wwPDB X-ray Structure Validation Summary Report (i)

Oct 15, 2023 – 06:36 PM EDT

PDB ID	:	8DH5
Title	:	T7 RNA polymerase elongation complex with unnatural base dPa-ATP mismatch
Authors	:	Oh, J.; Wang, D.
Deposited on	:	2022-06-24
Resolution	:	2.85 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:

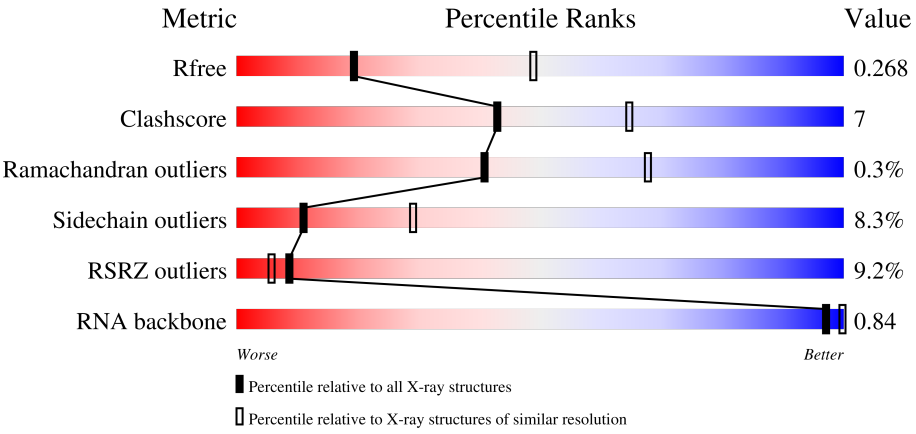
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.85 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.



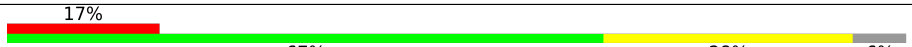
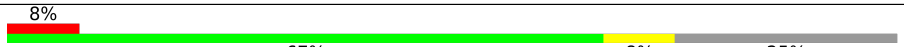
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3168 (2.90-2.82)
Clashscore	141614	3438 (2.90-2.82)
Ramachandran outliers	138981	3348 (2.90-2.82)
Sidechain outliers	138945	3351 (2.90-2.82)
RSRZ outliers	127900	3103 (2.90-2.82)
RNA backbone	3102	1088 (3.12-2.60)

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	B	883	3% 79% 18%
1	E	883	8% 71% 19% 9%
1	I	883	9% 71% 22% 5%
1	M	883	13% 61% 20% 16%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	A	18	
2	F	18	
2	J	18	
2	N	18	
3	C	12	
3	G	12	
3	K	12	
3	O	12	
4	D	9	
4	H	9	
4	L	9	
4	P	9	

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 27994 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 T7 RNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	858	Total	C	N	O	S	0	0	0
			6747	4298	1173	1240	36			
1	E	803	Total	C	N	O	S	0	0	0
			6217	3965	1072	1146	34			
1	I	837	Total	C	N	O	S	0	0	0
			6541	4172	1129	1205	35			
1	M	740	Total	C	N	O	S	0	0	0
			5711	3640	994	1045	32			

- Molecule 2 is a DNA chain called Template strand DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	16	Total	C	N	O	P	0	0	0
			325	155	60	94	16			
2	F	17	Total	C	N	O	P	0	0	0
			344	165	65	98	16			
2	J	16	Total	C	N	O	P	0	0	0
			325	155	60	94	16			
2	N	17	Total	C	N	O	P	0	0	0
			344	165	65	98	16			

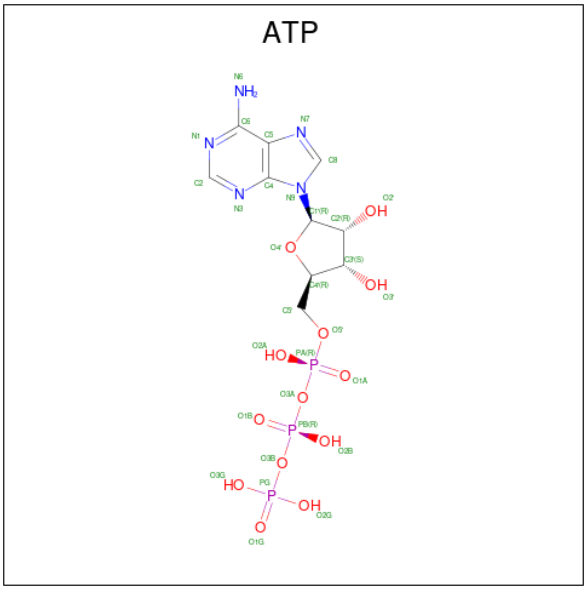
- Molecule 3 is a RNA chain called RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	9	Total	C	N	O	P	0	0	0
			193	86	35	63	9			
3	G	8	Total	C	N	O	P	0	0	0
			173	77	33	55	8			
3	K	8	Total	C	N	O	P	0	0	0
			173	77	33	55	8			
3	O	8	Total	C	N	O	P	0	0	0
			173	77	33	55	8			

- Molecule 4 is a DNA chain called Non-template strand DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	7	Total	C	N	O	P	0	0	0
			141	68	22	44	7			
4	H	8	Total	C	N	O	P	0	0	0
			160	77	25	50	8			
4	L	7	Total	C	N	O	P	0	0	0
			141	68	22	44	7			
4	P	8	Total	C	N	O	P	0	0	0
			160	77	25	50	8			

- Molecule 5 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	B	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
5	E	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
5	I	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
5	M	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

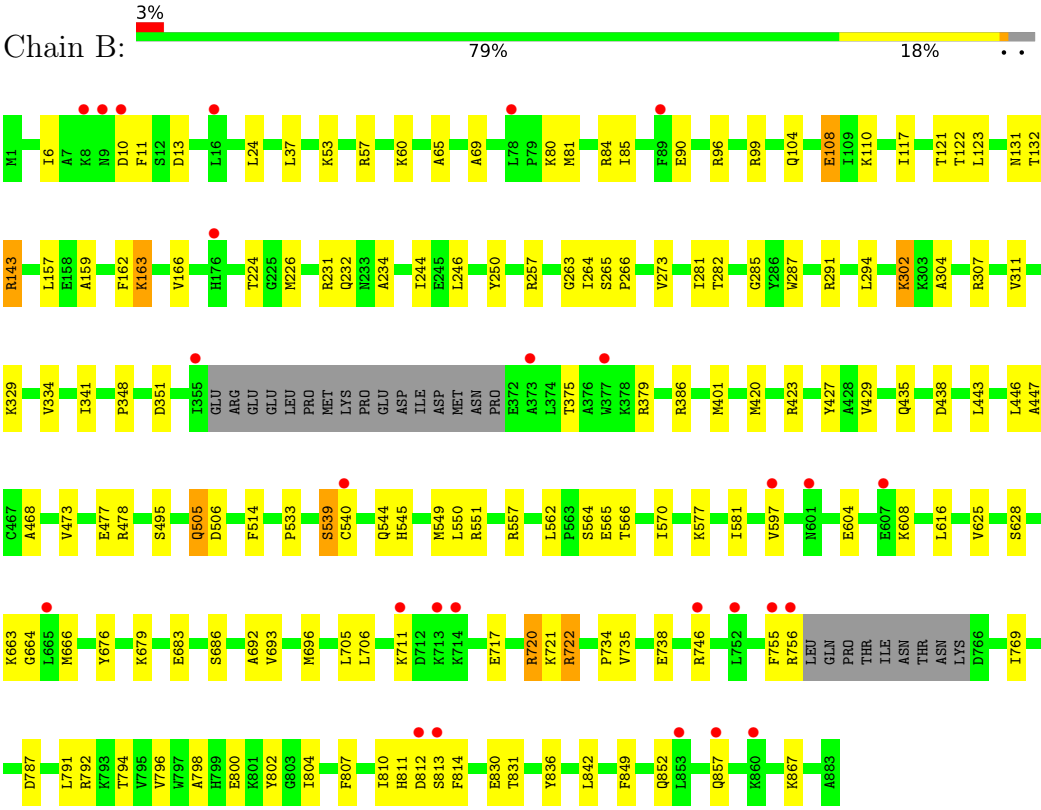
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	1	Total Mg 1 1	0	0
6	E	1	Total Mg 1 1	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

