wwPDB EM Validation Summary Report (i)

Nov 29, 2022 – 02:42 AM JST

PDB ID	:	7W59
EMDB ID	:	EMD-32317
Title	:	The cryo-EM structure of human pre-C*-I complex
Authors	:	Zhan, X.; Lu, Y.; Shi, Y.
Deposited on	:	2021-11-29
Resolution	:	3.60 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:

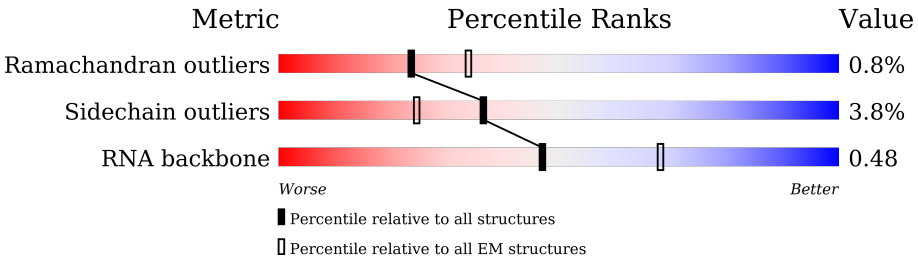
EMDB validation analysis	:	0.0.1.dev43
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

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 3.60 Å.

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



Metric	Whole archive (#Entries)	EM structures (#Entries)
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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

Mol	Chain	Length	Quality of chain
1	A	2335	79% 5% 15%
2	C	972	85% 11%
3	E	357	83% 16%
4	4	46	9% 15% 72%
5	G	174	9% 9% 30% 9% 53%
6	J	848	65% 33%
7	L	802	53% 46%
8	M	243	52% 47%

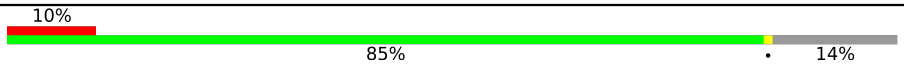
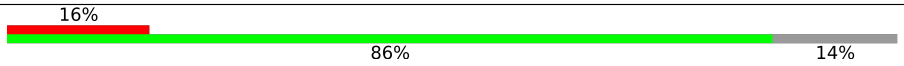
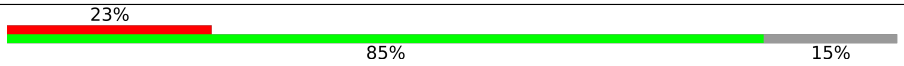
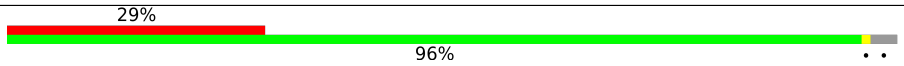
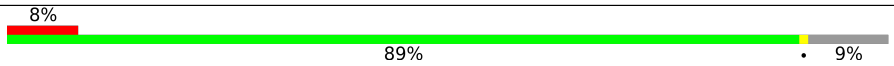
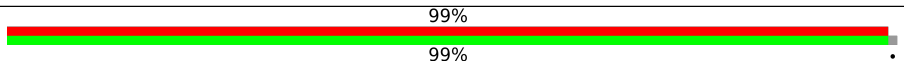
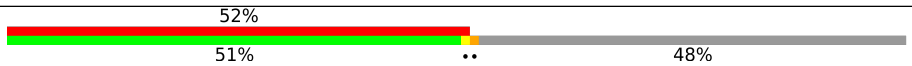
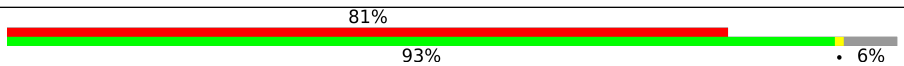
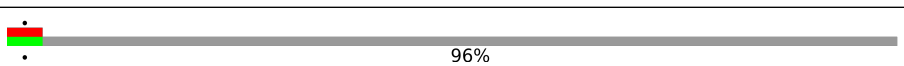
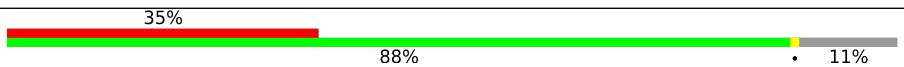
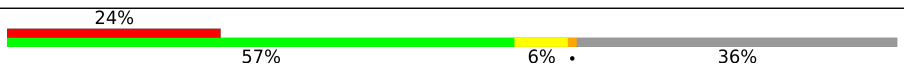
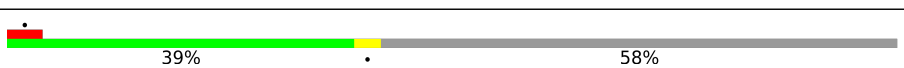
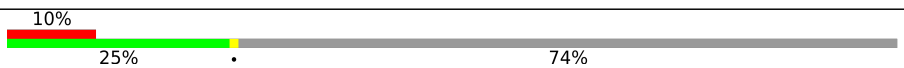
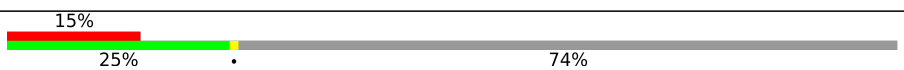
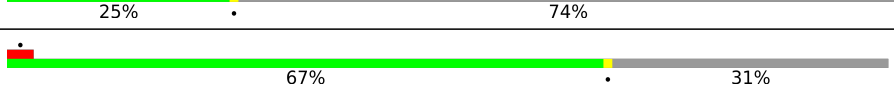
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	N	144	91% 6%
10	O	420	66% 32%
11	P	229	44% 51%
12	R	536	45% 7% 48%
13	S	166	94%
14	T	514	55% 39%
15	U	2752	97%
16	V	908	21% 49% 50%
17	W	579	84% 12%
18	B	117	54% 16% 28%
19	F	107	46% 43% 9%
20	H	188	6% 28% 43% 26%
21	b	240	6% 40% 58%
21	i	240	14% 35% 64%
22	X	254	17% 10% 73%
23	I	855	70% 28%
24	y	301	26% 74%
25	Y	1220	41% 54% 45%
26	a	126	5% 64% 36%
26	h	126	12% 64% 36%
27	c	119	67% 31%
27	j	119	21% 67% 31%
28	d	118	14% 81% 18%
28	k	118	24% 70% 28%
29	f	86	16% 85% 14%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
29	m	86	
30	e	92	
30	l	92	
31	g	76	
31	n	76	
32	v	146	
33	w	174	
34	u	411	
35	x	703	
36	Q	1485	
37	o	255	
38	p	225	
39	q	504	
39	r	504	
39	s	504	
39	t	504	
40	K	225	

2 Entry composition (i)

There are 45 unique types of molecules in this entry. The entry contains 92315 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 Pre-mRNA-processing-splicing factor 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	1984	16449	10601	2879	2899	70	0	0

- Molecule 2 is a protein called 116 kDa U5 small nuclear ribonucleoprotein component.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	C	862	6798	4347	1138	1281	32	0	0

- Molecule 3 is a protein called U5 small nuclear ribonucleoprotein 40 kDa protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	E	299	2338	1470	410	445	13	0	0

- Molecule 4 is a RNA chain called pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
4	4	13	276	123	50	90	13	0	0

- Molecule 5 is a RNA chain called pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
5	G	82	1510	666	210	552	82	0	0

- Molecule 6 is a protein called Crooked neck-like protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	J	568	3814	2376	717	715	6	0	0

- Molecule 7 is a protein called Cell division cycle 5-like protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	L	437	Total	C	N	O	S	0	0
			3015	1859	584	565	7		

- Molecule 8 is a protein called Pre-mRNA-splicing factor SYF2.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	M	130	Total	C	N	O	S	0	0
			1098	684	204	208	2		

- Molecule 9 is a protein called Protein BUD31 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	N	143	Total	C	N	O	S	0	0
			1184	746	217	209	12		

- Molecule 10 is a protein called Pre-mRNA-splicing factor RBM22.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	O	285	Total	C	N	O	S	0	0
			2296	1442	408	426	20		

- Molecule 11 is a protein called Spliceosome-associated protein CWC15 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	P	113	Total	C	N	O	S	0	0
			953	583	189	179	2		

- Molecule 12 is a protein called SNW domain-containing protein 1.

Mol	Chain	Residues	Atoms						AltConf	Trace
12	R	280	Total	C	N	O	P	S	0	0
			2243	1401	411	416	2	13		

- Molecule 13 is a protein called Peptidyl-prolyl cis-trans isomerase-like 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	S	159	Total	C	N	O	S	0	0
			1236	787	215	227	7		

- Molecule 14 is a protein called Pleiotropic regulator 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	T	312	Total	C	N	O	S	0	0
			2454	1550	446	450	8		

- Molecule 15 is a protein called Serine/arginine repetitive matrix protein 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	U	72	Total	C	N	O	S	0	0
			422	257	82	82	1		

- Molecule 16 is a protein called Pre-mRNA-splicing factor CWC22 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	V	452	Total	C	N	O	S	0	0
			2632	1639	492	495	6		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
17	W	509	Total	C	N	O	S	0	0
			4129	2628	715	762	24		

- Molecule 18 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	B	84	Total	C	N	O	P	0	0
			1768	792	295	597	84		

- Molecule 19 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	F	97	Total	C	N	O	P	0	0
			2075	928	381	669	97		

- Molecule 20 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	H	140	Total	C	N	O	P	0	0
			2966	1326	510	990	140		

- Molecule 21 is a protein called Small nuclear ribonucleoprotein-associated proteins B and B'.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	b	102	Total	C	N	O	S	0	0
			786	492	148	139	7		
21	i	86	Total	C	N	O	S	0	0
			690	434	126	123	7		

- Molecule 22 is a protein called PSME3-interacting protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	X	69	Total	C	N	O	S	0	0
			616	380	111	124	1		

- Molecule 23 is a protein called Pre-mRNA-splicing factor SYF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	I	617	Total	C	N	O	S	0	0
			3845	2380	721	733	11		

- Molecule 24 is a protein called Peptidyl-prolyl cis-trans isomerase E.

Mol	Chain	Residues	Atoms				AltConf	Trace
24	y	79	Total	C	N	O	0	0
			390	232	79	79		

- Molecule 25 is a protein called ATP-dependent RNA helicase DHX8.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	Y	667	Total	C	N	O	S	4	0
			3431	2057	680	693	1		

- Molecule 26 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	a	81	Total	C	N	O	S	0	0
			640	401	113	120	6		
26	h	81	Total	C	N	O	S	0	0
			633	397	112	118	6		

- Molecule 27 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	c	82	Total	C	N	O	S	0	0
			649	413	113	119	4		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
27	j	82	Total 649	C 413	N 113	O 119	S 4	0	0

- Molecule 28 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	d	97	Total 776	C 488	N 143	O 140	S 5	0	0
28	k	85	Total 688	C 432	N 125	O 126	S 5	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
29	f	74	Total 576	C 373	N 95	O 103	S 5	0	0
29	m	74	Total 576	C 373	N 95	O 103	S 5	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
30	e	79	Total 652	C 412	N 116	O 119	S 5	0	0
30	l	78	Total 644	C 407	N 115	O 118	S 4	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
31	g	74	Total 577	C 364	N 104	O 103	S 6	0	0
31	n	69	Total 538	C 342	N 96	O 94	S 6	0	0

- Molecule 32 is a protein called Protein mago nashi homolog.

Mol	Chain	Residues	Atoms				AltConf	Trace
32	v	144	Total 711	C 423	N 144	O 144	0	0

- Molecule 33 is a protein called RNA-binding protein 8A.

Mol	Chain	Residues	Atoms				AltConf	Trace
33	w	91	Total	C	N	O	0	0
			445	263	91	91		

- Molecule 34 is a protein called Eukaryotic initiation factor 4A-III.

Mol	Chain	Residues	Atoms				AltConf	Trace
34	u	386	Total	C	N	O	0	0
			1907	1135	386	386		

- Molecule 35 is a protein called Protein CASC3.

Mol	Chain	Residues	Atoms				AltConf	Trace
35	x	25	Total	C	N	O	0	0
			124	74	25	25		

- Molecule 36 is a protein called RNA helicase aquarius.

Mol	Chain	Residues	Atoms				AltConf	Trace
36	Q	1322	Total	C	N	O	4	0
			6562	3918	1322	1322		

- Molecule 37 is a protein called U2 small nuclear ribonucleoprotein A'.

Mol	Chain	Residues	Atoms					AltConf	Trace
37	o	162	Total	C	N	O	S	0	0
			1282	820	219	240	3		

- Molecule 38 is a protein called U2 small nuclear ribonucleoprotein B".

Mol	Chain	Residues	Atoms					AltConf	Trace
38	p	94	Total	C	N	O	S	0	0
			760	488	135	132	5		

- Molecule 39 is a protein called Pre-mRNA-processing factor 19.

Mol	Chain	Residues	Atoms				AltConf	Trace
39	q	132	Total	C	N	O	0	0
			659	395	132	132		
39	r	131	Total	C	N	O	0	0
			654	392	131	131		
39	s	132	Total	C	N	O	0	0
			659	395	132	132		

Continued on next page...

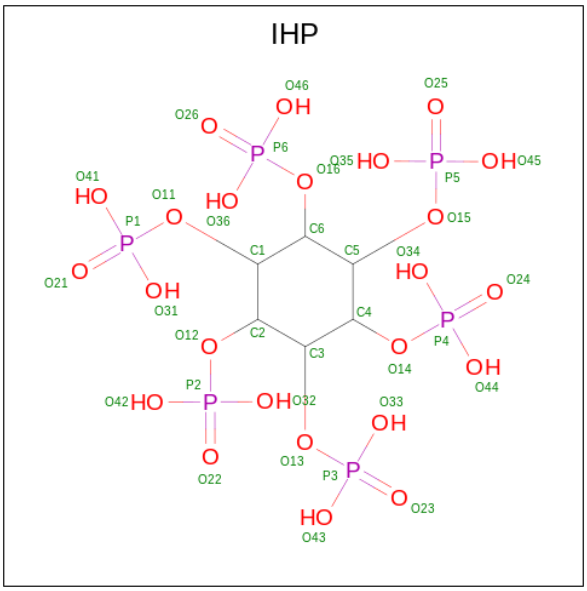
Continued from previous page...

Mol	Chain	Residues	Atoms				AltConf	Trace
39	t	131	Total	C	N	O	0	0
			654	392	131	131		

- Molecule 40 is a protein called Pre-mRNA-splicing factor SPF27.

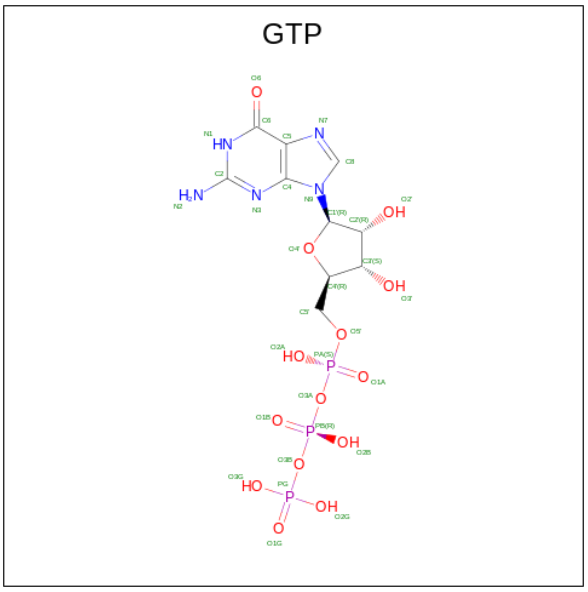
Mol	Chain	Residues	Atoms				AltConf	Trace
40	K	155	Total	C	N	O	0	0
			772	462	155	155		

- Molecule 41 is INOSITOL HEXAKISPHOSPHATE (three-letter code: IHP) (formula: $C_6H_{18}O_{24}P_6$) (labeled as "Ligand of Interest" by depositor).



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

- Molecule 42 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$) (labeled as "Ligand of Interest" by depositor).



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

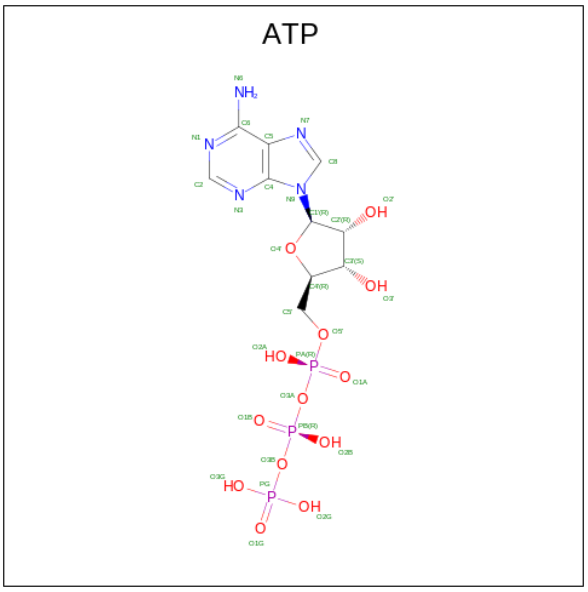
- Molecule 43 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
43	C	1	Total	Mg	0
			1	1	
43	F	6	Total	Mg	0
			6	6	
43	Q	2	Total	Mg	0
			2	2	

- Molecule 44 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
44	N	3	Total	Zn	0
			3	3	
44	O	3	Total	Zn	0
			3	3	

- Molecule 45 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$) (labeled as "Ligand of Interest" by depositor).



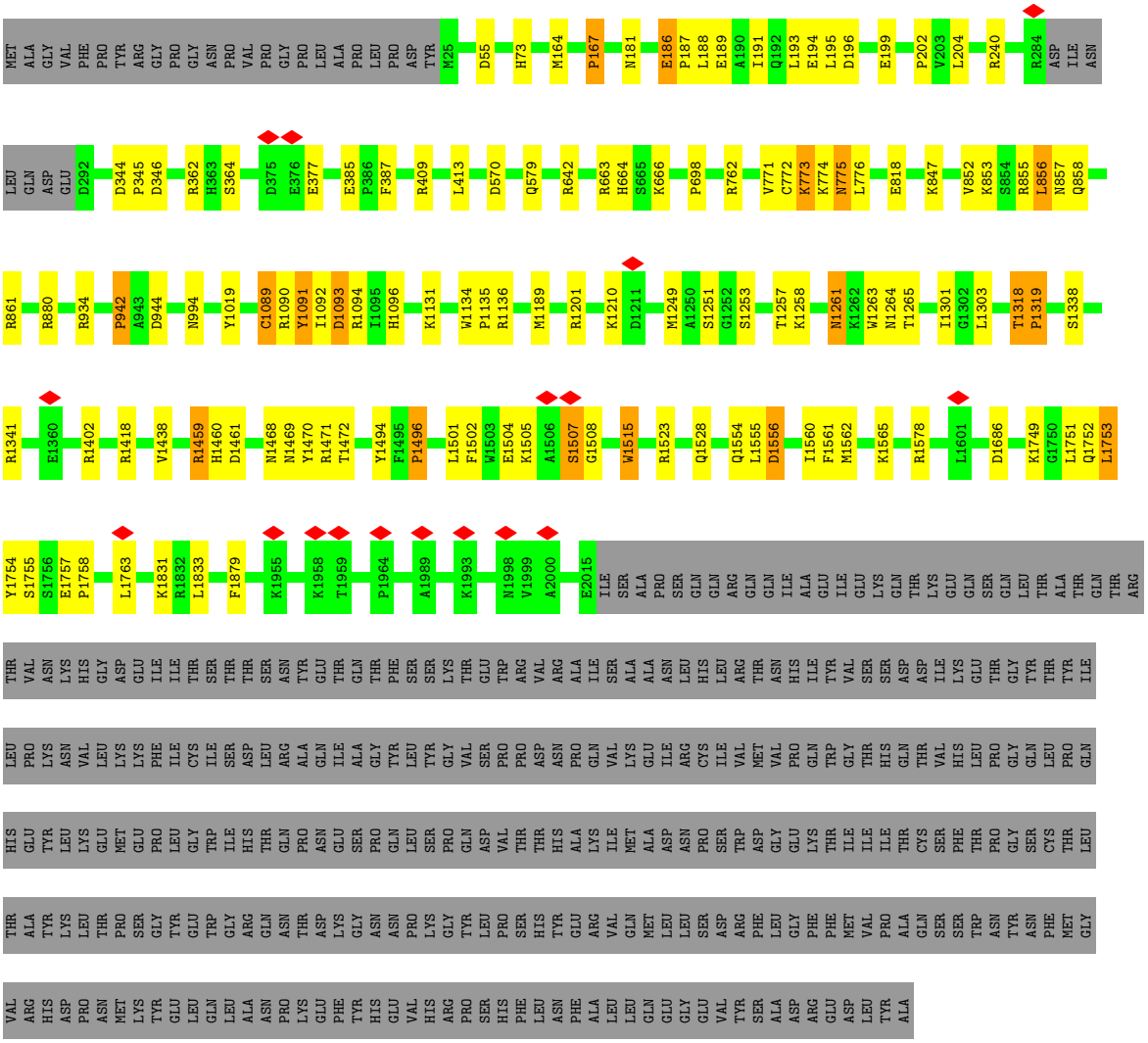
Mol	Chain	Residues	Atoms					AltConf
45	Q	1	Total	C	N	O	P	0
			31	10	5	13	3	

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.

