WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 18, 2024 – 09:48 PM EDT

PDB ID	:	4CFX
Title	:	Structure-based design of C8-substituted O6-cyclohexylmethoxyguanine CDK1 and 2 inhibitors.
Authors	:	Carbain, B.; Paterson, D.J.; Anscombe, E.; Campbell, A.; Cano, C.; Echalier, A.; Endicott, J.; Golding, B.T.; Haggerty, K.; Hardcastle, I.R.; Jewsbury, P.; Newell, D.R.; Noble, M.E.M.; Roche, C.; Wang, L.Z.; Griffin, R.
Deposited on	:	2013-11-19
Resolution	:	3.50 Å(reported)

The following versions of software and data (see references ①) were used in the production of this report:

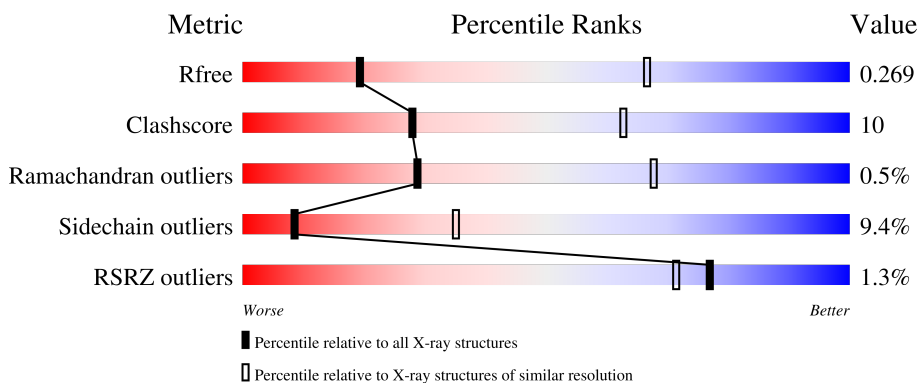|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.50 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		0.269
Clashscore		10
Ramachandran outliers		0.5%
Sidechain outliers		9.4%
RSRZ outliers		1.3%

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	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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	303	74% ... 20%
1	C	303	73% ... 21% ... 5%
2	B	260	73% ... 23%
2	D	260	77% ... 22%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 9356 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 CYCLIN-DEPENDENT KINASE 2.

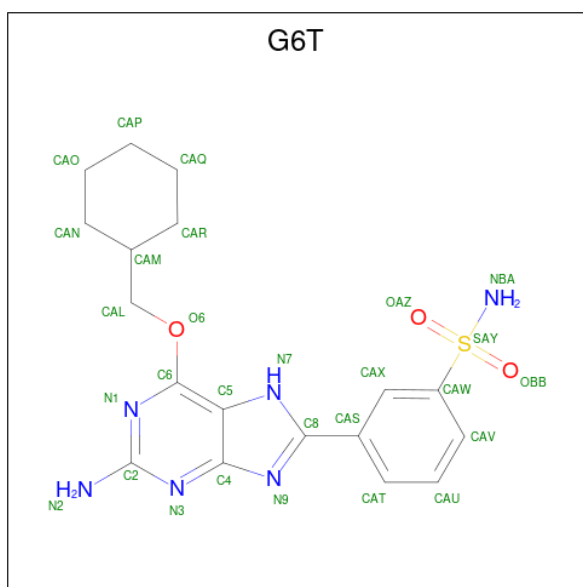| Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	299	Total 2460	C 1591	N 418	O 441	P 1	S 9	0	7	0
1	C	299	Total 2454	C 1588	N 417	O 439	P 1	S 9	0	7	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	GLY	-	expression tag	UNP P24941
A	-3	PRO	-	expression tag	UNP P24941
A	-2	LEU	-	expression tag	UNP P24941
A	-1	GLY	-	expression tag	UNP P24941
A	0	SER	-	expression tag	UNP P24941
C	-4	GLY	-	expression tag	UNP P24941
C	-3	PRO	-	expression tag	UNP P24941
C	-2	LEU	-	expression tag	UNP P24941
C	-1	GLY	-	expression tag	UNP P24941
C	0	SER	-	expression tag	UNP P24941

