



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 12, 2026 – 06:00 PM UTC

PDB ID : 1TWF / pdb\_00001twf  
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](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 : 4-5-2 with Phenix2.0  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtrriage (Phenix) : 2.0  
EDS : 3.0  
Buster-report : wwPDB partial adaption of 1.1.7 (2018)  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

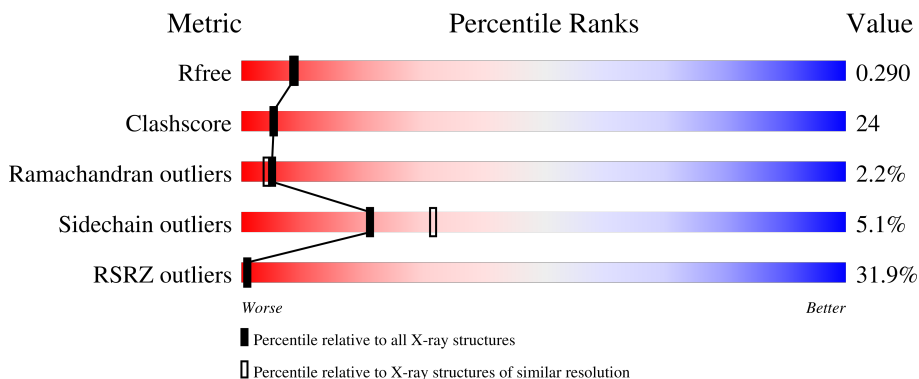
# 1 Overall quality at a glance

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))
$R_{free}$	180053	6319 (2.30-2.30)
Clashscore	190562	6919 (2.30-2.30)
Ramachandran outliers	187476	6854 (2.30-2.30)
Sidechain outliers	187428	6854 (2.30-2.30)
RSRZ outliers	180081	6325 (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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$ . 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	1733	
2	B	1224	
3	C	318	
4	E	215	
5	F	155	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
6	H	146	<p>70% 42% 41% 8% 9%</p>
7	I	122	<p>32% 67% 30%</p>
8	J	70	<p>17% 44% 40% 9% 7%</p>
9	K	120	<p>23% 52% 39% 5%</p>
10	L	70	<p>54% 31% 26% 9% 34%</p>

## 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
			Total	C	N	O	S			
1	A	1419	11154	7023	1952	2118	61	0	0	0

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

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

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

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

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

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

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

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

- 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
			Total	C	N	O	S			
6	H	133	1068	673	180	211	4	0	0	0

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

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

- 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
			Total	C	N	O	S			
8	J	65	532	339	93	94	6	0	0	0

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

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

- 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
			Total	C	N	O	S			
10	L	46	364	224	72	64	4	0	0	0

- Molecule 11 is ZINC ION (CCD ID: ZN) (formula: Zn).

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

*Continued on next page...*

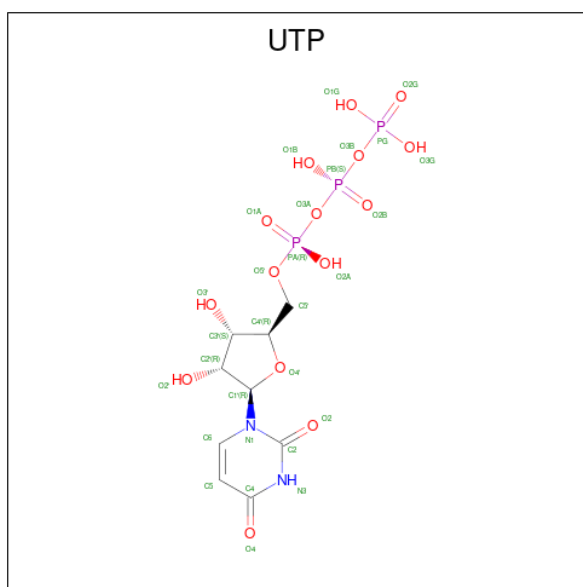
Continued from previous page...

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

- Molecule 12 is MANGANESE (II) ION (CCD ID: MN) (formula: Mn).

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

- Molecule 13 is URIDINE 5'-TRIPHOSPHATE (CCD ID: UTP) (formula: C<sub>9</sub>H<sub>15</sub>N<sub>2</sub>O<sub>15</sub>P<sub>3</sub>).

