WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 3, 2016 – 04:32 PM EDT

PDB ID	:	5GIP
Title	:	Crystal structure of box C/D RNP with 13 nt guide regions and 11 nt substrates
Authors	:	Yang, Z.; Lin, J.; Ye, K.
Deposited on	:	2016-06-24
Resolution	:	3.13 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027939
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027939

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.13 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.
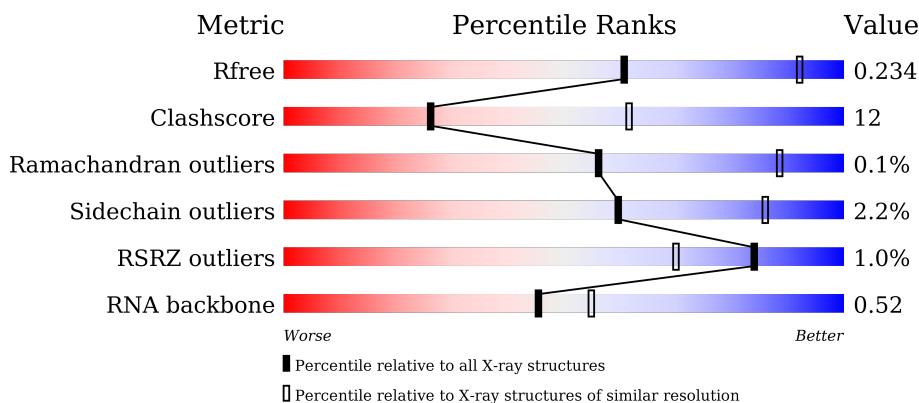
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1112 (3.16-3.08)
Clashscore	102246	1218 (3.16-3.08)
Ramachandran outliers	100387	1175 (3.16-3.08)
Sidechain outliers	100360	1175 (3.16-3.08)
RSRZ outliers	91569	1114 (3.16-3.08)
RNA backbone	2183	1010 (3.54-2.70)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=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 <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	388	
1	B	388	
1	K	388	
1	L	388	

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	C	130	2% 63% 29% · 6%
2	D	130	2% 69% 23% · 6%
2	M	130	2% 67% 26% · 6%
2	N	130	2% 65% 28% · 6%
3	E	232	68% 28% · ·
3	F	232	% 72% 23% · ·
3	O	232	73% 22% · ·
3	P	232	70% 25% · · ·
4	G	41	49% 24% · 24%
4	H	41	2% 44% 27% · 27%
4	Q	41	5% 49% 17% 10% 24%
4	R	41	37% 24% 10% 29%
5	I	11	73% 27%
5	J	11	73% 27%
5	S	11	64% 36%
5	T	11	73% 27%
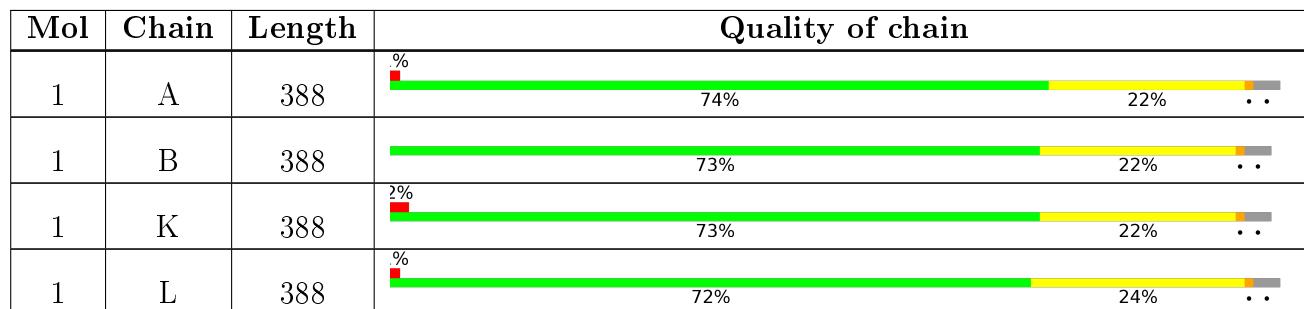
The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	SAH	O	301	-	-	X	-

