wwPDB X-ray Structure Validation Summary Report (i)

Aug 8, 2023 – 02:51 PM EDT

PDB ID	:	1Q9X
Title	:	Crystal structure of Enterobacteria phage RB69 gp43 DNA polymerase complexed with tetrahydrofuran containing DNA
Authors	:	Freisinger, E.; Grollman, A.P.; Miller, H.; Kisker, C.
Deposited on	:	2003-08-26
Resolution	:	2.69 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:

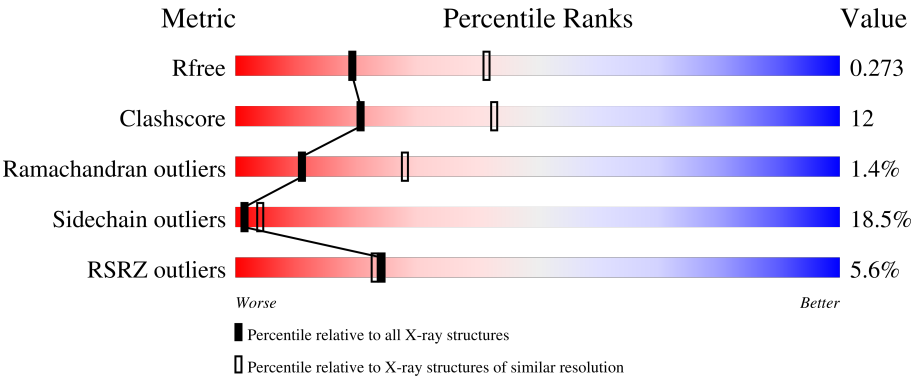
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 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 2.69 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	E	18	28% 22% 33% 44%
1	F	18	17% 33% 33% 33%
1	G	18	11% 28% 33% 39%
1	H	18	6% 33% 44% 22%
2	I	13	8% 54% 38%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	J	13	31% 38% 31%
2	K	13	8% 62% 31%
2	L	13	46% 38% 15%
3	A	903	11% 60% 32% 7%
3	B	903	7% 64% 29% 6%
3	C	903	3% 67% 25% 7%
3	D	903	2% 63% 29% 7%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	DGP	F	908	-	-	-	X
6	DGP	G	908	-	-	-	X
6	DGP	K	955	-	-	-	X

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 32756 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 DNA chain called 5'-GCGGACTGCTTAC(dideoxycytidine)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	18	Total	C	N	O	P	0	0	0
			370	175	74	104	17			
1	F	18	Total	C	N	O	P	0	0	0
			370	175	74	104	17			
1	G	18	Total	C	N	O	P	0	0	0
			370	175	74	104	17			
1	H	18	Total	C	N	O	P	0	0	0
			370	175	74	104	17			

- Molecule 2 is a DNA chain called 5'-AC(tetrahydrofuran)GGTAAGCAGTCCGCGG-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	13	Total	C	N	O	P	0	0	0
			263	126	48	77	12			
2	J	13	Total	C	N	O	P	0	0	0
			263	126	48	77	12			
2	K	13	Total	C	N	O	P	0	0	0
			263	126	48	77	12			
2	L	13	Total	C	N	O	P	0	0	0
			263	126	48	77	12			

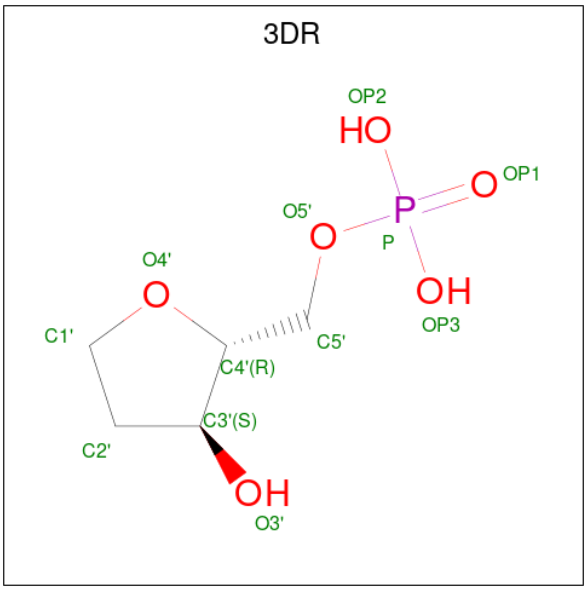
- Molecule 3 is a protein called DNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	903	Total	C	N	O	S	0	0	0
			7365	4730	1226	1376	33			
3	B	903	Total	C	N	O	S	0	0	0
			7365	4730	1226	1376	33			
3	C	903	Total	C	N	O	S	0	0	0
			7365	4730	1226	1376	33			
3	D	903	Total	C	N	O	S	0	0	0
			7365	4730	1226	1376	33			

There are 8 discrepancies between the modelled and reference sequences:

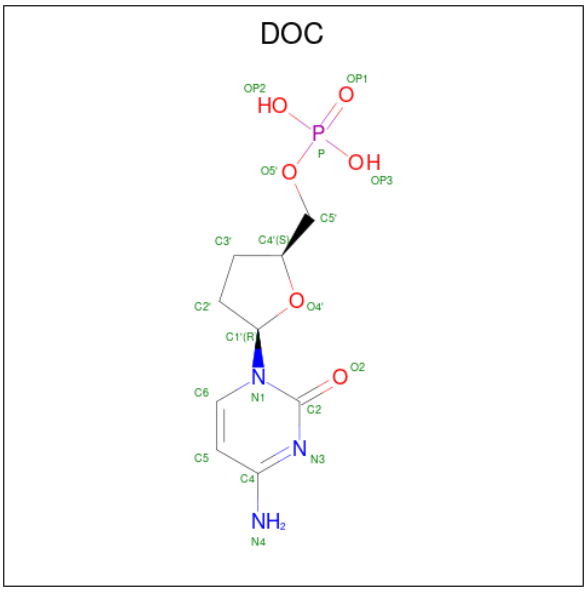
Chain	Residue	Modelled	Actual	Comment	Reference
A	222	ALA	ASP	engineered mutation	UNP Q38087
A	327	ALA	ASP	engineered mutation	UNP Q38087
B	222	ALA	ASP	engineered mutation	UNP Q38087
B	327	ALA	ASP	engineered mutation	UNP Q38087
C	222	ALA	ASP	engineered mutation	UNP Q38087
C	327	ALA	ASP	engineered mutation	UNP Q38087
D	222	ALA	ASP	engineered mutation	UNP Q38087
D	327	ALA	ASP	engineered mutation	UNP Q38087

- Molecule 4 is 1',2'-DIDEOXYRIBOFURANOSE-5'-PHOSPHATE (three-letter code: 3DR) (formula: $C_5H_{11}O_6P$).



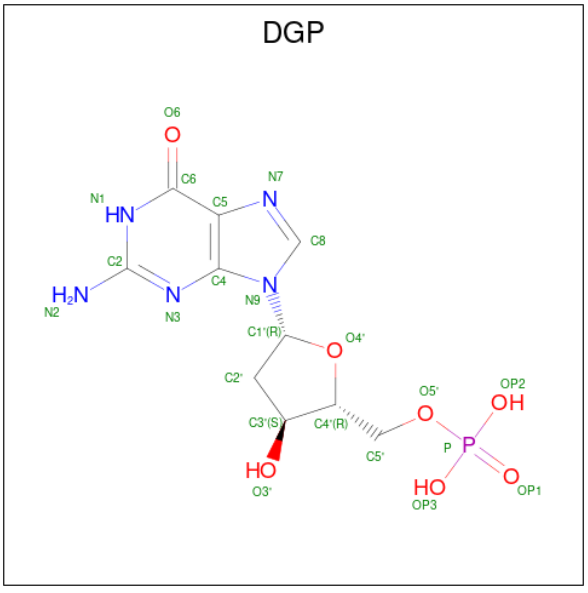
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	E	1	Total	C	O	P	0	0
			11	5	5	1		
4	F	1	Total	C	O	P	0	0
			11	5	5	1		
4	G	1	Total	C	O	P	0	0
			11	5	5	1		
4	H	1	Total	C	O	P	0	0
			11	5	5	1		

- Molecule 5 is 2',3'-DIDEOXYCYTIDINE-5'-MONOPHOSPHATE (three-letter code: DOC) (formula: $C_9H_{14}N_3O_6P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	I	1	Total	C	N	O	P	0	0
			18	9	3	5	1		
5	J	1	Total	C	N	O	P	0	0
			18	9	3	5	1		
5	K	1	Total	C	N	O	P	0	0
			18	9	3	5	1		
5	L	1	Total	C	N	O	P	0	0
			18	9	3	5	1		

- Molecule 6 is 2'-DEOXYGUANOSINE-5'-MONOPHOSPHATE (three-letter code: DGP) (formula: $C_{10}H_{14}N_5O_7P$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	F	1	Total C N O 19 10 5 4	0	0
6	G	1	Total O 1 1	0	0
6	G	1	Total C N O P 22 10 5 6 1	0	0
6	K	1	Total C N O P 22 10 5 6 1	0	0
6	H	1	Total C N O 19 10 5 4	0	0
6	L	1	Total C N O P 22 10 5 6 1	0	0

- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	2	Total Ca 2 2	0	0
7	B	2	Total Ca 2 2	0	0
7	C	2	Total Ca 2 2	0	0
7	D	2	Total Ca 2 2	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	E	11	Total O 11 11	0	0
8	I	11	Total O 11 11	0	0
8	F	8	Total O 8 8	0	0
8	J	6	Total O 6 6	0	0
8	G	7	Total O 7 7	0	0
8	K	3	Total O 3 3	0	0
8	H	8	Total O 8 8	0	0
8	L	3	Total O 3 3	0	0

Continued on next page...

Continued from previous page...

