wwPDB X-ray Structure Validation Summary Report (i)

Oct 6, 2022 – 04:52 pm BST

PDB ID	:	7ZX1
Title	:	Crystal structure of Pol theta polymerase domain in complex with compound 22
Authors	:	Krajewski, W.W.; Turnbull, A.P.; Willis, S.; Charles, M.; Stockley, M.; Heald, R.A.
Deposited on	:	2022-05-19
Resolution	:	2.83 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 (i) 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 (i)) were used in the production of this report:

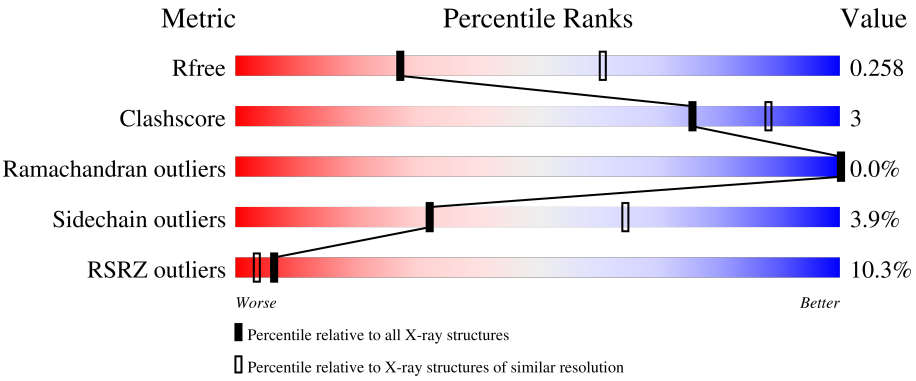
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.31.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.83 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



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

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=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	AAA	726	4% 79% 9% 12%
1	BBB	726	5% 78% 8% 13%
1	CCC	726	4% 79% 9% 12%
1	DDD	726	12% 77% 10% 13%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	EEE	726	13% 77% 10% 13%
1	FFF	726	18% 79% 9% 13%
2	GGG	16	56% 38% 6%
2	III	16	50% 44% 6%
2	KKK	16	44% 50% 6%
2	MMM	16	44% 50% 6%
2	OOO	16	62% 25% 12%
2	QQQ	16	6% 56% 38% 6%
3	HHH	13	38% 62%
3	JJJ	13	46% 54%
3	LLL	13	46% 54%
3	NNN	13	46% 54%
3	PPP	13	38% 62%
3	RRR	13	38% 62%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 32843 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	AAA	639	Total	C	N	O	S	0	0	0
			4884	3123	825	906	30			
1	BBB	629	Total	C	N	O	S	0	0	0
			4811	3075	811	896	29			
1	CCC	638	Total	C	N	O	S	0	0	0
			4856	3107	822	897	30			
1	DDD	633	Total	C	N	O	S	0	0	0
			4810	3072	808	901	29			
1	EEE	632	Total	C	N	O	S	0	0	0
			4824	3086	814	896	28			
1	FFF	634	Total	C	N	O	S	0	0	0
			4823	3086	813	896	28			

There are 276 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	2261	GLY	PRO	engineered mutation	UNP O75417
AAA	?	-	THR	deletion	UNP O75417
AAA	?	-	LEU	deletion	UNP O75417
AAA	?	-	VAL	deletion	UNP O75417
AAA	?	-	GLY	deletion	UNP O75417
AAA	?	-	GLU	deletion	UNP O75417
AAA	?	-	SER	deletion	UNP O75417
AAA	?	-	PRO	deletion	UNP O75417
AAA	?	-	PRO	deletion	UNP O75417
AAA	?	-	SER	deletion	UNP O75417
AAA	?	-	GLN	deletion	UNP O75417
AAA	?	-	ALA	deletion	UNP O75417
AAA	?	-	VAL	deletion	UNP O75417
AAA	?	-	GLY	deletion	UNP O75417
AAA	?	-	LYS	deletion	UNP O75417
AAA	?	-	GLY	deletion	UNP O75417
AAA	?	-	LEU	deletion	UNP O75417

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	?	-	LEU	deletion	UNP O75417
AAA	?	-	PRO	deletion	UNP O75417
AAA	?	-	MET	deletion	UNP O75417
AAA	?	-	GLY	deletion	UNP O75417
AAA	?	-	ARG	deletion	UNP O75417
AAA	?	-	GLY	deletion	UNP O75417
AAA	?	-	LYS	deletion	UNP O75417
AAA	?	-	TYR	deletion	UNP O75417
AAA	?	-	LYS	deletion	UNP O75417
AAA	?	-	LYS	deletion	UNP O75417
AAA	?	-	GLY	deletion	UNP O75417
AAA	?	-	PHE	deletion	UNP O75417
AAA	?	-	SER	deletion	UNP O75417
AAA	?	-	VAL	deletion	UNP O75417
AAA	?	-	ASN	deletion	UNP O75417
AAA	?	-	PRO	deletion	UNP O75417
AAA	?	-	ARG	deletion	UNP O75417
AAA	?	-	CYS	deletion	UNP O75417
AAA	?	-	GLN	deletion	UNP O75417
AAA	?	-	ALA	deletion	UNP O75417
AAA	?	-	GLN	deletion	UNP O75417
AAA	?	-	MET	deletion	UNP O75417
AAA	?	-	GLU	deletion	UNP O75417
AAA	?	-	GLU	deletion	UNP O75417
AAA	?	-	ARG	deletion	UNP O75417
AAA	?	-	ALA	deletion	UNP O75417
AAA	?	-	ALA	deletion	UNP O75417
AAA	?	-	ASP	deletion	UNP O75417
AAA	?	-	ARG	deletion	UNP O75417
BBB	2261	GLY	PRO	engineered mutation	UNP O75417
BBB	?	-	THR	deletion	UNP O75417
BBB	?	-	LEU	deletion	UNP O75417
BBB	?	-	VAL	deletion	UNP O75417
BBB	?	-	GLY	deletion	UNP O75417
BBB	?	-	GLU	deletion	UNP O75417
BBB	?	-	SER	deletion	UNP O75417
BBB	?	-	PRO	deletion	UNP O75417
BBB	?	-	PRO	deletion	UNP O75417
BBB	?	-	SER	deletion	UNP O75417
BBB	?	-	GLN	deletion	UNP O75417
BBB	?	-	ALA	deletion	UNP O75417
BBB	?	-	VAL	deletion	UNP O75417

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	?	-	GLY	deletion	UNP O75417
BBB	?	-	LYS	deletion	UNP O75417
BBB	?	-	GLY	deletion	UNP O75417
BBB	?	-	LEU	deletion	UNP O75417
BBB	?	-	LEU	deletion	UNP O75417
BBB	?	-	PRO	deletion	UNP O75417
BBB	?	-	MET	deletion	UNP O75417
BBB	?	-	GLY	deletion	UNP O75417
BBB	?	-	ARG	deletion	UNP O75417
BBB	?	-	GLY	deletion	UNP O75417
BBB	?	-	LYS	deletion	UNP O75417
BBB	?	-	TYR	deletion	UNP O75417
BBB	?	-	LYS	deletion	UNP O75417
BBB	?	-	LYS	deletion	UNP O75417
BBB	?	-	GLY	deletion	UNP O75417
BBB	?	-	PHE	deletion	UNP O75417
BBB	?	-	SER	deletion	UNP O75417
BBB	?	-	VAL	deletion	UNP O75417
BBB	?	-	ASN	deletion	UNP O75417
BBB	?	-	PRO	deletion	UNP O75417
BBB	?	-	ARG	deletion	UNP O75417
BBB	?	-	CYS	deletion	UNP O75417
BBB	?	-	GLN	deletion	UNP O75417
BBB	?	-	ALA	deletion	UNP O75417
BBB	?	-	GLN	deletion	UNP O75417
BBB	?	-	MET	deletion	UNP O75417
BBB	?	-	GLU	deletion	UNP O75417
BBB	?	-	GLU	deletion	UNP O75417
BBB	?	-	ARG	deletion	UNP O75417
BBB	?	-	ALA	deletion	UNP O75417
BBB	?	-	ALA	deletion	UNP O75417
BBB	?	-	ASP	deletion	UNP O75417
BBB	?	-	ARG	deletion	UNP O75417
CCC	2261	GLY	PRO	engineered mutation	UNP O75417
CCC	?	-	THR	deletion	UNP O75417
CCC	?	-	LEU	deletion	UNP O75417
CCC	?	-	VAL	deletion	UNP O75417
CCC	?	-	GLY	deletion	UNP O75417
CCC	?	-	GLU	deletion	UNP O75417
CCC	?	-	SER	deletion	UNP O75417
CCC	?	-	PRO	deletion	UNP O75417
CCC	?	-	PRO	deletion	UNP O75417

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
CCC	?	-	SER	deletion	UNP O75417
CCC	?	-	GLN	deletion	UNP O75417
CCC	?	-	ALA	deletion	UNP O75417
CCC	?	-	VAL	deletion	UNP O75417
CCC	?	-	GLY	deletion	UNP O75417
CCC	?	-	LYS	deletion	UNP O75417
CCC	?	-	GLY	deletion	UNP O75417
CCC	?	-	LEU	deletion	UNP O75417
CCC	?	-	LEU	deletion	UNP O75417
CCC	?	-	PRO	deletion	UNP O75417
CCC	?	-	MET	deletion	UNP O75417
CCC	?	-	GLY	deletion	UNP O75417
CCC	?	-	ARG	deletion	UNP O75417
CCC	?	-	GLY	deletion	UNP O75417
CCC	?	-	LYS	deletion	UNP O75417
CCC	?	-	TYR	deletion	UNP O75417
CCC	?	-	LYS	deletion	UNP O75417
CCC	?	-	LYS	deletion	UNP O75417
CCC	?	-	GLY	deletion	UNP O75417
CCC	?	-	PHE	deletion	UNP O75417
CCC	?	-	SER	deletion	UNP O75417
CCC	?	-	VAL	deletion	UNP O75417
CCC	?	-	ASN	deletion	UNP O75417
CCC	?	-	PRO	deletion	UNP O75417
CCC	?	-	ARG	deletion	UNP O75417
CCC	?	-	CYS	deletion	UNP O75417
CCC	?	-	GLN	deletion	UNP O75417
CCC	?	-	ALA	deletion	UNP O75417
CCC	?	-	GLN	deletion	UNP O75417
CCC	?	-	MET	deletion	UNP O75417
CCC	?	-	GLU	deletion	UNP O75417
CCC	?	-	GLU	deletion	UNP O75417
CCC	?	-	ARG	deletion	UNP O75417
CCC	?	-	ALA	deletion	UNP O75417
CCC	?	-	ALA	deletion	UNP O75417
CCC	?	-	ASP	deletion	UNP O75417
CCC	?	-	ARG	deletion	UNP O75417
DDD	2261	GLY	PRO	engineered mutation	UNP O75417
DDD	?	-	THR	deletion	UNP O75417
DDD	?	-	LEU	deletion	UNP O75417
DDD	?	-	VAL	deletion	UNP O75417
DDD	?	-	GLY	deletion	UNP O75417

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
DDD	?	-	GLU	deletion	UNP O75417
DDD	?	-	SER	deletion	UNP O75417
DDD	?	-	PRO	deletion	UNP O75417
DDD	?	-	PRO	deletion	UNP O75417
DDD	?	-	SER	deletion	UNP O75417
DDD	?	-	GLN	deletion	UNP O75417
DDD	?	-	ALA	deletion	UNP O75417
DDD	?	-	VAL	deletion	UNP O75417
DDD	?	-	GLY	deletion	UNP O75417
DDD	?	-	LYS	deletion	UNP O75417
DDD	?	-	GLY	deletion	UNP O75417
DDD	?	-	LEU	deletion	UNP O75417
DDD	?	-	LEU	deletion	UNP O75417
DDD	?	-	PRO	deletion	UNP O75417
DDD	?	-	MET	deletion	UNP O75417
DDD	?	-	GLY	deletion	UNP O75417
DDD	?	-	ARG	deletion	UNP O75417
DDD	?	-	GLY	deletion	UNP O75417
DDD	?	-	LYS	deletion	UNP O75417
DDD	?	-	TYR	deletion	UNP O75417
DDD	?	-	LYS	deletion	UNP O75417
DDD	?	-	LYS	deletion	UNP O75417
DDD	?	-	GLY	deletion	UNP O75417
DDD	?	-	PHE	deletion	UNP O75417
DDD	?	-	SER	deletion	UNP O75417
DDD	?	-	VAL	deletion	UNP O75417
DDD	?	-	ASN	deletion	UNP O75417
DDD	?	-	PRO	deletion	UNP O75417
DDD	?	-	ARG	deletion	UNP O75417
DDD	?	-	CYS	deletion	UNP O75417
DDD	?	-	GLN	deletion	UNP O75417
DDD	?	-	ALA	deletion	UNP O75417
DDD	?	-	GLN	deletion	UNP O75417
DDD	?	-	MET	deletion	UNP O75417
DDD	?	-	GLU	deletion	UNP O75417
DDD	?	-	GLU	deletion	UNP O75417
DDD	?	-	ARG	deletion	UNP O75417
DDD	?	-	ALA	deletion	UNP O75417
DDD	?	-	ALA	deletion	UNP O75417
DDD	?	-	ASP	deletion	UNP O75417
DDD	?	-	ARG	deletion	UNP O75417
EEE	2261	GLY	PRO	engineered mutation	UNP O75417

Continued on next page...

