



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Sep 3, 2017 – 02:14 PM EDT

PDB ID : 5IY6
EMDB ID: : EMD-3307
Title : Human holo-PIC in the closed state
Authors : He, Y.; Yan, C.; Fang, J.; Inouye, C.; Tjian, R.; Ivanov, I.; Nogales, E.
Deposited on : unknown
Resolution : 7.20 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report
for a publicly released PDB/EMDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20029824

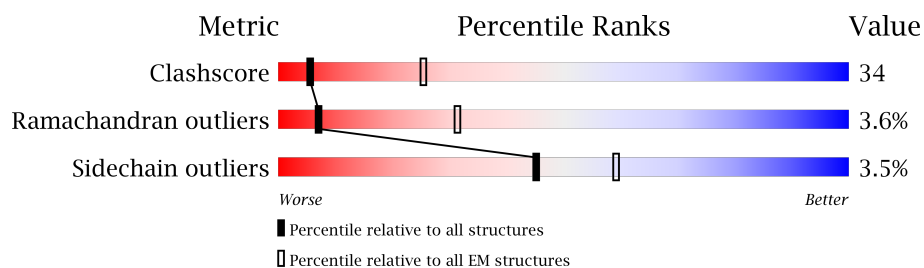
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 7.20 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.





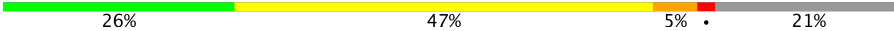








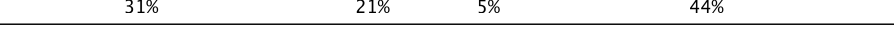
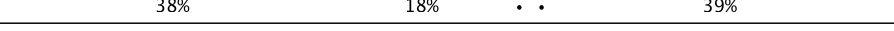
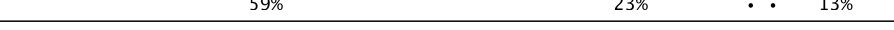

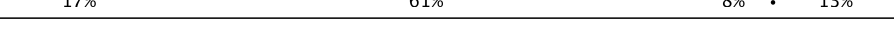




Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	1970	41% 31% . 26%
2	B	1174	50% 46% ..
3	C	275	46% 50% .
4	D	142	65% 26% 9%
5	E	210	50% 46% ..
6	F	127	35% 32% . 32%
7	G	172	65% 32% ..
8	H	150	47% 46% 6% .
9	I	125	41% 50% 9% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	J	67	
11	K	117	
12	L	58	
13	M	316	
14	N	376	
15	O	109	
16	P	339	
17	Q	439	
18	R	291	
19	S	517	
20	T	249	
21	U	301	
22	V	782	
23	W	760	
24	0	395	
25	1	71	
26	2	462	
27	3	308	
28	X	65	
29	Y	65	

2 Entry composition [i](#)

There are 31 unique types of molecules in this entry. The entry contains 61839 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	A	1454	Total	C	N	O	S	0	0
			11515	7234	2058	2150	73		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	1165	Total	C	N	O	S	0	0
			9317	5878	1637	1738	64		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	275	Total	C	N	O	S	0	0
			2213	1386	380	440	7		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	129	Total	C	N	O	S	0	0
			1062	665	179	214	4		

- Molecule 5 is a protein called DNA-directed RNA polymerase II subunit RPB5.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	210	Total	C	N	O	S	0	0
			1723	1088	301	325	9		

- Molecule 6 is a protein called DNA-directed RNA polymerase II subunit RPB6.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	86	Total	C	N	O	S	0	0
			689	437	120	127	5		

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

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

- Molecule 8 is a protein called DNA-directed RNA polymerase II subunit RPB8.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	150	Total	C	N	O	S	0	0
			1205	764	196	239	6		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
9	I	125	Total	C	N	O	S	0	0
			1013	626	177	198	12		

- Molecule 10 is a protein called DNA-directed RNA polymerase II subunit RPB10.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	67	Total	C	N	O	S	0	0
			533	345	90	92	6		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
11	K	117	Total	C	N	O	S	0	0
			937	604	154	177	2		

- Molecule 12 is a protein called DNA-directed RNA polymerase II subunit RPB12.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	46	Total	C	N	O	S	0	0
			388	241	75	66	6		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
13	M	260	Total	C	N	O	S	0	0
			2018	1265	360	376	17		

- Molecule 14 is a protein called Transcription initiation factor IIA subunit 1.

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

- Molecule 15 is a protein called Transcription initiation factor IIA subunit 2.

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

- Molecule 16 is a protein called TATA-box-binding protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	P	185	Total	C	N	O	S	0	0
			1462	946	257	252	7		

- Molecule 17 is a protein called General transcription factor IIE subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	Q	180	Total	C	N	O	S	0	0
			1484	938	262	273	11		

- Molecule 18 is a protein called Transcription initiation factor IIE subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	R	165	Total	C	N	O	S	0	0
			1357	865	235	253	4		

- Molecule 19 is a protein called General transcription factor IIF subunit 1.

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

- Molecule 20 is a protein called General transcription factor IIF subunit 2.

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

- Molecule 21 is a protein called Transcription elongation factor A protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	U	170	Total	C	N	O	S	0	0
			1343	818	247	263	15		

- Molecule 22 is a protein called TFIIH basal transcription factor complex helicase XPB subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	V	475	Total	C	N	O	S	0	0
			3855	2454	663	712	26		

- Molecule 23 is a protein called TFIIH basal transcription factor complex helicase XPD subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	W	665	Total	C	N	O	S	0	0
			5348	3415	932	975	26		

- Molecule 24 is a protein called General transcription factor IIH subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	0	188	Total	C	N	O	S	0	0
			1479	935	258	276	10		

- Molecule 25 is a protein called General transcription factor IIH subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	1	62	Total	C	N	O	S	0	0
			491	317	77	93	4		

- Molecule 26 is a protein called General transcription factor IIH subunit 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	2	274	Total	C	N	O	S	0	0
			2196	1417	377	392	10		

- Molecule 27 is a protein called General transcription factor IIH subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	3	193	Total	C	N	O	S	0	0
			1526	978	252	284	12		

- Molecule 28 is a DNA chain called SCP-X.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	X	65	Total	C	N	O	P	0	0
			1343	633	261	385	64		

- Molecule 29 is a DNA chain called SCP-Y.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	Y	65	Total	C	N	O	P	0	0
			1316	625	236	391	64		

- Molecule 30 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
30	B	1	Total	Mg	0
			1	1	
30	A	1	Total	Mg	0
			1	1	

