



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 15, 2024 – 04:33 PM EDT

PDB ID : 8SJ1
Title : Structure of ternary complex of cGAS with dsDNA and bound 3'-dATP
Authors : Wu, S.; Sohn, J.
Deposited on : 2023-04-17
Resolution : 2.81 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtrriage (Phenix) : 1.13
EDS : 2.36.1
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

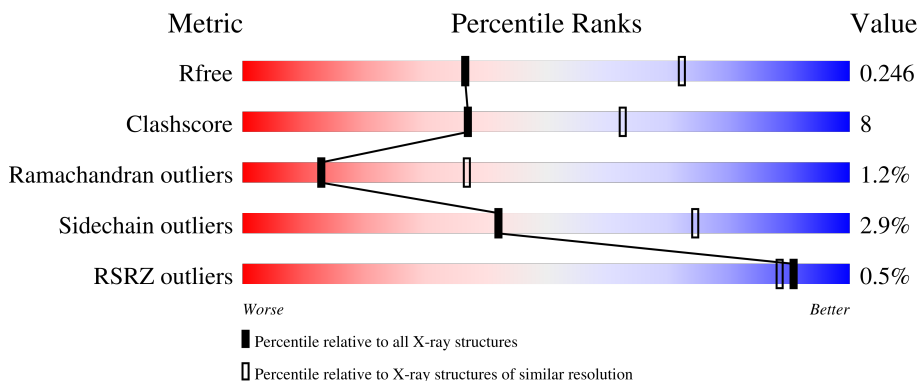
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION


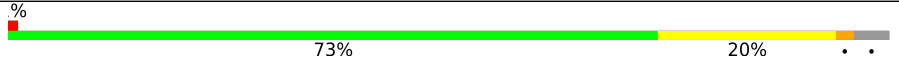
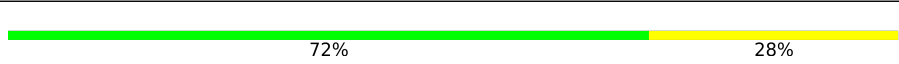
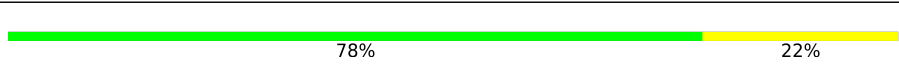
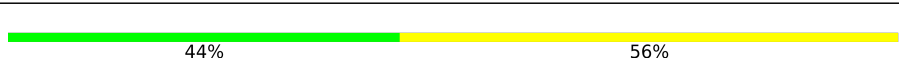
The reported resolution of this entry is 2.81 Å.

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.



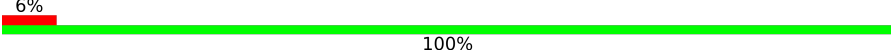
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3617 (2.84-2.80)
Clashscore	141614	4060 (2.84-2.80)
Ramachandran outliers	138981	3978 (2.84-2.80)
Sidechain outliers	138945	3980 (2.84-2.80)
RSRZ outliers	127900	3552 (2.84-2.80)

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 $\leq 5\%$. 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	364	 77% 19% ..
1	C	364	 73% 20% ..
2	E	18	 72% 28%
2	F	18	 78% 22%
2	I	18	 44% 56%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	J	18	 6% 100%

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 7372 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 Cyclic GMP-AMP synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	356	Total	C	N	O	S	0	0	0
			2947	1895	500	539	13			
1	C	348	Total	C	N	O	S	0	0	0
			2881	1855	487	526	13			

