W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

May 13, 2020 – 01:04 am BST

PDB ID	:	6FM0
Title	:	Deoxyguanylosuccinate synthase (DgsS) and ATP structure at 1.7 Angstrom resolution
Authors	:	Sleiman, D.; Loc'h, J.; Haouz, A.; Kaminski, P.A.
Deposited on	:	2018-01-29
Resolution	:	1.70 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	:	2.11
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 (i)

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

The reported resolution of this entry is 1.70 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.
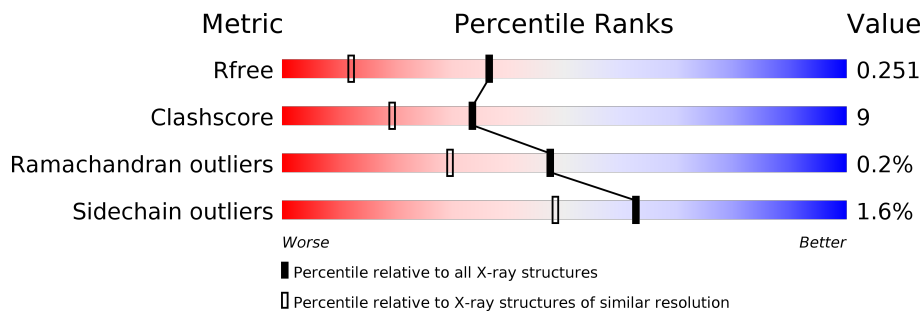
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)

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 >=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	363	67% 13% 19%
1	B	363	65% 14% • 20%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4540 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 Adenylosuccinate synthetase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	294	Total 2199	C 1401	N 373	O 405	S 20	0	0	0
1	B	290	Total 2130	C 1357	N 356	O 396	S 21	0	0	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP G3FFN6
A	-18	GLY	-	expression tag	UNP G3FFN6
A	-17	SER	-	expression tag	UNP G3FFN6
A	-16	SER	-	expression tag	UNP G3FFN6
A	-15	HIS	-	expression tag	UNP G3FFN6
A	-14	HIS	-	expression tag	UNP G3FFN6
A	-13	HIS	-	expression tag	UNP G3FFN6
A	-12	HIS	-	expression tag	UNP G3FFN6
A	-11	HIS	-	expression tag	UNP G3FFN6
A	-10	HIS	-	expression tag	UNP G3FFN6
A	-9	SER	-	expression tag	UNP G3FFN6
A	-8	SER	-	expression tag	UNP G3FFN6
A	-7	GLY	-	expression tag	UNP G3FFN6
A	-6	LEU	-	expression tag	UNP G3FFN6
A	-5	VAL	-	expression tag	UNP G3FFN6
A	-4	PRO	-	expression tag	UNP G3FFN6
A	-3	ARG	-	expression tag	UNP G3FFN6
A	-2	GLY	-	expression tag	UNP G3FFN6
A	-1	SER	-	expression tag	UNP G3FFN6
A	0	HIS	-	expression tag	UNP G3FFN6
A	1	MET	-	expression tag	UNP G3FFN6
A	2	GLU	-	expression tag	UNP G3FFN6
B	-19	MET	-	initiating methionine	UNP G3FFN6
B	-18	GLY	-	expression tag	UNP G3FFN6
B	-17	SER	-	expression tag	UNP G3FFN6

