



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 20, 2020 – 01:18 PM EST

PDB ID : 4WC4  
Title : tRNA-processing enzyme complex 2  
Authors : Yamashita, S.; Tomita, K.  
Deposited on : 2014-09-04  
Resolution : 3.50 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](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 : 1.8.0 (224370), CSD as540be (2019)  
Xtriage (Phenix) : 1.13  
EDS : 2.4  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.4

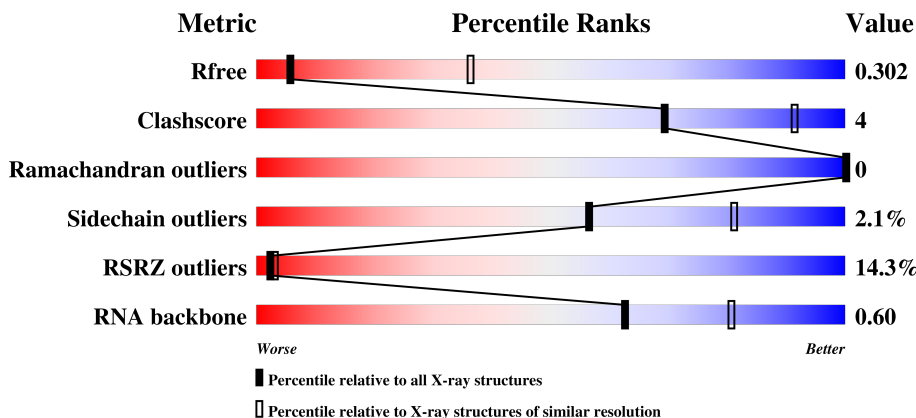
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.50 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))
$R_{free}$	111664	1391 (3.60-3.40)
Clashscore	122126	1485 (3.60-3.40)
Ramachandran outliers	120053	1446 (3.60-3.40)
Sidechain outliers	120020	1447 (3.60-3.40)
RSRZ outliers	108989	1303 (3.60-3.40)
RNA backbone	2636	1052 (4.10-2.90)

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. 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	396	 2% 81% 11% • 7%
2	B	74	 76% 58% 31% 11%

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4658 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 Poly A polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	367	Total	C	N	O	S	0	0	0
			3046	1978	527	534	7			

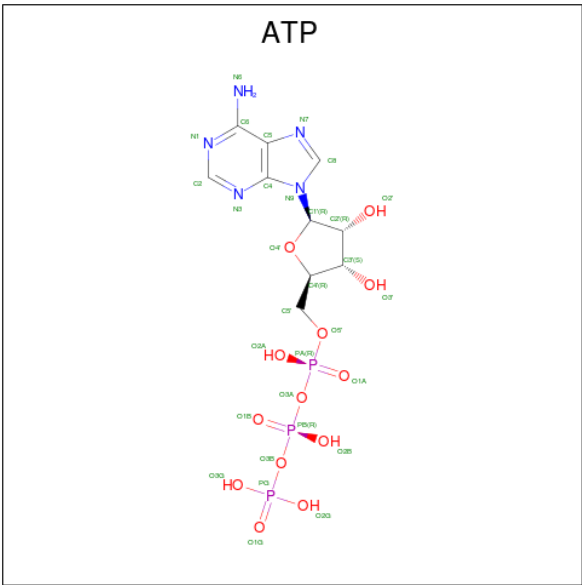
There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	expression tag	UNP O66728
A	384	LYS	-	expression tag	UNP O66728
A	385	LEU	-	expression tag	UNP O66728
A	386	ALA	-	expression tag	UNP O66728
A	387	ALA	-	expression tag	UNP O66728
A	388	ALA	-	expression tag	UNP O66728
A	389	LEU	-	expression tag	UNP O66728
A	390	GLU	-	expression tag	UNP O66728
A	391	HIS	-	expression tag	UNP O66728
A	392	HIS	-	expression tag	UNP O66728
A	393	HIS	-	expression tag	UNP O66728
A	394	HIS	-	expression tag	UNP O66728
A	395	HIS	-	expression tag	UNP O66728
A	396	HIS	-	expression tag	UNP O66728

- Molecule 2 is a RNA chain called RNA (74-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	74	Total	C	N	O	P	0	0	0
			1581	704	283	520	74			

- Molecule 3 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>5</sub>O<sub>13</sub>P<sub>3</sub>).