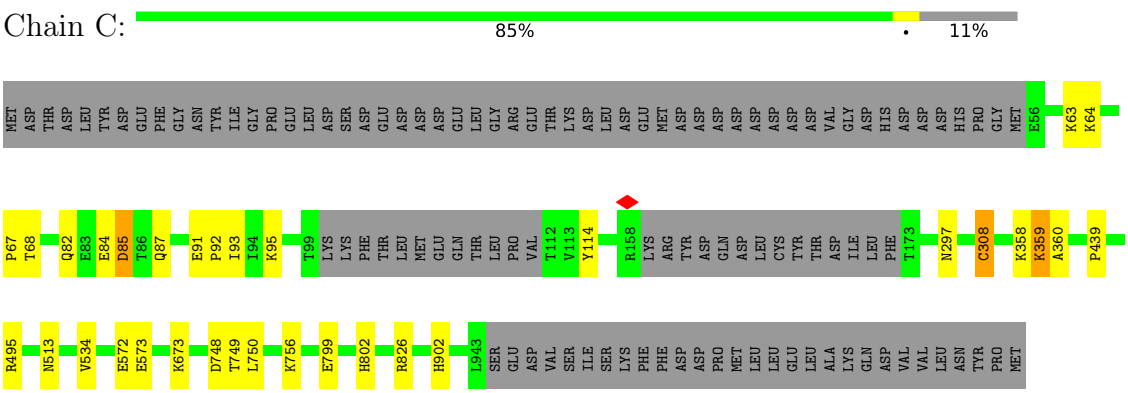
• Molecule 1: Pre-mRNA-processing-splicing factor 8

Chain A:



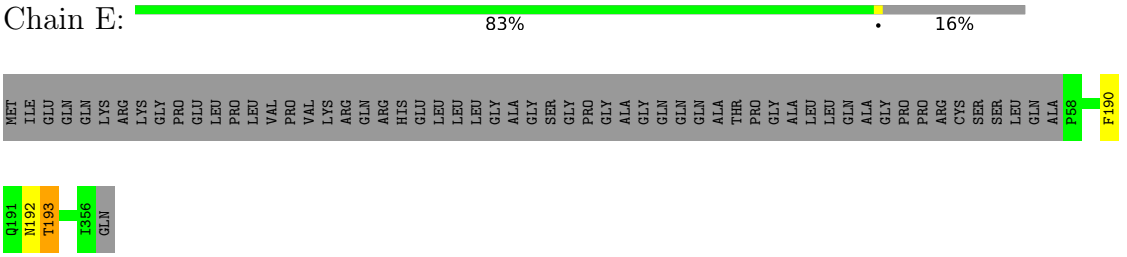
• Molecule 2: 116 kDa U5 small nuclear ribonucleoprotein component

Chain C:



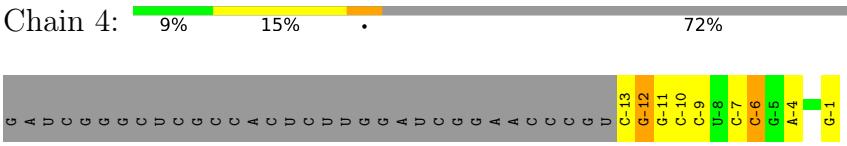
• Molecule 3: U5 small nuclear ribonucleoprotein 40 kDa protein

Chain E:



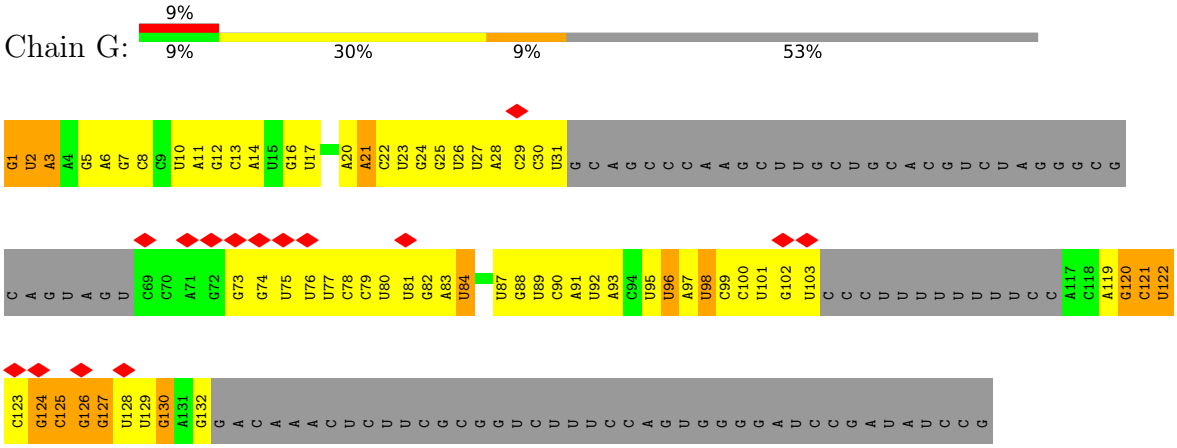
• Molecule 4: pre-mRNA

Chain 4:



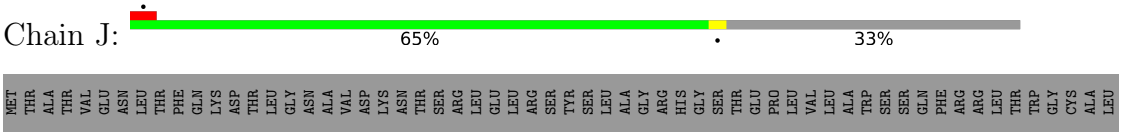
• Molecule 5: pre-mRNA

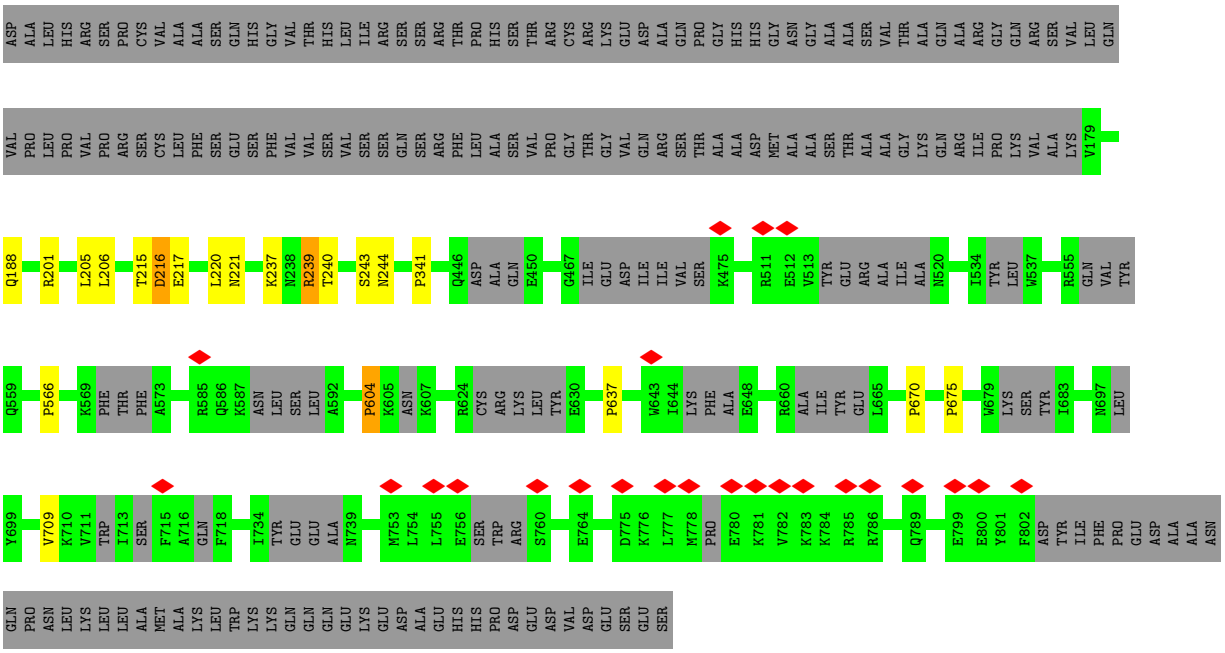
Chain G:



• Molecule 6: Crooked neck-like protein 1

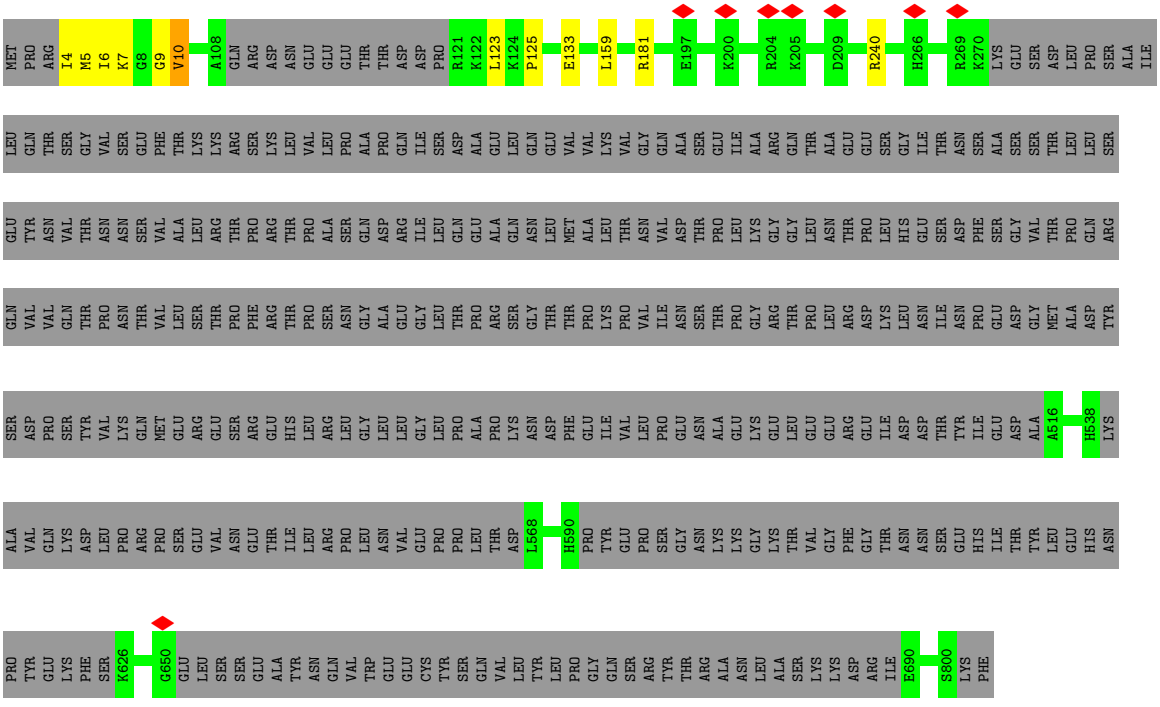
Chain J:





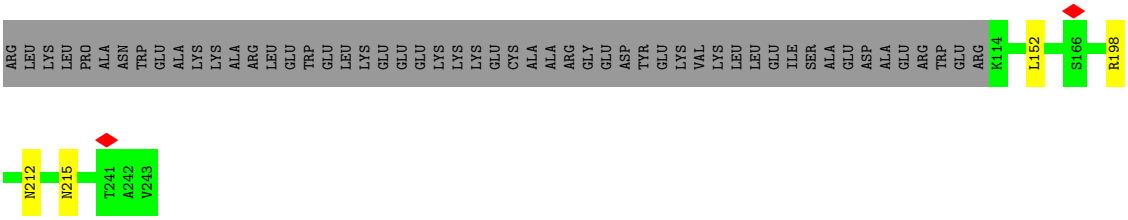
• Molecule 7: Cell division cycle 5-like protein

Chain L:



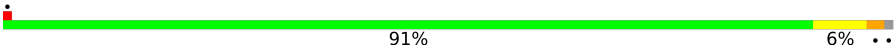
• Molecule 8: Pre-mRNA-splicing factor SYF2

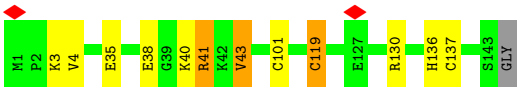
Chain M:



• Molecule 9: Protein BUD31 homolog

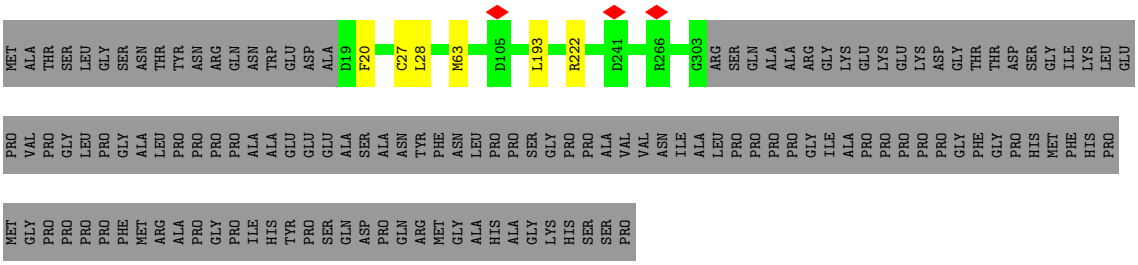
Chain N:





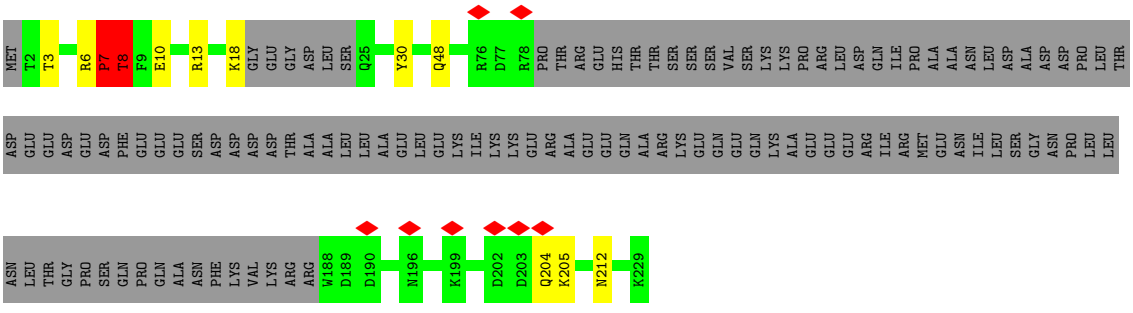
• Molecule 10: Pre-mRNA-splicing factor RBM22

Chain O:



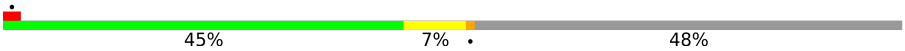
• Molecule 11: Spliceosome-associated protein CWC15 homolog

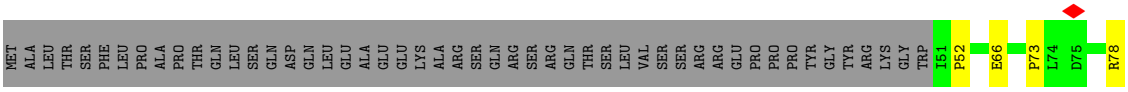
Chain P:

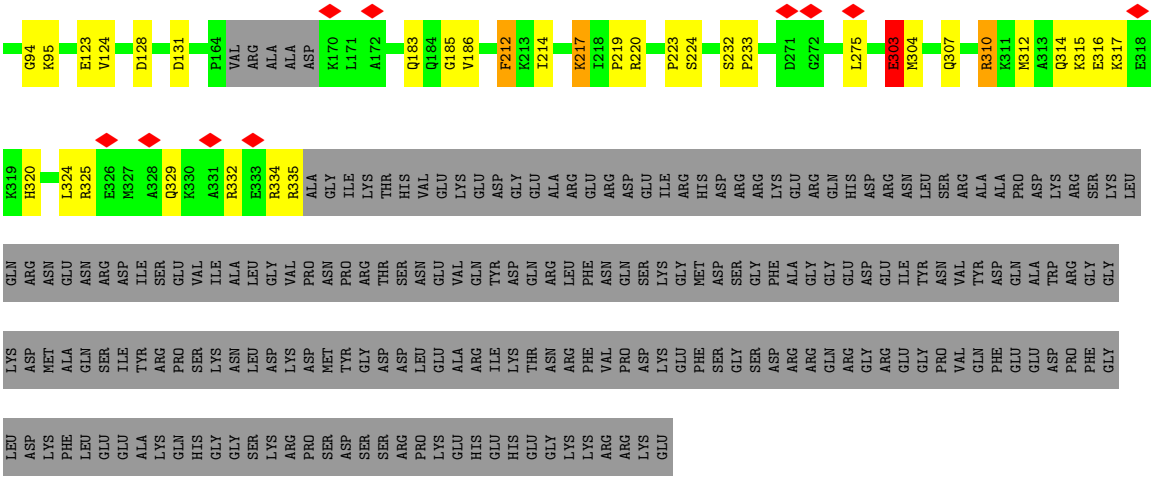


• Molecule 12: SNW domain-containing protein 1

Chain R:

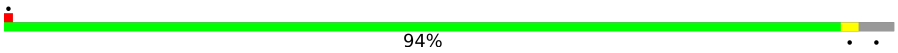


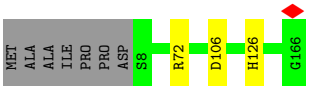




• Molecule 13: Peptidyl-prolyl cis-trans isomerase-like 1

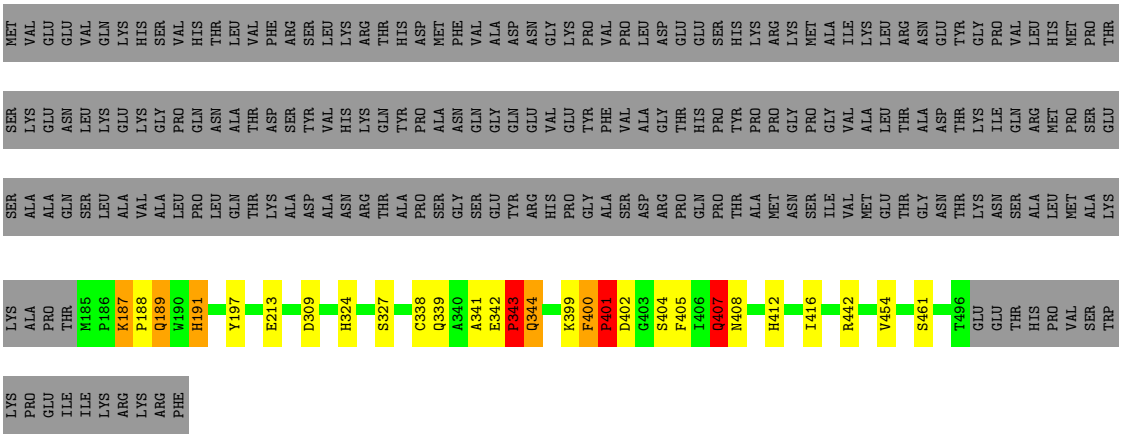
Chain S:





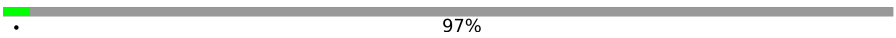
• Molecule 14: Pleiotropic regulator 1

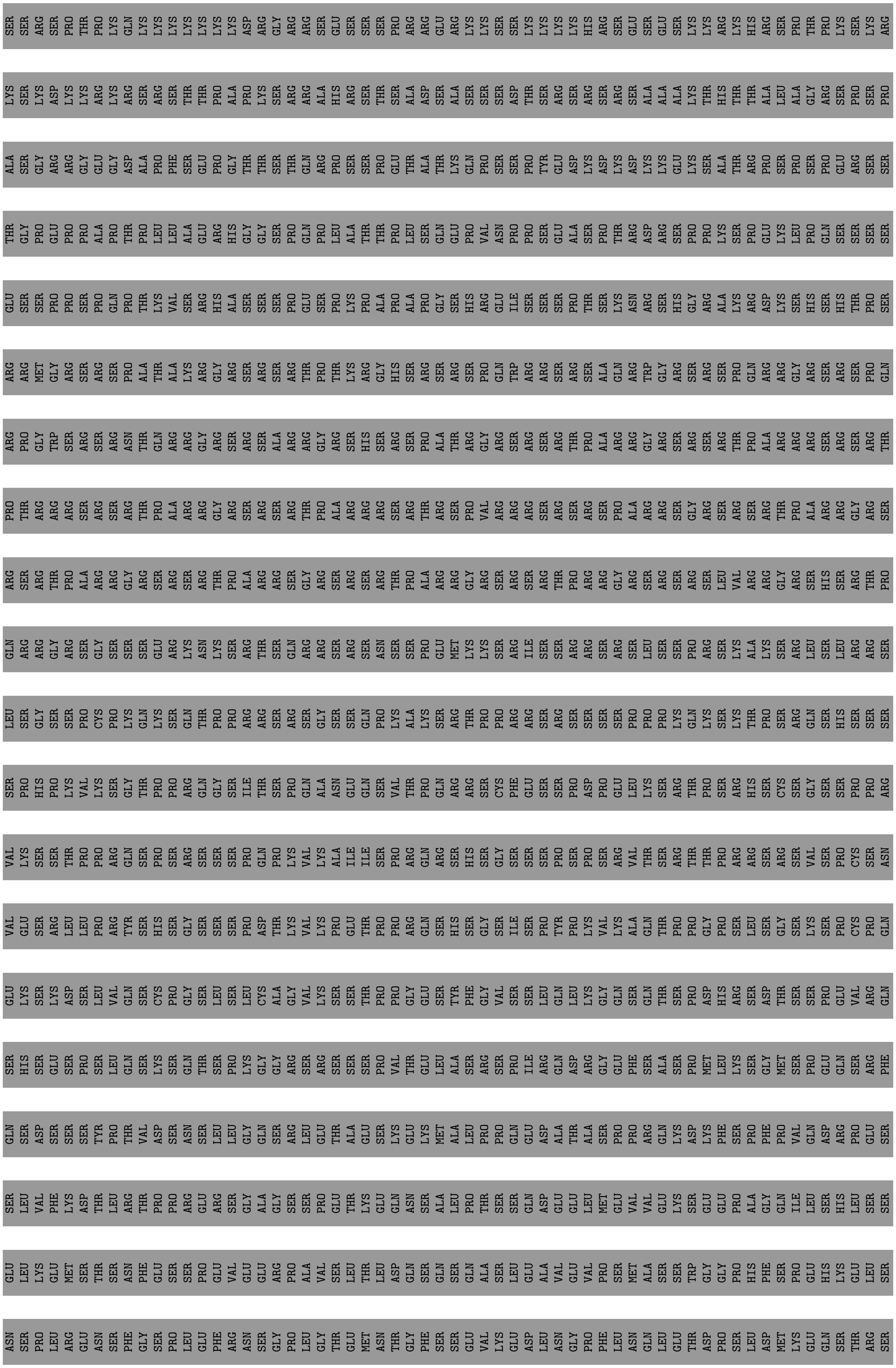
Chain T:

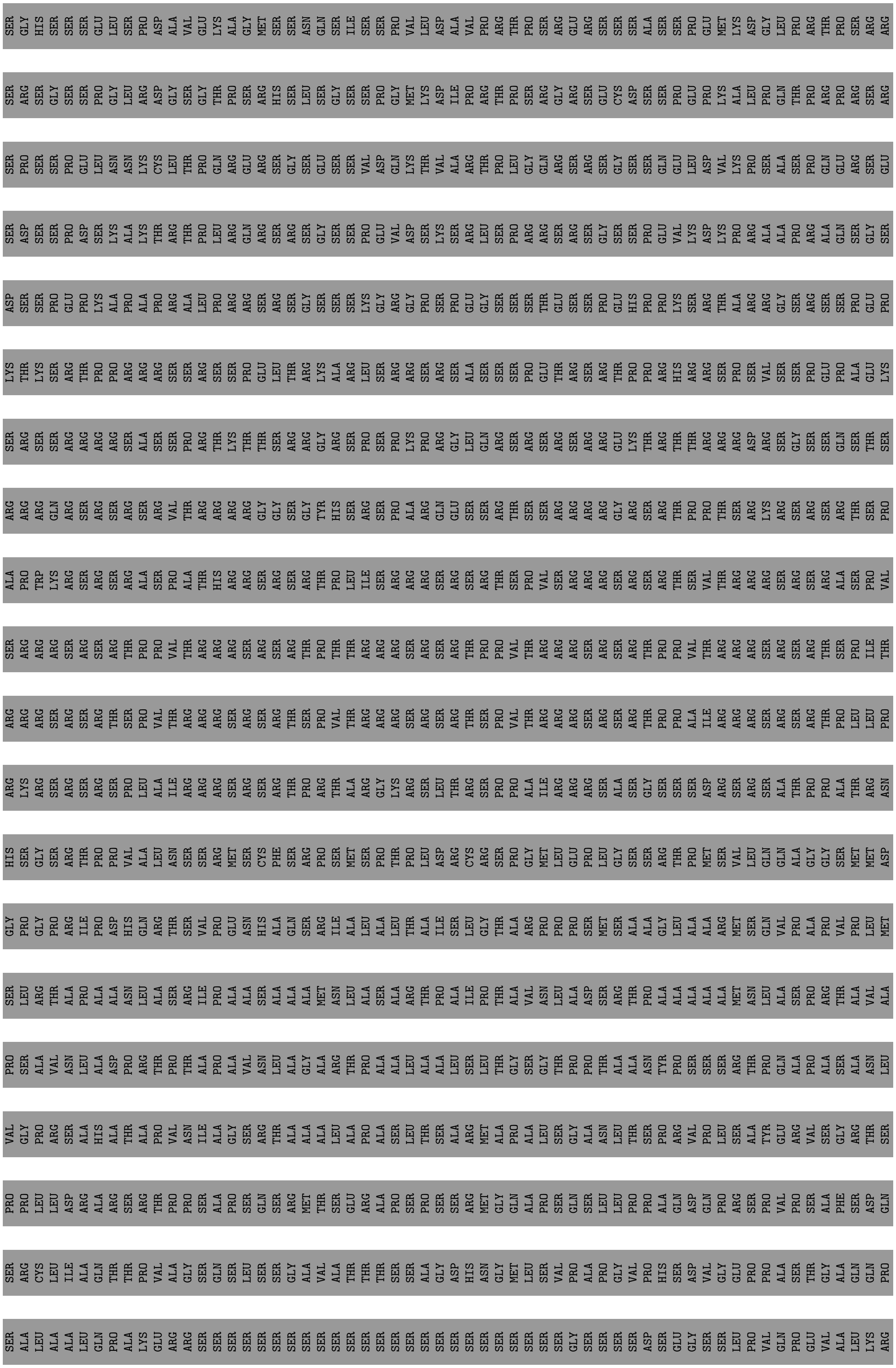


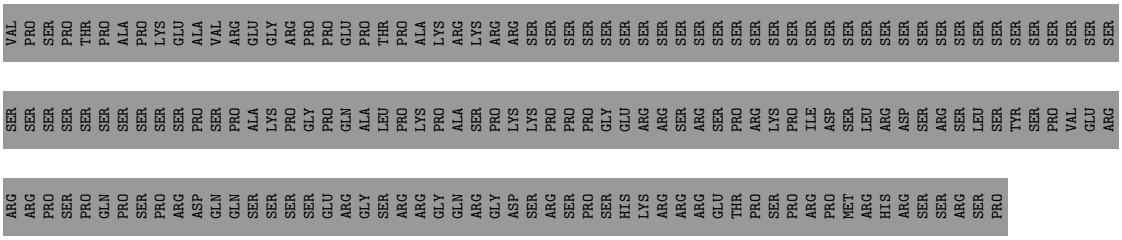
• Molecule 15: Serine/arginine repetitive matrix protein 2

Chain U:



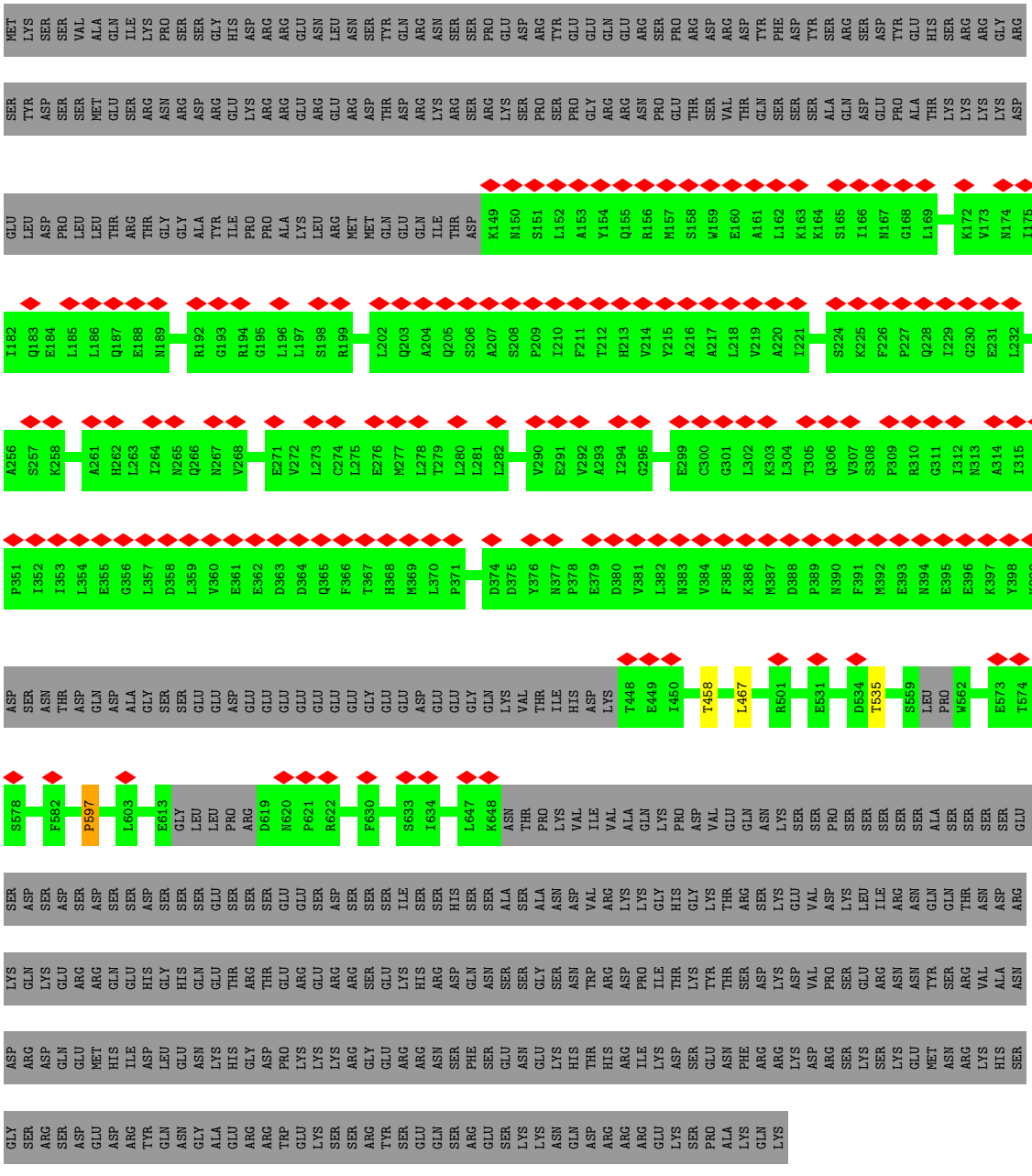

SER SER ARG SER PRO THR PRO LYS GLN LYS LYS LYS LYS LYS LYS LYS ASP ARG GLY ARG ARG SER GLU SER SER SER PRO ARG ARG GLU ARG LYS LYS SER SER LYS LYS LYS LYS HIS ARG SER GLU SER GLU SER LYS LYS ARG LYS HIS ARG SER PRO THR PRO LYS SER LYS ARG
LYS SER LYS ASP LYS LYS ARG LYS ARG SER ARG SER THR THR PRO ALA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


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



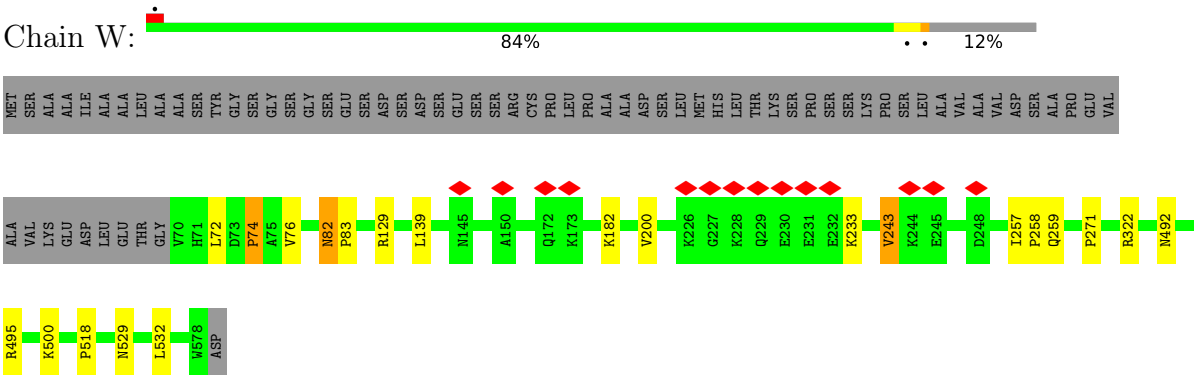
• Molecule 16: Pre-mRNA-splicing factor CWC22 homolog

Chain V:



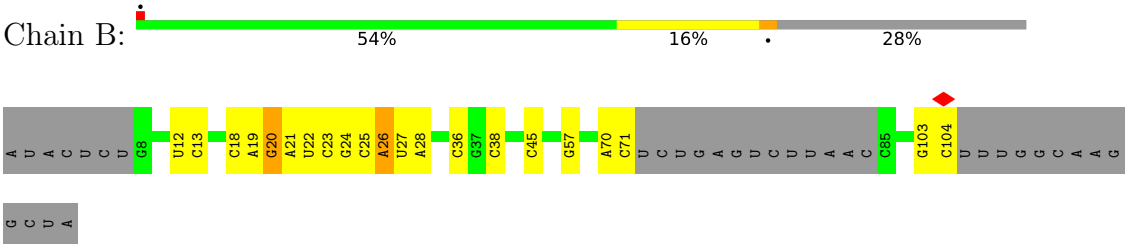
• Molecule 17: Pre-mRNA-processing factor 17

Chain W:



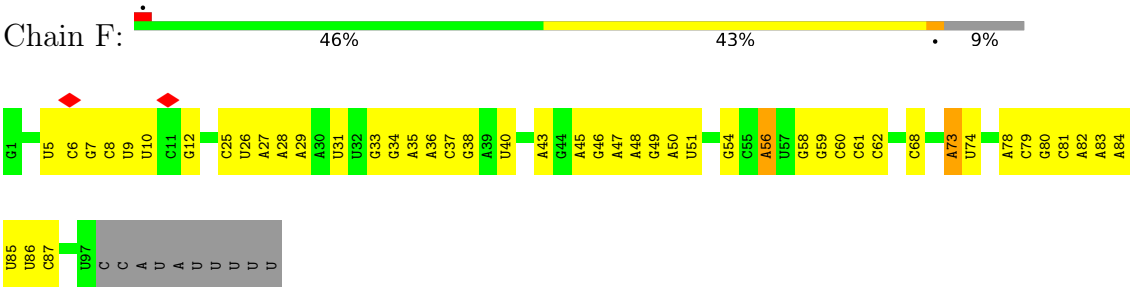
• Molecule 18: U5 snRNA

Chain B:



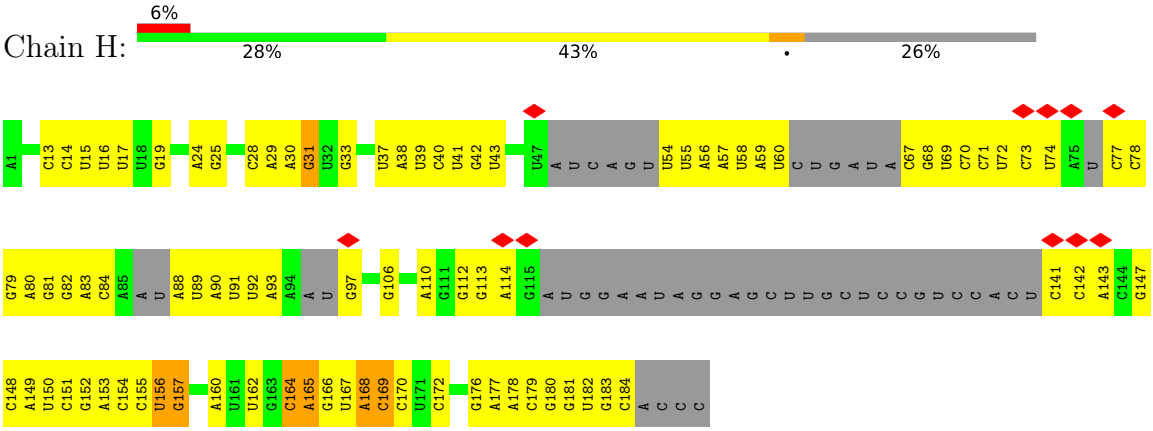
• Molecule 19: U6 snRNA

Chain F:



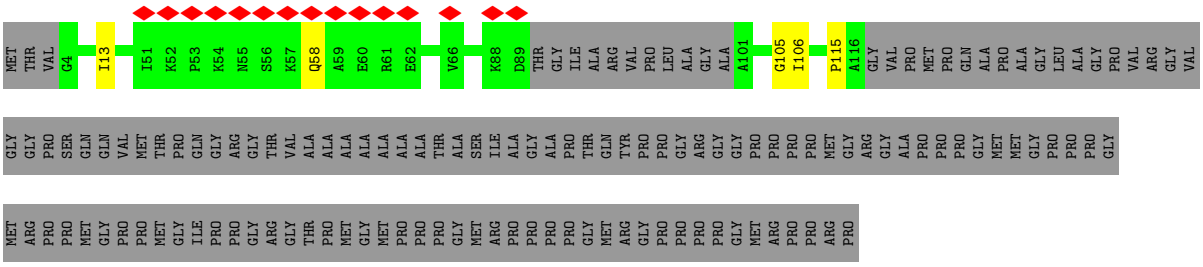
• Molecule 20: U2 snRNA

Chain H:



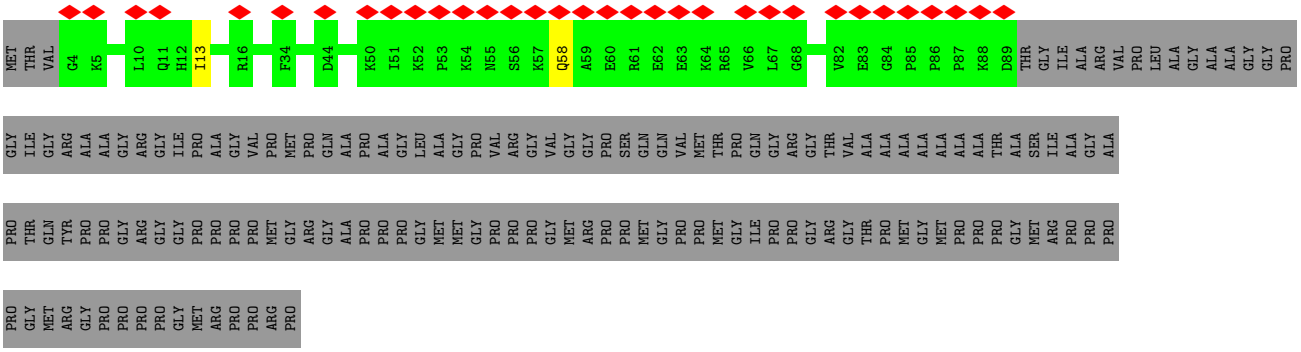
• Molecule 21: Small nuclear ribonucleoprotein-associated proteins B and B'

Chain b:



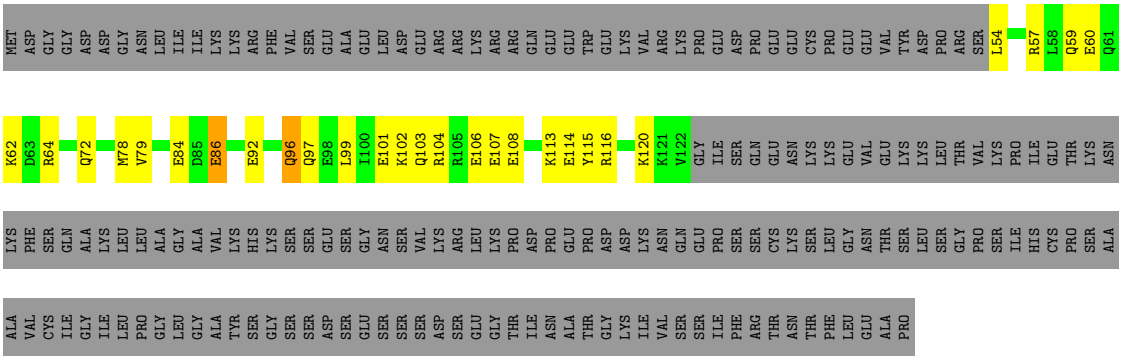
• Molecule 21: Small nuclear ribonucleoprotein-associated proteins B and B'

Chain i:



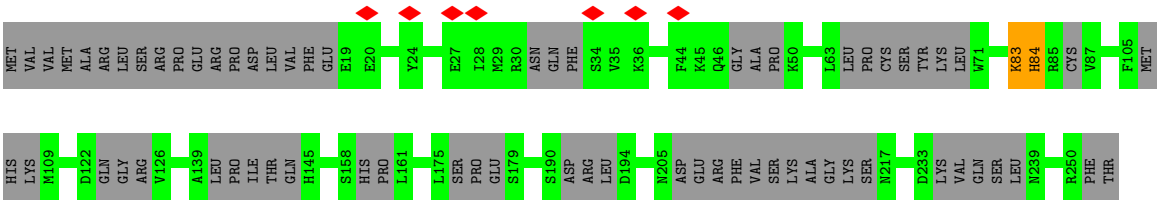
• Molecule 22: PSME3-interacting protein

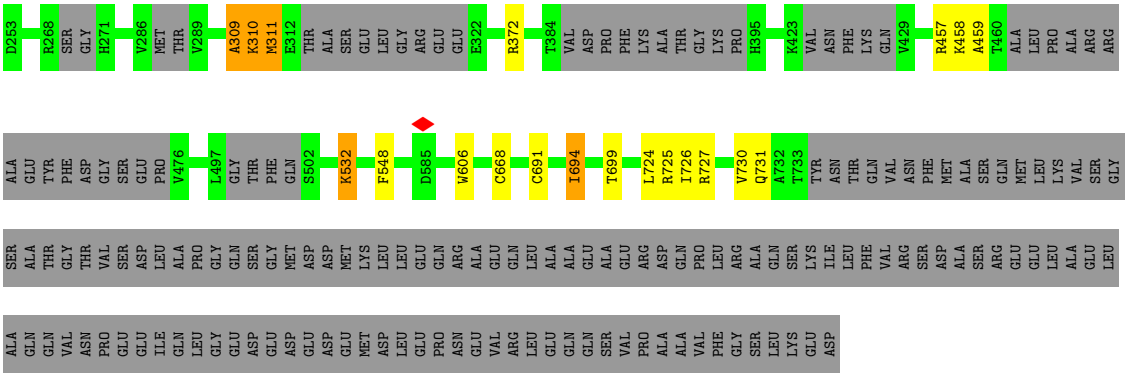
Chain X:



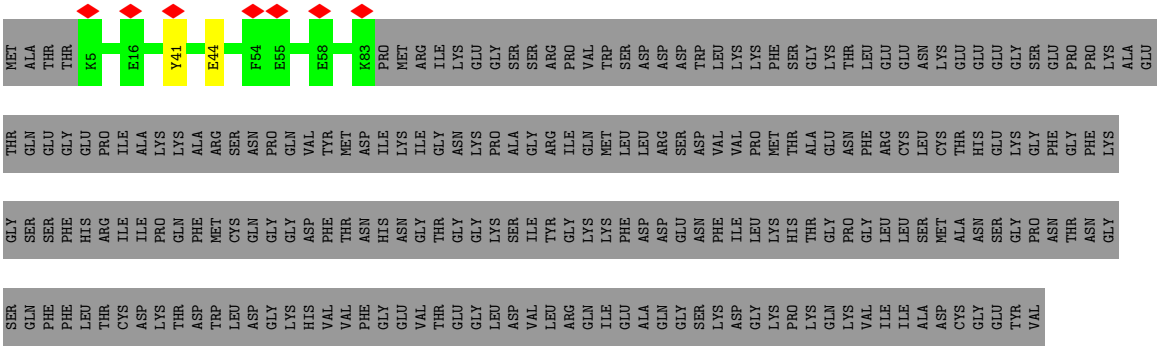
• Molecule 23: Pre-mRNA-splicing factor SYF1

Chain I:

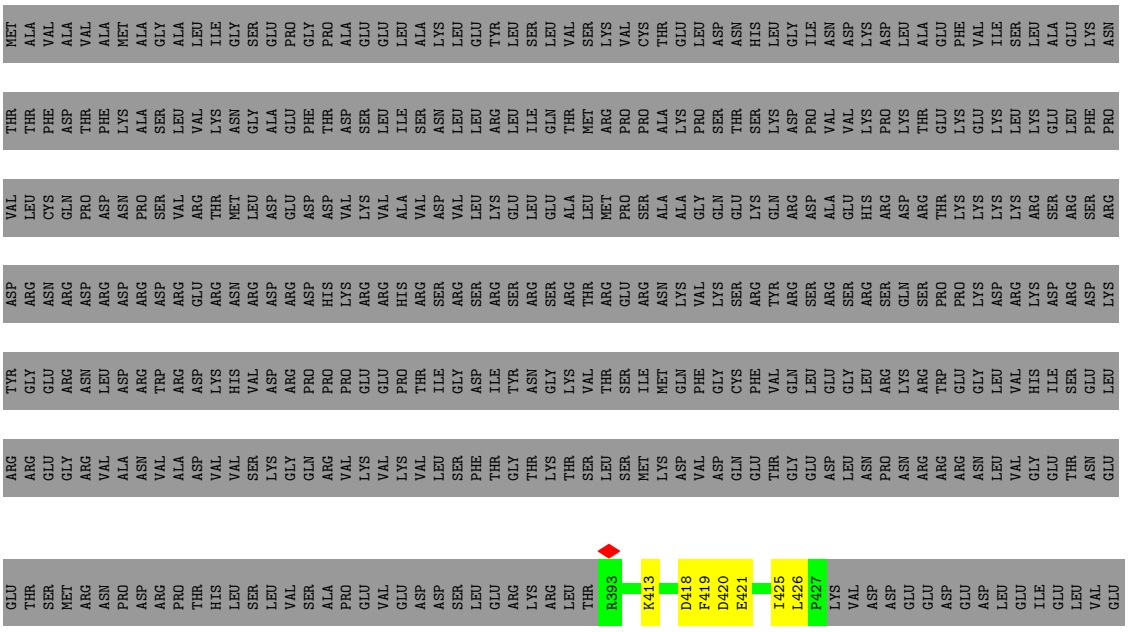


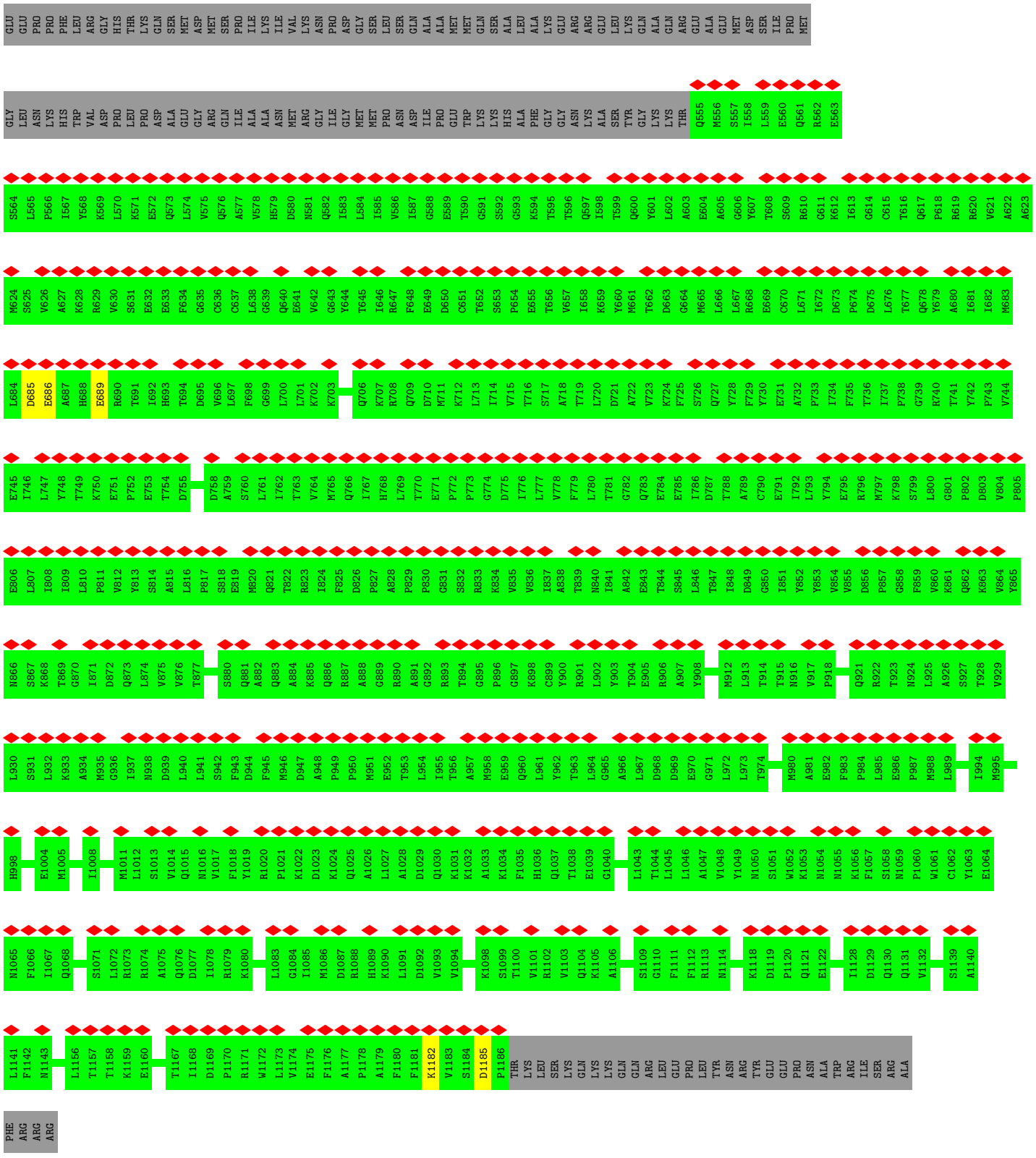


• Molecule 24: Peptidyl-prolyl cis-trans isomerase E



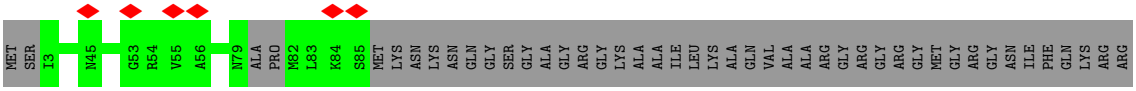
• Molecule 25: ATP-dependent RNA helicase DHX8



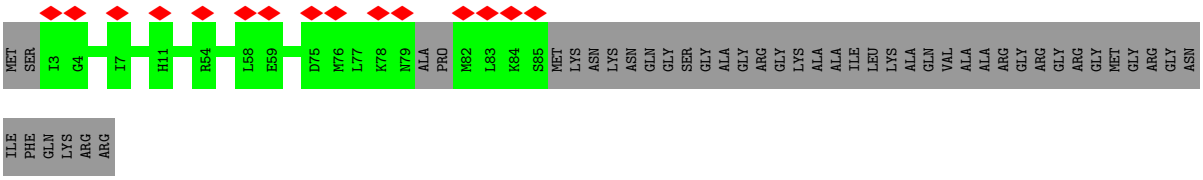


• Molecule 26: Small nuclear ribonucleoprotein Sm D3

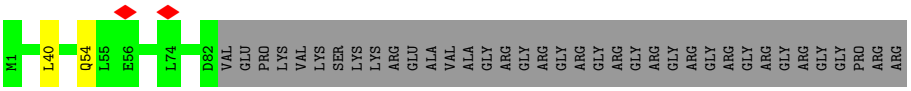
Chain a:



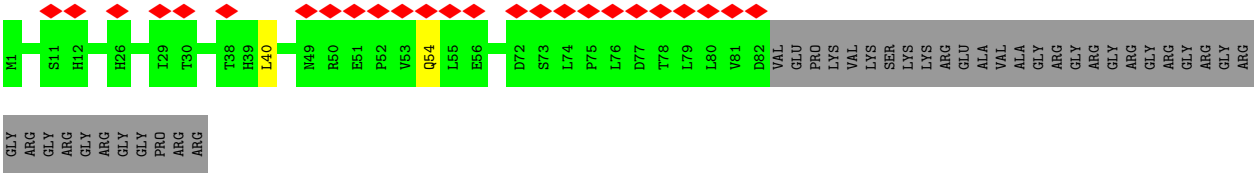
• Molecule 26: Small nuclear ribonucleoprotein Sm D3



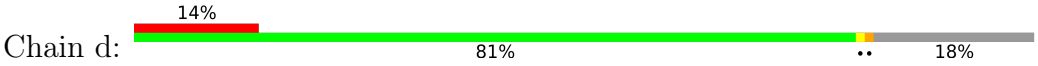
• Molecule 27: Small nuclear ribonucleoprotein Sm D1

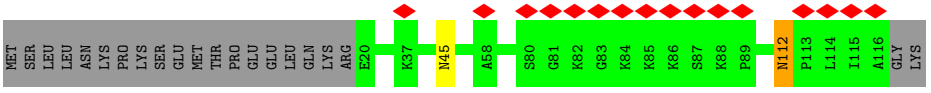


• Molecule 27: Small nuclear ribonucleoprotein Sm D1

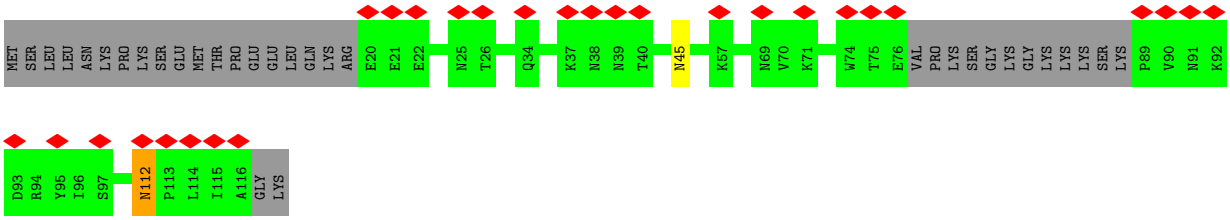


• Molecule 28: Small nuclear ribonucleoprotein Sm D2



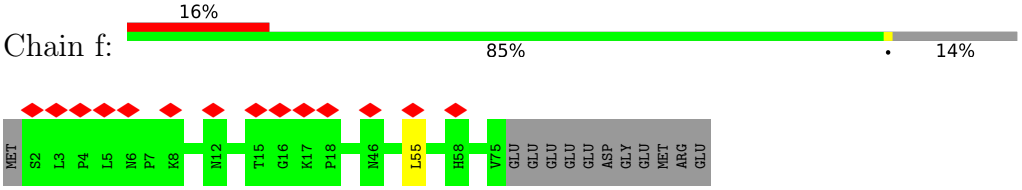


• Molecule 28: Small nuclear ribonucleoprotein Sm D2



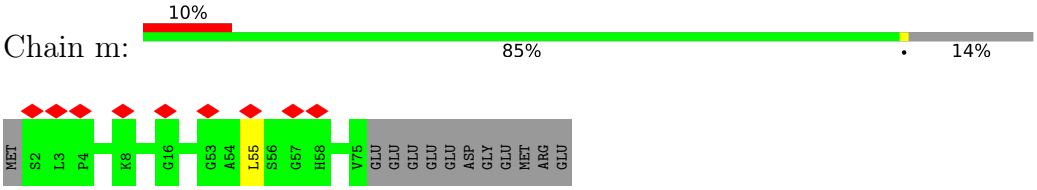
• Molecule 29: Small nuclear ribonucleoprotein F

Chain f:



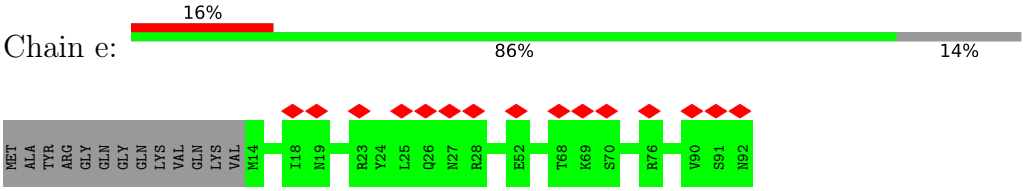
• Molecule 29: Small nuclear ribonucleoprotein F

Chain m:



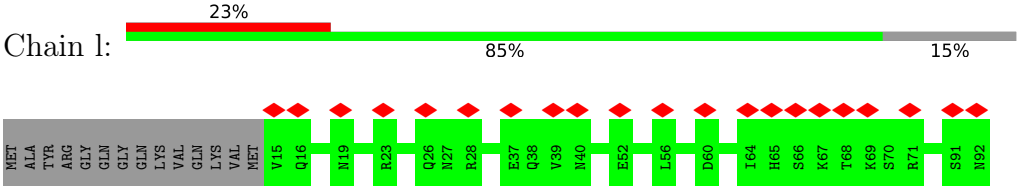
• Molecule 30: Small nuclear ribonucleoprotein E

Chain e:



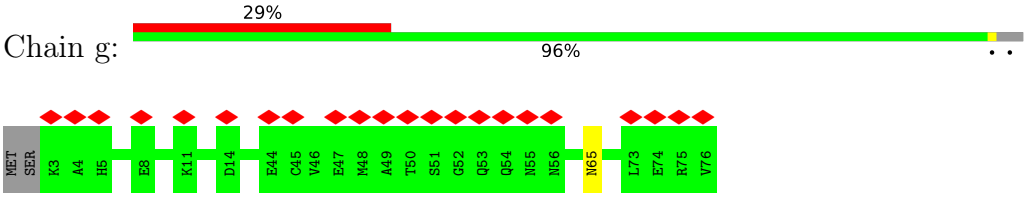
• Molecule 30: Small nuclear ribonucleoprotein E

Chain l:



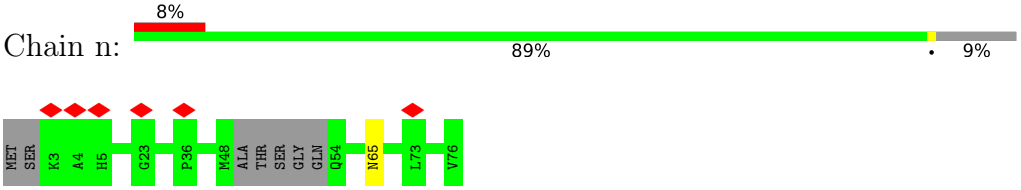
• Molecule 31: Small nuclear ribonucleoprotein G

Chain g:



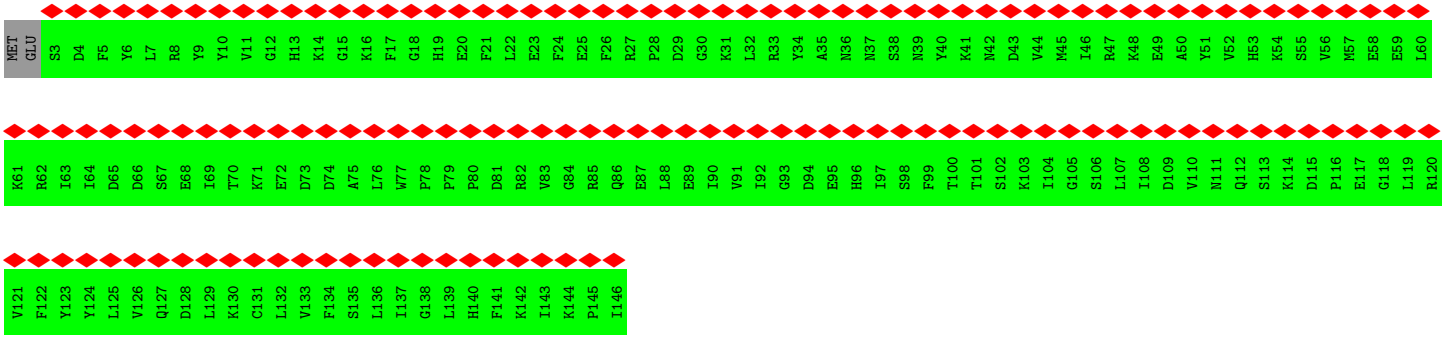
• Molecule 31: Small nuclear ribonucleoprotein G

Chain n:



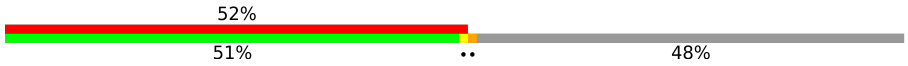
• Molecule 32: Protein mago nashi homolog

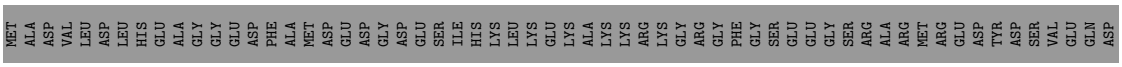
Chain v:

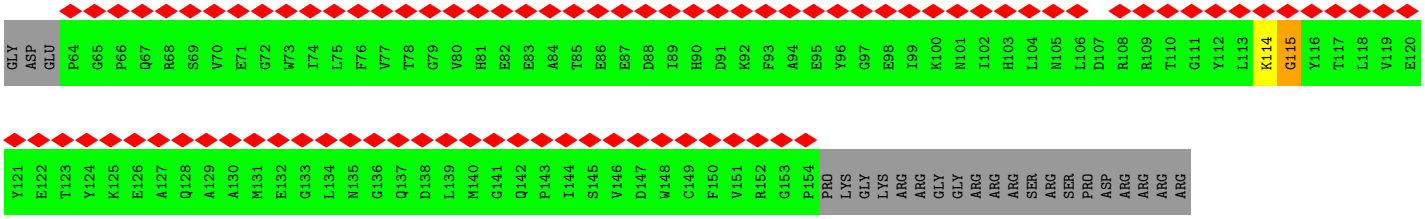


• Molecule 33: RNA-binding protein 8A

Chain w:

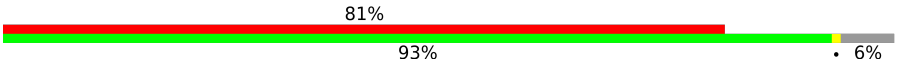


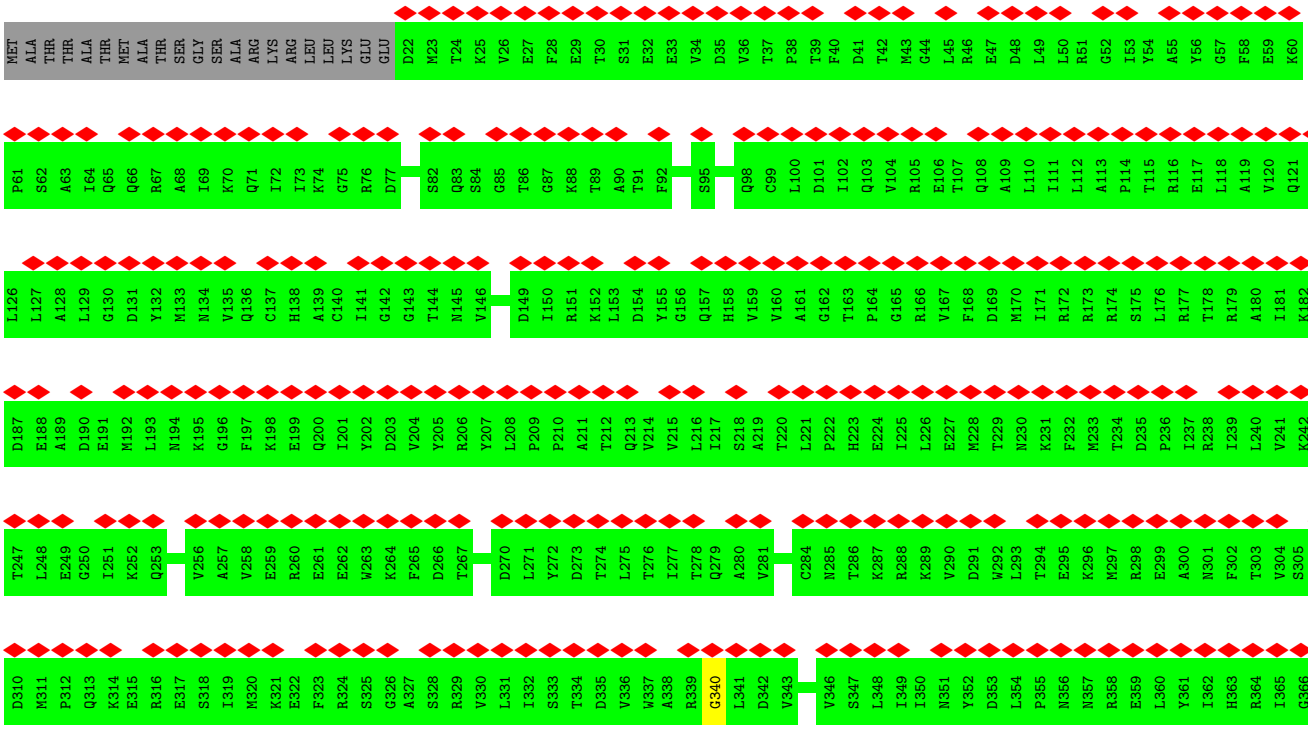


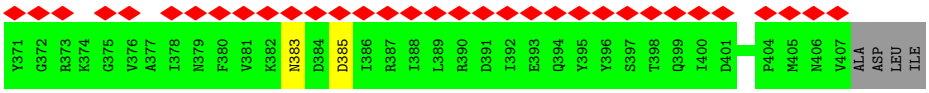


• Molecule 34: Eukaryotic initiation factor 4A-III

Chain u:

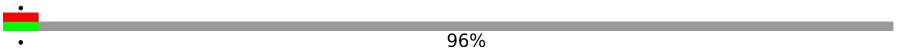


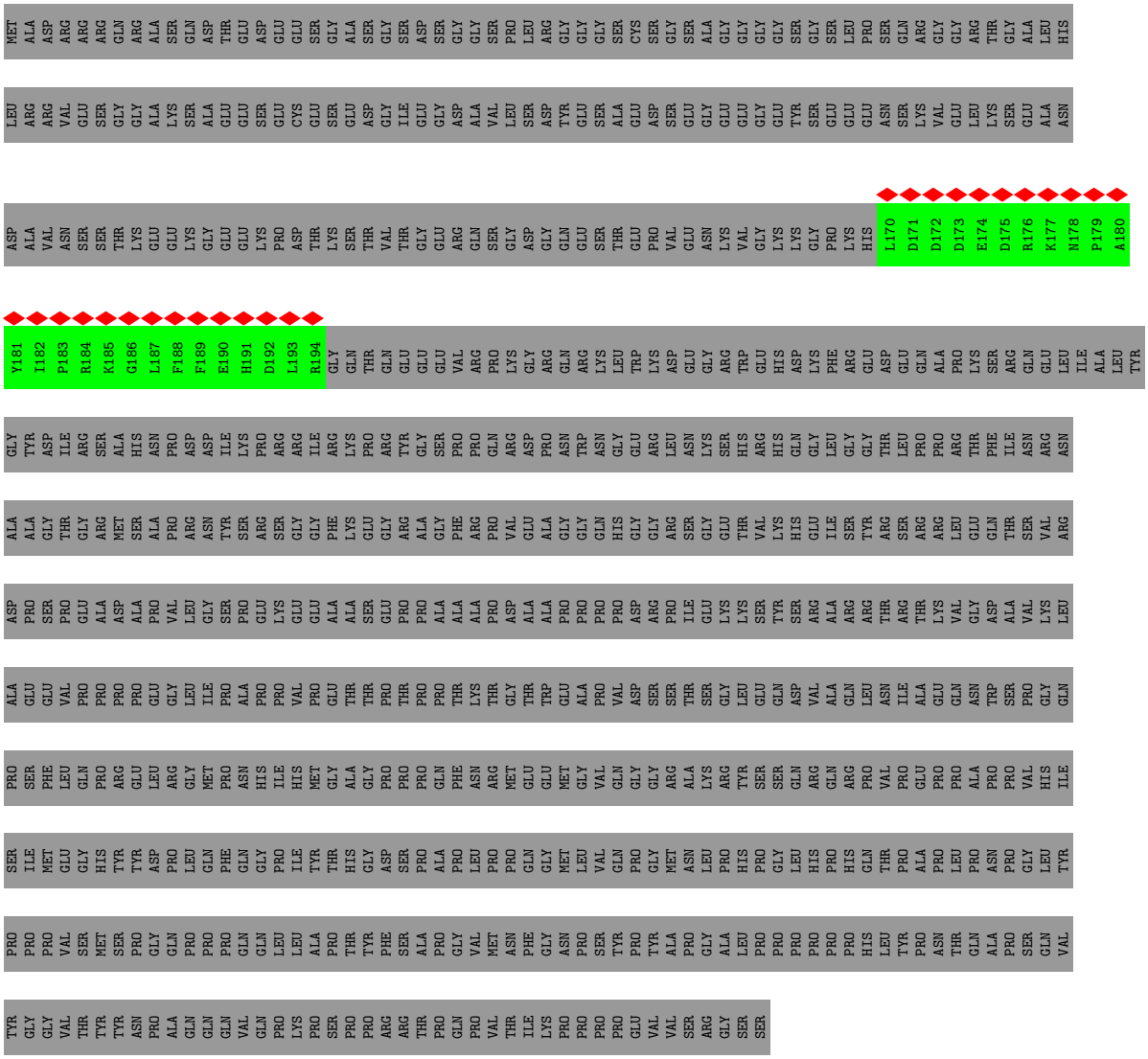




• Molecule 35: Protein CASC3

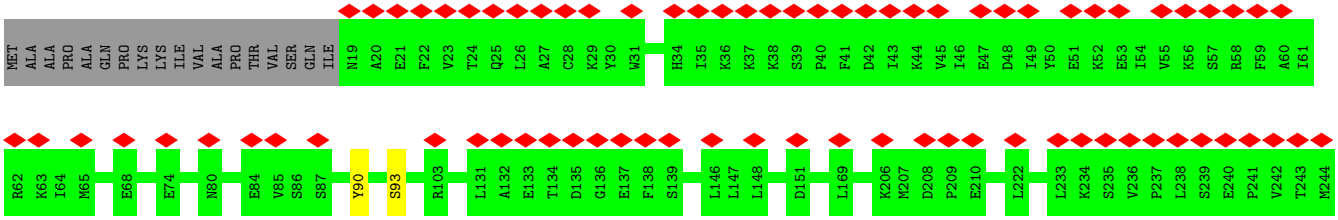
Chain x:

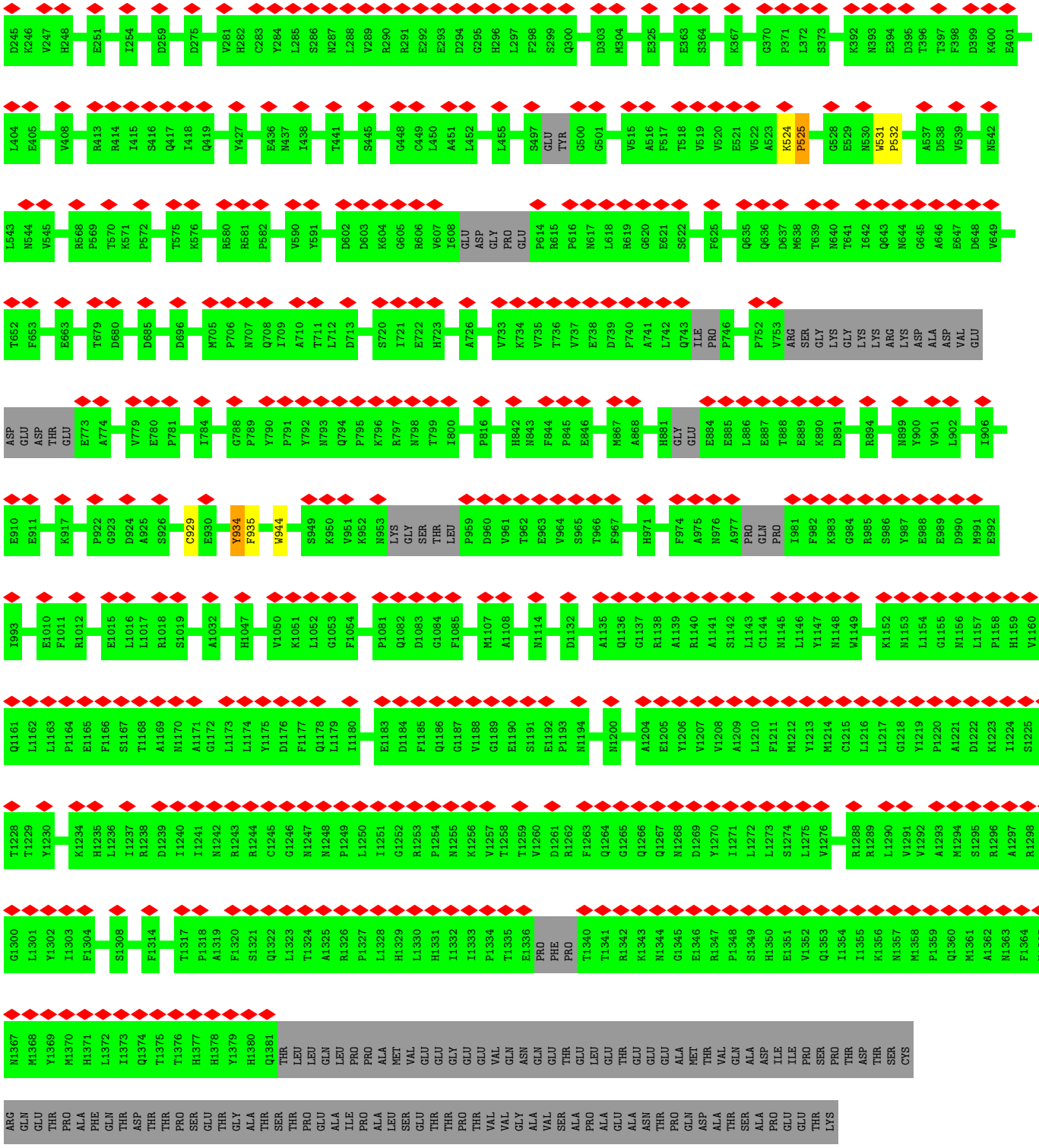




• Molecule 36: RNA helicase aquarius

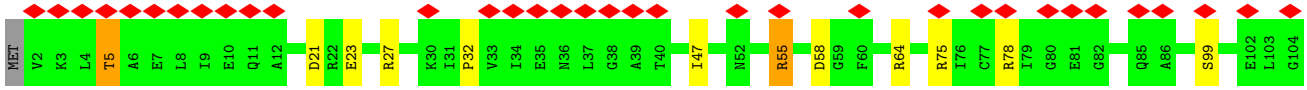
Chain Q:

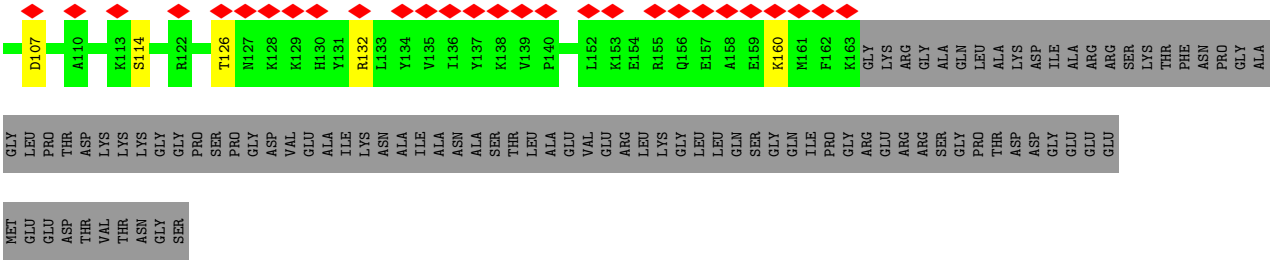




• Molecule 37: U2 small nuclear ribonucleoprotein A'

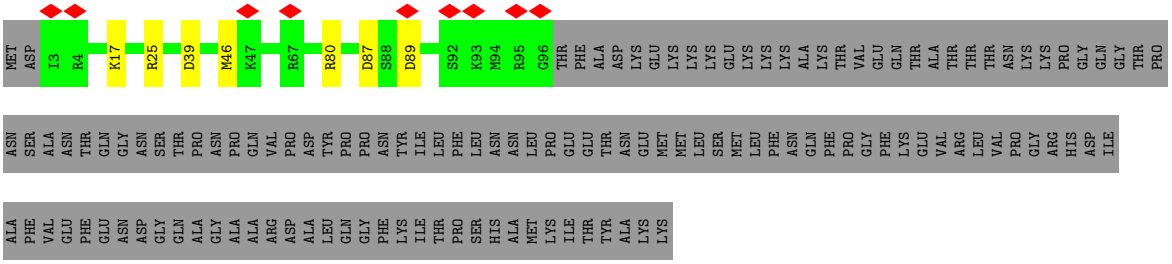
Chain o:





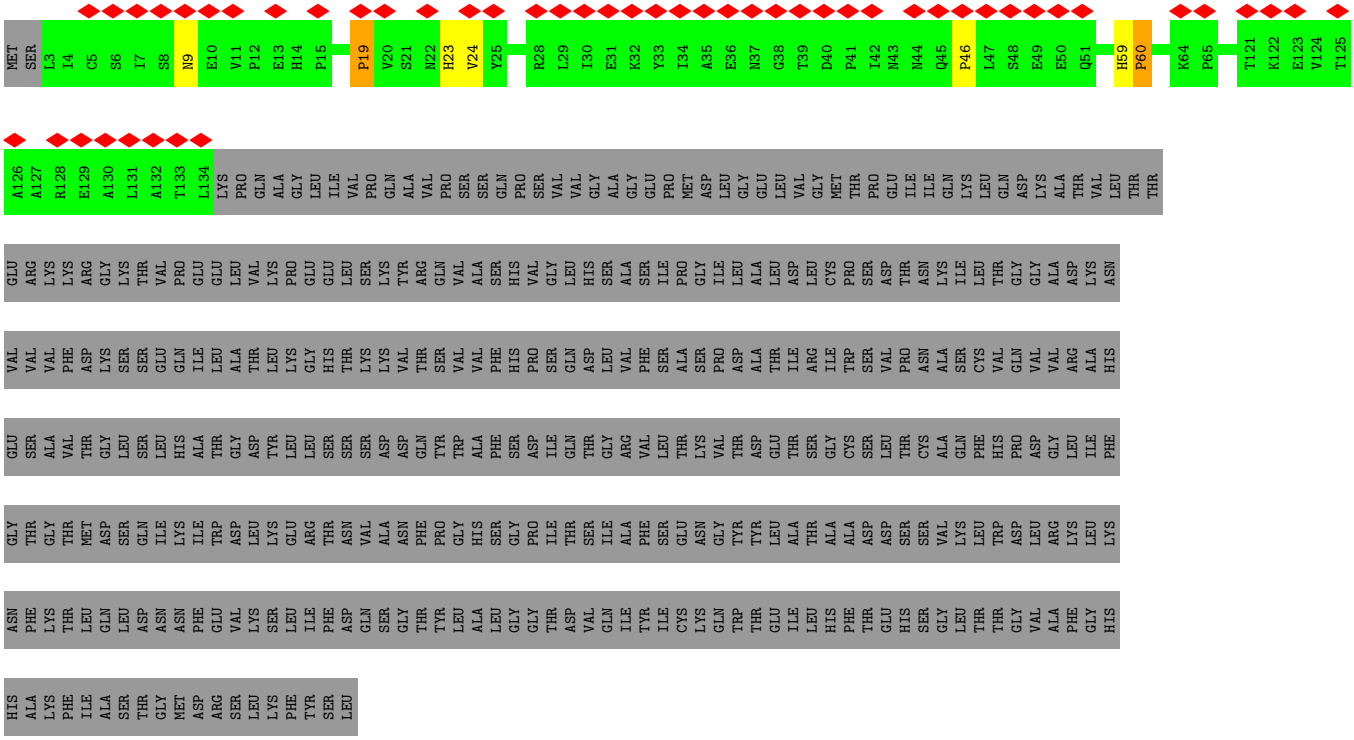
• Molecule 38: U2 small nuclear ribonucleoprotein B"

Chain p:



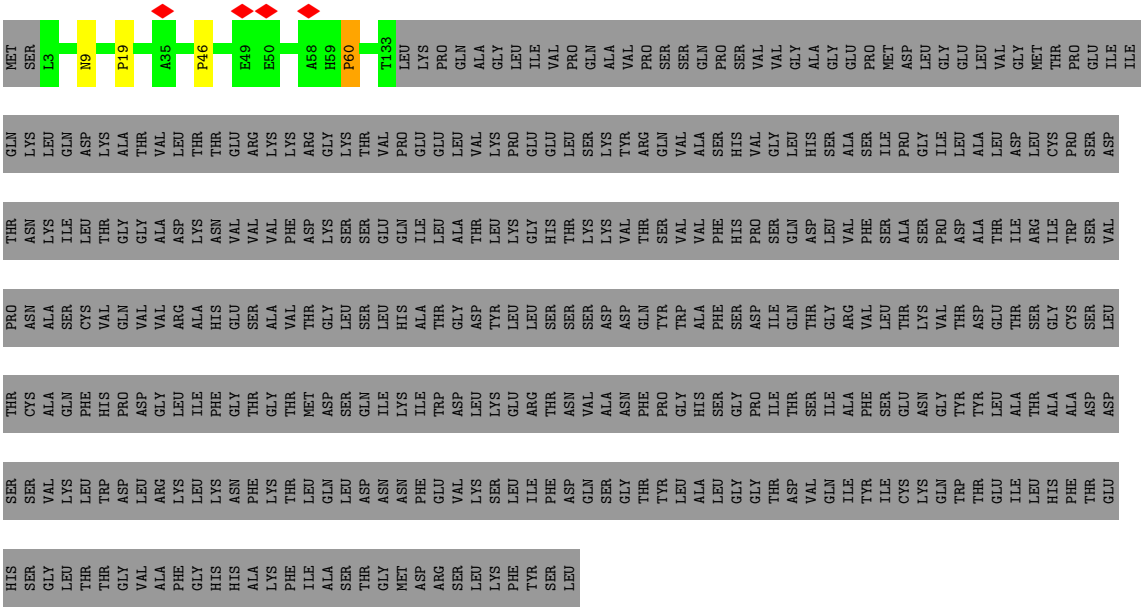
• Molecule 39: Pre-mRNA-processing factor 19

