wwPDB X-ray Structure Validation Summary Report (i)

Aug 30, 2023 – 10:59 AM EDT

PDB ID	:	3OLA
Title	:	Poliovirus polymerase elongation complex with 2'-deoxy-CTP
Authors	:	Gong, P.; Peersen, O.B.
Deposited on	:	2010-08-25
Resolution	:	2.55 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:

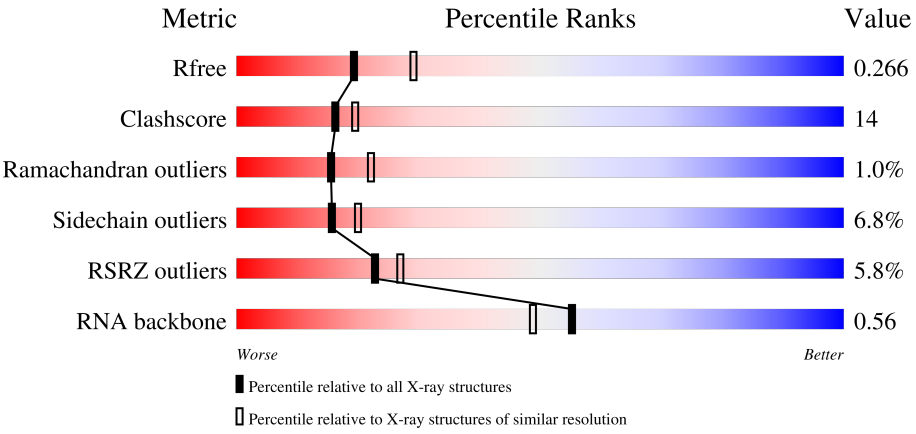
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 2.55 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	1284 (2.56-2.52)
Clashscore	141614	1332 (2.56-2.52)
Ramachandran outliers	138981	1315 (2.56-2.52)
Sidechain outliers	138945	1315 (2.56-2.52)
RSRZ outliers	127900	1272 (2.56-2.52)
RNA backbone	3102	1026 (2.88-2.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	471	6% 70% 25%
1	E	471	5% 72% 23%
1	I	471	7% 70% 24%
1	M	471	6% 72% 23%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	26	58% 35%
2	F	26	8% 54% 35%
2	J	26	27% 35% 31%
2	N	26	15% 46% 31%
3	C	15	40% 53% 7%
3	G	15	47% 47% 7%
4	D	9	22% 11% 67%
4	H	9	22% 11% 67%
4	L	9	22% 22% 56%
4	P	9	22% 22% 56%
5	K	14	29% 71%
5	O	14	7% 29% 71%

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
7	IPA	I	6029	-	-	-	X
8	PEG	A	7001	-	-	X	-
8	PEG	A	7003	-	-	X	-
8	PEG	I	7002	-	-	X	-

2 Entry composition (i)

There are 11 unique types of molecules in this entry. The entry contains 18452 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 Polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	461	Total	C	N	O	S	0	0	0
			3697	2370	610	695	22			
1	E	461	Total	C	N	O	S	0	0	0
			3697	2370	610	695	22			
1	I	461	Total	C	N	O	S	0	0	0
			3697	2370	610	695	22			
1	M	461	Total	C	N	O	S	0	0	0
			3697	2370	610	695	22			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	446	ASP	LEU	engineered mutation	UNP B3VQP5
A	462	GLY	-	expression tag	UNP B3VQP5
A	463	SER	-	expression tag	UNP B3VQP5
A	464	SER	-	expression tag	UNP B3VQP5
A	465	SER	-	expression tag	UNP B3VQP5
A	466	HIS	-	expression tag	UNP B3VQP5
A	467	HIS	-	expression tag	UNP B3VQP5
A	468	HIS	-	expression tag	UNP B3VQP5
A	469	HIS	-	expression tag	UNP B3VQP5
A	470	HIS	-	expression tag	UNP B3VQP5
A	471	HIS	-	expression tag	UNP B3VQP5
E	446	ASP	LEU	engineered mutation	UNP B3VQP5
E	462	GLY	-	expression tag	UNP B3VQP5
E	463	SER	-	expression tag	UNP B3VQP5
E	464	SER	-	expression tag	UNP B3VQP5
E	465	SER	-	expression tag	UNP B3VQP5
E	466	HIS	-	expression tag	UNP B3VQP5
E	467	HIS	-	expression tag	UNP B3VQP5
E	468	HIS	-	expression tag	UNP B3VQP5
E	469	HIS	-	expression tag	UNP B3VQP5
E	470	HIS	-	expression tag	UNP B3VQP5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	471	HIS	-	expression tag	UNP B3VQP5
I	446	ASP	LEU	engineered mutation	UNP B3VQP5
I	462	GLY	-	expression tag	UNP B3VQP5
I	463	SER	-	expression tag	UNP B3VQP5
I	464	SER	-	expression tag	UNP B3VQP5
I	465	SER	-	expression tag	UNP B3VQP5
I	466	HIS	-	expression tag	UNP B3VQP5
I	467	HIS	-	expression tag	UNP B3VQP5
I	468	HIS	-	expression tag	UNP B3VQP5
I	469	HIS	-	expression tag	UNP B3VQP5
I	470	HIS	-	expression tag	UNP B3VQP5
I	471	HIS	-	expression tag	UNP B3VQP5
M	446	ASP	LEU	engineered mutation	UNP B3VQP5
M	462	GLY	-	expression tag	UNP B3VQP5
M	463	SER	-	expression tag	UNP B3VQP5
M	464	SER	-	expression tag	UNP B3VQP5
M	465	SER	-	expression tag	UNP B3VQP5
M	466	HIS	-	expression tag	UNP B3VQP5
M	467	HIS	-	expression tag	UNP B3VQP5
M	468	HIS	-	expression tag	UNP B3VQP5
M	469	HIS	-	expression tag	UNP B3VQP5
M	470	HIS	-	expression tag	UNP B3VQP5
M	471	HIS	-	expression tag	UNP B3VQP5

