WORLDWIDE PDB PROTEIN DATA BANK

Full wwPDB EM Validation Report (i)

Oct 1, 2024 – 09:03 PM JST

PDB ID	:	8XI2
EMDB ID	:	EMD-38362
Title	:	Cryo-EM structure of the Chlamydomonas C* complex
Authors	:	Lu, Y.; Zhan, X.
Deposited on	:	2023-12-19
Resolution	:	2.60 Å(reported)

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.60 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	Percentile Ranks	Value
Ramachandran outliers		0.4%
Sidechain outliers		13.1%
RNA backbone		0.28

Worse   *Better*

■ Percentile relative to all structures
▯ Percentile relative to all EM structures
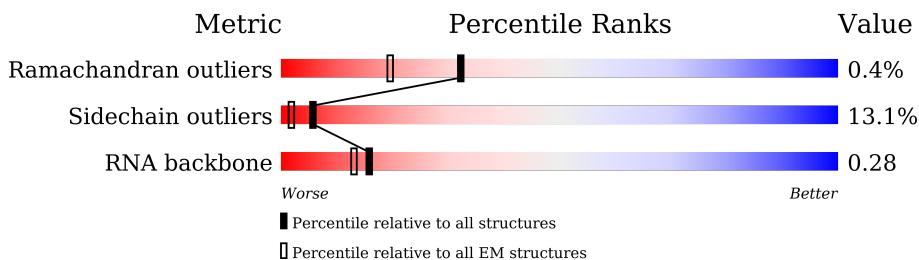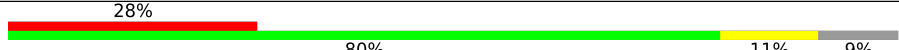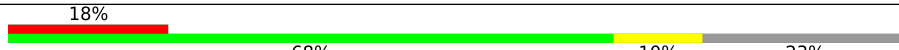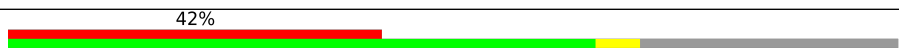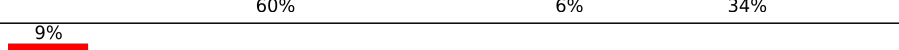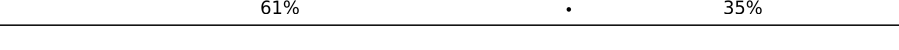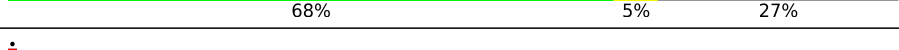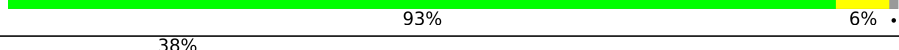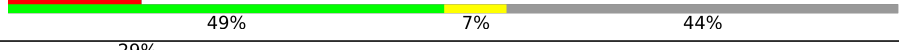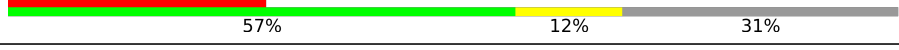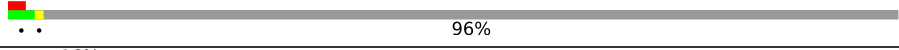
Metric	Whole archive (#Entries)	EM structures (#Entries)
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415
RNA backbone	6643	2191

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	2398	10% — 71% / 9% / 19%
2	B	111	5% — 29% / 46% / • / 23%
3	C	989	8% — 80% / 11% / 9%
4	E	362	78% / 10% / 12%
5	F	101	• — 27% / 58% / 6% / 9%
6	q	503	8% — 18% / • / 79%
6	r	503	6% — 22% / • / 74%
6	s	503	• — 13% / • / 86%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	t	503	10% 88%
7	a	127	7% 63% 8% 29%
8	g	77	9% 73% 12% 16%
9	e	87	34% 80% 11% 8%
10	f	85	34% 74% 9% 16%
11	d	110	28% 80% 11% 9%
12	c	114	18% 68% 10% 23%
13	b	265	6% 34% 63%
14	I	925	42% 66% 5% 29%
15	J	835	33% 67% 5% 29%
16	P	235	39% 5% 56%
17	M	563	13% 86%
18	T	518	60% 6% 34%
19	O	417	9% 61% 35%
20	N	233	68% 5% 27%
21	R	684	33% 5% 62%
22	H	192	7% 6% 14% 79%
23	S	157	93% 6%
24	W	576	38% 80% 15%
25	L	833	15% 49% 7% 44%
26	K	303	29% 57% 12% 31%
27	U	721	96%
28	V	928	16% 20% 80%
29	Q	1844	70% 70% 30%
30	5	7	71% 14% 14%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
31	3	173	6% 8% 9% · 80%

2 Entry composition ⓘ

There are 34 unique types of molecules in this entry. The entry contains 73326 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 MPN domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1931	Total 15194	C 9743	N 2699	O 2682	S 70	0	0

- Molecule 2 is a RNA chain called U5 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	86	Total 1817	C 812	N 307	O 612	P 86	0	0

- Molecule 3 is a protein called Elongation factor Tu, chloroplastic.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	897	Total 6973	C 4431	N 1203	O 1297	S 42	0	0

- Molecule 4 is a protein called U5-40K.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	E	318	Total 2441	C 1524	N 436	O 465	S 16	0	0

- Molecule 5 is a RNA chain called U6 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	F	92	Total 1985	C 885	N 369	O 637	P 94	0	0

- Molecule 6 is a protein called Pre-mRNA-processing factor 19.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	q	106	Total 831	C 522	N 145	O 160	S 4	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
6	r	131	Total 1021	C 638	N 183	O 196	S 4	0	0
6	s	70	Total 570	C 353	N 108	O 107	S 2	0	0
6	t	62	Total 504	C 312	N 95	O 96	S 1	0	0

- Molecule 7 is a protein called Small nuclear ribonucleoprotein Sm D3.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	a	90	Total 705	C 444	N 129	O 126	S 6	0	0

- Molecule 8 is a protein called Small nuclear ribonucleoprotein G.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	g	65	Total 512	C 323	N 90	O 96	S 3	0	0

- Molecule 9 is a protein called Small nuclear ribonucleoprotein E.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	e	80	Total 660	C 420	N 117	O 119	S 4	0	0

- Molecule 10 is a protein called Sm protein F.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	f	71	Total 559	C 358	N 91	O 106	S 4	0	0

- Molecule 11 is a protein called Small nuclear ribonucleoprotein Sm D2.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	d	100	Total 810	C 505	N 152	O 147	S 6	0	0

- Molecule 12 is a protein called Small nuclear ribonucleoprotein Sm D1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
12	c	88	684	435	121	124	4	0	0

- Molecule 13 is a protein called Sm domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
13	b	99	780	489	149	138	4	0	0

- Molecule 14 is a protein called Syf1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
14	I	661	4629	2900	855	861	13	0	0

- Molecule 15 is a protein called Crooked neck protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
15	J	595	4033	2511	759	753	10	0	0

- Molecule 16 is a protein called Cwf15/Cwc15 cell cycle control protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
16	P	104	867	536	175	156	0	0

- Molecule 17 is a protein called PPIase cyclophilin-type domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
17	M	80	644	402	121	120	1	0	0

- Molecule 18 is a protein called PLRG1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
18	T	344	2700	1705	488	490	17	0	0

- Molecule 19 is a protein called Rbm22.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	O	273	Total 2136	C 1337	N 376	O 405	S 18	0	0

- Molecule 20 is a protein called G10 protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	N	171	Total 1412	C 897	N 260	O 248	S 7	0	0

- Molecule 21 is a protein called SKI-interacting protein SKIP SNW domain-containing protein.

Mol	Chain	Residues	Atoms						AltConf	Trace
21	R	260	Total 1994	C 1251	N 361	O 377	P 2	S 3	0	0

- Molecule 22 is a RNA chain called U2 snRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	H	40	Total 838	C 375	N 134	O 289	P 40	0	0

- Molecule 23 is a protein called Peptidyl-prolyl cis-trans isomerase.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	S	155	Total 1198	C 754	N 213	O 225	S 6	0	0

- Molecule 24 is a protein called Prp17.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	W	487	Total 2826	C 1721	N 553	O 549	S 3	0	0

- Molecule 25 is a protein called Cdc5L.

Mol	Chain	Residues	Atoms					AltConf	Trace
25	L	468	Total 3795	C 2349	N 722	O 712	S 12	0	0

- Molecule 26 is a protein called Pre-mRNA-splicing factor SPF27.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	K	210	Total 1688	C 1045	N 318	O 321	S 4	0	0

- Molecule 27 is a protein called CWF21 domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
27	U	26	Total 191	C 119	N 34	O 36	S 2	0	0

- Molecule 28 is a protein called MI domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	V	186	Total 1020	C 621	N 198	O 200	S 1	0	0

- Molecule 29 is a protein called Intron-binding protein aquarius.

Mol	Chain	Residues	Atoms				AltConf	Trace
29	Q	1299	Total 6423	C 3825	N 1299	O 1299	0	0

- Molecule 30 is a RNA chain called RNA (5'-R(P*CP*CP*GP*AP*AP*CP*G)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
30	5	7	Total 150	C 67	N 29	O 47	P 7	0	0

- Molecule 31 is a RNA chain called RNA (173-mer).

Mol	Chain	Residues	Atoms					AltConf	Trace
31	3	34	Total 694	C 311	N 111	O 238	P 34	0	0

- Molecule 32 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).

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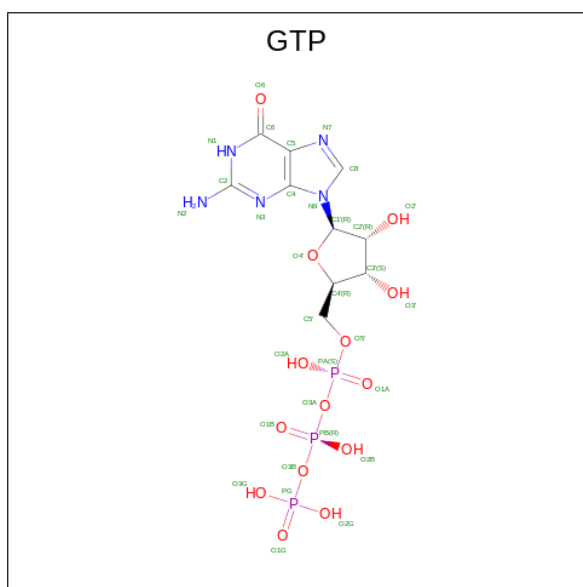
Mol	Chain	Residues	Atoms					AltConf
32	C	1	Total	C	N	O	P	0
			32	10	5	14	3	

- Molecule 33 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
33	C	1	Total	Mg	0
			1	1	
33	F	6	Total	Mg	0
			6	6	

- Molecule 34 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
34	O	2	Total	Zn	0
			2	2	
34	N	1	Total	Zn	0
			1	1	

3 Residue-property plots ⓘ

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

• Molecule 2: U5 snRNA

Chain B:

• Molecule 3: Elongation factor Tu, chloroplastic

Chain C:

• Molecule 4: U5-40K

Chain E:

• Molecule 5: U6 snRNA

Chain F:

• Molecule 6: Pre-mRNA-processing factor 19

Chain q:

M1 F2 C3 S4 I5 S6 G7 T8 V9 P10 E11 Q12 G13 T17 K18 K25 S26 L27 R33 V40 S46 T47 E48 D49 L50 L51 P52 L53 LYS VAL ASN LYS THR VAL LYS PRO ARG THR ALA PRO A66 L72 D78 E79 W80 M84 L85 H88 T92

Q96 T97 S102 H103 A104 A110 A111 T112 R113 V114 I115 A116 R117 L118 LEU LYS GLU ARG ASP GLU GLU ARG ALA ALA LEU ALA ASP MET LYS GLN GLN TYR ALA ALA ASP VAL ALA ALA ALA ALA ALA PRO ARG GLU GLY GLY GLY ILE GLU GLY GLY ARG ARG LYS ALA GLY ILE

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

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

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

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

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

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

• Molecule 6: Pre-mRNA-processing factor 19

Chain r: 6%   22%   •   74%

M1 F2 S6 Q12 G13 V14 K18 S19 G20 Y31 G36 T41 G42 E43 P44 L45 S46 T47 E48 D49 V55 N56 K57 T58 V59 T63 T67 S68 I69 P70 H77 D78 E79 W80 D81 E86 L87 H88 H95 S102 H103 A104 L105 Y106 Q107 T112

R113 V114 I115 A116 R117 L118 L119 K120 E121 A125 A128 L129 A130 D131 MET LYS GLN GLN TYR ALA ALA ASP VAL ALA ALA ALA ALA ALA PRO ARG GLU GLY GLY GLY ILE GLU GLY GLY ARG ARG LYS ALA GLY ILE ALA ALA ILE VAL ASP GLY VAL VAL THR ALA GLY VAL ASP MET THR GLU VAL ASN ALA

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

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

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

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

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

ALA LEU ARG TRP GLY PRO ASP ALA ALA ARG SER VAL LEU VAL GLY ALA ALA ASP HIS ASN LEU ARG VAL PHE SER SER PRO SER ALA

• Molecule 6: Pre-mRNA-processing factor 19

Chain s:　　13%　　　　　　　　　　　　　86%

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

PRO ARG THR ALA PRO ALA T67 G71 L72 L73 S74 M84 L85 R113 V114 I115 E121 R122 D123 A128 L129 A130 D131 Q134 Q135 Y136 ALA ALA ASP VAL ALA ALA ALA ALA PRO ARG GLY GLY GLU GLY GLY ARG ARG PRO LYS ALA ILE PRO GLU

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

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

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

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

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

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

• Molecule 6: Pre-mRNA-processing factor 19

Chain t:　　10%　　　　　　　　　　　　　88%

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

PRO ARG T63 A66 T67 S68 I69 P70 G71 L72 L73 S74 L75 D78 E79 W80 D81 A82 H88 T92 A104 L105 H108 R113 V114 I115 A116 R117 L118 L119 K120 E121 R122 D123 E124 ALA ARG ALA ALA LEU ALA ASP MET LYS GLN GLN TYR ALA ALA ASP VAL ALA ALA

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

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

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

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

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

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

ALA

• Molecule 7: Small nuclear ribonucleoprotein Sm D3

Chain a:

• Molecule 8: Small nuclear ribonucleoprotein G

Chain g:
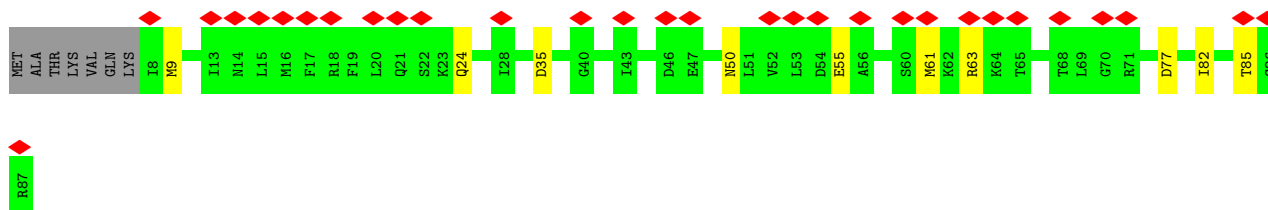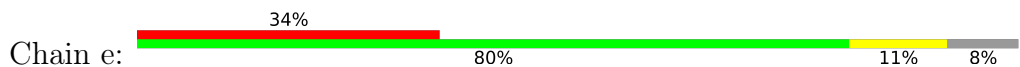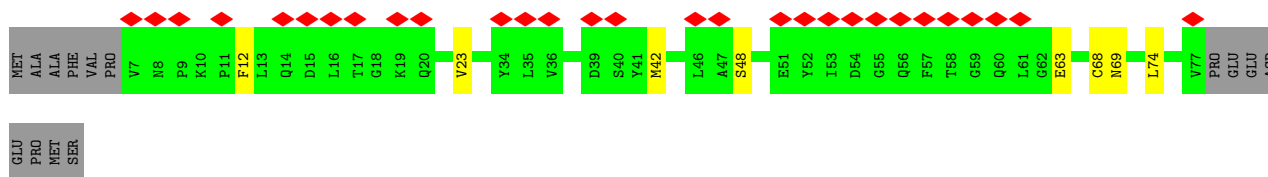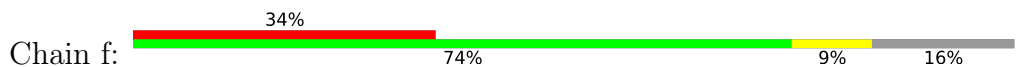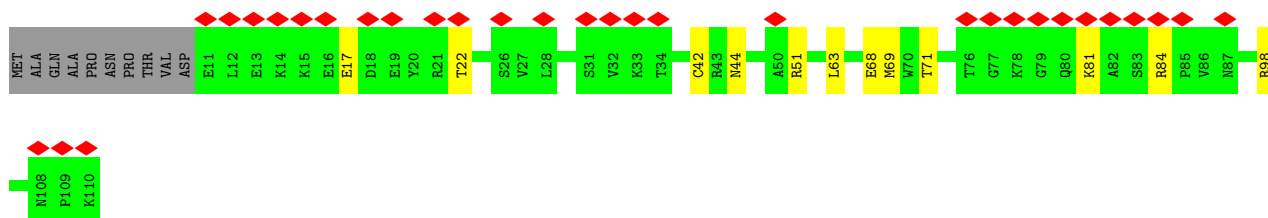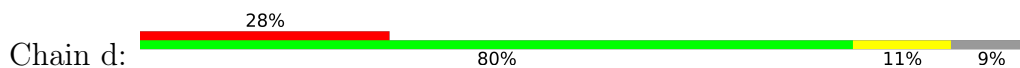
• Molecule 9: Small nuclear ribonucleoprotein E

Chain e:

