Full wwPDB EM Validation Report (i)

Nov 29, 2022 – 02:42 AM JST

PDB ID	:	7W59
EMDB ID	:	EMD-32317
Title	:	The cryo-EM structure of human pre-C*-I complex
Authors	:	Zhan, X.; Lu, Y.; Shi, Y.
Deposited on	:	2021-11-29
Resolution	:	3.60 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

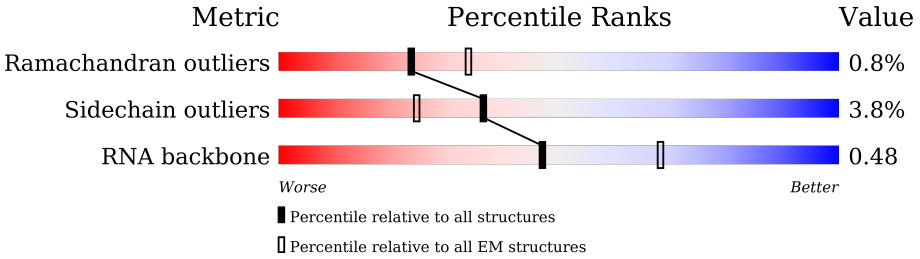
EMDB validation analysis	:	0.0.1.dev43
Mogul	:	1.8.5 (274361), CSD as541be (2020)
MolProbity	:	4.02b-467
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	1.9.9
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

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 3.60 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



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

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the EM map (all-atom inclusion < 40%). The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	2335	79% 5% • 15%
2	C	972	85% • 11%
3	E	357	83% • 16%
4	4	46	9% 15% • 72%
5	G	174	9% 9% 30% 9% 53%
6	J	848	65% • 33%
7	L	802	53% • 46%
8	M	243	52% • 47%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	N	144	91% 6%
10	O	420	66% 32%
11	P	229	44% 51%
12	R	536	45% 7% 48%
13	S	166	94%
14	T	514	55% 39%
15	U	2752	97%
16	V	908	21% 49% 50%
17	W	579	84% 12%
18	B	117	54% 16% 28%
19	F	107	46% 43% 9%
20	H	188	6% 28% 43% 26%
21	b	240	6% 40% 58%
21	i	240	14% 35% 64%
22	X	254	17% 10% 73%
23	I	855	70% 28%
24	y	301	26% 74%
25	Y	1220	41% 54% 45%
26	a	126	5% 64% 36%
26	h	126	12% 64% 36%
27	c	119	67% 31%
27	j	119	21% 67% 31%
28	d	118	14% 81% 18%
28	k	118	24% 70% 28%
29	f	86	16% 85% 14%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
29	m	86	10% 85% 14%
30	e	92	16% 86% 14%
30	l	92	23% 85% 15%
31	g	76	29% 96%
31	n	76	8% 89% 9%
32	v	146	99% 99%
33	w	174	52% 51% 48%
34	u	411	81% 93% 6%
35	x	703	96%
36	Q	1485	35% 88% 11%
37	o	255	24% 57% 6% 36%
38	p	225	39% 58%
39	q	504	10% 25% 74%
39	r	504	25% 74%
39	s	504	13% 24% 74%
39	t	504	15% 25% 74%
40	K	225	67% 31%

2 Entry composition (i)

There are 45 unique types of molecules in this entry. The entry contains 92315 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	1984	16449	10601	2879	2899	70	0	0

- Molecule 2 is a protein called 116 kDa U5 small nuclear ribonucleoprotein component.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	C	862	6798	4347	1138	1281	32	0	0

- Molecule 3 is a protein called U5 small nuclear ribonucleoprotein 40 kDa protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	E	299	2338	1470	410	445	13	0	0

- Molecule 4 is a RNA chain called pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
4	4	13	276	123	50	90	13	0	0

- Molecule 5 is a RNA chain called pre-mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
5	G	82	1510	666	210	552	82	0	0

- Molecule 6 is a protein called Crooked neck-like protein 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	J	568	3814	2376	717	715	6	0	0

- Molecule 7 is a protein called Cell division cycle 5-like protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	L	437	3015	1859	584	565	7	0	0

- Molecule 8 is a protein called Pre-mRNA-splicing factor SYF2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
8	M	130	1098	684	204	208	2	0	0

- Molecule 9 is a protein called Protein BUD31 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
9	N	143	1184	746	217	209	12	0	0

- Molecule 10 is a protein called Pre-mRNA-splicing factor RBM22.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	O	285	2296	1442	408	426	20	0	0

- Molecule 11 is a protein called Spliceosome-associated protein CWC15 homolog.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
11	P	113	953	583	189	179	2	0	0

- Molecule 12 is a protein called SNW domain-containing protein 1.

Mol	Chain	Residues	Atoms						AltConf	Trace
			Total	C	N	O	P	S		
12	R	280	2243	1401	411	416	2	13	0	0

- Molecule 13 is a protein called Peptidyl-prolyl cis-trans isomerase-like 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
13	S	159	1236	787	215	227	7	0	0

- Molecule 14 is a protein called Pleiotropic regulator 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
14	T	312	2454	1550	446	450	8	0	0

- Molecule 15 is a protein called Serine/arginine repetitive matrix protein 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
15	U	72	422	257	82	82	1	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
16	V	452	2632	1639	492	495	6	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
17	W	509	4129	2628	715	762	24	0	0

- Molecule 18 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
18	B	84	1768	792	295	597	84	0	0

- Molecule 19 is a RNA chain called U6 snRNA.

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

- Molecule 20 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
20	H	140	2966	1326	510	990	140	0	0

- Molecule 21 is a protein called Small nuclear ribonucleoprotein-associated proteins B and B'.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
21	b	102	786	492	148	139	7	0	0
21	i	86	690	434	126	123	7	0	0

- Molecule 22 is a protein called PSME3-interacting protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
22	X	69	616	380	111	124	1	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
23	I	617	3845	2380	721	733	11	0	0

- Molecule 24 is a protein called Peptidyl-prolyl cis-trans isomerase E.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
24	y	79	390	232	79	79	0	0

- Molecule 25 is a protein called ATP-dependent RNA helicase DHX8.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
25	Y	667	3431	2057	680	693	1	4	0

- Molecule 26 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
26	a	81	640	401	113	120	6	0	0
26	h	81	633	397	112	118	6	0	0

- Molecule 27 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
27	c	82	649	413	113	119	4	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
27	j	82	649	413	113	119	4	0	0

- Molecule 28 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
28	d	97	776	488	143	140	5	0	0
			Total	C	N	O	S		
28	k	85	688	432	125	126	5	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
29	f	74	576	373	95	103	5	0	0
			Total	C	N	O	S		
29	m	74	576	373	95	103	5	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
30	e	79	652	412	116	119	5	0	0
			Total	C	N	O	S		
30	l	78	644	407	115	118	4	0	0

- Molecule 31 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
31	g	74	577	364	104	103	6	0	0
			Total	C	N	O	S		
31	n	69	538	342	96	94	6	0	0

- Molecule 32 is a protein called Protein mago nashi homolog.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
32	v	144	711	423	144	144	0	0

- Molecule 33 is a protein called RNA-binding protein 8A.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
33	w	91	445	263	91	91	0	0

- Molecule 34 is a protein called Eukaryotic initiation factor 4A-III.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
34	u	386	1907	1135	386	386	0	0

- Molecule 35 is a protein called Protein CASC3.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
35	x	25	124	74	25	25	0	0

- Molecule 36 is a protein called RNA helicase aquarius.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
36	Q	1322	6562	3918	1322	1322	4	0

- Molecule 37 is a protein called U2 small nuclear ribonucleoprotein A'.

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

- Molecule 38 is a protein called U2 small nuclear ribonucleoprotein B".

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
38	p	94	760	488	135	132	5	0	0

- Molecule 39 is a protein called Pre-mRNA-processing factor 19.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
39	q	132	659	395	132	132	0	0
39	r	131	654	392	131	131	0	0
39	s	132	659	395	132	132	0	0

Continued on next page...

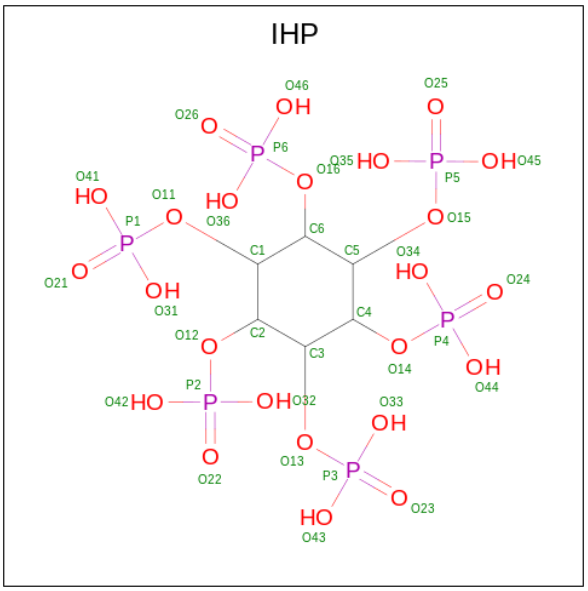
Continued from previous page...

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
39	t	131	654	392	131	131	0	0

- Molecule 40 is a protein called Pre-mRNA-splicing factor SPF27.

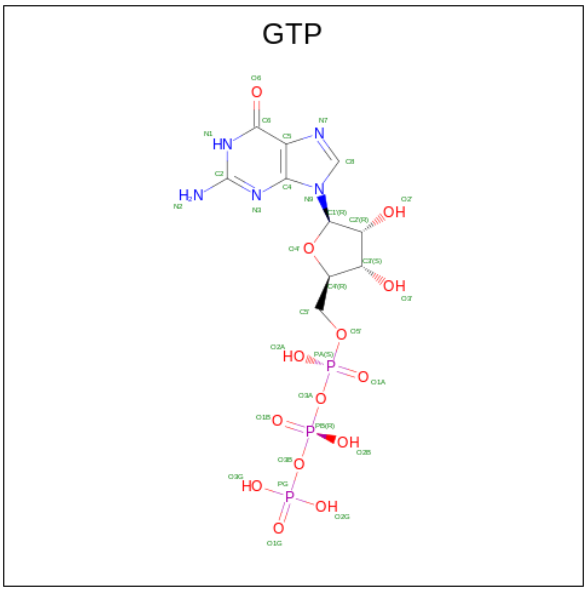
Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
40	K	155	772	462	155	155	0	0

- Molecule 41 is INOSITOL HEXAKISPHOSPHATE (three-letter code: IHP) (formula: $C_6H_{18}O_{24}P_6$) (labeled as "Ligand of Interest" by depositor).



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

- Molecule 42 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$) (labeled as "Ligand of Interest" by depositor).



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

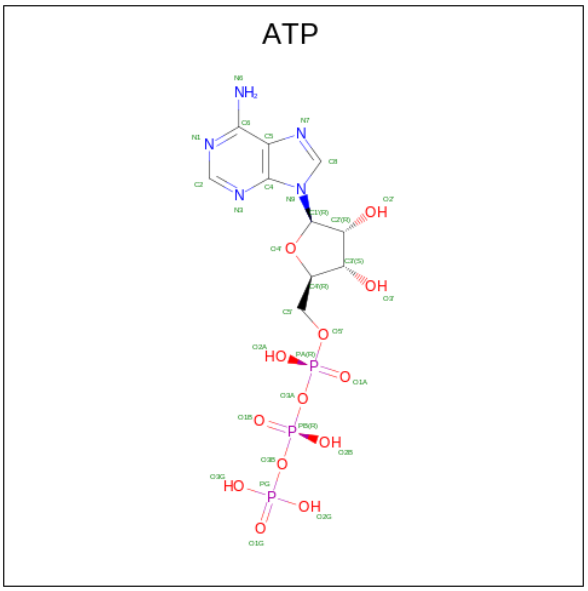
- Molecule 43 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
43	C	1	1	1	0
43	F	6	6	6	0
43	Q	2	2	2	0

- Molecule 44 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
			Total	Zn	
44	N	3	3	3	0
44	O	3	3	3	0

- Molecule 45 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$) (labeled as "Ligand of Interest" by depositor).



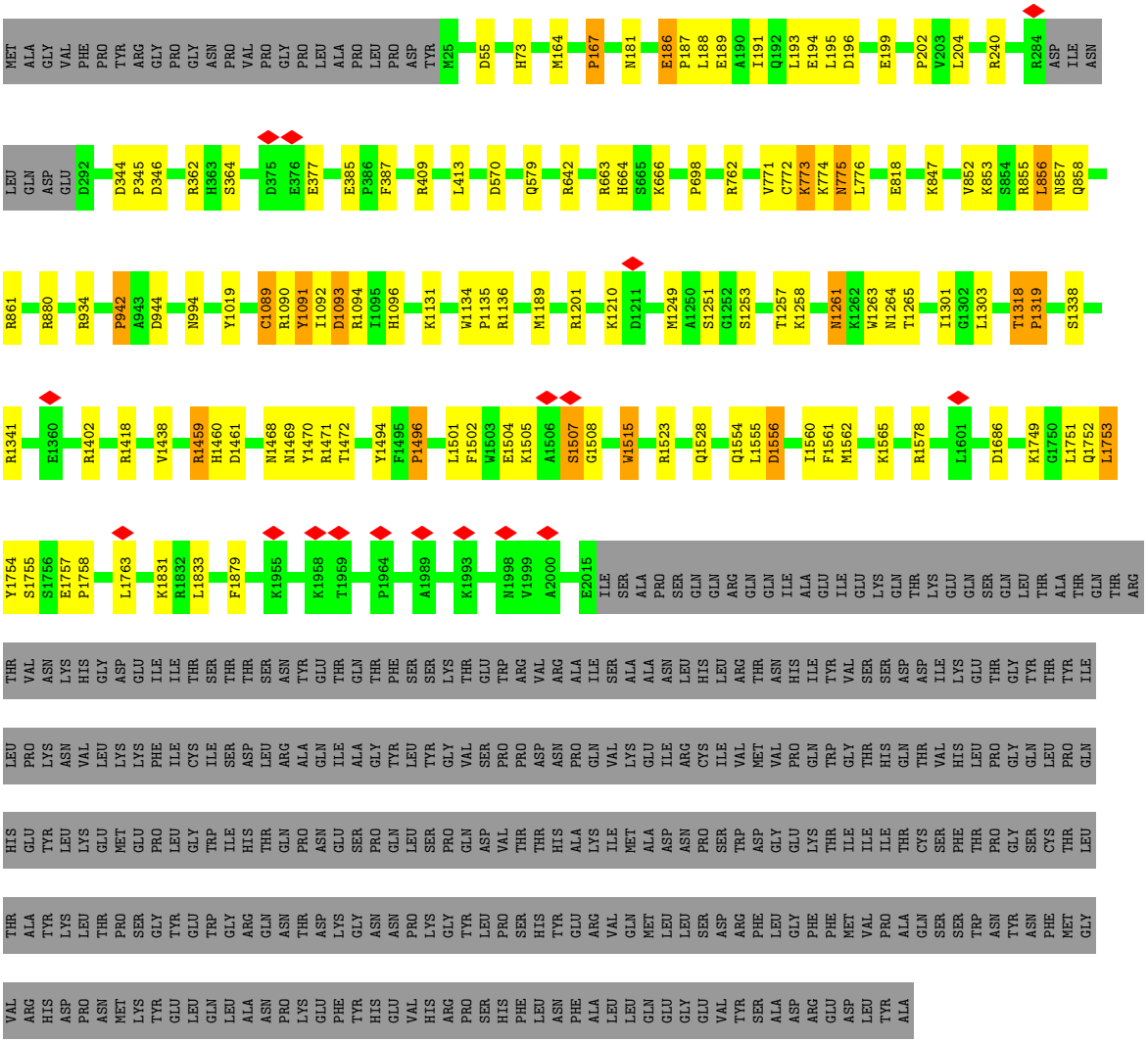
Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
45	Q	1	31	10	5	13	3	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and atom inclusion in map density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

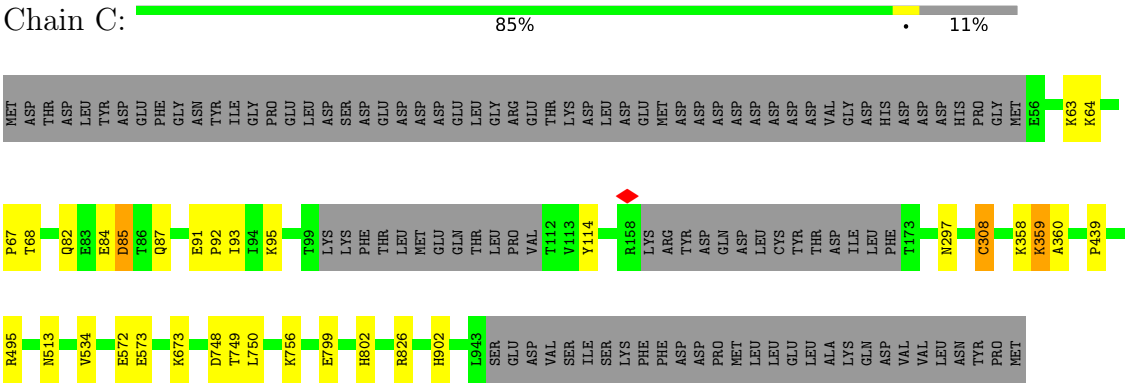
• Molecule 1: Pre-mRNA-processing-splicing factor 8

Chain A: 79% 5% 15%



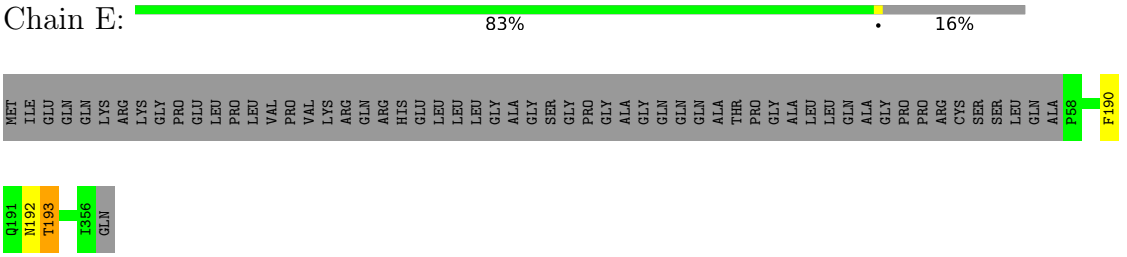
• Molecule 2: 116 kDa U5 small nuclear ribonucleoprotein component

Chain C:



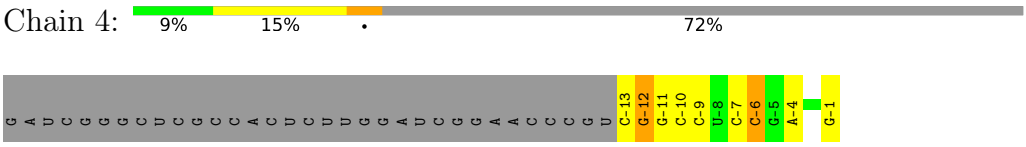
- Molecule 3: U5 small nuclear ribonucleoprotein 40 kDa protein

Chain E:



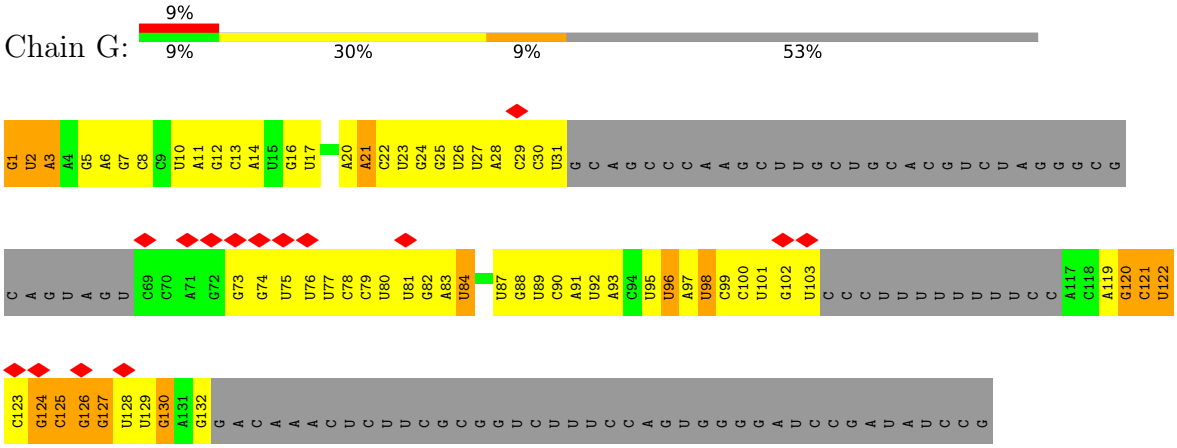
- Molecule 4: pre-mRNA

Chain 4:



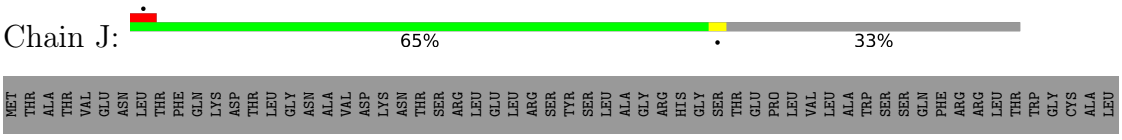
- Molecule 5: pre-mRNA

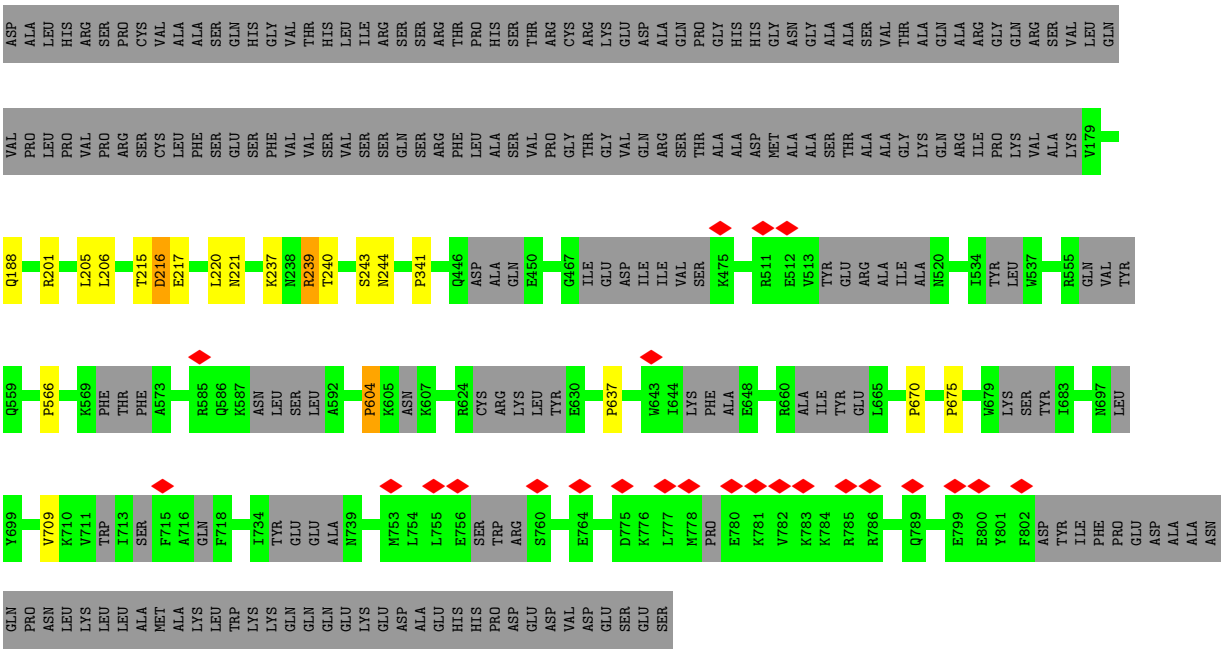
Chain G:



- Molecule 6: Crooked neck-like protein 1

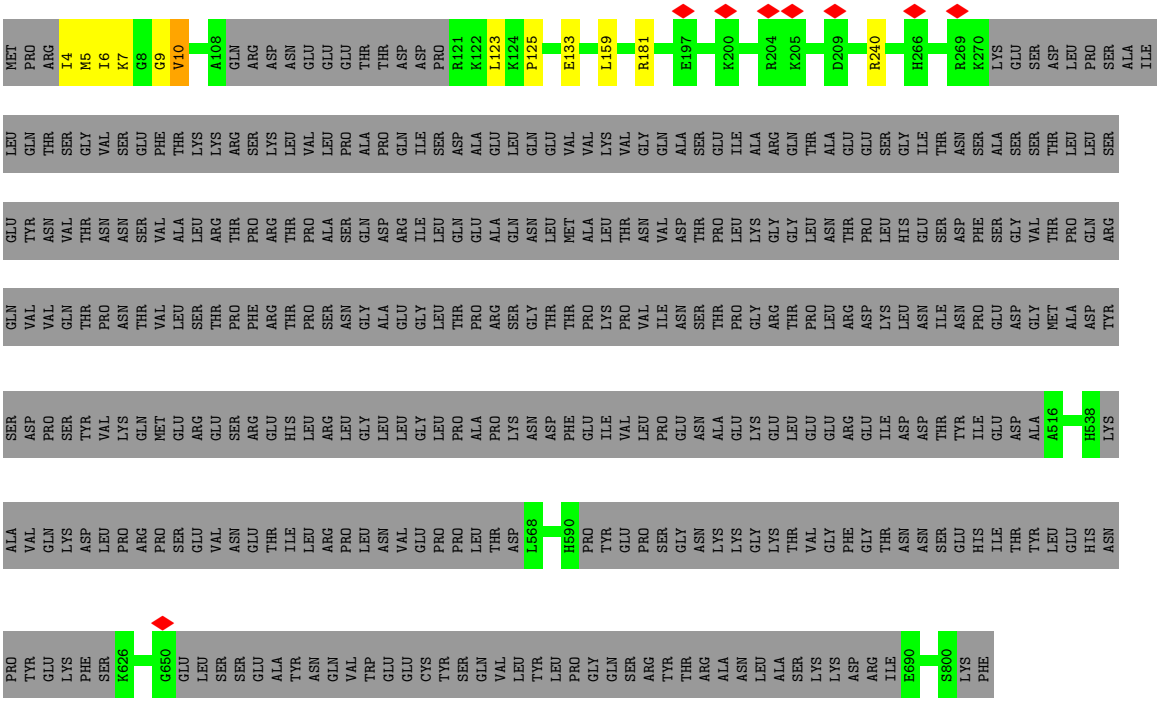
Chain J:





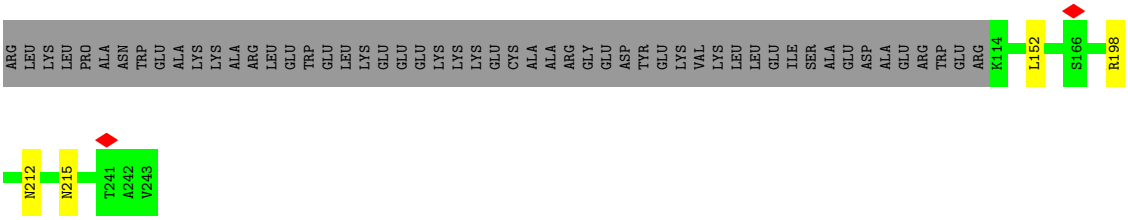
• Molecule 7: Cell division cycle 5-like protein

Chain L:



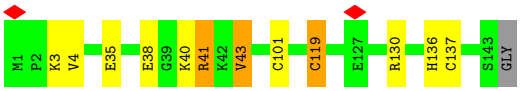
• Molecule 8: Pre-mRNA-splicing factor SYF2

Chain M:



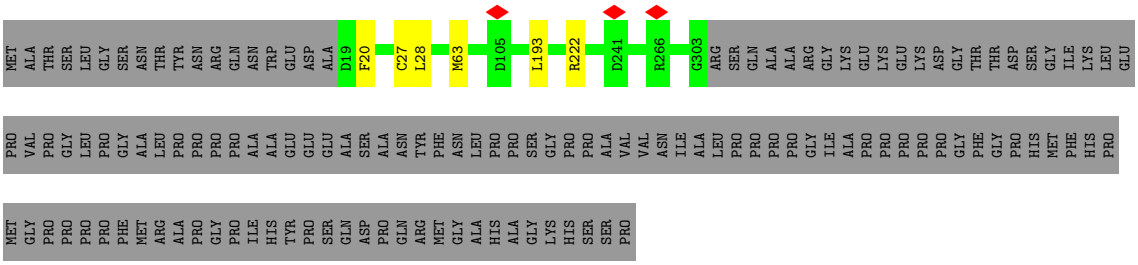
• Molecule 9: Protein BUD31 homolog

Chain N:



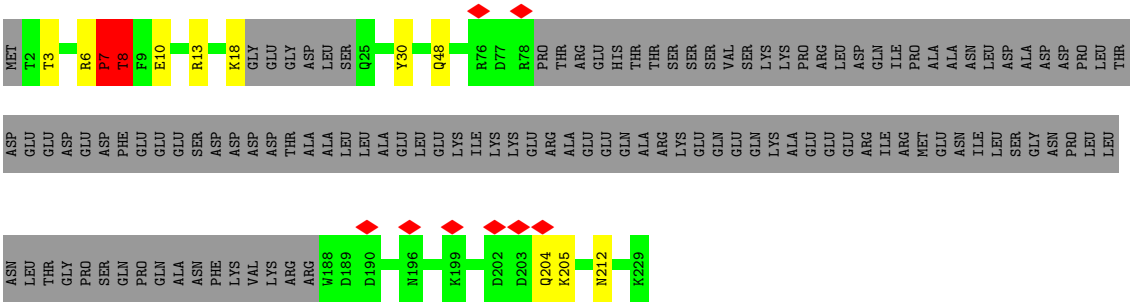
• Molecule 10: Pre-mRNA-splicing factor RBM22

Chain O:



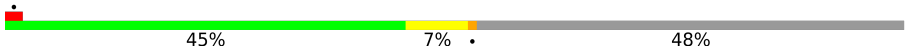
• Molecule 11: Spliceosome-associated protein CWC15 homolog

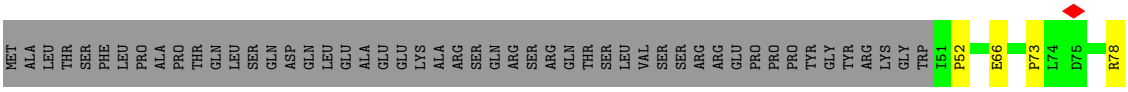
Chain P:

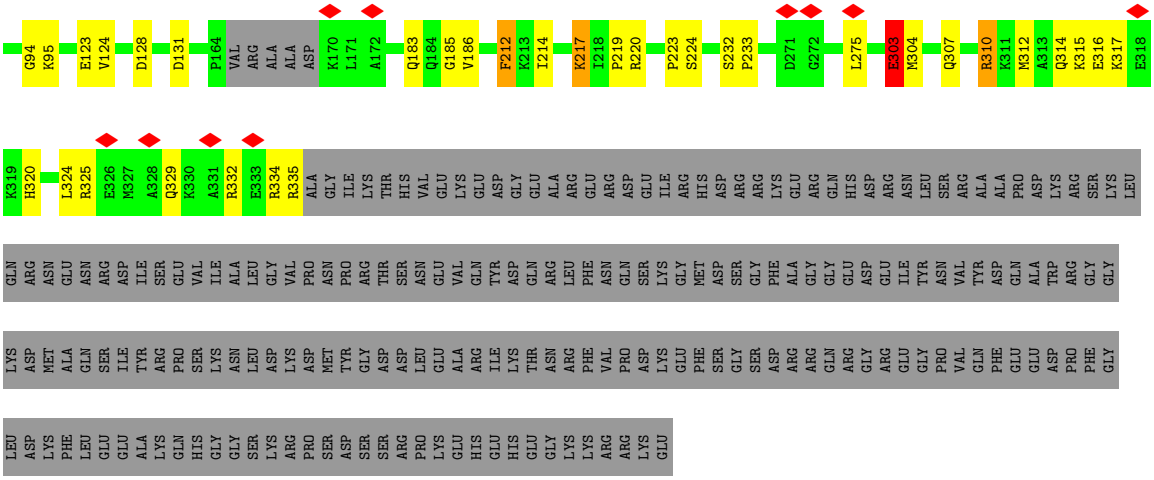


• Molecule 12: SNW domain-containing protein 1

Chain R:

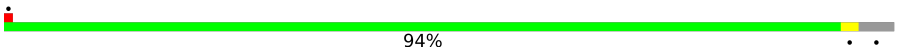


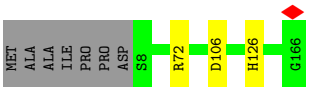




• Molecule 13: Peptidyl-prolyl cis-trans isomerase-like 1

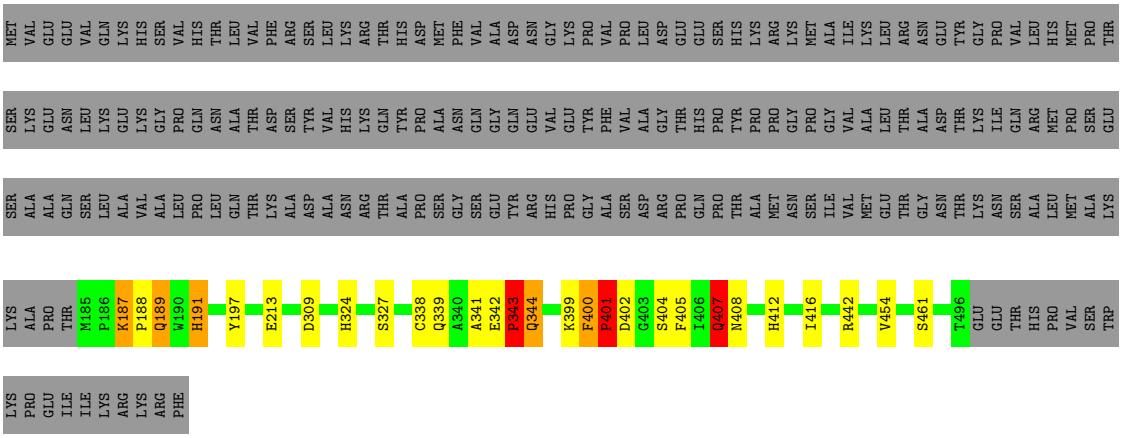
Chain S:





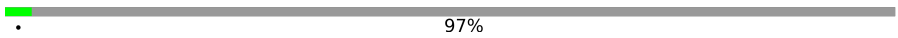
• Molecule 14: Pleiotropic regulator 1

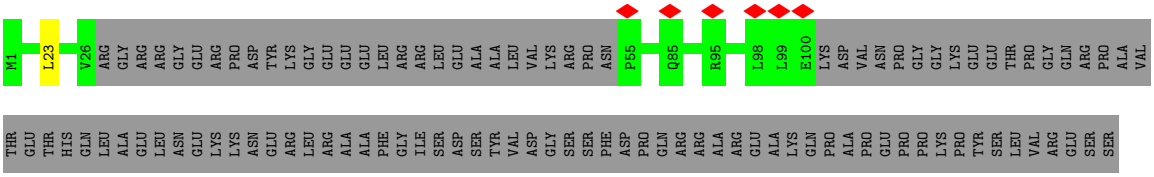
Chain T:



• Molecule 15: Serine/arginine repetitive matrix protein 2

Chain U:





SER SER ARG SER PRO THR PRO LYS GLN LYS LYS LYS LYS LYS LYS LYS ASP ARG GLY ARG ARG SER GLU SER SER SER PRO ARG ARG GLU ARG LYS LYS SER SER LYS LYS LYS LYS HIS ARG SER GLU SER GLU SER LYS LYS ARG LYS HIS ARG SER PRO THR PRO LYS SER LYS ARG

LYS SER LYS ASP LYS LYS ARG LYS ARG SER ARG SER THR THR PRO ALA PRO LYS SER ARG ARG ALA HIS ARG SER THR SER ALA ASP SER ALA SER SER SER ASP THR SER ARG SER ARG SER ARG SER ALA ALA ALA LYS THR HIS THR THR ALA LEU ALA GLY ARG SER PRO SER PRO

ALA SER GLY ARG ARG GLY GLU GLY ASP ALA PRO PHE SER GLU PRO GLY THR THR SER THR GLN ARG PRO SER SER PRO GLU THR ALA THR LYS GLN PRO SER SER PRO TYR GLU ASP LYS ASP LYS ASP LYS LYS GLU LYS SER ALA THR ARG PRO SER PRO SER PRO GLU ARG SER SER

THR GLY PRO GLU PRO PRO ALA PRO THR PRO LEU LEU ALA GLU ARG HIS GLY GLY SER PRO GLN PRO LEU ALA THR THR PRO LEU SER GLN GLU PRO VAL ASN PRO PRO SER GLU ALA SER PRO THR ARG ASP ARG SER PRO PRO LYS SER PRO GLU LYS LEU PRO GLN SER SER SER SER

GLU SER SER PRO PRO SER PRO GLN PRO THR LYS VAL SER ARG HIS ALA SER SER SER PRO GLU SER PRO LYS PRO ALA PRO ALA PRO GLY SER HIS ARG GLU ILE SER SER SER PRO THR SER LYS ASN ARG SER HIS GLY ARG ALA LYS ARG ASP LYS SER HIS SER HIS THR PRO SER

ARG ARG MET GLY ARG SER ARG SER PRO ALA THR ALA LYS ARG GLY ARG SER ARG SER ARG THR PRO THR LYS ARG GLY HIS SER ARG SER ARG SER PRO GLN TRP ARG ARG SER ARG SER ALA GLN ARG TRP GLY ARG SER ARG SER PRO GLN ARG ARG GLY ARG SER ARG SER PRO GLN

ARG PRO GLY TRP SER ARG SER ARG ASN THR GLN ARG ARG GLY ARG SER ARG SER ALA ARG ARG GLY ARG SER HIS SER ARG SER PRO ALA THR ARG GLY ARG SER ARG SER ARG THR PRO ALA ARG ARG GLY ARG SER ARG SER ARG THR PRO ALA ARG ARG ARG SER ARG SER ARG THR

PRO THR ARG ARG ARG SER ARG SER ARG THR PRO ALA ARG ARG GLY ARG SER ARG SER ARG THR PRO ALA ARG ARG ARG SER ARG THR ARG SER PRO VAL ARG ARG ARG SER ARG SER ARG SER PRO ALA ARG ARG SER GLY ARG SER ARG SER ARG THR PRO ALA ARG ARG GLY ARG SER

ARG SER ARG THR PRO ALA ARG ARG GLY ARG SER ARG SER ARG THR PRO ALA ARG ARG SER GLY ARG SER ARG SER ARG THR PRO ALA ARG ARG GLY ARG SER ARG SER ARG THR PRO ARG ARG GLY ARG SER ARG SER ARG SER LEU VAL ARG ARG GLY ARG SER HIS SER ARG THR PRO

GLN ARG ARG GLY ARG SER GLY SER SER SER GLU ARG LYS ASN LYS SER ARG THR SER GLN ARG ARG SER ARG SER ASN SER SER PRO GLU MET LYS LYS SER ARG ILE SER SER ARG ARG SER ARG SER LEU SER SER PRO ARG SER LYS ALA LYS SER ARG LEU SER LEU ARG ARG SER

LEU SER GLY SER SER PRO CYS PRO LYS GLN LYS SER GLN THR PRO PRO ARG ARG SER ARG SER GLY SER SER GLN PRO LYS ALA LYS SER ARG THR PRO PRO ARG ARG SER ARG SER SER SER SER PRO PRO PRO LYS GLN LYS SER LYS THR PRO SER ARG GLN SER HIS SER SER SER

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

VAL LYS SER SER THR PRO PRO ARG GLN SER PRO SER ARG SER SER SER PRO GLN PRO LYS VAL LYS ALA ILE ILE SER PRO ARG GLN ARG SER HIS SER GLY SER SER SER PRO SER PRO SER ARG VAL THR SER ARG THR THR PRO ARG ARG SER ARG SER VAL SER PRO CYS SER ASN

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

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

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

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

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

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

ASN SER PRO LEU ARG GLU ASN SER PHE GLY SER PRO LEU GLU PHE ARG ASN SER GLY PRO LEU GLY THR GLU MET ASN THR GLY PHE SER SER GLU VAL LYS GLU ASP LEU ASN GLY PRO PHE LEU ASN GLN LEU GLU THR ASP PRO SER LEU ASP MET LYS GLU GLN SER THR ARG SER

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

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

SER PRO SER SER PRO GLU LEU ASN ASN LYS CYS LEU THR PRO GLN ARG GLU ARG SER GLY SER GLU SER SER VAL ASP GLN LYS THR VAL ALA ARG THR PRO LEU GLY GLN ARG SER ARG SER GLY SER SER GLN GLU LEU ASP VAL LYS PRO SER ALA SER PRO GLN GLU ARG SER GLU

SER ASP SER SER PRO ASP SER LYS ALA LYS THR ARG THR PRO LEU ARG GLN ARG SER ARG SER GLY SER SER PRO GLU VAL ASP SER LYS SER ARG LEU SER PRO ARG ARG SER ARG SER GLY SER SER PRO GLU VAL LYS ASP LYS PRO ARG ALA ALA PRO ARG ALA GLN SER GLY SER

ASP SER SER PRO GLU PRO LYS ALA PRO ALA PRO ARG ALA LEU PRO ARG ARG SER ARG SER GLY SER SER SER LYS GLY ARG GLY PRO SER PRO GLU GLY SER SER SER THR GLU SER SER PRO GLU HIS PRO PRO LYS SER ARG THR ALA ARG ARG GLY SER ARG SER SER PRO GLU PRO

LYS THR LYS SER ARG THR PRO PRO ARG ARG ARG SER SER ARG SER SER PRO GLU LEU THR ARG LYS ALA ARG LEU SER ARG ARG SER ARG SER ALA SER SER SER PRO GLU THR ARG SER ARG THR PRO PRO ARG HIS ARG ARG SER PRO SER VAL SER SER PRO GLU PRO ALA GLU LYS

SER ARG SER SER ARG ARG ARG ARG SER ALA SER SER PRO ARG THR LYS THR THR SER ARG ARG GLY ARG SER PRO SER PRO LYS PRO ARG GLY LEU GLN ARG SER ARG SER ARG SER ARG ARG GLU LYS THR ARG THR THR ARG ARG ARG ASP ARG SER GLY SER SER GLN SER THR SER

ARG ARG ARG GLN ARG SER ARG SER ARG SER ARG VAL THR ARG ARG ARG ARG GLY GLY SER GLY TYR HIS SER ARG SER PRO ALA ARG GLN GLU SER SER ARG THR SER SER ARG ARG ARG ARG GLY ARG SER ARG THR PRO PRO THR SER ARG LYS ARG SER ARG SER ARG THR SER PRO

ALA PRO TRP LYS ARG SER ARG SER ARG ALA SER PRO ALA THR HIS ARG ARG SER ARG SER ARG THR PRO LEU ILE SER ARG ARG ARG SER ARG SER ARG THR SER PRO VAL SER ARG ARG ARG SER ARG SER ARG THR SER VAL THR ARG ARG ARG SER ARG SER ARG ALA SER PRO VAL

SER ARG ARG ARG SER ARG SER ARG THR PRO PRO VAL THR ARG ARG ARG SER ARG SER ARG THR PRO THR THR ARG ARG ARG SER ARG SER ARG THR PRO PRO VAL THR ARG ARG ARG SER ARG SER ARG THR PRO PRO VAL THR ARG ARG ARG SER ARG SER ARG THR SER PRO ILE THR

ARG ARG ARG SER ARG SER ARG THR SER PRO VAL THR ARG ARG ARG SER ARG SER ARG THR SER PRO VAL THR ARG ARG ARG SER ARG SER ARG THR SER PRO VAL THR ARG ARG ARG SER ARG SER ARG THR PRO PRO ALA ILE ARG ARG ARG SER ARG SER ARG THR PRO LEU LEU PRO

ARG LYS ARG SER ARG SER ARG SER PRO LEU ALA ILE ARG ARG ARG SER ARG SER ARG THR PRO ARG THR ALA ARG GLY LYS ARG SER LEU THR ARG SER PRO PRO ALA ILE ARG ARG ARG SER ALA SER GLY SER SER SER ASP ARG SER ARG SER ALA THR PRO PRO ALA THR ARG ASN

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

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

SER LEU ARG THR ALA PRO ALA ALA ASN LEU ALA SER ARG ILE PRO ALA ALA SER ALA ALA ALA MET ASN LEU ALA SER ALA ARG THR PRO ALA ILE PRO THR ALA VAL ASN LEU ALA ASP SER ARG THR PRO ALA ALA ALA ALA ALA MET ASN LEU ALA SER PRO ARG THR ALA VAL ALA

PRO SER ALA VAL ASN LEU ALA ASP PRO ARG THR PRO THR ALA PRO ALA VAL ASN LEU ALA GLY ALA ARG THR PRO ALA ALA LEU ALA ALA LEU SER LEU THR GLY SER GLY THR PRO PRO THR ALA ALA ASN TYR PRO SER SER SER ARG THR PRO GLN ALA PRO ALA SER ALA ASN LEU

VAL GLY PRO ARG SER ALA HIS ALA THR ALA PRO VAL ASN ILE ALA GLY SER ARG THR ALA ALA ALA LEU ALA PRO ALA SER LEU THR SER ALA ARG MET ALA PRO ALA LEU SER GLY ALA ASN LEU THR SER PRO ARG VAL PRO LEU SER ALA TYR GLU ARG VAL SER GLY ARG THR SER

PRO PRO LEU LEU ASP ARG ALA ARG SER ARG THR PRO PRO SER ALA PRO SER GLN SER ARG MET THR SER GLU ARG ALA PRO SER PRO SER SER ARG MET GLY GLN ALA PRO SER GLN SER LEU LEU PRO PRO ALA GLN ASP GLN PRO ARG SER PRO VAL PRO SER ALA PHE SER ASP GLN

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

SER ALA LEU ALA ALA LEU GLN PRO ALA LYS GLU ARG ARG SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER GLY SER SER SER SER ASP SER GLU GLY SER SER LEU PRO VAL GLN PRO GLU VAL ALA LEU LYS ARG

VAL PRO SER PRO THR PRO ALA PRO LYS GLU ALA VAL ARG GLU GLY ARG PRO PRO GLU PRO THR PRO ALA LYS ARG LYS ARG ARG SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER SER

SER SER SER SER SER SER SER SER SER SER PRO SER PRO ALA LYS PRO GLY PRO GLN ALA LEU PRO LYS PRO ALA SER PRO LYS LYS PRO PRO PRO GLY GLU ARG ARG SER ARG SER PRO ARG LYS PRO ILE ASP SER LEU ARG ASP SER ARG SER LEU SER TYR SER PRO VAL GLU ARG

ARG ARG PRO SER PRO GLN PRO SER PRO ARG ASP GLN GLN SER SER SER SER GLU ARG GLY SER ARG ARG GLY GLN ARG GLY ASP SER ARG SER PRO SER HIS LYS ARG ARG ARG GLU THR PRO SER PRO ARG PRO MET ARG HIS ARG SER SER ARG SER PRO

- Molecule 16: Pre-mRNA-splicing factor CWC22 homolog

Chain V:

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

SER TYR ASP SER SER MET GLU SER ARG ASN ARG ASP ARG GLU LYS ARG ARG GLU ARG GLU ARG ASP THR ASP ARG LYS ARG SER ARG LYS SER PRO SER PRO GLY ARG ARG ASN PRO GLU THR SER VAL THR GLN SER SER SER ALA GLN ASP GLU PRO ALA THR LYS LYS LYS LYS ASP

GLU LEU ASP PRO LEU LEU THR ARG THR GLY GLY ALA TYR ILE PRO PRO ALA LYS LEU ARG MET MET GLN GLU GLN ILE THR ASP K149 N150 S151 L152 A153 Y154 Q155 R156 M157 S158 W159 E160 A161 L162 K163 K164 S165 I166 N167 G168 L169 K172 V173 N174 I175 S176 N177 I178 S179 I180 I181

I182 Q183 E184 L185 L186 Q187 E188 N189 R192 G193 R194 G195 L196 L197 S198 R199 L202 Q203 A204 Q205 S206 A207 S208 P209 I210 F211 T212 H213 V214 Y215 A216 A217 L218 V219 A220 I221 S224 K225 F226 P227 Q228 I229 G230 E231 L232 R236 K243 R247 K250 Q251 L254 T255

A256 S257 K258 A261 H262 L263 I264 N265 Q266 N267 V268 E271 V272 L273 C274 L275 E276 M277 L278 T279 L280 L281 L282 V290 E291 V292 A293 I294 G295 E299 C300 G301 L302 K303 L304 T305 Q306 V307 S308 P309 R310 G311 I312 N313 A314 I315 F316 E325 D345 G346 F347 K348 D349 H350

P351 I352 I353 L354 E355 G356 L357 D358 L359 V360 E361 E362 D363 D364 Q365 F366 T367 H368 M369 L370 P371 D374 D375 Y376 N377 P378 E379 D380 V381 L382 N383 V384 F385 K386 M387 D388 P389 N390 F391 M392 E393 N394 E395 E396 K397 Y398 K399 A400 I401 K402 K403 E404 I405 L406 ASP GLU GLY ASP THR

ASP SER ASN THR ASP GLN ASP ALA GLY SER SER GLU GLU ASP GLU GLU GLU GLU GLU GLU GLU GLY GLU GLU ASP GLU GLU GLY GLN LYS VAL THR ILE HIS ASP LYS T448 E449 I450 T458 L467 R501 E531 D534 T535 S559 LEU PRO W562 E573 T574 S577

S578 F582 P597 L603 E613 GLY LEU LEU PRO ARG D619 N620 P621 R622 F630 S633 I634 L647 K648 ASN THR PRO LYS VAL ILE VAL ALA GLN LYS PRO ASP VAL GLU GLN ASN LYS SER SER PRO SER SER SER SER SER ALA SER SER SER SER GLU SER ASP SER

SER ASP SER ASP SER ASP SER SER ASP SER SER SER GLU SER SER SER GLU GLU SER ASP SER SER SER ILE SER SER HIS SER SER ALA SER ALA ASN ASP VAL ARG LYS LYS GLY HIS GLY LYS THR ARG SER LYS GLU VAL ASP LYS LEU ILE ARG ASN GLN GLN THR ASN ASP ARG

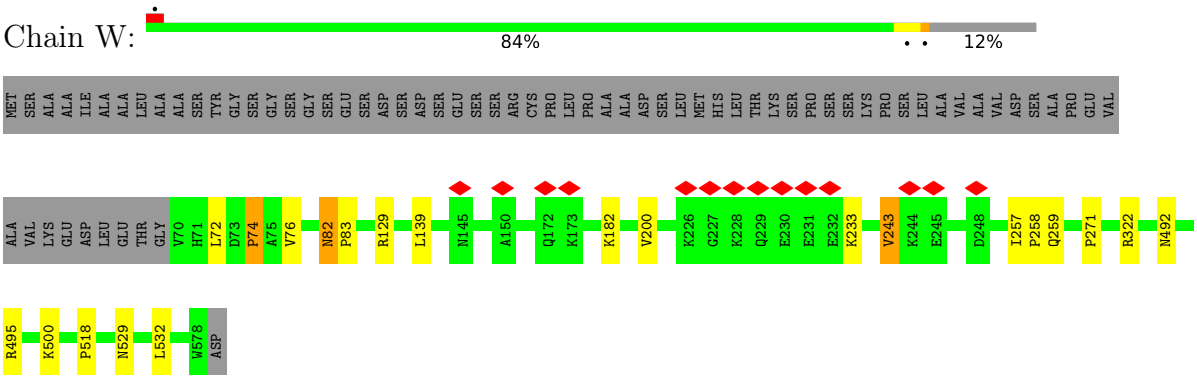
LYS GLN LYS GLU ARG ARG GLN GLU HIS GLY HIS GLN GLU THR ARG THR GLU ARG GLU ARG ARG SER GLU LYS HIS ARG ASP GLN ASN SER SER GLY SER ASN TRP ARG ASP PRO ILE THR LYS TYR THR SER ASP LYS ASP VAL PRO SER GLU ARG ASN ASN TYR SER ARG VAL ALA ASN