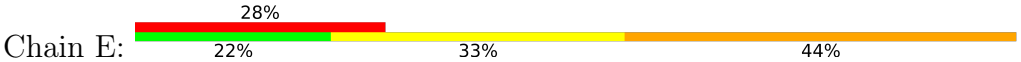
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	112	Total O 112 112	0	0
8	B	129	Total O 129 129	0	0
8	C	126	Total O 126 126	0	0
8	D	111	Total O 111 111	0	0

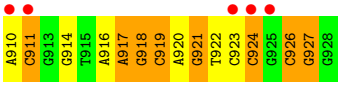
3 Residue-property plots (i)

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: 5'-GCGGACTGCTTAC(dideoxycytidine)-3'

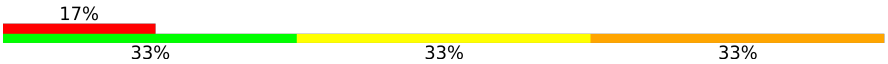
Chain E:

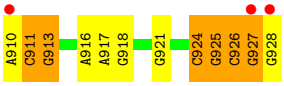




• Molecule 1: 5'-GCGGACTGCTTAC(dideoxycytidine)-3'

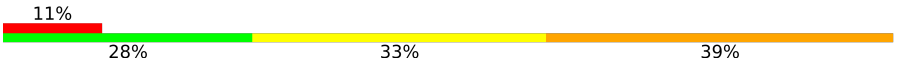
Chain F:

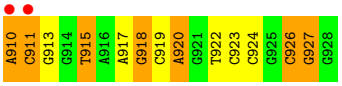




• Molecule 1: 5'-GCGGACTGCTTAC(dideoxycytidine)-3'

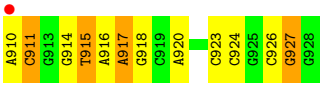
Chain G:





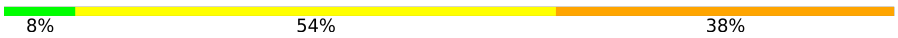
• Molecule 1: 5'-GCGGACTGCTTAC(dideoxycytidine)-3'

Chain H:



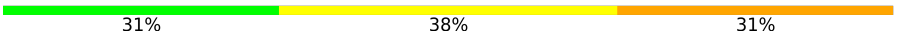
• Molecule 2: 5'-AC(tetrahydrofuran)GGTAAGCAGTCCGCGG-3'

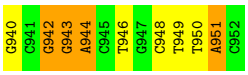
Chain I:



• Molecule 2: 5'-AC(tetrahydrofuran)GGTAAGCAGTCCGCGG-3'

Chain J:





• Molecule 2: 5'-AC(tetrahydrofuran)GGTAAGCAGTCCGCGG-3'

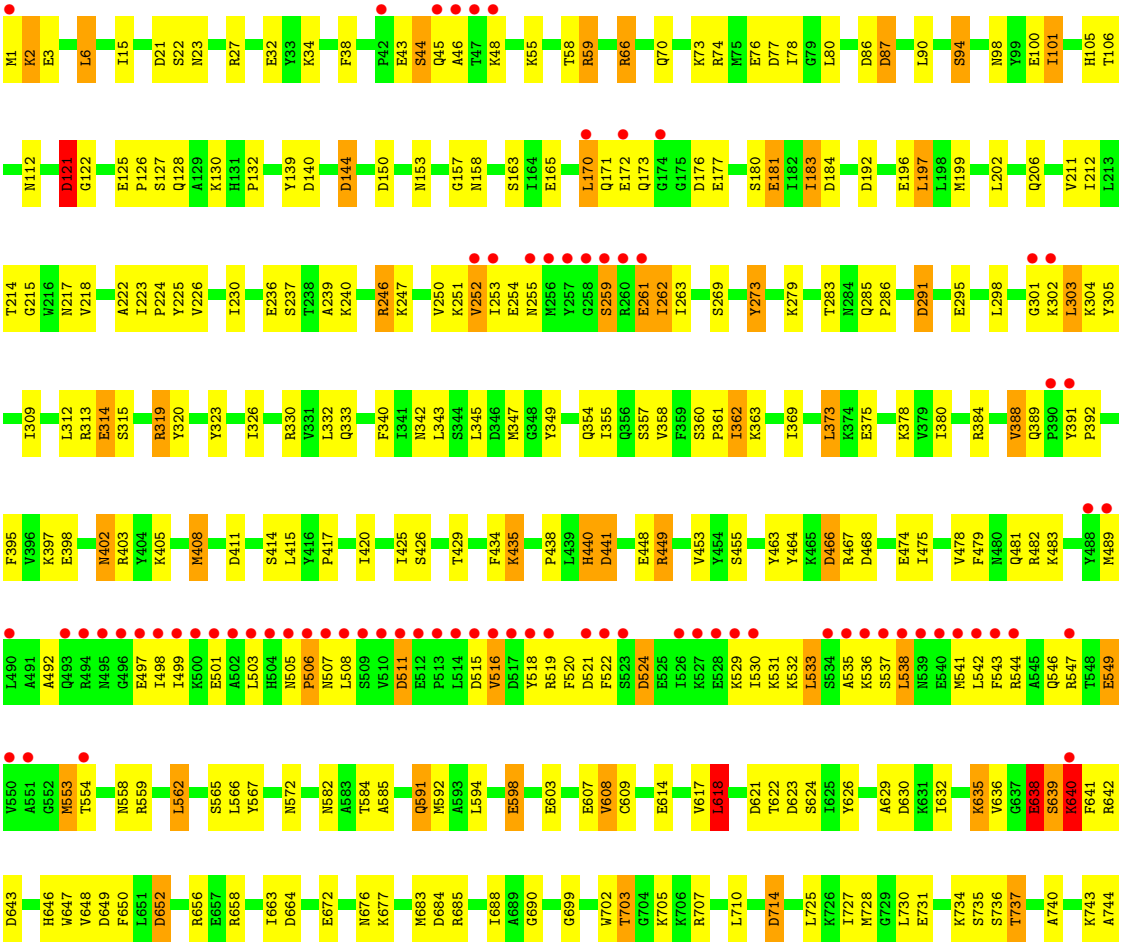
Chain K:

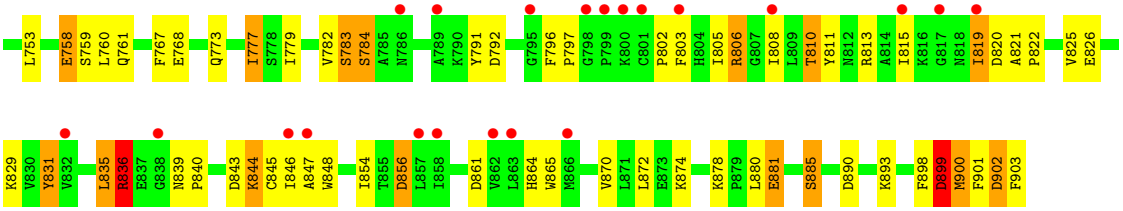
• Molecule 2: 5'-AC(tetrahydrofuran)GGTAAGCAGTCCGCGG-3'

Chain L:

• Molecule 3: DNA polymerase

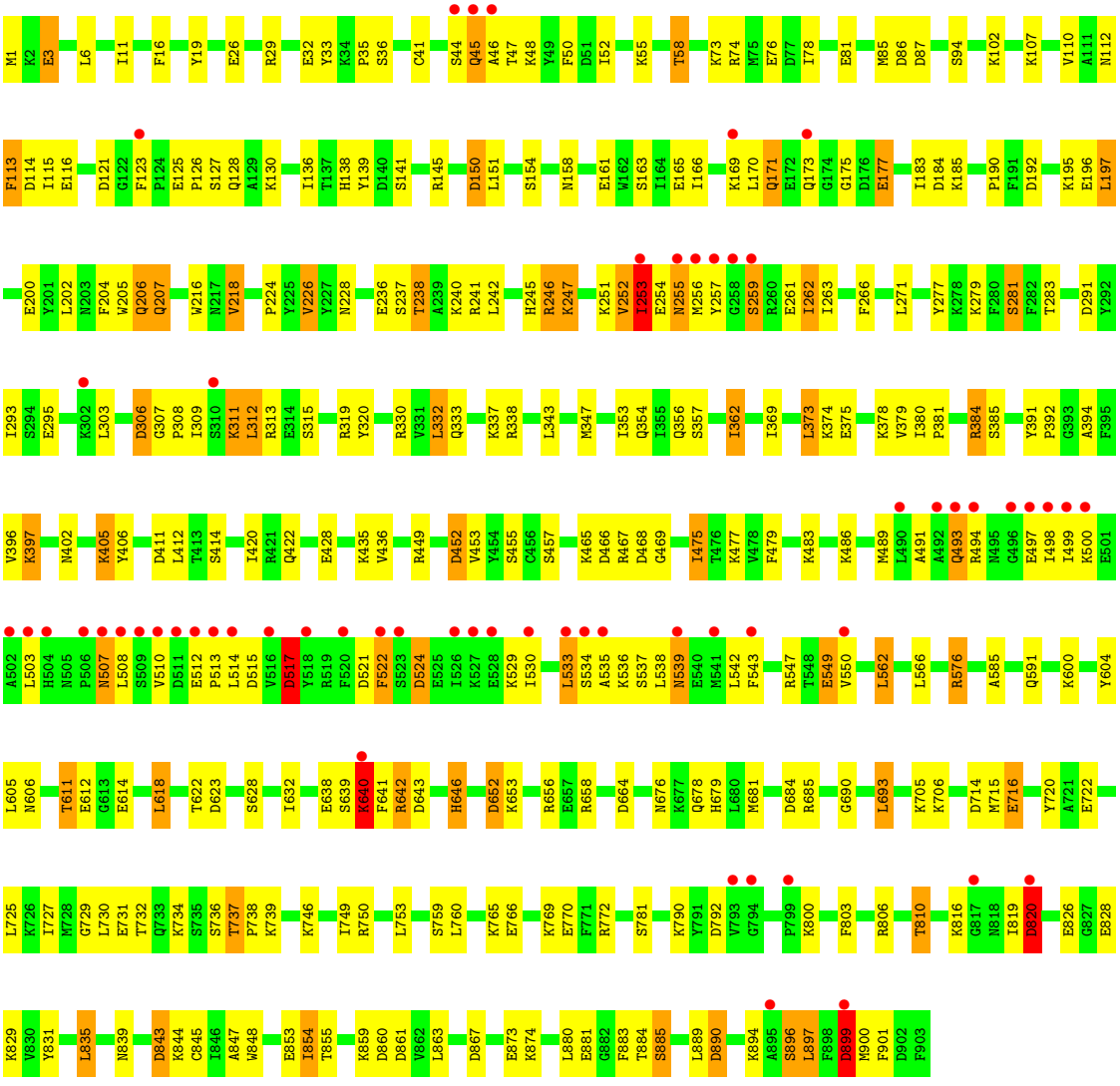
Chain A:





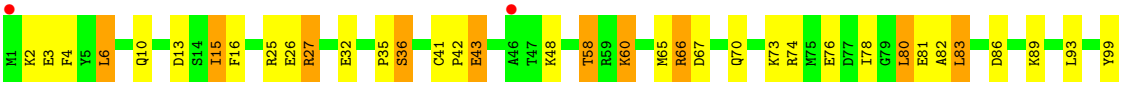
• Molecule 3: DNA polymerase

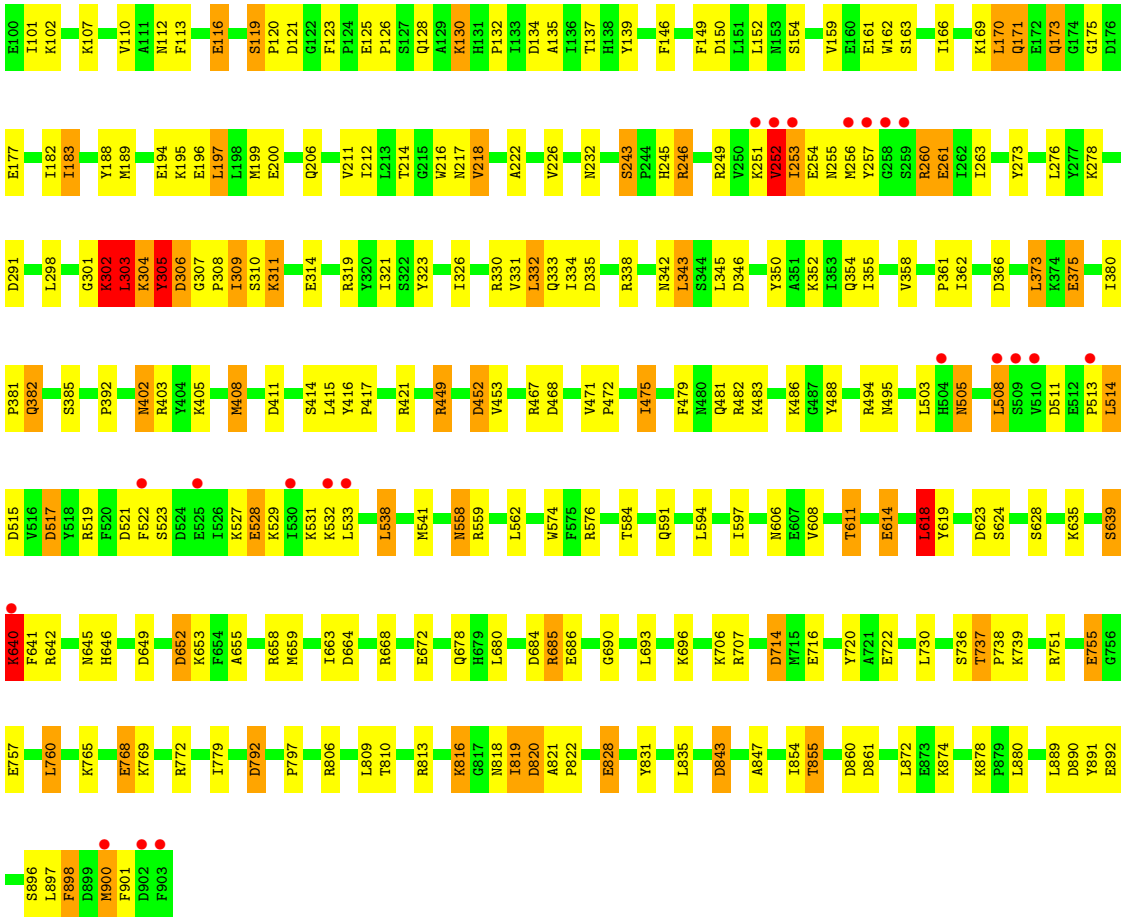
Chain B:



• Molecule 3: DNA polymerase

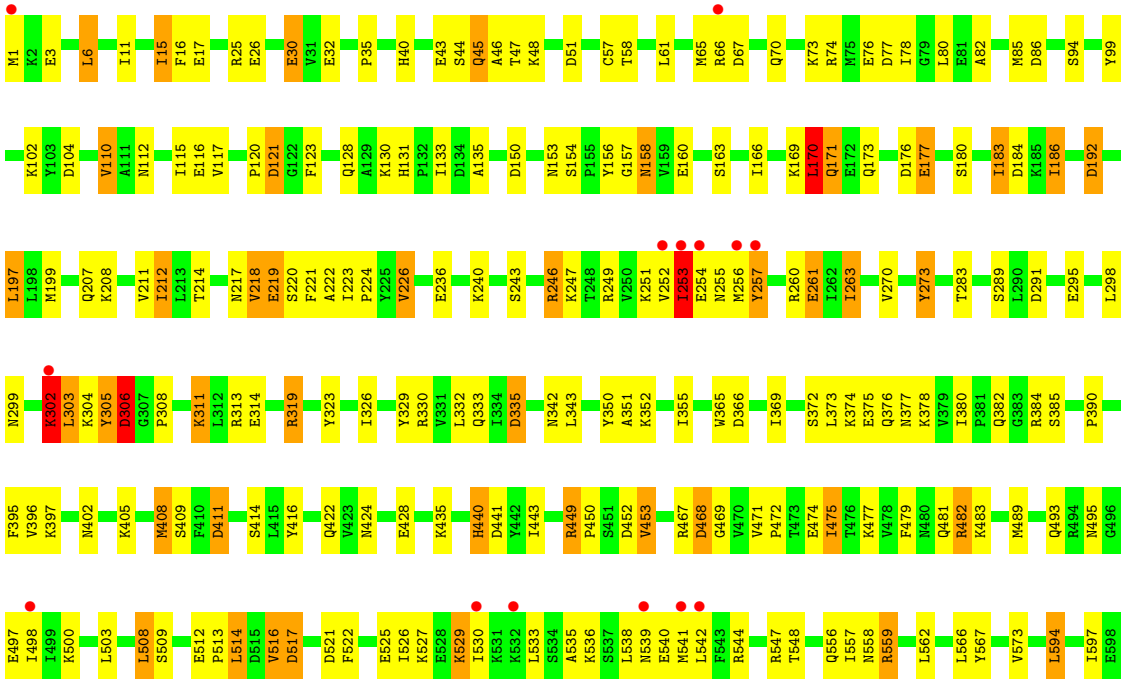
Chain C:

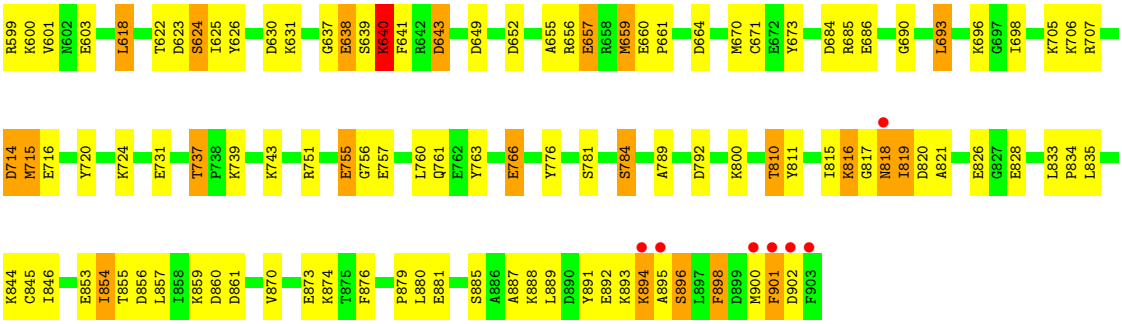




• Molecule 3: DNA polymerase

Chain D:




R599 K600 V601 N602 E603 L618 T622 D623 S624 I625 Y626 D630 K631 G637 E638 S639 K640 F641 R642 D643 D649 D652 A655 R656 E657 R658 M659 E660 P661 D664 M670 C671 E672 Y673 D684 R685 E686 G690 L693 K696 G697 I698 K705 K706 R707
D714 M715 E716 Y720 K724 E731 T737 P738 K739 K743 R751 E755 G756 E757 L760 Q761 E762 Y763 E766 Y776 S781 S784 A789 D792 K800 T810 Y811 I815 K816 G817 N818 I819 D820 A821 E826 G827 E828 L833 P834 L835
K844 C845 I846 E853 I854 T855 D856 L857 I858 K859 D860 D861 V870 E873 K874 T875 F876 P879 L880 E881 S885 A886 A887 K888 L889 D890 Y891 E892 K893 K894 A895 S896 L897 F898 D899 M900 F901 D902 P903

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	131.95Å 122.24Å 165.36Å 90.00° 96.85° 90.00°	Depositor
Resolution (Å)	47.00 – 2.69 45.07 – 2.70	Depositor EDS
% Data completeness (in resolution range)	94.7 (47.00-2.69) 95.1 (45.07-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$< I/\sigma(I) >$ [1]	2.48 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.215 , 0.288 0.208 , 0.273	Depositor DCC
R_{free} test set	6757 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	58.9	Xtriage
Anisotropy	0.324	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 45.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	32756	wwPDB-VP
Average B, all atoms (Å^2)	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.88% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: DOC, 3DR, DGP, CA

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	E	1.19	0/415	2.02	23/637 (3.6%)
1	F	1.07	0/415	1.77	8/637 (1.3%)
1	G	1.28	2/415 (0.5%)	2.14	18/637 (2.8%)
1	H	1.43	4/415 (1.0%)	1.90	13/637 (2.0%)
2	I	1.06	0/294	1.78	9/452 (2.0%)
2	J	1.15	0/294	1.90	7/452 (1.5%)
2	K	1.20	0/294	1.89	9/452 (2.0%)
2	L	1.25	1/294 (0.3%)	1.84	6/452 (1.3%)
3	A	0.56	0/7545	0.83	32/10196 (0.3%)
3	B	0.62	0/7545	0.83	28/10196 (0.3%)
3	C	0.65	0/7545	0.87	26/10196 (0.3%)
3	D	0.60	0/7545	0.83	27/10196 (0.3%)
All	All	0.68	7/33016 (0.0%)	0.99	206/45140 (0.5%)

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
3	B	0	1
3	C	0	1
All	All	0	2

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	H	926	DC	C4-C5	9.24	1.50	1.43
1	G	926	DC	C4-C5	6.59	1.48	1.43
1	H	926	DC	N1-C2	6.51	1.46	1.40
1	H	926	DC	N1-C6	6.47	1.41	1.37

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	926	DC	N1-C2	5.28	1.45	1.40

The worst 5 of 206 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	G	911	DC	C1'-O4'-C4'	-13.14	96.96	110.10
1	G	911	DC	O4'-C1'-N1	12.56	116.80	108.00
1	E	918	DG	O4'-C1'-N9	12.30	116.61	108.00
2	K	943	DG	O4'-C1'-N9	11.64	116.15	108.00
1	E	919	DC	O4'-C1'-N1	10.99	115.69	108.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	B	640	LYS	Peptide
3	C	252	VAL	Peptide

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	E	370	0	202	16	0
1	F	370	0	202	13	0
1	G	370	0	202	11	0
1	H	370	0	202	10	0
2	I	263	0	147	16	0
2	J	263	0	147	9	0
2	K	263	0	147	6	0
2	L	263	0	147	3	0
3	A	7365	0	7257	191	0
3	B	7365	0	7258	158	0
3	C	7365	0	7258	172	0
3	D	7365	0	7258	188	0
4	E	11	0	8	1	0
4	F	11	0	8	3	0
4	G	11	0	8	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	H	11	0	8	0	0
5	I	18	0	12	2	0
5	J	18	0	12	1	0
5	K	18	0	12	1	0
5	L	18	0	12	0	0
6	F	19	0	12	0	0
6	G	23	0	12	2	0
6	H	19	0	12	3	0
6	K	22	0	12	1	0
6	L	22	0	12	1	0
7	A	2	0	0	0	0
7	B	2	0	0	0	0
7	C	2	0	0	0	0
7	D	2	0	0	0	0
8	A	112	0	0	16	0
8	B	129	0	0	18	0
8	C	126	0	0	21	0
8	D	111	0	0	18	0
8	E	11	0	0	3	0
8	F	8	0	0	0	0
8	G	7	0	0	2	0
8	H	8	0	0	2	0
8	I	11	0	0	3	0
8	J	6	0	0	0	0
8	K	3	0	0	0	0
8	L	3	0	0	0	0
All	All	32756	0	30567	769	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 12.