- Molecule 2 is a protein called CYCLIN-A2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	260	Total 2143	C 1384	N 349	O 396	S 14	0	6	0
2	D	260	Total 2155	C 1390	N 351	O 400	S 14	0	8	0

- Molecule 3 is 3-[2-amino-6-(cyclohexylmethoxy)-7H-purin-8-yl]benzenesulfonamide (three-letter code: G6T) (formula: $C_{18}H_{22}N_6O_3S$).

G6T

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			28	18	6	3	1		
3	C	1	Total	C	N	O	S	0	0
			28	18	6	3	1		

• Molecule 4 is water.

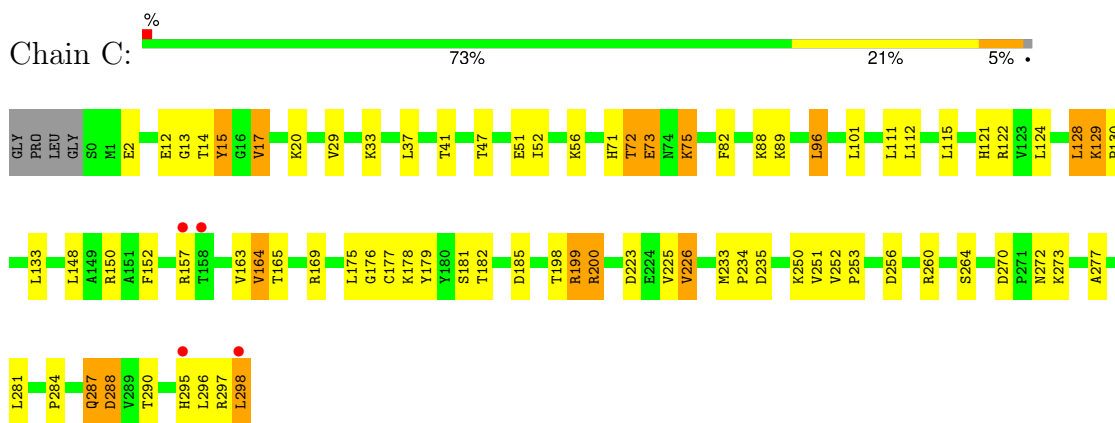| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	45	Total	O	0	0
			45	45		
4	B	16	Total	O	0	0
			16	16		
4	C	17	Total	O	0	0
			17	17		
4	D	10	Total	O	0	0
			10	10		

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: CYCLIN-DEPENDENT KINASE 2
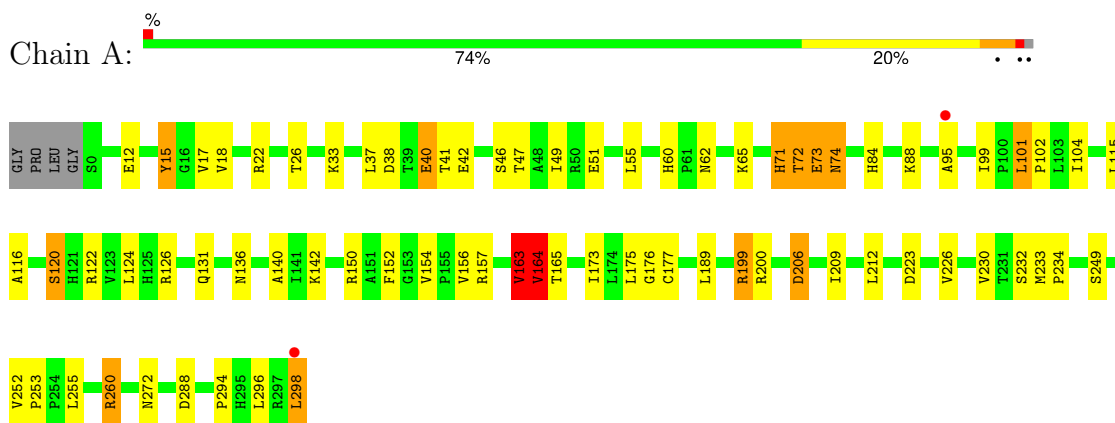
• Molecule 1: CYCLIN-DEPENDENT KINASE 2

