Full wwPDB X-ray Structure Validation Report (i)

Jan 2, 2024 – 02:17 pm GMT

PDB ID	:	5FMM
Title	:	crystal structure of the mid, cap-binding, mid-link and 627 domains from avian influenza a virus polymerase PB2 subunit bound to M7GTP
Authors	:	Thierry, E.; Pflug, A.; Hart, D.; Cusack, S.
Deposited on	:	2015-11-06
Resolution	:	2.40 Å(reported)

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

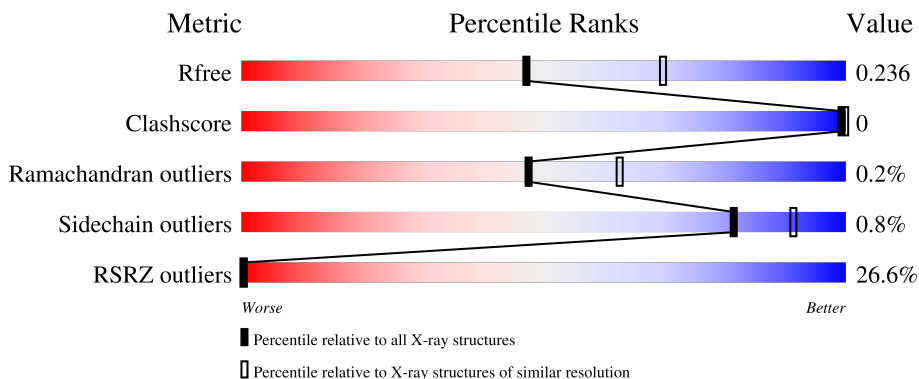|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 2.40 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		0.236
Clashscore		0
Ramachandran outliers		0.2%
Sidechain outliers		0.8%
RSRZ outliers		26.6%

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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	490	23% ... 86% ... · ... 13%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 3559 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 INFLUENZA A PB2 SUBUNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	425	Total	C	N	O	S	0	4	0
			3388	2126	611	627	24			

- Molecule 2 is 7N-METHYL-8-HYDROGUANOSINE-5'-TRIPHOSPHATE (three-letter code: MGT) (formula: $C_{11}H_{20}N_5O_{14}P_3$).
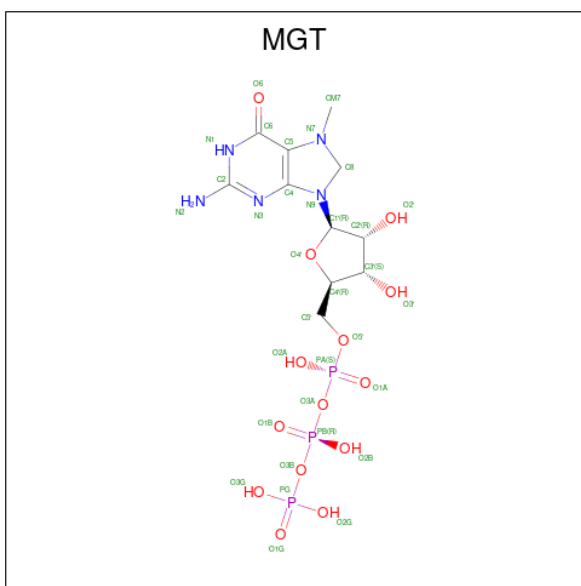
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			33	11	5	14	3		

- Molecule 3 is CITRATE ANION (three-letter code: FLC) (formula: $C_6H_5O_7$).

FLC
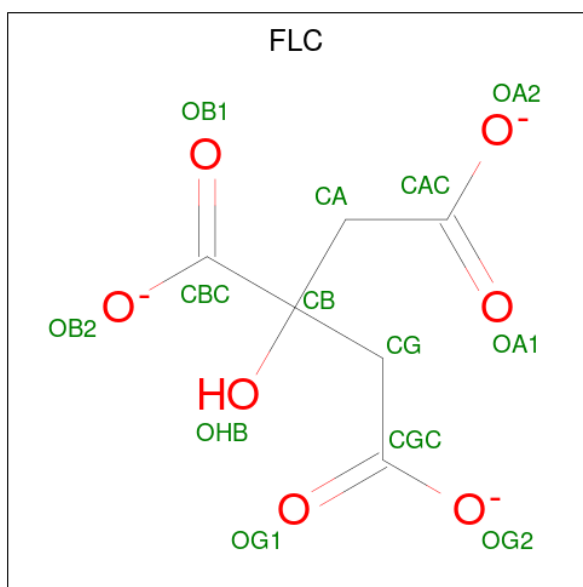
| Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 13	C 6	O 7	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	125	Total 125	O 125	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: INFLUENZA A PB2 SUBUNIT

