



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

Feb 21, 2017 – 03:36 PM EST

PDB ID : 5MQ0  
EMDB ID: : EMD-3541  
Title : Structure of a spliceosome remodeled for exon ligation  
Authors : Fica, S.M.; Oubridge, C.; Galej, W.P.; Wilkinson, M.E.; Newman, A.J.; Bai, X.-C.; Nagai, K.  
Deposited on : 2016-12-19  
Resolution : 4.17 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



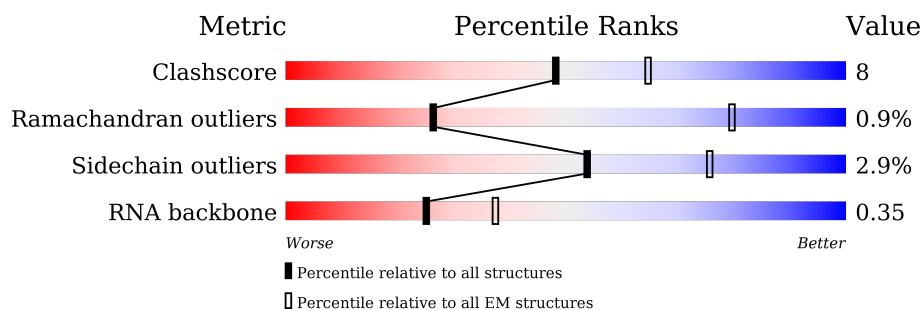
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 4.17 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














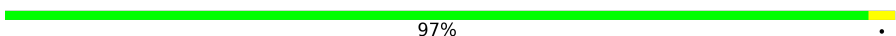











The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of 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	I	95	13% 19% . 64%
2	E	20	15% 50% 15% 20%
3	2	1175	8% . . 87%
4	6	112	36% 43% 10% 12%
5	5	179	32% 37% 9% 21%
6	A	2413	64% 15% 21%
7	C	1008	71% 16% 13%
8	H	577	52% 17% . 31%

*Continued on next page...*




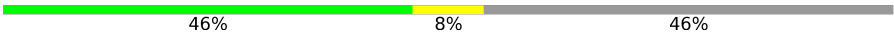


Continued from previous page...

Mol	Chain	Length	Quality of chain
9	J	451	
10	K	379	
11	L	157	
12	M	339	
13	N	364	
14	O	590	
15	P	175	
16	R	135	
17	S	687	
18	T	877	
19	a	251	
20	c	382	
21	o	455	
22	X	68	
23	y	215	
24	b	196	
24	k	196	
25	d	101	
25	n	101	
26	e	94	
26	p	94	
27	f	86	
27	q	86	
28	g	77	
28	r	77	

Continued on next page...



*Continued from previous page...*

Mol	Chain	Length	Quality of chain
29	h	146	
29	l	146	
30	j	110	
30	m	110	
31	V	1145	
32	W	238	
33	Y	111	
34	3	3	
35	s	175	
36	t	503	
36	u	503	
36	v	503	
36	w	503	



## 2 Entry composition

There are 41 unique types of molecules in this entry. The entry contains 77441 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 Yeast UBC4 gene for ubiquitin-conjugating enzyme.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	I	34	Total	C	N	O	P	0	0
			714	321	118	241	34		

- Molecule 2 is a RNA chain called 5'-EXON OF UBC4 PRE-MRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	E	16	Total	C	N	O	P	0	0
			346	155	66	109	16		

- Molecule 3 is a RNA chain called S.cerevisiae chromosome II reading frame ORF YBR230c.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	2	155	Total	C	N	O	P	0	0
			3271	1462	547	1107	155		

- Molecule 4 is a RNA chain called Saccharomyces cerevisiae strain T.52\_2H chromosome XII sequence.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	6	99	Total	C	N	O	P	0	0
			2108	944	375	690	99		

- Molecule 5 is a RNA chain called Saccharomyces cerevisiae strain WI\_C\_MBSP\_4 chromosome VII sequence.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	5	141	Total	C	N	O	P	0	0
			2999	1342	530	986	141		

- Molecule 6 is a protein called Pre-mRNA-splicing factor 8.



Mol	Chain	Residues	Atoms					AltConf	Trace
6	A	1914	Total	C	N	O	S	0	0
			15199	9832	2669	2645	53		

- Molecule 7 is a protein called Pre-mRNA-splicing factor SNU114.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	C	874	Total	C	N	O	S	0	0
			6562	4265	1104	1168	25		

- Molecule 8 is a protein called Pre-mRNA-splicing factor CWC22.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	401	Total	C	N	O	S	0	0
			3261	2104	544	595	18		

- Molecule 9 is a protein called Pre-mRNA-splicing factor PRP46.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	J	342	Total	C	N	O	S	0	0
			2690	1699	475	506	10		

- Molecule 10 is a protein called Pre-mRNA-processing protein 45.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	K	187	Total	C	N	O	S	0	0
			1458	908	269	276	5		

- Molecule 11 is a protein called Pre-mRNA-splicing factor BUD31.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	L	155	Total	C	N	O	S	0	0
			1162	737	217	198	10		

- Molecule 12 is a protein called Pre-mRNA-splicing factor CWC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	M	252	Total	C	N	O	S	0	0
			2016	1281	356	368	11		

- Molecule 13 is a protein called Pre-mRNA-splicing factor SLT11.



Mol	Chain	Residues	Atoms					AltConf	Trace
13	N	227	Total	C	N	O	S	0	0
			1798	1139	309	335	15		

- Molecule 14 is a protein called Pre-mRNA-splicing factor CEF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	O	211	Total	C	N	O	S	0	0
			1755	1102	320	327	6		

- Molecule 15 is a protein called Pre-mRNA-splicing factor CWC15.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	P	69	Total	C	N	O	S	0	0
			565	358	112	94	1		

- Molecule 16 is a protein called Pre-mRNA-splicing factor CWC21.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	R	108	Total	C	N	O		0	0
			614	369	121	124			

- Molecule 17 is a protein called Pre-mRNA-splicing factor CLF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	S	468	Total	C	N	O	S	0	0
			3229	2025	599	598	7		

- Molecule 18 is a protein called Pre-mRNA-splicing factor SYF1,PRE-MRNA-SPLICING FACTOR SYF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	T	633	Total	C	N	O		0	0
			3154	1888	633	633			

- Molecule 19 is a protein called Pre-mRNA-splicing factor 18.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	a	137	Total	C	N	O	S	0	0
			1119	726	194	196	3		

- Molecule 20 is a protein called Pre-mRNA-splicing factor SLU7.



Mol	Chain	Residues	Atoms					AltConf	Trace
20	c	103	Total	C	N	O	S	0	0
			786	498	142	144	2		

- Molecule 21 is a protein called Pre-mRNA-processing factor 17.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	o	313	Total	C	N	O	S	0	0
			2425	1537	429	451	8		

- Molecule 22 is a protein called UNKNOWN PROTEIN.

Mol	Chain	Residues	Atoms				AltConf	Trace
22	X	68	Total	C	N	O	0	0
			338	202	68	68		

- Molecule 23 is a protein called Pre-mRNA-splicing factor SYF2.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	y	83	Total	C	N	O	S	0	0
			679	420	125	133	1		

- Molecule 24 is a protein called Small nuclear ribonucleoprotein-associated protein B.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	b	80	Total	C	N	O	S	0	0
			631	403	114	111	3		
24	k	80	Total	C	N	O	S	0	0
			631	403	114	111	3		

- Molecule 25 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	d	82	Total	C	N	O	S	0	0
			625	399	109	115	2		
25	n	82	Total	C	N	O	S	0	0
			625	399	109	115	2		

- Molecule 26 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	e	75	Total	C	N	O	S	0	0
			575	379	92	101	3		

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms					AltConf	Trace
26	p	75	Total	C	N	O	S	0	0
			575	379	92	101	3		

- Molecule 27 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	f	72	Total	C	N	O	S	0	0
			573	368	101	103	1		
27	q	72	Total	C	N	O	S	0	0
			573	368	101	103	1		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
28	g	69	Total	C	N	O	S	0	0
			529	337	93	97	2		
28	r	69	Total	C	N	O	S	0	0
			526	336	93	95	2		

- Molecule 29 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	h	82	Total	C	N	O	S	0	0
			644	409	110	123	2		
29	l	79	Total	C	N	O	S	0	0
			618	393	107	116	2		

- Molecule 30 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	j	94	Total	C	N	O	S	0	0
			741	477	141	119	4		
30	m	93	Total	C	N	O	S	0	0
			726	468	136	118	4		

- Molecule 31 is a protein called Pre-mRNA-splicing factor ATP-dependent RNA helicase PRP22.

Mol	Chain	Residues	Atoms				AltConf	Trace
31	V	645	Total	C	N	O	0	0
			3189	1899	645	645		



- Molecule 32 is a protein called U2 small nuclear ribonucleoprotein A'.

Mol	Chain	Residues	Atoms				AltConf	Trace
32	W	164	Total	C	N	O	0	0
			816	488	164	164		

- Molecule 33 is a protein called U2 small nuclear ribonucleoprotein B''.

