



wwPDB X-ray Structure Validation Summary Report i

Feb 27, 2014 – 04:42 PM GMT

PDB ID : 2F7K
Title : Crystal Structure of Human Pyridoxal Kinase
Authors : Jiang, T.; Cao, P.; Gong, Y.; Tang, L.
Deposited on : 2005-12-01
Resolution : 2.80 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

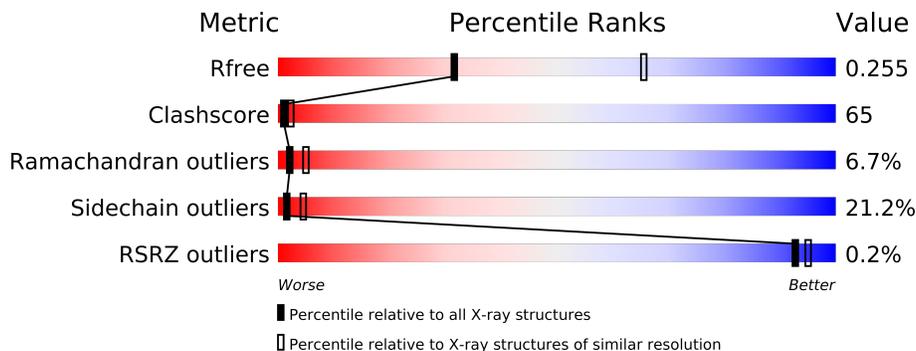
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.80 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} | 66092 | 1799 (2.80-2.80) |
| Clashscore | 79885 | 2295 (2.80-2.80) |
| Ramachandran outliers | 78287 | 2252 (2.80-2.80) |
| Sidechain outliers | 78261 | 2254 (2.80-2.80) |
| RSRZ outliers | 66119 | 1802 (2.80-2.80) |

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|------------------|
| 1 | A | 327 | |
| 1 | B | 327 | |

2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 5327 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Pyridoxal kinase.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|------|-----|-----|----|---------|---------|-------|
| | | | Total | C | N | O | S | | | |
| 1 | A | 323 | 2560 | 1608 | 458 | 478 | 16 | 0 | 0 | 0 |
| 1 | B | 323 | 2560 | 1608 | 458 | 478 | 16 | 0 | 0 | 0 |

There are 30 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|------------------|------------|
| A | -14 | SER | - | CLONING ARTIFACT | UNP O00764 |
| A | -13 | TYR | - | CLONING ARTIFACT | UNP O00764 |
| A | -12 | TYR | - | CLONING ARTIFACT | UNP O00764 |
| A | -11 | HIS | - | EXPRESSION TAG | UNP O00764 |
| A | -10 | HIS | - | EXPRESSION TAG | UNP O00764 |
| A | -9 | HIS | - | EXPRESSION TAG | UNP O00764 |
| A | -8 | HIS | - | EXPRESSION TAG | UNP O00764 |
| A | -7 | HIS | - | EXPRESSION TAG | UNP O00764 |
| A | -6 | HIS | - | EXPRESSION TAG | UNP O00764 |
| A | -5 | HIS | - | EXPRESSION TAG | UNP O00764 |
| A | -4 | GLU | - | CLONING ARTIFACT | UNP O00764 |
| A | -3 | GLY | - | CLONING ARTIFACT | UNP O00764 |
| A | -2 | VAL | - | CLONING ARTIFACT | UNP O00764 |
| A | -1 | ARG | - | CLONING ARTIFACT | UNP O00764 |
| A | 0 | THR | - | CLONING ARTIFACT | UNP O00764 |
| B | -14 | SER | - | CLONING ARTIFACT | UNP O00764 |
| B | -13 | TYR | - | CLONING ARTIFACT | UNP O00764 |
| B | -12 | TYR | - | CLONING ARTIFACT | UNP O00764 |
| B | -11 | HIS | - | EXPRESSION TAG | UNP O00764 |
| B | -10 | HIS | - | EXPRESSION TAG | UNP O00764 |
| B | -9 | HIS | - | EXPRESSION TAG | UNP O00764 |
| B | -8 | HIS | - | EXPRESSION TAG | UNP O00764 |
| B | -7 | HIS | - | EXPRESSION TAG | UNP O00764 |
| B | -6 | HIS | - | EXPRESSION TAG | UNP O00764 |
| B | -5 | HIS | - | EXPRESSION TAG | UNP O00764 |

Continued on next page...

Continued from previous page...

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|------------------|------------|
| B | -4 | GLU | - | CLONING ARTIFACT | UNP O00764 |
| B | -3 | GLY | - | CLONING ARTIFACT | UNP O00764 |
| B | -2 | VAL | - | CLONING ARTIFACT | UNP O00764 |
| B | -1 | ARG | - | CLONING ARTIFACT | UNP O00764 |
| B | 0 | THR | - | CLONING ARTIFACT | UNP O00764 |

- Molecule 2 is water.

| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|-----|-------|----------|--------------------|---------|---------|
| 2 | A | 96 | Total O 96 96 | 0 | 0 |
| 2 | B | 111 | Total O 111 111 | 0 | 0 |

