W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Oct 3, 2021 – 08:49 AM EDT

PDB ID	:	3LZJ
Title	:	RB69 DNA Polymerase (Y567A) ternary complex with dCTP Opposite 7,8-Dihydro-8-oxoguanine
Authors	:	Wang, M.; Beckman, J.; Blaha, G.; Wang, J.; Konigsberg, W.H.
Deposited on	:	2010-03-01
Resolution	:	2.05 Å(reported)

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

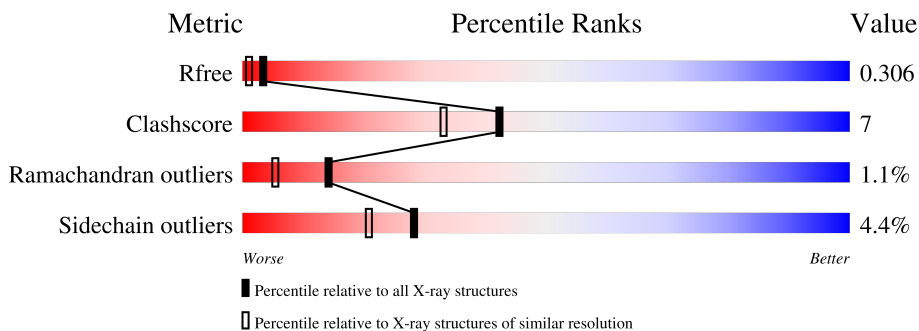|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

1   Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.05 Å.

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

Metric	Percentile Ranks	Value
Rfree		0.306
Clashscore		7
Ramachandran outliers		1.1%
Sidechain outliers		4.4%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1692 (2.04-2.04)
Clashscore	141614	1773 (2.04-2.04)
Ramachandran outliers	138981	1752 (2.04-2.04)
Sidechain outliers	138945	1752 (2.04-2.04)

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%

Mol	Chain	Length	Quality of chain
1	A	903	87% 11% •
2	T	18	56% 22% 22%
3	P	13	38% 46% 15%

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
2	8OG	T	4	X	-	-	-

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 8619 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	903	Total 7385	C 4742	N 1232	O 1378	S 33	0	2	0

There are 3 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
A	567	ALA	TYR	engineered mutation	UNP Q38087

- Molecule 2 is a DNA chain called DNA (5'-D(P*TP*CP*AP*(8OG)P*GP*TP*AP*AP*G P*CP*AP*GP*TP*CP*CP*GP*CP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	18	Total 372	C 175	N 71	O 108	P 18	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	P	13	Total 262	C 126	N 48	O 76	P 12	0	0	0

- Molecule 4 is CYTIDINE-5'-TRIPHOSPHATE (three-letter code: CTP) (formula: $C_9H_{16}N_3O_{14}P_3$).

CTP
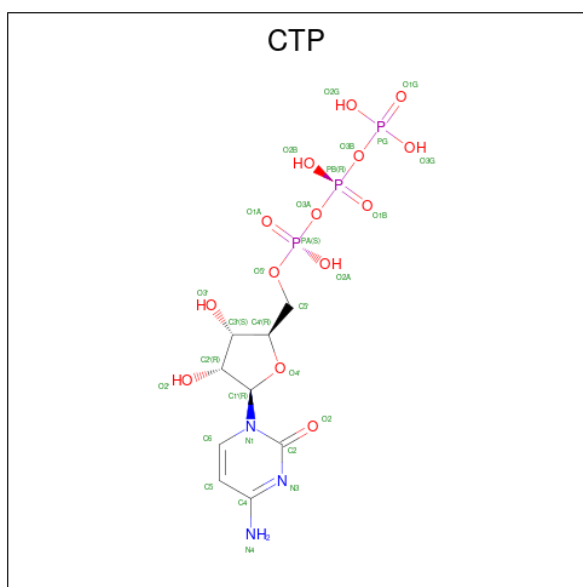
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
4	A	1	28	9	3	13	3	0	0

