

Dec 30, 2024 – 12:21 AM EST

P	DB ID	:	7ZSB
EM	IDB ID	:	EMD-14929
	Title	:	Yeast RNA polymerase II transcription pre-initiation complex with the $+1$
			nucleosome and NTP, complex C
A	Authors	:	Wang, H.; Cramer, P.
Depos	ited on	:	2022-05-06
Res	olution	:	6.60 Å(reported)
_			
	This is	a I	Full wwPDB EM Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org* A user guide is available at https://www.wwpdb.org/validation/2017/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 (i)) were used in the production of this report:

EMDB validation analysis	:	0.0.1.dev113
Mogul	:	2022.3.0, CSD as543be (2022)
MolProbity	:	4.02b-467
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

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 6.60 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	2	Percentile Ranks	Value
Ramachandran outliers			0
Sidechain outliers			0.3%
	Worse		Better
	Percentile relativ	ve to all structures	
	Percentile relativ	ve to all EM structures	
		Whole archive	EM structures
Metric			

\mathbf{Metric}	(# Entries)	(# Entries)
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

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 $\geq=3, 2, 1$ 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 EM map (all-atom inclusion < 40%). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain	
1	А	1733	82%	18%
2	В	1224	96%	•
3	С	347	76%	23%
4	D	221	75%	24%
5	Е	215	99%	
6	F	155	75% •	24%
7	G	177	97%	·
8	Н	146	95%	
9	Ι	122	95%	5%



Continue	nued fron	n previous	page	
Mol	Chain	Length	Quality of chain	
10	J	70	97%	
11	K	120	95%	• •
12	L	70	64%	36%
13	М	352	88%	12%
14	Ν	219	100%	
15	0	247	73%	27%
16	Q	738	30% 70%	
17	R	400	67%	33%
18	Т	219	100%	
19	U	171	• 62% •	37%
20	V	129	79%	• 19%
21	W	492	5% 62%	38%
22	Х	328	6 4% •	36%
23	0	778	97%	·
24	1	645	80%	19%
25	2	517	87%	13%
26	3	324	40% 60%	
27	4	341	89%	10%
28	5	76	83%	• 14%
29	6	464	82%	17%
30	7	843	73%	27%
31	a	135	72%	28%
31	е	135	73%	27%
32	b	102	79%	• 20%
32	f	102	78%	22%



~			
	P.		
\mathbf{I}	trom	nromanic	naae
	110110	$p_1 c_0 c_0 a_0$	puuc
	1		p

Mol	Chain	Length	Quality of chain		
33	С	129	5% 84%		16%
33	g	129	81%	•	18%
34	d	125	77%	•	22%
34	h	125	76%		24%



2 Entry composition (i)

There are 37 unique types of molecules in this entry. The entry contains 86808 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 II subunit RPB1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	А	1426	Total 11221	C 7070	N 1960	O 2129	S 62	0	0

• Molecule 2 is a protein called DNA-directed RNA polymerase II subunit RPB2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	В	1180	Total 9404	C 5946	N 1643	O 1760	${ m S}{55}$	0	0

• Molecule 3 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	С	266	Total 2092	C 1315	N 348	0 416	S 13	0	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
С	-28	MET	-	initiating methionine	UNP P16370
С	-27	GLY	-	expression tag	UNP P16370
С	-26	SER	-	expression tag	UNP P16370
С	-25	HIS	-	expression tag	UNP P16370
С	-24	HIS	-	expression tag	UNP P16370
С	-23	HIS	-	expression tag	UNP P16370
С	-22	HIS	-	expression tag	UNP P16370
С	-21	HIS	-	expression tag	UNP P16370
С	-20	HIS	-	expression tag	UNP P16370
С	-19	SER	-	expression tag	UNP P16370
С	-18	ASN	-	expression tag	UNP P16370
С	-17	SER	-	expression tag	UNP P16370
C	-16	GLY	-	expression tag	UNP P16370
С	-15	LEU	-	expression tag	UNP P16370
C	-14	ASN	-	expression tag	UNP P16370



Chain	Residue	Modelled	Actual	Comment	Reference
С	-13	ASP	-	expression tag	UNP P16370
С	-12	ILE	-	expression tag	UNP P16370
С	-11	PHE	-	expression tag	UNP P16370
С	-10	GLU	-	expression tag	UNP P16370
С	-9	ALA	-	expression tag	UNP P16370
С	-8	GLN	-	expression tag	UNP P16370
С	-7	LYS	-	expression tag	UNP P16370
С	-6	ILE	-	expression tag	UNP P16370
С	-5	GLU	-	expression tag	UNP P16370
С	-4	TRP	-	expression tag	UNP P16370
С	-3	HIS	-	expression tag	UNP P16370
С	-2	GLU	-	expression tag	UNP P16370
С	-1	ASP	-	expression tag	UNP P16370
С	0	THR	-	expression tag	UNP P16370
С	1	GLY	-	expression tag	UNP P16370
С	2	SER	-	expression tag	UNP P16370
С	3	SER	-	expression tag	UNP P16370

Continued from previous page...

• Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB4.

Mol	Chain	Residues		At	oms	AltConf	Trace		
4	D	167	Total 1343	C 829	N 242	O 270	${ m S} { m 2}$	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	214	Total 1752	C 1111	N 309	O 321	S 11	0	0

• Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues		At	oms	AltConf	Trace		
6	F	118	Total 977	C 620	N 161	0 193	${ m S} { m 3}$	0	0

• Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB7.

Mol	Chain	Residues		At	oms	AltConf	Trace		
7	G	171	Total 1339	C 861	N 222	O 248	S 8	0	0



Chain	Residue	Modelled	Actual	Comment	Reference
G	172	HIS	-	expression tag	UNP P34087
G	173	HIS	-	expression tag	UNP P34087
G	174	HIS	-	expression tag	UNP P34087
G	175	HIS	-	expression tag	UNP P34087
G	176	HIS	-	expression tag	UNP P34087
G	177	HIS	-	expression tag	UNP P34087

There are 6 discrepancies between the modelled and reference sequences:

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

Mol	Chain	Residues	Atoms					AltConf	Trace
8	Н	140	Total 1120	С 704	N 188	0 224	$\frac{S}{4}$	0	0

• Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues		\mathbf{A}	AltConf	Trace			
9	Ι	116	Total 944	C 581	N 172	0 181	S 10	0	0

• Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

Mol	Chain	Residues		At	AltConf	Trace			
10	J	69	Total 569	C 362	N 101	O 100	S 6	0	0

• Molecule 11 is a protein called DNA-directed RNA polymerase II subunit RPB11.

