



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 15, 2023 – 07:44 PM EDT

PDB ID : 8SX5
Title : GpppA dinucleotide binding to RNA CU template
Authors : Zhang, W.; Dantsu, Y.
Deposited on : 2023-05-19
Resolution : 1.95 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)
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

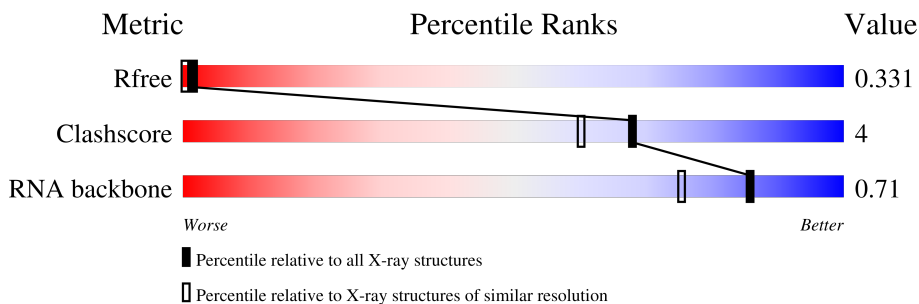
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.95 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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
RNA backbone	3102	1124 (2.50-1.42)

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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	14	71% 21% 7%
1	B	14	71% 21% 7%

2 Entry composition [i](#)

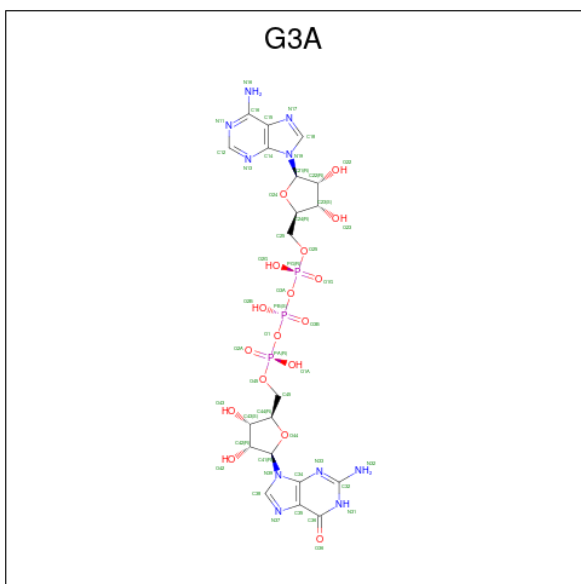
There are 3 unique types of molecules in this entry. The entry contains 706 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 RNA chain called RNA (5'-R*(TLN)P*(LCC)P*(LCC)P*(LCG)P*AP*CP*UP*UP*AP*AP*GP*UP*CP*G*(GA3))-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	A	14	Total	C	N	O	P	0	0	0
			299	139	50	97	13			
1	B	14	Total	C	N	O	P	0	0	0
			299	139	50	97	13			

- Molecule 2 is GUANOSINE-P3-ADENOSINE-5',5'-TRIPHOSPHATE (three-letter code: G3A) (formula: C₂₀H₂₇N₁₀O₁₇P₃) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	A	1	Total	C	N	O	P	0	0
			50	20	10	17	3		
2	B	1	Total	C	N	O	P	0	0
			50	20	10	17	3		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	2	Total O 2 2	0	0
3	B	6	Total O 6 6	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: RNA (5'-R*(TLN)P*(LCC)P*(LCC)P*(LCG)P*AP*CP*UP*UP*AP*AP*GP*UP*CP*G*(GA3))-3')

Chain A:  71% 21% 7%



- Molecule 1: RNA (5'-R*(TLN)P*(LCC)P*(LCC)P*(LCG)P*AP*CP*UP*UP*AP*AP*GP*UP*CP*G*(GA3))-3')

