



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 13, 2018 – 03:46 pm GMT

PDB ID : 4V4J  
Title : Interactions and Dynamics of the Shine-Dalgarno Helix in the 70S Ribosome.  
Authors : Korostelev, A.; Trakhanov, S.; Asahara, H.; Laurberg, M.; Noller, H.F.  
Deposited on : 2007-07-18  
Resolution : 3.83 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk31020  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk31020

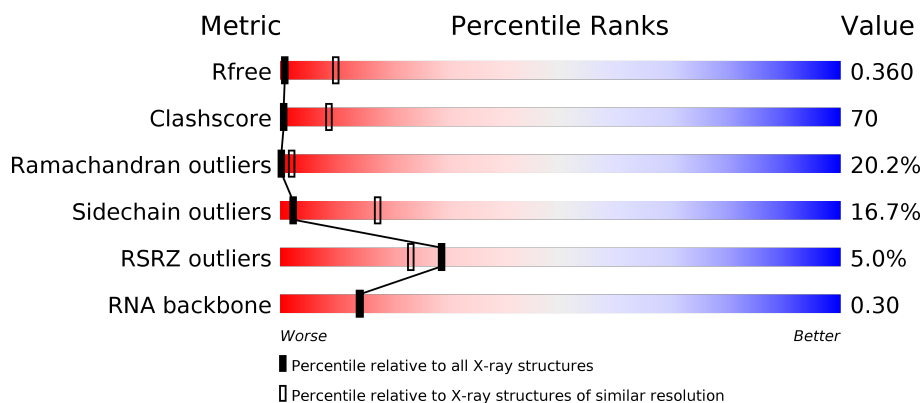
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.83 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))
$R_{free}$	111664	1018 (4.08-3.60)
Clashscore	122126	1087 (4.08-3.60)
Ramachandran outliers	120053	1050 (4.08-3.60)
Sidechain outliers	120020	1044 (4.08-3.60)
RSRZ outliers	108989	1002 (4.10-3.58)
RNA backbone	2636	1084 (4.76-2.90)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	w	2889	<div> <div>4%</div> <div>65%</div> <div>33%</div> <div>.</div> </div>
2	x	120	<div> <div>61%</div> <div>38%</div> <div>.</div> </div>
3	A	229	<div> <div>3%</div> <div>15%</div> <div>28%</div> <div>10%</div> <div>.</div> <div>45%</div> </div>
4	B	276	<div> <div>7%</div> <div>16%</div> <div>55%</div> <div>26%</div> <div>..</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	C	206	
6	D	205	
7	E	182	
8	F	180	
9	G	148	
10	H	140	
11	I	122	
12	J	150	
13	K	141	
14	L	118	
15	M	112	
16	N	146	
17	O	118	
18	P	101	
19	Q	113	
20	R	96	
21	S	110	
22	T	206	
23	U	85	
24	V	98	
25	W	72	
26	X	60	
27	Y	60	
28	Z	49	
29	a	65	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
30	b	37	
31	y	1522	
32	z	77	
33	2	76	
34	3	18	
35	c	256	
36	d	239	
37	e	209	
38	f	162	
39	g	101	
40	h	156	
41	i	138	
42	j	128	
43	k	105	
44	l	129	
45	m	132	
46	n	126	
47	o	61	
48	p	89	
49	q	88	
50	r	105	
51	s	88	
52	t	93	
53	u	106	
54	v	27	

## 2 Entry composition

There are 54 unique types of molecules in this entry. The entry contains 147125 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 RNA chain called 23S LARGE SUBUNIT RIBOSOMAL RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	w	2889	Total	C	N	O	P	0	0	0
			62213	27690	11624	20011	2888			

- Molecule 2 is a RNA chain called 5S LARGE SUBUNIT RIBOSOMAL RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	x	120	Total	C	N	O	P	0	0	0
			2573	1146	476	832	119			

- Molecule 3 is a protein called 50S ribosomal protein L1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	127	Total	C	N	O	S	0	0	0
			996	627	184	184	1			

- Molecule 4 is a protein called 50S ribosomal protein L2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	272	Total	C	N	O	S	0	0	0
			2115	1335	420	357	3			

- Molecule 5 is a protein called 50S ribosomal protein L3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	C	201	Total	C	N	O	S	0	0	0
			1541	974	295	267	5			

- Molecule 6 is a protein called 50S ribosomal protein L4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	D	194	Total	C	N	O	S	0	0	0
			1517	969	283	263	2			

- Molecule 7 is a protein called 50S ribosomal protein L5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	E	180	Total	C	N	O	S	0	0	0
			1468	938	267	259	4			

- Molecule 8 is a protein called 50S ribosomal protein L6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	F	173	Total	C	N	O	S	0	0	0
			1319	839	245	234	1			

- Molecule 9 is a protein called 50S ribosomal protein L9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	G	148	Total	C	N	O	S	0	0	0
			1156	737	204	214	1			

- Molecule 10 is a protein called 50S ribosomal protein L13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	H	138	Total	C	N	O	S	0	0	0
			1103	712	206	182	3			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	24	VAL	MET	CONFLICT	UNP Q72IN1

- Molecule 11 is a protein called 50S ribosomal protein L14.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	I	122	Total	C	N	O	S	0	0	0
			932	587	171	170	4			

- Molecule 12 is a protein called 50S ribosomal protein L15.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	J	146	Total	C	N	O	S	0	0	0
			1114	692	227	193	2			

- Molecule 13 is a protein called 50S ribosomal protein L16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	K	137	Total	C	N	O	S	0	0	0
			1089	698	207	177	7			

- Molecule 14 is a protein called 50S ribosomal protein L17.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	L	118	Total	C	N	O	S	0	0	0
			968	604	203	160	1			

- Molecule 15 is a protein called 50S ribosomal protein L18.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	M	106	Total	C	N	O	S	0	0	0
			846	534	168	144				

- Molecule 16 is a protein called 50S ribosomal protein L19.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
16	N	137	Total	C	N	O	S	0	0	0
			1143	713	234	195	1			

- Molecule 17 is a protein called 50S ribosomal protein L20.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
17	O	117	Total	C	N	O	S	0	0	0
			964	610	202	151	1			

- Molecule 18 is a protein called 50S ribosomal protein L21.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
18	P	101	Total	C	N	O	S	0	0	0
			779	501	142	135	1			

- Molecule 19 is a protein called 50S ribosomal protein L22.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
19	Q	109	Total	C	N	O	S	0	0	0
			868	547	170	150	1			

- Molecule 20 is a protein called 50S ribosomal protein L23.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
20	R	92	Total	C	N	O			
			725	471	131	123	0	0	0

- Molecule 21 is a protein called 50S ribosomal protein L24.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
21	S	103	Total	C	N	O	S		
			793	510	151	126	6	0	0

- Molecule 22 is a protein called 50S ribosomal protein L25.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
22	T	185	Total	C	N	O	S		
			1475	941	262	269	3	0	0

- Molecule 23 is a protein called 50S ribosomal protein L27.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
23	U	76	Total	C	N	O	S		
			605	376	126	102	1	0	0

- Molecule 24 is a protein called LSU ribosomal protein L28P.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
24	V	88	Total	C	N	O			
			694	435	141	118		0	0

- Molecule 25 is a protein called 50S ribosomal protein L29.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
25	W	62	Total	C	N	O	S		
			520	325	102	91	2	0	0

- Molecule 26 is a protein called LSU ribosomal protein L30P.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
26	X	60	Total	C	N	O	S		
			477	303	91	82	1	0	0

- Molecule 27 is a protein called 50S ribosomal protein L32.



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
27	Y	56	Total	C	N	O	S	0	0	0
			436	275	84	72	5			

- Molecule 28 is a protein called 50S ribosomal protein L34.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
28	Z	48	Total	C	N	O	S	0	0	0
			418	257	104	55	2			

- Molecule 29 is a protein called 50S ribosomal protein L35.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
29	a	63	Total	C	N	O	S	0	0	0
			507	326	101	78	2			

- Molecule 30 is a protein called 50S ribosomal protein L36.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
30	b	35	Total	C	N	O	S	0	0	0
			294	181	66	44	3			

- Molecule 31 is a RNA chain called 16S RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
31	y	1514	Total	C	N	O	P	0	0	0
			32546	14494	6022	10517	1513			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
y	466	G	C	CONFLICT	GB 155076

- Molecule 32 is a RNA chain called P-site tRNA<sup>f</sup>MET.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
32	z	77	Total	C	N	O	P	0	0	0
			1639	732	297	534	76			

- Molecule 33 is a RNA chain called E-site tRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
33	2	76	Total	C	N	O	P	0	0	0
			1621	725	293	528	75			

- Molecule 34 is a RNA chain called MRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
34	3	18	Total	C	N	O	P	0	0	0
			390	176	80	117	17			

- Molecule 35 is a protein called 30S ribosomal protein S2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
35	c	234	Total	C	N	O	S	0	0	0
			1900	1213	341	341	5			

- Molecule 36 is a protein called 30S ribosomal protein S3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
36	d	206	Total	C	N	O	S	0	0	0
			1612	1016	314	281	1			

- Molecule 37 is a protein called 30S ribosomal protein S4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
37	e	208	Total	C	N	O	S	0	0	0
			1703	1066	339	291	7			

- Molecule 38 is a protein called 30S ribosomal protein S5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
38	f	150	Total	C	N	O	S	0	0	0
			1146	724	217	201	4			

- Molecule 39 is a protein called 30S ribosomal protein S6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
39	g	101	Total	C	N	O	S	0	0	0
			843	531	155	154	3			

- Molecule 40 is a protein called 30S ribosomal protein S7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
40	h	155	Total	C	N	O	S	0	0	0
			1257	781	252	218	6			

- Molecule 41 is a protein called 30S ribosomal protein S8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
41	i	138	Total	C	N	O	S	0	0	0
			1116	705	215	193	3			

- Molecule 42 is a protein called 30S ribosomal protein S9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
42	j	127	Total	C	N	O	S	0	0	0
			1011	639	198	174				

- Molecule 43 is a protein called 30S ribosomal protein S10.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
43	k	98	Total	C	N	O	S	0	0	0
			794	499	156	138	1			

- Molecule 44 is a protein called 30S ribosomal protein S11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
44	l	116	Total	C	N	O	S	0	0	0
			864	537	164	160	3			

- Molecule 45 is a protein called 30S ribosomal protein S12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
45	m	124	Total	C	N	O	S	0	0	0
			970	611	195	163	1			

- Molecule 46 is a protein called 30S ribosomal protein S13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
46	n	125	Total	C	N	O	S	0	0	0
			997	617	207	171	2			

- Molecule 47 is a protein called 30S ribosomal protein S14 type Z.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
47	o	60	Total	C	N	O	S	0	0	0
			492	312	104	72	4			

- Molecule 48 is a protein called 30S ribosomal protein S15.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
48	p	88	Total	C	N	O	S	0	0	0
			734	459	147	126	2			

- Molecule 49 is a protein called 30S ribosomal protein S16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
49	q	83	Total	C	N	O	S	0	0	0
			700	443	139	117	1			

- Molecule 50 is a protein called 30S ribosomal protein S17.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
50	r	104	Total	C	N	O	S	0	0	0
			857	547	161	147	2			

- Molecule 51 is a protein called 30S ribosomal protein S18.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
51	s	81	Total	C	N	O	0	0	0
			668	423	135	110			

- Molecule 52 is a protein called 30S ribosomal protein S19.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
52	t	80	Total	C	N	O	S	0	0	0
			647	414	119	112	2			

- Molecule 53 is a protein called 30S ribosomal protein S20.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
53	u	99	Total	C	N	O	S	0	0	0
			762	469	162	129	2			

