



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

Mar 3, 2018 – 10:42 AM EST

PDB ID : 5TZS  
EMDB ID: : EMD-8473  
Title : Architecture of the yeast small subunit processome  
Authors : Chaker-Margot, M.; Barandun, J.; Hunziker, M.; Klinge, S.  
Deposited on : 2016-11-22  
Resolution : 5.10 Å(reported)

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

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

---

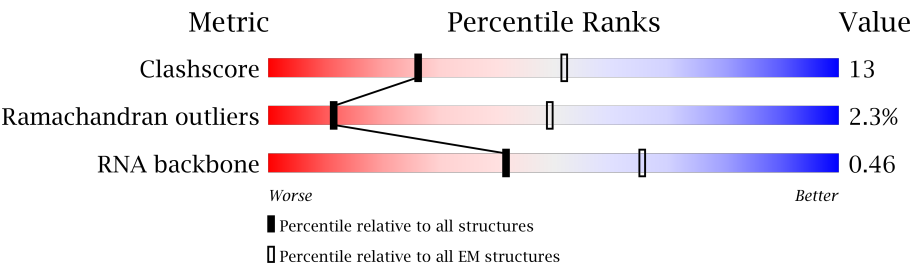
MolProbity : 4.02b-467  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030736

# 1 Overall quality at a glance i

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

The reported resolution of this entry is 5.10 Å.

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



Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	125131	1336
Ramachandran outliers	121729	1120
RNA backbone	3398	335

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	0	364	92% 7% .
2	1	1800	10% 11% 6% . 71%
3	2	126	19% 40% 37% .
4	3	236	76% 14% . 8%
5	5	261	76% 15% . 8%
6	6	225	74% 14% . 10%
7	7	190	85% 13% ..
8	8	200	73% 13% . 14%
9	9	197	81% 11% . 6%

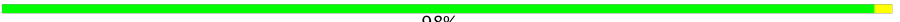
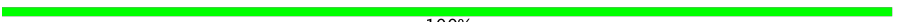








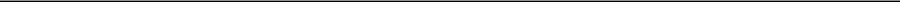

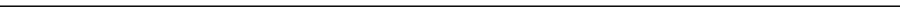


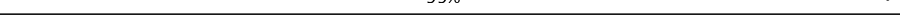
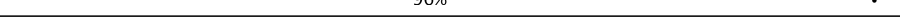
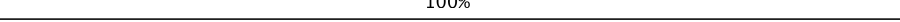

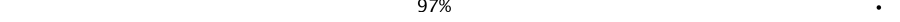


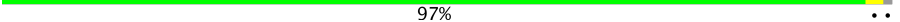

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
10	A	39	100%
11	B	108	95% 5%
12	C	143	75% 6% 20%
13	D	156	85% 9% 6%
14	E	130	94% . .
15	F	135	58% 8% . 33%
16	G	67	85% 7% 7%
17	H	544	91% 9%
18	I	176	90% . 6%
19	J	107	90% 10%
19	K	107	89% 7% . .
20	M	258	90% 10%
21	N	545	87% 13%
22	O	638	92% 8%
23	P	306	92% 8%
24	Q	710	86% 14%
25	R	717	97% .
26	S	250	88% 11% .
27	T	781	93% . .
28	U	284	98% .
29	V	263	93% 7%
30	W	104	84% 7% 10%
31	X	640	91% 9%
32	Y	641	84% 16%
33	Z	151	95% 5%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
34	a	312	98% 
35	b	341	100% 
36	c	221	99% 
37	d	216	99% 
38	e	126	97% 
38	f	126	90% 
39	g	573	63%  36%
40	h	367	97% 
41	i	511	92%  7%
42	j	252	84% 
42	k	252	86%  13%
43	l	124	100% 
44	m	156	99% 
45	n	160	100% 
46	o	175	99% 
47	p	924	96% 
48	q	372	100% 
49	r	145	56%  43%
50	s	290	97% 
50	t	290	89%  7%
50	u	290	94%  6%
51	v	580	97% 
52	y	507	93%  7%
53	z	76	100% 

## 2 Entry composition

There are 53 unique types of molecules in this entry. The entry contains 98451 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a RNA chain called 5' external transcribed spacer.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	0	364	Total	C	N	O	P	0	0
			4476	1871	43	2198	364		

- Molecule 2 is a RNA chain called 18S ribosomal RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	1	515	Total	C	N	O	P	0	0
			10978	4908	1957	3598	515		

- Molecule 3 is a RNA chain called U3 snoRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	2	126	Total	C	N	O	P	0	0
			2468	1095	388	859	126		

- Molecule 4 is a protein called rpS6\_ES6.

Mol	Chain	Residues	Atoms				AltConf	Trace
4	3	216	Total	C	N	O	0	0
			1063	631	216	216		

- Molecule 5 is a protein called rpS4\_ES4.

Mol	Chain	Residues	Atoms				AltConf	Trace
5	5	241	Total	C	N	O	0	0
			1185	703	241	241		

- Molecule 6 is a protein called rpS5\_US7.

Mol	Chain	Residues	Atoms				AltConf	Trace
6	6	202	Total	C	N	O	0	0
			1000	596	202	202		

- Molecule 7 is a protein called rpS7\_eS7.

Mol	Chain	Residues	Atoms				AltConf	Trace
7	7	186	Total	C	N	O	0	0
			923	551	186	186		

- Molecule 8 is a protein called rpS8\_eS8.

Mol	Chain	Residues	Atoms				AltConf	Trace
8	8	173	Total	C	N	O	0	0
			849	503	173	173		

- Molecule 9 is a protein called rpS9\_uS4.

Mol	Chain	Residues	Atoms				AltConf	Trace
9	9	185	Total	C	N	O	0	0
			915	545	185	185		

- Molecule 10 is a DNA chain called 5' domain-associated.

Mol	Chain	Residues	Atoms				AltConf	Trace
10	A	39	Total	C	O	P	0	0
			468	195	234	39		

- Molecule 11 is a DNA chain called 3' domain-associated.

Mol	Chain	Residues	Atoms				AltConf	Trace
11	B	108	Total	C	O	P	0	0
			1296	540	648	108		

- Molecule 12 is a protein called rpS16\_uS9.