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 26608 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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 C/D box methylation guide ribonucleoprotein complex aNOP56 subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	375	Total	C	N	O	S	0	0	0
			2993	1902	528	558	5			
1	B	375	Total	C	N	O	S	0	0	0
			2993	1902	528	558	5			
1	K	375	Total	C	N	O	S	0	0	0
			2993	1902	528	558	5			
1	L	375	Total	C	N	O	S	0	0	0
			2993	1902	528	558	5			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP A0A0E3MJI1
A	2	VAL	-	expression tag	UNP A0A0E3MJI1
A	3	LYS	-	expression tag	UNP A0A0E3MJI1
A	381	HIS	-	expression tag	UNP A0A0E3MJI1
A	382	HIS	-	expression tag	UNP A0A0E3MJI1
A	383	HIS	-	expression tag	UNP A0A0E3MJI1
A	384	HIS	-	expression tag	UNP A0A0E3MJI1
A	385	HIS	-	expression tag	UNP A0A0E3MJI1
A	386	HIS	-	expression tag	UNP A0A0E3MJI1
A	387	HIS	-	expression tag	UNP A0A0E3MJI1
A	388	HIS	-	expression tag	UNP A0A0E3MJI1
B	1	MET	-	initiating methionine	UNP A0A0E3MJI1
B	2	VAL	-	expression tag	UNP A0A0E3MJI1
B	3	LYS	-	expression tag	UNP A0A0E3MJI1
B	381	HIS	-	expression tag	UNP A0A0E3MJI1
B	382	HIS	-	expression tag	UNP A0A0E3MJI1
B	383	HIS	-	expression tag	UNP A0A0E3MJI1
B	384	HIS	-	expression tag	UNP A0A0E3MJI1
B	385	HIS	-	expression tag	UNP A0A0E3MJI1
B	386	HIS	-	vexpression tag	UNP A0A0E3MJI1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	387	HIS	-	expression tag	UNP A0A0E3MJI1
B	388	HIS	-	expression tag	UNP A0A0E3MJI1
K	1	MET	-	initiating methionine	UNP A0A0E3MJI1
K	2	VAL	-	expression tag	UNP A0A0E3MJI1
K	3	LYS	-	expression tag	UNP A0A0E3MJI1
K	381	HIS	-	expression tag	UNP A0A0E3MJI1
K	382	HIS	-	expression tag	UNP A0A0E3MJI1
K	383	HIS	-	expression tag	UNP A0A0E3MJI1
K	384	HIS	-	expression tag	UNP A0A0E3MJI1
K	385	HIS	-	expression tag	UNP A0A0E3MJI1
K	386	HIS	-	expression tag	UNP A0A0E3MJI1
K	387	HIS	-	expression tag	UNP A0A0E3MJI1
K	388	HIS	-	expression tag	UNP A0A0E3MJI1
L	1	MET	-	initiating methionine	UNP A0A0E3MJI1
L	2	VAL	-	expression tag	UNP A0A0E3MJI1
L	3	LYS	-	expression tag	UNP A0A0E3MJI1
L	381	HIS	-	expression tag	UNP A0A0E3MJI1
L	382	HIS	-	expression tag	UNP A0A0E3MJI1
L	383	HIS	-	expression tag	UNP A0A0E3MJI1
L	384	HIS	-	expression tag	UNP A0A0E3MJI1
L	385	HIS	-	expression tag	UNP A0A0E3MJI1
L	386	HIS	-	expression tag	UNP A0A0E3MJI1
L	387	HIS	-	expression tag	UNP A0A0E3MJI1
L	388	HIS	-	expression tag	UNP A0A0E3MJI1

- Molecule 2 is a protein called 50S ribosomal protein L7Ae.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	122	Total	C	N	O	S	0	0	0
			927	588	156	181	2			
2	D	122	Total	C	N	O	S	0	0	0
			927	588	156	181	2			
2	M	122	Total	C	N	O	S	0	0	0
			927	588	156	181	2			
2	N	122	Total	C	N	O	S	0	0	0
			927	588	156	181	2			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	1	MET	-	initiating methionine	UNP A0A0E3JZF7
C	2	ASP	-	expression tag	UNP A0A0E3JZF7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	3	ALA	-	expression tag	UNP A0A0E3JZF7
C	4	MET	-	expression tag	UNP A0A0E3JZF7
C	5	SER	-	expression tag	UNP A0A0E3JZF7
D	1	MET	-	initiating methionine	UNP A0A0E3JZF7
D	2	ASP	-	expression tag	UNP A0A0E3JZF7
D	3	ALA	-	expression tag	UNP A0A0E3JZF7
D	4	MET	-	expression tag	UNP A0A0E3JZF7
D	5	SER	-	expression tag	UNP A0A0E3JZF7
M	1	MET	-	initiating methionine	UNP A0A0E3JZF7
M	2	ASP	-	expression tag	UNP A0A0E3JZF7
M	3	ALA	-	expression tag	UNP A0A0E3JZF7
M	4	MET	-	expression tag	UNP A0A0E3JZF7
M	5	SER	-	expression tag	UNP A0A0E3JZF7
N	1	MET	-	initiating methionine	UNP A0A0E3JZF7
N	2	ASP	-	expression tag	UNP A0A0E3JZF7
N	3	ALA	-	expression tag	UNP A0A0E3JZF7
N	4	MET	-	expression tag	UNP A0A0E3JZF7
N	5	SER	-	expression tag	UNP A0A0E3JZF7