- Molecule 2 is a RNA chain called RNA (5'-R(*AP*AP*GP*UP*CP*UP*CP*CP*AP*GP*GP*UP*CP*UP*CP*UP*CP*GP*UP*CP*CP*GP*GP*AP*AP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	17	Total	C	N	O	P	0	0	0
			357	159	59	122	17			
2	F	17	Total	C	N	O	P	0	0	0
			357	159	59	122	17			
2	J	18	Total	C	N	O	P	0	0	0
			377	168	61	130	18			
2	N	18	Total	C	N	O	P	0	0	0
			377	168	61	130	18			

- Molecule 3 is DNA/RNA hybrid called DNA/RNA (5'-R(*GP*CP*CP*CP*GP*GP*AP*CP*GP*AP*GP*AP*GP*A)-D(P*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	15	Total	C	N	O	P	0	0	0
			322	145	65	98	14			
3	G	15	Total	C	N	O	P	0	0	0
			322	145	65	98	14			

- Molecule 4 is a RNA chain called RNA (5'-R(*GP*GP*GP*AP*GP*AP*UP*GP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	3	Total	C	N	O	P	0	0	0
			68	30	15	20	3			
4	H	3	Total	C	N	O	P	0	0	0
			68	30	15	20	3			
4	L	4	Total	C	N	O	P	0	0	0
			91	40	20	27	4			
4	P	4	Total	C	N	O	P	0	0	0
			91	40	20	27	4			

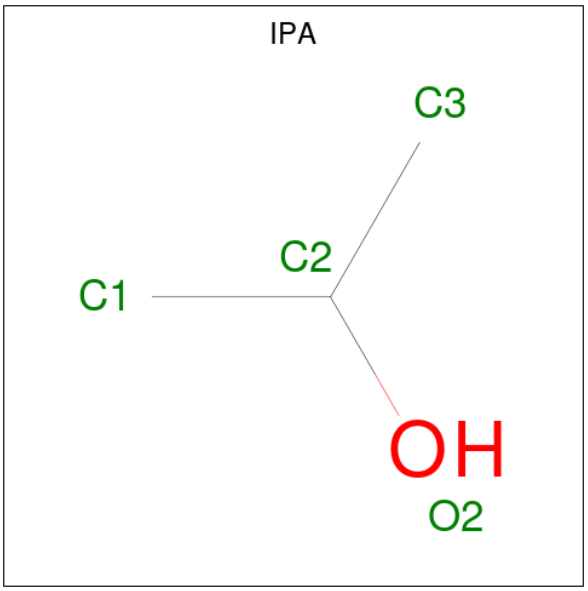
- Molecule 5 is a RNA chain called RNA (5'-R(*GP*CP*CP*CP*GP*GP*AP*CP*GP*AP*GP*AP*GP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	K	14	Total	C	N	O	P	0	0	0
			303	136	62	92	13			
5	O	14	Total	C	N	O	P	0	0	0
			303	136	62	92	13			

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

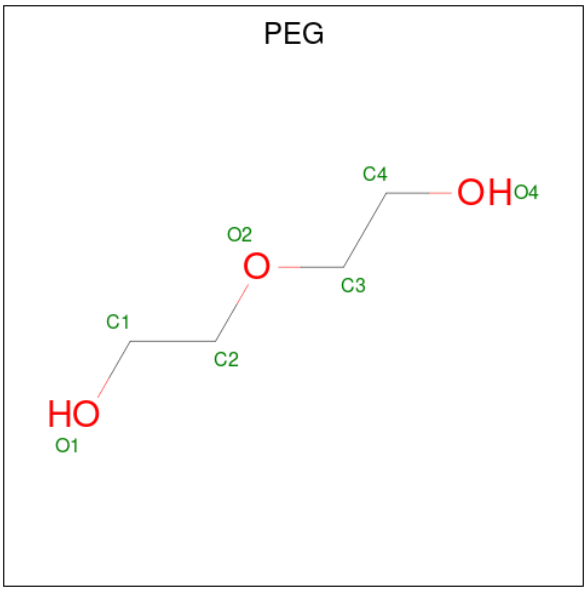
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Zn	0	0
			1	1		
6	E	1	Total	Zn	0	0
			1	1		
6	I	1	Total	Zn	0	0
			1	1		
6	M	1	Total	Zn	0	0
			1	1		

- Molecule 7 is ISOPROPYL ALCOHOL (three-letter code: IPA) (formula: C_3H_8O).



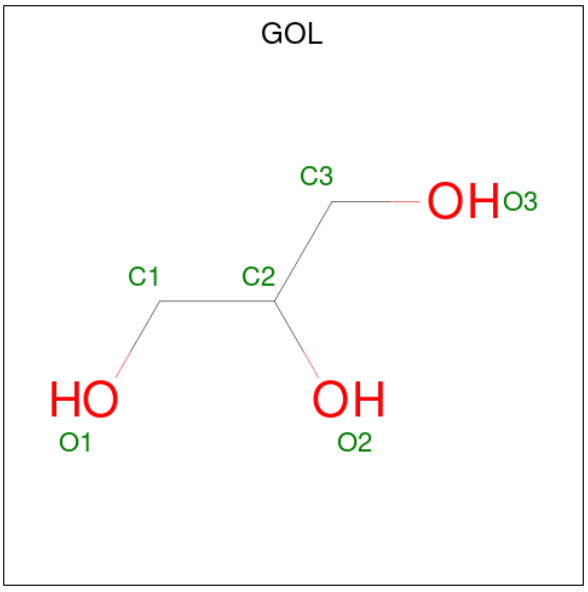
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C O 4 3 1	0	0
7	A	1	Total C O 4 3 1	0	0
7	A	1	Total C O 4 3 1	0	0
7	E	1	Total C O 4 3 1	0	0
7	E	1	Total C O 4 3 1	0	0
7	E	1	Total C O 4 3 1	0	0
7	I	1	Total C O 4 3 1	0	0
7	I	1	Total C O 4 3 1	0	0
7	M	1	Total C O 4 3 1	0	0
7	M	1	Total C O 4 3 1	0	0

