



## wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Dec 13, 2016 – 05:27 PM EST

PDB ID : 5GMK  
EMDB ID: : EMD-9525  
Title : Cryo-EM structure of the Catalytic Step I spliceosome (C complex) at 3.4 angstrom resolution  
Authors : Wan, R.; Yan, C.; Bai, R.; Huang, G.; Shi, Y.  
Deposited on : 2016-07-14  
Resolution : 3.40 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/EMValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

MolProbity : 4.02b-467  
Mogul : 1.7.1 (RC1), CSD as537be (2016)  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20028442

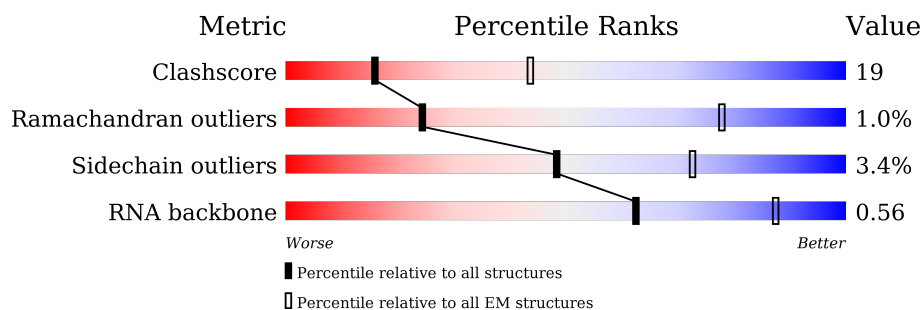
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.40 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.




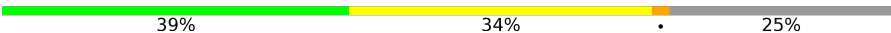
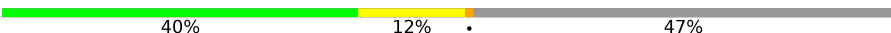






















Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686
RNA backbone	3027	244

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	2413	49% 29% . 21%
2	C	1008	54% 32% . 13%
3	D	214	34% 11% 9% . 45%
4	E	112	32% 35% 25% 8%
5	L	1175	. . 5% 89%
6	M	29	38% 34% 28%
7	B	13	31% 31% 38%
8	N	15	40% 20% 33% 7%


*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	J	135	
10	O	451	
11	P	379	
12	Q	364	
13	R	339	
14	S	175	
15	T	157	
16	Z	577	
17	c	587	
18	d	687	
19	F	278	
20	G	179	
21	H	235	
22	I	215	
23	v	859	
24	n	455	
25	o	503	
25	p	503	
25	q	503	
25	r	503	
26	t	175	
27	k	196	
27	s	196	
28	i	94	
28	u	94	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
29	h	86	 78% . 19%
29	w	86	 78% . 19%
30	j	77	 86% . . 10%
30	x	77	 86% . . 10%
31	l	101	 77% . 19%
31	y	101	 77% . 19%
32	m	146	 52% . 44%
32	z	146	 52% . 44%
33	e	110	 54% . . 41%
33	g	110	 79% 5% . 15%
34	a	111	 68% 5% 27%
35	b	238	 48% 7% . 43%

## 2 Entry composition

There are 38 unique types of molecules in this entry. The entry contains 75273 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a protein called Pre-mRNA-splicing factor 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1910	Total	C	N	O	S	0	0
			15775	10142	2709	2867	57		

- Molecule 2 is a protein called Pre-mRNA-splicing factor SNU114.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	878	Total	C	N	O	S	0	0
			7019	4529	1166	1295	29		

- Molecule 3 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	117	Total	C	N	O	P	0	0
			2465	1104	414	830	117		

- Molecule 4 is a RNA chain called U6 snRNA.

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

- Molecule 5 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	L	127	Total	C	N	O	P	0	0
			2673	1197	445	904	127		

- Molecule 6 is a RNA chain called Intron\_BPS.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	M	29	Total	C	N	O	P	0	0
			608	274	101	204	29		

- Molecule 7 is a RNA chain called 5'-Exon.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	B	13	Total	C	N	O	P	0	0
			275	124	47	91	13		

- Molecule 8 is a RNA chain called 5'-Splicing Site.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	N	15	Total	C	N	O	P	0	0
			312	140	45	112	15		

- Molecule 9 is a protein called Pre-mRNA-splicing factor CWC21.

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

- Molecule 10 is a protein called Pre-mRNA-splicing factor PRP46.

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

- Molecule 11 is a protein called Pre-mRNA-processing protein 45.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	P	201	Total	C	N	O	S	0	0
			1583	988	290	298	7		

- Molecule 12 is a protein called Pre-mRNA-splicing factor SLT11.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	Q	185	Total	C	N	O	S	0	0
			1472	930	256	271	15		

- Molecule 13 is a protein called Pre-mRNA-splicing factor CWC2.

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

- Molecule 14 is a protein called Pre-mRNA-splicing factor CWC15.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	S	69	Total	C	N	O	S	0	0
			560	351	112	96	1		

- Molecule 15 is a protein called Pre-mRNA-splicing factor BUD31.

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

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

Mol	Chain	Residues	Atoms					AltConf	Trace
16	Z	447	Total	C	N	O	S	0	0
			3651	2343	602	688	18		

- Molecule 17 is a protein called Pre-mRNA-splicing factor CEF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	c	436	Total	C	N	O	S	0	0
			2971	1841	549	573	8		

- Molecule 18 is a protein called Pre-mRNA-splicing factor CLF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	d	549	Total	C	N	O	S	0	0
			3590	2232	675	675	8		

- Molecule 19 is a protein called Protein CWC16.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	F	115	Total	C	N	O	S	0	0
			937	592	167	169	9		

- Molecule 20 is a protein called Pre-mRNA-splicing factor CWC25.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	G	41	Total	C	N	O	S	0	0
			342	215	63	63	1		

- Molecule 21 is a protein called Pre-mRNA-splicing factor ISY1.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	H	95	Total	C	N	O	S	0	0
			810	506	152	151	1		

- Molecule 22 is a protein called Pre-mRNA-splicing factor SYF2.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	I	102	Total	C	N	O	S	0	0
			822	504	152	165	1		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
23	v	673	Total	C	N	O	S	0	0
			3580	2190	683	706	1		

- Molecule 24 is a protein called Pre-mRNA-processing factor 17.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	n	299	Total	C	N	O	S	0	0
			1890	1175	340	369	6		

- Molecule 25 is a protein called Pre-mRNA-processing factor 19.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	o	126	Total	C	N	O	S	0	0
			830	525	134	169	2		
25	p	128	Total	C	N	O	S	0	0
			843	532	136	173	2		
25	q	387	Total	C	N	O	S	0	0
			2345	1471	402	464	8		
25	r	125	Total	C	N	O	S	0	0
			823	521	133	167	2		

- Molecule 26 is a protein called Pre-mRNA-splicing factor SNT309.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	t	156	Total	C	N	O	S	0	0
			926	585	160	180	1		

- Molecule 27 is a protein called Small nuclear ribonucleoprotein-associated protein B.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	k	80	Total	C	N	O	S	0	0
			631	403	114	111	3		
27	s	78	Total	C	N	O	S	0	0
			610	389	110	108	3		

- Molecule 28 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	i	75	Total	C	N	O	S	0	0
			575	379	92	101	3		
28	u	75	Total	C	N	O	S	0	0
			575	379	92	101	3		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
29	h	70	Total	C	N	O	S	0	0
			554	355	98	100	1		
29	w	70	Total	C	N	O	S	0	0
			554	355	98	100	1		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
30	j	69	Total	C	N	O	S	0	0
			529	337	93	97	2		
30	x	69	Total	C	N	O	S	0	0
			529	337	93	97	2		

- Molecule 31 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	l	82	Total	C	N	O	S	0	0
			625	399	109	115	2		
31	y	82	Total	C	N	O	S	0	0
			625	399	109	115	2		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
32	m	82	Total	C	N	O	S	0	0
			644	409	110	123	2		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					AltConf	Trace
32	z	82	Total	C	N	O	S	0	0
			644	409	110	123	2		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
33	g	94	Total	C	N	O	S	0	0
			741	477	141	119	4		
33	e	65	Total	C	N	O	S	0	0
			528	340	102	84	2		

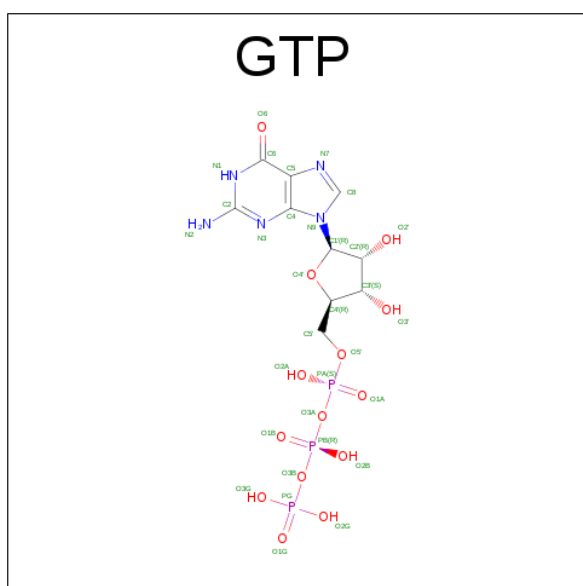
- Molecule 34 is a protein called U2 small nuclear ribonucleoprotein B''.

