



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 12:17 AM GMT

PDB ID : 2B9P  
Title : 50S ribosomal subunit from a crystal structure of the ribosome in complex with tRNAs and mRNA with a stop codon in the A-site. This file contains the 50S subunit from a crystal structure of the ribosome in complex with tRNAs and mRNA with a stop codon in the A-site and is described in remark 400.  
Authors : Petry, S.; Brodersen, D.E.; Murphy IV, F.V.; Dunham, C.M.; Selmer, M.; Tarry, M.J.; Kelley, A.C.; Ramakrishnan, V.  
Deposited on : 2005-10-12  
Resolution : 6.46 Å(reported)

This is a wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

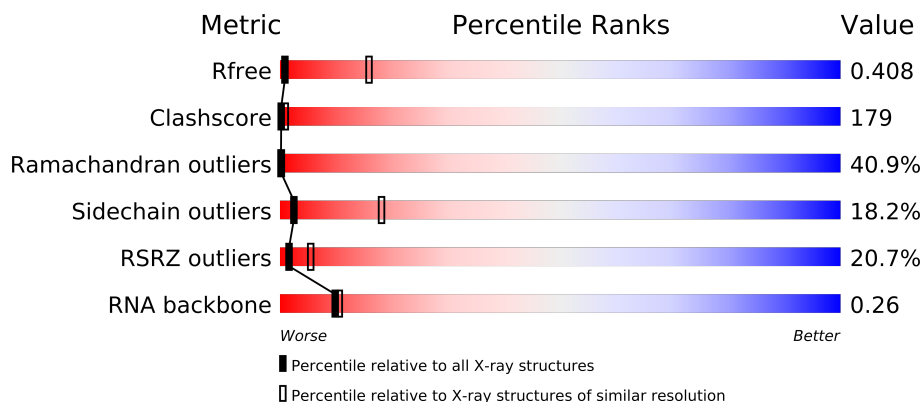
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 6.46 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))
$R_{free}$	66092	1095 (9.00-3.50)
Clashscore	79885	1033 (9.00-3.52)
Ramachandran outliers	78287	1289 (9.00-3.50)
Sidechain outliers	78261	1264 (9.00-3.50)
RSRZ outliers	66119	1094 (9.00-3.50)
RNA backbone	1838	1045 (10.00-2.80)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	B	123	
2	A	2916	
3	D	173	
4	E	338	
5	F	246	
6	G	176	
7	H	177	
8	I	149	
9	N	145	
10	O	122	
11	P	164	
12	Q	138	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
13	S	186	
14	T	66	
15	W	113	
16	X	84	
17	Y	119	
18	Z	253	
19	R	118	
20	U	118	
21	V	100	
22	2	70	
23	3	60	
24	0	91	
25	4	73	
26	5	60	
27	6	82	
28	7	47	
29	8	64	
30	9	36	
31	K	141	

## 2 Entry composition

There are 31 unique types of molecules in this entry. The entry contains 87269 atoms, of which 0 are hydrogen and 0 are deuterium.

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 5S ribosomal RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	123	Total	C	N	O	P	0	0	0
			2637	1175	488	852	122			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-1	A	-	INSERTION	GB 48271

- Molecule 2 is a RNA chain called 23S ribosomal RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	2814	Total	C	N	O	P	0	0	0
			60600	26974	11331	19482	2813			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	493	G	-	INSERTION	GB 48268

- Molecule 3 is a protein called 50S ribosomal protein L2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	173	Total	C	N	O	S	0	0	0
			1308	820	246	236	6			

- Molecule 4 is a protein called 50S ribosomal protein L3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	191	Total	C	N	O	S	0	0	0
			1507	940	290	273	4			

- Molecule 5 is a protein called 50S ribosomal protein L4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	189	Total	C	N	O	S	0	0	0
			1430	872	255	302	1			

- Molecule 6 is a protein called 50S ribosomal protein L5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	G	122	Total	C	N	O	S	0	0	0
			957	597	176	180	4			

- Molecule 7 is a protein called 50S ribosomal protein L6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	H	164	Total	C	N	O	S	0	0	0
			1251	787	225	237	2			

- Molecule 8 is a protein called 50S ribosomal protein L9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	I	148	Total	C	N	O	S	0	0	0
			1145	727	205	212	1			

- Molecule 9 is a protein called 50S ribosomal protein L13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	N	117	Total	C	N	O	S	0	0	0
			917	570	164	180	3			

- Molecule 10 is a protein called 50S ribosomal protein L14.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	O	122	Total	C	N	O	S	0	0	0
			937	585	180	169	3			

- Molecule 11 is a protein called 50S ribosomal protein L15.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
11	P	84	Total	C	N	O	0	0	0
			639	391	109	139			

- Molecule 12 is a protein called 50S ribosomal protein L16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	Q	138	Total	C	N	O	S	0	0	0
			1081	678	208	192	3			

- Molecule 13 is a protein called 50S ribosomal protein L18.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	S	113	Total	C	N	O	S	0	0	0
			866	536	165	164	1			

- Molecule 14 is a protein called 50S ribosomal protein L19.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	T	52	Total	C	N	O	S	0	0	0
			406	242	74	85	5			

- Molecule 15 is a protein called 50S ribosomal protein L22.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	W	108	Total	C	N	O		0	0	0
			860	542	169	149				

- Molecule 16 is a protein called 50S ribosomal protein L23.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
16	X	76	Total	C	N	O	S	0	0	0
			602	366	102	131	3			

- Molecule 17 is a protein called 50S ribosomal protein 24.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
17	Y	110	Total	C	N	O		0	0	0
			879	531	166	182				

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
18	Z	177	Total	C	N	O	S	0	0	0
			1360	859	238	257	6			

- Molecule 19 is a protein called 50S ribosomal protein L17.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
19	R	105	Total	C	N	O			
			855	536	174	145	0	0	0

- Molecule 20 is a protein called 50S ribosomal protein L20.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
20	U	117	Total	C	N	O	S		
			978	608	210	159	1	0	0

- Molecule 21 is a protein called 50S ribosomal protein L21.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
21	V	100	Total	C	N	O	S		
			787	495	146	145	1	0	0

- Molecule 22 is a protein called 50S ribosomal protein L29.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
22	2	64	Total	C	N	O	S		
			494	301	93	99	1	0	0

- Molecule 23 is a protein called 50S ribosomal protein L30.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
23	3	60	Total	C	N	O	S		
			477	303	91	82	1	0	0

- Molecule 24 is a protein called 50S ribosomal protein L27.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
24	0	86	Total	C	N	O	S		
			641	402	124	114	1	0	0

- Molecule 25 is a protein called 50S ribosomal protein L31.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
25	4	73	Total	C	N	O	S		
			604	382	110	108	4	0	0

- Molecule 26 is a protein called 50S ribosomal protein L32.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
26	5	58	Total	C	N	O	S	0	0	0
			457	281	94	77	5			

- Molecule 27 is a protein called 50S ribosomal protein L33.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
27	6	53	Total	C	N	O	S	0	0	0
			431	274	80	76	1			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
28	7	46	Total	C	N	O	S	0	0	0
			383	230	91	60	2			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
29	8	63	Total	C	N	O	S	0	0	0
			496	312	101	78	5			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
30	9	35	Total	C	N	O	S	0	0	0
			285	172	64	45	4			

- Molecule 31 is a protein called 50S ribosomal protein L11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
31	K	133	Total	C	N	O	S	0	0	0
			999	642	169	182	6			

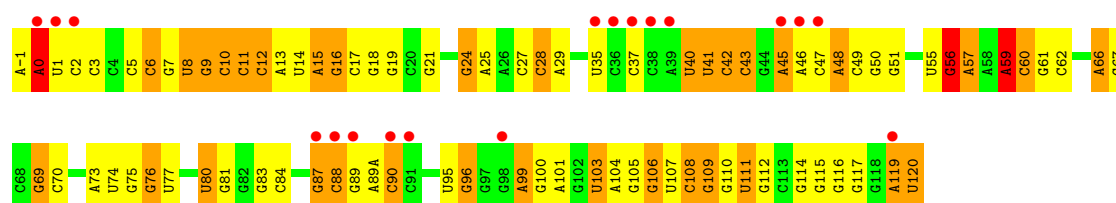


### 3 Residue-property plots

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.