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
5	A	4	4	4	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	A	514	514	514	0	0
6	T	38	Total 38	O 38	0	0
6	P	16	Total 16	O 16	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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.
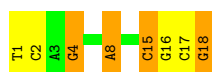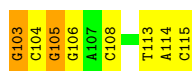
- Molecule 1: DNA polymerase
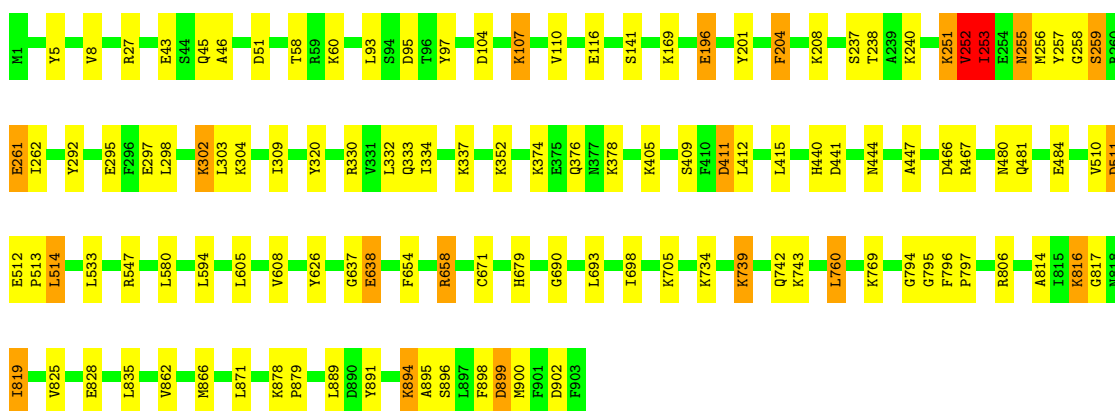
- Molecule 2: DNA (5'-D(P*TP*CP*AP*(8OG)P*GP*TP*AP*AP*GP*CP*AP*GP*TP*CP*CP*GP*CP*G)-3')

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

4   Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	76.39Å 121.09Å 122.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.81 – 2.05 38.79 – 2.05	Depositor EDS
% Data completeness (in resolution range)	90.8 (38.81-2.05) 90.8 (38.79-2.05)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.26 (at 2.05Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.205 , 0.251 0.265 , 0.306	Depositor DCC
R_{free} test set	3317 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	32.1	Xtriage
Anisotropy	1.077	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 36.1	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.014 for -h,l,k	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	8619	wwPDB-VP
Average B, all atoms (Å^2)	11.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.82% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 8OG, CTP, CA, DOC

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	0.47	0/7569	0.57	0/10227
2	T	0.78	0/390	1.53	8/597 (1.3%)
3	P	0.77	0/273	1.41	5/420 (1.2%)
All	All	0.51	0/8232	0.71	13/11244 (0.1%)

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
2	T	2	0

There are no bond length outliers.

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	T	18	DG	O4'-C4'-C3'	-7.93	101.24	106.00
3	P	108	DC	O4'-C1'-N1	6.51	112.56	108.00
2	T	18	DG	O4'-C1'-N9	6.49	112.54	108.00
2	T	8	DA	O4'-C1'-N9	6.32	112.42	108.00
2	T	18	DG	C1'-O4'-C4'	-6.18	103.92	110.10

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	T	4	8OG	C4',C3'

There are no planarity outliers.

5.2　Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	7385	0	7282	102	0
2	T	372	0	200	3	0
3	P	262	0	148	5	0
4	A	28	0	10	0	0
5	A	4	0	0	0	0
6	A	514	0	0	11	0
6	P	16	0	0	0	0
6	T	38	0	0	0	0
All	All	8619	0	7640	108	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 7.

The worst 5 of 108 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:251:LYS:HB3	1:A:252:VAL:HA	1.16	1.15
1:A:898:PHE:HA	1:A:899:ASP:HB3	1.29	1.15
1:A:510:VAL:HA	1:A:511:ASP:CB	1.95	0.97
1:A:251:LYS:HB3	1:A:252:VAL:CA	1.96	0.95
1:A:258:GLY:HA3	1:A:259:SER:HB2	1.46	0.94

There are no symmetry-related clashes.

