Full wwPDB EM Validation Report ⓘ

Jun 3, 2024 – 09:20 PM JST

PDB ID : 8H6K
EMDB ID : EMD-34507
Title : Cryo-EM structure of human exon-defined spliceosome in the mature B state.
Authors : Zhang, W.; Zhan, X.; Zhang, X.; Bai, R.; Lei, J.; Yan, C.; Shi, Y.
Deposited on : 2022-10-18
Resolution : 2.70 Å(reported)

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

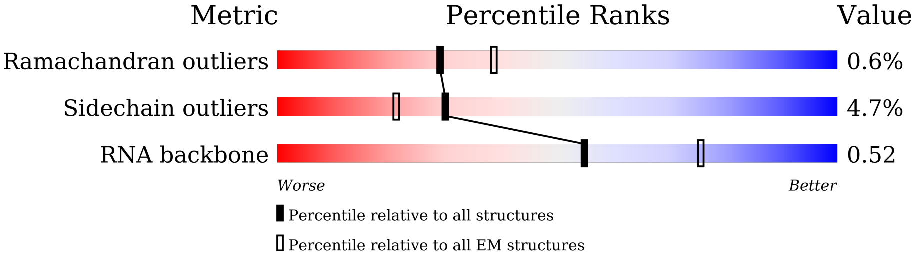
EMDB validation analysis : FAILED
Mogul : 1.8.5 (274361), CSD as541be (2020)
MolProbity : 4.02b-467
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ : FAILED
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36.2

1 Overall quality at a glance

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

The reported resolution of this entry is 2.70 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%

Mol	Chain	Length	Quality of chain
1	A	144	
2	5A	117	
3	5B	2335	
4	5C	972	
5	5D	2136	
6	5E	357	
7	2a	231	
7	4a	231	
7	5a	231	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	2b	119	69% 31%
8	4b	119	69% 31%
8	5b	119	69% 31%
9	2c	118	71% • 28%
9	4c	118	63% 37%
9	5c	118	81% • 18%
10	2d	86	86% 14%
10	4d	86	83% 17%
10	5d	86	86% 14%
11	2e	92	86% 14%
11	4e	92	85% 15%
11	5e	92	86% 14%
12	2f	76	89% 11%
12	4f	76	96% •
12	5f	76	95% 5%
13	2g	126	63% 37%
13	4g	126	56% 44%
13	5g	126	61% 39%
14	6A	107	47% 9% 44%
15	6a	95	91% • 5%
16	6b	102	70% • 27%
17	6c	139	53% • 47%
18	6d	91	78% • 21%
19	6e	80	85% • 12%
20	6f	103	63% 37%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
21	6g	96	63% • 36%
22	4A	145	64% 25% 11%
23	4B	683	36% • 63%
24	4C	522	76% 5% 18%
25	4D	499	71% • 25%
26	4E	128	93% • •
27	4F	142	94% 6% •
28	4G	941	82% • 15%
29	4H	177	95% 5%
30	4I	376	19% • 80%
31	4J	800	18% • 81%
32	4K	439	41% • 57%
33	4L	312	51% 5% 44%
34	4M	73	92% 8%
35	4N	199	35% 6% 60%
36	4Z	513	82% 18%
37	2A	188	33% 21% • 42%
38	2B	255	63% • 36%
39	2C	225	42% 58%
40	2D	793	25% • 74%
41	2E	464	18% • 80%
42	2F	501	83% • 16%
43	2G	1304	78% • 20%
44	2H	895	23% • 76%
45	2I	1217	94% • •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
46	2J	424	
47	2K	125	
48	2L	110	
49	2M	86	

2 Entry composition ⓘ

There are 53 unique types of molecules in this entry. The entry contains 98459 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a RNA chain called pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	A	64	1334	597	209	464	64	0	0

- Molecule 2 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	5A	115	2420	1084	403	818	115	0	0

- Molecule 3 is a protein called Pre-mRNA-processing-splicing factor 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	5B	2253	18639	11991	3249	3318	81	0	0

- Molecule 4 is a protein called 116 kDa U5 small nuclear ribonucleoprotein component.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	5C	818	6430	4108	1085	1205	32	0	0

- Molecule 5 is a protein called U5 small nuclear ribonucleoprotein 200 kDa helicase.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	5D	1696	13633	8715	2329	2519	70	0	0

- Molecule 6 is a protein called U5 small nuclear ribonucleoprotein 40 kDa protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
6	5E	299	1196	598	299	299	0	0

- Molecule 7 is a protein called Isoform SM-B of Small nuclear ribonucleoprotein-associated proteins B and B'.

Mol	Chain	Residues	Atoms				AltConf	Trace
7	5a	84	Total	C	N	O	0	0
			336	168	84	84		
7	4a	64	Total	C	N	O	0	0
			256	128	64	64		
7	2a	86	Total	C	N	O	0	0
			344	172	86	86		

- Molecule 8 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms				AltConf	Trace
8	5b	82	Total	C	N	O	0	0
			328	164	82	82		
8	4b	82	Total	C	N	O	0	0
			334	170	82	82		
8	2b	82	Total	C	N	O	0	0
			328	164	82	82		

- Molecule 9 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms				AltConf	Trace
9	5c	97	Total	C	N	O	0	0
			388	194	97	97		
9	4c	74	Total	C	N	O	0	0
			300	152	74	74		
9	2c	85	Total	C	N	O	0	0
			340	170	85	85		

- Molecule 10 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms				AltConf	Trace
10	5d	74	Total	C	N	O	0	0
			296	148	74	74		
10	4d	71	Total	C	N	O	0	0
			292	150	71	71		
10	2d	74	Total	C	N	O	0	0
			296	148	74	74		

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

Mol	Chain	Residues	Atoms				AltConf	Trace
11	5e	79	Total 316	C 158	N 79	O 79	0	0
11	4e	78	Total 314	C 158	N 78	O 78	0	0
11	2e	79	Total 316	C 158	N 79	O 79	0	0

- Molecule 12 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms				AltConf	Trace
12	5f	72	Total 288	C 144	N 72	O 72	0	0
12	4f	73	Total 298	C 152	N 73	O 73	0	0
12	2f	68	Total 272	C 136	N 68	O 68	0	0

- Molecule 13 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms				AltConf	Trace
13	5g	77	Total 308	C 154	N 77	O 77	0	0
13	4g	71	Total 288	C 146	N 71	O 71	0	0
13	2g	80	Total 320	C 160	N 80	O 80	0	0

- Molecule 14 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	6A	60	Total 1273	C 568	N 235	O 410	P 60	0	0

- Molecule 15 is a protein called U6 snRNA-associated Sm-like protein LSm2.

Mol	Chain	Residues	Atoms				AltConf	Trace
15	6a	90	Total 360	C 180	N 90	O 90	0	0

- Molecule 16 is a protein called U6 snRNA-associated Sm-like protein LSm3.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
16	6b	74	296	148	74	74	0	0

- Molecule 17 is a protein called U6 snRNA-associated Sm-like protein LSm4.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
17	6c	74	296	148	74	74	0	0

- Molecule 18 is a protein called U6 snRNA-associated Sm-like protein LSm5.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
18	6d	72	288	144	72	72	0	0

- Molecule 19 is a protein called U6 snRNA-associated Sm-like protein LSm6.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
19	6e	70	280	140	70	70	0	0

- Molecule 20 is a protein called U6 snRNA-associated Sm-like protein LSm7.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
20	6f	65	260	130	65	65	0	0

- Molecule 21 is a protein called U6 snRNA-associated Sm-like protein LSm8.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
21	6g	61	244	122	61	61	0	0

- Molecule 22 is a RNA chain called U4 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
22	4A	129	2744	1225	472	917	130	0	0

- Molecule 23 is a protein called U4/U6 small nuclear ribonucleoprotein Prp3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
23	4B	256	2076	1316	385	367	8	0	0

- Molecule 24 is a protein called U4/U6 small nuclear ribonucleoprotein Prp4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
24	4C	426	3370	2118	612	620	20	0	0

- Molecule 25 is a protein called U4/U6 small nuclear ribonucleoprotein Prp31.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
25	4D	376	2874	1788	524	550	12	0	0

- Molecule 26 is a protein called NHP2-like protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
26	4E	124	962	608	171	178	5	0	0

- Molecule 27 is a protein called Thioredoxin-like protein 4A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
27	4F	141	1169	751	194	214	10	0	0

- Molecule 28 is a protein called Pre-mRNA-processing factor 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
28	4G	801	5504	3419	1043	1026	16	0	0

- Molecule 29 is a protein called Peptidyl-prolyl cis-trans isomerase H.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
29	4H	169	844	506	169	169	0	0

- Molecule 30 is a protein called WW domain-binding protein 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	4I	75	Total 494	C 304	N 96	O 91	S 3	0	0

- Molecule 31 is a protein called U4/U6.U5 tri-snRNP-associated protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	4J	152	Total 1144	C 709	N 204	O 229	S 2	0	0

- Molecule 32 is a protein called Microfibrillar-associated protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
32	4K	188	Total 1192	C 741	N 219	O 230	S 2	0	0

- Molecule 33 is a protein called Pre-mRNA-splicing factor 38A.

Mol	Chain	Residues	Atoms					AltConf	Trace
33	4L	175	Total 1452	C 934	N 244	O 265	S 9	0	0

- Molecule 34 is a protein called Ubiquitin-like protein 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
34	4M	73	Total 599	C 383	N 103	O 109	S 4	0	0

- Molecule 35 is a protein called Zinc finger matrin-type protein 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
35	4N	80	Total 640	C 397	N 116	O 120	S 7	0	0

- Molecule 36 is a protein called WD40 repeat-containing protein SMU1.

Mol	Chain	Residues	Atoms				AltConf	Trace
36	4Z	420	Total 2092	C 1252	N 420	O 420	0	0

- Molecule 37 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
37	2A	109	2311	1032	396	774	109	0	0

- Molecule 38 is a protein called U2 small nuclear ribonucleoprotein A'.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
38	2B	162	648	324	162	162	0	0

- Molecule 39 is a protein called U2 small nuclear ribonucleoprotein B".

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
39	2C	94	376	188	94	94	0	0

- Molecule 40 is a protein called Splicing factor 3A subunit 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
40	2D	204	1134	644	238	250	2	0	0

- Molecule 41 is a protein called Splicing factor 3A subunit 2.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
41	2E	94	376	188	94	94	0	0

- Molecule 42 is a protein called Splicing factor 3A subunit 3.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
42	2F	423	1693	847	423	423	0	0

- Molecule 43 is a protein called Splicing factor 3B subunit 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
43	2G	1048	4192	2096	1048	1048	0	0

- Molecule 44 is a protein called Splicing factor 3B subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
44	2H	213	Total	C	N	O	S	0	0
			959	510	220	226	3		

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

Mol	Chain	Residues	Atoms				AltConf	Trace
45	2I	1168	Total	C	N	O	0	0
			4672	2336	1168	1168		

- Molecule 46 is a protein called Splicing factor 3B subunit 4.

Mol	Chain	Residues	Atoms				AltConf	Trace
46	2J	78	Total	C	N	O	0	0
			312	156	78	78		

- Molecule 47 is a protein called Splicing factor 3B subunit 6.

Mol	Chain	Residues	Atoms				AltConf	Trace
47	2K	108	Total	C	N	O	0	0
			432	216	108	108		

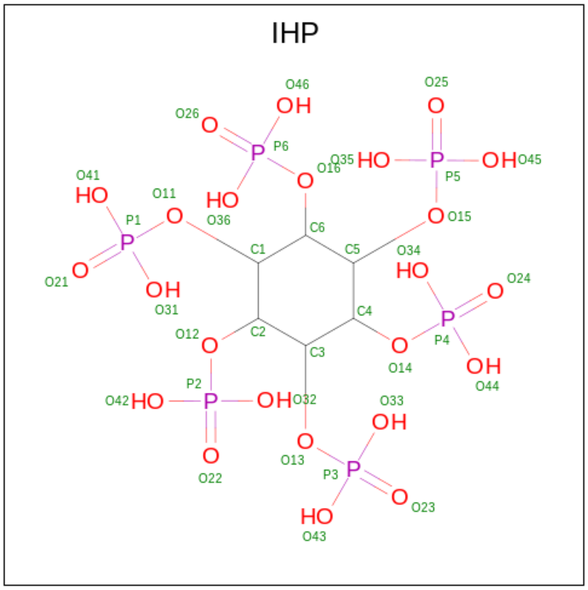
- Molecule 48 is a protein called PHD finger-like domain-containing protein 5A.

Mol	Chain	Residues	Atoms				AltConf	Trace
48	2L	89	Total	C	N	O	0	0
			356	178	89	89		

- Molecule 49 is a protein called Splicing factor 3B subunit 5.

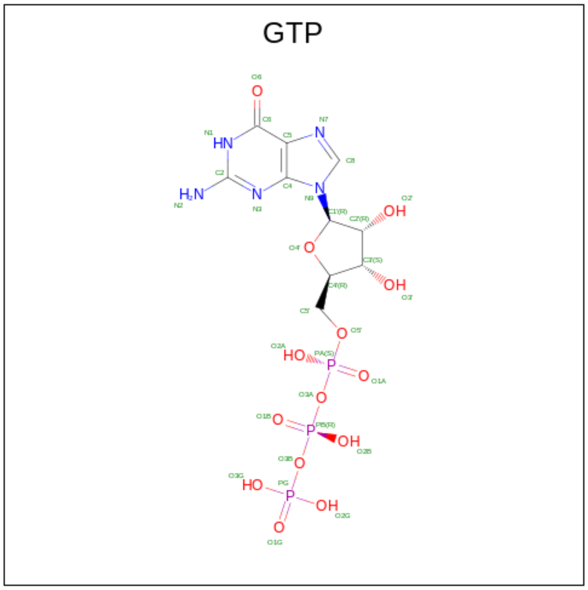
Mol	Chain	Residues	Atoms				AltConf	Trace
49	2M	66	Total	C	N	O	0	0
			264	132	66	66		

- Molecule 50 is INOSITOL HEXAKISPHOSPHATE (three-letter code: IHP) (formula: $C_6H_{18}O_{24}P_6$).



Mol	Chain	Residues	Atoms				AltConf
			Total	C	O	P	
50	5B	1	36	6	24	6	0

- Molecule 51 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).



Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
51	5C	1	32	10	5	14	3	0

- Molecule 52 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	AltConf
52	5C	1	Total Mg 1 1	0

- Molecule 53 is ZINC ION (three-letter code: ZN) (formula: Zn).