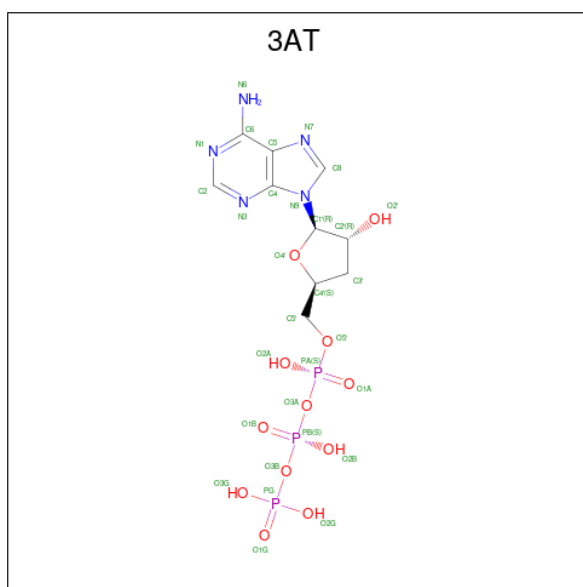
There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	144	GLY	-	expression tag	UNP Q8C6L5
A	145	THR	-	expression tag	UNP Q8C6L5
A	146	GLY	-	expression tag	UNP Q8C6L5
C	144	GLY	-	expression tag	UNP Q8C6L5
C	145	THR	-	expression tag	UNP Q8C6L5
C	146	GLY	-	expression tag	UNP Q8C6L5

- Molecule 2 is a DNA chain called Palindromic DNA18.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	E	18	Total	C	N	O	P	0	0	0
			366	177	66	106	17			
2	F	18	Total	C	N	O	P	0	0	0
			366	177	66	106	17			
2	I	18	Total	C	N	O	P	0	0	0
			366	177	66	106	17			
2	J	18	Total	C	N	O	P	0	0	0
			366	177	66	106	17			

- Molecule 3 is 3'-DEOXYADENOSINE-5'-TRIPHOSPHATE (three-letter code: 3AT) (formula: C₁₀H₁₆N₅O₁₂P₃) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
3	A	1	30	10	5	12	3	0	0
3	C	1	30	10	5	12	3	0	0

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
4	A	1	1	1	0	0
4	C	1	1	1	0	0

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
5	A	1	1	1	0	0
5	C	1	1	1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	A	6	6	6	0	0

Continued on next page...

Continued from previous page...

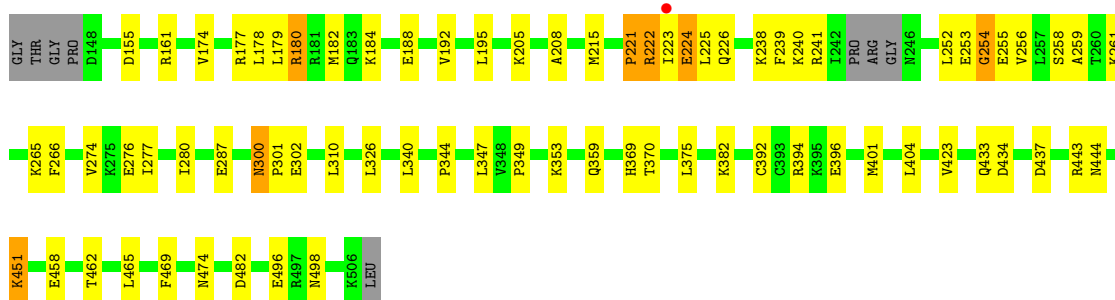
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	C	8	Total O 8 8	0	0
6	E	1	Total O 1 1	0	0
6	I	1	Total O 1 1	0	0

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 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.

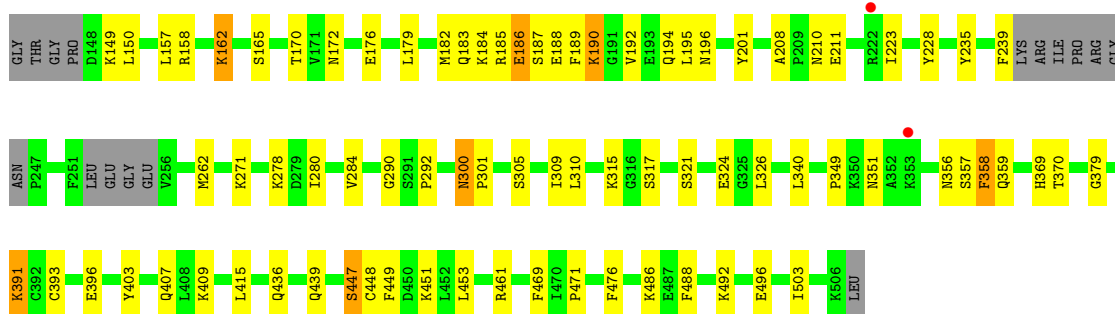
- Molecule 1: Cyclic GMP-AMP synthase

Chain A: 