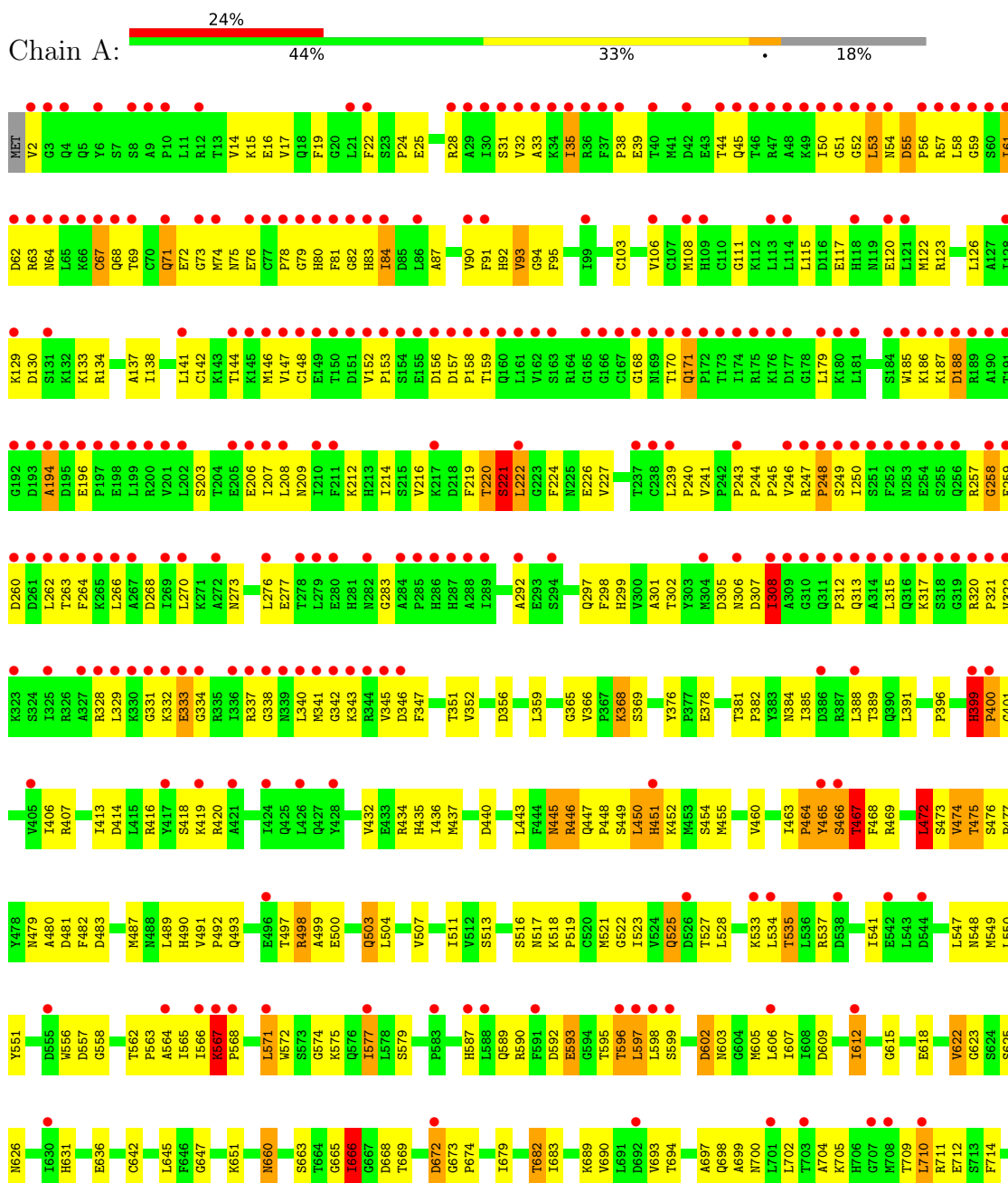


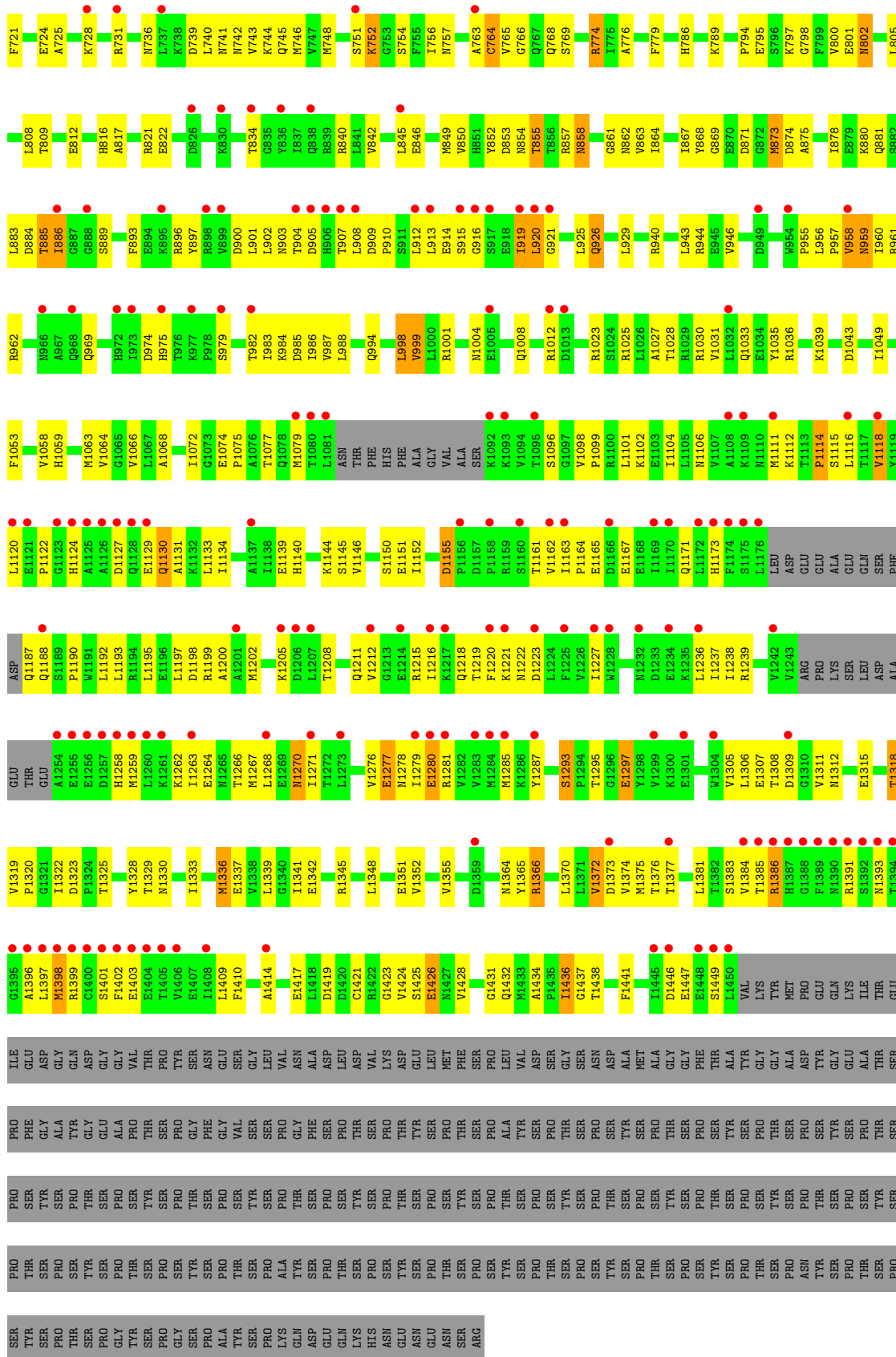
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
13	B	1	29	9	2	15	3	0	0

