



## Full wwPDB EM Validation Report ⓘ

Nov 20, 2022 – 02:49 pm GMT

PDB ID : 6QDV  
EMDB ID : EMD-4525  
Title : Human post-catalytic P complex spliceosome  
Authors : Fica, S.M.; Oubridge, C.; Wilkinson, M.E.; Newman, A.J.; Nagai, K.  
Deposited on : 2019-01-03  
Resolution : 3.30 Å (reported)

This is a Full wwPDB EM Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

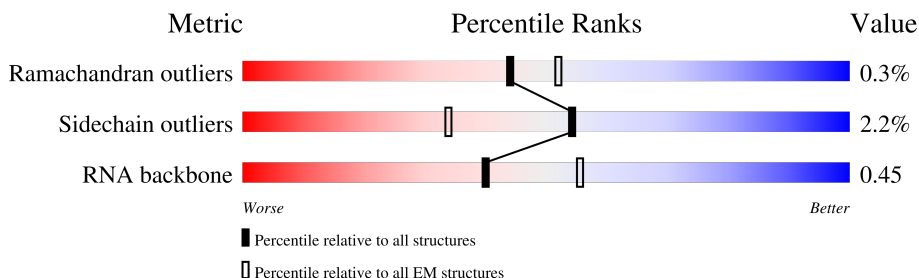
EMDB validation analysis : 0.0.1.dev43  
Mogul : 1.8.4, CSD as541be (2020)  
MolProbity : 4.02b-467  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : 1.9.9  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.31.2

# 1 Overall quality at a glance i

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

The reported resolution of this entry is 3.30 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  $\geq 3$ , 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  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the EM map (all-atom inclusion  $< 40\%$ ). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	2	189	
2	5	116	
3	6	106	
4	7	390	
5	8	91	
6	9	144	
7	A	2335	
8	B	1722	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	C	899	12% 99%
10	D	123	54% 99%
11	E	14	7% 57% 29% 14%
12	F	122	18% 100%
13	G	60	17% 97%
14	H	908	30% 50% 49%
15	I	113	9% 19% 15% 63%
16	J	320	100%
17	K	295	26% 99%
18	L	144	6% 96%
19	M	289	39% 99%
20	N	306	15% 99%
21	O	802	27% 55% 45%
22	P	229	8% 46% 54%
23	R	26	92% 8%
24	S	848	37% 66% 33%
25	T	855	73% 73% 25%
26	U	1485	89% 86% 11%
27	V	1220	55% 55% 42%
28	W	162	99% 99%
29	Y	92	100% 100%
30	Z	30	7% 100%
31	b	82	59% 83% 17%
31	k	82	90% 100%
32	c	586	7% 46% 54%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
33	d	84	39% 100%
33	n	84	49% 99%
34	e	81	96% 98%
34	p	81	88% 100%
35	f	72	99% 99%
35	q	72	97% 100%
36	g	73	81% 100%
36	r	73	59% 99%
37	h	80	94% 100%
37	l	80	96% 100%
38	i	164	22% 100%
39	j	118	81% 81% 19%
39	m	118	81% 81% 19%
40	o	513	12% 99%
41	s	225	73% 74% 25%
42	t	504	25% 25% 75%
42	u	504	23% 23% 77%
42	v	504	25% 25% 75%
42	w	504	23% 23% 77%
43	y	144	20% 99%
44	z	34	100%

## 2 Entry composition

There are 50 unique types of molecules in this entry. The entry contains 121152 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 U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	2	120	2535	1135	428	852	120	0	0

- Molecule 2 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	5	75	1579	708	264	532	75	0	0

- Molecule 3 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
3	6	97	2075	928	381	669	97	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	7	390	3130	1976	546	589	19	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
7	?	-	GLU	deletion	UNP P38919
7	405	MET	-	expression tag	UNP P38919
7	406	ASN	-	expression tag	UNP P38919
7	407	VAL	-	expression tag	UNP P38919
7	408	ALA	-	expression tag	UNP P38919
7	409	ASP	-	expression tag	UNP P38919
7	410	LEU	-	expression tag	UNP P38919
7	411	ILE	-	expression tag	UNP P38919

- Molecule 5 is a protein called RNA-binding protein 8A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	8	91	730	463	122	142	3	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
8	151	VAL	-	expression tag	UNP Q9Y5S9
8	152	ARG	-	expression tag	UNP Q9Y5S9
8	153	GLY	-	expression tag	UNP Q9Y5S9
8	154	PRO	-	expression tag	UNP Q9Y5S9

- Molecule 6 is a protein called Protein mago nashi homolog 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	9	144	1196	772	200	221	3	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
9	146	LYS	-	expression tag	UNP Q96A72
9	147	PRO	-	expression tag	UNP Q96A72
9	148	ILE	-	expression tag	UNP Q96A72

- Molecule 7 is a protein called Pre-mRNA-processing-splicing factor 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	A	2250	18655	12009	3256	3309	81	0	0

- Molecule 8 is a protein called U5 small nuclear ribonucleoprotein 200 kDa helicase.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
8	B	1722	13846	8848	2369	2557	72	0	0

- Molecule 9 is a protein called 116 kDa U5 small nuclear ribonucleoprotein component.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
9	C	899	7116	4553	1184	1345	34	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	?	-	SER	deletion	UNP Q15029

- Molecule 10 is a protein called PRKR-interacting protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	D	123	1013	635	193	180	5	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	140	LEU	-	expression tag	UNP Q9H875
D	141	ALA	-	expression tag	UNP Q9H875
D	142	LYS	-	expression tag	UNP Q9H875

- Molecule 11 is a RNA chain called Ligated exons: MINX mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
11	E	14	296	132	52	98	14	0	0

- Molecule 12 is a protein called Cactin.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
12	F	122	1084	712	197	173	2	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	654	ALA	GLU	conflict	UNP Q8WUQ7

- Molecule 13 is a protein called Protein FAM32A.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	G	60	Total	C	N	O	S	0	0
			504	314	96	92	2		

- Molecule 14 is a protein called Pre-mRNA-splicing factor CWC22 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	H	459	Total	C	N	O	S	0	0
			3713	2380	634	678	21		

- Molecule 15 is a RNA chain called Intron lariat: MINX RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	I	42	Total	C	N	O	P	0	0
			872	390	148	292	42		

- Molecule 16 is a protein called Pleiotropic regulator 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	J	320	Total	C	N	O	S	0	0
			2523	1594	457	464	8		

- Molecule 17 is a protein called SNW domain-containing protein 1.

Mol	Chain	Residues	Atoms						AltConf	Trace
17	K	295	Total	C	N	O	P	S	0	0
			2360	1479	431	435	2	13		

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	?	-	GLU	deletion	UNP Q13573
K	330	LYS	-	expression tag	UNP Q13573
K	331	ALA	-	expression tag	UNP Q13573
K	332	ARG	-	expression tag	UNP Q13573
K	333	GLU	-	expression tag	UNP Q13573
K	334	ARG	-	expression tag	UNP Q13573
K	335	ARG	-	expression tag	UNP Q13573

- Molecule 18 is a protein called Protein BUD31 homolog.



Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
18	L	144	1188	748	218	210	12	0	0

- Molecule 19 is a protein called Pre-mRNA-splicing factor RBM22.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
19	M	289	2318	1455	416	428	19	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	143	ALA	THR	conflict	UNP Q9NW64
M	144	ALA	SER	conflict	UNP Q9NW64
M	145	ALA	ASP	conflict	UNP Q9NW64
M	146	ALA	MET	conflict	UNP Q9NW64

- Molecule 20 is a protein called U5 small nuclear ribonucleoprotein 40 kDa protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
20	N	306	2394	1501	422	457	14	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
N	?	-	ALA	deletion	UNP Q96DI7

- Molecule 21 is a protein called Cell division cycle 5-like protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
21	O	441	3416	2116	648	639	13	0	0

- Molecule 22 is a protein called Spliceosome-associated protein CWC15 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
22	P	106	888	544	174	168	2	0	0

- Molecule 23 is a protein called Serine/arginine repetitive matrix protein 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
23	R	26	193	120	36	36	1	0	0

- Molecule 24 is a protein called Crooked neck-like protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
24	S	570	3965	2482	740	737	6	0	0

- Molecule 25 is a protein called Pre-mRNA-splicing factor SYF1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
25	T	639	4003	2479	748	763	13	0	0

- Molecule 26 is a protein called Intron-binding protein aquarius.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
26	U	1322	10885	6989	1879	1963	54	4	0

- Molecule 27 is a protein called ATP-dependent RNA helicase DHX8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
27	V	713	2995	1538	722	734	1	0	0

- Molecule 28 is a protein called U2 small nuclear ribonucleoprotein A'.

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

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
W	163	LYS	-	expression tag	UNP P09661

- Molecule 29 is a protein called U2 small nuclear ribonucleoprotein B'.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
29	Y	92	745	480	130	130	5	0	0

- Molecule 30 is a protein called NF-kappa-B-activating protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
30	Z	30	230	140	43	45	2	0	0

- Molecule 31 is a protein called Small nuclear ribonucleoprotein-associated proteins B and B'.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
31	b	68	545	347	95	96	7	0	0
31	k	82	664	419	121	117	7	0	0

- Molecule 32 is a protein called Pre-mRNA-splicing factor SLU7.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
32	c	269	2215	1392	397	418	8	0	0

- Molecule 33 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
33	d	84	658	412	116	124	6	0	0
33	n	83	652	409	115	122	6	0	0

- Molecule 34 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
34	e	79	651	413	115	118	5	0	0
34	p	81	669	424	119	121	5	0	0

- Molecule 35 is a protein called Small nuclear ribonucleoprotein F.

Mol	Chain	Residues	Atoms					AltConf	Trace
35	f	72	Total	C	N	O	S	0	0
			562	364	93	100	5		
35	q	72	Total	C	N	O	S	0	0
			562	364	93	100	5		

- Molecule 36 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					AltConf	Trace
36	g	73	Total	C	N	O	S	0	0
			568	358	102	102	6		
36	r	73	Total	C	N	O	S	0	0
			568	358	102	102	6		

- Molecule 37 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					AltConf	Trace
37	h	80	Total	C	N	O	S	0	0
			634	404	111	115	4		
37	l	80	Total	C	N	O	S	0	0
			634	404	111	115	4		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
h	79	LEU	-	expression tag	UNP P62314
h	80	LEU	-	expression tag	UNP P62314
l	79	LEU	-	expression tag	UNP P62314
l	80	LEU	-	expression tag	UNP P62314

- Molecule 38 is a protein called Peptidyl-prolyl cis-trans isomerase-like 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
38	i	164	Total	C	N	O	S	0	0
			1270	810	220	233	7		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
i	165	SER	-	expression tag	UNP Q9Y3C6
i	166	GLY	-	expression tag	UNP Q9Y3C6

- Molecule 39 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
39	j	95	Total	C	N	O	S	0	0
			774	486	141	142	5		
39	m	95	Total	C	N	O	S	0	0
			774	486	141	142	5		

- Molecule 40 is a protein called Pre-mRNA-processing factor 17.

Mol	Chain	Residues	Atoms					AltConf	Trace
40	o	513	Total	C	N	O	S	0	0
			4157	2643	719	771	24		

- Molecule 41 is a protein called Pre-mRNA-splicing factor SPF27.

Mol	Chain	Residues	Atoms					AltConf	Trace
41	s	169	Total	C	N	O	S	0	0
			1402	872	257	264	9		

- Molecule 42 is a protein called Pre-mRNA-processing factor 19.

Mol	Chain	Residues	Atoms					AltConf	Trace
42	t	125	Total	C	N	O	S	0	0
			988	618	176	190	4		
42	u	118	Total	C	N	O	S	0	0
			938	586	167	181	4		
42	v	125	Total	C	N	O	S	0	0
			988	618	176	190	4		
42	w	118	Total	C	N	O	S	0	0
			938	586	167	181	4		

- Molecule 43 is a protein called Pre-mRNA-splicing factor SYF2.

Mol	Chain	Residues	Atoms					AltConf	Trace
43	y	144	Total	C	N	O	S	0	0
			1218	758	225	233	2		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
y	240	GLY	-	expression tag	UNP O95926
y	241	THR	-	expression tag	UNP O95926
y	242	ALA	-	expression tag	UNP O95926
y	243	VAL	-	expression tag	UNP O95926

- Molecule 44 is a protein called Replication stress response regulator SDE2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
44	z	34	280	166	59	53	2	0	0

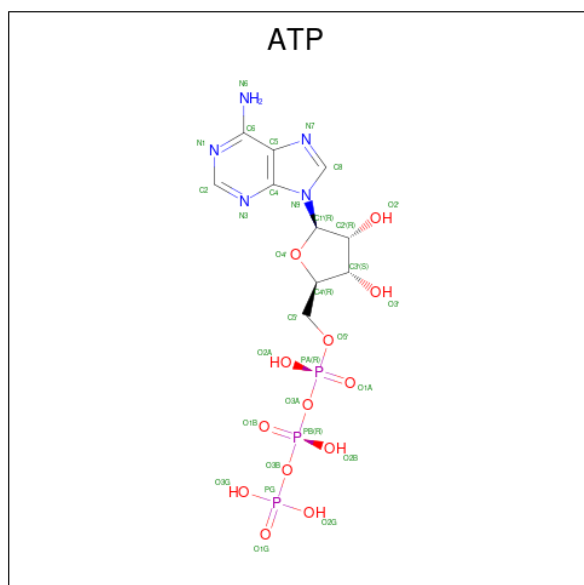
- Molecule 45 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
45	6	5	Total	Mg	0
			5	5	
45	7	1	Total	Mg	0
			1	1	
45	C	1	Total	Mg	0
			1	1	

- Molecule 46 is POTASSIUM ION (three-letter code: K) (formula: K).

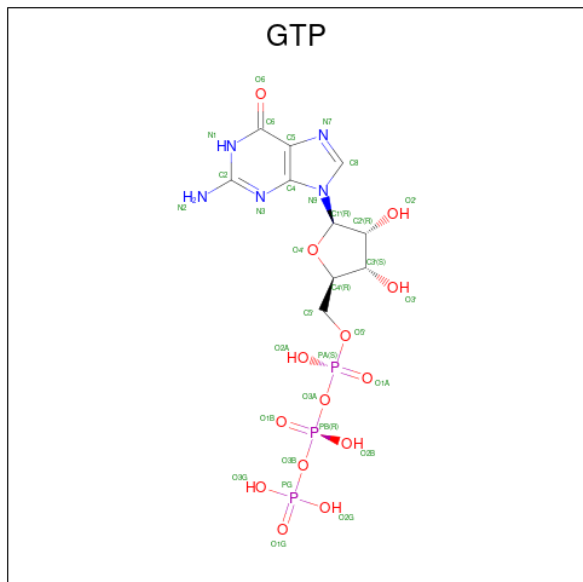
Mol	Chain	Residues	Atoms		AltConf
46	6	1	Total	K	0
			1	1	

- Molecule 47 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>5</sub>O<sub>13</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
47	7	1	31	10	5	13	3	0

- Molecule 48 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula:  $C_{10}H_{16}N_5O_{14}P_3$ ).

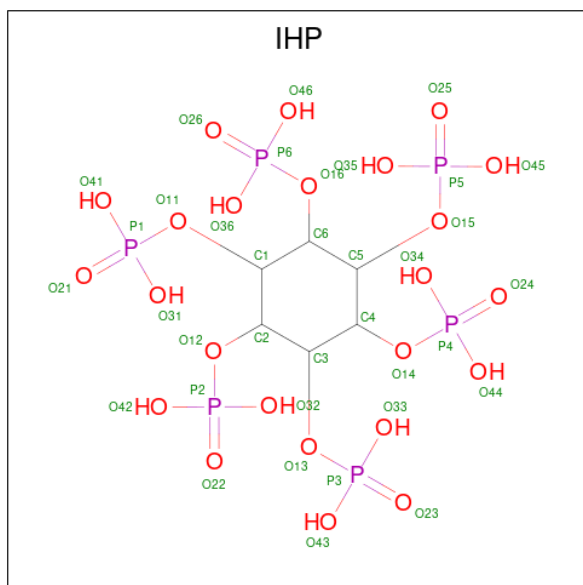


Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
48	C	1	32	10	5	14	3	0

- Molecule 49 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
			Total	Zn	
49	L	3	3	3	0
49	M	3	3	3	0
49	c	1	1	1	0

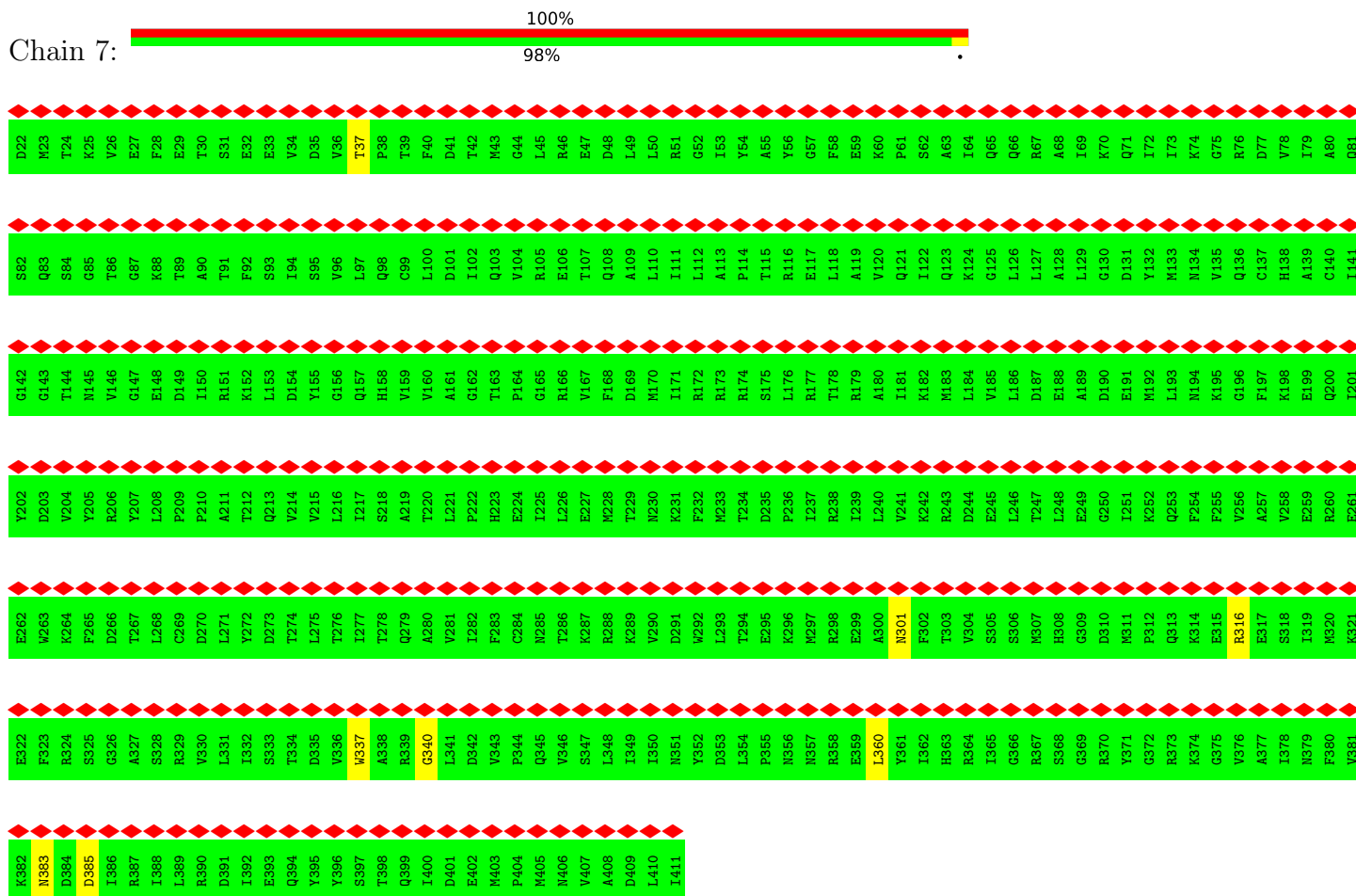
- Molecule 50 is INOSITOL HEXAKISPHOSPHATE (three-letter code: IHP) (formula:  $C_6H_{18}O_{24}P_6$ ).