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-16	SER	-	expression tag	UNP G3FFN6
B	-15	HIS	-	expression tag	UNP G3FFN6
B	-14	HIS	-	expression tag	UNP G3FFN6
B	-13	HIS	-	expression tag	UNP G3FFN6
B	-12	HIS	-	expression tag	UNP G3FFN6
B	-11	HIS	-	expression tag	UNP G3FFN6
B	-10	HIS	-	expression tag	UNP G3FFN6
B	-9	SER	-	expression tag	UNP G3FFN6
B	-8	SER	-	expression tag	UNP G3FFN6
B	-7	GLY	-	expression tag	UNP G3FFN6
B	-6	LEU	-	expression tag	UNP G3FFN6
B	-5	VAL	-	expression tag	UNP G3FFN6
B	-4	PRO	-	expression tag	UNP G3FFN6
B	-3	ARG	-	expression tag	UNP G3FFN6
B	-2	GLY	-	expression tag	UNP G3FFN6
B	-1	SER	-	expression tag	UNP G3FFN6
B	0	HIS	-	expression tag	UNP G3FFN6
B	1	MET	-	expression tag	UNP G3FFN6
B	2	GLU	-	expression tag	UNP G3FFN6

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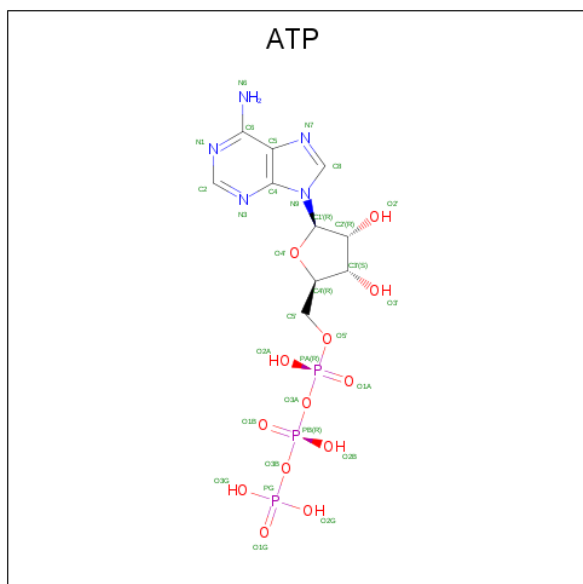
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	A	1	31	10	5	13	3	0	0

Continued from previous page...

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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	67	Total	O	0	0
			67	67		
3	B	82	Total	O	0	0
			82	82		

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. 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: Adenylosuccinate synthetase

Chain A:

- Molecule 1: Adenylosuccinate synthetase

Chain B:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	158.86Å 60.41Å 91.03Å 90.00° 114.76° 90.00°	Depositor
Resolution (Å)	45.60 – 1.70 44.87 – 1.70	Depositor EDS
% Data completeness (in resolution range)	97.0 (45.60-1.70) 48.6 (44.87-1.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.33 (at 1.70Å)	Xtriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
R, R_{free}	(Not available) , (Not available) 0.224 , 0.251	Depositor DCC
R_{free} test set	353 reflections (0.84%)	wwPDB-VP
Wilson B-factor (Å^2)	36.7	Xtriage
Anisotropy	0.403	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 57.3	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	0.96	EDS
Total number of atoms	4540	wwPDB-VP
Average B, all atoms (Å^2)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 72.17 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.3083e-06. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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

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	0.41	0/2246	0.61	0/3058
1	B	0.52	0/2172	0.61	0/2956
All	All	0.47	0/4418	0.61	0/6014

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	2199	0	2056	32	0
1	B	2130	0	1970	40	0
2	A	31	0	12	2	0
2	B	31	0	12	2	0
3	A	67	0	0	3	0
3	B	82	0	0	4	0
All	All	4540	0	4050	72	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 9.

The worst 5 of 72 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:279:MET:O	1:A:283:MET:HG2	1.80	0.82
1:B:4:VAL:HG22	1:B:222:LEU:HA	1.65	0.78
1:A:187:GLN:NE2	1:A:202:THR:HG21	1.99	0.77
1:B:37:ASN:ND2	3:B:501:HOH:O	2.22	0.71
1:A:332:THR:HG22	1:A:334:ASN:H	1.56	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	286/363 (79%)	274 (96%)	12 (4%)	0	100	100
1	B	280/363 (77%)	270 (96%)	9 (3%)	1 (0%)	34	18
All	All	566/726 (78%)	544 (96%)	21 (4%)	1 (0%)	47	30

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	187	GLN

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	218/306 (71%)	215 (99%)	3 (1%)	67	53
1	B	207/306 (68%)	203 (98%)	4 (2%)	57	41
All	All	425/612 (69%)	418 (98%)	7 (2%)	62	48

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	17	LYS
1	B	281	LYS
1	B	81	ASN
1	A	147	ASN
1	B	235	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	43	HIS
1	A	287	GLN
1	B	278	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 (i)

There are no carbohydrates in this entry.