The worst 5 of 769 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)
3:A:253:ILE:HG22	3:A:261:GLU:HB3	1.28	1.13
3:C:514:LEU:H	3:C:541:MET:HE2	1.11	1.12
3:C:171:GLN:NE2	3:C:177:GLU:HB2	1.66	1.10
3:C:89:LYS:HE2	3:C:354:GLN:HE22	1.13	1.06
3:A:253:ILE:HG22	3:A:261:GLU:CB	1.91	1.01

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
3	A	901/903 (100%)	805 (89%)	81 (9%)	15 (2%)	9 23
3	B	901/903 (100%)	830 (92%)	62 (7%)	9 (1%)	15 37
3	C	901/903 (100%)	827 (92%)	60 (7%)	14 (2%)	9 24
3	D	901/903 (100%)	823 (91%)	64 (7%)	14 (2%)	9 24
All	All	3604/3612 (100%)	3285 (91%)	267 (7%)	52 (1%)	11 28

5 of 52 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	22	SER
3	A	45	GLN
3	A	639	SER
3	A	640	LYS
3	A	902	ASP

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
3	A	798/800 (100%)	636 (80%)	162 (20%)	1 3
3	B	798/800 (100%)	640 (80%)	158 (20%)	1 3
3	C	798/800 (100%)	669 (84%)	129 (16%)	2 6
3	D	798/800 (100%)	656 (82%)	142 (18%)	2 4
All	All	3192/3200 (100%)	2601 (82%)	591 (18%)	1 4

5 of 591 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	66	ARG
3	D	800	LYS
3	D	163	SER
3	D	61	LEU
3	D	402	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 78 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	558	ASN
3	D	495	ASN
3	C	818	ASN
3	D	333	GLN
3	D	679	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 22 ligands modelled in this entry, 1 is modelled with single atom and 8 are monoatomic - leaving 13 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	$\#\|Z\|>2$
4	3DR	G	912	1	8,11,12	0.88	0	9,14,17	1.31	2 (22%)
5	DOC	L	953	2	16,19,20	0.52	0	20,26,29	1.07	1 (5%)
5	DOC	J	953	2	16,19,20	0.69	0	20,26,29	1.22	2 (10%)
4	3DR	F	912	1	8,11,12	0.48	0	9,14,17	0.96	0
6	DGP	K	955	-	18,24,25	1.19	3 (16%)	19,35,38	1.39	4 (21%)
5	DOC	K	953	2	16,19,20	0.49	0	20,26,29	1.95	3 (15%)
6	DGP	G	908	6,7	18,24,25	1.19	2 (11%)	19,35,38	0.86	0
4	3DR	E	912	1	8,11,12	0.64	0	9,14,17	0.98	0
6	DGP	F	908	-	18,21,25	1.03	1 (5%)	19,31,38	1.35	3 (15%)
6	DGP	L	955	-	18,24,25	1.20	2 (11%)	19,35,38	0.69	0
6	DGP	H	908	-	18,21,25	1.24	2 (11%)	19,31,38	1.14	1 (5%)
5	DOC	I	953	2	16,19,20	0.50	0	20,26,29	0.86	1 (5%)
4	3DR	H	912	1	8,11,12	0.62	0	9,14,17	0.83	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	3DR	G	912	1	-	1/3/15/16	0/1/1/1
5	DOC	L	953	2	-	0/7/18/19	0/2/2/2
5	DOC	J	953	2	-	0/7/18/19	0/2/2/2
4	3DR	F	912	1	-	0/3/15/16	0/1/1/1
6	DGP	K	955	-	-	1/3/21/22	0/3/3/3
5	DOC	K	953	2	-	2/7/18/19	0/2/2/2
6	DGP	G	908	6,7	-	2/3/21/22	0/3/3/3
4	3DR	E	912	1	-	2/3/15/16	0/1/1/1
6	DGP	F	908	-	-	2/2/18/22	0/3/3/3
6	DGP	L	955	-	-	1/3/21/22	0/3/3/3
6	DGP	H	908	-	-	0/2/18/22	0/3/3/3
5	DOC	I	953	2	-	2/7/18/19	0/2/2/2
4	3DR	H	912	1	-	2/3/15/16	0/1/1/1

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	L	955	DGP	C5-C6	-2.99	1.41	1.47
6	G	908	DGP	C8-N7	-2.90	1.30	1.35

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	H	908	DGP	C8-N7	-2.78	1.30	1.35
6	L	955	DGP	C8-N7	-2.76	1.30	1.35
6	G	908	DGP	C5-C6	-2.61	1.42	1.47

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	K	953	DOC	C4'-O4'-C1'	-7.11	103.09	109.81
6	K	955	DGP	C4'-O4'-C1'	-3.21	101.71	109.45
5	K	953	DOC	O4'-C1'-C2'	-3.07	103.34	106.67
5	J	953	DOC	C4'-O4'-C1'	-3.07	106.91	109.81
6	F	908	DGP	O6-C6-N1	-2.95	117.17	120.65

There are no chirality outliers.

5 of 15 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	I	953	DOC	C3'-C4'-C5'-O5'
5	I	953	DOC	O4'-C4'-C5'-O5'
5	K	953	DOC	O4'-C4'-C5'-O5'
6	F	908	DGP	C3'-C4'-C5'-O5'
6	F	908	DGP	O4'-C4'-C5'-O5'

There are no ring outliers.

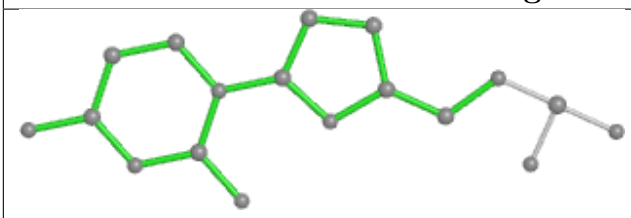
10 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	G	912	3DR	3	0
5	J	953	DOC	1	0
4	F	912	3DR	3	0
6	K	955	DGP	1	0
5	K	953	DOC	1	0
6	G	908	DGP	1	0
4	E	912	3DR	1	0
6	L	955	DGP	1	0
6	H	908	DGP	3	0
5	I	953	DOC	2	0

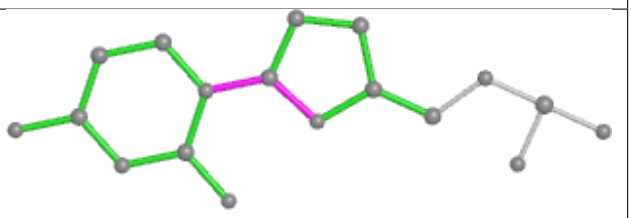
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will

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.