Chain A:
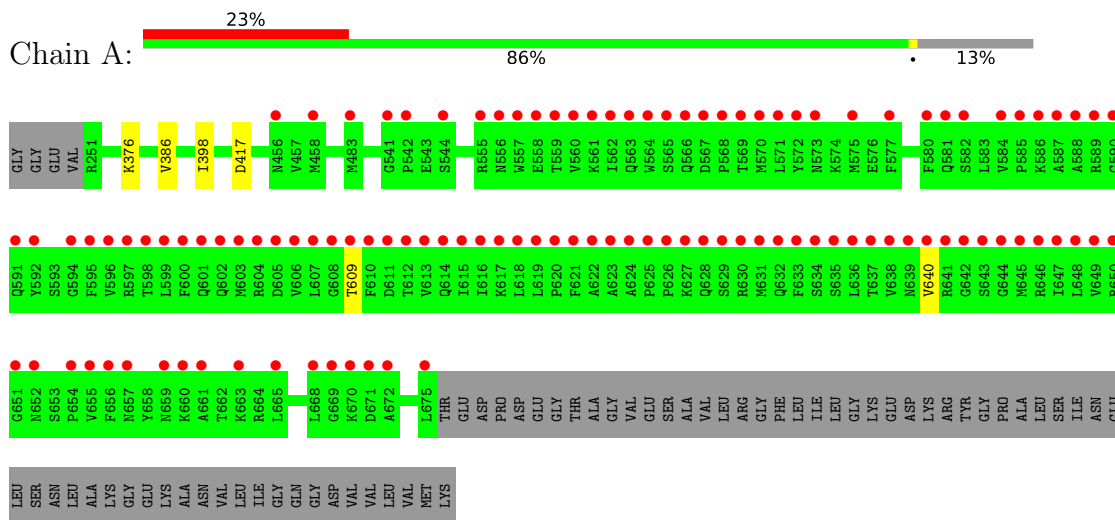
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	97.27Å 97.27Å 126.22Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	84.24 – 2.40 48.64 – 2.30	Depositor EDS
% Data completeness (in resolution range)	100.0 (84.24-2.40) 100.0 (48.64-2.30)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.70 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.8.0131	Depositor
R, R_{free}	0.212 , 0.234 0.217 , 0.236	Depositor DCC
R_{free} test set	1286 reflections (4.12%)	wwPDB-VP
Wilson B-factor (Å^2)	43.9	Xtriage
Anisotropy	0.358	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 41.2	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.026 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3559	wwPDB-VP
Average B, all atoms (Å^2)	76.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.83% 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: FLC, MGT

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	0.37	0/3441	0.61	0/4640

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	3388	0	3448	1	0
2	A	33	0	16	0	0
3	A	13	0	5	0	0
4	A	125	0	0	0	0
All	All	3559	0	3469	1	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 0.

All (1) 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:386:VAL:HG11	1:A:398:ILE:HD11	1.84	0.59

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	427/490 (87%)	414 (97%)	12 (3%)	1 (0%)	47	62

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	640	VAL

5.3.2 Protein sidechains ⓘ

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	376/423 (89%)	373 (99%)	3 (1%)	81	91

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	376	LYS
1	A	417	ASP
1	A	609	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3   RNA ⓘ

There are no RNA molecules in this entry.

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

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 ⓘ

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	MGT	A	1676	-	30,35,35	1.30	3 (10%)	42,56,56	2.14	12 (28%)
3	FLC	A	1677	-	12,12,12	0.95	0	17,17,17	1.48	3 (17%)

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	MGT	A	1676	-	-	3/22/50/50	0/3/3/3
3	FLC	A	1677	-	-	3/16/16/16	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1676	MGT	C4-N9	-4.60	1.32	1.37
2	A	1676	MGT	C5-C4	2.85	1.47	1.38
2	A	1676	MGT	C6-N1	-2.19	1.34	1.38