Chain q:



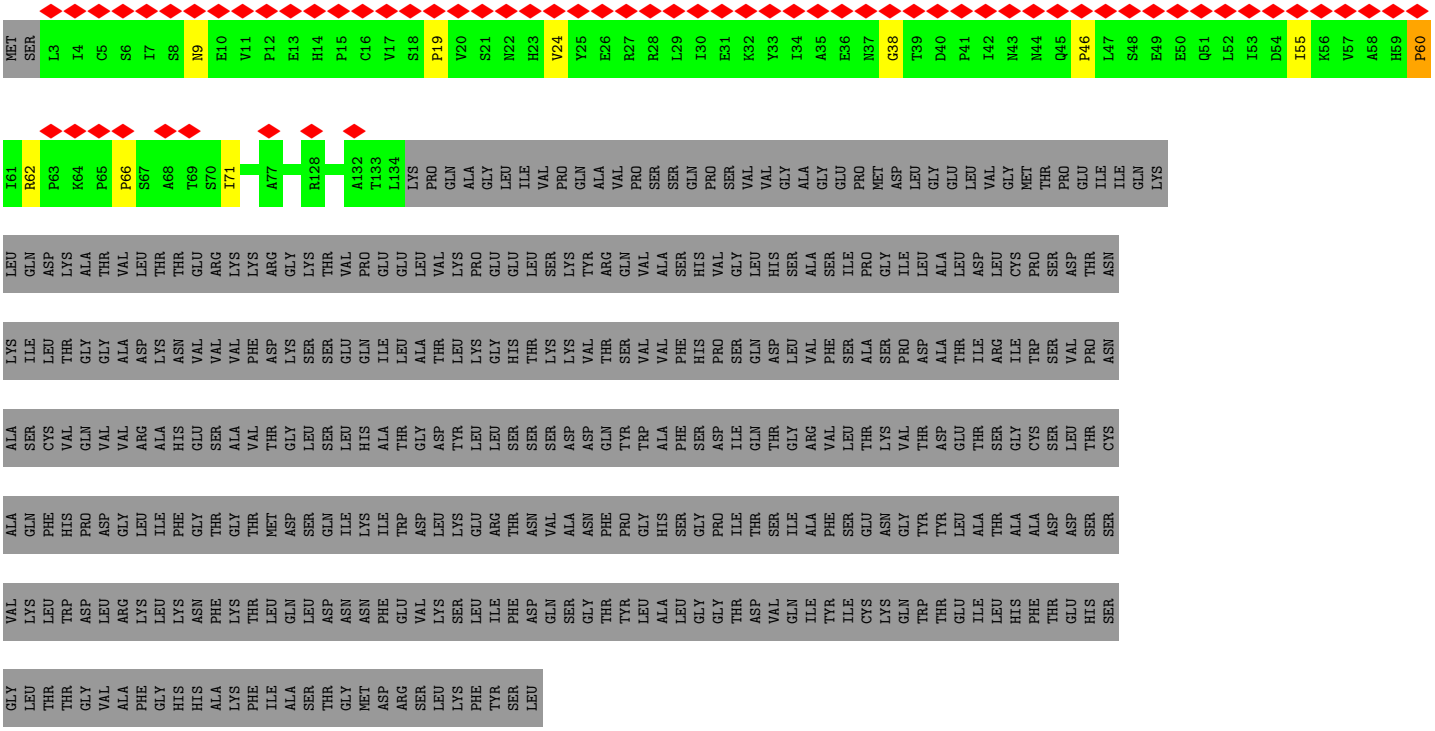
• Molecule 39: Pre-mRNA-processing factor 19

Chain r:



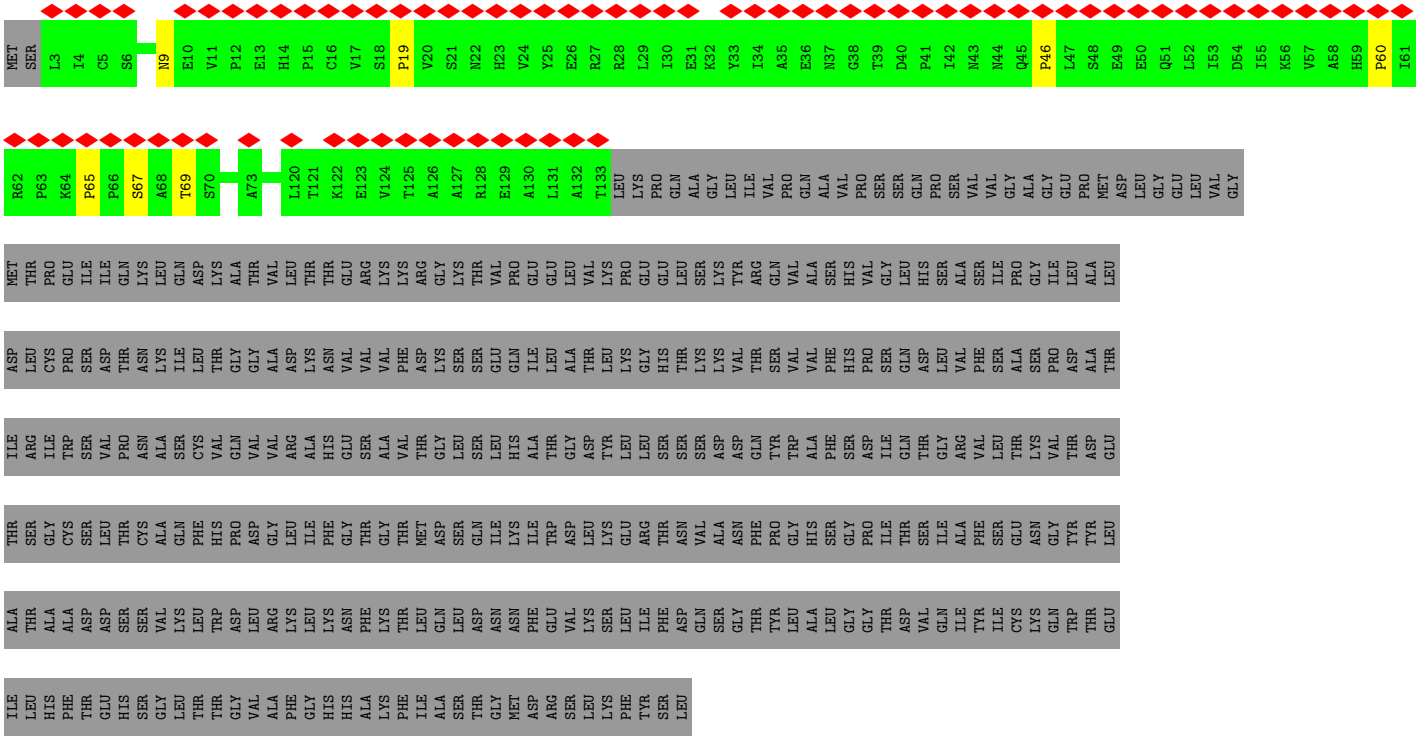
• Molecule 39: Pre-mRNA-processing factor 19

Chain s:



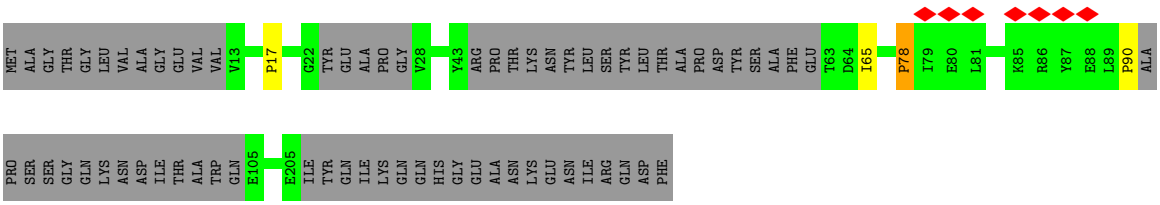
• Molecule 39: Pre-mRNA-processing factor 19

Chain t:



• Molecule 40: Pre-mRNA-splicing factor SPF27

Chain K:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	192274	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	50	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	3.470	Depositor
Minimum map value	-1.736	Depositor
Average map value	0.012	Depositor
Map value standard deviation	0.092	Depositor
Recommended contour level	0.38	Depositor
Map size (Å)	535.2, 535.2, 535.2	wwPDB
Map dimensions	400, 400, 400	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.338, 1.338, 1.338	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, ATP, ZN, IHP, GTP, SEP

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	A	0.81	21/16897 (0.1%)	0.87	44/22917 (0.2%)
2	C	0.61	1/6950 (0.0%)	0.76	8/9443 (0.1%)
3	E	0.57	0/2392	0.71	2/3242 (0.1%)
4	4	1.34	0/307	1.53	7/476 (1.5%)
5	G	0.62	3/1674 (0.2%)	1.27	24/2594 (0.9%)
6	J	0.60	2/3856 (0.1%)	0.64	8/5234 (0.2%)
7	L	0.48	0/3046	0.63	2/4115 (0.0%)
8	M	0.51	0/1119	0.64	0/1497
9	N	0.82	3/1210 (0.2%)	0.81	0/1622
10	O	0.59	1/2344 (0.0%)	0.69	3/3163 (0.1%)
11	P	0.65	0/967	0.92	3/1285 (0.2%)
12	R	0.69	1/2262 (0.0%)	0.91	9/3031 (0.3%)
13	S	0.52	0/1268	0.72	2/1714 (0.1%)
14	T	0.99	5/2519 (0.2%)	1.03	16/3433 (0.5%)
15	U	0.49	0/424	0.59	0/582
16	V	0.34	0/2642	0.56	1/3602 (0.0%)
17	W	0.47	0/4237	0.73	4/5723 (0.1%)
18	B	0.61	1/1970 (0.1%)	0.84	4/3060 (0.1%)
19	F	0.49	0/2323	0.78	2/3619 (0.1%)
20	H	0.95	26/3305 (0.8%)	1.63	112/5130 (2.2%)
21	b	0.55	0/797	0.81	0/1062
21	i	0.53	0/700	0.82	0/933
22	X	0.83	0/621	1.39	7/822 (0.9%)
23	I	0.44	0/3871	0.87	9/5283 (0.2%)
24	y	0.29	0/389	0.64	0/540
25	Y	0.55	1/3436 (0.0%)	0.77	4/4774 (0.1%)
26	a	0.47	0/646	0.68	0/867
26	h	0.47	0/639	0.68	0/857
27	c	0.55	0/657	0.77	0/888
27	j	0.56	0/657	0.77	0/888
28	d	0.69	0/786	0.86	0/1053
28	k	0.69	0/696	0.86	0/935

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
29	f	0.82	0/588	0.83	0/795
29	m	0.81	0/588	0.83	0/795
30	e	0.62	0/660	0.83	0/886
30	l	0.61	0/652	0.84	0/876
31	g	0.52	0/584	0.78	0/779
31	n	0.53	0/544	0.80	0/725
32	v	0.30	0/710	0.65	0/987
33	w	0.31	0/444	0.78	2/614 (0.3%)
34	u	0.32	0/1906	0.69	0/2653
35	x	0.34	0/123	0.70	0/170
36	Q	0.35	4/6565 (0.1%)	0.52	5/9143 (0.1%)
37	o	0.59	0/1299	1.63	17/1761 (1.0%)
38	p	0.56	0/774	1.35	6/1035 (0.6%)
39	q	0.35	0/658	0.63	3/919 (0.3%)
39	r	0.32	0/653	0.59	3/912 (0.3%)
39	s	0.34	0/658	0.66	3/919 (0.3%)
39	t	0.35	0/653	0.59	3/912 (0.3%)
40	K	0.37	0/768	0.48	2/1067 (0.2%)
All	All	0.63	69/94434 (0.1%)	0.86	315/130332 (0.2%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	15
2	C	0	5
6	J	0	3
9	N	0	3
10	O	0	1
11	P	0	3
12	R	0	6
17	W	0	1
23	I	0	5
28	d	0	1
28	k	0	1
36	Q	0	1
All	All	0	45

The worst 5 of 69 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	345	PRO	N-CA	12.23	1.68	1.47
36	Q	525	PRO	N-CA	11.95	1.67	1.47
2	C	92	PRO	N-CA	11.82	1.67	1.47
25	Y	426	LEU	C-N	8.48	1.50	1.34
5	G	21	A	O3'-P	-8.44	1.51	1.61

The worst 5 of 315 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1515	TRP	N-CA-CB	15.62	138.72	110.60
37	o	55	ARG	NE-CZ-NH1	15.18	127.89	120.30
1	A	772	CYS	CA-CB-SG	-13.83	89.10	114.00
37	o	55	ARG	CD-NE-CZ	13.69	142.76	123.60
1	A	856	LEU	C-N-CA	13.32	155.00	121.70

There are no chirality outliers.

5 of 45 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	186	GLU	Mainchain
1	A	196	ASP	Mainchain
1	A	377	GLU	Peptide
1	A	55	ASP	Peptide
1	A	73	HIS	Peptide

5.2 Too-close contacts (i)

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

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	1980/2335 (85%)	1807 (91%)	160 (8%)	13 (1%)	22 61
2	C	856/972 (88%)	790 (92%)	64 (8%)	2 (0%)	47 79
3	E	297/357 (83%)	276 (93%)	20 (7%)	1 (0%)	41 75
6	J	529/848 (62%)	491 (93%)	31 (6%)	7 (1%)	12 50
7	L	425/802 (53%)	408 (96%)	15 (4%)	2 (0%)	29 68
8	M	128/243 (53%)	117 (91%)	11 (9%)	0	100 100
9	N	141/144 (98%)	125 (89%)	14 (10%)	2 (1%)	11 48
10	O	283/420 (67%)	259 (92%)	23 (8%)	1 (0%)	34 71
11	P	107/229 (47%)	88 (82%)	16 (15%)	3 (3%)	5 34
12	R	274/536 (51%)	247 (90%)	23 (8%)	4 (2%)	10 47
13	S	157/166 (95%)	148 (94%)	9 (6%)	0	100 100
14	T	310/514 (60%)	278 (90%)	25 (8%)	7 (2%)	6 38
15	U	68/2752 (2%)	60 (88%)	8 (12%)	0	100 100
16	V	444/908 (49%)	431 (97%)	12 (3%)	1 (0%)	47 79
17	W	507/579 (88%)	433 (85%)	68 (13%)	6 (1%)	13 51
21	b	98/240 (41%)	93 (95%)	2 (2%)	3 (3%)	4 32
21	i	84/240 (35%)	82 (98%)	2 (2%)	0	100 100
22	X	67/254 (26%)	64 (96%)	3 (4%)	0	100 100
23	I	575/855 (67%)	562 (98%)	11 (2%)	2 (0%)	41 75
24	y	77/301 (26%)	74 (96%)	1 (1%)	2 (3%)	5 35
25	Y	667/1220 (55%)	642 (96%)	23 (3%)	2 (0%)	41 75
26	a	77/126 (61%)	76 (99%)	1 (1%)	0	100 100
26	h	77/126 (61%)	76 (99%)	1 (1%)	0	100 100
27	c	80/119 (67%)	77 (96%)	3 (4%)	0	100 100
27	j	80/119 (67%)	77 (96%)	3 (4%)	0	100 100
28	d	95/118 (80%)	91 (96%)	4 (4%)	0	100 100
28	k	81/118 (69%)	78 (96%)	3 (4%)	0	100 100
29	f	72/86 (84%)	69 (96%)	3 (4%)	0	100 100
29	m	72/86 (84%)	69 (96%)	3 (4%)	0	100 100
30	e	77/92 (84%)	76 (99%)	1 (1%)	0	100 100
30	l	76/92 (83%)	75 (99%)	1 (1%)	0	100 100
31	g	72/76 (95%)	70 (97%)	2 (3%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
31	n	65/76 (86%)	63 (97%)	2 (3%)	0	100 100
32	v	142/146 (97%)	138 (97%)	4 (3%)	0	100 100
33	w	89/174 (51%)	87 (98%)	1 (1%)	1 (1%)	14 53
34	u	384/411 (93%)	372 (97%)	9 (2%)	3 (1%)	19 59
35	x	23/703 (3%)	22 (96%)	1 (4%)	0	100 100
36	Q	1308/1485 (88%)	1278 (98%)	27 (2%)	3 (0%)	47 79
37	o	160/255 (63%)	146 (91%)	12 (8%)	2 (1%)	12 50
38	p	92/225 (41%)	90 (98%)	2 (2%)	0	100 100
39	q	130/504 (26%)	117 (90%)	7 (5%)	6 (5%)	2 23
39	r	129/504 (26%)	118 (92%)	9 (7%)	2 (2%)	9 46
39	s	130/504 (26%)	114 (88%)	8 (6%)	8 (6%)	1 17
39	t	129/504 (26%)	116 (90%)	9 (7%)	4 (3%)	4 32
40	K	147/225 (65%)	139 (95%)	5 (3%)	3 (2%)	7 41
All	All	11861/21789 (54%)	11109 (94%)	662 (6%)	90 (1%)	24 59

5 of 90 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	167	PRO
1	A	857	ASN
1	A	942	PRO
1	A	1831	LYS
6	J	217	GLU

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	1780/2108 (84%)	1702 (96%)	78 (4%)	28 63
2	C	761/866 (88%)	740 (97%)	21 (3%)	43 72
3	E	256/300 (85%)	255 (100%)	1 (0%)	91 97

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
6	J	241/751 (32%)	236 (98%)	5 (2%)	53 78
7	L	218/709 (31%)	209 (96%)	9 (4%)	30 64
8	M	117/209 (56%)	113 (97%)	4 (3%)	37 69
9	N	130/130 (100%)	123 (95%)	7 (5%)	22 57
10	O	255/361 (71%)	254 (100%)	1 (0%)	91 97
11	P	101/203 (50%)	94 (93%)	7 (7%)	15 49
12	R	236/457 (52%)	213 (90%)	23 (10%)	8 36
13	S	129/134 (96%)	128 (99%)	1 (1%)	81 91
14	T	268/441 (61%)	253 (94%)	15 (6%)	21 56
15	U	21/2432 (1%)	20 (95%)	1 (5%)	25 60
16	V	98/838 (12%)	95 (97%)	3 (3%)	40 71
17	W	448/502 (89%)	434 (97%)	14 (3%)	40 71
21	b	83/177 (47%)	81 (98%)	2 (2%)	49 75
21	i	77/177 (44%)	75 (97%)	2 (3%)	46 74
22	X	68/230 (30%)	46 (68%)	22 (32%)	0 2
23	I	198/749 (26%)	185 (93%)	13 (7%)	16 51
25	Y	32/1085 (3%)	26 (81%)	6 (19%)	1 9
26	a	72/101 (71%)	72 (100%)	0	100 100
26	h	70/101 (69%)	70 (100%)	0	100 100
27	c	77/101 (76%)	75 (97%)	2 (3%)	46 74
27	j	77/101 (76%)	75 (97%)	2 (3%)	46 74
28	d	90/110 (82%)	88 (98%)	2 (2%)	52 77
28	k	80/110 (73%)	78 (98%)	2 (2%)	47 75
29	f	63/74 (85%)	62 (98%)	1 (2%)	62 83
29	m	63/74 (85%)	62 (98%)	1 (2%)	62 83
30	e	74/84 (88%)	74 (100%)	0	100 100
30	l	73/84 (87%)	73 (100%)	0	100 100
31	g	64/66 (97%)	63 (98%)	1 (2%)	62 83
31	n	59/66 (89%)	58 (98%)	1 (2%)	60 82
37	o	139/218 (64%)	135 (97%)	4 (3%)	42 72
38	p	82/195 (42%)	79 (96%)	3 (4%)	34 66

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	6600/14344 (46%)	6346 (96%)	254 (4%)	36 66

5 of 254 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
8	M	198	ARG
23	I	724	LEU
12	R	307	GLN
23	I	694	ILE
28	d	112	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 92 such sidechains are listed below:

Mol	Chain	Res	Type
13	S	87	HIS
17	W	549	HIS
13	S	126	HIS
17	W	71	HIS
22	X	72	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
18	B	82/117 (70%)	17 (20%)	4 (4%)
19	F	96/107 (89%)	45 (46%)	16 (16%)
20	H	133/188 (70%)	34 (25%)	10 (7%)
4	4	13/46 (28%)	8 (61%)	3 (23%)
5	G	80/174 (45%)	63 (78%)	19 (23%)
All	All	404/632 (63%)	167 (41%)	52 (12%)

5 of 167 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
4	4	-12	G
4	4	-11	G
4	4	-10	C
4	4	-9	C
4	4	-7	C