Continued from previous page...

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

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
EEE	?	-	ALA	deletion	UNP O75417
EEE	?	-	ASP	deletion	UNP O75417
EEE	?	-	ARG	deletion	UNP O75417
FFF	2261	GLY	PRO	engineered mutation	UNP O75417
FFF	?	-	THR	deletion	UNP O75417
FFF	?	-	LEU	deletion	UNP O75417
FFF	?	-	VAL	deletion	UNP O75417
FFF	?	-	GLY	deletion	UNP O75417
FFF	?	-	GLU	deletion	UNP O75417
FFF	?	-	SER	deletion	UNP O75417
FFF	?	-	PRO	deletion	UNP O75417
FFF	?	-	PRO	deletion	UNP O75417
FFF	?	-	SER	deletion	UNP O75417
FFF	?	-	GLN	deletion	UNP O75417
FFF	?	-	ALA	deletion	UNP O75417
FFF	?	-	VAL	deletion	UNP O75417
FFF	?	-	GLY	deletion	UNP O75417
FFF	?	-	LYS	deletion	UNP O75417
FFF	?	-	GLY	deletion	UNP O75417
FFF	?	-	LEU	deletion	UNP O75417
FFF	?	-	LEU	deletion	UNP O75417
FFF	?	-	PRO	deletion	UNP O75417
FFF	?	-	MET	deletion	UNP O75417
FFF	?	-	GLY	deletion	UNP O75417
FFF	?	-	ARG	deletion	UNP O75417
FFF	?	-	GLY	deletion	UNP O75417
FFF	?	-	LYS	deletion	UNP O75417
FFF	?	-	TYR	deletion	UNP O75417
FFF	?	-	LYS	deletion	UNP O75417
FFF	?	-	LYS	deletion	UNP O75417
FFF	?	-	GLY	deletion	UNP O75417
FFF	?	-	PHE	deletion	UNP O75417
FFF	?	-	SER	deletion	UNP O75417
FFF	?	-	VAL	deletion	UNP O75417
FFF	?	-	ASN	deletion	UNP O75417
FFF	?	-	PRO	deletion	UNP O75417
FFF	?	-	ARG	deletion	UNP O75417
FFF	?	-	CYS	deletion	UNP O75417
FFF	?	-	GLN	deletion	UNP O75417
FFF	?	-	ALA	deletion	UNP O75417
FFF	?	-	GLN	deletion	UNP O75417
FFF	?	-	MET	deletion	UNP O75417

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
FFF	?	-	GLU	deletion	UNP O75417
FFF	?	-	GLU	deletion	UNP O75417
FFF	?	-	ARG	deletion	UNP O75417
FFF	?	-	ALA	deletion	UNP O75417
FFF	?	-	ALA	deletion	UNP O75417
FFF	?	-	ASP	deletion	UNP O75417
FFF	?	-	ARG	deletion	UNP O75417

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	GGG	16	Total	C	N	O	P	0	0	0
			324	154	59	95	16			
2	III	16	Total	C	N	O	P	0	0	0
			324	154	59	95	16			
2	KKK	16	Total	C	N	O	P	0	0	0
			324	154	59	95	16			
2	MMM	16	Total	C	N	O	P	0	0	0
			324	154	59	95	16			
2	OOO	16	Total	C	N	O	P	0	0	0
			324	154	59	95	16			
2	QQQ	16	Total	C	N	O	P	0	0	0
			324	154	59	95	16			

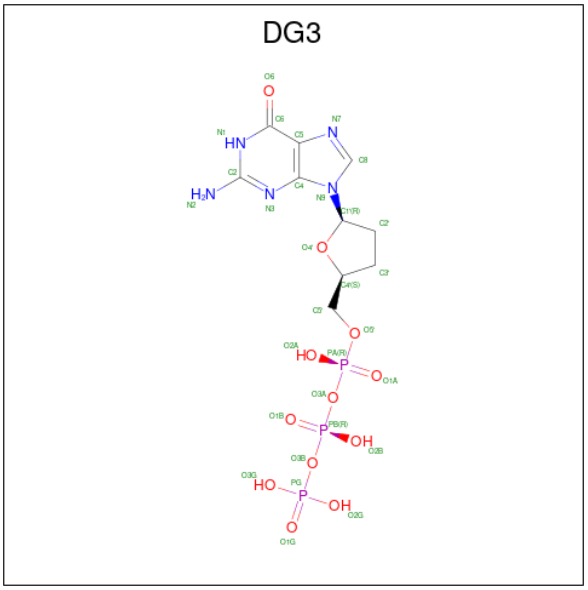
- Molecule 3 is a DNA chain called DNA (5'-D(*GP*CP*GP*GP*CP*TP*GP*TP*CP*AP* TP*TP*(DDG))-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	HHH	13	Total	C	N	O	P	0	0	0
			264	127	47	78	12			
3	JJJ	13	Total	C	N	O	P	0	0	0
			264	127	47	78	12			
3	LLL	13	Total	C	N	O	P	0	0	0
			264	127	47	78	12			
3	NNN	13	Total	C	N	O	P	0	0	0
			264	127	47	78	12			
3	PPP	13	Total	C	N	O	P	0	0	0
			264	127	47	78	12			
3	RRR	13	Total	C	N	O	P	0	0	0
			264	127	47	78	12			

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	AAA	1	Total Mg 1 1	0	0
4	BBB	1	Total Mg 1 1	0	0
4	CCC	1	Total Mg 1 1	0	0
4	DDD	1	Total Mg 1 1	0	0
4	EEE	1	Total Mg 1 1	0	0
4	FFF	1	Total Mg 1 1	0	0

- Molecule 5 is 2'-3'-DIDEOXYGUANOSINE-5'-TRIPHOSPHATE (three-letter code: DG3) (formula: $C_{10}H_{16}N_5O_{12}P_3$) (labeled as "Ligand of Interest" by depositor).



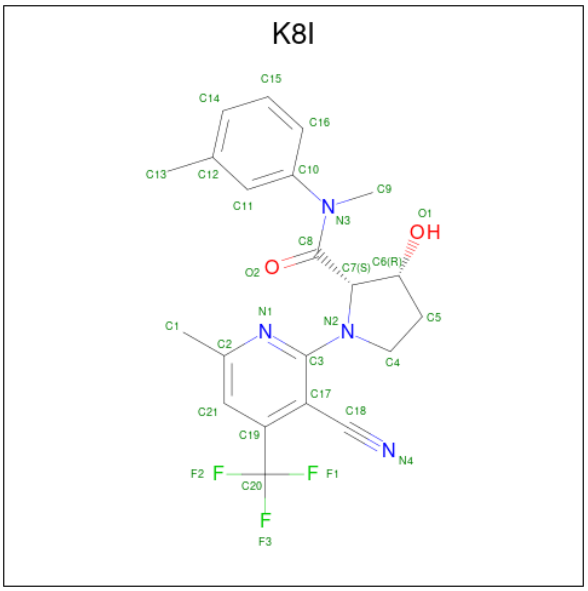
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	AAA	1	Total C N O P 30 10 5 12 3	0	0
5	BBB	1	Total C N O P 30 10 5 12 3	0	0
5	DDD	1	Total C N O P 30 10 5 12 3	0	0
5	EEE	1	Total C N O P 30 10 5 12 3	0	0

Continued on next page...

Continued from previous page...

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

- Molecule 6 is (2 {S},3 {R})-1-[3-cyano-6-methyl-4-(trifluoromethyl)pyridin-2-yl]- {N}-methyl- {N}-(3-methylphenyl)-3-oxidanyl-pyrrolidine-2-carboxamide (three-letter code: K8I) (formula: $C_{21}H_{21}F_3N_4O_2$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	AAA	1	Total	C	F	N	O	0	0
			30	21	3	4	2		
6	BBB	1	Total	C	F	N	O	0	0
			30	21	3	4	2		
6	CCC	1	Total	C	F	N	O	0	0
			30	21	3	4	2		
6	DDD	1	Total	C	F	N	O	0	0
			30	21	3	4	2		

- Molecule 7 is water.

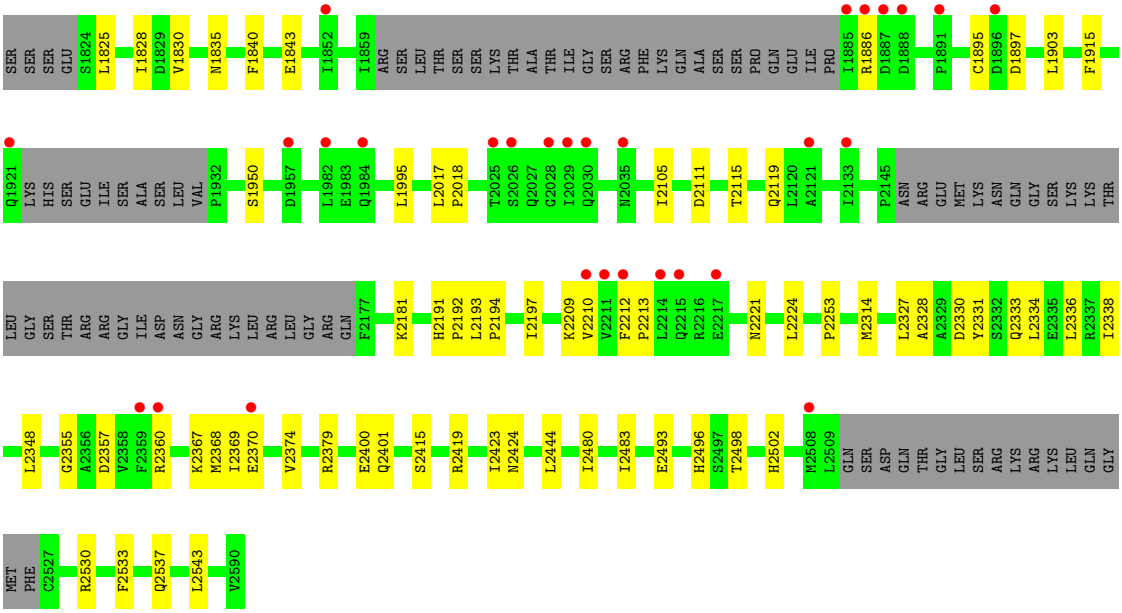
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	AAA	1	Total	O	0	0
			1	1		

3 Residue-property plots (i)

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.