Chain B:  71% 21% 7%



4 Data and refinement statistics i

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, α , β , γ	44.52Å 44.52Å 85.52Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	28.52 – 1.95 28.51 – 1.95	Depositor EDS
% Data completeness (in resolution range)	90.4 (28.52-1.95) 90.5 (28.51-1.95)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	10.93 (at 1.95Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.284 , 0.323 0.294 , 0.331	Depositor DCC
R_{free} test set	301 reflections (4.36%)	wwPDB-VP
Wilson B-factor (Å ²)	28.9	Xtrriage
Anisotropy	0.154	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 28.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.22$	Xtrriage
Estimated twinning fraction	0.375 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	706	wwPDB-VP
Average B, all atoms (Å ²)	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.27% of the height of the origin peak. No significant pseudotranslation is detected.*

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: LCC, G3A, TLN, LCG

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	0.67	0/236	0.89	0/365
1	B	0.69	0/236	0.94	0/365
All	All	0.68	0/472	0.91	0/730

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

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	299	0	157	2	0
1	B	299	0	157	0	0
2	A	50	0	24	3	0
2	B	50	0	24	0	0
3	A	2	0	0	0	0
3	B	6	0	0	0	0
All	All	706	0	362	4	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.

All (4) 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:1:LCC:H42	2:A:101:G3A:HN31	1.57	0.53
1:A:1:LCC:H41	1:A:1:LCC:H5M1	1.58	0.46
2:A:101:G3A:H23	2:A:101:G3A:H41	2.00	0.44
2:A:101:G3A:H23	2:A:101:G3A:C41	2.49	0.42

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

There are no protein molecules in this entry.

5.3.2 Protein sidechains [i](#)

There are no protein molecules in this entry.

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	9/14 (64%)	0	0
1	B	10/14 (71%)	0	1 (10%)
All	All	19/28 (67%)	0	1 (5%)

There are no RNA backbone outliers to report.

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	B	1	LCC

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

8 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
1	LCG	A	4	1	19,27,28	2.46	5 (26%)	22,42,45	1.84	8 (36%)
1	LCC	A	1	1	20,21,25	2.53	7 (35%)	30,33,40	1.76	6 (20%)
1	LCC	A	3	1	21,24,25	2.34	9 (42%)	28,37,40	1.67	5 (17%)
1	TLN	A	2	1	21,24,25	1.48	4 (19%)	30,37,40	2.24	9 (30%)
1	LCG	B	4	1	19,27,28	2.56	7 (36%)	22,42,45	1.98	8 (36%)
1	TLN	B	2	1	21,24,25	2.42	6 (28%)	30,37,40	2.30	12 (40%)
1	LCC	B	1	1	20,21,25	1.69	5 (25%)	30,33,40	2.32	9 (30%)
1	LCC	B	3	1	21,24,25	2.65	5 (23%)	28,37,40	1.82	9 (32%)

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
1	LCG	A	4	1	-	0/4/35/36	0/5/4/4
1	LCC	A	1	1	-	3/7/32/36	0/4/3/3
1	LCC	A	3	1	-	0/8/35/36	0/4/3/3
1	TLN	A	2	1	-	0/8/35/36	0/4/3/3
1	LCG	B	4	1	-	2/4/35/36	0/5/4/4
1	TLN	B	2	1	-	0/8/35/36	0/4/3/3
1	LCC	B	1	1	-	0/7/32/36	0/4/3/3
1	LCC	B	3	1	-	0/8/35/36	0/4/3/3

The worst 5 of 48 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	4	LCG	C5-C6	-7.70	1.31	1.47
1	B	3	LCC	C4'-C3'	-6.60	1.45	1.53
1	A	4	LCG	C5-C6	-6.49	1.34	1.47
1	B	3	LCC	O2'-C6'	5.99	1.54	1.43
1	B	3	LCC	O3'-C3'	5.94	1.54	1.42

The worst 5 of 66 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1	LCC	O4'-C1'-N1	6.05	116.56	108.77
1	B	2	TLN	C5M-C5-C4	5.22	124.51	118.77
1	A	2	TLN	C4-N3-C2	-5.15	120.68	127.35
1	B	2	TLN	C5-C4-N3	4.93	119.52	115.31
1	A	2	TLN	N3-C2-N1	4.93	121.44	114.89

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	1	LCC	C3'-C4'-C5'-O5'
1	A	1	LCC	C6'-C4'-C5'-O5'
1	A	1	LCC	O4'-C4'-C5'-O5'
1	B	4	LCG	O4'-C4'-C5'-O5'
1	B	4	LCG	C6'-C4'-C5'-O5'

