



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 06:09 am BST

PDB ID : 5WU4  
Title : Crystal structure of human Tut1 bound with MgATP, form II  
Authors : Yamashita, S.; Tomita, K.  
Deposited on : 2016-12-16  
Resolution : 2.80 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](mailto: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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
buster-report : 1.1.7 (2018)  
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) : 2.11

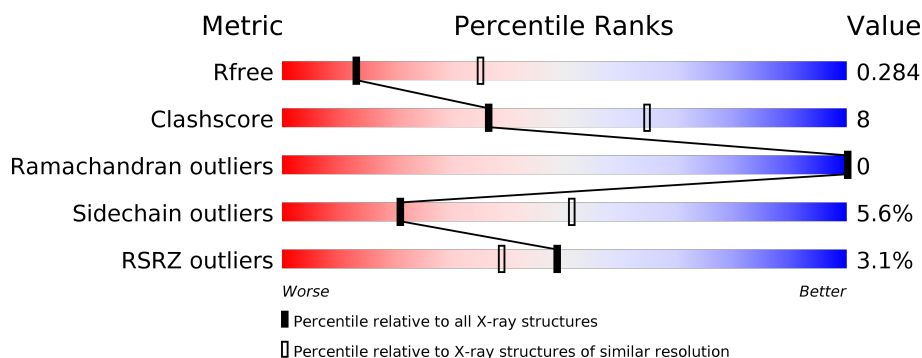
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 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))
$R_{free}$	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (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  $\geq 3$ , 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  $\leq 5\%$ . 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	573	<div> <div>3%</div> <div>68%</div> <div>15%</div> <div>15%</div> </div>
1	B	573	<div> <div>2%</div> <div>64%</div> <div>19%</div> <div>16%</div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 7595 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 Speckle targeted PIP5K1A-regulated poly(A) polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	485	Total	C	N	O	S	0	0	0
			3770	2385	687	684	14			
1	B	483	Total	C	N	O	S	0	0	0
			3757	2378	685	680	14			

