wwPDB EM Validation Summary Report

Mar 20, 2024 – 04:04 PM JST

PDB ID	:	7DCO
EMDB ID	:	EMD-30637
Title	:	Cryo-EM structure of the activated spliceosome (Bact complex) at an atomic resolution of 2.5 angstrom
Authors	:	Bai, R.; Wan, R.; Yan, C.; Qi, J.; Zhang, P.; Lei, J.; Shi, Y.
Deposited on	:	2020-10-26
Resolution	:	2.50 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 ⓘ 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 ⓘ) were used in the production of this report:

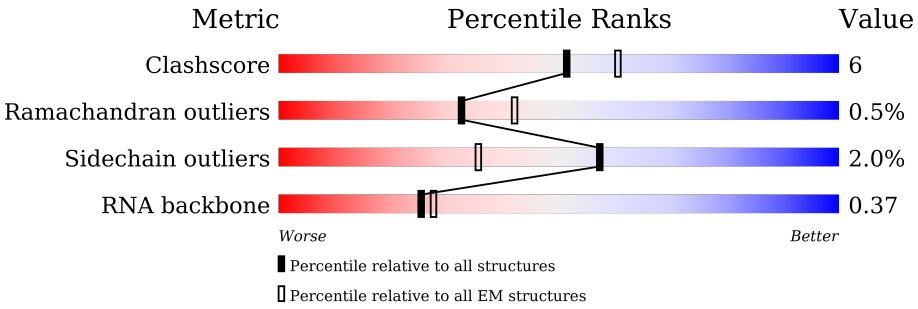
EMDB validation analysis	:	0.0.1.dev70
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.13
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.50 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.



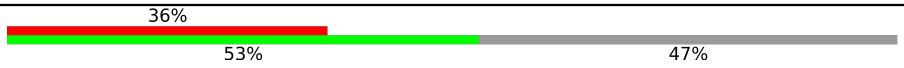
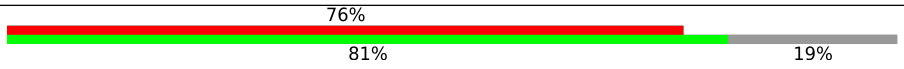
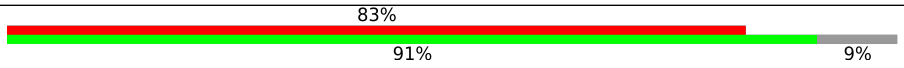
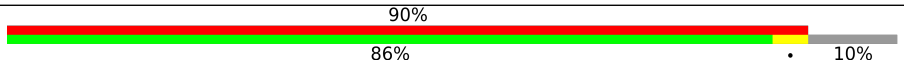
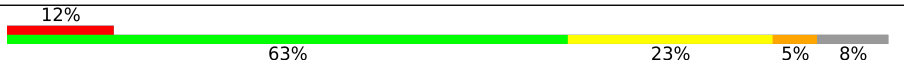
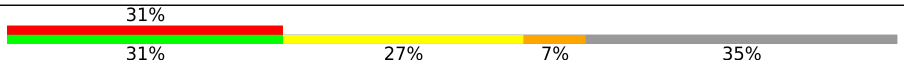
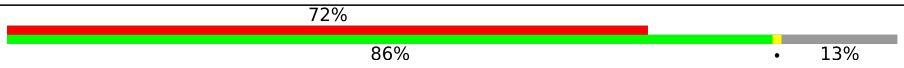
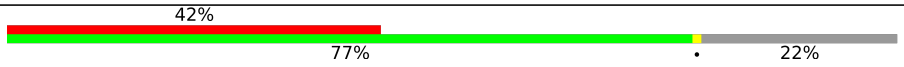
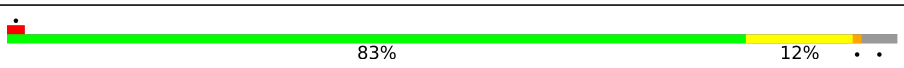
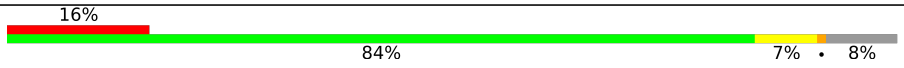
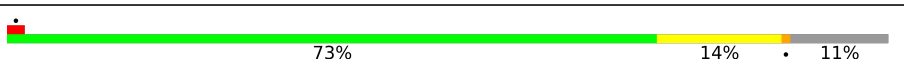
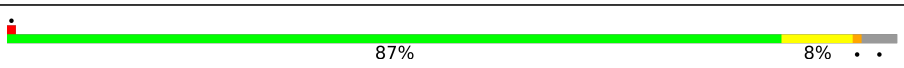
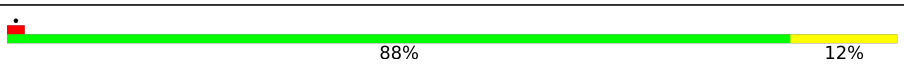
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

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	A	2413	80% 11% 9%
2	B	214	43% 40% 32% 12% 16%
3	C	1008	14% 77% 13% 9%
4	D	2163	44% 74% 10% 15%
5	d	101	40% 78% 22%
6	a	196	18% 37% 63%
6	h	196	40% 40% 60%

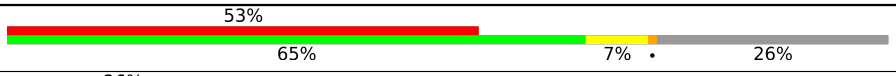
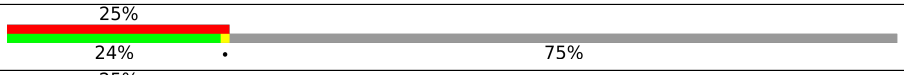
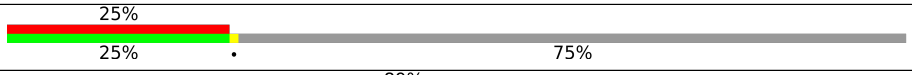
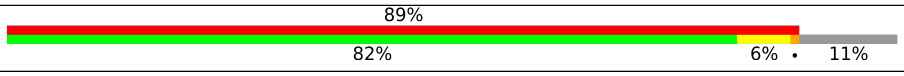
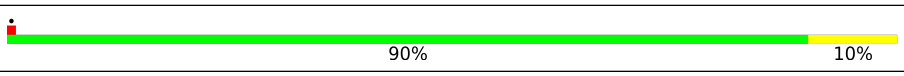
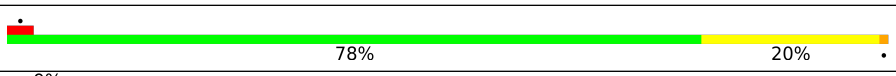
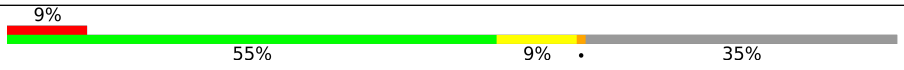
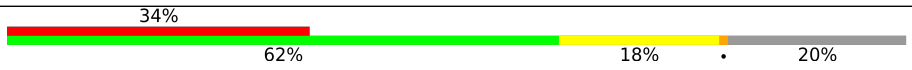
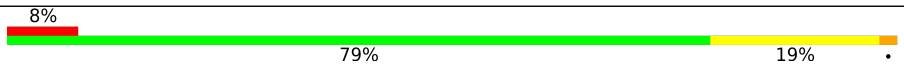
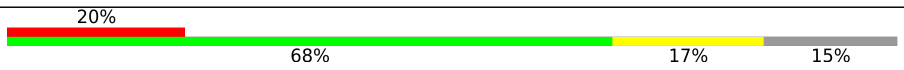
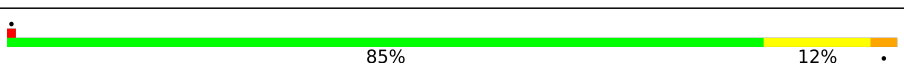
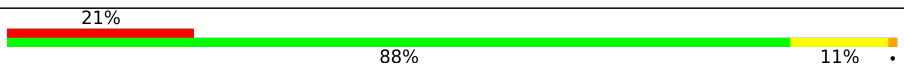
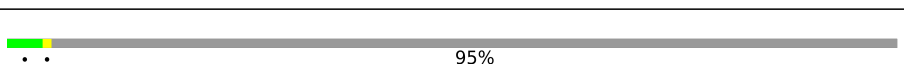
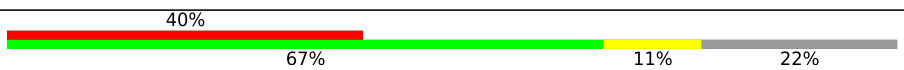
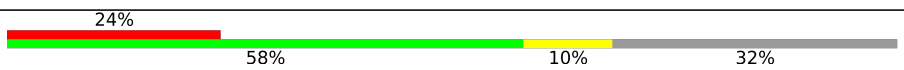
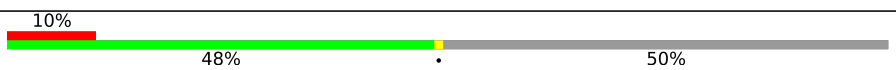
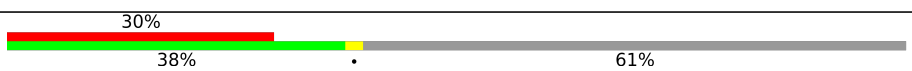
Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	b	146	
7	m	146	
8	c	110	
8	n	110	
9	e	94	
9	i	94	
10	f	86	
10	j	86	
11	g	77	
11	k	77	
12	F	112	
13	G	162	
14	H	1175	
15	o	238	
16	p	111	
17	l	81	
18	u	530	
19	w	280	
20	v	266	
21	1	971	
22	2	238	
23	3	1361	
24	4	213	
25	5	107	
26	6	84	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
27	L	590	
28	q	503	
28	r	503	
28	s	503	
28	t	503	
29	K	175	
30	N	157	
31	T	337	
32	P	379	
33	Q	364	
34	R	261	
35	S	175	
36	Y	204	
37	X	128	
38	Z	121	
39	W	455	
40	U	135	
41	V	577	
42	M	259	
43	z	301	
44	y	185	
45	J	687	
46	I	859	
47	x	876	

2 Entry composition (i)

There are 53 unique types of molecules in this entry. The entry contains 129875 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 PRP8 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	2205	18135	11656	3091	3324	64	0	0

- Molecule 2 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	B	179	3795	1699	660	1258	178	0	0

- Molecule 3 is a protein called SNU114 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	918	7328	4725	1218	1355	30	0	0

- Molecule 4 is a protein called BRR2 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	1828	14666	9388	2437	2784	57	0	0

- Molecule 5 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
5	d	79	316	158	79	79	0	0

- Molecule 6 is a protein called BJ4_G0014900.mRNA.1.CDS.1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
6	a	73	292	146	73	73	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	AltConf	Trace
6	h	78	Total C N O S 610 389 110 108 3	0	0

- Molecule 7 is a protein called SMD1 isoform 1.

Mol	Chain	Residues	Atoms	AltConf	Trace
7	b	77	Total C N O 308 154 77 77	0	0
7	m	82	Total C N O S 644 409 110 123 2	0	0

- Molecule 8 is a protein called BJ4_G0037700.mRNA.1.CDS.1.

Mol	Chain	Residues	Atoms	AltConf	Trace
8	c	90	Total C N O 360 180 90 90	0	0
8	n	65	Total C N O S 528 340 102 84 2	0	0

- Molecule 9 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms	AltConf	Trace
9	e	72	Total C N O 288 144 72 72	0	0
9	i	75	Total C N O S 575 379 92 101 3	0	0

- Molecule 10 is a protein called Sm protein F.

Mol	Chain	Residues	Atoms	AltConf	Trace
10	f	70	Total C N O 280 140 70 70	0	0
10	j	70	Total C N O S 554 355 98 100 1	0	0

- Molecule 11 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms	AltConf	Trace
11	g	70	Total C N O 280 140 70 70	0	0
11	k	69	Total C N O S 529 337 93 97 2	0	0

- Molecule 12 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	F	103	Total	C	N	O	P	0	0
			2192	982	391	716	103		

- Molecule 13 is a RNA chain called pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	G	105	Total	C	N	O	P	0	0
			2099	942	330	723	104		

- Molecule 14 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	H	169	Total	C	N	O	P	0	0
			3566	1594	595	1208	169		

- Molecule 15 is a protein called HLJ1_G0053790.mRNA.1.CDS.1.

Mol	Chain	Residues	Atoms				AltConf	Trace
15	o	135	Total	C	N	O	0	0
			841	538	142	161		

- Molecule 16 is a protein called BJ4_G0027490.mRNA.1.CDS.1.

Mol	Chain	Residues	Atoms				AltConf	Trace
16	p	75	Total	C	N	O	0	0
			476	310	83	83		

- Molecule 17 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	l	81	Total	C	N	O	S	0	0
			611	390	106	113	2		

- Molecule 18 is a protein called PRP9 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	u	461	Total	C	N	O	S	0	0
			3899	2477	675	732	15		

- Molecule 19 is a protein called Pre-mRNA-splicing factor PRP21.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	w	127	Total	C	N	O	S	0	0
			1084	689	193	196	6		

- Molecule 20 is a protein called PRP11 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	v	207	Total	C	N	O	S	0	0
			1621	1014	281	319	7		

- Molecule 21 is a protein called HSH155 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	1	930	Total	C	N	O	S	0	0
			7376	4723	1262	1348	43		

- Molecule 22 is a protein called HLJ1_G0043010.mRNA.1.CDS.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	2	220	Total	C	N	O	S	0	0
			1793	1158	311	314	10		

- Molecule 23 is a protein called RSE1 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	3	1207	Total	C	N	O	S	0	0
			9599	6134	1613	1801	51		

- Molecule 24 is a protein called HSH49 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	4	174	Total	C	N	O	S	0	0
			1433	932	240	259	2		

- Molecule 25 is a protein called BJ4_G0056610.mRNA.1.CDS.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	5	103	Total	C	N	O	S	0	0
			814	503	154	143	14		

- Molecule 26 is a protein called RDS3 complex subunit 10.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	6	84	Total	C	N	O	S	0	0
			693	429	130	132	2		

- Molecule 27 is a protein called Pre-mRNA-splicing factor CEF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	L	435	Total	C	N	O	S	0	0
			2901	1799	538	557	7		

- Molecule 28 is a protein called Pre-mRNA-processing factor 19.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	s	124	Total	C	N	O	S	0	0
			819	518	132	167	2		
28	t	128	Total	C	N	O	S	0	0
			843	533	136	172	2		
28	q	129	Total	C	N	O	S	0	0
			850	537	137	174	2		
28	r	125	Total	C	N	O	S	0	0
			823	521	133	167	2		

- Molecule 29 is a protein called SNT309 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	K	155	Total	C	N	O	S	0	0
			920	581	159	179	1		

- Molecule 30 is a protein called BUD31 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	N	157	Total	C	N	O	S	0	0
			1291	808	240	232	11		

- Molecule 31 is a protein called HLJ1_G0054350.mRNA.1.CDS.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	T	337	Total	C	N	O	S	0	0
			2646	1669	466	501	10		

- Molecule 32 is a protein called Pre-mRNA-processing protein 45.

Mol	Chain	Residues	Atoms					AltConf	Trace
32	P	246	Total	C	N	O	S	0	0
			1978	1233	359	380	6		

- Molecule 33 is a protein called Pre-mRNA-splicing factor SLT11.

Mol	Chain	Residues	Atoms					AltConf	Trace
33	Q	292	Total	C	N	O	S	0	0
			2301	1461	399	426	15		

- Molecule 34 is a protein called Pre-mRNA-splicing factor CWC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
34	R	261	Total	C	N	O	S	0	0
			2089	1320	369	388	12		

- Molecule 35 is a protein called Pre-mRNA-splicing factor CWC15.

Mol	Chain	Residues	Atoms					AltConf	Trace
35	S	71	Total	C	N	O	S	0	0
			578	361	117	99	1		

- Molecule 36 is a protein called Pre-mRNA leakage protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
36	Y	174	Total	C	N	O	S	0	0
			1386	868	233	275	10		

- Molecule 37 is a protein called SX2_G0027210.mRNA.1.CDS.1.

Mol	Chain	Residues	Atoms				AltConf	Trace
37	X	128	Total	C	N	O	0	0
			1051	662	181	208		

- Molecule 38 is a protein called Pre-mRNA-splicing factor CWC26.

Mol	Chain	Residues	Atoms					AltConf	Trace
38	Z	121	Total	C	N	O	S	0	0
			920	575	161	182	2		

- Molecule 39 is a protein called CDC40 isoform 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
39	W	23	Total	C	N	O	0	0
			195	122	41	32		

- Molecule 40 is a protein called Pre-mRNA-splicing factor CWC21.

Mol	Chain	Residues	Atoms				AltConf	Trace
40	U	27	Total	C	N	O	0	0
			190	112	38	40		

- Molecule 41 is a protein called CWC22 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
41	V	450	Total	C	N	O	S	0	0
			3660	2346	605	691	18		

- Molecule 42 is a protein called Pre-mRNA-splicing factor CWC24.

Mol	Chain	Residues	Atoms					AltConf	Trace
42	M	176	Total	C	N	O	S	0	0
			1360	852	235	260	13		

- Molecule 43 is a protein called CWC27 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
43	z	150	Total	C	N	O	S	0	0
			1224	789	206	223	6		

- Molecule 44 is a protein called Pre-mRNA-splicing factor SPP2.

Mol	Chain	Residues	Atoms					AltConf	Trace
44	y	73	Total	C	N	O	S	0	0
			572	371	93	107	1		

- Molecule 45 is a protein called CLF1 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
45	J	554	Total	C	N	O	S	0	0
			3595	2231	680	676	8		

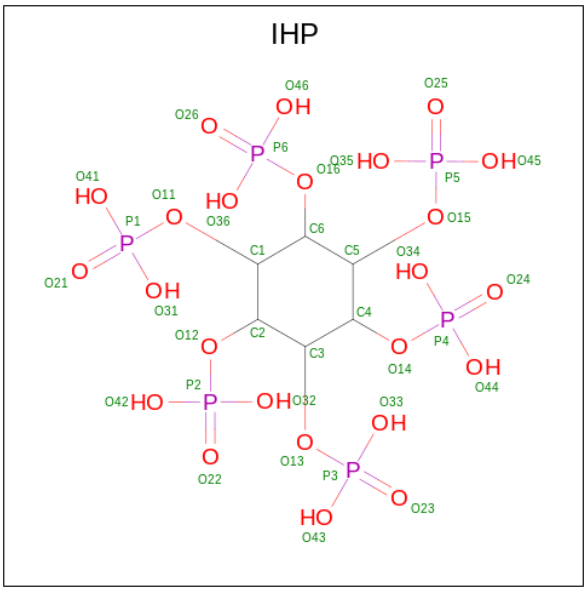
- Molecule 46 is a protein called SYF1 isoform 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
46	I	580	Total	C	N	O	S	0	0
			3101	1897	590	613	1		

- Molecule 47 is a protein called Pre-mRNA-splicing factor ATP-dependent RNA helicase-like protein PRP2.

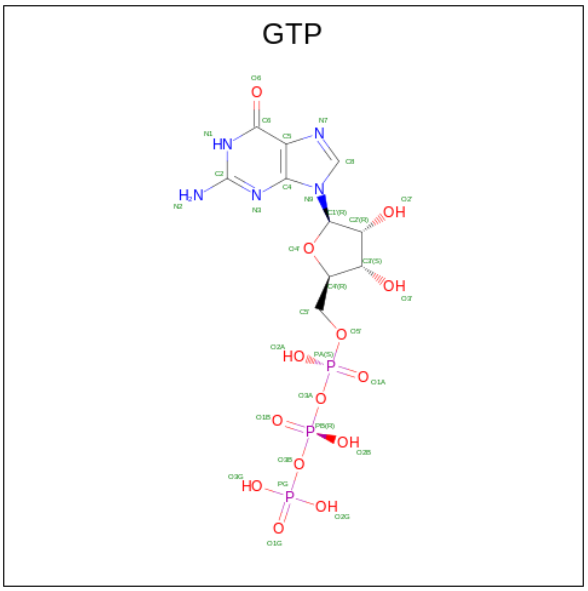
Mol	Chain	Residues	Atoms					AltConf	Trace
47	x	653	Total	C	N	O	S	0	0
			5193	3313	891	956	33		

- Molecule 48 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
48	A	1	Total	C	O	P	0
			36	6	24	6	

- Molecule 49 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
49	C	1	Total	C	N	O	P	0
			32	10	5	14	3	

- Molecule 50 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
50	C	1	Total	Mg	0
			1	1	
50	F	4	Total	Mg	0
			4	4	
50	3	1	Total	Mg	0
			1	1	

- Molecule 51 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
51	C	1	Total	Ca	0
			1	1	

- Molecule 52 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
52	u	2	Total	Zn	0
			2	2	

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	AltConf
52	v	1	Total Zn 1 1	0
52	5	3	Total Zn 3 3	0
52	N	3	Total Zn 3 3	0
52	Q	2	Total Zn 2 2	0
52	R	1	Total Zn 1 1	0
52	M	3	Total Zn 3 3	0

- Molecule 53 is water.

Mol	Chain	Residues	Atoms	AltConf
53	A	582	Total O 582 582	0
53	B	127	Total O 127 127	0
53	C	8	Total O 8 8	0
53	D	2	Total O 2 2	0
53	F	109	Total O 109 109	0
53	G	56	Total O 56 56	0
53	H	70	Total O 70 70	0
53	v	29	Total O 29 29	0
53	1	185	Total O 185 185	0
53	2	51	Total O 51 51	0
53	3	218	Total O 218 218	0
53	5	53	Total O 53 53	0
53	6	42	Total O 42 42	0

Continued on next page...

Continued from previous page...

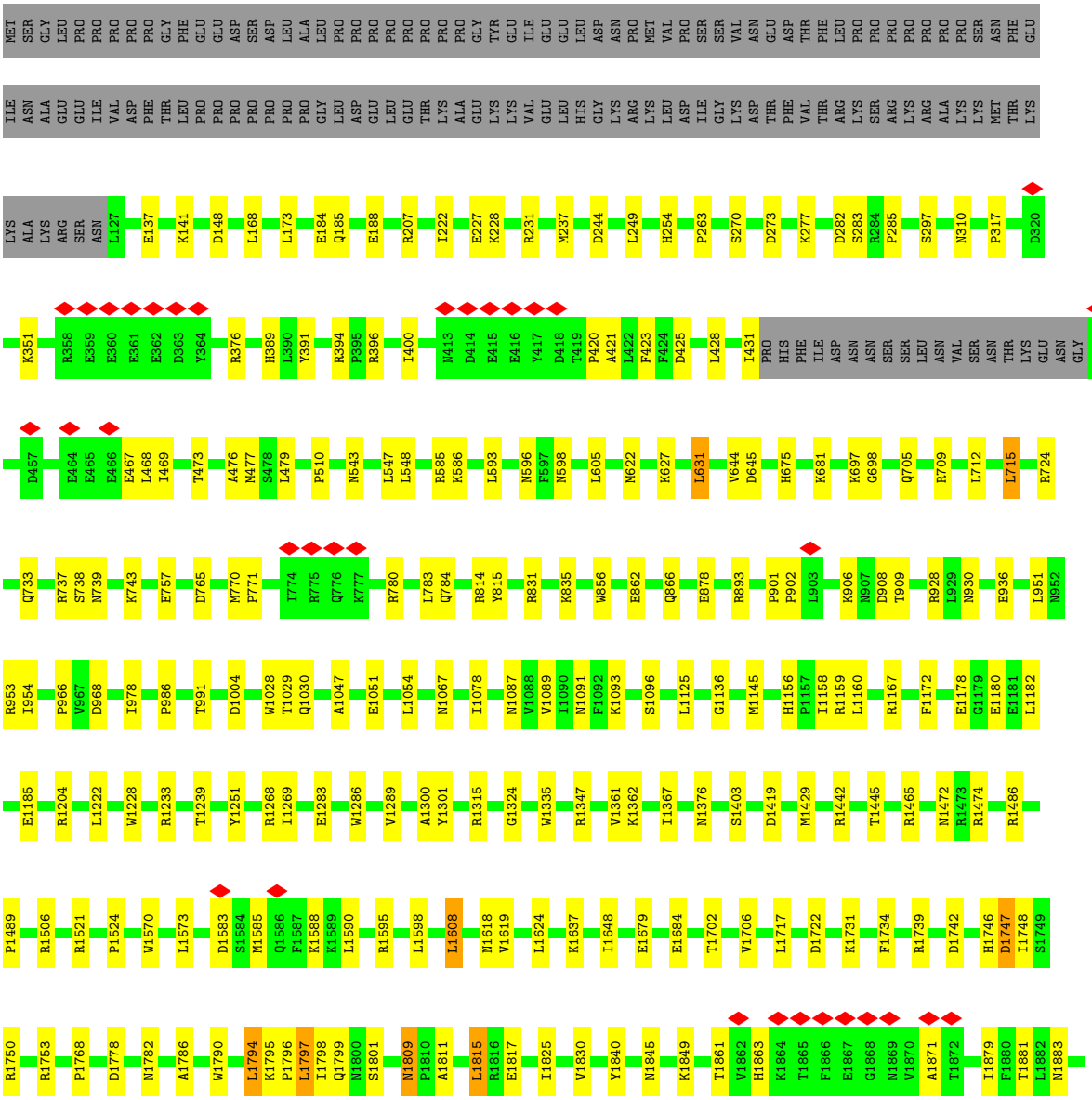
Mol	Chain	Residues	Atoms	AltConf
53	L	18	Total O 18 18	0
53	N	12	Total O 12 12	0
53	T	39	Total O 39 39	0
53	P	26	Total O 26 26	0
53	R	8	Total O 8 8	0
53	S	9	Total O 9 9	0
53	X	13	Total O 13 13	0
53	Z	9	Total O 9 9	0
53	U	6	Total O 6 6	0
53	V	21	Total O 21 21	0
53	M	21	Total O 21 21	0