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	1	LCC	2	0

5.5 Carbohydrates [i](#)

There are no monosaccharides 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	G3A	B	101	-	44,55,55	2.33	11 (25%)	48,86,86	1.68	11 (22%)
2	G3A	A	101	-	44,55,55	2.37	13 (29%)	48,86,86	2.00	10 (20%)

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	G3A	B	101	-	-	6/24/64/64	0/6/6/6
2	G3A	A	101	-	-	4/24/64/64	0/6/6/6

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	101	G3A	C35-C36	-7.56	1.32	1.47
2	A	101	G3A	C35-C36	-7.10	1.33	1.47
2	B	101	G3A	PA-O1A	-6.09	1.26	1.55
2	A	101	G3A	PA-O2A	-5.89	1.30	1.50
2	B	101	G3A	O44-C41	5.88	1.49	1.41

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	101	G3A	PG-O3A-PB	-5.92	112.51	132.83
2	A	101	G3A	N13-C12-N11	-5.80	119.61	128.68
2	B	101	G3A	N13-C12-N11	-5.55	120.00	128.68
2	A	101	G3A	C32-N31-C36	-4.19	117.38	125.10
2	A	101	G3A	C35-C36-N31	3.90	120.83	113.95

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

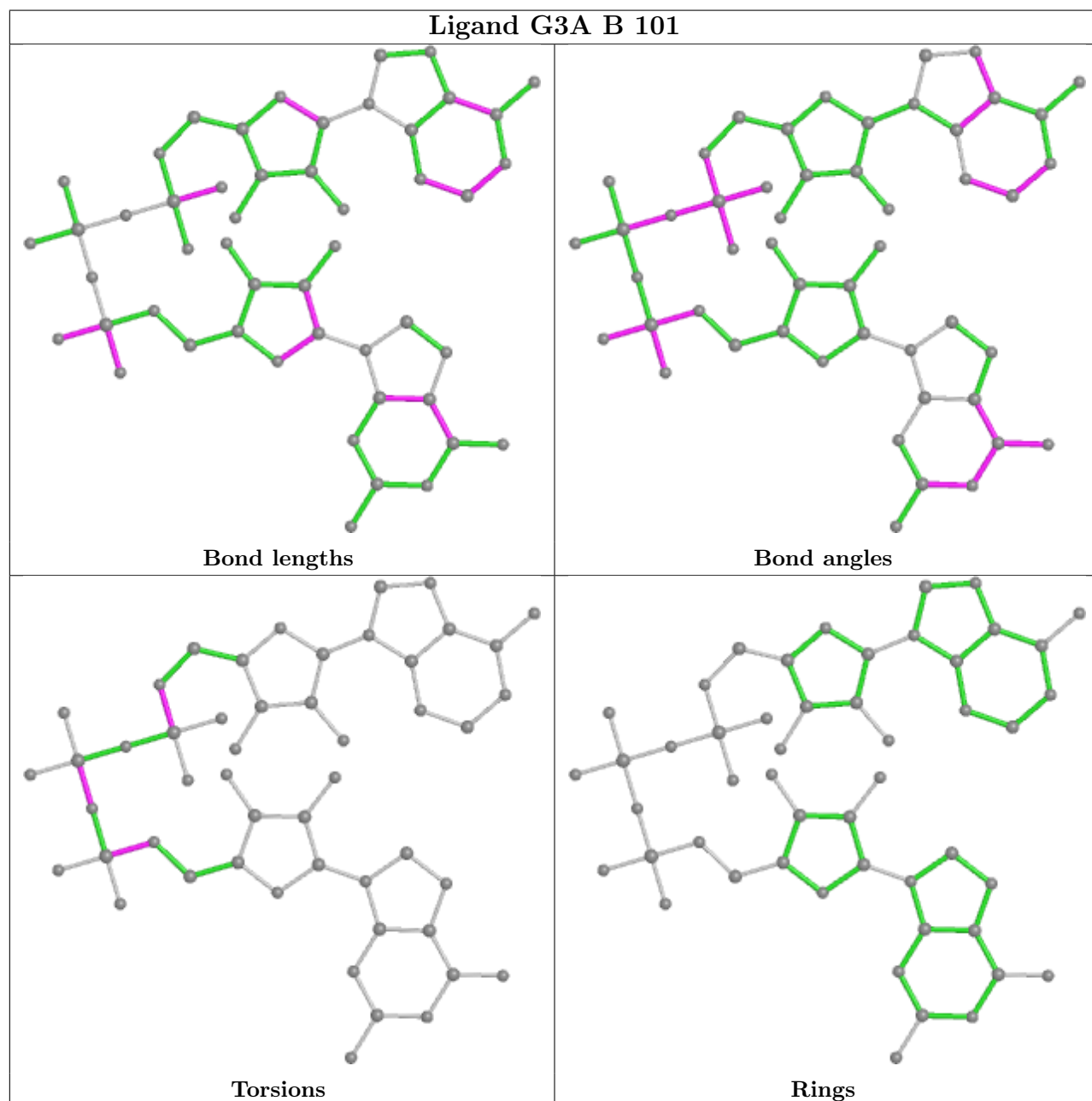
Mol	Chain	Res	Type	Atoms
2	A	101	G3A	O44-C44-C45-O45
2	B	101	G3A	C45-O45-PA-O2A
2	B	101	G3A	C25-O25-PG-O1G
2	A	101	G3A	C43-C44-C45-O45
2	B	101	G3A	PA-O1-PB-O3B