There are 366 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	140	MET	-	initiating methionine	UNP Q9H6E5
A	?	-	ALA	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5
A	?	-	GLU	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	?	-	LEU	deletion	UNP Q9H6E5
A	?	-	ASP	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	?	-	ALA	deletion	UNP Q9H6E5
A	?	-	LEU	deletion	UNP Q9H6E5
A	?	-	ALA	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5
A	?	-	LEU	deletion	UNP Q9H6E5
A	?	-	ASP	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5
A	?	-	GLN	deletion	UNP Q9H6E5
A	?	-	ALA	deletion	UNP Q9H6E5
A	?	-	LEU	deletion	UNP Q9H6E5
A	?	-	ALA	deletion	UNP Q9H6E5
A	?	-	CYS	deletion	UNP Q9H6E5
A	?	-	THR	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	ALA	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5
A	?	-	ASP	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	?	-	GLN	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5
A	?	-	ALA	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5
A	?	-	GLN	deletion	UNP Q9H6E5
A	?	-	ASP	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	?	-	GLU	deletion	UNP Q9H6E5
A	?	-	ALA	deletion	UNP Q9H6E5
A	?	-	LEU	deletion	UNP Q9H6E5
A	?	-	ASP	deletion	UNP Q9H6E5
A	?	-	PHE	deletion	UNP Q9H6E5
A	?	-	GLU	deletion	UNP Q9H6E5
A	?	-	THR	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	?	-	LEU	deletion	UNP Q9H6E5
A	?	-	ALA	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5
A	?	-	GLN	deletion	UNP Q9H6E5
A	?	-	THR	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5
A	?	-	ASP	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	?	-	ALA	deletion	UNP Q9H6E5
A	?	-	LEU	deletion	UNP Q9H6E5
A	?	-	ALA	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	?	-	GLU	deletion	UNP Q9H6E5
A	?	-	THR	deletion	UNP Q9H6E5
A	?	-	LEU	deletion	UNP Q9H6E5
A	?	-	ALA	deletion	UNP Q9H6E5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	SER	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5
A	?	-	GLN	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	372	ALA	CYS	engineered mutation	UNP Q9H6E5
A	415	ALA	CYS	engineered mutation	UNP Q9H6E5
A	501	ALA	CYS	engineered mutation	UNP Q9H6E5
A	504	SER	CYS	engineered mutation	UNP Q9H6E5
A	?	-	GLU	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	?	-	GLN	deletion	UNP Q9H6E5
A	?	-	GLY	deletion	UNP Q9H6E5
A	?	-	GLY	deletion	UNP Q9H6E5
A	?	-	THR	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	?	-	LYS	deletion	UNP Q9H6E5
A	?	-	ARG	deletion	UNP Q9H6E5
A	?	-	LEU	deletion	UNP Q9H6E5
A	?	-	LYS	deletion	UNP Q9H6E5
A	?	-	VAL	deletion	UNP Q9H6E5
A	?	-	ASP	deletion	UNP Q9H6E5
A	?	-	GLY	deletion	UNP Q9H6E5
A	?	-	GLN	deletion	UNP Q9H6E5
A	?	-	LYS	deletion	UNP Q9H6E5
A	?	-	ASN	deletion	UNP Q9H6E5
A	?	-	CYS	deletion	UNP Q9H6E5
A	?	-	CYS	deletion	UNP Q9H6E5
A	?	-	GLU	deletion	UNP Q9H6E5
A	?	-	GLU	deletion	UNP Q9H6E5
A	?	-	GLY	deletion	UNP Q9H6E5
A	?	-	LYS	deletion	UNP Q9H6E5
A	?	-	GLU	deletion	UNP Q9H6E5
A	?	-	GLU	deletion	UNP Q9H6E5
A	?	-	GLN	deletion	UNP Q9H6E5
A	?	-	GLN	deletion	UNP Q9H6E5
A	?	-	GLY	deletion	UNP Q9H6E5
A	?	-	CYS	deletion	UNP Q9H6E5
A	?	-	ALA	deletion	UNP Q9H6E5
A	?	-	GLY	deletion	UNP Q9H6E5
A	?	-	ASP	deletion	UNP Q9H6E5
A	?	-	GLY	deletion	UNP Q9H6E5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	GLY	deletion	UNP Q9H6E5
A	?	-	GLU	deletion	UNP Q9H6E5
A	?	-	ASP	deletion	UNP Q9H6E5
A	?	-	ARG	deletion	UNP Q9H6E5
A	?	-	VAL	deletion	UNP Q9H6E5
A	?	-	GLU	deletion	UNP Q9H6E5
A	?	-	GLU	deletion	UNP Q9H6E5
A	?	-	MET	deletion	UNP Q9H6E5
A	?	-	VAL	deletion	UNP Q9H6E5
A	?	-	ILE	deletion	UNP Q9H6E5
A	?	-	GLU	deletion	UNP Q9H6E5
A	?	-	VAL	deletion	UNP Q9H6E5
A	?	-	GLY	deletion	UNP Q9H6E5
A	?	-	GLU	deletion	UNP Q9H6E5
A	?	-	MET	deletion	UNP Q9H6E5
A	?	-	VAL	deletion	UNP Q9H6E5
A	?	-	GLN	deletion	UNP Q9H6E5
A	?	-	ASP	deletion	UNP Q9H6E5
A	?	-	TRP	deletion	UNP Q9H6E5
A	?	-	ALA	deletion	UNP Q9H6E5
A	?	-	MET	deletion	UNP Q9H6E5
A	?	-	GLN	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5
A	?	-	GLY	deletion	UNP Q9H6E5
A	?	-	GLN	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5
A	?	-	GLY	deletion	UNP Q9H6E5
A	?	-	ASP	deletion	UNP Q9H6E5
A	?	-	LEU	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5
A	?	-	LEU	deletion	UNP Q9H6E5
A	?	-	THR	deletion	UNP Q9H6E5
A	?	-	THR	deletion	UNP Q9H6E5
A	?	-	GLY	deletion	UNP Q9H6E5
A	?	-	LYS	deletion	UNP Q9H6E5
A	?	-	HIS	deletion	UNP Q9H6E5
A	?	-	GLY	deletion	UNP Q9H6E5
A	?	-	ALA	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5
A	?	-	GLY	deletion	UNP Q9H6E5
A	?	-	GLU	deletion	UNP Q9H6E5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	GLU	deletion	UNP Q9H6E5
A	?	-	GLY	deletion	UNP Q9H6E5
A	?	-	GLN	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	?	-	HIS	deletion	UNP Q9H6E5
A	?	-	ALA	deletion	UNP Q9H6E5
A	?	-	ALA	deletion	UNP Q9H6E5
A	?	-	LEU	deletion	UNP Q9H6E5
A	?	-	ALA	deletion	UNP Q9H6E5
A	?	-	GLU	deletion	UNP Q9H6E5
A	?	-	ARG	deletion	UNP Q9H6E5
A	?	-	GLY	deletion	UNP Q9H6E5
A	?	-	PRO	deletion	UNP Q9H6E5
A	?	-	LYS	deletion	UNP Q9H6E5
A	?	-	GLY	deletion	UNP Q9H6E5
A	?	-	HIS	deletion	UNP Q9H6E5
A	?	-	GLU	deletion	UNP Q9H6E5
A	?	-	ALA	deletion	UNP Q9H6E5
A	?	-	ALA	deletion	UNP Q9H6E5
A	?	-	GLN	deletion	UNP Q9H6E5
A	?	-	GLU	deletion	UNP Q9H6E5
A	?	-	TRP	deletion	UNP Q9H6E5
A	?	-	SER	deletion	UNP Q9H6E5
A	875	LEU	-	expression tag	UNP Q9H6E5
A	876	GLU	-	expression tag	UNP Q9H6E5
A	877	HIS	-	expression tag	UNP Q9H6E5
A	878	HIS	-	expression tag	UNP Q9H6E5
A	879	HIS	-	expression tag	UNP Q9H6E5
A	880	HIS	-	expression tag	UNP Q9H6E5
A	881	HIS	-	expression tag	UNP Q9H6E5
A	882	HIS	-	expression tag	UNP Q9H6E5
B	140	MET	-	initiating methionine	UNP Q9H6E5
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5
B	?	-	GLU	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	?	-	LEU	deletion	UNP Q9H6E5
B	?	-	ASP	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	LEU	deletion	UNP Q9H6E5
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5
B	?	-	LEU	deletion	UNP Q9H6E5
B	?	-	ASP	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5
B	?	-	GLN	deletion	UNP Q9H6E5
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	LEU	deletion	UNP Q9H6E5
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	CYS	deletion	UNP Q9H6E5
B	?	-	THR	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5
B	?	-	ASP	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	?	-	GLN	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5
B	?	-	GLN	deletion	UNP Q9H6E5
B	?	-	ASP	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	?	-	GLU	deletion	UNP Q9H6E5
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	LEU	deletion	UNP Q9H6E5
B	?	-	ASP	deletion	UNP Q9H6E5
B	?	-	PHE	deletion	UNP Q9H6E5
B	?	-	GLU	deletion	UNP Q9H6E5
B	?	-	THR	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	?	-	LEU	deletion	UNP Q9H6E5