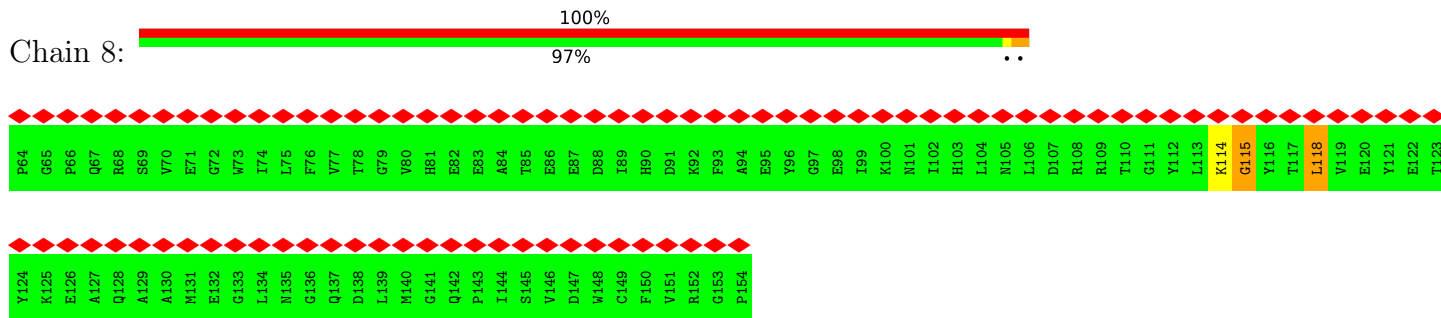
Mol	Chain	Residues	Atoms				AltConf
			Total	C	O	P	
50	c	1	36	6	24	6	0



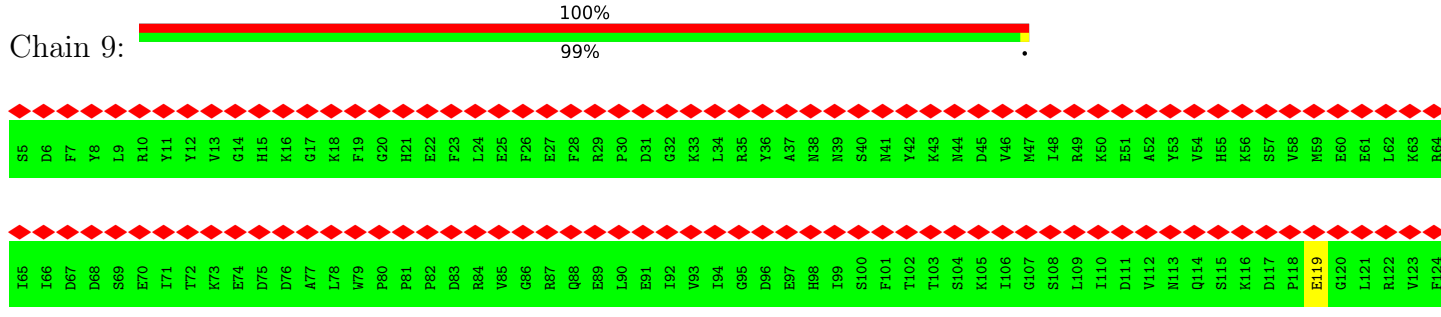


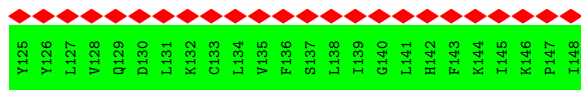


• Molecule 5: RNA-binding protein 8A

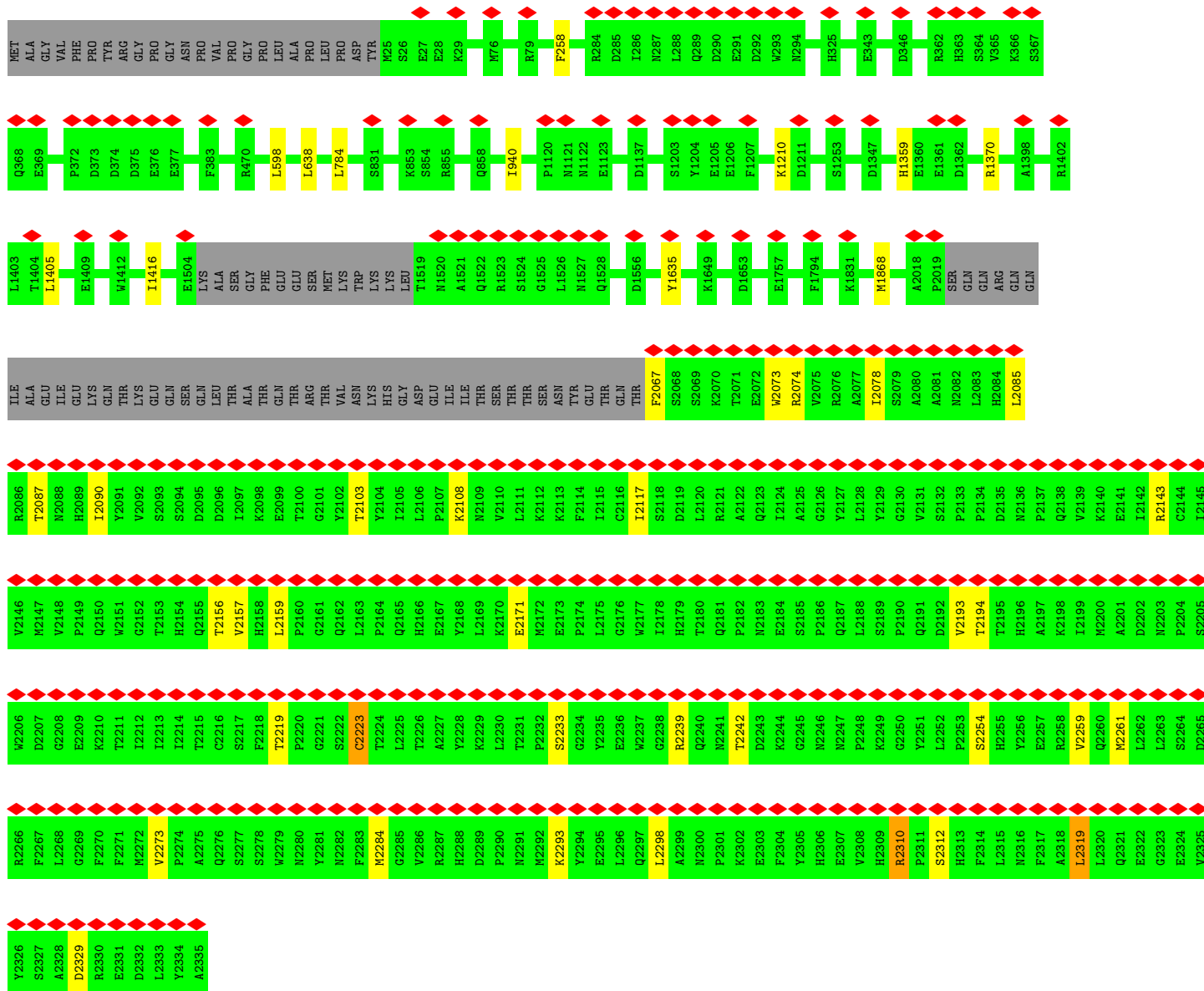


• Molecule 6: Protein mago nashi homolog 2

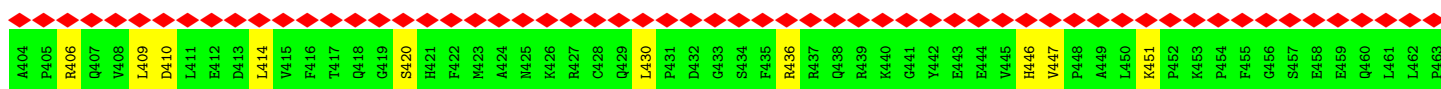
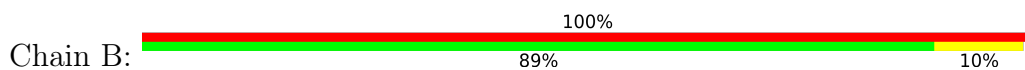




• Molecule 7: Pre-mRNA-processing-splicing factor 8

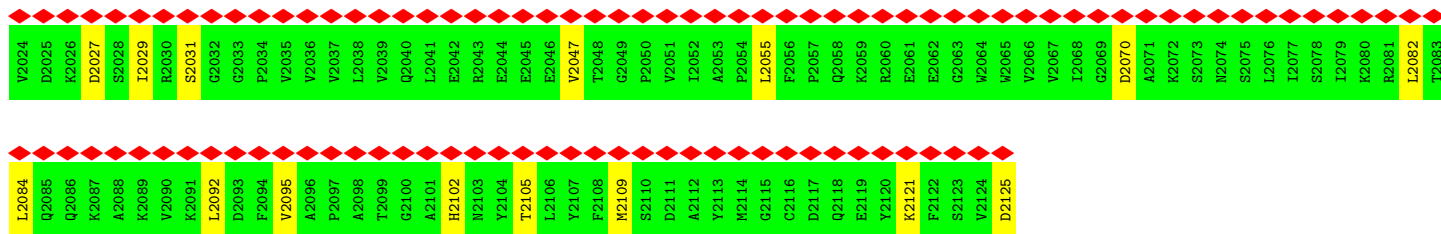


• Molecule 8: U5 small nuclear ribonucleoprotein 200 kDa helicase

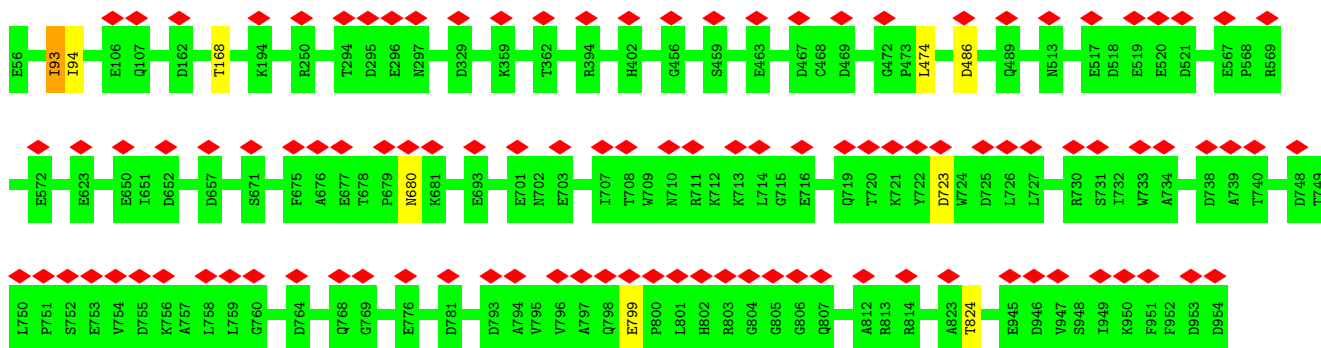


V464	E465	K466	L467	P468	K469	Y470	A471	Q472	A473	G474	F475	E476	G477	F478	K479	T480	L481	N482	R483	I484	Q485	S486	K487	L488	Y489	R490	A491	A492	L493	E494	T495	D496	E497	N498	L499	L500	L501	C502	A503	P504	T505	G506	A507	G508	K509	T510	N511	E512	V513	L514	M515	C516	L517	L518	R519	E520	I521	G522	K523
H524	I525	N526	M527	D528	G529	T530	I531	V532	V533	D534	D535	F536	K537	I538	I539	Y540	I541	A542	P543	L544	R545	S546	L547	V548	G549	E550	M551	V552	G553	S554	F555	G556	K557	R558	L559	A560	T561	Y562	G563	I564	T565	V566	A567	E568	K569	T570	N571	D572	V573	L574	M575	C576	K577	E578	E579	I580	E581	G582	T583
Q584	I585	I586	V587	C588	T589	P590	E591	K592	M593	D594	I595	I596	N597	R598	K599	G600	G601	E602	R603	T604	Y605	T606	Q607	L608	V609	R610	L611	I612	G613	L614	D615	E616	L617	H618	L619	L620	G621	D622	D623	G624	G625	P626	V627	L628	E629	A630	L631	G632	V633	H634	A635	I636	R637	M638	E639	E640	M641	T642	Q643
E644	D645	V646	R647	L648	I649	G650	L651	S652	A653	T654	L655	P656	N657	Y658	E659	D660	V661	A662	T663	F664	L665	S666	V667	D668	P669	A670	K671	G672	L673	F674	Y675	F676	D677	M678	S679	F680	L681	P682	V683	L684	P685	E686	T688	G689	N690	G691	I692	V693	E694	K695	K696	A697	I698	K699	R700	M641	Q702	I703	
M704	N705	E706	I707	V708	Y709	E710	K711	I712	M713	E714	H715	A716	G717	K718	N719	Q720	V721	L722	V723	F724	Y725	H726	S727	K728	E729	E730	T731	G732	K733	T734	A735	R736	I738	R739	D740	M741	C742	L743	E744	K745	D746	T747	L748	G749	L750	F751	L752	L753	E754	G755	F756	A757	S758	T759	E760	L762	R763		
T764	E765	A766	E767	Q768	C769	K770	N771	L772	E773	L774	K775	D776	L777	L778	P779	Y780	G781	F782	A783	I784	H785	H786	A787	G788	M789	T790	R791	V792	D793	R794	T795	L796	V797	E798	P799	L800	F801	A802	D803	K804	H805	I806	Q807	V808	L809	N810	S811	T812	A813	T814	L815	A816	M817	G818	N819	N820	P822	A823	
H824	T825	V826	I827	L828	K829	G830	T831	Q832	H833	L834	S835	P836	E837	K838	G839	R840	M841	T842	E843	L844	G845	A846	L847	D848	I849	L850	Q851	M852	L853	G854	R855	A856	G857	R858	P859	Q860	D862	T863	K864	G865	E866	Q867	L868	T870	T871	S872	H873	G874	E875	L876	Q877	Y878	V879	L880	L882	L883			
H884	Q885	Q886	L887	P888	L889	E890	S891	Q892	M893	Y894	S895	K896	L897	P898	D899	N900	L901	N902	A903	E904	L905	Y906	L907	G908	N909	Y910	Q911	N912	A913	K914	D915	A916	V917	N918	Y919	L920	Y922	A923	Y924	L925	Y926	I927	R928	N929	L930	R931	S932	P933	T934	L935	Y936	Q937	I938	S939	H940	D942	L943		
K944	G945	D946	P947	L948	L949	D950	Q951	R952	R953	L954	D955	L956	V957	H958	T959	A960	A961	L962	M963	L964	D965	K966	N967	N968	L969	V970	K971	Y972	D973	K974	K975	T976	G977	N978	F979	Q980	G981	T982	E983	L984	G985	R986	I987	A988	S989	H990	Y991	Y992	I993	T994	N995	D996	T997	V998	Q999	T1000	Y1001	M1002	Q1003
L1004	L1005	K1006	P1007	T1008	L1009	L1010	E1011	L1012	E1013	L1014	F1015	R1016	V1017	F1018	S1019	L1020	S1021	S1022	E1023	F1024	K1025	M1026	I1027	T1028	V1029	R1030	E1031	E1032	E1033	K1034	L1035	E1036	Q1038	K1039	L1040	L1041	E1042	R1043	V1044	P1045	L1046	P1047	V1048	L1049	E1050	S1051	L1052	E1053	E1054	P1055	S1056	L1057	K1058	T1059	M1060	L1062	L1063		
Q1064	A1065	F1066	I1067	S1068	Q1069	L1070	K1071	L1072	E1073	G1074	F1075	A1076	L1077	M1078	A1079	D1080	M1081	V1082	Y1083	L1084	T1085	Q1086	S1087	G1088	O1089	R1090	L1091	M1092	R1093	A1094	I1095	F1096	E1097	I1098	V1099	L1100	N1101	R1102	G1103	W1104	A1105	Q1106	L1107	T1108	D1109	K1110	T1111	L1112	L1114	C1115	K1116	L1117	I1118	D1119	K1120	R1121	M1122	W1123	
Q1124	S1125	M1126	C1127	P1128	L1129	R1130	Q1131	F1132	R1133	K1134	L1135	P1136	E1137	E1138	V1139	V1140	K1141	I1143	E1144	K1145	K1146	M1147	F1148	P1149	F1150	E1151	R1152	L1153	Y1154	D1155	L1156	M1157	H1158	M1159	E1160	I1161	G1162	E1163	L1164	I1165	R1166	M1167	P1168	K1169	M1170	G1171	K1172	T1173	I1174	H1175	K1176	Y1177	V1178	H1179	L1180	F1181	P1182	K1183	
L1184	E1185	L1186	S1187	V1188	H1189	L1190	Q1191	F1192	I1193	T1194	R1195	S1196	L1197	L1198	K1199	V1200	E1201	L1202	T1203	I1204	T1205	P1206	D1207	F1208	Q1209	W1210	D1211	E1212	K1213	V1214	H1215	G1216	S1217	S1218	E1219	A1220	F1221	W1222	I1223	L1224	V1225	E1226	D1227	V1228	D1229	S1230	E1231	V1232	I1233	L1234	H1235	H1236	E1237	Y1238	L1239	L1240	L1241	K1242	A1243

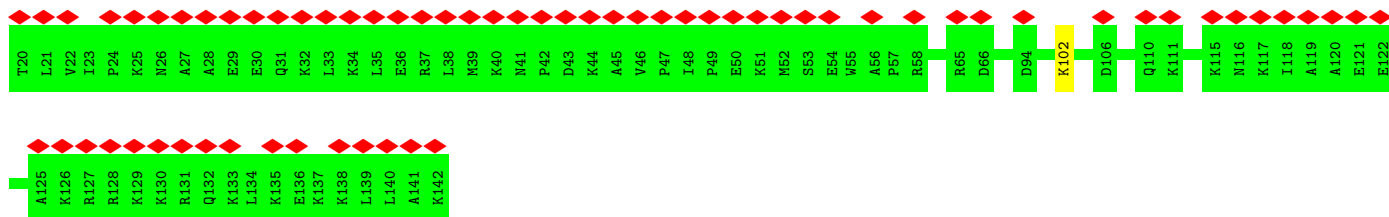
P1964	H1965	F1966	T1967	S1968	H1969	I1971	R1972	K1973	C1974	T1975	D1976	K1977	G1978	V1979	E1980	S1981	V1982	F1983	D1984	M1985	M1986	V1987	M1988	E1989	D1990	E1991	E1992	R1993	M1994	A1995	L1996	L1997	Q1998	L1999	T2000	D2001	S2002	Q2003	I2004	A2005	D2006	V2007	A2008	F2010	C2011	M2012	R2013	V2014	P2015	M2016	I2017	E2018	L2019	S2020	K1960	K1961	Q1962	L1963											
L1904	S1905	A1906	E1907	L1908	Q1909	S1910	D1911	T1912	E1913	E1914	L1915	L1916	S1917	K1918	A1919	I1920	R1921	L1922	I1923	Q1924	A1925	C1926	V1927	I1928	V1929	L1930	S1931	S1932	G1933	G1934	M1935	L1936	S1937	P1938	A1939	L1940	A1941	A1942	M1943	E1944	L1945	A1946	Q1947	M1948	V1949	T1950	Q1951	M1952	M1953	A1954	S1955	K1956	D1957	S1958	Y1959	L1960	K1961	Q1962	L1963										
H1784	L1785	S1786	E1787	L1788	V1789	E1790	Q1791	T1792	L1793	S1794	D1795	L1796	E1797	Q1798	S1799	K1800	I1801	I1802	S1803	I1804	E1805	D1806	E1807	M1808	I1809	Q1810	A1811	P1812	L1813	M1814	L1815	G1816	M1817	I1818	I1818	A1819	A1820	I1821	Y1822	Y1823	I1824	M1825	Y1826	T1827	I1828	E1829	E1830	L1831	F1832	S1833	M1834	S1835	L1836	M1837	A1838	K1839	T1840	V1842	R1843										
G1844	L1845	I1846	E1847	I1848	I1849	S1850	M1851	A1852	A1853	E1854	Y1855	E1856	N1857	I1858	P1859	I1860	R1861	H1862	H1863	E1864	D1865	N1866	L1867	L1868	R1869	Q1870	L1871	A1872	Q1873	K1874	L1875	P1876	L1877	K1878	L1878	L1879	L1880	N1881	P1882	F1883	F1884	N1885	D1886	P1887	H1888	V1889	K1890	L1891	M1892	L1893	L1894	L1895	Q1896	A1897	H1898	Y1899	L1900	S1901	M1902	Q1903									
H1774	E1775	S1776	H1777	L1778	Y1779	H1780	T1781	M1782	L1783	Q1784	H1785	F1786	N1787	A1788	E1789	I1790	I1791	Y1792	D1793	I1794	V1795	E1796	M1797	E1798	R1799	D1800	D1801	V1802	I1803	Y1804	L1805	F1806	D1807	S1808	E1809	F1810	E1811	C1812	V1813	I1814	M1815	C1816	Q1817	R1818	S1819	K1820	L1821	M1822	M1823	L1824	M1825	Q1826	Q1827	R1828	L1829	L1830	L1831	L1832	L1833	L1834	L1835	L1836	L1837	L1838	L1839	L1840	L1841	L1842	L1843
L1604	S1605	D1606	S1607	T1608	L1609	K1610	E1611	T1612	L1613	L1614	M1615	G1616	L1617	G1618	V1619	L1620	H1621	E1622	G1623	L1624	S1625	P1626	M1627	Q1628	R1629	R1630	L1631	V1632	E1633	Q1634	L1635	F1636	L1637	S1638	G1639	A1640	L1641	Q1642	L1643	V1644	L1645	A1646	L1647	R1648	S1649	L1650	K1651	M1652	G1653	M1654	M1655	V1656	A1657	L1658	H1659	Y1660	L1661	L1662	L1663										
M1664	D1665	T1666	Q1667	Y1668	Y1669	M1670	G1671	K1672	L1673	H1674	A1675	Y1676	Y1677	D1678	V1679	P1680	I1681	Y1682	D1683	V1684	L1685	Q1686	M1687	V1688	G1689	H1690	A1691	M1692	R1693	P1694	L1695	Q1696	D1697	D1698	E1699	G1700	R1701	C1702	V1703	I1704	M1705	C1706	Q1707	G1708	S1709	K1710	K1711	D1712	F1713	F1714	K1715	K1716	F1717	L1718	Y1719	E1720	L1721	L1722	P1723										
V1724	E1725	S1726	H1727	L1728	D1729	H1730	C1731	M1732	H1733	D1734	H1735	F1736	N1737	A1738	E1739	I1740	V1741	T1742	K1743	T1744	I1745	E1746	M1747	K1748	Q1749	D1750	A1751	V1752	D1753	Y1754	L1755	T1756	V1757	T1758	F1759	L1760	Y1761	R1762	R1763	M1764	T1765	Q1766	N1767	P1768	N1769	Y1770	Y1771	M1772	L1773	Q1774	G1775	I1776	S1777	H1778	R1779	H1780	S1782	L1783											
H1784	L1785	S1786	E1787	L1788	V1789	E1790	Q1791	T1792	L1793	S1794	D1795	L1796	E1797	Q1798	S1799	K1800	I1801	I1802	S1803	I1804	E1805	D1806	E1807	M1808	I1809	Q1810	A1811	P1812	L1813	M1814	L1815	G1816	M1817	I1818	I1818	A1819	A1820	I1821	Y1822	Y1823	I1824	M1825	Y1826	T1827	I1828	E1829	E1830	L1831	F1832	S1833	M1834	S1835	L1836	M1837	A1838	K1839	T1840	V1842	R1843										
G1844	L1845	I1846	E1847	I1848	I1849	S1850	M1851	A1852	A1853	E1854	Y1855	E1856	N1857	I1858	P1859	I1860	R1861	H1862	H1863	E1864	D1865	N1866	L1867	L1868	R1869	Q1870	L1871	A1872	Q1873	K1874	L1875	P1876	L1877	K1878	L1878	L1879	L1880	N1881	P1882	F1883	F1884	N1885	D1886	P1887	H1888	V1889	K1890	L1891	M1892	L1893	L1894	L1895	Q1896	A1897	H1898	Y1899	L1900	S1901	M1902	Q1903									
L1904	S1905	A1906	E1907	L1908	Q1909	S1910	D1911	T1912	E1913	E1914	L1915	L1916	S1917	K1918	A1919	I1920	R1921	L1922	I1923	Q1924	A1925	C1926	V1927	I1928	V1929	L1930	S1931	S1932	G1933	G1934	M1935	L1936	S1937	P1938	A1939	L1940	A1941	A1942	M1943	E1944	L1945	A1946	Q1947	M1948	V1949	T1950	Q1951	M1952	M1953	A1954	S1955	K1956	D1957	S1958	Y1959	L1960	K1961	Q1962	L1963										
H1244	Y1245	A1246	Q1247	D1248	E1249	H1250	L1251	T1252	L1253	F1254	F1255	V1256	P1257	V1258	F1259	E1260	P1261	L1262	P1263	L1264	Q1265	Y1266	F1267	I1268	R1269	V1270	V1271	S1272	D1273	R1274	M1275	L1276	S1277	C1278	E1279	T1280	Q1281	L1282	P1283	V1284	S1285	F1286	R1287	H1288	L1289	I1290	L1291	P1292	E1293	K1294	Y1295	P1296	P1297	P1298	T1299	E1300	L1302	D1303											
L1304	Q1305	R1306	L1307	L1308	V1309	S1310	A1311	L1312	R1313	M1314	S1315	A1316	F1317	E1318	F1319	E1320	Y1321	Q1322	D1323	K1324	F1325	P1326	F1327	F1328	M1329	P1330	I1331	Q1332	T1333	Q1334	V1335	F1336	N1337	T1338	V1339	Y1340	N1341	S1342	D1343	D1344	N1345	V1346	F1347	V1348	G1349	A1350	P1351	T1352	G1353	S1354	G1355	K1356	T1357	I1358	C1359	A1360	I1361	F1362	A1363										
I1364	L1365	R1366	M1367	L1368	L1369	Q1370	S1371	L1372	E1373	G1374	R1375	C1376	F1377	Y1378	I1379	T1380	P1381	M1382	E1383	L1384	L1385	A1386	E1387	Q1388	V1389	Y1390	M1391	D1392	W1393	Y1394	E1395	L1396	F1397	L1398	Q1398	D1399	R1400	L1401	M1402	K1403	K1404	V1405	V1406	L1407	L1408	L1409	G1410	E1411	T1412	S1413	T1414	D1415	L1416	K1417	L1418	L1419	G1420	K1421	G1422	M1423									
I1424	I1425	I1426	S1427	T1428	A1429	E1430	K1431	W1432	D1433	I1434	L1435	S1436	R1437	R1438	W1439	K1440	Q1441	R1442	K1443	L1444	V1445	Q1446	I1447	I1448	I1449	L1450	F1451	V1452	W1453	D1454	E1455	V1456	H1457	L1458	I1459	G1460	L1461	E1462	M1463	G1464	V1465	V1466	L1467	E1468	V1469	I1470	G1471	S1472	R1473	M1474	R1475	Y1476	Y1477	S1478	S1479	Q1480	I1481	E1482	R1483										
P1484	I1485	R1486	I1487	V1488	A1489	L1490	S1491	S1492	L1493	L1494	S1495	N1496	A1497	K1498	D1499	V1500	A1501	H1502	W1503	L1504	G1505	C1506	S1507	A1508	I1509	S1510	T1511	F1512	M1513	F1514	H1515	P1516	L1517	V1518	R1519	P1520	P1522	L1523	E1524	L1525	H1526	I1527	Q1528	G1529	F1530	M1531	I1532	S1533	H1534	V1535	Q1536	T1537	R1538	L1539	L1540	S1541	M1542	A1543											
K1544	P1545	V1546	Y1547	H1548	A1549	L1550	T1551	H1552	H1553	S1554	P1555	K1556	P1557	P1558	V1559	L1560	V1561	F1562	V1563	P1564	S1565	R1566	K1567	Q1568	R1569	L1570	L1571	L1572	A1573	I1574	D1575	L1576	L1577	L1578	T1579	C1580	A1581	A1582	D1583	I1584	Q1585	R1586	Q1587	L1588	F1589	L1590	H1591	C1592	L1593	E1594	L1595	D1596	L1597	I1598	P1599	Y1600	L1601	E1602	L1603										