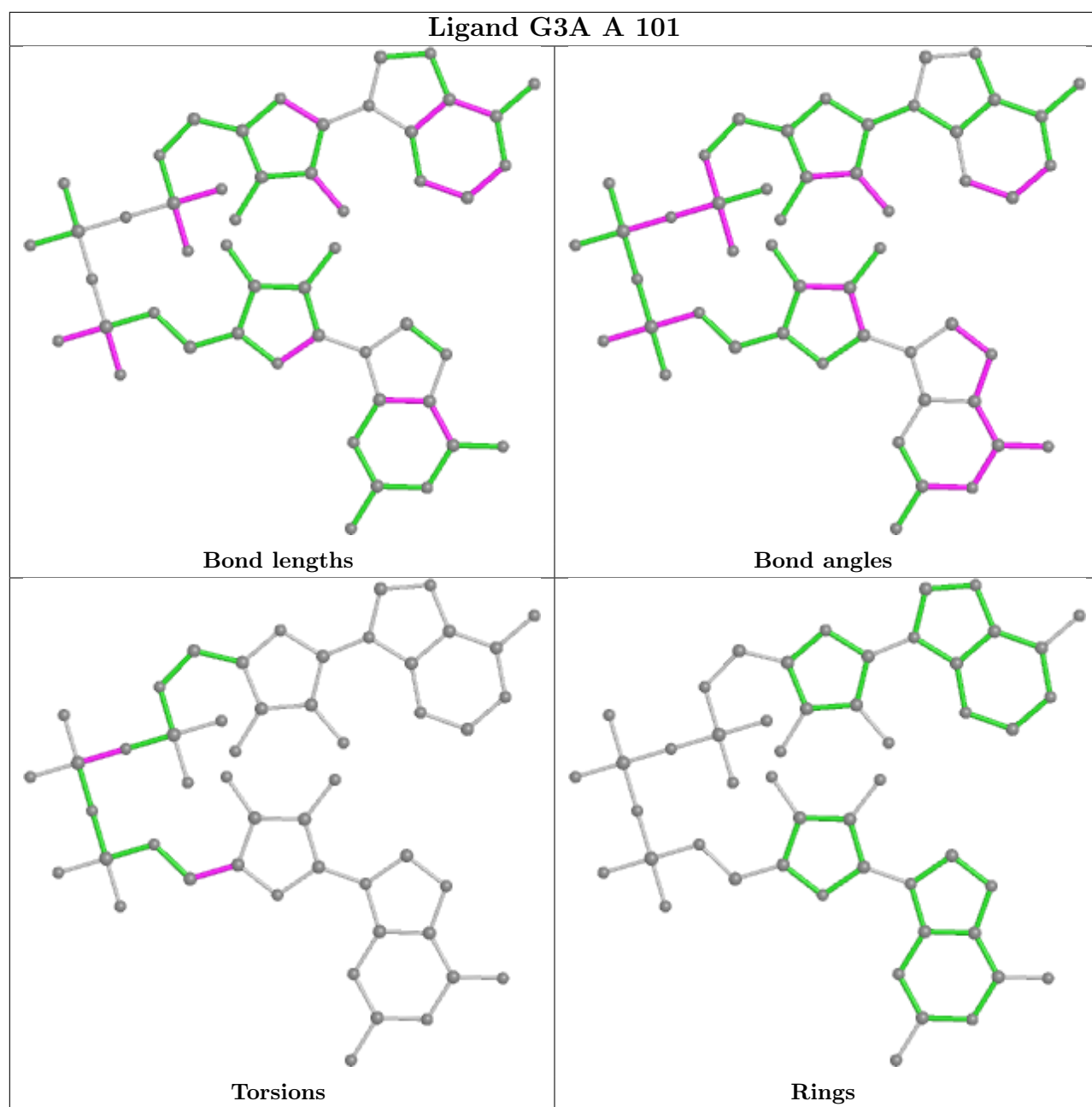
There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	101	G3A	3	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 [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