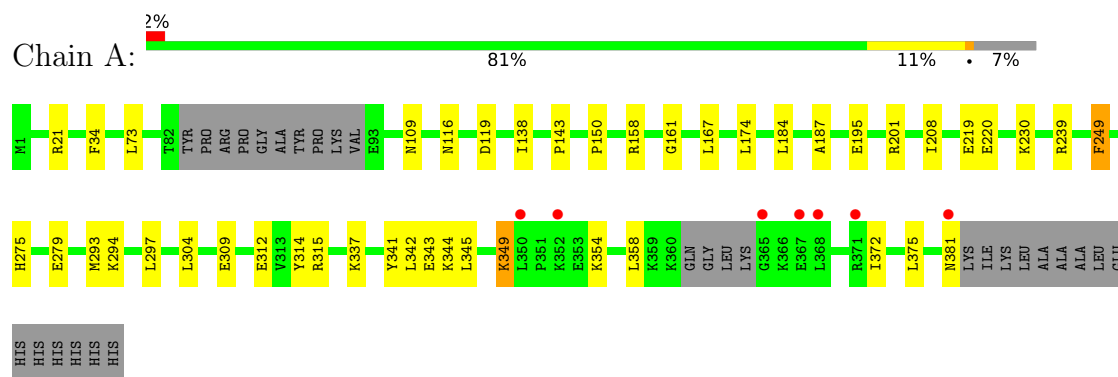


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

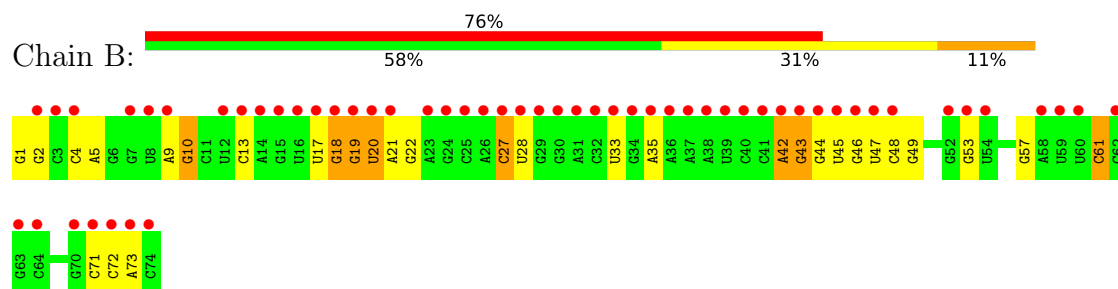
### 3 Residue-property plots [i](#)

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: Poly A polymerase



#### • Molecule 2: RNA (74-MER)



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	59.31Å 111.76Å 127.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.88 – 3.50 19.88 – 3.50	Depositor EDS
% Data completeness (in resolution range)	99.7 (19.88-3.50) 99.9 (19.88-3.50)	Depositor EDS
$R_{merge}$	0.22	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.08 (at 3.52Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.277 , 0.302 0.278 , 0.302	Depositor DCC
$R_{free}$ test set	558 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	95.7	Xtriage
Anisotropy	0.784	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 82.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	4658	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	224.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.79% 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

### 5.1 Standard geometry

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

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	0.21	0/3097	0.36	0/4150
2	B	0.28	1/1766 (0.1%)	0.67	0/2750
All	All	0.24	1/4863 (0.0%)	0.51	0/6900

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	G	OP3-P	-10.60	1.48	1.61

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts

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	3046	0	3205	24	0
2	B	1581	0	800	11	0
3	A	31	0	12	0	0
All	All	4658	0	4017	34	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 4.