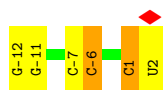
• Molecule 9: 116 kDa U5 small nuclear ribonucleoprotein component



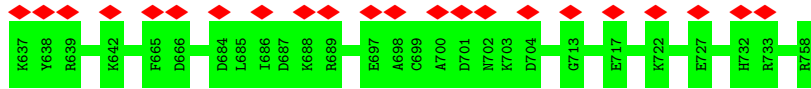
• Molecule 10: PRKR-interacting protein 1



• Molecule 11: Ligated exons: MINX mRNA



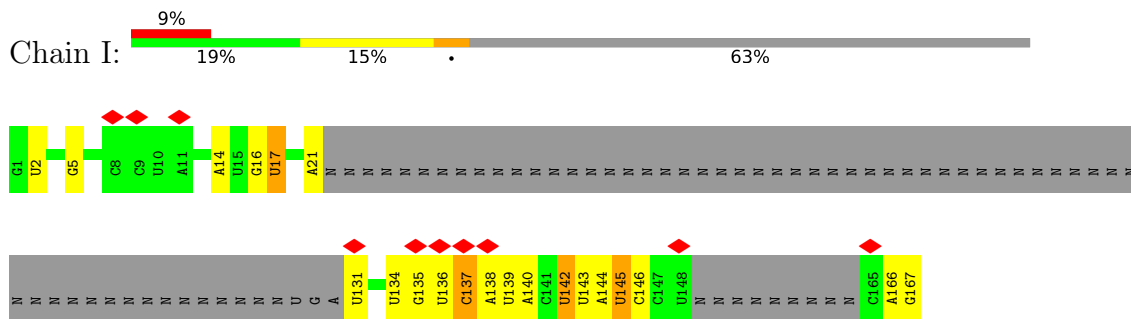
• Molecule 12: Cactin



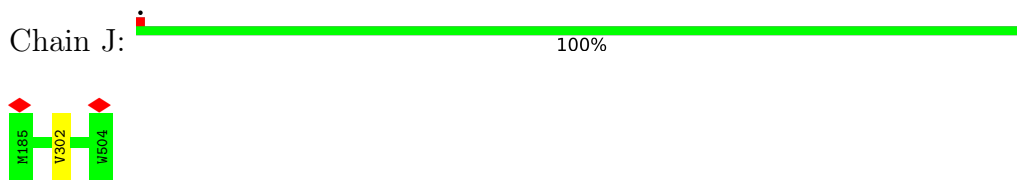
• Molecule 13: Protein FAM32A



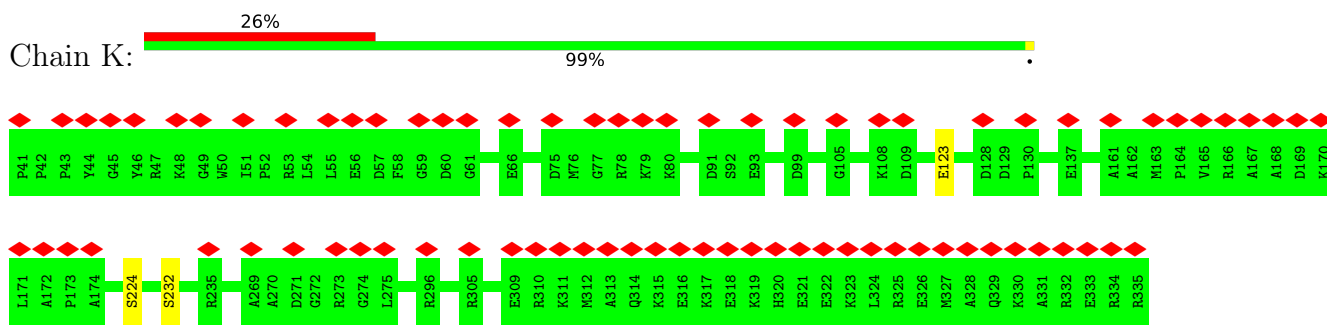
● Molecule 15: Intron lariat: MINX RNA



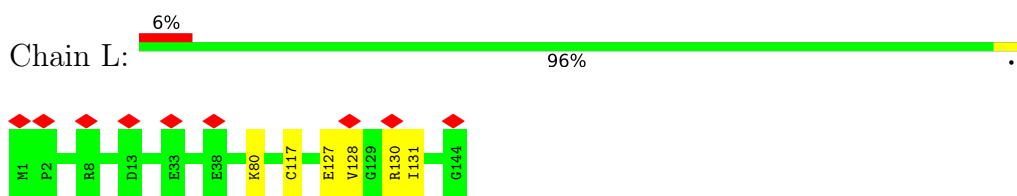
● Molecule 16: Pleiotropic regulator 1



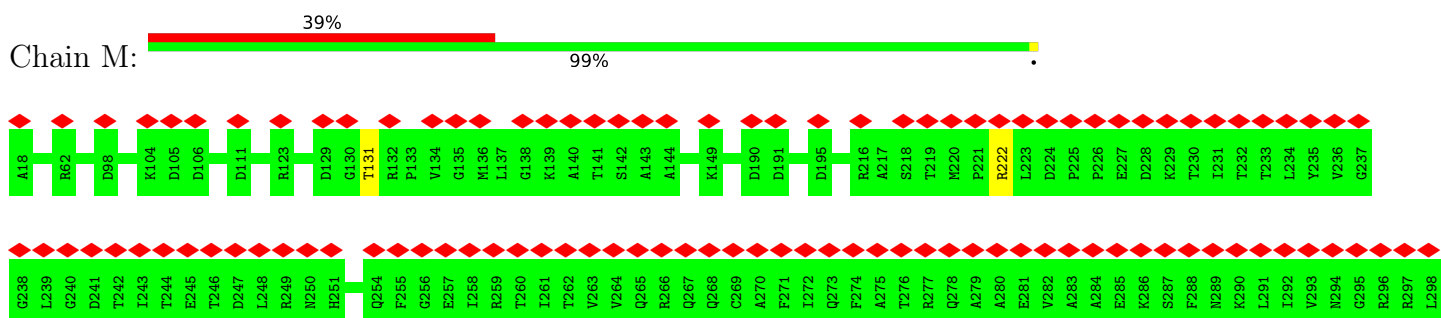
● Molecule 17: SNW domain-containing protein 1



● Molecule 18: Protein BUD31 homolog

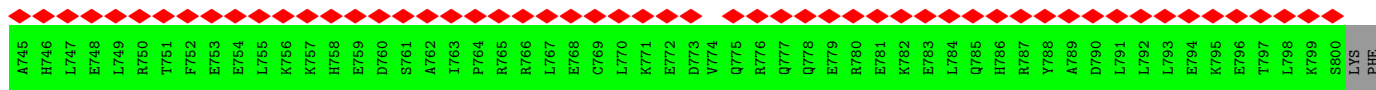


● Molecule 19: Pre-mRNA-splicing factor RBM22

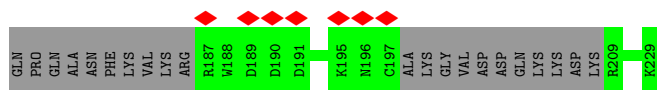
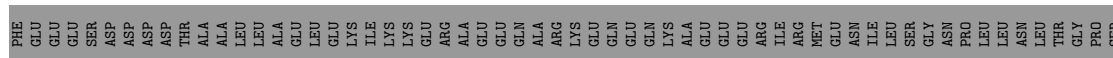
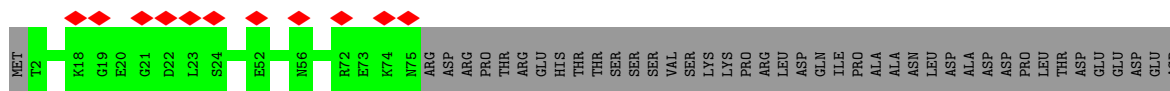




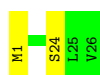




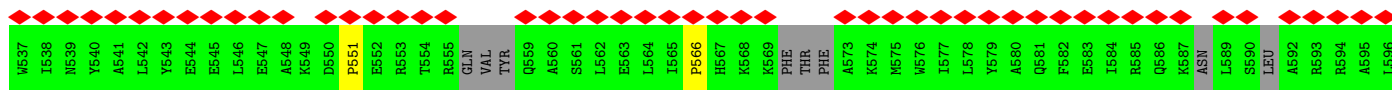
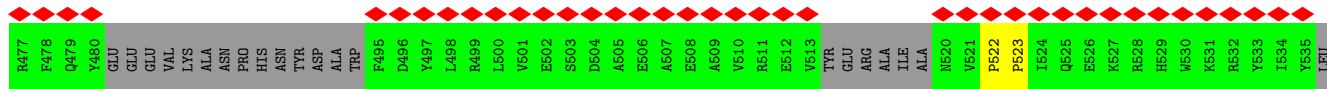
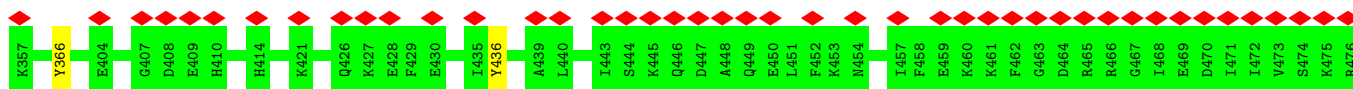
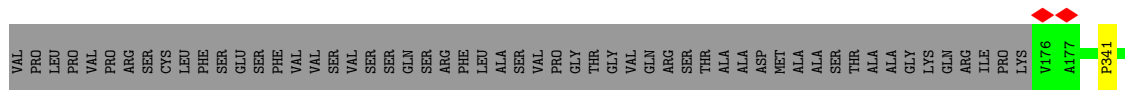
• Molecule 22: Spliceosome-associated protein CWC15 homolog

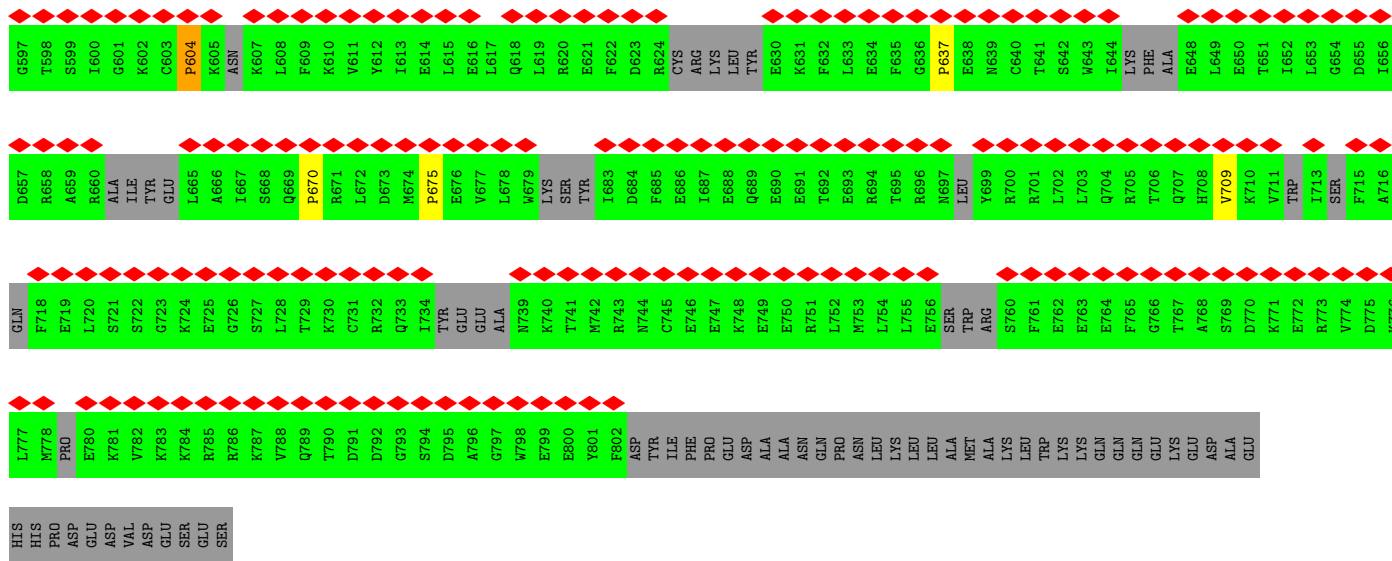


• Molecule 23: Serine/arginine repetitive matrix protein 2

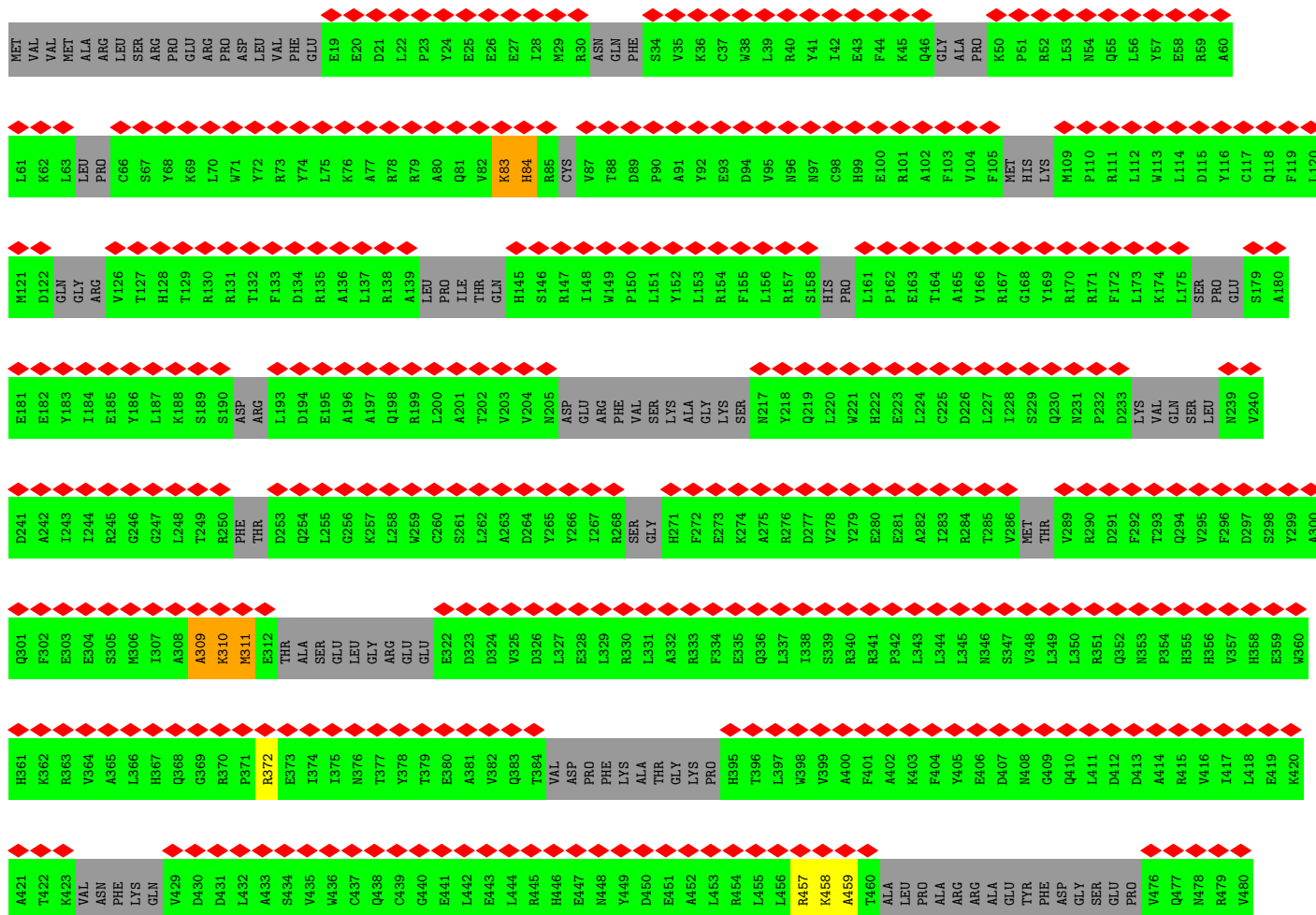


• Molecule 24: Crooked neck-like protein 1



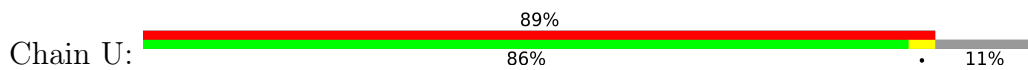


● Molecule 25: Pre-mRNA-splicing factor SYF1



Y481	K482	S483	L484	K485	V486	W487	S488	M489	L490	A491	D492	L493	E494	E495	S496	L497	GLY	THR	PHE	GLN	S502	T503	K504	A505	V506	Y507	D508	R509	I510	L511	D512	L513	R514	I515	A516	T517	P518	Q519	I520	V521	I522	N523	Y524	A525	M526	F527	L528	E529	E530	H531	K532	Y533	F534	E535	E536	S537	F538	K539	A540	
Y541	E542	R543	G544	I545	S546	L547	F548	K549	W550	V553	S554	D555	I556	W557	S558	T559	Y560	L561	T562	K563	F564	I565	A566	R567	Y568	G569	G570	R571	K572	L573	E574	R575	A576	R577	D578	L579	F580	E581	Q582	A583	L584	D585	G586	C587	F588	P589	K590	Y591	A592	K593	T594	L595	Y596	L597	L598	Y599	A600	Q601		
L602	E603	E604	E605	W606	G607	L608	A609	R610	H611	A612	M613	V614	A615	Y616	E617	R618	T620	R621	A622	V623	E624	P625	A626	Q627	Q628	Y629	D630	M631	F632	N633	I634	Y635	I636	K637	R638	A639	A640	E641	I642	Y643	G644	V645	F646	T647	T648	D709	R649	G650	I651	Y652	Q653	K654	A655	I656	E657	V658	L659	S660	D661	
E662	H663	A664	R665	E666	M667	C668	L669	R670	F671	A672	D673	M674	E675	C676	K677	L678	G679	E680	I681	D682	R683	A684	R685	A686	I687	Y688	Y689	F690	C691	S692	Q693	I694	C695	D696	P697	R698	T699	T700	G701	A702	F703	W704	Q705	L706	T707	K708	D709	F710	E711	V712	Q653	K654	A655	I656	E657	V658	L659	S660	K721	
E722	M723	L724	R725	I726	R727	R728	S729	V730	Q731	A732	T733	Y734	E735	T736	Q737	V738	N739	F740	M741	A742	S743	Q744	M745	L746	K747	V748	S749	GLY	SER	ALA	THR	GLY	THR	VAL	SER	ASP	GLU	ASP	GLU	LEU	ASP	ALA	PRO	GLY	ASP	ASN	ASP	VAL	VAL	LYS	LEU	GLU	GLN	ARG	ALA	GLU	GLN	LEU	ALA	ALA
GLU	ALA	GLU	PRO	ASP	GLN	PRO	LYS	VAL	ALA	SER	LYS	GLN	ILE	LEU	PHE	VAL	ARG	GLN	VAL	ARG	ASP	ALA	SER	ASP	ALA	ARG	ARG	PRO	GLU	GLU	ILE	GLN	LEU	GLY	THR	ASP	GLU	ASP	GLU	LEU	ASP	ALA	PRO	GLY	ASP	ASN	ASP	VAL	VAL	LYS	LEU	GLU	GLN	ARG	ALA	GLU	GLN	LEU	ALA	ALA
GLN	SER	VAL	PRO	ALA	ALA	VAL	PHE	GLY	SER	LEU	LEU	VAL	GLN	ASP	ASP	ARG	ARG	GLU	GLU	LEU	ALA	LEU	ALA	GLN	VAL	ASN	PRO	GLU	GLU	ILE	GLN	LEU	GLY	THR	ASP	GLU	ASP	GLU	LEU	ASP	ALA	PRO	GLY	ASP	ASN	ASP	VAL	VAL	LYS	LEU	GLU	GLN	ARG	ALA	GLU	GLN	LEU	ALA	ALA	