- Molecule 3 is a protein called Fibrillarin-like rRNA/tRNA 2'-O-methyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	227	Total	C	N	O	S	0	0	0
			1829	1175	309	341	4			
3	F	227	Total	C	N	O	S	0	0	0
			1829	1175	309	341	4			
3	O	227	Total	C	N	O	S	0	0	0
			1829	1175	309	341	4			
3	P	227	Total	C	N	O	S	0	0	0
			1829	1175	309	341	4			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	1	MET	-	initiating methionine	UNP A0A0E3JUC9
E	2	ALA	-	expression tag	UNP A0A0E3JUC9
F	1	MET	-	initiating methionine	UNP A0A0E3JUC9
F	2	ALA	-	expression tag	UNP A0A0E3JUC9
O	1	MET	-	initiating methionine	UNP A0A0E3JUC9
O	2	ALA	-	expression tag	UNP A0A0E3JUC9
P	1	MET	-	initiating methionine	UNP A0A0E3JUC9
P	2	ALA	-	expression tag	UNP A0A0E3JUC9

• Molecule 4 is a RNA chain called C/D RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	G	31	Total	C	N	O	P	0	0	0
			662	296	118	217	31			
4	H	30	Total	C	N	O	P	0	0	0
			642	287	116	209	30			
4	Q	31	Total	C	N	O	P	0	0	0
			662	296	118	217	31			
4	R	29	Total	C	N	O	P	0	0	0
			622	278	113	202	29			

• Molecule 5 is a RNA chain called substrate.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	I	11	Total	C	N	O	P	0	0	0
			230	104	39	77	10			
5	J	11	Total	C	N	O	P	0	0	0
			230	104	39	77	10			
5	S	11	Total	C	N	O	P	0	0	0
			230	104	39	77	10			
5	T	11	Total	C	N	O	P	0	0	0
			230	104	39	77	10			

• Molecule 6 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula: $C_{14}H_{20}N_6O_5S$).
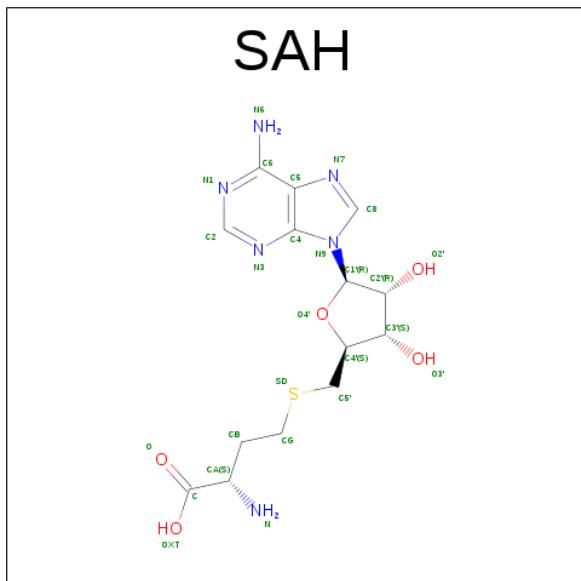
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	E	1	Total 26	C 14	N 6	O 5	S 1	0	0
6	F	1	Total 26	C 14	N 6	O 5	S 1	0	0
6	O	1	Total 26	C 14	N 6	O 5	S 1	0	0
6	P	1	Total 26	C 14	N 6	O 5	S 1	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron 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 dot above a residue indicates a poor fit to the electron density (RSRZ > 2). 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: C/D box methylation guide ribonucleoprotein complex aNOP56 subunit
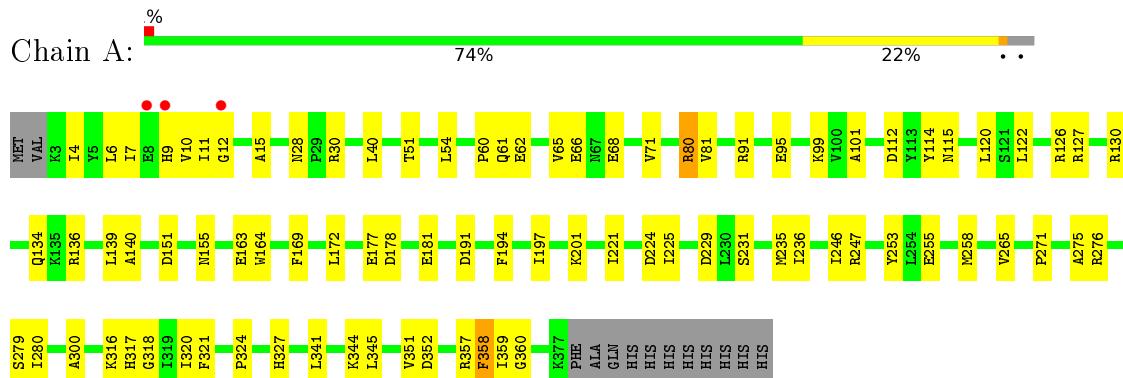
• Molecule 1: C/D box methylation guide ribonucleoprotein complex aNOP56 subunit

• Molecule 1: C/D box methylation guide ribonucleoprotein complex aNOP56 subunit

• Molecule 1: C/D box methylation guide ribonucleoprotein complex aNOP56 subunit

Chain L:

• Molecule 2: 50S ribosomal protein L7Ae

Chain C:

• Molecule 2: 50S ribosomal protein L7Ae

Chain D:

• Molecule 2: 50S ribosomal protein L7Ae

Chain M:

• Molecule 2: 50S ribosomal protein L7Ae

Chain N:

MET ASP ALA MET SER LYS A7 D17 K21 A25 V26 R27 K30 E31 S32 K37 G38 T39 K52 L53 V54 I55 Q61 P62 L72 C73 D74 K77 Y80 V81 G90 E91 A92 C93 G94 L95 Q96 V97 A98 T99 A100 S101 A102 L105 E106 P107 G108 E109

L113 I117 I118 K119 R120 V121 N122 E123 I124 K125 G126 K127 T128 SER SER
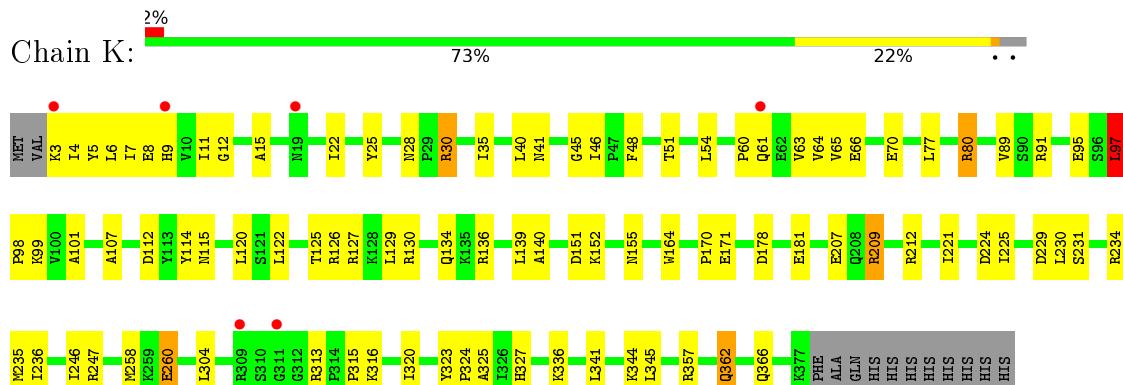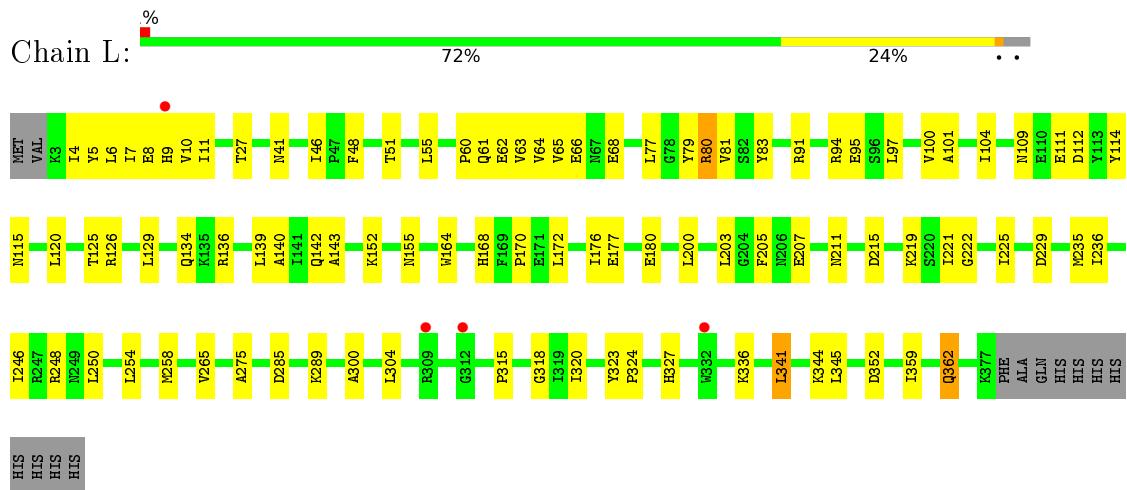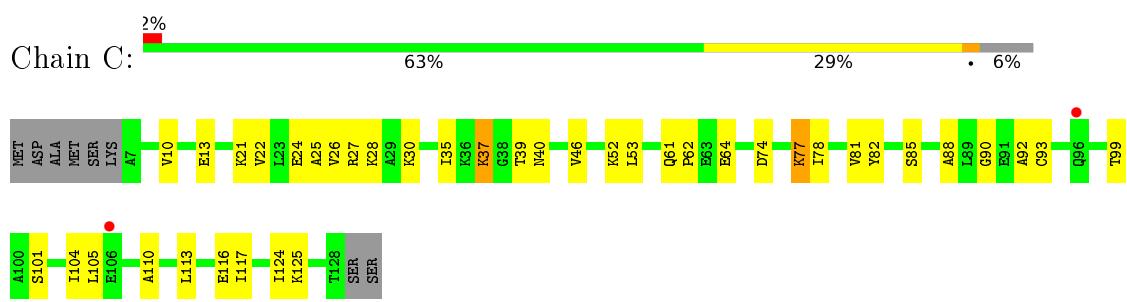
• Molecule 3: Fibrillarin-like rRNA/tRNA 2'-O-methyltransferase

