



wwPDB X-ray Structure Validation Summary Report ⓘ

May 23, 2020 – 02:11 pm BST

PDB ID : 1TWF
Title : RNA polymerase II complexed with UTP at 2.3 Å resolution
Authors : Westover, K.D.; Bushnell, D.A.; Kornberg, R.D.
Deposited on : 2004-06-30
Resolution : 2.30 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 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 : 2.11
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) : 2.11

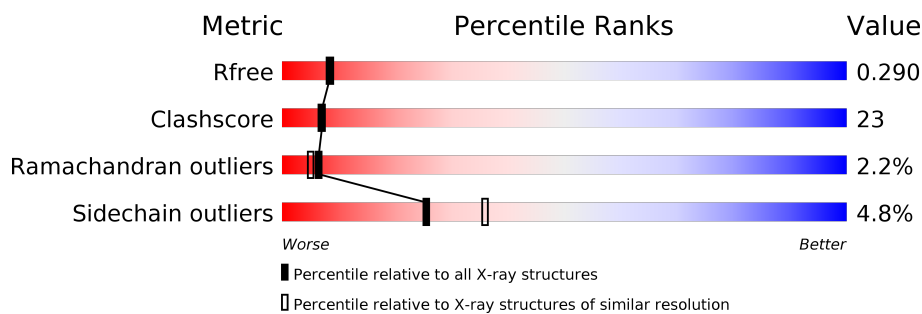
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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.





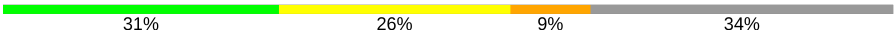
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)

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 on 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	1733	<div> <div>47%</div> <div>31%</div> <div>•</div> <div>18%</div> </div>
2	B	1224	<div> <div>51%</div> <div>36%</div> <div>•</div> <div>11%</div> </div>
3	C	318	<div> <div>44%</div> <div>37%</div> <div>•</div> <div>16%</div> </div>
4	E	215	<div> <div>66%</div> <div>33%</div> <div>•</div> </div>
5	F	155	<div> <div>33%</div> <div>21%</div> <div>•</div> <div>46%</div> </div>
6	H	146	<div> <div>45%</div> <div>40%</div> <div>7%</div> <div>9%</div> </div>
7	I	122	<div> <div>68%</div> <div>29%</div> <div>•</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	J	70	 44% 40% 9% 7%
9	K	120	 54% 38% 5% 5%
10	L	70	 31% 26% 9% 34%

2 Entry composition [i](#)

There are 13 unique types of molecules in this entry. The entry contains 28318 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 II largest subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1419	Total	C	N	O	S	0	0	0
			11154	7023	1952	2118	61			

- Molecule 2 is a protein called DNA-directed RNA polymerase II 140 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1094	Total	C	N	O	S	0	0	0
			8711	5525	1519	1614	53			

- Molecule 3 is a protein called DNA-directed RNA polymerase II 45 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	266	Total	C	N	O	S	0	0	0
			2095	1317	348	417	13			

- Molecule 4 is a protein called DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	215	Total	C	N	O	S	0	0	0
			1760	1116	310	322	12			

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	84	Total	C	N	O	S	0	0	0
			679	434	115	127	3			

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	H	133	Total	C	N	O	S	0	0	0
			1068	673	180	211	4			

- Molecule 7 is a protein called DNA-directed RNA polymerase II 14.2 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	I	122	Total	C	N	O	S	0	0	0
			997	613	182	191	11			

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III 8.3 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	J	65	Total	C	N	O	S	0	0	0
			532	339	93	94	6			

- Molecule 9 is a protein called DNA-directed RNA polymerase II 13.6 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	K	114	Total	C	N	O	S	0	0	0
			919	590	156	171	2			

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	L	46	Total	C	N	O	S	0	0	0
			364	224	72	64	4			

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	J	1	Total	Zn	0	0
			1	1		
11	B	1	Total	Zn	0	0
			1	1		
11	I	2	Total	Zn	0	0
			2	2		
11	C	1	Total	Zn	0	0
			1	1		
11	A	2	Total	Zn	0	0
			2	2		

Continued on next page...