• Molecule 26: Intron-binding protein aquarius

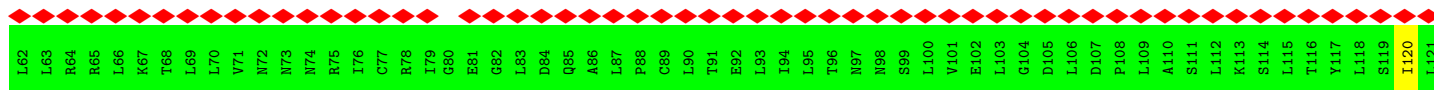


MET	ALA	ALA	PRO	ALA	GLN	PRO	LYS	VAL	ALA	PRO	THR	VAL	SER	GLN	ILE	M19	A20	E21	F22	V23	T24	Q25	A27	C28	K29	Y30	W31	A32	P33	H34	I35	K36	K37	K38	S39	P40	F41	D42	I43	K44	V45	I46	E47	D48	I49	Y50	E51	K52	E53	I54	V55	K56	S57	R58	F59	A60			
I61	R62	K63	I64	M65	L66	L67	E68	F69	S70	Q71	Y72	L73	E74	M75	Y76	L77	M78	N79	N80	Y81	S82	P83	E84	W85	S86	S87	K88	A89	Y90	L91	N92	S93	I94	C95	C96	M97	V98	N99	E100	K101	F102	R103	E104	M105	V106	P107	A108	W109	E110	I111	V112	K113	K114	K115	P116	D117	H118	F119	P120
F121	F122	F123	K124	I125	I126	L127	K128	A129	A130	L131	A132	E133	L134	D135	G136	E137	F138	S139	L140	H141	E142	Q143	T144	V145	L146	L147	L148	F149	L150	D151	H152	C153	M155	S156	L157	E158	V159	D160	L161	L162	R163	S164	Q165	V166	Q167	Q168	L169	I170	S171	L172	P173	M174	W175	M176	G177	L178	Q179	L180	
A181	R182	L183	E184	E185	L186	L187	K188	K189	T190	P191	K192	L193	R194	K195	F196	W197	N198	L199	A200	K201	M202	M203	D204	E205	K206	N207	D208	P209	E210	A211	R212	E213	A215	Y216	Q217	Q218	R219	R220	F221	L222	S223	Q224	L225	L226	Q227	K228	F229	D230	S231	V232	L233	K234	S235	V236	P237	L238	S239	E240	
F241	V242	T243	W244	D245	K246	V247	H248	Y249	G310	E251	R252	F253	L254	E255	L256	N257	D258	D259	L260	A320	E261	A262	L263	L264	F265	T266	R267	R268	W269	F270	M271	T272	L273	D275	D276	Q277	H278	L279	L280	F281	H282	C283	Y284	L285	S286	N287	L288	V289	R290	R291	E292	E293	D294	G295	H296	L297	F298	S299	Q300
L301	L302	D303	K304	L305	K306	F307	Y308	T309	G310	F311	E312	L313	N314	D315	Q316	T317	G318	N319	A320	L321	T322	E323	N324	K325	K326	T327	T328	L329	H330	Y331	D332	R333	L334	T335	S336	L337	Q338	R339	A340	F341	A342	C343	H344	F345	P346	E347	L348	Y349	D350	F351	A352	L353	S354	N355	V356	A357	E358	V359	D360

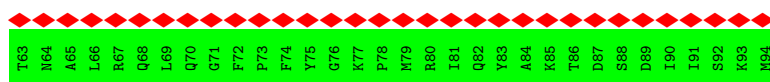
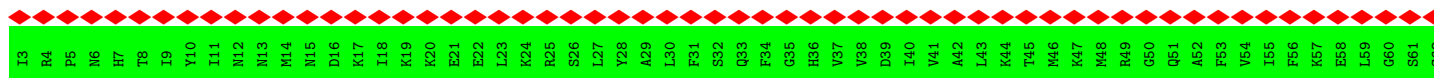
T361	R362	E363	S364	L365	V366	K367	F368	F369	G370	P371	L372	S373	S374	N375	T376	L377	H378	Q379	V380	A381	S382	Y383	L384	C385	L386	L387	P388	L389	L390	P391	K392	N393	N394	D395	T396	T397	F398	D399	K400	E401	F402	L403	L404	E405	L406	L407	V408	S409	R410	H411	E412	R413	R414	I415	S416	Q417	I418	Q419	Q420	
L421	N422	Q423	M424	P425	L426	Y427	P428	T429	E430	K431	I432	I433	W434	D435	E436	N437	I438	V439	P440	T441	E442	Y443	Y444	S445	S446	E447	G448	C449	L450	A451	L452	P453	E454	L455	N456	L457	L458	F459	L460	T461	L462	H463	D464	Y465	L466	L467	R468	N469	M470	N471	L472	F473	R474	L475	E476	S477	T478	Y479	E480	
I481	R482	Q483	D484	I485	E486	D487	S488	F489	S490	R491	M492	K493	P494	W495	Q496	GLU	TYR	G500	G501	V502	V503	F504	G505	G506	W507	A508	R509	M510	A511	Q512	P513	I514	V515	A516	F517	T518	V519	V520	E521	V522	A523	K524	P525	N526	I527	G528	E529	N530	M531	P532	T533	R534	V535	E536	A537	D538	V539	T540		
I541	N542	L543	N544	V545	R546	D547	H548	I549	K550	D551	E552	M553	E554	G555	L556	R557	K558	H559	D560	G561	V562	F563	L564	I565	T566	V567	R568	P569	T570	K571	P572	Y573	G574	T575	K576	F577	T578	R579	R580	R581	V582	P582	F583	I584	E585	Q586	V587	G588	L589	V590	Y591	V592	R593	G594	C595	E596	I597	Q598	V599	M600
L601	D602	D603	K604	G605	R606	V607	I608	GLU	ASP	GLY	PRO	GLU	P614	R615	P616	N617	L618	R619	G620	E621	S622	R623	T624	F625	R626	F627	F628	L629	D630	P631	M632	Q633	Y634	Q635	Q636	M637	T639	M640	T641	I642	Q643	M644	G645	A646	Q647	D648	V649	R650	E651	T652	F653	M654	I655	M656	M657	R658	R659	G599	M600	
P661	K662	E663	N664	N665	F666	K667	A668	V669	L670	E671	T672	I673	R674	N675	L676	M677	N678	T679	D680	C681	V682	F683	P684	D685	M686	L687	H688	D689	I690	I691	L692	G693	G694	G695	D696	P697	S698	S699	A700	H701	V702	S703	K704	M705	P706	N707	Q708	I709	A710	T711	L712	D713	F714	M715	D716	T717	F718	L719	S720	
I721	E722	H723	L724	K725	A726	S727	F728	P729	G730	H731	N732	K733	R734	N735	V736	T737	E738	D739	P740	A741	L742	Q743	I744	PRO	F746	F747	R748	I749	T750	F751	F752	V753	ARG	SER	GLY	LYS	LYS	ARG	LYS	ASP	ALA	ASP	VAL	GLU	ASP	GLU	THR	GLU	E773	A774	K775	T776	L777	I778	V779	E780				
P781	H782	V783	I784	P785	N786	R787	G788	P789	Y790	P791	Y792	N793	Q794	P795	K796	R797	N798	T799	I800	Q801	F802	T803	H804	T805	Q806	L807	E808	A809	I810	R811	A812	G813	M814	Q815	P816	G817	L818	T819	M820	V821	V822	G823	P824	P825	G826	T827	G828	K829	K830	D831	V832	A833	V834	Q835	I836	I837	S838	N839	I840	
H841	H842	N843	F844	P845	E846	R847	R848	T849	L850	I851	V852	T853	H854	S855	N856	Q857	A858	L859	N860	Q861	L862	F863	E864	K865	M866	M867	A868	L869	D870	I871	D872	E873	R874	H875	L876	L877	R878	L879	G880	H881	GLY	GLU	E884	E885	L886	E887	T888	E889	K890	D891	F892	S893	K894	Y895	G896	R897	V898	N899	Y900	
V901	L902	A903	R904	R905	I906	E907	L908	L909	E910	E911	V912	K913	R914	L915	Q916	K917	S918	L919	G920	V921	P922	G923	D924	A925	S926	Y927	T928	C929	E930	T931	A932	G933	V934	F935	F936	L937	Y938	Q939	V940	M941	S942	R943	W944	E945	E946	E947	Y947	I948	S949	K950	V951	K952	N953	LYS	GLY	SER	THR	LEU	P959	D960
V961	T962	E963	V964	S965	T966	F967	F968	P969	F970	H971	E972	Y973	F974	A975	N976	A977	PRO	GLN	PRO	I981	F982	K983	G984	R985	S986	Y987	E988	E989	D990	M991	E992	I993	A994	E995	G996	C997	F998	R999	H1000	I1001	K1002	K1003	I1004	F1005	T1006	Q1007	L1008	E1009	I1010	F1011	F1012	A1013	S1014	E1015	A1016	L1017	R1018	S1019	G1020	
L1021	D1022	R1023	S1024	K1025	Y1026	L1027	L1028	V1029	K1030	E1031	A1032	K1033	F1034	I1035	A1036	M1037	T1038	C1039	T1040	H1041	A1042	A1043	L1044	K1045	R1046	H1047	D1048	L1049	Y1050	K1051	L1052	G1053	F1054	K1055	Y1056	D1057	N1058	I1059	L1060	M1061	E1062	E1063	A1064	R1065	Q1066	Q1067	I1067	L1068	E1069	I1070	E1071	T1072	F1073	I1074	P1075	L1076	L1077	L1078	Q1079	N1080
P1081	Q1082	D1083	G1084	F1085	S1086	R1087	L1088	K1089	R1090	W1091	I1092	M1093	I1094	G1095	D1096	H1097	H1098	Q1099	L1100	P1101	P1102	V1103	I1104	K1105	N1106	M1107	A1108	F1109	Q1110	K1111	Y1112	S1113	N1114	M1115	E1116	Q1117	S1118	L1119	F1120	T1121	R1122	F1123	V1124	R1125	V1126	G1127	V1128	P1129	T1130	V1131	D1132	L1133	L1134	A1135	Q1136	L1137	R1138	A1139	R1140	







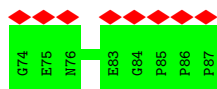
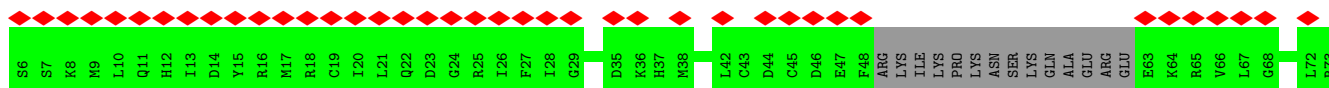
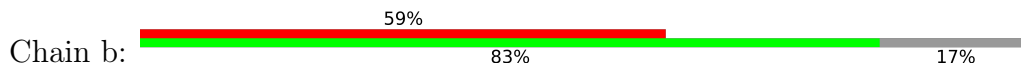
• Molecule 29: U2 small nuclear ribonucleoprotein B'



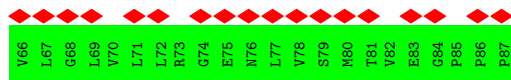
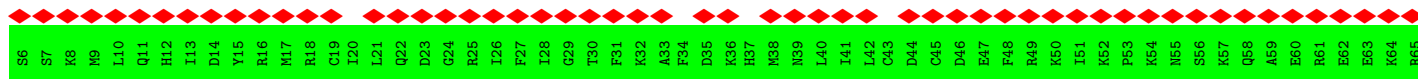
• Molecule 30: NF-kappa-B-activating protein



• Molecule 31: Small nuclear ribonucleoprotein-associated proteins B and B'

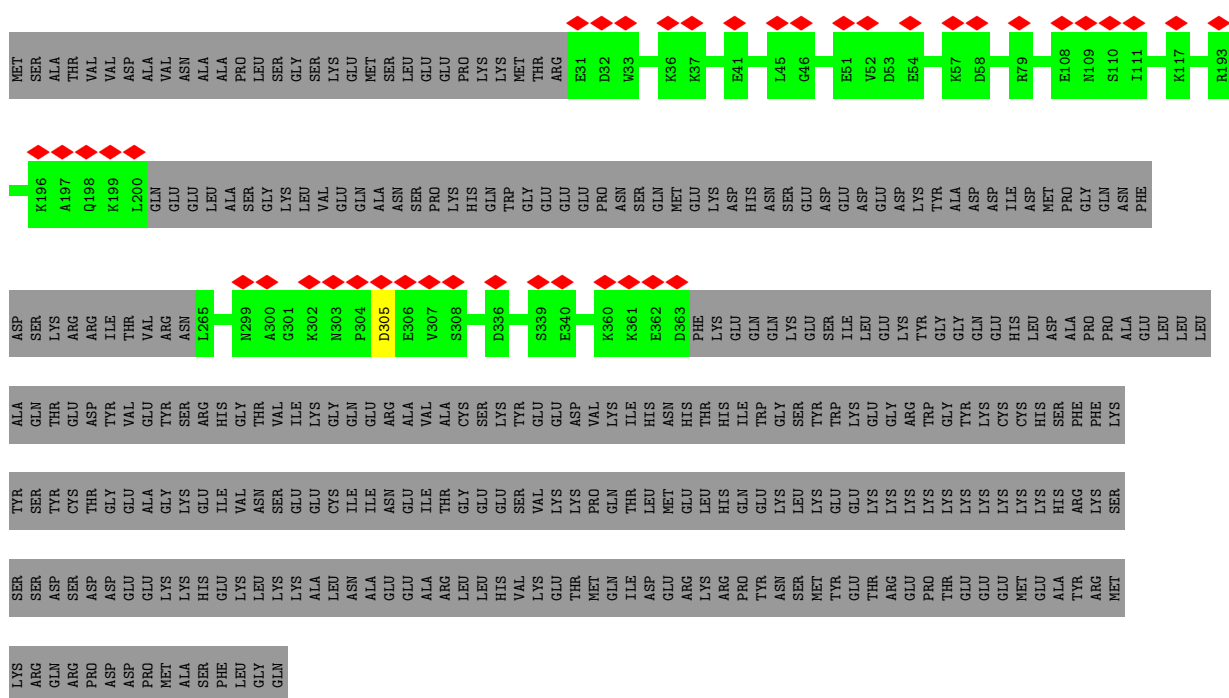


• Molecule 31: Small nuclear ribonucleoprotein-associated proteins B and B'

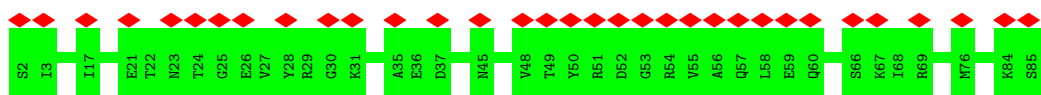
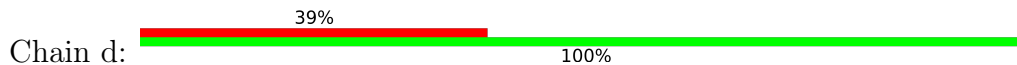


• Molecule 32: Pre-mRNA-splicing factor SLU7

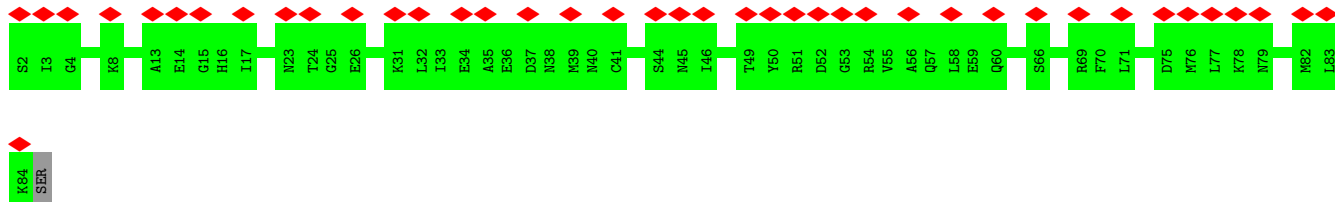




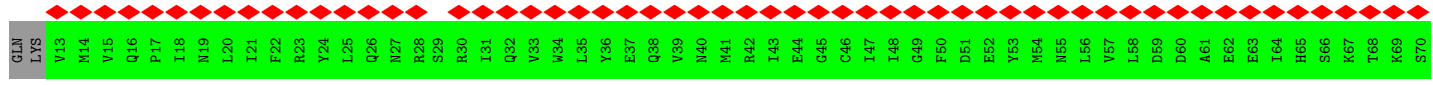
• Molecule 33: Small nuclear ribonucleoprotein Sm D3

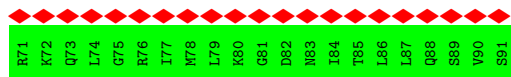


• Molecule 33: Small nuclear ribonucleoprotein Sm D3

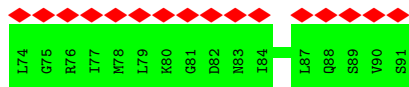
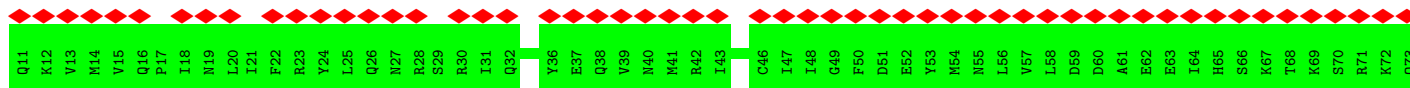
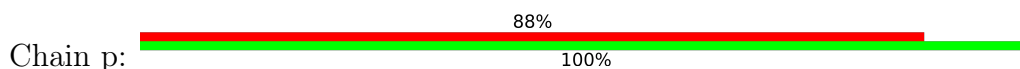


• Molecule 34: Small nuclear ribonucleoprotein E

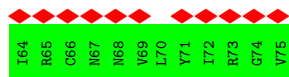
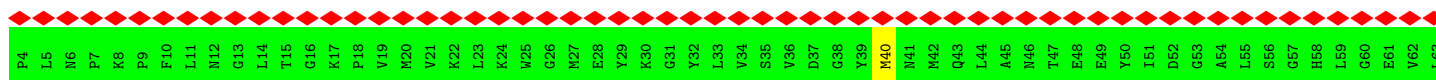




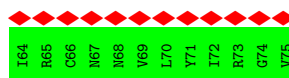
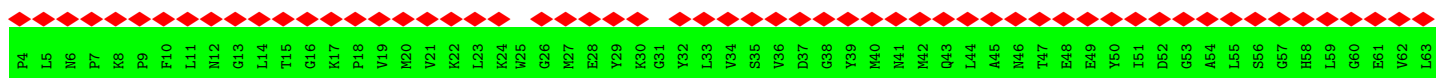
• Molecule 34: Small nuclear ribonucleoprotein E



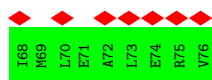
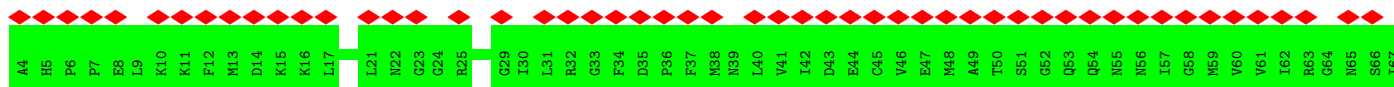
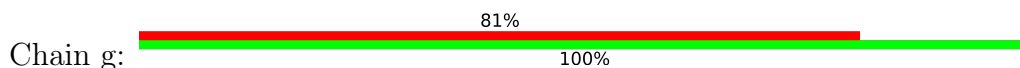
• Molecule 35: Small nuclear ribonucleoprotein F



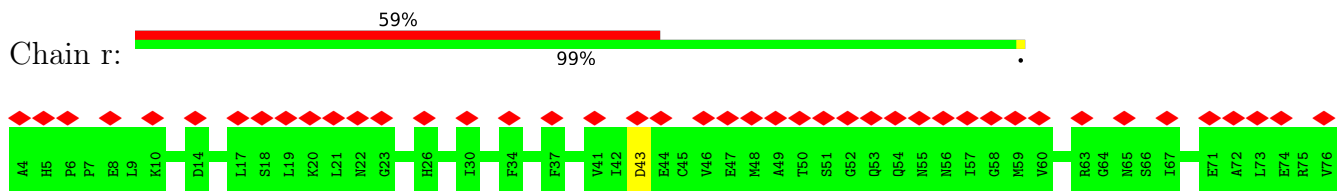
• Molecule 35: Small nuclear ribonucleoprotein F



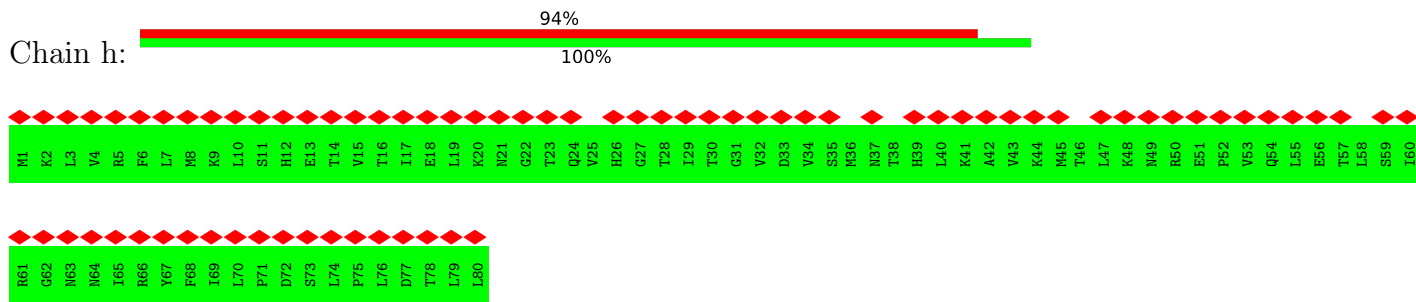
• Molecule 36: Small nuclear ribonucleoprotein G



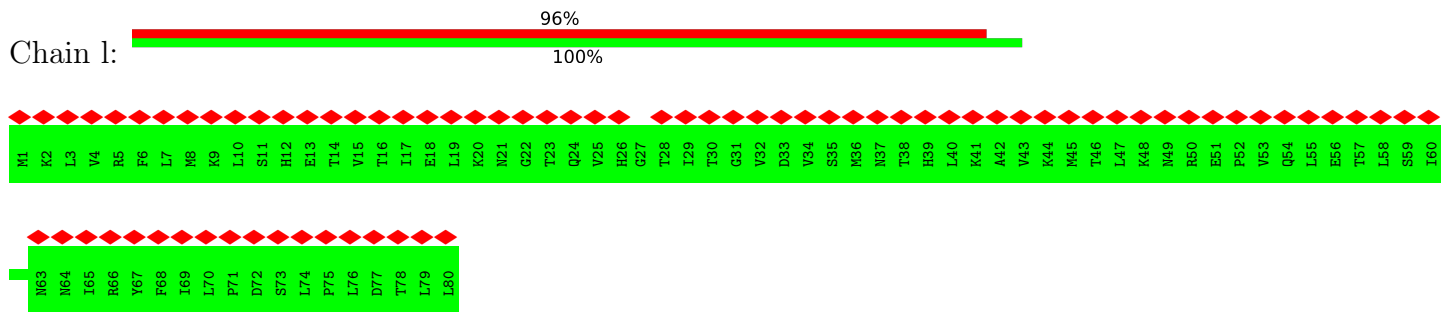
• Molecule 36: Small nuclear ribonucleoprotein G



