

# wwPDB EM Validation Summary Report (i)

#### Jun 26, 2024 - 04:51 PM JST

PDB ID	:	7XTD
EMDB ID	:	EMD-33447
Title	:	RNA polymerase II elongation complex transcribing a nucleosome (EC58oct)
Authors	:	Ehara, H.; Kujirai, T.; Shirouzu, M.; Kurumizaka, H.; Sekine, S.
Deposited on	:	2022-05-16
Resolution	:	3.90 Å(reported)

This is a wwPDB EM 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/EMValidationReportHelp 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 (1)) were used in the production of this report:

:	0.0.1. dev 92
:	4.02b-467
:	20191225.v01 (using entries in the PDB archive December 25th 2019)
:	1.9.13
:	Engh & Huber (2001)
:	Parkinson et al. (1996)
:	2.37.1
	: : : : :

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:  $ELECTRON\ MICROSCOPY$ 

The reported resolution of this entry is 3.90 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)	${ m EM~structures} \ (\#{ m Entries})$
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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 <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the EM map (all-atom inclusion < 40%). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain	
1	А	1743	80%	19%
2	В	1227	95%	5%
3	С	304	<b>●</b> 87%	13%
4	D	186	94%	6%
5	Е	214	99%	•
6	F	155	54% 46%	
7	G	171	15%	
8	Н	145	92%	8%



Mol	Chain	Length	Quality of chain	
0	т	115	30%	
9	1	110	97%	•
10	J	72	93%	7%
11	Κ	118	96%	•
1.2			6%	
12	L	72	62%	38%
13	М	113	56%	43%
14	Ν	198	34% 35% 20% •	43%
15	Р	19	21%	9/ 59/
10	1	10	36%	
16	Т	198	43% 18% ·	38%
17	V	108	37%	·
18	W	911	25%	41%
19	m	1503	37%	21%
20	n	417	13%	7%
			59%	
21	q	1084	86%	14%
22	r	544	49%	51%
23	u	459	45%	55%
		200	70%	
24	V	396	88%	12%
25	х	395	52%	48%
26	a	139	55%	45%
		100	64%	
26	е	139	64% ·	35%
27	b	106	78%	22%
27	f	106	73%	• 26%
28	с	133	77%	23%
		100	53%	
28	g	133	52% • 73%	47%
29	d	129	74%	26%
29	h	129	72%	28%

Continued from previous page...



Continued from previous page...

Mol	Chain	Length	Quality of chain							
	_		46%							
30	j	1008	46%	54%						
			81%		L					
31	k	531	79%	•	19%					



# 2 Entry composition (i)

There are 33 unique types of molecules in this entry. The entry contains 81668 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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.

Mol	Chain	Residues		Α	AltConf	Trace			
1	А	1404	Total 11064	C 6975	N 1930	O 2089	S 70	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	В	1164	Total 9284	C 5848	N 1639	O 1739	S 58	0	0

• Molecule 3 is a protein called RNA polymerase II third largest subunit B44, part of central core.

Mol	Chain	Residues		At	AltConf	Trace			
3	С	263	Total 2098	C 1319	N 354	0 413	S 12	0	0

• Molecule 4 is a protein called RNA polymerase II subunit B32.

Mol	Chain	Residues		At	oms	AltConf	Trace		
4	D	174	Total 1349	C 828	N 244	0 274	${ m S} { m 3}$	0	0

• Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues		At	AltConf	Trace			
5	Е	213	Total 1741	C 1094	N 312	0 325	S 10	0	0

• Molecule 6 is a protein called RNA polymerase subunit ABC23, common to RNA polymerases I, II, and III.



Mol	Chain	Residues		At	oms	AltConf	Trace		
6	F	84	Total 677	C 429	N 114	0 131	${ m S} { m 3}$	0	0

• Molecule 7 is a protein called RNA polymerase II subunit.

Mol	Chain	Residues		At	AltConf	Trace			
7	G	171	Total 1325	C 858	N 214	0 248	${ m S}{ m 5}$	0	0

• Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues		At	AltConf	Trace			
8	Н	133	Total 1053	C 671	N 169	O 209	$\frac{S}{4}$	0	0

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

Mol	Chain	Residues		$\mathbf{A}^{\dagger}$	AltConf	Trace			
9	Ι	111	Total 917	C 565	N 161	0 180	S 11	0	0

• Molecule 10 is a protein called RNA polymerase subunit ABC10-beta, common to RNA polymerases I, II, and III.

Mol	Chain	Residues		Atc	$\mathbf{ms}$	AltConf	Trace		
10	T	67	Total	С	Ν	Ο	$\mathbf{S}$	0	0
10	5	07	554	355	97	96	6	0	0

• Molecule 11 is a protein called RNA polymerase II subunit B12.5.