Mol	Chain	Residues	Atoms				AltConf	Trace
33	Y	84	Total	C	N	O	0	0
			416	248	84	84		

- Molecule 34 is a RNA chain called 3'-EXON OF UBC4 PRE-MRNA, BOUND BY PRP22 HELICASE.

Mol	Chain	Residues	Atoms					AltConf	Trace
34	3	3	Total	C	N	O	P	0	0
			60	27	6	24	3		

- Molecule 35 is a protein called Pre-mRNA-splicing factor SNT309.

Mol	Chain	Residues	Atoms				AltConf	Trace
35	s	110	Total	C	N	O	0	0
			548	328	110	110		

- Molecule 36 is a protein called Pre-mRNA-processing factor 19.

Mol	Chain	Residues	Atoms				AltConf	Trace
36	t	438	Total	C	N	O	0	0
			2171	1295	438	438		
36	u	116	Total	C	N	O	0	0
			578	346	116	116		
36	v	118	Total	C	N	O	0	0
			588	352	118	118		
36	w	435	Total	C	N	O	0	0
			2156	1286	435	435		

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

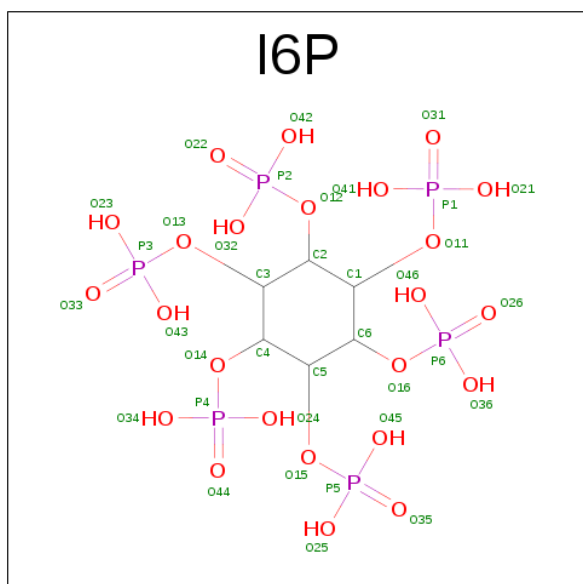
Mol	Chain	Residues	Atoms		AltConf
37	6	3	Total	Mg	0
			3	3	

- Molecule 38 is POTASSIUM ION (three-letter code: K) (formula: K).



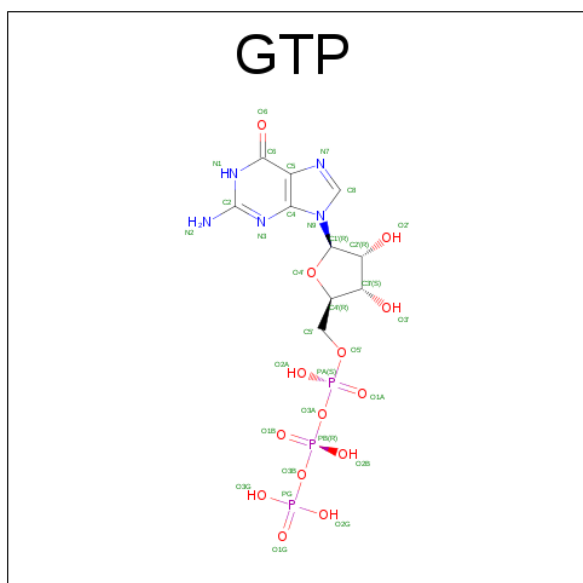
Mol	Chain	Residues	Atoms	AltConf
38	6	2	Total K 2 2	0

- Molecule 39 is INOSITOL 1,2,3,4,5,6-HEXAKISPHOSPHATE (three-letter code: I6P) (formula:  $\text{C}_6\text{H}_{18}\text{O}_{24}\text{P}_6$ ).



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

- Molecule 40 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula:  $\text{C}_{10}\text{H}_{16}\text{N}_5\text{O}_{14}\text{P}_3$ ).





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

- Molecule 41 is ZINC ION (three-letter code: ZN) (formula: Zn).

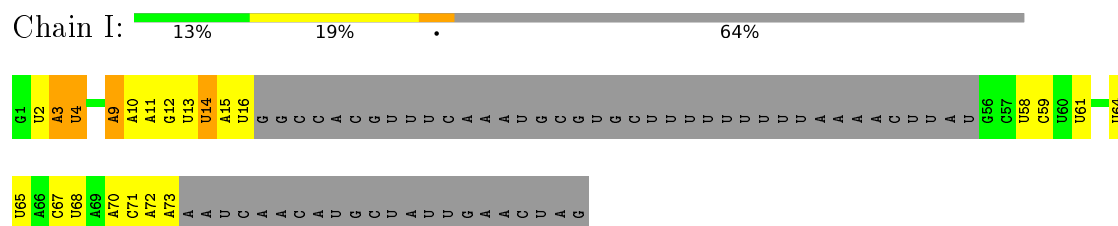
Mol	Chain	Residues	Atoms		AltConf
41	L	3	Total	Zn	0
			3	3	
41	N	2	Total	Zn	0
			2	2	
41	M	1	Total	Zn	0
			1	1	



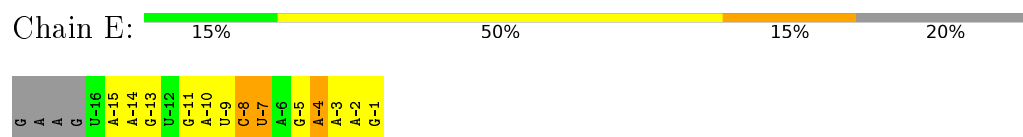
### 3 Residue-property plots

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.

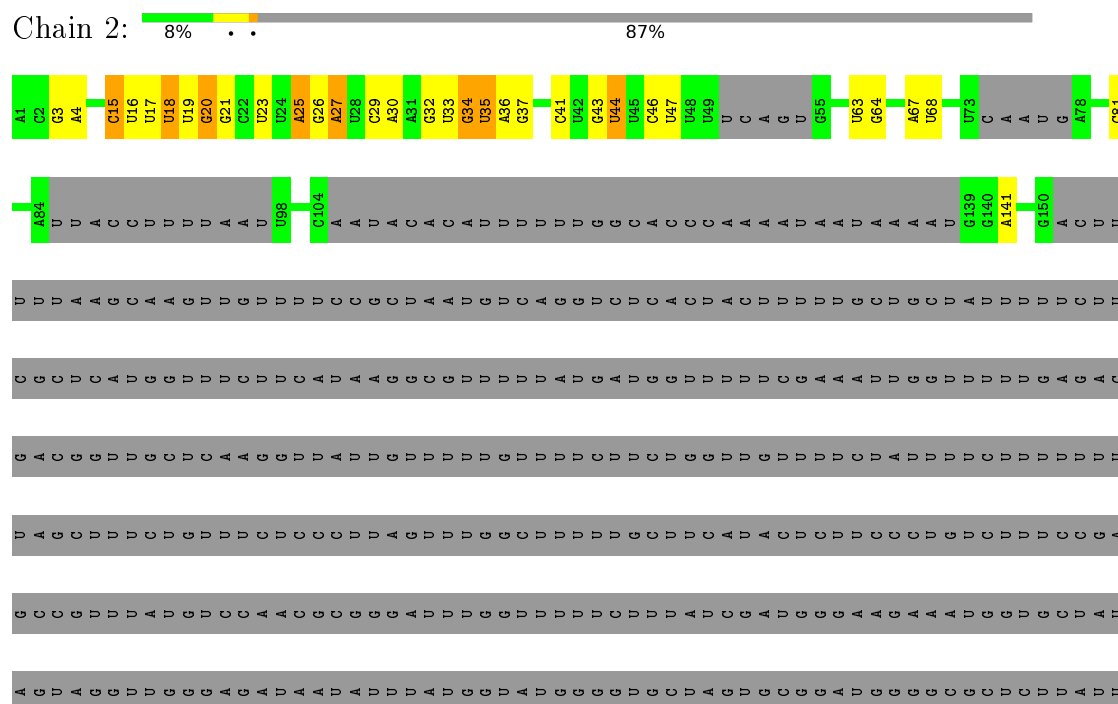
- Molecule 1: Yeast UBC4 gene for ubiquitin-conjugating enzyme