ASP ARG ASP GLN GLU MET HIS ILE ASP LEU GLU ASN LYS HIS GLY ASP PRO LYS LYS LYS ARG GLY GLU ARG ARG ASN SER PHE SER GLU ASN GLU LYS HIS THR HIS ARG ILE LYS ASP SER GLU ASN PHE ARG ARG LYS ASP ARG SER LYS SER LYS GLU MET ASN ARG LYS HIS SER

GLY SER ARG SER ASP GLU ASP ARG TYR GLN ASN GLY ALA GLU ARG ARG TRP GLU LYS SER SER ARG TYR SER GLU GLN SER ARG GLU SER LYS LYS ASN GLN ASP ARG ARG ARG GLU LYS SER PRO ALA LYS GLN LYS

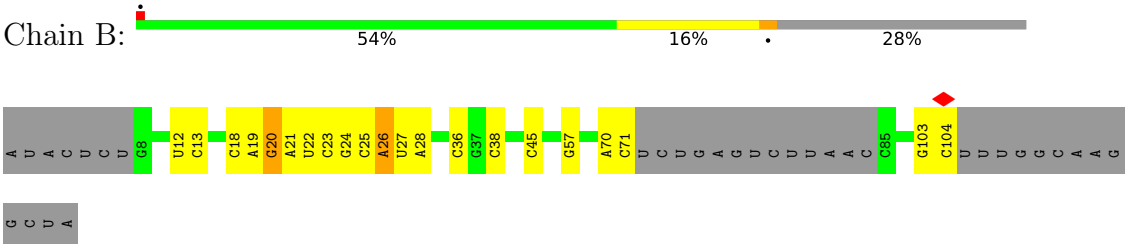
- Molecule 17: Pre-mRNA-processing factor 17

Chain W:



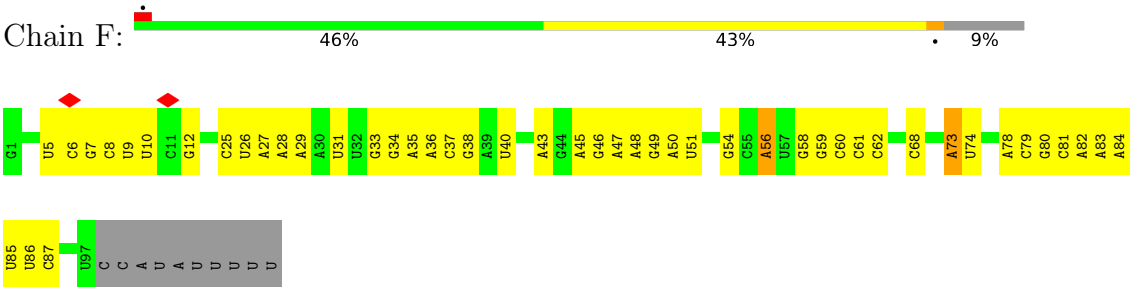
- Molecule 18: U5 snRNA

Chain B:



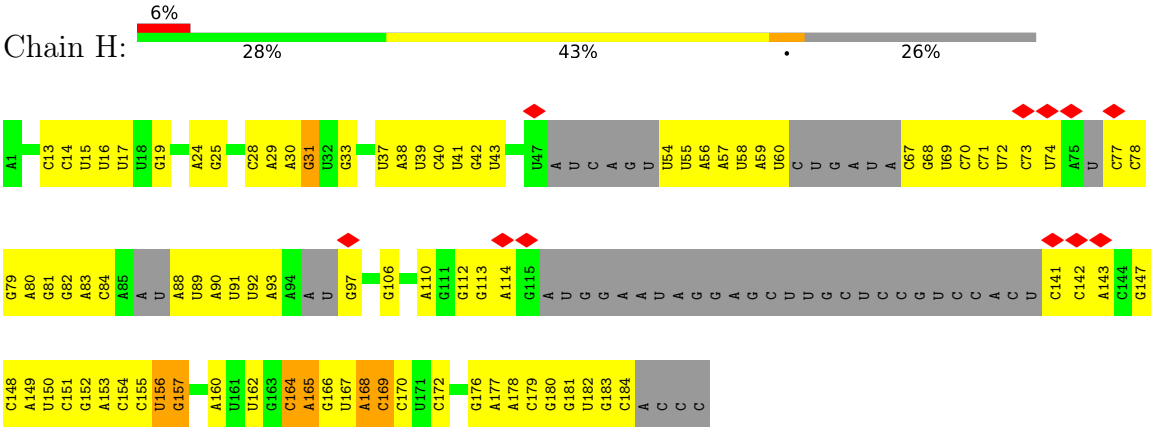
- Molecule 19: U6 snRNA

Chain F:



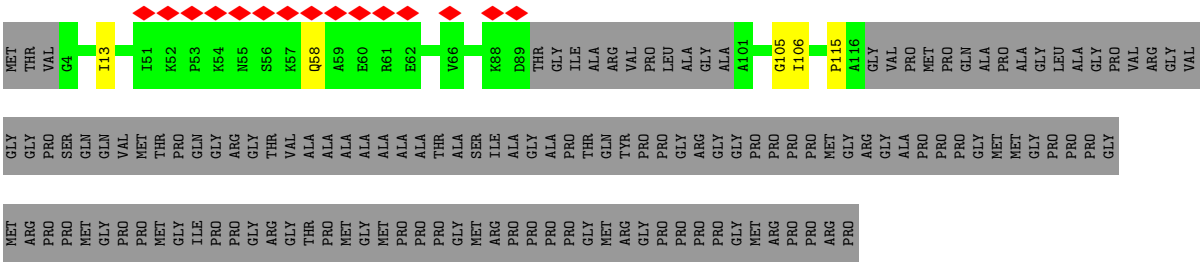
- Molecule 20: U2 snRNA

Chain H:



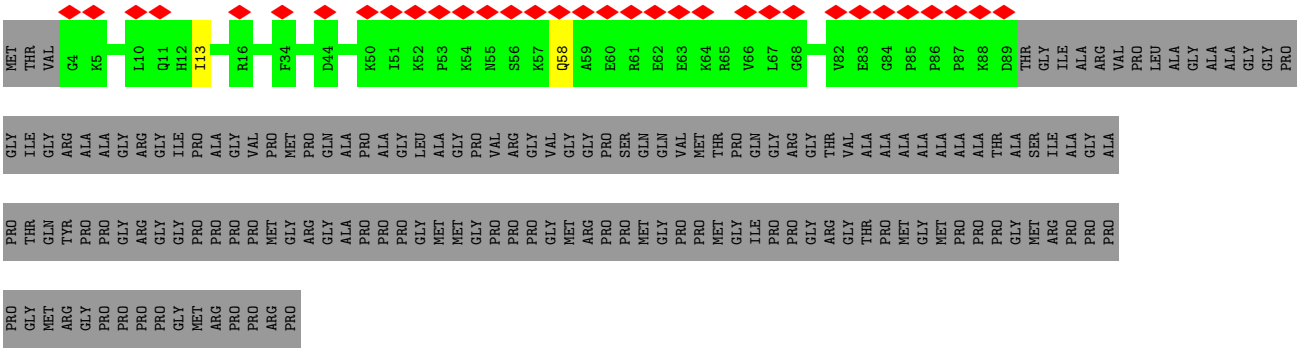
- Molecule 21: Small nuclear ribonucleoprotein-associated proteins B and B'

Chain b:



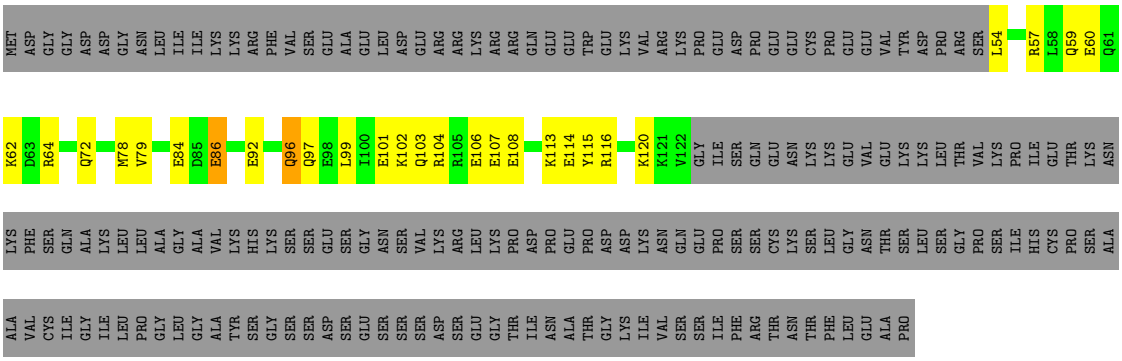
• Molecule 21: Small nuclear ribonucleoprotein-associated proteins B and B'

Chain i:



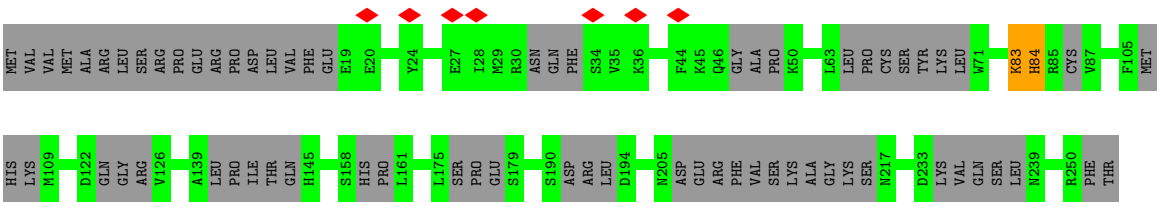
• Molecule 22: PSME3-interacting protein

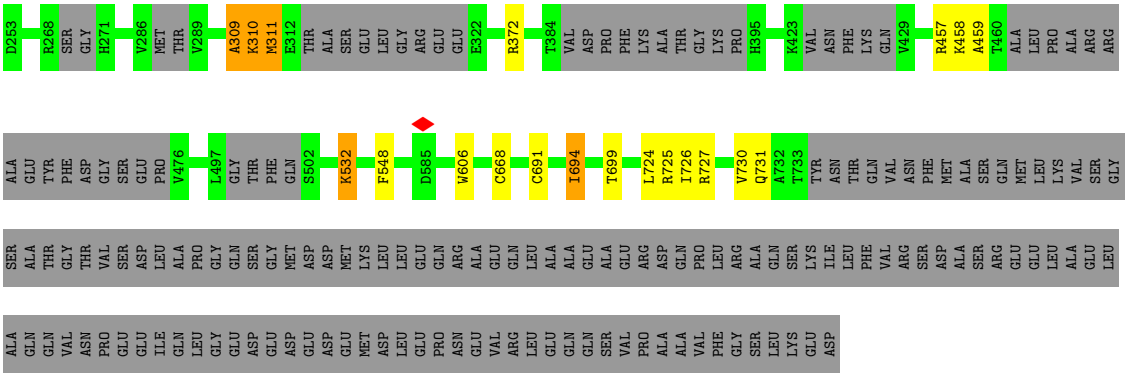
Chain X:



• Molecule 23: Pre-mRNA-splicing factor SYF1

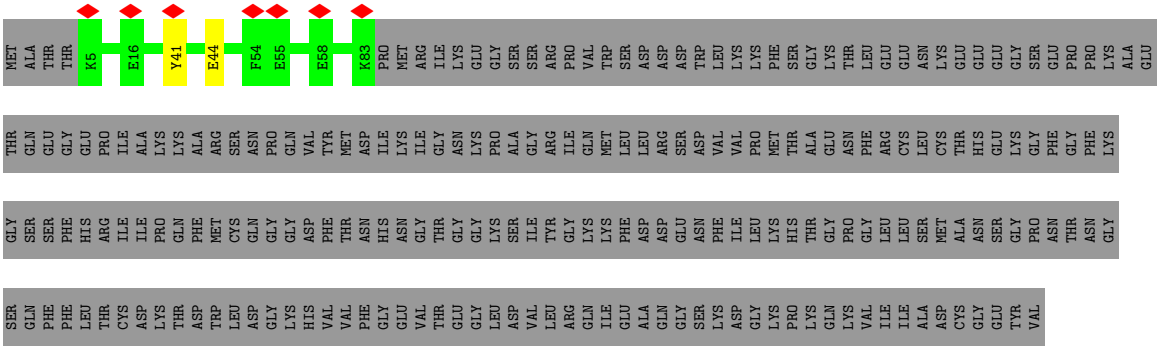
Chain I:





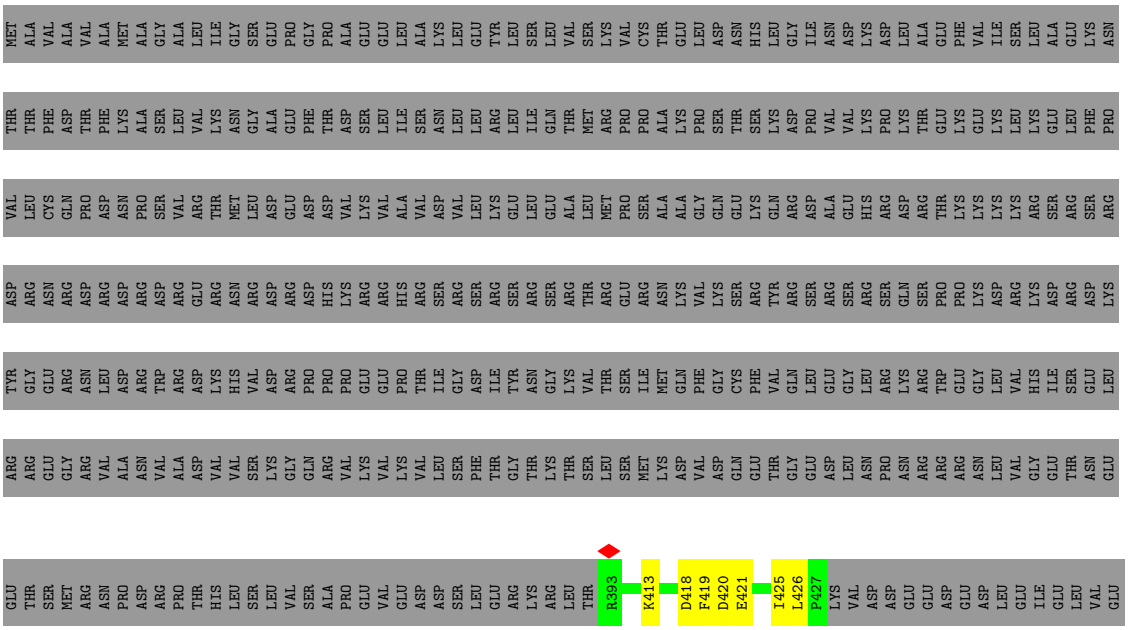
- Molecule 24: Peptidyl-prolyl cis-trans isomerase E

Chain y:



- Molecule 25: ATP-dependent RNA helicase DHX8

Chain Y:



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

GLY LEU ASN LYS HIS TRP VAL ASP PRO LEU PRO ASP ALA GLU GLY ARG GLN ILE ALA ALA ASN MET ARG GLY ILE GLY MET MET PRO ASN ASP ILE PRO GLU TRP LYS LYS HIS ALA PHE GLY GLY ASN LYS ALA SER TYR GLY LYS LYS THR Q555 M556 S557 I558 L559 E560 Q561 R562 E563

S564 L565 P566 I567 Y568 K569 L570 K571 E572 Q573 L574 V575 Q576 A577 V578 H579 D580 N581 Q582 I583 L584 I585 V586 I587 G588 E589 T590 G591 S592 G593 K594 T595 T596 Q597 I598 T599 Q600 Y601 L602 A603 E604 A605 G606 Y607 T608 S609 R610 G611 K612 I613 G614 C615 T616 Q617 P618 R619 R620 V621 A622 A623

M624 S625 V626 A627 K628 R629 V630 S631 E632 E633 F634 G635 C636 C637 L638 G639 Q640 E641 V642 G643 Y644 T645 I646 R647 F648 E649 D650 C651 T652 S653 P654 E655 T656 V657 I658 K659 Y660 M661 T662 D663 G664 M665 L666 L667 R668 E669 C670 L671 I672 D673 P674 D675 L676 T677 Q678 Y679 A680 I681 I682 M683

L684 D685 E686 A687 H688 E689 R690 T691 I692 H693 T694 D695 V696 L697 F698 G699 L700 L701 K702 K703 Q706 K707 R708 Q709 D710 M711 K712 L713 I714 V715 T716 S717 A718 T719 L720 D721 A722 V723 K724 F725 S726 Q727 Y728 F729 Y730 E731 A732 P733 I734 F735 T736 I737 P738 G739 R740 T741 Y742 P743 V744

E745 I746 L747 Y748 T749 K750 E751 P752 E753 T754 D755 D758 A759 S760 L761 I762 T763 V764 M765 Q766 I767 H768 L769 T770 E771 P772 P773 G774 D775 I776 L777 V778 F779 L780 T781 G782 Q783 E784 E785 I786 D787 T788 A789 C790 E791 I792 L793 Y794 E795 R796 M797 K798 S799 L800 G801 P802 D803 V804 P805

E806 L807 I808 I809 L810 P811 V812 Y813 S814 A815 L816 P817 S818 E819 M820 Q821 T822 R823 I824 F825 D826 P827 A828 P829 P830 G831 S832 R833 K834 V835 V836 I837 A838 T839 N840 I841 A842 E843 T844 S845 L846 T847 I848 D849 G850 I851 Y852 Y853 V854 V855 D856 P857 G858 F859 V860 K861 Q862 K863 V864 Y865

N866 S867 K868 T869 G870 I871 D872 Q873 L874 V875 V876 T877 S880 Q881 A882 Q883 A884 K885 Q886 R887 A888 G889 R890 A891 G892 R893 T894 G895 P896 G897 K898 C899 Y900 R901 L902 Y903 T904 E905 R906 A907 Y908 M912 L913 T914 T915 N916 V917 P918 Q921 R922 T923 N924 L925 A926 S927 T928 V929

L930 S931 L932 K933 A934 M935 G936 I937 N938 D939 L940 L941 S942 F943 D944 F945 M946 D947 A948 P949 P950 M951 E952 T953 L954 I955 T956 A957 M958 E959 Q960 L961 Y962 T963 L964 G965 A966 L967 D968 D969 E970 G971 L972 L973 T974 M980 A981 E982 F983 P984 L985 E986 P987 M988 L989 I994 M995

H998 E1004 M1005 I1008 M1011 L1012 S1013 V1014 Q1015 N1016 V1017 F1018 Y1019 R1020 P1021 K1022 D1023 K1024 Q1025 A1026 L1027 A1028 D1029 Q1030 K1031 K1032 A1033 K1034 F1035 H1036 Q1037 T1038 E1039 G1040 L1043 T1044 L1045 L1046 A1047 V1048 Y1049 N1050 S1051 W1052 K1053 N1054 N1055 K1056 F1057 S1058 N1059 P1060 W1061 C1062 Y1063 E1064

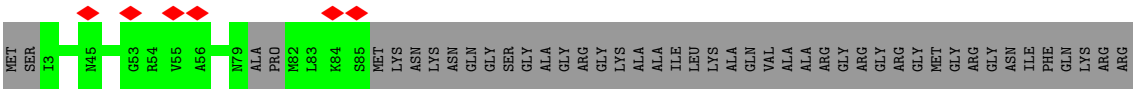
N1065 F1066 I1067 Q1068 S1071 L1072 R1073 R1074 A1075 Q1076 D1077 I1078 R1079 K1080 L1083 G1084 I1085 M1086 D1087 R1088 H1089 K1090 L1091 D1092 V1093 V1094 K1098 S1099 T1100 V1101 R1102 V1103 Q1104 K1105 A1106 S1109 G1110 F1111 F1112 R1113 N1114 K1118 D1119 P1120 Q1121 E1122 I1128 D1129 Q1130 Q1131 V1132 S1139 A1140

L1141 F1142 N1143 L1156 T1157 T1158 K1159 E1160 T1167 I1168 D1169 P1170 R1171 W1172 L1173 V1174 E1175 F1176 A1177 P1178 A1179 F1180 F1181 K1182 V1183 S1184 D1185 P1186 THR LYS LEU SER LYS GLN LYS LYS GLN GLN GLN ARG LEU GLU PRO LEU TYR ASN ARG TYR GLU GLU GLU PRO ASN ASN ALA TRP ARG ILE SER ARG ALA

PHE ARG ARG ARG

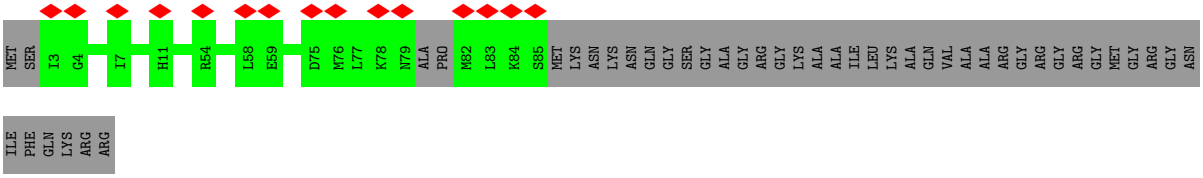
- Molecule 26: Small nuclear ribonucleoprotein Sm D3

Chain a: 5% 64% 36%



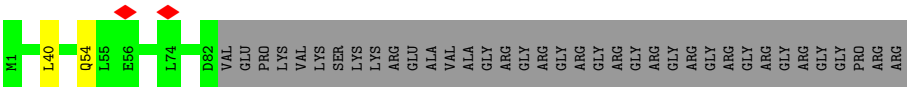
• Molecule 26: Small nuclear ribonucleoprotein Sm D3

Chain h:



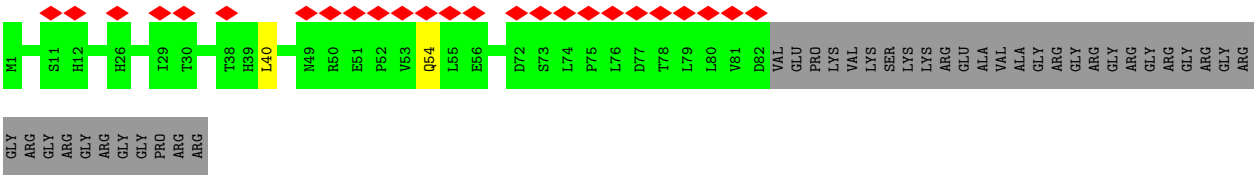
• Molecule 27: Small nuclear ribonucleoprotein Sm D1

Chain c:



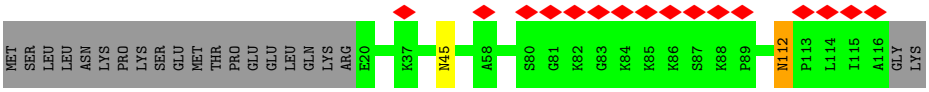
• Molecule 27: Small nuclear ribonucleoprotein Sm D1

Chain j:



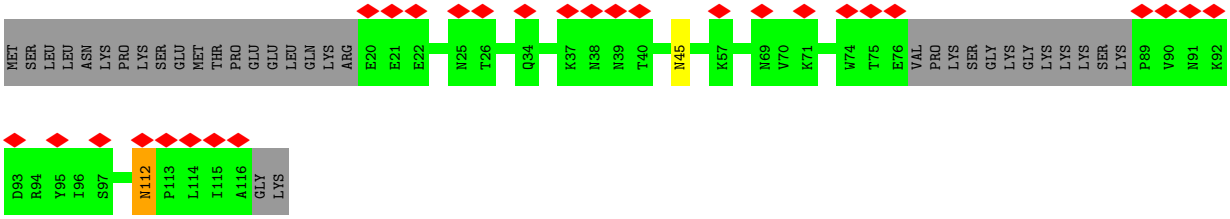
• Molecule 28: Small nuclear ribonucleoprotein Sm D2

Chain d:



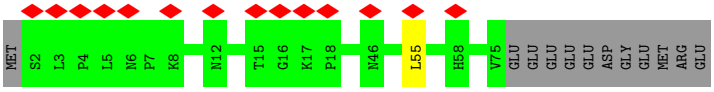
• Molecule 28: Small nuclear ribonucleoprotein Sm D2

Chain k:



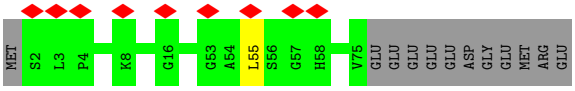
• Molecule 29: Small nuclear ribonucleoprotein F

Chain f:



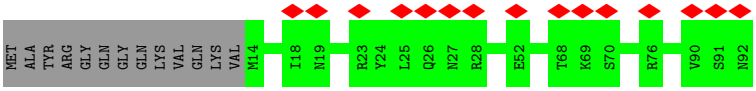
• Molecule 29: Small nuclear ribonucleoprotein F

Chain m:



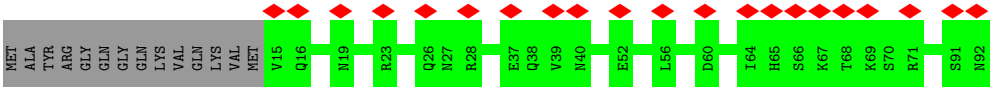
• Molecule 30: Small nuclear ribonucleoprotein E

Chain e:



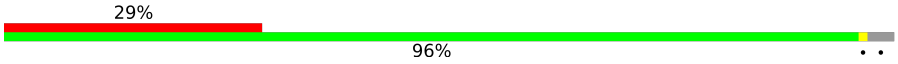
• Molecule 30: Small nuclear ribonucleoprotein E

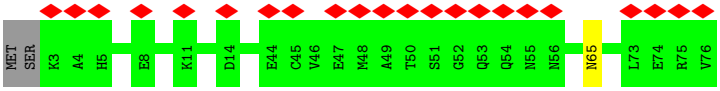
Chain l:



• Molecule 31: Small nuclear ribonucleoprotein G

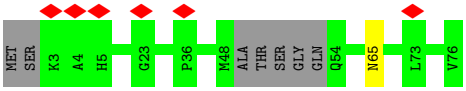
Chain g:





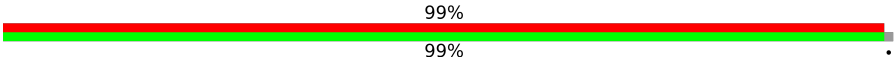
• Molecule 31: Small nuclear ribonucleoprotein G

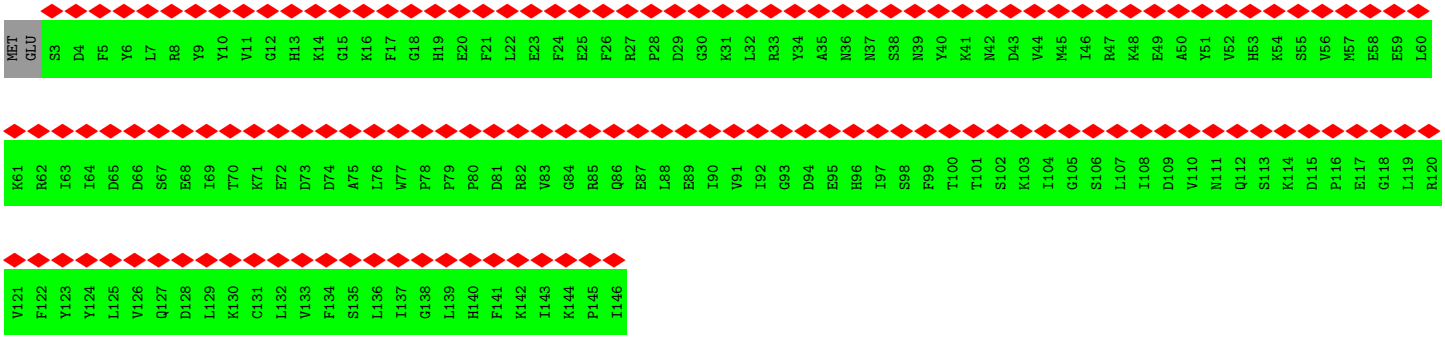
Chain n:



• Molecule 32: Protein mago nashi homolog

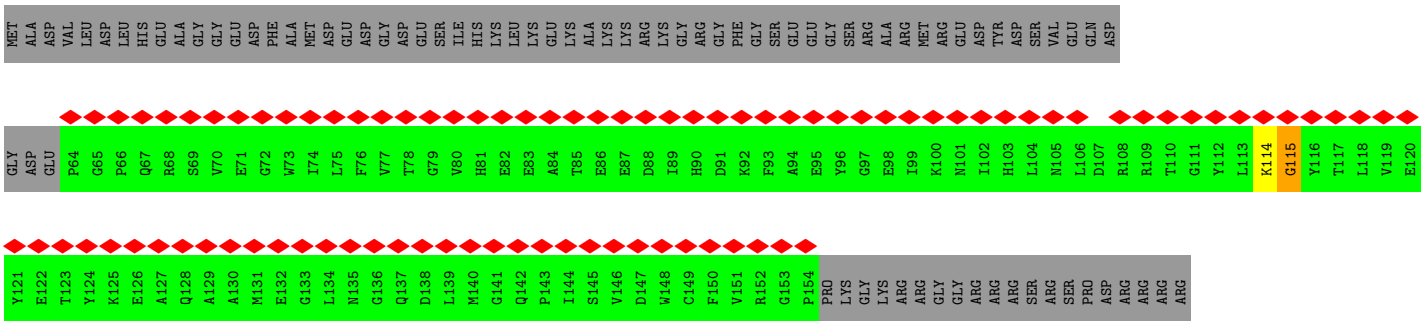
Chain v:





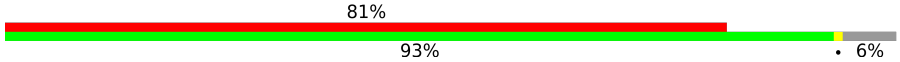
- Molecule 33: RNA-binding protein 8A

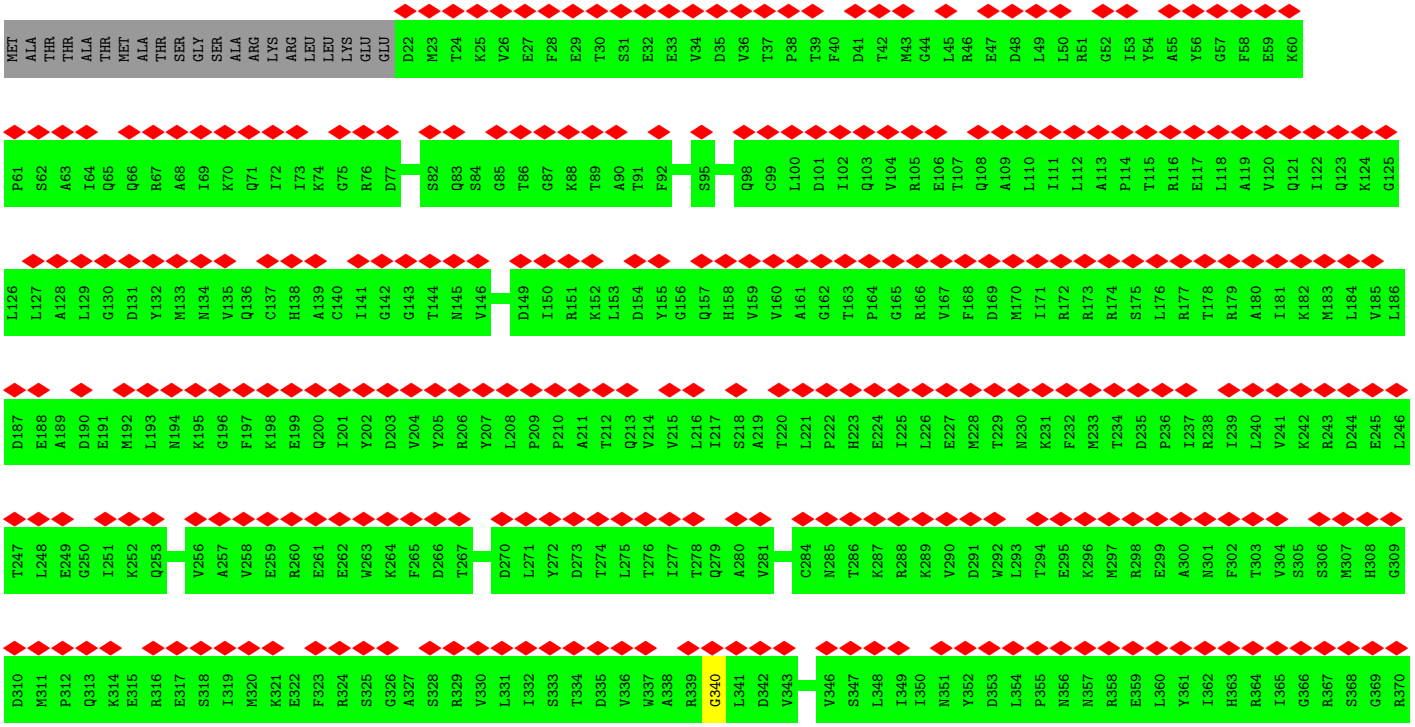
Chain w:

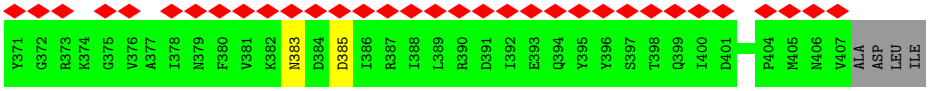


- Molecule 34: Eukaryotic initiation factor 4A-III

Chain u:

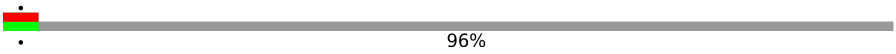


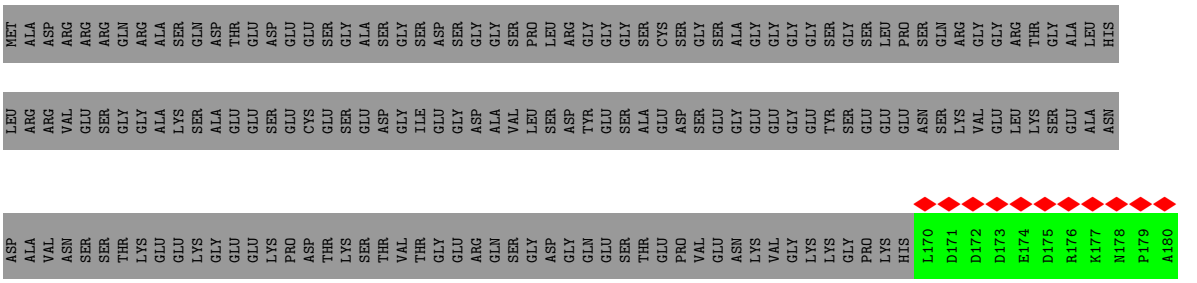




• Molecule 35: Protein CASC3

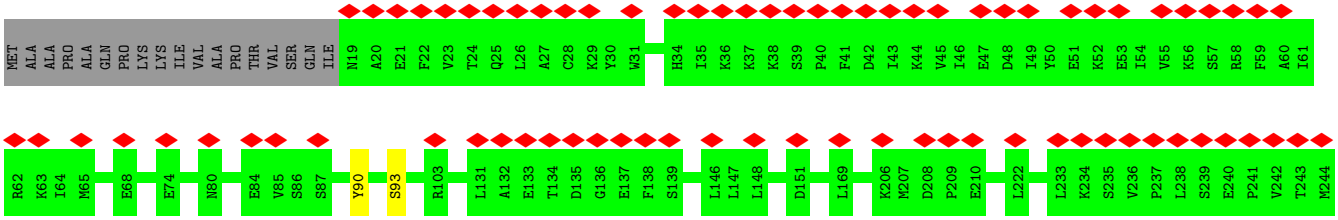
Chain x:

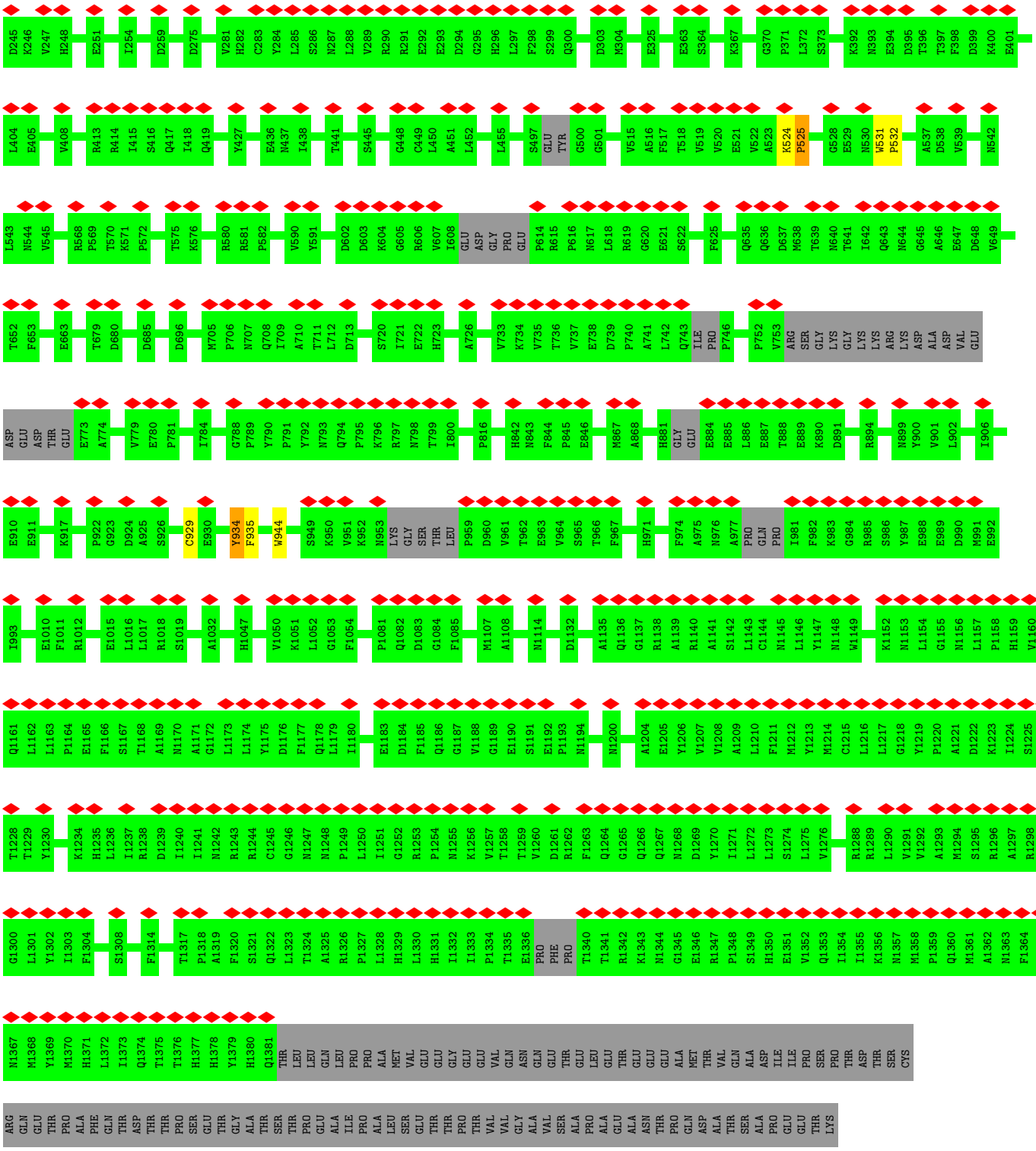




• Molecule 36: RNA helicase aquarius

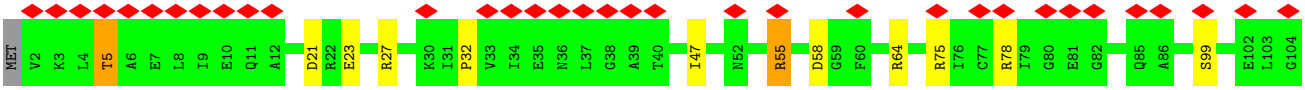
Chain Q:

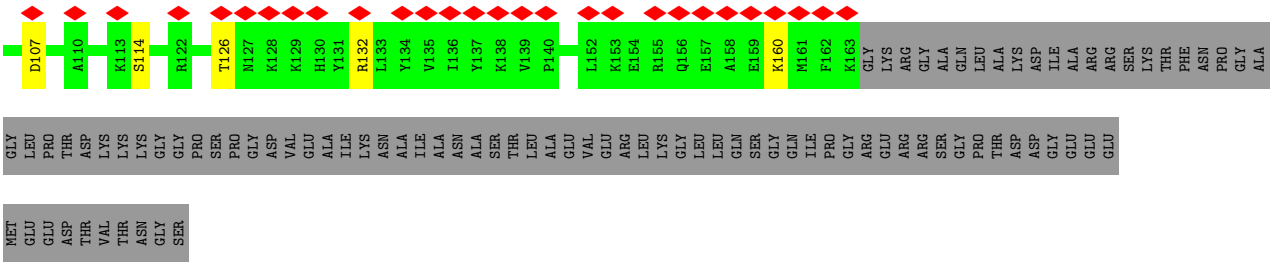




- Molecule 37: U2 small nuclear ribonucleoprotein A'

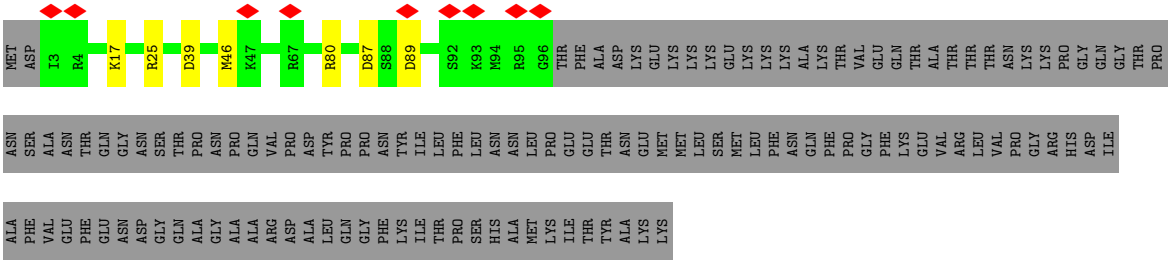
Chain o:





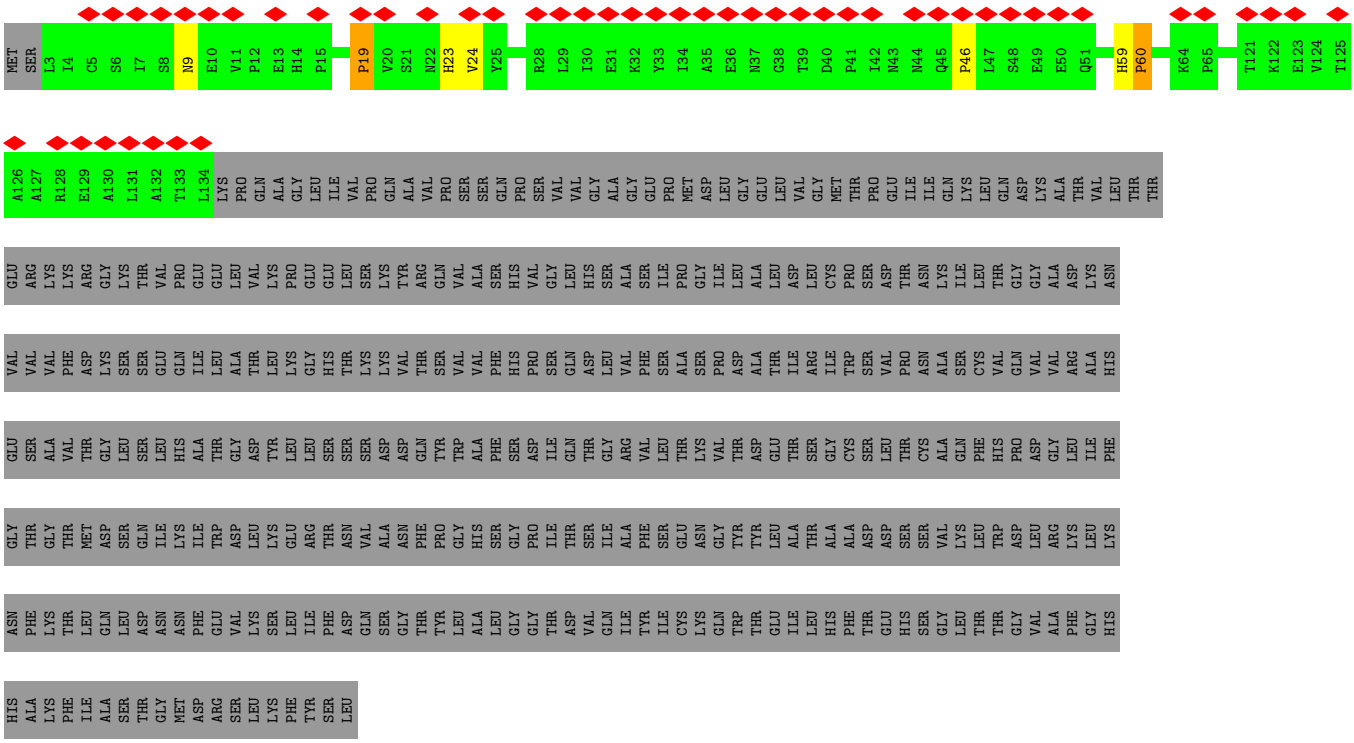
• Molecule 38: U2 small nuclear ribonucleoprotein B”

Chain p:



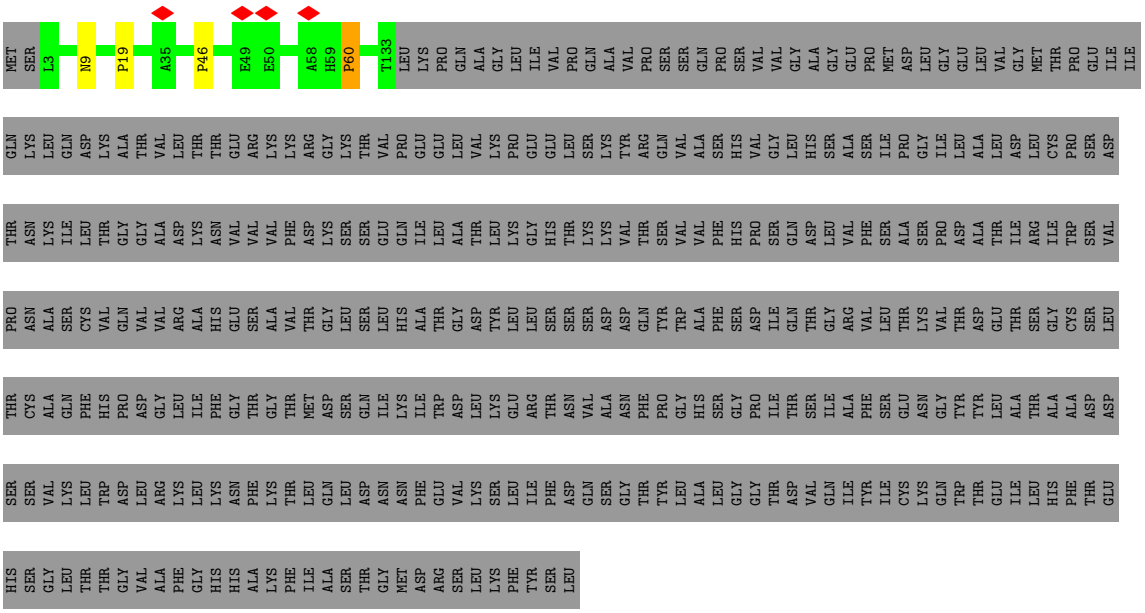
• Molecule 39: Pre-mRNA-processing factor 19

Chain q:



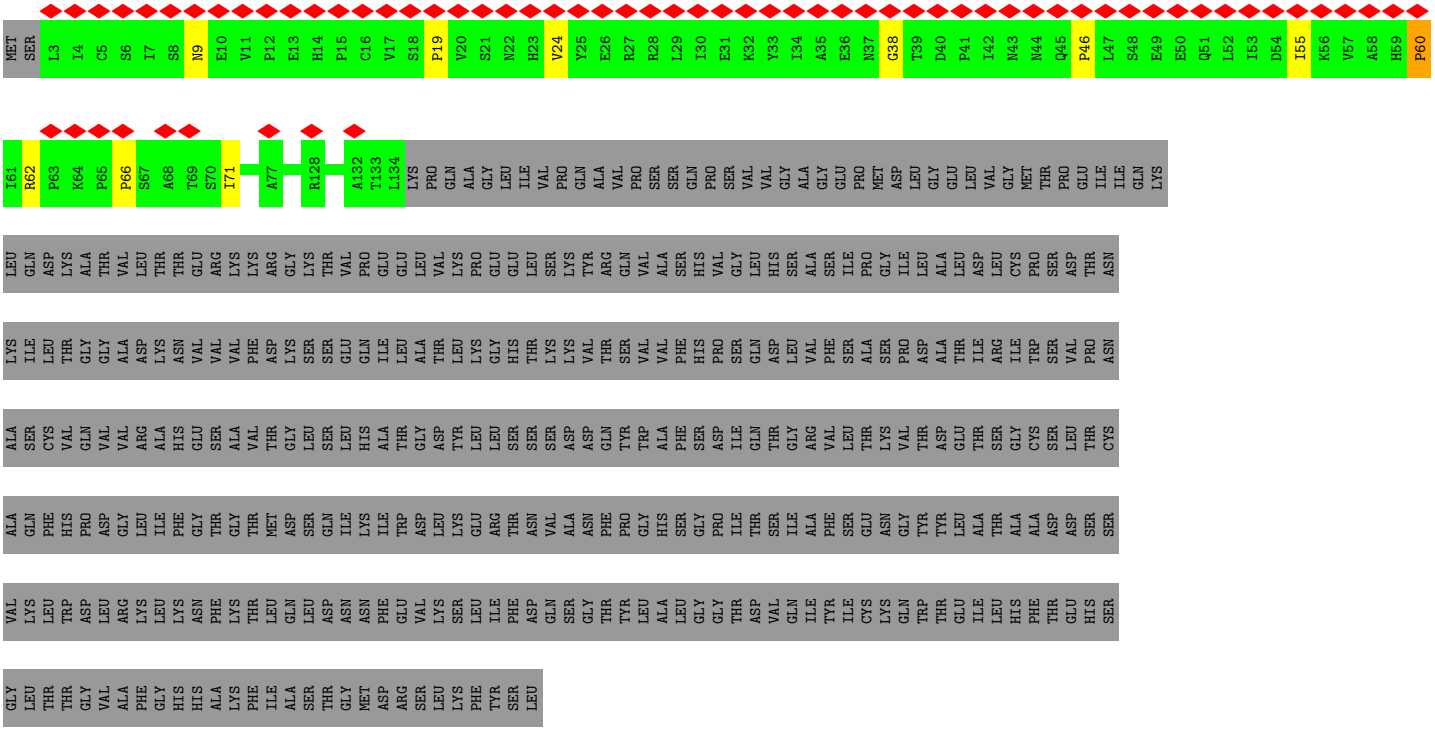
• Molecule 39: Pre-mRNA-processing factor 19