Mol	Chain	Residues		At	oms	AltConf	Trace		
11	K	113	Total 932	C 599	N 160	O 169	${S \atop 4}$	0	0

• Molecule 12 is a protein called RNA polymerase subunit ABC10-alpha.

Mol	Chain	Residues		Atc	$\mathbf{ms}$	AltConf	Trace		
12	L	45	Total 359	C 221	N 72	O 61	${ m S}{ m 5}$	0	0

• Molecule 13 is a protein called Transcription elongation factor 1 homolog.



Mol	Chain	Residues		Atc	$\mathbf{ms}$	AltConf	Trace		
12	М	64	Total	С	Ν	Ο	$\mathbf{S}$	0	0
10	111	04	505	318	82	99	6	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
М	-2	GLY	-	expression tag	UNP C4QZ45
М	-1	PRO	-	expression tag	UNP C4QZ45
М	0	GLY	-	expression tag	UNP C4QZ45

• Molecule 14 is a DNA chain called DNA (198-MER).

Mol	Chain	Residues		A	AltConf	Trace			
14	Ν	112	Total 2305	C 1099	N 389	O 705	Р 112	0	0

• Molecule 15 is a RNA chain called RNA (5'-R(P\*UP\*GP\*UP\*AP\*AP\*UP\*CP\*CP\*CP\*C P\*UP\*UP\*GP\*GP\*CP\*GP\*GP\*UP\*U)-3').

Mol	Chain	Residues		At	$\mathbf{oms}$	AltConf	Trace		
15	Р	19	Total 399	C 178	N 64	0 138	Р 19	0	0

• Molecule 16 is a DNA chain called DNA (198-MER).

Mol	Chain	Residues		$\mathbf{A}$	AltConf	Trace			
16	Т	123	Total 2512	C 1189	N 497	O 703	Р 123	0	0

• Molecule 17 is a protein called Transcription elongation factor SPT4.

Mol	Chain	Residues		At	oms	AltConf	Trace		
17	V	106	Total 824	C 512	N 150	0 155	S 7	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
V	7	MET	-	initiating methionine	UNP C4R0E6

• Molecule 18 is a protein called Transcription elongation factor SPT5.



Mol	Chain	Residues	Atoms					AltConf	Trace
18	W	533	Total 4232	C 2666	N 752	O 812	$\frac{S}{2}$	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
W	-2	GLY	-	expression tag	UNP C4R370
W	-1	PRO	-	expression tag	UNP C4R370
W	0	GLY	-	expression tag	UNP C4R370

• Molecule 19 is a protein called Transcription elongation factor Spt6.

Mol	Chain	Residues		Α	AltConf	Trace			
10		1107	Total	С	Ν	Ο	$\mathbf{S}$	0	0
19	111	1107	9730	6162	1663	1877	28	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
m	-2	GLY	-	expression tag	UNP C4R7H2
m	-1	PRO	-	expression tag	UNP C4R7H2
m	0	GLY	-	expression tag	UNP C4R7H2

• Molecule 20 is a protein called Protein that interacts with Spt6p and copurifies with Spt5p and RNA polymerase II.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	n	139	Total 1115	C 716	N 193	O 202	${S \atop 4}$	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
n	-2	GLY	-	expression tag	UNP C4R7L8
n	-1	PRO	-	expression tag	UNP C4R7L8
n	0	GLY	-	expression tag	UNP C4R7L8

• Molecule 21 is a protein called Component of the Paf1p complex.

Mol	Chain	Residues		Α	AltConf	Trace			
21	q	930	Total 7552	C 4805	N 1283	O 1439	S 25	0	0