- Molecule 8 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total C O 7 4 3	0	0
8	A	1	Total C O 7 4 3	0	0
8	I	1	Total C O 7 4 3	0	0

- Molecule 9 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



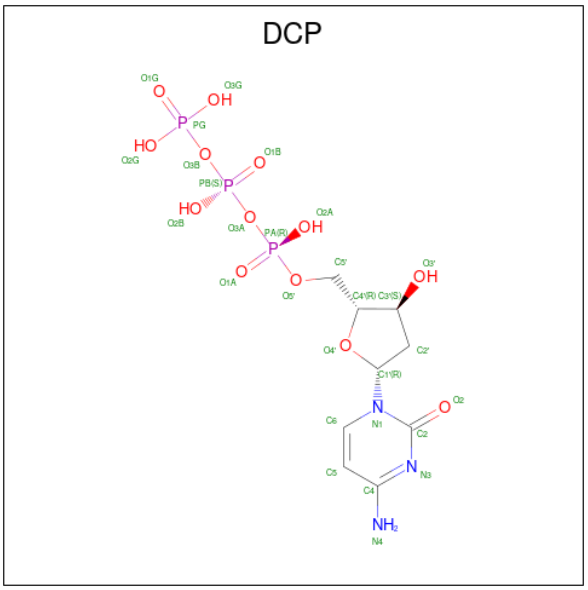
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	1	Total C O 6 3 3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	I	1	Total C O 6 3 3	0	0

- Molecule 10 is 2'-DEOXYCYTIDINE-5'-TRIPHOSPHATE (three-letter code: DCP) (formula: $C_9H_{16}N_3O_{13}P_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	K	1	Total C N O P 24 9 3 10 2	0	0
10	O	1	Total C N O P 24 9 3 10 2	0	0

- Molecule 11 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	A	94	Total O 94 94	0	0
11	B	26	Total O 26 26	0	0
11	C	9	Total O 9 9	0	0
11	D	1	Total O 1 1	0	0
11	E	100	Total O 100 100	0	0
11	F	19	Total O 19 19	0	0

Continued on next page...

Continued from previous page...

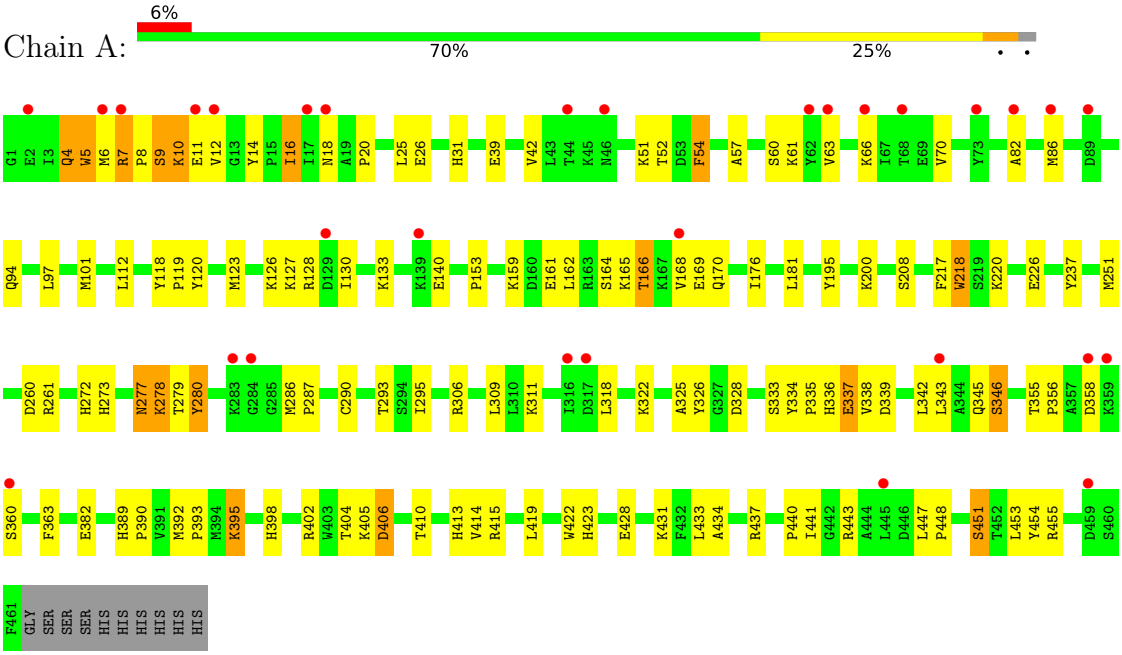
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	G	4	Total O 4 4	0	0
11	I	85	Total O 85 85	0	0
11	J	23	Total O 23 23	0	0
11	K	11	Total O 11 11	0	0
11	L	5	Total O 5 5	0	0
11	M	89	Total O 89 89	0	0
11	N	28	Total O 28 28	0	0
11	O	8	Total O 8 8	0	0
11	P	1	Total O 1 1	0	0

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.

