wwPDB EM Validation Summary Report ⓘ

Jun 5, 2021 – 06:38 am BST

PDB ID	:	7NVY
EMDB ID	:	EMD-12617
Title	:	RNA polymerase II pre-initiation complex with closed promoter DNA in proximal position
Authors	:	Aibara, S.; Schilbach, S.; Cramer, P.
Deposited on	:	2021-03-16
Resolution	:	7.30 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 ⓘ symbol.

The following versions of software and data (see references ⓘ) were used in the production of this report:

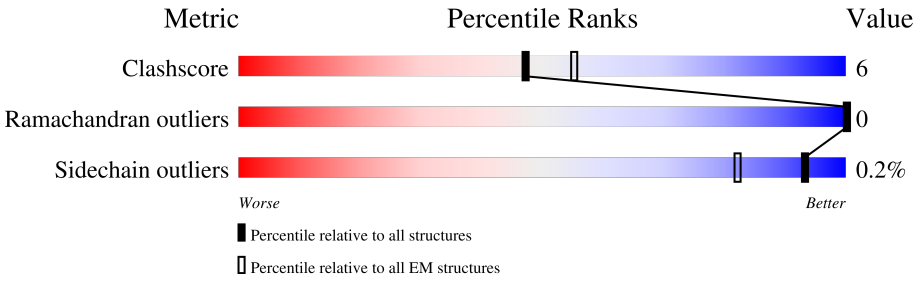
EMDB validation analysis	:	FAILED
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.19

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 7.30 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.



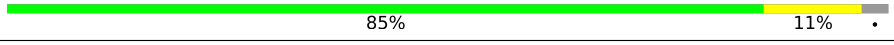
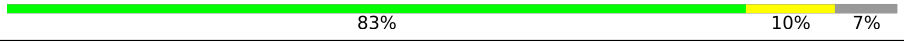
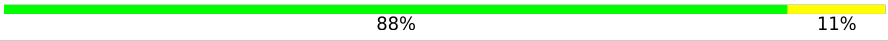
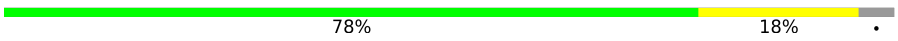
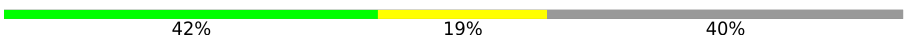
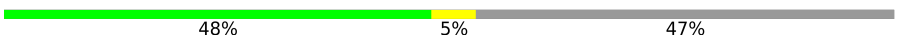
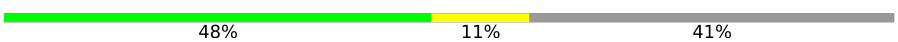
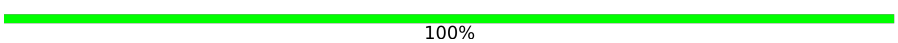
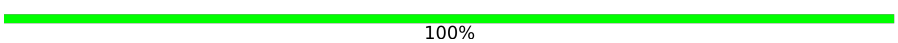
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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%

Mol	Chain	Length	Quality of chain
1	0	760	72% 21% 6%
2	1	548	40% 8% 52%
3	2	462	71% 13% 16%
4	3	309	33% 15% 52%
5	4	308	66% 19% 15%
6	5	71	82% 11% 7%
7	6	395	73% 15% 12%
8	7	782	64% 13% 23%
9	A	1970	60% 12% 28%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	B	1174	
11	C	275	
12	D	142	
13	E	210	
14	F	127	
15	G	172	
16	H	150	
17	I	125	
18	J	67	
19	K	117	
20	L	58	
21	M	316	
22	N	106	
23	O	339	
24	Q	517	
25	R	249	
26	T	106	
27	U	376	
28	V	109	
29	W	439	
30	X	291	
31	Y	19	
32	Z	8	

2 Entry composition

There are 35 unique types of molecules in this entry. The entry contains 67474 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 TFIIH basal transcription factor complex helicase XPD subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	0	714	5751	3683	999	1040	29	0	0

- Molecule 2 is a protein called General transcription factor IIH subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	1	265	2167	1382	378	395	12	0	0

- Molecule 3 is a protein called General transcription factor IIH subunit 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	2	390	3158	2050	545	551	12	0	0

- Molecule 4 is a protein called CDK-activating kinase assembly factor MAT1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	3	149	1225	763	212	240	10	0	0

- Molecule 5 is a protein called General transcription factor IIH subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	4	263	2066	1323	344	380	19	0	0

- Molecule 6 is a protein called General transcription factor IIH subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	5	66	523	337	83	100	3	0	0

- Molecule 7 is a protein called General transcription factor IIH subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	6	347	Total	C	N	O	S	0	0
			2732	1726	471	508	27		

- Molecule 8 is a protein called General transcription and DNA repair factor IIH helicase subunit XPB.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	7	605	Total	C	N	O	S	0	0
			4890	3127	848	885	30		

- Molecule 9 is a protein called RPB1.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	A	1423	Total	C	N	O	S	0	0
			11274	7092	2016	2094	72		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
10	B	1136	Total	C	N	O	S	0	0
			9076	5739	1597	1676	64		

- Molecule 11 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	C	257	Total	C	N	O	S	0	0
			2059	1294	351	408	6		

- Molecule 12 is a protein called RPOL4c domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	D	128	Total	C	N	O	S	0	0
			1050	656	178	212	4		

- Molecule 13 is a protein called DNA-directed RNA polymerase II subunit E.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	E	209	Total	C	N	O	S	0	0
			1721	1089	300	324	8		

- Molecule 14 is a protein called DNA-directed RNA polymerase II subunit F.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	F	79	Total	C	N	O	S	0	0
			636	406	108	117	5		

- Molecule 15 is a protein called DNA-directed RNA polymerase II subunit RPB7.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	G	171	Total	C	N	O	S	0	0
			1351	875	219	249	8		

- Molecule 16 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	H	148	Total	C	N	O	S	0	0
			1186	750	194	237	5		

- Molecule 17 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	I	114	Total	C	N	O	S	0	0
			928	571	166	180	11		

- Molecule 18 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	J	64	Total	C	N	O	S	0	0
			507	328	86	87	6		

- Molecule 19 is a protein called RNA_pol_L_2 domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	K	115	Total	C	N	O	S	0	0
			920	593	152	173	2		

- Molecule 20 is a protein called RNA polymerase II subunit K.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	L	44	Total	C	N	O	S	0	0
			373	231	72	64	6		

- Molecule 21 is a protein called Transcription initiation factor IIB.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	M	252	Total	C	N	O	S	0	0
			1953	1224	346	366	17		

