wwPDB EM Validation Summary Report (i)

Dec 18, 2022 – 08:49 am GMT

PDB ID	:	7ABI
EMDB ID	:	EMD-11697
Title	:	Human pre-Bact-2 spliceosome
Authors	:	Townsend, C.; Kastner, B.; Leelaram, M.N.; Bertram, K.; Stark, H.; Luehrmann, R.
Deposited on	:	2020-09-07
Resolution	:	8.00 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

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

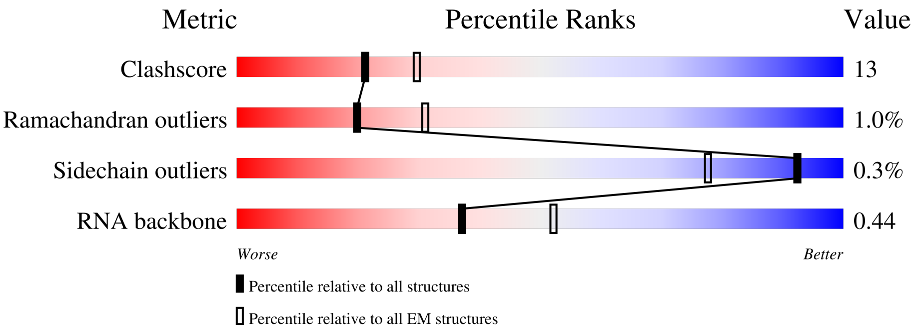
EMDB validation analysis	:	0.0.1.dev43
Mogul	:	1.8.4, CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 8.00 Å.

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



Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
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	D	357	7% / 82% • 15%
2	E	1217	• / 89% 8% •
3	b	86	5% / 85% 15%
3	i	86	• / 84% 16%
4	W	255	15% / 62% • 36%
5	w	424	• / 18% 82%
6	d	76	• / 91% 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	k	76	• 96% •
7	U	1485	52% 84% 5% • 11%
8	x	86	91% • 8%
9	6	106	5% 31% 25% 25% 6% 14%
10	B	225	40% 70% • 25%
11	z	125	86% 14%
12	q	73	• 100%
13	s	2136	9% 81% 19%
14	v	536	• 38% 62%
15	3	619	• 7% • 91%
16	0	396	• 36% • 62%
17	Q	144	• 87% 8% • •
18	1	322	34% • • 62%
19	L	802	14% 86%
20	N	199	• 26% • 72%
21	S	563	24% 61% • 37%
22	r	972	91% • 8%
23	R	229	6% 19% 81%
24	Y	904	10% • 89%
25	7	393	12% 59% • 40%
26	M	855	42% 76% 6% 17%
27	K	439	28% 72%
28	O	848	• 28% • 71%
29	y	110	88% • 9%
30	2	188	24% 41% 30% 15% • 13%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
31	G	514	55% 6% 38%
32	a	118	66% 34%
32	h	118	13% 81% 19%
33	o	301	33% 74% 24%
34	t	520	20% 48% 51%
35	f	240	27% 73%
35	m	240	34% 66%
36	Z	324	5% 8% 5% 82%
37	e	126	62% 38%
37	l	126	8% 66% 34%
38	A	2335	80% 9% 7%
39	8	579	97%
40	c	92	85% 15%
40	j	92	88% 12%
41	I	312	54% 44%
42	g	119	8% 78% 22%
42	n	119	29% 67% 33%
43	P	420	5% 54% 44%
44	5	116	28% 51% 19%
45	p	793	9% 15% 84%
46	F	464	10% 33% 64%
47	4	501	49% 81% 16%
48	u	1304	72% 28%
49	T	895	18% 80%

2 Entry composition (i)

There are 52 unique types of molecules in this entry. The entry contains 85564 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 U5 small nuclear ribonucleoprotein 40 kDa protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
1	D	302	1506	902	302	302	0	0

- Molecule 2 is a protein called Splicing factor 3B subunit 3.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
2	E	1177	5926	3572	1177	1177	0	0

- Molecule 3 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
3	b	73	364	218	73	73	0	0
3	i	72	359	215	72	72	0	0

- Molecule 4 is a protein called U2 small nuclear ribonucleoprotein A'.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
4	W	162	816	492	162	162	0	0

- Molecule 5 is a protein called Splicing factor 3B subunit 4.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
5	w	78	391	235	78	78	0	0

- Molecule 6 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms				AltConf	Trace
6	d	69	Total 344	C 206	N 69	O 69	0	0
6	k	73	Total 364	C 218	N 73	O 73	0	0

- Molecule 7 is a protein called Intron-binding protein aquarius.

Mol	Chain	Residues	Atoms				AltConf	Trace
7	U	1329	Total 6730	C 4072	N 1329	O 1329	0	0

- Molecule 8 is a protein called Splicing factor 3B subunit 5.

Mol	Chain	Residues	Atoms				AltConf	Trace
8	x	79	Total 397	C 239	N 79	O 79	0	0

- Molecule 9 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	6	91	Total 1947	C 871	N 359	O 626	P 91	0	0

- Molecule 10 is a protein called U2 small nuclear ribonucleoprotein B".

Mol	Chain	Residues	Atoms				AltConf	Trace
10	B	169	Total 851	C 513	N 169	O 169	0	0

- Molecule 11 is a protein called Splicing factor 3B subunit 6.

Mol	Chain	Residues	Atoms				AltConf	Trace
11	z	108	Total 544	C 328	N 108	O 108	0	0

- Molecule 12 is a protein called Ubiquitin-like protein 5.

Mol	Chain	Residues	Atoms				AltConf	Trace
12	q	73	Total 360	C 214	N 73	O 73	0	0

- Molecule 13 is a protein called U5 small nuclear ribonucleoprotein 200 kDa helicase.

Mol	Chain	Residues	Atoms				AltConf	Trace
13	s	1722	Total 8688	C 5244	N 1722	O 1722	0	0

- Molecule 14 is a protein called SNW domain-containing protein 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
14	v	206	Total 1053	C 641	N 206	O 206	0	0

- Molecule 15 is a protein called BUD13 homolog.

Mol	Chain	Residues	Atoms				AltConf	Trace
15	3	55	Total 283	C 173	N 55	O 55	0	0

- Molecule 16 is a protein called Smad nuclear-interacting protein 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
16	0	150	Total 761	C 461	N 150	O 150	0	0

- Molecule 17 is a protein called Protein BUD31 homolog.

Mol	Chain	Residues	Atoms				AltConf	Trace
17	Q	138	Total 695	C 419	N 138	O 138	0	0

- Molecule 18 is a protein called RNA-binding motif protein, X-linked 2.

Mol	Chain	Residues	Atoms				AltConf	Trace
18	1	122	Total 607	C 363	N 122	O 122	0	0

- Molecule 19 is a protein called Cell division cycle 5-like protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
19	L	116	Total 582	C 350	N 116	O 116	0	0

- Molecule 20 is a protein called Zinc finger matrin-type protein 2.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
20	N	56	277	165	56	56	0	0

- Molecule 21 is a protein called Beta-catenin-like protein 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
21	S	355	1771	1061	355	355	0	0

- Molecule 22 is a protein called 116 kDa U5 small nuclear ribonucleoprotein component.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
22	r	895	4531	2741	895	895	0	0

- Molecule 23 is a protein called Spliceosome-associated protein CWC15 homolog.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
23	R	44	219	131	44	44	0	0

- Molecule 24 is a protein called Serine/arginine repetitive matrix protein 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
24	Y	95	478	288	95	95	0	0

- Molecule 25 is a protein called DNA/RNA-binding protein KIN17.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
25	7	236	1178	706	236	236	0	0

- Molecule 26 is a protein called Pre-mRNA-splicing factor SYF1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
26	M	706	3553	2141	706	706	0	0

- Molecule 27 is a protein called Microfibrillar-associated protein 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
27	K	123	614	368	123	123	0	0

- Molecule 28 is a protein called Crooked neck-like protein 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
28	O	248	1243	747	248	248	0	0

- Molecule 29 is a protein called PHD finger-like domain-containing protein 5A.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
29	y	100	498	298	100	100	0	0

- Molecule 30 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
30	2	164	3479	1554	600	1161	164	0	0

- Molecule 31 is a protein called Pleiotropic regulator 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
31	G	320	1604	964	320	320	0	0

- Molecule 32 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
32	h	95	482	292	95	95	0	0
			Total	C	N	O		
32	a	78	393	237	78	78	0	0

- Molecule 33 is a protein called Peptidyl-prolyl cis-trans isomerase E.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
33	o	229	1135	677	229	229	0	0

- Molecule 34 is a protein called RING-type E3 ubiquitin-protein ligase PPIL2.

Mol	Chain	Residues	Atoms				AltConf	Trace
34	t	253	Total 1272	C 766	N 253	O 253	0	0

- Molecule 35 is a protein called Small nuclear ribonucleoprotein-associated proteins B and B'.

Mol	Chain	Residues	Atoms				AltConf	Trace
35	m	82	Total 413	C 249	N 82	O 82	0	0
35	f	64	Total 319	C 191	N 64	O 64	0	0

- Molecule 36 is a RNA chain called MINX M3 pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
36	Z	57	Total 1201	C 538	N 203	O 403	P 57	0	0

- Molecule 37 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms				AltConf	Trace
37	l	83	Total 415	C 249	N 83	O 83	0	0
37	e	78	Total 390	C 234	N 78	O 78	0	0

- Molecule 38 is a protein called Pre-mRNA-processing-splicing factor 8.

Mol	Chain	Residues	Atoms				AltConf	Trace
38	A	2163	Total 10968	C 6642	N 2163	O 2163	0	0

- Molecule 39 is a protein called Pre-mRNA-processing factor 17.

Mol	Chain	Residues	Atoms				AltConf	Trace
39	8	18	Total 92	C 56	N 18	O 18	0	0

- Molecule 40 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms				AltConf	Trace
40	j	81	Total 403	C 241	N 81	O 81	0	0
40	c	78	Total 388	C 232	N 78	O 78	0	0

- Molecule 41 is a protein called Pre-mRNA-splicing factor 38A.

Mol	Chain	Residues	Atoms				AltConf	Trace
41	I	176	Total 883	C 531	N 176	O 176	0	0

- Molecule 42 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms				AltConf	Trace
42	n	80	Total 402	C 242	N 80	O 80	0	0
42	g	93	Total 469	C 283	N 93	O 93	0	0

- Molecule 43 is a protein called Pre-mRNA-splicing factor RBM22.

Mol	Chain	Residues	Atoms				AltConf	Trace
43	P	237	Total 1194	C 720	N 237	O 237	0	0

- Molecule 44 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
44	5	114	Total 2397	C 1074	N 399	O 810	P 114	0	0

- Molecule 45 is a protein called Splicing factor 3A subunit 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
45	p	124	Total 631	C 383	N 124	O 124	0	0

- Molecule 46 is a protein called Splicing factor 3A subunit 2.

Mol	Chain	Residues	Atoms				AltConf	Trace
46	F	165	Total 835	C 505	N 165	O 165	0	0

- Molecule 47 is a protein called Splicing factor 3A subunit 3.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
47	4	421	2110	1268	421	421	0	0

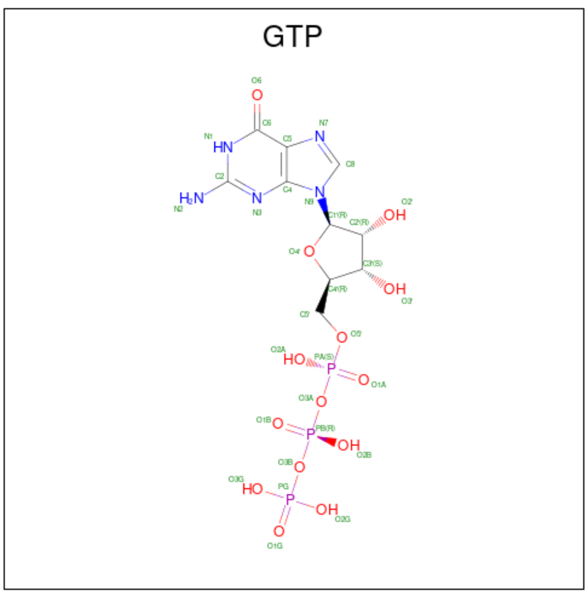
- Molecule 48 is a protein called Splicing factor 3B subunit 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
48	u	941	4748	2866	941	941	0	0

- Molecule 49 is a protein called Splicing factor 3B subunit 2.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
49	T	183	942	576	183	183	0	0

- Molecule 50 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).

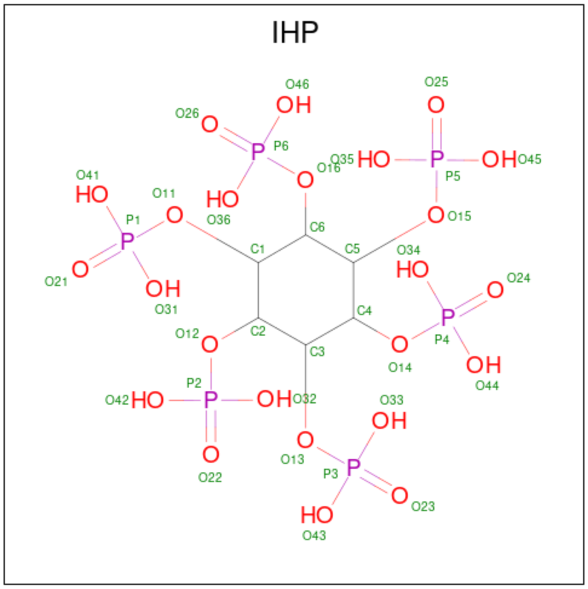


Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
50	r	1	32	10	5	14	3	0

- Molecule 51 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
51	r	1	1	1	0

- Molecule 52 is INOSITOL HEXAKISPHOSPHATE (three-letter code: IHP) (formula: $C_6H_{18}O_{24}P_6$).



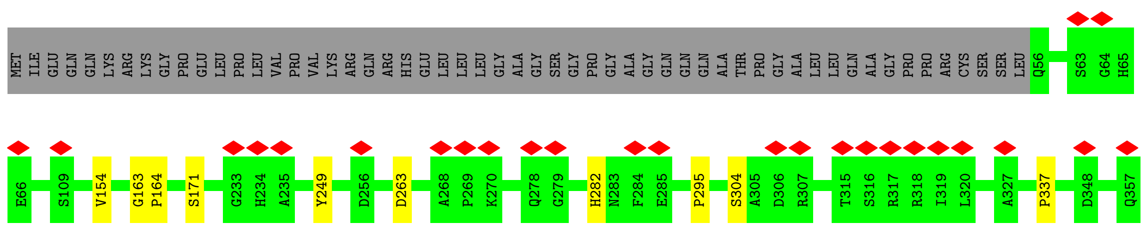
Mol	Chain	Residues	Atoms				AltConf
			Total	C	O	P	
52	A	1	36	6	24	6	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion in map density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

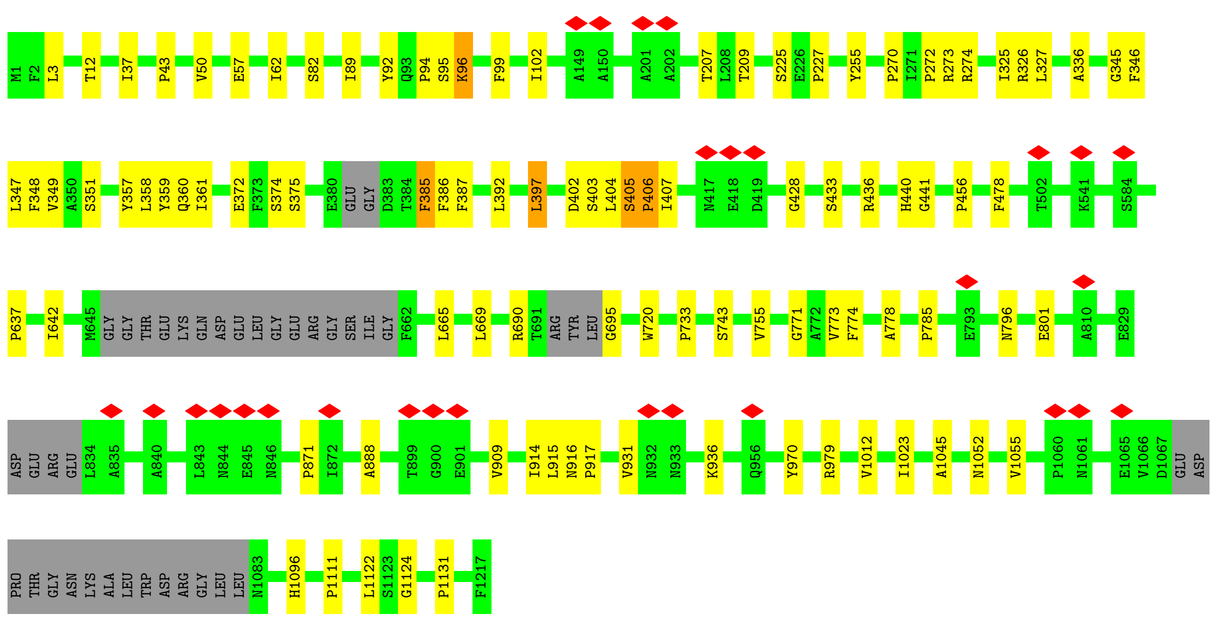
- Molecule 1: U5 small nuclear ribonucleoprotein 40 kDa protein

Chain D:



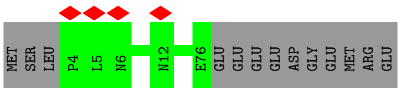
- Molecule 2: Splicing factor 3B subunit 3

Chain E:

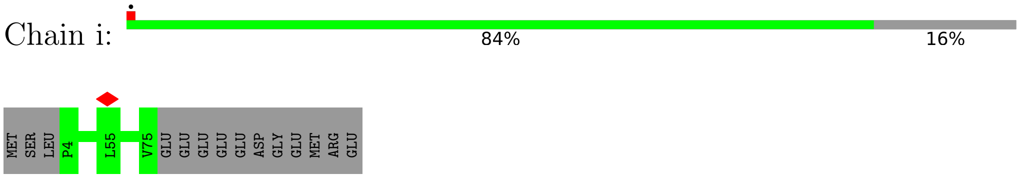


- Molecule 3: Small nuclear ribonucleoprotein F

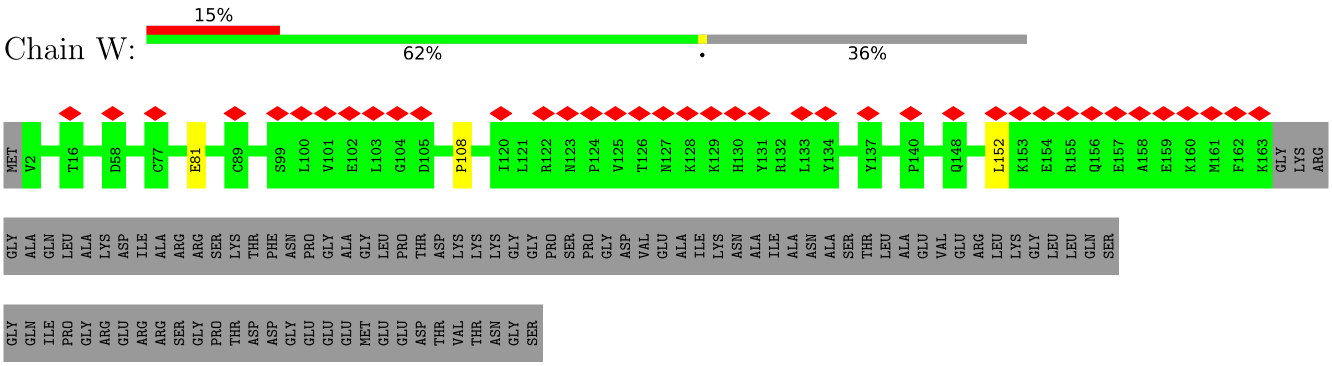
Chain b:



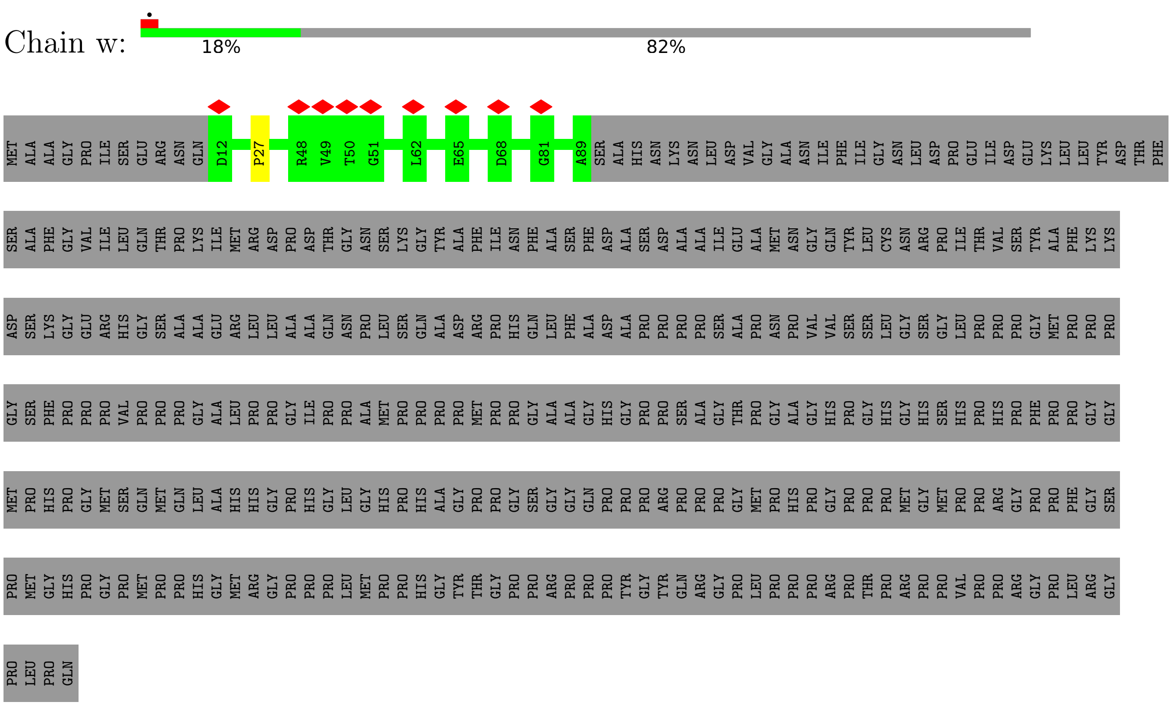
• Molecule 3: Small nuclear ribonucleoprotein F



