WORLDWIDE PDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Oct 9, 2023 – 10:22 AM EDT

PDB ID	:	8E23
Title	:	Human DNA polymerase theta in complex with allosteric inhibitor
Authors	:	Mader, P.; Pau, V.P.T.; Sicheri, F.
Deposited on	:	2022-08-13
Resolution	:	2.59 Å(reported)

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

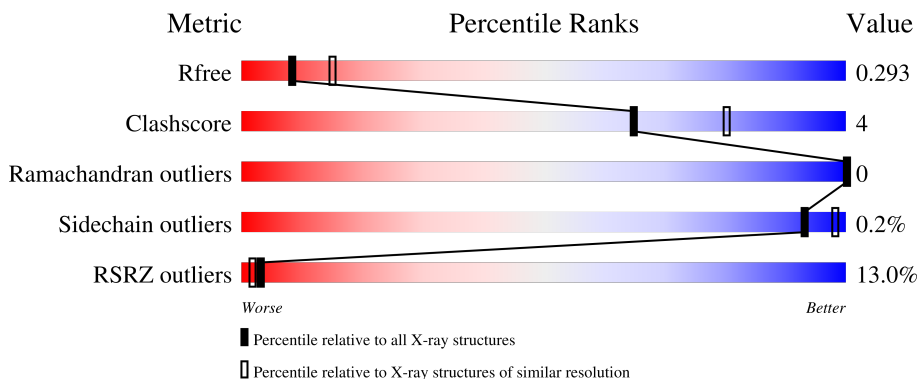|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.59 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		0.293
Clashscore		4
Ramachandran outliers		0
Sidechain outliers		0.2%
RSRZ outliers		13.0%

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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	668	5% / 88% / 7% / 5%
1	D	668	20% / 79% / 12% / 10%
2	B	17	6% / 82% / 6% / 12%
2	E	17	53% / 18% / 29%
3	C	13	8% / 62% / 23% / 15%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	13	46% 15% 38%

2 Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 10562 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 DNA polymerase theta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	637	Total	C	N	O	S	0	0	0
			4980	3171	855	927	27			
1	D	603	Total	C	N	O	S	0	0	0
			4478	2819	775	862	22			

There are 214 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1817	GLY	-	expression tag	UNP O75417
A	?	-	ALA	deletion	UNP O75417
A	?	-	THR	deletion	UNP O75417
A	?	-	ILE	deletion	UNP O75417
A	?	-	GLY	deletion	UNP O75417
A	?	-	SER	deletion	UNP O75417
A	?	-	ARG	deletion	UNP O75417
A	?	-	PHE	deletion	UNP O75417
A	?	-	LYS	deletion	UNP O75417
A	?	-	GLN	deletion	UNP O75417
A	?	-	ALA	deletion	UNP O75417
A	?	-	SER	deletion	UNP O75417
A	?	-	SER	deletion	UNP O75417
A	?	-	PRO	deletion	UNP O75417
A	?	-	GLN	deletion	UNP O75417
A	?	-	GLU	deletion	UNP O75417
A	?	-	ILE	deletion	UNP O75417
A	?	-	PRO	deletion	UNP O75417
A	?	-	ILE	deletion	UNP O75417
A	?	-	ARG	deletion	UNP O75417
A	?	-	ASP	deletion	UNP O75417
A	?	-	ASP	deletion	UNP O75417
A	?	-	GLN	deletion	UNP O75417
A	?	-	LYS	deletion	UNP O75417
A	?	-	HIS	deletion	UNP O75417

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	SER	deletion	UNP O75417
A	?	-	GLU	deletion	UNP O75417
A	?	-	ILE	deletion	UNP O75417
A	?	-	SER	deletion	UNP O75417
A	?	-	ALA	deletion	UNP O75417
A	?	-	SER	deletion	UNP O75417
A	?	-	LEU	deletion	UNP O75417
A	?	-	VAL	deletion	UNP O75417
A	?	-	PRO	deletion	UNP O75417
A	?	-	PRO	deletion	UNP O75417
A	?	-	MET	deletion	UNP O75417
A	?	-	LYS	deletion	UNP O75417
A	?	-	ASN	deletion	UNP O75417
A	?	-	GLN	deletion	UNP O75417
A	?	-	GLY	deletion	UNP O75417
A	?	-	SER	deletion	UNP O75417
A	?	-	LYS	deletion	UNP O75417
A	?	-	LYS	deletion	UNP O75417
A	?	-	THR	deletion	UNP O75417
A	?	-	LEU	deletion	UNP O75417
A	?	-	GLY	deletion	UNP O75417
A	?	-	SER	deletion	UNP O75417
A	?	-	THR	deletion	UNP O75417
A	?	-	ARG	deletion	UNP O75417
A	?	-	ARG	deletion	UNP O75417
A	?	-	GLY	deletion	UNP O75417
A	?	-	ILE	deletion	UNP O75417
A	?	-	ASP	deletion	UNP O75417
A	?	-	ASN	deletion	UNP O75417
A	?	-	GLY	deletion	UNP O75417
A	?	-	ARG	deletion	UNP O75417
A	?	-	LYS	deletion	UNP O75417
A	?	-	LEU	deletion	UNP O75417
A	?	-	ARG	deletion	UNP O75417
A	?	-	VAL	deletion	UNP O75417
A	?	-	GLY	deletion	UNP O75417
A	?	-	GLU	deletion	UNP O75417
A	?	-	SER	deletion	UNP O75417
A	?	-	PRO	deletion	UNP O75417
A	?	-	PRO	deletion	UNP O75417
A	?	-	SER	deletion	UNP O75417
A	?	-	GLN	deletion	UNP O75417

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	ALA	deletion	UNP O75417
A	?	-	VAL	deletion	UNP O75417
A	?	-	GLY	deletion	UNP O75417
A	?	-	LYS	deletion	UNP O75417
A	?	-	GLY	deletion	UNP O75417
A	?	-	LEU	deletion	UNP O75417
A	?	-	LEU	deletion	UNP O75417
A	?	-	PRO	deletion	UNP O75417
A	?	-	MET	deletion	UNP O75417
A	?	-	GLY	deletion	UNP O75417
A	?	-	ARG	deletion	UNP O75417
A	?	-	GLY	deletion	UNP O75417
A	?	-	LYS	deletion	UNP O75417
A	?	-	TYR	deletion	UNP O75417
A	?	-	LYS	deletion	UNP O75417
A	?	-	LYS	deletion	UNP O75417
A	?	-	GLY	deletion	UNP O75417
A	?	-	PHE	deletion	UNP O75417
A	?	-	SER	deletion	UNP O75417
A	?	-	VAL	deletion	UNP O75417
A	?	-	ASN	deletion	UNP O75417
A	?	-	PRO	deletion	UNP O75417
A	?	-	ARG	deletion	UNP O75417
A	?	-	CYS	deletion	UNP O75417
A	?	-	GLN	deletion	UNP O75417
A	?	-	ALA	deletion	UNP O75417
A	?	-	GLN	deletion	UNP O75417
A	?	-	MET	deletion	UNP O75417
A	?	-	GLU	deletion	UNP O75417
A	?	-	GLU	deletion	UNP O75417
A	?	-	ARG	deletion	UNP O75417
A	?	-	ALA	deletion	UNP O75417
A	?	-	LEU	deletion	UNP O75417
A	?	-	SER	deletion	UNP O75417
A	?	-	ARG	deletion	UNP O75417
A	?	-	LYS	deletion	UNP O75417
A	?	-	ARG	deletion	UNP O75417
A	?	-	LYS	deletion	UNP O75417
A	?	-	LEU	deletion	UNP O75417
A	?	-	GLN	deletion	UNP O75417
D	1817	GLY	-	expression tag	UNP O75417
D	?	-	ALA	deletion	UNP O75417

Continued on next page...