Mol	Chain	Residues		At	oms	AltConf	Trace		
11	K	115	Total 924	C 593	N 157	0 172	${ m S} { m 2}$	0	0

• Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues		Ato	\mathbf{ms}	AltConf	Trace		
12	L	45	Total 359	C 221	N 71	O 63	$\frac{S}{4}$	0	0

• Molecule 13 is a protein called Transcription initiation factor IIB.



Mol	Chain	Residues		At	AltConf	Trace			
13	М	310	Total 2379	C 1504	N 408	O 449	S 18	0	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
М	346	LYS	-	expression tag	UNP P29055
М	347	HIS	-	expression tag	UNP P29055
М	348	HIS	-	expression tag	UNP P29055
М	349	HIS	-	expression tag	UNP P29055
М	350	HIS	-	expression tag	UNP P29055
М	351	HIS	-	expression tag	UNP P29055
М	352	HIS	-	expression tag	UNP P29055

• Molecule 14 is a DNA chain called Non-template DNA (219-MER).

Mol	Chain	Residues		Α	toms			AltConf	Trace
14	Ν	219	Total 4463	C 2131	N 794	O 1320	Р 218	0	0

• Molecule 15 is a protein called TATA-box-binding protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	О	181	Total 1422	C 925	N 243	0 248	S 6	0	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
0	241	LYS	-	expression tag	UNP P13393
0	242	HIS	-	expression tag	UNP P13393
0	243	HIS	-	expression tag	UNP P13393
0	244	HIS	-	expression tag	UNP P13393
0	245	HIS	-	expression tag	UNP P13393
0	246	HIS	-	expression tag	UNP P13393
0	247	HIS	-	expression tag	UNP P13393

• Molecule 16 is a protein called Transcription initiation factor IIF subunit alpha.

Mol	Chain	Residues		Ate	oms			AltConf	Trace
16	Q	221	Total 1871	C 1179	N 346	O 339	S 7	0	0



There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Q	-2	GLY	-	expression tag	UNP P41895
Q	-1	PRO	-	expression tag	UNP P41895
Q	0	GLY	-	expression tag	UNP P41895

• Molecule 17 is a protein called Transcription initiation factor IIF subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	R	268	Total 2230	C 1409	N 392	0 419	S 10	0	0

• Molecule 18 is a DNA chain called Template DNA (219-MER).

Mol	Chain	Residues		Α	AltConf	Trace			
18	Т	219	Total 4510	C 2144	N 844	O 1304	Р 218	0	0

• Molecule 19 is a protein called Transcription initiation factor IIA large subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	U	107	Total 885	C 559	N 147	0 176	${ m S} { m 3}$	0	0

• Molecule 20 is a protein called Transcription initiation factor IIA subunit 2.

Mol	Chain	Residues		At	oms			AltConf	Trace
20	V	104	Total 815	С 511	N 136	0 164	$\frac{S}{4}$	0	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
V	123	LYS	-	expression tag	UNP P32774
V	124	HIS	-	expression tag	UNP P32774
V	125	HIS	-	expression tag	UNP P32774
V	126	HIS	-	expression tag	UNP P32774
V	127	HIS	-	expression tag	UNP P32774
V	128	HIS	-	expression tag	UNP P32774
V	129	HIS	-	expression tag	UNP P32774

• Molecule 21 is a protein called Transcription initiation factor IIE subunit alpha.



Mol	Chain	Residues		At	oms			AltConf	Trace
21	W	304	Total 2473	C 1558	N 431	O 477	${ m S} 7$	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
W	483	ALA	-	expression tag	UNP P36100
W	484	ALA	-	expression tag	UNP P36100
W	485	ALA	-	expression tag	UNP P36100
W	486	LEU	-	expression tag	UNP P36100
W	487	GLU	-	expression tag	UNP P36100
W	488	HIS	-	expression tag	UNP P36100
W	489	HIS	-	expression tag	UNP P36100
W	490	HIS	-	expression tag	UNP P36100
W	491	HIS	-	expression tag	UNP P36100
W	492	HIS	-	expression tag	UNP P36100

• Molecule 22 is a protein called Transcription initiation factor IIE subunit beta.

Mol	Chain	Residues		At	oms			AltConf	Trace
22	Х	211	Total 1708	C 1089	N 293	O 320	S 6	0	0

• Molecule 23 is a protein called General transcription and DNA repair factor IIH helicase subunit XPD.

Mol	Chain	Residues		Α	toms			AltConf	Trace
23	0	752	Total 6091	C 3882	N 1029	0 1142	S 38	0	0

• Molecule 24 is a protein called General transcription and DNA repair factor IIH subunit TFB1.

Mol	Chain	Residues		At	\mathbf{oms}			AltConf	Trace
24	1	522	Total 4214	C 2660	N 734	O 798	S 22	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
1	-2	GLY	-	expression tag	UNP P32776
1	-1	GLY	-	expression tag	UNP P32776
1	0	SER	-	expression tag	UNP P32776



• Molecule 25 is a protein called General transcription and DNA repair factor IIH subunit TFB2.

Mol	Chain	Residues		At	oms			AltConf	Trace
25	2	452	Total 3647	C 2354	N 600	O 677	S 16	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
2	-3	GLY	-	expression tag	UNP Q02939
2	-2	PRO	-	expression tag	UNP Q02939
2	-1	GLY	-	expression tag	UNP Q02939
2	0	SER	-	expression tag	UNP Q02939

• Molecule 26 is a protein called RNA polymerase II transcription factor B subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	3	131	Total 1089	C 692	N 180	0 209	S 8	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
3	-2	GLY	-	expression tag	UNP Q03290
3	-1	PRO	-	expression tag	UNP Q03290
3	0	HIS	-	expression tag	UNP Q03290

• Molecule 27 is a protein called General transcription and DNA repair factor IIH subunit TFB4.