### 3 Residue-property plots i

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: DNA-directed RNA polymerase II largest subunit

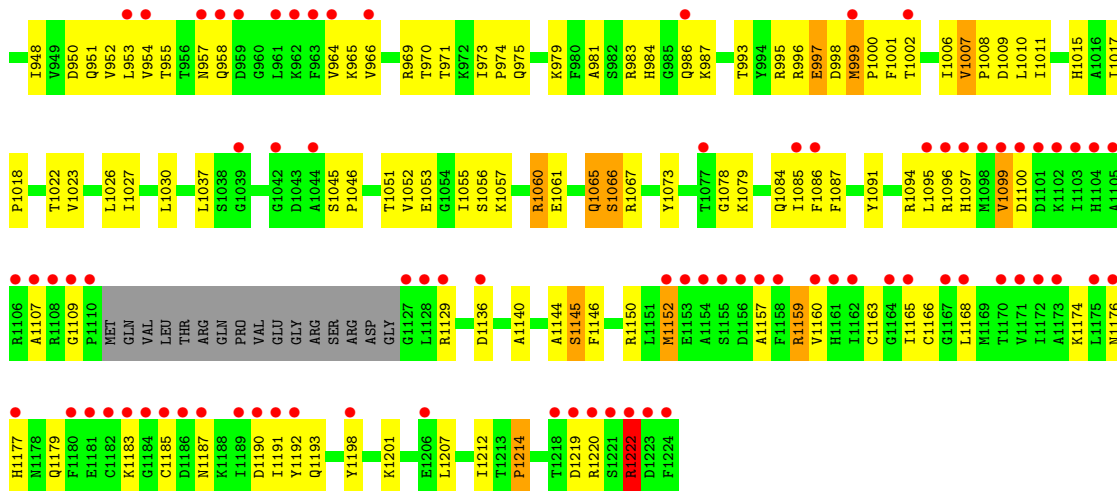




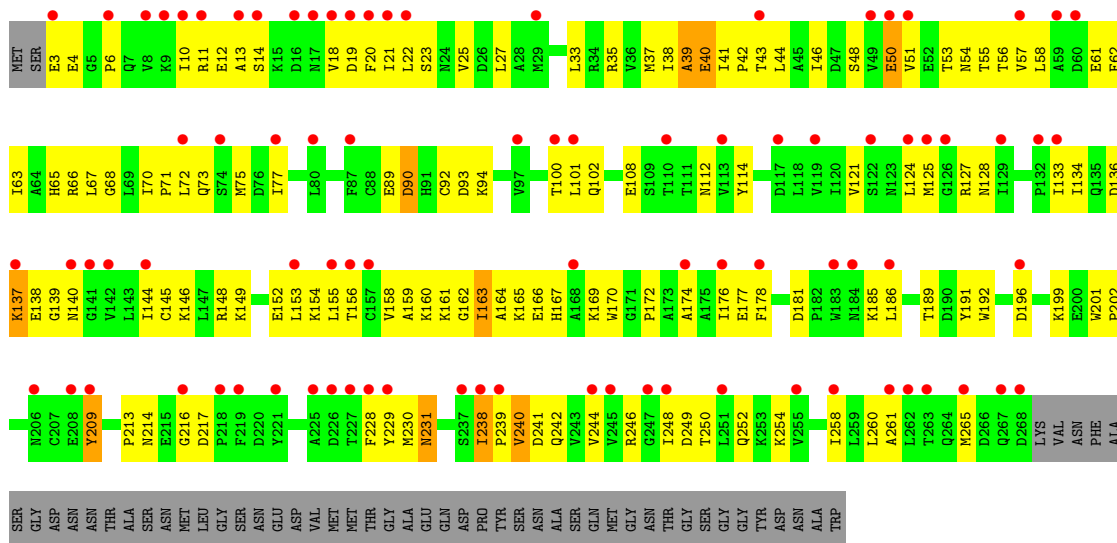
● Molecule 2: DNA-directed RNA polymerase II 140 kDa polypeptide