Chain E: 68% 28%

MET ALA GLU VAL I5 N14 C18 E19 D22 F25 N31 L32 V33 R42 K45 E50 Y51 R52 E53 W54 R58 S59 K60 L61 A62 K67 G68 L69 K70 T71 K76 G77 T78 L81 S92 H93 V94 S95 D96 I97 I98 E99 L100 K103 F109

R112 R115 E116 L117 L118 Q122 R123 N126 L130 P137 Q138 S139 Y140 K141 S142 V147 D148 V149 L150 Y151 V152 D153 F169 F170 L171 K182 A183 R184 V188 P192 K193 E194 I195 N207 F208 E209 T210 I211 Q212 I213 I214 N215 L216 D217 K221 D222 H223

A224 I225 K229 Y230 K231 GLY
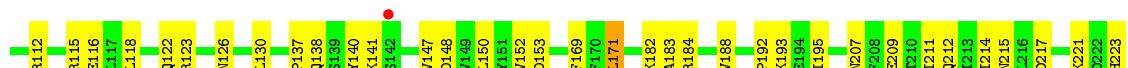
• Molecule 3: Fibrillarin-like rRNA/tRNA 2'-O-methyltransferase

Chain F: 72% 23%

MET ALA GLU VAL I5 T6 N14 I15 Y16 F20 N21 D22 R30 N31 L32 R42 Y46 R52 E53 W54 N55 R58 S59 K60 L61 A62 G63 A64 K67 G68 L69 K70 T71 N72 P73 K76 L81 A86 D96 E99 L100 N101 G102 K103 E108 F109

S110 P111 R112 V113 E116 L117 L118 Q122 N126 P137 Q138 S139 Y140 K141 S142 V147 D148 V149 L150 Y151 V152 D153 N166 F169 F170 L171 K172 K182 A183 R184 N207 Q212 I213 I214 N215 L216 D217 P218 Y219 D220 K221 D222 H223 K231 GLY

• Molecule 3: Fibrillarin-like rRNA/tRNA 2'-O-methyltransferase

Chain O: 73% 22%

MET ALA GLU VAL I5 Q9 T10 N14 C18 F25 C28 N31 L32 V33 E41 R42 L43 I44 E53 W54 F57 R58 S59 K60 L61 A62 K67 K70 T71 N72 P73 G77 T78 K79 V80 L81 S92 S95 L100 N101 G102 K103 F109

R112 R115 L118 Q122 N126 P129 L130 P137 Y140 K141 S142 V147 D148 V149 L150 Y151 V152 F169 F170 L171 M177 L178 L179 K182 E209 Q212 D217 P218 Y219 D220 K221 D222 H223 Y230 K231 GLY

• Molecule 3: Fibrillarin-like rRNA/tRNA 2'-O-methyltransferase

Chain P: 70% 25%

MET ALA GLU VAL I5 Q9 T10 N14 E19 F20 F25 N31 L32 V33 V38 Y39 R42 W54 R58 S59 K60 L61 A62 K67 G68 L69 K76 K79 V80 L81 H93 I97 L100 N101 G102 K103 F109 S110 P111 R112 V113 V114 R115
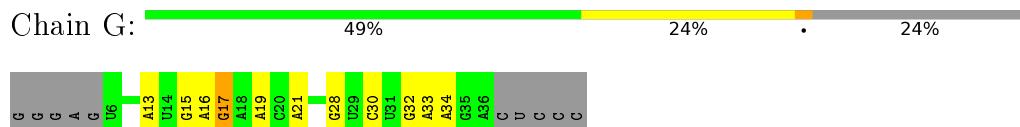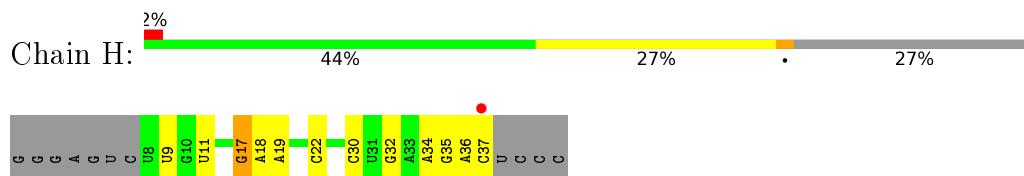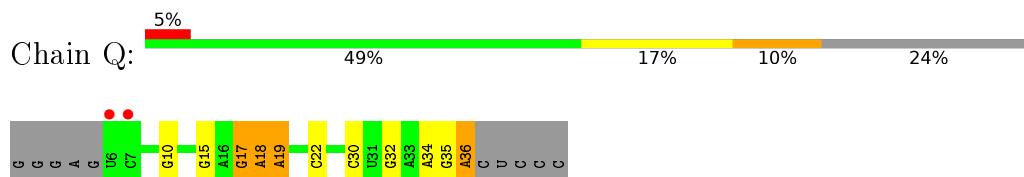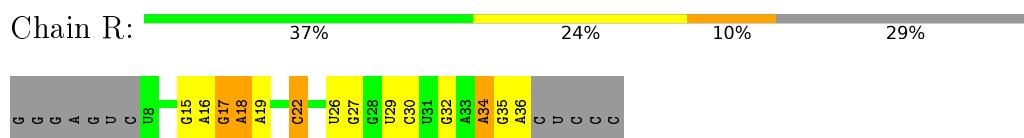
Q122 N126 P129 F136 P137 Q138 S139 Y140 K141 S142 V147 L150 Y151 V152 T160 L171 R184 K190 I195 E199 N207 F208 E209 T210 I211 Q212 I213 I214 N215 L216 D217 K221 D222 H223 A224 I225 S228 K229 Y230 K231 GLY