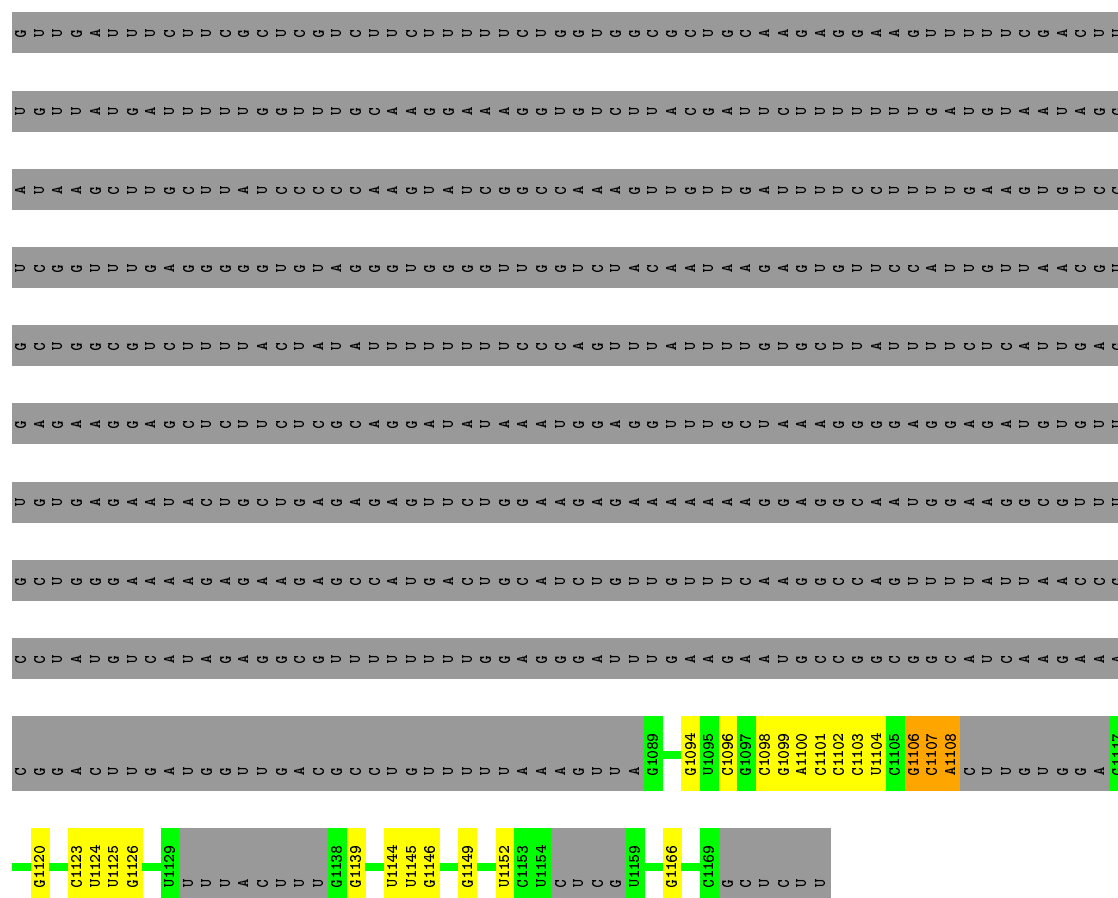
- Molecule 2: 5'-EXON OF UBC4 PRE-MRNA



- Molecule 3: *S.cerevisiae* chromosome II reading frame ORF YBR230c

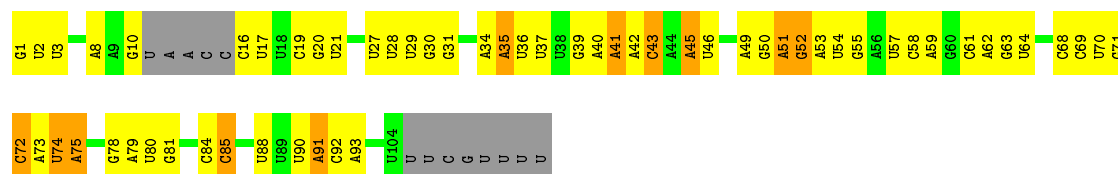






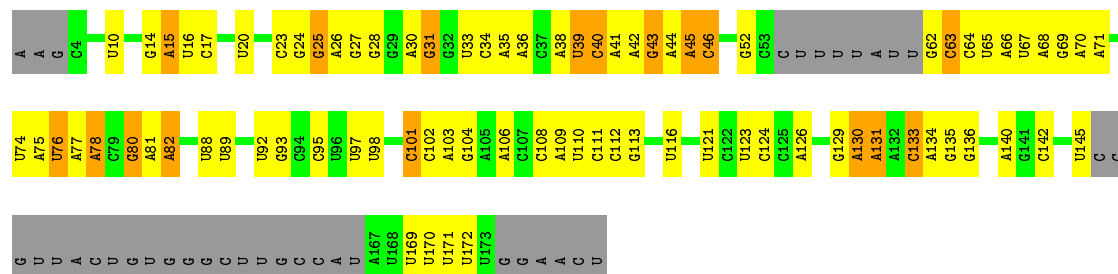
- Molecule 4: *Saccharomyces cerevisiae* strain T.52\_2H chromosome XII sequence

Chain 6: 36% 43% 10% 12%



- Molecule 5: *Saccharomyces cerevisiae* strain WI\_C\_Mbsp\_4 chromosome VII sequence

Chain 5: 32% 37% 9% 21%



- Molecule 6: Pre-mRNA-splicing factor 8



Chain A:

[illegible]



[illegible]

- Molecule 7: Pre-mRNA-splicing factor SNU114

Chain C:  71% 16% 13%

ASN	GLY	V691	ASN	R360	L193	GLU	NET
	LEU	S692	GLU	R360	M194	LEU	GLY
	VAL	S693	ASP	R376	G195	GLY	GLY
	PRO	V825	GLU	I377	S196	ILE	ASP
		N830	ILE	ASP	T197	GLU	SER
		M834	SER	GLU	Y382	LEU	LEU
	GLU	I839	ARG	ASP	Y382	C200	PHE
		P840	GLY	LEU	D632	I210	HIS
		I841	GLY	GLY	F385	I211	PRO
		P842	GLU	ASN	L400	I212	THR
ASN	K843	ASN	V539	L400	I213	GLY	
	T852	LEU	V539	D416	I214	LYS	
	THR	K852	PRO	T542	P417	E73	LEU
		L857	GLY	GLY	Q418	T95	ILE
PRO	P860	LEU	L544	Q418	T95	VAL	
	I861	SER	L545	P419	I105	ASP	
	Y862	SER	L545	F420	F106	PRO	
	GLY	I866	ASN	V550	L421	T107	PHE
P867		ASN	V550	R422	T116	ASP	
I868		THR	Q560	R422	T117	SER	
P871		LEU	V561	V424	T118	GLY	
LEU	I886	LEU	V562	Q426	M126	GLU	
	P887	GLY	L563	L427	A127	SER	
	I888	LYS	L564	L428	R132	LEU	
	P890	GLY	L567	R430	I133	ASP	
ASN	A870	GLN	S568	V436	V241	GLU	
	P871	CYS	S568	D437	V242	GLN	
	L879	LEU	V571	A438	L252	GLN	
	I890	ASP	M586	Q444	I139	THR	
GLY	E901	GLY	K590	P445	V261	GLN	
	V902	ILE	ILE	Q445	T267	THR	
	V908	NET	V602	K461	T272	ASN	
	A912	ASP	K617	G479	L273	THR	
GLU	I920	ASN	L622	W493	M150	PHE	
	S921	GLU	L622	W493	D151	GLU	
	M927	GLY	P630	R487	L152	GLY	
	C928	THR	P630	R487	L153	SER	
ASN	Q929	ASN	L633	L493	G301	GLY	
	W933	VAL	L633	L493	A314	ASN	
	A957	GLY	L644	I501	S315	ASN	
	P957	ASN	L645	L502	K317	GLU	
THR	P987	VAL	D653	R510	L318	ILE	
	L997	THR	C954	R510	G319	GLU	
	ALA	THR	L655	L515	F320	SER	
	LEU	GLN	D658	HIS	T321	ARG	
GLY	LEU	GLN	ASP	T322	L178	GLN	
	ARG	THR	ILE	T323	K182	LEU	
	SER	GLY	SER	K324	R187	THR	
	LYS	GLY	LYS	K325	G188	LEU	
LYS	ARG	THR	THR	V331	S189	GLY	
	THR	THR	GLU	R347	S190	SER	
	VAL	THR	THR	R347	I191	LYS	
	VAL	THR	THR	R347	I191	LYS	

- Molecule 8: Pre-mRNA-splicing factor CWC22

Chain H:  52% 17% . 31%

[illegible]



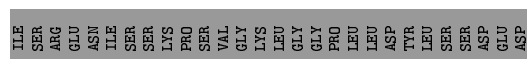
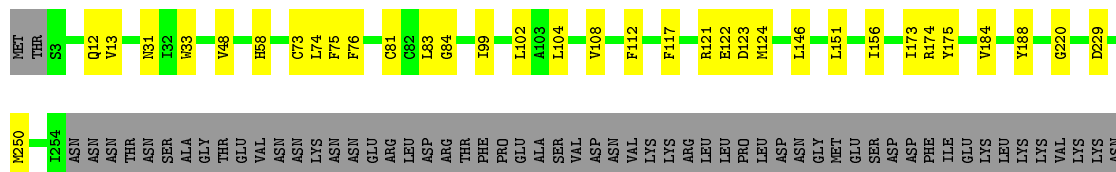