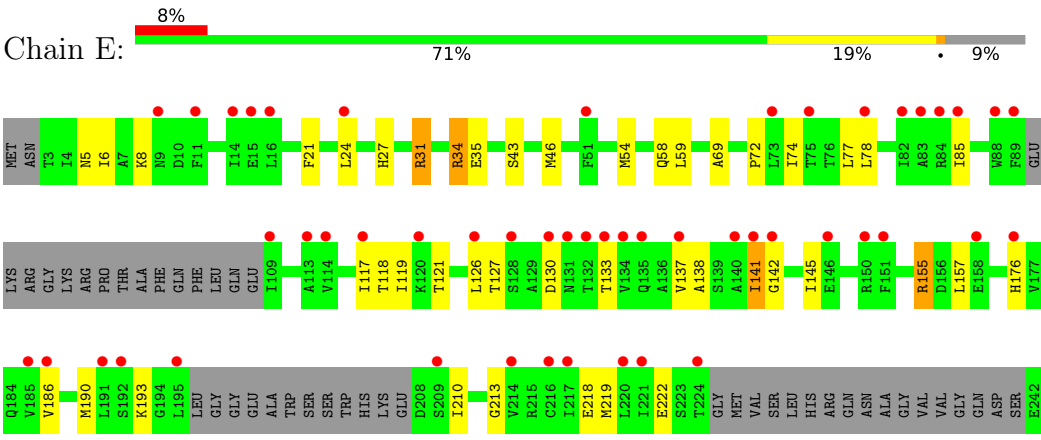
• Molecule 1: T7 RNA polymerase

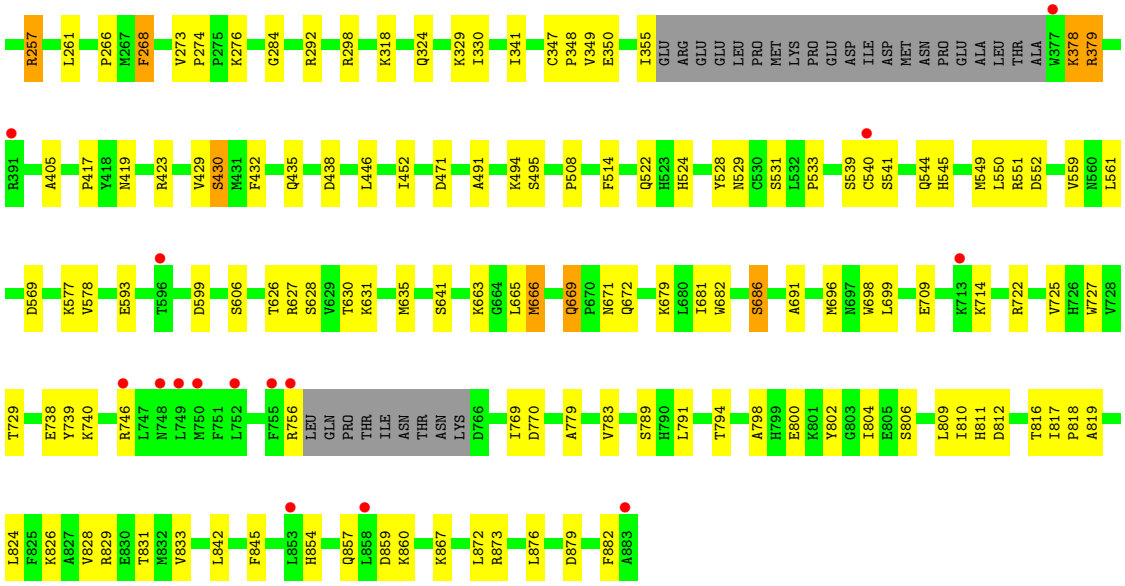
Chain B:



• Molecule 1: T7 RNA polymerase

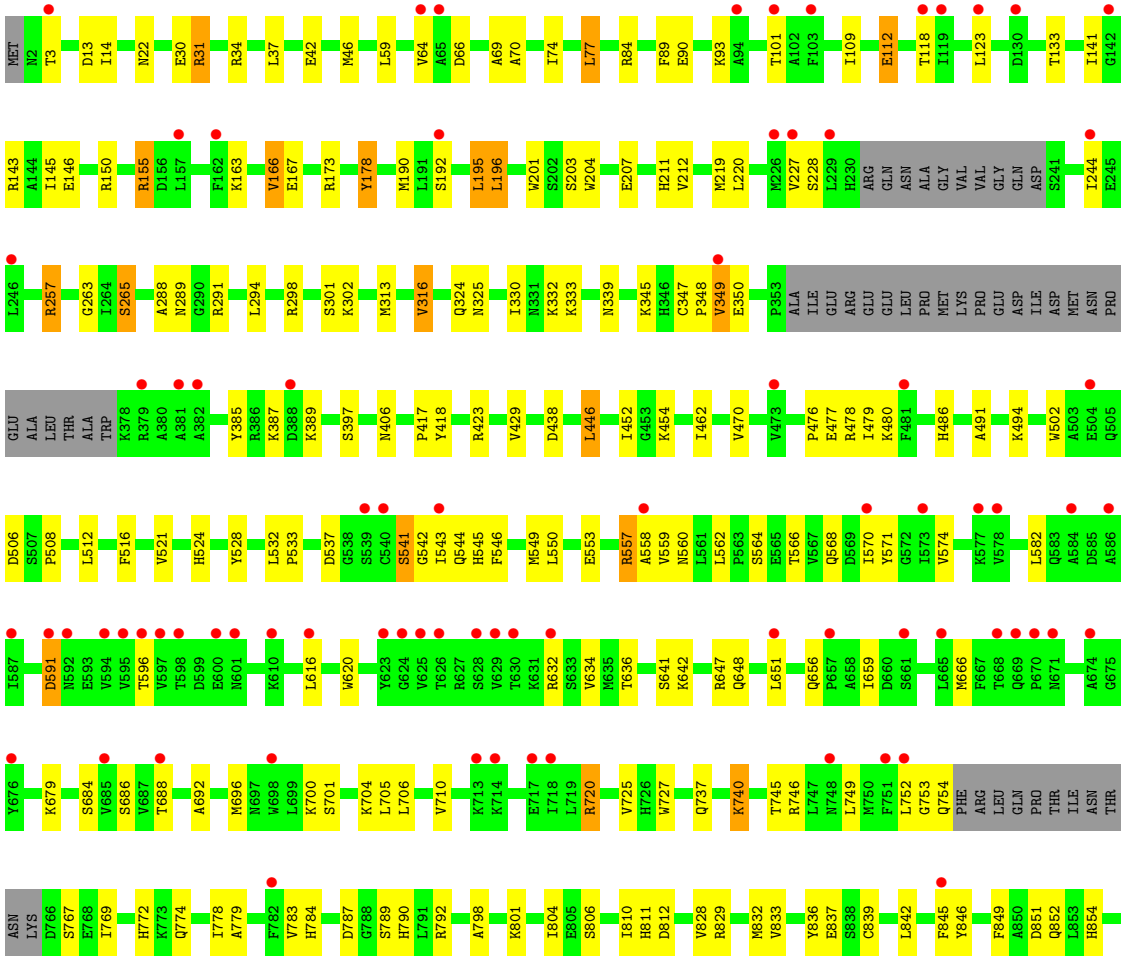
Chain E:

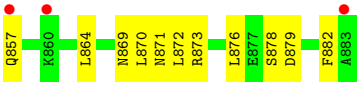




• Molecule 1: T7 RNA polymerase

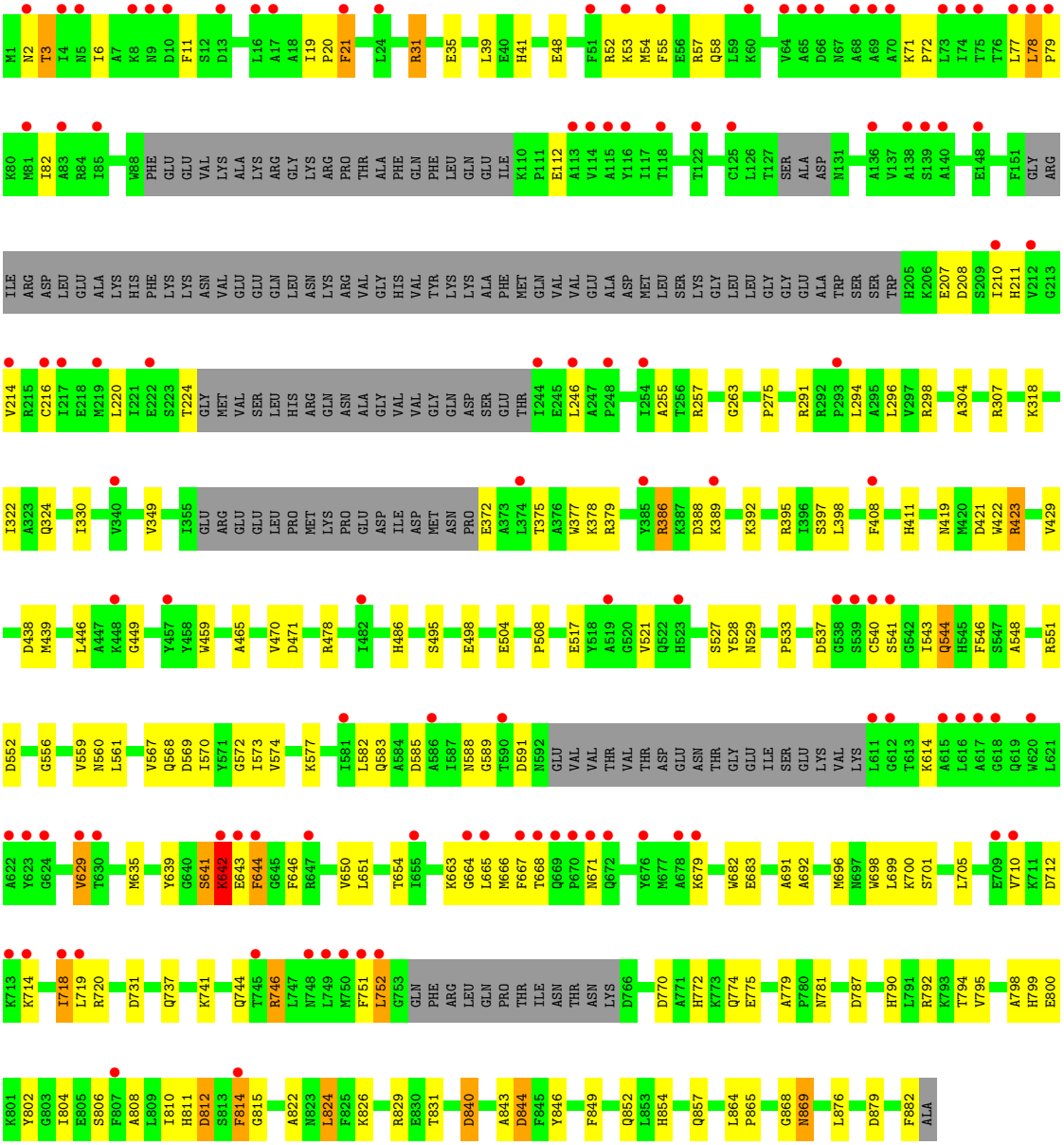
Chain I:





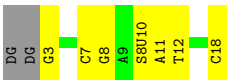
• Molecule 1: T7 RNA polymerase

Chain M:



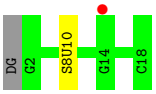
• Molecule 2: Template strand DNA

Chain A:



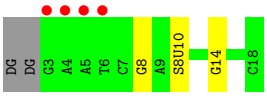
• Molecule 2: Template strand DNA

Chain F:



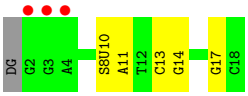
• Molecule 2: Template strand DNA

Chain J:



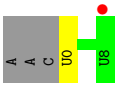
• Molecule 2: Template strand DNA

Chain N:



• Molecule 3: RNA

Chain C:

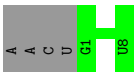


• Molecule 3: RNA

Chain G:

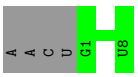
• Molecule 3: RNA

Chain K:



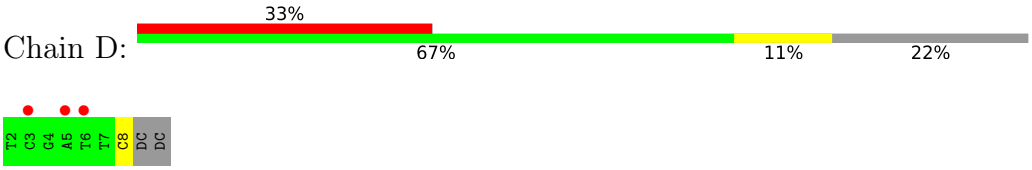
• Molecule 3: RNA

Chain O:



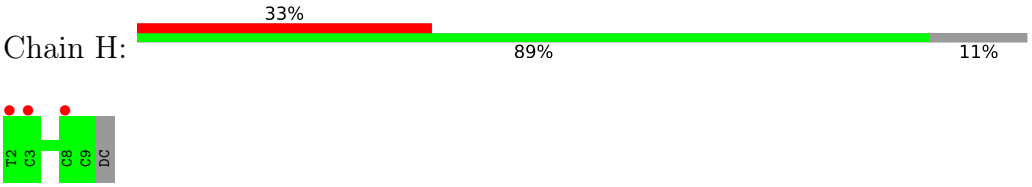
• Molecule 4: Non-template strand DNA

Chain D:



• Molecule 4: Non-template strand DNA

Chain H:



• Molecule 4: Non-template strand DNA

Chain L:

• Molecule 4: Non-template strand DNA

Chain P:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	78.54Å 86.31Å 201.20Å 89.79° 85.39° 69.49°	Depositor
Resolution (Å)	47.43 – 2.85 47.43 – 2.85	Depositor EDS
% Data completeness (in resolution range)	98.4 (47.43-2.85) 98.4 (47.43-2.85)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.09 (at 2.86Å)	Xtriage
Refinement program	PHENIX 1.19	Depositor
R, R_{free}	0.234 , 0.267 0.234 , 0.268	Depositor DCC
R_{free} test set	1951 reflections (1.72%)	wwPDB-VP
Wilson B-factor (Å^2)	76.7	Xtriage
Anisotropy	0.302	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 67.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	27994	wwPDB-VP
Average B, all atoms (Å^2)	108.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.99% 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.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ATP, S8U, MG

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	B	0.28	0/6898	0.51	1/9330 (0.0%)
1	E	0.28	0/6352	0.53	0/8604
1	I	0.31	1/6688 (0.0%)	0.58	3/9052 (0.0%)
1	M	0.33	1/5838 (0.0%)	0.57	1/7906 (0.0%)
2	A	0.52	0/343	0.84	0/524
2	F	0.52	0/365	0.83	0/559
2	J	0.50	0/343	0.83	0/524
2	N	0.56	0/365	0.93	0/559
3	C	0.23	0/215	0.86	0/333
3	G	0.18	0/193	0.73	0/299
3	K	0.21	0/193	0.74	0/299
3	O	0.18	0/193	0.76	0/299
4	D	0.44	0/156	0.99	0/238
4	H	0.48	0/177	1.00	0/270
4	L	0.51	0/156	1.04	0/238
4	P	0.47	0/177	0.95	0/270
All	All	0.32	2/28652 (0.0%)	0.59	5/39304 (0.0%)

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	E	0	1
1	I	0	1
1	M	0	1
All	All	0	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	I	31	ARG	CB-CG	-6.45	1.35	1.52
1	M	644	PHE	CB-CG	-5.06	1.42	1.51

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	I	195	LEU	CA-CB-CG	12.33	143.66	115.30
1	I	31	ARG	NE-CZ-NH1	9.99	125.30	120.30
1	I	31	ARG	CG-CD-NE	-7.87	95.27	111.80
1	B	787	ASP	CB-CG-OD1	6.71	124.34	118.30
1	M	21	PHE	CB-CG-CD2	5.04	124.33	120.80

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	155	ARG	Sidechain
1	I	31	ARG	Sidechain
1	M	642	LYS	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	B	6747	0	6712	74	0
1	E	6217	0	6100	79	0
1	I	6541	0	6466	104	0
1	M	5711	0	5565	105	0
2	A	325	0	168	4	0
2	F	344	0	180	0	0
2	J	325	0	168	2	0
2	N	344	0	180	5	0
3	C	193	0	96	1	0
3	G	173	0	86	0	0
3	K	173	0	86	0	0
3	O	173	0	86	0	0
4	D	141	0	81	1	0
4	H	160	0	92	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	L	141	0	81	1	0
4	P	160	0	92	2	0
5	B	31	0	12	0	0
5	E	31	0	12	0	0
5	I	31	0	12	0	0
5	M	31	0	12	0	0
6	B	1	0	0	0	0
6	E	1	0	0	0	0
All	All	27994	0	26287	369	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 7.

The worst 5 of 369 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)
1:B:304:ALA:HA	1:B:307:ARG:HE	1.51	0.75
1:I:347:CYS:HB2	1:I:350:GLU:HG2	1.72	0.72
1:E:72:PRO:HG3	1:E:257:ARG:HD3	1.72	0.71
1:I:324:GLN:HE21	1:I:418:TYR:H	1.39	0.68
1:I:294:LEU:HD11	1:I:429:VAL:HG21	1.77	0.67

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	B	852/883 (96%)	824 (97%)	25 (3%)	3 (0%)	34 62
1	E	791/883 (90%)	762 (96%)	24 (3%)	5 (1%)	25 53
1	I	829/883 (94%)	798 (96%)	30 (4%)	1 (0%)	51 79

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	M	724/883 (82%)	703 (97%)	19 (3%)	2 (0%)	41 68
All	All	3196/3532 (90%)	3087 (97%)	98 (3%)	11 (0%)	41 68

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	I	811	HIS
1	B	811	HIS
1	E	141	ILE
1	E	142	GLY
1	E	539	SER

5.3.2 Protein sidechains (i)

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	B	702/729 (96%)	661 (94%)	41 (6%)	20 46
1	E	634/729 (87%)	579 (91%)	55 (9%)	10 27
1	I	677/729 (93%)	612 (90%)	65 (10%)	8 22
1	M	573/729 (79%)	519 (91%)	54 (9%)	8 23
All	All	2586/2916 (89%)	2371 (92%)	215 (8%)	11 29

5 of 215 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	265	SER
1	I	648	GLN
1	M	668	THR
1	I	325	ASN
1	I	506	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	I	406	ASN
1	M	560	ASN
1	M	781	ASN
1	M	772	HIS
1	B	726	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	C	7/12 (58%)	0	0
3	G	6/12 (50%)	0	0
3	K	6/12 (50%)	0	0
3	O	6/12 (50%)	0	0
All	All	25/48 (52%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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

4 non-standard protein/DNA/RNA residues are 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	$\#\|Z\| > 2$
2	S8U	F	10	2	15,19,20	4.34	7 (46%)	16,26,29	1.32	2 (12%)
2	S8U	N	10	2	15,19,20	4.37	7 (46%)	16,26,29	1.40	3 (18%)
2	S8U	A	10	2	15,19,20	4.36	7 (46%)	16,26,29	1.33	2 (12%)
2	S8U	J	10	2	15,19,20	4.37	7 (46%)	16,26,29	1.21	1 (6%)