• Molecule 4: C/D RNA

Chain G:

• Molecule 4: C/D RNA

Chain H:

• Molecule 4: C/D RNA

Chain Q:

• Molecule 4: C/D RNA

Chain R:

• Molecule 5: substrate

Chain I:
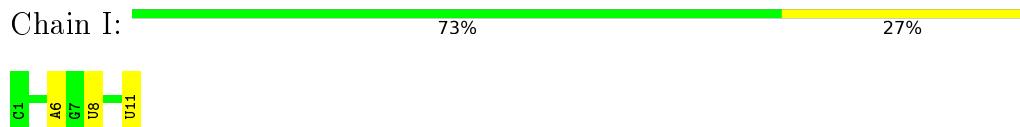
• Molecule 5: substrate
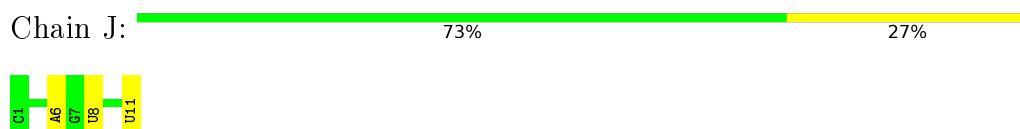
Chain J:

• Molecule 5: substrate

Chain S:

• Molecule 5: substrate

Chain T:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	97.92Å 97.72Å 128.79Å 108.98° 104.42° 89.86°	Depositor
Resolution (Å)	24.86 – 3.13 24.86 – 3.13	Depositor EDS
% Data completeness (in resolution range)	96.9 (24.86-3.13) 86.4 (24.86-3.13)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	5.22 (at 3.11Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.178 , 0.234 0.177 , 0.234	Depositor DCC
R_{free} test set	2005 reflections (2.70%)	DCC
Wilson B-factor (Å^2)	68.3	Xtriage
Anisotropy	0.356	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 35.5	EDS
L-test for twinning[2]	$<|L|> = 0.50, <L^2> = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	26608	wwPDB-VP
Average B, all atoms (Å^2)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.76% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $<|L|>$, $<L^2>$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: SAH

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	0.61	3/3043 (0.1%)	0.69	0/4106
1	B	0.72	13/3043 (0.4%)	0.66	0/4106
1	K	0.65	7/3043 (0.2%)	0.66	1/4106 (0.0%)
1	L	0.65	6/3043 (0.2%)	0.64	0/4106
2	C	0.51	1/936 (0.1%)	0.63	0/1260
2	D	0.56	1/936 (0.1%)	0.64	0/1260
2	M	0.52	0/936	0.65	0/1260
2	N	0.53	1/936 (0.1%)	0.63	0/1260
3	E	0.86	12/1863 (0.6%)	0.69	0/2521
3	F	0.78	8/1863 (0.4%)	0.69	1/2521 (0.0%)
3	O	0.65	5/1863 (0.3%)	0.70	0/2521
3	P	0.79	11/1863 (0.6%)	0.71	1/2521 (0.0%)
4	G	0.70	0/740	1.20	1/1151 (0.1%)
4	H	0.73	0/718	1.21	2/1117 (0.2%)
4	Q	0.67	1/740 (0.1%)	1.22	1/1151 (0.1%)
4	R	0.65	0/696	1.23	2/1083 (0.2%)
5	I	0.69	0/256	1.24	3/397 (0.8%)
5	J	0.74	0/256	1.25	1/397 (0.3%)
5	S	0.76	0/256	1.53	6/397 (1.5%)
5	T	0.83	0/256	1.33	3/397 (0.8%)
All	All	0.68	69/27286 (0.3%)	0.79	22/37638 (0.1%)