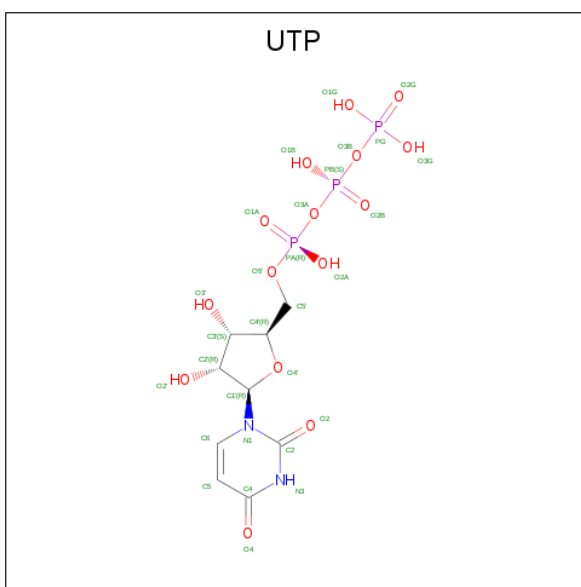
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	L	1	Total	Zn	0	0
			1	1		

- Molecule 12 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	A	2	Total	Mn	0	0
			2	2		

- Molecule 13 is URIDINE 5'-TRIPHOSPHATE (three-letter code: UTP) (formula: C₉H₁₅N₂O₁₅P₃).

