wwPDB EM Validation Summary Report ⓘ

Dec 26, 2023 – 11:31 AM JST

PDB ID	:	8WAK
EMDB ID	:	EMD-37395
Title	:	Structure of transcribing complex 2 (TC2), the initially transcribing complex with Pol II positioned 2nt downstream of TSS.
Authors	:	Chen, X.; Liu, W.; Wang, Q.; Wang, X.; Ren, Y.; Qu, X.; Li, W.; Xu, Y.
Deposited on	:	2023-09-07
Resolution	:	5.47 Å(reported)

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

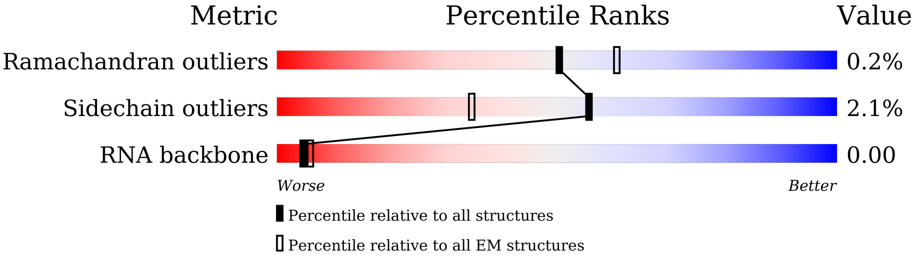
EMDB validation analysis	:	0.0.1.dev70
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance

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

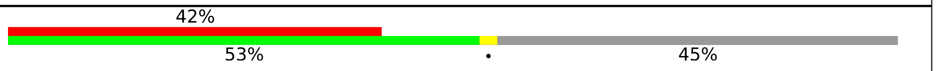
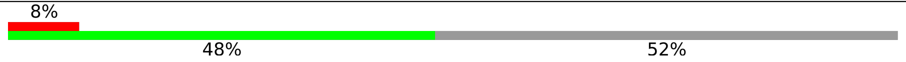
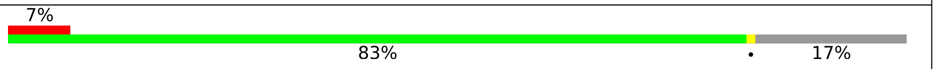
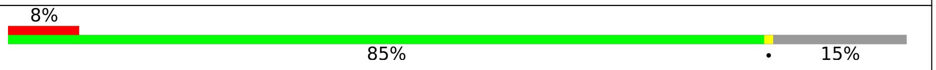
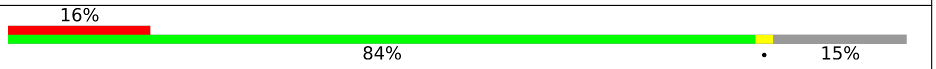
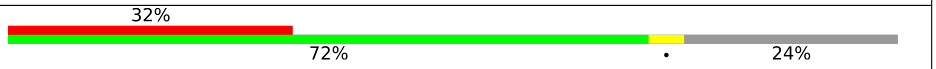
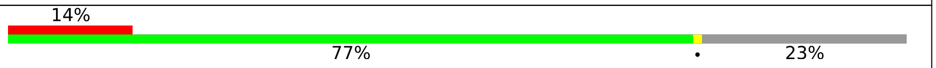
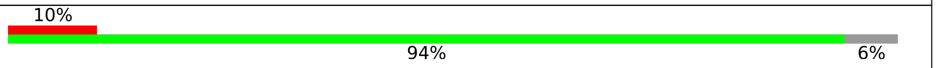
The reported resolution of this entry is 5.47 Å.

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)
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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

Mol	Chain	Length	Quality of chain
1	0	309	
2	1	548	
3	2	395	
4	3	308	
5	4	462	
6	5	71	
7	6	782	
8	7	760	

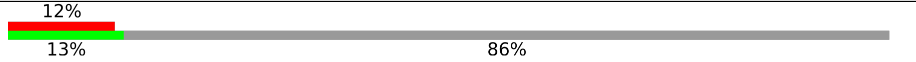
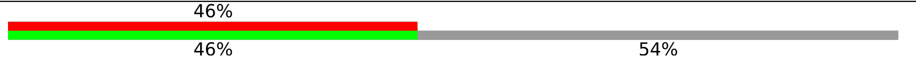
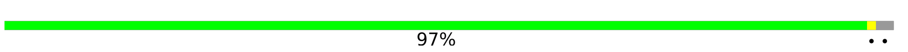
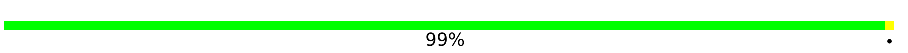
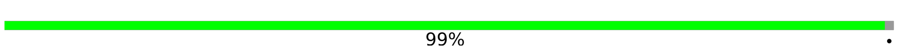
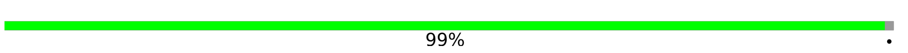
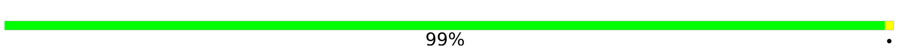
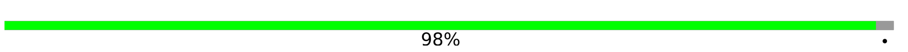
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	N	8	50% 50%
10	Z	2	50% 50%
11	A	1872	24% 28% • 70%
12	B	1199	9% 79% • 20%
13	D	1085	9% 14% • 85%
13	d	1085	15% 14% 85%
14	E	800	12% 66% • 32%
14	e	800	52% 66% • 33%
15	F	677	39% 58% • 39%
15	f	677	19% 58% • 40%
16	G	349	41% 39% • 59%
17	H	310	25% 64% • 33%
18	I	264	23% 45% 55%
18	i	264	39% 46% 54%
19	J	218	16% 38% • 59%
19	j	218	41% 42% • 56%
20	L	161	43% 44% • 53%
20	l	161	66% 66% • 34%
21	O	109	• 89% • 9%
22	P	339	53% 47%
23	Q	376	• 30% 70%
24	R	316	87% • 11%
25	S	517	30% • 66%
26	T	249	88% • 11%
27	U	439	• 41% 58%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
28	V	291	
29	X	99	
30	Y	99	
31	c	929	
32	k	211	
33	m	124	
34	o	1970	
35	p	1174	
36	q	275	
37	r	142	
38	s	210	
39	t	127	
40	u	172	
41	v	150	
42	w	125	
43	x	67	
44	y	117	
45	z	58	

2 Entry composition

There are 49 unique types of molecules in this entry. The entry contains 108642 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	171	1420	882	249	279	10	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 Alpha-amanitin.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
9	N	8	64	39	10	14	1	0	0

- Molecule 10 is a RNA chain called RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
10	Z	2	55	21	10	20	4	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	164	1366	851	256	255	4	0	0

Continued on next page...

Continued from previous page...

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

- Molecule 23 is a protein called Transcription initiation factor IIA beta chain.

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

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

Mol	Chain	Residues	Atoms					AltConf	Trace
24	R	282	Total	C	N	O	S	0	0
			2167	1353	384	412	18		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
25	S	176	Total	C	N	O	S	0	0
			1461	921	268	268	4		

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

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

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

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

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

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

- Molecule 29 is a DNA chain called non-template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	X	63	Total	C	N	O	P	0	0
			1301	616	245	377	63		

- Molecule 30 is a DNA chain called template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	Y	77	Total	C	N	O	P	0	0
			1571	747	288	459	77		

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

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

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

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

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

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

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

Mol	Chain	Residues	Atoms					AltConf	Trace
34	o	1452	Total 11510	C 7235	N 2056	O 2146	S 73	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
35	p	1149	Total 9175	C 5796	N 1614	O 1701	S 64	0	0

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

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

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

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

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

Mol	Chain	Residues	Atoms					AltConf	Trace
38	s	209	Total 1720	C 1089	N 300	O 323	S 8	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
39	t	82	Total 657	C 418	N 113	O 121	S 5	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
40	u	171	Total 1351	C 875	N 219	O 249	S 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
41	v	148	Total 1186	C 750	N 194	O 237	S 5	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
42	w	114	Total 927	C 571	N 166	O 179	S 11	0	0

- Molecule 43 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

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

- Molecule 44 is a protein called DNA-directed RNA polymerase II subunit RPB11-a.

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

- Molecule 45 is a protein called RPB12.

Mol	Chain	Residues	Atoms					AltConf	Trace
45	z	44	Total 372	C 231	N 72	O 63	S 6	0	0

- Molecule 46 is ZINC ION (three-letter code: ZN) (formula: Zn).

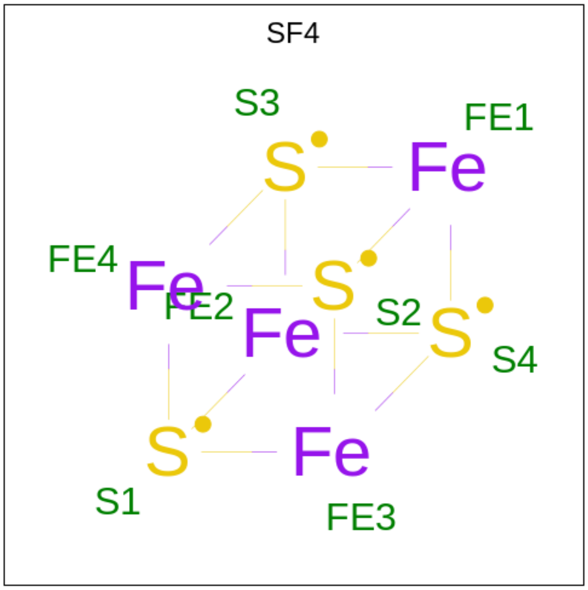
Mol	Chain	Residues	Atoms		AltConf
46	0	2	Total 2	Zn 2	0
46	2	3	Total 3	Zn 3	0
46	3	2	Total 2	Zn 2	0

Continued on next page...

Continued from previous page...

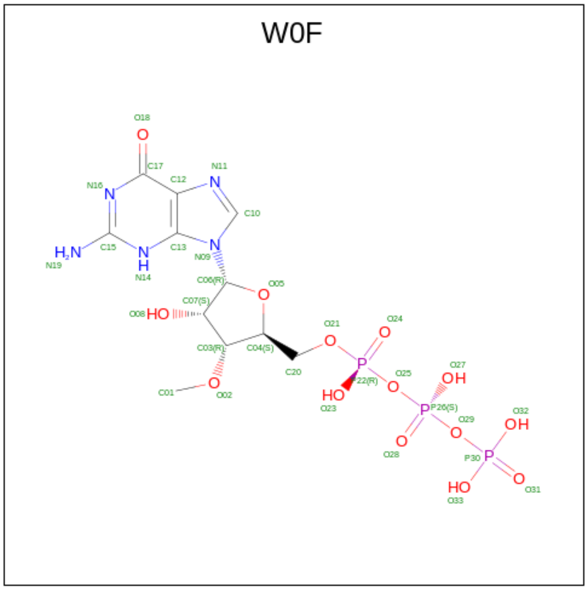
Mol	Chain	Residues	Atoms	AltConf
46	R	1	Total Zn 1 1	0
46	U	1	Total Zn 1 1	0
46	o	2	Total Zn 2 2	0
46	p	1	Total Zn 1 1	0
46	q	1	Total Zn 1 1	0
46	w	2	Total Zn 2 2	0
46	x	1	Total Zn 1 1	0
46	z	1	Total Zn 1 1	0

- Molecule 47 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe_4S_4) (labeled as "Ligand of Interest" by depositor).



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

- Molecule 48 is [[(2 {S},3 {R},4 {S},5 {R})-5-(2-azanyl-6-oxidanylidene-3 {H}-purin-9-yl)-3-methoxy-4-oxidanyl-oxolan-2-yl]methoxy-oxidanyl-phosphoryl] phosphono hydrogen phosphate (three-letter code: W0F) (formula: $C_{11}H_{18}N_5O_{14}P_3$).



Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
48	Z	1	33	11	5	14	3	0

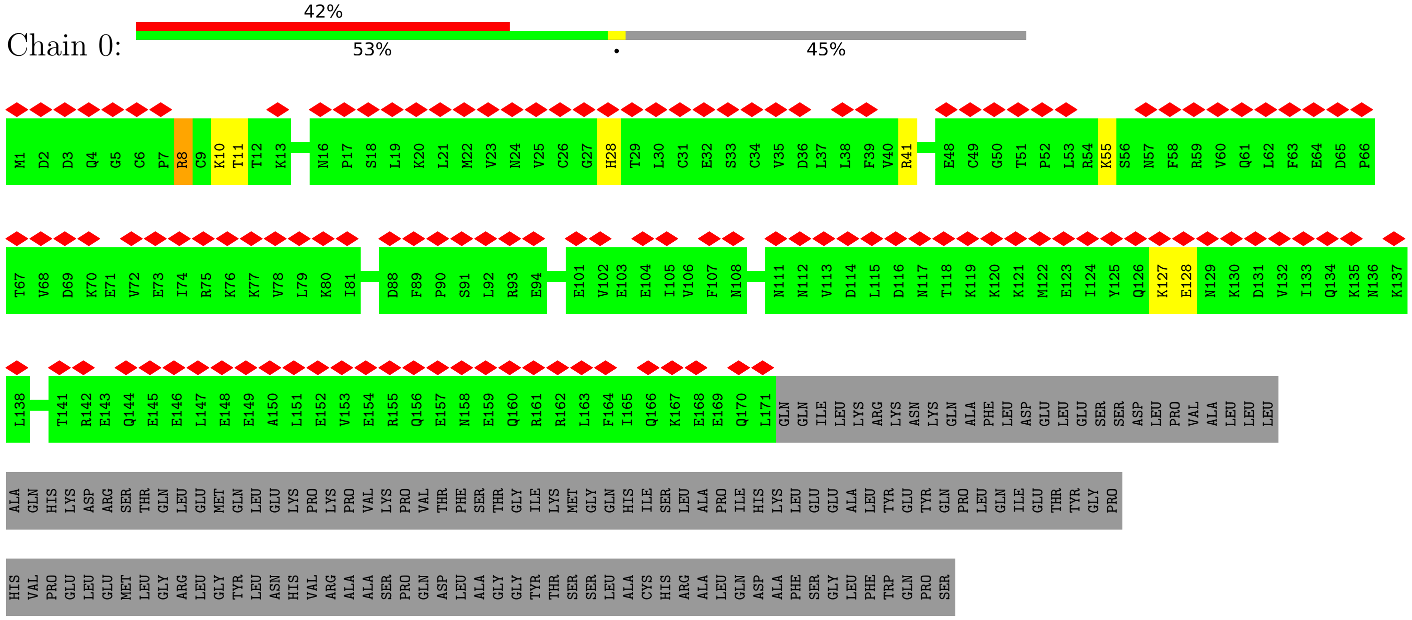
- Molecule 49 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

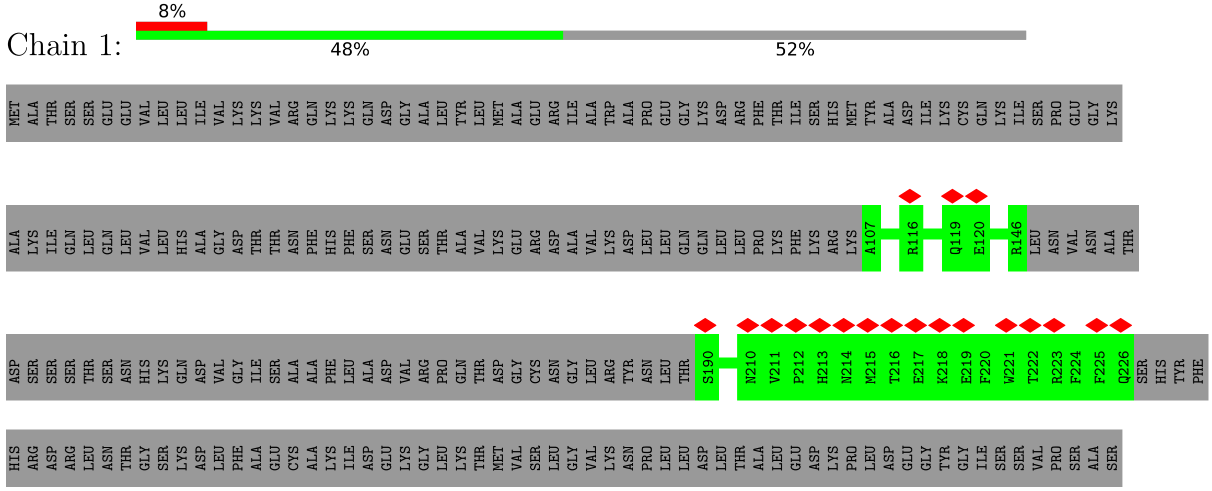
3 Residue-property plots ⓘ

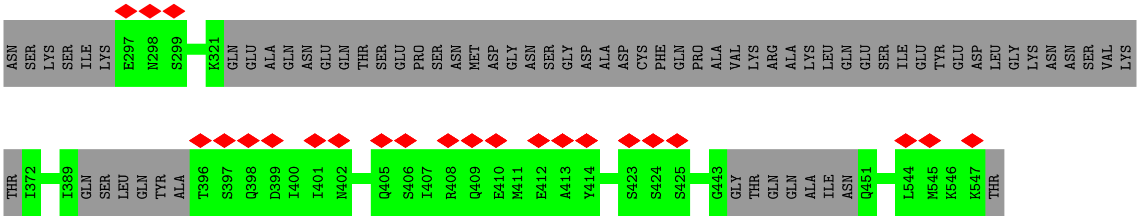
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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



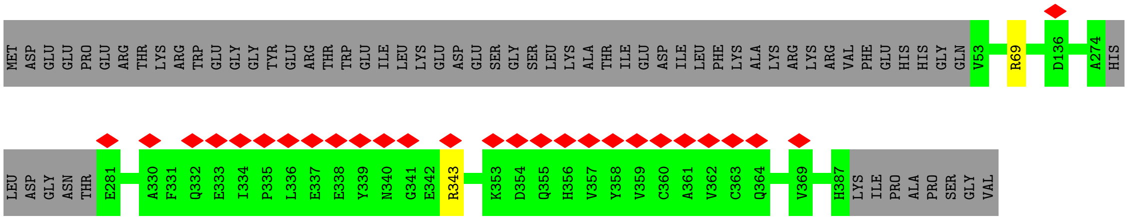
- Molecule 2: General transcription factor IIH subunit 1





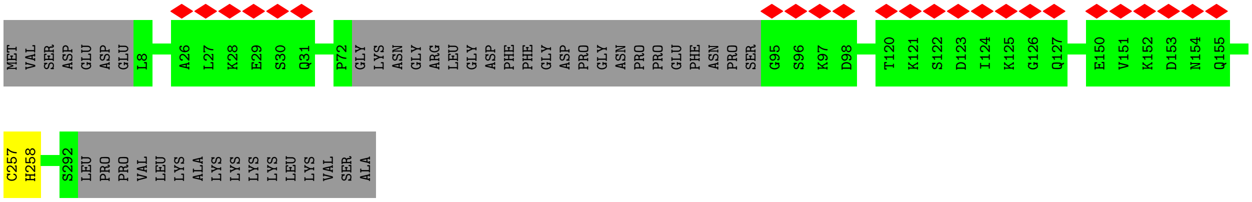
- Molecule 3: General transcription factor IIH subunit 2

Chain 2:



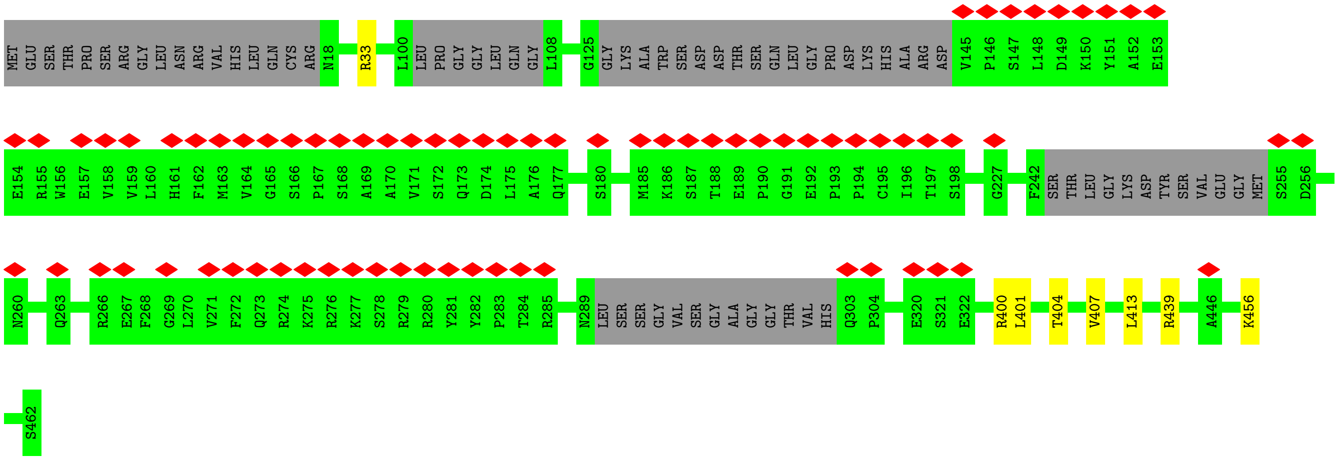
- Molecule 4: General transcription factor IIH subunit 3

Chain 3:



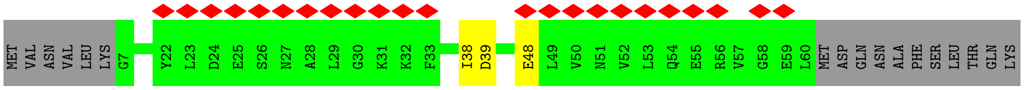
- Molecule 5: General transcription factor IIH subunit 4

Chain 4:



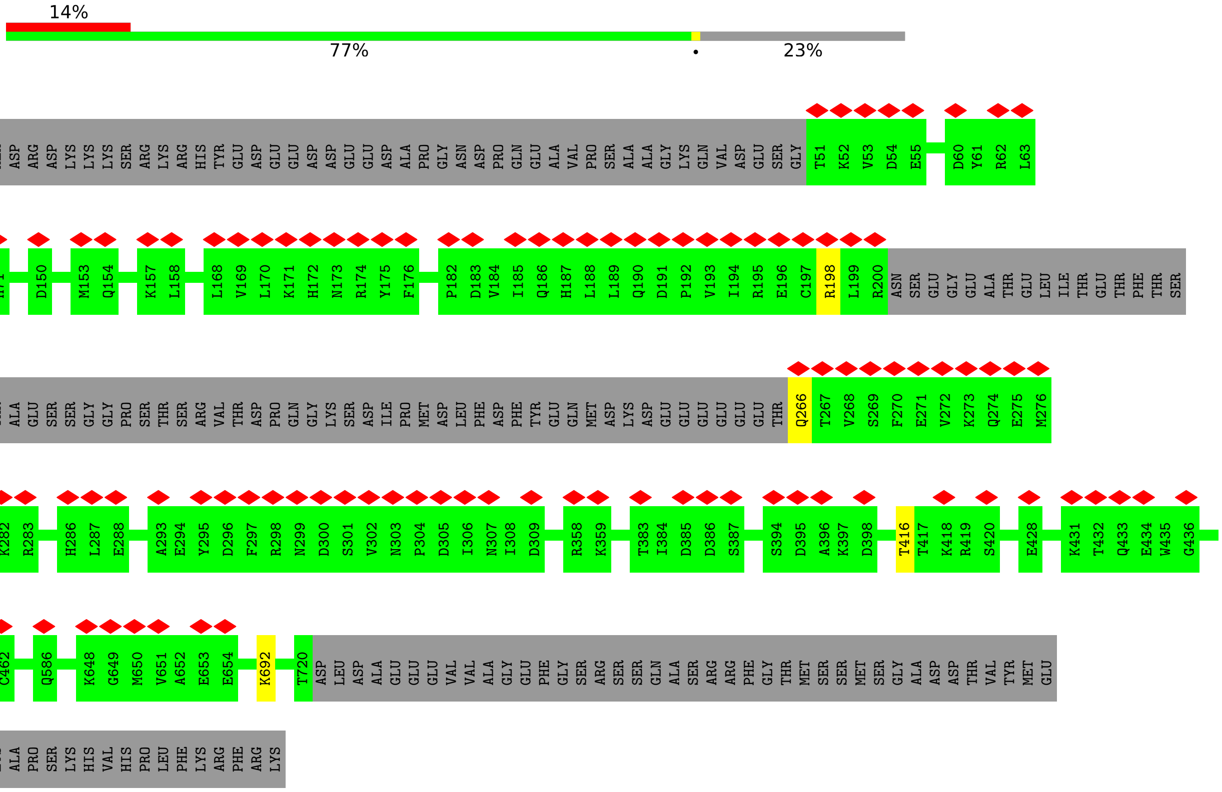
- Molecule 6: General transcription factor IIH subunit 5

Chain 5:



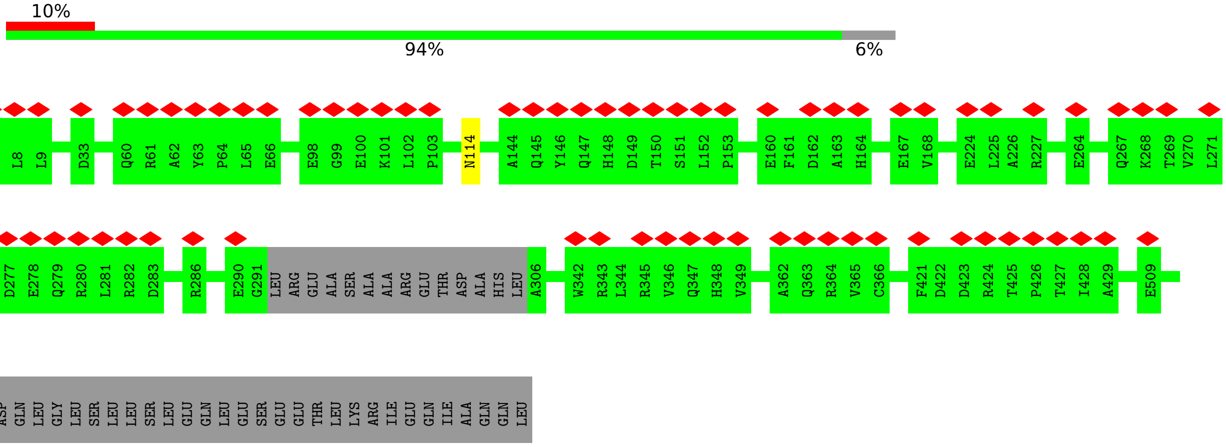
- Molecule 7: General transcription and DNA repair factor IIH helicase subunit XPB

Chain 6:



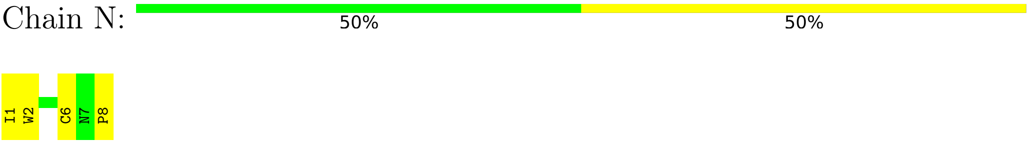
- Molecule 8: General transcription and DNA repair factor IIH helicase subunit XPD

Chain 7:



- Molecule 9: Alpha-amanitin

Chain N:



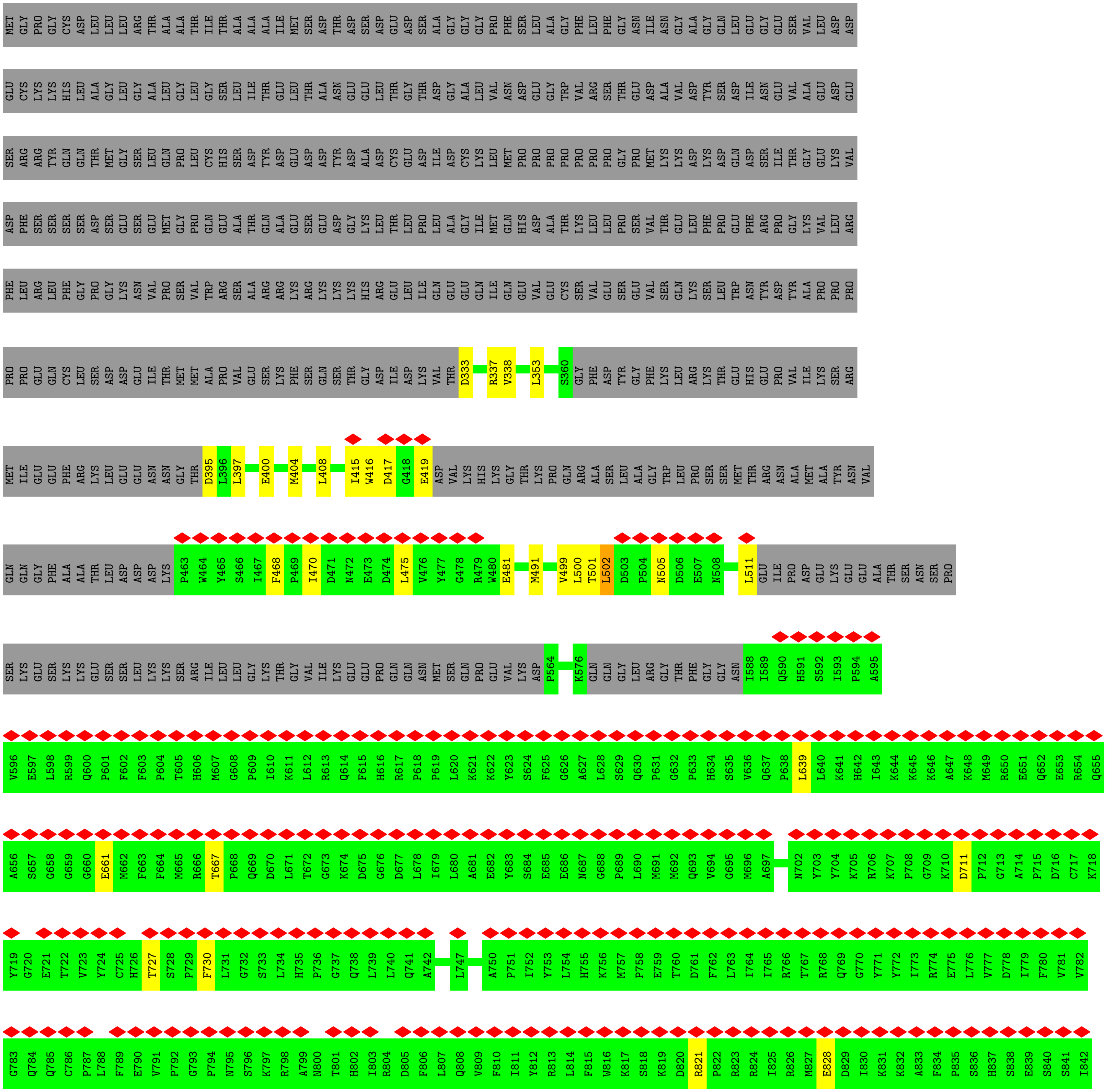
- Molecule 10: RNA

Chain Z:

ATP1 G2

- Molecule 11: Transcription initiation factor TFIID subunit 1

Chain A:



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

GLU CYS LYS LYS HIS LEU ALA GLY LEU GLY ALA LEU GLY LEU GLY SER LEU ILE THR GLU LEU THR ALA ASN GLU GLU LEU THR GLY THR ASP GLY ALA LEU VAL ASN ASP GLU GLY TRP VAL ARG SER THR GLU ASP ALA VAL ASP TYR SER ASP ILE ASN GLU VAL ALA GLU ASP GLU

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

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

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

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 R337 V338 L353 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 PHE ARG LYS LEU GLU GLU ASN ASN GLY THR D395 L396 L397 E400 M404 L408 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 M491 V499 L500 T501 L502 D503 P504 N505 D506 E507 N508 L511 GLU ILE PRO ASP GLU LYS GLU GLU ALA THR SER ASN SER PRO

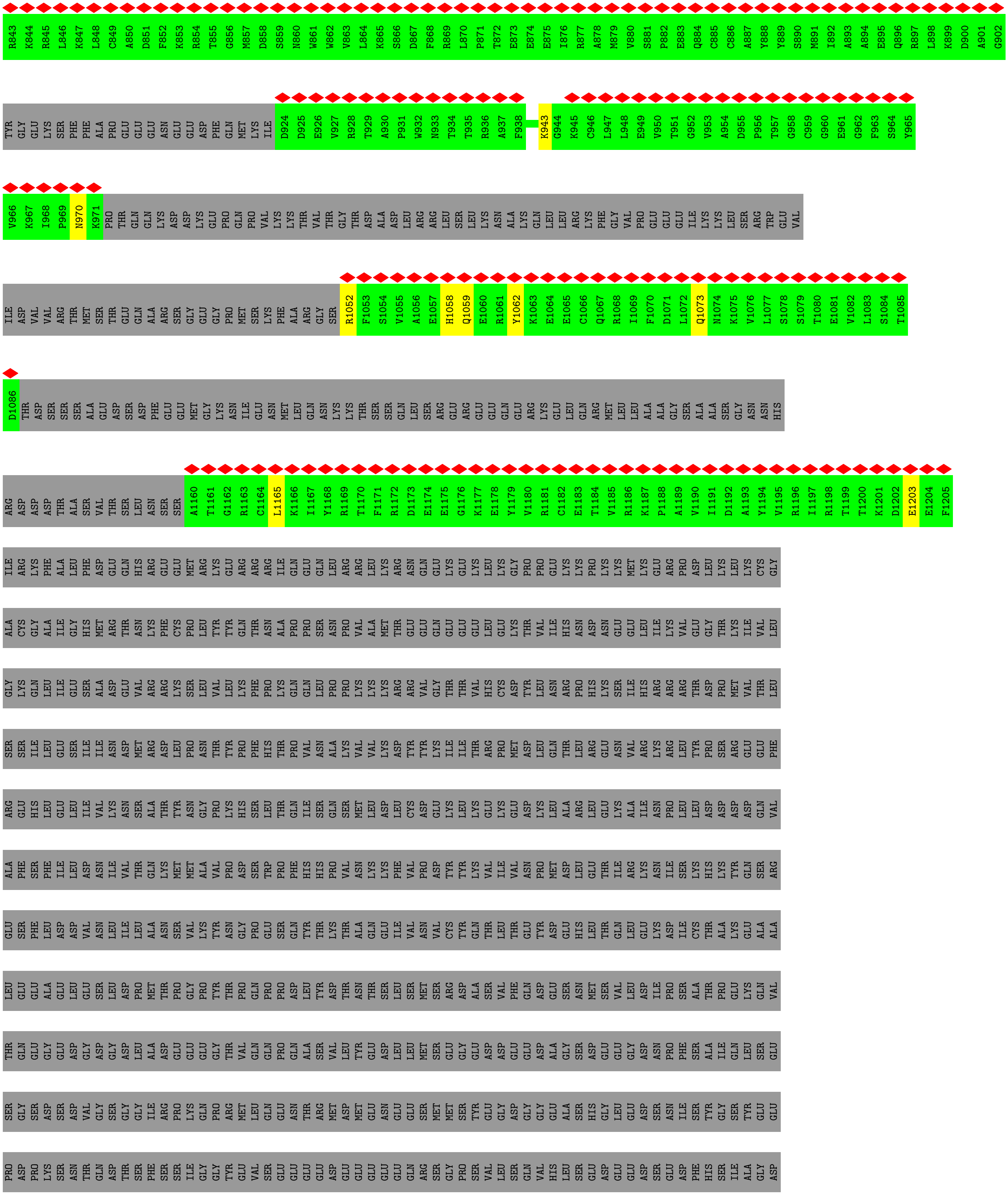
SER LYS GLU SER LYS LYS GLU SER SER LEU LYS LYS SER ARG ILE LEU LEU GLY LYS THR GLY VAL ILE LYS GLU PRO GLN GLN ASN MET SER GLN PRO GLU VAL LYS ASP P564 K576 GLN GLN GLY LEU ARG GLY THR PHE GLY GLY ASN I588 I589 Q590 H591 S592 I593 P594 A595

V596 E597 L598 R599 Q600 P601 F602 F603 P604 T605 H606 M607 G608 P609 I610 K611 L612 R613 Q614 F615 H616 R617 P618 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 E685 E686 N687 G688 P689 L690 M691 M692 Q693 V694 G695 M696 A697 N702 Y703 Y704 K705 R706 K707 P708 G709 K710 D711 P712 G713 A714 P715 D716 C717 K718

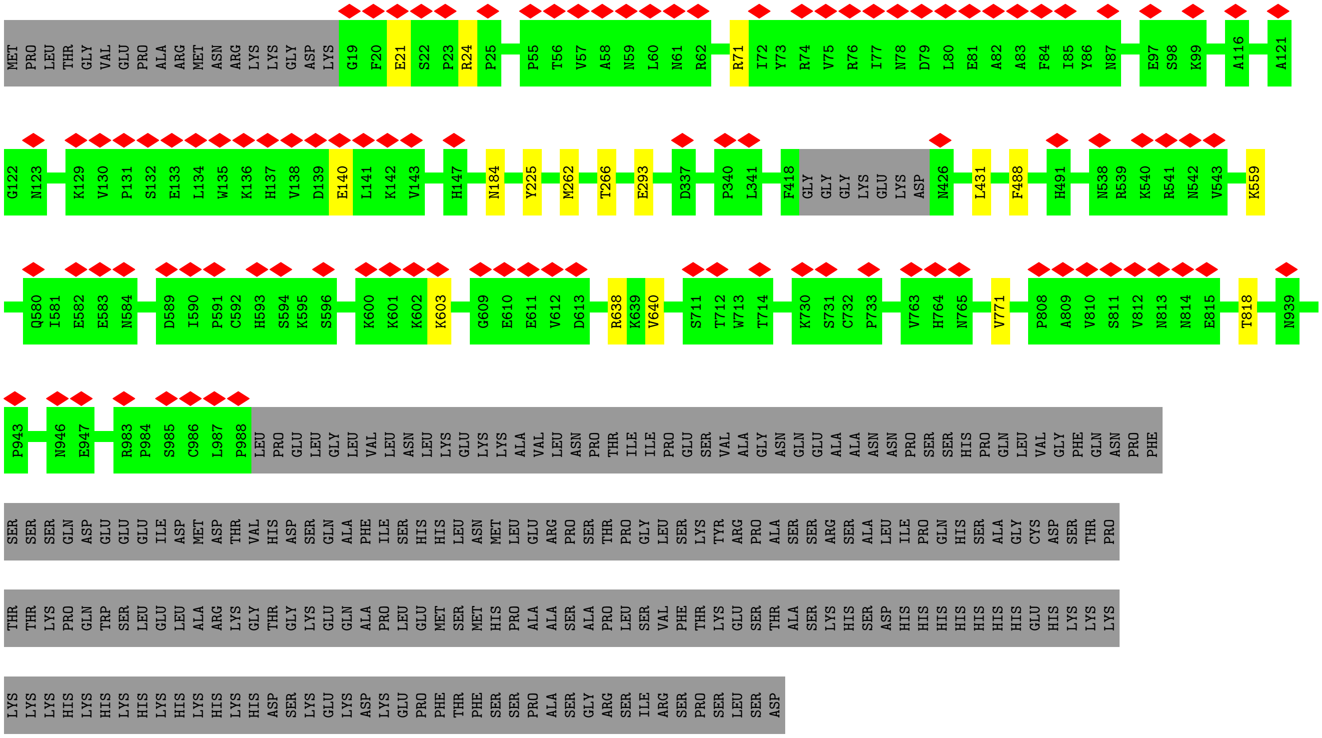
Y719 G720 E721 T722 V723 Y724 C725 H726 T727 S728 P729 F730 L731 G732 S733 L734 H735 P736 G737 Q738 L739 L740 Q741 A742 L747 A750 P751 I752 Y753 L754 H755 K756 M757 P758 E759 T760 D761 F762 L763 I764 I765 R766 T767 R768 Q769 G770 Y771 Y772 I773 R774 E775 L776 V777 D778 I779 F780 V781 V782

G783 Q784 Q785 C786 P787 L788 F789 E790 V791 P792 G793 P794 N795 S796 K797 R798 A799 N800 T801 H802 I803 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 C885 C886 A887 Y888 Y889 S890 M891 I892 A893 A894 E895 Q896 R897 L898 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 R936 A937 F938 K943 G944 K945 C946 L947 L948 E949 V950 T951 G952 V953 A954 D955 P956 T957 G958 C959 G960 E961 G962 F963 S964 Y965
V966 K967 I968 P969 N970 K971 PRO THR GLN GLN LYS ASP ASP LYS 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 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 ASP THR ALA SER VAL THR SER LEU ASN SER SER A1160 T1161 G1162 R1163 C1164 L1165 K1166 I1167 Y1168 R1169 T1170 F1171 R1172 D1173 E1174 E1175 G1176 K1177 E1178 Y1179 V1180 R1181 C1182 E1183 T1184 V1185 R1186 K1187 P1188 A1189 V1190 I1191 D1192 A1193 Y1194 V1195 R1196 I1197 R1198 T1199 T1200 K1201 D1202 E1203 E1204 F1205
ILE ARG LYS PHE ALA LEU PHE ASP GLU GLN HIS ARG GLU GLU MET ARG LYS GLU ARG ARG ILE GLN GLU GLN LEU ARG ARG LEU LYS ARG ASN GLN GLU LYS GLU LYS LEU LYS GLY PRO PRO GLU LYS LYS PRO LYS LYS MET LYS GLU ARG PRO ASP LEU LYS LEU LYS CYS GLY
ALA CYS GLY ALA ILE GLY HIS MET ARG THR ASN LYS PHE CYS PRO LEU TYR TYR GLN THR ASN ALA PRO PRO SER ASN PRO VAL ALA MET THR GLU GLU GLN GLU GLU GLU LEU GLU LYS THR VAL ILE HIS ASN ASP ASN GLU LEU ILE LYS VAL GLU GLY THR LYS ILE VAL LEU
GLY LYS GLN LEU ILE GLU SER ALA ASP GLU VAL ARG ARG LYS SER LEU VAL LEU LYS PHE PRO LYS GLN GLN LEU PRO PRO LYS LYS ARG ARG VAL GLY THR THR VAL HIS CYS ASP TYR LEU ASN ARG PRO HIS LYS SER ILE HIS ARG ARG ARG THR ASP PRO MET VAL THR LEU
SER SER ILE LEU GLU SER ILE ILE ASN ASP MET ARG ASP LEU PRO ASN THR TYR PRO PHE HIS THR PRO VAL ASN ALA LYS VAL VAL LYS ASP TYR TYR LYS ILE ILE THR ARG PRO MET ASP LEU GLN THR LEU ARG GLU ASN VAL ARG LYS ARG LEU TYR PRO SER ARG GLU GLU PHE
ARG GLU HIS LEU GLU LEU ILE VAL LYS ASN SER ALA THR TYR ASN GLY PRO LYS HIS SER LEU THR GLN ILE SER GLN SER MET LEU ASP LEU CYS ASP GLU LYS LEU LYS GLU LYS GLU ASP LYS LEU ALA ARG LEU GLU LYS ALA ILE ASN PRO LEU LEU ASP ASP ASP GLN VAL
ALA PHE SER PHE ILE LEU ASP ASN ILE VAL THR GLN LYS MET MET ALA VAL PRO ASP SER TRP PRO PHE HIS HIS PRO VAL ASN LYS LYS PHE VAL PRO ASP TYR TYR LYS VAL ILE VAL ASN PRO MET ASP LEU GLU THR ILE ARG LYS ASN ILE SER LYS HIS LYS TYR GLN SER ARG
GLU SER PHE LEU ASP ASP VAL ASN LEU ILE LEU ALA ASN SER VAL LYS TYR ASN GLY PRO GLU SER GLN TYR THR LYS THR ALA GLN GLU ILE VAL ASN VAL CYS TYR GLN THR LEU THR GLU TYR ASP GLU HIS LEU THR GLN LEU GLU LYS ASP ILE CYS THR ALA LYS GLU ALA ALA
LEU GLU GLU ALA GLU LEU GLU SER LEU ASP PRO MET THR PRO GLY PRO TYR THR PRO GLN PRO PRO ASP LEU TYR ASP THR ASN THR SER LEU SER MET SER ARG ASP ALA SER VAL PHE GLN ASP GLU SER ASN MET SER VAL LEU ASP ILE PRO SER ALA THR PRO GLU LYS GLN VAL
THR GLN GLU GLY GLU ASP GLY ASP GLY ASP LEU ALA ASP GLU GLU GLU GLY THR VAL GLN GLN PRO GLN ALA SER VAL LEU TYR GLU ASP LEU LEU MET SER GLU GLY GLU ASP ASP GLU GLU ASP ALA GLY SER ASP GLU GLU GLY ASP ASN PRO PHE SER ALA ILE GLN LEU SER GLU
SER GLY SER ASP SER ASP VAL GLY SER GLY GLY ILE ARG PRO LYS GLN PRO ARG MET LEU GLN GLU ASN THR ARG MET ASP MET GLU ASN GLU GLU SER MET MET SER TYR GLU GLY ASP GLY GLY GLU ALA SER HIS GLY LEU GLU ASP SER ASN ILE SER TYR GLY SER TYR GLU GLU
PRO ASP PRO LYS SER ASN THR GLN ASP THR SER PHE SER SER ILE GLY GLY TYR GLU VAL SER GLU GLU GLU GLU ASP GLU GLU GLU GLU GLU GLN ARG SER GLY PRO SER VAL LEU SER GLN VAL HIS LEU SER GLU ASP GLU ASP SER GLU ASP PHE HIS SER ILE ALA GLY ASP