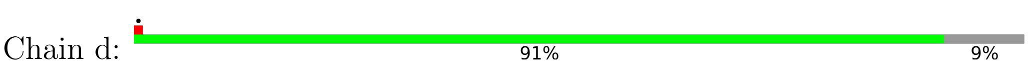
• Molecule 4: U2 small nuclear ribonucleoprotein A'

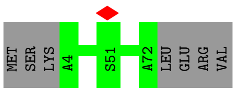


• Molecule 5: Splicing factor 3B subunit 4



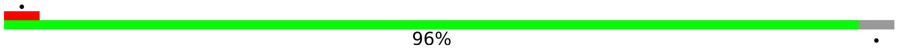
• Molecule 6: Small nuclear ribonucleoprotein G

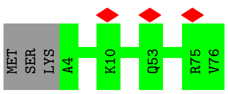




• Molecule 6: Small nuclear ribonucleoprotein G

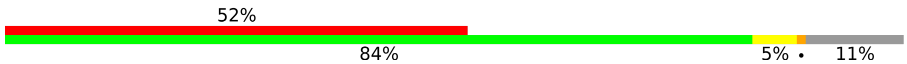
Chain k:

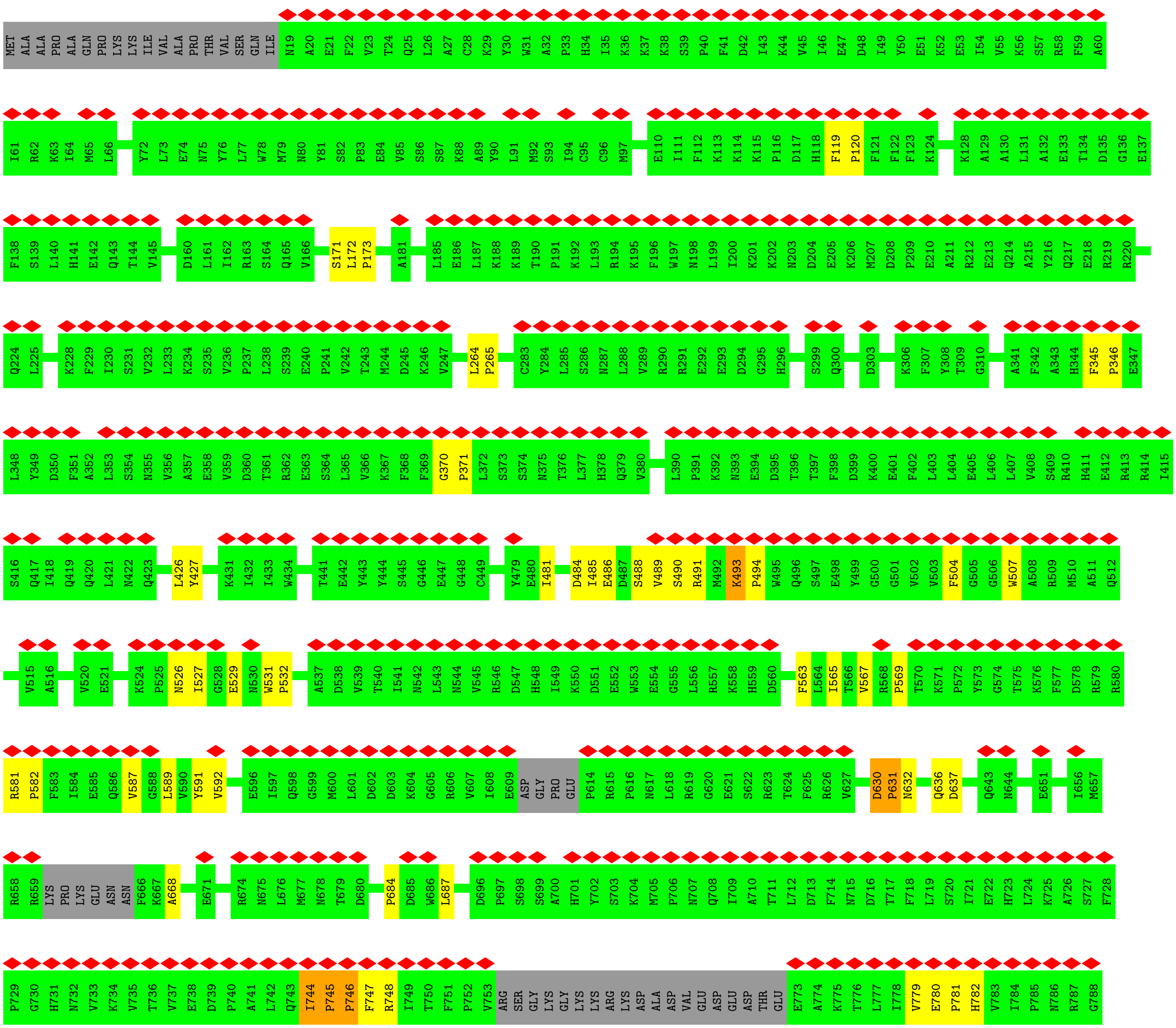


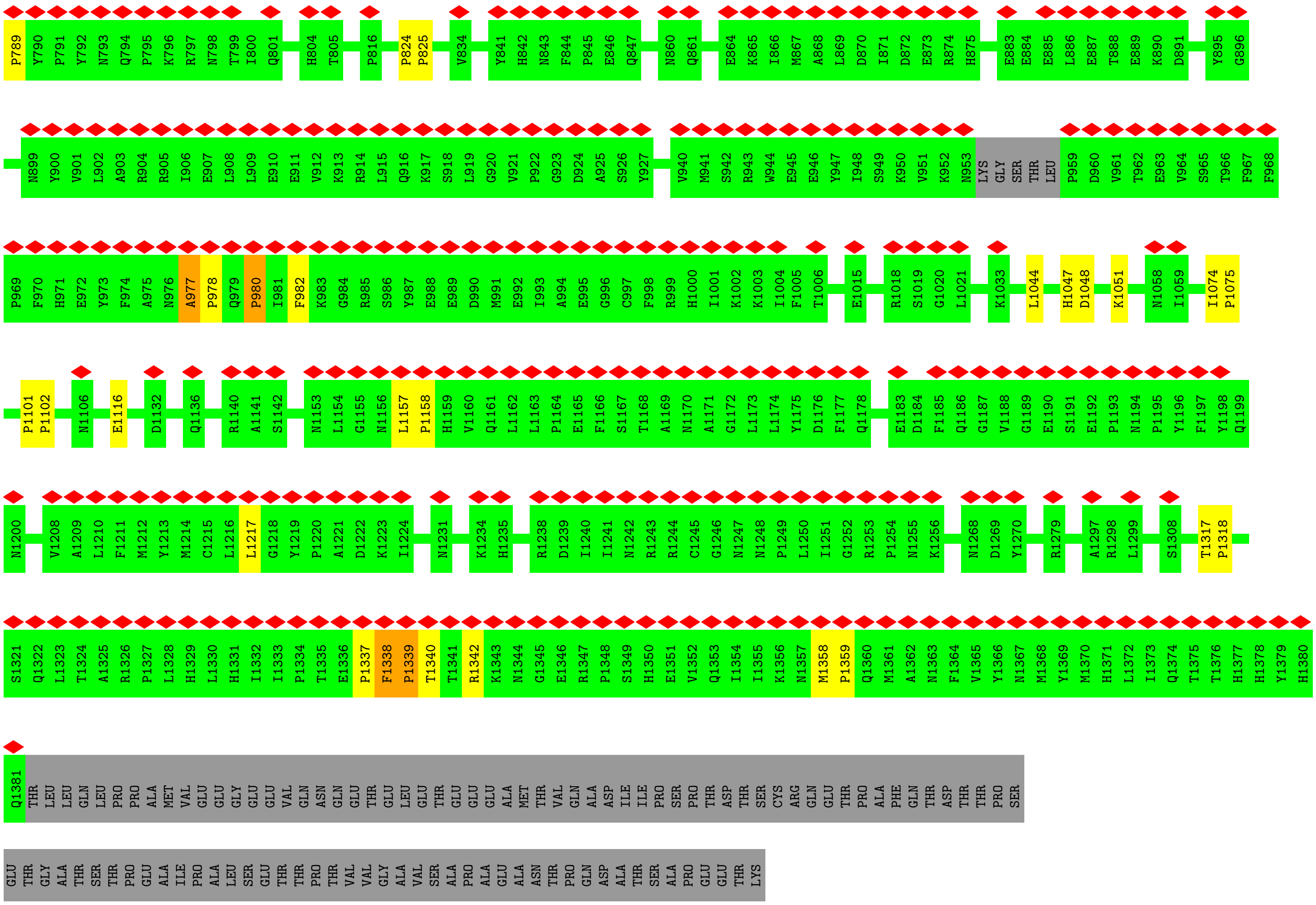


• Molecule 7: Intron-binding protein aquarius

Chain U:

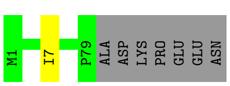






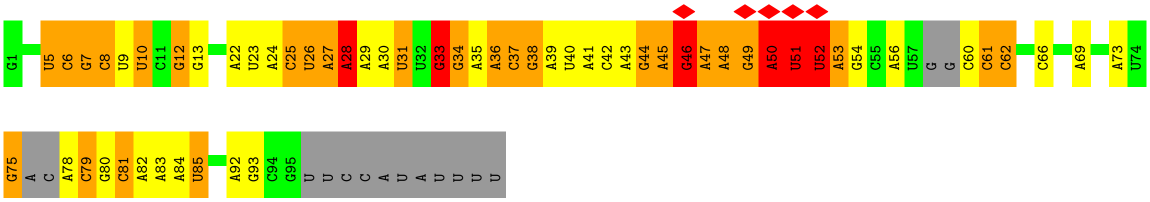
• Molecule 8: Splicing factor 3B subunit 5

Chain x:



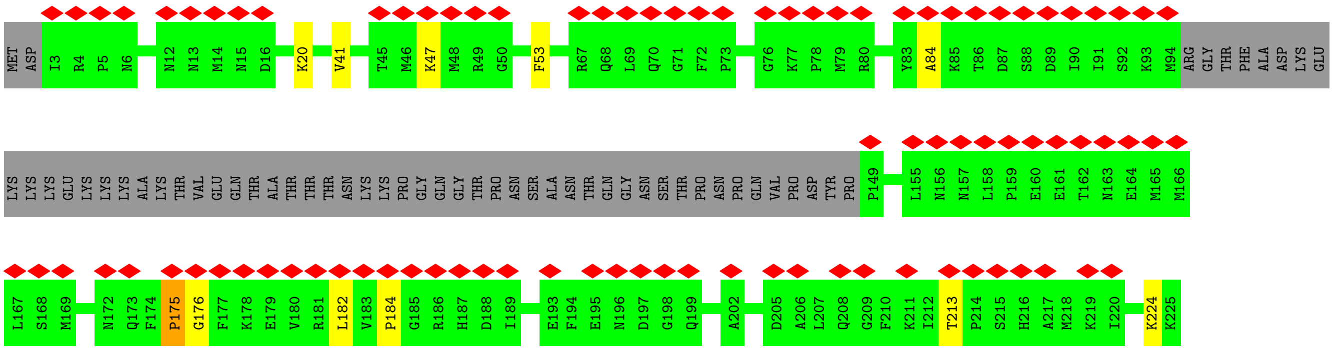
• Molecule 9: U6 snRNA

Chain 6:

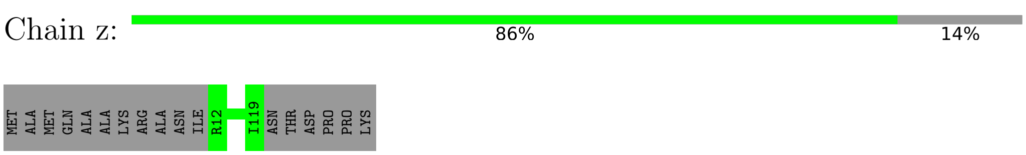


• Molecule 10: U2 small nuclear ribonucleoprotein B”

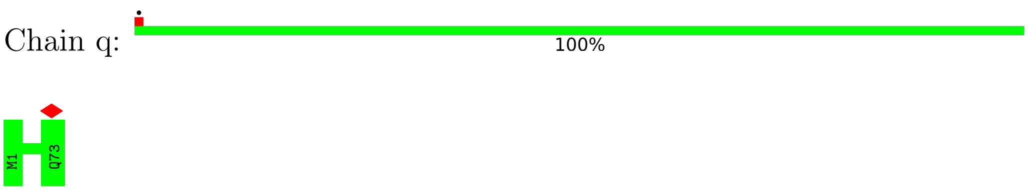
Chain B:



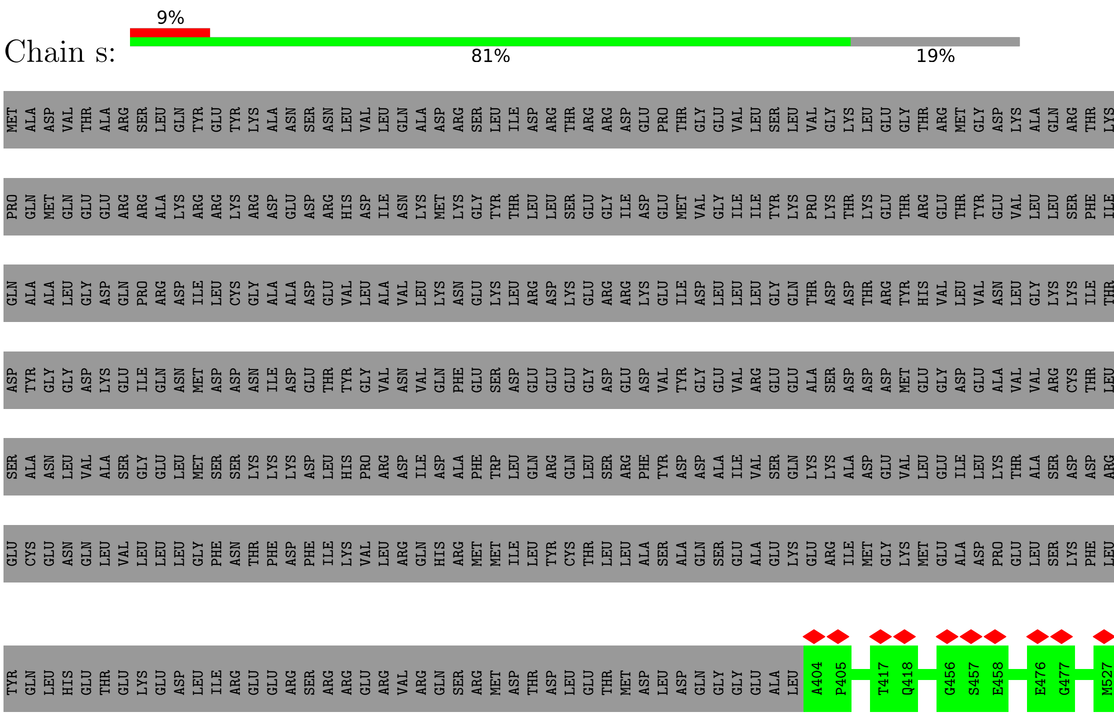
- Molecule 11: Splicing factor 3B subunit 6

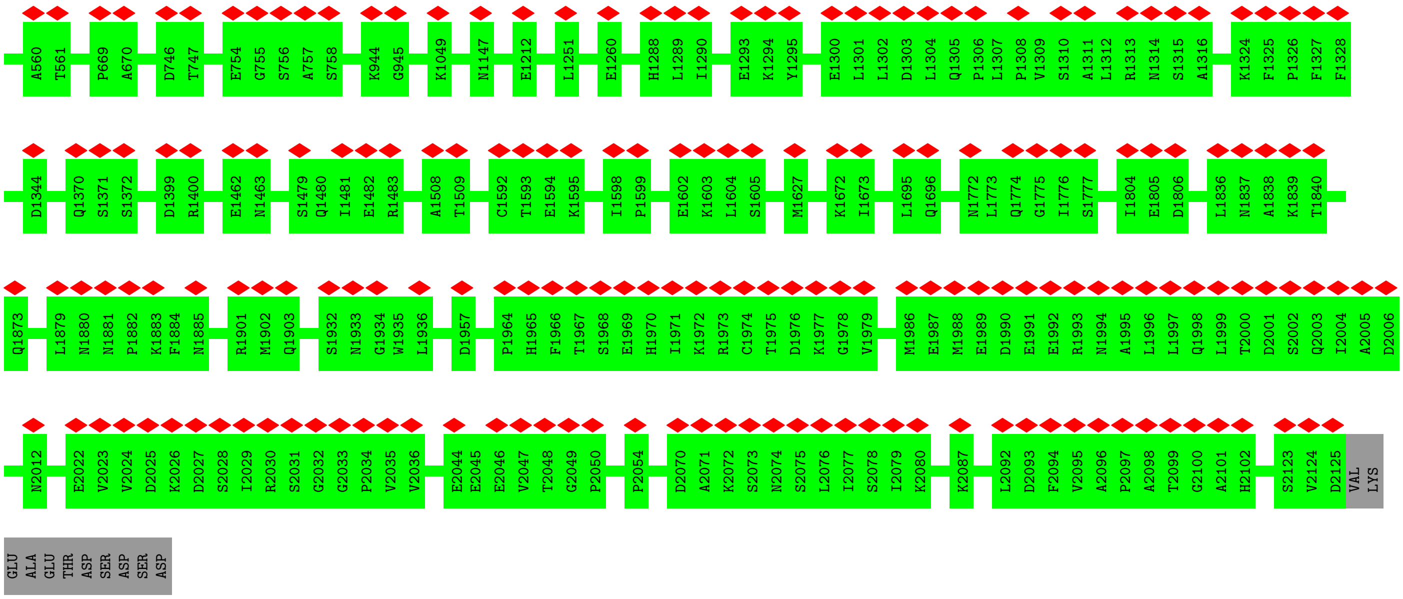


- Molecule 12: Ubiquitin-like protein 5



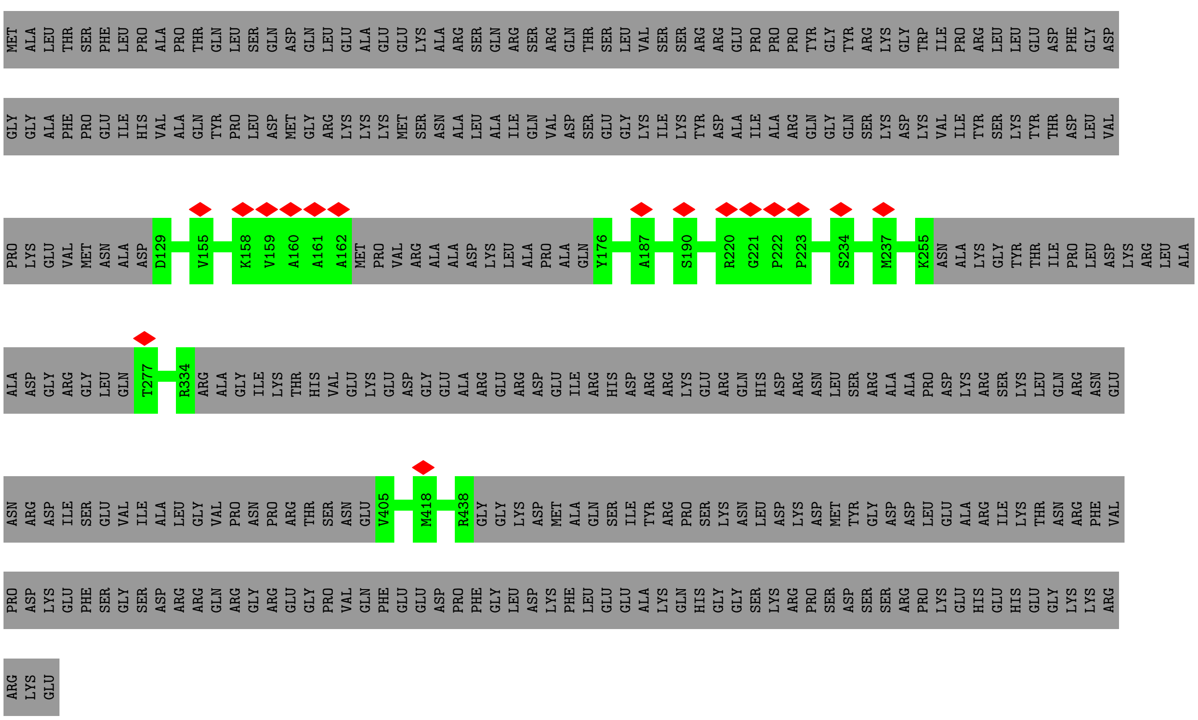
- Molecule 13: U5 small nuclear ribonucleoprotein 200 kDa helicase