Chain r:



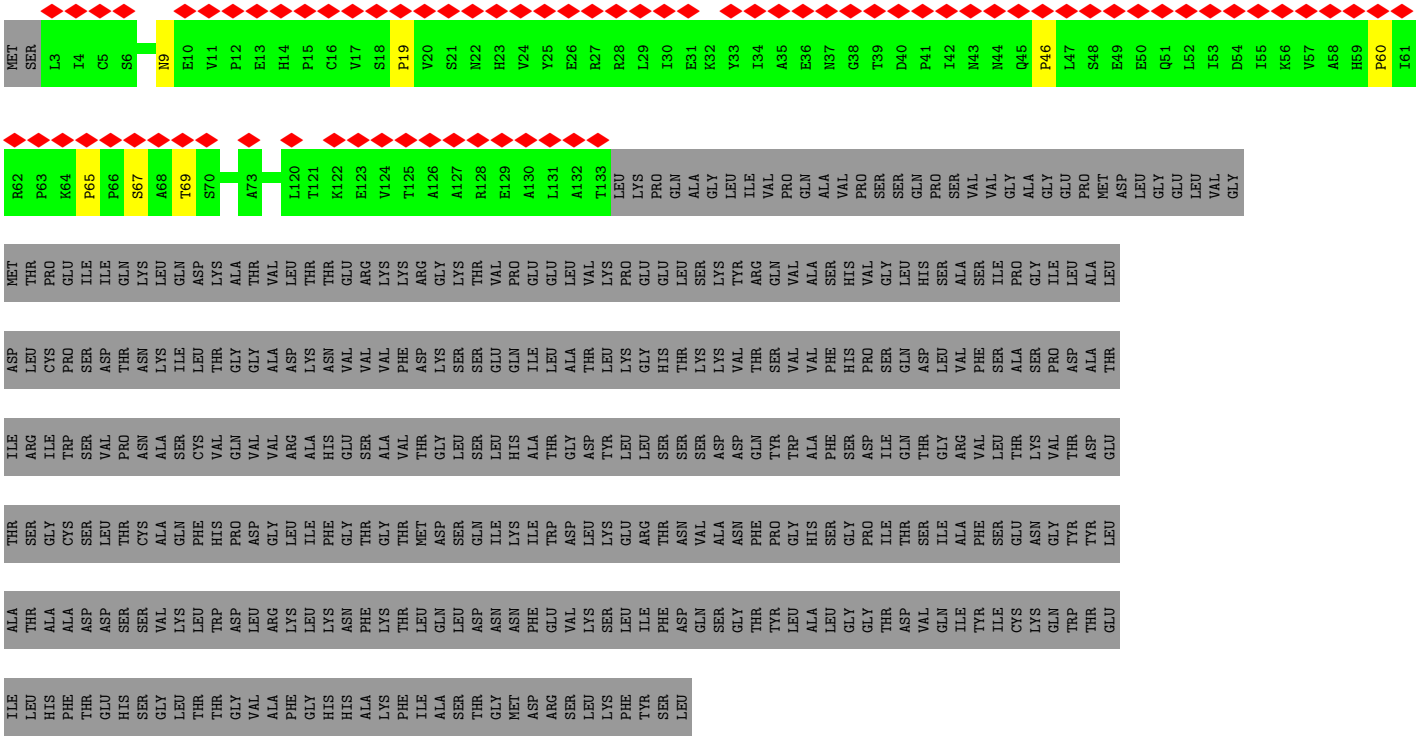
• Molecule 39: Pre-mRNA-processing factor 19

Chain s:



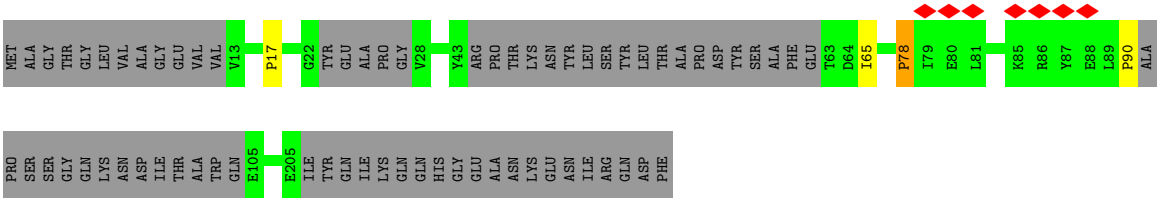
• Molecule 39: Pre-mRNA-processing factor 19

Chain t:



- Molecule 40: Pre-mRNA-splicing factor SPF27

Chain K:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	192274	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	50	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor
Maximum map value	3.470	Depositor
Minimum map value	-1.736	Depositor
Average map value	0.012	Depositor
Map value standard deviation	0.092	Depositor
Recommended contour level	0.38	Depositor
Map size (Å)	535.2, 535.2, 535.2	wwPDB
Map dimensions	400, 400, 400	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	1.338, 1.338, 1.338	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: MG, ATP, ZN, IHP, GTP, SEP

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\|>5$	RMSZ	$\#\|Z\|>5$
1	A	0.81	21/16897 (0.1%)	0.87	44/22917 (0.2%)
2	C	0.61	1/6950 (0.0%)	0.76	8/9443 (0.1%)
3	E	0.57	0/2392	0.71	2/3242 (0.1%)
4	4	1.34	0/307	1.53	7/476 (1.5%)
5	G	0.62	3/1674 (0.2%)	1.27	24/2594 (0.9%)
6	J	0.60	2/3856 (0.1%)	0.64	8/5234 (0.2%)
7	L	0.48	0/3046	0.63	2/4115 (0.0%)
8	M	0.51	0/1119	0.64	0/1497
9	N	0.82	3/1210 (0.2%)	0.81	0/1622
10	O	0.59	1/2344 (0.0%)	0.69	3/3163 (0.1%)
11	P	0.65	0/967	0.92	3/1285 (0.2%)
12	R	0.69	1/2262 (0.0%)	0.91	9/3031 (0.3%)
13	S	0.52	0/1268	0.72	2/1714 (0.1%)
14	T	0.99	5/2519 (0.2%)	1.03	16/3433 (0.5%)
15	U	0.49	0/424	0.59	0/582
16	V	0.34	0/2642	0.56	1/3602 (0.0%)
17	W	0.47	0/4237	0.73	4/5723 (0.1%)
18	B	0.61	1/1970 (0.1%)	0.84	4/3060 (0.1%)
19	F	0.49	0/2323	0.78	2/3619 (0.1%)
20	H	0.95	26/3305 (0.8%)	1.63	112/5130 (2.2%)
21	b	0.55	0/797	0.81	0/1062
21	i	0.53	0/700	0.82	0/933
22	X	0.83	0/621	1.39	7/822 (0.9%)
23	I	0.44	0/3871	0.87	9/5283 (0.2%)
24	y	0.29	0/389	0.64	0/540
25	Y	0.55	1/3436 (0.0%)	0.77	4/4774 (0.1%)
26	a	0.47	0/646	0.68	0/867
26	h	0.47	0/639	0.68	0/857
27	c	0.55	0/657	0.77	0/888
27	j	0.56	0/657	0.77	0/888
28	d	0.69	0/786	0.86	0/1053
28	k	0.69	0/696	0.86	0/935

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
29	f	0.82	0/588	0.83	0/795
29	m	0.81	0/588	0.83	0/795
30	e	0.62	0/660	0.83	0/886
30	l	0.61	0/652	0.84	0/876
31	g	0.52	0/584	0.78	0/779
31	n	0.53	0/544	0.80	0/725
32	v	0.30	0/710	0.65	0/987
33	w	0.31	0/444	0.78	2/614 (0.3%)
34	u	0.32	0/1906	0.69	0/2653
35	x	0.34	0/123	0.70	0/170
36	Q	0.35	4/6565 (0.1%)	0.52	5/9143 (0.1%)
37	o	0.59	0/1299	1.63	17/1761 (1.0%)
38	p	0.56	0/774	1.35	6/1035 (0.6%)
39	q	0.35	0/658	0.63	3/919 (0.3%)
39	r	0.32	0/653	0.59	3/912 (0.3%)
39	s	0.34	0/658	0.66	3/919 (0.3%)
39	t	0.35	0/653	0.59	3/912 (0.3%)
40	K	0.37	0/768	0.48	2/1067 (0.2%)
All	All	0.63	69/94434 (0.1%)	0.86	315/130332 (0.2%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	15
2	C	0	5
6	J	0	3
9	N	0	3
10	O	0	1
11	P	0	3
12	R	0	6
17	W	0	1
23	I	0	5
28	d	0	1
28	k	0	1
36	Q	0	1
All	All	0	45

All (69) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	345	PRO	N-CA	12.23	1.68	1.47
36	Q	525	PRO	N-CA	11.95	1.67	1.47
2	C	92	PRO	N-CA	11.82	1.67	1.47
25	Y	426	LEU	C-N	8.48	1.50	1.34
5	G	21	A	O3'-P	-8.44	1.51	1.61
36	Q	531	TRP	C-N	8.33	1.50	1.34
1	A	942	PRO	C-O	-8.10	1.07	1.23
1	A	1319	PRO	N-CD	-7.83	1.36	1.47
14	T	191	HIS	C-N	7.77	1.49	1.34
20	H	77	C	C1'-N1	7.43	1.59	1.48
20	H	142	C	C1'-N1	7.42	1.59	1.48
1	A	1319	PRO	C-O	-7.31	1.08	1.23
5	G	98	U	O3'-P	-7.23	1.52	1.61
18	B	103	G	C1'-N9	-7.03	1.37	1.46
20	H	55	U	C1'-N1	6.99	1.59	1.48
20	H	92	U	C1'-N1	6.96	1.59	1.48
20	H	54	U	C1'-N1	6.95	1.59	1.48
9	N	101	CYS	CB-SG	-6.92	1.70	1.82
20	H	72	U	C1'-N1	6.92	1.59	1.48
20	H	89	U	C1'-N1	6.91	1.59	1.48
20	H	74	U	C1'-N1	6.90	1.59	1.48
20	H	69	U	C1'-N1	6.90	1.59	1.48
20	H	60	U	C1'-N1	6.85	1.59	1.48
20	H	91	U	C1'-N1	6.82	1.58	1.48
20	H	58	U	C1'-N1	6.82	1.58	1.48
20	H	150	U	C1'-N1	6.81	1.58	1.48
20	H	182	U	C1'-N1	6.75	1.58	1.48
20	H	151	C	C1'-N1	6.46	1.58	1.48
20	H	141	C	C1'-N1	6.42	1.58	1.48
20	H	73	C	C1'-N1	6.40	1.58	1.48
20	H	97	G	C1'-N9	-6.39	1.38	1.46
1	A	942	PRO	N-CD	-6.37	1.39	1.47
20	H	184	C	C1'-N1	6.37	1.58	1.48
20	H	71	C	C1'-N1	6.34	1.58	1.48
20	H	148	C	C1'-N1	6.30	1.58	1.48
20	H	67	C	C1'-N1	6.29	1.58	1.48
1	A	1508	GLY	C-O	6.28	1.33	1.23
20	H	78	C	C1'-N1	6.28	1.58	1.48
20	H	70	C	C1'-N1	6.27	1.58	1.48
20	H	84	C	C1'-N1	6.25	1.58	1.48
1	A	1472	THR	C-O	-6.12	1.11	1.23
1	A	1470	TYR	C-O	-5.98	1.11	1.23
1	A	776	LEU	C-O	-5.97	1.12	1.23

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	G	3	A	O3'-P	-5.89	1.54	1.61
14	T	404	SER	CA-CB	-5.88	1.44	1.52
12	R	219	PRO	C-O	-5.83	1.11	1.23
1	A	187	PRO	N-CD	-5.82	1.39	1.47
36	Q	524	LYS	C-N	5.81	1.45	1.34
36	Q	934	TYR	C-O	5.75	1.34	1.23
14	T	327	SER	CA-CB	-5.71	1.44	1.52
10	O	27	CYS	CB-SG	-5.62	1.72	1.81
6	J	243	SER	CA-CB	-5.58	1.44	1.52
14	T	343	PRO	C-O	-5.58	1.12	1.23
1	A	934	ARG	C-O	-5.49	1.12	1.23
1	A	1318	THR	C-O	-5.49	1.12	1.23
1	A	1096	HIS	C-O	-5.49	1.12	1.23
1	A	1093	ASP	C-O	-5.46	1.12	1.23
1	A	771	VAL	C-O	-5.45	1.13	1.23
1	A	1089	CYS	C-O	-5.39	1.13	1.23
9	N	137	CYS	CB-SG	-5.39	1.73	1.81
1	A	167	PRO	N-CD	-5.34	1.40	1.47
6	J	239	ARG	C-O	-5.30	1.13	1.23
9	N	119	CYS	CB-SG	-5.29	1.73	1.81
1	A	344	ASP	C-N	5.21	1.44	1.34
20	H	110	A	C1'-N9	-5.13	1.39	1.46
1	A	1438	VAL	CB-CG1	-5.12	1.42	1.52
1	A	773	LYS	C-O	-5.03	1.13	1.23
14	T	461	SER	CA-CB	-5.02	1.45	1.52
1	A	1090	ARG	C-O	-5.00	1.13	1.23

All (315) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1515	TRP	N-CA-CB	15.62	138.72	110.60
37	o	55	ARG	NE-CZ-NH1	15.18	127.89	120.30
1	A	772	CYS	CA-CB-SG	-13.83	89.10	114.00
37	o	55	ARG	CD-NE-CZ	13.69	142.76	123.60
1	A	856	LEU	C-N-CA	13.32	155.00	121.70
5	G	120	G	O4'-C1'-N9	-13.02	97.78	108.20
11	P	7	PRO	N-CA-CB	-12.82	87.91	103.30
13	S	72	ARG	CB-CA-C	-12.80	84.80	110.40
1	A	1470	TYR	CB-CA-C	-12.56	85.28	110.40
1	A	187	PRO	N-CD-CG	-12.34	84.69	103.20
20	H	167	U	C5-C4-O4	11.90	133.04	125.90
2	C	85	ASP	CB-CA-C	-11.81	86.77	110.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
37	o	55	ARG	NE-CZ-NH2	-11.51	114.55	120.30
5	G	130	G	N9-C1'-C2'	-11.45	99.11	114.00
37	o	75	ARG	NE-CZ-NH1	-11.31	114.65	120.30
1	A	1261	ASN	CB-CA-C	-11.14	88.12	110.40
14	T	401	PRO	CA-N-CD	-11.05	96.03	111.50
4	4	-12	G	N9-C1'-C2'	-11.04	99.65	114.00
1	A	186	GLU	CB-CA-C	-10.72	88.95	110.40
14	T	343	PRO	CA-N-CD	-10.27	97.12	111.50
20	H	164	C	N1-C2-O2	-10.09	112.85	118.90
1	A	1264	ASN	CB-CA-C	-9.98	90.43	110.40
22	X	115	TYR	CB-CA-C	9.94	130.29	110.40
20	H	31	G	N9-C1'-C2'	-9.82	101.20	112.00
1	A	187	PRO	N-CA-CB	-9.60	91.78	103.30
5	G	130	G	C4'-C3'-O3'	9.46	131.93	113.00
1	A	1318	THR	CA-CB-OG1	-9.46	89.14	109.00
1	A	934	ARG	CG-CD-NE	-9.39	92.08	111.80
5	G	1	G	N9-C1'-C2'	-9.24	101.84	112.00
1	A	167	PRO	CA-N-CD	-9.23	98.58	111.50
22	X	96	GLN	CB-CA-C	-9.20	92.00	110.40
38	p	25	ARG	NE-CZ-NH1	9.14	124.87	120.30
5	G	122	U	O5'-P-OP2	-9.05	97.56	105.70
20	H	162	U	N3-C2-O2	-8.96	115.93	122.20
19	F	56	A	O5'-P-OP1	-8.96	97.64	105.70
22	X	107	GLU	CB-CA-C	-8.86	92.67	110.40
38	p	80	ARG	CD-NE-CZ	8.86	136.00	123.60
2	C	114	TYR	CB-CA-C	-8.81	92.78	110.40
14	T	339	GLN	CB-CA-C	-8.66	93.08	110.40
1	A	1090	ARG	CB-CG-CD	-8.61	89.21	111.60
4	4	-6	C	C2-N1-C1'	8.50	128.15	118.80
18	B	20	G	N9-C1'-C2'	8.50	125.05	114.00
1	A	942	PRO	N-CA-C	-8.41	90.24	112.10
1	A	1754	TYR	CB-CA-C	-8.38	93.64	110.40
14	T	343	PRO	CB-CA-C	8.35	132.87	112.00
1	A	1507	SER	O-C-N	-8.33	109.04	123.20
20	H	164	C	C5'-C4'-O4'	-8.32	99.12	109.10
5	G	3	A	N9-C1'-C2'	-8.31	102.86	112.00
1	A	1459	ARG	CB-CG-CD	-8.25	90.15	111.60
20	H	169	C	P-O3'-C3'	8.21	129.55	119.70
23	I	83	LYS	O-C-N	-8.16	109.64	122.70
23	I	84	HIS	O-C-N	-8.13	109.69	122.70
23	I	310	LYS	O-C-N	-8.12	109.71	122.70
23	I	311	MET	O-C-N	-8.10	109.74	122.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	345	PRO	CA-N-CD	-8.09	100.17	111.50
20	H	166	G	O4'-C1'-N9	8.08	114.66	108.20
23	I	309	ALA	O-C-N	-8.06	109.81	122.70
1	A	1091	TYR	CB-CG-CD2	-8.01	116.19	121.00
20	H	167	U	N3-C4-O4	-8.00	113.80	119.40
1	A	1090	ARG	CG-CD-NE	-7.96	95.07	111.80
12	R	310	ARG	CB-CA-C	-7.96	94.48	110.40
36	Q	935	PHE	CB-CA-C	7.95	126.30	110.40
10	O	193	LEU	CA-CB-CG	7.87	133.40	115.30
5	G	1	G	C1'-C2'-O2'	-7.86	87.01	110.60
2	C	68	THR	CB-CA-C	-7.82	90.48	111.60
5	G	125	C	P-O3'-C3'	7.75	129.00	119.70
37	o	107	ASP	CB-CG-OD1	7.66	125.19	118.30
1	A	1472	THR	CB-CA-C	-7.65	90.95	111.60
20	H	167	U	N1-C2-O2	7.57	128.10	122.80
2	C	92	PRO	CA-N-CD	-7.54	100.95	111.50
20	H	164	C	P-O3'-C3'	7.50	128.70	119.70
5	G	120	G	P-O3'-C3'	7.44	128.63	119.70
37	o	27	ARG	NE-CZ-NH2	-7.41	116.59	120.30
20	H	164	C	N3-C2-O2	7.41	127.09	121.90
1	A	663	ARG	CG-CD-NE	-7.39	96.27	111.80
20	H	69	U	OP2-P-O3'	7.28	121.22	105.20
20	H	114	A	OP2-P-O3'	7.25	121.15	105.20
20	H	73	C	OP2-P-O3'	7.24	121.14	105.20
20	H	141	C	OP2-P-O3'	7.23	121.10	105.20
20	H	183	G	OP2-P-O3'	7.23	121.10	105.20
20	H	77	C	OP2-P-O3'	7.22	121.09	105.20
20	H	148	C	OP2-P-O3'	7.22	121.09	105.20
20	H	82	G	OP2-P-O3'	7.21	121.07	105.20
20	H	92	U	OP2-P-O3'	7.21	121.07	105.20
20	H	88	A	OP2-P-O3'	7.21	121.07	105.20
20	H	113	G	OP2-P-O3'	7.21	121.07	105.20
20	H	70	C	OP2-P-O3'	7.21	121.07	105.20
20	H	149	A	OP2-P-O3'	7.21	121.06	105.20
20	H	55	U	OP2-P-O3'	7.21	121.06	105.20
20	H	182	U	OP2-P-O3'	7.21	121.05	105.20
20	H	81	G	OP2-P-O3'	7.21	121.05	105.20
20	H	54	U	OP2-P-O3'	7.20	121.04	105.20
20	H	59	A	OP2-P-O3'	7.20	121.04	105.20
20	H	91	U	OP2-P-O3'	7.20	121.04	105.20
20	H	72	U	OP2-P-O3'	7.20	121.03	105.20
20	H	68	G	OP2-P-O3'	7.20	121.03	105.20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
20	H	71	C	OP2-P-O3'	7.20	121.03	105.20
20	H	93	A	OP2-P-O3'	7.19	121.03	105.20
20	H	78	C	OP2-P-O3'	7.19	121.02	105.20
20	H	89	U	OP2-P-O3'	7.19	121.01	105.20
20	H	180	G	OP2-P-O3'	7.19	121.01	105.20
20	H	58	U	OP2-P-O3'	7.18	121.00	105.20
20	H	90	A	OP2-P-O3'	7.18	121.00	105.20
20	H	74	U	OP2-P-O3'	7.18	121.00	105.20
20	H	79	G	OP2-P-O3'	7.18	120.99	105.20
20	H	57	A	OP2-P-O3'	7.18	120.99	105.20
20	H	56	A	OP2-P-O3'	7.17	120.98	105.20
20	H	67	C	OP2-P-O3'	7.17	120.98	105.20
20	H	181	G	OP2-P-O3'	7.17	120.98	105.20
20	H	84	C	OP2-P-O3'	7.17	120.97	105.20
20	H	150	U	OP2-P-O3'	7.17	120.97	105.20
20	H	80	A	OP2-P-O3'	7.15	120.94	105.20
14	T	344	GLN	N-CA-CB	7.14	123.46	110.60
20	H	83	A	OP2-P-O3'	7.14	120.90	105.20
39	s	60	PRO	N-CA-CB	7.11	111.83	103.30
20	H	168	A	P-O5'-C5'	-7.10	109.54	120.90
20	H	167	U	N3-C2-O2	-7.06	117.26	122.20
14	T	187	LYS	C-N-CD	-7.02	105.15	120.60
5	G	127	G	C4'-C3'-O3'	6.96	126.92	113.00
5	G	121	C	C4'-C3'-O3'	6.95	126.89	113.00
1	A	775	ASN	CA-CB-CG	-6.92	98.17	113.40
1	A	1461	ASP	CB-CA-C	-6.90	96.61	110.40
5	G	122	U	N1-C1'-C2'	-6.86	104.45	112.00
20	H	183	G	O3'-P-O5'	-6.84	91.01	104.00
3	E	193	THR	CA-CB-OG1	-6.83	94.65	109.00
20	H	82	G	O3'-P-O5'	-6.82	91.03	104.00
20	H	77	C	O3'-P-O5'	-6.81	91.05	104.00
20	H	59	A	O3'-P-O5'	-6.81	91.06	104.00
20	H	113	G	O3'-P-O5'	-6.81	91.06	104.00
20	H	91	U	O3'-P-O5'	-6.81	91.07	104.00
20	H	78	C	O3'-P-O5'	-6.80	91.07	104.00
20	H	141	C	O3'-P-O5'	-6.80	91.08	104.00
20	H	150	U	O3'-P-O5'	-6.80	91.09	104.00
20	H	180	G	O3'-P-O5'	-6.79	91.09	104.00
20	H	73	C	O3'-P-O5'	-6.79	91.09	104.00
20	H	74	U	O3'-P-O5'	-6.79	91.09	104.00
20	H	56	A	O3'-P-O5'	-6.79	91.10	104.00
20	H	148	C	O3'-P-O5'	-6.79	91.10	104.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
20	H	68	G	O3'-P-O5'	-6.79	91.11	104.00
20	H	182	U	O3'-P-O5'	-6.78	91.11	104.00
20	H	57	A	O3'-P-O5'	-6.78	91.12	104.00
20	H	89	U	O3'-P-O5'	-6.78	91.12	104.00
20	H	54	U	O3'-P-O5'	-6.78	91.12	104.00
18	B	104	C	C2'-C3'-O3'	-6.78	94.59	109.50
20	H	93	A	O3'-P-O5'	-6.78	91.12	104.00
20	H	181	G	O3'-P-O5'	-6.78	91.12	104.00
20	H	79	G	O3'-P-O5'	-6.77	91.13	104.00
20	H	55	U	O3'-P-O5'	-6.77	91.14	104.00
20	H	72	U	O3'-P-O5'	-6.77	91.14	104.00
20	H	80	A	O3'-P-O5'	-6.77	91.14	104.00
20	H	90	A	O3'-P-O5'	-6.77	91.14	104.00
1	A	1496	PRO	CB-CA-C	-6.77	95.08	112.00
20	H	149	A	O3'-P-O5'	-6.76	91.15	104.00
20	H	58	U	O3'-P-O5'	-6.76	91.15	104.00
20	H	88	A	O3'-P-O5'	-6.76	91.16	104.00
1	A	1265	THR	CA-CB-OG1	-6.76	94.81	109.00
20	H	83	A	O3'-P-O5'	-6.76	91.16	104.00
20	H	92	U	O3'-P-O5'	-6.76	91.16	104.00
20	H	69	U	O3'-P-O5'	-6.75	91.17	104.00
20	H	81	G	O3'-P-O5'	-6.75	91.18	104.00
20	H	84	C	O3'-P-O5'	-6.75	91.18	104.00
37	o	23	GLU	OE1-CD-OE2	-6.75	115.20	123.30
14	T	399	LYS	CB-CA-C	-6.74	96.92	110.40
20	H	71	C	O3'-P-O5'	-6.74	91.19	104.00
20	H	114	A	O3'-P-O5'	-6.74	91.20	104.00
20	H	155	C	P-O3'-C3'	6.74	127.78	119.70
20	H	70	C	O3'-P-O5'	-6.73	91.22	104.00
20	H	67	C	O3'-P-O5'	-6.72	91.23	104.00
1	A	775	ASN	CB-CA-C	-6.72	96.96	110.40
20	H	166	G	C8-N9-C4	-6.70	103.72	106.40
39	t	60	PRO	N-CA-CB	6.69	111.33	103.30
16	V	467	LEU	CA-CB-CG	6.64	130.58	115.30
38	p	89	ASP	CB-CG-OD1	6.62	124.26	118.30
5	G	127	G	N9-C1'-C2'	-6.59	104.75	112.00
39	q	46	PRO	N-CA-CB	6.57	111.19	103.30
39	s	46	PRO	N-CA-CB	6.57	111.19	103.30
39	r	46	PRO	N-CA-CB	6.49	111.09	103.30
1	A	167	PRO	N-CD-CG	-6.48	93.48	103.20
2	C	67	PRO	CB-CA-C	-6.47	95.84	112.00
40	K	90	PRO	N-CA-CB	6.46	111.05	103.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
39	q	60	PRO	N-CA-CB	6.45	111.04	103.30
33	w	114	LYS	N-CA-C	-6.42	93.65	111.00
7	L	159	LEU	CA-CB-CG	6.40	130.02	115.30
1	A	1094	ARG	CB-CA-C	-6.37	97.66	110.40
5	G	84	U	C2'-C3'-O3'	-6.36	95.51	109.50
36	Q	929	CYS	O-C-N	-6.36	112.53	122.70
37	o	5	THR	N-CA-CB	-6.32	98.29	110.30
5	G	122	U	C5'-C4'-C3'	-6.31	105.91	116.00
20	H	165	A	O4'-C1'-N9	-6.30	103.16	108.20
5	G	124	G	C2'-C3'-O3'	6.30	123.78	113.70
1	A	1261	ASN	CA-CB-CG	-6.29	99.56	113.40
40	K	78	PRO	N-CA-CB	6.28	110.84	103.30
2	C	308	CYS	CA-CB-SG	6.26	125.28	114.00
7	L	9	GLY	C-N-CA	6.24	137.31	121.70
20	H	166	G	N9-C4-C5	6.24	107.90	105.40
22	X	108	GLU	CB-CA-C	-6.24	97.92	110.40
22	X	64	ARG	CB-CA-C	-6.19	98.01	110.40
14	T	401	PRO	CB-CA-C	6.17	127.42	112.00
5	G	21	A	O3'-P-O5'	6.14	115.67	104.00
4	4	-6	C	C6-N1-C1'	-6.14	113.44	120.80
1	A	1523	ARG	N-CA-CB	-6.12	99.59	110.60
12	R	320	HIS	CB-CA-C	6.11	122.61	110.40
20	H	162	U	N1-C2-O2	6.10	127.07	122.80
12	R	310	ARG	CB-CG-CD	-6.08	95.79	111.60
37	o	132	ARG	CD-NE-CZ	6.03	132.04	123.60
12	R	223	PRO	CB-CA-C	-6.02	96.96	112.00
10	O	193	LEU	CB-CG-CD2	-6.01	100.78	111.00
20	H	166	G	N3-C4-C5	-5.99	125.60	128.60
1	A	1091	TYR	N-CA-CB	-5.99	99.83	110.60
6	J	220	LEU	CA-CB-CG	5.98	129.05	115.30
37	o	78	ARG	NE-CZ-NH2	-5.97	117.31	120.30
22	X	103	GLN	CB-CA-C	-5.97	98.47	110.40
1	A	663	ARG	CB-CA-C	-5.96	98.48	110.40
14	T	400	PHE	CB-CA-C	5.95	122.29	110.40
37	o	21	ASP	CB-CG-OD1	5.94	123.65	118.30
39	t	46	PRO	N-CA-CB	5.94	110.42	103.30
6	J	670	PRO	N-CA-CB	5.93	110.42	103.30
6	J	240	THR	CA-CB-OG1	-5.92	96.58	109.00
38	p	25	ARG	CD-NE-CZ	5.91	131.88	123.60
6	J	675	PRO	N-CA-CB	5.88	110.36	103.30
37	o	64	ARG	NE-CZ-NH1	5.88	123.24	120.30
6	J	637	PRO	N-CA-CB	5.87	110.35	103.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
23	I	458	LYS	O-C-N	5.87	132.10	122.70
20	H	168	A	C5'-C4'-C3'	-5.87	106.61	116.00
20	H	172	C	P-O3'-C3'	5.86	126.73	119.70
37	o	27	ARG	CB-CA-C	-5.85	98.69	110.40
6	J	604	PRO	N-CA-CB	5.85	110.32	103.30
23	I	459	ALA	O-C-N	5.84	132.04	122.70
18	B	26	A	P-O5'-C5'	-5.83	111.58	120.90
5	G	21	A	P-O3'-C3'	-5.82	112.72	119.70
20	H	164	C	C5-C4-N4	-5.80	116.14	120.20
39	q	19	PRO	N-CA-CB	5.80	110.26	103.30
14	T	342	GLU	CB-CA-C	5.79	121.97	110.40
11	P	6	ARG	CB-CA-C	-5.78	98.84	110.40
23	I	457	ARG	O-C-N	5.77	131.93	122.70
20	H	167	U	O3'-P-O5'	-5.74	93.09	104.00
6	J	566	PRO	N-CA-CB	5.73	110.17	103.30
39	r	19	PRO	N-CA-CB	5.72	110.16	103.30
12	R	217	LYS	CB-CA-C	-5.71	98.98	110.40
17	W	139	LEU	CA-CB-CG	5.69	128.38	115.30
12	R	214	ILE	CG1-CB-CG2	-5.68	98.90	111.40
1	A	187	PRO	N-CA-C	5.68	126.86	112.10
36	Q	934	TYR	C-N-CA	-5.68	107.51	121.70
37	o	75	ARG	NH1-CZ-NH2	5.67	125.64	119.40
39	t	19	PRO	N-CA-CB	5.65	110.08	103.30
20	H	156	U	P-O3'-C3'	-5.65	112.92	119.70
4	4	-12	G	C4'-C3'-O3'	5.62	124.24	113.00
20	H	166	G	C6-N1-C2	-5.59	121.74	125.10
5	G	2	U	N1-C1'-C2'	-5.57	105.87	112.00
14	T	324	HIS	CB-CA-C	-5.57	99.26	110.40
14	T	213	GLU	N-CA-CB	-5.56	100.59	110.60
37	o	58	ASP	N-CA-CB	-5.56	100.59	110.60
36	Q	93	SER	N-CA-CB	5.56	118.84	110.50
39	s	19	PRO	N-CA-CB	5.51	109.91	103.30
36	Q	944	TRP	C-N-CA	-5.48	108.00	121.70
17	W	532	LEU	CB-CG-CD1	-5.47	101.70	111.00
19	F	73	A	P-O3'-C3'	5.46	126.26	119.70
39	r	60	PRO	N-CA-CB	5.45	109.84	103.30
14	T	407	GLN	CB-CA-C	-5.44	99.52	110.40
18	B	20	G	O4'-C1'-N9	5.44	112.55	108.20
1	A	1091	TYR	CB-CG-CD1	5.42	124.25	121.00
22	X	86	GLU	CB-CG-CD	5.42	128.85	114.20
11	P	8	THR	CB-CA-C	5.38	126.14	111.60
20	H	157	G	O4'-C1'-N9	-5.38	103.89	108.20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
20	H	106	G	O5'-P-OP1	5.37	117.15	110.70
1	A	1556	ASP	CB-CA-C	-5.37	99.67	110.40
5	G	122	U	C2'-C3'-O3'	5.36	122.27	113.70
12	R	303	GLU	CB-CA-C	5.34	121.07	110.40
20	H	156	U	OP2-P-O3'	5.33	116.94	105.20
2	C	91	GLU	O-C-N	-5.33	110.97	121.10
12	R	275	LEU	CA-CB-CG	5.33	127.56	115.30
37	o	47	ILE	N-CA-CB	5.32	123.03	110.80
14	T	189	GLN	CB-CA-C	5.31	121.01	110.40
13	S	106	ASP	CB-CG-OD1	5.30	123.07	118.30
1	A	186	GLU	CB-CG-CD	-5.30	99.89	114.20
14	T	309	ASP	O-C-N	5.29	131.17	122.70
12	R	304	MET	CB-CG-SD	-5.29	96.53	112.40
1	A	1318	THR	N-CA-CB	-5.28	100.28	110.30
17	W	82	ASN	C-N-CD	-5.27	109.01	120.60
23	I	694	ILE	C-N-CA	5.26	134.85	121.70
38	p	39	ASP	CB-CG-OD2	-5.26	113.56	118.30
1	A	1753	LEU	N-CA-CB	-5.26	99.88	110.40
20	H	162	U	C2-N3-C4	-5.26	123.84	127.00
6	J	239	ARG	CG-CD-NE	-5.26	100.76	111.80
4	4	-6	C	N1-C2-O2	5.25	122.05	118.90
1	A	1301	ILE	CG1-CB-CG2	-5.25	99.85	111.40
4	4	-6	C	O4'-C1'-N1	5.24	112.39	108.20
14	T	309	ASP	CB-CG-OD1	5.24	123.02	118.30
3	E	192	ASN	C-N-CA	5.22	134.76	121.70
25	Y	685	ASP	C-N-CA	5.22	134.75	121.70
20	H	157	G	P-O5'-C5'	-5.21	112.56	120.90
17	W	243	VAL	CG1-CB-CG2	-5.21	102.56	110.90
10	O	28	LEU	CB-CG-CD2	-5.20	102.16	111.00
20	H	156	U	C4'-C3'-C2'	5.20	107.80	102.60
33	w	115	GLY	N-CA-C	5.19	126.07	113.10
1	A	1469	ASN	CB-CA-C	-5.18	100.04	110.40
5	G	96	U	N1-C1'-C2'	-5.18	106.30	112.00
1	A	187	PRO	CA-N-CD	5.17	118.94	111.70
5	G	2	U	C3'-C2'-C1'	5.17	105.63	101.50
20	H	170	C	N3-C4-C5	-5.16	119.83	121.90
38	p	39	ASP	CB-CG-OD1	5.15	122.94	118.30
20	H	160	A	P-O5'-C5'	-5.15	112.66	120.90
1	A	1303	LEU	CA-CB-CG	-5.14	103.47	115.30
5	G	121	C	C4'-C3'-C2'	5.11	107.71	102.60
20	H	176	G	OP1-P-OP2	-5.11	111.93	119.60
25	Y	686	GLU	C-N-CA	5.11	134.47	121.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	G	126	G	O5'-P-OP2	-5.10	101.11	105.70
20	H	160	A	C4'-C3'-C2'	-5.10	97.50	102.60
25	Y	689	GLU	N-CA-C	-5.08	97.27	111.00
25	Y	426	LEU	C-N-CD	-5.08	109.43	120.60
4	4	-6	C	N3-C2-O2	-5.08	118.35	121.90
1	A	167	PRO	N-CA-C	-5.07	98.91	112.10
2	C	359	LYS	CA-CB-CG	5.04	124.48	113.40
37	o	99	SER	N-CA-CB	-5.03	102.96	110.50
20	H	170	C	O4'-C1'-C2'	-5.02	100.78	105.80