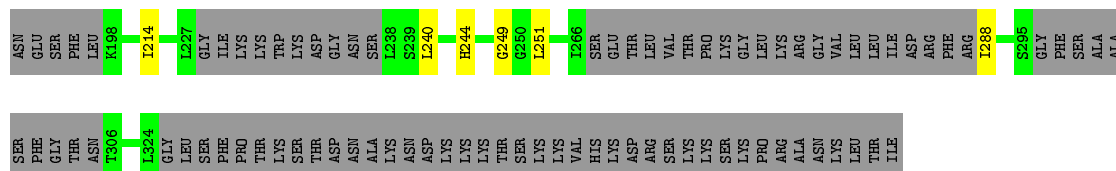
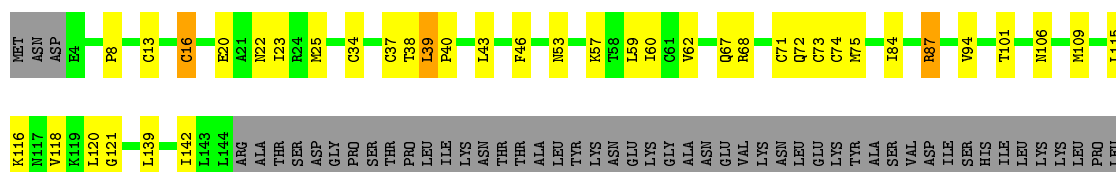
- Molecule 12: Pre-mRNA-splicing factor CWC2

Chain M:



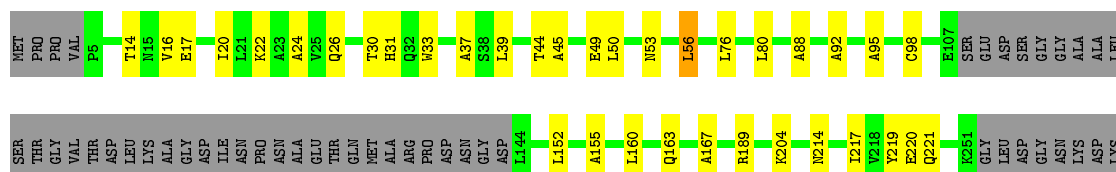
- Molecule 13: Pre-mRNA-splicing factor SLT11

Chain N:

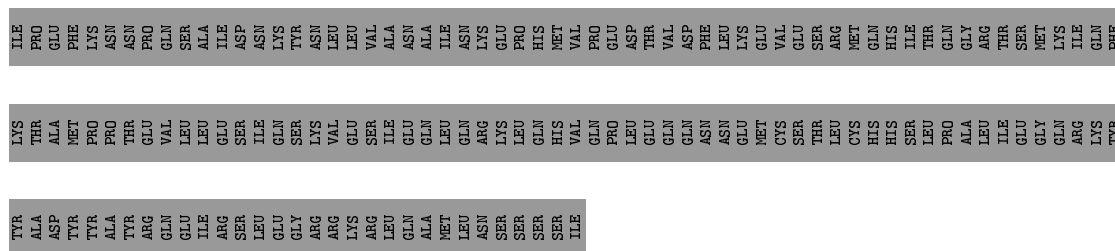


- Molecule 14: Pre-mRNA-splicing factor CEF1

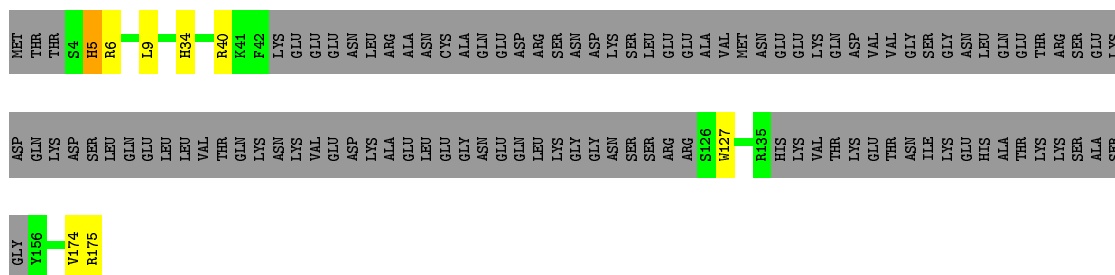
Chain O:







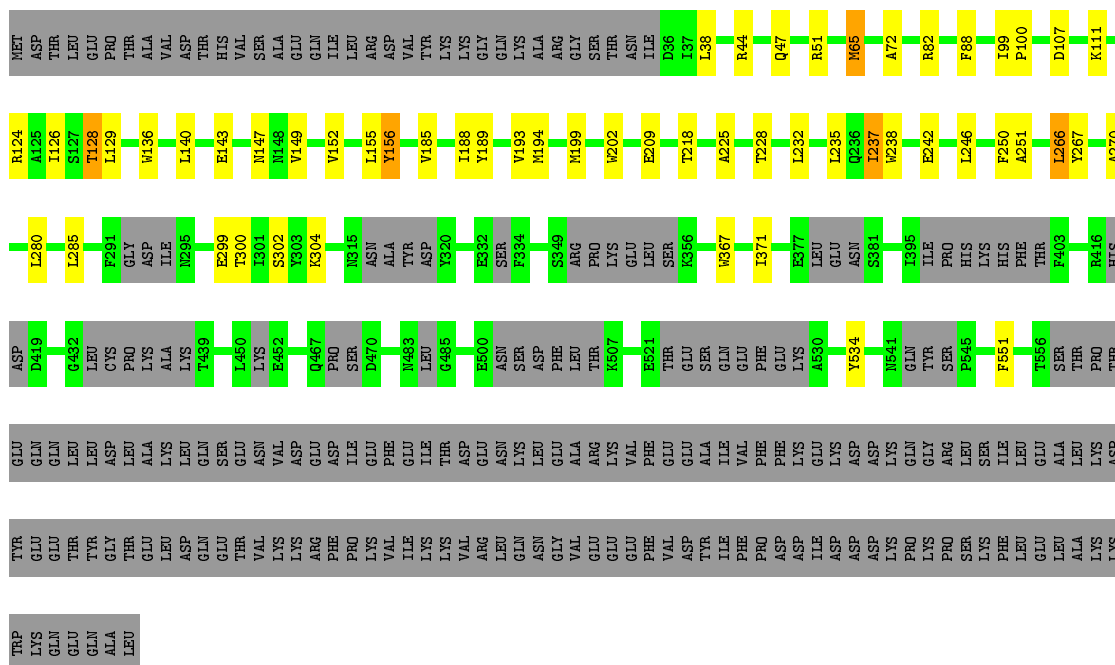
- Molecule 15: Pre-mRNA-splicing factor CWC15



- Molecule 16: Pre-mRNA-splicing factor CWC21



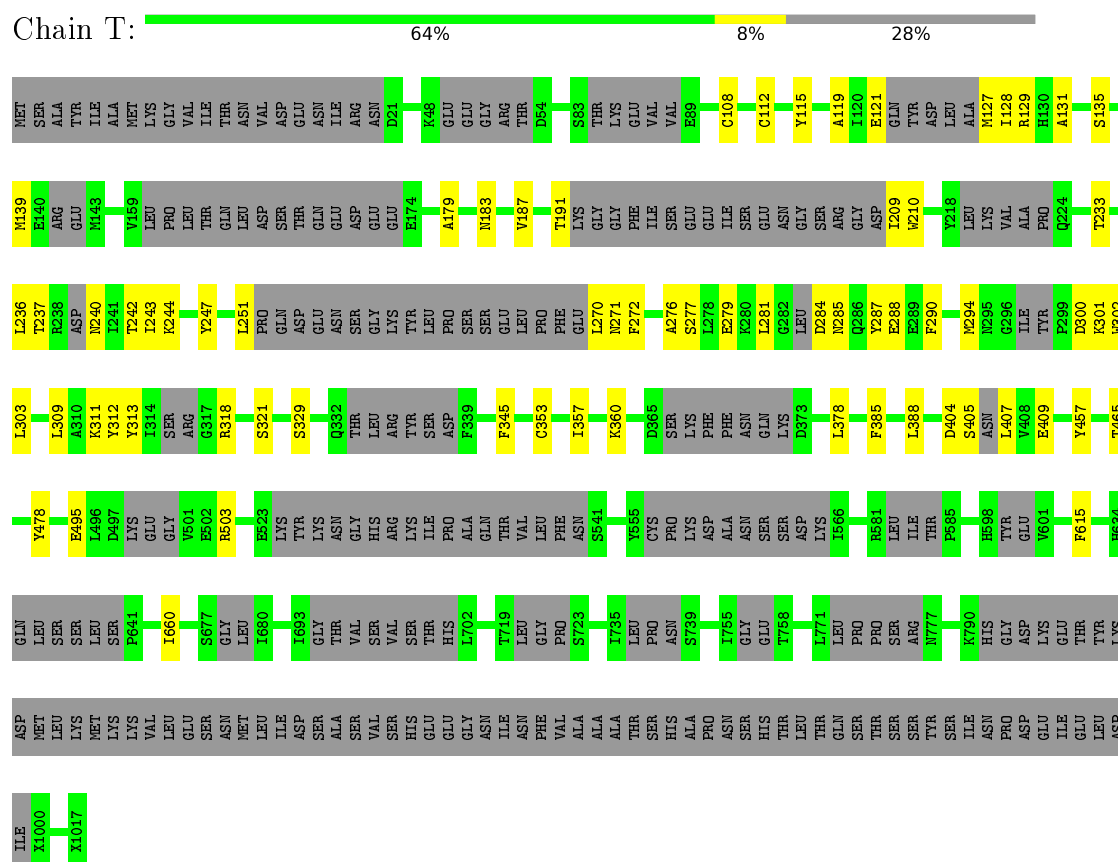
- Molecule 17: Pre-mRNA-splicing factor CLF1





