wwPDB EM Validation Summary Report (i)

Dec 21, 2024 – 01:11 PM EST

PDB ID	:	9DTR
EMDB ID	:	EMD-47157
Title	:	Structure of the yeast post-catalytic P complex spliceosome at 2.3 Angstrom resolution
Authors	:	Wilkinson, M.E.; Hoskins, A.A.
Deposited on	:	2024-10-01
Resolution	:	2.31 Å(reported)
Based on initial model	:	6EXN

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

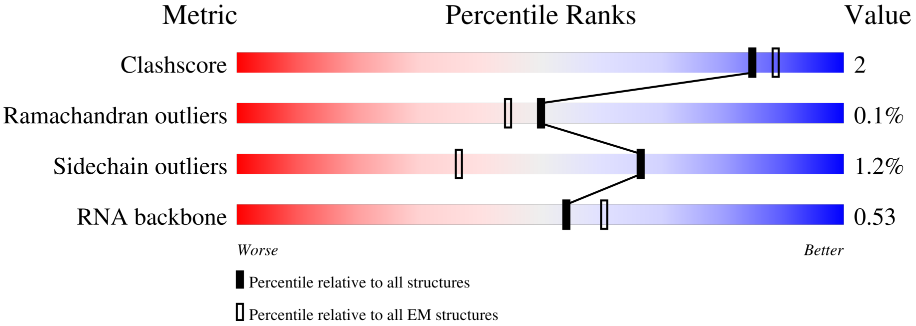
EMDB validation analysis	:	0.0.1.dev113
Mogul	:	2022.3.0, CSD as543be (2022)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

1 Overall quality at a glance ⓘ

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

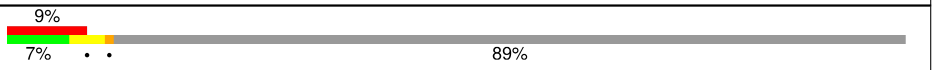
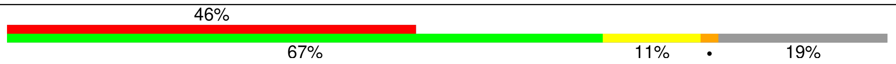
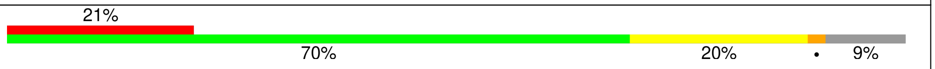
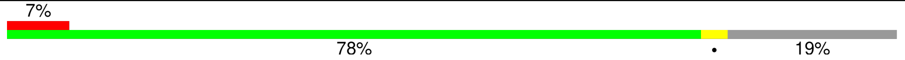
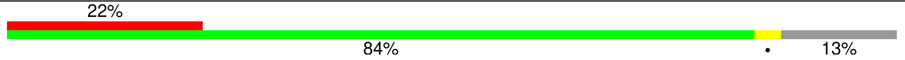
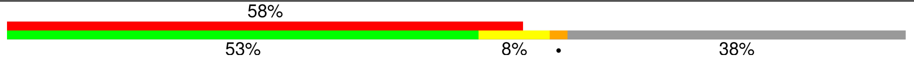
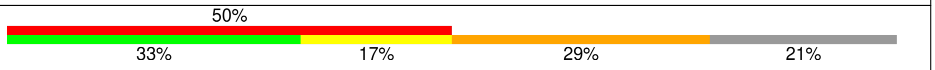
The reported resolution of this entry is 2.31 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



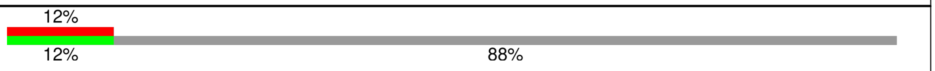
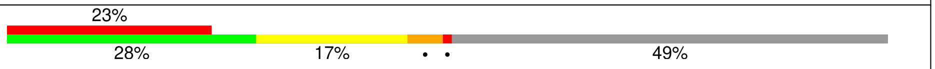
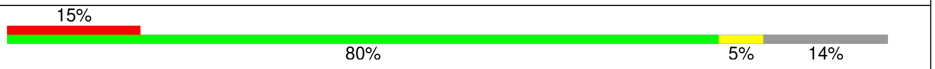
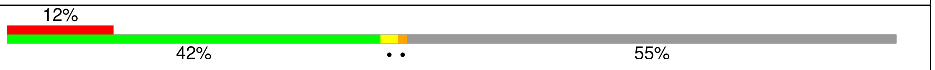
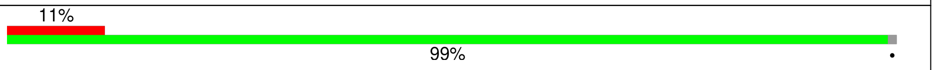
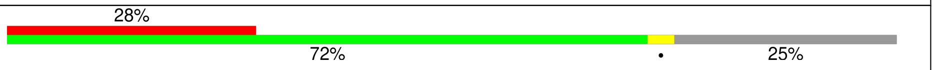
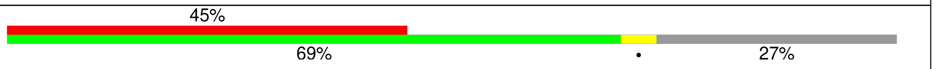
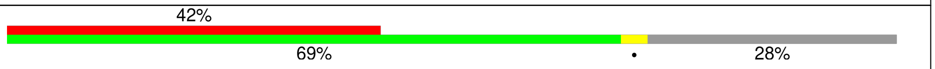
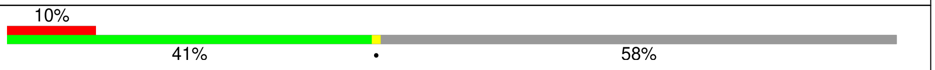
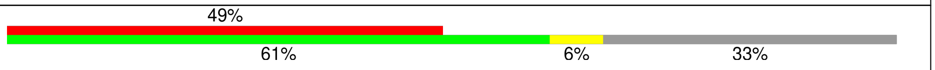
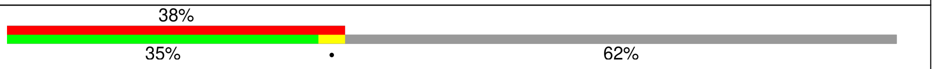
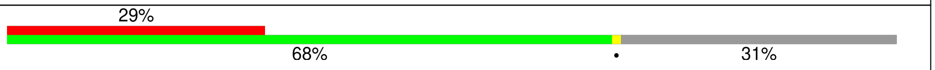
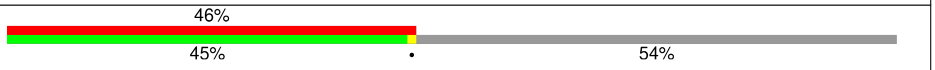
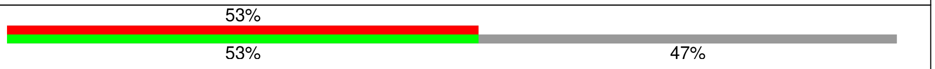
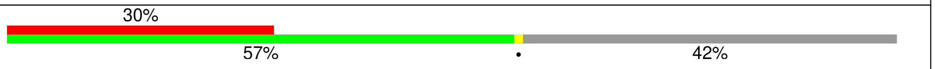
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415
RNA backbone	6643	2191

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	2	1175	
2	5	214	
3	6	112	
4	A	2413	
5	C	1008	
6	D	173	
7	E	42	

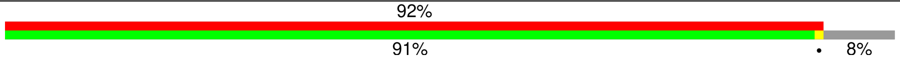
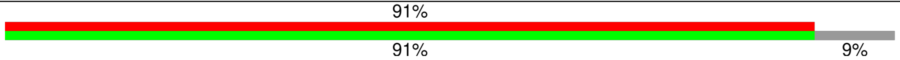
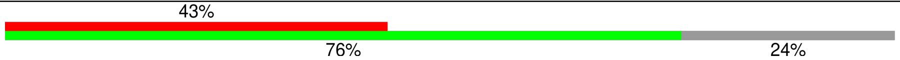
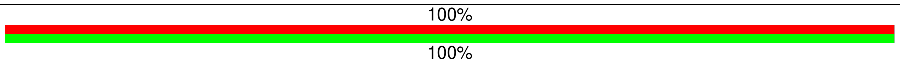
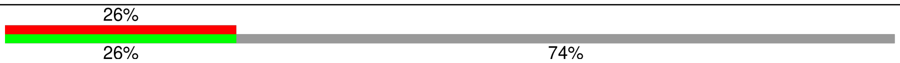
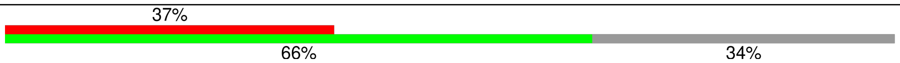
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	G	235	
9	H	577	
10	I	95	
11	J	451	
12	K	379	
13	L	157	
14	M	339	
15	N	364	
16	O	590	
17	P	175	
18	R	135	
19	S	687	
20	T	859	
21	V	1145	
22	W	238	
23	Y	111	
24	Z	140	
25	a	251	
26	b	196	
26	k	196	
27	c	382	
28	d	101	
28	n	101	
29	e	94	
29	p	94	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
30	f	86	
30	q	86	
31	g	77	
31	r	77	
32	h	146	
32	l	146	
33	j	110	
33	m	110	
34	o	455	
35	s	175	
36	t	503	
36	u	503	
36	v	503	
36	w	503	
37	y	215	

2 Entry composition

There are 43 unique types of molecules in this entry. The entry contains 91230 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 RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	2	132	2779	1242	456	949	132	0	0

- Molecule 2 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	5	173	3671	1642	638	1218	173	0	0

- Molecule 3 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
3	6	102	2170	972	386	710	102	0	0

- Molecule 4 is a protein called Pre-mRNA-splicing factor 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	A	1950	16079	10335	2766	2919	59	0	0

- Molecule 5 is a protein called Pre-mRNA-splicing factor SNU114.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	C	879	7040	4551	1173	1289	27	0	0

- Molecule 6 is a protein called Protein FYV6.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	D	108	914	573	168	172	1	0	0

- Molecule 7 is a RNA chain called UBC4 mRNA spliced exons.

Mol	Chain	Residues	Atoms	AltConf	Trace
7	E	33	Total C N O P 703 315 126 229 33	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	21	G	C	conflict	GB NM_001178430
E	22	C	A	conflict	GB NM_001178430

- Molecule 8 is a protein called Pre-mRNA-splicing factor ISY1.

Mol	Chain	Residues	Atoms	AltConf	Trace
8	G	28	Total C N O 139 83 28 28	0	0

- Molecule 9 is a protein called Pre-mRNA-splicing factor CWC22.

Mol	Chain	Residues	Atoms	AltConf	Trace
9	H	460	Total C N O S 3762 2415 622 707 18	0	0

- Molecule 10 is a RNA chain called UBC4 lariat-intron.

Mol	Chain	Residues	Atoms	AltConf	Trace
10	I	48	Total C N O P 1012 455 172 337 48	0	0

- Molecule 11 is a protein called Pre-mRNA-splicing factor PRP46.

Mol	Chain	Residues	Atoms	AltConf	Trace
11	J	386	Total C N O S 2990 1887 530 562 11	0	0

- Molecule 12 is a protein called Pre-mRNA-processing protein 45.

Mol	Chain	Residues	Atoms	AltConf	Trace
12	K	170	Total C N O S 1355 847 249 254 5	0	0

- Molecule 13 is a protein called Pre-mRNA-splicing factor BUD31.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
13	L	156	1283	803	239	231	10	0	0

- Molecule 14 is a protein called Pre-mRNA-splicing factor CWC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
14	M	255	2048	1297	362	378	11	0	0

- Molecule 15 is a protein called Pre-mRNA-splicing factor SLT11.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
15	N	264	2092	1331	364	382	15	0	0

- Molecule 16 is a protein called Pre-mRNA-splicing factor CEF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
16	O	424	2918	1805	546	560	7	0	0

- Molecule 17 is a protein called Pre-mRNA-splicing factor CWC15.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
17	P	74	607	382	120	104	1	0	0

- Molecule 18 is a protein called Pre-mRNA-splicing factor CWC21.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
18	R	91	510	305	101	103	1	0	0

- Molecule 19 is a protein called Pre-mRNA-splicing factor CLF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
19	S	604	4683	2996	818	855	14	0	0

- Molecule 20 is a protein called Pre-mRNA-splicing factor SYF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
20	T	707	5543	3563	922	1038	20	0	0

- Molecule 21 is a protein called Pre-mRNA-splicing factor ATP-dependent RNA helicase PRP22.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
21	V	764	6096	3860	1061	1145	30	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
V	635	ALA	SER	engineered mutation	UNP P24384

- Molecule 22 is a protein called U2 small nuclear ribonucleoprotein A'.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
22	W	170	1383	866	253	257	7	0	0

- Molecule 23 is a protein called U2 small nuclear ribonucleoprotein B".

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
23	Y	86	697	448	122	124	3	0	0

- Molecule 24 is a protein called Pre-mRNA-splicing factor NTC20.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
24	Z	53	424	262	77	82	3	0	0

- Molecule 25 is a protein called Pre-mRNA-splicing factor 18.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
25	a	172	1380	884	245	247	4	0	0

- Molecule 26 is a protein called Small nuclear ribonucleoprotein-associated protein B.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
26	b	90	725	461	131	130	3	0	0
26	k	104	843	533	157	150	3	0	0

- Molecule 27 is a protein called Pre-mRNA-splicing factor SLU7.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
27	c	223	1893	1183	344	358	8	0	0

- Molecule 28 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
28	d	81	624	398	107	117	2	0	0
28	n	80	615	392	105	116	2	0	0

- Molecule 29 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
29	e	86	665	432	106	124	3	0	0
29	p	83	646	420	102	121	3	0	0

- Molecule 30 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
30	f	74	592	381	103	107	1	0	0
30	q	74	592	381	103	107	1	0	0

- Molecule 31 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
31	g	71	551	348	96	105	2	0	0
31	r	70	544	344	95	103	2	0	0

- Molecule 32 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					AltConf	Trace
32	h	107	Total 826	C 522	N 145	O 157	S 2	0	0
32	l	107	Total 826	C 522	N 145	O 157	S 2	0	0

- Molecule 33 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
33	j	91	Total 747	C 478	N 135	O 130	S 4	0	0
33	m	91	Total 747	C 478	N 135	O 130	S 4	0	0

- Molecule 34 is a protein called Pre-mRNA-processing factor 17.

Mol	Chain	Residues	Atoms					AltConf	Trace
34	o	348	Total 2821	C 1794	N 501	O 517	S 9	0	0

- Molecule 35 is a protein called Pre-mRNA-splicing factor SNT309.

Mol	Chain	Residues	Atoms				AltConf	Trace
35	s	175	Total 870	C 519	N 175	O 176	0	0

- Molecule 36 is a protein called Pre-mRNA-processing factor 19.

Mol	Chain	Residues	Atoms				AltConf	Trace
36	t	132	Total 658	C 394	N 132	O 132	0	0
36	u	132	Total 658	C 394	N 132	O 132	0	0
36	v	132	Total 658	C 394	N 132	O 132	0	0
36	w	132	Total 658	C 394	N 132	O 132	0	0

- Molecule 37 is a protein called Pre-mRNA-splicing factor SYF2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
37	y	141	1075	667	198	209	1	0	0

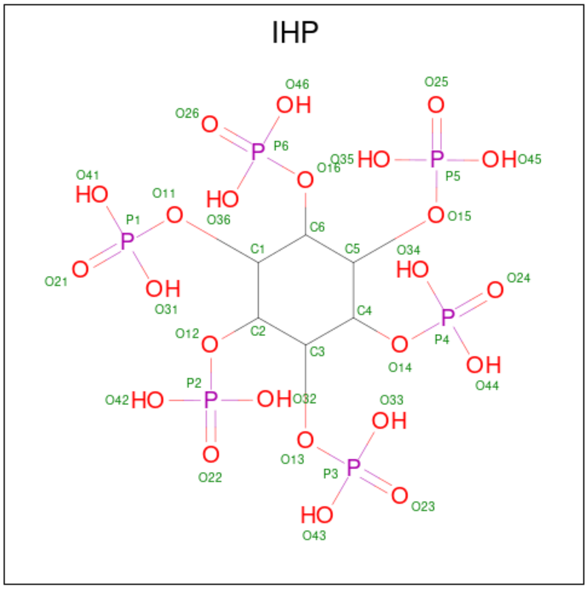
- Molecule 38 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
38	6	1	1	1	0
			Total	Mg	
38	C	1	1	1	0

- Molecule 39 is POTASSIUM ION (three-letter code: K) (formula: K).

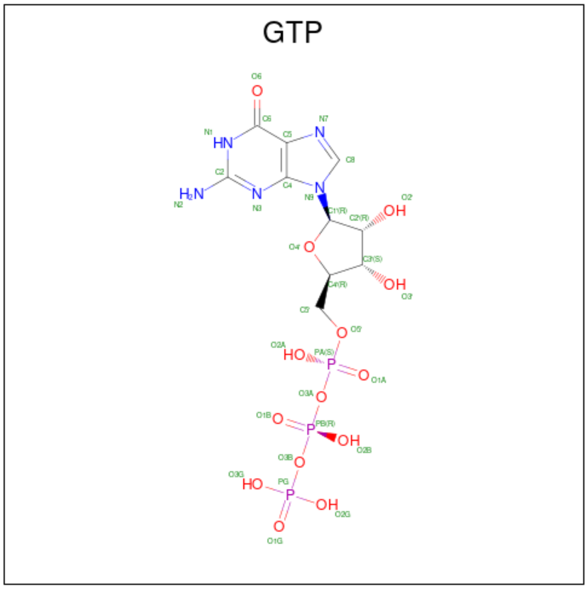
Mol	Chain	Residues	Atoms		AltConf
			Total	K	
39	6	3	3	3	0
			Total	K	
39	E	1	1	1	0

- Molecule 40 is INOSITOL HEXAKISPHOSPHATE (three-letter code: IHP) (formula: $C_6H_{18}O_{24}P_6$).



Mol	Chain	Residues	Atoms				AltConf
			Total	C	O	P	
40	A	1	36	6	24	6	0
			Total	C	O	P	
40	S	1	36	6	24	6	0

- Molecule 41 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).



Mol	Chain	Residues	Atoms	AltConf
41	C	1	Total C N O P 32 10 5 14 3	0

- Molecule 42 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	AltConf
42	L	3	Total Zn 3 3	0
42	M	1	Total Zn 1 1	0
42	N	2	Total Zn 2 2	0
42	c	1	Total Zn 1 1	0

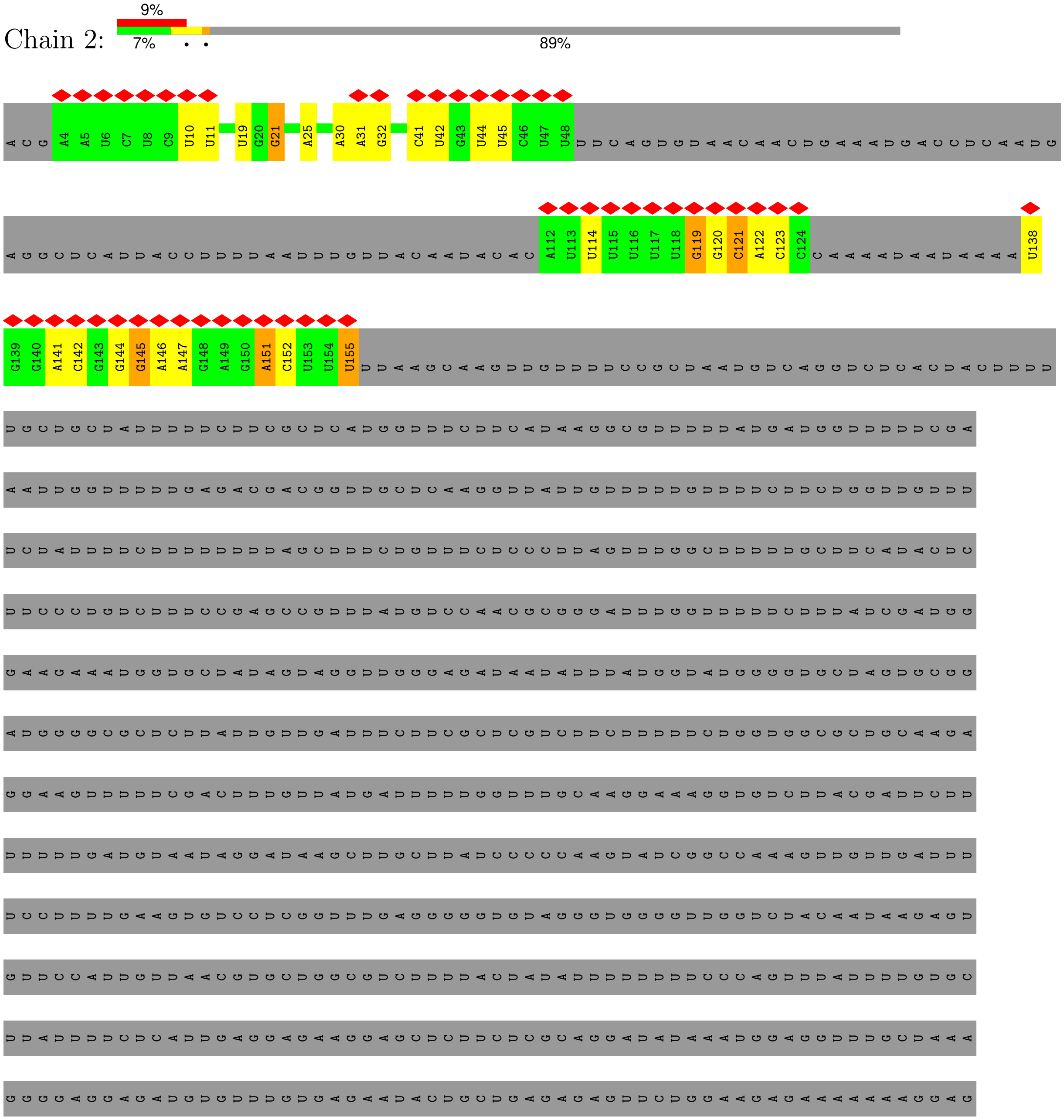
- Molecule 43 is water.