- Molecule 22 is a DNA chain called Non-template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	N	64	Total	C	N	O	P	0	0
			1318	624	243	388	63		

- Molecule 23 is a protein called TATA-box-binding protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	O	179	Total	C	N	O	S	0	0
			1422	923	251	241	7		

- Molecule 24 is a protein called General transcription factor IIF subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	Q	138	Total	C	N	O	S	0	0
			1138	719	208	208	3		

- Molecule 25 is a protein called General transcription factor IIF subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	R	222	Total	C	N	O	S	0	0
			1788	1127	320	338	3		

- Molecule 26 is a DNA chain called Template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	T	64	Total	C	N	O	P	0	0
			1303	616	245	378	64		

- Molecule 27 is a protein called Transcription initiation factor IIA subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	U	113	Total	C	N	O	S	0	0
			930	585	152	189	4		

- Molecule 28 is a protein called Transcription initiation factor IIA subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	V	99	Total	C	N	O	S	0	0
			806	510	142	151	3		

- Molecule 29 is a protein called General transcription factor IIE subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	W	202	Total	C	N	O	S	0	0
			1659	1042	299	307	11		

- Molecule 30 is a protein called Transcription initiation factor IIE subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	X	171	Total	C	N	O	S	0	0
			1403	895	243	261	4		

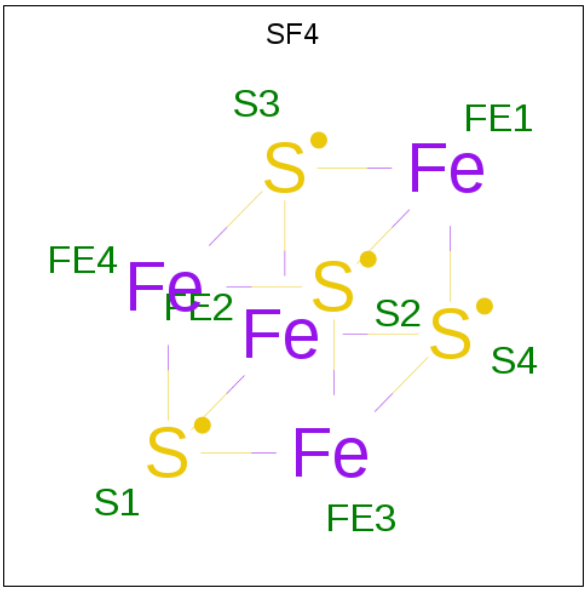
- Molecule 31 is a protein called Unassigned peptide, likely TFIIE-beta.

Mol	Chain	Residues	Atoms				AltConf	Trace
31	Y	19	Total	C	N	O	0	0
			95	57	19	19		

- Molecule 32 is a protein called Unassigned peptide, likely XPB.

Mol	Chain	Residues	Atoms				AltConf	Trace
32	Z	8	Total	C	N	O	0	0
			40	24	8	8		

- Molecule 33 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe_4S_4).



Mol	Chain	Residues	Atoms	AltConf
33	0	1	Total Fe S 8 4 4	0

- Molecule 34 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	AltConf
34	3	2	Total Zn 2 2	0
34	4	2	Total Zn 2 2	0
34	6	3	Total Zn 3 3	0
34	A	2	Total Zn 2 2	0
34	B	1	Total Zn 1 1	0
34	C	1	Total Zn 1 1	0
34	I	2	Total Zn 2 2	0
34	J	1	Total Zn 1 1	0
34	L	1	Total Zn 1 1	0
34	M	1	Total Zn 1 1	0
34	W	1	Total Zn 1 1	0

- Molecule 35 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

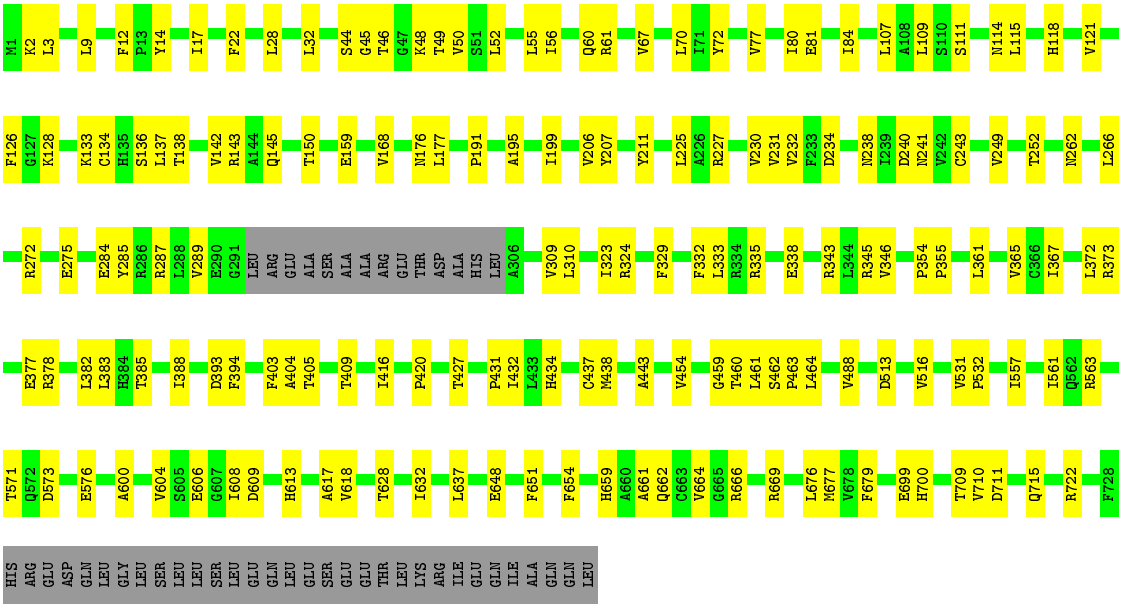
Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
35	A	1	1	1	0

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. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

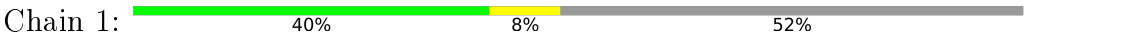
- Molecule 1: TFIIH basal transcription factor complex helicase XPD subunit

