



# wwPDB X-ray Structure Validation Summary Report i

Sep 15, 2014 – 08:12 PM EDT

PDB ID : 1VY1  
Title : Crystal Structure of Unmodified tRNA Proline (CGG) Bound to Codon CCG on the Ribosome  
Authors : Maehigashi, T.; Dunkle, J.A.; Dunham, C.M.  
Deposited on : 2014-03-25  
Resolution : 3.68 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>

---

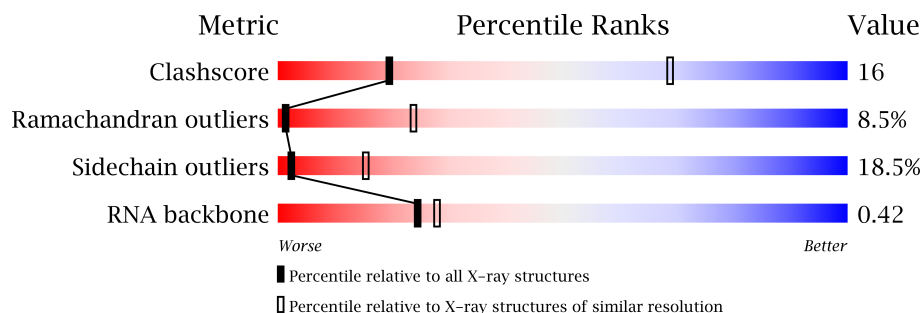
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.16 November 2013  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 21963  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable23489

# 1 Overall quality at a glance

The reported resolution of this entry is 3.68 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))
Clashscore	79885	1006 (3.92-3.44)
Ramachandran outliers	78287	1218 (3.96-3.40)
Sidechain outliers	78261	1216 (3.96-3.40)
RNA backbone	1838	1008 (4.52-2.80)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	2916	
2	B	122	
3	D	276	
4	E	206	
5	F	210	
6	G	182	
7	H	180	
8	I	148	
9	N	140	
10	O	122	
11	P	150	
12	Q	141	
13	R	118	
14	S	112	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
15	T	146	
16	U	118	
17	V	101	
18	W	113	
19	X	96	
20	Y	110	
21	Z	206	
22	0	85	
23	1	98	
24	2	72	
25	3	60	
26	4	71	
27	5	60	
28	6	54	
29	7	49	
30	8	65	
31	9	37	
32	a	3	

## 2 Entry composition

There are 33 unique types of molecules in this entry. The entry contains 92250 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 23S rRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	2882	Total	C	N	O	P	0	0	0
			62071	27627	11611	19952	2881			

- Molecule 2 is a RNA chain called 5S rRNA.

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

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	205	Total	C	N	O	S	0	0	0
			1568	991	300	271	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	202	Total	C	N	O	S	0	0	0
			1585	1011	297	275	2			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	G	181	Total	C	N	O	S	0	0	0
			1474	942	268	260	4			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	H	170	Total	C	N	O	S	0	0	0
			1307	829	245	232	1			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	I	146	Total	C	N	O	S	0	0	0
			1136	726	201	208	1			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	N	138	Total	C	N	O	S	0	0	0
			1104	712	206	182	4			

- 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
			933	588	171	170	4			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	P	150	Total	C	N	O	S	0	0	0
			1145	712	232	198	3			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	Q	141	Total	C	N	O	S	0	0	0
			1122	715	212	188	7			

- Molecule 13 is a protein called 50S ribosomal protein L17.

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

- Molecule 14 is a protein called 50S ribosomal protein L18.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
14	S	111	Total	C	N	O	0	0	0
			882	556	176	150			

- Molecule 15 is a protein called 50S ribosomal protein L19.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	T	137	Total	C	N	O	S	0	0	0
			1141	710	234	196	1			

- Molecule 16 is a protein called 50S ribosomal protein L20.

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

- Molecule 17 is a protein called 50S ribosomal protein L21.