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1676	MGT	N9-C4-N3	7.26	136.32	125.47
2	A	1676	MGT	N9-C8-N7	-5.02	96.20	103.38
2	A	1676	MGT	C5-C4-N3	-4.42	119.71	128.13
2	A	1676	MGT	C2-N3-C4	4.00	119.43	112.30
3	A	1677	FLC	OB2-CBC-CB	3.70	119.47	113.05
2	A	1676	MGT	O4'-C1'-N9	3.40	113.94	109.30
2	A	1676	MGT	C5-C6-N1	3.09	116.43	110.99
2	A	1676	MGT	PA-O3A-PB	-2.55	124.09	132.83
2	A	1676	MGT	O6-C6-C5	-2.41	121.63	127.54
3	A	1677	FLC	OB1-CBC-CB	-2.29	119.01	122.25
2	A	1676	MGT	N2-C2-N3	-2.21	115.44	119.73
2	A	1676	MGT	N2-C2-N1	2.10	121.18	116.71
2	A	1676	MGT	C6-C5-C4	-2.09	118.32	122.62
3	A	1677	FLC	OA1-CAC-CA	-2.07	116.89	122.94
2	A	1676	MGT	CM7-N7-C5	2.06	131.72	126.40

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1676	MGT	PB-O3B-PG-O1G
3	A	1677	FLC	CG-CB-CBC-OB1
3	A	1677	FLC	CB-CA-CAC-OA2
2	A	1676	MGT	PB-O3B-PG-O2G
2	A	1676	MGT	PB-O3B-PG-O3G
3	A	1677	FLC	CB-CA-CAC-OA1