Full wwPDB X-ray Structure Validation Report

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	?	-	THR	deletion	UNP O75417
D	?	-	ILE	deletion	UNP O75417
D	?	-	GLY	deletion	UNP O75417
D	?	-	SER	deletion	UNP O75417
D	?	-	ARG	deletion	UNP O75417
D	?	-	PHE	deletion	UNP O75417
D	?	-	LYS	deletion	UNP O75417
D	?	-	GLN	deletion	UNP O75417
D	?	-	ALA	deletion	UNP O75417
D	?	-	SER	deletion	UNP O75417
D	?	-	SER	deletion	UNP O75417
D	?	-	PRO	deletion	UNP O75417
D	?	-	GLN	deletion	UNP O75417
D	?	-	GLU	deletion	UNP O75417
D	?	-	ILE	deletion	UNP O75417
D	?	-	PRO	deletion	UNP O75417
D	?	-	ILE	deletion	UNP O75417
D	?	-	ARG	deletion	UNP O75417
D	?	-	ASP	deletion	UNP O75417
D	?	-	ASP	deletion	UNP O75417
D	?	-	GLN	deletion	UNP O75417
D	?	-	LYS	deletion	UNP O75417
D	?	-	HIS	deletion	UNP O75417
D	?	-	SER	deletion	UNP O75417
D	?	-	GLU	deletion	UNP O75417
D	?	-	ILE	deletion	UNP O75417
D	?	-	SER	deletion	UNP O75417
D	?	-	ALA	deletion	UNP O75417
D	?	-	SER	deletion	UNP O75417
D	?	-	LEU	deletion	UNP O75417
D	?	-	VAL	deletion	UNP O75417
D	?	-	PRO	deletion	UNP O75417
D	?	-	PRO	deletion	UNP O75417
D	?	-	MET	deletion	UNP O75417
D	?	-	LYS	deletion	UNP O75417
D	?	-	ASN	deletion	UNP O75417
D	?	-	GLN	deletion	UNP O75417
D	?	-	GLY	deletion	UNP O75417
D	?	-	SER	deletion	UNP O75417
D	?	-	LYS	deletion	UNP O75417
D	?	-	LYS	deletion	UNP O75417
D	?	-	THR	deletion	UNP O75417

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	?	-	LEU	deletion	UNP O75417
D	?	-	GLY	deletion	UNP O75417
D	?	-	SER	deletion	UNP O75417
D	?	-	THR	deletion	UNP O75417
D	?	-	ARG	deletion	UNP O75417
D	?	-	ARG	deletion	UNP O75417
D	?	-	GLY	deletion	UNP O75417
D	?	-	ILE	deletion	UNP O75417
D	?	-	ASP	deletion	UNP O75417
D	?	-	ASN	deletion	UNP O75417
D	?	-	GLY	deletion	UNP O75417
D	?	-	ARG	deletion	UNP O75417
D	?	-	LYS	deletion	UNP O75417
D	?	-	LEU	deletion	UNP O75417
D	?	-	ARG	deletion	UNP O75417
D	?	-	VAL	deletion	UNP O75417
D	?	-	GLY	deletion	UNP O75417
D	?	-	GLU	deletion	UNP O75417
D	?	-	SER	deletion	UNP O75417
D	?	-	PRO	deletion	UNP O75417
D	?	-	PRO	deletion	UNP O75417
D	?	-	SER	deletion	UNP O75417
D	?	-	GLN	deletion	UNP O75417
D	?	-	ALA	deletion	UNP O75417
D	?	-	VAL	deletion	UNP O75417
D	?	-	GLY	deletion	UNP O75417
D	?	-	LYS	deletion	UNP O75417
D	?	-	GLY	deletion	UNP O75417
D	?	-	LEU	deletion	UNP O75417
D	?	-	LEU	deletion	UNP O75417
D	?	-	PRO	deletion	UNP O75417
D	?	-	MET	deletion	UNP O75417
D	?	-	GLY	deletion	UNP O75417
D	?	-	ARG	deletion	UNP O75417
D	?	-	GLY	deletion	UNP O75417
D	?	-	LYS	deletion	UNP O75417
D	?	-	TYR	deletion	UNP O75417
D	?	-	LYS	deletion	UNP O75417
D	?	-	LYS	deletion	UNP O75417
D	?	-	GLY	deletion	UNP O75417
D	?	-	PHE	deletion	UNP O75417
D	?	-	SER	deletion	UNP O75417

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	?	-	VAL	deletion	UNP O75417
D	?	-	ASN	deletion	UNP O75417
D	?	-	PRO	deletion	UNP O75417
D	?	-	ARG	deletion	UNP O75417
D	?	-	CYS	deletion	UNP O75417
D	?	-	GLN	deletion	UNP O75417
D	?	-	ALA	deletion	UNP O75417
D	?	-	GLN	deletion	UNP O75417
D	?	-	MET	deletion	UNP O75417
D	?	-	GLU	deletion	UNP O75417
D	?	-	GLU	deletion	UNP O75417
D	?	-	ARG	deletion	UNP O75417
D	?	-	ALA	deletion	UNP O75417
D	?	-	LEU	deletion	UNP O75417
D	?	-	SER	deletion	UNP O75417
D	?	-	ARG	deletion	UNP O75417
D	?	-	LYS	deletion	UNP O75417
D	?	-	ARG	deletion	UNP O75417
D	?	-	LYS	deletion	UNP O75417
D	?	-	LEU	deletion	UNP O75417
D	?	-	GLN	deletion	UNP O75417

- Molecule 2 is a DNA chain called DNA (5'-D(*CP*GP*TP*CP*CP*AP*AP*TP*GP*AP* CP*AP*GP*CP*CP*GP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	15	Total	C	N	O	P	0	0	0
			301	144	57	86	14			
2	E	12	Total	C	N	O	P	0	0	0
			241	116	46	68	11			

- Molecule 3 is a DNA chain called DNA (5'-D(*GP*C*GP*GP*CP*TP*GP*TP*CP*AP*T P*TP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	11	Total	C	N	O	P	0	0	0
			223	108	39	66	10			
3	F	8	Total	C	N	O	P	0	0	0
			163	79	26	50	8			

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total 1	Mg 1	0	0
4	D	1	Total 1	Mg 1	0	0

- Molecule 5 is 2'-3'-DIDEOXYGUANOSINE-5'-TRIPHOSPHATE (three-letter code: DG3) (formula: $C_{10}H_{16}N_5O_{12}P_3$).