5.6   Ligand geometry (i)

2 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	$\#|Z| > 2$
2	ATP	A	401	-	26,33,33	1.51	6 (23%)	31,52,52	1.79	9 (29%)
2	ATP	B	401	-	26,33,33	0.72	0	31,52,52	1.06	3 (9%)

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	401	-	-	5/18/38/38	0/3/3/3
2	ATP	B	401	-	-	4/18/38/38	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	ATP	PB-O1B	-2.89	1.40	1.50
2	A	401	ATP	PB-O2B	-2.88	1.41	1.55
2	A	401	ATP	C5-N7	-2.43	1.30	1.39
2	A	401	ATP	O4'-C1'	2.25	1.44	1.41
2	A	401	ATP	C2'-C1'	-2.18	1.50	1.53

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	401	ATP	O5'-C5'-C4'	5.38	127.49	108.99
2	A	401	ATP	N3-C2-N1	-3.08	123.87	128.68
2	B	401	ATP	PA-O3A-PB	-3.03	122.44	132.83
2	A	401	ATP	O3G-PG-O3B	2.80	114.02	104.64
2	A	401	ATP	C4-C5-N7	-2.70	106.58	109.40

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	401	ATP	C5'-O5'-PA-O1A
2	A	401	ATP	C5'-O5'-PA-O2A

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	B	401	ATP	C5'-O5'-PA-O1A
2	A	401	ATP	O4'-C4'-C5'-O5'
2	A	401	ATP	C3'-C4'-C5'-O5'

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	401	ATP	2	0
2	B	401	ATP	2	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 401