• Molecule 10: Sm protein F

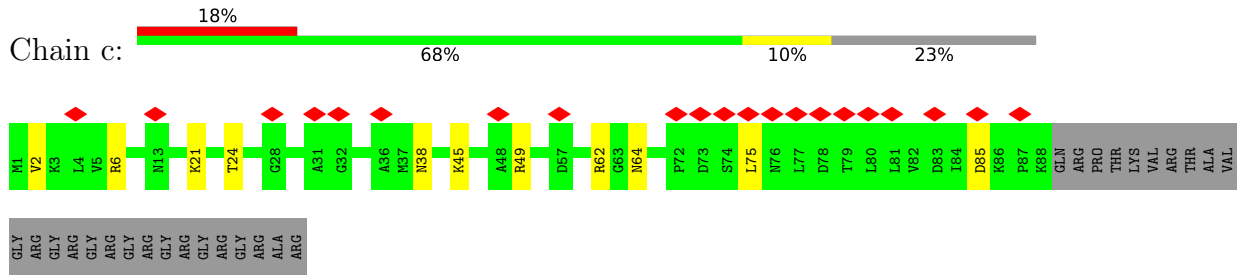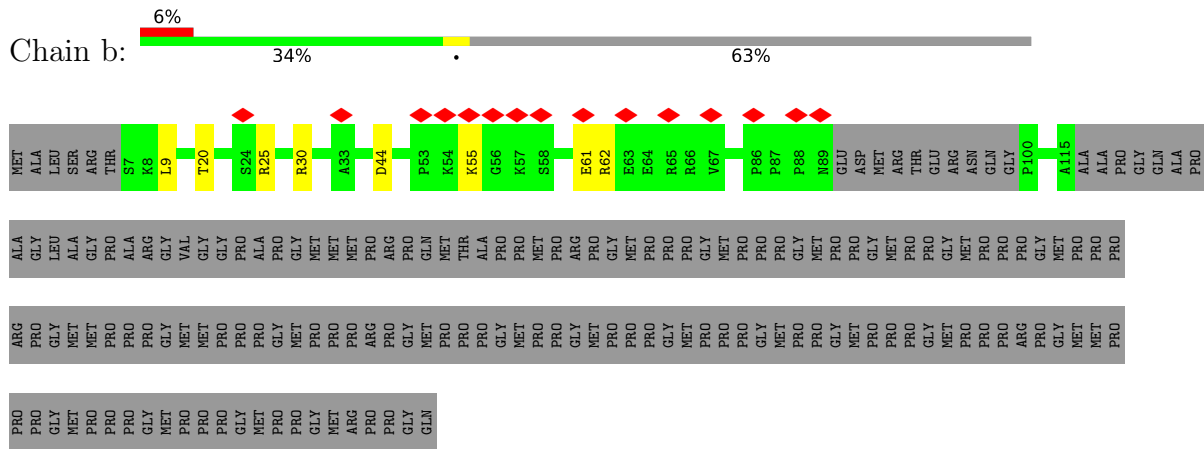Chain f:
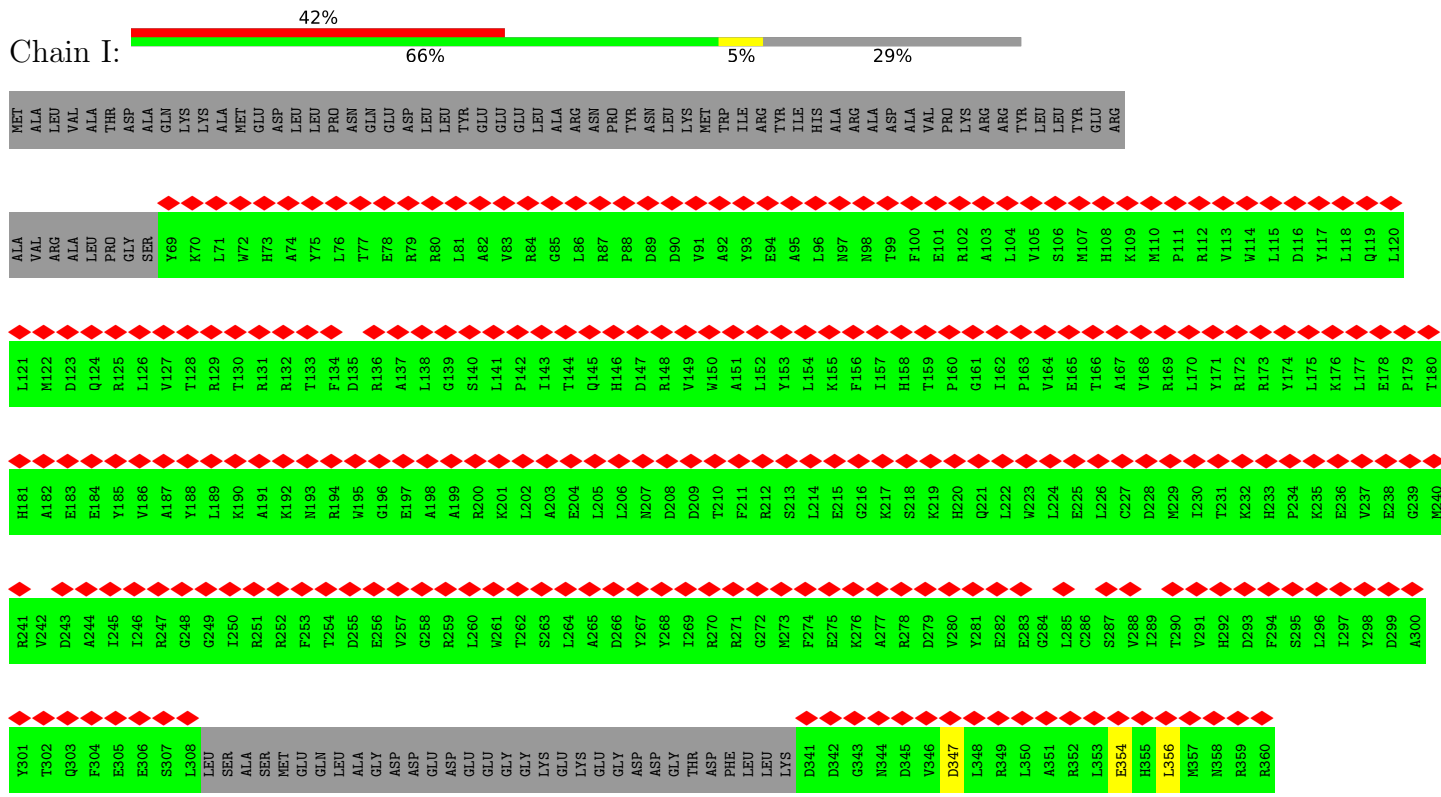
• Molecule 11: Small nuclear ribonucleoprotein Sm D2

Chain d:

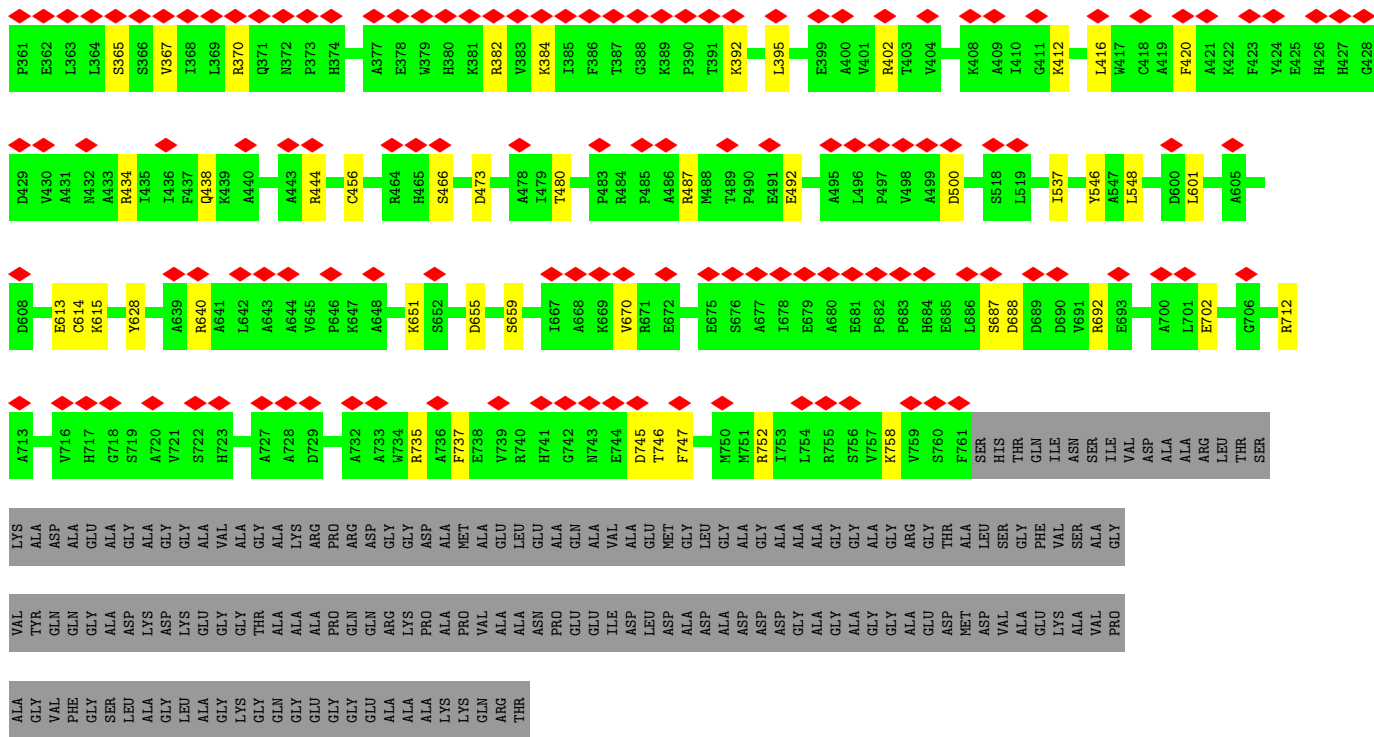• Molecule 12: Small nuclear ribonucleoprotein Sm D1

Chain c:

18%

68% 10% 23%

M1 V2 K3 L4 V5 R6 N13 K21 T24 G28 A31 G32 A36 M37 N38 K45 A48 R49 D57 R62 G63 N64 P72 D73 S74 L75 N76 L77 D78 T79 L80 L81 V82 D83 I84 D85 K86 P87 K88 GLN ARG PRO THR LYS VAL ARG THR ALA VAL

GLY ARG GLY ARG GLY ARG GLY ARG GLY ARG GLY ARG GLY ALA ARG

• Molecule 13: Sm domain-containing protein

Chain b:

6%

34% . 63%

MET ALA LEU SER ARG THR S7 K8 L9 T20 S24 R25 R30 A33 D44 P53 K54 K55 G56 K57 S58 E61 R62 E63 E64 R65 R66 V67 P86 P87 P88 N89 GLU ASP MET ARG THR GLU ARG ASN GLN GLY P100 A115 ALA ALA PRO GLY GLN ALA PRO

ALA GLY LEU ALA GLY PRO ALA ARG GLY VAL GLY GLY PRO ALA PRO GLY MET MET MET PRO ARG PRO GLY GLN MET THR ALA PRO PRO MET PRO ARG PRO GLY MET PRO PRO GLY MET PRO PRO GLY MET PRO PRO GLY MET PRO GLY MET PRO PRO MET PRO PRO GLY MET PRO PRO PRO GLY MET PRO PRO GLY MET PRO PRO GLY MET PRO PRO

ARG PRO GLY MET MET PRO PRO GLY MET MET PRO PRO PRO GLY MET PRO PRO PRO ARG PRO GLY MET PRO PRO PRO GLY MET PRO PRO GLY MET PRO PRO PRO GLY MET PRO PRO PRO GLY MET PRO PRO GLY MET PRO PRO GLY MET PRO PRO PRO GLY MET PRO PRO GLY MET PRO PRO PRO ARG PRO GLY MET MET PRO

PRO PRO GLY MET PRO PRO PRO GLY MET PRO PRO PRO GLY MET PRO PRO PRO GLY MET ARG PRO PRO GLY GLN

• Molecule 14: Syf1

Chain I:

42%

66% 5% 29%

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

ALA VAL ARG ALA LEU PRO GLY SER Y69 K70 L71 W72 H73 A74 Y75 L76 T77 E78 R79 R80 L81 A82 V83 R84 G85 L86 R87 P88 D89 D90 V91 A92 Y93 E94 A95 L96 N97 N98 T99 F100 E101 R102 A103 L104 V105 S106 M107 H108 K109 M110 P111 R112 V113 W114 L115 D116 Y117 L118 Q119 L120

L121 M122 D123 Q124 R125 L126 V127 T128 R129 T130 R131 R132 T133 F134 D135 R136 A137 L138 G139 S140 L141 P142 I143 T144 Q145 H146 D147 R148 V149 W150 A151 L152 Y153 L154 K155 F156 I157 H158 T159 P160 G161 I162 P163 V164 E165 T166 A167 V168 R169 L170 Y171 R172 R173 Y174 L175 K176 L177 E178 P179 T180

H181 A182 E183 E184 Y185 V186 A187 Y188 L189 K190 A191 K192 N193 R194 W195 G196 E197 A198 A199 K200 K201 L202 A203 E204 L205 L206 N207 D208 D209 T210 F211 R212 S213 L214 E215 G216 K217 S218 K219 H220 Q221 L222 W223 L224 E225 L226 C227 D228 M229 I230 T231 K232 H233 P234 K235 E236 V237 E238 G239 M240

R241 V242 D243 A244 I245 I246 R247 G248 G249 I250 R251 R252 F253 T254 D255 E256 V257 G258 R259 L260 W261 T262 S263 L264 A265 D266 Y267 Y268 I269 R270 R271 G272 M273 F274 E275 K276 A277 R278 D279 V280 Y281 E282 E283 G284 L285 C286 S287 V288 I289 T290 V291 H292 D293 F294 S295 L296 I297 Y298 D299 A300

Y301 T302 Q303 F304 E305 E306 S307 L308 LEU SER ALA SER MET GLU GLN LEU ALA GLY ASP ASP GLU ASP GLU GLU GLY GLY LYS GLU LEU GLY GLY ASP ASP GLY THR ASP PHE LEU LEU LYS D341 D342 G343 N344 D345 V346 D347 L348 R349 L350 A351 R352 L353 E354 H355 L356 M357 N358 R359 R360

• Molecule 15: Crooked neck protein

Chain J:

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

PRO GLN ALA GLN GLY GLY ARG GLY GLY ALA GLU GLU GLU GLU PRO SER ALA SER ASP ALA ALA GLN ALA PHE ALA ALA ALA MET SER ARG PHE ALA ALA GLY GLU ALA GLY GLU LYS GLN GLU GLU VAL ARG ALA GLY GLU VAL GLU GLU GLY ARG PRO ALA GLN GLU GLU GLN ASP ARG PRO ALA ARG GLY GLY

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

ALA ALA GLY GLU ALA MET ALA GLY GLU ASP SER

• Molecule 16: Cwf15/Cwc15 cell cycle control protein

Chain P: 39%   5%   56%

MET THR THR ALA HIS ARG PRO THR W9 I13 E16 GLU GLN GLY GLY MET ARG ILE PHE K25 K40 Q48 E62 R68 G76 T77 D78 THR SER PHE GLU ASP GLU ARG LYS ARG ASP LEU GLU LEU LEU GLN SER ALA PRO GLU GLY GLY ALA ARG GLN LEU

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

GLU GLU GLU LEU ARG ALA GLY ASN PRO LEU LEU ALA ILE GLY ALA ALA GLY ALA ALA GLY ALA ASP VAL SER PHE ASN ILE LYS R194 D199 N204 K211 Q231 R232 Y233 I234 R235

• Molecule 17: PPIase cyclophilin-type domain-containing protein

Chain M: 13%   86%

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

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

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

LEU LYS LEU LYS ALA GLU GLU LYS GLU GLU ALA ALA GLY ALA LYS ALA GLY SER LEU LYS SER ASP PRO ASN GLU LEU ASN PRO ASP GLN ASP ALA SER LEU ALA ARG LEU LYS ALA LEU ALA LYS GLN ALA ALA LYS ALA ARG ALA PRO GLY ALA GLY LYS SER ILE ALA ALA SER

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

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

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

ALA ALA LEU LYS TYR GLU GLU LYS ARG ASP LYS LYS GLY GLY PRO THR GLY TRP GLY VAL PHE A441 L445 Y446 Q447 E464 R468 A469 A472 D473 P474 E475 F476 TYR ARG GLY LEU ALA ASP S482 L483 K484 M499 E505 K510 ASP GLU PHE SER ARG ARG ARG ALA HIS ASN

ASP GLY SER ASP VAL ASP PHE ILE N529 D530 R531 F535 K538 F543 GLY GLN THR THR THR GLU ILE LYS GLY ASN LEU GLU ARG GLY THR ALA LEU PRO ASP ARG

• Molecule 18: PLRG1

Chain T:

60% 6% 34%

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

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

LEU ILE ASP ASN ILE PRO ALA ALA VAL ALA LYS LYS GLY ASP PRO GLY ALA GLY GLN LEU THR LEU TYR GLN PRO GLY GLY PRO ALA ALA PRO GLY ALA ASP ALA ALA HIS GLN ARG ALA VAL MET ALA VAL ILE ALA ASP LYS GLY GLY **N175** **R190** **M199**

V202 **C212** **V213** **D216** **G247** **D270** **V283** **S286** **L306** **D312** **S313** **V314** **K323** **H353** **S377** **M383** **F388** **E396** **S414** **N426** **R446** **C450** **S462** **L463** **E464** **S465** **L469** **E485** **E495** **V496** **E497** **D498** **R510** **D514** **F518**

• Molecule 19: Rbm22

Chain O:

9%

61% • 35%

MET ALA ALA ARG LEU LEU ARG ASP PRO GLU ALA ASP GLY TRP GLU GLU ARG SER **D18** **E24** **M35** **R49** **V68** **I69** **C70** **C83** **S110** **E111** **A124** GLY THR LEU GLY GLY GLY VAL GLY GLY ALA SER SER **S138** **Y139** **A140** **L151** **S154** **Y157** **R164** **V165**

C180 **A192** **N206** **T224** **P225** **P226** **E227** **D228** **T233** **L234** **Y235** **V236** **G237** **G256** **E257** **L258** **A259** **S260** **V261** **R262** **K263** **M264** **D265** **V266** **K267** **S268** **C269** **V272** **T273** **Y274** **T275** **S278** **A279** **A280** **E281** **K282** **A283** **A284** **G289** **L292** **G295** **A296** **R297** **V298** **G303** ARG PRO PRO

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

GLY GLY GLY PRO GLY GLY GLY PRO GLY MET ALA PRO PRO ARG PRO MET GLY TYR GLY ALA PRO PRO PRO TYR MET PRO PRO ARG PRO MET VAL SER SER ALA SER MET GLN PRO GLN GLN HIS GLN

• Molecule 20: G10 protein

Chain N:

68% 5% 27%

MET SER LEU **R4** **K10** **E21** **V22** **I23** **K41** **L46** **S72** **E82** **S118** **T134** **S146** **L161** ASP GLU GLU ASP SER **A167** **K168** **L169** **W176** **A179** GLY THR GLY GLU PRO GLY ASP GLY MET GLY GLY GLY GLY GLY GLY PRO GLY PRO SER ARG LYS ARG

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

• Molecule 21: SKI-interacting protein SKIP SNW domain-containing protein

Chain R:

33% 5% 62%

MET SER LEU LEU GLY THR LEU PRO ALA PRO GLN GLN LYS HIS ALA ALA PRO ILE ALA PRO VAL PRO ALA SER ALA ALA ILE VAL SER ALA SER PHE LYS GLU **P35** **L39** **K42** **I83** **K88** **V97** **R105** **D110** **L114** **K117** **V118** **D119** **R128** **E146**

K152 **L153** **S154** **S155** **P158** **T159** **K160** **L161** **G164** **S168** **T174** **D195** **V211** **S217** **V220** **L223** **H224** **S225** **T231** **V232** **K233** **Q236** **L257** **D258** **K259** **D264** **G265** ARG GLY LEU GLN GLN THR GLN ILE **N274** **T286** **A299** **S300** **I301** **Q302** LYS GLU LEU ALA

• Molecule 22: U2 snRNA

Chain H:

• Molecule 23: Peptidyl-prolyl cis-trans isomerase

Chain S:

• Molecule 24: Prp17

Chain W:

• Molecule 25: Cdc5L

E648 L649 E650 A651 L652 Q653 S654 A655 A656 E657 D660 R661 E662 V663 A664 Y665 L666 R667 S668 A669 W670 D671 H672 A673 K674 L675 S676 P677 D678 D679 Y680 S681 S686 D690 Y693 L694 P695 S696 R697 Q698 R699 Y700 E701 R702 ALA ALA SER ALA THR ASN VAL ASP ARG ILE ASP S714

L715 K716 S717 E718 F719 D720 A724 E727 A730 K731 A734 E737 G738 K739 L740 G741 L742 G746 L747 Q748 R749 D753 G756 R757 A763 V773 A777 L778 H779 E780 R781 A792 T800 R816 R817 R818 D819 E820 L821 A822 A823 A824 L825 A826 Q827 K828

R829 A830 A831 A832 S833

- Molecule 26: Pre-mRNA-splicing factor SPF27

Chain K:

29%
57% 12% 31%

MET ALA ALA LYS THR SER GLN ALA THR LEU ALA LEU GLU TYR THR ALA HIS GLY ALA ALA LYS GLY TRP ARG ALA N26 E27 H28 L29 V30 D31 D37 S38 V39 P40 P41 E42 L43 K44 P45 E48 A49 E52 E53 E54 K55 R56 R57 S58 T59 P62 S63 D64 Y65 E68

M69 P75 K76 F77 D78 P81 E86 Y87 E88 N92 K93 E94 P95 M96 A97 S101 V102 R103 Y104 R105 L106 E107 P108 P109 P110 Q111 A112 R113 R114 G115 D116 V117 G118 A119 W120 K121 S122 S123 L124 D125 N126 A127 A128 A129 Q130 L131 E132 H133 L136 L145 G148 D149

K150 A151 W152 R153 A154 D159 E160 A161 G165 L166 E167 A168 T177 D178 G179 L180 N181 L186 Q187 Q188 H189 A190 K197 Y202 L203 S204 Q205 K208 N209 A210 D211 R214 A215 C216 D217 R218 L219 E220 D221 A222 V223 A224 A225 M226 E227 A228 E229 L230 D231 T232 H233 I234

R235 ALA ASP THR GLY ALA ALA GLU ALA GLU ALA GLU ALA GLY ALA ALA GLU GLN GLN GLY PRO GLY ALA ALA GLY GLU ALA ALA GLU ALA ALA GLY ALA ASP LYS PRO ALA ALA ALA ASP GLY SER ASP THR ASP MET ALA ASP GLY ALA ALA ALA ALA GLU GLY ALA ALA ALA ASN

GLY THR HIS ALA GLY ALA GLU SER

- Molecule 27: CWF21 domain-containing protein

Chain U:

96%

M1 Y2 N3 G4 I5 G6 L7 V8 T9 P10 R11 G12 S13 G14 T15 V19 K23 F24 N25 M26 ARG GLY PRO PRO GLN VAL ARG THR ASP ALA PRO LYS ASP GLY PRO ARG ARG VAL LYS GLN PRO ASN GLN GLU ILE LEU ASP HIS LYS LYS ARG ASP ILE GLU LEU LYS VAL

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

LEU LYS ALA ALA LEU GLY LEU LYS GLU ASP VAL ARG ALA GLU GLY LYS ASP ALA PHE ASN ARG GLU GLU LEU GLN GLU LYS ILE GLU LYS ASP ARG ILE ALA GLU ARG GLU LYS ARG GLU VAL LYS ARG GLU ASP ARG GLU ASP ALA ASP ARG GLU LYS ARG GLU GLU GLU LYS GLN ALA ALA ARG ARG GLU GLN GLU ARG ALA ARG

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

ALA GLY ARG ASP ASP ARG ASP SER GLY ARG GLY ASP ALA ALA GLU GLU ARG GLY ARG SER LYS ASP ALA LYS LYS GLY PRO LYS ASP GLY LYS LYS GLU GLU LYS GLY LYS LYS GLU GLU LYS ARG ARG ARG ARG SER PRO SER SER SER SER SER GLY SER SER SER GLY

• Molecule 28: MI domain-containing protein

Chain V:

16%

20%   •   80%

T650 N651 L652 V653 N654 L655 R656 R657 T658 I659

D668 F669 A672 G673 G680 I681 Q682 P683 G684 Q685 E686 I687 E688 L689 V690 T691 M692 I693 I694 E695 C696 C697 S698 Q699 E700 R701 T702 F703 L710 A711 Q712 R713 F714 C715 Y716 L717 N718 R719 A720 Y721 A722 E723 Q724 F725 E726 D727 C728 F729 R730 K731 Q732 Y733 A734 V735 I736 H737 R738

L739 E740 T741 N742 K743 L744 R745 N746 V747 S748 C749 L750 F751 A752 H753 L754 L755 A756 T757 D758 A759 L760 P761 W762 S763 S764 L765 S766 C767 I768 M769 L770 T771 E772 E773 E774 T775 T776 S777 S778 S779 R780 I781 F782 I783 K784 Y785 L786 F787 Q788 E789 L790 S791 S792 T793 M794 G795 L796 V797 K798

L799 N800 Q801 R802 L803 N804 D805 P806 A807 F808 S809 P810 W811 F812 E813 G814 L815 F816 P817 R818 D819 S820 L821 G822 N823 M824 R825 F826 S827 I828 N829 F830 F831 T832 S833 I834 G835 LEU GLY GLY VAL THR ASP VAL MET ARG GLU GLN TYR LYS SER ALA VAL ALA GLN ALA ALA ALA GLN

GLN GLN GLN VAL ALA LYS ARG MET LEU GLY SER GLY SER GLY SER ASP GLY SER SER SER SER ASP SER SER SER ASP SER SER SER ASP ASP SER SER SER ASP SER SER SER ASP ARG GLY ARG ARG SER LYS GLY LYS GLY LYS GLY LYS LYS SER

ASP LYS LYS LYS SER SER LYS ARG ARG

• Molecule 29: Intron-binding protein aquarius

Chain Q:
70%
70%   •   30%

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

SER GLY PRO MET GLY ARG GLY GLN GLY ARG GLY PHE GLY ARG GLY GLN GLN GLN PHE MET GLY PRO MET GLY ALA GLY ARG PHE GLY TRP PRO ARG GLN PHE MET PRO GLN GLN GLN GLN PRO ARG GLN ARG PRO ARG GLY GLY LEU THR LEU THR GLU ILE A118 E119 D120

K121 L122 T123 K124 T125 A126 K127 Q128 E129 W130 S131 N132 E133 A134 R135 A136 Q137 E138 N139 P140 P141 A142 F143 K144 P145 E146 L147 V148 Q149 T150 I151 Y152 K153 Q154 E155 L156 G157 G158 A159 S160 K161 R162 A163 P164 G165 H166 K167 R168 I169 M170 L171 L172 E173 I174 S175 Q176 Y177 L178 E179 N180

Y181 L182 W183 P184 N185 F186 D187 V188 E189 K190 A191 T192 F193 E194 H195 V196 M197 S198 L199 V200 L201 M202 V203 N204 E205 K206 F207 R208 E209 G210 V211 P212 A213 W214 T215 C216 F217 H218 T219 R220 E221 D222 A223 F224 P225 G226 F227 E228 K229 R230 V231 L232 S233 L234 K235 D236 G237 R238 E239 E240

E241 L242 K243 L244 H245 E246 R247 T248 A249 Y250 V251 L252 F253 M254 I255 R256 S257 F258 Q259 S260 L261 E262 D263 E264 M265 V266 R267 A268 Q269 V270 L271 R272 L273 V274 S275 L276 P277 L278 W279 H280 A281 L282 S283 P284 G285 R286 L287 Q288 L289 E290 L291 H292 A293 H294 E295 A296 L297 A298 K299 H300

W301 K302 A303 A304 A305 K306 K307 E308 A309 K310 ALA ALA LYS LYS ALA LYS LYS GLY GLU GLY GLY GLU GLY GLY SER TYR VAL PRO VAL THR GLN ARG PRO E336 T337 R338 F339 L340 P341 A342 L343 M344 D345 E346 F347 L348 A349 V350 L351 D352 Q353 V354 V355 V356 P357 PRO LYS LEU

PRO GLY CYS TYR P365 S366 L367 N368 R369 G370 A371 L372 L373 Y374 C375 E376 R377 F378 L379 E380 F381 L382 I383 D384 L385 L386 S387 Q388 L389 P390 T391 R392 R393 F394 V395 R396 T397 L398 I399 D400 D401 R402 Q403 L404 L405 V406 K407 V408 R409 M410 S411 G412 L413 F414 K415 Y416 E417 VAL PRO PRO

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

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

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

GLY ARG GLY ARG GLY ARG GLY ARG GLY GLN GLY GLN GLY ARG GLY ALA SER ALA GLY ARG ARG ARG GLY ARG SER ALA ALA GLU GLY GLY ASP GLY SER GLY ARG GLY GLY PRO ALA PRO PRO PRO SER ASN | L642 L643 Y644 R645 Q646 L647 V648 D649 L650 F651 S652 Y653 Y654 M655 S656 F657 P658 I659 N660

D661 H662 T663 G664 E665 P666 L667 T668 D669 D670 E671 V672 N673 A674 A675 H676 Y677 E678 K679 V680 C681 Q682 F683 Q684 R685 L686 C687 F688 K689 H690 W691 Q692 G693 V694 E695 A696 M697 Q698 E699 L700 A701 L702 S703 H704 C705 G706 A707 V708 E709 A710 R711 D712 T713 L714 R715 R716 H717 L718 A719 S720

L721 T722 G723 E724 Q725 V726 R727 E728 L729 V730 C731 R732 Q733 L734 R735 L736 V737 G738 E739 D740 D741 P742 W743 A744 A745 D746 G747 A748 F749 L750 L751 E752 V753 L754 L755 A756 A757 Y758 E759 R760 R761 R762 S763 Q764 R765 E766 V767 V768 N769 E770 M771 P772 L773 Y774 P775 T776 E777 G778 L779 L780

W781 D782 E783 S784 Q785 I786 P787 A788 S789 S790 E791 H792 Y793 T794 G795 E796 G797 A798 L799 A800 L801 P802 K803 L804 N805 L806 Q807 F808 L809 T810 V811 A812 D813 Y814 L815 L816 R817 S818 F819 H820 L821 F822 R823 L824 E825 A826 T827 Y828 E829 V830 R831 E832 D833 L834 A835 D836 V837 L838 G839 R840

V841 G842 A843 Y844 T845 G846 ASP ASP G849 R850 T851 R852 F853 A854 G855 W856 A857 R858 M859 A860 L861 P862 L863 T864 S865 F866 K867 V868 T869 E870 V871 R872 K873 P874 N875 V876 G877 E878 A879 K880 P881 A882 G883 V884 T885 A886 N887 V888 V889 I890 D891 T892 R893 P894 L895 R896 G897 D898 V899 R900

S901 E902 W903 D904 E905 L906 K907 Q908 H909 D910 V911 L912 F913 L914 L915 T916 I917 R918 P919 P920 D921 ALA HIS ALA ALA ALA ALA HIS ALA ASP GLY ARG THR PRO ASN P937 A938 E939 K940 F941 G942 L943 V944 Y945 V946 R947 G948 C949 E950 V951 I952 E953 L954 R955 D956 E957 G958 G959 K960

L961 M962 ASN ASP PHE THR GLY ARG VAL ARG ARG ASP GLU VAL LYS PRO PRO GLU | G979 T980 A981 R982 T983 V984 T985 V986 A987 L988 D989 T990 A991 Q992 Y993 Q994 I995 D996 M997 N998 T999 M1000 A1001 R1002 H1003 K1004 S1005 E1006 D1007 P1008 Y1009 A1010 T1011 F1012 N1013 L1014 L1015 M1016 R1017 R1018 K1019 P1020

K1021 E1022 N1023 N1024 F1025 K1026 A1027 V1028 L1029 E1030 S1031 I1032 R1033 D1034 L1035 M1036 N1037 D1038 D1039 T1040 A1041 V1042 I1043 P1044 P1045 W1046 L1047 H1048 D1049 V1050 F1051 L1052 G1053 Y1054 G1055 D1056 P1057 A1058 A1059 A1060 Q1061 ALA ARG VAL PRO GLY THR ALA ALA ALA ALA ALA ALA ALA GLU GLY GLY THR

GLY GLY GLU GLY VAL GLY GLY GLN | A1089 P1090 L1091 R1092 T1093 V1094 D1095 F1096 G1097 D1098 T1099 F1100 L1101 D1102 A1103 Q1104 H1105 V1106 V1107 E1108 A1109 F1110 P1111 Q1112 F1113 K1114 V1115 S1116 F1117 V1118 N1119 K1120 S1121 G1122 K1123 A1124 V1125 P1126 A1127 P1128 P1129 F1130 R1131 I1132 T1133 F1134 P1135 T1136 GLN PRO LEU PRO

GLU ASP ALA SER THR SER GLY ALA ASP ALA ALA ALA ALA ALA GLU ALA ALA ALA GLY SER PHE PRO ALA PRO ALA ASP PRO ALA ASN PRO | A1173 A1174 G1175 E1176 L1177 V1178 V1179 E1180 A1181 Y1182 V1183 P1184 P1185 D1186 P1187 G1188 P1189 Y1190 P1191 Q1192 D1193 Q1194 P1195 R1196 R1197 M1198 A1199 V1200

R1201 F1202 T1203 P1204 V1205 Q1206 V1207 E1208 A1209 I1210 A1211 S1212 G1213 V1214 Q1215 P1216 G1217 L1218 T1219 M1220 V1221 V1222 G1223 P1224 P1225 G1226 T1227 G1228 K1229 T1230 D1231 T1232 A1233 V1234 Q1235 V1236 M1237 T1238 C1239 L1240 Y1241 H1242 N1243 C1244 P1245 G1246 Q1247 R1248 T1249 L1250 L1251 I1252 T1253 H1254 S1255 N1256 Q1257 A1258 L1259 N1260

D1261 L1262 F1263 S1264 K1265 I1266 M1267 E1268 R1269 D1270 V1271 P1272 E1273 R1274 Y1275 L1276 L1277 R1278 L1279 G1280 M1281 GLY GLU | A1284 E1285 L1286 D1287 T1288 E1289 Q1290 D1291 F1292 S1293 R1294 V1295 G1296 R1297 V1298 N1299 A1300 M1301 L1302 A1303 R1304 R1305 L1306 E1307 L1308 L1309 A1310 E1311 V1312 E1313 K1314 M1315 A1316 R1317 Q1318 L1319 G1320

• Molecule 30: RNA (5'-R(P*CP*CP*GP*AP*AP*CP*G)-3')

Chain 5: 71% 14% 14%

C-7 C-6 C-2 G-1

• Molecule 31: RNA (173-mer)

Chain 3: 6% 8% 9% 80%

GGUUGAGGACAAACUCUUCGCGGUCUUUCCAGUGGGGAUCCCGAUAUCCCG

4   Experimental information ⓘ

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	518369	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{Å}^2)$	50.00	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	1500	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV (4k x 4k)	Depositor
Maximum map value	2.593	Depositor
Minimum map value	-1.565	Depositor
Average map value	0.002	Depositor
Map value standard deviation	0.038	Depositor
Recommended contour level	0.08	Depositor
Map size (Å)	404.80002, 404.80002, 404.80002	wwPDB
Map dimensions	440, 440, 440	wwPDB
Map angles (°)	90.0, 90.0, 90.0	wwPDB
Pixel spacing (Å)	0.92, 0.92, 0.92	Depositor

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GTP, MG, SEP, G5J, ZN

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$ >5	RMSZ	#$|Z|$ >5
1	A	0.86	19/15581 (0.1%)	0.67	7/21167 (0.0%)
2	B	1.22	6/2024 (0.3%)	1.16	5/3144 (0.2%)
3	C	0.41	0/7126	0.57	1/9683 (0.0%)
4	E	0.40	0/2496	0.58	1/3382 (0.0%)
5	F	1.60	43/2188 (2.0%)	1.31	17/3409 (0.5%)
6	q	0.30	0/847	0.57	0/1148
6	r	0.31	0/1040	0.58	0/1411
6	s	0.30	0/580	0.59	0/784
6	t	0.33	0/514	0.62	0/698
7	a	0.30	0/716	0.53	0/960
8	g	0.31	0/516	0.56	0/691
9	e	0.28	0/667	0.52	0/891
10	f	0.29	0/567	0.56	0/764
11	d	0.28	0/821	0.50	0/1097
12	c	0.29	0/691	0.61	0/933
13	b	0.30	0/792	0.60	0/1059
14	I	0.28	0/4712	0.48	0/6427
15	J	0.65	0/4092	0.52	1/5574 (0.0%)
16	P	0.74	1/882 (0.1%)	0.60	0/1173
17	M	0.39	0/654	0.56	0/874
18	T	1.14	5/2773 (0.2%)	0.81	0/3765
19	O	0.54	0/2179	0.59	0/2944
20	N	0.82	0/1446	0.64	2/1944 (0.1%)
21	R	0.64	0/2020	0.59	0/2736
22	H	1.25	4/932 (0.4%)	1.29	4/1447 (0.3%)
23	S	0.43	0/1224	0.56	0/1652
24	W	0.44	2/2852 (0.1%)	0.57	2/3908 (0.1%)
25	L	0.46	0/3847	0.57	1/5158 (0.0%)
26	K	0.29	0/1717	0.55	3/2320 (0.1%)
27	U	0.32	0/194	0.60	0/259
28	V	0.28	0/1023	0.42	0/1413
29	Q	0.28	0/6411	0.48	1/8913 (0.0%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
30	5	1.29	0/167	1.52	4/258 (1.6%)
31	3	0.55	0/771	1.19	4/1193 (0.3%)
All	All	0.69	80/75062 (0.1%)	0.68	53/103179 (0.1%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	11
3	C	0	1
6	s	0	1
6	t	0	1
25	L	0	1
28	V	0	1
All	All	0	16