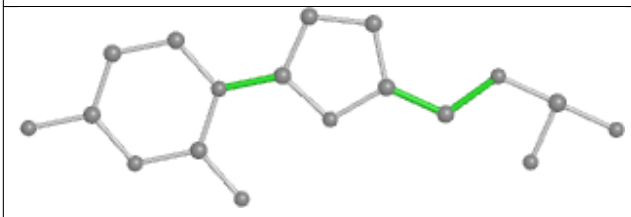
Ligand DOC L 953



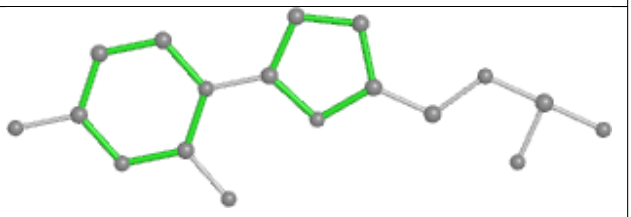
Bond lengths



Bond angles

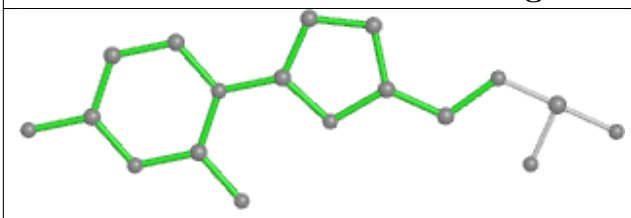


Torsions

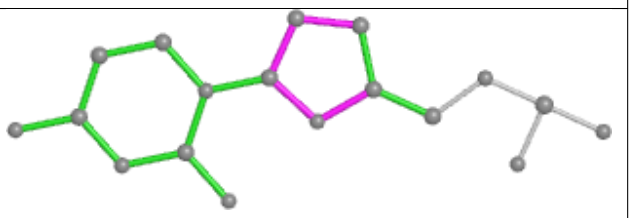


Rings

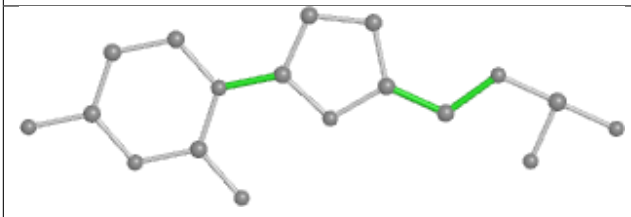
Ligand DOC J 953



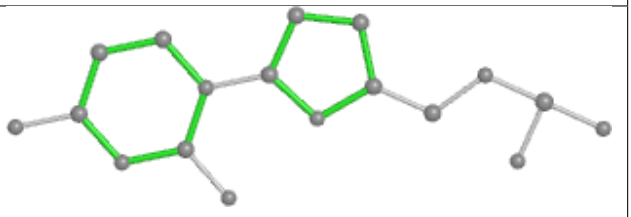
Bond lengths



Bond angles

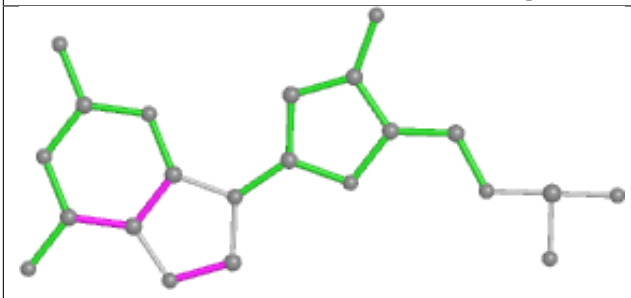


Torsions

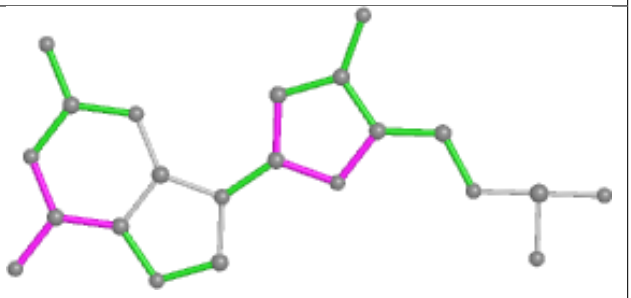


Rings

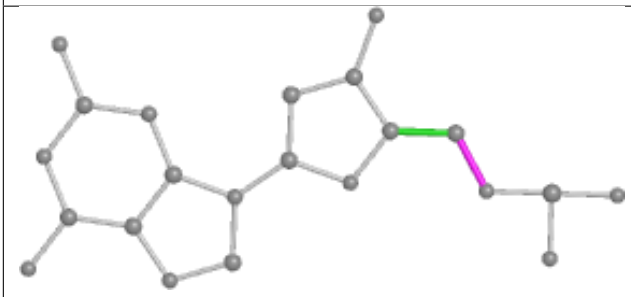
Ligand DGP K 955



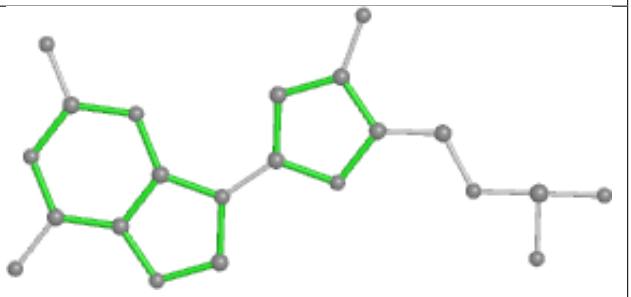
Bond lengths



Bond angles

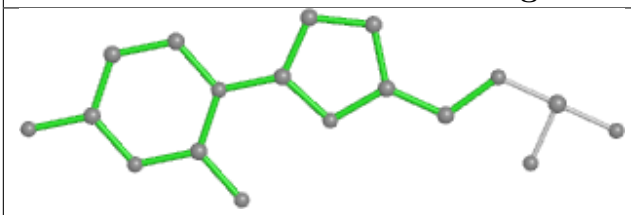


Torsions

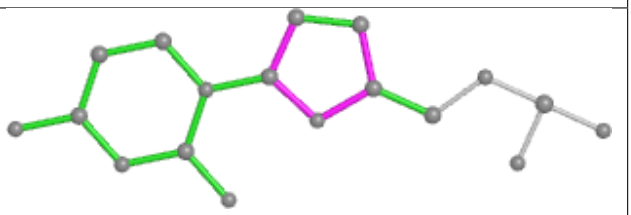


Rings

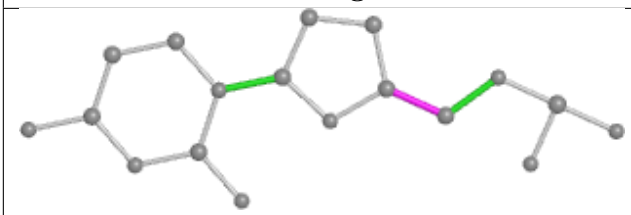
Ligand DOC K 953



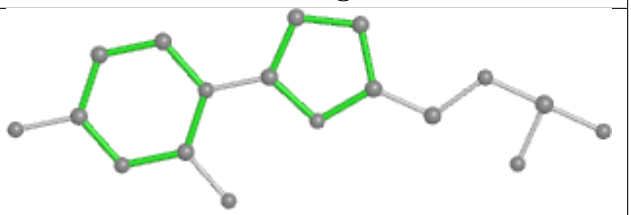
Bond lengths



Bond angles

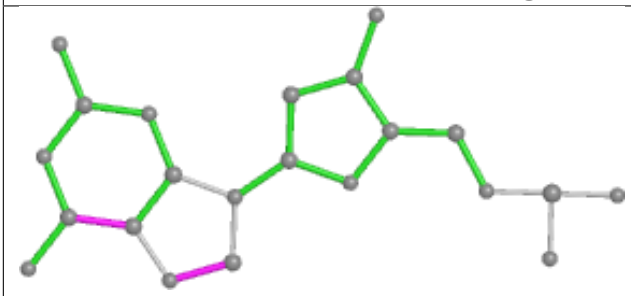


Torsions

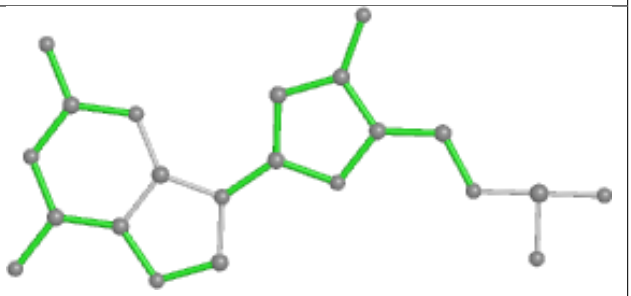


Rings

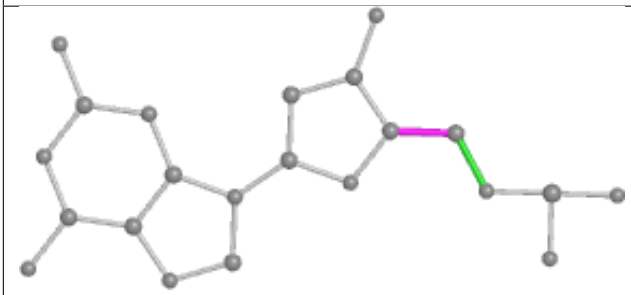
Ligand DGP G 908



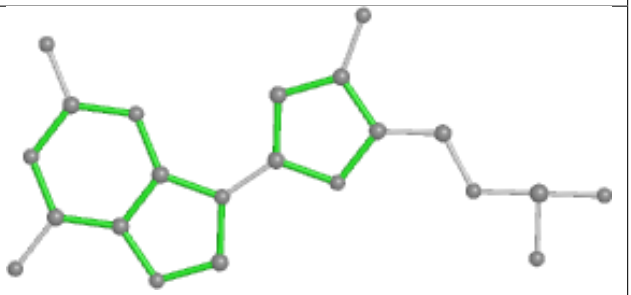
Bond lengths



Bond angles

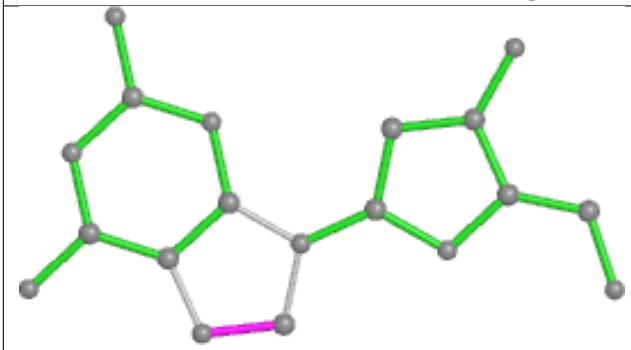


Torsions

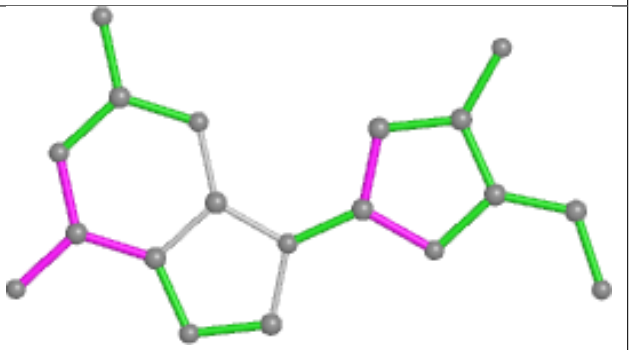


Rings

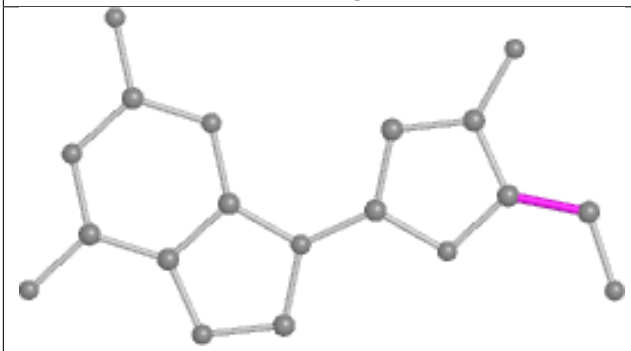
Ligand DGP F 908



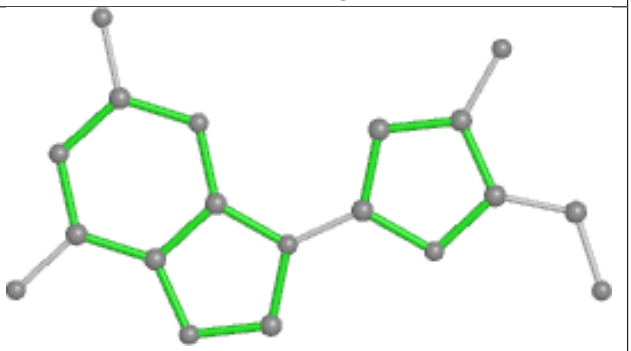
Bond lengths



Bond angles

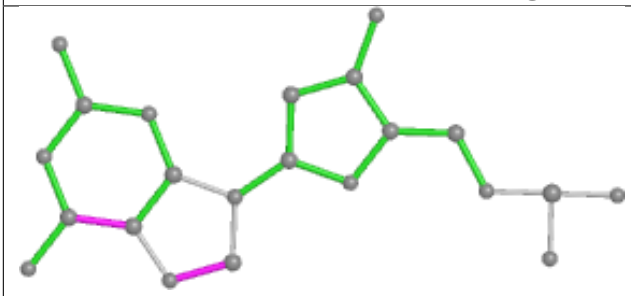


Torsions

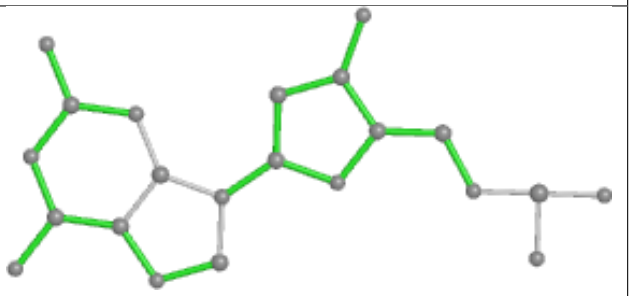


Rings

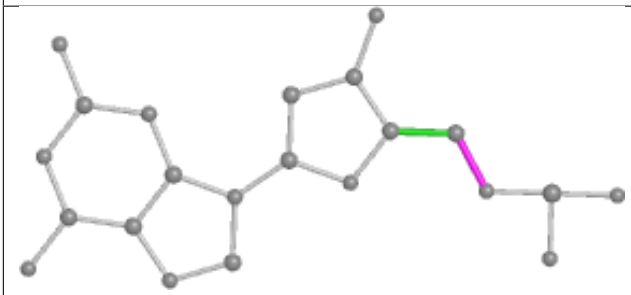
Ligand DGP L 955



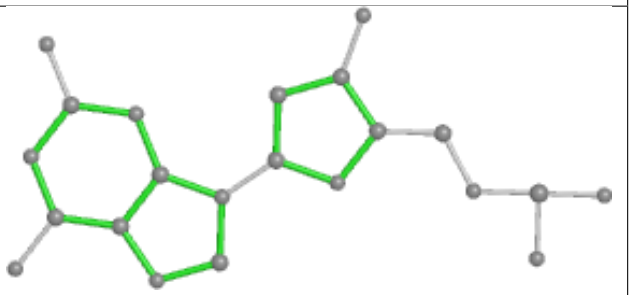
Bond lengths



Bond angles

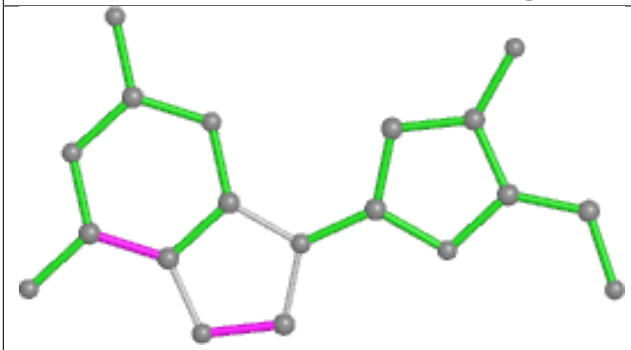


Torsions

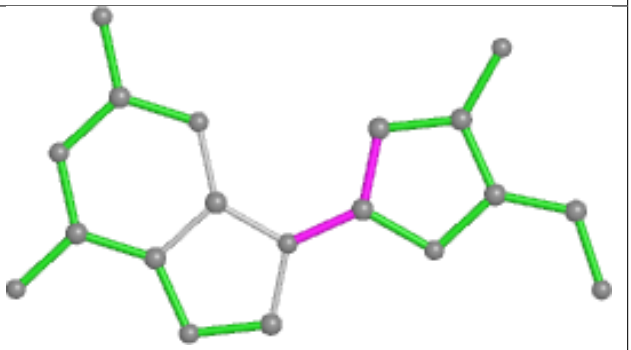


Rings

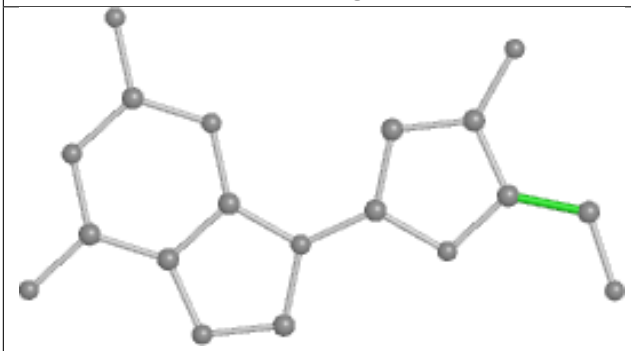
Ligand DGP H 908



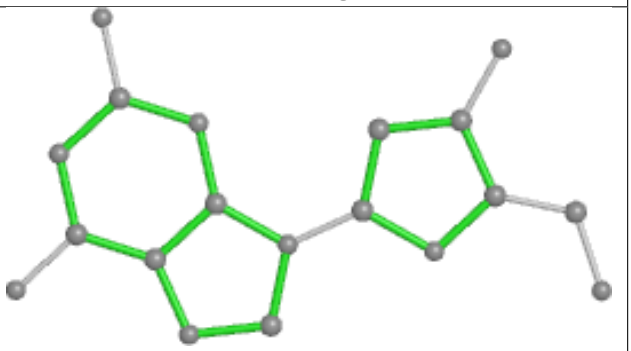
Bond lengths



Bond angles

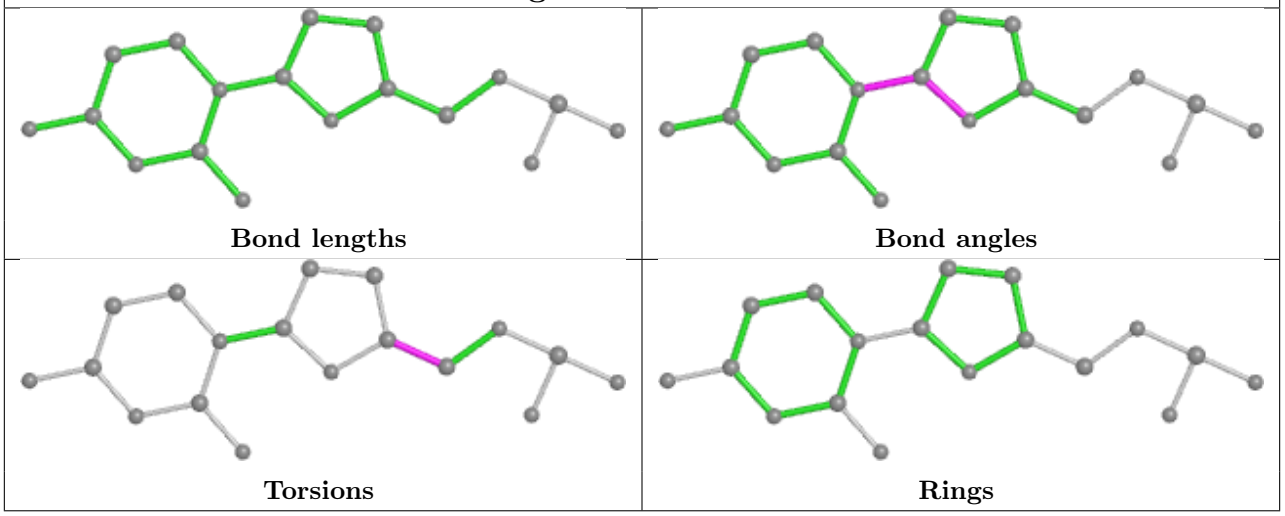


Torsions



Rings

Ligand DOC I 953



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	H	1
1	G	1
1	E	1
1	F	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	H	911:DC	O3'	913:DG	P	7.44
1	G	911:DC	O3'	913:DG	P	6.78
1	E	911:DC	O3'	913:DG	P	6.67
1	F	911:DC	O3'	913:DG	P	6.67