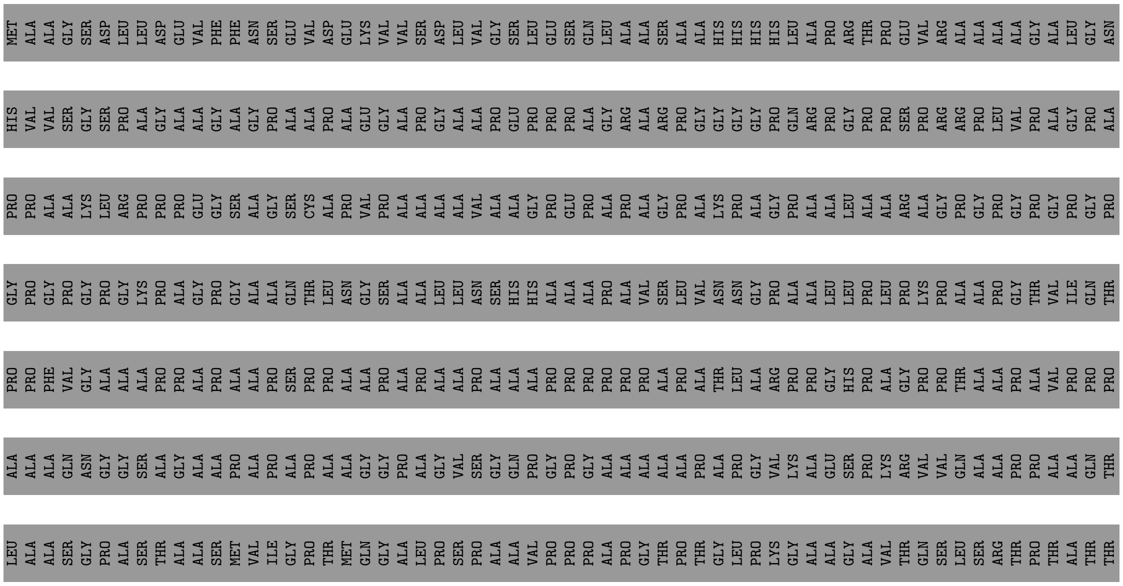
- Molecule 12: Transcription initiation factor TFIID subunit 2

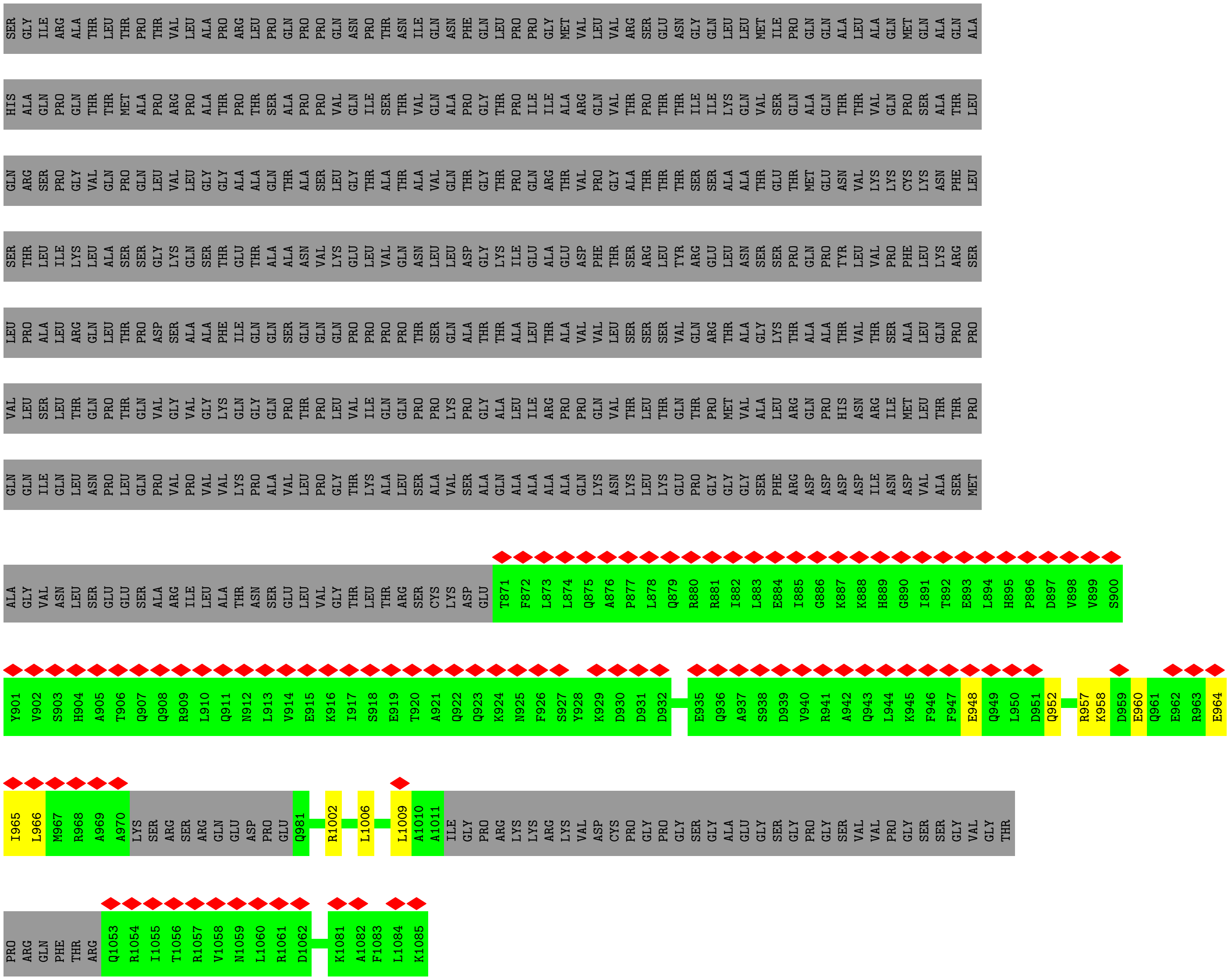
Chain B:



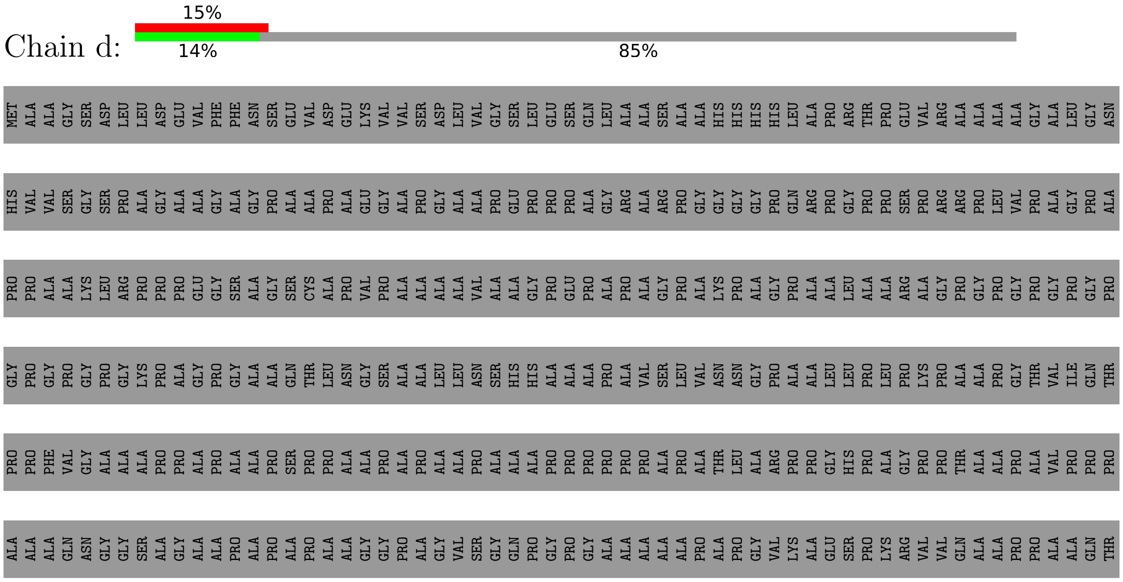
- Molecule 13: Transcription initiation factor TFIID subunit 4

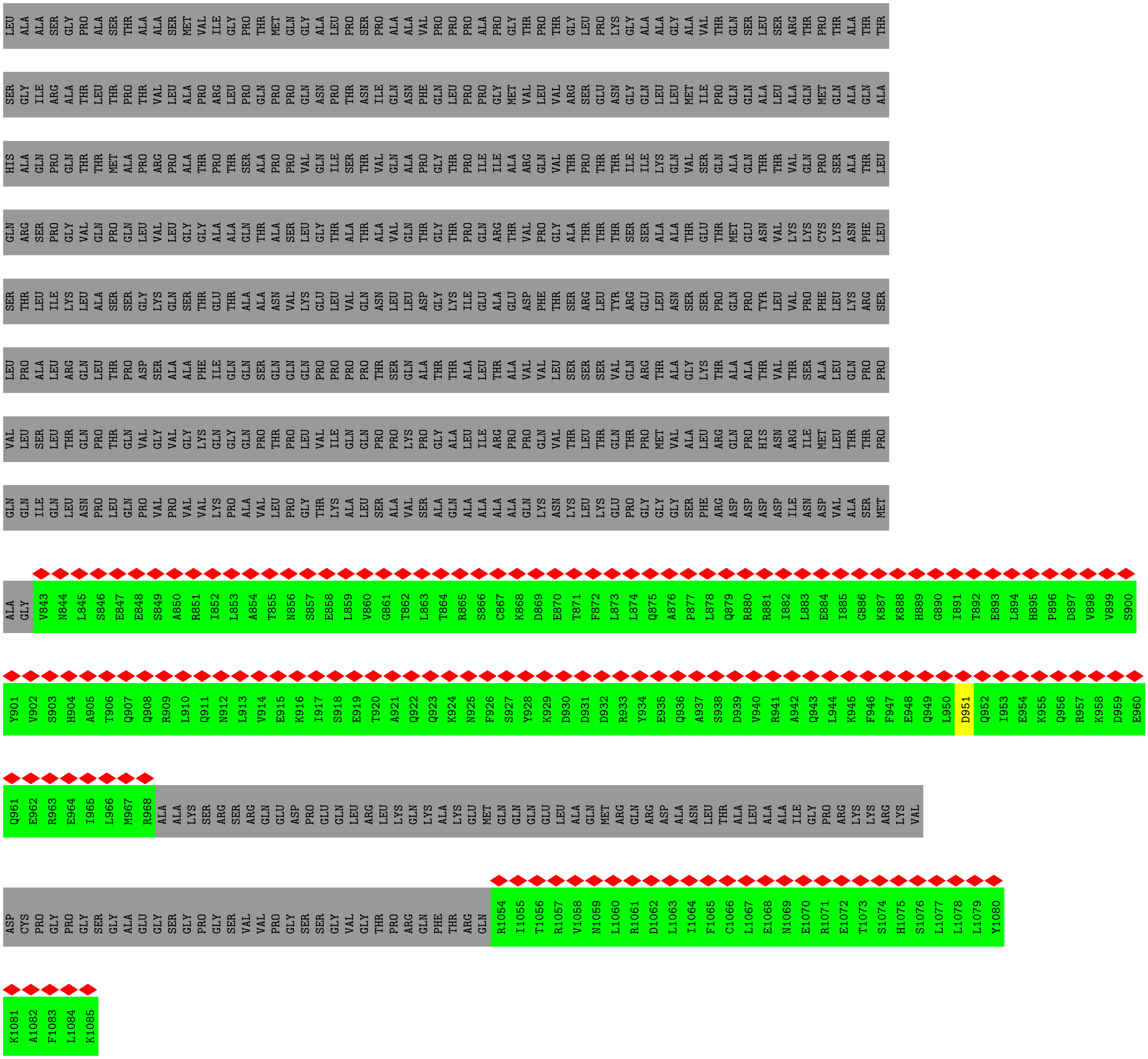
Chain D:



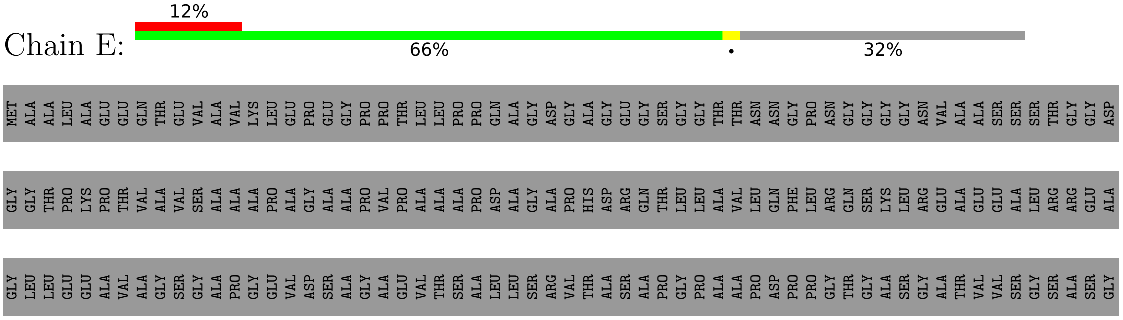


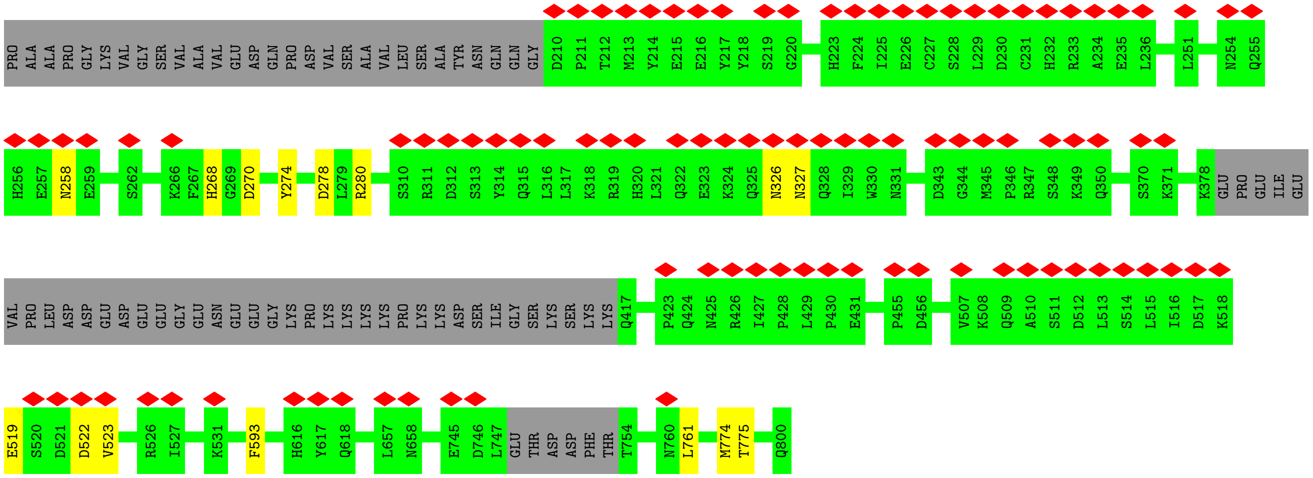
- Molecule 13: Transcription initiation factor TFIID subunit 4





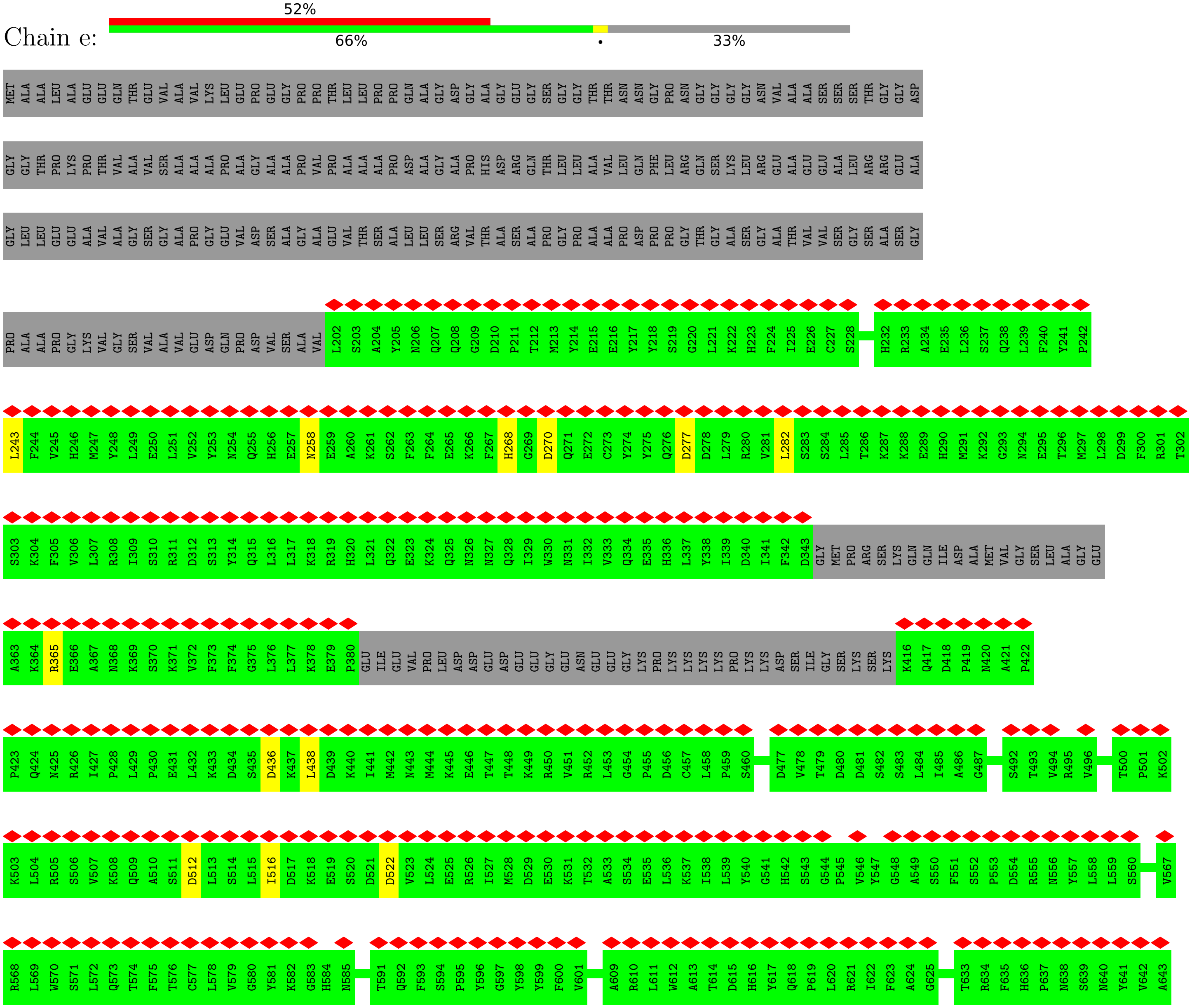
- Molecule 14: Transcription initiation factor TFIID subunit 5