All (34) 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:B:27:C:H42	2:B:43:G:H1	1.44	0.64
1:A:161:GLY:HA3	1:A:208:ILE:HG12	1.86	0.57
1:A:349:LYS:HB2	1:A:349:LYS:NZ	2.20	0.57
1:A:184:LEU:HD23	1:A:220:GLU:HB2	1.87	0.56
2:B:42:A:H5'	2:B:43:G:OP2	2.05	0.56
1:A:158:ARG:NE	1:A:195:GLU:OE1	2.40	0.55
1:A:372:ILE:HA	1:A:375:LEU:HD12	1.87	0.55
2:B:18:G:O2'	2:B:57:G:N2	2.39	0.55
1:A:293:MET:HG3	1:A:297:LEU:HD22	1.89	0.55
1:A:293:MET:HA	1:A:297:LEU:HB2	1.90	0.53
1:A:304:LEU:O	1:A:337:LYS:NZ	2.42	0.53
2:B:9:A:O2'	2:B:10:G:N7	2.41	0.53
1:A:309:GLU:OE2	1:A:381:ASN:ND2	2.42	0.52
1:A:354:LYS:O	1:A:358:LEU:HG	2.11	0.51
1:A:138:ILE:HB	1:A:167:LEU:HD23	1.93	0.51
2:B:33:U:H2'	2:B:35:A:OP2	2.15	0.47
1:A:219:GLU:N	1:A:219:GLU:OE1	2.42	0.47
1:A:201:ARG:HH22	2:B:2:G:H4'	1.80	0.47
1:A:358:LEU:HD13	1:A:375:LEU:HD13	1.97	0.47
1:A:150:PRO:HB2	1:A:187:ALA:HB2	1.97	0.46
1:A:312:GLU:HG3	1:A:315:ARG:HH21	1.82	0.45
1:A:341:TYR:HD1	1:A:345:LEU:HD22	1.83	0.44
1:A:116:ASN:HB2	1:A:119:ASP:OD1	2.18	0.44
1:A:143:PRO:HA	1:A:174:LEU:HD13	2.00	0.44
1:A:34:PHE:HE1	1:A:73:LEU:HD22	1.82	0.44
1:A:275:HIS:NE2	1:A:279:GLU:OE2	2.51	0.44
2:B:27:C:N3	2:B:43:G:N2	2.52	0.43
1:A:249:PHE:HE2	1:A:342:LEU:HD22	1.84	0.43
2:B:4:C:H2'	2:B:5:A:O4'	2.19	0.42
1:A:314:TYR:HB2	1:A:345:LEU:HD21	2.01	0.42
2:B:27:C:H2'	2:B:28:U:C6	2.54	0.42
2:B:53:G:H1	2:B:61:C:H42	1.68	0.41
1:A:21:ARG:NH1	1:A:109:ASN:OD1	2.53	0.41
2:B:19:G:H4'	2:B:20:U:C5'	2.51	0.40

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	361/396 (91%)	345 (96%)	16 (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	326/349 (93%)	319 (98%)	7 (2%)	56	81

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	230	LYS
1	A	239	ARG
1	A	249	PHE
1	A	294	LYS
1	A	343	GLU
1	A	344	LYS
1	A	349	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	73/74 (98%)	21 (28%)	2 (2%)

All (21) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	10	G
2	B	13	C
2	B	17	U
2	B	18	G
2	B	19	G
2	B	20	U
2	B	21	A
2	B	22	G
2	B	27	C
2	B	42	A
2	B	43	G
2	B	44	G
2	B	45	U
2	B	46	G
2	B	47	U
2	B	48	C
2	B	49	G
2	B	61	C
2	B	71	C
2	B	72	C
2	B	73	A

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	B	19	G
2	B	45	U

## 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

1 ligand is modelled in this entry.

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	ATP	A	1500	-	26,33,33	1.01	1 (3%)	27,52,52	1.63	5 (18%)

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	ATP	A	1500	-	-	6/18/38/38	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1500	ATP	C5-C4	3.27	1.47	1.40

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1500	ATP	PB-O3B-PG	-3.86	120.30	132.57
3	A	1500	ATP	PA-O3A-PB	-3.84	120.36	132.57
3	A	1500	ATP	N3-C2-N1	-3.04	123.79	128.68
3	A	1500	ATP	C4-C5-N7	-2.78	106.50	109.40
3	A	1500	ATP	O3G-PG-O2G	2.04	115.51	107.57

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1500	ATP	C5'-O5'-PA-O2A
3	A	1500	ATP	C5'-O5'-PA-O3A