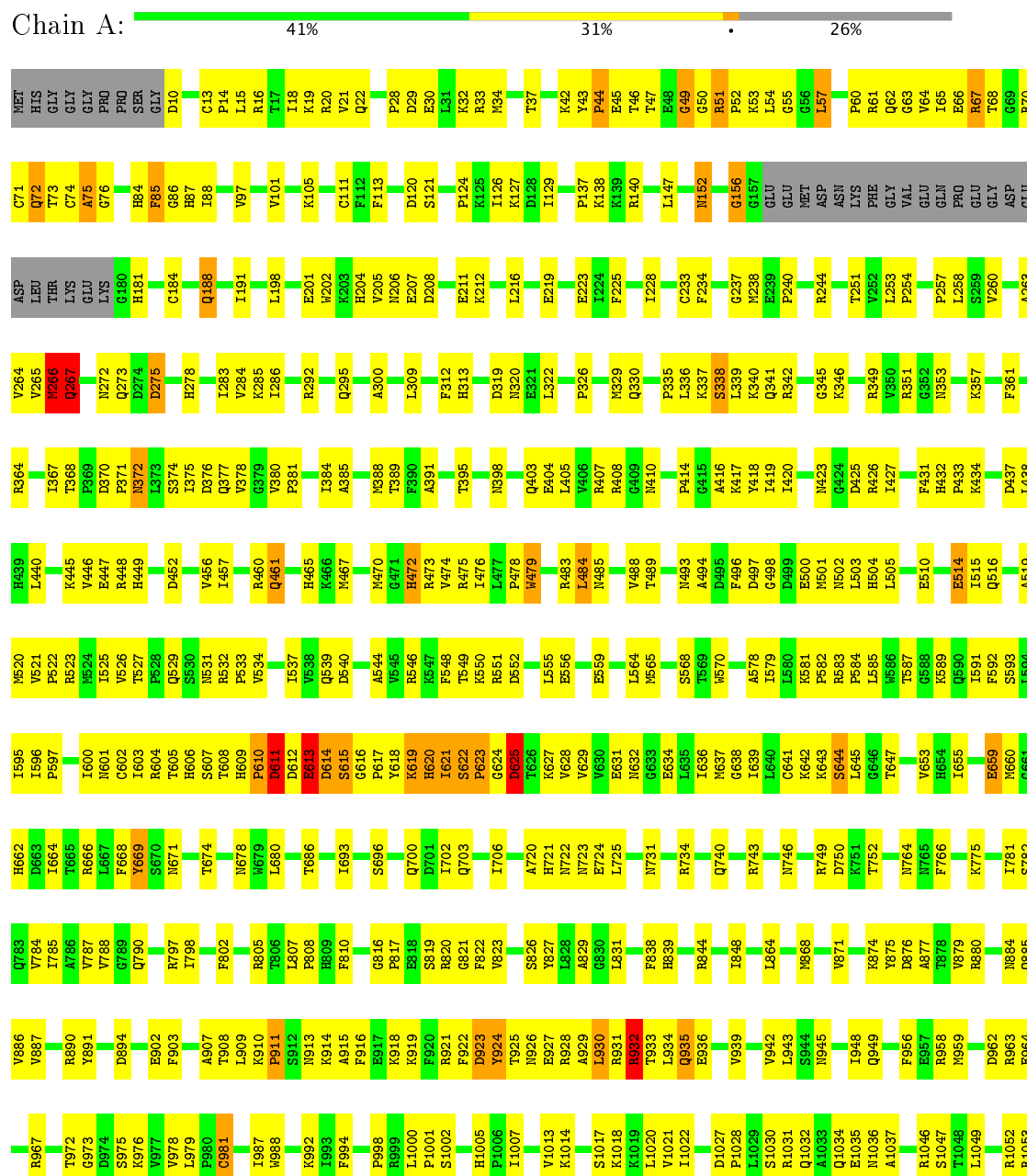
- Molecule 31 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
31	J	1	Total	Zn	0
			1	1	
31	Q	1	Total	Zn	0
			1	1	
31	B	1	Total	Zn	0
			1	1	
31	I	2	Total	Zn	0
			2	2	
31	C	1	Total	Zn	0
			1	1	
31	A	2	Total	Zn	0
			2	2	
31	U	1	Total	Zn	0
			1	1	
31	L	1	Total	Zn	0
			1	1	
31	M	1	Total	Zn	0
			1	1	

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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: DNA-directed RNA polymerase II subunit RPB1



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

I152	P153	I154	M155	L156	L157	Y160	C161	M164	G165	L166	R169	L170	L171	C172	E173	L174	M175	E176	P181	G182	G183	Y184	G189	K192	V193	L194	E198	A201	V205	A209	K210	K211	A216	R222	S223	C224	L225	E226	N227	S228	R230	T232	K236	P237										
V78	E79	P81	P82	R83	Y84	L85	K86	F88	E89	Q90	I91	Y92	L93	S94	T97	H98	W99	E100	R101	D102	P107	M108	M109	P110	N111	E112	R116	N117	Y120	S121	A122	P123	L124	Y125	V126	D127	L128	T129	K130	T131	V132	K134	E135	G136	E137	E138	Q139	Q145	F148					
Met	Tyr	Asp	Ala	Asp	Glu	Asp	Met	Met	Gln	Y110	D111	E112	L121	W122	Q123	E124	A125	C126	W127	L128	V129	I130	E136	K137	V140	R141	Q142	S146	F150	I151	Q152	M153	S154	V155	Q156	R157	E160	D161	A162	P163	P164	I165	D166	K167	Q168	A169	T170	A171	H172	H173	A174	S175	G176	E177

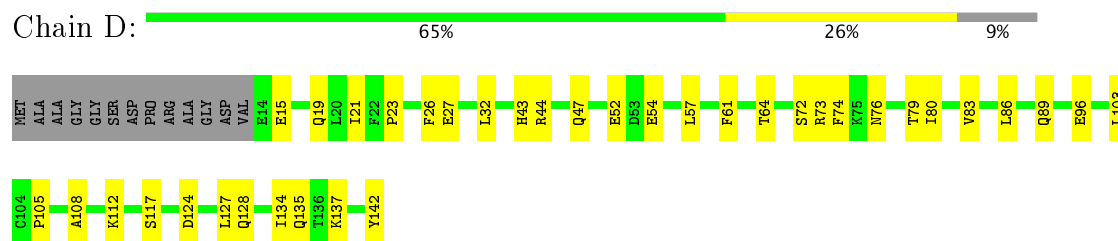
N1142	Q1068	G901	R834	V759	I665	Q500	D428	L318	S238
P1152	R1072	G902	E835	T760	D666	L501	F429	L318	M239
Y1153	Q1073	D805	R836	T761	E669	H502	N430	G322	L240
M1163	P1074	D806	R837	S762	E670	L505	L431		A241
I1167	M1075	T910	G839	S763	E671	L506	E432	A326	R242
	S1079	G915	R840	E765	T672	V609	L433	T331	G243
	R1080	Y916	H841	T766	E673	S602	A434	A335	G244
	G1081		H842	L767	E674	L604	I435		Q245
	G1082	I921	Y845	F769	A676		R438	K334	K248
	G1083	R922			Q883	A519	I439	R335	K249
	L1084	V923	L848	P773	E684	V520	I440	I336	S250
	R1085	R924	D849	T776	E685	G521	D442	V342	A251
	F1086		D850	I776	E686	L522		K345	G253
	G1087	R927	D851	N777	A688	V523	K445	E346	Q254
	E1090	I928	G852	S778		K524	Y446		R255
		P929		I779	T692	N525		P849	T256
			P856	V780		L526	G451		V257
		G932			H695	A527	I452	T359	A258
		D933	R860	Y789	E696	R620	I453	K360	T259
		R934	S861	N789	E697	L621	G454	Y363	K264
		P935	G862	Q790	I698	C622	D455		P268
		A936	D863	E791		R623	Q456		L269
			D864	D792		R624			I270
		H939			I711				L271
		E1026	R866	I795	P712	L626	A458	P368	T272
	Q1101	Y1027	E867	N796	F713	L627	H460	V369	L276
	F1102	L1028	G868	N797	P714	V628		R371	I271
	L1103	Y1029	R869	K798	D715	E629		L372	G277
		N1030	T870	S799	I716	E546		L373	F273
	R1106	G1031		A799		R546		L374	R274
	L1107	F1032	R873	R807	N717	E547	V466		A275
		G946	R874	D802	Q718	R632	S467	G378	L276
	A1110	Q948	R875	R803	S719	L634	I471	R379	G277
	S1111	Y949	R876	G804	P720	L635	L472	D884	F278
	D1112		E877	F805	N722	E551	L473	R385	V279
	P1113	Q1040	D878	S806	T723	L553	T474	R386	S280
	Y1114	I1041	E879	R807	H639	E554		H387	D281
	Q1115		L880	S808	N728	E555	S477	Y388	
	V1116		S881	R809	G729	I556			L284
	H1117	P1045	S882	F810	K730				
	L1118	T1046	T883	R811		L643	H481	R392	T289
	C1119	Y1047	R884	R812	Y736	R646	L482	L395	L290
		Y1048	R885	S813	T737		R483		D291
	G1123	Q1049	R886	Y814	T738	K566	R484	R405	D294
	I1124	R1050	E887		N739	L567	I485	G406	P295
	M1125	L1051	P888	Q817		V651	S487	M407	
	A1126	K1052	R889		R743	S652	P488	F408	E299
		H1053	A971	K821	T746	W653	I489	K409	M300
			R890	G822		D654	G490	I410	V301
	T1130	D1056	D891	R823	L751	D655	R491	L411	K302
	R1131	K1057	P973	S824	Y752		D492	L412	P303
	T1132	D1058	S974	R825	T753	L583	R494	Y418	E307
	H1133		R975	Q825	P754	E588	A496	A419	A308
		S1061	F895		V661		K497	R425	F309
	T1134	R1062	R896	F829	T756	V662	P498		
	Y1135	A1063	R897	E830	K755		R499		Q312
	C1137		R898	K331	P757		P498		
	R1138	P1066	S899	P832	T758		R499		
	C1139	I1067	E900	T832					