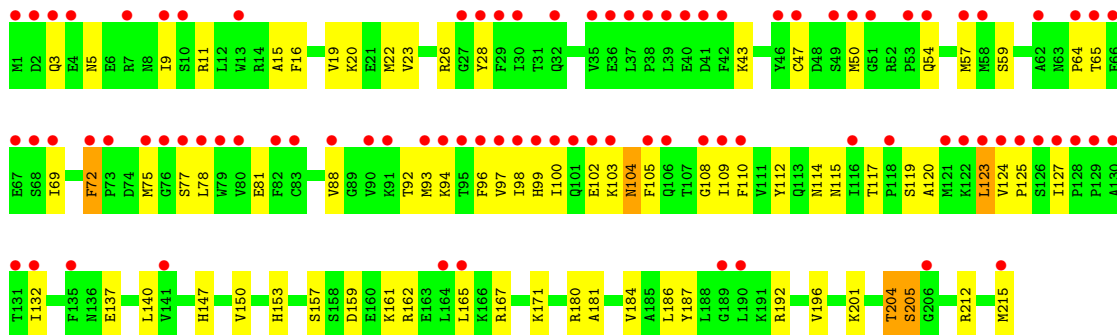
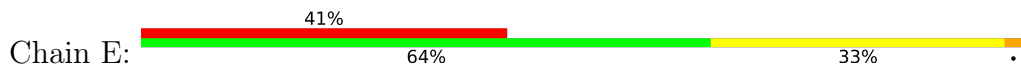
THR	THR	ALA	F203	L269	E346	R434	M499	S732	E810	N881
GLU	GLU	ILE	I204	K270	K347	T436	I502	H733	Y811	T882
ASP	ASP	ASP	N206	P274	R348	V436	GLY	H734	Y812	L883
LEU	LEU	VAL	G207	P274	E437	E437	ARG	A735	R813	R884
ALA	ALA	PRO	V211	I276	Y351	GLU	ASP	T736	F814	M885
ASN	ASN	GLY	L212	L280	T355	HIS	GLY	W586	P818	K886
SER	SER	ARG	L212	P281	K357	ASP	GLY	H587	L887	H887
GLU	GLU	LEU	Q215	L282	Q357	PHE	LYS	G588	N822	G888
LYS	LYS	LEU	E216	V283	K358	ASN	LEU	H589	E742	T889
TYR	TYR	TYR	R217	L284	E359	MET	ASP	I743	A823	D890
ASP	ASP	TYR	S218	L285	F360	GLU	GLU	H744	I824	I891
GLU	GLU	LEU	A219	F286	L361	L446	GLU	P745	V825	K892
PRO	PRO	ALA	I222	R287	L446	A447	Q513	M747	I827	L893
GLY	GLY	GLU	V225	A288	T365	I448	L514	I748	D894	D894
TYR	TYR	GLU	F226	L289	Q366	N449	H515	L749	A828	D895
Y96	Y96	SER	K227	G290	Q367	A450	T517	G750	C929	I899
Y97	Y97	GLU	K228	D294	L367	K451	H518	I751	Y830	A900
Y98	Y98	GLU	A229	L291	E368	K452	L521	V751	N834	P901
Y99	Y99	ASP	A230	L292	G369	T452	V522	S754	Q835	P901
P100	P100	SER	P231	L293	F370	T453	E526	I755	S838	R904
V102	V102	GLU	E231	D294	E371	L457	S527	I756	M839	V905
N103	N103	SER	L234	E296	S372	K458	P528	F758	I940	S906
E104	E104	GLY	L234	L297	K374	Y459	Q531	D760	M842	G907
S105	S105	K164	S235	L298	F377	A460	E531	K775	N842	E908
D106	D106	V165	H236	E299	L378	A461	M615	Q763	Q843	I911
G107	G107	F166	V237	H300	K381	L462	L616	S764	E896	I912
V108	V108	I167	A238	G305	N382	T463	R617	P765	E897	G913
T109	T109	C168	E239	V305	N383	G464	K637	R765	S700	K914
H110	H110	R169	L240	L311	R384	M465	M638	K775	S700	K914
A111	A111	L171	R241	E312	R385	W466	N639	Q776	I703	T915
L112	L112	I172	S242	M313	L386	G467	S540	A777	A704	T916
P114	P114	M173	A244	L314	L387	GLU	M541	M778	M705	P917
Q115	Q115	L174	L244	K315	C388	LYS	S543	Q706	Q706	R918
E116	E116	R175	E245	P316	A389	ALA	C544	F707	F707	S919
A117	A117	S176	K246	C317	L390	MET	E544	E708	E708	P920
R118	R118	K177	G247	S248	G467	SER	I545	D709	D709	ASP
L119	L119	M178	R249	D320	A460	ARG	S546	L710	L710	GLU
R120	R120	L181	F250	I324	A461	ARG	V547	E711	E711	LEU
N121	N121	S182	S251	L324	L405	A477	G548	F712	F712	GLY
L122	L122	E183	S252	R327	L416	G478	T549	ALA	ALA	GLN
T123	T123	A184	T253	E328	F417	V479	D550	GLU	GLU	ARG
Y124	Y124	T185	L254	T329	L420	L483	M552	ALA	ALA	THR
S125	S125	E186	Q255	A330	F421	N484	P553	ASN	ASN	THR
G126	G126	L186	V256	L331	K422	R485	F557	GLU	GLU	TYR
S127	S127	K191	K257	D332	K423	Y486	L558	ASN	ASN	HIS
L128	L128	L192	L258	F333	L424	T487	L558	ASP	ASP	HIS
F129	F129	L192	Y259	F333	L424	T487	L558	LEU	LEU	K934
V130	V130	K193	G260	R336	T425	K649	W561	D722	D722	R935
D131	D131	E194	R261	R337	K426	L491	G562	E650	E650	R935
V132	V132	C195	E262	G338	D427	L492	M563	L492	L492	D936
K133	K133	P196	G263	T339	L428	S493	E564	K652	K652	A937
L134	L134	F197	S264	A340	F429	H494	P565	P725	P725	S938
R135	R135	G200	S265	T343	K430	L495	L566	W653	W653	T939
T136	T136	G201	A266	K434	Y431	R496	E567	R654	R654	P940
Y137	Y137	E138	R267	K344	N432	R497	E567	H657	H657	L941
THR	THR	Y202	T268	K345	Q433	T498	P571	L658	L658	T944