*Continued on next page...*



*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5
B	?	-	GLN	deletion	UNP Q9H6E5
B	?	-	THR	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5
B	?	-	ASP	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	LEU	deletion	UNP Q9H6E5
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	?	-	GLU	deletion	UNP Q9H6E5
B	?	-	THR	deletion	UNP Q9H6E5
B	?	-	LEU	deletion	UNP Q9H6E5
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5
B	?	-	GLN	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	372	ALA	CYS	engineered mutation	UNP Q9H6E5
B	415	ALA	CYS	engineered mutation	UNP Q9H6E5
B	501	ALA	CYS	engineered mutation	UNP Q9H6E5
B	504	SER	CYS	engineered mutation	UNP Q9H6E5
B	?	-	GLU	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	?	-	GLN	deletion	UNP Q9H6E5
B	?	-	GLY	deletion	UNP Q9H6E5
B	?	-	GLY	deletion	UNP Q9H6E5
B	?	-	THR	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	?	-	LYS	deletion	UNP Q9H6E5
B	?	-	ARG	deletion	UNP Q9H6E5
B	?	-	LEU	deletion	UNP Q9H6E5
B	?	-	LYS	deletion	UNP Q9H6E5
B	?	-	VAL	deletion	UNP Q9H6E5
B	?	-	ASP	deletion	UNP Q9H6E5
B	?	-	GLY	deletion	UNP Q9H6E5
B	?	-	GLN	deletion	UNP Q9H6E5
B	?	-	LYS	deletion	UNP Q9H6E5
B	?	-	ASN	deletion	UNP Q9H6E5
B	?	-	CYS	deletion	UNP Q9H6E5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	CYS	deletion	UNP Q9H6E5
B	?	-	GLU	deletion	UNP Q9H6E5
B	?	-	GLU	deletion	UNP Q9H6E5
B	?	-	GLY	deletion	UNP Q9H6E5
B	?	-	LYS	deletion	UNP Q9H6E5
B	?	-	GLU	deletion	UNP Q9H6E5
B	?	-	GLU	deletion	UNP Q9H6E5
B	?	-	GLN	deletion	UNP Q9H6E5
B	?	-	GLN	deletion	UNP Q9H6E5
B	?	-	GLY	deletion	UNP Q9H6E5
B	?	-	CYS	deletion	UNP Q9H6E5
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	GLY	deletion	UNP Q9H6E5
B	?	-	ASP	deletion	UNP Q9H6E5
B	?	-	GLY	deletion	UNP Q9H6E5
B	?	-	GLY	deletion	UNP Q9H6E5
B	?	-	GLU	deletion	UNP Q9H6E5
B	?	-	ASP	deletion	UNP Q9H6E5
B	?	-	ARG	deletion	UNP Q9H6E5
B	?	-	VAL	deletion	UNP Q9H6E5
B	?	-	GLU	deletion	UNP Q9H6E5
B	?	-	GLU	deletion	UNP Q9H6E5
B	?	-	MET	deletion	UNP Q9H6E5
B	?	-	VAL	deletion	UNP Q9H6E5
B	?	-	ILE	deletion	UNP Q9H6E5
B	?	-	GLU	deletion	UNP Q9H6E5
B	?	-	VAL	deletion	UNP Q9H6E5
B	?	-	GLY	deletion	UNP Q9H6E5
B	?	-	GLU	deletion	UNP Q9H6E5
B	?	-	MET	deletion	UNP Q9H6E5
B	?	-	VAL	deletion	UNP Q9H6E5
B	?	-	GLN	deletion	UNP Q9H6E5
B	?	-	ASP	deletion	UNP Q9H6E5
B	?	-	TRP	deletion	UNP Q9H6E5
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	MET	deletion	UNP Q9H6E5
B	?	-	GLN	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5
B	?	-	GLY	deletion	UNP Q9H6E5
B	?	-	GLN	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5