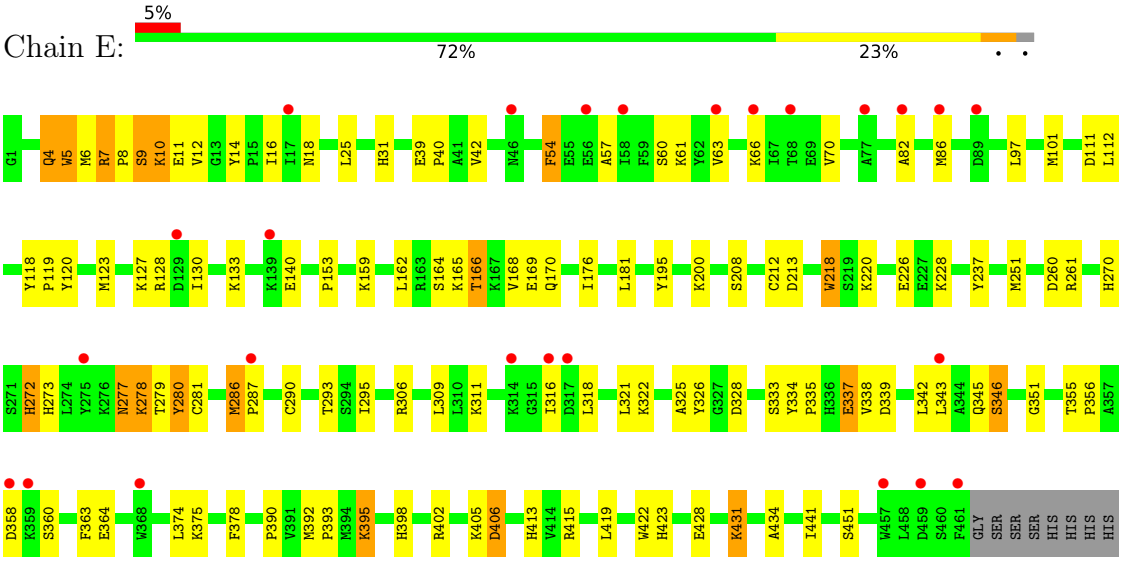
• Molecule 1: Polymerase

Chain A:



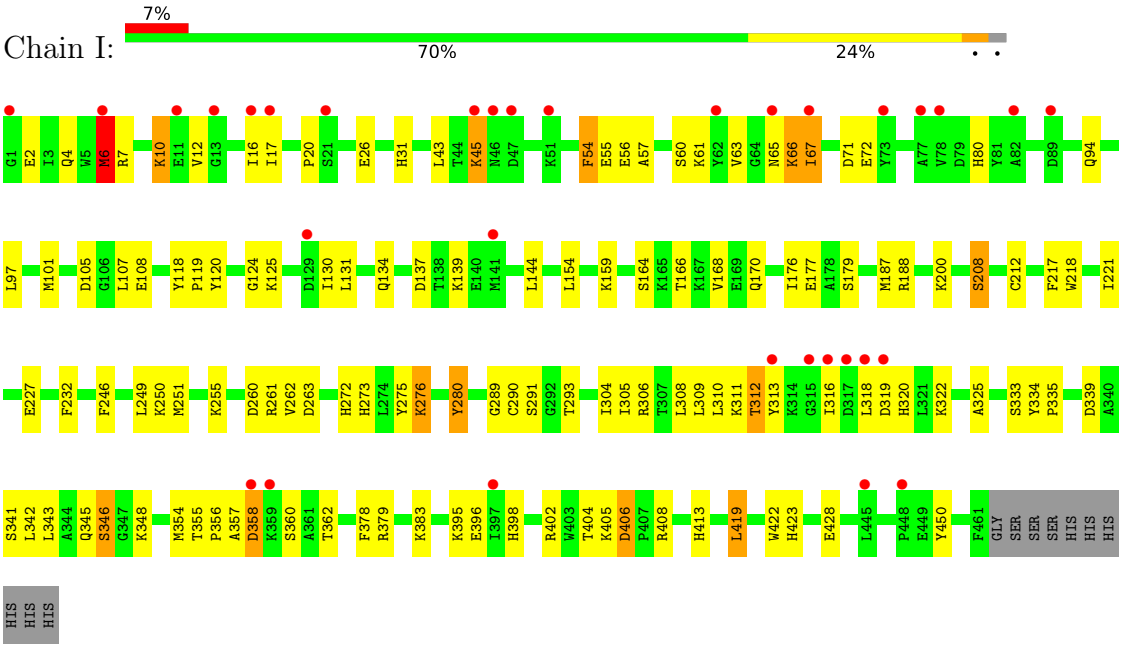
• Molecule 1: Polymerase

Chain E:



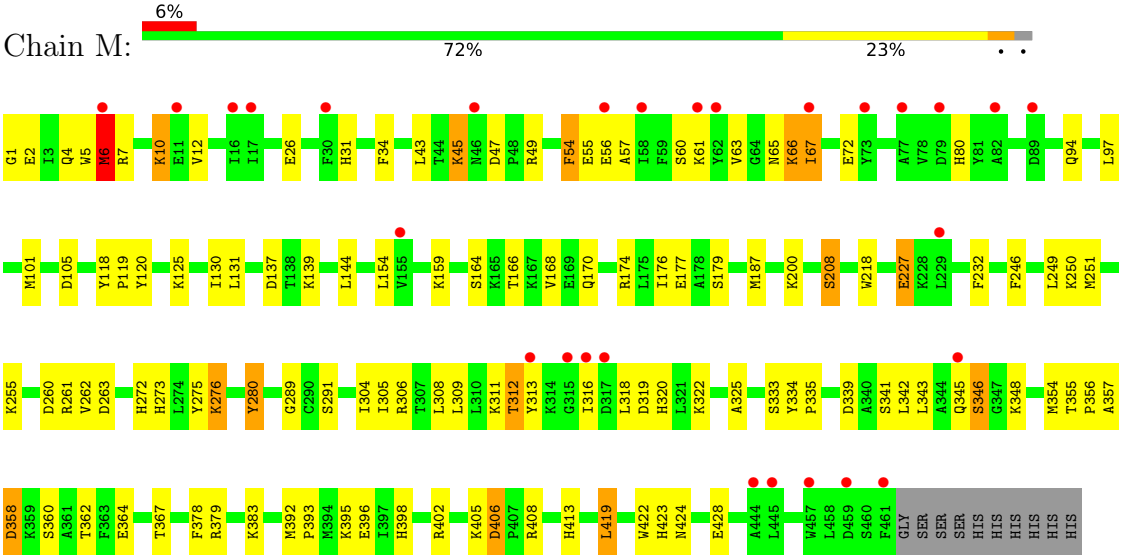
• Molecule 1: Polymerase

Chain I:



• Molecule 1: Polymerase

Chain M:



• Molecule 2: RNA (5'-R(*AP*AP*GP*UP*CP*UP*CP*CP*AP*GP*GP*UP*CP*UP*CP*UP *CP*GP*UP*CP*CP*GP*GP*AP*AP*A)-3')