5.3　Torsion angles ⓘ

5.3.1　Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	903/903 (100%)	855 (95%)	38 (4%)	10 (1%)	14 5

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	253	ILE
1	A	259	SER
1	A	511	ASP
1	A	638	GLU
1	A	251	LYS

5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	801/799 (100%)	766 (96%)	35 (4%)	28 21

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	769	LYS
1	A	816	LYS
1	A	871	LEU
1	A	295	GLU
1	A	261	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	376	GLN
1	A	444	ASN
1	A	823	GLN
1	A	761	GLN
1	A	773	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
3	DOC	P	115	2,3	14,19,20	0.69	0	13,26,29	2.60	5 (38%)
2	8OG	T	4	2	18,25,26	2.78	4 (22%)	21,37,40	3.36	12 (57%)

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	DOC	P	115	2,3	-	1/4/18/19	0/2/2/2
2	8OG	T	4	2	2/2/4/4	1/3/21/22	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	T	4	8OG	C3'-C4'	-7.20	1.33	1.53
2	T	4	8OG	C5'-C4'	-5.97	1.33	1.51
2	T	4	8OG	O3'-C3'	-4.80	1.33	1.43
2	T	4	8OG	O6-C6	4.61	1.36	1.24

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	T	4	8OG	C2'-C1'-N9	-9.63	105.49	116.01
3	P	115	DOC	C4'-O4'-C1'	-6.35	103.81	109.81

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	P	115	DOC	C2-N3-C4	4.75	121.15	116.34
2	T	4	8OG	N3-C2-N1	-4.39	121.36	127.22
2	T	4	8OG	O3'-C3'-C4'	3.99	125.38	110.10

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	T	4	8OG	C4'
2	T	4	8OG	C3'

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	T	4	8OG	O4'-C4'-C5'-O5'
3	P	115	DOC	C3'-C4'-C5'-O5'

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 5 ligands modelled in this entry, 4 are monoatomic - leaving 1 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	CTP	A	904	5	23,29,30	0.83	0	30,45,47	1.44	2 (6%)

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	CTP	A	904	5	-	5/19/34/38	0/2/2/2

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	904	CTP	C2-N3-C4	5.90	122.33	116.34
4	A	904	CTP	C5-C4-N3	-2.48	118.85	121.72

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	904	CTP	PA-O3A-PB-O2B
4	A	904	CTP	PB-O3B-PG-O1G
4	A	904	CTP	PB-O3B-PG-O2G
4	A	904	CTP	PB-O3B-PG-O3G
4	A	904	CTP	PA-O3A-PB-O1B

There are no ring outliers.

No monomer is involved in short contacts.

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 CTP A 904
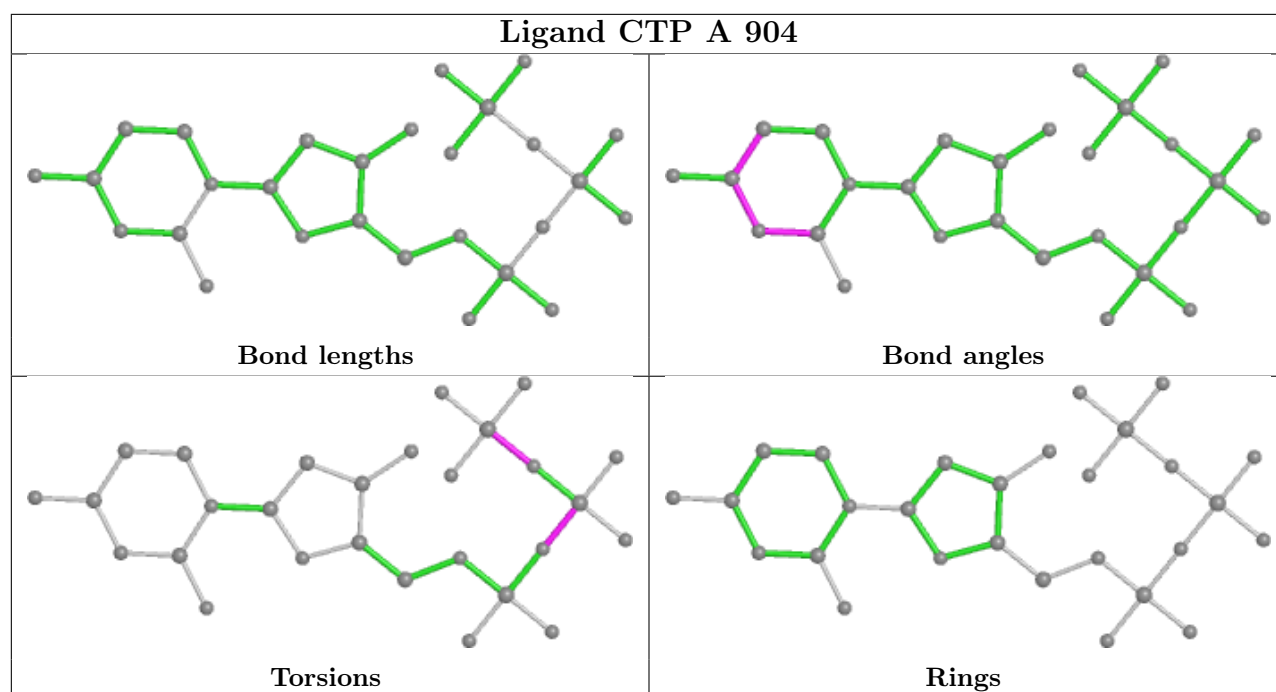
5.7　Other polymers ⓘ