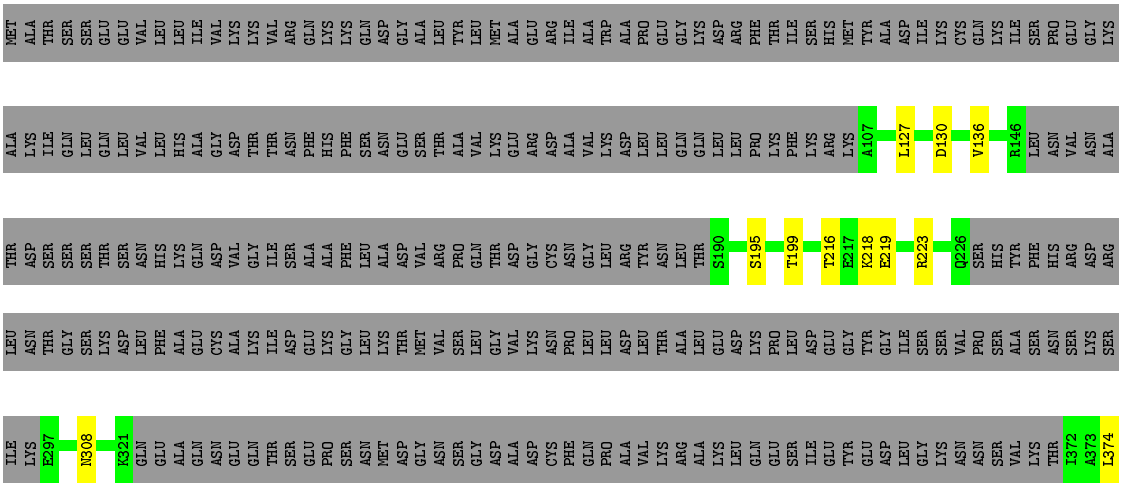
Chain 0:

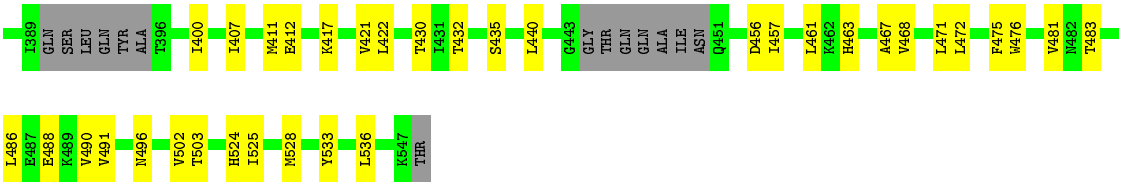


- Molecule 2: General transcription factor IIH subunit 1

Chain 1:

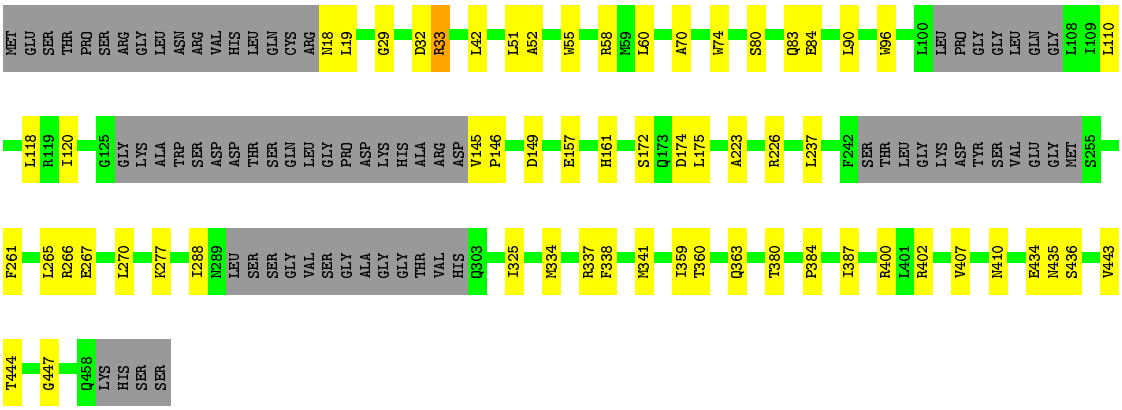






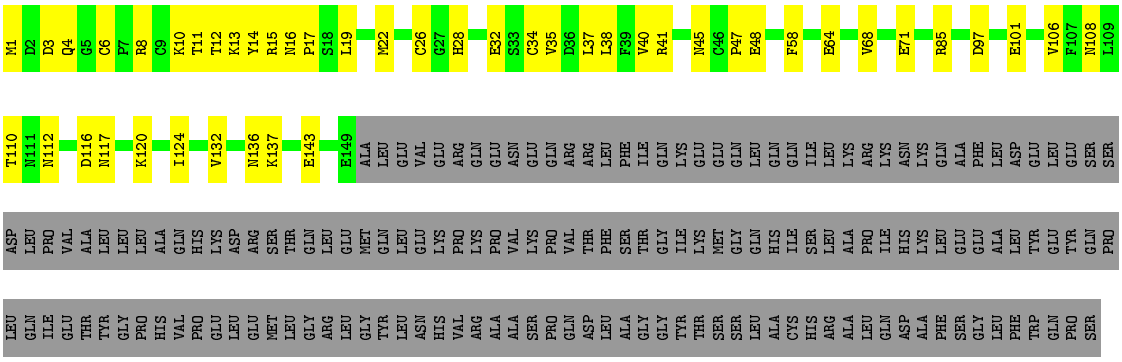
- Molecule 3: General transcription factor IIH subunit 4

Chain 2:



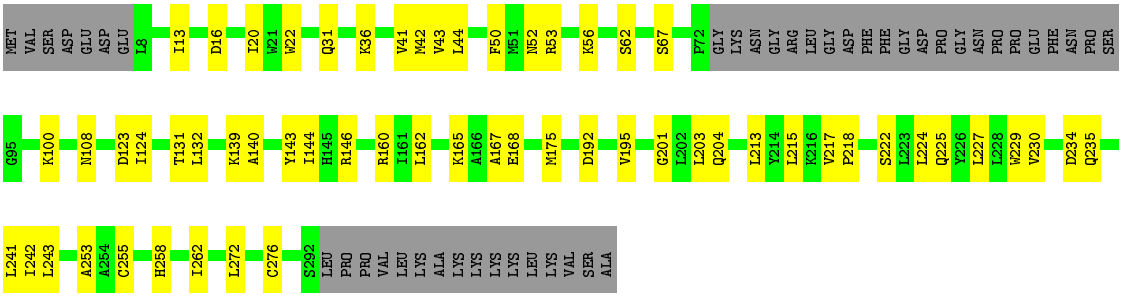
- Molecule 4: CDK-activating kinase assembly factor MAT1

Chain 3:



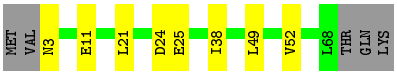
- Molecule 5: General transcription factor IIH subunit 3

Chain 4:



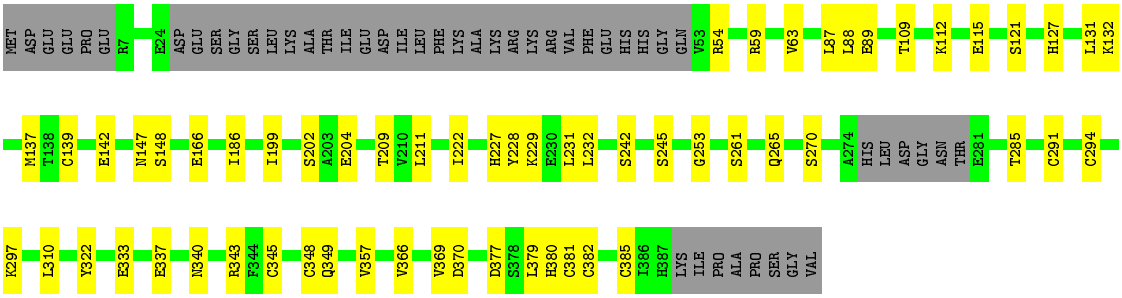
• Molecule 6: General transcription factor IIH subunit 5

Chain 5:



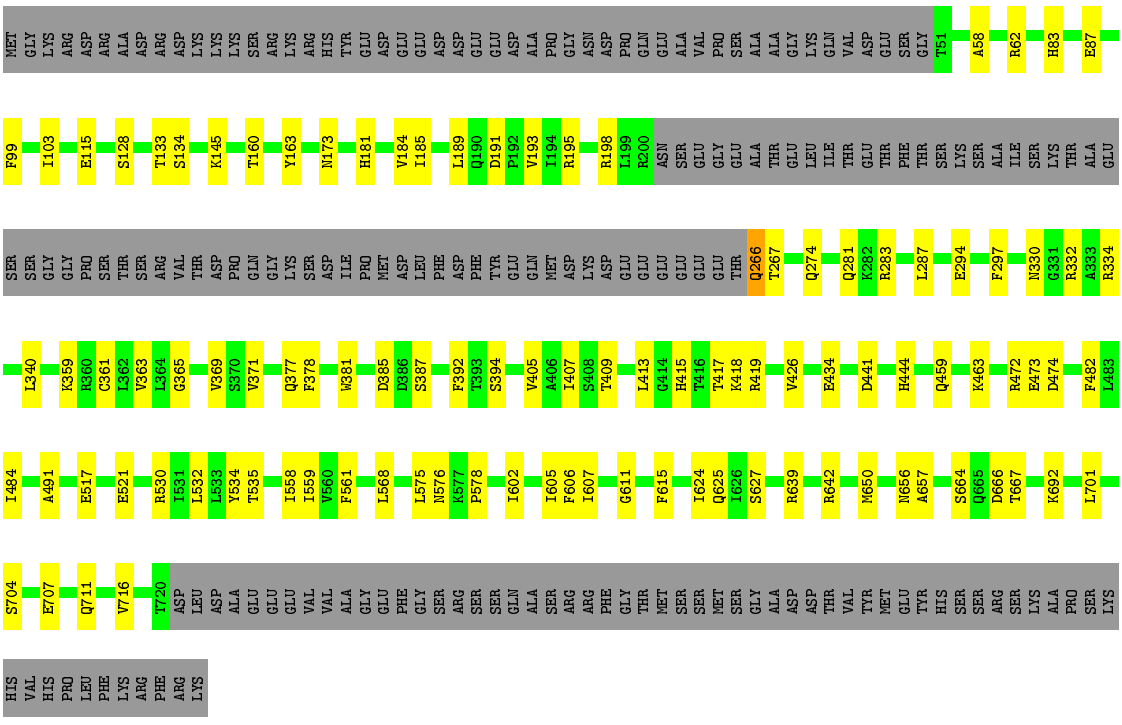
• Molecule 7: General transcription factor IIH subunit 2

Chain 6:



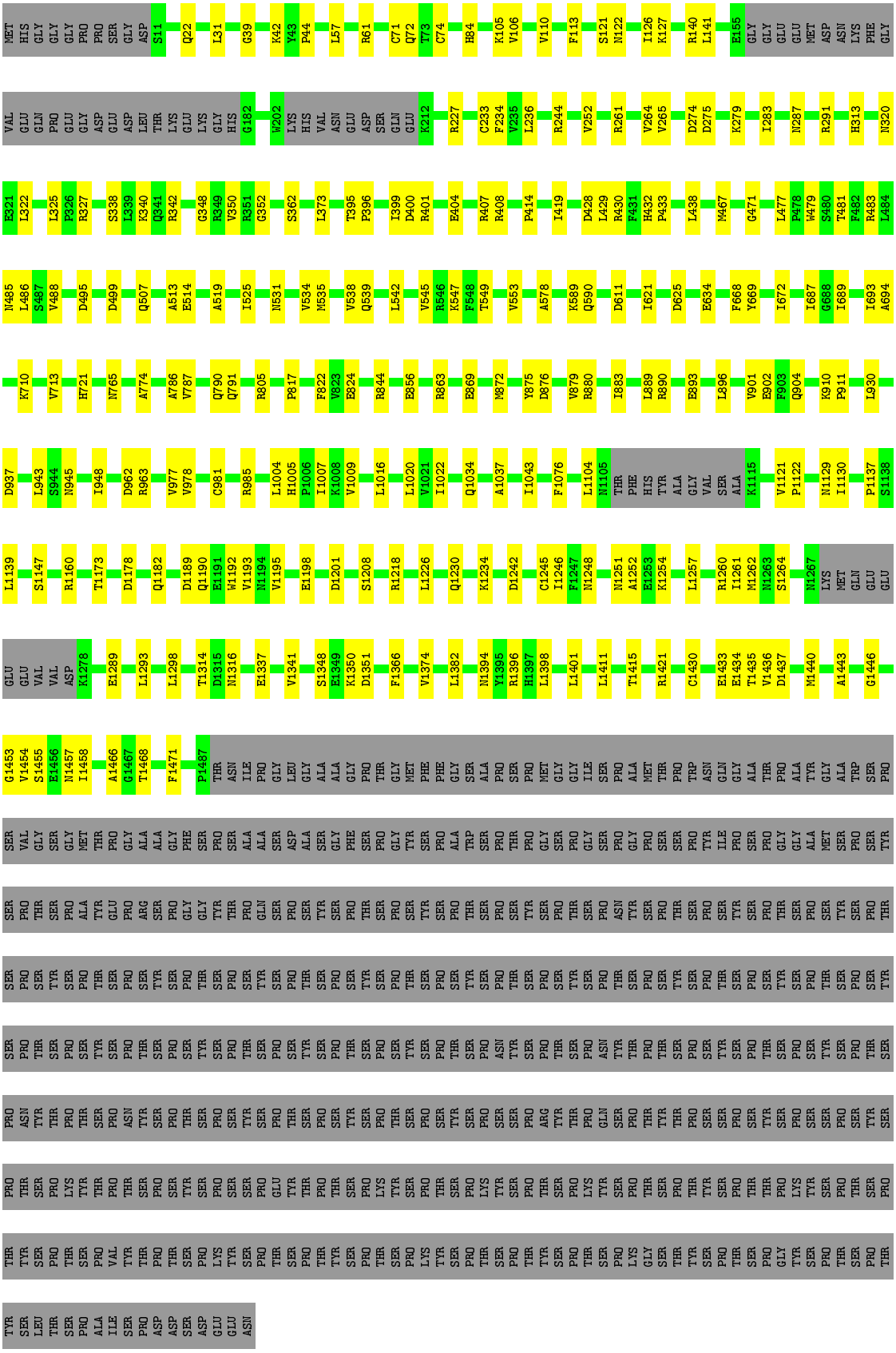
• Molecule 8: General transcription and DNA repair factor IIH helicase subunit XPB