Mol	Chain	Residues	Atoms				AltConf	Trace
12	C	115	Total	C	N	O	0	0
			566	336	115	115		

- Molecule 13 is a protein called rpS11\_uS17.

Mol	Chain	Residues	Atoms				AltConf	Trace
13	D	146	Total	C	N	O	0	0
			721	429	146	146		

- Molecule 14 is a protein called rpS22\_uS8.

Mol	Chain	Residues	Atoms				AltConf	Trace
14	E	127	Total	C	N	O	0	0
			624	370	127	127		

- Molecule 15 is a protein called rpS24\_eS24.

Mol	Chain	Residues	Atoms				AltConf	Trace
15	F	90	Total	C	N	O	0	0
			444	264	90	90		

- Molecule 16 is a protein called rpS28\_eS28.

Mol	Chain	Residues	Atoms				AltConf	Trace
16	G	62	Total	C	N	O	0	0
			306	182	62	62		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	17	ALA	GLY	conflict	UNP Q3E7X9

- Molecule 17 is a protein called Utp4.

Mol	Chain	Residues	Atoms				AltConf	Trace
17	H	544	Total	C	N	O	0	0
			2680	1592	544	544		

- Molecule 18 is a protein called UtpA\_CTD1.

Mol	Chain	Residues	Atoms				AltConf	Trace
18	I	176	Total	C	N	O	0	0
			880	528	176	176		

- Molecule 19 is a protein called UtpA\_CTD2.

Mol	Chain	Residues	Atoms				AltConf	Trace
19	J	107	Total	C	N	O	0	0
			535	321	107	107		
19	K	105	Total	C	N	O	0	0
			525	315	105	105		

- Molecule 20 is a protein called Beta-propeller 2.

Mol	Chain	Residues	Atoms				AltConf	Trace
20	M	258	Total	C	N	O	0	0
			1275	759	258	258		

- Molecule 21 is a protein called Utp17.

Mol	Chain	Residues	Atoms				AltConf	Trace
21	N	545	Total	C	N	O	0	0
			2678	1588	545	545		

- Molecule 22 is a protein called Utp1.

Mol	Chain	Residues	Atoms				AltConf	Trace
22	O	638	Total	C	N	O	0	0
			3153	1877	638	638		

- Molecule 23 is a protein called Utp6.

Mol	Chain	Residues	Atoms				AltConf	Trace
23	P	306	Total	C	N	O	0	0
			1530	918	306	306		

- Molecule 24 is a protein called Utp12.

Mol	Chain	Residues	Atoms				AltConf	Trace
24	Q	710	Total	C	N	O	0	0
			3503	2083	710	710		

- Molecule 25 is a protein called Utp13.

Mol	Chain	Residues	Atoms				AltConf	Trace
25	R	717	Total	C	N	O	0	0
			3539	2105	717	717		

- Molecule 26 is a protein called Utp18.

Mol	Chain	Residues	Atoms				AltConf	Trace
26	S	250	Total	C	N	O	0	0
			1228	728	250	250		

- Molecule 27 is a protein called Utp21.



Mol	Chain	Residues	Atoms				AltConf	Trace
27	T	749	Total	C	N	O	0	0
			3691	2193	749	749		

- Molecule 28 is a protein called Beta-propeller 5.

Mol	Chain	Residues	Atoms				AltConf	Trace
28	U	284	Total	C	N	O	0	0
			1398	830	284	284		

- Molecule 29 is a protein called Enp2.

Mol	Chain	Residues	Atoms				AltConf	Trace
29	V	263	Total	C	N	O	0	0
			1298	772	263	263		

- Molecule 30 is a protein called UtpA\_CTD4.

Mol	Chain	Residues	Atoms				AltConf	Trace
30	W	104	Total	C	N	O	0	0
			520	312	104	104		

- Molecule 31 is a protein called Kre33.

Mol	Chain	Residues	Atoms				AltConf	Trace
31	X	640	Total	C	N	O	0	0
			3155	1875	640	640		

- Molecule 32 is a protein called Kre33.

Mol	Chain	Residues	Atoms				AltConf	Trace
32	Y	641	Total	C	N	O	0	0
			3160	1878	641	641		

- Molecule 33 is a protein called Imp3.

Mol	Chain	Residues	Atoms				AltConf	Trace
33	Z	151	Total	C	N	O	0	0
			748	446	151	151		

- Molecule 34 is a protein called Nop56.

Mol	Chain	Residues	Atoms				AltConf	Trace
34	a	312	Total	C	N	O	0	0
			1544	920	312	312		

- Molecule 35 is a protein called Nop58.

Mol	Chain	Residues	Atoms				AltConf	Trace
35	b	341	Total	C	N	O	0	0
			1687	1005	341	341		

- Molecule 36 is a protein called Nop1.

Mol	Chain	Residues	Atoms				AltConf	Trace
36	c	221	Total	C	N	O	0	0
			1088	646	221	221		

- Molecule 37 is a protein called Nop1.

Mol	Chain	Residues	Atoms				AltConf	Trace
37	d	216	Total	C	N	O	0	0
			1064	632	216	216		

- Molecule 38 is a protein called Snu13.

Mol	Chain	Residues	Atoms				AltConf	Trace
38	e	122	Total	C	N	O	0	0
			606	362	122	122		
38	f	114	Total	C	N	O	0	0
			566	338	114	114		

- Molecule 39 is a protein called Ribosomal RNA-processing protein 9.

Mol	Chain	Residues	Atoms				AltConf	Trace
39	g	365	Total	C	N	O	0	0
			1799	1069	365	365		

- Molecule 40 is a protein called RNA 3'-terminal phosphate cyclase-like protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
40	h	355	Total	C	N	O	0	0
			1742	1032	355	355		

- Molecule 41 is a protein called Bms1,Ribosome biogenesis protein BMS1,Bms1.

Mol	Chain	Residues	Atoms				AltConf	Trace
41	i	475	Total	C	N	O	0	0
			2347	1398	475	474		

- Molecule 42 is a protein called Ribosomal RNA small subunit methyltransferase NEP1.