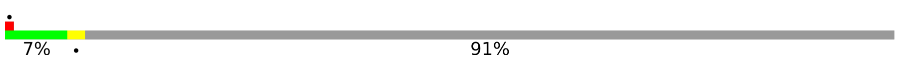
• Molecule 14: SNW domain-containing protein 1

Chain v:



• Molecule 15: BUD13 homolog

Chain 3:



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

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

ARG HIS ASP SER SER ASP THR SER PRO PRO ARG ARG ALA ARG HIS ASP SER PRO ASP PRO SER PRO PRO ARG ARG PRO GLN HIS ASN SER SER GLY ALA SER PRO ARG ARG VAL ARG HIS ASP SER PRO ASP PRO SER PRO PRO ARG ARG ALA ARG HIS GLY SER SER ASP ILE SER SER

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

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

HIS LYS GLN SER PRO GLY HIS GLN ASP SER ASP SER ASP LEU SER PRO PRO ARG ASN ARG PRO ARG HIS ARG SER SER ASP SER ASP LEU SER PRO PRO ARG ARG ARG GLN ARG THR LYS SER SER ASP SER ASP LEU SER PRO PRO ARG ARG SER GLN PRO PRO GLY LYS LYS ALA ALA

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

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

GLU GLY ASP PRO MET ALA ASN PHE ILE LYS LYS ASN LYS ALA LYS GLU ASN LYS ASN LYS LYS VAL R563 P564 S567 G568 P569 F576 ASN ILE W579 P580 G581 Y582 R583 W584 N591 G592 F593 E608 A611 D618 M619

- Molecule 16: Smad nuclear-interacting protein 1

Chain 0:

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

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

ASP ARG GLN HIS ARG GLU PRO SER GLU GLN GLN HIS ARG ARG ALA ARG ASN SER ASP ARG ASP ARG HIS ARG GLY HIS SER HIS GLN ARG ARG THR SER ASN GLU ARG PRO GLY SER GLY GLN GLY GLN GLY ARG ASP ARG ASP THR GLN ASN LEU GLN ALA GLN GLU GLU GLU ARG GLU

PHE TYR ASN ALA ARG ARG ARG GLU HIS ARG GLN ARG ASN ASP VAL GLY GLY GLY GLY SER GLU SER GLN GLU LEU VAL PRO ARG PRO GLY GLY ASN LYS GLU LYS GLU VAL PRO ALA K221 E222 F226 E227 P255 L270 P271 H276 D317 G334 L370 HIS GLU

SER SER ASP THR SER GLU ILE ASP ARG LYS ASP ASP GLU ASP GLU GLU GLU GLU GLU GLU VAL SER ASP SER

- Molecule 17: Protein BUD31 homolog

Chain Q:

MET PRO K3 P11 E19 P20 R41 S45 L46 W47 P48 R109 D110 T111 N116 C117 V128 C134 R140 GLY CYS SER GLY

- Molecule 18: RNA-binding motif protein, X-linked 2

Chain 1: 34% 62%

MET N2 P3 E49 G50 D51 C54 S57 Q58 D70 K71 K72 T73 G74 K75 D95 N96 E123 ILE ASP ASP VAL THR ARG GLN LEU GLN GLU LYS GLY CYS GLY ALA ARG THR PRO SER PRO SER LEU SER GLU SER SER GLU ASP GLU LYS PRO THR LYS LYS HIS LYS

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

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

TYR ASN GLY ARG SER GLU GLY ARG SER TYR ARG SER ARG SER ARG SER ARG ASP LYS SER HIS ARG HIS LYS ARG ALA ARG ARG SER ARG GLU ARG GLU SER SER ASN PRO SER ASP ARG TRP ARG HIS

- Molecule 19: Cell division cycle 5-like protein

Chain L: 14% 86%

MET PRO ARG ILE MET ILE LYS G8 N29 A86 P87 R110 ASP ASN GLU GLU GLU THR THR ASP ASP PRO ARG LYS LEU LYS PRO GLY GLU ILE ASP PRO ASN PRO GLU THR LYS PRO ALA ARG PRO ASP PRO ILE ASP MET ASP GLU D147 L159 ALA ASN THR GLN GLY LYS

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

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

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

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

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

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

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

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

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

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

ARG LEU GLU CYS LEU LYS GLU ASP VAL GLN ARG GLN GLN GLU ARG GLU LYS GLU LEU GLN HIS ARG TYR ALA ASP LEU LEU LEU GLU LYS GLU THR LEU LYS SER LYS PHE

- Molecule 20: Zinc finger matrin-type protein 2

Chain N: 26% 72%

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

LYS LEU GLY LYS THR ILE VAL ILE THR LYS THR THR PRO GLN SER GLU MET GLY G79 C82 D86 D97 Q105 M110 E134 GLU LYS GLN LYS ASP TYR ASP PHE GLU GLU ARG MET LYS GLU LEU ARG GLU GLU GLU LYS ALA LYS ALA TYR LYS LYS GLU LYS

GLN LYS GLU LYS LYS ARG ARG ALA GLU GLU ASP LEU THR PHE GLU GLU ASP ASP GLU MET ALA ALA VAL MET GLY PHE SER GLY PHE GLY SER THR LYS LYS SER TYR

- Molecule 21: Beta-catenin-like protein 1

Chain S: 24% 61% • 37%

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

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

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

ALA GLU VAL LEU ILE ASP ALA LEU VAL ASP GLY GLN VAL VAL ALA LEU LEU VAL GLN ASN LEU GLU ARG LEU ASP GLU SER VAL K209 E210 E211 A212 D213 G214 V215 H216 N217 T218 I221 N224 E227 F228 R229 P230 E231 Q238 Q239 K247 K252 M253 P254 F255 D256

I268 D272 E278 L279 E282 F296 K297 R298 H299 N300 P301 S302 T303 A304 E305 K332 G333 E334 L341 R344 E345 K346 K347 I348 H371 K372 D375 I376 R380 T381 I382 F383 P384 K388 S389 P390 R391 K392 I393 K394 K395 V396 G397 T398 T399 E400 K401 E402

R419 G420 Q421 T431 E432 N433 D434 E443 L444 H445 F446 K447 Y448 L449 G450 A451 M452 Q453 V454 A455 D456 K457 H464 D465 M466 V467 R468 R469 G470 E471 I472 I473 D474 N475 D476 T477 E478 E479 R484 R485 G489 L490 F491 V492 L493 Q494 H495 I496 N507 V508 P509 Q510 Q517

I518 L519 N520 M521 R522 G523 S524 S525 I526 K527 I528 H531 K534 E535 Y536 A537 E538 N539 D542 G543 R544 S545 P546 E547 F548 R549 E550 N551 E552 Q553 K554 F563

- Molecule 22: 116 kDa U5 small nuclear ribonucleoprotein component

Chain r: 91% • 8%

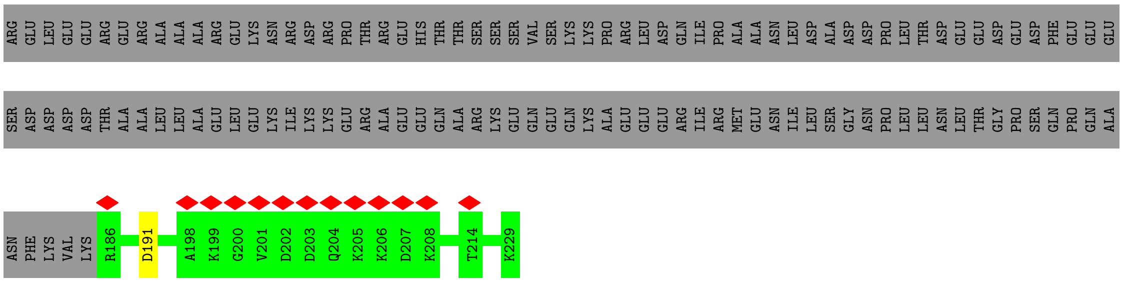
MET ASP THR ASP LEU TYR ASP GLU PHE GLY ASN TYR ILE GLY PRO GLU LEU ASP SER ASP GLU ASP ASP ASP GLU LEU GLY ARG GLU THR LYS ASP LEU ASP GLU MET ASP ASP ASP ASP ASP ASP ASP ASP VAL GLY ASP HIS ASP ASP ASP HIS PRO GLY MET GLU VAL VAL LEU HIS

GLU ASP K63 E83 E84 D85 K98 W344 Y636 K941 V947 S948 D954 D955 P956 M957 LEU LEU GLU LEU ALA LYS GLN ASP VAL VAL LEU ASN TYR PRO MET

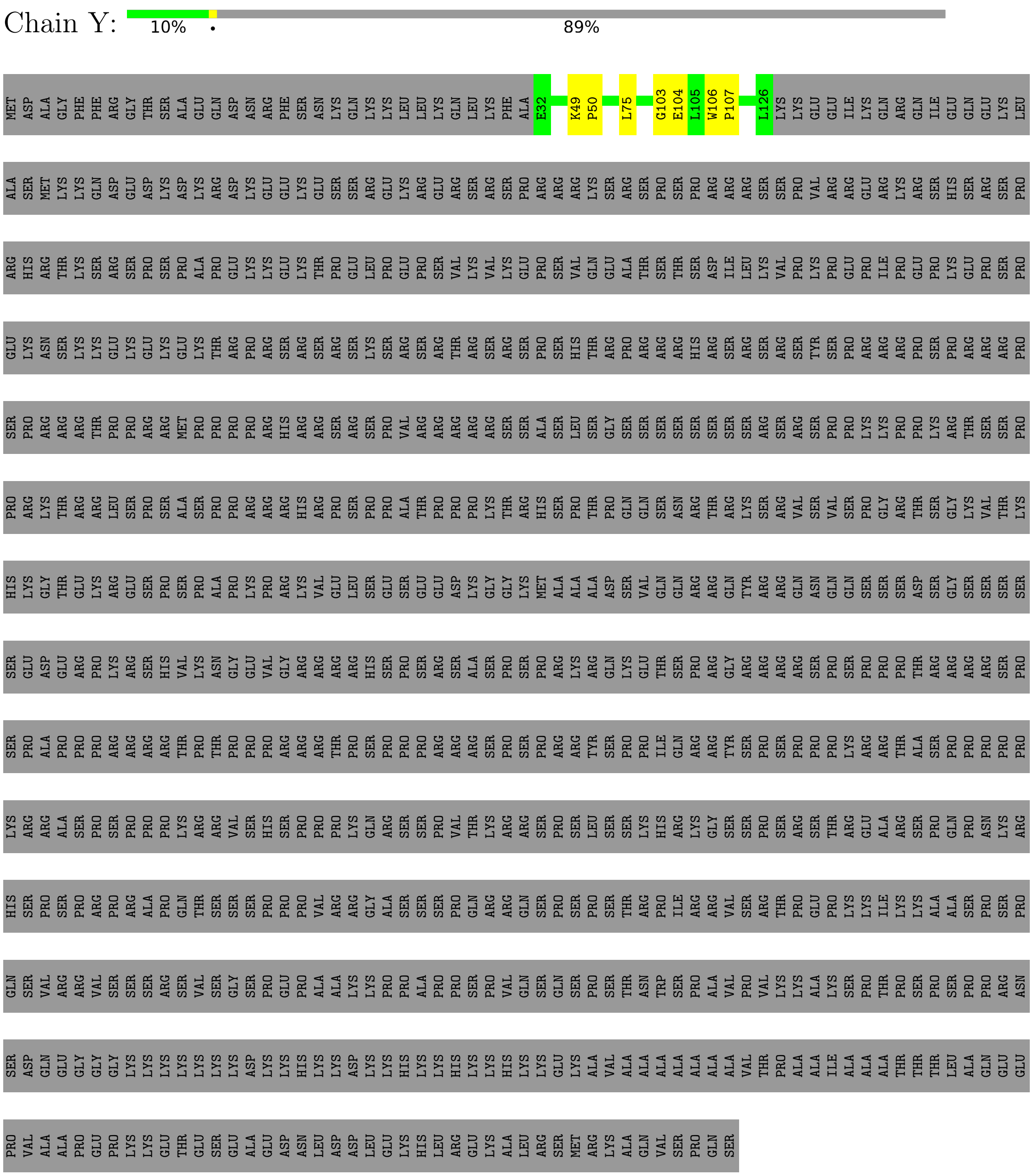
- Molecule 23: Spliceosome-associated protein CWC15 homolog

Chain R: 6% 19% 81%

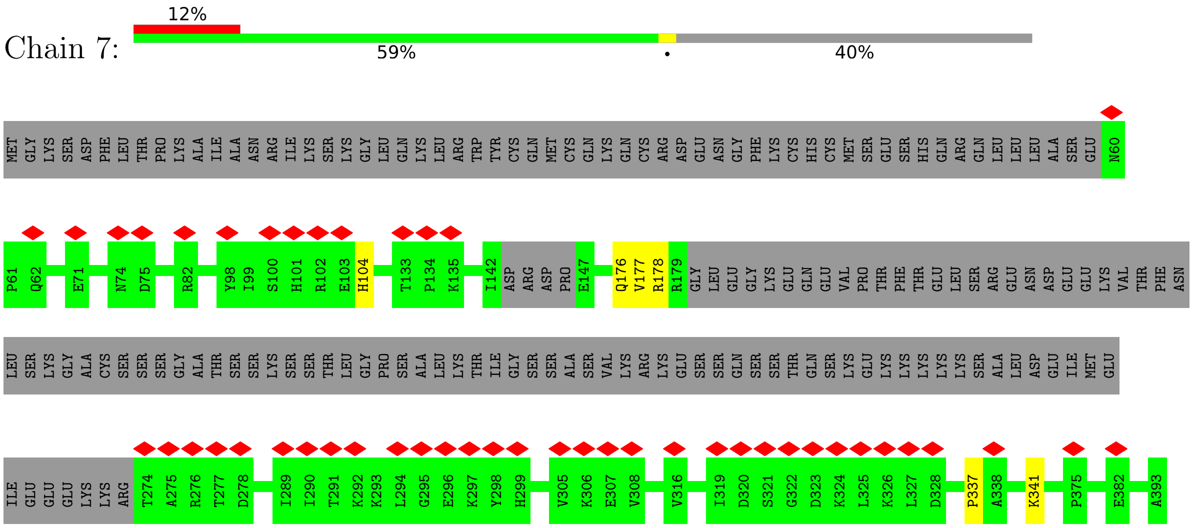
MET THR THR ALA ALA ARG PRO THR PHE GLU PRO ALA ARG GLY GLY ARG GLY LYS GLY GLU GLY ASP LEU SER GLN LEU SER LYS GLN TYR SER SER ARG ASP LEU PRO SER HIS THR LYS ILE LYS TYR ARG GLN THR THR GLN ASP ALA PRO GLU GLU VAL ARG ASN ARG ASP PHE ARG



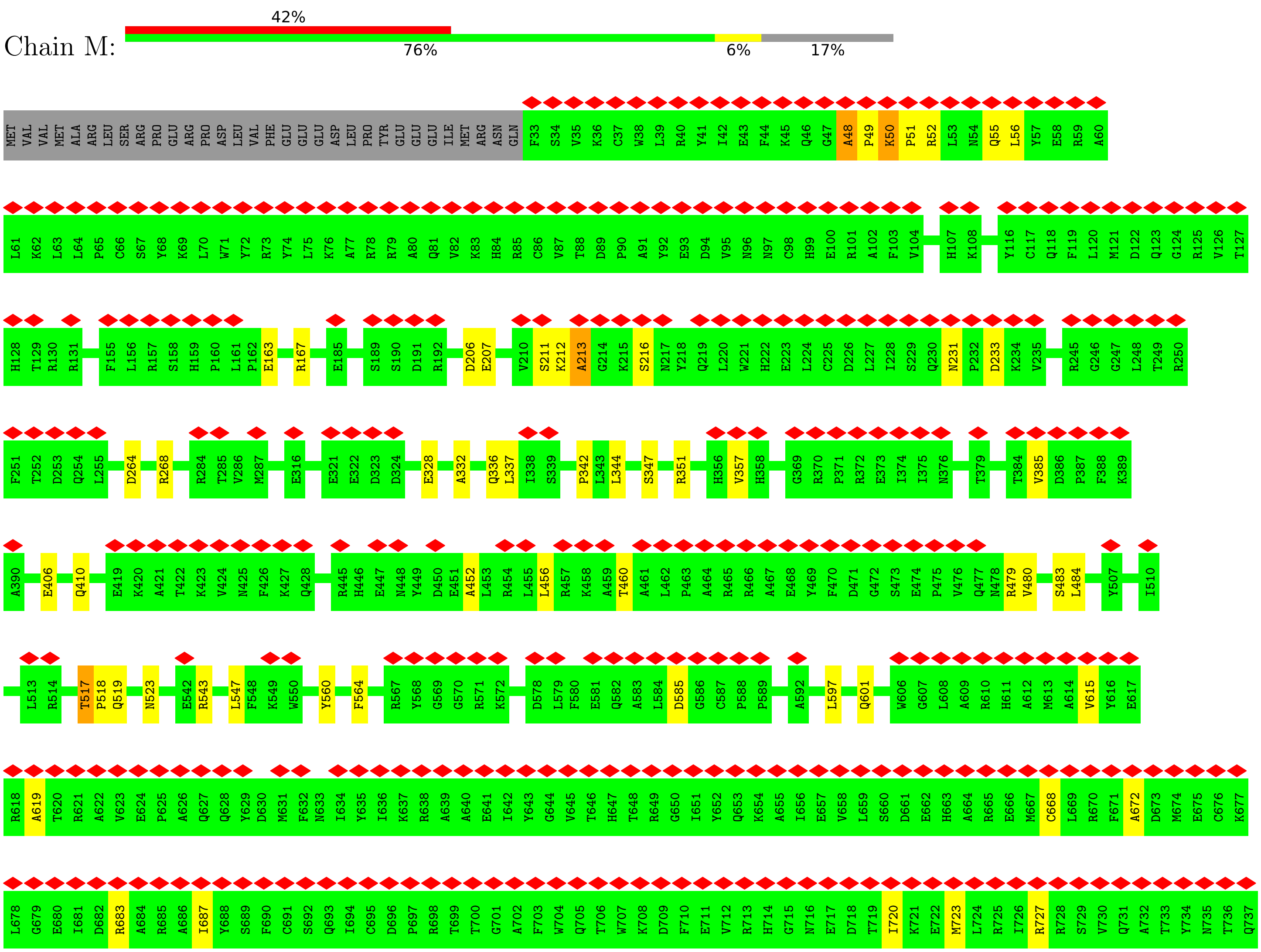
• Molecule 24: Serine/arginine repetitive matrix protein 1

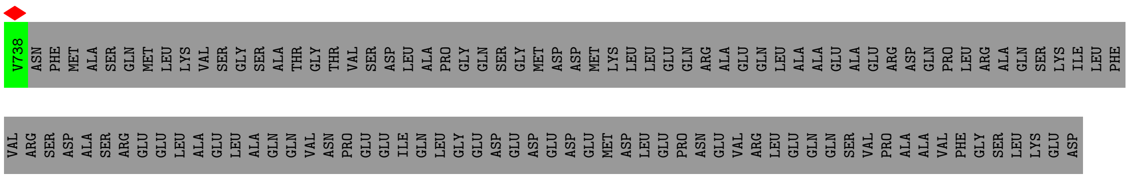


• Molecule 25: DNA/RNA-binding protein KIN17



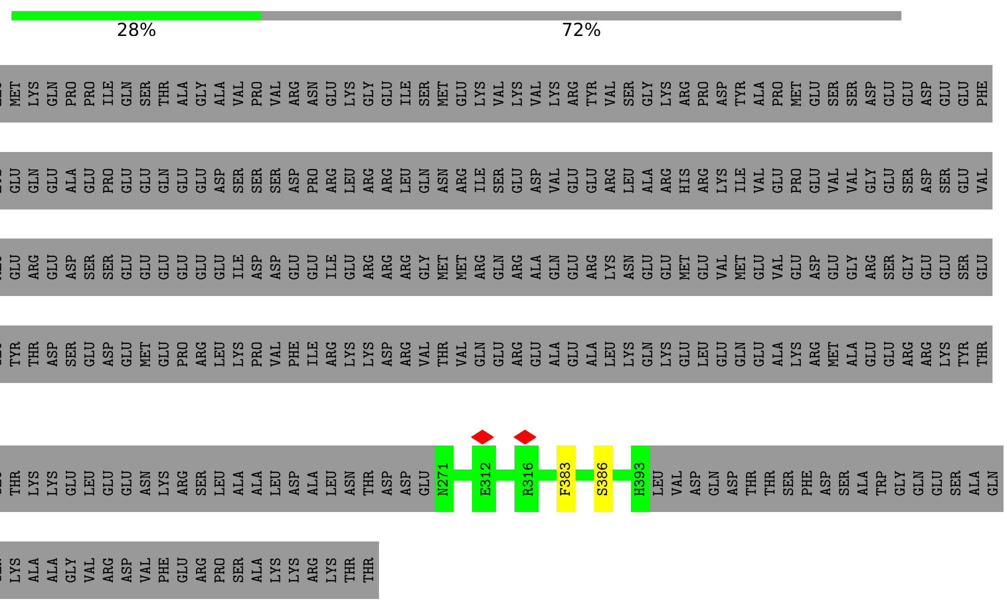
• Molecule 26: Pre-mRNA-splicing factor SYF1