• Molecule 3: DNA-directed RNA polymerase II subunit RPB3

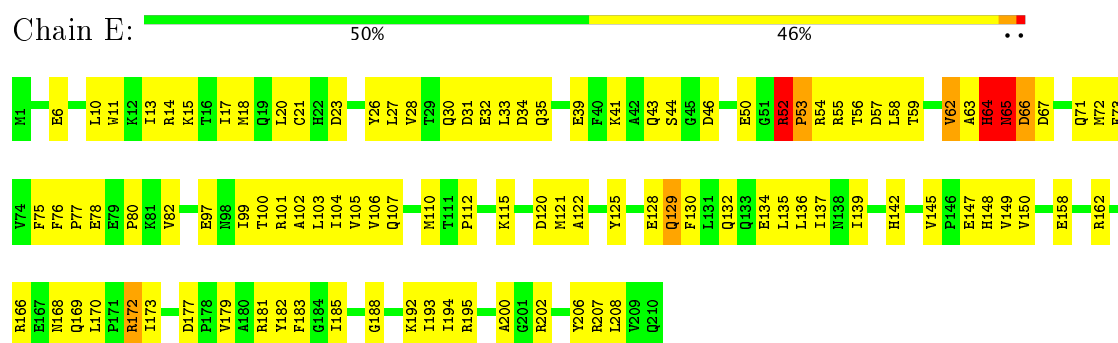
Chain C:  46% 50%

A218	N223	R228	F229	Y230	Y231	N232	V233	E234	S235	S238	L239	R240	P241	I244	A248	K253	D258	Q262	E266	D270	N275	L149	I150	V151	K152	R153	L154	K155	E158	L159	R160	L161	A163	Y164	A165	K166	K167	G168	F169	G170	K171	E172	H173	W176	N177	P178	T179	A183	F184	E185	Y186	N190	R193	H194	T195	Y197	P198	K199	P200	E201	E202	N203	P204	K205	S209	E210	L211	D212	E213	D214	E215	S216	Q217	V19	K20	F21	I22	E23	E24	N25	T26	D27	R35	R36	V37	F38	S39	A40	P43	I44	L45	A46	R47	D48	W49	V50	D53	A54	N55	S56	S57	V58	D61	E62	F63	R67	L68																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														</

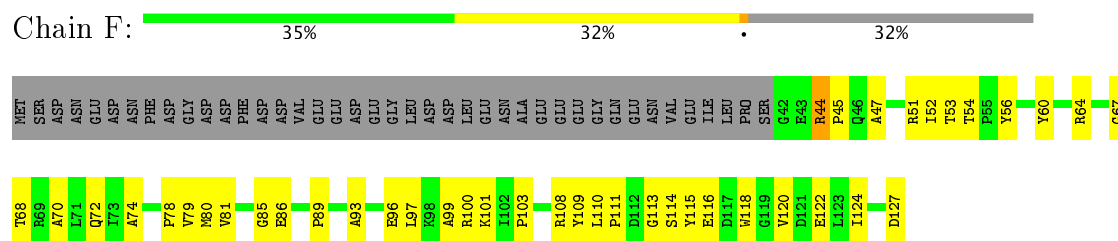
- Molecule 4: DNA-directed RNA polymerase II subunit RPB4



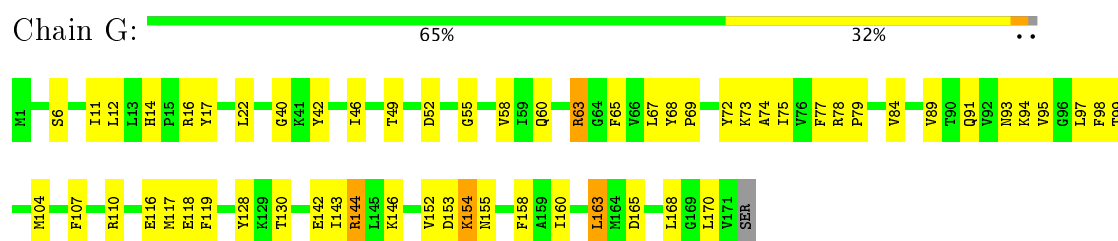
- Molecule 5: DNA-directed RNA polymerase II subunit RPB5



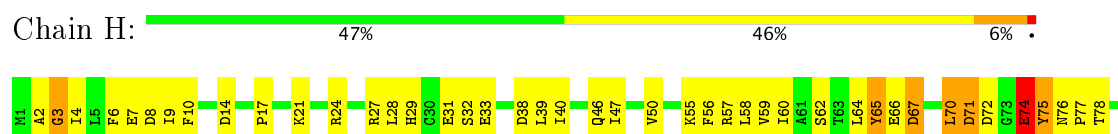
- Molecule 6: DNA-directed RNA polymerase II subunit RPB6



- Molecule 7: DNA-directed RNA polymerase II subunit RPB7



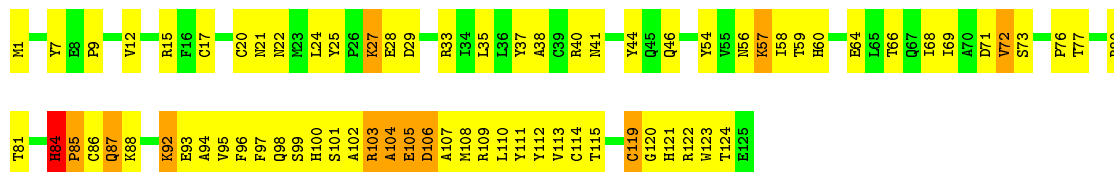
- Molecule 8: DNA-directed RNA polymerase II subunit RPB8