Mol	Chain	Residues	Atoms				AltConf	Trace
34	a	81	Total	C	N	O	0	0
			513	332	89	92		

- Molecule 35 is a protein called U2 small nuclear ribonucleoprotein A'.

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

- Molecule 36 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula:  $C_{10}H_{16}N_5O_{14}P_3$ ).



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

- Molecule 37 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
37	C	1	Total	Mg	0
			1	1	
37	E	5	Total	Mg	0
			5	5	

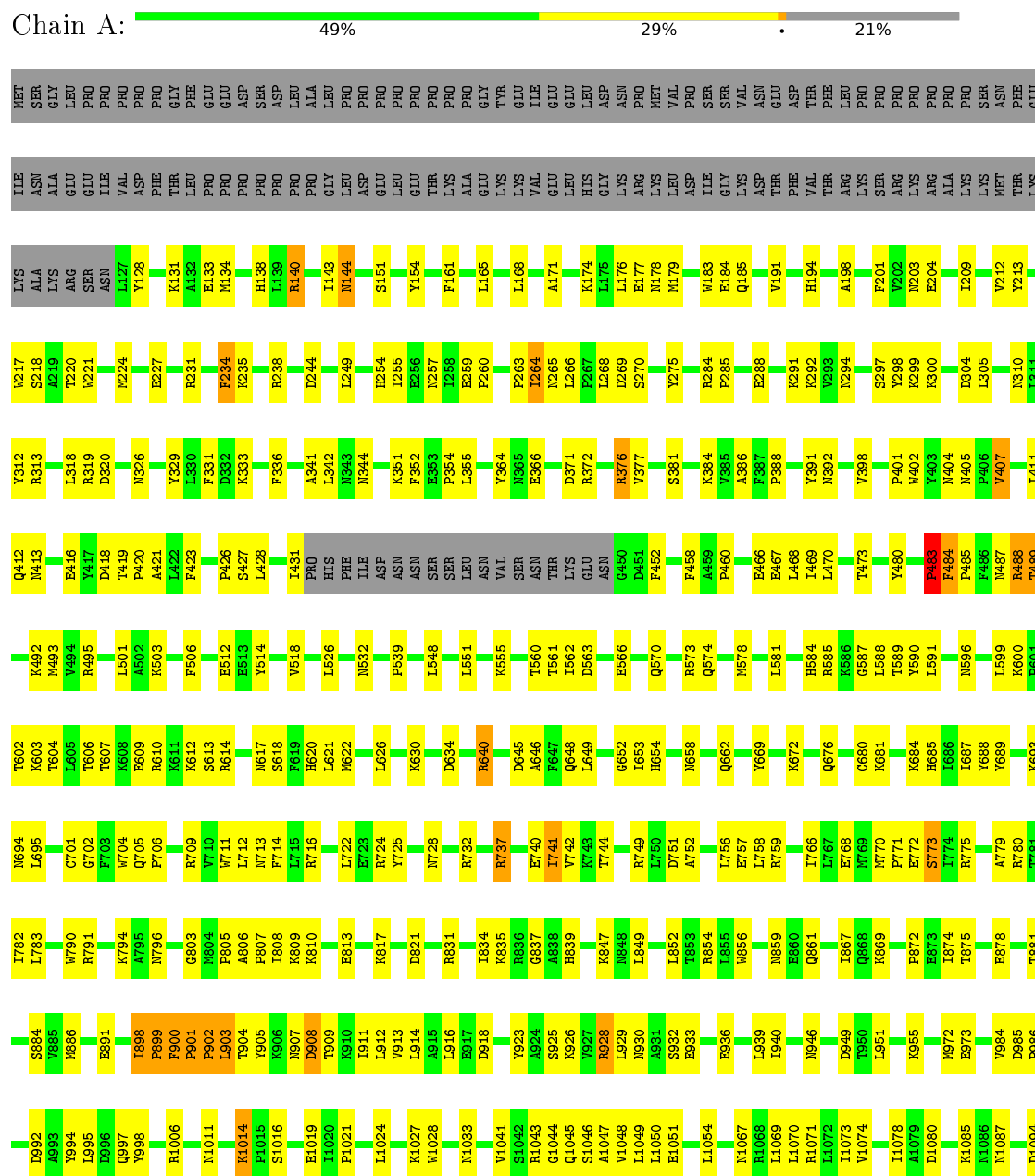
- Molecule 38 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
38	R	1	Total	Zn	0
			1	1	
38	Q	2	Total	Zn	0
			2	2	
38	T	3	Total	Zn	0
			3	3	
38	F	1	Total	Zn	0
			1	1	

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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-splicing factor 8



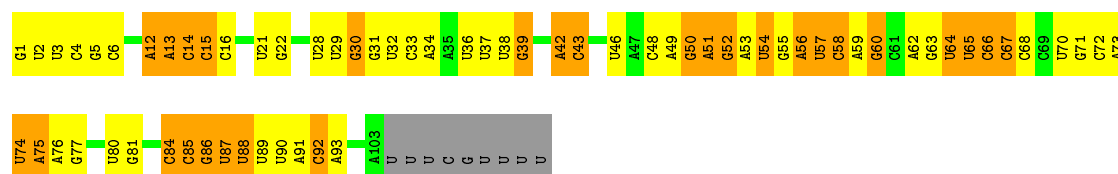
[illegible]

- Molecule 2: Pre-mRNA-splicing factor SNU114

G924	S824	C654	ASP	R441	V245	L153	GLU	MET
I925	K746	L655	GLU	T346	T246	V154	GLU	GLU
G926	G929	L656	ASP	P445	P247	V147	GLY	GLY
M927	N830	G657	E530	F446	L348	S157	ILE	ASP
G928	K749	D658	E533	E447	K349	H158	SER	ASP
M933	H833	L659	E534	N451	Y353	Q251	LEU	LEU
H934	R752	A661	T534	K452	R360	L252	E87	PHE
K935	T753	S662	S536	H455	T361	Q255	P89	GLU
L945	E754	K665	V539	L458	K362	F265	G71	GLY
P952	G755	E667	T542	K461	F363	N268	E75	ASN
K955	R758	L668	G543	K461	E365	K269	V76	LEU
P956	S759	K669	L544	S462	N366	L270	L77	ILE
A957	L760	L670	L545	T463	D271	D271	M78	GLY
M968	R763	S671	R548	E465	R272	R176	E79	VAL
R972	N764	T675	P552	L468	T372	V177	T80	ASP
R973	V765	V676	V553	H471	F376	L286	N82	PHE
K974	S767	S680	B557	V472	L378	N289	P89	ASP
G975	F768	C681	E560	L473	S379	H290	K99	GLU
ILE	D885	S682	G561	D477	P380	I291	L100	GLU
SER	T867	L686	N683	H477	L384	A293	Q101	SER
THR	H868	G771	V562	L485	F385	N294	E102	VAL
GLY	H869	PHE	L564	V486	S386	I295	H103	LEU
PHE	V876	ALA	L567	R487	S387	N296	T104	ASP
SER	R883	ILE	G491	K391	F304	K206	I105	GLU
ASN	R884	PRO	V571	L492	S305	S207	Q108	GLU
GLY	G885	VAL	ASN	L493	K393	R208	L109	GLN
PRO	S886	SER	A575	V499	K397	N209	K110	GLN
THR	R887	SER	T577	R500	I311	I210	T116	THR
LEU	H889	ILE	Y578	L501	N407	F212	R117	ASN
LVS	G896	S696	P592	D503	A314	L213	N119	PHE
THR	T897	R697	E700	Q506	S315	D214	R120	GLU
ILE	G898	E701	V602	R510	T316	A215	M123	GLY
ALA	G794	L702	L607	GLN	G319	H218	M126	SER
LEU	P906	L703	B613	LVS	T321	N220	ASN	ASN
THR	V907	P704	E614	Q418	F221	K222	E131	ASN
ALA	N908	G705	E614	ARG	P419	K324	R132	GLU
GLN	E910	L706	S626	GLN	F420	K325	I133	ILE
LEU	S911	S707	K627	LEU	E326	T233	I134	GLU
GLU	A912	A712	V628	HIS	L425	A231	N135	SER
ASN	G914	E713	V629	ILE	F429	S232	V136	ARG
GLY	E915	T727	T634	LVS	S335	L234	G137	GLN
LEU	E910	E310	G646	THR	S338	T236	L142	LEU
PRO	R919	P812	N647	THR	K340	V238	H144	GLY
LVS	T922	Q817	M652	SER	I341	I239	S144	SER
LVS	H923	T732	P852	ASN	D342	D240	T152	LVS