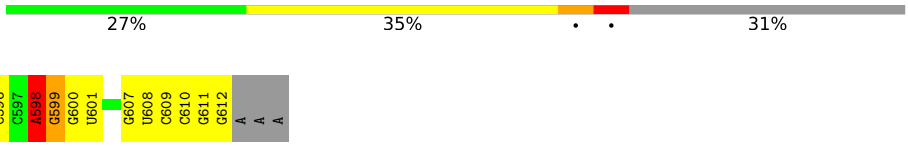
Chain B:

• Molecule 2: RNA (5'-R(*AP*AP*GP*UP*CP*UP*CP*CP*AP*GP*GP*UP*CP*UP*CP*UP *CP*GP*UP*CP*CP*GP*GP*AP*AP*A)-3')

Chain F:

• Molecule 2: RNA (5'-R(*AP*AP*GP*UP*CP*UP*CP*CP*AP*GP*GP*UP*CP*UP*CP*UP *CP*GP*UP*CP*CP*GP*GP*AP*AP*A)-3')

Chain J:



• Molecule 2: RNA (5'-R(*AP*AP*GP*UP*CP*UP*CP*CP*AP*GP*GP*UP*CP*UP*CP*UP *CP*GP*UP*CP*CP*GP*GP*AP*AP*A)-3')

Chain N:

• Molecule 3: DNA/RNA (5'-R(*GP*CP*CP*CP*GP*GP*AP*CP*GP*AP*GP*AP*GP*A)-D(P*C)-3')

Chain C:

• Molecule 3: DNA/RNA (5'-R(*GP*CP*CP*CP*GP*GP*AP*CP*GP*AP*GP*AP*GP*A)-D(P*C)-3')

Chain G:

• Molecule 4: RNA (5'-R(*GP*GP*GP*AP*GP*AP*UP*GP*A)-3')

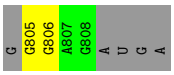
Chain D:

• Molecule 4: RNA (5'-R(*GP*GP*GP*AP*GP*AP*UP*GP*A)-3')

Chain H:

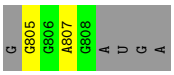
• Molecule 4: RNA (5'-R(*GP*GP*GP*AP*GP*AP*UP*GP*A)-3')

Chain L:



• Molecule 4: RNA (5'-R(*GP*GP*GP*AP*GP*AP*UP*GP*A)-3')

Chain P:

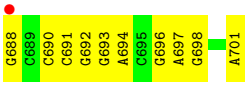


• Molecule 5: RNA (5'-R(*GP*CP*CP*CP*GP*GP*AP*CP*GP*AP*GP*AP*GP*A)-3')

Chain K:

• Molecule 5: RNA (5'-R(*GP*CP*CP*CP*GP*GP*AP*CP*GP*AP*GP*AP*GP*A)-3')

Chain O:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	60.56Å 60.56Å 192.83Å 83.57° 83.56° 77.59°	Depositor
Resolution (Å)	46.71 – 2.55 46.96 – 2.35	Depositor EDS
% Data completeness (in resolution range)	96.7 (46.71-2.55) 96.4 (46.96-2.35)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.45 (at 2.34Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.4_486)	Depositor
R, R_{free}	0.226 , 0.283 0.209 , 0.266	Depositor DCC
R_{free} test set	5344 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	61.7	Xtriage
Anisotropy	0.184	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 44.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.62$, $< L^2 > = 0.48$	Xtriage
Estimated twinning fraction	0.450 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	18452	wwPDB-VP
Average B, all atoms (Å^2)	70.0	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 47.32 % 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 1.0091e-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: DCP, GOL, ZN, IPA, PEG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
1	A	0.46	0/3787	0.63	0/5122
1	E	0.46	0/3787	0.63	1/5122 (0.0%)
1	I	0.47	0/3787	0.64	0/5122
1	M	0.47	0/3787	0.64	0/5122
2	B	0.68	0/396	1.01	0/614
2	F	0.64	0/396	0.96	0/614
2	J	0.68	0/418	1.06	1/648 (0.2%)
2	N	0.69	0/418	1.00	1/648 (0.2%)
3	C	0.67	0/361	1.04	1/562 (0.2%)
3	G	0.65	0/361	1.02	0/562
4	D	0.31	0/76	0.58	0/117
4	H	0.32	0/76	0.53	0/117
4	L	0.27	0/102	0.54	0/158
4	P	0.30	0/102	0.55	0/158
5	K	0.71	0/340	1.16	0/530
5	O	0.67	0/340	1.10	0/530
All	All	0.50	0/18534	0.72	4/25746 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	213	ASP	CB-CG-OD1	5.58	123.32	118.30
2	J	598	A	C3'-C2'-C1'	5.14	105.61	101.50
3	C	697	A	C2-N3-C4	5.13	113.17	110.60
2	N	598	A	C3'-C2'-C1'	5.05	105.54	101.50

There are no chirality outliers.

There are no planarity outliers.

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	A	3697	0	3658	98	0
1	E	3697	0	3658	88	0
1	I	3697	0	3658	99	0
1	M	3697	0	3658	97	0
2	B	357	0	184	26	0
2	F	357	0	184	21	0
2	J	377	0	194	18	0
2	N	377	0	194	19	0
3	C	322	0	167	9	0
3	G	322	0	167	7	0
4	D	68	0	34	13	0
4	H	68	0	34	7	0
4	L	91	0	45	4	0
4	P	91	0	45	4	0
5	K	303	0	156	10	0
5	O	303	0	156	7	0
6	A	1	0	0	0	0
6	E	1	0	0	0	0
6	I	1	0	0	0	0
6	M	1	0	0	0	0
7	A	12	0	24	2	0
7	E	12	0	24	3	0
7	I	8	0	16	0	0
7	M	8	0	16	4	0
8	A	14	0	20	12	0
8	I	7	0	10	9	0
9	A	6	0	8	0	0
9	I	6	0	8	0	0
10	K	24	0	12	2	0
10	O	24	0	12	3	0
11	A	94	0	0	1	0
11	B	26	0	0	3	0
11	C	9	0	0	0	0
11	D	1	0	0	0	0
11	E	100	0	0	4	0
11	F	19	0	0	0	0
11	G	4	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
11	I	85	0	0	2	0
11	J	23	0	0	2	0
11	K	11	0	0	0	0
11	L	5	0	0	0	0
11	M	89	0	0	1	0
11	N	28	0	0	1	0
11	O	8	0	0	0	0
11	P	1	0	0	0	0
All	All	18452	0	16342	493	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 14.