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 MGT A 1676

Bond lengths

Bond angles

Torsions

Rings
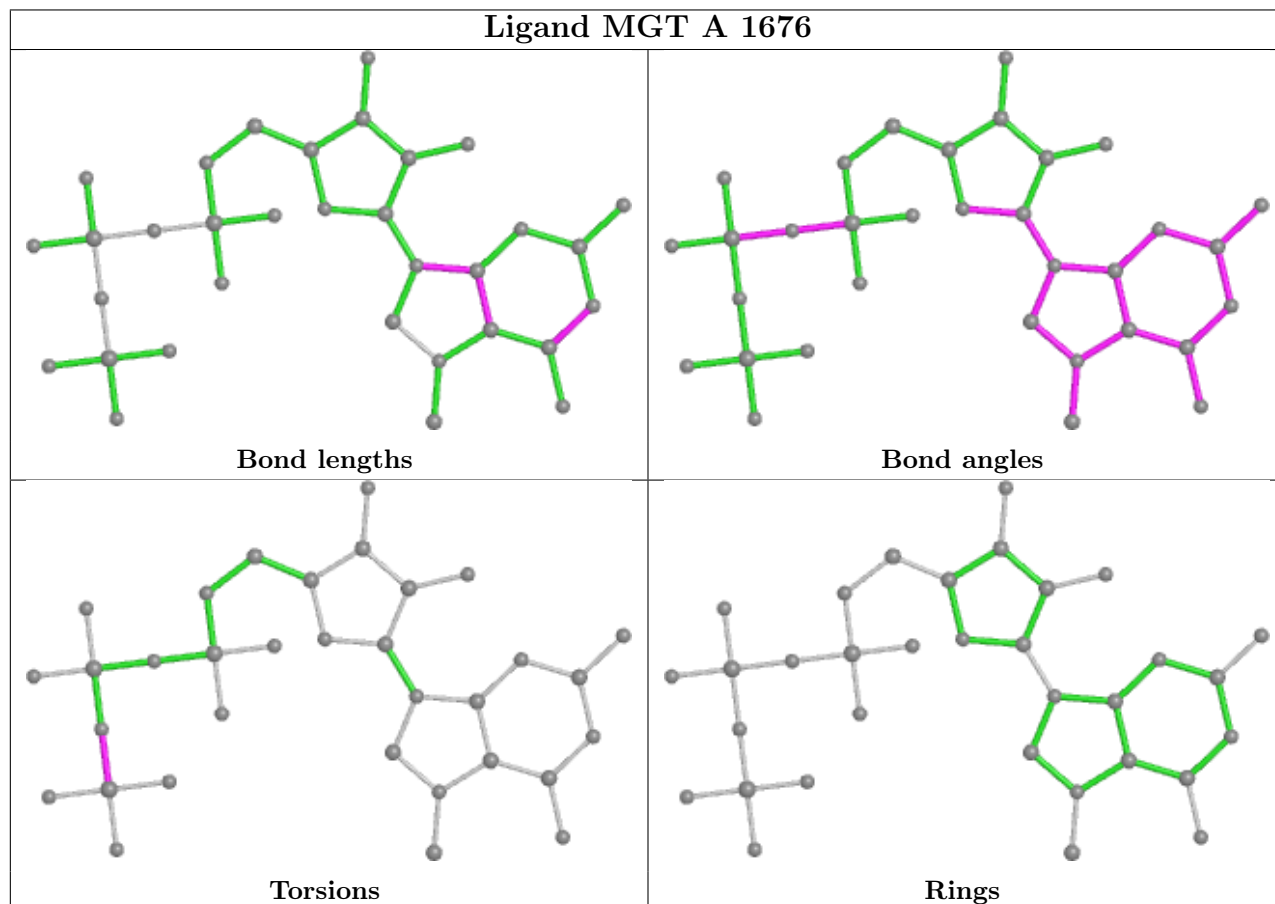
5.7 Other polymers ⓘ

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 ⓘ

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	425/490 (86%)	1.42	113 (26%)	0	0	31, 51, 167, 188	0

All (113) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	565	SER	13.2
1	A	628	GLN	13.0
1	A	632	GLN	11.9
1	A	629	SER	11.0
1	A	640	VAL	10.9
1	A	669	GLY	9.9
1	A	627	LYS	9.5
1	A	607	LEU	9.1
1	A	654	PRO	8.5
1	A	643	SER	8.0
1	A	650	ARG	7.8
1	A	610	PHE	7.6
1	A	601	GLN	7.5
1	A	623	ALA	7.1
1	A	563	GLN	7.1
1	A	598	THR	7.1
1	A	618	LEU	6.9
1	A	631	MET	6.8
1	A	648	LEU	6.8
1	A	635	SER	6.7
1	A	564	TRP	6.6
1	A	637	THR	6.5
1	A	625	PRO	6.5
1	A	605	ASP	6.5
1	A	655	VAL	6.5
1	A	634	SER	6.4
1	A	630	ARG	6.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	646	ARG	6.2
1	A	582	SER	6.1
1	A	647	ILE	6.0
1	A	568	PRO	6.0
1	A	644	GLY	5.9
1	A	594	GLY	5.8
1	A	542	PRO	5.7
1	A	585	PRO	5.7
1	A	626	PRO	5.7
1	A	600	PHE	5.7
1	A	638	VAL	5.5
1	A	604	ARG	5.1
1	A	562	ILE	5.1
1	A	596	VAL	5.0
1	A	572	TYR	5.0
1	A	566	GLN	4.9
1	A	584	VAL	4.9
1	A	645	MET	4.7
1	A	587	ALA	4.6
1	A	569	THR	4.6
1	A	458	MET	4.6
1	A	591	GLN	4.6
1	A	649	VAL	4.6
1	A	659	ASN	4.5
1	A	573	ASN	4.5
1	A	580	PHE	4.5
1	A	633	PHE	4.5
1	A	620	PRO	4.4
1	A	557	TRP	4.3
1	A	663	LYS	4.2
1	A	561	LYS	4.2
1	A	592	TYR	4.2
1	A	558	GLU	4.2
1	A	541	GLY	4.1
1	A	660	LYS	4.0
1	A	622	ALA	3.9
1	A	606	VAL	3.9
1	A	639	ASN	3.9
1	A	575	MET	3.9
1	A	595	PHE	3.9
1	A	642	GLY	3.8
1	A	641	ARG	3.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	611	ASP	3.7
1	A	675	LEU	3.7
1	A	636	LEU	3.5
1	A	614	GLN	3.5
1	A	624	ALA	3.5
1	A	570	MET	3.5
1	A	599	LEU	3.4
1	A	577	PHE	3.3
1	A	651	GLY	3.3
1	A	567	ASP	3.2
1	A	603	MET	3.2
1	A	616	ILE	3.2
1	A	597	ARG	3.2
1	A	661	ALA	3.1
1	A	617	LYS	3.1
1	A	613	VAL	3.1
1	A	670	LYS	3.1
1	A	615	ILE	3.0
1	A	621	PHE	2.9
1	A	560	VAL	2.8
1	A	589	ARG	2.8
1	A	456	ASN	2.8
1	A	586	LYS	2.8
1	A	657	ASN	2.8
1	A	559	THR	2.7
1	A	588	ALA	2.6
1	A	609	THR	2.6
1	A	571	LEU	2.5
1	A	652	ASN	2.5
1	A	656	PHE	2.5
1	A	672	ALA	2.5
1	A	665	LEU	2.5
1	A	612	THR	2.4
1	A	602	GLN	2.4
1	A	556	ASN	2.4
1	A	581	GLN	2.4
1	A	668	LEU	2.4
1	A	555	ARG	2.3
1	A	671	ASP	2.3
1	A	590	GLY	2.3
1	A	544	SER	2.2
1	A	619	LEU	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	483	MET	2.2
1	A	608	GLY	2.1