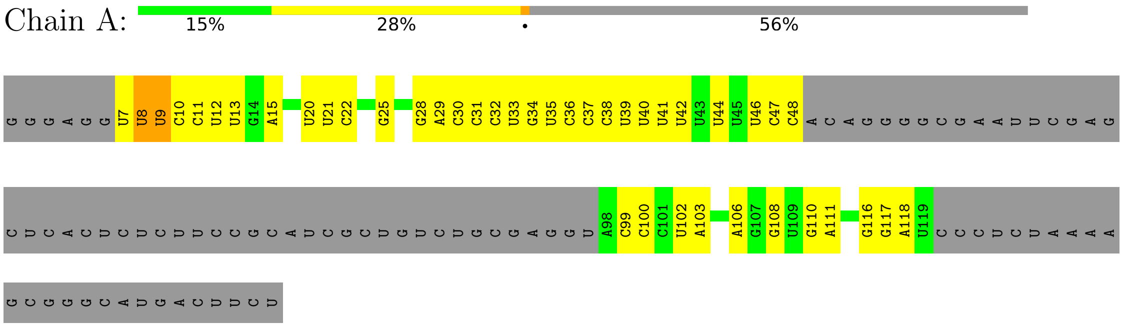
Mol	Chain	Residues	Atoms	AltConf
53	4I	1	Total Zn 1 1	0
53	4N	1	Total Zn 1 1	0

3 Residue-property plots ⓘ

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

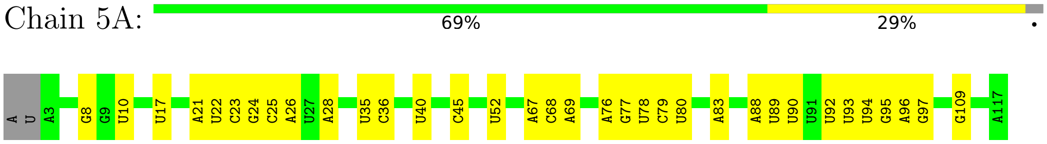
- Molecule 1: pre-mRNA

Chain A:



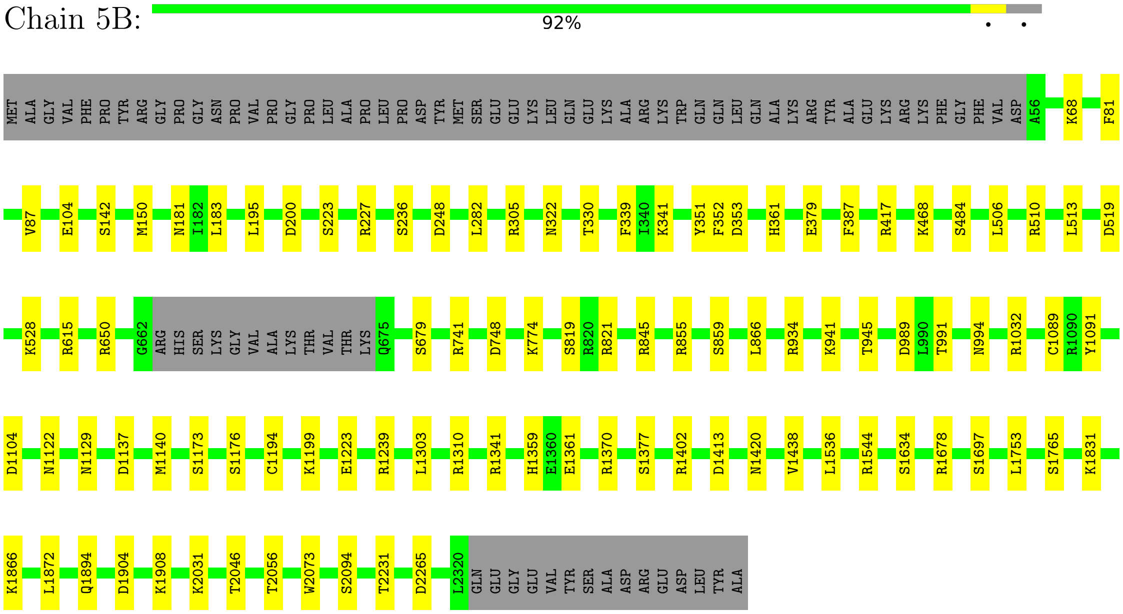
- Molecule 2: U5 snRNA

Chain 5A:



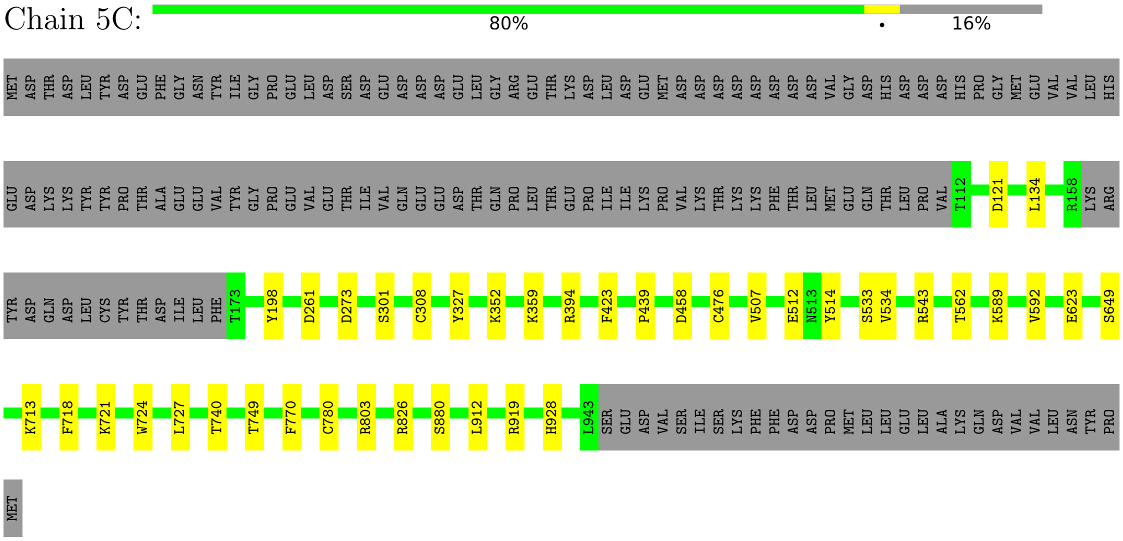
- Molecule 3: Pre-mRNA-processing-splicing factor 8

Chain 5B:



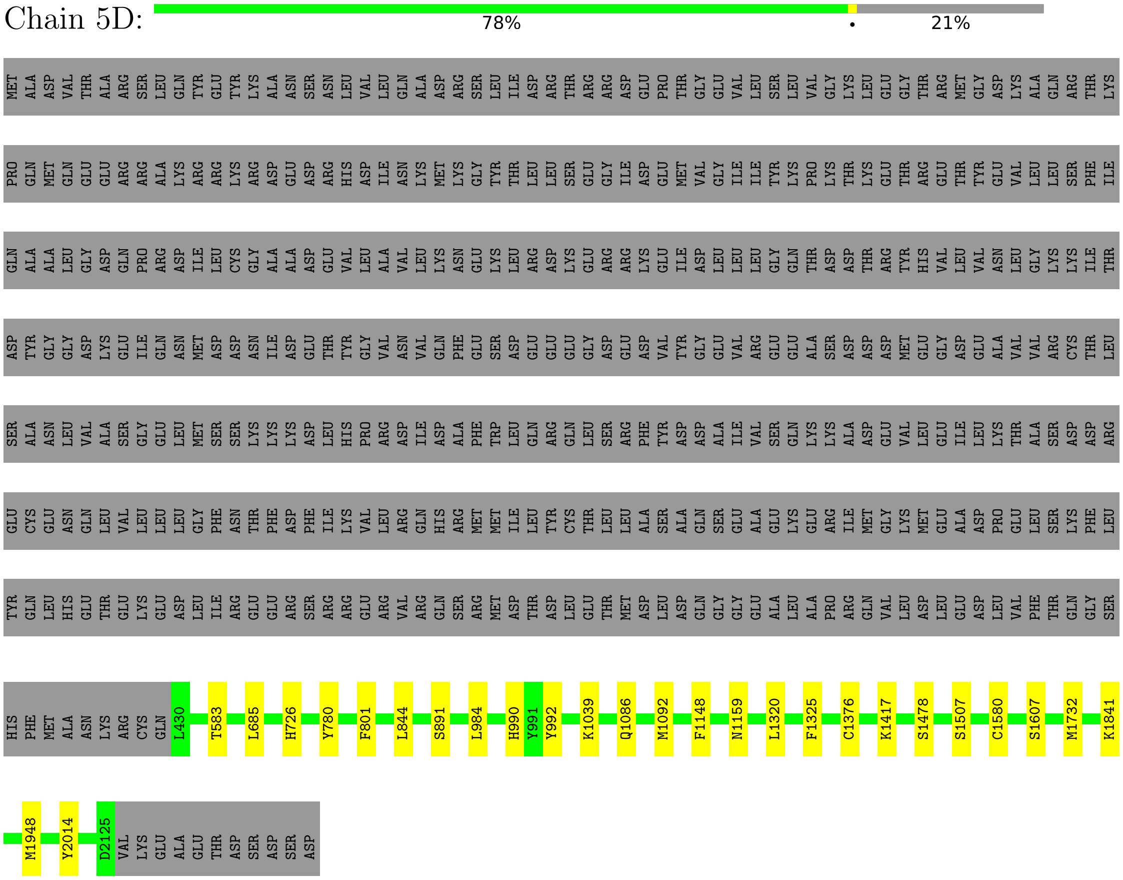
- Molecule 4: 116 kDa U5 small nuclear ribonucleoprotein component

Chain 5C:



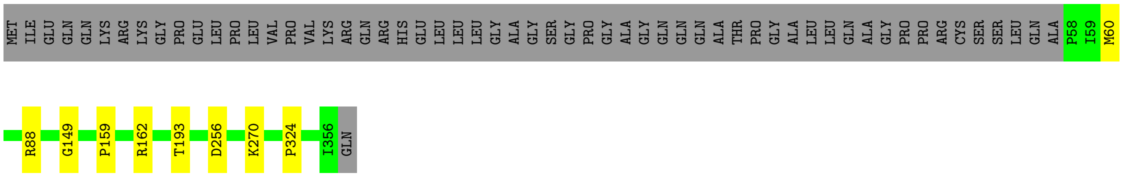
- Molecule 5: U5 small nuclear ribonucleoprotein 200 kDa helicase

Chain 5D:



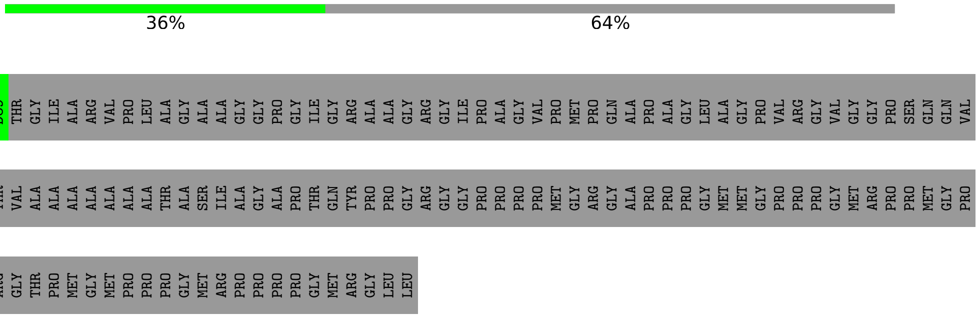
- Molecule 6: U5 small nuclear ribonucleoprotein 40 kDa protein

Chain 5E:



- Molecule 7: Isoform SM-B of Small nuclear ribonucleoprotein-associated proteins B and B’

Chain 5a:

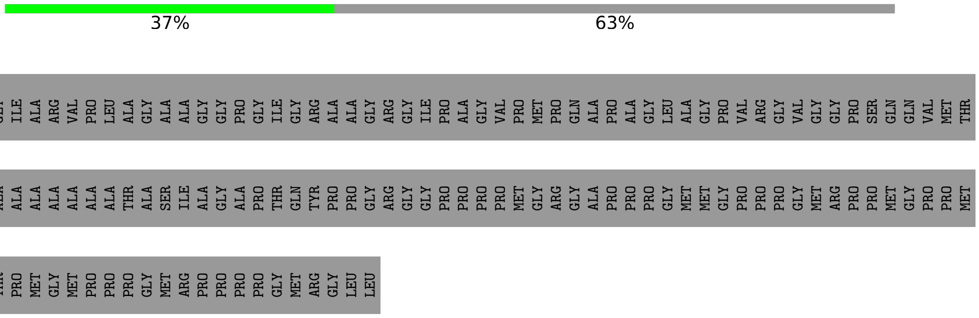


- Molecule 7: Isoform SM-B of Small nuclear ribonucleoprotein-associated proteins B and B’

Chain 4a:

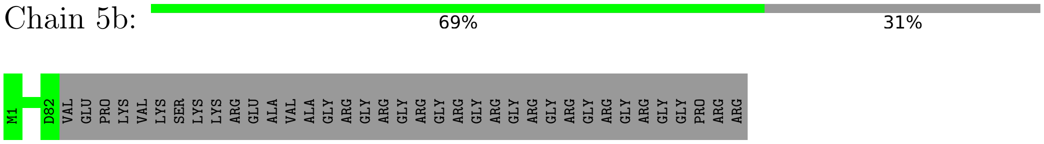
- Molecule 7: Isoform SM-B of Small nuclear ribonucleoprotein-associated proteins B and B’

Chain 2a:



- Molecule 8: Small nuclear ribonucleoprotein Sm D1

Chain 5b:



- Molecule 8: Small nuclear ribonucleoprotein Sm D1

Chain 4b:

- Molecule 8: Small nuclear ribonucleoprotein Sm D1

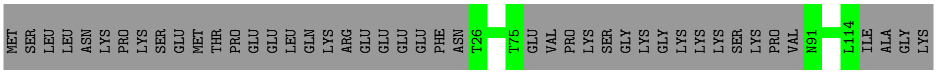
Chain 2b:

- Molecule 9: Small nuclear ribonucleoprotein Sm D2

Chain 5c:

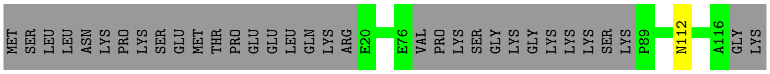
- Molecule 9: Small nuclear ribonucleoprotein Sm D2

Chain 4c:



- Molecule 9: Small nuclear ribonucleoprotein Sm D2

Chain 2c:



- Molecule 10: Small nuclear ribonucleoprotein F

Chain 5d:

- Molecule 10: Small nuclear ribonucleoprotein F

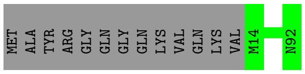
Chain 4d:

- Molecule 10: Small nuclear ribonucleoprotein F

Chain 2d:

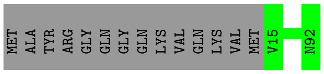
- Molecule 11: Small nuclear ribonucleoprotein E

Chain 5e:



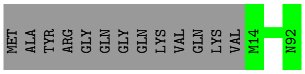
- Molecule 11: Small nuclear ribonucleoprotein E

Chain 4e:



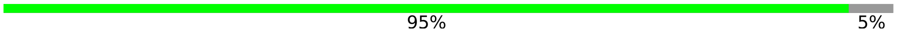
- Molecule 11: Small nuclear ribonucleoprotein E

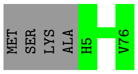
Chain 2e:



- Molecule 12: Small nuclear ribonucleoprotein G

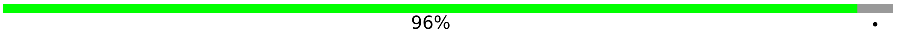
Chain 5f:

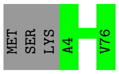




- Molecule 12: Small nuclear ribonucleoprotein G

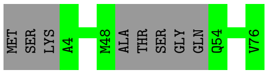
Chain 4f:





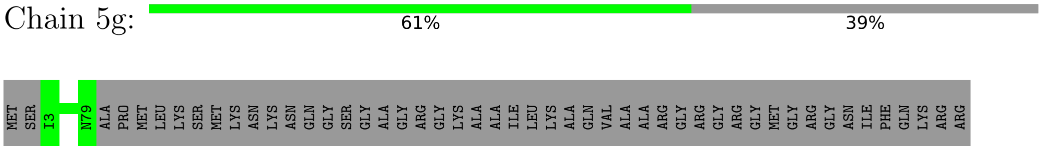
- Molecule 12: Small nuclear ribonucleoprotein G

Chain 2f:



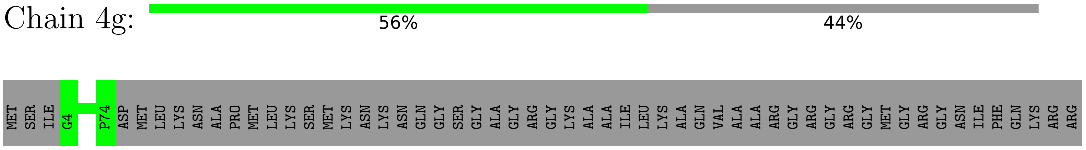
- Molecule 13: Small nuclear ribonucleoprotein Sm D3

Chain 5g:



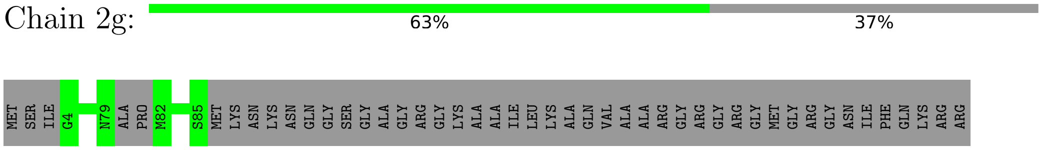
- Molecule 13: Small nuclear ribonucleoprotein Sm D3

Chain 4g:



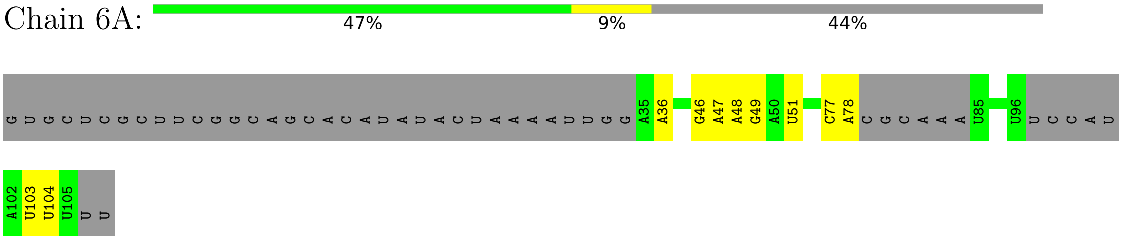
- Molecule 13: Small nuclear ribonucleoprotein Sm D3

Chain 2g:



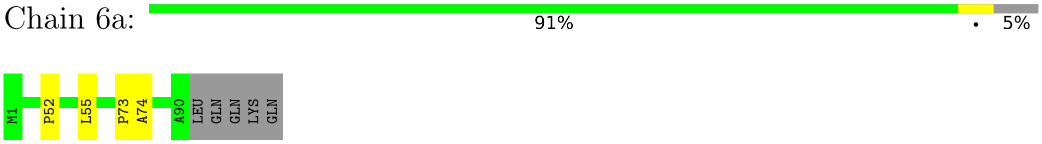
- Molecule 14: U6 snRNA

Chain 6A:



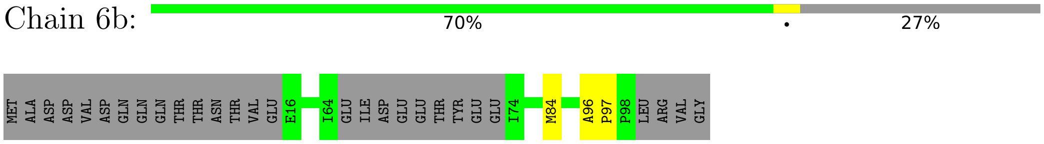
- Molecule 15: U6 snRNA-associated Sm-like protein LSm2

Chain 6a:



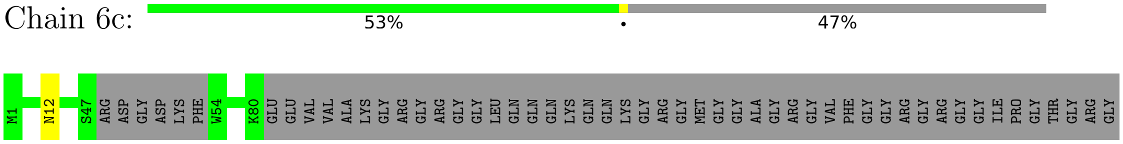
- Molecule 16: U6 snRNA-associated Sm-like protein LSm3

Chain 6b:

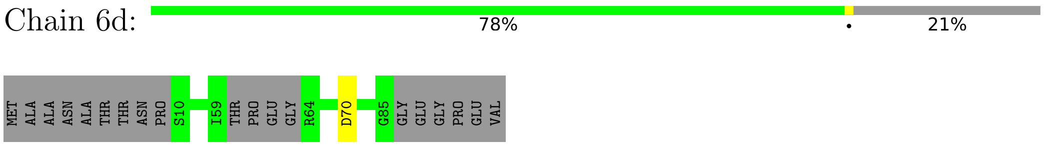


- Molecule 17: U6 snRNA-associated Sm-like protein LSm4

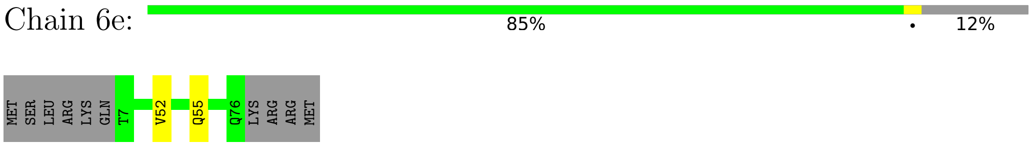
Chain 6c:



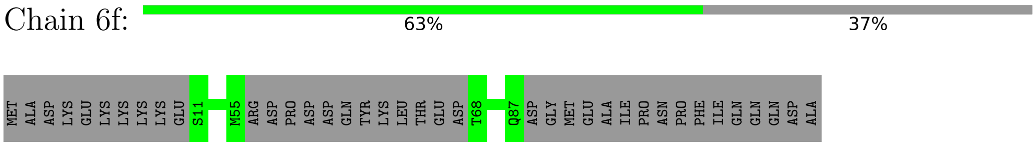
- Molecule 18: U6 snRNA-associated Sm-like protein LSm5



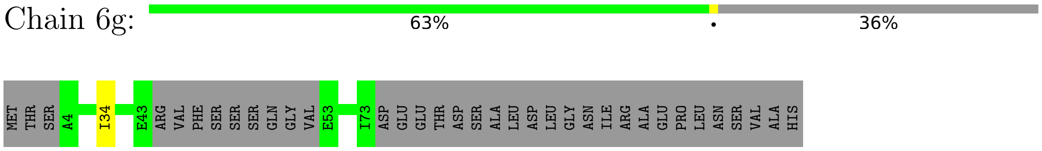
- Molecule 19: U6 snRNA-associated Sm-like protein LSm6



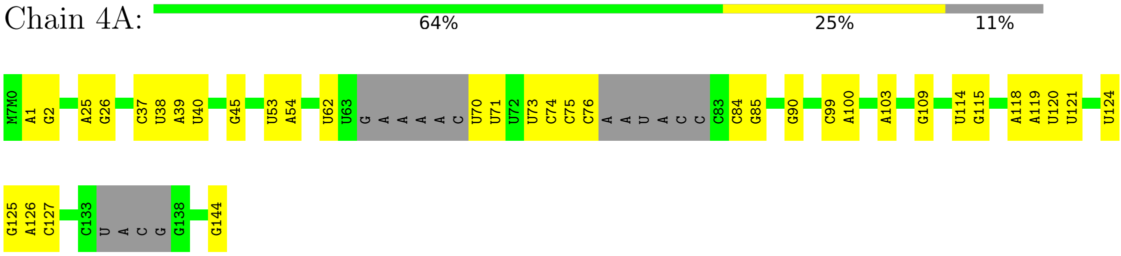
- Molecule 20: U6 snRNA-associated Sm-like protein LSm7



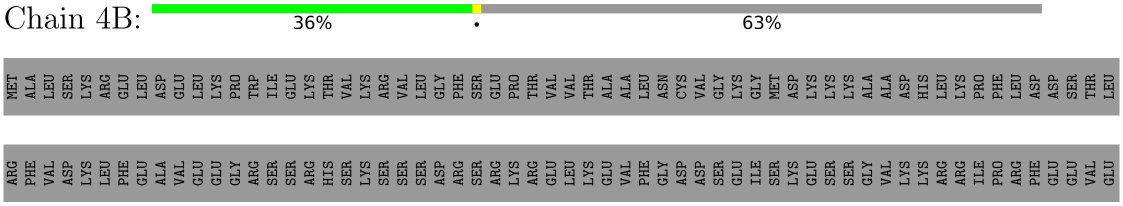
- Molecule 21: U6 snRNA-associated Sm-like protein LSm8

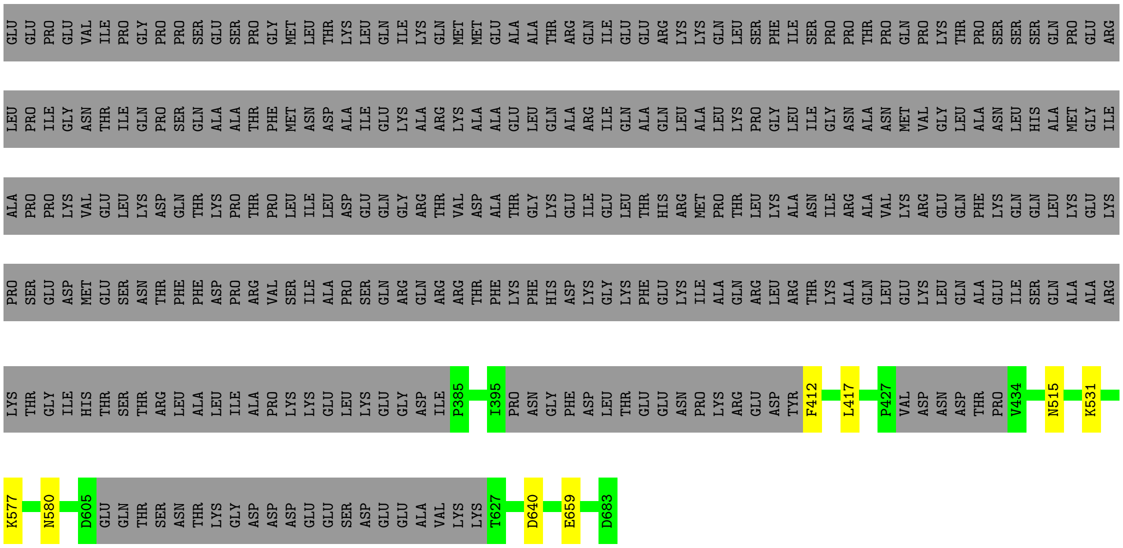


- Molecule 22: U4 snRNA



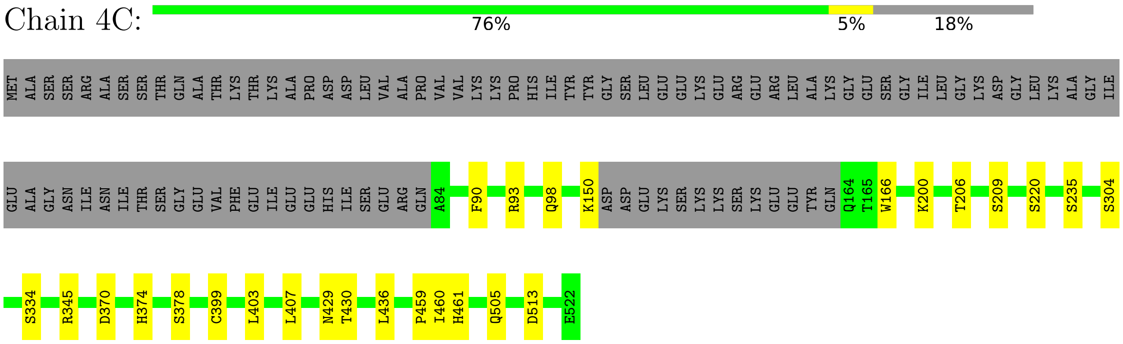
- Molecule 23: U4/U6 small nuclear ribonucleoprotein Prp3





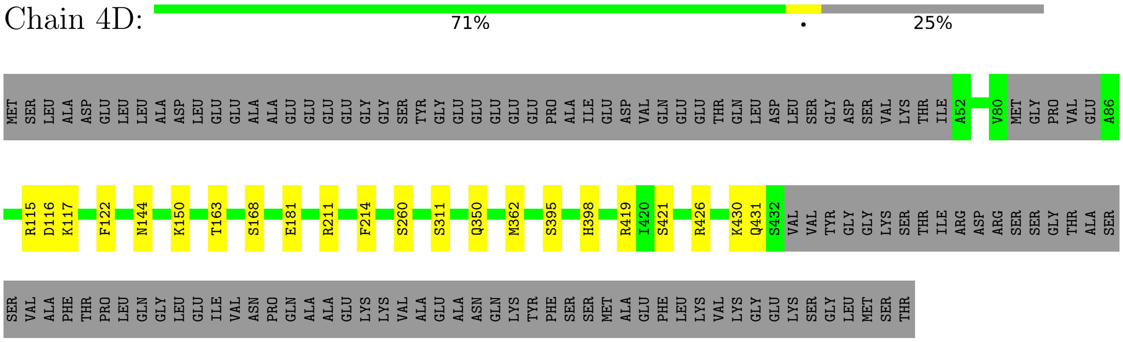
- Molecule 24: U4/U6 small nuclear ribonucleoprotein Prp4

Chain 4C:



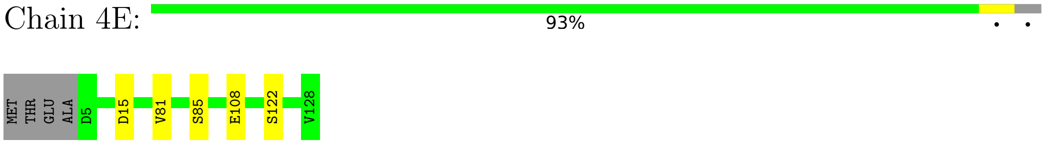
- Molecule 25: U4/U6 small nuclear ribonucleoprotein Prp31

Chain 4D:



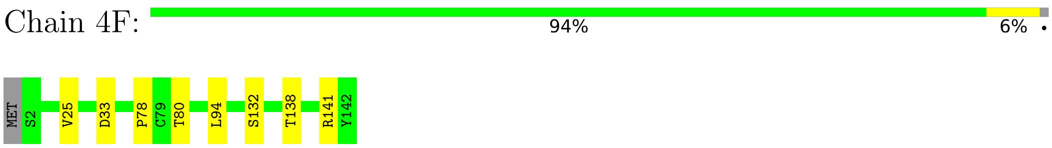
- Molecule 26: NHP2-like protein 1

Chain 4E:



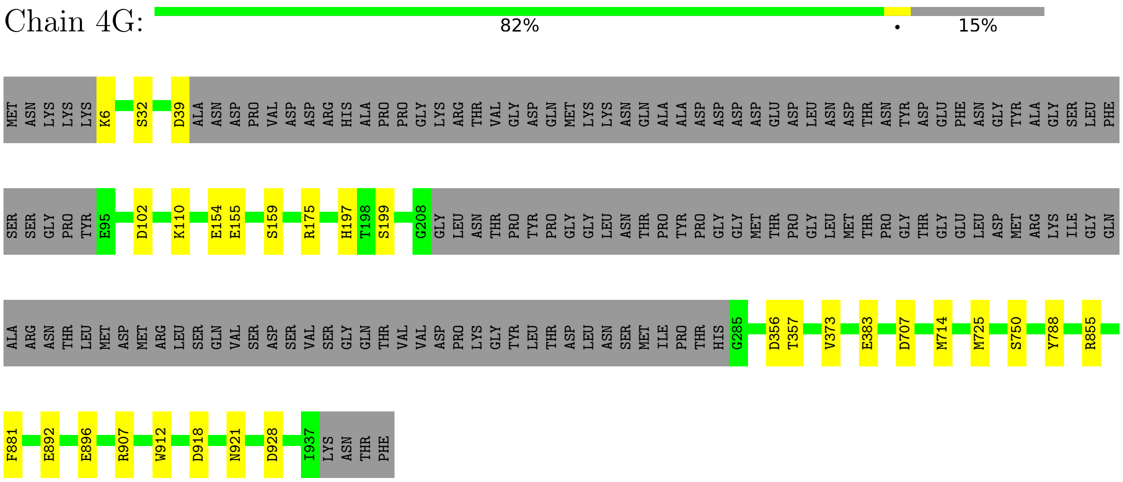
- Molecule 27: Thioredoxin-like protein 4A

Chain 4F:



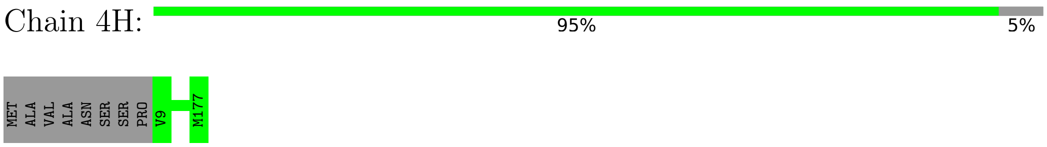
- Molecule 28: Pre-mRNA-processing factor 6

Chain 4G:



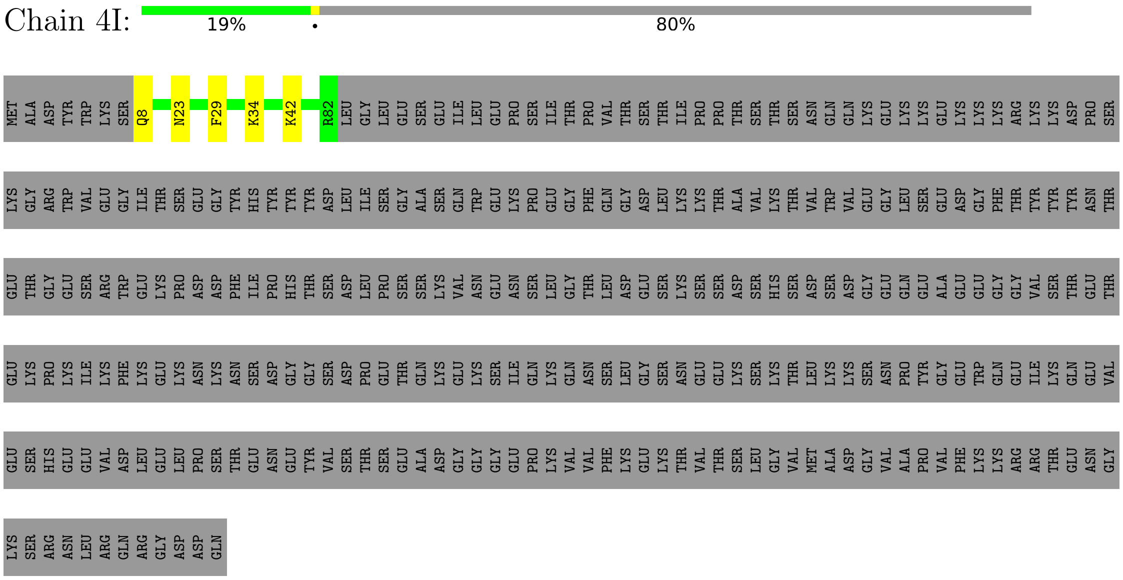
- Molecule 29: Peptidyl-prolyl cis-trans isomerase H

Chain 4H:



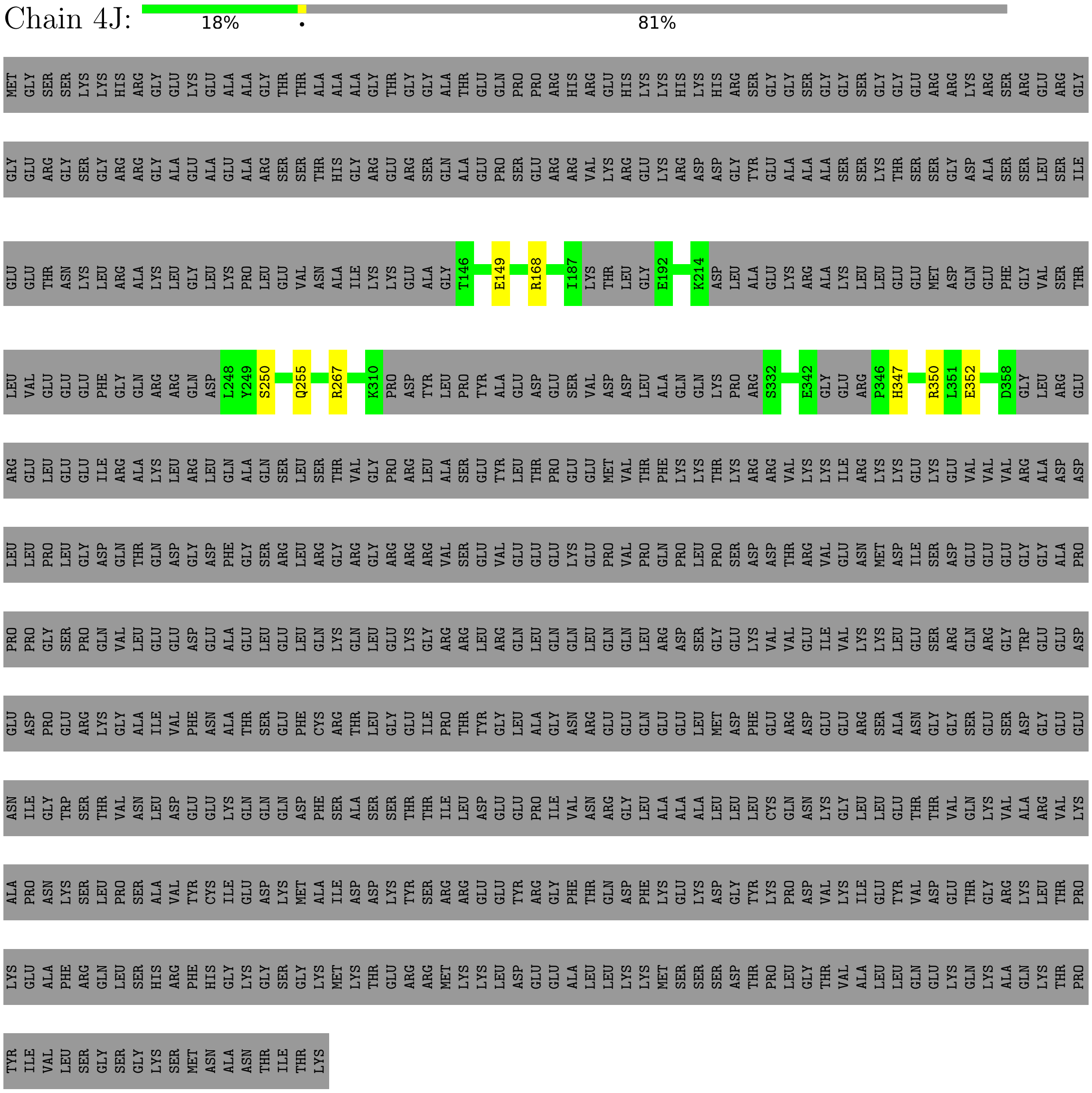
- Molecule 30: WW domain-binding protein 4

Chain 4I:



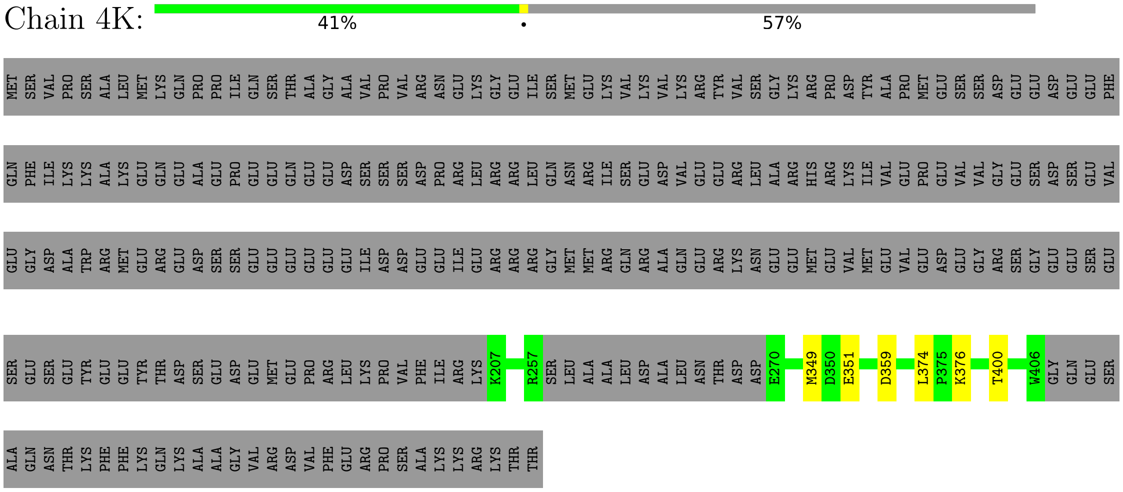
- Molecule 31: U4/U6.U5 tri-snRNP-associated protein 1

Chain 4J:



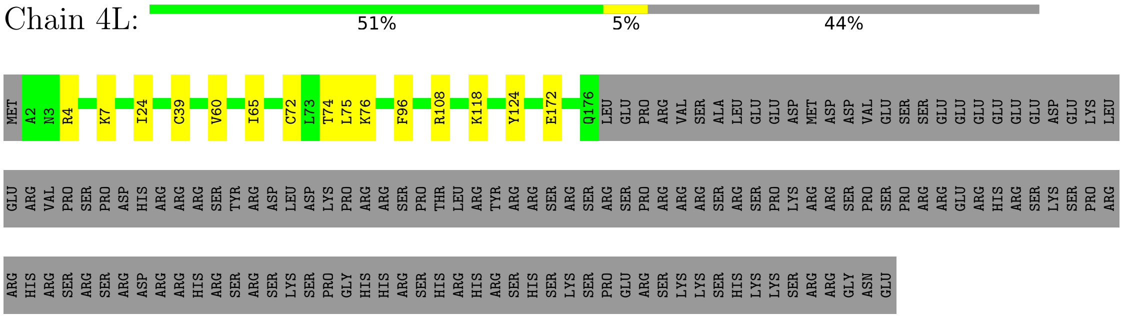
- Molecule 32: Microfibrillar-associated protein 1

Chain 4K:



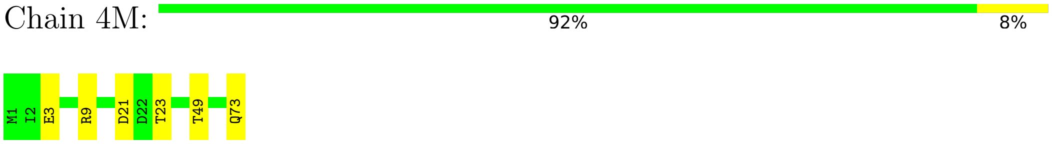
- Molecule 33: Pre-mRNA-splicing factor 38A

Chain 4L:



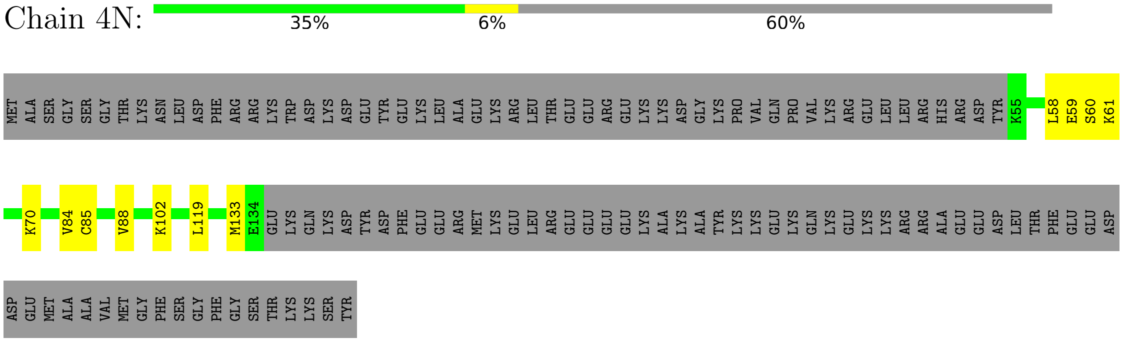
- Molecule 34: Ubiquitin-like protein 5

Chain 4M:



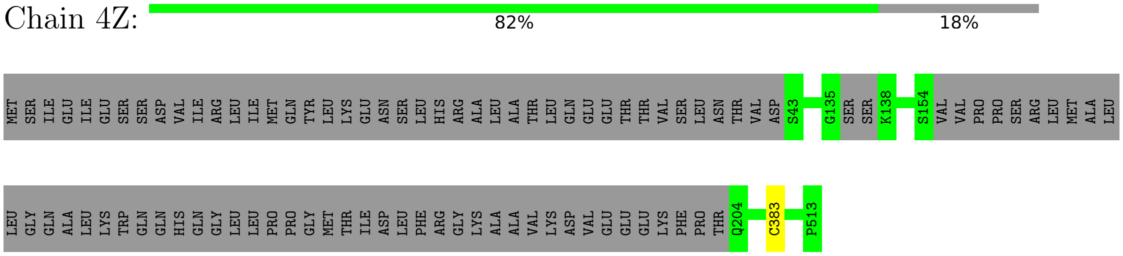
- Molecule 35: Zinc finger matrin-type protein 2

Chain 4N:



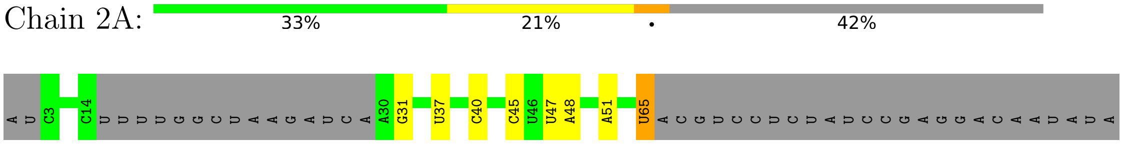
- Molecule 36: WD40 repeat-containing protein SMU1

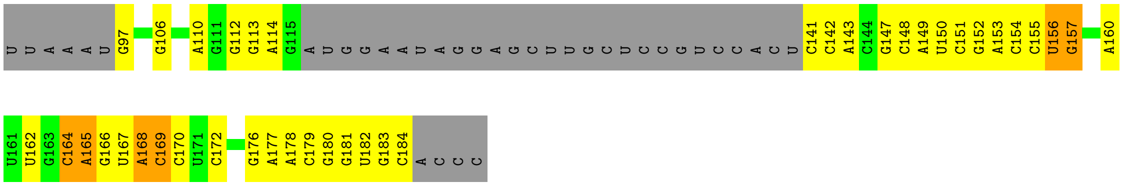
Chain 4Z:



- Molecule 37: U2 snRNA

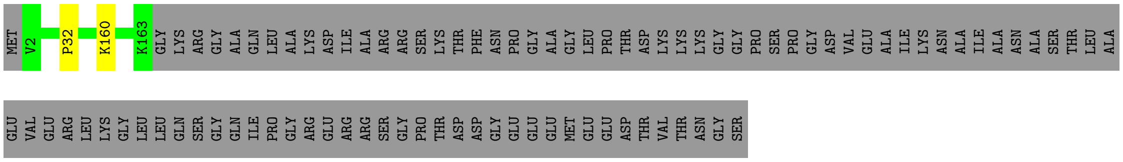
Chain 2A:





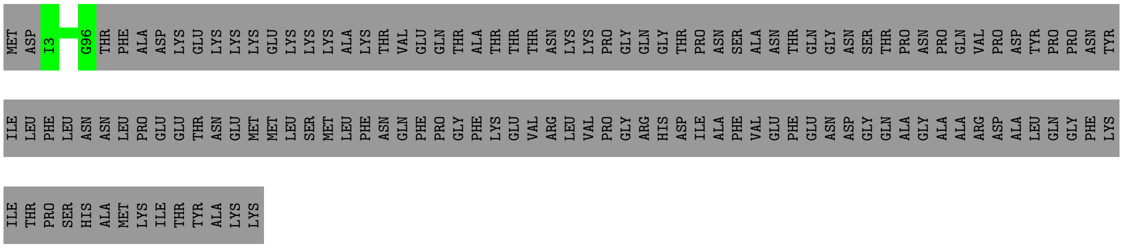
- Molecule 38: U2 small nuclear ribonucleoprotein A'

Chain 2B:



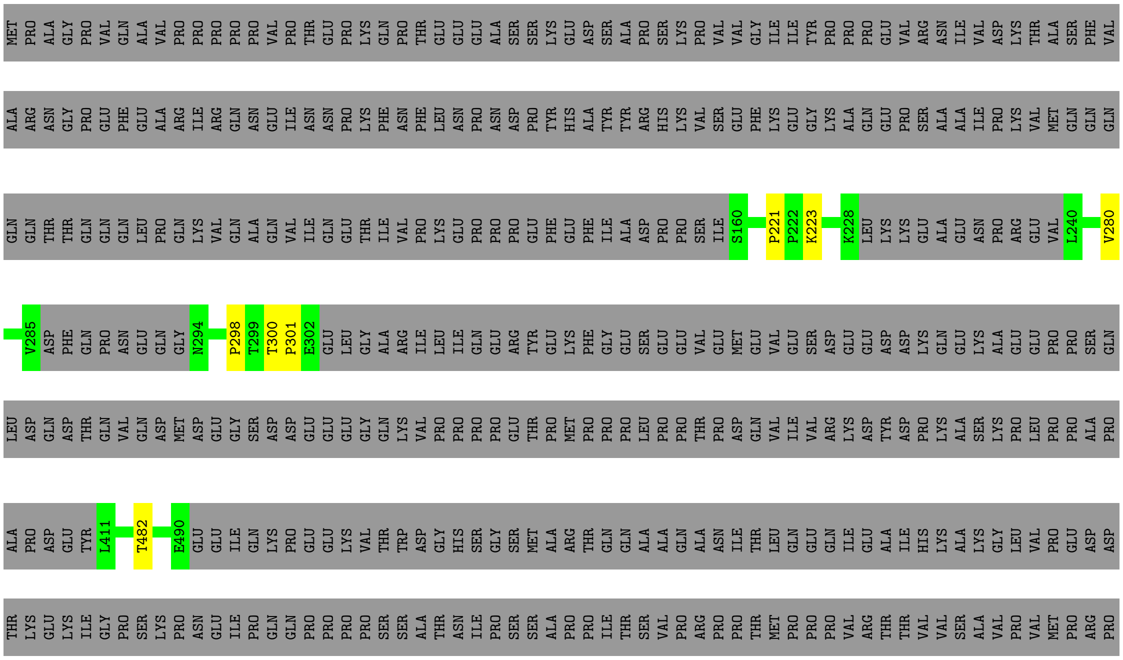
- Molecule 39: U2 small nuclear ribonucleoprotein B"

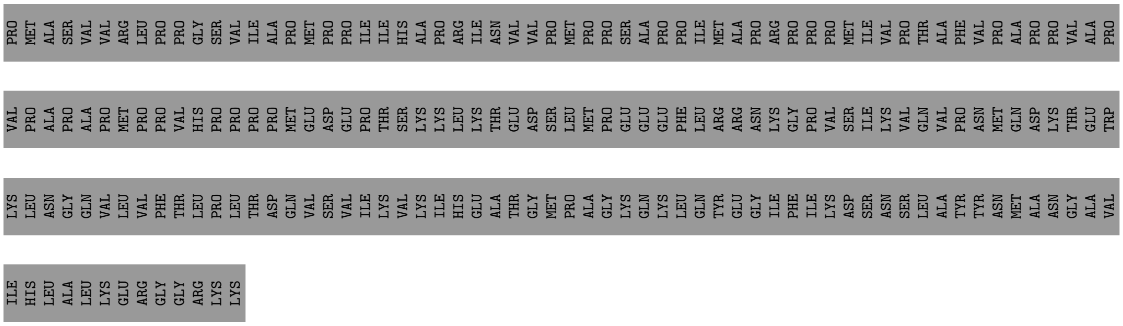
Chain 2C:



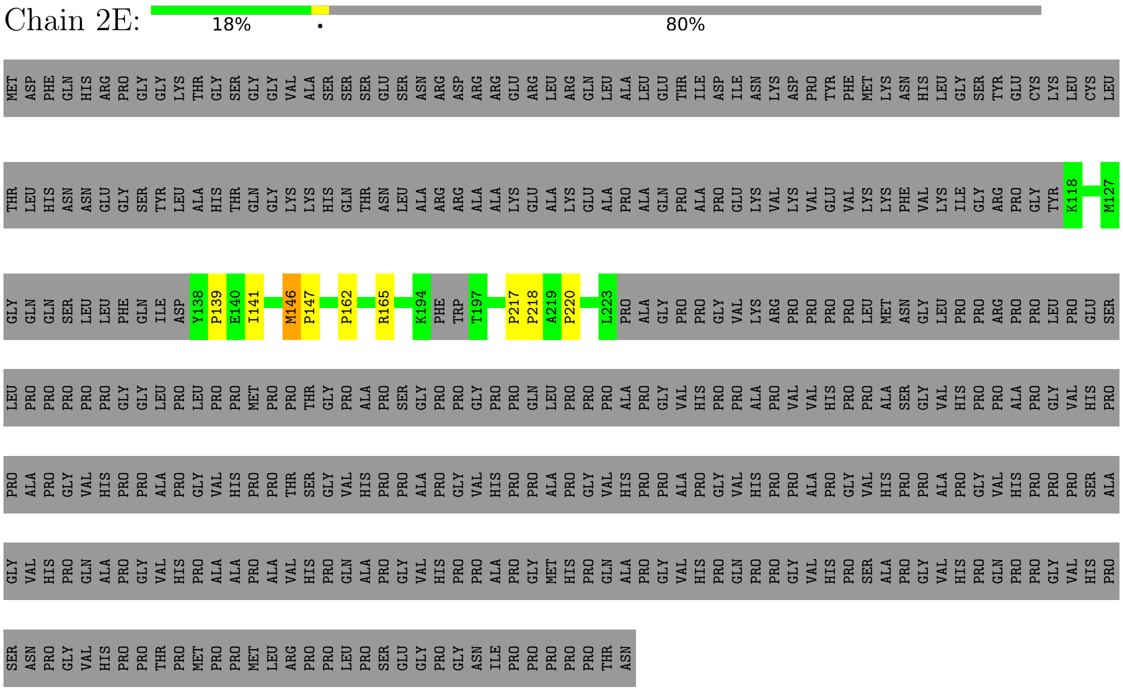
- Molecule 40: Splicing factor 3A subunit 1

Chain 2D:

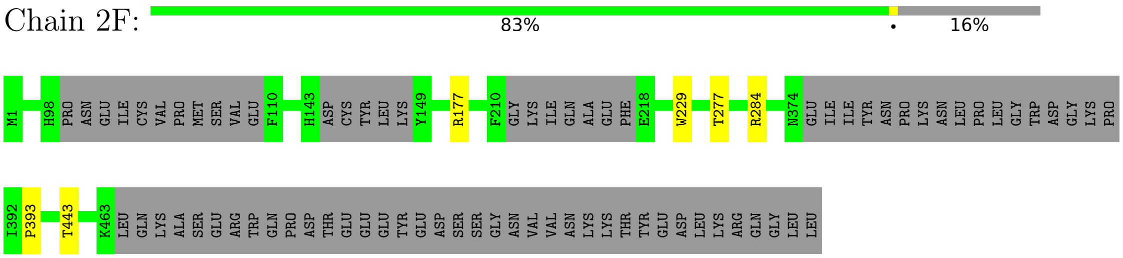




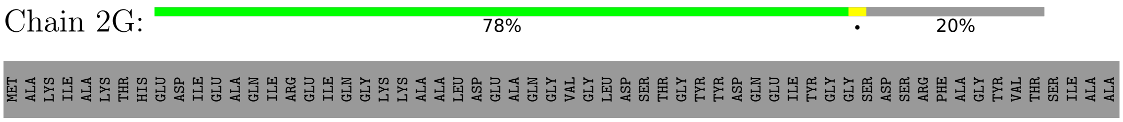
- Molecule 41: Splicing factor 3A subunit 2

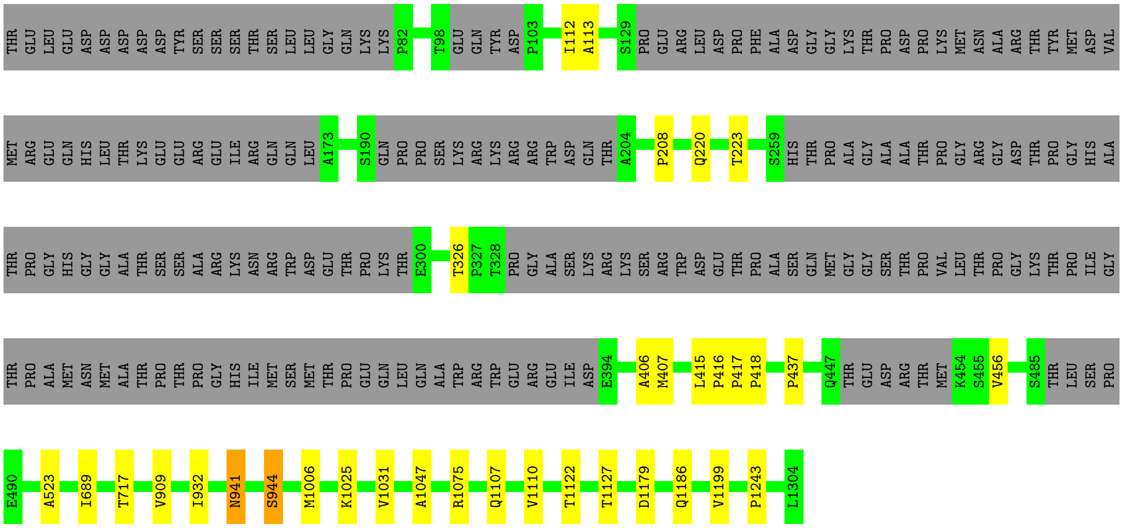


- Molecule 42: Splicing factor 3A subunit 3



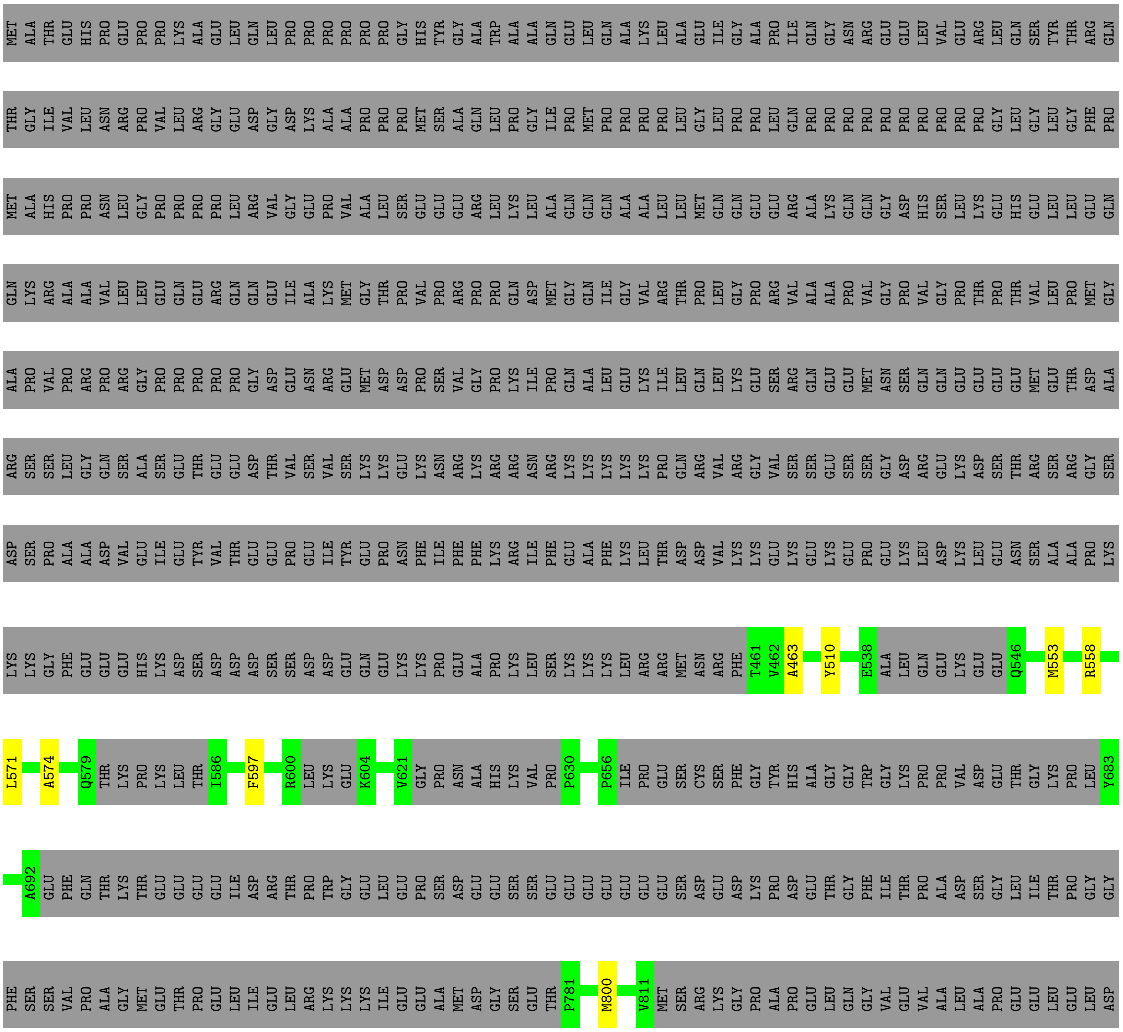
- Molecule 43: Splicing factor 3B subunit 1





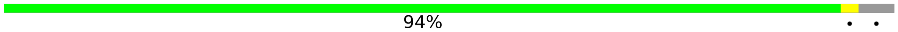
- Molecule 44: Splicing factor 3B subunit 2

