



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 5, 2023 – 04:37 AM EDT

PDB ID : 6UU3  
Title : E. coli sigma-S transcription initiation complex with a 4-nt RNA and a CTP ("Old" crystal soaked with GTP, ATP, CTP, and ddTTP for 30 minutes)  
Authors : Zuo, Y.; De, S.; Steitz, T.A.  
Deposited on : 2019-10-30  
Resolution : 4.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : **FAILED**  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : **FAILED**  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.35.1

## 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 4.00 Å.

There are no overall percentile quality scores available for this entry.

MolProbity and EDS failed to run properly - the sequence quality summary graphics cannot be shown.

## 2 Entry composition

There are 12 unique types of molecules in this entry. The entry contains 29064 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-directed RNA polymerase subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	230	1787	1112	317	352	6	0	0	0
1	BBB	228	1767	1100	312	349	6	0	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	-6	ALA	-	expression tag	UNP A0A377D9Q8
AAA	-5	HIS	-	expression tag	UNP A0A377D9Q8
AAA	-4	HIS	-	expression tag	UNP A0A377D9Q8
AAA	-3	HIS	-	expression tag	UNP A0A377D9Q8
AAA	-2	HIS	-	expression tag	UNP A0A377D9Q8
AAA	-1	HIS	-	expression tag	UNP A0A377D9Q8
AAA	0	HIS	-	expression tag	UNP A0A377D9Q8
BBB	-6	ALA	-	expression tag	UNP A0A377D9Q8
BBB	-5	HIS	-	expression tag	UNP A0A377D9Q8
BBB	-4	HIS	-	expression tag	UNP A0A377D9Q8
BBB	-3	HIS	-	expression tag	UNP A0A377D9Q8
BBB	-2	HIS	-	expression tag	UNP A0A377D9Q8
BBB	-1	HIS	-	expression tag	UNP A0A377D9Q8
BBB	0	HIS	-	expression tag	UNP A0A377D9Q8

- Molecule 2 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	CCC	1341	10576	6636	1842	2055	43	0	0	0

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	DDD	1362	10568	6633	1887	1998	50	0	0	0

- Molecule 4 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	EEE	79	627	382	118	126	1	0	0	0

- Molecule 5 is a protein called RNA polymerase sigma factor RpoS.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	FFF	277	2253	1411	415	423	4	0	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
FFF	2	GLY	SER	conflict	UNP P13445
FFF	33	GLU	GLN	conflict	UNP P13445
FFF	329	LEU	-	expression tag	UNP P13445
FFF	330	GLU	-	expression tag	UNP P13445
FFF	331	HIS	-	expression tag	UNP P13445
FFF	332	HIS	-	expression tag	UNP P13445
FFF	333	HIS	-	expression tag	UNP P13445
FFF	334	HIS	-	expression tag	UNP P13445
FFF	335	HIS	-	expression tag	UNP P13445
FFF	336	HIS	-	expression tag	UNP P13445

- Molecule 6 is a DNA chain called Synthetic DNA 50-MER (promoter non-template strand).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
6	111	30	619	294	114	181	30	0	0	0

- Molecule 7 is a DNA chain called Synthetic DNA 50-MER (promoter template strand).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
7	222	36	738	352	137	214	35	0	0	0

- Molecule 8 is a RNA chain called RNA 4-mer.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
8	333	3	77	30	15	27	5	0	0	0

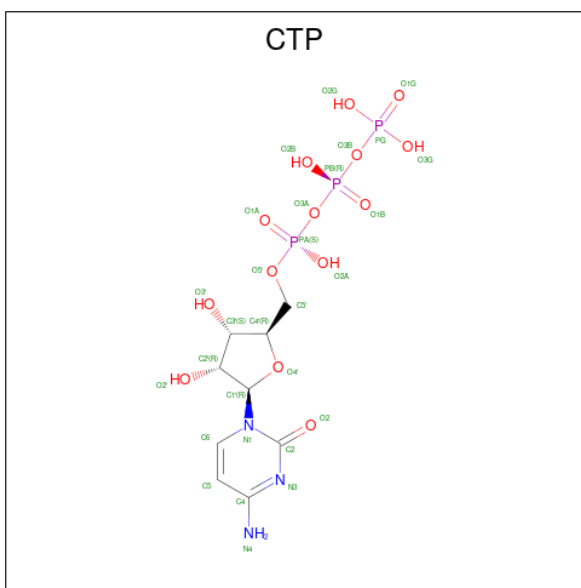
- Molecule 9 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	DDD	2	Total	Zn	0	0
			2	2		

