



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 19, 2023 – 01:18 AM EDT

PDB ID : 2NVZ
Title : RNA Polymerase II elongation complex with UTP, updated 11/2006
Authors : Wang, D.; Bushnell, D.A.; Westover, K.D.; Kaplan, C.D.; Kornberg, R.D.
Deposited on : 2006-11-14
Resolution : 4.30 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 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 : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36
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) : 2.36

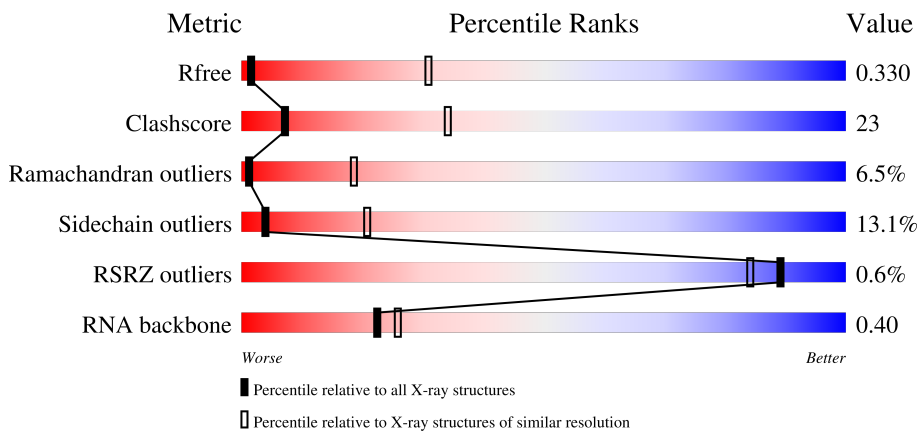
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.30 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.









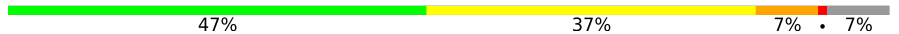

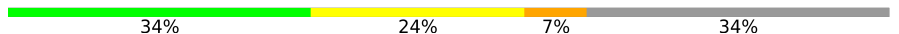
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1014 (4.80-3.80)
Clashscore	141614	1077 (4.80-3.80)
Ramachandran outliers	138981	1029 (4.80-3.80)
Sidechain outliers	138945	1012 (4.80-3.80)
RSRZ outliers	127900	1075 (4.90-3.70)
RNA backbone	3102	1058 (5.60-3.00)

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 ≥ 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	R	10	
2	T	28	
3	N	14	
4	A	1733	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	B	1224	
6	C	318	
7	E	215	
8	F	155	
9	H	146	
10	I	122	
11	J	70	
12	K	120	
13	L	70	

2 Entry composition [i](#)

There are 16 unique types of molecules in this entry. The entry contains 29002 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 RNA chain called 5'-R(*AP*UP*CP*GP*AP*GP*AP*GP*GP*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	R	10	216	98	45	64	9	0	0	0

- Molecule 2 is a DNA chain called 28-MER DNA template strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	T	28	566	271	104	164	27	0	0	0

- Molecule 3 is a DNA chain called 5'-D(*CP*TP*GP*CP*TP*TP*AP*TP*CP*GP*GP*TP*AP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	N	14	284	137	49	85	13	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	A	1398	10984	6930	1924	2069	61	0	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerase II 140 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	B	1096	8701	5508	1518	1620	55	0	0	0

- Molecule 6 is a protein called DNA-directed RNA polymerase II 45 kDa polypeptide.

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

- Molecule 7 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			
7	E	193	1594	1016	283	287	8	0	0	0

- Molecule 8 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			
8	F	83	670	428	114	125	3	0	0	0

- Molecule 9 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			
9	H	133	1068	673	180	211	4	0	0	0

- Molecule 10 is a protein called DNA-directed RNA polymerase II subunit 9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
10	I	119	971	596	179	186	10	0	0	0

- Molecule 11 is a protein called DNA-directed RNA polymerases I/II/III subunit 10.

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

- Molecule 12 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			
12	K	114	919	590	156	171	2	0	0	0

- Molecule 13 is a protein called DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide.

tide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	L	46	Total	C	N	O	S	0	0	0
			363	224	72	63	4			

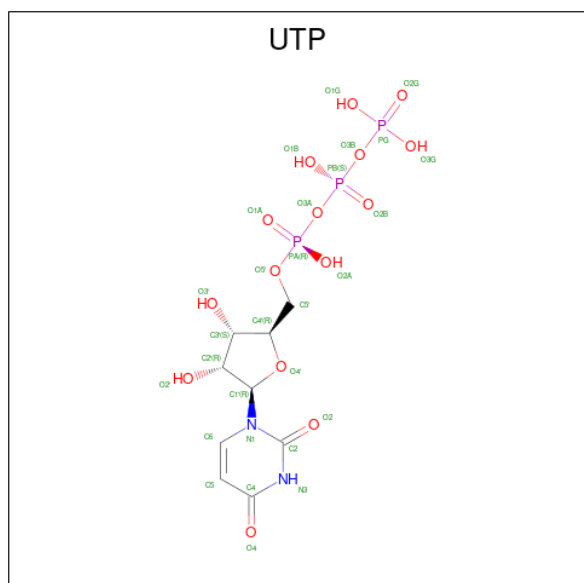
- Molecule 14 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 15 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	A	2	Total	Mg	0	0
			2	2		

- Molecule 16 is URIDINE 5'-TRIPHOSPHATE (three-letter code: UTP) (formula: C₉H₁₅N₂O₁₅P₃).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
16	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.