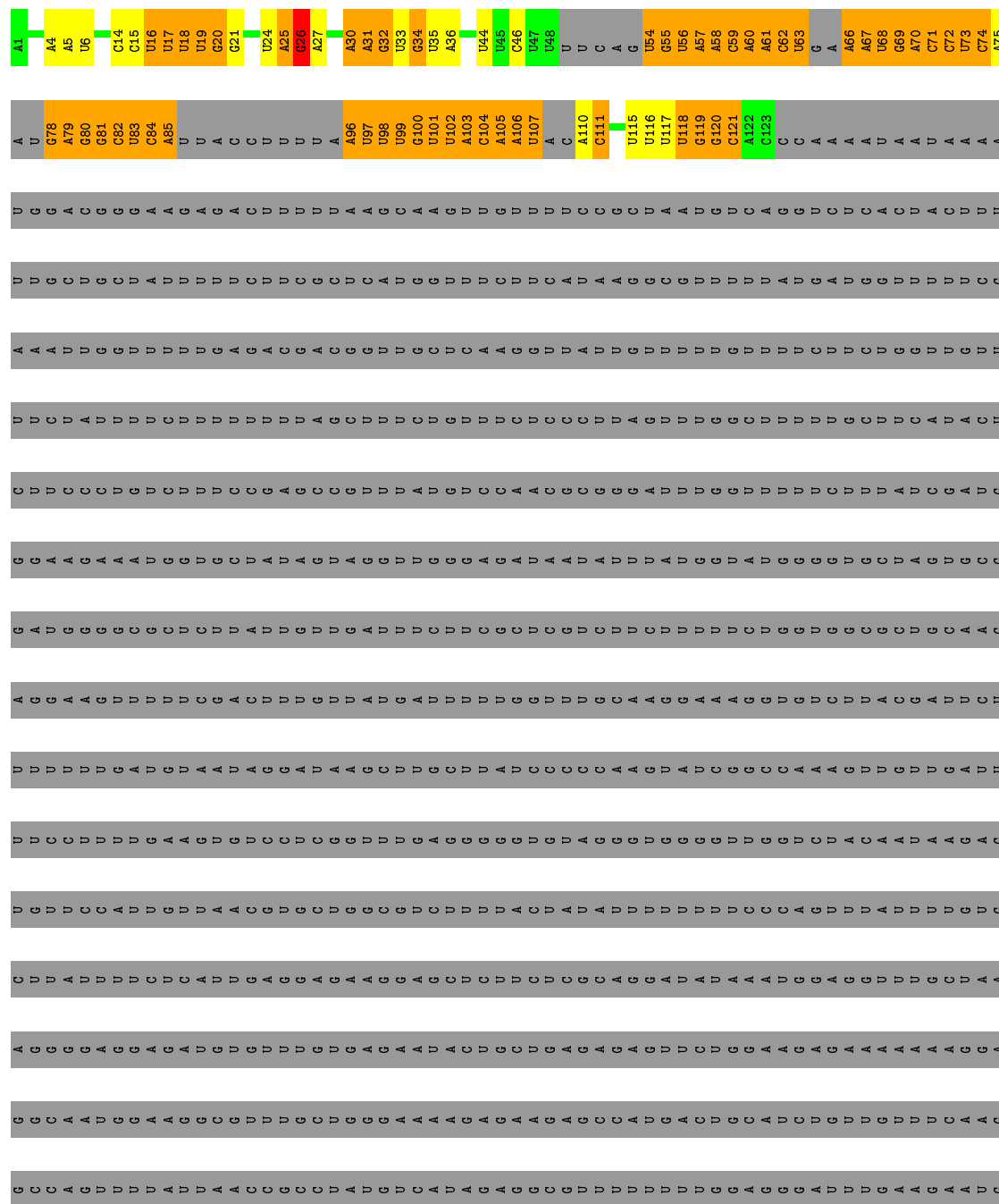
[illegible]



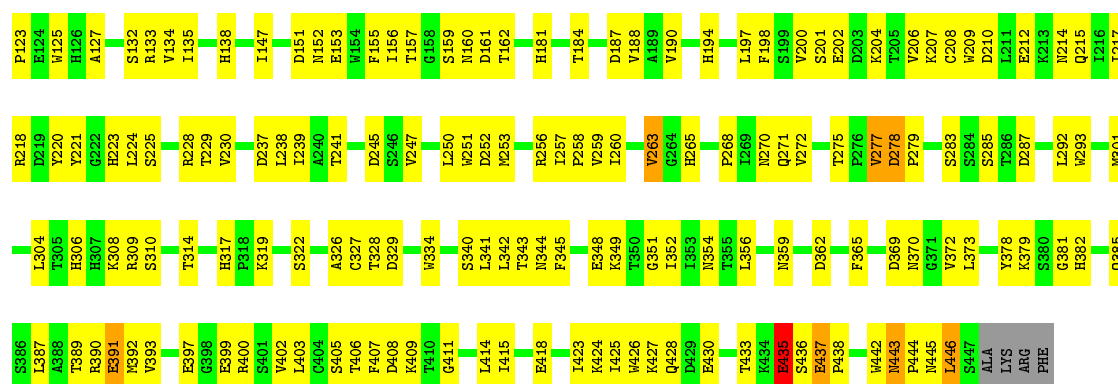



• Molecule 5: U2 snRNA

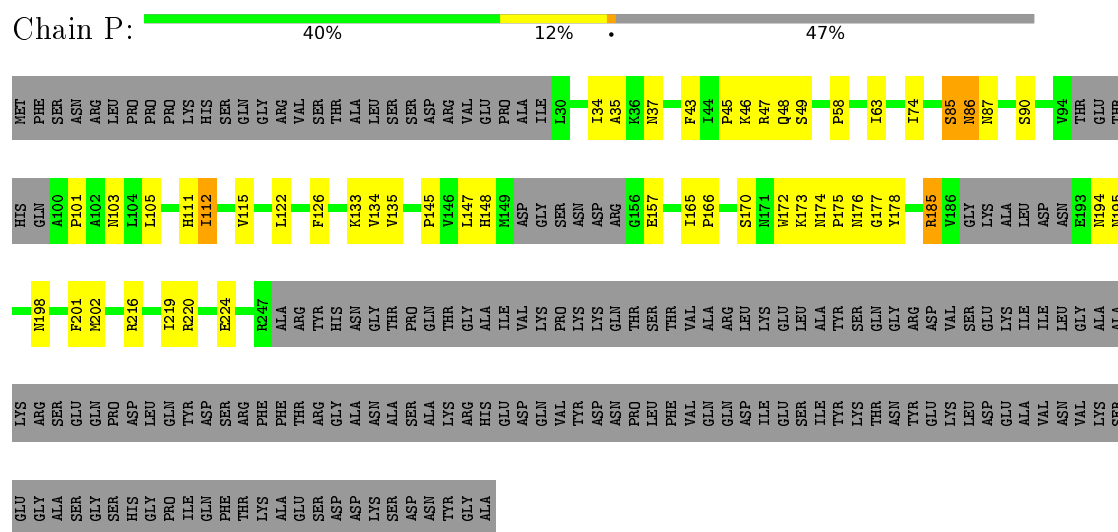
Chain L: . . 5% 89%



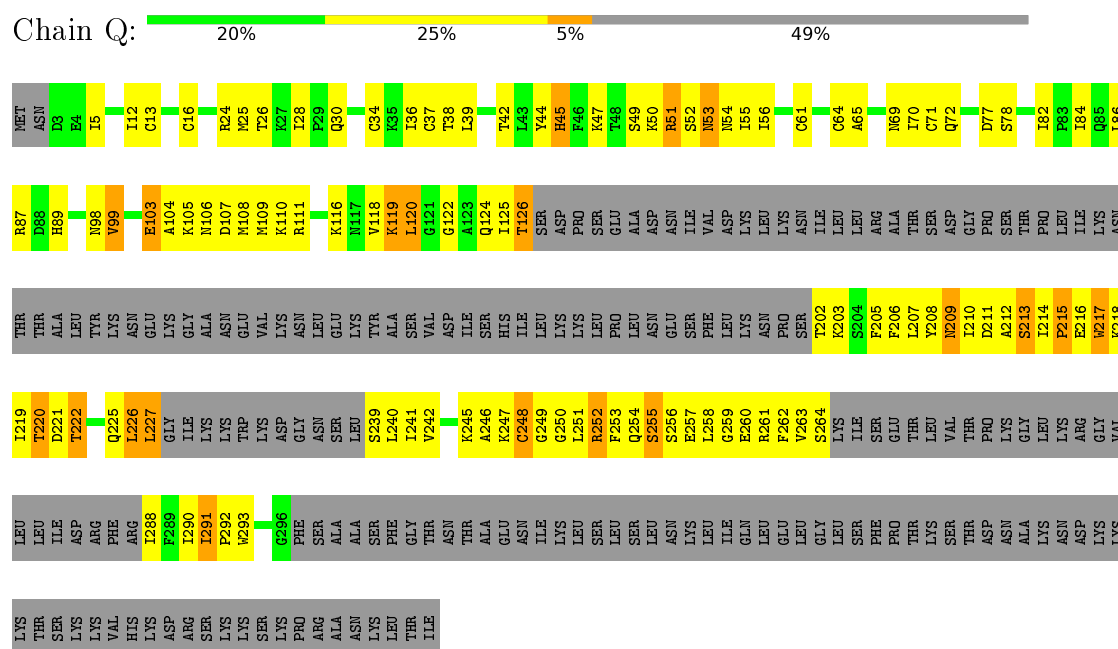




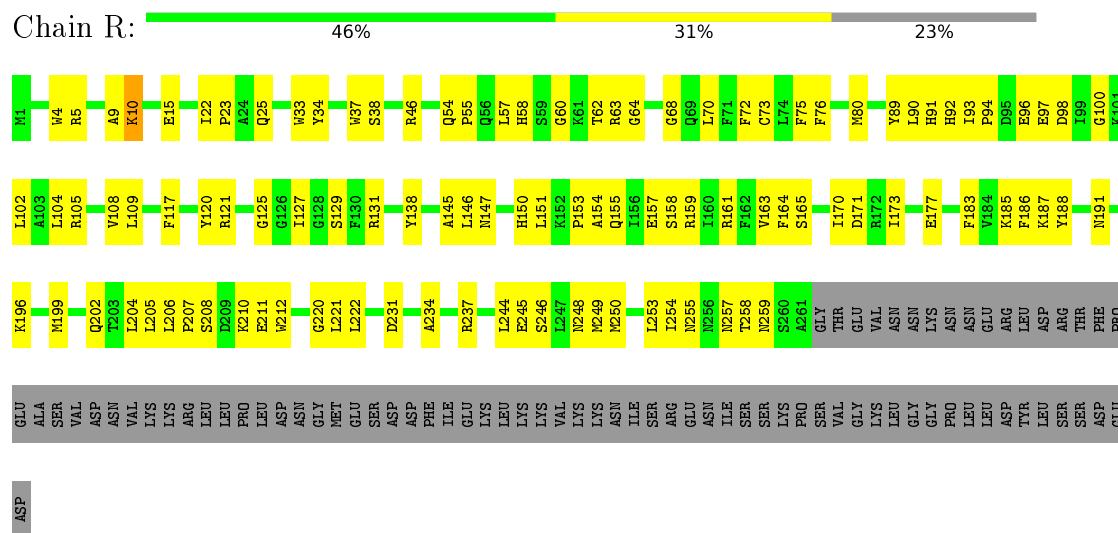
• Molecule 11: Pre-mRNA-processing protein 45



