wwPDB EM Validation Summary Report (i)

Aug 6, 2026 – 08:19 AM JST

PDB ID	:	7EGB / pdb_00007egb
EMDB ID	:	EMD-31111
Title	:	TFIID-based holo PIC on SCP promoter
Authors	:	Chen, X.; Wu, Z.; Hou, H.; Qi, Y.; Wang, X.; Li, J.; Xu, Y.
Deposited on	:	2021-03-24
Resolution	:	3.30 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

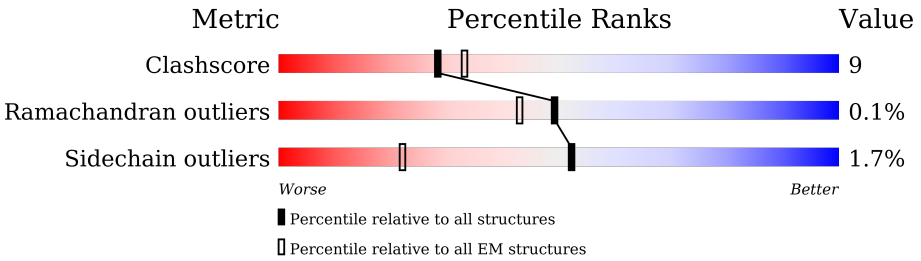
EMDB validation analysis	:	0.0.1.dev133
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4-5-2 with Phenix2.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
EM percentile statistics	:	202505.v01 (Using data in the EMDB archive up until May 2025)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.50

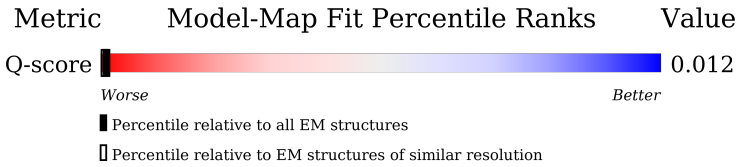
1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.30 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)	Similar EM resolution (#Entries, resolution range(Å))
Clashscore	229148	23984	-
Ramachandran outliers	224038	23583	-
Sidechain outliers	223484	23102	-
Q-score	-	25397	15087 (2.80 - 3.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the EM map (all-atom inclusion < 40%). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	0	309	69%; 53%, 22%, •, 24%
2	1	548	27%; 40%, 8%, 52%
3	2	395	32%; 68%, 15%, 17%
4	3	308	17%; 65%, 19%, •, 15%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	4	462	36% 67% 17% 15%
6	5	71	51% 52% 23% 24%
7	6	782	23% 58% 19% 23%
8	7	760	37% 72% 22% 6%
9	8	346	74% 58% 28% 14%
10	9	323	60% 60% 25% 13%
11	A	1872	25% 24% 5% 70%
12	B	1199	29% 68% 11% 20%
13	D	1085	7% 12% 85%
13	d	1085	15% 12% 85%
14	E	800	39% 54% 13% 32%
14	e	800	67% 56% 11% 33%
15	F	677	39% 43% 15% 39%
15	f	677	43% 48% 11% 40%
16	G	349	40% 36% 59%
17	H	310	22% 52% 13% 33%
18	I	264	24% 39% 6% 55%
18	i	264	46% 42% 54%
19	J	218	7% 33% 8% 59%
19	j	218	44% 36% 7% 56%
20	L	161	32% 32% 12% 53%
20	l	161	66% 59% 7% 34%
21	O	109	40% 70% 21% 9%
22	P	339	9% 47% 6% 47%
23	Q	376	11% 25% 5% 70%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
24	R	316	32% 70% 10% 20%
25	S	517	10% 18% 79%
26	T	249	49% 76% 12% 11%
27	U	439	20% 23% 18% 58%
28	V	291	22% 24% 32% 43%
29	X	94	35% 56% 17% 27%
30	Y	69	54% 83% 17%
31	c	929	14% 12% 86%
32	k	211	46% 43% 54%
33	m	124	70% 48% 17% 6% 30%
34	o	1970	38% 60% 12% 28%
35	p	1174	46% 81% 15%
36	q	275	36% 85% 9% 7%
37	r	142	37% 67% 23% 10%
38	s	210	48% 87% 13%
39	t	127	39% 57% 6% 38%
40	u	172	26% 65% 34%
41	v	150	50% 91% 8%
42	w	125	53% 78% 14% 9%
43	x	67	48% 87% 9%
44	y	117	37% 89% 9%
45	z	58	34% 64% 12% 24%

2 Entry composition (i)

There are 48 unique types of molecules in this entry. The entry contains 112409 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 CDK-activating kinase assembly factor MAT1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	0	236	1932	1211	336	373	12	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 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	2	329	2567	1621	440	479	27	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	3	263	2066	1323	344	380	19	0	0

- Molecule 5 is a protein called General transcription factor IIH subunit 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	4	394	3189	2068	552	557	12	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	54	428	277	67	82	2	0	0

- Molecule 7 is a protein called General transcription and DNA repair factor IIH helicase subunit XPB.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	6	605	4890	3127	848	885	30	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
8	7	714	5751	3683	999	1040	29	0	0

- Molecule 9 is a protein called Cyclin-dependent kinase 7.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
9	8	298	2376	1537	404	423	12	0	0

- Molecule 10 is a protein called Cyclin-H.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	9	281	2293	1465	394	416	18	0	0

- Molecule 11 is a protein called Transcription initiation factor TFIID subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
11	A	559	4580	2924	795	834	27	0	0

- Molecule 12 is a protein called Transcription initiation factor TFIID subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
12	B	963	7796	5011	1315	1412	58	0	0

- Molecule 13 is a protein called Transcription initiation factor TFIID subunit 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
13	D	160	1337	834	249	251	3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
13	d	158	1307	814	238	252	3	0	0

- Molecule 14 is a protein called Transcription initiation factor TFIID subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
14	E	547	4376	2774	759	822	21	0	0
14	e	539	4327	2746	748	814	19	0	0

- Molecule 15 is a protein called Transcription initiation factor TFIID subunit 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
15	F	412	3143	1994	548	583	18	0	0
15	f	403	3081	1954	533	576	18	0	0

- Molecule 16 is a protein called Transcription initiation factor TFIID subunit 7.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
16	G	144	1171	742	215	210	4	0	0

- Molecule 17 is a protein called Transcription initiation factor TFIID subunit 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
17	H	209	1633	1034	283	311	5	0	0

- Molecule 18 is a protein called Transcription initiation factor TFIID subunit 9.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
18	I	120	959	610	166	177	6	0	0
18	i	121	967	615	167	178	7	0	0

- Molecule 19 is a protein called Transcription initiation factor TFIID subunit 10.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
19	J	90	720	466	115	135	4	0	0
19	j	95	759	488	124	143	4	0	0

- Molecule 20 is a protein called Transcription initiation factor TFIID subunit 12.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
20	L	76	622	388	109	122	3	0	0
20	l	107	876	547	158	166	5	0	0

- Molecule 21 is a protein called Transcription initiation factor IIA subunit 2.

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

- Molecule 22 is a protein called TATA-box-binding protein.

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

- Molecule 23 is a protein called Transcription initiation factor IIA subunit 1.

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

- Molecule 24 is a protein called Transcription initiation factor IIB.

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

- Molecule 25 is a protein called General transcription factor IIF subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
25	S	108	872	558	153	159	2	0	0

- Molecule 26 is a protein called General transcription factor IIF subunit 2.

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

- Molecule 27 is a protein called General transcription factor IIE subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
27	U	184	1520	957	272	280	11	0	0

- Molecule 28 is a protein called Transcription initiation factor IIE subunit beta.

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

- Molecule 29 is a DNA chain called DNA (69-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
29	X	69	1429	672	279	409	69	0	0

- Molecule 30 is a DNA chain called DNA (69-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
30	Y	69	1400	664	248	419	69	0	0

- Molecule 31 is a protein called Transcription initiation factor TFIID subunit 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
31	c	127	1011	638	174	193	6	0	0

- Molecule 32 is a protein called Transcription initiation factor TFIID subunit 11.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
32	k	98	785	499	142	139	5	0	0

- Molecule 33 is a protein called Transcription initiation factor TFIID subunit 13.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
33	m	87	724	456	131	131	6	0	0

- Molecule 34 is a protein called DNA-directed RNA polymerase II subunit RPB1.

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

- Molecule 35 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
35	p	1135	9068	5735	1596	1673	64	0	0

- Molecule 36 is a protein called DNA-directed RNA polymerase II subunit RPB3.

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

- Molecule 37 is a protein called DNA-directed RNA polymerase II subunit RPB4.

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

- Molecule 38 is a protein called DNA-directed RNA polymerase II subunit E.

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

- Molecule 39 is a protein called DNA-directed RNA polymerase II subunit F.

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

- Molecule 40 is a protein called DNA-directed RNA polymerase II subunit RPB7.

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

- Molecule 41 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

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

- Molecule 42 is a protein called DNA-directed RNA polymerase II subunit RPB9.

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

- Molecule 43 is a protein called RPB10.

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

- Molecule 44 is a protein called RNA_pol_L_2 domain-containing protein.

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

- Molecule 45 is a protein called RPB12.

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

- Molecule 46 is ZINC ION (CCD ID: ZN) (formula: Zn).

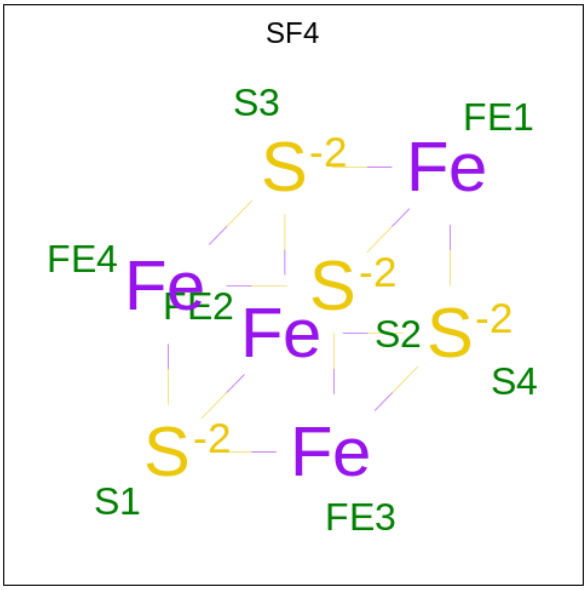
Mol	Chain	Residues	Atoms		AltConf
			Total	Zn	
46	0	2	2	2	0
46	2	3	3	3	0
46	3	2	2	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			AltConf
			Total	Zn		
46	R	1	1	1		0
46	U	1	1	1		0
46	o	2	2	2		0
46	p	1	1	1		0
46	q	1	1	1		0
46	w	2	2	2		0
46	x	1	1	1		0
46	z	1	1	1		0

- Molecule 47 is IRON/SULFUR CLUSTER (CCD ID: SF4) (formula: Fe_4S_4).



Mol	Chain	Residues	Atoms			AltConf
			Total	Fe	S	
47	7	1	8	4	4	0

- Molecule 48 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

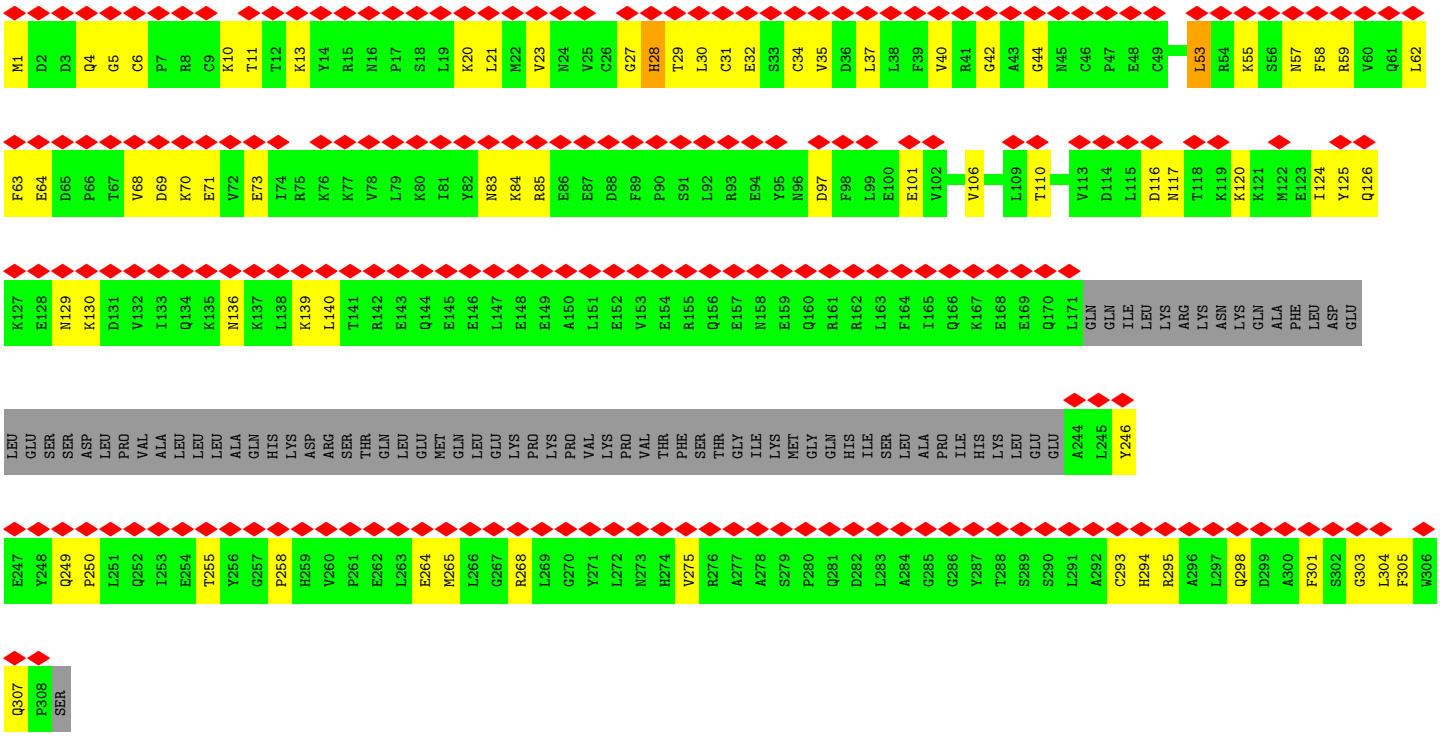
Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
48	o	1	1	1	0

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 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.

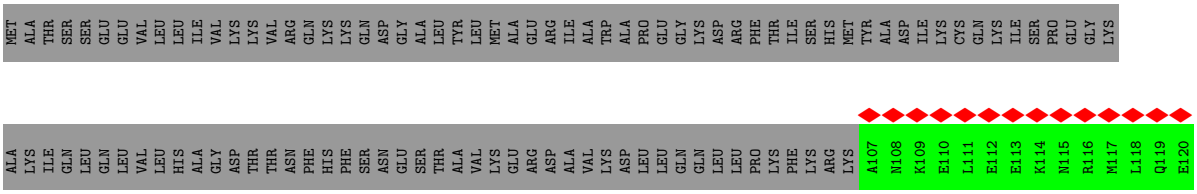
- Molecule 1: CDK-activating kinase assembly factor MAT1

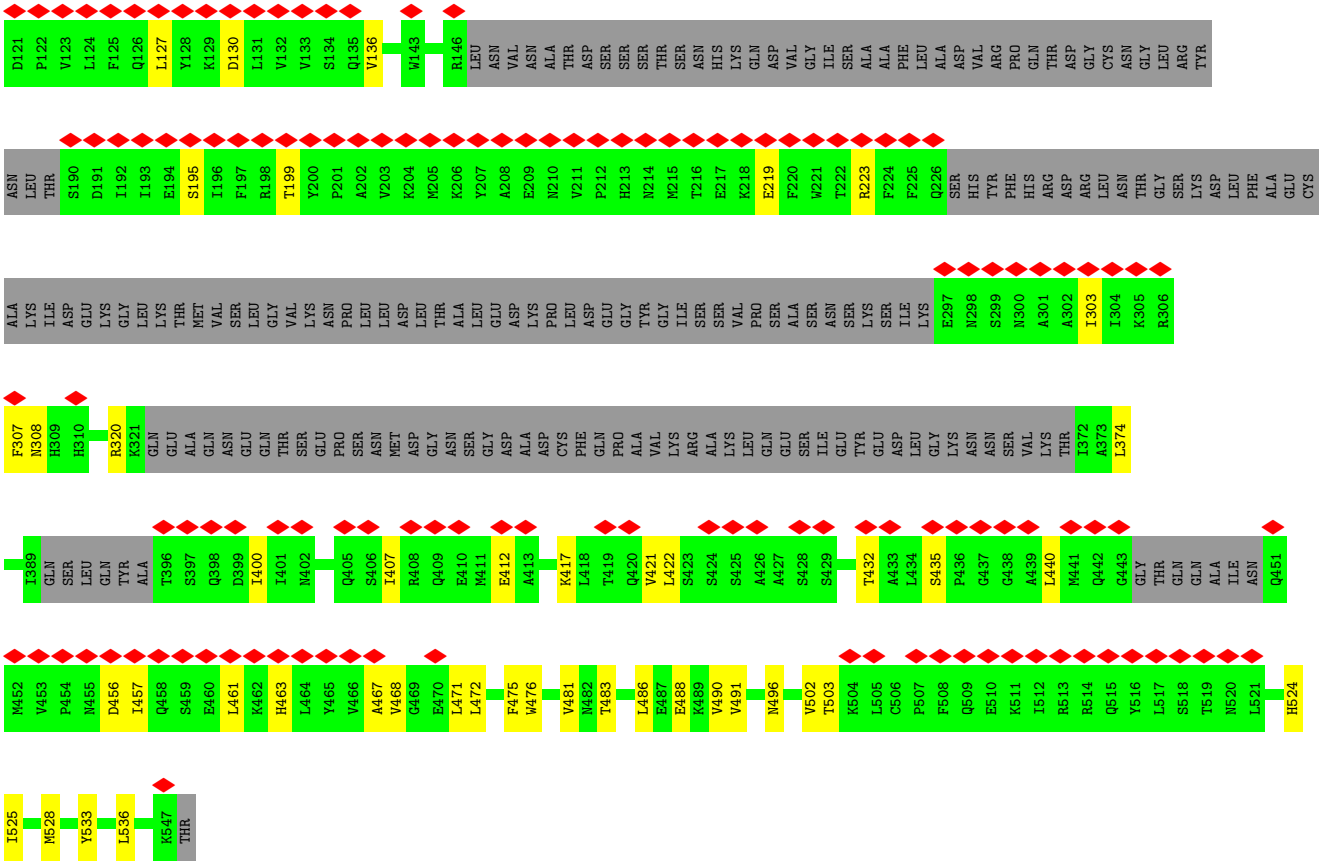
Chain 0:



- Molecule 2: General transcription factor IIH subunit 1

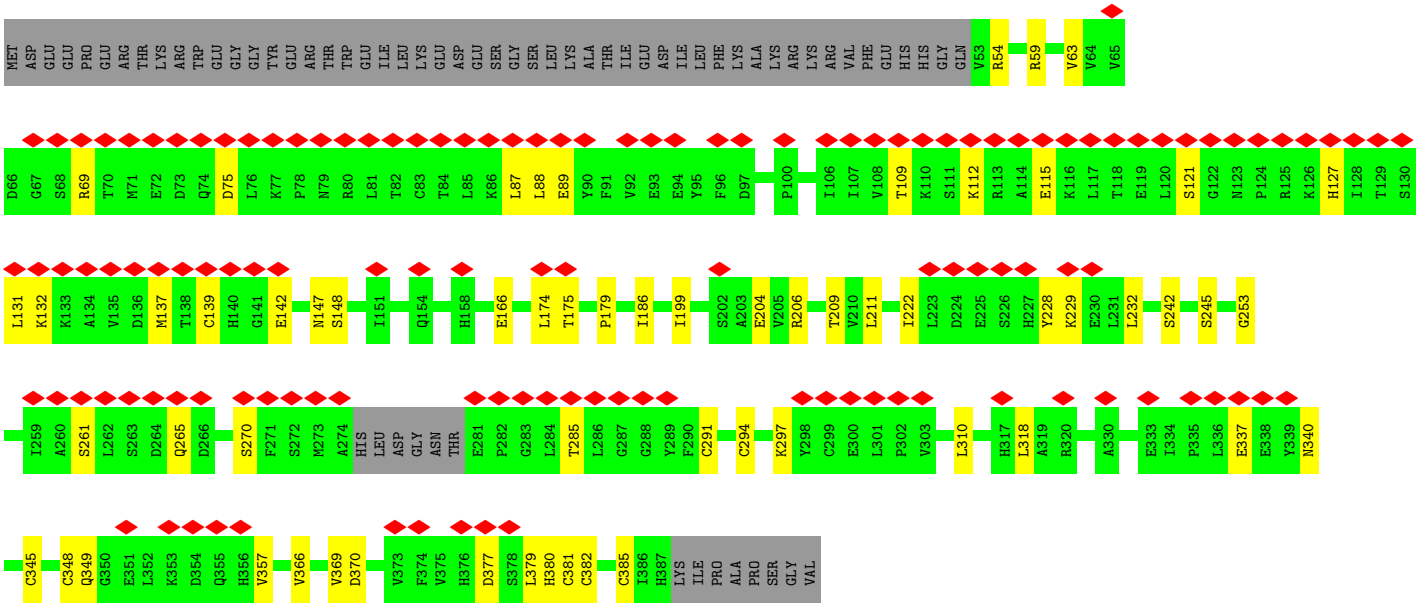
Chain 1:





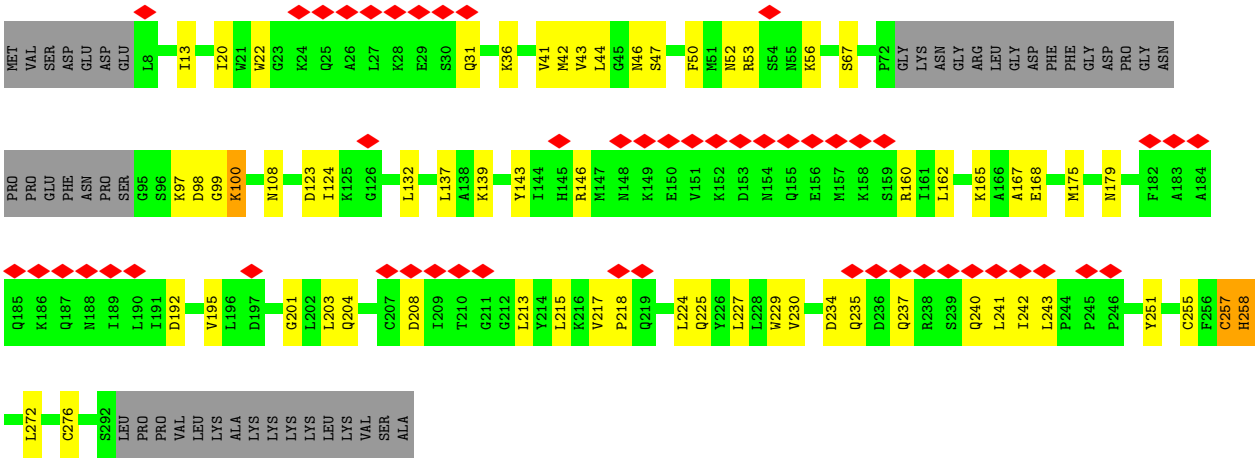
- Molecule 3: General transcription factor IIH subunit 2

Chain 2:



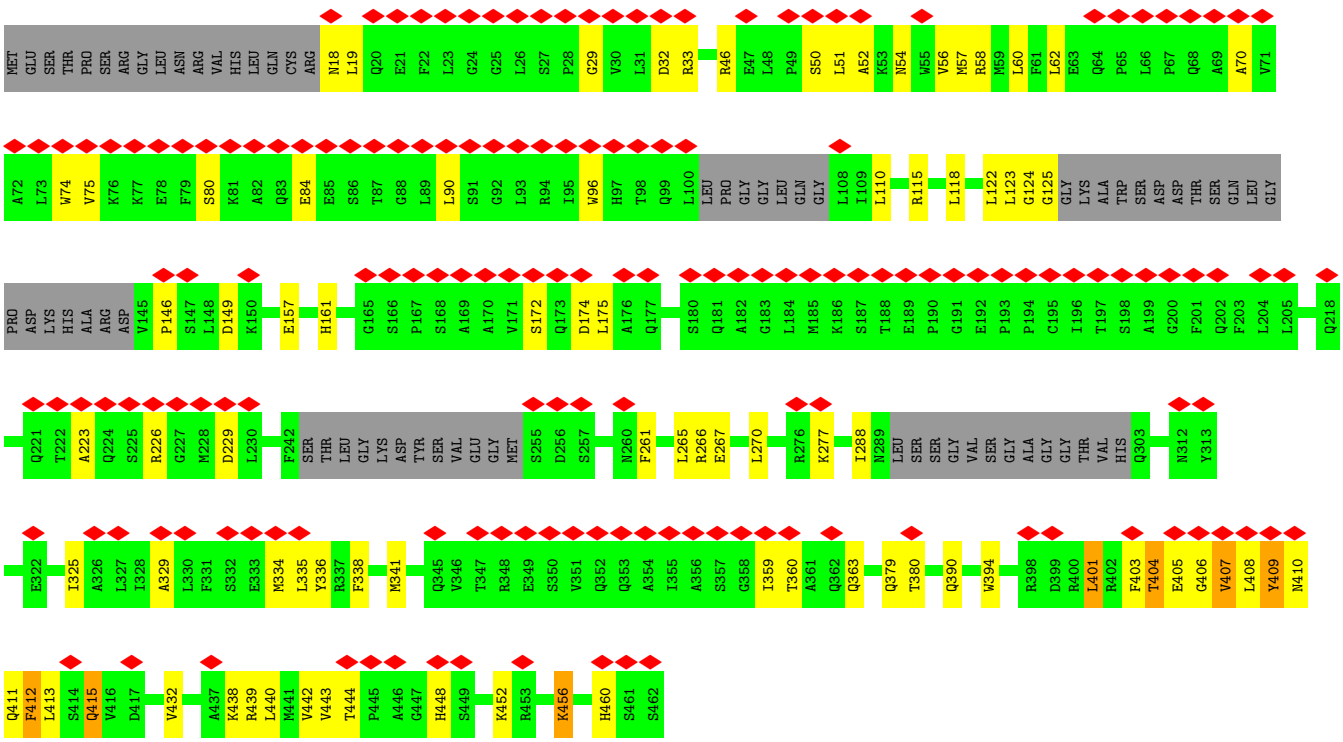
- Molecule 4: General transcription factor IIH subunit 3