Mol	Chain	Residues	Atoms				AltConf	Trace
42	j	211	Total	C	N	O	0	0
			1047	625	211	211		
42	k	218	Total	C	N	O	0	0
			1081	645	218	218		

- Molecule 43 is a protein called Utp24.

Mol	Chain	Residues	Atoms				AltConf	Trace
43	l	124	Total	C	N	O	0	0
			613	365	124	124		

- Molecule 44 is a protein called Imp4.

Mol	Chain	Residues	Atoms				AltConf	Trace
44	m	156	Total	C	N	O	0	0
			775	463	156	156		

- Molecule 45 is a protein called Utp30.

Mol	Chain	Residues	Atoms				AltConf	Trace
45	n	160	Total	C	N	O	0	0
			791	471	160	160		

- Molecule 46 is a protein called Unassigned KH domain.

Mol	Chain	Residues	Atoms				AltConf	Trace
46	o	173	Total	C	N	O	0	0
			858	512	173	173		

- Molecule 47 is a protein called Utp20.

Mol	Chain	Residues	Atoms				AltConf	Trace
47	p	924	Total	C	N	O	0	0
			4620	2772	924	924		

- Molecule 48 is a protein called Repeat protein 2.

Mol	Chain	Residues	Atoms				AltConf	Trace
48	q	372	Total	C	N	O	0	0
			1860	1116	372	372		

- Molecule 49 is a protein called 40S ribosomal protein S23-A.

Mol	Chain	Residues	Atoms				AltConf	Trace
49	r	82	Total	C	N	O	0	0
			402	238	82	82		

- Molecule 50 is a protein called Beta-propeller 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
50	s	290	Total	C	N	O	0	0
			1429	849	290	290		
50	t	269	Total	C	N	O	0	0
			1076	538	269	269		
50	u	274	Total	C	N	O	0	0
			1346	798	274	274		

- Molecule 51 is a protein called Repeat protein 1.

Mol	Chain	Residues	Atoms				AltConf	Trace
51	v	577	Total	C	N	O	0	0
			2885	1731	577	577		

- Molecule 52 is a protein called Unassigned protein helices.

Mol	Chain	Residues	Atoms				AltConf	Trace
52	y	507	Total	C	N	O	0	0
			2535	1521	507	507		

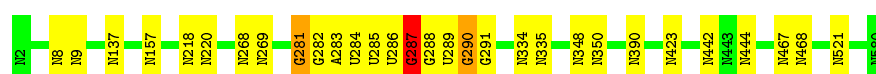
- Molecule 53 is a DNA chain called Unassigned RNA helices.

Mol	Chain	Residues	Atoms				AltConf	Trace
53	z	76	Total	C	O	P	0	0
			912	380	456	76		