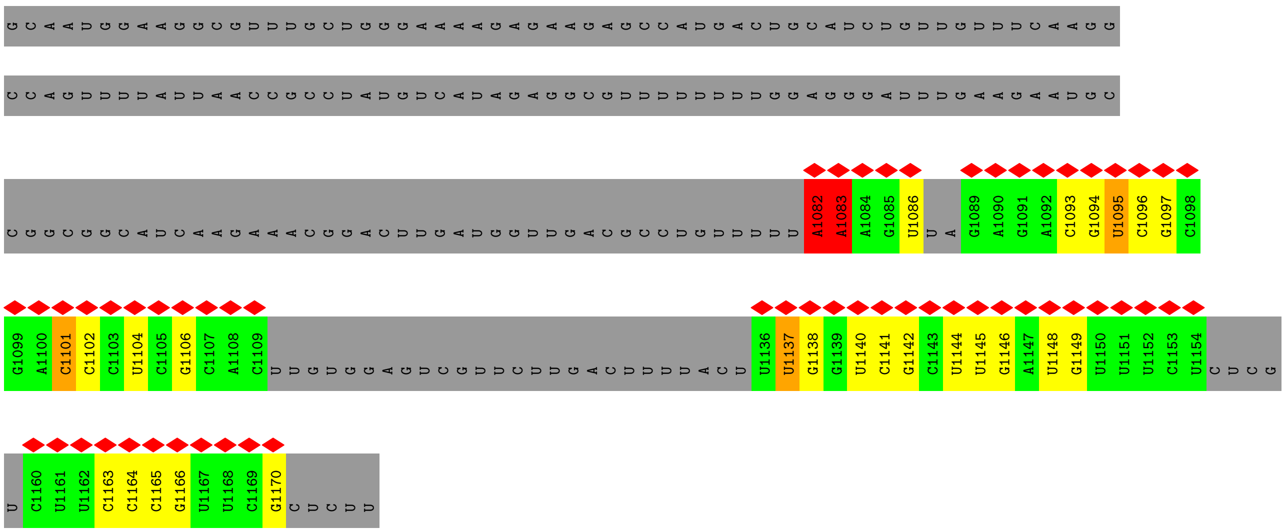
Mol	Chain	Residues	Atoms	AltConf
43	6	1	Total O 1 1	0

3 Residue-property plots ⓘ

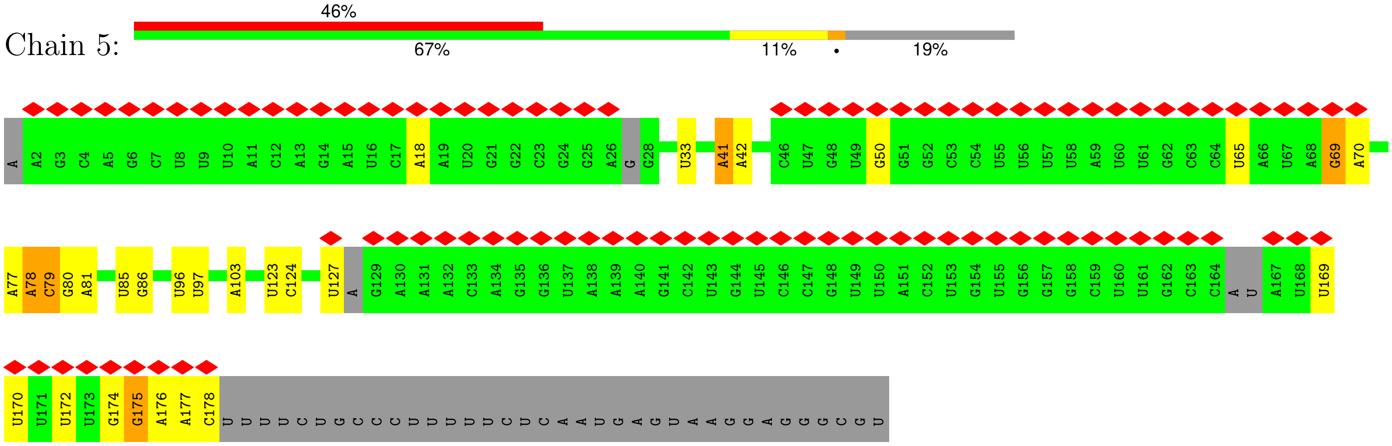
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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: U2 snRNA

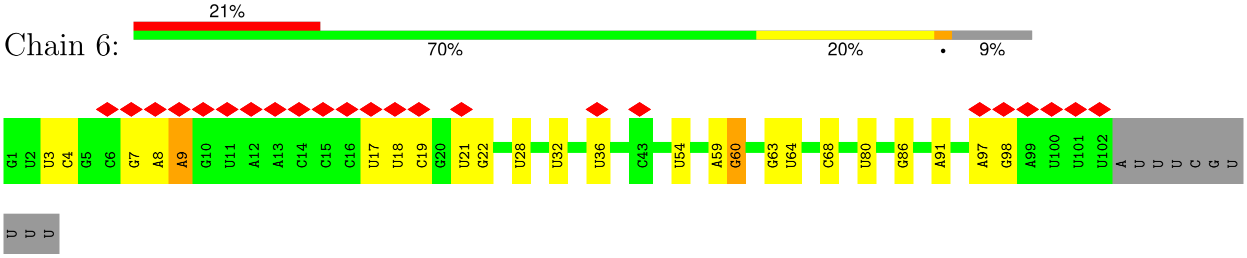




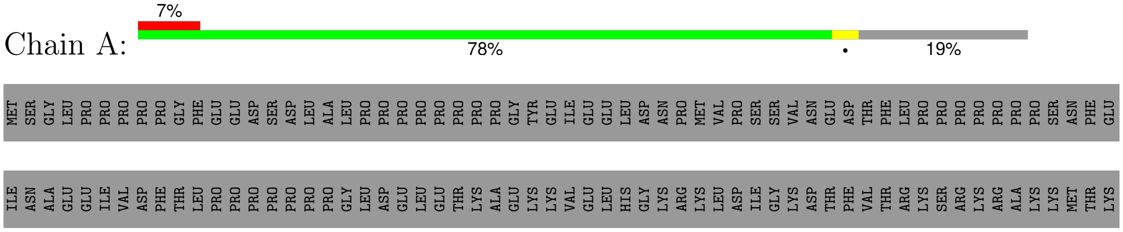
- Molecule 2: U5 snRNA

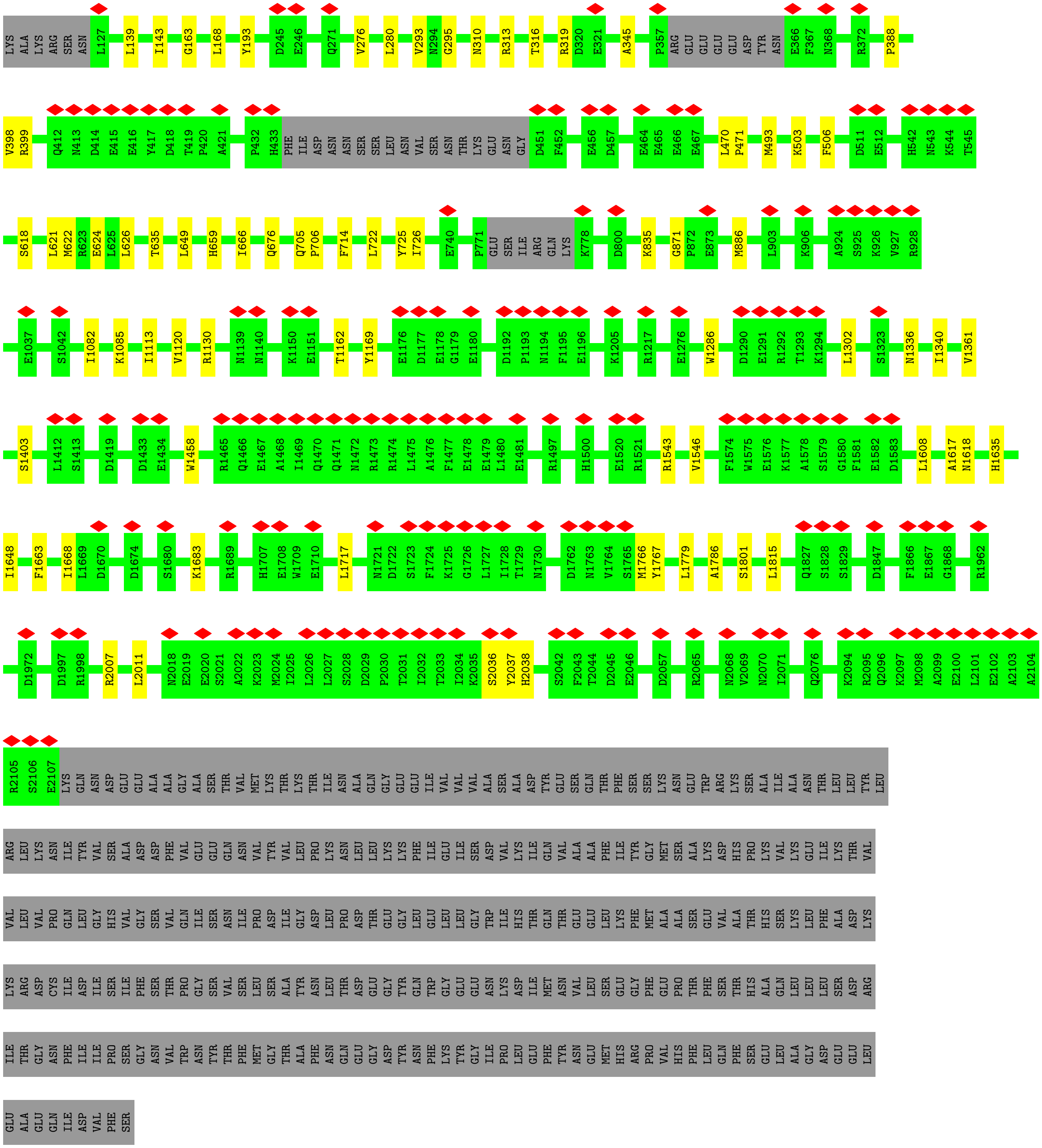


- Molecule 3: U6 snRNA



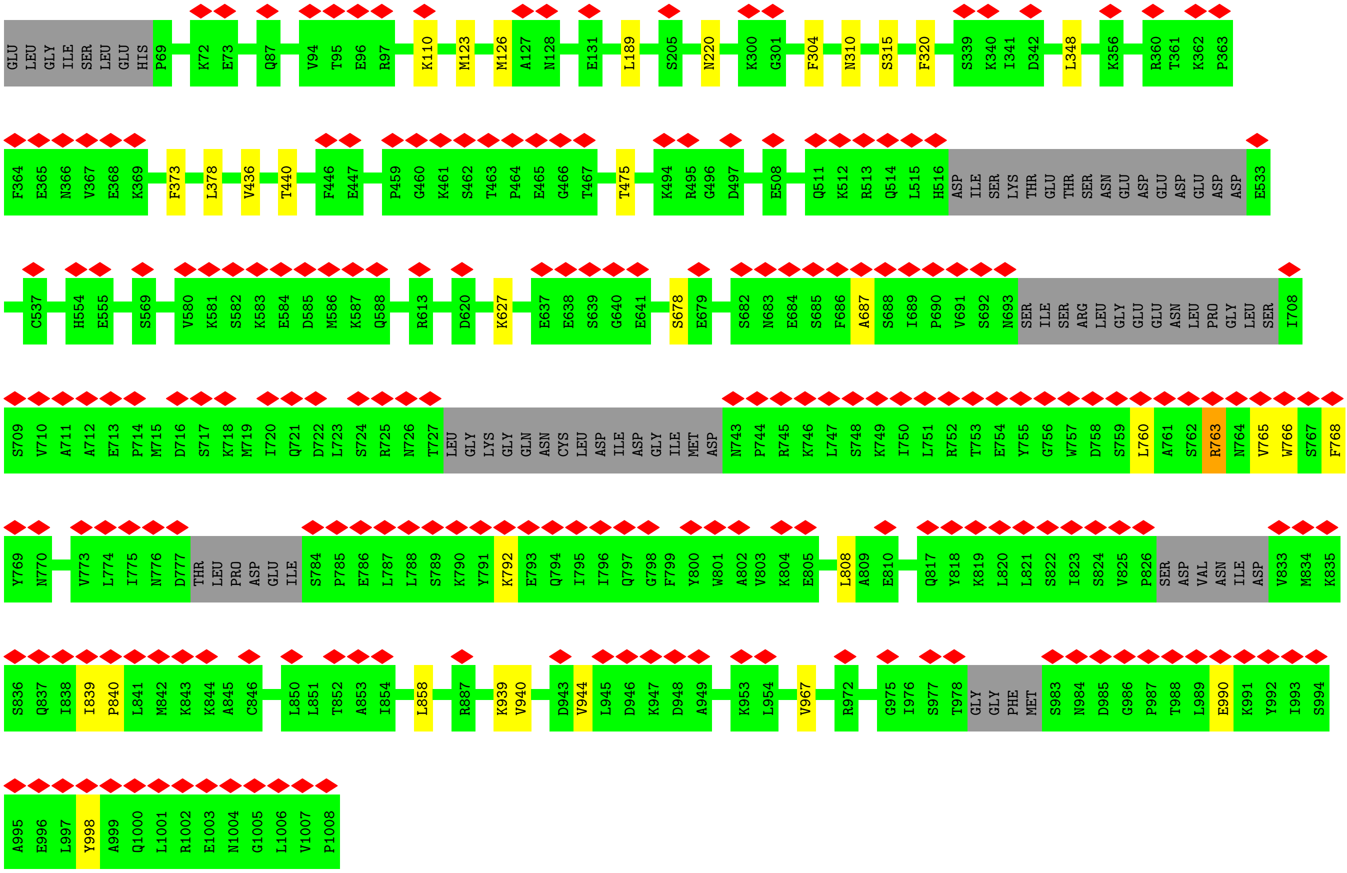
- Molecule 4: Pre-mRNA-splicing factor 8



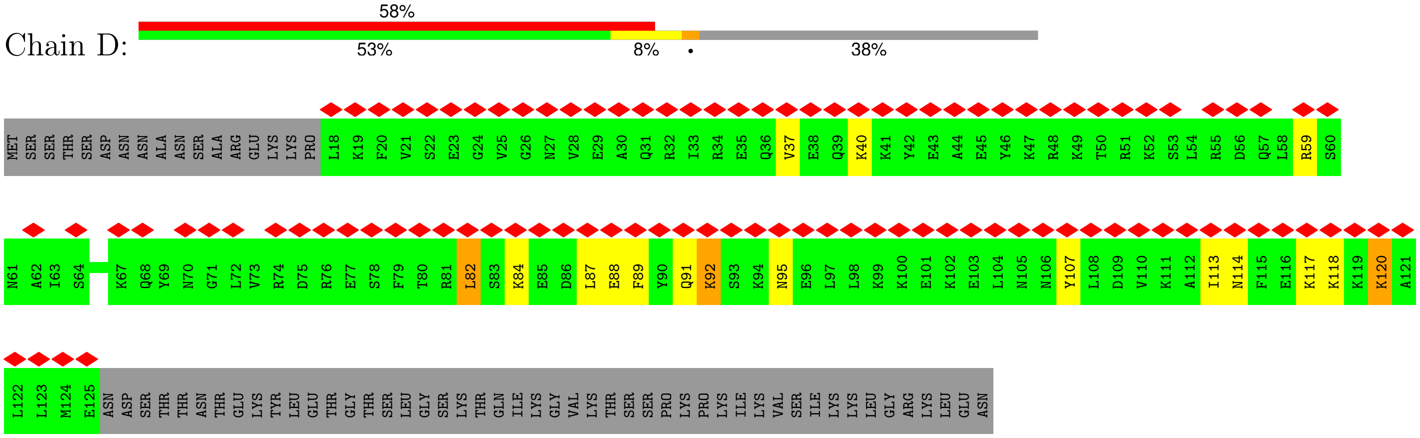


- Molecule 5: Pre-mRNA-splicing factor SNU114

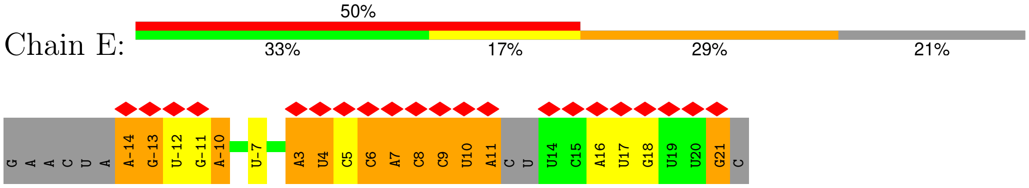
Chain C:



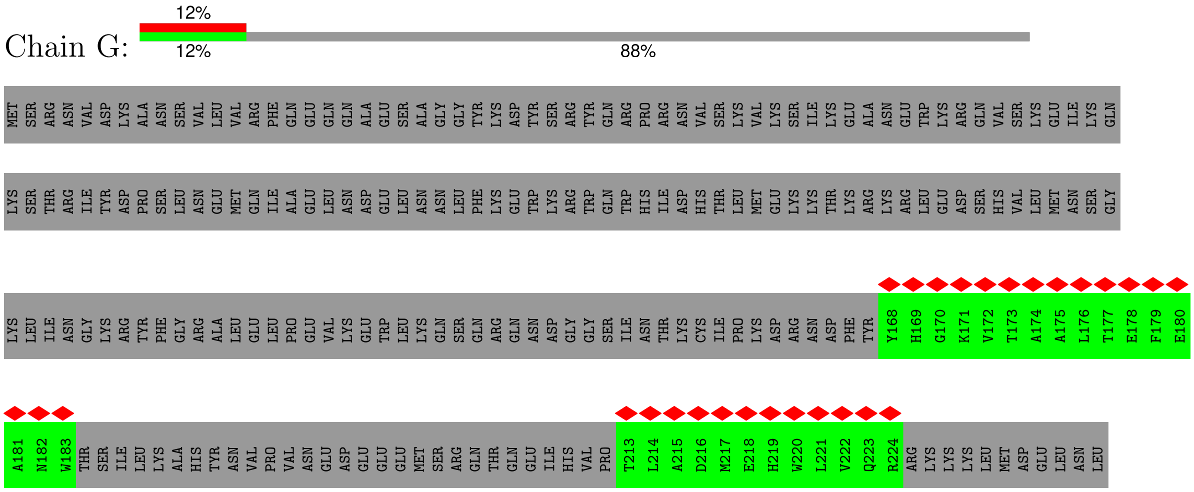
- Molecule 6: Protein FYV6



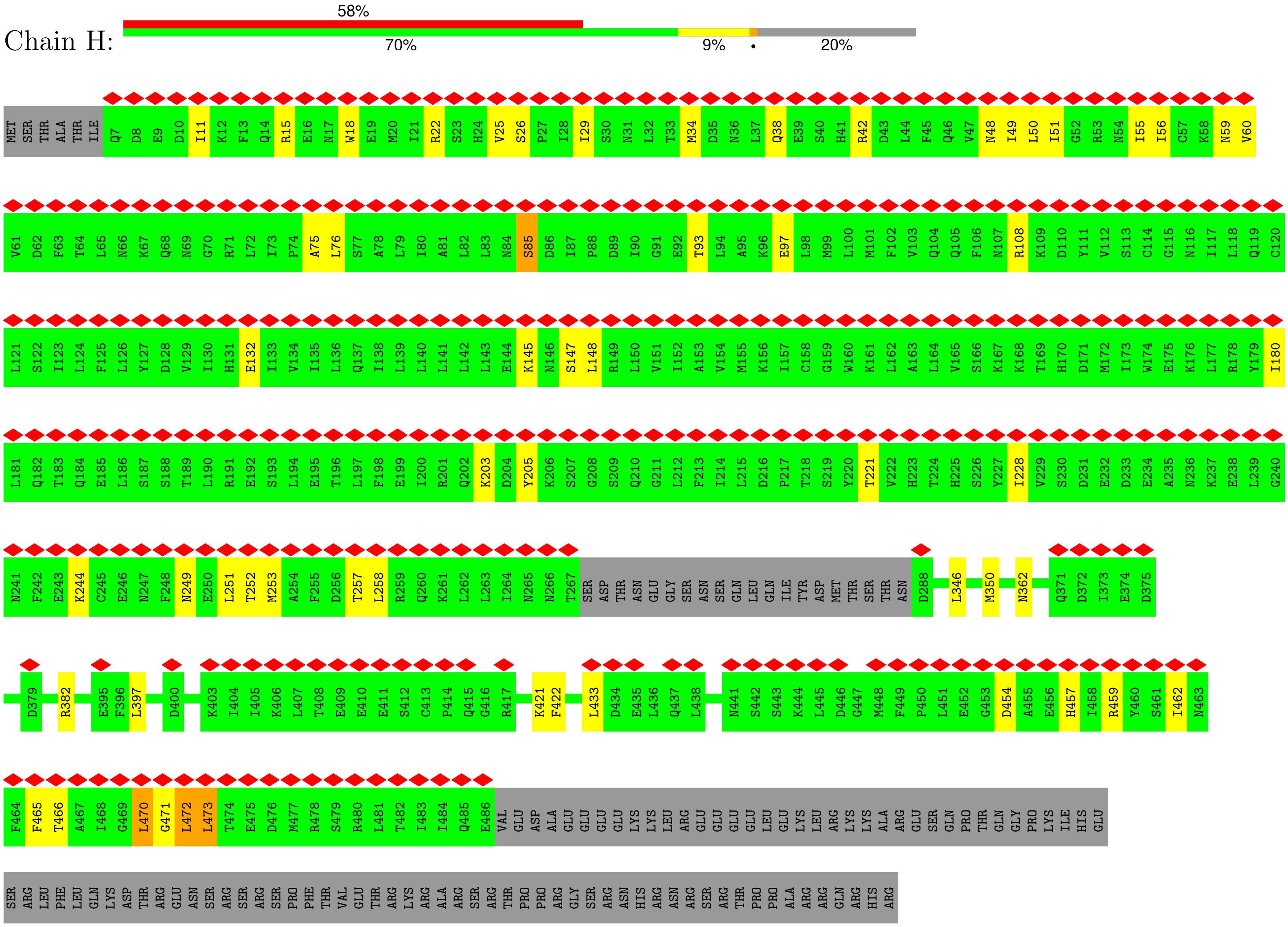
- Molecule 7: UBC4 mRNA spliced exons



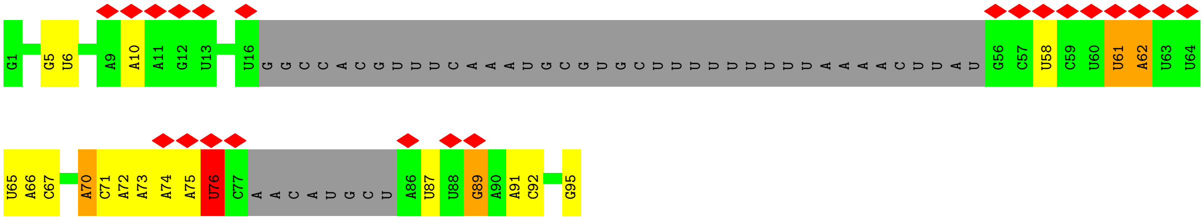
- Molecule 8: Pre-mRNA-splicing factor ISY1



- Molecule 9: Pre-mRNA-splicing factor CWC22

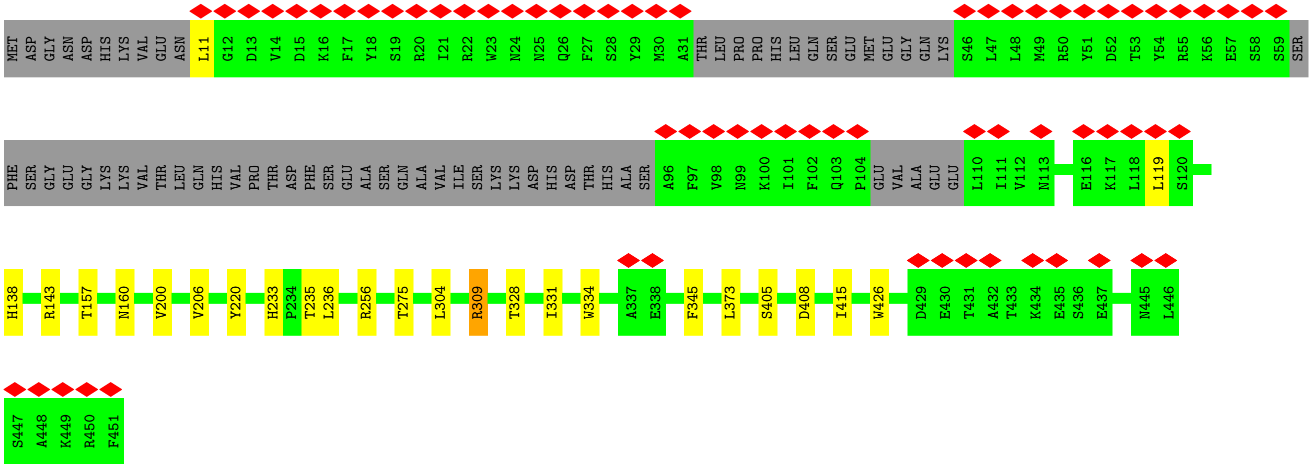


- Molecule 10: UBC4 lariat-intron



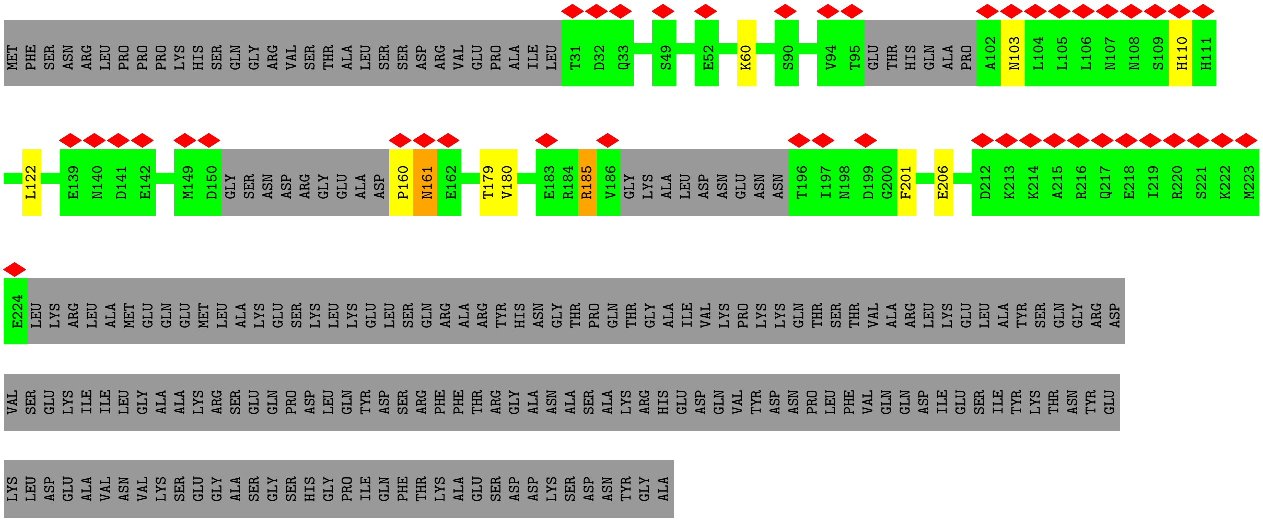
- Molecule 11: Pre-mRNA-splicing factor PRP46

Chain J:



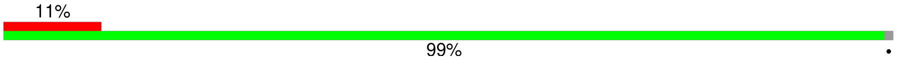
- Molecule 12: Pre-mRNA-processing protein 45

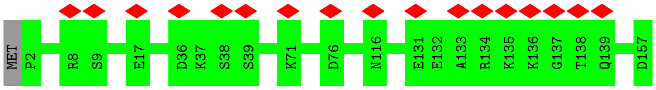
Chain K:



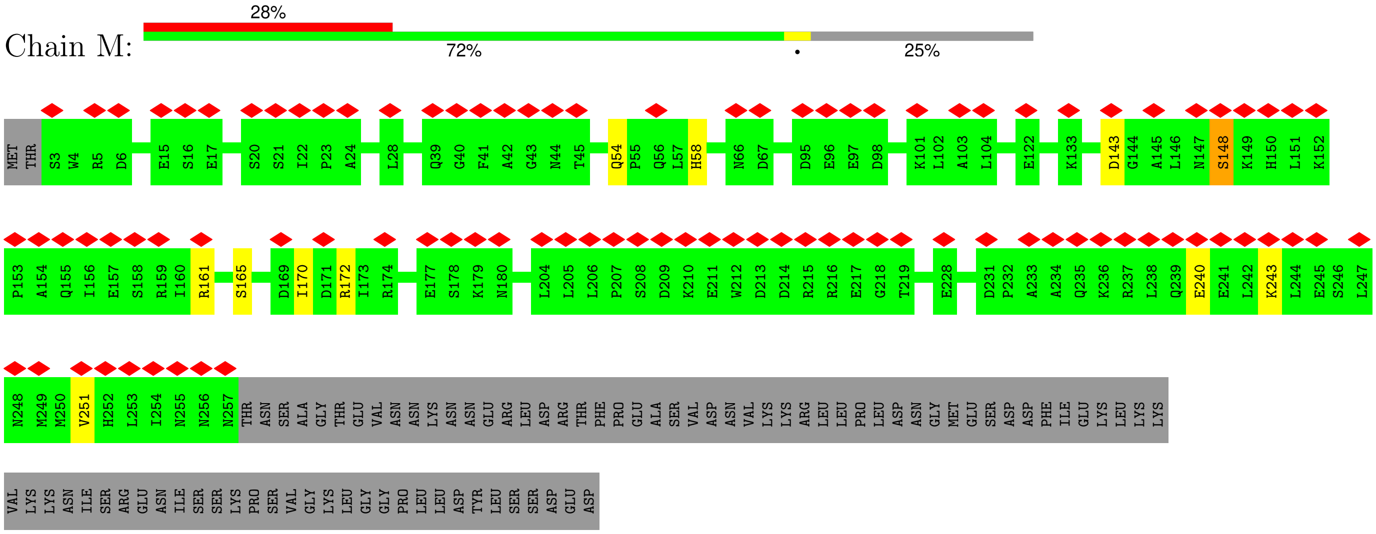
- Molecule 13: Pre-mRNA-splicing factor BUD31

Chain L:

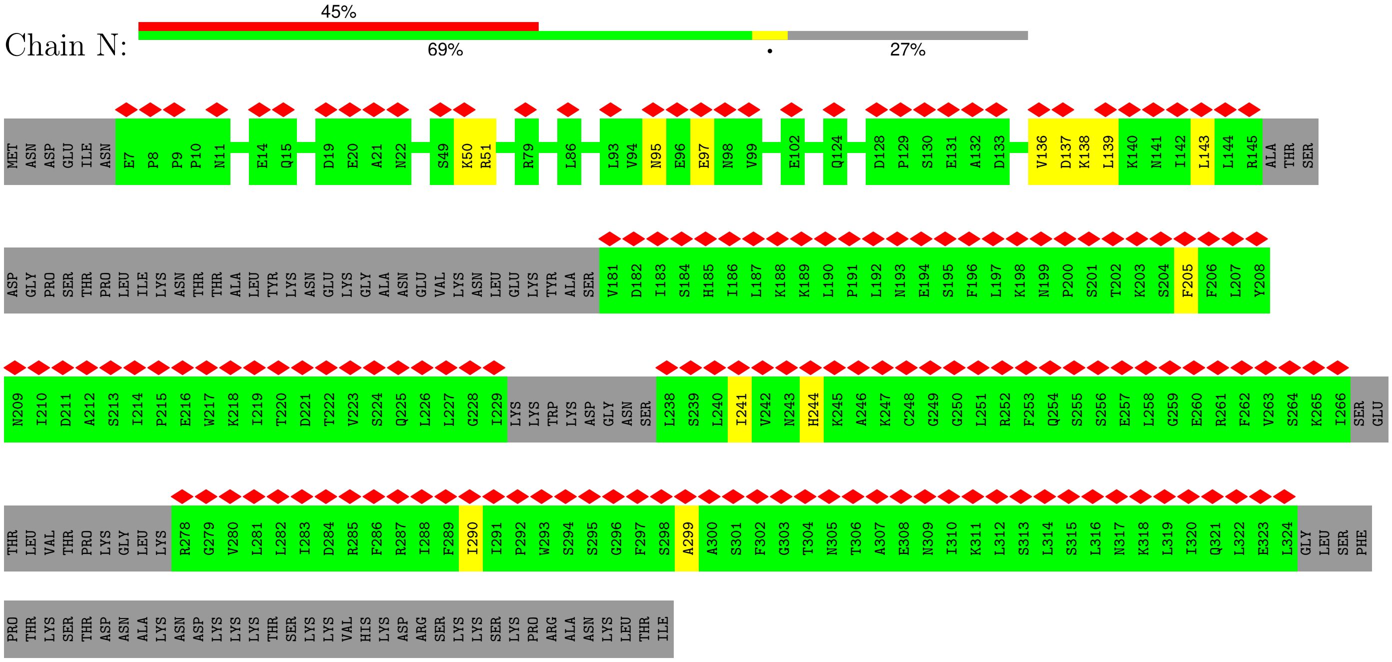




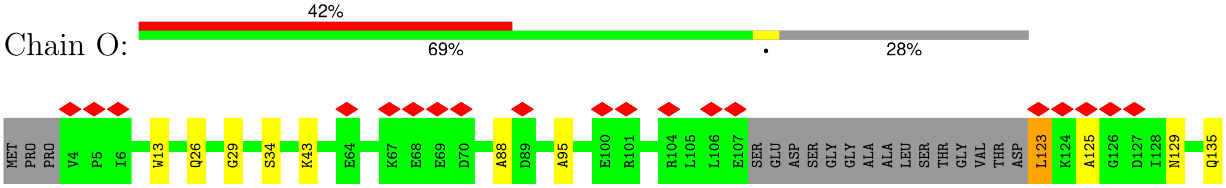
- Molecule 14: Pre-mRNA-splicing factor CWC2

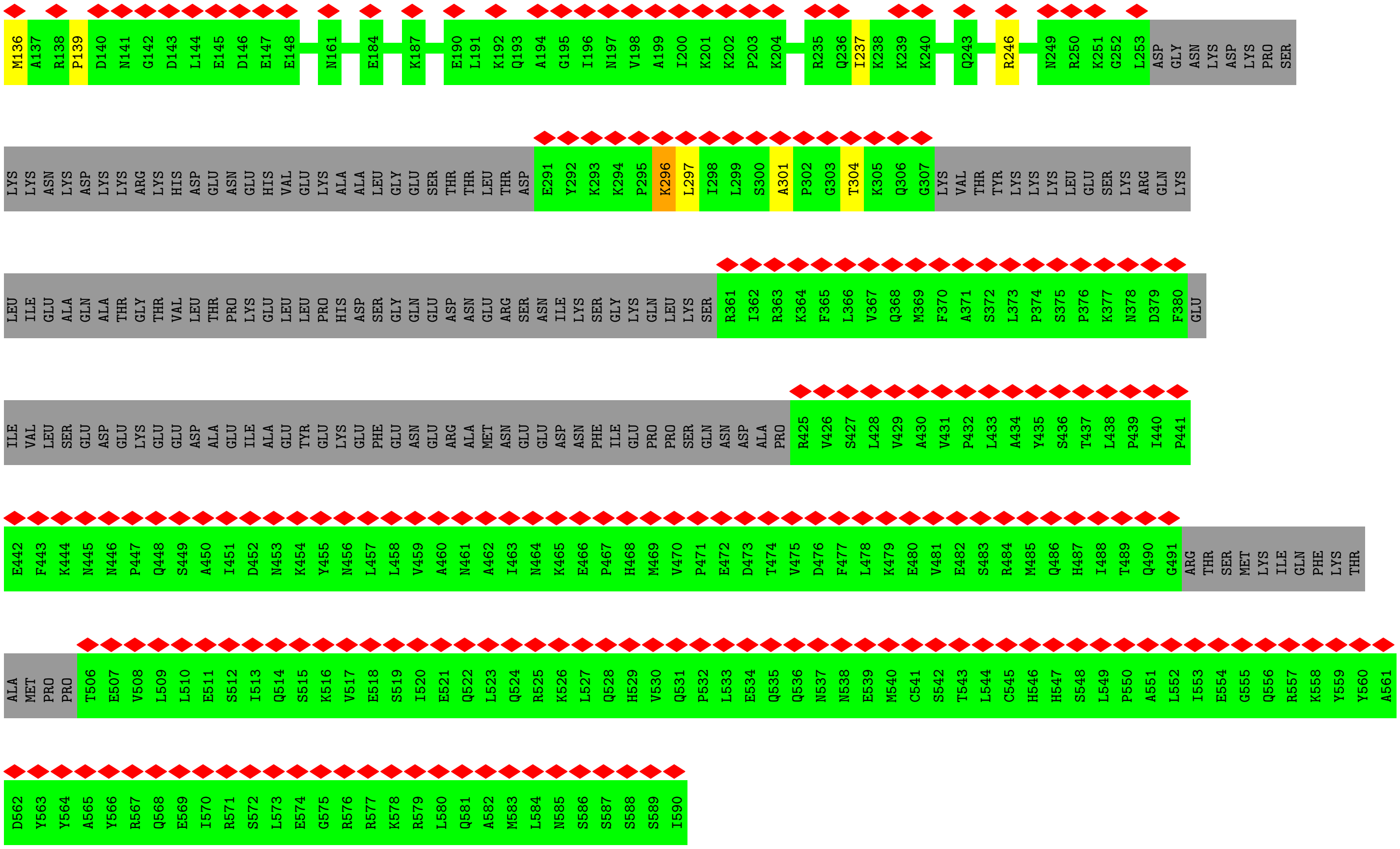


- Molecule 15: Pre-mRNA-splicing factor SLT11



- Molecule 16: Pre-mRNA-splicing factor CEF1



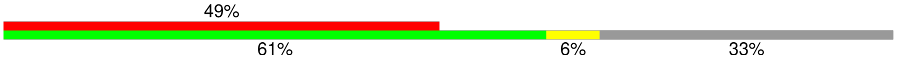


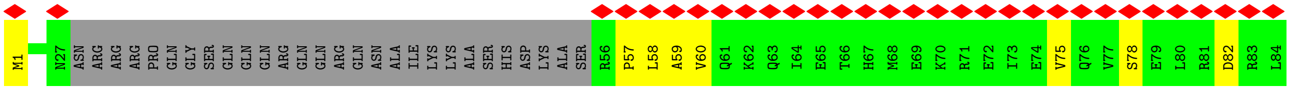
- Molecule 17: Pre-mRNA-splicing factor CWC15

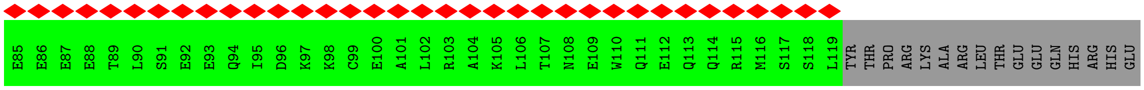
Chain P:

- Molecule 18: Pre-mRNA-splicing factor CWC21

Chain R:

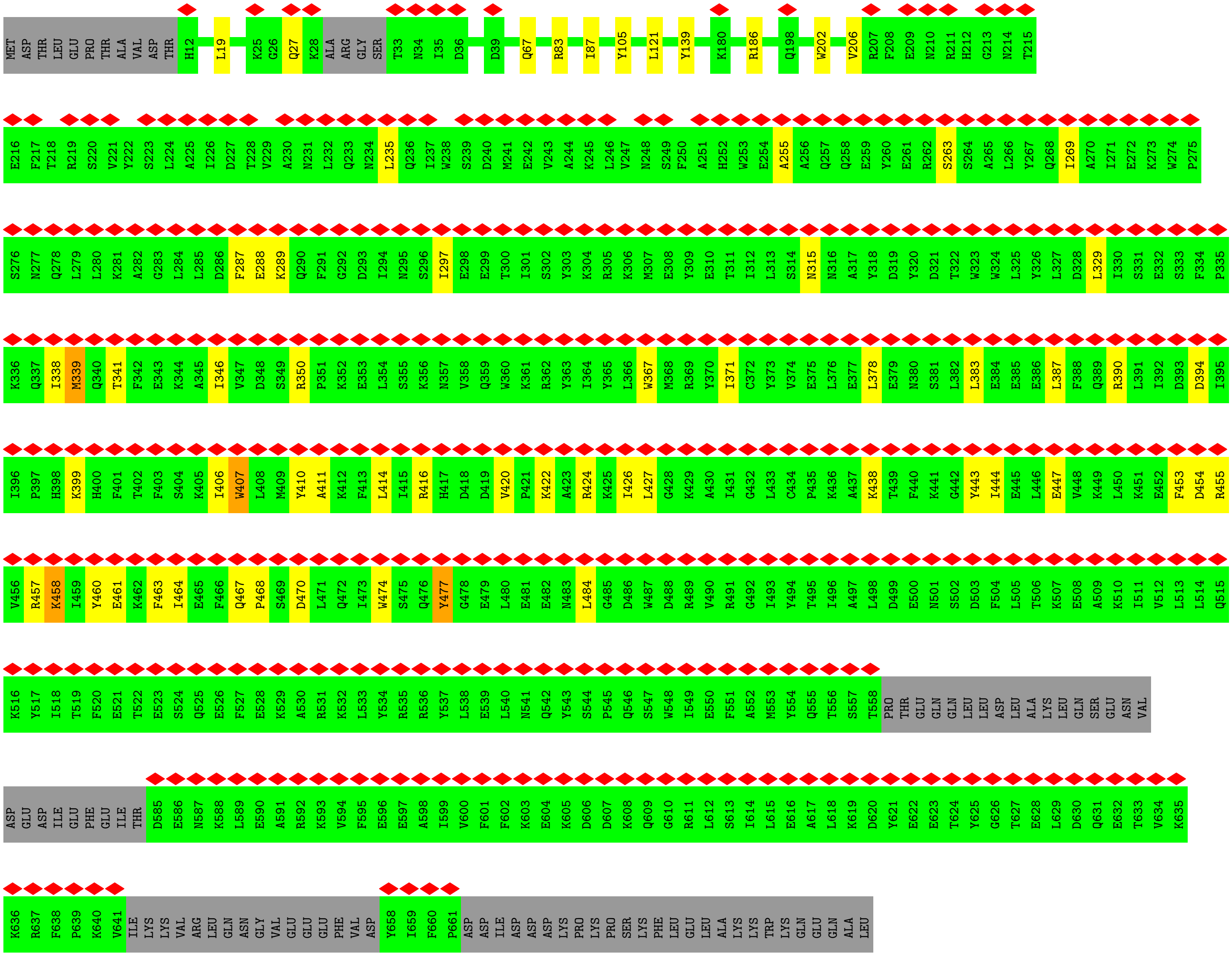






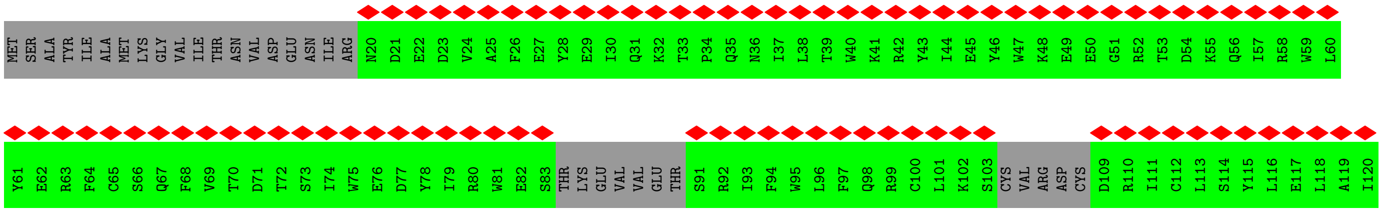
- Molecule 19: Pre-mRNA-splicing factor CLF1

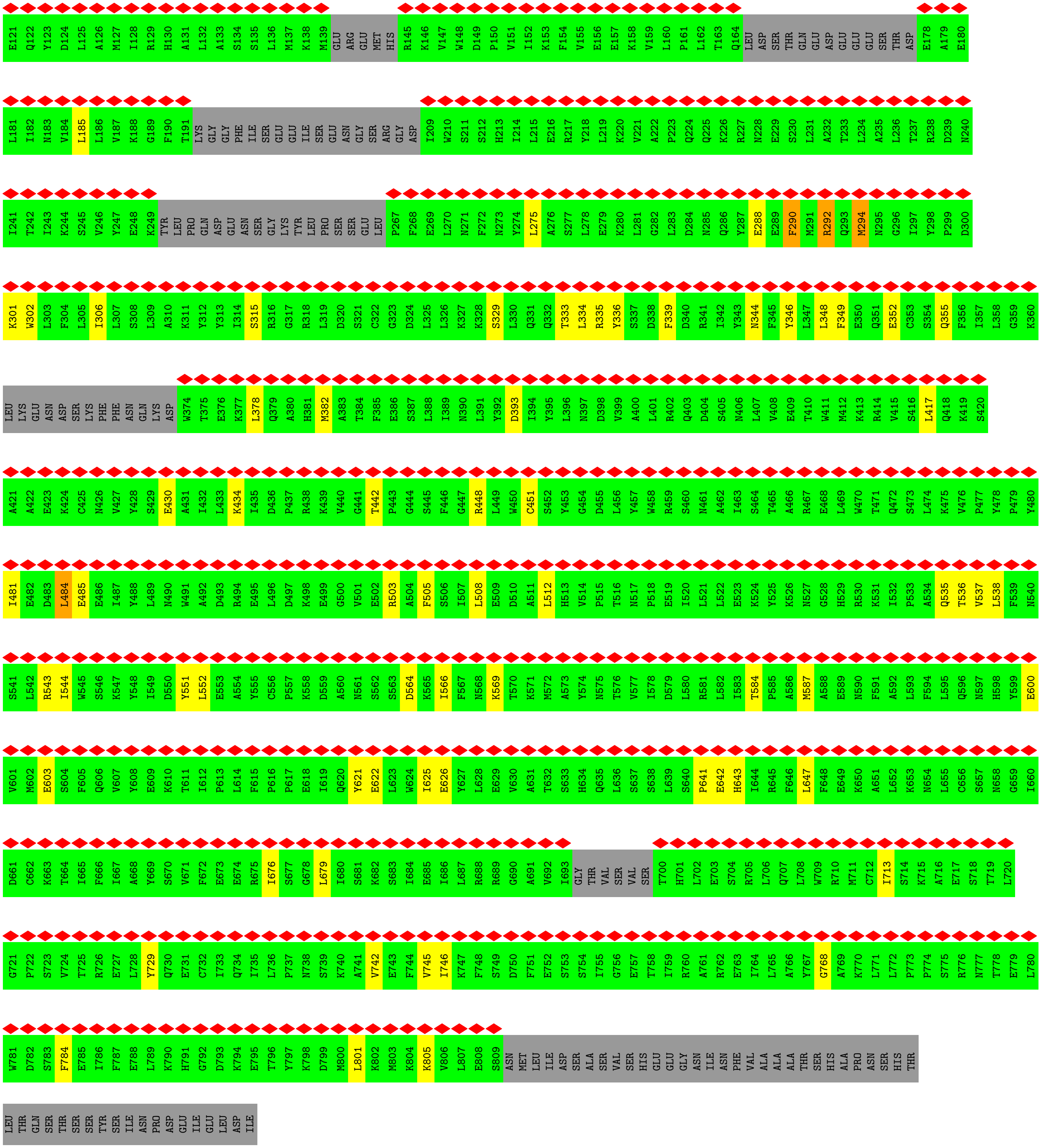
Chain S:



- Molecule 20: Pre-mRNA-splicing factor SYF1

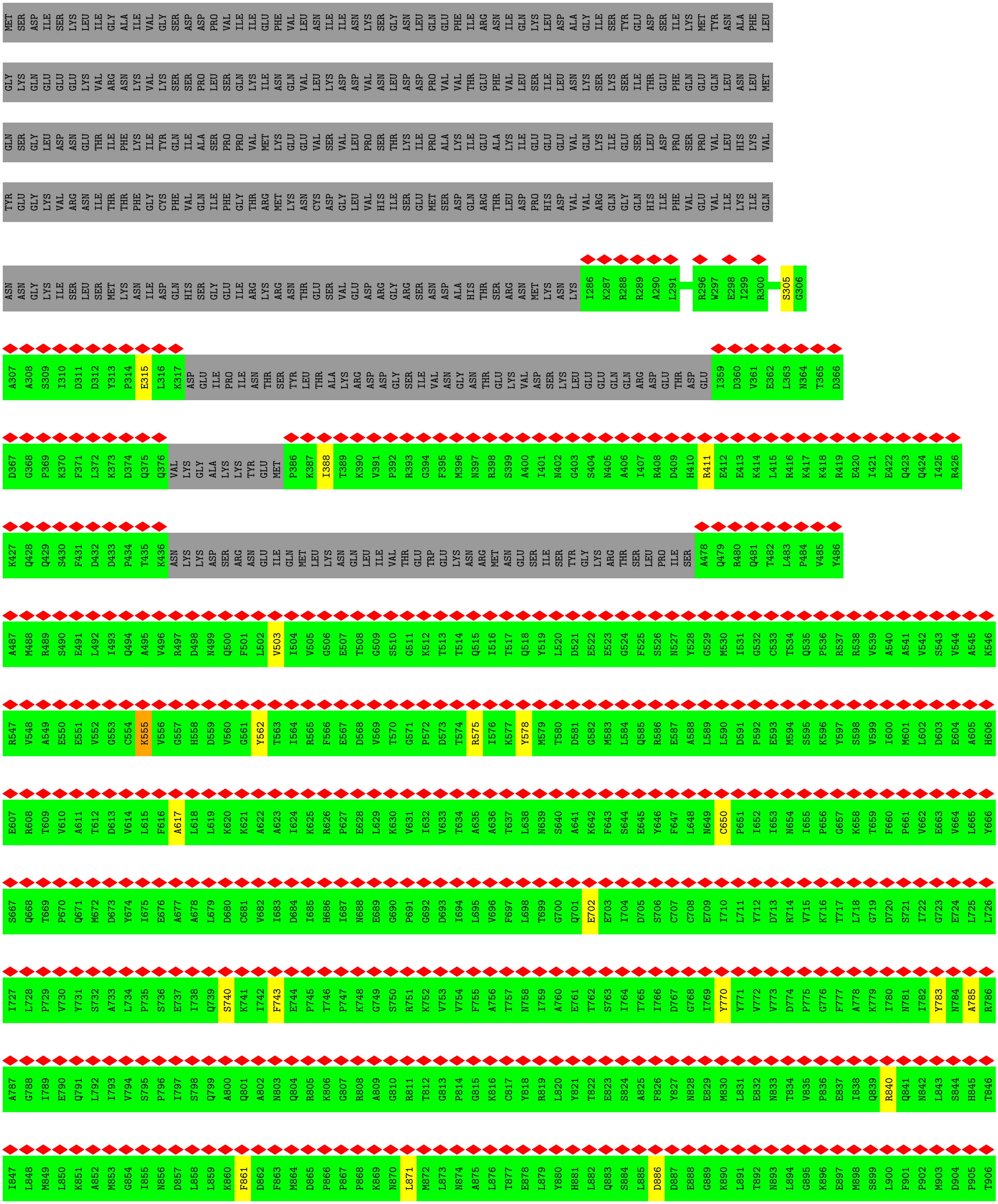
Chain T:





- Molecule 21: Pre-mRNA-splicing factor ATP-dependent RNA helicase PRP22

Chain V:



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

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

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

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

ASN ASN GLY LYS ILE SER LEU SER MET LYS ASN ILE ASP GLN HIS SER GLY GLU ILE ARG LYS ARG ASN THR GLU SER VAL GLU ASP ARG GLY ARG SER ASN ASP ALA HIS THR SER ARG ASN MET LYS ASN LYS I286 K287 R288 R289 A290 L291 R296 W297 E298 I299 R300 S305 G306

A307 A308 S309 I310 D311 D312 Y313 P314 E315 L316 K317 ASP GLU ILE PRO ILE ASN THR SER TYR LEU THR ALA LYS ARG ASP ASP GLY SER ILE VAL ASN GLY ASN THR GLU LYS VAL ASP SER LYS LEU GLU GLU GLN GLN ARG ASP GLU THR ASP GLU I359 D360 V361 E362 L363 N364 T365 D366

D367 G368 P369 K370 F371 L372 K373 D374 Q375 Q376 VAL LYS GLY ALA LYS LYS TYR GLU MET P386 K387 I388 T389 K390 V391 P392 R393 G394 F395 M396 N397 R398 S399 A400 I401 N402 G403 S404 N405 A406 I407 R408 D409 H410 R411 E412 E413 K414 L415 R416 K417 K418 R419 E420 I421 E422 Q423 Q424 I425 R426