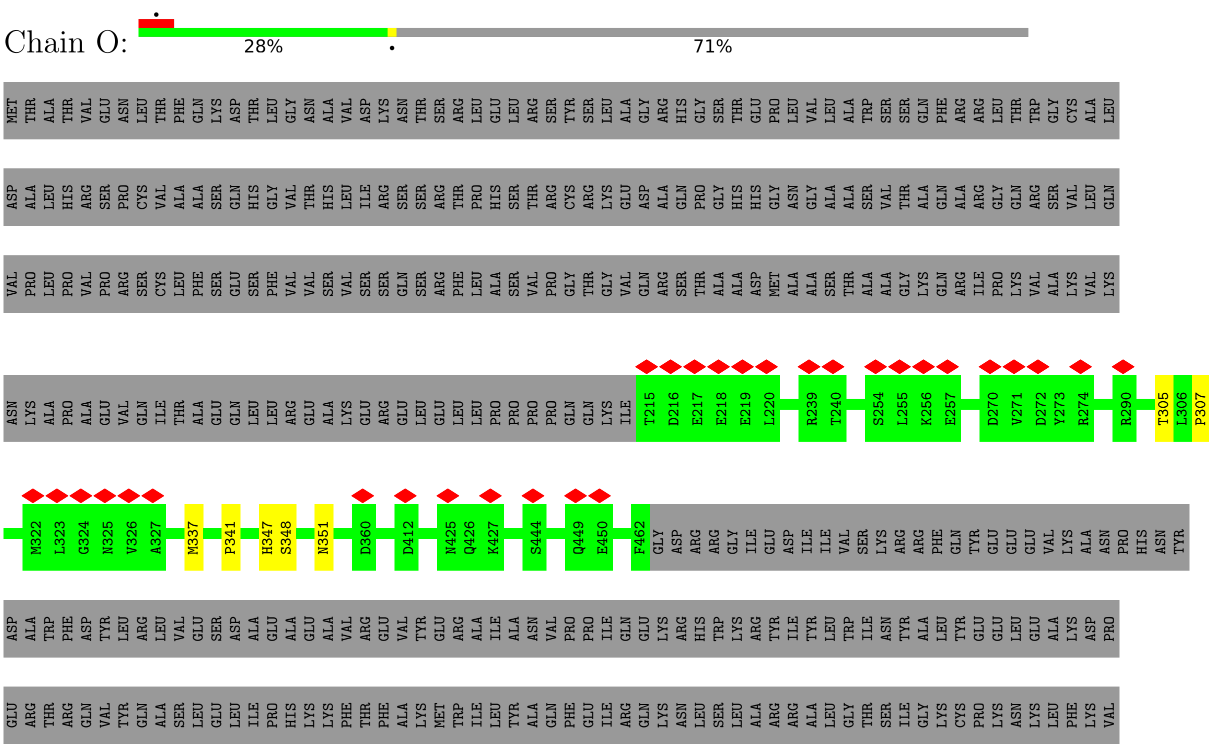
• Molecule 27: Microfibrillar-associated protein 1

Chain K:



• Molecule 28: Crooked neck-like protein 1

Chain O:



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

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

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

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

- Molecule 29: PHD finger-like domain-containing protein 5A

Chain y: 88% • 9%

MET A2 K3 H4 E101 ARG LYS LYS TYR GLY PHE LYS LYS ARG

- Molecule 30: U2 snRNA

Chain 2: 24% 41% 30% 15% • 13%

A U C3 C14 U15 U16 U17 U18 G19 G20 A23 A24 G25 A26 U27 C28 A29 A30 G31 U32 G33 U34 A35 G36 U37 A38 U39 C40 U41 G42 U43 U44 C45 U46 U47 A48 A51 U55 A56 A57 U58 A59 U60 C61 U62 G63 A64 U65 A C G68 U69 C70 C71 U72 C73

U A U C C G79 A80 G81 G82 A83 C84 A85 A U A U A U U A A A U G97 U101 U102 U103 G111 G112 G113 A114 G115 A116 U117 G118 G119 A120 A121 U122 A123 G124 G125 A126 G127 C128 U129 U130 G131 C132 U133 C134 C135 G136 U137 C138 C141 U142

C143 C144 A145 C146 G147 U150 C151 G152 A153 G157 G158 U159 A160 U161 U162 G163 C164 A165 G166 U167 A168 C169 C170 U171 C172 C173 A174 G175 G176 A177 G181 C184 A C C C

- Molecule 31: Pleiotropic regulator 1

Chain G: 55% 6% 38%

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

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

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

K181 M185 P186 K187 P188 L196 V249 G265 P298 Q339 A340 A341 E342 P343 Q344 H371 K372 K373 S374 H381 P382 R383 T386 P392 K396 K399 N408 I416 N422 S423 A431 H437 R449 Q455 P456 A467 A480 Y489

H500 PRO VAL SER TRP LYS PRO GLU ILE ILE LYS ARG LYS ARG PHE

• Molecule 32: Small nuclear ribonucleoprotein Sm D2

Chain h: 13% 81% 19%

MET SER LEU LEU ASN LYS PRO LYS SER GLU MET THR P13 E14 E15 K18 R19 E22 K37 N38 N39 T40 P78 LYS SER GLY LYS GLY LYS LYS LYS S87 V90 N91 K92 Y95 K98 I115 ALA GLY LYS

• Molecule 32: Small nuclear ribonucleoprotein Sm D2

Chain a: 66% 34%

MET SER LEU LEU ASN LYS PRO LYS SER GLU MET THR PRO GLU GLU LEU GLN LYS ARG GLU GLU GLU GLU PHE ASN T26 V77 PRO LYS SER GLY LYS GLY LYS LYS SER LYS P89 L114 ILE ALA GLY LYS

• Molecule 33: Peptidyl-prolyl cis-trans isomerase E

Chain o: 33% 74% 24%

M1 A2 T3 T4 K5 R6 Y9 V10 G11 G12 L13 A14 E15 E16 V17 K20 V21 L22 H23 A24 A25 F26 I27 P28 F29 G30 D31 I32 T33 D34 I35 P38 L39 D40 Y41 E42 T43 E44 K45 H46 R47 G48 F49 E55 L56 A57 N68 E69 S70 E71 L72 F73 G74 R75 T76

I77 R78 V79 N80 L81 A82 LYS PRO MET ARG ILE LYS GLU GLY SER SER ARG PRO VAL TRP SER ASP ASP ASP TRP LEU LYS LYS PHE SER GLY LYS THR LEU GLU GLU ASN LYS GLU GLU GLU GLY SER GLU PRO PRO LYS ALA GLU THR GLN GLU GLY GLU PRO ILE ALA LYS LYS ALA

ARG SER ASN PRO Q141 V142 L159 L160 ARG SER ASP VAL VAL P166 M167 E170 N171 F172 R173 C174 L175 C176 T177 H178 E179 K180 G181 F196 F203 G208 I214 Y215 K218 PHE ASP ASP GLU ASN PHE ILE LEU LYS H228 T229 G240 P241 N242 T243 K254 T255 D256

W257 L258 D259 G260 K261 T269 E270 S283 K284 D285 G286 K287 P288 K289 Q290 K291 E299 Y300 V301

• Molecule 34: RING-type E3 ubiquitin-protein ligase PPIL2

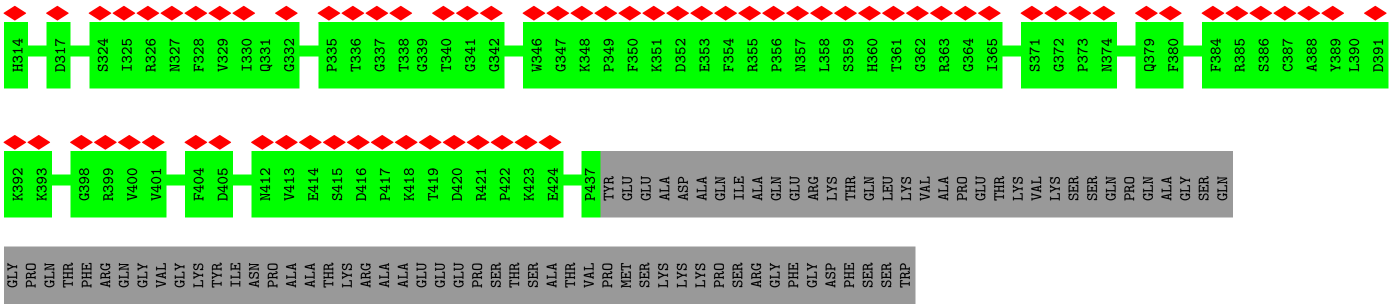
Chain t: 20% 48% 51%

MET GLY LYS ARG GLN HIS GLN LYS ASP LYS MET TYR ILE THR CYS ALA GLU TYR THR HIS PHE TYR GLY LYS LYS PRO ASP LEU PRO GLN THR ASN PHE ARG ARG LEU PRO PHE ASP HIS CYS S43 L44 S45 L46 Q47 P48 F49 V50 D57 V60 P68 K71 K72 Y73

P77 S78 N79 G80 R86 I89 N92 F93 SER LYS ASN SER GLU GLY LYS TYR HIS CYS PRO VAL LEU PHE THR VAL PHE THR ASN ASN THR HIS I116 G123 N135 I136 K137 A138 D143 Q159 ASP PRO THR ASN LEU ASP LYS PHE ASN VAL SER ASN PHE TYR

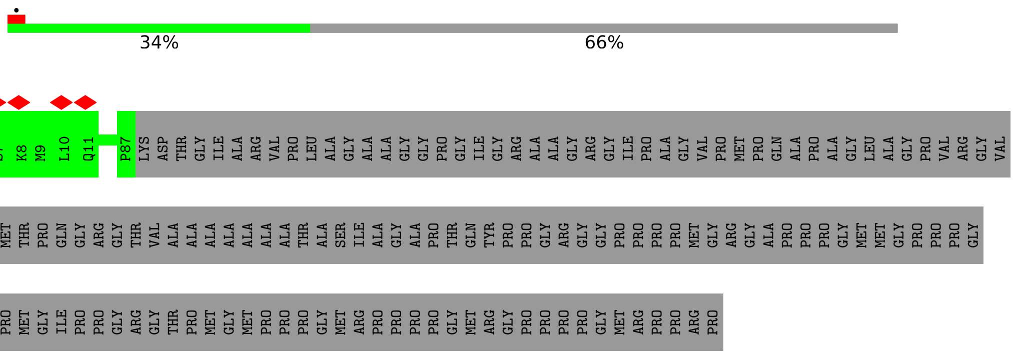
HIS VAL LYS ASN ASN MET LYS ILE ILE ASP PRO ASP GLU GLU LYS ALA LYS GLN ASP PRO SER TYR TYR LEU LYS ASN THR ASN ALA GLU THR ARG GLU THR LEU GLN GLU LEU TYR LYS GLU PHE LYS GLY ASP GLU ILE LEU ALA ALA THR MET LYS ALA PRO GLU LYS LYS LYS VAL

ASP LYS LEU ASN ALA ALA HIS TYR SER THR GLY LYS VAL SER ALA SER PHE THR SER THR ALA MET VAL PRO GLU THR THR HIS GLU ALA ALA ALA ILE ASP GLU ASP VAL LEU ARG TYR GLN PHE VAL LYS LYS G280 N287 K288 E294 L295 H296 C297 D298 L299 K312 K313



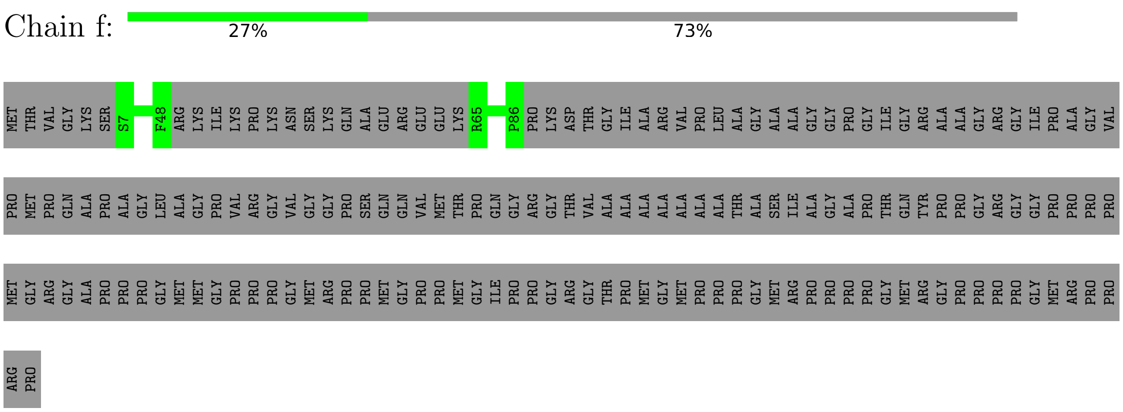
• Molecule 35: Small nuclear ribonucleoprotein-associated proteins B and B'

Chain m:



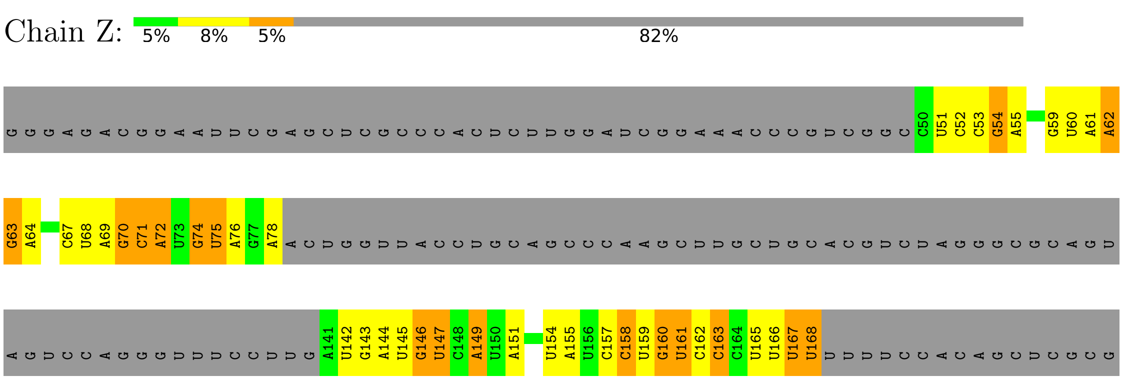
• Molecule 35: Small nuclear ribonucleoprotein-associated proteins B and B'

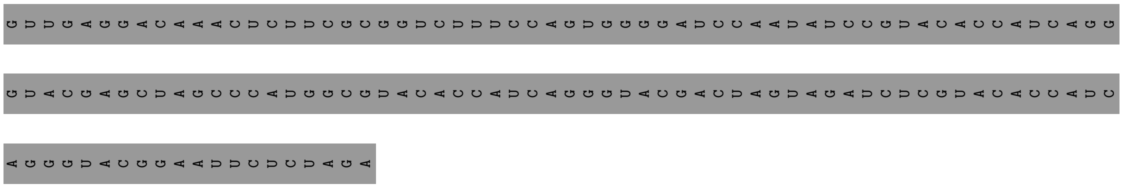
Chain f:



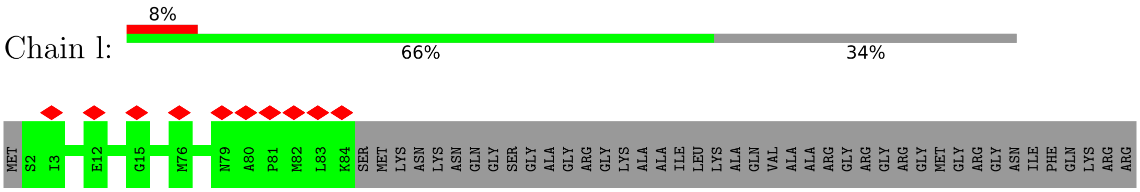
• Molecule 36: MINX M3 pre-mRNA

Chain Z:

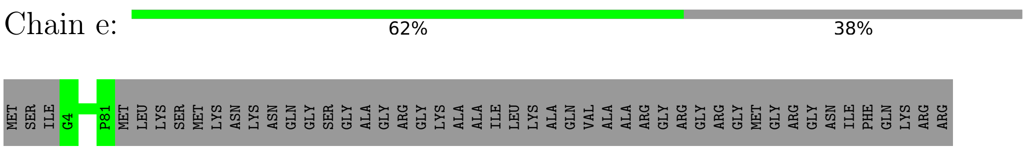




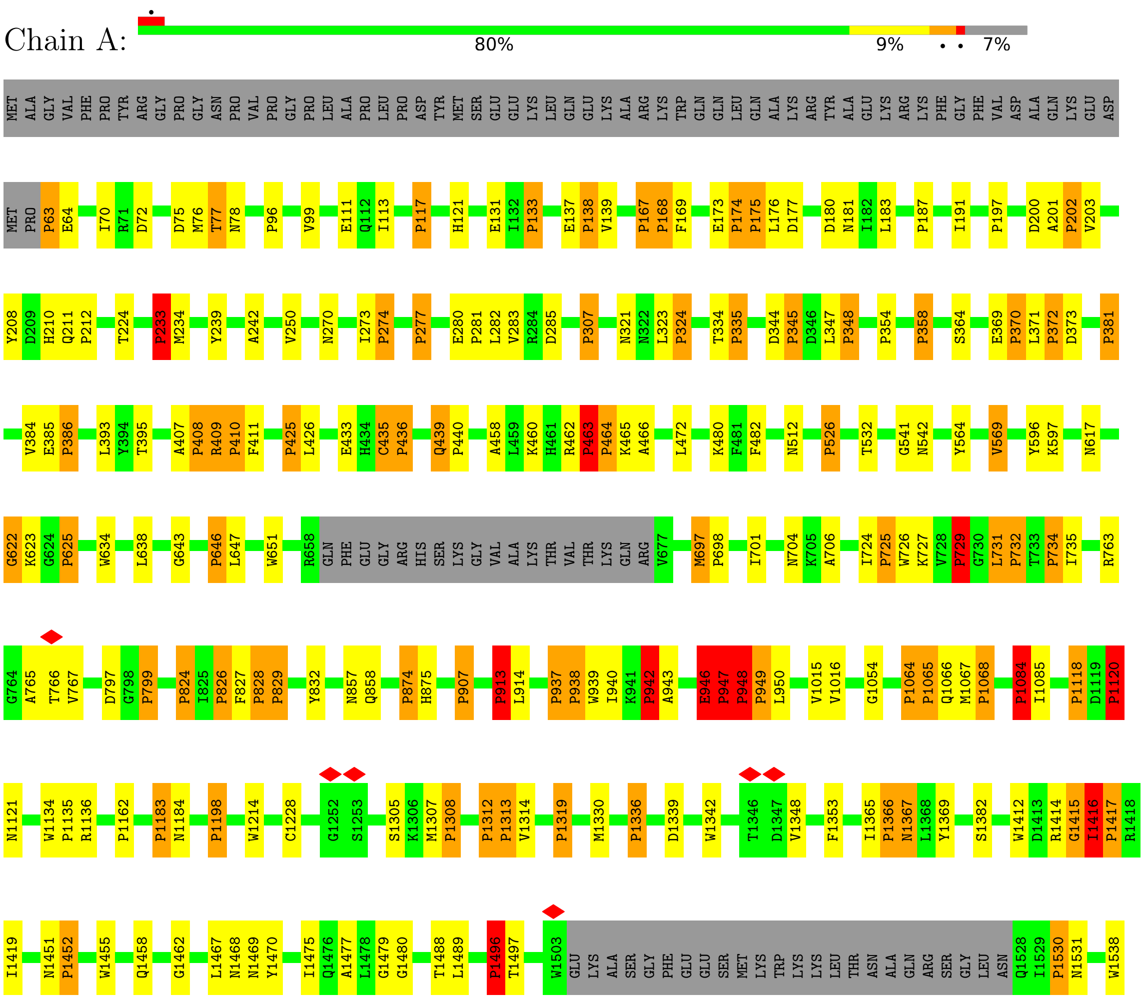
• Molecule 37: Small nuclear ribonucleoprotein Sm D3

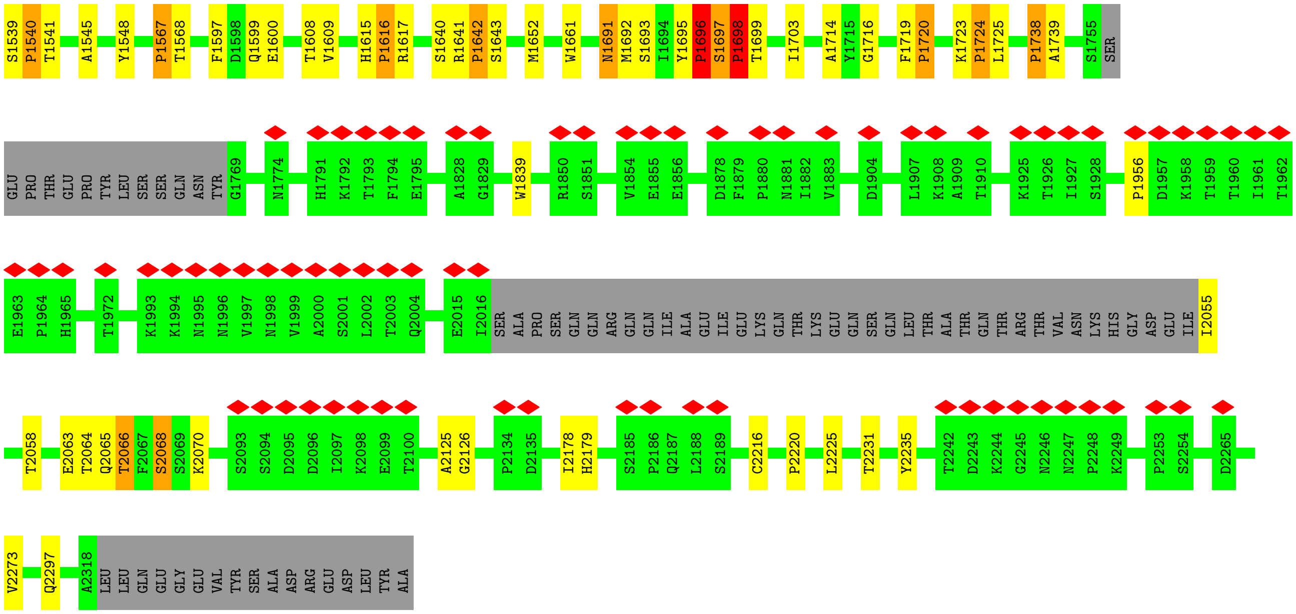


• Molecule 37: Small nuclear ribonucleoprotein Sm D3



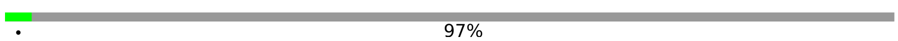
• Molecule 38: Pre-mRNA-processing-splicing factor 8