Mol	Chain	Residues		At	oms			AltConf	Trace
27	4	306	Total 2372	C 1511	N 396	0 451	S 14	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
4	-2	SER	-	expression tag	UNP Q12004
4	-1	ASN	-	expression tag	UNP Q12004
4	0	ALA	-	expression tag	UNP Q12004

• Molecule 28 is a protein called General transcription and DNA repair factor IIH subunit TFB5.



Mol	Chain	Residues	Atoms					AltConf	Trace
28	5	65	Total 514	C 326	N 90	O 95	${ m S} { m 3}$	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
5	-3	GLY	-	expression tag	UNP Q3E7C1
5	-2	PRO	-	expression tag	UNP Q3E7C1
5	-1	GLY	-	expression tag	UNP Q3E7C1
5	0	SER	-	expression tag	UNP Q3E7C1

• Molecule 29 is a protein called General transcription and DNA repair factor IIH subunit SSL1.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	6	383	Total 3019	C 1915	N 523	O 552	S 29	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
6	-2	GLY	-	expression tag	UNP Q04673
6	-1	GLY	-	expression tag	UNP Q04673
6	0	SER	-	expression tag	UNP Q04673

• Molecule 30 is a protein called General transcription and DNA repair factor IIH helicase subunit XPB.

Mol	Chain	Residues	Atoms				AltConf	Trace	
30	7	616	Total 4961	C 3158	N 861	0 915	S 27	0	0

• Molecule 31 is a protein called Histone H3.2.

Mol	Chain	Residues	Atoms			AltConf	Trace		
31	a	97	Total 801	C 506	N 155	0 138	$\begin{array}{c} \mathrm{S} \\ \mathrm{2} \end{array}$	0	0
31	е	98	Total 810	C 512	N 157	0 139	${ m S} { m 2}$	0	0

There are 4 discrepancies between the modelled and reference sequences:



Chain	Residue	Modelled	Actual	Comment	Reference
a	102	ALA	GLY	conflict	UNP P84233
a	110	ALA	CYS	engineered mutation	UNP P84233
e	102	ALA	GLY	conflict	UNP P84233
e	110	ALA	CYS	engineered mutation	UNP P84233

• Molecule 32 is a protein called Histone H4.

Mol	Chain	Residues	Atoms				AltConf	Trace	
20	h	80	Total	С	Ν	0	S	0	0
32	3 <u>2</u> D	02	653	412	127	113	1	0	0
20	f	80	Total	С	Ν	0	S	0	0
32 I	80	638	401	125	111	1	0	0	

• Molecule 33 is a protein called Histone H2A.

Mol	Chain	Residues	Atoms			AltConf	Trace	
33	С	109	Total 843	C 531	N 167	0 145	0	0
33	g	106	Total 818	C 516	N 160	O 142	0	0

• Molecule 34 is a protein called Histone H2B 1.1.

Mol	Chain	Residues	Atoms				AltConf	Trace	
34	d	97	Total 767	C 481	N 142	0 142	$\frac{S}{2}$	0	0
34	h	95	Total 745	C 469	N 134	0 140	S 2	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
d	29	THR	SER	conflict	UNP P02281
h	29	THR	SER	conflict	UNP P02281

• Molecule 35 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	AltConf
35	А	2	Total Zn 2 2	0
35	В	1	Total Zn 1 1	0



Continued from previous page...

Mol	Chain	Residues	Atoms	AltConf
35	С	1	Total Zn 1 1	0
35	Ι	2	Total Zn 2 2	0
35	J	1	Total Zn 1 1	0
35	L	1	Total Zn 1 1	0
35	М	1	Total Zn 1 1	0
35	W	1	Total Zn 1 1	0
35	3	2	Total Zn 2 2	0
35	4	1	Total Zn 1 1	0
35	6	4	Total Zn 4 4	0

• Molecule 36 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	AltConf
36	А	1	Total Mg 1 1	0

• Molecule 37 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe_4S_4).





Mol	Chain	Residues	Atoms	AltConf
37	0	1	Total Fe S 8 4 4	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 II subunit RPB1



• Molecule 2: DNA-directed RNA polymerase II subunit RPB2





THR GLIX GLIX GLIX GLIX ASP ASP ASS ASS ASS ASS CLIX TTRR ASS SER ASS SER ASS ASS TTRR TTRR TTRR TTRR

• Molecule 4: DNA-directed RNA polymerase II subunit RPB4

Chain D:	75%	24%
MET ASN VAL SER THR SER FHE PHE GLN	THR Rd 1 H 4 H 4 H 4 CLN CLN CLN CLN CLN CLN CLN CLN CLN CLN	GLU GLU ASP ASP ASN ASN ASP ASP ASP PHE MET HTS SER
L220 TYR		
• Molecule 5	5: DNA-directed RNA polymerases I, II, and II	I subunit RPABC1
Chain E:	99%	·
MET D2 N5 M215 M215		
• Molecule 6	5: DNA-directed RNA polymerases I, II, and II	I subunit RPABC2
Chain F:	75% •	24%
MET SER ASP TYR GLU GLU ALA PHE ASN	ASP ASP ALSP ALSP F35 F35 F35 F35 ALSP ALSA ALA ALA ALA ALA ALA ALA ALA ALA AL	D77 L.165
• Molecule 7	7: DNA-directed RNA polymerase II subunit R	PB7
Chain G:	97%	
M G124 H171 H15 H15 H15 H15	HIS	
• Molecule 8	8: DNA-directed RNA polymerases I, II, and II	I subunit RPABC3
Chain H:	95%	
MET S2 P69 ASN ASN SER SER	R75 R77 R130 R146	
• Molecule 9	: DNA-directed RNA polymerase II subunit R	PB9
Chain I:	95%	5%
MET T2 K117 ARG GLN PHE SER		

• Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5



Chain J: 97%	••
A1 A5P A5P	
• Molecule 11: DNA-directed RNA polymerase II subunit RPB11	
Chain K: 95%	•••
M15 D5 A116 ASP ALA PHE	
• Molecule 12: DNA-directed RNA polymerases I, II, and III subunit R	PABC4
Chain L: 64% 36%	_
MET SER ARG ALV GLY GLY ALA ALA ALA ALA ALA ALA ALA ALA ALA A	
• Molecule 13: Transcription initiation factor IIB	
Chain M: 88% 12	20%
MET MET THR ARG GLU SERR ARG ARG ARG ARG ARG ARG ARG ARG ARG A	
• Molecule 14: Non-template DNA (219-MER)	
Chain N: 100%	_
There are no outlier residues recorded for this chain.	
• Molecule 15: TATA-box-binding protein	
Chain O: 73% 27%	_
MET ASP ALA ASP GLU GLU GLU CYS CLU CYS CLU VAL LYS CLU ASN ASN ASN ASN ASN ASN ASN ASN ASN ASN	JER GLU LYS ASP THR SER ALA ALA
M2400 HISSIN SIN SIN SIN SIN SIN SIN SIN SIN SI	
\bullet Molecule 16: Transcription initiation factor IIF subunit alpha	
Chain Q: 30% 70%	_
GLY MET ARG ARG ARG ARG ARG ARG ARG ARG ARG ARG	G NU G TYR ALA GLU GLU GLU GLU



Chain U:

MET

LYS MET LEU I EU	GLN	GLY VAL GT II	ALA ASP	ALA GLY	ARG SER	ASN VAL	LYS VAL	LYS ASP	GLU D94	K168	LYS AI,A	GLU	GLU	SER	PRO	SER	GLY GLY	MET ASN	LYS SER	GLY	VAL	SER	ASN ASN	THR VAL	LYS	GLY	GLN	THR PRO TUB	VAL	
ASP SER VAL THB	LYS ASP	ASN THR AI A	ASN GLY	VAL ASN	SER SER	ILE PRO	THR VAL	THR GLY	SER	VAL PRO	PRO ALA	SER	THR	VAL	ALA	GLU	SER	GLY LEU	SER	GLY	THR	SER	ALA ASN	GLY LEU	ASP	ASN	SER	THR ALA ASM	LEU	
ALA ASN GLY	PRO	VAL THR r vs	GLU	ASP ALA	GLY PRO	ALA GLU	ASP PRO	THR LYS	VAL GLY	MET VAL	LYS TYR	ASP	TAS TAS	VAL	ASN	PRO	GLU	GLU GLU	GLY THR	MET	PRO	LEU ALA	ASP VAL	ALA PRO	ASP	ULL GLY	ARG	ALA LYS Abr	GLY	
ASN L325 K11	D429	N430 T431	G432 T433	RA AG	A447	V448 A449	D450	V IV	MET	GLU	ASP ASP	ARG ASP	ASP ASN	SER GLU	VAL GLU	LEU	TYR	GLU	GLU	ALA ASP	ASP	GLU GLU	PRO TEO	ILE	ASP GLY	ASN GLU	GLN	ASN LYS	GLU SER	GLU GLU
ARG ILE LYS I VS	GLU MET	GLN GLN	ASN	GLY	LEU ARG	ASP GLU	GLU ALA	PRO SER	GLU ASN	GLU GLU	ASP	LEU	GLY	LYS	TLE	GLU	ASP GLY	GLU ARG	ILE	LYS	LEU	GLN	THR GLU	LEU ALA	ALA	TYR	NER S	GLU GLU	GLU	
ILE ASN PRO TVP	LEU SER	GLU SER ASD	1LE GLU	ASN LYS	GLU ASN	GLU SER	PRO VAL	LYS	GLU GLU	ASP SER	ASP THR	LEU	TYS	TYS And	SER	PRO	LYS	GLN	LYS	ALA	ASN	ALA HIS	VAL HIS	GLU	PRO THR	TEU	VAL	LYS SER	LYS	
ASN CYS VAL TIF	TLE	CLY GLY ASP	SY1 SY1	LEU	LYS SER	PHE PRO	GLY	GLU TRP	ASN PRO	GLN THR	THR	ALA	ASP	SER	ASN	SER	ASN THR	VAL PRO	SER	ILE	GLN	GLU GLU	GLY LEU	ASN SER	THR	ALA	ARG	GLU GLU	PRO	
ALA PRO THR	THR GLU	LYS ASP TI F	TLE	ALA ILE	GLY ASP	GLY LYS	VAL ASN	ILE LYS	GLU	GLY LYS	PHE TLE	ARG	LYS	PRO PRO	ALA	ASN	LYS	LEU MET	PHE AL.A	ILE	LYS	LYS LEU	CYS ARG	LYS VAL	GLY	ASP	MET	GLU LEU	LYS	
GLU																														
• Mc	oleci	ıle	17:	Tra	ans	scri	ipt	ion	in	iti	ati	on	fa	ctc	or I	IF	su	ıbu	mi	t ł	oet	a								
Chai	n R	: -							6	7%												33	3%				-			
MET SER SER	SER	GLY ALA PRO	ALA LEU	SER ASN	ASN SER	THR ASN	SER VAL	ALA LYS	GLU	SER GLY	ASN	SER	ASP	TYR	SER	GLU	GLU E38	Q144	GLN ARG	LYS	GLU	GLU	ALA ASP	PR0 GLU	LYS	ARG	ALA	TYR LEU 1 VS	LYS	
GLN GLU ARG	GLU	LEU LYS T VS	GLN	GLN	GLN	ARG ARG	ASN ASN	ARG LYS	LYS PHE	ASN HIS	ARG VAL	MET	ASP	ASP	чьт R198	M270	L.358	GLU	GLU	ARG	ALA	LEU	GLU GLU	LEU ALA	ASP	NTD	GLY	ALA	ASP	
ASN ALA GLN	ASP ALA	GLU ALA ASD	CEU GEU	ASP GLU	GLU	MET GLU	ASP VAL	VAL																						
• Mc	olect	ıle	18:	Te	mp	olat	te İ	DN	A	(2	19-	M	ER	t)																
Chai	n T	: -											1	00%													-			
Ther	e ai	e n	0.0	utli	er	res	sid	ues	re	eco	rde	ed	for	$^{\mathrm{th}}$	is	cha	ain	•												



