wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 22, 2023 – 11:37 AM JST

PDB ID	:	7W9S
Title	:	Crystal structure of the enterovirus 71 polymerase elongation complex (C1S3 form)
Authors	:	Wang, M.; Gong, P.
Deposited on	:	2021-12-10
Resolution	:	2.53 Å(reported)

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

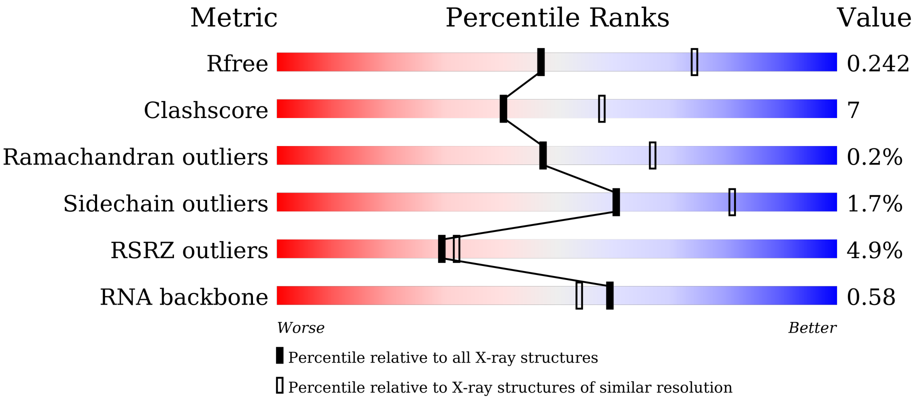
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance

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

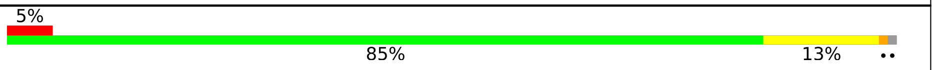
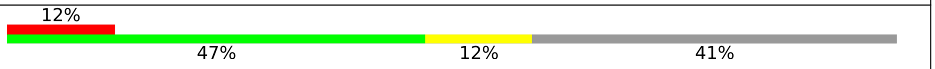
The reported resolution of this entry is 2.53 Å.

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	5743 (2.54-2.50)
Clashscore	141614	6463 (2.54-2.50)
Ramachandran outliers	138981	6335 (2.54-2.50)
Sidechain outliers	138945	6337 (2.54-2.50)
RSRZ outliers	127900	5630 (2.54-2.50)
RNA backbone	3102	1020 (2.86-2.18)

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	468	
2	B	35	
3	C	17	

2 Entry composition ⓘ

There are 8 unique types of molecules in this entry. The entry contains 4048 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 Genome polyprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	463	3514	2267	586	643	18	0	0	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	291	MET	CYS	engineered mutation	UNP E5RPG3
A	463	HIS	-	expression tag	UNP E5RPG3
A	464	HIS	-	expression tag	UNP E5RPG3
A	465	HIS	-	expression tag	UNP E5RPG3
A	466	HIS	-	expression tag	UNP E5RPG3
A	467	HIS	-	expression tag	UNP E5RPG3
A	468	HIS	-	expression tag	UNP E5RPG3

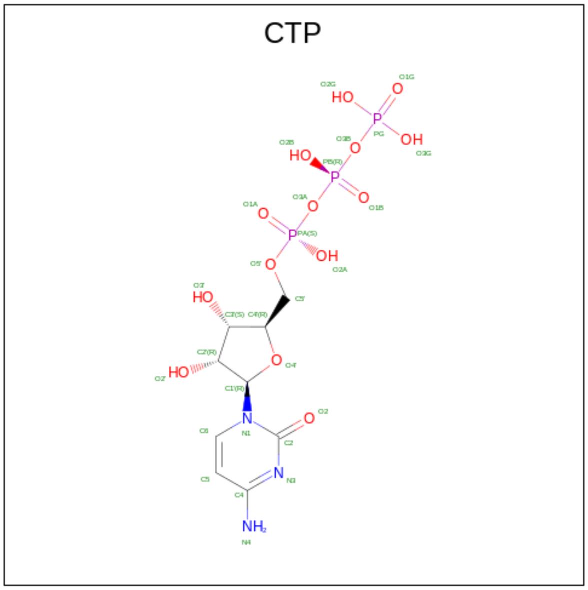
- Molecule 2 is a RNA chain called RNA (35-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	12	249	111	37	89	12	0	0	0