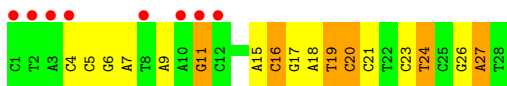
- Molecule 1: 5'-R(*AP*UP*CP*GP*AP*GP*AP*GP*GP*A)-3'

Chain R: 




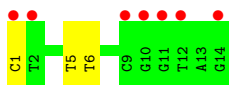
- Molecule 2: 28-MER DNA template strand

Chain T: 



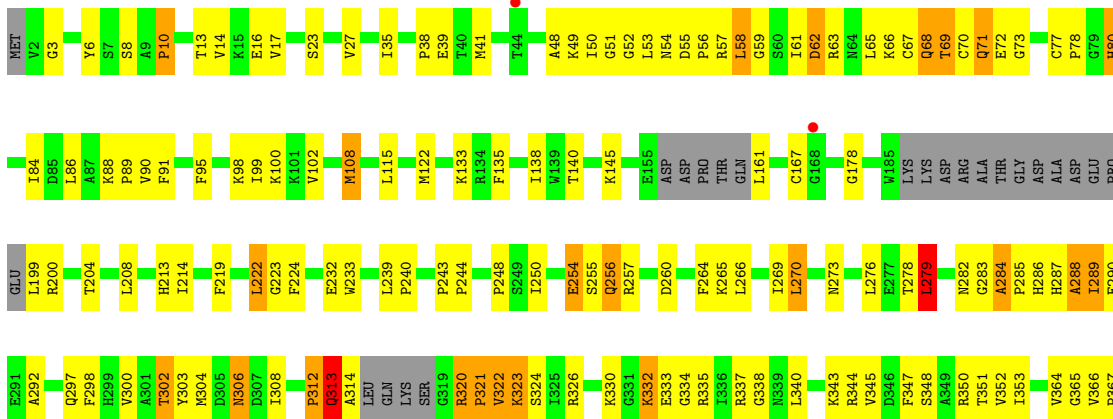
- Molecule 3: 5'-D(*CP*TP*GP*CP*TP*TP*AP*TP*CP*GP*GP*TP*AP*G)-3'

Chain N: 

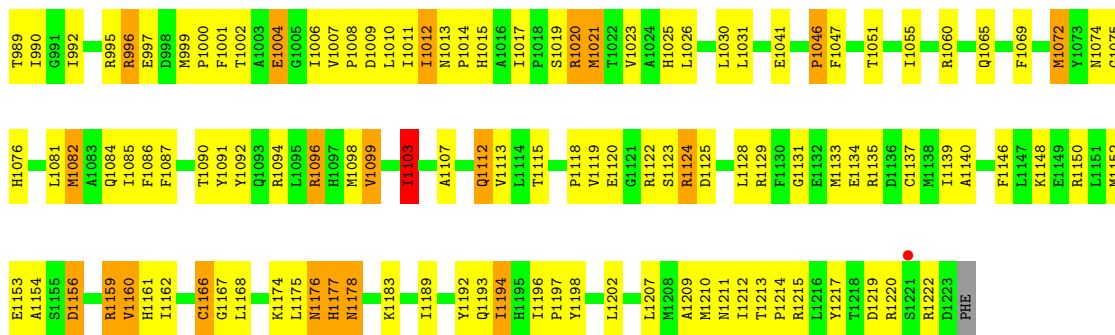


- Molecule 4: DNA-directed RNA polymerase II largest subunit

Chain A: 