The worst 5 of 493 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:382:GLU:HB2	8:A:7001:PEG:H21	1.25	1.18
1:I:66:LYS:HG3	1:I:67:ILE:H	1.13	1.08
1:M:66:LYS:HG3	1:M:67:ILE:H	1.13	1.05
2:J:610:C:H3'	11:J:504:HOH:O	1.69	0.93
1:I:10:LYS:H	1:I:10:LYS:HD3	1.34	0.92

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	459/471 (98%)	433 (94%)	22 (5%)	4 (1%)	17	24
1	E	459/471 (98%)	431 (94%)	24 (5%)	4 (1%)	17	24
1	I	459/471 (98%)	420 (92%)	33 (7%)	6 (1%)	12	16

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	M	459/471 (98%)	419 (91%)	35 (8%)	5 (1%)	14 19
All	All	1836/1884 (98%)	1703 (93%)	114 (6%)	19 (1%)	15 22

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	7	ARG
1	E	7	ARG
1	I	6	MET
1	I	66	LYS
1	M	6	MET

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	A	403/412 (98%)	374 (93%)	29 (7%)	14 18
1	E	403/412 (98%)	374 (93%)	29 (7%)	14 18
1	I	403/412 (98%)	377 (94%)	26 (6%)	17 23
1	M	403/412 (98%)	378 (94%)	25 (6%)	18 24
All	All	1612/1648 (98%)	1503 (93%)	109 (7%)	16 20

5 of 109 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	451	SER
1	I	261	ARG
1	M	280	TYR
1	I	7	ARG
1	I	67	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	M	269	ASN
1	M	320	HIS
1	E	269	ASN
1	E	272	HIS
1	E	277	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	16/26 (61%)	1 (6%)	0
2	F	16/26 (61%)	1 (6%)	0
2	J	17/26 (65%)	3 (17%)	1 (5%)
2	N	17/26 (65%)	3 (17%)	1 (5%)
3	C	13/15 (86%)	1 (7%)	0
3	G	13/15 (86%)	1 (7%)	0
4	D	2/9 (22%)	1 (50%)	0
4	H	2/9 (22%)	1 (50%)	0
4	L	3/9 (33%)	0	0
4	P	3/9 (33%)	0	0
5	K	13/14 (92%)	0	0
5	O	13/14 (92%)	0	0
All	All	128/198 (64%)	12 (9%)	2 (1%)

5 of 12 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	598	A
3	C	690	C
4	D	807	A
2	F	598	A
3	G	690	C

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	J	598	A
2	N	598	A

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 21 ligands modelled in this entry, 4 are monoatomic - leaving 17 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	$\#\|Z\| > 2$
7	IPA	M	6028	-	3,3,3	0.51	0	3,3,3	0.40	0
7	IPA	E	6007	-	3,3,3	0.62	0	3,3,3	0.29	0
7	IPA	A	6013	-	3,3,3	0.60	0	3,3,3	0.24	0
8	PEG	A	7003	-	6,6,6	0.51	0	5,5,5	0.32	0
9	GOL	I	8001	-	5,5,5	0.34	0	5,5,5	0.22	0
7	IPA	M	6015	-	3,3,3	0.57	0	3,3,3	0.38	0
10	DCP	O	4001	-	23,25,29	0.94	1 (4%)	35,38,45	1.17	3 (8%)
8	PEG	A	7001	-	6,6,6	0.47	0	5,5,5	0.37	0
7	IPA	I	6029	-	3,3,3	0.57	0	3,3,3	0.25	0
10	DCP	K	4002	-	23,25,29	0.99	2 (8%)	35,38,45	1.05	2 (5%)
9	GOL	A	8002	6	5,5,5	0.43	0	5,5,5	0.35	0
7	IPA	I	6014	-	3,3,3	0.61	0	3,3,3	0.30	0
7	IPA	A	6016	-	3,3,3	0.56	0	3,3,3	0.34	0
7	IPA	E	6008	-	3,3,3	0.57	0	3,3,3	0.25	0
7	IPA	A	6017	-	3,3,3	0.54	0	3,3,3	0.30	0
7	IPA	E	6010	-	3,3,3	0.56	0	3,3,3	0.39	0
8	PEG	I	7002	-	6,6,6	0.57	0	5,5,5	0.17	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
8	PEG	A	7003	-	-	2/4/4/4	-
10	DCP	O	4001	-	-	1/16/28/34	0/2/2/2

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	PEG	A	7001	-	-	1/4/4/4	-
10	DCP	K	4002	-	-	4/16/28/34	0/2/2/2
9	GOL	A	8002	6	-	2/4/4/4	-
9	GOL	I	8001	-	-	4/4/4/4	-
8	PEG	I	7002	-	-	1/4/4/4	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	K	4002	DCP	PB-O3B	2.33	1.63	1.54
10	K	4002	DCP	C6-C5	2.22	1.40	1.35
10	O	4001	DCP	PB-O3B	2.19	1.63	1.54

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
10	O	4001	DCP	O5'-PA-O1A	-2.97	97.46	109.07
10	K	4002	DCP	PA-O3A-PB	-2.35	124.77	132.83
10	K	4002	DCP	O4'-C1'-N1	2.17	111.74	107.86
10	O	4001	DCP	PA-O3A-PB	-2.15	125.44	132.83
10	O	4001	DCP	O2B-PB-O1B	2.00	118.52	110.68

There are no chirality outliers.

5 of 15 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	I	8001	GOL	O1-C1-C2-O2
9	I	8001	GOL	O1-C1-C2-C3
9	I	8001	GOL	C1-C2-C3-O3
10	K	4002	DCP	O4'-C4'-C5'-O5'
10	K	4002	DCP	C5'-O5'-PA-O1A

There are no ring outliers.