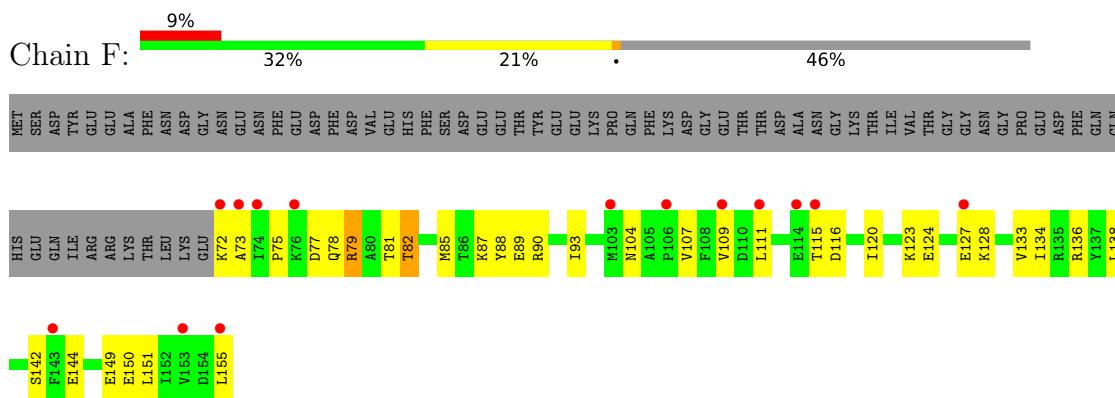
● Molecule 3: DNA-directed RNA polymerase II 45 kDa polypeptide



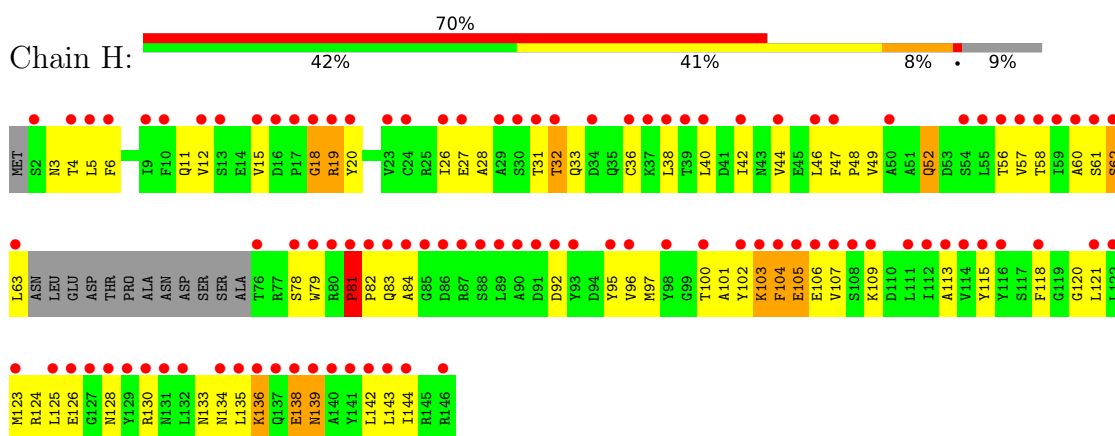
● Molecule 4: DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide



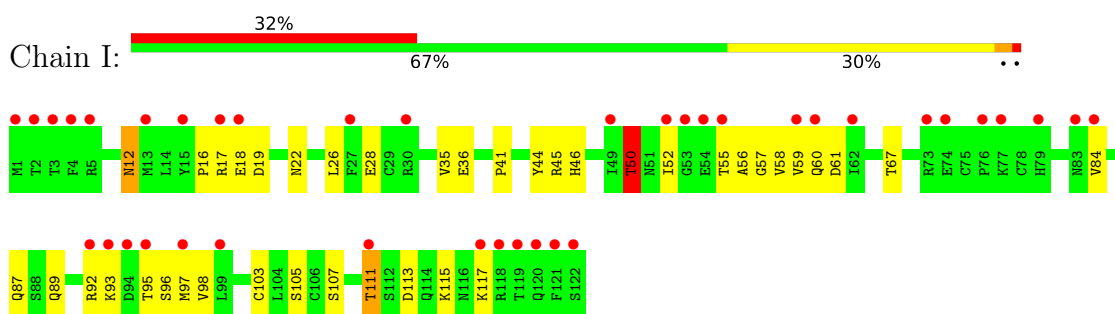
- Molecule 5: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide



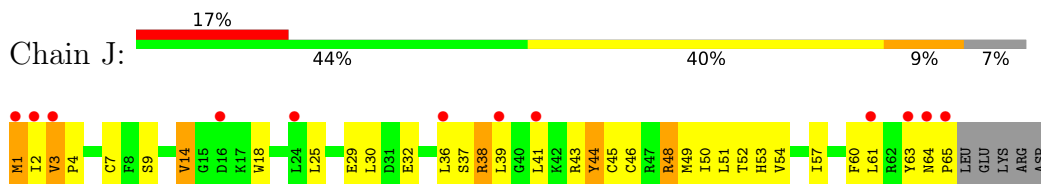
- Molecule 6: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide



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

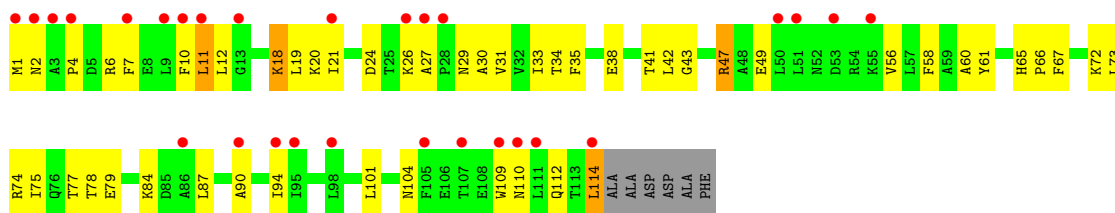


- Molecule 8: DNA-directed RNA polymerases I, II, and III 8.3 kDa polypeptide

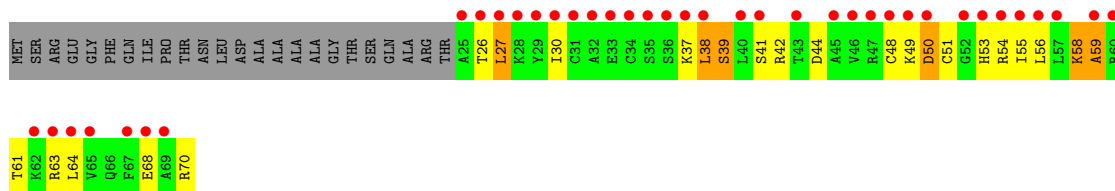
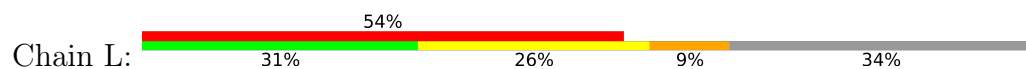


- Molecule 9: DNA-directed RNA polymerase II 13.6 kDa polypeptide





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



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	123.00Å 223.00Å 374.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 2.30 40.00 – 2.30	Depositor EDS
% Data completeness (in resolution range)	(Not available) (40.00-2.30) 92.0 (40.00-2.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.95 (at 2.20Å)	Xtrriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.247 , 0.294 0.293 , 0.290	Depositor DCC
$R_{free}$ test set	5166 reflections (2.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	47.4	Xtrriage
Anisotropy	0.331	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 43.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	28318	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	65.0	wwPDB-VP

Xtrriage'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.*

<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.

## 5 Model quality i

### 5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: UTP, MN, 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	0.55	2/11352 (0.0%)	0.99	42/15352 (0.3%)
2	B	0.52	0/8882	0.96	27/11976 (0.2%)
3	C	0.48	0/2133	0.94	5/2891 (0.2%)
4	E	0.51	1/1796 (0.1%)	0.93	4/2416 (0.2%)
5	F	0.53	0/691	0.96	2/933 (0.2%)
6	H	0.35	0/1086	0.90	6/1470 (0.4%)
7	I	0.43	0/1016	0.90	1/1365 (0.1%)
8	J	0.53	0/541	0.88	1/727 (0.1%)
9	K	0.44	0/937	0.82	1/1265 (0.1%)
10	L	0.38	0/366	0.76	1/485 (0.2%)
All	All	0.52	3/28800 (0.0%)	0.96	90/38880 (0.2%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	746	MET	SD-CE	-5.20	1.66	1.79
4	E	150	VAL	CA-CB	5.09	1.58	1.53
1	A	507	VAL	CA-CB	5.05	1.60	1.54

The worst 5 of 90 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	829	CYS	N-CA-C	-8.82	94.35	108.73
1	A	886	ILE	N-CA-C	8.40	119.21	110.72
1	A	779	PHE	N-CA-C	-8.39	97.50	110.17
1	A	999	VAL	N-CA-C	-8.35	105.10	112.12
1	A	333	GLU	N-CA-C	-8.09	103.13	113.16

There are no chirality outliers.

There are no planarity outliers.

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

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	602	0
2	B	8711	0	8738	449	0
3	C	2095	0	2051	142	0
4	E	1760	0	1788	58	0
5	F	679	0	701	36	0
6	H	1068	0	1040	63	0
7	I	997	0	953	43	0
8	J	532	0	542	55	0
9	K	919	0	929	63	0
10	L	364	0	389	32	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	1376	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 24.

The worst 5 of 1376 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 [i](#)

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

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%)	4	3
2	B	1074/1224 (88%)	950 (88%)	110 (10%)	14 (1%)	9	10
3	C	264/318 (83%)	236 (89%)	24 (9%)	4 (2%)	8	8
4	E	213/215 (99%)	189 (89%)	22 (10%)	2 (1%)	14	17
5	F	82/155 (53%)	76 (93%)	5 (6%)	1 (1%)	10	12
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%)	7	7
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%)	5	4

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 [i](#)