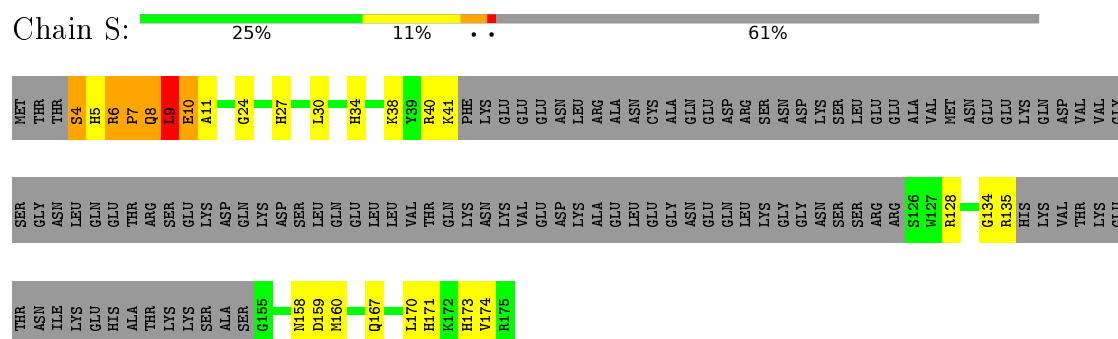
• Molecule 12: Pre-mRNA-splicing factor SLT11



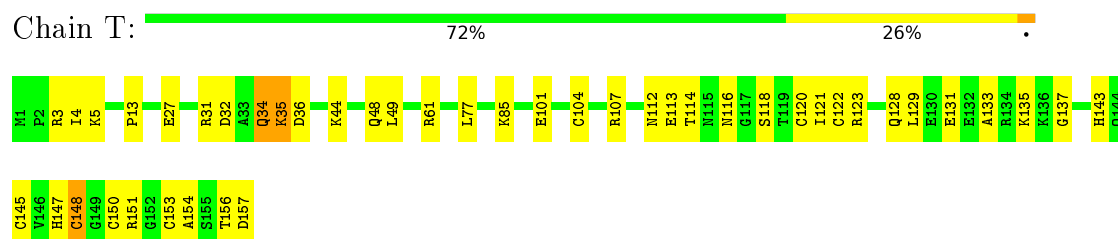
- Molecule 13: Pre-mRNA-splicing factor CWC2



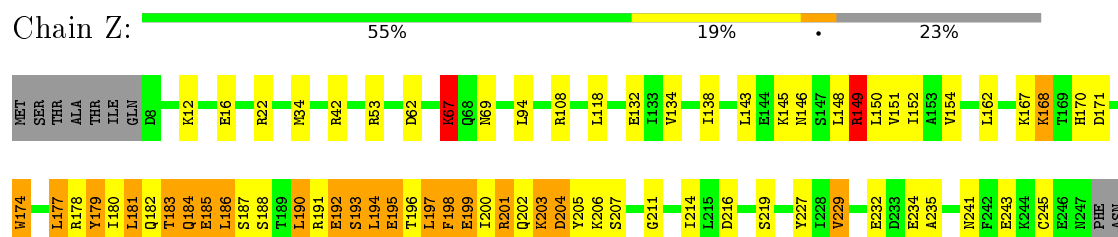
- Molecule 14: Pre-mRNA-splicing factor CWC15

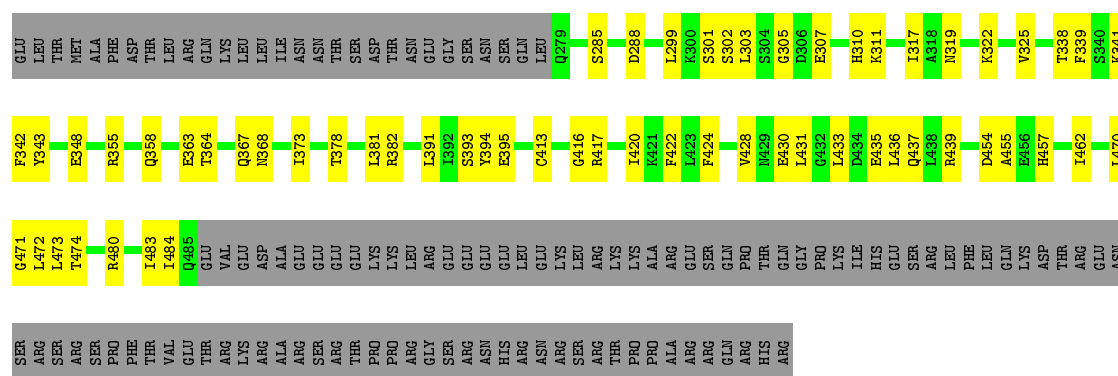


- Molecule 15: Pre-mRNA-splicing factor BUD31



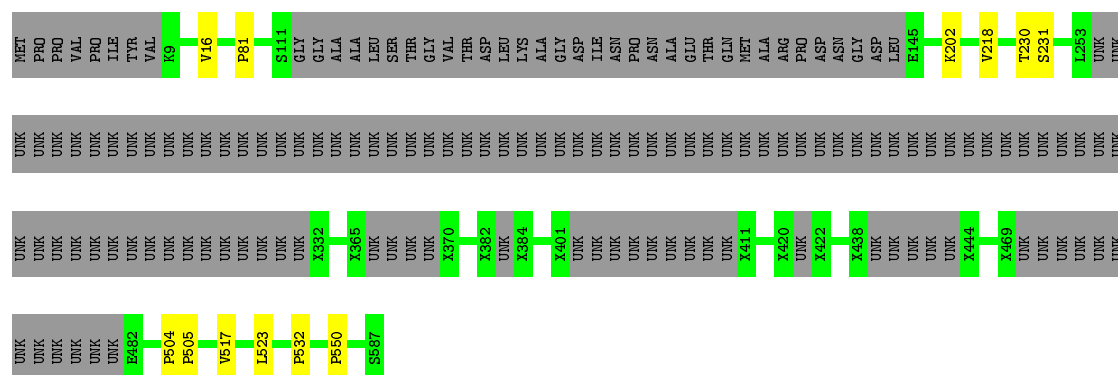
- Molecule 16: Pre-mRNA-splicing factor CWC22





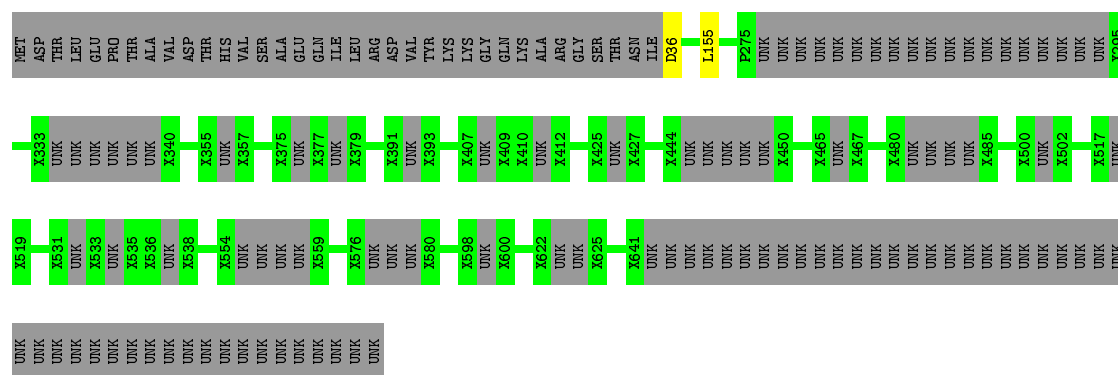
- Molecule 17: Pre-mRNA-splicing factor CEF1

Chain c:  72% . 26%



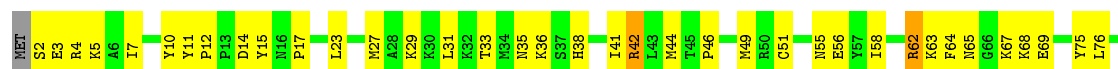
- Molecule 18: Pre-mRNA-splicing factor CLF1