*Continued on next page...*

*Continued from previous page...*

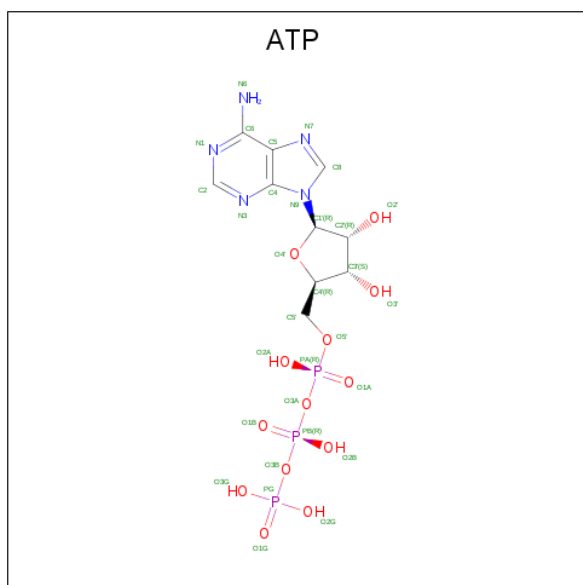
Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	GLY	deletion	UNP Q9H6E5
B	?	-	ASP	deletion	UNP Q9H6E5
B	?	-	LEU	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5
B	?	-	LEU	deletion	UNP Q9H6E5
B	?	-	THR	deletion	UNP Q9H6E5
B	?	-	THR	deletion	UNP Q9H6E5
B	?	-	GLY	deletion	UNP Q9H6E5
B	?	-	LYS	deletion	UNP Q9H6E5
B	?	-	HIS	deletion	UNP Q9H6E5
B	?	-	GLY	deletion	UNP Q9H6E5
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5
B	?	-	GLY	deletion	UNP Q9H6E5
B	?	-	GLU	deletion	UNP Q9H6E5
B	?	-	GLU	deletion	UNP Q9H6E5
B	?	-	GLY	deletion	UNP Q9H6E5
B	?	-	GLN	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	?	-	HIS	deletion	UNP Q9H6E5
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	LEU	deletion	UNP Q9H6E5
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	GLU	deletion	UNP Q9H6E5
B	?	-	ARG	deletion	UNP Q9H6E5
B	?	-	GLY	deletion	UNP Q9H6E5
B	?	-	PRO	deletion	UNP Q9H6E5
B	?	-	LYS	deletion	UNP Q9H6E5
B	?	-	GLY	deletion	UNP Q9H6E5
B	?	-	HIS	deletion	UNP Q9H6E5
B	?	-	GLU	deletion	UNP Q9H6E5
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	ALA	deletion	UNP Q9H6E5
B	?	-	GLN	deletion	UNP Q9H6E5
B	?	-	GLU	deletion	UNP Q9H6E5
B	?	-	TRP	deletion	UNP Q9H6E5
B	?	-	SER	deletion	UNP Q9H6E5
B	875	LEU	-	expression tag	UNP Q9H6E5
B	876	GLU	-	expression tag	UNP Q9H6E5
B	877	HIS	-	expression tag	UNP Q9H6E5

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	878	HIS	-	expression tag	UNP Q9H6E5
B	879	HIS	-	expression tag	UNP Q9H6E5
B	880	HIS	-	expression tag	UNP Q9H6E5
B	881	HIS	-	expression tag	UNP Q9H6E5
B	882	HIS	-	expression tag	UNP Q9H6E5

- Molecule 2 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula:  $C_{10}H_{16}N_5O_{13}P_3$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	B	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Mg	0	0
			1	1		
3	A	1	Total	Mg	0	0
			1	1		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total 1	Cl 1	0	0
4	A	1	Total 1	Cl 1	0	0

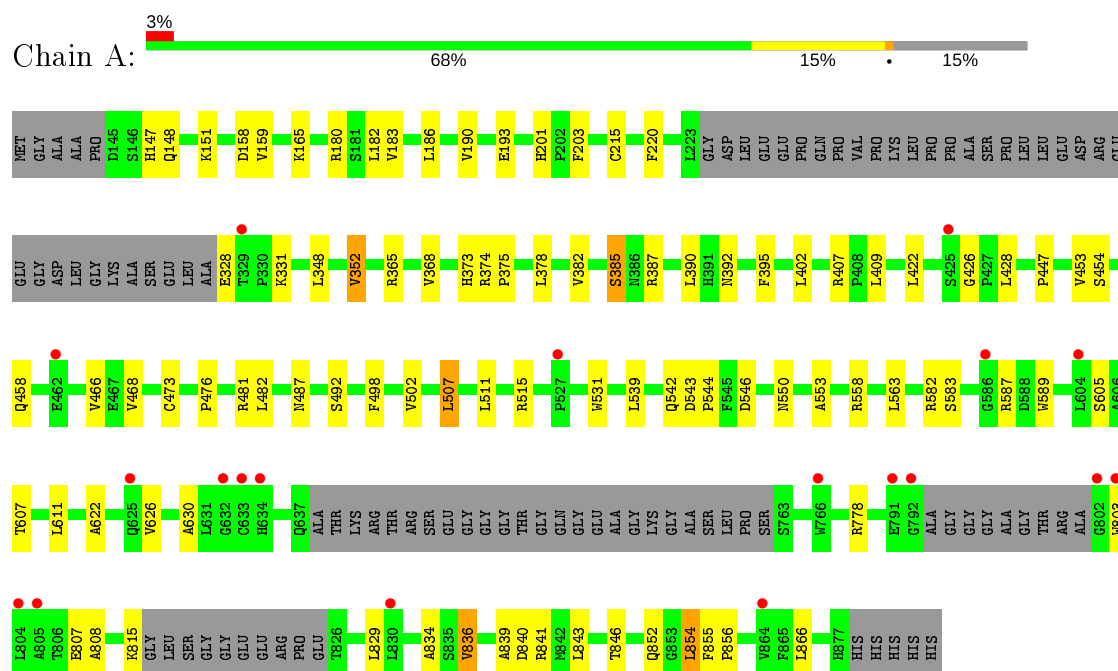
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total 1	O 1	0	0
5	B	1	Total 1	O 1	0	0