Bond lengths

Bond angles

Torsions

Rings

Ligand ATP B 401

Bond lengths

Bond angles

Torsions

Rings
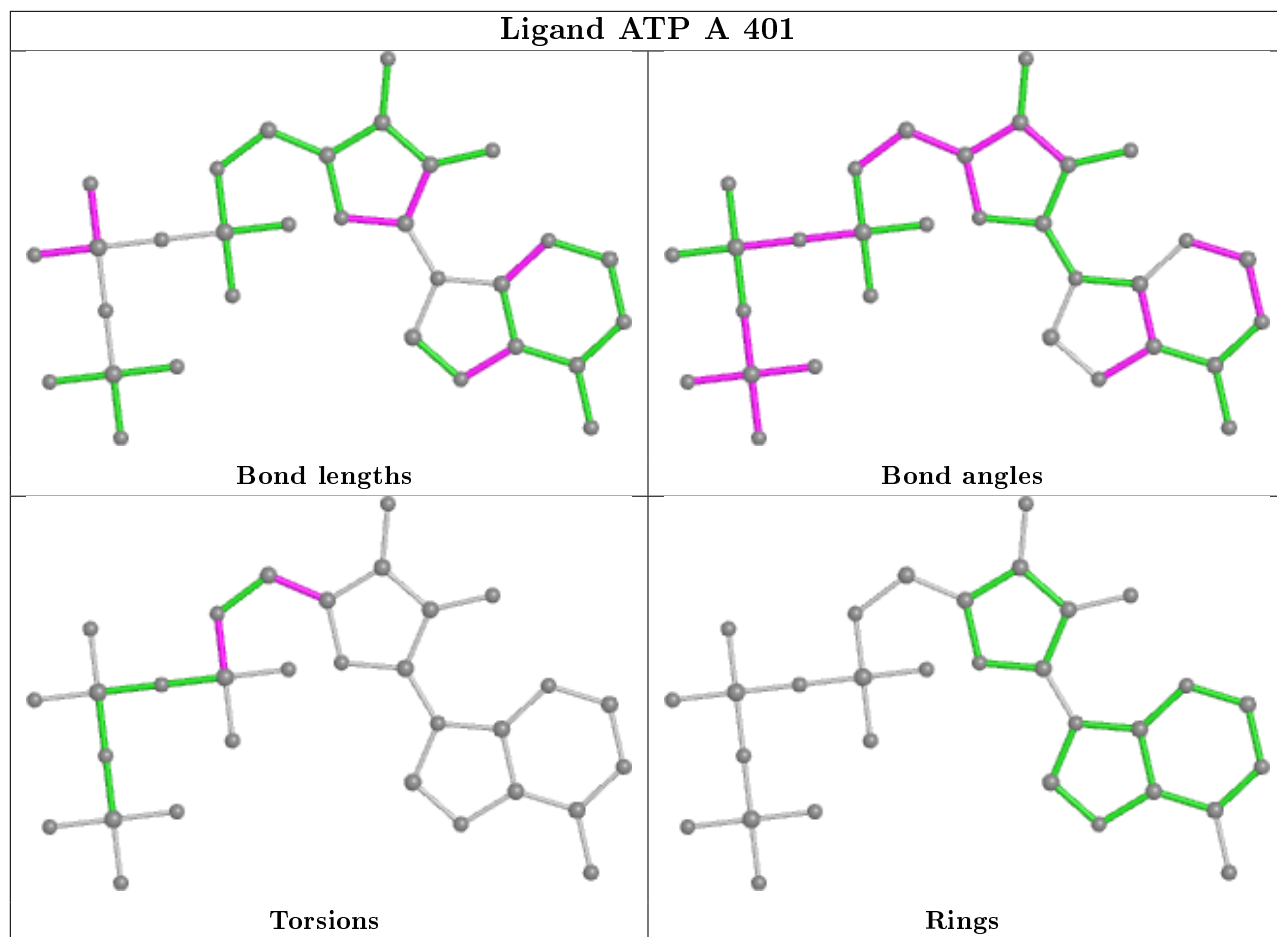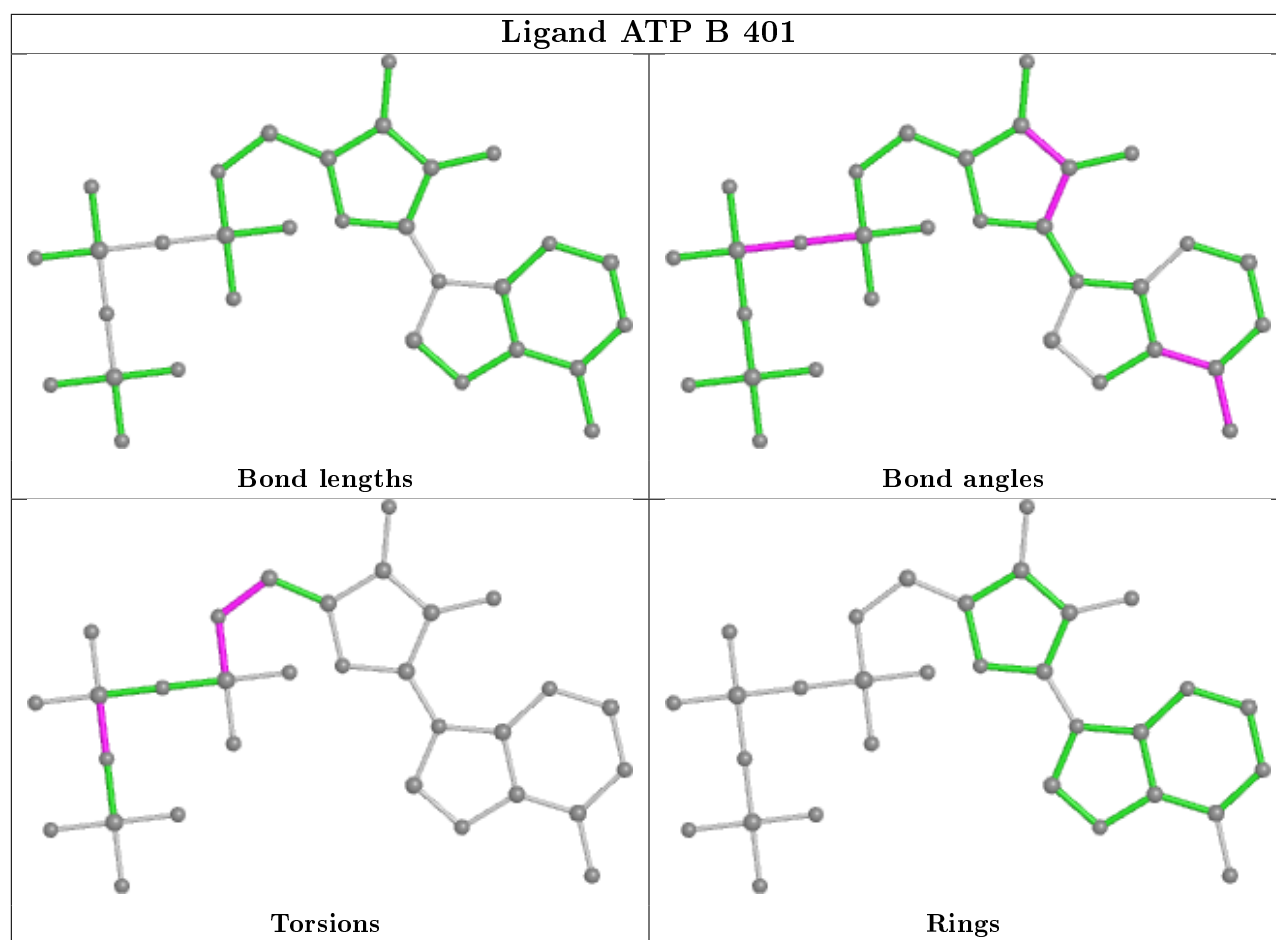
5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.
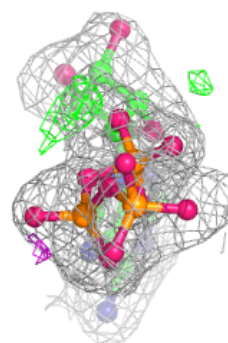
6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