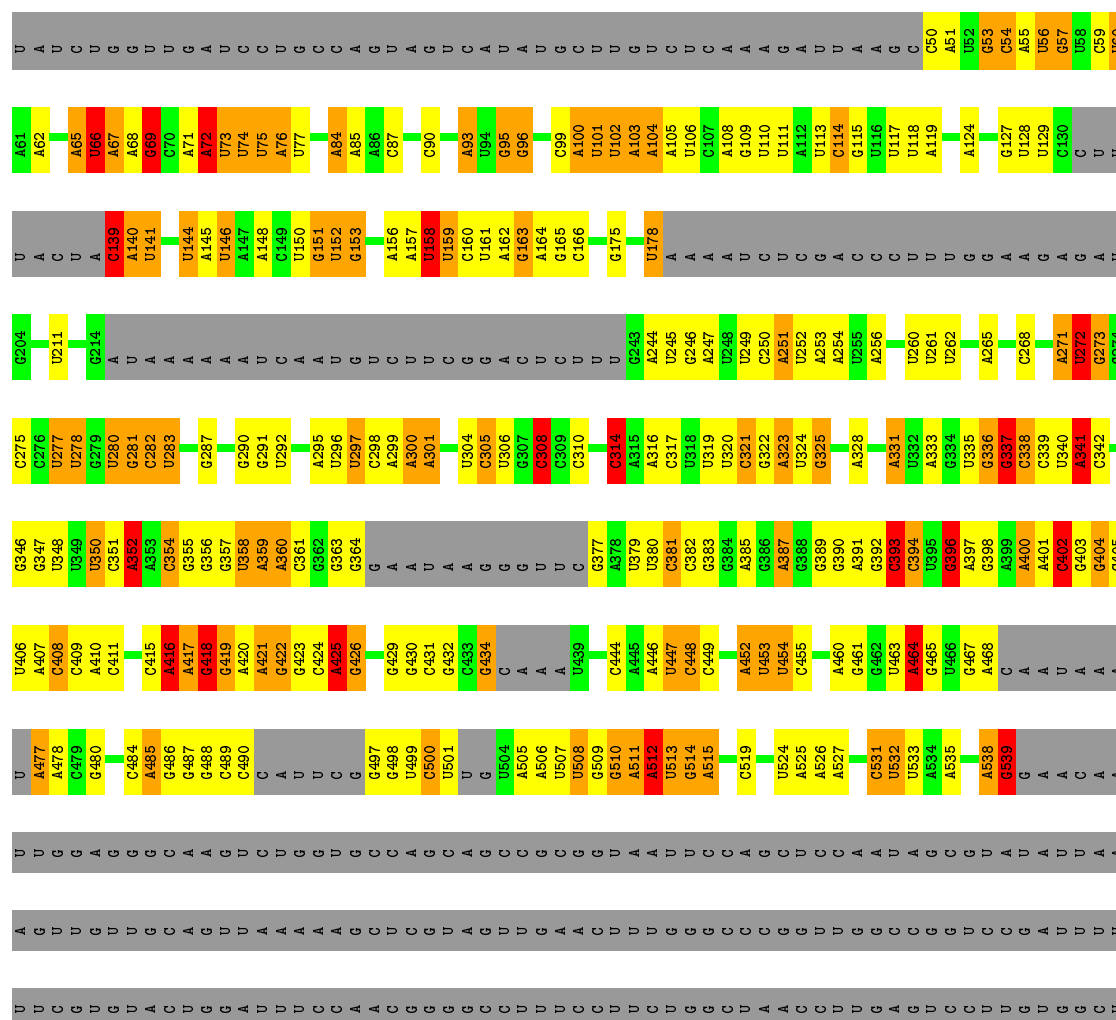
### 3 Residue-property plots

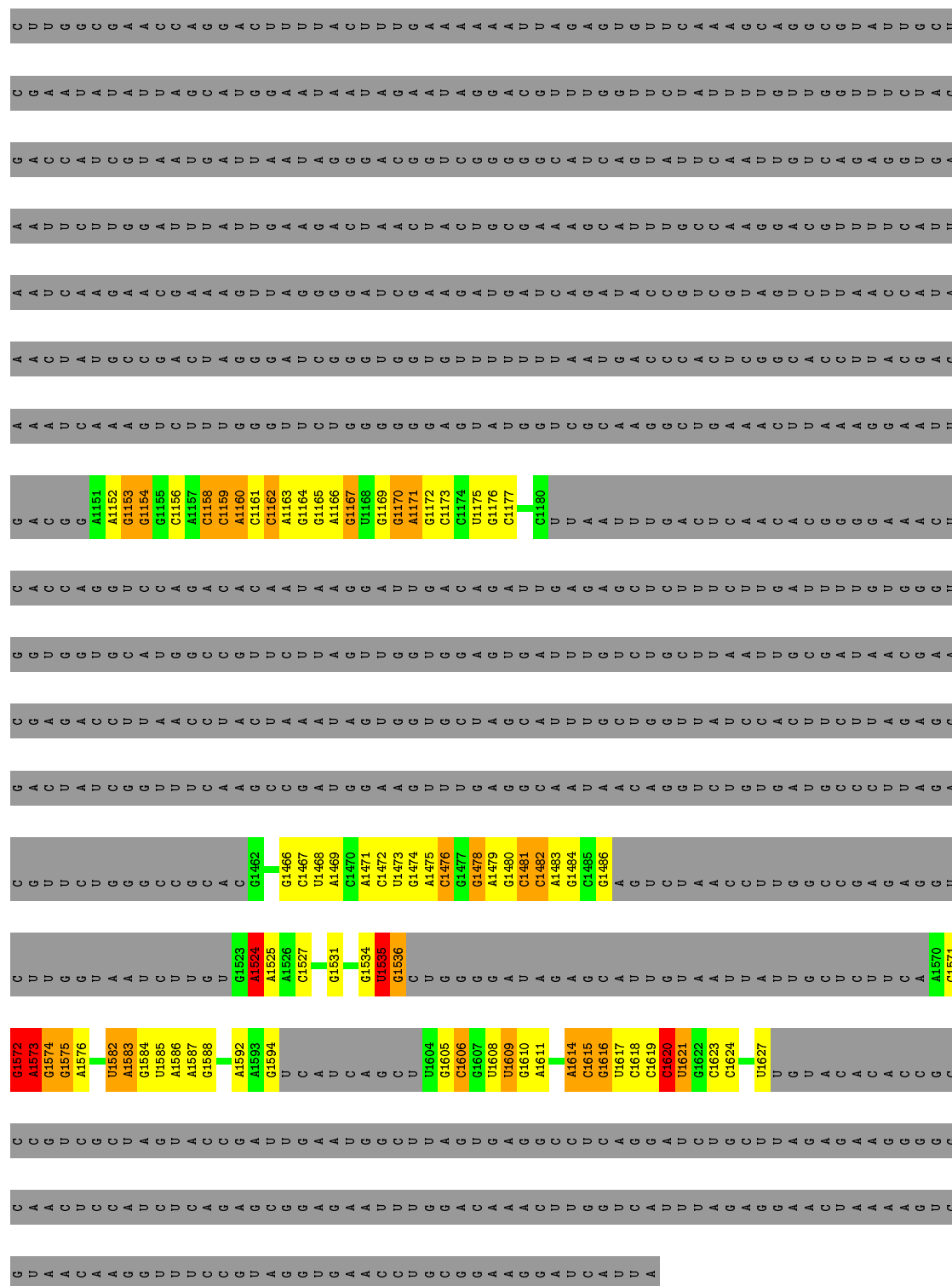
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according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

- Molecule 1: 5' external transcribed spacer




- Molecule 2: 18S ribosomal RNA



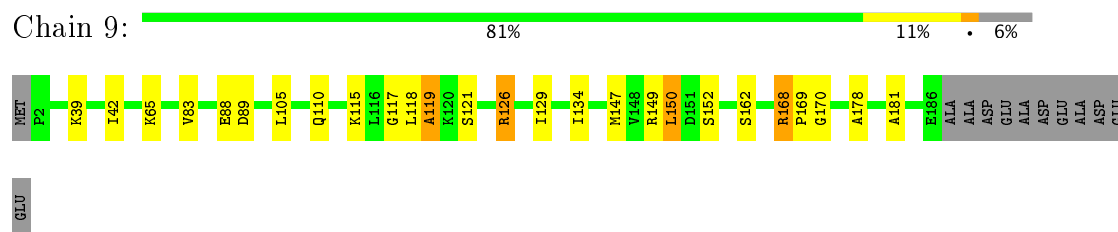


- Molecule 3: U3 snoRNA

Chain 2:  19% 40% 37% .





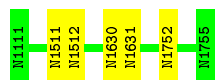


- Molecule 10: 5' domain-associated

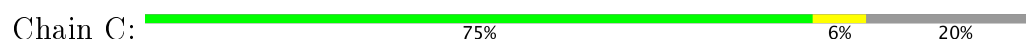


There are no outlier residues recorded for this chain.