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%)	1175 (95%)	64 (5%)	21	31
2	B	950/1061 (90%)	895 (94%)	55 (6%)	18	26
3	C	234/274 (85%)	226 (97%)	8 (3%)	32	49
4	E	197/197 (100%)	193 (98%)	4 (2%)	48	67
5	F	74/137 (54%)	71 (96%)	3 (4%)	27	41
6	H	117/128 (91%)	115 (98%)	2 (2%)	53	72
7	I	116/116 (100%)	110 (95%)	6 (5%)	21	31
8	J	60/65 (92%)	53 (88%)	7 (12%)	5	6
9	K	99/102 (97%)	94 (95%)	5 (5%)	21	32
10	L	40/57 (70%)	36 (90%)	4 (10%)	7	9
All	All	3126/3657 (86%)	2968 (95%)	158 (5%)	21	32

5 of 158 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	1159	ARG
8	J	3	VAL
3	C	23	SER
4	E	204	THR
9	K	18	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 98 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	770	GLN
3	C	65	HIS
2	B	957	ASN
2	B	1161	HIS
3	C	151	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides 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	29,30,30	1.38	4 (13%)	43,47,47	1.10	3 (6%)

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 (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	B	3571	UTP	PA-O3A	4.21	1.64	1.59
13	B	3571	UTP	C2-N1	3.59	1.44	1.38
13	B	3571	UTP	C1'-N1	2.29	1.54	1.47
13	B	3571	UTP	PB-O1B	-2.12	1.45	1.55

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
13	B	3571	UTP	O1B-PB-O3A	3.20	115.93	107.27
13	B	3571	UTP	O4'-C1'-C2'	-2.76	100.70	106.62
13	B	3571	UTP	O3A-PB-O2B	-2.49	103.21	110.70

There are no chirality outliers.

All (5) torsion outliers are listed below:

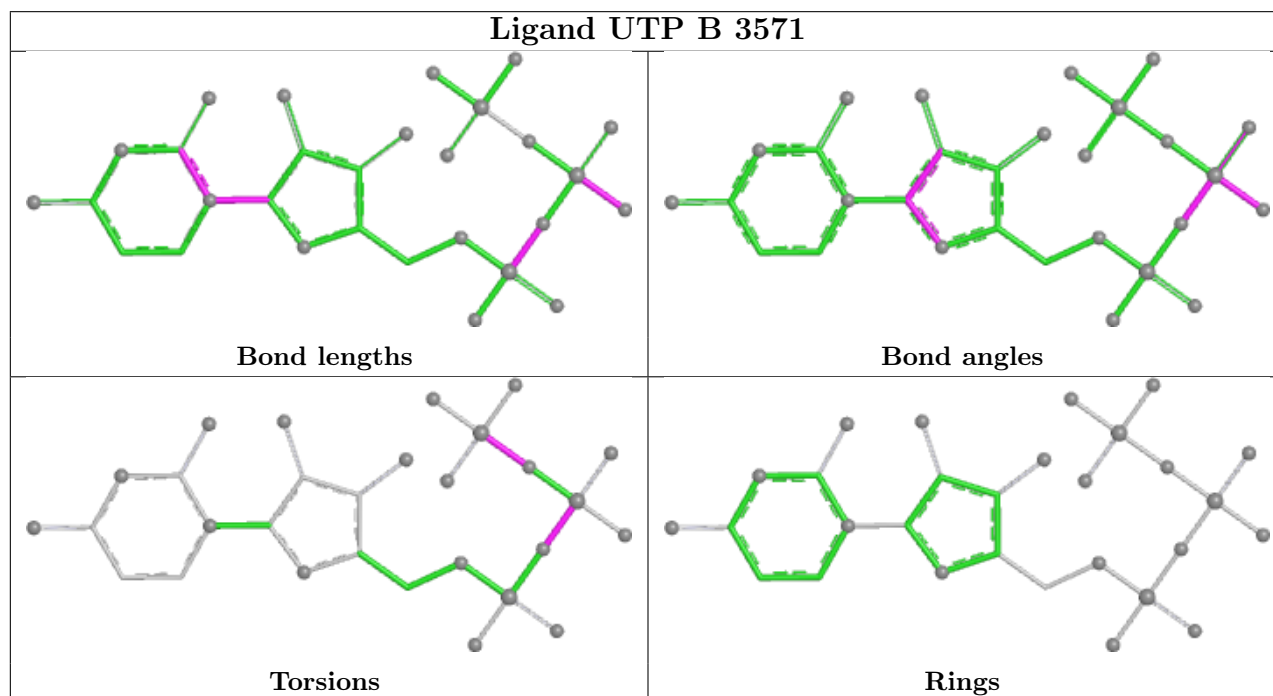
Mol	Chain	Res	Type	Atoms
13	B	3571	UTP	PA-O3A-PB-O2B
13	B	3571	UTP	PB-O3B-PG-O2G
13	B	3571	UTP	PA-O3A-PB-O1B
13	B	3571	UTP	PB-O3B-PG-O1G
13	B	3571	UTP	PB-O3B-PG-O3G

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 [i](#)

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

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	1419/1733 (81%)	1.54	417 (29%) 1 1	15, 54, 147, 169	0
2	B	1094/1224 (89%)	1.49	311 (28%) 1 1	16, 50, 129, 158	0
3	C	266/318 (83%)	1.70	87 (32%) 1 1	27, 58, 95, 144	0
4	E	215/215 (100%)	1.79	88 (40%) 0 0	23, 67, 113, 149	0
5	F	84/155 (54%)	1.25	14 (16%) 4 5	20, 49, 77, 98	0
6	H	133/146 (91%)	3.08	102 (76%) 0 0	63, 99, 141, 150	0
7	I	122/122 (100%)	1.72	39 (31%) 1 1	41, 64, 108, 130	0
8	J	65/70 (92%)	1.33	12 (18%) 3 4	29, 49, 80, 94	0
9	K	114/120 (95%)	1.61	28 (24%) 2 2	31, 67, 86, 97	0
10	L	46/70 (65%)	3.07	38 (82%) 0 0	52, 105, 132, 137	0
All	All	3558/4173 (85%)	1.63	1136 (31%) 1 1	15, 56, 135, 169	0