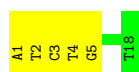
- Molecule 1: Cyclic GMP-AMP synthase

Chain C: 




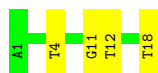
- Molecule 2: Palindromic DNA18

Chain E: 

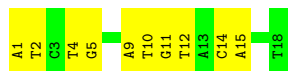


- Molecule 2: Palindromic DNA18

Chain F: 



- Molecule 2: Palindromic DNA18



- Molecule 2: Palindromic DNA18



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	77.18Å 97.58Å 142.28Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.74 – 2.81 28.74 – 2.81	Depositor EDS
% Data completeness (in resolution range)	77.6 (28.74-2.81) 77.6 (28.74-2.81)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.87 (at 2.80Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.188 , 0.240 0.192 , 0.246	Depositor DCC
R_{free} test set	1093 reflections (5.22%)	wwPDB-VP
Wilson B-factor (Å ²)	55.7	Xtrriage
Anisotropy	0.248	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 41.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7372	wwPDB-VP
Average B, all atoms (Å ²)	58.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 27.78 % 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 2.0709e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical 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: 3AT, MG, ZN

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	0.53	0/3009	0.69	2/4037 (0.0%)
1	C	0.45	0/2942	0.63	0/3946
2	E	0.90	0/410	1.10	0/631
2	F	0.85	0/410	1.01	0/631
2	I	0.86	0/410	1.08	0/631
2	J	0.93	0/410	1.08	0/631
All	All	0.60	0/7591	0.78	2/10507 (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	A	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	375	LEU	CA-CB-CG	-5.51	102.62	115.30
1	A	221	PRO	CA-N-CD	-5.24	104.17	111.50

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	300	ASN	Peptide

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	2947	0	2991	50	0
1	C	2881	0	2922	47	0
2	E	366	0	206	4	0
2	F	366	0	206	4	0
2	I	366	0	206	7	0
2	J	366	0	206	0	0
3	A	30	0	12	1	0
3	C	30	0	12	0	0
4	A	1	0	0	0	0
4	C	1	0	0	0	0
5	A	1	0	0	0	0
5	C	1	0	0	0	0
6	A	6	0	0	0	0
6	C	8	0	0	0	0
6	E	1	0	0	0	0
6	I	1	0	0	1	0
All	All	7372	0	6761	109	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 8.

The worst 5 of 109 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:340:LEU:HD12	1:A:370:THR:HG23	1.43	0.97
1:C:391:LYS:HD2	1:C:391:LYS:H	1.42	0.84
1:C:149:LYS:HG3	1:C:150:LEU:H	1.49	0.78
1:A:222:ARG:HE	1:A:241:ARG:H	1.33	0.76
1:C:172:ASN:O	1:C:176:GLU:HG3	1.88	0.74

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	352/364 (97%)	331 (94%)	17 (5%)	4 (1%)	14	39
1	C	342/364 (94%)	325 (95%)	13 (4%)	4 (1%)	13	37
All	All	694/728 (95%)	656 (94%)	30 (4%)	8 (1%)	13	37

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	301	PRO
1	C	186	GLU
1	A	224	GLU
1	C	356	ASN
1	C	300	ASN

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	330/335 (98%)	321 (97%)	9 (3%)	44	77
1	C	323/335 (96%)	313 (97%)	10 (3%)	40	72
All	All	653/670 (98%)	634 (97%)	19 (3%)	42	74

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	317	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	447	SER
1	C	461	ARG
1	C	403	TYR
1	A	498	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	436	GLN
1	C	250	HIS
1	A	498	ASN
1	A	474	ASN
1	C	194	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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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
3	3AT	A	601	5	25,32,32	0.76	0	28,50,50	0.83	1 (3%)
3	3AT	C	601	5	25,32,32	0.58	0	28,50,50	0.80	2 (7%)

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	3AT	A	601	5	-	2/18/34/34	0/3/3/3
3	3AT	C	601	5	-	3/18/34/34	0/3/3/3

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	601	3AT	C5-C6-N6	2.28	123.82	120.35
3	A	601	3AT	C5-C6-N6	2.23	123.75	120.35
3	C	601	3AT	PB-O3B-PG	2.10	140.03	132.83