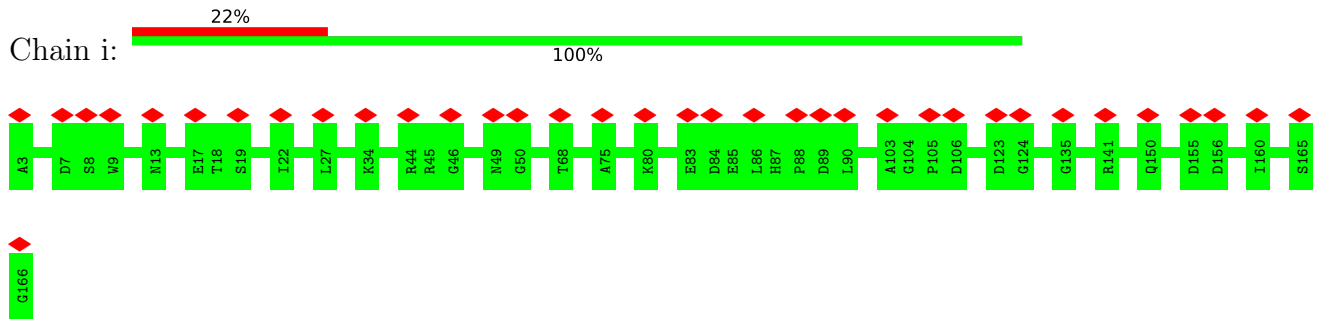
• Molecule 37: Small nuclear ribonucleoprotein Sm D1



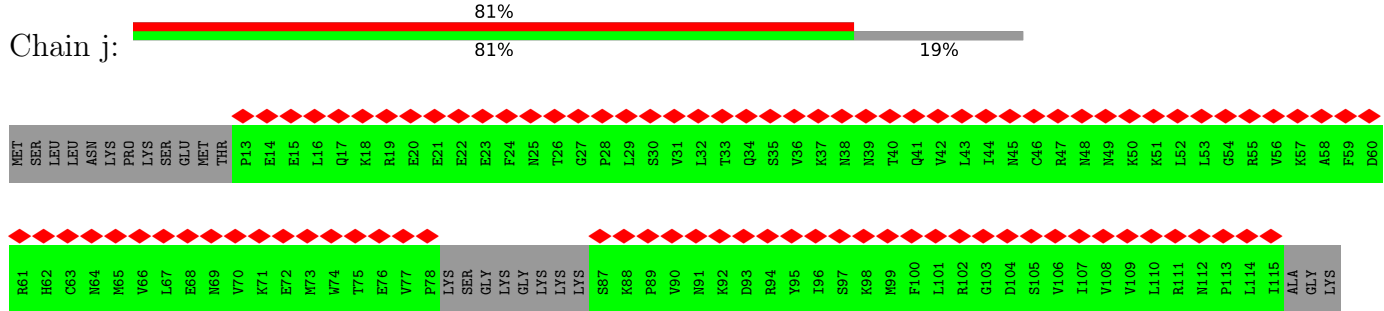
• Molecule 37: Small nuclear ribonucleoprotein Sm D1



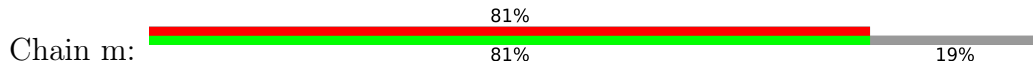
• Molecule 38: Peptidyl-prolyl cis-trans isomerase-like 1



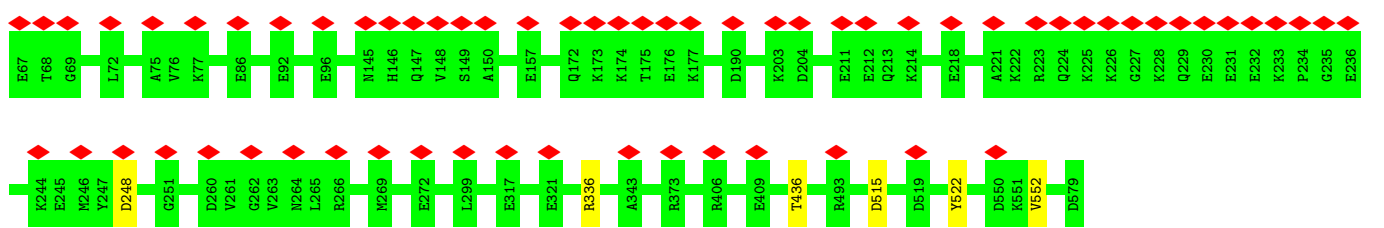
• Molecule 39: Small nuclear ribonucleoprotein Sm D2



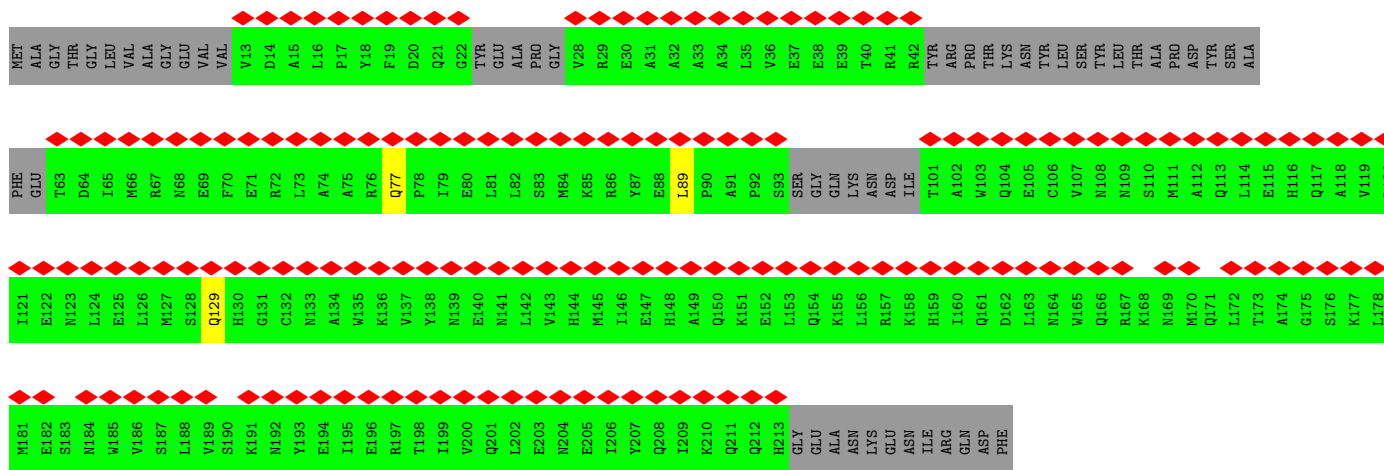
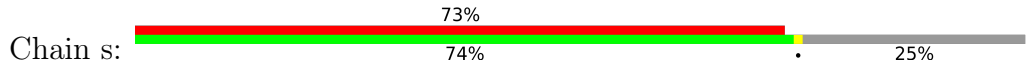
• Molecule 39: Small nuclear ribonucleoprotein Sm D2



• Molecule 40: Pre-mRNA-processing factor 17



• Molecule 41: Pre-mRNA-splicing factor SPF27



• Molecule 42: Pre-mRNA-processing factor 19





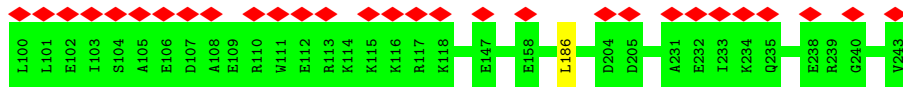


ASP GLY LEU LEU LEU ILE PHE HIS GLY THR GLY THR THR MET ASP ASP SER SER GLN ILE LYS LYS TRP ASP LEU LYS LYS LEU GLU ARG THR THR ASP ASN VAL VAL ALA ASN GLY PHE THR TYR PRO GLY HIS SER SER GLY PRO ILE THR ASP THR SER ILE ILE THR VAL SER ILE ALA PHE SER ILE SER CYS GLU LYS ASN GLY TYR TYR THR TYR THR LEU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

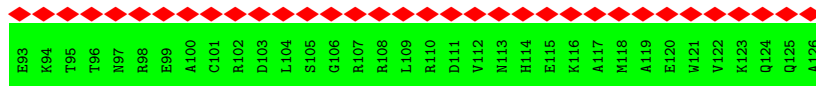
LEU ARG LYS LEU LYS PHE ASN PHE LYS THR ASP ASN ASN PHE ASP VAL SER LYS SER LEU ILE ARG THR THR ASP ASN VAL SER SER GLY PHE THR TYR LEU ALA LEU GLY PRO THR ASP VAL SER ILE ALA TYR PHE ILE SER CYS GLU LYS ASN GLN TRP THR TYR LEU ILE ALA THR THR HIS PHE THR ASP ASP HIS SER SER SER SER VAL LYS LEU THR TRP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

VAL ALA PHE GLY HIS HIS ALA LYS PHE ILE ALA SER THR ASP THR GLY MET ASP ARG SER LEU LYS PHE TYR SER LEU

● Molecule 43: Pre-mRNA-splicing factor SYF2



● Molecule 44: Replication stress response regulator SDE2



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	103860	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	53	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	135000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	0.177	Depositor
Minimum map value	-0.101	Depositor
Average map value	0.000	Depositor
Map value standard deviation	0.005	Depositor
Recommended contour level	0.024	Depositor
Map size (Å)	492.00003, 492.00003, 492.00003	wwPDB
Map dimensions	410, 410, 410	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.2, 1.2, 1.2	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: K, GTP, ATP, SEP, IHP, 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  >5	RMSZ	# Z  >5
1	2	0.56	0/2827	1.12	16/4393 (0.4%)
2	5	0.58	0/1760	1.27	25/2733 (0.9%)
3	6	0.53	0/2323	0.98	3/3619 (0.1%)
4	7	0.38	0/3179	0.65	0/4291
5	8	0.36	0/748	0.71	3/1012 (0.3%)
6	9	0.38	0/1225	0.60	0/1648
7	A	0.43	1/19172 (0.0%)	0.58	7/26014 (0.0%)
8	B	0.37	0/14140	0.58	2/19159 (0.0%)
9	C	0.39	0/7277	0.62	4/9887 (0.0%)
10	D	0.34	0/1030	0.59	0/1371
11	E	0.60	0/329	1.18	4/510 (0.8%)
12	F	0.32	0/1129	0.53	0/1525
13	G	0.27	0/513	0.55	1/683 (0.1%)
14	H	0.33	0/3779	0.50	0/5087
15	I	0.41	0/971	1.20	10/1504 (0.7%)
16	J	0.46	0/2592	0.63	0/3535
17	K	0.35	0/2387	0.57	0/3205
18	L	0.45	1/1214 (0.1%)	0.58	0/1627
19	M	0.36	0/2366	0.57	1/3193 (0.0%)
20	N	0.31	0/2448	0.58	0/3316
21	O	0.32	0/3457	0.50	0/4627
22	P	0.34	0/902	0.54	0/1201
23	R	0.58	1/196 (0.5%)	0.53	0/265
24	S	0.38	0/4013	0.55	8/5432 (0.1%)
25	T	0.41	0/4031	0.81	8/5500 (0.1%)
26	U	0.43	35/11155 (0.3%)	0.41	0/15095
27	V	0.47	0/3000	0.93	0/3777
28	W	0.31	0/1299	0.62	0/1761
29	Y	0.38	0/759	0.50	0/1016
30	Z	0.32	0/232	0.48	0/307
31	b	0.33	0/553	0.53	0/739
31	k	0.45	0/674	0.55	0/899

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
32	c	0.36	0/2268	0.53	0/3052
33	d	0.39	0/666	0.55	0/897
33	n	0.47	0/660	0.61	0/889
34	e	0.29	0/659	0.54	0/885
34	p	0.47	0/677	0.60	0/908
35	f	0.30	0/574	0.55	0/775
35	q	0.49	0/574	0.59	0/775
36	g	0.34	0/575	0.59	0/768
36	r	0.47	0/575	0.62	0/768
37	h	0.28	0/642	0.53	0/867
37	l	0.41	0/642	0.55	0/867
38	i	0.31	0/1304	0.57	0/1767
39	j	0.28	0/784	0.51	0/1053
39	m	0.42	0/784	0.56	0/1053
40	o	0.37	0/4265	0.60	0/5761
41	s	0.35	0/1423	0.51	0/1914
42	t	0.32	0/1004	0.50	0/1365
42	u	0.34	0/953	0.49	0/1295
42	v	0.34	0/1004	0.51	0/1365
42	w	0.31	0/953	0.49	0/1295
43	y	0.33	0/1241	0.56	1/1662 (0.1%)
44	z	0.44	0/282	0.50	0/375
All	All	0.40	38/124189 (0.0%)	0.64	93/169287 (0.1%)

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
7	A	0	4
8	B	0	1
9	C	0	2
17	K	0	1
18	L	0	2
25	T	0	5
27	V	0	15
35	f	0	1
41	s	0	3
All	All	0	34

All (38) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	2223	CYS	CB-SG	-6.69	1.70	1.82
18	L	117	CYS	CB-SG	-6.56	1.71	1.82
23	R	24	SER	CA-CB	-5.95	1.44	1.52
26	U	1107	MET	CG-SD	5.78	1.96	1.81
26	U	244	MET	CG-SD	5.75	1.96	1.81
26	U	79	MET	CG-SD	5.74	1.96	1.81
26	U	176	MET	CG-SD	5.74	1.96	1.81
26	U	1294	MET	CG-SD	5.73	1.96	1.81
26	U	991	MET	CG-SD	5.72	1.96	1.81
26	U	1358	MET	CG-SD	5.70	1.96	1.81
26	U	424	MET	CG-SD	5.67	1.96	1.81
26	U	705	MET	CG-SD	5.65	1.95	1.81
26	U	941	MET	CG-SD	5.60	1.95	1.81
26	U	257	MET	CG-SD	5.58	1.95	1.81
26	U	510	MET	CG-SD	5.58	1.95	1.81
26	U	1368	MET	CG-SD	5.55	1.95	1.81
26	U	207	MET	CG-SD	5.54	1.95	1.81
26	U	1212	MET	CG-SD	5.53	1.95	1.81
26	U	638	MET	CG-SD	5.52	1.95	1.81
26	U	1370	MET	CG-SD	5.52	1.95	1.81
26	U	1214	MET	CG-SD	5.52	1.95	1.81
26	U	600	MET	CG-SD	5.48	1.95	1.81
26	U	1361	MET	CG-SD	5.48	1.95	1.81
26	U	174	MET	CG-SD	5.47	1.95	1.81
26	U	677	MET	CG-SD	5.47	1.95	1.81
26	U	657	MET	CG-SD	5.45	1.95	1.81
26	U	92	MET	CG-SD	5.34	1.95	1.81
26	U	97	MET	CG-SD	5.31	1.95	1.81
26	U	1037	MET	CG-SD	5.30	1.95	1.81
26	U	492	MET	CG-SD	5.29	1.95	1.81
26	U	1093	MET	CG-SD	5.23	1.94	1.81
26	U	326	MET	CG-SD	5.23	1.94	1.81
26	U	1061	MET	CG-SD	5.22	1.94	1.81
26	U	65	MET	CG-SD	5.21	1.94	1.81
26	U	304	MET	CG-SD	5.20	1.94	1.81
26	U	1115	MET	CG-SD	5.14	1.94	1.81
26	U	814	MET	CG-SD	5.11	1.94	1.81
26	U	867	MET	CG-SD	5.07	1.94	1.81

All (93) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	5	86	C	N1-C2-O2	14.03	127.32	118.90

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	5	86	C	C2-N1-C1'	11.76	131.73	118.80
2	5	86	C	N3-C2-O2	-11.61	113.78	121.90
1	2	50	C	N1-C2-O2	10.30	125.08	118.90
1	2	50	C	C2-N1-C1'	9.37	129.10	118.80
2	5	36	C	N1-C2-O2	8.89	124.24	118.90
1	2	50	C	N3-C2-O2	-8.47	115.97	121.90
2	5	86	C	C6-N1-C2	-8.14	117.04	120.30
25	T	84	HIS	O-C-N	-8.14	109.67	122.70
25	T	83	LYS	O-C-N	-8.11	109.73	122.70
25	T	309	ALA	O-C-N	-8.09	109.76	122.70
25	T	311	MET	O-C-N	-8.09	109.76	122.70
25	T	310	LYS	O-C-N	-8.07	109.78	122.70
2	5	86	C	C6-N1-C1'	-8.00	111.20	120.80
43	y	186	LEU	CB-CG-CD1	-7.85	97.66	111.00
2	5	36	C	C2-N1-C1'	7.69	127.26	118.80
7	A	598	LEU	CB-CG-CD1	-7.19	98.78	111.00
7	A	1405	LEU	CA-CB-CG	7.16	131.76	115.30
2	5	36	C	C6-N1-C2	-7.13	117.45	120.30
11	E	-6	C	N1-C2-O2	7.12	123.17	118.90
2	5	36	C	N3-C2-O2	-7.02	116.99	121.90
9	C	486	ASP	CB-CG-OD1	6.99	124.59	118.30
1	2	102	U	OP2-P-O3'	6.93	120.44	105.20
2	5	23	C	C2-N1-C1'	6.80	126.28	118.80
15	I	2	U	N1-C2-O2	6.71	127.50	122.80
11	E	-6	C	C2-N1-C1'	6.68	126.15	118.80
2	5	95	G	P-O3'-C3'	6.65	127.68	119.70
15	I	2	U	N3-C2-O2	-6.64	117.56	122.20
1	2	102	U	P-O3'-C3'	6.62	127.65	119.70
2	5	23	C	N1-C2-O2	6.55	122.83	118.90
1	2	105	G	P-O3'-C3'	6.52	127.53	119.70
1	2	50	C	C6-N1-C1'	-6.51	112.98	120.80
5	8	114	LYS	N-CA-C	-6.42	93.67	111.00
5	8	118	LEU	CA-CB-CG	6.26	129.70	115.30
1	2	54	U	N3-C2-O2	-6.21	117.85	122.20
15	I	142	U	OP1-P-O3'	6.20	118.84	105.20
8	B	1276	LEU	CA-CB-CG	6.15	129.44	115.30
1	2	54	U	N1-C2-O2	6.09	127.07	122.80
1	2	50	C	C6-N1-C2	-6.00	117.90	120.30
24	S	637	PRO	N-CA-CB	5.94	110.43	103.30
24	S	523	PRO	N-CA-CB	5.92	110.40	103.30
9	C	723	ASP	CB-CG-OD1	5.91	123.62	118.30
24	S	675	PRO	N-CA-CB	5.90	110.39	103.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
24	S	604	PRO	N-CA-CB	5.89	110.37	103.30
9	C	93	ILE	CG1-CB-CG2	-5.88	98.46	111.40
24	S	670	PRO	N-CA-CB	5.87	110.35	103.30
25	T	458	LYS	O-C-N	5.86	132.07	122.70
25	T	459	ALA	O-C-N	5.83	132.03	122.70
2	5	86	C	C5-C6-N1	5.83	123.91	121.00
25	T	457	ARG	O-C-N	5.82	132.02	122.70
9	C	474	LEU	CA-CB-CG	5.80	128.65	115.30
2	5	92	U	P-O3'-C3'	5.77	126.63	119.70
7	A	598	LEU	CA-CB-CG	5.77	128.57	115.30
24	S	551	PRO	N-CA-CB	5.70	110.14	103.30
2	5	36	C	C5-C6-N1	5.69	123.84	121.00
11	E	-6	C	N3-C2-O2	-5.68	117.92	121.90
24	S	566	PRO	N-CA-CB	5.66	110.09	103.30
2	5	23	C	N3-C2-O2	-5.65	117.94	121.90
24	S	522	PRO	N-CA-CB	5.61	110.03	103.30
15	I	145	U	O4'-C1'-N1	5.59	112.67	108.20
1	2	156	U	N1-C2-O2	5.59	126.71	122.80
2	5	95	G	N3-C4-C5	-5.52	125.84	128.60
2	5	11	U	C2-N1-C1'	5.45	124.24	117.70
7	A	2310	ARG	CG-CD-NE	5.40	123.14	111.80
15	I	17	U	C2-N1-C1'	5.39	124.17	117.70
15	I	142	U	P-O3'-C3'	5.38	126.16	119.70
2	5	11	U	C5-C6-N1	5.34	125.37	122.70
2	5	55	C	C6-N1-C2	-5.34	118.16	120.30
15	I	131	U	N1-C2-O2	5.34	126.54	122.80
2	5	12	U	N3-C2-O2	-5.32	118.48	122.20
1	2	13	C	N1-C2-O2	5.31	122.09	118.90
2	5	43	U	N1-C2-O2	5.31	126.52	122.80
11	E	1	C	C6-N1-C2	-5.28	118.19	120.30
8	B	501	LEU	CA-CB-CG	5.27	127.41	115.30
7	A	638	LEU	CB-CG-CD1	-5.25	102.07	111.00
1	2	156	U	N3-C2-O2	-5.23	118.54	122.20
1	2	154	C	N3-C2-O2	-5.22	118.25	121.90
5	8	115	GLY	N-CA-C	5.22	126.16	113.10
13	G	56	ASP	CB-CG-OD2	5.19	122.97	118.30
19	M	131	THR	C-N-CA	5.18	134.66	121.70
15	I	17	U	N1-C2-O2	5.17	126.42	122.80
2	5	40	U	N1-C2-O2	5.16	126.41	122.80
2	5	11	U	N1-C2-O2	5.15	126.40	122.80
3	6	50	A	P-O3'-C3'	5.14	125.86	119.70
3	6	37	C	C6-N1-C2	-5.13	118.25	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	2	156	U	C2-N1-C1'	5.12	123.85	117.70
7	A	784	LEU	CA-CB-CG	-5.07	103.65	115.30
15	I	137	C	P-O3'-C3'	5.07	125.78	119.70
3	6	61	C	C6-N1-C2	-5.05	118.28	120.30
1	2	49	U	N3-C2-O2	-5.04	118.67	122.20
15	I	131	U	N3-C2-O2	-5.04	118.67	122.20
2	5	22	U	C2-N1-C1'	5.03	123.74	117.70
7	A	2319	LEU	CB-CG-CD2	5.03	119.55	111.00

There are no chirality outliers.

