wwPDB X-ray Structure Validation Summary Report (i)

Nov 15, 2023 – 10:17 AM JST

PDB ID	:	6J18
Title	:	ATPase
Authors	:	Wang, S.H.; Li, J.; Rao, Z.H.
Deposited on	:	2018-12-28
Resolution	:	2.00 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:

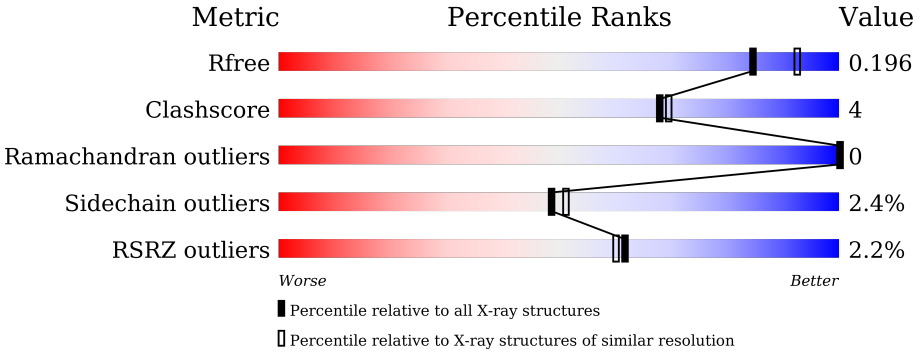
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber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

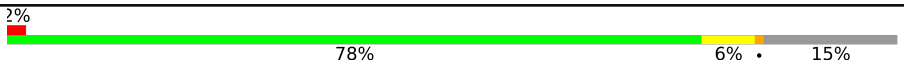
The reported resolution of this entry is 2.00 Å.

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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	A	321	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2333 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 ESX-5 secretion system protein EccC5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	272	2087	1317	385	376	9	0	3	0

There are 45 discrepancies between the modelled and reference sequences:

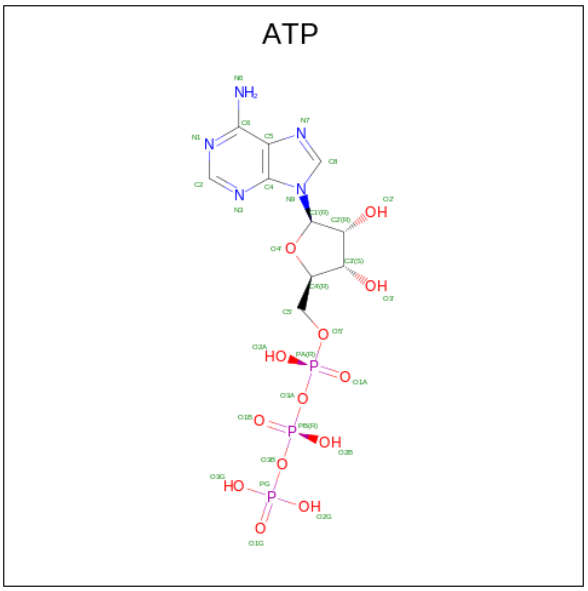
Chain	Residue	Modelled	Actual	Comment	Reference
A	1079	MET	-	initiating methionine	UNP P9WNA5
A	1080	LYS	-	expression tag	UNP P9WNA5
A	1081	TYR	-	expression tag	UNP P9WNA5
A	1082	LEU	-	expression tag	UNP P9WNA5
A	1083	LEU	-	expression tag	UNP P9WNA5
A	1084	PRO	-	expression tag	UNP P9WNA5
A	1085	THR	-	expression tag	UNP P9WNA5
A	1086	ALA	-	expression tag	UNP P9WNA5
A	1087	ALA	-	expression tag	UNP P9WNA5
A	1088	ALA	-	expression tag	UNP P9WNA5
A	1089	GLY	-	expression tag	UNP P9WNA5
A	1090	LEU	-	expression tag	UNP P9WNA5
A	1091	LEU	-	expression tag	UNP P9WNA5
A	1092	LEU	-	expression tag	UNP P9WNA5
A	1093	LEU	-	expression tag	UNP P9WNA5
A	1094	ALA	-	expression tag	UNP P9WNA5
A	1095	ALA	-	expression tag	UNP P9WNA5
A	1096	GLN	-	expression tag	UNP P9WNA5
A	1097	PRO	-	expression tag	UNP P9WNA5
A	1098	ALA	-	expression tag	UNP P9WNA5
A	1099	MET	-	expression tag	UNP P9WNA5
A	1100	ALA	-	expression tag	UNP P9WNA5
A	1101	MET	-	expression tag	UNP P9WNA5
A	1102	ASP	-	expression tag	UNP P9WNA5
A	1103	SER	-	expression tag	UNP P9WNA5
A	1104	SER	-	expression tag	UNP P9WNA5
A	1105	GLY	-	expression tag	UNP P9WNA5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	1106	LEU	-	expression tag	UNP P9WNA5
A	1107	GLU	-	expression tag	UNP P9WNA5
A	1108	VAL	-	expression tag	UNP P9WNA5
A	1109	LEU	-	expression tag	UNP P9WNA5
A	1110	PHE	-	expression tag	UNP P9WNA5
A	1111	GLN	-	expression tag	UNP P9WNA5
A	1112	GLY	-	expression tag	UNP P9WNA5
A	1113	PRO	-	expression tag	UNP P9WNA5
A	1114	GLY	-	expression tag	UNP P9WNA5
A	1115	SER	-	expression tag	UNP P9WNA5
A	1392	LEU	-	expression tag	UNP P9WNA5
A	1393	GLU	-	expression tag	UNP P9WNA5
A	1394	HIS	-	expression tag	UNP P9WNA5
A	1395	HIS	-	expression tag	UNP P9WNA5
A	1396	HIS	-	expression tag	UNP P9WNA5
A	1397	HIS	-	expression tag	UNP P9WNA5
A	1398	HIS	-	expression tag	UNP P9WNA5
A	1399	HIS	-	expression tag	UNP P9WNA5