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

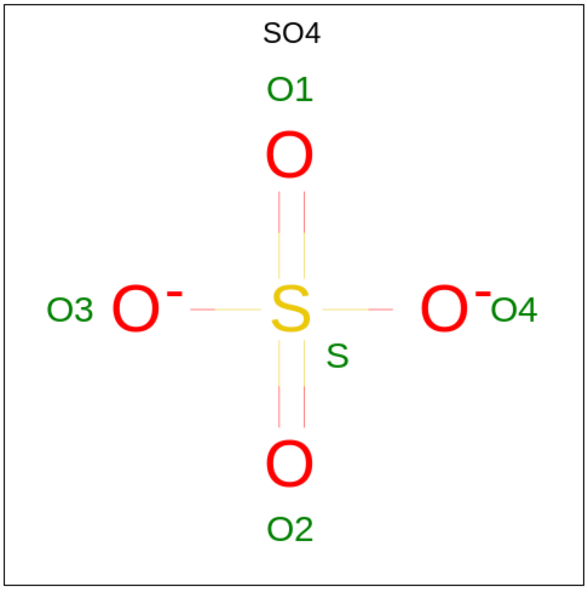
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	C	10	222	99	48	65	10	0	0	0

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



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
4	A	1	29	9	3	14	3	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
5	A	1	5	4	1	0	0
5	A	1	5	4	1	0	0
5	A	1	5	4	1	0	0

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

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

- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	2	Total Mg 2 2	0	0

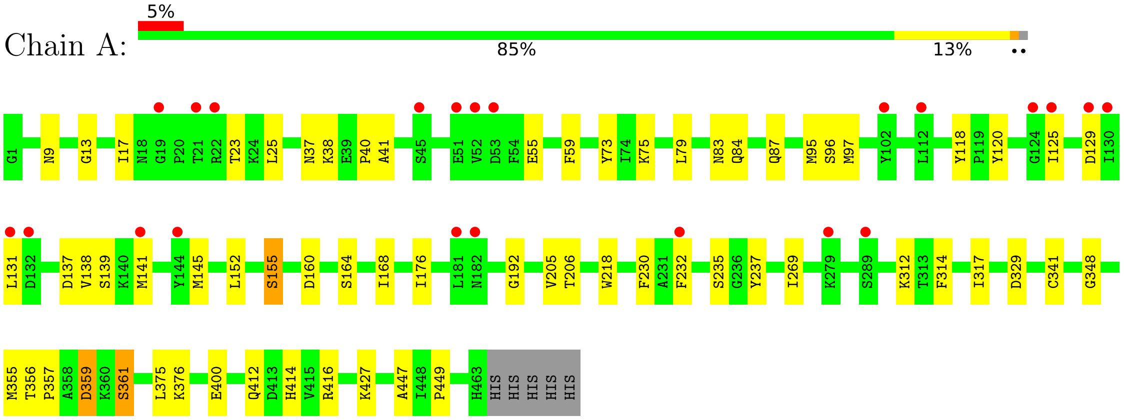
- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	15	Total O 15 15	0	0
8	B	1	Total O 1 1	0	0

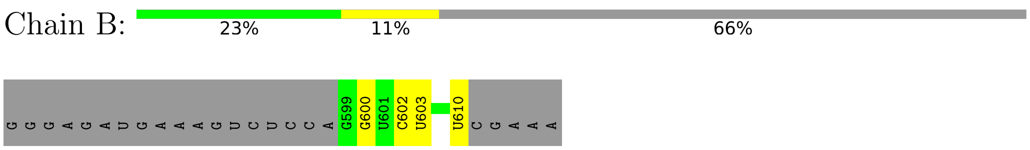
3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