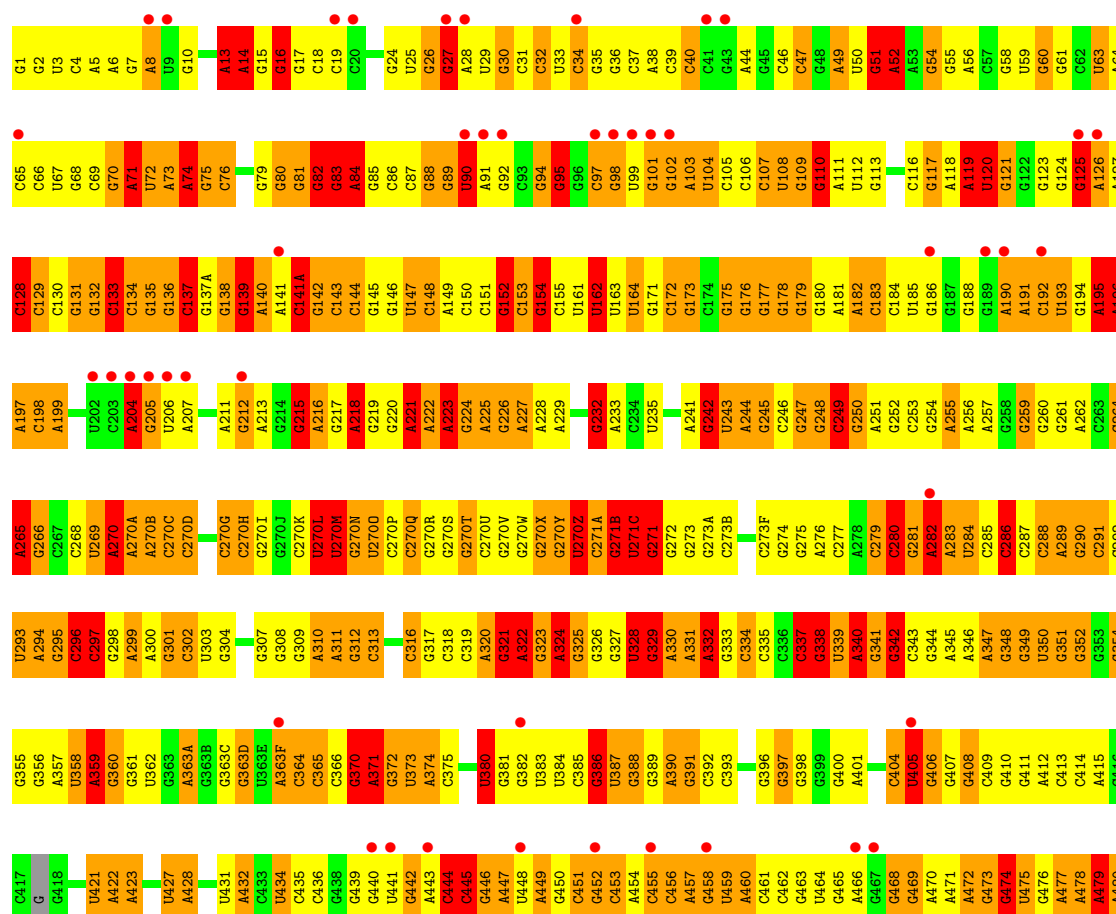
#### • Molecule 1: 5S ribosomal RNA

Chain B: 



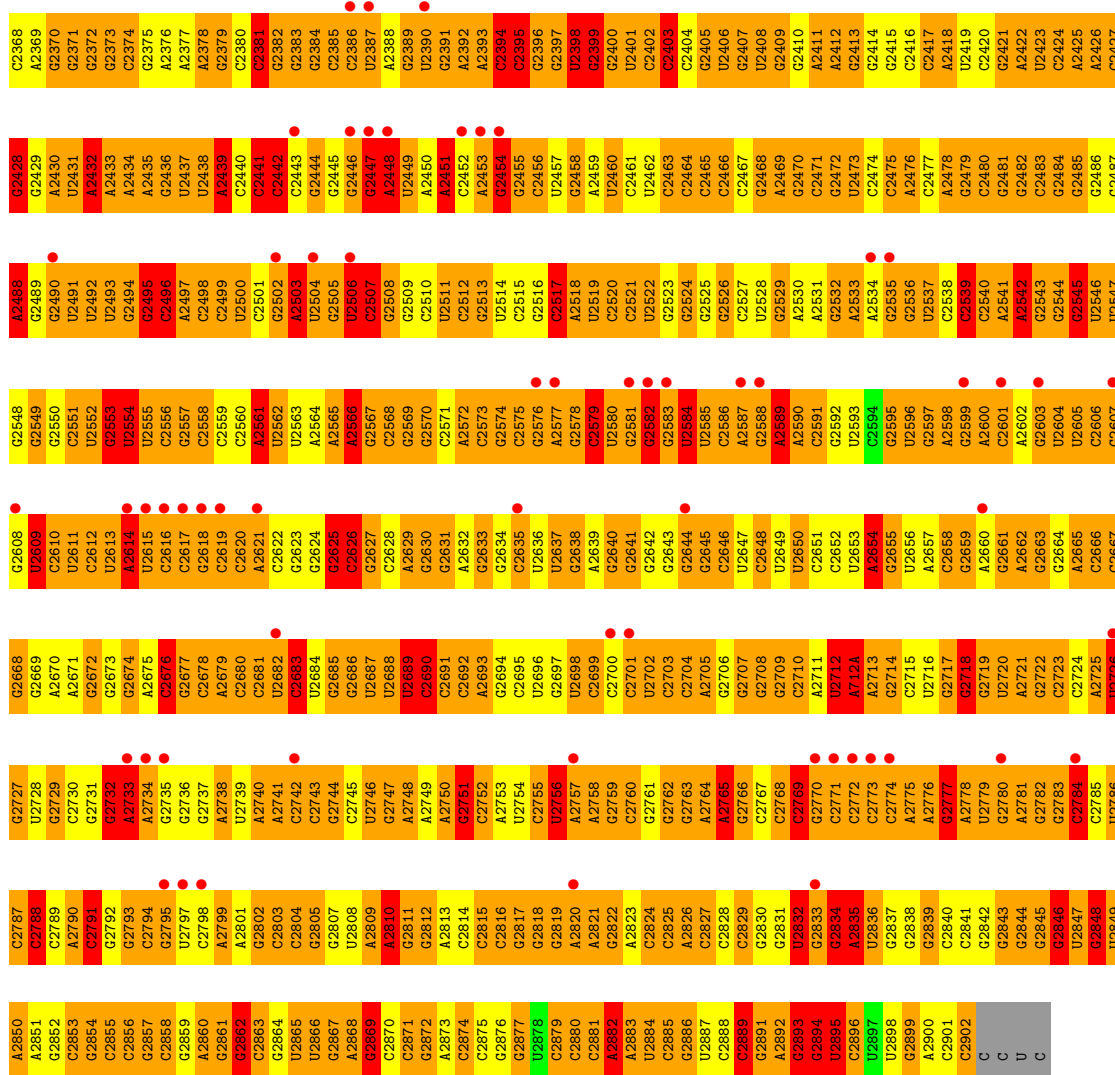
#### • Molecule 2: 23S ribosomal RNA

Chain A: 



