Full wwPDB X-ray Structure Validation Report (i)

Oct 5, 2023 – 02:02 AM EDT

PDB ID	:	6UU8
Title	:	E. coli mutant sigma-S transcription initiation complex with a 7-nt RNA ("Fresh" mutant crystal soaked with GTP, UTP, and CTP for 30 minutes)
Authors	:	Zuo, Y.; De, S.; Steitz, T.A.
Deposited on	:	2019-10-30
Resolution	:	4.40 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 (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	:	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 4.40 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 (i)

There are 11 unique types of molecules in this entry. The entry contains 28931 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 P0A7Z4
AAA	-5	HIS	-	expression tag	UNP P0A7Z4
AAA	-4	HIS	-	expression tag	UNP P0A7Z4
AAA	-3	HIS	-	expression tag	UNP P0A7Z4
AAA	-2	HIS	-	expression tag	UNP P0A7Z4
AAA	-1	HIS	-	expression tag	UNP P0A7Z4
AAA	0	HIS	-	expression tag	UNP P0A7Z4
BBB	-6	ALA	-	expression tag	UNP P0A7Z4
BBB	-5	HIS	-	expression tag	UNP P0A7Z4
BBB	-4	HIS	-	expression tag	UNP P0A7Z4
BBB	-3	HIS	-	expression tag	UNP P0A7Z4
BBB	-2	HIS	-	expression tag	UNP P0A7Z4
BBB	-1	HIS	-	expression tag	UNP P0A7Z4
BBB	0	HIS	-	expression tag	UNP P0A7Z4

- 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	268	2186	1370	405	407	4	0	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
FFF	2	GLY	SER	conflict	UNP A0A377K1M2
FFF	219	GLY	ILE	engineered mutation	UNP A0A377K1M2
FFF	221	ALA	SER	engineered mutation	UNP A0A377K1M2
FFF	329	LEU	-	expression tag	UNP A0A377K1M2
FFF	330	GLU	-	expression tag	UNP A0A377K1M2
FFF	331	HIS	-	expression tag	UNP A0A377K1M2
FFF	332	HIS	-	expression tag	UNP A0A377K1M2
FFF	333	HIS	-	expression tag	UNP A0A377K1M2
FFF	334	HIS	-	expression tag	UNP A0A377K1M2
FFF	335	HIS	-	expression tag	UNP A0A377K1M2
FFF	336	HIS	-	expression tag	UNP A0A377K1M2

- 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	29	595	283	107	176	29	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	32	652	312	117	192	31	0	0	0

- Molecule 8 is a RNA chain called RNA 7-mer (de novo synthesized).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
8	333	7	160	67	27	57	9	0	0	0

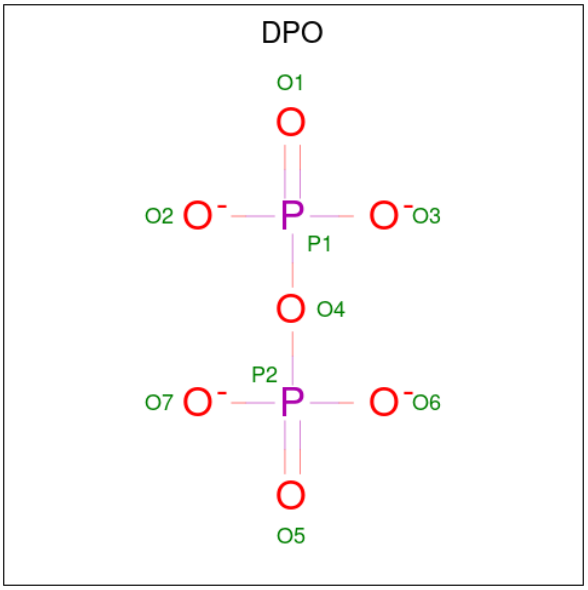
- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
9	CCC	1	1	1	0	0
9	DDD	1	1	1	0	0

- Molecule 10 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 11 is DIPHOSPHATE (three-letter code: DPO) (formula: O_7P_2).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	P		
11	DDD	1	9	7	2	0	0

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 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	133.46Å 154.50Å 235.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.38 – 4.40	Depositor
% Data completeness (in resolution range)	99.0 (49.38-4.40)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.07 (at 4.45Å)	Xtriage
Refinement program	REFMAC 5.8.0257	Depositor
R, R_{free}	0.357 , 0.380	Depositor
Wilson B-factor (Å^2)	163.9	Xtriage
Anisotropy	0.949	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.30$, $< L^2 > = 0.14$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	28931	wwPDB-VP
Average B, all atoms (Å^2)	293.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.27% 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 (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 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	$\#\|Z\| > 2$
11	DPO	DDD	1504	-	6,8,8	0.72	0	13,13,13	0.81	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
11	DPO	DDD	1504	-	-	0/6/6/6	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

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 (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.