Chain	Residue	Modelled	Actual	Comment	Reference
q	-39	MET	-	initiating methionine	UNP C4R6B2
q	-38	LYS	-	expression tag	UNP C4R6B2
q	-37	ASP	-	expression tag	UNP C4R6B2
q	-36	HIS	-	expression tag	UNP C4R6B2
q	-35	LEU	-	expression tag	UNP C4R6B2
q	-34	ILE	-	expression tag	UNP C4R6B2
q	-33	HIS	-	expression tag	UNP C4R6B2
q	-32	ASN	-	expression tag	UNP C4R6B2
q	-31	HIS	-	expression tag	UNP C4R6B2
q	-30	HIS	-	expression tag	UNP C4R6B2
q	-29	LYS	-	expression tag	UNP C4R6B2
q	-28	HIS	-	expression tag	UNP C4R6B2
q	-27	GLU	-	expression tag	UNP C4R6B2
q	-26	HIS	-	expression tag	UNP C4R6B2
q	-25	ALA	-	expression tag	UNP C4R6B2
q	-24	HIS	-	expression tag	UNP C4R6B2
q	-23	ALA	-	expression tag	UNP C4R6B2
q	-22	GLU	-	expression tag	UNP C4R6B2
q	-21	HIS	-	expression tag	UNP C4R6B2
q	-20	ASP	-	expression tag	UNP C4R6B2
q	-19	TYR	-	expression tag	UNP C4R6B2
q	-18	LYS	-	expression tag	UNP C4R6B2
q	-17	ASP	-	expression tag	UNP C4R6B2
q	-16	ASP	-	expression tag	UNP C4R6B2
q	-15	ASP	-	expression tag	UNP C4R6B2
q	-14	ASP	-	expression tag	UNP C4R6B2
q	-13	LYS	-	expression tag	UNP C4R6B2
q	-12	GLU	-	expression tag	UNP C4R6B2
q	-11	HIS	-	expression tag	UNP C4R6B2
q	-10	LEU	-	expression tag	UNP C4R6B2
q	-9	TYR	-	expression tag	UNP C4R6B2
q	-8	PHE	-	expression tag	UNP C4R6B2
q	-7	GLN	-	expression tag	UNP C4R6B2
q	-6	GLY	-	expression tag	UNP C4R6B2
q	-5	SER	-	expression tag	UNP C4R6B2
q	-4	SER	-	expression tag	UNP C4R6B2
q	-3	GLY	-	expression tag	UNP C4R6B2
q	-2	SER	-	expression tag	UNP C4R6B2
q	-1	SER	-	expression tag	UNP C4R6B2
q	0	GLY	-	expression tag	UNP C4R6B2

There are 40 discrepancies between the modelled and reference sequences:

• Molecule 22 is a protein called RNAPII-associated chromatin remodeling Paf1 complex sub-



unit.

Mol	Chain	Residues		At	AltConf	Trace			
22	r	266	Total 2139	C 1342	N 374	O 412	S 11	0	0

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
r	-29	MET	-	initiating methionine	UNP F2QQ42
r	-28	LYS	-	expression tag	UNP F2QQ42
r	-27	ASP	-	expression tag	UNP F2QQ42
r	-26	HIS	-	expression tag	UNP F2QQ42
r	-25	LEU	-	expression tag	UNP F2QQ42
r	-24	ILE	-	expression tag	UNP F2QQ42
r	-23	HIS	-	expression tag	UNP F2QQ42
r	-22	ASN	-	expression tag	UNP F2QQ42
r	-21	HIS	-	expression tag	UNP F2QQ42
r	-20	HIS	-	expression tag	UNP F2QQ42
r	-19	LYS	-	expression tag	UNP F2QQ42
r	-18	HIS	-	expression tag	UNP F2QQ42
r	-17	GLU	-	expression tag	UNP F2QQ42
r	-16	HIS	-	expression tag	UNP F2QQ42
r	-15	ALA	-	expression tag	UNP F2QQ42
r	-14	HIS	-	expression tag	UNP F2QQ42
r	-13	ALA	-	expression tag	UNP F2QQ42
r	-12	GLU	-	expression tag	UNP F2QQ42
r	-11	HIS	-	expression tag	UNP F2QQ42
r	-10	LEU	-	expression tag	UNP F2QQ42
r	-9	TYR	-	expression tag	UNP F2QQ42
r	-8	PHE	-	expression tag	UNP F2QQ42
r	-7	GLN	-	expression tag	UNP F2QQ42
r	-6	GLY	-	expression tag	UNP F2QQ42
r	-5	SER	-	expression tag	UNP F2QQ42
r	-4	SER	-	expression tag	UNP F2QQ42
r	-3	GLY	-	expression tag	UNP F2QQ42
r	-2	SER	-	expression tag	UNP F2QQ42
r	-1	SER	-	expression tag	UNP F2QQ42
r	0	GLY	-	expression tag	UNP $F2QQ42$

• Molecule 23 is a protein called Leo1.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	u	208	Total 1707	C 1063	N 304	O 337	${ m S} { m 3}$	0	0