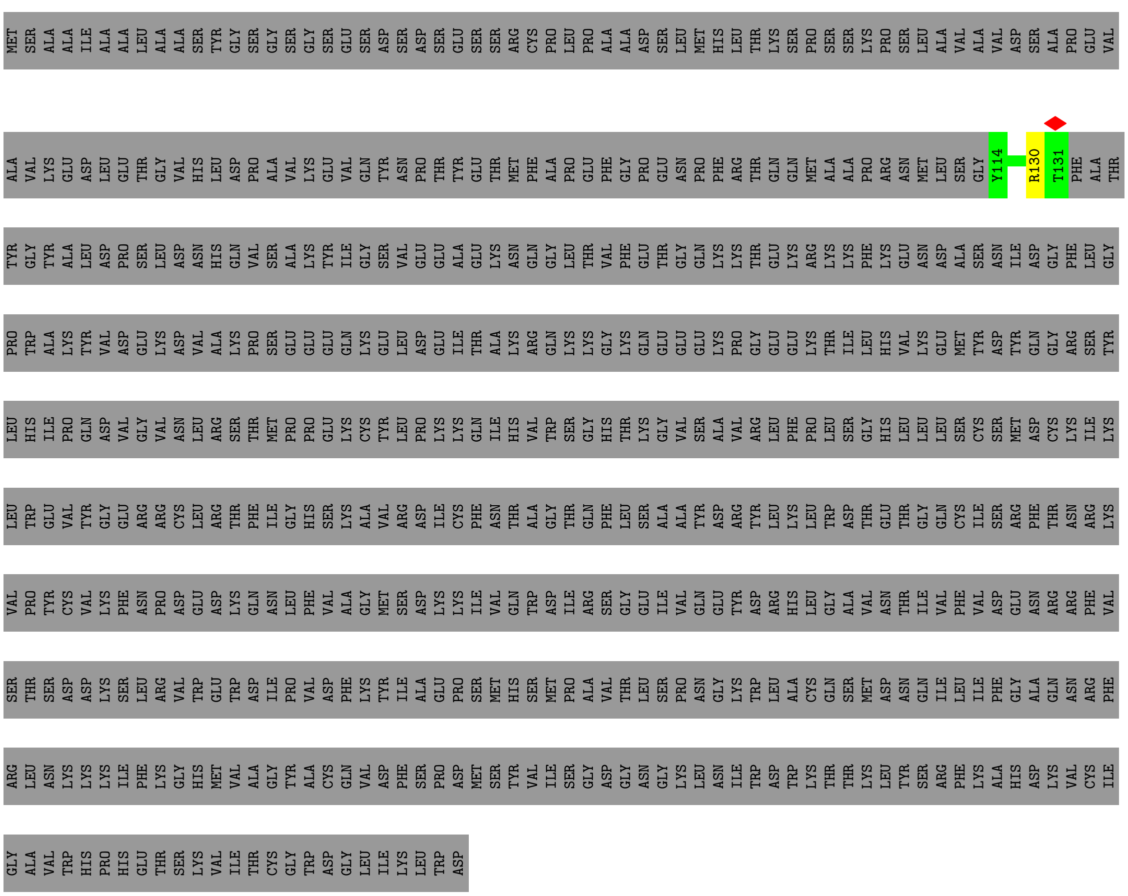




• Molecule 39: Pre-mRNA-processing factor 17

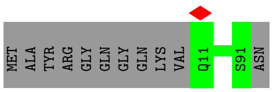
Chain 8:





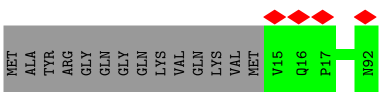
• Molecule 40: Small nuclear ribonucleoprotein E

Chain j:



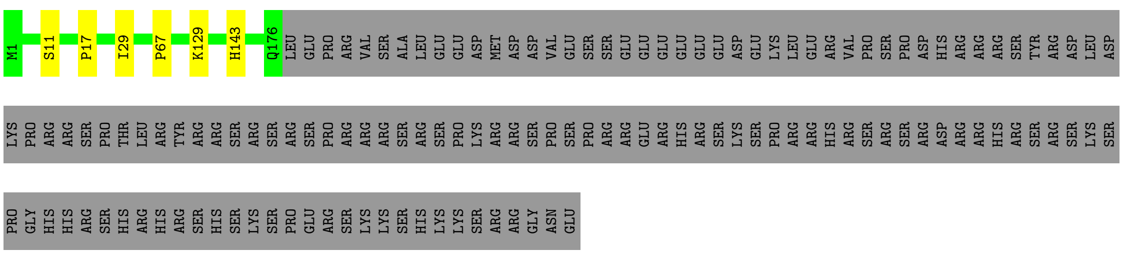
• Molecule 40: Small nuclear ribonucleoprotein E

Chain c:



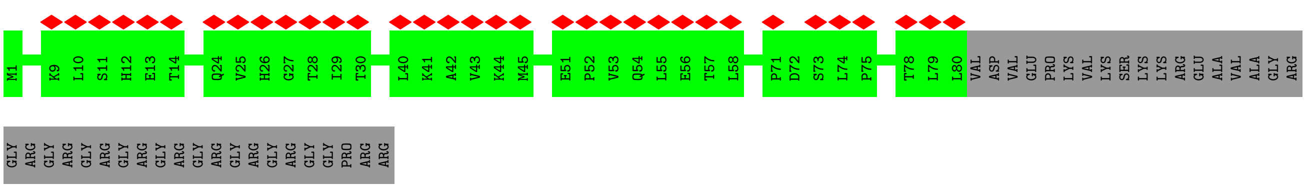
• Molecule 41: Pre-mRNA-splicing factor 38A

Chain I:



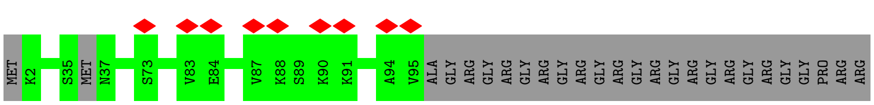
• Molecule 42: Small nuclear ribonucleoprotein Sm D1

Chain n:



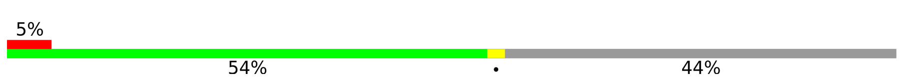
• Molecule 42: Small nuclear ribonucleoprotein Sm D1

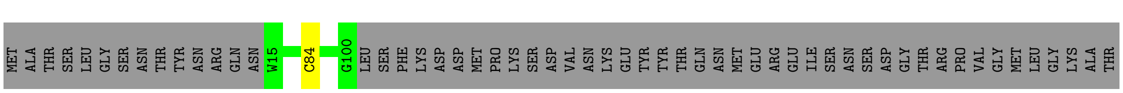
Chain g:



• Molecule 43: Pre-mRNA-splicing factor RBM22

Chain P:





SER THR SER D145 S166 F167 P180 D188 I198 M220 PRO ARG LEU ASP PRO PRO GLU ASP LYS THR ILE THR THR L234 G240 D241 T242 Q254 F255 R259 Q265 R266 A275 T276 R277 Q278 A279 A280 E281 V282 A283 I292 V293 N294 G295 R296 R297 R304

S305 Q306 A307 A308 ARG GLY LYS GLU LYS GLU LYS ASP GLY THR THR ASP SER GLY ILE LYS LEU GLU PRO VAL PRO GLY LEU PRO GLY ALA LEU PRO PRO PRO ALA ALA GLU GLU GLU ALA SER ALA ASN TYR PHE ASN LEU PRO PRO SER GLY PRO PRO ALA VAL VAL ASN ILE ALA

LEU PRO PRO PRO PRO GLY ILE ALA PRO PRO PRO PRO PRO GLY PHE GLY PRO HIS MET PHE HIS PRO MET GLY PRO PRO PRO PRO PHE MET ARG ALA PRO GLY PRO ILE HIS TYR PRO SER GLN ASP PRO GLN ARG MET GLY ALA HIS ALA GLY LYS HIS SER SER PRO

• Molecule 44: U5 snRNA

Chain 5: 28% 51% 19%

A U A3 C4 U5 C6 U7 G8 G9 U10 U11 U12 C13 U14 C15 U16 U17 C18 A19 G20 A21 U22 C23 G24 C25 A26 U27 A28 A29 A30 U31 C32 U33 U34 U35 C36 G37 C38 C39 U40 U41 U42 U43 A44 C45 U46 A47 A48 A49 G50 A51 U52 U53 U54 C55 C56 G57 U58 G59 G60

A61 G62 A63 G64 G65 C68 A69 U72 C73 U74 G75 A76 G77 U78 C79 U80 U81 A82 A83 C84 C85 A88 U89 U90 U91 U92 U93 U94 G95 A96 G97 C98 C99 U100 G107 G108 C109 A110 A111 A116

• Molecule 45: Splicing factor 3A subunit 1

Chain p: 9% 15% 84%

MET PRO ALA GLY PRO VAL GLN ALA VAL PRO PRO PRO PRO PRO VAL PRO THR GLU PRO LYS GLN PRO THR GLU GLU GLU ALA SER SER LYS GLU ASP SER ALA PRO SER LYS PRO VAL VAL GLY ILE ILE TYR PRO PRO P47 E48 V49 R50 K80 P86 N87 D88 P89 Y90 H91 A92

G104 K105 A106 Q107 S110 ALA ALA ILE PRO LYS VAL MET GLN GLN GLN GLN GLN THR THR GLN GLN GLN LEU PRO GLN LYS VAL GLN ALA GLN VAL ILE GLN GLU THR ILE VAL PRO LYS GLU PRO PRO PRO GLU PHE GLU PHE ILE ALA ASP PRO PRO SER ILE S160 A161 F162 D163

L164 D165 V166 V167 K168 L169 T170 A171 Q172 F173 V174 A175 R176 N177 G178 R179 Q180 F181 L182 T183 Q184 L185 M186 Q187 K188 E189 Q190 R191 N192 Y193 Q194 F195 D196 F197 L198 R199 P200 Q201 H202 S203 L204 F205 N206 Y207 F208 T209 K210 L211 V212 E213 Q214 Y215 T216 K217 I218 L219 ILE PRO PRO LYS

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

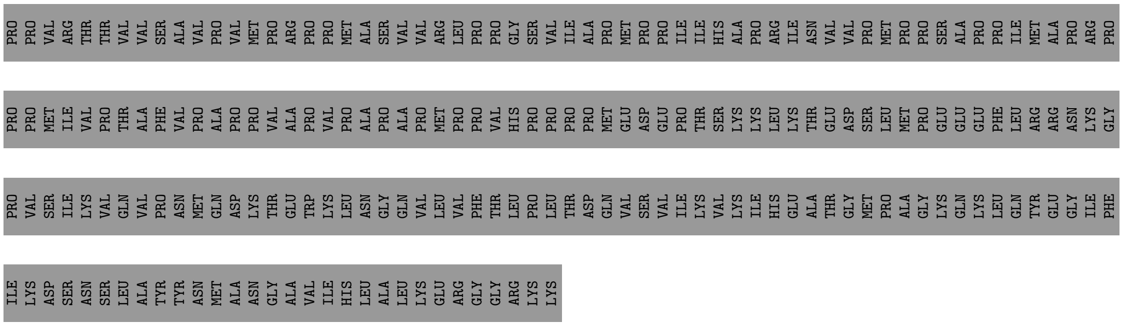
THR VAL ASP PHE GLN PRO ASN GLU GLN GLY ASN PHE PRO PRO PRO THR THR PRO GLU GLU LEU GLY ALA ARG ILE LEU ILE GLN GLU ARG TYR GLU LYS PHE GLY GLU SER GLU GLU VAL GLU MET GLU VAL GLU SER ASP GLU GLU ASP ASP LYS GLN GLU LYS ALA GLU GLU PRO PRO

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

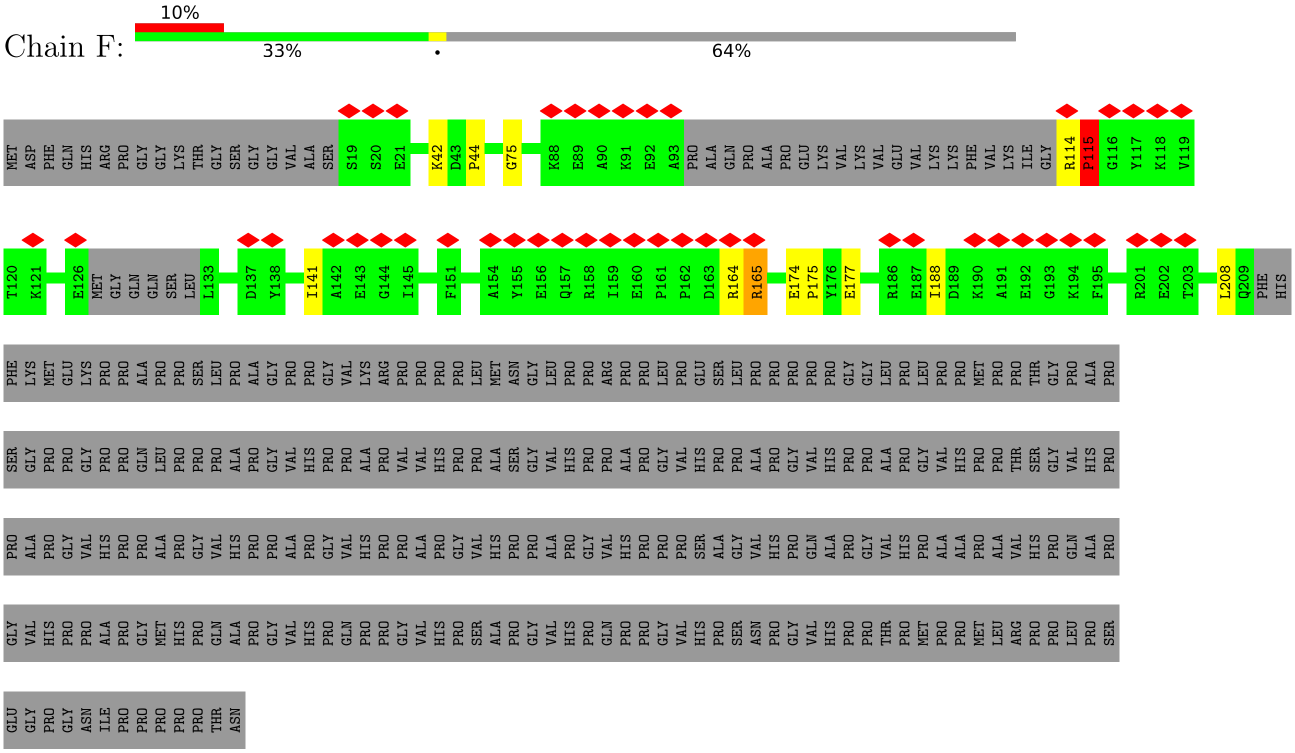
ALA PRO ALA PRO ASP GLU TYR LEU VAL SER PRO ILE THR GLY GLU LYS ILE PRO ALA SER LYS MET GLN GLU HIS MET ARG ILE GLY LEU LEU ASP PRO ARG TRP LEU GLU GLN ARG ASP ARG SER ILE ARG GLU LYS GLN SER ASP ASP GLU VAL TYR ALA PRO GLY LEU ASP ILE GLU

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

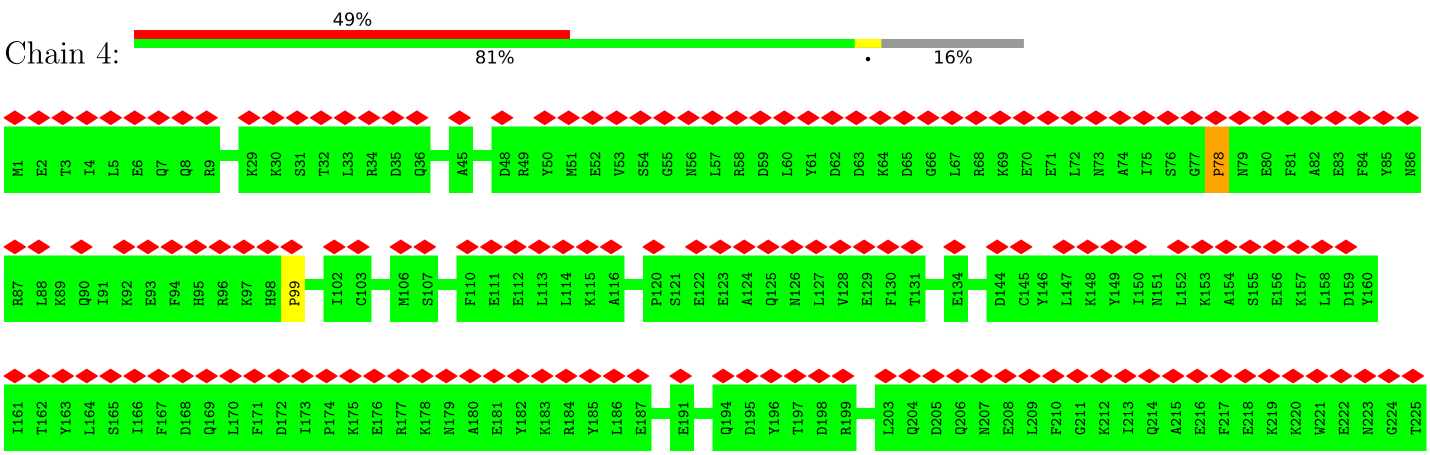
GLU GLN ILE GLU ALA ILE HIS LYS ALA LYS GLY LEU VAL PRO GLU ASP ASP THR LYS GLU LYS ILE GLY PRO SER LYS PRO ASN GLU ILE PRO GLN GLN PRO PRO PRO PRO SER SER ALA THR ASN ILE PRO SER SER ALA PRO PRO ILE THR SER VAL PRO ARG PRO PRO THR MET PRO

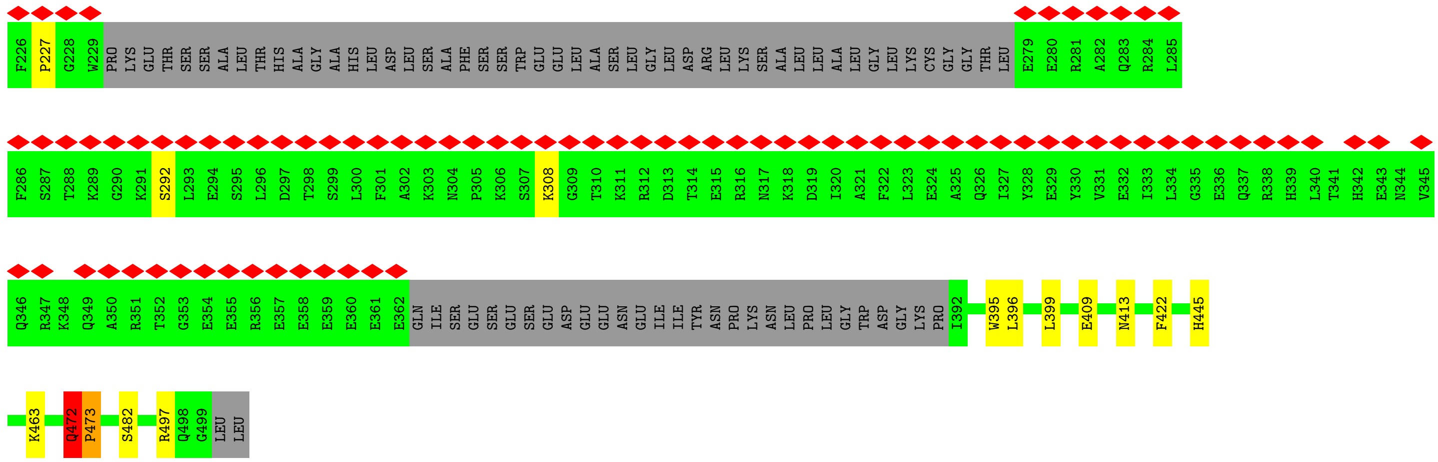


• Molecule 46: Splicing factor 3A subunit 2



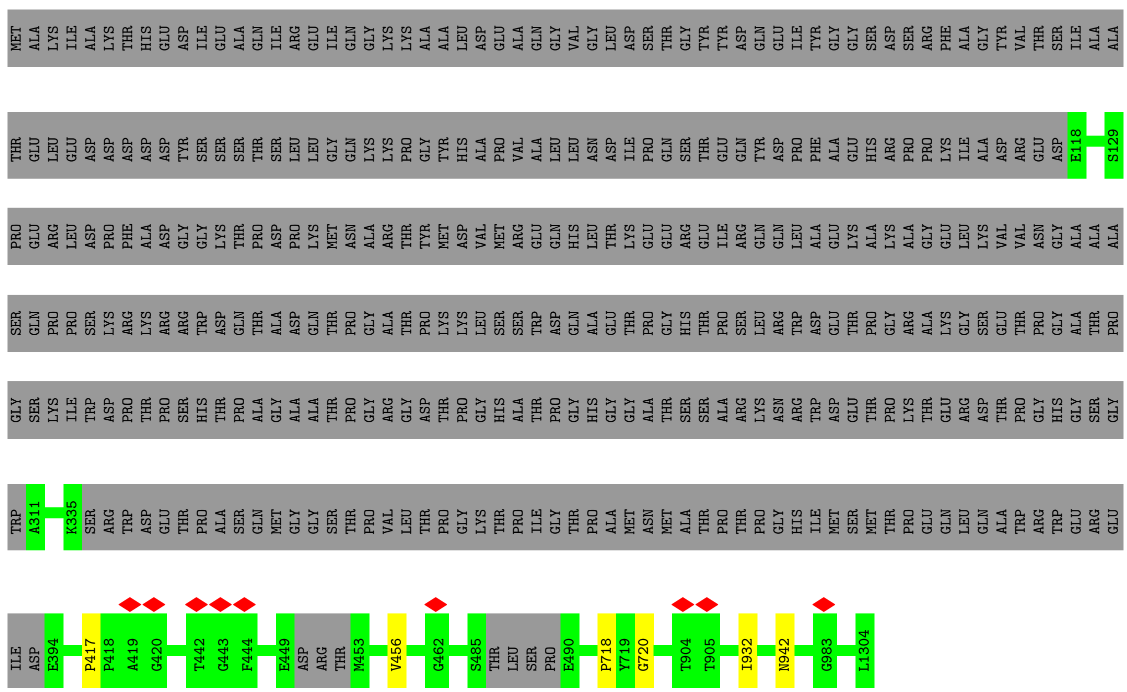
• Molecule 47: Splicing factor 3A subunit 3