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

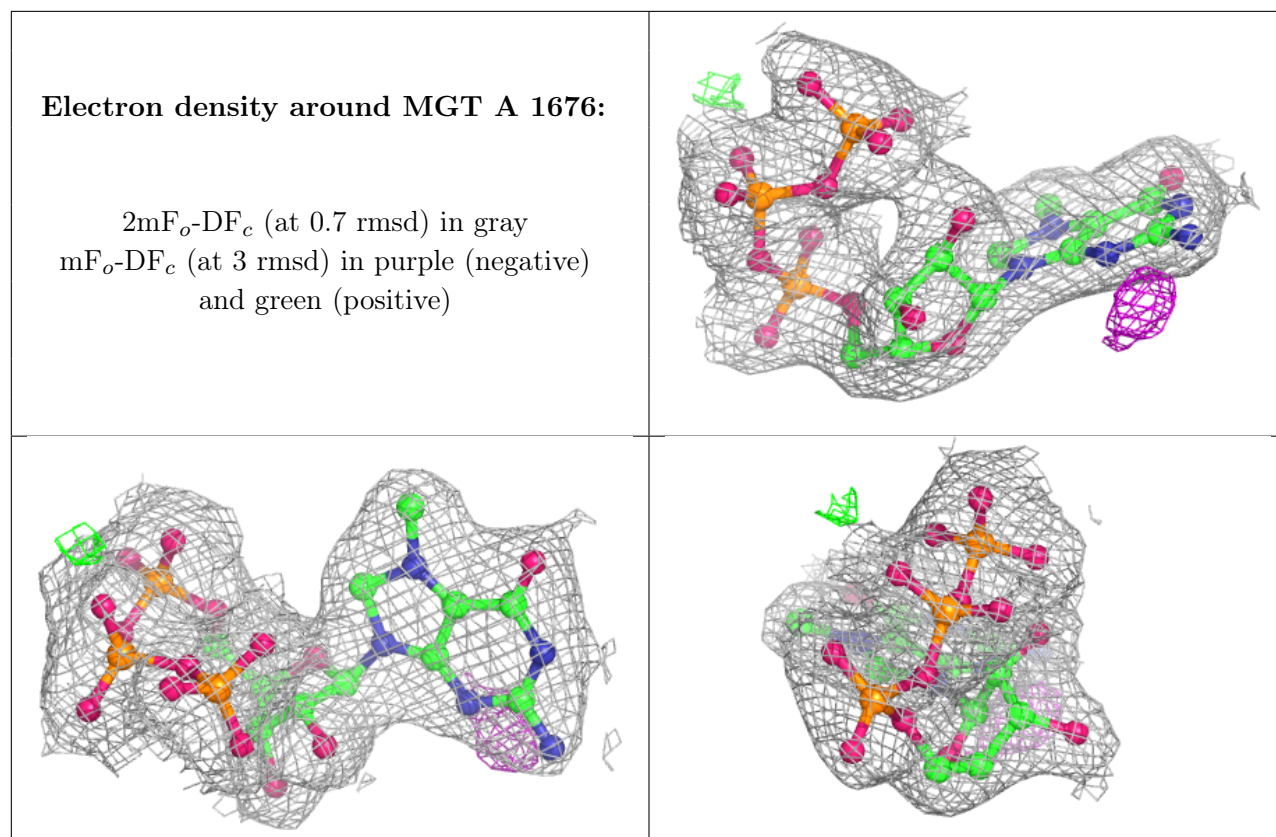There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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
3	FLC	A	1677	13/13	0.77	0.17	71,81,83,85	0
2	MGT	A	1676	33/33	0.97	0.14	42,45,48,49	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 MGT A 1676:

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

6.5 Other polymers ⓘ

There are no such residues in this entry.