There are no chirality outliers.

All (45) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1019	TYR	Peptide
1	A	1201	ARG	Peptide
1	A	1210	LYS	Peptide
1	A	1338	SER	Peptide
1	A	1507	SER	Mainchain
1	A	1686	ASP	Peptide
1	A	186	GLU	Mainchain
1	A	1879	PHE	Peptide
1	A	196	ASP	Mainchain
1	A	377	GLU	Peptide
1	A	385	GLU	Peptide
1	A	55	ASP	Peptide
1	A	698	PRO	Peptide
1	A	73	HIS	Peptide
1	A	775	ASN	Sidechain
2	C	308	CYS	Peptide
2	C	358	LYS	Peptide
2	C	360	ALA	Peptide
2	C	534	VAL	Peptide
2	C	902	HIS	Peptide
23	I	309	ALA	Mainchain
23	I	310	LYS	Mainchain
23	I	311	MET	Mainchain
23	I	83	LYS	Mainchain
23	I	84	HIS	Mainchain
6	J	215	THR	Peptide
6	J	216	ASP	Peptide
6	J	237	LYS	Mainchain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
9	N	136	HIS	Peptide
9	N	3	LYS	Peptide
9	N	4	VAL	Peptide
10	O	63	MET	Peptide
11	P	204	GLN	Peptide
11	P	30	TYR	Peptide
11	P	48	GLN	Peptide
36	Q	90	TYR	Mainchain
12	R	183	GLN	Peptide
12	R	185	GLY	Peptide
12	R	212	PHE	Peptide
12	R	303	GLU	Mainchain
12	R	66	GLU	Peptide
12	R	94	GLY	Peptide
17	W	518	PRO	Peptide
28	d	112	ASN	Peptide
28	k	112	ASN	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	
1	A	1980/2335 (85%)	1807 (91%)	160 (8%)	13 (1%)	22	61
2	C	856/972 (88%)	790 (92%)	64 (8%)	2 (0%)	47	79
3	E	297/357 (83%)	276 (93%)	20 (7%)	1 (0%)	41	75
6	J	529/848 (62%)	491 (93%)	31 (6%)	7 (1%)	12	50
7	L	425/802 (53%)	408 (96%)	15 (4%)	2 (0%)	29	68

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
8	M	128/243 (53%)	117 (91%)	11 (9%)	0	100	100
9	N	141/144 (98%)	125 (89%)	14 (10%)	2 (1%)	11	48
10	O	283/420 (67%)	259 (92%)	23 (8%)	1 (0%)	34	71
11	P	107/229 (47%)	88 (82%)	16 (15%)	3 (3%)	5	34
12	R	274/536 (51%)	247 (90%)	23 (8%)	4 (2%)	10	47
13	S	157/166 (95%)	148 (94%)	9 (6%)	0	100	100
14	T	310/514 (60%)	278 (90%)	25 (8%)	7 (2%)	6	38
15	U	68/2752 (2%)	60 (88%)	8 (12%)	0	100	100
16	V	444/908 (49%)	431 (97%)	12 (3%)	1 (0%)	47	79
17	W	507/579 (88%)	433 (85%)	68 (13%)	6 (1%)	13	51
21	b	98/240 (41%)	93 (95%)	2 (2%)	3 (3%)	4	32
21	i	84/240 (35%)	82 (98%)	2 (2%)	0	100	100
22	X	67/254 (26%)	64 (96%)	3 (4%)	0	100	100
23	I	575/855 (67%)	562 (98%)	11 (2%)	2 (0%)	41	75
24	y	77/301 (26%)	74 (96%)	1 (1%)	2 (3%)	5	35
25	Y	667/1220 (55%)	642 (96%)	23 (3%)	2 (0%)	41	75
26	a	77/126 (61%)	76 (99%)	1 (1%)	0	100	100
26	h	77/126 (61%)	76 (99%)	1 (1%)	0	100	100
27	c	80/119 (67%)	77 (96%)	3 (4%)	0	100	100
27	j	80/119 (67%)	77 (96%)	3 (4%)	0	100	100
28	d	95/118 (80%)	91 (96%)	4 (4%)	0	100	100
28	k	81/118 (69%)	78 (96%)	3 (4%)	0	100	100
29	f	72/86 (84%)	69 (96%)	3 (4%)	0	100	100
29	m	72/86 (84%)	69 (96%)	3 (4%)	0	100	100
30	e	77/92 (84%)	76 (99%)	1 (1%)	0	100	100
30	l	76/92 (83%)	75 (99%)	1 (1%)	0	100	100
31	g	72/76 (95%)	70 (97%)	2 (3%)	0	100	100
31	n	65/76 (86%)	63 (97%)	2 (3%)	0	100	100
32	v	142/146 (97%)	138 (97%)	4 (3%)	0	100	100
33	w	89/174 (51%)	87 (98%)	1 (1%)	1 (1%)	14	53
34	u	384/411 (93%)	372 (97%)	9 (2%)	3 (1%)	19	59

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
35	x	23/703 (3%)	22 (96%)	1 (4%)	0	100	100
36	Q	1308/1485 (88%)	1278 (98%)	27 (2%)	3 (0%)	47	79
37	o	160/255 (63%)	146 (91%)	12 (8%)	2 (1%)	12	50
38	p	92/225 (41%)	90 (98%)	2 (2%)	0	100	100
39	q	130/504 (26%)	117 (90%)	7 (5%)	6 (5%)	2	23
39	r	129/504 (26%)	118 (92%)	9 (7%)	2 (2%)	9	46
39	s	130/504 (26%)	114 (88%)	8 (6%)	8 (6%)	1	17
39	t	129/504 (26%)	116 (90%)	9 (7%)	4 (3%)	4	32
40	K	147/225 (65%)	139 (95%)	5 (3%)	3 (2%)	7	41
All	All	11861/21789 (54%)	11109 (94%)	662 (6%)	90 (1%)	24	59

All (90) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	167	PRO
1	A	857	ASN
1	A	942	PRO
1	A	1831	LYS
6	J	217	GLU
7	L	10	VAL
7	L	125	PRO
9	N	41	ARG
9	N	43	VAL
11	P	7	PRO
11	P	8	THR
12	R	233	PRO
14	T	187	LYS
14	T	188	PRO
14	T	341	ALA
14	T	343	PRO
14	T	401	PRO
17	W	82	ASN
17	W	83	PRO
17	W	259	GLN
23	I	532	LYS
24	y	44	GLU
34	u	383	ASN
36	Q	934	TYR
39	q	24	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
39	q	59	HIS
39	q	60	PRO
39	s	9	ASN
39	s	55	ILE
39	s	60	PRO
39	s	66	PRO
39	s	71	ILE
39	t	9	ASN
39	t	69	THR
40	K	78	PRO
1	A	204	LEU
1	A	1092	ILE
2	C	439	PRO
6	J	709	VAL
14	T	344	GLN
25	Y	1185	ASP
33	w	115	GLY
34	u	340	GLY
34	u	385	ASP
37	o	160	LYS
39	q	9	ASN
39	q	19	PRO
39	s	24	VAL
1	A	1515	TRP
2	C	63	LYS
6	J	188	GLN
12	R	52	PRO
12	R	186	VAL
16	V	597	PRO
24	y	41	TYR
39	q	23	HIS
39	r	9	ASN
39	t	67	SER
1	A	570	ASP
1	A	1134	TRP
1	A	1135	PRO
3	E	193	THR
6	J	206	LEU
6	J	341	PRO
6	J	604	PRO
11	P	205	LYS
12	R	73	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
14	T	407	GLN
17	W	74	PRO
17	W	258	PRO
25	Y	1182	LYS
36	Q	525	PRO
37	o	32	PRO
39	t	65	PRO
40	K	65	ILE
1	A	364	SER
1	A	1494	TYR
1	A	1758	PRO
21	b	105	GLY
21	b	106	ILE
39	s	62	ARG
6	J	216	ASP
10	O	20	PHE
23	I	372	ARG
36	Q	532	PRO
21	b	115	PRO
17	W	271	PRO
40	K	17	PRO
39	s	38	GLY
39	r	60	PRO

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	1780/2108 (84%)	1702 (96%)	78 (4%)	28	63
2	C	761/866 (88%)	740 (97%)	21 (3%)	43	72
3	E	256/300 (85%)	255 (100%)	1 (0%)	91	97
6	J	241/751 (32%)	236 (98%)	5 (2%)	53	78
7	L	218/709 (31%)	209 (96%)	9 (4%)	30	64
8	M	117/209 (56%)	113 (97%)	4 (3%)	37	69

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
9	N	130/130 (100%)	123 (95%)	7 (5%)	22	57
10	O	255/361 (71%)	254 (100%)	1 (0%)	91	97
11	P	101/203 (50%)	94 (93%)	7 (7%)	15	49
12	R	236/457 (52%)	213 (90%)	23 (10%)	8	36
13	S	129/134 (96%)	128 (99%)	1 (1%)	81	91
14	T	268/441 (61%)	253 (94%)	15 (6%)	21	56
15	U	21/2432 (1%)	20 (95%)	1 (5%)	25	60
16	V	98/838 (12%)	95 (97%)	3 (3%)	40	71
17	W	448/502 (89%)	434 (97%)	14 (3%)	40	71
21	b	83/177 (47%)	81 (98%)	2 (2%)	49	75
21	i	77/177 (44%)	75 (97%)	2 (3%)	46	74
22	X	68/230 (30%)	46 (68%)	22 (32%)	0	2
23	I	198/749 (26%)	185 (93%)	13 (7%)	16	51
25	Y	32/1085 (3%)	26 (81%)	6 (19%)	1	9
26	a	72/101 (71%)	72 (100%)	0	100	100
26	h	70/101 (69%)	70 (100%)	0	100	100
27	c	77/101 (76%)	75 (97%)	2 (3%)	46	74
27	j	77/101 (76%)	75 (97%)	2 (3%)	46	74
28	d	90/110 (82%)	88 (98%)	2 (2%)	52	77
28	k	80/110 (73%)	78 (98%)	2 (2%)	47	75
29	f	63/74 (85%)	62 (98%)	1 (2%)	62	83
29	m	63/74 (85%)	62 (98%)	1 (2%)	62	83
30	e	74/84 (88%)	74 (100%)	0	100	100
30	l	73/84 (87%)	73 (100%)	0	100	100
31	g	64/66 (97%)	63 (98%)	1 (2%)	62	83
31	n	59/66 (89%)	58 (98%)	1 (2%)	60	82
37	o	139/218 (64%)	135 (97%)	4 (3%)	42	72
38	p	82/195 (42%)	79 (96%)	3 (4%)	34	66
All	All	6600/14344 (46%)	6346 (96%)	254 (4%)	36	66