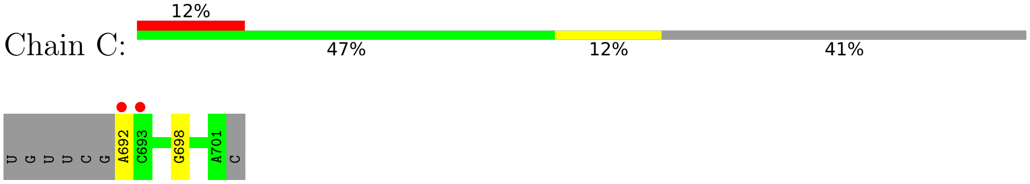
- Molecule 1: Genome polyprotein



- Molecule 2: RNA (35-MER)



- Molecule 3: RNA (5’-R(*UP*GP*UP*UP*CP*GP*AP*CP*GP*AP*GP*AP*GP*AP*GP*AP*CP*CP*U)-3’)



4 Data and refinement statistics

Property	**Value**	**Source**
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	63.99Å 77.66Å 150.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.94 – 2.53 46.93 – 2.50	Depositor EDS
% Data completeness (in resolution range)	95.5 (46.94-2.53) 93.7 (46.93-2.50)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.58 (at 2.51Å)	Xtriage
Refinement program	PHENIX 1.10_2155	Depositor
R, R_{free}	0.200 , 0.241 0.201 , 0.242	Depositor DCC
R_{free} test set	1300 reflections (5.18%)	wwPDB-VP
Wilson B-factor (Å^2)	68.9	Xtriage
Anisotropy	0.327	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 51.9	EDS
L-test for twinning[2]	$< \vert L \vert > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4048	wwPDB-VP
Average B, all atoms (Å^2)	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.23% 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: SO4, CTP, MG, ZN

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	0.48	0/3607	0.61	0/4920
2	B	0.57	0/275	1.10	0/425
3	C	0.71	0/250	1.12	0/389
All	All	0.51	0/4132	0.70	0/5734

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

Mol	**Chain**	**#Chirality outliers**	**#Planarity outliers**
1	A	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	**Chain**	**Res**	**Type**	**Group**
1	A	232	PHE	Peptide

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	3514	0	3324	52	0
2	B	249	0	128	3	0
3	C	222	0	111	2	0
4	A	29	0	12	1	0
5	A	15	0	0	0	0
6	A	1	0	0	0	0
7	A	2	0	0	0	0
8	A	15	0	0	0	0
8	B	1	0	0	0	0
All	All	4048	0	3575	54	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 54 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:23:THR:HG22	1:A:25:LEU:H	1.32	0.92
2:B:610:U:O4	3:C:692:A:N6	2.19	0.71
1:A:23:THR:HG22	1:A:25:LEU:N	2.09	0.66
1:A:17:ILE:HD13	1:A:155:SER:HB3	1.79	0.64
1:A:137:ASP:OD2	1:A:139:SER:N	2.35	0.59

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	461/468 (98%)	446 (97%)	14 (3%)	1 (0%)	47	67

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	359	ASP

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	356/412 (86%)	350 (98%)	6 (2%)	60 81

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	235	SER
1	A	361	SER
1	A	427	LYS
1	A	155	SER
1	A	96	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	412	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	11/35 (31%)	0	0
3	C	9/17 (52%)	0	0
All	All	20/52 (38%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 7 ligands modelled in this entry, 3 are monoatomic - leaving 4 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	501	7	26,30,30	0.82	0	39,47,47	0.93	1 (2%)
5	SO4	A	504	-	4,4,4	0.17	0	6,6,6	0.36	0
5	SO4	A	502	-	4,4,4	0.15	0	6,6,6	0.22	0
5	SO4	A	503	-	4,4,4	0.24	0	6,6,6	0.20	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	CTP	A	501	7	-	6/22/38/38	0/2/2/2