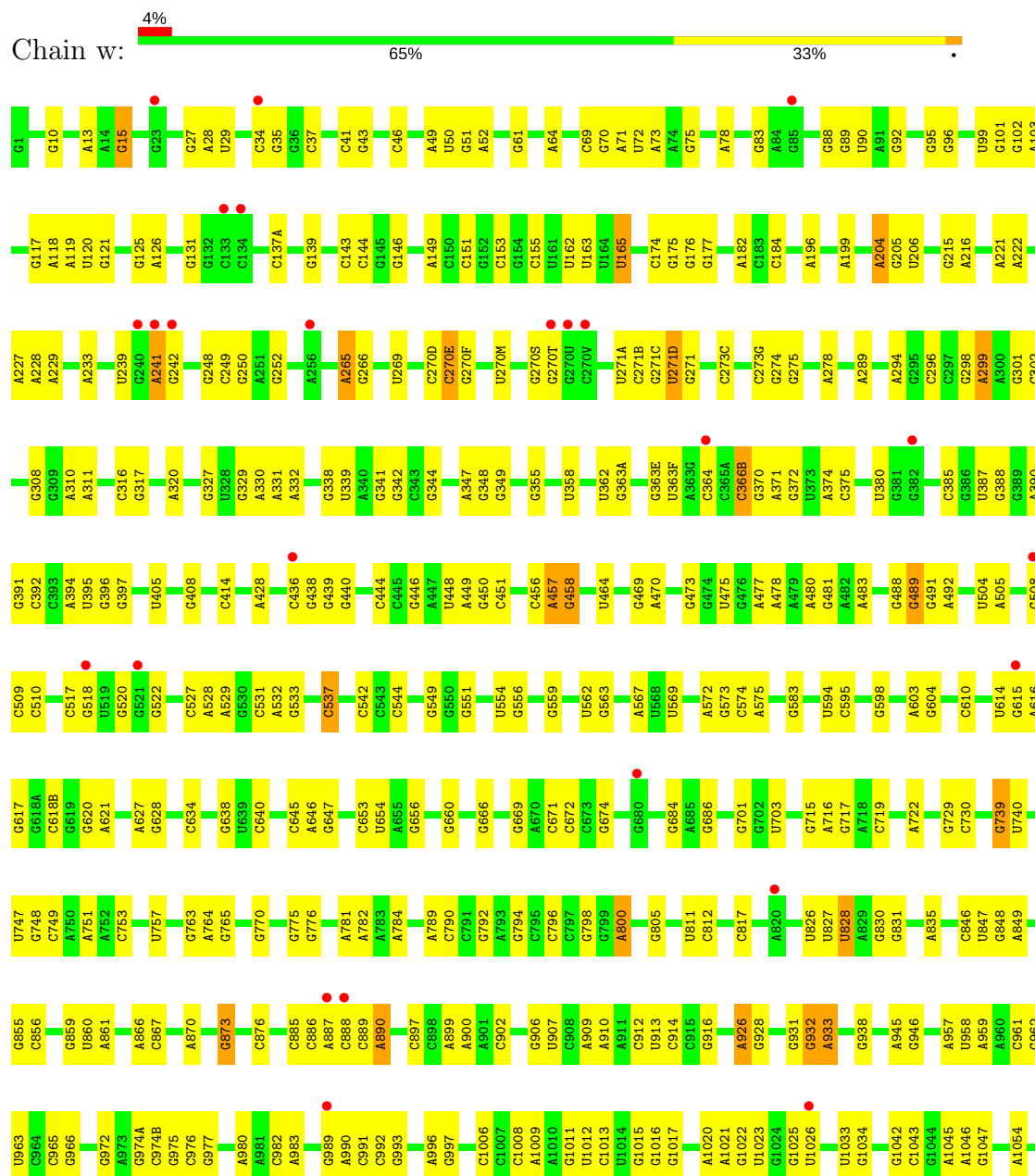
- Molecule 54 is a protein called 30S ribosomal protein Thx.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
54	v	24	Total	C	N	O	0	0	0
			208	128	50	30			

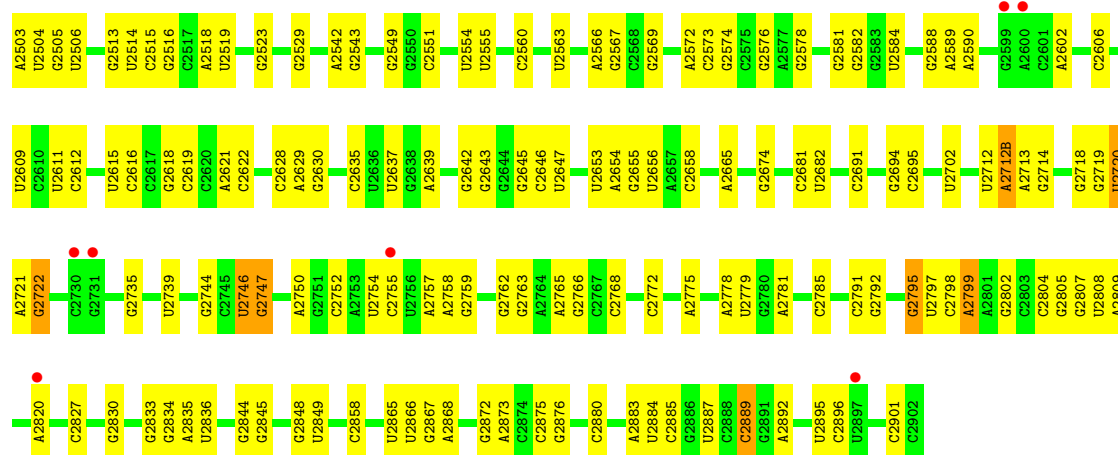
### 3 Residue-property plots

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 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: 23S LARGE SUBUNIT RIBOSOMAL RNA

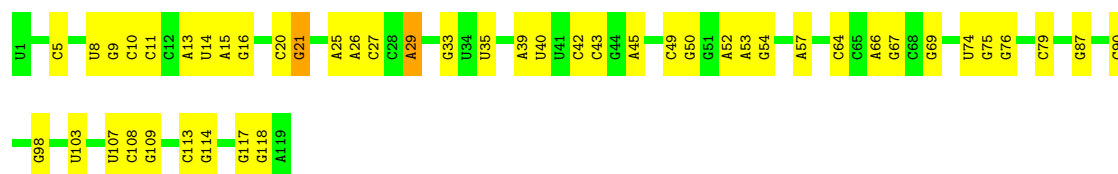






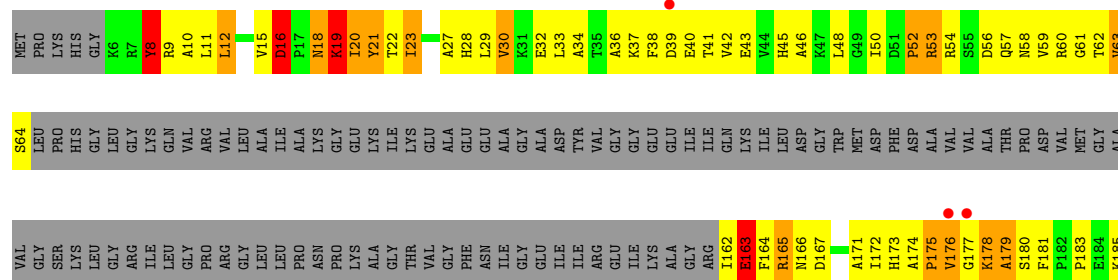
• Molecule 2: 5S LARGE SUBUNIT RIBOSOMAL RNA