All (254) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	164	MET
1	A	181	ASN
1	A	188	LEU
1	A	189	GLU
1	A	191	ILE
1	A	193	LEU
1	A	194	GLU
1	A	195	LEU
1	A	199	GLU
1	A	202	PRO
1	A	240	ARG
1	A	346	ASP
1	A	362	ARG
1	A	387	PHE
1	A	409	ARG
1	A	413	LEU
1	A	579	GLN
1	A	642	ARG
1	A	664	HIS
1	A	666	LYS
1	A	762	ARG
1	A	773	LYS
1	A	774	LYS
1	A	818	GLU
1	A	847	LYS
1	A	852	VAL
1	A	853	LYS
1	A	855	ARG
1	A	856	LEU
1	A	858	GLN
1	A	861	ARG
1	A	880	ARG
1	A	944	ASP
1	A	994	ASN
1	A	1089	CYS
1	A	1091	TYR
1	A	1093	ASP
1	A	1131	LYS
1	A	1136	ARG
1	A	1189	MET
1	A	1249	MET
1	A	1251	SER
1	A	1253	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1257	THR
1	A	1258	LYS
1	A	1261	ASN
1	A	1263	TRP
1	A	1318	THR
1	A	1319	PRO
1	A	1341	ARG
1	A	1402	ARG
1	A	1418	ARG
1	A	1459	ARG
1	A	1460	HIS
1	A	1468	ASN
1	A	1471	ARG
1	A	1496	PRO
1	A	1501	LEU
1	A	1502	PHE
1	A	1504	GLU
1	A	1505	LYS
1	A	1528	GLN
1	A	1554	GLN
1	A	1555	LEU
1	A	1556	ASP
1	A	1560	ILE
1	A	1561	PHE
1	A	1562	MET
1	A	1565	LYS
1	A	1578	ARG
1	A	1749	LYS
1	A	1751	LEU
1	A	1752	GLN
1	A	1753	LEU
1	A	1755	SER
1	A	1757	GLU
1	A	1763	LEU
1	A	1833	LEU
2	C	64	LYS
2	C	82	GLN
2	C	84	GLU
2	C	85	ASP
2	C	87	GLN
2	C	93	ILE
2	C	95	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	C	297	ASN
2	C	359	LYS
2	C	495	ARG
2	C	513	ASN
2	C	572	GLU
2	C	573	GLU
2	C	673	LYS
2	C	748	ASP
2	C	749	THR
2	C	750	LEU
2	C	756	LYS
2	C	799	GLU
2	C	802	HIS
2	C	826	ARG
3	E	190	PHE
6	J	201	ARG
6	J	205	LEU
6	J	221	ASN
6	J	239	ARG
6	J	244	ASN
7	L	4	ILE
7	L	5	MET
7	L	6	ILE
7	L	7	LYS
7	L	10	VAL
7	L	123	LEU
7	L	133	GLU
7	L	181	ARG
7	L	240	ARG
8	M	152	LEU
8	M	198	ARG
8	M	212	ASN
8	M	215	ASN
9	N	35	GLU
9	N	38	GLU
9	N	40	LYS
9	N	41	ARG
9	N	43	VAL
9	N	119	CYS
9	N	130	ARG
10	O	222	ARG
11	P	3	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
11	P	7	PRO
11	P	8	THR
11	P	10	GLU
11	P	13	ARG
11	P	18	LYS
11	P	212	ASN
12	R	78	ARG
12	R	95	LYS
12	R	123	GLU
12	R	124	VAL
12	R	128	ASP
12	R	131	ASP
12	R	212	PHE
12	R	217	LYS
12	R	220	ARG
12	R	303	GLU
12	R	307	GLN
12	R	310	ARG
12	R	312	MET
12	R	314	GLN
12	R	315	LYS
12	R	316	GLU
12	R	317	LYS
12	R	324	LEU
12	R	325	ARG
12	R	329	GLN
12	R	332	ARG
12	R	334	ARG
12	R	335	ARG
13	S	126	HIS
14	T	189	GLN
14	T	191	HIS
14	T	197	TYR
14	T	338	CYS
14	T	343	PRO
14	T	400	PHE
14	T	401	PRO
14	T	402	ASP
14	T	405	PHE
14	T	407	GLN
14	T	408	ASN
14	T	412	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
14	T	416	ILE
14	T	442	ARG
14	T	454	VAL
15	U	23	LEU
16	V	458	THR
16	V	535	THR
16	V	597	PRO
17	W	72	LEU
17	W	74	PRO
17	W	76	VAL
17	W	129	ARG
17	W	182	LYS
17	W	200	VAL
17	W	233	LYS
17	W	243	VAL
17	W	257	ILE
17	W	322	ARG
17	W	492	ASN
17	W	495	ARG
17	W	500	LYS
17	W	529	ASN
21	b	13	ILE
21	b	58	GLN
22	X	54	LEU
22	X	57	ARG
22	X	59	GLN
22	X	60	GLU
22	X	62	LYS
22	X	72	GLN
22	X	78	MET
22	X	79	VAL
22	X	84	GLU
22	X	86	GLU
22	X	92	GLU
22	X	96	GLN
22	X	97	GLN
22	X	99	LEU
22	X	101	GLU
22	X	102	LYS
22	X	104	ARG
22	X	106	GLU
22	X	113	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
22	X	114	GLU
22	X	116	ARG
22	X	120	LYS
23	I	532	LYS
23	I	548	PHE
23	I	606	TRP
23	I	668	CYS
23	I	691	CYS
23	I	694	ILE
23	I	699	THR
23	I	724	LEU
23	I	725	ARG
23	I	726	ILE
23	I	727	ARG
23	I	730	VAL
23	I	731	GLN
25	Y	413	LYS
25	Y	418	ASP
25	Y	419	PHE
25	Y	420	ASP
25	Y	421	GLU
25	Y	425	ILE
27	c	40	LEU
27	c	54	GLN
28	d	45	ASN
28	d	112	ASN
29	f	55	LEU
31	g	65	ASN
21	i	13	ILE
21	i	58	GLN
27	j	40	LEU
27	j	54	GLN
28	k	45	ASN
28	k	112	ASN
29	m	55	LEU
31	n	65	ASN
37	o	5	THR
37	o	55	ARG
37	o	114	SER
37	o	126	THR
38	p	17	LYS
38	p	46	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
38	p	87	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (92) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	57	GLN
1	A	78	ASN
1	A	105	ASN
1	A	181	ASN
1	A	361	HIS
1	A	457	ASN
1	A	467	GLN
1	A	495	GLN
1	A	584	HIS
1	A	703	GLN
1	A	704	ASN
1	A	711	GLN
1	A	792	HIS
1	A	793	ASN
1	A	994	ASN
1	A	1014	ASN
1	A	1096	HIS
1	A	1124	ASN
1	A	1148	ASN
1	A	1359	HIS
1	A	1460	HIS
1	A	1487	HIS
1	A	1528	GLN
1	A	1717	ASN
1	A	1766	GLN
2	C	60	HIS
2	C	137	HIS
2	C	208	HIS
2	C	210	ASN
2	C	245	HIS
2	C	297	ASN
2	C	451	HIS
2	C	513	ASN
2	C	680	ASN
2	C	802	HIS
2	C	856	HIS
3	E	101	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	E	192	ASN
3	E	253	ASN
6	J	212	GLN
6	J	221	ASN
6	J	244	ASN
6	J	275	ASN
6	J	331	GLN
6	J	347	HIS
7	L	175	GLN
7	L	186	GLN
7	L	266	HIS
8	M	212	ASN
9	N	54	HIS
10	O	113	ASN
10	O	196	GLN
10	O	294	ASN
12	R	189	ASN
12	R	279	HIS
12	R	307	GLN
12	R	320	HIS
13	S	87	HIS
13	S	111	GLN
13	S	126	HIS
14	T	297	HIS
14	T	384	HIS
14	T	437	HIS
14	T	455	GLN
16	V	474	HIS
17	W	71	HIS
17	W	82	ASN
17	W	264	ASN
17	W	472	ASN
17	W	492	ASN
17	W	529	ASN
17	W	549	HIS
21	b	22	GLN
21	b	76	ASN
22	X	61	GLN
22	X	72	GLN
22	X	96	GLN
23	I	601	GLN
26	a	60	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
27	c	64	ASN
28	d	69	ASN
29	f	58	HIS
31	g	65	ASN
26	h	60	GLN
21	i	22	GLN
21	i	76	ASN
27	j	64	ASN
28	k	69	ASN
29	m	58	HIS
31	n	65	ASN
37	o	130	HIS
38	p	7	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
18	B	82/117 (70%)	17 (20%)	4 (4%)
19	F	96/107 (89%)	45 (46%)	16 (16%)
20	H	133/188 (70%)	34 (25%)	10 (7%)
4	4	13/46 (28%)	8 (61%)	3 (23%)
5	G	80/174 (45%)	63 (78%)	19 (23%)
All	All	404/632 (63%)	167 (41%)	52 (12%)

All (167) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
4	4	-12	G
4	4	-11	G
4	4	-10	C
4	4	-9	C
4	4	-7	C
4	4	-6	C
4	4	-4	A
4	4	-1	G
5	G	2	U
5	G	3	A
5	G	5	G
5	G	6	A
5	G	7	G
5	G	8	C

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	G	10	U
5	G	11	A
5	G	12	G
5	G	13	C
5	G	14	A
5	G	17	U
5	G	21	A
5	G	22	C
5	G	23	U
5	G	24	G
5	G	25	G
5	G	26	U
5	G	27	U
5	G	28	A
5	G	29	C
5	G	30	C
5	G	31	U
5	G	73	G
5	G	74	G
5	G	75	U
5	G	76	U
5	G	77	U
5	G	78	C
5	G	79	C
5	G	80	U
5	G	81	U
5	G	82	G
5	G	83	A
5	G	84	U
5	G	87	U
5	G	88	G
5	G	89	U
5	G	90	C
5	G	91	A
5	G	92	U
5	G	93	A
5	G	96	U
5	G	97	A
5	G	98	U
5	G	99	C
5	G	100	C
5	G	101	U

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	G	102	G
5	G	103	U
5	G	119	A
5	G	120	G
5	G	121	C
5	G	122	U
5	G	123	C
5	G	124	G
5	G	125	C
5	G	126	G
5	G	127	G
5	G	128	U
5	G	129	U
5	G	130	G
5	G	132	G
18	B	12	U
18	B	13	C
18	B	19	A
18	B	20	G
18	B	21	A
18	B	22	U
18	B	23	C
18	B	24	G
18	B	25	C
18	B	26	A
18	B	28	A
18	B	36	C
18	B	38	C
18	B	45	C
18	B	57	G
18	B	70	A
18	B	71	C
19	F	6	C
19	F	7	G
19	F	8	C
19	F	9	U
19	F	10	U
19	F	12	G
19	F	25	C
19	F	26	U
19	F	27	A
19	F	28	A

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
19	F	29	A
19	F	31	U
19	F	33	G
19	F	34	G
19	F	35	A
19	F	36	A
19	F	37	C
19	F	38	G
19	F	40	U
19	F	43	A
19	F	45	A
19	F	46	G
19	F	47	A
19	F	48	A
19	F	49	G
19	F	51	U
19	F	54	G
19	F	56	A
19	F	59	G
19	F	60	C
19	F	61	C
19	F	62	C
19	F	68	C
19	F	73	A
19	F	74	U
19	F	78	A
19	F	79	C
19	F	80	G
19	F	81	C
19	F	82	A
19	F	83	A
19	F	84	A
19	F	85	U
19	F	86	U
19	F	87	C
20	H	13	C
20	H	14	C
20	H	15	U
20	H	16	U
20	H	17	U
20	H	19	G
20	H	24	A

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
20	H	25	G
20	H	28	C
20	H	29	A
20	H	30	A
20	H	31	G
20	H	33	G
20	H	37	U
20	H	39	U
20	H	40	C
20	H	41	U
20	H	42	G
20	H	43	U
20	H	112	G
20	H	143	A
20	H	147	G
20	H	152	G
20	H	153	A
20	H	154	C
20	H	156	U
20	H	157	G
20	H	164	C
20	H	165	A
20	H	168	A
20	H	169	C
20	H	177	A
20	H	178	A
20	H	179	C

All (52) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
4	4	-13	C
4	4	-12	G
4	4	-11	G
5	G	1	G
5	G	16	G
5	G	20	A
5	G	21	A
5	G	22	C
5	G	23	U
5	G	90	C
5	G	95	U

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	G	96	U
5	G	97	A
5	G	98	U
5	G	100	C
5	G	119	A
5	G	120	G
5	G	124	G
5	G	125	C
5	G	126	G
5	G	128	U
5	G	129	U
18	B	18	C
18	B	19	A
18	B	20	G
18	B	27	U
19	F	5	U
19	F	7	G
19	F	25	C
19	F	26	U
19	F	33	G
19	F	34	G
19	F	35	A
19	F	36	A
19	F	45	A
19	F	50	A
19	F	58	G
19	F	59	G
19	F	73	A
19	F	81	C
19	F	84	A
19	F	86	U
20	H	15	U
20	H	28	C
20	H	29	A
20	H	30	A
20	H	38	A
20	H	39	U
20	H	40	C
20	H	156	U
20	H	164	C
20	H	168	A

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	$\#\|Z\| > 2$
12	SEP	R	232	12	8,9,10	1.55	1 (12%)	8,12,14	1.71	2 (25%)
12	SEP	R	224	12	8,9,10	1.69	3 (37%)	8,12,14	2.06	3 (37%)

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
12	SEP	R	232	12	-	1/5/8/10	-
12	SEP	R	224	12	-	0/5/8/10	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
12	R	232	SEP	P-O1P	3.39	1.61	1.50
12	R	224	SEP	P-OG	-2.84	1.51	1.60
12	R	224	SEP	CB-CA	-2.13	1.45	1.52
12	R	224	SEP	P-O2P	-2.05	1.46	1.54

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
12	R	224	SEP	OG-CB-CA	-4.52	103.74	108.14
12	R	232	SEP	P-OG-CB	-4.07	107.09	118.30
12	R	232	SEP	OG-CB-CA	2.34	110.42	108.14
12	R	224	SEP	O3P-P-O2P	2.30	116.42	107.64
12	R	224	SEP	O3P-P-OG	-2.14	101.03	106.73

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
12	R	232	SEP	N-CA-CB-OG

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	$\#\|Z\| > 2$
45	ATP	Q	1501	43	26,33,33	1.72	8 (30%)	31,52,52	1.85	10 (32%)
42	GTP	C	1500	43	26,34,34	1.50	3 (11%)	32,54,54	1.95	7 (21%)
41	IHP	A	3000	-	36,36,36	0.72	0	54,60,60	1.05	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
45	ATP	Q	1501	43	-	4/18/38/38	0/3/3/3
42	GTP	C	1500	43	-	1/18/38/38	0/3/3/3
41	IHP	A	3000	-	-	3/30/54/54	0/1/1/1

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
42	C	1500	GTP	C5-C6	-4.84	1.37	1.47
45	Q	1501	ATP	C2'-C1'	-3.63	1.48	1.53
45	Q	1501	ATP	C4-N3	3.47	1.40	1.35
45	Q	1501	ATP	C6-N6	3.34	1.46	1.34
45	Q	1501	ATP	C2'-C3'	-2.76	1.45	1.53
42	C	1500	GTP	C5-C4	-2.41	1.36	1.43
42	C	1500	GTP	O4'-C4'	-2.25	1.40	1.45
45	Q	1501	ATP	O2'-C2'	-2.24	1.37	1.43
45	Q	1501	ATP	O3'-C3'	-2.17	1.37	1.43
45	Q	1501	ATP	C3'-C4'	-2.16	1.47	1.53
45	Q	1501	ATP	C2-N3	2.14	1.35	1.32

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
42	C	1500	GTP	PA-O3A-PB	-5.63	113.49	132.83
45	Q	1501	ATP	PB-O3B-PG	-5.46	114.08	132.83
42	C	1500	GTP	PB-O3B-PG	-4.37	117.82	132.83
45	Q	1501	ATP	N3-C2-N1	-4.17	122.16	128.68
42	C	1500	GTP	C5-C6-N1	3.72	120.52	113.95
42	C	1500	GTP	C2-N1-C6	-3.48	118.70	125.10
42	C	1500	GTP	C3'-C2'-C1'	3.20	105.80	100.98
42	C	1500	GTP	C8-N7-C5	2.83	108.39	102.99
45	Q	1501	ATP	PA-O3A-PB	-2.59	123.92	132.83
45	Q	1501	ATP	C1'-N9-C4	-2.52	122.22	126.64
45	Q	1501	ATP	O2G-PG-O1G	-2.40	101.27	110.68
45	Q	1501	ATP	O2A-PA-O1A	-2.35	100.62	112.24
45	Q	1501	ATP	O2G-PG-O3B	2.25	112.19	104.64
42	C	1500	GTP	O6-C6-C5	-2.21	120.05	124.37
45	Q	1501	ATP	O2B-PB-O1B	-2.08	101.95	112.24
45	Q	1501	ATP	O5'-C5'-C4'	2.07	116.13	108.99
45	Q	1501	ATP	O3G-PG-O3B	2.04	111.49	104.64

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
41	A	3000	IHP	C3-O13-P3-O43
42	C	1500	GTP	O4'-C4'-C5'-O5'
45	Q	1501	ATP	C5'-O5'-PA-O1A
45	Q	1501	ATP	C5'-O5'-PA-O2A
41	A	3000	IHP	C3-C4-O14-P4

Continued on next page...

Continued from previous page...

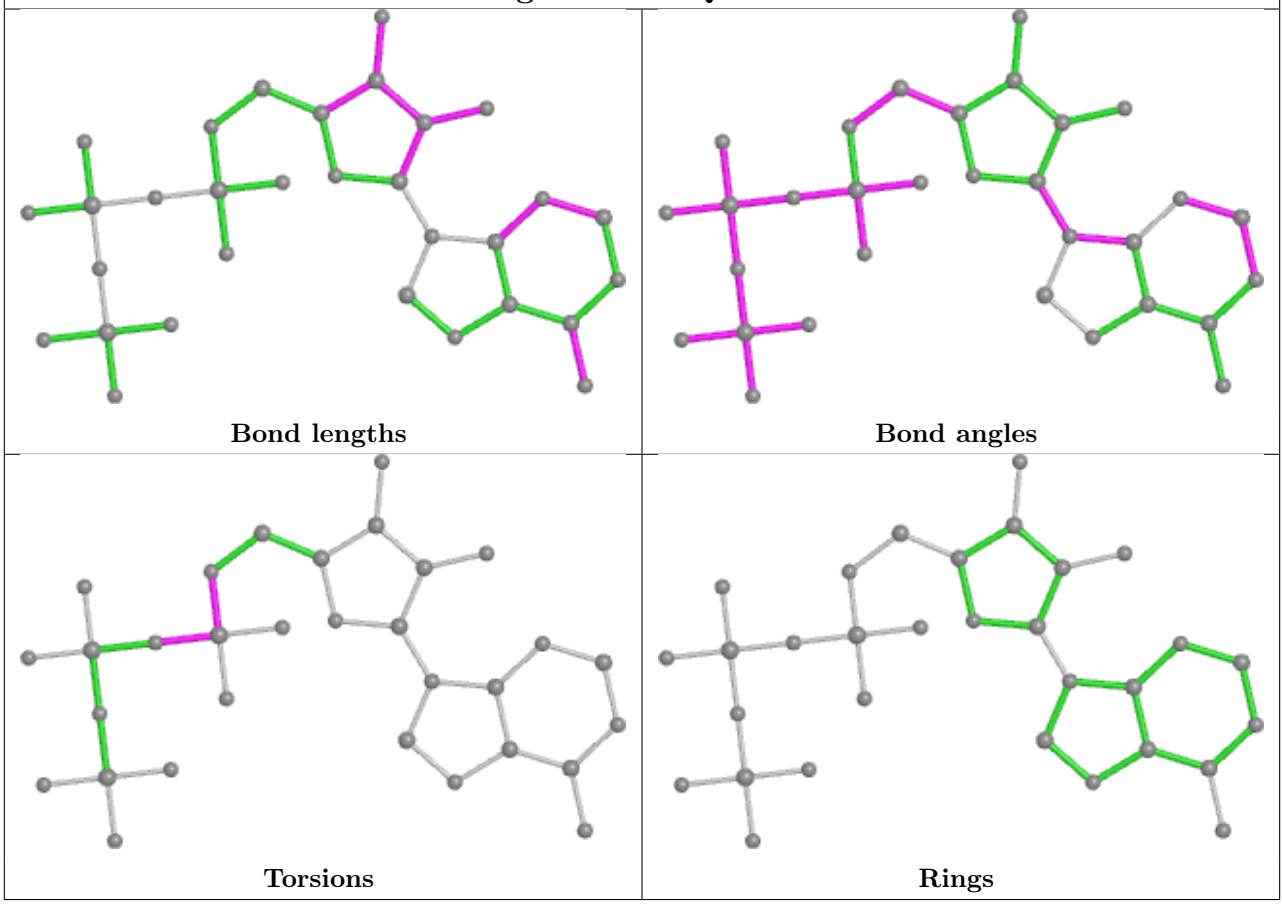
Mol	Chain	Res	Type	Atoms
45	Q	1501	ATP	PB-O3A-PA-O2A
41	A	3000	IHP	C4-O14-P4-O44
45	Q	1501	ATP	C5'-O5'-PA-O3A

There are no ring outliers.

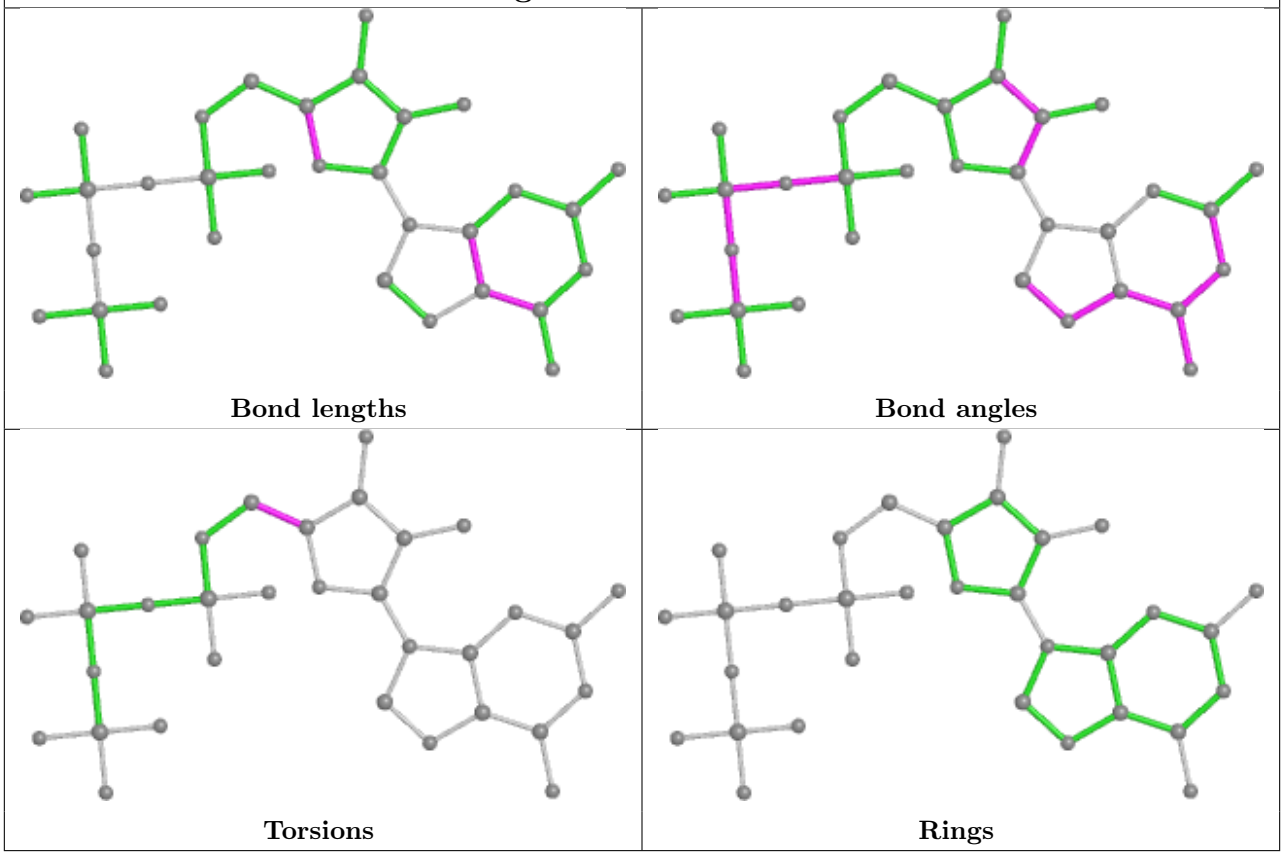
No monomer is involved in short contacts.

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

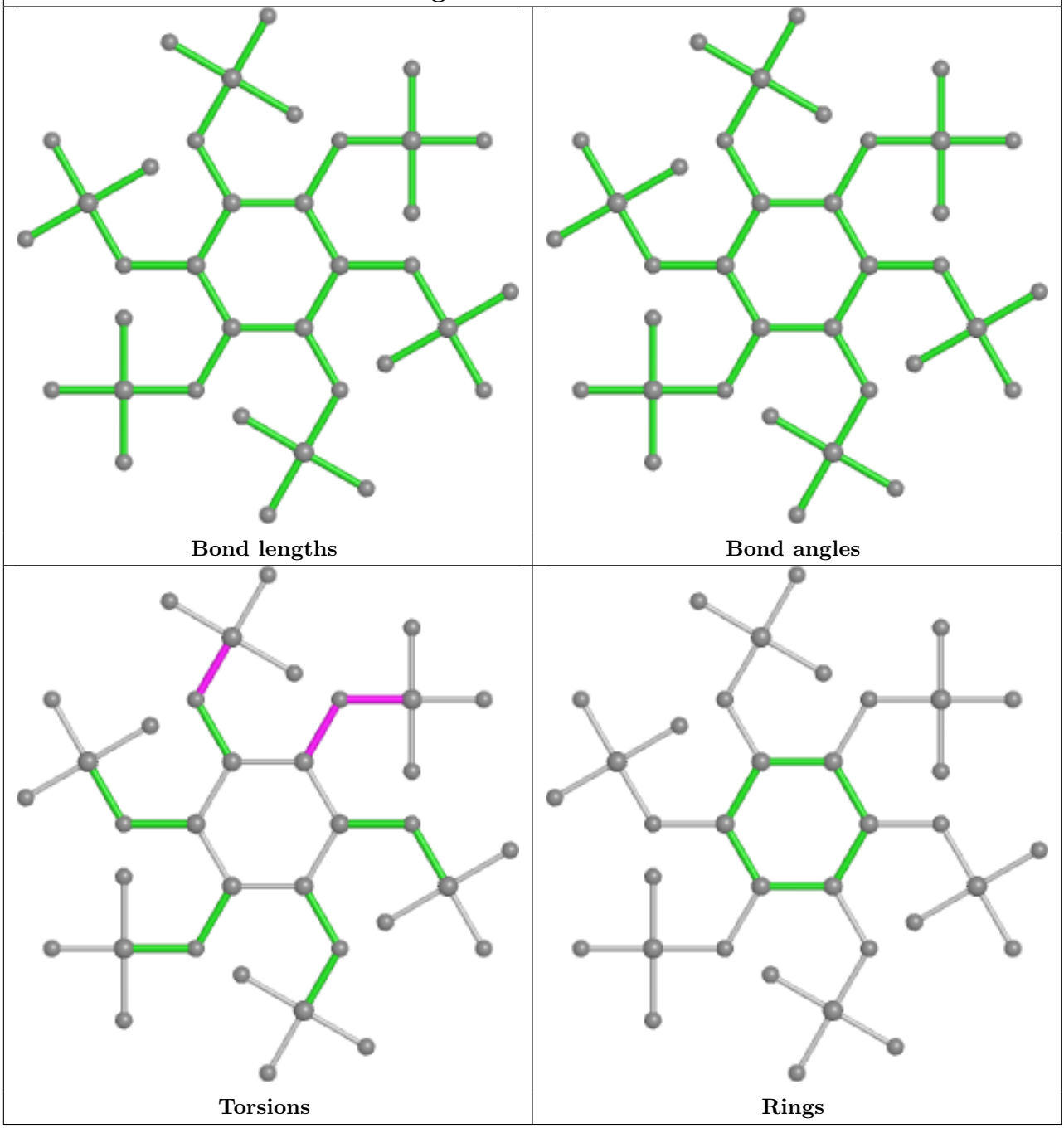
Ligand ATP Q 1501



Ligand GTP C 1500



Ligand IHP A 3000



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.

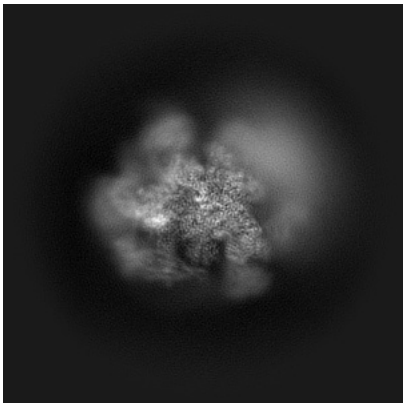
6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-32317. These allow visual inspection of the internal detail of the map and identification of artifacts.