• Molecule 2: CYCLIN-A2
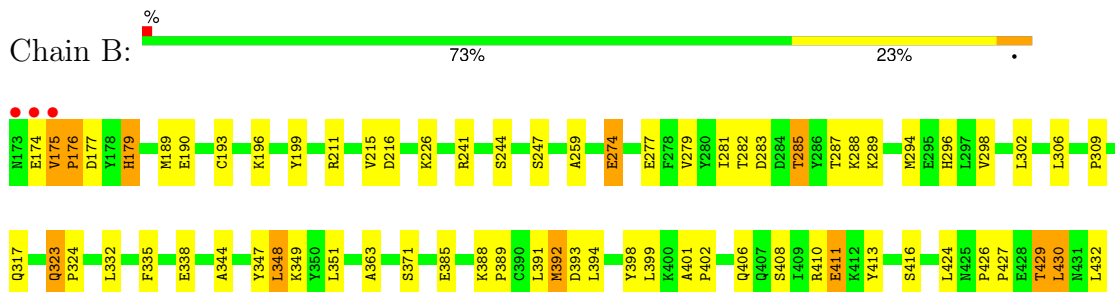
• Molecule 2: CYCLIN-A2

Chain D:

2%

77%

22%

N173 E174 V175 P176 D177 Y178 H179 E180 D181 C193 Y199 S209 M210 R211 V215 D216 V219 E220 E223 E224 N229 L232 V248 L263 A264 S265 E274 E277 Y280 I281 T282 D283 Q290 M294 E295 H296 L297 V298 P309 N312 T316

N326 C327 E330 F335 L336 G337 E338 L339 I342 D345 P346 Y347 L362 P373 E374 S375 R378 L394 H395 Q396 T397 Y398 L399 K400 A401 P402 Q403 H404 A405 Q406 Q407 E411 E428 T429 L430 N431 L432

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	74.66Å 136.56Å 149.03Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.34 – 3.50 49.20 – 3.50	Depositor EDS
% Data completeness (in resolution range)	98.0 (50.34-3.50) 98.0 (49.20-3.50)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.30 (at 3.48Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.204 , 0.269 0.200 , 0.269	Depositor DCC
R_{free} test set	976 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	53.2	Xtriage
Anisotropy	0.751	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 38.9	EDS
L-test for twinning[2]	$< |L| > = 0.44, < L^2 > = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	9356	wwPDB-VP
Average B, all atoms (Å^2)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.09% 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: TPO, G6T

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	0.58	0/2510	0.81	4/3405 (0.1%)
1	C	0.54	0/2504	0.76	0/3397
2	B	0.56	0/2193	0.75	0/2977
2	D	0.53	0/2205	0.73	0/2993
All	All	0.55	0/9412	0.76	4/12772 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	298	LEU	CA-CB-CG	6.05	129.21	115.30
1	A	164	VAL	N-CA-C	-6.01	94.77	111.00
1	A	163	VAL	N-CA-C	-5.80	95.33	111.00
1	A	163	VAL	C-N-CA	5.04	134.31	121.70

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	2460	0	2497	50	0
1	C	2454	0	2492	58	0
2	B	2143	0	2156	45	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	2155	0	2164	38	0
3	A	28	0	22	1	0
3	C	28	0	22	0	0
4	A	45	0	0	3	0
4	B	16	0	0	0	0
4	C	17	0	0	0	0
4	D	10	0	0	0	0
All	All	9356	0	9353	184	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 10.