*Continued on next page...*

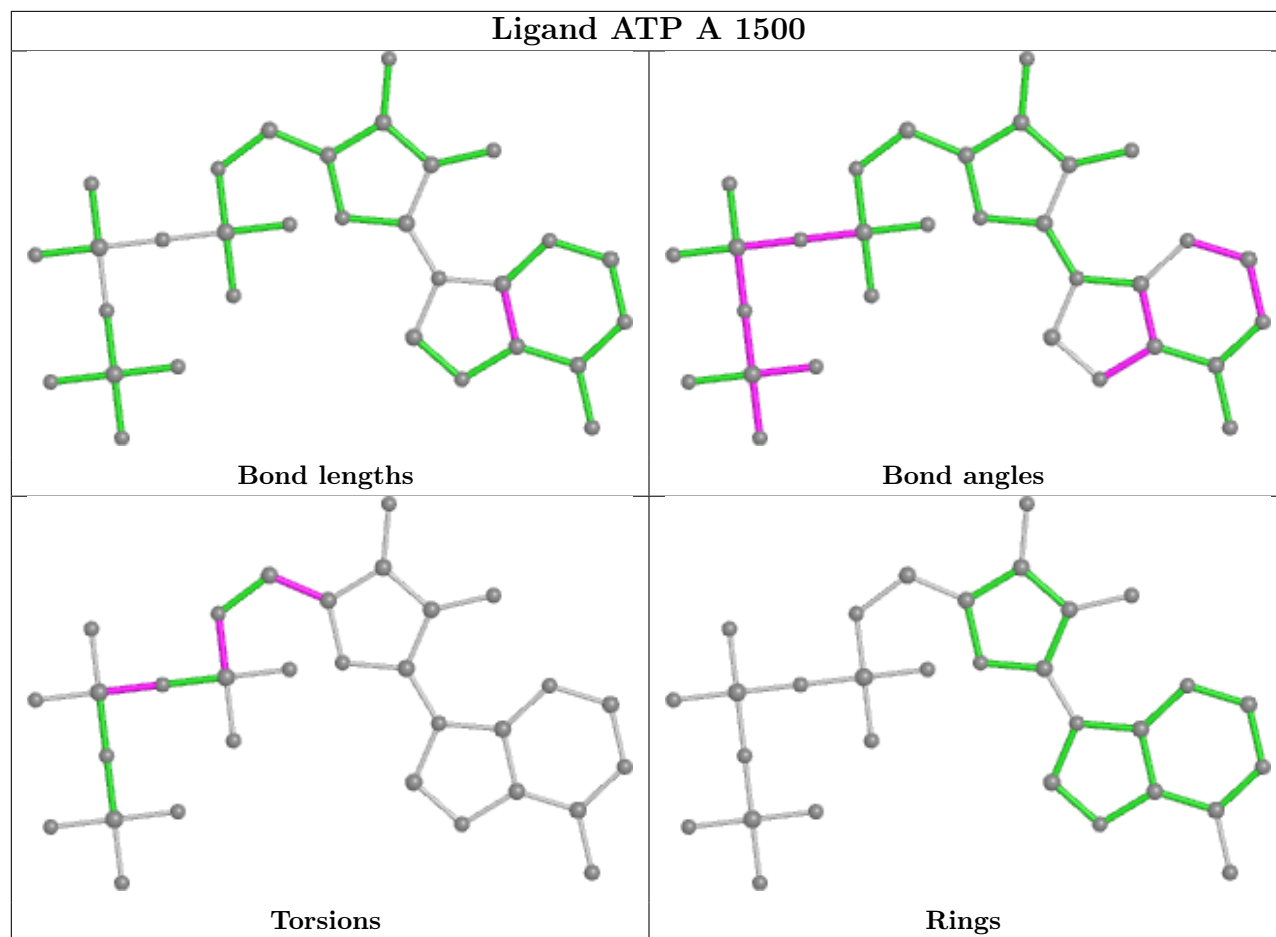
*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
3	A	1500	ATP	PA-O3A-PB-O2B
3	A	1500	ATP	C5'-O5'-PA-O1A
3	A	1500	ATP	O4'-C4'-C5'-O5'
3	A	1500	ATP	PA-O3A-PB-O1B

There are no ring outliers.

No monomer is involved in short contacts.

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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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	367/396 (92%)	-0.13	7 (1%) 66 60	76, 138, 326, 481	0
2	B	74/74 (100%)	3.48	56 (75%) 0 0	238, 324, 430, 473	0
All	All	441/470 (93%)	0.48	63 (14%) 2 3	76, 162, 370, 481	0

All (63) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	20	U	12.7
2	B	32	C	10.0
2	B	29	G	9.5
2	B	36	A	8.9
2	B	37	A	8.4
2	B	53	G	7.9
2	B	33	U	7.4
2	B	72	C	7.3
2	B	9	A	7.2
2	B	44	G	6.1
2	B	13	C	6.0
1	A	365	GLY	5.9
2	B	31	A	5.9
2	B	8	U	5.6
1	A	367	GLU	5.5
2	B	52	G	5.4
2	B	17	U	5.3
2	B	30	G	5.2
2	B	60	U	5.2
2	B	46	G	4.9
2	B	73	A	4.5
2	B	19	G	4.4
2	B	41	C	4.2
2	B	48	C	4.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	45	U	4.1
2	B	24	G	4.0
1	A	368	LEU	3.9
2	B	34	G	3.9
2	B	54	U	3.6
2	B	25	C	3.6
2	B	35	A	3.6
2	B	21	A	3.4
2	B	71	C	3.3
2	B	43	G	3.2
2	B	62	C	3.1
2	B	28	U	3.0
2	B	59	U	2.9
2	B	42	A	2.9
2	B	64	C	2.9
1	A	352	LYS	2.8
2	B	38	A	2.8
2	B	70	G	2.8
2	B	26	A	2.8
2	B	12	U	2.7
2	B	47	U	2.7
2	B	63	G	2.7
2	B	27	C	2.7
2	B	40	C	2.6
2	B	39	U	2.5
1	A	381	ASN	2.5
2	B	7	G	2.4
2	B	15	G	2.4
2	B	58	A	2.4
2	B	18	G	2.3
2	B	2	G	2.2
2	B	4	C	2.2
2	B	16	U	2.2
2	B	3	C	2.2
2	B	14	A	2.1
1	A	350	LEU	2.1
1	A	371	ARG	2.1
2	B	23	A	2.0
2	B	74	C	2.0