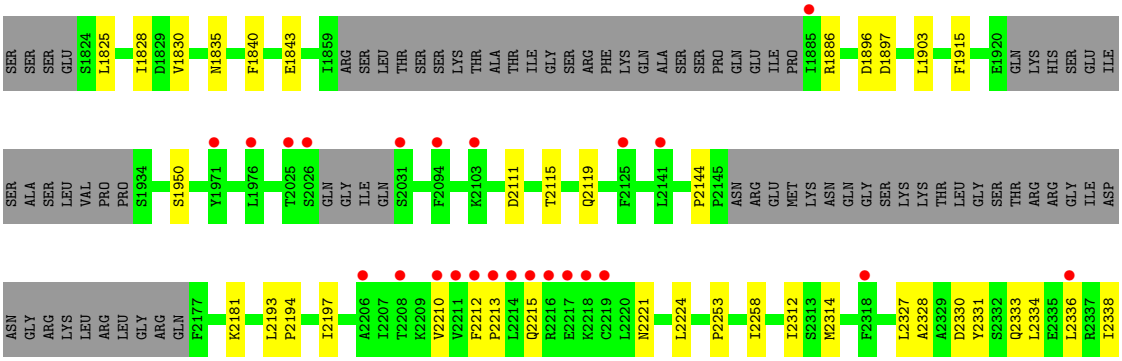
• Molecule 1: DNA polymerase theta

Chain AAA:



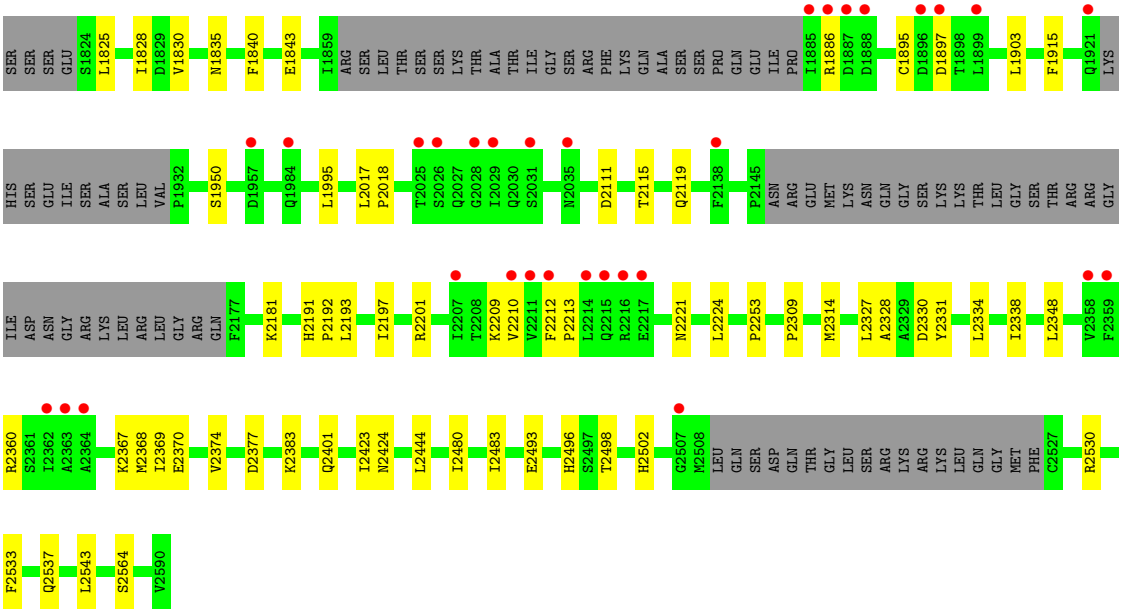
• Molecule 1: DNA polymerase theta

Chain BBB:



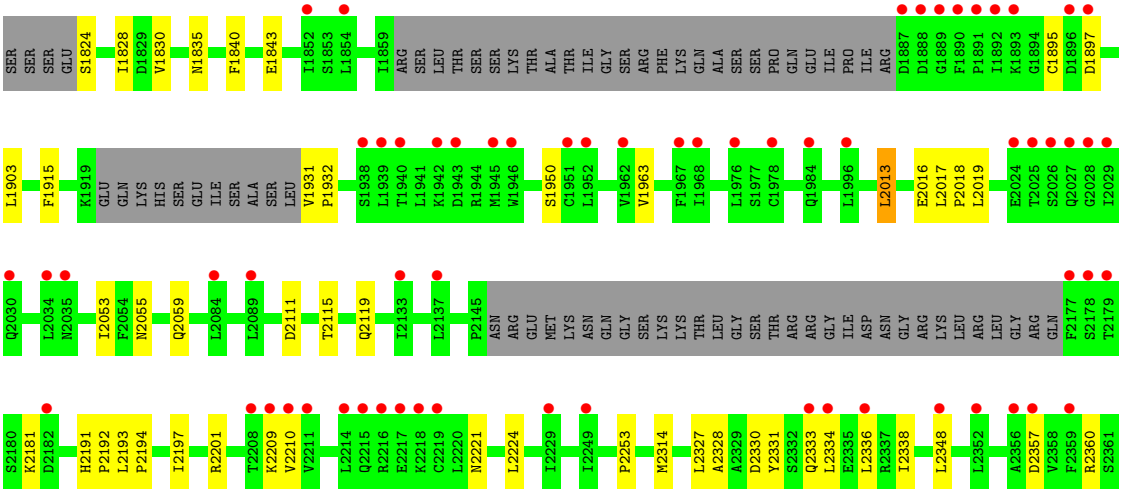
• Molecule 1: DNA polymerase theta

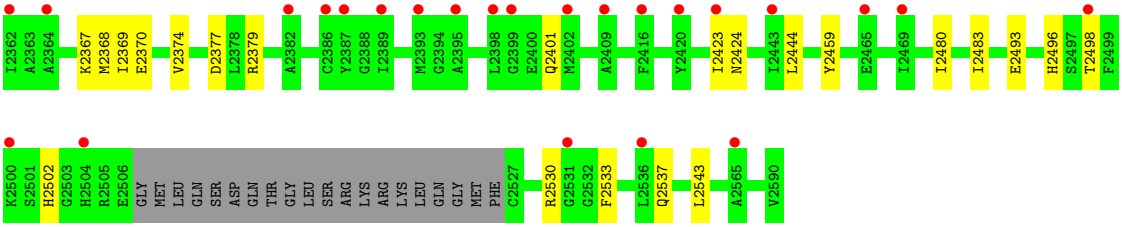
Chain CCC:



• Molecule 1: DNA polymerase theta

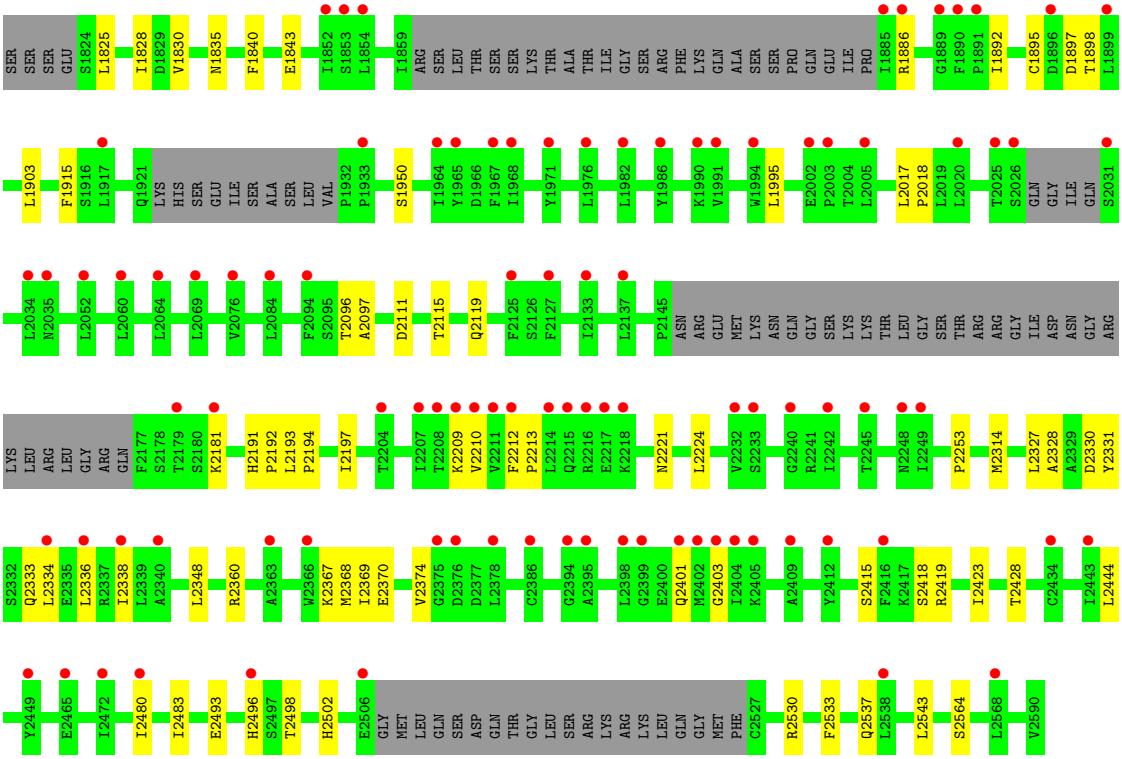
Chain DDD:





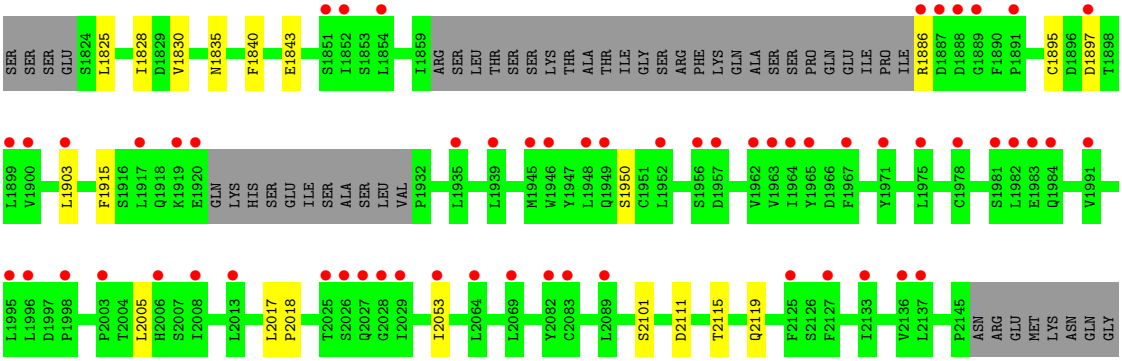
• Molecule 1: DNA polymerase theta

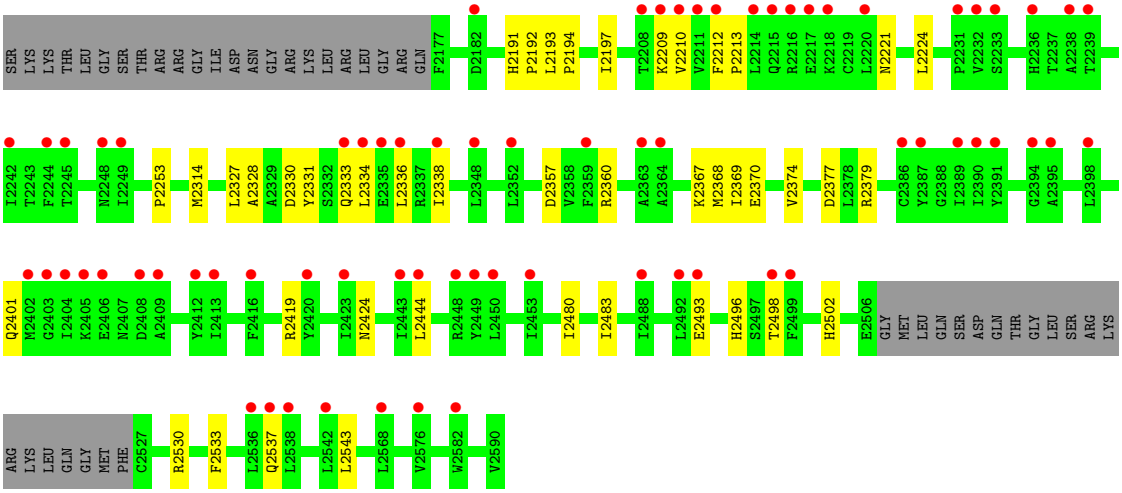
Chain EEE:



• Molecule 1: DNA polymerase theta

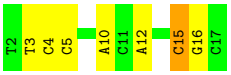
Chain FFF:





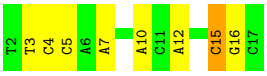
• Molecule 2: DNA (5'-D(P*TP*TP*CP*CP*AP*AP*TP*GP*AP*CP*AP*GP*CP*CP*GP*C) -3')

Chain GGG:



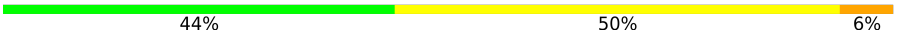
• Molecule 2: DNA (5'-D(P*TP*TP*CP*CP*AP*AP*TP*GP*AP*CP*AP*GP*CP*CP*GP*C) -3')

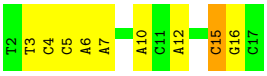
Chain III:



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

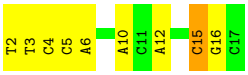
Chain KKK:





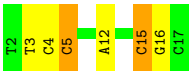
• Molecule 2: DNA (5'-D(P*TP*TP*CP*CP*AP*AP*TP*GP*AP*CP*AP*GP*CP*CP*GP*C) -3')

Chain MMM:



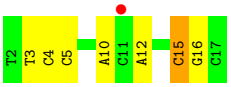
• Molecule 2: DNA (5'-D(P*TP*TP*CP*CP*AP*AP*TP*GP*AP*CP*AP*GP*CP*CP*GP*C) -3')

Chain OOO:



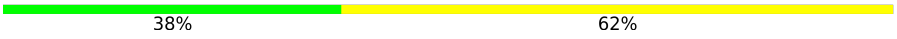
• Molecule 2: DNA (5'-D(P*TP*TP*CP*CP*AP*AP*TP*GP*AP*CP*AP*GP*CP*CP*GP*C) -3')