K427 Q428 Q429 S430 F431 D432 D433 P434 T435 K436 ASN LYS LYS ASP SER ARG ASN GLU ILE GLN MET LEU LYS ASN GLN LEU ILE VAL THR GLU TRP GLU LYS ASN ARG MET ASN GLU SER ILE SER TYR GLY LYS ARG THR SER LEU PRO ILE SER A478 Q479 R480 Q481 T482 L483 P484 V485 Y486

A487 Q488 R489 S490 E491 L492 I493 Q494 A495 V496 R497 D498 N499 Q500 F501 L502 V503 I504 V505 G506 E507 T508 G509 S510 G511 K512 T513 T514 Q515 I516 T517 Q518 Y519 L520 D521 E522 E523 G524 F525 S526 N527 Y528 G529 M530 I531 G532 C533 T534 Q535 P536 R537 R538 V539 A540 A541 V542 S543 V544 A545 K546

R547 V548 A549 E550 E551 V552 G553 C554 K555 V556 G557 H558 D559 V560 G561 Y562 T563 I564 R565 F566 E567 D568 V569 T570 G571 P572 D573 T574 R575 I576 K577 Y578 M579 T580 D581 G582 M583 L584 Q585 R586 E587 A588 L589 L590 D591 P592 E593 M594 S595 K596 Y597 S598 V599 I600 M601 L602 D603 E604 A605 H606

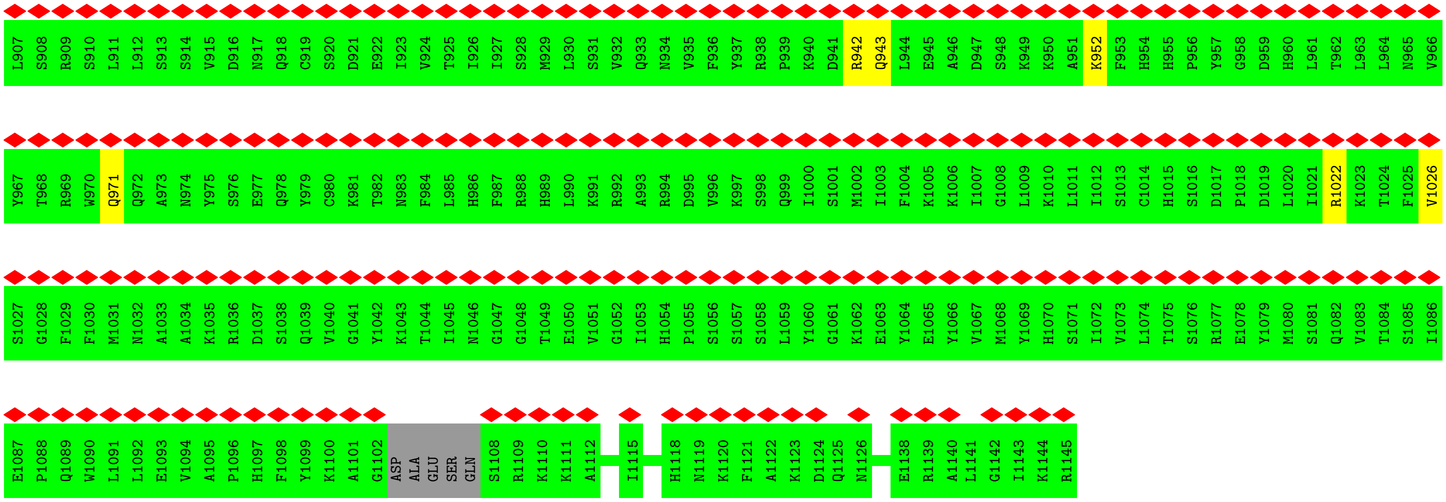
E607 R608 T609 V610 A611 T612 D613 V614 L615 F616 A617 L618 L619 K620 K621 A622 A623 I624 K625 R626 P627 E628 L629 K630 V631 I632 V633 T634 A635 A636 T637 L638 N639 S640 A641 K642 F643 S644 E645 Y646 F647 L648 N649 C650 P651 I652 I653 N654 I655 P656 G657 K658 T659 F660 P661 V662 E663 V664 L665 Y666

S667 Q668 T669 P670 Q671 M672 D673 Y674 I675 E676 A677 A678 L679 D680 C681 V682 I683 D684 I685 H686 I687 N688 E689 G690 P691 G692 D693 I694 L695 V696 F697 L698 T699 G700 Q701 E702 E703 I704 D705 S706 C707 C708 E709 I710 L711 Y712 D713 R714 V715 K716 T717 L718 G719 D720 S721 I722 G723 E724 L725 L726

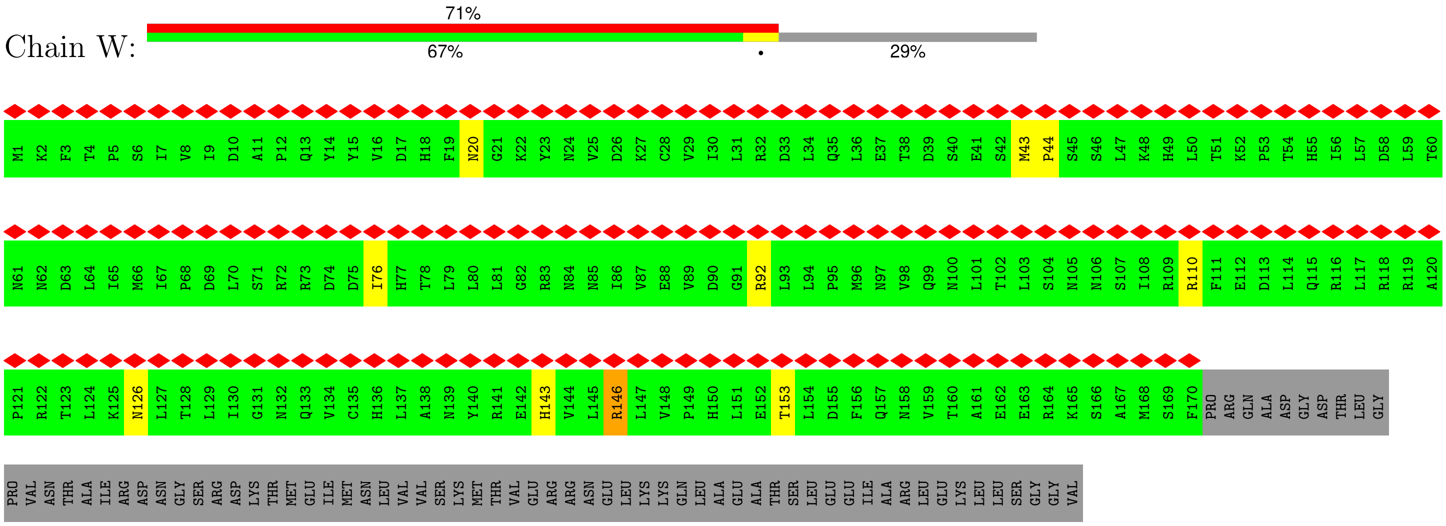
I727 L728 P729 V730 Y731 S732 A733 L734 P735 S736 E737 I738 Q739 S740 K741 I742 F743 E744 P745 T746 P747 K748 G749 S750 R751 K752 V753 V754 F755 A756 T757 N758 I759 A760 E761 T762 S763 I764 T765 I766 D767 G768 I769 Y770 Y771 V772 V773 D774 P775 G776 F777 A778 K779 I780 N781 I782 Y783 N784 A785 R786

A787 G788 I789 E790 Q791 L792 I793 V794 S795 P796 I797 S798 Q799 A800 Q801 A802 N803 Q804 R805 K806 G807 R808 A809 G810 R811 T812 G813 P814 G815 K816 C817 Y818 R819 L820 Y821 T822 E823 S824 A825 F826 Y827 N828 E829 M830 L831 E832 N833 T834 V835 P836 E837 I838 Q839 R840 Q841 N842 L843 S844 H845 T846

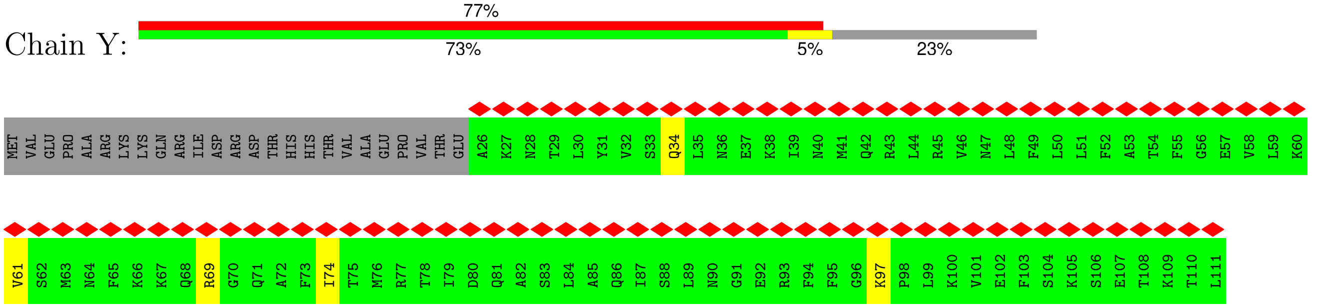
I847 L848 M849 L850 K851 A852 M853 G854 I855 N856 D857 L858 L859 K860 F861 D862 F863 M864 D865 P866 P867 P868 K869 N870 L871 M872 L873 N874 A875 L876 T877 E878 L879 Y880 H881 L882 Q883 S884 L885 D886 D887 E888 G889 K890 L891 T892 N893 L894 G895 K896 E897 M898 S899 L900 F901 P902 M903 D904 P905 T906



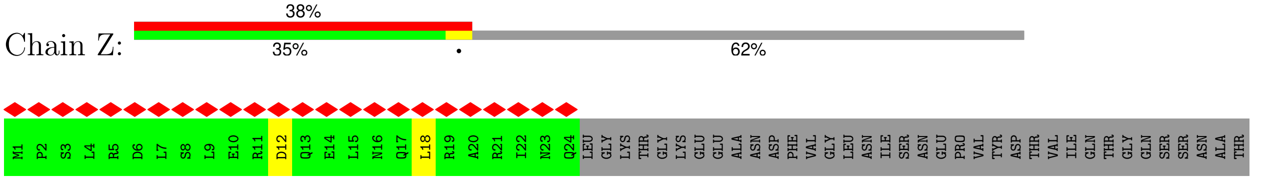
- Molecule 22: U2 small nuclear ribonucleoprotein A'

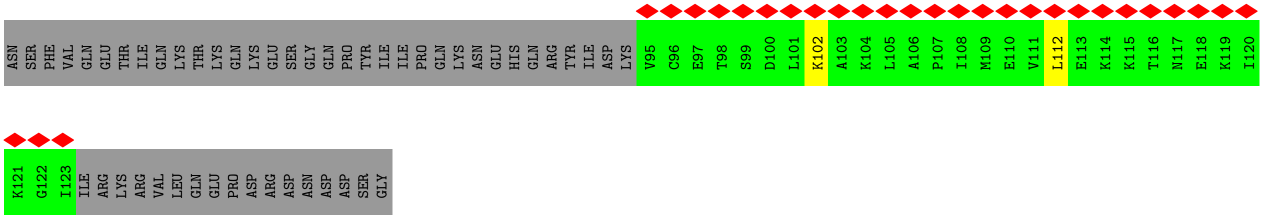


- Molecule 23: U2 small nuclear ribonucleoprotein B"



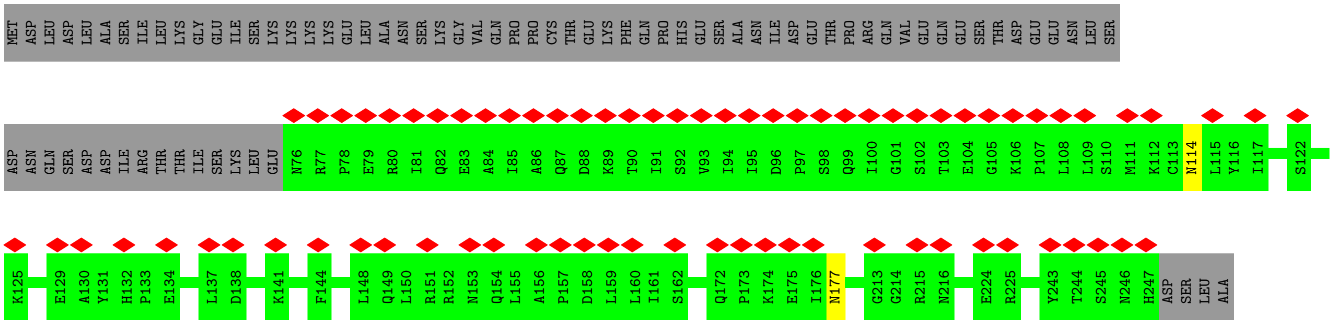
- Molecule 24: Pre-mRNA-splicing factor NTC20





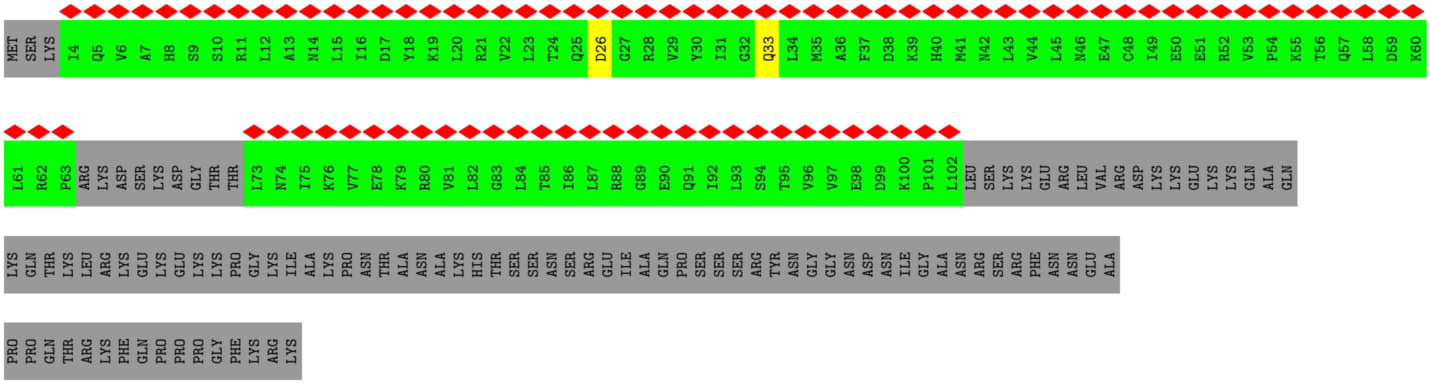
- Molecule 25: Pre-mRNA-splicing factor 18

Chain a:



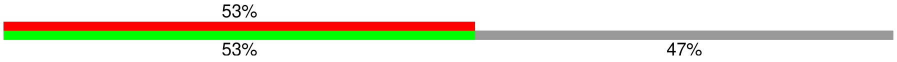
- Molecule 26: Small nuclear ribonucleoprotein-associated protein B

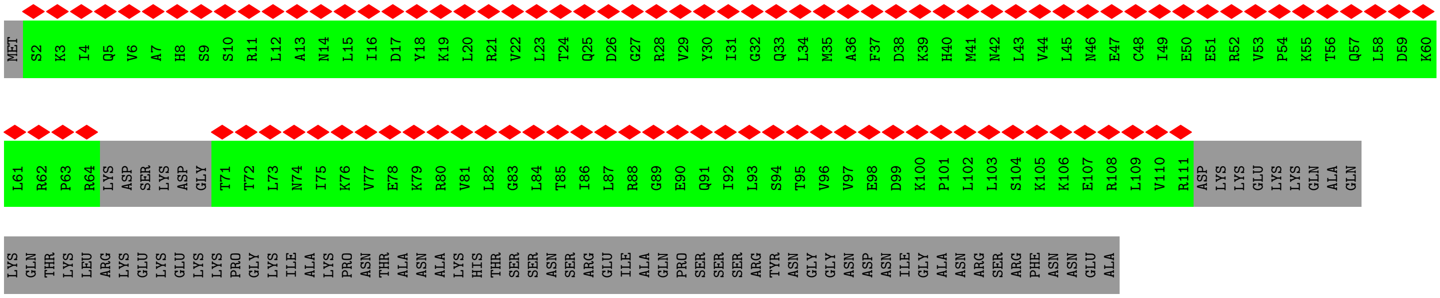
Chain b:



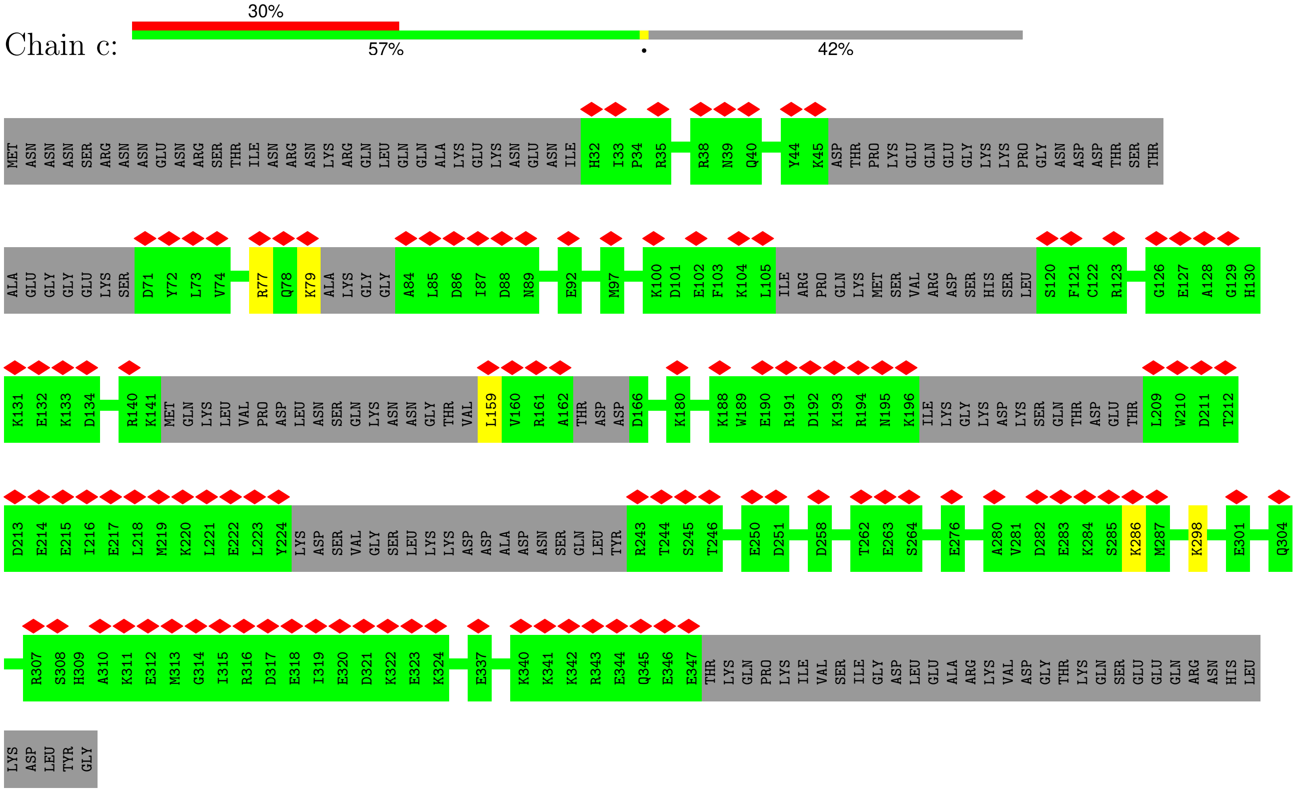
- Molecule 26: Small nuclear ribonucleoprotein-associated protein B

Chain k:

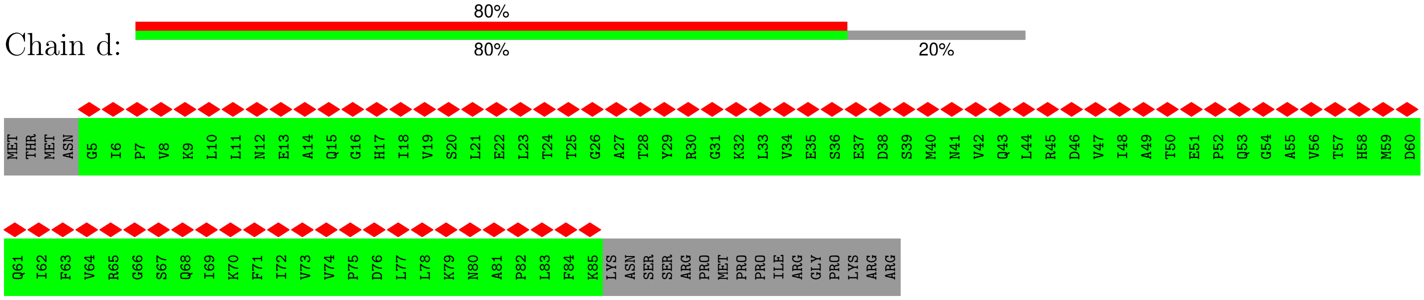




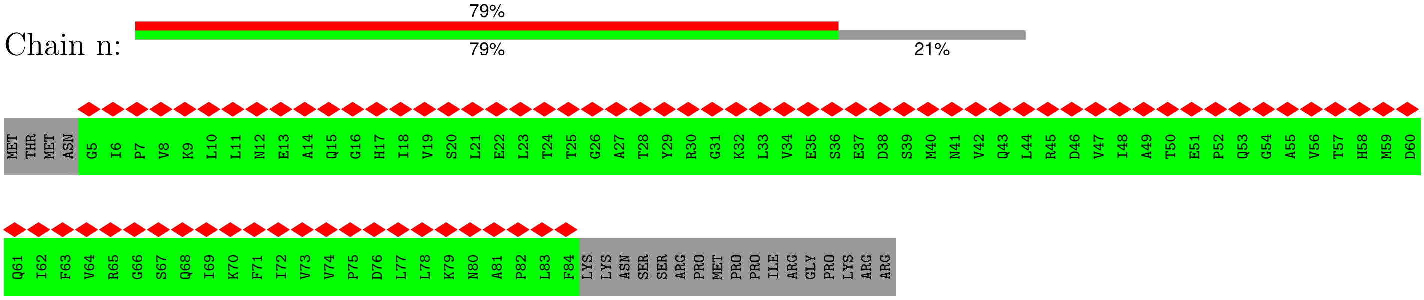
- Molecule 27: Pre-mRNA-splicing factor SLU7



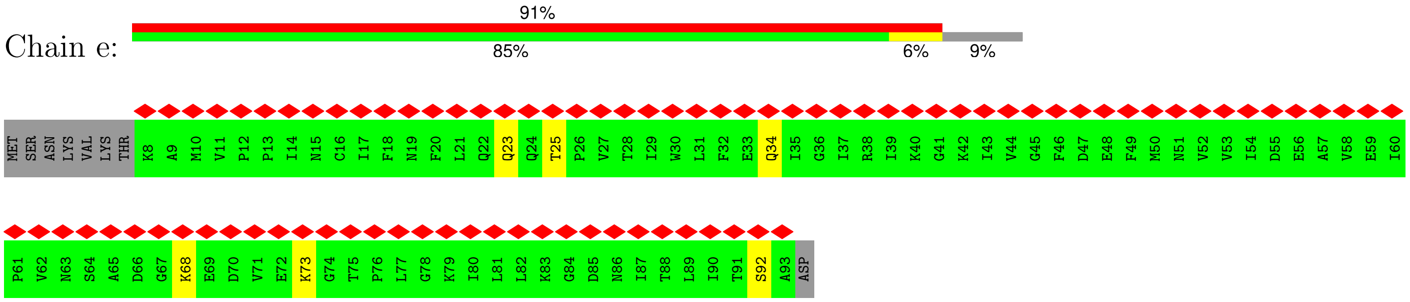
- Molecule 28: Small nuclear ribonucleoprotein Sm D3



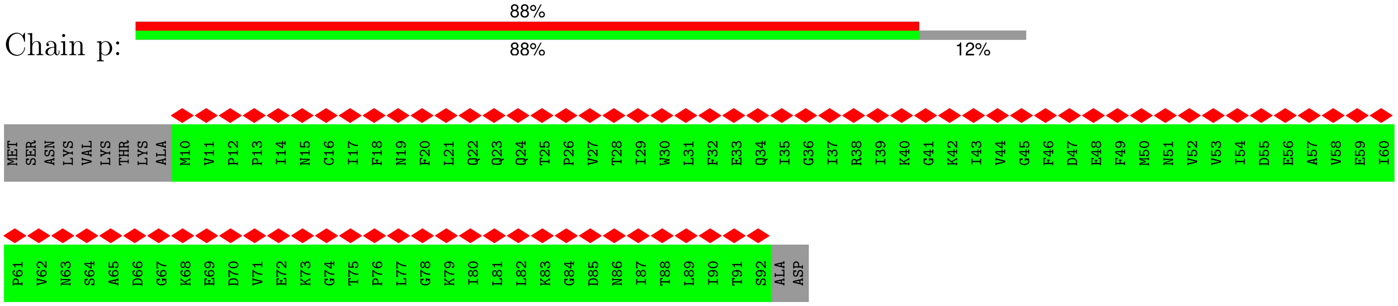
- Molecule 28: Small nuclear ribonucleoprotein Sm D3