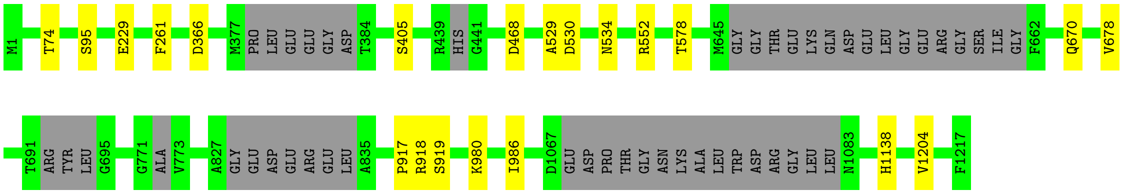
Chain 2H:



- Molecule 45: Splicing factor 3B subunit 3

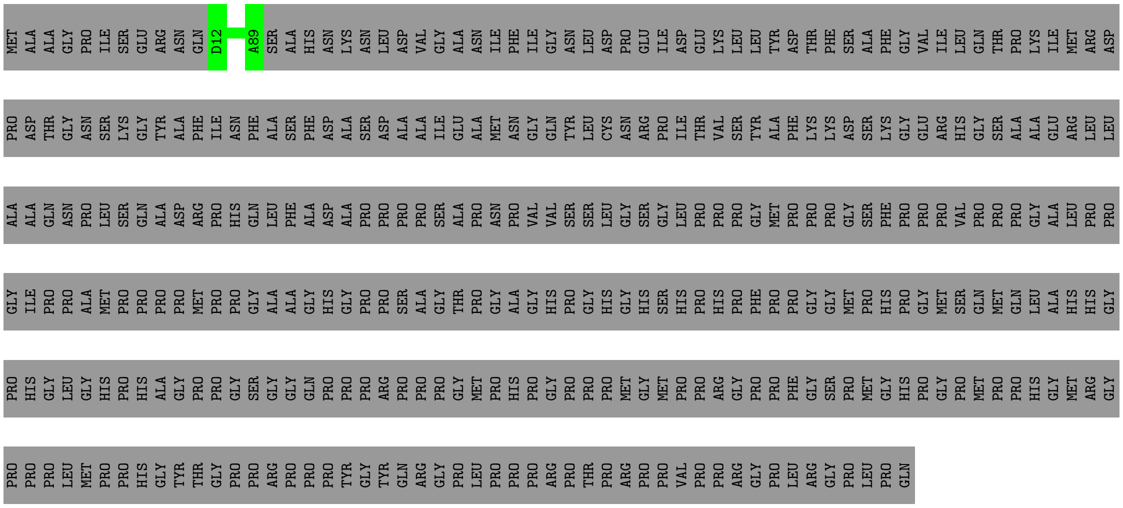
Chain 2I:





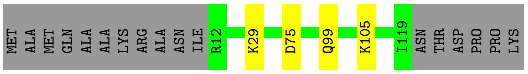
- Molecule 46: Splicing factor 3B subunit 4

Chain 2J:



- Molecule 47: Splicing factor 3B subunit 6

Chain 2K:

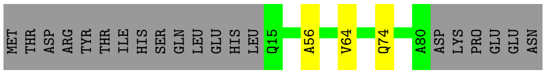


- Molecule 48: PHD finger-like domain-containing protein 5A

Chain 2L:

- Molecule 49: Splicing factor 3B subunit 5

Chain 2M:


MET
THR
ASP
ARG
TYR
THR
ILE
HIS
SER
GLN
LEU
GLU
HIS
LEU
Q15
A56
V64
Q74
A80
ASP
LYS
PRO
GLU
GLU
ASN

4 Experimental information

Property	**Value**	**Source**
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	414060	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{Å}^2)$	50	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	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: IHP, M7M, MG, GTP, ZN

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	0.65	3/1481 (0.2%)	0.92	0/2297
2	5A	0.29	0/2698	0.86	3/4195 (0.1%)
3	5B	0.26	0/19154	0.51	1/26000 (0.0%)
4	5C	0.26	0/6573	0.52	1/8929 (0.0%)
5	5D	0.26	0/13923	0.49	1/18868 (0.0%)
6	5E	0.67	0/1195	0.71	0/1492
7	2a	0.50	0/343	0.69	0/427
7	4a	0.22	0/254	0.48	0/314
7	5a	0.50	0/335	0.68	0/417
8	2b	0.56	0/327	0.68	0/407
8	4b	0.22	0/333	0.48	0/416
8	5b	0.57	0/327	0.67	0/407
9	2c	0.70	0/338	0.73	0/419
9	4c	0.23	0/298	0.48	0/370
9	5c	0.69	0/387	0.72	0/482
10	2d	0.77	0/295	0.76	0/367
10	4d	0.24	0/291	0.49	0/363
10	5d	0.77	0/295	0.76	0/367
11	2e	0.64	0/315	0.75	0/392
11	4e	0.22	0/313	0.49	0/390
11	5e	0.65	0/315	0.74	0/392
12	2f	0.55	0/270	0.63	0/334
12	4f	0.24	0/297	0.51	0/371
12	5f	0.54	0/287	0.61	0/357
13	2g	0.47	0/318	0.56	0/394
13	4g	0.23	0/287	0.49	0/358
13	5g	0.46	0/307	0.55	0/382
14	6A	0.24	0/1423	0.77	0/2211
15	6a	0.43	0/359	0.67	0/447
16	6b	0.46	0/294	0.75	0/364
17	6c	0.34	0/294	0.61	0/364
18	6d	0.43	0/286	0.59	0/354

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
19	6e	0.43	0/279	0.72	0/347
20	6f	0.38	0/258	0.61	0/319
21	6g	0.41	0/242	0.64	0/299
22	4A	0.24	0/3025	0.76	2/4702 (0.0%)
23	4B	0.25	0/2114	0.50	0/2836
24	4C	0.25	0/3452	0.52	0/4675
25	4D	0.24	0/2912	0.50	0/3924
26	4E	0.24	0/974	0.47	0/1316
27	4F	0.29	0/1198	0.54	1/1620 (0.1%)
28	4G	0.24	0/5592	0.49	1/7615 (0.0%)
29	4H	0.26	0/853	0.46	0/1188
30	4I	0.25	0/502	0.44	0/683
31	4J	0.25	0/1149	0.52	0/1542
32	4K	0.24	0/1209	0.43	0/1655
33	4L	0.25	0/1481	0.52	0/1995
34	4M	0.23	0/609	0.49	0/819
35	4N	0.27	0/646	0.52	0/859
36	4Z	0.24	0/2100	0.45	0/2926
37	2A	0.86	11/2576 (0.4%)	1.43	55/4003 (1.4%)
38	2B	0.63	0/647	1.42	0/807
39	2C	0.61	0/375	1.20	0/467
40	2D	0.23	0/1139	0.49	0/1477
41	2E	0.22	0/373	0.58	1/461 (0.2%)
42	2F	0.25	0/1688	0.47	0/2102
43	2G	1.04	4/4184 (0.1%)	0.83	2/5216 (0.0%)
44	2H	0.65	0/957	0.67	0/1209
45	2I	0.85	0/4664	0.76	0/5816
46	2J	0.62	0/311	0.64	0/387
47	2K	0.79	0/431	0.79	0/537
48	2L	0.74	0/355	0.68	0/442
49	2M	1.01	0/263	0.77	0/327
All	All	0.44	18/100770 (0.0%)	0.64	68/136218 (0.0%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
3	5B	0	1
9	2c	0	1
9	5c	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
42	2F	0	1
43	2G	0	11
44	2H	0	3
45	2I	0	11
47	2K	0	1
49	2M	0	1
All	All	0	31

All (18) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
43	2G	407	MET	N-CA	12.36	1.71	1.46
43	2G	406	ALA	C-N	7.94	1.52	1.34
37	2A	142	C	C1'-N1	7.32	1.59	1.48
43	2G	1243	PRO	N-CA	-7.11	1.35	1.47
37	2A	182	U	C1'-N1	6.94	1.59	1.48
37	2A	150	U	C1'-N1	6.74	1.58	1.48
37	2A	151	C	C1'-N1	6.53	1.58	1.48
37	2A	97	G	C1'-N9	-6.41	1.37	1.46
37	2A	141	C	C1'-N1	6.38	1.58	1.48
37	2A	184	C	C1'-N1	6.35	1.58	1.48
37	2A	148	C	C1'-N1	6.33	1.58	1.48
43	2G	944	SER	N-CA	-5.72	1.34	1.46
37	2A	65	U	C1'-N1	5.54	1.57	1.48
37	2A	48	A	C1'-N9	-5.48	1.39	1.46
1	A	9	U	C1'-N1	5.14	1.56	1.48
1	A	8	U	C1'-N1	5.13	1.56	1.48
1	A	7	U	C1'-N1	5.13	1.56	1.48
37	2A	110	A	C1'-N9	-5.08	1.39	1.46

All (68) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
37	2A	167	U	C5-C4-O4	11.61	132.87	125.90
43	2G	406	ALA	C-N-CA	10.28	147.39	121.70
37	2A	164	C	N1-C2-O2	-10.12	112.83	118.90
3	5B	1194	CYS	CA-CB-SG	9.09	130.36	114.00
37	2A	162	U	N3-C2-O2	-8.97	115.92	122.20
37	2A	164	C	C5'-C4'-O4'	-8.21	99.25	109.10
37	2A	169	C	P-O3'-C3'	8.20	129.54	119.70
37	2A	166	G	O4'-C1'-N9	8.03	114.62	108.20
37	2A	167	U	N3-C4-O4	-7.89	113.88	119.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
37	2A	164	C	P-O3'-C3'	7.54	128.75	119.70
37	2A	167	U	N1-C2-O2	7.41	127.98	122.80
37	2A	164	C	N3-C2-O2	7.38	127.07	121.90
37	2A	113	G	OP2-P-O3'	7.26	121.17	105.20
37	2A	149	A	OP2-P-O3'	7.26	121.16	105.20
37	2A	141	C	OP2-P-O3'	7.24	121.13	105.20
37	2A	114	A	OP2-P-O3'	7.24	121.12	105.20
37	2A	183	G	OP2-P-O3'	7.22	121.08	105.20
37	2A	181	G	OP2-P-O3'	7.21	121.07	105.20
37	2A	182	U	OP2-P-O3'	7.21	121.07	105.20
37	2A	180	G	OP2-P-O3'	7.21	121.06	105.20
37	2A	150	U	OP2-P-O3'	7.20	121.05	105.20
37	2A	148	C	OP2-P-O3'	7.18	121.00	105.20
37	2A	168	A	P-O5'-C5'	-7.14	109.47	120.90
2	5A	77	G	P-O3'-C3'	-7.04	111.26	119.70
37	2A	167	U	N3-C2-O2	-6.96	117.33	122.20
37	2A	180	G	O3'-P-O5'	-6.83	91.02	104.00
37	2A	149	A	O3'-P-O5'	-6.83	91.03	104.00
43	2G	406	ALA	CA-C-O	-6.82	105.77	120.10
37	2A	155	C	P-O3'-C3'	6.82	127.88	119.70
37	2A	182	U	O3'-P-O5'	-6.81	91.06	104.00
37	2A	183	G	O3'-P-O5'	-6.79	91.11	104.00
37	2A	148	C	O3'-P-O5'	-6.78	91.12	104.00
37	2A	141	C	O3'-P-O5'	-6.77	91.14	104.00
37	2A	150	U	O3'-P-O5'	-6.75	91.17	104.00
37	2A	113	G	O3'-P-O5'	-6.75	91.19	104.00
37	2A	181	G	O3'-P-O5'	-6.74	91.19	104.00
37	2A	114	A	O3'-P-O5'	-6.73	91.21	104.00
4	5C	308	CYS	CA-CB-SG	6.22	125.20	114.00
37	2A	165	A	O4'-C1'-N9	-6.19	103.25	108.20
37	2A	166	G	N9-C4-C5	6.14	107.86	105.40
37	2A	166	G	C8-N9-C4	-6.11	103.96	106.40
37	2A	162	U	N1-C2-O2	6.07	127.05	122.80
37	2A	166	G	N3-C4-C5	-6.01	125.60	128.60
37	2A	168	A	C5'-C4'-C3'	-5.93	106.51	116.00
28	4G	707	ASP	C-N-CA	5.89	136.43	121.70
37	2A	172	C	P-O3'-C3'	5.80	126.66	119.70
2	5A	76	A	P-O3'-C3'	-5.78	112.77	119.70
37	2A	156	U	P-O3'-C3'	-5.77	112.78	119.70
41	2E	146	MET	C-N-CA	5.75	146.17	122.00
5	5D	583	THR	C-N-CA	5.75	136.07	121.70
37	2A	167	U	O3'-P-O5'	-5.74	93.09	104.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
37	2A	164	C	C5-C4-N4	-5.72	116.20	120.20
27	4F	78	PRO	N-CD-CG	-5.55	94.88	103.20
37	2A	157	G	O4'-C1'-N9	-5.50	103.80	108.20
37	2A	166	G	C6-N1-C2	-5.46	121.83	125.10
37	2A	106	G	O5'-P-OP1	5.41	117.19	110.70
22	4A	1	A	OP1-P-OP2	5.36	127.64	119.60
37	2A	156	U	OP2-P-O3'	5.28	116.82	105.20
37	2A	160	A	P-O5'-C5'	-5.27	112.46	120.90
37	2A	170	C	O4'-C1'-C2'	-5.23	100.57	105.80
22	4A	70	U	C2-N1-C1'	5.20	123.94	117.70
37	2A	164	C	C6-N1-C2	5.17	122.37	120.30
37	2A	157	G	P-O5'-C5'	-5.15	112.66	120.90
37	2A	170	C	N3-C4-C5	-5.14	119.84	121.90
37	2A	156	U	C4'-C3'-C2'	5.08	107.69	102.60
2	5A	23	C	C2-N1-C1'	5.07	124.38	118.80
37	2A	162	U	C2-N3-C4	-5.03	123.98	127.00
37	2A	176	G	N9-C4-C5	5.01	107.41	105.40

There are no chirality outliers.