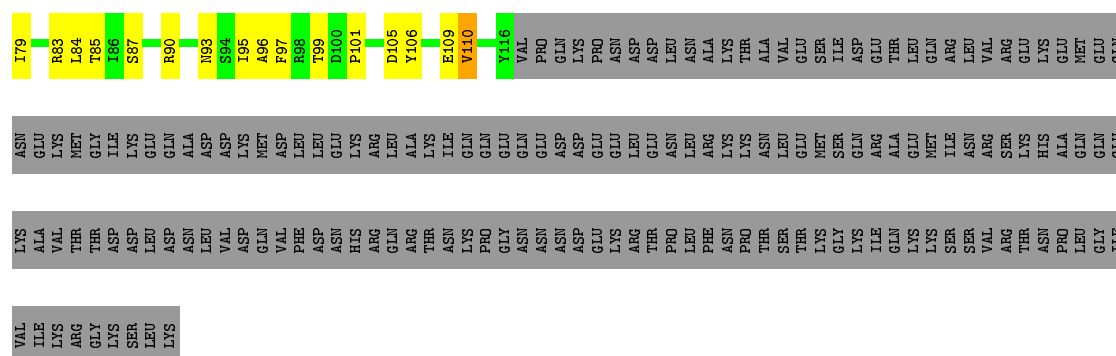
Chain d:  80% 20%



- Molecule 19: Protein CWC16

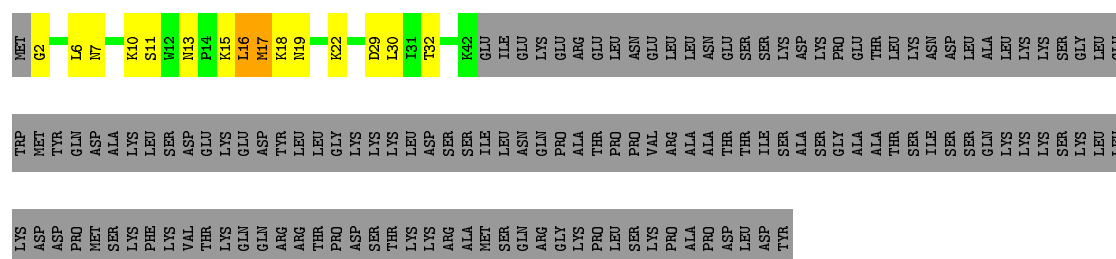
Chain F:  22% 18% . 59%





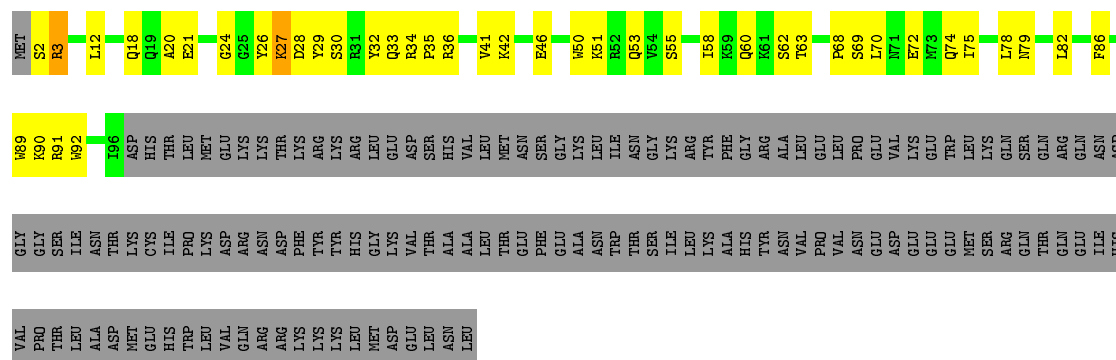
• Molecule 20: Pre-mRNA-splicing factor CWC25

Chain G: 15% 7% 77%



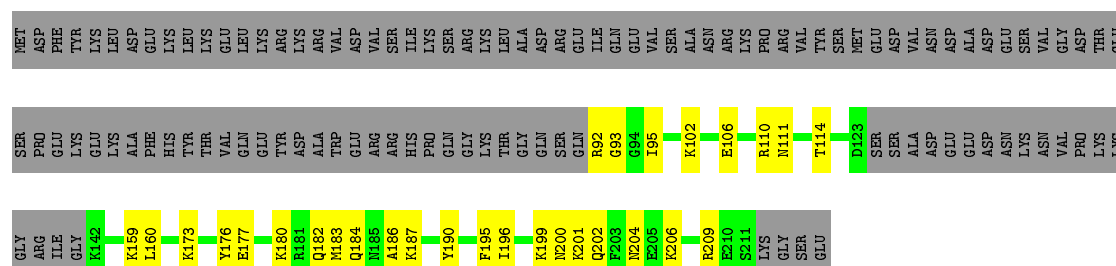
• Molecule 21: Pre-mRNA-splicing factor ISY1

Chain H: 23% 17% 60%



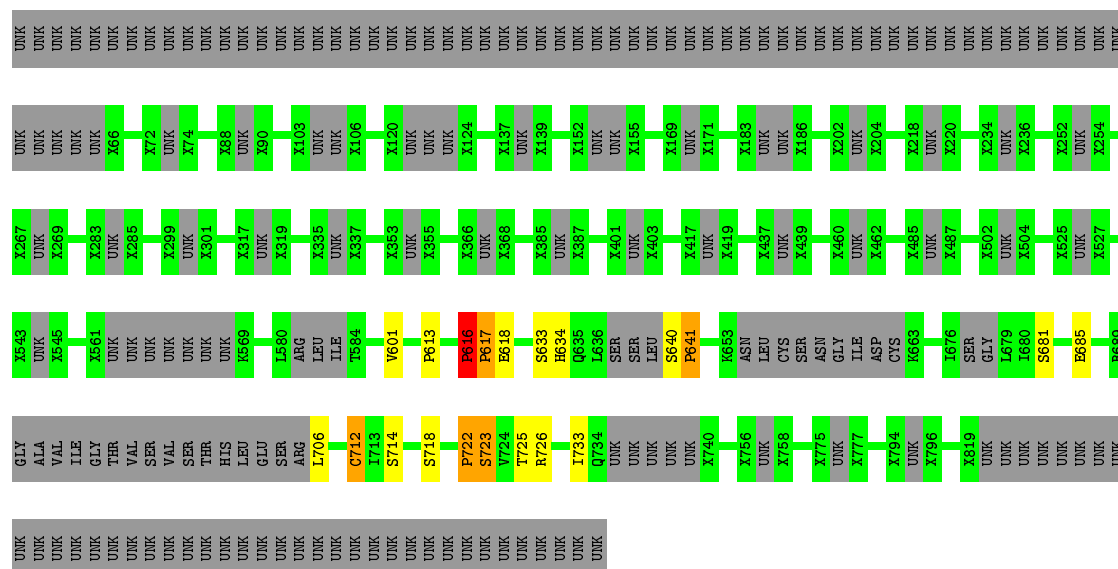
• Molecule 22: Pre-mRNA-splicing factor SYF2

Chain I: 34% 13% 53%



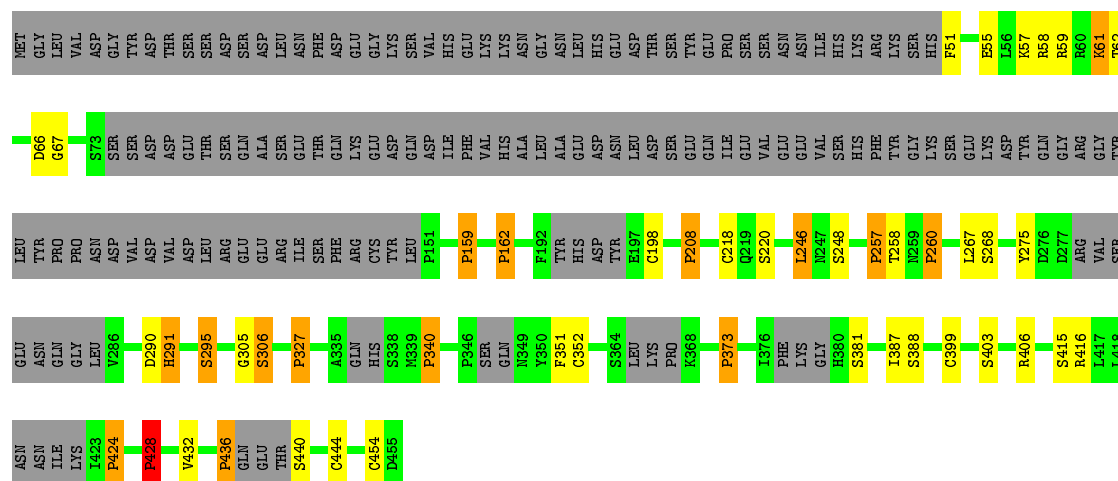
- Molecule 23: Pre-mRNA-splicing factor SYF1

Chain v:  76% .. 22%



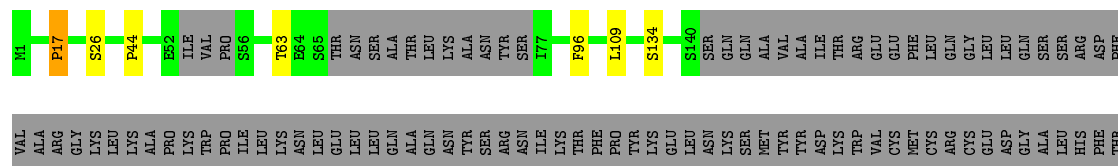
- Molecule 24: Pre-mRNA-processing factor 17

