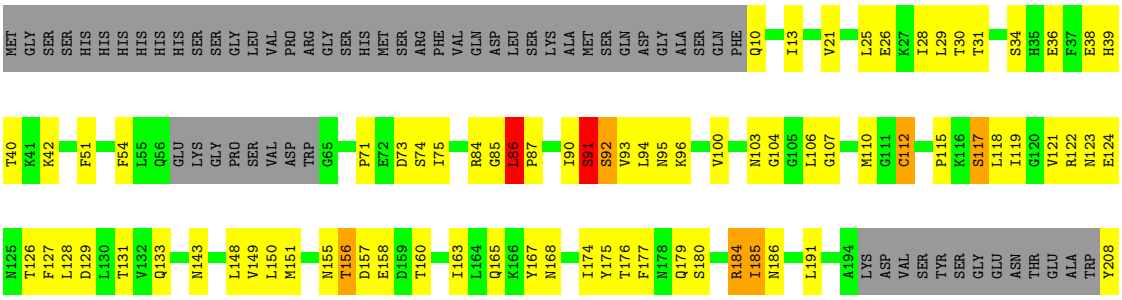
• Molecule 1: UTP--glucose-1-phosphate uridylyltransferase

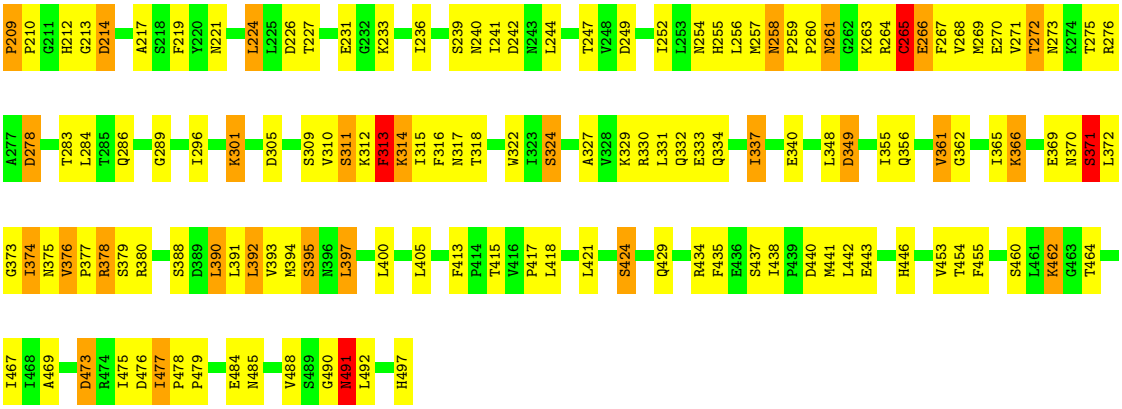
Chain A: 51% 32% 5% • 12%




• Molecule 1: UTP--glucose-1-phosphate uridylyltransferase