Chain	Residue	Modelled	Actual	Comment	Reference
u	-29	MET	-	initiating methionine	UNP C4R3K1
u	-28	LYS	-	expression tag	UNP C4R3K1
u	-27	ASP	-	expression tag	UNP C4R3K1
u	-26	HIS	-	expression tag	UNP C4R3K1
u	-25	LEU	-	expression tag	UNP C4R3K1
u	-24	ILE	-	expression tag	UNP C4R3K1
u	-23	HIS	-	expression tag	UNP C4R3K1
u	-22	ASN	-	expression tag	UNP C4R3K1
u	-21	HIS	-	expression tag	UNP C4R3K1
u	-20	HIS	-	expression tag	UNP C4R3K1
u	-19	LYS	-	expression tag	UNP C4R3K1
u	-18	HIS	-	expression tag	UNP C4R3K1
u	-17	GLU	-	expression tag	UNP C4R3K1
u	-16	HIS	-	expression tag	UNP C4R3K1
u	-15	ALA	-	expression tag	UNP C4R3K1
u	-14	HIS	-	expression tag	UNP C4R3K1
u	-13	ALA	-	expression tag	UNP C4R3K1
u	-12	GLU	-	expression tag	UNP C4R3K1
u	-11	HIS	-	expression tag	UNP C4R3K1
u	-10	LEU	-	expression tag	UNP C4R3K1
u	-9	TYR	-	expression tag	UNP C4R3K1
u	-8	PHE	-	expression tag	UNP C4R3K1
u	-7	GLN	-	expression tag	UNP C4R3K1
u	-6	GLY	-	expression tag	UNP C4R3K1
u	-5	SER	-	expression tag	UNP C4R3K1
u	-4	SER	-	expression tag	UNP C4R3K1
u	-3	GLY	-	expression tag	UNP C4R3K1
u	-2	SER	-	expression tag	UNP C4R3K1
u	-1	SER	-	expression tag	UNP C4R3K1
u	0	GLY	-	expression tag	UNP C4R3K1

There are 30 discrepancies between the modelled and reference sequences:

• Molecule 24 is a protein called RNAP II-associated protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	v	349	Total 2878	C 1835	N 510	O 528	${ m S}{ m 5}$	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
V	-2	GLY	-	expression tag	UNP C4R997



Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
V	-1	SER	-	expression tag	UNP C4R997
V	0	ALA	-	expression tag	UNP C4R997

• Molecule 25 is a protein called Constituent of Paf1 complex with RNA polymerase II, Paf1p, Hpr1p, Ctr9, Leo1, Rtf1 and Ccr4p.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	х	205	Total 1682	C 1086	N 287	O 307	${ m S} { m 2}$	0	0

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
X	-29	MET	-	initiating methionine	UNP C4R1E6
X	-28	LYS	-	expression tag	UNP C4R1E6
X	-27	ASP	-	expression tag	UNP C4R1E6
X	-26	HIS	-	expression tag	UNP C4R1E6
X	-25	LEU	-	expression tag	UNP C4R1E6
X	-24	ILE	-	expression tag	UNP C4R1E6
X	-23	HIS	-	expression tag	UNP C4R1E6
X	-22	ASN	-	expression tag	UNP C4R1E6
X	-21	HIS	-	expression tag	UNP C4R1E6
X	-20	HIS	-	expression tag	UNP C4R1E6
X	-19	LYS	-	expression tag	UNP C4R1E6
X	-18	HIS	-	expression tag	UNP C4R1E6
X	-17	GLU	-	expression tag	UNP C4R1E6
X	-16	HIS	-	expression tag	UNP C4R1E6
X	-15	ALA	-	expression tag	UNP C4R1E6
X	-14	HIS	-	expression tag	UNP C4R1E6
X	-13	ALA	-	expression tag	UNP C4R1E6
X	-12	GLU	-	expression tag	UNP C4R1E6
X	-11	HIS	-	expression tag	UNP C4R1E6
X	-10	LEU	-	expression tag	UNP C4R1E6
X	-9	TYR	-	expression tag	UNP C4R1E6
X	-8	PHE	-	expression tag	UNP C4R1E6
X	-7	GLN	-	expression tag	UNP C4R1E6
X	-6	GLY	-	expression tag	UNP C4R1E6
X	-5	SER	-	expression tag	UNP C4R1E6
X	-4	SER	-	expression tag	UNP C4R1E6
X	-3	GLY	-	expression tag	UNP C4R1E6
X	-2	SER	-	expression tag	UNP C4R1E6
X	-1	SER	-	expression tag	UNP C4R1E6



Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
Х	0	GLY	-	expression tag	UNP C4R1E6

• Molecule 26 is a protein called Histone H3.3.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	76	76	Total	С	Ν	0	S	0	0
20 a	10	615	390	115	108	2	0	0	
26	0	00	Total	С	Ν	0	S	0	0
20 e	90	735	464	140	129	2	0	0	

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference	
a	-3	GLY	-	expression tag	UNP P84243	
a	-2	SER	-	expression tag	UNP P84243	
a	-1	HIS	-	expression tag	UNP P84243	
e	-3	GLY	-	expression tag	UNP P84243	
e	-2	SER	-	expression tag	UNP P84243	
e	-1	HIS	-	expression tag	UNP P84243	

• Molecule 27 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	b	83	Total 662	C 418	N 129	0 114	S 1	0	0
27	f	78	Total 619	C 391	N 120	0 107	S 1	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference	
b	-3	GLY	-	expression tag	UNP P62805	
b	-2	SER	-	expression tag	UNP P62805	
b	-1	HIS	-	expression tag	UNP P62805	
f	-3	GLY	-	expression tag	UNP P62805	
f	-2	SER	-	expression tag	UNP P62805	
f	-1	HIS	-	expression tag	UNP P62805	

• Molecule 28 is a protein called Histone H2A type 1-B/E.