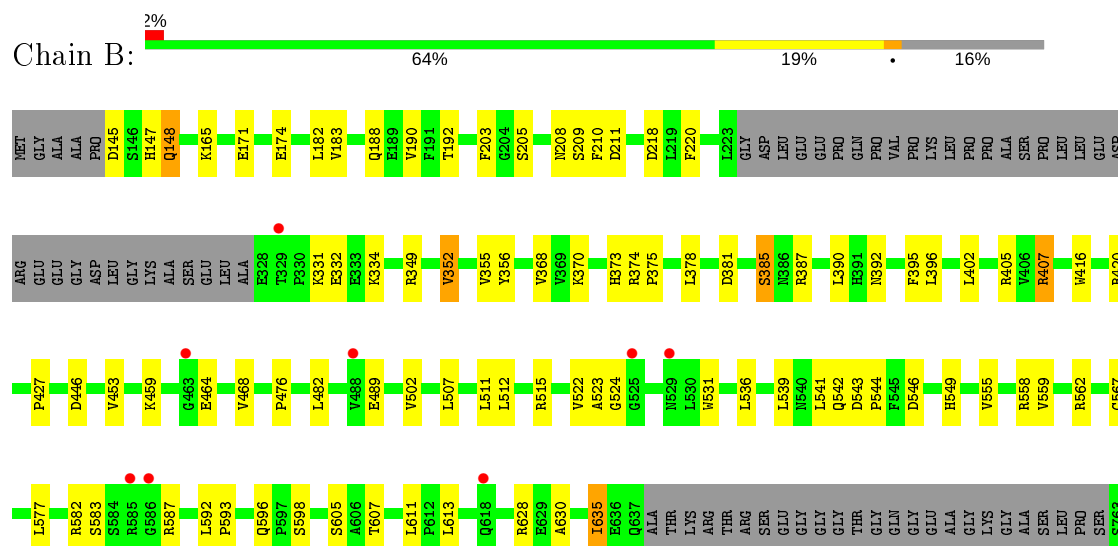
### 3 Residue-property plots [i](#)

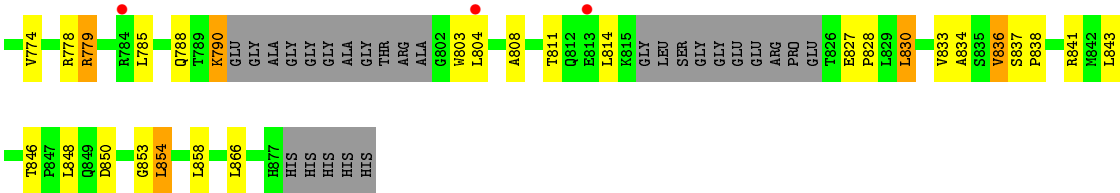
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Speckle targeted PIP5K1A-regulated poly(A) polymerase



- Molecule 1: Speckle targeted PIP5K1A-regulated poly(A) polymerase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	77.66Å 88.56Å 93.82Å 90.00° 98.66° 90.00°	Depositor
Resolution (Å)	19.98 – 2.80 46.85 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.8 (19.98-2.80) 99.8 (46.85-2.80)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.51 (at 2.81Å)	Xtriage
Refinement program	PHENIX 1.8.4_1496	Depositor
R, $R_{free}$	0.241 , 0.282 0.245 , 0.284	Depositor DCC
$R_{free}$ test set	1556 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	64.3	Xtriage
Anisotropy	0.151	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 43.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	7595	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	82.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.45% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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.



## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, ATP, CL

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	0.22	0/3853	0.41	0/5235
1	B	0.22	0/3840	0.41	0/5218
All	All	0.22	0/7693	0.41	0/10453

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	A	0	1
1	B	0	2
All	All	0	3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	385	SER	Peptide
1	B	385	SER	Peptide
1	B	446	ASP	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	3770	0	3785	48	0
1	B	3757	0	3776	65	0
2	A	31	0	12	7	0
2	B	31	0	12	5	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
All	All	7595	0	7585	120	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 8.

The worst 5 of 120 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:A:1001:ATP:H8	2:A:1001:ATP:H5'1	1.27	0.97
2:B:1001:ATP:H5'2	2:B:1001:ATP:H8	1.31	0.93
2:A:1001:ATP:H5'1	2:A:1001:ATP:C8	2.05	0.90
1:B:402:LEU:HD21	1:B:482:LEU:HD12	1.70	0.71
2:B:1001:ATP:H5'2	2:B:1001:ATP:C8	2.21	0.70

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	475/573 (83%)	454 (96%)	21 (4%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	473/573 (82%)	453 (96%)	20 (4%)	0	100	100
All	All	948/1146 (83%)	907 (96%)	41 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	410/469 (87%)	390 (95%)	20 (5%)	25	57
1	B	409/469 (87%)	383 (94%)	26 (6%)	17	45
All	All	819/938 (87%)	773 (94%)	46 (6%)	21	51

5 of 46 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	147	HIS
1	B	334	LYS
1	B	830	LEU
1	B	148	GLN
1	B	174	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	776	GLN