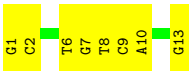
Chain QQQ:



• Molecule 3: DNA (5'-D(*GP*CP*GP*GP*CP*TP*GP*TP*CP*AP*TP*TP*(DDG))-3')

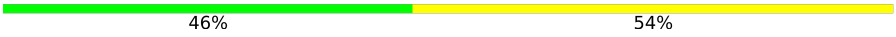
Chain HHH:





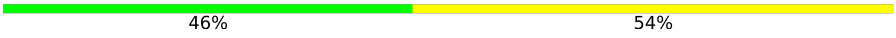
• Molecule 3: DNA (5'-D(*GP*CP*GP*GP*CP*TP*GP*TP*CP*AP*TP*TP*(DDG))-3')

Chain JJJ:



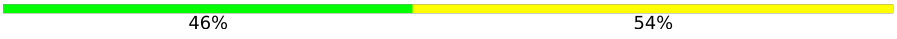
• Molecule 3: DNA (5'-D(*GP*CP*GP*GP*CP*TP*GP*TP*CP*AP*TP*TP*(DDG))-3')

Chain LLL:



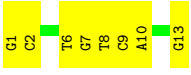
• Molecule 3: DNA (5'-D(*GP*CP*GP*GP*CP*TP*GP*TP*CP*AP*TP*TP*(DDG))-3')

Chain NNN:



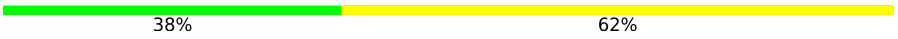
• Molecule 3: DNA (5'-D(*GP*CP*GP*GP*CP*TP*GP*TP*CP*AP*TP*TP*(DDG))-3')

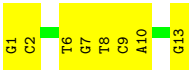
Chain PPP:



• Molecule 3: DNA (5'-D(*GP*CP*GP*GP*CP*TP*GP*TP*CP*AP*TP*TP*(DDG))-3')

Chain RRR:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	59.59Å 172.29Å 289.21Å 90.00° 91.22° 90.00°	Depositor
Resolution (Å)	50.01 – 2.83 110.75 – 2.83	Depositor EDS
% Data completeness (in resolution range)	97.5 (50.01-2.83) 97.5 (110.75-2.83)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.38 (at 2.82Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.228 , 0.260 0.229 , 0.258	Depositor DCC
R_{free} test set	6878 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	79.6	Xtriage
Anisotropy	0.345	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.014 for h,-k,-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	32843	wwPDB-VP
Average B, all atoms (Å^2)	103.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 33.89 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 7.4445e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: DG3, DDG, K8I, MG

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	AAA	0.71	0/4980	0.77	0/6751
1	BBB	0.69	0/4904	0.76	0/6649
1	CCC	0.69	0/4952	0.76	0/6717
1	DDD	0.70	0/4905	0.76	0/6659
1	EEE	0.68	0/4919	0.75	0/6672
1	FFF	0.68	0/4919	0.75	0/6676
2	GGG	1.03	0/362	1.47	6/555 (1.1%)
2	III	1.01	1/362 (0.3%)	1.45	6/555 (1.1%)
2	KKK	1.02	0/362	1.44	8/555 (1.4%)
2	MMM	1.01	0/362	1.44	7/555 (1.3%)
2	OOO	0.95	0/362	1.44	5/555 (0.9%)
2	QQQ	0.95	0/362	1.46	6/555 (1.1%)
3	HHH	1.02	0/271	1.46	7/417 (1.7%)
3	JJJ	1.00	0/271	1.45	6/417 (1.4%)
3	LLL	1.01	0/271	1.47	6/417 (1.4%)
3	NNN	0.97	0/271	1.43	6/417 (1.4%)
3	PPP	0.93	0/271	1.46	7/417 (1.7%)
3	RRR	0.91	0/271	1.45	7/417 (1.7%)
All	All	0.73	1/33377 (0.0%)	0.88	77/45956 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	BBB	0	1
1	FFF	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	III	7	DA	P-O5'	5.32	1.65	1.59

The worst 5 of 77 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	GGG	3	DT	P-O3'-C3'	-9.60	108.18	119.70
3	LLL	7	DG	P-O3'-C3'	-9.16	108.71	119.70
3	NNN	7	DG	P-O3'-C3'	-9.06	108.83	119.70
2	QQQ	3	DT	P-O3'-C3'	-8.99	108.91	119.70
2	III	15	DC	P-O3'-C3'	-8.64	109.33	119.70

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	BBB	2215	GLN	Mainchain
1	FFF	2101	SER	Mainchain

5.2 Too-close contacts (i)

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	AAA	4884	0	4714	31	0
1	BBB	4811	0	4625	28	0
1	CCC	4856	0	4667	27	0
1	DDD	4810	0	4593	37	0
1	EEE	4824	0	4632	32	0
1	FFF	4823	0	4617	28	0
2	GGG	324	0	180	1	0
2	III	324	0	180	1	0
2	KKK	324	0	180	1	0
2	MMM	324	0	180	2	0
2	OOO	324	0	180	3	0
2	QQQ	324	0	180	1	0
3	HHH	264	0	149	0	0
3	JJJ	264	0	149	0	0
3	LLL	264	0	149	0	0
3	NNN	264	0	149	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	PPP	264	0	149	0	0
3	RRR	264	0	149	0	0
4	AAA	1	0	0	0	0
4	BBB	1	0	0	0	0
4	CCC	1	0	0	0	0
4	DDD	1	0	0	0	0
4	EEE	1	0	0	0	0
4	FFF	1	0	0	0	0
5	AAA	30	0	12	0	0
5	BBB	30	0	12	0	0
5	DDD	30	0	12	0	0
5	EEE	30	0	12	2	0
5	FFF	30	0	12	0	0
5	KKK	30	0	12	1	0
6	AAA	30	0	0	2	0
6	BBB	30	0	0	2	0
6	CCC	30	0	0	2	0
6	DDD	30	0	0	1	0
7	AAA	1	0	0	0	0
All	All	32843	0	29894	196	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 3.

The worst 5 of 196 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:DDD:2013:LEU:HD11	1:DDD:2055:ASN:OD1	1.49	1.13
5:EEE:2602:DG3:N2	2:OOO:5:DC:C2	2.53	0.76
1:DDD:2210:VAL:HG21	1:DDD:2253:PRO:HD3	1.71	0.72
1:BBB:2210:VAL:HG21	1:BBB:2253:PRO:HD3	1.73	0.69
1:EEE:2210:VAL:HG21	1:EEE:2253:PRO:HD3	1.75	0.69

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	AAA	629/726 (87%)	612 (97%)	16 (2%)	1 (0%)	47 76
1	BBB	617/726 (85%)	600 (97%)	17 (3%)	0	100 100
1	CCC	628/726 (86%)	607 (97%)	21 (3%)	0	100 100
1	DDD	623/726 (86%)	604 (97%)	19 (3%)	0	100 100
1	EEE	620/726 (85%)	599 (97%)	21 (3%)	0	100 100
1	FFF	624/726 (86%)	604 (97%)	20 (3%)	0	100 100
All	All	3741/4356 (86%)	3626 (97%)	114 (3%)	1 (0%)	100 100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	2355	GLY

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	AAA	502/638 (79%)	483 (96%)	19 (4%)	33 65
1	BBB	494/638 (77%)	475 (96%)	19 (4%)	33 65
1	CCC	494/638 (77%)	473 (96%)	21 (4%)	29 60
1	DDD	491/638 (77%)	472 (96%)	19 (4%)	32 64
1	EEE	493/638 (77%)	473 (96%)	20 (4%)	30 63
1	FFF	490/638 (77%)	471 (96%)	19 (4%)	32 64

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2964/3828 (77%)	2847 (96%)	117 (4%)	32 64

5 of 117 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	CCC	2537	GLN
1	FFF	2330	ASP
1	DDD	2314	MET
1	FFF	2314	MET
1	FFF	1835	ASN

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

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

6 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	$\#\|Z\| > 2$
3	DDG	NNN	13	3,2	17,23,24	1.18	1 (5%)	15,33,36	1.32	3 (20%)
3	DDG	PPP	13	3,2	17,23,24	1.20	2 (11%)	15,33,36	1.45	3 (20%)
3	DDG	JJJ	13	3,2	17,23,24	1.18	1 (5%)	15,33,36	1.36	3 (20%)
3	DDG	HHH	13	3,2	17,23,24	1.36	3 (17%)	15,33,36	1.39	3 (20%)
3	DDG	LLL	13	3,2	17,23,24	1.31	1 (5%)	15,33,36	1.71	4 (26%)
3	DDG	RRR	13	3,2	17,23,24	1.22	1 (5%)	15,33,36	1.43	3 (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
3	DDG	NNN	13	3,2	-	2/3/18/19	0/3/3/3
3	DDG	PPP	13	3,2	-	2/3/18/19	0/3/3/3
3	DDG	JJJ	13	3,2	-	2/3/18/19	0/3/3/3
3	DDG	HHH	13	3,2	-	2/3/18/19	0/3/3/3
3	DDG	LLL	13	3,2	-	2/3/18/19	0/3/3/3
3	DDG	RRR	13	3,2	-	2/3/18/19	0/3/3/3

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	RRR	13	DDG	C6-N1	-3.72	1.32	1.37
3	LLL	13	DDG	C6-N1	-3.48	1.32	1.37
3	PPP	13	DDG	C6-N1	-3.37	1.32	1.37
3	JJJ	13	DDG	C6-N1	-3.30	1.33	1.37
3	NNN	13	DDG	C6-N1	-2.86	1.33	1.37

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	LLL	13	DDG	O4'-C1'-C2'	3.06	109.99	106.67
3	NNN	13	DDG	C8-N7-C5	2.83	108.38	102.99
3	LLL	13	DDG	C8-N7-C5	2.82	108.36	102.99
3	JJJ	13	DDG	C8-N7-C5	2.72	108.18	102.99
3	PPP	13	DDG	C8-N7-C5	2.69	108.12	102.99

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	HHH	13	DDG	O4'-C4'-C5'-O5'
3	JJJ	13	DDG	O4'-C4'-C5'-O5'
3	LLL	13	DDG	O4'-C4'-C5'-O5'
3	NNN	13	DDG	O4'-C4'-C5'-O5'
3	PPP	13	DDG	O4'-C4'-C5'-O5'

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 16 ligands modelled in this entry, 6 are monoatomic - leaving 10 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	AAA	2602	4	25,32,32	0.96	1 (4%)	28,50,50	0.98	2 (7%)
5	DG3	DDD	2602	4	25,32,32	1.00	2 (8%)	28,50,50	1.10	2 (7%)
6	K8I	BBB	2603	-	31,32,32	0.84	1 (3%)	39,48,48	0.68	1 (2%)
6	K8I	CCC	2602	-	31,32,32	0.88	1 (3%)	39,48,48	0.75	2 (5%)
6	K8I	DDD	2603	-	31,32,32	0.92	1 (3%)	39,48,48	0.83	2 (5%)
5	DG3	BBB	2602	4	25,32,32	1.01	2 (8%)	28,50,50	1.23	4 (14%)
5	DG3	FFF	2602	4	25,32,32	0.95	2 (8%)	28,50,50	1.14	2 (7%)
5	DG3	KKK	101	4	25,32,32	1.02	3 (12%)	28,50,50	0.99	1 (3%)
6	K8I	AAA	2603	-	31,32,32	0.97	1 (3%)	39,48,48	0.84	1 (2%)
5	DG3	EEE	2602	4	25,32,32	0.98	1 (4%)	28,50,50	1.42	3 (10%)

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	AAA	2602	4	-	5/18/31/31	0/3/3/3
5	DG3	DDD	2602	4	-	7/18/31/31	0/3/3/3
6	K8I	BBB	2603	-	-	0/22/37/37	0/3/3/3
6	K8I	CCC	2602	-	-	1/22/37/37	0/3/3/3
6	K8I	DDD	2603	-	-	0/22/37/37	0/3/3/3
5	DG3	BBB	2602	4	-	3/18/31/31	0/3/3/3
5	DG3	FFF	2602	4	-	5/18/31/31	0/3/3/3

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	DG3	KKK	101	4	-	6/18/31/31	0/3/3/3
6	K8I	AAA	2603	-	-	0/22/37/37	0/3/3/3
5	DG3	EEE	2602	4	-	2/18/31/31	0/3/3/3

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	AAA	2603	K8I	C10-N3	-4.32	1.36	1.43
6	CCC	2602	K8I	C10-N3	-3.85	1.37	1.43
6	DDD	2603	K8I	C10-N3	-3.81	1.37	1.43
6	BBB	2603	K8I	C10-N3	-3.72	1.37	1.43
5	KKK	101	DG3	C5-C6	-3.16	1.41	1.47

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	EEE	2602	DG3	PA-O3A-PB	-4.33	117.96	132.83
6	DDD	2603	K8I	C17-C18-N4	-3.30	169.92	177.40
5	BBB	2602	DG3	PA-O3A-PB	-3.15	122.03	132.83
5	FFF	2602	DG3	PA-O3A-PB	-2.88	122.96	132.83
5	EEE	2602	DG3	PB-O3B-PG	-2.59	123.95	132.83

There are no chirality outliers.