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	E	18/18 (100%)	1.33	5 (27%) 0 0	76, 92, 130, 152	0
1	F	18/18 (100%)	0.68	3 (16%) 1 1	39, 75, 149, 176	0
1	G	18/18 (100%)	0.23	2 (11%) 5 4	36, 54, 90, 102	0
1	H	18/18 (100%)	0.25	1 (5%) 24 23	44, 67, 128, 131	0
2	I	13/13 (100%)	0.47	0 100 100	77, 95, 100, 102	0
2	J	13/13 (100%)	-0.05	0 100 100	41, 68, 121, 135	0
2	K	13/13 (100%)	-0.31	0 100 100	36, 55, 77, 83	0
2	L	13/13 (100%)	-0.19	0 100 100	43, 65, 95, 102	0
3	A	903/903 (100%)	0.45	97 (10%) 6 4	30, 64, 147, 260	0
3	B	903/903 (100%)	0.15	59 (6%) 18 17	30, 53, 121, 222	0
3	C	903/903 (100%)	-0.05	23 (2%) 57 59	28, 47, 88, 138	0
3	D	903/903 (100%)	0.01	21 (2%) 60 62	30, 57, 93, 170	0
All	All	3736/3736 (100%)	0.15	211 (5%) 24 23	28, 56, 113, 260	0

The worst 5 of 211 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	510	VAL	14.7
3	B	508	LEU	12.9
3	A	509	SER	12.0
3	D	903	PHE	10.5
3	A	543	PHE	10.4

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 (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

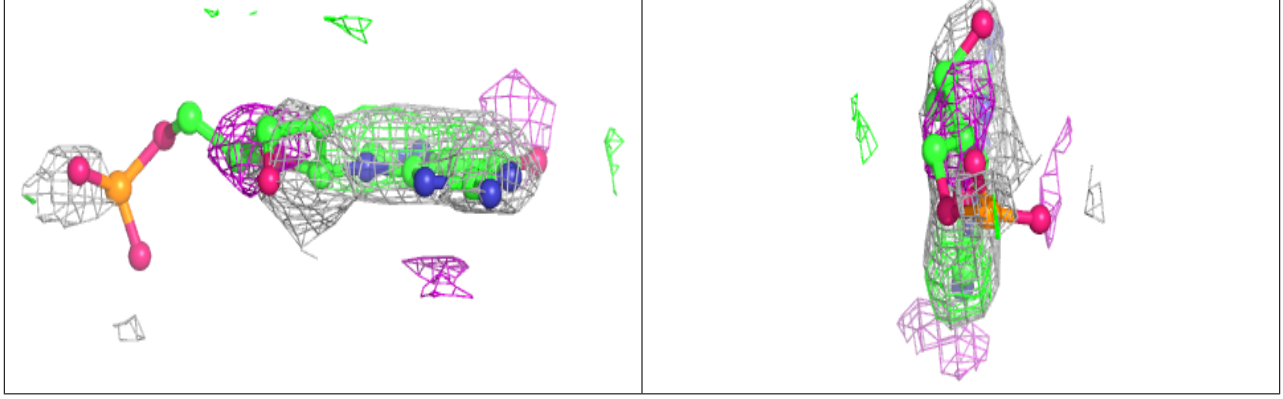
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
6	DGP	K	955	22/23	0.61	0.56	51,60,71,72	15
7	CA	B	1004	1/1	0.68	0.23	96,96,96,96	0
6	DGP	L	955	22/23	0.69	0.35	58,66,76,76	15
7	CA	D	1008	1/1	0.69	0.26	79,79,79,79	0
6	DGP	G	908	22/23	0.77	0.41	70,72,83,85	0
6	DGP	F	908	19/23	0.78	0.44	72,75,81,81	0
6	DGP	H	908	19/23	0.81	0.39	69,71,75,75	0
7	CA	A	1001	1/1	0.82	0.14	91,91,91,91	0
7	CA	C	1005	1/1	0.83	0.14	75,75,75,75	0
7	CA	B	1003	1/1	0.84	0.08	54,54,54,54	0
4	3DR	E	912	11/12	0.85	0.21	92,98,109,109	0
7	CA	C	1006	1/1	0.87	0.10	57,57,57,57	0
7	CA	D	1007	1/1	0.89	0.10	55,55,55,55	0
4	3DR	H	912	11/12	0.93	0.21	61,66,81,82	0
5	DOC	I	953	18/19	0.93	0.14	74,81,85,89	0
6	DGP	G	907	1/23	0.94	0.61	84,84,84,84	0
4	3DR	G	912	11/12	0.94	0.17	51,55,68,71	0
4	3DR	F	912	11/12	0.96	0.17	57,59,74,77	0
5	DOC	K	953	18/19	0.97	0.16	32,36,40,40	0
5	DOC	L	953	18/19	0.97	0.15	39,43,51,52	0
5	DOC	J	953	18/19	0.97	0.15	39,42,46,46	0
7	CA	A	1002	1/1	0.99	0.14	67,67,67,67	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.