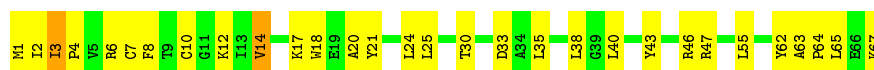
• Molecule 9: DNA-directed RNA polymerase II subunit RPB9

Chain I: 41% 50% 9%



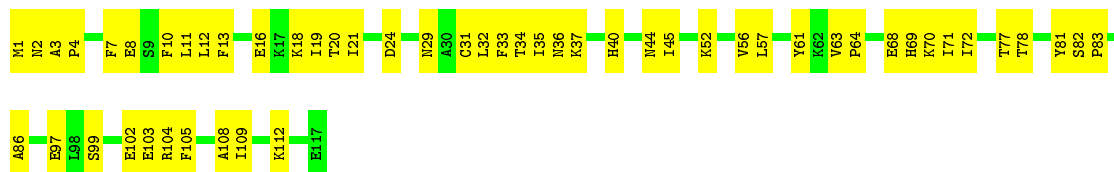
• Molecule 10: DNA-directed RNA polymerase II subunit RPB10

Chain J: 55% 42%



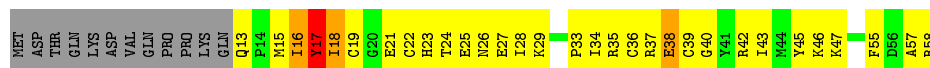
• Molecule 11: DNA-directed RNA polymerase II subunit RPB11-a

Chain K: 55% 45%



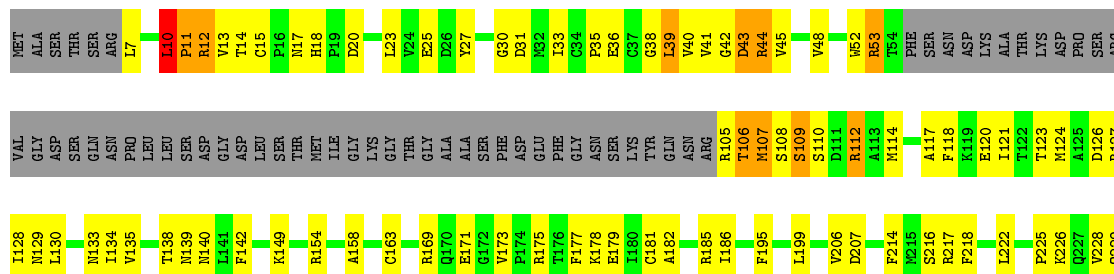
• Molecule 12: DNA-directed RNA polymerase II subunit RPB12

Chain L: 26% 47% 5% 21%



• Molecule 13: Transcription initiation factor IIB

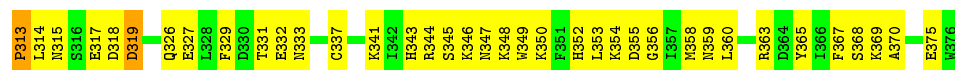
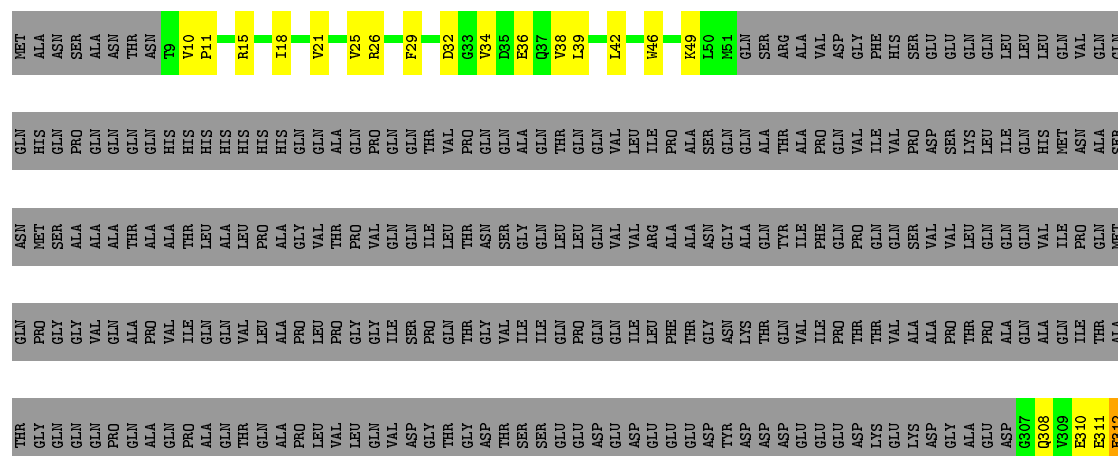
Chain M: 48% 30% 18%





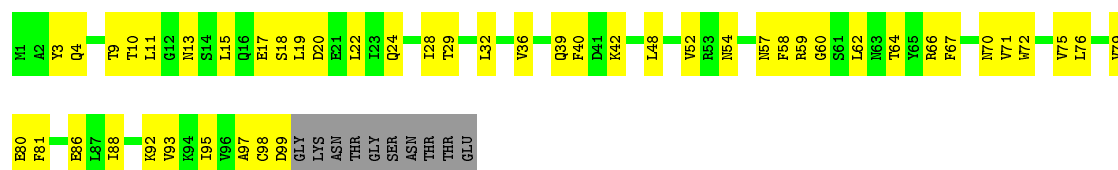
• Molecule 14: Transcription initiation factor IIA subunit 1

Chain N: 15% 14% 70%



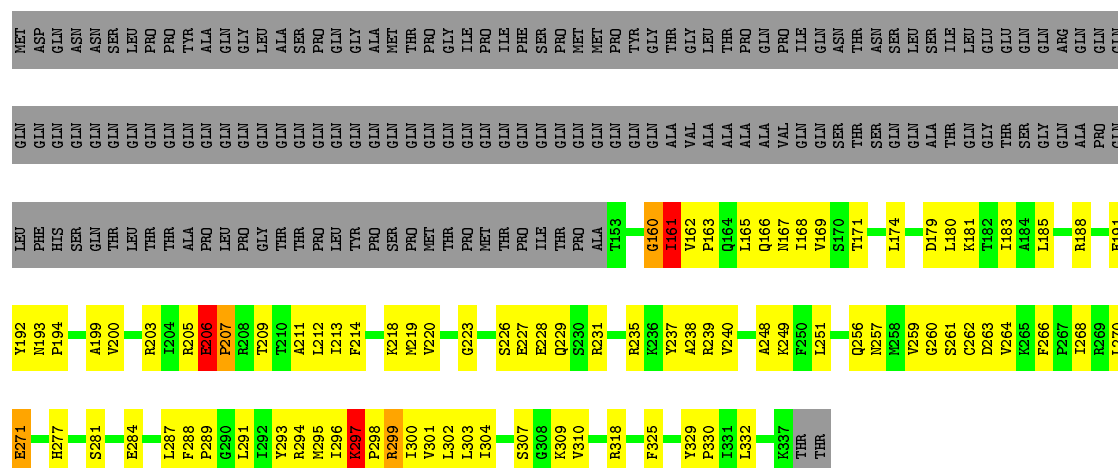
• Molecule 15: Transcription initiation factor IIA subunit 2