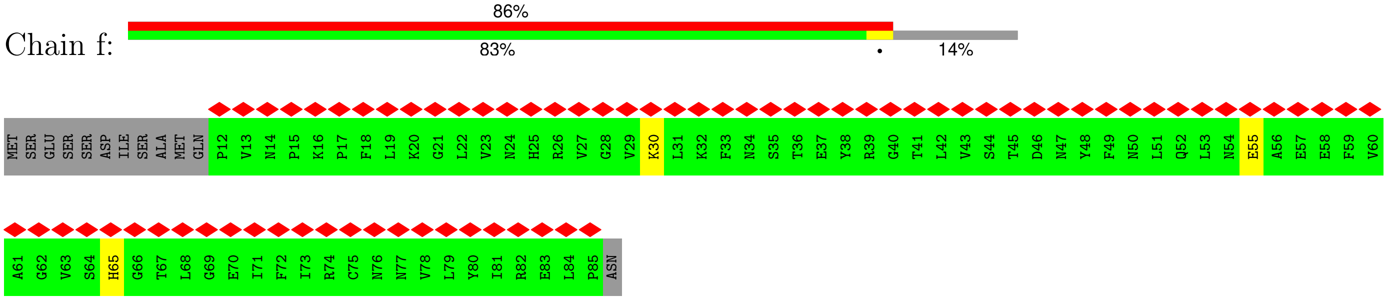
- Molecule 29: Small nuclear ribonucleoprotein E



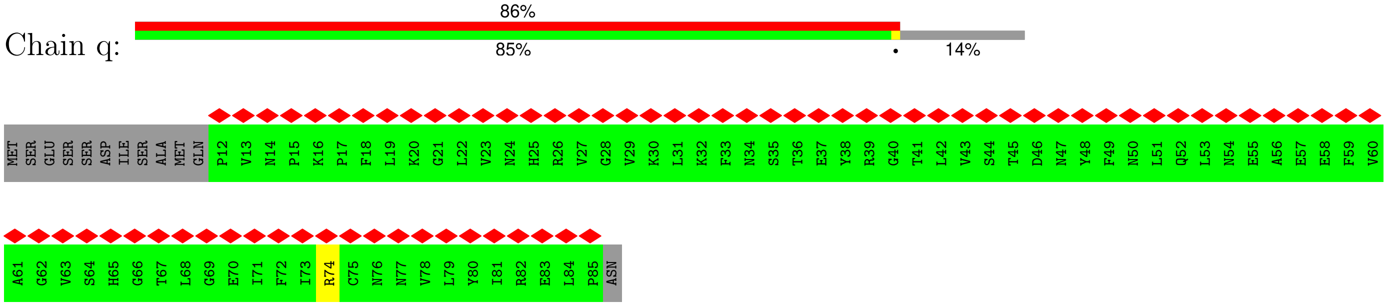
- Molecule 29: Small nuclear ribonucleoprotein E



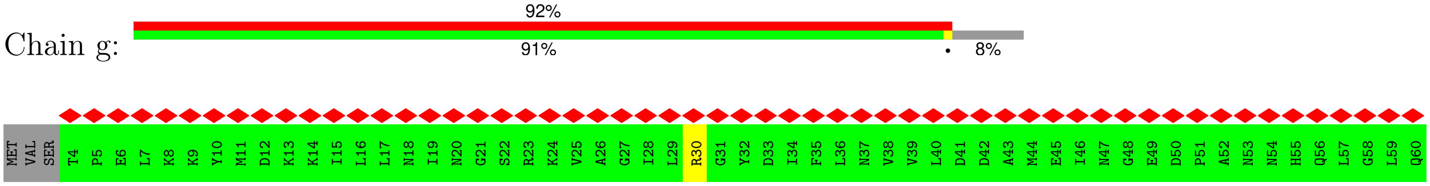
- Molecule 30: Small nuclear ribonucleoprotein F

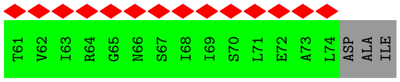


- Molecule 30: Small nuclear ribonucleoprotein F



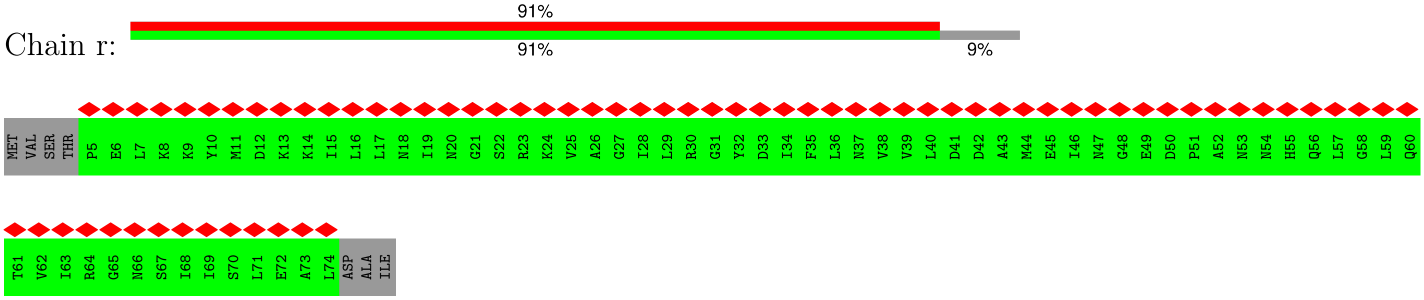
- Molecule 31: Small nuclear ribonucleoprotein G





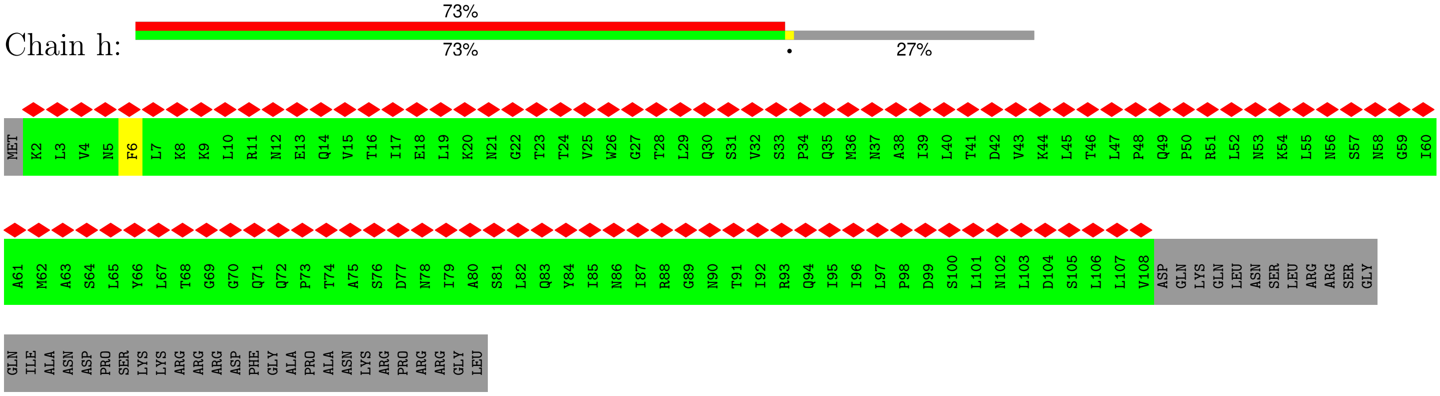
- Molecule 31: Small nuclear ribonucleoprotein G

Chain r:



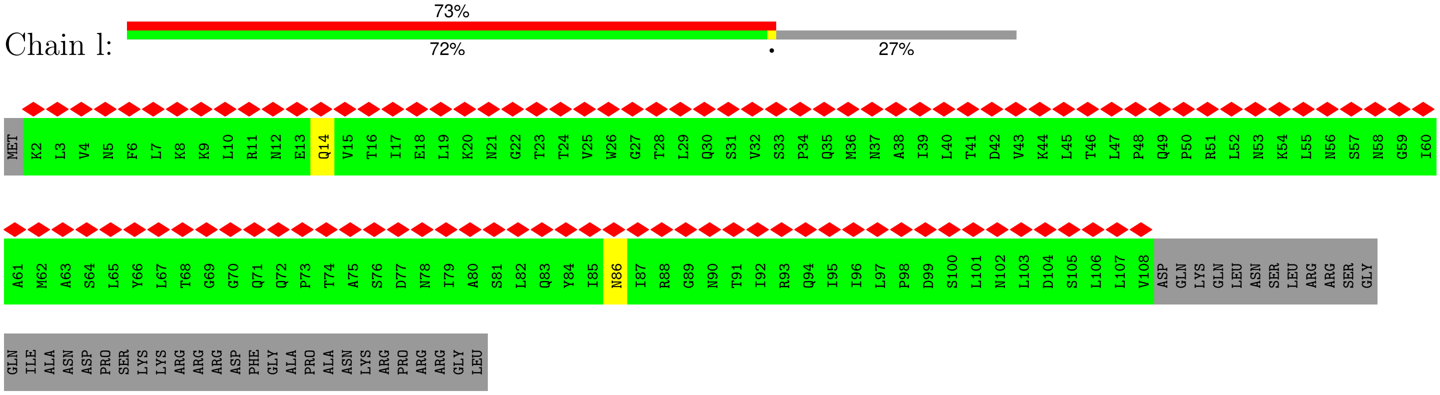
- Molecule 32: Small nuclear ribonucleoprotein Sm D1

Chain h:



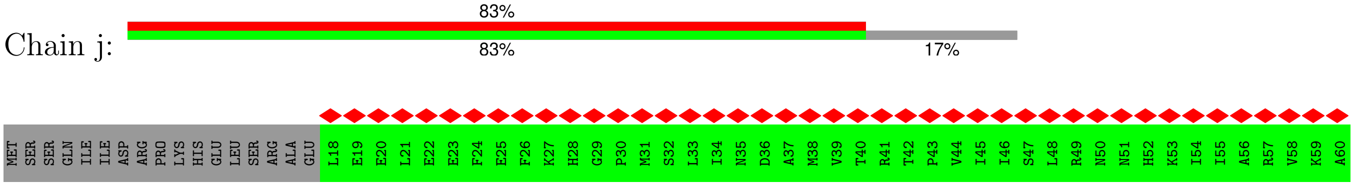
- Molecule 32: Small nuclear ribonucleoprotein Sm D1

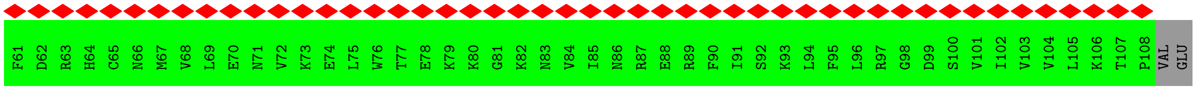
Chain l:



- Molecule 33: Small nuclear ribonucleoprotein Sm D2

Chain j:



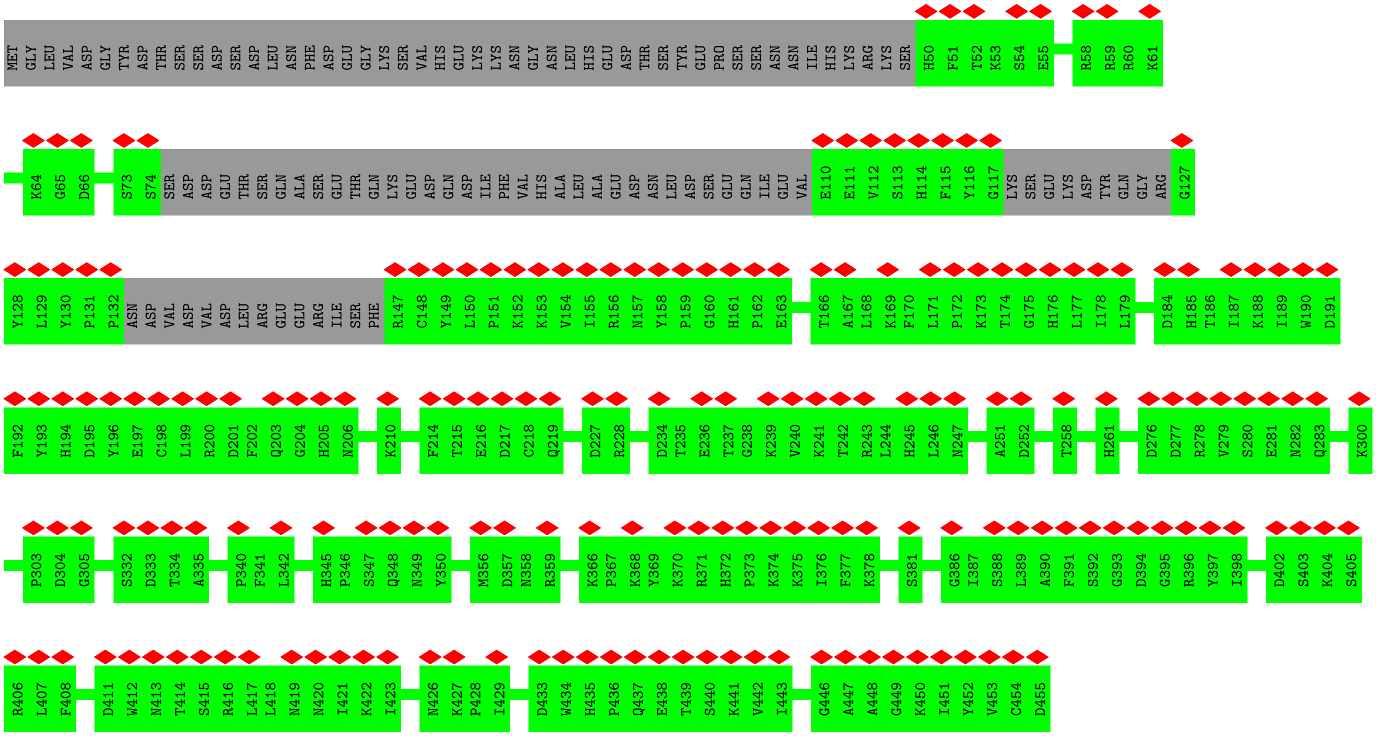


- Molecule 33: Small nuclear ribonucleoprotein Sm D2

Chain m:

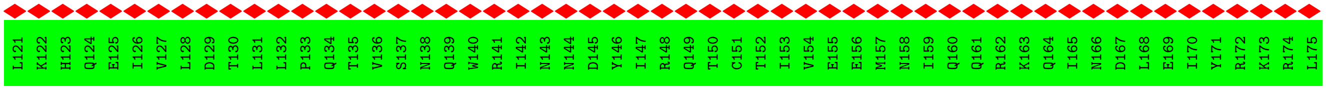
- Molecule 34: Pre-mRNA-processing factor 17

Chain o:

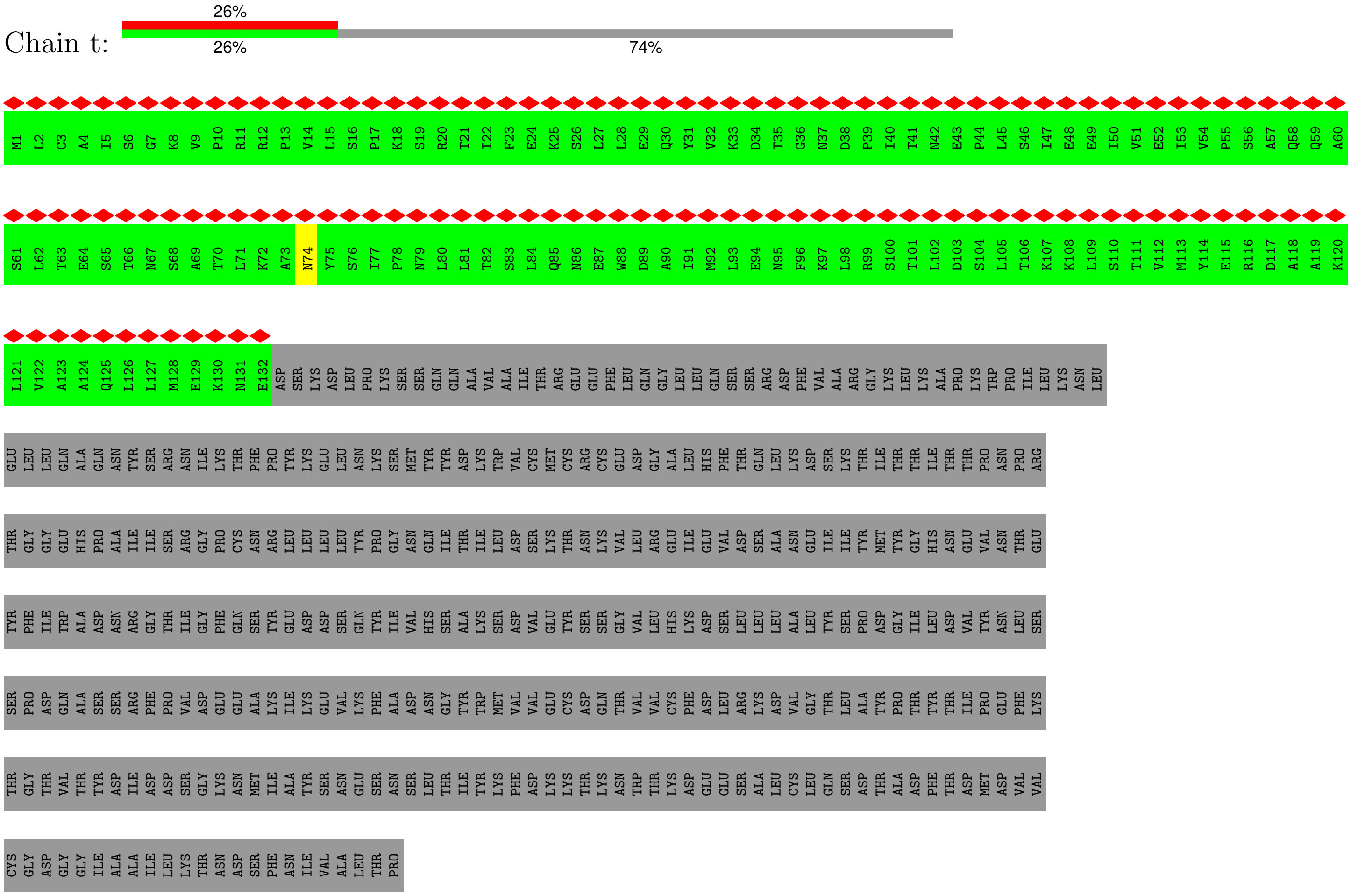


- Molecule 35: Pre-mRNA-splicing factor SNT309

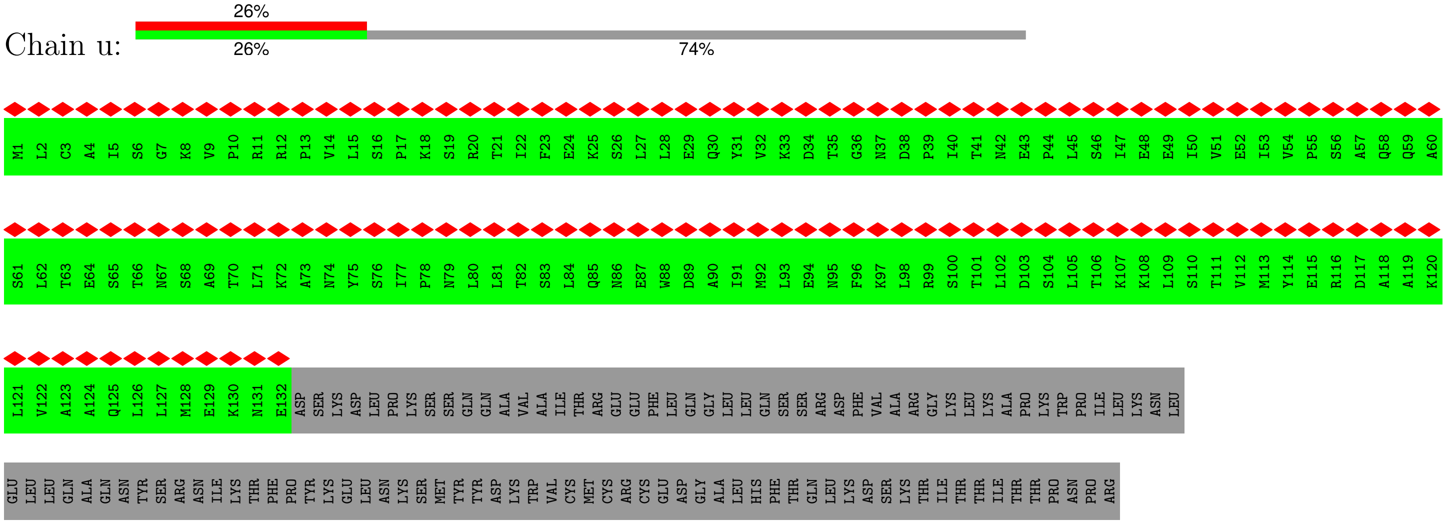
Chain s:



- Molecule 36: Pre-mRNA-processing factor 19



- Molecule 36: Pre-mRNA-processing factor 19



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

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

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

THR GLY THR VAL THR TYR ASP ILE ASP ASP SER GLY LYS ASN MET ILE ALA TYR SER ASN GLU SER ASN SER LEU THR ILE TYR LYS PHE ASP LYS LYS THR LYS ASN TRP THR LYS ASP GLU GLU SER ALA LEU CYS LEU GLN SER ASP THR ALA ASP PHE THR ASP MET ASP VAL VAL

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

• Molecule 36: Pre-mRNA-processing factor 19

Chain v:

M1 L2 C3 A4 I5 S6 G7 K8 V9 P10 R11 R12 P13 V14 L15 S16 P17 K18 S19 R20 T21 I22 F23 E24 K25 S26 L27 L28 E29 Q30 Y31 V32 K33 D34 T35 G36 N37 D38 P39 I40 T41 N42 E43 P44 L45 S46 I47 E48 E49 I50 V51 E52 I53 V54 P55 S56 A57 Q58 Q59 A60

S61 L62 T63 E64 S65 T66 N67 S68 A69 T70 L71 K72 A73 N74 Y75 S76 I77 P78 N79 L80 L81 T82 S83 L84 Q85 N86 E87 W88 D89 A90 I91 M92 L93 E94 N95 F96 K97 L98 R99 S100 T101 L102 D103 S104 L105 T106 K107 K108 L109 S110 T111 V112 M113 Y114 E115 R116 D117 A118 A119 K120

L121 V122 A123 A124 Q125 L126 L127 M128 E129 K130 N131 E132 ASP SER LYS ASP LEU PRO LYS SER SER GLN GLN ALA VAL ALA ILE THR ARG GLU GLU PHE LEU GLN GLY LEU LEU GLN SER SER ARG ASP PHE VAL ALA ARG GLY LYS LEU LYS ALA PRO LYS TRP PRO ILE LEU LYS ASN LEU

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

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

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

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

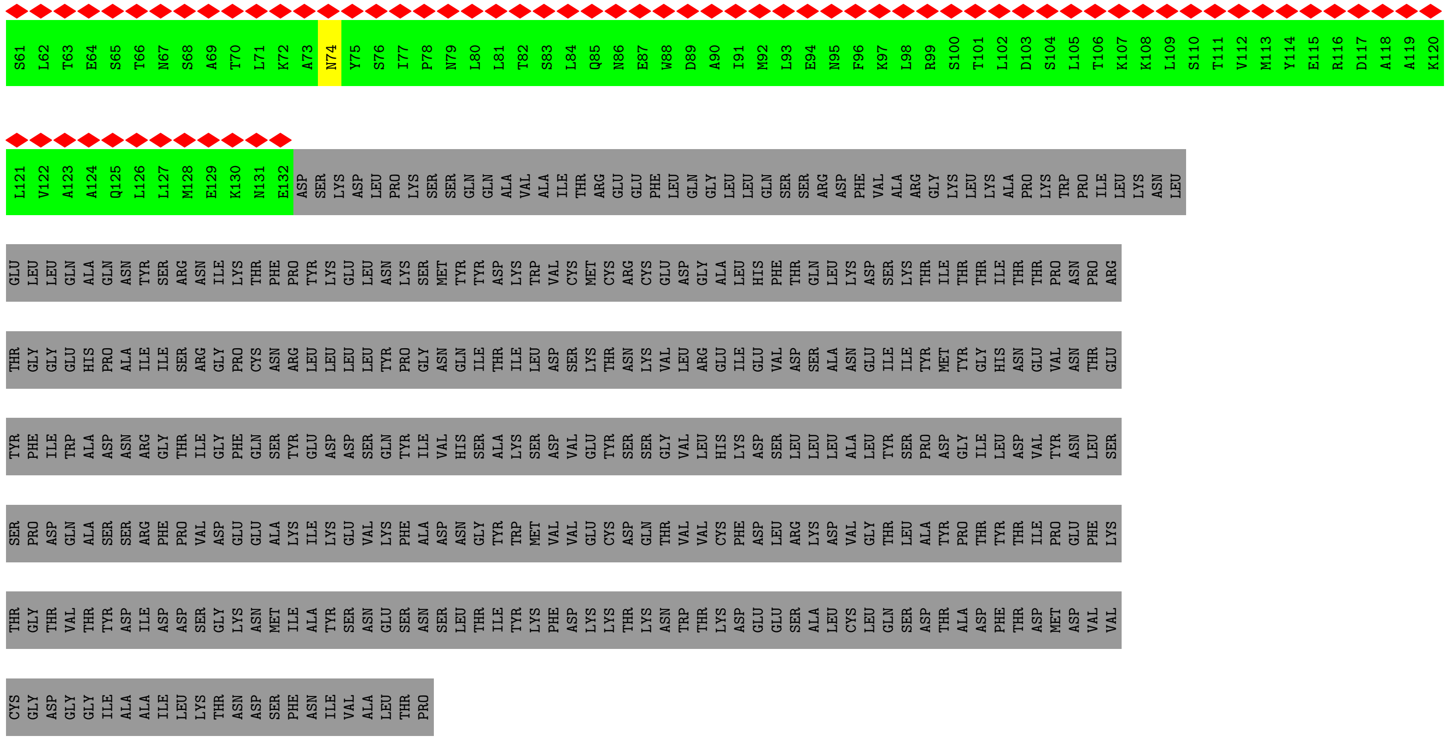
THR GLY THR VAL THR TYR ASP ILE ASP ASP SER GLY LYS ASN MET ILE ALA TYR SER ASN GLU SER ASN SER LEU THR ILE TYR LYS PHE ASP LYS LYS THR LYS ASN TRP THR LYS ASP GLU GLU SER ALA LEU CYS LEU GLN SER ASP THR ALA ASP PHE THR ASP MET ASP VAL VAL