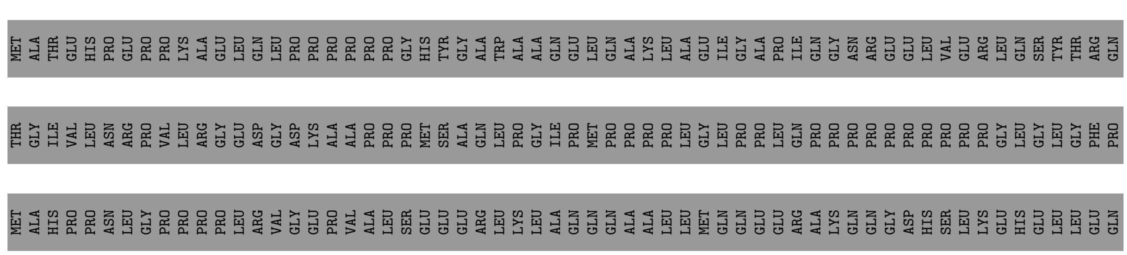
- Molecule 48: Splicing factor 3B subunit 1

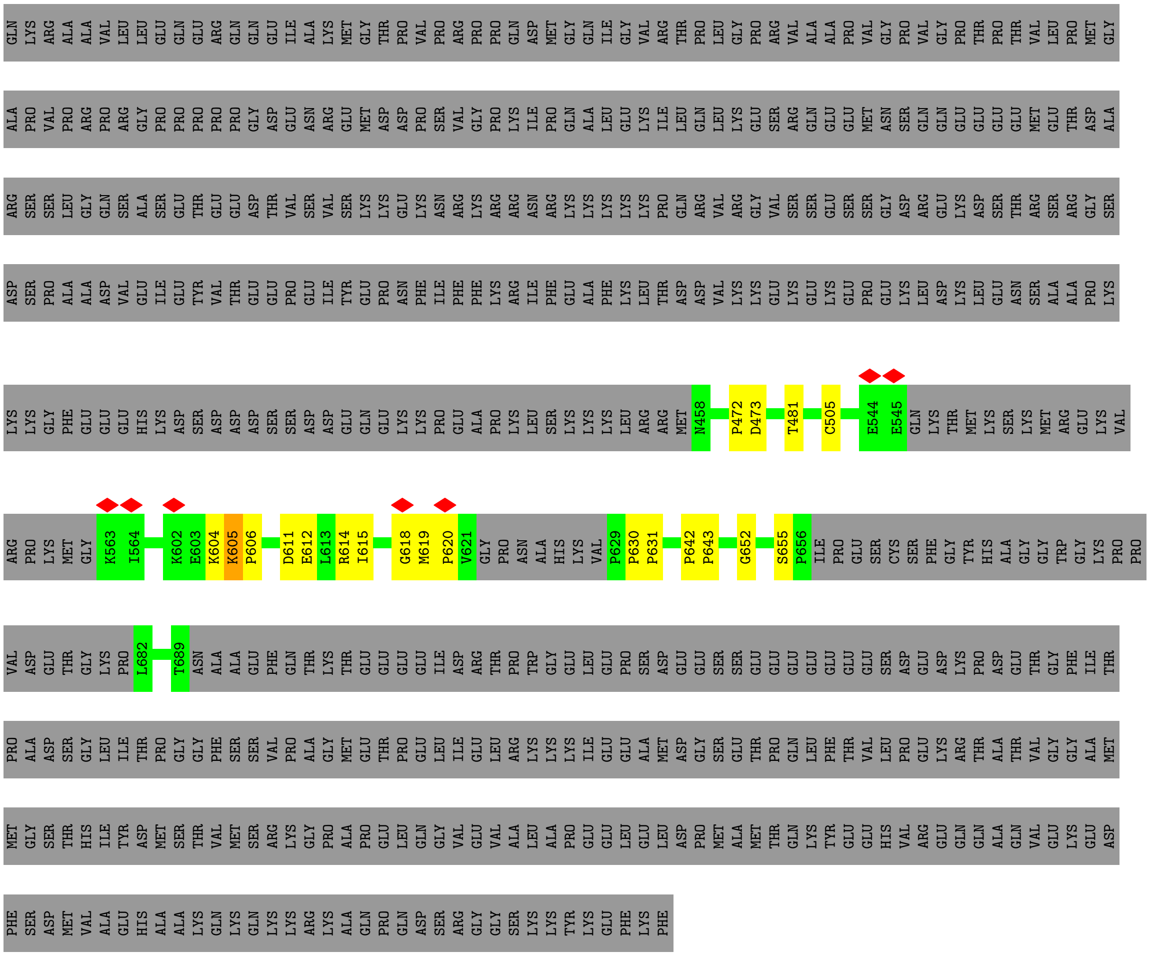
Chain u:



- Molecule 49: Splicing factor 3B subunit 2

Chain T:




GLN LYS ARG ALA ALA VAL LEU LEU GLU GLN GLU ARG GLN GLN GLU ILE ALA LYS MET GLY THR PRO VAL PRO ARG PRO PRO GLN ASP MET GLY GLN ILE GLY VAL ARG THR PRO LEU GLY PRO ARG VAL ALA ALA PRO VAL GLY PRO VAL GLY PRO THR PRO THR VAL LEU PRO MET GLY
ALA PRO VAL PRO ARG PRO ARG GLY PRO PRO PRO PRO PRO GLY ASP GLU ASN ARG GLU MET ASP ASP PRO SER VAL GLY PRO LYS ILE PRO GLN ALA LEU GLU LYS ILE LEU GLN LEU LYS GLU SER ARG GLN GLU GLU MET ASN SER GLN GLN GLU GLU GLU GLU MET GLU THR ASP ALA
ARG SER SER LEU GLY GLN SER ALA SER GLU THR GLU GLU ASP THR VAL SER VAL SER LYS LYS GLU LYS ASN ARG LYS ARG ARG ASN ARG LYS LYS LYS LYS LYS PRO GLN ARG VAL ARG GLY VAL SER SER GLU SER SER GLY ASP ARG GLU LYS ASP SER THR ARG SER ARG GLY SER
ASP SER PRO ALA ALA ASP VAL GLU ILE GLU TYR VAL THR GLU GLU PRO GLU ILE TYR GLU PRO ASN PHE ILE PHE PHE LYS ARG ILE PHE GLU ALA PHE LYS LEU THR ASP ASP VAL LYS LYS GLU LYS GLU LYS GLU PRO GLU LYS LEU ASP LYS LEU GLU ASN SER ALA ALA PRO LYS
LYS LYS GLY PHE GLU GLU GLU HIS LYS ASP SER ASP ASP ASP SER SER ASP ASP GLU GLN GLU LYS LYS PRO GLU ALA PRO LYS LEU SER LYS LYS LYS LEU ARG ARG MET N458 P472 D473 T481 C505 E544 E545 GLN LYS THR MET LYS SER LYS MET ARG GLU LYS VAL
ARG PRO LYS MET GLY K563 I564 K602 E603 K604 K605 P606 D611 E612 L613 R614 I615 G618 M619 P620 V621 GLY PRO ASN ALA HIS LYS VAL P629 P630 P631 P642 P643 G652 S655 P656 ILE PRO GLU SER CYS SER PHE GLY TYR HIS ALA GLY GLY TRP GLY LYS PRO PRO
VAL ASP GLU THR GLY LYS PRO L682 T689 ASN ALA ALA GLU PHE GLN THR LYS THR GLU GLU GLU GLU ILE ASP ARG THR PRO TRP GLY GLU LEU GLU PRO SER ASP GLU GLU SER SER GLU GLU GLU GLU GLU GLU GLU SER ASP GLU ASP LYS PRO ASP GLU THR GLY PHE ILE THR
PRO ALA ASP SER GLY LEU ILE THR PRO GLY GLY PHE SER SER VAL PRO ALA GLY MET GLU THR PRO GLU LEU ILE GLU LEU ARG LYS LYS ILE GLU GLU ALA MET ASP GLY SER GLU THR PRO GLN LEU PHE THR VAL LEU PRO GLU LYS ARG THR ALA THR VAL GLY GLY ALA MET
MET GLY SER THR HIS ILE TYR ASP MET SER THR VAL MET SER ARG LYS GLY PRO ALA PRO GLU LEU GLN GLY VAL GLU VAL ALA LEU ALA PRO GLU GLU LEU GLU LEU ASP PRO MET ALA MET THR GLN LYS TYR GLU GLU HIS VAL ARG GLU GLN GLN ALA GLN VAL GLU LYS GLU ASP
PHE SER ASP MET VAL ALA GLU HIS ALA ALA LYS GLN LYS GLN LYS LYS ARG LYS ALA GLN PRO GLN ASP SER ARG GLY GLY SER LYS LYS TYR LYS GLU PHE LYS PHE

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	39336	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{Å}^2$)	2.27	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	FEI FALCON III (4k x 4k)	Depositor
Maximum map value	0.055	Depositor
Minimum map value	-0.015	Depositor
Average map value	0.001	Depositor
Map value standard deviation	0.004	Depositor
Recommended contour level	0.013	Depositor
Map size (Å)	445.44, 445.44, 445.44	wwPDB
Map dimensions	384, 384, 384	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.16, 1.16, 1.16	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: MG, IHP, GTP

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	D	0.25	0/1515	0.48	0/2113
2	E	0.31	0/5980	0.49	0/8363
3	b	0.26	0/367	0.52	0/509
3	i	0.36	0/362	0.55	0/502
4	W	0.27	0/821	0.53	0/1149
5	w	0.65	0/394	0.68	0/549
6	d	0.27	0/346	0.54	0/481
6	k	0.39	0/366	0.60	0/509
7	U	0.29	1/6796 (0.0%)	0.47	3/9527 (0.0%)
8	x	0.49	0/399	0.54	0/557
9	6	0.44	1/2178 (0.0%)	0.91	7/3388 (0.2%)
10	B	0.67	0/857	0.82	1/1196 (0.1%)
11	z	0.26	0/548	0.49	0/766
12	q	0.37	0/359	0.57	0/498
13	s	0.24	0/8766	0.44	0/12286
14	v	0.23	0/1064	0.44	0/1491
15	3	0.25	0/288	0.46	0/404
16	0	0.27	0/770	0.50	0/1079
17	Q	0.44	1/700 (0.1%)	0.52	0/979
18	1	0.30	0/609	0.48	0/848
19	L	0.25	0/583	0.47	0/813
20	N	0.22	0/276	0.42	0/383
21	S	0.28	0/1778	0.40	0/2482
22	r	0.31	0/4585	0.53	3/6429 (0.0%)
23	R	0.24	0/218	0.48	0/303
24	Y	0.23	0/481	0.77	2/672 (0.3%)
25	7	0.86	0/1181	1.04	0/1645
26	M	0.28	0/3576	0.46	2/5009 (0.0%)
27	K	0.24	0/615	0.42	0/858
28	O	0.23	0/1245	0.45	0/1742
29	y	0.26	0/501	0.60	1/697 (0.1%)
30	2	0.44	5/3882 (0.1%)	1.04	28/6040 (0.5%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
31	G	0.32	0/1616	0.61	3/2258 (0.1%)
32	a	0.24	0/394	0.46	0/548
32	h	0.33	0/485	0.54	0/677
33	o	0.48	0/1139	0.61	0/1578
34	t	0.51	0/1283	0.88	3/1790 (0.2%)
35	f	0.25	0/319	0.46	0/442
35	m	0.36	0/416	0.55	0/581
36	Z	0.46	1/1338 (0.1%)	0.87	0/2077
37	e	0.25	0/392	0.50	0/546
37	l	0.39	0/417	0.60	0/581
38	A	0.51	3/11086 (0.0%)	0.82	83/15554 (0.5%)
39	8	0.31	0/92	0.52	0/128
40	c	0.23	0/388	0.50	0/540
40	j	0.36	0/403	0.55	0/561
41	I	0.24	0/888	0.43	0/1241
42	g	0.24	0/471	0.48	0/657
42	n	0.34	0/404	0.55	0/564
43	P	0.43	1/1203 (0.1%)	0.67	0/1680
44	5	0.23	0/2672	0.79	5/4154 (0.1%)
45	p	0.43	0/636	0.82	0/890
46	F	0.43	0/840	0.79	1/1174 (0.1%)
47	4	0.52	0/2119	1.02	2/2960 (0.1%)
48	u	0.32	0/4792	0.50	0/6707
49	T	0.56	3/957 (0.3%)	0.64	2/1341 (0.1%)
All	All	0.38	16/87156 (0.0%)	0.66	146/123496 (0.1%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
38	A	1366	PRO	N-CA	13.70	1.70	1.47
43	P	188	ASP	C-N	8.60	1.50	1.34
49	T	605	LYS	C-N	8.48	1.50	1.34
38	A	1416	ILE	C-N	8.47	1.50	1.34
7	U	493	LYS	C-N	8.46	1.50	1.34

The worst 5 of 146 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
47	4	472	GLN	C-N-CD	-16.12	85.13	120.60
47	4	422	PHE	N-CA-CB	-15.71	82.32	110.60
24	Y	75	LEU	N-CA-CB	-14.01	82.37	110.40
9	6	52	U	C2'-C3'-O3'	10.62	132.86	109.50

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
38	A	372	PRO	CA-N-CD	-9.45	98.27	111.50

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	1506	0	737	5	0
2	E	5926	0	2964	82	0
3	b	364	0	181	0	0
3	i	359	0	179	0	0
4	W	816	0	386	2	0
5	w	391	0	197	0	0
6	d	344	0	168	0	0
6	k	364	0	176	0	0
7	U	6730	0	3268	110	0
8	x	397	0	191	0	0
9	6	1947	0	984	173	0
10	B	851	0	423	9	0
11	z	544	0	264	0	0
12	q	360	0	159	0	0
13	s	8688	0	4221	0	0
14	v	1053	0	550	0	0
15	3	283	0	158	23	0
16	0	761	0	375	13	0
17	Q	695	0	327	31	0
18	1	607	0	277	21	0
19	L	582	0	288	6	0
20	N	277	0	114	3	0
21	S	1771	0	820	8	0
22	r	4531	0	2256	0	0
23	R	219	0	91	3	0
24	Y	478	0	226	9	0
25	7	1178	0	537	8	0
26	M	3553	0	1739	40	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
27	K	614	0	284	5	0
28	O	1243	0	570	12	0
29	y	498	0	241	0	0
30	2	3479	0	1754	97	0
31	G	1604	0	795	37	0
32	a	393	0	176	0	0
32	h	482	0	220	0	0
33	o	1135	0	572	0	0
34	t	1272	0	618	0	0
35	f	319	0	144	0	0
35	m	413	0	193	0	0
36	Z	1201	0	607	111	0
37	e	390	0	188	0	0
37	l	415	0	198	0	0
38	A	10968	0	5356	428	0
39	8	92	0	46	1	0
40	c	388	0	167	0	0
40	j	403	0	173	0	0
41	I	883	0	414	3	0
42	g	469	0	214	0	0
42	n	402	0	184	0	0
43	P	1194	0	589	12	0
44	5	2397	0	1214	107	0
45	p	631	0	309	0	0
46	F	835	0	412	14	0
47	4	2110	0	987	25	0
48	u	4748	0	2402	0	0
49	T	942	0	490	18	0
50	r	32	0	12	0	0
51	r	1	0	0	0	0
52	A	36	0	6	0	0
All	All	85564	0	41791	1196	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 13.

The worst 5 of 1196 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)
38:A:270:ASN:CB	38:A:282:LEU:HA	1.33	1.57
9:6:28:A:H5”	17:Q:41:ARG:CA	1.36	1.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:E:407:ILE:CB	2:E:1122:LEU:CB	1.81	1.53
38:A:466:ALA:CB	44:5:20:G:C6	1.93	1.50
38:A:948:PRO:HD2	38:A:949:PRO:CD	1.40	1.50

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	300/357 (84%)	292 (97%)	8 (3%)	0	100	100
2	E	1165/1217 (96%)	1126 (97%)	27 (2%)	12 (1%)	15	55
3	b	71/86 (83%)	71 (100%)	0	0	100	100
3	i	70/86 (81%)	69 (99%)	1 (1%)	0	100	100
4	W	160/255 (63%)	147 (92%)	13 (8%)	0	100	100
5	w	76/424 (18%)	68 (90%)	7 (9%)	1 (1%)	12	48
6	d	67/76 (88%)	65 (97%)	2 (3%)	0	100	100
6	k	71/76 (93%)	69 (97%)	2 (3%)	0	100	100
7	U	1319/1485 (89%)	1273 (96%)	36 (3%)	10 (1%)	19	60
8	x	77/86 (90%)	66 (86%)	10 (13%)	1 (1%)	12	48
10	B	165/225 (73%)	158 (96%)	4 (2%)	3 (2%)	8	40
11	z	106/125 (85%)	90 (85%)	16 (15%)	0	100	100
12	q	71/73 (97%)	64 (90%)	7 (10%)	0	100	100
13	s	1720/2136 (80%)	1691 (98%)	29 (2%)	0	100	100
14	v	198/536 (37%)	181 (91%)	17 (9%)	0	100	100
15	3	53/619 (9%)	48 (91%)	5 (9%)	0	100	100
16	0	148/396 (37%)	145 (98%)	3 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
17	Q	136/144 (94%)	127 (93%)	8 (6%)	1 (1%)	22	63
18	1	120/322 (37%)	111 (92%)	6 (5%)	3 (2%)	5	32
19	L	112/802 (14%)	108 (96%)	4 (4%)	0	100	100
20	N	54/199 (27%)	47 (87%)	7 (13%)	0	100	100
21	S	353/563 (63%)	344 (98%)	8 (2%)	1 (0%)	41	77
22	r	893/972 (92%)	813 (91%)	76 (8%)	4 (0%)	34	72
23	R	42/229 (18%)	38 (90%)	4 (10%)	0	100	100
24	Y	93/904 (10%)	92 (99%)	1 (1%)	0	100	100
25	7	230/393 (58%)	225 (98%)	4 (2%)	1 (0%)	34	72
26	M	704/855 (82%)	597 (85%)	96 (14%)	11 (2%)	9	44
27	K	121/439 (28%)	96 (79%)	25 (21%)	0	100	100
28	O	246/848 (29%)	234 (95%)	12 (5%)	0	100	100
29	y	98/110 (89%)	93 (95%)	3 (3%)	2 (2%)	7	38
31	G	318/514 (62%)	292 (92%)	21 (7%)	5 (2%)	9	44
32	a	74/118 (63%)	73 (99%)	1 (1%)	0	100	100
32	h	91/118 (77%)	86 (94%)	5 (6%)	0	100	100
33	o	221/301 (73%)	196 (89%)	19 (9%)	6 (3%)	5	31
34	t	247/520 (48%)	234 (95%)	12 (5%)	1 (0%)	34	72
35	f	60/240 (25%)	60 (100%)	0	0	100	100
35	m	80/240 (33%)	74 (92%)	6 (8%)	0	100	100
37	e	76/126 (60%)	76 (100%)	0	0	100	100
37	l	81/126 (64%)	76 (94%)	5 (6%)	0	100	100
38	A	2153/2335 (92%)	1791 (83%)	296 (14%)	66 (3%)	4	27
39	8	16/579 (3%)	15 (94%)	1 (6%)	0	100	100
40	c	76/92 (83%)	75 (99%)	1 (1%)	0	100	100
40	j	79/92 (86%)	77 (98%)	2 (2%)	0	100	100
41	I	174/312 (56%)	166 (95%)	8 (5%)	0	100	100
42	g	89/119 (75%)	88 (99%)	1 (1%)	0	100	100
42	n	78/119 (66%)	75 (96%)	3 (4%)	0	100	100
43	P	231/420 (55%)	191 (83%)	39 (17%)	1 (0%)	34	72
45	p	120/793 (15%)	115 (96%)	3 (2%)	2 (2%)	9	42