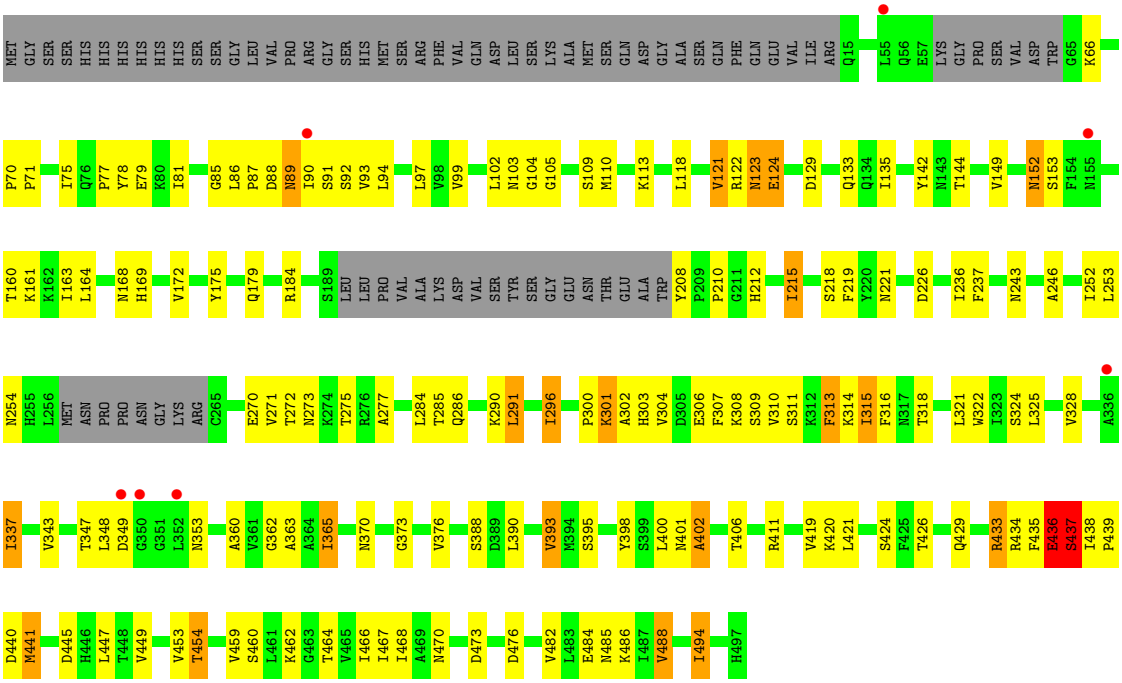
Chain B: 47% 34% 6% • 12%





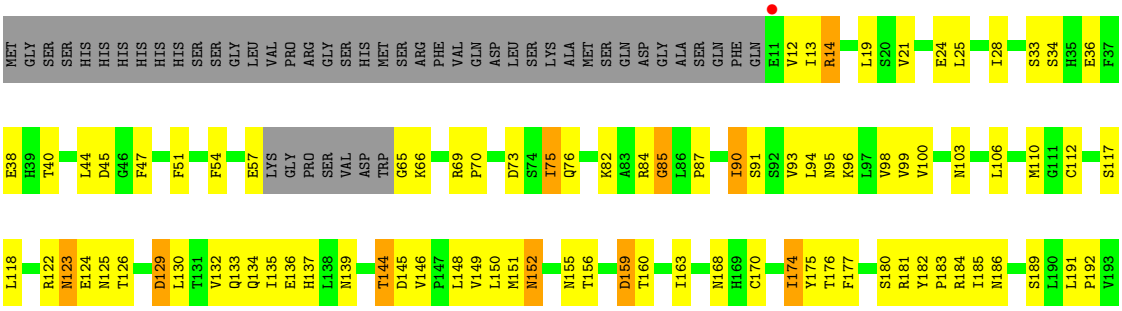
• Molecule 1: UTP--glucose-1-phosphate uridylyltransferase

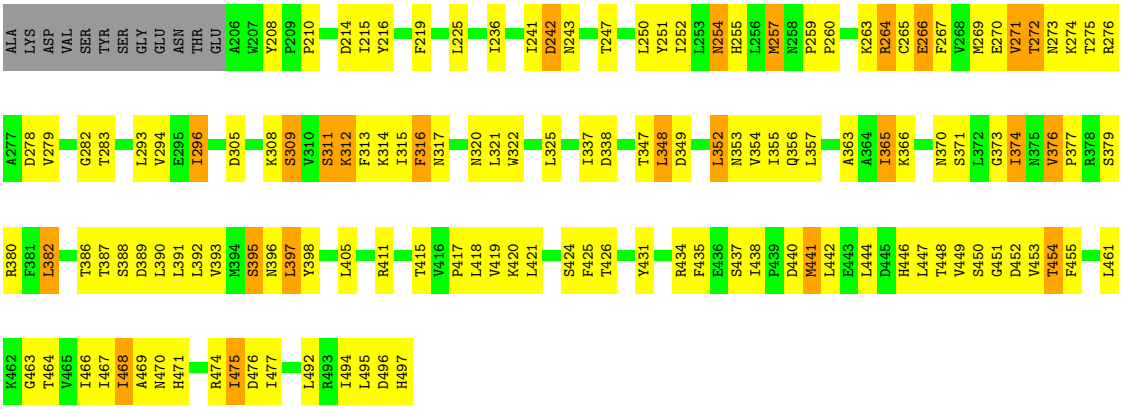
Chain C:



• Molecule 1: UTP--glucose-1-phosphate uridylyltransferase

Chain D:




ALA LYS ASP VAL SER TYR SER GLY GLU ASN THR GLU A206 W207 Y208 P209 P210 D214 I215 Y216 F219 L225 I236 I241 D242 N243 T247 L250 Y251 I252 L253 N254 H255 L256 M257 N258 P259 P260 K263 R264 C265 E266 F267 V268 M269 E270 V271 T272 N273 K274 T275 R276
A277 D278 V279 G282 T283 L293 V294 E295 I296 D305 K308 S309 V310 S311 K312 F313 K314 I315 F316 N317 N320 L321 W322 L325 I337 D338 T347 L348 D349 L352 N353 V354 I355 Q356 L357 A363 A364 I365 K366 N370 S371 L372 G373 I374 N375 V376 P377 R378 S379
R380 F381 L382 T386 T387 S388 D389 L390 L391 L392 V393 M394 S395 N396 L397 Y398 L405 R411 T415 Y416 P417 L418 V419 K420 L421 S424 F425 T426 Y431 R434 F435 E436 S437 I438 P439 D440 M441 L442 E443 L444 D445 H446 L447 T448 V449 S450 G451 D452 V453 T454 F455 L461
K462 G463 T464 V465 I466 I467 I468 A469 N470 H471 R474 I475 D476 I477 L492 R493 I494 L495 D496 H497

# 4 Data and refinement statistics (i)

| Property | Value | Source |
|---|---|---|
| Space group | P 31 2 1 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 140.44Å 140.44Å 311.72Å<br>90.00° 90.00° 120.00° | Depositor |
| Resolution (Å) | 20.00 – 3.60<br>20.00 – 3.60 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 94.0 (20.00-3.60)<br>94.0 (20.00-3.60) | Depositor<br>EDS |
| $R_{merge}$ | (Not available) | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 2.37 (at 3.62Å) | Xtriage |
| Refinement program | REFMAC 5.5.0102 | Depositor |
| R, $R_{free}$ | 0.247 , 0.304<br>0.267 , 0.315 | Depositor<br>DCC |
| $R_{free}$ test set | 1985 reflections (5.05%) | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 113.6 | Xtriage |
| Anisotropy | 0.002 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.22 , 57.3 | EDS |
| L-test for twinning[2] | $< \|L\| > = 0.49$, $< L^2 > = 0.33$ | Xtriage |
| Estimated twinning fraction | 0.022 for -h,-k,l | Xtriage |
| $F_o$,$F_c$ correlation | 0.92 | EDS |
| Total number of atoms | 13938 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 144.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.12% 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.

# 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 | 0.61 | 2/3602 (0.1%) | 0.72 | 1/4894 (0.0%) |
| 1 | B | 0.58 | 0/3596 | 0.79 | 1/4889 (0.0%) |
| 1 | C | 0.51 | 0/3353 | 0.66 | 0/4578 |
| 1 | D | 0.67 | 0/3630 | 0.83 | 3/4936 (0.1%) |
| All | All | 0.60 | 2/14181 (0.0%) | 0.75 | 5/19297 (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 | B | 0 | 1 |

All (2) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 1 | A | 311 | SER | CB-OG | 10.91 | 1.56 | 1.42 |
| 1 | A | 379 | SER | CB-OG | 7.32 | 1.51 | 1.42 |

All (5) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 1 | D | 382 | LEU | CA-CB-CG | 5.99 | 129.08 | 115.30 |
| 1 | D | 33 | SER | N-CA-C | -5.80 | 95.35 | 111.00 |
| 1 | A | 444 | LEU | CA-CB-CG | 5.35 | 127.60 | 115.30 |
| 1 | B | 86 | LEU | CA-CB-CG | 5.28 | 127.45 | 115.30 |
| 1 | D | 33 | SER | C-N-CA | 5.09 | 134.43 | 121.70 |

There are no chirality outliers.