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.

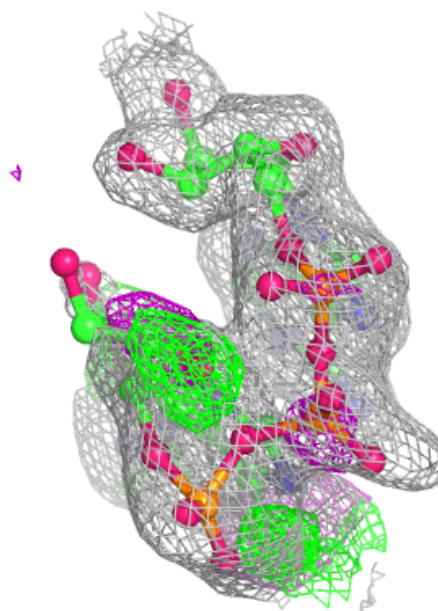
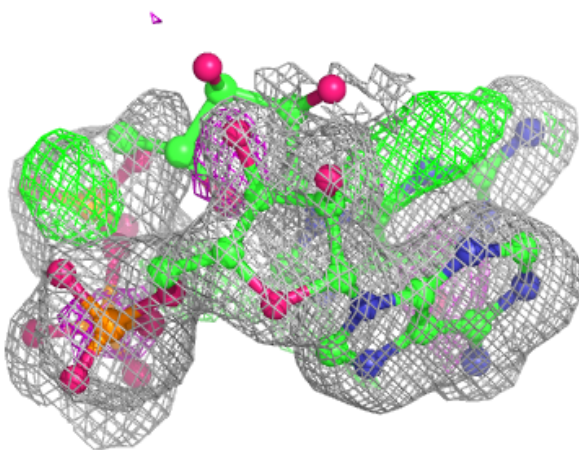
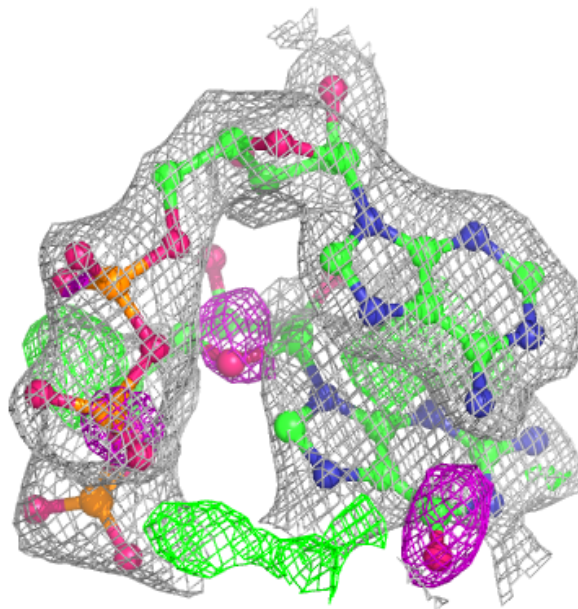
6.4 Ligands