Chain x: 61% 38%



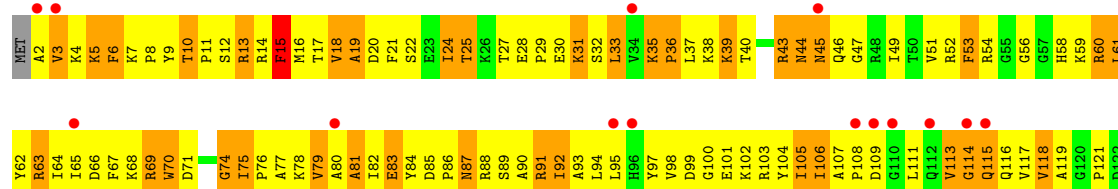
• Molecule 3: 50S ribosomal protein L1

Chain A: 3% 15% 28% 10% 45%

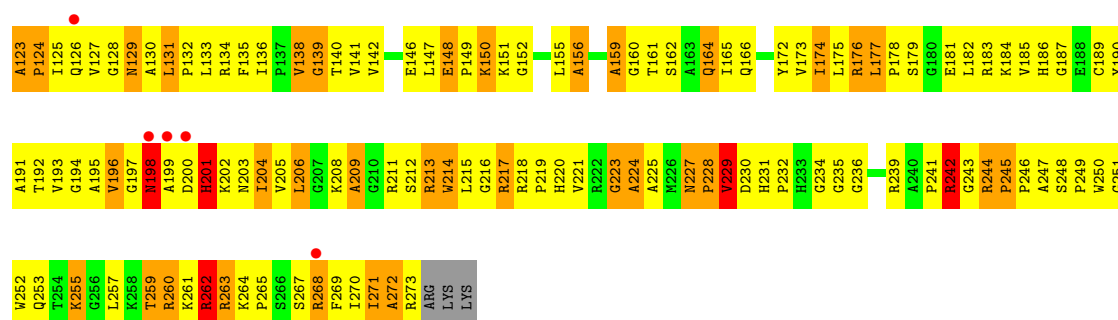


• Molecule 4: 50S ribosomal protein L2

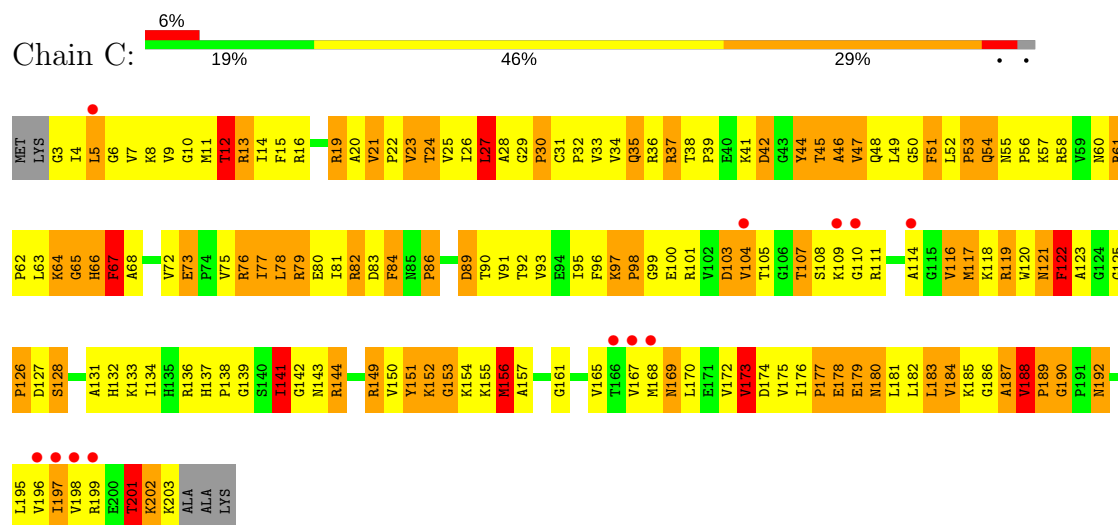
Chain B: 7% 16% 55% 26%



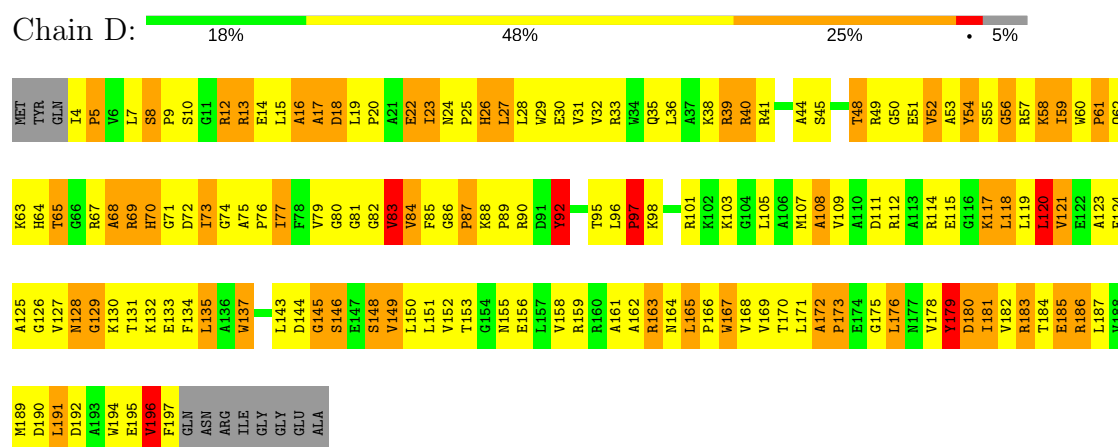




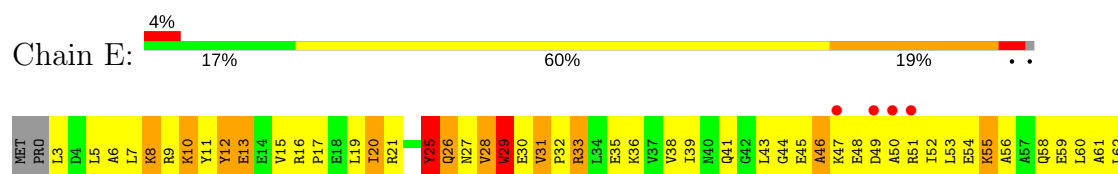
• Molecule 5: 50S ribosomal protein L3

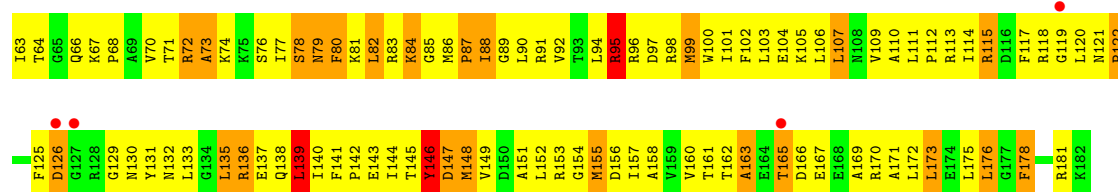


• Molecule 6: 50S ribosomal protein L4

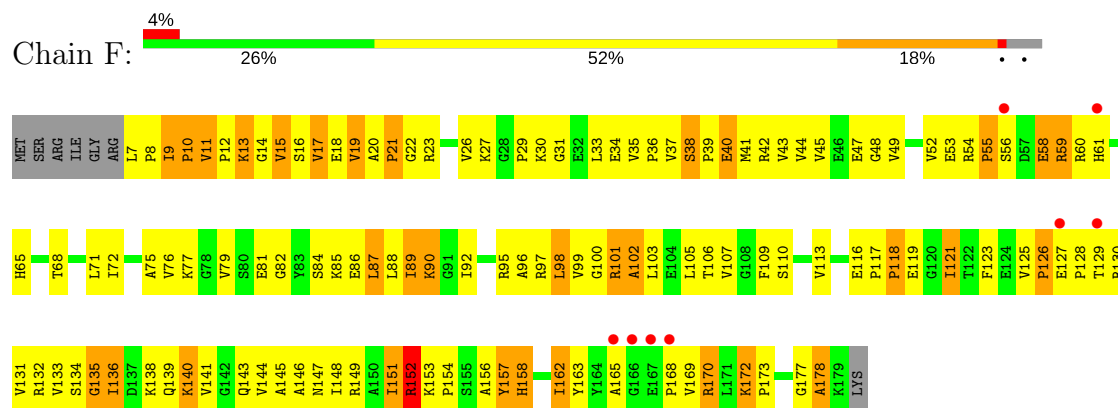


• Molecule 7: 50S ribosomal protein L5

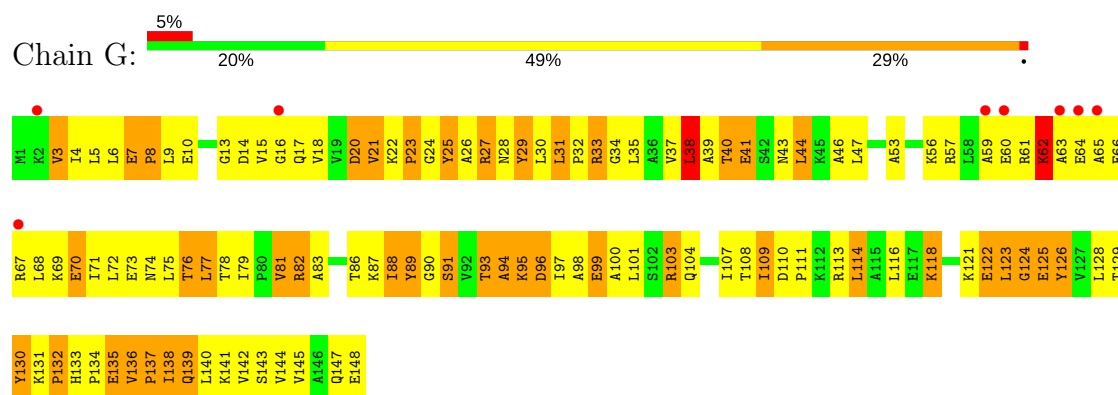




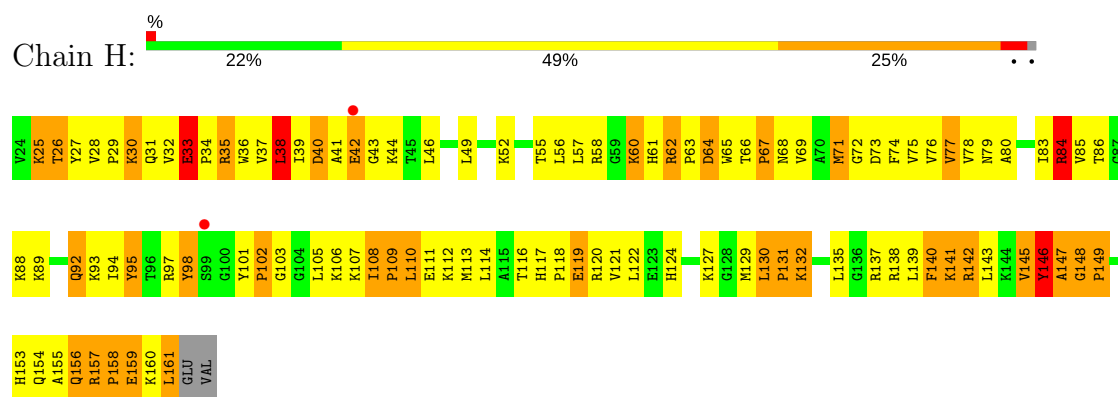
• Molecule 8: 50S ribosomal protein L6



• Molecule 9: 50S ribosomal protein L9

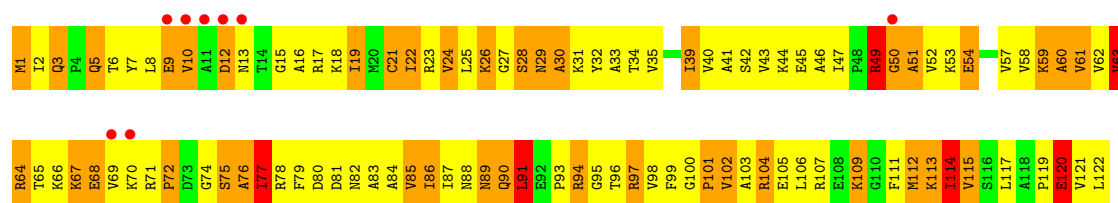


• Molecule 10: 50S ribosomal protein L13

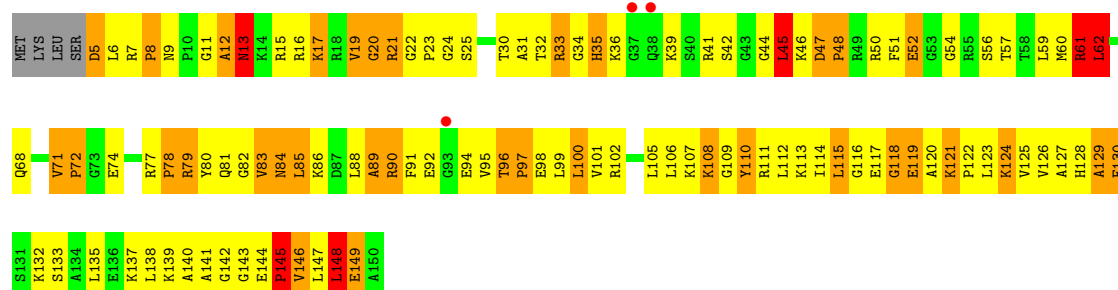
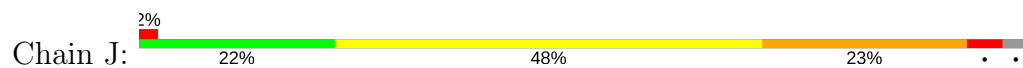


• Molecule 11: 50S ribosomal protein L14

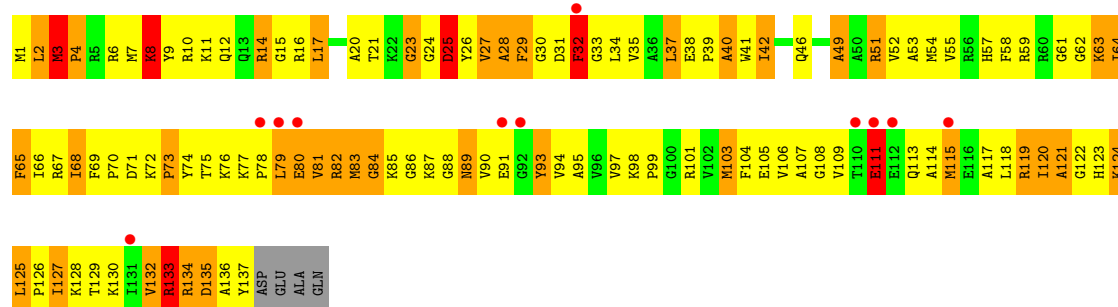
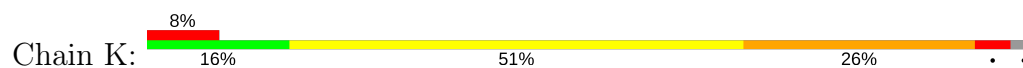




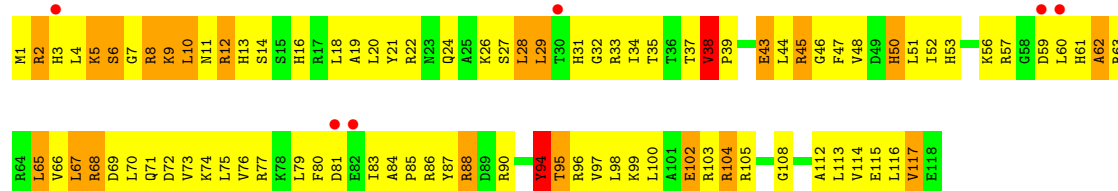
• Molecule 12: 50S ribosomal protein L15



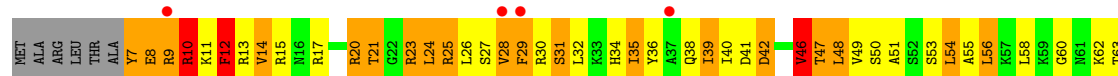
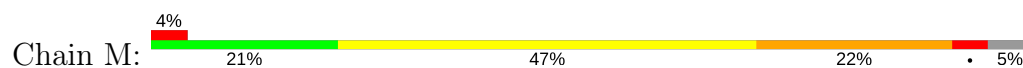
• Molecule 13: 50S ribosomal protein L16