- Molecule 10 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

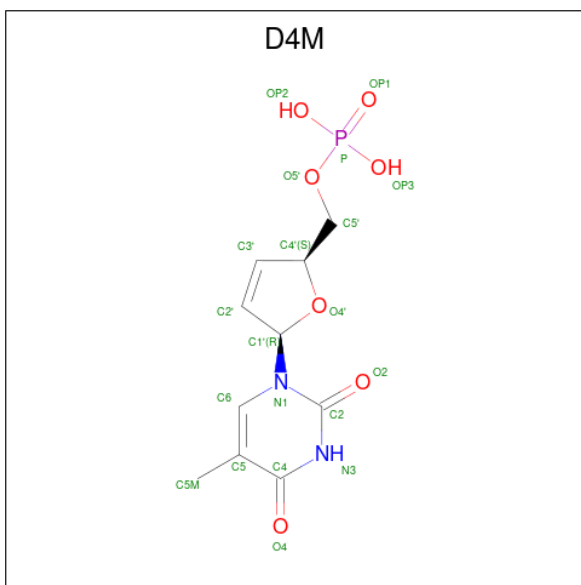
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	DDD	2	Total	Mg	0	0
			2	2		

- Molecule 11 is CYTIDINE-5'-TRIPHOSPHATE (three-letter code: CTP) (formula: C<sub>9</sub>H<sub>16</sub>N<sub>3</sub>O<sub>14</sub>P<sub>3</sub>).



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

- Molecule 12 is [(5R)-5-(5-METHYL-2,4-DIOXO-3,4-DIHYDROPYRIMIDIN-1(2H)-YL)-2,5-DIHYDROFURAN-2-YL]METHYL DIHYDROGEN PHOSPHATE (three-letter code: D4M) (formula: C<sub>10</sub>H<sub>13</sub>N<sub>2</sub>O<sub>7</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
12	333	1	19	10	2	6	1	0	0

MolProbity and EDS failed to run properly - this section is therefore empty.

### 3 Data and refinement statistics

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	132.56Å 153.58Å 231.21Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.08 – 4.00	Depositor
% Data completeness (in resolution range)	98.7 (49.08-4.00)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.07 (at 4.00Å)	Xtrriage
Refinement program	REFMAC 5.8.0257	Depositor
R, $R_{free}$	0.303 , 0.380	Depositor
Wilson B-factor (Å <sup>2</sup> )	175.8	Xtrriage
Anisotropy	0.600	Xtrriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.32$ , $\langle L^2 \rangle = 0.16$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	29064	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	299.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.77% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 4 Model quality [i](#)

### 4.1 Standard geometry [i](#)

MolProbity failed to run properly - this section is therefore empty.

### 4.2 Too-close contacts [i](#)

MolProbity failed to run properly - this section is therefore empty.

### 4.3 Torsion angles [i](#)

#### 4.3.1 Protein backbone [i](#)

MolProbity failed to run properly - this section is therefore empty.

#### 4.3.2 Protein sidechains [i](#)

MolProbity failed to run properly - this section is therefore empty.

#### 4.3.3 RNA [i](#)

MolProbity failed to run properly - this section is therefore empty.