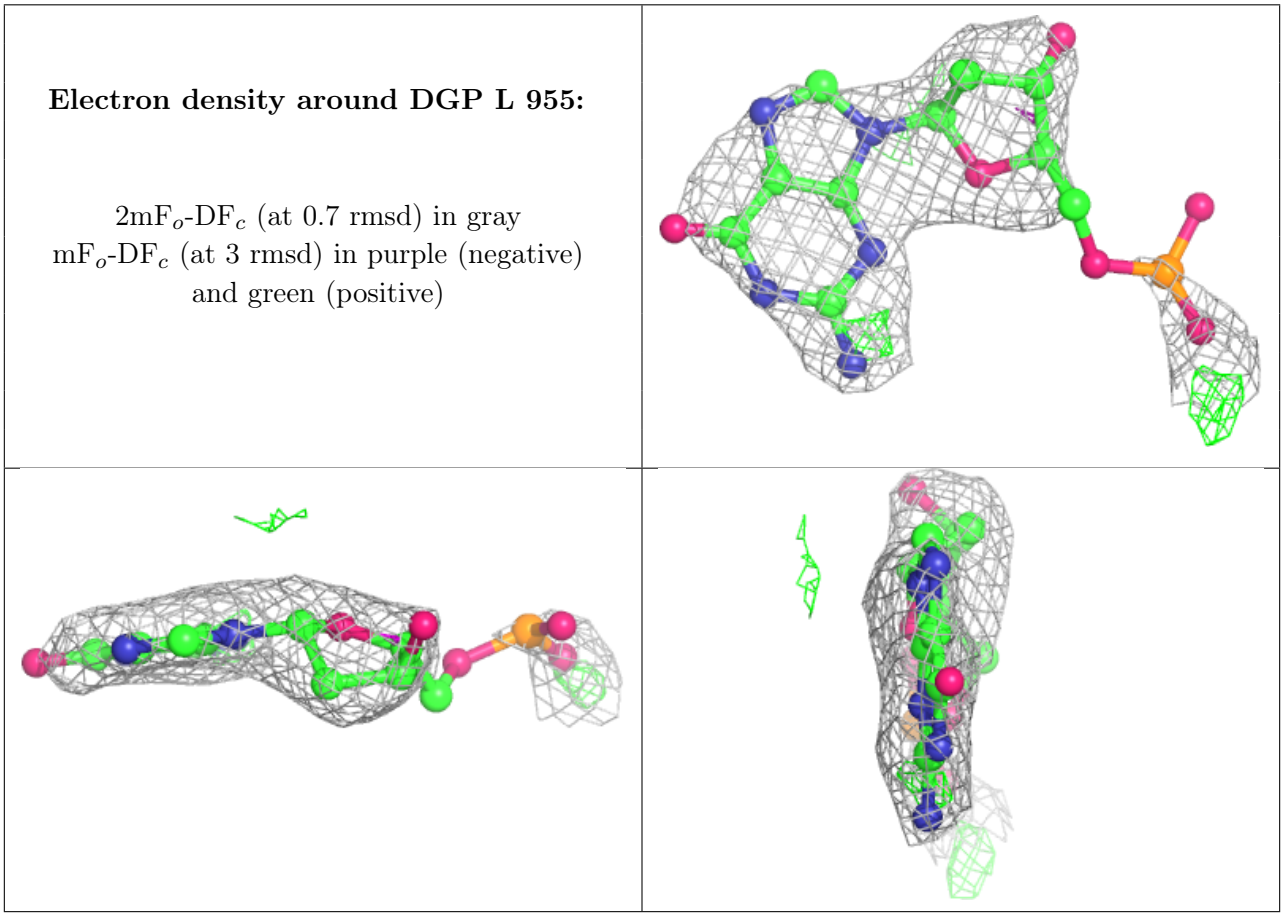
Electron density around DGP K 955:

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



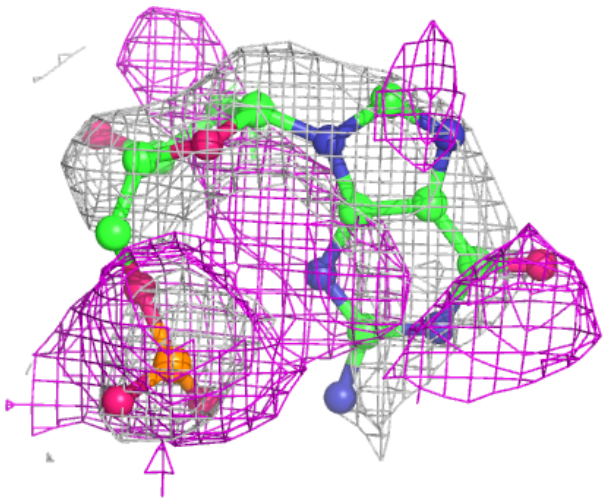
Electron density around DGP L 955:

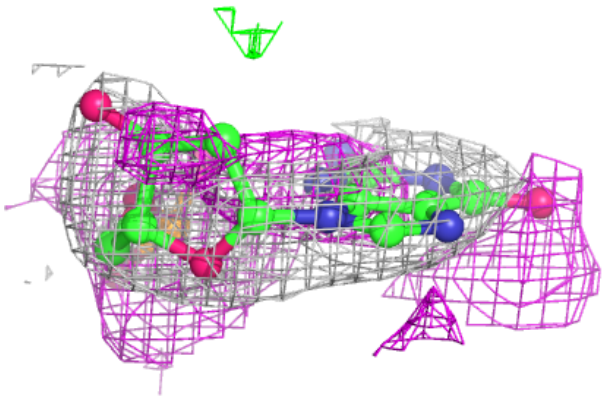
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

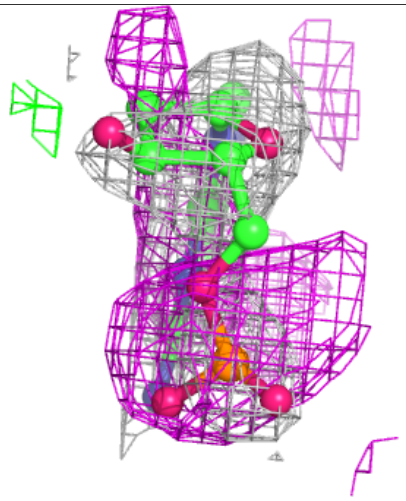


Electron density around DGP G 908:

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

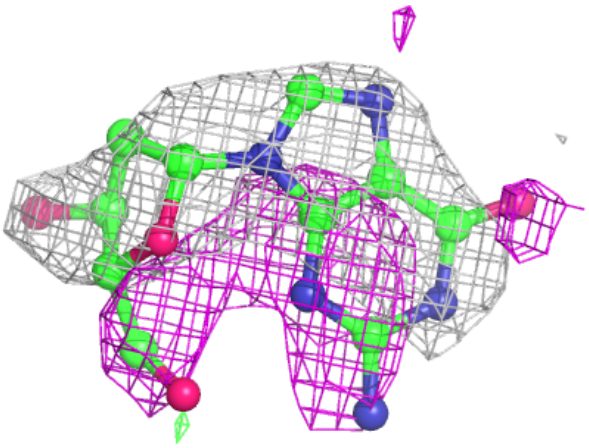


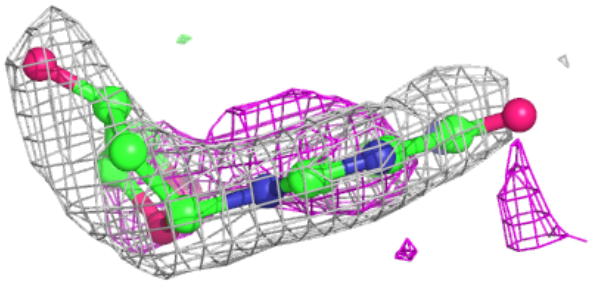


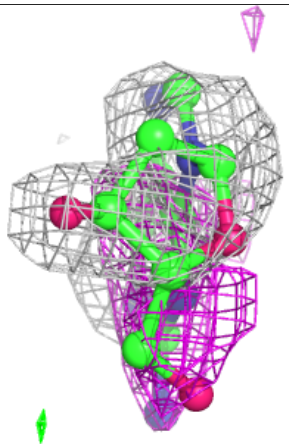


Electron density around DGP F 908:

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

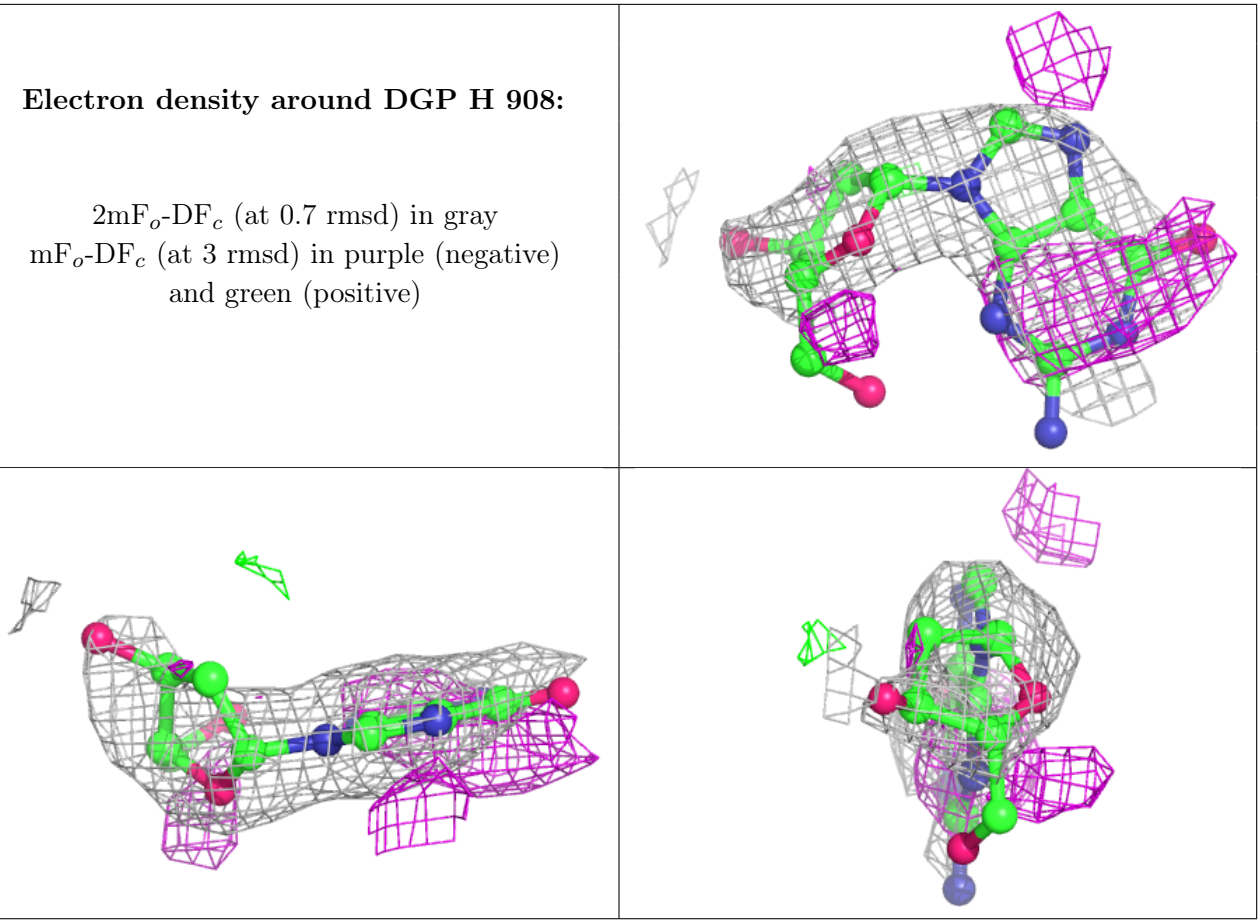






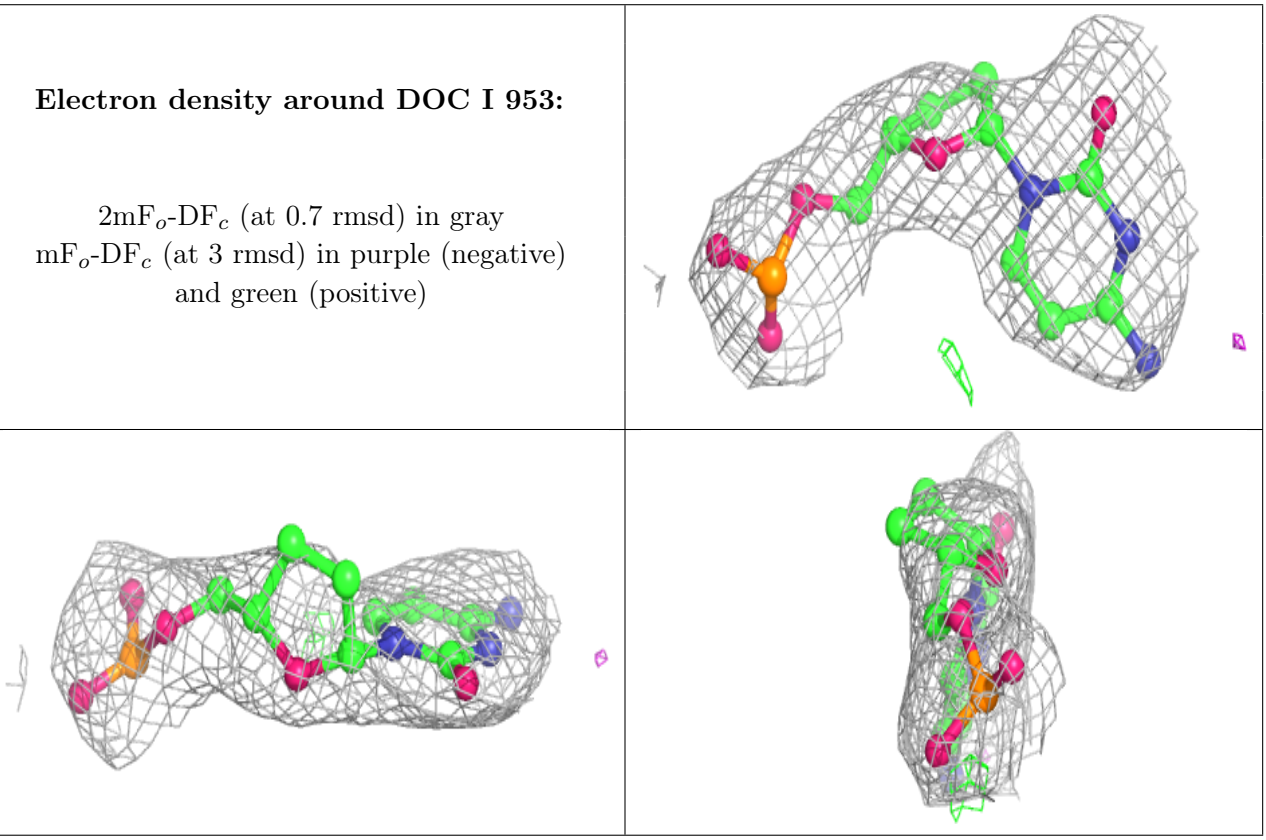
Electron density around DGP H 908:

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



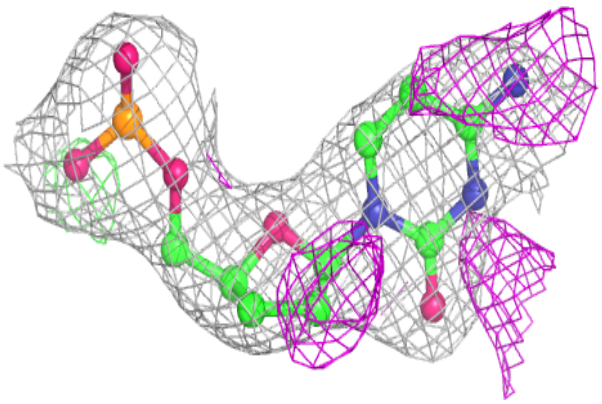
Electron density around DOC I 953:

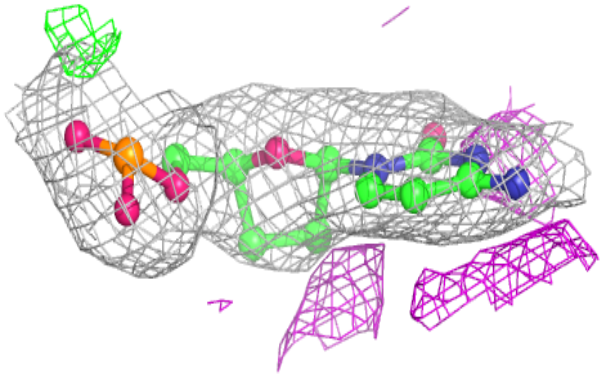
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

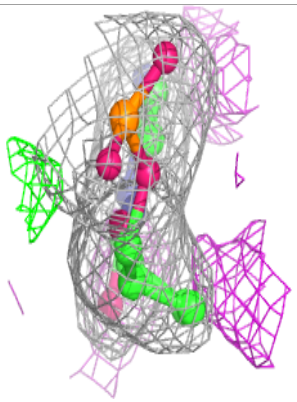


Electron density around DOC K 953:

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

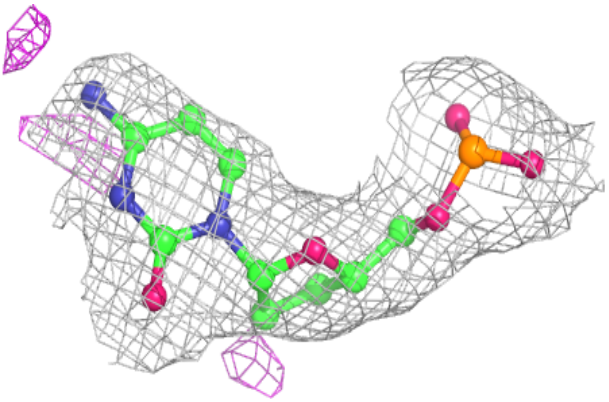


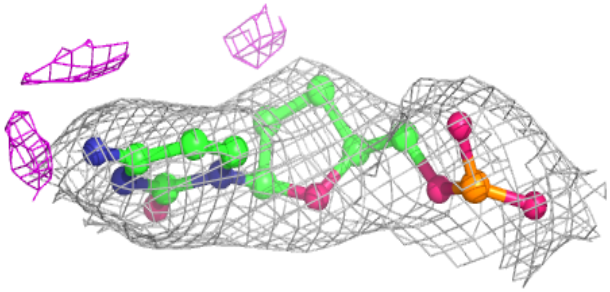


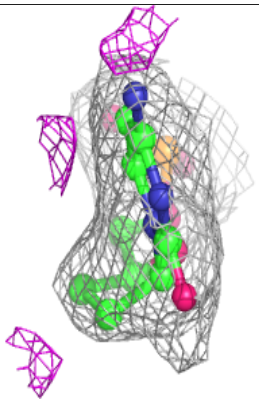


Electron density around DOC L 953:

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

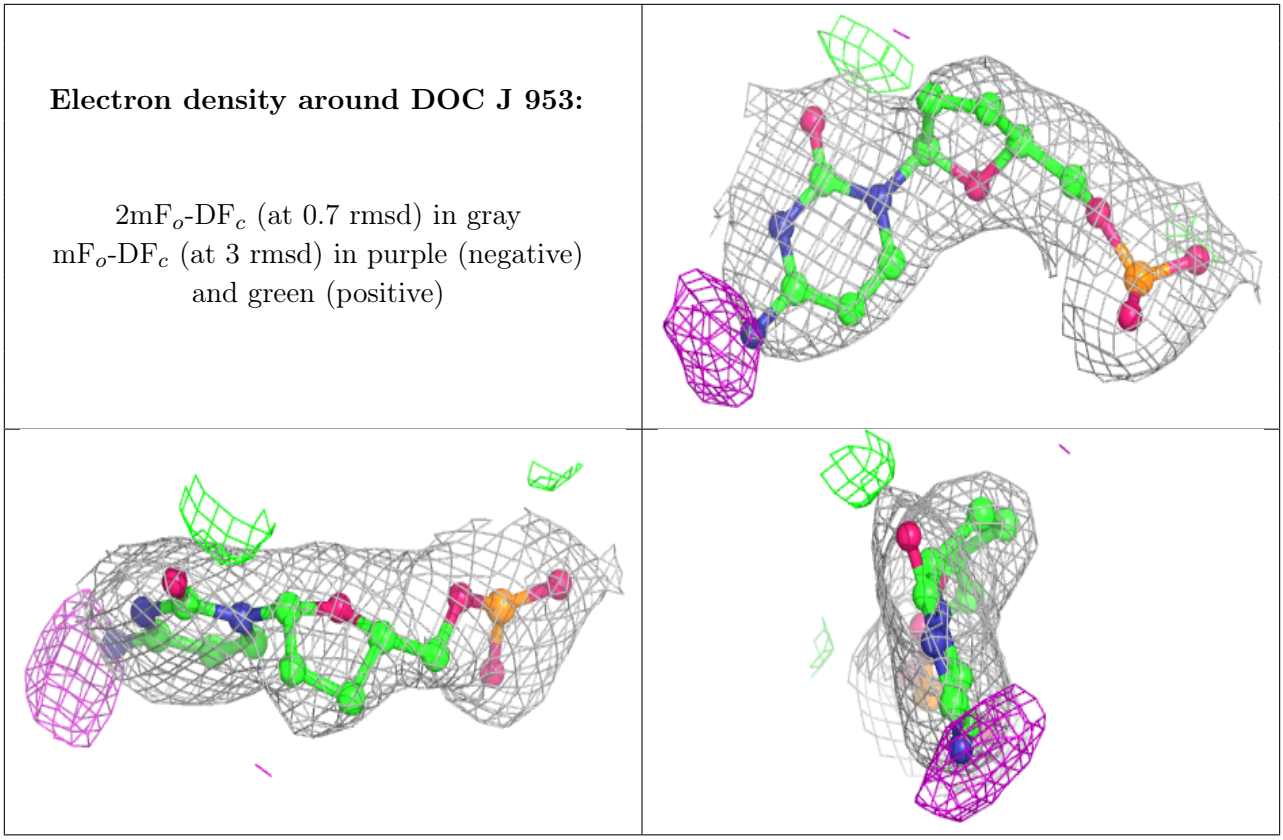






Electron density around DOC J 953:

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



6.5 Other polymers (i)

There are no such residues in this entry.