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

• Molecule 36: Pre-mRNA-processing factor 19

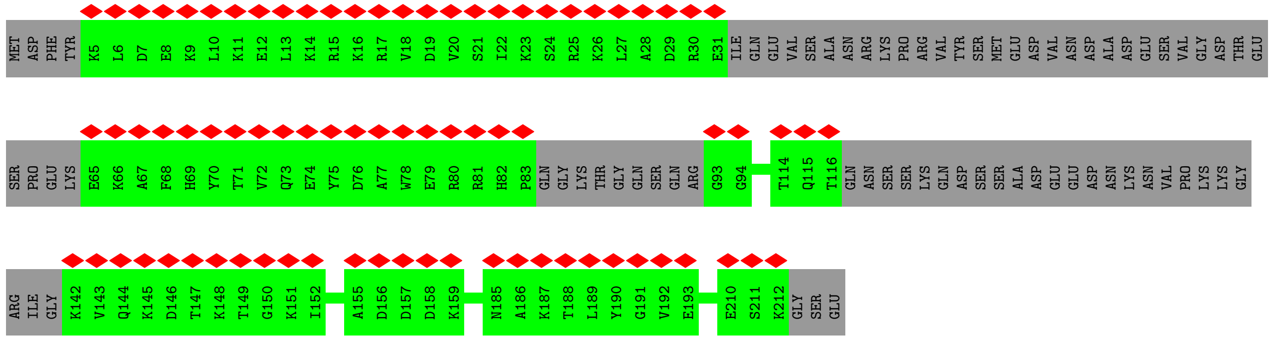
Chain w:

M1 L2 C3 A4 I5 S6 G7 K8 V9 P10 R11 R12 P13 V14 L15 S16 P17 K18 S19 R20 T21 I22 F23 E24 K25 S26 L27 L28 E29 Q30 Y31 V32 K33 D34 T35 G36 N37 D38 P39 I40 T41 N42 E43 P44 L45 S46 I47 E48 E49 I50 V51 E52 I53 V54 P55 S56 A57 Q58 Q59 A60



- Molecule 37: Pre-mRNA-splicing factor SYF2

Chain y:



4 Experimental information

Property	**Value**	**Source**
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	209005	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{Å}^2)$	40	Depositor
Minimum defocus (nm)	1300	Depositor
Maximum defocus (nm)	3100	Depositor
Magnification	130000	Depositor
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor
Maximum map value	0.138	Depositor
Minimum map value	-0.054	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.004	Depositor
Recommended contour level	0.02	Depositor
Map size (Å)	464.63998, 464.63998, 464.63998	wwPDB
Map dimensions	480, 480, 480	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.968, 0.968, 0.968	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: ZN, GTP, IHP, PSU, MG, K

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	2	0.59	0/3090	1.14	10/4792 (0.2%)
2	5	0.73	0/4098	1.05	6/6373 (0.1%)
3	6	0.82	0/2405	0.98	0/3744
4	A	0.38	0/16489	0.59	0/22348
5	C	0.47	0/7189	0.62	0/9728
6	D	0.35	0/922	0.63	0/1220
7	E	0.68	0/785	1.39	17/1218 (1.4%)
8	G	0.24	0/137	0.52	0/188
9	H	0.46	0/3824	0.58	0/5154
10	I	0.66	0/1128	1.13	7/1747 (0.4%)
11	J	0.37	0/3052	0.64	0/4143
12	K	0.34	0/1375	0.60	0/1854
13	L	0.37	0/1307	0.60	0/1748
14	M	0.34	0/2094	0.61	0/2815
15	N	0.30	0/2124	0.57	0/2860
16	O	0.33	0/2945	0.54	0/3992
17	P	0.35	0/623	0.66	0/832
18	R	0.31	0/510	0.50	0/698
19	S	0.36	0/4780	0.54	0/6483
20	T	0.36	0/5658	0.52	0/7669
21	V	0.43	0/6210	0.61	0/8377
22	W	0.31	0/1406	0.62	0/1905
23	Y	0.31	0/706	0.62	0/941
24	Z	0.31	0/424	0.58	0/564
25	a	0.31	0/1409	0.57	0/1911
26	b	0.36	0/731	0.63	0/983
26	k	0.33	0/849	0.63	0/1138
27	c	0.37	0/1925	0.63	0/2565
28	d	0.40	0/633	0.61	0/857
28	n	0.33	0/624	0.57	0/846
29	e	0.36	0/676	0.59	0/916
29	p	0.32	0/657	0.60	0/891

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
30	f	0.37	0/605	0.60	0/819
30	q	0.35	0/605	0.60	0/819
31	g	0.35	0/556	0.60	0/749
31	r	0.32	0/549	0.62	0/738
32	h	0.33	0/835	0.59	0/1135
32	l	0.31	0/835	0.59	0/1135
33	j	0.36	0/759	0.62	0/1019
33	m	0.31	0/759	0.60	0/1019
34	o	0.31	0/2903	0.62	0/3921
35	s	0.25	0/869	0.40	0/1210
36	t	0.25	0/657	0.43	0/917
36	u	0.25	0/657	0.41	0/917
36	v	0.25	0/657	0.40	0/917
36	w	0.25	0/657	0.40	0/917
37	y	0.33	0/1085	0.52	0/1453
All	All	0.43	0/93773	0.68	40/129185 (0.0%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
11	J	0	1
12	K	0	1
19	S	0	1
21	V	0	1
22	W	0	1
27	c	0	1
All	All	0	6

There are no bond length outliers.

The worst 5 of 40 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	2	1082	A	N1-C2-N3	9.71	134.15	129.30
7	E	6	C	OP1-P-OP2	-9.25	105.73	119.60
1	2	1082	A	C2-N3-C4	-8.11	106.54	110.60
7	E	3	A	OP1-P-O3’	8.05	122.92	105.20
1	2	155	U	N1-C2-O2	-7.95	117.23	122.80

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
11	J	309	ARG	Sidechain
12	K	185	ARG	Sidechain
19	S	186	ARG	Sidechain
21	V	840	ARG	Sidechain
22	W	92	ARG	Sidechain

5.2 Too-close contacts ⓘ

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	2	2779	0	1411	24	0
2	5	3671	0	1856	6	0
3	6	2170	0	1095	14	0
4	A	16079	0	16079	43	0
5	C	7040	0	7241	21	0
6	D	914	0	953	19	0
7	E	703	0	356	2	0
8	G	139	0	67	0	0
9	H	3762	0	3825	38	0
10	I	1012	0	510	5	0
11	J	2990	0	2906	12	0
12	K	1355	0	1383	7	0
13	L	1283	0	1301	0	0
14	M	2048	0	2011	9	0
15	N	2092	0	2162	9	0
16	O	2918	0	2467	15	0
17	P	607	0	596	2	0
18	R	510	0	327	6	0
19	S	4683	0	4267	39	0
20	T	5543	0	5182	41	0
21	V	6096	0	6158	10	0
22	W	1383	0	1407	2	0
23	Y	697	0	733	2	0
24	Z	424	0	455	2	0
25	a	1380	0	1425	0	0
26	b	725	0	778	0	0
26	k	843	0	917	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
27	c	1893	0	1849	0	0
28	d	624	0	647	0	0
28	n	615	0	634	0	0
29	e	665	0	690	0	0
29	p	646	0	667	0	0
30	f	592	0	594	0	0
30	q	592	0	594	0	0
31	g	551	0	572	0	0
31	r	544	0	566	0	0
32	h	826	0	874	0	0
32	l	826	0	874	0	0
33	j	747	0	784	0	0
33	m	747	0	784	0	0
34	o	2821	0	2743	0	0
35	s	870	0	363	0	0
36	t	658	0	291	0	0
36	u	658	0	291	0	0
36	v	658	0	291	0	0
36	w	658	0	291	0	0
37	y	1075	0	1004	0	0
38	6	1	0	0	0	0
38	C	1	0	0	0	0
39	6	3	0	0	0	0
39	E	1	0	0	0	0
40	A	36	0	6	1	0
40	S	36	0	6	1	0
41	C	32	0	12	0	0
42	L	3	0	0	0	0
42	M	1	0	0	0	0
42	N	2	0	0	0	0
42	c	1	0	0	0	0
43	6	1	0	0	0	0
All	All	91230	0	83295	296	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 2.

The worst 5 of 296 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)
6:D:84:LYS:HA	6:D:87:LEU:HD12	1.50	0.92

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
14:M:161:ARG:HH21	15:N:299:ALA:HB2	1.40	0.86
1:2:1141:C:H2'	1:2:1142:G:C8	2.20	0.76
11:J:143:ARG:HH11	11:J:160:ASN:HD21	1.35	0.74
5:C:763:ARG:HH22	18:R:82:ASP:CB	2.01	0.73

There are no symmetry-related clashes.

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	
4	A	1942/2413 (80%)	1888 (97%)	54 (3%)	0	100	100
5	C	865/1008 (86%)	833 (96%)	32 (4%)	0	100	100
6	D	106/173 (61%)	106 (100%)	0	0	100	100
8	G	24/235 (10%)	24 (100%)	0	0	100	100
9	H	456/577 (79%)	449 (98%)	7 (2%)	0	100	100
11	J	378/451 (84%)	360 (95%)	18 (5%)	0	100	100
12	K	162/379 (43%)	157 (97%)	5 (3%)	0	100	100
13	L	154/157 (98%)	152 (99%)	2 (1%)	0	100	100
14	M	253/339 (75%)	248 (98%)	5 (2%)	0	100	100
15	N	256/364 (70%)	247 (96%)	9 (4%)	0	100	100
16	O	412/590 (70%)	407 (99%)	5 (1%)	0	100	100
17	P	68/175 (39%)	67 (98%)	1 (2%)	0	100	100
18	R	87/135 (64%)	82 (94%)	3 (3%)	2 (2%)	5	4
19	S	596/687 (87%)	586 (98%)	10 (2%)	0	100	100
20	T	689/859 (80%)	671 (97%)	18 (3%)	0	100	100
21	V	754/1145 (66%)	722 (96%)	31 (4%)	1 (0%)	48	60

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
22	W	168/238 (71%)	152 (90%)	15 (9%)	1 (1%)	22	26
23	Y	84/111 (76%)	81 (96%)	3 (4%)	0	100	100
24	Z	49/140 (35%)	48 (98%)	1 (2%)	0	100	100
25	a	170/251 (68%)	165 (97%)	5 (3%)	0	100	100
26	b	86/196 (44%)	85 (99%)	1 (1%)	0	100	100
26	k	100/196 (51%)	97 (97%)	3 (3%)	0	100	100
27	c	207/382 (54%)	196 (95%)	11 (5%)	0	100	100
28	d	79/101 (78%)	78 (99%)	1 (1%)	0	100	100
28	n	78/101 (77%)	77 (99%)	1 (1%)	0	100	100
29	e	84/94 (89%)	81 (96%)	3 (4%)	0	100	100
29	p	81/94 (86%)	80 (99%)	1 (1%)	0	100	100
30	f	72/86 (84%)	70 (97%)	2 (3%)	0	100	100
30	q	72/86 (84%)	70 (97%)	2 (3%)	0	100	100
31	g	69/77 (90%)	65 (94%)	4 (6%)	0	100	100
31	r	68/77 (88%)	66 (97%)	2 (3%)	0	100	100
32	h	105/146 (72%)	104 (99%)	1 (1%)	0	100	100
32	l	105/146 (72%)	102 (97%)	3 (3%)	0	100	100
33	j	89/110 (81%)	88 (99%)	1 (1%)	0	100	100
33	m	89/110 (81%)	88 (99%)	1 (1%)	0	100	100
34	o	340/455 (75%)	323 (95%)	17 (5%)	0	100	100
35	s	173/175 (99%)	166 (96%)	7 (4%)	0	100	100
36	t	130/503 (26%)	127 (98%)	2 (2%)	1 (1%)	16	20
36	u	130/503 (26%)	129 (99%)	1 (1%)	0	100	100
36	v	130/503 (26%)	129 (99%)	1 (1%)	0	100	100
36	w	130/503 (26%)	124 (95%)	5 (4%)	1 (1%)	16	20
37	y	133/215 (62%)	132 (99%)	1 (1%)	0	100	100
All	All	10223/15286 (67%)	9922 (97%)	295 (3%)	6 (0%)	50	60

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
18	R	59	ALA
21	V	943	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
18	R	57	PRO
36	t	74	ASN
36	w	74	ASN

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	A	1770/2182 (81%)	1765 (100%)	5 (0%)	91	95
5	C	795/910 (87%)	792 (100%)	3 (0%)	89	94
6	D	100/159 (63%)	97 (97%)	3 (3%)	36	51
9	H	429/538 (80%)	421 (98%)	8 (2%)	52	68
11	J	318/397 (80%)	315 (99%)	3 (1%)	75	86
12	K	152/328 (46%)	147 (97%)	5 (3%)	33	47
13	L	140/141 (99%)	140 (100%)	0	100	100
14	M	219/296 (74%)	218 (100%)	1 (0%)	86	93
15	N	243/332 (73%)	241 (99%)	2 (1%)	79	88
16	O	219/525 (42%)	214 (98%)	5 (2%)	45	62
17	P	61/151 (40%)	61 (100%)	0	100	100
18	R	21/121 (17%)	21 (100%)	0	100	100
19	S	433/633 (68%)	422 (98%)	11 (2%)	42	59
20	T	551/786 (70%)	541 (98%)	10 (2%)	54	70
21	V	675/1028 (66%)	666 (99%)	9 (1%)	65	78
22	W	161/219 (74%)	156 (97%)	5 (3%)	35	50
23	Y	77/100 (77%)	76 (99%)	1 (1%)	65	78
24	Z	49/128 (38%)	47 (96%)	2 (4%)	26	38
25	a	152/225 (68%)	150 (99%)	2 (1%)	65	78
26	b	83/176 (47%)	81 (98%)	2 (2%)	44	60
26	k	97/176 (55%)	97 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
27	c	204/346 (59%)	200 (98%)	4 (2%)	50	67
28	d	70/89 (79%)	70 (100%)	0	100	100
28	n	69/89 (78%)	69 (100%)	0	100	100
29	e	75/83 (90%)	69 (92%)	6 (8%)	10	12
29	p	74/83 (89%)	74 (100%)	0	100	100
30	f	66/77 (86%)	63 (96%)	3 (4%)	23	34
30	q	66/77 (86%)	65 (98%)	1 (2%)	60	75
31	g	61/66 (92%)	60 (98%)	1 (2%)	58	73
31	r	60/66 (91%)	60 (100%)	0	100	100
32	h	96/129 (74%)	95 (99%)	1 (1%)	73	84
32	l	96/129 (74%)	94 (98%)	2 (2%)	48	65
33	j	85/103 (82%)	85 (100%)	0	100	100
33	m	85/103 (82%)	83 (98%)	2 (2%)	44	60
34	o	314/413 (76%)	314 (100%)	0	100	100
37	y	102/193 (53%)	102 (100%)	0	100	100
All	All	8268/11597 (71%)	8171 (99%)	97 (1%)	66	80

5 of 97 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
21	V	388	ILE
24	Z	12	ASP
21	V	575	ARG
22	W	20	ASN
26	b	26	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 60 such sidechains are listed below:

Mol	Chain	Res	Type
11	J	306	HIS
32	l	86	ASN
16	O	82	ASN
32	l	78	ASN
37	y	200	ASN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	2	126/1175 (10%)	26 (20%)	5 (3%)
10	I	45/95 (47%)	13 (28%)	6 (13%)
2	5	169/214 (78%)	16 (9%)	4 (2%)
3	6	101/112 (90%)	9 (8%)	0
7	E	32/42 (76%)	14 (43%)	6 (18%)
All	All	473/1638 (28%)	78 (16%)	21 (4%)

5 of 78 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	2	19	U
1	2	21	G
1	2	25	A
1	2	31	A
1	2	32	G

5 of 21 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
7	E	16	A
10	I	70	A
10	I	91	A
10	I	71	C
10	I	61	U

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

1 non-standard protein/DNA/RNA residue is 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	$\#\|Z\| > 2$
3	PSU	6	28	3	18,21,22	1.06	1 (5%)	21,30,33	2.08	5 (23%)

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
3	PSU	6	28	3	-	1/7/25/26	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	6	28	PSU	C6-C5	3.01	1.38	1.35

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	6	28	PSU	N1-C2-N3	5.26	120.72	115.17
3	6	28	PSU	C4-N3-C2	-5.09	119.36	126.37
3	6	28	PSU	C6-C5-C4	3.07	120.25	118.17
3	6	28	PSU	O2-C2-N1	-2.79	119.91	122.79
3	6	28	PSU	C6-N1-C2	-2.56	120.32	122.69

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	6	28	PSU	O4'-C1'-C5-C4

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 16 ligands modelled in this entry, 13 are monoatomic - leaving 3 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	$\#|Z| > 2$
41	GTP	C	2501	38	29,34,34	1.28	2 (6%)	35,54,54	1.44	5 (14%)
40	IHP	S	701	-	36,36,36	0.69	0	60,60,60	0.85	2 (3%)
40	IHP	A	2500	-	36,36,36	0.82	0	60,60,60	0.60	0

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
41	GTP	C	2501	38	-	2/18/38/38	0/3/3/3
40	IHP	S	701	-	-	2/30/54/54	0/1/1/1
40	IHP	A	2500	-	-	1/30/54/54	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
41	C	2501	GTP	C5-C6	-4.53	1.38	1.47
41	C	2501	GTP	C5-C4	-2.17	1.37	1.43

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
41	C	2501	GTP	C8-N7-C5	3.84	109.09	102.55
41	C	2501	GTP	C5-C6-N1	3.10	119.99	114.07
41	C	2501	GTP	C2-N1-C6	-2.94	119.73	125.11
41	C	2501	GTP	O2B-PB-O3A	2.81	114.88	107.27
40	S	701	IHP	C6-C1-C2	2.78	116.53	110.43

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
40	S	701	IHP	C1-O11-P1-O21
40	A	2500	IHP	C3-O13-P3-O43
40	S	701	IHP	C4-O14-P4-O34

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
41	C	2501	GTP	PA-O3A-PB-O1B
41	C	2501	GTP	PB-O3B-PG-O3G

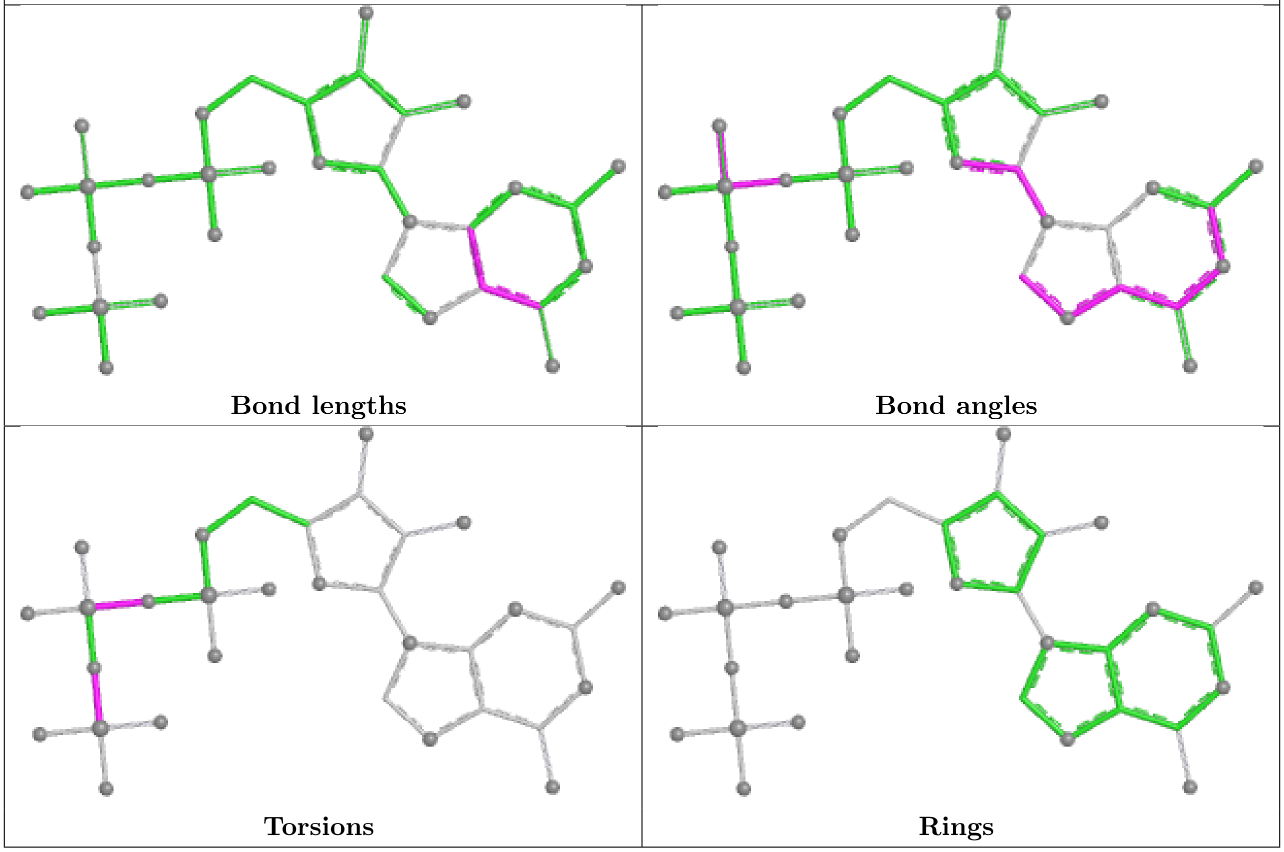
There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
40	S	701	IHP	1	0
40	A	2500	IHP	1	0