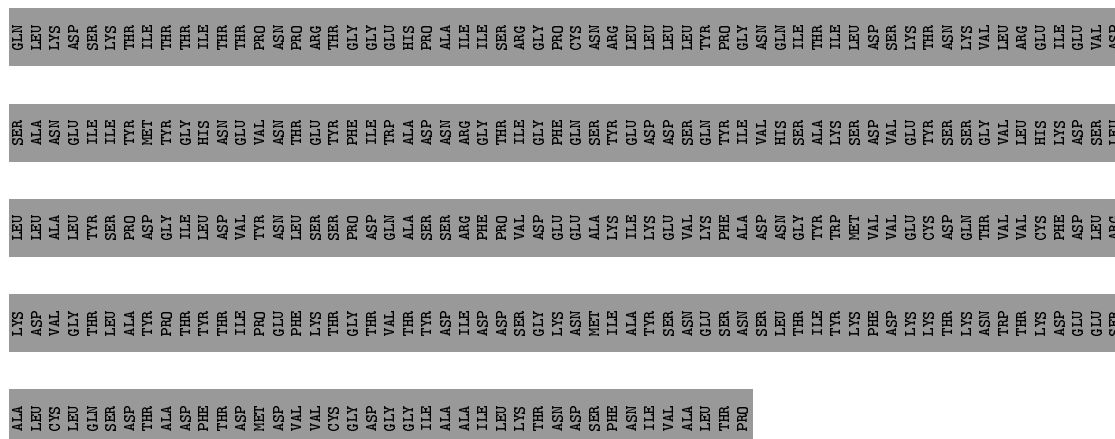
Chain n:  55% 7% • 34%



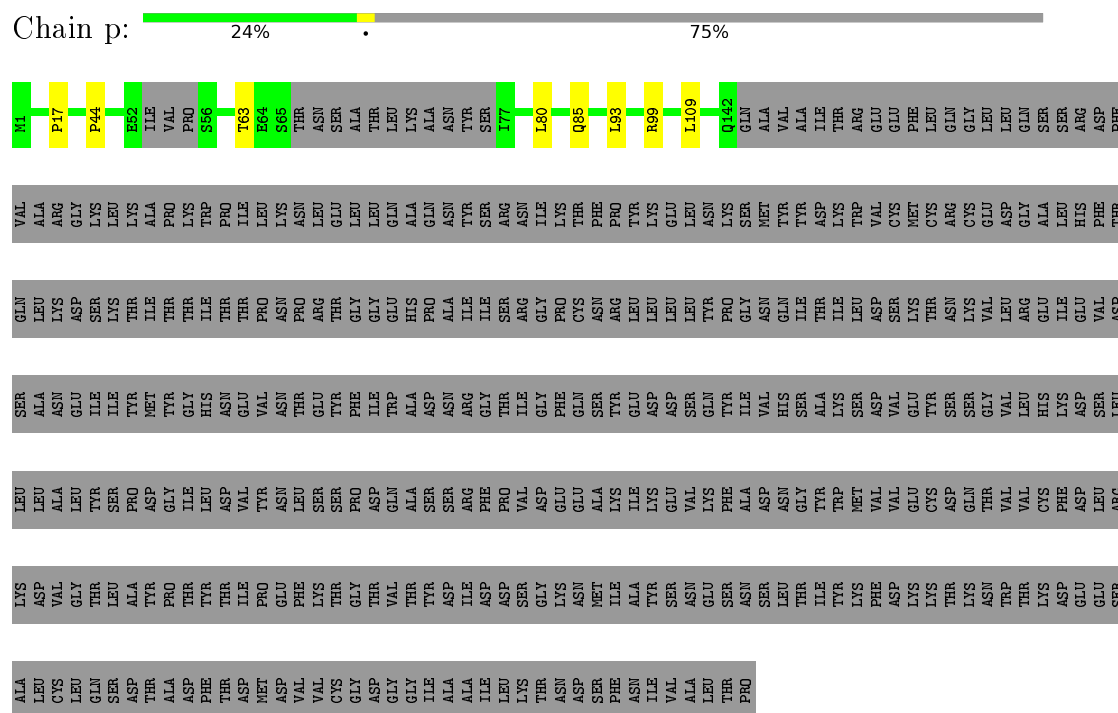
- Molecule 25: Pre-mRNA-processing factor 19

Chain o:  24% 1% 75%

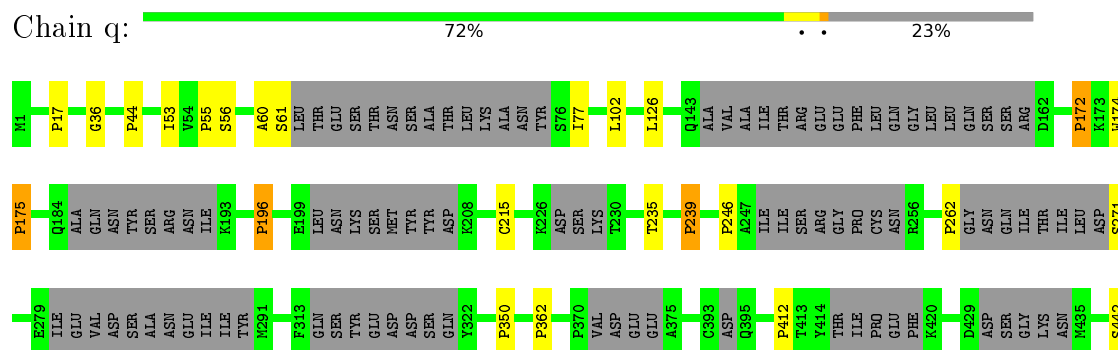


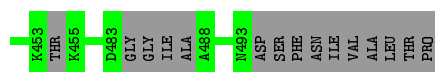


- Molecule 25: Pre-mRNA-processing factor 19



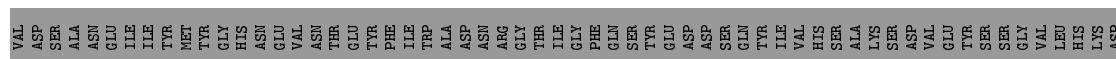
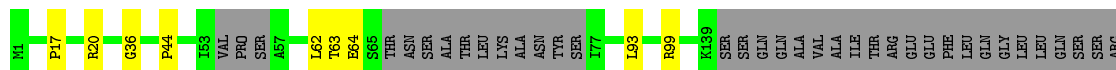
- Molecule 25: Pre-mRNA-processing factor 19





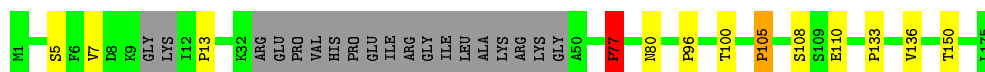
- Molecule 25: Pre-mRNA-processing factor 19

Chain r: 23% 75%



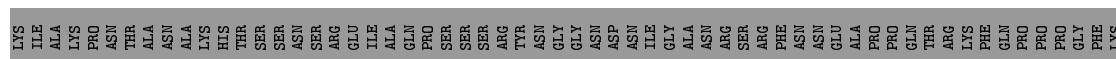
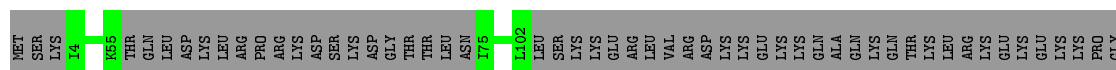
- Molecule 26: Pre-mRNA-splicing factor SNT309

Chain t: 82% 6% 11%



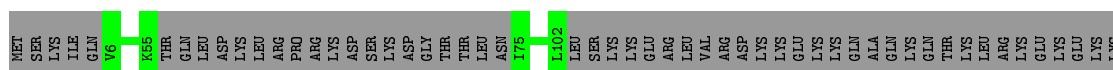
- Molecule 27: Small nuclear ribonucleoprotein-associated protein B

Chain k: 41% 59%



- Molecule 27: Small nuclear ribonucleoprotein-associated protein B

Chain s: 40% 60%



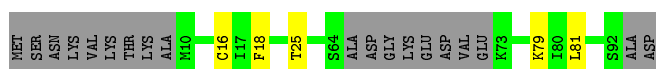
- Molecule 28: Small nuclear ribonucleoprotein E

Chain i: 74% 5% 20%



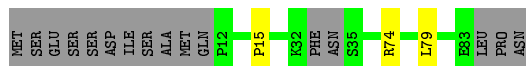
- Molecule 28: Small nuclear ribonucleoprotein E

Chain u: 74% 5% 20%



- Molecule 29: Small nuclear ribonucleoprotein F

Chain h: 78% • 19%



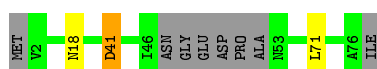
- Molecule 29: Small nuclear ribonucleoprotein F

Chain w: 78% • 19%



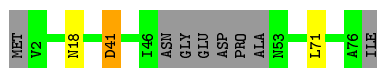
- Molecule 30: Small nuclear ribonucleoprotein G

Chain j: 86% • • 10%




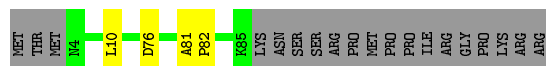
- Molecule 30: Small nuclear ribonucleoprotein G

Chain x: 86% • • 10%




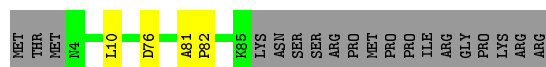
- Molecule 31: Small nuclear ribonucleoprotein Sm D3