Chain 3: 17% 65% 19% • 15%



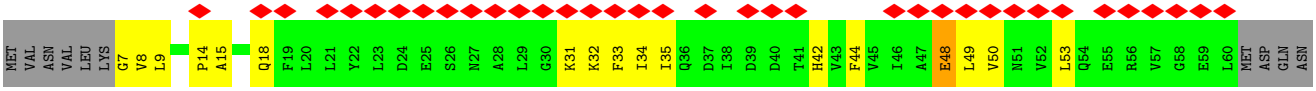
- Molecule 5: General transcription factor IIH subunit 4

Chain 4: 36% 67% 17% • 15%



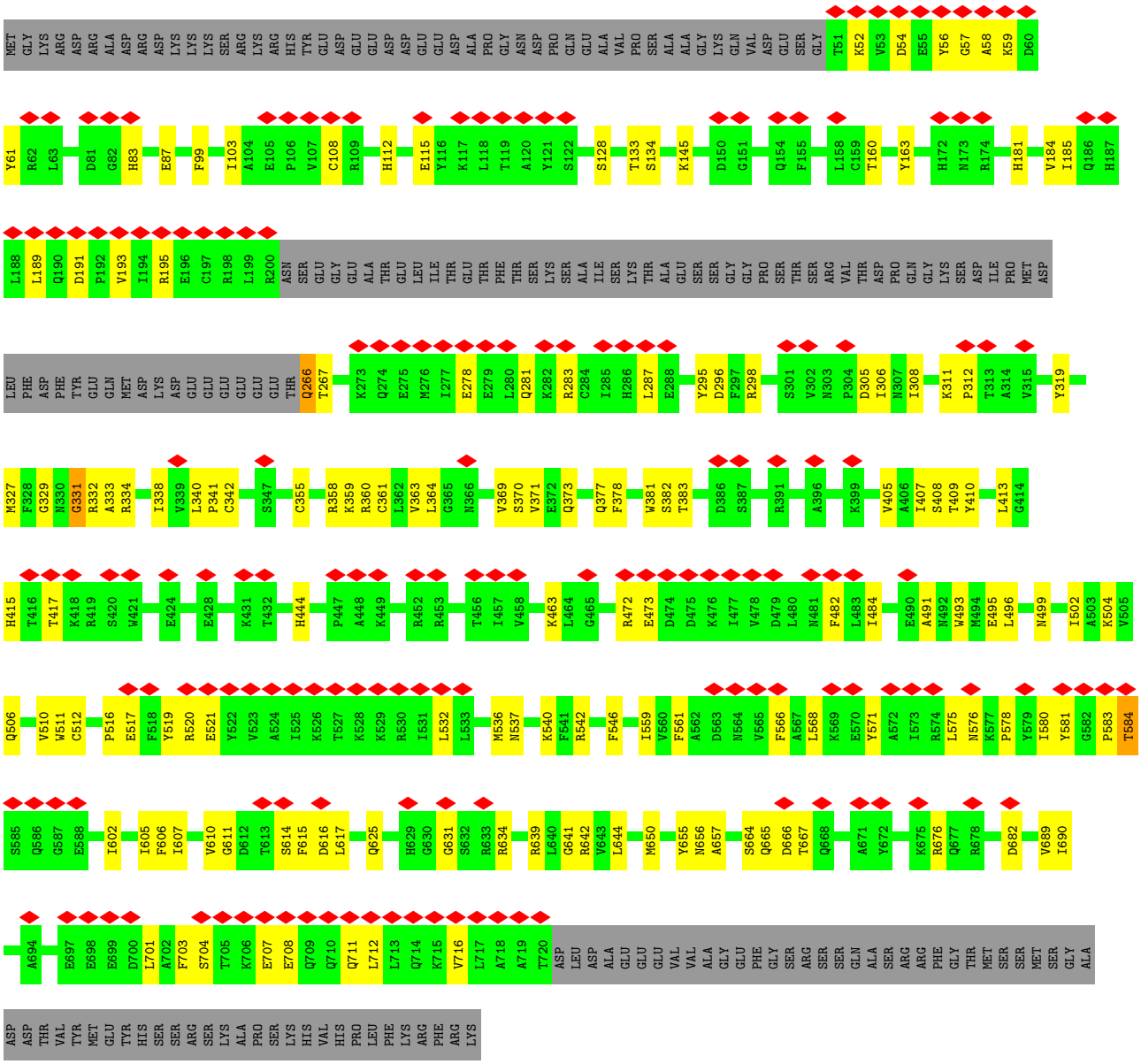
- Molecule 6: General transcription factor IIH subunit 5

Chain 5: 51% 52% 23% • 24%



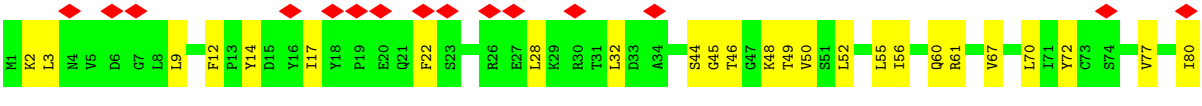
• Molecule 7: General transcription and DNA repair factor IIH helicase subunit XPB

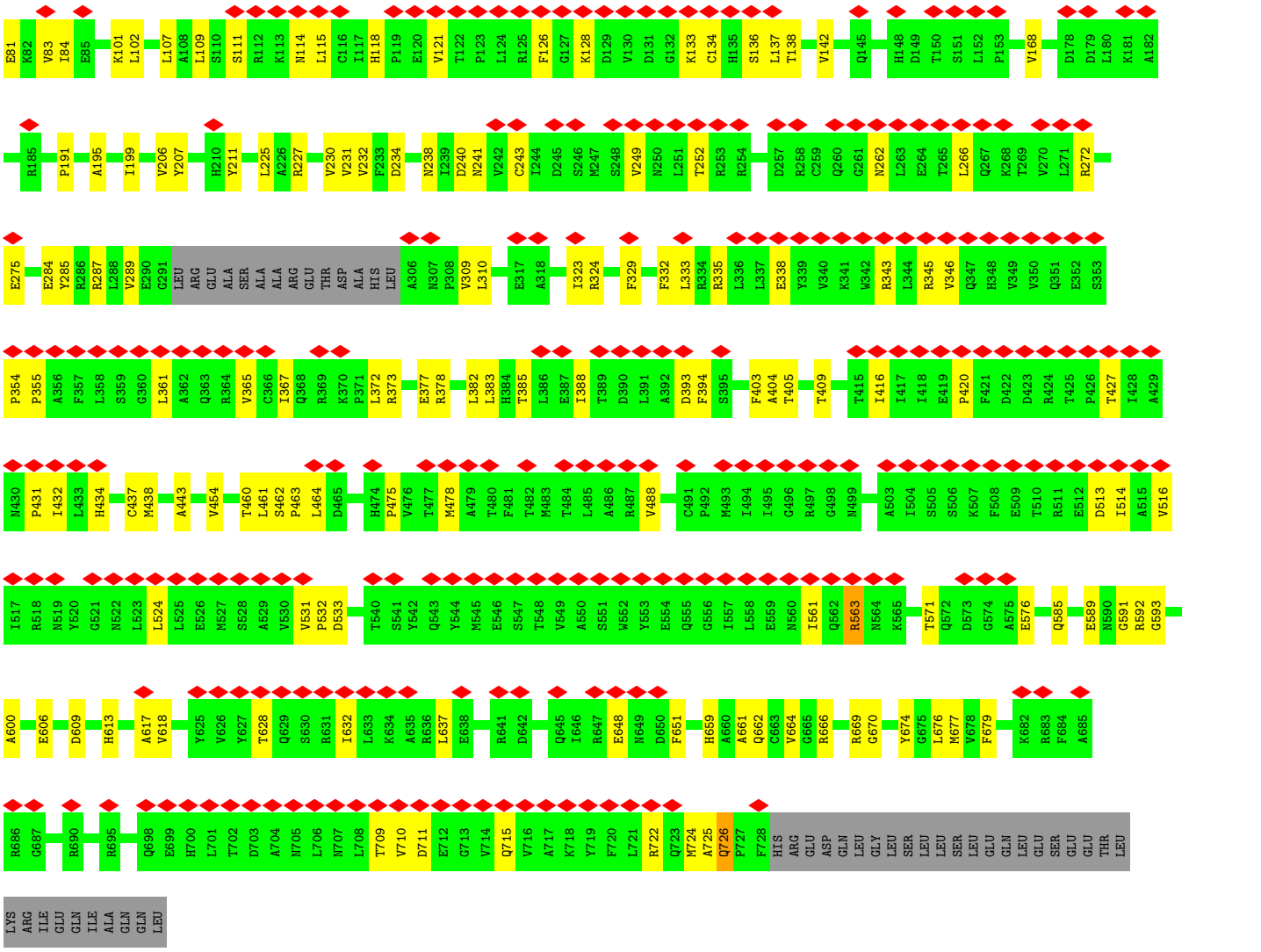
Chain 6:



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

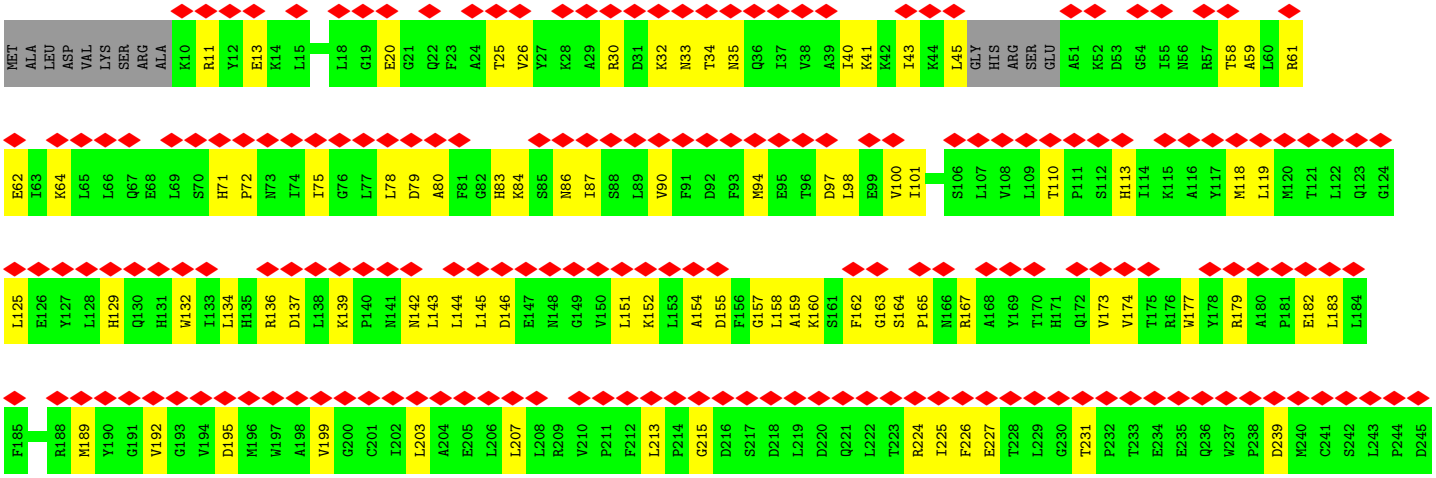
Chain 7:

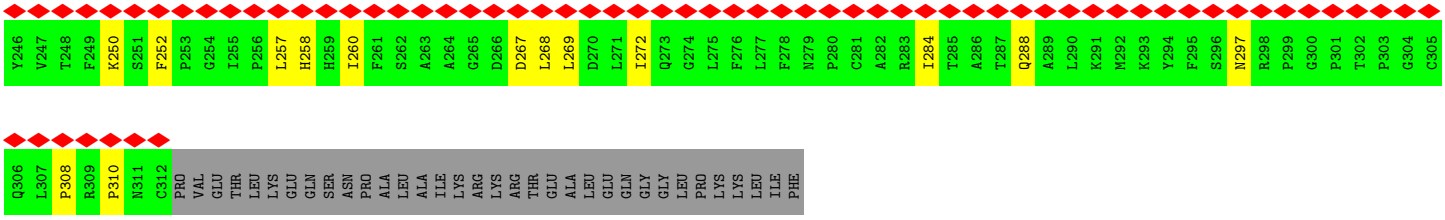




- Molecule 9: Cyclin-dependent kinase 7

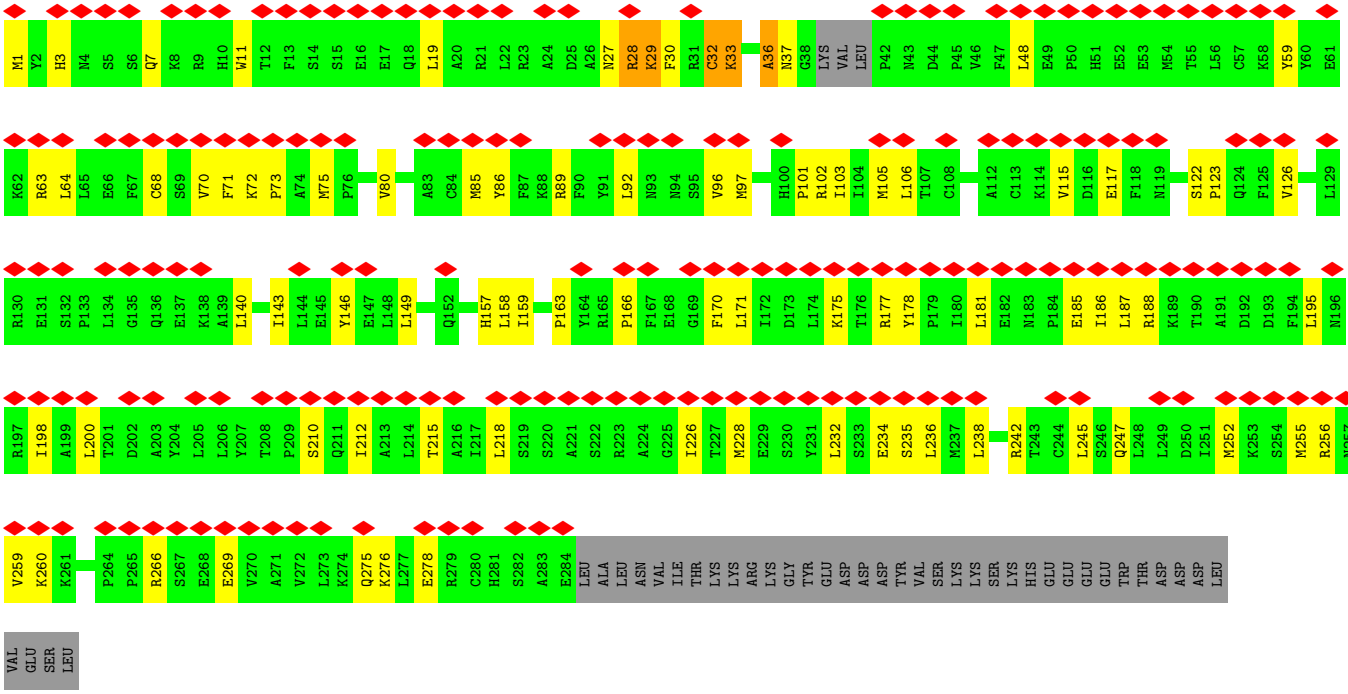
Chain 8:





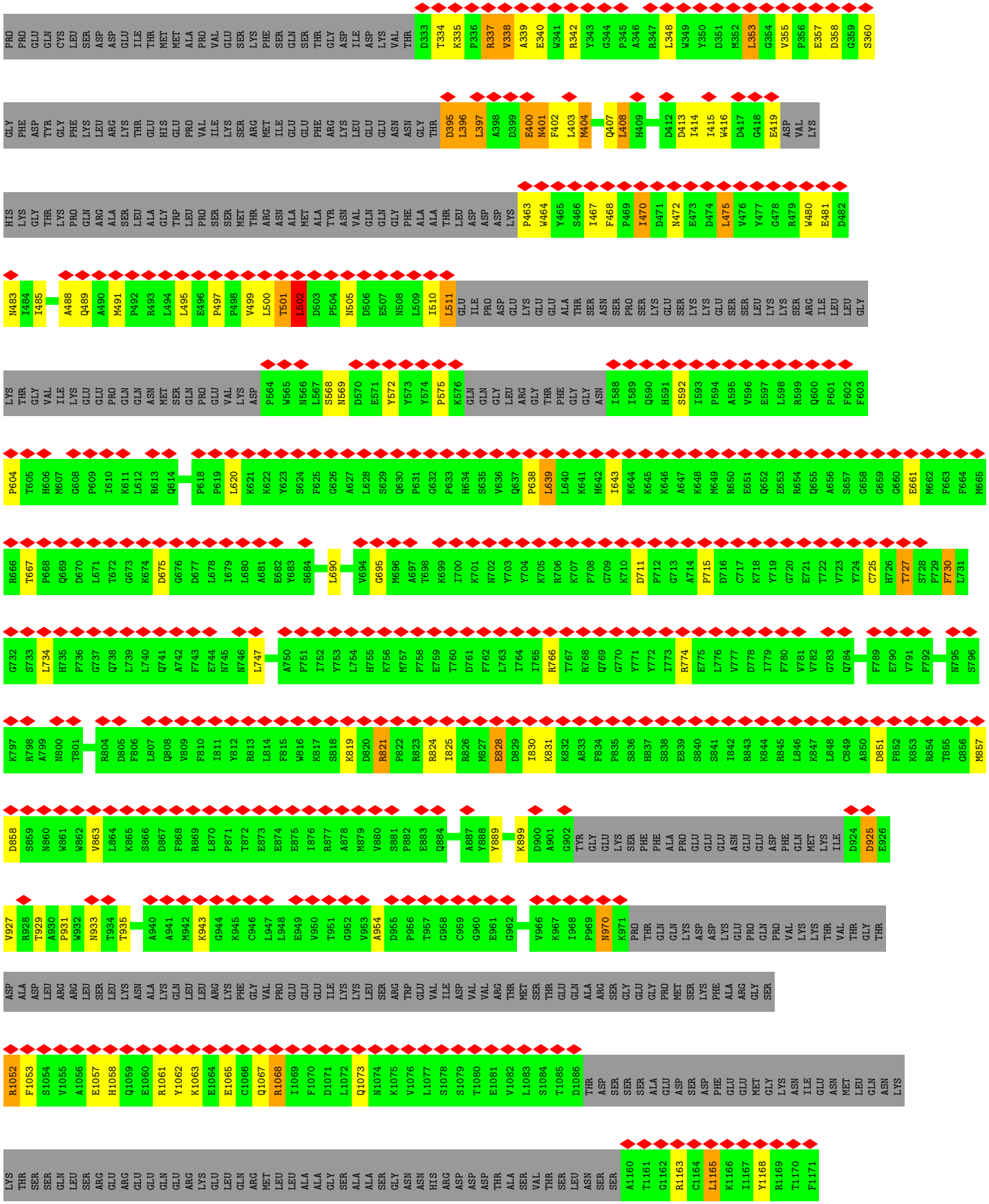
• Molecule 10: Cyclin-H

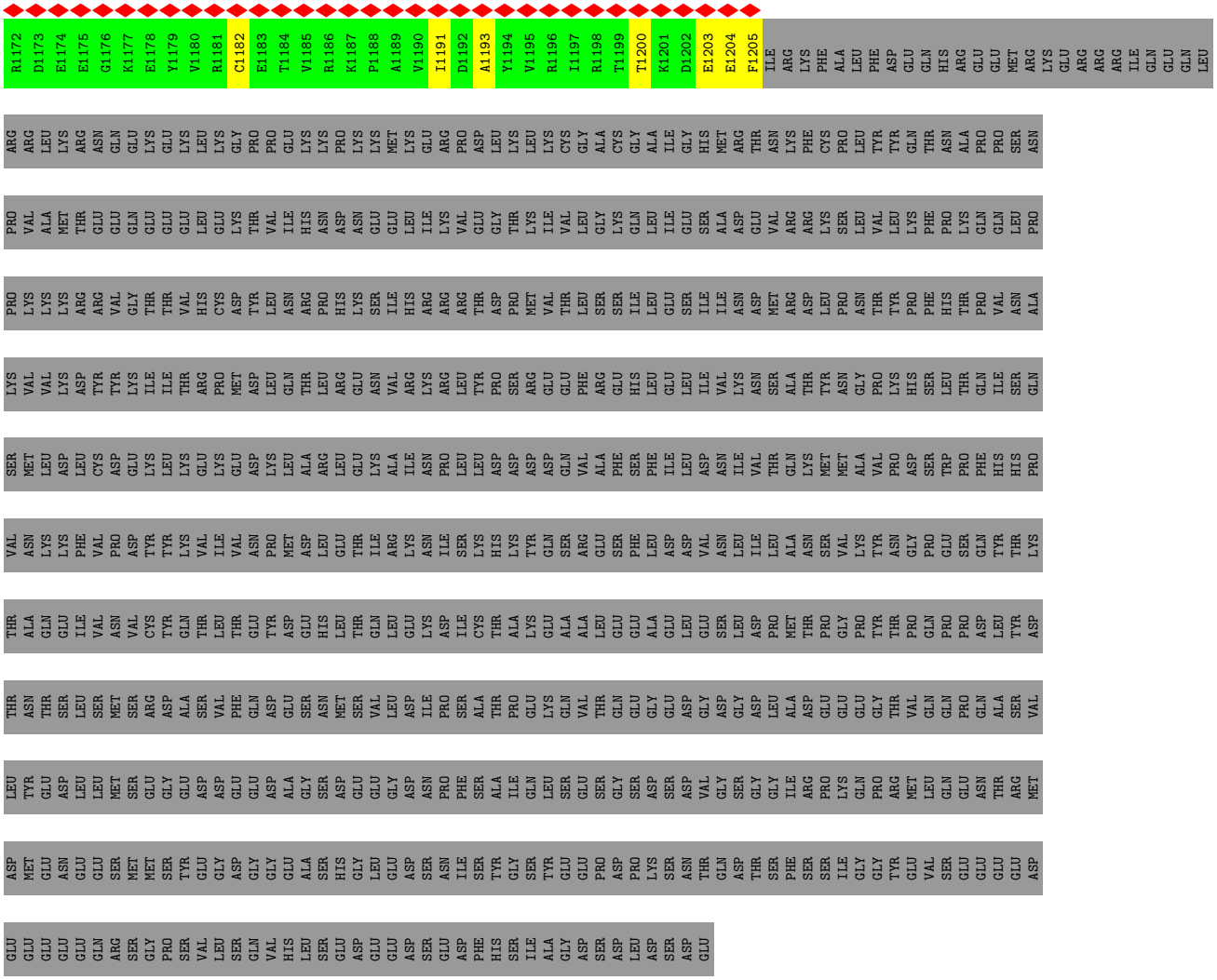
Chain 9:



• Molecule 11: Transcription initiation factor TFIID subunit 1

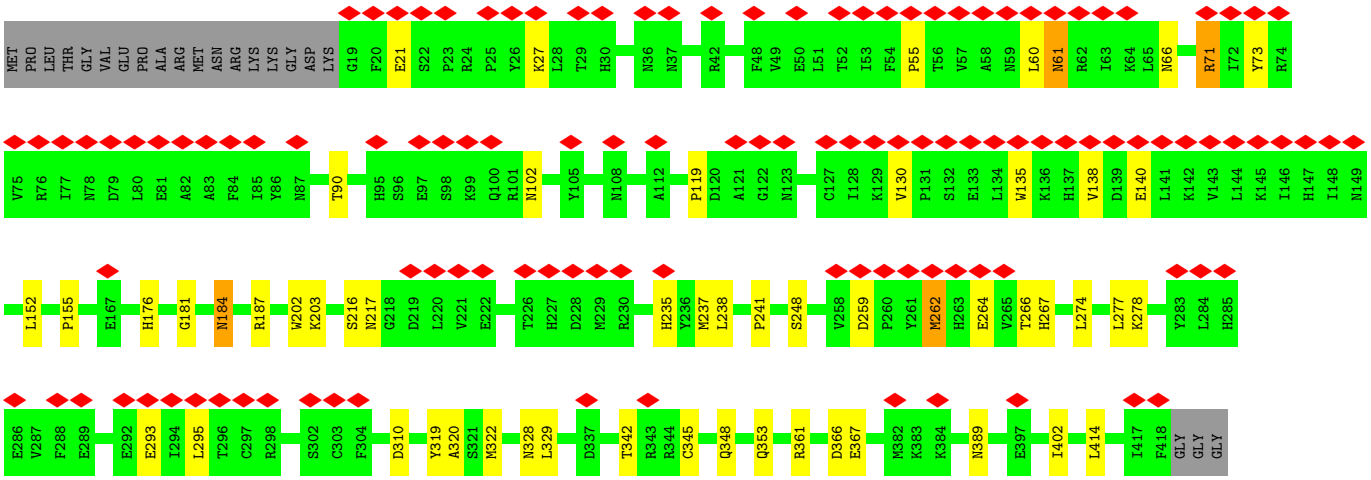
Chain A:

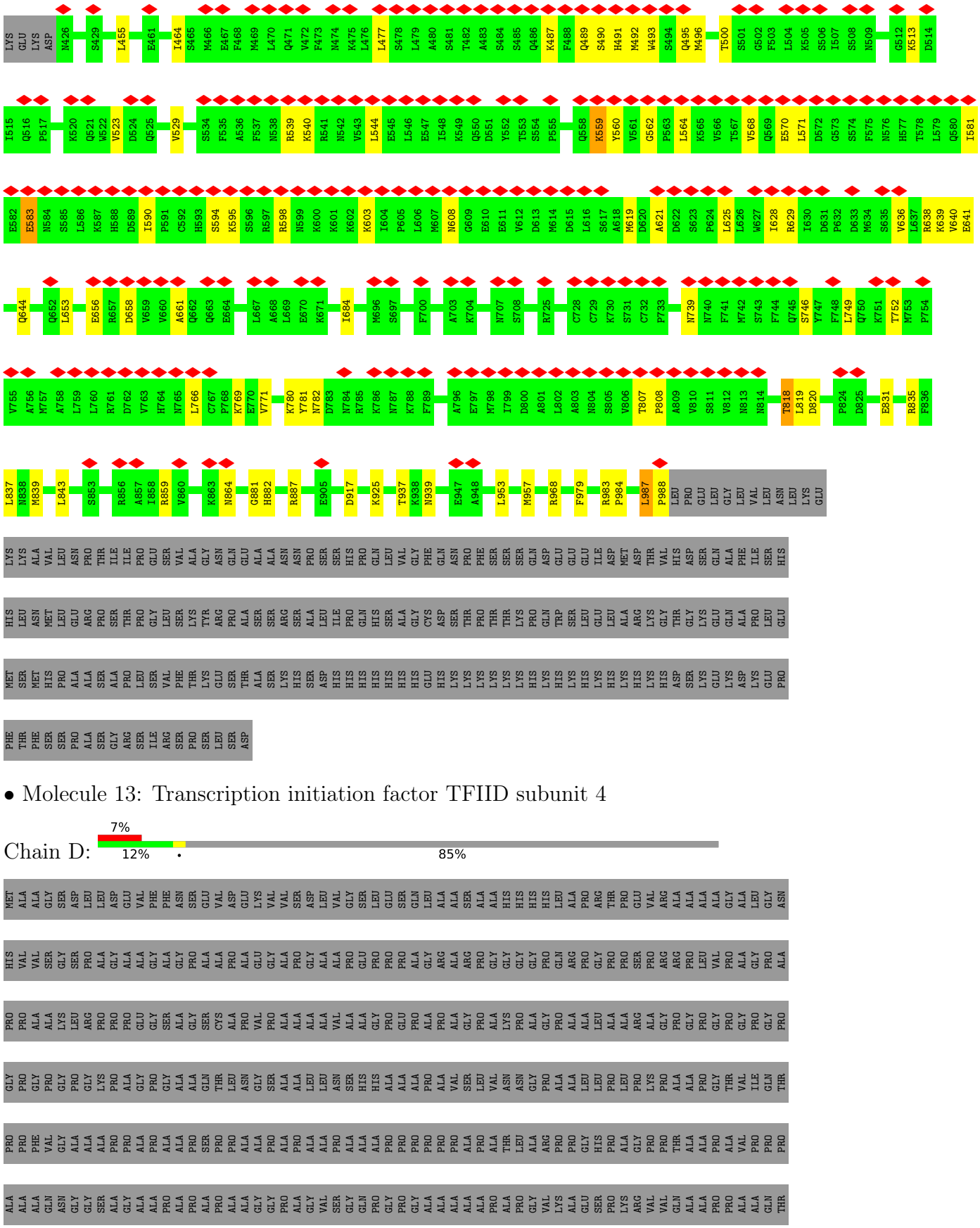

PRO PRO GLU GLN CYS LEU SER ASP ASP GLU ILE THR MET MET ALA PRO VAL GLU SER LYS PHE SER GLN SER THR GLY ASP ILE ASP LYS VAL THR D333 T334 K335 P336 R337 V338 A339 E340 W341 R342 Y343 G344 P345 A346 R347 L348 W349 Y350 D351 M352 L353 G354 V355 P356 E357 D358 G359 S360
GLY PHE ASP TYR GLY PHE LYS LEU ARG LYS THR GLU HIS GLU PRO VAL ILE LYS SER ARG MET ILE GLU GLU GLU PHE ARG LYS LEU GLU GLU ASN ASN GLY THR D395 L396 L397 A398 D399 E400 N401 F402 L403 M404 Q407 L408 H409 D412 D413 I414 I415 W416 D417 G418 E419 ASP VAL LYS
HIS LYS GLY THR LYS PRO GLN ARG ALA SER LEU ALA GLY TRP LEU PRO SER SER MET THR ARG ASN ALA MET ALA TYR ASN VAL GLN GLN GLY PHE ALA ALA THR LEU ASP ASP ASP LYS P463 W464 Y465 S466 I467 F468 P469 I470 D471 N472 E473 D474 L475 V476 Y477 G478 R479 W480 E481 D482
N483 I484 I485 A488 Q489 A490 M491 P492 R493 L494 L495 E496 P497 P498 V499 L500 T501 L502 D503 P504 N505 D506 E507 N508 L509 I510 L511 GLU ILE PRO ASP GLU LYS GLU GLU ALA THR SER ASN SER PRO SER LYS GLU SER SER LYS LYS GLU SER SER LEU LYS LYS SER ARG ILE LEU LEU GLY
LYS THR GLY VAL ILE LYS GLU GLU PRO GLN GLN ASN MET SER GLN PRO GLU VAL LYS ASP P564 W565 N566 L567 S568 N569 D570 E571 Y572 Y573 Y574 P575 K576 GLN GLN GLY LEU ARG GLY THR PHE GLY GLY ASN I588 I589 Q590 H591 S592 I593 P594 A595 V596 E597 L598 R599 Q600 P601 F602 F603
P604 T605 H606 M607 G608 P609 I610 K611 L612 R613 Q614 F618 P619 L620 K621 K622 Y623 S624 F625 G626 A627 L628 S629 Q630 P631 G632 P633 H634 S635 V636 Q637 P638 L639 L640 K641 H642 I643 K644 K645 K646 A647 K648 M649 R650 E651 Q652 E653 R654 Q655 A656 S657 G658 G659 G660 E661 M662 F663 F664 M665
R666 T667 P668 Q669 D670 L671 T672 G673 K674 D675 G676 D677 L678 I679 L680 A681 E682 Y683 S684 L690 V694 G695 M696 A697 T698 K699 I700 K701 N702 Y703 Y704 K705 R706 K707 P708 G709 K710 D711 P712 G713 A714 P715 D716 C717 K718 Y719 G720 E721 T722 V723 Y724 G725 H726 T727 S728 P729 F730 L731
G732 S733 L734 H735 P736 G737 Q738 L739 L740 Q741 A742 F743 E744 N745 N746 L747 A750 P751 I752 Y753 L754 H755 K756 M757 P758 E759 T760 D761 F762 L763 I764 I765 R766 T767 R768 Q769 G770 Y771 Y772 I773 R774 E775 L776 V777 D778 I779 F780 V781 V782 G783 Q784 F789 E790 V791 F792 N795 S796
K797 R798 A799 N800 T801 R804 D805 F806 L807 Q808 V809 F810 I811 Y812 R813 L814 F815 W816 K817 S818 K819 D820 R821 P822 R823 R824 I825 R826 M827 E828 D829 I830 K831 K832 A833 F834 P835 S836 H837 S838 E839 S840 S841 I842 R843 K844 R845 L846 K847 L848 C849 A850 D851 F852 K853 R854 T855 G856 M857
D858 S859 N860 W861 W862 V863 L864 K865 S866 D867 F868 R869 L870 P871 T872 E873 E874 E875 I876 R877 A878 M879 V880 S881 P882 E883 Q884 A887 Y888 Y889 K899 D900 A901 G902 TYR GLY GLU LYS SER PHE PHE ALA PRO GLU GLU GLU ASN GLU GLU ASP PHE GLN MET LYS ILE D924 D925 E926
V927 R928 T929 A930 P931 W932 N933 T934 T935 A940 A941 M942 K943 G944 K945 C946 L947 L948 E949 V950 T951 G952 V953 A954 D955 P956 T957 G958 C959 G960 E961 G962 V966 K967 I968 P969 N970 K971 PRO THR GLN GLN GLN LYS ASP ASP LYS GLU GLU PRO GLN PRO VAL LYS LYS THR VAL THR GLY THR
ASP ALA ASP LEU ARG ARG LEU SER LEU LYS ASN ALA LYS GLN LEU LEU ARG LYS PHE GLY VAL PRO GLU GLU GLU ILE LYS LYS LEU SER ARG TRP GLU VAL ILE ASP VAL VAL ARG THR MET SER THR GLU GLN ALA ARG SER GLY GLU GLY PRO MET SER LYS PHE ALA ARG GLY SER
R1052 F1053 S1054 V1055 A1056 E1057 H1058 Q1059 E1060 R1061 Y1062 K1063 E1064 E1065 C1066 Q1067 R1068 I1069 F1070 D1071 L1072 Q1073 N1074 K1075 V1076 L1077 S1078 S1079 T1080 E1081 V1082 L1083 S1084 T1085 D1086 THR ASP SER SER SER ALA GLU ASP SER ASP PHE GLU GLU MET GLY LYS ASN ILE GLU ASN MET LEU GLN ASN LYS
LYS THR SER SER GLN LEU SER ARG GLU ARG GLU GLU GLN GLU ARG LYS GLU LEU GLN ARG MET LEU LEU ALA ALA GLY SER ALA ALA SER GLY ASN ASN HIS ARG ASP ASP THR ALA SER VAL THR SER LEU ASN SER A1160 T1161 G1162 R1163 C1164 L1165 K1166 I1167 Y1168 R1169 T1170 F1171



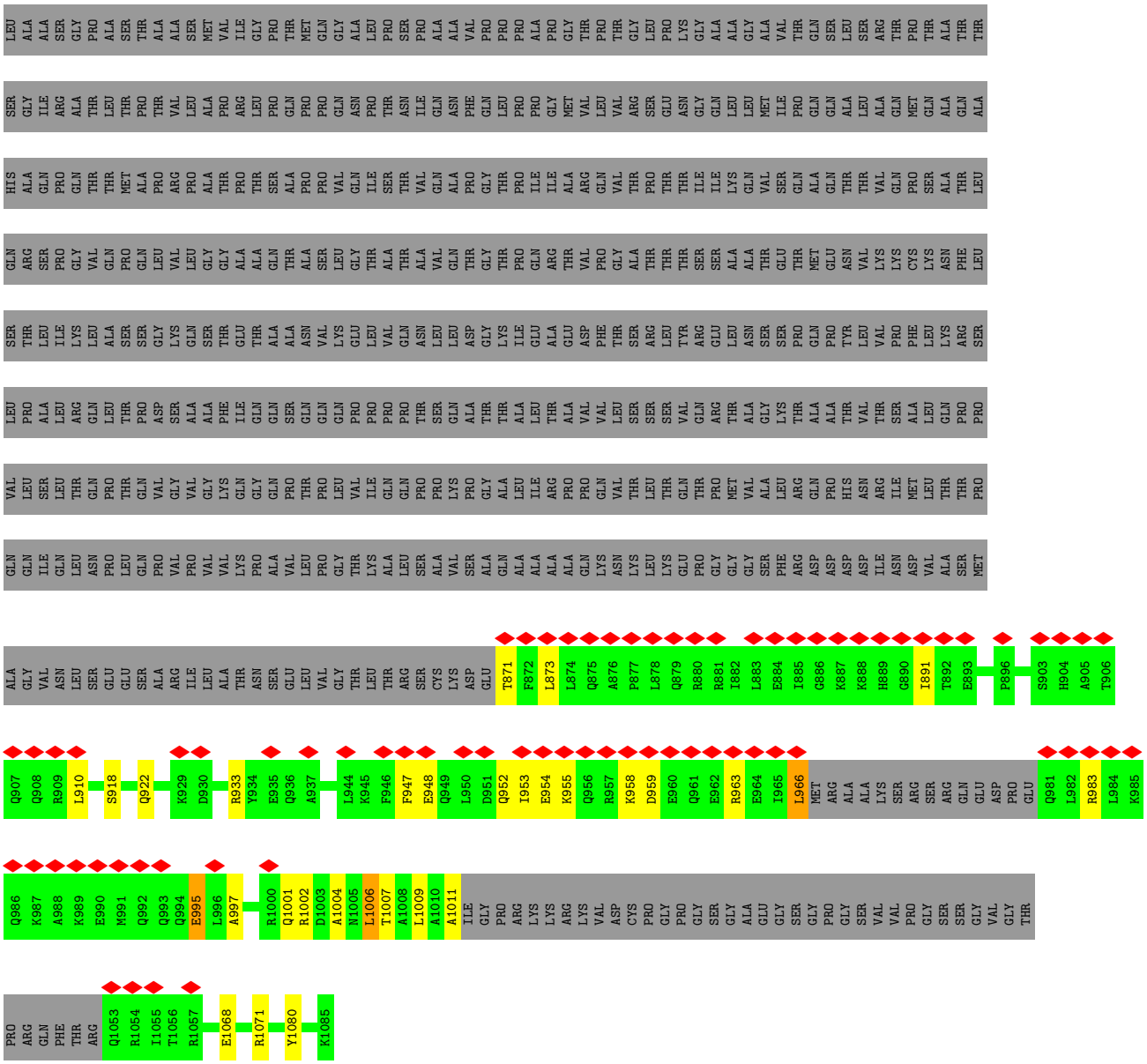
• Molecule 12: Transcription initiation factor TFIID subunit 2

Chain B:





• Molecule 13: Transcription initiation factor TFIID subunit 4



- Molecule 13: Transcription initiation factor TFIID subunit 4

Chain d:

ALA ALA ALA GLN ASN GLY GLY SER ALA GLY ALA ALA PRO ALA PRO ALA ALA GLY GLY PRO ALA ALA GLY VAL SER GLY GLN PRO GLY PRO GLY ALA ALA ALA ALA PRO ALA PRO GLY VAL LYS ALA GLU SER PRO LYS ARG VAL VAL GLN ALA ALA PRO ALA ALA THR

LEU ALA ALA SER GLY PRO ALA SER THR ALA ALA SER MET VAL ILE GLY PRO THR MET GLN GLY ALA LEU PRO SER PRO ALA ALA VAL PRO PRO PRO ALA PRO GLY THR PRO THR GLY LEU PRO LYS GLY ALA ALA GLY ALA VAL THR GLN SER LEU SER ARG THR PRO THR ALA THR THR

SER GLY ILE ARG ALA THR LEU THR PRO THR VAL LEU ALA PRO ARG LEU PRO GLN PRO PRO GLN ASN PRO THR ASN ILE GLN ASN PHE GLN LEU PRO PRO PRO GLY GLY MET VAL LEU VAL ARG SER GLU ASN GLY GLN LEU LEU MET ILE PRO GLN GLN GLN ALA LEU ALA GLN MET GLN ALA GLN

HIS ALA GLN PRO GLN THR THR MET ALA PRO ARG PRO ALA ALA THR PRO THR SER ALA PRO PRO VAL GLN ILE SER THR VAL GLN ALA PRO GLY THR PRO ILE ILE ALA ARG GLN VAL THR PRO THR THR ILE ILE LYS GLN VAL SER GLN ALA GLN THR THR VAL GLN PRO SER ALA THR LEU

GLN ARG SER PRO GLY VAL GLN PRO GLN LEU VAL LEU GLY GLY GLY ALA ALA GLN THR ALA SER LEU GLY THR ALA THR ALA VAL GLN THR GLY THR PRO GLN ARG THR VAL PRO GLY ALA THR THR THR SER SER ALA ALA THR GLU THR MET GLU ASN VAL LYS LYS CYS LYS ASN PHE LEU

SER THR LEU ILE LYS LEU ALA SER SER GLY LYS GLN SER THR GLU THR ALA ALA ASN VAL LYS GLU LEU VAL GLN ASN LEU LEU ASP GLY LYS ILE GLU ALA GLU ASP PHE THR SER ARG LEU TYR ARG GLU LEU ASN SER SER PRO GLN PRO TYR LEU VAL PRO PHE LEU LYS ARG SER

LEU PRO ALA LEU ARG GLN LEU THR PRO ASP SER ALA ALA PHE ILE GLN GLN SER GLN GLN GLN PRO PRO PRO THR SER GLN ALA THR THR ALA LEU THR ALA VAL VAL LEU SER SER SER VAL GLN ARG THR ALA ALA ALA THR VAL THR SER ALA LEU GLN PRO PRO

VAL LEU SER LEU THR GLN PRO THR GLN VAL GLY VAL GLY LYS GLN GLY GLN PRO THR PRO LEU VAL ILE GLN GLN PRO PRO LYS PRO GLY ALA LEU ILE ARG PRO PRO GLN VAL THR LEU THR GLN THR PRO MET VAL ALA LEU LEU VAL ARG ILE MET LEU THR THR PRO

GLN GLN ILE GLN LEU ASN PRO LEU GLN PRO VAL PRO VAL VAL LYS PRO ALA VAL LEU PRO GLY THR LYS ALA LEU SER ALA VAL SER ALA GLN ALA ALA GLN LYS ASN LYS LEU LYS GLU PRO GLY GLY SER PHE ARG ASP ASP ASP ILE ASN ASP VAL ALA SER MET

ALA GLY V843 N844 L845 S846 E847 E848 S849 A850 R851 I852 L853 A854 T855 N856 S857 E858 L859 V860 G861 T862 L863 T864 R865 S866 C867 K868 D869 E870 T871 F872 L873 L874 Q875 A876 P877 L878 Q879 R880 R881 I882 L883 E884 I885 G886 K887 K888 H889 G890 I891 T892 E893 L894 H895 P896 D897 V898 V899 S900

Y901 V902 S903 H904 A905 T906 Q907 Q908 R909 L910 Q911 N912 L913 V914 E915 K916 I917 S918 E919 T920 A921 Q922 Q923 K924 N925 F926 S927 Y928 K929 D930 D931 D932 R933 Y934 E935 Q936 A937 S938 D939 V940 R941 A942 Q943 L944 K945 F946 F947 E948 Q949 L950 D951 Q952 I953 E954 K955 Q956 R957 K958 D959 E960

Q961 E962 R963 E964 I965 L966 M967 R968 ALA ALA LYS SER ARG SER ARG GLN GLU ASP PRO GLU GLN GLN LEU ARG LEU LYS GLN LYS ALA LYS GLU MET GLN GLN GLN GLU LEU ALA GLN MET ARG GLN ARG ASP ALA ASN LEU THR ALA LEU ALA ALA ILE GLY PRO ARG LYS LYS ARG LYS VAL

ASP CYS PRO GLY PRO GLY SER GLY ALA GLU GLY SER GLY PRO GLY SER VAL VAL PRO GLY SER SER GLY VAL GLY THR PRO ARG GLN PHE THR ARG GLN R1054 I1055 T1056 R1057 V1058 N1059 L1060 R1061 D1062 L1063 I1064 F1065 C1066 L1067 E1068 N1069 E1070 R1071 E1072 T1073 S1074 H1075 S1076 L1077 L1078 L1079 Y1080

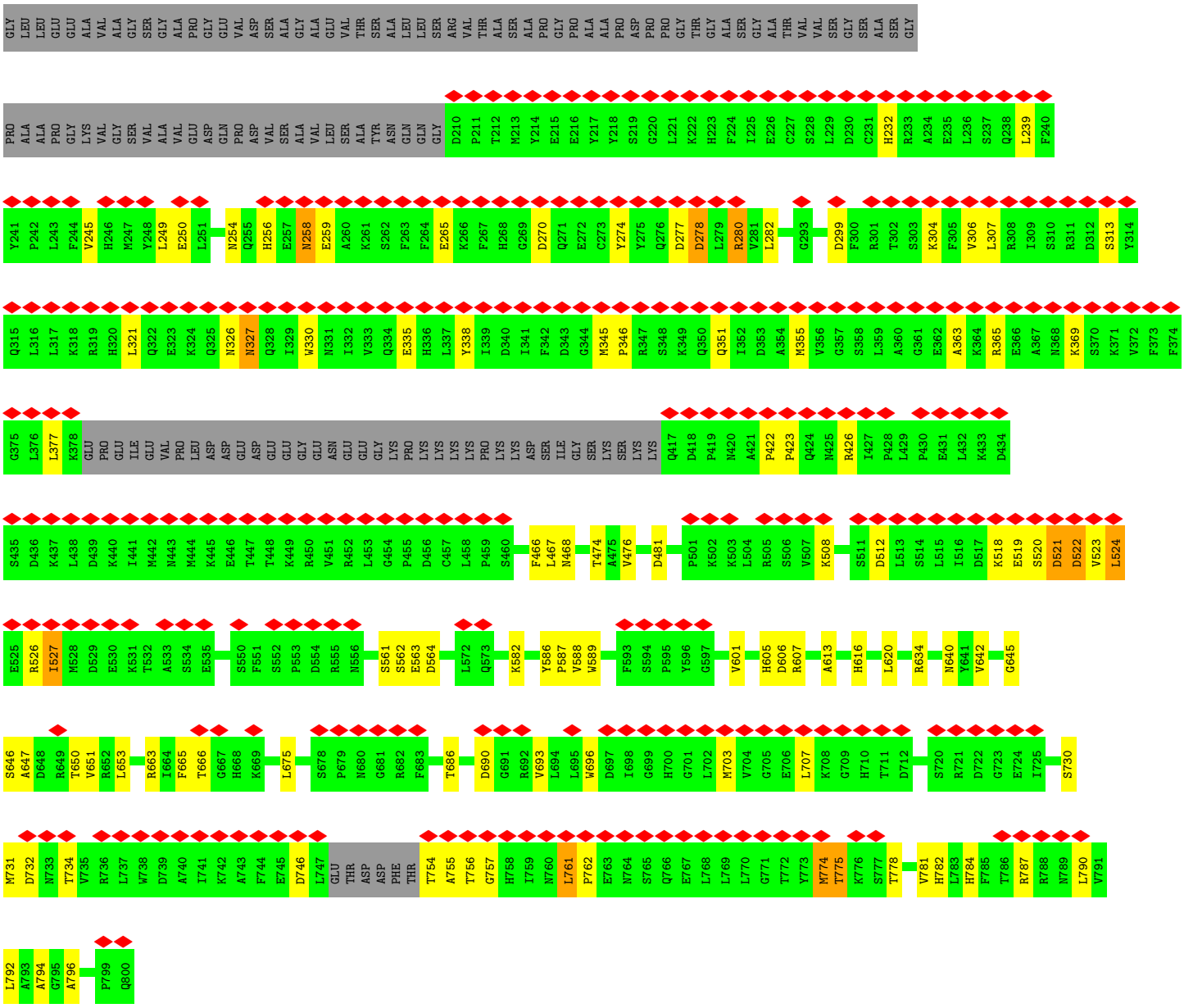
K1081 A1082 F1083 L1084 K1085

- Molecule 14: Transcription initiation factor TFIID subunit 5

Chain E:

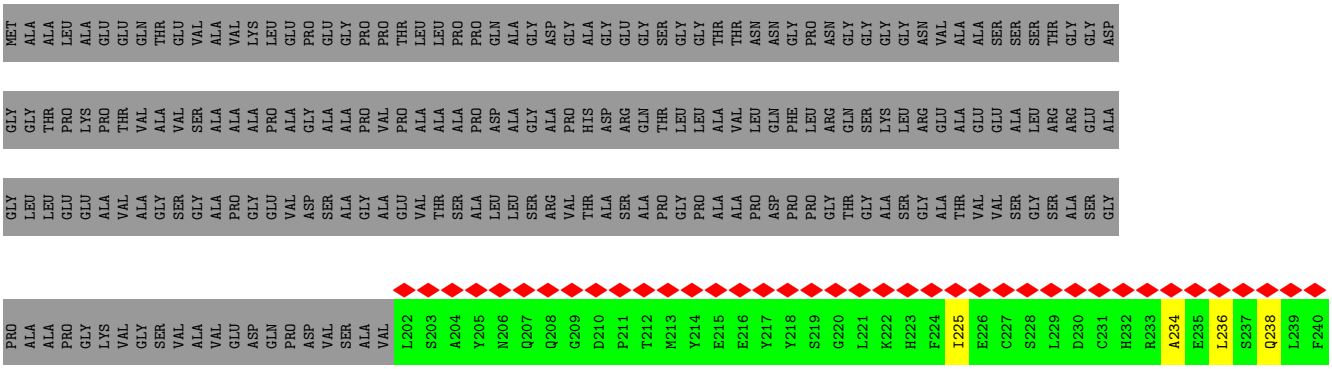
MET ALA ALA LEU ALA GLU GLU GLN THR GLU VAL ALA VAL LYS LEU GLU PRO GLU GLY PRO PRO THR LEU LEU PRO PRO GLN ALA GLY ASP GLY ALA GLY GLU GLY SER GLY GLY THR THR ASN ASN GLY PRO ASN GLY GLY GLY GLY ASN VAL ALA ALA SER SER SER THR GLY GLY

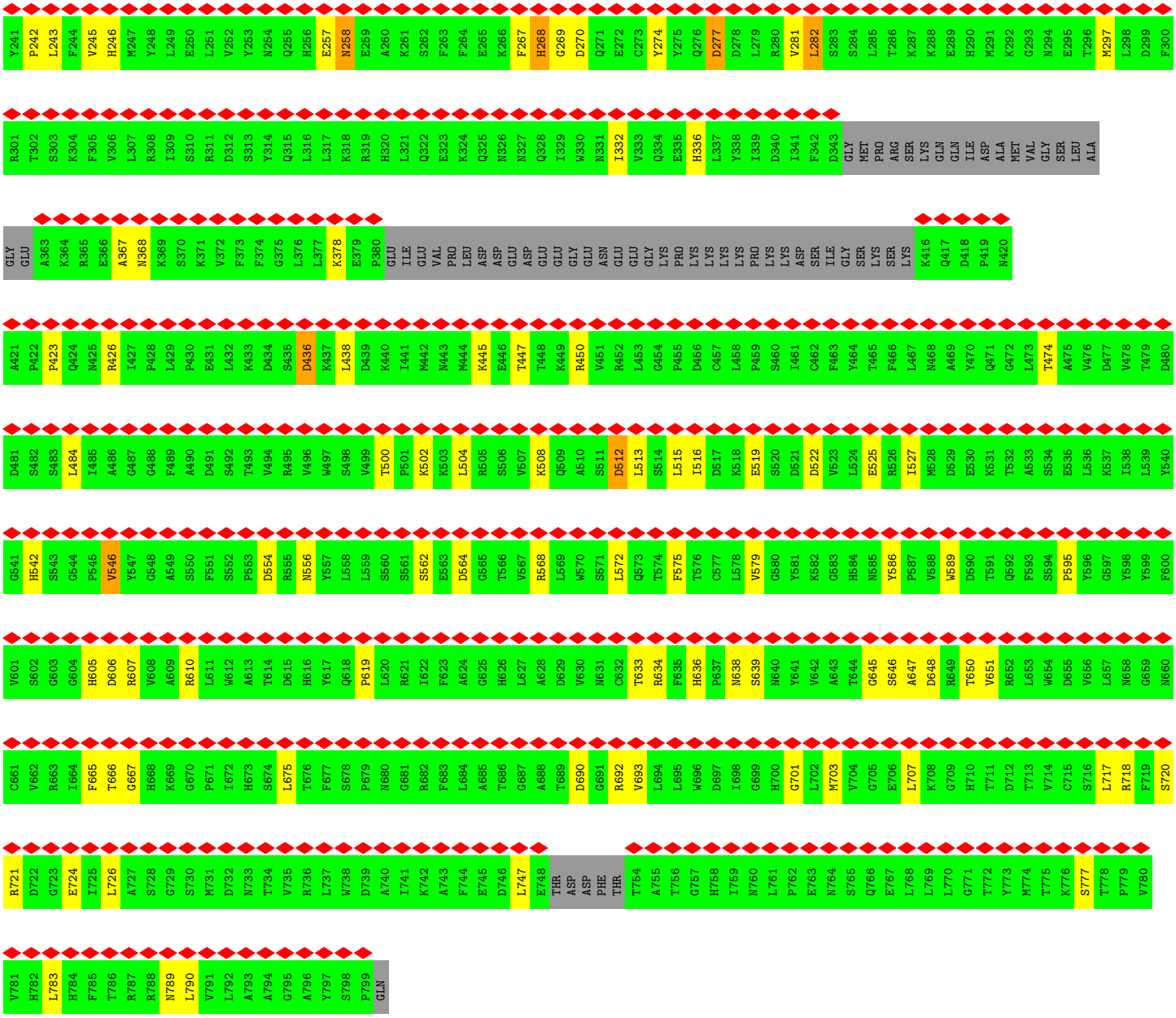
GLY GLY THR PRO LYS PRO THR VAL ALA VAL SER ALA ALA ALA PRO ALA GLY ALA ALA PRO VAL PRO ALA ALA PRO ASP ALA GLY ALA PRO HIS ASP ARG GLN THR LEU LEU ALA VAL LEU GLN PHE LEU ARG GLN SER LYS LEU ARG GLU ALA GLU GLU ALA LEU ARG ARG GLU ALA