37%

• Molecule 19: Transcription initiation factor IIA large subunit

62%



• Molecule 24: General transcription and DNA repair factor IIH subunit TFB1



Chain 1:	80%		19%
GLY GLY SER M1 S2 S43 S43 S43 S43 S43 K47	L48 Q49 P55 P55 A55 A55 A55 C66 A5P C1U SER LVS SER LVS SER LVS SER LVS SER ASP ASP ASP ASP ASP ASP ASP ASP ASP ASS ASS	GLU GLU GLU CLY ASN VAL VAL VAL PRO P84 P84 P84 P84 P86 P85	D98 E119 ARC ARC ARC ARC ALU SELU ALA ALA ALA ALA ALA ALA ALA ALA PRO
MET SER SER SER SER VAL THR THR THR THR PRO	THR THR HIS THR HIS THR THR PRO CLY CLY ASN TLE TLEU TLEU TLEU TLEU TLEU ASN ASN ASN ASN ASN ASN ASN ASN ASN ASN	N232 V240 ALA SER SER CLU CLU N245 R306 R306 R313	7393 ARG ARG ARG ARN LEU GLN GLN GLN SER SER ILE ILE
THR ASN ASP GLU GLU ASP ASP ASP M413 1446	ASP ASP ALA ALA ALA ASP SER SER ALA ASP ASP ASP ASP ASN ASA ASN ASN ASN ASN ASN ASN ASN ASN	LEU GLY SGLY SGLY SGL VAL ASP ASP ASN GLA GLA GLA	NG 39 ASN ASN ASN
• Molecule 25:	General transcription and DNA	A repair factor IIH	subunit TFB2
Chain 2:	87%		13%
GLY PRO GLY SER MET ASP ASP ASP 1282	LYS LYS TILE ARG ARG SER ALA ALA ALA ALA ALA ALA ALA GLU GLU GLU SER ASN ASN ASN ASN ASN ASN ASN ASN ASN ASN	ASN ASP ASP ASP ALA ASP CITY SER SER THR THR THR THR TTE	THR ASP ASP ASP ASP LYS LYS CLYS CLYS CLYS ASN ASN
GLN ASP ASP B421 ₽421 P424 N425 R507	LYS LYS LYS GLM GLM		
• Molecule 26:	RNA polymerase II transcription	on factor B subuni	t 3
Chain 3:	40%	60%	
GLY PRO HIS MET MET LEU MET ASP GLU GLU GLU	N9 L139 L139 LVS LVS LVS LLU LEU ASN ASN ASN ASN ASN ASN ASN ASN ASN ASN	ALM ARG GLN CLYS PHE CLU CLU CLU CLU CLU CLU CLU CLU CLU CLU	LEU LEU GLU GLN ILE GLU GLU GLU MET
ASN LYS CLU CLYS GLU THR THR THR LYS CLU TLE VAL	ARG SER THR THR THR THR ASP ASP ASP ASP ASP ASP ASP ASP ASP ASP	LYS SER SER ALA ARG ARG LYS LYS LLY GLU GLU GLU ARG ARG	LEU LEV ASW ASW ASW PRO PRO PRO ASW SER ASW VAL ASW
VAL GLN ASN SER ARG LEU LEU LYS ALA ALA VAL	PHE THR PRO PHE GLY GLY ASP ASP ARG ARG ARG ARG CLU CLU CVS SER VAL CVS ARG ARG ARG ARG ARG ARG ARG ARG ARG ARG	PRO PHE LIYE LIYE ASP ASP CLU HIS ANG CLU PHE TLE	SER GLY ASN THR ASN TYR ALA ALA ALA ALA ANG VAL
LEU THR GLU GLU MET MET CLEU CLY CYS VAL	LLE GLU GLU LLEU LLEU		
• Molecule 27:	General transcription and DNA	A repair factor IIH	subunit TFB4
Chain 4:	89%		10%
SER ASN MET MET ASP ALA ALA TLE SER PRO PRO	PHE LYS HIS ALA ALA ALA ALA ARG STR CLW CLW CLW CLW CLW CLW THR THR THR THR THR THR THR THR THR THR	LYS LYS LYS LYS LYS VAL THR LYS PRO	
• Molecule 28:	General transcription and DNA	A repair factor IIH	subunit TFB5
Chain 5:	83%		14%
GLY PRO CLY SER MET M2 R5 R5 L21 L21	MG6 ASP GLU GLU ASN GLN GLN		



• Molecule 29: General transcription and DNA repair factor IIH subunit SSL1

	Chain 6:	82%	17%	
--	----------	-----	-----	--

																						e	ი		
LYS	LYS	LYS	ARG	LEU	SER	ASN	ARG	ASN	LEU	GLN	GLY	SER	ASN	G72	V89	ASP	ASP	GLU	GLY	ASP	M95	L41	V45	E	THR

• Molecule 30: General transcription and DNA repair factor IIH helicase subunit XPB