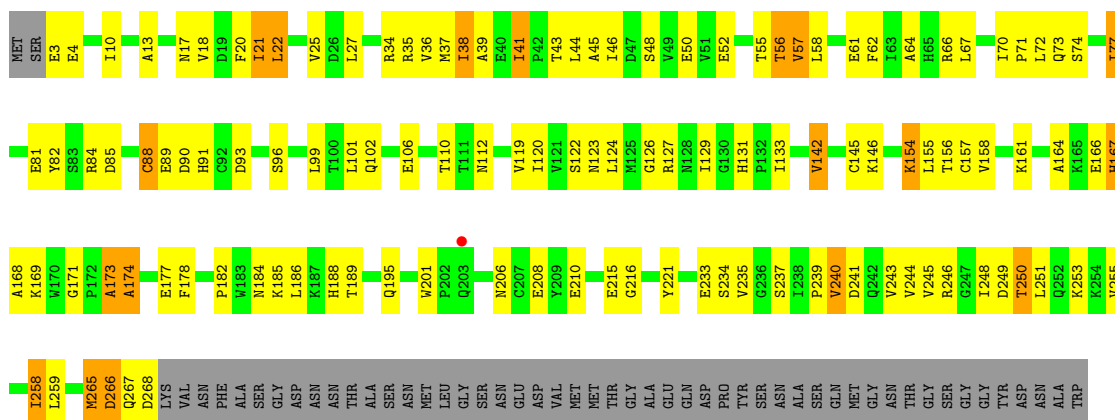


SER	ALA	MET	V1374	M1267	ASP	L1105	C1019	L928	L848	G786	R674	Q589	Q510	L443	I370
PRO	ASP	PRO	M1375	L1268	Q1187	M1106	Q1187	L929	M849	Q787	T675	R890	Q510	F444	I370
TYR	TYR	GLU	T1376	E1269	Q1188	V1107	R1025	D930	D853	Q788	T680	F591	S513	N445	T373
SER	GLY	GLN	M1270	M1270	S1189	T1113	T1028	E931	T855	R774	E681	T896	Q515	R446	L374
PRO	ALA	LYS	V1384	I1271	P1190	T1114	R1029	E932	T856	I775	T682	L597	S516	Q448	L375
THR	THR	THR	T1365	V1276	W1191	S1115	R1030	L936	R857	A776	I683	L598	M517	S449	E378
SER	SER	GLU	G1388	E1277	L1193	L1116	E1034	R857	M858	F777	A694	L606	K516	L450	V379
PRO	PRO	ILE	F1389	R1194	R1194	T1117	E1034	K941	R859	R782	K687	G610	M521	K452	V380
SER	PHE	GLU	M1390	L1281	W1035	V1118	F942	R942	S859	R783	G522	G610	G522	M453	T381
TYR	TYR	ASP	L1391	V1282	E1196	V1119	R1036	L860	L860	W783	D692	Q611	I524	S484	P382
SER	ALA	GLY	S1392	V1283	L1197	V1119	R1036	L860	L860	W783	D692	Q611	I524	S484	P382
PRO	TYR	GLN	M1393	M1284	D1198	L1120	R944	L860	L860	W783	D692	Q611	I524	S484	Y383
THR	GLY	ASP	T1394	W1284	D1198	L1120	R944	L860	L860	W783	D692	Q611	I524	S484	Y383
SER	GLU	GLY	L1199	Y1287	K1199	D1127	E1050	E945	L864	S788	K695	I613	Q525	M456	Y383
PRO	GLU	GLY	A1200	Y1287	K1199	D1127	E1050	E945	L864	S788	K695	I613	Q525	M456	Y383
PRO	ALA	GLY	A1201	Y1287	K1199	D1127	E1050	E945	L864	S788	K695	I613	Q525	M456	Y383
SER	PRO	VAL	M1202	T1295	M1202	A1131	V1057	D949	Y868	Y792	Q698	V617	D526	A457	R387
TYR	THR	THR	M1203	T1295	M1202	A1131	V1057	D949	Y868	Y792	Q698	V617	D526	A457	R387
SER	SER	PRO	W1304	W1304	D1206	I1134	V1058	L956	E899	S793	A699	E618	L528	R459	L388
PRO	PRO	TYR	V1406	W1305	D1206	I1134	V1058	L956	E899	S793	A699	E618	L528	R459	L388
THR	GLY	SER	L1409	L1306	G1213	T1141	V1064	P958	D871	E795	N700	K619	L534	V460	T389
PRO	GLY	ASN	F1410	T1308	G1213	T1142	V1064	P958	D871	E795	N700	K619	L534	V460	T389
SER	VAL	SER	E1411	T1308	G1213	T1143	V1066	R961	G872	S796	L702	S625	L536	Y465	Q390
TYR	SER	GLY	A1414	V1311	R1215	L1143	V1066	R961	G872	S796	L702	S625	L536	Y465	Q390
SER	SER	LEU	S1415	M1312	I1216	S1150	L1067	R961	G872	S796	L702	S625	L536	Y465	Q390
PRO	PRO	VAL	A1416	M1312	I1216	S1150	L1067	R961	G872	S796	L702	S625	L536	Y465	Q390
THR	GLY	ASN	D1223	V1319	D1223	I1152	A1069	Q965	L883	M802	T709	I630	E542	F468	Y404
SER	PHE	ALA	I1224	V1319	D1223	I1152	A1069	Q965	L883	M802	T709	I630	E542	F468	Y404
PRO	SER	ASP	I1227	I1322	I1227	Y1153	Q1073	Q969	T886	F814	R711	V633	D544	L470	Y404
SER	PRO	ASP	M1228	I1323	I1228	Y1154	G1073	Q969	T886	F814	R711	V633	D544	L470	Y404
TYR	PRO	LEU	D1420	P1324	M1228	D1155	E1074	F971	A892	H816	S713	R635	Q545	N472	I406
SER	THR	ASP	C1421	P1324	M1228	D1155	E1074	F971	A892	H816	S713	R635	Q545	N472	I406
SER	SER	VAL	V1424	T1325	D1231	P1158	P1075	H972	E894	A817	F714	F636	V546	L471	R407
PRO	PRO	LYS	A1425	T1326	D1231	P1158	P1075	H972	E894	A817	F714	F636	V546	L471	R407
THR	THR	ASP	E1426	I1327	ASN	R1159	Q1078	D974	R896	G819	L722	Q640	W552	N479	Y478
SER	PRO	GLU	M1427	I1327	ASN	R1159	Q1078	D974	R896	G819	L722	Q640	W552	N479	Y478
PRO	SER	LEU	V1428	T1329	LYS	V1162	M1079	T976	V899	E822	K728	V641	W556	A480	Y417
SER	PRO	MET	I1429	T1329	LYS	V1162	M1079	T976	V899	E822	K728	V641	W556	A480	Y417
THR	THR	PHE	L1430	I1333	L1236	I1163	L1081	P978	D900	G893	R731	C642	D481	D481	Y417
SER	SER	SER	L1431	D1334	I1237	P1164	M1082	S979	L901	L824	L645	L645	I560	F462	R420
PRO	PRO	PRO	A1434	D1334	I1237	P1164	M1082	S979	L901	L824	L645	L645	I560	F462	R420
TYR	ALA	LEU	P1435	D1335	I1237	P1164	M1082	S979	L901	L824	L645	L645	I560	F462	R420
SER	VAL	VAL	G1437	M1336	V1242	F1084	F1084	N903	N903	I826	N736	D826	P561	D423	D423
PRO	PRO	ASP	T1438	M1336	V1242	F1084	F1084	N903	N903	I826	N736	D826	P561	D423	D423
SER	PRO	ASP	F1441	G1340	ARG	I1169	F1086	T907	L908	A828	D739	I649	I565	M487	Q425
THR	THR	GLY	D1442	I1341	ARG	I1169	F1086	T907	L908	A828	D739	I649	I565	M487	Q425
SER	THR	THR	V1443	I1341	ARG	I1169	F1086	T907	L908	A828	D739	I649	I565	M487	Q425
PRO	SER	ALA	M1444	I1342	THR	L1170	A1087	D909	D909	K830	N741	K651	I566	M488	Q427
SER	PRO	ALA	I1445	I1342	THR	L1170	A1087	D909	D909	K830	N741	K651	I566	M488	Q427
TYR	THR	GLY	ASP	M1364	GLU	Q1171	V1089	L993	P910	T831	N742	V652	K567	M488	Q427
PRO	THR	GLY	GLU	Y1365	GLU	L1172	A1090	L994	P910	T831	N742	V652	K567	M488	Q427
PRO	PRO	PHE	GLU	R1366	GLU	L1173	S1091	E996	E996	E833	V743	P568	K569	H490	G429
THR	SER	THR	ALA	K1261	ALA	F1174	K1092	N996	E996	E833	V743	P568	K569	H490	G429
SER	TYR	ALA	L1370	K1262	GLU	S1175	K1093	L997	E914	G835	Q745	L657	K569	H490	G429
PRO	PRO	VAL	L1371	K1262	GLU	S1175	K1093	L997	E914	G835	Q745	L657	K569	H490	G429
TYR	PRO	VAL	L1372	K1262	GLU	S1175	K1093	L997	E914	G835	Q745	L657	K569	H490	G429
TYR	THR	TYR	D1373	K1262	GLU	S1175	K1093	L997	E914	G835	Q745	L657	K569	H490	G429



