Full wwPDB X-ray Structure Validation Report ⓘ

Oct 2, 2023 – 11:44 AM EDT

PDB ID	:	6N61
Title	:	Escherichia coli RNA polymerase sigma70-holoenzyme bound to upstream fork promoter DNA and Capistruin
Authors	:	Braffman, N.; Hauver, J.; Campbell, E.A.; Darst, S.A.
Deposited on	:	2018-11-24
Resolution	:	3.25 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*
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)
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
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 (i)

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

The reported resolution of this entry is 3.25 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 13 unique types of molecules in this entry. The entry contains 28613 atoms, of which 23 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
1	A	228	Total	C	N	O	S	0	0	0
			1709	1067	300	336	6			
1	B	217	Total	C	N	O	S	0	0	0
			1658	1035	290	327	6			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	235	GLU	-	expression tag	UNP P0A7Z4
A	236	VAL	-	expression tag	UNP P0A7Z4
A	237	LEU	-	expression tag	UNP P0A7Z4
A	238	PHE	-	expression tag	UNP P0A7Z4
A	239	GLN	-	expression tag	UNP P0A7Z4
B	235	GLU	-	expression tag	UNP P0A7Z4
B	236	VAL	-	expression tag	UNP P0A7Z4
B	237	LEU	-	expression tag	UNP P0A7Z4
B	238	PHE	-	expression tag	UNP P0A7Z4
B	239	GLN	-	expression tag	UNP P0A7Z4

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1332	Total	C	N	O	S	0	0	0
			10489	6581	1829	2036	43			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1239	Total	C	N	O	S	0	0	0
			9649	6061	1733	1807	48			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	1	VAL	-	expression tag	UNP P0A8T7
D	1408	LEU	-	expression tag	UNP P0A8T7
D	1409	GLU	-	expression tag	UNP P0A8T7

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	395	Total	C	N	O	S	0	0	0
			3197	1993	578	603	23			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	149	ASN	ASP	conflict	UNP Q0P6L9

- Molecule 6 is a DNA chain called non-template strand DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	N	29	Total	C	N	O	P	0	0	0
			595	284	106	176	29			

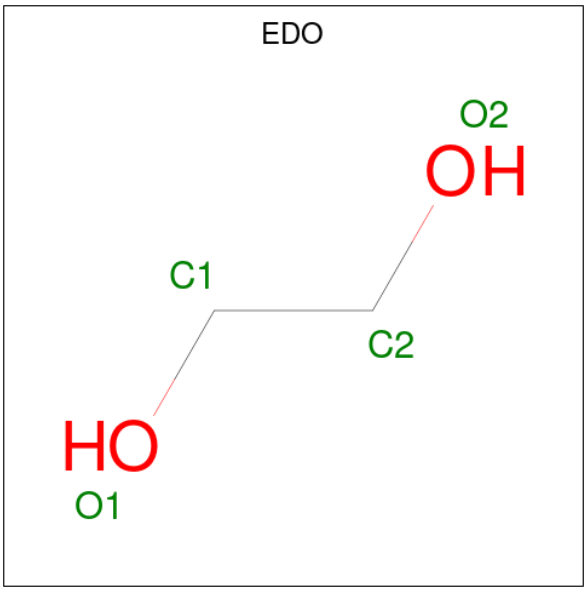
- Molecule 7 is a DNA chain called template strand DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	T	24	Total	C	N	O	P	0	0	0
			492	233	94	141	24			

- Molecule 8 is a protein called Capistruin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
8	I	17	Total	C	N	O	0	0	0
			126	80	24	22			

- Molecule 9 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	C	1	Total	C	H	O	0	0
			10	2	6	2		

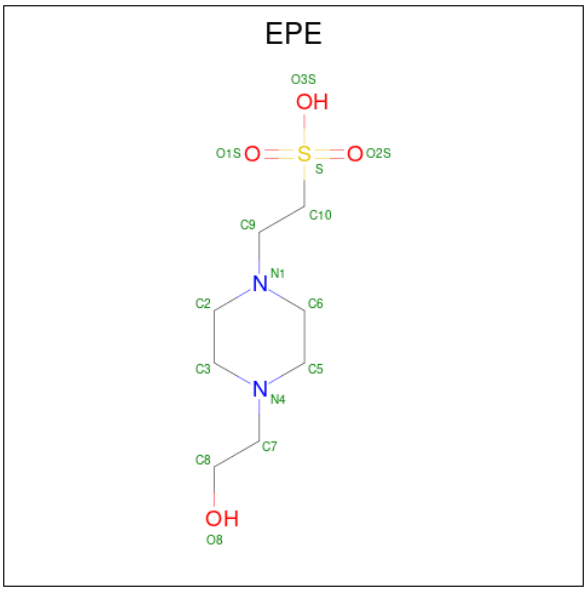
- Molecule 10 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	D	1	Total	Mg	0	0
			1	1		

- Molecule 11 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	D	2	Total	Zn	0	0
			2	2		

- Molecule 12 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
12	D	1	Total	C	H	N	O	S	0	0
			32	8	17	2	4	1		

- Molecule 13 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	A	1	Total	O	0	0
			1	1		
13	B	2	Total	O	0	0
			2	2		
13	C	13	Total	O	0	0
			13	13		
13	D	3	Total	O	0	0
			3	3		
13	E	1	Total	O	0	0
			1	1		
13	F	5	Total	O	0	0
			5	5		
13	T	1	Total	O	0	0
			1	1		

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

3 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	172.89Å 172.89Å 385.01Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.45 – 3.25	Depositor
% Data completeness (in resolution range)	99.4 (49.45-3.25)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.07 (at 3.19Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.273 , 0.316	Depositor
Wilson B-factor (Å^2)	99.1	Xtriage
Anisotropy	0.118	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	28613	wwPDB-VP
Average B, all atoms (Å^2)	115.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.65% 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.

4 Model quality

4.1 Standard geometry

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

4.2 Too-close contacts

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

4.3 Torsion angles

4.3.1 Protein backbone

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

4.3.2 Protein sidechains

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

4.3.3 RNA

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

4.4 Non-standard residues in protein, DNA, RNA chains

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

4.5 Carbohydrates

There are no monosaccharides in this entry.

4.6 Ligand geometry

Of 5 ligands modelled in this entry, 3 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	EPE	D	1504	-	15,15,15	0.90	1 (6%)	18,20,20	2.09	6 (33%)
9	EDO	C	1401	-	3,3,3	0.44	0	2,2,2	0.58	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
12	EPE	D	1504	-	-	3/9/19/19	0/1/1/1
9	EDO	C	1401	-	-	0/1/1/1	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
12	D	1504	EPE	C10-S	3.09	1.81	1.77

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
12	D	1504	EPE	C7-N4-C5	4.36	122.38	111.23
12	D	1504	EPE	O2S-S-C10	3.98	111.71	106.92
12	D	1504	EPE	C5-N4-C3	3.13	115.88	108.83
12	D	1504	EPE	C7-N4-C3	2.85	118.52	111.23
12	D	1504	EPE	O3S-S-C10	2.64	110.04	105.77
12	D	1504	EPE	C6-N1-C2	2.39	114.21	108.83

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
12	D	1504	EPE	C8-C7-N4-C3
12	D	1504	EPE	C9-C10-S-O3S
12	D	1504	EPE	C9-C10-S-O2S

There are no ring outliers.

No monomer is involved in short contacts.

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

5.1 Protein, DNA and RNA chains

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

5.2 Non-standard residues in protein, DNA, RNA chains

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

5.3 Carbohydrates

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

5.4 Ligands

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

5.5 Other polymers

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