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	501	CTP	PB-O3B-PG	-2.76	123.34	132.83

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	501	CTP	PB-O3B-PG-O2G
4	A	501	CTP	O4'-C4'-C5'-O5'
4	A	501	CTP	PA-O3A-PB-O2B
4	A	501	CTP	C3'-C4'-C5'-O5'
4	A	501	CTP	PB-O3B-PG-O3G

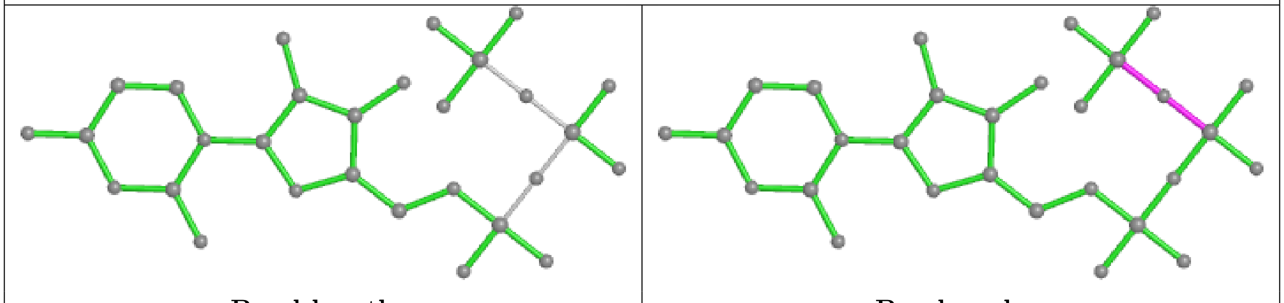
There are no ring outliers.

1 monomer is involved in 1 short contact:

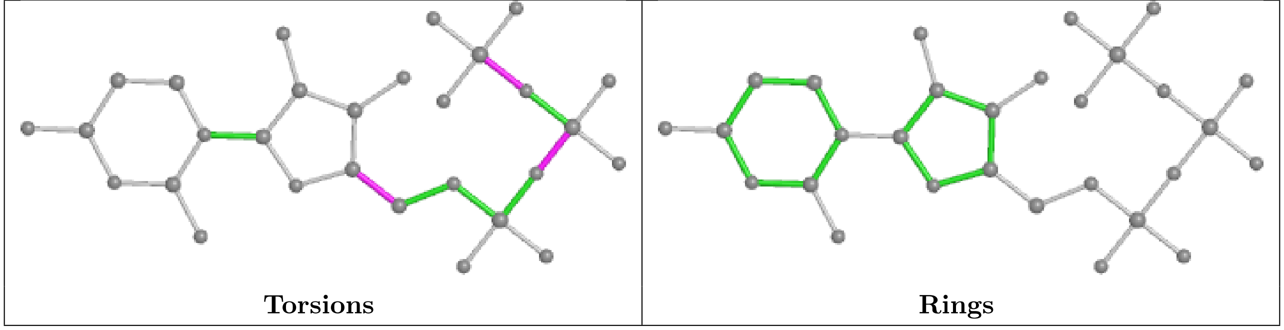
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	501	CTP	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand CTP A 501



Bond lengths **Bond angles**



Torsions **Rings**

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	463/468 (98%)	0.29	22 (4%)	30	33	50, 67, 89, 99	0
2	B	12/35 (34%)	0.21	0	100	100	58, 67, 115, 127	0
3	C	10/17 (58%)	0.89	2 (20%)	1	1	56, 69, 127, 134	0
All	All	485/520 (93%)	0.30	24 (4%)	29	32	50, 67, 92, 134	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	131	LEU	6.2
1	A	125	ILE	3.6
1	A	22	ARG	3.6
1	A	52	VAL	3.3
1	A	45	SER	3.3