All (80) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
24	W	89	PRO	N-CA	13.08	1.69	1.47
5	F	58	G	C8-N7	-7.87	1.26	1.30
5	F	58	G	C5-C4	-7.64	1.33	1.38
5	F	60	G	C8-N7	-7.22	1.26	1.30
5	F	57	U	C2-N3	-7.16	1.32	1.37
5	F	65	G	C5-C4	-7.13	1.33	1.38
5	F	47	G	C2-N3	-7.11	1.27	1.32
5	F	60	G	C6-N1	-6.90	1.34	1.39
5	F	58	G	N9-C8	-6.88	1.33	1.37
22	H	21	G	C6-N1	-6.84	1.34	1.39
5	F	62	A	N7-C5	-6.76	1.35	1.39
1	A	1092	TYR	CD2-CE2	-6.74	1.29	1.39
5	F	68	G	C8-N7	-6.66	1.26	1.30
2	B	56	U	C2-N3	-6.52	1.33	1.37
5	F	62	A	C8-N7	-6.50	1.26	1.31
5	F	63	A	C6-N1	-6.47	1.31	1.35
5	F	58	G	C6-O6	-6.38	1.18	1.24
1	A	1083	TYR	CD1-CE1	-6.29	1.29	1.39
1	A	190	GLU	CB-CG	-6.17	1.40	1.52
5	F	49	A	C5-C4	-6.11	1.34	1.38
2	B	34	C	N3-C4	-6.05	1.29	1.33

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	F	63	A	N9-C8	-6.03	1.32	1.37
1	A	190	GLU	CG-CD	-5.96	1.43	1.51
5	F	65	G	N1-C2	-5.89	1.33	1.37
5	F	60	G	N1-C2	-5.88	1.33	1.37
18	T	426	ASN	CB-CG	-5.86	1.37	1.51
24	W	88	GLY	C-N	5.86	1.45	1.34
5	F	60	G	N7-C5	-5.84	1.35	1.39
5	F	65	G	N9-C4	-5.80	1.33	1.38
5	F	64	G	C5-C4	-5.79	1.34	1.38
5	F	58	G	N1-C2	-5.75	1.33	1.37
1	A	1083	TYR	CD2-CE2	-5.71	1.30	1.39
1	A	154	VAL	CB-CG1	-5.70	1.40	1.52
2	B	54	G	C2-N3	-5.67	1.28	1.32
5	F	62	A	C5-C6	-5.67	1.35	1.41
2	B	51	A	C6-N1	-5.63	1.31	1.35
1	A	957	GLU	CB-CG	-5.60	1.41	1.52
5	F	64	G	C2-N3	-5.55	1.28	1.32
5	F	59	C	N3-C4	-5.51	1.30	1.33
18	T	485	GLU	CG-CD	-5.49	1.43	1.51
1	A	179	TYR	CD1-CE1	-5.49	1.31	1.39
16	P	233	TYR	CD2-CE2	-5.43	1.31	1.39
1	A	858	TYR	CD1-CE1	-5.39	1.31	1.39
5	F	47	G	C5-C4	-5.39	1.34	1.38
5	F	64	G	C8-N7	-5.38	1.27	1.30
5	F	56	C	N3-C4	-5.36	1.30	1.33
5	F	62	A	C5-C4	-5.36	1.34	1.38
18	T	202	VAL	CB-CG2	-5.36	1.41	1.52
1	A	742	GLU	CB-CG	-5.36	1.42	1.52
22	H	22	C	N3-C4	-5.34	1.30	1.33
5	F	48	C	N3-C4	-5.32	1.30	1.33
22	H	21	G	C2-N3	-5.31	1.28	1.32
1	A	283	TYR	CD1-CE1	-5.25	1.31	1.39
1	A	644	TYR	CD2-CE2	-5.23	1.31	1.39
22	H	25	A	N7-C5	-5.23	1.36	1.39
5	F	68	G	C5-C4	-5.22	1.34	1.38
5	F	59	C	N1-C6	-5.20	1.34	1.37
1	A	237	TYR	CD1-CE1	-5.20	1.31	1.39
1	A	179	TYR	CD2-CE2	-5.20	1.31	1.39
5	F	60	G	N9-C4	-5.19	1.33	1.38
1	A	1044	GLU	CB-CG	-5.18	1.42	1.52
5	F	60	G	C5-C4	-5.17	1.34	1.38
18	T	283	VAL	CB-CG2	-5.17	1.42	1.52

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
18	T	495	GLU	CB-CG	-5.16	1.42	1.52
1	A	169	TRP	CE3-CZ3	-5.15	1.29	1.38
2	B	55	A	N9-C4	-5.14	1.34	1.37
5	F	60	G	N9-C8	-5.14	1.34	1.37
1	A	973	TYR	CD2-CE2	-5.13	1.31	1.39
5	F	57	U	N3-C4	-5.13	1.33	1.38
5	F	64	G	C6-N1	-5.13	1.35	1.39
1	A	1340	GLU	CG-CD	-5.11	1.44	1.51
5	F	63	A	C8-N7	-5.08	1.27	1.31
2	B	52	G	C6-N1	-5.07	1.36	1.39
5	F	17	G	C8-N7	-5.04	1.27	1.30
5	F	47	G	N9-C4	-5.04	1.33	1.38
5	F	56	C	N1-C6	-5.03	1.34	1.37
5	F	55	C	N3-C4	-5.03	1.30	1.33
1	A	634	ASP	CB-CG	-5.01	1.41	1.51
5	F	59	C	C2-N3	-5.00	1.31	1.35
5	F	69	A	C5-C4	-5.00	1.35	1.38

All (53) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
22	H	22	C	O5'-P-OP1	-18.02	89.07	110.70
20	N	46	LEU	CA-CB-CG	7.61	132.81	115.30
1	A	843	LEU	CA-CB-CG	7.12	131.67	115.30
5	F	39	G	C5-C6-O6	-6.80	124.52	128.60
5	F	68	G	N9-C4-C5	-6.78	102.69	105.40
24	W	89	PRO	CA-N-CD	-6.59	102.28	111.50
4	E	357	LEU	CA-CB-CG	6.53	130.32	115.30
1	A	711	LEU	CA-CB-CG	6.43	130.09	115.30
5	F	39	G	N1-C6-O6	6.41	123.74	119.90
5	F	52	G	C4-C5-N7	6.39	113.36	110.80
20	N	169	LEU	CA-CB-CG	6.37	129.96	115.30
31	3	20	A	P-O3'-C3'	6.31	127.27	119.70
5	F	65	G	C8-N9-C4	6.21	108.89	106.40
5	F	68	G	C4-C5-N7	6.11	113.24	110.80
5	F	58	G	C8-N9-C4	6.08	108.83	106.40
5	F	58	G	C5-C6-N1	6.05	114.52	111.50
22	H	4	C	N3-C2-O2	-6.03	117.68	121.90
25	L	740	LEU	CA-CB-CG	5.99	129.08	115.30
1	A	303	LEU	CA-CB-CG	5.98	129.05	115.30
2	B	70	U	C2-N1-C1'	5.96	124.85	117.70
5	F	46	A	N1-C6-N6	-5.83	115.10	118.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
26	K	62	PRO	N-CA-CB	5.80	110.26	103.30
1	A	1156	LEU	CA-CB-CG	5.65	128.29	115.30
29	Q	788	ALA	N-CA-C	-5.64	95.76	111.00
30	5	-6	C	C2-N1-C1'	5.59	124.95	118.80
2	B	29	U	N3-C2-O2	-5.52	118.34	122.20
5	F	7	C	C2-N1-C1'	5.52	124.87	118.80
30	5	-2	C	C2-N3-C4	-5.51	117.14	119.90
1	A	749	LEU	CA-CB-CG	5.51	127.96	115.30
30	5	-6	C	N1-C2-O2	5.47	122.18	118.90
3	C	415	GLY	N-CA-C	-5.43	99.53	113.10
22	H	21	G	N7-C8-N9	5.43	115.81	113.10
5	F	54	C	N1-C2-O2	-5.35	115.69	118.90
5	F	47	G	N3-C4-C5	5.33	131.26	128.60
24	W	91	HIS	CB-CA-C	5.31	121.01	110.40
2	B	29	U	O4'-C1'-N1	5.25	112.40	108.20
15	J	21	ASP	CB-CG-OD2	5.25	123.03	118.30
5	F	58	G	C4-C5-C6	-5.24	115.66	118.80
5	F	26	G	N3-C4-N9	-5.20	122.88	126.00
5	F	19	U	C2-N1-C1'	5.19	123.93	117.70
1	A	1004	VAL	C-N-CA	-5.18	108.75	121.70
31	3	8	C	C5-C6-N1	5.18	123.59	121.00
5	F	63	A	N1-C6-N6	-5.17	115.50	118.60
5	F	47	G	C8-N9-C4	5.16	108.46	106.40
31	3	13	C	C6-N1-C2	-5.15	118.24	120.30
22	H	21	G	C8-N9-C4	-5.13	104.35	106.40
2	B	56	U	C2-N3-C4	-5.13	123.92	127.00
30	5	-2	C	N3-C2-O2	-5.12	118.31	121.90
1	A	1154	ARG	NE-CZ-NH2	-5.10	117.75	120.30
26	K	106	LEU	CA-CB-CG	5.09	127.02	115.30
31	3	16	U	C2-N1-C1'	5.09	123.81	117.70
2	B	49	A	O4'-C1'-N9	5.07	112.25	108.20
26	K	219	LEU	CA-CB-CG	5.00	126.81	115.30

There are no chirality outliers.

All (16) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1004	VAL	Peptide
1	A	1005	LYS	Peptide
1	A	1045	PHE	Peptide
1	A	1338	TYR	Peptide
1	A	1470	ASP	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	1480	ILE	Peptide
1	A	1701	TRP	Peptide
1	A	1751	TYR	Peptide
1	A	1783	PHE	Peptide
1	A	532	LYS	Peptide
1	A	686	GLY	Peptide
3	C	69	GLU	Peptide
25	L	670	TRP	Peptide
28	V	703	PHE	Peptide
6	s	67	THR	Peptide
6	t	67	THR	Peptide

5.2   Too-close contacts ⓘ

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

5.3   Torsion angles ⓘ

5.3.1   Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	1923/2398 (80%)	1730 (90%)	182 (10%)	11 (1%)	22	43
3	C	895/989 (90%)	816 (91%)	78 (9%)	1 (0%)	48	71
4	E	316/362 (87%)	289 (92%)	26 (8%)	1 (0%)	37	59
6	q	102/503 (20%)	102 (100%)	0	0	100	100
6	r	129/503 (26%)	120 (93%)	8 (6%)	1 (1%)	16	34
6	s	68/503 (14%)	65 (96%)	3 (4%)	0	100	100
6	t	60/503 (12%)	56 (93%)	4 (7%)	0	100	100
7	a	88/127 (69%)	86 (98%)	2 (2%)	0	100	100
8	g	63/77 (82%)	60 (95%)	3 (5%)	0	100	100
9	e	78/87 (90%)	74 (95%)	4 (5%)	0	100	100

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
10	f	69/85 (81%)	68 (99%)	1 (1%)	0	100	100
11	d	98/110 (89%)	97 (99%)	1 (1%)	0	100	100
12	c	86/114 (75%)	82 (95%)	4 (5%)	0	100	100
13	b	95/265 (36%)	89 (94%)	6 (6%)	0	100	100
14	I	657/925 (71%)	641 (98%)	16 (2%)	0	100	100
15	J	587/835 (70%)	551 (94%)	30 (5%)	6 (1%)	13	29
16	P	98/235 (42%)	90 (92%)	8 (8%)	0	100	100
17	M	74/563 (13%)	68 (92%)	6 (8%)	0	100	100
18	T	342/518 (66%)	326 (95%)	15 (4%)	1 (0%)	37	59
19	O	269/417 (64%)	245 (91%)	24 (9%)	0	100	100
20	N	167/233 (72%)	154 (92%)	13 (8%)	0	100	100
21	R	254/684 (37%)	240 (94%)	14 (6%)	0	100	100
23	S	153/157 (98%)	137 (90%)	16 (10%)	0	100	100
24	W	475/576 (82%)	431 (91%)	38 (8%)	6 (1%)	10	21
25	L	454/833 (54%)	429 (94%)	25 (6%)	0	100	100
26	K	208/303 (69%)	191 (92%)	16 (8%)	1 (0%)	25	47
27	U	24/721 (3%)	23 (96%)	1 (4%)	0	100	100
28	V	184/928 (20%)	177 (96%)	7 (4%)	0	100	100
29	Q	1275/1844 (69%)	1217 (96%)	49 (4%)	9 (1%)	19	38
All	All	9291/16398 (57%)	8654 (93%)	600 (6%)	37 (0%)	32	52

All (37) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	389	VAL
1	A	449	GLU
15	J	26	ALA
15	J	27	GLU
29	Q	893	ARG
1	A	246	PRO
3	C	422	THR
6	r	58	THR
15	J	37	GLN
15	J	40	ASP
15	J	53	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
26	K	62	PRO
29	Q	155	GLU
29	Q	158	GLY
29	Q	719	ALA
29	Q	789	SER
1	A	247	LEU
1	A	528	PRO
4	E	47	GLY
24	W	86	VAL
29	Q	154	GLN
1	A	1266	MET
1	A	1717	ASP
1	A	2011	PRO
29	Q	153	LYS
1	A	1426	ASP
1	A	1718	GLN
24	W	87	LEU
24	W	92	PRO
24	W	91	HIS
29	Q	898	ASP
29	Q	1365	VAL
1	A	1822	PRO
15	J	52	PRO
24	W	84	ALA
24	W	198	GLN
18	T	247	GLY

5.3.2　Protein sidechains　ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	1505/2124 (71%)	1306 (87%)	199 (13%)	3	6
3	C	753/825 (91%)	647 (86%)	106 (14%)	3	5
4	E	264/300 (88%)	231 (88%)	33 (12%)	3	7
6	q	94/398 (24%)	79 (84%)	15 (16%)	2	3

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles		
6	r	113/398 (28%)	93 (82%)	20 (18%)	1	2	
6	s	59/398 (15%)	54 (92%)	5 (8%)	8	18	
6	t	54/398 (14%)	45 (83%)	9 (17%)	2	3	
7	a	76/97 (78%)	66 (87%)	10 (13%)	3	6	
8	g	58/67 (87%)	49 (84%)	9 (16%)	2	3	
9	e	73/79 (92%)	63 (86%)	10 (14%)	3	5	
10	f	62/75 (83%)	54 (87%)	8 (13%)	3	6	
11	d	92/100 (92%)	80 (87%)	12 (13%)	3	6	
12	c	78/95 (82%)	67 (86%)	11 (14%)	3	5	
13	b	84/211 (40%)	76 (90%)	8 (10%)	7	14	
14	I	357/743 (48%)	308 (86%)	49 (14%)	3	5	
15	J	260/676 (38%)	228 (88%)	32 (12%)	4	8	
16	P	88/188 (47%)	77 (88%)	11 (12%)	3	7	
17	M	63/424 (15%)	56 (89%)	7 (11%)	5	10	
18	T	291/416 (70%)	263 (90%)	28 (10%)	7	14	
19	O	228/318 (72%)	210 (92%)	18 (8%)	10	21	
20	N	148/193 (77%)	137 (93%)	11 (7%)	11	24	
21	R	209/507 (41%)	179 (86%)	30 (14%)	2	5	
23	S	128/130 (98%)	119 (93%)	9 (7%)	12	27	
24	W	118/474 (25%)	95 (80%)	23 (20%)	1	2	
25	L	388/648 (60%)	333 (86%)	55 (14%)	2	5	
26	K	177/221 (80%)	144 (81%)	33 (19%)	1	2	
27	U	20/572 (4%)	16 (80%)	4 (20%)	1	2	
28	V	28/746 (4%)	24 (86%)	4 (14%)	2	5	
All	All	5868/11821 (50%)	5099 (87%)	769 (13%)	6	6	

All (769) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	85	GLN
1	A	95	LYS
1	A	96	TRP
1	A	108	ARG
1	A	114	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	127	LYS
1	A	130	ARG
1	A	137	SER
1	A	145	ARG
1	A	146	VAL
1	A	148	LEU
1	A	167	MET
1	A	175	VAL
1	A	177	VAL
1	A	181	ILE
1	A	185	ILE
1	A	229	ASP
1	A	276	LYS
1	A	279	ASN
1	A	288	LEU
1	A	296	LEU
1	A	299	LEU
1	A	303	LEU
1	A	305	THR
1	A	310	ARG
1	A	311	ASN
1	A	321	PHE
1	A	331	ILE
1	A	342	ARG
1	A	368	THR
1	A	380	ASN
1	A	385	VAL
1	A	386	ARG
1	A	387	LEU
1	A	391	HIS
1	A	396	LEU
1	A	478	ARG
1	A	489	PRO
1	A	492	ASN
1	A	493	CYS
1	A	498	HIS
1	A	502	THR
1	A	503	TYR
1	A	522	LEU
1	A	529	LYS
1	A	530	SER
1	A	540	LEU

Continued from previous page...