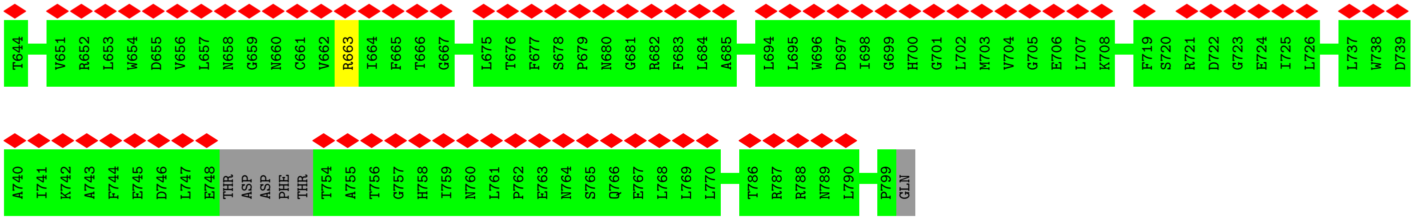




- Molecule 14: Transcription initiation factor TFIID subunit 5

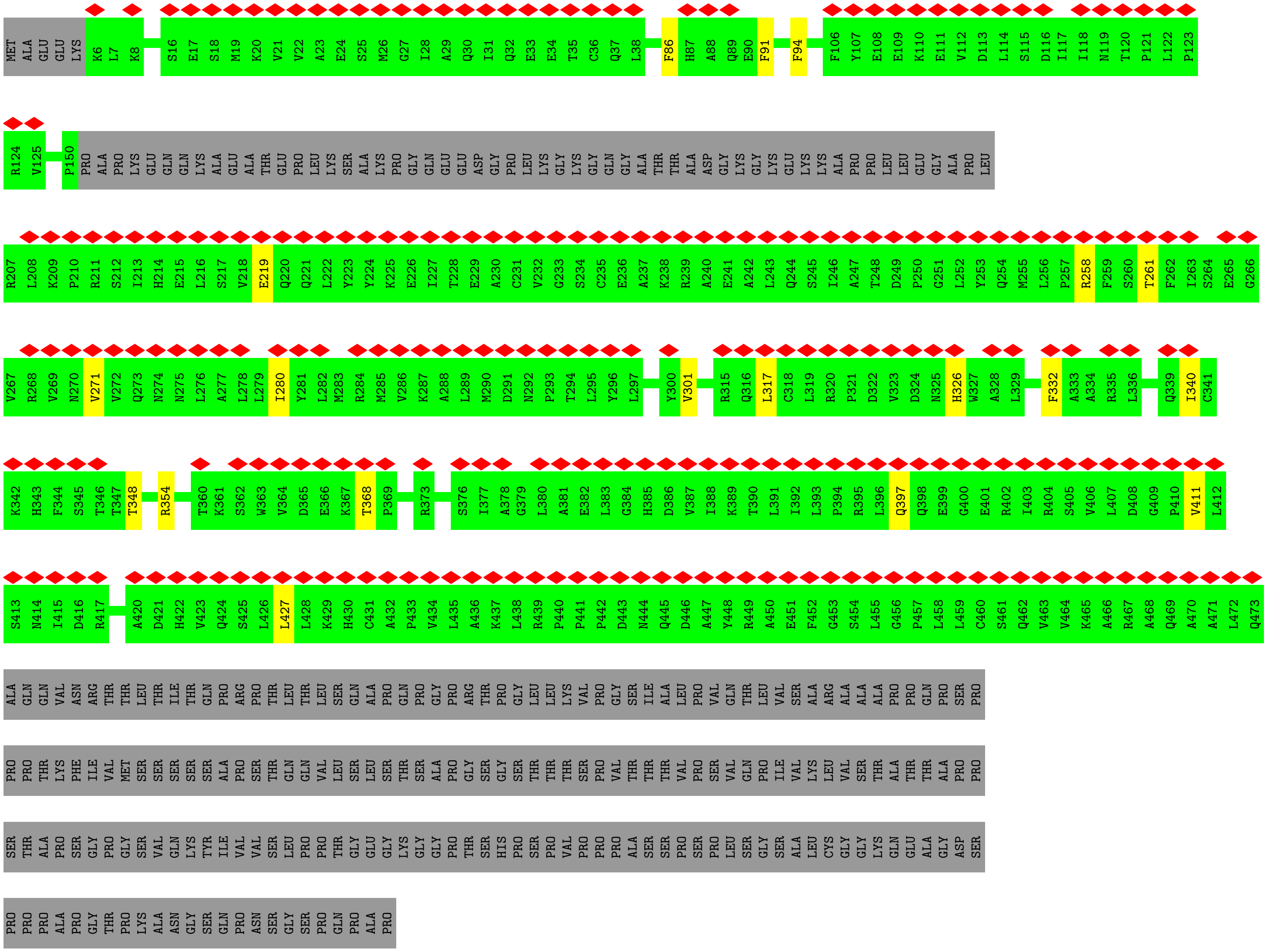
Chain e:





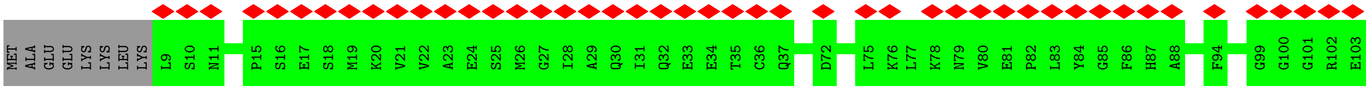
• Molecule 15: Transcription initiation factor TFIID subunit 6

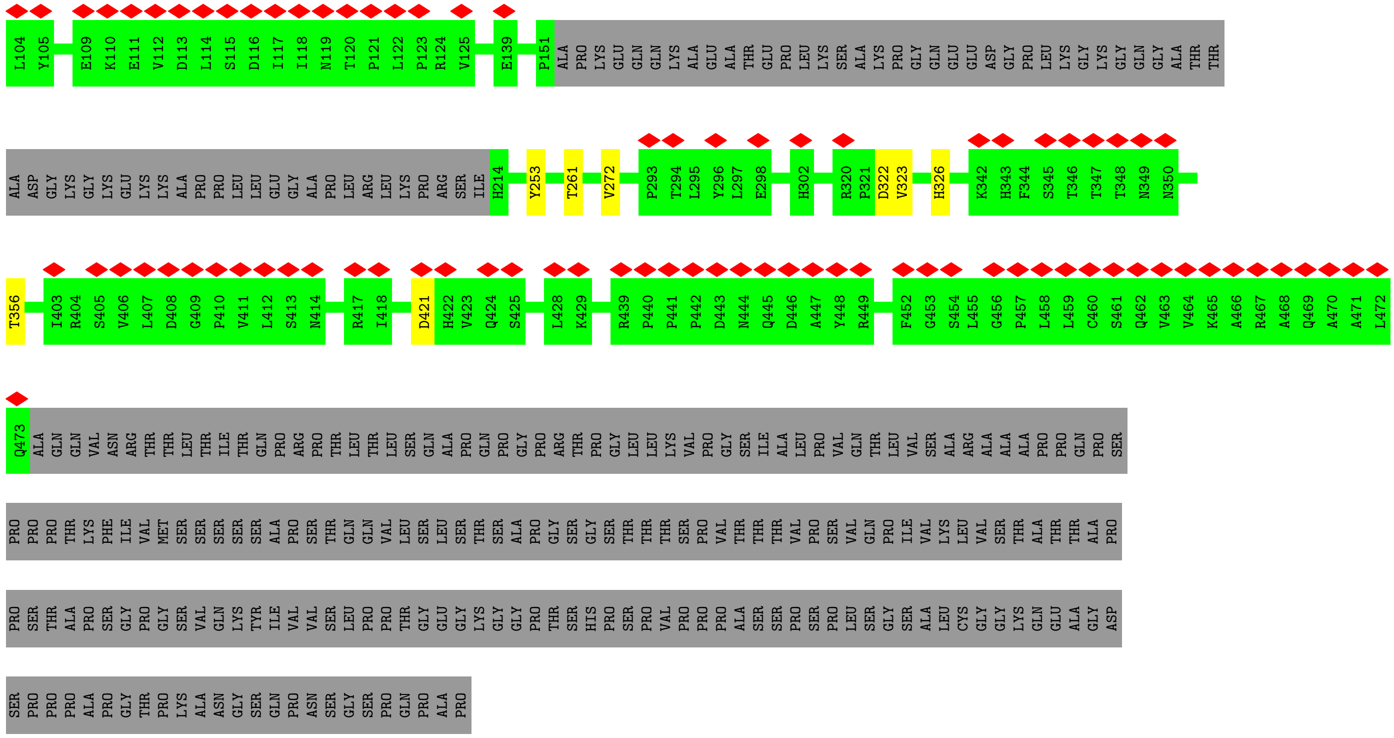
Chain F:



• Molecule 15: Transcription initiation factor TFIID subunit 6

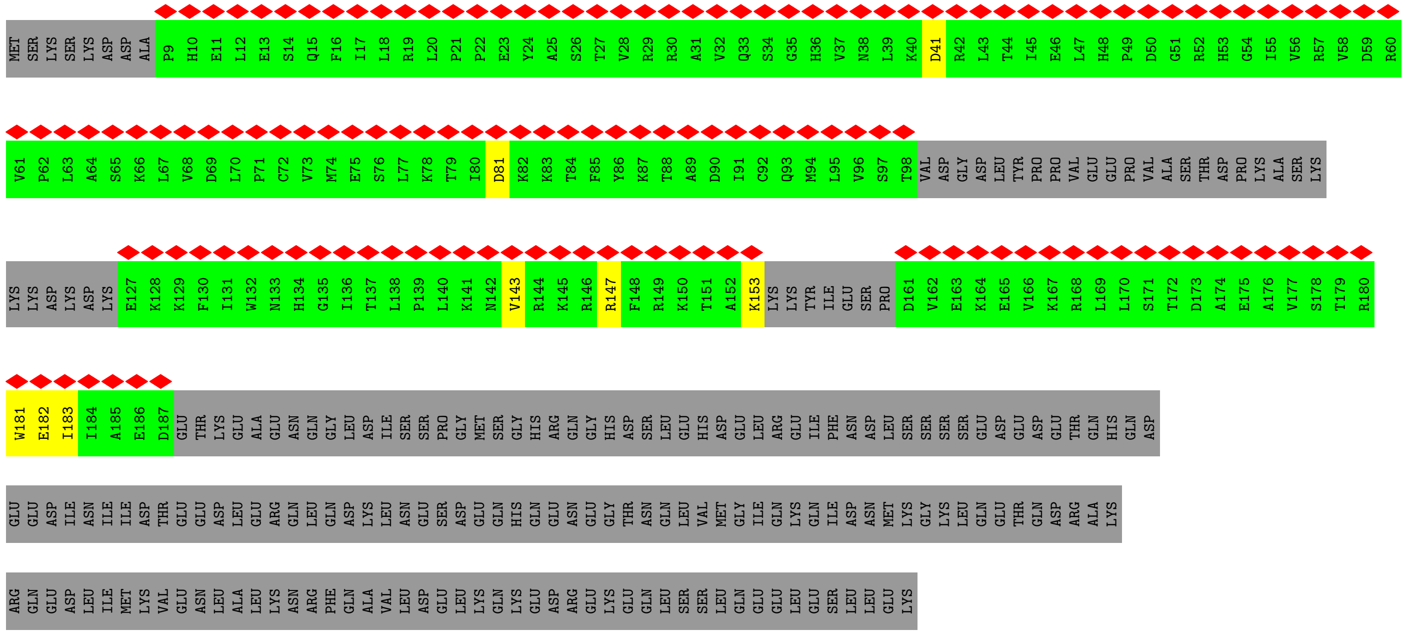
Chain f:





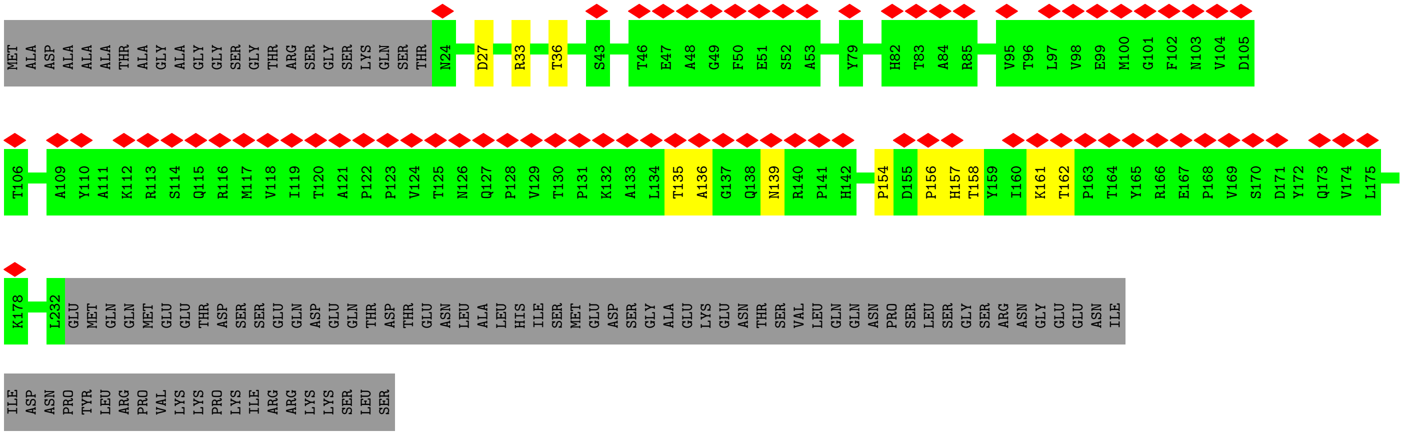
- Molecule 16: Transcription initiation factor TFIID subunit 7

Chain G:

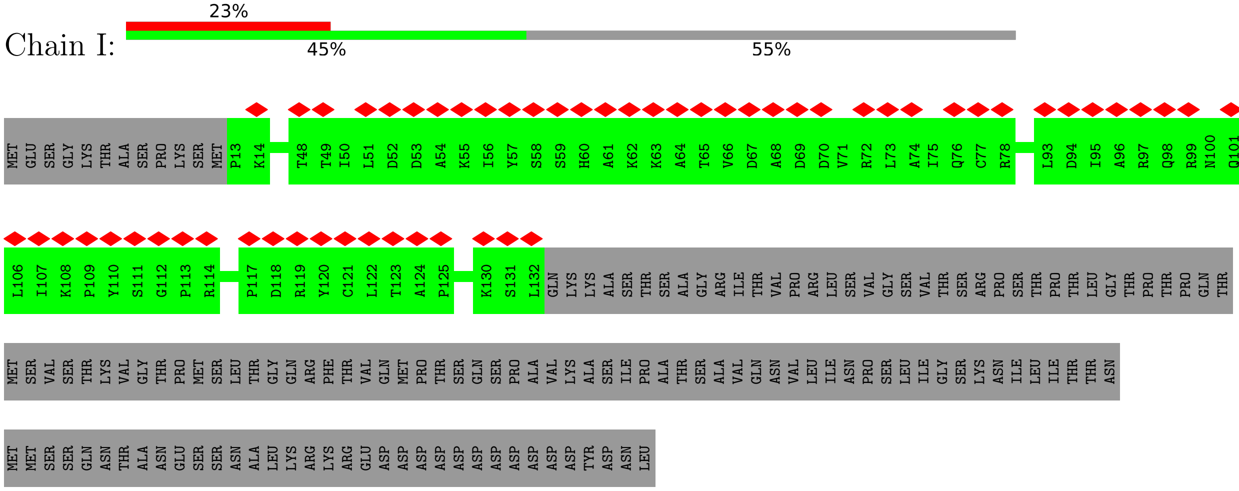


- Molecule 17: Transcription initiation factor TFIID subunit 8

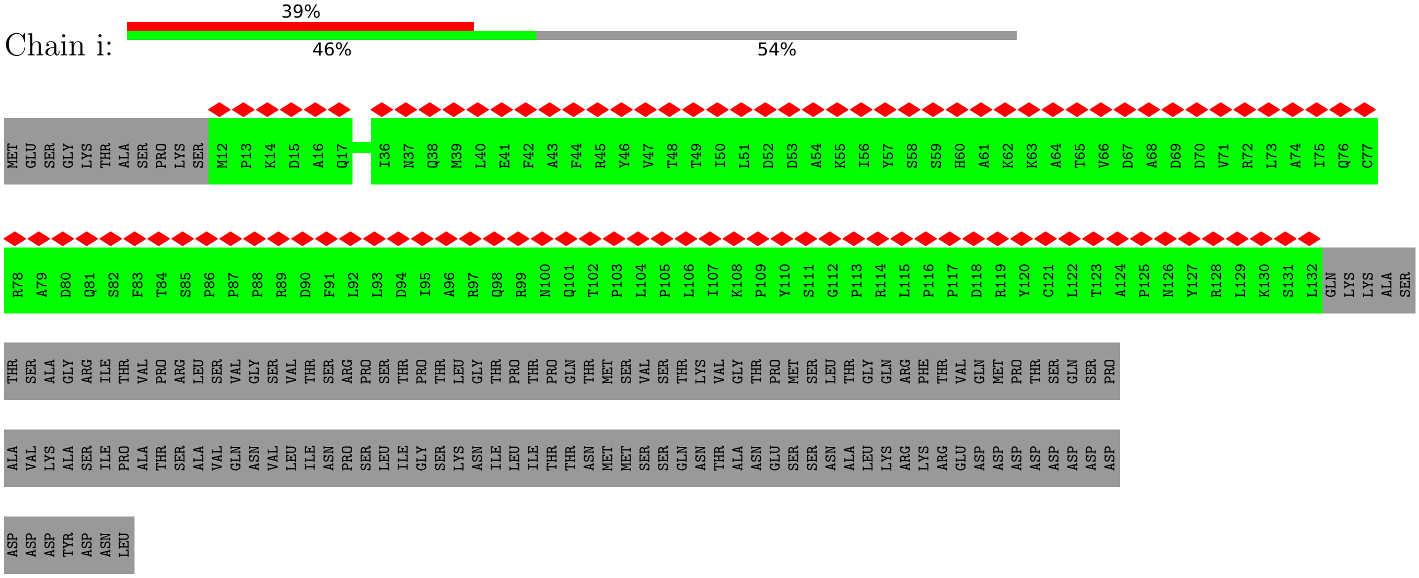
Chain H:



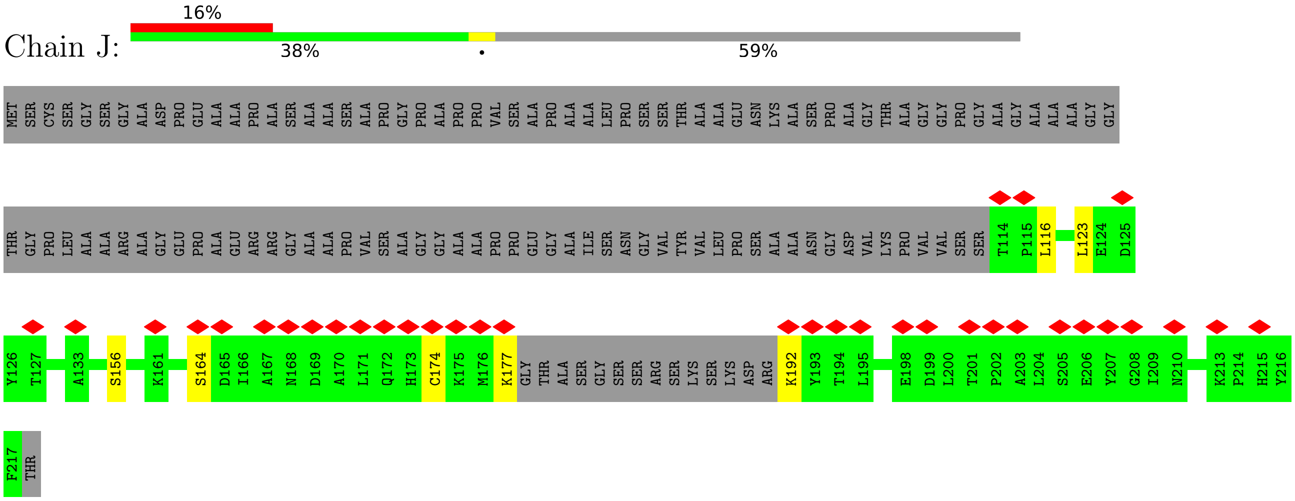
- Molecule 18: Transcription initiation factor TFIID subunit 9



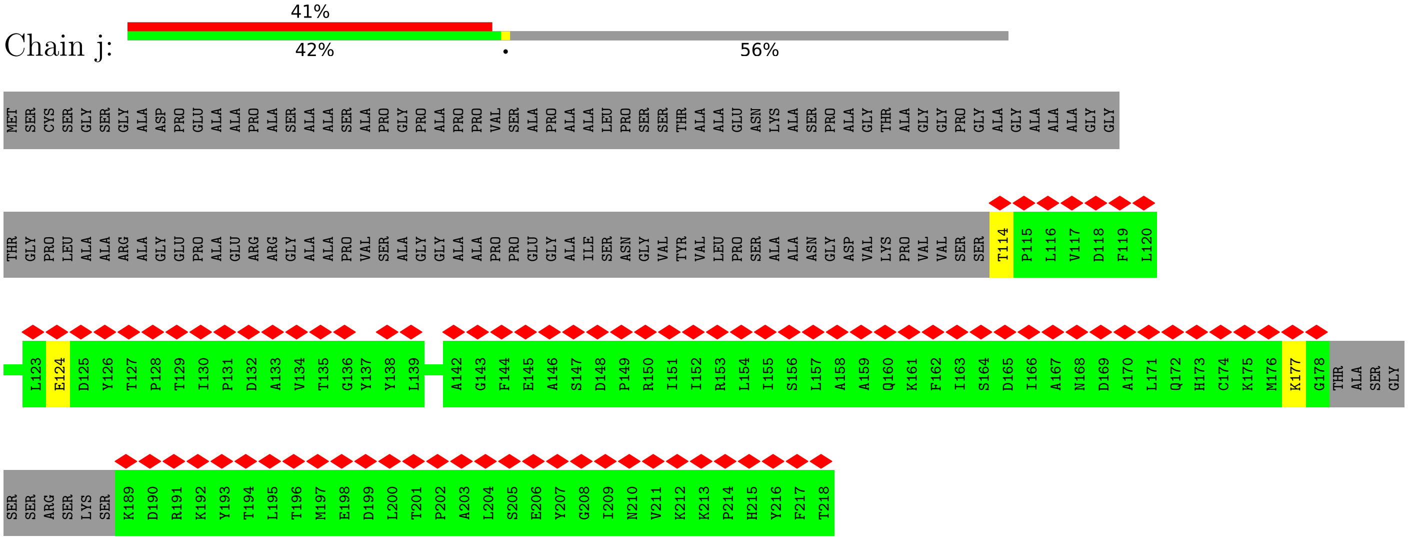
- Molecule 18: Transcription initiation factor TFIID subunit 9