- Molecule 2 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
			Total	C	N	O	P		
2	A	1	31	10	5	13	3	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand

of Interest" by depositor).

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

- Molecule 4 is water.

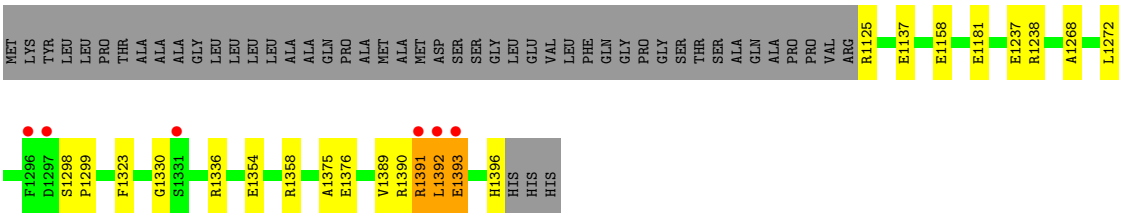
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	214	214	214	0	0

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: ESX-5 secretion system protein EccC5

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α, β, γ	89.16Å 89.16Å 62.64Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	38.61 – 2.00 48.64 – 2.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (38.61-2.00) 100.0 (48.64-2.00)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.29 (at 2.00Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.167 , 0.200 0.168 , 0.196	Depositor DCC
R_{free} test set	934 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å^2)	27.5	Xtriage
Anisotropy	0.111	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 37.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.056 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2333	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.22% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: ATP, 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	A	0.29	0/2147	0.44	0/2920

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	2087	0	2086	16	0
2	A	31	0	12	0	0
3	A	1	0	0	0	0
4	A	214	0	0	1	0
All	All	2333	0	2098	16	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.

The worst 5 of 16 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:A:1389:VAL:HB	1:A:1392:LEU:HD13	1.63	0.80
1:A:1137:GLU:HG2	4:A:1678:HOH:O	1.83	0.78
1:A:1389:VAL:HB	1:A:1392:LEU:CD1	2.28	0.63
1:A:1391:ARG:O	1:A:1393:GLU:HG2	1.99	0.63
1:A:1323:PHE:CE2	1:A:1354:GLU:HG3	2.34	0.62

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	272/321 (85%)	266 (98%)	6 (2%)	0	100 100

There are no Ramachandran outliers to report.

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	A	215/248 (87%)	210 (98%)	5 (2%)	50 53

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1376	GLU
1	A	1390	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1391	ARG
1	A	1392	LEU
1	A	1393	GLU

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

Mol	Chain	Res	Type
1	A	1396	HIS

5.3.3 RNA (i)

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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	1401	3	26,33,33	2.05	10 (38%)	31,52,52	1.26	4 (12%)

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	1401	3	-	0/18/38/38	0/3/3/3

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1401	ATP	PG-O3G	-3.82	1.40	1.54
2	A	1401	ATP	PB-O2B	-3.32	1.39	1.55
2	A	1401	ATP	PG-O1G	-3.18	1.40	1.50
2	A	1401	ATP	PA-O2A	-3.13	1.40	1.55
2	A	1401	ATP	PG-O2G	-2.78	1.44	1.54

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1401	ATP	N3-C2-N1	-2.60	124.61	128.68
2	A	1401	ATP	C3'-C2'-C1'	2.52	104.78	100.98
2	A	1401	ATP	C4-C5-N7	-2.36	106.94	109.40
2	A	1401	ATP	PB-O3B-PG	-2.25	125.11	132.83

There are no chirality outliers.