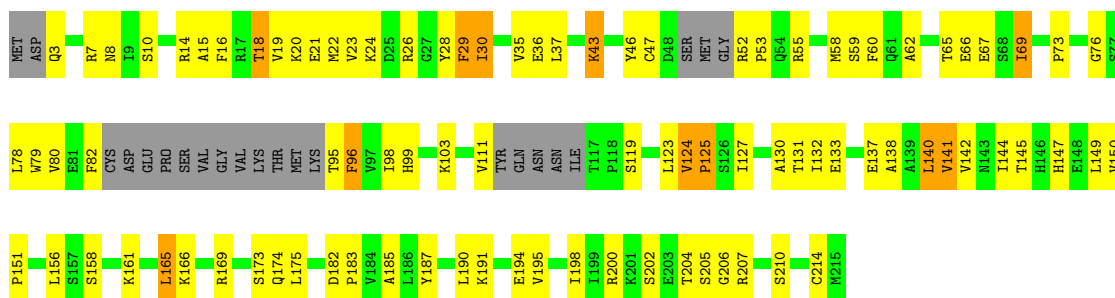
- Molecule 6: DNA-directed RNA polymerase II 45 kDa polypeptide

Chain C: 45% 33% 6% 16%



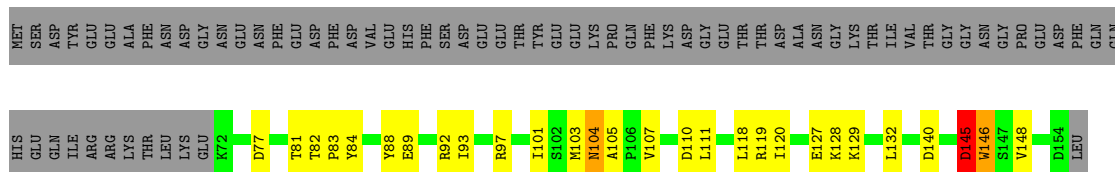
- Molecule 7: DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide

Chain E: 47% 38% 5% 10%

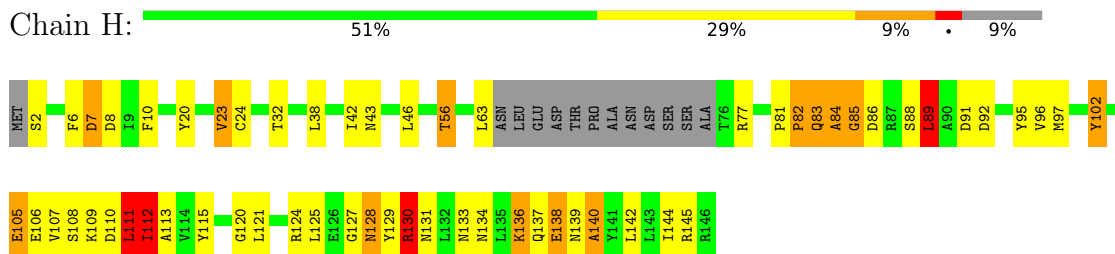


- Molecule 8: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide

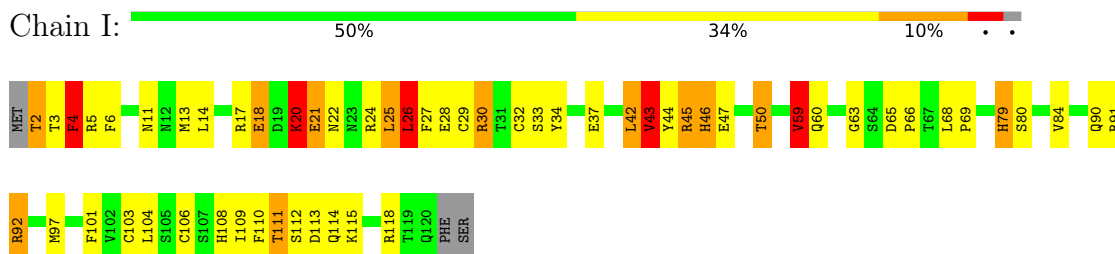
Chain F: 35% 16% 2% 46%



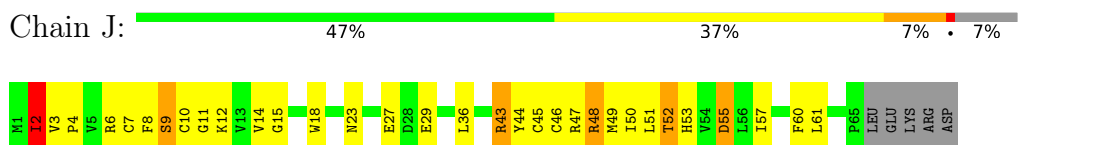
- Molecule 9: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide



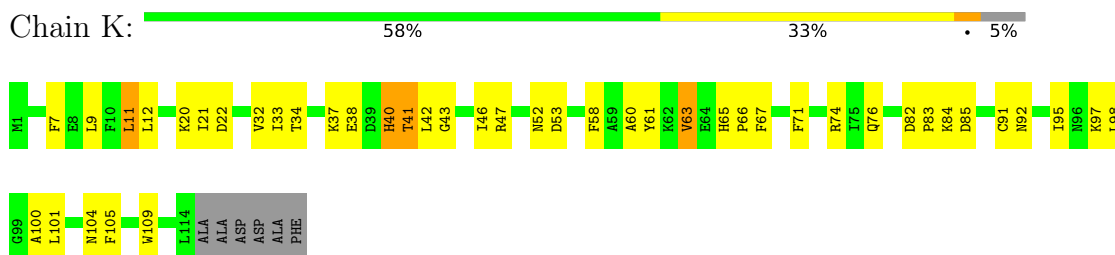
- Molecule 10: DNA-directed RNA polymerase II subunit 9



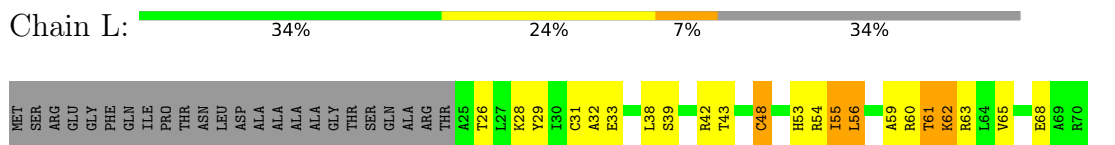
- Molecule 11: DNA-directed RNA polymerases I/II/III subunit 10



- Molecule 12: DNA-directed RNA polymerase II 13.6 kDa polypeptide



- Molecule 13: DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	169.65Å 222.34Å 194.32Å 90.00° 101.67° 90.00°	Depositor
Resolution (Å)	40.00 – 4.30 39.94 – 4.30	Depositor EDS
% Data completeness (in resolution range)	(Not available) (40.00-4.30) 86.4 (39.94-4.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.16	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.67 (at 4.28Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.270 , 0.332 0.277 , 0.330	Depositor DCC
R_{free} test set	4179 reflections (10.09%)	wwPDB-VP
Wilson B-factor (Å ²)	97.0	Xtrriage
Anisotropy	0.409	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.25 , 60.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	29002	wwPDB-VP
Average B, all atoms (Å ²)	126.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.04% 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 i