• Molecule 14: Transcription initiation factor TFIID subunit 5

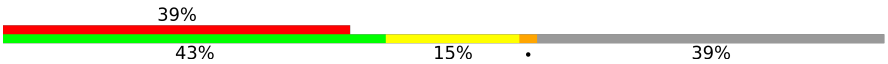
Chain e:

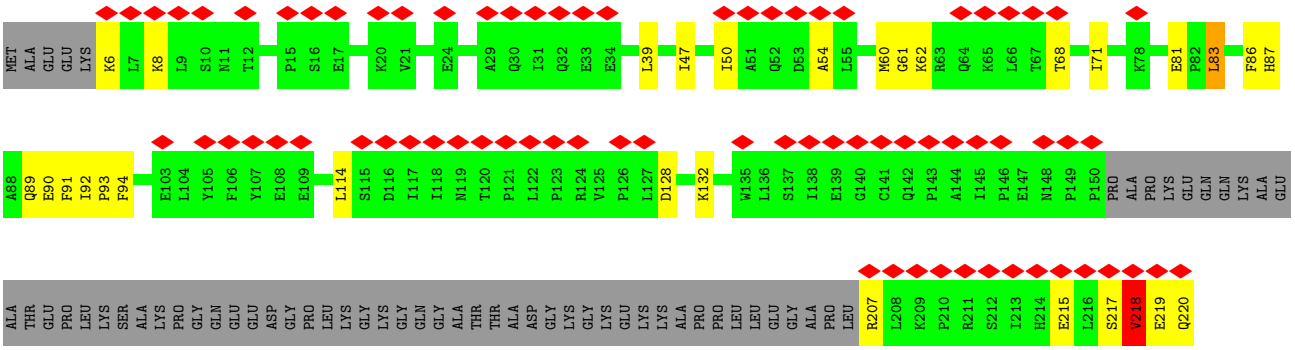




- Molecule 15: Transcription initiation factor TFIID subunit 6

Chain F:





Q221 L222 Y223 Y224 K225 E226 I227 T228 E229 A230 C231 V232 G233 S234 C235 E236 A237 K238 R239 A240 E241 A242 L243 Q244 S245 I246 A247 T248 D249 P250 G251 L252 Y253 Q254 M255 L256 P257 R258 F259 S260 T261 F262 I263 S264 E265 G266 V267 R268 V269 N270 V271 V272 Q273 N274 N275 L276 A277 L278 L279 I280

Y281 L282 M283 R284 M285 V286 K287 A288 L289 M290 D291 N292 P293 T294 L295 Y296 L297 E298 K299 Y300 V301 H302 E303 L304 A307 V308 M309 T310 C311 I312 V313 S314 R315 Q316 L317 C318 L319 R320 P321 D322 V323 D324 N325 H326 W327 A328 L329 R330 D331 F332 A333 A334 R335 L336 V337 I340 H343

F344 S345 T346 T347 T348 N349 N350 R354 F359 T360 K361 V364 D365 E366 K367 T368 P369 W370 T371 A378 G379 L380 A381 E382 L383 G384 H385 D386 V387 I388 K389 T390 L391 I392 L393 P394 R395 L396 Q397 Q398 E399 G400 E401 R402 I403 R404 S405 V406 L407 D408 G409 P410 V411 L412 S413 N414

I415 D416 R417 I418 G419 A420 Q424 S425 L426 L427 L428 K429 A432 P433 V434 L435 A436 K437 L438 R439 P440 P441 P442 D443 N444 Q445 D446 A447 Y448 R449 A450 E451 F452 G453 G456 P457 L458 L459 C460 S461 Q462 V463 V464 K465 A466 R467 A468 Q469 A470 A471 L472 Q473 ALA GLN GLN VAL ASN

ARG THR THR LEU THR ILE THR GLN PRO ARG PRO THR LEU THR LEU SER GLN ALA PRO GLN PRO GLY PRO ARG THR PRO GLY LEU LEU LYS VAL PRO GLY SER ILE ILE ALA LEU PRO VAL GLN THR LEU VAL SER ALA ARG ALA ALA ALA PRO PRO GLN PRO SER PRO PRO PRO THR LYS PHE

ILE VAL MET SER SER SER SER SER ALA PRO SER THR GLN GLN VAL LEU SER LEU SER THR SER ALA PRO GLY SER GLY SER THR THR THR SER PRO VAL THR THR THR VAL PRO SER VAL GLN PRO ILE VAL LYS LEU VAL SER THR ALA THR THR ALA PRO PRO SER THR ALA PRO SER

GLY PRO GLY SER VAL GLN LYS TYR ILE VAL VAL SER LEU PRO PRO THR GLY GLU GLY LYS GLY GLY PRO THR SER HIS PRO SER PRO VAL PRO PRO PRO ALA SER SER PRO SER PRO LEU SER GLY SER ALA LEU CYS GLY GLY LYS GLN GLU ALA GLY ASP SER PRO PRO PRO ALA PRO

GLY THR PRO LYS ALA ASN GLY SER GLN PRO ASN SER GLY SER PRO GLN PRO ALA PRO

- Molecule 15: Transcription initiation factor TFIID subunit 6

Chain f:

MET ALA GLU GLU LYS LYS LEU LYS L9 S10 N11 T12 V13 L14 P15 S16 E17 S18 M19 K20 V21 V22 A23 E24 S25 M26 G27 I28 A29 Q30 I31 Q32 E33 E34 T35 C36 Q37 L38 L39 T40 D41 E42 V43 S44 Y45 R46 I47 K48 E49 I50 A51 Q52 D53 A54 L55 K56 F57 M58 H59 M60

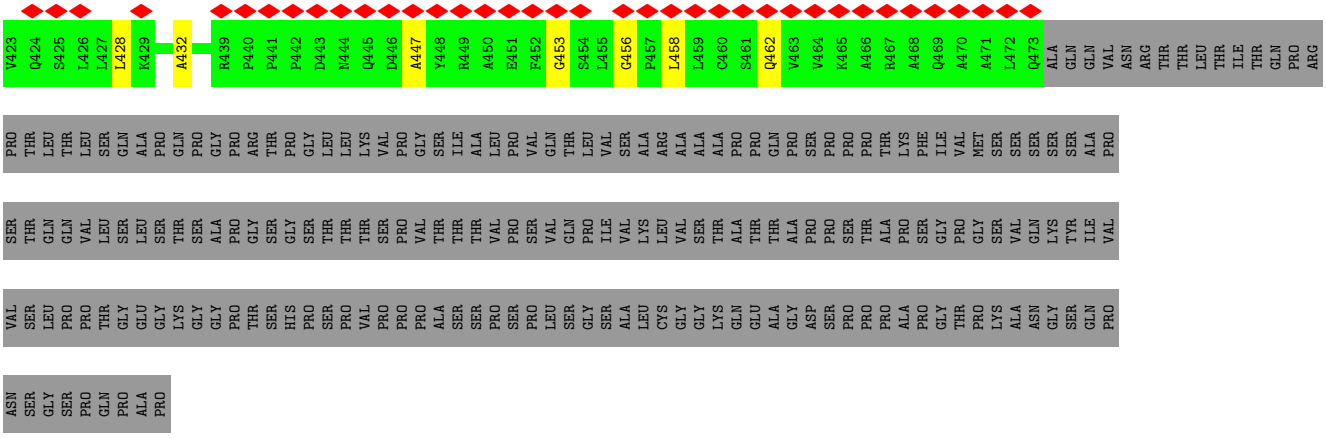
G61 K62 R63 Q64 K65 L66 T67 T68 S69 D70 I71 D72 Y73 A74 L75 K76 L77 K78 N79 V80 E81 P82 L83 Y84 G85 F86 H87 A88 Q89 E90 F91 I92 P93 F94 R95 F96 A97 S98 G99 G100 G101 R102 E103 L104 Y105 F106 Y107 E108 E109 K110 E111 V112 D113 L114 S115 D116 I117 I118 N119 T120

P121 L122 P123 R124 V125 P126 L127 D128 V129 C130 L131 K132 A133 H134 W135 L136 S137 I138 E139 G140 C141 Q142 P143 A144 I145 P146 E147 N148 P149 P150 P151 ALA PRO LYS GLU GLN GLN LYS ALA GLU ALA THR GLU PRO LEU LYS SER ALA LYS PRO GLY GLN GLU GLU ASP GLY PRO LEU LYS GLY

LYS GLY GLN GLY ALA THR THR ALA ASP GLY LYS GLY LYS GLU LYS LYS ALA PRO PRO LEU LEU GLU GLY ALA PRO LEU ARG LEU LYS PRO ARG SER ILE H214 E215 L216 S217 V218 E219 Q220 Y223 Y224 K225 E226 E229 A230 C231 V232 G233 S234 C235 E236 A237 K238 R239 A240 E241 A242

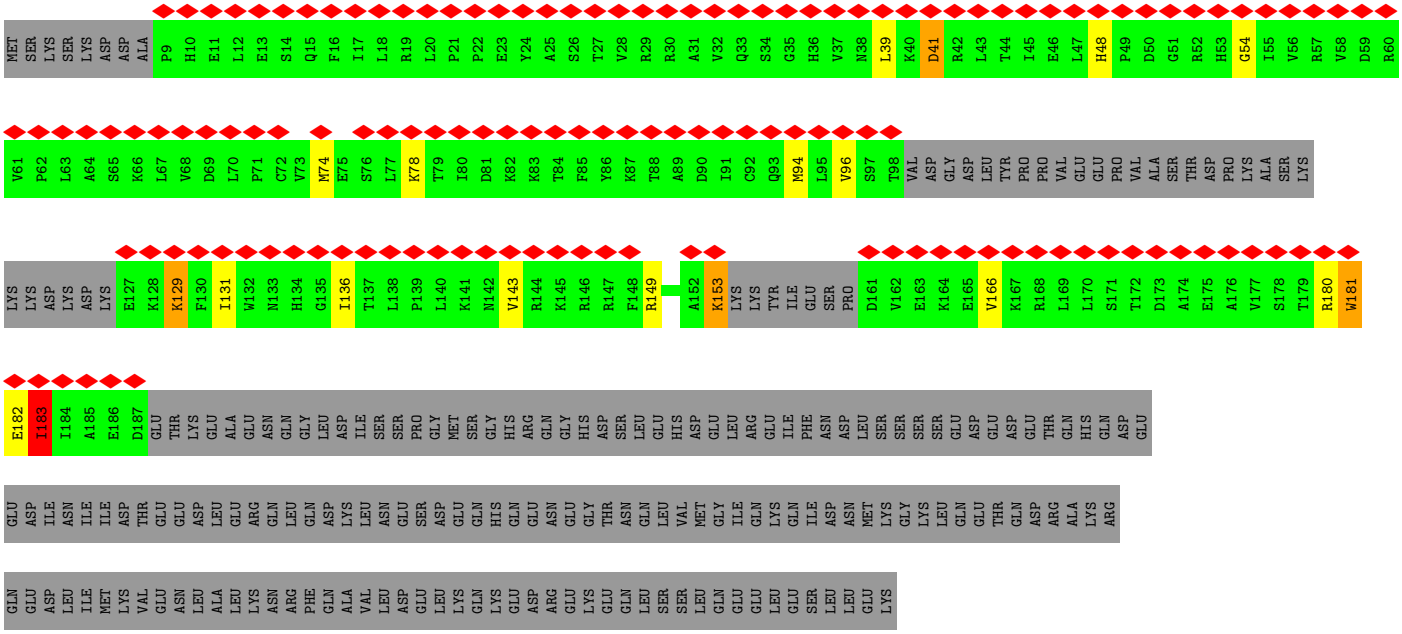
Y253 Q254 M255 L256 P257 R258 T261 R268 V269 N270 V271 V272 Q273 N274 N275 L276 A277 L278 L279 I280 Y281 L282 M283 R284 E298 K299 Y300 V301 H302 E303 L304 I305 V308 M309 T310 C311 R315 Q316 L317 C318 L319 R320 P321 D322 V323 D324 N325 H326 W327 A328 L329 R330 D331 F332

A333 L336 T346 T347 T348 N349 N350 I351 Q352 S353 R354 I355 T356 K357 K361 V364 D365 E366 K367 T368 P369 W370 T371 T390 L391 P394 R395 L396 Q397 Q398 E399 G400 E401 R402 I403 R404 S405 V406 L407 D408 G409 P410 V411 L412 S413 N414 I415 D416 R417 I418 G419 A420 D421 H422



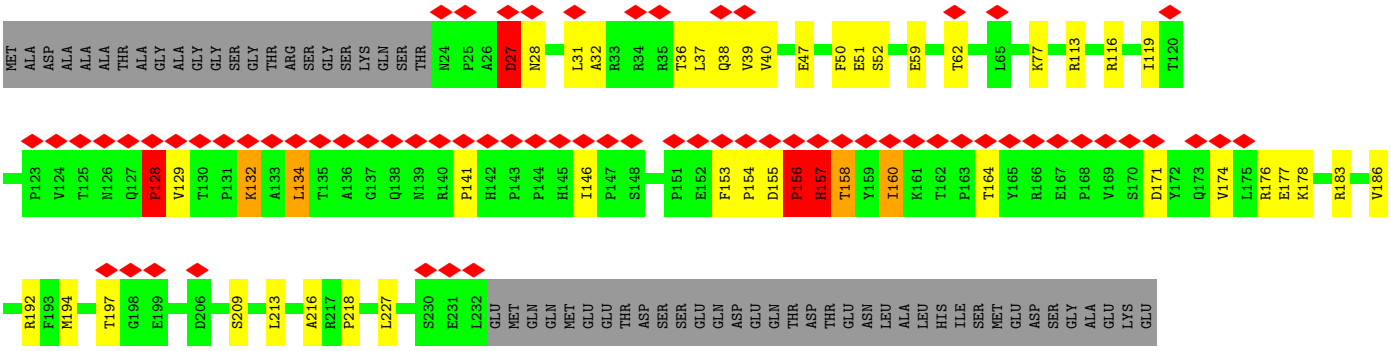
• Molecule 16: Transcription initiation factor TFIID subunit 7

Chain G:

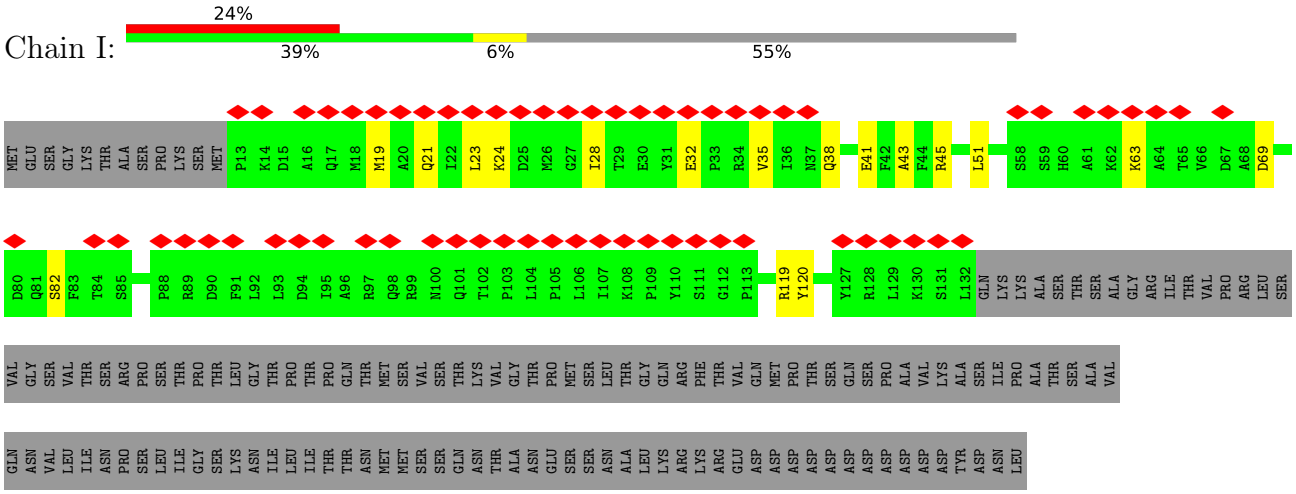


• Molecule 17: Transcription initiation factor TFIID subunit 8

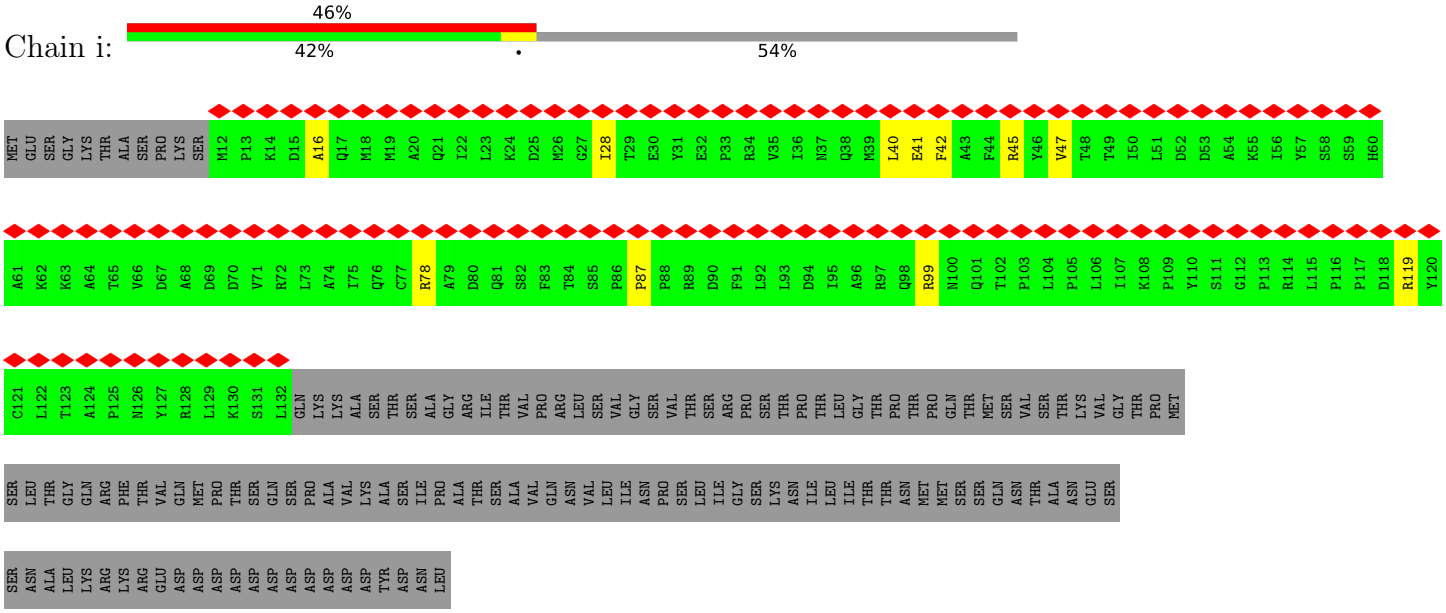
Chain H:



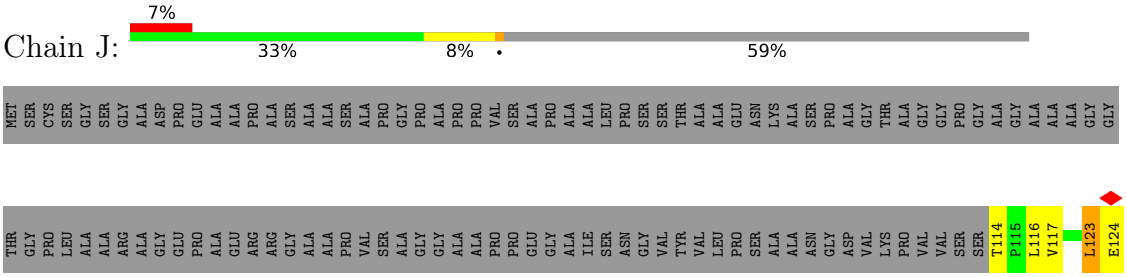
- Molecule 18: Transcription initiation factor TFIID subunit 9

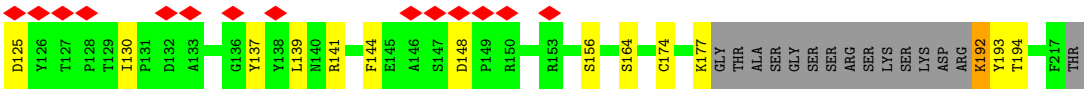


- Molecule 18: Transcription initiation factor TFIID subunit 9



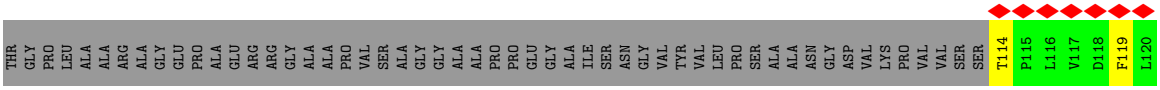
- Molecule 19: Transcription initiation factor TFIID subunit 10

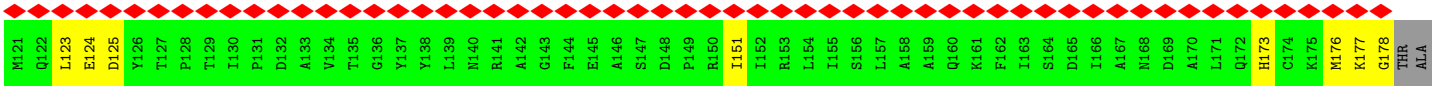


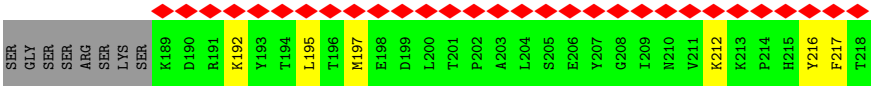


• Molecule 19: Transcription initiation factor TFIID subunit 10

Chain j:

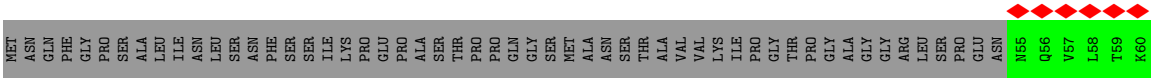


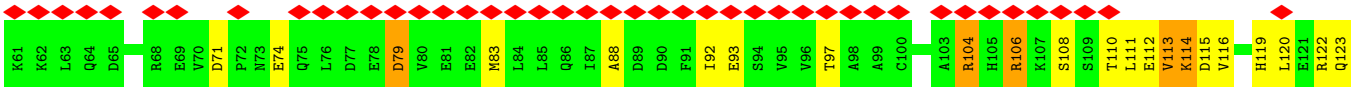


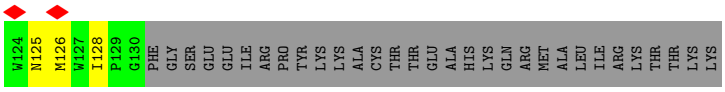


• Molecule 20: Transcription initiation factor TFIID subunit 12

Chain L:

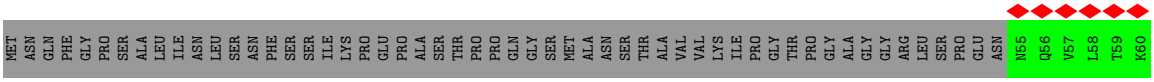


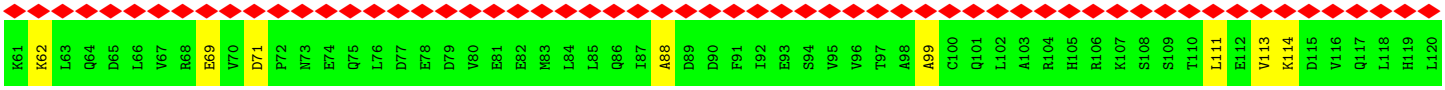


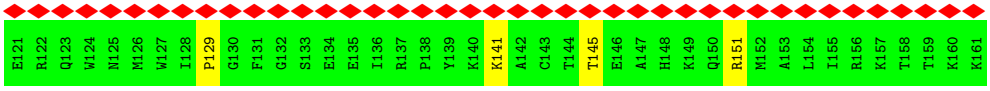


• Molecule 20: Transcription initiation factor TFIID subunit 12

Chain l:

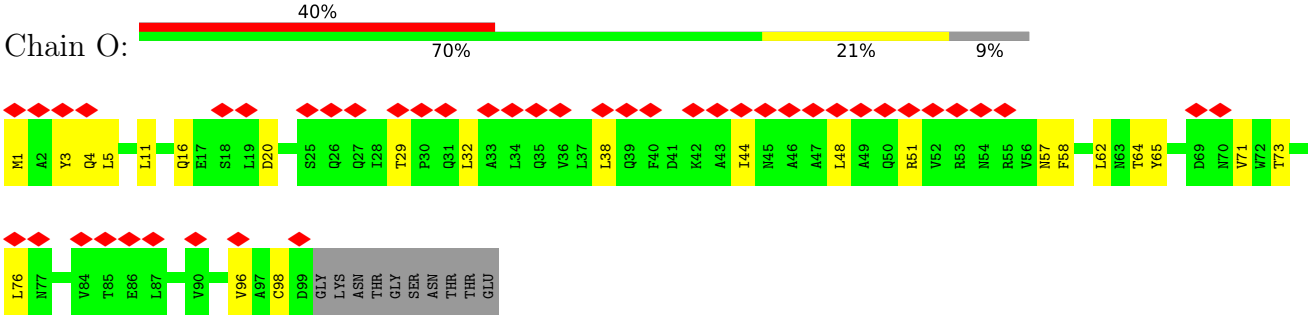






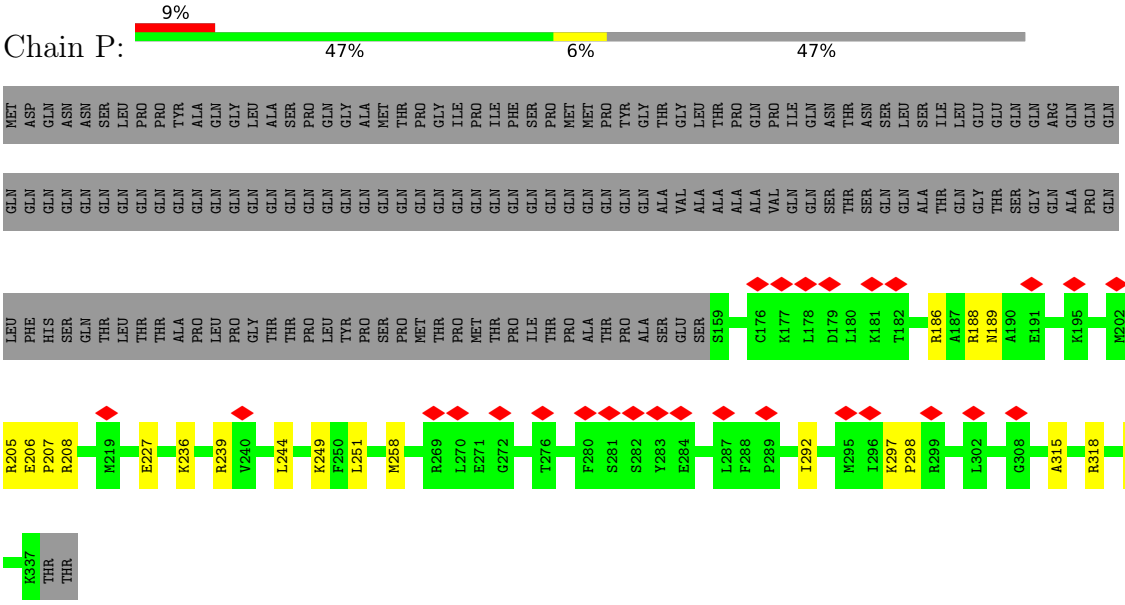
- Molecule 21: Transcription initiation factor IIA subunit 2