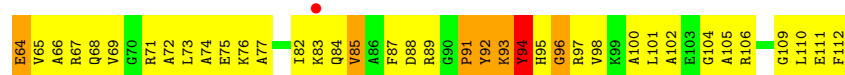


• Molecule 14: 50S ribosomal protein L17

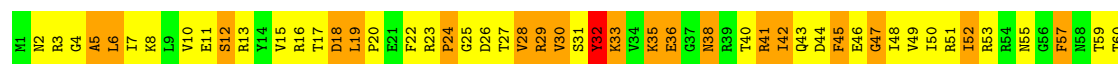
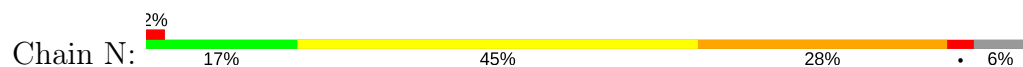


• Molecule 15: 50S ribosomal protein L18

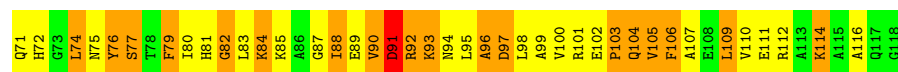




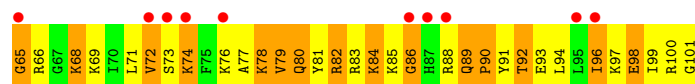
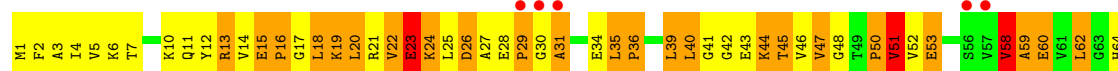
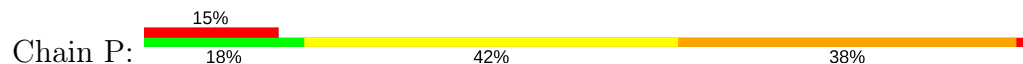
• Molecule 16: 50S ribosomal protein L19



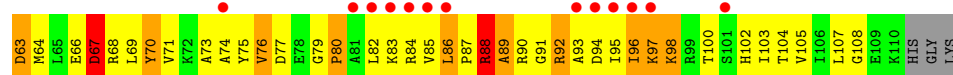
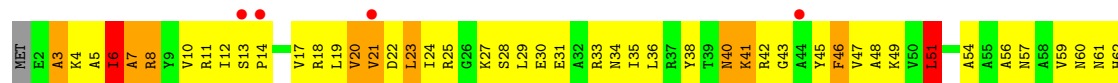
• Molecule 17: 50S ribosomal protein L20



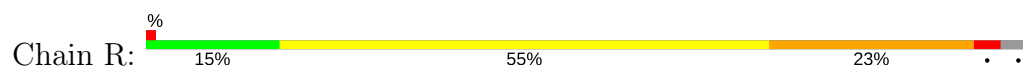
• Molecule 18: 50S ribosomal protein L21



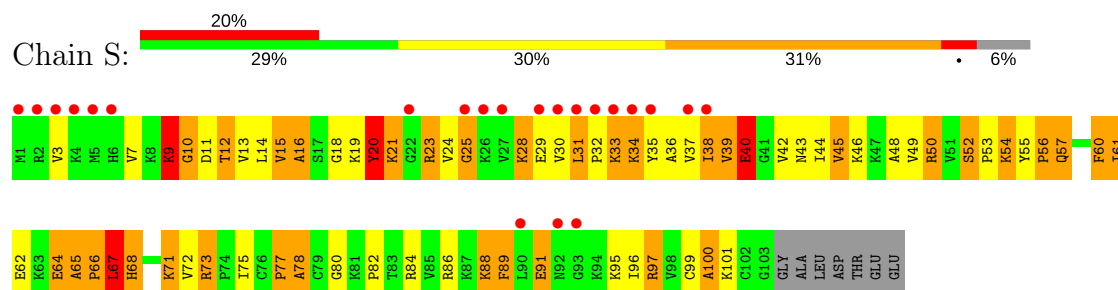
• Molecule 19: 50S ribosomal protein L22



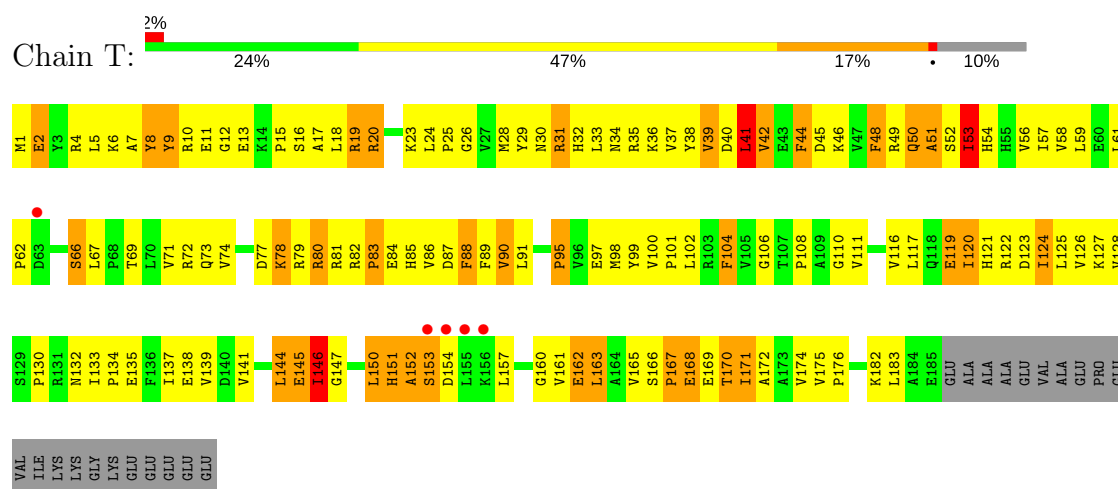
• Molecule 20: 50S ribosomal protein L23



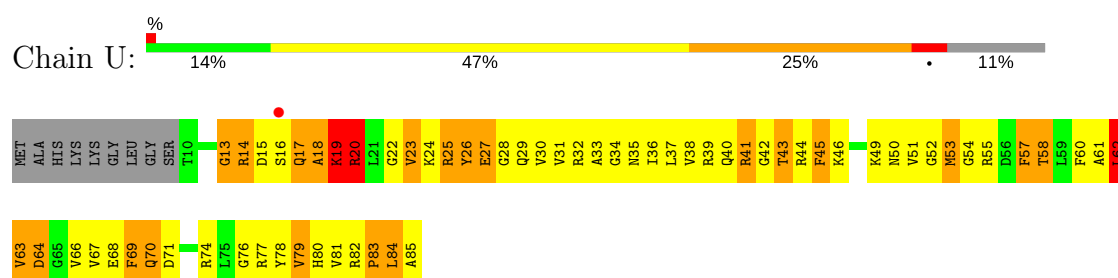
- Molecule 21: 50S ribosomal protein L24



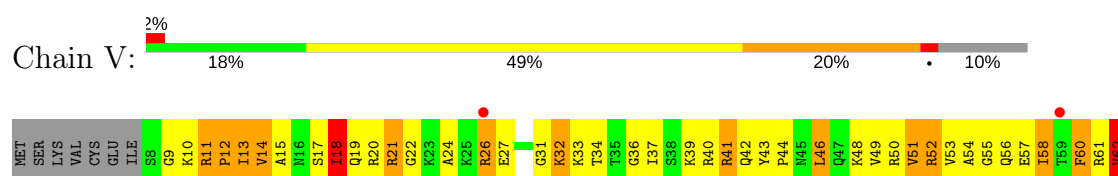
- Molecule 22: 50S ribosomal protein L25

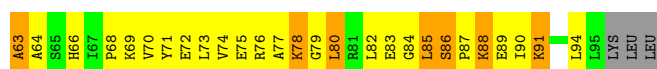


- Molecule 23: 50S ribosomal protein L27

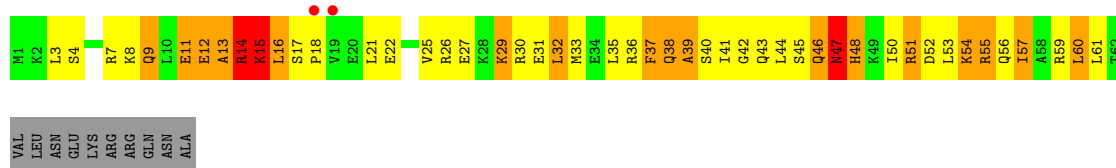
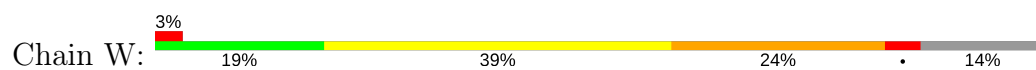


- Molecule 24: LSU ribosomal protein L28P

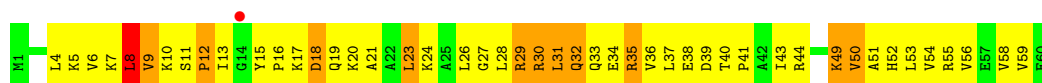




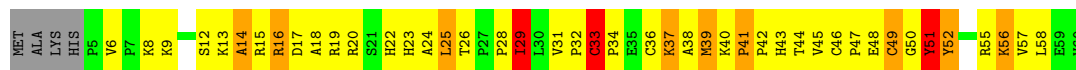
- Molecule 25: 50S ribosomal protein L29



- Molecule 26: LSU ribosomal protein L30P



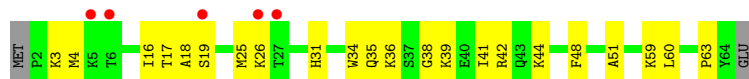
- Molecule 27: 50S ribosomal protein L32