4 Data and refinement statistics

| Property | Value | Source |
|---|---|------------------|
| Space group | P 43 | Depositor |
| Cell constants a, b, c, α , β , γ | 52.50Å 52.50Å 301.00Å 90.00° 90.00° 90.00° | Depositor |
| Resolution (Å) | 8.00 – 2.80 7.99 – 2.80 | Depositor EDS |
| % Data completeness (in resolution range) | 81.0 (8.00-2.80) 81.0 (7.99-2.80) | Depositor EDS |
| R_{merge} | 0.03 | Depositor |
| R_{sym} | (Not available) | Depositor |
| $\langle I/\sigma(I) \rangle$ ¹ | 4.94 (at 2.78Å) | Xtriage |
| Refinement program | CNS | Depositor |
| R, R_{free} | 0.228 , 0.271 0.213 , 0.255 | Depositor DCC |
| R_{free} test set | 707 reflections (4.58%) | DCC |
| Wilson B-factor (Å ²) | 68.3 | Xtriage |
| Anisotropy | 0.089 | Xtriage |
| Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²) | 0.40 , 75.0 | EDS |
| Estimated twinning fraction | 0.190 for h,-k,-l | Xtriage |
| L-test for twinning | $\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.36$ | Xtriage |
| Outliers | 0 of 15454 reflections | Xtriage |
| F_o, F_c correlation | 0.93 | EDS |
| Total number of atoms | 5327 | wwPDB-VP |
| Average B, all atoms (Å ²) | 44.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.72% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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 | 0.46 | 0/2612 | 0.87 | 10/3540 (0.3%) |
| 1 | B | 0.45 | 0/2612 | 0.83 | 3/3540 (0.1%) |
| All | All | 0.45 | 0/5224 | 0.85 | 13/7080 (0.2%) |

There are no bond length outliers.

The worst 5 of 13 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(°) | Ideal(°) |
|-----|-------|-----|------|----------|-------|-------------|----------|
| 1 | A | 142 | PRO | CA-N-CD | -7.94 | 100.39 | 111.50 |
| 1 | A | 149 | PRO | CA-N-CD | -7.75 | 100.66 | 111.50 |
| 1 | B | 120 | TRP | CA-CB-CG | 6.54 | 126.12 | 113.70 |
| 1 | A | -3 | GLY | N-CA-C | -6.51 | 96.83 | 113.10 |
| 1 | A | 192 | SER | N-CA-C | -5.85 | 95.22 | 111.00 |

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 1 | A | 2560 | 0 | 2551 | 354 | 0 |
| 1 | B | 2560 | 0 | 2551 | 334 | 0 |
| 2 | A | 96 | 0 | 0 | 0 | 0 |
| 2 | B | 111 | 0 | 0 | 3 | 0 |
| All | All | 5327 | 0 | 5102 | 660 | 0 |

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 65.

The worst 5 of 660 close contacts within the same asymmetric unit are listed below.

| Atom-1 | Atom-2 | Distance(Å) | Clash(Å) |
|-----------------|------------------|-------------|----------|
| 1:A:199:LEU:HB2 | 1:A:223:ILE:HB | 1.38 | 1.05 |
| 1:A:72:ASN:HD21 | 1:A:74:MET:HG3 | 1.22 | 1.04 |
| 1:A:221:MET:HG2 | 1:A:309:ALA:HA | 1.39 | 1.02 |
| 1:A:147:ILE:H | 1:A:147:ILE:CD1 | 1.74 | 1.00 |
| 1:A:147:ILE:H | 1:A:147:ILE:HD13 | 1.24 | 1.00 |

There are no symmetry-related clashes.

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 | 321/327 (98%) | 229 (71%) | 65 (20%) | 27 (8%) | 1 3 |
| 1 | B | 321/327 (98%) | 240 (75%) | 65 (20%) | 16 (5%) | 3 10 |
| All | All | 642/654 (98%) | 469 (73%) | 130 (20%) | 43 (7%) | 2 5 |

5 of 43 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | A | -9 | HIS |
| 1 | A | 59 | SER |
| 1 | A | 109 | VAL |
| 1 | A | 195 | GLY |
| 1 | A | 210 | 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 | 285/289 (99%) | 226 (79%) | 59 (21%) | 2 | 5 |
| 1 | B | 285/289 (99%) | 223 (78%) | 62 (22%) | 1 | 4 |
| All | All | 570/578 (99%) | 449 (79%) | 121 (21%) | 1 | 4 |

5 of 121 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | A | 292 | ARG |
| 1 | B | 5 | CYS |
| 1 | B | 273 | CYS |
| 1 | A | 305 | ILE |
| 1 | B | -8 | HIS |

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | A | 268 | GLN |
| 1 | B | -10 | HIS |
| 1 | B | 209 | ASN |
| 1 | A | 288 | GLN |
| 1 | A | 308 | GLN |

5.3.3 RNA [i](#)

There are no RNA chains 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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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 [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, 95th 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 | 323/327 (98%) | -0.21 | 1 (0%) 91 93 | 36, 44, 50, 56 | 0 |
| 1 | B | 323/327 (98%) | -0.20 | 0 100 100 | 35, 44, 50, 55 | 0 |
| All | All | 646/654 (98%) | -0.20 | 1 (0%) 93 95 | 35, 44, 50, 56 | 0 |

All (1) RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|-----|-------|-----|------|------|
| 1 | A | 186 | THR | 2.1 |

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 carbohydrates 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.