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

- Molecule 18 is a protein called 50S ribosomal protein L22.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
18	W	113	Total	C	N	O	S	0	0	0
			900	566	177	155	2			

- Molecule 19 is a protein called 50S ribosomal protein L23.

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

- Molecule 20 is a protein called 50S ribosomal protein L24.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
20	Y	102	Total	C	N	O	S	0	0	0
			785	505	150	125	5			

- Molecule 21 is a protein called 50S ribosomal protein L25.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
21	Z	183	Total	C	N	O	S	0	0	0
			1461	933	260	265	3			

- Molecule 22 is a protein called 50S ribosomal protein L27.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
22	0	82	Total	C	N	O	S	0	0	0
			648	401	138	108	1			

- Molecule 23 is a protein called 50S ribosomal protein L28.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
23	1	97	Total	C	N	O	S	0	0	0
			763	481	150	131	1			

- Molecule 24 is a protein called 50S ribosomal protein L29.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
24	2	69	Total	C	N	O	S	0	0	0
			581	358	118	104	1			

- Molecule 25 is a protein called 50S ribosomal protein L30.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
25	3	59	Total	C	N	O	0	0	0
			469	298	90	81			

- Molecule 26 is a protein called 50S ribosomal protein L31.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
26	4	71	Total	C	N	O	S	0	0	0
			581	364	108	104	5			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
27	5	59	Total	C	N	O	S	0	0	0
			459	288	90	76	5			

- Molecule 28 is a protein called 50S ribosomal protein L33.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
28	6	49	Total	C	N	O	S	0	0	0
			424	264	87	69	4			

- Molecule 29 is a protein called 50S ribosomal protein L34.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
29	7	49	Total	C	N	O	S	0	0	0
			430	263	108	57	2			

- Molecule 30 is a protein called 50S ribosomal protein L35.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
30	8	64	Total	C	N	O	S	0	0	0
			517	331	102	82	2			

- Molecule 31 is a protein called 50S ribosomal protein L36.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
31	9	37	Total	C	N	O	S	0	0	0
			307	188	68	47	4			

- Molecule 32 is a RNA chain called tRNA acceptor end mimic.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
32	a	3	Total	C	N	O	P	0	0	0
			74	40	13	19	2			

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
33	P	2	Total	Mg	0	0
			2	2		
33	0	1	Total	Mg	0	0
			1	1		
33	D	1	Total	Mg	0	0
			1	1		
33	E	2	Total	Mg	0	0
			2	2		
33	B	2	Total	Mg	0	0
			2	2		
33	A	248	Total	Mg	0	0
			248	248		

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
33	5	1	Total 1	Mg 1	0	0
33	8	1	Total 1	Mg 1	0	0
33	F	1	Total 1	Mg 1	0	0