## 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](#)

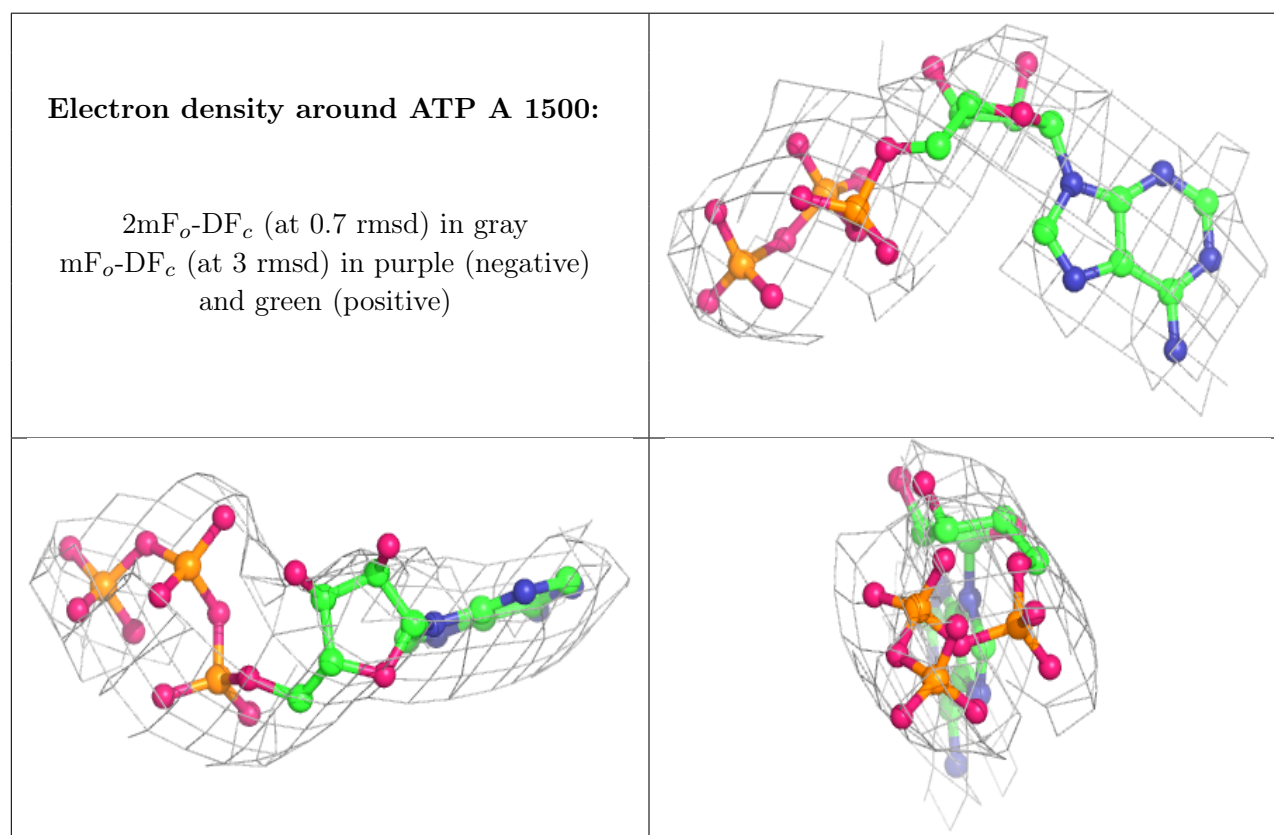
There are no carbohydrates in this entry.

## 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	ATP	A	1500	31/31	0.94	0.16	79,150,212,225	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.



## 6.5 Other polymers [i](#)

There are no such residues in this entry.