- Molecule 28: 50S ribosomal protein L34



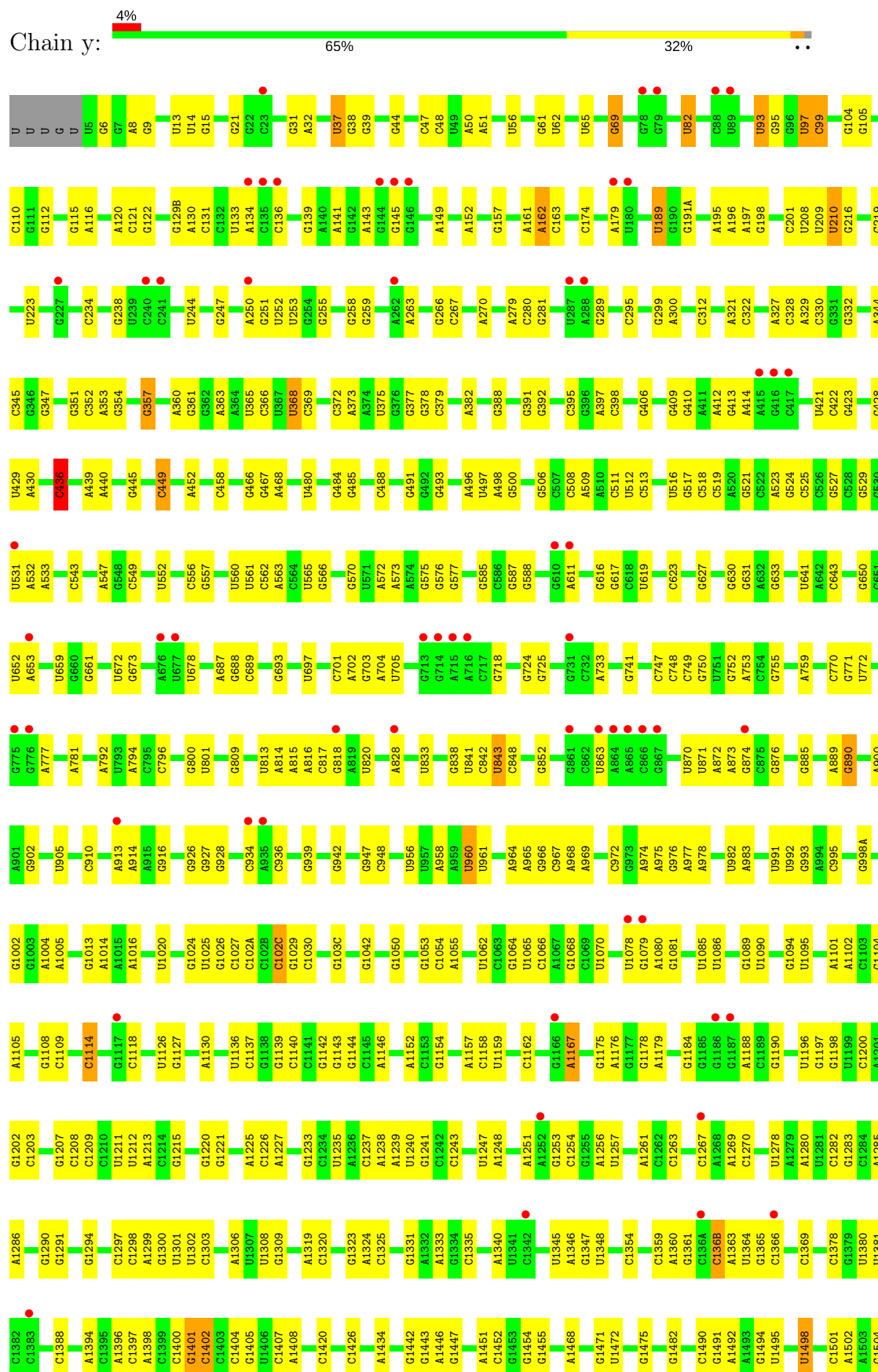
- Molecule 29: 50S ribosomal protein L35



- Molecule 30: 50S ribosomal protein L36



- Molecule 31: 16S RNA

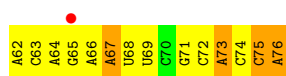
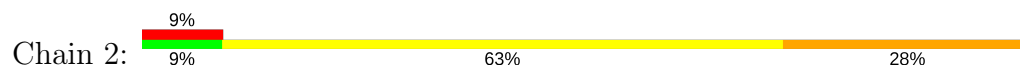




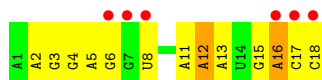
• Molecule 32: P-site tRNA<sup>f</sup>MET



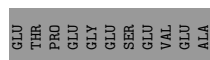
• Molecule 33: E-site tRNA



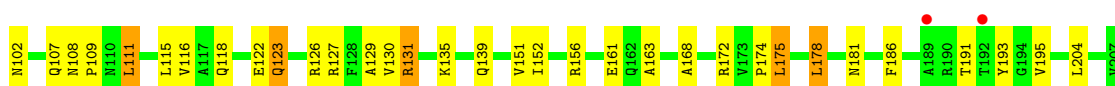
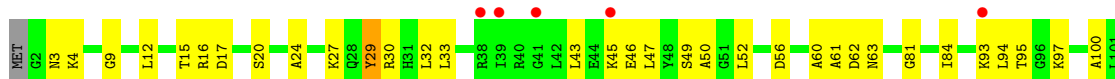
• Molecule 34: MRNA



• Molecule 35: 30S ribosomal protein S2



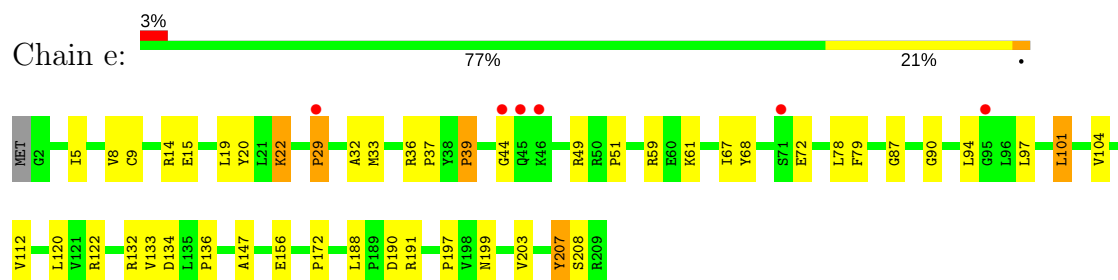
• Molecule 36: 30S ribosomal protein S3



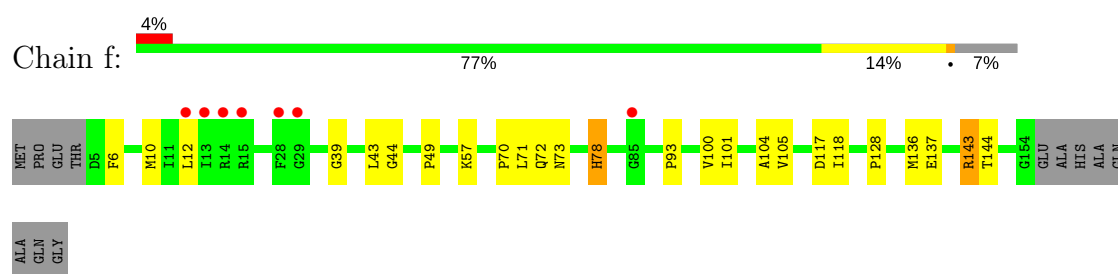


ILE  
GLY  
GLN  
LYS  
PRO  
LYS  
ALA  
ARG  
PRO  
GLU  
LEU  
PRO  
LYS  
ALA  
GLU  
GLU  
ARG  
PRO  
ARG  
ARG  
ARG  
ARG  
PRO  
ALA  
VAL  
ARG  
VAL  
LYS  
LYS  
GLU  
GLU

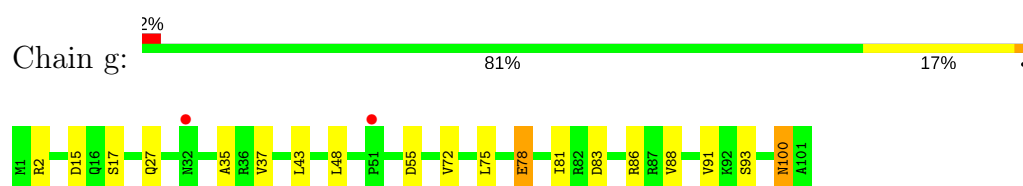
• Molecule 37: 30S ribosomal protein S4



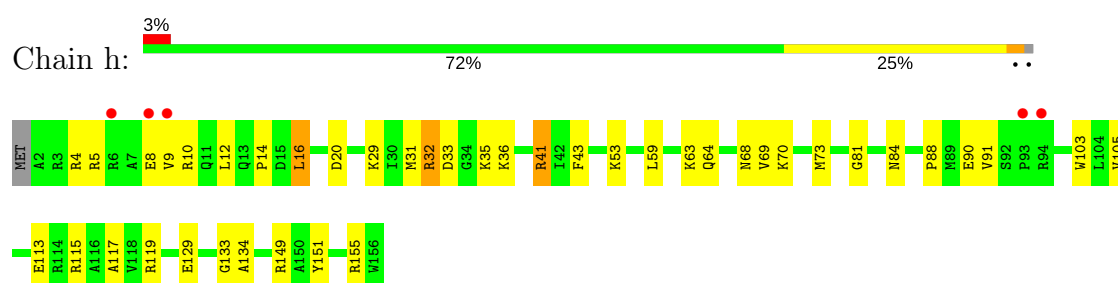
• Molecule 38: 30S ribosomal protein S5



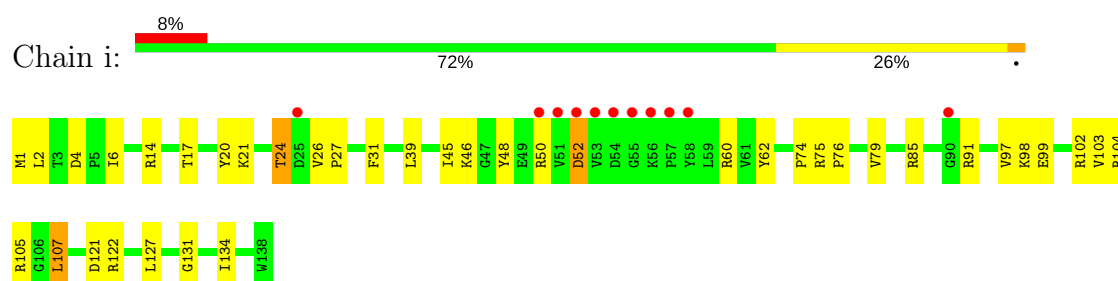
• Molecule 39: 30S ribosomal protein S6



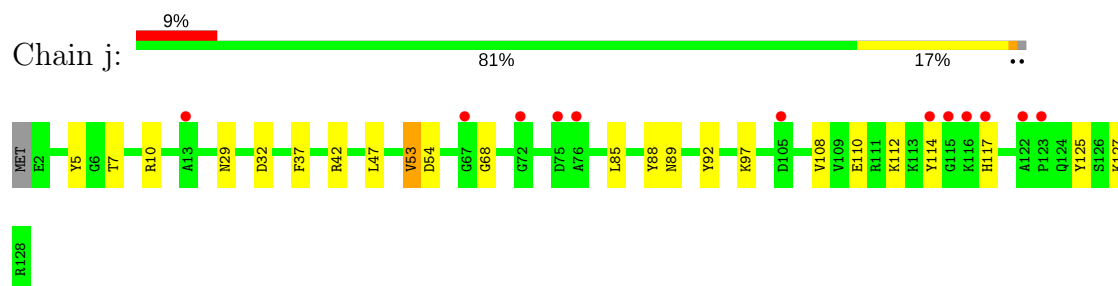
• Molecule 40: 30S ribosomal protein S7



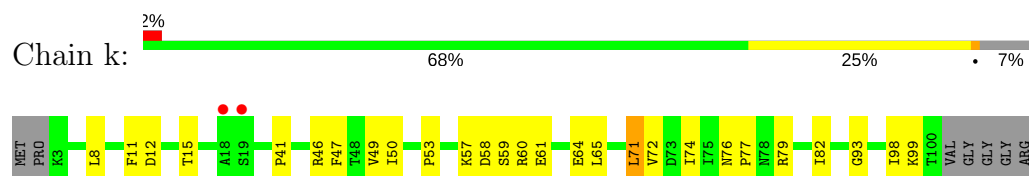
• Molecule 41: 30S ribosomal protein S8



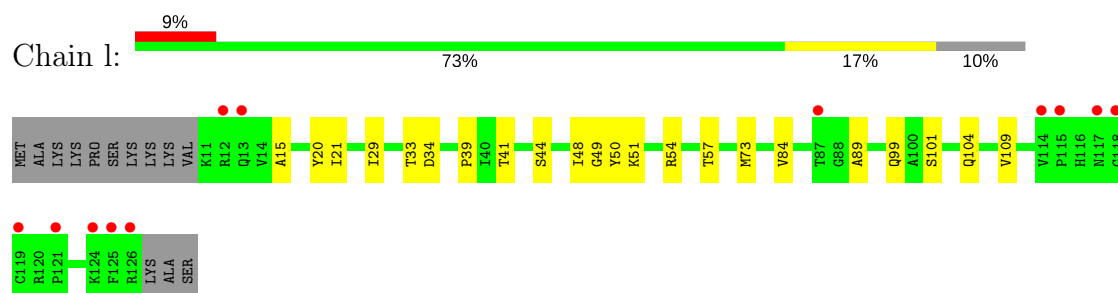
- Molecule 42: 30S ribosomal protein S9



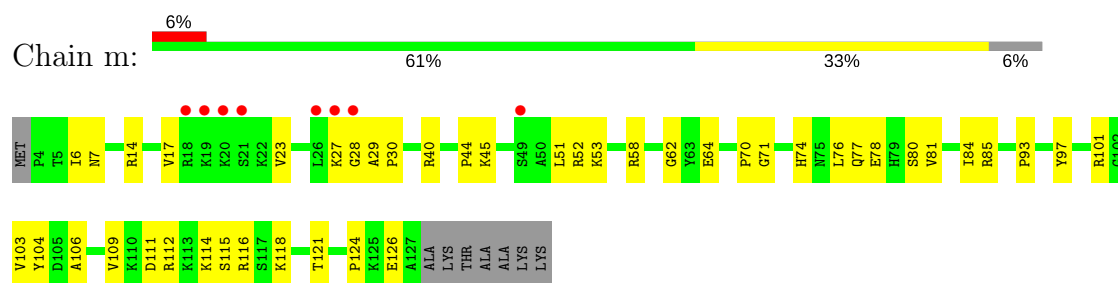
- Molecule 43: 30S ribosomal protein S10