All (34) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
7	A	1210	LYS	Peptide
7	A	1416	ILE	Peptide
7	A	1635	TYR	Peptide
7	A	940	ILE	Peptide
8	B	430	LEU	Peptide
9	C	799	GLU	Peptide
9	C	93	ILE	Peptide
17	K	123	GLU	Peptide
18	L	127	GLU	Peptide
18	L	128	VAL	Peptide
25	T	309	ALA	Mainchain
25	T	310	LYS	Mainchain
25	T	311	MET	Mainchain
25	T	83	LYS	Mainchain
25	T	84	HIS	Mainchain
27	V	1098	LYS	Peptide
27	V	1127	LEU	Peptide
27	V	1163	ARG	Peptide
27	V	1187	THR	Peptide
27	V	515	ASP	Peptide
27	V	550	TYR	Peptide
27	V	563	GLU	Peptide
27	V	636	CYS	Peptide
27	V	662	THR	Peptide
27	V	811	PRO	Peptide
27	V	855	VAL	Peptide
27	V	868	LYS	Peptide
27	V	938	ASN	Peptide
27	V	963	THR	Peptide

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Group
27	V	993	LEU	Peptide
35	f	40	MET	Peptide
41	s	129	GLN	Peptide
41	s	77	GLN	Peptide
41	s	89	LEU	Peptide

## 5.2 Too-close contacts [i](#)

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

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	7	388/390 (100%)	376 (97%)	9 (2%)	3 (1%)	19	51
5	8	89/91 (98%)	87 (98%)	1 (1%)	1 (1%)	14	45
6	9	142/144 (99%)	138 (97%)	4 (3%)	0	100	100
7	A	2244/2335 (96%)	2105 (94%)	139 (6%)	0	100	100
8	B	1720/1722 (100%)	1633 (95%)	84 (5%)	3 (0%)	47	77
9	C	897/899 (100%)	828 (92%)	67 (8%)	2 (0%)	47	77
10	D	121/123 (98%)	116 (96%)	5 (4%)	0	100	100
12	F	120/122 (98%)	107 (89%)	13 (11%)	0	100	100
13	G	58/60 (97%)	56 (97%)	2 (3%)	0	100	100
14	H	455/908 (50%)	440 (97%)	15 (3%)	0	100	100
16	J	318/320 (99%)	303 (95%)	15 (5%)	0	100	100
17	K	291/295 (99%)	271 (93%)	20 (7%)	0	100	100
18	L	142/144 (99%)	134 (94%)	8 (6%)	0	100	100
19	M	287/289 (99%)	271 (94%)	16 (6%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
20	N	304/306 (99%)	283 (93%)	19 (6%)	2 (1%)	22	54
21	O	429/802 (54%)	416 (97%)	12 (3%)	1 (0%)	47	77
22	P	100/229 (44%)	97 (97%)	3 (3%)	0	100	100
23	R	24/26 (92%)	17 (71%)	7 (29%)	0	100	100
24	S	531/848 (63%)	500 (94%)	28 (5%)	3 (1%)	25	57
25	T	597/855 (70%)	584 (98%)	12 (2%)	1 (0%)	47	77
26	U	1308/1485 (88%)	1283 (98%)	25 (2%)	0	100	100
27	V	709/1220 (58%)	617 (87%)	60 (8%)	32 (4%)	2	15
28	W	160/162 (99%)	147 (92%)	13 (8%)	0	100	100
29	Y	90/92 (98%)	88 (98%)	2 (2%)	0	100	100
30	Z	28/30 (93%)	27 (96%)	1 (4%)	0	100	100
31	b	64/82 (78%)	62 (97%)	2 (3%)	0	100	100
31	k	80/82 (98%)	74 (92%)	6 (8%)	0	100	100
32	c	265/586 (45%)	248 (94%)	17 (6%)	0	100	100
33	d	82/84 (98%)	77 (94%)	5 (6%)	0	100	100
33	n	81/84 (96%)	76 (94%)	5 (6%)	0	100	100
34	e	77/81 (95%)	75 (97%)	2 (3%)	0	100	100
34	p	79/81 (98%)	77 (98%)	2 (2%)	0	100	100
35	f	70/72 (97%)	69 (99%)	1 (1%)	0	100	100
35	q	70/72 (97%)	69 (99%)	1 (1%)	0	100	100
36	g	71/73 (97%)	68 (96%)	3 (4%)	0	100	100
36	r	71/73 (97%)	69 (97%)	2 (3%)	0	100	100
37	h	78/80 (98%)	76 (97%)	2 (3%)	0	100	100
37	l	78/80 (98%)	75 (96%)	3 (4%)	0	100	100
38	i	162/164 (99%)	149 (92%)	13 (8%)	0	100	100
39	j	91/118 (77%)	87 (96%)	4 (4%)	0	100	100
39	m	91/118 (77%)	86 (94%)	5 (6%)	0	100	100
40	o	511/513 (100%)	465 (91%)	45 (9%)	1 (0%)	47	77
41	s	161/225 (72%)	148 (92%)	13 (8%)	0	100	100
42	t	121/504 (24%)	118 (98%)	3 (2%)	0	100	100
42	u	114/504 (23%)	113 (99%)	1 (1%)	0	100	100

Continued on next page...



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
42	v	121/504 (24%)	119 (98%)	2 (2%)	0	100	100
42	w	114/504 (23%)	113 (99%)	1 (1%)	0	100	100
43	y	142/144 (99%)	138 (97%)	4 (3%)	0	100	100
44	z	32/34 (94%)	30 (94%)	2 (6%)	0	100	100
All	All	14378/18759 (77%)	13605 (95%)	724 (5%)	49 (0%)	44	71

All (49) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	7	383	ASN
8	B	957	VAL
8	B	1584	ILE
20	N	59	ILE
27	V	531	MET
27	V	610	ARG
27	V	855	VAL
27	V	944	ASP
27	V	995	MET
27	V	1090	LYS
27	V	1098	LYS
27	V	1099	SER
27	V	1129	ASP
27	V	1145	GLN
27	V	1164	GLU
4	7	340	GLY
4	7	385	ASP
5	8	115	GLY
9	C	94	ILE
24	S	709	VAL
27	V	517	GLU
27	V	845	SER
27	V	870	GLY
27	V	1182	LYS
27	V	524	ASN
27	V	550	TYR
27	V	806	GLU
27	V	994	ILE
27	V	1056	LYS
20	N	58	PRO
24	S	604	PRO
27	V	590	THR

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type
27	V	663	ASP
27	V	722	ALA
27	V	835	VAL
27	V	1055	ASN
40	o	248	ASP
9	C	680	ASN
24	S	341	PRO
27	V	811	PRO
27	V	1157	THR
25	T	372	ARG
27	V	523	ALA
27	V	856	ASP
21	O	215	PRO
27	V	738	PRO
27	V	521	ILE
27	V	1137	PRO
8	B	585	ILE

### 5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	7	345/345 (100%)	340 (99%)	5 (1%)	67	82
5	8	76/76 (100%)	75 (99%)	1 (1%)	69	82
6	9	132/132 (100%)	131 (99%)	1 (1%)	81	89
7	A	2033/2108 (96%)	1996 (98%)	37 (2%)	59	78
8	B	1541/1541 (100%)	1363 (88%)	178 (12%)	5	22
9	C	799/799 (100%)	797 (100%)	2 (0%)	92	96
10	D	106/106 (100%)	105 (99%)	1 (1%)	78	87
12	F	110/110 (100%)	110 (100%)	0	100	100
13	G	54/55 (98%)	53 (98%)	1 (2%)	57	77
14	H	410/838 (49%)	403 (98%)	7 (2%)	60	78
16	J	276/276 (100%)	275 (100%)	1 (0%)	91	95

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
17	K	246/247 (100%)	246 (100%)	0	100	100
18	L	130/130 (100%)	127 (98%)	3 (2%)	50	73
19	M	254/254 (100%)	253 (100%)	1 (0%)	91	95
20	N	263/263 (100%)	263 (100%)	0	100	100
21	O	322/709 (45%)	321 (100%)	1 (0%)	92	96
22	P	94/203 (46%)	94 (100%)	0	100	100
23	R	21/21 (100%)	20 (95%)	1 (5%)	25	56
24	S	275/751 (37%)	273 (99%)	2 (1%)	84	90
25	T	213/749 (28%)	210 (99%)	3 (1%)	67	82
26	U	1202/1336 (90%)	1193 (99%)	9 (1%)	84	90
27	V	31/1085 (3%)	31 (100%)	0	100	100
28	W	139/147 (95%)	138 (99%)	1 (1%)	84	90
29	Y	81/82 (99%)	81 (100%)	0	100	100
30	Z	25/25 (100%)	25 (100%)	0	100	100
31	b	62/75 (83%)	62 (100%)	0	100	100
31	k	75/75 (100%)	75 (100%)	0	100	100
32	c	236/520 (45%)	235 (100%)	1 (0%)	91	95
33	d	74/74 (100%)	74 (100%)	0	100	100
33	n	73/74 (99%)	73 (100%)	0	100	100
34	e	74/76 (97%)	74 (100%)	0	100	100
34	p	76/76 (100%)	76 (100%)	0	100	100
35	f	61/61 (100%)	61 (100%)	0	100	100
35	q	61/61 (100%)	61 (100%)	0	100	100
36	g	63/63 (100%)	63 (100%)	0	100	100
36	r	63/63 (100%)	62 (98%)	1 (2%)	62	79
37	h	75/75 (100%)	75 (100%)	0	100	100
37	l	75/75 (100%)	75 (100%)	0	100	100
38	i	133/133 (100%)	133 (100%)	0	100	100
39	j	91/110 (83%)	91 (100%)	0	100	100
39	m	91/110 (83%)	91 (100%)	0	100	100
40	o	451/451 (100%)	446 (99%)	5 (1%)	73	85

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
41	s	152/196 (78%)	152 (100%)	0	100	100
42	t	111/435 (26%)	111 (100%)	0	100	100
42	u	106/435 (24%)	106 (100%)	0	100	100
42	v	111/435 (26%)	111 (100%)	0	100	100
42	w	106/435 (24%)	105 (99%)	1 (1%)	78	87
43	y	129/129 (100%)	129 (100%)	0	100	100
44	z	29/29 (100%)	29 (100%)	0	100	100
All	All	11756/16654 (71%)	11493 (98%)	263 (2%)	54	74

All (263) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	7	37	THR
4	7	301	ASN
4	7	316	ARG
4	7	337	TRP
4	7	360	LEU
5	8	118	LEU
6	9	119	GLU
7	A	258	PHE
7	A	1359	HIS
7	A	1370	ARG
7	A	1868	MET
7	A	2067	PHE
7	A	2073	TRP
7	A	2074	ARG
7	A	2078	ILE
7	A	2085	LEU
7	A	2087	THR
7	A	2090	ILE
7	A	2103	THR
7	A	2108	LYS
7	A	2117	ILE
7	A	2143	ARG
7	A	2156	THR
7	A	2157	VAL
7	A	2159	LEU
7	A	2171	GLU
7	A	2193	VAL
7	A	2194	THR

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
7	A	2219	THR
7	A	2223	CYS
7	A	2233	SER
7	A	2239	ARG
7	A	2242	THR
7	A	2254	SER
7	A	2259	VAL
7	A	2261	MET
7	A	2273	VAL
7	A	2284	MET
7	A	2293	LYS
7	A	2298	LEU
7	A	2310	ARG
7	A	2312	SER
7	A	2319	LEU
7	A	2329	ASP
8	B	406	ARG
8	B	409	LEU
8	B	410	ASP
8	B	414	LEU
8	B	420	SER
8	B	436	ARG
8	B	446	HIS
8	B	447	VAL
8	B	451	LYS
8	B	467	LEU
8	B	475	PHE
8	B	488	LEU
8	B	495	THR
8	B	500	LEU
8	B	501	LEU
8	B	505	THR
8	B	533	VAL
8	B	535	ASP
8	B	547	LEU
8	B	550	GLU
8	B	558	ARG
8	B	566	VAL
8	B	572	ASP
8	B	576	CYS
8	B	578	GLU
8	B	584	GLN

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
8	B	591	GLU
8	B	595	ILE
8	B	602	GLU
8	B	610	ARG
8	B	614	LEU
8	B	623	ASP
8	B	637	ARG
8	B	643	GLN
8	B	673	LEU
8	B	677	ASP
8	B	690	VAL
8	B	693	THR
8	B	712	ILE
8	B	728	ARG
8	B	743	LEU
8	B	759	THR
8	B	763	ARG
8	B	773	GLU
8	B	775	LYS
8	B	782	PHE
8	B	786	HIS
8	B	807	GLN
8	B	810	VAL
8	B	820	ASN
8	B	837	GLU
8	B	849	ILE
8	B	850	LEU
8	B	855	ARG
8	B	868	ILE
8	B	869	LEU
8	B	877	GLN
8	B	885	GLN
8	B	887	LEU
8	B	894	VAL
8	B	897	LEU
8	B	900	MET
8	B	901	LEU
8	B	910	VAL
8	B	920	LEU
8	B	934	THR
8	B	941	ASP
8	B	942	ASP

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
8	B	952	ARG
8	B	957	VAL
8	B	972	TYR
8	B	975	LYS
8	B	992	TYR
8	B	1016	ARG
8	B	1020	LEU
8	B	1028	THR
8	B	1030	ARG
8	B	1062	LEU
8	B	1063	LEU
8	B	1087	SER
8	B	1100	LEU
8	B	1101	ASN
8	B	1102	ARG
8	B	1125	SER
8	B	1135	LEU
8	B	1143	ILE
8	B	1165	ILE
8	B	1166	ARG
8	B	1186	LEU
8	B	1187	SER
8	B	1224	LEU
8	B	1225	VAL
8	B	1234	LEU
8	B	1240	LEU
8	B	1241	LEU
8	B	1244	LYS
8	B	1248	ASP
8	B	1250	HIS
8	B	1262	LEU
8	B	1278	CYS
8	B	1287	ARG
8	B	1301	LEU
8	B	1312	LEU
8	B	1320	LEU
8	B	1337	ASN
8	B	1368	LEU
8	B	1375	ARG
8	B	1399	ASP
8	B	1406	VAL
8	B	1408	LEU

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
8	B	1413	SER
8	B	1419	LEU
8	B	1421	LYS
8	B	1425	ILE
8	B	1430	GLU
8	B	1436	SER
8	B	1441	GLN
8	B	1442	ARG
8	B	1443	LYS
8	B	1455	GLU
8	B	1456	VAL
8	B	1474	MET
8	B	1477	ILE
8	B	1480	GLN
8	B	1481	ILE
8	B	1482	GLU
8	B	1492	SER
8	B	1567	LYS
8	B	1580	CYS
8	B	1629	ARG
8	B	1655	ASN
8	B	1682	TYR
8	B	1683	ASP
8	B	1707	GLN
8	B	1713	PHE
8	B	1728	LEU
8	B	1734	ASP
8	B	1742	THR
8	B	1747	ASN
8	B	1756	THR
8	B	1762	ARG
8	B	1779	ARG
8	B	1781	LEU
8	B	1788	LEU
8	B	1817	MET
8	B	1823	TYR
8	B	1826	TYR
8	B	1829	ILE
8	B	1834	MET
8	B	1840	THR
8	B	1842	VAL
8	B	1863	HIS

*Continued on next page...*



*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
8	B	1865	ASP
8	B	1936	LEU
8	B	1956	LYS
8	B	1957	ASP
8	B	1969	GLU
8	B	1970	HIS
8	B	1988	MET
8	B	1996	LEU
8	B	1999	LEU
8	B	2000	THR
8	B	2017	ILE
8	B	2027	ASP
8	B	2029	ILE
8	B	2031	SER
8	B	2047	VAL
8	B	2055	LEU
8	B	2070	ASP
8	B	2082	LEU
8	B	2084	LEU
8	B	2092	LEU
8	B	2095	VAL
8	B	2102	HIS
8	B	2105	THR
8	B	2109	MET
8	B	2121	LYS
8	B	2125	ASP
9	C	168	THR
9	C	824	THR
10	D	102	LYS
13	G	79	LYS
14	H	156	ARG
14	H	259	PHE
14	H	333	GLN
14	H	344	LYS
14	H	387	MET
14	H	481	PHE
14	H	494	LEU
16	J	302	VAL
18	L	80	LYS
18	L	130	ARG
18	L	131	ILE
19	M	222	ARG

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
21	O	261	LYS
23	R	1	MET
24	S	366	TYR
24	S	436	TYR
25	T	548	PHE
25	T	606	TRP
25	T	724	LEU
26	U	90	TYR
26	U	544	ASN
26	U	881	HIS
26	U	960	ASP
26	U	986	SER
26	U	1021	LEU
26	U	1206	TYR
26	U	1283	HIS
26	U	1343	LYS
28	W	120	ILE
32	c	305	ASP
40	o	336	ARG
40	o	436	THR
40	o	515	ASP
40	o	522	TYR
40	o	552	VAL
36	r	43	ASP
42	w	119	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (117) such sidechains are listed below:

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
4	7	301	ASN
4	7	394	GLN
7	A	160	HIS
7	A	210	HIS
7	A	270	ASN
7	A	434	HIS
7	A	505	ASN
7	A	664	HIS
7	A	675	GLN
7	A	755	HIS
7	A	775	ASN
7	A	792	HIS
7	A	834	HIS

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
7	A	994	ASN
7	A	1014	ASN
7	A	1024	HIS
7	A	1096	HIS
7	A	1129	ASN
7	A	1159	ASN
7	A	1169	GLN
7	A	1296	GLN
7	A	1337	GLN
7	A	1394	GLN
7	A	1468	ASN
7	A	1476	GLN
7	A	1586	HIS
7	A	1599	GLN
7	A	1658	GLN
7	A	1717	ASN
7	A	1946	ASN
7	A	1947	ASN
7	A	2123	GLN
7	A	2300	ASN
7	A	2306	HIS
8	B	425	ASN
9	C	82	GLN
9	C	137	HIS
9	C	154	HIS
9	C	175	GLN
9	C	477	HIS
9	C	557	GLN
9	C	680	ASN
9	C	924	GLN
14	H	171	ASN
14	H	474	HIS
14	H	499	GLN
14	H	553	HIS
14	H	600	ASN
16	J	191	HIS
16	J	216	ASN
16	J	269	GLN
16	J	417	ASN
16	J	446	ASN
16	J	451	HIS
16	J	500	HIS

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
17	K	84	ASN
19	M	82	GLN
20	N	101	ASN
20	N	165	GLN
20	N	225	ASN
21	O	245	GLN
22	P	56	ASN
23	R	20	GLN
24	S	221	ASN
24	S	250	GLN
24	S	311	GLN
25	T	601	GLN
25	T	739	ASN
25	T	744	GLN
26	U	393	ASN
28	W	43	GLN
28	W	72	ASN
29	Y	15	ASN
31	b	76	ASN
32	c	84	HIS
32	c	85	GLN
32	c	154	HIS
32	c	156	GLN
32	c	171	ASN
32	c	344	GLN
35	f	41	ASN
37	h	64	ASN
38	i	31	HIS
38	i	148	ASN
39	j	39	ASN
39	j	91	ASN
31	k	76	ASN
37	l	64	ASN
39	m	49	ASN
33	n	16	HIS
33	n	40	ASN
33	n	45	ASN
40	o	145	ASN
40	o	147	GLN
40	o	250	GLN
40	o	402	GLN
40	o	462	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
40	o	479	GLN
34	p	19	ASN
34	p	88	GLN
35	q	68	ASN
36	r	26	HIS
41	s	192	ASN
42	t	9	ASN
42	t	94	GLN
42	t	95	GLN
42	u	45	GLN
42	u	79	GLN
42	v	14	HIS
42	v	43	ASN
42	v	79	GLN
42	v	101	GLN
42	w	22	ASN
42	w	37	ASN
43	y	131	GLN
43	y	172	HIS
43	y	215	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	2	117/189 (61%)	30 (25%)	5 (4%)
11	E	14/14 (100%)	5 (35%)	1 (7%)
15	I	39/113 (34%)	19 (48%)	2 (5%)
2	5	73/116 (62%)	24 (32%)	2 (2%)
3	6	96/106 (90%)	37 (38%)	5 (5%)
All	All	339/538 (63%)	115 (33%)	15 (4%)

All (115) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	2	16	U
1	2	17	U
1	2	19	G
1	2	20	G
1	2	24	A
1	2	25	G
1	2	29	A

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	2	30	A
1	2	32	U
1	2	40	C
1	2	41	U
1	2	46	U
1	2	47	U
1	2	51	A
1	2	53	U
1	2	54	U
1	2	96	A
1	2	97	G
1	2	99	U
1	2	100	U
1	2	101	U
1	2	102	U
1	2	103	U
1	2	104	G
1	2	105	G
1	2	106	A
1	2	109	U
1	2	157	G
1	2	171	U
1	2	178	A
2	5	10	U
2	5	11	U
2	5	20	G
2	5	21	A
2	5	34	U
2	5	35	U
2	5	36	C
2	5	38	C
2	5	39	C
2	5	42	U
2	5	45	C
2	5	47	A
2	5	52	U
2	5	53	U
2	5	57	G
2	5	69	A
2	5	86	C
2	5	89	U
2	5	91	U

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
2	5	92	U
2	5	93	U
2	5	94	U
2	5	95	G
2	5	96	A
3	6	6	C
3	6	7	G
3	6	8	C
3	6	9	U
3	6	11	C
3	6	12	G
3	6	26	U
3	6	28	A
3	6	29	A
3	6	33	G
3	6	34	G
3	6	35	A
3	6	37	C
3	6	38	G
3	6	41	A
3	6	44	G
3	6	45	A
3	6	46	G
3	6	48	A
3	6	49	G
3	6	51	U
3	6	54	G
3	6	56	A
3	6	58	G
3	6	59	G
3	6	60	C
3	6	61	C
3	6	62	C
3	6	67	G
3	6	68	C
3	6	74	U
3	6	79	C
3	6	81	C
3	6	84	A
3	6	85	U
3	6	87	C
3	6	88	G

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
11	E	-11	G
11	E	-7	C
11	E	-6	C
11	E	1	C
11	E	2	U
15	I	5	G
15	I	14	A
15	I	16	G
15	I	17	U
15	I	21	A
15	I	134	U
15	I	135	G
15	I	136	U
15	I	137	C
15	I	138	A
15	I	139	U
15	I	140	A
15	I	142	U
15	I	143	U
15	I	144	A
15	I	145	U
15	I	146	C
15	I	166	A
15	I	167	G

All (15) RNA pucker outliers are listed below:

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	2	39	U
1	2	46	U
1	2	95	A
1	2	102	U
1	2	105	G
2	5	92	U
2	5	95	G
3	6	5	U
3	6	33	G
3	6	37	C
3	6	50	A
3	6	58	G
11	E	-12	G
15	I	137	C

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type
15	I	142	U

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

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

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
17	SEP	K	232	-	8,9,10	1.57	1 (12%)	8,12,14	1.47	2 (25%)
17	SEP	K	224	-	8,9,10	1.54	1 (12%)	8,12,14	1.58	1 (12%)

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
17	SEP	K	232	-	-	3/5/8/10	-
17	SEP	K	224	-	-	0/5/8/10	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
17	K	232	SEP	P-O1P	3.39	1.61	1.50
17	K	224	SEP	P-O1P	3.34	1.61	1.50

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
17	K	224	SEP	P-OG-CB	-3.83	107.75	118.30
17	K	232	SEP	P-OG-CB	-2.85	110.45	118.30
17	K	232	SEP	OG-CB-CA	2.46	110.54	108.14

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
17	K	232	SEP	CB-OG-P-O2P
17	K	232	SEP	CB-OG-P-O3P
17	K	232	SEP	CB-OG-P-O1P

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 18 ligands modelled in this entry, 15 are monoatomic - leaving 3 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
47	ATP	7	702	45	26,33,33	0.92	1 (3%)	31,52,52	1.55	5 (16%)
48	GTP	C	1500	45	26,34,34	1.24	1 (3%)	32,54,54	1.70	7 (21%)
50	IHP	c	601	-	36,36,36	0.74	0	54,60,60	0.91	3 (5%)

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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
47	ATP	7	702	45	-	0/18/38/38	0/3/3/3
48	GTP	C	1500	45	-	5/18/38/38	0/3/3/3
50	IHP	c	601	-	-	6/30/54/54	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
48	C	1500	GTP	C5-C6	-4.41	1.38	1.47
47	7	702	ATP	O4'-C1'	2.08	1.44	1.41

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
48	C	1500	GTP	PA-O3A-PB	-4.43	117.62	132.83
47	7	702	ATP	PB-O3B-PG	-4.34	117.95	132.83
47	7	702	ATP	N3-C2-N1	-4.09	122.29	128.68
48	C	1500	GTP	PB-O3B-PG	-3.51	120.78	132.83
48	C	1500	GTP	C5-C6-N1	3.47	120.07	113.95
47	7	702	ATP	C4-C5-N7	-3.29	105.97	109.40
50	c	601	IHP	C5-C4-C3	3.07	117.13	110.41
48	C	1500	GTP	C8-N7-C5	2.99	108.69	102.99
48	C	1500	GTP	C2-N1-C6	-2.98	119.62	125.10
50	c	601	IHP	C6-C1-C2	-2.95	103.96	110.41
48	C	1500	GTP	C3'-C2'-C1'	2.85	105.26	100.98
47	7	702	ATP	PA-O3A-PB	-2.73	123.47	132.83
48	C	1500	GTP	O6-C6-C5	-2.38	119.72	124.37
47	7	702	ATP	O4'-C1'-C2'	-2.30	103.57	106.93
50	c	601	IHP	C4-C3-C2	2.06	114.92	110.41

There are no chirality outliers.