5.1 Standard geometry i

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

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	R	0.81	0/243	1.33	0/378
2	T	0.87	0/634	1.63	17/975 (1.7%)
3	N	0.77	0/317	1.35	1/488 (0.2%)
4	A	0.45	1/11180 (0.0%)	0.67	3/15117 (0.0%)
5	B	0.45	0/8866	0.65	1/11956 (0.0%)
6	C	0.43	0/2133	0.60	0/2891
7	E	0.44	0/1625	0.61	0/2182
8	F	0.43	0/682	0.61	1/922 (0.1%)
9	H	0.47	0/1086	0.77	2/1470 (0.1%)
10	I	0.48	0/989	0.74	0/1331
11	J	0.47	0/541	0.62	0/727
12	K	0.46	0/937	0.60	0/1265
13	L	0.53	0/365	0.79	0/485
All	All	0.47	1/29598 (0.0%)	0.72	25/40187 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
4	A	0	9
5	B	0	1
9	H	0	2
All	All	0	12

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1158	PRO	CG-CD	5.94	1.70	1.50

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
2	T	23	DC	O4'-C4'-C3'	-9.30	100.42	106.00
2	T	16	DC	O4'-C1'-N1	9.29	114.50	108.00
2	T	11	DG	O4'-C1'-N9	8.90	114.23	108.00
2	T	23	DC	O4'-C1'-N1	8.24	113.77	108.00
2	T	20	DC	O4'-C4'-C3'	-8.15	101.11	106.00

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	A	1068	ALA	Peptide
4	A	1069	ALA	Peptide
4	A	1070	GLN	Peptide
4	A	451	HIS	Peptide
4	A	974	ASP	Peptide

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	R	216	0	109	6	0
2	T	566	0	316	18	0
3	N	284	0	161	2	0
4	A	10984	0	11069	568	0
5	B	8701	0	8728	479	0
6	C	2095	0	2051	92	0
7	E	1594	0	1622	52	0
8	F	670	0	690	15	0
9	H	1068	0	1040	52	0
10	I	971	0	928	63	0
11	J	532	0	542	38	0
12	K	919	0	929	32	0
13	L	363	0	387	13	0
14	A	2	0	0	0	0
14	B	1	0	0	0	0
14	C	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
14	I	2	0	0	0	0
14	J	1	0	0	0	0
14	L	1	0	0	0	0
15	A	2	0	0	0	0
16	B	29	0	11	2	0
All	All	29002	0	28583	1299	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 1299 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)
4:A:701:LEU:HA	4:A:702:LEU:CB	1.68	1.20
4:A:975:HIS:HB3	4:A:976:THR:OG1	1.41	1.19
4:A:335:ARG:HD2	5:B:1202:LEU:HD12	1.27	1.16
4:A:1167:GLU:CB	4:A:1168:GLU:HA	1.75	1.15
5:B:1019:SER:HB2	5:B:1020:ARG:HB2	1.24	1.15

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
4	A	1384/1733 (80%)	1055 (76%)	224 (16%)	105 (8%)	1 15
5	B	1074/1224 (88%)	855 (80%)	161 (15%)	58 (5%)	2 22
6	C	264/318 (83%)	217 (82%)	35 (13%)	12 (4%)	2 24
7	E	185/215 (86%)	146 (79%)	26 (14%)	13 (7%)	1 17
8	F	81/155 (52%)	62 (76%)	15 (18%)	4 (5%)	2 23

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
9	H	129/146 (88%)	87 (67%)	30 (23%)	12 (9%)	0	12
10	I	117/122 (96%)	73 (62%)	32 (27%)	12 (10%)	0	9
11	J	63/70 (90%)	56 (89%)	4 (6%)	3 (5%)	2	23
12	K	112/120 (93%)	98 (88%)	13 (12%)	1 (1%)	17	56
13	L	44/70 (63%)	29 (66%)	11 (25%)	4 (9%)	1	12
All	All	3453/4173 (83%)	2678 (78%)	551 (16%)	224 (6%)	1	18

5 of 224 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	63	ARG
4	A	223	GLY
4	A	284	ALA
4	A	313	GLN
4	A	404	TYR

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	
4	A	1218/1520 (80%)	1055 (87%)	163 (13%)	4	20
5	B	951/1061 (90%)	827 (87%)	124 (13%)	4	21
6	C	234/274 (85%)	207 (88%)	27 (12%)	5	24
7	E	177/197 (90%)	160 (90%)	17 (10%)	8	30
8	F	73/137 (53%)	65 (89%)	8 (11%)	6	25
9	H	117/128 (91%)	103 (88%)	14 (12%)	5	23
10	I	113/116 (97%)	92 (81%)	21 (19%)	1	10
11	J	60/65 (92%)	51 (85%)	9 (15%)	3	17
12	K	99/102 (97%)	88 (89%)	11 (11%)	6	25
13	L	40/57 (70%)	31 (78%)	9 (22%)	1	6
All	All	3082/3657 (84%)	2679 (87%)	403 (13%)	4	21

5 of 403 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	B	740	HIS
5	B	1222	ARG
13	L	62	LYS
5	B	844	SER
5	B	1060	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 78 such sidechains are listed below:

Mol	Chain	Res	Type
5	B	1025	HIS
8	F	78	GLN
5	B	1177	HIS
6	C	224	GLN
10	I	116	ASN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	8/10 (80%)	1 (12%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	8	G

There are no RNA pucker outliers to report.