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	601	3AT	PB-O3B-PG-O2G
3	A	601	3AT	PB-O3B-PG-O1G
3	C	601	3AT	PB-O3A-PA-O1A
3	A	601	3AT	PA-O3A-PB-O2B
3	C	601	3AT	PB-O3A-PA-O2A

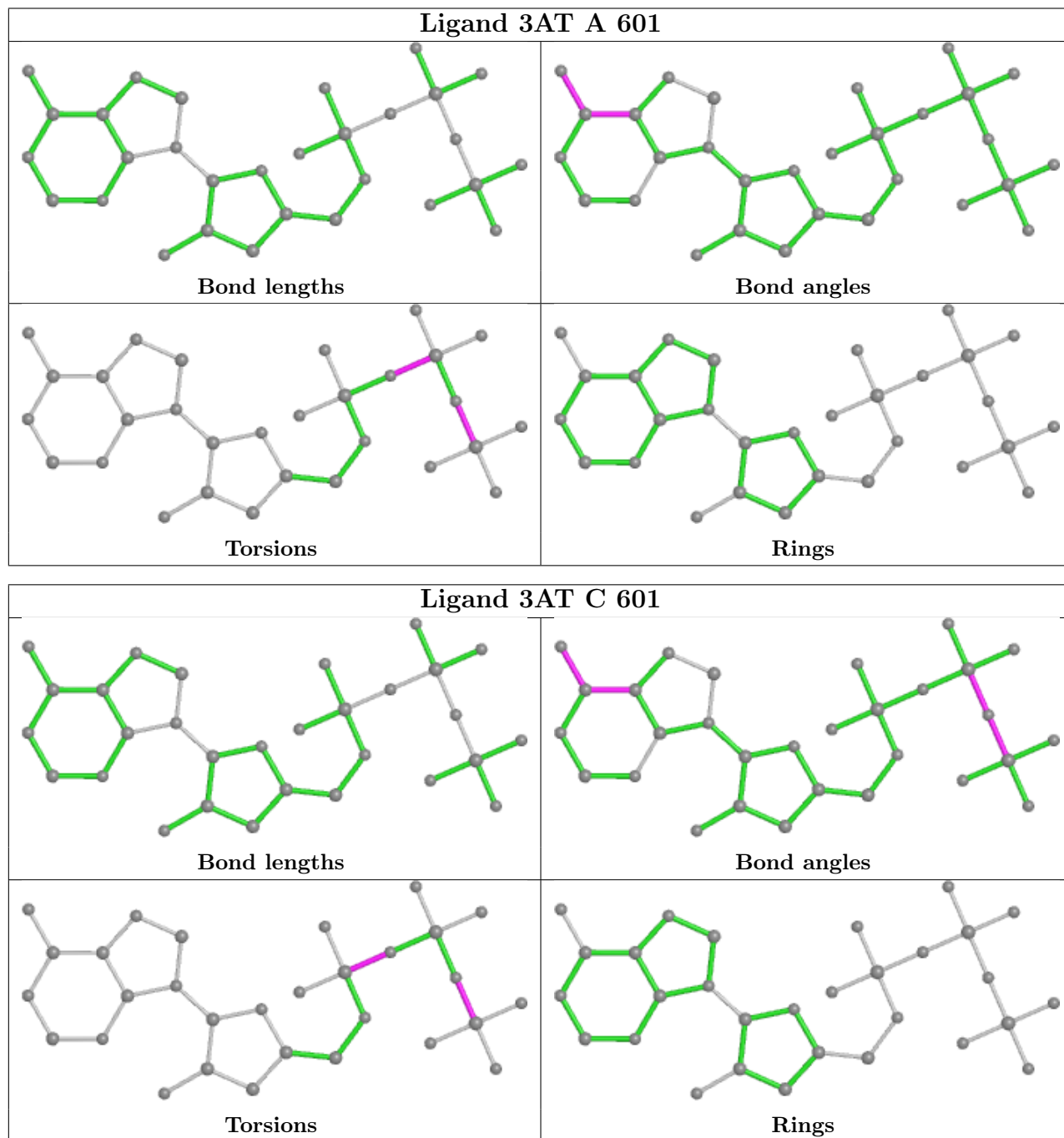
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	601	3AT	1	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 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