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	S8U	F	10	2	-	0/3/23/24	0/2/2/2
2	S8U	N	10	2	-	0/3/23/24	0/2/2/2
2	S8U	A	10	2	-	2/3/23/24	0/2/2/2
2	S8U	J	10	2	-	2/3/23/24	0/2/2/2

The worst 5 of 28 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	10	S8U	C2'-C3'	-11.43	1.22	1.52
2	J	10	S8U	C2'-C3'	-11.33	1.22	1.52
2	N	10	S8U	C2'-C3'	-11.33	1.22	1.52
2	F	10	S8U	C2'-C3'	-11.21	1.23	1.52
2	J	10	S8U	O4'-C4'	-8.02	1.27	1.45

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	J	10	S8U	O14-C13-C12	-3.14	119.72	124.17
2	A	10	S8U	O14-C13-C12	-3.05	119.84	124.17
2	N	10	S8U	O14-C13-C12	-2.96	119.97	124.17
2	F	10	S8U	O14-C13-C12	-2.92	120.04	124.17
2	N	10	S8U	C2'-C3'-C4'	2.68	108.35	102.76

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	10	S8U	O4'-C4'-C5'-O5'
2	A	10	S8U	C3'-C4'-C5'-O5'
2	J	10	S8U	O4'-C4'-C5'-O5'
2	J	10	S8U	C3'-C4'-C5'-O5'

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 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	$\#\|Z\| > 2$
5	ATP	E	901	6	26,33,33	0.60	0	31,52,52	0.73	2 (6%)
5	ATP	I	901	-	26,33,33	0.58	0	31,52,52	0.80	2 (6%)
5	ATP	B	901	6	26,33,33	0.60	0	31,52,52	0.75	2 (6%)
5	ATP	M	901	-	26,33,33	0.59	0	31,52,52	0.78	2 (6%)

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
5	ATP	E	901	6	-	4/18/38/38	0/3/3/3
5	ATP	I	901	-	-	7/18/38/38	0/3/3/3
5	ATP	B	901	6	-	5/18/38/38	0/3/3/3
5	ATP	M	901	-	-	4/18/38/38	0/3/3/3

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	I	901	ATP	C5-C6-N6	2.31	123.86	120.35
5	M	901	ATP	C5-C6-N6	2.30	123.85	120.35
5	B	901	ATP	C5-C6-N6	2.29	123.84	120.35
5	E	901	ATP	C5-C6-N6	2.27	123.80	120.35
5	M	901	ATP	PB-O3B-PG	2.11	140.06	132.83

There are no chirality outliers.

5 of 20 torsion outliers are listed below:

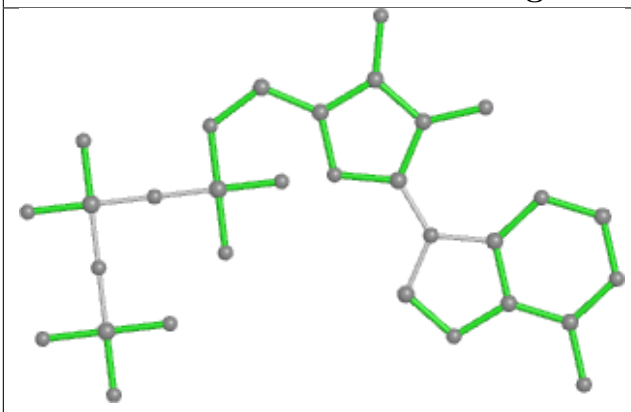
Mol	Chain	Res	Type	Atoms
5	E	901	ATP	C5'-O5'-PA-O2A
5	E	901	ATP	C5'-O5'-PA-O3A
5	E	901	ATP	O4'-C4'-C5'-O5'
5	E	901	ATP	C3'-C4'-C5'-O5'
5	I	901	ATP	C5'-O5'-PA-O1A

There are no ring outliers.

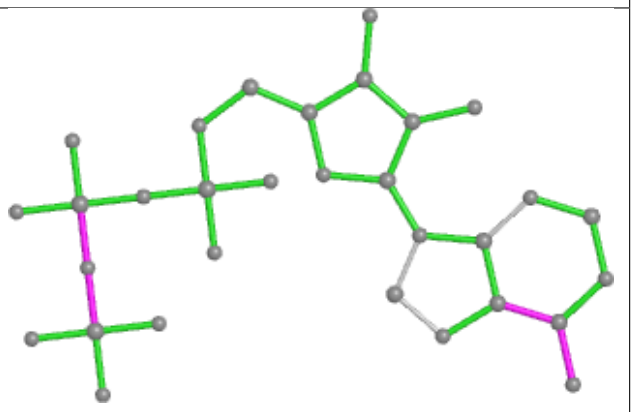
No monomer is involved in short contacts.

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 then 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.