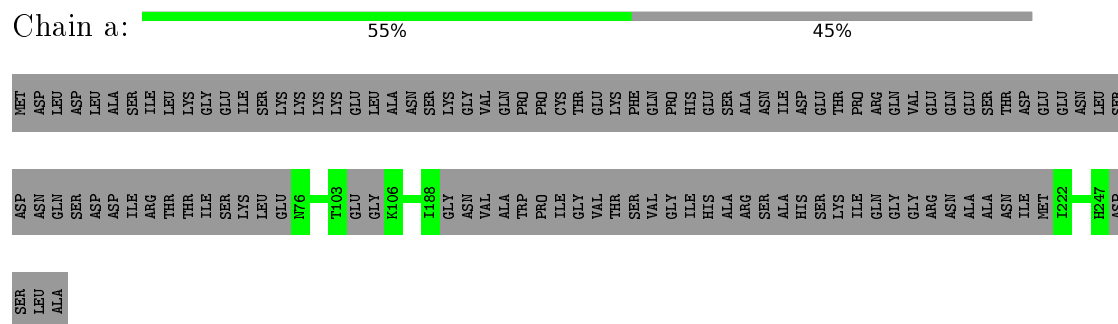
• Molecule 18: Pre-mRNA-splicing factor SYF1,PRE-MRNASPLICING FACTOR SYF1

Chain T:



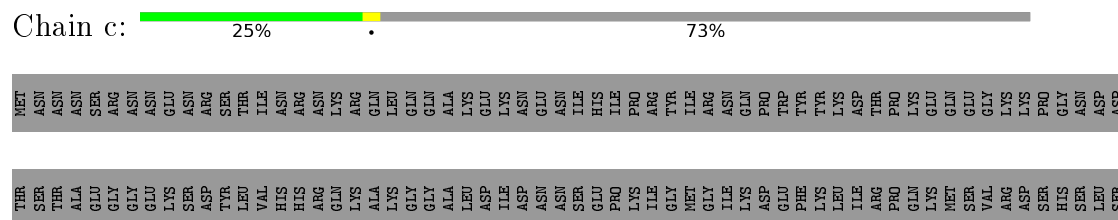
• Molecule 19: Pre-mRNA-splicing factor 18

Chain a:

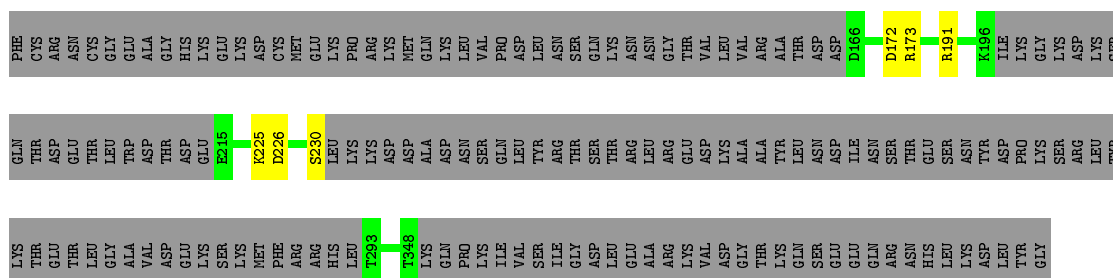


• Molecule 20: Pre-mRNA-splicing factor SLU7

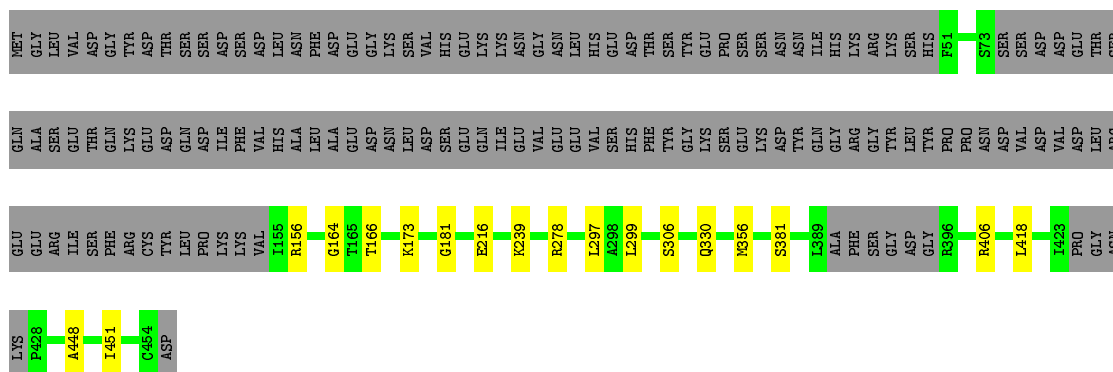
Chain c:







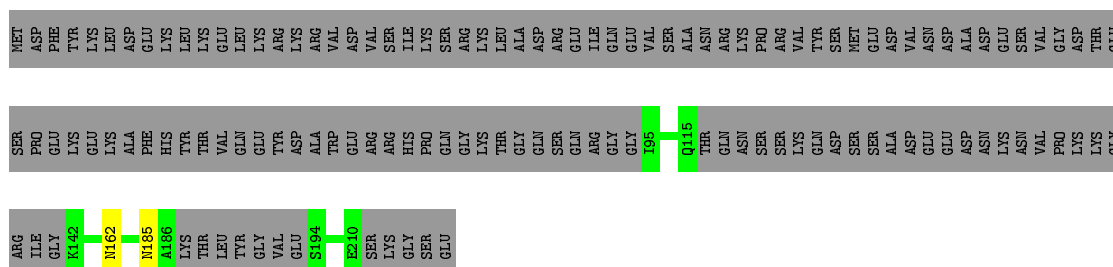
- Molecule 21: Pre-mRNA-processing factor 17



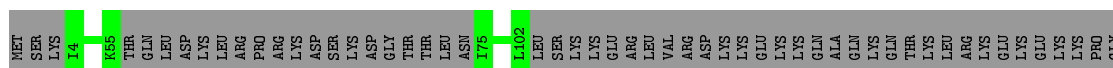
- Molecule 22: UNKNOWN PROTEIN



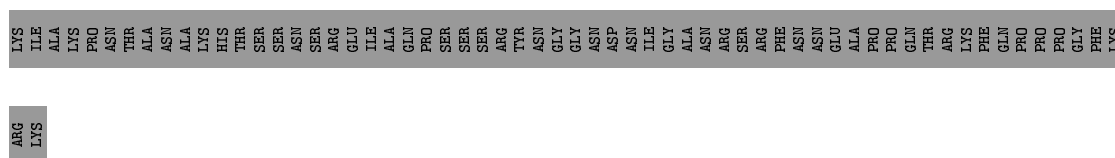
- Molecule 23: Pre-mRNA-splicing factor SYF2



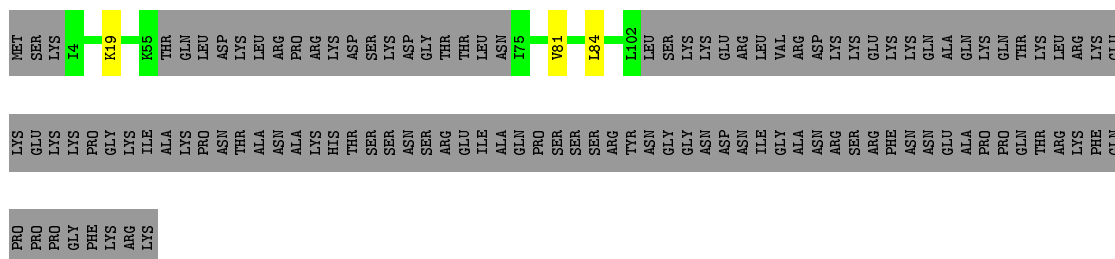
- Molecule 24: Small nuclear ribonucleoprotein-associated protein B



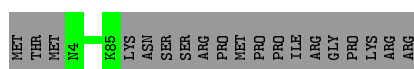
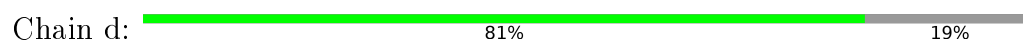




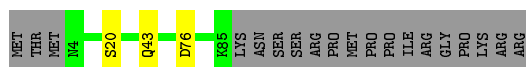
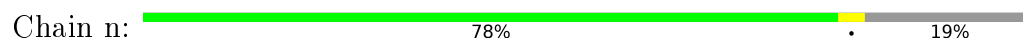
- Molecule 24: Small nuclear ribonucleoprotein-associated protein B