Chain l:  77% 19%



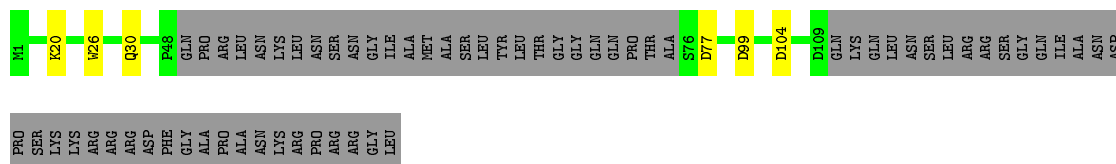
- Molecule 31: Small nuclear ribonucleoprotein Sm D3

Chain y:  77% 19%



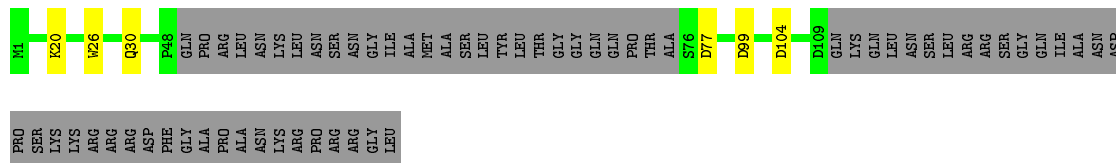
- Molecule 32: Small nuclear ribonucleoprotein Sm D1

Chain m:  52% 44%




- Molecule 32: Small nuclear ribonucleoprotein Sm D1

Chain z:  52% 44%



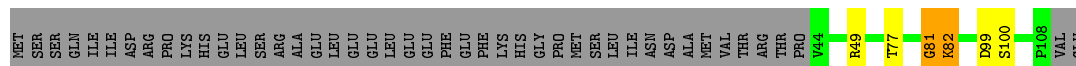
- Molecule 33: Small nuclear ribonucleoprotein Sm D2

Chain g:  79% 5% 15%



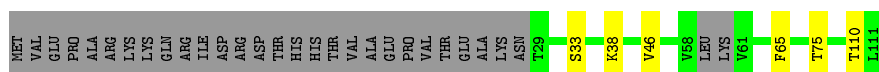
- Molecule 33: Small nuclear ribonucleoprotein Sm D2

Chain e:  54% 41%



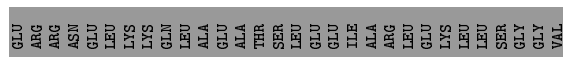
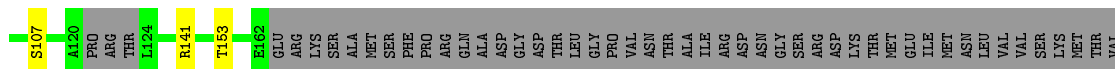
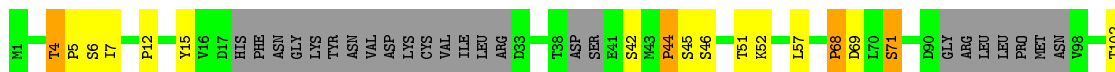
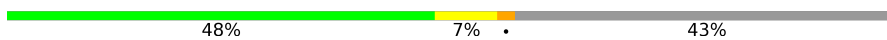
- Molecule 34: U2 small nuclear ribonucleoprotein B"

Chain a:  68% 5% 27%



- Molecule 35: U2 small nuclear ribonucleoprotein A'

Chain b:



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	161066	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	Not provided	Depositor

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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

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	$\# Z  > 2$	RMSZ	$\# Z  > 2$
1	A	0.48	3/16178 (0.0%)	0.60	6/21929 (0.0%)
10	O	0.57	0/2704	0.63	1/3676 (0.0%)
11	P	0.42	2/1604 (0.1%)	0.52	0/2160
12	Q	0.47	1/1496 (0.1%)	0.62	0/2014
13	R	0.44	0/2135	0.56	0/2871
14	S	0.46	1/574 (0.2%)	0.56	0/766
15	T	0.51	1/1315 (0.1%)	0.61	0/1759
16	Z	0.50	0/3712	0.85	7/5004 (0.1%)
17	c	0.41	1/2405 (0.0%)	0.50	0/3218
18	d	0.49	0/2107	0.54	1/2852 (0.0%)
19	F	0.37	0/955	0.55	0/1277
2	C	0.47	1/7168 (0.0%)	0.57	3/9707 (0.0%)
20	G	0.27	0/346	0.51	0/462
21	H	0.33	0/826	0.54	0/1108
22	I	0.34	0/826	0.47	0/1097
23	v	1.04	8/905 (0.9%)	0.76	6/1214 (0.5%)
24	n	1.48	19/1900 (1.0%)	0.89	15/2537 (0.6%)
25	o	0.40	0/835	0.53	0/1126
25	p	0.40	0/848	0.55	0/1143
25	q	0.44	0/2342	0.65	0/3139
25	r	0.39	0/828	0.54	1/1117 (0.1%)
26	t	0.42	0/924	0.56	2/1244 (0.2%)
27	k	0.37	0/636	0.61	0/856
27	s	0.37	0/615	0.61	0/829
28	i	0.42	0/585	0.62	0/795
28	u	0.42	0/585	0.62	0/795
29	h	0.44	0/564	0.65	1/761 (0.1%)
29	w	0.44	0/564	0.65	1/761 (0.1%)
3	D	0.73	0/2747	0.93	2/4267 (0.0%)
30	j	0.37	0/532	0.60	0/715
30	x	0.37	0/532	0.60	0/715
31	l	0.40	0/634	0.70	0/859

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >2	RMSZ	# Z  >2
31	y	0.40	0/634	0.70	0/859
32	m	0.41	0/649	0.61	0/880
32	z	0.41	0/649	0.61	0/880
33	e	0.43	0/535	0.67	2/717 (0.3%)
33	g	0.45	0/753	0.69	2/1013 (0.2%)
34	a	0.82	4/514 (0.8%)	1.32	2/686 (0.3%)
35	b	1.03	9/839 (1.1%)	1.65	11/1127 (1.0%)
4	E	0.76	1/2452 (0.0%)	0.96	3/3817 (0.1%)
5	L	0.94	33/2974 (1.1%)	1.51	88/4610 (1.9%)
6	M	0.73	2/678 (0.3%)	1.01	3/1051 (0.3%)
7	B	1.01	1/307 (0.3%)	0.95	0/475
8	N	1.27	5/346 (1.4%)	1.29	10/535 (1.9%)
9	J	0.41	0/191	0.57	0/254
All	All	0.59	92/72448 (0.1%)	0.75	167/99677 (0.2%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	8
10	O	0	6
11	P	0	1
12	Q	0	3
16	Z	0	5
17	c	0	1
2	C	0	2
20	G	0	1
24	n	0	4
30	j	0	1
30	x	0	1
31	l	0	2
31	y	0	2
33	e	0	2
33	g	0	2
All	All	0	41

The worst 5 of 92 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
24	n	399	CYS	CB-SG	-24.10	1.41	1.82

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
24	n	444	CYS	CB-SG	-22.82	1.43	1.82
24	n	454	CYS	CB-SG	-19.98	1.48	1.82
24	n	218	CYS	CB-SG	-19.56	1.49	1.82
24	n	352	CYS	CB-SG	-18.84	1.50	1.82

The worst 5 of 167 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	483	PRO	O-C-N	-28.27	77.47	122.70
6	M	502	C	O3'-P-O5'	-12.66	79.95	104.00
5	L	1110	U	C5-C4-O4	12.01	133.11	125.90
6	M	502	C	P-O3'-C3'	11.50	133.50	119.70
5	L	1107	C	N1-C2-O2	-10.12	112.83	118.90

There are no chirality outliers.

5 of 41 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1014	LYS	Peptide
1	A	257	ASN	Peptide
1	A	483	PRO	Mainchain
1	A	539	PRO	Peptide
1	A	773	SER	Peptide

## 5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	15775	0	15756	707	0
2	C	7019	0	7201	312	0
3	D	2465	0	1251	53	0
4	E	2192	0	1106	112	0
5	L	2673	0	1360	281	0
6	M	608	0	307	44	0
7	B	275	0	137	20	0
8	N	312	0	156	18	0
9	J	190	0	186	15	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	O	2646	0	2639	131	0
11	P	1583	0	1608	79	0
12	Q	1472	0	1485	173	0
13	R	2089	0	2053	116	0
14	S	560	0	545	68	0
15	T	1291	0	1312	44	0
16	Z	3651	0	3707	180	0
17	c	2971	0	2336	0	0
18	d	3590	0	2359	0	0
19	F	937	0	962	54	0
20	G	342	0	360	16	0
21	H	810	0	799	37	0
22	I	822	0	845	62	0
23	v	3580	0	1298	0	0
24	n	1890	0	1407	0	0
25	o	830	0	663	0	0
25	p	843	0	675	0	0
25	q	2345	0	1636	0	0
25	r	823	0	654	0	0
26	t	926	0	606	0	0
27	k	631	0	670	0	0
27	s	610	0	640	0	0
28	i	575	0	597	0	0
28	u	575	0	597	0	0
29	h	554	0	556	0	0
29	w	554	0	556	0	0
30	j	529	0	557	0	0
30	x	529	0	557	0	0
31	l	625	0	647	0	0
31	y	625	0	647	0	0
32	m	644	0	686	0	0
32	z	644	0	686	0	0
33	e	528	0	573	0	0
33	g	741	0	778	0	0
34	a	513	0	402	0	0
35	b	841	0	614	0	0
36	C	32	0	12	1	0
37	C	1	0	0	0	0
37	E	5	0	0	0	0
38	F	1	0	0	0	0
38	Q	2	0	0	0	0
38	R	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
38	T	3	0	0	0	0
All	All	75273	0	65184	2096	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 19.