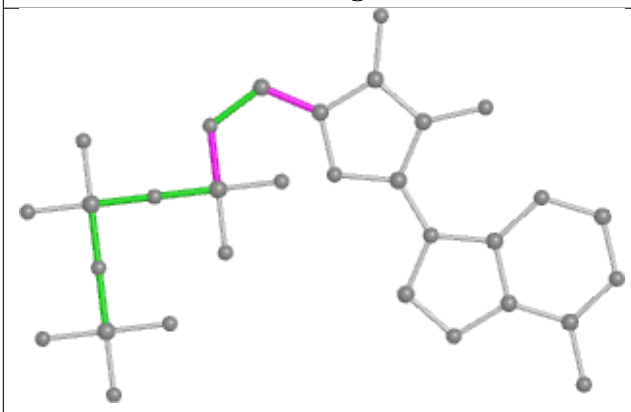
Ligand ATP E 901



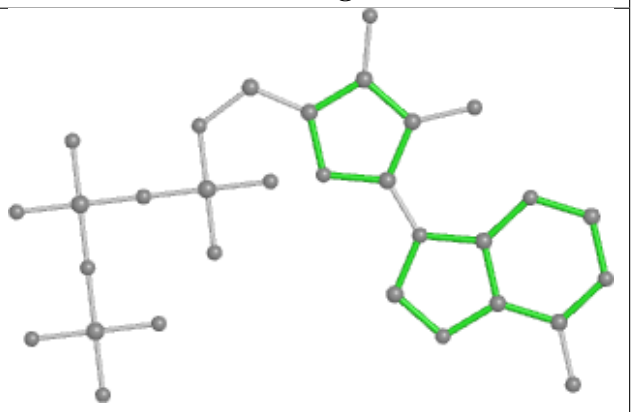
Bond lengths



Bond angles

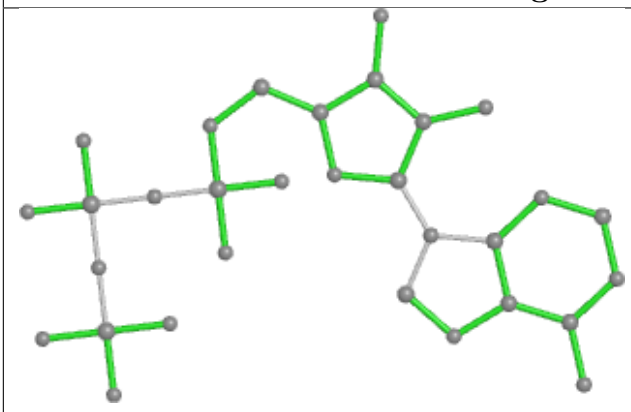


Torsions

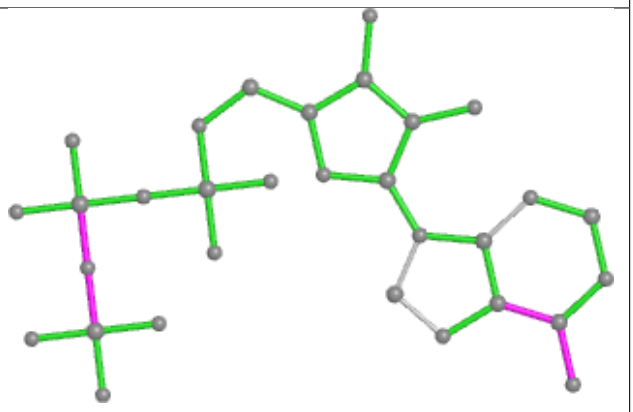


Rings

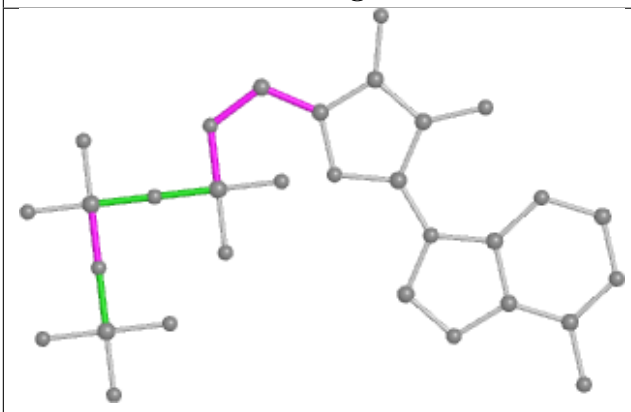
Ligand ATP I 901



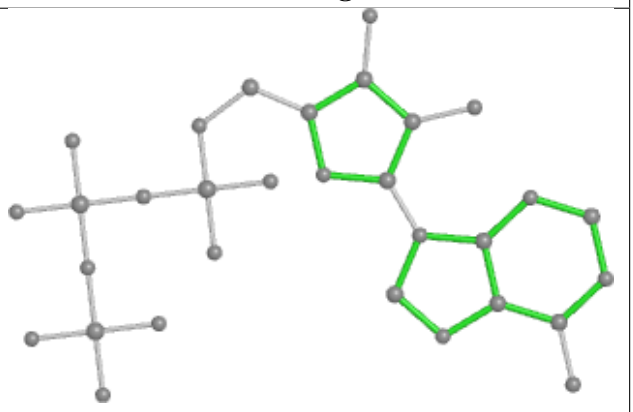
Bond lengths



Bond angles

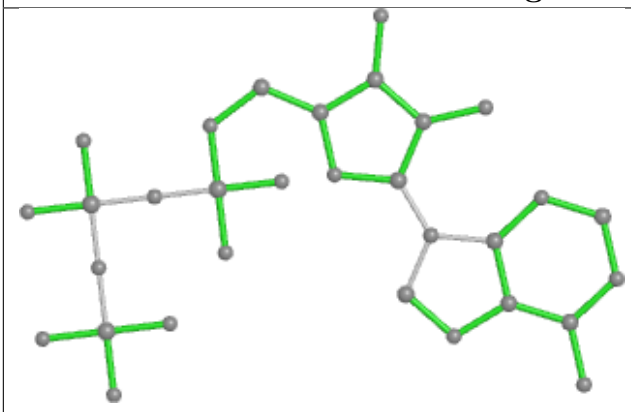


Torsions

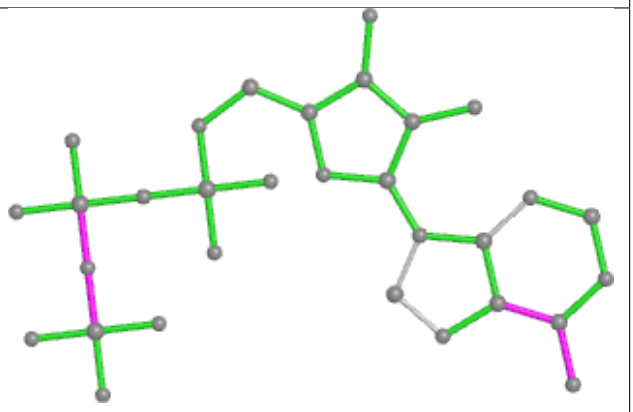


Rings

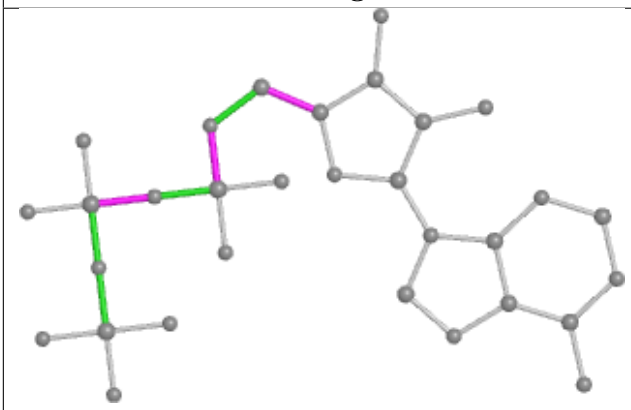
Ligand ATP B 901



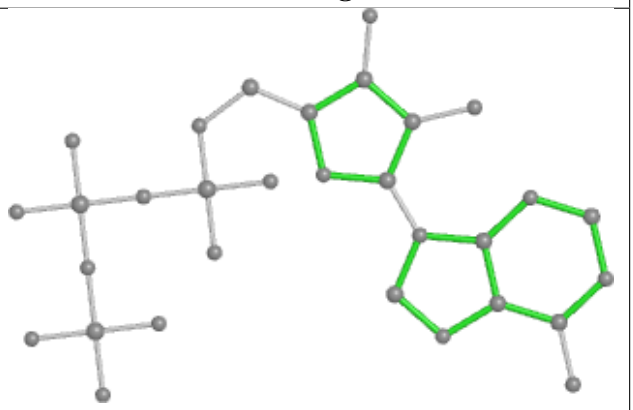
Bond lengths



Bond angles

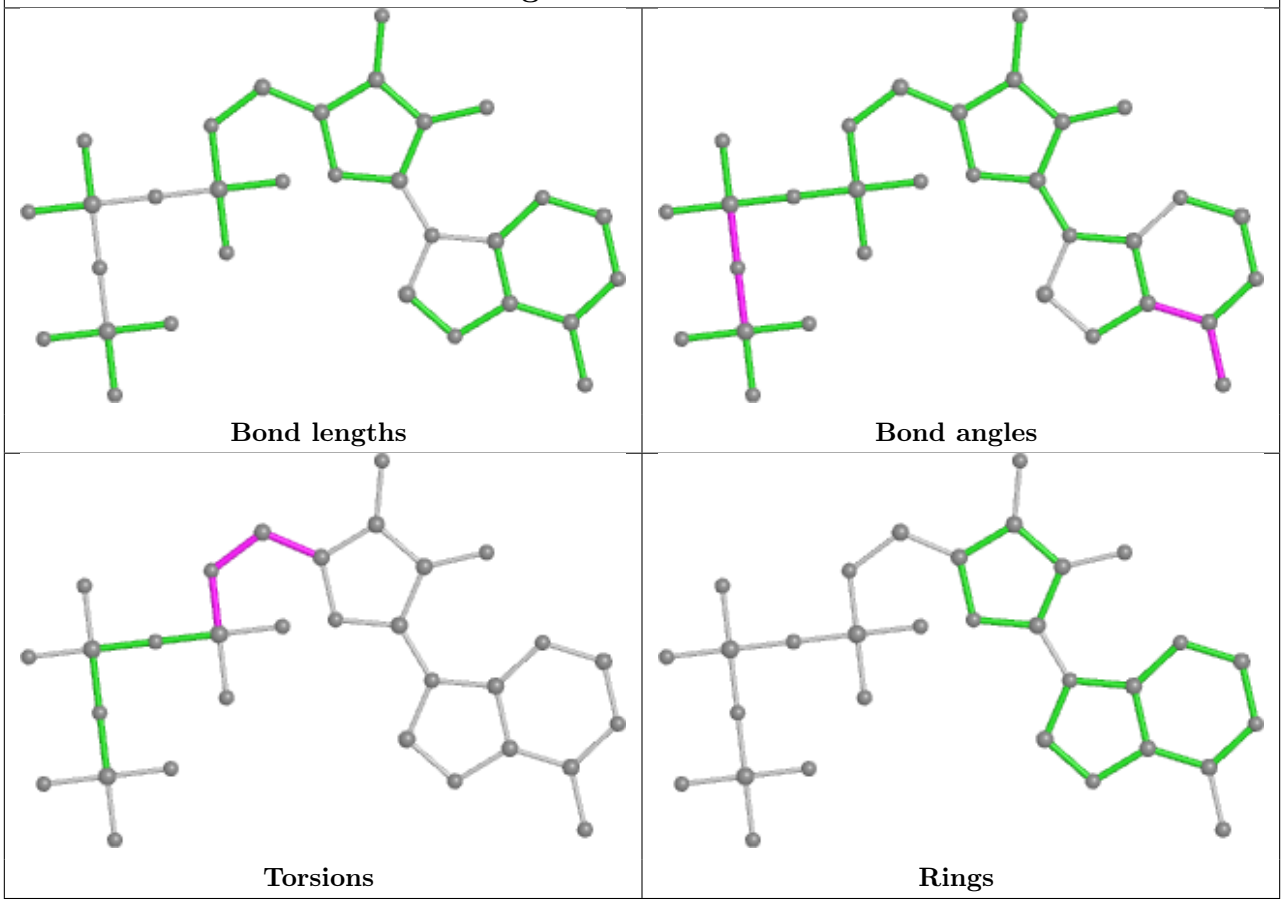


Torsions



Rings

Ligand ATP M 901



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^{th} 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	B	858/883 (97%)	0.24	27 (3%) 49 44	38, 71, 139, 239	0
1	E	803/883 (90%)	0.49	69 (8%) 10 7	44, 81, 176, 231	0
1	I	837/883 (94%)	0.52	82 (9%) 7 5	54, 110, 175, 249	0
1	M	740/883 (83%)	0.79	113 (15%) 2 1	73, 119, 190, 247	0
2	A	15/18 (83%)	0.64	0 100 100	56, 68, 237, 260	0
2	F	16/18 (88%)	0.41	1 (6%) 20 15	66, 118, 308, 326	0
2	J	15/18 (83%)	1.25	4 (26%) 0 0	79, 108, 334, 384	0
