wwPDB X-ray Structure Validation Summary Report (i)

Nov 16, 2023 – 02:03 AM JST

PDB ID	:	6KID
Title	:	Crystal structure of human leucyl-tRNA synthetase, ATP-bound form
Authors	:	Kim, S.; Son, J.; Kim, S.; Hwang, K.Y.
Deposited on	:	2019-07-18
Resolution	:	3.15 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

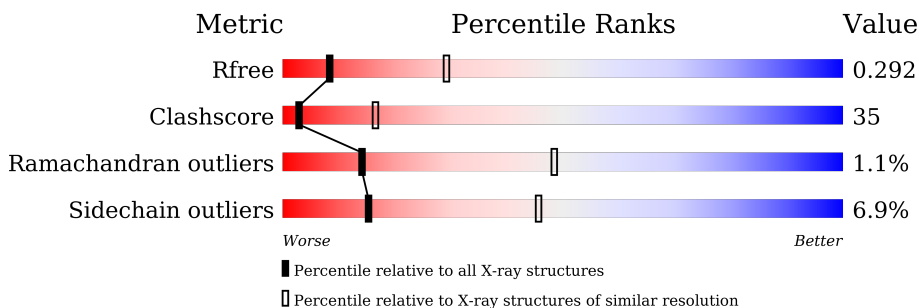|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh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

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 3.15 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	Percentile Ranks	Value
Rfree		0.292
Clashscore		35
Ramachandran outliers		1.1%
Sidechain outliers		6.9%

Worse   *Better*

▌Percentile relative to all X-ray structures

◻ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1665 (3.20-3.12)
Clashscore	141614	1804 (3.20-3.12)
Ramachandran outliers	138981	1770 (3.20-3.12)
Sidechain outliers	138945	1769 (3.20-3.12)

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%

Mol	Chain	Length	Quality of chain
1	A	1188	37% 43% 5% 15%

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 8353 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 Leucine–tRNA ligase, cytoplasmic.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1004	8103	5213	1346	1491	53	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MET	-	initiating methionine	UNP Q9P2J5
A	-10	ARG	-	expression tag	UNP Q9P2J5
A	-9	GLY	-	expression tag	UNP Q9P2J5
A	-8	SER	-	expression tag	UNP Q9P2J5
A	-7	HIS	-	expression tag	UNP Q9P2J5
A	-6	HIS	-	expression tag	UNP Q9P2J5
A	-5	HIS	-	expression tag	UNP Q9P2J5
A	-4	HIS	-	expression tag	UNP Q9P2J5
A	-3	HIS	-	expression tag	UNP Q9P2J5
A	-2	HIS	-	expression tag	UNP Q9P2J5
A	-1	GLY	-	expression tag	UNP Q9P2J5
A	0	SER	-	expression tag	UNP Q9P2J5

- 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).

ATP

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

- Molecule 3 is LEUCINE (three-letter code: LEU) (formula: $C_6H_{13}NO_2$).

LEU
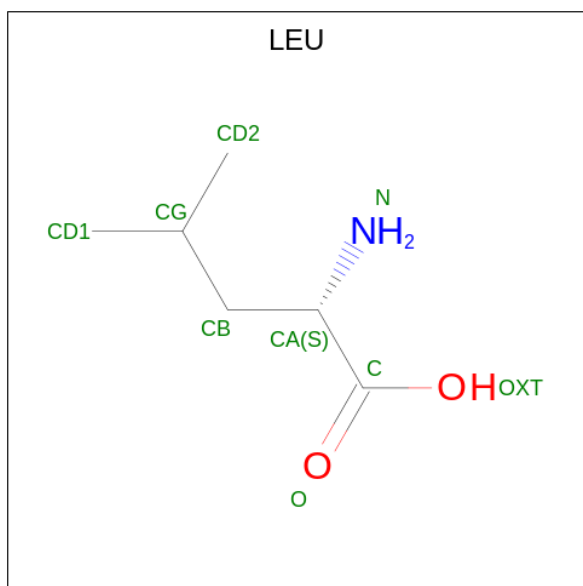
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			9	6	1	2		

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).