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 monosaccharides in this entry.

5.6 Ligand geometry

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
16	UTP	B	3000	15	22,30,30	2.08	4 (18%)	27,47,47	1.39	4 (14%)

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
16	UTP	B	3000	15	-	6/20/38/38	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
16	B	3000	UTP	C6-N1	5.25	1.42	1.35
16	B	3000	UTP	C4-N3	5.06	1.41	1.33
16	B	3000	UTP	O4'-C1'	3.87	1.46	1.41
16	B	3000	UTP	PG-O2G	3.32	1.61	1.50

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
16	B	3000	UTP	PB-O3B-PG	-3.37	121.26	132.83
16	B	3000	UTP	O1G-PG-O3B	2.57	113.25	104.64
16	B	3000	UTP	PB-O3A-PA	-2.46	124.38	132.83
16	B	3000	UTP	C5'-C4'-C3'	-2.01	107.63	115.18

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

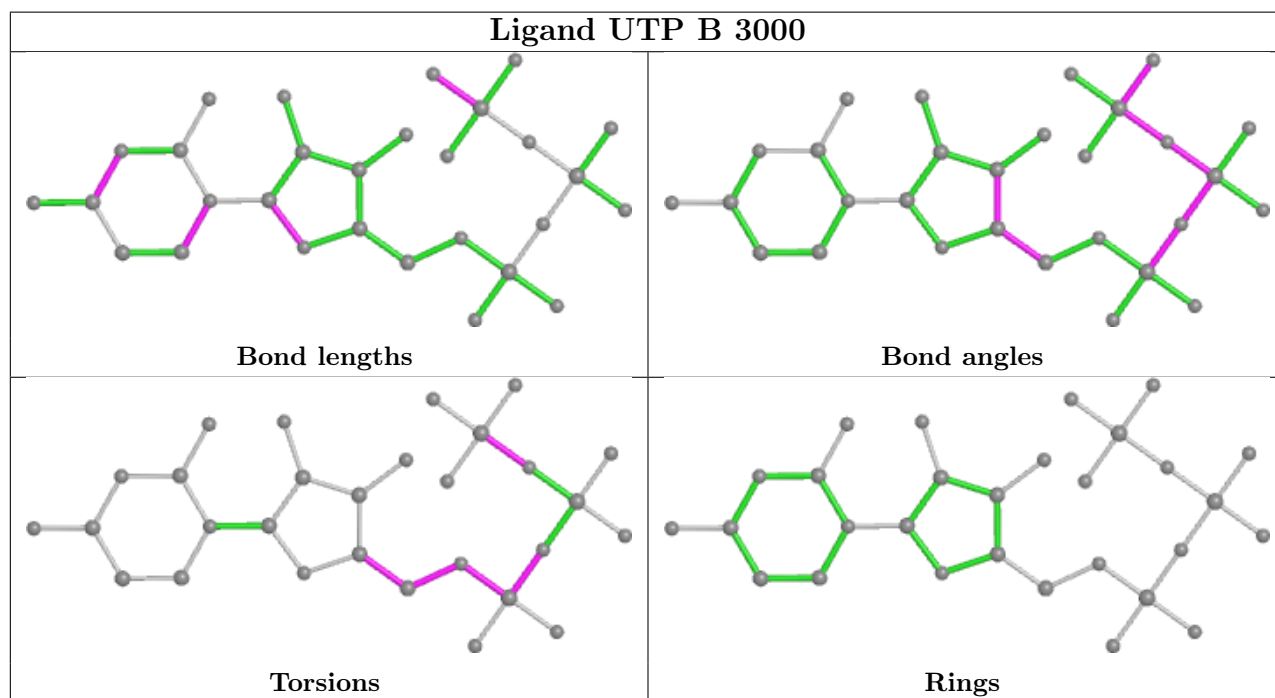
Mol	Chain	Res	Type	Atoms
16	B	3000	UTP	C5'-O5'-PA-O1A
16	B	3000	UTP	C4'-C5'-O5'-PA
16	B	3000	UTP	O4'-C4'-C5'-O5'
16	B	3000	UTP	C3'-C4'-C5'-O5'
16	B	3000	UTP	PB-O3A-PA-O5'

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
16	B	3000	UTP	2	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, 95th 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 ²)	Q<0.9
1	R	10/10 (100%)	-0.36	0 100 100	89, 111, 136, 141	0
2	T	28/28 (100%)	0.73	8 (28%) 0 0	81, 200, 285, 291	0
3	N	14/14 (100%)	1.65	7 (50%) 0 0	265, 273, 290, 292	0
4	A	1398/1733 (80%)	-0.38	4 (0%) 94 90	84, 118, 171, 184	0
5	B	1096/1224 (89%)	-0.38	1 (0%) 95 95	86, 114, 151, 165	0
6	C	266/318 (83%)	-0.42	1 (0%) 92 87	99, 118, 146, 150	0
7	E	193/215 (89%)	-0.32	0 100 100	98, 130, 163, 167	0
8	F	83/155 (53%)	-0.36	0 100 100	111, 127, 138, 144	0
9	H	133/146 (91%)	-0.26	0 100 100	120, 141, 176, 181	0
10	I	119/122 (97%)	-0.22	0 100 100	118, 149, 168, 173	0
11	J	65/70 (92%)	-0.56	0 100 100	103, 119, 134, 136	0
12	K	114/120 (95%)	-0.34	0 100 100	99, 118, 134, 136	0
13	L	46/70 (65%)	-0.06	0 100 100	137, 178, 191, 193	0
All	All	3565/4225 (84%)	-0.35	21 (0%) 89 84	81, 120, 168, 292	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	T	2	DT	4.9
2	T	3	DA	3.7
2	T	10	DA	3.2
3	N	11	DG	3.2
2	T	1	DC	3.1