The worst 5 of 2096 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:426:PRO:HB3	16:Z:201:ARG:CG	1.38	1.50
1:A:1327:THR:CG2	5:L:35:U:H5'	1.41	1.48
1:A:1327:THR:HG22	5:L:35:U:C5'	1.47	1.44
5:L:15:C:H1'	5:L:16:U:C5	1.50	1.42
16:Z:192:GLU:CA	16:Z:195:GLU:OE2	1.70	1.37

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	1902/2413 (79%)	1744 (92%)	149 (8%)	9 (0%)	34	75
2	C	872/1008 (86%)	796 (91%)	72 (8%)	4 (0%)	34	75
9	J	25/135 (18%)	20 (80%)	5 (20%)	0	100	100
10	O	335/451 (74%)	299 (89%)	31 (9%)	5 (2%)	13	54
11	P	193/379 (51%)	174 (90%)	17 (9%)	2 (1%)	19	63
12	Q	177/364 (49%)	154 (87%)	21 (12%)	2 (1%)	17	61
13	R	259/339 (76%)	241 (93%)	18 (7%)	0	100	100
14	S	63/175 (36%)	55 (87%)	5 (8%)	3 (5%)	3	25

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
15	T	155/157 (99%)	137 (88%)	16 (10%)	2 (1%)	15	57
16	Z	443/577 (77%)	390 (88%)	44 (10%)	9 (2%)	9	48
17	c	312/587 (53%)	291 (93%)	17 (5%)	4 (1%)	15	57
18	d	238/687 (35%)	221 (93%)	17 (7%)	0	100	100
19	F	113/278 (41%)	104 (92%)	8 (7%)	1 (1%)	21	65
20	G	39/179 (22%)	37 (95%)	1 (3%)	1 (3%)	7	42
21	H	93/235 (40%)	78 (84%)	14 (15%)	1 (1%)	17	61
22	I	98/215 (46%)	95 (97%)	3 (3%)	0	100	100
23	v	121/859 (14%)	110 (91%)	7 (6%)	4 (3%)	5	37
24	n	279/455 (61%)	242 (87%)	29 (10%)	8 (3%)	6	40
25	o	120/503 (24%)	115 (96%)	4 (3%)	1 (1%)	24	67
25	p	122/503 (24%)	116 (95%)	6 (5%)	0	100	100
25	q	355/503 (71%)	327 (92%)	16 (4%)	12 (3%)	5	36
25	r	119/503 (24%)	111 (93%)	5 (4%)	3 (2%)	7	43
26	t	150/175 (86%)	134 (89%)	13 (9%)	3 (2%)	9	48
27	k	76/196 (39%)	69 (91%)	7 (9%)	0	100	100
27	s	74/196 (38%)	67 (90%)	7 (10%)	0	100	100
28	i	71/94 (76%)	65 (92%)	6 (8%)	0	100	100
28	u	71/94 (76%)	65 (92%)	6 (8%)	0	100	100
29	h	66/86 (77%)	61 (92%)	4 (6%)	1 (2%)	13	54
29	w	66/86 (77%)	61 (92%)	4 (6%)	1 (2%)	13	54
30	j	65/77 (84%)	64 (98%)	1 (2%)	0	100	100
30	x	65/77 (84%)	64 (98%)	1 (2%)	0	100	100
31	l	80/101 (79%)	70 (88%)	9 (11%)	1 (1%)	15	57
31	y	80/101 (79%)	70 (88%)	9 (11%)	1 (1%)	15	57
32	m	78/146 (53%)	74 (95%)	4 (5%)	0	100	100
32	z	78/146 (53%)	74 (95%)	4 (5%)	0	100	100
33	e	63/110 (57%)	58 (92%)	4 (6%)	1 (2%)	12	53
33	g	92/110 (84%)	85 (92%)	6 (6%)	1 (1%)	17	61
34	a	77/111 (69%)	75 (97%)	2 (3%)	0	100	100
35	b	125/238 (52%)	111 (89%)	12 (10%)	2 (2%)	12	53

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	7810/13649 (57%)	7124 (91%)	604 (8%)	82 (1%)	24	63

5 of 82 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	483	PRO
2	C	364	PHE
2	C	463	THR
16	Z	206	LYS
20	G	17	MET

### 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	
1	A	1736/2182 (80%)	1721 (99%)	15 (1%)	84	94
2	C	794/910 (87%)	789 (99%)	5 (1%)	90	96
9	J	21/121 (17%)	21 (100%)	0	100	100
10	O	295/397 (74%)	292 (99%)	3 (1%)	82	93
11	P	173/328 (53%)	172 (99%)	1 (1%)	90	96
12	Q	171/332 (52%)	154 (90%)	17 (10%)	10	39
13	R	224/296 (76%)	223 (100%)	1 (0%)	93	97
14	S	56/151 (37%)	53 (95%)	3 (5%)	27	67
15	T	141/141 (100%)	140 (99%)	1 (1%)	88	95
16	Z	417/538 (78%)	382 (92%)	35 (8%)	14	49
17	c	212/316 (67%)	206 (97%)	6 (3%)	51	83
18	d	219/249 (88%)	218 (100%)	1 (0%)	92	96
19	F	105/256 (41%)	103 (98%)	2 (2%)	65	87
20	G	39/163 (24%)	39 (100%)	0	100	100
21	H	88/216 (41%)	87 (99%)	1 (1%)	80	92
22	I	92/193 (48%)	92 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
23	v	58/152 (38%)	48 (83%)	10 (17%)	2	13
24	n	124/413 (30%)	102 (82%)	22 (18%)	2	12
25	o	60/451 (13%)	53 (88%)	7 (12%)	7	30
25	p	62/451 (14%)	54 (87%)	8 (13%)	5	26
25	q	124/451 (28%)	107 (86%)	17 (14%)	4	23
25	r	60/451 (13%)	55 (92%)	5 (8%)	14	49
26	t	40/165 (24%)	29 (72%)	11 (28%)	0	3
27	k	70/176 (40%)	70 (100%)	0	100	100
27	s	67/176 (38%)	67 (100%)	0	100	100
28	i	65/83 (78%)	60 (92%)	5 (8%)	16	53
28	u	65/83 (78%)	60 (92%)	5 (8%)	16	53
29	h	61/77 (79%)	60 (98%)	1 (2%)	70	89
29	w	61/77 (79%)	60 (98%)	1 (2%)	70	89
30	j	58/66 (88%)	55 (95%)	3 (5%)	29	68
30	x	58/66 (88%)	55 (95%)	3 (5%)	29	68
31	l	69/89 (78%)	67 (97%)	2 (3%)	50	82
31	y	69/89 (78%)	67 (97%)	2 (3%)	50	82
32	m	77/129 (60%)	71 (92%)	6 (8%)	16	52
32	z	77/129 (60%)	71 (92%)	6 (8%)	16	52
33	e	59/103 (57%)	55 (93%)	4 (7%)	20	60
33	g	79/103 (77%)	74 (94%)	5 (6%)	22	62
34	a	26/100 (26%)	25 (96%)	1 (4%)	40	76
35	b	47/219 (22%)	44 (94%)	3 (6%)	22	62
All	All	6319/11088 (57%)	6101 (97%)	218 (3%)	48	79

5 of 218 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
23	v	725	THR
24	n	436	PRO
30	x	71	LEU
24	n	51	PHE
24	n	162	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 87 such sidechains are listed below:

Mol	Chain	Res	Type
11	P	174	ASN
13	R	191	ASN
28	u	34	GLN
12	Q	53	ASN
13	R	91	HIS

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	D	114/214 (53%)	30 (26%)	3 (2%)
4	E	102/112 (91%)	34 (33%)	7 (6%)
5	L	120/1175 (10%)	25 (20%)	5 (4%)
6	M	28/29 (96%)	7 (25%)	0
7	B	12/13 (92%)	5 (41%)	0
8	N	14/15 (93%)	7 (50%)	1 (7%)
All	All	390/1558 (25%)	108 (27%)	16 (4%)

5 of 108 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	D	29	G
3	D	31	G
3	D	32	G
3	D	33	U
3	D	42	A

5 of 16 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
4	E	64	U
4	E	66	C
5	L	117	U
4	E	56	A
5	L	1107	C

## 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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 14 ligands modelled in this entry, 13 are monoatomic - leaving 1 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	$\# Z  > 2$
36	GTP	C	1500	37	26,34,34	1.04	1 (3%)	29,54,54	1.60	5 (17%)

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
36	GTP	C	1500	37	-	0/18/38/38	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
36	C	1500	GTP	C6-N1	2.41	1.37	1.33

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
36	C	1500	GTP	N3-C2-N1	-5.17	120.52	127.56
36	C	1500	GTP	C5-C6-N1	-2.87	119.77	123.52
36	C	1500	GTP	C2'-C1'-N9	-2.47	106.86	113.47
36	C	1500	GTP	C4'-O4'-C1'	2.07	111.83	109.64
36	C	1500	GTP	C6-N1-C2	3.03	119.44	115.88

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
36	C	1500	GTP	1	0

## 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.