PO4

O3

O⁻

O2 O⁻ — P ═ O O1

P

O⁻

O4

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 5	O 4	P 1	0	0
4	A	1	Total 5	O 4	P 1	0	0

• Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	200	Total 200	O 200	0	0

3 Residue-property plots ⓘ

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. 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. 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: Leucine–tRNA ligase, cytoplasmic

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, α, β, γ	137.22Å 137.22Å 433.32Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	42.89 – 3.15 49.01 – 2.97	Depositor EDS
% Data completeness (in resolution range)	91.0 (42.89-3.15) 88.2 (49.01-2.97)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.40 (at 2.96Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.245 , 0.293 0.246 , 0.292	Depositor DCC
R_{free} test set	1832 reflections (4.09%)	wwPDB-VP
Wilson B-factor (Å^2)	32.6	Xtriage
Anisotropy	0.345	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.20 , 34.6	EDS
L-test for twinning[2]	$< |L| > = 0.31, < L^2 > = 0.15$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.64	EDS
Total number of atoms	8353	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.44% 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: PO4, 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	0.81	6/8303 (0.1%)	0.98	21/11222 (0.2%)

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	248	CYS	CB-SG	-10.06	1.65	1.82
1	A	305	CYS	CB-SG	-8.84	1.67	1.82
1	A	326	CYS	CB-SG	-6.48	1.71	1.82
1	A	546	CYS	CB-SG	-6.41	1.71	1.82
1	A	201	TRP	CB-CG	5.20	1.59	1.50

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	249	MET	CG-SD-CE	10.73	117.38	100.20
1	A	847	LEU	CA-CB-CG	8.86	135.69	115.30
1	A	282	LEU	CA-CB-CG	-8.23	96.36	115.30
1	A	575	ARG	NE-CZ-NH1	-8.07	116.27	120.30
1	A	249	MET	CB-CG-SD	7.01	133.44	112.40

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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	8103	0	8062	562	0
2	A	31	0	11	4	0
3	A	9	0	10	0	0
4	A	10	0	0	0	0
5	A	200	0	0	22	0
All	All	8353	0	8083	564	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 35.

The worst 5 of 564 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:1017:LYS:NZ	1:A:1038:VAL:CG1	1.68	1.55
1:A:1017:LYS:NZ	1:A:1038:VAL:HG12	1.03	1.36
1:A:1017:LYS:CE	1:A:1038:VAL:HG12	1.70	1.19
1:A:629:GLN:HE21	1:A:650:ILE:HD11	0.99	1.14
1:A:629:GLN:NE2	1:A:650:ILE:HD11	1.65	1.11

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	998/1188 (84%)	810 (81%)	177 (18%)	11 (1%)	14	48

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	587	GLU
1	A	978	SER
1	A	62	GLY

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	164	LEU
1	A	354	GLY

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	885/1047 (84%)	825 (93%)	60 (7%)	16	46

5 of 60 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	531	TYR
1	A	995	MET
1	A	645	PHE
1	A	978	SER
1	A	1052	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	629	GLN
1	A	729	GLN
1	A	892	ASN
1	A	793	ASN
1	A	513	GLN

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 ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

4 ligands 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	#$|Z| > 2$
4	PO4	A	1203	-	4,4,4	0.97	0	6,6,6	0.85	0
3	LEU	A	1202	-	7,8,8	1.05	1 (14%)	9,10,10	1.34	2 (22%)
2	ATP	A	1201	1	26,33,33	1.22	4 (15%)	31,52,52	1.78	7 (22%)
4	PO4	A	1204	-	4,4,4	0.71	0	6,6,6	0.67	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	LEU	A	1202	-	-	2/8/8/8	-
2	ATP	A	1201	1	-	6/18/38/38	0/3/3/3

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1201	ATP	C5-C4	2.63	1.47	1.40
2	A	1201	ATP	C2-N3	2.62	1.36	1.32
3	A	1202	LEU	OXT-C	-2.30	1.23	1.30
2	A	1201	ATP	C2'-C1'	-2.12	1.50	1.53
2	A	1201	ATP	O4'-C1'	2.09	1.44	1.41

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1201	ATP	N6-C6-N1	4.30	127.49	118.57
2	A	1201	ATP	N3-C2-N1	-3.34	123.47	128.68
2	A	1201	ATP	C5-C6-N6	-3.27	115.39	120.35
2	A	1201	ATP	C2'-C3'-C4'	3.11	108.69	102.64
2	A	1201	ATP	O3'-C3'-C2'	-3.06	101.91	111.82