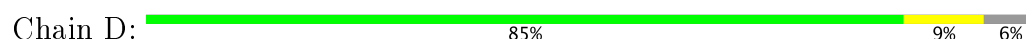
- Molecule 11: 3' domain-associated



- Molecule 12: rpS16\_uS9



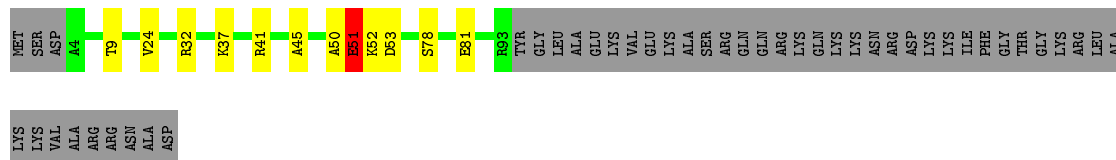
- Molecule 13: rpS11\_uS17



- Molecule 14: rpS22\_uS8




- Molecule 15: rpS24\_eS24





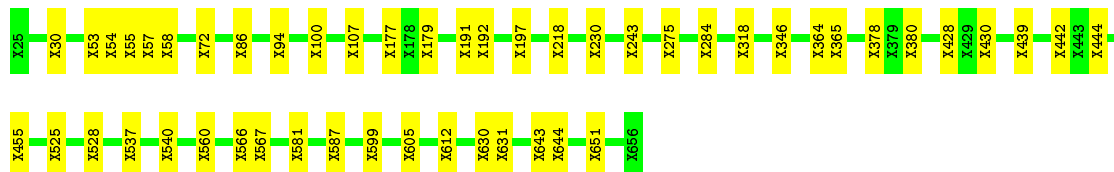
- Molecule 16: rpS28\_eS28

Chain G:  85% 7% 7%




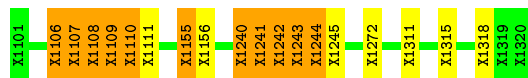
- Molecule 17: Utp4

Chain H:  91% 9%




- Molecule 18: UtpA\_CTD1

Chain I:  90% 6%



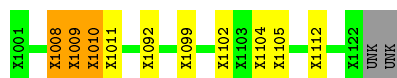
- Molecule 19: UtpA\_CTD2

Chain J:  90% 10%



- Molecule 19: UtpA\_CTD2

Chain K:  89% 7%




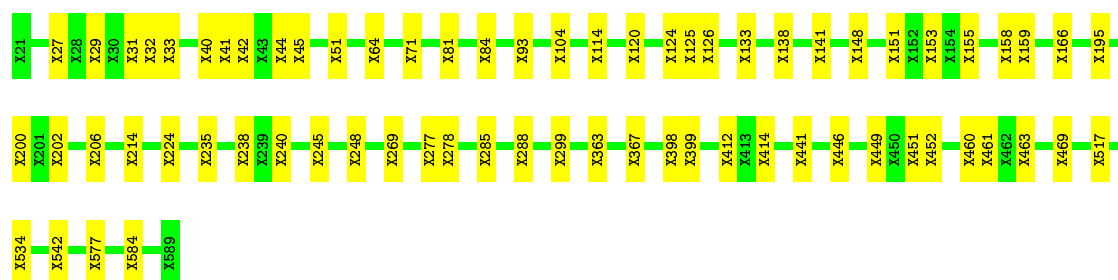
- Molecule 20: Beta-propeller 2

Chain M:  90% 10%



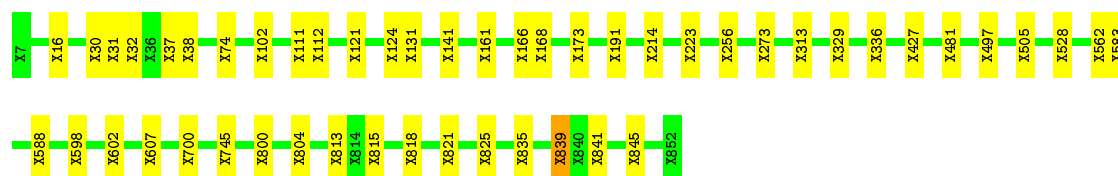
- Molecule 21: Utp17

Chain N:  87% 13%



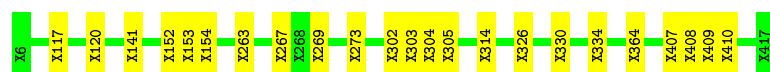
• Molecule 22: Utp1

Chain O: 92% 8%



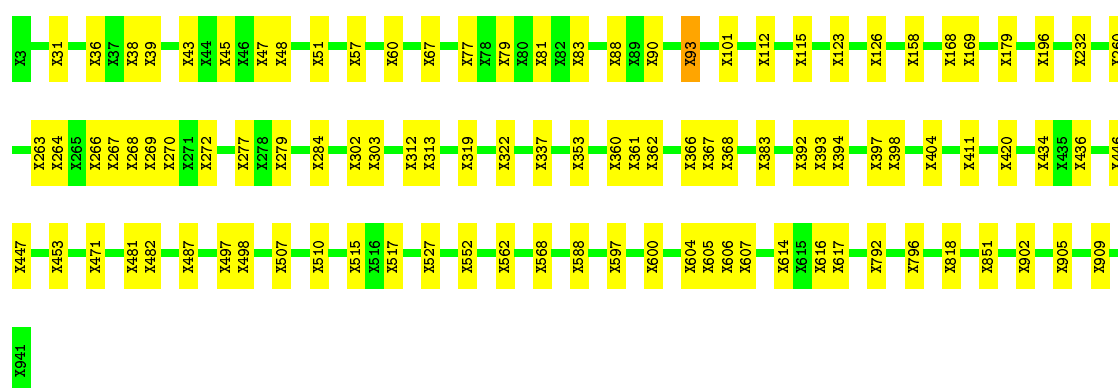
• Molecule 23: Utp6

Chain P: 92% 8%



• Molecule 24: Utp12

Chain Q: 86% 14%

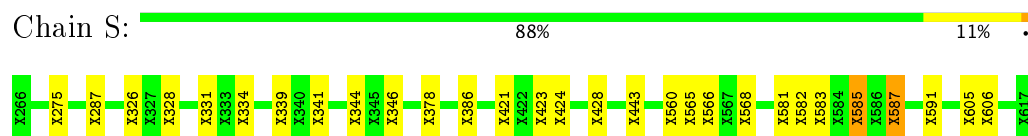


• Molecule 25: Utp13

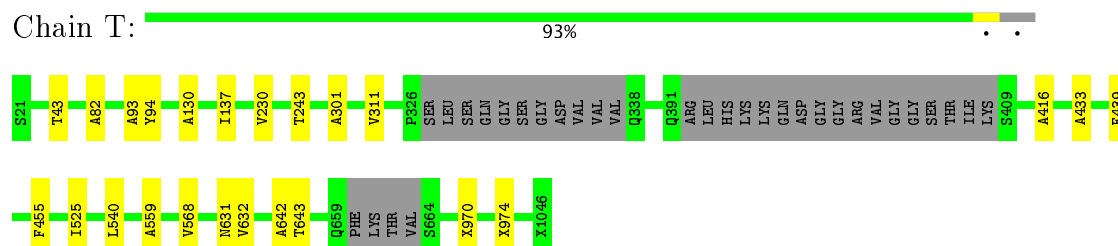
Chain R: 97% 3%



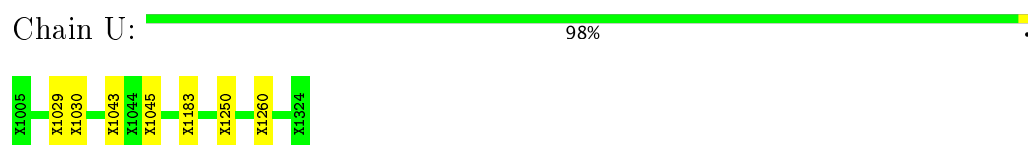
• Molecule 26: Utp18



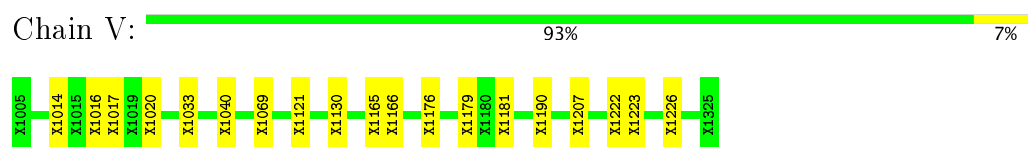