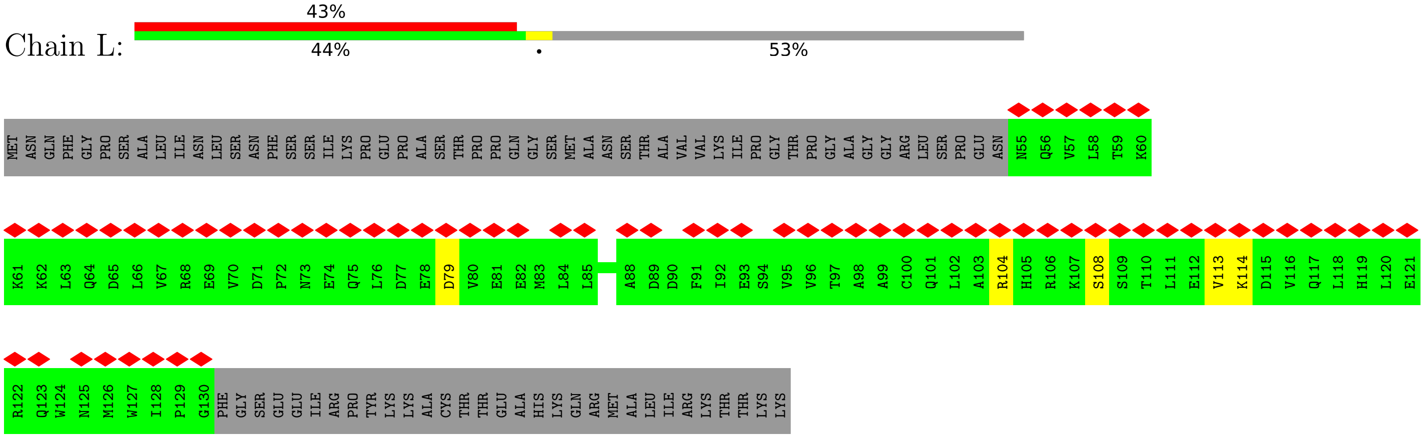
- Molecule 19: Transcription initiation factor TFIID subunit 10



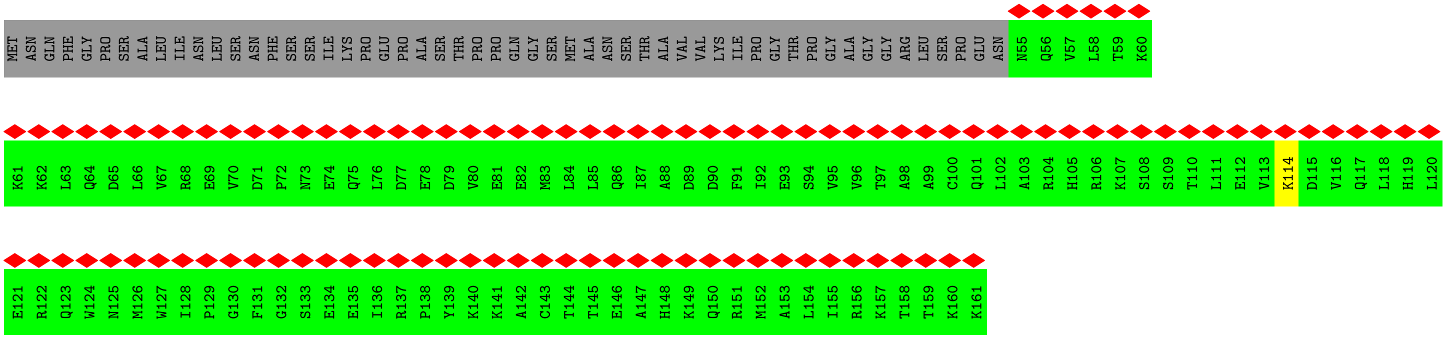
- Molecule 19: Transcription initiation factor TFIID subunit 10



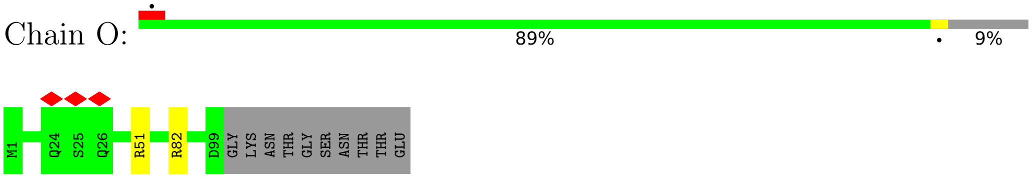
- Molecule 20: Transcription initiation factor TFIID subunit 12



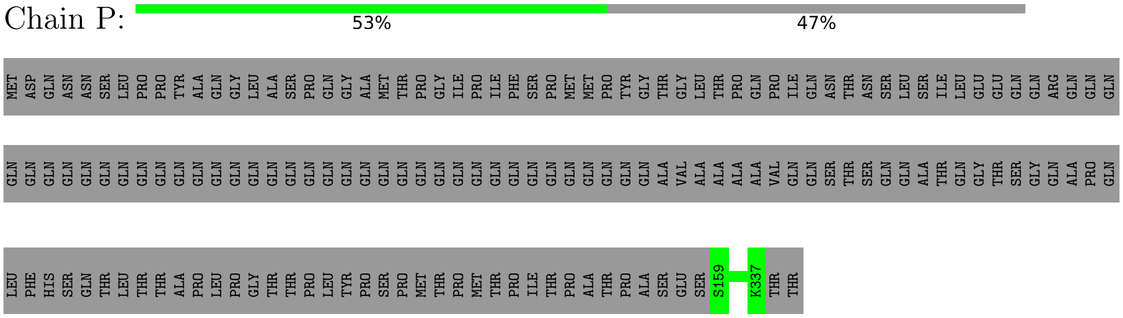
- Molecule 20: Transcription initiation factor TFIID subunit 12



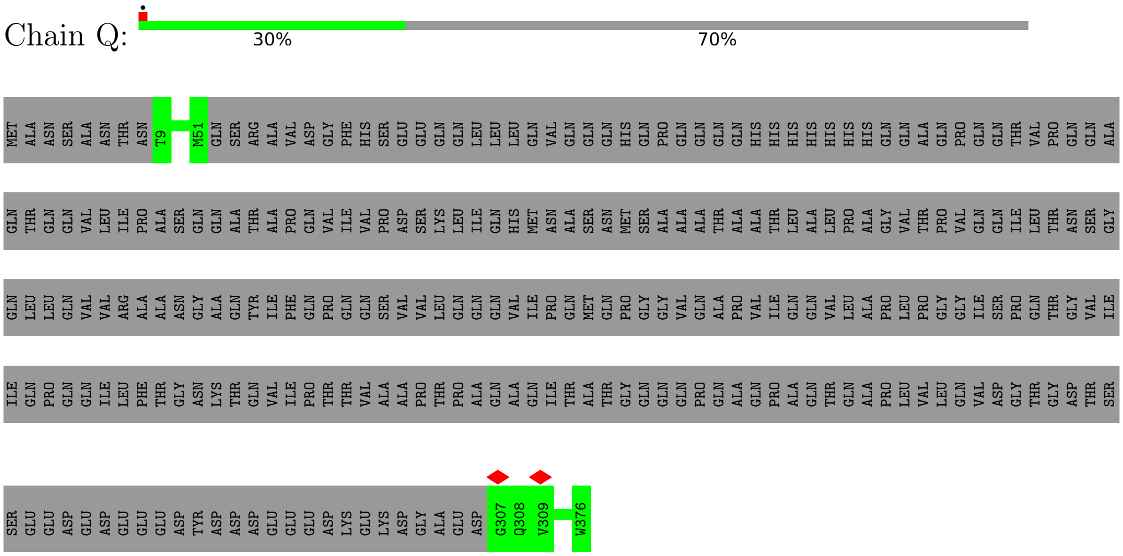
- Molecule 21: Transcription initiation factor IIA subunit 2



- Molecule 22: TATA-box-binding protein

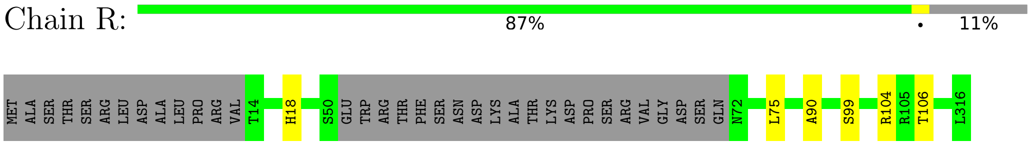


- Molecule 23: Transcription initiation factor IIA beta chain



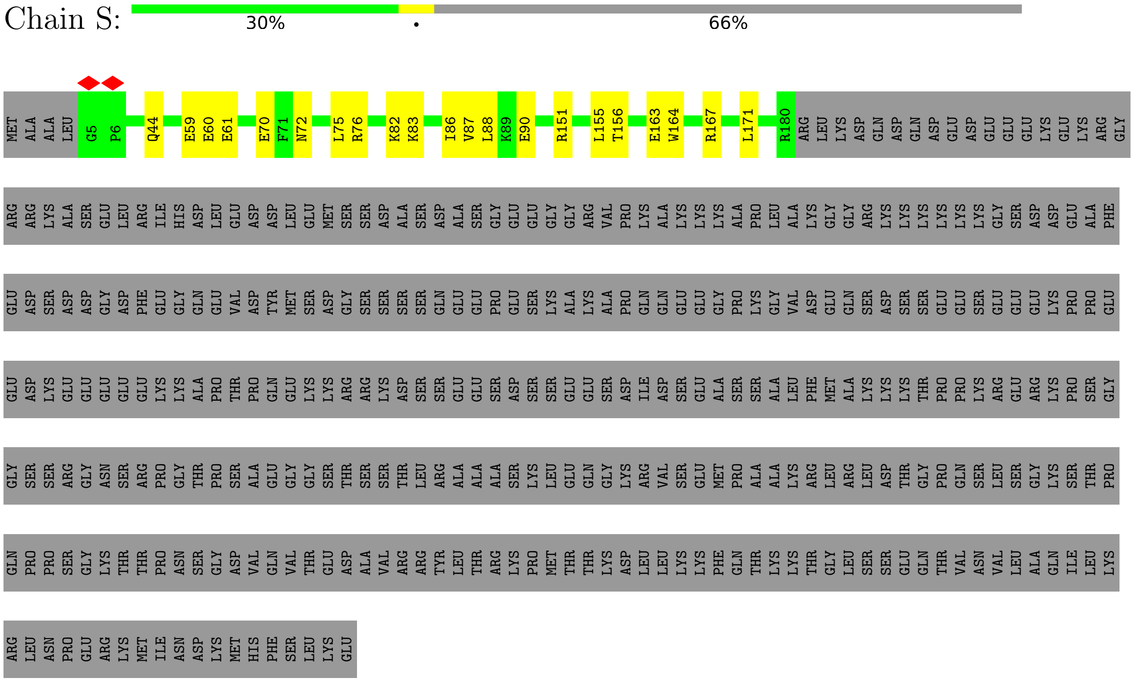
- Molecule 24: Transcription initiation factor IIB

Chain R:



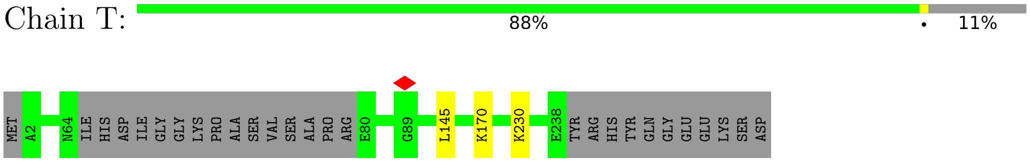
- Molecule 25: General transcription factor IIF subunit 1

Chain S:



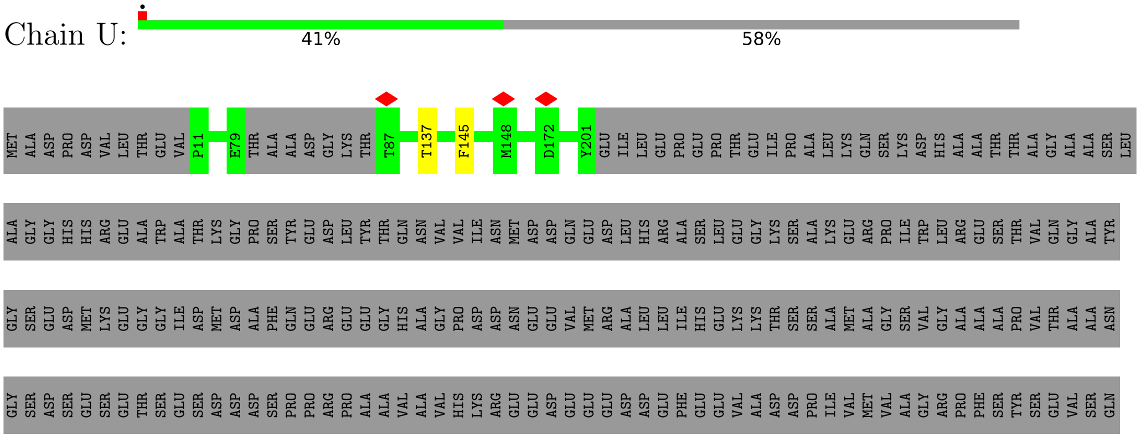
- Molecule 26: General transcription factor IIF subunit 2

Chain T:

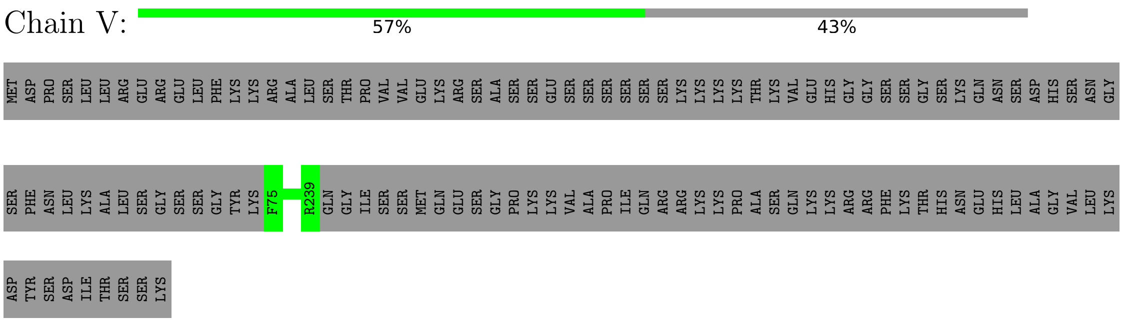


- Molecule 27: General transcription factor IIE subunit 1

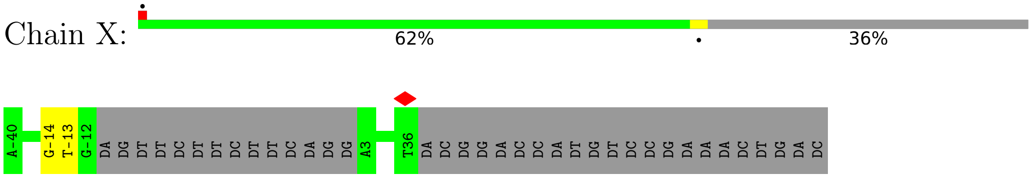
Chain U:



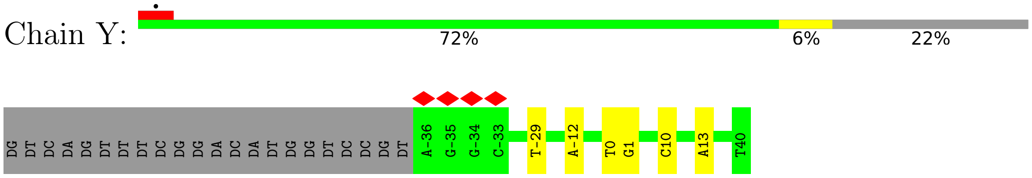
- Molecule 28: Transcription initiation factor IIE subunit beta



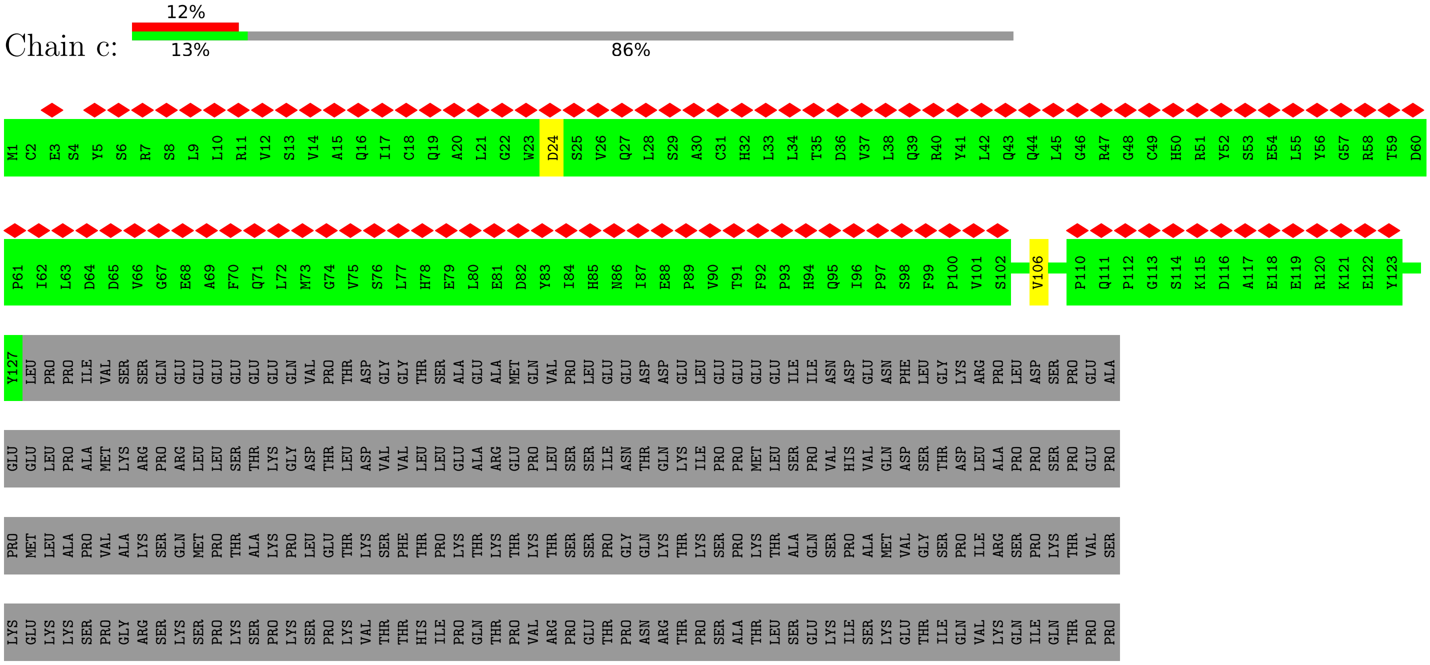
- Molecule 29: non-template DNA



- Molecule 30: template DNA



- Molecule 31: Transcription initiation factor TFIID subunit 3



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

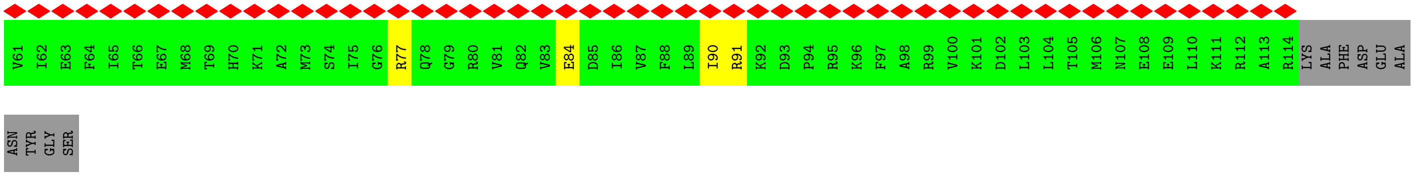
Y121 E122 M123 Y124 R125 R126 S127 A128 F129 P130 K131 A132 A133 I134 K135 R136 L137 I138 Q139 S140 I141 T142 G143 T144 S145 V146 S147 Q148 N149 V150 V151 I152 A153 M154 S155 G156 I157 S158 K159 V160 F161 V162 G163 E164 V165 V166 E167 E168 A169 L170 D171 V172 C173 E174 K175 W176 G177 E178 M179 P180

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

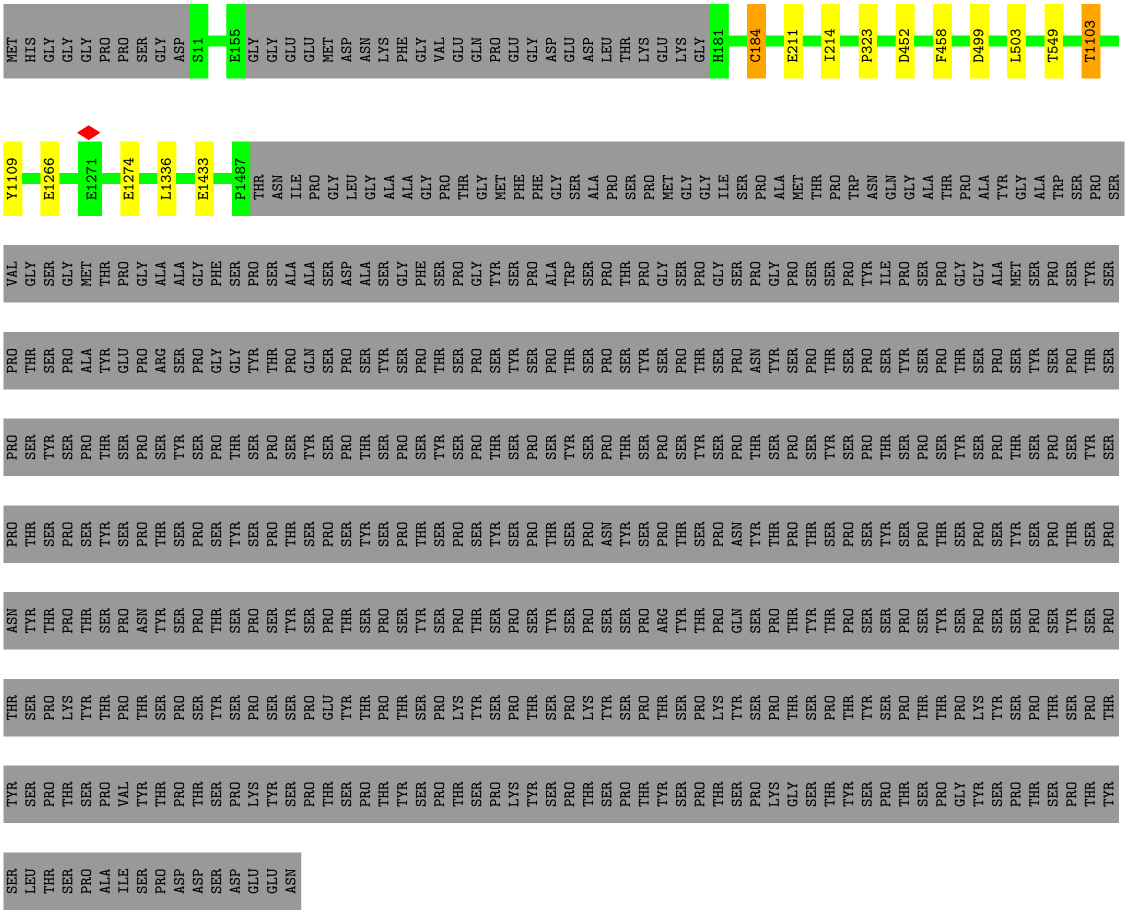
Chain m:

MET ALA ASP GLU GLU GLU ASP PRO THR PHE GLU GLU GLU ASN GLU GLU ILE GLY GLY GLY ALA GLU GLY GLY GLN GLY LYS R28 K29 R30 L31 F32 S33 K34 E35 L36 R37 C38 M39 M40 Y41 G42 F43 G44 D45 D46 Q47 N48 P49 Y50 T51 E52 S53 V54 D55 I56 L57 E58 D59 L60



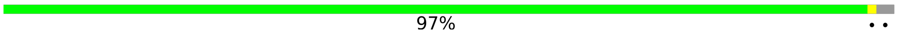
- Molecule 34: DNA-directed RNA polymerase subunit

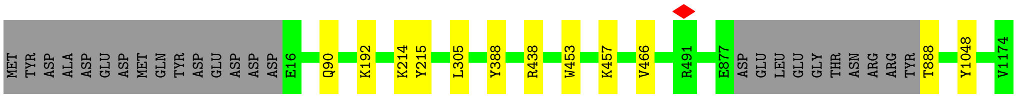
Chain o:



- Molecule 35: DNA-directed RNA polymerase subunit beta

Chain p:



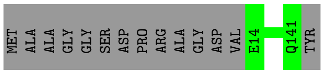


- Molecule 36: DNA-directed RNA polymerase II subunit RPB3

Chain q:

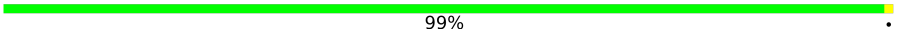
- Molecule 37: DNA-directed RNA polymerase II subunit RPB4

Chain r:



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

Chain s:

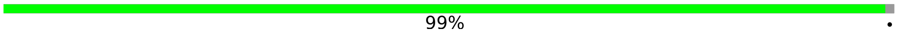


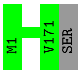
- Molecule 39: DNA-directed RNA polymerase II subunit F

Chain t:

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

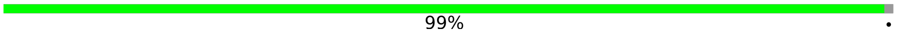
Chain u:

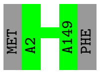




- Molecule 41: DNA-directed RNA polymerases I, II, and III subunit RPABC3

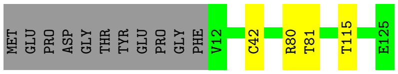
Chain v:





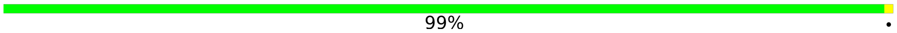
- Molecule 42: DNA-directed RNA polymerase II subunit RPB9

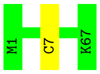
Chain w:



- Molecule 43: DNA-directed RNA polymerases I, II, and III subunit RPABC5

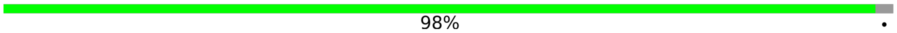
Chain x:

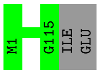




- Molecule 44: DNA-directed RNA polymerase II subunit RPB11-a

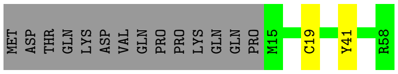
Chain y:





- Molecule 45: RPB12

Chain z:



4 Experimental information ⓘ

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	6727	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 $(e^-/\text{Å}^2)$	50	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor
Maximum map value	3.990	Depositor
Minimum map value	-0.875	Depositor
Average map value	0.041	Depositor
Map value standard deviation	0.187	Depositor
Recommended contour level	0.477	Depositor
Map size (Å)	426.88, 426.88, 426.88	wwPDB
Map dimensions	160, 160, 160	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	2.668, 2.668, 2.668	Depositor

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: HYP, TRX, MG, ZN, W0F, ILX, SF4, CSX, G2L, ATP

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	0.36	0/1437	0.59	3/1926 (0.2%)
2	1	0.26	0/2210	0.41	0/2975
3	2	0.27	0/2624	0.43	0/3555
4	3	0.30	0/2103	0.46	0/2846
5	4	0.31	0/3262	0.46	1/4418 (0.0%)
6	5	0.35	0/433	0.57	0/585
7	6	0.26	0/4994	0.41	0/6745
8	7	0.26	0/5875	0.41	0/7955
9	N	0.46	0/22	1.01	0/26
11	A	0.50	0/4698	0.66	0/6345
12	B	0.46	0/7993	0.61	0/10836
13	D	0.42	0/1379	0.63	1/1843 (0.1%)
13	d	0.39	0/1321	0.53	0/1772
14	E	0.39	0/4482	0.59	0/6069
14	e	0.44	0/4433	0.60	0/6004
15	F	0.48	0/3201	0.69	0/4347
15	f	0.41	0/3140	0.63	0/4268
16	G	0.51	0/1190	0.62	0/1601
17	H	0.46	0/1673	0.64	0/2285
18	I	0.27	0/981	0.47	0/1332
18	i	0.29	0/989	0.46	0/1343
19	J	0.57	0/736	0.69	0/998
19	j	0.54	0/775	0.63	0/1049
20	L	0.48	0/630	0.71	0/852
20	l	0.44	0/888	0.55	0/1194
21	O	0.28	0/816	0.44	0/1105
22	P	0.26	0/1448	0.50	0/1948
23	Q	0.26	0/945	0.43	0/1274
24	R	0.29	0/2199	0.52	1/2967 (0.0%)
25	S	0.27	0/1496	0.52	0/2013
26	T	0.26	0/1817	0.47	0/2445
27	U	0.32	0/1545	0.55	0/2075

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
28	V	0.28	0/1380	0.53	0/1854
29	X	0.67	0/1460	1.08	4/2251 (0.2%)
30	Y	0.71	1/1761 (0.1%)	1.04	5/2713 (0.2%)
31	c	0.39	0/1035	0.54	0/1406
32	k	0.31	0/799	0.48	0/1070
33	m	0.59	0/733	0.64	0/977
34	o	0.34	0/11723	0.52	1/15830 (0.0%)
35	p	0.37	0/9358	0.52	0/12633
36	q	0.38	0/2102	0.52	0/2857
37	r	0.24	0/1064	0.42	0/1428
38	s	0.33	0/1751	0.51	0/2366
39	t	0.34	0/667	0.51	0/901
40	u	0.30	0/1382	0.48	0/1874
41	v	0.36	0/1207	0.52	0/1628
42	w	0.30	0/948	0.49	0/1284
43	x	0.38	0/542	0.51	0/730
44	y	0.35	0/939	0.48	0/1271
45	z	0.35	0/377	0.61	0/500
All	All	0.38	1/110963 (0.0%)	0.56	16/150569 (0.0%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
24	R	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
30	Y	-12	DA	C1'-N9	-5.52	1.39	1.47

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
30	Y	0	DT	O4'-C4'-C3'	-8.44	100.94	106.00
1	0	55	LYS	CB-CA-C	-7.71	94.97	110.40
30	Y	1	DG	O4'-C4'-C3'	-7.41	101.53	104.50
29	X	-14	DG	O4'-C1'-N9	6.71	112.69	108.00
30	Y	13	DA	O4'-C1'-N9	6.49	112.54	108.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
24	R	106	THR	Peptide

5.2 Too-close contacts ⓘ

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

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	169/309 (55%)	156 (92%)	13 (8%)	0	100	100
2	1	253/548 (46%)	243 (96%)	10 (4%)	0	100	100
3	2	325/395 (82%)	312 (96%)	13 (4%)	0	100	100
4	3	259/308 (84%)	253 (98%)	6 (2%)	0	100	100
5	4	384/462 (83%)	365 (95%)	19 (5%)	0	100	100
6	5	52/71 (73%)	48 (92%)	2 (4%)	2 (4%)	3	24
7	6	601/782 (77%)	573 (95%)	27 (4%)	1 (0%)	47	81
8	7	710/760 (93%)	683 (96%)	27 (4%)	0	100	100
9	N	4/8 (50%)	3 (75%)	1 (25%)	0	100	100
11	A	543/1872 (29%)	519 (96%)	23 (4%)	1 (0%)	47	81
12	B	959/1199 (80%)	911 (95%)	48 (5%)	0	100	100
13	D	158/1085 (15%)	153 (97%)	5 (3%)	0	100	100
13	d	154/1085 (14%)	150 (97%)	4 (3%)	0	100	100
14	E	541/800 (68%)	518 (96%)	22 (4%)	1 (0%)	47	81
14	e	531/800 (66%)	482 (91%)	48 (9%)	1 (0%)	47	81
15	F	408/677 (60%)	392 (96%)	15 (4%)	1 (0%)	47	81

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
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%)	13 (6%)	5 (2%)	6	33
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%)	6	35
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%)	95 (98%)	2 (2%)	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	278/316 (88%)	261 (94%)	15 (5%)	2 (1%)	22	62
25	S	174/517 (34%)	161 (92%)	12 (7%)	1 (1%)	25	65
26	T	218/249 (88%)	211 (97%)	7 (3%)	0	100	100
27	U	180/439 (41%)	153 (85%)	26 (14%)	1 (1%)	25	65
28	V	163/291 (56%)	140 (86%)	23 (14%)	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	1448/1970 (74%)	1399 (97%)	45 (3%)	4 (0%)	41	76
35	p	1145/1174 (98%)	1109 (97%)	36 (3%)	0	100	100
36	q	253/275 (92%)	246 (97%)	7 (3%)	0	100	100
37	r	126/142 (89%)	126 (100%)	0	0	100	100
38	s	207/210 (99%)	203 (98%)	4 (2%)	0	100	100
39	t	80/127 (63%)	79 (99%)	1 (1%)	0	100	100
40	u	169/172 (98%)	167 (99%)	2 (1%)	0	100	100
41	v	146/150 (97%)	142 (97%)	4 (3%)	0	100	100
42	w	112/125 (90%)	108 (96%)	4 (4%)	0	100	100
43	x	65/67 (97%)	62 (95%)	3 (5%)	0	100	100
44	y	113/117 (97%)	110 (97%)	3 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
45	z	42/58 (72%)	40 (95%)	2 (5%)	0	100	100
All	All	12996/22270 (58%)	12412 (96%)	562 (4%)	22 (0%)	50	81

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
6	5	48	GLU
14	E	523	VAL
17	H	156	PRO
25	S	87	VAL
34	o	1103	THR

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	0	164/283 (58%)	157 (96%)	7 (4%)	29	54
2	1	241/484 (50%)	241 (100%)	0	100	100
3	2	295/352 (84%)	293 (99%)	2 (1%)	84	90
4	3	234/272 (86%)	232 (99%)	2 (1%)	78	88
5	4	346/399 (87%)	339 (98%)	7 (2%)	55	73
6	5	48/64 (75%)	47 (98%)	1 (2%)	53	72
7	6	536/688 (78%)	533 (99%)	3 (1%)	86	91
8	7	624/664 (94%)	623 (100%)	1 (0%)	93	96
9	N	2/2 (100%)	2 (100%)	0	100	100
11	A	496/1665 (30%)	455 (92%)	41 (8%)	11	36
12	B	876/1083 (81%)	859 (98%)	17 (2%)	57	75
13	D	147/815 (18%)	137 (93%)	10 (7%)	16	42
13	d	146/815 (18%)	145 (99%)	1 (1%)	84	90
14	E	480/657 (73%)	466 (97%)	14 (3%)	42	64
14	e	475/657 (72%)	463 (98%)	12 (2%)	47	68

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
15	F	328/574 (57%)	310 (94%)	18 (6%)	21	48
15	f	322/574 (56%)	314 (98%)	8 (2%)	47	68
16	G	132/322 (41%)	124 (94%)	8 (6%)	18	45
17	H	181/270 (67%)	174 (96%)	7 (4%)	32	57
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%)	71	84
20	L	71/141 (50%)	66 (93%)	5 (7%)	15	41
20	l	98/141 (70%)	97 (99%)	1 (1%)	76	86
21	O	90/98 (92%)	88 (98%)	2 (2%)	52	71
22	P	154/293 (53%)	154 (100%)	0	100	100
23	Q	105/324 (32%)	105 (100%)	0	100	100
24	R	238/268 (89%)	236 (99%)	2 (1%)	81	89
25	S	154/448 (34%)	134 (87%)	20 (13%)	4	19
26	T	196/218 (90%)	193 (98%)	3 (2%)	65	80
27	U	167/373 (45%)	166 (99%)	1 (1%)	86	91
28	V	150/261 (58%)	150 (100%)	0	100	100
31	c	113/833 (14%)	111 (98%)	2 (2%)	59	77
32	k	87/182 (48%)	87 (100%)	0	100	100
33	m	80/106 (76%)	73 (91%)	7 (9%)	10	33
34	o	1280/1749 (73%)	1268 (99%)	12 (1%)	78	88
35	p	1004/1027 (98%)	992 (99%)	12 (1%)	71	84
36	q	234/252 (93%)	228 (97%)	6 (3%)	46	67
37	r	118/126 (94%)	118 (100%)	0	100	100
38	s	191/192 (100%)	189 (99%)	2 (1%)	76	86
39	t	71/111 (64%)	71 (100%)	0	100	100
40	u	152/153 (99%)	152 (100%)	0	100	100
41	v	129/131 (98%)	129 (100%)	0	100	100
42	w	103/112 (92%)	99 (96%)	4 (4%)	32	57
43	x	56/56 (100%)	55 (98%)	1 (2%)	59	77

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
44	y	104/106 (98%)	104 (100%)	0	100	100
45	z	41/55 (74%)	39 (95%)	2 (5%)	25	51
All	All	11634/19174 (61%)	11385 (98%)	249 (2%)	56	72

5 of 249 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
15	F	326	HIS
42	w	81	THR
19	J	177	LYS
42	w	42	CYS
14	e	512	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 126 such sidechains are listed below:

Mol	Chain	Res	Type
18	I	21	GLN
42	w	50	ASN
28	V	83	ASN
42	w	41	ASN
14	e	616	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
10	Z	0/2	-	-

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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

5 non-standard protein/DNA/RNA residues are modelled in this entry.

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	$\#\|Z\| > 2$
9	ILX	N	1	9	8,9,10	0.62	0	9,11,13	1.32	2 (22%)
9	CSX	N	6	9	3,6,7	1.04	0	1,6,8	2.26	1 (100%)
9	HYP	N	8	9	6,8,9	0.76	0	5,10,12	2.19	2 (40%)
9	TRX	N	2	9	14,16,17	0.97	0	15,22,24	1.98	4 (26%)
10	G2L	Z	2	10,30	19,26,30	1.02	1 (5%)	18,38,44	1.34	3 (16%)

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
9	ILX	N	1	9	-	7/11/12/14	-
9	CSX	N	6	9	-	1/1/5/7	-
9	HYP	N	8	9	-	0/0/11/13	0/1/1/1
9	TRX	N	2	9	-	0/4/6/8	0/2/2/2
10	G2L	Z	2	10,30	-	2/5/27/31	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	Z	2	G2L	C6-N1	-2.42	1.34	1.37

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
9	N	2	TRX	CG-CB-CA	5.77	123.45	114.53
9	N	8	HYP	O-C-CA	-3.30	116.12	124.78
9	N	8	HYP	CB-CG-CD	3.06	107.01	103.27
9	N	1	ILX	OD1-CD1-CG1	-2.89	104.77	111.07
9	N	2	TRX	CB-CG-CD1	-2.89	124.40	127.97

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	N	1	ILX	C-CA-CB-CG2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
9	N	1	ILX	OD1-CD1-CG1-CB
9	N	1	ILX	OD1-CD1-CG1-OG1
9	N	6	CSX	N-CA-CB-SG
10	Z	2	G2L	C3'-C4'-C5'-O5'

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 20 ligands modelled in this entry, 18 are monoatomic - leaving 2 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	-	-	-		
48	W0F	Z	2500	-	25,35,35	3.22	7 (28%)	31,55,55	2.54	12 (38%)

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
48	W0F	Z	2500	-	-	3/20/40/40	0/3/3/3

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
48	Z	2500	W0F	C03-C04	-9.40	1.27	1.52
48	Z	2500	W0F	O05-C06	-9.31	1.28	1.41
48	Z	2500	W0F	O05-C04	6.87	1.60	1.45
48	Z	2500	W0F	C07-C03	-2.91	1.46	1.52
48	Z	2500	W0F	C20-C04	2.44	1.59	1.51

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
48	Z	2500	W0F	O05-C06-C07	-6.28	97.74	106.93
48	Z	2500	W0F	O05-C04-C03	-5.68	92.69	104.87
48	Z	2500	W0F	O05-C04-C20	4.79	125.15	109.37
48	Z	2500	W0F	O21-C20-C04	3.94	122.56	108.99
48	Z	2500	W0F	O32-P30-O29	3.91	117.73	104.64

There are no chirality outliers.

All (3) torsion outliers are listed below:

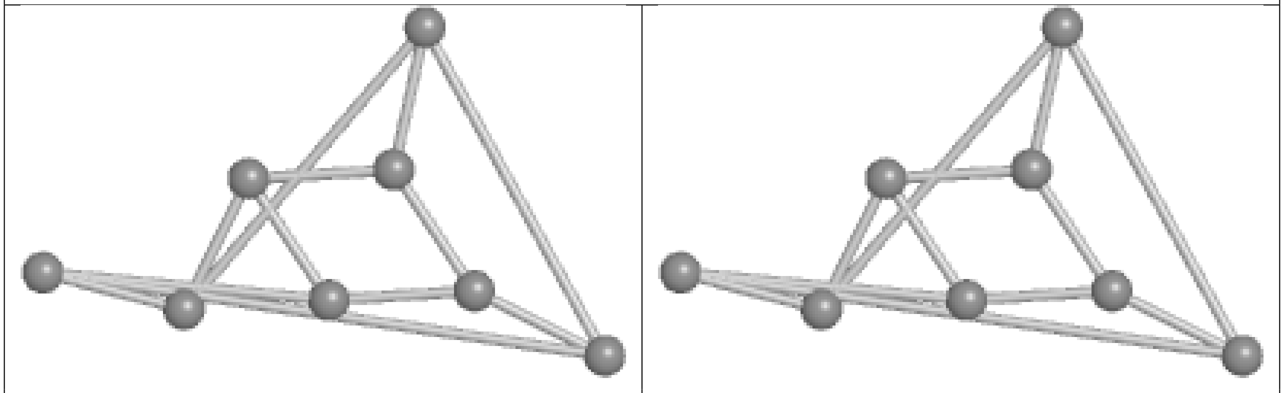
Mol	Chain	Res	Type	Atoms
48	Z	2500	W0F	C20-O21-P22-O24
48	Z	2500	W0F	C20-O21-P22-O25
48	Z	2500	W0F	C03-C04-C20-O21

There are no ring outliers.

No monomer is involved in short contacts.

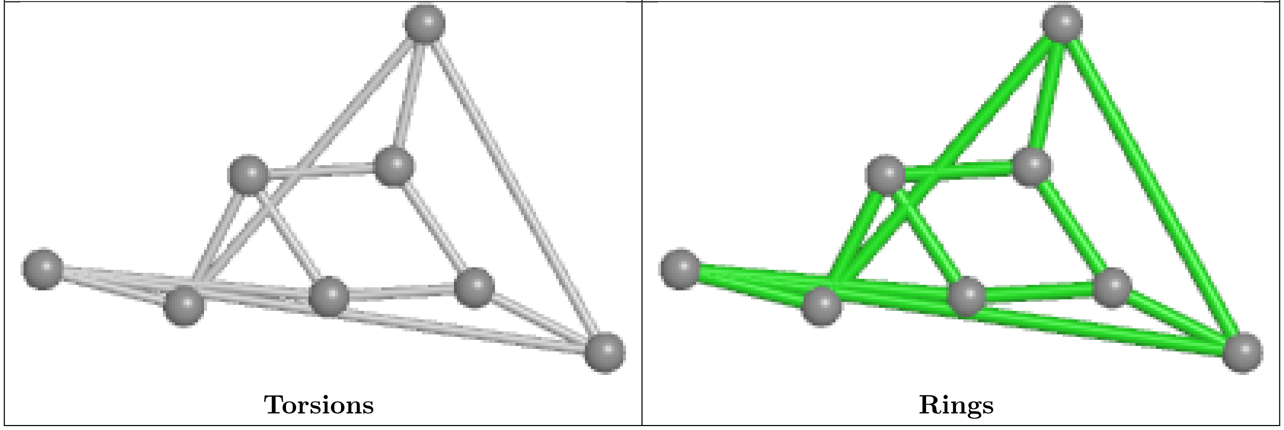
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand SF4 7 1000



Bond lengths

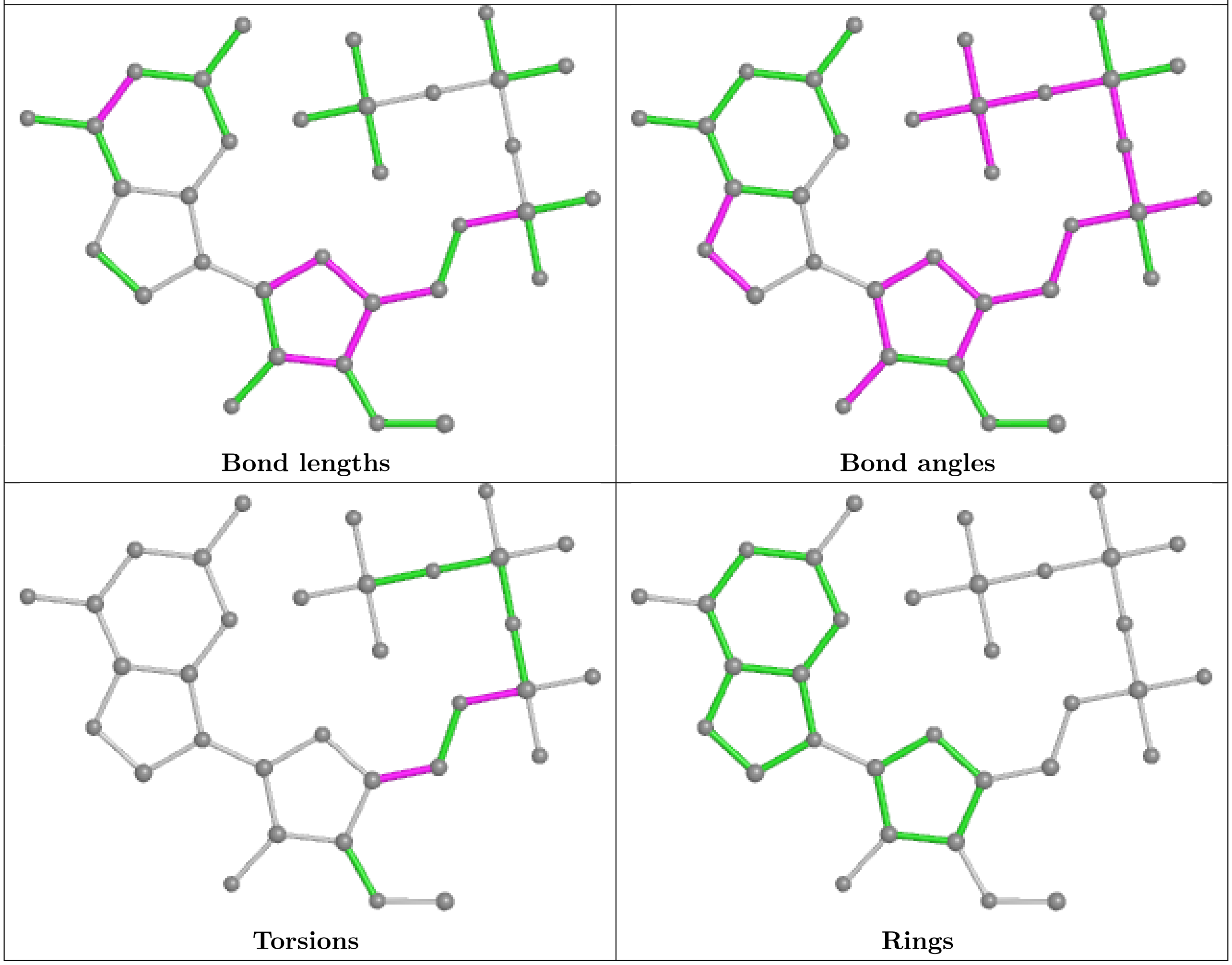
Bond angles



Torsions

Rings

Ligand W0F Z 2500



Bond lengths

Bond angles

Torsions

Rings

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.