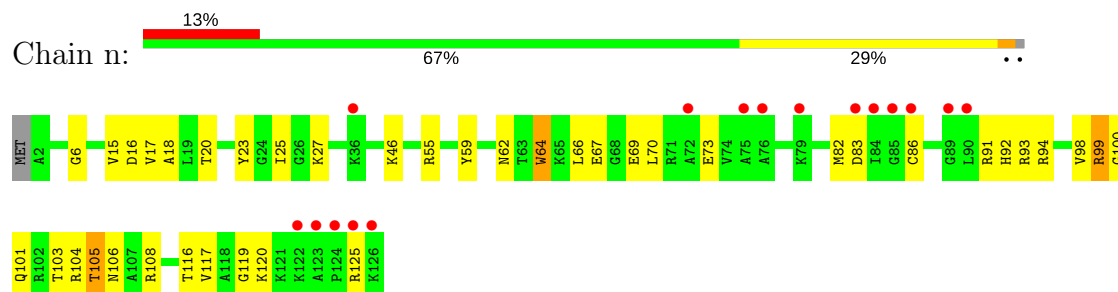
- Molecule 44: 30S ribosomal protein S11



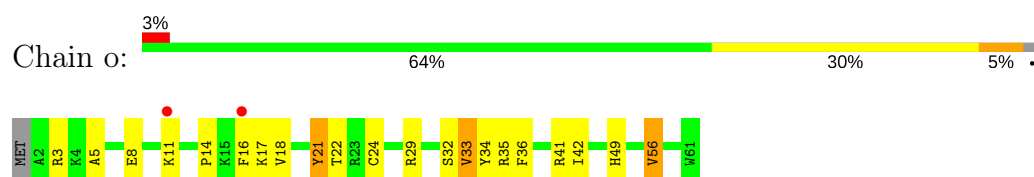
- Molecule 45: 30S ribosomal protein S12



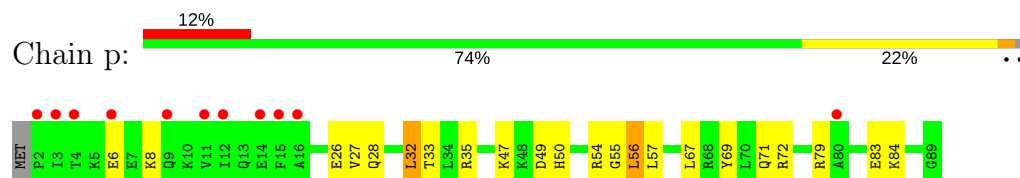
- Molecule 46: 30S ribosomal protein S13



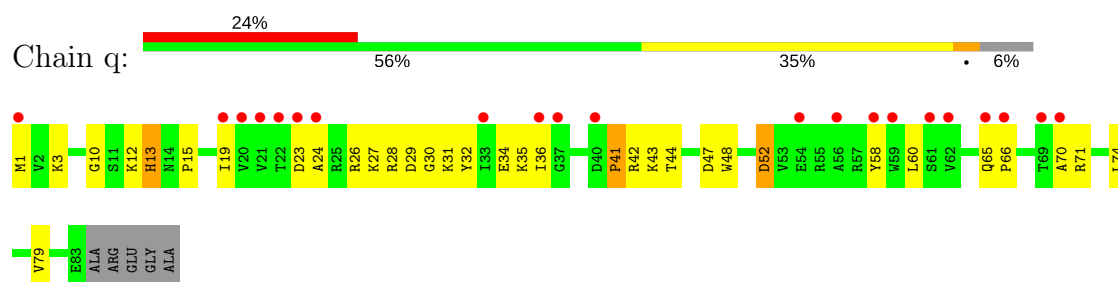
- Molecule 47: 30S ribosomal protein S14 type Z



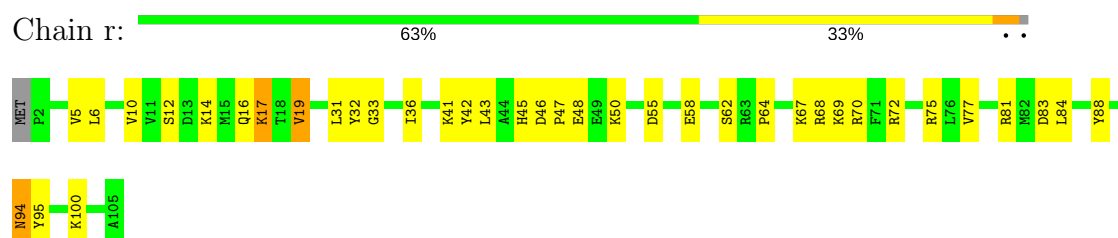
- Molecule 48: 30S ribosomal protein S15



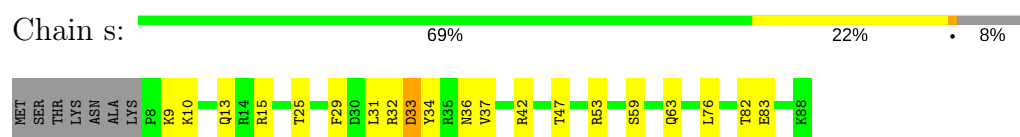
- Molecule 49: 30S ribosomal protein S16



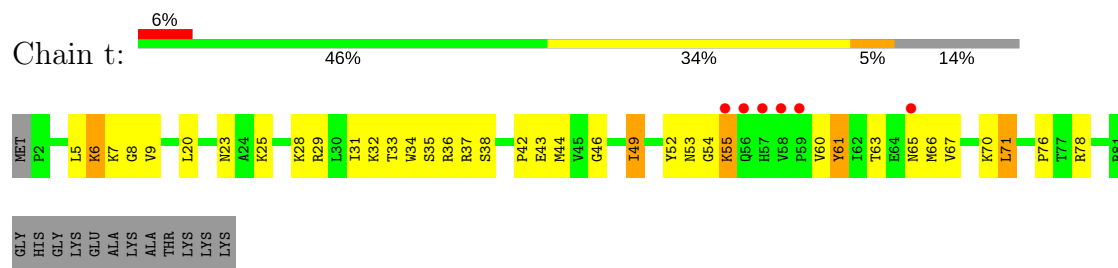
- Molecule 50: 30S ribosomal protein S17



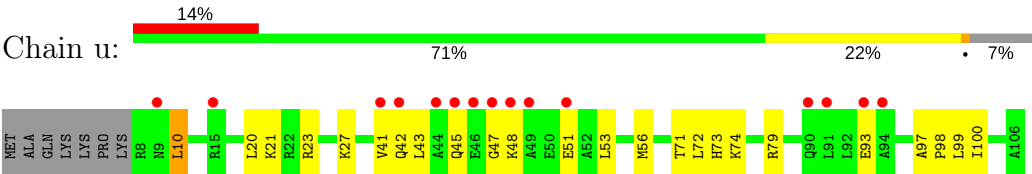
- Molecule 51: 30S ribosomal protein S18



- Molecule 52: 30S ribosomal protein S19



- Molecule 53: 30S ribosomal protein S20



• Molecule 54: 30S ribosomal protein Thx



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	507.21Å 507.21Å 692.51Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 3.83 78.70 – 3.70	Depositor EDS
% Data completeness (in resolution range)	81.7 (30.00-3.83) 68.4 (78.70-3.70)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.43 (at 3.67Å)	Xtriage
Refinement program	REFMAC 5.2.0019, CNS	Depositor
R, $R_{free}$	0.327 , 0.351 0.340 , 0.360	Depositor DCC
$R_{free}$ test set	10300 reflections (2.90%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.5	Xtriage
Anisotropy	0.111	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.07 , -110.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.22$ , $\langle L^2 \rangle = 0.07$	Xtriage
Estimated twinning fraction	0.267 for $-1/2^*h+1/2^*k-1/2^*l, 1/2^*h-1/2^*k-1/2^*l, -h-k$ 0.260 for $-1/2^*h-1/2^*k+1/2^*l, -1/2^*h-1/2^*k-1/2^*l, h-k$	Xtriage
$F_o, F_c$ correlation	0.69	EDS
Total number of atoms	147125	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	1.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.08% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: MA6, 2MG, 5MC, M2G, 7MG, PSU

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	w	1.00	42/69679 (0.1%)	1.03	92/108779 (0.1%)
2	x	0.65	1/2878 (0.0%)	0.90	1/4490 (0.0%)
3	A	0.53	0/1015	0.63	0/1369
4	B	0.53	0/2165	0.70	0/2919
5	C	0.56	0/1574	0.69	0/2125
6	D	0.58	0/1551	0.69	0/2101
7	E	0.58	0/1492	0.72	1/2006 (0.0%)
8	F	0.56	0/1345	0.70	0/1819
9	G	0.51	0/1171	0.70	0/1583
10	H	0.54	0/1130	0.71	1/1525 (0.1%)
11	I	0.57	0/942	0.74	1/1268 (0.1%)
12	J	0.57	0/1131	0.76	1/1504 (0.1%)
13	K	0.58	0/1110	0.74	1/1483 (0.1%)
14	L	0.49	0/982	0.69	0/1312
15	M	0.51	0/856	0.63	0/1138
16	N	0.56	0/1157	0.72	0/1544
17	O	0.53	0/982	0.64	0/1306
18	P	0.58	0/790	0.69	0/1057
19	Q	0.51	0/878	0.74	1/1179 (0.1%)
20	R	0.60	0/739	0.75	0/993
21	S	0.61	0/806	0.70	0/1074
22	T	0.54	0/1507	0.66	0/2045
23	U	0.56	0/613	0.75	0/816
24	V	0.64	0/701	0.71	0/932
25	W	0.53	0/522	0.75	0/690
26	X	0.51	0/482	0.66	0/646
27	Y	0.53	0/449	0.69	0/606
28	Z	0.52	0/426	0.65	0/561
29	a	0.56	0/515	0.70	0/679
30	b	0.60	0/297	0.63	0/392
31	y	0.72	23/36178 (0.1%)	0.93	44/56463 (0.1%)
32	z	0.62	0/1831	0.88	0/2853

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
33	2	0.65	0/1791	0.85	0/2791
34	3	0.63	0/439	0.90	0/684
35	c	0.56	0/1935	0.67	0/2609
36	d	0.51	0/1636	0.68	0/2205
37	e	0.55	0/1733	0.68	1/2318 (0.0%)
38	f	0.60	0/1162	0.68	0/1564
39	g	0.55	0/856	0.70	0/1154
40	h	0.54	0/1276	0.63	0/1709
41	i	0.51	0/1136	0.68	0/1527
42	j	0.58	0/1029	0.66	0/1378
43	k	0.59	0/807	0.66	0/1085
44	l	0.58	0/879	0.64	0/1187
45	m	0.55	0/986	0.75	0/1320
46	n	0.54	0/1008	0.68	0/1347
47	o	0.52	0/501	0.67	0/664
48	p	0.50	0/745	0.64	0/992
49	q	0.57	0/716	0.70	0/963
50	r	0.55	0/870	0.71	0/1159
51	s	0.60	0/675	0.67	0/894
52	t	0.59	0/661	0.68	1/890 (0.1%)
53	u	0.46	0/764	0.65	0/1006
54	v	0.58	0/212	0.60	0/277
All	All	0.81	66/159711 (0.0%)	0.92	145/238980 (0.1%)

The worst 5 of 66 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	w	41	C	O3'-P	43.37	2.13	1.61
1	w	1506	C	O3'-P	40.09	2.09	1.61
1	w	489	G	O3'-P	39.72	2.08	1.61
1	w	1448(B)	A	O3'-P	38.61	2.07	1.61
1	w	436	C	O3'-P	36.62	2.05	1.61

The worst 5 of 145 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	w	2712(B)	A	P-O3'-C3'	-23.05	92.04	119.70
1	w	489	G	P-O3'-C3'	-18.72	97.23	119.70
31	y	97	U	P-O3'-C3'	-16.92	99.40	119.70
1	w	41	C	P-O3'-C3'	-15.29	101.35	119.70
1	w	1712	C	P-O3'-C3'	-15.07	101.61	119.70

There are no chirality outliers.