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1201	ATP	C5'-O5'-PA-O1A
2	A	1201	ATP	C5'-O5'-PA-O2A
2	A	1201	ATP	C5'-O5'-PA-O3A
3	A	1202	LEU	CA-CB-CG-CD1
3	A	1202	LEU	CA-CB-CG-CD2

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1201	ATP	4	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 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 1201

Bond lengths

Bond angles

Torsions

Rings
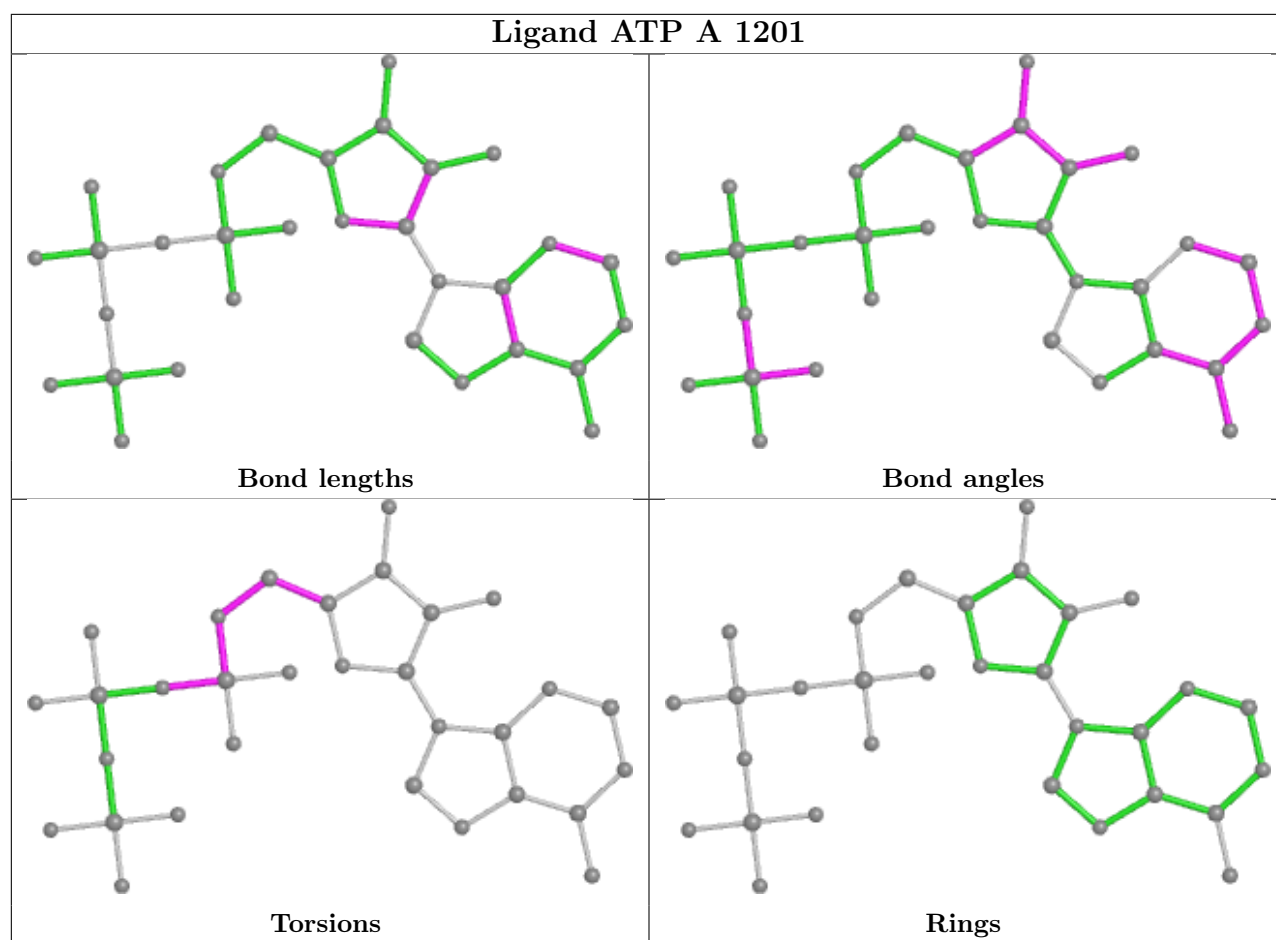
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