Mol	Chain	Residues	Atoms				AltConf	Trace
28	С	103	Total 796	C 502	N 155	O 139	0	0
28	g	70	Total 544	C 342	N 109	O 93	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
с	-3	GLY	-	expression tag	UNP P04908
с	-2	SER	-	expression tag	UNP P04908
с	-1	HIS	-	expression tag	UNP P04908
g	-3	GLY	-	expression tag	UNP P04908
g	-2	SER	-	expression tag	UNP P04908
g	-1	HIS	-	expression tag	UNP P04908

• Molecule 29 is a protein called Histone H2B type 1-J.

Mol	Chain	Residues		At	oms			AltConf	Trace
20	d	05	Total	С	Ν	0	S	0	0
29	29 a	90	746	468	136	140	2	0	0
20	h	03	Total	С	Ν	0	S	0	0
29	11	90	725	456	130	137	2	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
d	-6	GLY	-	expression tag	UNP P06899
d	-5	SER	-	expression tag	UNP P06899
d	-4	HIS	-	expression tag	UNP P06899
h	-6	GLY	-	expression tag	UNP P06899
h	-5	SER	-	expression tag	UNP P06899
h	-4	HIS	-	expression tag	UNP P06899

• Molecule 30 is a protein called FACT complex subunit.

Mol	Chain	Residues	Atoms			AltConf	Trace		
30	j	466	Total 3780	C 2401	N 658	0 708	S 13	0	0

There are 3 discrepancies between the modelled and reference sequences:



Chain	Residue	Modelled	Actual	Comment	Reference
j	-2	GLY	-	expression tag	UNP C4QYQ8
j	-1	PRO	-	expression tag	UNP C4QYQ8
j	0	GLY	-	expression tag	UNP C4QYQ8

• Molecule 31 is a protein called FACT complex subunit POB3.

Mol	Chain	Residues	Atoms			AltConf	Trace		
31	k	429	Total 3502	C 2215	N 613	O 664	S 10	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
k	-2	GLY	-	expression tag	UNP F2QNN8
k	-1	PRO	-	expression tag	UNP F2QNN8
k	0	GLY	-	expression tag	UNP F2QNN8

• Molecule 32 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	AltConf
32	А	2	Total Zn 2 2	0
32	В	1	Total Zn 1 1	0
32	С	1	Total Zn 1 1	0
32	Ι	2	Total Zn 2 2	0
32	J	1	Total Zn 1 1	0
32	L	1	Total Zn 1 1	0
32	М	1	Total Zn 1 1	0
32	V	1	Total Zn 1 1	0

• Molecule 33 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	AltConf
33	А	1	Total Mg 1 1	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 atom inclusion in map 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 diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). 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 subunit



• Molecule 2: DNA-directed RNA polymerase subunit beta







• Molecule 3: RNA polymerase II third largest subunit B44, part of central core



Chain H:

8%

92%



• Molecule 9: DNA-directed RNA polymerase subunit



 $\bullet$  Molecule 10: RNA polymerase subunit ABC10-beta, common to RNA polymerases I, II, and III







#### DG DA DT

• Molecule 15: RNA (5'-R(P\*UP\*GP\*UP\*AP\*AP\*UP\*CP\*CP\*CP\*CP\*UP\*UP\*GP\*GP\*CP\* GP\*GP\*UP\*U)-3')











#### ALA GLY ALA ASP ASP GLY TYR SER ASN ALA ALA ALA ALA ATA TYR

 $\bullet$  Molecule 20: Protein that interacts with Spt6p and copurifies with Spt5p and RNA polymerase II











• Molecule 26: Histone H3.3



• Molecule 28: Histone H2A type 1-B/E







Chain k:

19%





# 4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	38828	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{\AA}^2)$	51	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	0.096	Depositor
Minimum map value	-0.052	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.003	Depositor
Recommended contour level	0.009	Depositor
Map size (Å)	356.16, 356.16, 356.16	wwPDB
Map dimensions	240, 240, 240	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.484, 1.484, 1.484	Depositor