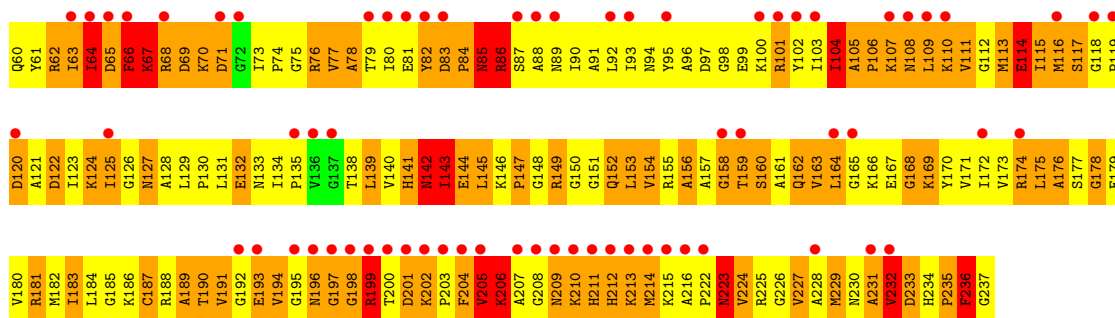


G2308	C2248	A2057	C1896	C1834	G1860	U1798	G1728	C1656	A1529	C1463	G1401
A2309	U2249	A2058	G1997	G1935	G1861	G1799	A1729	C1657	G1530	C1466	C1402
A2310	G2250	A2059	G1998	G1936	G1862	C1800	U1730	C1658	C1531	G1467	C1403
A2311	G2251	G	C1999	A1937		C1801	G1731	U1659	C1532	C1468	C1404
C2312	G2252	U	G2000	A1938	C1870	A1802	A1732	C1660	G1533	C1469	U1405
C2313	G2253	U	G2001	U1939	C1871	A1803	G1733	G1661	G1534	C1470	U1406
C2314	C2254	G	C2002	U1940	A1872	C1804	C1734	C1662	U1535	G1471	C1407
C2315	G2255	G	G2003	C1941	G1878	U1806	C1735	C1663	A1536	A1472	C1408
C2316	G2256	A	C2004	C1942	C1879	C1806	G1741	A1664	C1537	C1409	C1409
C2317	U2257	G	C2006	U1943	C1880	G1807	C1742	A1665	G1538	G1473	G1410
C2318	G2258	C	C2007	U1944	C1881	U1808	G1743	G1666	G1539	C1474	C1411
C2319	G2259	C	U2008	G1945	C1882	A1809	G1746	G1667	U1540	C1475	A1412
C2320	G2260	U	G2009	U1946	G1883	A1810	G1747	A1668	U1541	G1476	
G2321	C2261	G	G2010	C1947	A1884	C1811		A1669	G1542	G1477	U1415
A2322	U2262	A	U2011	G1948	A1885	A1812	C1751	C1670	G1543	G1478	G1416
C2323	G2263	G	G2012	G1949	C1886	G1813	C1752	U1671	C1544	G1479	C1417
C2324	G2264	A	A2013	G1950	C1887	G1814	G1753	C1672	G1545	G1480	G1418
G2325	U2265	A	A2014	U1951	C1888	A1815	G1754	U1673	A545A	A1482	A1419
C2326	A2266	C	A2015	U1952	A1889	G1816	A1755	G1674	C1546	G1483	U1420
A2267	C2267	C	U2016	A1953	A1890	C1817	G1756	C1675	U1547	G1484	G1421
A2268	A2268	C	U2017	G1954	C1891	U1818	G1757	A1676	C1548	G1485	G1422
A2269	A2269	C	G2018	U1955		U1819	G1758	A1677	G1552	U1489	G1423
G2330	G2270	C	A2019	U1956	C1895	U1820	A1759	G1678	G1553	A1490	G1424
G2331	G2271	G	G2020	C1957	G1896	A1821	A1760	U1679	A1554	G1491	G1425
G2332	U2272	C	C2021	C1958		G1822	C1761	U1680	G1555	G1492	G1426
A2333	A2273	C	U2022	G1959	G1899	G1823	A1762	G1681	G1556	C1493	A1427
G2334	A2274	C	G2023	A1960	A1900	G1824	G1763	G1682	G1557	C1494	G1428
C2335	C2275	U	G2024	C1961	A1901	A1825	G1764	C1683	C1558	A1494	G1429
A2336	G2276	C	C2025	C1962	C1902	G1826	C1765	C1684	G1559	A1495	C1430
G2337	G2277	G	U2086	U1963	G1903	C1827	U1766	C1685	G1560	A1496	U1431
G2338	G2278	G	G2087	G1964	G1904	G1828	G1767	C1686	G1561	U1497	C1432
G2339	G2279	G	G2088	C1965	C1905	A1829	U1768	G1687	G1562	C1498	U1433
G2340	U2280	U	U2089	C1966	G1906	C1830	G1769	U1688	G1563	C1499	A1434
G2341	C2281	G	A2030	C1967	G1907	G1831	G1770	A1689	C1564	G1500	C1437
G2342	G2282	G	A2031	G1968	C1908	C1832	G1771	A1690	A1565	C1502	U1438
C2343	C2283	G	G2032	A1969	C1909	U1833	G1772	C1691	A1566	A1439	C1439
U2344	G2284	G	A2033	A1970	G1910	U1834	A1773	U1692	A1567	G1440	C1441
C2345	C2285	G	U2034	A1971	U1911	G1835	C1774	U1693	G1568	G1442	G1443
A2346	A2286	G	G2035	A1972	A1912	C1836	U1775	C1694	A1569	A444A	
C2347	A2287	G	C2036	A1973	A1913	C1837	G1776	G1695	A1570		
U2348	A2288	A	G2037	C1974	C1914	C1838	U1777	G1696	A1571	C1506	
G2349	G2289	G	G2038	G1975	U1915	G1839	U1778	G1697	A1572	A1507	
U2350	U2290	G	C2039	U1976	A1916	G1840	U1779	A1698	A1573	C1508	
C2351	G2291	C	C2040	A1977	U1917	U1841	A1780	G1699	G1574	A1511	
A2352	C2292	G	U2041	A1978	A1918	G1842	C1781	A1700	C1575	G1512	
C2353	G2293	C	A2042	C1979	A1919	C1843	G1782	A1701	U1576	C1513	
G2354	C2294	C	C2043	G1980	C1920	G1844	A1783	G1702	U1577	U1514	
C2355	G2295	G	C2044	A1981	G1921		A1784	G1703	U1578	C1515	
C2356	U2296	G	C2045	C1982	G1922	A1847	A1785	G1704	U1579	U1516	
U2357	G2297	U	G2046	C1983	U1923	A1848	A1786	G1705	A1580	G1517	
C2358	A2298	G	U2047	G1984	C1924	A1849	A1787	U1706	G1581	C1518	
C2359	G2299	A	G2048		C1925	U1851	G1788	G1646	G1582	G1519	
A2360	G2300	A	C2049	C1988	U1926	C1852	A1789	C1647	C1583	U1520	
A2361	C2301	A	C2050	G1989	A1927	A1853	C1708	C1648	C1584	U1454	
G2362	G2302	A	A2051	U1990	A1928	A1854	G1709	G1649	G1585	G1455	
C2363	C2303	U	G2052	U1991	U1929	G1855	A1791	G1650	A1586	G1456	
G2364	U2244	C	G2053	G1992	C1930	G1856	G1792	G1651	U1523	A1457	
A2365	G2305	C	A2054	U1993	U1931	G1857		A1652	G1524	C1458	
C2366	G2306	G	C2055	C1994	A1932	G1858	C1795	G1653	G1525	G1459	
A2367	G2307	C	A	U1995	G1933	A1859	G1796	G1654	G1526	A1460	
							C1797	U1727	A1528	C1462	



