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

Oct 5, 2023 – 06:51 AM EDT

| | | |
|---|---|---|
| PDB ID | : | 6VGE |
| Title | : | Crystal structure of the DNA binding domains of human transcription factor ERG, human Runx2 bound to core binding factor beta (Cbfb), in complex with 16mer DNA CAGAGGATGTGGCTTC |
| Authors | : | Hou, C.; Tsodikov, O.V. |
| Deposited on | : | 2020-01-07 |
| Resolution | : | 4.25 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 | : | FAILED |
| 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 (i)

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

The reported resolution of this entry is 4.25 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 (i)

There are 5 unique types of molecules in this entry. The entry contains 3361 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 Transcriptional regulator ERG.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | | |
| 1 | A | 98 | 806 | 512 | 143 | 146 | 5 | 0 | 0 | 0 |

There are 14 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| A | 302 | GLY | - | expression tag | UNP P11308 |
| A | 303 | PRO | - | expression tag | UNP P11308 |
| A | 304 | HIS | - | expression tag | UNP P11308 |
| A | 305 | MET | - | expression tag | UNP P11308 |
| A | 420 | SER | - | expression tag | UNP P11308 |
| A | 421 | TYR | - | expression tag | UNP P11308 |
| A | 422 | HIS | - | expression tag | UNP P11308 |
| A | 423 | ALA | - | expression tag | UNP P11308 |
| A | 424 | HIS | - | expression tag | UNP P11308 |
| A | 425 | PRO | - | expression tag | UNP P11308 |
| A | 426 | GLN | - | expression tag | UNP P11308 |
| A | 427 | LYS | - | expression tag | UNP P11308 |
| A | 428 | MET | - | expression tag | UNP P11308 |
| A | 429 | ASN | - | expression tag | UNP P11308 |

- Molecule 2 is a DNA chain called DNA (5'-D(P*CP*AP*GP*AP*GP*GP*AP*TP*GP*TP*GP*GP*CP*TP*TP*C)-3').

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | P | | | |
| 2 | B | 16 | 332 | 157 | 62 | 97 | 16 | 0 | 0 | 0 |

- Molecule 3 is a DNA chain called DNA (5'-D(P*GP*AP*AP*GP*CP*CP*AP*CP*AP*TP*CP*CP*TP*CP*TP*G)-3').

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | P | | | |
| 3 | C | 16 | 324 | 154 | 59 | 95 | 16 | 0 | 0 | 0 |

- Molecule 4 is a protein called Runt-related transcription factor 2.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | | |
| 4 | D | 117 | 901 | 567 | 164 | 166 | 4 | 0 | 0 | 0 |

- Molecule 5 is a protein called Core-binding factor subunit beta.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | | |
| 5 | G | 121 | 998 | 622 | 181 | 189 | 6 | 0 | 0 | 0 |

There are 14 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| G | -13 | MET | - | initiating methionine | UNP Q13951 |
| G | -12 | GLY | - | expression tag | UNP Q13951 |
| G | -11 | SER | - | expression tag | UNP Q13951 |
| G | -10 | SER | - | expression tag | UNP Q13951 |
| G | -9 | HIS | - | expression tag | UNP Q13951 |
| G | -8 | HIS | - | expression tag | UNP Q13951 |
| G | -7 | HIS | - | expression tag | UNP Q13951 |
| G | -6 | HIS | - | expression tag | UNP Q13951 |
| G | -5 | HIS | - | expression tag | UNP Q13951 |
| G | -4 | HIS | - | expression tag | UNP Q13951 |
| G | -3 | SER | - | expression tag | UNP Q13951 |
| G | -2 | GLN | - | expression tag | UNP Q13951 |
| G | -1 | ASP | - | expression tag | UNP Q13951 |
| G | 0 | PRO | - | expression tag | UNP Q13951 |

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 62 2 2 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 103.92Å 103.92Å 322.91Å<br>90.00° 90.00° 120.00° | Depositor |
| Resolution (Å) | 35.00 – 4.25 | Depositor |
| % Data completeness<br>(in resolution range) | 95.0 (35.00-4.25) | Depositor |
| $R_{merge}$ | 0.10 | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 1.70 (at 4.28Å) | Xtriage |
| Refinement program | REFMAC 5.8.0238 | Depositor |
| R, $R_{free}$ | 0.312 , 0.317 | Depositor |
| Wilson B-factor (Å$^2$) | 194.8 | Xtriage |
| Anisotropy | 0.477 | Xtriage |
| L-test for twinning[2] | $< \|L\| > = 0.44$, $< L^2 > = 0.27$ | Xtriage |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| Total number of atoms | 3361 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 303.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.82% 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 (i)

## 4.1 Standard geometry (i)

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

## 4.2 Too-close contacts (i)

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

## 4.3 Torsion angles (i)

### 4.3.1 Protein backbone (i)

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

### 4.3.2 Protein sidechains (i)

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

### 4.3.3 RNA (i)

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

## 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)

There are no ligands in this entry.

## 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 (i)

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

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

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

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

## 5.3 Carbohydrates (i)

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

## 5.4 Ligands (i)

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

## 5.5 Other polymers (i)

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