### 5.3.3 RNA ⓘ

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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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	$\# Z  > 2$
2	ATP	A	1001	3	26,33,33	1.12	1 (3%)	31,52,52	1.81	6 (19%)
2	ATP	B	1001	3	26,33,33	1.02	1 (3%)	31,52,52	1.83	7 (22%)

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
2	ATP	A	1001	3	-	7/18/38/38	0/3/3/3
2	ATP	B	1001	3	-	3/18/38/38	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1001	ATP	C2'-C1'	-3.12	1.49	1.53
2	B	1001	ATP	C5-C4	2.07	1.46	1.40

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1001	ATP	PB-O3B-PG	-5.11	115.30	132.83
2	A	1001	ATP	PB-O3B-PG	-4.72	116.64	132.83
2	A	1001	ATP	PA-O3A-PB	-4.40	117.74	132.83
2	A	1001	ATP	O4'-C1'-C2'	-4.10	100.93	106.93
2	B	1001	ATP	PA-O3A-PB	-3.96	119.24	132.83

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

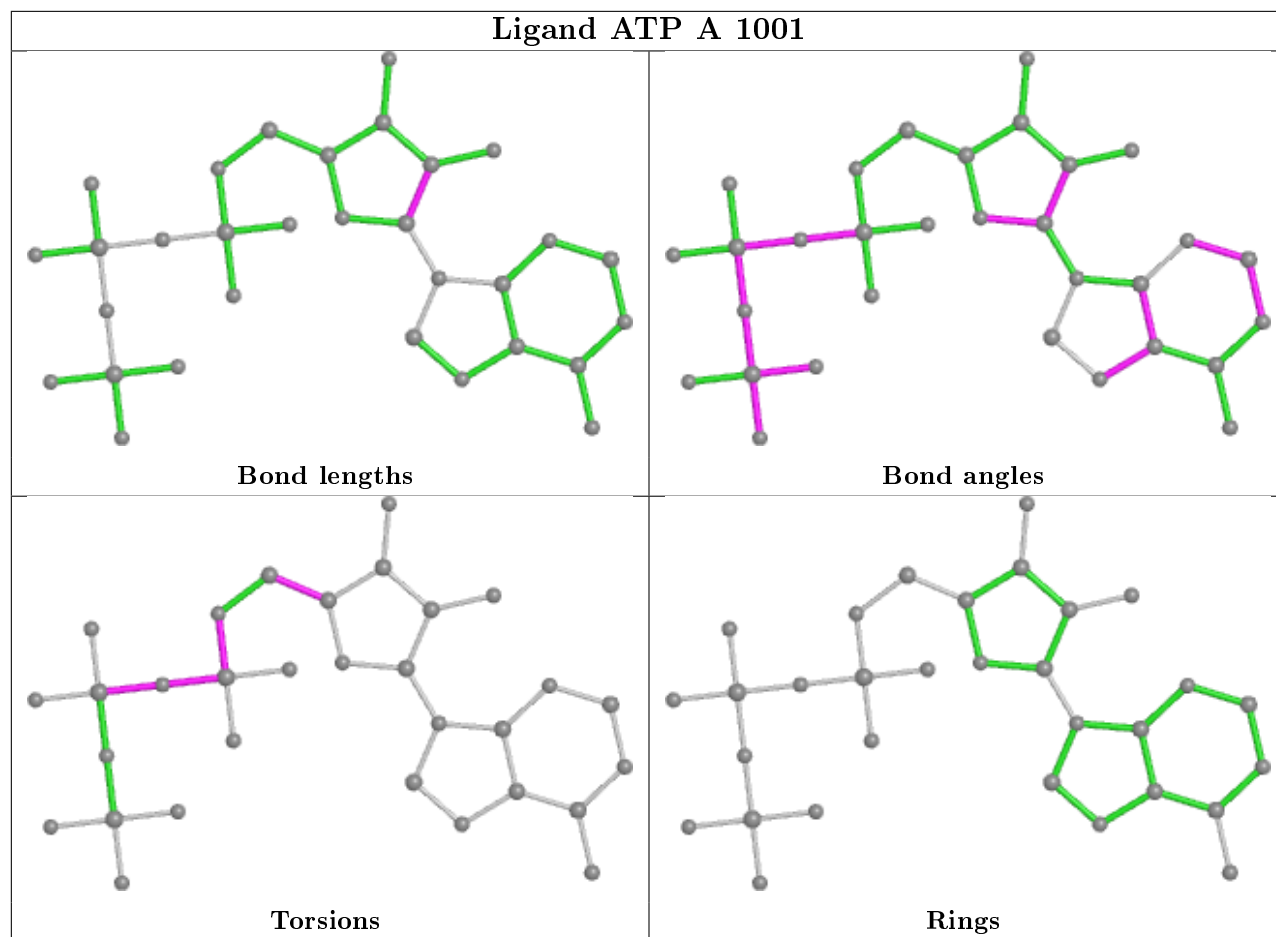
Mol	Chain	Res	Type	Atoms
2	A	1001	ATP	C5'-O5'-PA-O2A
2	B	1001	ATP	C3'-C4'-C5'-O5'
2	B	1001	ATP	O4'-C4'-C5'-O5'
2	A	1001	ATP	PB-O3A-PA-O5'
2	B	1001	ATP	PB-O3A-PA-O5'