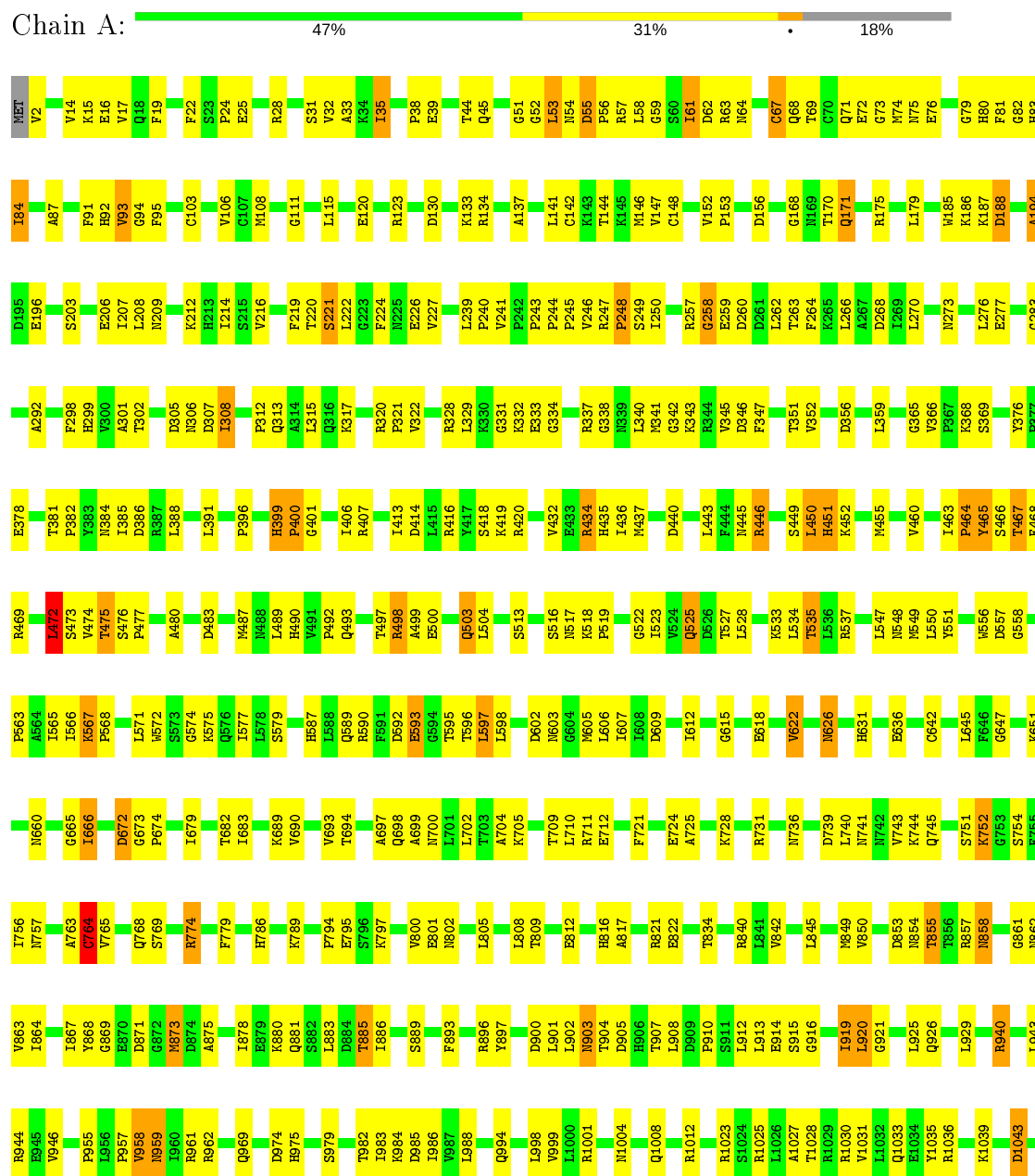


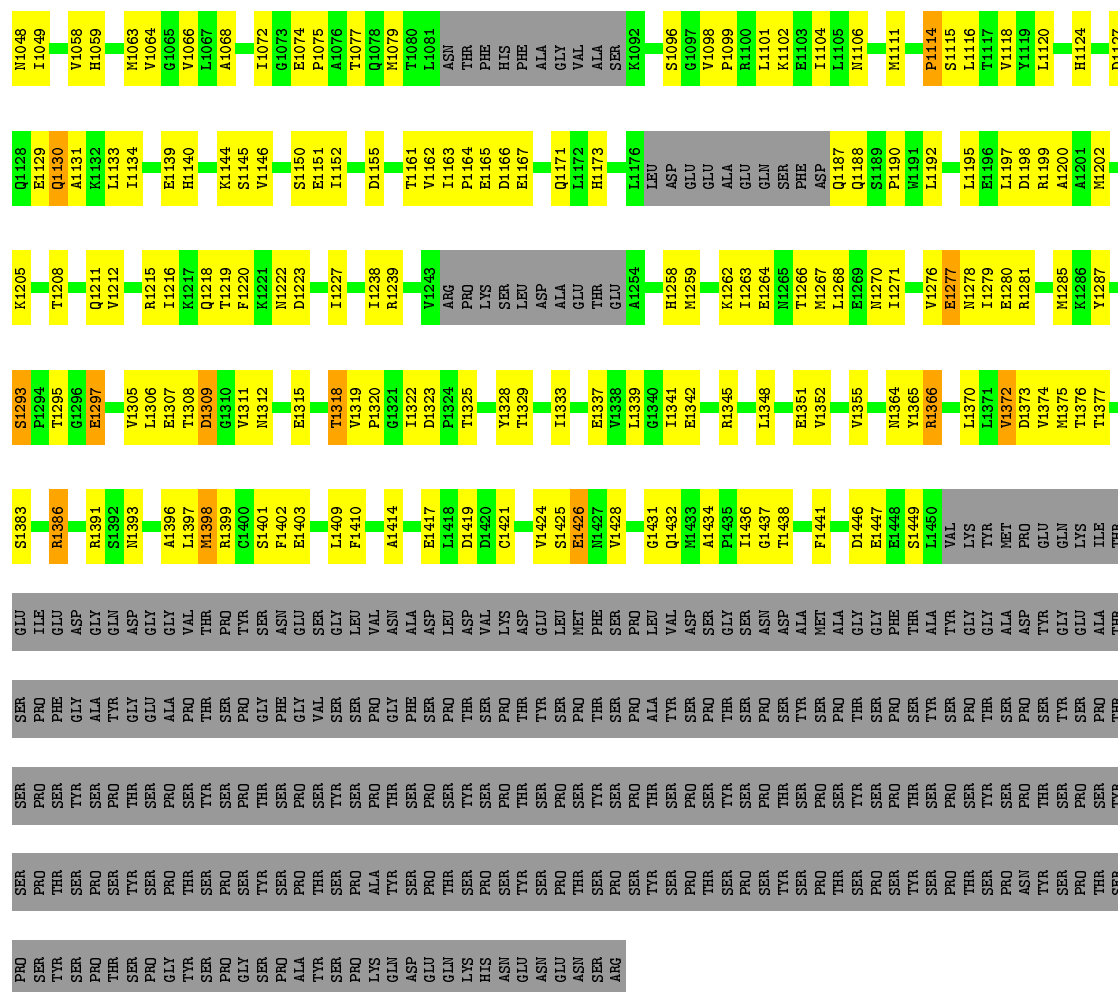
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
13	B	1	Total	C	N	O	P	
			29	9	2	15	3	0

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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.