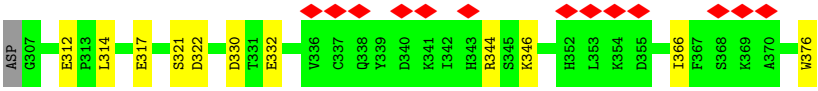
Chain O:



- Molecule 22: TATA-box-binding protein

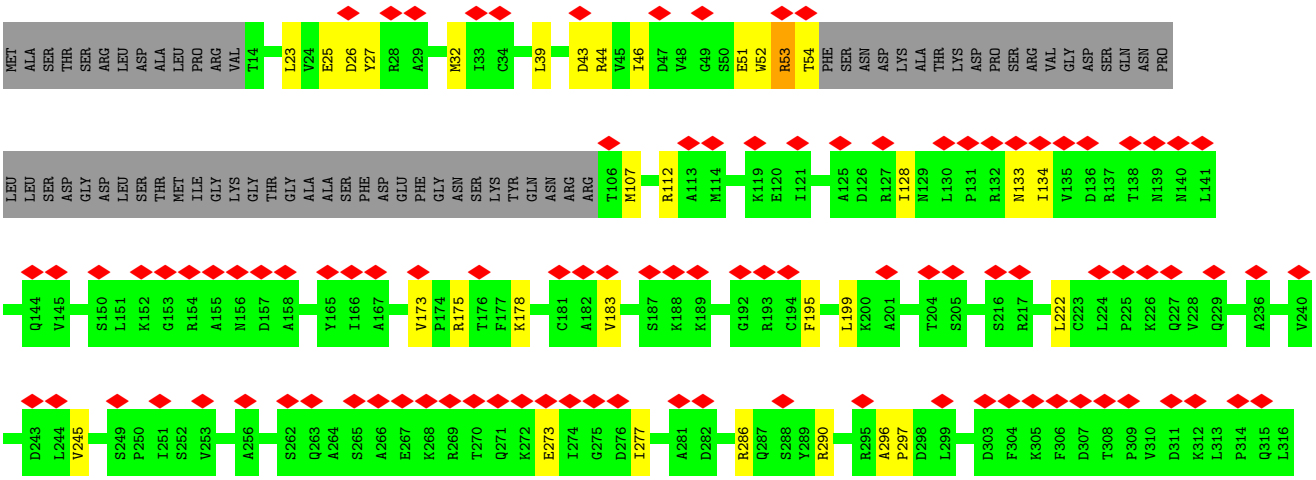
Chain P:


9%
47%
6%
47%
MET ASP GLN ASN ASN SER LEU PRO PRO TYR ALA GLN GLY LEU ALA SER PRO GLN GLY ALA MET THR PRO GLY ILE PRO ILE PHE SER PRO MET MET PRO TYR GLY THR GLY LEU THR PRO GLN PRO ILE GLN ASN THR ASN SER LEU SER ILE LEU GLU GLU GLN GLN ARG GLN GLN GLN
GLN GLN GLN



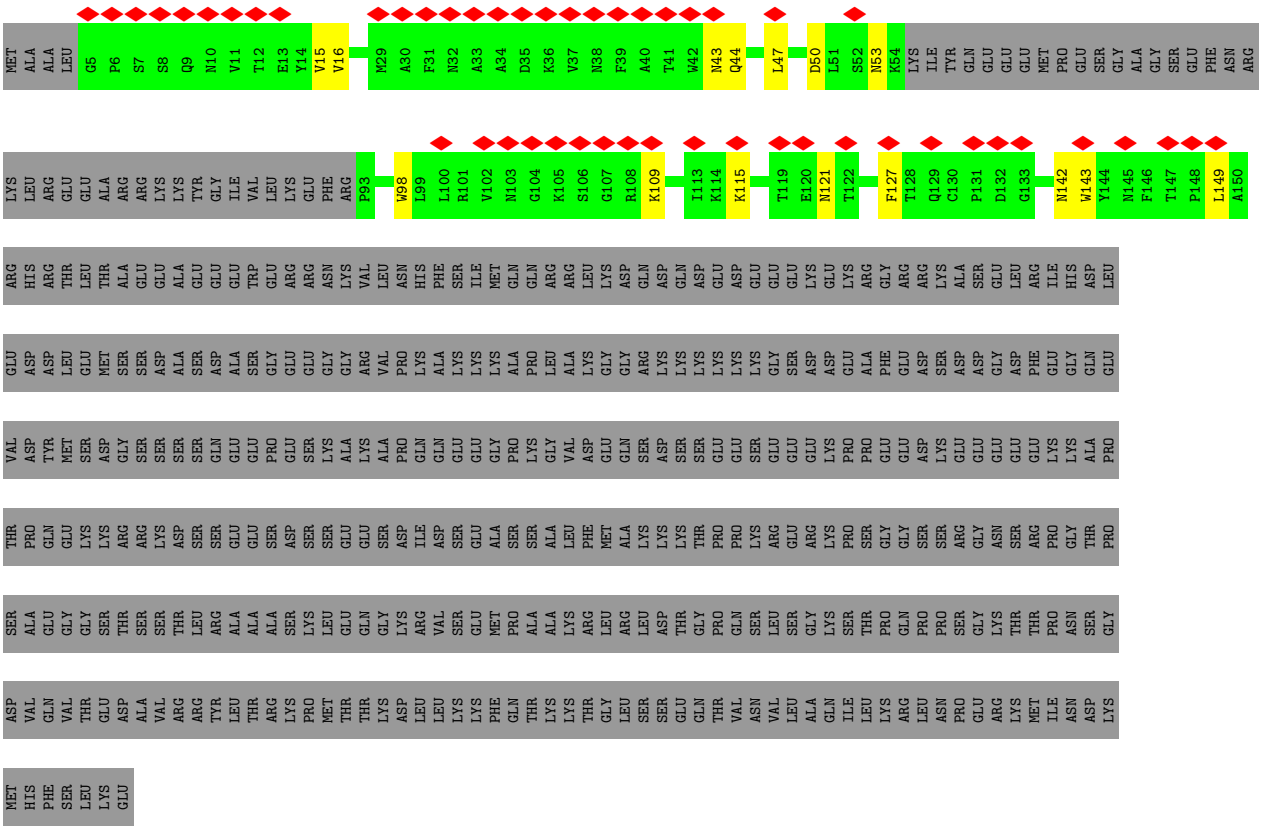
- Molecule 24: Transcription initiation factor IIB

Chain R:



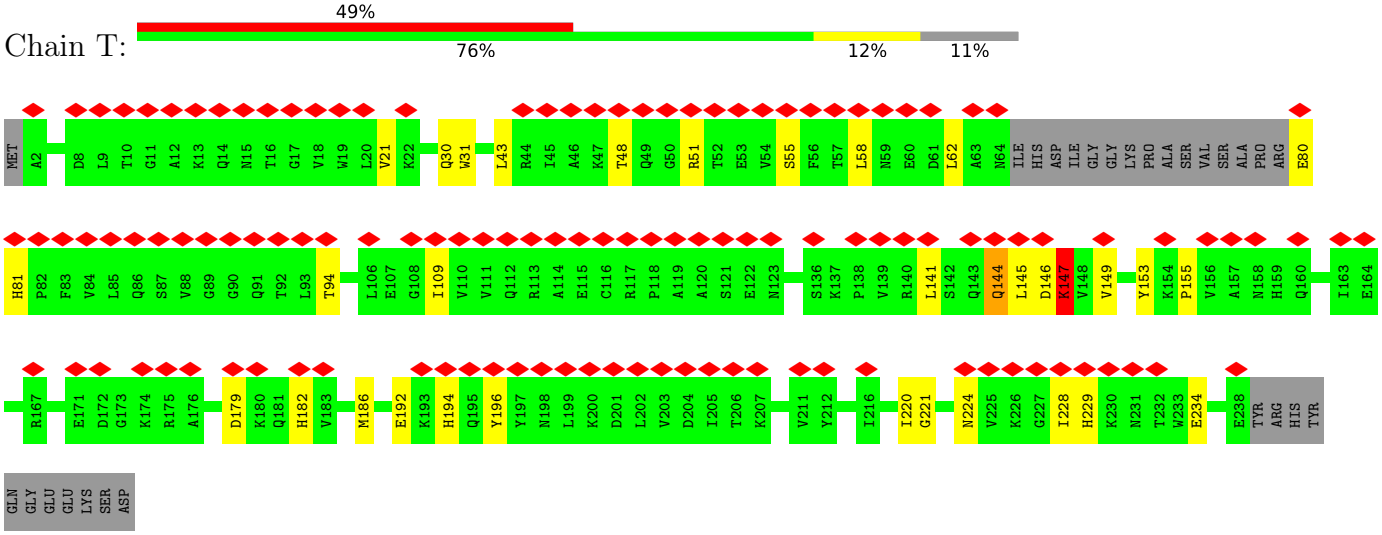
- Molecule 25: General transcription factor IIF subunit 1

Chain S:



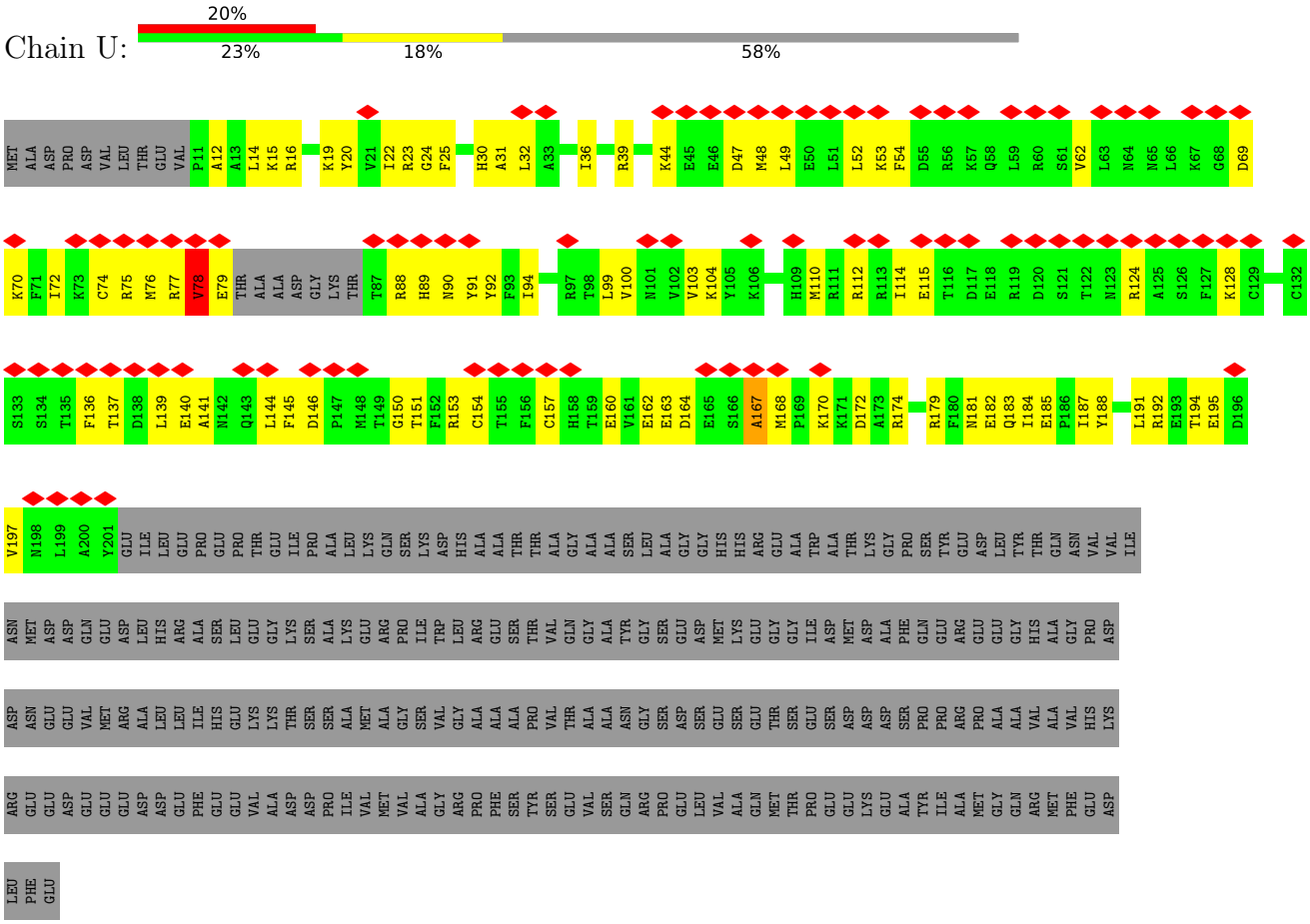
• Molecule 26: General transcription factor IIF subunit 2

Chain T:



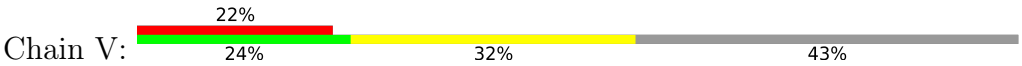
• Molecule 27: General transcription factor IIE subunit 1

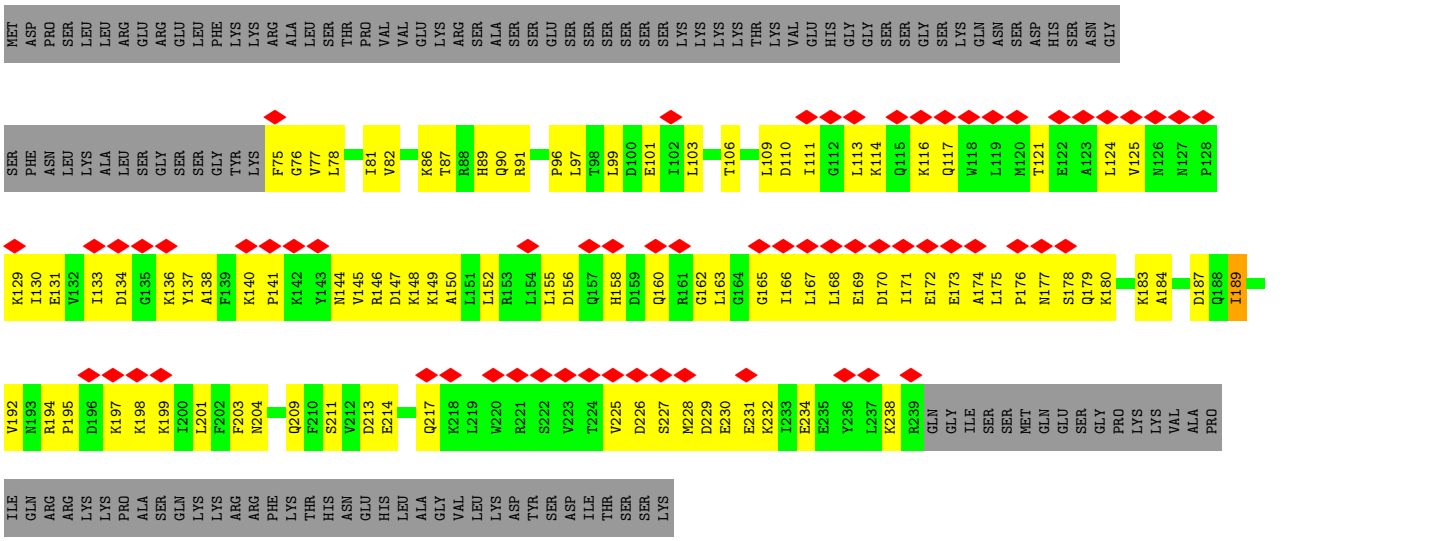
Chain U:



• Molecule 28: Transcription initiation factor IIE subunit beta

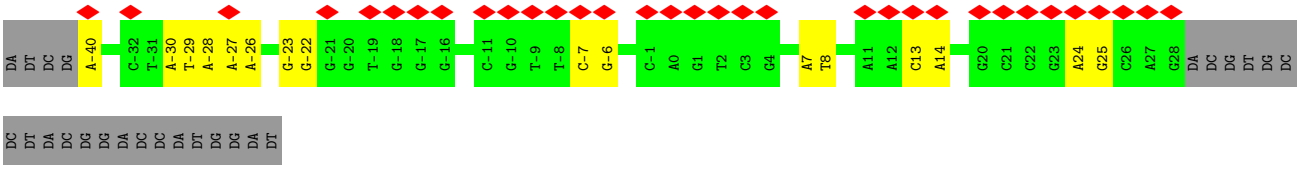
Chain V:





- Molecule 29: DNA (69-MER)

Chain X:

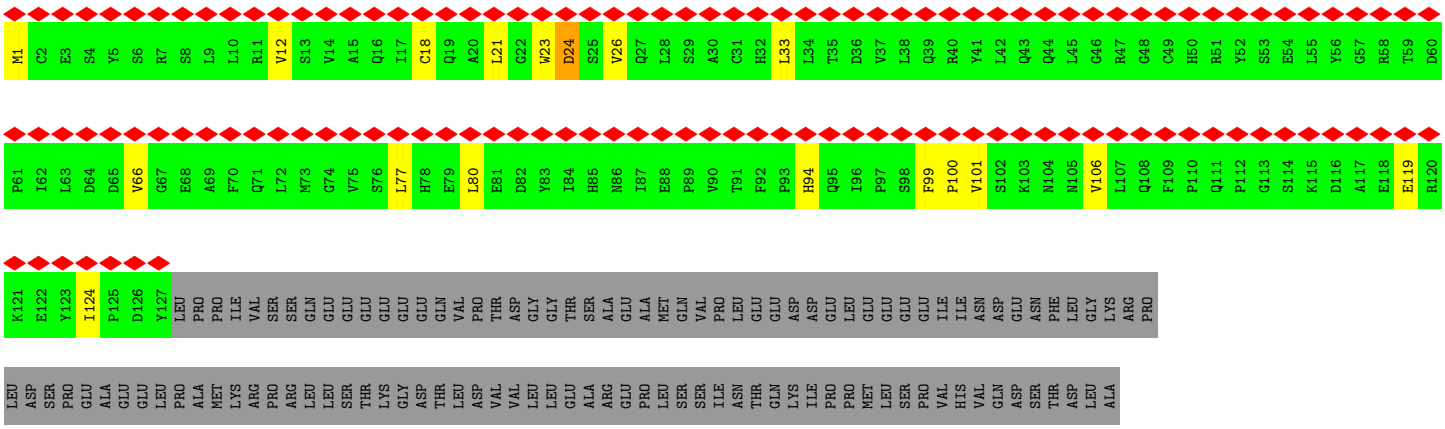


- Molecule 30: DNA (69-MER)

Chain Y:

- Molecule 31: Transcription initiation factor TFIID subunit 3

Chain c:



PRO PRO SER PRO GLU PRO PRO MET LEU ALA PRO VAL ALA LYS SER GLN MET PRO THR ALA LYS PRO LEU GLU THR LYS SER PHE THR PRO LYS THR LYS THR LYS THR SER SER PRO GLY GLN LYS THR LYS SER PRO LYS THR ALA GLN SER PRO ALA MET VAL GLY SER PRO ILE ARG

SER PRO LYS THR VAL SER LYS GLU LYS LYS SER PRO GLY ARG SER LYS SER PRO LYS SER PRO LYS SER PRO LYS VAL THR THR HIS ILE PRO GLN THR PRO VAL ARG PRO GLU THR PRO ASN ARG THR PRO SER ALA THR LEU SER GLU LYS ILE SER LYS GLU THR ILE GLN VAL LYS

GLN ILE GLN THR PRO PRO ASP ALA GLY LYS LYS LEU ASN SER GLU ASN GLN PRO LYS LYS ALA VAL VAL ALA ASP LYS THR ILE GLU ALA SER ILE ASP ALA VAL ILE ALA ARG ALA CYS ALA GLU ARG GLU PRO ASP PRO PHE GLU PHE SER SER GLY SER GLU SER GLU LYS ALA LEU GLY PHE

THR SER PRO LYS ARG ILE SER GLY PRO GLU CYS THR THR PRO LYS ALA SER THR SER ALA ASN ASN PHE THR LYS SER GLY SER THR PRO LEU PRO LEU GLY GLY THR SER SER SER SER ASP ASN SER TRP THR MET ASP ALA SER ILE ASP GLU VAL VAL ARG LYS ALA LYS LEU GLY

THR PRO SER ASN MET PRO PRO ASN PHE PRO TYR ILE SER SER PRO SER VAL PRO PRO THR PRO GLU PRO LEU HIS LYS VAL TYR GLU GLU LYS THR LYS LEU PRO SER SER VAL GLU VAL LYS LYS LYS LEU LYS LYS GLU LEU LYS THR LYS LYS GLU LYS GLN ARG

ASP ARG GLU ARG GLU LYS ASP LYS ASN LYS ASP LYS SER LYS GLU LYS ASP LYS VAL LYS GLU LYS GLU LYS ASP LYS GLU THR GLY ARG GLU THR LYS TYR PRO TRP LYS GLU PHE LEU LYS GLU GLU GLU ALA ASP PRO TYR LYS PHE LYS ILE LYS GLU PHE GLU ASP VAL ASP PRO

LYS VAL LYS LEU LYS ASP GLY LEU VAL ARG LYS GLU LYS GLU LYS HIS LYS ASP LYS LYS LYS ASP ARG GLU LYS GLY LYS LYS ASP LYS ASP LYS GLU LYS GLU LYS ASP LYS GLY ARG GLU ASP LYS MET LYS ALA PRO ALA PRO PRO LEU VAL LEU PRO PRO LYS GLU

LEU ALA LEU PRO LEU PHE SER PRO ALA THR ALA SER ARG VAL PRO ALA MET LEU PRO SER LEU LEU PRO VAL LEU PRO GLU LYS LEU PHE GLU GLU LYS GLU LYS VAL LYS GLU LYS GLU LYS LYS LYS ASP LYS LYS LYS LYS LYS LYS LYS LYS GLU

LYS GLU ARG GLU LYS GLU LYS ARG GLU ARG GLU LYS ARG GLU LYS GLU LYS GLU LYS HIS LYS HIS GLU GLU LYS ILE LYS VAL GLU PRO VAL ALA LEU ALA PRO SER PRO VAL ILE PRO ARG LEU THR LEU ARG VAL GLY ALA GLY GLN ASP LYS ILE VAL ILE SER LYS VAL VAL PRO ALA

PRO GLU ALA LYS PRO ALA PRO SER GLN ASN ARG PRO LYS THR PRO PRO PRO ALA PRO ALA PRO PRO GLY PRO MET LEU VAL SER PRO ALA PRO VAL PRO LEU PRO LEU LEU ALA GLN ALA ALA ALA GLY PRO ALA LEU LEU PRO SER PRO GLY PRO ALA ALA SER GLY ALA SER ALA

LYS ALA PRO VAL ARG SER VAL VAL THR GLU THR VAL SER THR TYR VAL ILE ARG ASP GLU TRP GLY ASN GLN ILE TRP ILE CYS PRO GLY CYS ASN LYS PRO ASP ASP GLY SER PRO MET ILE GLY CYS ASP ASP CYS ASP ASP TRP TYR HIS TRP PRO CYS VAL GLY ILE MET THR ALA

PRO PRO GLU GLU MET GLN TRP PHE CYS PRO LYS CYS ALA ASN LYS LYS LYS ASP LYS LYS HIS LYS LYS ARG LYS HIS ARG ALA HIS

- Molecule 32: Transcription initiation factor TFIID subunit 11

Chain k:

MET ASP ASP ALA HIS GLU SER PRO SER ASP LYS GLY GLY GLU THR GLY GLU SER ASP GLU THR ALA ALA VAL PRO GLY ASP PRO GLY ALA THR ASP THR ASP GLY ILE PRO GLU GLU THR ASP GLY ASP ALA ASP VAL ASP LEU LYS GLU ALA ALA ALA GLU GLU GLY GLU LEU GLU SER

GLN ASP VAL SER ASP LEU THR THR VAL GLU ARG GLU ASP SER SER LEU LEU ASN PRO ALA ALA LYS LYS LEU LYS ILE ASP THR LYS GLU LYS LYS GLU LYS LYS GLN LYS VAL ASP GLU ASP GLU ILE GLN LYS MET GLN ILE LEU VAL SER SER PHE S114 E115 E116 Q117 L118 N119 R120

Y121 E122 M123 Y124 R125 R126 S127 A128 F129 P130 K131 A132 A133 I134 K135 R136 L137 I138 Q139 S140 I141 T142 G143 T144 S145 V146 S147 Q148 N149 V150 V151 I152 A153 M154 S155 G156 I157 S158 K159 V160 F161 V162 G163 E164 V165 V166 E167 E168 A169 L170 D171 V172 C173 E174 K175 W176 G177 E178 M179 P180