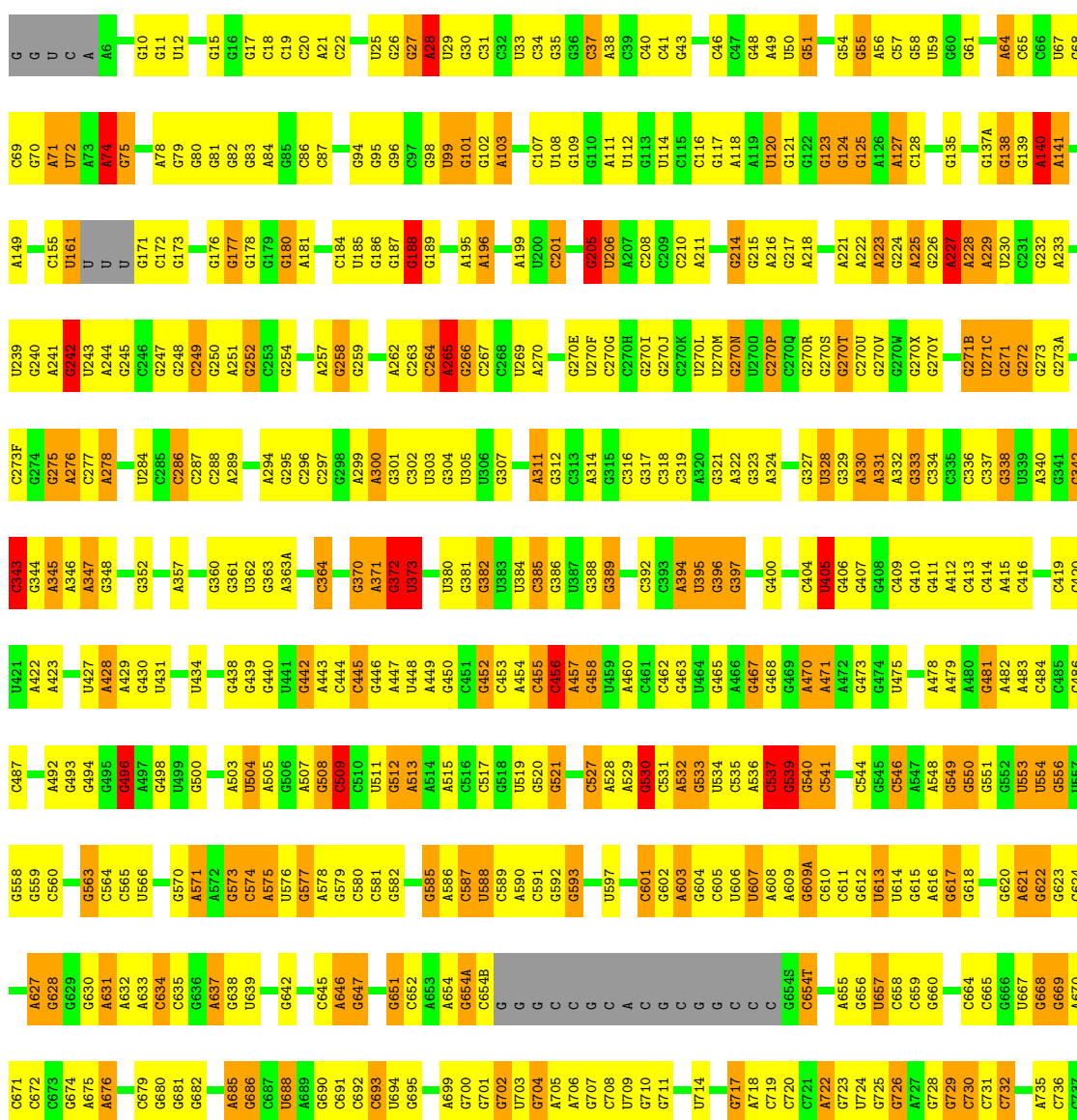
### 3 Residue-property plots

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.