Chain 7:



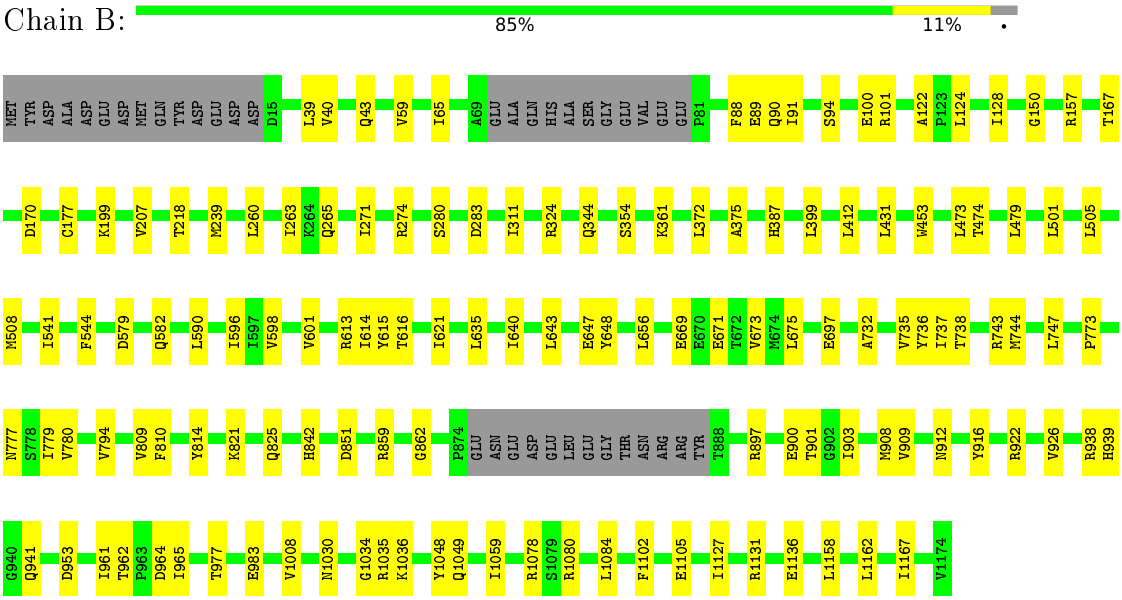
• Molecule 9: RPB1

Chain A:



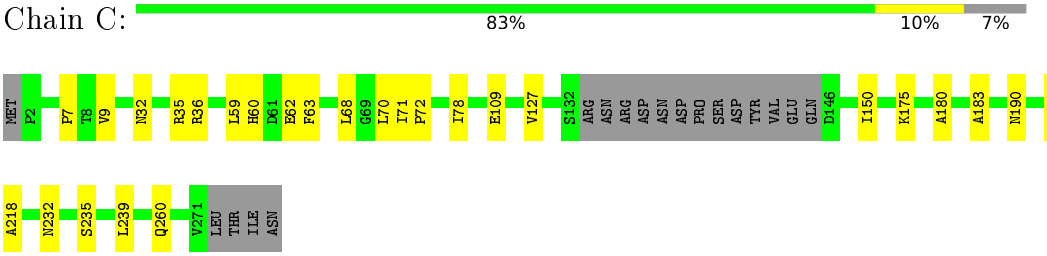
- Molecule 10: DNA-directed RNA polymerase subunit beta

Chain B:



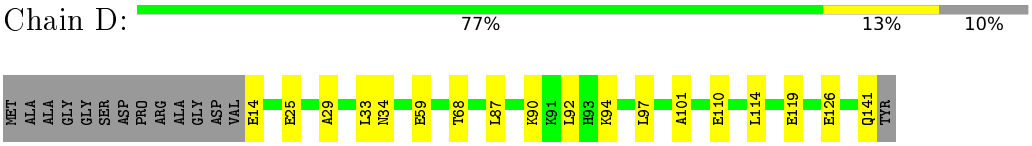
• Molecule 11: DNA-directed RNA polymerase II subunit RPB3

Chain C:



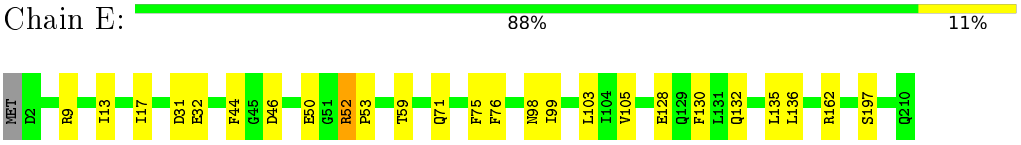
• Molecule 12: RPOL4c domain-containing protein

Chain D:



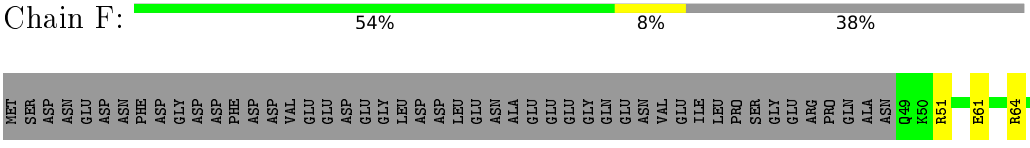
• Molecule 13: DNA-directed RNA polymerase II subunit E

Chain E:



• Molecule 14: DNA-directed RNA polymerase II subunit F

Chain F:

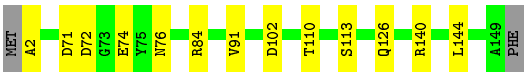


• Molecule 15: DNA-directed RNA polymerase II subunit RPB7

Chain G:

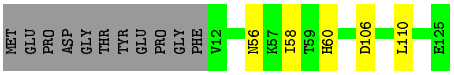
• Molecule 16: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H:



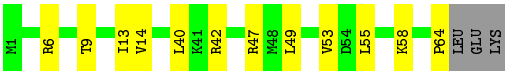
• Molecule 17: DNA-directed RNA polymerase II subunit RPB9

Chain I:



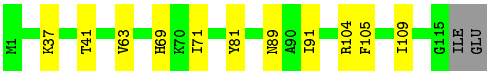
• Molecule 18: DNA-directed RNA polymerases I, II, and III subunit RPABC5

Chain J:



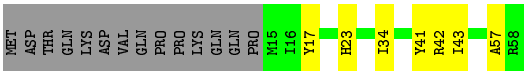
• Molecule 19: RNA_pol_L_2 domain-containing protein

Chain K:



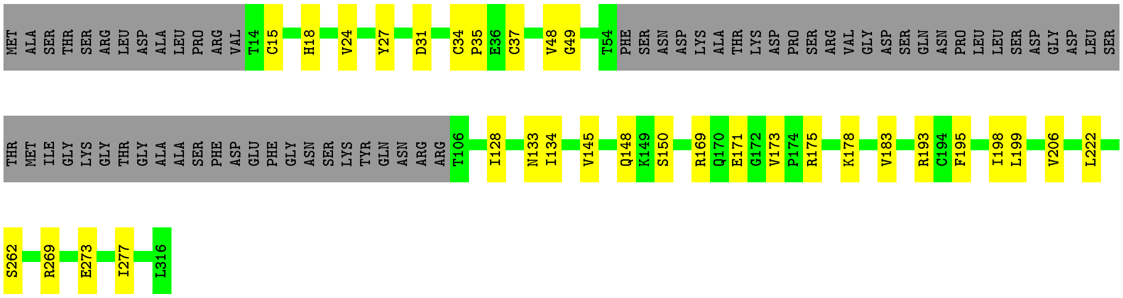
• Molecule 20: RNA polymerase II subunit K

Chain L:



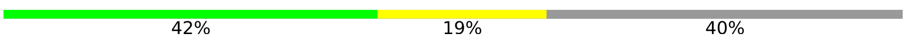
• Molecule 21: Transcription initiation factor IIB

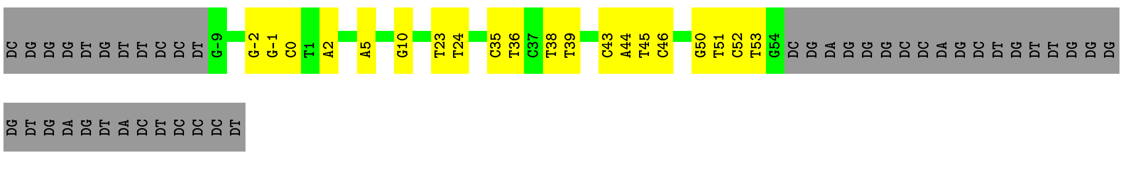
Chain M:



• Molecule 22: Non-template DNA

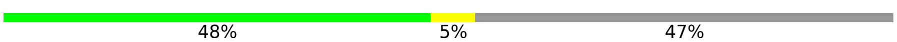
Chain N:

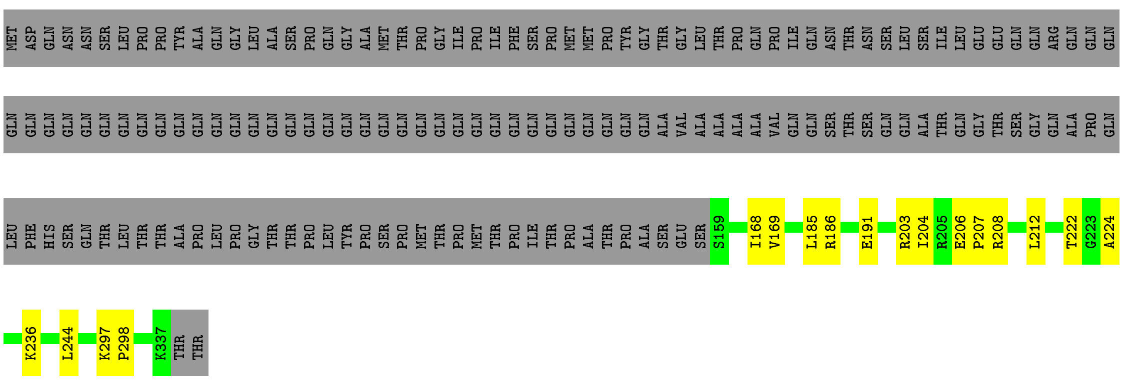




• Molecule 23: TATA-box-binding protein

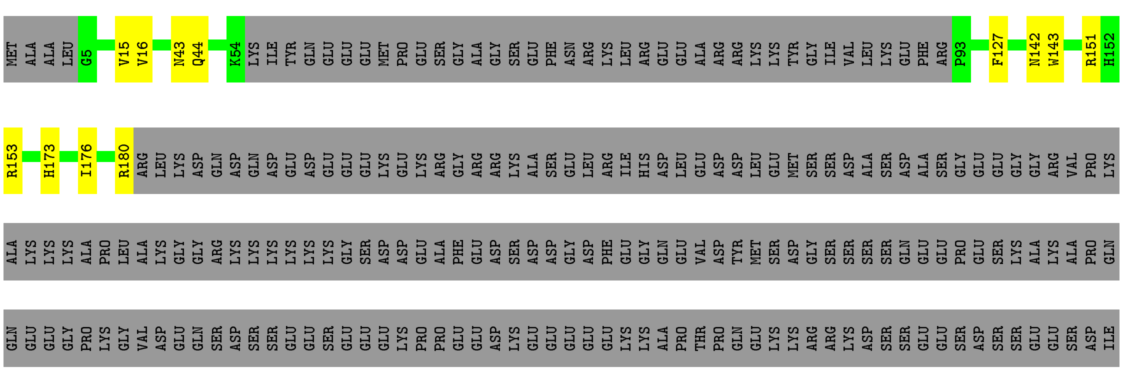
Chain O:





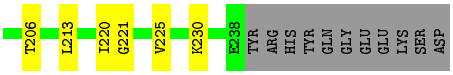
• Molecule 24: General transcription factor IIF subunit 1

Chain Q:



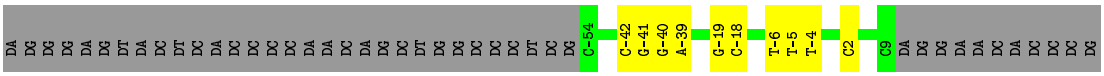
• Molecule 25: General transcription factor IIF subunit 2