### 4.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 4.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 4.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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
12	D4M	333	101	-	17,20,21	2.30	6 (35%)	23,28,31	2.65	12 (52%)
11	CTP	DDD	1505	10	26,30,30	0.48	0	39,47,47	0.74	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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
12	D4M	333	101	-	-	0/7/18/19	0/2/2/2
11	CTP	DDD	1505	10	-	5/22/38/38	0/2/2/2

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
12	333	101	D4M	C2'-C3'	6.66	1.51	1.32
12	333	101	D4M	C6-C5	2.55	1.38	1.34
12	333	101	D4M	C4-C5	2.48	1.48	1.44
12	333	101	D4M	C6-N1	-2.13	1.34	1.38
12	333	101	D4M	C1'-C2'	2.13	1.52	1.49
12	333	101	D4M	C2-N1	2.08	1.41	1.38

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
12	333	101	D4M	C4-N3-C2	-4.88	121.04	127.35
12	333	101	D4M	C5-C4-N3	4.57	119.21	115.31
12	333	101	D4M	O4-C4-C5	-4.35	119.86	124.90
12	333	101	D4M	N3-C2-N1	4.27	120.55	114.89
12	333	101	D4M	O4'-C4'-C3'	3.83	108.44	103.06
12	333	101	D4M	O4'-C1'-C2'	3.55	108.86	104.91
12	333	101	D4M	C1'-C2'-C3'	-3.21	105.78	109.68
12	333	101	D4M	C5-C6-N1	-3.08	120.17	123.34
12	333	101	D4M	C2'-C1'-N1	-2.99	110.16	113.42
12	333	101	D4M	C5'-C4'-C3'	-2.83	110.43	115.61
12	333	101	D4M	C5M-C5-C4	2.20	121.18	118.77
12	333	101	D4M	O2-C2-N1	-2.08	120.02	122.79

There are no chirality outliers.

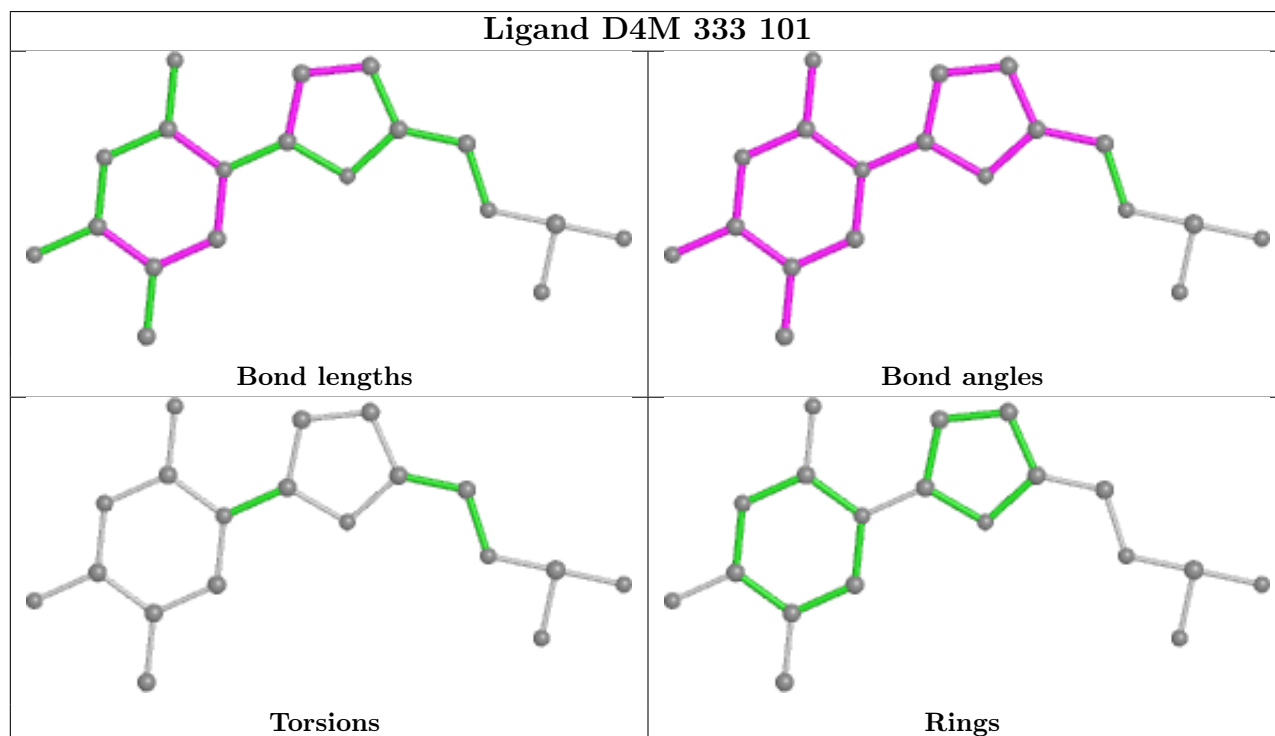
All (5) torsion outliers are listed below:

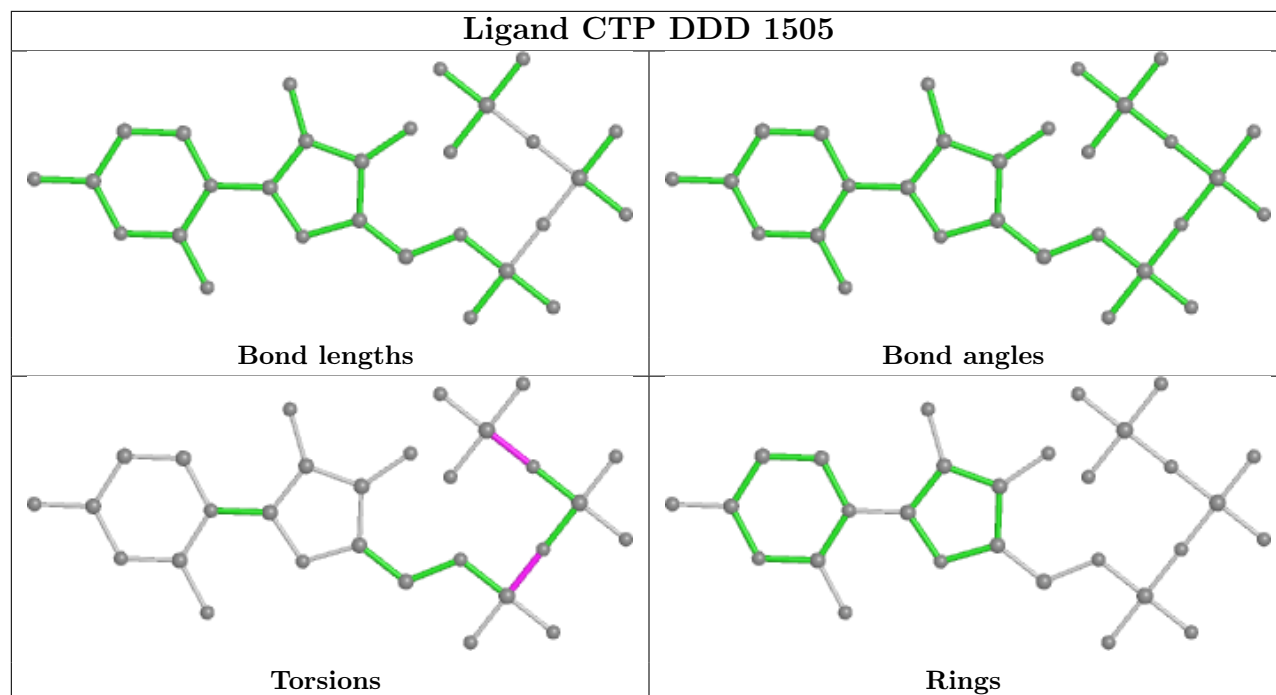
Mol	Chain	Res	Type	Atoms
11	DDD	1505	CTP	PB-O3A-PA-O2A
11	DDD	1505	CTP	PB-O3B-PG-O1G
11	DDD	1505	CTP	PB-O3A-PA-O1A
11	DDD	1505	CTP	PB-O3B-PG-O2G
11	DDD	1505	CTP	PB-O3B-PG-O3G

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 than 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.





#### 4.7 Other polymers [i](#)

There are no such residues in this entry.

#### 4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 5 Fit of model and data [i](#)

### 5.1 Protein, DNA and RNA chains [i](#)

EDS failed to run properly - this section is therefore empty.

### 5.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS failed to run properly - this section is therefore empty.

### 5.3 Carbohydrates [i](#)

EDS failed to run properly - this section is therefore empty.

### 5.4 Ligands [i](#)

EDS failed to run properly - this section is therefore empty.

### 5.5 Other polymers [i](#)

EDS failed to run properly - this section is therefore empty.