The worst 5 of 184 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:C:177[A]:CYS:SG	1:C:233:MET:SD	2.44	1.16
1:C:200:ARG:HG3	1:C:200:ARG:HH11	1.13	1.07
1:C:13:GLY:O	1:C:14:THR:HG22	1.60	1.00
1:A:95:ALA:HA	1:A:199:ARG:HD2	1.45	0.94
1:A:72:THR:CG2	1:A:73:GLU:HG2	2.03	0.88

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	304/303 (100%)	285 (94%)	17 (6%)	2 (1%)	22	61
1	C	303/303 (100%)	282 (93%)	19 (6%)	2 (1%)	22	61
2	B	264/260 (102%)	257 (97%)	5 (2%)	2 (1%)	19	58
2	D	266/260 (102%)	261 (98%)	5 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1137/1126 (101%)	1085 (95%)	46 (4%)	6 (0%)	29 68

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	176	PRO
1	C	164	VAL
1	A	15	TYR
1	C	297	ARG
2	B	175	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	271/265 (102%)	242 (89%)	29 (11%)	6 30
1	C	270/265 (102%)	239 (88%)	31 (12%)	5 26
2	B	240/234 (103%)	219 (91%)	21 (9%)	10 38
2	D	242/234 (103%)	222 (92%)	20 (8%)	11 40
All	All	1023/998 (102%)	922 (90%)	101 (10%)	8 33

5 of 101 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	72	THR
1	C	182	THR
2	D	428	GLU
1	C	75	LYS
1	C	128	LEU

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	C	287	GLN
2	D	179	HIS
2	D	396	GLN
2	D	296	HIS
1	C	161	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 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	TPO	C	160	1	8,10,11	0.77	0	10,14,16	1.07	0
1	TPO	A	160	1	8,10,11	0.68	0	10,14,16	1.19	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
1	TPO	C	160	1	-	0/9/11/13	-
1	TPO	A	160	1	-	0/9/11/13	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	$\#|Z| > 2$
3	G6T	C	1299	-	28,31,31	2.59	8 (28%)	35,45,45	2.10	9 (25%)
3	G6T	A	1299	-	28,31,31	2.69	8 (28%)	35,45,45	2.45	10 (28%)

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	G6T	C	1299	-	-	6/15/23/23	0/4/4/4
3	G6T	A	1299	-	-	4/15/23/23	0/4/4/4

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1299	G6T	SAY-NBA	-7.62	1.45	1.60
3	C	1299	G6T	SAY-NBA	-6.48	1.47	1.60
3	A	1299	G6T	CAS-C8	-5.63	1.32	1.48
3	C	1299	G6T	OAZ-SAY	5.62	1.53	1.43
3	C	1299	G6T	CAS-C8	-5.37	1.33	1.48

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1299	G6T	OBB-SAY-OAZ	-6.82	108.31	118.80
3	C	1299	G6T	OBB-SAY-OAZ	-5.66	110.09	118.80
3	A	1299	G6T	N3-C2-N1	-5.62	120.06	127.21
3	A	1299	G6T	O6-C6-N1	5.34	124.79	120.23
3	C	1299	G6T	N3-C2-N1	-5.12	120.70	127.21