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 [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	356/364 (97%)	-0.52	1 (0%) 94 93	28, 47, 79, 115	0
1	C	348/364 (95%)	-0.40	2 (0%) 89 86	31, 55, 93, 112	0
2	E	18/18 (100%)	-0.06	0 100 100	39, 68, 114, 115	0
2	F	18/18 (100%)	-0.24	0 100 100	40, 65, 110, 126	0
2	I	18/18 (100%)	0.01	0 100 100	39, 56, 115, 116	0
2	J	18/18 (100%)	-0.17	1 (5%) 24 16	43, 69, 102, 102	0
All	All	776/800 (97%)	-0.43	4 (0%) 91 88	28, 52, 95, 126	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	353	LYS	2.7
1	A	223	ILE	2.6
1	C	222	ARG	2.6
2	J	18	DT	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 monosaccharides in this entry.

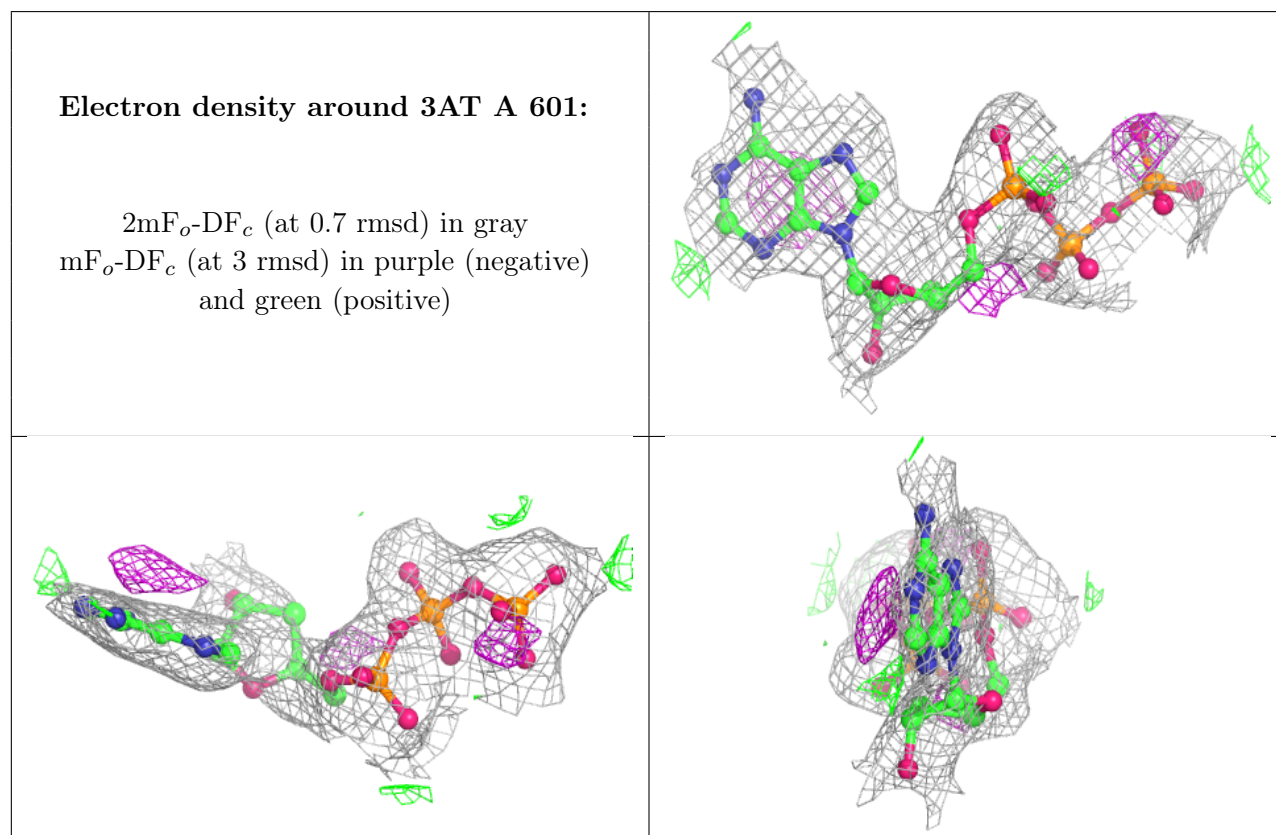
6.4 Ligands [i](#)