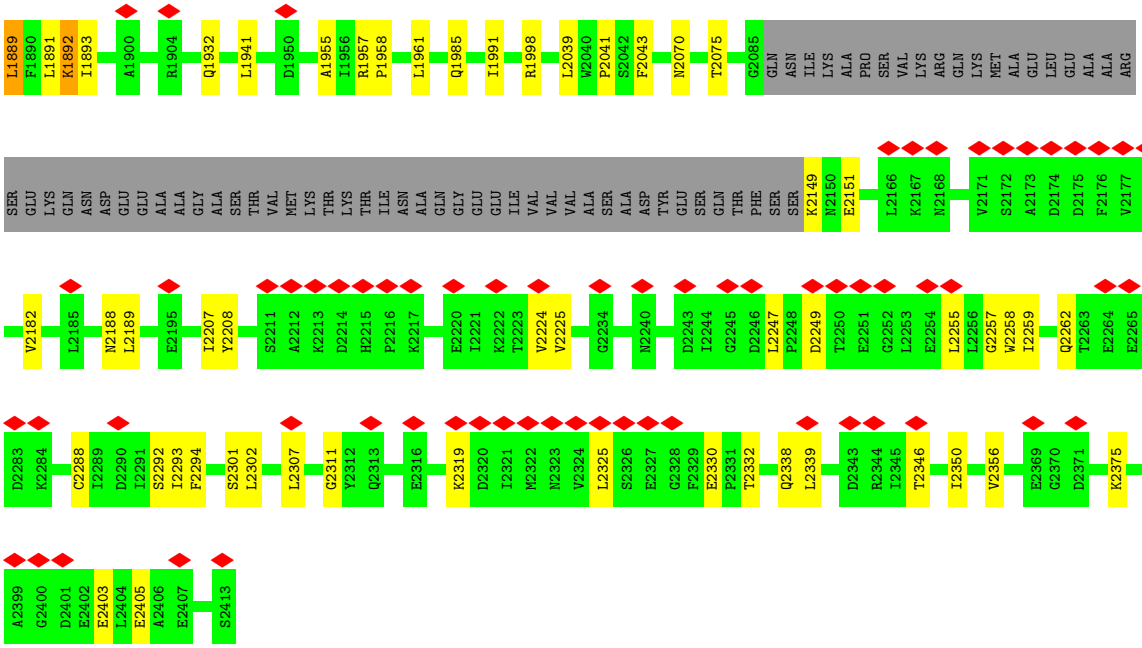
3 Residue-property plots (i)

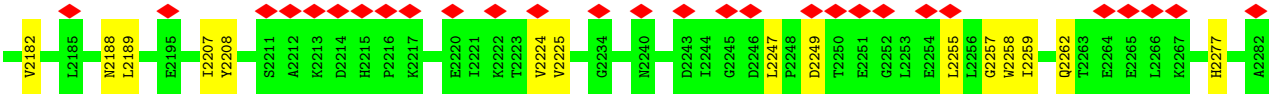
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion 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: PRP8 isoform 1

Chain A:

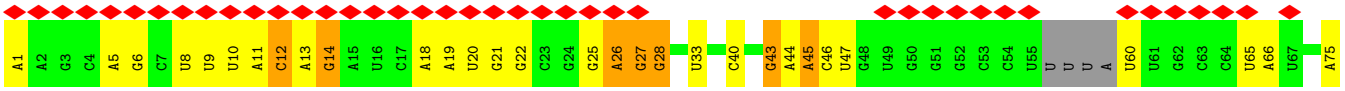


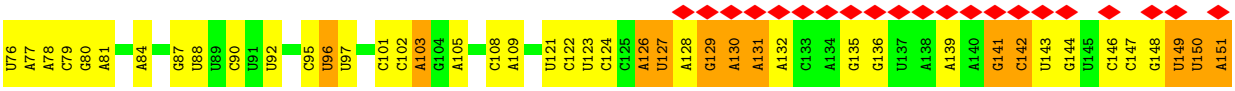


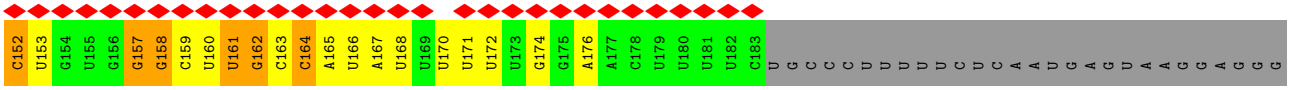


• Molecule 2: U5 snRNA

Chain B:

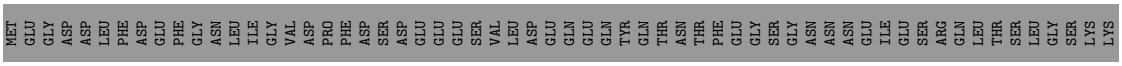


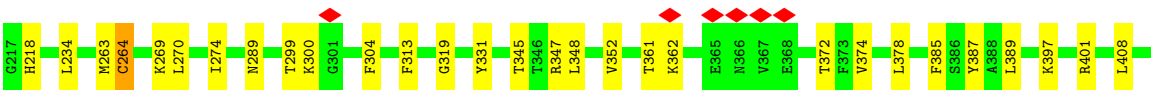


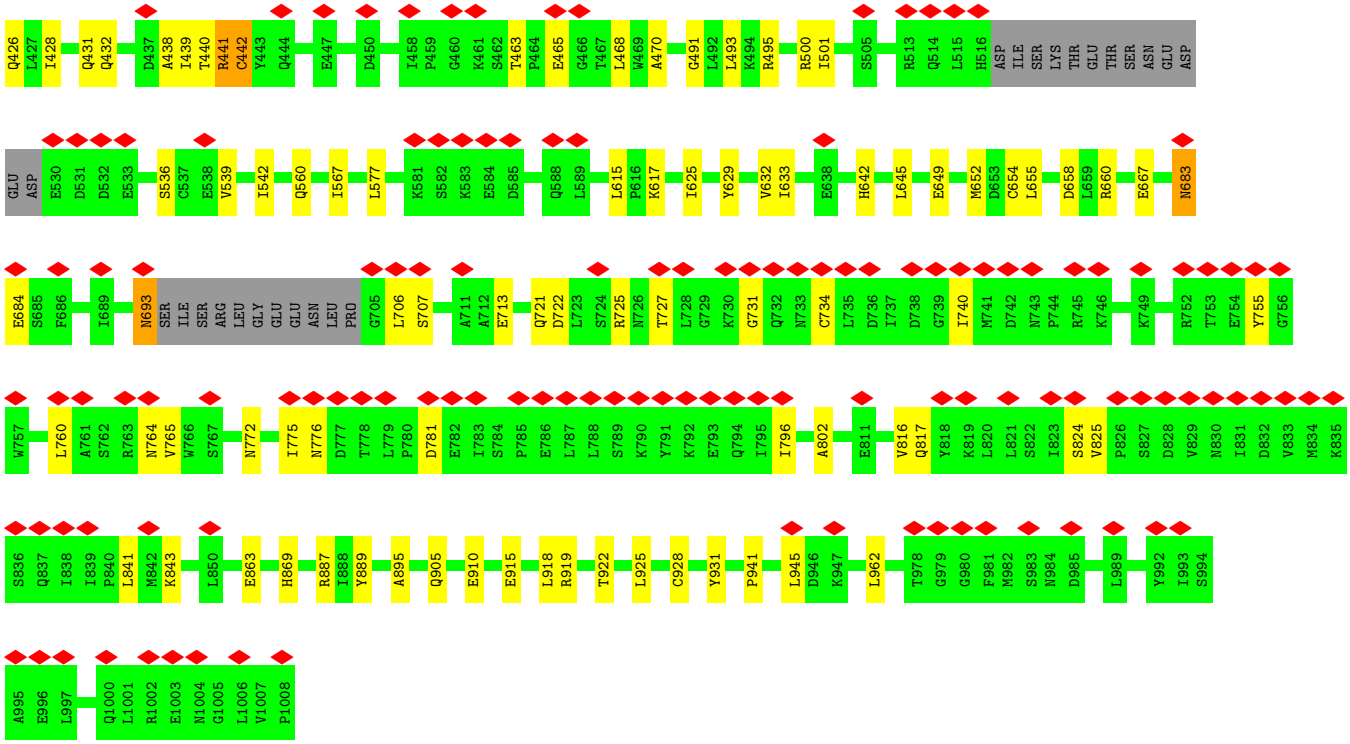


• Molecule 3: SNU114 isoform 1

Chain C:

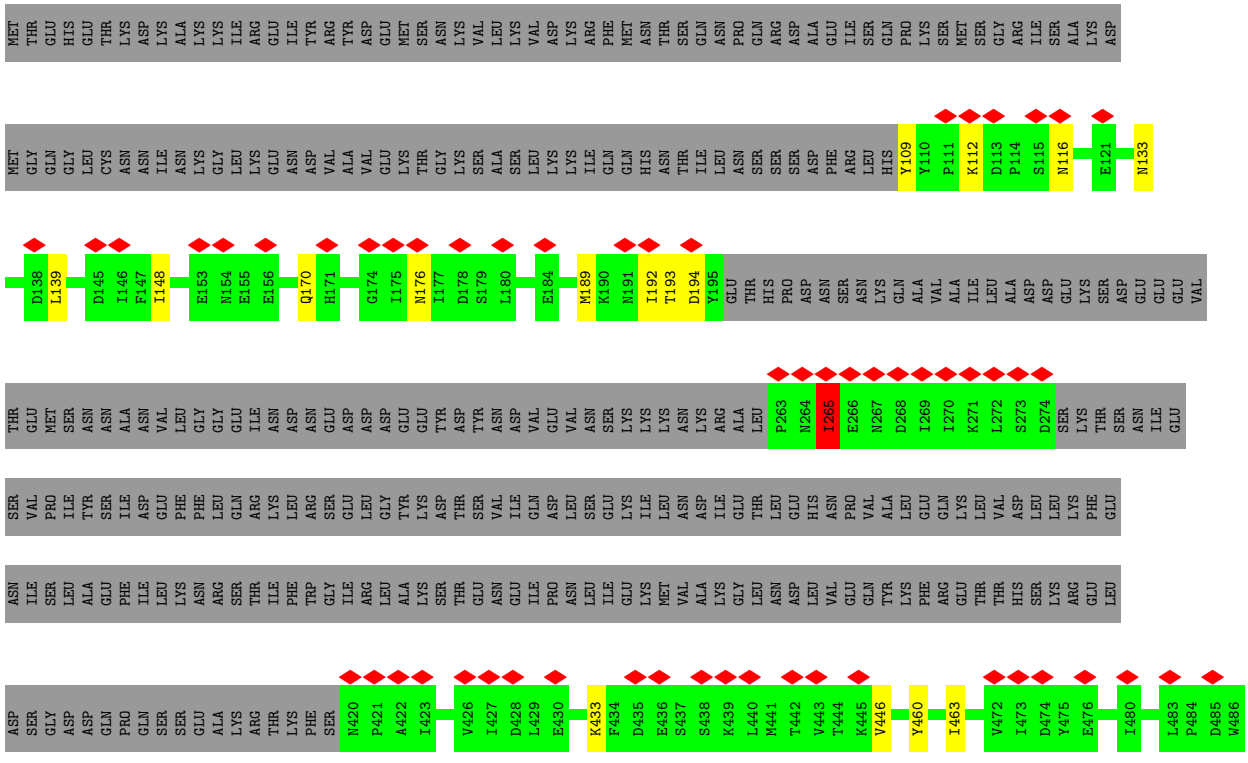


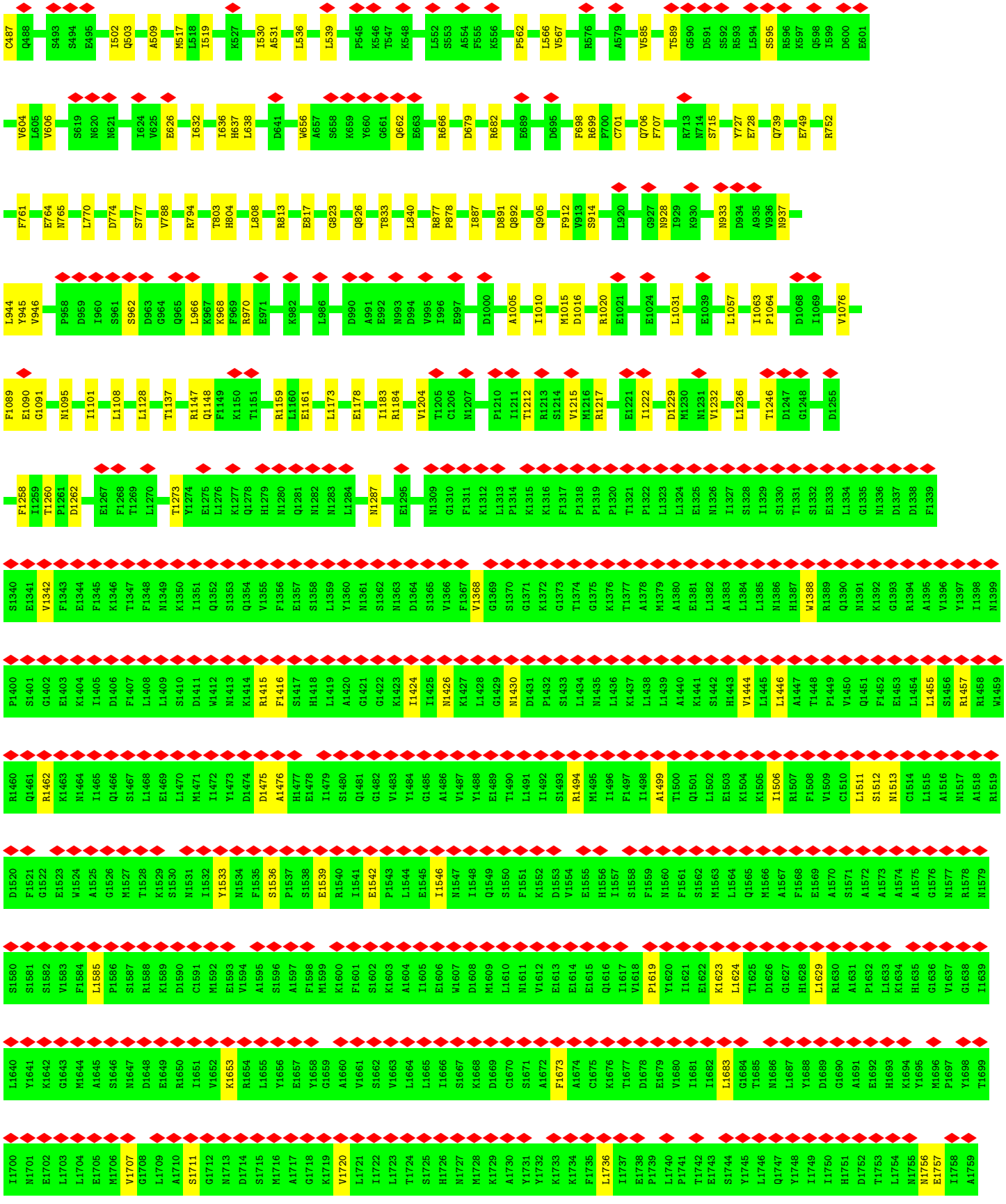


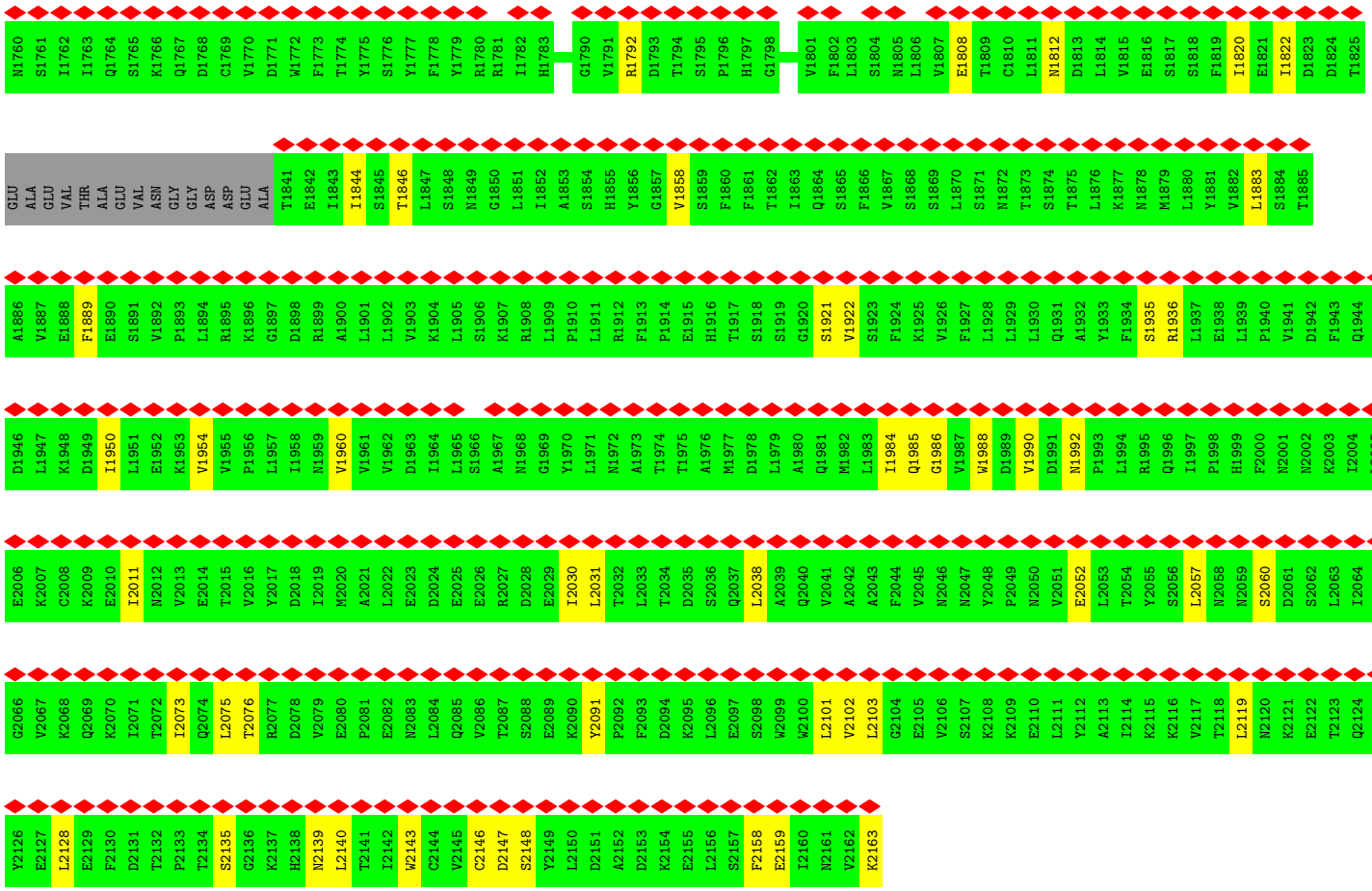


• Molecule 4: BRR2 isoform 1

Chain D:

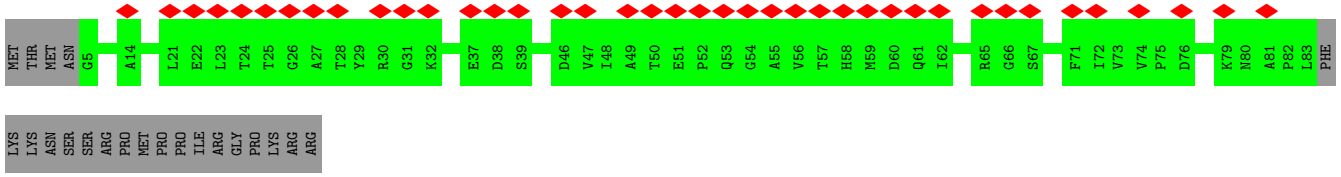



C487 Q488 S493 S494 E495 I502 Q503 A509 M517 L518 I519 K527 I530 A531 L536 L539 P545 K546 T547 K548 L552 S553 A554 F555 K556 P562 L566 V567 R576 A579 V585 T589 G590 D591 S592 R593 L594 S595 R596 K597 Q598 I599 D600 E601
V604 L605 V606 S619 N620 N621 I624 V625 E626 I632 I636 H637 L638 D641 W656 A657 S658 K659 Y660 G661 Q662 E663 R666 D679 R682 E689 D695 F698 R699 P700 C701 Q706 F707 R713 N714 S715 Y727 E728 Q739 E749 R752
F761 E764 N765 L770 D774 S777 V788 R794 T803 H804 L808 R813 E817 G823 Q826 T833 L840 R877 P878 I887 D891 Q892 Q905 F912 V913 S914 L920 G927 N928 I929 K930 N933 D934 A935 V936 N937
L944 Y945 V946 P958 D959 I960 S961 S962 D963 G964 Q965 L966 K967 K968 F969 R970 E971 K982 L986 D990 A991 E992 N993 D994 V995 I996 E997 D1000 A1005 I1010 M1015 D1016 R1020 E1021 E1024 L1031 E1039 L1057 I1063 P1064 D1068 I1069 V1076
F1089 E1090 G1091 N1095 I1101 L1108 L1128 T1137 R1147 Q1148 F1149 K1150 T1151 R1159 L1160 E1161 L1173 E1178 I1183 R1184 V1204 T1205 C1206 N1207 P1210 I1211 T1212 R1213 S1214 V1215 M1216 R1217 E1221 I1222 D1229 M1230 N1231 V1232 L1236 T1246 D1247 G1248 D1255
F1258 I1259 T1260 P1261 D1262 E1267 F1268 T1269 L1270 T1273 Y1274 E1275 L1276 K1277 Q1278 H1279 N1280 Q1281 N1282 N1283 L1284 N1287 E1295 N1309 G1310 F1311 K1312 L1313 P1314 K1315 K1316 F1317 P1318 P1319 P1320 T1321 P1322 L1323 L1324 E1325 N1326 I1327 S1328 I1329 S1330 T1331 S1332 E1333 L1334 G1335 N1336 D1337 D1338 F1339
S1340 E1341 V1342 F1343 E1344 F1345 K1346 T1347 F1348 N1349 K1350 I1351 Q1352 S1353 Q1354 V1355 F1356 E1357 S1358 L1359 Y1360 N1361 S1362 N1363 D1364 S1365 V1366 F1367 V1368 G1369 S1370 G1371 K1372 G1373 T1374 G1375 K1376 T1377 A1378 M1379 A1380 E1381 L1382 A1383 L1384 L1385 N1386 H1387 W1388 R1389 Q1390 N1391 K1392 G1393 R1394 A1395 V1396 Y1397 I1398 N1399
P1400 S1401 G1402 E1403 K1404 I1405 D1406 F1407 L1408 L1409 S1410 D1411 W1412 N1413 K1414 R1415 F1416 S1417 H1418 L1419 A1420 G1421 G1422 K1423 I1424 I1425 N1426 K1427 L1428 G1429 N1430 D1431 P1432 S1433 L1434 N1435 L1436 K1437 L1438 L1439 A1440 K1441 S1442 H1443 V1444 L1445 L1446 A1447 T1448 P1449 V1450 Q1451 F1452 E1453 L1454 L1455 S1456 R1457 R1458 W1459
R1460 Q1461 R1462 K1463 N1464 I1465 Q1466 S1467 L1468 E1469 L1470 M1471 I1472 Y1473 D1474 D1475 A1476 H1477 E1478 I1479 S1480 Q1481 G1482 V1483 Y1484 G1485 A1486 V1487 Y1488 E1489 T1490 L1491 I1492 S1493 R1494 M1495 I1496 F1497 I1498 A1499 T1500 Q1501 L1502 E1503 K1504 K1505 I1506 R1507 F1508 V1509 C1510 L1511 S1512 N1513 C1514 L1515 A1516 N1517 A1518 R1519
D1520 F1521 G1522 E1523 W1524 A1525 G1526 M1527 T1528 K1529 S1530 N1531 I1532 Y1533 N1534 F1535 S1536 P1537 S1538 E1539 R1540 I1541 E1542 P1543 L1544 E1545 I1546 N1547 I1548 Q1549 S1550 F1551 K1552 D1553 V1554 E1555 H1556 I1557 S1558 F1559 N1560 F1561 S1562 M1563 L1564 Q1565 M1566 A1567 F1568 E1569 A1570 S1571 A1572 A1573 A1574 A1575 G1576 N1577 R1578 N1579
S1580 S1581 S1582 V1583 F1584 L1585 P1586 S1587 R1588 K1589 D1590 C1591 M1592 E1593 V1594 A1595 S1596 A1597 F1598 M1599 K1600 F1601 S1602 K1603 A1604 I1605 E1606 W1607 D1608 M1609 L1610 N1611 V1612 E1613 E1614 E1615 Q1616 I1617 V1618 P1619 Y1620 I1621 E1622 K1623 L1624 T1625 D1626 G1627 H1628 L1629 R1630 A1631 P1632 L1633 K1634 H1635 G1636 V1637 G1638 I1639
L1640 Y1641 K1642 G1643 M1644 A1645 S1646 N1647 D1648 E1649 R1650 I1651 V1652 K1653 R1654 L1655 Y1656 E1657 Y1658 G1659 A1660 V1661 S1662 V1663 L1664 L1665 I1666 S1667 K1668 D1669 C1670 S1671 A1672 F1673 A1674 C1675 K1676 T1677 D1678 E1679 V1680 I1681 I1682 L1683 G1684 T1685 N1686 L1687 Y1688 D1689 G1690 A1691 E1692 H1693 K1694 Y1695 M1696 P1697 Y1698 T1699
I1700 N1701 E1702 L1703 L1704 E1705 M1706 V1707 G1708 L1709 A1710 S1711 G1712 N1713 D1714 S1715 M1716 A1717 G1718 K1719 V1720 L1721 I1722 L1723 T1724 S1725 H1726 N1727 M1728 K1729 A1730 Y1731 Y1732 K1733 K1734 F1735 L1736 I1737 E1738 P1739 L1740 P1741 T1742 E1743 S1744 Y1745 L1746 Q1747 Y1748 I1749 I1750 H1751 D1752 T1753 L1754 N1755 N1756 E1757 I1758 A1759