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

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

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

6.3 Carbohydrates (i)

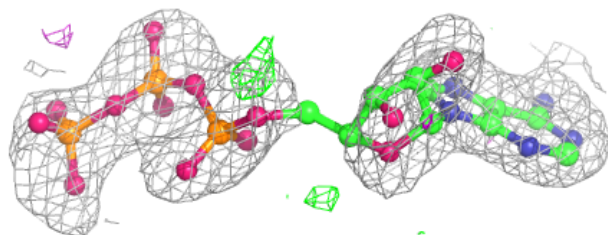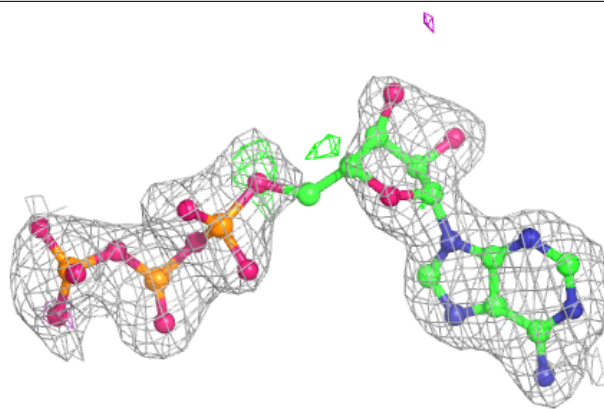Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands (i)