There are no planarity outliers.

## 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	w	62213	0	31375	0	0
2	x	2573	0	1306	0	0
3	A	996	0	1013	124	0
4	B	2115	0	2195	368	0
5	C	1541	0	1599	276	0
6	D	1517	0	1565	225	0
7	E	1468	0	1529	224	0
8	F	1319	0	1399	148	0
9	G	1156	0	1239	188	6
10	H	1103	0	1177	181	0
11	I	932	0	994	188	0
12	J	1114	0	1187	155	0
13	K	1089	0	1156	198	0
14	L	968	0	1033	140	0
15	M	846	0	902	143	0
16	N	1143	0	1211	186	0
17	O	964	0	1022	150	0
18	P	779	0	852	155	0
19	Q	868	0	929	121	0
20	R	725	0	778	106	0
21	S	793	0	890	76	0
22	T	1475	0	1504	190	0
23	U	605	0	628	123	0
24	V	694	0	764	128	0
25	W	520	0	575	72	0
26	X	477	0	529	74	0
27	Y	436	0	460	62	0
28	Z	418	0	467	31	0
29	a	507	0	576	0	0
30	b	294	0	323	0	0
31	y	32546	0	16450	0	6
32	z	1639	0	837	0	0
33	2	1621	0	821	150	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
34	3	390	0	198	12	0
35	c	1900	0	1951	0	0
36	d	1612	0	1677	0	0
37	e	1703	0	1767	0	0
38	f	1146	0	1207	0	0
39	g	843	0	857	0	0
40	h	1257	0	1296	0	0
41	i	1116	0	1177	0	0
42	j	1011	0	1043	0	0
43	k	794	0	840	0	0
44	l	864	0	881	0	0
45	m	970	0	1057	0	0
46	n	997	0	1072	0	0
47	o	492	0	533	0	0
48	p	734	0	771	0	0
49	q	700	0	720	0	0
50	r	857	0	930	0	0
51	s	668	0	748	0	0
52	t	647	0	673	0	0
53	u	762	0	859	0	0
54	v	208	0	221	0	0
All	All	147125	0	99763	4100	6

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 70.

The worst 5 of 4100 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)
11:I:97:ARG:HE	11:I:97:ARG:HA	1.04	1.15
4:B:65:ILE:HD12	4:B:105:ILE:HG13	1.18	1.15
15:M:17:ARG:HE	15:M:89:ARG:HD2	1.06	1.14
14:L:96:ARG:HG3	14:L:117:VAL:HG23	1.31	1.11
22:T:10:ARG:HG2	22:T:37:VAL:HA	1.33	1.11

The worst 5 of 6 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
9:G:10:GLU:O	31:y:436:C:OP1[4_555]	1.78	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
9:G:89:TYR:O	31:y:357:G:O2'[4_555]	2.00	0.20
9:G:90:GLY:O	31:y:368:U:OP2[4_555]	2.11	0.09
9:G:90:GLY:O	31:y:368:U:OP1[4_555]	2.14	0.06
9:G:91:SER:O	31:y:368:U:OP1[4_555]	2.17	0.03

## 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 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	
3	A	123/229 (54%)	72 (58%)	31 (25%)	20 (16%)	0	4
4	B	270/276 (98%)	134 (50%)	64 (24%)	72 (27%)	0	0
5	C	199/206 (97%)	99 (50%)	47 (24%)	53 (27%)	0	1
6	D	192/205 (94%)	91 (47%)	45 (23%)	56 (29%)	0	0
7	E	178/182 (98%)	97 (54%)	55 (31%)	26 (15%)	0	5
8	F	171/180 (95%)	90 (53%)	40 (23%)	41 (24%)	0	1
9	G	146/148 (99%)	83 (57%)	35 (24%)	28 (19%)	0	2
10	H	136/140 (97%)	62 (46%)	45 (33%)	29 (21%)	0	2
11	I	120/122 (98%)	62 (52%)	24 (20%)	34 (28%)	0	0
12	J	144/150 (96%)	63 (44%)	46 (32%)	35 (24%)	0	1
13	K	135/141 (96%)	56 (42%)	41 (30%)	38 (28%)	0	0
14	L	116/118 (98%)	64 (55%)	30 (26%)	22 (19%)	0	2
15	M	104/112 (93%)	58 (56%)	22 (21%)	24 (23%)	0	1
16	N	135/146 (92%)	56 (42%)	40 (30%)	39 (29%)	0	0
17	O	115/118 (98%)	66 (57%)	28 (24%)	21 (18%)	0	3
18	P	99/101 (98%)	39 (39%)	28 (28%)	32 (32%)	0	0
19	Q	107/113 (95%)	64 (60%)	23 (22%)	20 (19%)	0	2
20	R	90/96 (94%)	29 (32%)	36 (40%)	25 (28%)	0	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
21	S	101/110 (92%)	28 (28%)	31 (31%)	42 (42%)	0	0
22	T	183/206 (89%)	115 (63%)	45 (25%)	23 (13%)	0	7
23	U	74/85 (87%)	31 (42%)	21 (28%)	22 (30%)	0	0
24	V	86/98 (88%)	41 (48%)	29 (34%)	16 (19%)	0	2
25	W	60/72 (83%)	33 (55%)	11 (18%)	16 (27%)	0	0
26	X	58/60 (97%)	31 (53%)	20 (34%)	7 (12%)	0	7
27	Y	54/60 (90%)	27 (50%)	16 (30%)	11 (20%)	0	2
28	Z	46/49 (94%)	31 (67%)	10 (22%)	5 (11%)	0	9
29	a	61/65 (94%)	28 (46%)	20 (33%)	13 (21%)	0	2
30	b	33/37 (89%)	16 (48%)	10 (30%)	7 (21%)	0	2
35	c	232/256 (91%)	142 (61%)	59 (25%)	31 (13%)	0	5
36	d	204/239 (85%)	102 (50%)	60 (29%)	42 (21%)	0	2
37	e	206/209 (99%)	120 (58%)	60 (29%)	26 (13%)	0	7
38	f	148/162 (91%)	102 (69%)	32 (22%)	14 (10%)	1	13
39	g	99/101 (98%)	72 (73%)	19 (19%)	8 (8%)	1	16
40	h	153/156 (98%)	96 (63%)	30 (20%)	27 (18%)	0	3
41	i	136/138 (99%)	84 (62%)	33 (24%)	19 (14%)	0	5
42	j	125/128 (98%)	81 (65%)	29 (23%)	15 (12%)	0	7
43	k	96/105 (91%)	55 (57%)	25 (26%)	16 (17%)	0	4
44	l	114/129 (88%)	74 (65%)	28 (25%)	12 (10%)	0	10
45	m	122/132 (92%)	67 (55%)	35 (29%)	20 (16%)	0	4
46	n	123/126 (98%)	70 (57%)	35 (28%)	18 (15%)	0	5
47	o	58/61 (95%)	30 (52%)	16 (28%)	12 (21%)	0	2
48	p	86/89 (97%)	51 (59%)	24 (28%)	11 (13%)	0	6
49	q	81/88 (92%)	34 (42%)	25 (31%)	22 (27%)	0	0
50	r	102/105 (97%)	54 (53%)	24 (24%)	24 (24%)	0	1
51	s	79/88 (90%)	45 (57%)	21 (27%)	13 (16%)	0	4
52	t	78/93 (84%)	30 (38%)	26 (33%)	22 (28%)	0	0
53	u	97/106 (92%)	47 (48%)	35 (36%)	15 (16%)	0	4
54	v	22/27 (82%)	8 (36%)	6 (27%)	8 (36%)	0	0
All	All	5697/6163 (92%)	3030 (53%)	1515 (27%)	1152 (20%)	0	2

5 of 1152 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	8	TYR
3	A	16	ASP
3	A	63	VAL
3	A	176	VAL
3	A	179	ALA

### 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 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	
3	A	106/181 (59%)	86 (81%)	20 (19%)	1	12
4	B	214/218 (98%)	179 (84%)	35 (16%)	2	18
5	C	163/166 (98%)	130 (80%)	33 (20%)	1	10
6	D	154/162 (95%)	133 (86%)	21 (14%)	4	25
7	E	154/156 (99%)	130 (84%)	24 (16%)	3	20
8	F	142/148 (96%)	129 (91%)	13 (9%)	10	40
9	G	124/124 (100%)	100 (81%)	24 (19%)	1	11
10	H	117/119 (98%)	96 (82%)	21 (18%)	2	14
11	I	100/100 (100%)	74 (74%)	26 (26%)	0	4
12	J	112/116 (97%)	91 (81%)	21 (19%)	1	12
13	K	108/111 (97%)	84 (78%)	24 (22%)	1	8
14	L	101/101 (100%)	89 (88%)	12 (12%)	6	30
15	M	84/88 (96%)	71 (84%)	13 (16%)	3	20
16	N	121/128 (94%)	101 (84%)	20 (16%)	2	17
17	O	93/94 (99%)	77 (83%)	16 (17%)	2	16
18	P	82/82 (100%)	64 (78%)	18 (22%)	1	8
19	Q	89/92 (97%)	75 (84%)	14 (16%)	3	20
20	R	74/78 (95%)	62 (84%)	12 (16%)	2	18
21	S	86/91 (94%)	73 (85%)	13 (15%)	3	21

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
22	T	163/179 (91%)	136 (83%)	27 (17%)	2	17
23	U	61/67 (91%)	54 (88%)	7 (12%)	6	31
24	V	73/83 (88%)	61 (84%)	12 (16%)	2	18
25	W	58/67 (87%)	47 (81%)	11 (19%)	1	12
26	X	52/52 (100%)	46 (88%)	6 (12%)	6	31
27	Y	49/52 (94%)	40 (82%)	9 (18%)	2	13
28	Z	41/42 (98%)	34 (83%)	7 (17%)	2	16
29	a	53/55 (96%)	44 (83%)	9 (17%)	2	17
30	b	33/34 (97%)	31 (94%)	2 (6%)	20	57
35	c	202/220 (92%)	169 (84%)	33 (16%)	2	18
36	d	160/188 (85%)	130 (81%)	30 (19%)	1	12
37	e	180/181 (99%)	154 (86%)	26 (14%)	3	23
38	f	115/123 (94%)	102 (89%)	13 (11%)	6	32
39	g	90/90 (100%)	77 (86%)	13 (14%)	3	23
40	h	126/127 (99%)	108 (86%)	18 (14%)	3	24
41	i	119/119 (100%)	96 (81%)	23 (19%)	1	11
42	j	98/99 (99%)	89 (91%)	9 (9%)	10	40
43	k	88/92 (96%)	76 (86%)	12 (14%)	4	25
44	l	88/99 (89%)	78 (89%)	10 (11%)	6	31
45	m	104/109 (95%)	80 (77%)	24 (23%)	1	7
46	n	100/101 (99%)	75 (75%)	25 (25%)	0	5
47	o	49/50 (98%)	37 (76%)	12 (24%)	1	6
48	p	79/80 (99%)	66 (84%)	13 (16%)	2	17
49	q	72/74 (97%)	57 (79%)	15 (21%)	1	9
50	r	96/97 (99%)	79 (82%)	17 (18%)	2	15
51	s	71/77 (92%)	63 (89%)	8 (11%)	6	32
52	t	71/80 (89%)	52 (73%)	19 (27%)	0	4
53	u	76/82 (93%)	66 (87%)	10 (13%)	4	26
54	v	19/22 (86%)	16 (84%)	3 (16%)	3	20
All	All	4810/5096 (94%)	4007 (83%)	803 (17%)	2	17