Chain 7:	73%	27%
MET THR ASP VAL GLV GLV GLV GLV GLV SER SER CLYS SER CLYS SER CLYS SER CLYS SER CLYS SER CLYS SER CLYS CLYS CLYS CLYS CLYS CLYS CLYS CLYS	MET GLY GLY PHE PHE PHE SER ASP ASP ASP ASP ASP ASP ASP ASP ASP ASP	ASP ASP ASP ASP ASP ASP ASP CLU CLU CLV CLV CLV CLV CLV CLV CLV CLV CLV CLV
THR CLEU LEU LEU LYS LYS PRO LYS LYS SER LYS SER ARG ARG ARG ARG ARG ARG ARG ARG ARG AR	ASP SER MET MET ASP ASP ASP ASP ASP ASP ASS ASP ASP ASP	P.253 P.253 LEU ARG ARF ARF ARF ARF CLN CLN CLN CLN CLN CLN CLN CLN CLN CLN
GL N GL N GL N GL N GL N GL N GL A ALA ALA ASN ASN ASN ASN ASN	VAL VAL ALJ VAL ALA ALA ALA ALA ALA ALA ALV GLV GLV GLU GLU GLU ASP ASP ASP ASP ASP ASP ASP ASP ASP ASP	ALA V313 P358 8359 8359 8359 8363 8363 8363 8363 8363 8363 8363 836
GLY ILE VAL VAL VAL ASP ASP ASP ASP ASP ASP ASP ASS ASS ASS	HIS HIS ARG SPHE SPHE SPHE SPHE ARG ALA ARG CLV CLV SER ARG CLV CLV CLV SER ARG CLV SER ALA ARG CLV SER ARG CLV SCA SCA SCA SCA SCA SCA SCA SCA SCA SCA	MET ALA ALA ALA ALA ALA MET ALU CLU CLU CLU CLU CLU CLU CLU CLU MIS30
Y837 TYR LVS LVS LVS LVS LVS LVS		
• Molecule 31: Histone I	H3.2	
Chain a:	72%	28%
ALA ARG THR THR THR GLM GLM ARG CLY CLYS CLY CLYS CLY ALA ARG CLY CLYS CLY CLYS CLY CLYS CLY CLYS CLY CLYS CLY CLYS CLUA	GLN THR ALA THR ALA ALA ALA ALA ALA ALA ALA ALA ALA AL	
• Molecule 31: Histone I	H3.2	
Chain e:	73%	27%
ALA ARG LYS LYS CYS GJN ARG CJN CYS SER THR THR THR CYS CJN ARG ALA ALA ARG ALA	LEU LEU ALEU LYS LYS ALA ALA ALA ALA ALA ALA ALA ALA CLY VAL LYS K37 K37 K37 K37 K37 K37	
• Molecule 32: Histone I	H4	
Chain b:	79%	• 20%
SER GLY ARG CLY CLY CLY CLY CLY CLY CLY CLY CLY CLY	ARG 1275 7021 7022 7102	
• Molecule 32: Histone I	H4	



Chain f:	78%	22%
SER GLY GLY GLY CLYS CLY CLY CLY CLY CLY CLY CLY CLY CLY CLY	GLY ALA LYS ARG ARG ARG LYS VAL LVAL LEU R23 G102	
• Molecule 33: Hi	stone H2A	
Chain c:	84%	16%
SER ARG ARG CLY CLY CLY CLY CLY CLY CLY THR THI	V114 L115 L116 P117 K116 K118 K118 K118 SER SER SER SER LVS SER LVS SER LVS SER LVS	
• Molecule 33: Hi	stone H2A	
Chain g:	81%	• 18%
SER GLY ARG CLY CLY CLY CLY CLY GLY CLY ALA ALA ALA	K116 L51 L51 L178 K116 C1U C178 SER SER SER SER SER L17S SER L17S	
• Molecule 34: Hi	stone H2B 1.1	
Chain d:	77%	• 22%
PRO GLU PRO ALA LYS SER ALA ALA PRO PRO LYS LYS GLY	SER LYS LYS VAL VAL THR THR GLN GLN GLY LYS LYS LYS K113 K113 K122	
• Molecule 34: Hi	stone H2B 1.1	
Chain h:	76%	24%
PRO GLU PRO LYS SER ALA PRO PRO PRO LYS CLY GLY	SER LLYS ALA ALA THR THR CLYS CLYS CLYS ARG ARG K122 K122	



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	82942	Depositor
Resolution determination method	FSC 0.5 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE	Depositor
	CORRECTION	
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{\AA}^2)$	41	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	0.351	Depositor
Minimum map value	-0.171	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.007	Depositor
Recommended contour level	0.007	Depositor
Map size (Å)	377.99997, 377.99997, 377.99997	wwPDB
Map dimensions	360, 360, 360	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.05, 1.05, 1.05	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, SF4

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	Bond	lengths	Bond angles							
WIOI	Ullalli	RMSZ	# Z > 5	RMSZ	# Z > 5						
1	А	0.31	0/11422	0.54	0/15445						
2	В	0.34	0/9589	0.56	1/12934~(0.0%)						
3	С	0.34	0/2130	0.54	0/2887						
4	D	0.25	0/1351	0.55	0/1811						
5	Е	0.30	0/1788	0.57	1/2406~(0.0%)						
6	F	0.33	0/995	0.63	1/1340~(0.1%)						
7	G	0.29	0/1367	0.56	0/1844						
8	Н	0.34	0/1139	0.58	0/1544						
9	Ι	0.31	0/962	0.57	0/1295						
10	J	0.37	0/578	0.65	0/775						
11	Κ	0.34	0/942	0.62	0/1272						
12	L	0.37	0/361	0.72	0/478						
13	М	0.28	0/2408	0.54	1/3241~(0.0%)						
14	Ν	0.53	0/4999	0.96	0/7709						
15	0	0.29	0/1449	0.58	0/1952						
16	Q	0.26	0/1907	0.58	0/2556						
17	R	0.25	0/2270	0.56	1/3052~(0.0%)						
18	Т	0.53	0/5067	0.94	0/7824						
19	U	0.25	0/898	0.56	1/1212~(0.1%)						
20	V	0.25	0/822	0.55	0/1109						
21	W	0.25	0/2513	0.53	0/3388						
22	Х	0.24	0/1739	0.53	0/2339						
23	0	0.28	0/6209	0.57	1/8384~(0.0%)						
24	1	0.25	0/4277	0.51	0/5755						
25	2	0.27	0/3717	0.57	0/5028						
26	3	0.30	0/1109	0.55	0/1492						
27	4	0.30	0/2411	0.59	1/3260~(0.0%)						
28	5	0.27	0/520	0.66	1/701~(0.1%)						
29	6	0.29	0/3082	0.57	1/4165~(0.0%)						
30	7	0.27	0/5067	0.55	1/6853~(0.0%)						
31	a	0.26	0/813	0.58	0/1091						
31	е	0.28	0/822	0.60	0/1103						