DG3
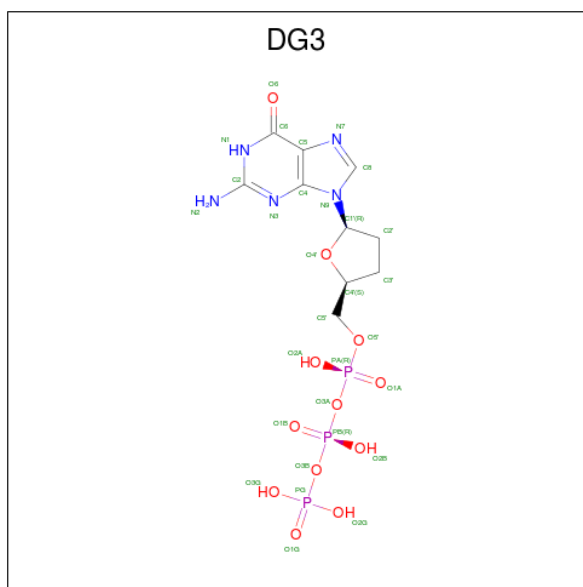
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total 30	C 10	N 5	O 12	P 3	0	0
5	D	1	Total 30	C 10	N 5	O 12	P 3	0	0

- Molecule 6 is N-methyl-N-phenyl[(3aM)-3-(trifluoromethyl)cyclopenta[c]pyrazol-2(1H)-yl] ethanethioamide (three-letter code: UAF) (formula: $C_{16}H_{14}F_3N_3S$) (labeled as "Ligand of Interest" by depositor).

UAF
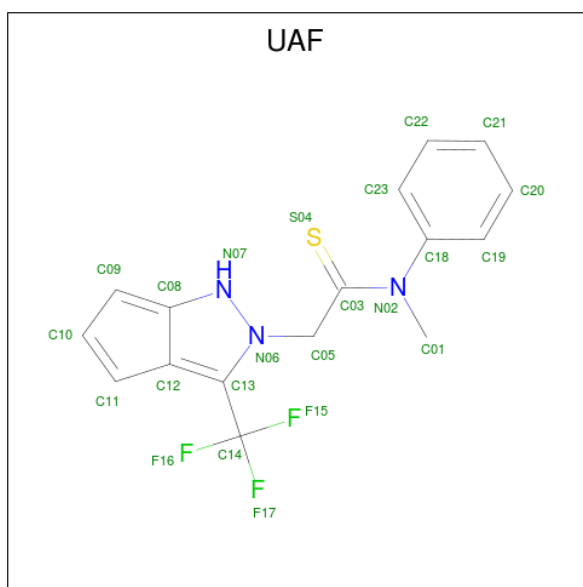
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	F	N	S		
6	A	1	23	16	3	3	1	0	0

• Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	67	Total 67	O 67	0	0
7	C	4	Total 4	O 4	0	0
7	D	19	Total 19	O 19	0	0
7	E	1	Total 1	O 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: DNA polymerase theta
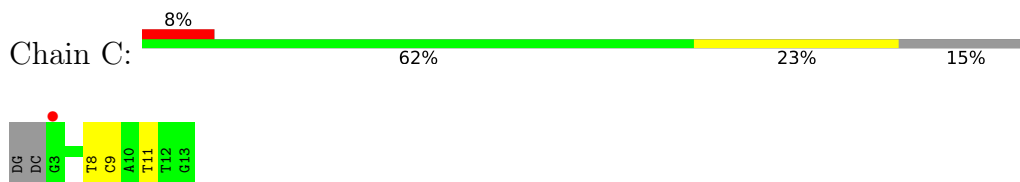

• Molecule 1: DNA polymerase theta

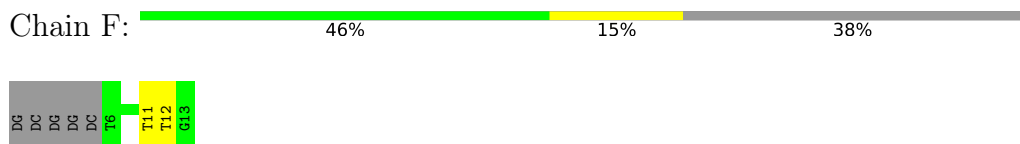● Molecule 2: DNA (5'-D(*CP*GP*TP*CP*CP*AP*AP*TP*GP*AP*CP*AP*GP*CP*CP*GP*C)-3')

● Molecule 2: DNA (5'-D(*CP*GP*TP*CP*CP*AP*AP*TP*GP*AP*CP*AP*GP*CP*CP*GP*C)-3')

● Molecule 3: DNA (5'-D(*GP*C*GP*GP*CP*TP*GP*TP*CP*AP*TP*TP*G)-3')

● Molecule 3: DNA (5'-D(*GP*C*GP*GP*CP*TP*GP*TP*CP*AP*TP*TP*G)-3')

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	246.85Å　　69.32Å　　147.43Å 90.00°　　122.55°　　90.00°	Depositor
Resolution (Å)	73.53　－　2.59 104.04　－　2.59	Depositor EDS
% Data completeness (in resolution range)	98.8 (73.53-2.59) 98.8 (104.04-2.59)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.23 (at 2.58Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.247　,　0.295 0.244　,　0.293	Depositor DCC
R_{free} test set	1963 reflections (3.02%)	wwPDB-VP
Wilson B-factor (Å^2)	84.0	Xtriage
Anisotropy	0.228	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 77.4	EDS
L-test for twinning[2]	$< |L| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	10562	wwPDB-VP
Average B, all atoms (Å^2)	104.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.93% 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: MG, DG3, UAF

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	0.23	0/5074	0.43	0/6856
1	D	0.23	0/4549	0.43	0/6163
2	B	0.47	0/337	0.80	0/517
2	E	0.51	0/270	0.85	0/414
3	C	0.46	0/249	0.95	0/383
3	F	0.45	0/181	1.03	0/277
All	All	0.27	0/10660	0.50	0/14610