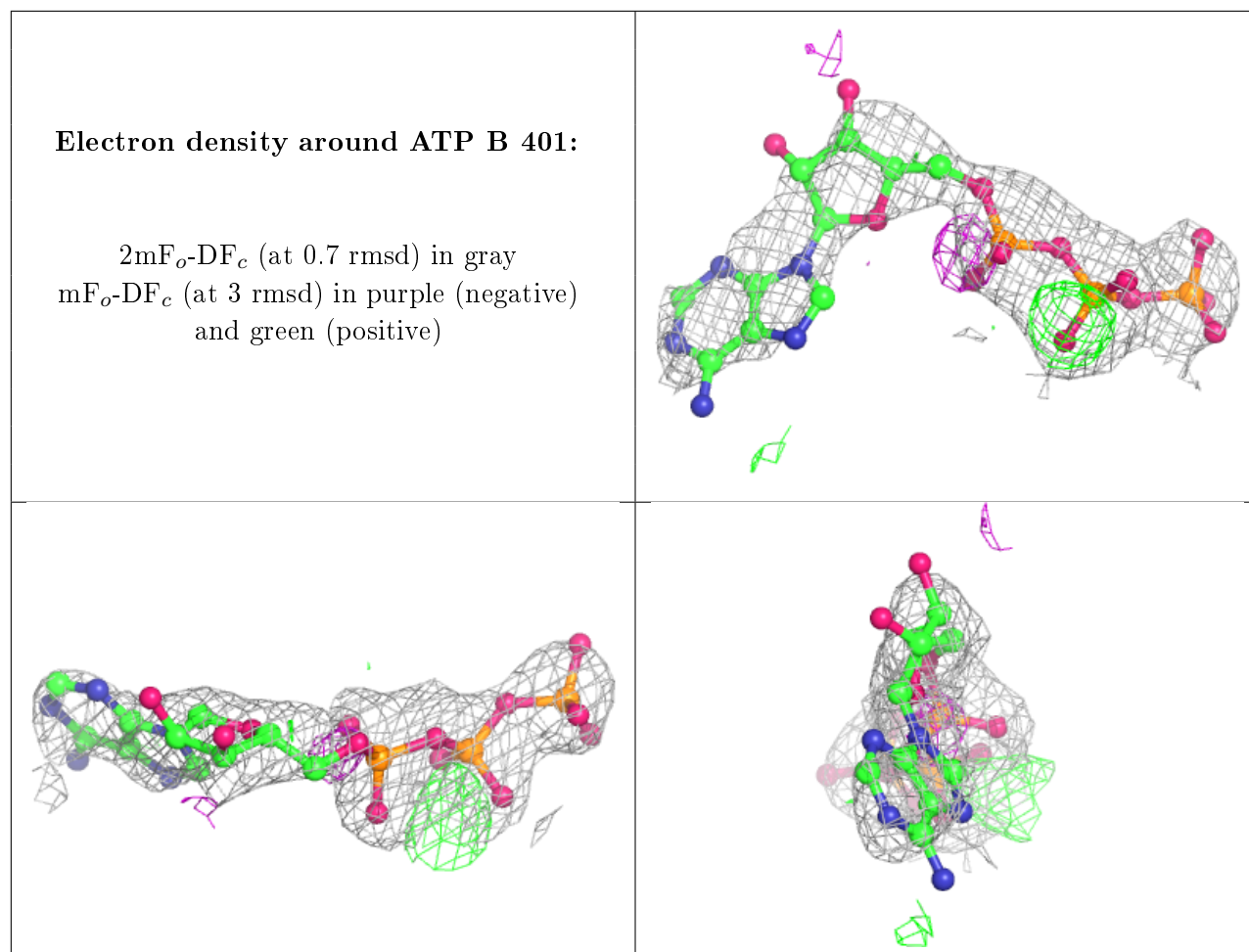Unable to reproduce the depositors R factor - this section is therefore empty.

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 401:

$2mF_o$-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 401:

$2mF_o$-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.