- Molecule 27: Utp21



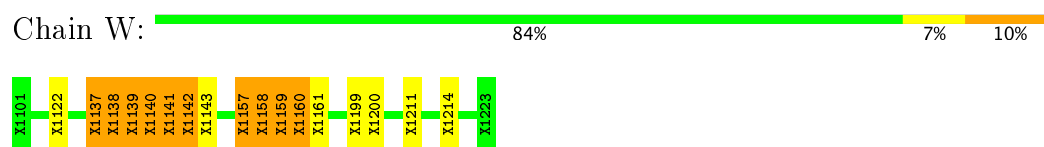
- Molecule 28: Beta-propeller 5



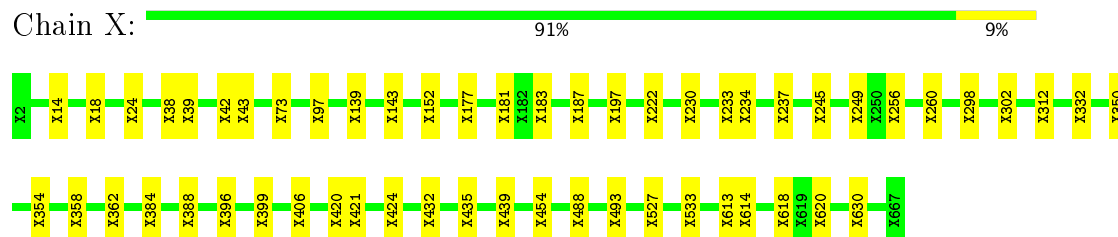
- Molecule 29: Enp2



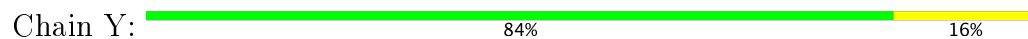
- Molecule 30: UtpA\_CTD4

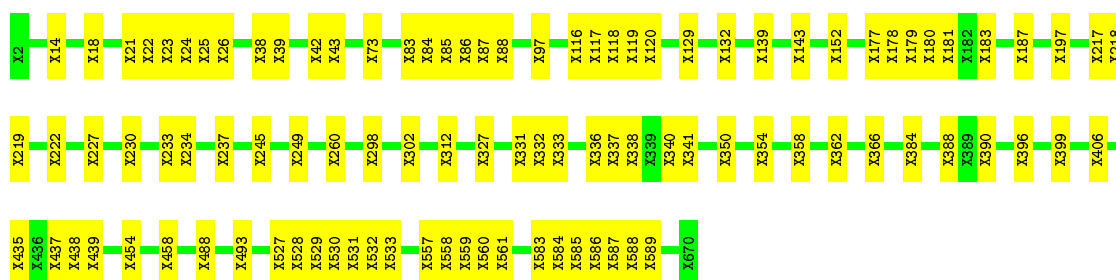


- Molecule 31: Kre33



- Molecule 32: Kre33





• Molecule 33: Imp3

Chain Z:  95% 5%



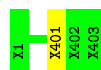
• Molecule 34: Nop56

Chain a:  98% .



• Molecule 35: Nop58

Chain b:  100%



• Molecule 36: Nop1

Chain c:  99% .



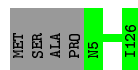
• Molecule 37: Nop1

Chain d:  99% .



• Molecule 38: Snu13

Chain e:  97% .



• Molecule 38: Snu13

MET
SER
ALA
PRO
ASN
PRO
LYS
ALA
PHE
PRO
LEU
ALA
D13
I126

- Chain g:  63% . 36%

[illegible]

- Chain h:  97%

MET
SER
SER
SER
ALA
PRO
K7
N361
THR
SER
LYS
LYS
ILE
ALA

- Chain i:  92% • 7%

X68  
 X277  
 C569  
 GLU  
 ASP  
 ASP  
 ASP  
 SER  
 LYS  
 ASP  
 GLU  
 SER  
 SER  
 ILE  
 GLU  
 GLU  
 ASP  
 VAL  
 ASP  
 ASP  
 ASP  
 PHE  
 PHE  
 ARG  
 LYS  
 LYS  
 ASP  
 GLY  
 THR  
 VAL  
 THR  
 LYS  
 GLU  
 GLY  
 ASN  
 LYS  
 ASP  
 HIS  
 ALA  
 V60A  
 X705  
 X706  
 X784  
 X809  
 Y1015