2	N	16/18 (88%)	1.10	3 (18%) 1 1	102, 171, 312, 351	0
3	C	9/12 (75%)	1.21	1 (11%) 5 3	58, 63, 107, 180	0
3	G	8/12 (66%)	0.80	0 100 100	68, 97, 144, 171	0
3	K	8/12 (66%)	1.01	0 100 100	89, 96, 122, 140	0
3	O	8/12 (66%)	0.58	0 100 100	102, 121, 174, 189	0
4	D	7/9 (77%)	1.70	3 (42%) 0 0	208, 233, 257, 280	0
4	H	8/9 (88%)	1.23	3 (37%) 0 0	192, 213, 236, 262	0
4	L	7/9 (77%)	2.02	2 (28%) 0 0	260, 294, 323, 332	0
4	P	8/9 (88%)	1.17	2 (25%) 0 0	210, 246, 265, 265	0
All	All	3363/3688 (91%)	0.52	310 (9%) 9 6	38, 98, 184, 384	0

The worst 5 of 310 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	M	65	ALA	6.7
1	M	620	TRP	6.5
1	E	140	ALA	6.5
1	M	670	PRO	6.5
2	N	4	DA	6.0

6.2 Non-standard residues in protein, DNA, RNA chains (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^{th} 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(Å^2)	Q<0.9
2	S8U	N	10	18/19	0.92	0.21	105,130,157,163	0
2	S8U	J	10	18/19	0.93	0.26	89,118,142,143	0
2	S8U	A	10	18/19	0.95	0.21	51,69,79,83	0
2	S8U	F	10	18/19	0.95	0.19	67,74,84,88	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

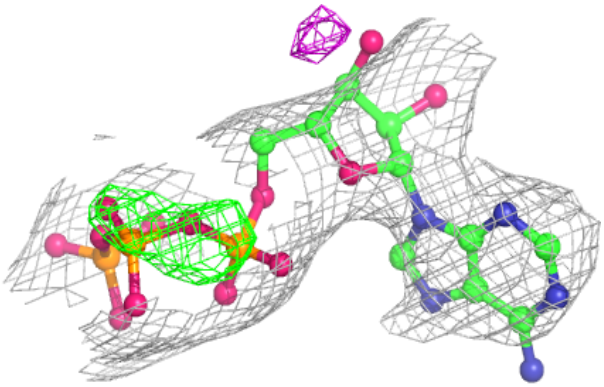
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