Chain R:



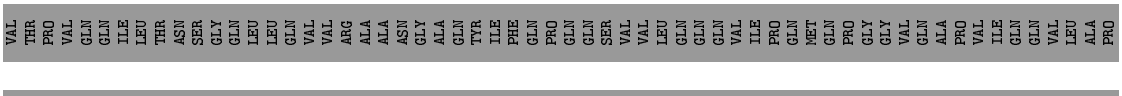
• Molecule 26: Template DNA

Chain T:



• Molecule 27: Transcription initiation factor IIA subunit 1

Chain U:

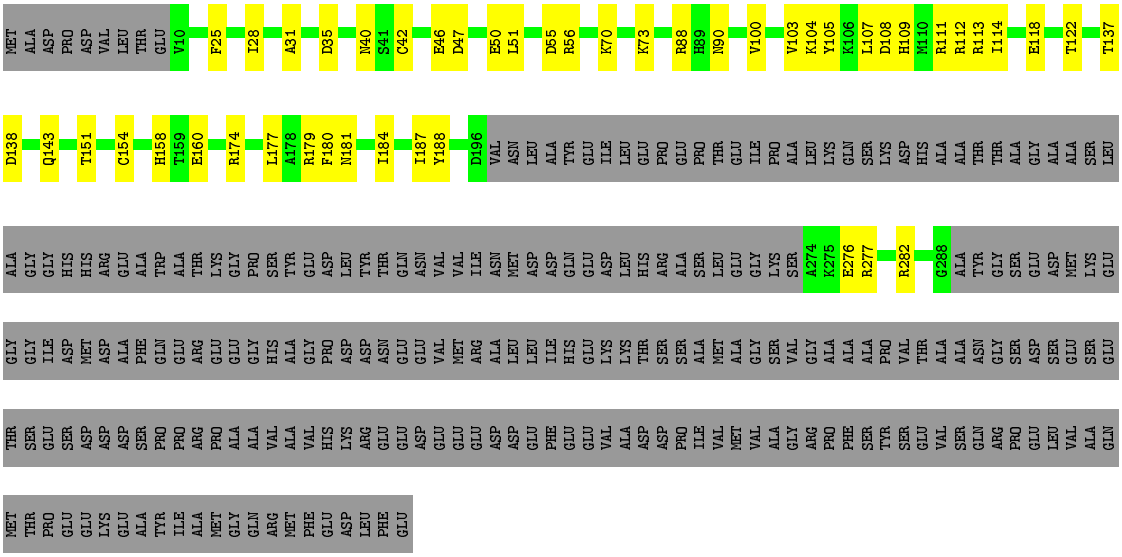


• Molecule 28: Transcription initiation factor IIA subunit 2

Chain V:

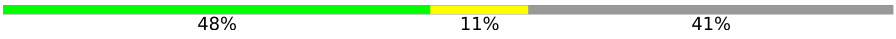
• Molecule 29: General transcription factor IIE subunit 1

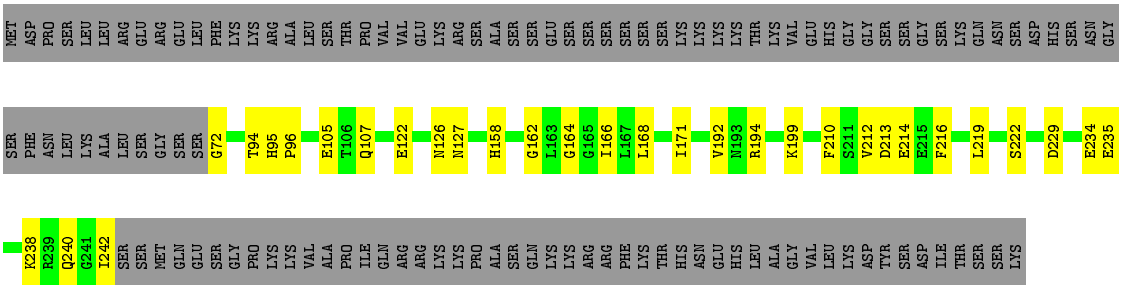
Chain W:



• Molecule 30: Transcription initiation factor IIE subunit beta

Chain X:





• Molecule 31: Unassigned peptide, likely TFIIE-beta

Chain Y:

There are no outlier residues recorded for this chain.

• Molecule 32: Unassigned peptide, likely XPB

Chain Z:

There are no outlier residues recorded for this chain.

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	15226	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	41.1	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

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).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|>5$	RMSZ	#$\|Z\|>5$
1	0	0.25	0/5875	0.41	0/7955
2	1	0.25	0/2210	0.40	0/2975
3	2	0.26	0/3230	0.42	0/4376
4	3	0.27	0/1241	0.44	0/1665
5	4	0.28	0/2103	0.46	0/2846
6	5	0.24	0/529	0.40	0/714
7	6	0.24	0/2793	0.41	0/3780
8	7	0.24	0/4994	0.40	0/6745
9	A	0.24	0/11479	0.41	0/15496
10	B	0.24	0/9257	0.41	0/12493
11	C	0.24	0/2102	0.42	0/2857
12	D	0.24	0/1064	0.37	0/1428
13	E	0.24	0/1752	0.41	0/2366
14	F	0.23	0/646	0.40	0/871
15	G	0.25	0/1382	0.41	0/1874
16	H	0.24	0/1207	0.44	0/1628
17	I	0.23	0/949	0.44	0/1284
18	J	0.25	0/516	0.40	0/696
19	K	0.25	0/939	0.39	0/1271
20	L	0.30	0/378	0.42	0/500
21	M	0.24	0/1983	0.41	0/2679
22	N	0.52	0/1478	0.89	0/2283
23	O	0.25	0/1448	0.42	0/1948
24	Q	0.24	0/1167	0.40	0/1576
25	R	0.24	0/1817	0.41	0/2445
26	T	0.53	0/1461	0.85	0/2249
27	U	0.23	0/945	0.42	0/1274
28	V	0.24	0/816	0.41	0/1105
29	W	0.27	0/1686	0.46	0/2266
30	X	0.25	0/1427	0.41	0/1916
All	All	0.26	0/68874	0.45	0/93561