# 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: MG, 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).

Mal	Chain	Bo	ond lengths	Bond angles		
	Unam	RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	А	0.32	0/11267	0.49	0/15222	
2	В	0.34	0/9464	0.53	2/12763~(0.0%)	
3	С	0.33	0/2139	0.49	0/2895	
4	D	0.24	0/1361	0.46	0/1837	
5	Ε	0.30	0/1773	0.50	0/2385	
6	F	0.33	0/687	0.50	0/931	
7	G	0.28	0/1354	0.48	0/1837	
8	Н	0.33	0/1070	0.49	0/1444	
9	Ι	0.24	0/934	0.50	0/1257	
10	J	0.36	0/563	0.51	0/753	
11	Κ	0.32	0/953	0.50	0/1291	
12	L	0.37	0/365	0.62	0/484	
13	М	0.25	0/513	0.41	0/693	
14	Ν	1.13	23/2575~(0.9%)	1.32	34/3978~(0.9%)	
15	Р	0.43	0/443	0.86	2/687~(0.3%)	
16	Т	1.14	17/2827~(0.6%)	1.29	41/4352~(0.9%)	
17	V	0.24	0/840	0.50	0/1140	
18	W	0.25	0/4300	0.49	0/5812	
19	m	0.24	0/9925	0.45	0/13424	
20	n	0.24	0/1132	0.42	0/1526	
21	q	0.24	0/7689	0.40	0/10368	
22	r	0.30	0/2169	0.53	0/2901	
23	u	0.34	0/1740	0.57	2/2347~(0.1%)	
24	V	0.25	0/2944	0.47	0/3973	
25	Х	0.26	0/1716	0.45	0/2310	
26	a	0.29	0/622	0.54	0/834	
26	е	0.33	0/743	0.61	1/996~(0.1%)	
27	b	0.30	0/669	0.59	0/894	
27	f	0.34	0/626	0.63	1/837~(0.1%)	
28	с	0.29	0/806	0.55	0/1089	
28	g	0.34	0/552	0.59	1/743~(0.1%)	
29	d	0.30	0/757	0.51	0/1015	



Mal	Chain	Bo	ond lengths	Bond angles		
IVIOI	Unam	RMSZ	# Z  > 5	RMSZ	# Z  > 5	
29	h	0.28	0/736	0.47	0/990	
30	j	0.68	0/3855	0.86	4/5191~(0.1%)	
31	k	0.68	0/3579	0.88	9/4833~(0.2%)	
All	All	0.44	40/83688~(0.0%)	0.63	97/114032~(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
14	Ν	0	5
16	Т	0	2
30	j	0	1
31	k	0	2
All	All	0	10

The worst 5 of 40 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
14	Ν	-40	DC	C1'-N1	6.95	1.58	1.49
16	Т	30	DT	C1'-N1	6.67	1.57	1.49
16	Т	37	DC	C1'-N1	6.60	1.57	1.49
14	Ν	-96	DT	C1'-N1	6.50	1.57	1.49
14	Ν	-27	DC	C1'-N1	6.47	1.57	1.49

The worst 5 of 97 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Ζ	$Observed(^{o})$	$\operatorname{Ideal}(^{o})$
30	j	684	ARG	NE-CZ-NH1	9.40	125.00	120.30
30	j	836	ARG	NE-CZ-NH1	9.39	125.00	120.30
16	Т	105	DA	N1-C6-N6	-9.10	113.14	118.60
16	Т	31	DT	O4'-C1'-N1	9.04	114.32	108.00
15	Р	10	U	OP1-P-O3'	8.79	124.55	105.20

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
14	Ν	-101	DG	Sidechain
14	Ν	-102	DT	Sidechain
	-	0	7	



Continued from previous page...

Mol	Chain	Res	Type	Group
14	Ν	-108	DC	Sidechain
14	Ν	-111	DG	Sidechain
14	Ν	-114	DT	Sidechain

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

Due to software issues we are unable to calculate clashes - this section is therefore empty.

#### 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 PDB entries followed by that with respect to all EM entries.

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	Perce	entiles
1	А	1392/1743~(80%)	1354 (97%)	37~(3%)	1 (0%)	51	84
2	В	1154/1227~(94%)	1117 (97%)	37~(3%)	0	100	100
3	С	261/304~(86%)	259 (99%)	2(1%)	0	100	100
4	D	170/186~(91%)	166 (98%)	4 (2%)	0	100	100
5	E	211/214~(99%)	205 (97%)	6 (3%)	0	100	100
6	F	82/155~(53%)	80 (98%)	2(2%)	0	100	100
7	G	169/171~(99%)	166 (98%)	3(2%)	0	100	100
8	Н	129/145~(89%)	125 (97%)	4 (3%)	0	100	100
9	Ι	109/115~(95%)	107 (98%)	2(2%)	0	100	100
10	J	65/72~(90%)	65 (100%)	0	0	100	100
11	K	111/118~(94%)	110 (99%)	1 (1%)	0	100	100
12	L	43/72~(60%)	41 (95%)	2(5%)	0	100	100
13	М	62/113~(55%)	62 (100%)	0	0	100	100
17	V	104/108~(96%)	100 (96%)	4 (4%)	0	100	100
18	W	527/911~(58%)	506 (96%)	21 (4%)	0	100	100
19	m	1179/1503~(78%)	1157 (98%)	22 (2%)	0	100	100