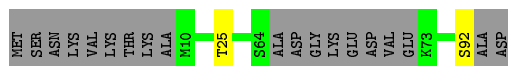
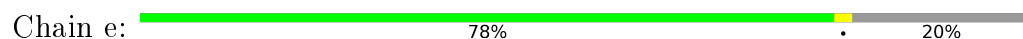
- Molecule 25: Small nuclear ribonucleoprotein Sm D3



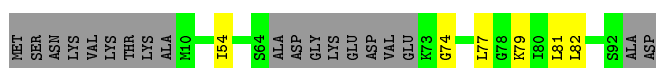
- Molecule 25: Small nuclear ribonucleoprotein Sm D3



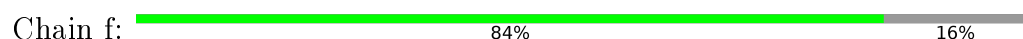
- Molecule 26: Small nuclear ribonucleoprotein E



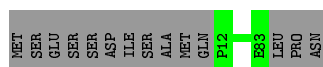
- Molecule 26: Small nuclear ribonucleoprotein E



- Molecule 27: Small nuclear ribonucleoprotein F

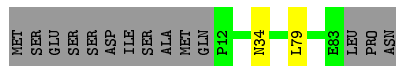






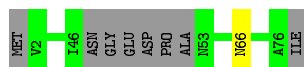
- Molecule 27: Small nuclear ribonucleoprotein F

Chain q: 81% 16%



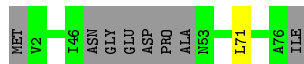
- Molecule 28: Small nuclear ribonucleoprotein G

Chain g: 88% 10%



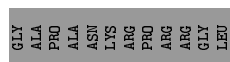
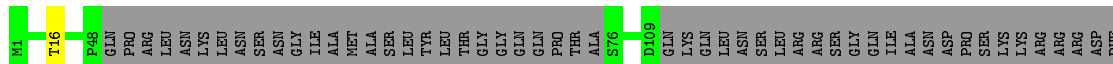
- Molecule 28: Small nuclear ribonucleoprotein G

Chain r: 88% 10%



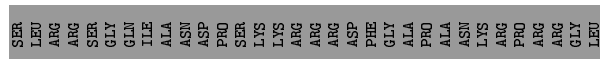
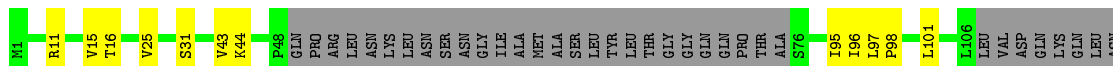
- Molecule 29: Small nuclear ribonucleoprotein Sm D1

Chain h: 55% 44%



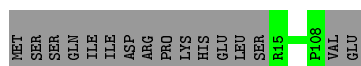
- Molecule 29: Small nuclear ribonucleoprotein Sm D1

Chain l: 46% 8% 46%



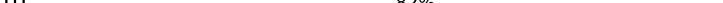
- Molecule 30: Small nuclear ribonucleoprotein Sm D2

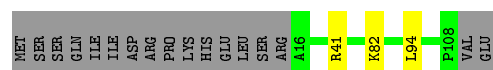
Chain j: 85% 15%



- Molecule 30: Small nuclear ribonucleoprotein Sm D2

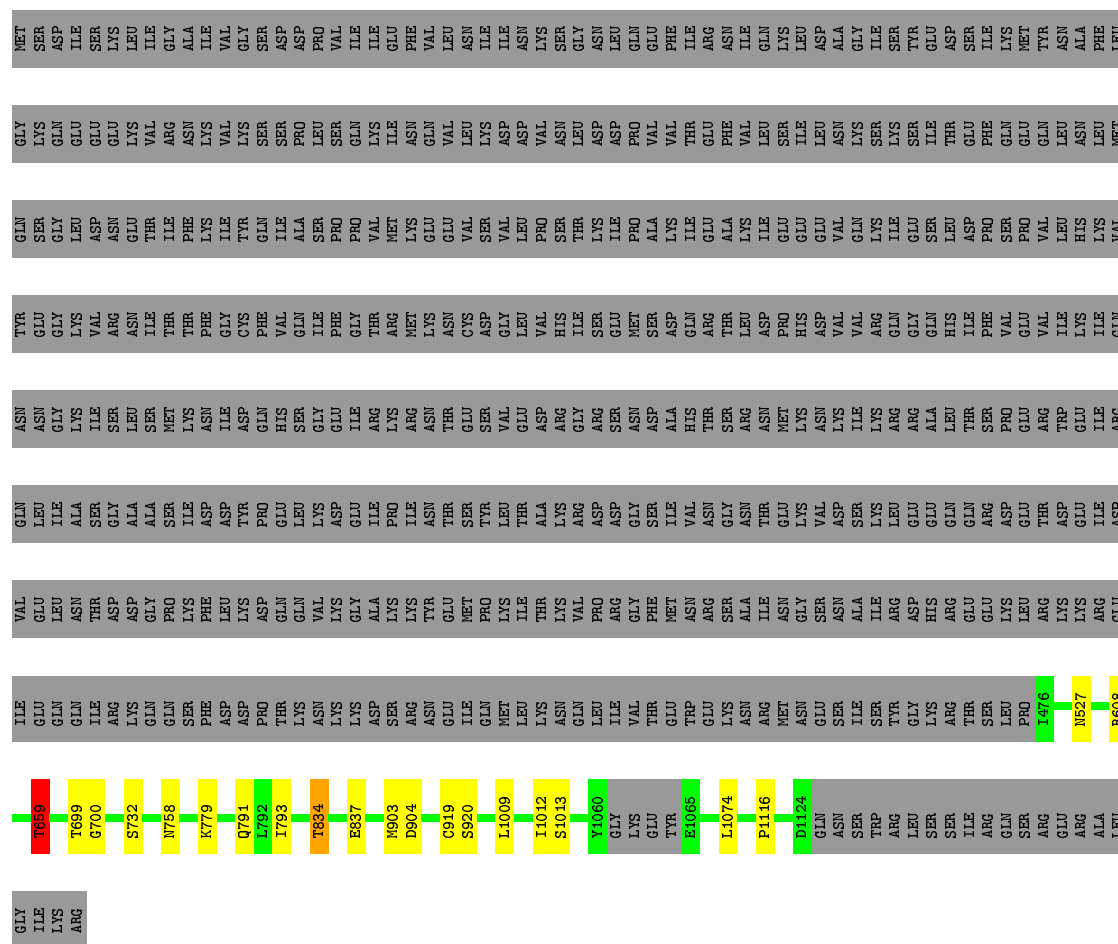


Chain m:  82% • 15%



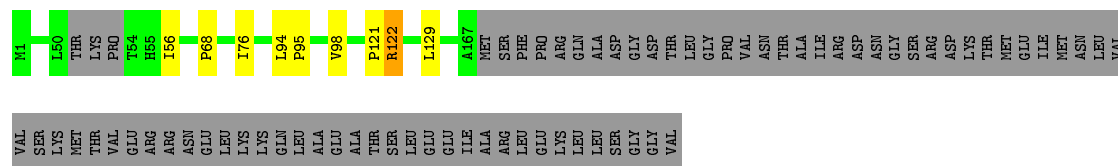
- Molecule 31: Pre-mRNA-splicing factor ATP-dependent RNA helicase PRP22

Chain V:  54% . 44%



- Molecule 32: U2 small nuclear ribonucleoprotein A'

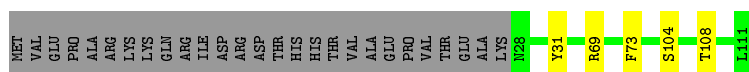
Chain W:  65% . 31%



- Molecule 33: U2 small nuclear ribonucleoprotein B''

Chain Y:  71% 5% 24%





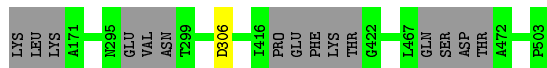
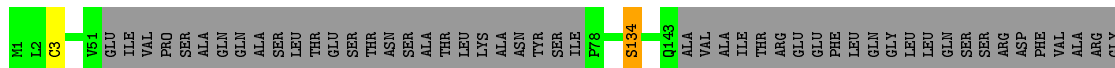
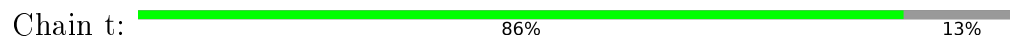
- Molecule 34: 3'-EXON OF UBC4 PRE-MRNA, BOUND BY PRP22 HELICASE



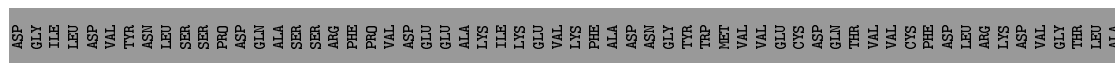
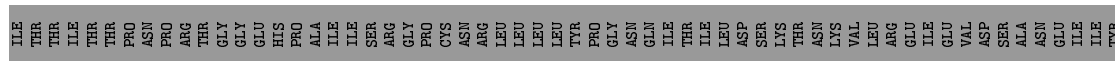
- Molecule 35: Pre-mRNA-splicing factor SNT309