- Chain j:  84% 16%

MET	VAL	GLU	ASP	SER	ARG	VAL	ARG	ASP	ALA	LEU	LYS	GLY	GLN	LYS	ALA	LEU	PRO	GLN	ALA	P28	155	SER	SER	ASN	GLY	PRO	GLY	ASP	LYS	Y65	L77	LVS	LVS	MET	GLY	ARG	D83	P52
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- Chain k:  86% • 13%

MET	VAL	GLU	ASP	SER	ARG	VAL	ARG	ASP	ALA	LEU	LYS	GLY	GLY	ASP	GLN	LYS	ALA	LEU	PRO	ALA	SER	SER	LEU	VAL	PRO	GLN	GLN	ALA	P28	I55	SER	SER	ASN	GLY	PRO	GLY	GLY	D63	D83	I84	I252
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------

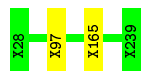
- Molecule 43: Utp24

Chain l:  100%

There are no outlier residues recorded for this chain.

- Molecule 44: Imp4

Chain m:  99%



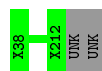
- Molecule 45: Utp30

Chain n:  100%

There are no outlier residues recorded for this chain.

- Molecule 46: Unassigned KH domain

Chain o:  99%



- Molecule 47: Utp20

Chain p:  96%



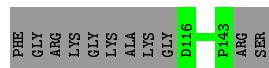
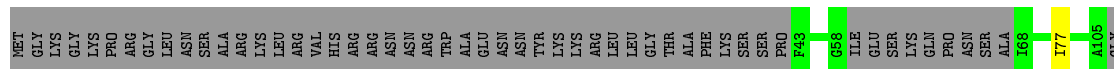
- Molecule 48: Repeat protein 2

Chain q:  100%

There are no outlier residues recorded for this chain.

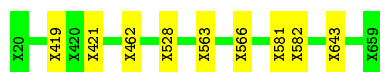
- Molecule 49: 40S ribosomal protein S23-A

Chain r:  56% 43%



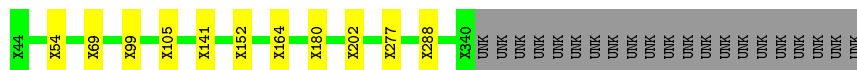
- Molecule 50: Beta-propeller 1

Chain s:  97%



- Molecule 50: Beta-propeller 1

Chain t: 89% 7%



- Molecule 50: Beta-propeller 1

Chain u: 94% 6%



- Molecule 51: Repeat protein 1

Chain v: 97% 2%



- Molecule 52: Unassigned protein helices

Chain y: 93% 7%



- Molecule 53: Unassigned RNA helices

Chain z: 100%

There are no outlier residues recorded for this chain.

## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	33813	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	1.56	Depositor
Minimum defocus (nm)	600	Depositor
Maximum defocus (nm)	2600	Depositor
Magnification	22500	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor



## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# $ Z  > 2$	RMSZ	# $ Z  > 2$
1	0	1.11	0/268	1.42	5/417 (1.2%)
12	C	0.50	0/565	0.77	0/784
13	D	0.30	0/720	0.62	0/1001
14	E	0.34	0/623	0.63	0/864
15	F	0.38	0/443	0.73	0/615
16	G	0.46	0/305	0.80	0/423
2	1	1.25	39/12260 (0.3%)	1.71	368/19055 (1.9%)
27	T	0.51	0/3055	0.69	0/4243
3	2	1.66	26/2443 (1.1%)	1.69	73/3787 (1.9%)
38	e	0.25	0/605	0.56	0/843
38	f	0.26	0/565	0.57	0/787
39	g	0.41	0/1793	0.64	2/2485 (0.1%)
4	3	0.57	0/1061	0.81	1/1472 (0.1%)
40	h	0.40	0/1741	0.59	0/2416
41	i	0.32	0/265	0.50	0/367
42	j	0.36	0/1044	0.62	0/1452
42	k	0.39	0/1079	0.57	0/1502
49	r	0.37	0/399	0.58	0/549
5	5	0.63	0/1182	0.92	0/1638
6	6	0.70	2/998 (0.2%)	1.22	7/1388 (0.5%)
7	7	0.51	0/922	0.85	0/1285
8	8	0.65	0/847	0.93	0/1173
9	9	0.56	0/914	0.82	0/1272
All	All	0.95	67/34097 (0.2%)	1.28	456/49818 (0.9%)

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
12	C	0	1
15	F	0	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
16	G	0	1
17	H	0	4
18	I	0	11
19	J	0	2
19	K	0	3
20	M	0	3
21	N	0	4
22	O	0	12
24	Q	0	4
26	S	0	4
27	T	0	2
28	U	0	1
3	2	0	1
30	W	0	10
31	X	0	2
33	Z	0	2
34	a	0	6
35	b	0	1
36	c	0	3
37	d	0	3
39	g	0	2
41	i	0	5
44	m	0	2
47	p	0	38
50	s	0	9
50	t	0	11
50	u	0	1
51	v	0	13
52	y	0	33
6	6	0	3
9	9	0	3
All	All	0	202

The worst 5 of 67 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	2	308	U	O3'-P	-63.84	0.84	1.61
2	1	1615	C	O3'-P	-41.94	1.10	1.61
6	6	145	ASP	C-N	-12.45	1.05	1.34
2	1	337	G	C2-N2	10.63	1.45	1.34
3	2	118	A	C1'-N9	-10.57	1.32	1.46

The worst 5 of 456 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	1	1615	C	OP2-P-O3'	-30.61	37.85	105.20
6	6	145	ASP	O-C-N	-22.20	87.17	122.70
3	2	118	A	O4'-C1'-N9	-21.15	91.28	108.20
2	1	1615	C	OP1-P-O3'	17.13	142.90	105.20
2	1	1615	C	O3'-P-O5'	-16.35	72.93	104.00

There are no chirality outliers.