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

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

6.3 Carbohydrates ⓘ

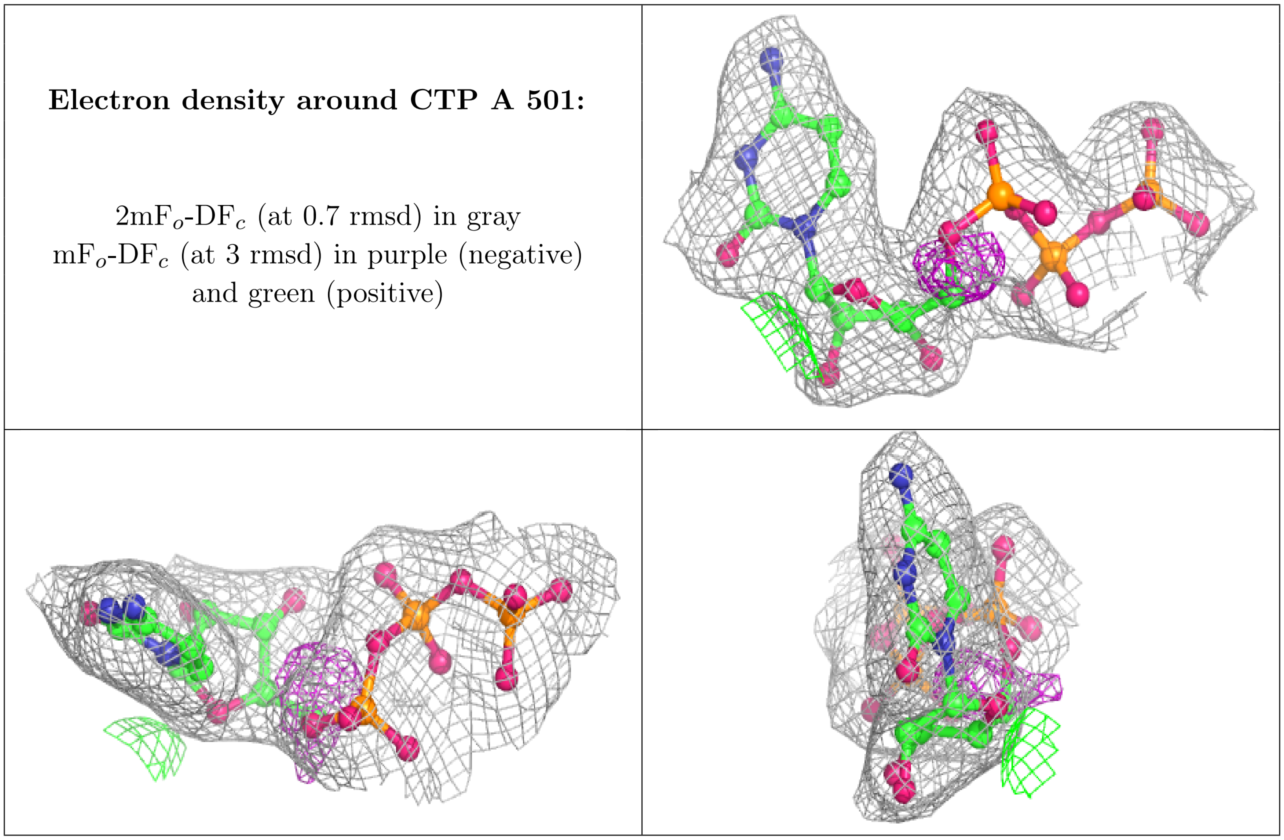
There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	SO4	A	503	5/5	0.82	0.31	74,76,88,90	5
7	MG	A	506	1/1	0.92	0.20	65,65,65,65	0
4	CTP	A	501	29/29	0.96	0.24	56,63,76,83	0
7	MG	A	507	1/1	0.97	0.23	65,65,65,65	0
5	SO4	A	502	5/5	0.98	0.10	72,74,77,81	0
5	SO4	A	504	5/5	0.98	0.16	52,54,64,68	5
6	ZN	A	505	1/1	0.99	0.15	61,61,61,61	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 CTP A 501:

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

6.5 Other polymers

There are no such residues in this entry.