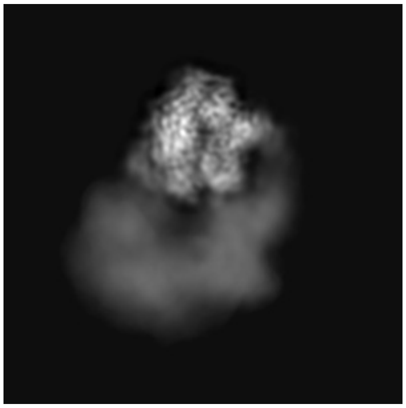
6 Map visualisation ⓘ

This section contains visualisations of the EMDB entry EMD-37395. These allow visual inspection of the internal detail of the map and identification of artifacts.

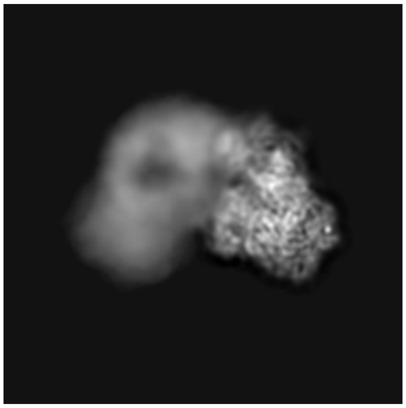
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections ⓘ

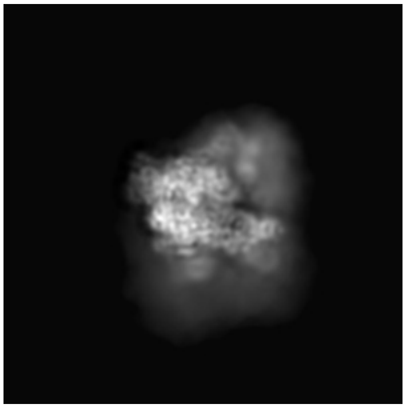
6.1.1 Primary map



X

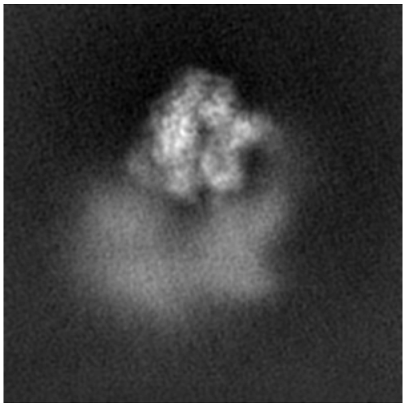


Y

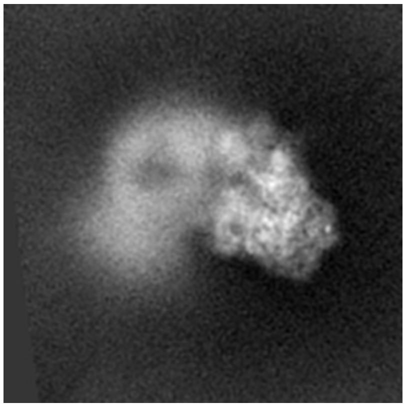


Z

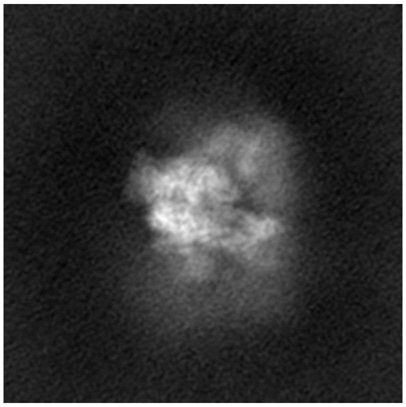
6.1.2 Raw map



X



Y

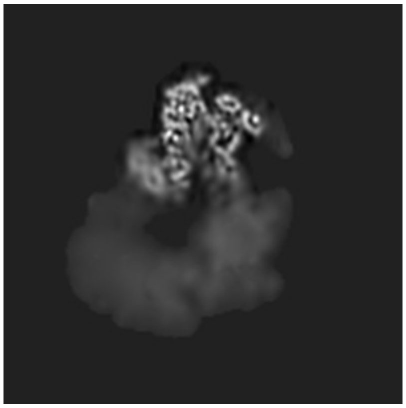


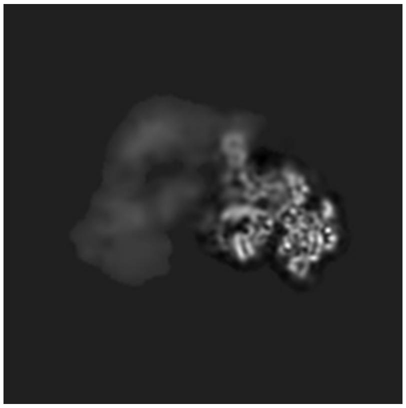
Z

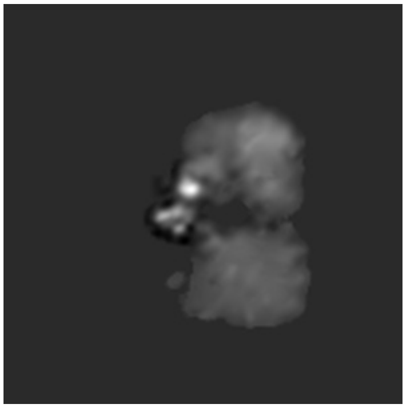
The images above show the map projected in three orthogonal directions.

6.2 Central slices

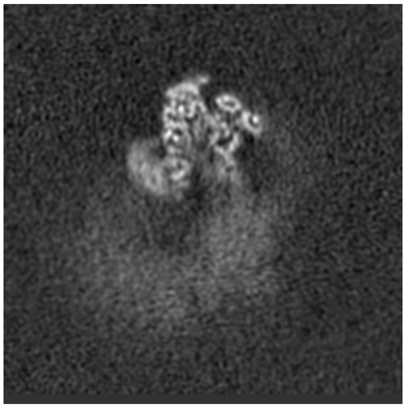
6.2.1 Primary map

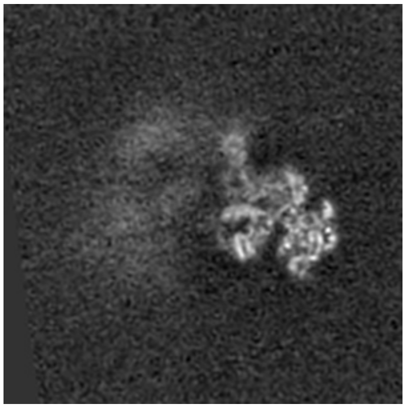

X Index: 80

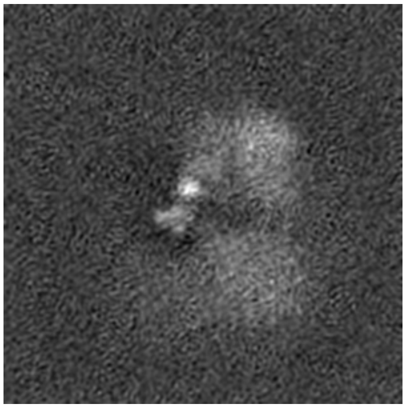

Y Index: 80


Z Index: 80

6.2.2 Raw map


X Index: 80

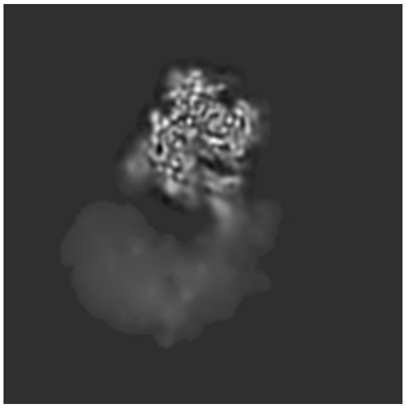

Y Index: 80

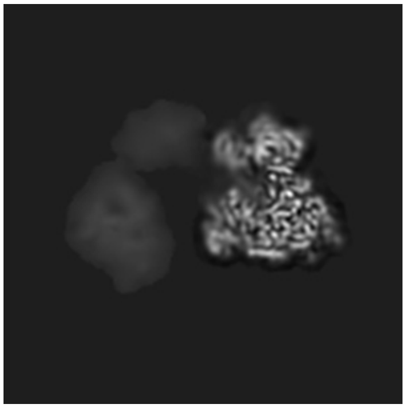

Z Index: 80

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

6.3 Largest variance slices

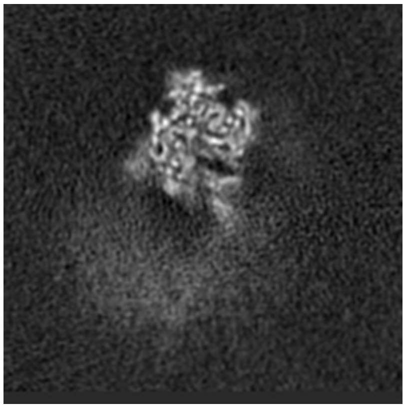
6.3.1 Primary map

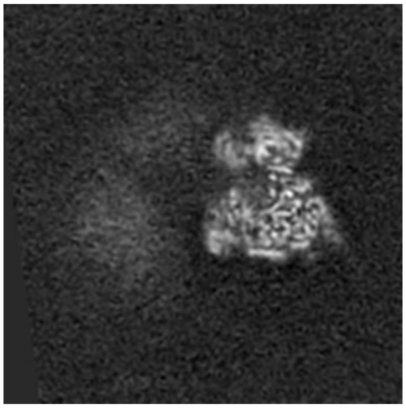

X Index: 73

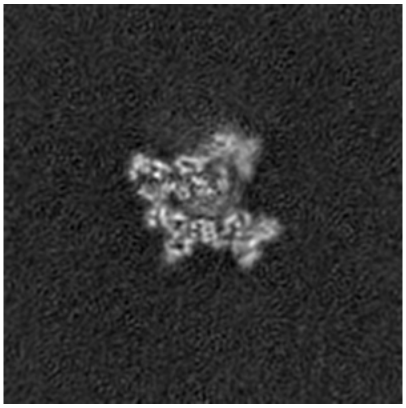

Y Index: 71


Z Index: 111

6.3.2 Raw map


X Index: 73

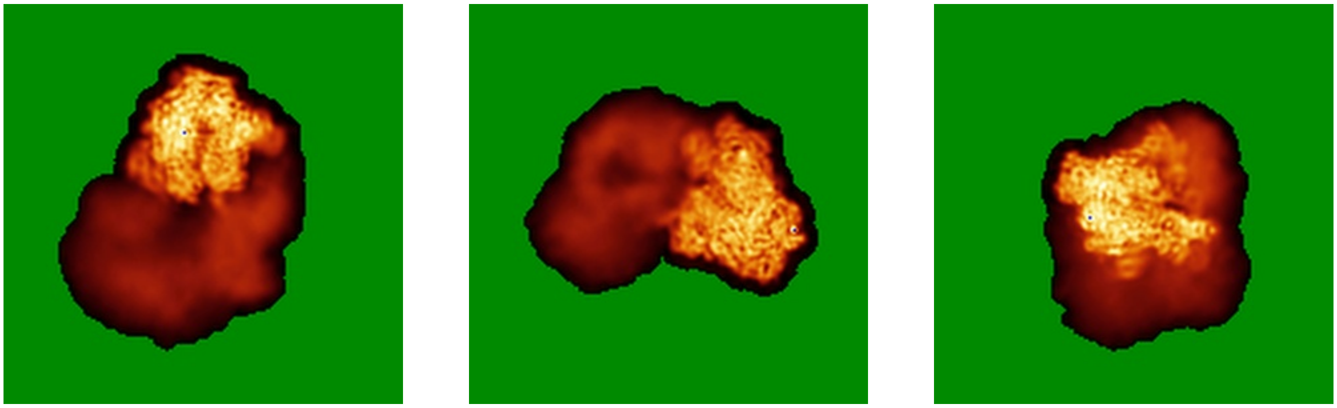

Y Index: 71


Z Index: 111

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

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

6.4.1 Primary map

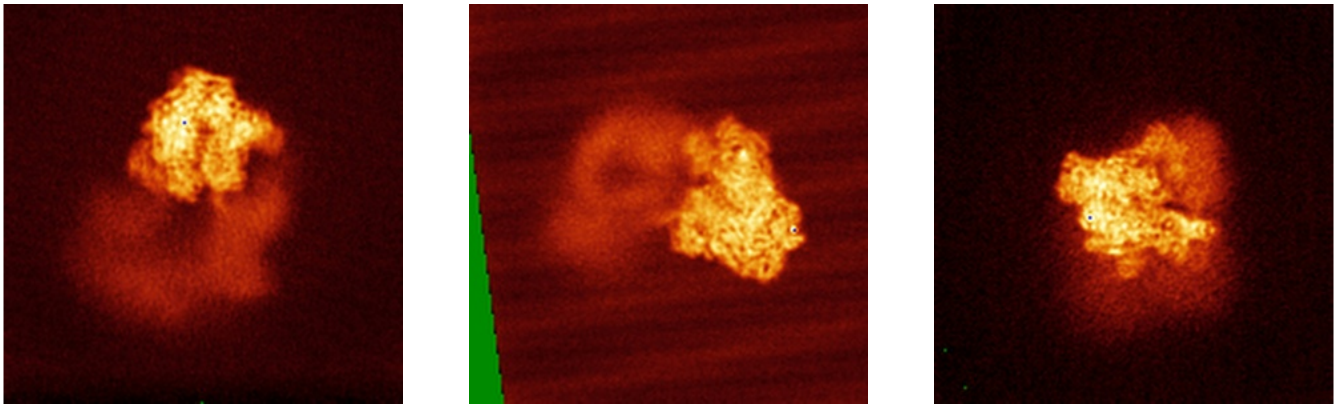


X

Y

Z

6.4.2 Raw map



X

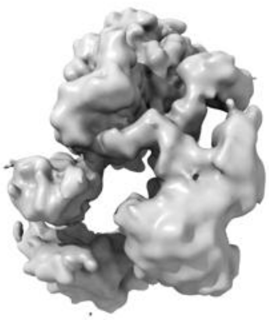
Y

Z

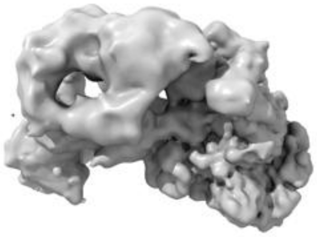
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views

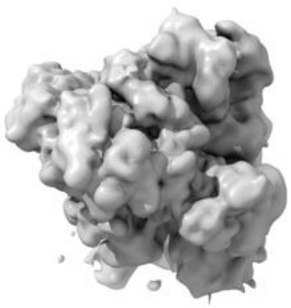
6.5.1 Primary map



X



Y

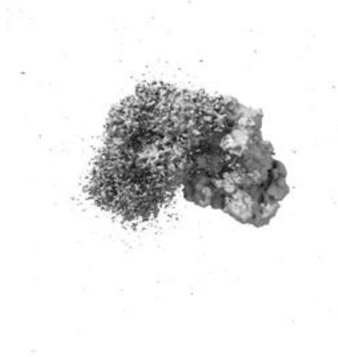


Z

The images above show the 3D surface view of the map at the recommended contour level 0.477. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.5.2 Raw map

X



Y

Z

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

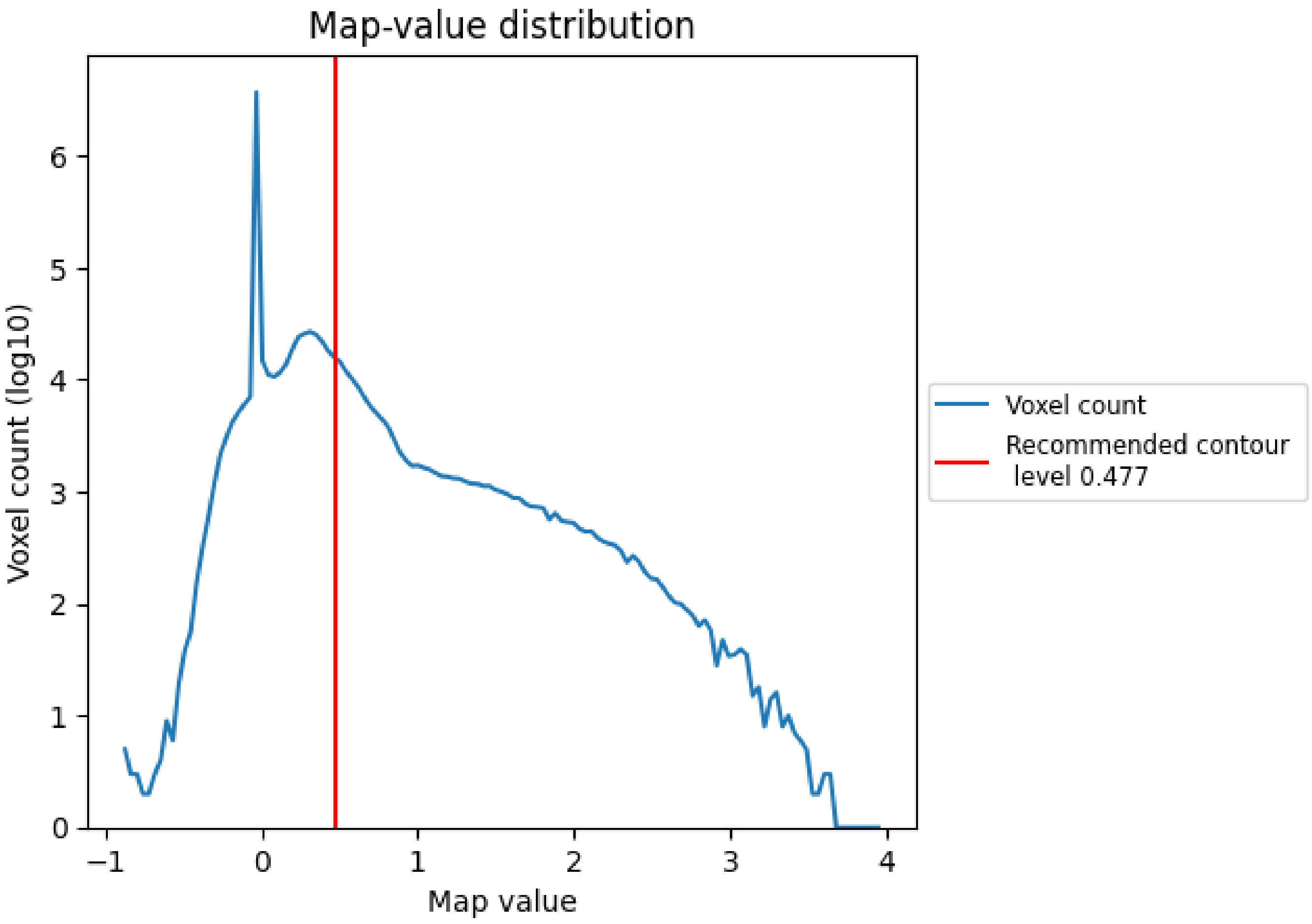
6.6 Mask visualisation

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

7 Map analysis ⓘ

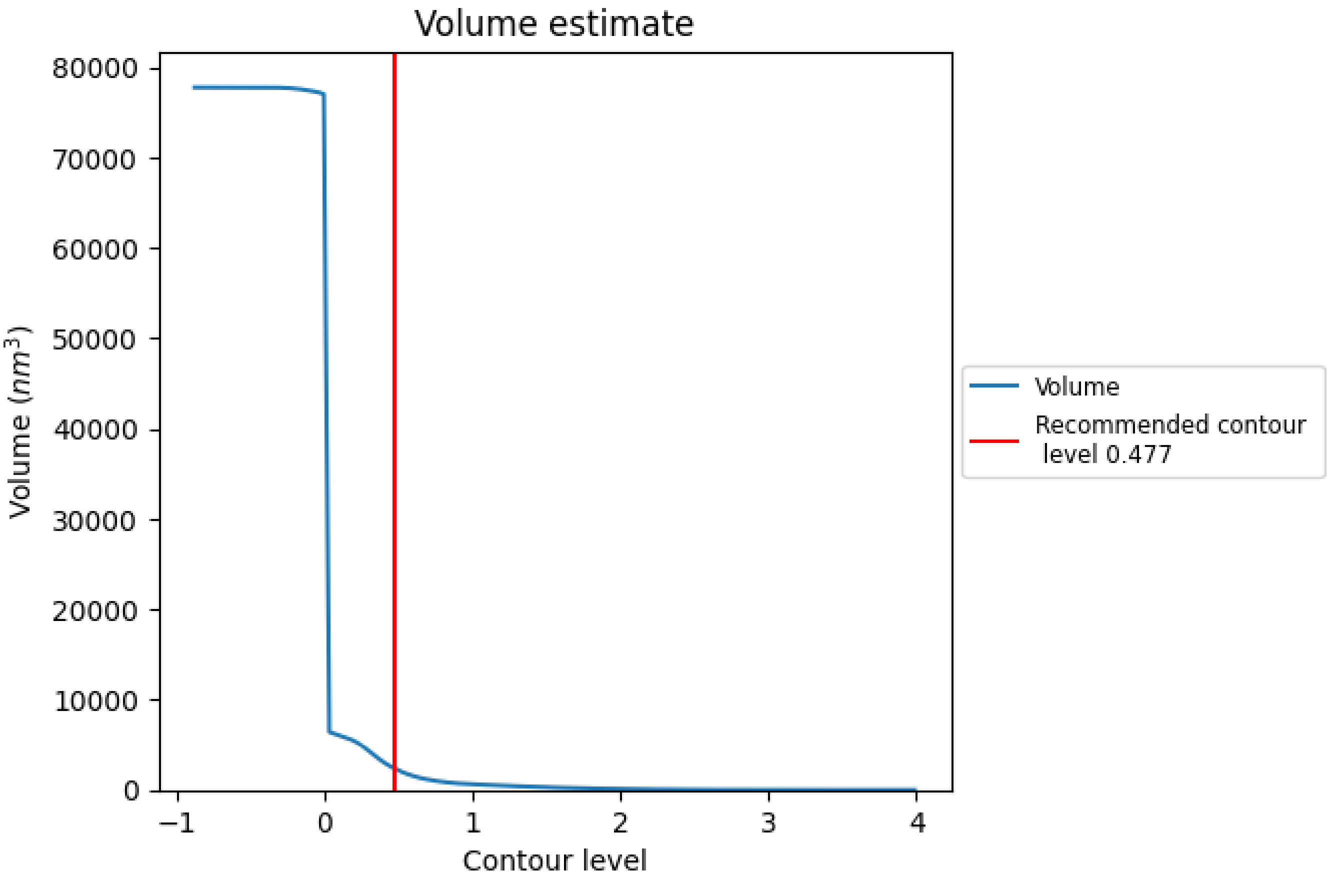
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution ⓘ



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.