5 of 52 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
19	F	25	C
19	F	50	A
20	H	156	U
19	F	26	U
19	F	35	A

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

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

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
12	SEP	R	232	12	8,9,10	1.55	1 (12%)	8,12,14	1.71	2 (25%)
12	SEP	R	224	12	8,9,10	1.69	3 (37%)	8,12,14	2.06	3 (37%)

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
12	SEP	R	232	12	-	1/5/8/10	-
12	SEP	R	224	12	-	0/5/8/10	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
12	R	232	SEP	P-O1P	3.39	1.61	1.50
12	R	224	SEP	P-OG	-2.84	1.51	1.60
12	R	224	SEP	CB-CA	-2.13	1.45	1.52
12	R	224	SEP	P-O2P	-2.05	1.46	1.54

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
12	R	224	SEP	OG-CB-CA	-4.52	103.74	108.14
12	R	232	SEP	P-OG-CB	-4.07	107.09	118.30
12	R	232	SEP	OG-CB-CA	2.34	110.42	108.14
12	R	224	SEP	O3P-P-O2P	2.30	116.42	107.64
12	R	224	SEP	O3P-P-OG	-2.14	101.03	106.73

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
12	R	232	SEP	N-CA-CB-OG

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 18 ligands modelled in this entry, 15 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	$\#\|Z\| > 2$
45	ATP	Q	1501	43	26,33,33	1.72	8 (30%)	31,52,52	1.85	10 (32%)
42	GTP	C	1500	43	26,34,34	1.50	3 (11%)	32,54,54	1.95	7 (21%)
41	IHP	A	3000	-	36,36,36	0.72	0	54,60,60	1.05	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
45	ATP	Q	1501	43	-	4/18/38/38	0/3/3/3
42	GTP	C	1500	43	-	1/18/38/38	0/3/3/3
41	IHP	A	3000	-	-	3/30/54/54	0/1/1/1

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
42	C	1500	GTP	C5-C6	-4.84	1.37	1.47
45	Q	1501	ATP	C2'-C1'	-3.63	1.48	1.53
45	Q	1501	ATP	C4-N3	3.47	1.40	1.35
45	Q	1501	ATP	C6-N6	3.34	1.46	1.34
45	Q	1501	ATP	C2'-C3'	-2.76	1.45	1.53

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
42	C	1500	GTP	PA-O3A-PB	-5.63	113.49	132.83
45	Q	1501	ATP	PB-O3B-PG	-5.46	114.08	132.83
42	C	1500	GTP	PB-O3B-PG	-4.37	117.82	132.83
45	Q	1501	ATP	N3-C2-N1	-4.17	122.16	128.68
42	C	1500	GTP	C5-C6-N1	3.72	120.52	113.95

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
41	A	3000	IHP	C3-O13-P3-O43
42	C	1500	GTP	O4'-C4'-C5'-O5'
45	Q	1501	ATP	C5'-O5'-PA-O1A
45	Q	1501	ATP	C5'-O5'-PA-O2A
41	A	3000	IHP	C3-C4-O14-P4

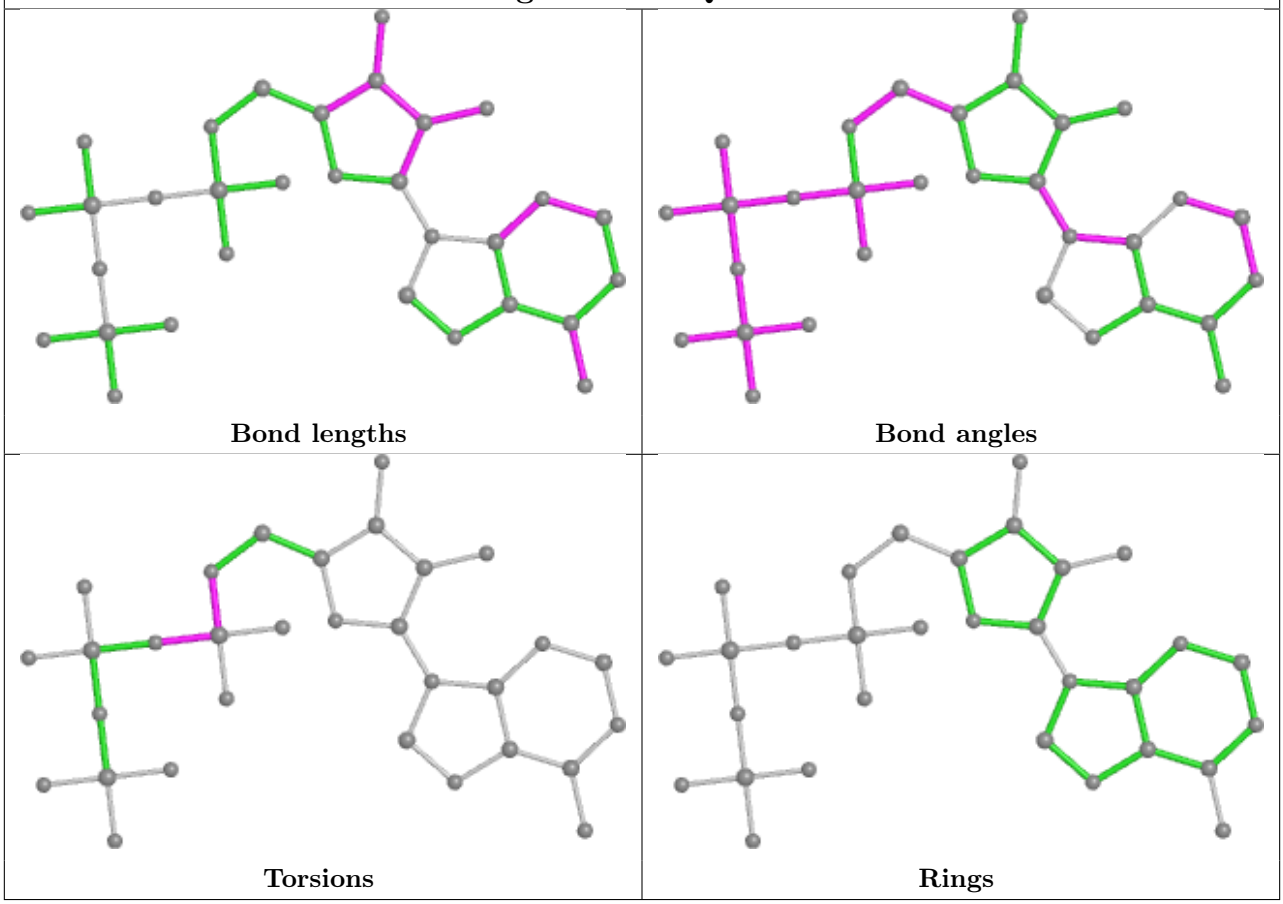
There are no ring outliers.

No monomer is involved in short contacts.

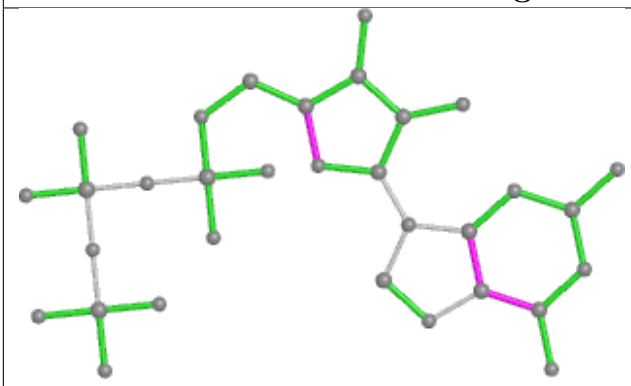
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

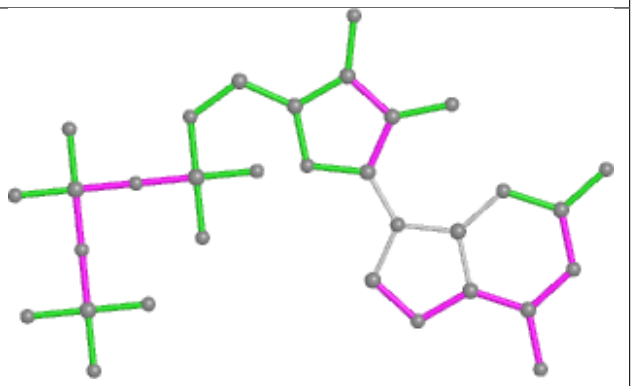
Ligand ATP Q 1501



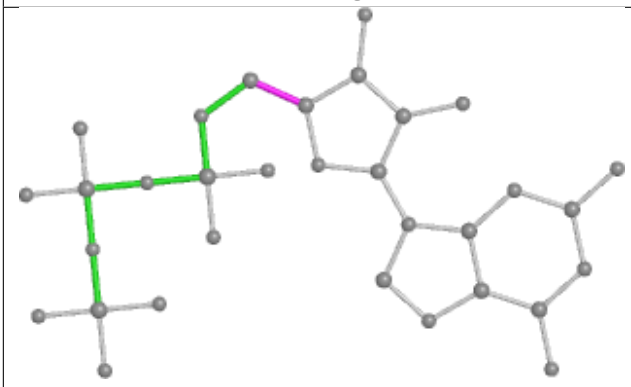
Ligand GTP C 1500



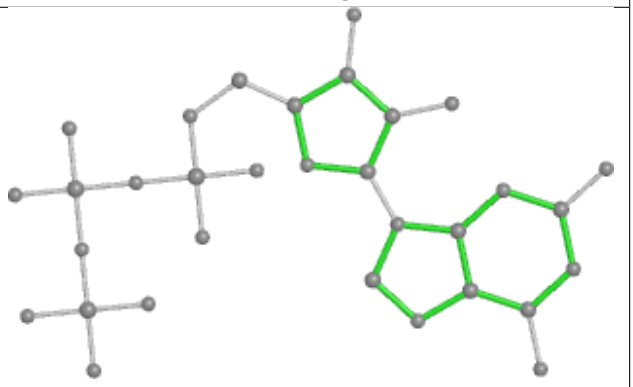
Bond lengths



Bond angles

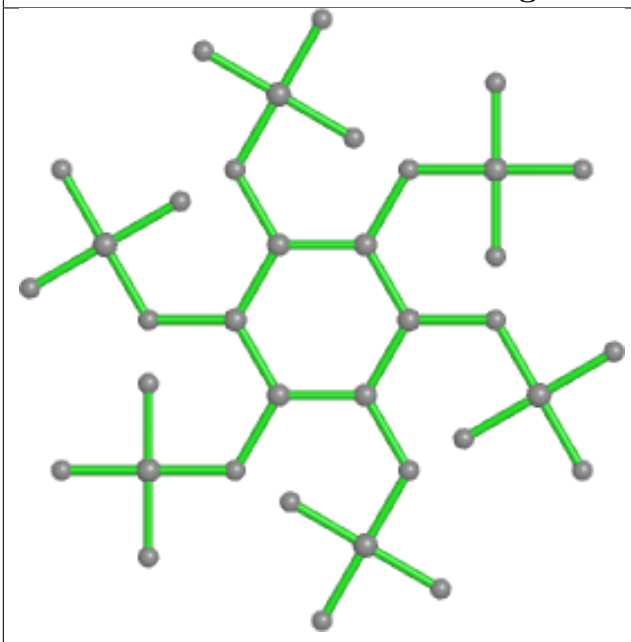


Torsions

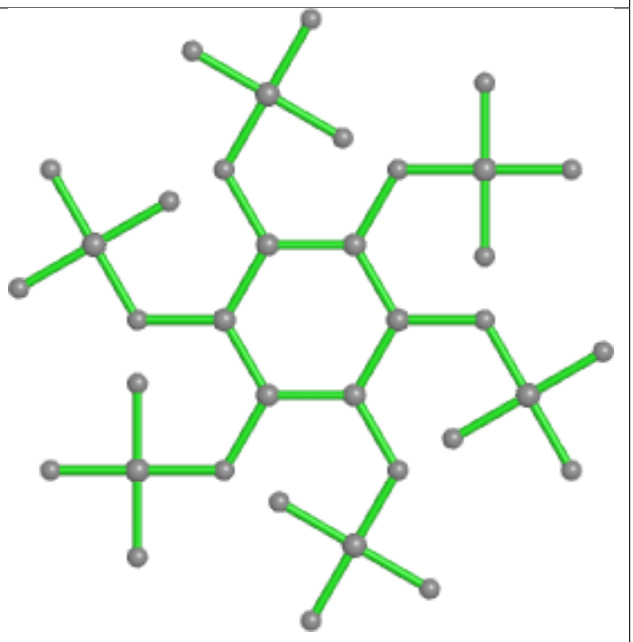


Rings

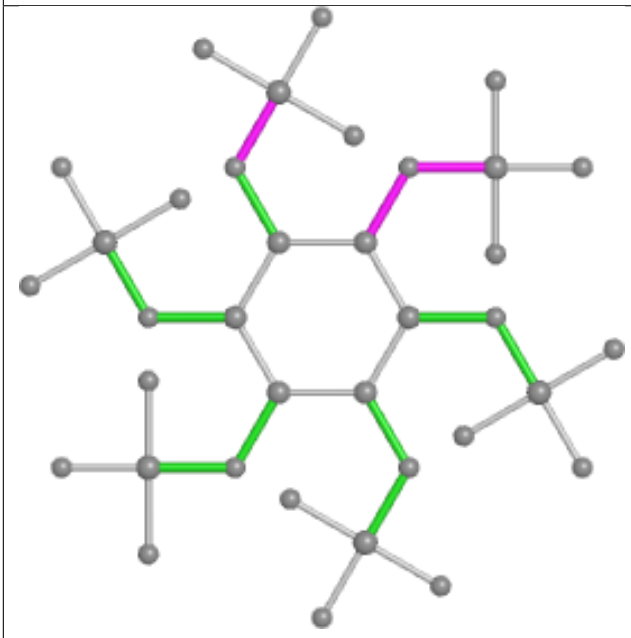
Ligand IHP A 3000



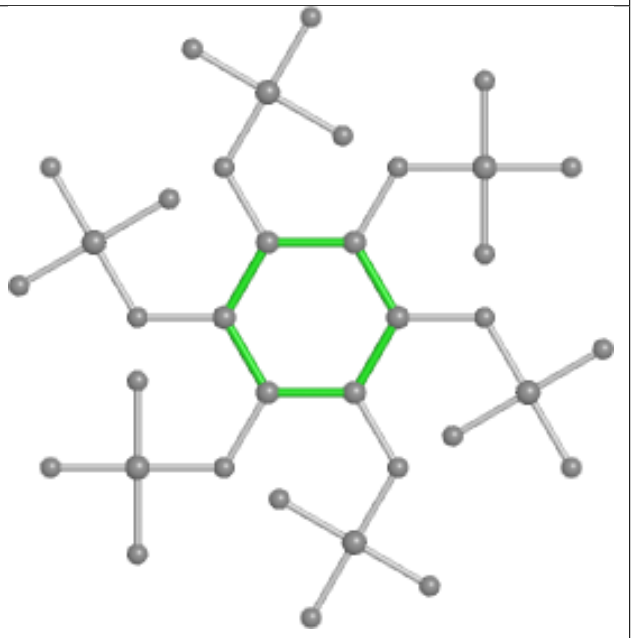
Bond lengths



Bond angles



Torsions



Rings

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.

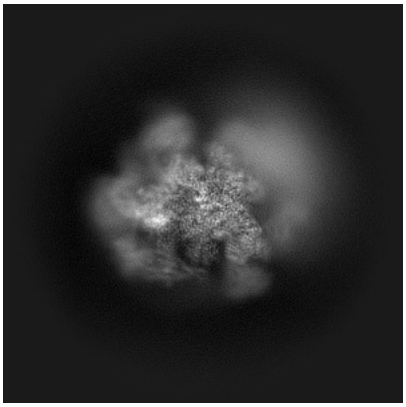
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-32317. These allow visual inspection of the internal detail of the map and identification of artifacts.

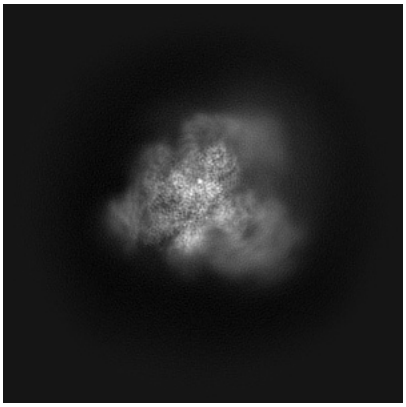
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

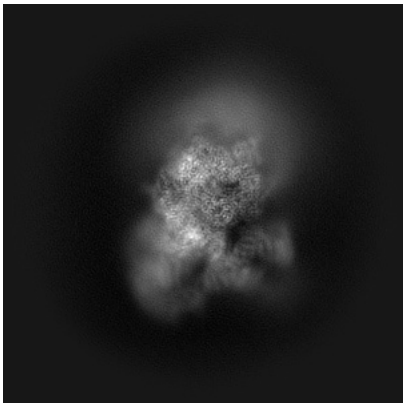
6.1.1 Primary map



X



Y

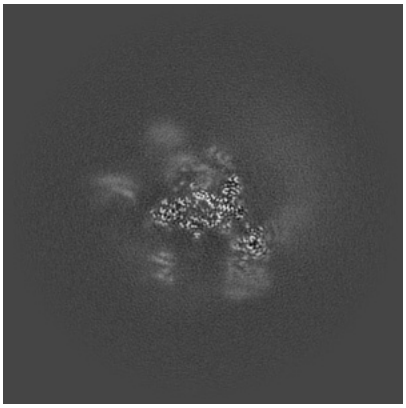


Z

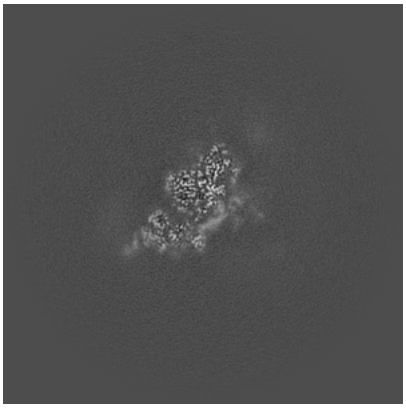
The images above show the map projected in three orthogonal directions.