Mol	Chain	Res	Type
1	A	547	GLN
1	A	548	SER
1	A	549	THR
1	A	587	ASN
1	A	593	THR
1	A	595	THR
1	A	596	THR
1	A	599	ARG
1	A	610	LEU
1	A	618	THR
1	A	619	LYS
1	A	633	VAL
1	A	634	ASP
1	A	640	ASP
1	A	643	GLN
1	A	675	LEU
1	A	691	VAL
1	A	706	ARG
1	A	708	ILE
1	A	712	LEU
1	A	714	ARG
1	A	729	SER
1	A	732	ILE
1	A	738	LYS
1	A	756	ASP
1	A	800	GLU
1	A	809	GLN
1	A	834	THR
1	A	866	THR
1	A	884	ARG
1	A	899	ASP
1	A	909	ARG
1	A	923	SER
1	A	935	TYR
1	A	946	LYS
1	A	952	GLN
1	A	954	SER
1	A	978	LEU
1	A	980	LYS
1	A	984	CYS
1	A	986	LEU
1	A	987	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	993	GLU
1	A	999	LEU
1	A	1004	VAL
1	A	1005	LYS
1	A	1015	LEU
1	A	1024	ASN
1	A	1027	THR
1	A	1042	GLN
1	A	1043	SER
1	A	1059	ARG
1	A	1085	ASP
1	A	1087	SER
1	A	1091	SER
1	A	1094	LEU
1	A	1120	THR
1	A	1133	ASN
1	A	1142	LYS
1	A	1153	SER
1	A	1157	ASN
1	A	1163	PHE
1	A	1164	ARG
1	A	1188	ASN
1	A	1196	LYS
1	A	1201	ASP
1	A	1213	LEU
1	A	1224	ARG
1	A	1230	THR
1	A	1231	THR
1	A	1240	SER
1	A	1245	ASP
1	A	1246	ASN
1	A	1248	ASN
1	A	1265	ARG
1	A	1283	VAL
1	A	1285	THR
1	A	1289	THR
1	A	1319	THR
1	A	1321	THR
1	A	1322	LYS
1	A	1331	LEU
1	A	1336	THR
1	A	1340	GLU

Continued from previous page...

Mol	Chain	Res	Type
1	A	1348	LEU
1	A	1362	ARG
1	A	1365	ILE
1	A	1402	SER
1	A	1417	PHE
1	A	1421	MET
1	A	1428	LEU
1	A	1434	ARG
1	A	1441	SER
1	A	1447	GLU
1	A	1452	GLU
1	A	1456	LYS
1	A	1457	LYS
1	A	1459	GLU
1	A	1461	LYS
1	A	1465	ARG
1	A	1468	THR
1	A	1470	ASP
1	A	1472	LEU
1	A	1490	ASP
1	A	1491	ARG
1	A	1493	THR
1	A	1502	VAL
1	A	1512	THR
1	A	1513	LYS
1	A	1514	MET
1	A	1538	VAL
1	A	1545	VAL
1	A	1557	THR
1	A	1558	TYR
1	A	1590	LEU
1	A	1592	GLN
1	A	1593	ILE
1	A	1600	LEU
1	A	1601	TRP
1	A	1605	THR
1	A	1608	ARG
1	A	1630	ILE
1	A	1632	THR
1	A	1633	LEU
1	A	1635	ILE
1	A	1658	CYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1665	LEU
1	A	1679	HIS
1	A	1683	SER
1	A	1700	LYS
1	A	1701	TRP
1	A	1702	ASN
1	A	1705	LYS
1	A	1707	SER
1	A	1712	THR
1	A	1725	TRP
1	A	1726	VAL
1	A	1728	VAL
1	A	1729	GLN
1	A	1732	TRP
1	A	1747	LYS
1	A	1754	ASP
1	A	1756	MET
1	A	1757	SER
1	A	1783	PHE
1	A	1786	ILE
1	A	1807	LEU
1	A	1810	ARG
1	A	1813	LYS
1	A	1816	GLN
1	A	1817	LEU
1	A	1819	SER
3	C	67	LEU
3	C	69	GLU
3	C	76	SER
3	C	92	GLU
3	C	109	LYS
3	C	111	GLU
3	C	112	THR
3	C	113	LEU
3	C	118	LEU
3	C	119	ARG
3	C	127	LEU
3	C	135	GLU
3	C	140	VAL
3	C	150	LYS
3	C	162	HIS
3	C	166	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	C	171	GLN
3	C	176	LYS
3	C	197	MET
3	C	200	SER
3	C	204	GLU
3	C	210	SER
3	C	212	LEU
3	C	218	CYS
3	C	269	MET
3	C	306	CYS
3	C	313	GLU
3	C	319	VAL
3	C	330	SER
3	C	337	GLN
3	C	339	PHE
3	C	355	GLU
3	C	356	PHE
3	C	358	LYS
3	C	365	TYR
3	C	384	ARG
3	C	397	MET
3	C	416	LEU
3	C	419	TYR
3	C	421	ARG
3	C	429	THR
3	C	465	LYS
3	C	482	MET
3	C	485	CYS
3	C	508	ASP
3	C	525	LYS
3	C	527	LEU
3	C	537	GLU
3	C	550	TYR
3	C	560	ARG
3	C	566	LEU
3	C	572	VAL
3	C	585	ASP
3	C	591	VAL
3	C	606	LYS
3	C	610	GLU
3	C	616	GLU
3	C	620	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	C	622	GLU
3	C	634	CYS
3	C	636	THR
3	C	640	GLU
3	C	643	GLU
3	C	649	THR
3	C	654	LEU
3	C	658	MET
3	C	662	ARG
3	C	676	VAL
3	C	690	LYS
3	C	700	ASN
3	C	709	LEU
3	C	710	ASP
3	C	721	ASN
3	C	725	ASP
3	C	729	LYS
3	C	747	ARG
3	C	749	LEU
3	C	763	LEU
3	C	770	GLU
3	C	772	ASN
3	C	773	LYS
3	C	781	ASP
3	C	784	ILE
3	C	785	GLN
3	C	787	PHE
3	C	789	TRP
3	C	809	LEU
3	C	810	ASP
3	C	826	ILE
3	C	839	MET
3	C	843	ARG
3	C	849	TYR
3	C	855	THR
3	C	858	ASP
3	C	865	ASN
3	C	874	VAL
3	C	878	VAL
3	C	895	VAL
3	C	901	GLU
3	C	902	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	C	905	ARG
3	C	932	SER
3	C	936	ARG
3	C	938	LEU
3	C	948	ARG
3	C	961	SER
4	E	53	ASN
4	E	54	LEU
4	E	72	ARG
4	E	76	ASP
4	E	96	ASP
4	E	103	VAL
4	E	104	ILE
4	E	118	THR
4	E	127	SER
4	E	130	LYS
4	E	157	CYS
4	E	181	ARG
4	E	184	LYS
4	E	200	PHE
4	E	202	ASP
4	E	212	VAL
4	E	214	ASN
4	E	223	ARG
4	E	228	ASP
4	E	232	SER
4	E	243	THR
4	E	256	SER
4	E	265	GLU
4	E	266	TRP
4	E	277	CYS
4	E	284	HIS
4	E	288	PHE
4	E	300	ASP
4	E	302	SER
4	E	306	CYS
4	E	311	ARG
4	E	333	SER
4	E	337	VAL
6	q	1	MET
6	q	3	CYS
6	q	12	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
6	q	18	LYS
6	q	33	ARG
6	q	46	SER
6	q	48	GLU
6	q	50	LEU
6	q	72	LEU
6	q	78	ASP
6	q	92	THR
6	q	97	THR
6	q	113	ARG
6	q	114	VAL
6	q	117	ARG
6	r	2	PHE
6	r	6	SER
6	r	12	GLN
6	r	14	VAL
6	r	31	TYR
6	r	41	THR
6	r	45	LEU
6	r	55	VAL
6	r	57	LYS
6	r	59	VAL
6	r	63	THR
6	r	67	THR
6	r	88	HIS
6	r	95	HIS
6	r	106	TYR
6	r	113	ARG
6	r	115	ILE
6	r	119	LEU
6	r	120	LYS
6	r	129	LEU
6	s	74	SER
6	s	84	MET
6	s	85	LEU
6	s	113	ARG
6	s	131	ASP
6	t	74	SER
6	t	75	LEU
6	t	80	TRP
6	t	88	HIS
6	t	105	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
6	t	108	HIS
6	t	117	ARG
6	t	119	LEU
6	t	123	ASP
7	a	19	GLU
7	a	34	ARG
7	a	47	GLN
7	a	57	ASP
7	a	59	ARG
7	a	61	THR
7	a	66	LEU
7	a	68	ILE
7	a	89	LYS
7	a	90	ARG
8	g	21	LEU
8	g	24	ASN
8	g	27	VAL
8	g	38	MET
8	g	53	ASP
8	g	56	MET
8	g	60	ARG
8	g	65	GLN
8	g	70	LEU
9	e	9	MET
9	e	24	GLN
9	e	35	ASP
9	e	50	ASN
9	e	55	GLU
9	e	61	MET
9	e	63	ARG
9	e	77	ASP
9	e	82	ILE
9	e	85	THR
10	f	12	PHE
10	f	23	VAL
10	f	42	MET
10	f	48	SER
10	f	63	GLU
10	f	68	CYS
10	f	69	ASN
10	f	74	LEU
11	d	17	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
11	d	22	THR
11	d	42	CYS
11	d	44	ASN
11	d	51	ARG
11	d	63	LEU
11	d	68	GLU
11	d	69	MET
11	d	71	THR
11	d	81	LYS
11	d	84	ARG
11	d	98	ARG
12	c	2	VAL
12	c	6	ARG
12	c	21	LYS
12	c	24	THR
12	c	38	ASN
12	c	45	LYS
12	c	49	ARG
12	c	62	ARG
12	c	64	ASN
12	c	75	LEU
12	c	85	ASP
13	b	9	LEU
13	b	20	THR
13	b	25	ARG
13	b	30	ARG
13	b	44	ASP
13	b	55	LYS
13	b	61	GLU
13	b	62	ARG
14	I	347	ASP
14	I	354	GLU
14	I	356	LEU
14	I	365	SER
14	I	367	VAL
14	I	370	ARG
14	I	382	ARG
14	I	384	LYS
14	I	392	LYS
14	I	395	LEU
14	I	402	ARG
14	I	412	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
14	I	416	LEU
14	I	420	PHE
14	I	434	ARG
14	I	438	GLN
14	I	444	ARG
14	I	456	CYS
14	I	466	SER
14	I	473	ASP
14	I	480	THR
14	I	487	ARG
14	I	492	GLU
14	I	500	ASP
14	I	537	ILE
14	I	546	TYR
14	I	548	LEU
14	I	601	LEU
14	I	613	GLU
14	I	614	CYS
14	I	615	LYS
14	I	628	TYR
14	I	640	ARG
14	I	651	LYS
14	I	655	ASP
14	I	659	SER
14	I	670	VAL
14	I	687	SER
14	I	688	ASP
14	I	692	ARG
14	I	702	GLU
14	I	712	ARG
14	I	735	ARG
14	I	737	PHE
14	I	745	ASP
14	I	746	THR
14	I	747	PHE
14	I	752	ARG
14	I	758	LYS
15	J	24	ILE
15	J	38	LEU
15	J	54	GLU
15	J	57	GLU
15	J	61	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
15	J	65	GLN
15	J	71	ARG
15	J	77	ASN
15	J	96	ARG
15	J	126	ARG
15	J	139	VAL
15	J	140	SER
15	J	144	ARG
15	J	167	GLN
15	J	175	PHE
15	J	196	ASP
15	J	202	PHE
15	J	213	LYS
15	J	223	MET
15	J	237	ARG
15	J	242	LEU
15	J	244	GLU
15	J	247	GLN
15	J	248	THR
15	J	249	GLU
15	J	255	PHE
15	J	257	GLU
15	J	273	ARG
15	J	279	ILE
15	J	288	TYR
15	J	311	LYS
15	J	313	ARG
16	P	13	ILE
16	P	40	LYS
16	P	48	GLN
16	P	62	GLU
16	P	68	ARG
16	P	194	ARG
16	P	199	ASP
16	P	204	ASN
16	P	211	LYS
16	P	231	GLN
16	P	234	ILE
17	M	447	GLN
17	M	464	GLU
17	M	484	LYS
17	M	499	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
17	M	531	ARG
17	M	535	PHE
17	M	538	LYS
18	T	190	ARG
18	T	199	MET
18	T	212	CYS
18	T	213	VAL
18	T	216	ASP
18	T	270	ASP
18	T	286	SER
18	T	306	LEU
18	T	312	ASP
18	T	314	VAL
18	T	323	LYS
18	T	353	HIS
18	T	377	SER
18	T	383	MET
18	T	388	PHE
18	T	396	GLU
18	T	414	SER
18	T	446	ARG
18	T	450	CYS
18	T	462	SER
18	T	464	GLU
18	T	465	SER
18	T	469	LEU
18	T	485	GLU
18	T	497	GLU
18	T	498	ASP
18	T	510	ARG
18	T	514	ASP
19	O	24	GLU
19	O	35	MET
19	O	49	ARG
19	O	68	VAL
19	O	70	CYS
19	O	83	CYS
19	O	110	SER
19	O	111	GLU
19	O	151	LEU
19	O	154	SER
19	O	157	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
19	O	164	ARG
19	O	165	VAL
19	O	180	CYS
19	O	206	ASN
19	O	224	THR
19	O	233	THR
19	O	282	LYS
20	N	10	LYS
20	N	21	GLU
20	N	23	ILE
20	N	41	LYS
20	N	72	SER
20	N	82	GLU
20	N	118	SER
20	N	134	THR
20	N	146	SER
20	N	169	LEU
20	N	176	TRP
21	R	39	LEU
21	R	42	LYS
21	R	83	ILE
21	R	88	LYS
21	R	97	VAL
21	R	105	ARG
21	R	110	ASP
21	R	114	LEU
21	R	117	LYS
21	R	118	VAL
21	R	119	ASP
21	R	128	ARG
21	R	146	GLU
21	R	152	LYS
21	R	154	SER
21	R	155	SER
21	R	161	LEU
21	R	168	SER
21	R	174	THR
21	R	195	ASP
21	R	211	VAL
21	R	220	VAL
21	R	223	LEU
21	R	231	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
21	R	233	LYS
21	R	236	GLN
21	R	257	LEU
21	R	259	LYS
21	R	264	ASP
21	R	286	THR
23	S	9	ASP
23	S	28	CYS
23	S	73	LYS
23	S	80	ARG
23	S	85	THR
23	S	91	SER
23	S	133	ARG
23	S	134	LEU
23	S	146	VAL
24	W	47	THR
24	W	50	VAL
24	W	51	LEU
24	W	65	LEU
24	W	66	GLN
24	W	72	LYS
24	W	77	LEU
24	W	81	ASP
24	W	87	LEU
24	W	89	PRO
24	W	91	HIS
24	W	102	LEU
24	W	107	VAL
24	W	112	ASP
24	W	118	ILE
24	W	122	GLN
24	W	134	GLU
24	W	139	ILE
24	W	167	THR
24	W	171	ARG
24	W	202	SER
24	W	212	GLU
24	W	225	LYS
25	L	24	LYS
25	L	35	SER
25	L	38	VAL
25	L	61	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
25	L	70	LEU
25	L	103	LEU
25	L	120	ARG
25	L	122	LEU
25	L	130	ASN
25	L	148	LYS
25	L	152	SER
25	L	161	THR
25	L	171	ARG
25	L	175	LEU
25	L	180	ARG
25	L	183	GLN
25	L	184	LEU
25	L	217	LEU
25	L	220	GLN
25	L	237	GLN
25	L	541	ASP
25	L	543	ARG
25	L	545	ARG
25	L	548	ARG
25	L	552	GLU
25	L	553	ARG
25	L	563	LYS
25	L	571	ARG
25	L	596	LEU
25	L	599	ARG
25	L	614	HIS
25	L	619	TYR
25	L	622	LYS
25	L	654	SER
25	L	660	ASP
25	L	665	TYR
25	L	675	LEU
25	L	696	SER
25	L	699	ARG
25	L	702	ARG
25	L	718	GLU
25	L	720	ASP
25	L	731	LYS
25	L	748	GLN
25	L	749	ARG
25	L	757	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
25	L	779	HIS
25	L	780	GLU
25	L	781	ARG
25	L	800	THR
25	L	816	ARG
25	L	817	ARG
25	L	818	ARG
25	L	828	LYS
25	L	833	SER
26	K	28	HIS
26	K	29	LEU
26	K	55	LYS
26	K	58	SER
26	K	59	THR
26	K	65	TYR
26	K	68	GLU
26	K	69	MET
26	K	77	PHE
26	K	86	GLU
26	K	87	TYR
26	K	94	GLU
26	K	104	TYR
26	K	105	ARG
26	K	114	ARG
26	K	116	ASP
26	K	122	SER
26	K	123	SER
26	K	130	GLN
26	K	131	LEU
26	K	136	LEU
26	K	145	LEU
26	K	149	ASP
26	K	159	ASP
26	K	177	THR
26	K	188	GLN
26	K	197	LYS
26	K	202	TYR
26	K	205	GLN
26	K	208	LYS
26	K	214	ARG
26	K	216	CYS
26	K	226	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
27	U	2	TYR
27	U	15	THR
27	U	19	VAL
27	U	23	LYS
28	V	656	ARG
28	V	658	THR
28	V	668	ASP
28	V	669	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (156) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	100	ASN
1	A	124	HIS
1	A	132	HIS
1	A	269	HIS
1	A	279	ASN
1	A	311	ASN
1	A	379	ASN
1	A	391	HIS
1	A	461	ASN
1	A	481	HIS
1	A	492	ASN
1	A	496	GLN
1	A	512	GLN
1	A	521	ASN
1	A	581	HIS
1	A	625	ASN
1	A	627	GLN
1	A	648	HIS
1	A	674	HIS
1	A	728	GLN
1	A	739	GLN
1	A	744	HIS
1	A	755	HIS
1	A	809	GLN
1	A	852	GLN
1	A	855	GLN
1	A	895	ASN
1	A	939	HIS
1	A	948	HIS
1	A	952	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	968	HIS
1	A	1024	ASN
1	A	1025	ASN
1	A	1033	ASN
1	A	1078	ASN
1	A	1133	ASN
1	A	1147	HIS
1	A	1157	ASN
1	A	1160	HIS
1	A	1175	GLN
1	A	1194	ASN
1	A	1223	ASN
1	A	1246	ASN
1	A	1282	ASN
1	A	1284	HIS
1	A	1291	GLN
1	A	1310	GLN
1	A	1325	ASN
1	A	1346	GLN
1	A	1357	ASN
1	A	1360	GLN
1	A	1401	GLN
1	A	1437	GLN
1	A	1464	ASN
1	A	1504	GLN
1	A	1521	HIS
1	A	1524	HIS
1	A	1616	GLN
1	A	1639	GLN
1	A	1702	ASN
1	A	1801	ASN
3	C	68	HIS
3	C	133	ASN
3	C	147	HIS
3	C	214	ASN
3	C	337	GLN
3	C	388	GLN
3	C	502	GLN
3	C	757	GLN
3	C	760	ASN
3	C	865	ASN
3	C	915	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	C	944	GLN
4	E	100	ASN
4	E	275	ASN
4	E	284	HIS
4	E	330	HIS
6	q	83	ASN
6	q	95	HIS
6	q	99	GLN
6	r	21	HIS
6	r	56	ASN
6	r	108	HIS
6	t	88	HIS
6	t	107	GLN
7	a	16	HIS
7	a	47	GLN
7	a	84	ASN
8	g	65	GLN
9	e	21	GLN
9	e	27	GLN
9	e	50	ASN
9	e	78	ASN
11	d	44	ASN
11	d	58	HIS
11	d	80	GLN
12	c	65	ASN
12	c	76	ASN
14	I	374	HIS
14	I	393	GLN
14	I	427	HIS
14	I	432	ASN
14	I	445	HIS
14	I	453	HIS
14	I	465	HIS
14	I	512	HIS
14	I	545	ASN
14	I	553	GLN
14	I	743	ASN
15	J	154	HIS
15	J	207	GLN
15	J	224	GLN
15	J	247	GLN
16	P	38	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
16	P	48	GLN
17	M	495	ASN
17	M	540	GLN
18	T	261	HIS
18	T	347	GLN
18	T	353	HIS
19	O	45	HIS
19	O	60	ASN
19	O	71	GLN
19	O	119	GLN
19	O	149	GLN
19	O	152	GLN
20	N	29	GLN
20	N	51	HIS
20	N	111	ASN
20	N	119	HIS
20	N	138	GLN
20	N	172	ASN
21	R	79	GLN
21	R	150	GLN
21	R	235	GLN
21	R	236	GLN
21	R	274	ASN
23	S	30	ASN
23	S	98	ASN
23	S	138	GLN
24	W	66	GLN
24	W	91	HIS
24	W	114	HIS
24	W	125	HIS
24	W	195	GLN
24	W	198	GLN
25	L	29	GLN
25	L	130	ASN
25	L	568	GLN
25	L	688	HIS
26	K	26	ASN
26	K	28	HIS
26	K	134	GLN
26	K	187	GLN
26	K	233	HIS
28	V	674	HIS

