Full wwPDB X-ray Structure Validation Report ⓘ

Oct 1, 2023 – 10:40 PM EDT

PDB ID	:	6MJK
Title	:	Crystal Structure of dUTP pyrophosphatase protein, from Naegleria fowleri in complex with deoxyuridine
Authors	:	Seattle Structural Genomics Center for Infectious Disease (SSGCID)
Deposited on	:	2018-09-21
Resolution	:	1.45 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references ⓘ) 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 (i)

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

The reported resolution of this entry is 1.45 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 1343 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 dUTP pyrophosphatase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	144	Total	C	N	O	S	0	9	0
			1125	719	188	214	4			

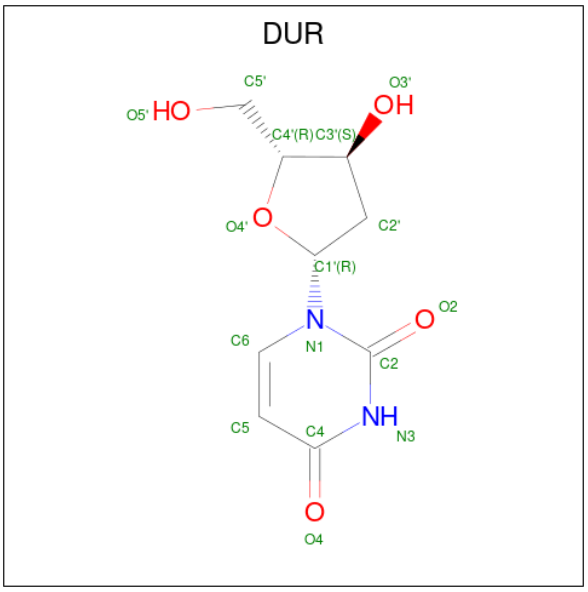
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	MET	-	expression tag	UNP A0A1Z0YU86
A	-6	ALA	-	expression tag	UNP A0A1Z0YU86
A	-5	HIS	-	expression tag	UNP A0A1Z0YU86
A	-4	HIS	-	expression tag	UNP A0A1Z0YU86
A	-3	HIS	-	expression tag	UNP A0A1Z0YU86
A	-2	HIS	-	expression tag	UNP A0A1Z0YU86
A	-1	HIS	-	expression tag	UNP A0A1Z0YU86
A	0	HIS	-	expression tag	UNP A0A1Z0YU86

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

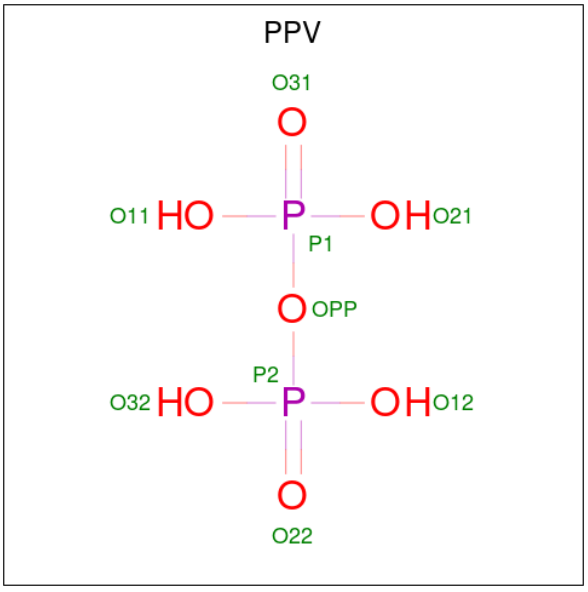
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Mg	0	0
			2	2		

- Molecule 3 is 2'-DEOXYURIDINE (three-letter code: DUR) (formula: $C_9H_{12}N_2O_5$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	1
			32	18	4	10		

- Molecule 4 is PYROPHOSPHATE (three-letter code: PPV) (formula: $H_4O_7P_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	P	0	0
			9	7	2		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	166	Total O 175 175	0	9

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 21 3	Depositor
Cell constants a, b, c, α, β, γ	74.23Å 74.23Å 74.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.20 – 1.45	Depositor
% Data completeness (in resolution range)	100.0 (33.20-1.45)	Depositor
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.27 (at 1.45Å)	Xtriage
Refinement program	PHENIX (1.14rc2_3191: ???)	Depositor
R, R_{free}	0.134 , 0.170	Depositor
Wilson B-factor (Å^2)	14.3	Xtriage
Anisotropy	0.000	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.040 for l,-k,h	Xtriage
Total number of atoms	1343	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.66% 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

Of 5 ligands modelled in this entry, 2 are monoatomic - leaving 3 for Mogul analysis.

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	$\#\|Z\| > 2$
3	DUR	A	203[A]	-	17,17,17	0.19	0	24,24,24	0.24	0
4	PPV	A	204	2	6,8,8	0.78	0	13,13,13	1.16	1 (7%)
3	DUR	A	203[B]	-	17,17,17	0.16	0	24,24,24	0.24	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
3	DUR	A	203[A]	-	-	0/6/18/18	0/2/2/2
4	PPV	A	204	2	-	2/6/6/6	-
3	DUR	A	203[B]	-	-	0/6/18/18	0/2/2/2

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	A	204	PPV	P2-OPP-P1	-3.04	122.41	132.83

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	204	PPV	P2-OPP-P1-O11
4	A	204	PPV	P2-OPP-P1-O21

There are no ring outliers.

No monomer is involved in short contacts.

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.