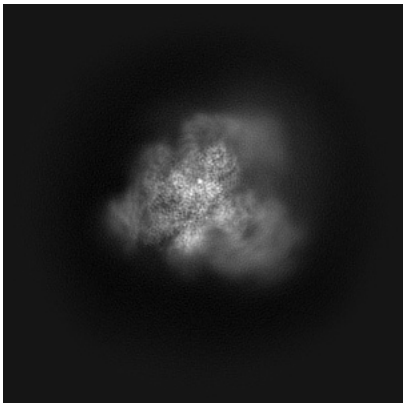
No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections (i)

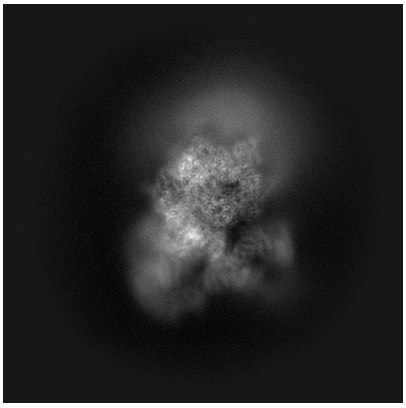
6.1.1 Primary map



X



Y

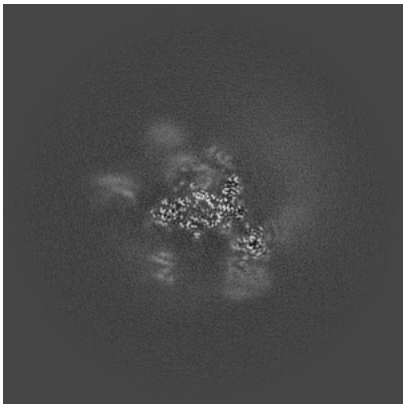


Z

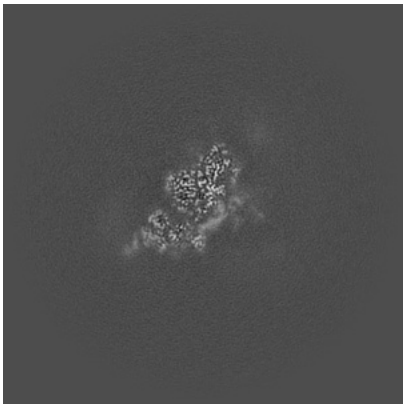
The images above show the map projected in three orthogonal directions.

6.2 Central slices (i)

6.2.1 Primary map



X Index: 200



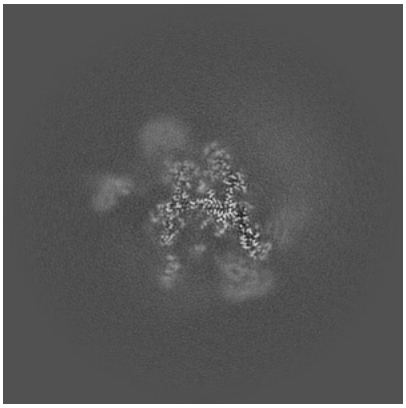
Y Index: 200

Z Index: 200

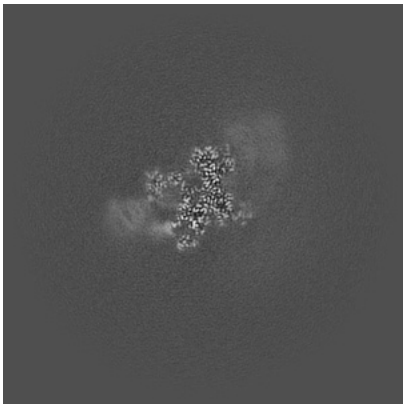
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices (i)

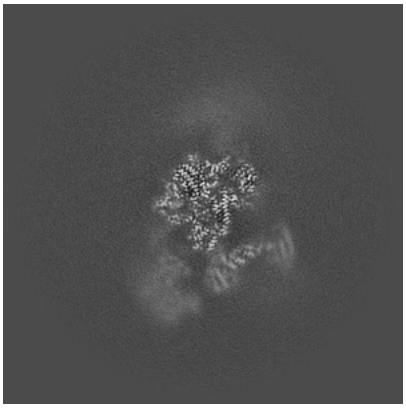
6.3.1 Primary map



X Index: 193



Y Index: 227



Z Index: 197

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 0.38. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

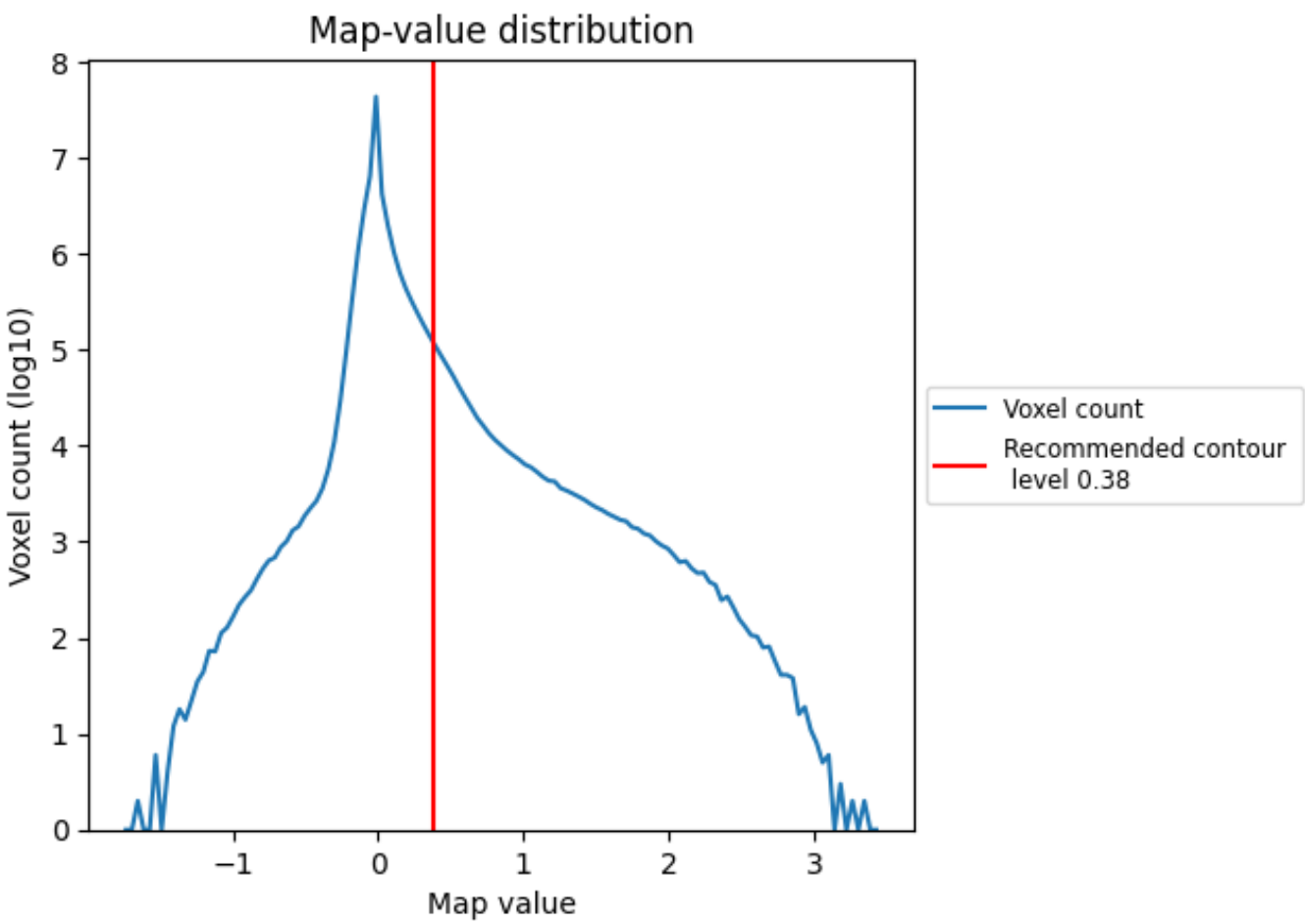
6.5 Mask visualisation (i)

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

7 Map analysis (i)

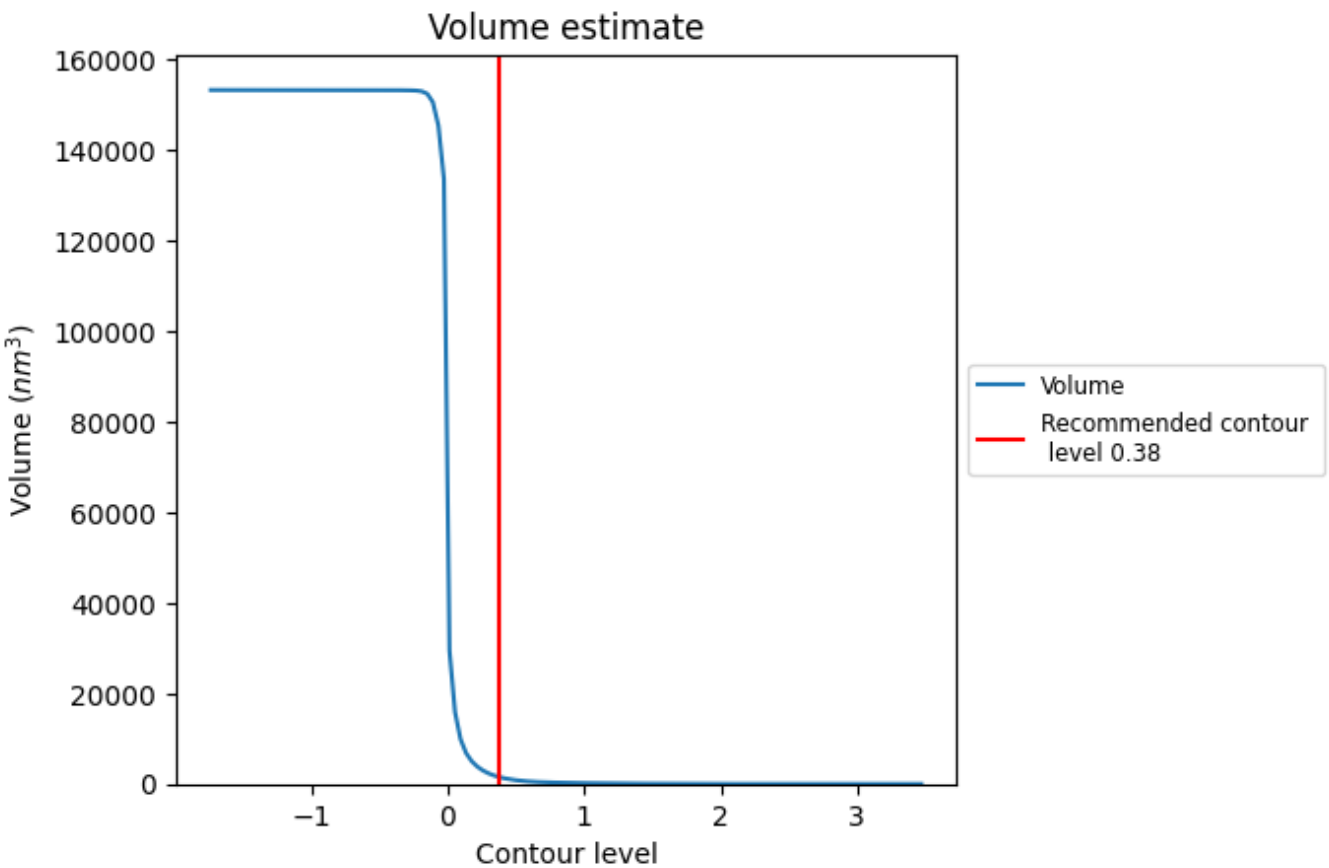
This section contains the results of statistical analysis of the map.

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.

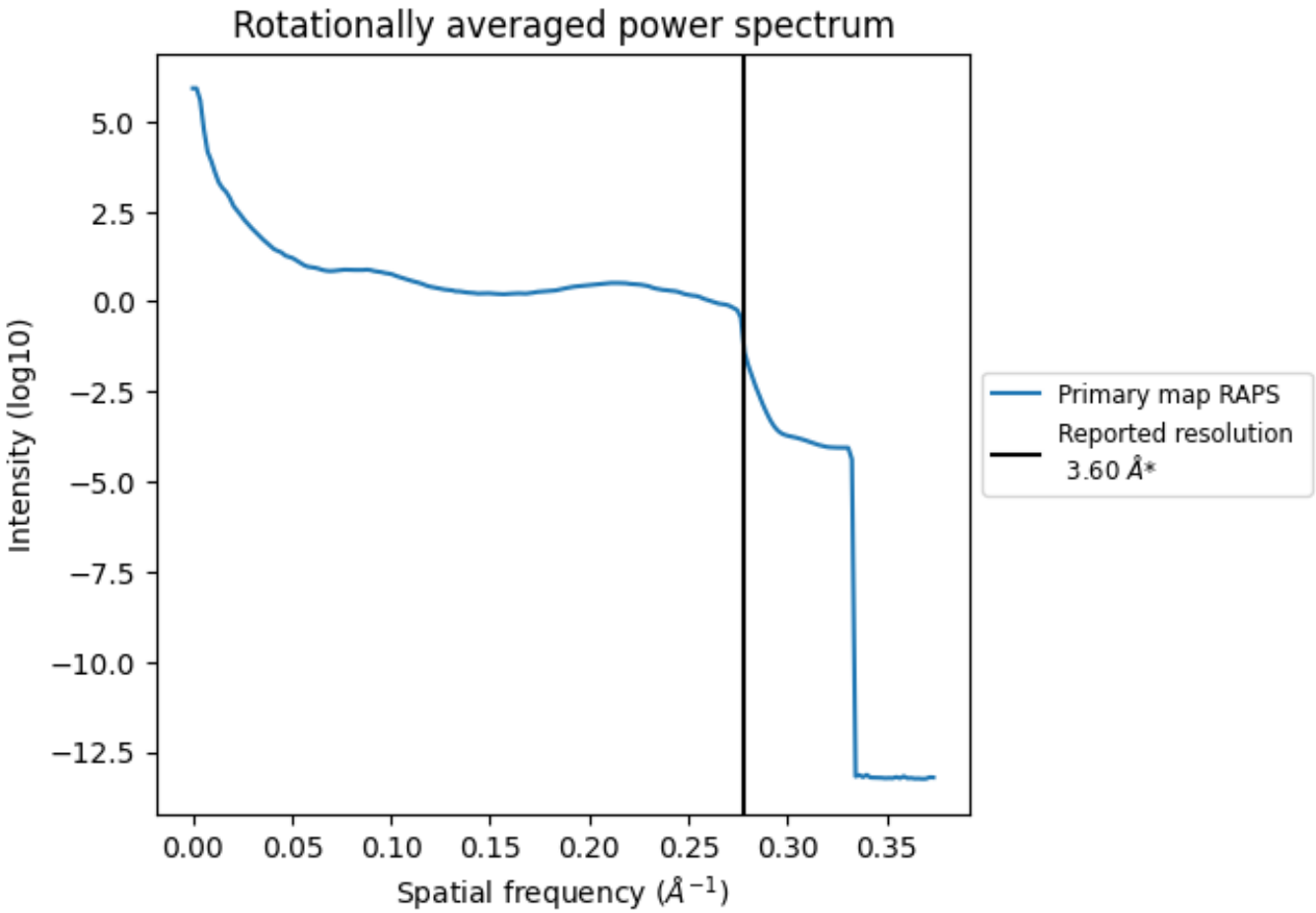
7.2 Volume estimate (i)



The volume at the recommended contour level is 1511 nm^3; this corresponds to an approximate mass of 1365 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 0.278 Å^{-1}

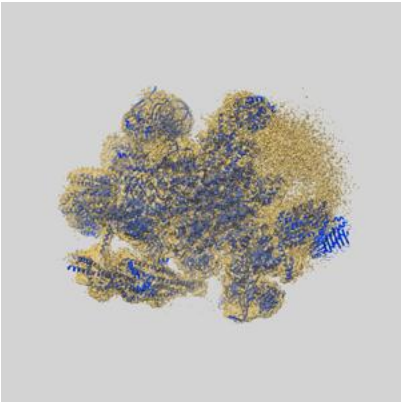
8 Fourier-Shell correlation (i)

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

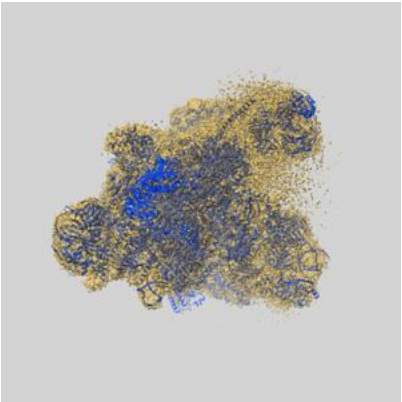
9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-32317 and PDB model 7W59. Per-residue inclusion information can be found in section 3 on page 14.

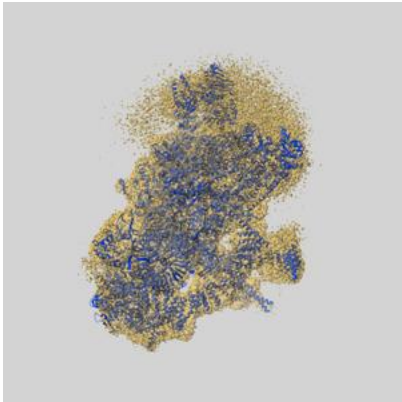
9.1 Map-model overlay (i)



X



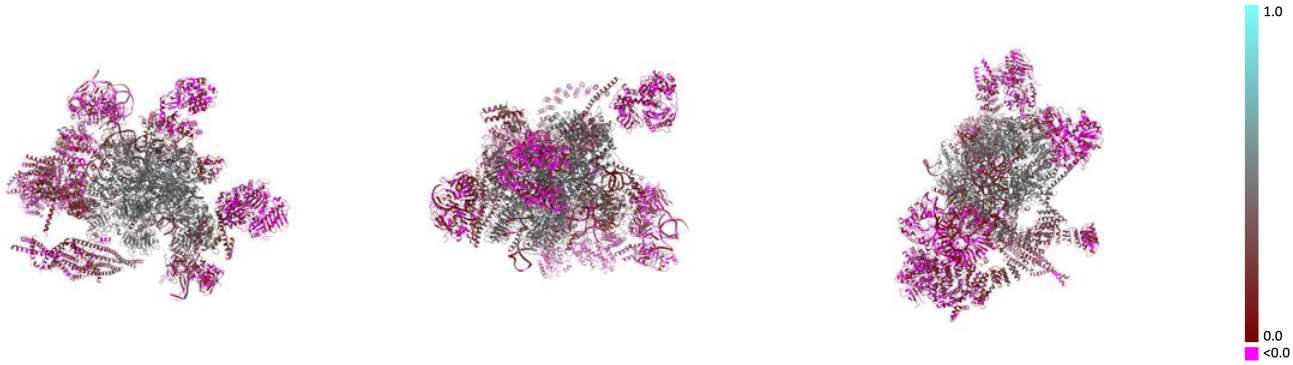
Y



Z

The images above show the 3D surface view of the map at the recommended contour level 0.38 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.

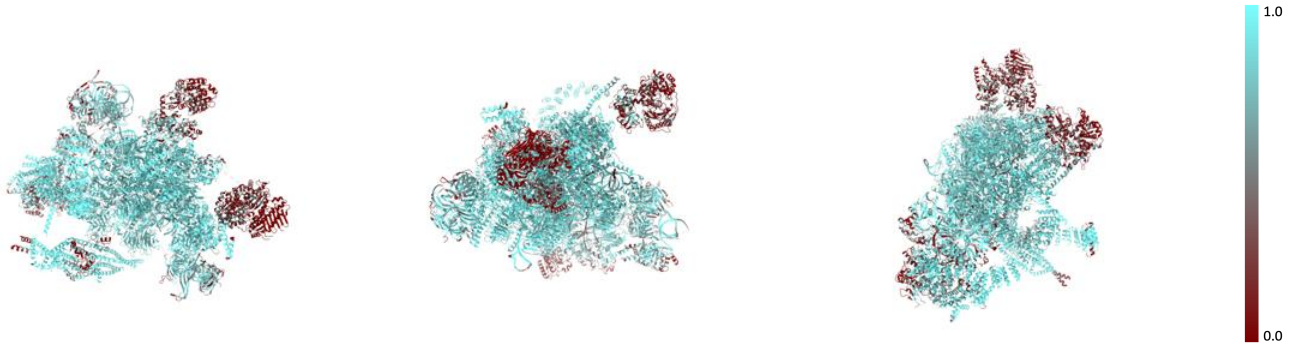
9.2 Q-score mapped to coordinate model (i)



X Y Z

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

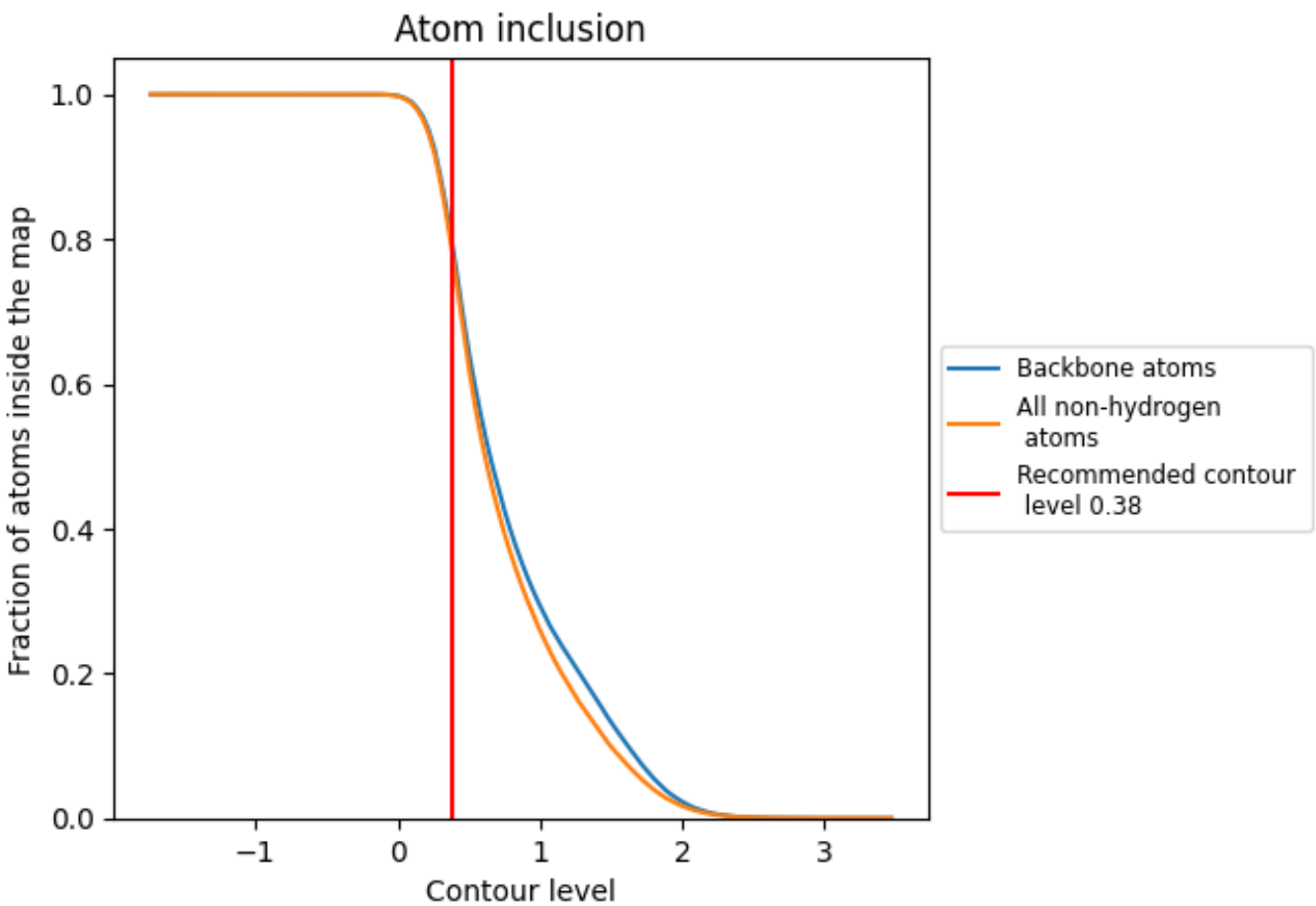
9.3 Atom inclusion mapped to coordinate model (i)



X Y Z

The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.38).

9.4 Atom inclusion (i)



At the recommended contour level, 80% of all backbone atoms, 78% of all non-hydrogen atoms, are inside the map.

9.5 Map-model fit summary (i)

The table lists the average atom inclusion at the recommended contour level (0.38) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.7843	0.2680
4	0.9928	0.4660
A	0.9084	0.4360
B	0.9412	0.3070
C	0.9125	0.4070
E	0.9423	0.4320
F	0.9505	0.3510
G	0.7397	0.1590
H	0.7977	0.1190
I	0.9146	0.2190
J	0.9069	0.3200
K	0.9469	0.1950
L	0.8986	0.3480
M	0.8759	0.4000
N	0.9384	0.4560
O	0.8938	0.3870
P	0.8342	0.4200
Q	0.5798	0.0360
R	0.8754	0.4190
S	0.9275	0.4120
T	0.9658	0.5060
U	0.8798	0.3190
V	0.5731	0.1630
W	0.8580	0.3230
X	0.9316	0.4390
Y	0.2981	0.0460
a	0.7687	0.1890
b	0.7536	0.1470
c	0.8625	0.0820
d	0.7677	0.0280
e	0.6855	0.1080
f	0.7668	0.0650
g	0.5852	0.0950
h	0.7065	0.0390
i	0.5378	0.0180

Continued on next page...

Continued from previous page...

Chain	Atom inclusion	Q-score
j	0.6141	0.0150
k	0.6202	0.0240
l	0.7022	0.0070
m	0.7986	0.0160
n	0.7585	0.0690
o	0.5350	0.0260
p	0.7863	0.0300
q	0.6085	0.0490
r	0.9327	0.1960
s	0.4962	0.1310
t	0.4144	0.0610
u	0.1987	0.0220
v	0.0070	-0.0000
w	0.0135	-0.0170
x	0.0645	0.0020
y	0.8000	0.1020