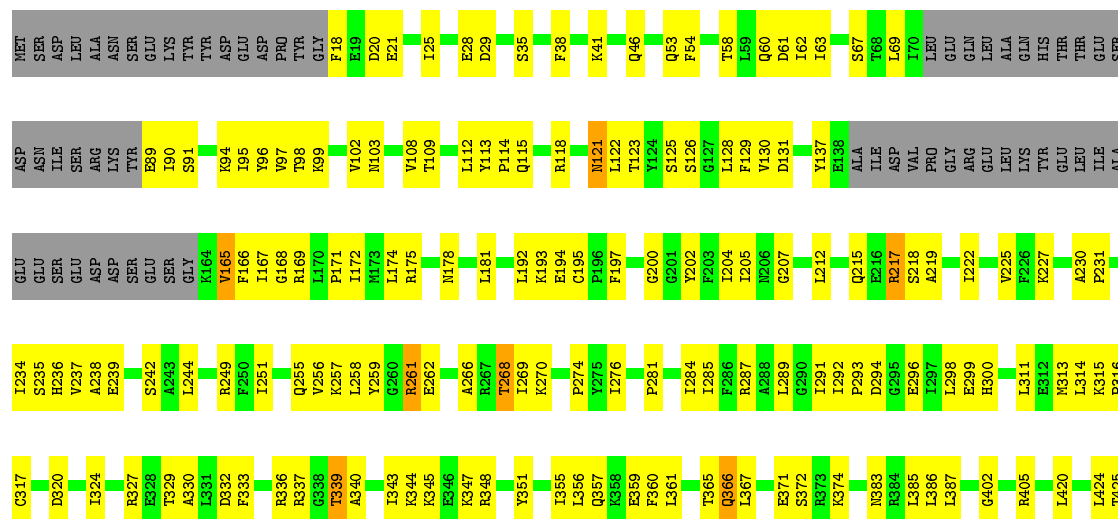
- Molecule 1: DNA-directed RNA polymerase II largest subunit