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	4980	0	4940	29	0
1	D	4478	0	4136	49	0
2	B	301	0	169	1	0
2	E	241	0	136	5	0
3	C	223	0	125	2	0
3	F	163	0	91	2	0
4	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	1	0	0	0	0
5	A	30	0	12	0	0
5	D	30	0	12	1	0
6	A	23	0	0	0	0
7	A	67	0	0	2	0
7	C	4	0	0	0	0
7	D	19	0	0	0	0
7	E	1	0	0	0	0
All	All	10562	0	9621	82	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 (82) 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:D:2393:MET:HA	2:E:4:DC:H5'	1.73	0.69
1:A:1852:ILE:HB	1:A:1903:LEU:HD11	1.73	0.69
1:D:2241:ARG:HH21	1:D:2474:GLN:HB2	1.58	0.68
1:A:1854:LEU:HD23	1:A:1903:LEU:HD13	1.74	0.67
1:D:2241:ARG:NH2	1:D:2335:GLU:OE2	2.28	0.66
1:D:1852:ILE:HB	1:D:1903:LEU:HD11	1.76	0.66
1:A:1842:LYS:HE3	1:D:1937:PRO:HG2	1.80	0.64
1:D:2466:ARG:HH11	2:E:4:DC:H5"	1.63	0.64
1:D:2218:LYS:HA	1:D:2229:ILE:HG12	1.83	0.61
1:D:2004:THR:H	1:D:2007:SER:HB3	1.67	0.59
1:A:1953:ARG:HE	1:A:1983:GLU:HG3	1.69	0.58
1:D:1854:LEU:HD23	1:D:1903:LEU:HD13	1.86	0.58
1:D:2143:LEU:HD22	1:D:2190:LEU:HD13	1.86	0.57
1:D:2256:PHE:O	1:D:2312:ILE:N	2.39	0.56
1:D:2332:SER:HB2	1:D:2337:ARG:HH21	1.71	0.55
1:D:2538:LEU:HB2	1:D:2541:GLU:HB3	1.88	0.55
1:D:2184:LEU:HB3	1:D:2198:LEU:HD11	1.89	0.55
1:D:1918:GLN:O	1:D:1944:ARG:NH2	2.36	0.54
1:D:2315:ARG:HB3	1:D:2582:TRP:HB3	1.87	0.54
1:D:2098:GLU:HB2	1:D:2317:ALA:HB2	1.89	0.54
1:D:2241:ARG:NE	1:D:2474:GLN:OE1	2.43	0.51
1:D:2431:VAL:HG13	1:D:2453:ILE:HD11	1.91	0.51
2:E:5:DC:H2'	2:E:6:DA:H8	1.76	0.51
1:D:1903:LEU:O	1:D:1915:PHE:N	2.38	0.51
1:A:2185:ASN:HA	1:A:2188:LYS:HD3	1.93	0.51

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:2254:ARG:HH21	1:D:2379:ARG:HH21	1.58	0.50
1:D:2332:SER:O	1:D:2337:ARG:NE	2.41	0.50
3:F:11:DT:H2'	3:F:12:DT:C6	2.47	0.50
1:A:2348:LEU:HD22	1:A:2423:ILE:HD11	1.93	0.50
1:D:2466:ARG:NH1	2:E:4:DC:H5"	2.28	0.49
1:D:1940:THR:OG1	1:D:1943:ASP:OD1	2.23	0.49
1:D:2093:GLY:HA2	1:D:2228:ARG:HG2	1.94	0.49
1:A:2028:GLY:O	1:A:2030:GLN:N	2.43	0.48
1:A:2084:LEU:HD11	1:A:2242:ILE:HD13	1.96	0.47
1:A:2181:LYS:NZ	3:C:11:DT:OP2	2.45	0.47
1:A:2538:LEU:HB3	1:A:2541:GLU:HB3	1.95	0.47
1:D:2247:PRO:HG2	1:D:2249:ILE:HD11	1.97	0.47
1:D:2499:PHE:CE2	1:D:2505:ARG:HG2	2.50	0.47
1:A:2341:HIS:CD2	1:A:2568:LEU:HD22	2.49	0.47
1:D:2333:GLN:NE2	5:D:2601:DG3:O1G	2.47	0.47
2:E:5:DC:H2'	2:E:6:DA:C8	2.50	0.47
1:D:2439:PHE:HB2	1:D:2447:ARG:HB3	1.96	0.47
1:D:2013:LEU:HD11	1:D:2055:ASN:HB2	1.96	0.46
1:A:2441:GLN:HG2	1:A:2447:ARG:HG2	1.97	0.46
1:D:2191:HIS:O	1:D:2194:PRO:HD2	2.14	0.46
1:D:2398:LEU:HD23	1:D:2409:ALA:HB1	1.96	0.46
1:A:1941:LEU:O	1:A:1945:MET:HG2	2.16	0.45
1:A:2338:ILE:HD11	1:A:2480:ILE:HD12	1.98	0.45
1:D:1989:PRO:HB2	1:D:2008:ILE:HD13	1.98	0.45
1:D:2114:GLU:HG2	1:D:2127:PHE:CE1	2.51	0.45
1:A:1965:TYR:HD1	1:A:1989:PRO:HG2	1.82	0.45
1:A:2446:ARG:NH1	1:A:2479:ASP:OD1	2.50	0.44
1:D:2315:ARG:HD3	1:D:2582:TRP:CD2	2.52	0.44
1:A:2369:ILE:HD11	1:A:2374:VAL:HG22	1.98	0.44
3:F:11:DT:H2'	3:F:12:DT:H6	1.82	0.44
1:D:1968:ILE:HG12	1:D:2085:ALA:HB2	2.00	0.44
1:A:2446:ARG:NH2	7:A:2707:HOH:O	2.51	0.43
1:D:2385:ILE:HG23	1:D:2398:LEU:HD11	2.00	0.43
1:D:2434:CYS:HA	1:D:2437:ASP:O	2.18	0.43
1:D:1994:TRP:HZ2	1:D:2240:GLY:HA2	1.82	0.43
1:A:1994:TRP:HZ2	1:A:2240:GLY:HA2	1.84	0.43
1:A:2128:THR:HB	2:B:12:DA:H5'	2.01	0.43
1:D:2009:VAL:HG21	1:D:2052:LEU:HD13	2.00	0.43
1:A:2191:HIS:O	1:A:2194:PRO:HD2	2.19	0.42
1:A:2315:ARG:HB3	1:A:2582:TRP:HB3	2.00	0.42
1:A:2367:LYS:HG3	1:A:2378:LEU:HD21	2.01	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1832:SER:HB2	1:A:1935:LEU:HD12	2.02	0.42
1:A:2577:LYS:HG2	1:A:2587:ASP:HA	2.02	0.42
1:D:2501:SER:O	1:D:2505:ARG:HG3	2.18	0.42
1:D:2017:LEU:N	1:D:2018:PRO:HD2	2.34	0.42
1:A:1994:TRP:CZ2	1:A:2240:GLY:HA2	2.55	0.41
1:D:1825:LEU:HD23	1:D:1827:ILE:HD11	2.02	0.41
1:D:2457:ASN:HB3	1:D:2460:ARG:HB2	2.02	0.41
1:D:2193:LEU:HB3	1:D:2194:PRO:HD3	2.03	0.41
1:D:2369:ILE:HG12	1:D:2370:GLU:H	1.85	0.41
1:A:1968:ILE:O	1:A:1972:LYS:HG3	2.20	0.41
3:C:8:DT:H2"	3:C:9:DC:C6	2.56	0.41
1:D:2194:PRO:O	1:D:2198:LEU:HD12	2.21	0.41
1:D:2501:SER:HB3	1:D:2504:HIS:HB2	2.02	0.41
1:A:2250:GLN:HG2	1:A:2538:LEU:HD12	2.04	0.40
1:A:2380:GLN:NE2	7:A:2709:HOH:O	2.54	0.40
1:D:2087:LEU:HD12	1:D:2534:PHE:CE2	2.57	0.40