5 of 29 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	FFF	2602	DG3	PB-O3B-PG-O2G
5	KKK	101	DG3	C5'-O5'-PA-O2A
5	DDD	2602	DG3	PB-O3B-PG-O1G
5	AAA	2602	DG3	PG-O3B-PB-O1B
5	DDD	2602	DG3	PG-O3B-PB-O1B

There are no ring outliers.

6 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	BBB	2603	K8I	2	0
6	CCC	2602	K8I	2	0
6	DDD	2603	K8I	1	0
5	KKK	101	DG3	1	0

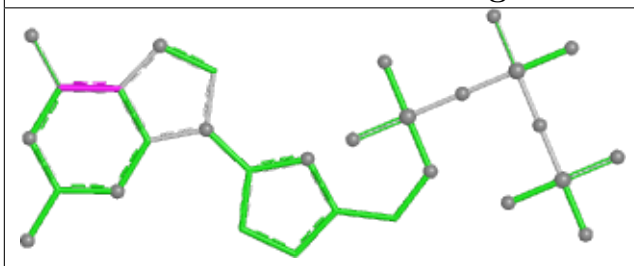
Continued on next page...

Continued from previous page...

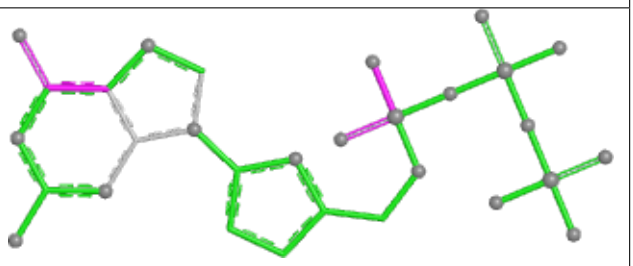
Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	AAA	2603	K8I	2	0
5	EEE	2602	DG3	2	0

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.

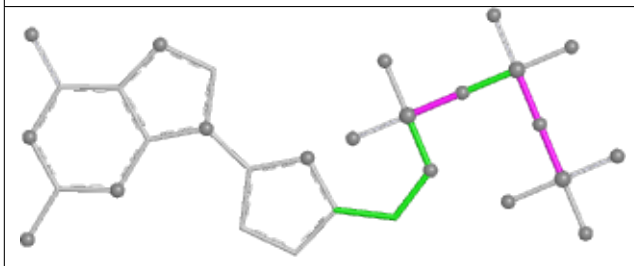
Ligand DG3 AAA 2602



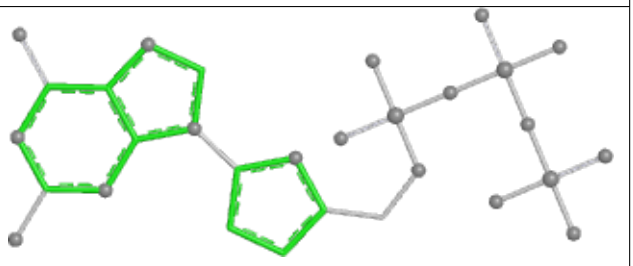
Bond lengths



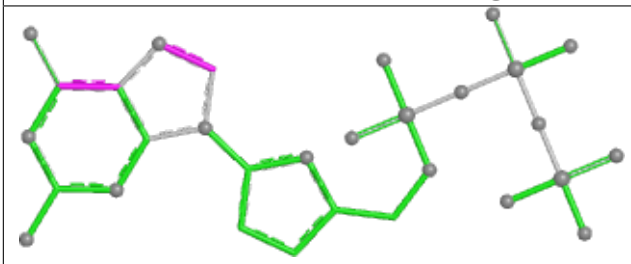
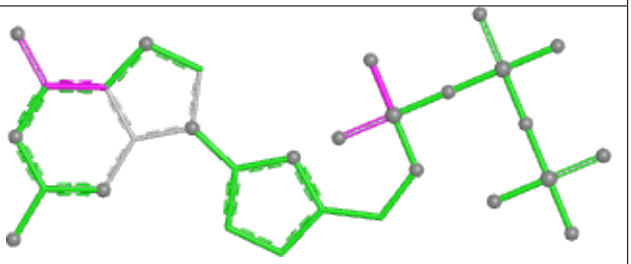
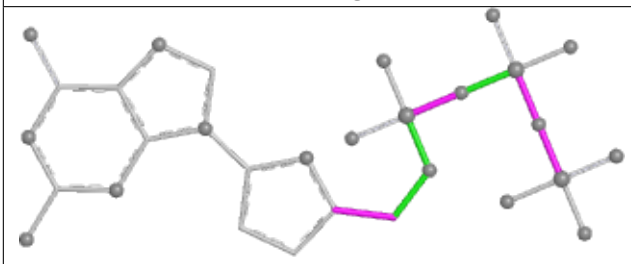
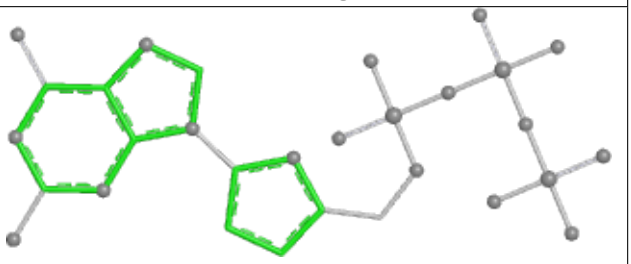
Bond angles

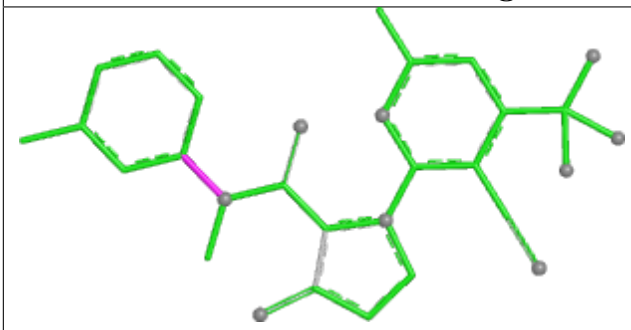
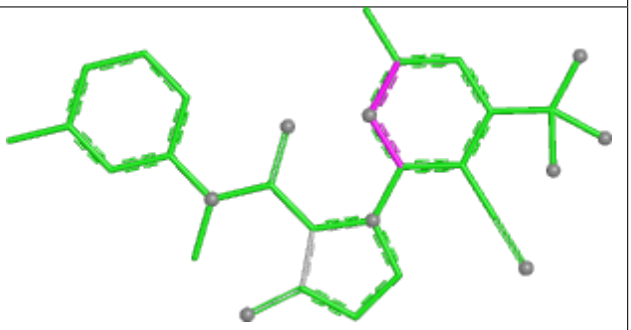
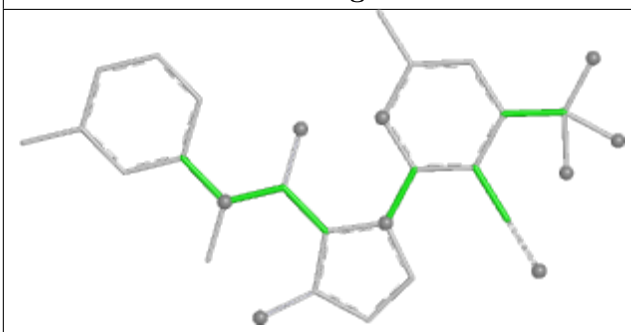
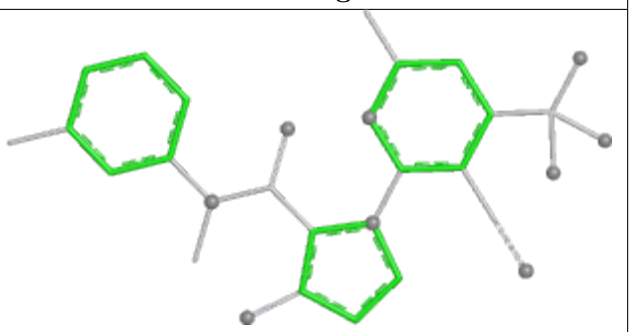


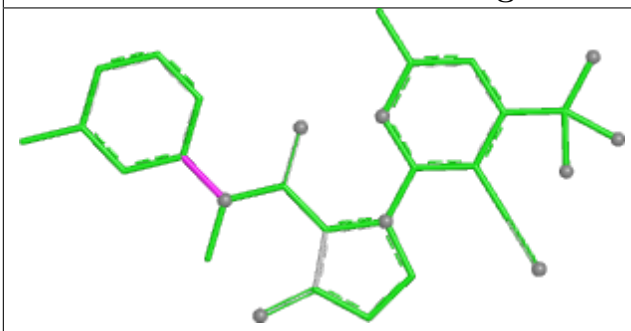
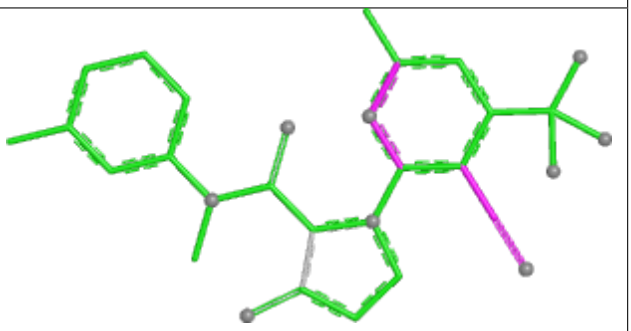
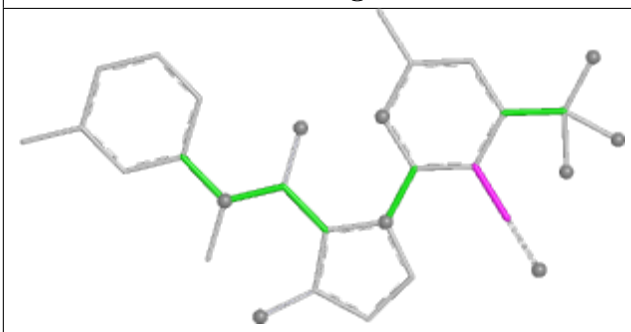
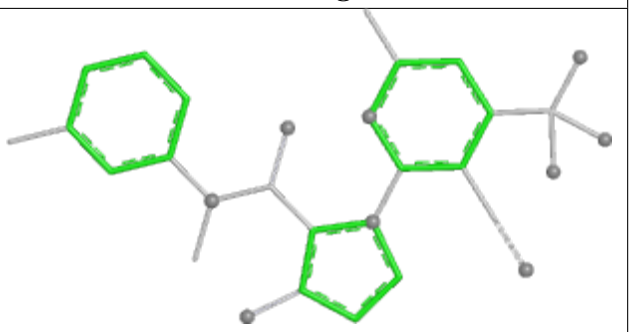
Torsions