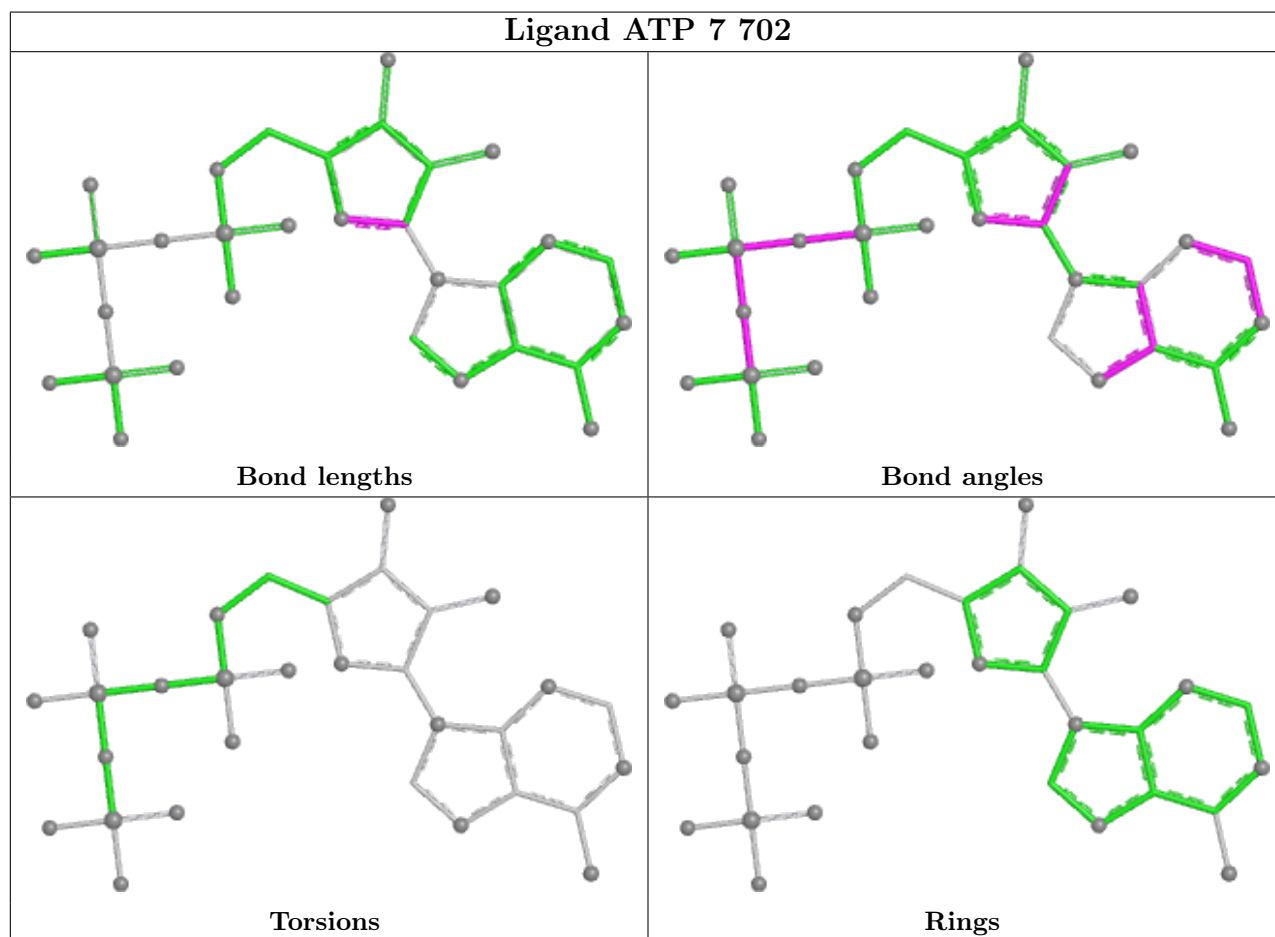
All (11) torsion outliers are listed below:

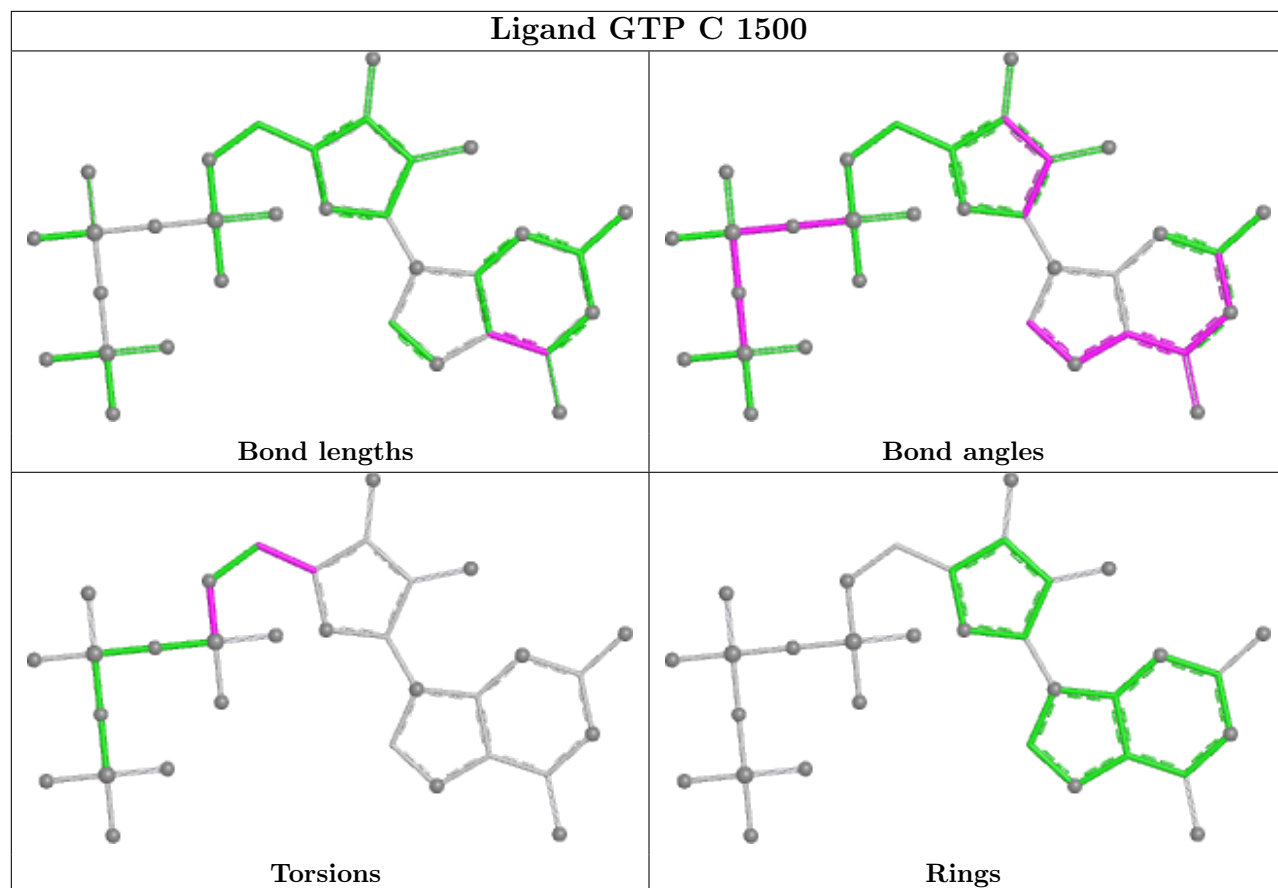
Mol	Chain	Res	Type	Atoms
48	C	1500	GTP	C5'-O5'-PA-O3A
48	C	1500	GTP	C5'-O5'-PA-O1A
48	C	1500	GTP	C5'-O5'-PA-O2A
50	c	601	IHP	C2-O12-P2-O42
50	c	601	IHP	C5-O15-P5-O25
50	c	601	IHP	C6-O16-P6-O26
48	C	1500	GTP	C3'-C4'-C5'-O5'
48	C	1500	GTP	O4'-C4'-C5'-O5'
50	c	601	IHP	C1-O11-P1-O31
50	c	601	IHP	C1-O11-P1-O41
50	c	601	IHP	C6-O16-P6-O46

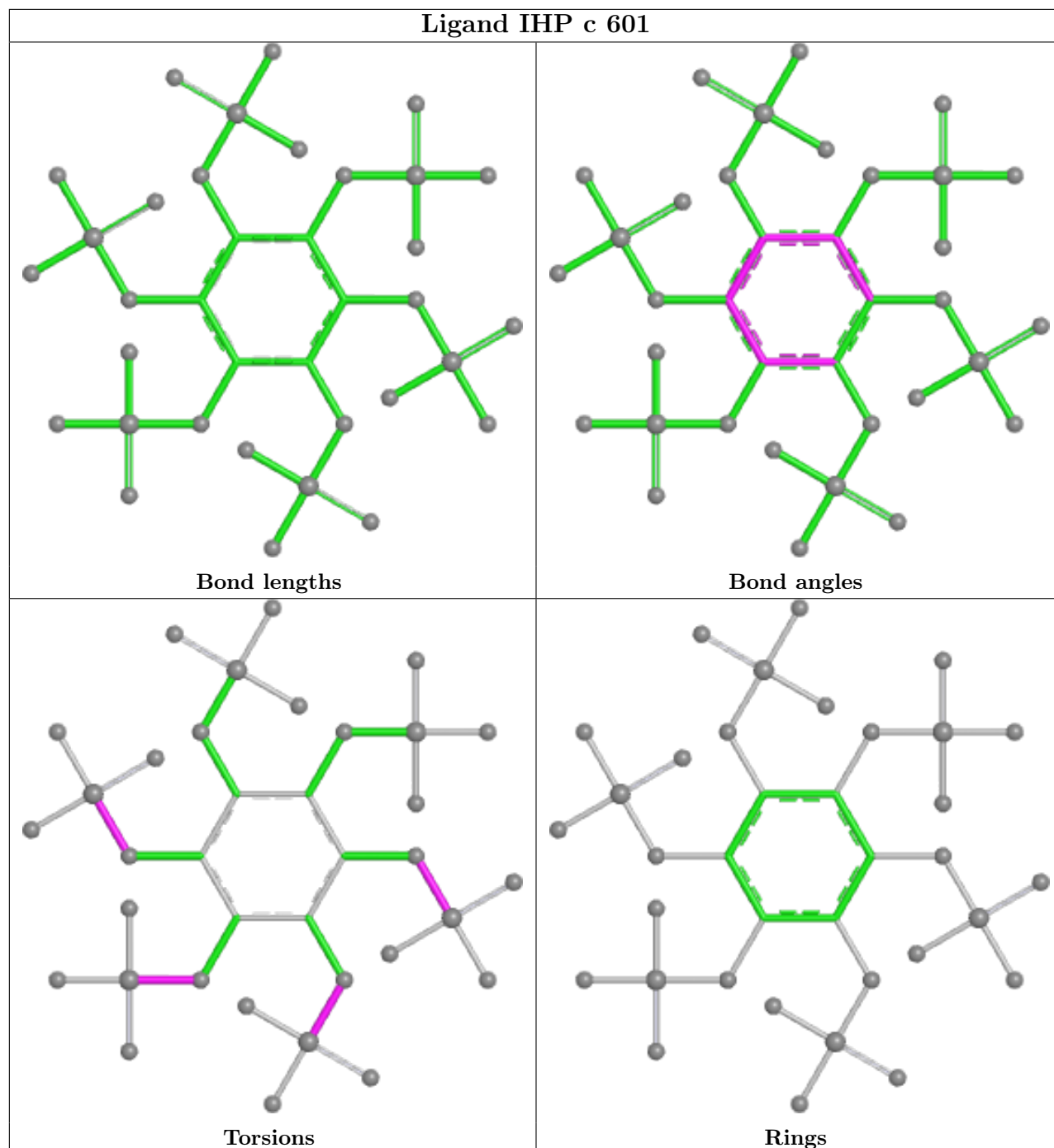
There are no ring outliers.

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 than 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.







## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

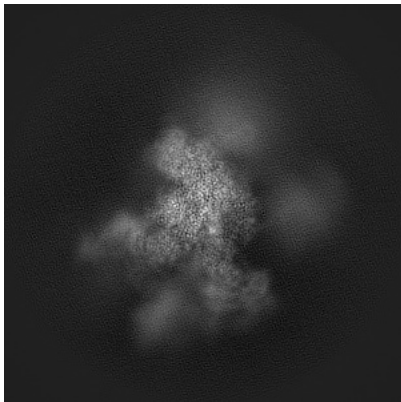
## 6 Map visualisation [i](#)

This section contains visualisations of the EMDB entry EMD-4525. These allow visual inspection of the internal detail of the map and identification of artifacts.

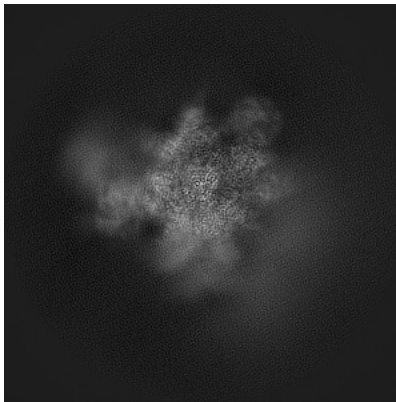
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

### 6.1 Orthogonal projections [i](#)

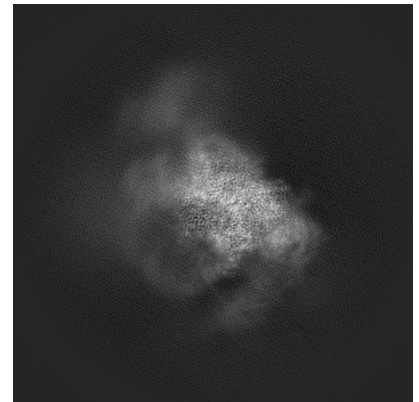
#### 6.1.1 Primary map



X



Y

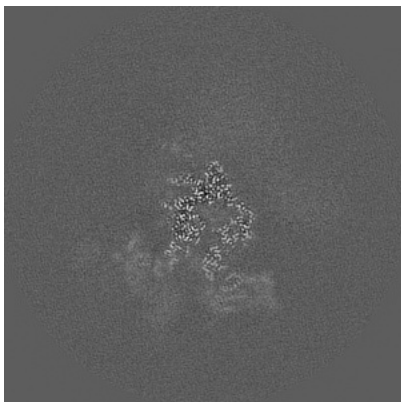


Z

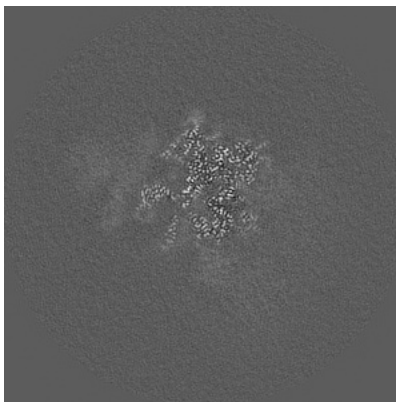
The images above show the map projected in three orthogonal directions.

### 6.2 Central slices [i](#)

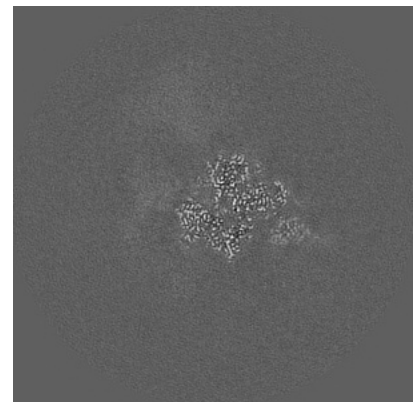
#### 6.2.1 Primary map



X Index: 205



Y Index: 205

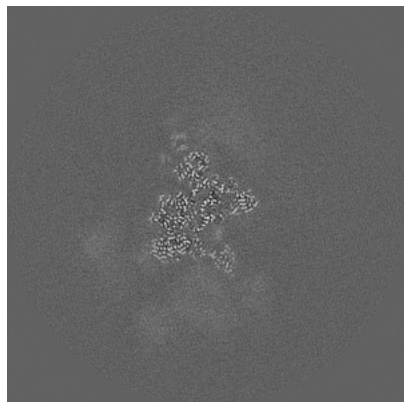


Z Index: 205

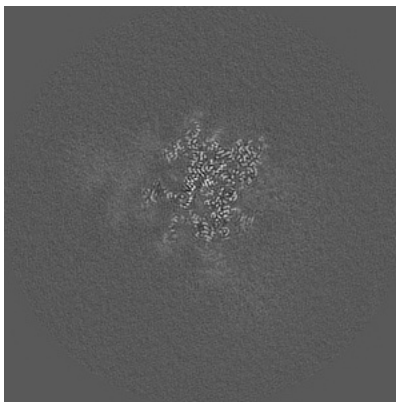
The images above show central slices of the map in three orthogonal directions.

## 6.3 Largest variance slices [i](#)

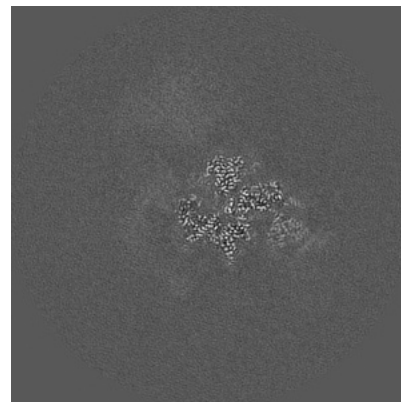
### 6.3.1 Primary map



X Index: 231



Y Index: 203



Z Index: 202

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

## 6.4 Orthogonal surface views [i](#)

### 6.4.1 Primary map



X



Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.024. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.



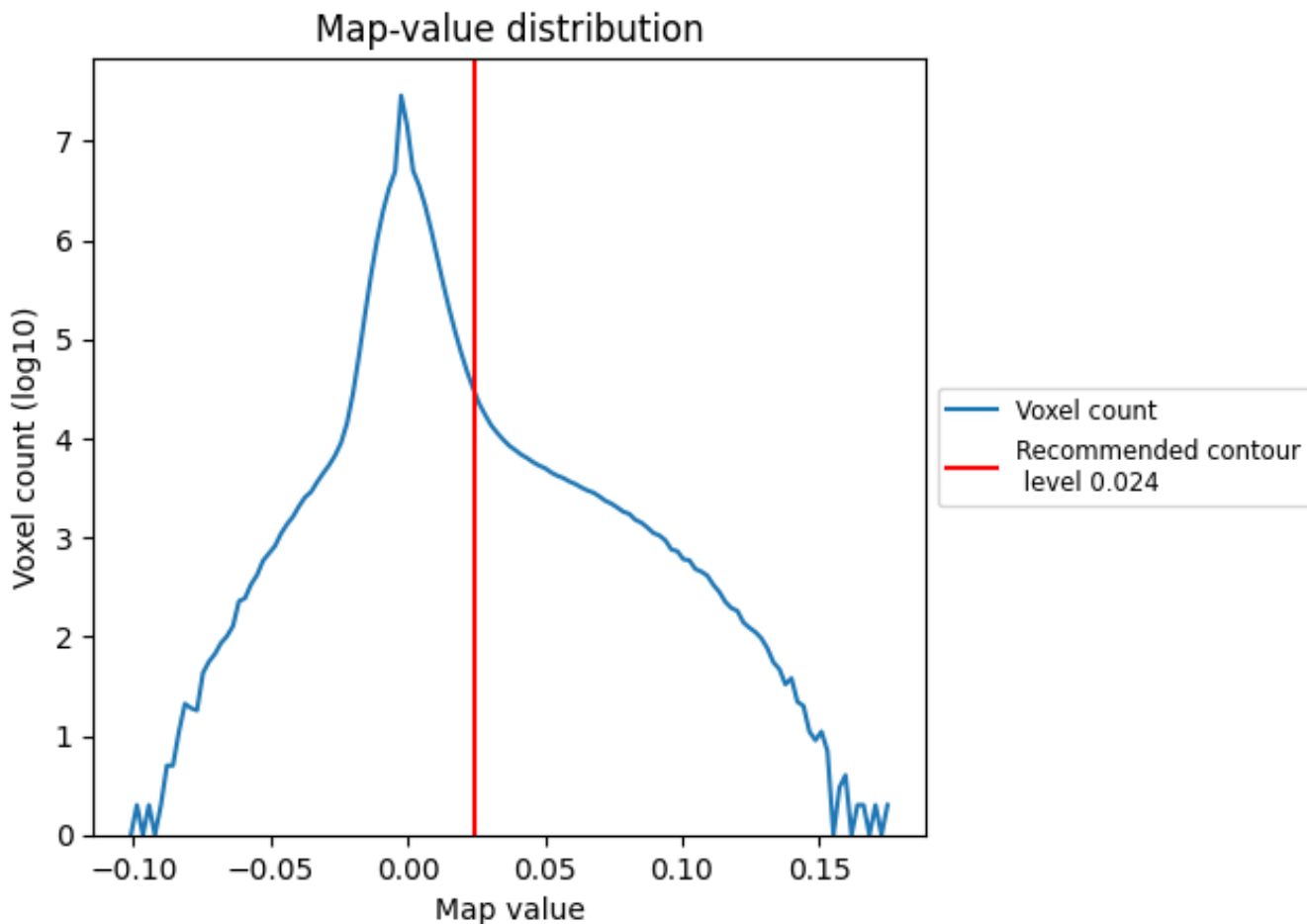
## 6.5 Mask visualisation

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

## 7 Map analysis [i](#)

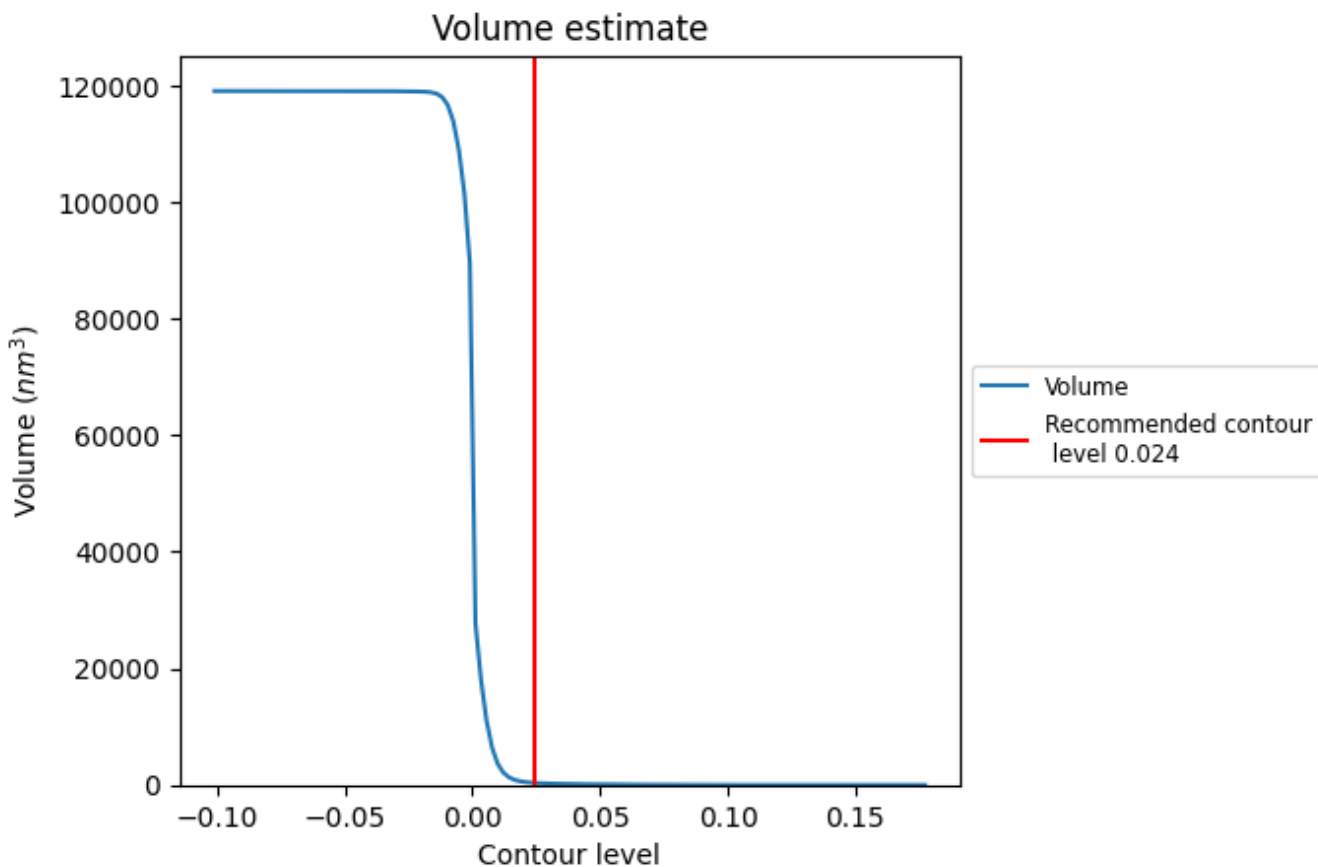
This section contains the results of statistical analysis of the map.