Mol	Chain	Analysed	Favoured	Allowed	Outliers	Perce	ntiles
20	n	137/417~(33%)	136~(99%)	1 (1%)	0	100	100
21	q	928/1084~(86%)	922 (99%)	6 (1%)	0	100	100
22	r	260/544~(48%)	254~(98%)	6 (2%)	0	100	100
23	u	206/459~(45%)	204 (99%)	2(1%)	0	100	100
24	v	341/396~(86%)	327~(96%)	14 (4%)	0	100	100
25	х	201/395~(51%)	200 (100%)	1 (0%)	0	100	100
26	a	74/139~(53%)	70~(95%)	4(5%)	0	100	100
26	е	88/139~(63%)	87~(99%)	1 (1%)	0	100	100
27	b	81/106~(76%)	78~(96%)	3~(4%)	0	100	100
27	f	76/106~(72%)	73~(96%)	3~(4%)	0	100	100
28	с	101/133~(76%)	100 (99%)	1 (1%)	0	100	100
28	g	68/133~(51%)	66~(97%)	2(3%)	0	100	100
29	d	93/129~(72%)	89~(96%)	4 (4%)	0	100	100
29	h	91/129~(70%)	90~(99%)	1 (1%)	0	100	100
30	j	462/1008~(46%)	449 (97%)	12 (3%)	1 (0%)	47	79
31	k	423/531~(80%)	411 (97%)	12 (3%)	0	100	100
All	All	9398/13005~(72%)	9176~(98%)	220 (2%)	2 (0%)	100	100

Continued from previous page...

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	А	960	VAL
30	j	543	LYS

#### 5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent side chain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

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	А	1219/1528~(80%)	1216 (100%)	3~(0%)	93 96		
2	В	1018/1077~(94%)	1016 (100%)	2~(0%)	93 96		



Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers Percen		ntiles
3	С	$236/264~(\overline{89\%})$	236 (100%)	0	100	100
4	D	149/160~(93%)	149 (100%)	0	100	100
5	Ε	196/197~(100%)	194 (99%)	2 (1%)	76	86
6	F	75/137~(55%)	75~(100%)	0	100	100
7	G	148/148 (100%)	148 (100%)	0	100	100
8	Н	120/130~(92%)	120 (100%)	0	100	100
9	Ι	106/109~(97%)	106 (100%)	0	100	100
10	J	61/66~(92%)	61 (100%)	0	100	100
11	Κ	104/109~(95%)	104 (100%)	0	100	100
12	L	38/56~(68%)	38~(100%)	0	100	100
13	М	61/99~(62%)	60~(98%)	1 (2%)	62	79
17	V	90/92~(98%)	90 (100%)	0	100	100
18	W	480/796~(60%)	480 (100%)	0	100	100
19	m	1087/1354~(80%)	1086 (100%)	1 (0%)	93	97
20	n	125/361~(35%)	125 (100%)	0	100	100
21	q	824/962~(86%)	824 (100%)	0	100	100
22	r	239/485~(49%)	239 (100%)	0	100	100
23	u	192/406~(47%)	191 (100%)	1 (0%)	88	93
24	v	325/369~(88%)	325~(100%)	0	100	100
25	х	190/354~(54%)	190 (100%)	0	100	100
26	a	64/112~(57%)	64 (100%)	0	100	100
26	е	77/112~(69%)	77~(100%)	0	100	100
27	b	68/81~(84%)	68 (100%)	0	100	100
27	f	63/81~(78%)	63 (100%)	0	100	100
28	с	82/102~(80%)	82 (100%)	0	100	100
28	g	54/102~(53%)	54 (100%)	0	100	100
29	d	81/107 (76%)	81 (100%)	0	100	100
29	h	79/107 (74%)	79 (100%)	0	100	100
30	j	417/910 (46%)	417 (100%)	0	100	100
31	k	392/474~(83%)	392 (100%)	0	100	100
All	All	8460/11447~(74%)	8450 (100%)	10 (0%)	93	97



 $5~{\rm of}~10$  residues with a non-rotameric side chain are listed below:

Mol	Chain	$\mathbf{Res}$	Type
13	М	40	LYS
19	m	514	LYS
23	u	124	ARG
2	В	463	LYS
2	В	904	ARG

Sometimes side chains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such side chains are listed below:

Mol	Chain	$\mathbf{Res}$	Type
25	Х	296	ASN
27	b	25	ASN
31	k	141	ASN
30	j	750	GLN
31	k	127	ASN

#### 5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
15	Р	19/19~(100%)	10~(52%)	1 (5%)

5 of 10 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
15	Р	-6	G
15	Р	-5	U
15	Р	-4	А
15	Р	-3	А
15	Р	-2	U

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
15	Р	-7	U

#### 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 (i)

Of 11 ligands modelled in this entry, 11 are monoatomic - leaving 0 for Mogul analysis.