5.3.3   RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	83/111 (74%)	48 (57%)	2 (2%)
22	H	39/192 (20%)	28 (71%)	3 (7%)
30	5	6/7 (85%)	1 (16%)	0
31	3	33/173 (19%)	16 (48%)	6 (18%)
5	F	90/101 (89%)	48 (53%)	5 (5%)
All	All	251/584 (42%)	141 (56%)	16 (6%)

All (141) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	8	A
2	B	9	C
2	B	10	G
2	B	11	A
2	B	15	G
2	B	16	A
2	B	17	A
2	B	20	C
2	B	27	A
2	B	28	A
2	B	35	U
2	B	39	G
2	B	40	C
2	B	41	C
2	B	42	U
2	B	48	U
2	B	49	A
2	B	52	G
2	B	53	A
2	B	58	C
2	B	59	G
2	B	62	A
2	B	63	G
2	B	64	U
2	B	65	U
2	B	66	U
2	B	68	G
2	B	69	C
2	B	70	U
2	B	71	C
2	B	72	A

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	73	C
2	B	74	G
2	B	75	C
2	B	76	A
2	B	77	G
2	B	79	C
2	B	80	G
2	B	88	A
2	B	89	U
2	B	90	C
2	B	91	U
2	B	92	U
2	B	94	U
2	B	95	U
2	B	96	G
2	B	97	U
2	B	98	G
5	F	4	C
5	F	5	U
5	F	6	U
5	F	7	C
5	F	8	G
5	F	9	G
5	F	13	A
5	F	14	A
5	F	19	U
5	F	20	A
5	F	21	A
5	F	22	A
5	F	24	U
5	F	26	G
5	F	27	G
5	F	29	A
5	F	37	G
5	F	40	A
5	F	41	A
5	F	42	G
5	F	47	G
5	F	51	G
5	F	61	C
5	F	62	A
5	F	67	U

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	F	68	G
5	F	71	A
5	F	72	C
5	F	73	G
5	F	74	C
5	F	75	A
5	F	76	U
5	F	77	A
5	F	78	A
5	F	79	A
5	F	80	U
5	F	81	C
5	F	82	G
5	F	83	A
5	F	84	G
5	F	85	A
5	F	86	A
5	F	87	A
5	F	88	U
5	F	89	G
5	F	90	G
5	F	91	U
5	F	92	U
22	H	2	U
22	H	3	A
22	H	4	C
22	H	5	C
22	H	6	U
22	H	7	U
22	H	8	U
22	H	9	C
22	H	10	U
22	H	11	C
22	H	12	G
22	H	13	G
22	H	14	C
22	H	16	U
22	H	17	U
22	H	18	U
22	H	19	U
22	H	20	G
22	H	21	G

Continued from previous page...

Mol	Chain	Res	Type
22	H	24	C
22	H	25	A
22	H	26	G
22	H	30	A
22	H	31	U
22	H	32	G
22	H	34	G
22	H	38	U
22	H	40	U
30	5	-6	C
31	3	2	U
31	3	3	A
31	3	7	G
31	3	8	C
31	3	12	G
31	3	13	C
31	3	14	A
31	3	17	A
31	3	19	U
31	3	20	A
31	3	21	C
31	3	22	U
31	3	91	A
31	3	96	U
31	3	98	U
31	3	100	C

All (16) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	B	16	A
2	B	97	U
5	F	19	U
5	F	26	G
5	F	76	U
5	F	78	A
5	F	80	U
22	H	6	U
22	H	30	A
22	H	31	U
31	3	16	U
31	3	20	A

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
31	3	21	C
31	3	90	C
31	3	95	U
31	3	97	A

5.4   Non-standard residues in protein, DNA, RNA chains ⓘ

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	$\#|Z| > 2$
21	SEP	R	225	21	8,9,10	1.19	1 (12%)	8,12,14	1.31	2 (25%)
21	SEP	R	217	21	8,9,10	1.16	1 (12%)	8,12,14	1.38	1 (12%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
21	SEP	R	225	21	-	0/5/8/10	-
21	SEP	R	217	21	-	1/5/8/10	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
21	R	217	SEP	P-O1P	2.38	1.58	1.50
21	R	225	SEP	P-O1P	2.18	1.57	1.50

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
21	R	217	SEP	O2P-P-OG	3.08	114.92	106.73

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
21	R	225	SEP	OG-CB-CA	2.72	110.79	108.14
21	R	225	SEP	O3P-P-OG	2.09	112.30	106.73

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
21	R	217	SEP	CA-CB-OG-P

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 11 ligands modelled in this entry, 10 are monoatomic - leaving 1 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
32	GTP	C	1001	33	26,34,34	1.11	3 (11%)	32,54,54	0.89	1 (3%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
32	GTP	C	1001	33	-	5/18/38/38	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
32	C	1001	GTP	C5-C6	-2.67	1.42	1.47
32	C	1001	GTP	C8-N7	-2.23	1.31	1.35
32	C	1001	GTP	C5-C4	-2.16	1.37	1.43

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
32	C	1001	GTP	C3'-C2'-C1'	2.08	104.11	100.98

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
32	C	1001	GTP	C5'-O5'-PA-O1A
32	C	1001	GTP	C5'-O5'-PA-O2A
32	C	1001	GTP	PA-O3A-PB-O3B
32	C	1001	GTP	PG-O3B-PB-O2B
32	C	1001	GTP	C5'-O5'-PA-O3A

There are no ring outliers.

No monomer is involved in short contacts.

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

Ligand GTP C 1001
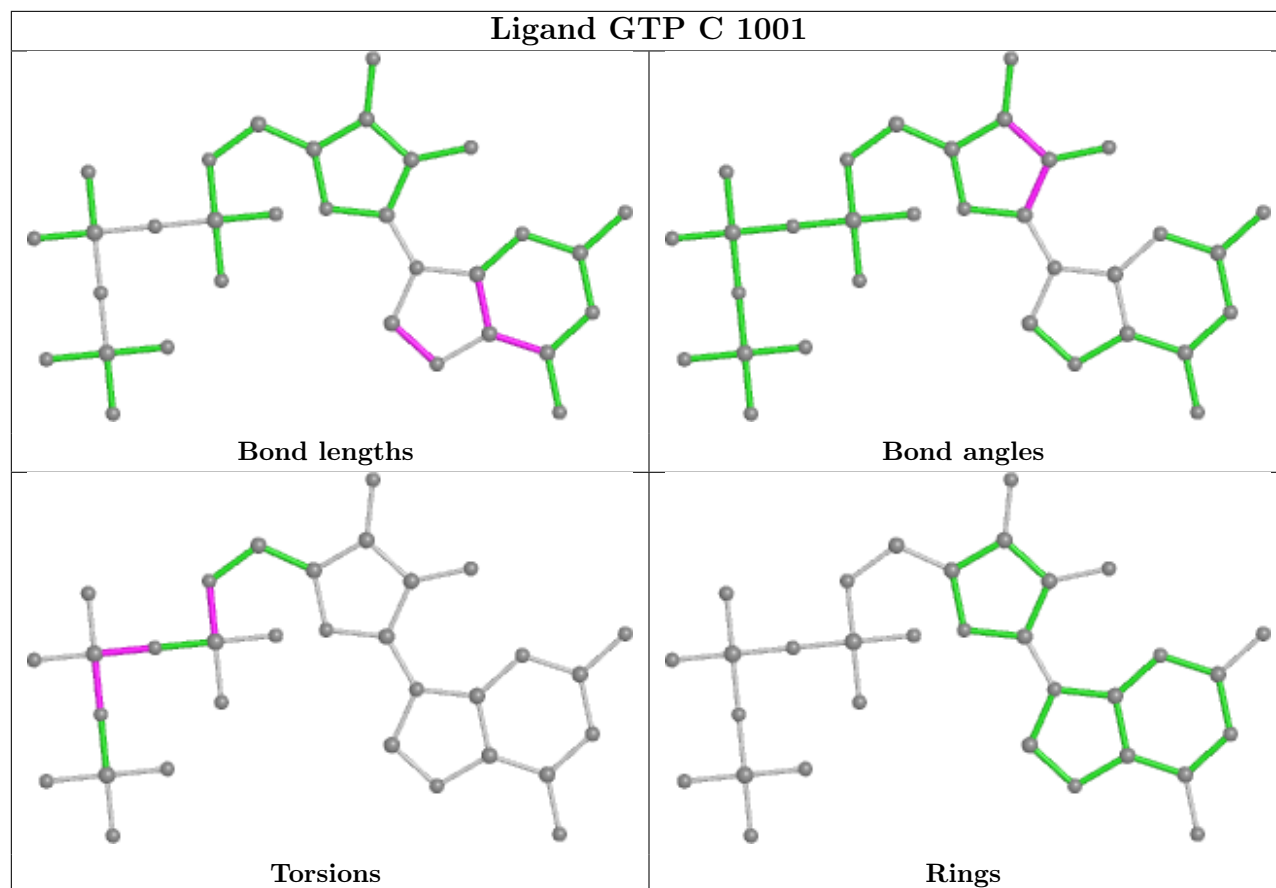
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Map visualisation ⓘ

This section contains visualisations of the EMDB entry EMD-38362. These allow visual inspection of the internal detail of the map and identification of artifacts.

Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections ⓘ
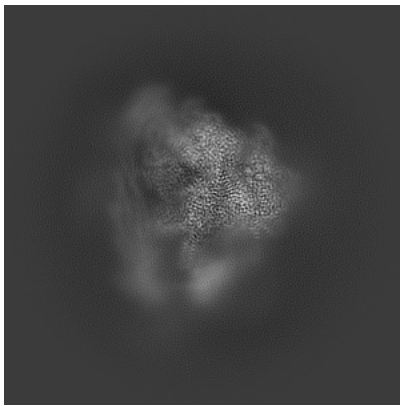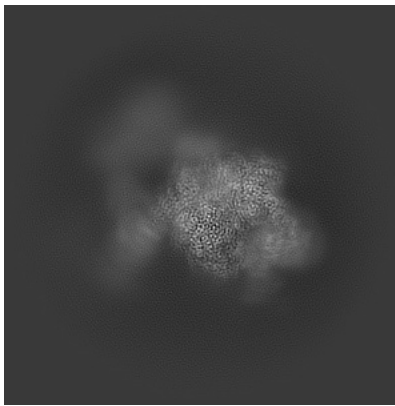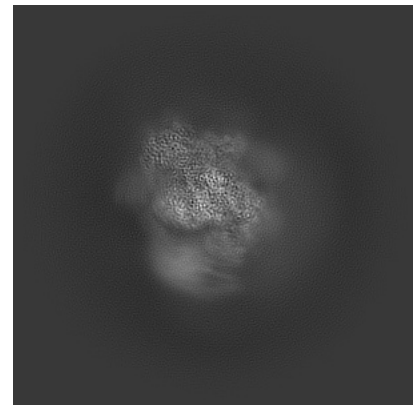
6.1.1 Primary map

X   Y   Z

6.1.2 Raw map

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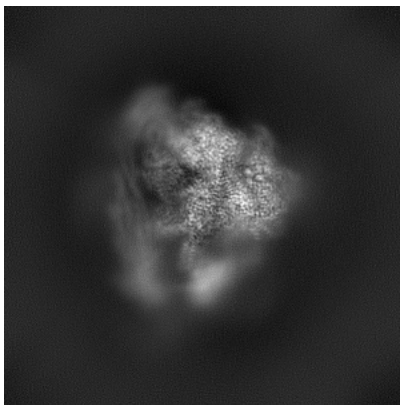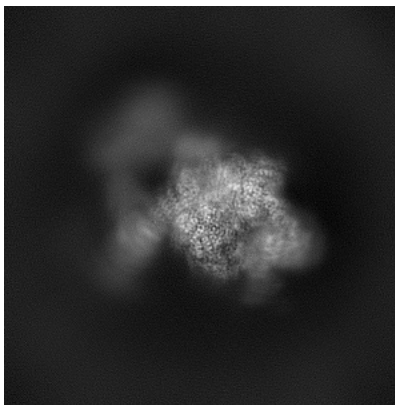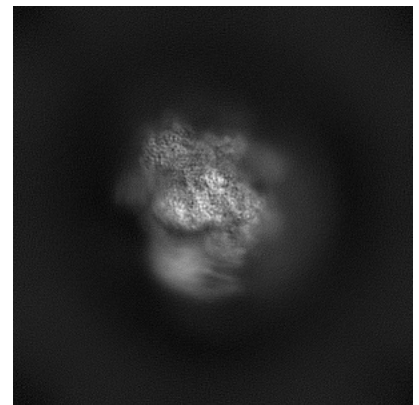
The images above show the map projected in three orthogonal directions.

6.2　Central slices ⓘ
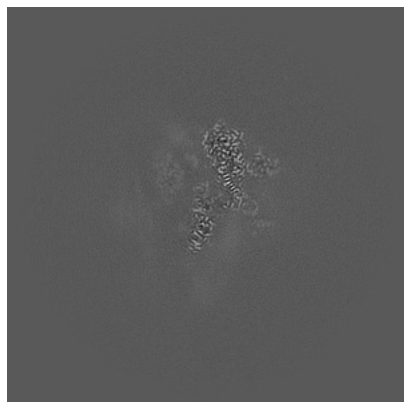
6.2.1　Primary map
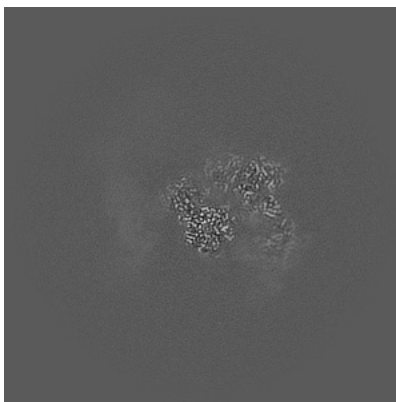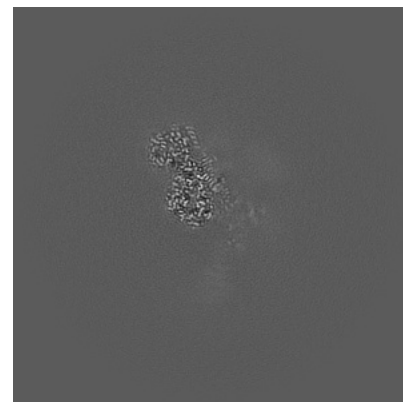
| X Index: 220 | Y Index: 220 | Z Index: 220 |

6.2.2　Raw map
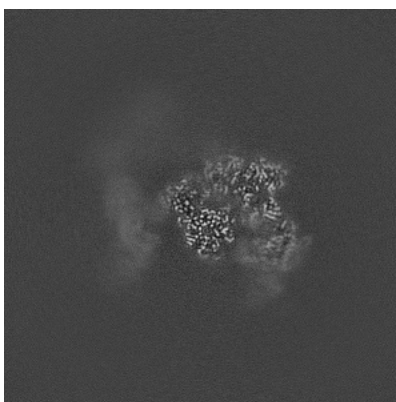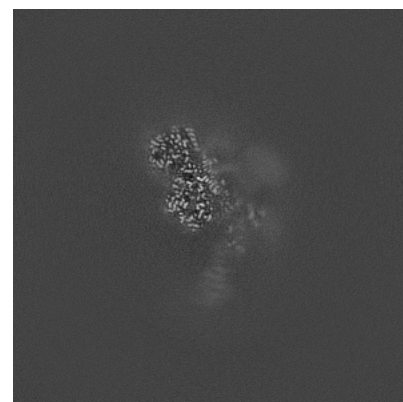
| X Index: 220 | Y Index: 220 | Z Index: 220 |