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	627/668 (94%)	612 (98%)	15 (2%)	0	100	100
1	D	575/668 (86%)	554 (96%)	21 (4%)	0	100	100
All	All	1202/1336 (90%)	1166 (97%)	36 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	539/586 (92%)	539 (100%)	0	100	100
1	D	446/586 (76%)	444 (100%)	2 (0%)	91	97
All	All	985/1172 (84%)	983 (100%)	2 (0%)	93	98

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	1949	GLN
1	D	2198	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	1969	GLN
1	D	2030	GLN

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 ⓘ

Of 5 ligands modelled in this entry, 2 are monoatomic - leaving 3 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	#$|Z| > 2$
5	DG3	A	2601	4	25,32,32	0.87	1 (4%)	28,50,50	1.50	5 (17%)
5	DG3	D	2601	4	25,32,32	0.87	1 (4%)	28,50,50	1.52	5 (17%)
6	UAF	A	2602	-	20,25,25	2.52	8 (40%)	22,37,37	1.68	8 (36%)

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
5	DG3	A	2601	4	-	4/18/31/31	0/3/3/3
5	DG3	D	2601	4	-	5/18/31/31	0/3/3/3
6	UAF	A	2602	-	-	3/17/18/18	0/3/3/3

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	2602	UAF	C05-N06	4.93	1.56	1.47
6	A	2602	UAF	C10-C09	4.08	1.56	1.39
6	A	2602	UAF	C10-C11	4.08	1.56	1.39
6	A	2602	UAF	C03-S04	-3.76	1.58	1.66
6	A	2602	UAF	C13-N06	3.69	1.42	1.37
6	A	2602	UAF	C18-N02	3.65	1.50	1.43
6	A	2602	UAF	C09-C08	2.70	1.54	1.42
6	A	2602	UAF	C01-N02	2.40	1.50	1.46
5	D	2601	DG3	C6-N1	-2.35	1.34	1.37
5	A	2601	DG3	C6-N1	-2.32	1.34	1.37

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	2602	UAF	C11-C12-C08	3.69	109.31	106.18
5	D	2601	DG3	PB-O3B-PG	-3.50	120.81	132.83
5	A	2601	DG3	PA-O3A-PB	-3.23	121.75	132.83
5	A	2601	DG3	PB-O3B-PG	-3.17	121.94	132.83
6	A	2602	UAF	C13-N06-N07	-2.96	108.70	112.10
5	D	2601	DG3	PA-O3A-PB	-2.92	122.80	132.83
5	D	2601	DG3	C3'-C2'-C1'	2.87	106.09	102.78
6	A	2602	UAF	F17-C14-C13	2.79	116.81	112.13
5	A	2601	DG3	C3'-C2'-C1'	2.67	105.87	102.78
6	A	2602	UAF	C23-C18-N02	2.60	123.49	120.10
5	D	2601	DG3	C8-N7-C5	2.35	107.46	102.99
5	A	2601	DG3	C8-N7-C5	2.32	107.41	102.99
5	D	2601	DG3	C5-C6-N1	2.32	118.04	113.95
5	A	2601	DG3	C5-C6-N1	2.32	118.04	113.95
6	A	2602	UAF	F15-C14-C13	2.17	115.75	112.13
6	A	2602	UAF	C10-C09-C08	-2.04	105.72	108.39
6	A	2602	UAF	C05-C03-S04	-2.01	116.67	120.11
6	A	2602	UAF	C05-N06-N07	2.00	123.41	117.77

There are no chirality outliers.

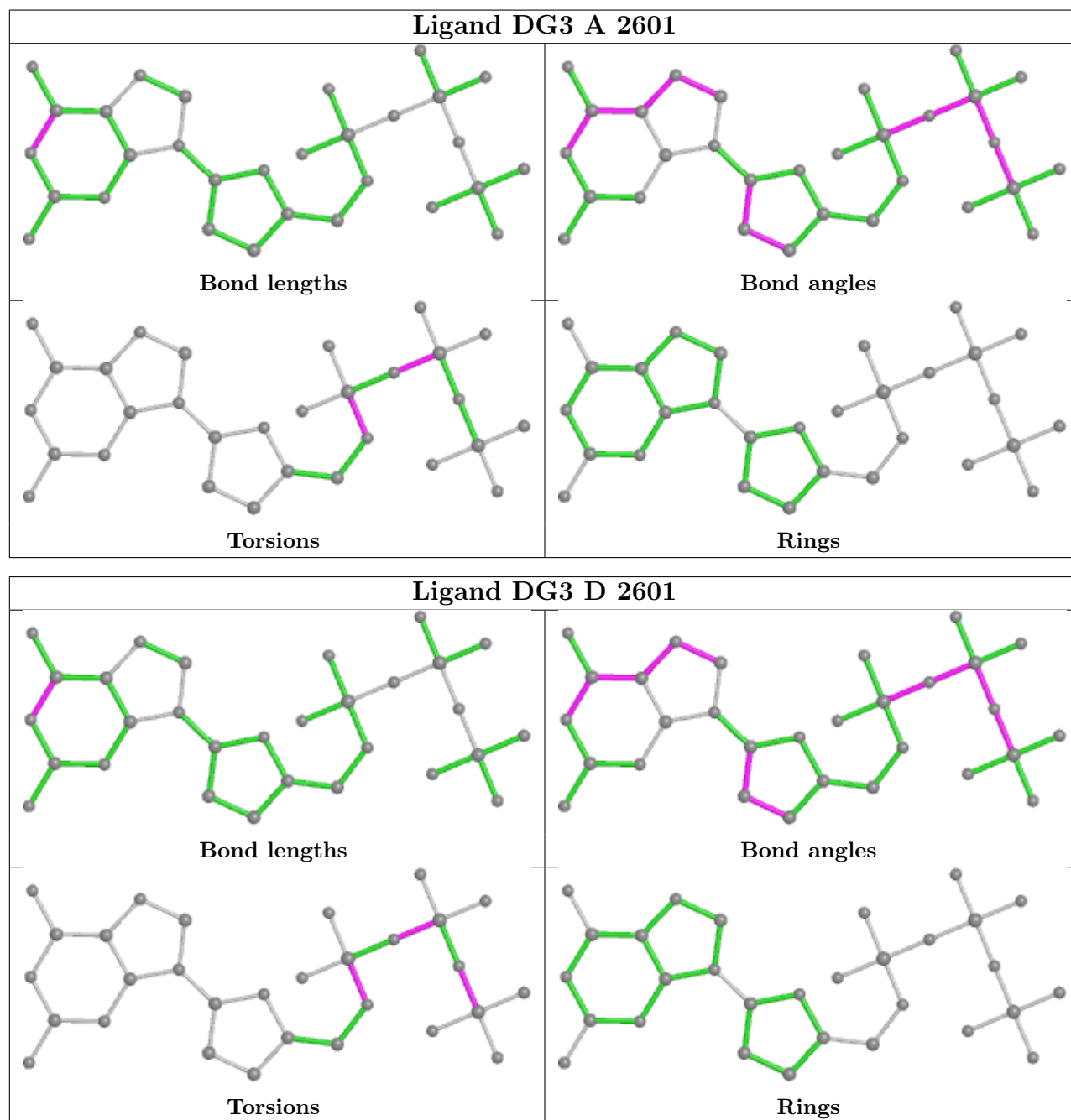All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	2601	DG3	C5'-O5'-PA-O1A
5	A	2601	DG3	C5'-O5'-PA-O2A
5	D	2601	DG3	C5'-O5'-PA-O1A
5	D	2601	DG3	C5'-O5'-PA-O2A
6	A	2602	UAF	C12-C13-C14-F16
5	D	2601	DG3	PB-O3B-PG-O2G
5	A	2601	DG3	C5'-O5'-PA-O3A
5	D	2601	DG3	C5'-O5'-PA-O3A
5	A	2601	DG3	PA-O3A-PB-O2B
5	D	2601	DG3	PA-O3A-PB-O1B
6	A	2602	UAF	C12-C13-C14-F15
6	A	2602	UAF	C12-C13-C14-F17