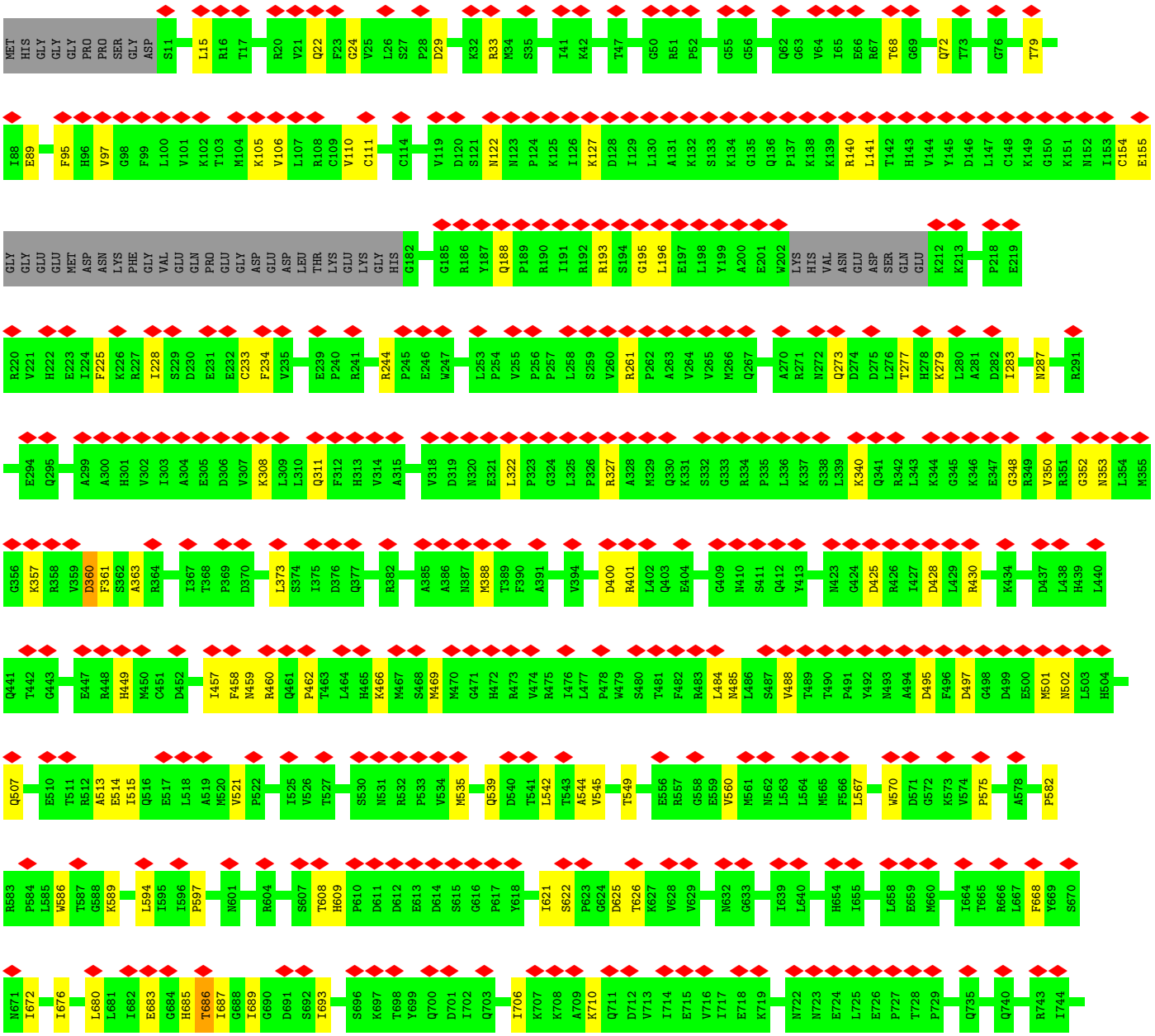
P181 L182 Q183 P184 K185 H186 M187 R188 E189 A190 V191 R192 R193 L194 K195 S196 K197 G198 Q199 I200 P201 N202 S203 K204 H205 K206 K207 I208 I209 F210 F211

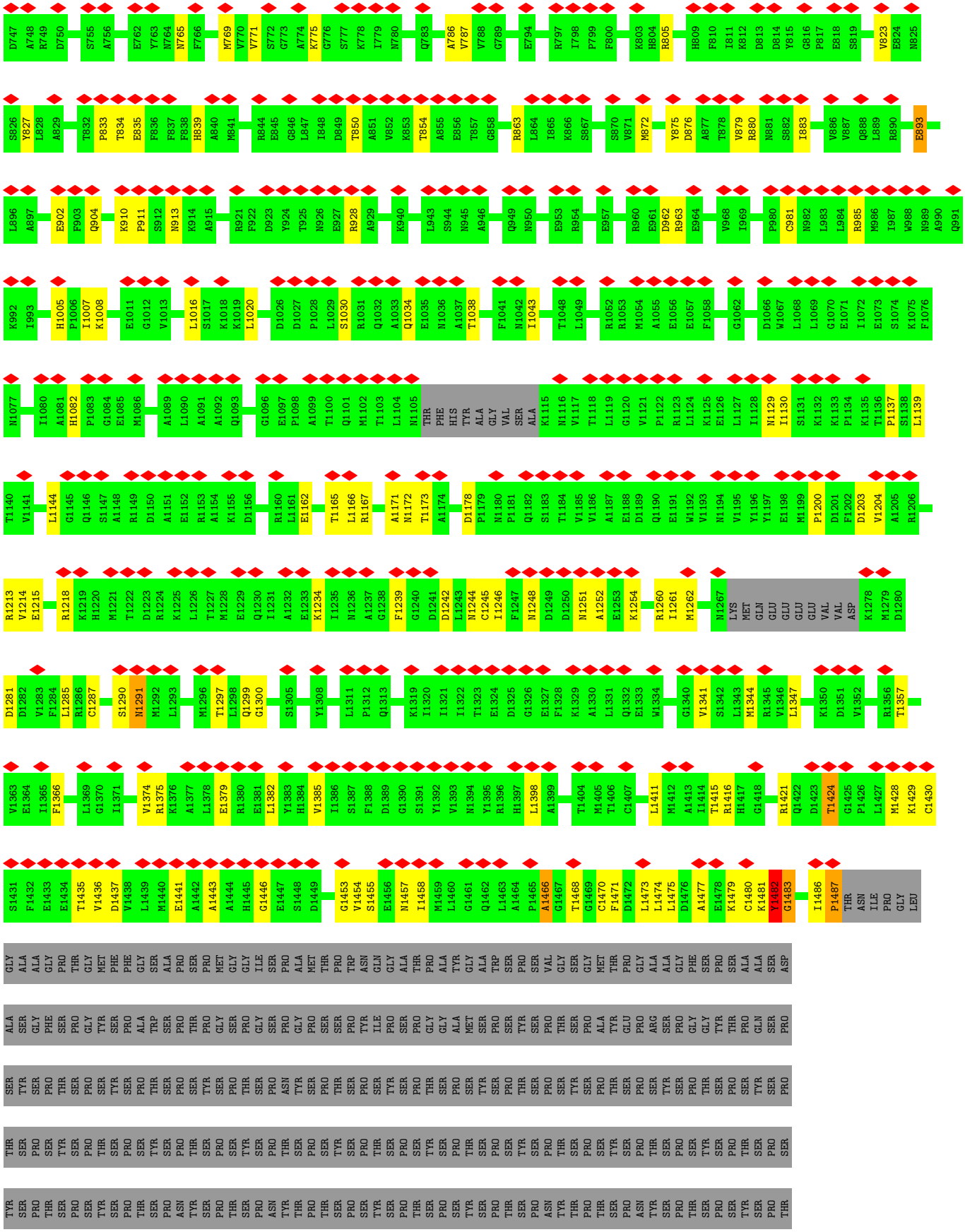
- Molecule 33: Transcription initiation factor TFIID subunit 13

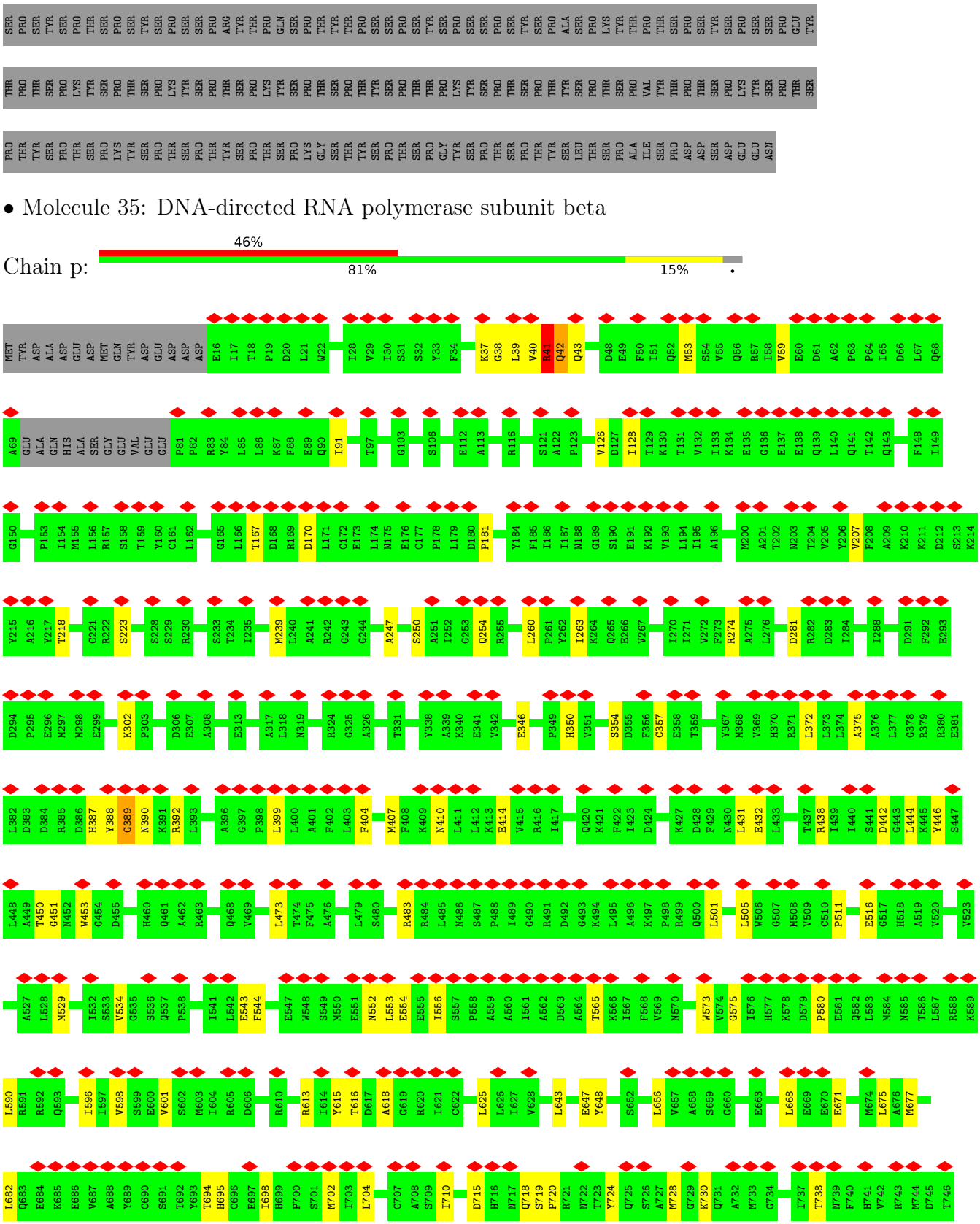
Chain m:

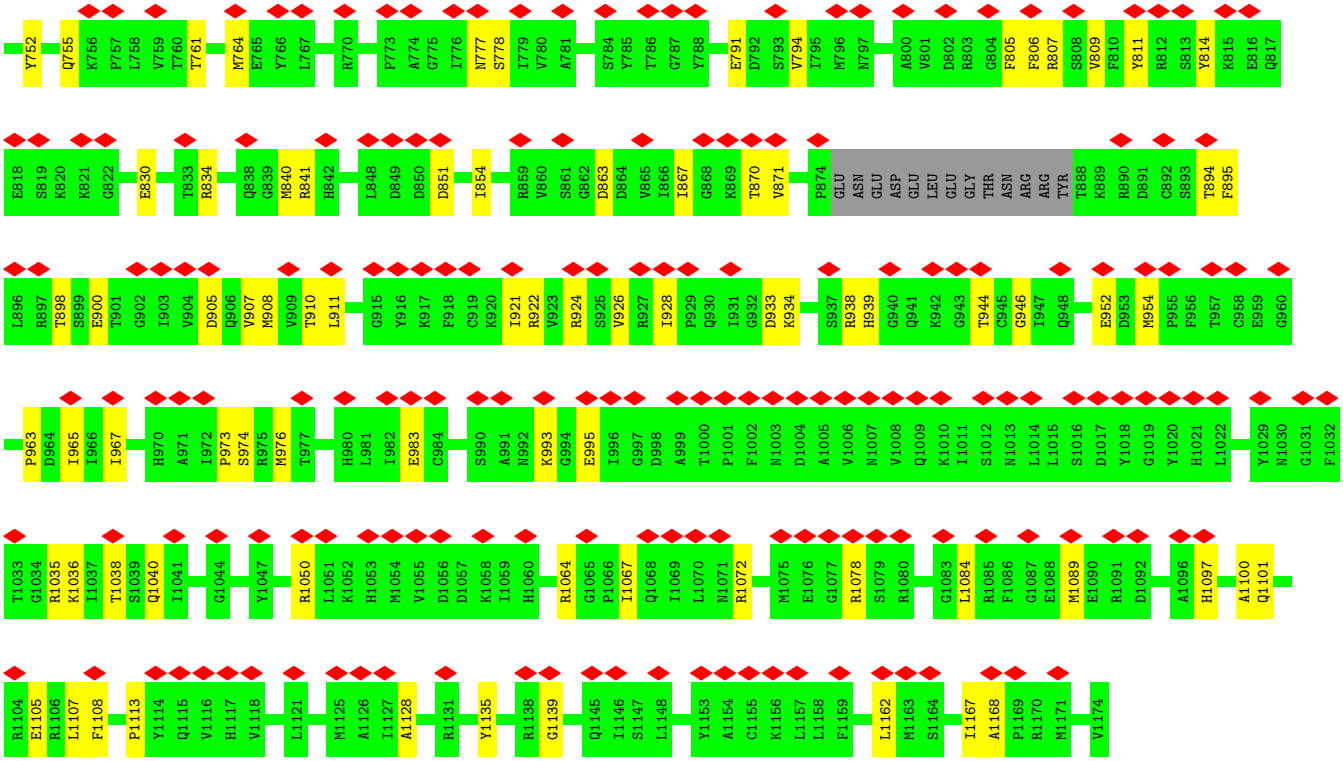
- Molecule 34: DNA-directed RNA polymerase II subunit RPB1

Chain o:



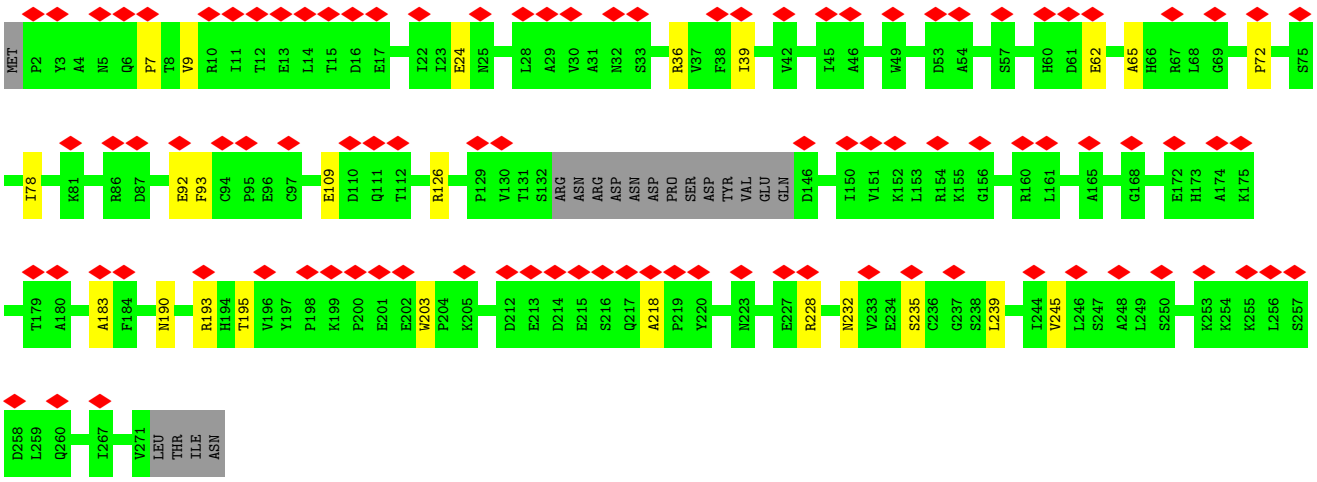

D747 A748 R749 D750 S755 A756 E762 Y763 N764 N765 F766 M769 V770 V771 S772 G773 A774 K775 G776 S777 K778 I779 N780 Q783 A786 V787 V788 G789 E794 R797 I798 P799 F800 K803 H804 R805 H809 F810 I811 K812 D813 D814 Y815 G816 P817 E818 S819 V823 E824 N825
S826 Y827 L828 A829 T832 P833 T834 E835 F836 F837 F838 H839 A840 M841 R844 E845 G846 L847 I848 D849 T850 A851 V852 K853 T854 A855 E856 T857 G858 R863 L864 I865 K866 S867 S870 V871 M872 Y875 D876 A877 T878 V879 R880 N881 S882 I883 V886 V887 Q888 L889 R890 E893
L896 A897 E902 F903 Q904 K910 P911 S912 N913 K914 A915 R921 F922 D923 Y924 T925 N926 E927 R928 A929 K940 L943 S944 N945 A946 Q949 N950 E953 R954 E957 R960 E961 D962 R963 E964 V968 I969 P980 C981 N982 L983 L984 R985 M986 I987 W988 N989 A990 Q991
K992 I993 H1005 P1006 I1007 K1008 E1011 G1012 V1013 L1016 S1017 K1018 K1019 L1020 D1026 D1027 P1028 L1029 S1030 R1031 Q1032 A1033 Q1034 E1035 N1036 A1037 T1038 F1041 N1042 I1043 T1048 L1049 R1052 R1053 M1054 A1055 E1056 E1057 F1058 G1062 D1066 W1067 L1068 L1069 G1070 E1071 I1072 E1073 S1074 K1075 F1076
N1077 I1080 A1081 H1082 P1083 G1084 E1085 M1086 A1089 L1090 A1091 A1092 Q1093 G1096 E1097 P1098 A1099 T1100 Q1101 M1102 T1103 L1104 N1105 THR PHE HIS TYR ALA GLY VAL SER ALA K1115 N1116 V1117 T1118 L1119 G1120 V1121 P1122 R1123 L1124 K1125 E1126 L1127 I1128 N1129 I1130 S1131 K1132 K1133 P1134 K1135 T1136 P1137 S1138 L1139
T1140 V1141 L1144 G1145 Q1146 S1147 A1148 R1149 D1150 A1151 E1152 R1153 A1154 K1155 D1156 R1160 L1161 E1162 T1165 L1166 R1167 A1171 N1172 T1173 A1174 D1178 P1179 N1180 P1181 Q1182 S1183 T1184 V1185 V1186 A1187 E1188 D1189 Q1190 E1191 W1192 V1193 N1194 V1195 Y1196 Y1197 E1198 M1199 P1200 D1201 F1202 D1203 V1204 A1205 R1206
R1213 V1214 E1215 R1218 K1219 H1220 M1221 T1222 D1223 R1224 K1225 L1226 T1227 M1228 E1229 Q1230 I1231 A1232 E1233 K1234 I1235 N1236 A1237 G1238 F1239 G1240 D1241 D1242 L1243 N1244 C1245 I1246 F1247 N1248 D1249 D1250 N1251 A1252 E1253 K1254 R1260 I1261 M1262 N1267 LYS MET GLN GLU GLU GLU GLU VAL VAL ASP K1278 M1279 D1280
D1281 D1282 V1283 F1284 L1285 R1286 C1287 S1290 N1291 M1292 L1293 M1296 T1297 L1298 Q1299 G1300 S1305 Y1308 L1311 P1312 Q1313 K1319 I1320 I1321 I1322 T1323 E1324 D1325 G1326 E1327 F1328 K1329 A1330 L1331 Q1332 E1333 W1334 G1340 V1341 S1342 L1343 M1344 R1345 V1346 L1347 K1350 D1351 V1352 R1356 T1357
V1363 E1364 I1365 F1366 L1369 G1370 I1371 V1374 R1375 K1376 A1377 L1378 E1379 R1380 E1381 L1382 Y1383 H1384 V1385 I1386 S1387 F1388 D1389 G1390 S1391 Y1392 V1393 N1394 Y1395 R1396 H1397 L1398 A1399 T1404 M1405 T1406 C1407 L1411 M1412 A1413 I1414 T1415 R1416 H1417 G1418 R1421 Q1422 D1423 T1424 G1425 P1426 L1427 M1428 K1429 C1430
S1431 F1432 E1433 E1434 T1435 V1436 D1437 V1438 L1439 M1440 E1441 A1442 A1443 A1444 A1444 H1445 G1446 E1447 S1448 D1449 G1453 V1454 S1455 E1456 N1457 I1458 M1459 L1460 G1461 Q1462 L1463 A1464 P1465 A1466 G1467 T1468 G1469 C1470 F1471 D1472 L1473 L1474 L1475 D1476 A1477 E1478 K1479 C1480 K1481 Y1482 G1483 I1486 P1487 THR ASN ILE PRO GLY LEU
GLY ALA ALA GLY PRO THR GLY MET PHE PHE GLY SER ALA PRO SER PRO MET GLY GLY ILE SER PRO ALA MET THR PRO TRP ASN GLN GLY ALA THR PRO GLY ALA TYR GLY ALA TRP TYR GLU ARG SER PRO GLY PHE SER PRO GLY TYR SER PRO ASN ASP
ALA SER GLY PHE SER PRO GLY PRO SER PRO ALA TRP SER PRO THR PRO GLY SER PRO SER PRO TYR ILE PRO SER PRO GLY GLY ALA MET SER PRO SER THR SER PRO ALA THR GLU PRO ARG SER TYR SER PRO GLY GLY TYR THR PRO GLN SER PRO
SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO ASN TYR SER PRO SER PRO TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER


SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO SER SER PRO ARG TYR THR PRO GLN SER PRO THR TYR TYR THR PRO SER SER PRO SER TYR SER PRO SER SER SER PRO SER TYR SER PRO SER PRO ALA SER PRO LYS TYR THR PRO THR SER PRO SER TYR SER PRO SER TYR
THR PRO THR SER PRO LYS TYR SER PRO SER TYR SER PRO THR SER PRO LYS TYR SER PRO THR SER PRO LYS TYR SER PRO THR SER PRO THR SER PRO THR TYR SER PRO THR SER PRO THR TYR SER PRO THR SER PRO VAL TYR THR PRO THR SER PRO LYS TYR SER PRO THR SER
PRO THR TYR SER PRO THR SER PRO LYS TYR SER PRO THR SER PRO THR SER PRO LYS GLY SER THR TYR SER PRO THR SER PRO GLY TYR SER PRO THR SER PRO THR TYR SER LEU THR SER PRO ALA ILE SER PRO ASP ASP SER ASP GLU GLU ASN
Chain p:
46%
81%
15%
MET TYR ASP ALA ASP GLU ASP MET GLN TYR ASP GLU ASP ASP ASP E16 I17 T18 P19 D20 L21 W22 I28 V29 I30 S31 S32 Y33 F34 K37 G38 L39 V40 R41 Q42 Q43 D48 E49 F50 I51 Q52 M53 S54 V55 Q56 R57 I58 V59 E60 D61 A62 P63 P64 I65 D66 L67 Q68
A69 GLU ALA GLN HIS ALA SER GLY GLU VAL GLU GLU P81 P82 R83 Y84 L85 L86 K87 F88 E89 Q90 I91 T97 G103 S106 E112 A113 R116 S121 A122 P123 V126 D127 I128 T129 K130 T131 V132 I133 K134 E135 G136 E137 E138 Q139 L140 Q141 T142 Q143 F148 I149
G150 P153 I154 M155 L156 R157 S158 T159 Y160 C161 L162 G165 L166 T167 D168 R169 D170 L171 C172 E173 L174 N175 E176 C177 P178 L179 D180 P181 Y184 F185 I186 I187 N188 G189 S190 E191 K192 V193 L194 I195 A196 M200 A201 T202 N203 T204 V205 Y206 V207 F208 A209 K210 K211 D212 S213 K214
Y215 A216 Y217 T218 C221 R222 S223 S228 S229 R230 S233 T234 I235 M239 L240 A241 R242 G243 G244 A247 S250 A251 I252 G253 Q254 R255 L260 P261 Y262 I263 K264 Q265 E266 V267 I270 I271 V272 F273 R274 A275 L276 D281 R282 D283 I284 I288 D291 F292 E293
D294 P295 E296 M297 M298 E299 K302 P303 D306 E307 A308 E313 A317 L318 N319 R324 G325 A326 T331 Y338 A339 K340 E341 V342 E346 P349 H350 V351 S354 D355 F356 C357 E358 T359 Y367 M368 V369 H370 R371 L372 L373 L374 A375 A376 L377 G378 R379 R380 E381
L382 D383 D384 R385 D386 H387 Y388 G389 N390 K391 R392 L393 A396 G397 P398 L399 L400 A401 F402 L403 F404 M407 F408 K409 N410 L411 L412 K413 E414 V415 R416 I417 Q420 K421 F422 I423 D424 K427 D428 F429 N430 L431 E432 L433 T437 R438 I439 I440 S441 D442 G443 L444 K445 Y446 S447
L448 A449 T450 G451 N452 W453 G454 D455 H460 Q461 A462 R463 Q468 V469 L473 T474 F475 A476 L479 S480 R483 R484 L485 N486 S487 P488 I489 G490 R491 D492 G493 K494 L495 A496 K497 P498 R499 Q500 L501 L505 W506 G507 M508 V509 C510 P511 E516 G517 H518 A519 V520 V523
A527 L528 M529 I532 S533 V534 G535 S536 Q537 P538 I541 L542 E543 F544 E547 W548 S549 M550 E551 N552 L553 E554 E555 I556 S557 P558 A559 A560 I561 A562 D563 A564 T565 K566 I567 F568 V569 N570 W573 V574 G575 I576 H577 K578 D579 P580 E581 Q582 L583 M584 N585 T586 L587 R588 K589
L590 R591 R592 Q593 I596 I597 V598 S599 E600 V601 S602 M603 I604 R605 D606 R610 R613 I614 Y615 T616 D617 A618 G619 R620 I621 C622 L625 L626 I627 V628 L643 E647 Y648 S652 L656 V657 A658 S659 G660 E663 L668 E669 E670 E671 M674 L675 A676 M677
L682 Q683 E684 K685 E686 V687 A688 Y689 C690 S691 T692 Y693 T694 H695 C696 E697 I698 H699 P700 S701 M702 I703 L704 C707 A708 S709 I710 D715 H716 N717 Q718 S719 P720 R721 N722 T723 Y724 Q725 S726 A727 M728 G729 K730 Q731 A732 M733 G734 I737 T738 N739 F740 H741 V742 R743 M744 D745 T746