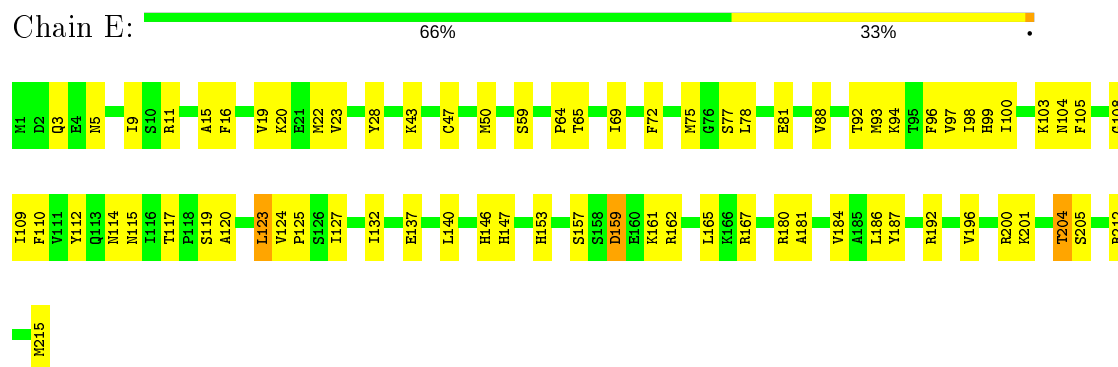




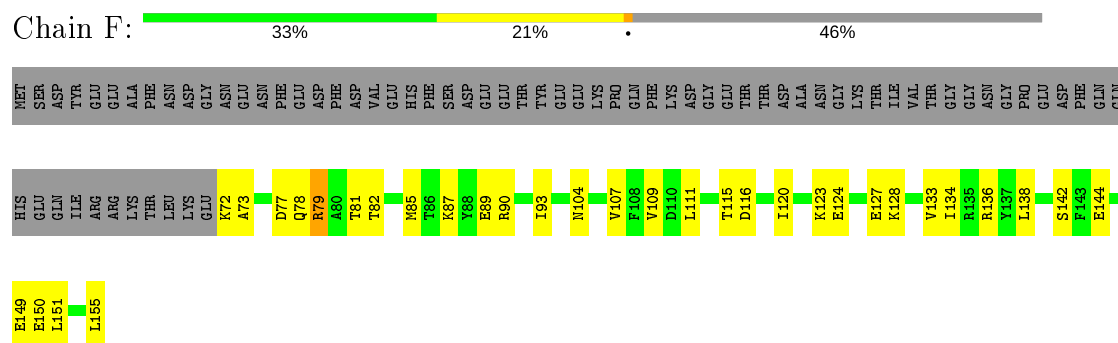
● Molecule 2: DNA-directed RNA polymerase II 140 kDa polypeptide

Chain B: 51% 36% 11%

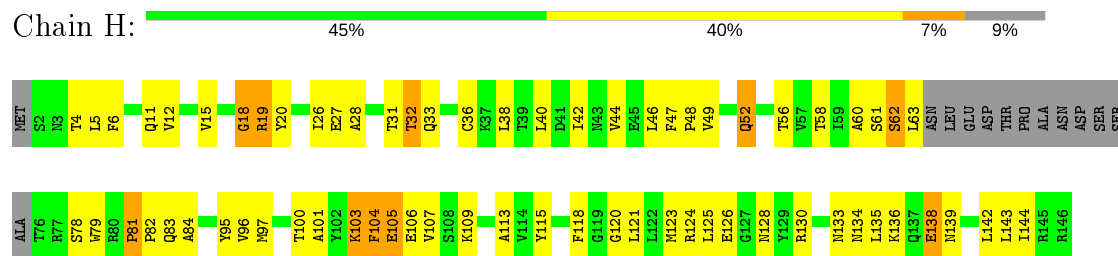




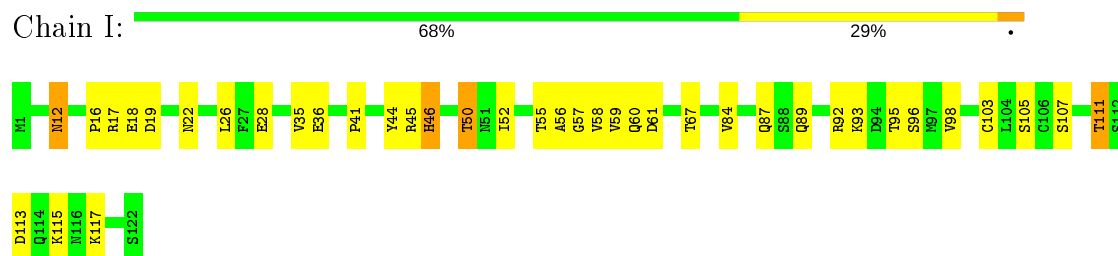
- Molecule 5: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide



- Molecule 6: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide

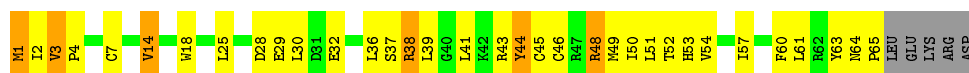


- Molecule 7: DNA-directed RNA polymerase II 14.2 kDa polypeptide



- Molecule 8: DNA-directed RNA polymerases I, II, and III 8.3 kDa polypeptide





- Molecule 9: DNA-directed RNA polymerase II 13.6 kDa polypeptide

Chain K: 54% 38% 5%



- Molecule 10: DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide

Chain L: 31% 26% 9% 34%



4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	123.00 Å 223.00 Å 374.00 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 2.30 39.69 – 2.19	Depositor EDS
% Data completeness (in resolution range)	(Not available) (40.00-2.30) 90.5 (39.69-2.19)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.95 (at 2.20 Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.247 , 0.294 0.292 , 0.290	Depositor DCC
R_{free} test set	5166 reflections (2.00%)	wwPDB-VP
Wilson B-factor (Å ²)	47.4	Xtriage
Anisotropy	0.331	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 46.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	28318	wwPDB-VP
Average B, all atoms (Å ²)	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.09% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, MN, UTP