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	D	2601	DG3	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 DG3 A 2601

Bond lengths

Bond angles

Torsions

Rings

Ligand DG3 D 2601

Bond lengths

Bond angles

Torsions

Rings

Ligand UAF A 2602

Bond lengths

Bond angles

Torsions

Rings
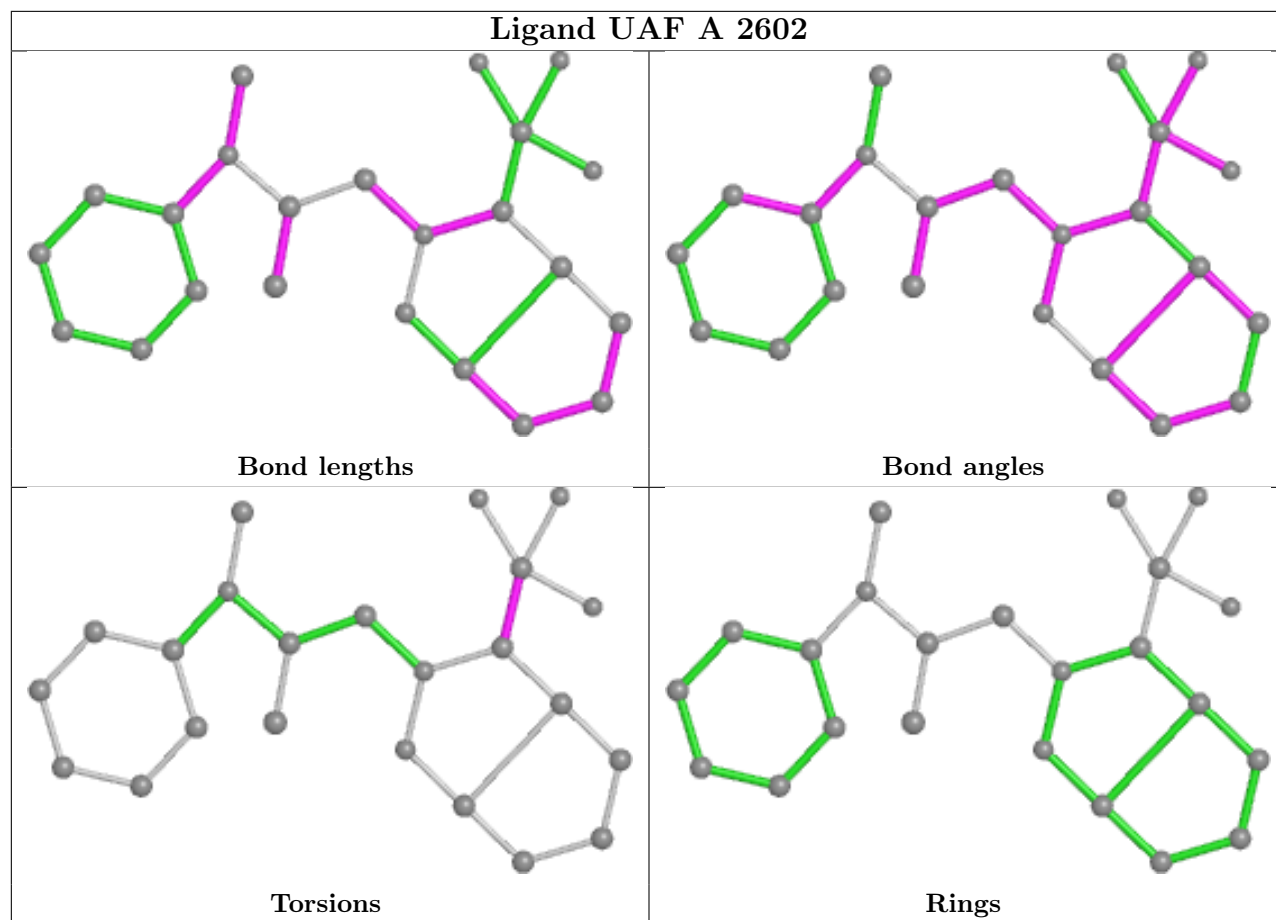
5.7 Other polymers ⓘ

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	637/668 (95%)	0.71	34 (5%)	26	20	48, 77, 119, 187	0
1	D	603/668 (90%)	1.18	131 (21%)	0	0	66, 129, 187, 246	0
2	B	15/17 (88%)	0.77	1 (6%)	17	13	63, 91, 200, 233	0
2	E	12/17 (70%)	0.13	0	100	100	99, 119, 171, 175	0
3	C	11/13 (84%)	0.75	1 (9%)	9	6	62, 125, 215, 217	0
3	F	8/13 (61%)	0.57	0	100	100	130, 136, 175, 186	0
All	All	1286/1396 (92%)	0.93	167 (12%)	3	2	48, 94, 177, 246	0