The worst 5 of 1136 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	340	LEU	10.9
1	A	1176	LEU	10.5
6	H	104	PHE	10.2
1	A	1175	SER	9.9
10	L	27	LEU	9.0

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

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

### 6.3 Carbohydrates [i](#)

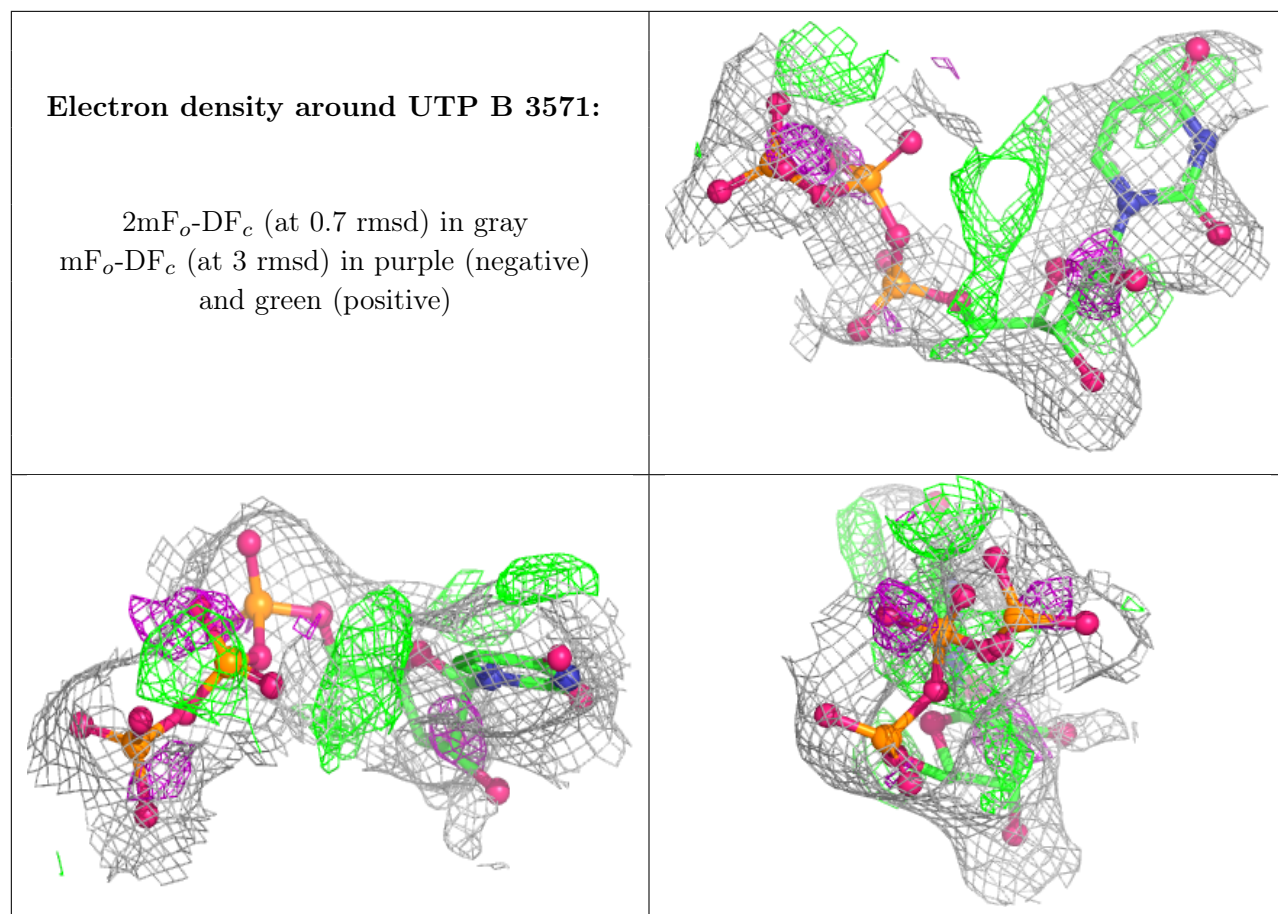
There are no oligosaccharides in this entry.

### 6.4 Ligands [i](#)

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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
13	UTP	B	3571	29/29	0.79	0.14	53,61,64,65	0
11	ZN	A	3006	1/1	0.90	0.08	53,53,53,53	0
11	ZN	I	3003	1/1	0.93	0.05	58,58,58,58	0
11	ZN	A	3008	1/1	0.94	0.08	97,97,97,97	0
11	ZN	B	3007	1/1	0.94	0.06	55,55,55,55	0
11	ZN	J	3001	1/1	0.95	0.05	45,45,45,45	0
11	ZN	L	3005	1/1	0.95	0.06	89,89,89,89	0
11	ZN	C	3002	1/1	0.95	0.05	48,48,48,48	0
12	MN	A	3010	1/1	0.96	0.04	40,40,40,40	0
11	ZN	I	3004	1/1	0.96	0.06	74,74,74,74	0
12	MN	A	3009	1/1	0.97	0.12	44,44,44,44	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.



## 6.5 Other polymers [i](#)

There are no such residues in this entry.