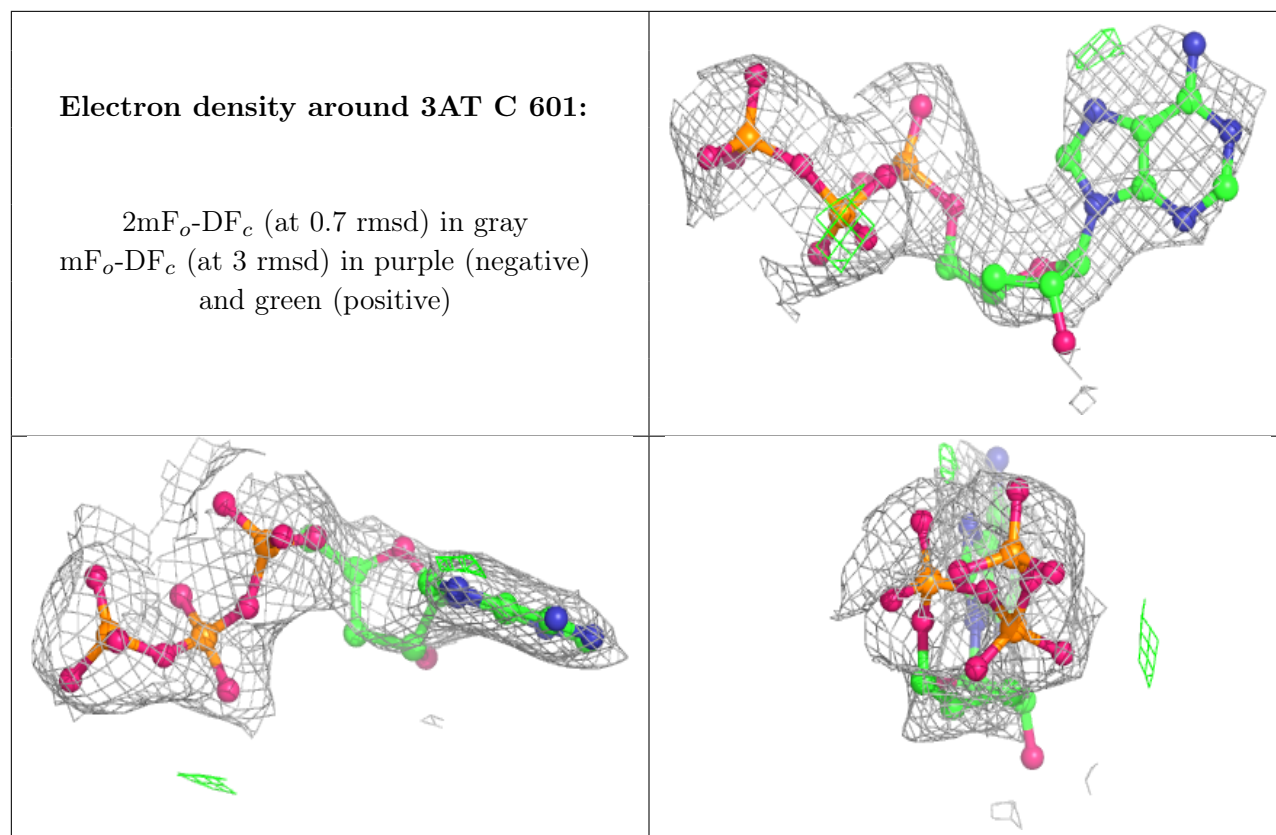
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,

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
5	MG	A	603	1/1	0.92	0.15	43,43,43,43	0
3	3AT	A	601	30/30	0.93	0.15	48,80,96,97	0
3	3AT	C	601	30/30	0.95	0.17	60,93,107,108	0
5	MG	C	603	1/1	0.97	0.18	52,52,52,52	0
4	ZN	A	602	1/1	0.99	0.19	49,49,49,49	0
4	ZN	C	602	1/1	1.00	0.10	39,39,39,39	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.