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	0.43	1/11352 (0.0%)	0.66	2/15352 (0.0%)
2	B	0.41	0/8882	0.64	1/11976 (0.0%)
3	C	0.41	0/2133	0.60	0/2891
4	E	0.41	0/1796	0.63	1/2416 (0.0%)
5	F	0.44	0/691	0.63	0/933
6	H	0.30	0/1086	0.58	0/1470
7	I	0.40	0/1016	0.60	0/1365
8	J	0.41	0/541	0.65	0/727
9	K	0.39	0/937	0.56	0/1265
10	L	0.37	0/366	0.55	0/485
All	All	0.41	1/28800 (0.0%)	0.64	4/38880 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	764	CYS	CB-SG	-7.44	1.69	1.82

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	472	LEU	CA-CB-CG	-7.72	97.54	115.30
2	B	829	CYS	N-CA-C	-6.17	94.35	111.00
4	E	200	ARG	NE-CZ-NH2	-6.02	117.29	120.30
1	A	779	PHE	N-CA-C	-5.00	97.50	111.00

There are no chirality outliers.

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	11154	0	11224	555	0
2	B	8711	0	8738	427	0
3	C	2095	0	2051	139	0
4	E	1760	0	1788	53	0
5	F	679	0	701	34	0
6	H	1068	0	1040	57	0
7	I	997	0	953	40	0
8	J	532	0	542	53	0
9	K	919	0	929	62	0
10	L	364	0	389	33	0
11	A	2	0	0	0	0
11	B	1	0	0	0	0
11	C	1	0	0	0	0
11	I	2	0	0	0	0
11	J	1	0	0	0	0
11	L	1	0	0	0	0
12	A	2	0	0	0	0
13	B	29	0	11	1	0
All	All	28318	0	28366	1291	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 23.

The worst 5 of 1291 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:855:THR:HG21	1:A:857:ARG:HE	1.09	1.14
6:H:130:ARG:HA	6:H:133:ASN:HD22	1.11	1.11
1:A:1364:ASN:ND2	1:A:1366:ARG:HH11	1.53	1.05
1:A:351:THR:HG22	1:A:352:VAL:H	1.20	1.05
1:A:1364:ASN:HD21	1:A:1366:ARG:NH1	1.58	0.99

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	1411/1733 (81%)	1250 (89%)	123 (9%)	38 (3%)	5	3
2	B	1074/1224 (88%)	950 (88%)	110 (10%)	14 (1%)	12	12
3	C	264/318 (83%)	236 (89%)	24 (9%)	4 (2%)	10	10
4	E	213/215 (99%)	189 (89%)	22 (10%)	2 (1%)	17	20
5	F	82/155 (53%)	76 (93%)	5 (6%)	1 (1%)	13	14
6	H	129/146 (88%)	93 (72%)	21 (16%)	15 (12%)	0	0
7	I	120/122 (98%)	103 (86%)	17 (14%)	0	100	100
8	J	63/70 (90%)	59 (94%)	3 (5%)	1 (2%)	9	9
9	K	112/120 (93%)	106 (95%)	6 (5%)	0	100	100
10	L	44/70 (63%)	25 (57%)	16 (36%)	3 (7%)	1	0
All	All	3512/4173 (84%)	3087 (88%)	347 (10%)	78 (2%)	6	5

5 of 78 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	55	ASP
1	A	464	PRO
1	A	465	TYR
1	A	466	SER
1	A	567	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	1239/1520 (82%)	1180 (95%)	59 (5%)	25	36
2	B	950/1061 (90%)	903 (95%)	47 (5%)	25	35
3	C	234/274 (85%)	226 (97%)	8 (3%)	37	51
4	E	197/197 (100%)	191 (97%)	6 (3%)	41	57
5	F	74/137 (54%)	71 (96%)	3 (4%)	30	43
6	H	117/128 (91%)	115 (98%)	2 (2%)	60	76
7	I	116/116 (100%)	109 (94%)	7 (6%)	19	26
8	J	60/65 (92%)	52 (87%)	8 (13%)	4	4
9	K	99/102 (97%)	93 (94%)	6 (6%)	18	25
10	L	40/57 (70%)	35 (88%)	5 (12%)	4	5
All	All	3126/3657 (86%)	2975 (95%)	151 (5%)	25	36