The worst 5 of 69 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	42	ARG	CZ-NH2	-9.99	1.20	1.33
3	E	42	ARG	NE-CZ	-9.88	1.20	1.33
1	B	188	ARG	NE-CZ	-9.82	1.20	1.33
3	E	42	ARG	CZ-NH1	-9.44	1.20	1.33
1	L	80	ARG	NE-CZ	-9.25	1.21	1.33

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	S	11	U	C2-N1-C1'	9.63	129.26	117.70
5	S	11	U	C6-N1-C1'	-7.65	110.48	121.20
5	S	11	U	N1-C2-O2	7.22	127.86	122.80
5	T	11	U	N3-C4-O4	-7.19	114.37	119.40
5	S	11	U	O4'-C1'-N1	6.90	113.72	108.20

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

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	A	2993	0	3055	83	0
1	B	2993	0	3055	85	0
1	K	2993	0	3055	95	0
1	L	2993	0	3055	98	0
2	C	927	0	981	33	0
2	D	927	0	981	31	0
2	M	927	0	981	26	0
2	N	927	0	981	28	0
3	E	1829	0	1862	54	0
3	F	1829	0	1862	54	0
3	O	1829	0	1862	56	0
3	P	1829	0	1862	57	0
4	G	662	0	333	6	0
4	H	642	0	323	4	0
4	Q	662	0	333	7	0
4	R	622	0	312	9	0
5	I	230	0	119	2	0
5	J	230	0	119	2	0
5	S	230	0	119	2	0
5	T	230	0	119	2	0
6	E	26	0	19	2	0
6	F	26	0	19	7	0
6	O	26	0	19	12	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	P	26	0	19	7	0
All	All	26608	0	25445	640	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 12.

The worst 5 of 640 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)
3:P:54:TRP:CH2	3:P:61:LEU:HD12	1.90	1.06
1:K:139:LEU:HD11	1:L:225:ILE:CD1	1.94	0.97
1:K:139:LEU:CD1	1:L:225:ILE:HD11	1.95	0.96
1:K:139:LEU:HD11	1:L:225:ILE:HD11	0.99	0.96
1:B:130:ARG:HH12	1:B:134:GLN:HE21	1.15	0.95

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 X-ray entries followed by that with respect to entries of similar resolution.

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	373/388 (96%)	361 (97%)	12 (3%)	0	100	100
1	B	373/388 (96%)	361 (97%)	12 (3%)	0	100	100
1	K	373/388 (96%)	361 (97%)	12 (3%)	0	100	100
1	L	373/388 (96%)	362 (97%)	11 (3%)	0	100	100
2	C	120/130 (92%)	117 (98%)	3 (2%)	0	100	100
2	D	120/130 (92%)	117 (98%)	3 (2%)	0	100	100
2	M	120/130 (92%)	117 (98%)	3 (2%)	0	100	100
2	N	120/130 (92%)	117 (98%)	3 (2%)	0	100	100

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	E	225/232 (97%)	205 (91%)	19 (8%)	1 (0%)	39	77
3	F	225/232 (97%)	208 (92%)	17 (8%)	0	100	100
3	O	225/232 (97%)	207 (92%)	17 (8%)	1 (0%)	39	77
3	P	225/232 (97%)	208 (92%)	16 (7%)	1 (0%)	39	77
All	All	2872/3000 (96%)	2741 (95%)	128 (4%)	3 (0%)	56	88

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	P	14	ASN
3	E	14	ASN
3	O	14	ASN

5.3.2   Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

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	318/330 (96%)	314 (99%)	4 (1%)	76	91
1	B	318/330 (96%)	312 (98%)	6 (2%)	65	88
1	K	318/330 (96%)	311 (98%)	7 (2%)	60	86
1	L	318/330 (96%)	314 (99%)	4 (1%)	76	91
2	C	100/107 (94%)	99 (99%)	1 (1%)	82	93
2	D	100/107 (94%)	99 (99%)	1 (1%)	82	93
2	M	100/107 (94%)	99 (99%)	1 (1%)	82	93
2	N	100/107 (94%)	99 (99%)	1 (1%)	82	93
3	E	202/205 (98%)	196 (97%)	6 (3%)	48	81
3	F	202/205 (98%)	195 (96%)	7 (4%)	43	79
3	O	202/205 (98%)	193 (96%)	9 (4%)	34	71
3	P	202/205 (98%)	194 (96%)	8 (4%)	38	75
All	All	2480/2568 (97%)	2425 (98%)	55 (2%)	60	86

5 of 55 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	171	LEU
1	K	260	GLU
3	P	103	LYS
1	K	61	GLN
1	K	97	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 60 such sidechains are listed below:

Mol	Chain	Res	Type
1	K	9	HIS
1	K	362	GLN
3	P	31	ASN
1	K	155	ASN
1	L	9	HIS

5.3.3   RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
4	G	30/41 (73%)	5 (16%)	0
4	H	29/41 (70%)	6 (20%)	0
4	Q	30/41 (73%)	6 (20%)	0
4	R	28/41 (68%)	6 (21%)	1 (3%)
5	I	10/11 (90%)	0	0
5	J	10/11 (90%)	0	0
5	S	10/11 (90%)	0	0
5	T	10/11 (90%)	0	0
All	All	157/208 (75%)	23 (14%)	1 (0%)