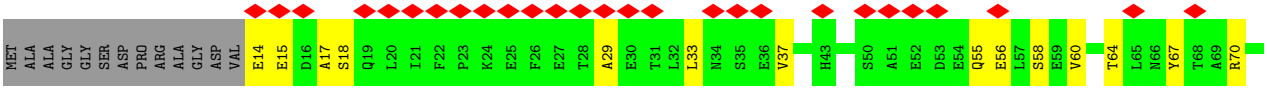
- Molecule 36: DNA-directed RNA polymerase II subunit RPB3

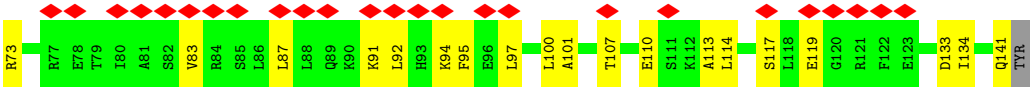
Chain q:



- Molecule 37: DNA-directed RNA polymerase II subunit RPB4

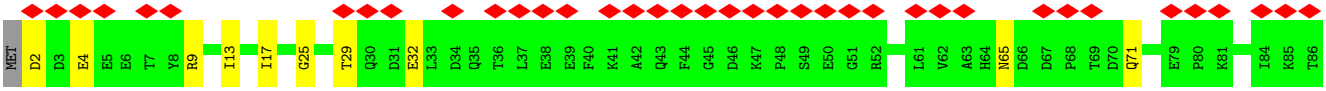
Chain r:

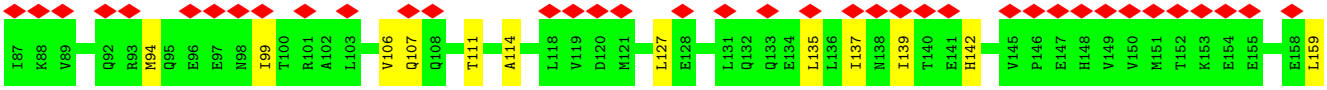


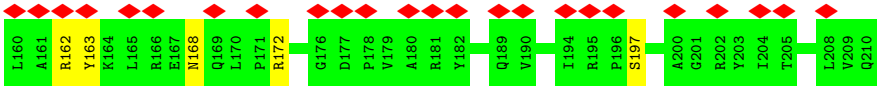


• Molecule 38: DNA-directed RNA polymerase II subunit E

Chain s:

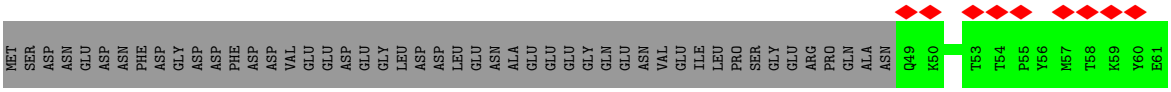


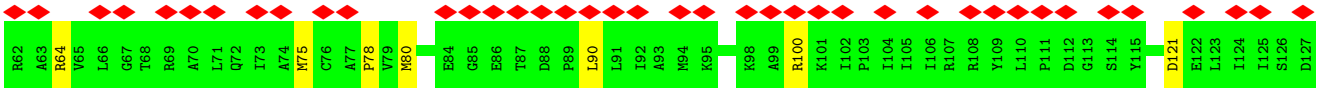




• Molecule 39: DNA-directed RNA polymerase II subunit F

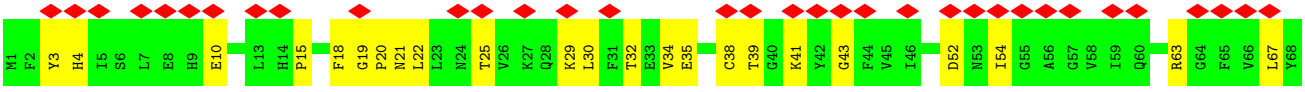
Chain t:

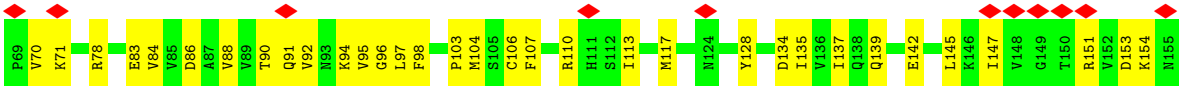


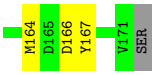


• Molecule 40: DNA-directed RNA polymerase II subunit RPB7

Chain u:

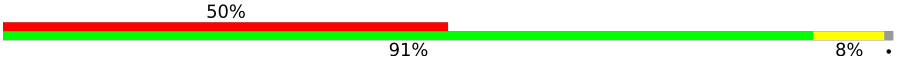


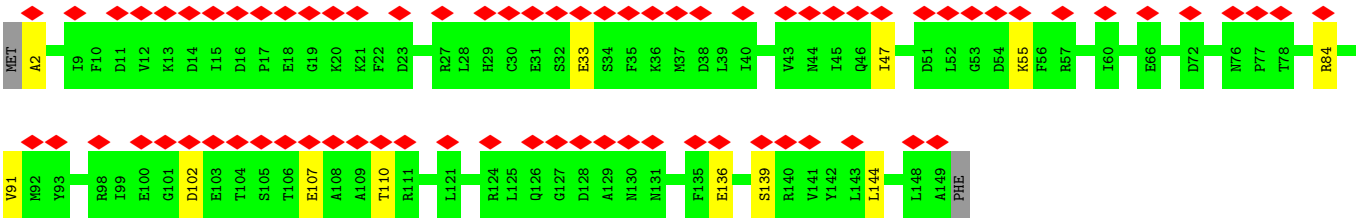




• Molecule 41: DNA-directed RNA polymerases I, II, and III subunit RPABC3

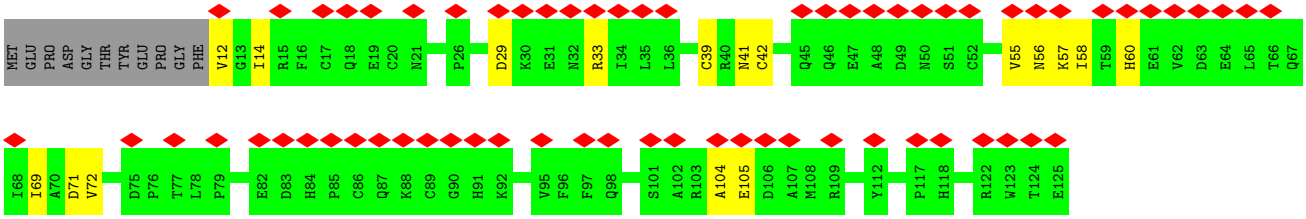
Chain v:





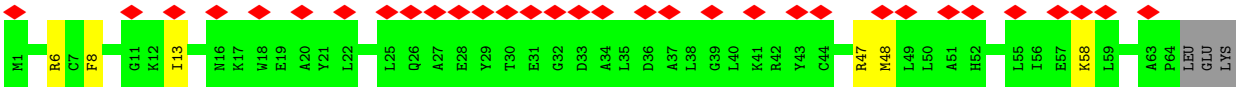
• Molecule 42: DNA-directed RNA polymerase II subunit RPB9

Chain w:



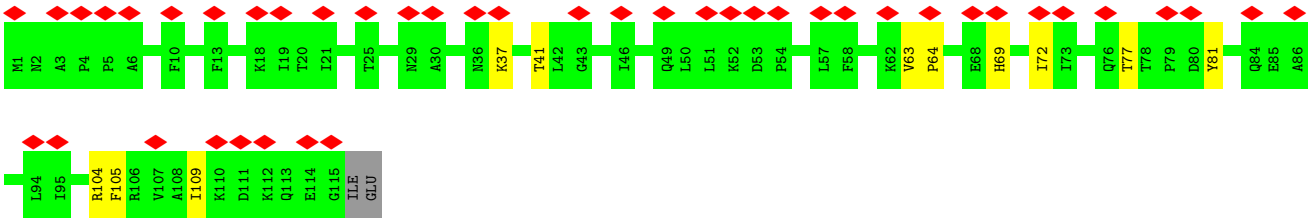
• Molecule 43: RPB10

Chain x:



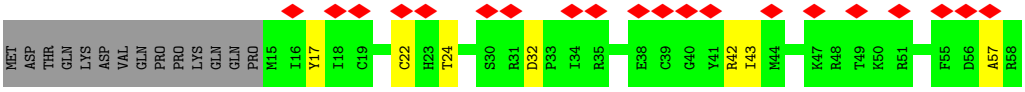
• Molecule 44: RNA_pol_L_2 domain-containing protein

Chain y:



• Molecule 45: RPB12

Chain z:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	92806	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	50.0	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	6.811	Depositor
Minimum map value	-4.382	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.084	Depositor
Recommended contour level	0.3	Depositor
Map size (Å)	542.72, 542.72, 542.72	wwPDB
Map dimensions	512, 512, 512	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.06, 1.06, 1.06	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: SF4, MG, ZN

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
1	0	0.39	0/1965	0.62	0/2646
2	1	0.23	0/2210	0.39	0/2975
3	2	0.16	0/2624	0.28	0/3555
4	3	0.30	0/2103	0.47	1/2846 (0.0%)
5	4	0.31	0/3262	0.51	2/4418 (0.0%)
6	5	0.34	0/433	0.60	0/585
7	6	0.29	0/4994	0.48	2/6745 (0.0%)
8	7	0.21	0/5875	0.37	1/7955 (0.0%)
9	8	0.35	0/2434	0.55	1/3300 (0.0%)
10	9	0.46	0/2342	0.72	6/3159 (0.2%)
11	A	0.63	0/4698	0.95	5/6345 (0.1%)
12	B	0.46	1/7993 (0.0%)	0.69	6/10836 (0.1%)
13	D	0.54	0/1350	0.81	2/1805 (0.1%)
13	d	0.48	0/1321	0.71	2/1772 (0.1%)
14	E	0.45	1/4482 (0.0%)	0.72	3/6069 (0.0%)
14	e	0.55	1/4433 (0.0%)	0.77	0/6004
15	F	0.65	1/3201 (0.0%)	1.07	13/4347 (0.3%)
15	f	0.47	1/3140 (0.0%)	0.80	6/4268 (0.1%)
16	G	0.68	0/1190	0.94	1/1601 (0.1%)
17	H	0.63	3/1673 (0.2%)	1.00	14/2285 (0.6%)
18	I	0.21	0/981	0.41	0/1332
18	i	0.21	0/989	0.40	0/1343
19	J	0.83	0/736	1.15	2/998 (0.2%)
19	j	0.73	0/775	0.96	0/1049
20	L	0.68	0/630	1.13	1/852 (0.1%)
20	l	0.59	0/888	0.88	0/1194
21	O	0.25	0/816	0.43	0/1105
22	P	0.16	0/1448	0.28	0/1948
23	Q	0.18	0/945	0.35	0/1274
24	R	0.23	0/1983	0.38	0/2679
25	S	0.12	0/896	0.25	0/1213
26	T	0.20	0/1817	0.39	1/2445 (0.0%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
27	U	0.41	0/1545	0.68	1/2075 (0.0%)
28	V	0.24	0/1380	0.58	0/1854
29	X	0.28	0/1607	0.58	0/2481
30	Y	0.29	0/1565	0.59	0/2410
31	c	0.48	0/1035	0.67	0/1406
32	k	0.29	0/799	0.53	1/1070 (0.1%)
33	m	0.85	0/733	1.18	1/977 (0.1%)
34	o	0.37	8/11479 (0.1%)	0.55	21/15496 (0.1%)
35	p	0.40	6/9249 (0.1%)	0.39	7/12482 (0.1%)
36	q	0.15	0/2102	0.29	0/2857
37	r	0.14	0/1064	0.22	0/1428
38	s	0.29	1/1752 (0.1%)	0.53	3/2366 (0.1%)
39	t	0.16	0/646	0.27	0/871
40	u	0.20	0/1382	0.28	0/1874
41	v	0.14	0/1207	0.28	0/1628
42	w	0.23	0/949	0.41	0/1284
43	x	0.16	0/516	0.25	0/696
44	y	0.22	0/939	0.37	0/1271
45	z	0.17	0/378	0.28	0/500
All	All	0.41	23/114954 (0.0%)	0.62	103/155974 (0.1%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	0	0	1
26	T	0	1
27	U	0	1
All	All	0	3

The worst 5 of 23 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
35	p	704	LEU	C-N	20.97	1.56	1.33
35	p	806	PHE	C-N	16.04	1.56	1.33
35	p	928	ILE	C-N	13.37	1.50	1.33
34	o	388	MET	C-N	12.26	1.50	1.33
34	o	1466	ALA	C-N	12.03	1.51	1.33

The worst 5 of 103 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
34	o	1082	HIS	CA-C-N	16.32	136.59	119.78
34	o	1082	HIS	C-N-CA	16.32	136.59	119.78
38	s	139	ILE	O-C-N	-15.11	106.61	121.87
34	o	1483	GLY	N-CA-C	14.43	129.23	112.50
17	H	156	PRO	N-CA-C	-13.59	90.10	111.03

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	0	126	GLN	Mainchain
26	T	144	GLN	Mainchain
27	U	78	VAL	Peptide

5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	0	1932	0	1905	86	0
2	1	2167	0	2175	35	0
3	2	2567	0	2539	49	0
4	3	2066	0	2097	89	0
5	4	3189	0	3243	131	0
6	5	428	0	434	45	0
7	6	4890	0	4949	214	0
8	7	5751	0	5794	115	0
9	8	2376	0	2402	73	0
10	9	2293	0	2315	76	0
11	A	4580	0	4502	97	0
12	B	7796	0	7751	113	0
13	D	1337	0	1358	44	0
13	d	1307	0	1318	16	0
14	E	4376	0	4256	78	0
14	e	4327	0	4197	64	0
15	F	3143	0	3093	103	0
15	f	3081	0	3012	70	0
16	G	1171	0	1222	17	0
17	H	1633	0	1621	42	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
18	I	959	0	969	13	0
18	i	967	0	977	10	0
19	J	720	0	719	12	0
19	j	759	0	759	14	0
20	L	622	0	620	21	0
20	l	876	0	893	10	0
21	O	806	0	818	22	0
22	P	1422	0	1514	29	0
23	Q	930	0	888	31	0
24	R	1953	0	1987	30	0
25	S	872	0	848	9	0
26	T	1788	0	1819	40	0
27	U	1520	0	1520	88	0
28	V	1357	0	1381	78	0
29	X	1429	0	770	13	0
30	Y	1400	0	775	9	0
31	c	1011	0	975	17	0
32	k	785	0	818	7	0
33	m	724	0	744	26	0
34	o	11274	0	11406	175	0
35	p	9068	0	9112	157	0
36	q	2059	0	2007	18	0
37	r	1050	0	1033	30	0
38	s	1721	0	1737	18	0
39	t	636	0	665	18	0
40	u	1351	0	1358	65	0
41	v	1186	0	1147	9	0
42	w	928	0	861	16	0
43	x	507	0	523	7	0
44	y	920	0	942	11	0
45	z	373	0	378	6	0
46	0	2	0	0	0	0
46	2	3	0	0	0	0
46	3	2	0	0	0	0
46	R	1	0	0	0	0
46	U	1	0	0	0	0
46	o	2	0	0	0	0
46	p	1	0	0	0	0
46	q	1	0	0	0	0
46	w	2	0	0	0	0
46	x	1	0	0	0	0
46	z	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
47	7	8	0	0	0	0
48	o	1	0	0	0	0
All	All	112409	0	111146	2117	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 9.

The worst 5 of 2117 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)
7:6:308:ILE:CG2	7:6:358:ARG:HA	1.25	1.62
6:5:33:PHE:CE1	6:5:49:LEU:HD12	1.44	1.51
6:5:31:LYS:HB3	12:B:540:LYS:NZ	1.31	1.39
4:3:50:PHE:CZ	5:4:118:LEU:HD11	1.57	1.37
6:5:31:LYS:CB	12:B:540:LYS:HZ3	1.37	1.37

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	232/309 (75%)	215 (93%)	17 (7%)	0	100 100
2	1	253/548 (46%)	245 (97%)	8 (3%)	0	100 100
3	2	325/395 (82%)	313 (96%)	12 (4%)	0	100 100
4	3	259/308 (84%)	251 (97%)	8 (3%)	0	100 100
5	4	384/462 (83%)	367 (96%)	17 (4%)	0	100 100
6	5	52/71 (73%)	47 (90%)	4 (8%)	1 (2%)	6 28
7	6	601/782 (77%)	576 (96%)	23 (4%)	2 (0%)	36 65
8	7	710/760 (93%)	683 (96%)	26 (4%)	1 (0%)	48 75

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
9	8	294/346 (85%)	277 (94%)	17 (6%)	0	100	100
10	9	277/323 (86%)	268 (97%)	9 (3%)	0	100	100
11	A	543/1872 (29%)	519 (96%)	23 (4%)	1 (0%)	43	71
12	B	959/1199 (80%)	911 (95%)	48 (5%)	0	100	100
13	D	154/1085 (14%)	148 (96%)	6 (4%)	0	100	100
13	d	154/1085 (14%)	149 (97%)	5 (3%)	0	100	100
14	E	541/800 (68%)	517 (96%)	23 (4%)	1 (0%)	43	71
14	e	531/800 (66%)	482 (91%)	48 (9%)	1 (0%)	43	71
15	F	408/677 (60%)	389 (95%)	18 (4%)	1 (0%)	43	71
15	f	399/677 (59%)	380 (95%)	19 (5%)	0	100	100
16	G	138/349 (40%)	135 (98%)	3 (2%)	0	100	100
17	H	207/310 (67%)	189 (91%)	14 (7%)	4 (2%)	6	28
18	I	118/264 (45%)	115 (98%)	3 (2%)	0	100	100
18	i	119/264 (45%)	115 (97%)	4 (3%)	0	100	100
19	J	86/218 (39%)	82 (95%)	4 (5%)	0	100	100
19	j	91/218 (42%)	85 (93%)	4 (4%)	2 (2%)	5	26
20	L	74/161 (46%)	72 (97%)	2 (3%)	0	100	100
20	l	105/161 (65%)	100 (95%)	5 (5%)	0	100	100
21	O	97/109 (89%)	97 (100%)	0	0	100	100
22	P	177/339 (52%)	175 (99%)	2 (1%)	0	100	100
23	Q	109/376 (29%)	102 (94%)	7 (6%)	0	100	100
24	R	248/316 (78%)	246 (99%)	2 (1%)	0	100	100
25	S	104/517 (20%)	104 (100%)	0	0	100	100
26	T	218/249 (88%)	211 (97%)	6 (3%)	1 (0%)	24	55
27	U	180/439 (41%)	152 (84%)	26 (14%)	2 (1%)	11	39
28	V	163/291 (56%)	139 (85%)	24 (15%)	0	100	100
31	c	125/929 (14%)	116 (93%)	9 (7%)	0	100	100
32	k	96/211 (46%)	91 (95%)	5 (5%)	0	100	100
33	m	85/124 (68%)	82 (96%)	3 (4%)	0	100	100
34	o	1413/1970 (72%)	1369 (97%)	44 (3%)	0	100	100
35	p	1129/1174 (96%)	1089 (96%)	40 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
36	q	253/275 (92%)	241 (95%)	12 (5%)	0	100	100
37	r	126/142 (89%)	122 (97%)	4 (3%)	0	100	100
38	s	207/210 (99%)	203 (98%)	4 (2%)	0	100	100
39	t	77/127 (61%)	76 (99%)	1 (1%)	0	100	100
40	u	169/172 (98%)	166 (98%)	3 (2%)	0	100	100
41	v	146/150 (97%)	143 (98%)	3 (2%)	0	100	100
42	w	112/125 (90%)	103 (92%)	9 (8%)	0	100	100
43	x	62/67 (92%)	61 (98%)	1 (2%)	0	100	100
44	y	113/117 (97%)	112 (99%)	1 (1%)	0	100	100
45	z	42/58 (72%)	41 (98%)	1 (2%)	0	100	100
All	All	13465/22931 (59%)	12871 (96%)	577 (4%)	17 (0%)	49	75

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
6	5	48	GLU
14	E	523	VAL
17	H	128	PRO
14	e	522	ASP
19	j	124	GLU

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	216/283 (76%)	213 (99%)	3 (1%)	59	73
2	1	241/484 (50%)	241 (100%)	0	100	100
3	2	295/352 (84%)	295 (100%)	0	100	100
4	3	234/272 (86%)	229 (98%)	5 (2%)	47	67
5	4	346/399 (87%)	337 (97%)	9 (3%)	40	64
6	5	48/64 (75%)	48 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
7	6	536/688 (78%)	535 (100%)	1 (0%)	87	88
8	7	624/664 (94%)	623 (100%)	1 (0%)	87	88
9	8	259/299 (87%)	258 (100%)	1 (0%)	84	84
10	9	256/296 (86%)	254 (99%)	2 (1%)	73	79
11	A	496/1665 (30%)	452 (91%)	44 (9%)	9	32
12	B	876/1083 (81%)	862 (98%)	14 (2%)	55	72
13	D	145/815 (18%)	139 (96%)	6 (4%)	27	55
13	d	146/815 (18%)	144 (99%)	2 (1%)	59	73
14	E	480/657 (73%)	467 (97%)	13 (3%)	39	63
14	e	475/657 (72%)	462 (97%)	13 (3%)	39	63
15	F	328/574 (57%)	303 (92%)	25 (8%)	12	37
15	f	322/574 (56%)	312 (97%)	10 (3%)	35	60
16	G	132/322 (41%)	124 (94%)	8 (6%)	17	45
17	H	181/270 (67%)	173 (96%)	8 (4%)	25	54
18	I	106/235 (45%)	106 (100%)	0	100	100
18	i	107/235 (46%)	107 (100%)	0	100	100
19	J	79/154 (51%)	72 (91%)	7 (9%)	9	32
19	j	83/154 (54%)	82 (99%)	1 (1%)	63	75
20	L	71/141 (50%)	66 (93%)	5 (7%)	14	41
20	l	98/141 (70%)	97 (99%)	1 (1%)	68	76
21	O	90/98 (92%)	90 (100%)	0	100	100
22	P	154/293 (53%)	154 (100%)	0	100	100
23	Q	105/324 (32%)	105 (100%)	0	100	100
24	R	215/268 (80%)	214 (100%)	1 (0%)	81	83
25	S	93/448 (21%)	93 (100%)	0	100	100
26	T	196/218 (90%)	193 (98%)	3 (2%)	57	72
27	U	167/373 (45%)	163 (98%)	4 (2%)	43	65
28	V	150/261 (58%)	149 (99%)	1 (1%)	76	80
31	c	113/833 (14%)	111 (98%)	2 (2%)	51	70
32	k	87/182 (48%)	87 (100%)	0	100	100
33	m	80/106 (76%)	71 (89%)	9 (11%)	5	21

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
34	o	1254/1748 (72%)	1247 (99%)	7 (1%)	78	81
35	p	993/1027 (97%)	991 (100%)	2 (0%)	87	88
36	q	234/252 (93%)	234 (100%)	0	100	100
37	r	118/126 (94%)	118 (100%)	0	100	100
38	s	191/192 (100%)	191 (100%)	0	100	100
39	t	69/111 (62%)	69 (100%)	0	100	100
40	u	152/153 (99%)	151 (99%)	1 (1%)	76	80
41	v	129/131 (98%)	129 (100%)	0	100	100
42	w	103/112 (92%)	103 (100%)	0	100	100
43	x	53/56 (95%)	53 (100%)	0	100	100
44	y	104/106 (98%)	104 (100%)	0	100	100
45	z	41/55 (74%)	41 (100%)	0	100	100
All	All	12071/19766 (61%)	11862 (98%)	209 (2%)	52	71

5 of 209 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
15	F	326	HIS
19	J	164	SER
34	o	1291	ASN
15	F	368	THR
16	G	182	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 191 such sidechains are listed below:

Mol	Chain	Res	Type
22	P	229	GLN
15	f	134	HIS
25	S	32	ASN
31	c	50	HIS
32	k	202	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).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
47	SF4	7	1000	8	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
47	SF4	7	1000	8	-	-	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.

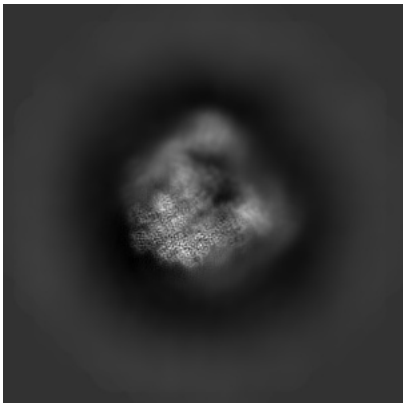
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-31111. These allow visual inspection of the internal detail of the map and identification of artifacts.