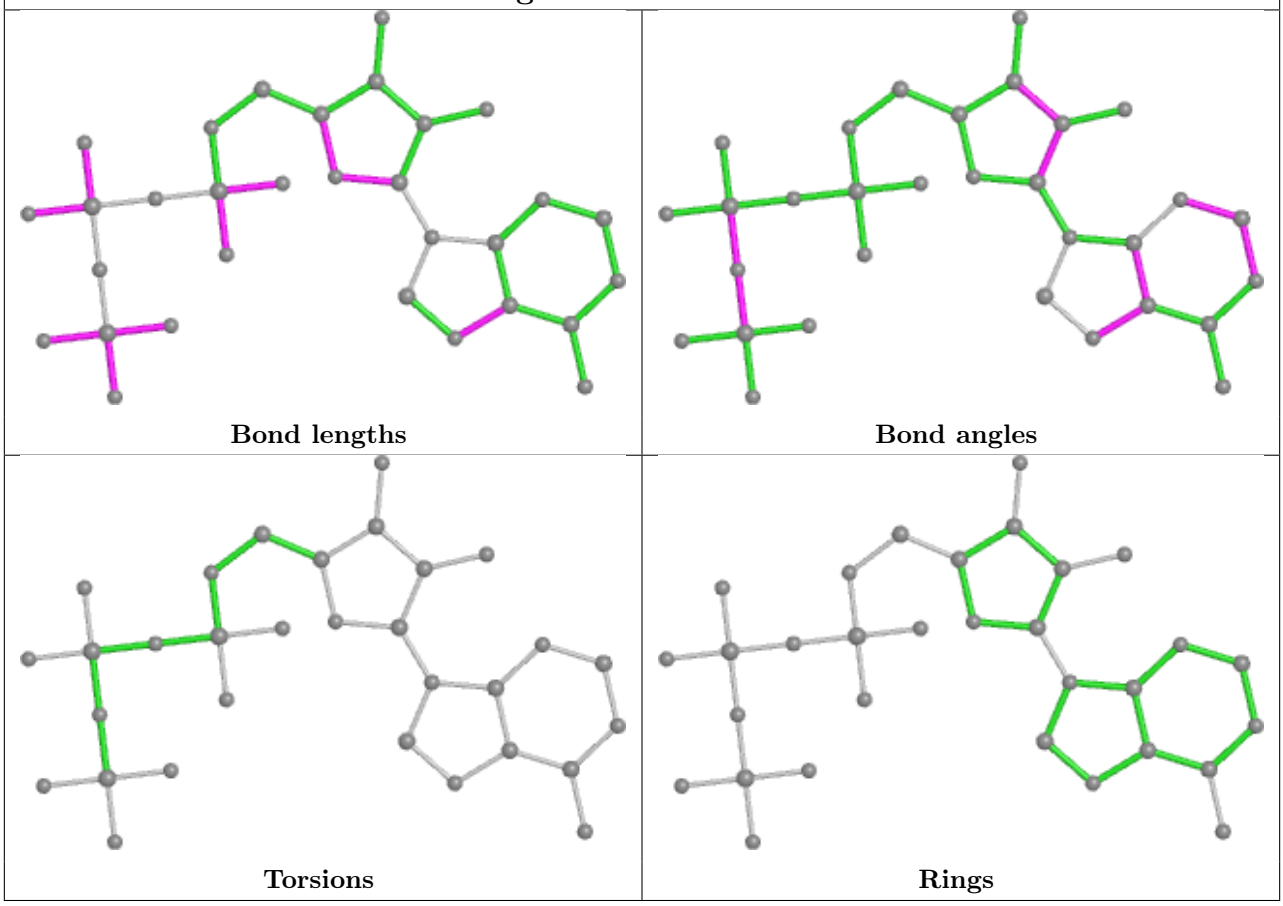
There are no torsion outliers.

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 A 1401



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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	272/321 (84%)	-0.50	6 (2%) 62 60	20, 29, 60, 80	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1296	PHE	4.6
1	A	1331	SER	3.4
1	A	1391	ARG	3.2
1	A	1297	ASP	2.8
1	A	1393	GLU	2.5

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 monosaccharides in this entry.