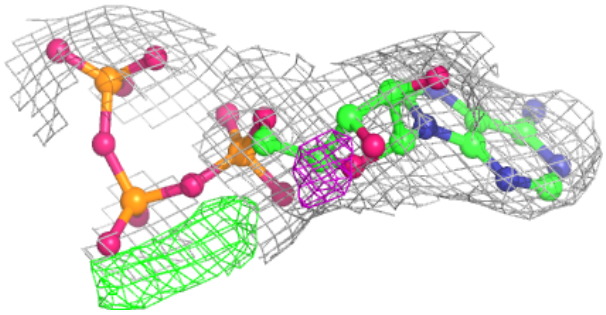
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	MG	B	902	1/1	0.87	0.85	107,107,107,107	0
5	ATP	M	901	31/31	0.89	0.31	116,135,149,151	0
5	ATP	I	901	31/31	0.90	0.24	109,131,159,170	0
5	ATP	B	901	31/31	0.95	0.24	51,70,95,106	0
5	ATP	E	901	31/31	0.96	0.20	43,91,107,114	0
6	MG	E	902	1/1	0.98	0.32	58,58,58,58	0

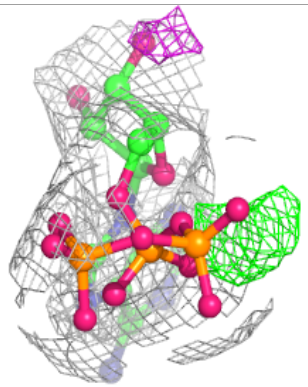
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around ATP M 901:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

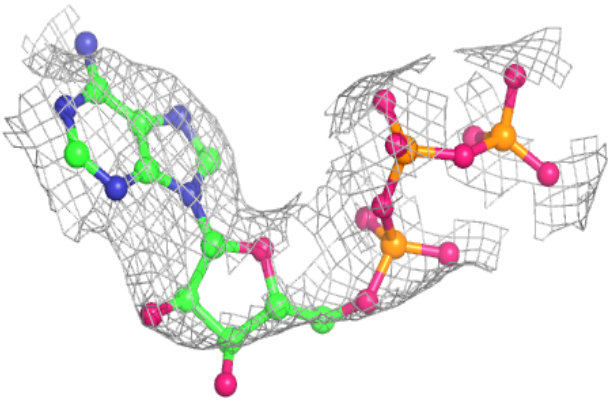


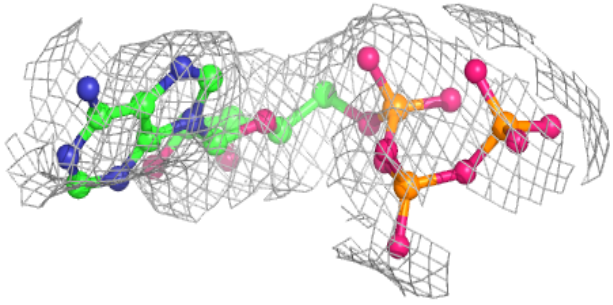


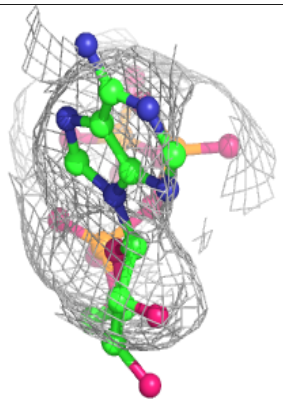


Electron density around ATP I 901:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

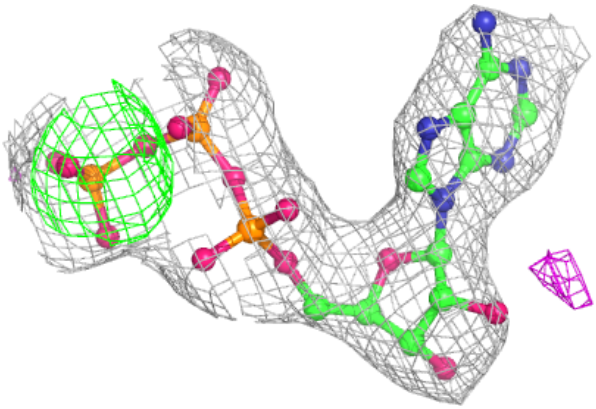


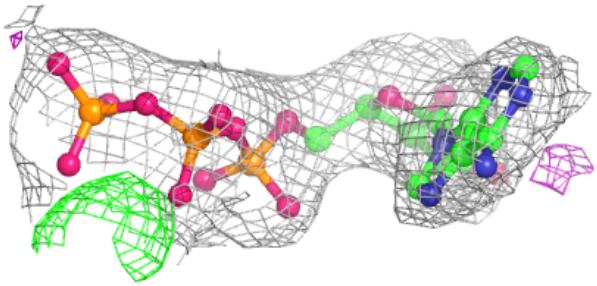


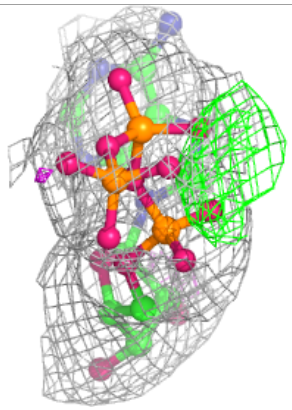


Electron density around ATP B 901:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

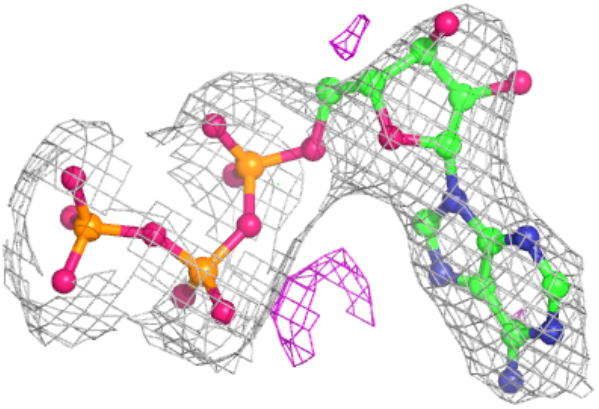


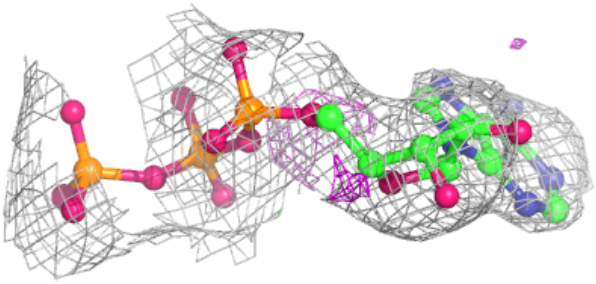


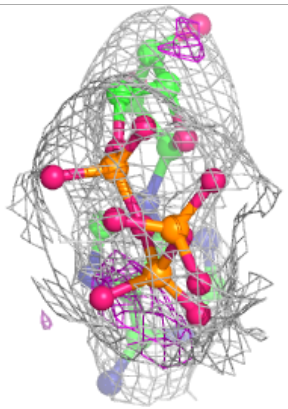


Electron density around ATP E 901:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)







6.5 Other polymers (i)

There are no such residues in this entry.