5 of 23 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
4	G	17	G
4	G	19	A
4	G	30	C
4	G	32	G
4	G	34	A

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
4	R	17	G

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

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

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

4 ligands 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	#$|Z| > 2$
6	SAH	E	301	-	22,28,28	1.14	2 (9%)	18,40,40	2.77	4 (22%)
6	SAH	F	301	-	22,28,28	1.07	2 (9%)	18,40,40	2.89	2 (11%)
6	SAH	O	301	-	22,28,28	1.14	2 (9%)	18,40,40	3.07	4 (22%)
6	SAH	P	301	-	22,28,28	1.18	3 (13%)	18,40,40	3.16	6 (33%)

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
6	SAH	E	301	-	-	0/7/31/31	0/3/3/3
6	SAH	F	301	-	-	0/7/31/31	0/3/3/3
6	SAH	O	301	-	-	0/7/31/31	0/3/3/3
6	SAH	P	301	-	-	0/7/31/31	0/3/3/3

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	P	301	SAH	C2'-C1'	-2.11	1.50	1.53
6	F	301	SAH	C2-N1	2.10	1.38	1.33
6	O	301	SAH	C2-N1	2.12	1.38	1.33
6	P	301	SAH	C2-N1	2.27	1.38	1.33
6	E	301	SAH	C2-N1	2.52	1.38	1.33

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	P	301	SAH	N3-C2-N1	-11.78	119.62	128.87
6	O	301	SAH	N3-C2-N1	-11.77	119.63	128.87
6	F	301	SAH	N3-C2-N1	-11.15	120.11	128.87
6	E	301	SAH	N3-C2-N1	-10.25	120.82	128.87
6	P	301	SAH	C4'-O4'-C1'	-3.53	105.90	109.64

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 28 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	E	301	SAH	2	0
6	F	301	SAH	7	0
6	O	301	SAH	12	0
6	P	301	SAH	7	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å²)	Q<0.9
1	A	375/388 (96%)	-0.29	3 (0%)	87	75	31, 58, 89, 122	0
1	B	375/388 (96%)	-0.29	1 (0%)	94	89	32, 59, 97, 119	0
1	K	375/388 (96%)	-0.26	6 (1%)	74	55	37, 66, 100, 128	0
1	L	375/388 (96%)	-0.26	4 (1%)	82	67	38, 67, 93, 106	0
2	C	122/130 (93%)	-0.23	2 (1%)	74	55	53, 82, 104, 113	0
2	D	122/130 (93%)	-0.23	2 (1%)	74	55	46, 73, 93, 103	0
2	M	122/130 (93%)	-0.15	3 (2%)	61	39	53, 74, 101, 122	0
2	N	122/130 (93%)	-0.05	3 (2%)	61	39	53, 76, 105, 112	0
3	E	227/232 (97%)	-0.24	1 (0%)	93	86	44, 64, 85, 92	0
3	F	227/232 (97%)	-0.20	2 (0%)	85	73	40, 64, 88, 103	0
3	O	227/232 (97%)	-0.21	1 (0%)	93	86	47, 67, 85, 101	0
3	P	227/232 (97%)	-0.27	1 (0%)	93	86	42, 57, 78, 106	0
4	G	31/41 (75%)	-0.18	0	100	100	37, 56, 102, 123	0
4	H	30/41 (73%)	-0.23	1 (3%)	50	26	32, 61, 92, 114	0
4	Q	31/41 (75%)	-0.06	2 (6%)	22	8	38, 63, 102, 121	0
4	R	29/41 (70%)	-0.23	0	100	100	41, 57, 86, 113	0
5	I	11/11 (100%)	-0.35	0	100	100	39, 43, 53, 53	0
5	J	11/11 (100%)	-0.48	0	100	100	34, 41, 48, 49	0
5	S	11/11 (100%)	-0.39	0	100	100	39, 45, 49, 59	0
5	T	11/11 (100%)	-0.26	0	100	100	41, 46, 50, 57	0
All	All	3061/3208 (95%)	-0.24	32 (1%)	84	70	31, 65, 96, 128	0

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	96	GLN	3.8
3	P	142	SER	3.5
2	M	109	GLU	3.2
4	H	37	C	3.1
3	E	142	SER	2.9

6.2 Non-standard residues in protein, DNA, RNA chains (i)

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

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
6	SAH	O	301	26/26	0.91	0.20	-0.11	50,58,66,69	0
6	SAH	E	301	26/26	0.95	0.18	-0.15	50,54,69,71	0
6	SAH	F	301	26/26	0.94	0.19	-0.34	43,52,69,74	0
6	SAH	P	301	26/26	0.93	0.18	-0.38	42,52,64,68	0

6.5 Other polymers (i)

There are no such residues in this entry.