Chain O: 48% 43% 9%

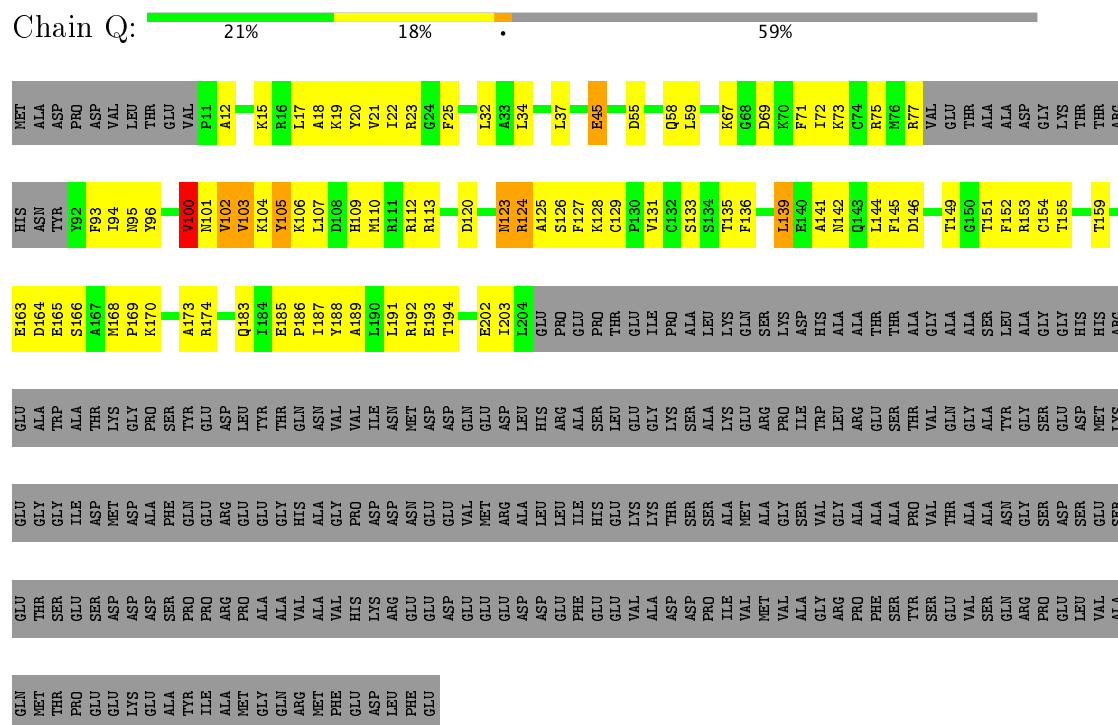


• Molecule 16: TATA-box-binding protein

Chain P: 29% 24% 45%

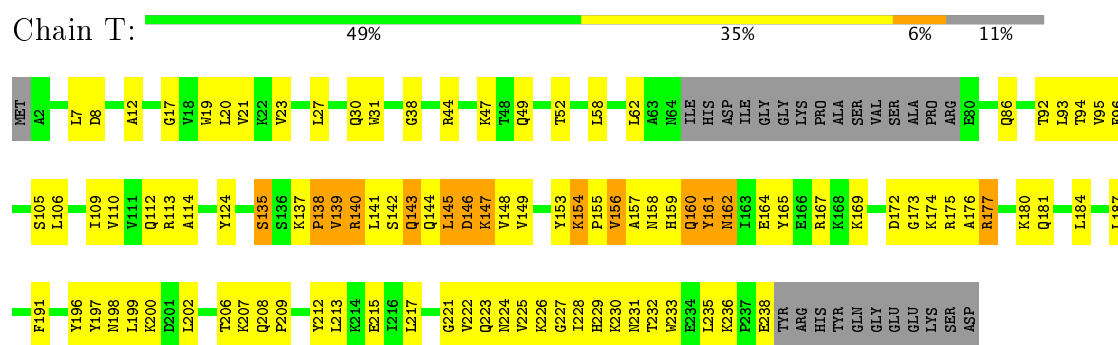


• Molecule 17: General transcription factor IIE subunit 1

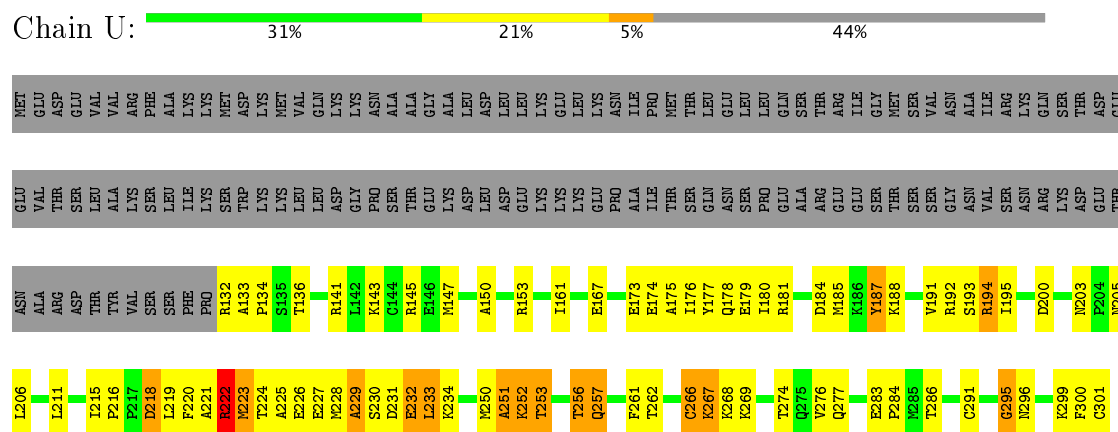




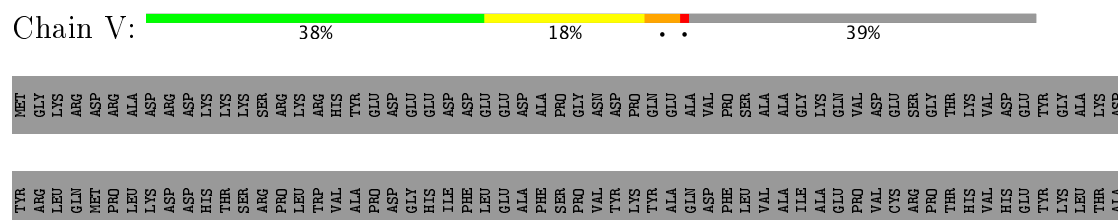
- Molecule 20: General transcription factor IIF subunit 2