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](#)

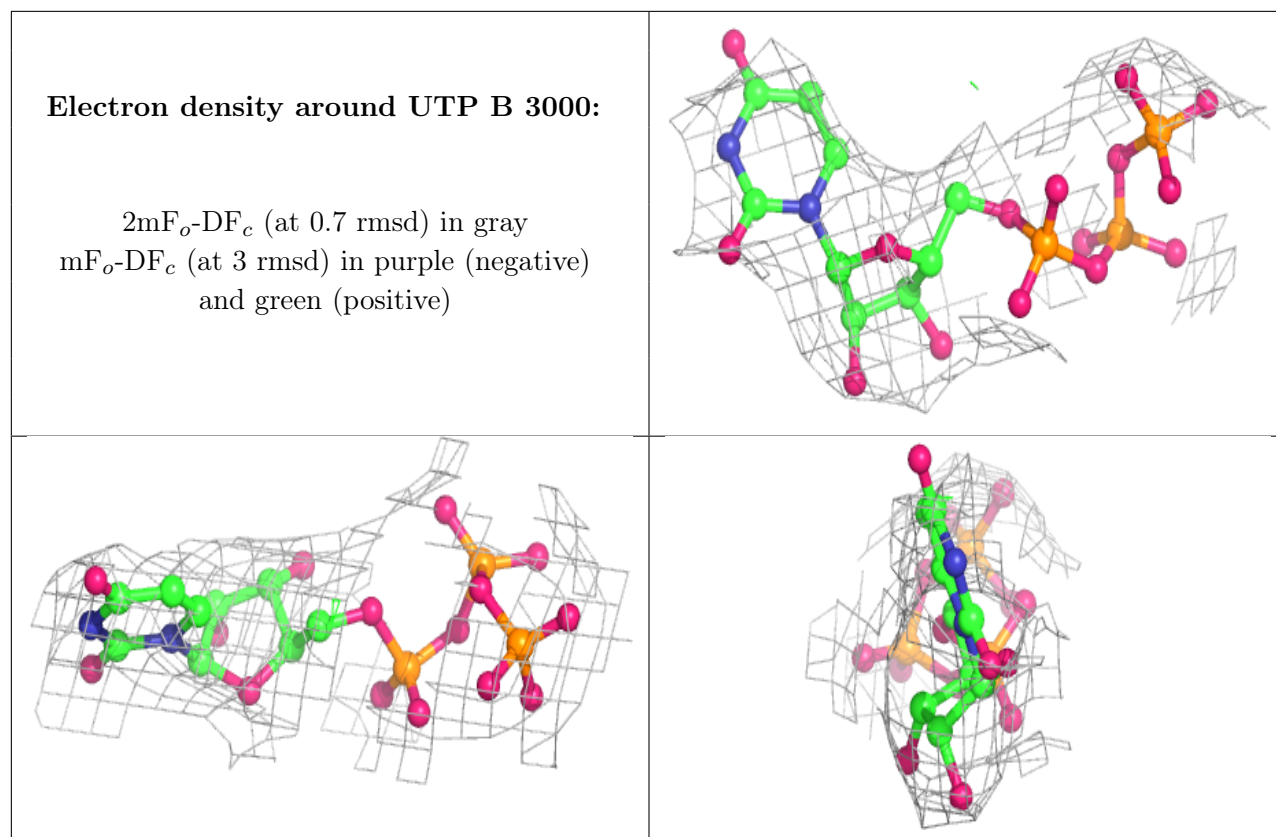
There are no monosaccharides in this entry.

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, 95th 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(\AA^2)	Q<0.9
15	MG	A	2001	1/1	0.84	0.14	97,97,97,97	0
15	MG	A	2002	1/1	0.88	0.19	102,102,102,102	0
14	ZN	A	1734	1/1	0.92	0.07	177,177,177,177	0
14	ZN	I	204	1/1	0.95	0.04	157,157,157,157	0
14	ZN	L	105	1/1	0.95	0.11	181,181,181,181	0
14	ZN	A	1735	1/1	0.95	0.05	165,165,165,165	0
14	ZN	I	203	1/1	0.95	0.09	121,121,121,121	0
16	UTP	B	3000	29/29	0.96	0.16	110,112,113,113	0
14	ZN	B	1307	1/1	0.97	0.05	147,147,147,147	0
14	ZN	J	101	1/1	0.97	0.12	112,112,112,112	0
14	ZN	C	319	1/1	0.98	0.03	117,117,117,117	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.