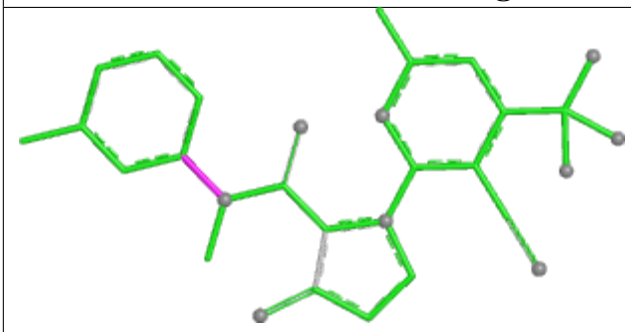
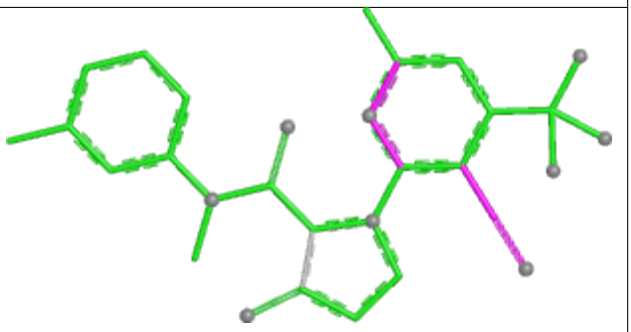
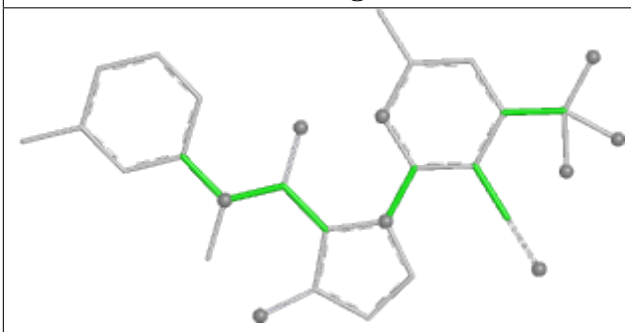
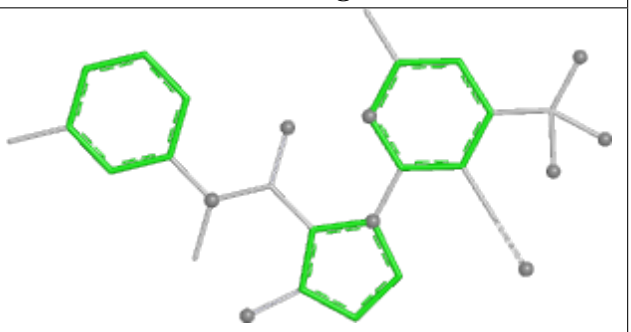


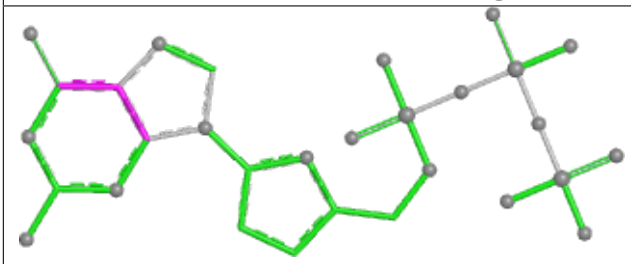
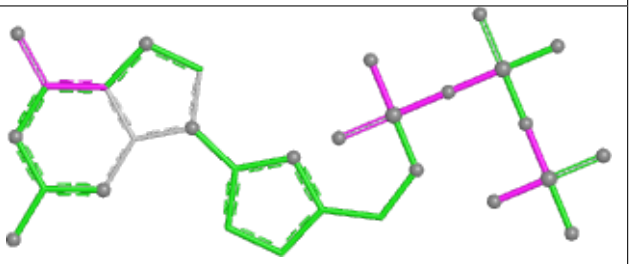
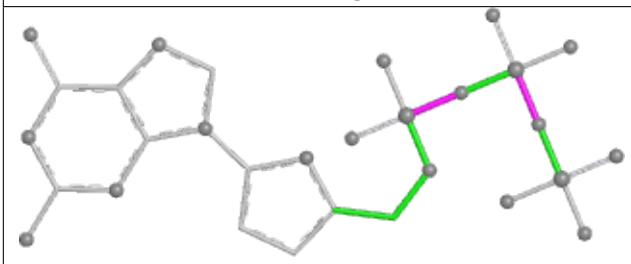
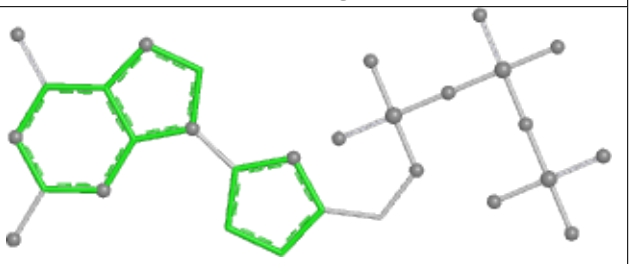
Rings

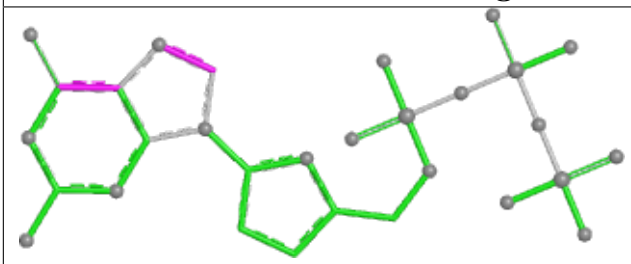
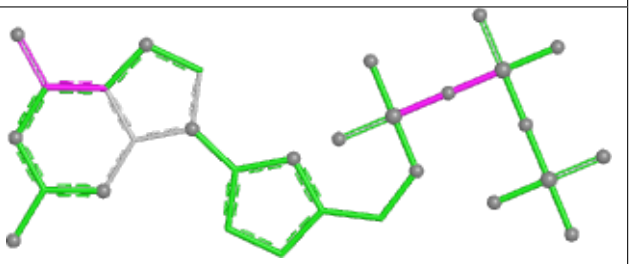
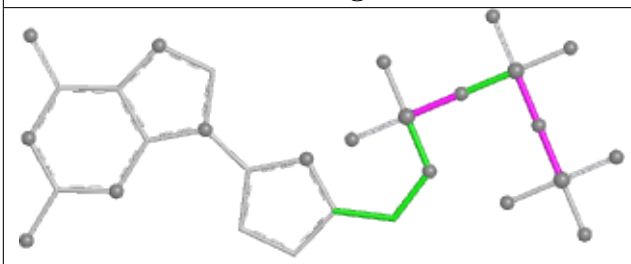
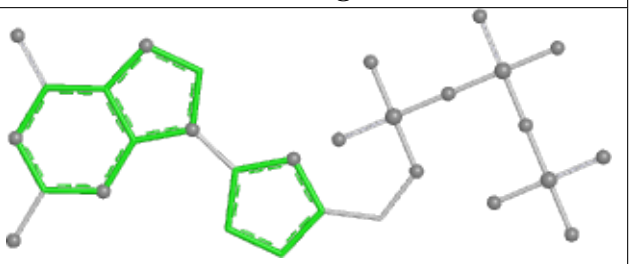
Ligand DG3 DDD 2602	
 Bond lengths	 Bond angles
 Torsions	 Rings

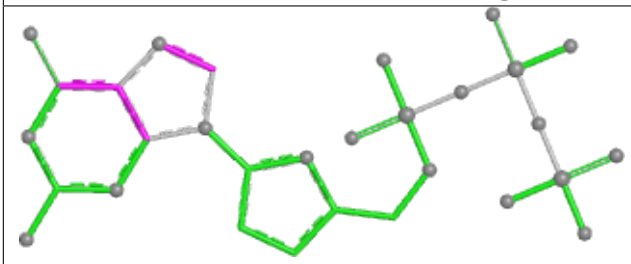
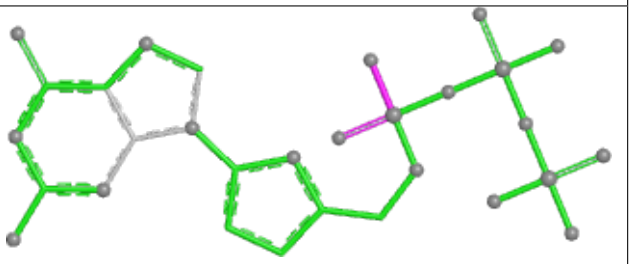
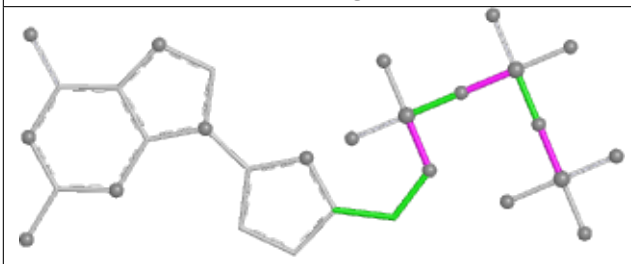
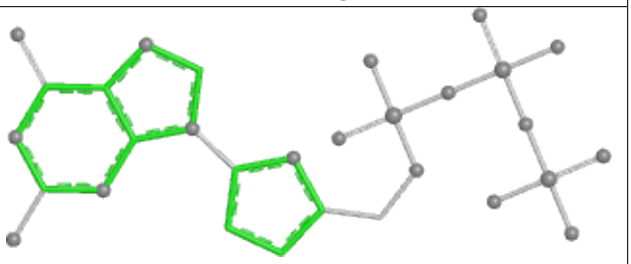
Ligand K8I BBB 2603	
 Bond lengths	 Bond angles
 Torsions	 Rings

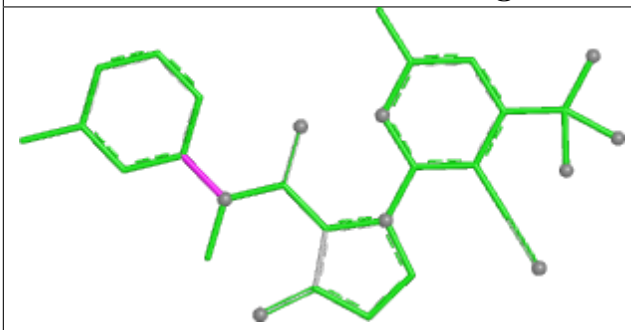
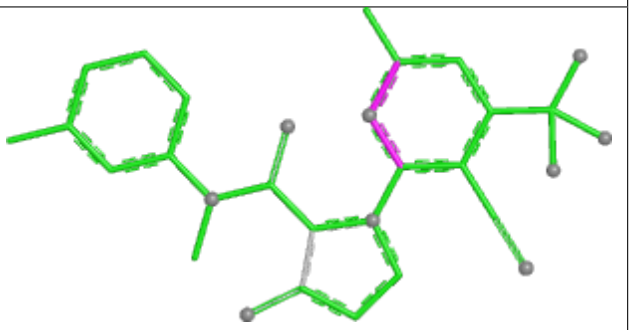
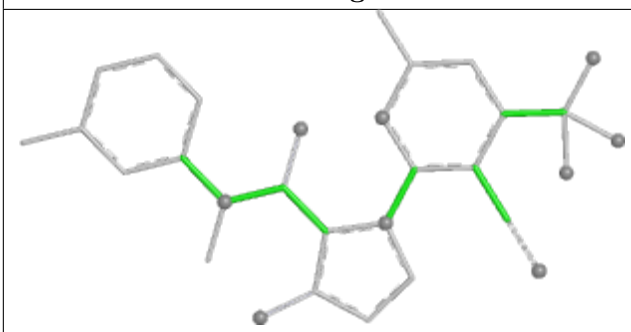
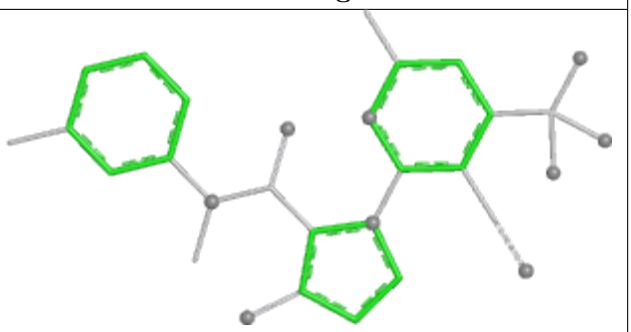
Ligand K8I CCC 2602	
 Bond lengths	 Bond angles
 Torsions	 Rings

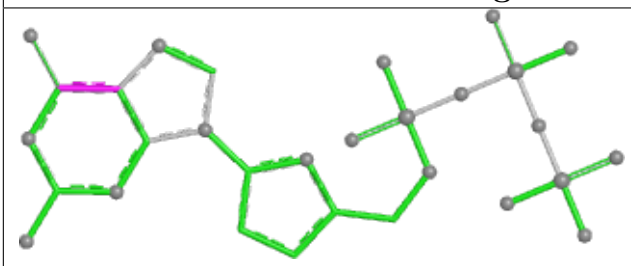
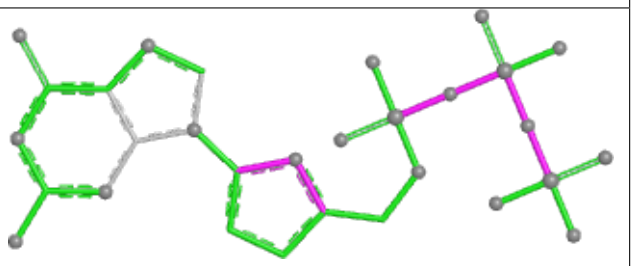
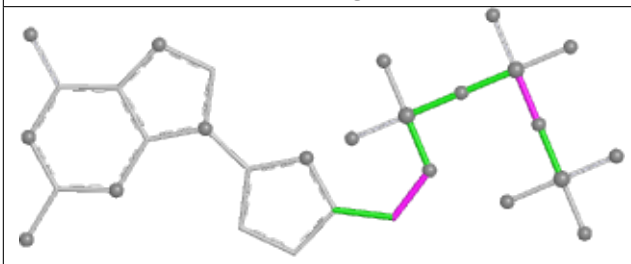
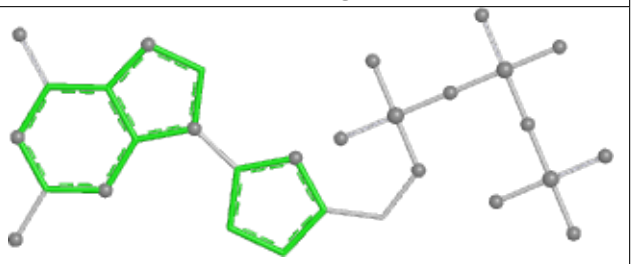
Ligand K8I DDD 2603	
 Bond lengths	 Bond angles
 Torsions	 Rings