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 GTP C 2501



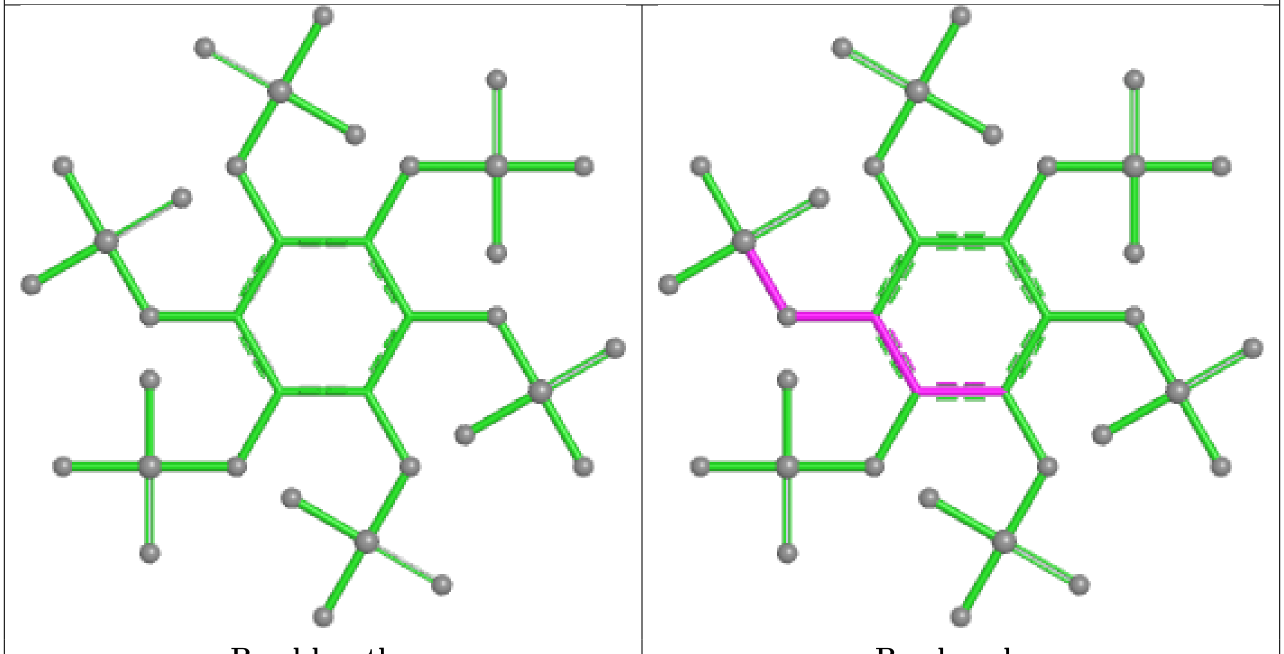
Bond lengths

Bond angles

Torsions

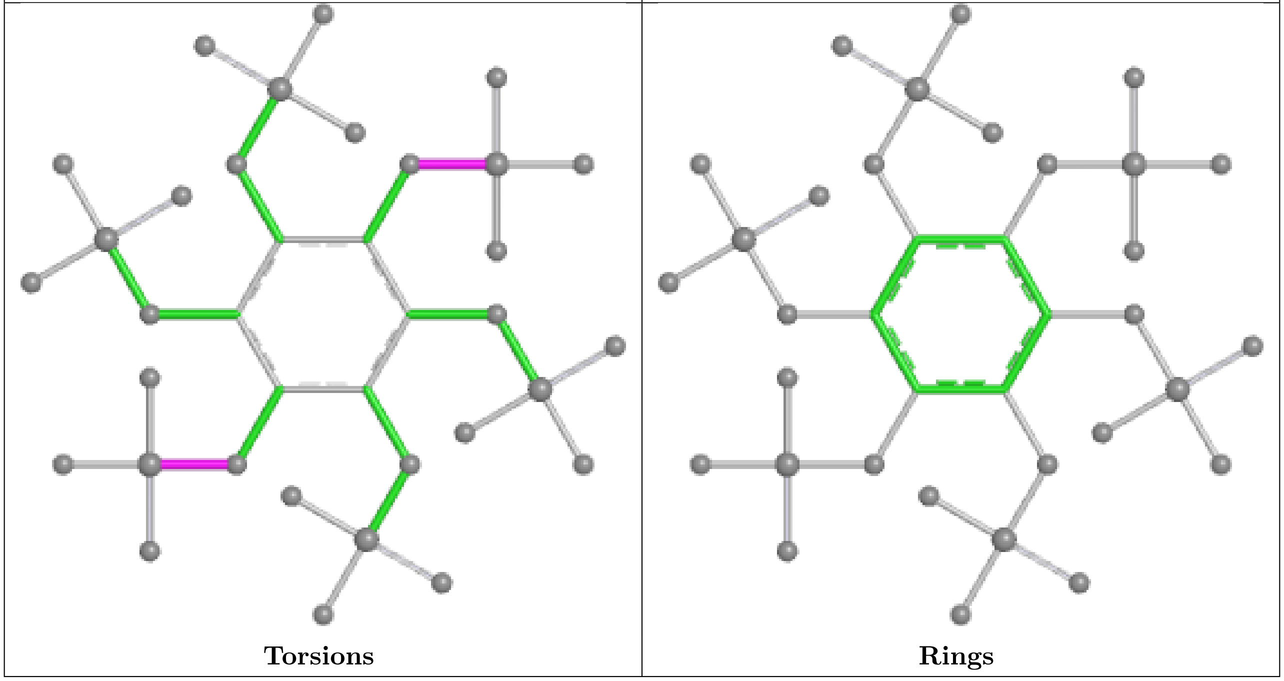
Rings

Ligand IHP S 701



Bond lengths

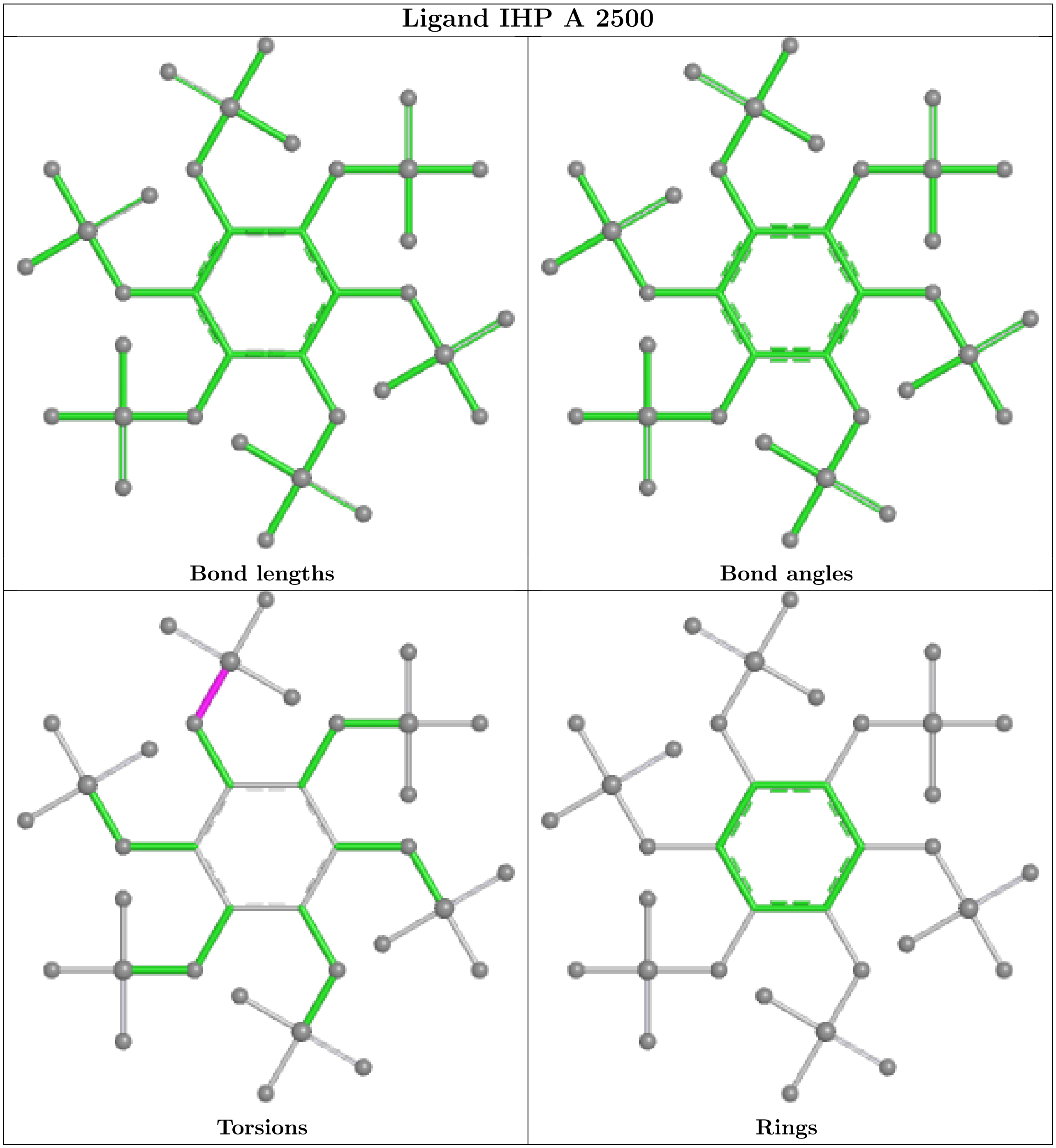
Bond angles



Torsions

Rings

Ligand IHP A 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.

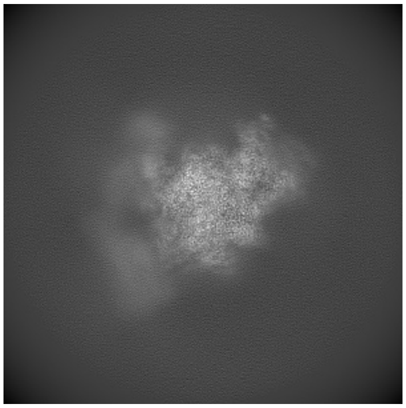
6 Map visualisation ⓘ

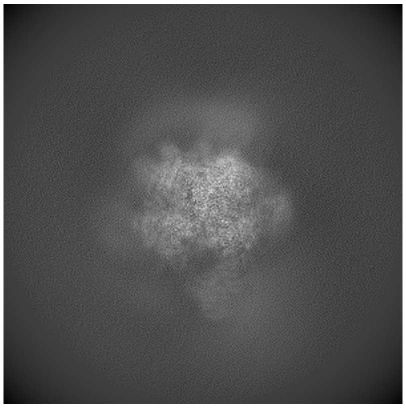
This section contains visualisations of the EMDB entry EMD-47157. These allow visual inspection of the internal detail of the map and identification of artifacts.

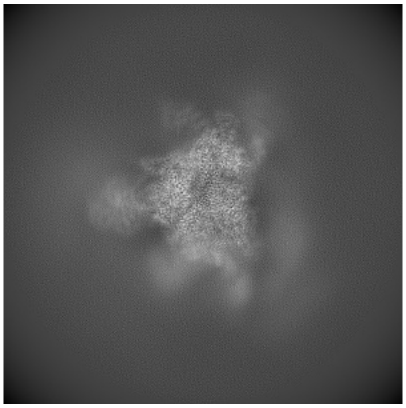
Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections ⓘ

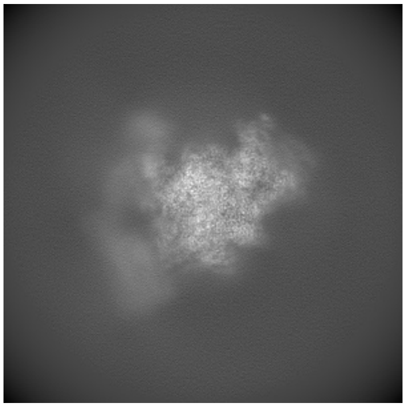
6.1.1 Primary map

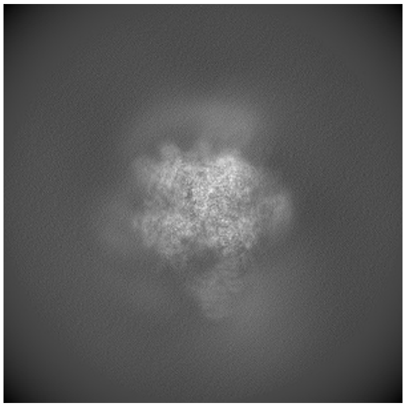

X

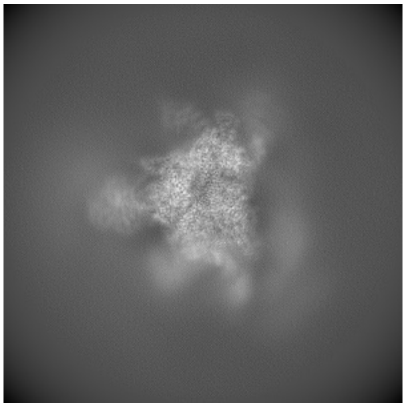

Y


Z

6.1.2 Raw map


X

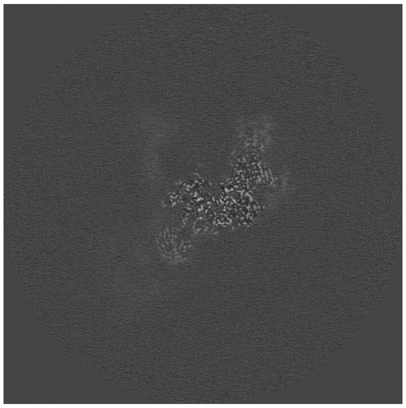

Y

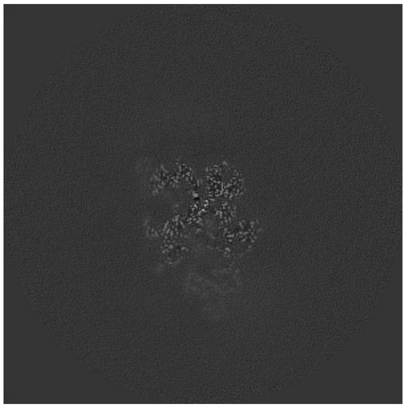

Z

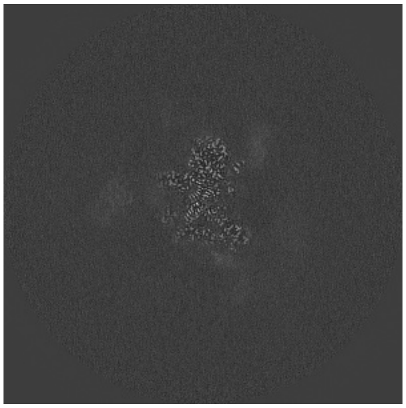
The images above show the map projected in three orthogonal directions.

6.2 Central slices

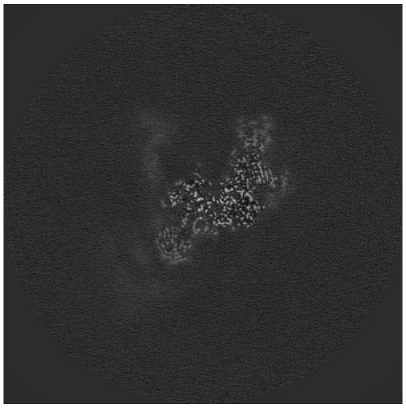
6.2.1 Primary map

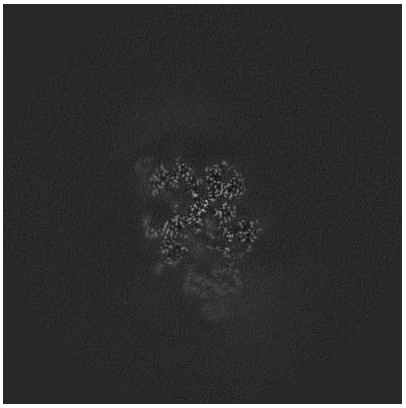

X Index: 240

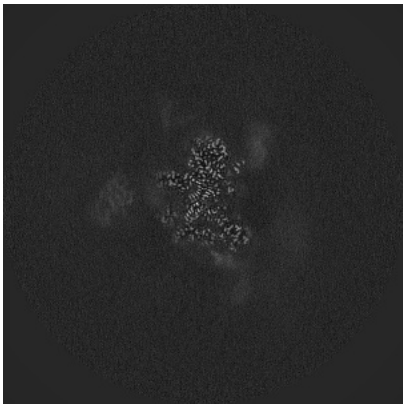

Y Index: 240


Z Index: 240

6.2.2 Raw map


X Index: 240

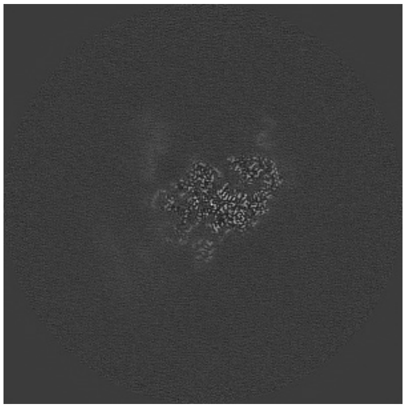

Y Index: 240


Z Index: 240

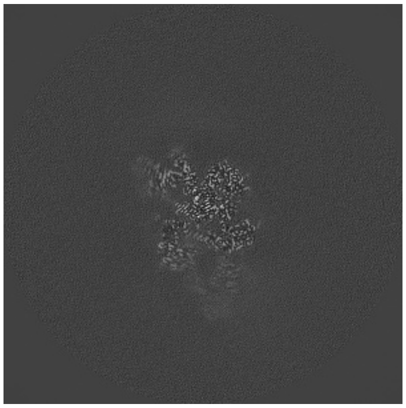
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices ⓘ

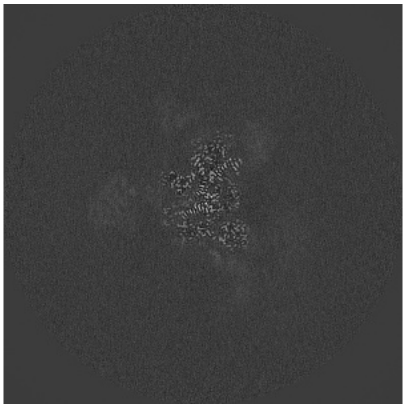
6.3.1 Primary map



X Index: 253

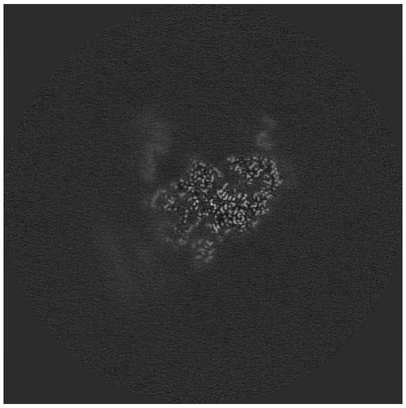


Y Index: 234

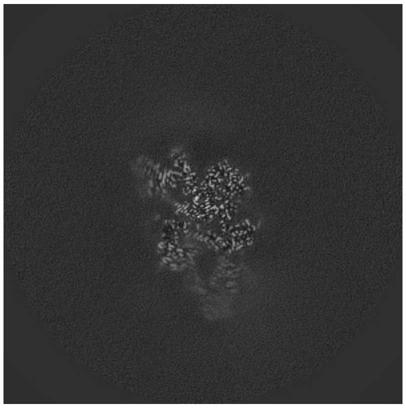


Z Index: 244

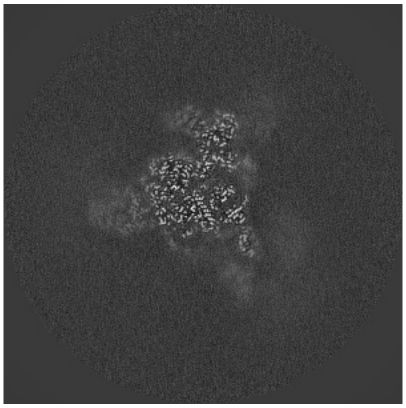
6.3.2 Raw map



X Index: 253



Y Index: 234

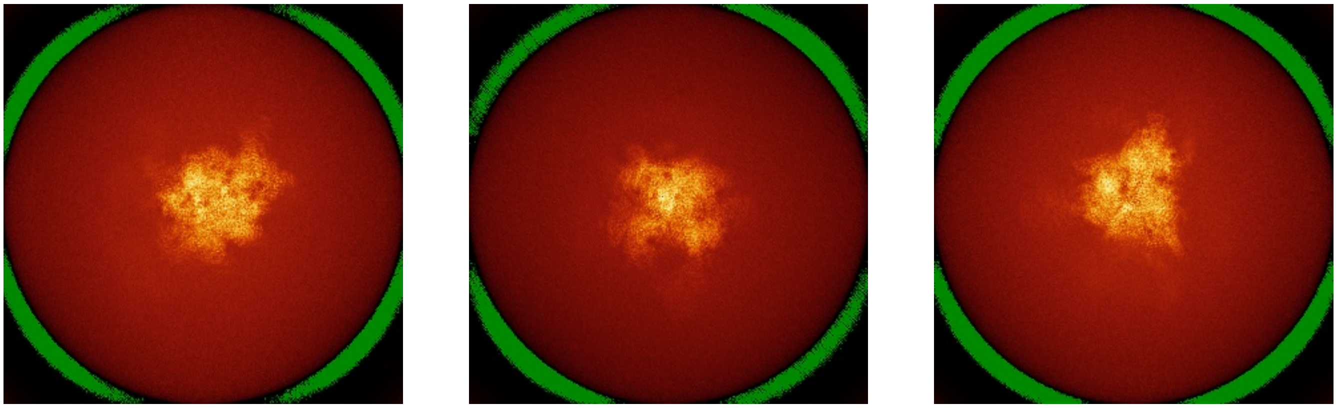


Z Index: 265

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

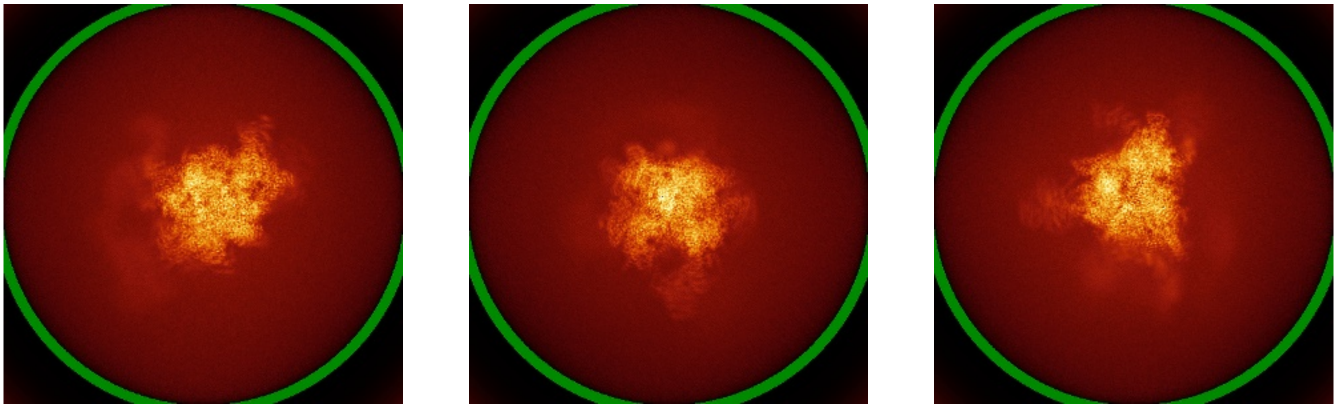
6.4 Orthogonal standard-deviation projections (False-color)

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

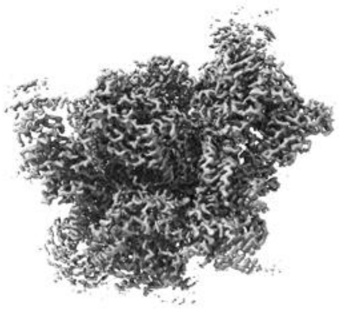
X

Y

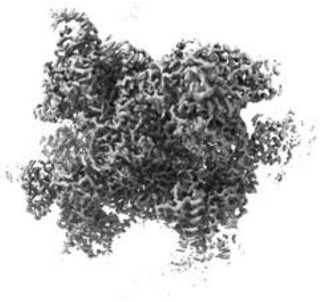
Z

The images above show the 3D surface view of the map at the recommended contour level 0.02. 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.