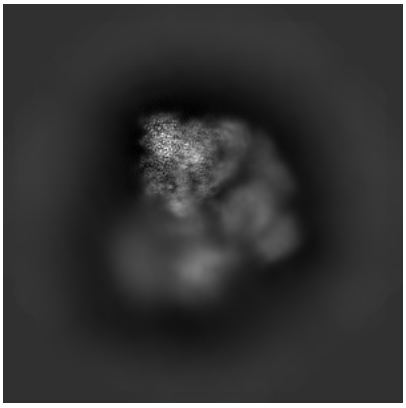
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

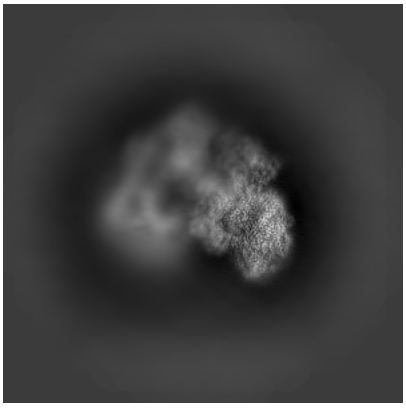
6.1.1 Primary map



X



Y

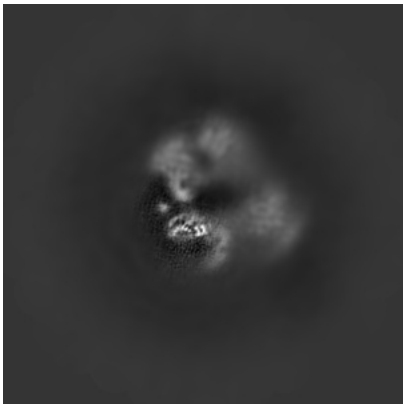


Z

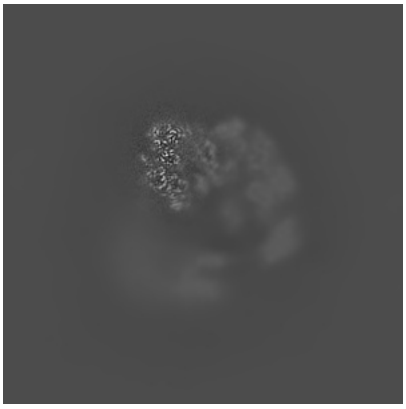
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

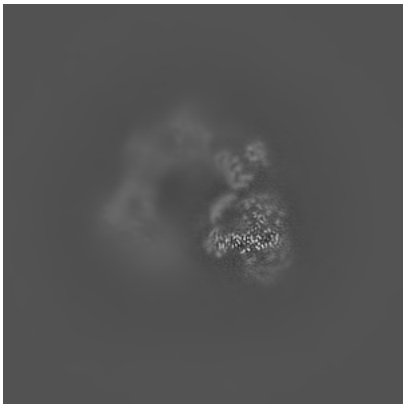
6.2.1 Primary map



X Index: 256



Y Index: 256

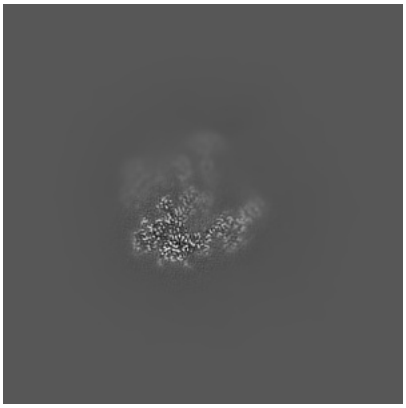


Z Index: 256

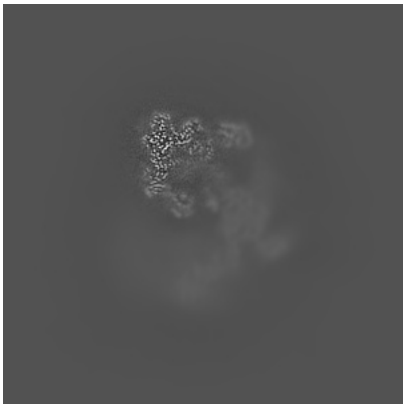
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

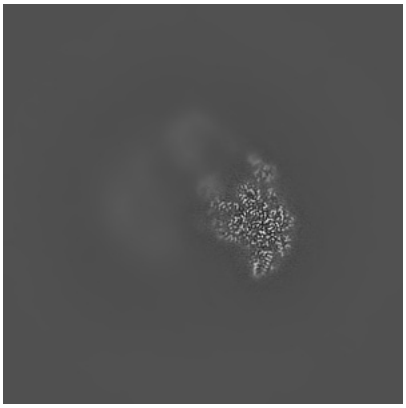
6.3.1 Primary map



X Index: 331



Y Index: 229

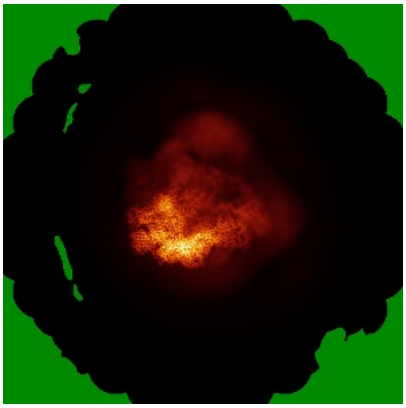


Z Index: 206

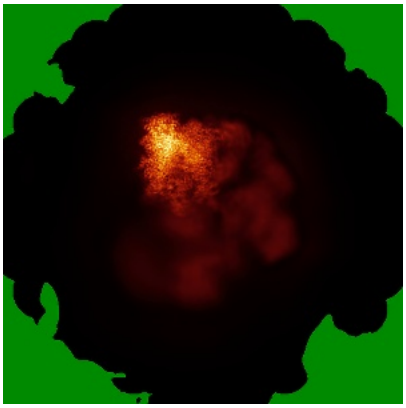
The images above show the largest variance slices of the map in three orthogonal directions.

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

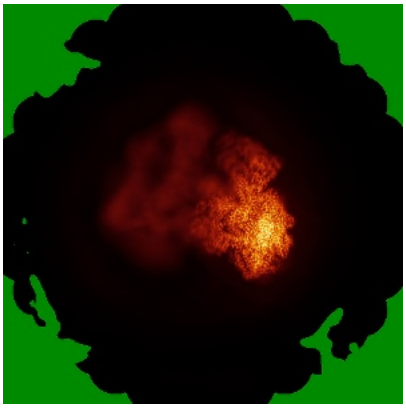
6.4.1 Primary map



X



Y



Z

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

X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.3. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

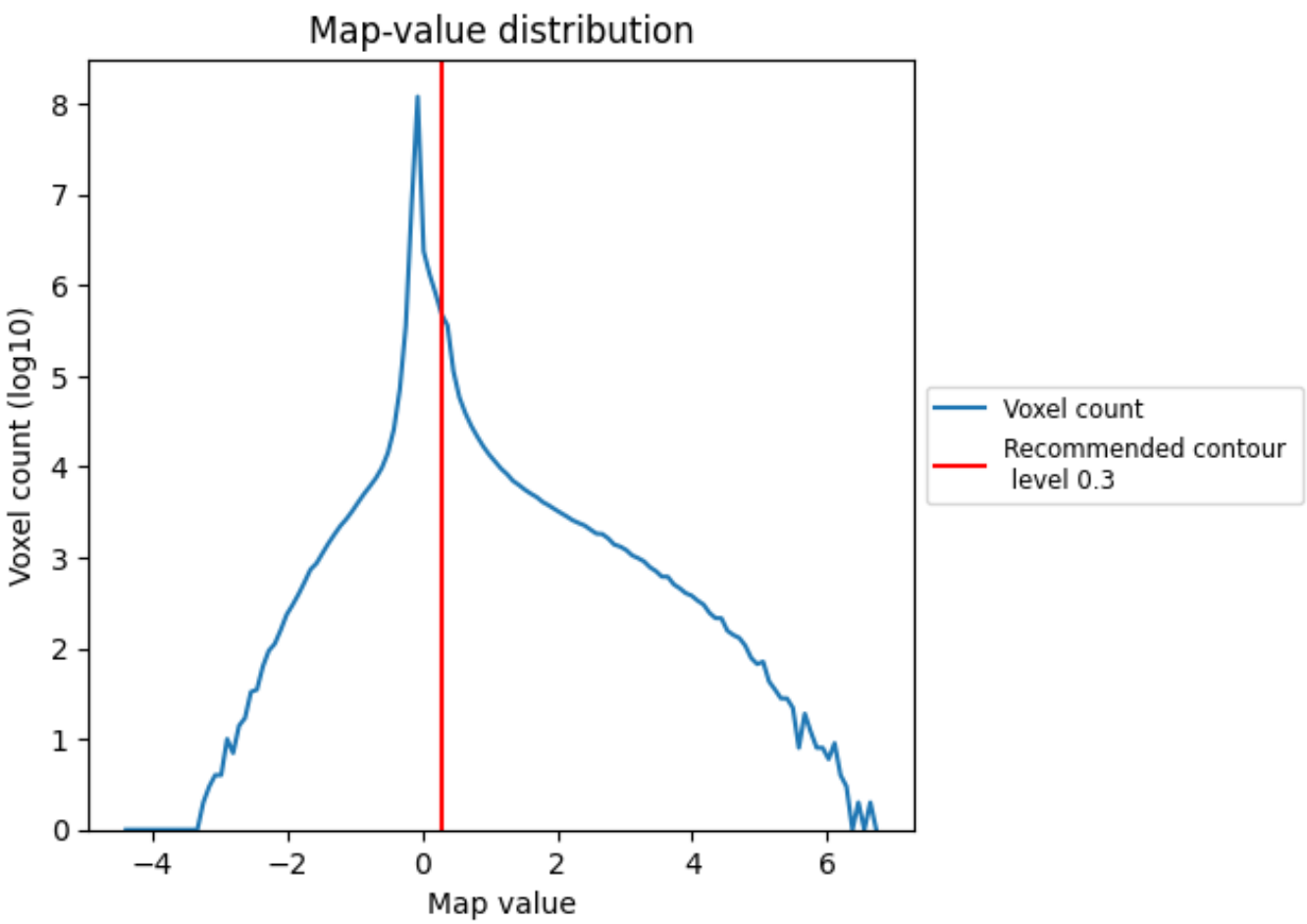
6.6 Mask visualisation (i)

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

7 Map analysis (i)

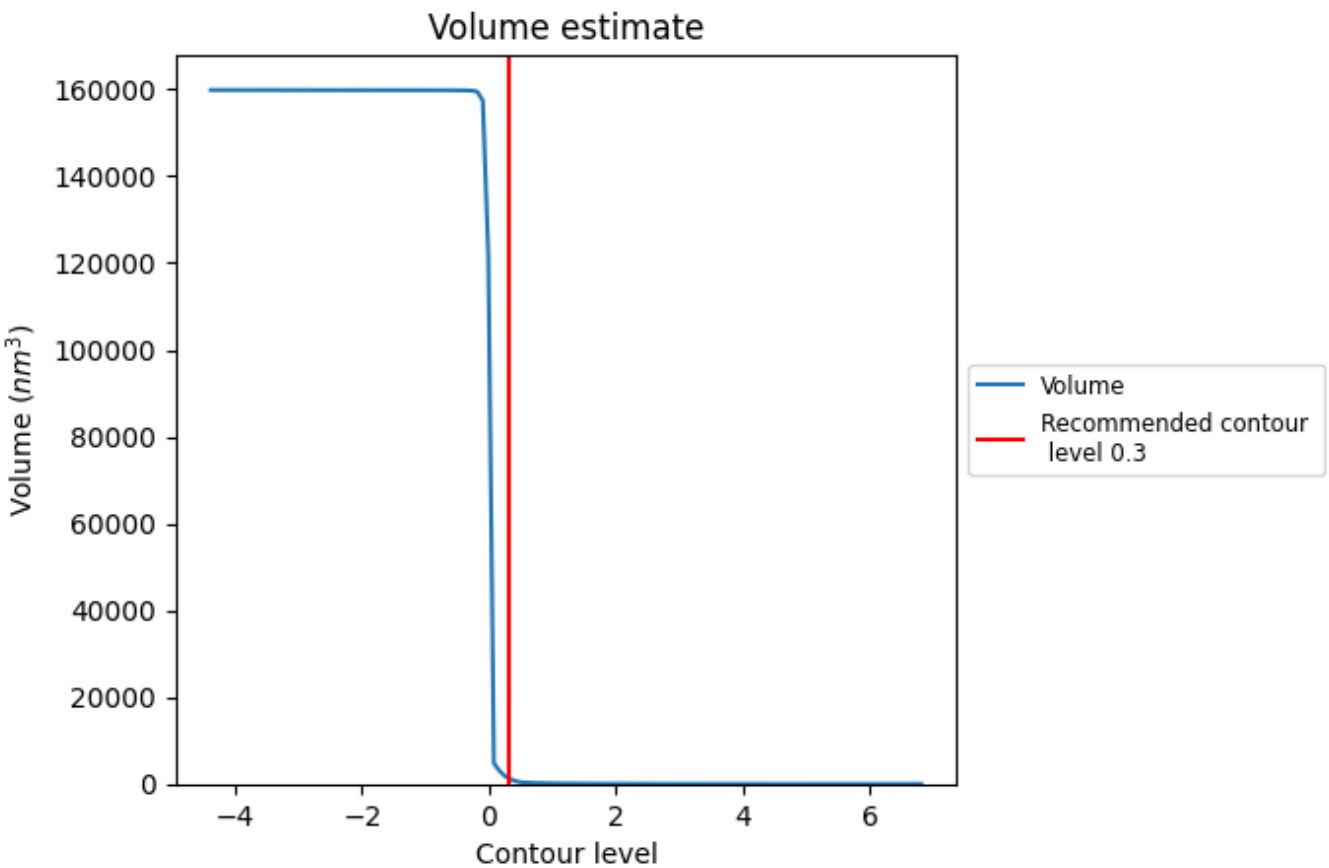
This section contains the results of statistical analysis of the map.

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.

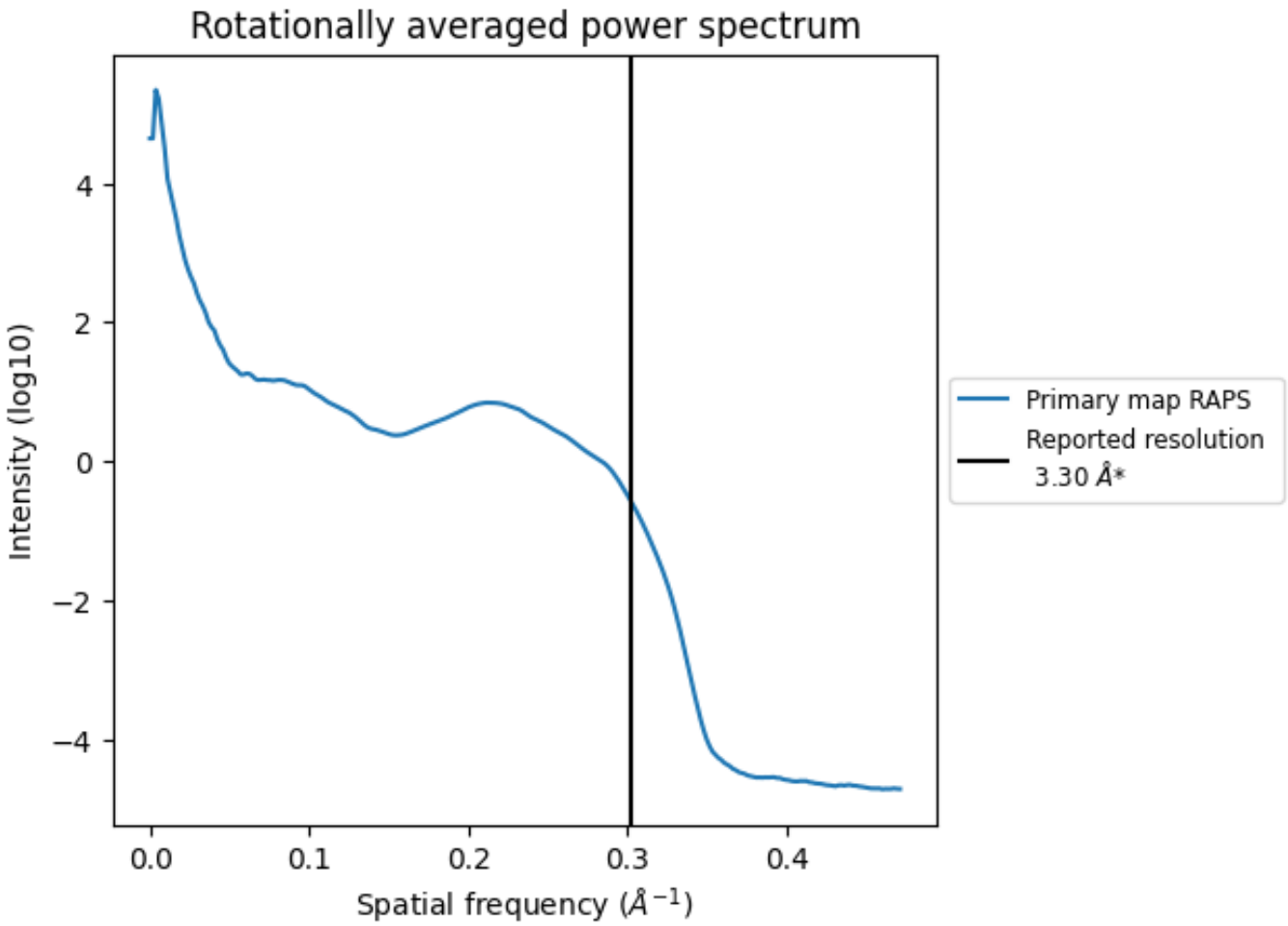
7.2 Volume estimate (i)



The volume at the recommended contour level is 1427 nm^3; this corresponds to an approximate mass of 1289 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 0.303 Å^{-1}

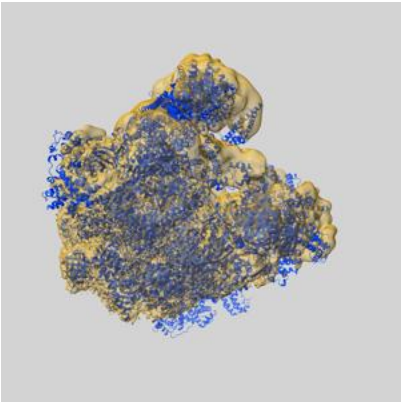
8 Fourier-Shell correlation (i)

This section was not generated. No FSC curve or half-maps provided.

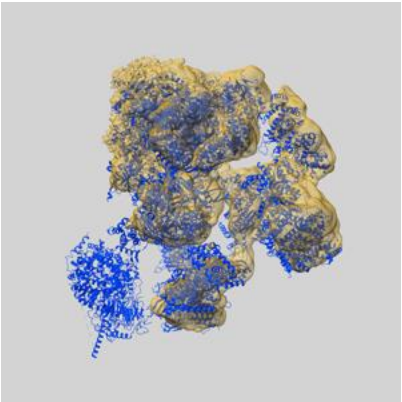
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-31111 and PDB model 7EGB. Per-residue inclusion information can be found in section 3 on page 14.

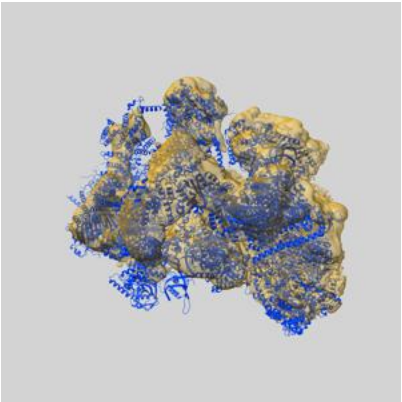
9.1 Map-model overlay (i)



X



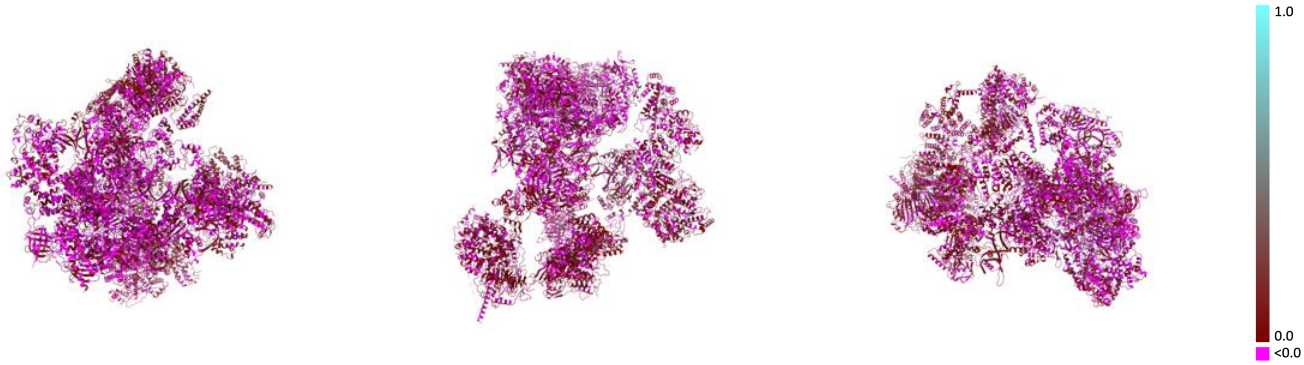
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.3 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.

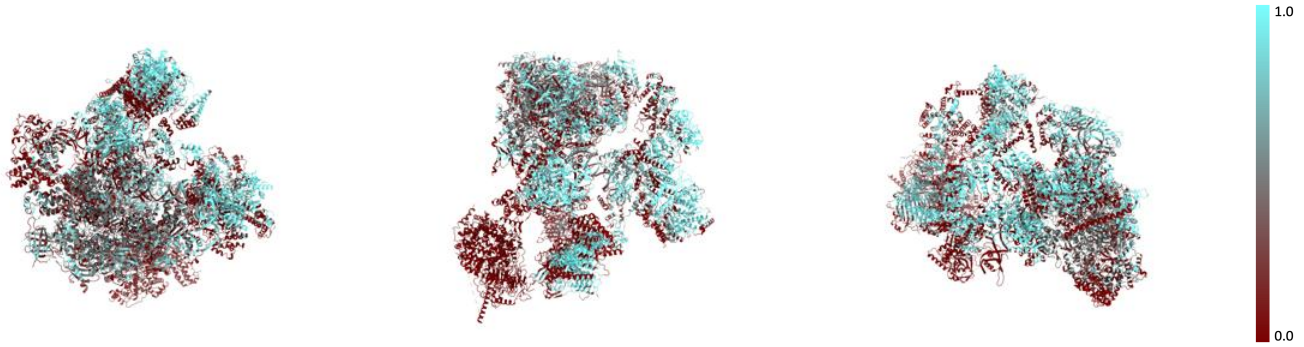
9.2 Q-score mapped to coordinate model (i)



X Y Z

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.

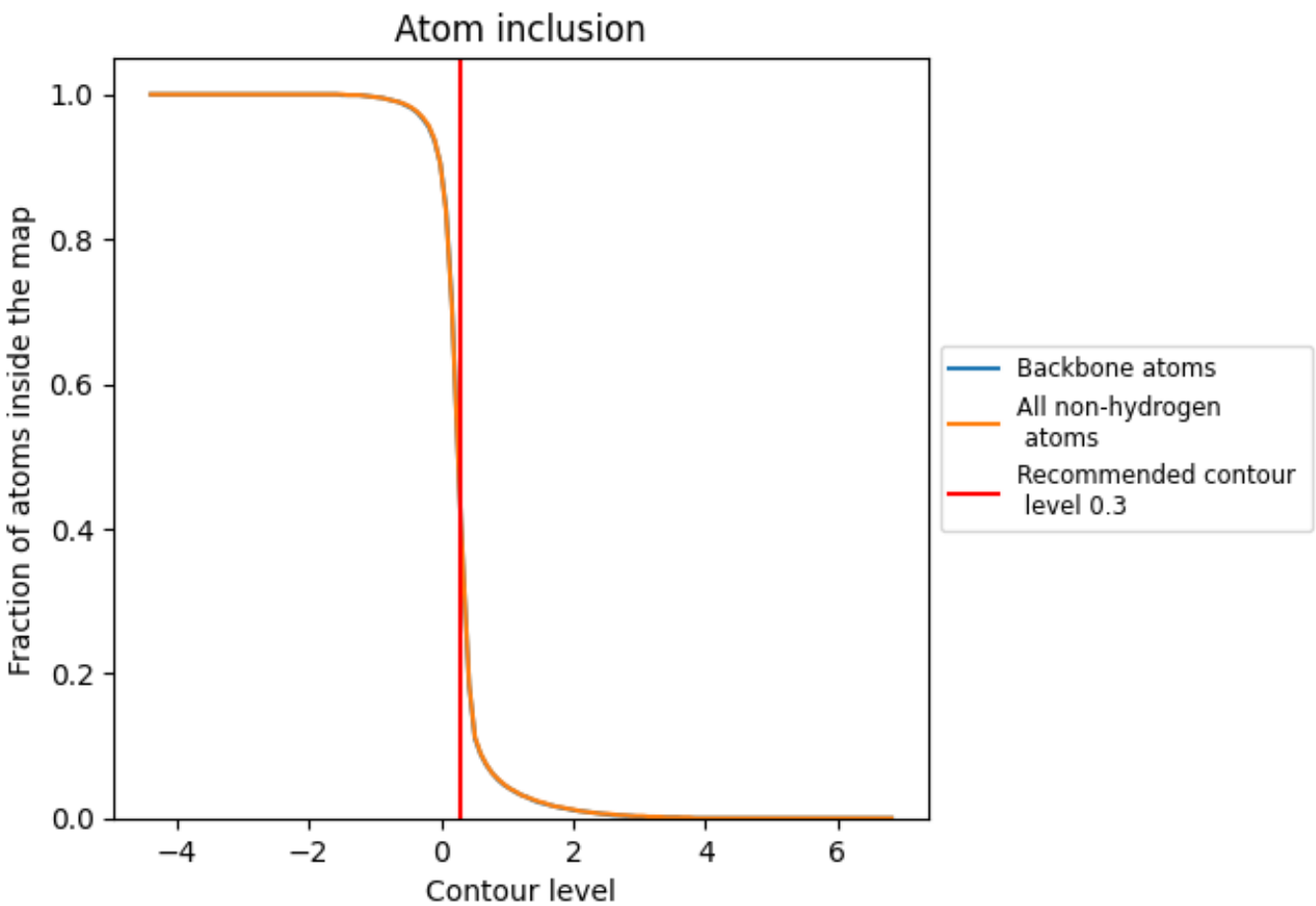
9.3 Atom inclusion mapped to coordinate model (i)



X Y Z

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 (0.3).

9.4 Atom inclusion (i)



At the recommended contour level, 41% of all backbone atoms, 41% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary (i)

The table lists the average atom inclusion at the recommended contour level (0.3) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.4140	0.0120
0	0.0890	-0.0020
1	0.4320	0.0430
2	0.5810	0.0260
3	0.7830	0.0500
4	0.5720	0.0510
5	0.3110	-0.0010
6	0.6730	0.0420
7	0.5860	0.0310
8	0.1610	0.0120
9	0.2780	0.0070
A	0.1510	0.0250
B	0.6230	0.0280
D	0.5050	0.0280
E	0.4020	0.0310
F	0.3590	0.0460
G	0.0300	0.0170
H	0.6440	0.0390
I	0.4490	-0.0180
J	0.8020	0.0260
L	0.3170	-0.0230
O	0.5370	-0.0150
P	0.7170	-0.0160
Q	0.5570	0.0160
R	0.4940	-0.0010
S	0.5220	0.0290
T	0.4100	0.0270
U	0.4650	-0.0380
V	0.6180	0.0190
X	0.4470	0.0500
Y	0.4070	0.0480
c	0.0000	0.0340
d	0.0000	0.0100
e	0.0000	0.0340
f	0.2680	0.0380

Continued on next page...

Continued from previous page...

Chain	Atom inclusion	Q-score
i	0.0000	0.0300
j	0.0000	0.0070
k	0.0000	0.0130
l	0.0000	0.0230
m	0.0000	0.0150
o	0.4030	-0.0240
p	0.4300	-0.0300
q	0.4700	0.0000
r	0.5270	-0.0370
s	0.4330	-0.0260
t	0.3740	-0.0310
u	0.6640	-0.0100
v	0.4060	-0.0110
w	0.3780	-0.0350
x	0.4220	0.0140
y	0.4850	-0.0190
z	0.4710	0.0040