6.2 Central slices

6.2.1 Primary map



X Index: 200



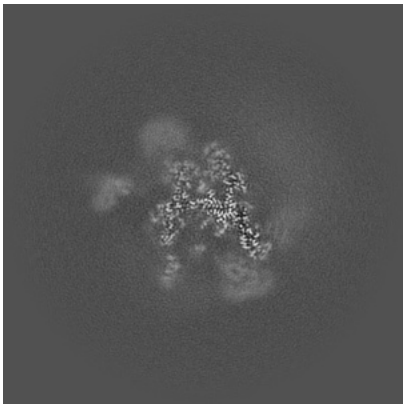
Y Index: 200

Z Index: 200

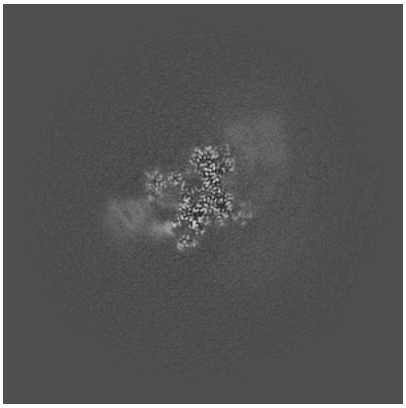
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

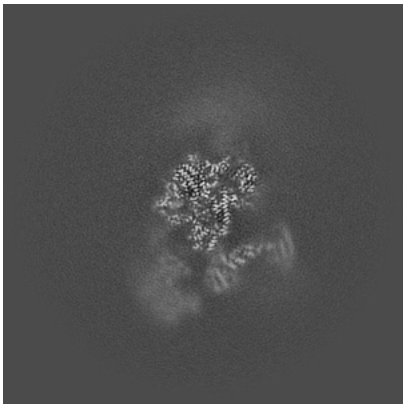
6.3.1 Primary map



X Index: 193



Y Index: 227



Z Index: 197

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 0.38. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

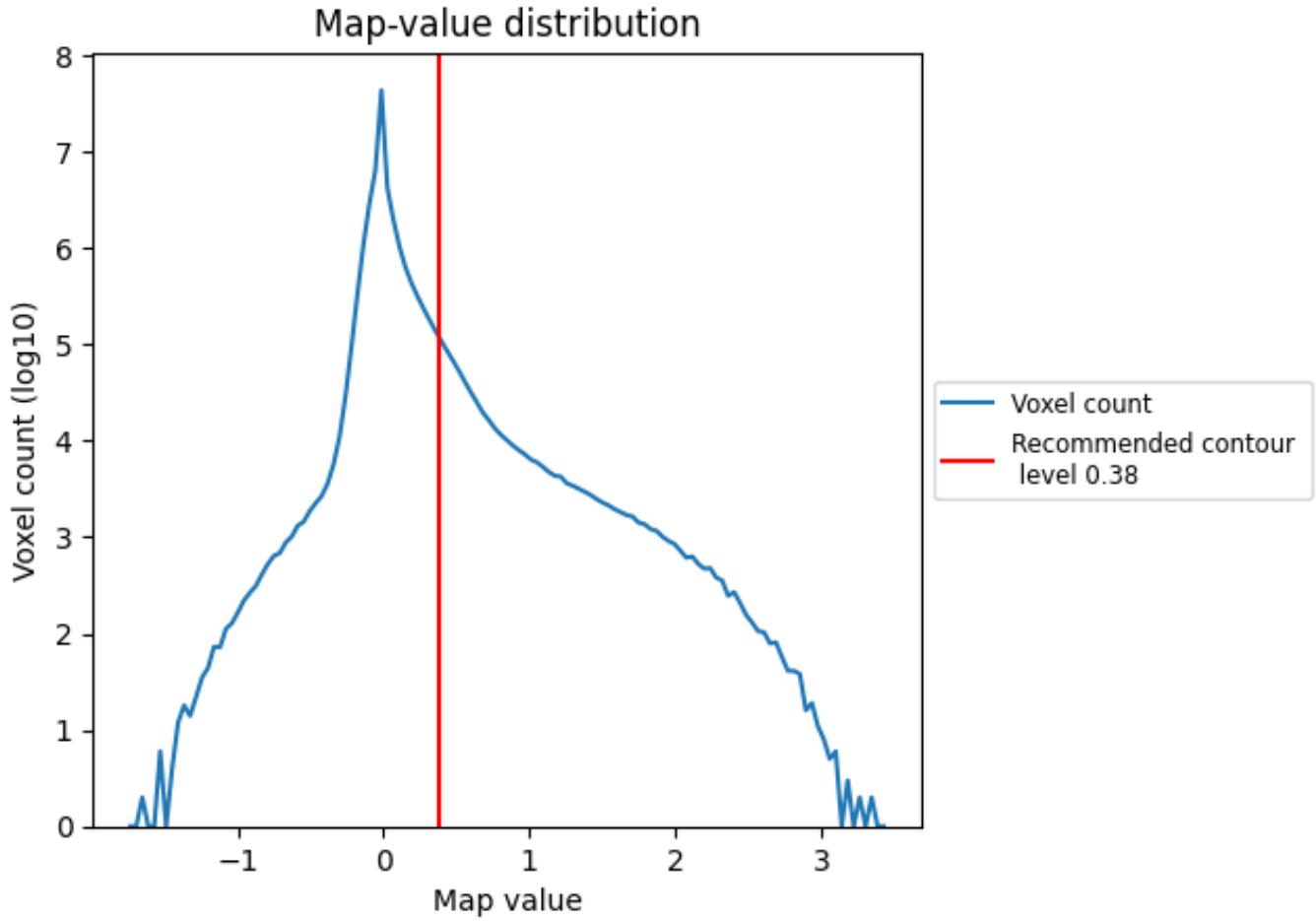
6.5 Mask visualisation (i)

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

7 Map analysis

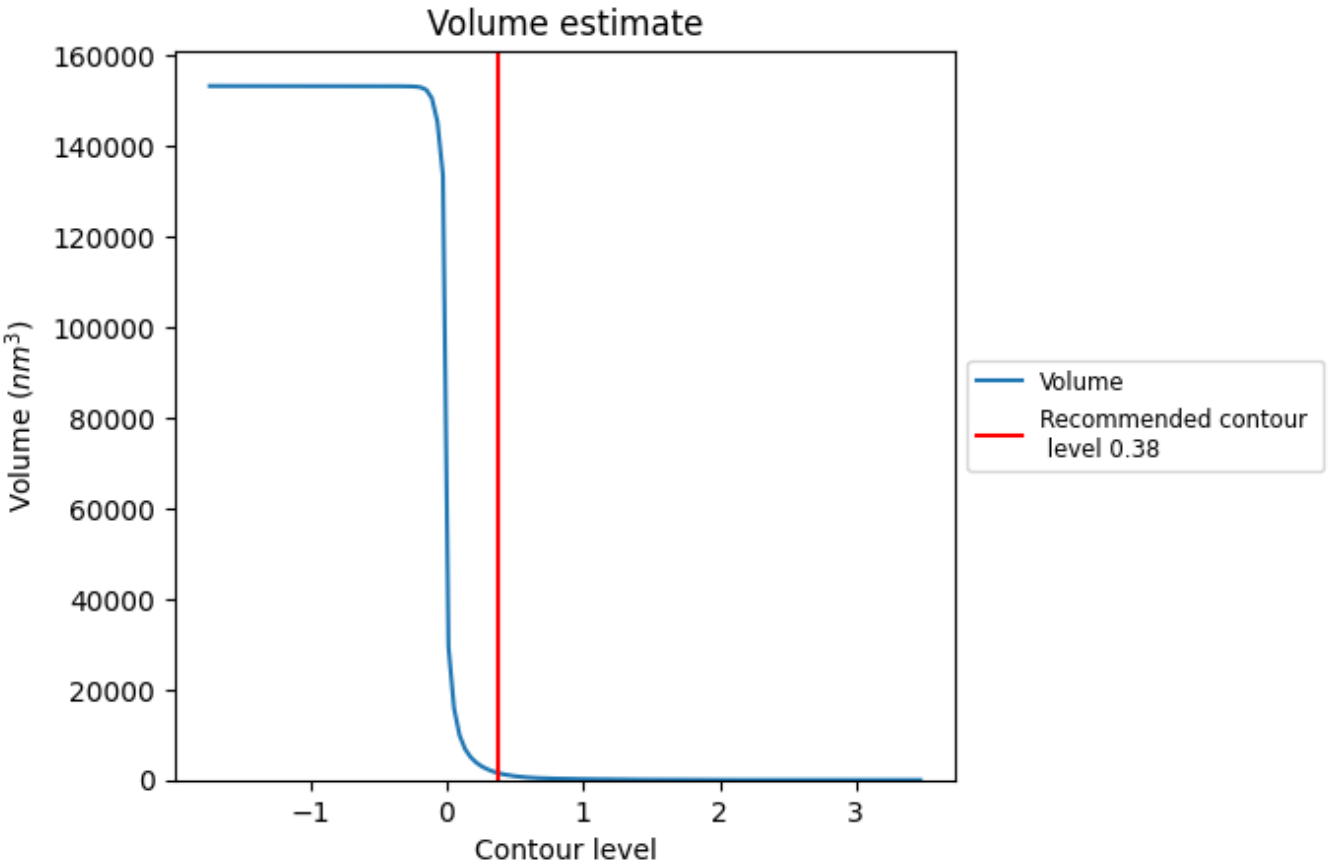
This section contains the results of statistical analysis of the map.

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.

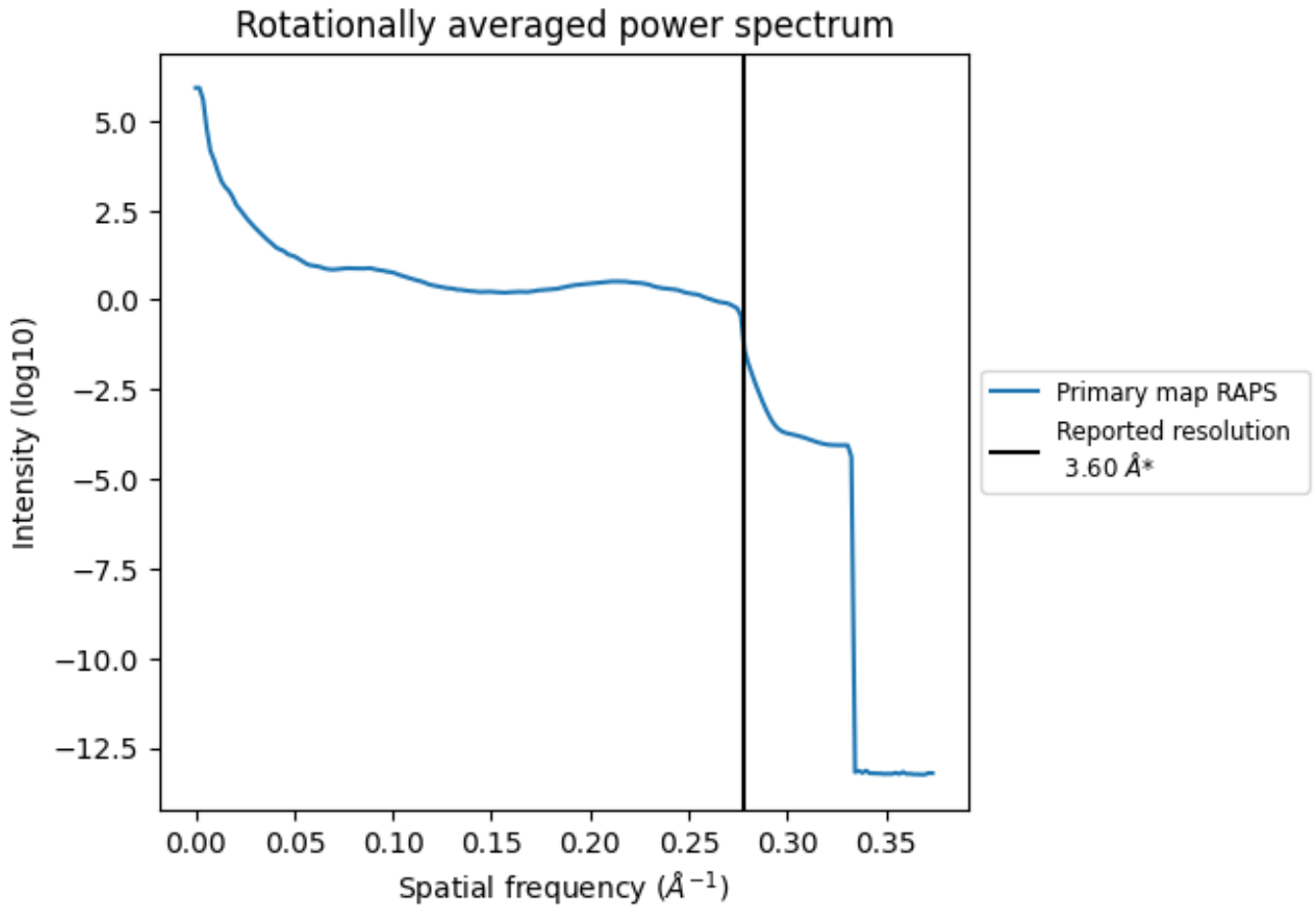
7.2 Volume estimate (i)



The volume at the recommended contour level is 1511 nm^3; this corresponds to an approximate mass of 1365 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 0.278 $Å^{-1}$

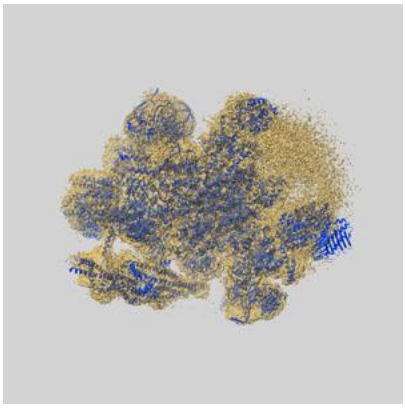
8 Fourier-Shell correlation

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

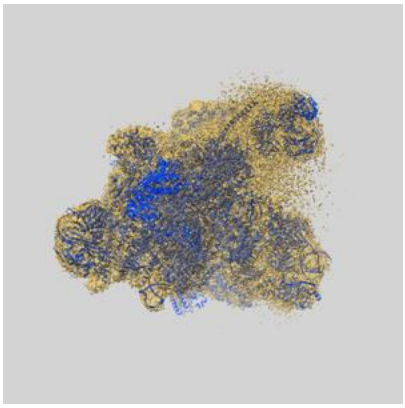
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-32317 and PDB model 7W59. Per-residue inclusion information can be found in section 3 on page 14.

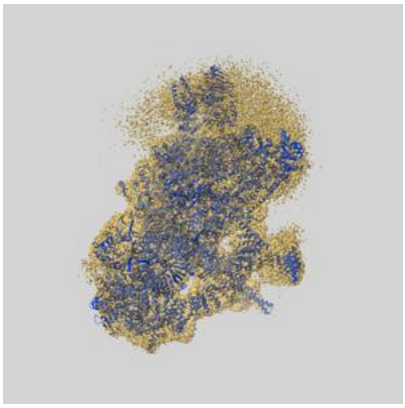
9.1 Map-model overlay



X



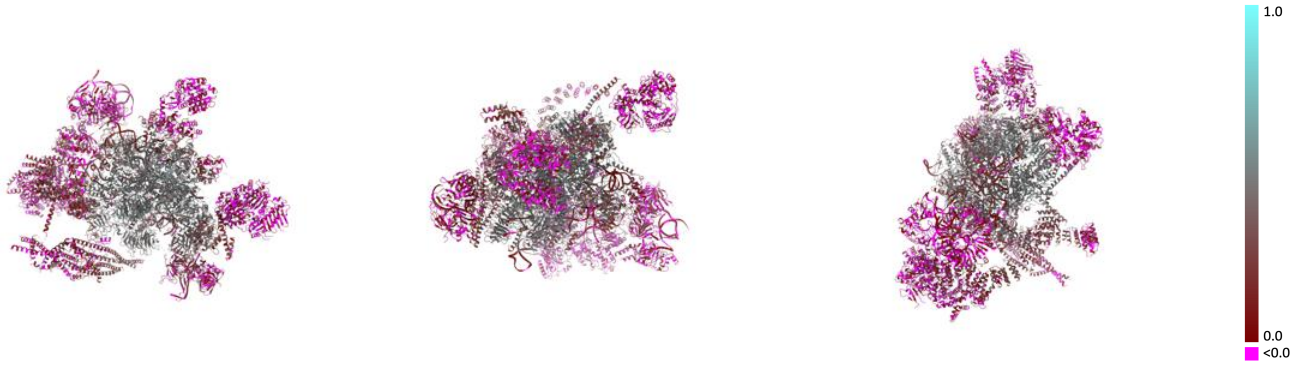
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.38 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.

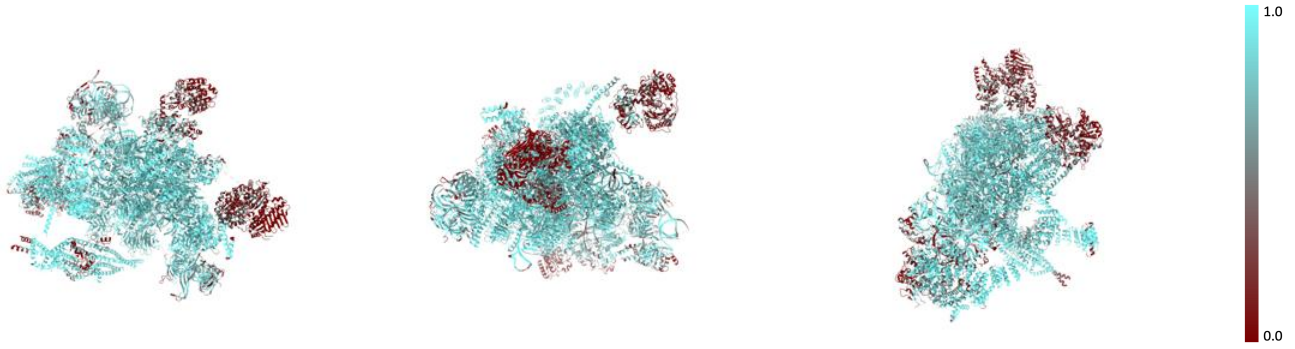
9.2 Q-score mapped to coordinate model (i)



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.

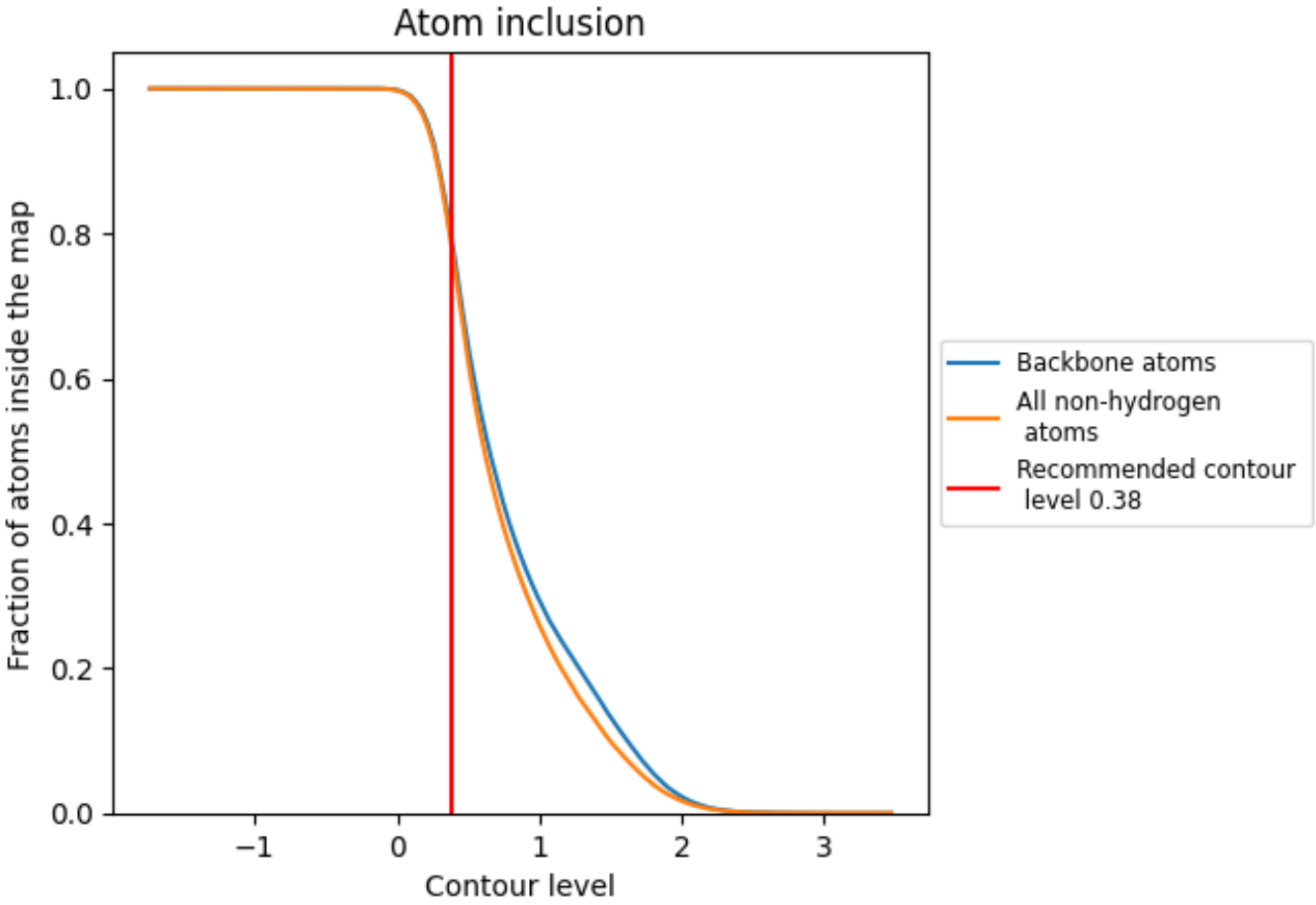
9.3 Atom inclusion mapped to coordinate model (i)



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 (0.38).

9.4 Atom inclusion (i)



At the recommended contour level, 80% of all backbone atoms, 78% 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 (0.38) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7843	0.2680
4	0.9928	0.4660
A	0.9084	0.4360
B	0.9412	0.3070
C	0.9125	0.4070
E	0.9423	0.4320
F	0.9505	0.3510
G	0.7397	0.1590
H	0.7977	0.1190
I	0.9146	0.2190
J	0.9069	0.3200
K	0.9469	0.1950
L	0.8986	0.3480
M	0.8759	0.4000
N	0.9384	0.4560
O	0.8938	0.3870
P	0.8342	0.4200
Q	0.5798	0.0360
R	0.8754	0.4190
S	0.9275	0.4120
T	0.9658	0.5060
U	0.8798	0.3190
V	0.5731	0.1630
W	0.8580	0.3230
X	0.9316	0.4390
Y	0.2981	0.0460
a	0.7687	0.1890
b	0.7536	0.1470
c	0.8625	0.0820
d	0.7677	0.0280
e	0.6855	0.1080
f	0.7668	0.0650
g	0.5852	0.0950
h	0.7065	0.0390
i	0.5378	0.0180

Continued on next page...

Continued from previous page...

Chain	Atom inclusion	Q-score
j	0.6141	0.0150
k	0.6202	0.0240
l	0.7022	0.0070
m	0.7986	0.0160
n	0.7585	0.0690
o	0.5350	0.0260
p	0.7863	0.0300
q	0.6085	0.0490
r	0.9327	0.1960
s	0.4962	0.1310
t	0.4144	0.0610
u	0.1987	0.0220
v	0.0070	-0.0000
w	0.0135	-0.0170
x	0.0645	0.0020
y	0.8000	0.1020