- Molecule 21: Transcription elongation factor A protein 1



- Molecule 22: TFIIH basal transcription factor complex helicase XPB subunit



SER	V663	F550	A448	N366	LVS	HIS	TYR
	ME1	S664	F550	S387	SER	PRO	
SER	Q665	R553	F451	A368	ASP	LEU	TYR
GLY	D666	R553	R452	V369	I244	VAL	
ALA	T667	R554	R453	S370	P245	ILE	ALA
ASP				V371	M246	GLN	
THR	Y672	V560	I457	I457	D247	HIS	VAL
ASP	S673	F566	V458	V458	L248	LEU	
VAL	T674	A567	Q459	Q459	F249	LEU	GLY
TYR		L568	A460	F378	D250	GLN	
MET	R678		H461		F251	ASP	GLU
GLU	F679				Y252	PRO	
TYR		Y571	L470	S382	E253	VAL	THR
HIS	D682		R471		Q254	ILE	
SER	Y685	Y581	R472	D385		ARG	ASP
SER	S686	G582	R473	D386		GLU	
ARG	F687	H590	D474		K257	ILE	THR
SER	R688		D475		E264	ARG	
LVS		F606	V478	F392		LEU	TYR
ALA	R689	I607	D479	S394	V272	LEU	
PRO		S608				ASN	ARG
LVS	A702	K609			K282	SER	
LVS	F703	V610	F482	D398	R283	GLU	LEU
HIS					C284	GLY	
VAL	K706	G611	Y489	S404		LEU	SER
THR		D612	E490	V405	Y289	GLU	
HIS	L712	T613	A491	A406		ALA	THR
PRO	L713	S614	M492		F297	GLU	
LEU		F615		I407	R298	LEU	PRO
PHE		D616	M499	S408		LEU	
LVS	A718			T409		ILE	ASP
ALA	THR			V410	L310	THR	
THR	ALA	H629	I502	S411	K311	GLU	GLY
ASP	ASP	G630	A503	M412		GLU	
LEU	LEU	G631	K504	L413	V315	THR	MET
ASP	ASP	S632		L414	L316	PHE	
ALA	ALA	R633	M514	H415		THR	ILE
GLU	GLU	R634	S515	T416	Y319	SER	
GLU	GLU	Q635	P516	T417	Q320	LEU	LVS
GLU	GLU		E517	K418	E321	ALA	
VAL	VAL	Q638	F518	R419		ILE	THR
VAL	VAL	R639	Y519	S420	R325	SER	
ALA	ALA	L640	R520	W421		LYS	SER
GLY	GLY	G641	E521		F328	THR	
GLY	GLY	R642	Y522	R425		ALA	GLY
PHE	PHE	V643	Y523	V426	R332	GLU	
GLY	GLY	L644	A524	W427	A333	SER	VAL
SER	SER	R645		F428	R334	SER	
ARG	ARG	A646	K528	W429		GLY	LEU
SER	SER	R647	E529	L430	G343	GLY	
SER	SER	G648	R530	T431	A344	VAL	VAL
GLN	GLN	G649	I531	T432		PRO	
ALA	ALA	M650		Q433		LYS	HIS
SER	SER	V651	Y534	F434	A353	THR	
ARG	ARG		T535	W435	V357	ARG	ARG
ARG	ARG	P536	N537	W436	K358	VAL	
PHE	PHE	V655				THR	PHE
GLY	GLY	THR	P538	H444	K359	ASP	
THR	THR		N539	T445	C361	PRO	GLU
MET	MET			I446	L362	GLN	
SER	SER		P540	D447		GLY	CYS

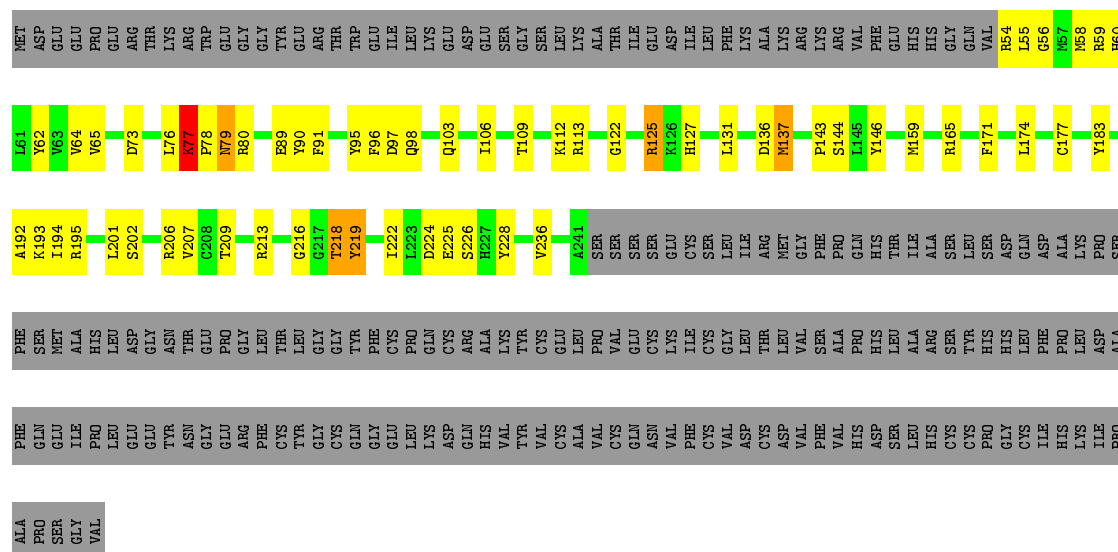
- Molecule 23: TFIIH basal transcription factor complex helicase XPD subunit

Chain W:  59% 23% . . 13%

Q726	R636	D533	K412	R343	S208	Y94	MET
PRO	PHI	PHI	P419	R344	Y209	R95	LVS
PHI	L640	T540	R420	R345	Y210	K101	ASH
ARG	R641		P421	R346	Y211		VAL
GLU		Q543	P422	GLN	D219	S110	ASP
ASP	P644	Y544	D422	HIS	E224	R111	GLY
GLN	Q645	D423	D422	VAL		S112	LEU
GLY	P646	R424	R424	GLN	R227	C116	LEU
LEU	R647	Y551	T425	VAL			VAL
		Y552		GLN			THR
		Y553		GLU			THR
	D650	E554	T428	GLU	V232	T122	PHI
			A429	PRO	F233	P123	PRO
	P654	R563	A430	PRO	D234	L124	THR
	D655	P431	P431	ALA	R125	R125	
	A656	L566	I432	PHI	F126		
	P657		I433	LEU			
	R658	D573	H434	LEU			
	R659			SER	R233	D129	M25
	A660	L581	D439	GLY	R234	Y130	R26
		E582		LEU		D131	E27
		X583	F448	ALA	R238	Y141	L28
	C663	Y584	E449	ARG	V142	R143	R29
			R450	VAL		Q145	R30
	R666			CYS	V270		D33
	R669	H590	V454	ILE	L271		
		G591	I455	GLN	R272		G36
	D673	R592	I456	ARG		D149	H37
	P674	G593		LVS	R230		
	P675	A594	L461	PRO	L281	L152	E41
	L676	L595		LEU	R282		
	P677	L596	Y467	ARG	D283	C155	G46
	P678	L597	P468	PHI	E284	R156	G47
	P679	S598		CYS	V285	F157	
	A680	Y599	D472	ALA	R286	Y158	V50
	D681	A600		GLU	R287		S51
	A682	R601	R487	ARG	L288	D162	L52
			V488	LEU			
	P683	E606	C489	ARG	R293	R166	A58
	P684	G607	P492	SER			Y59
	A685	L608	P493	LEU	R299	L174	Q60
	R686	D609	H493	LEU	D302	Y175	R61
		F610	I494	HIS	R176	A62	A63
	R690		I495	THR	L177	P64	Y63
				LEU	V309		
	P695	G614	N499	GLU		R186	
	P696	L615	D500	ILE	L315		V67
	P697	R616		THR	G316	W189	
		A617		LEU	E317	C190	L70
			T504	THR		P191	I71
				ALA	G321	Y192	Y72
		P623	T510	ASP		F193	C73
	D711	R624	R511	ASP	T325	L194	S74
		P625		F304		A195	R75
		P626	R518		F329	R196	T76
	V714	P627	N519	F403		Y197	V77
				S408	F332	V205	L87
			G521	T409	L333		
		S630	Y520	R334	R334	V206	R88
	R722	R631	N522	Y410	R336	V207	
			L523	A411	R336		