There are no chirality outliers.
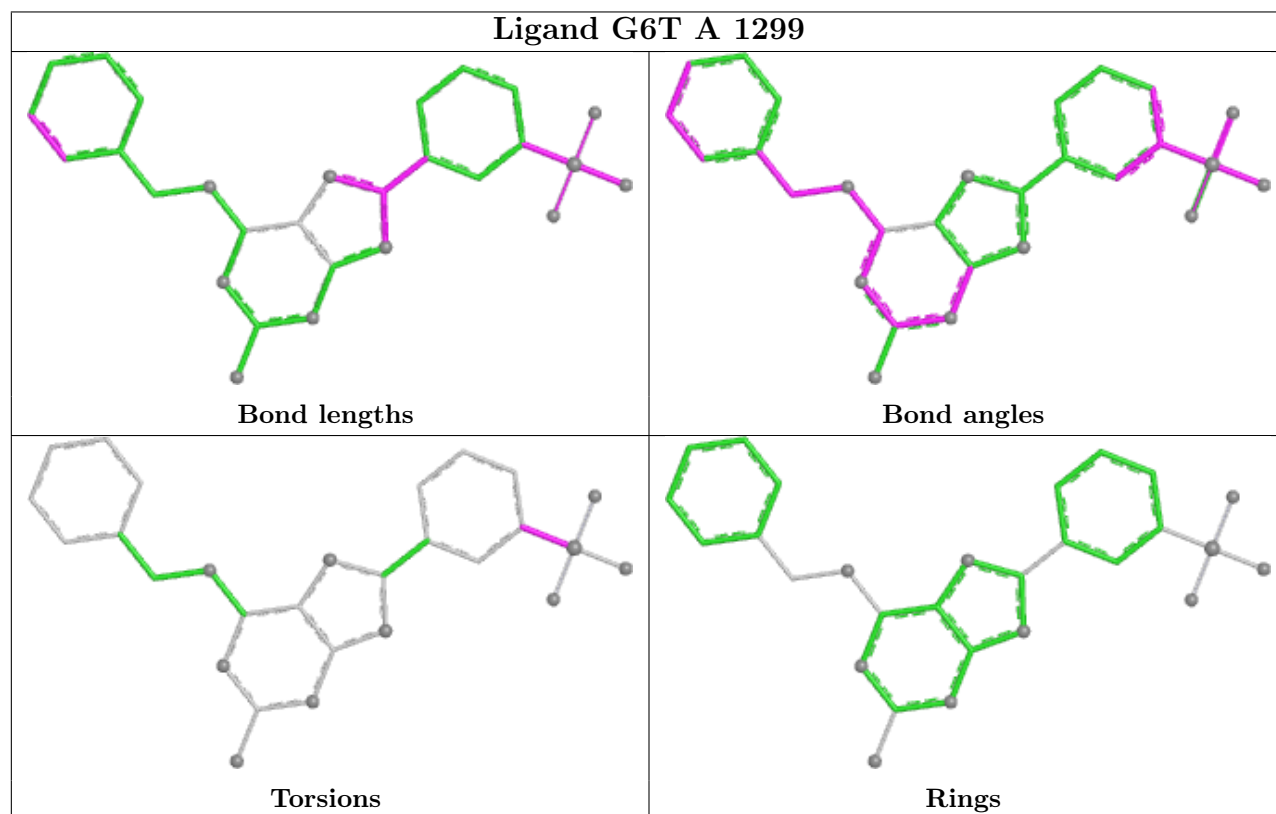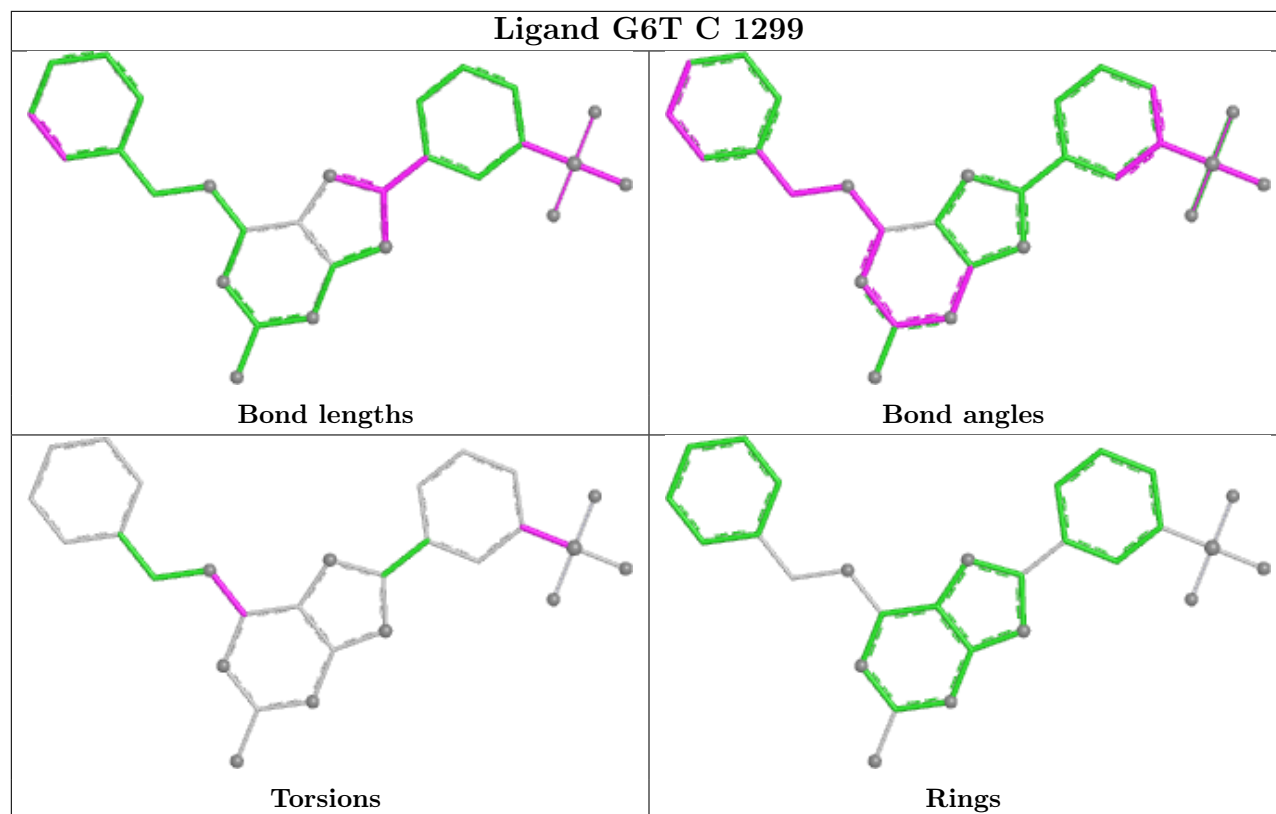
5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	1299	G6T	N1-C6-O6-CAL
3	C	1299	G6T	C5-C6-O6-CAL
3	A	1299	G6T	CAV-CAW-SAY-NBA
3	A	1299	G6T	CAX-CAW-SAY-NBA
3	C	1299	G6T	CAV-CAW-SAY-NBA