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

6.3 Largest variance slices (i)

6.3.1 Primary map
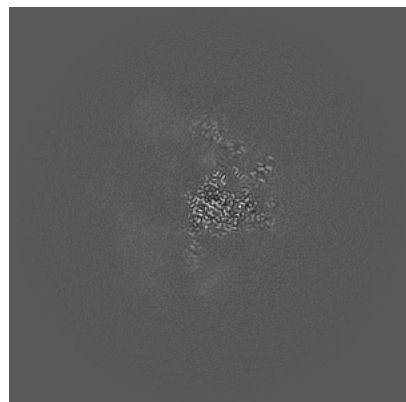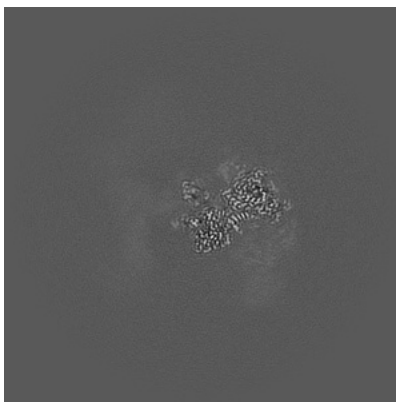
| X Index: 203 | Y Index: 234 | Z Index: 222 |
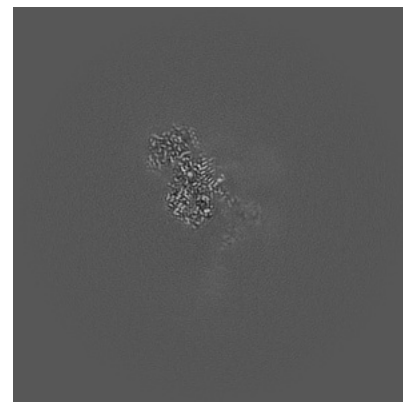
6.3.2 Raw map
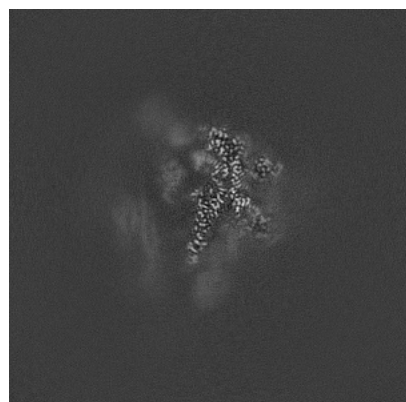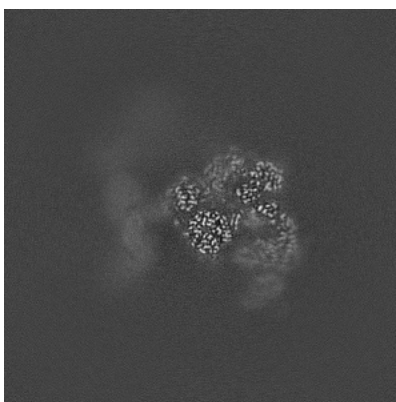
| X Index: 213 | Y Index: 224 | Z Index: 222 |
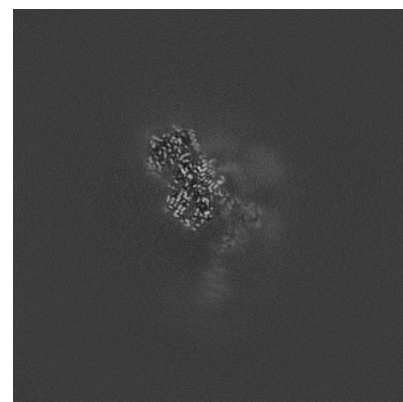
The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal standard-deviation projections (False-color) ⓘ

6.4.1 Primary map
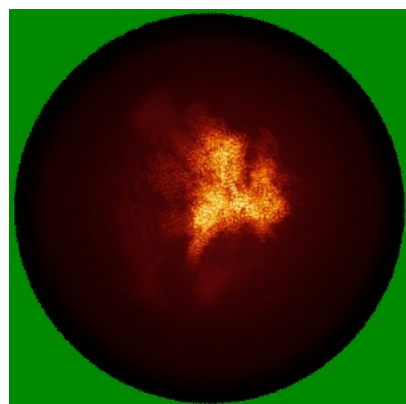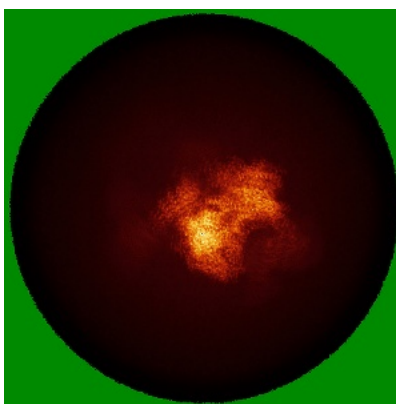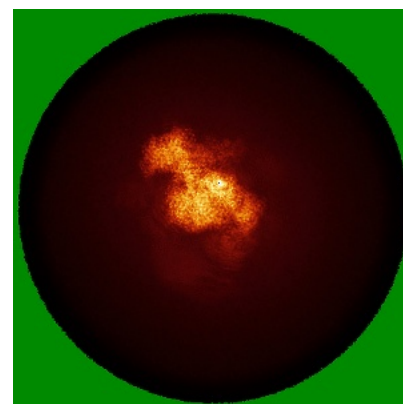
| X | Y | Z |

6.4.2 Raw map

| X | Y | Z |
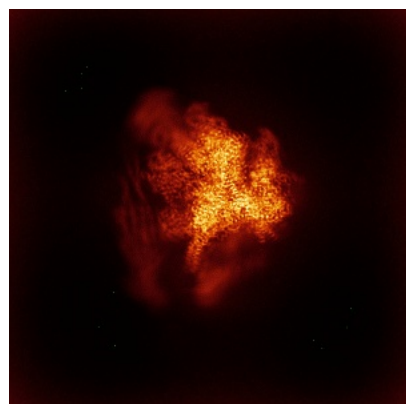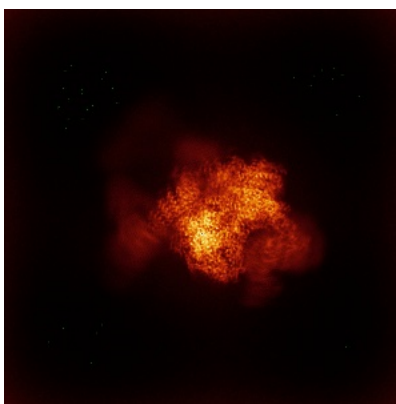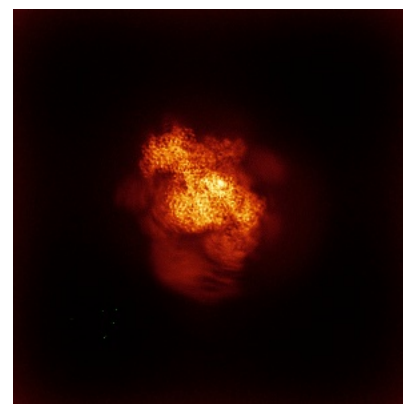
The images above show the map standard deviation projections with false color in three orthogonal directions. Minimum values are shown in green, max in blue, and dark to light orange shades represent small to large values respectively.

6.5 Orthogonal surface views ⓘ

6.5.1 Primary map

X Y Z
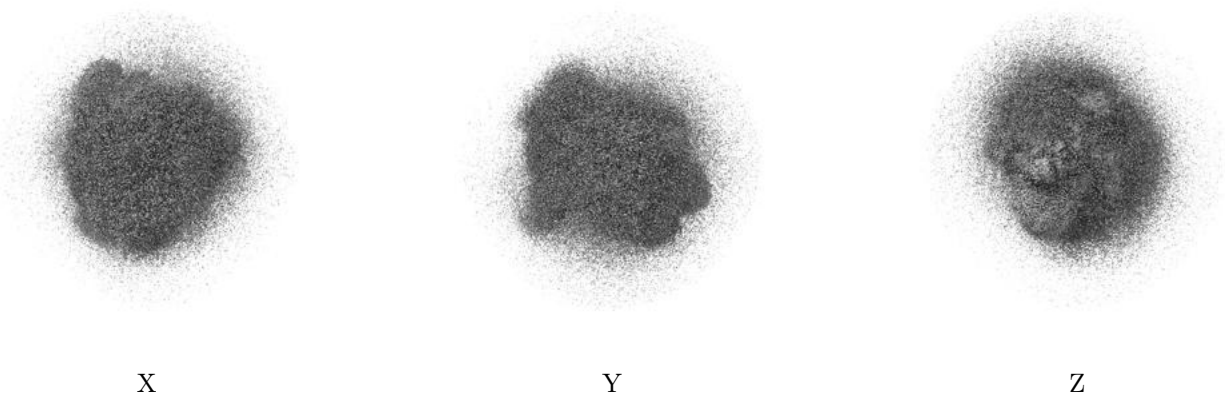
The images above show the 3D surface view of the map at the recommended contour level 0.08. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.

6.5.2 Raw map

X Y Z
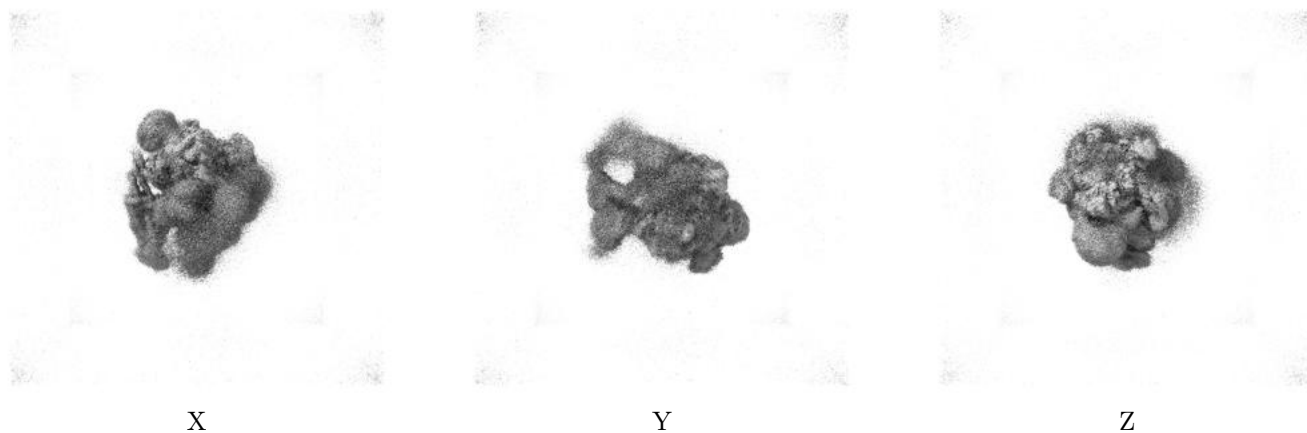
These images show the 3D surface of the raw map. The raw map's contour level was selected so that its surface encloses the same volume as the primary map does at its recommended contour level.

6.6 Mask visualisation ⓘ

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

7 Map analysis ⓘ

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

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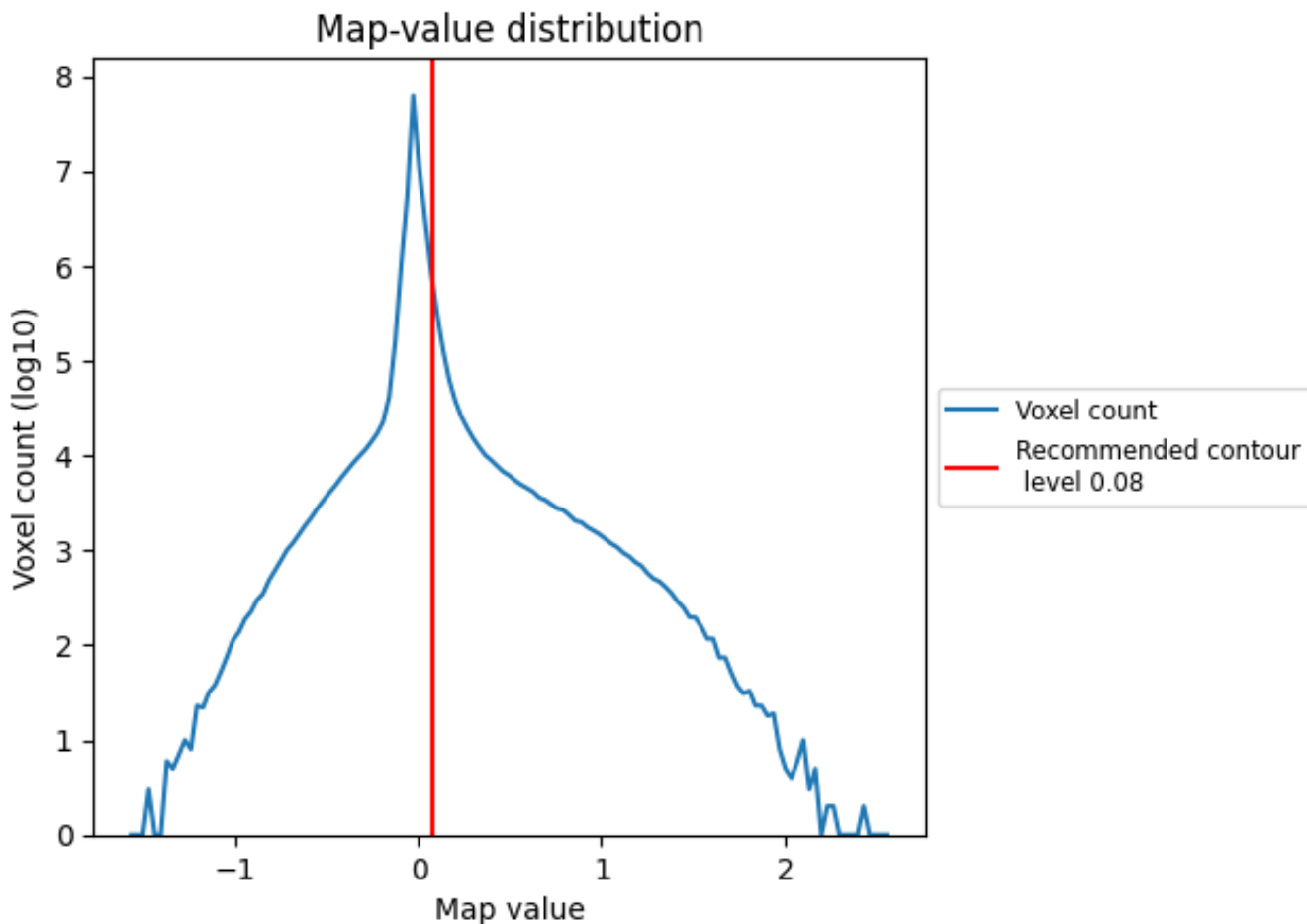
The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.

7.2　Volume estimate ⓘ
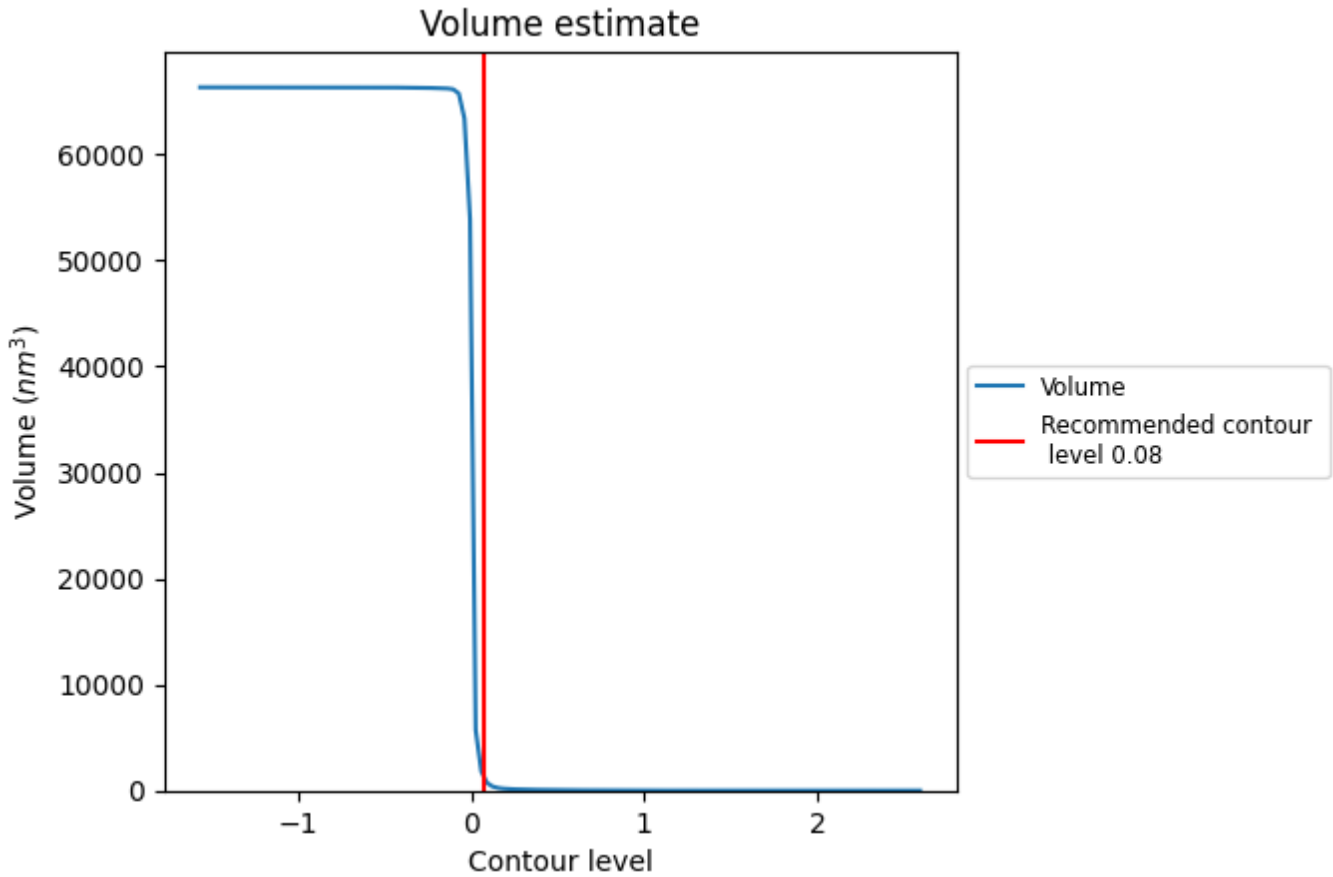
The volume at the recommended contour level is 1131 nm^3; this corresponds to an approximate mass of 1022 kDa.