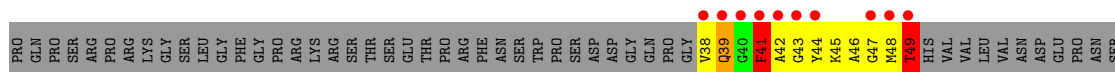
• Molecule 3: 50S ribosomal protein L2

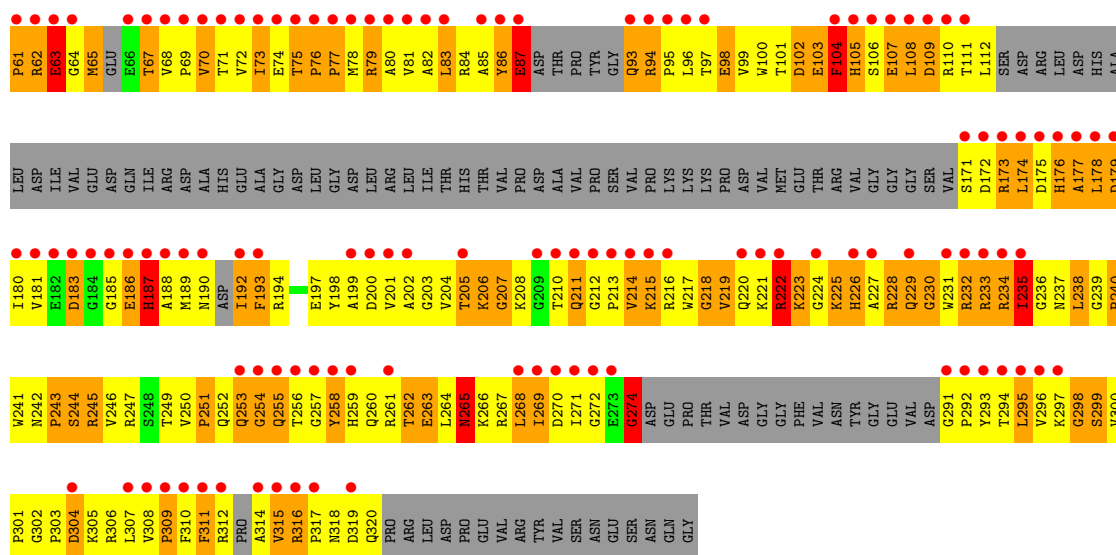
Chain D:



• Molecule 4: 50S ribosomal protein L3

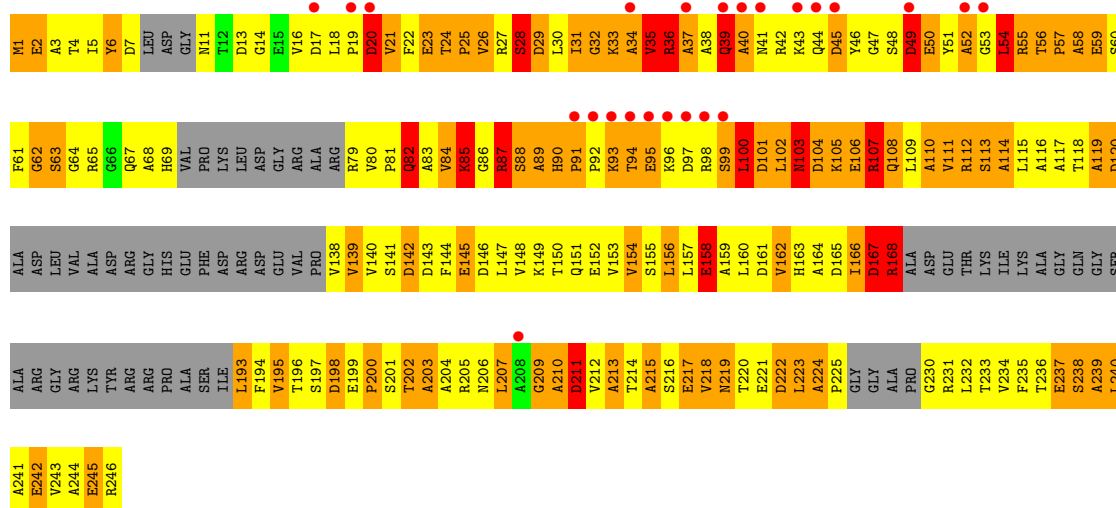
Chain E:





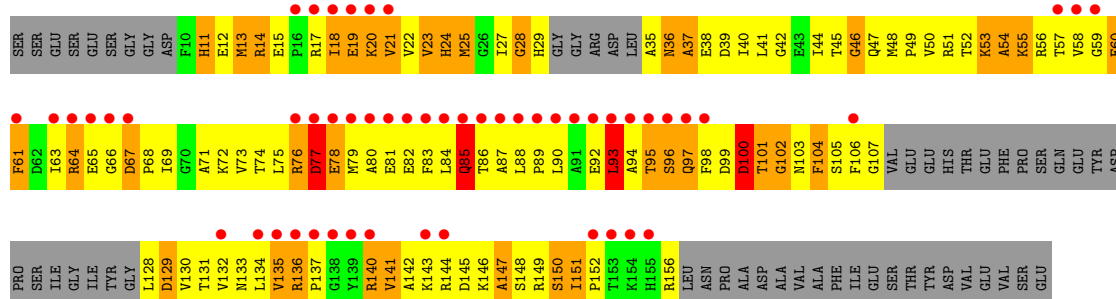
### • Molecule 5: 50S ribosomal protein L4

Chain F:



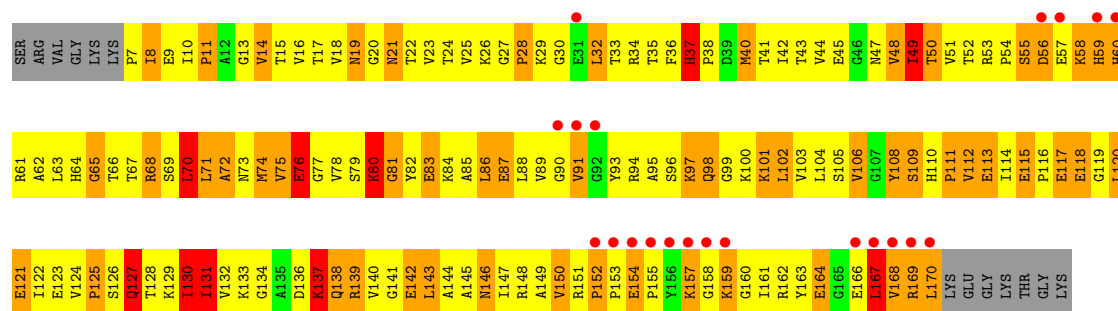
### • Molecule 6: 50S ribosomal protein L5

Chain G:



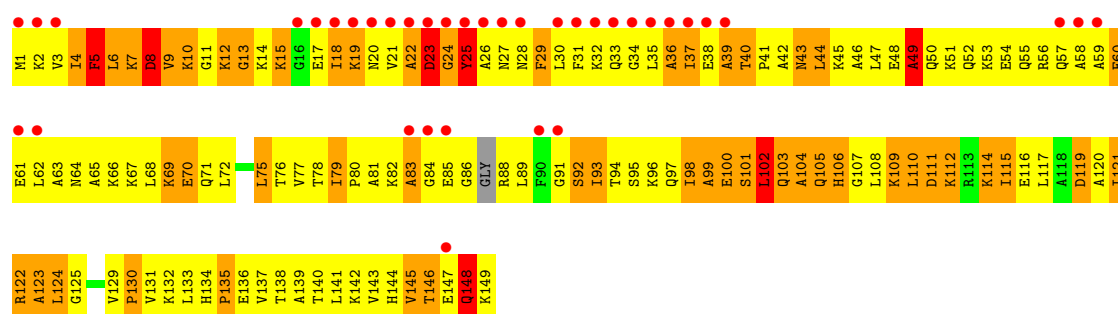
### • Molecule 7: 50S ribosomal protein L6