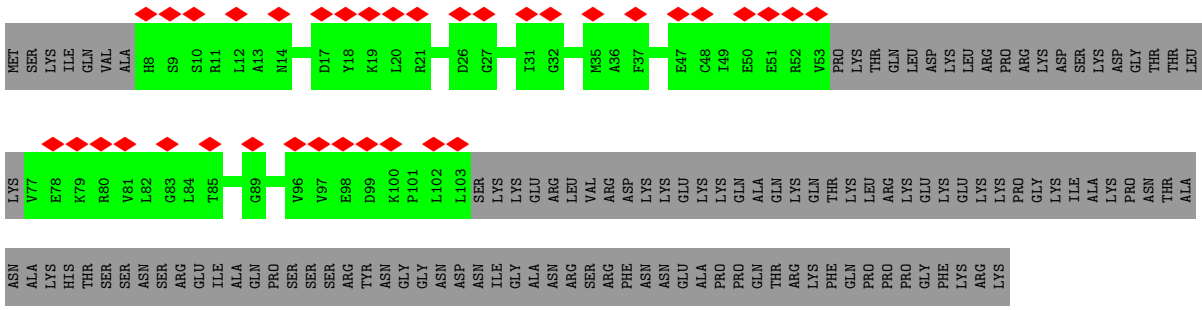
• Molecule 5: Small nuclear ribonucleoprotein Sm D3

Chain d:



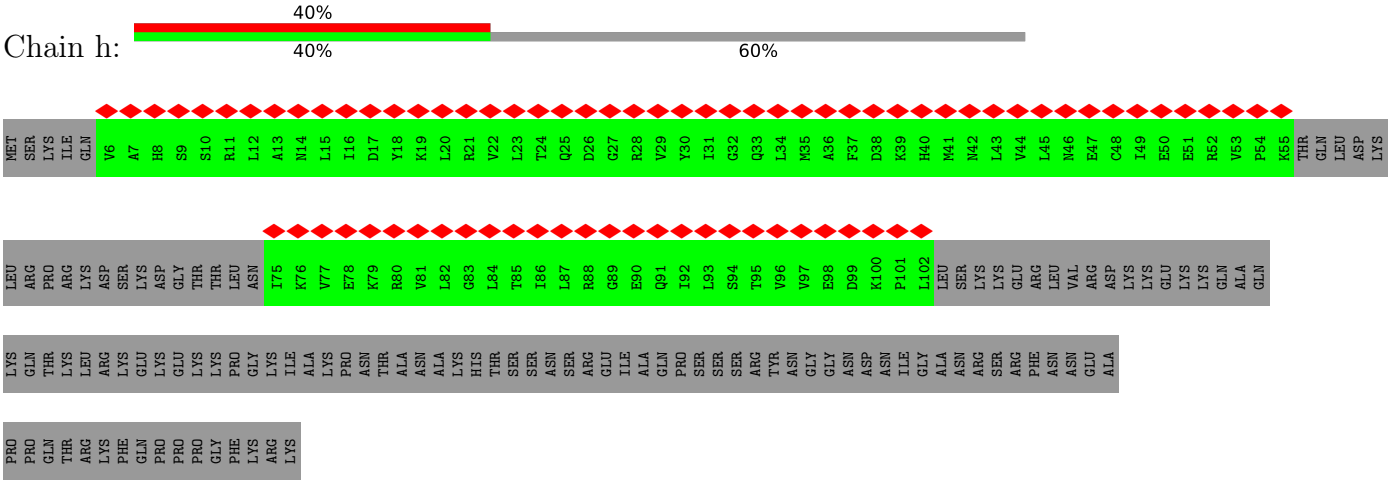
• Molecule 6: BJ4_G0014900.mRNA.1.CDS.1

Chain a:



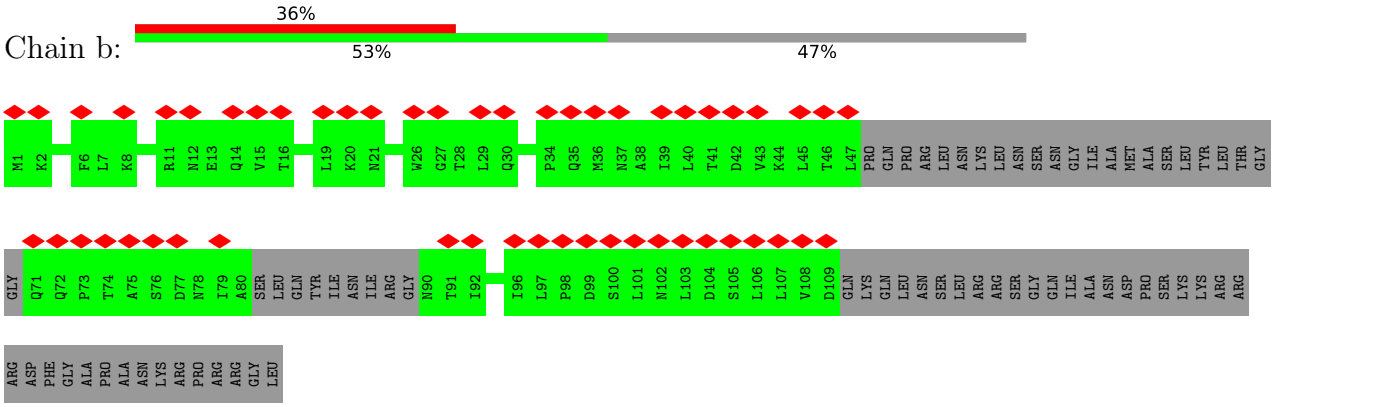
• Molecule 6: BJ4_G0014900.mRNA.1.CDS.1

Chain h:



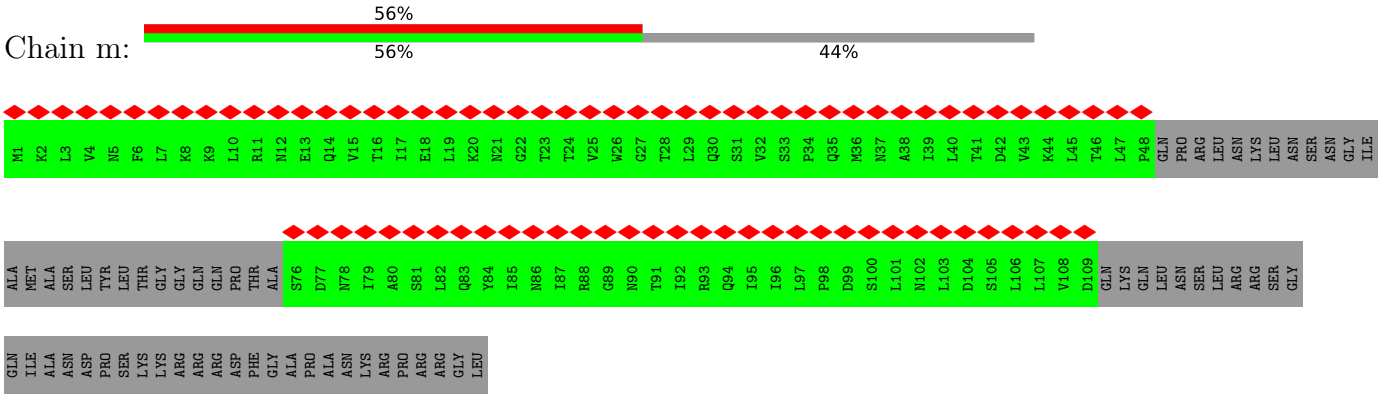
• Molecule 7: SMD1 isoform 1

Chain b:



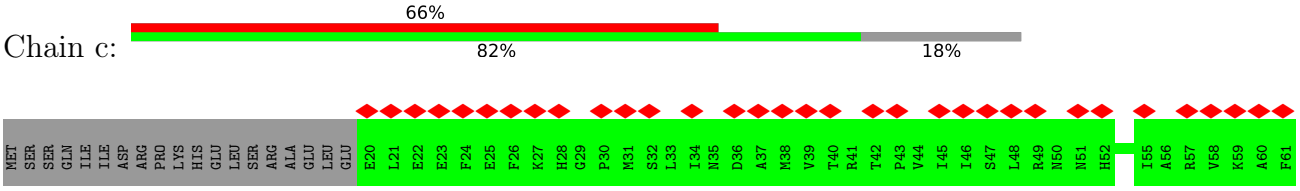
• Molecule 7: SMD1 isoform 1

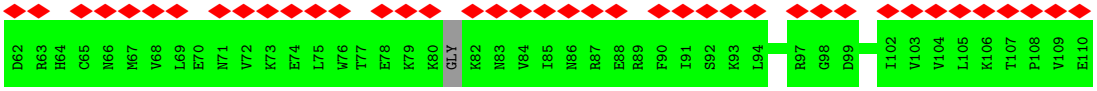
Chain m:



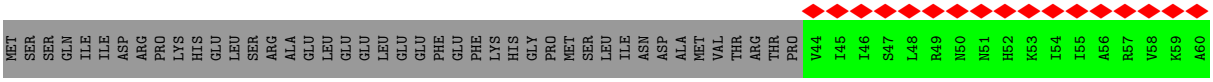
• Molecule 8: BJ4_G0037700.mRNA.1.CDS.1

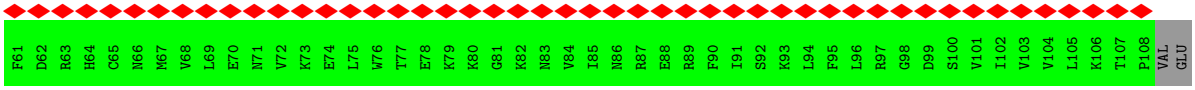
Chain c:



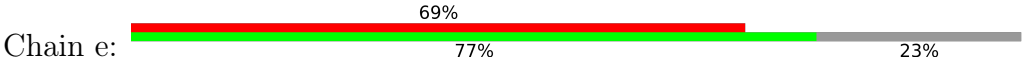


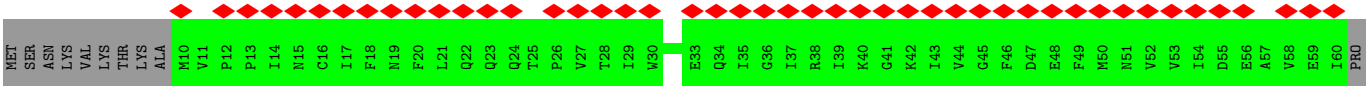
• Molecule 8: BJ4_G0037700.mRNA.1.CDS.1

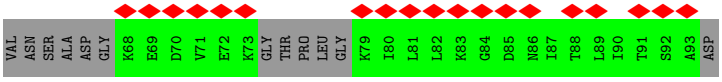




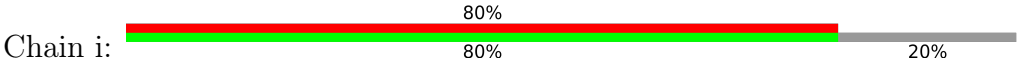
• Molecule 9: Small nuclear ribonucleoprotein E

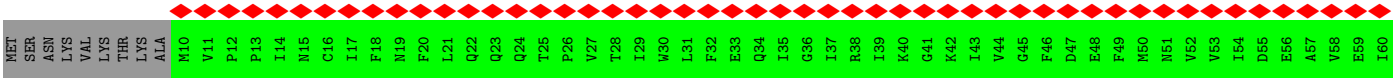


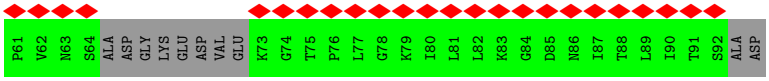




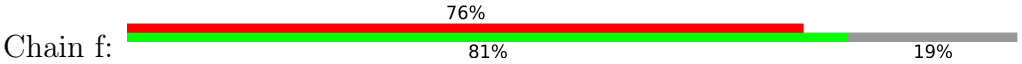
• Molecule 9: Small nuclear ribonucleoprotein E

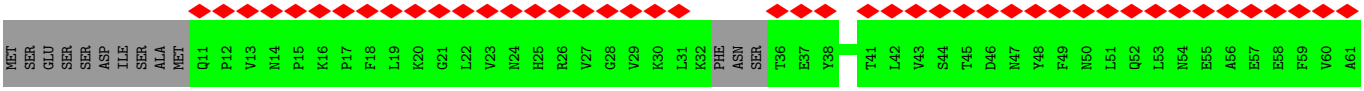


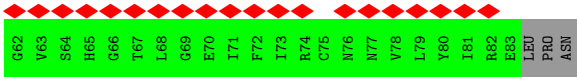




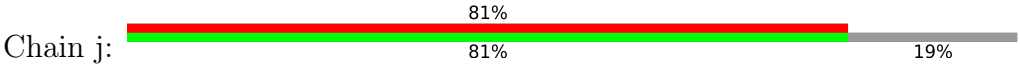
• Molecule 10: Sm protein F

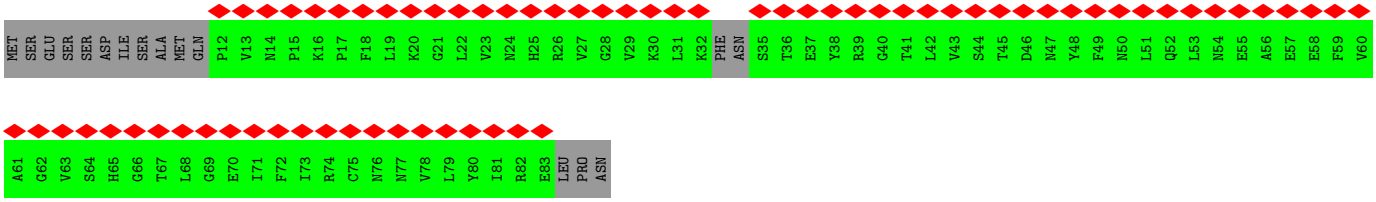




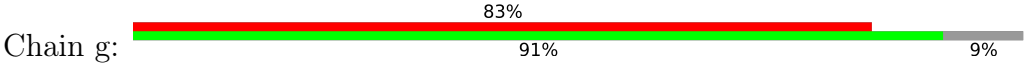


• Molecule 10: Sm protein F

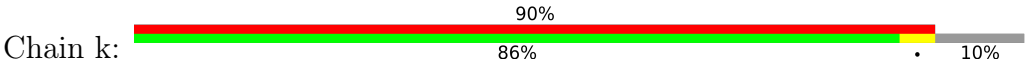




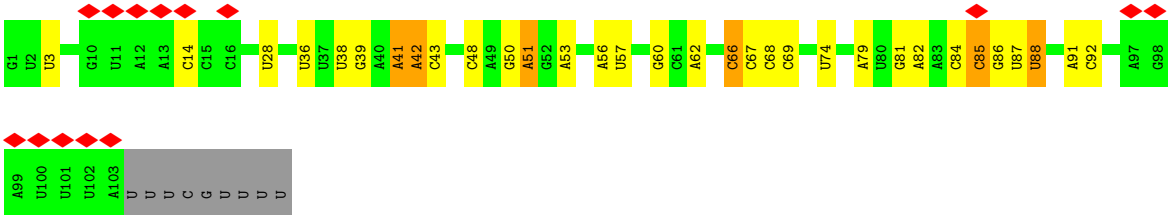
• Molecule 11: Small nuclear ribonucleoprotein G



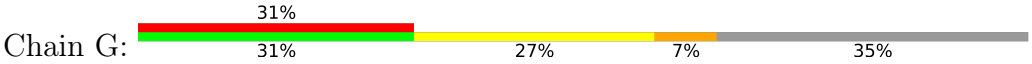
• Molecule 11: Small nuclear ribonucleoprotein G

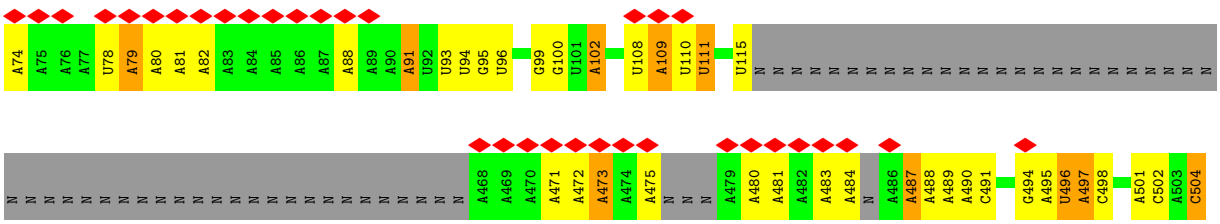


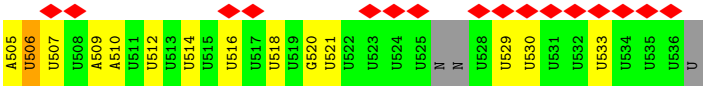
• Molecule 12: U6 snRNA



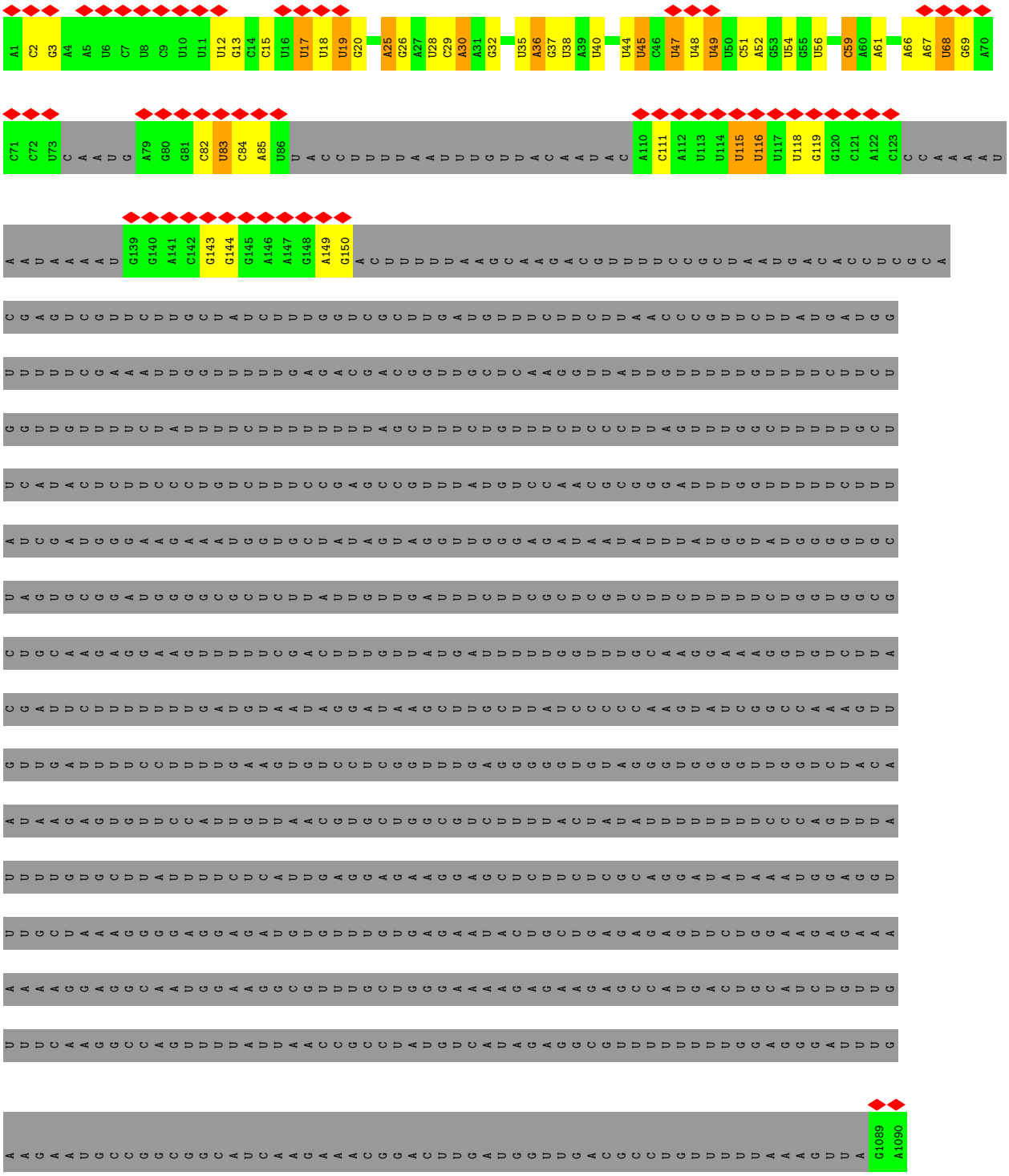
• Molecule 13: pre-mRNA

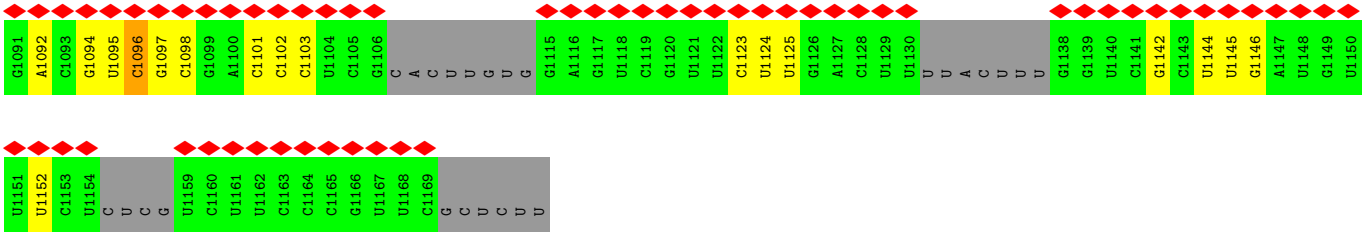






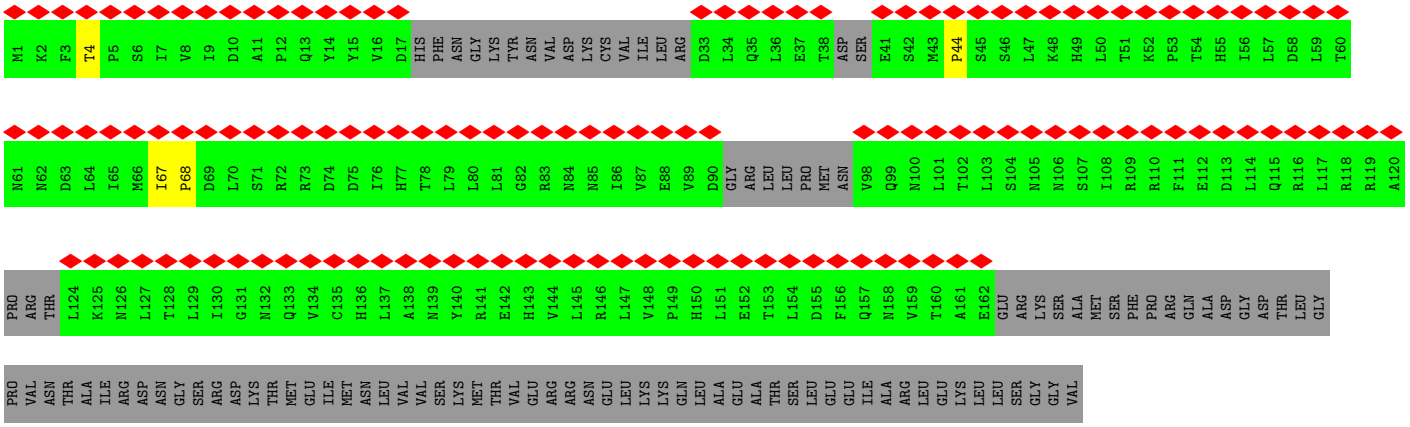
- Molecule 14: U2 snRNA





• Molecule 15: HLJ1_G0053790.mRNA.1.CDS.1

Chain o:

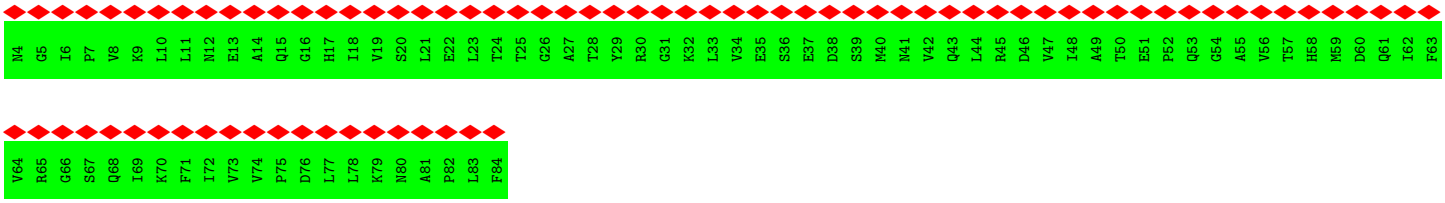


• Molecule 16: BJ4_G0027490.mRNA.1.CDS.1

Chain p:

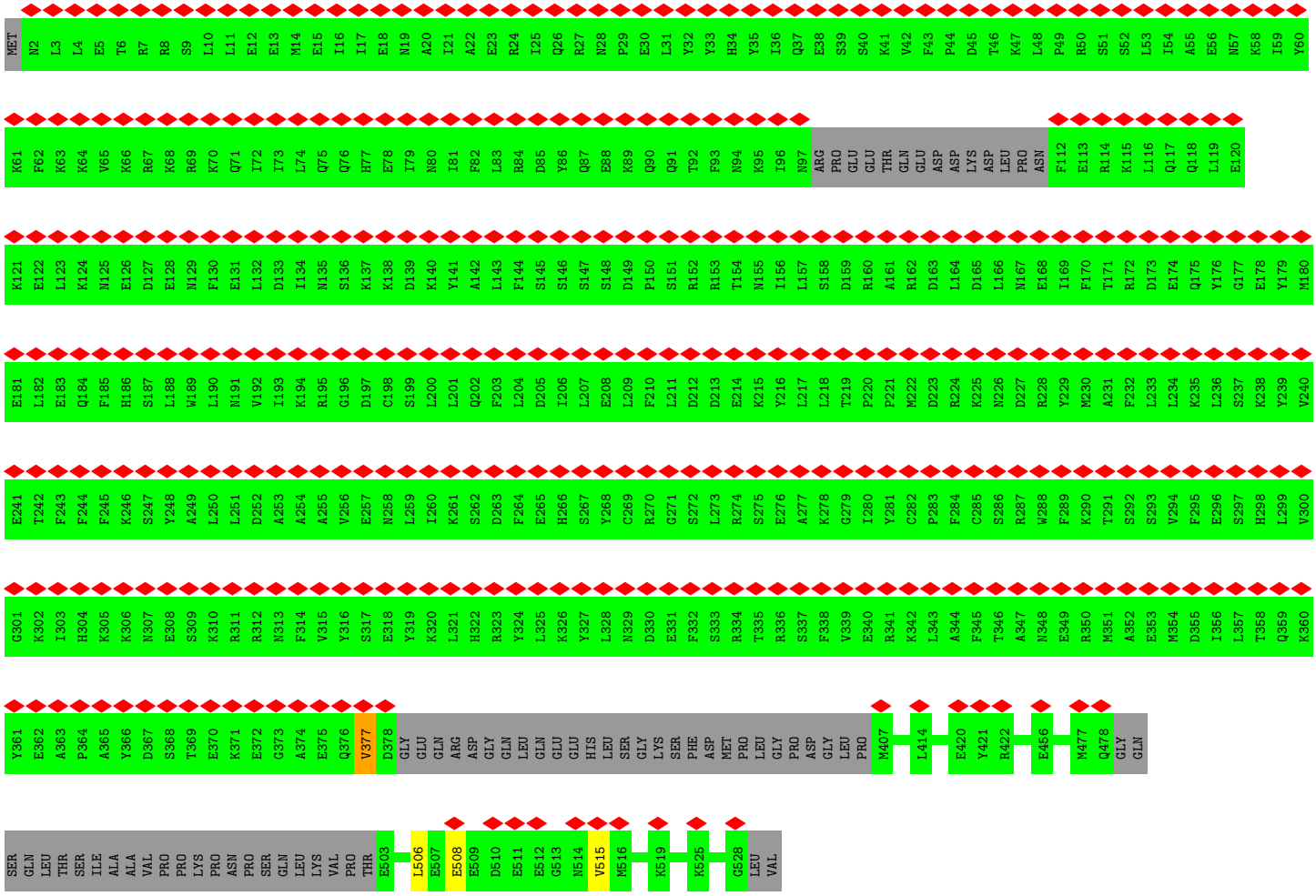
• Molecule 17: Small nuclear ribonucleoprotein Sm D3