6.4 Ligands (i)

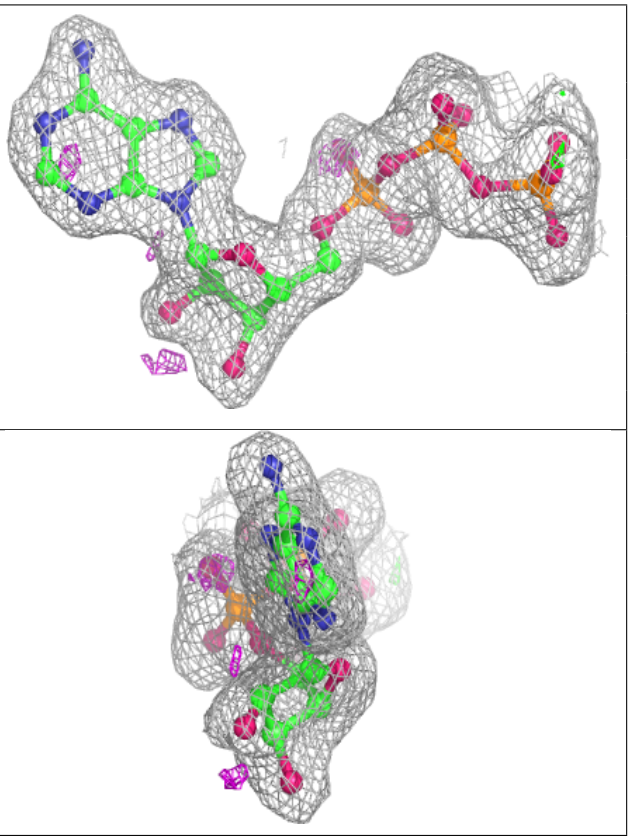
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists 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	ATP	A	1401	31/31	0.98	0.07	20,24,30,34	0
3	MG	A	1402	1/1	0.99	0.05	22,22,22,22	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.

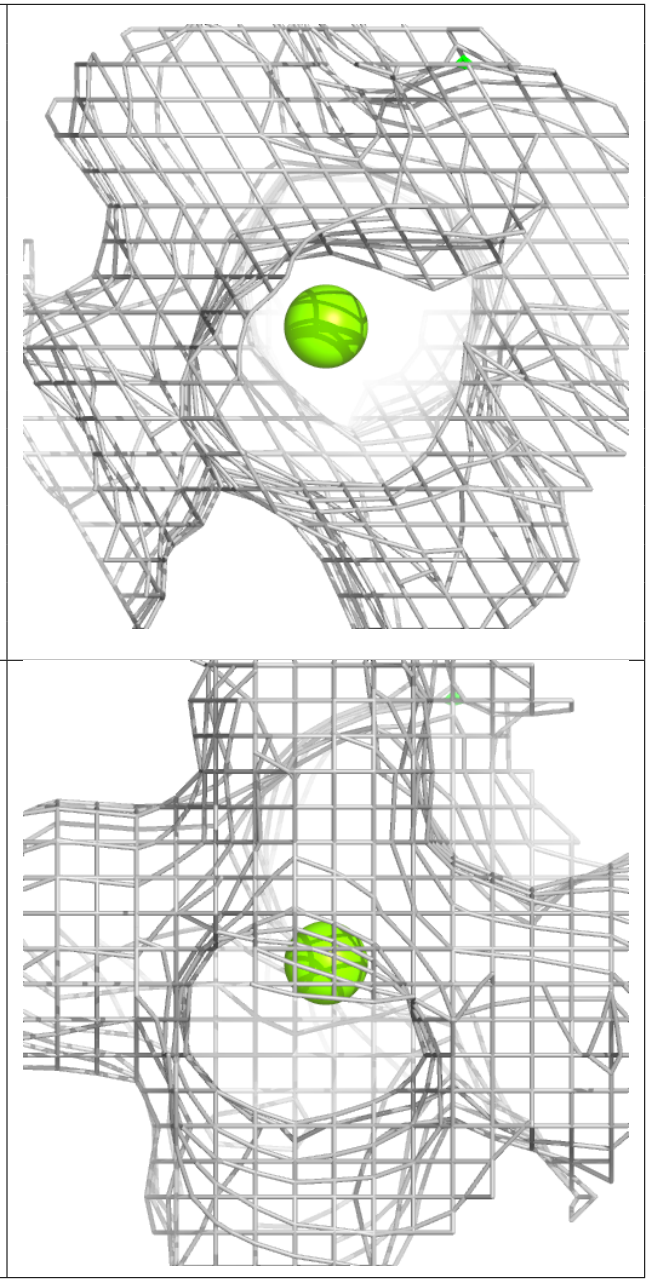
Electron density around ATP A 1401:

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



Electron density around MG A 1402:

$2mF_o$-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.