5 of 151 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	313	MET
2	B	644	GLU
8	J	48	ARG
2	B	357	GLN
2	B	496	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 84 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	255	GLN
2	B	587	HIS
7	I	116	ASN
2	B	363	HIS
2	B	484	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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 [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 11 ligands modelled in this entry, 10 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
13	UTP	B	3571	12	26,30,30	2.20	5 (19%)	34,47,47	1.30	4 (11%)

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
13	UTP	B	3571	12	-	5/22/38/38	0/2/2/2

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	B	3571	UTP	C6-C5	-5.93	1.36	1.52
13	B	3571	UTP	C2-N1	5.92	1.44	1.35
13	B	3571	UTP	C6-N1	-4.12	1.39	1.47
13	B	3571	UTP	C1'-N1	3.94	1.54	1.46
13	B	3571	UTP	PB-O1B	-2.09	1.45	1.55

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
13	B	3571	UTP	C5-C6-N1	3.53	123.25	111.61
13	B	3571	UTP	O4'-C1'-C2'	-2.72	100.70	106.64
13	B	3571	UTP	O4-C4-C5	2.09	126.63	122.17
13	B	3571	UTP	O4'-C1'-N1	2.02	112.05	109.30

There are no chirality outliers.

All (5) torsion outliers are listed below:

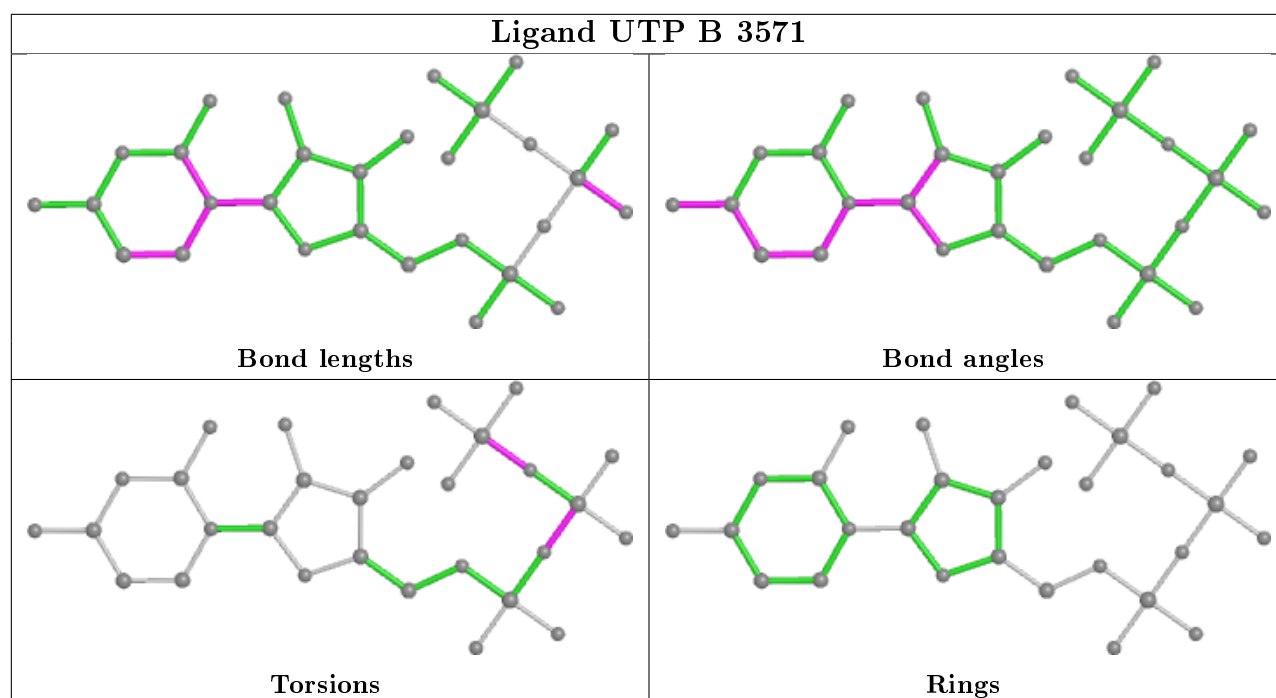
Mol	Chain	Res	Type	Atoms
13	B	3571	UTP	PA-O3A-PB-O2B
13	B	3571	UTP	PB-O3B-PG-O2G
13	B	3571	UTP	PA-O3A-PB-O1B
13	B	3571	UTP	PB-O3B-PG-O1G
13	B	3571	UTP	PB-O3B-PG-O3G

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
13	B	3571	UTP	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 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.