Chain l:

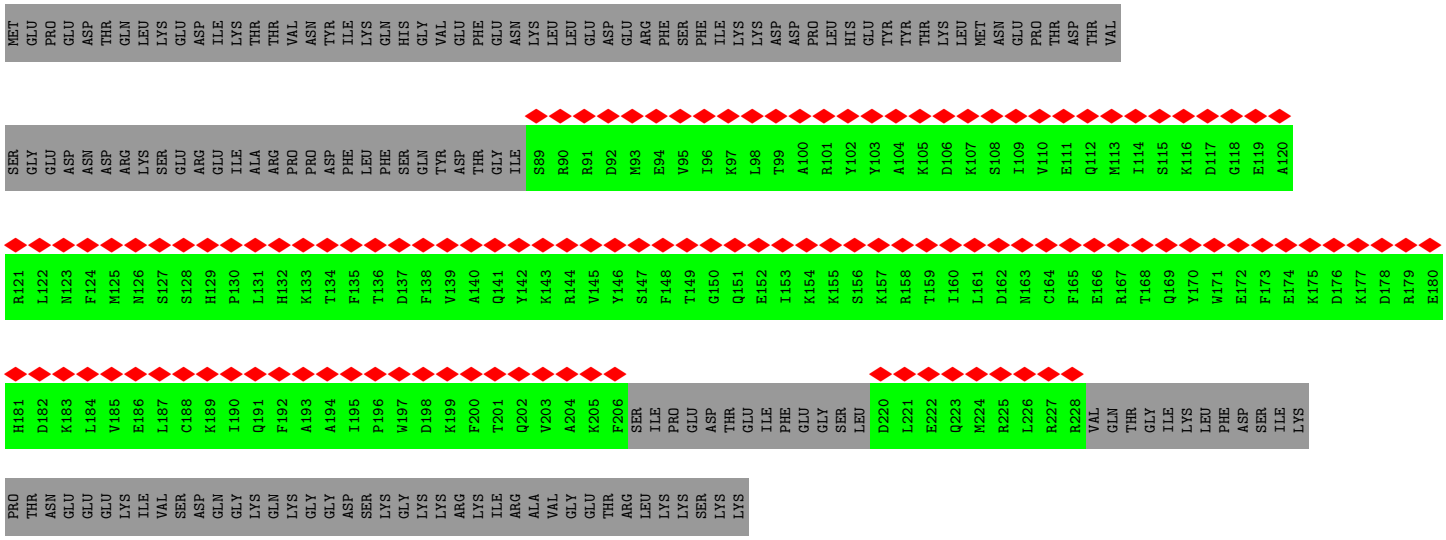


• Molecule 18: PRP9 isoform 1

Chain u:

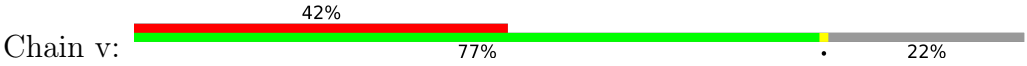


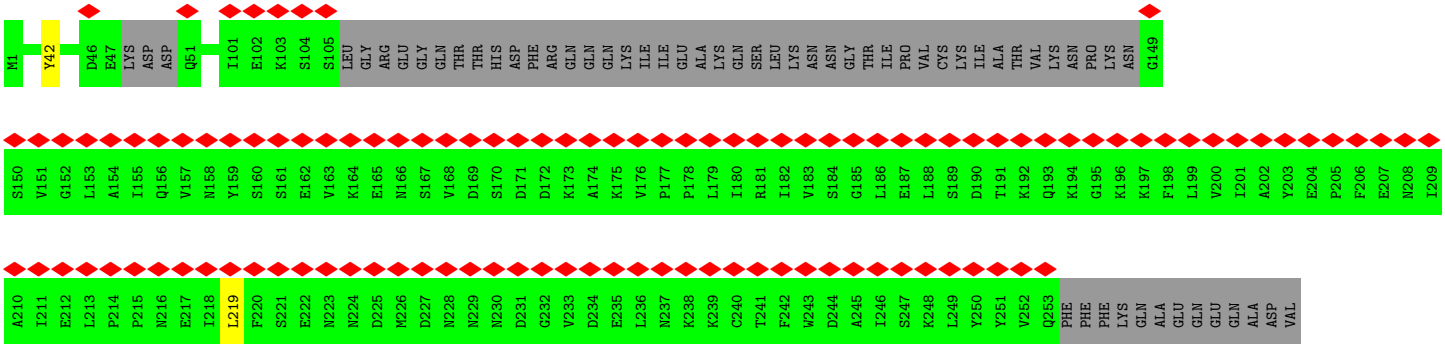
• Molecule 19: Pre-mRNA-splicing factor PRP21



• Molecule 20: PRP11 isoform 1

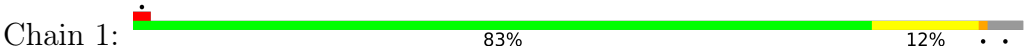
Chain v:

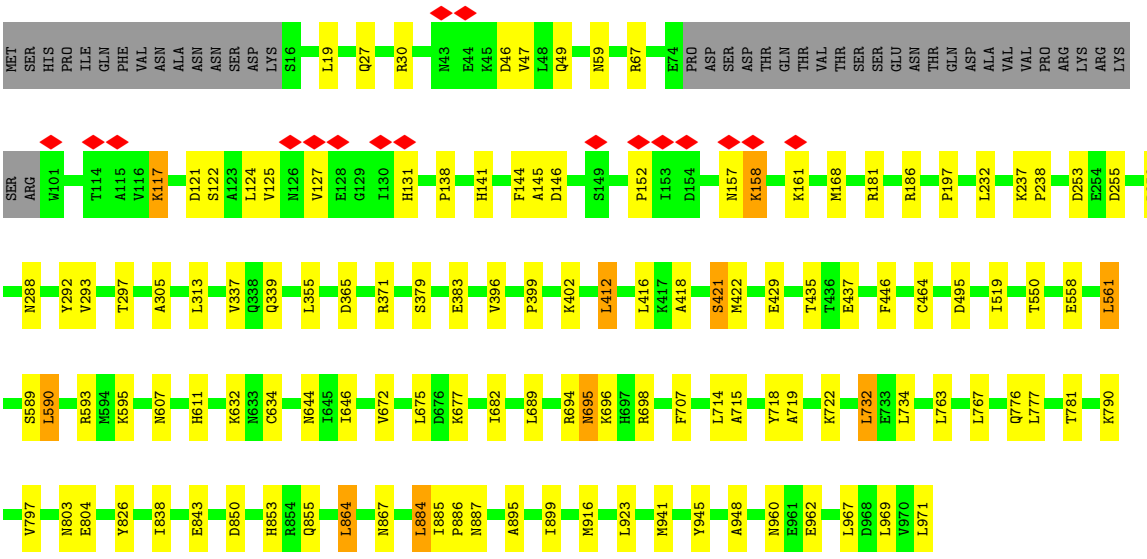




• Molecule 21: HSH155 isoform 1

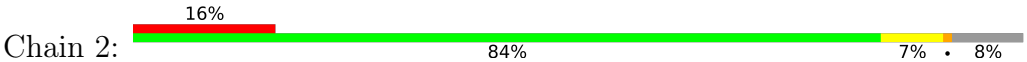
Chain 1:

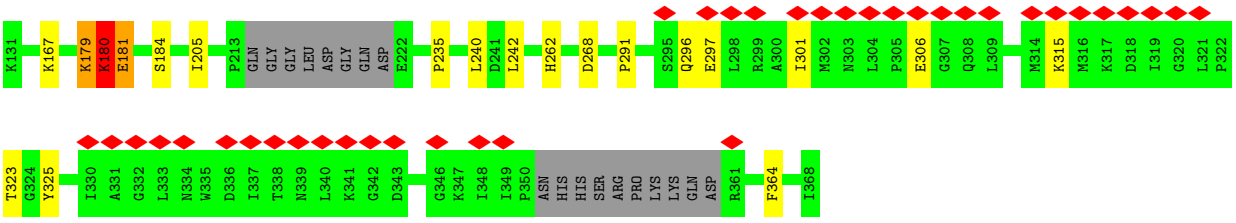




• Molecule 22: HLJ1_G0043010.mRNA.1.CDS.1

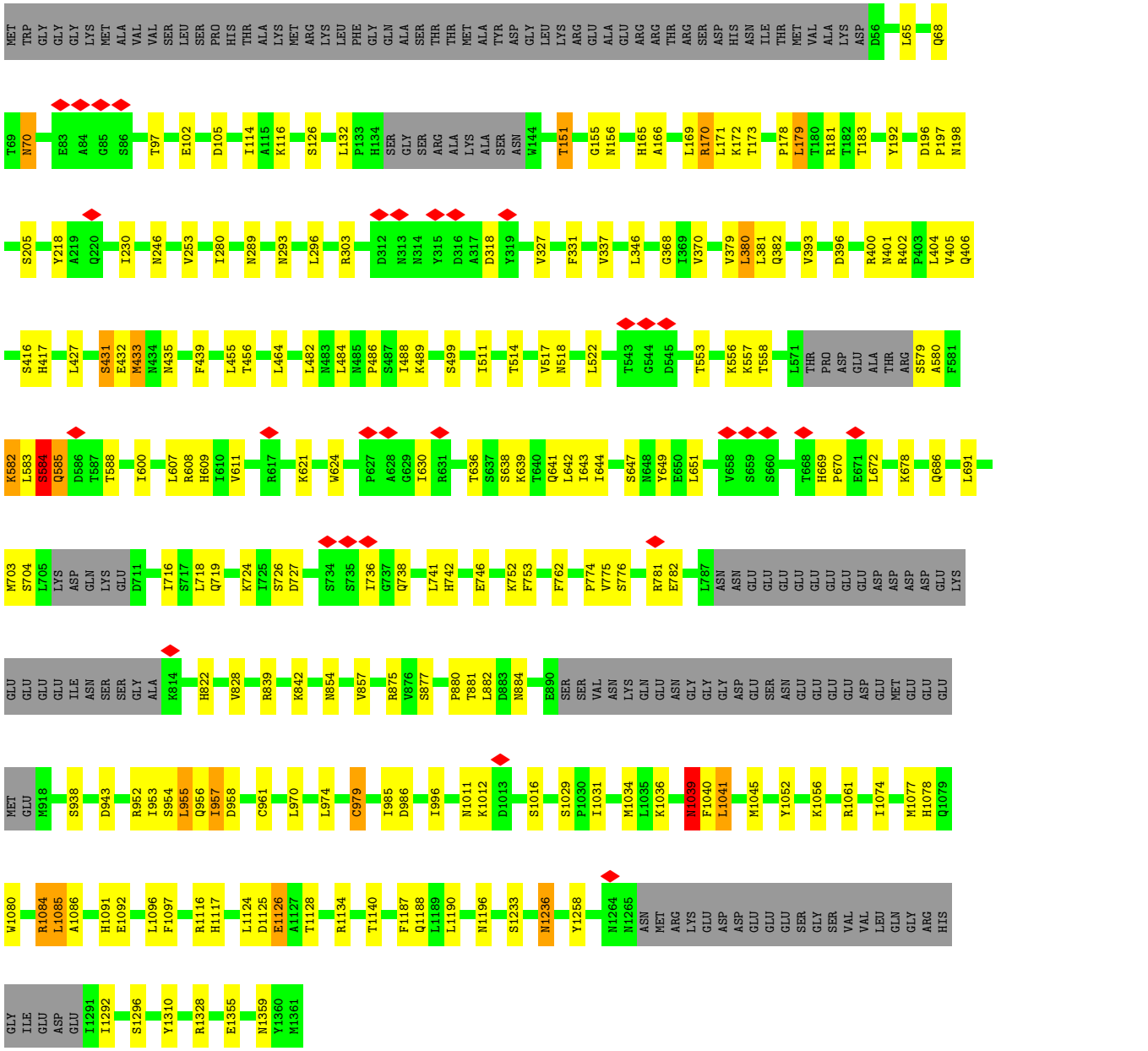
Chain 2:





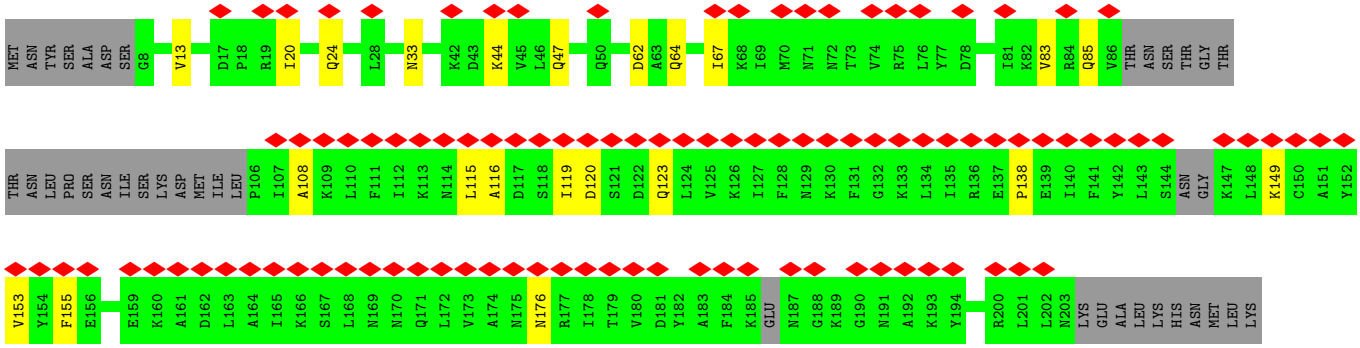
• Molecule 23: RSE1 isoform 1

Chain 3:



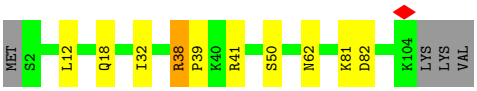
• Molecule 24: HSH49 isoform 1

Chain 4:



• Molecule 25: BJ4_G0056610.mRNA.1.CDS.1

Chain 5:

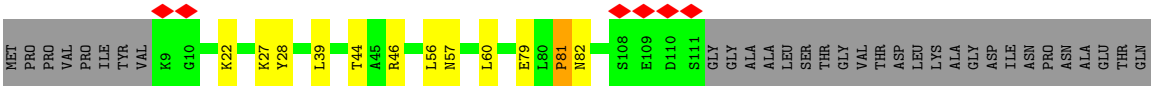


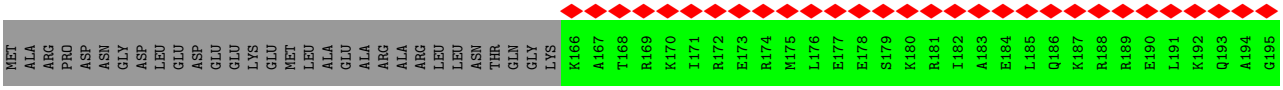
• Molecule 26: RDS3 complex subunit 10

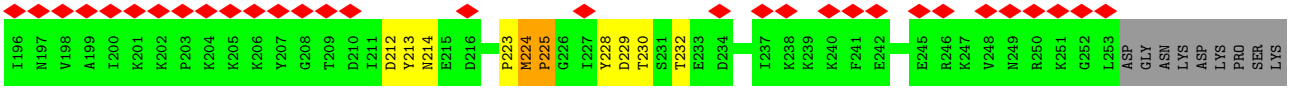
Chain 6:

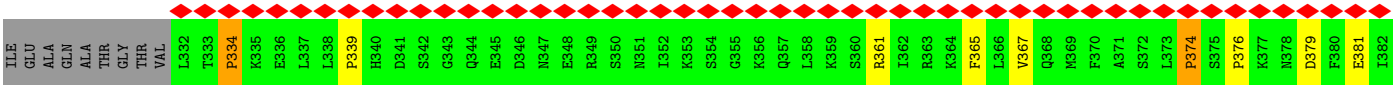
• Molecule 27: Pre-mRNA-splicing factor CEF1

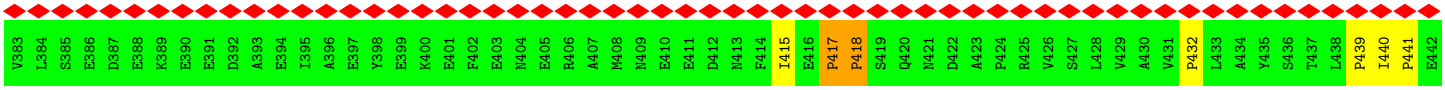
Chain L:

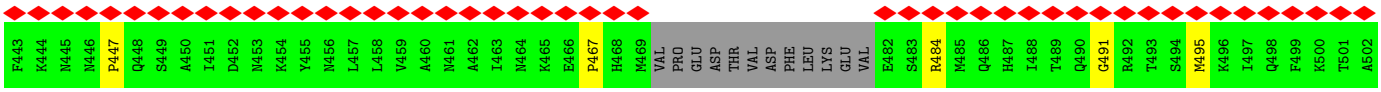


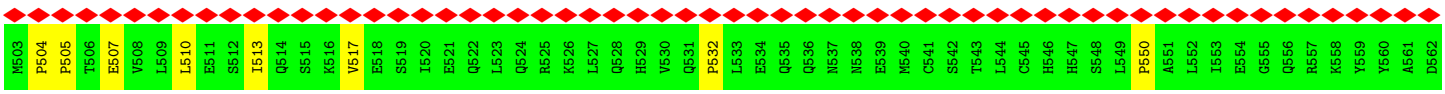


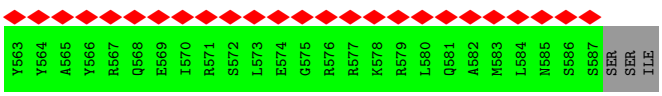






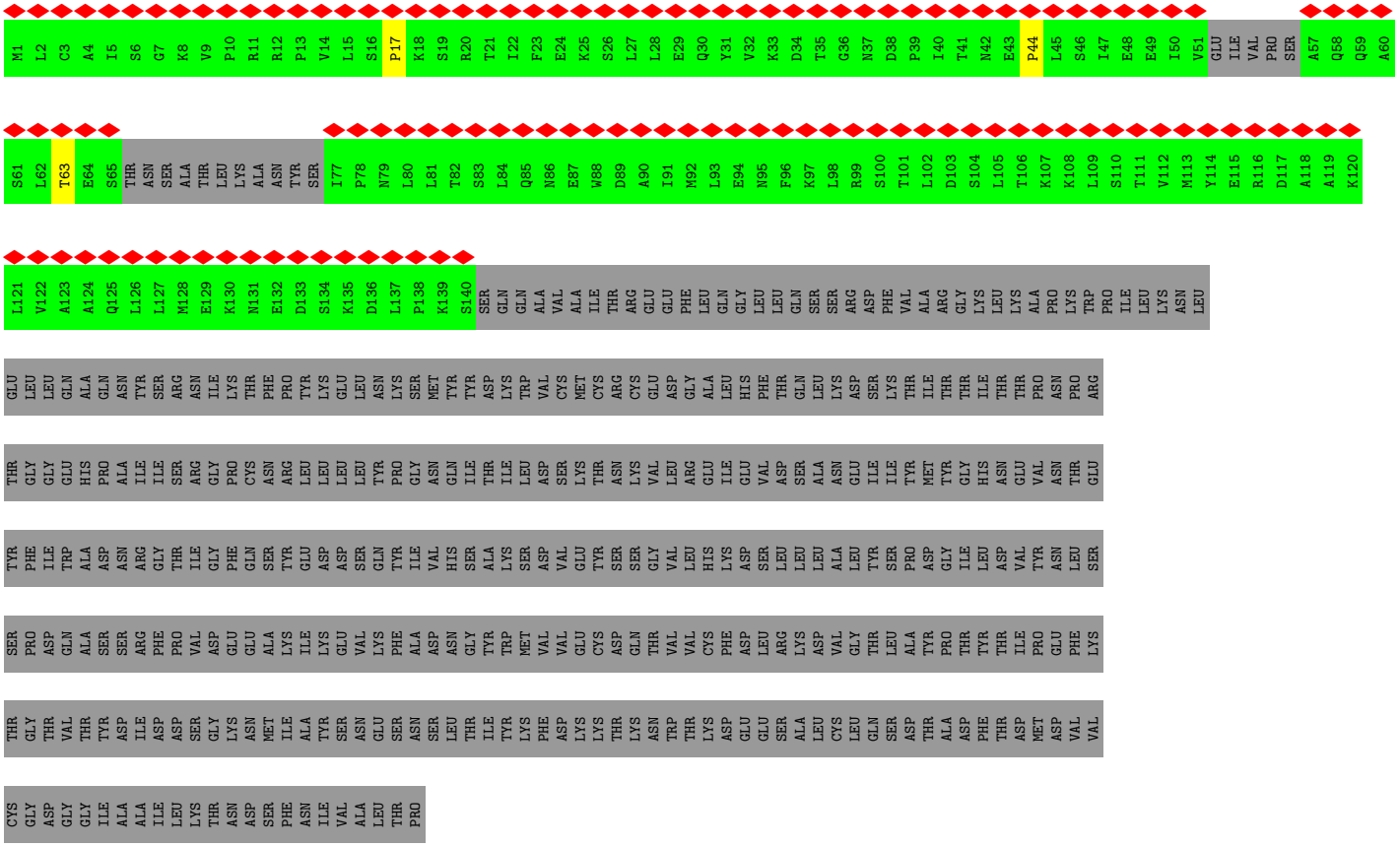






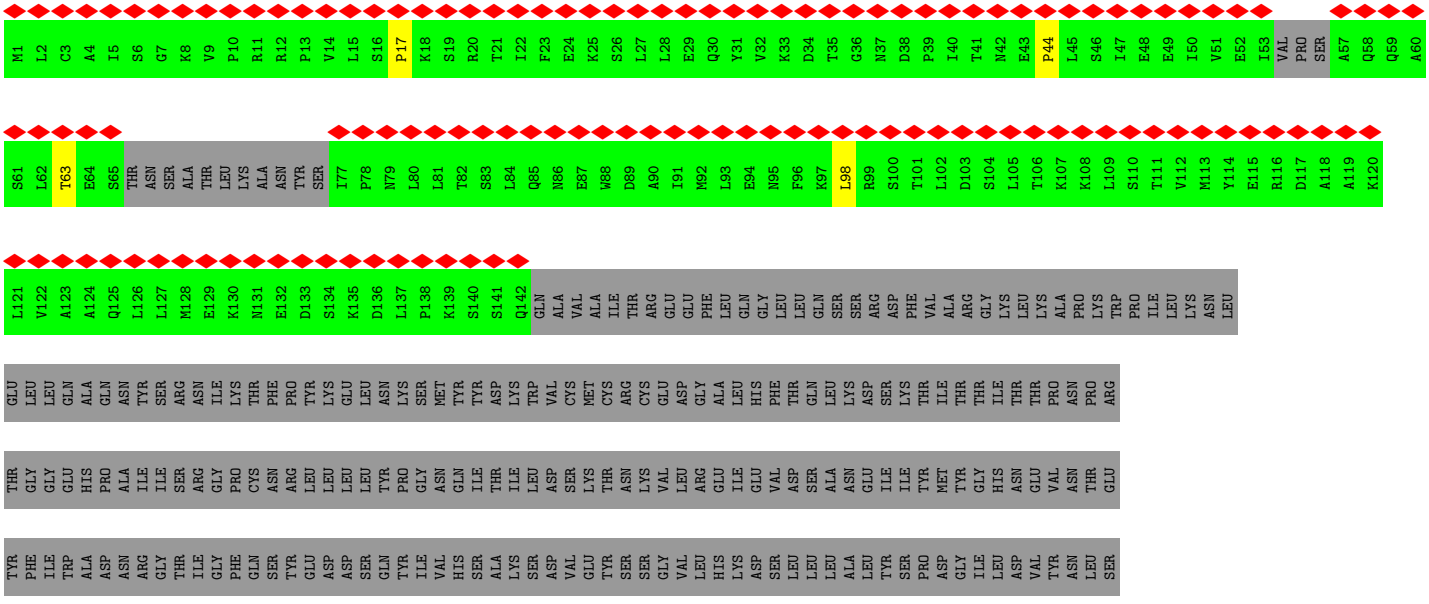
• Molecule 28: Pre-mRNA-processing factor 19

Chain s:



• Molecule 28: Pre-mRNA-processing factor 19

Chain t:



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

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

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

• Molecule 28: Pre-mRNA-processing factor 19

Chain q:

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

S61 LEU THR GLU SER THR ASN SER ALA THR LEU LYS ALA ASN TYR S76 I77 P78 N79 L80 L81 T82 S83 L84 Q85 N86 E87 W88 D89 A90 I91 M92 L93 E94 N95 F96 K97 L98 R99 S100 T101 L102 D103 S104 L105 T106 K107 K108 L109 S110 T111 V112 M113 Y114 E115 R116 D117 A118 A119 K120

L121 V122 A123 A124 Q125 L126 L127 M128 E129 K130 N131 E132 D133 S134 K135 D136 L137 P138 K139 S140 S141 Q142 Q143 ALA VAL ALA ILE THR ARG GLU GLU PHE LEU GLN GLY LEU LEU GLN SER SER ARG ASP PHE VAL ALA ARG GLY LYS LEU LYS ALA PRO LYS TRP PRO ILE LEU LYS ASN LEU

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

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

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

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

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

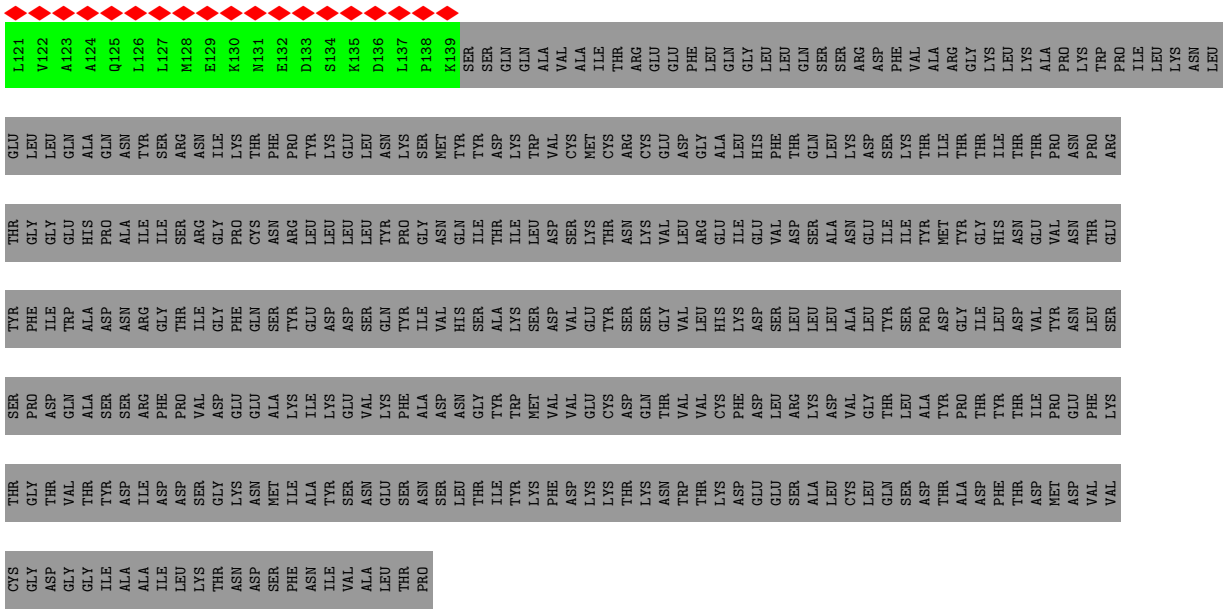
CYS GLY ASP GLY GLY ILE ALA ALA ILE LEU LYS THR ASN ASP SER PHE ASN ILE VAL ALA LEU THR PRO

• Molecule 28: Pre-mRNA-processing factor 19

Chain r:

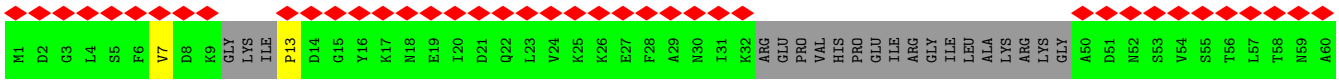
M1 L2 C3 A4 I5 S6 G7 K8 V9 P10 R11 R12 P13 V14 L15 S16 P17 K18 S19 R20 T21 I22 F23 E24 K25 S26 L27 L28 E29 Q30 Y31 V32 K33 D34 T35 G36 N37 D38 P39 I40 T41 N42 E43 P44 L45 S46 I47 E48 E49 I50 V51 E52 I53 VAL PRO SER A57 Q58 Q59 A60

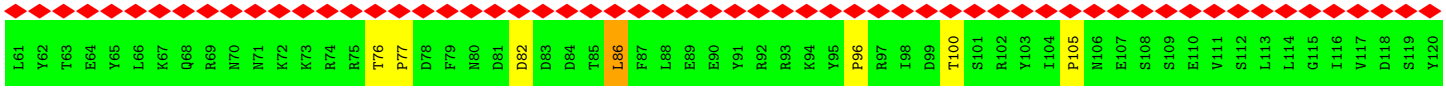
S61 L62 T63 E64 S65 THR ASN SER ALA THR LEU LYS ALA ASN TYR SER I77 P78 N79 L80 L81 T82 S83 L84 Q85 N86 E87 W88 D89 A90 I91 M92 L93 E94 N95 F96 K97 L98 R99 S100 T101 L102 D103 S104 L105 T106 K107 K108 L109 S110 T111 V112 M113 Y114 E115 R116 D117 A118 A119 K120

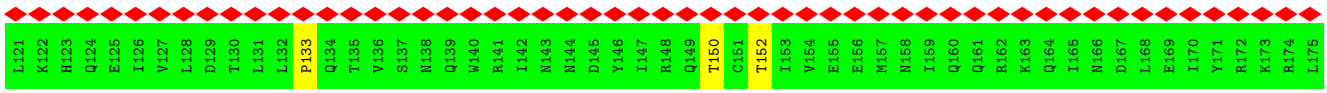


• Molecule 29: SNT309 isoform 1

Chain K:

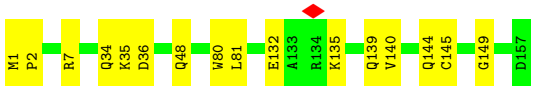






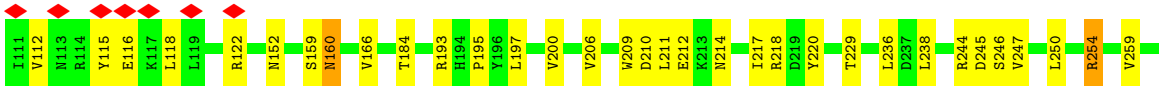
• Molecule 30: BUD31 isoform 1

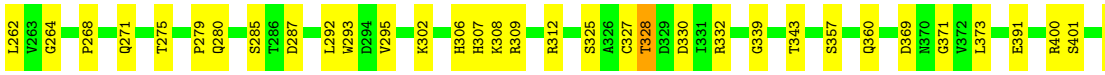
Chain N:

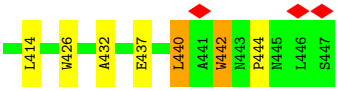


• Molecule 31: HLJ1_G0054350.mRNA.1.CDS.1

Chain T:

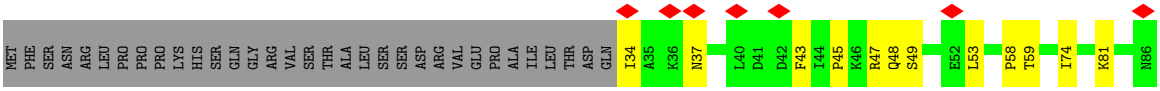


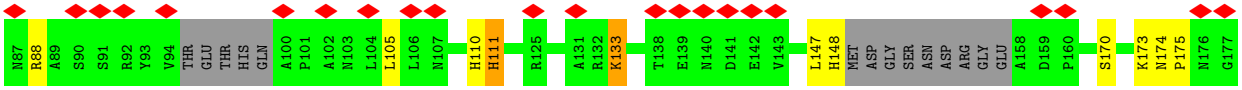


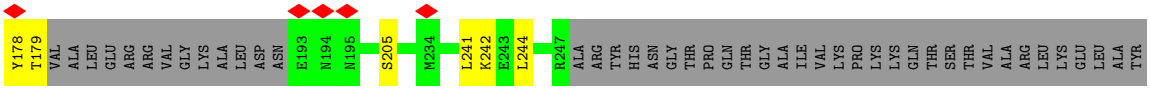


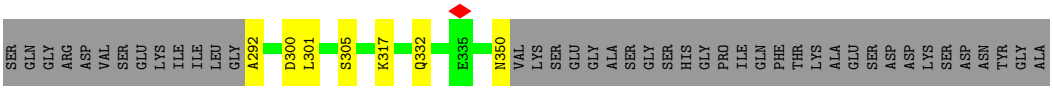
• Molecule 32: Pre-mRNA-processing protein 45

Chain P:



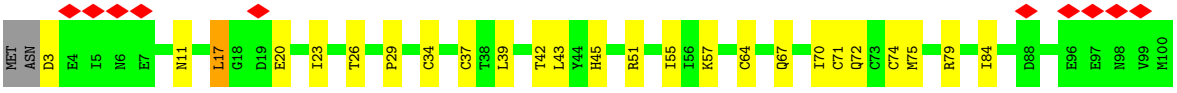


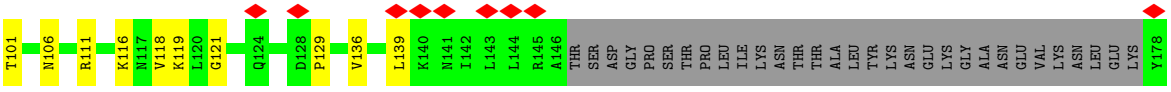


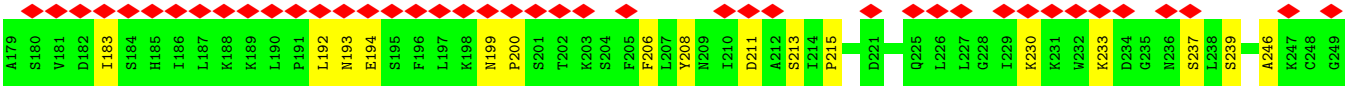


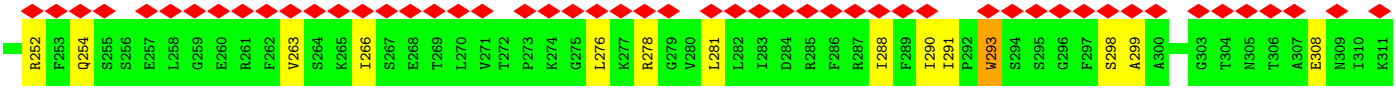
• Molecule 33: Pre-mRNA-splicing factor SLT11

Chain Q:



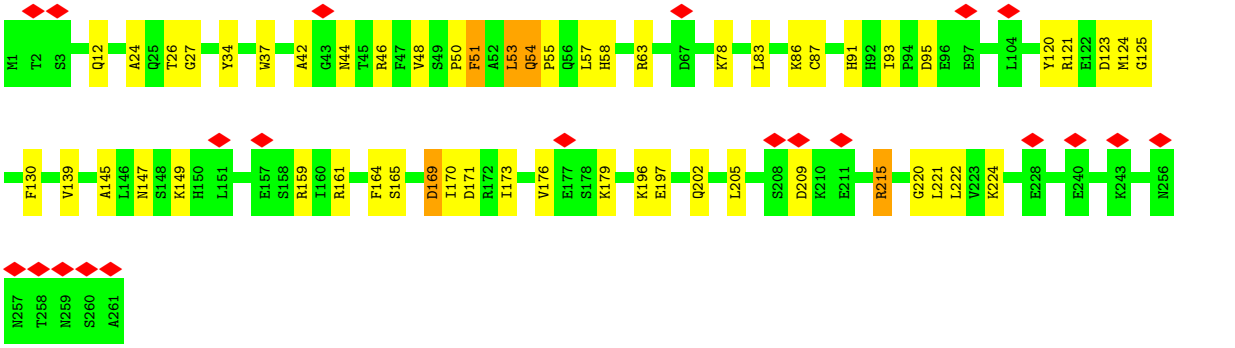




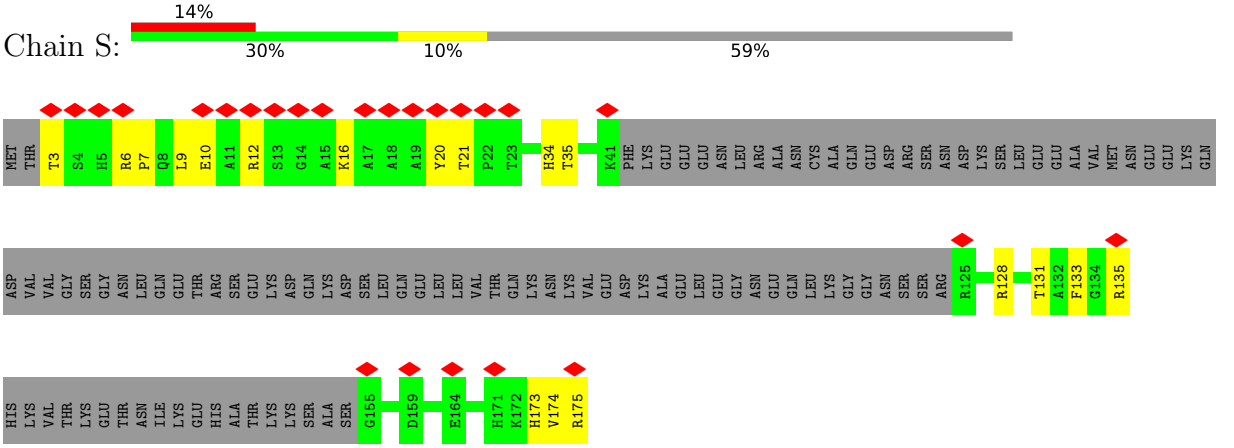


• Molecule 34: Pre-mRNA-splicing factor CWC2

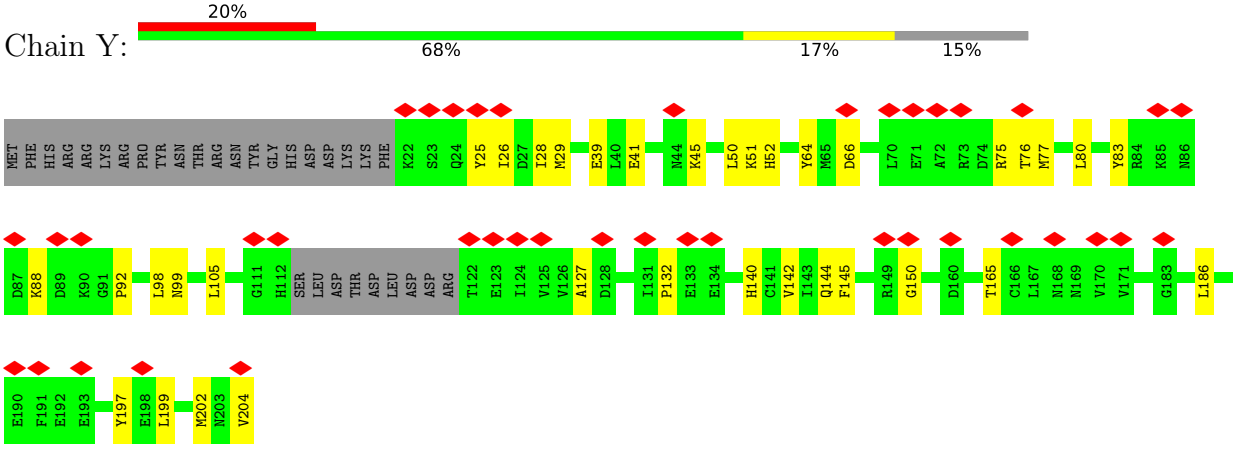
Chain R:



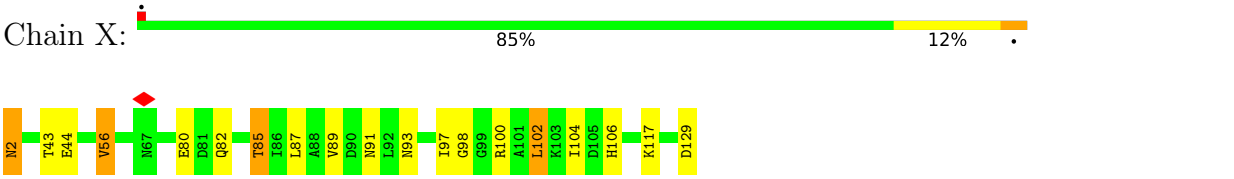
• Molecule 35: Pre-mRNA-splicing factor CWC15



• Molecule 36: Pre-mRNA leakage protein 1

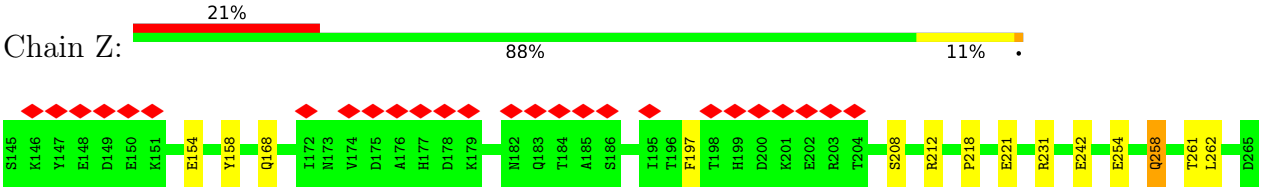


• Molecule 37: SX2_G0027210.mRNA.1.CDS.1



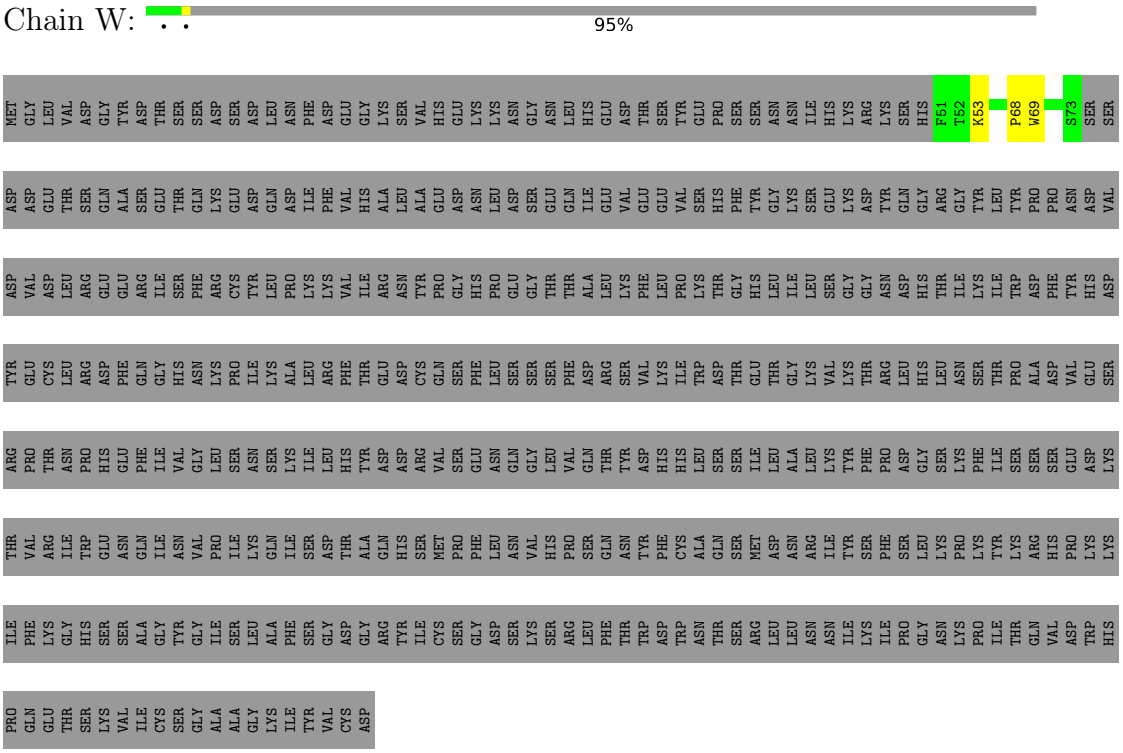
• Molecule 38: Pre-mRNA-splicing factor CWC26

Chain Z:



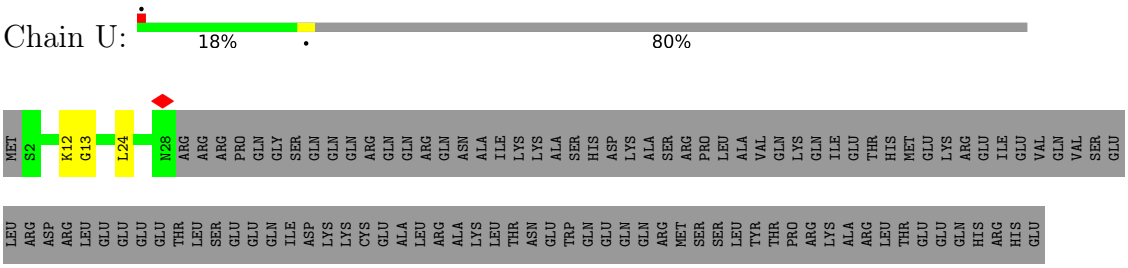
• Molecule 39: CDC40 isoform 1

Chain W:



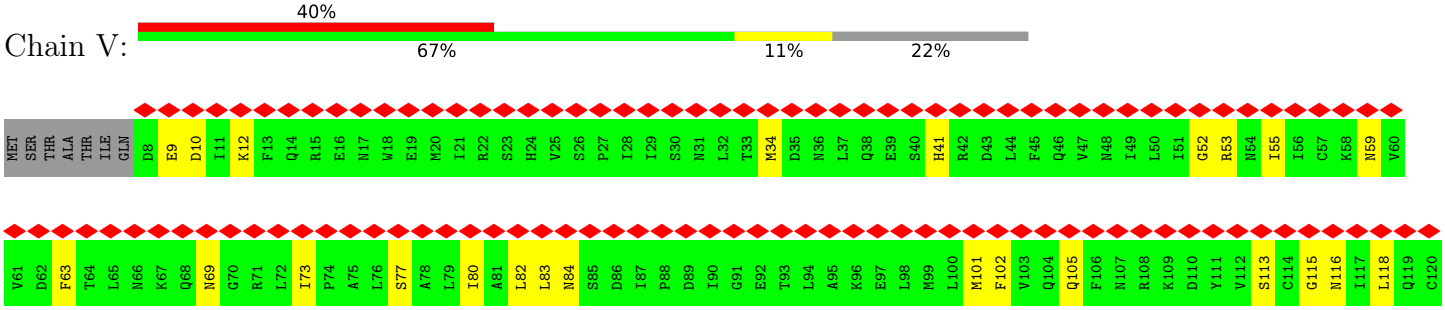
• Molecule 40: Pre-mRNA-splicing factor CWC21

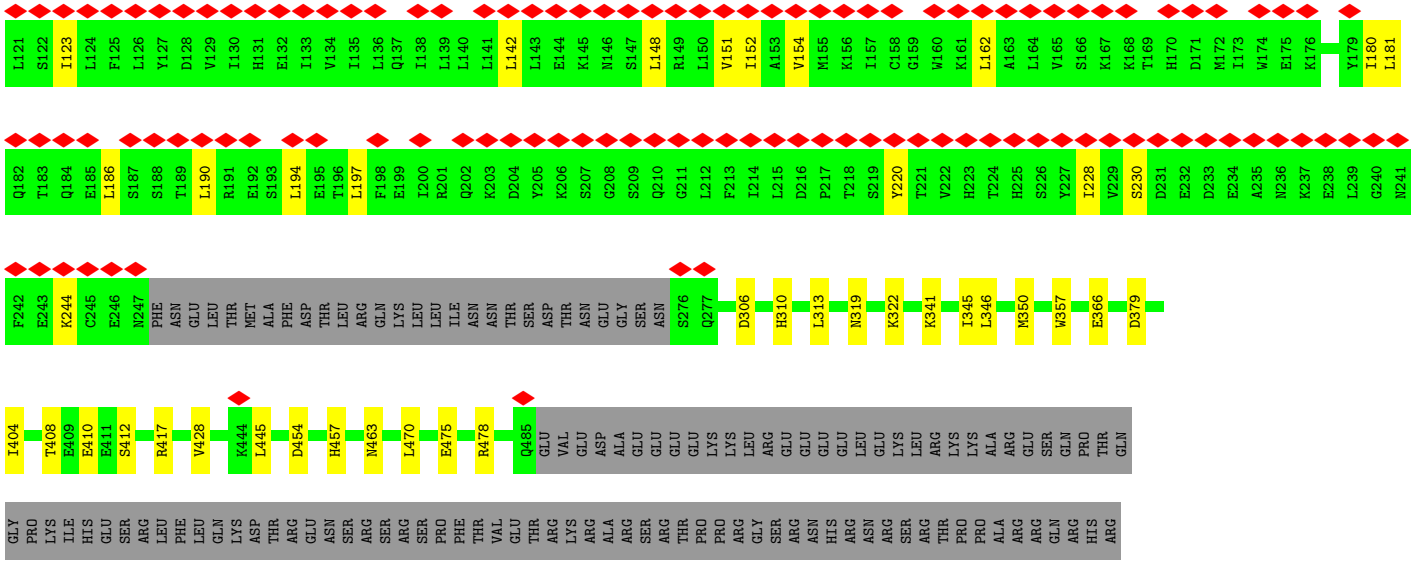
Chain U:



• Molecule 41: CWC22 isoform 1

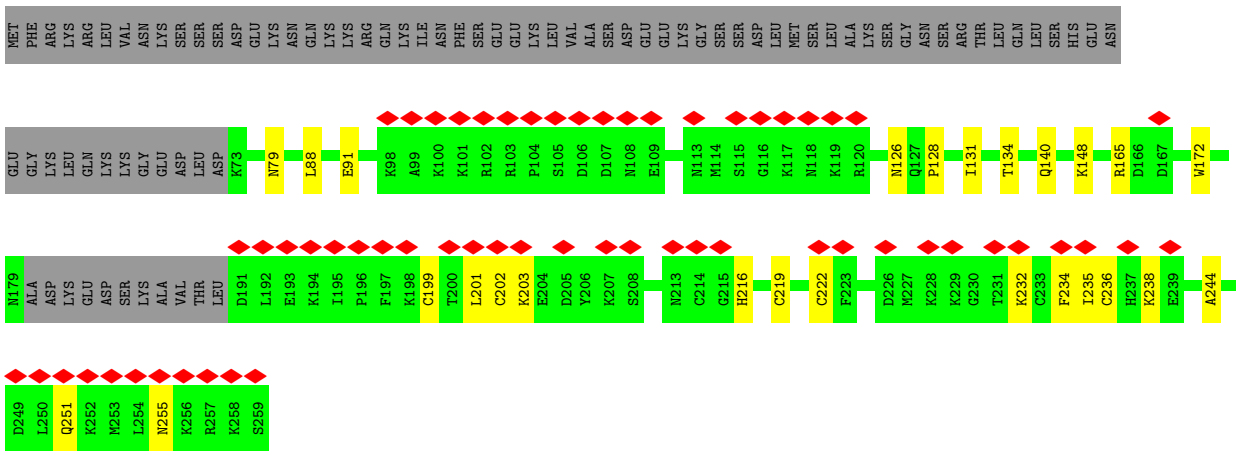
Chain V:





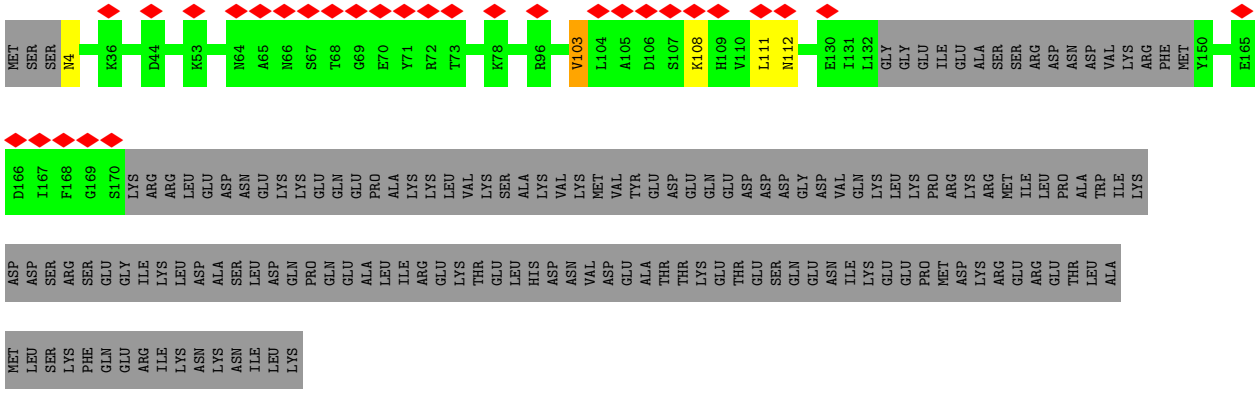
• Molecule 42: Pre-mRNA-splicing factor CWC24

Chain M:



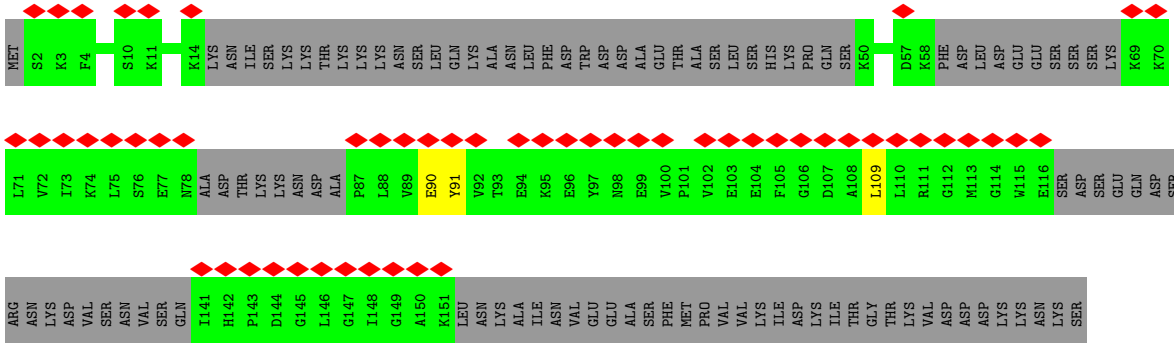
• Molecule 43: CWC27 isoform 1

Chain z:



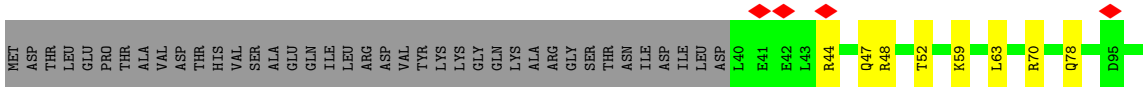
• Molecule 44: Pre-mRNA-splicing factor SPP2

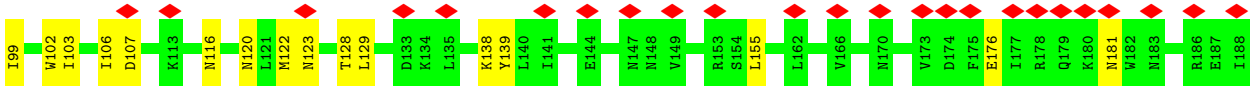
Chain y:

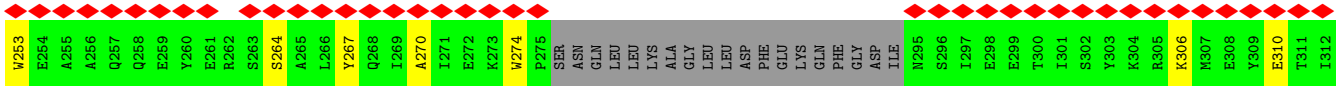


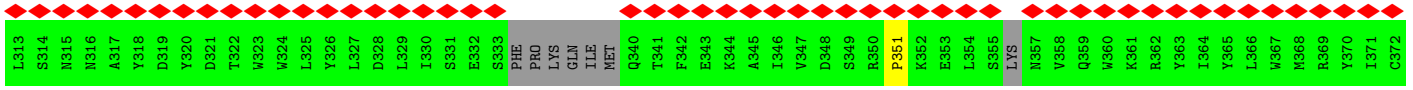
• Molecule 45: CLF1 isoform 1

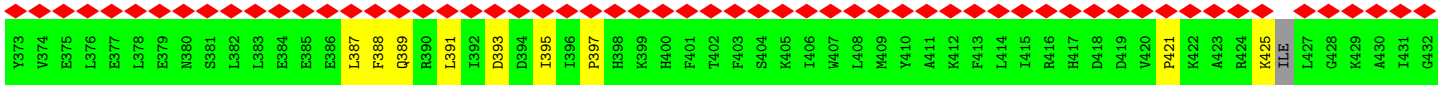
Chain J:

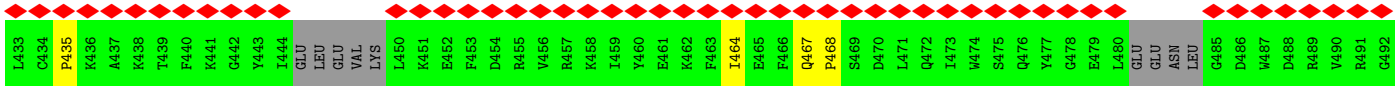


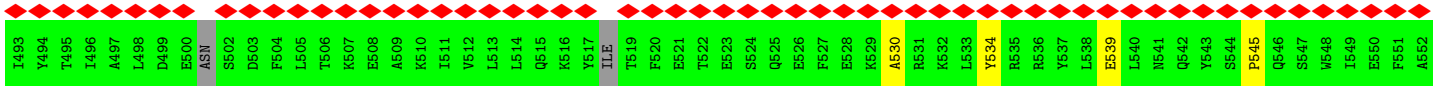


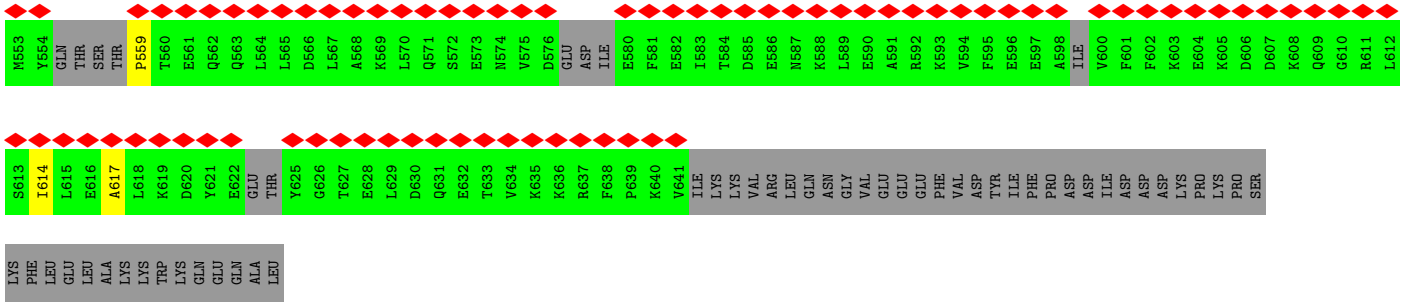






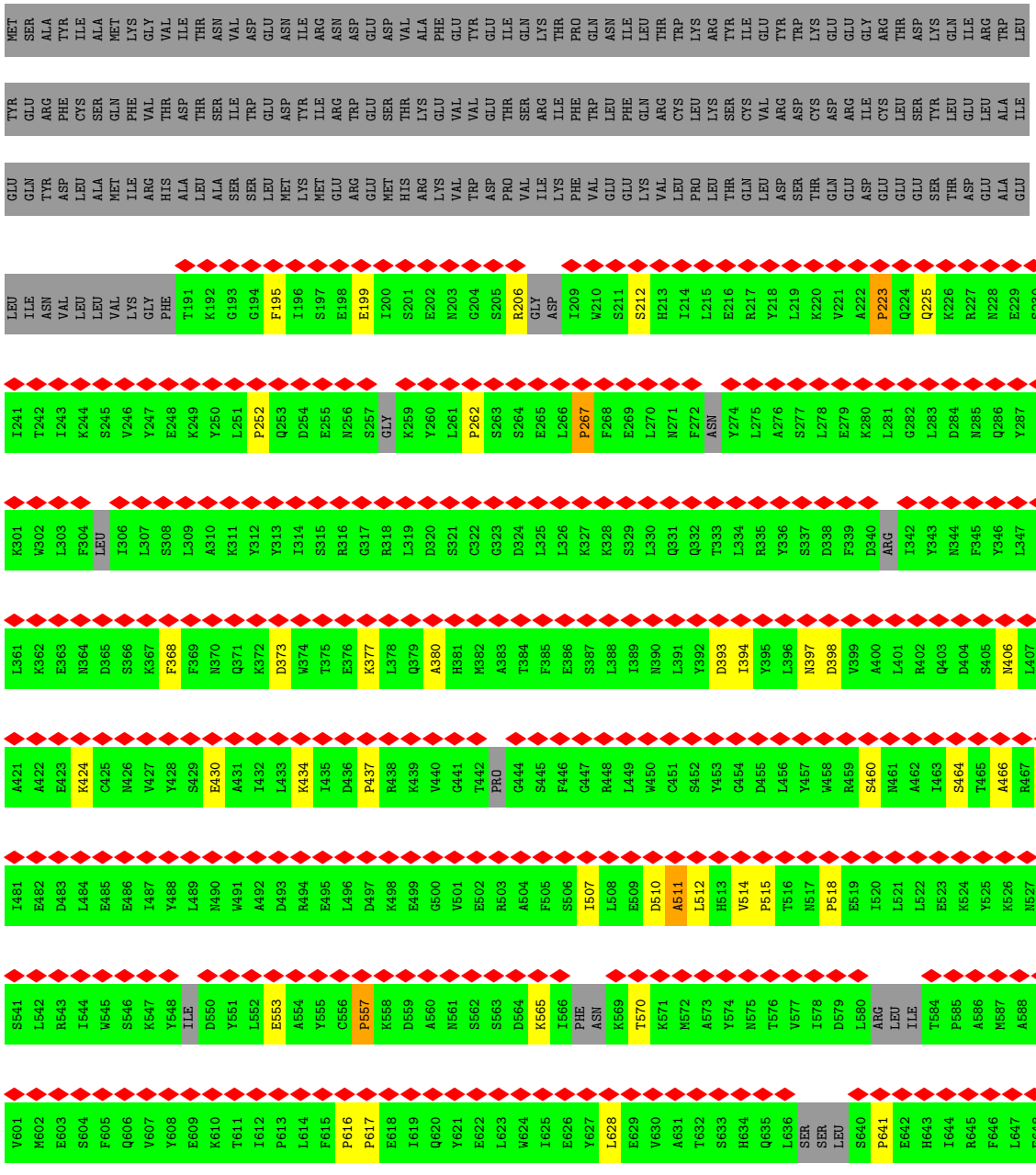


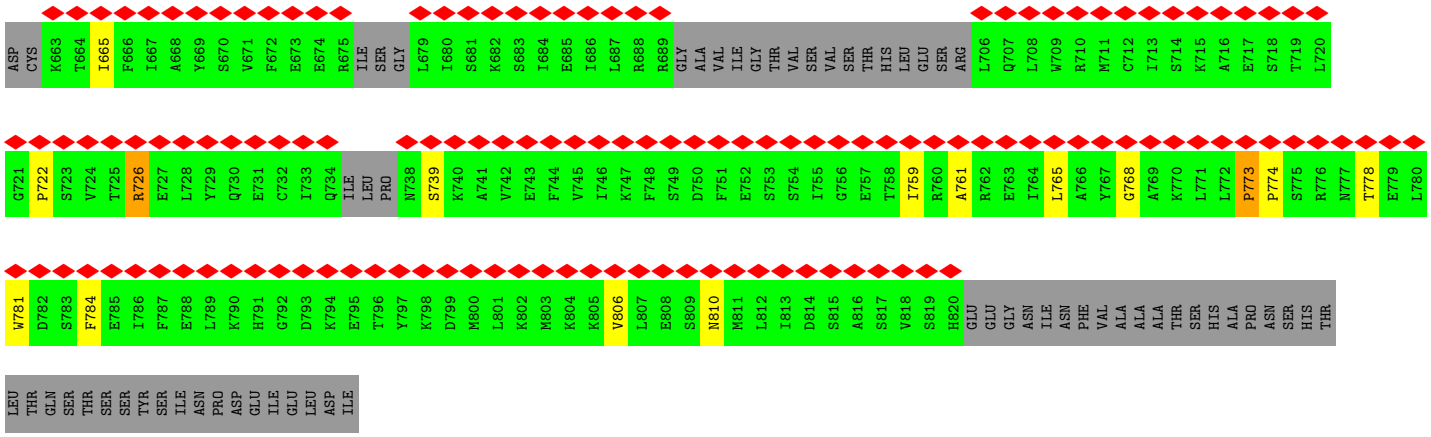




• Molecule 46: SYF1 isoform 1

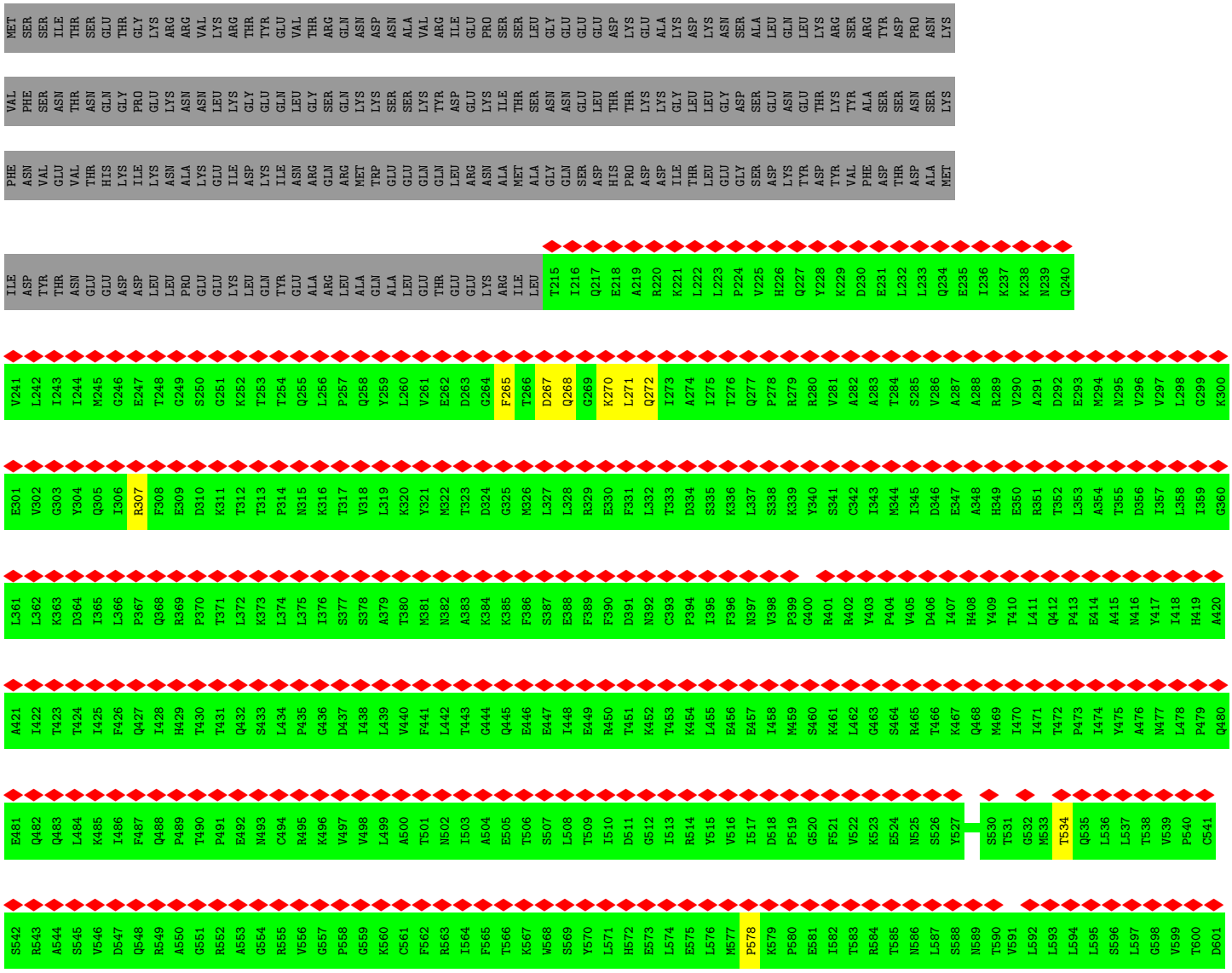
Chain I:

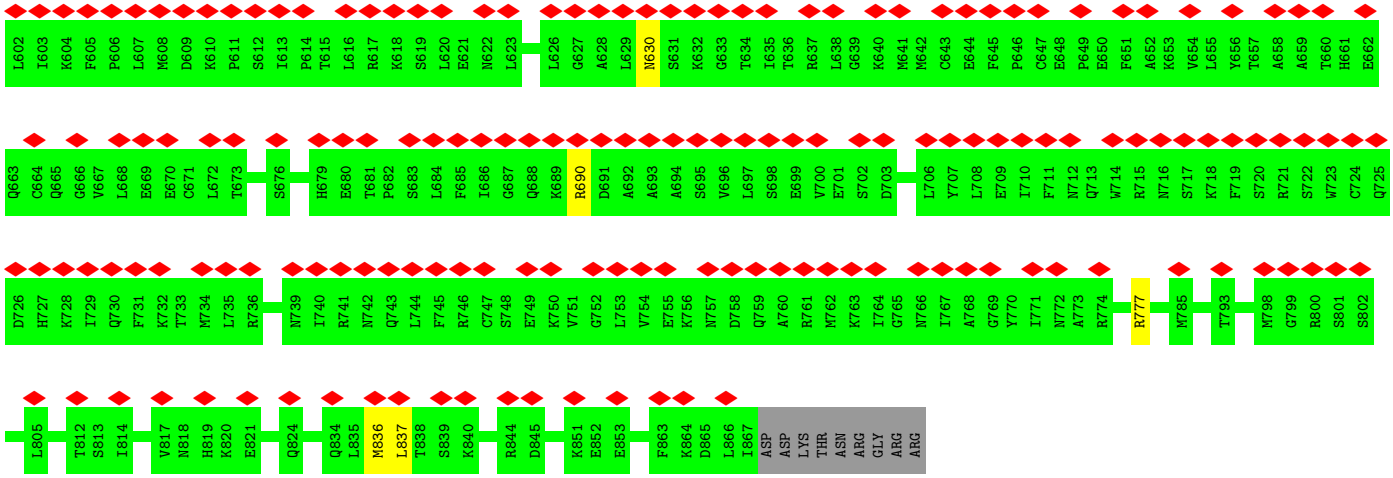




• Molecule 47: Pre-mRNA-splicing factor ATP-dependent RNA helicase-like protein PRP2

Chain x:




L602 I603 K604 F605 P606 L607 M608 D609 K610 P611 S612 I613 P614 T615 L616 R617 K618 S619 L620 E621 N622 L623
L626 G627 A628 L629 N630 S631 K632 G633 T634 I635 T636 R637 L638 G639 K640 M641 M642 C643 E644 F645 P646 C647 E648 P649 E650 F651 A652 K653 V654 L655 Y656 T657 A658 A659 T660 H661 E662
Q663 G664 Q665 G666 V667 L668 E669 E670 C671 L672 T673
S676
H679 E680 T681 P682 S683 L684 F685 I686 G687 Q688 K689 R690 D691 A692 A693 A694 S695 V696 L697 S698 E699 V700 E701 S702 D703
L706 Y707 L708 E709 I710 F711 N712 Q713 W714 R715 N716 S717 K718 F719 S720 R721 S722 W723 C724 Q725
D726 H727 K728 I729 Q730 F731 K732 T733 M734 L735 R736
N739 I740 R741 N742 Q743 L744 F745 R746 C747 S748 E749 K750 V751 G752 L753 V754 E755 K756 N757 D758 Q759 A760 R761 M762 K763 I764 G765 N766 I767 A768 G769 Y770 I771 N772 A773 R774
R777
M785
T793
M798 G799 R800 S801 S802
L805
T812 S813 I814
V817 N818 H819 K820 E821
Q824
Q834 L835 M836 L837 T838 S839 K840
R844 D845
K851 E852 E853
F863 K864 D865 L866 I867
ASP ASP LYS THR ASN ARG GLY ARG ARG

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	705371	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 ($e^-/Å^2$)	50	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.251	Depositor
Minimum map value	-0.119	Depositor
Average map value	-0.000	Depositor
Map value standard deviation	0.006	Depositor
Recommended contour level	0.015	Depositor
Map size (Å)	423.99997, 423.99997, 423.99997	wwPDB
Map dimensions	400, 400, 400	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.06, 1.06, 1.06	Depositor

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: CA, MG, ZN, GTP, IHP

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	0.61	0/18595	0.69	0/25207
2	B	0.60	0/4239	0.90	0/6598
3	C	0.41	0/7484	0.64	0/10135
4	D	0.33	0/14978	0.54	0/20302
5	d	0.24	0/315	0.50	0/392
6	a	0.23	0/290	0.50	0/359
6	h	0.29	0/615	0.62	0/829
7	b	0.23	0/305	0.52	0/376
7	m	0.29	0/649	0.54	0/880
8	c	0.23	0/358	0.51	0/444
8	n	0.30	0/535	0.62	0/717
9	e	0.22	0/285	0.48	0/351
9	i	0.31	0/585	0.57	0/795
10	f	0.24	0/278	0.57	0/344
10	j	0.30	0/564	0.65	0/761
11	g	0.22	0/277	0.54	0/341
11	k	0.28	0/532	0.61	0/715
12	F	0.60	0/2452	0.89	0/3817
13	G	0.48	0/2338	0.85	0/3621
14	H	0.43	0/3969	0.89	0/6159
15	o	0.25	0/839	0.52	0/1127
16	p	0.27	0/478	0.49	0/640
17	l	0.31	0/620	0.57	0/841
18	u	0.34	0/3976	0.53	0/5327
19	w	0.29	0/1105	0.47	0/1475
20	v	0.51	0/1647	0.67	0/2213
21	1	0.60	0/7522	0.72	1/10203 (0.0%)
22	2	0.54	0/1840	0.65	0/2484
23	3	0.57	0/9789	0.73	0/13273
24	4	0.35	0/1457	0.57	0/1959
25	5	0.69	0/827	0.76	0/1105
26	6	0.70	0/702	0.79	0/939

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
27	L	0.39	0/2924	0.62	11/3954 (0.3%)
28	q	0.27	0/856	0.54	0/1155
28	r	0.26	0/828	0.50	0/1117
28	s	0.26	0/824	0.49	0/1111
28	t	0.26	0/848	0.51	0/1143
29	K	0.24	0/918	0.38	0/1236
30	N	0.55	0/1315	0.69	0/1759
31	T	0.56	0/2704	0.72	0/3676
32	P	0.47	0/2008	0.67	0/2703
33	Q	0.37	0/2339	0.69	0/3154
34	R	0.42	0/2135	0.61	0/2871
35	S	0.43	0/592	0.70	0/790
36	Y	0.35	0/1408	0.58	0/1900
37	X	0.58	0/1071	0.66	0/1445
38	Z	0.53	0/936	0.61	0/1264
39	W	0.40	0/200	0.62	0/264
40	U	0.46	0/191	0.72	0/254
41	V	0.44	0/3720	0.63	0/5016
42	M	0.56	0/1385	0.64	0/1862
43	z	0.42	0/1252	0.64	0/1692
44	y	0.37	0/577	0.64	0/765
45	J	0.33	0/3635	0.57	7/4962 (0.1%)
46	I	0.28	0/3095	0.57	14/4242 (0.3%)
47	x	0.53	0/5290	0.61	0/7155
All	All	0.48	0/131496	0.67	33/180219 (0.0%)

There are no bond length outliers.