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
46	F	159/464 (34%)	145 (91%)	10 (6%)	4 (2%)	5	32
47	4	415/501 (83%)	370 (89%)	36 (9%)	9 (2%)	6	35
48	u	931/1304 (71%)	822 (88%)	103 (11%)	6 (1%)	25	66
49	T	175/895 (20%)	159 (91%)	15 (9%)	1 (1%)	25	66
All	All	14983/25366 (59%)	13804 (92%)	1028 (7%)	151 (1%)	20	55

5 of 151 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	E	405	SER
2	E	914	ILE
7	U	744	ILE
7	U	980	PRO
7	U	1338	PHE

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	10/300 (3%)	10 (100%)	0	100	100
2	E	60/1051 (6%)	60 (100%)	0	100	100
3	b	4/74 (5%)	4 (100%)	0	100	100
3	i	4/74 (5%)	4 (100%)	0	100	100
4	W	6/218 (3%)	6 (100%)	0	100	100
5	w	4/336 (1%)	4 (100%)	0	100	100
6	d	3/66 (4%)	3 (100%)	0	100	100
6	k	3/66 (4%)	3 (100%)	0	100	100
7	U	71/1336 (5%)	70 (99%)	1 (1%)	67	80
8	x	3/77 (4%)	3 (100%)	0	100	100
10	B	8/195 (4%)	8 (100%)	0	100	100
11	z	5/109 (5%)	5 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
13	s	79/1908 (4%)	79 (100%)	0	100	100
14	v	15/459 (3%)	15 (100%)	0	100	100
15	3	6/545 (1%)	6 (100%)	0	100	100
16	0	10/349 (3%)	10 (100%)	0	100	100
17	Q	6/130 (5%)	6 (100%)	0	100	100
18	1	3/291 (1%)	3 (100%)	0	100	100
19	L	3/709 (0%)	3 (100%)	0	100	100
21	S	8/510 (2%)	8 (100%)	0	100	100
22	r	55/866 (6%)	55 (100%)	0	100	100
24	Y	4/831 (0%)	4 (100%)	0	100	100
25	7	6/354 (2%)	6 (100%)	0	100	100
26	M	24/749 (3%)	24 (100%)	0	100	100
27	K	2/395 (0%)	2 (100%)	0	100	100
28	O	3/751 (0%)	3 (100%)	0	100	100
29	y	4/95 (4%)	4 (100%)	0	100	100
31	G	13/441 (3%)	13 (100%)	0	100	100
32	a	3/110 (3%)	3 (100%)	0	100	100
32	h	5/110 (4%)	5 (100%)	0	100	100
33	o	8/252 (3%)	8 (100%)	0	100	100
34	t	14/456 (3%)	14 (100%)	0	100	100
35	f	2/177 (1%)	2 (100%)	0	100	100
35	m	4/177 (2%)	4 (100%)	0	100	100
37	e	3/101 (3%)	3 (100%)	0	100	100
37	l	3/101 (3%)	3 (100%)	0	100	100
38	A	123/2108 (6%)	123 (100%)	0	100	100
39	8	1/502 (0%)	1 (100%)	0	100	100
40	c	1/84 (1%)	1 (100%)	0	100	100
40	j	1/84 (1%)	1 (100%)	0	100	100
41	I	6/293 (2%)	6 (100%)	0	100	100
42	g	4/101 (4%)	4 (100%)	0	100	100
42	n	3/101 (3%)	3 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
43	P	12/361 (3%)	12 (100%)	0	100	100
45	p	7/709 (1%)	7 (100%)	0	100	100
46	F	8/382 (2%)	8 (100%)	0	100	100
47	4	12/446 (3%)	11 (92%)	1 (8%)	11	34
48	u	49/1104 (4%)	49 (100%)	0	100	100
49	T	19/776 (2%)	19 (100%)	0	100	100
All	All	710/21820 (3%)	708 (100%)	2 (0%)	92	95

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	U	980	PRO
47	4	78	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
30	2	161/188 (85%)	40 (24%)	3 (1%)
36	Z	55/324 (16%)	26 (47%)	5 (9%)
44	5	113/116 (97%)	39 (34%)	3 (2%)
9	6	88/106 (83%)	38 (43%)	13 (14%)
All	All	417/734 (56%)	143 (34%)	24 (5%)

5 of 143 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
9	6	6	C
9	6	7	G
9	6	8	C
9	6	10	U
9	6	12	G

5 of 24 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
30	2	46	U
36	Z	74	G
36	Z	70	G
36	Z	160	G
9	6	46	G

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 1 is 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	$\#\|Z\| > 2$
52	IHP	A	3001	-	36,36,36	1.53	6 (16%)	54,60,60	0.70	1 (1%)
50	GTP	r	1500	22,51	26,34,34	1.10	1 (3%)	32,54,54	1.86	7 (21%)

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
52	IHP	A	3001	-	-	7/30/54/54	0/1/1/1
50	GTP	r	1500	22,51	-	6/18/38/38	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)
52	A	3001	IHP	P2-O12	3.61	1.66	1.59
50	r	1500	GTP	C6-N1	-3.56	1.32	1.37
52	A	3001	IHP	P3-O13	3.30	1.65	1.59
52	A	3001	IHP	P1-O11	3.24	1.65	1.59
52	A	3001	IHP	P5-O15	3.14	1.65	1.59

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
50	r	1500	GTP	C5-C6-N1	4.18	121.33	113.95
50	r	1500	GTP	O6-C6-C5	-4.10	116.36	124.37
50	r	1500	GTP	PA-O3A-PB	-3.91	119.43	132.83
50	r	1500	GTP	PB-O3B-PG	-3.62	120.40	132.83
50	r	1500	GTP	C2-N1-C6	-2.99	119.59	125.10

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

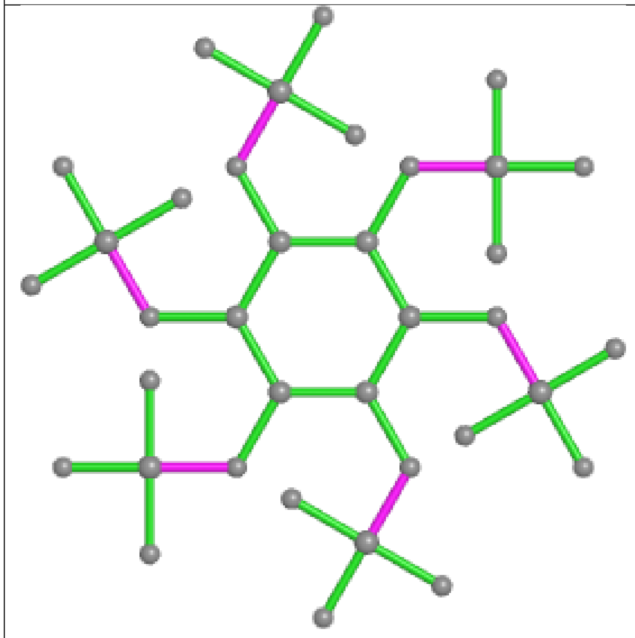
Mol	Chain	Res	Type	Atoms
50	r	1500	GTP	C5'-O5'-PA-O3A
50	r	1500	GTP	C5'-O5'-PA-O1A
50	r	1500	GTP	C5'-O5'-PA-O2A
50	r	1500	GTP	O4'-C4'-C5'-O5'
52	A	3001	IHP	C1-C2-O12-P2

There are no ring outliers.

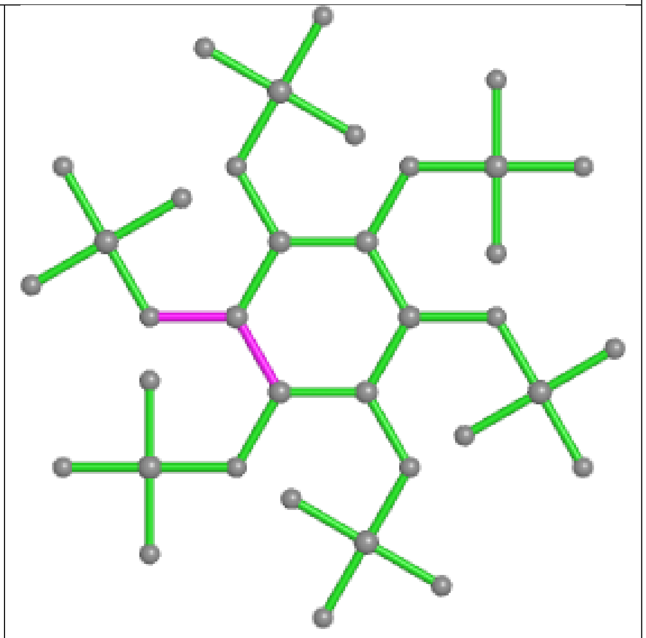
No monomer is involved in short contacts.

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.

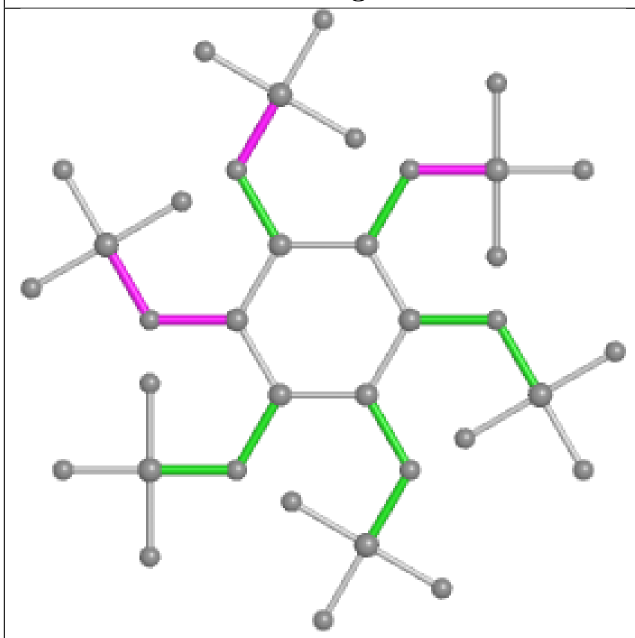
Ligand IHP A 3001



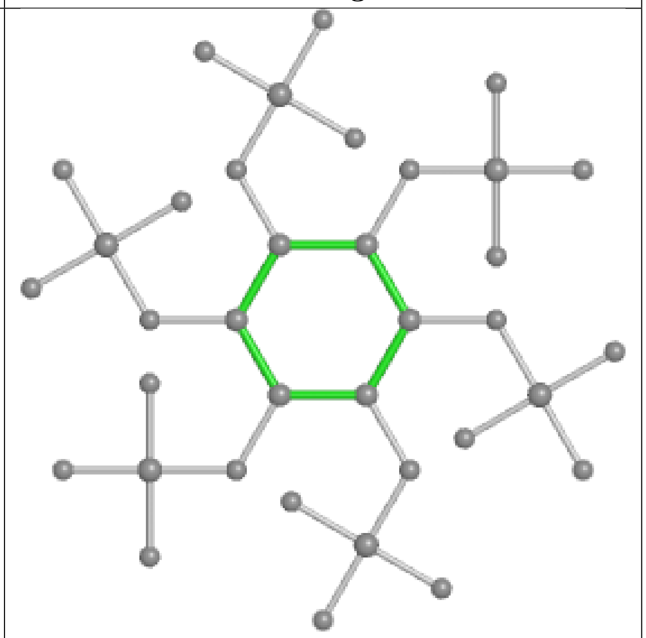
Bond lengths



Bond angles

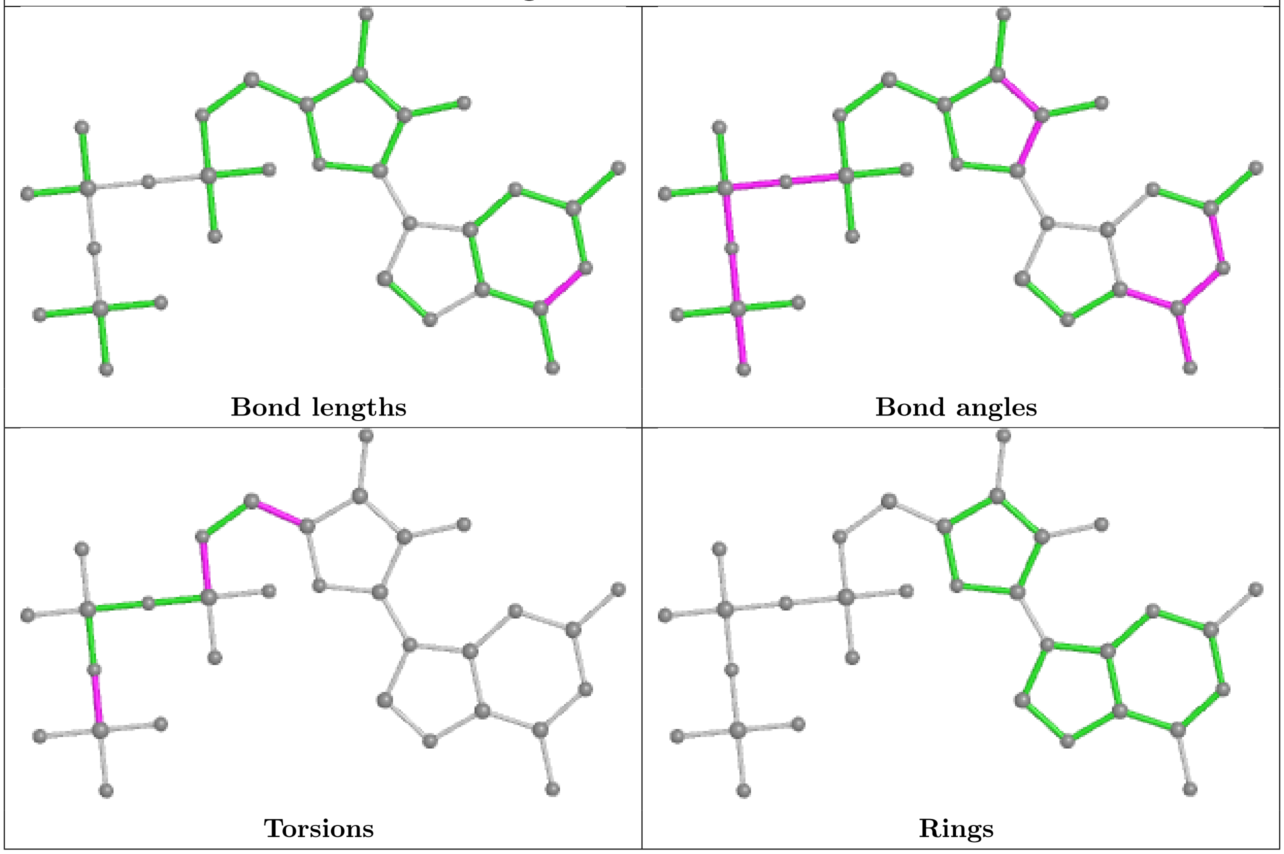


Torsions



Rings

Ligand GTP r 1500



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

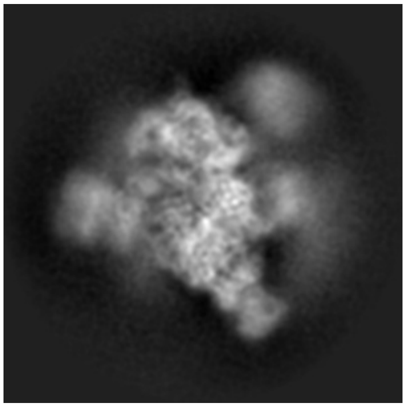
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-11697. These allow visual inspection of the internal detail of the map and identification of artifacts.

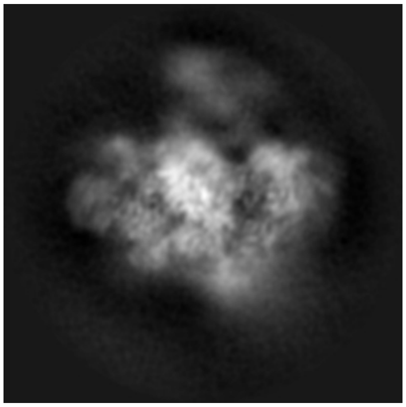
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

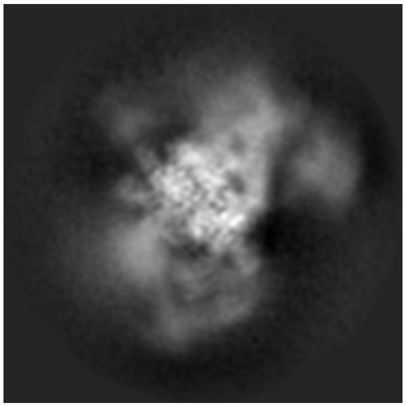
6.1.1 Primary map



X



Y

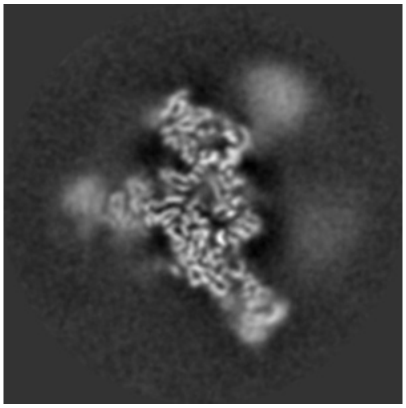


Z

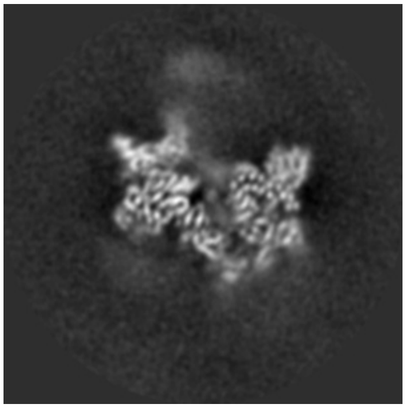
The images above show the map projected in three orthogonal directions.

6.2 Central slices

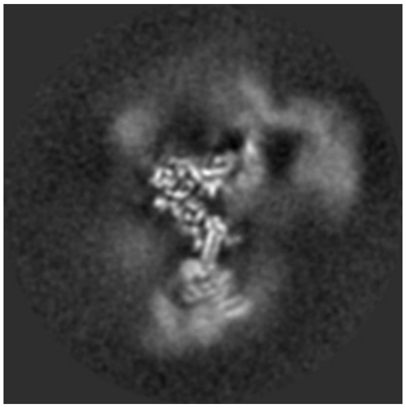
6.2.1 Primary map



X Index: 192



Y Index: 192

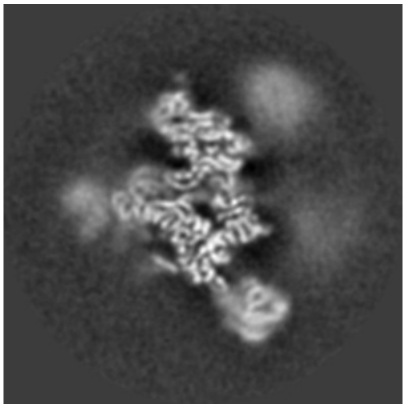


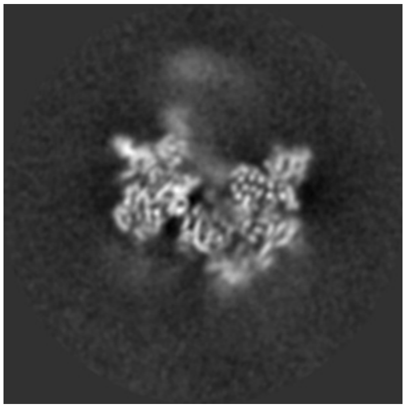
Z Index: 192

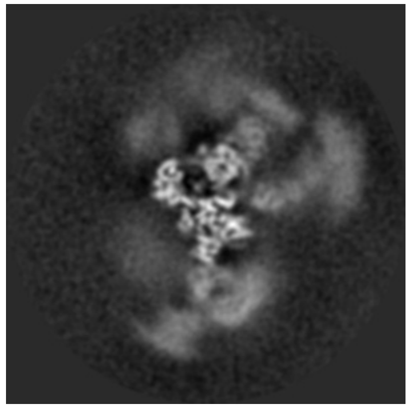
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

6.3.1 Primary map


X Index: 198


Y Index: 195

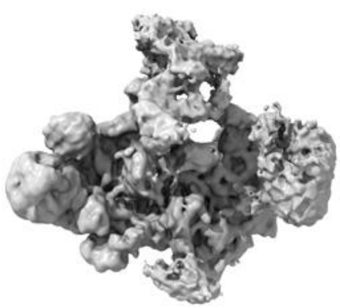

Z Index: 175

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

6.4 Orthogonal surface views (i)

6.4.1 Primary map


X


Y


Z

The images above show the 3D surface view of the map at the recommended contour level 0.013. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