Mal Chain		Bond	Bond lengths		ond angles
	Unam	RMSZ	# Z > 5	RMSZ	# Z > 5
32	b	0.29	0/660	0.63	0/883
32	f	0.26	0/645	0.58	0/862
33	с	0.27	0/853	0.55	0/1149
33	g	0.27	0/828	0.62	1/1117~(0.1%)
34	d	0.26	0/778	0.56	0/1043
34	h	0.31	0/756	0.60	0/1015
All	All	0.33	0/89243	0.62	12/122314~(0.0%)

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
11	K	0	1

There are no bond length outliers.

All (1	12)	bond	angle	outliers	are	listed	below:
--------	-----	------	-------	----------	-----	--------	--------

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
33	g	51	LEU	CA-CB-CG	6.54	130.34	115.30
2	В	1125	ASP	CB-CG-OD1	5.97	123.67	118.30
27	4	293	LEU	CA-CB-CG	5.94	128.97	115.30
17	R	270	MET	CA-CB-CG	5.85	123.24	113.30
13	М	250	MET	CB-CG-SD	-5.80	95.00	112.40
23	0	547	MET	CA-CB-CG	5.65	122.91	113.30
29	6	413	LEU	CA-CB-CG	5.63	128.25	115.30
28	5	21	LEU	CA-CB-CG	5.38	127.68	115.30
19	U	38	LEU	CA-CB-CG	5.24	127.35	115.30
6	F	77	ASP	CB-CG-OD2	5.22	123.00	118.30
5	Е	2	ASP	CB-CG-OD2	5.09	122.88	118.30
30	7	765	LEU	CA-CB-CG	5.02	126.84	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
11	Κ	5	ASP	Peptide



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	1418/1733~(82%)	1379 (97%)	39~(3%)	0	100	100
2	В	1168/1224~(95%)	1121 (96%)	47 (4%)	0	100	100
3	С	264/347~(76%)	254 (96%)	10 (4%)	0	100	100
4	D	163/221~(74%)	163 (100%)	0	0	100	100
5	Е	212/215~(99%)	204 (96%)	8 (4%)	0	100	100
6	F	114/155~(74%)	109 (96%)	5 (4%)	0	100	100
7	G	169/177~(96%)	165 (98%)	4 (2%)	0	100	100
8	Н	136/146~(93%)	134 (98%)	2 (2%)	0	100	100
9	Ι	114/122~(93%)	109 (96%)	5 (4%)	0	100	100
10	J	67/70~(96%)	65~(97%)	2 (3%)	0	100	100
11	K	113/120~(94%)	112 (99%)	1 (1%)	0	100	100
12	L	43/70~(61%)	40 (93%)	3 (7%)	0	100	100
13	М	306/352~(87%)	298 (97%)	8 (3%)	0	100	100
15	Ο	179/247~(72%)	175 (98%)	4 (2%)	0	100	100
16	Q	215/738~(29%)	208 (97%)	7 (3%)	0	100	100
17	R	264/400~(66%)	258~(98%)	6 (2%)	0	100	100
19	U	101/171~(59%)	98~(97%)	3 (3%)	0	100	100
20	V	100/129~(78%)	99~(99%)	1 (1%)	0	100	100
21	W	296/492~(60%)	291 (98%)	5 (2%)	0	100	100
22	Х	207/328~(63%)	204 (99%)	3 (1%)	0	100	100
23	0	750/778~(96%)	730 (97%)	20 (3%)	0	100	100



Mol	Chain	Analysed	Favoured	Allowed	Outliers	Perce	ntiles
24	1	508/645~(79%)	501 (99%)	7 (1%)	0	100	100
25	2	448/517~(87%)	441 (98%)	7 (2%)	0	100	100
26	3	129/324~(40%)	122~(95%)	7 (5%)	0	100	100
27	4	302/341~(89%)	290 (96%)	12 (4%)	0	100	100
28	5	63/76~(83%)	62~(98%)	1 (2%)	0	100	100
29	6	379/464~(82%)	363~(96%)	16 (4%)	0	100	100
30	7	610/843~(72%)	590~(97%)	20 (3%)	0	100	100
31	a	95/135~(70%)	93~(98%)	2(2%)	0	100	100
31	е	96/135~(71%)	96 (100%)	0	0	100	100
32	b	80/102~(78%)	79~(99%)	1 (1%)	0	100	100
32	f	78/102~(76%)	77~(99%)	1 (1%)	0	100	100
33	с	107/129~(83%)	105~(98%)	2(2%)	0	100	100
33	g	104/129~(81%)	100 (96%)	4 (4%)	0	100	100
34	d	95/125~(76%)	94 (99%)	1 (1%)	0	100	100
34	h	93/125 (74%)	89 (96%)	4 (4%)	0	100	100
All	All	9586/12427~(77%)	9318 (97%)	268 (3%)	0	100	100

Continued from previous page...

There are no Ramachandran outliers to report.