## Chain H:



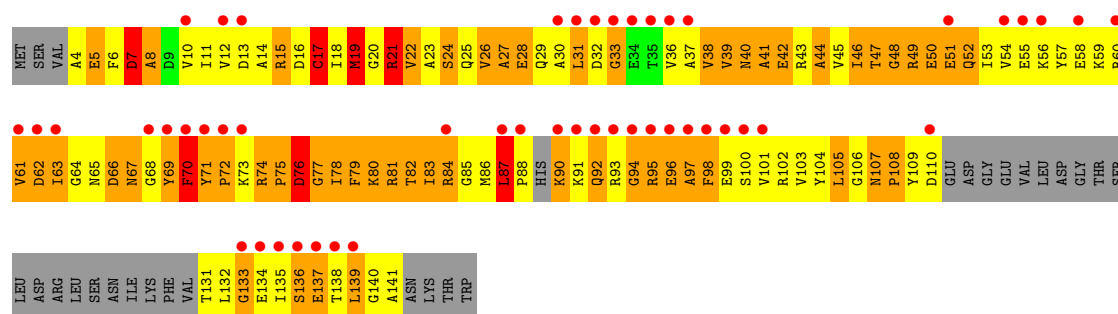
## • Molecule 8: 50S ribosomal protein L9

## Chain I:



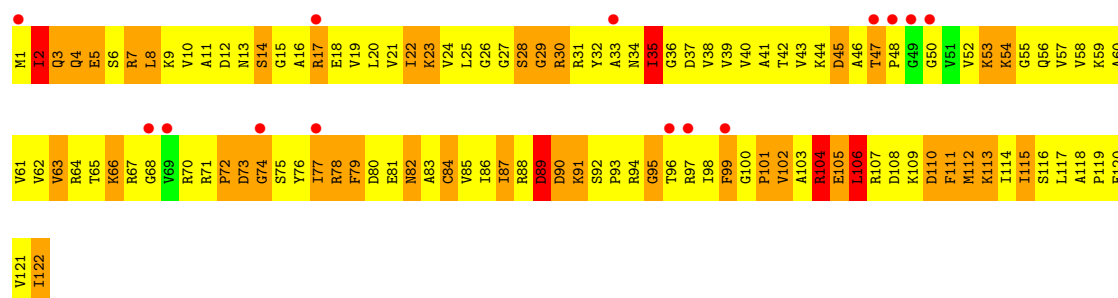
## • Molecule 9: 50S ribosomal protein L13

## Chain N:



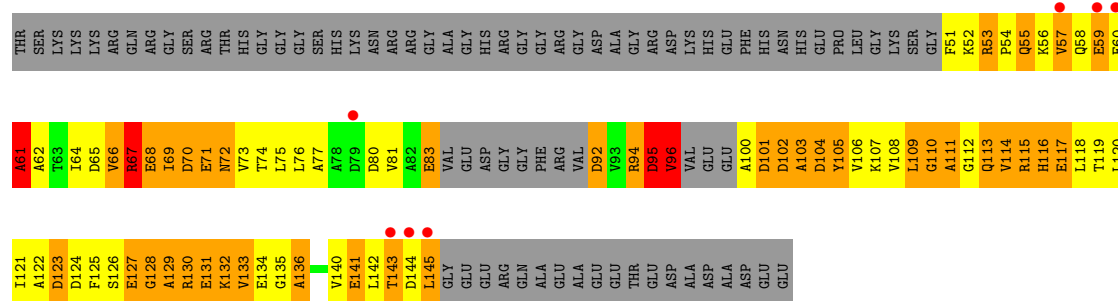
## • Molecule 10: 50S ribosomal protein L14

## Chain O:



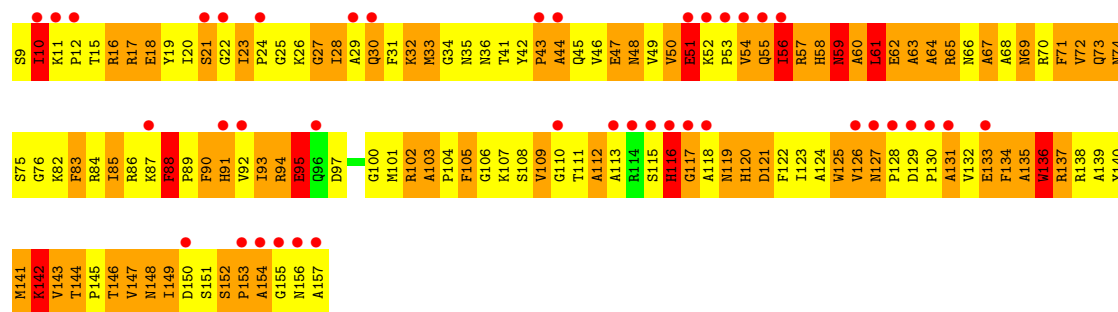
## • Molecule 11: 50S ribosomal protein L15

Chain P:



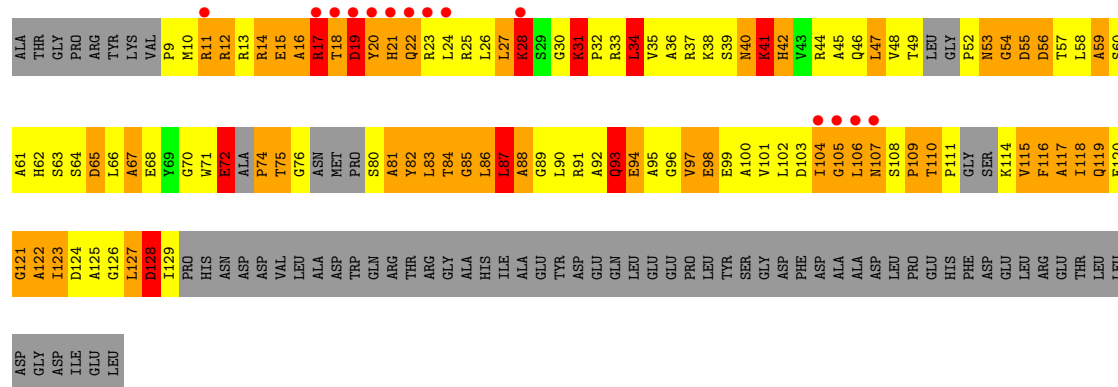
- Molecule 12: 50S ribosomal protein L16

Chain Q:



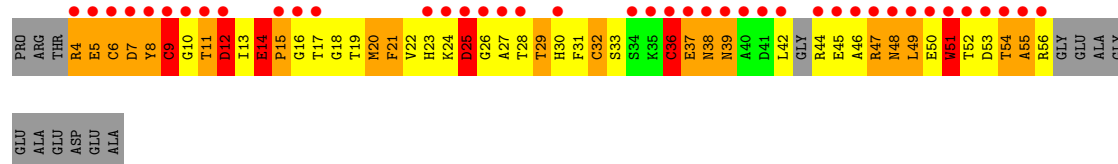
- Molecule 13: 50S ribosomal protein L18

Chain S:



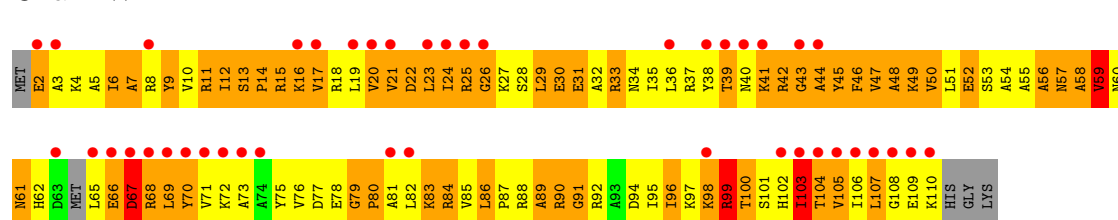
- Molecule 14: 50S ribosomal protein L19

Chain T:



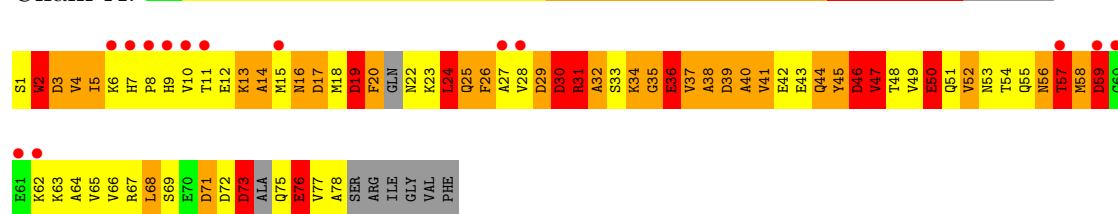
- Molecule 15: 50S ribosomal protein L22

Chain W:



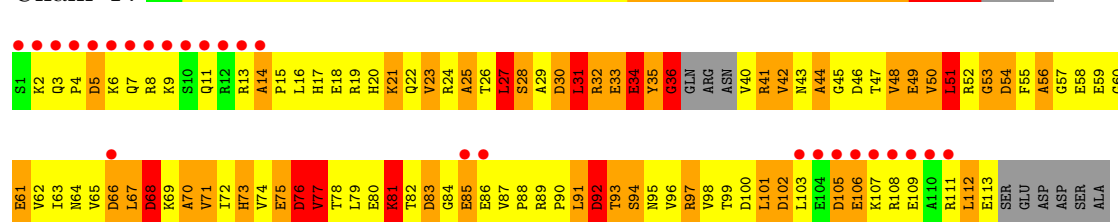
- Molecule 16: 50S ribosomal protein L23

Chain X:



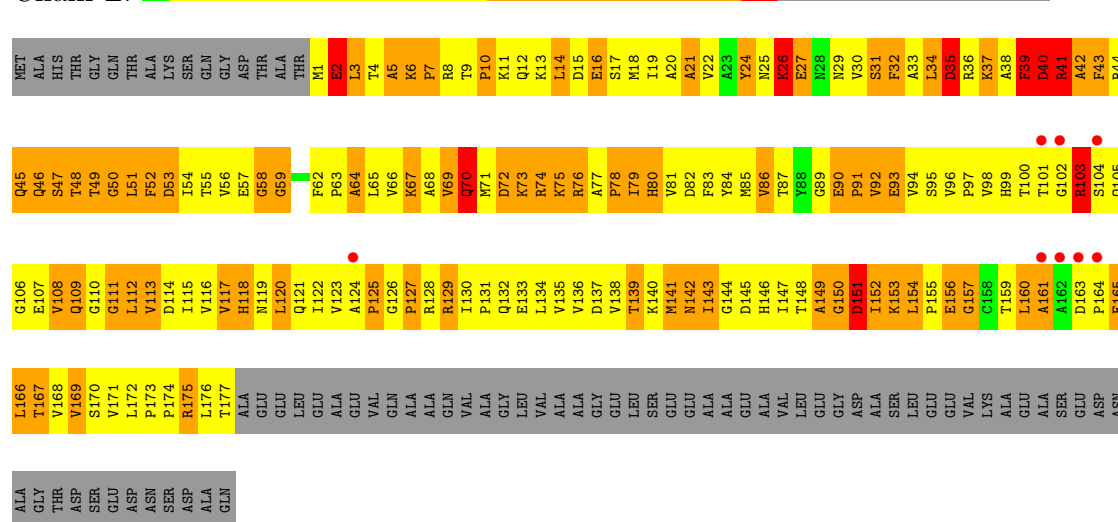
- Molecule 17: 50S ribosomal protein 24

Chain Y:



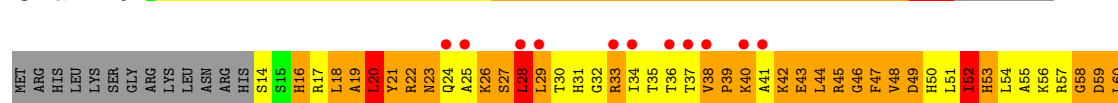
- Molecule 18: 50S ribosomal protein CTC

Chain Z:

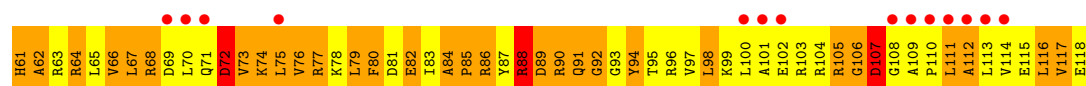


- Molecule 19: 50S ribosomal protein L17

Chain R:







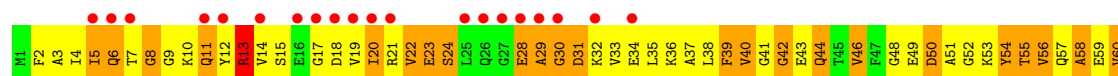
• Molecule 20: 50S ribosomal protein L20

Chain U:



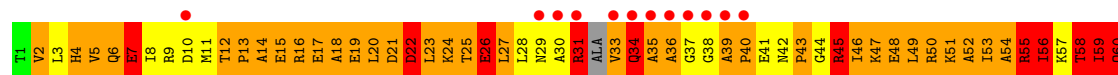
• Molecule 21: 50S ribosomal protein L21

Chain V:



• Molecule 22: 50S ribosomal protein L29

Chain 2:



• Molecule 23: 50S ribosomal protein L30

Chain 3:



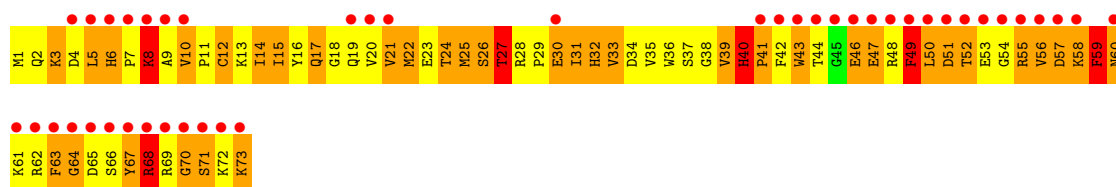
• Molecule 24: 50S ribosomal protein L27

Chain 0:



• Molecule 25: 50S ribosomal protein L31

Chain 4:



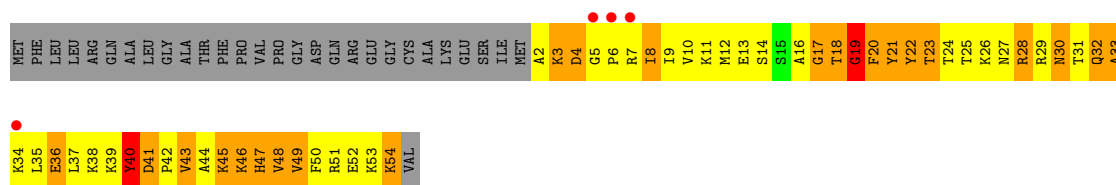
● Molecule 26: 50S ribosomal protein L32

Chain 5:



● Molecule 27: 50S ribosomal protein L33

Chain 6:



● Molecule 28: 50S ribosomal protein L34

Chain 7:



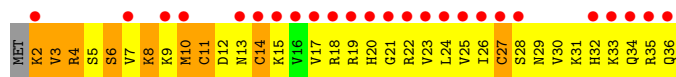
● Molecule 29: 50S ribosomal protein L35

Chain 8:



● Molecule 30: 50S ribosomal protein L36

Chain 9:



● Molecule 31: 50S ribosomal protein L11

Chain K:





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	517.41 Å   517.41 Å   365.38 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	40.00 – 6.46 99.57 – 6.00	Depositor EDS
% Data completeness (in resolution range)	99.9 (40.00-6.46) 99.7 (99.57-6.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.13	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.87 (at 6.19 Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.354   ,   0.361 0.413   ,   0.408	Depositor DCC
$R_{free}$ test set	6060 reflections (4.93%)	DCC
Wilson B-factor (Å <sup>2</sup> )	240.8	Xtriage
Anisotropy	0.222	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.07 , 35.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	0 of 123103 reflections	Xtriage
$F_o, F_c$ correlation	0.79	EDS
Total number of atoms	87269	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	289.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.58% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 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  >5	RMSZ	# Z  >5
1	B	1.25	6/2950 (0.2%)	1.44	25/4602 (0.5%)
2	A	1.17	153/67839 (0.2%)	1.46	906/105818 (0.9%)
3	D	0.38	0/1328	0.60	0/1783
4	E	0.65	4/1540 (0.3%)	1.07	7/2078 (0.3%)
5	F	0.76	3/1444 (0.2%)	0.83	1/1954 (0.1%)
6	G	0.25	0/971	0.46	0/1304
7	H	0.54	1/1272 (0.1%)	0.80	3/1721 (0.2%)
8	I	0.32	0/1156	0.71	3/1544 (0.2%)
9	N	0.35	0/927	0.55	0/1245
10	O	0.32	0/946	0.57	0/1269
11	P	1.59	3/643 (0.5%)	1.31	5/870 (0.6%)
12	Q	0.32	0/1106	0.53	0/1490
13	S	1.13	3/877 (0.3%)	0.69	1/1179 (0.1%)
14	T	0.39	0/412	0.70	0/554
15	W	0.75	3/869 (0.3%)	0.75	4/1166 (0.3%)
16	X	0.49	1/608 (0.2%)	1.04	3/820 (0.4%)
17	Y	0.26	0/887	0.83	3/1195 (0.3%)
18	Z	0.32	1/1385 (0.1%)	0.62	3/1883 (0.2%)
19	R	0.30	0/867	0.50	0/1162
20	U	0.64	1/994 (0.1%)	0.74	3/1323 (0.2%)
21	V	0.82	1/796 (0.1%)	0.91	3/1058 (0.3%)
22	2	0.37	0/497	1.00	2/668 (0.3%)
23	3	0.31	0/482	0.50	0/646
24	0	0.38	1/649 (0.2%)	0.87	3/860 (0.3%)
25	4	0.89	2/620 (0.3%)	0.54	0/831
26	5	0.36	0/469	0.90	3/629 (0.5%)
27	6	0.32	0/438	0.55	1/583 (0.2%)
28	7	0.38	0/387	0.64	0/509
29	8	0.91	2/503 (0.4%)	0.92	3/657 (0.5%)
30	9	0.33	0/286	0.59	0/375
31	K	0.27	1/1014 (0.1%)	0.44	0/1363
All	All	1.06	186/95162 (0.2%)	1.33	982/143139 (0.7%)

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
3	D	0	1
4	E	0	3
5	F	0	4
7	H	0	1
8	I	0	1
11	P	0	1
13	S	0	1
16	X	0	1
17	Y	0	1
22	2	0	1
26	5	0	1
31	K	0	1
All	All	0	17

The worst 5 of 186 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	2199	A	O3'-P	-71.04	0.75	1.61
2	A	1546	C	O3'-P	-51.62	0.99	1.61
2	A	2196	C	O3'-P	-48.04	1.03	1.61
2	A	1545	A	O3'-P	45.79	2.16	1.61
2	A	2454	G	O3'-P	-44.69	1.07	1.61

The worst 5 of 982 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	712(A)	A	P-O3'-C3'	-43.62	67.36	119.70
2	A	2199	A	O3'-P-O5'	-43.07	22.17	104.00
1	B	24	G	P-O3'-C3'	32.35	158.52	119.70
2	A	1545	A	P-O3'-C3'	-28.07	86.01	119.70
4	E	49	THR	O-C-N	-27.52	68.81	121.10

There are no chirality outliers.

5 of 17 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	D	213	LYS	Peptide
4	E	274	GLY	Mainchain,Peptide
4	E	49	THR	Mainchain

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
5	F	103	ASN	Mainchain
5	F	28	SER	Mainchain

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	2637	0	1339	219	1
2	A	60600	0	30522	11003	138
3	D	1308	0	1346	1086	0
4	E	1507	0	1478	1144	3
5	F	1430	0	1357	1085	0
6	G	957	0	952	662	0
7	H	1251	0	1291	754	0
8	I	1145	0	1225	625	0
9	N	917	0	896	771	2
10	O	937	0	993	612	0
11	P	639	0	605	482	0
12	Q	1081	0	1047	920	0
13	S	866	0	866	677	0
14	T	406	0	359	163	0
15	W	860	0	909	557	0
16	X	602	0	558	460	0
17	Y	879	0	860	755	0
18	Z	1360	0	1378	902	0
19	R	855	0	904	579	0
20	U	978	0	996	895	0
21	V	787	0	782	635	0
22	2	494	0	504	393	0
23	3	477	0	527	460	0
24	0	641	0	661	531	0
25	4	604	0	587	470	0
26	5	457	0	456	279	0
27	6	431	0	454	289	0
28	7	383	0	409	382	0
29	8	496	0	539	347	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
30	9	285	0	312	198	0
31	K	999	0	1064	573	0
All	All	87269	0	56176	25597	142

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 179.

The worst 5 of 25597 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
26:5:33:CYS:SG	26:5:36:CYS:HB2	1.24	1.69
2:A:2470:G:C2	2:A:2471:C:C5	1.81	1.68
28:7:30:ILE:HA	28:7:33:ARG:CD	1.21	1.67
2:A:2712:U:C6	2:A:712(A):A:C8	1.77	1.67
2:A:2580:U:C6	2:A:2581:G:C8	1.82	1.66

The worst 5 of 142 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	Distance(Å)	Clash(Å)
2:A:6:A:O4'	2:A:2902:C:C2'[8_554]	0.72	1.48
2:A:2899:G:N1	2:A:2901:C:C4[8_554]	0.79	1.41
2:A:6:A:C4'	2:A:2902:C:C2'[8_554]	0.97	1.23
2:A:2900:A:N7	2:A:2900:A:N6[8_554]	1.03	1.17
2:A:3:U:O4	2:A:2899:G:O2'[8_554]	1.09	1.11

## 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	D	169/173 (98%)	60 (36%)	34 (20%)	75 (44%)	0	0
4	E	183/338 (54%)	89 (49%)	34 (19%)	60 (33%)	0	0
5	F	179/246 (73%)	51 (28%)	47 (26%)	81 (45%)	0	0
6	G	116/176 (66%)	46 (40%)	31 (27%)	39 (34%)	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
7	H	162/177 (92%)	74 (46%)	39 (24%)	49 (30%)	0	1
8	I	144/149 (97%)	71 (49%)	29 (20%)	44 (31%)	0	1
9	N	111/145 (77%)	34 (31%)	21 (19%)	56 (50%)	0	0
10	O	120/122 (98%)	61 (51%)	27 (22%)	32 (27%)	0	2
11	P	82/164 (50%)	28 (34%)	21 (26%)	33 (40%)	0	0
12	Q	130/138 (94%)	38 (29%)	35 (27%)	57 (44%)	0	0
13	S	105/186 (56%)	36 (34%)	20 (19%)	49 (47%)	0	0
14	T	48/66 (73%)	17 (35%)	13 (27%)	18 (38%)	0	0
15	W	104/113 (92%)	41 (39%)	16 (15%)	47 (45%)	0	0
16	X	72/84 (86%)	26 (36%)	18 (25%)	28 (39%)	0	0
17	Y	108/119 (91%)	49 (45%)	20 (18%)	39 (36%)	0	0
18	Z	175/253 (69%)	52 (30%)	53 (30%)	70 (40%)	0	0
19	R	103/118 (87%)	35 (34%)	20 (19%)	48 (47%)	0	0
20	U	115/118 (98%)	22 (19%)	23 (20%)	70 (61%)	0	0
21	V	96/100 (96%)	39 (41%)	25 (26%)	32 (33%)	0	0
22	2	62/70 (89%)	8 (13%)	9 (14%)	45 (73%)	0	0
23	3	58/60 (97%)	24 (41%)	13 (22%)	21 (36%)	0	0
24	0	84/91 (92%)	32 (38%)	17 (20%)	35 (42%)	0	0
25	4	71/73 (97%)	21 (30%)	16 (22%)	34 (48%)	0	0
26	5	56/60 (93%)	16 (29%)	18 (32%)	22 (39%)	0	0
27	6	51/82 (62%)	21 (41%)	9 (18%)	21 (41%)	0	0
28	7	44/47 (94%)	4 (9%)	7 (16%)	33 (75%)	0	0
29	8	61/64 (95%)	22 (36%)	10 (16%)	29 (48%)	0	0
30	9	33/36 (92%)	14 (42%)	9 (27%)	10 (30%)	0	1
31	K	129/141 (92%)	73 (57%)	18 (14%)	38 (30%)	0	1
All	All	2971/3709 (80%)	1104 (37%)	652 (22%)	1215 (41%)	0	0