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.

#### 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 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-33447. These allow visual inspection of the internal detail of the map and identification of artifacts.

Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

#### 6.1 Orthogonal projections (i)

#### 6.1.1 Primary map



6.1.2 Raw map



The images above show the map projected in three orthogonal directions.



#### 6.2 Central slices (i)

#### 6.2.1 Primary map



X Index: 120





Z Index: 120

#### 6.2.2 Raw map



X Index: 120

Y Index: 120



The images above show central slices of the map in three orthogonal directions.



#### 6.3 Largest variance slices (i)

#### 6.3.1 Primary map



X Index: 141



Y Index: 125



Z Index: 137

#### 6.3.2 Raw map



X Index: 141

Y Index: 121



The images above show the largest variance slices of the map in three orthogonal directions.



#### 6.4 Orthogonal standard-deviation projections (False-color) (i)

#### 6.4.1 Primary map



6.4.2 Raw map



The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.



#### 6.5 Orthogonal surface views (i)

6.5.1 Primary map



The images above show the 3D surface view of the map at the recommended contour level 0.009. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

#### 6.5.2 Raw map



These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

#### 6.6 Mask visualisation (i)

This section was not generated. No masks/segmentation were deposited.



# 7 Map analysis (i)

This section contains the results of statistical analysis of the map.

#### 7.1 Map-value distribution (i)



The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.



#### 7.2 Volume estimate (i)



The volume at the recommended contour level is 526  $\rm nm^3;$  this corresponds to an approximate mass of 475 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.



#### 7.3 Rotationally averaged power spectrum (i)



\*Reported resolution corresponds to spatial frequency of 0.256  ${\rm \AA^{-1}}$ 



## 8 Fourier-Shell correlation (i)

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

#### 8.1 FSC (i)



\*Reported resolution corresponds to spatial frequency of 0.256  $\mathrm{\AA^{-1}}$ 



### 8.2 Resolution estimates (i)

$\begin{bmatrix} Bosolution ostimato (Å) \end{bmatrix}$	Estimation criterion (FSC cut-off)		
Resolution estimate (A)	0.143	0.5	Half-bit
Reported by author	3.90	-	-
Author-provided FSC curve	3.85	4.64	3.91
Unmasked-calculated*	4.19	7.59	4.34

\*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.



# 9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-33447 and PDB model 7XTD. Per-residue inclusion information can be found in section 3 on page 16.

#### 9.1 Map-model overlay (i)



The images above show the 3D surface view of the map at the recommended contour level 0.009 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.



#### 9.2 Q-score mapped to coordinate model (i)



The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

#### 9.3 Atom inclusion mapped to coordinate model (i)



The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.009).



#### 9.4 Atom inclusion (i)



At the recommended contour level, 60% of all backbone atoms, 47% of all non-hydrogen atoms, are inside the map.



#### 9.5 Map-model fit summary (i)

The table lists the average atom inclusion at the recommended contour level (0.009) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	$\mathbf{Q} extsf{-score}$
All	0.4720	0.2210
А	0.8070	0.4220
В	0.7830	0.4370
С	0.8350	0.4550
D	0.5840	0.2180
Е	0.8350	0.4030
F	0.8500	0.4680
G	0.6570	0.3200
Н	0.8420	0.4480
Ι	0.5620	0.2560
J	0.8260	0.4530
K	0.8580	0.4640
L	0.7900	0.4030
М	0.4520	0.1740
N	0.3410	0.0870
Р	0.7520	0.3230
Т	0.3710	0.1150
V	0.5460	0.1410
W	0.4460	0.2030
a	0.0440	0.0430
b	0.0130	0.0010
с	0.0060	0.0020
d	0.0190	0.0010
e	0.0280	0.0190
f	0.0290	-0.0230
g	0.0040	0.0030
h	0.0040	0.0010
j	0.0480	0.0270
k	0.0060	0.0010
m	0.4210	0.1310
n	0.4850	0.1470
q	0.2850	0.0790
r	0.2050	0.0950
u	0.3130	0.1560
V	0.2020	0.0990
X	0.4440	0.2670