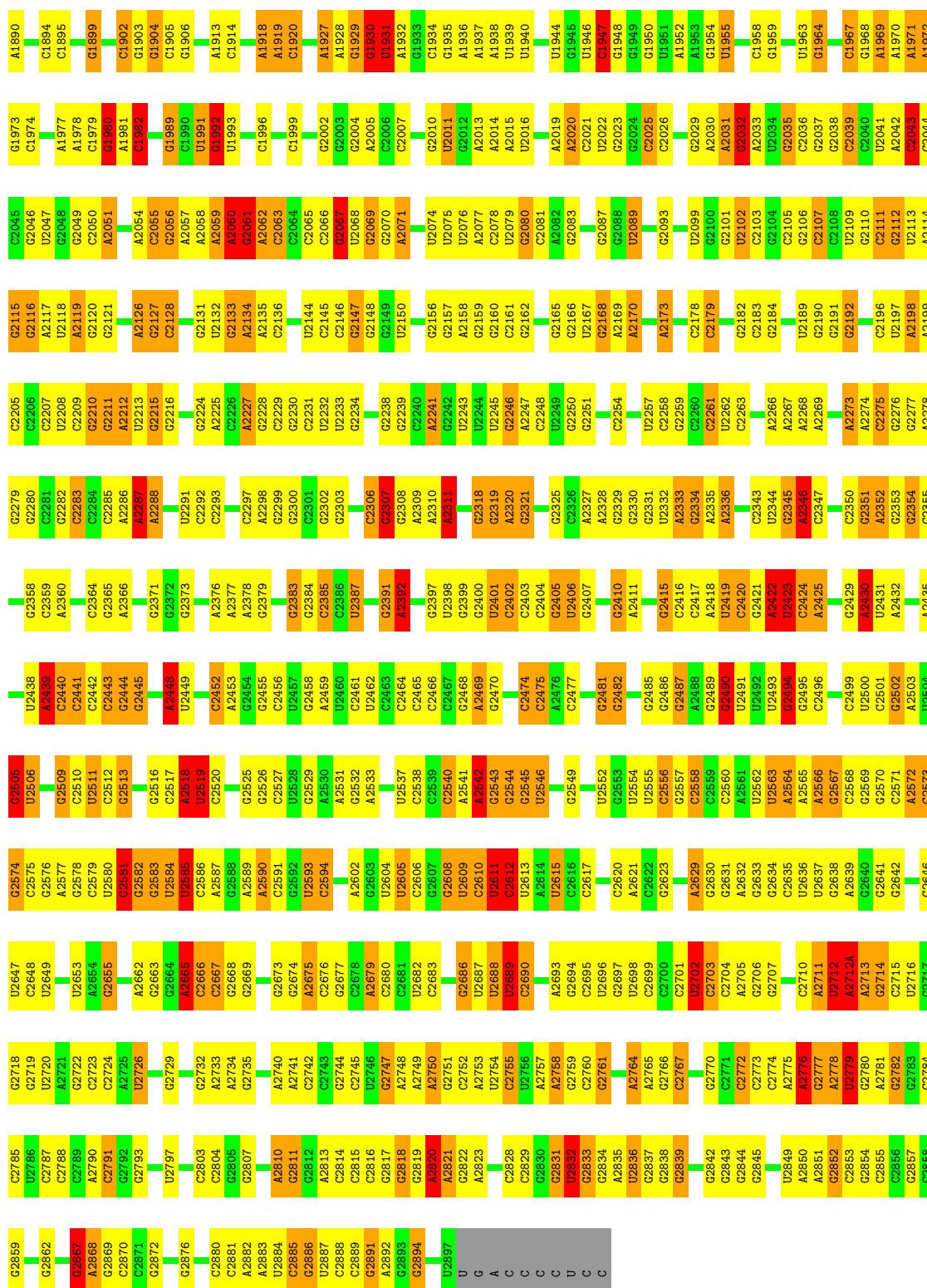
Note EDS was not executed.

#### • Molecule 1: 23S rRNA

Chain A:



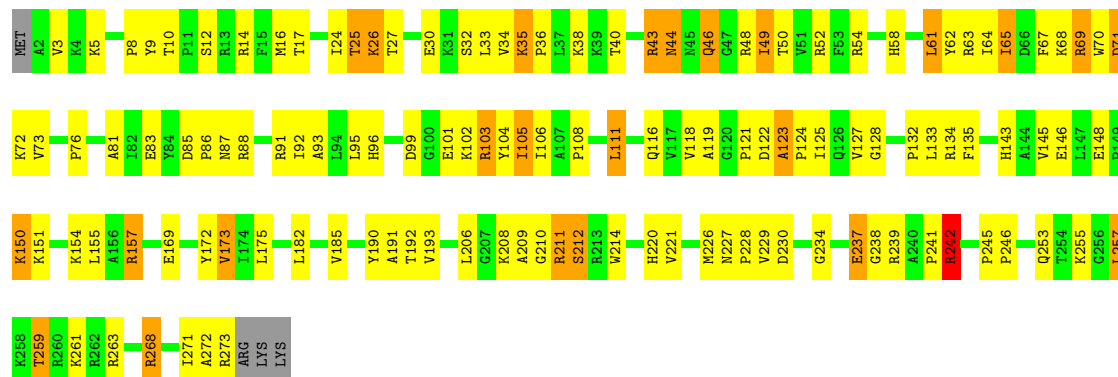
C1806	A1732	C1646	A1580	G1440	A1365	A1286	G1212	G1138	A1069	C1007	U943	A872	A804	G738
A1810	G1733	G1647	G1581	G1441	A1366	A1287	A1213	G1139	A1070	A1010	G944	G873	G805	G739
G1811	C1735	G1648	A1582	G1442	A1367	U1288	G1366	U1141	C1072	G1014	A945	G879	C806	G740
G1814	C1741	G1650	C1515	A1444A	C1370	C1291	G1219	A1142A	G1074	U1012	G947	G880	G808	G741
A1815	G1743	A1652	G1519	C1445	G1371	U1292	A1220	A1143	C1075	U1014	G948	G881	G809	G742
G1816	G1746	G1653	U1520	G1448	U1372	C1293	C1221	A1143	C1076	G1015	G952	G882	U810	G743
A1818	C1751	A1654	G1589	A1449	G1377	C1297	C1222	A1148	A1077	G1016	A953	G883	U811	A746
A1819	C1752	A1655	U1521	A1494	G1378	C1298	C1223	G1149	U1078	G1017	G954	G884	C812	U747
U1820	G1753	C1656	U1523	G1494A	A1379	C1299	G1224	G1149	C1079	G1018	G955	C885	U813	A750
G1754	C1754	C1657	U1524	C1450	A1378	U1300	G1225	G1151	C1080	U1019	G956	A887	C814	A751
A1821	C1755	A1658	G1525	C1451	G1380	U1301	G1226	G1152	U1081	A1020	A957	C888	C815	A752
G1822	A1756	C1658	G1526	A1455	G1381	A1302	A1227	C1153	U1082	A1021	U958	C889	C817	C753
A1825	G1756	C1663	U1527	U1454	G1382	C1303	A1228	A1155	U1083	G1022	A959	A990	C818	C754
A1826	A1757	A1664	A1528	G1455	C1383	G1302	G1229	A1155	A1084	U1023	A960	G892	A819	C755
C1827	A1762	G1667	G1529	A1460	A1384	C1306	G1229A	G1158	A1085	G1024	G962	C961	U822	C758
G1828	G1763	A1668	U1530	G1461	G1385	U1306	G1236	C1161	A1086	G1025	U963	A896	G823	C759
A1829	C1764	A1669	C1532	C1462	C1386	G1309	A1237	G1162	G1087	U1026	G964	C897	A824	G760
U1833	C1765	C1604	G1533	C1463	G1389	G1310	A1238	U1167	C1091	A1027	C965	C898	C825	A761
U1834	G1766	G1605	U1534	G1464	U1390	G1311	G1239	G1168	G1092	U1029	G966	A899	U826	U762
U1835	C1767	G1606	G1535	G1465	U1391	U1312	U1240	A1174	C1093	G1030	G967	A900	U827	G763
G1839	G1769	A1608	A1536	G1466	A1392	U1313	A1241	G1169	U1094	U1033	U969	A901	U828	A764
G1840	C1771	A1609	G1537	C1467	A1393	G1314	G1245	G1170	U1095	G1034	C970	C902	A829	G765
U1843	G1772	A1610	G1538	A1468	U1394	G1315	G1246	G1171	A1096	U1035	C971	C903	G830	C766
G1846	C1773	G1607	U1539	A1469	A1395	U1316	A1246	G1172	C1097	U1036	G972	C904	G831	U767
U1847	G1774	A1611	G1540	G1470	U1396	A1317	A1247	A1173	G1103	G1042	G977	C915	C840	G775