5 of 803 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
20	R	65	ARG
27	Y	49	CYS
49	q	60	LEU
21	S	50	ARG
23	U	19	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 134 such sidechains are listed below:

Mol	Chain	Res	Type
22	T	65	GLN
35	c	135	GLN
48	p	13	GLN
22	T	75	ASN
25	W	47	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	w	2888/2889 (99%)	983 (34%)	0
2	x	119/120 (99%)	47 (39%)	0
31	y	1511/1522 (99%)	507 (33%)	0
32	z	76/77 (98%)	20 (26%)	0
33	2	75/76 (98%)	22 (29%)	0
34	3	17/18 (94%)	5 (29%)	0
All	All	4686/4702 (99%)	1584 (33%)	0

5 of 1584 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	w	10	G
1	w	13	A
1	w	15	G
1	w	27	G
1	w	28	A

There are no RNA pucker outliers to report.

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

11 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$
33	PSU	2	55	33	16,21,22	1.27	3 (18%)	20,30,33	3.56	9 (45%)
31	2MG	y	1207	31	19,26,27	2.22	5 (26%)	19,38,41	4.20	6 (31%)
31	5MC	y	1400	31	14,22,23	1.15	2 (14%)	16,32,35	0.96	1 (6%)
31	5MC	y	1404	31	14,22,23	1.36	2 (14%)	16,32,35	1.70	1 (6%)
31	5MC	y	1407	31	14,22,23	1.22	2 (14%)	16,32,35	1.08	0
31	MA6	y	1518	31	16,26,27	1.26	1 (6%)	17,38,41	2.52	6 (35%)
31	MA6	y	1519	31	16,26,27	1.99	3 (18%)	17,38,41	2.37	6 (35%)
31	PSU	y	516	31	16,21,22	1.34	2 (12%)	20,30,33	3.54	6 (30%)
31	7MG	y	527	31	20,26,27	2.33	5 (25%)	23,39,42	2.29	4 (17%)
31	M2G	y	966	31	20,27,28	2.32	4 (20%)	20,40,43	2.91	6 (30%)
31	5MC	y	967	31	14,22,23	1.09	1 (7%)	16,32,35	0.96	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
33	PSU	2	55	33	-	0/7/25/26	0/2/2/2
31	2MG	y	1207	31	-	0/5/27/28	0/3/3/3
31	5MC	y	1400	31	-	0/3/25/26	0/2/2/2
31	5MC	y	1404	31	-	0/3/25/26	0/2/2/2
31	5MC	y	1407	31	-	0/3/25/26	0/2/2/2
31	MA6	y	1518	31	-	0/7/29/30	0/3/3/3
31	MA6	y	1519	31	-	0/7/29/30	0/3/3/3
31	PSU	y	516	31	-	0/7/25/26	0/2/2/2
31	7MG	y	527	31	-	0/7/37/38	0/3/3/3
31	M2G	y	966	31	-	0/7/29/30	0/3/3/3
31	5MC	y	967	31	-	0/3/25/26	0/2/2/2

The worst 5 of 30 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
33	2	55	PSU	C5-C1'	-3.36	1.49	1.52
31	y	527	7MG	C8-N9	-3.33	1.37	1.45
33	2	55	PSU	O5'-C5'	-2.08	1.41	1.44
31	y	1518	MA6	C2-N3	2.02	1.35	1.32
31	y	1400	5MC	O4'-C1'	2.14	1.44	1.41

The worst 5 of 45 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
31	y	1207	2MG	C4'-O4'-C1'	-13.37	95.89	109.83
31	y	516	PSU	N1-C2-N3	-10.55	119.93	128.41
33	2	55	PSU	N1-C2-N3	-9.78	120.55	128.41
31	y	1207	2MG	C5-C6-N1	-8.45	111.45	123.47
31	y	966	M2G	C5-C6-N1	-8.30	111.67	123.47

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
33	2	55	PSU	3	0

## 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
1	w	35
31	y	9

The worst 5 of 44 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	w	41:C	O3'	43:G	P	2.13
1	w	1506:C	O3'	1508:A	P	2.09
1	w	489:G	O3'	491:G	P	2.08
1	w	1448(B):A	O3'	1449:G	P	2.07
1	w	436:C	O3'	438:G	P	2.05

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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, 95<sup>th</sup> 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(Å <sup>2</sup> )	Q<0.9
1	w	2889/2889 (100%)	0.03	107 (3%) 41 31	1, 1, 1, 1	0
2	x	120/120 (100%)	-0.22	0 100 100	1, 1, 1, 1	0
3	A	127/229 (55%)	-0.37	6 (4%) 31 25	1, 1, 1, 1	0
4	B	272/276 (98%)	-0.14	19 (6%) 16 12	1, 1, 1, 1	0
5	C	201/206 (97%)	-0.13	12 (5%) 22 16	1, 1, 1, 1	0
6	D	194/205 (94%)	-0.64	0 100 100	1, 1, 1, 1	0
7	E	180/182 (98%)	-0.30	8 (4%) 34 27	1, 1, 1, 1	0
8	F	173/180 (96%)	-0.46	8 (4%) 32 26	1, 1, 1, 1	0
9	G	148/148 (100%)	-0.21	8 (5%) 26 20	1, 1, 1, 1	0
10	H	138/140 (98%)	-0.43	2 (1%) 75 66	1, 1, 1, 1	0
11	I	122/122 (100%)	-0.17	8 (6%) 18 13	1, 1, 1, 1	0
12	J	146/150 (97%)	-0.51	3 (2%) 63 53	1, 1, 1, 1	0
13	K	137/141 (97%)	-0.08	11 (8%) 12 10	1, 1, 1, 1	0
14	L	118/118 (100%)	-0.24	6 (5%) 28 23	1, 1, 1, 1	0
15	M	106/112 (94%)	-0.16	5 (4%) 31 25	1, 1, 1, 1	0
16	N	137/146 (93%)	-0.24	3 (2%) 62 51	1, 1, 1, 1	0
17	O	117/118 (99%)	-0.40	4 (3%) 45 34	1, 1, 1, 1	0
18	P	101/101 (100%)	0.17	15 (14%) 2 2	1, 1, 1, 1	0
19	Q	109/113 (96%)	0.43	17 (15%) 2 2	1, 1, 1, 1	0
20	R	92/96 (95%)	-0.49	1 (1%) 80 72	1, 1, 1, 1	0
21	S	103/110 (93%)	0.56	22 (21%) 1 1	1, 1, 1, 1	0
22	T	185/206 (89%)	-0.41	5 (2%) 54 43	1, 1, 1, 1	0
23	U	76/85 (89%)	-0.41	1 (1%) 77 68	1, 1, 1, 1	0
24	V	88/98 (89%)	-0.30	2 (2%) 60 50	1, 1, 1, 1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
25	W	62/72 (86%)	-0.12	2 (3%) 47 36	1, 1, 1, 1	0
26	X	60/60 (100%)	-0.14	1 (1%) 70 60	1, 1, 1, 1	0
27	Y	56/60 (93%)	-0.72	0 100 100	1, 1, 1, 1	0
28	Z	48/49 (97%)	-0.69	0 100 100	1, 1, 1, 1	0
29	a	63/65 (96%)	-0.07	5 (7%) 12 10	1, 1, 1, 1	0
30	b	35/37 (94%)	0.97	11 (31%) 0 0	1, 1, 1, 1	0
31	y	1504/1522 (98%)	0.04	62 (4%) 37 29	1, 1, 1, 1	0
32	z	77/77 (100%)	0.23	4 (5%) 27 22	1, 1, 1, 1	0
33	2	75/76 (98%)	0.63	7 (9%) 8 7	1, 1, 1, 1	0
34	3	18/18 (100%)	1.14	6 (33%) 0 0	1, 1, 1, 1	0
35	c	234/256 (91%)	-0.28	10 (4%) 35 27	1, 1, 1, 1	0
36	d	206/239 (86%)	-0.50	7 (3%) 45 34	1, 1, 1, 1	0
37	e	208/209 (99%)	-0.40	6 (2%) 51 39	1, 1, 1, 1	0
38	f	150/162 (92%)	-0.49	7 (4%) 31 25	1, 1, 1, 1	0
39	g	101/101 (100%)	-0.56	2 (1%) 65 55	1, 1, 1, 1	0
40	h	155/156 (99%)	-0.37	5 (3%) 47 36	1, 1, 1, 1	0
41	i	138/138 (100%)	-0.04	11 (7%) 12 10	1, 1, 1, 1	0
42	j	127/128 (99%)	-0.03	12 (9%) 8 7	1, 1, 1, 1	0
43	k	98/105 (93%)	-0.32	2 (2%) 65 55	1, 1, 1, 1	0
44	l	116/129 (89%)	0.07	12 (10%) 6 5	1, 1, 1, 1	0
45	m	124/132 (93%)	-0.25	8 (6%) 19 13	1, 1, 1, 1	0
46	n	125/126 (99%)	0.14	16 (12%) 3 4	1, 1, 1, 1	0
47	o	60/61 (98%)	-0.50	2 (3%) 46 35	1, 1, 1, 1	0
48	p	88/89 (98%)	0.18	11 (12%) 4 4	1, 1, 1, 1	0
49	q	83/88 (94%)	1.02	21 (25%) 0 0	1, 1, 1, 1	0
50	r	104/105 (99%)	-0.59	0 100 100	1, 1, 1, 1	0
51	s	81/88 (92%)	-0.65	0 100 100	1, 1, 1, 1	0
52	t	80/93 (86%)	-0.11	6 (7%) 14 10	1, 1, 1, 1	0
53	u	99/106 (93%)	0.44	15 (15%) 2 2	1, 1, 1, 1	0
54	v	24/27 (88%)	-0.51	0 100 100	1, 1, 1, 1	0
All	All	10478/10865 (96%)	-0.11	524 (5%) 29 23	1, 1, 1, 1	0

The worst 5 of 524 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
34	3	7	G	9.1
31	y	145	G	8.0
41	i	52	ASP	7.4
1	w	2399	G	7.0
44	l	118	GLY	7.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	B-factors( $\text{\AA}^2$ )	Q<0.9
33	PSU	2	55	20/21	0.82	0.24	1,1,1,1	0
31	PSU	y	516	20/21	0.92	0.17	1,1,1,1	0
31	M2G	y	966	25/26	0.92	0.14	1,1,1,1	0
31	MA6	y	1518	24/25	0.93	0.17	1,1,1,1	0
31	5MC	y	1407	21/22	0.93	0.16	1,1,1,1	0
31	MA6	y	1519	24/25	0.93	0.16	1,1,1,1	0
31	7MG	y	527	24/25	0.93	0.23	1,1,1,1	0
31	5MC	y	1404	21/22	0.94	0.20	1,1,1,1	0
31	5MC	y	967	21/22	0.94	0.14	1,1,1,1	0
31	2MG	y	1207	24/25	0.95	0.16	1,1,1,1	0
31	5MC	y	1400	21/22	0.97	0.24	1,1,1,1	0

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.