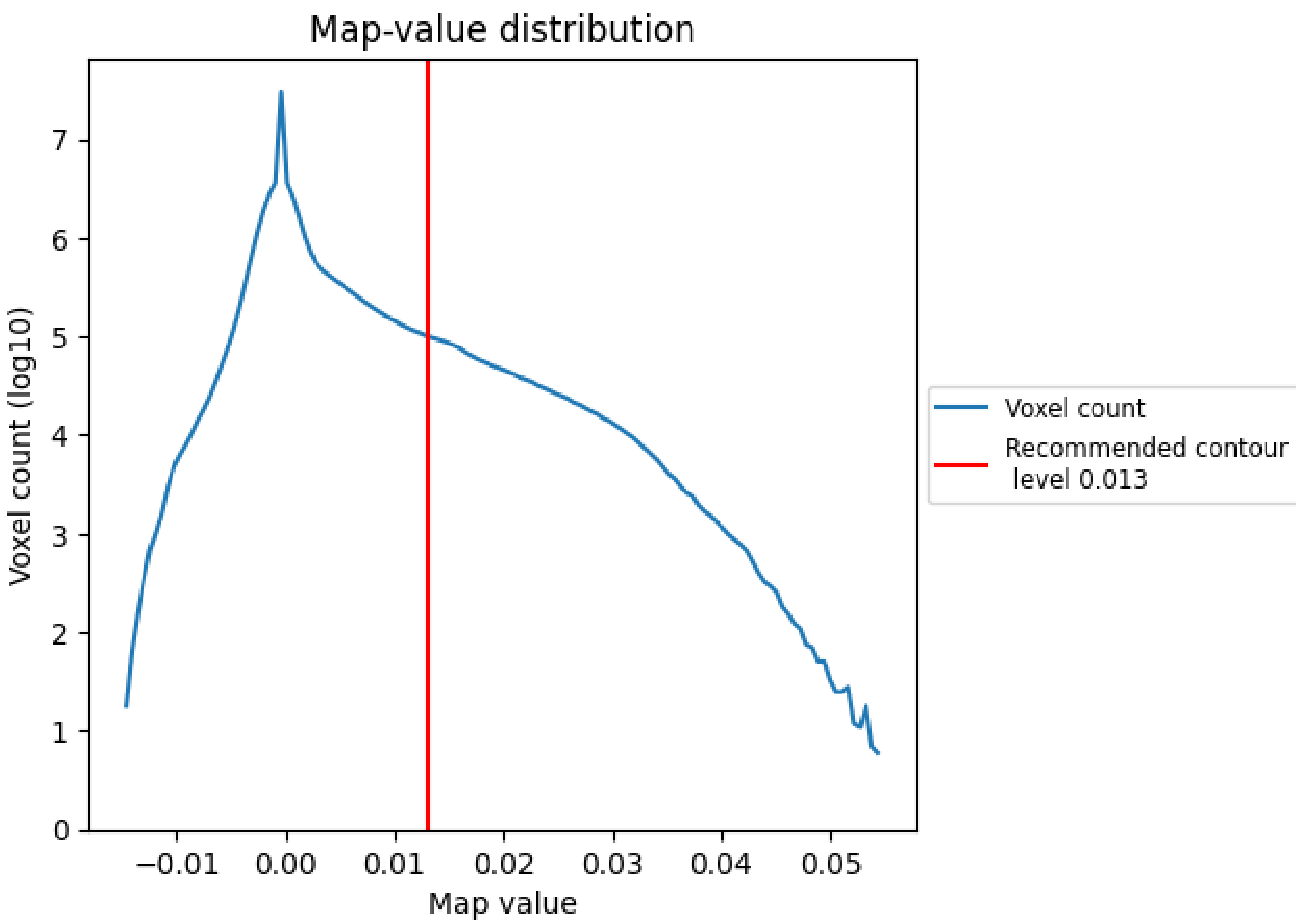
6.5 Mask visualisation (i)

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

7 Map analysis (i)

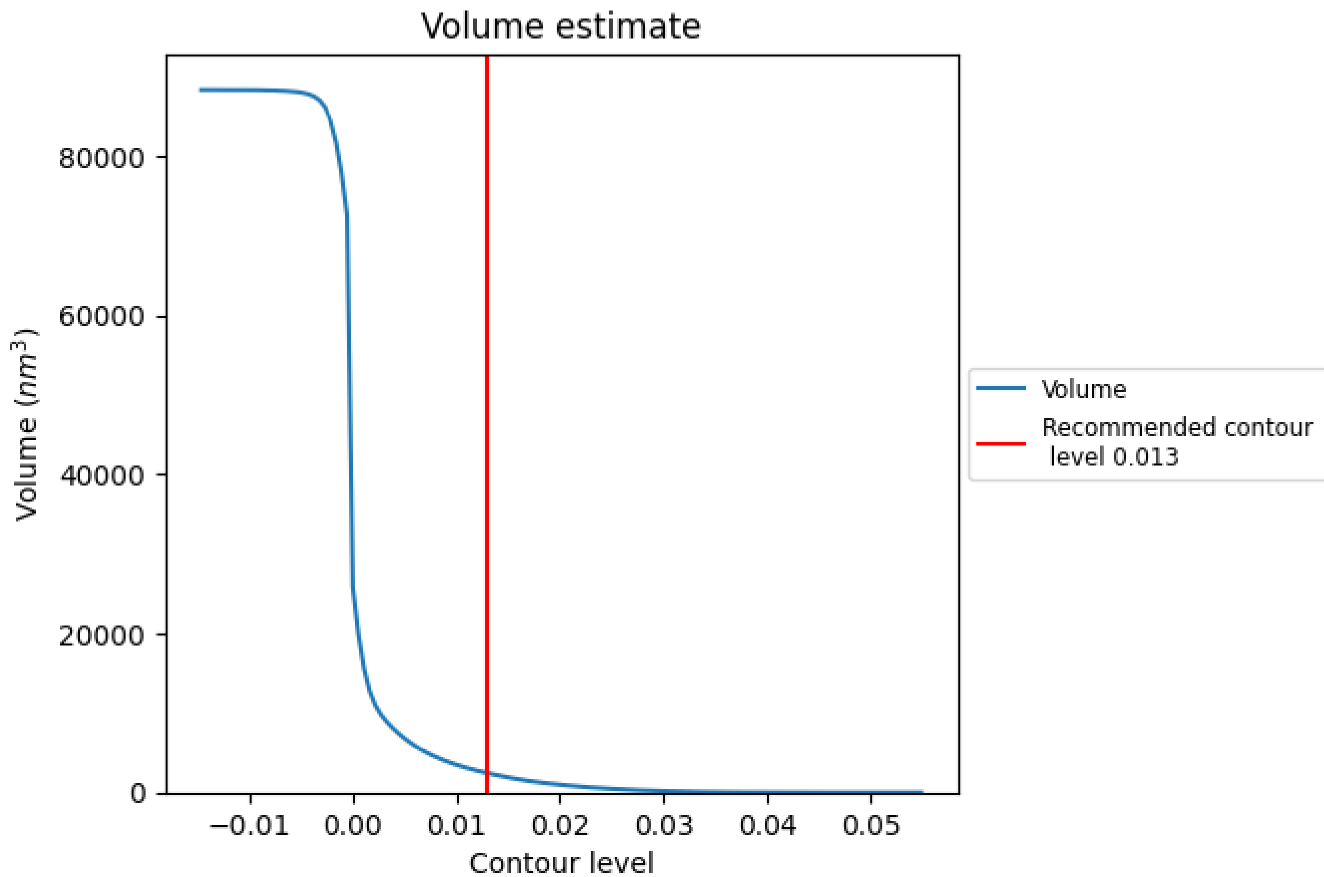
This section contains the results of statistical analysis of the map.

7.1 Map-value distribution (i)



The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

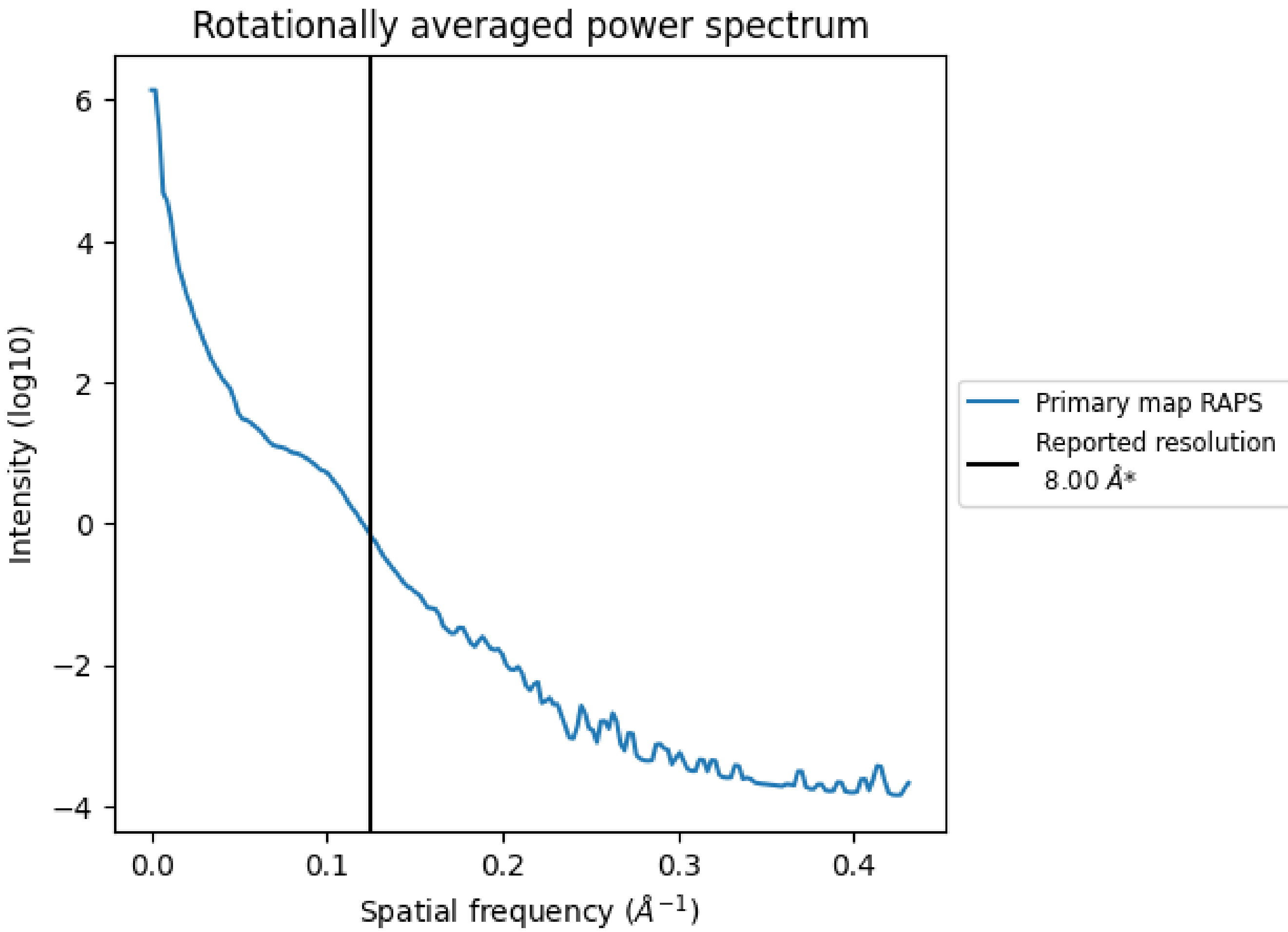
7.2 Volume estimate (i)



The volume at the recommended contour level is 2441 nm^3; this corresponds to an approximate mass of 2205 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

7.3 Rotationally averaged power spectrum (i)



*Reported resolution corresponds to spatial frequency of 0.125 Å^{-1}

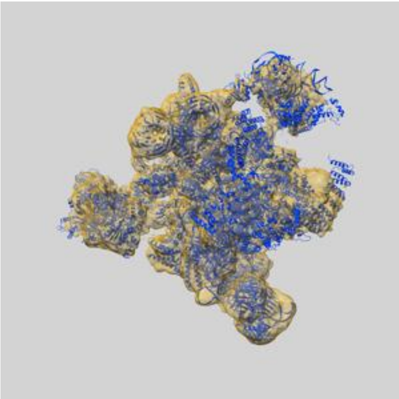
8 Fourier-Shell correlation (i)

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

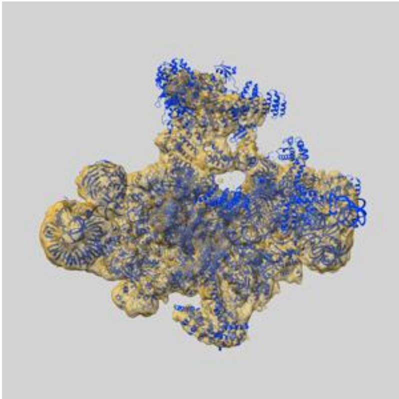
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-11697 and PDB model 7ABI. Per-residue inclusion information can be found in section 3 on page 14.

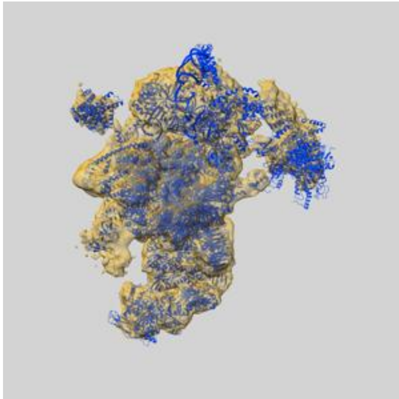
9.1 Map-model overlay (i)



X



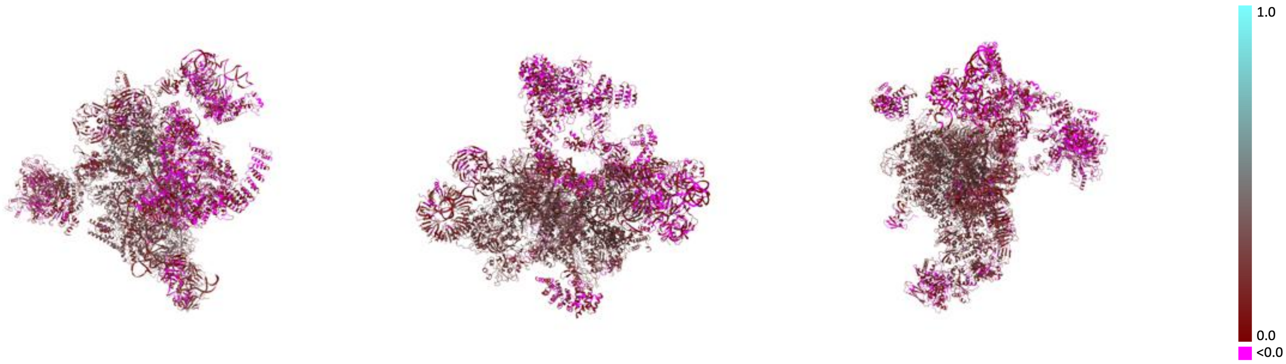
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.013 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.

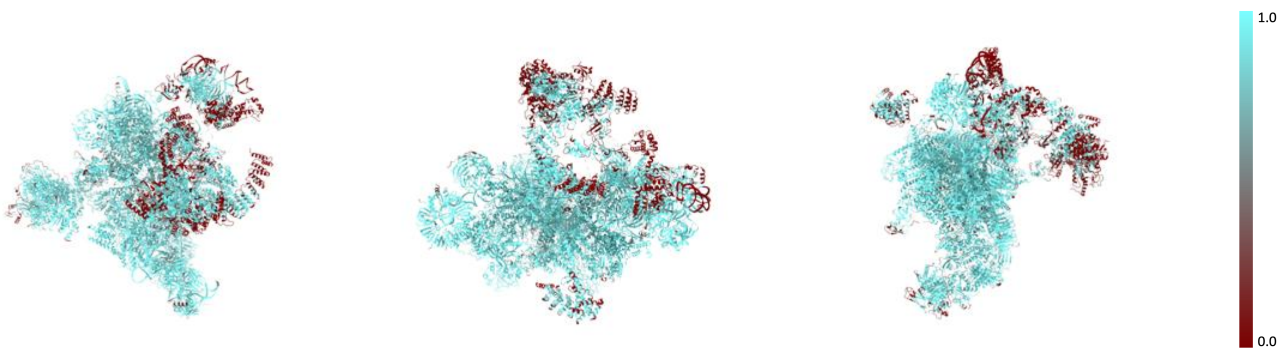
9.2 Q-score mapped to coordinate model (i)



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.

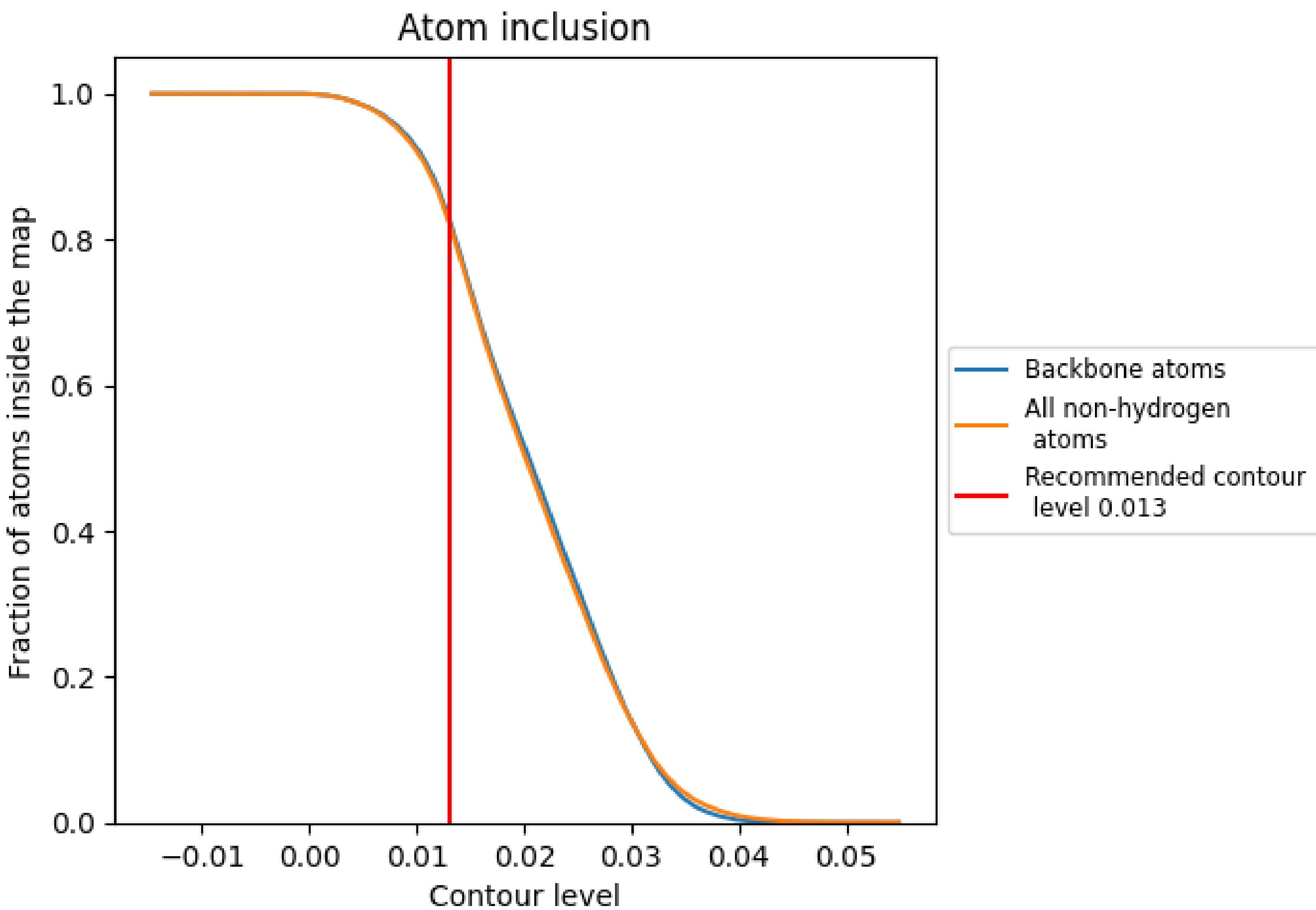
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.013).

9.4 Atom inclusion (i)



At the recommended contour level, 83% of all backbone atoms, 83% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary (i)

The table lists the average atom inclusion at the recommended contour level (0.013) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.8250	0.1520
0	0.9750	0.1750
1	0.9918	0.2300
2	0.6789	0.0990
3	0.9081	0.2390
4	0.4118	0.0810
5	0.9791	0.1700
6	0.9183	0.1750
7	0.7708	0.1220
8	0.9565	0.1530
A	0.9545	0.2510
B	0.4536	0.0970
D	0.8997	0.0580
E	0.9593	0.1460
F	0.7329	0.1450
G	0.9813	0.2200
I	0.9943	0.2500
K	0.9788	0.2720
L	0.9759	0.2470
M	0.4827	0.0430
N	0.9639	0.2250
O	0.8721	0.1650
P	0.9037	0.1410
Q	0.9914	0.2330
R	0.7169	0.2070
S	0.6126	0.0610
T	0.9437	0.2090
U	0.4184	0.0250
W	0.7451	0.0360
Y	0.9958	0.2410
Z	0.9784	0.2070
a	0.9822	0.1340
b	0.9423	0.1000
c	0.9459	0.1440
d	0.9738	0.1730

Continued on next page...

Continued from previous page...

Chain	Atom inclusion	Q-score
e	1.0000	0.2160
f	0.9969	0.1840
g	0.8934	0.1490
h	0.8174	0.0190
i	0.9694	0.0310
j	0.9752	0.0390
k	0.9643	0.0070
l	0.8554	0.0330
m	0.9298	0.0350
n	0.5423	-0.0150
o	0.5374	0.0480
p	0.3946	0.0320
q	0.9861	0.2730
r	0.9888	0.2430
s	0.8771	0.1020
t	0.5621	0.1020
u	0.9802	0.2790
v	0.8917	0.2330
w	0.8798	0.1230
x	1.0000	0.2590
y	0.9960	0.2310
z	0.9853	0.2420