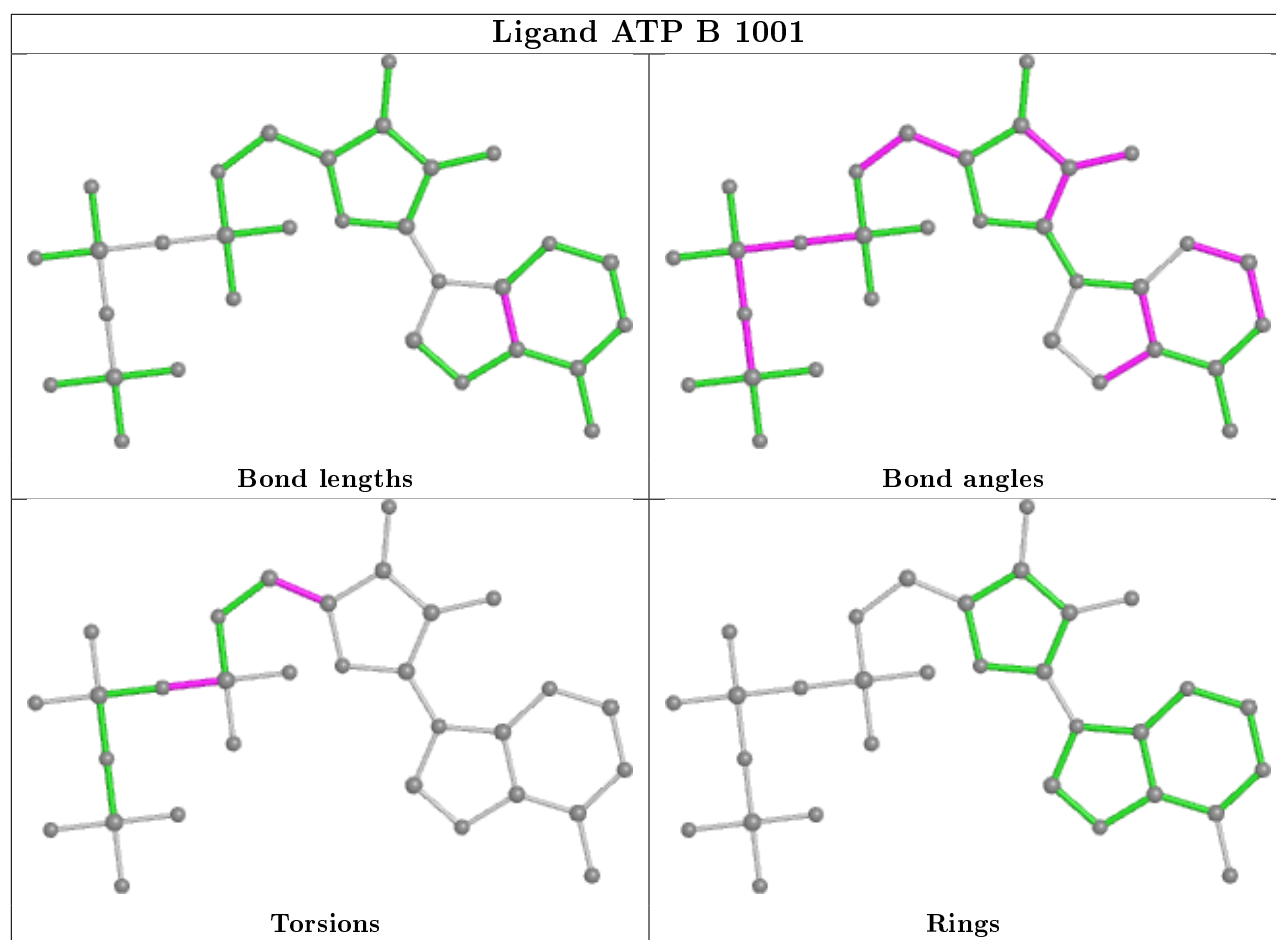
There are no ring outliers.

2 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1001	ATP	7	0
2	B	1001	ATP	5	0

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.





## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	485/573 (84%)	0.34	19 (3%) 39 29	35, 75, 139, 183	0
1	B	483/573 (84%)	0.31	11 (2%) 60 51	36, 75, 135, 171	0
All	All	968/1146 (84%)	0.32	30 (3%) 49 39	35, 75, 138, 183	0

The worst 5 of 30 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	586	GLY	7.3
1	A	792	GLY	5.3
1	A	791	GLU	5.2
1	B	329	THR	4.9
1	A	803	TRP	4.6