5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

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.

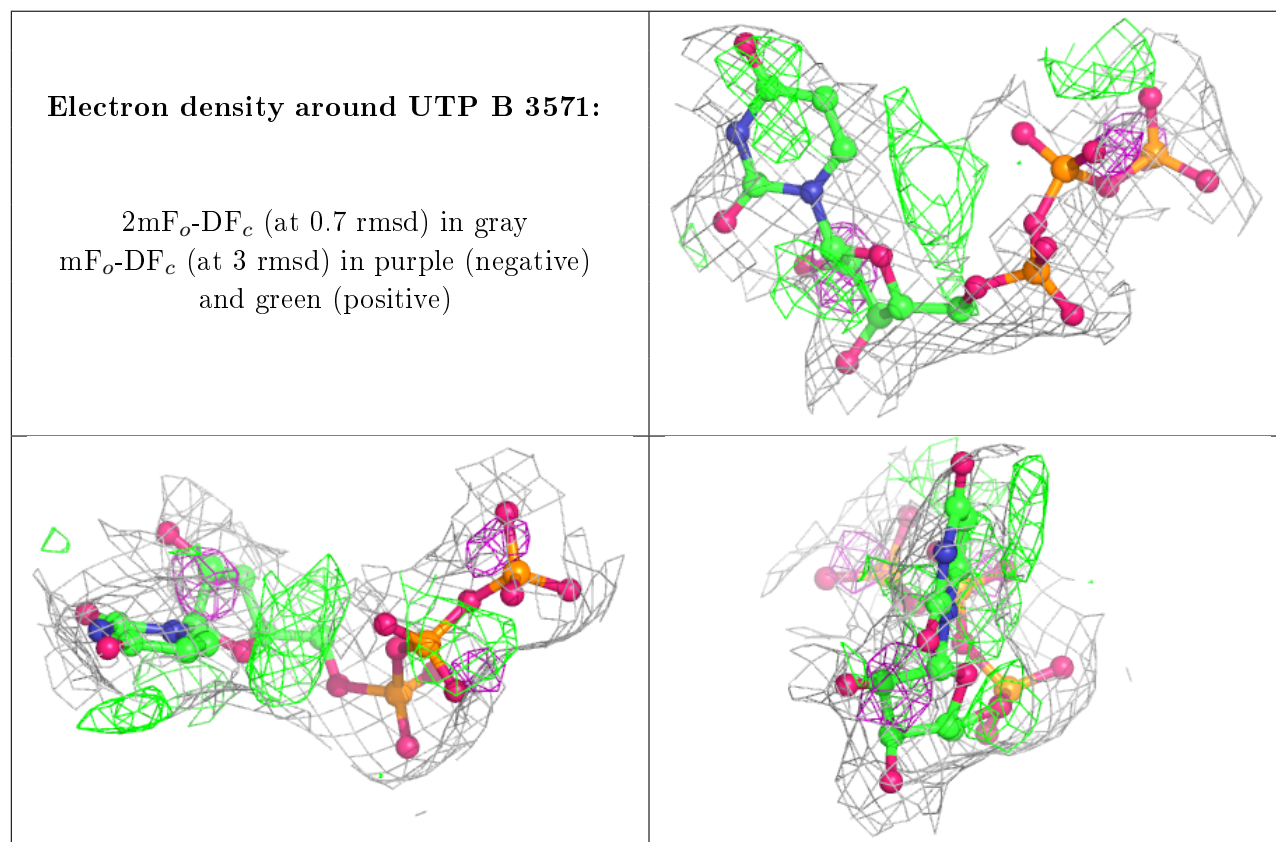
6.3 Carbohydrates ⓘ

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.



6.5 Other polymers ⓘ

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