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
27	L	417	PRO	N-CA-CB	6.75	111.40	103.30
27	L	418	PRO	N-CA-CB	6.61	111.23	103.30
46	I	557	PRO	N-CA-CB	6.23	110.77	103.30
27	L	334	PRO	N-CA-CB	6.16	110.69	103.30
46	I	773	PRO	N-CA-CB	6.08	110.60	103.30

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	A	18135	0	18054	191	0
2	B	3795	0	1919	66	0
3	C	7328	0	7505	104	0
4	D	14666	0	14661	120	0
5	d	316	0	86	0	0
6	a	292	0	78	0	0
6	h	610	0	640	0	0
7	b	308	0	78	0	0
7	m	644	0	686	0	0
8	c	360	0	89	0	0
8	n	528	0	573	0	0
9	e	288	0	74	0	0
9	i	575	0	597	0	0
10	f	280	0	77	0	0
10	j	554	0	556	0	0
11	g	280	0	79	0	0
11	k	529	0	557	0	0
12	F	2192	0	1106	20	0
13	G	2099	0	1055	34	0
14	H	3566	0	1809	51	0
15	o	841	0	614	0	0
16	p	476	0	378	0	0
17	l	611	0	627	0	0
18	u	3899	0	3826	0	0
19	w	1084	0	1081	0	0
20	v	1621	0	1596	0	0
21	1	7376	0	7521	80	0
22	2	1793	0	1827	16	0
23	3	9599	0	9679	142	0
24	4	1433	0	1461	14	0
25	5	814	0	809	5	0
26	6	693	0	705	11	0
27	L	2901	0	2325	27	0
28	q	850	0	682	0	0
28	r	823	0	654	0	0
28	s	819	0	657	0	0
28	t	843	0	672	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
29	K	920	0	603	4	0
30	N	1291	0	1312	8	0
31	T	2646	0	2639	53	0
32	P	1978	0	1981	37	0
33	Q	2301	0	2366	45	0
34	R	2089	0	2053	47	0
35	S	578	0	565	27	0
36	Y	1386	0	1353	27	0
37	X	1051	0	1015	23	0
38	Z	920	0	841	14	0
39	W	195	0	198	2	0
40	U	190	0	186	3	0
41	V	3660	0	3706	39	0
42	M	1360	0	1280	16	0
43	z	1224	0	1217	0	0
44	y	572	0	597	0	0
45	J	3595	0	2639	39	0
46	I	3101	0	1639	33	0
47	x	5193	0	5347	0	0
48	A	36	0	6	2	0
49	C	32	0	12	2	0
50	3	1	0	0	0	0
50	C	1	0	0	0	0
50	F	4	0	0	0	0
51	C	1	0	0	0	0
52	5	3	0	0	0	0
52	M	3	0	0	0	0
52	N	3	0	0	0	0
52	Q	2	0	0	0	0
52	R	1	0	0	0	0
52	u	2	0	0	0	0
52	v	1	0	0	0	0
53	1	185	0	0	4	0
53	2	51	0	0	1	0
53	3	218	0	0	7	0
53	5	53	0	0	0	0
53	6	42	0	0	1	0
53	A	582	0	0	9	0
53	B	127	0	0	2	0
53	C	8	0	0	0	0
53	D	2	0	0	0	0
53	F	109	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
53	G	56	0	0	1	0
53	H	70	0	0	0	0
53	L	18	0	0	0	0
53	M	21	0	0	0	0
53	N	12	0	0	0	0
53	P	26	0	0	2	0
53	R	8	0	0	0	0
53	S	9	0	0	0	0
53	T	39	0	0	0	0
53	U	6	0	0	0	0
53	V	21	0	0	1	0
53	X	13	0	0	1	0
53	Z	9	0	0	0	0
53	v	29	0	0	0	0
All	All	129875	0	116948	1119	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 6.

The worst 5 of 1119 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)
1:A:784:GLN:HG2	14:H:19:U:C5	1.67	1.28
46:I:466:ALA:CB	46:I:470:TRP:CB	2.20	1.19
27:L:224:MET:SD	27:L:225:PRO:HD3	1.82	1.19
27:L:224:MET:SD	27:L:225:PRO:CD	2.33	1.17
46:I:466:ALA:HB3	46:I:470:TRP:CB	1.76	1.14

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	A	2199/2413 (91%)	2127 (97%)	68 (3%)	4 (0%)	47 68
3	C	912/1008 (90%)	877 (96%)	33 (4%)	2 (0%)	47 68
4	D	1820/2163 (84%)	1759 (97%)	60 (3%)	1 (0%)	51 73
5	d	77/101 (76%)	69 (90%)	8 (10%)	0	100 100
6	a	69/196 (35%)	64 (93%)	5 (7%)	0	100 100
6	h	74/196 (38%)	70 (95%)	4 (5%)	0	100 100
7	b	71/146 (49%)	69 (97%)	2 (3%)	0	100 100
7	m	78/146 (53%)	77 (99%)	1 (1%)	0	100 100
8	c	86/110 (78%)	79 (92%)	7 (8%)	0	100 100
8	n	63/110 (57%)	57 (90%)	6 (10%)	0	100 100
9	e	66/94 (70%)	63 (96%)	3 (4%)	0	100 100
9	i	71/94 (76%)	70 (99%)	1 (1%)	0	100 100
10	f	66/86 (77%)	59 (89%)	7 (11%)	0	100 100
10	j	66/86 (77%)	65 (98%)	1 (2%)	0	100 100
11	g	64/77 (83%)	61 (95%)	3 (5%)	0	100 100
11	k	65/77 (84%)	64 (98%)	1 (2%)	0	100 100
15	o	125/238 (52%)	115 (92%)	8 (6%)	2 (2%)	9 17
16	p	73/111 (66%)	73 (100%)	0	0	100 100
17	l	79/81 (98%)	73 (92%)	6 (8%)	0	100 100
18	u	453/530 (86%)	423 (93%)	29 (6%)	1 (0%)	47 68
19	w	123/280 (44%)	117 (95%)	6 (5%)	0	100 100
20	v	201/266 (76%)	185 (92%)	16 (8%)	0	100 100
21	1	926/971 (95%)	898 (97%)	23 (2%)	5 (0%)	29 48
22	2	214/238 (90%)	208 (97%)	4 (2%)	2 (1%)	17 31
23	3	1193/1361 (88%)	1124 (94%)	65 (5%)	4 (0%)	41 61
24	4	166/213 (78%)	161 (97%)	5 (3%)	0	100 100
25	5	101/107 (94%)	99 (98%)	2 (2%)	0	100 100
26	6	82/84 (98%)	81 (99%)	0	1 (1%)	13 24
27	L	427/590 (72%)	391 (92%)	26 (6%)	10 (2%)	6 10
28	q	125/503 (25%)	116 (93%)	9 (7%)	0	100 100
28	r	119/503 (24%)	114 (96%)	5 (4%)	0	100 100
28	s	118/503 (24%)	114 (97%)	4 (3%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
28	t	122/503 (24%)	118 (97%)	4 (3%)	0	100	100
29	K	149/175 (85%)	140 (94%)	7 (5%)	2 (1%)	12	21
30	N	155/157 (99%)	148 (96%)	6 (4%)	1 (1%)	25	43
31	T	335/337 (99%)	317 (95%)	14 (4%)	4 (1%)	13	24
32	P	236/379 (62%)	228 (97%)	5 (2%)	3 (1%)	12	21
33	Q	288/364 (79%)	265 (92%)	20 (7%)	3 (1%)	15	28
34	R	259/261 (99%)	241 (93%)	17 (7%)	1 (0%)	34	54
35	S	65/175 (37%)	61 (94%)	4 (6%)	0	100	100
36	Y	170/204 (83%)	163 (96%)	6 (4%)	1 (1%)	25	43
37	X	126/128 (98%)	122 (97%)	4 (3%)	0	100	100
38	Z	119/121 (98%)	117 (98%)	2 (2%)	0	100	100
39	W	21/455 (5%)	20 (95%)	1 (5%)	0	100	100
40	U	25/135 (18%)	24 (96%)	1 (4%)	0	100	100
41	V	446/577 (77%)	427 (96%)	19 (4%)	0	100	100
42	M	172/259 (66%)	163 (95%)	8 (5%)	1 (1%)	25	43
43	z	146/301 (48%)	126 (86%)	18 (12%)	2 (1%)	11	20
44	y	63/185 (34%)	57 (90%)	6 (10%)	0	100	100
45	J	528/687 (77%)	493 (93%)	31 (6%)	4 (1%)	19	35
46	I	544/859 (63%)	491 (90%)	36 (7%)	17 (3%)	4	5
47	x	651/876 (74%)	617 (95%)	33 (5%)	1 (0%)	47	68
All	All	14992/20820 (72%)	14260 (95%)	660 (4%)	72 (0%)	32	48

5 of 72 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	91	VAL
22	2	268	ASP
23	3	585	GLN
27	L	367	VAL
27	L	417	PRO

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	1997/2182 (92%)	1950 (98%)	47 (2%)	49 74
3	C	828/910 (91%)	813 (98%)	15 (2%)	59 81
4	D	1651/1955 (84%)	1649 (100%)	2 (0%)	93 98
6	h	67/176 (38%)	67 (100%)	0	100 100
7	m	77/129 (60%)	77 (100%)	0	100 100
8	n	59/103 (57%)	59 (100%)	0	100 100
9	i	65/83 (78%)	65 (100%)	0	100 100
10	j	61/77 (79%)	61 (100%)	0	100 100
11	k	58/66 (88%)	55 (95%)	3 (5%)	23 44
15	o	44/219 (20%)	42 (96%)	2 (4%)	27 51
16	p	23/100 (23%)	22 (96%)	1 (4%)	29 53
17	l	67/70 (96%)	67 (100%)	0	100 100
18	u	426/492 (87%)	422 (99%)	4 (1%)	78 92
19	w	118/259 (46%)	118 (100%)	0	100 100
20	v	182/236 (77%)	180 (99%)	2 (1%)	73 89
21	1	817/867 (94%)	788 (96%)	29 (4%)	36 62
22	2	197/212 (93%)	194 (98%)	3 (2%)	65 85
23	3	1113/1244 (90%)	1082 (97%)	31 (3%)	43 70
24	4	154/189 (82%)	154 (100%)	0	100 100
25	5	93/97 (96%)	90 (97%)	3 (3%)	39 65
26	6	76/76 (100%)	75 (99%)	1 (1%)	69 87
27	L	195/525 (37%)	187 (96%)	8 (4%)	30 55
28	q	65/451 (14%)	62 (95%)	3 (5%)	27 50
28	r	59/451 (13%)	56 (95%)	3 (5%)	24 45
28	s	61/451 (14%)	58 (95%)	3 (5%)	25 47
28	t	62/451 (14%)	58 (94%)	4 (6%)	17 33
29	K	37/165 (22%)	27 (73%)	10 (27%)	0 0
30	N	141/141 (100%)	137 (97%)	4 (3%)	43 70
31	T	295/295 (100%)	288 (98%)	7 (2%)	49 74

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
32	P	218/328 (66%)	215 (99%)	3 (1%)	67 86
33	Q	265/332 (80%)	261 (98%)	4 (2%)	65 85
34	R	224/224 (100%)	218 (97%)	6 (3%)	44 71
35	S	58/151 (38%)	57 (98%)	1 (2%)	60 82
36	Y	156/186 (84%)	156 (100%)	0	100 100
37	X	114/114 (100%)	108 (95%)	6 (5%)	22 43
38	Z	88/107 (82%)	87 (99%)	1 (1%)	73 89
39	W	20/413 (5%)	20 (100%)	0	100 100
40	U	21/121 (17%)	21 (100%)	0	100 100
41	V	416/538 (77%)	410 (99%)	6 (1%)	67 86
42	M	145/237 (61%)	142 (98%)	3 (2%)	53 78
43	z	134/273 (49%)	130 (97%)	4 (3%)	41 68
44	y	63/167 (38%)	60 (95%)	3 (5%)	25 48
45	J	215/633 (34%)	215 (100%)	0	100 100
46	I	58/786 (7%)	53 (91%)	5 (9%)	10 20
47	x	586/789 (74%)	573 (98%)	13 (2%)	52 77
All	All	11869/18071 (66%)	11629 (98%)	240 (2%)	57 79

5 of 240 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
23	3	296	LEU
44	y	109	LEU
27	L	44	THR
44	y	90	GLU
47	x	690	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 170 such sidechains are listed below:

Mol	Chain	Res	Type
23	3	1236	ASN
35	S	34	HIS
24	4	191	ASN
32	P	246	GLN
41	V	116	ASN

5.3.3 RNA

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
12	F	102/112 (91%)	18 (17%)	1 (0%)
13	G	99/162 (61%)	45 (45%)	4 (4%)
14	H	161/1175 (13%)	40 (24%)	1 (0%)
2	B	177/214 (82%)	58 (32%)	6 (3%)
All	All	539/1663 (32%)	161 (29%)	12 (2%)

5 of 161 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	9	U
2	B	12	C
2	B	13	A
2	B	14	G
2	B	18	A

5 of 12 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
13	G	109	A
13	G	472	A
14	H	36	A
13	G	480	A
2	B	130	A

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

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 24 ligands modelled in this entry, 22 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The

Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
49	GTP	C	1101	50	26,34,34	1.29	2 (7%)	32,54,54	1.52	6 (18%)
48	IHP	A	3000	-	36,36,36	0.77	0	54,60,60	0.93	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
49	GTP	C	1101	50	-	6/18/38/38	0/3/3/3
48	IHP	A	3000	-	-	6/30/54/54	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
49	C	1101	GTP	C5-C6	-4.34	1.38	1.47
49	C	1101	GTP	C5-C4	-2.05	1.37	1.43

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
49	C	1101	GTP	PB-O3B-PG	-4.06	118.88	132.83
49	C	1101	GTP	C5-C6-N1	2.99	119.22	113.95
49	C	1101	GTP	C8-N7-C5	2.92	108.56	102.99
49	C	1101	GTP	C2-N1-C6	-2.85	119.86	125.10
49	C	1101	GTP	PA-O3A-PB	-2.18	125.33	132.83

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
48	A	3000	IHP	C2-O12-P2-O22
48	A	3000	IHP	C4-O14-P4-O24
48	A	3000	IHP	C6-O16-P6-O26
49	C	1101	GTP	C5'-O5'-PA-O3A
49	C	1101	GTP	C5'-O5'-PA-O2A

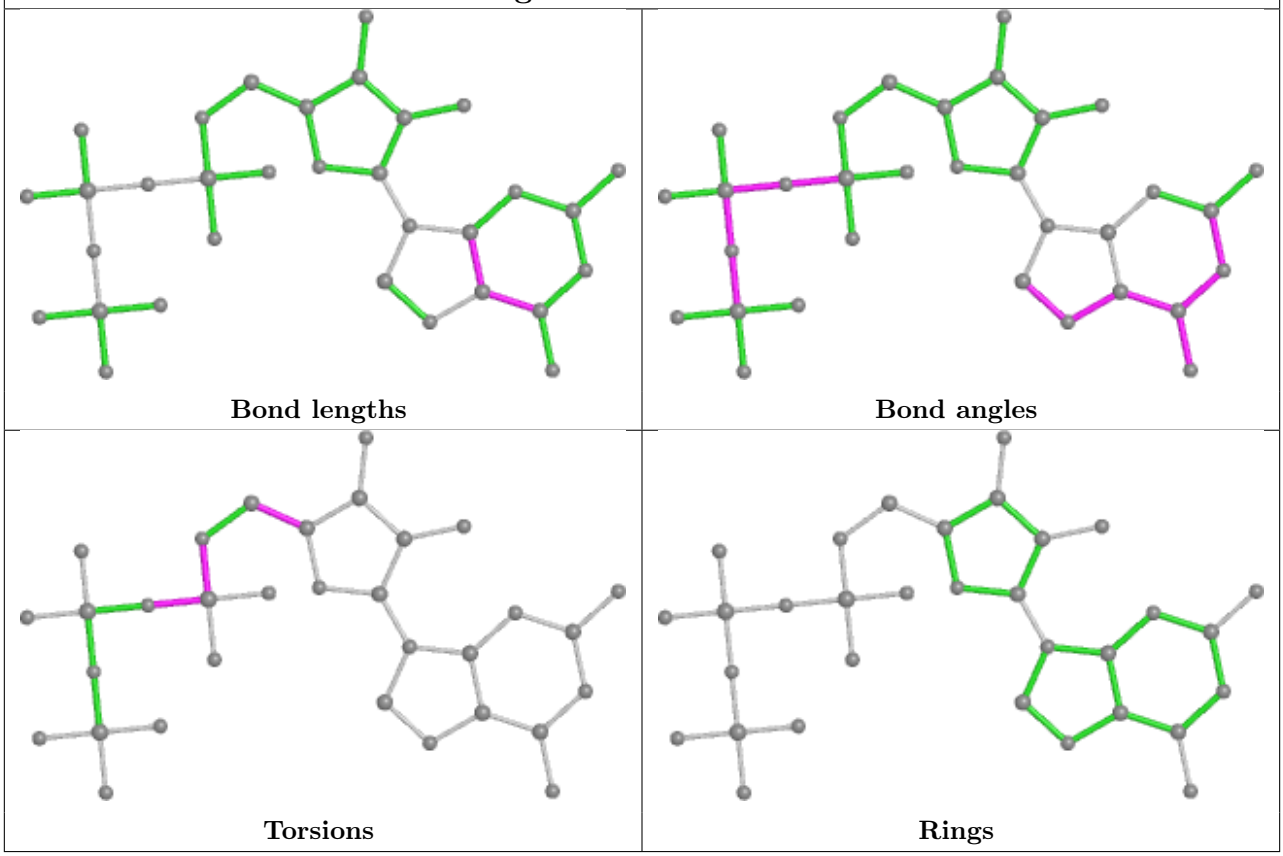
There are no ring outliers.

2 monomers are involved in 4 short contacts:

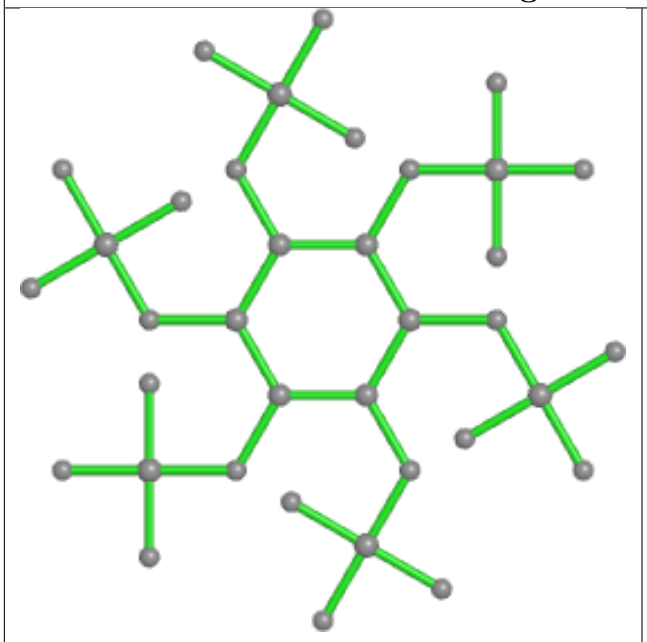
Mol	Chain	Res	Type	Clashes	Symm-Clashes
49	C	1101	GTP	2	0
48	A	3000	IHP	2	0

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

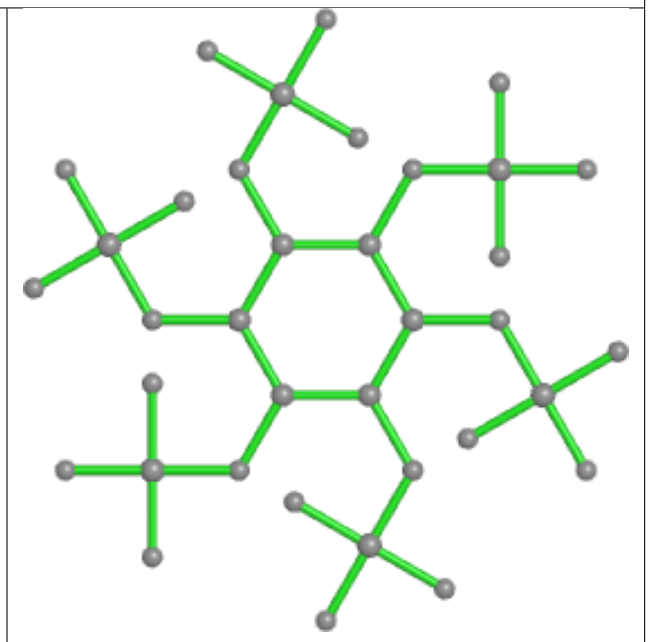
Ligand GTP C 1101



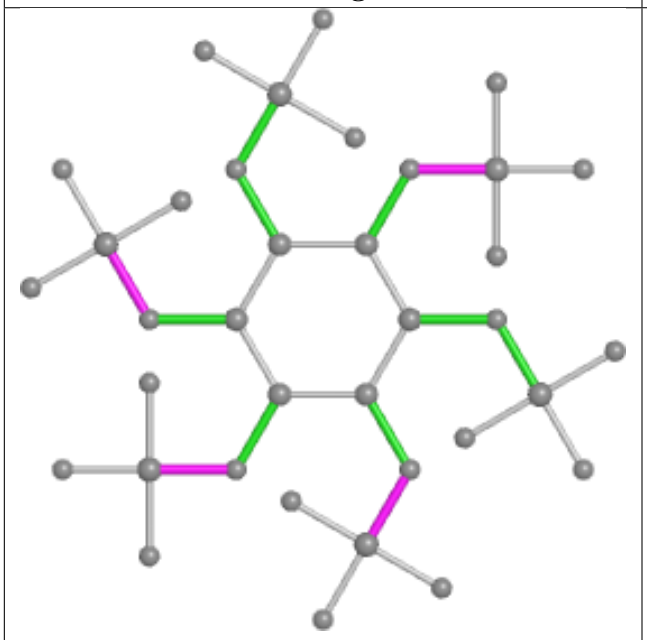
Ligand IHP A 3000



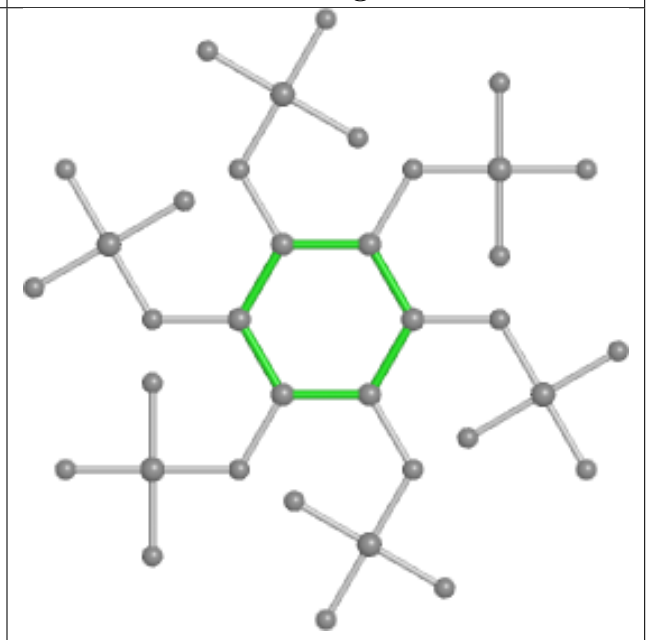
Bond lengths



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.

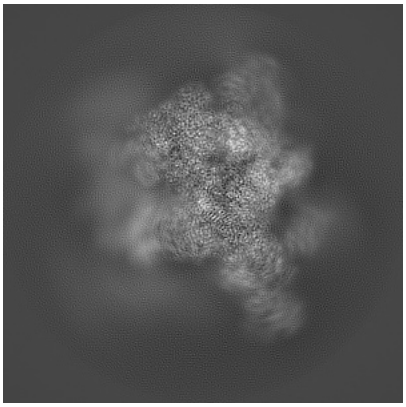
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-30637. These allow visual inspection of the internal detail of the map and identification of artifacts.

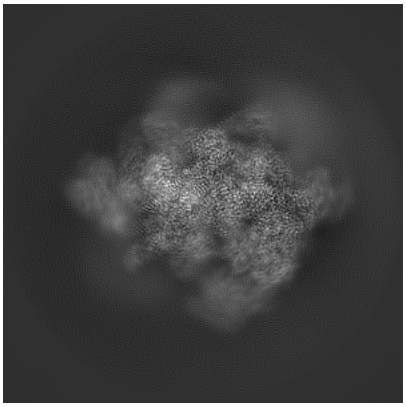
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

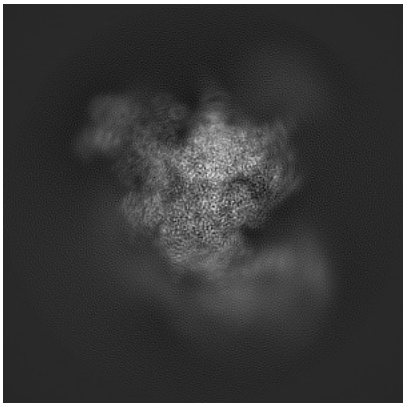
6.1.1 Primary map



X



Y

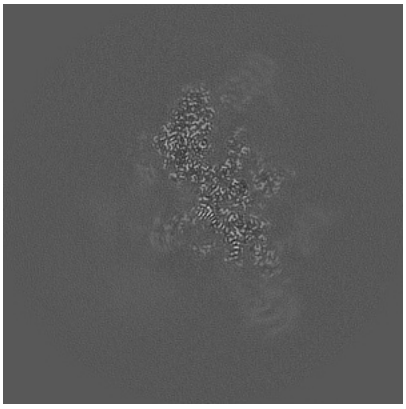


Z

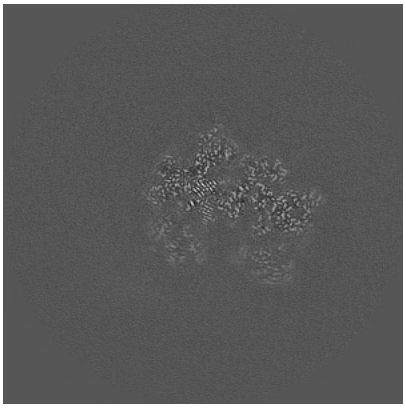
The images above show the map projected in three orthogonal directions.