All (167) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	2368	MET	4.8
1	D	2340	ALA	4.8
1	D	2344	HIS	4.7
1	D	2439	PHE	4.6
1	D	2249	ILE	4.5
1	D	2328	ALA	4.4
1	D	2453	ILE	4.3
1	D	2411	CYS	4.3
1	D	2205	ASN	4.2
1	D	2252	VAL	4.2
1	D	2375	GLY	4.1
1	D	2541	GLU	4.1
1	D	2478	ALA	4.0
1	D	2193	LEU	4.0
1	D	2203	ILE	3.9
1	D	2402	MET	3.9
1	D	2489	GLN	3.9
1	D	2376	ASP	3.9
1	D	2546	VAL	3.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	2326	ILE	3.8
1	D	2318	PHE	3.8
1	D	2576	VAL	3.7
1	D	2212	PHE	3.7
1	A	2355	GLY	3.7
1	D	2394	GLY	3.7
1	D	2110	LEU	3.6
1	D	2582	TRP	3.6
1	D	2314	MET	3.6
1	D	2355	GLY	3.5
1	D	1859	ILE	3.5
1	D	2359	PHE	3.5
1	D	2392	GLY	3.5
1	D	2440	VAL	3.5
1	D	2137	LEU	3.4
1	D	2538	LEU	3.4
1	D	2125	PHE	3.4
1	D	2494	THR	3.4
1	D	2389	ILE	3.4
1	D	2480	ILE	3.4
1	D	2562	MET	3.3
1	D	2444	LEU	3.3
1	D	2186	LYS	3.3
1	D	2488	ILE	3.3
1	D	2256	PHE	3.3
1	D	2578	ILE	3.3
1	D	2091	GLY	3.2
1	D	1939	LEU	3.2
1	D	2423	ILE	3.2
1	D	2385	ILE	3.2
1	D	2561	GLU	3.2
1	D	2559	LYS	3.2
1	D	2588	PHE	3.1
1	A	1838	GLN	3.1
1	A	2590	VAL	3.1
1	D	2443	ILE	3.0
1	D	2037	GLY	3.0
1	D	2551	VAL	3.0
1	D	2102	GLN	3.0
1	D	2023	MET	3.0
1	D	2386	CYS	3.0
1	D	2123	HIS	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	2329	ALA	2.9
1	D	2177	PHE	2.9
1	D	2103	LYS	2.9
1	D	2210	VAL	2.9
1	A	2039	GLU	2.9
1	D	2477	ALA	2.9
1	D	2141	LEU	2.8
1	D	2222	PRO	2.8
1	D	2441	GLN	2.8
1	D	2493	GLU	2.8
1	D	2503	GLY	2.8
1	D	2374	VAL	2.7
1	D	2331	TYR	2.7
1	D	2403	GLY	2.6
1	D	2473	VAL	2.6
1	D	2484	ALA	2.6
1	A	1899	LEU	2.6
1	D	2534	PHE	2.6
1	A	2031	SER	2.6
1	D	2145	PRO	2.5
1	D	2544	TYR	2.5
1	D	2327	LEU	2.5
1	D	2450	LEU	2.5
1	D	2183	VAL	2.5
1	D	2190	LEU	2.5
1	D	2449	TYR	2.5
1	D	2099	CYS	2.4
1	A	2136	VAL	2.4
1	D	1965	TYR	2.4
1	D	2387	TYR	2.4
1	D	2558	VAL	2.4
1	D	2184	LEU	2.4
1	A	2332	SER	2.4
1	D	2247	PRO	2.4
1	D	2419	ARG	2.4
1	A	2258	ILE	2.4
2	B	15	DC	2.4
1	D	2226	MET	2.4
1	D	1959	GLU	2.4
1	A	2499	PHE	2.4
1	D	2502	HIS	2.4
1	A	2413	ILE	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	2115	THR	2.4
1	D	1824	SER	2.4
1	D	2325	SER	2.4
1	D	2209	LYS	2.3
1	A	1917	LEU	2.3
1	D	2113	ILE	2.3
1	A	1865	SER	2.3
1	A	2143	LEU	2.3
1	D	2542	LEU	2.3
1	D	2485	THR	2.3
1	D	2196	LEU	2.3
1	A	2570	VAL	2.3
1	D	2416	PHE	2.3
1	D	2565	ALA	2.3
1	D	2581	SER	2.3
1	D	2189	ALA	2.3
1	A	1903	LEU	2.3
1	A	2458	PRO	2.3
1	A	1919	LYS	2.3
1	A	1939	LEU	2.3
1	D	1955	GLU	2.3
1	D	2136	VAL	2.2
1	D	2469	ILE	2.2
1	A	2187	LEU	2.2
1	D	2224	LEU	2.2
1	D	2589	ASP	2.2
1	A	2358	VAL	2.2
1	D	2124	SER	2.2
1	D	2207	ILE	2.2
1	D	1825	LEU	2.2
3	C	3	DG	2.2
1	A	1920	GLU	2.2
1	D	2568	LEU	2.2
1	D	2178	SER	2.2
1	A	2473	VAL	2.2
1	A	2354	THR	2.2
1	A	2183	VAL	2.2
1	A	2211	VAL	2.2
1	D	2507	GLY	2.2
1	D	2536	LEU	2.2
1	D	2199	GLU	2.2
1	D	2574	VAL	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1825	LEU	2.1
1	D	1952	LEU	2.1
1	D	2013	LEU	2.1
1	D	2330	ASP	2.1
1	A	2106	MET	2.1
1	D	2065	GLN	2.1
1	D	2211	VAL	2.1
1	D	2557	ILE	2.1
1	A	2198	LEU	2.1
1	A	2495	PHE	2.1
1	D	2020	LEU	2.1
1	D	2563	GLU	2.1
1	D	2343	SER	2.1
1	D	2048	VAL	2.1
1	A	2312	ILE	2.1
1	A	2326	ILE	2.1
1	D	2341	HIS	2.1
1	D	2447	ARG	2.1
1	A	2443	ILE	2.0
1	D	2420	TYR	2.0
1	D	2068	ASN	2.0
1	A	1964	ILE	2.0