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1299	G6T	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 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 G6T C 1299

Bond lengths

Bond angles

Torsions

Rings

Ligand G6T A 1299

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	298/303 (98%)	-0.40	2 (0%)	87	83	26, 37, 63, 78	0
1	C	298/303 (98%)	0.06	4 (1%)	77	71	44, 59, 80, 89	0
2	B	260/260 (100%)	-0.15	3 (1%)	79	73	27, 41, 57, 67	0
2	D	260/260 (100%)	-0.04	5 (1%)	66	61	38, 59, 93, 107	0
All	All	1116/1126 (99%)	-0.13	14 (1%)	77	71	26, 49, 82, 107	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	174	GLU	5.6
1	C	298	LEU	3.3
2	B	174	GLU	3.2
2	D	177	ASP	2.8
1	C	157	ARG	2.7

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

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
1	TPO	C	160	11/12	0.93	0.31	50,53,54,54	0
1	TPO	A	160	11/12	0.98	0.27	37,39,40,41	0

6.3 Carbohydrates ⓘ

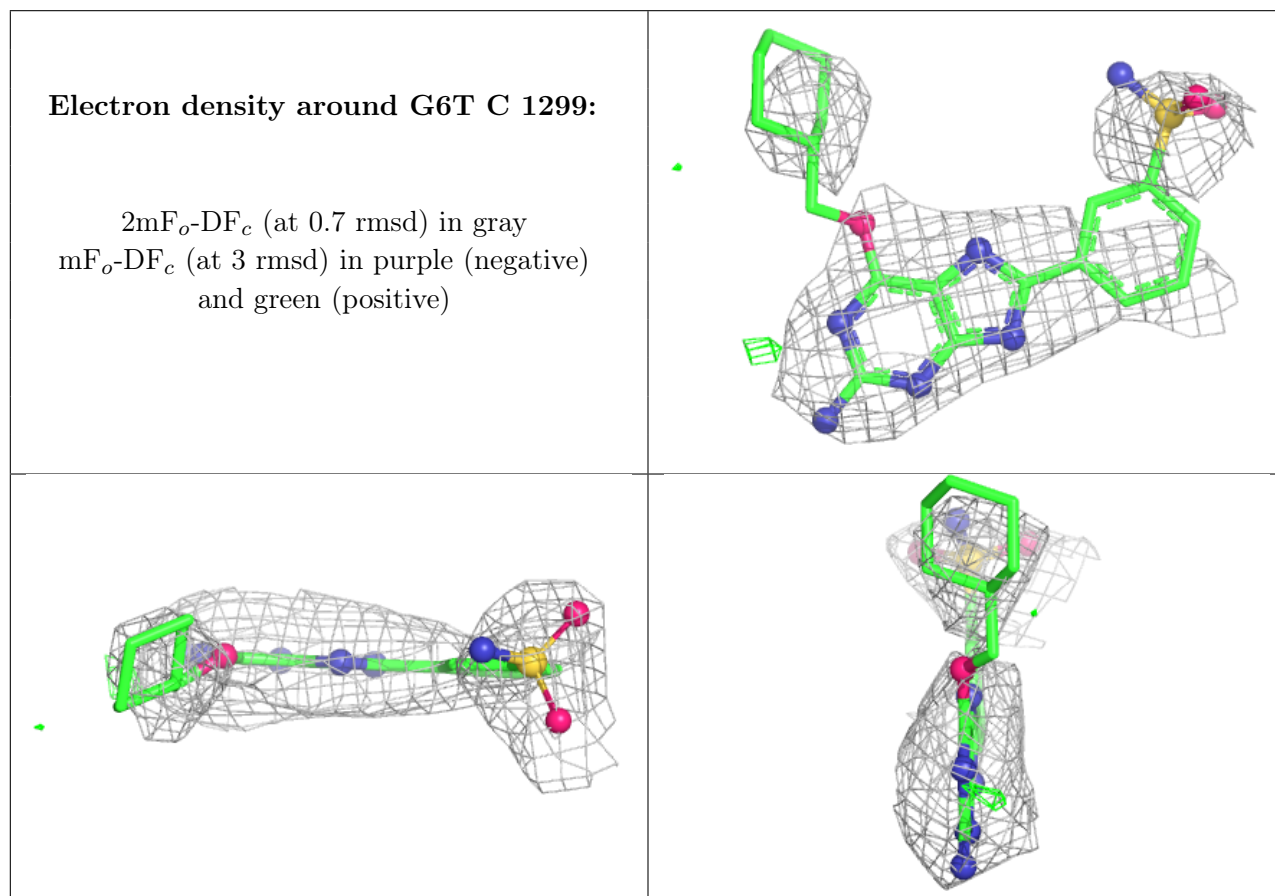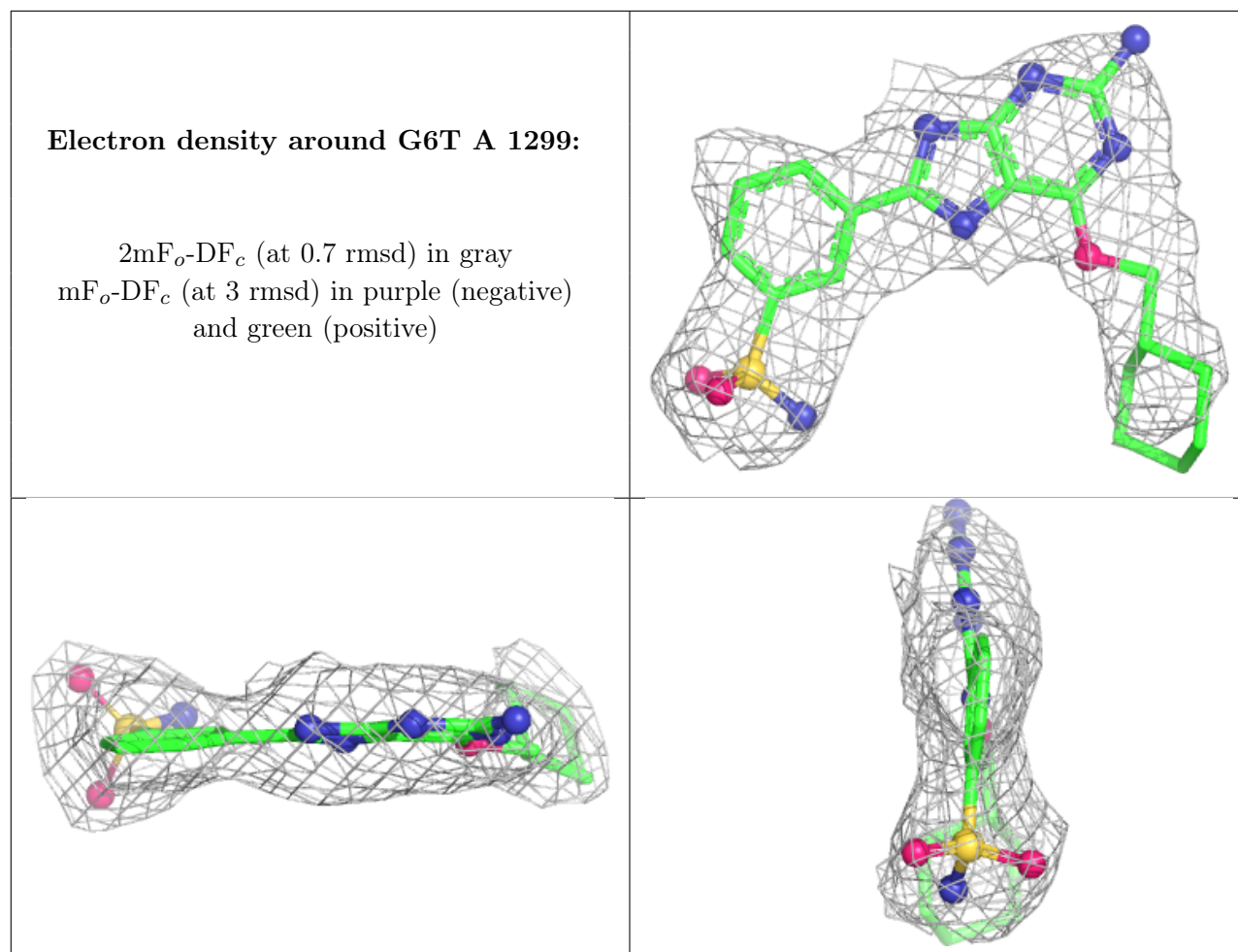There are no monosaccharides in this entry.

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	G6T	C	1299	28/28	0.87	0.37	102,123,139,140	0
3	G6T	A	1299	28/28	0.91	0.31	99,107,121,125	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 G6T C 1299:

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

Electron density around G6T A 1299:

$2mF_o$-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 ⓘ

There are no such residues in this entry.