12 monomers are involved in 35 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	M	6028	IPA	1	0
7	E	6007	IPA	1	0
7	A	6013	IPA	1	0
8	A	7003	PEG	5	0

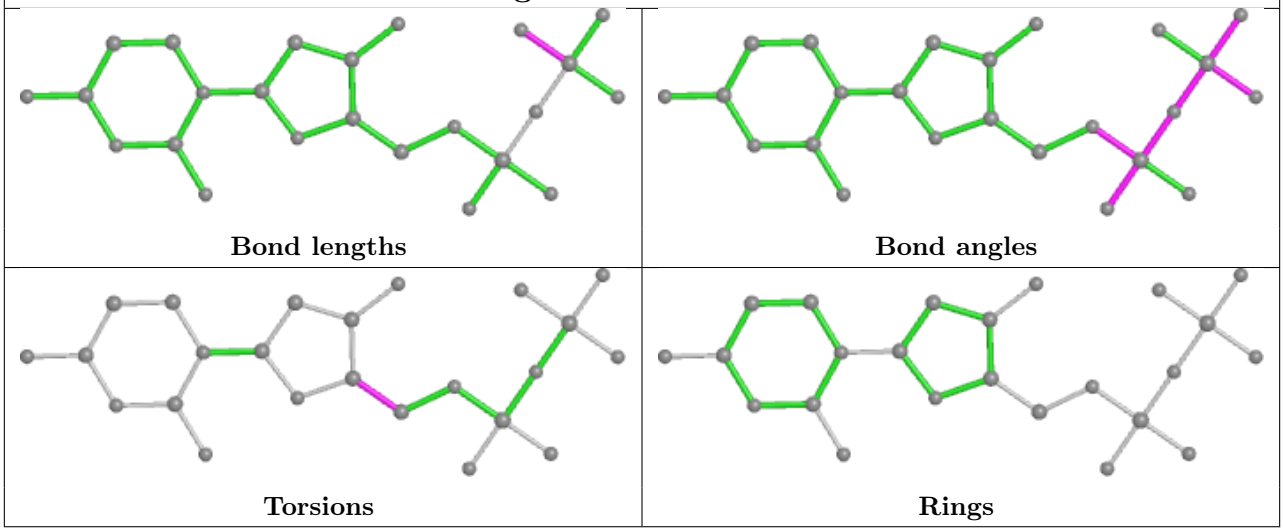
Continued on next page...

Continued from previous page...

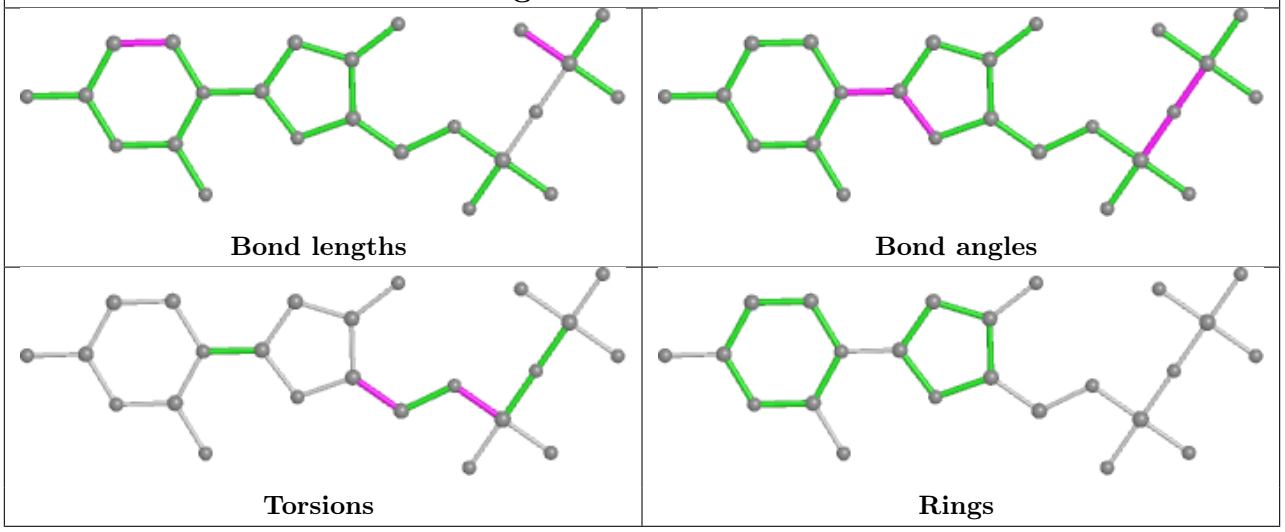
Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	M	6015	IPA	3	0
10	O	4001	DCP	3	0
8	A	7001	PEG	7	0
10	K	4002	DCP	2	0
7	E	6008	IPA	1	0
7	A	6017	IPA	1	0
7	E	6010	IPA	1	0
8	I	7002	PEG	9	0

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 DCP O 4001



Ligand DCP K 4002



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	A	461/471 (97%)	0.23	30 (6%) 18 22	34, 60, 100, 132	0
1	E	461/471 (97%)	0.24	25 (5%) 25 30	34, 60, 100, 132	0
1	I	461/471 (97%)	0.29	32 (6%) 16 20	32, 62, 107, 127	0
1	M	461/471 (97%)	0.23	28 (6%) 21 25	33, 63, 107, 127	0
2	B	17/26 (65%)	-0.59	0 100 100	46, 68, 162, 170	0
2	F	17/26 (65%)	-0.65	0 100 100	47, 67, 162, 170	0
2	J	18/26 (69%)	-0.56	0 100 100	42, 85, 153, 175	0
2	N	18/26 (69%)	-0.54	0 100 100	41, 85, 152, 175	0
3	C	15/15 (100%)	-0.64	0 100 100	50, 68, 134, 151	0
3	G	15/15 (100%)	-0.68	0 100 100	51, 68, 133, 151	0
4	D	3/9 (33%)	0.38	0 100 100	133, 133, 142, 153	0
4	H	3/9 (33%)	0.91	0 100 100	133, 133, 143, 152	0
4	L	4/9 (44%)	-0.12	0 100 100	114, 117, 127, 163	0
4	P	4/9 (44%)	0.39	0 100 100	114, 117, 126, 163	0
5	K	14/14 (100%)	-0.58	0 100 100	42, 69, 162, 168	0
5	O	14/14 (100%)	-0.47	1 (7%) 16 19	42, 68, 162, 168	0
All	All	1986/2082 (95%)	0.20	116 (5%) 23 27	32, 62, 112, 175	0