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	0	5751	0	5794	109	0
2	1	2167	0	2175	36	0
3	2	3158	0	3213	45	0
4	3	1225	0	1224	48	0
5	4	2066	0	2098	50	0
6	5	523	0	530	8	0
7	6	2732	0	2698	45	0
8	7	4890	0	4949	68	0
9	A	11274	0	11406	160	0
10	B	9076	0	9116	90	0
11	C	2059	0	2007	25	0
12	D	1050	0	1033	12	0
13	E	1721	0	1737	20	0
14	F	636	0	665	10	0
15	G	1351	0	1358	11	0
16	H	1186	0	1147	8	0
17	I	928	0	859	5	0
18	J	507	0	523	15	0
19	K	920	0	942	11	0
20	L	373	0	378	7	0
21	M	1953	0	1987	20	0
22	N	1318	0	721	12	0
23	O	1422	0	1514	12	0
24	Q	1138	0	1103	6	0
25	R	1788	0	1819	19	0
26	T	1303	0	714	10	0
27	U	930	0	888	16	0
28	V	806	0	818	13	0
29	W	1659	0	1665	60	0
30	X	1403	0	1428	30	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
31	Y	95	0	24	0	0
32	Z	40	0	10	0	0
33	0	8	0	0	0	0
34	3	2	0	0	0	0
34	4	2	0	0	0	0
34	6	3	0	0	0	0
34	A	2	0	0	0	0
34	B	1	0	0	0	0
34	C	1	0	0	0	0
34	I	2	0	0	0	0
34	J	1	0	0	0	0
34	L	1	0	0	0	0
34	M	1	0	0	0	0
34	W	1	0	0	0	0
35	A	1	0	0	0	0
All	All	67474	0	66543	854	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 6.

The worst 5 of 854 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)
5:4:255:CYS:SG	5:4:258:HIS:ND1	2.34	1.00
4:3:11:THR:HA	29:W:111:ARG:CZ	1.92	0.99
3:2:120:ILE:O	5:4:100:LYS:NZ	2.06	0.87
1:0:571:THR:OG1	1:0:573:ASP:OD1	1.93	0.86
5:4:62:SER:OG	5:4:132:LEU:O	1.95	0.84

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 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	Percentiles	
1	0	710/760 (93%)	682 (96%)	28 (4%)	0	100	100
2	1	253/548 (46%)	245 (97%)	8 (3%)	0	100	100
3	2	380/462 (82%)	365 (96%)	15 (4%)	0	100	100
4	3	147/309 (48%)	136 (92%)	11 (8%)	0	100	100
5	4	259/308 (84%)	255 (98%)	4 (2%)	0	100	100
6	5	64/71 (90%)	63 (98%)	1 (2%)	0	100	100
7	6	341/395 (86%)	329 (96%)	12 (4%)	0	100	100
8	7	601/782 (77%)	577 (96%)	24 (4%)	0	100	100
9	A	1413/1970 (72%)	1378 (98%)	35 (2%)	0	100	100
10	B	1130/1174 (96%)	1097 (97%)	33 (3%)	0	100	100
11	C	253/275 (92%)	248 (98%)	5 (2%)	0	100	100
12	D	126/142 (89%)	125 (99%)	1 (1%)	0	100	100
13	E	207/210 (99%)	201 (97%)	6 (3%)	0	100	100
14	F	77/127 (61%)	76 (99%)	1 (1%)	0	100	100
15	G	169/172 (98%)	165 (98%)	4 (2%)	0	100	100
16	H	146/150 (97%)	143 (98%)	3 (2%)	0	100	100
17	I	112/125 (90%)	107 (96%)	5 (4%)	0	100	100
18	J	62/67 (92%)	60 (97%)	2 (3%)	0	100	100
19	K	113/117 (97%)	112 (99%)	1 (1%)	0	100	100
20	L	42/58 (72%)	42 (100%)	0	0	100	100
21	M	248/316 (78%)	242 (98%)	6 (2%)	0	100	100
23	O	177/339 (52%)	174 (98%)	3 (2%)	0	100	100
24	Q	134/517 (26%)	130 (97%)	4 (3%)	0	100	100
25	R	218/249 (88%)	213 (98%)	5 (2%)	0	100	100
27	U	109/376 (29%)	102 (94%)	7 (6%)	0	100	100
28	V	97/109 (89%)	95 (98%)	2 (2%)	0	100	100
29	W	198/439 (45%)	195 (98%)	3 (2%)	0	100	100
30	X	169/291 (58%)	163 (96%)	6 (4%)	0	100	100
All	All	7955/10858 (73%)	7720 (97%)	235 (3%)	0	100	100

There are no Ramachandran outliers to report.

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 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	0	624/664 (94%)	624 (100%)	0	100 100
2	1	241/484 (50%)	241 (100%)	0	100 100
3	2	342/399 (86%)	341 (100%)	1 (0%)	92 95
4	3	143/283 (50%)	142 (99%)	1 (1%)	84 90
5	4	234/272 (86%)	234 (100%)	0	100 100
6	5	59/64 (92%)	59 (100%)	0	100 100
7	6	311/352 (88%)	310 (100%)	1 (0%)	92 95
8	7	536/688 (78%)	533 (99%)	3 (1%)	86 92
9	A	1254/1749 (72%)	1253 (100%)	1 (0%)	93 97
10	B	994/1027 (97%)	991 (100%)	3 (0%)	92 95
11	C	234/252 (93%)	234 (100%)	0	100 100
12	D	118/126 (94%)	118 (100%)	0	100 100
13	E	191/192 (100%)	189 (99%)	2 (1%)	76 86
14	F	69/111 (62%)	69 (100%)	0	100 100
15	G	152/153 (99%)	151 (99%)	1 (1%)	84 90
16	H	129/131 (98%)	129 (100%)	0	100 100
17	I	103/112 (92%)	103 (100%)	0	100 100
18	J	53/56 (95%)	53 (100%)	0	100 100
19	K	104/106 (98%)	104 (100%)	0	100 100
20	L	41/55 (74%)	41 (100%)	0	100 100
21	M	215/268 (80%)	215 (100%)	0	100 100
23	O	154/293 (53%)	154 (100%)	0	100 100
24	Q	121/448 (27%)	120 (99%)	1 (1%)	81 89
25	R	196/218 (90%)	195 (100%)	1 (0%)	88 93
27	U	105/324 (32%)	105 (100%)	0	100 100
28	V	90/98 (92%)	89 (99%)	1 (1%)	73 84

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
29	W	182/373 (49%)	181 (100%)	1 (0%)	88 93
30	X	154/261 (59%)	154 (100%)	0	100 100
All	All	7149/9559 (75%)	7132 (100%)	17 (0%)	93 96

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
25	R	230	LYS
29	W	56	ARG
10	B	199	LYS
10	B	897	ARG
10	B	1131	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
7	6	227	HIS
9	A	945	ASN
9	A	1005	HIS
4	3	16	ASN
1	0	114	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 monosaccharides 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).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
33	SF4	0	1000	1	0,12,12	0.00	-	-		

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
33	SF4	0	1000	1	-	-	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.