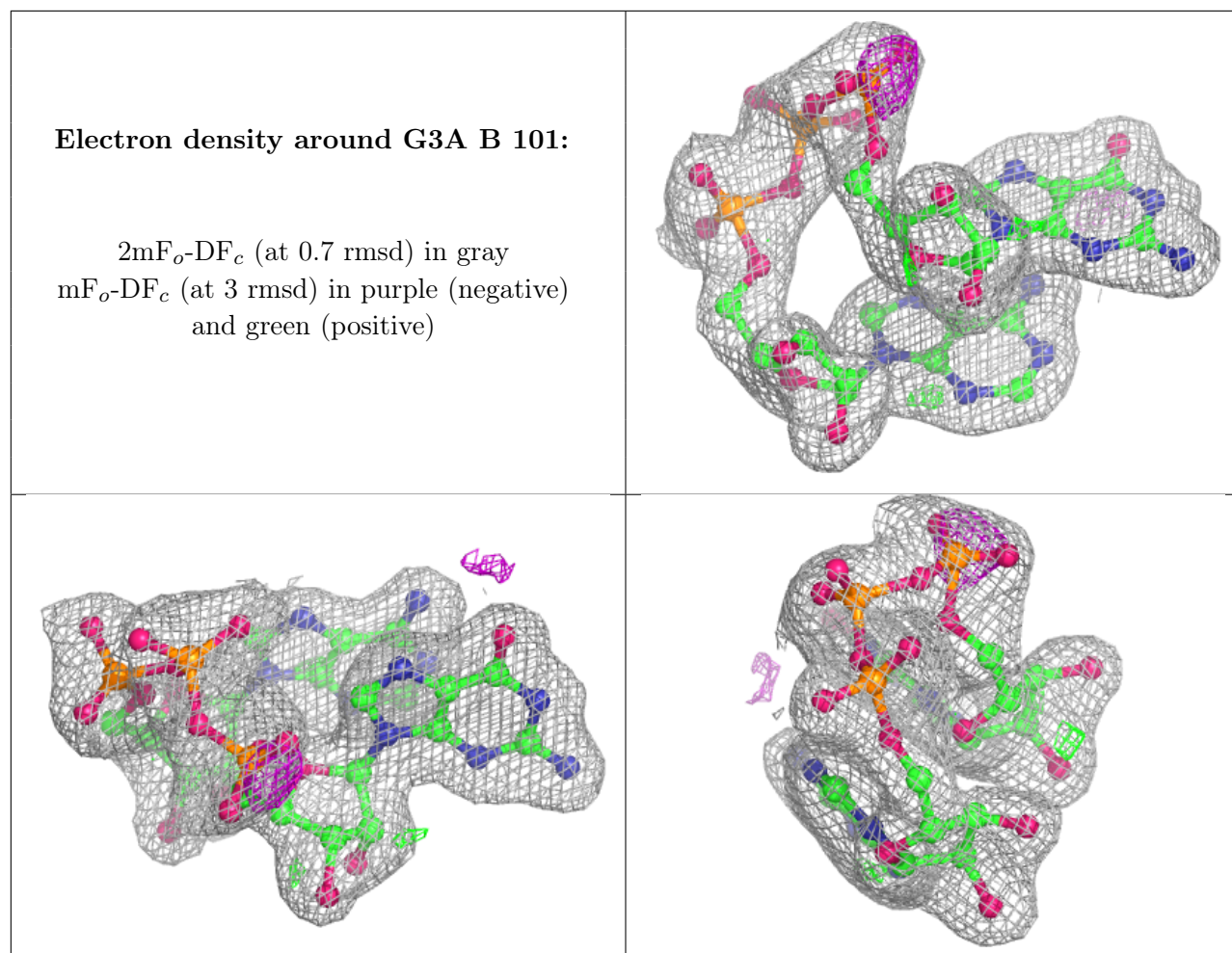
Unable to reproduce the depositors R factor - this section is therefore empty.

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 G3A A 101:

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





6.5 Other polymers [i](#)

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