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.
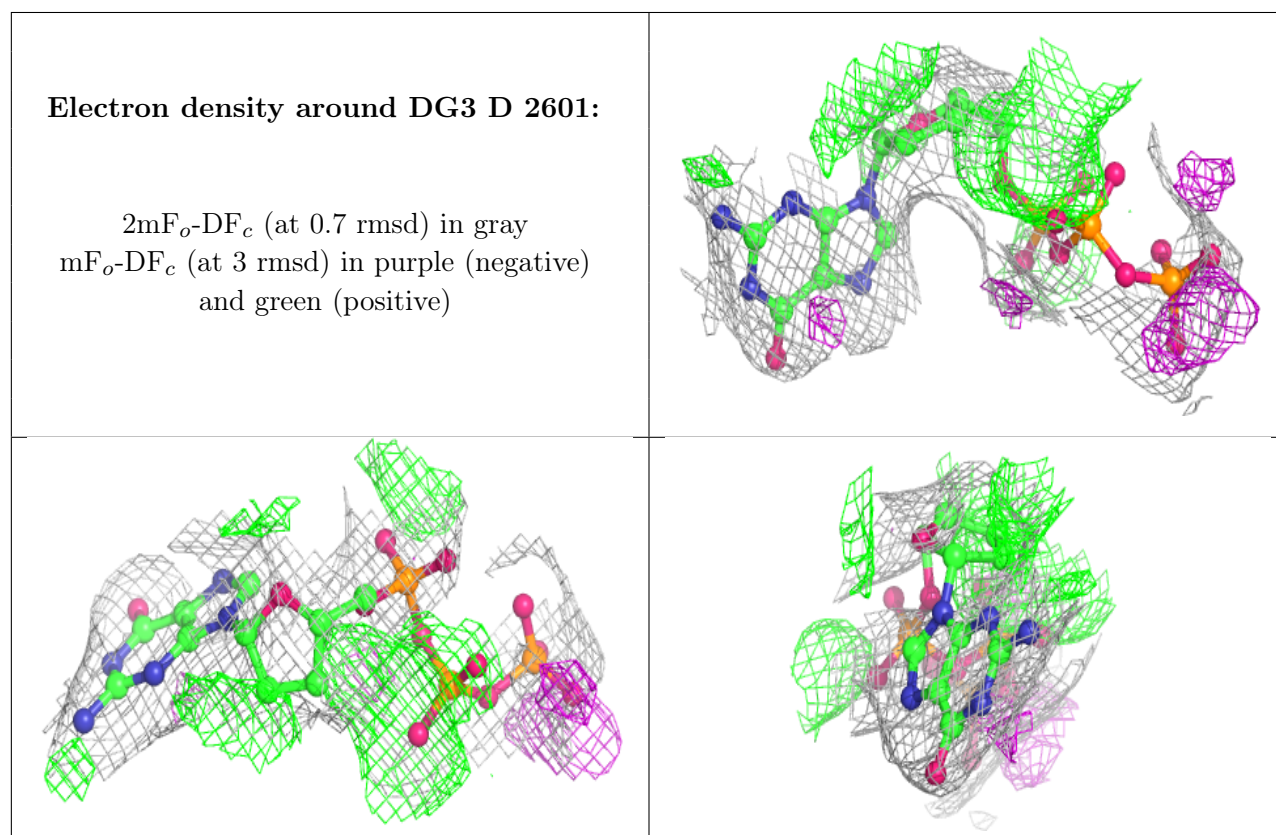
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	MG	D	2600	1/1	0.59	0.09	146,146,146,146	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	DG3	D	2601	30/30	0.79	0.16	109,138,158,164	0
4	MG	A	2600	1/1	0.80	0.17	62,62,62,62	0
5	DG3	A	2601	30/30	0.97	0.22	47,63,78,84	0
6	UAF	A	2602	23/23	0.98	0.26	58,70,78,82	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.
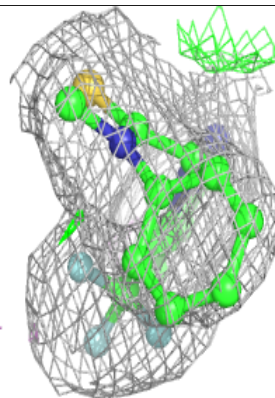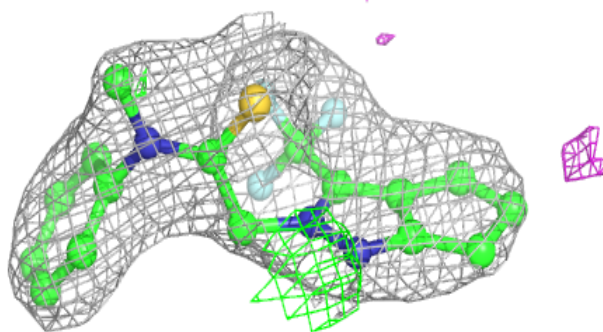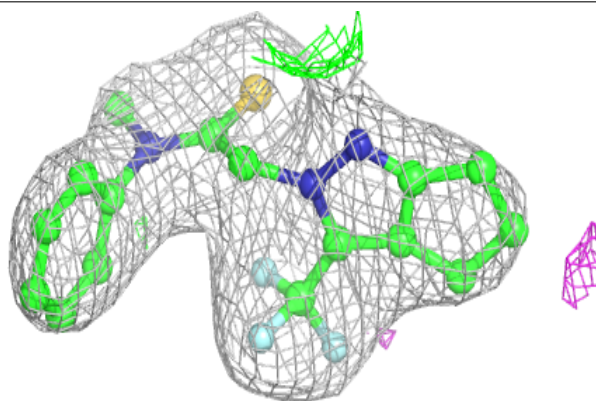

Electron density around DG3 D 2601:

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

Electron density around DG3 A 2601:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
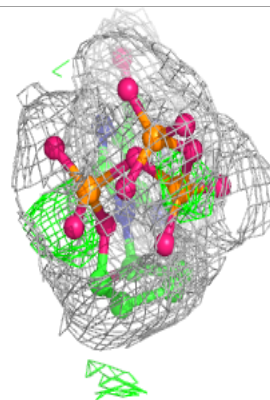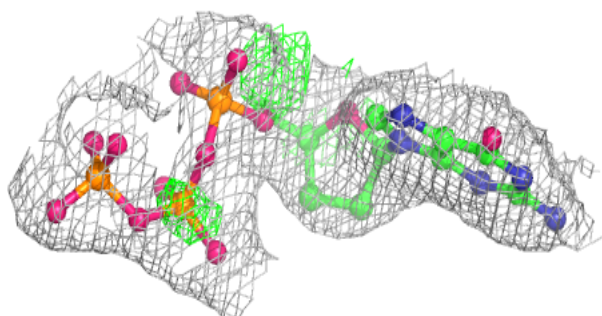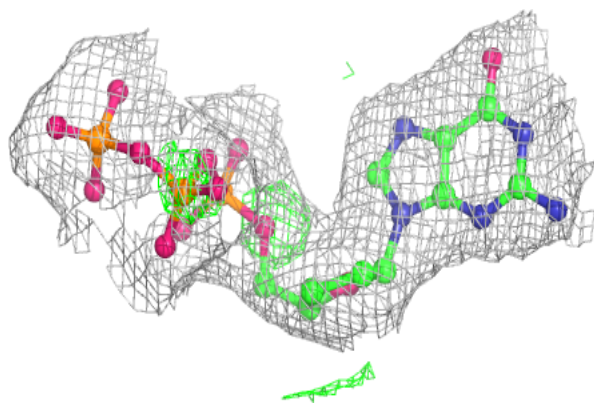
Electron density around UAF A 2602:

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

6.5 Other polymers (i)

There are no such residues in this entry.