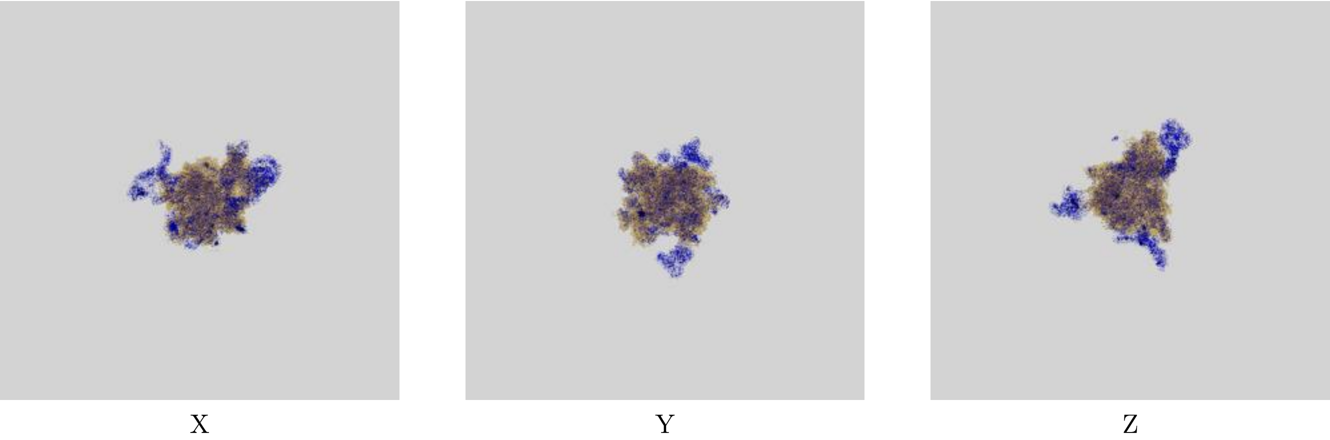
6.6 Mask visualisation (i)

This section shows the 3D surface view of the primary map at 50% transparency overlaid with the specified mask at 0% transparency

A mask typically either:

- Encompasses the whole structure
- Separates out a domain, a functional unit, a monomer or an area of interest from a larger structure

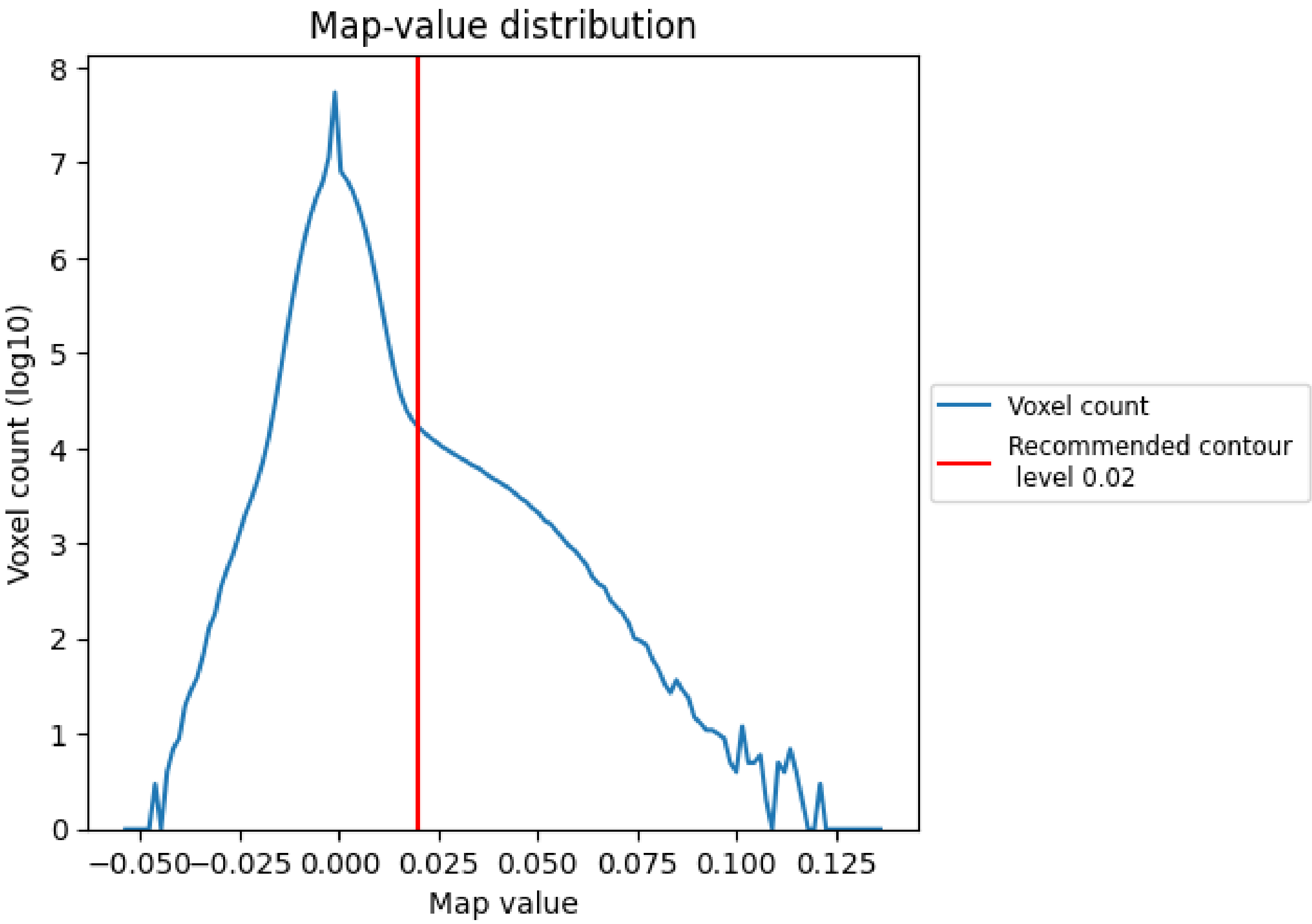
6.6.1 emd_47157_msk_1.map (i)


X
Y
Z

7 Map analysis ⓘ

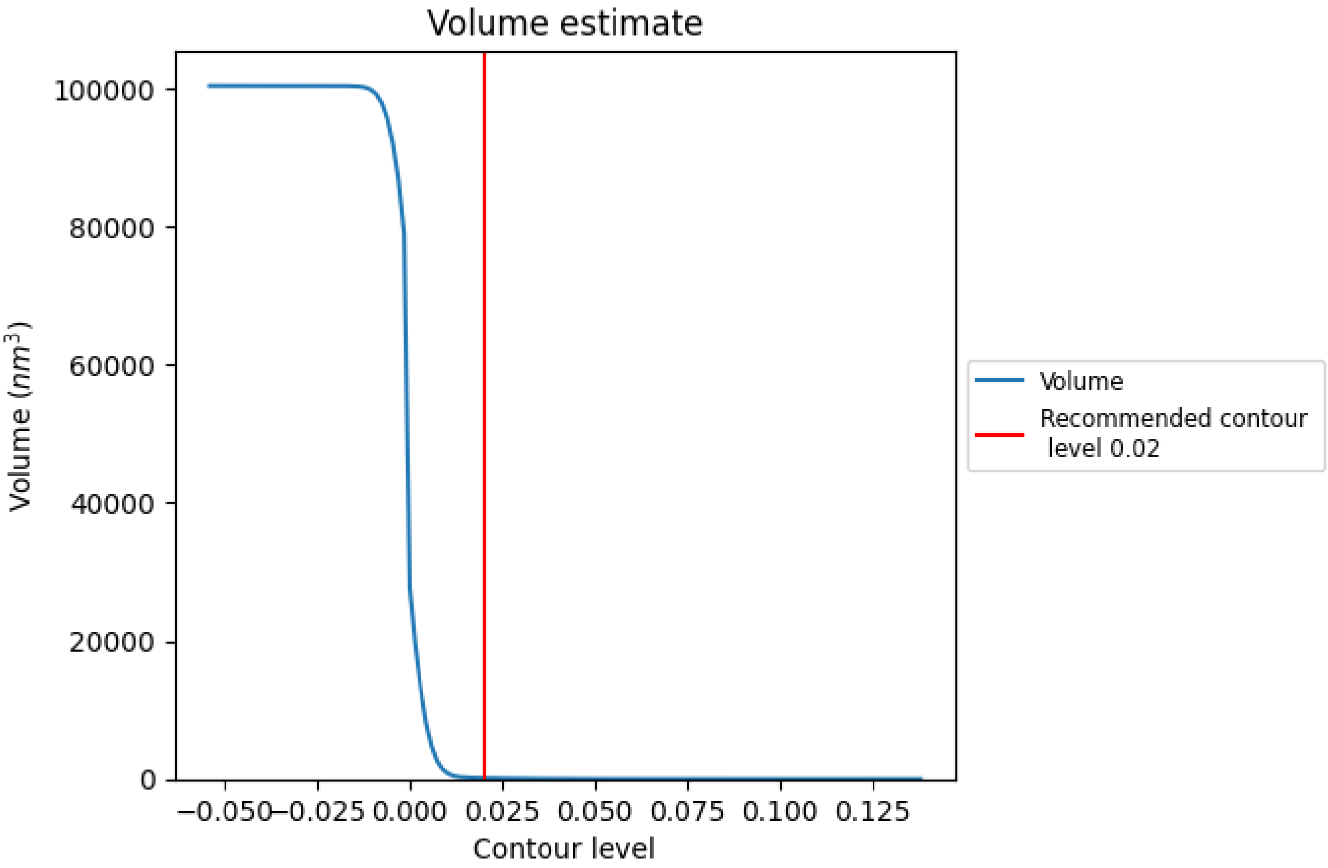
This section contains the results of statistical analysis of the map.

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.

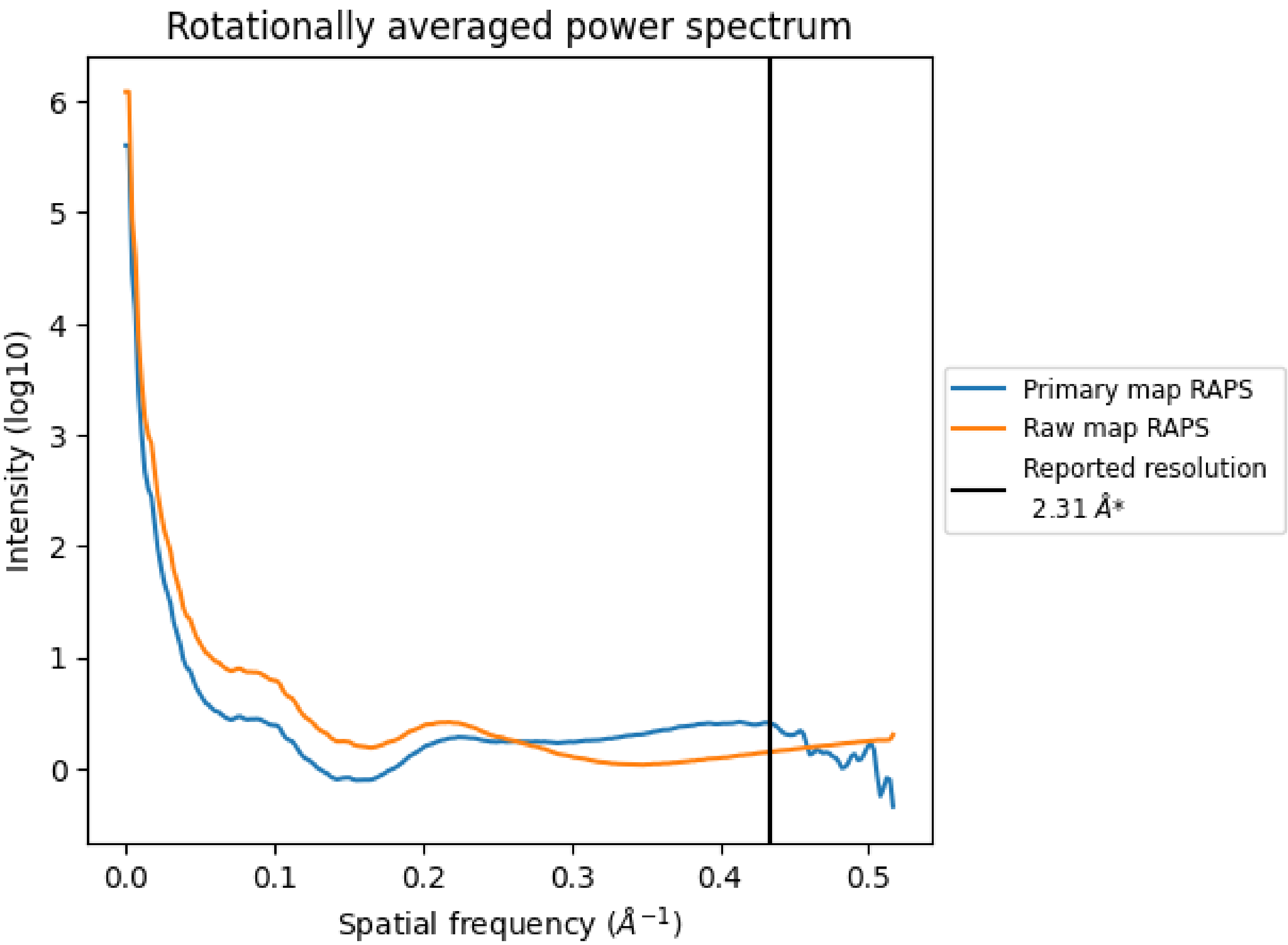
7.2 Volume estimate ⓘ



The volume at the recommended contour level is 146 nm^3; this corresponds to an approximate mass of 132 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 ⓘ

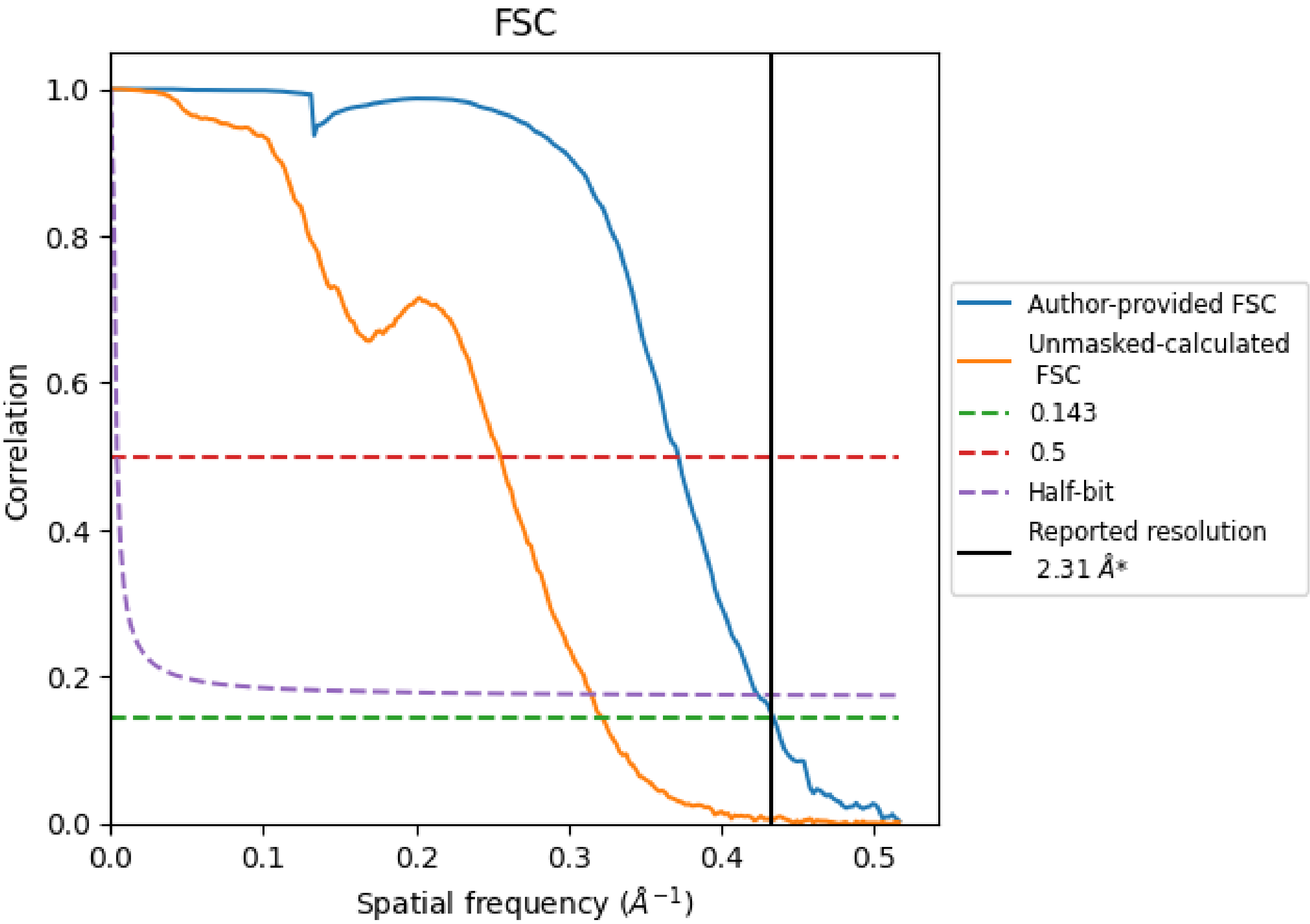


*Reported resolution corresponds to spatial frequency of 0.433 $Å^{-1}$

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 0.433 $\AA^{-1}$

8.2 Resolution estimates (i)

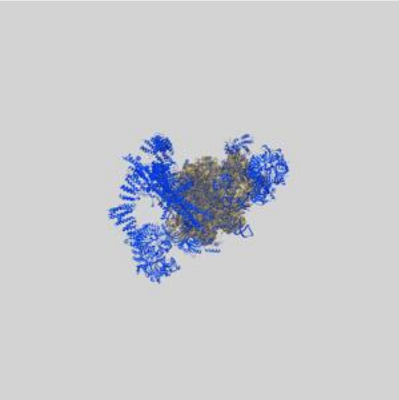
Resolution estimate (Å)	**Estimation criterion (FSC cut-off)**		
	0.143	0.5	Half-bit
Reported by author	2.31	-	-
Author-provided FSC curve	2.30	2.69	2.36
Unmasked-calculated*	3.10	3.92	3.18

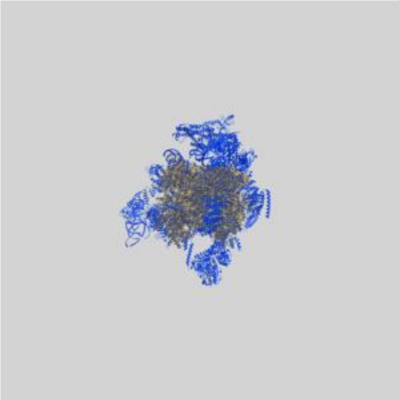
*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 3.10 differs from the reported value 2.31 by more than 10 %

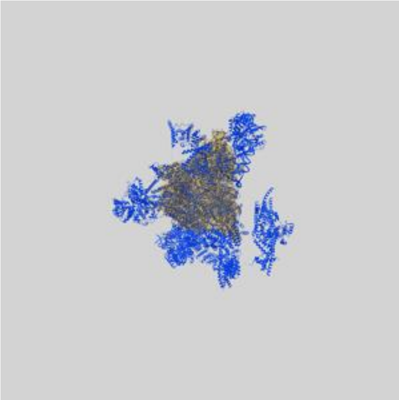
9 Map-model fit ⓘ

This section contains information regarding the fit between EMDB map EMD-47157 and PDB model 9DTR. Per-residue inclusion information can be found in section 3 on page 13.

9.1 Map-model overlay ⓘ

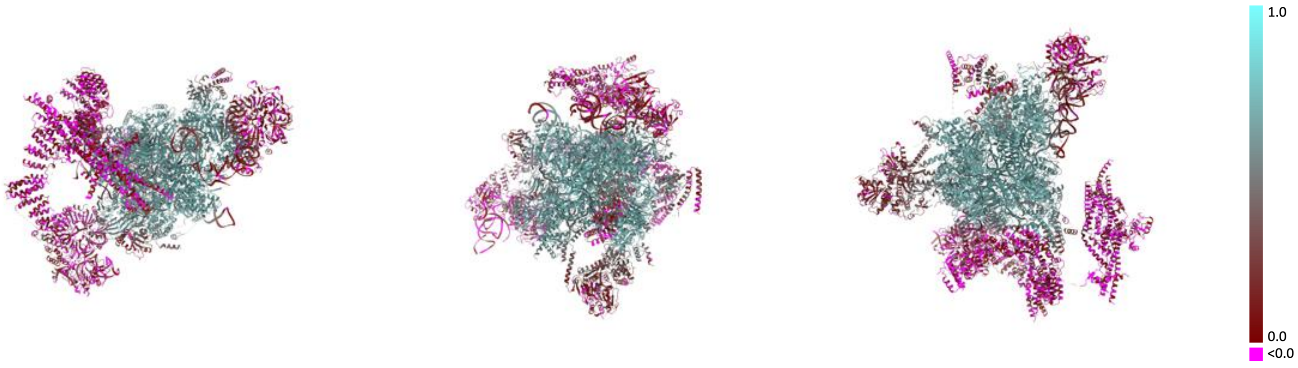

X


Y


Z

The images above show the 3D surface view of the map at the recommended contour level 0.02 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.

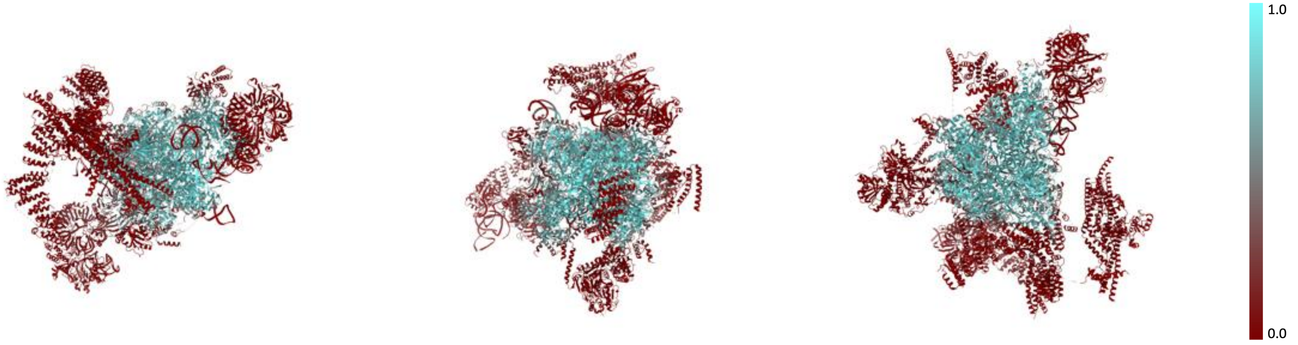
9.2 Q-score mapped to coordinate model (i)



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.

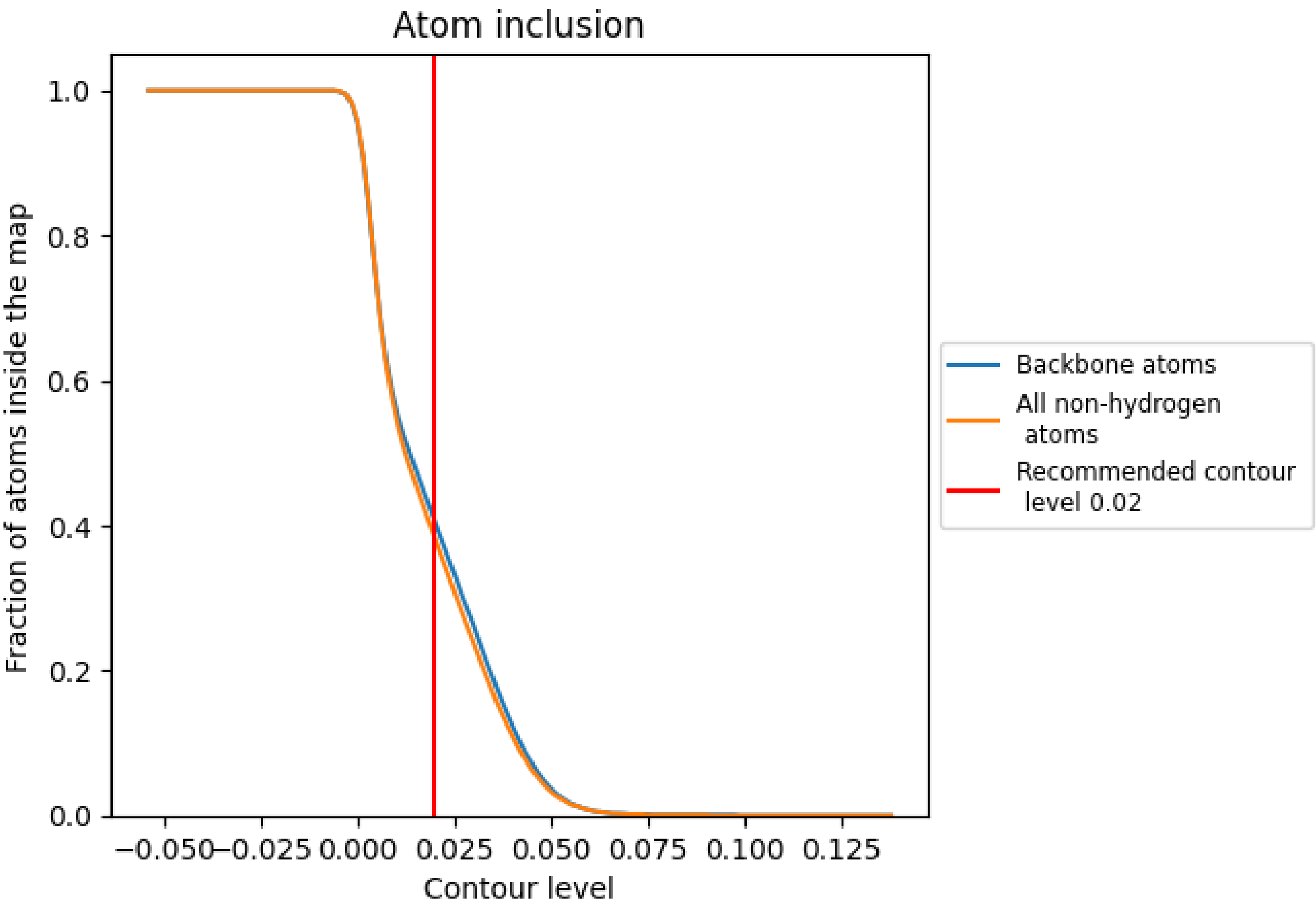
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.02).

9.4 Atom inclusion ⓘ



At the recommended contour level, 41% of all backbone atoms, 38% 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 (0.02) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.3830	0.4190
2	0.1890	0.1990
5	0.4020	0.3860
6	0.6940	0.6010
A	0.8090	0.6900
C	0.6380	0.6380
D	0.1020	0.3640
E	0.3640	0.4890
G	0.0000	0.2440
H	0.2550	0.3880
I	0.4820	0.5290
J	0.7740	0.6530
K	0.6140	0.6310
L	0.7560	0.6820
M	0.5670	0.6180
N	0.3400	0.4580
O	0.4200	0.4350
P	0.6700	0.6760
R	0.3160	0.3600
S	0.3100	0.3250
T	0.0000	0.0850
V	0.0330	0.2770
W	0.0000	0.0680
Y	0.0000	0.0490
Z	0.0000	0.0240
a	0.4950	0.5960
b	0.0000	0.1840
c	0.3990	0.5810
d	0.0000	0.2890
e	0.0000	0.0740
f	0.0000	0.0680
g	0.0000	0.1370
h	0.0000	0.1120
j	0.0000	0.0920
k	0.0000	0.0590

Continued on next page...

Continued from previous page...

Chain	Atom inclusion	Q-score
l	0.0000	0.0350
m	0.0000	0.0140
n	0.0000	0.0770
o	0.3940	0.5910
p	0.0000	0.0230
q	0.0000	0.0210
r	0.0000	0.0590
s	0.0000	0.0340
t	0.0000	0.0360
u	0.0000	0.0180
v	0.0000	0.0450
w	0.0000	0.0390
y	0.4080	0.4650