The worst 5 of 116 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	M	316	ILE	8.7
1	I	316	ILE	8.0
1	A	358	ASP	7.2
1	I	82	ALA	6.8
1	A	316	ILE	6.6

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

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)

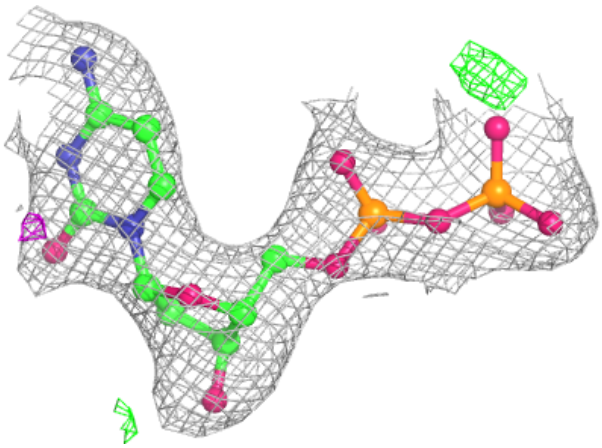
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

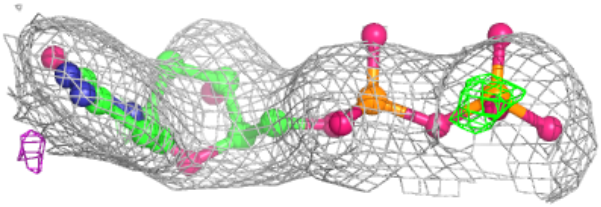
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
7	IPA	I	6029	4/4	0.75	0.40	108,117,120,121	0
9	GOL	A	8002	6/6	0.78	0.17	97,100,108,110	0
8	PEG	A	7001	7/7	0.82	0.20	91,101,108,108	0
9	GOL	I	8001	6/6	0.84	0.20	63,84,91,94	0
7	IPA	E	6008	4/4	0.85	0.21	90,91,103,103	0
7	IPA	M	6015	4/4	0.85	0.18	54,77,81,82	0
7	IPA	M	6028	4/4	0.85	0.16	75,86,88,90	0
10	DCP	K	4002	24/28	0.85	0.18	38,67,135,159	0
10	DCP	O	4001	24/28	0.88	0.14	36,60,128,150	0
7	IPA	E	6010	4/4	0.92	0.22	70,76,80,84	0
7	IPA	A	6013	4/4	0.93	0.41	66,78,82,83	0
7	IPA	A	6017	4/4	0.93	0.17	77,85,87,89	0
8	PEG	I	7002	7/7	0.94	0.15	60,64,82,85	0
8	PEG	A	7003	7/7	0.95	0.20	52,66,85,89	0
7	IPA	A	6016	4/4	0.95	0.10	71,73,83,85	0
7	IPA	E	6007	4/4	0.96	0.31	57,66,71,83	0
7	IPA	I	6014	4/4	0.96	0.24	54,78,79,86	0
6	ZN	M	2001	1/1	0.97	0.11	68,68,68,68	1
6	ZN	E	2003	1/1	0.97	0.04	79,79,79,79	1
6	ZN	A	2004	1/1	0.98	0.04	80,80,80,80	1
6	ZN	I	2002	1/1	0.98	0.15	70,70,70,70	1

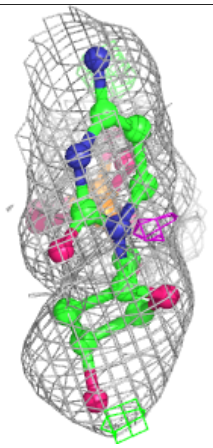
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around DCP K 4002:

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

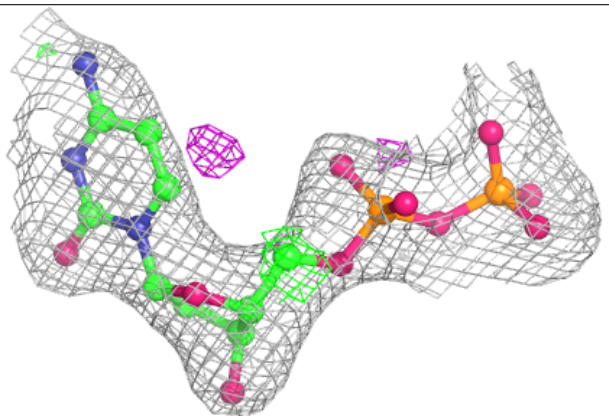


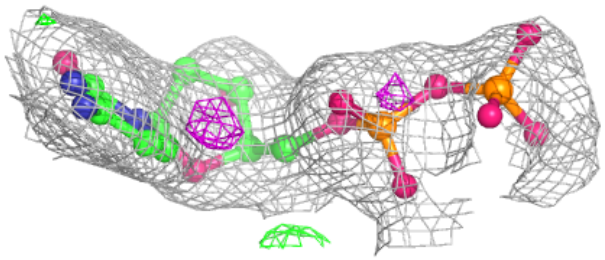


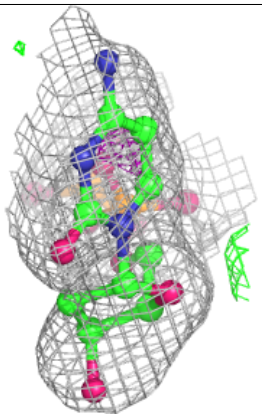


Electron density around DCP O 4001:

$2mF_o$-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.