All (31) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
42	2F	443	THR	Peptide
43	2G	1025	LYS	Peptide
43	2G	1122	THR	Peptide
43	2G	1127	THR	Peptide
43	2G	1179	ASP	Peptide
43	2G	1199	VAL	Peptide
43	2G	220	GLN	Peptide
43	2G	415	LEU	Peptide,Mainchain
43	2G	689	ILE	Peptide
43	2G	941	ASN	Peptide
43	2G	944	SER	Peptide
44	2H	553	MET	Peptide
44	2H	558	ARG	Peptide
44	2H	571	LEU	Peptide
45	2I	261	PHE	Peptide
45	2I	366	ASP	Peptide
45	2I	468	ASP	Peptide
45	2I	530	ASP	Peptide
45	2I	534	ASN	Peptide
45	2I	552	ARG	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
45	2I	670	GLN	Peptide
45	2I	678	VAL	Peptide
45	2I	74	THR	Peptide
45	2I	980	LYS	Peptide
45	2I	986	ILE	Peptide
47	2K	29	LYS	Peptide
49	2M	74	GLN	Peptide
9	2c	112	ASN	Peptide
3	5B	941	LYS	Peptide
9	5c	112	ASN	Peptide

5.2 Too-close contacts ⓘ

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

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	5B	2249/2335 (96%)	2156 (96%)	93 (4%)	0	100	100
4	5C	814/972 (84%)	755 (93%)	58 (7%)	1 (0%)	51	78
5	5D	1694/2136 (79%)	1618 (96%)	75 (4%)	1 (0%)	51	78
6	5E	297/357 (83%)	272 (92%)	16 (5%)	9 (3%)	4	10
7	2a	84/231 (36%)	82 (98%)	2 (2%)	0	100	100
7	4a	60/231 (26%)	57 (95%)	3 (5%)	0	100	100
7	5a	82/231 (36%)	80 (98%)	2 (2%)	0	100	100
8	2b	80/119 (67%)	77 (96%)	3 (4%)	0	100	100
8	4b	80/119 (67%)	76 (95%)	4 (5%)	0	100	100
8	5b	80/119 (67%)	77 (96%)	3 (4%)	0	100	100
9	2c	81/118 (69%)	78 (96%)	3 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
9	4c	70/118 (59%)	68 (97%)	2 (3%)	0	100	100
9	5c	95/118 (80%)	91 (96%)	4 (4%)	0	100	100
10	2d	72/86 (84%)	68 (94%)	4 (6%)	0	100	100
10	4d	69/86 (80%)	67 (97%)	2 (3%)	0	100	100
10	5d	72/86 (84%)	69 (96%)	3 (4%)	0	100	100
11	2e	77/92 (84%)	76 (99%)	1 (1%)	0	100	100
11	4e	76/92 (83%)	70 (92%)	6 (8%)	0	100	100
11	5e	77/92 (84%)	76 (99%)	1 (1%)	0	100	100
12	2f	64/76 (84%)	62 (97%)	2 (3%)	0	100	100
12	4f	71/76 (93%)	67 (94%)	4 (6%)	0	100	100
12	5f	70/76 (92%)	68 (97%)	2 (3%)	0	100	100
13	2g	76/126 (60%)	75 (99%)	1 (1%)	0	100	100
13	4g	69/126 (55%)	69 (100%)	0	0	100	100
13	5g	75/126 (60%)	73 (97%)	2 (3%)	0	100	100
15	6a	88/95 (93%)	77 (88%)	7 (8%)	4 (4%)	2	5
16	6b	70/102 (69%)	64 (91%)	3 (4%)	3 (4%)	2	5
17	6c	70/139 (50%)	63 (90%)	6 (9%)	1 (1%)	11	28
18	6d	68/91 (75%)	63 (93%)	4 (6%)	1 (2%)	10	26
19	6e	68/80 (85%)	64 (94%)	2 (3%)	2 (3%)	4	10
20	6f	61/103 (59%)	56 (92%)	5 (8%)	0	100	100
21	6g	57/96 (59%)	52 (91%)	4 (7%)	1 (2%)	8	21
23	4B	248/683 (36%)	235 (95%)	13 (5%)	0	100	100
24	4C	422/522 (81%)	393 (93%)	28 (7%)	1 (0%)	47	73
25	4D	372/499 (74%)	353 (95%)	19 (5%)	0	100	100
26	4E	122/128 (95%)	115 (94%)	7 (6%)	0	100	100
27	4F	139/142 (98%)	134 (96%)	5 (4%)	0	100	100
28	4G	795/941 (84%)	747 (94%)	48 (6%)	0	100	100
29	4H	167/177 (94%)	158 (95%)	9 (5%)	0	100	100
30	4I	73/376 (19%)	73 (100%)	0	0	100	100
31	4J	142/800 (18%)	136 (96%)	6 (4%)	0	100	100
32	4K	184/439 (42%)	174 (95%)	10 (5%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
33	4L	173/312 (55%)	165 (95%)	6 (4%)	2 (1%)	13	32
34	4M	71/73 (97%)	66 (93%)	5 (7%)	0	100	100
35	4N	78/199 (39%)	77 (99%)	0	1 (1%)	12	30
36	4Z	414/513 (81%)	401 (97%)	12 (3%)	1 (0%)	47	73
38	2B	160/255 (63%)	146 (91%)	12 (8%)	2 (1%)	12	30
39	2C	92/225 (41%)	90 (98%)	2 (2%)	0	100	100
40	2D	196/793 (25%)	180 (92%)	10 (5%)	6 (3%)	4	9
41	2E	88/464 (19%)	63 (72%)	16 (18%)	9 (10%)	0	0
42	2F	413/501 (82%)	367 (89%)	41 (10%)	5 (1%)	13	32
43	2G	1032/1304 (79%)	844 (82%)	166 (16%)	22 (2%)	7	18
44	2H	199/895 (22%)	179 (90%)	16 (8%)	4 (2%)	7	19
45	2I	1152/1217 (95%)	1053 (91%)	89 (8%)	10 (1%)	17	40
46	2J	76/424 (18%)	75 (99%)	1 (1%)	0	100	100
47	2K	106/125 (85%)	85 (80%)	18 (17%)	3 (3%)	5	11
48	2L	87/110 (79%)	74 (85%)	13 (15%)	0	100	100
49	2M	64/86 (74%)	55 (86%)	7 (11%)	2 (3%)	4	9
All	All	14181/21253 (67%)	13204 (93%)	886 (6%)	91 (1%)	29	50

All (91) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	5D	1086	GLN
6	5E	193	THR
15	6a	55	LEU
16	6b	84	MET
18	6d	70	ASP
19	6e	52	VAL
19	6e	55	GLN
24	4C	459	PRO
36	4Z	383	CYS
40	2D	301	PRO
41	2E	139	PRO
41	2E	141	ILE
41	2E	146	MET
41	2E	162	PRO
41	2E	165	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
41	2E	218	PRO
42	2F	284	ARG
43	2G	208	PRO
43	2G	416	PRO
43	2G	418	PRO
43	2G	456	VAL
43	2G	717	THR
43	2G	941	ASN
43	2G	1107	GLN
45	2I	405	SER
45	2I	919	SER
47	2K	99	GLN
47	2K	105	LYS
15	6a	74	ALA
16	6b	97	PRO
17	6c	12	ASN
38	2B	160	LYS
40	2D	223	LYS
40	2D	280	VAL
42	2F	277	THR
43	2G	113	ALA
43	2G	1110	VAL
44	2H	597	PHE
45	2I	917	PRO
6	5E	60	MET
6	5E	88	ARG
6	5E	256	ASP
33	4L	39	CYS
42	2F	177	ARG
42	2F	393	PRO
44	2H	510	TYR
6	5E	162	ARG
16	6b	96	ALA
38	2B	32	PRO
40	2D	300	THR
43	2G	112	ILE
43	2G	437	PRO
43	2G	523	ALA
43	2G	909	VAL
43	2G	1006	MET
44	2H	463	ALA
44	2H	574	ALA

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
45	2I	529	ALA
45	2I	578	THR
47	2K	75	ASP
6	5E	159	PRO
15	6a	73	PRO
35	4N	85	CYS
41	2E	147	PRO
41	2E	217	PRO
43	2G	1047	ALA
43	2G	1075	ARG
43	2G	1186	GLN
45	2I	95	SER
45	2I	229	GLU
6	5E	270	LYS
21	6g	34	ILE
33	4L	65	ILE
41	2E	220	PRO
43	2G	326	THR
43	2G	932	ILE
45	2I	918	ARG
45	2I	1138	HIS
49	2M	56	ALA
6	5E	149	GLY
43	2G	417	PRO
40	2D	221	PRO
43	2G	223	THR
42	2F	229	TRP
45	2I	1204	VAL
4	5C	439	PRO
6	5E	324	PRO
40	2D	298	PRO
43	2G	1031	VAL
49	2M	64	VAL
15	6a	52	PRO

5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	5B	2033/2108 (96%)	1938 (95%)	95 (5%)	26	54
4	5C	717/866 (83%)	678 (95%)	39 (5%)	22	47
5	5D	1517/1908 (80%)	1492 (98%)	25 (2%)	62	85
23	4B	225/599 (38%)	217 (96%)	8 (4%)	35	64
24	4C	362/442 (82%)	336 (93%)	26 (7%)	14	34
25	4D	299/424 (70%)	277 (93%)	22 (7%)	13	32
26	4E	108/111 (97%)	103 (95%)	5 (5%)	27	54
27	4F	129/130 (99%)	122 (95%)	7 (5%)	22	47
28	4G	417/792 (53%)	389 (93%)	28 (7%)	16	37
29	4H	10/148 (7%)	10 (100%)	0	100	100
30	4I	32/333 (10%)	27 (84%)	5 (16%)	2	7
31	4J	112/681 (16%)	104 (93%)	8 (7%)	14	34
32	4K	66/395 (17%)	60 (91%)	6 (9%)	9	21
33	4L	159/293 (54%)	146 (92%)	13 (8%)	11	26
34	4M	66/66 (100%)	60 (91%)	6 (9%)	9	21
35	4N	75/181 (41%)	65 (87%)	10 (13%)	4	9
36	4Z	11/450 (2%)	11 (100%)	0	100	100
40	2D	70/709 (10%)	69 (99%)	1 (1%)	67	86
44	2H	26/776 (3%)	25 (96%)	1 (4%)	33	62
All	All	6434/11412 (56%)	6129 (95%)	305 (5%)	30	54

All (305) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	5B	68	LYS
3	5B	81	PHE
3	5B	87	VAL
3	5B	104	GLU
3	5B	142	SER
3	5B	150	MET
3	5B	181	ASN
3	5B	183	LEU
3	5B	195	LEU
3	5B	200	ASP
3	5B	223	SER
3	5B	227	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	5B	236	SER
3	5B	248	ASP
3	5B	282	LEU
3	5B	305	ARG
3	5B	322	ASN
3	5B	330	THR
3	5B	339	PHE
3	5B	341	LYS
3	5B	351	TYR
3	5B	352	PHE
3	5B	353	ASP
3	5B	361	HIS
3	5B	379	GLU
3	5B	387	PHE
3	5B	417	ARG
3	5B	468	LYS
3	5B	484	SER
3	5B	506	LEU
3	5B	510	ARG
3	5B	513	LEU
3	5B	519	ASP
3	5B	528	LYS
3	5B	615	ARG
3	5B	650	ARG
3	5B	679	SER
3	5B	741	ARG
3	5B	748	ASP
3	5B	774	LYS
3	5B	819	SER
3	5B	821	ARG
3	5B	845	ARG
3	5B	855	ARG
3	5B	859	SER
3	5B	866	LEU
3	5B	934	ARG
3	5B	945	THR
3	5B	989	ASP
3	5B	991	THR
3	5B	994	ASN
3	5B	1032	ARG
3	5B	1089	CYS
3	5B	1091	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	5B	1104	ASP
3	5B	1122	ASN
3	5B	1129	ASN
3	5B	1137	ASP
3	5B	1140	MET
3	5B	1173	SER
3	5B	1176	SER
3	5B	1199	LYS
3	5B	1223	GLU
3	5B	1239	ARG
3	5B	1303	LEU
3	5B	1310	ARG
3	5B	1341	ARG
3	5B	1359	HIS
3	5B	1361	GLU
3	5B	1370	ARG
3	5B	1377	SER
3	5B	1402	ARG
3	5B	1413	ASP
3	5B	1420	ASN
3	5B	1438	VAL
3	5B	1536	LEU
3	5B	1544	ARG
3	5B	1634	SER
3	5B	1678	ARG
3	5B	1697	SER
3	5B	1753	LEU
3	5B	1765	SER
3	5B	1831	LYS
3	5B	1866	LYS
3	5B	1872	LEU
3	5B	1894	GLN
3	5B	1904	ASP
3	5B	1908	LYS
3	5B	2031	LYS
3	5B	2046	THR
3	5B	2056	THR
3	5B	2073	TRP
3	5B	2094	SER
3	5B	2231	THR
3	5B	2265	ASP
4	5C	121	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	5C	134	LEU
4	5C	198	TYR
4	5C	261	ASP
4	5C	273	ASP
4	5C	301	SER
4	5C	327	TYR
4	5C	352	LYS
4	5C	359	LYS
4	5C	394	ARG
4	5C	423	PHE
4	5C	458	ASP
4	5C	476	CYS
4	5C	507	VAL
4	5C	512	GLU
4	5C	514	TYR
4	5C	533	SER
4	5C	534	VAL
4	5C	543	ARG
4	5C	562	THR
4	5C	589	LYS
4	5C	592	VAL
4	5C	623	GLU
4	5C	649	SER
4	5C	713	LYS
4	5C	718	PHE
4	5C	721	LYS
4	5C	724	TRP
4	5C	727	LEU
4	5C	740	THR
4	5C	749	THR
4	5C	770	PHE
4	5C	780	CYS
4	5C	803	ARG
4	5C	826	ARG
4	5C	880	SER
4	5C	912	LEU
4	5C	919	ARG
4	5C	928	HIS
5	5D	685	LEU
5	5D	726	HIS
5	5D	780	TYR
5	5D	801	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	5D	844	LEU
5	5D	891	SER
5	5D	984	LEU
5	5D	990	HIS
5	5D	992	TYR
5	5D	1039	LYS
5	5D	1092	MET
5	5D	1148	PHE
5	5D	1159	ASN
5	5D	1320	LEU
5	5D	1325	PHE
5	5D	1376	CYS
5	5D	1417	LYS
5	5D	1478	SER
5	5D	1507	SER
5	5D	1580	CYS
5	5D	1607	SER
5	5D	1732	MET
5	5D	1841	LYS
5	5D	1948	MET
5	5D	2014	TYR
23	4B	412	PHE
23	4B	417	LEU
23	4B	515	ASN
23	4B	531	LYS
23	4B	577	LYS
23	4B	580	ASN
23	4B	640	ASP
23	4B	659	GLU
24	4C	90	PHE
24	4C	93	ARG
24	4C	98	GLN
24	4C	150	LYS
24	4C	166	TRP
24	4C	200	LYS
24	4C	206	THR
24	4C	209	SER
24	4C	220	SER
24	4C	235	SER
24	4C	304	SER
24	4C	334	SER
24	4C	345	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
24	4C	370	ASP
24	4C	374	HIS
24	4C	378	SER
24	4C	399	CYS
24	4C	403	LEU
24	4C	407	LEU
24	4C	429	ASN
24	4C	430	THR
24	4C	436	LEU
24	4C	460	ILE
24	4C	461	HIS
24	4C	505	GLN
24	4C	513	ASP
25	4D	115	ARG
25	4D	116	ASP
25	4D	117	LYS
25	4D	122	PHE
25	4D	144	ASN
25	4D	150	LYS
25	4D	163	THR
25	4D	168	SER
25	4D	181	GLU
25	4D	211	ARG
25	4D	214	PHE
25	4D	260	SER
25	4D	311	SER
25	4D	350	GLN
25	4D	362	MET
25	4D	395	SER
25	4D	398	HIS
25	4D	419	ARG
25	4D	421	SER
25	4D	426	ARG
25	4D	430	LYS
25	4D	431	GLN
26	4E	15	ASP
26	4E	81	VAL
26	4E	85	SER
26	4E	108	GLU
26	4E	122	SER
27	4F	25	VAL
27	4F	33	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
27	4F	80	THR
27	4F	94	LEU
27	4F	132	SER
27	4F	138	THR
27	4F	141	ARG
28	4G	6	LYS
28	4G	32	SER
28	4G	39	ASP
28	4G	102	ASP
28	4G	110	LYS
28	4G	154	GLU
28	4G	155	GLU
28	4G	159	SER
28	4G	175	ARG
28	4G	197	HIS
28	4G	199	SER
28	4G	356	ASP
28	4G	357	THR
28	4G	373	VAL
28	4G	383	GLU
28	4G	714	MET
28	4G	725	MET
28	4G	750	SER
28	4G	788	TYR
28	4G	855	ARG
28	4G	881	PHE
28	4G	892	GLU
28	4G	896	GLU
28	4G	907	ARG
28	4G	912	TRP
28	4G	918	ASP
28	4G	921	ASN
28	4G	928	ASP
30	4I	8	GLN
30	4I	23	ASN
30	4I	29	PHE
30	4I	34	LYS
30	4I	42	LYS
31	4J	149	GLU
31	4J	168	ARG
31	4J	250	SER
31	4J	255	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
31	4J	267	ARG
31	4J	347	HIS
31	4J	350	ARG
31	4J	352	GLU
32	4K	349	MET
32	4K	351	GLU
32	4K	359	ASP
32	4K	374	LEU
32	4K	376	LYS
32	4K	400	THR
33	4L	4	ARG
33	4L	7	LYS
33	4L	24	ILE
33	4L	60	VAL
33	4L	72	CYS
33	4L	74	THR
33	4L	75	LEU
33	4L	76	LYS
33	4L	96	PHE
33	4L	108	ARG
33	4L	118	LYS
33	4L	124	TYR
33	4L	172	GLU
34	4M	3	GLU
34	4M	9	ARG
34	4M	21	ASP
34	4M	23	THR
34	4M	49	THR
34	4M	73	GLN
35	4N	58	LEU
35	4N	59	GLU
35	4N	60	SER
35	4N	61	LYS
35	4N	70	LYS
35	4N	84	VAL
35	4N	88	VAL
35	4N	102	LYS
35	4N	119	LEU
35	4N	133	MET
40	2D	482	THR
44	2H	800	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (36) such sidechains are listed below:

Mol	Chain	Res	Type
3	5B	65	HIS
3	5B	73	HIS
3	5B	105	ASN
3	5B	361	HIS
3	5B	434	HIS
3	5B	509	HIS
3	5B	542	ASN
3	5B	545	HIS
3	5B	1450	GLN
3	5B	1487	HIS
3	5B	1552	GLN
3	5B	1563	HIS
3	5B	1944	HIS
3	5B	1966	HIS
3	5B	2166	HIS
4	5C	140	HIS
4	5C	245	HIS
4	5C	627	HIS
5	5D	785	HIS
5	5D	911	GLN
5	5D	1515	HIS
23	4B	511	HIS
23	4B	668	HIS
24	4C	282	HIS
24	4C	322	HIS
24	4C	421	HIS
25	4D	111	HIS
25	4D	270	HIS
26	4E	17	HIS
26	4E	28	GLN
26	4E	87	GLN
27	4F	89	HIS
28	4G	741	HIS
28	4G	773	ASN
31	4J	261	HIS
33	4L	81	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	62/144 (43%)	37 (59%)	12 (19%)

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
14	6A	56/107 (52%)	10 (17%)	1 (1%)
2	5A	114/117 (97%)	30 (26%)	5 (4%)
22	4A	124/145 (85%)	33 (26%)	2 (1%)
37	2A	105/188 (55%)	22 (20%)	3 (2%)
All	All	461/701 (65%)	132 (28%)	23 (4%)

All (132) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	8	U
1	A	9	U
1	A	10	C
1	A	11	C
1	A	12	U
1	A	13	U
1	A	15	A
1	A	20	U
1	A	21	U
1	A	22	C
1	A	25	G
1	A	29	A
1	A	30	C
1	A	31	C
1	A	32	C
1	A	33	U
1	A	34	G
1	A	35	U
1	A	36	C
1	A	37	C
1	A	39	U
1	A	41	U
1	A	42	U
1	A	44	U
1	A	46	U
1	A	47	C
1	A	48	C
1	A	100	C
1	A	102	U
1	A	103	A
1	A	106	A
1	A	108	G
1	A	110	G

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	111	A
1	A	116	G
1	A	117	G
1	A	118	A
2	5A	8	G
2	5A	10	U
2	5A	17	U
2	5A	21	A
2	5A	22	U
2	5A	24	G
2	5A	25	C
2	5A	26	A
2	5A	28	A
2	5A	35	U
2	5A	36	C
2	5A	40	U
2	5A	45	C
2	5A	52	U
2	5A	68	C
2	5A	69	A
2	5A	78	U
2	5A	79	C
2	5A	80	U
2	5A	83	A
2	5A	88	A
2	5A	89	U
2	5A	90	U
2	5A	92	U
2	5A	93	U
2	5A	94	U
2	5A	95	G
2	5A	96	A
2	5A	97	G
2	5A	109	G
14	6A	36	A
14	6A	46	G
14	6A	47	A
14	6A	48	A
14	6A	49	G
14	6A	51	U
14	6A	77	C
14	6A	78	A

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
14	6A	103	U
14	6A	104	U
22	4A	2	G
22	4A	25	A
22	4A	26	G
22	4A	37	C
22	4A	38	U
22	4A	39	A
22	4A	40	U
22	4A	45	G
22	4A	53	U
22	4A	54	A
22	4A	62	U
22	4A	71	U
22	4A	73	U
22	4A	74	C
22	4A	75	C
22	4A	76	C
22	4A	84	C
22	4A	85	G
22	4A	90	G
22	4A	100	A
22	4A	103	A
22	4A	109	G
22	4A	114	U
22	4A	115	G
22	4A	118	A
22	4A	119	A
22	4A	120	U
22	4A	121	U
22	4A	124	U
22	4A	125	G
22	4A	126	A
22	4A	127	C
22	4A	144	G
37	2A	31	G
37	2A	37	U
37	2A	40	C
37	2A	45	C
37	2A	47	U
37	2A	51	A
37	2A	65	U

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
37	2A	112	G
37	2A	143	A
37	2A	147	G
37	2A	152	G
37	2A	153	A
37	2A	154	C
37	2A	156	U
37	2A	157	G
37	2A	164	C
37	2A	165	A
37	2A	168	A
37	2A	169	C
37	2A	177	A
37	2A	178	A
37	2A	179	C

All (23) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	9	U
1	A	21	U
1	A	28	G
1	A	33	U
1	A	35	U
1	A	36	C
1	A	38	C
1	A	40	U
1	A	41	U
1	A	99	C
1	A	116	G
1	A	117	G
2	5A	67	A
2	5A	78	U
2	5A	79	C
2	5A	94	U
2	5A	96	A
14	6A	77	C
22	4A	99	C
22	4A	114	U
37	2A	156	U
37	2A	164	C
37	2A	168	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 5 ligands modelled in this entry, 3 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	$\#\|Z\| > 2$
51	GTP	5C	1500	52	26,34,34	1.11	2 (7%)	32,54,54	1.53	7 (21%)
50	IHP	5B	3000	-	36,36,36	0.76	0	54,60,60	1.13	2 (3%)

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
51	GTP	5C	1500	52	-	8/18/38/38	0/3/3/3
50	IHP	5B	3000	-	-	3/30/54/54	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
51	5C	1500	GTP	C5-C6	-3.96	1.39	1.47
51	5C	1500	GTP	C2-N3	2.20	1.38	1.33

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
51	5C	1500	GTP	C5-C6-N1	3.27	119.73	113.95
51	5C	1500	GTP	PB-O3B-PG	-3.15	122.02	132.83
51	5C	1500	GTP	PA-O3A-PB	-3.11	122.14	132.83
51	5C	1500	GTP	C8-N7-C5	3.09	108.87	102.99
51	5C	1500	GTP	C2-N1-C6	-2.91	119.73	125.10
51	5C	1500	GTP	C3'-C2'-C1'	2.79	105.18	100.98
50	5B	3000	IHP	C5-C4-C3	2.47	115.82	110.41
50	5B	3000	IHP	C6-C5-C4	2.25	115.35	110.41
51	5C	1500	GTP	O6-C6-C5	-2.10	120.28	124.37

There are no chirality outliers.

All (11) torsion outliers are listed below:

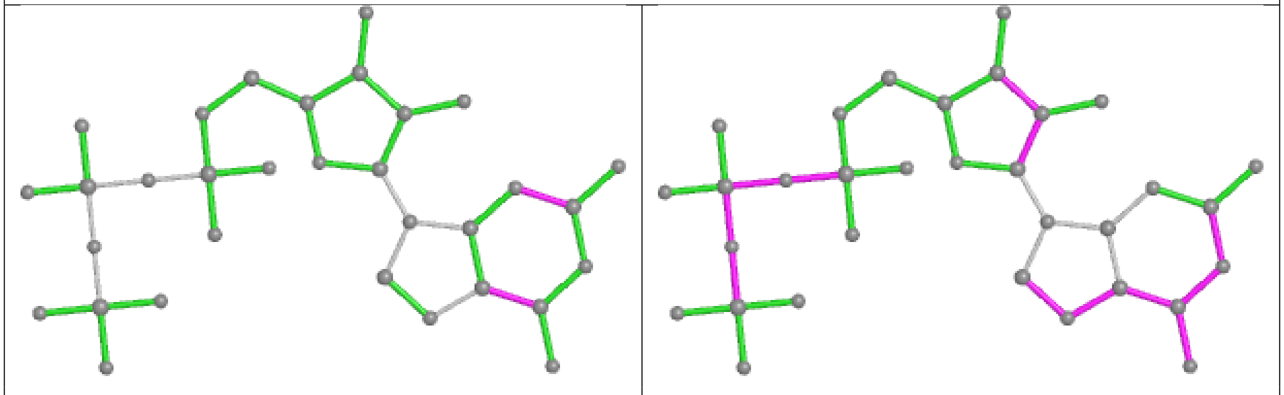
Mol	Chain	Res	Type	Atoms
50	5B	3000	IHP	C4-O14-P4-O44
51	5C	1500	GTP	C5'-O5'-PA-O1A
51	5C	1500	GTP	C5'-O5'-PA-O2A
51	5C	1500	GTP	PB-O3A-PA-O5'
51	5C	1500	GTP	C4'-C5'-O5'-PA
51	5C	1500	GTP	PB-O3B-PG-O1G
51	5C	1500	GTP	PB-O3B-PG-O2G
51	5C	1500	GTP	PB-O3B-PG-O3G
50	5B	3000	IHP	C1-O11-P1-O31
50	5B	3000	IHP	C5-O15-P5-O35
51	5C	1500	GTP	C5'-O5'-PA-O3A

There are no ring outliers.

No monomer is involved in short contacts.

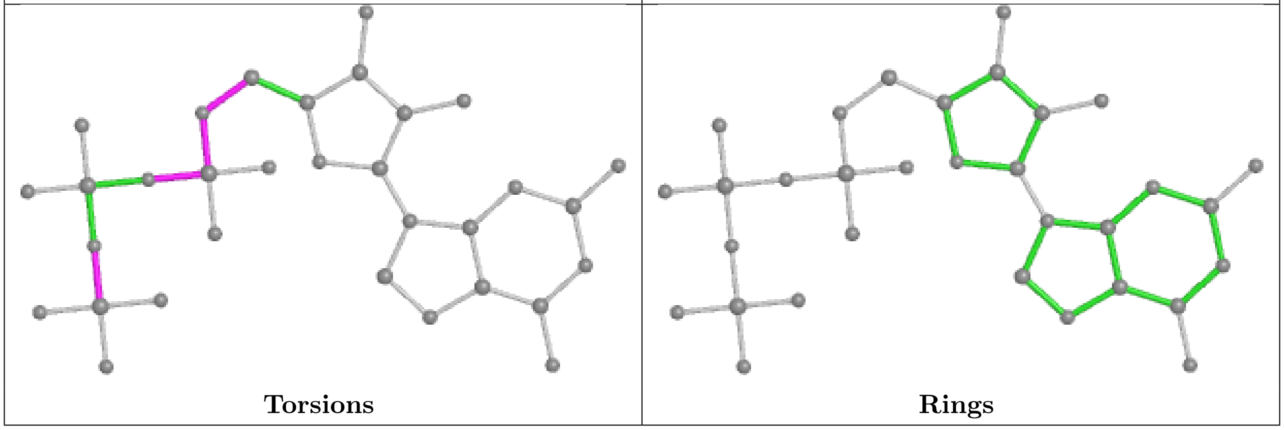
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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 GTP 5C 1500



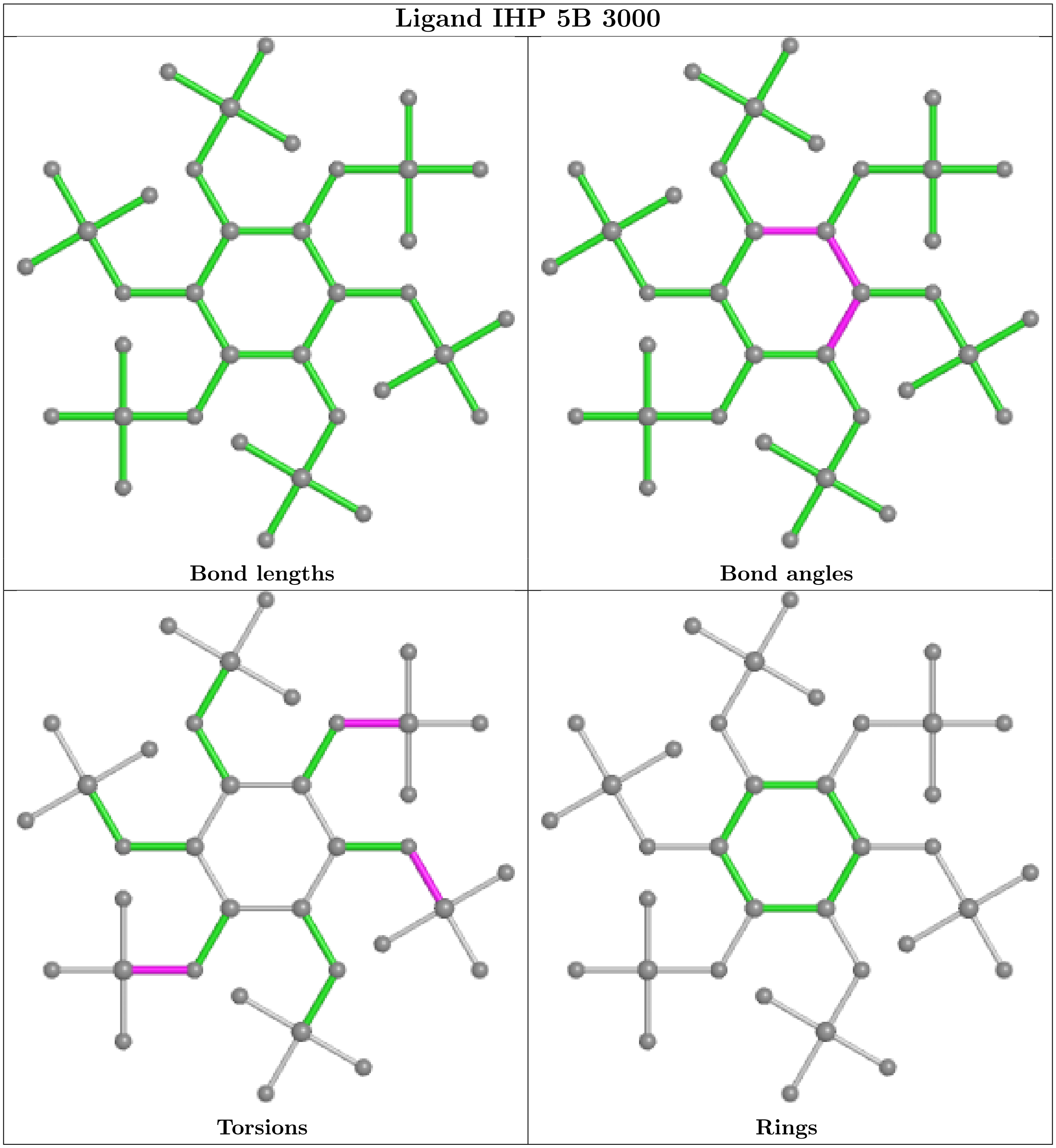
Bond lengths

Bond angles



Torsions

Rings



Ligand IHP 5B 3000

Bond lengths

Bond angles

Torsions

Rings

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.