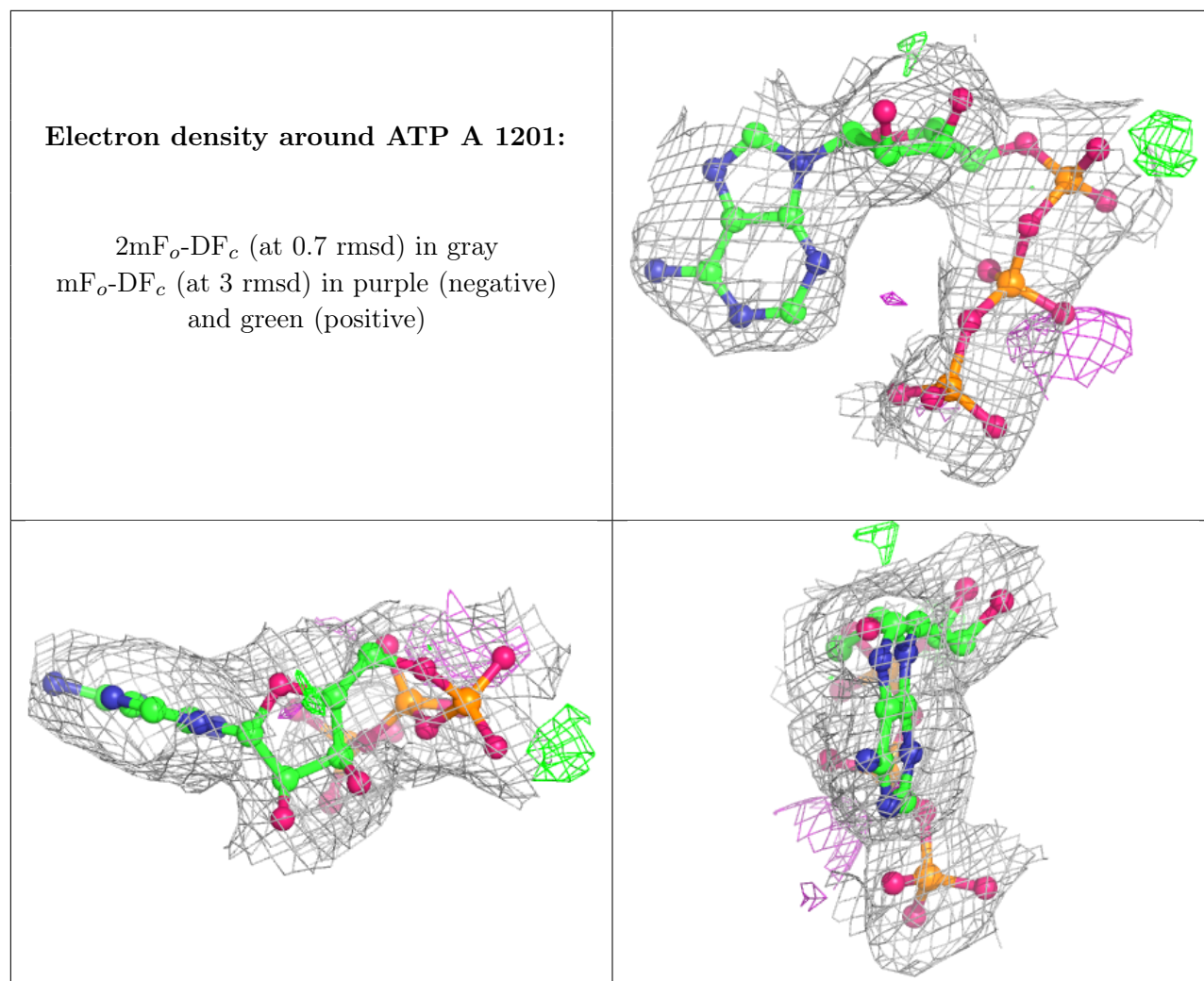Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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 1201:

$2mF_o$-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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.