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 ⓘ
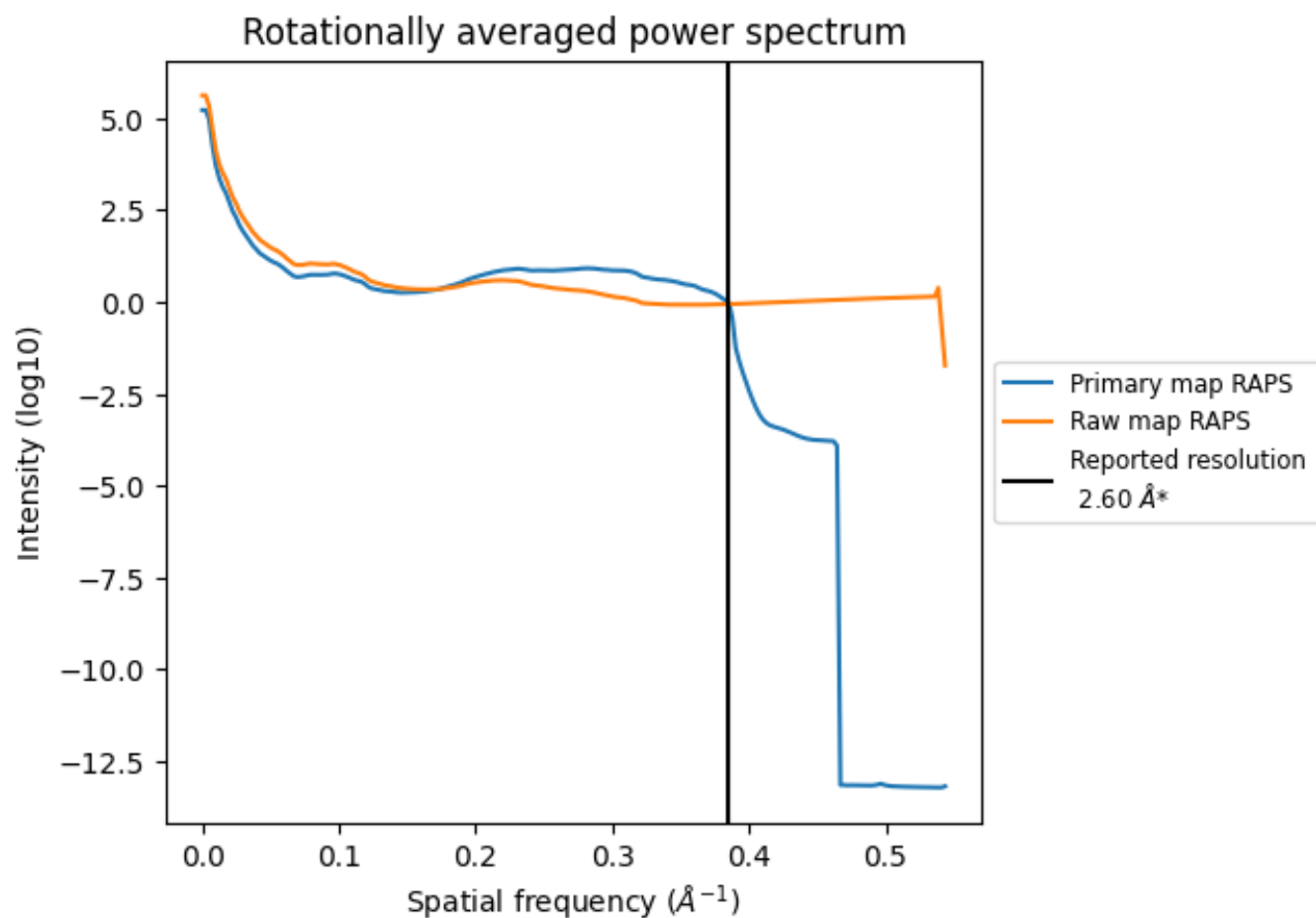
*Reported resolution corresponds to spatial frequency of 0.385 Å^{-1}

8 Fourier-Shell correlation (i)

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

8.1 FSC (i)
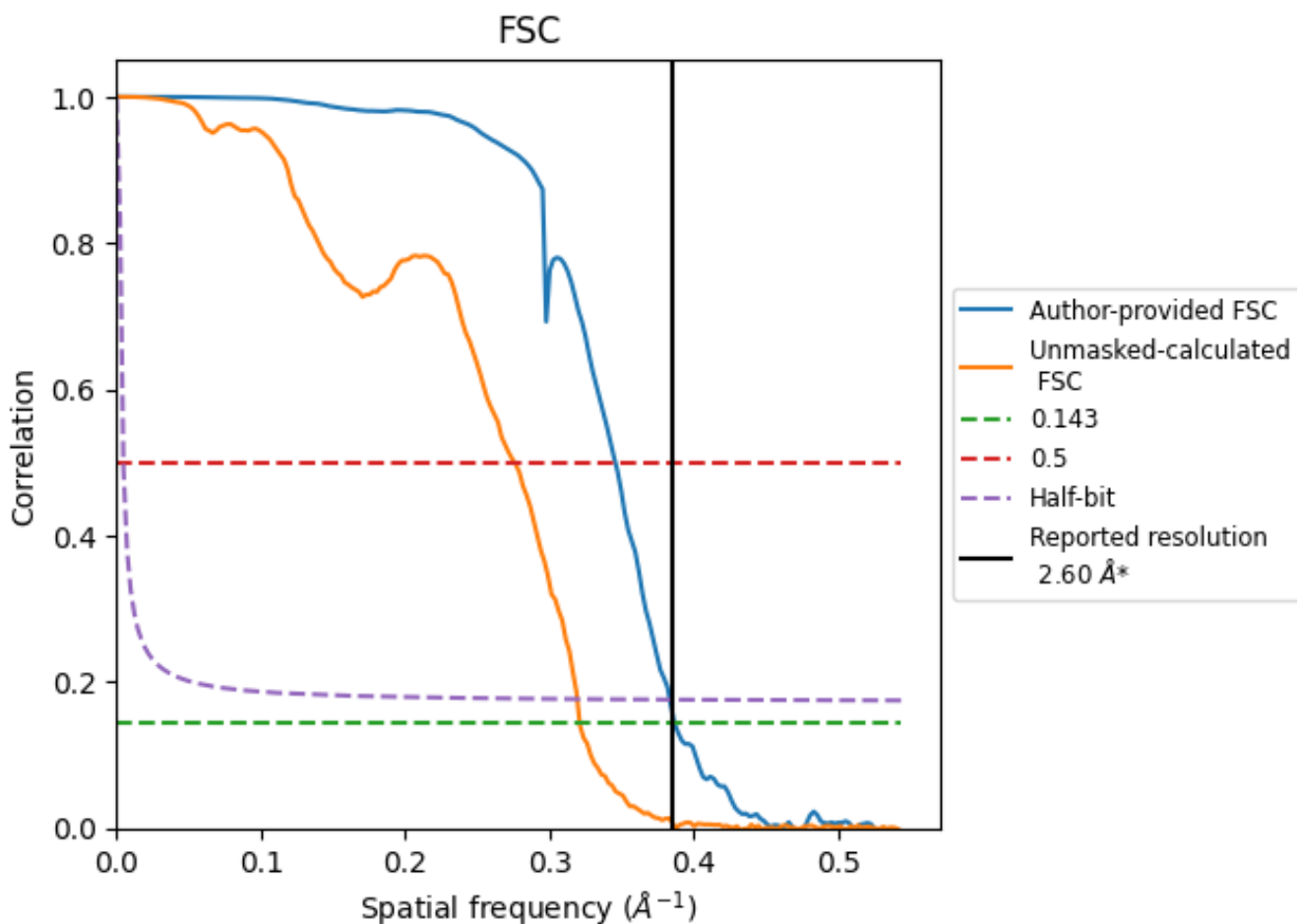
*Reported resolution corresponds to spatial frequency of 0.385 Å^{-1}

8.2 Resolution estimates ⓘ

Resolution estimate (Å)	Estimation criterion (FSC cut-off)		
	0.143	0.5	Half-bit
Reported by author	2.60	-	-
Author-provided FSC curve	2.59	2.89	2.61
Unmasked-calculated*	3.12	3.63	3.13

*Resolution estimate based on FSC curve calculated by comparison of deposited half-maps. The value from deposited half-maps intersecting FSC 0.143 CUT-OFF 3.12 differs from the reported value 2.6 by more than 10 %

9 Map-model fit (i)

This section contains information regarding the fit between EMDB map EMD-38362 and PDB model 8XI2. Per-residue inclusion information can be found in section 3 on page 11.

9.1 Map-model overlay (i)

| X | Y | Z |
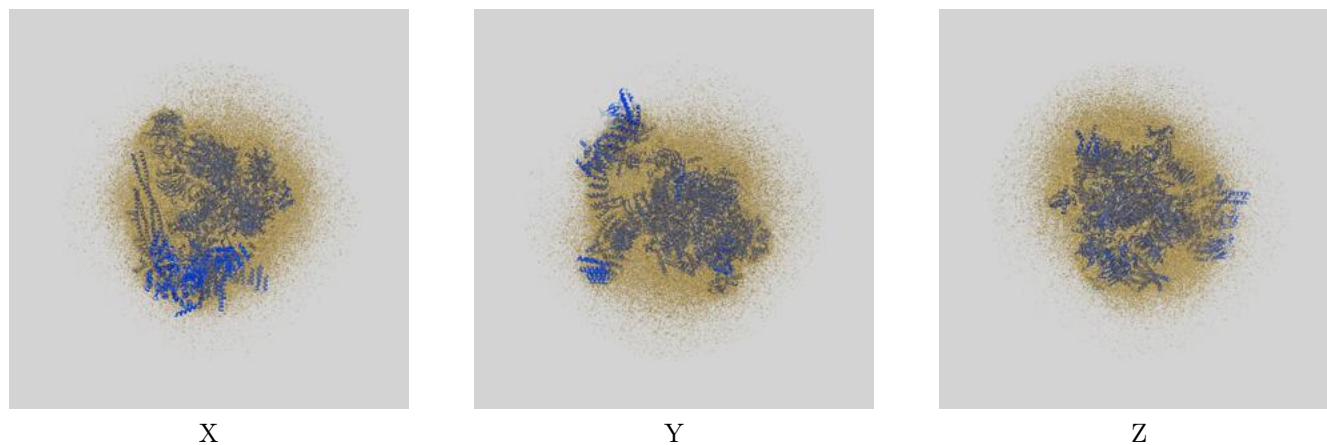
The images above show the 3D surface view of the map at the recommended contour level 0.08 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.

9.2　Q-score mapped to coordinate model ⓘ
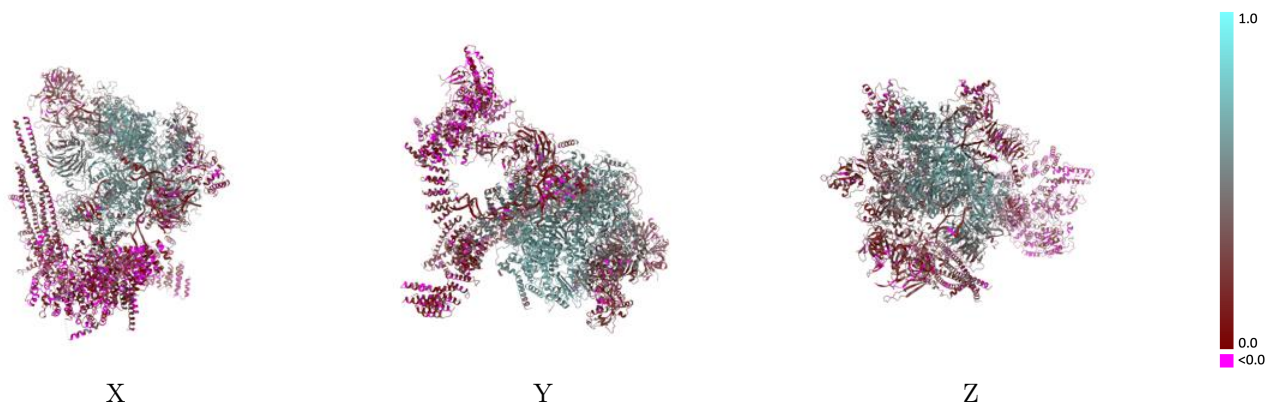
The images above show the model with each residue coloured according its Q-score. This shows their resolvability in the map with higher Q-score values reflecting better resolvability. Please note: Q-score is calculating the resolvability of atoms, and thus high values are only expected at resolutions 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 ⓘ
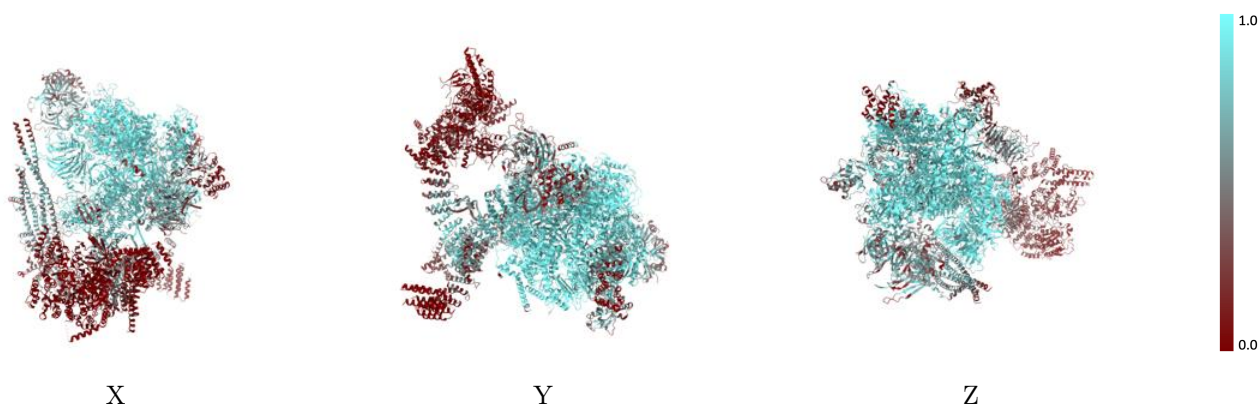
The images above show the model with each residue coloured according to its atom inclusion. This shows to what extent they are inside the map at the recommended contour level (0.08).

9.4 Atom inclusion ⓘ
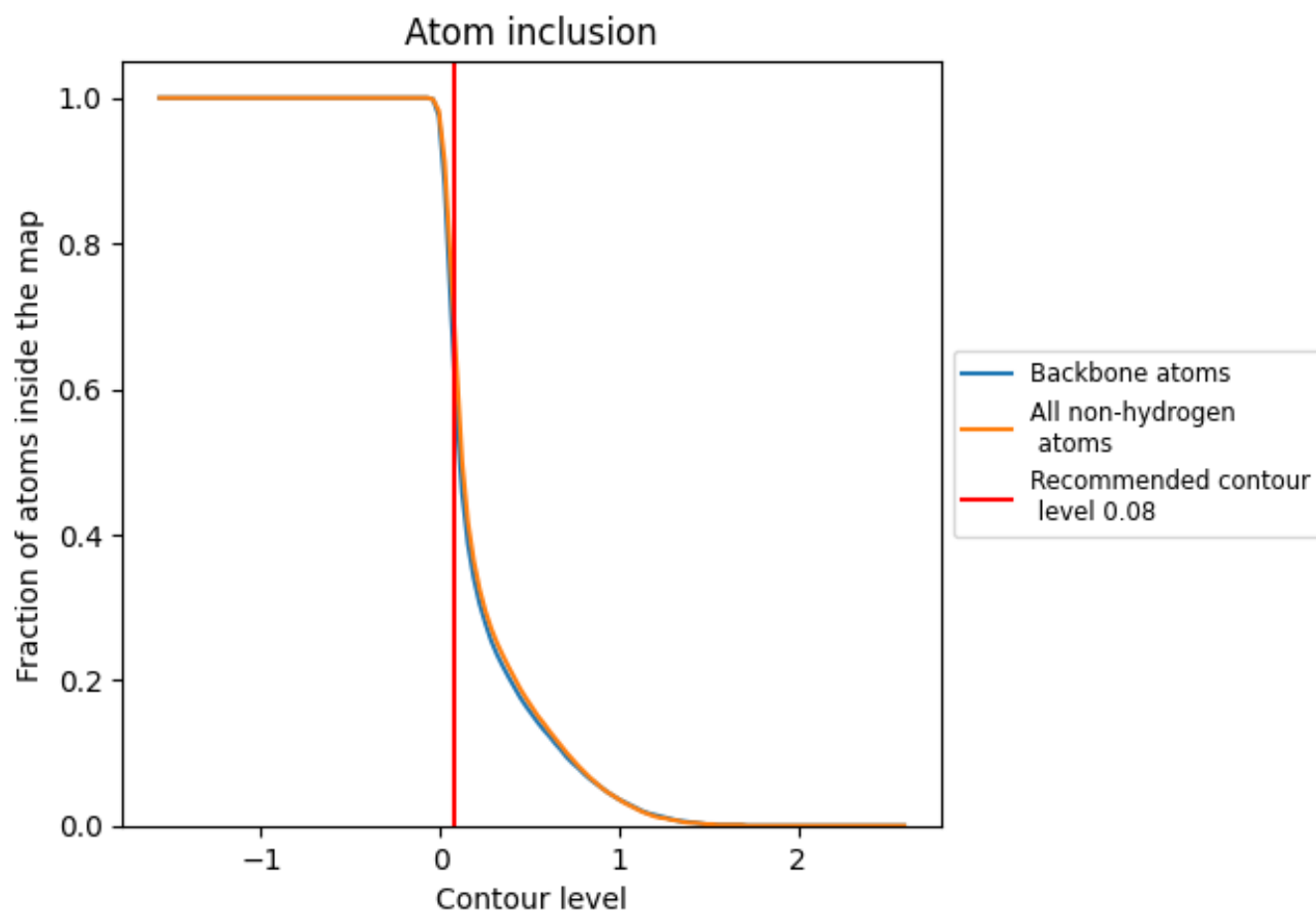
At the recommended contour level, 62% of all backbone atoms, 68% of all non-hydrogen atoms, are inside the map.

9.5　Map-model fit summary ⓘ

The table lists the average atom inclusion at the recommended contour level (0.08) and Q-score for the entire model and for each chain.

Chain	Atom inclusion	Q-score
All	0.6850	0.3810
3	0.7060	0.2340
5	0.9530	0.6210
A	0.8830	0.5470
B	0.8690	0.3780
C	0.7990	0.3620
E	0.9280	0.4580
F	0.9130	0.4530
H	0.6630	0.2840
I	0.4210	0.1850
J	0.6150	0.3610
K	0.4830	0.2310
L	0.6460	0.3640
M	0.7050	0.3600
N	0.9860	0.6290
O	0.7890	0.4460
P	0.9320	0.5920
Q	0.0410	0.1050
R	0.9330	0.5620
S	0.9540	0.5290
T	0.9960	0.6770
U	0.4390	0.3290
V	0.2680	0.2470
W	0.5890	0.3040
a	0.7190	0.2760
b	0.6710	0.2880
c	0.5790	0.2130
d	0.5490	0.2310
e	0.5280	0.2160
f	0.4730	0.2010
g	0.6610	0.2630
q	0.4630	0.2050
r	0.5790	0.2140
s	0.6220	0.2370
t	0.5040	0.2050

1.0

0.0
<0.0