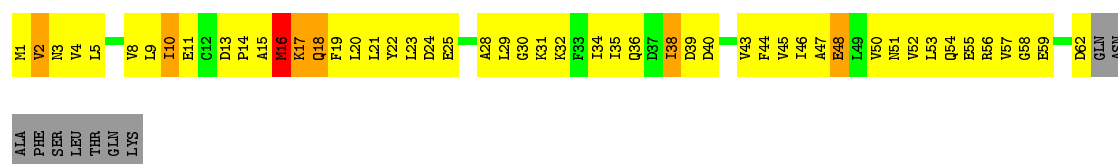
- Molecule 24: General transcription factor IIH subunit 2

Chain 0: 



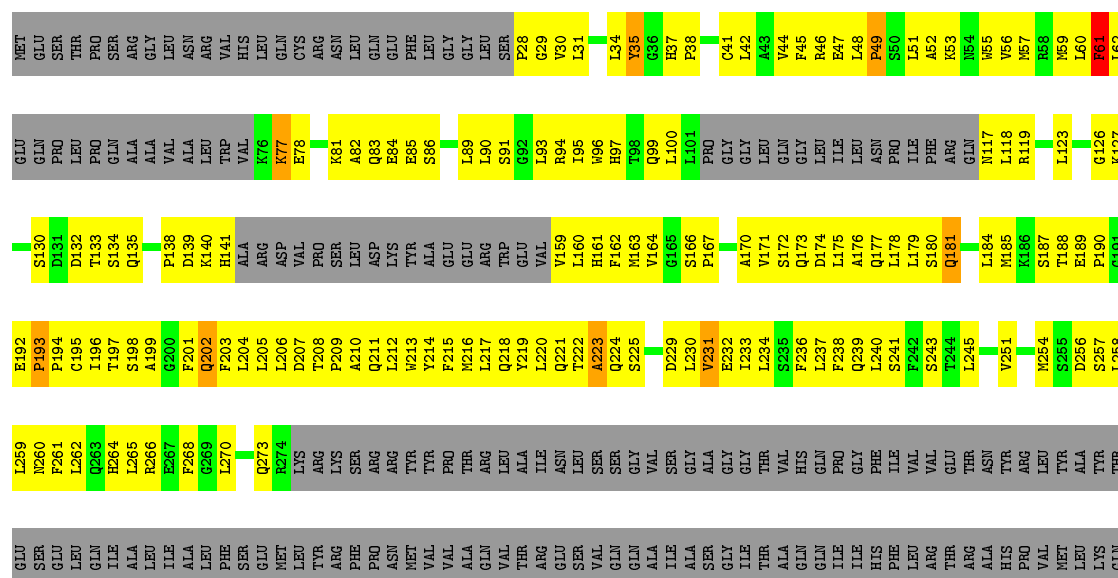
- Molecule 25: General transcription factor IIH subunit 5

Chain 1: 



- Molecule 26: General transcription factor IIH subunit 4

Chain 2: 





4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	34728	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 ($e^-/\text{\AA}^2$)	42	Depositor
Minimum defocus (nm)	2000	Depositor
Maximum defocus (nm)	4000	Depositor
Magnification	27500	Depositor
Image detector	GATAN K2 SUMMIT (4k 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: ZN, MG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z > 2$	RMSZ	$\# Z > 2$
1	A	0.33	5/11727 (0.0%)	0.66	23/15833 (0.1%)
10	J	0.27	0/542	0.56	0/730
11	K	0.26	0/956	0.52	0/1294
12	L	0.28	0/394	0.65	1/524 (0.2%)
13	M	0.26	0/2049	0.69	2/2769 (0.1%)
14	N	0.31	0/945	0.58	1/1274 (0.1%)
15	O	0.26	0/816	0.54	0/1105
16	P	0.26	0/1489	0.60	2/2005 (0.1%)
17	Q	0.28	0/1507	0.59	2/2023 (0.1%)
18	R	0.44	0/1380	0.87	2/1854 (0.1%)
19	S	0.29	0/1167	0.53	1/1576 (0.1%)
2	B	0.29	1/9503 (0.0%)	0.63	4/12831 (0.0%)
20	T	0.27	0/1817	0.57	0/2445
21	U	0.33	0/1358	0.66	2/1820 (0.1%)
22	V	1.40	13/3931 (0.3%)	1.87	96/5298 (1.8%)
23	W	1.50	22/5460 (0.4%)	2.00	154/7390 (2.1%)
24	0	1.49	5/1506 (0.3%)	1.95	43/2038 (2.1%)
25	1	0.83	0/496	1.15	1/669 (0.1%)
26	2	0.88	0/2243	1.18	8/3024 (0.3%)
27	3	0.85	0/1548	1.22	6/2090 (0.3%)
28	X	1.32	18/1510 (1.2%)	1.75	58/2332 (2.5%)
29	Y	1.26	15/1472 (1.0%)	1.71	52/2267 (2.3%)
3	C	0.27	0/2259	0.67	2/3073 (0.1%)
4	D	0.28	0/1077	0.51	0/1446
5	E	0.27	0/1753	0.66	2/2368 (0.1%)
6	F	0.25	0/700	0.51	0/946
7	G	0.27	0/1382	0.55	0/1874
8	H	0.26	0/1227	0.64	1/1654 (0.1%)
9	I	0.25	0/1038	0.90	1/1407 (0.1%)
All	All	0.75	79/63252 (0.1%)	1.09	464/85959 (0.5%)

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
1	A	0	1
17	Q	0	1
18	R	0	8
20	T	0	1
22	V	0	8
23	W	0	11
24	0	0	1
25	1	0	1
26	2	0	8
28	X	0	8
29	Y	0	6
All	All	0	54

The worst 5 of 79 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
29	Y	51	DC	O3'-P	-14.92	1.43	1.61
28	X	53	DA	P-O5'	-9.87	1.49	1.59
28	X	61	DA	C5'-C4'	8.65	1.60	1.51
29	Y	38	DT	P-O5'	8.07	1.67	1.59
23	W	158	TYR	CE1-CZ	8.04	1.49	1.38

The worst 5 of 464 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	I	84	HIS	C-N-CD	-24.67	66.32	120.60
24	0	77	LYS	C-N-CD	-21.82	72.60	120.60
3	C	6	GLN	C-N-CD	-21.15	74.06	120.60
13	M	10	LEU	C-N-CD	-20.90	74.63	120.60
27	3	71	TYR	C-N-CD	-20.67	75.13	120.60

There are no chirality outliers.