- Molecule 36: Pre-mRNA-processing factor 19



- Molecule 36: Pre-mRNA-processing factor 19

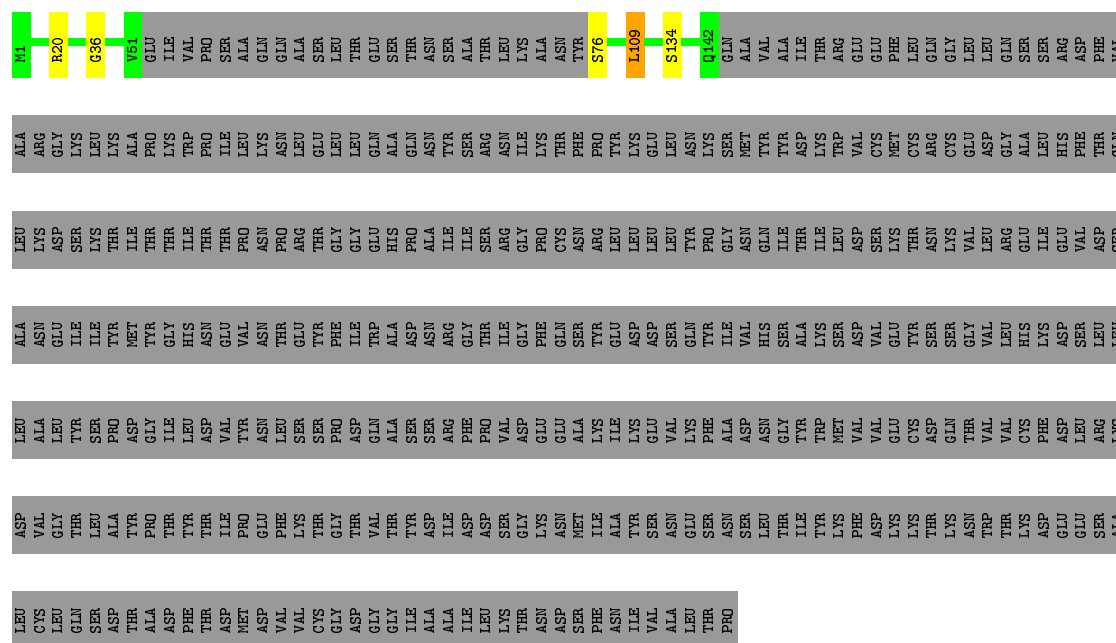




THR ALA ASP PHE THR ASP MET ASP VAL VAL CYS GLY ASP GLY GLY ILE ALA ALA ILE LEU LYS THR THR ASN ASP SER PHE ASN ILE VAL ALA LEU THR PRO

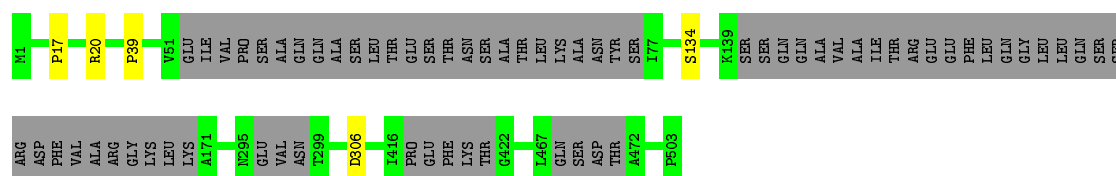
- Molecule 36: Pre-mRNA-processing factor 19

Chain v:  22% . 77%



- Molecule 36: Pre-mRNA-processing factor 19

Chain w:  85% • 14%





## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	29527	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)	500	Depositor
Maximum defocus (nm)	4500	Depositor
Magnification	81000	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, MG, K, I6P, 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	$\# Z  > 2$	RMSZ	$\# Z  > 2$
1	I	0.33	0/795	0.83	1/1231 (0.1%)
10	K	0.44	0/1480	0.75	0/2000
11	L	0.46	0/1186	0.72	0/1606
12	M	0.41	0/2062	0.66	0/2772
13	N	0.41	0/1823	0.71	0/2456
14	O	0.46	0/1781	0.78	0/2385
15	P	0.39	0/580	0.66	0/776
16	R	0.41	0/617	0.68	0/848
17	S	0.47	0/3269	0.76	0/4446
18	T	0.42	0/3036	0.71	0/4197
19	a	0.38	0/1141	0.61	0/1546
2	E	0.34	0/388	0.78	0/603
20	c	0.45	1/798 (0.1%)	0.60	0/1074
21	o	0.41	0/2491	0.64	0/3384
23	y	0.34	0/681	0.54	0/902
24	b	0.36	0/636	0.63	0/856
24	k	0.35	0/636	0.58	0/856
25	d	0.35	0/634	0.56	0/859
25	n	0.37	0/634	0.53	0/859
26	e	0.41	0/585	0.61	0/795
26	p	0.40	0/585	0.56	0/795
27	f	0.40	0/585	0.57	0/791
27	q	0.40	0/585	0.61	0/791
28	g	0.50	0/532	0.61	0/715
28	r	0.35	0/529	0.50	0/711
29	h	0.38	0/649	0.54	0/880
29	l	0.40	0/623	0.64	0/844
3	2	0.27	0/3639	0.72	0/5643
30	j	0.36	0/753	0.57	0/1013
30	m	0.37	0/738	0.61	0/995
31	V	0.57	2/3186 (0.1%)	1.53	9/4434 (0.2%)
32	W	0.31	0/814	0.53	0/1134



Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >2	RMSZ	# Z  >2
33	Y	0.32	0/415	0.55	0/577
34	3	0.19	0/65	0.65	0/98
35	s	0.57	0/546	0.80	0/760
36	t	0.46	0/2165	0.71	3/3010 (0.1%)
36	u	0.57	0/576	0.78	0/802
36	v	0.59	0/586	0.89	3/816 (0.4%)
36	w	0.47	0/2150	0.68	2/2989 (0.1%)
4	6	0.33	0/2357	0.72	1/3667 (0.0%)
5	5	0.32	0/3351	0.73	0/5213
6	A	0.46	0/15598	0.73	0/21212
7	C	0.42	0/6703	0.69	0/9138
8	H	0.48	0/3314	0.77	0/4463
9	J	0.47	0/2749	0.74	0/3735
All	All	0.43	3/79046 (0.0%)	0.76	19/109677 (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
15	P	0	1
21	o	0	1
24	k	0	1
31	V	0	3
35	s	0	2
36	t	0	2
36	v	0	1
36	w	0	1
7	C	0	1
9	J	0	1
All	All	0	14

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
31	V	1009	LEU	C-N	7.26	1.50	1.34
20	c	230	SER	C-O	6.14	1.35	1.23
31	V	834	THR	C-N	-5.27	1.22	1.34

The worst 5 of 19 bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
31	V	834	THR	C-N-CA	-46.64	5.11	121.70
31	V	834	THR	CA-C-N	-38.77	31.91	117.20
31	V	1009	LEU	O-C-N	27.39	166.52	122.70
31	V	1009	LEU	CA-C-N	-22.51	67.67	117.20
31	V	1009	LEU	C-N-CA	-18.95	74.32	121.70

There are no chirality outliers.

5 of 14 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
7	C	105	ILE	Peptide
9	J	194	HIS	Peptide
15	P	5	HIS	Peptide
31	V	659	THR	Mainchain
21	o	239	LYS	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	I	714	0	361	6	0
2	E	346	0	173	11	0
3	2	3271	0	1660	26	0
4	6	2108	0	1063	29	0
5	5	2999	0	1515	34	0
6	A	15199	0	14954	262	0
7	C	6562	0	6486	105	0
8	H	3261	0	3323	91	0
9	J	2690	0	2690	85	0
10	K	1458	0	1468	20	0
11	L	1162	0	1111	19	0
12	M	2016	0	1985	32	0
13	N	1798	0	1842	38	0
14	O	1755	0	1794	24	0
15	P	565	0	555	7	0
16	R	614	0	390	8	0
17	S	3229	0	2573	37	0
18	T	3154	0	1331	94	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
19	a	1119	0	1164	0	0
20	c	786	0	719	0	0
21	o	2425	0	2253	0	0
22	X	338	0	70	1	0
23	y	679	0	706	0	0
24	b	631	0	670	0	0
24	k	631	0	665	0	0
25	d	625	0	647	0	0
25	n	625	0	647	0	0
26	e	575	0	597	0	0
26	p	575	0	597	0	0
27	f	573	0	572	0	0
27	q	573	0	572	0	0
28	g	529	0	557	0	0
28	r	526	0	555	0	0
29	h	644	0	686	0	0
29	l	618	0	660	0	0
30	j	741	0	778	0	0
30	m	726	0	754	0	0
31	V	3189	0	1412	12	0
32	W	816	0	340	1	0
33	Y	416	0	182	14	0
34	3	60	0	31	10	0
35	s	548	0	219	0	0
36	t	2171	0	945	0	0
36	u	578	0	246	0	0
36	v	588	0	250	0	0
36	w	2156	0	938	0	0
37	6	3	0	0	0	0
38	6	2	0	0	0	0
39	A	36	0	6	0	0
40	C	32	0	12	0	0
41	L	3	0	0	0	0
41	M	1	0	0	0	0
41	N	2	0	0	0	0
All	All	77441	0	63724	846	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 8.