5 of 1215 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	64	ILE
3	D	66	PHE
3	D	67	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	D	82	TYR
3	D	85	ASN

### 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	D	135/135 (100%)	99 (73%)	36 (27%)	1	6
4	E	156/284 (55%)	128 (82%)	28 (18%)	2	19
5	F	152/193 (79%)	124 (82%)	28 (18%)	2	17
6	G	102/147 (69%)	93 (91%)	9 (9%)	14	57
7	H	137/147 (93%)	111 (81%)	26 (19%)	2	16
8	I	119/119 (100%)	98 (82%)	21 (18%)	3	20
9	N	95/121 (78%)	80 (84%)	15 (16%)	4	27
10	O	101/101 (100%)	81 (80%)	20 (20%)	2	14
11	P	67/126 (53%)	56 (84%)	11 (16%)	3	24
12	Q	110/110 (100%)	83 (76%)	27 (24%)	1	8
13	S	89/149 (60%)	73 (82%)	16 (18%)	2	18
14	T	44/52 (85%)	30 (68%)	14 (32%)	0	4
15	W	88/92 (96%)	74 (84%)	14 (16%)	4	26
16	X	67/73 (92%)	44 (66%)	23 (34%)	0	3
17	Y	97/105 (92%)	80 (82%)	17 (18%)	3	20
18	Z	151/203 (74%)	130 (86%)	21 (14%)	5	33
19	R	89/101 (88%)	71 (80%)	18 (20%)	2	14
20	U	96/97 (99%)	68 (71%)	28 (29%)	0	5
21	V	79/79 (100%)	69 (87%)	10 (13%)	6	37
22	2	51/56 (91%)	37 (72%)	14 (28%)	0	6
23	3	52/52 (100%)	47 (90%)	5 (10%)	12	52
24	0	64/67 (96%)	57 (89%)	7 (11%)	9	46
25	4	66/66 (100%)	54 (82%)	12 (18%)	2	18

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
26	5	51/53 (96%)	43 (84%)	8 (16%)	4	27
27	6	46/69 (67%)	39 (85%)	7 (15%)	4	28
28	7	39/40 (98%)	31 (80%)	8 (20%)	2	13
29	8	50/51 (98%)	39 (78%)	11 (22%)	1	11
30	9	34/35 (97%)	30 (88%)	4 (12%)	8	41
31	K	108/113 (96%)	104 (96%)	4 (4%)	45	85
All	All	2535/3036 (84%)	2073 (82%)	462 (18%)	2	18

5 of 462 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
12	Q	85	ILE
15	W	84	ARG
26	5	33	CYS
12	Q	134	PHE
13	S	94	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 80 such sidechains are listed below:

Mol	Chain	Res	Type
12	Q	148	ASN
17	Y	7	GLN
26	5	48	ASN
13	S	21	HIS
15	W	61	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	B	122/123 (99%)	45 (36%)	3 (2%)
2	A	2780/2916 (95%)	1487 (53%)	360 (12%)
All	All	2902/3039 (95%)	1532 (52%)	363 (12%)

5 of 1532 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	B	0	A
1	B	1	U

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	2	C
1	B	3	C
1	B	6	C

5 of 363 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	A	1329	U
2	A	1668	A
2	A	2717	G
2	A	1420	U
2	A	1544	C

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

There are no ligands in this entry.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 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	B	123/123 (100%)	0.73	18 (14%) 3 9	267, 323, 388, 388	0
2	A	2814/2916 (96%)	0.74	400 (14%) 3 9	65, 238, 393, 400	0
3	D	173/173 (100%)	1.90	67 (38%) 1 3	392, 392, 400, 400	0
4	E	191/338 (56%)	3.53	128 (67%) 0 3	400, 400, 400, 400	0
5	F	189/246 (76%)	0.56	24 (12%) 4 11	393, 396, 396, 396	0
6	G	122/176 (69%)	2.23	53 (43%) 1 3	400, 400, 400, 400	0
7	H	164/177 (92%)	0.70	21 (12%) 4 11	400, 400, 400, 400	0
8	I	148/149 (99%)	1.09	37 (25%) 1 5	400, 400, 400, 400	0
9	N	117/145 (80%)	1.99	49 (41%) 1 3	400, 400, 400, 400	0
10	O	122/122 (100%)	0.59	14 (11%) 5 13	400, 400, 400, 400	0
11	P	84/164 (51%)	0.44	7 (8%) 11 19	395, 395, 400, 400	0
12	Q	138/138 (100%)	1.37	40 (28%) 1 4	393, 393, 393, 393	0
13	S	113/186 (60%)	0.55	14 (12%) 5 12	278, 400, 400, 400	0
14	T	52/66 (78%)	4.04	41 (78%) 0 2	400, 400, 400, 400	0
15	W	108/113 (95%)	1.98	42 (38%) 1 3	278, 395, 400, 400	0
16	X	76/84 (90%)	0.81	14 (18%) 2 6	400, 400, 400, 400	0
17	Y	110/119 (92%)	1.13	26 (23%) 1 5	400, 400, 400, 400	0
18	Z	177/253 (69%)	0.19	8 (4%) 32 33	396, 398, 398, 398	0
19	R	105/118 (88%)	1.05	25 (23%) 1 5	400, 400, 400, 400	0
20	U	117/118 (99%)	0.33	5 (4%) 34 33	391, 391, 400, 400	0
21	V	100/100 (100%)	1.83	36 (36%) 1 4	400, 400, 400, 400	0
22	2	64/70 (91%)	0.50	12 (18%) 2 6	400, 400, 400, 400	0
23	3	60/60 (100%)	0.21	4 (6%) 17 24	398, 398, 398, 398	0
24	0	86/91 (94%)	1.78	32 (37%) 1 4	396, 400, 400, 400	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
25	4	73/73 (100%)	3.52	43 (58%) 0 3	396, 397, 397, 397	0
26	5	58/60 (96%)	0.87	9 (15%) 3 8	400, 400, 400, 400	0
27	6	53/82 (64%)	0.24	4 (7%) 14 21	400, 400, 400, 400	0
28	7	46/47 (97%)	1.63	16 (34%) 1 4	396, 396, 396, 396	0
29	8	63/64 (98%)	2.37	32 (50%) 0 3	400, 400, 400, 400	0
30	9	35/36 (97%)	4.51	25 (71%) 0 3	400, 400, 400, 400	0
31	K	133/141 (94%)	-0.02	1 (0%) 83 72	392, 400, 400, 400	0
All	All	6014/6748 (89%)	1.05	1247 (20%) 1 5	65, 392, 400, 400	0

The worst 5 of 1247 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
30	9	24	LEU	17.0
4	E	174	LEU	13.4
30	9	25	VAL	13.2
25	4	47	GLU	12.8
30	9	34	GLN	12.7

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

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

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.