5 of 202 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	2	118	A	Sidechain
6	6	145	ASP	Mainchain
6	6	44	ASN	Peptide
6	6	99	MET	Peptide
9	9	88	GLU	Peptide

## 5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	0	4476	0	2976	90	0
2	1	10978	0	5538	366	0
3	2	2468	0	1299	290	0
4	3	1063	0	481	38	0
5	5	1185	0	530	65	0
6	6	1000	0	473	24	0
7	7	923	0	406	6	0
8	8	849	0	407	54	0
9	9	915	0	421	12	0
10	A	468	0	318	0	0
11	B	1296	0	874	9	0
12	C	566	0	265	9	0
13	D	721	0	316	27	0
14	E	624	0	285	2	0
15	F	444	0	206	6	0
16	G	306	0	134	7	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
17	H	2680	0	588	25	0
18	I	880	0	200	15	0
19	J	535	0	118	8	0
19	K	525	0	117	10	0
20	M	1275	0	303	35	0
21	N	2678	0	575	40	0
22	O	3153	0	715	20	0
23	P	1530	0	339	21	0
24	Q	3503	0	746	65	0
25	R	3539	0	778	12	0
26	S	1228	0	295	17	0
27	T	3691	0	1493	11	0
28	U	1398	0	303	3	0
29	V	1298	0	300	11	0
30	W	520	0	116	13	0
31	X	3155	0	637	37	0
32	Y	3160	0	638	72	0
33	Z	748	0	151	6	0
34	a	1544	0	319	0	0
35	b	1687	0	344	0	0
36	c	1088	0	213	0	0
37	d	1064	0	215	0	0
38	e	606	0	291	0	0
38	f	566	0	271	0	0
39	g	1799	0	805	0	0
40	h	1742	0	785	0	0
41	i	2347	0	569	0	0
42	j	1047	0	454	0	0
42	k	1081	0	470	0	0
43	l	613	0	129	0	0
44	m	775	0	180	0	0
45	n	791	0	173	0	0
46	o	858	0	176	0	0
47	p	4620	0	1041	0	0
48	q	1860	0	419	0	0
49	r	402	0	178	0	0
50	s	1429	0	317	0	0
50	t	1076	0	45	0	0
50	u	1346	0	300	0	0
51	v	2885	0	641	0	0
52	y	2535	0	545	0	0
53	z	912	0	610	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	98451	0	32831	1174	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 13.

The worst 5 of 1174 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:1:252:U:H5"	5:5:133:LYS:CA	1.32	1.56
23:P:117:UNK:HA	23:P:141:UNK:CB	22.97	1.52
2:1:252:U:C5'	5:5:133:LYS:HA	1.40	1.47
32:Y:332:UNK:C	32:Y:333:UNK:N	1.79	1.45
3:2:320:G:H2'	3:2:321:C:C6	1.49	1.45

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

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

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	3	212/236 (90%)	180 (85%)	20 (9%)	12 (6%)	2	24
5	5	235/261 (90%)	199 (85%)	19 (8%)	17 (7%)	1	19
6	6	198/225 (88%)	148 (75%)	36 (18%)	14 (7%)	1	19
7	7	184/190 (97%)	143 (78%)	23 (12%)	18 (10%)	1	12
8	8	169/200 (84%)	150 (89%)	14 (8%)	5 (3%)	5	38
9	9	183/197 (93%)	152 (83%)	17 (9%)	14 (8%)	1	18
12	C	113/143 (79%)	107 (95%)	6 (5%)	0	100	100
13	D	144/156 (92%)	125 (87%)	19 (13%)	0	100	100
14	E	125/130 (96%)	113 (90%)	11 (9%)	1 (1%)	22	66

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
15	F	88/135 (65%)	72 (82%)	13 (15%)	3 (3%)	4	36
16	G	60/67 (90%)	53 (88%)	7 (12%)	0	100	100
27	T	613/781 (78%)	565 (92%)	48 (8%)	0	100	100
38	e	120/126 (95%)	118 (98%)	2 (2%)	0	100	100
38	f	112/126 (89%)	110 (98%)	2 (2%)	0	100	100
39	g	353/573 (62%)	337 (96%)	15 (4%)	1 (0%)	44	81
40	h	353/367 (96%)	340 (96%)	13 (4%)	0	100	100
41	i	50/511 (10%)	48 (96%)	2 (4%)	0	100	100
42	j	205/252 (81%)	190 (93%)	15 (7%)	0	100	100
42	k	214/252 (85%)	201 (94%)	11 (5%)	2 (1%)	20	63
49	r	76/145 (52%)	66 (87%)	9 (12%)	1 (1%)	14	56
All	All	3807/5073 (75%)	3417 (90%)	302 (8%)	88 (2%)	11	44

5 of 88 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	3	153	VAL
4	3	173	PRO
4	3	174	LYS
5	5	24	SER
5	5	95	THR

### 5.3.2 Protein sidechains ⓘ

There are no protein residues with a non-rotameric sidechain to report in this entry.

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	0	10/364 (2%)	1 (10%)	0
2	1	496/1800 (27%)	119 (23%)	22 (4%)
3	2	95/126 (75%)	20 (21%)	3 (3%)
All	All	601/2290 (26%)	140 (23%)	25 (4%)

5 of 140 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	0	287	G
2	1	57	G
2	1	60	U
2	1	67	A
2	1	68	A

5 of 25 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	1	352	A
2	1	417	A
3	2	317	A
2	1	400	A
2	1	512	A

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

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

#### 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

#### 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

#### 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
47	p	58
1	0	45
51	v	31

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Number of breaks
22	O	27
48	q	23
25	R	20
52	y	18
24	Q	18
17	H	16
41	i	16
23	P	15
26	S	13
21	N	11
29	V	11
20	M	11
18	I	11
45	n	10
11	B	9
3	2	8
2	1	8
31	X	7
34	a	7
44	m	7
53	z	7
32	Y	7
30	W	6
50	t	6
35	b	6
50	s	6
19	K	5
19	J	5
10	A	5
28	U	5
50	u	4
6	6	3
27	T	3
43	l	3
46	o	2
37	d	2
36	c	1

The worst 5 of 476 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	y	204:UNK	C	213:UNK	N	135.82
1	y	180:UNK	C	185:UNK	N	133.64

*Continued on next page...*



*Continued from previous page...*

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	y	618:UNK	C	625:UNK	N	128.71
1	z	950:N	O3'	1110:N	P	127.24
1	y	144:UNK	C	150:UNK	N	117.41