The worst 5 of 846 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)
18:T:119:ALA:HB2	18:T:131:ALA:CB	1.25	1.63
18:T:119:ALA:CB	18:T:131:ALA:HB3	1.40	1.48
18:T:237:THR:N	18:T:240:ASN:N	1.73	1.34
18:T:300:ASP:CB	18:T:303:LEU:CB	2.06	1.33
18:T:119:ALA:HB1	18:T:128:ILE:CA	1.58	1.33

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	
6	A	1906/2413 (79%)	1733 (91%)	161 (8%)	12 (1%)	30	74
7	C	864/1008 (86%)	771 (89%)	87 (10%)	6 (1%)	26	71
8	H	393/577 (68%)	353 (90%)	36 (9%)	4 (1%)	19	65
9	J	340/451 (75%)	295 (87%)	39 (12%)	6 (2%)	11	55
10	K	183/379 (48%)	163 (89%)	17 (9%)	3 (2%)	12	57
11	L	153/157 (98%)	136 (89%)	16 (10%)	1 (1%)	26	71
12	M	250/339 (74%)	236 (94%)	12 (5%)	2 (1%)	24	70
13	N	217/364 (60%)	191 (88%)	22 (10%)	4 (2%)	11	55
14	O	207/590 (35%)	193 (93%)	11 (5%)	3 (1%)	14	59
15	P	63/175 (36%)	56 (89%)	7 (11%)	0	100	100
16	R	104/135 (77%)	91 (88%)	12 (12%)	1 (1%)	19	65
17	S	438/687 (64%)	415 (95%)	20 (5%)	3 (1%)	26	71
18	T	557/877 (64%)	532 (96%)	22 (4%)	3 (0%)	34	77
19	a	131/251 (52%)	123 (94%)	8 (6%)	0	100	100
20	c	97/382 (25%)	87 (90%)	8 (8%)	2 (2%)	9	52
21	o	305/455 (67%)	251 (82%)	46 (15%)	8 (3%)	7	48

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
23	y	77/215 (36%)	76 (99%)	1 (1%)	0	100	100
24	b	76/196 (39%)	74 (97%)	2 (3%)	0	100	100
24	k	76/196 (39%)	68 (90%)	7 (9%)	1 (1%)	15	61
25	d	80/101 (79%)	73 (91%)	7 (9%)	0	100	100
25	n	80/101 (79%)	74 (92%)	6 (8%)	0	100	100
26	e	71/94 (76%)	62 (87%)	9 (13%)	0	100	100
26	p	71/94 (76%)	66 (93%)	4 (6%)	1 (1%)	14	59
27	f	70/86 (81%)	61 (87%)	9 (13%)	0	100	100
27	q	70/86 (81%)	65 (93%)	5 (7%)	0	100	100
28	g	65/77 (84%)	63 (97%)	2 (3%)	0	100	100
28	r	65/77 (84%)	57 (88%)	8 (12%)	0	100	100
29	h	78/146 (53%)	71 (91%)	7 (9%)	0	100	100
29	l	75/146 (51%)	65 (87%)	8 (11%)	2 (3%)	6	47
30	j	92/110 (84%)	83 (90%)	9 (10%)	0	100	100
30	m	91/110 (83%)	84 (92%)	7 (8%)	0	100	100
31	V	639/1145 (56%)	602 (94%)	34 (5%)	3 (0%)	34	77
32	W	160/238 (67%)	117 (73%)	35 (22%)	8 (5%)	3	32
33	Y	82/111 (74%)	76 (93%)	6 (7%)	0	100	100
35	s	106/175 (61%)	92 (87%)	8 (8%)	6 (6%)	2	28
36	t	426/503 (85%)	417 (98%)	9 (2%)	0	100	100
36	u	112/503 (22%)	104 (93%)	8 (7%)	0	100	100
36	v	114/503 (23%)	108 (95%)	3 (3%)	3 (3%)	7	48
36	w	423/503 (84%)	414 (98%)	7 (2%)	2 (0%)	34	77
All	All	9407/14756 (64%)	8598 (91%)	725 (8%)	84 (1%)	26	67

5 of 84 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
6	A	487	ASN
6	A	742	VAL
6	A	1620	TYR
7	C	568	SER
7	C	901	GLU



### 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	
6	A	1577/2182 (72%)	1549 (98%)	28 (2%)	66	87
7	C	681/910 (75%)	662 (97%)	19 (3%)	51	80
8	H	366/538 (68%)	357 (98%)	9 (2%)	55	82
9	J	299/397 (75%)	280 (94%)	19 (6%)	22	61
10	K	159/328 (48%)	152 (96%)	7 (4%)	35	72
11	L	112/141 (79%)	105 (94%)	7 (6%)	22	62
12	M	214/296 (72%)	210 (98%)	4 (2%)	65	86
13	N	211/332 (64%)	205 (97%)	6 (3%)	51	80
14	O	187/525 (36%)	181 (97%)	6 (3%)	46	78
15	P	56/151 (37%)	54 (96%)	2 (4%)	42	76
16	R	25/121 (21%)	24 (96%)	1 (4%)	38	73
17	S	230/633 (36%)	221 (96%)	9 (4%)	39	74
18	T	1/786 (0%)	1 (100%)	0	100	100
19	a	125/225 (56%)	125 (100%)	0	100	100
20	c	71/346 (20%)	68 (96%)	3 (4%)	36	72
21	o	256/413 (62%)	247 (96%)	9 (4%)	43	77
23	y	76/193 (39%)	74 (97%)	2 (3%)	54	81
24	b	70/176 (40%)	70 (100%)	0	100	100
24	k	70/176 (40%)	69 (99%)	1 (1%)	74	89
25	d	69/89 (78%)	69 (100%)	0	100	100
25	n	69/89 (78%)	66 (96%)	3 (4%)	35	72
26	e	65/83 (78%)	63 (97%)	2 (3%)	47	79
26	p	65/83 (78%)	60 (92%)	5 (8%)	16	55
27	f	63/77 (82%)	63 (100%)	0	100	100
27	q	63/77 (82%)	61 (97%)	2 (3%)	46	78
28	g	58/66 (88%)	57 (98%)	1 (2%)	68	88

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
28	r	57/66 (86%)	56 (98%)	1 (2%)	66	87
29	h	77/129 (60%)	76 (99%)	1 (1%)	76	90
29	l	73/129 (57%)	63 (86%)	10 (14%)	4	29
30	j	79/103 (77%)	79 (100%)	0	100	100
30	m	77/103 (75%)	74 (96%)	3 (4%)	39	74
All	All	5601/9963 (56%)	5441 (97%)	160 (3%)	54	79

5 of 160 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	J	443	ASN
12	M	13	VAL
30	m	82	LYS
9	J	450	ARG
10	K	202	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 59 such sidechains are listed below:

Mol	Chain	Res	Type
7	C	830	ASN
10	K	195	ASN
29	l	14	GLN
8	H	320	ASN
9	J	443	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	I	32/95 (33%)	14 (43%)	4 (12%)
2	E	15/20 (75%)	5 (33%)	1 (6%)
3	2	146/1175 (12%)	49 (33%)	5 (3%)
34	3	2/3 (66%)	2 (100%)	0
4	6	98/112 (87%)	37 (37%)	3 (3%)
5	5	138/179 (77%)	61 (44%)	3 (2%)
All	All	431/1584 (27%)	168 (38%)	16 (3%)

5 of 168 RNA backbone outliers are listed below:



Mol	Chain	Res	Type
1	I	2	U
1	I	3	A
1	I	4	U
1	I	10	A
1	I	11	A

5 of 16 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
3	2	1123	C
3	2	1124	U
4	6	92	C
3	2	17	U
5	5	27	G

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

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 13 ligands modelled in this entry, 11 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	$\# Z  > 2$
39	I6P	A	3001	-	30,36,36	0.66	0	60,60,60	1.17	6 (10%)
40	GTP	C	1101	-	26,34,34	1.06	2 (7%)	29,54,54	1.95	8 (27%)

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
39	I6P	A	3001	-	-	0/30/54/54	0/1/1/1
40	GTP	C	1101	-	-	0/18/38/38	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
40	C	1101	GTP	C5-C4	2.68	1.46	1.40
40	C	1101	GTP	C6-C5	3.34	1.48	1.41

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
40	C	1101	GTP	C1'-N9-C4	-3.90	122.45	126.81
40	C	1101	GTP	C6-C5-C4	-3.78	116.54	120.86
40	C	1101	GTP	N3-C2-N1	-3.71	122.51	127.56
40	C	1101	GTP	C5-C6-N1	-3.49	118.96	123.52
39	A	3001	I6P	O13-P3-O33	-2.46	101.61	107.48

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
22	X	1
18	T	1
31	V	1



All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	X	27:UNK	C	86:UNK	N	8.48
1	V	903:MET	C	904:ASP	N	2.35
1	T	300:ASP	C	301:LYS	N	2.16