### 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.

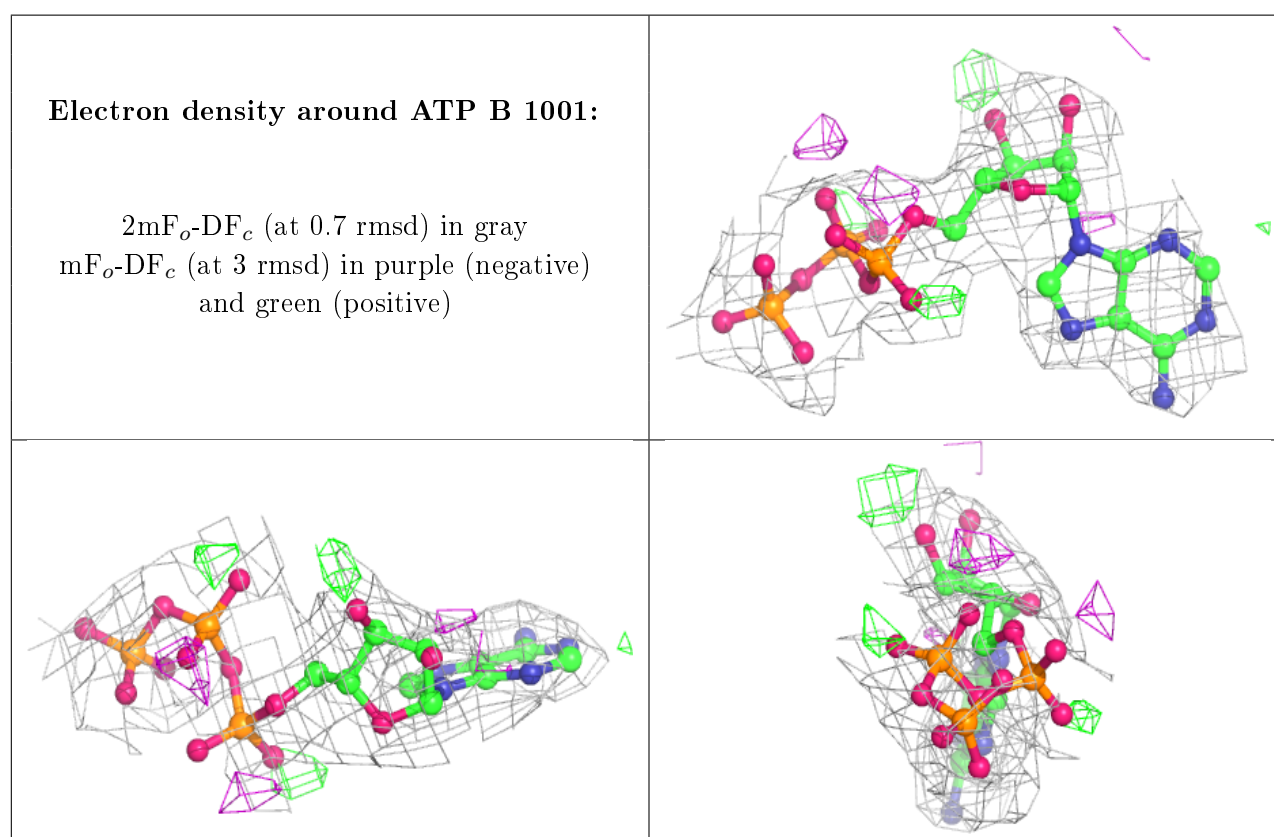
### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.



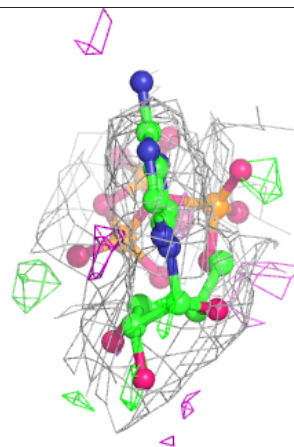
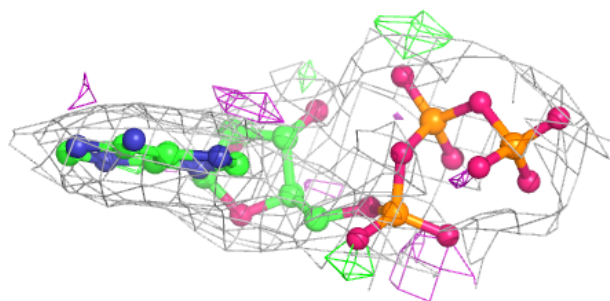
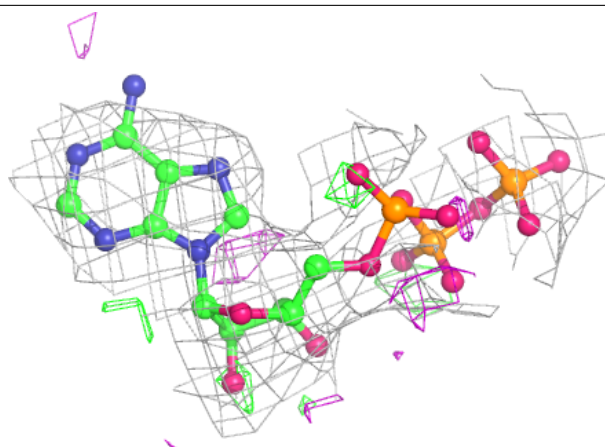
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	MG	B	1002	1/1	0.66	0.19	161,161,161,161	0
2	ATP	B	1001	31/31	0.72	0.25	140,143,270,270	0
3	MG	A	1002	1/1	0.72	0.10	129,129,129,129	0
2	ATP	A	1001	31/31	0.77	0.22	142,145,239,263	0
4	CL	B	1003	1/1	0.92	0.13	78,78,78,78	0
4	CL	A	1003	1/1	0.93	0.16	91,91,91,91	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



**Electron density around ATP A 1001:**

2mF<sub>o</sub>-DF<sub>c</sub> (at 0.7 rmsd) in gray  
mF<sub>o</sub>-DF<sub>c</sub> (at 3 rmsd) in purple (negative)  
and green (positive)



## 6.5 Other polymers ⓘ

There are no such residues in this entry.