5 of 54 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	85	PHE	Peptide
17	Q	100	VAL	Mainchain
18	R	204	ASN	Mainchain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
18	R	205	ASP	Mainchain,Peptide
18	R	206	LYS	Mainchain

5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	11515	0	11609	697	0
2	B	9317	0	9305	579	0
3	C	2213	0	2153	156	0
4	D	1062	0	1042	24	0
5	E	1723	0	1744	117	0
6	F	689	0	715	46	0
7	G	1351	0	1358	53	0
8	H	1205	0	1167	91	0
9	I	1013	0	932	93	0
10	J	533	0	553	51	0
11	K	937	0	959	47	0
12	L	388	0	393	70	0
13	M	2018	0	2059	132	0
14	N	930	0	888	68	0
15	O	806	0	818	50	0
16	P	1462	0	1549	112	0
17	Q	1484	0	1496	230	0
18	R	1357	0	1377	299	0
19	S	1138	0	1103	39	0
20	T	1788	0	1819	171	0
21	U	1343	0	1338	100	0
22	V	3855	0	3872	221	0
23	W	5348	0	5372	179	0
24	0	1479	0	1524	39	0
25	1	491	0	507	239	0
26	2	2196	0	2206	595	0
27	3	1526	0	1561	471	0
28	X	1343	0	725	35	0
29	Y	1316	0	730	36	0
30	A	1	0	0	0	0
30	B	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
31	A	2	0	0	0	0
31	B	1	0	0	0	0
31	C	1	0	0	0	0
31	I	2	0	0	0	0
31	J	1	0	0	0	0
31	L	1	0	0	0	0
31	M	1	0	0	0	0
31	Q	1	0	0	0	0
31	U	1	0	0	0	0
All	All	61839	0	60874	4132	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 34.

The worst 5 of 4132 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)
23:W:421:PHE:CD1	23:W:431:PRO:HG3	1.15	1.65
5:E:27:LEU:HD12	5:E:64:HIS:CD2	1.29	1.64
27:3:59:VAL:HG12	27:3:71:TYR:CD1	1.24	1.64
5:E:27:LEU:HB2	5:E:64:HIS:CD2	1.33	1.63
22:V:315:VAL:HG13	23:W:500:ASP:CB	1.21	1.63

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	A	1450/1970 (74%)	1305 (90%)	97 (7%)	48 (3%)	4	35
2	B	1163/1174 (99%)	1049 (90%)	76 (6%)	38 (3%)	4	35

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	273/275 (99%)	241 (88%)	18 (7%)	14 (5%)	2	26
4	D	127/142 (89%)	118 (93%)	8 (6%)	1 (1%)	22	67
5	E	208/210 (99%)	195 (94%)	7 (3%)	6 (3%)	5	38
6	F	84/127 (66%)	78 (93%)	4 (5%)	2 (2%)	7	42
7	G	169/172 (98%)	158 (94%)	10 (6%)	1 (1%)	28	71
8	H	148/150 (99%)	123 (83%)	13 (9%)	12 (8%)	1	16
9	I	123/125 (98%)	100 (81%)	14 (11%)	9 (7%)	1	19
10	J	65/67 (97%)	53 (82%)	9 (14%)	3 (5%)	3	28
11	K	115/117 (98%)	112 (97%)	3 (3%)	0	100	100
12	L	44/58 (76%)	37 (84%)	3 (7%)	4 (9%)	1	15
13	M	256/316 (81%)	236 (92%)	12 (5%)	8 (3%)	5	36
14	N	109/376 (29%)	100 (92%)	5 (5%)	4 (4%)	4	33
15	O	97/109 (89%)	90 (93%)	7 (7%)	0	100	100
16	P	183/339 (54%)	170 (93%)	8 (4%)	5 (3%)	6	40
17	Q	176/439 (40%)	159 (90%)	11 (6%)	6 (3%)	4	35
18	R	163/291 (56%)	128 (78%)	22 (14%)	13 (8%)	1	17
19	S	134/517 (26%)	123 (92%)	7 (5%)	4 (3%)	5	37
20	T	218/249 (88%)	191 (88%)	17 (8%)	10 (5%)	3	28
21	U	168/301 (56%)	136 (81%)	21 (12%)	11 (6%)	1	22
22	V	473/782 (60%)	400 (85%)	46 (10%)	27 (6%)	2	24
23	W	661/760 (87%)	567 (86%)	69 (10%)	25 (4%)	4	32
24	0	186/395 (47%)	168 (90%)	13 (7%)	5 (3%)	6	40
25	1	60/71 (84%)	53 (88%)	5 (8%)	2 (3%)	4	35
26	2	264/462 (57%)	246 (93%)	14 (5%)	4 (2%)	12	53
27	3	187/308 (61%)	175 (94%)	9 (5%)	3 (2%)	11	51
All	All	7304/10302 (71%)	6511 (89%)	528 (7%)	265 (4%)	7	33

5 of 265 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	67	ARG
1	A	205	VAL
1	A	266	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	267	GLN
1	A	531	ASN

5.3.2 Protein sidechains ⓘ

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	A	1279/1748 (73%)	1247 (98%)	32 (2%)	53	77
2	B	1020/1028 (99%)	997 (98%)	23 (2%)	56	79
3	C	252/252 (100%)	246 (98%)	6 (2%)	54	78
4	D	119/126 (94%)	118 (99%)	1 (1%)	85	92
5	E	192/192 (100%)	186 (97%)	6 (3%)	45	71
6	F	74/111 (67%)	74 (100%)	0	100	100
7	G	152/153 (99%)	149 (98%)	3 (2%)	60	82
8	H	131/131 (100%)	127 (97%)	4 (3%)	45	71
9	I	112/112 (100%)	106 (95%)	6 (5%)	26	58
10	J	56/56 (100%)	56 (100%)	0	100	100
11	K	106/106 (100%)	105 (99%)	1 (1%)	82	91
12	L	43/55 (78%)	43 (100%)	0	100	100
13	M	222/268 (83%)	212 (96%)	10 (4%)	32	63
14	N	105/324 (32%)	104 (99%)	1 (1%)	80	90
15	O	90/98 (92%)	89 (99%)	1 (1%)	78	89
16	P	159/293 (54%)	154 (97%)	5 (3%)	45	71
17	Q	164/373 (44%)	157 (96%)	7 (4%)	33	64
18	R	150/261 (58%)	138 (92%)	12 (8%)	14	45
19	S	121/448 (27%)	118 (98%)	3 (2%)	53	77
20	T	196/218 (90%)	187 (95%)	9 (5%)	31	62
21	U	148/266 (56%)	139 (94%)	9 (6%)	22	55
22	V	422/688 (61%)	403 (96%)	19 (4%)	32	63

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
23	W	577/664 (87%)	541 (94%)	36 (6%)	21	54
24	0	171/352 (49%)	163 (95%)	8 (5%)	30	62
25	1	56/64 (88%)	52 (93%)	4 (7%)	17	49
26	2	238/399 (60%)	229 (96%)	9 (4%)	38	67
27	3	171/272 (63%)	159 (93%)	12 (7%)	18	50
All	All	6526/9058 (72%)	6299 (96%)	227 (4%)	45	69

5 of 227 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
17	Q	123	ASN
20	T	162	ASN
26	2	202	GLN
17	Q	193	GLU
18	R	212	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 61 such sidechains are listed below:

Mol	Chain	Res	Type
9	I	98	GLN
22	V	286	HIS
27	3	148	ASN
9	I	121	HIS
18	R	177	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 13 ligands modelled in this entry, 13 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 ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.