There are no such residues in this entry.

5.8　Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

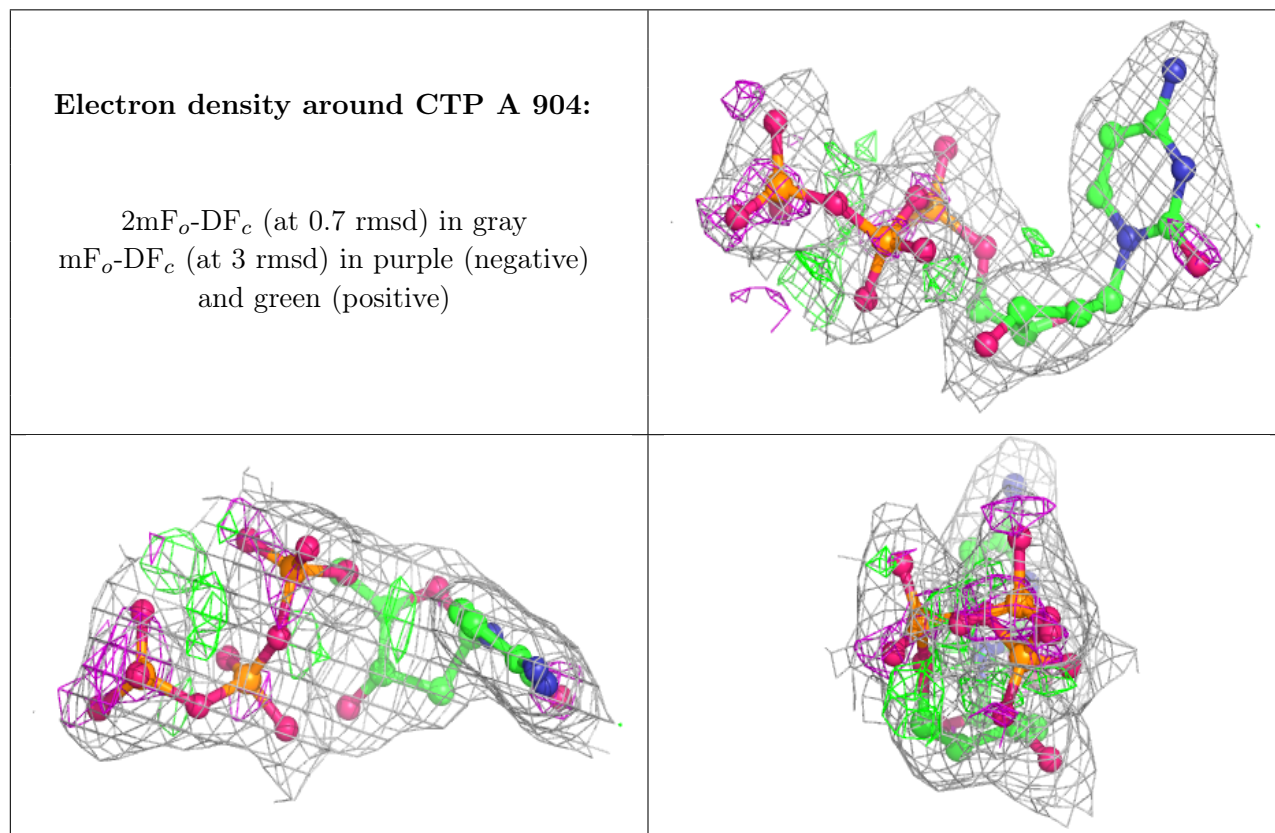Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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 CTP A 904:

$2mF_o$-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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.