6.2 Central slices

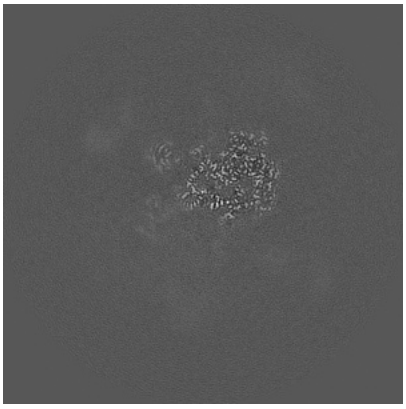
6.2.1 Primary map



X Index: 200



Y Index: 200

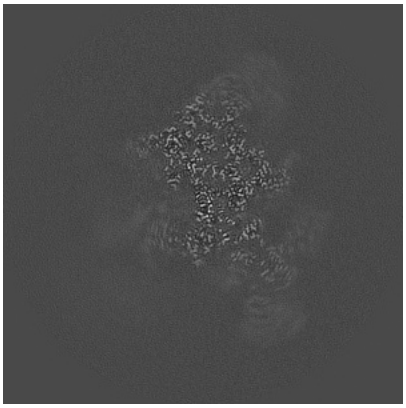


Z Index: 200

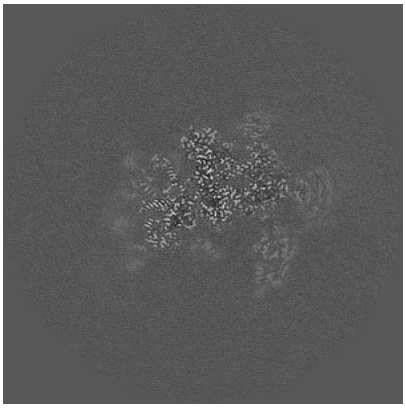
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

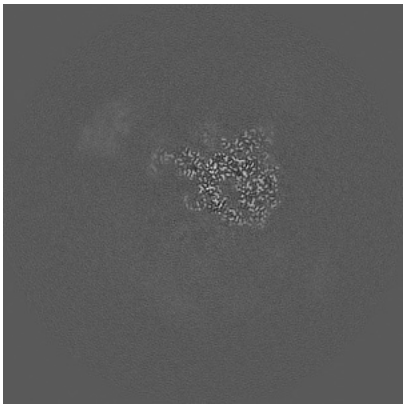
6.3.1 Primary map



X Index: 211



Y Index: 233

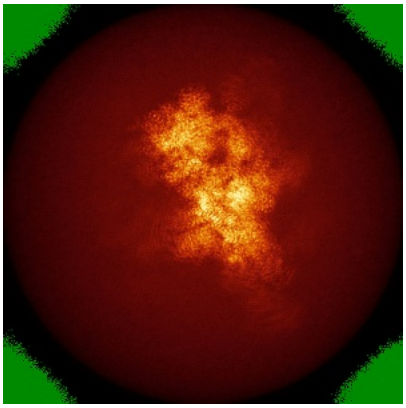


Z Index: 209

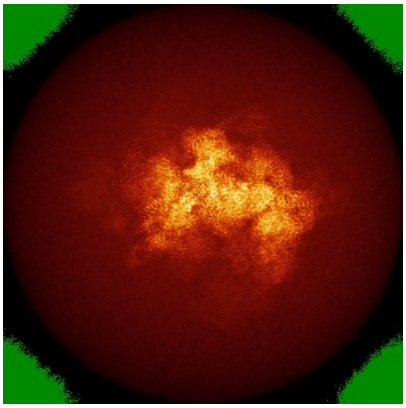
The images above show the largest variance slices of the map in three orthogonal directions.

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

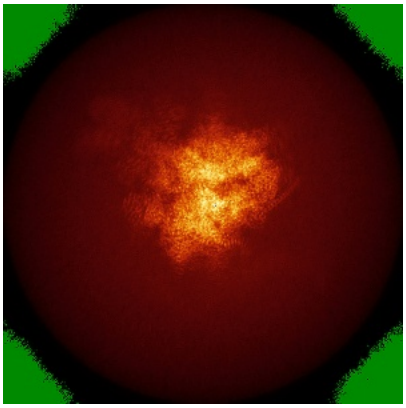
6.4.1 Primary map



X



Y



Z

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

6.5 Orthogonal surface views (i)

6.5.1 Primary map

X Y Z

The images above show the 3D surface view of the map at the recommended contour level 0.015. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

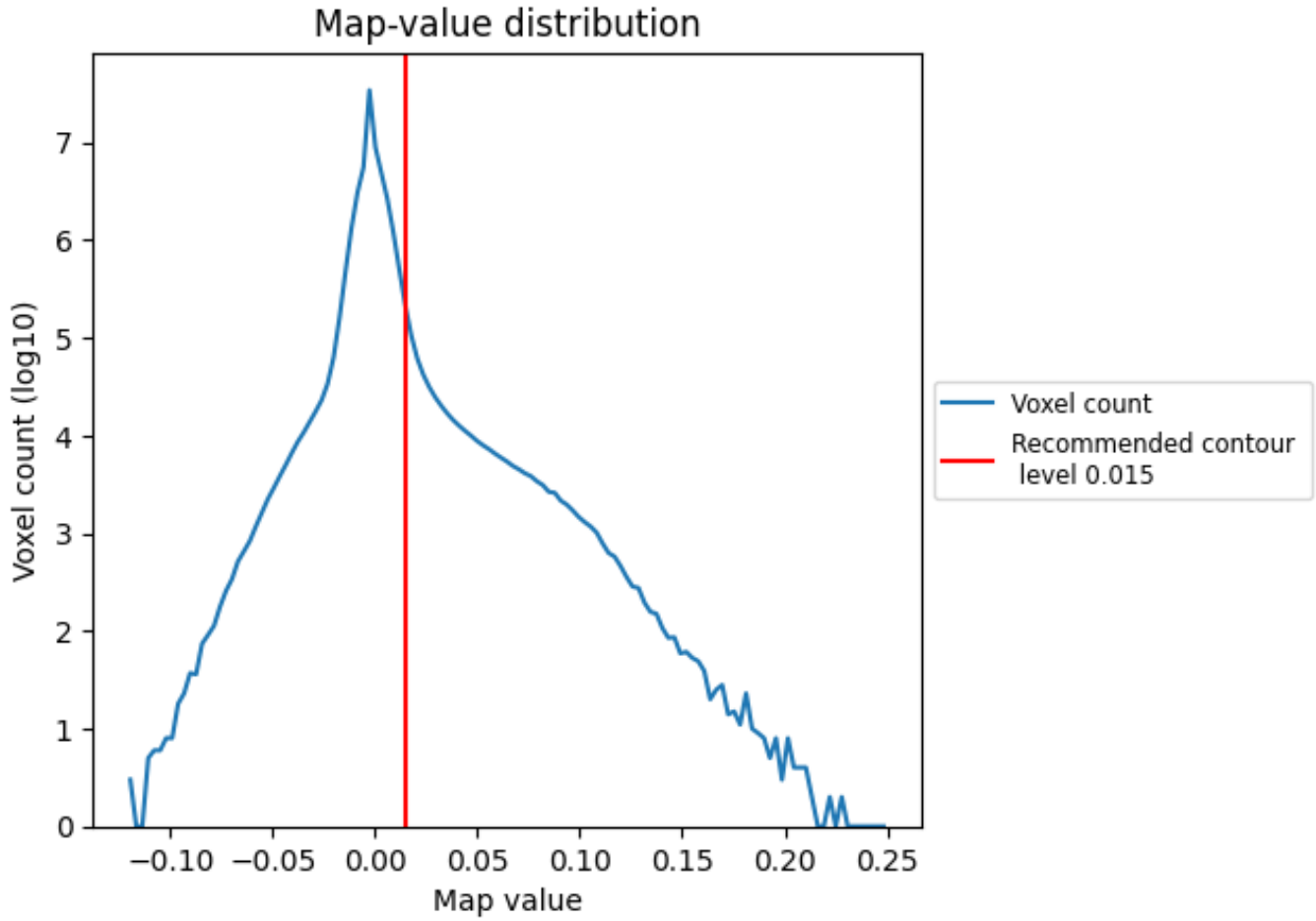
6.6 Mask visualisation (i)

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

7 Map analysis

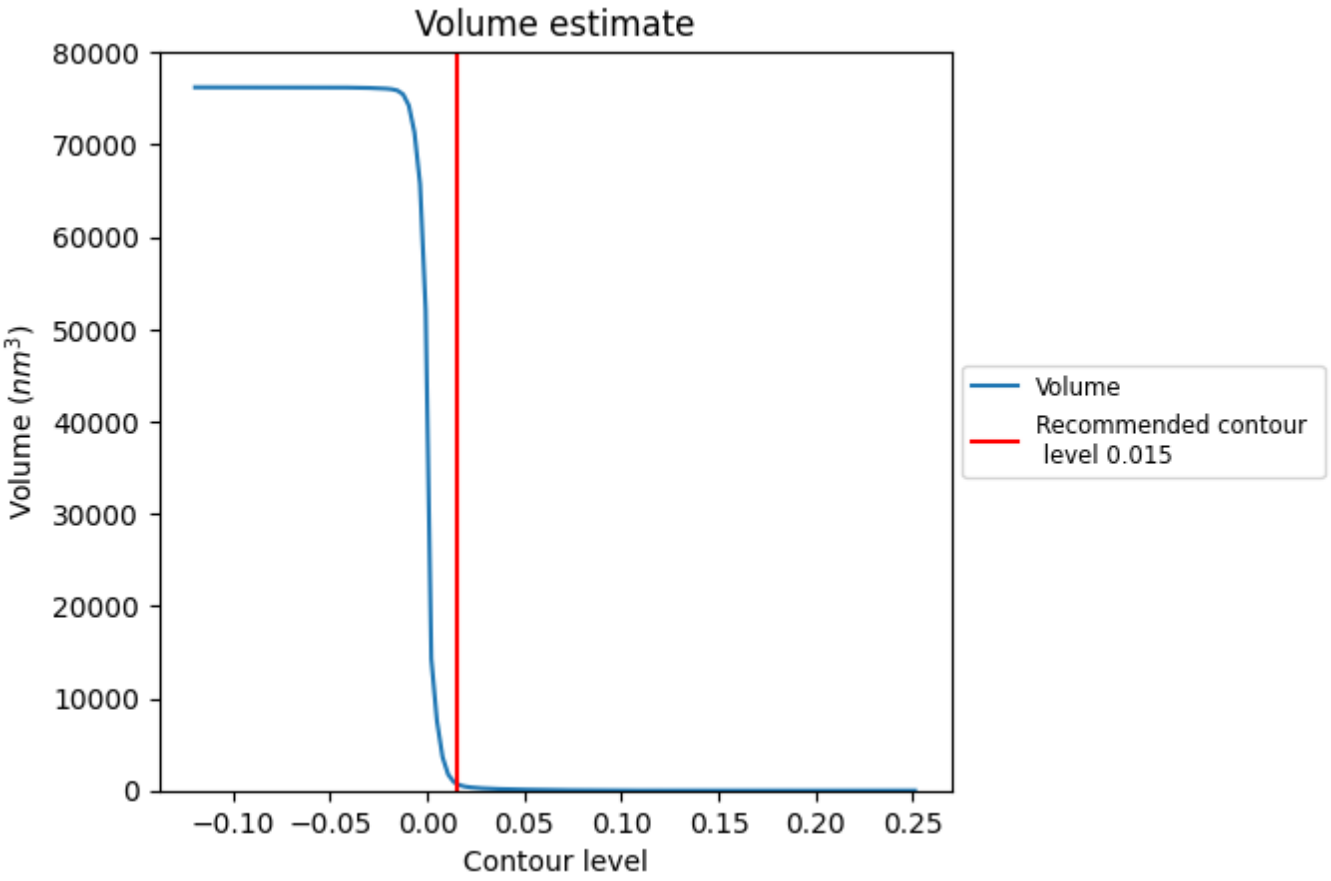
This section contains the results of statistical analysis of the map.

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.

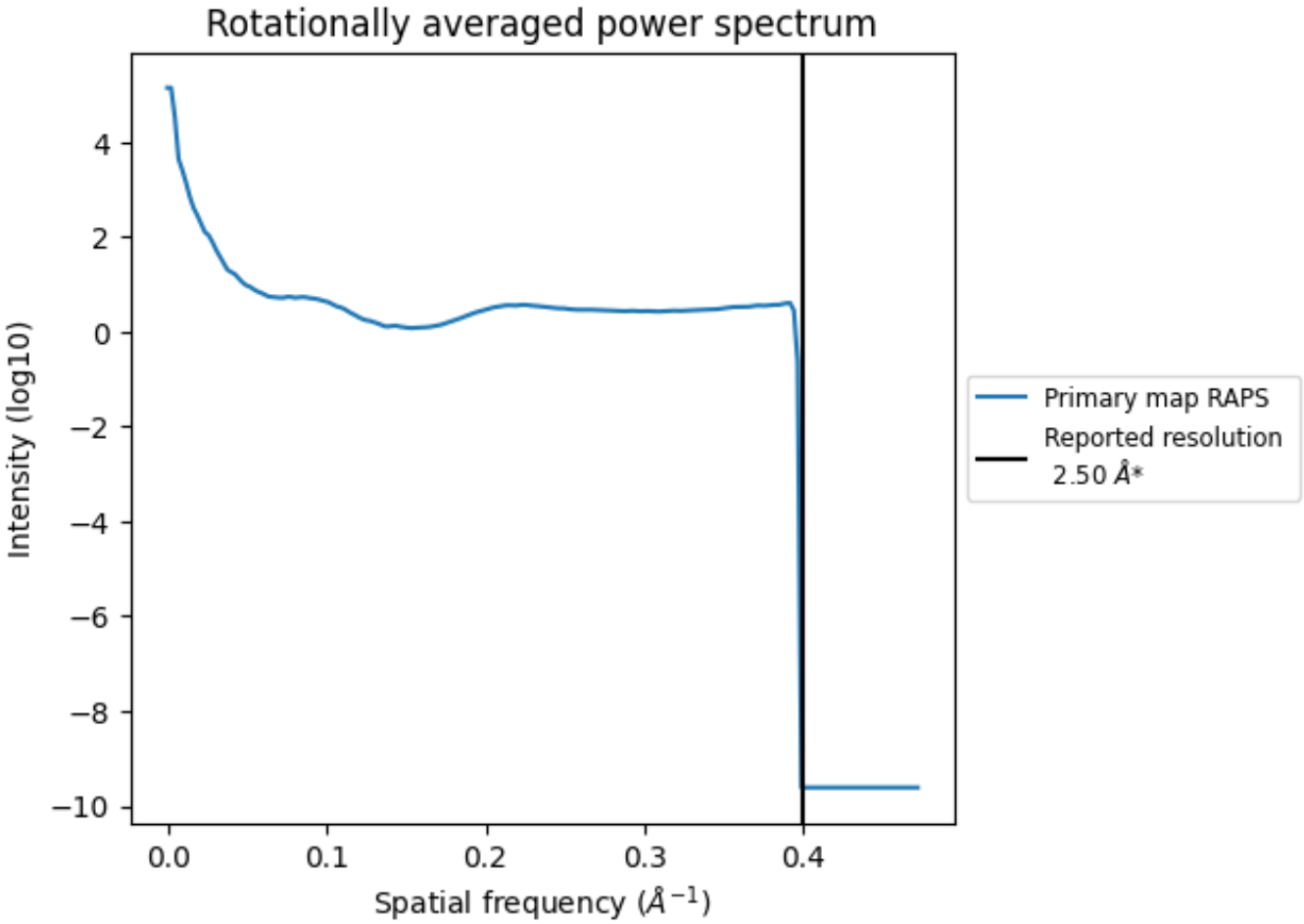
7.2 Volume estimate (i)



The volume at the recommended contour level is 787 nm^3; this corresponds to an approximate mass of 711 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 0.400 Å^{-1}

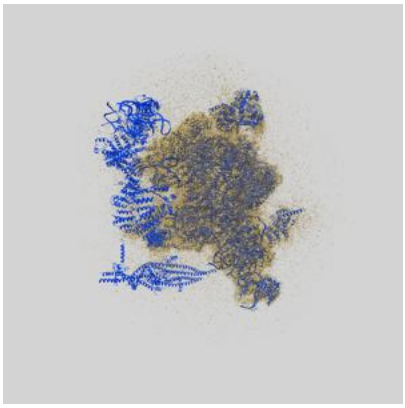
8 Fourier-Shell correlation (i)

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

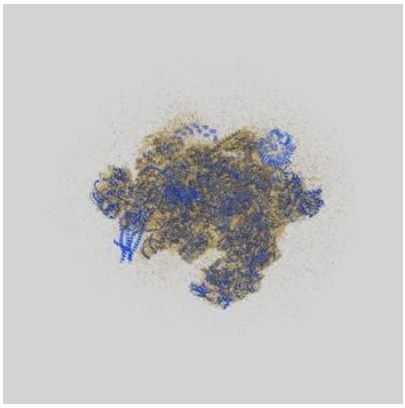
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-30637 and PDB model 7DCO. Per-residue inclusion information can be found in section 3 on page 16.

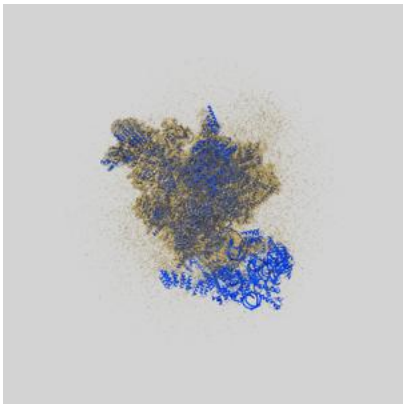
9.1 Map-model overlay



X



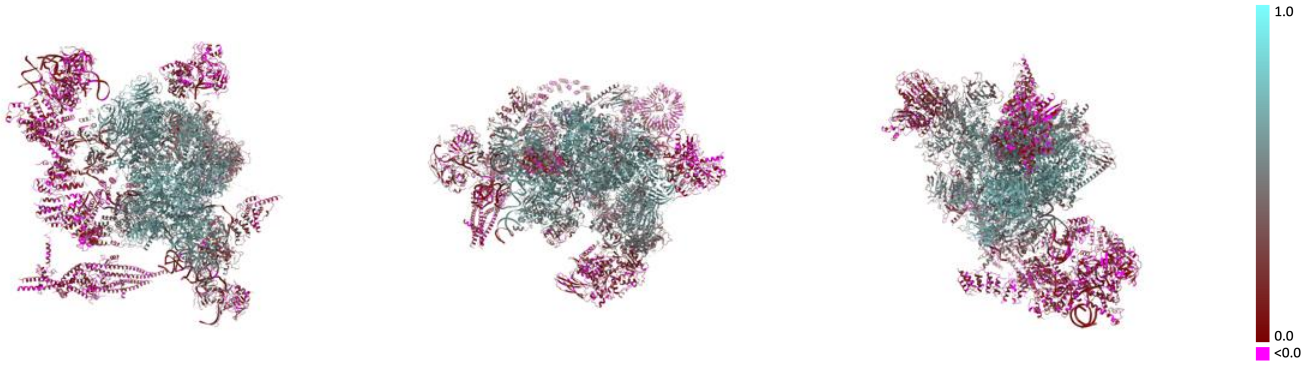
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.015 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

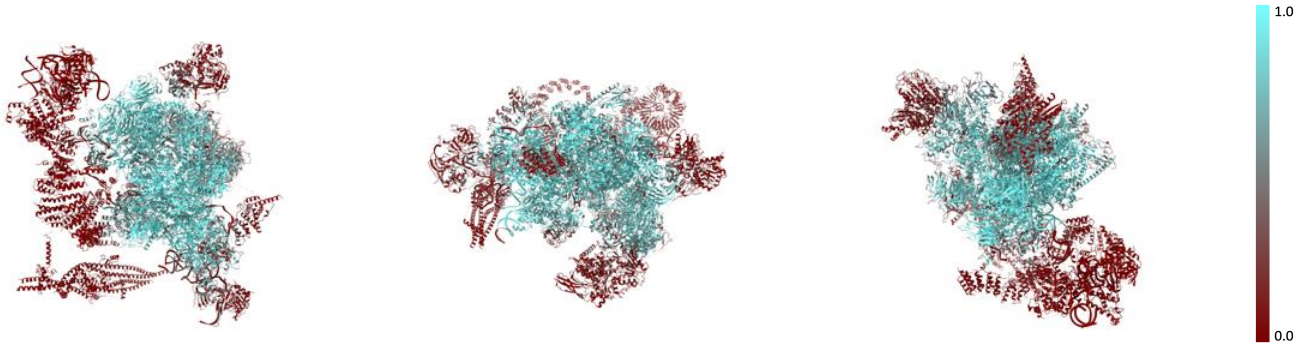
9.2 Q-score mapped to coordinate model



X Y Z

The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions at which atoms can be resolved. Low Q-score values may therefore be expected for many entries.

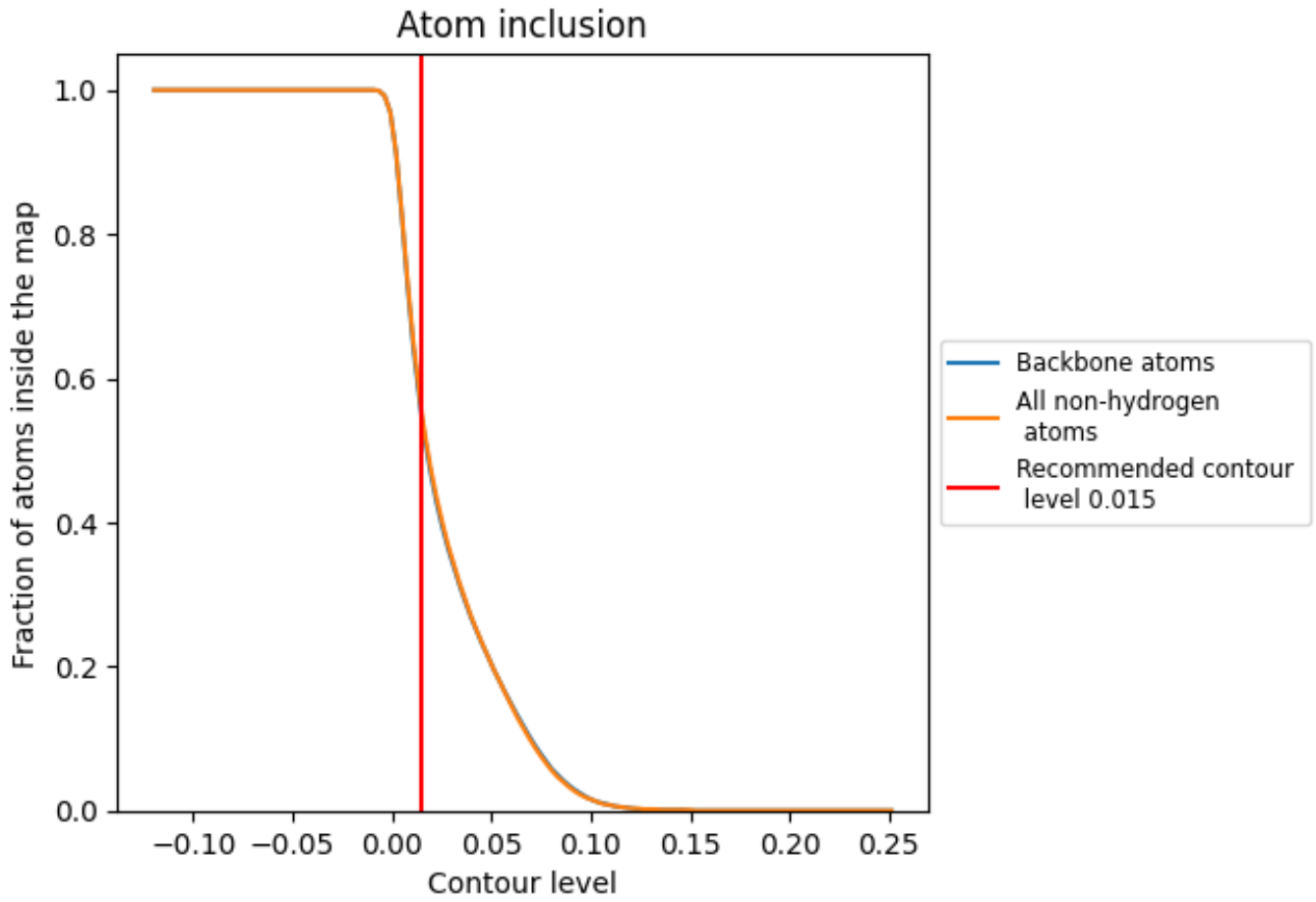
9.3 Atom inclusion mapped to coordinate model



X Y Z

The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.015).

9.4 Atom inclusion (i)



At the recommended contour level, 55% of all backbone atoms, 55% 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 (0.015) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.5530	0.4080
1	0.9220	0.6380
2	0.7810	0.5440
3	0.8940	0.6140
4	0.3420	0.2650
5	0.9540	0.6630
6	0.9480	0.6510
A	0.8840	0.6170
B	0.5340	0.3560
C	0.7090	0.4860
D	0.4180	0.3250
F	0.8190	0.5270
G	0.5350	0.3840
H	0.2790	0.2360
I	0.0170	0.1150
J	0.2970	0.2660
K	0.0020	0.0860
L	0.3340	0.3050
M	0.6130	0.4670
N	0.8920	0.5960
P	0.7290	0.5100
Q	0.5000	0.3620
R	0.7500	0.5030
S	0.5910	0.4930
T	0.9080	0.6170
U	0.8660	0.6060
V	0.4750	0.3730
W	0.7890	0.4970
X	0.8940	0.6040
Y	0.6090	0.4020
Z	0.7450	0.5220
a	0.3660	0.2980
b	0.2730	0.2640
c	0.1780	0.2200
d	0.3990	0.2990

Continued on next page...

Continued from previous page...

Chain	Atom inclusion	Q-score
e	0.1290	0.2060
f	0.1210	0.1970
g	0.1110	0.1900
h	0.0020	0.1030
i	0.0040	0.0620
j	0.0020	0.1110
k	0.0020	0.0870
l	0.0050	0.1030
m	0.0020	0.0930
n	0.0040	0.0790
o	0.0040	0.1290
p	0.0110	0.1330
q	0.0040	0.1050
r	0.0020	0.1400
s	0.0050	0.1100
t	0.0020	0.1190
u	0.1370	0.1700
v	0.4320	0.3700
w	0.0030	0.0800
x	0.1680	0.0830
y	0.2280	0.1810
z	0.6640	0.4540