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	1248/1520~(82%)	1247 (100%)	1 (0%)	92 95
2	В	1024/1061~(96%)	1022 (100%)	2~(0%)	92 94
3	С	234/299~(78%)	233 (100%)	1 (0%)	89 91
4	D	149/200~(74%)	148 (99%)	1 (1%)	81 87
5	Ε	196/197~(100%)	195 (100%)	1 (0%)	86 89
6	F	107/137~(78%)	107~(100%)	0	100 100



a 1	e		
Continued	trom	previous	page
	5	1	1 5

Mol	Chain	Analysed	Rotameric	Outliers	Perce	ntiles
7	G	152/158~(96%)	152~(100%)	0	100	100
8	Н	123/128~(96%)	121 (98%)	2(2%)	58	74
9	Ι	110/116~(95%)	110 (100%)	0	100	100
10	J	64/65~(98%)	63~(98%)	1 (2%)	58	74
11	Κ	99/102~(97%)	99 (100%)	0	100	100
12	L	40/57~(70%)	40 (100%)	0	100	100
13	М	267/306~(87%)	267 (100%)	0	100	100
15	О	153/212~(72%)	153 (100%)	0	100	100
16	Q	204/642~(32%)	202 (99%)	2 (1%)	73	82
17	R	252/363~(69%)	252 (100%)	0	100	100
19	U	99/154~(64%)	99 (100%)	0	100	100
20	V	94/115~(82%)	92~(98%)	2 (2%)	48	66
21	W	275/436~(63%)	274 (100%)	1 (0%)	89	91
22	Х	193/295~(65%)	191 (99%)	2 (1%)	73	82
23	0	684/707~(97%)	684 (100%)	0	100	100
24	1	483/590~(82%)	480 (99%)	3 (1%)	84	88
25	2	414/470 (88%)	414 (100%)	0	100	100
26	3	125/305~(41%)	125 (100%)	0	100	100
27	4	271/302~(90%)	271 (100%)	0	100	100
28	5	59/68~(87%)	58 (98%)	1 (2%)	56	72
29	6	346/419~(83%)	346 (100%)	0	100	100
30	7	548/737 (74%)	547 (100%)	1 (0%)	92	94
31	a	84/109~(77%)	84 (100%)	0	100	100
31	е	85/109 (78%)	85 (100%)	0	100	100
32	b	67/78~(86%)	66~(98%)	1 (2%)	60	75
32	f	65/78~(83%)	65 (100%)	0	100	100
33	с	86/101 (85%)	86 (100%)	0	100	100
33	g	84/101 (83%)	83 (99%)	1 (1%)	67	79
34	d	83/105 (79%)	82 (99%)	1 (1%)	67	79
34	h	81/105 (77%)	81 (100%)	0	100	100
All	All	8648/10947 (79%)	8624 (100%)	24 (0%)	90	92



Mol	Chain	Res	Type
1	А	175	ARG
2	В	476	ARG
2	В	1183	LYS
3	С	84	ARG
4	D	14	ARG
5	Е	5	ASN
8	Н	77	ARG
8	Н	130	ARG
10	J	42	LYS
16	Q	29	ARG
16	Q	411	LYS
20	V	12	ARG
20	V	112	ARG
21	W	115	LYS
22	Х	174	LYS
22	Х	276	ARG
24	1	232	ASN
24	1	306	ARG
24	1	313	ARG
28	5	5	ARG
30	7	363	ARG
32	b	92	ARG
34	d	113	LYS
33	g	36	LYS

All (24) residues with a non-rotameric sidechain are listed below:

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (18) such sidechains are listed below:

Mol	Chain	Res	Type
1	А	171	GLN
1	А	510	GLN
1	А	736	ASN
2	В	776	GLN
2	В	1097	HIS
2	В	1141	HIS
4	D	132	GLN
6	F	63	GLN
13	М	127	GLN
13	М	303	GLN
17	R	328	HIS
20	V	17	ASN
22	Х	270	GLN



Continued from previous page...

Mol	Chain	Res	Type
25	2	498	ASN
27	4	246	GLN
30	7	634	GLN
30	7	761	GLN
34	h	60	ASN

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 19 ligands modelled in this entry, 18 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).

Mal	Type	Chain	Dog	Tink	Link Bond lengths			Bond angles		
Moi Type Chain Kes		LIIIK	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z >2		
37	SF4	0	801	23	0,12,12	-	-	-		

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
37	SF4	0	801	23	-	-	0/6/5/5

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-14929. 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.



6.2Central slices (i)

Primary map 6.2.1







Y Index: 180

Z Index: 180

6.2.2Raw map



X Index: 180

Y Index: 180

Z Index: 180

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: 178



Y Index: 74



Z Index: 246

6.3.2 Raw map



X Index: 180

Y Index: 74



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 0.007. 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 2948 nm^3 ; this corresponds to an approximate mass of 2663 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 0.152 ${\rm \AA^{-1}}$



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 0.152 ${\rm \AA^{-1}}$



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	-	6.60	-	
Author-provided FSC curve	4.14	7.08	4.23	
Unmasked-calculated*	3.31	4.07	3.39	

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.5 CUT-OFF 4.07 differs from the reported value 6.6 by more than 10 %



9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-14929 and PDB model 7ZSB. 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 0.007 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 (0.007).



9.4 Atom inclusion (i)



At the recommended contour level, 99% of all backbone atoms, 98% of all non-hydrogen atoms, are inside the map.



1.0

0.0 <0.0

9.5 Map-model fit summary (i)

The table lists the average atom inclusion at the recommended contour level (0.007) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	$\mathbf{Q} extsf{-score}$
All	0.9750	0.2770
0	0.9900	0.2590
1	0.9490	0.1230
2	0.9770	0.1580
3	0.9970	0.1650
4	0.9890	0.2200
5	0.9940	0.1310
6	0.9760	0.1950
7	0.9760	0.1060
А	0.9840	0.4720
В	0.9820	0.5070
С	0.9890	0.5300
D	0.9560	0.1680
E	0.9870	0.4170
F	0.9560	0.4040
G	0.9650	0.2790
Н	0.9910	0.4870
I	0.9860	0.3880
J	0.9910	0.5400
K	0.9790	0.5130
L	0.9830	0.4640
М	0.9750	0.3620
N	0.9920	0.1640
0	0.9910	0.1850
Q	0.9280	0.2070
R	0.9780	0.1830
Т	0.9940	0.1580
U	0.9440	0.0420
V	0.9500	0.0520
W	0.8900	0.0740
Х	0.9080	0.0550
a	0.9820	0.1370
b	0.9780	0.1540
с	0.9260	0.1880
d	0.9770	0.1650



Continued from previous page...

Chain	Atom inclusion	Q-score
е	0.9770	0.1800
f	0.9940	0.1840
g	0.9800	0.1730
h	0.9920	0.1670