### 7.1 Map-value distribution [i](#)



The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

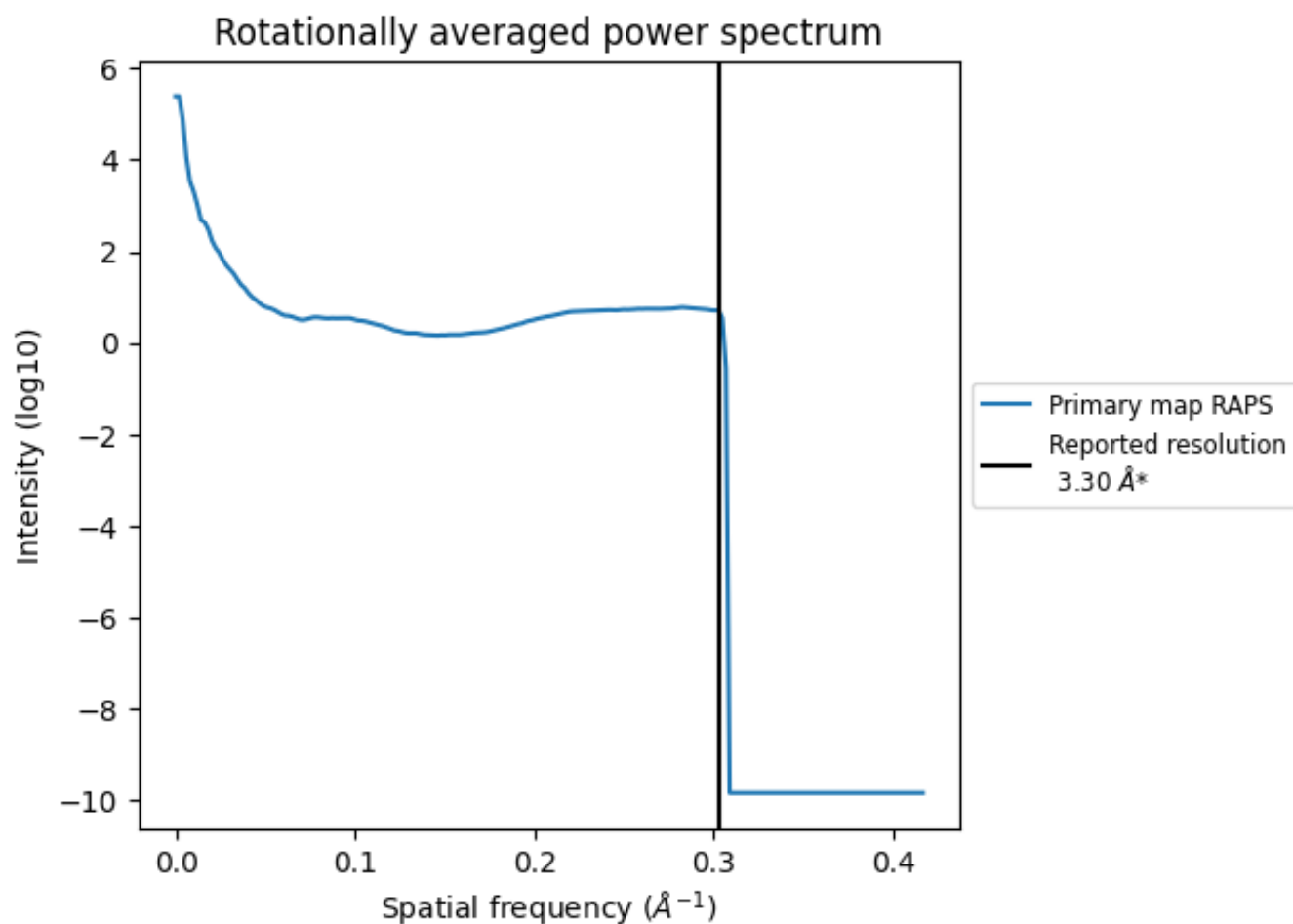
## 7.2 Volume estimate [i](#)



The volume at the recommended contour level is 356 nm<sup>3</sup>; this corresponds to an approximate mass of 321 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.

### 7.3 Rotationally averaged power spectrum [\(i\)](#)

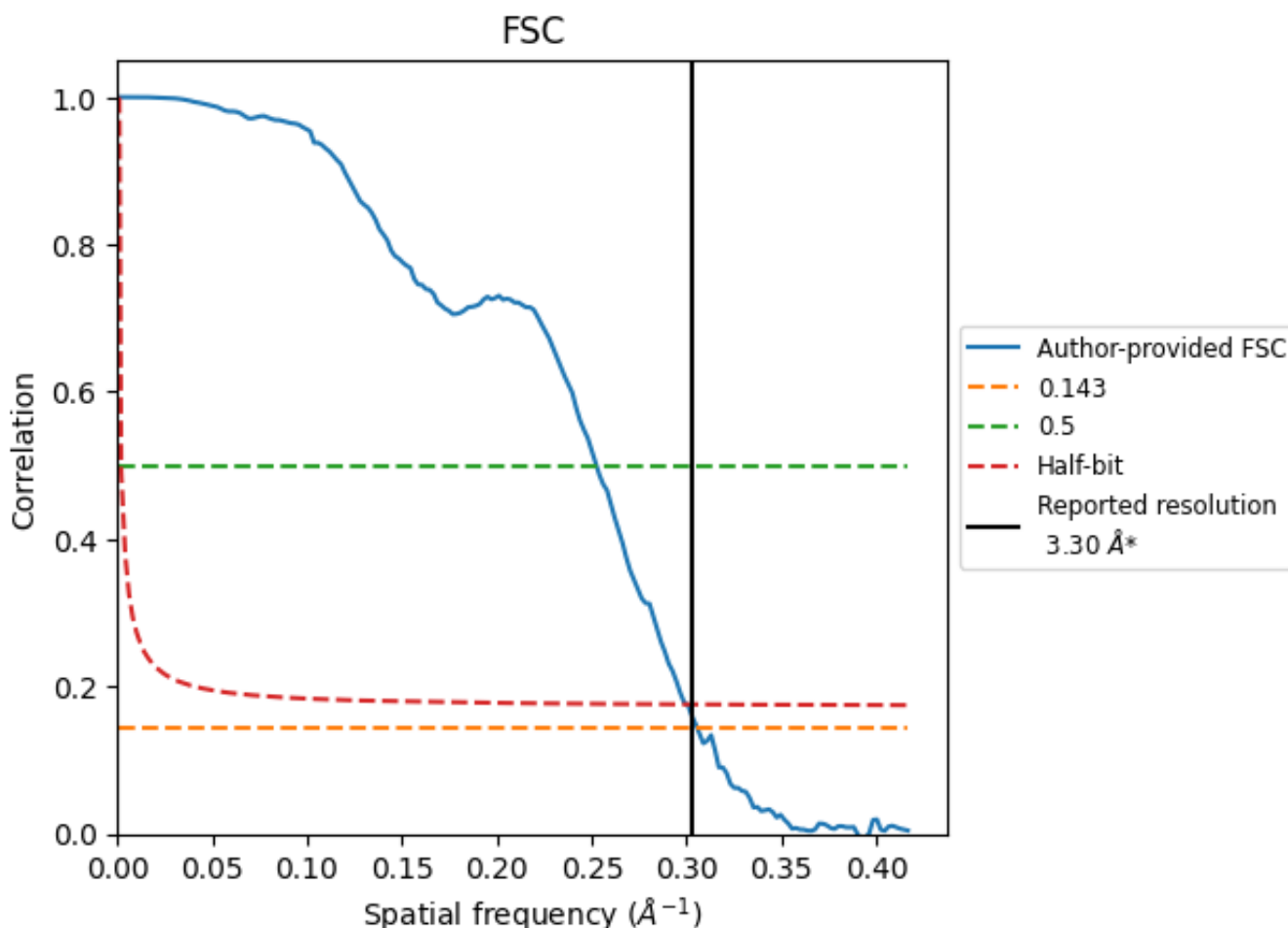


\*Reported resolution corresponds to spatial frequency of  $0.303 \text{\AA}^{-1}$

## 8 Fourier-Shell correlation [\(i\)](#)

Fourier-Shell Correlation (FSC) is the most commonly used method to estimate the resolution of single-particle and subtomogram-averaged maps. The shape of the curve depends on the imposed symmetry, mask and whether or not the two 3D reconstructions used were processed from a common reference. The reported resolution is shown as a black line. A curve is displayed for the half-bit criterion in addition to lines showing the 0.143 gold standard cut-off and 0.5 cut-off.

### 8.1 FSC [\(i\)](#)



\*Reported resolution corresponds to spatial frequency of 0.303 Å<sup>-1</sup>

## 8.2 Resolution estimates [i](#)

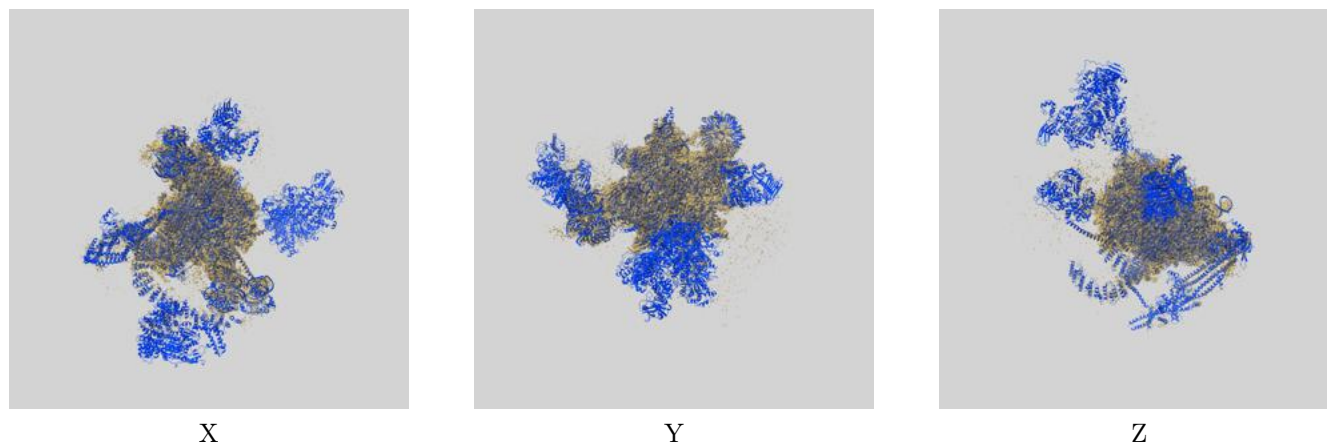
Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	3.30	-	-
Author-provided FSC curve	3.27	3.96	3.33
Unmasked-calculated*	-	-	-

\*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps.

## 9 Map-model fit [i](#)

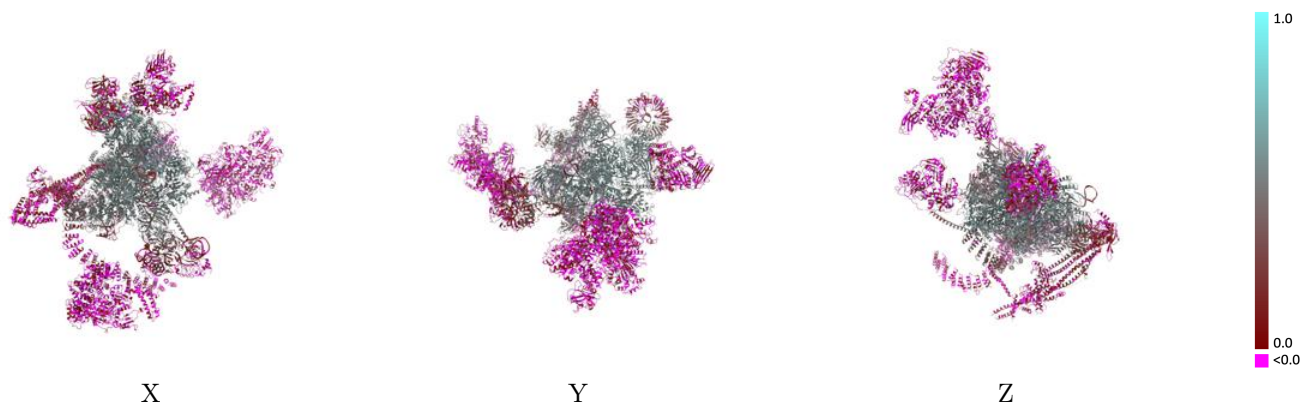
This section contains information regarding the fit between EMDB map EMD-4525 and PDB model 6QDV. Per-residue inclusion information can be found in section 3 on page 17.

### 9.1 Map-model overlay [i](#)



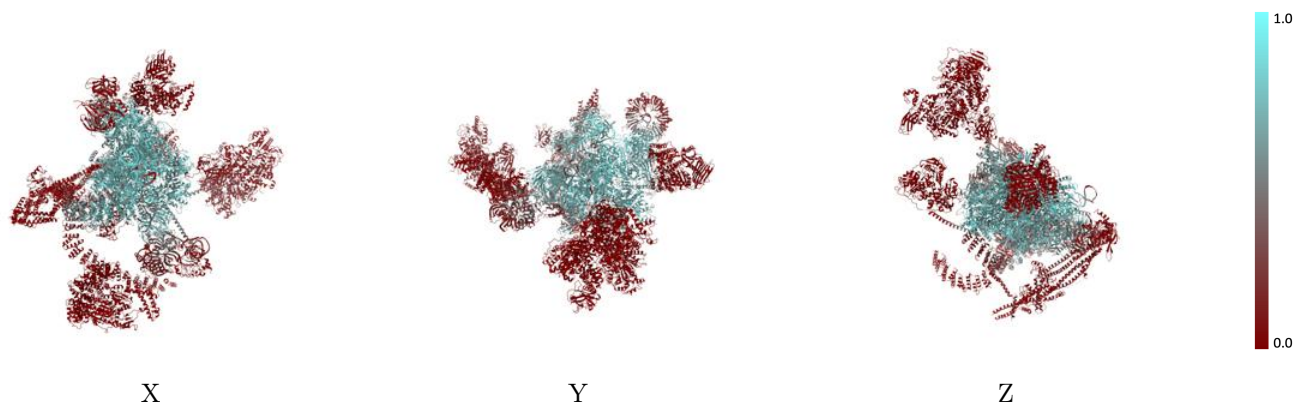
The images above show the 3D surface view of the map at the recommended contour level 0.024 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

## 9.2 Q-score mapped to coordinate model [i](#)



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

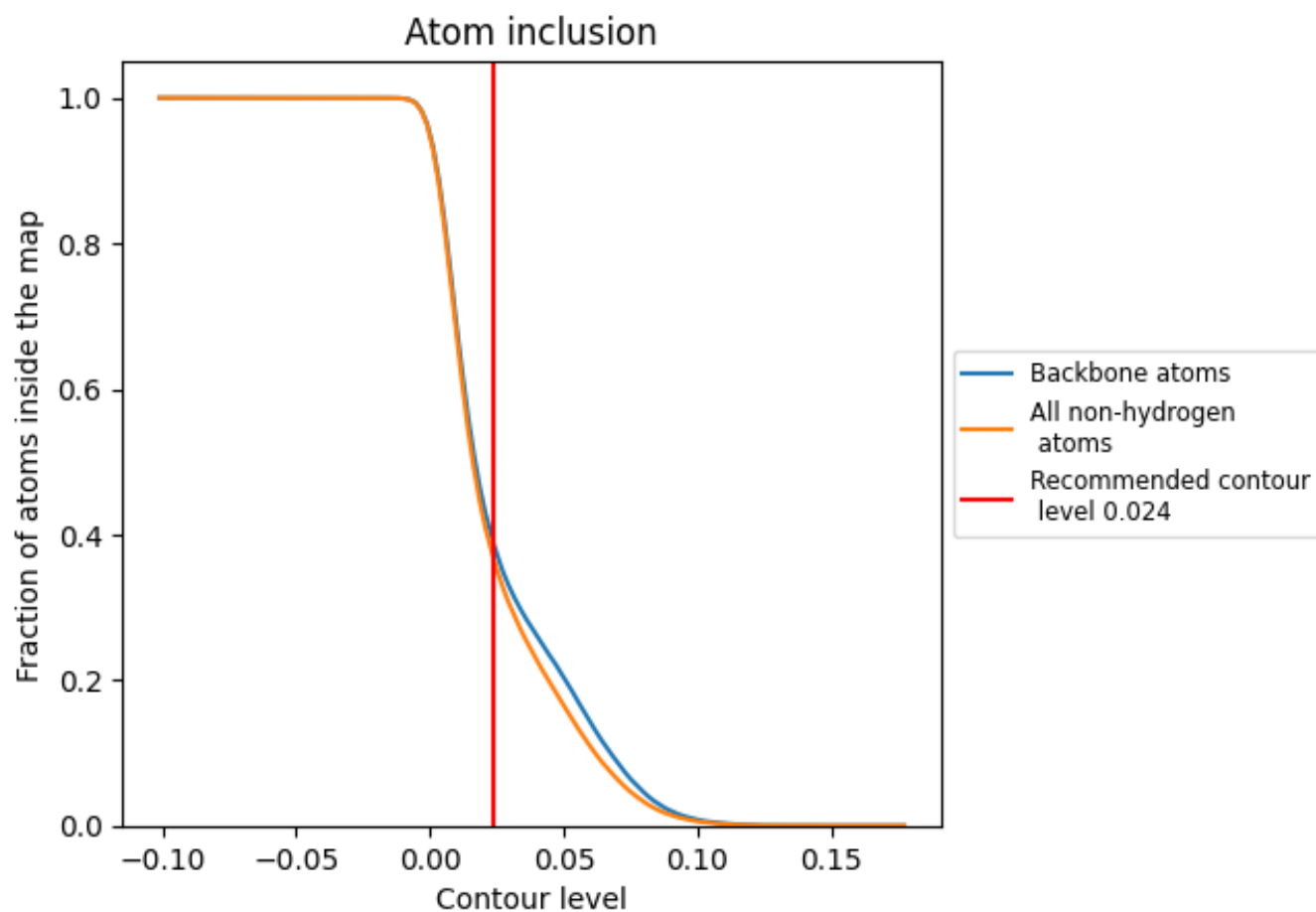
## 9.3 Atom inclusion mapped to coordinate model [i](#)



The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.024).









































































## 9.4 Atom inclusion [i](#)



At the recommended contour level, 38% of all backbone atoms, 36% of all non-hydrogen atoms, are inside the map.

## 9.5 Map-model fit summary









































The table lists the average atom inclusion at the recommended contour level (0.024) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	 0.3639	 0.2780
2	 0.4154	 0.2640
5	 0.7511	 0.4540
6	 0.7381	 0.4570
7	 0.0049	 0.0600
8	 0.0000	 -0.0230
9	 0.0000	 0.0230
A	 0.6844	 0.4780
B	 0.0007	 0.0020
C	 0.6848	 0.5010
D	 0.4081	 0.3490
E	 0.8412	 0.5300
F	 0.6278	 0.4670
G	 0.5772	 0.5060
H	 0.3249	 0.2530
I	 0.6044	 0.3950
J	 0.8354	 0.5600
K	 0.5775	 0.4700
L	 0.7396	 0.5240
M	 0.5155	 0.4410
N	 0.6243	 0.4720
O	 0.4296	 0.3550
P	 0.6534	 0.5150
R	 0.8342	 0.5500
S	 0.4421	 0.3340
T	 0.0662	 0.1560
U	 0.0014	 -0.0010
V	 0.0732	 0.0630
W	 0.0215	 0.0440
Y	 0.0479	 0.0930
Z	 0.6516	 0.5040
b	 0.2809	 0.2660
c	 0.6516	 0.4960
d	 0.4341	 0.3790
e	 0.0740	 0.1380



*Continued on next page...*

Continued from previous page...

Chain	Atom inclusion	Q-score
f	 0.0417	 0.1010
g	 0.2161	 0.2880
h	 0.1072	 0.1720
i	 0.5726	 0.4500
j	 0.0501	 0.0950
k	 0.2065	 0.2370
l	 0.1232	 0.1500
m	 0.0607	 0.0830
n	 0.3693	 0.3190
o	 0.6617	 0.4960
p	 0.2236	 0.2290
q	 0.1141	 0.1190
r	 0.3482	 0.3330
s	 0.0730	 0.1200
t	 0.0258	 0.0690
u	 0.0163	 0.0240
v	 0.0237	 0.1160
w	 0.0098	 0.0760
y	 0.6404	 0.4890
z	 0.0112	 0.1960