All (1) planarity outliers are listed below:

| Mol | Chain | Res | Type | Group |
|---|---|---|---|---|
| 1 | B | 91 | SER | Peptide |

## 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 | 3539 | 0 | 3474 | 134 | 0 |
| 1 | B | 3534 | 0 | 3471 | 171 | 0 |
| 1 | C | 3299 | 0 | 3094 | 104 | 0 |
| 1 | D | 3566 | 0 | 3511 | 184 | 0 |
| All | All | 13938 | 0 | 13550 | 581 | 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 581 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:B:265:CYS:HA | 1:B:266:GLU:CB | 1.83 | 1.07 |
| 1:C:218:SER:HA | 1:C:221:ASN:HB2 | 1.37 | 1.02 |
| 1:C:91:SER:H | 1:C:92:SER:HA | 1.19 | 1.00 |
| 1:D:87:PRO:HG3 | 1:D:254:ASN:HB2 | 1.47 | 0.96 |
| 1:D:129:ASP:O | 1:D:133:GLN:HG3 | 1.64 | 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 | 459/528 (87%) | 376 (82%) | 64 (14%) | 19 (4%) | 3 | 26 |
| 1 | B | 461/528 (87%) | 377 (82%) | 58 (13%) | 26 (6%) | 2 | 19 |
| 1 | C | 442/528 (84%) | 363 (82%) | 65 (15%) | 14 (3%) | 4 | 31 |
| 1 | D | 462/528 (88%) | 384 (83%) | 65 (14%) | 13 (3%) | 5 | 34 |
| All | All | 1824/2112 (86%) | 1500 (82%) | 252 (14%) | 72 (4%) | 3 | 27 |

5 of 72 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 34 | SER |
| 1 | A | 257 | MET |
| 1 | A | 263 | LYS |
| 1 | B | 265 | CYS |
| 1 | B | 266 | GLU |

### 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 | 381/470 (81%) | 345 (91%) | 36 (9%) | 8 | 38 |
| 1 | B | 380/470 (81%) | 331 (87%) | 49 (13%) | 4 | 24 |
| 1 | C | 341/470 (73%) | 310 (91%) | 31 (9%) | 9 | 39 |
| 1 | D | 385/470 (82%) | 331 (86%) | 54 (14%) | 3 | 21 |
| All | All | 1487/1880 (79%) | 1317 (89%) | 170 (11%) | 5 | 29 |

5 of 170 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | C | 494 | ILE |
| 1 | D | 278 | ASP |
| 1 | D | 45 | ASP |
| 1 | D | 176 | THR |
| 1 | D | 338 | ASP |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | C | 152 | ASN |
| 1 | C | 319 | ASN |
| 1 | D | 446 | HIS |
| 1 | C | 286 | GLN |
| 1 | C | 429 | GLN |

#### 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)

There are no ligands in this entry.

### 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)

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 | 465/528 (88%) | -0.41 | 2 (0%) 92 86 | 80, 145, 200, 335 | 0 |
| 1 | B | 467/528 (88%) | -0.52 | 0 100 100 | 72, 130, 184, 231 | 0 |
| 1 | C | 450/528 (85%) | -0.24 | 7 (1%) 72 57 | 85, 190, 241, 272 | 0 |
| 1 | D | 468/528 (88%) | -0.60 | 1 (0%) 95 91 | 58, 114, 164, 227 | 0 |
| All | All | 1850/2112 (87%) | -0.44 | 10 (0%) 91 83 | 58, 135, 218, 335 | 0 |

The worst 5 of 10 RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|---|---|---|---|---|
| 1 | C | 349 | ASP | 3.0 |
| 1 | A | 208 | TYR | 2.7 |
| 1 | C | 55 | LEU | 2.7 |
| 1 | C | 336 | ALA | 2.5 |
| 1 | A | 209 | PRO | 2.4 |

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

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

## 6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

## 6.4 Ligands (i)

There are no ligands in this entry.

## 6.5 Other polymers (i)

There are no such residues in this entry.