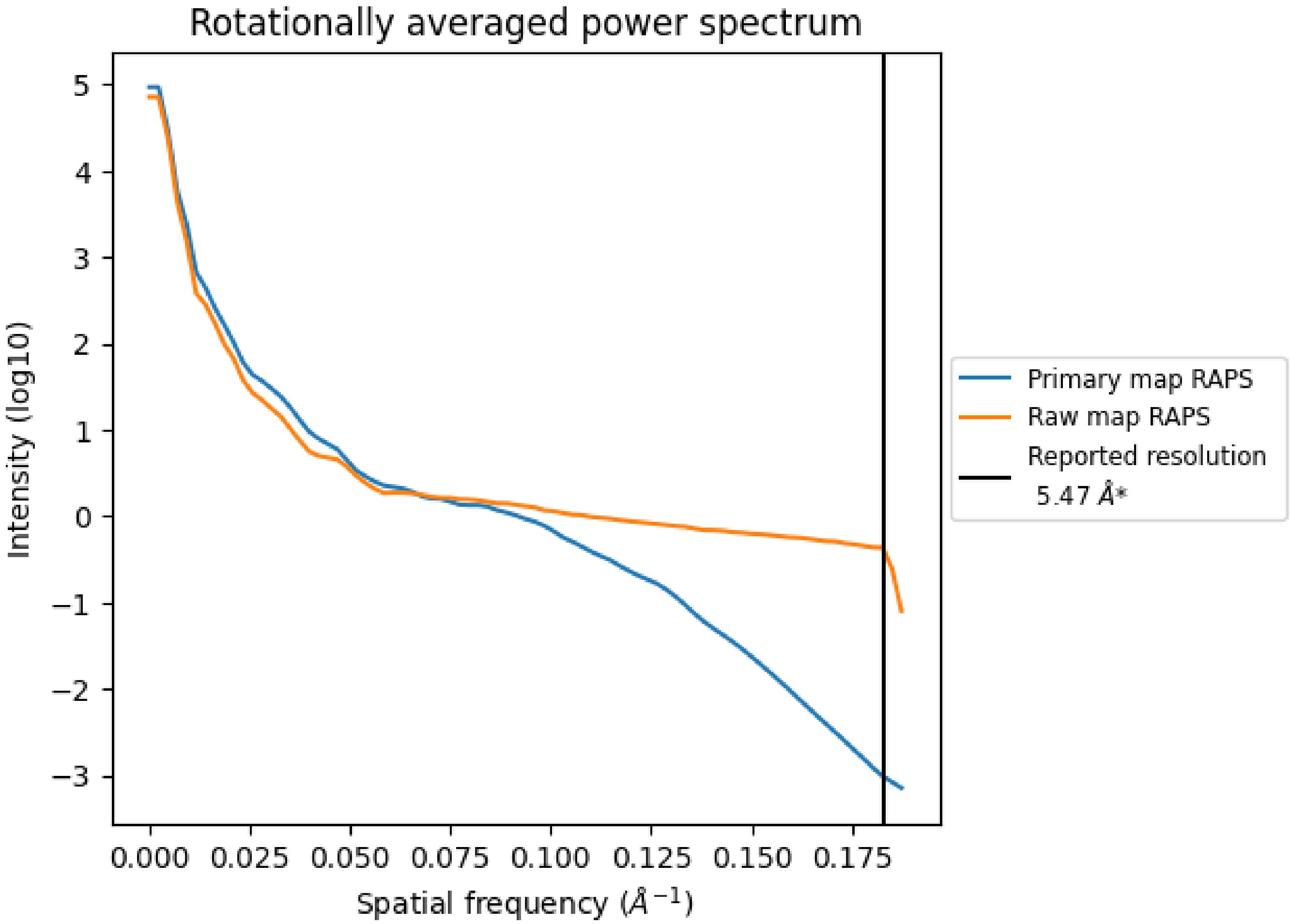
7.2 Volume estimate ⓘ



The volume at the recommended contour level is 2372 nm^3; this corresponds to an approximate mass of 2142 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 ⓘ

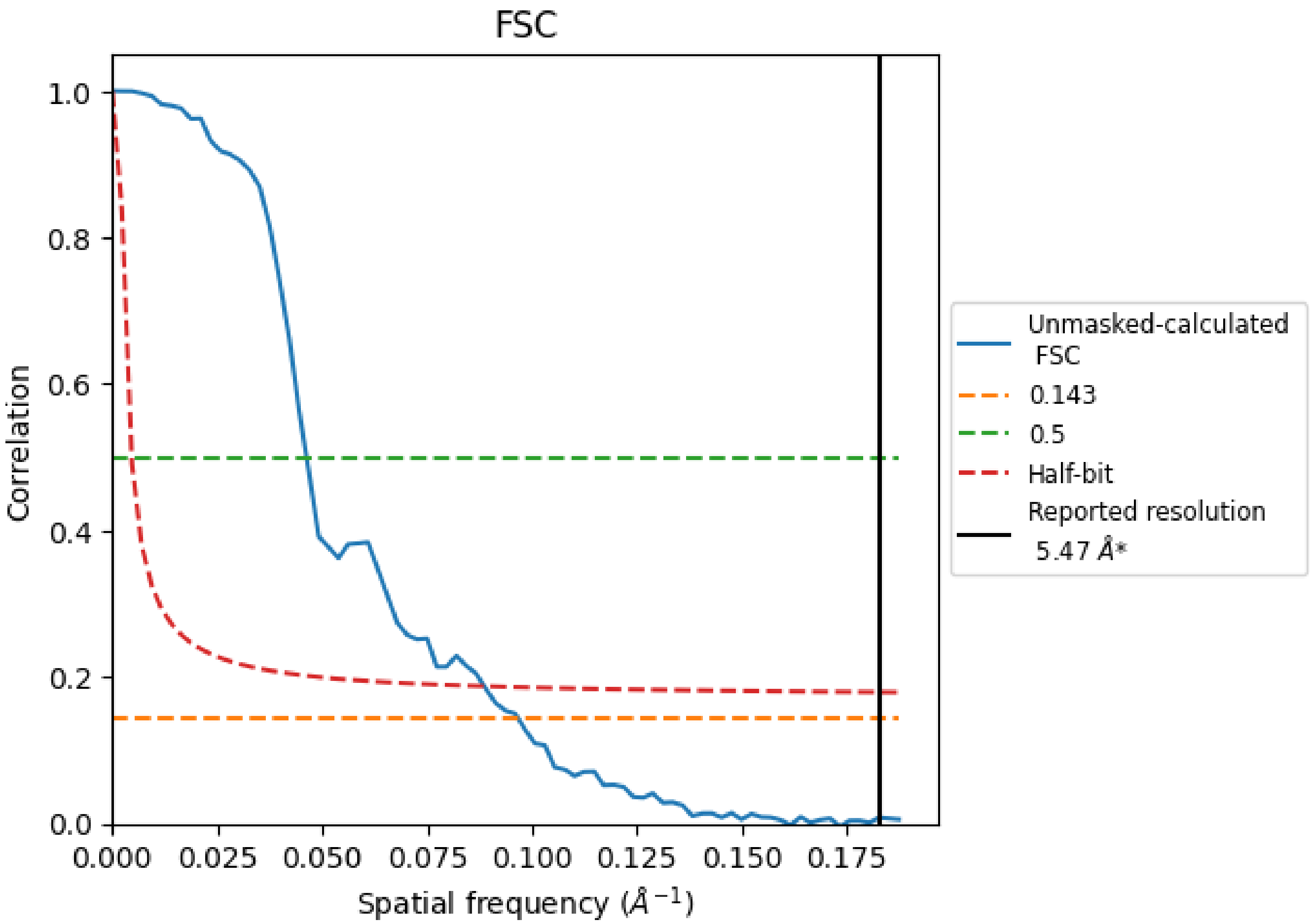


*Reported resolution corresponds to spatial frequency of 0.183 Å^{-1}

8 Fourier-Shell correlation ⓘ

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

8.1 FSC ⓘ



*Reported resolution corresponds to spatial frequency of 0.183 Å^{-1}

8.2 Resolution estimates (i)

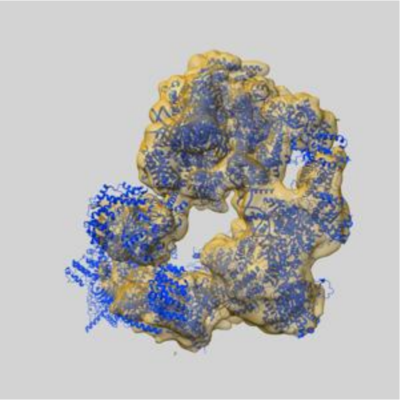
Resolution estimate (Å)	**Estimation criterion (FSC cut-off)**		
	0.143	0.5	Half-bit
Reported by author	5.47	-	-
Author-provided FSC curve	-	-	-
Unmasked-calculated*	10.33	21.65	11.29

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 10.33 differs from the reported value 5.47 by more than 10 %

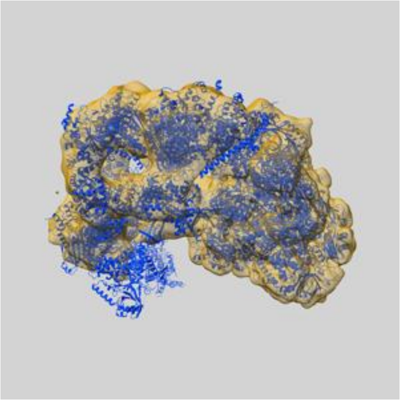
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-37395 and PDB model 8WAK. Per-residue inclusion information can be found in section 3 on page 14.

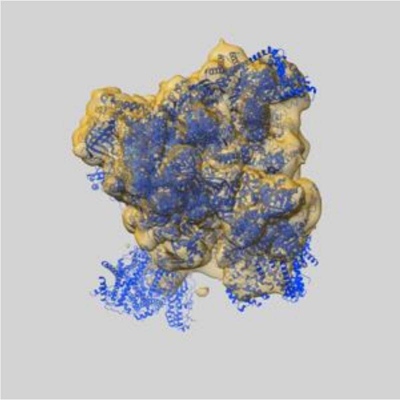
9.1 Map-model overlay



X



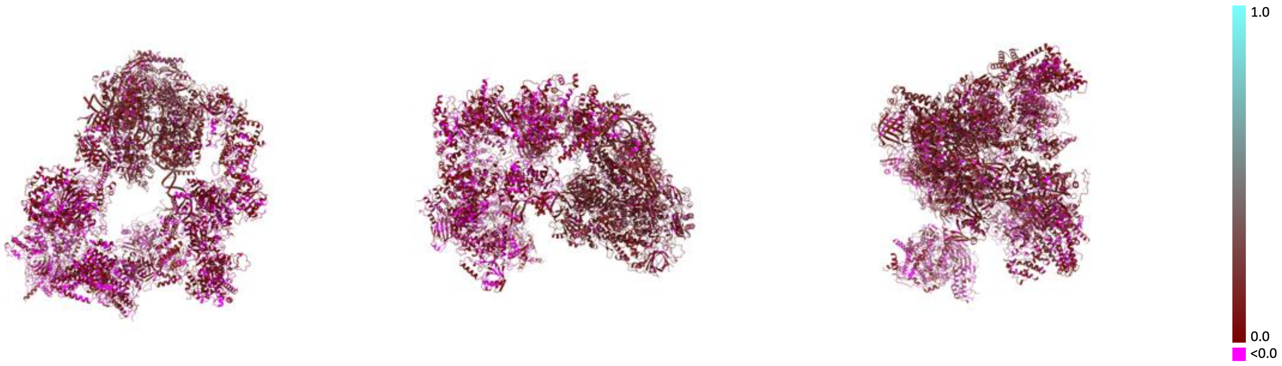
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.477 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.

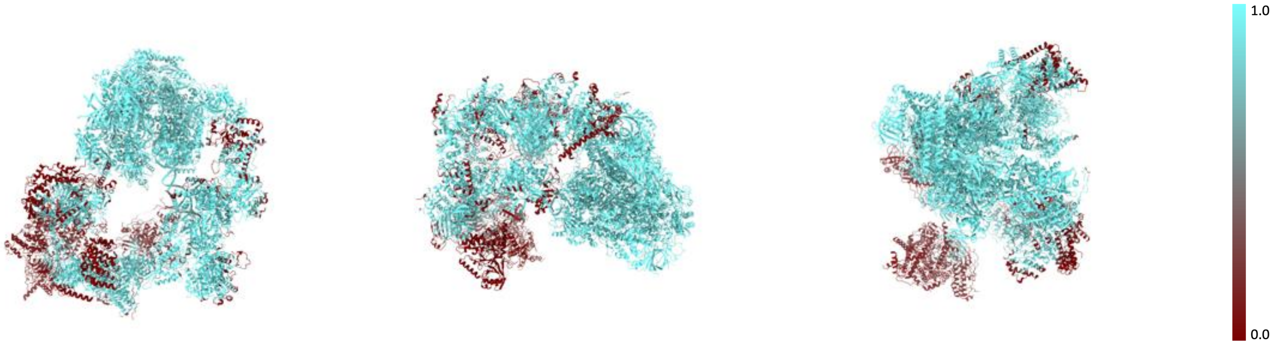
9.2 Q-score mapped to coordinate model ⓘ



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.

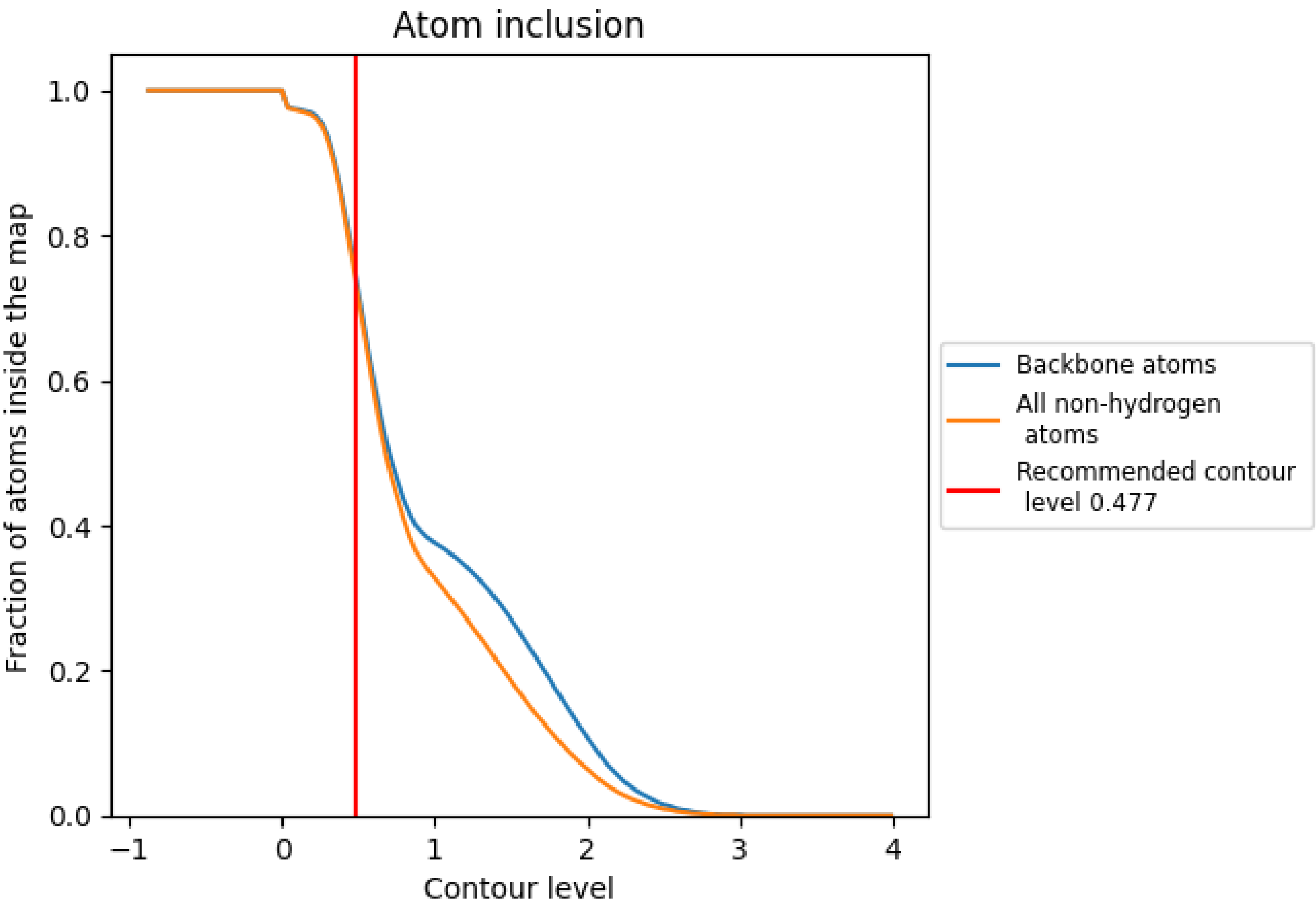
9.3 Atom inclusion mapped to coordinate model ⓘ



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 (0.477).

9.4 Atom inclusion ⓘ



At the recommended contour level, 75% of all backbone atoms, 74% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary ⓘ

The table lists the average atom inclusion at the recommended contour level (0.477) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7430	0.0790
0	0.1990	0.0570
1	0.8310	0.0510
2	0.9070	0.0560
3	0.8880	0.0360
4	0.7980	0.0540
5	0.5430	0.0320
6	0.7990	0.0370
7	0.8640	0.0540
A	0.1940	0.0270
B	0.8830	0.0400
D	0.3660	0.0370
E	0.7930	0.0460
F	0.3630	0.0230
G	0.0050	-0.0180
H	0.6150	0.0330
I	0.4590	0.0240
J	0.6000	0.0310
L	0.0590	0.0510
N	1.0000	0.1710
O	0.9530	0.0810
P	0.9990	0.1320
Q	0.9530	0.0880
R	0.9810	0.1380
S	0.9750	0.1110
T	0.9840	0.1290
U	0.9760	0.1120
V	0.9870	0.1180
X	0.9780	0.1750
Y	0.9250	0.1430
Z	0.7500	0.0690
c	0.1170	0.0100
d	0.0000	0.0100
e	0.1930	0.0110
f	0.6810	0.0390

Continued on next page...

Continued from previous page...

Chain	Atom inclusion	Q-score
i	🟥 0.1480	🟥 0.0210
j	🟥 0.0510	🟥 0.0500
k	🟥 0.0000	🟥 0.0060
l	🟥 0.0050	🟥 0.0380
m	🟥 0.0000	🟪 -0.0210
o	🟦 0.9740	🟥 0.1470
p	🟦 0.9640	🟥 0.1360
q	🟦 0.9810	🟥 0.1440
r	🟦 0.9900	🟥 0.1420
s	🟦 0.9820	🟥 0.1560
t	🟦 0.9880	🟥 0.1560
u	🟦 0.9940	🟥 0.1270
v	🟦 0.9870	🟥 0.1500
w	🟦 0.9800	🟥 0.1240
x	🟦 0.9540	🟥 0.1250
y	🟦 0.9830	🟥 0.1460
z	🟦 0.9920	🟥 0.1700