Ligand DG3 BBB 2602	
 Bond lengths	 Bond angles
 Torsions	 Rings

Ligand DG3 FFF 2602	
 Bond lengths	 Bond angles
 Torsions	 Rings

Ligand DG3 KKK 101	
 Bond lengths	 Bond angles
 Torsions	 Rings

Ligand K8I AAA 2603	
 Bond lengths	 Bond angles
 Torsions	 Rings

Ligand DG3 EEE 2602	
 Bond lengths	 Bond angles
 Torsions	 Rings

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

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, 95^{th} 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	AAA	639/726 (88%)	0.47	29 (4%) 33 23	39, 74, 135, 198	0
1	BBB	629/726 (86%)	0.49	33 (5%) 27 18	45, 84, 140, 194	0
1	CCC	638/726 (87%)	0.45	31 (4%) 29 20	45, 87, 145, 197	0
1	DDD	633/726 (87%)	0.81	88 (13%) 2 1	56, 103, 176, 219	0
1	EEE	632/726 (87%)	0.80	96 (15%) 2 1	62, 114, 169, 208	0
1	FFF	634/726 (87%)	0.97	131 (20%) 1 0	73, 129, 177, 227	0
2	GGG	16/16 (100%)	-0.22	0 100 100	43, 94, 155, 158	0
2	III	16/16 (100%)	-0.12	0 100 100	60, 112, 173, 177	0
2	KKK	16/16 (100%)	-0.17	0 100 100	54, 110, 172, 174	0
2	MMM	16/16 (100%)	-0.05	0 100 100	61, 100, 168, 179	0
2	OOO	16/16 (100%)	-0.14	0 100 100	93, 148, 225, 232	0
2	QQQ	16/16 (100%)	0.01	1 (6%) 20 12	93, 154, 226, 237	0
3	HHH	12/13 (92%)	-0.24	0 100 100	41, 100, 155, 160	0
3	JJJ	12/13 (92%)	-0.25	0 100 100	50, 134, 160, 176	0
3	LLL	12/13 (92%)	-0.35	0 100 100	58, 107, 160, 162	0
3	NNN	12/13 (92%)	-0.15	0 100 100	57, 117, 153, 155	0
3	PPP	12/13 (92%)	-0.35	0 100 100	74, 171, 238, 248	0
3	RRR	12/13 (92%)	0.11	0 100 100	83, 168, 218, 220	0
All	All	3973/4530 (87%)	0.63	409 (10%) 6 3	39, 100, 168, 248	0

The worst 5 of 409 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	EEE	2031	SER	14.3
1	DDD	2026	SER	12.2
1	DDD	2028	GLY	12.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	BBB	2212	PHE	9.3
1	FFF	2212	PHE	9.2

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} 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	DDG	RRR	13	21/22	0.95	0.17	72,85,98,100	0
3	DDG	NNN	13	21/22	0.96	0.23	48,55,66,69	0
3	DDG	PPP	13	21/22	0.97	0.19	61,76,95,97	0
3	DDG	LLL	13	21/22	0.97	0.18	42,51,60,62	0
3	DDG	HHH	13	21/22	0.98	0.22	30,38,43,44	0
3	DDG	JJJ	13	21/22	0.99	0.19	43,51,55,57	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} 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	AAA	2601	1/1	0.62	0.11	59,59,59,59	0
4	MG	EEE	2601	1/1	0.83	0.07	99,99,99,99	0
6	K8I	DDD	2603	30/30	0.85	0.51	94,114,136,144	0
4	MG	DDD	2601	1/1	0.87	0.07	84,84,84,84	0
4	MG	CCC	2601	1/1	0.91	0.12	59,59,59,59	0
4	MG	BBB	2601	1/1	0.91	0.15	66,66,66,66	0
6	K8I	CCC	2602	30/30	0.92	0.34	80,95,111,117	0
6	K8I	BBB	2603	30/30	0.93	0.35	78,93,115,119	0
6	K8I	AAA	2603	30/30	0.94	0.30	71,95,113,119	0
4	MG	FFF	2601	1/1	0.94	0.05	111,111,111,111	0

Continued on next page...

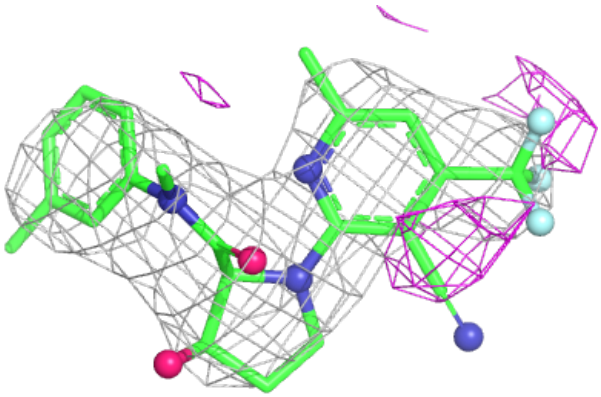
Continued from previous page...

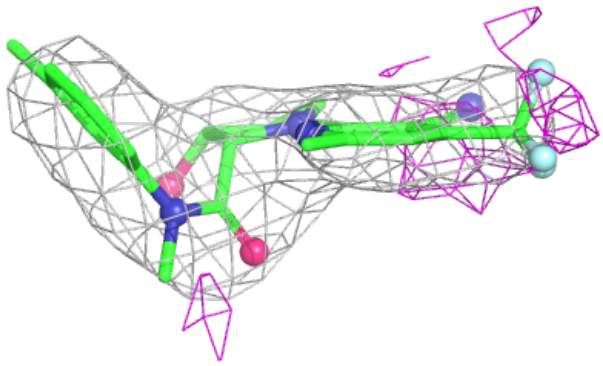
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	DG3	FFF	2602	30/30	0.95	0.17	89,102,119,121	0
5	DG3	EEE	2602	30/30	0.96	0.16	71,91,108,109	0
5	DG3	DDD	2602	30/30	0.96	0.17	51,63,86,93	0
5	DG3	BBB	2602	30/30	0.97	0.20	50,62,73,75	0
5	DG3	KKK	101	30/30	0.97	0.17	41,49,83,91	0
5	DG3	AAA	2602	30/30	0.98	0.18	39,44,64,76	0

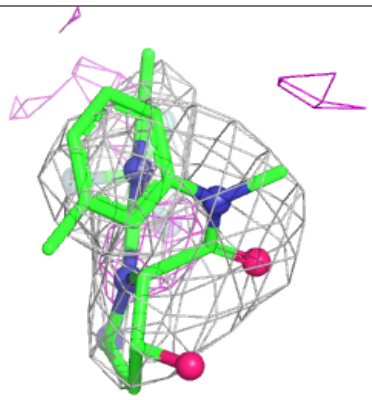
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around K8I DDD 2603:

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

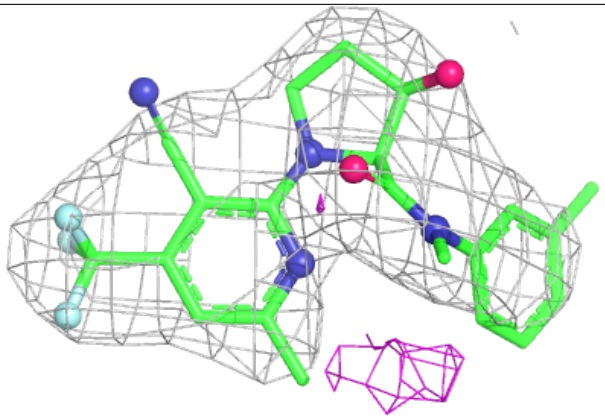


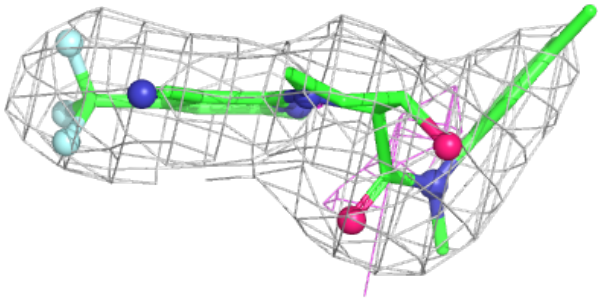


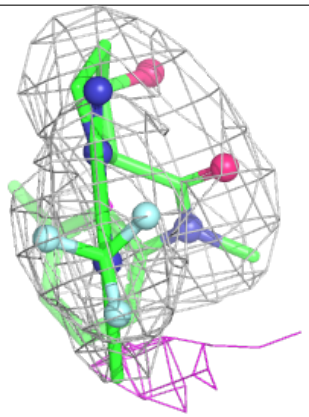


Electron density around K8I CCC 2602:

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

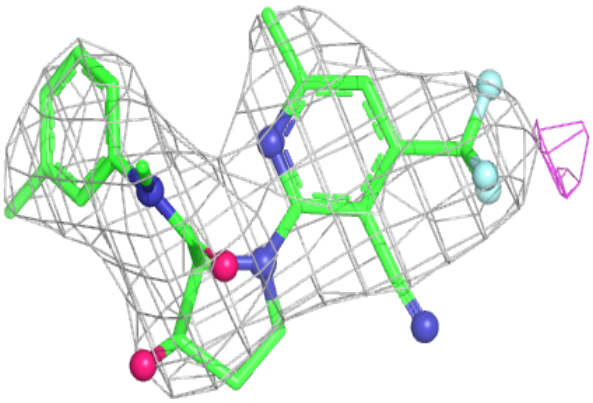


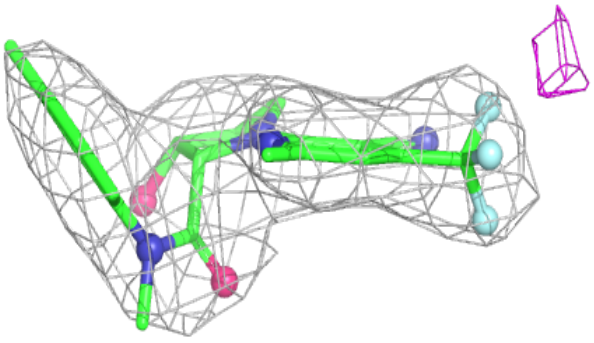


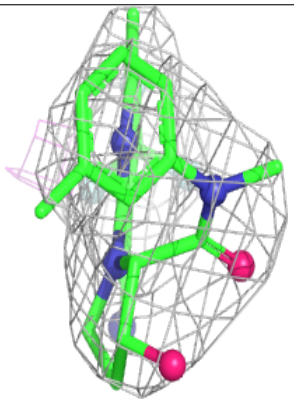


Electron density around K8I BBB 2603:

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

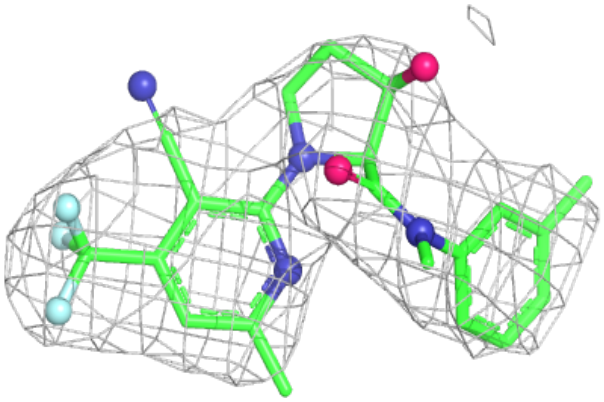


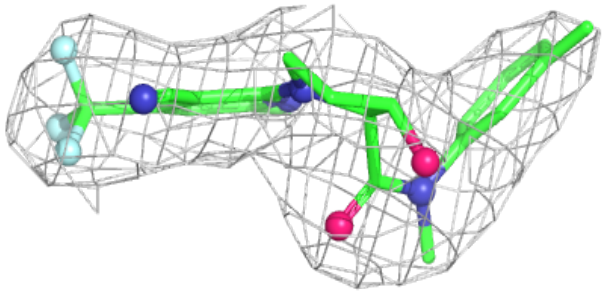


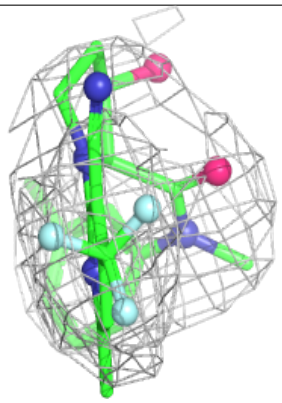


Electron density around K8I AAA 2603:

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

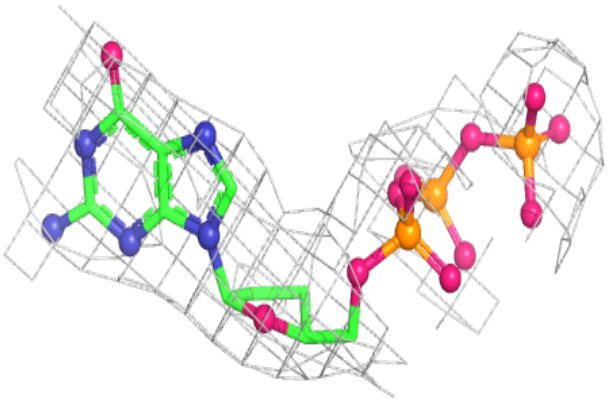


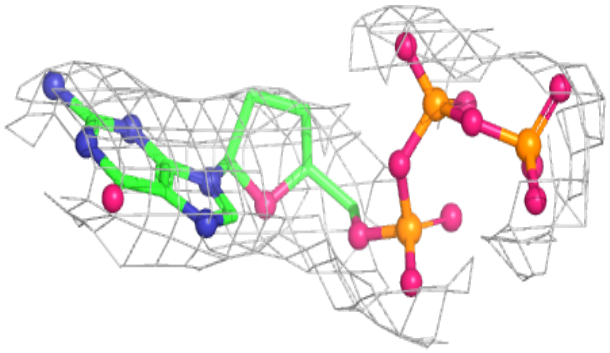


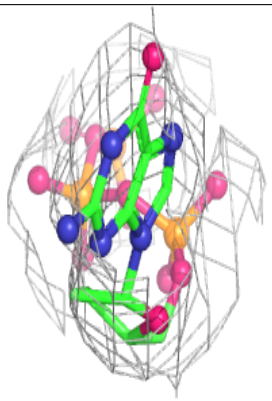


Electron density around DG3 FFF 2602:

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

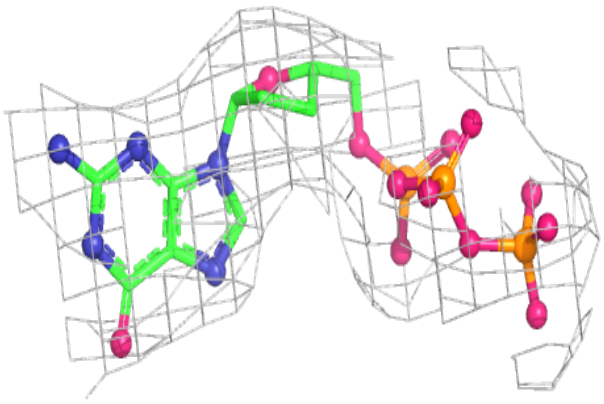


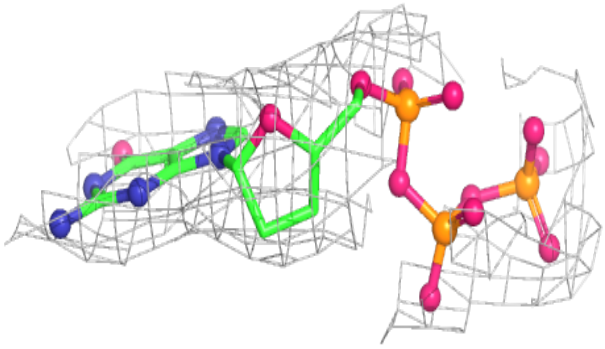


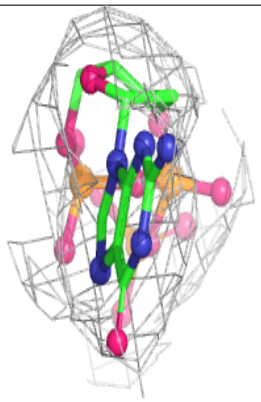


Electron density around DG3 EEE 2602:

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

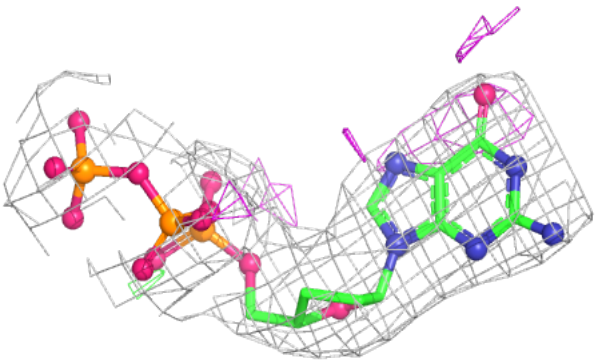


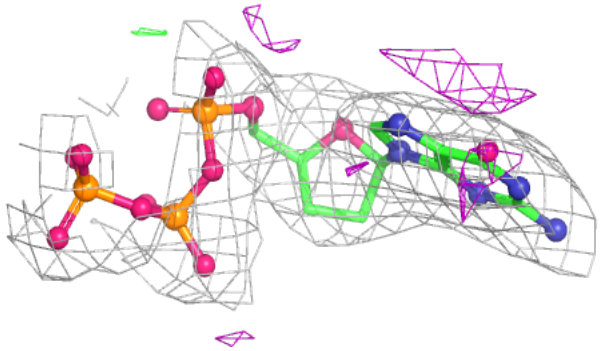


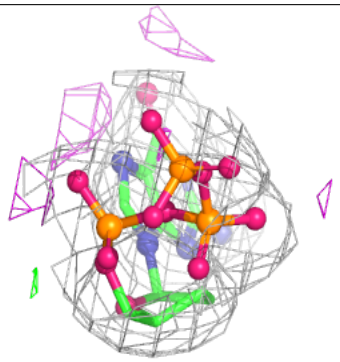


Electron density around DG3 DDD 2602:

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

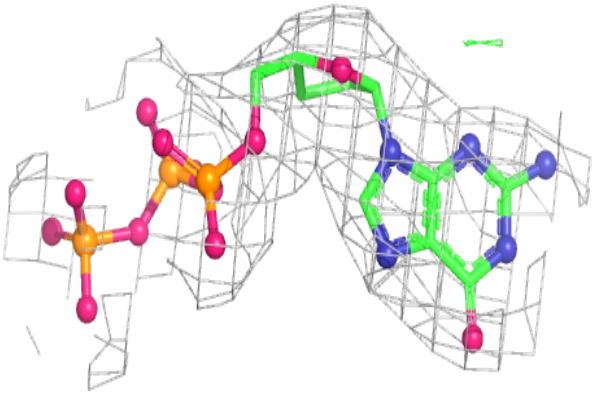


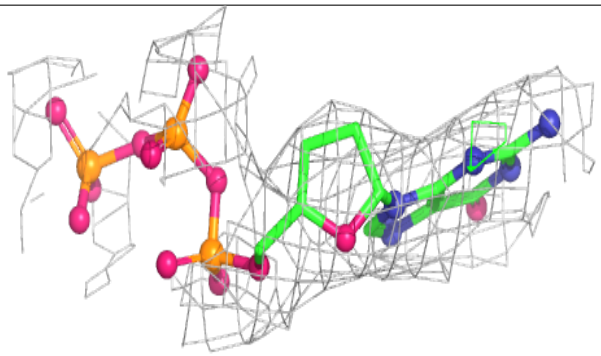


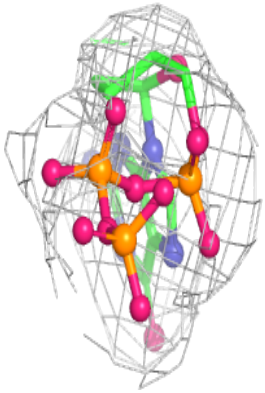


Electron density around DG3 BBB 2602:

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

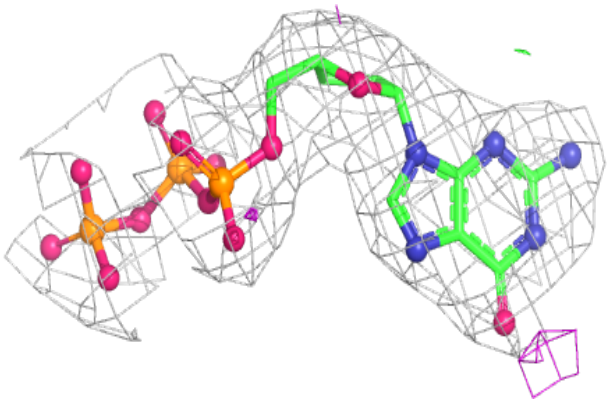


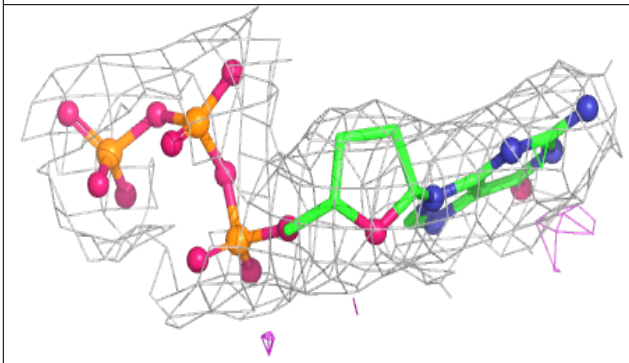


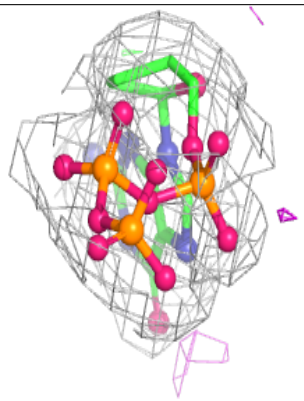


Electron density around DG3 KKK 101:

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

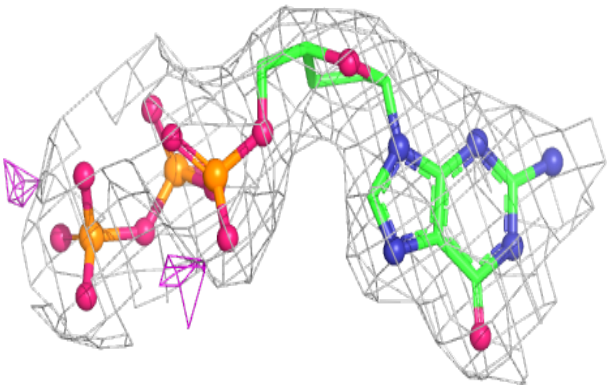


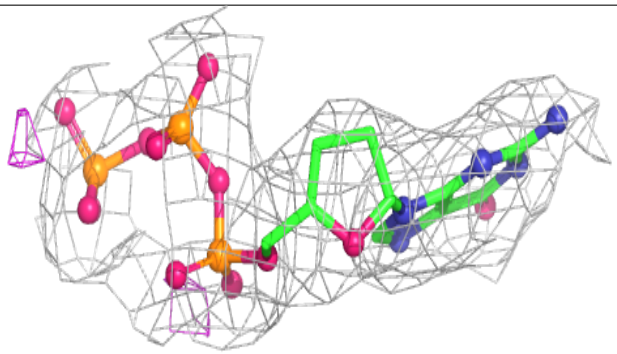


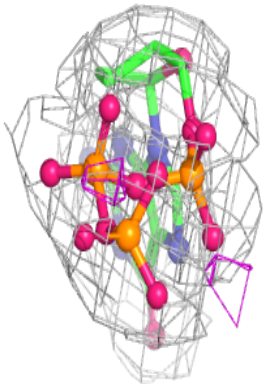


Electron density around DG3 AAA 2602:

$2mF_o$-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.