A1848	C1775	G1612	U1541	A1471	U1397	G1318	G1248	G1174	U1104	C1043	G978	C916	A841	G776
G1849	G1776	G1613	G1542	A1472	C1398	G1319	G1252	G1175	G1105	G1044	G979	G916	G842	A777
U1850	C1777	A1614	A1543	G1473	G1401	C1320	G1253	G1176	U1101	A1045	A980	A917	G843	G778
A1853	G1778	G1615	C1544	C1474	C1402	A1321	A1253	A1177	C1102	C1040	G975	A910	C838	U779
A1854	C1779	A1616	A1545	G1480	C1403	U1322	A1254	C1178	A1103	C1041	G976	C914	U839	A774
G1855	G1780	G1617	U1546	U1482	C1404	U1323	U1255	C1179	C1104	G1047	G977	C920	G847	A781
G1856	C1781	G1618	C1547	G1483	U1405	G1328	G1256	G1180	U106	C1049	C985	G921	G848	A782
A1857	G1782	A1619	G1548	G1484	U1406	U1329	C1257	G1181	G1114	A1050	C986	U922	A849	A783
G1858	C1783	U1620	C1549	G1485	C1407	C1330	G1259	A1182	G1115	C1051	G987	C923	A784	A784
A1859	A1784	U1621	A1553	A1486	C1408	U1331	G1260	G1183	C1116	G1052	A988	C924	G852	G785
G1860	G1785	G1622	A1554	G1487	C1409	A1332	C1261	C1184	A1111	C1053	G989	G925	G853	C786
U1864	C1786	U1623	C1555	U1488	G1410	G1332	G1262	C1185	G1112	A1054	A990	A926	G854	C787
G1865	A1787	G1624	G1556	G1489	C1411	C1333	A1262	G1186	U1113	G1048	C985	G928	G855	A788
A1869	G1788	G1625	C1557	U1489	C1412	U1334	U1263	G1187	G1114	C1049	C986	G929	C856	A789
G1880	C1789	G1626	G1558	A1490	A1412	U1335	G1264	U1188	C1115	C1052	G987	U930	C857	C790
U1884	G1790	G1627	A1559	G1491	G1416	U1336	A1265	A1189	C1116	G1053	G989	G931	U858	C791
G1885	A1791	U1630	G1560	G1492	C1417	G1337	G1266	G1190	G1120	C1054	A990	G932	G859	G792
C1870	G1792	C1630A	G1561	C1493	G1418	G1338	U1267	G1191	C1121	A1055	C991	A926	U860	A793
A1871	C1793	G1631	A1562	A1495	U1419	U1341	A1268	G1192	G1122	G1056	C992	G929	G861	C797
A1872	G1794	G1632	G1563	A1496	U1420	A1342	C1270	G1195	C1123	A1057	G993	U930	A861	G798
G1878	C1795	A1634	G1568	U1497	G1421	G1343	G1271	C1196	C1126	G1058	C994	G931	G862	G799
A1882	U1716	G1635	A1569	C1501	G1422	G1344	U1272	C1201	A1127	U1059	G995	G932	A863	G799
G1883	G1717	C1636	U1570	G1502	G1423	U1349	U1273	G1202	A1128	C1061	A996	A933	G864	A800
A1884	C1718	U1637	A1571	U1503	G1424	A1349	A1278	G1203	A1129	U1062	C997	G935	A866	G801
A1885	G1719	G1638	U1572	G1504	G1425	U1352	G1279	A1204	U1130	G1063	A1000	C936	G865	A802
A1886	C1720	C1639	G1573	C1505	A1426	A1353	U1282	U1205	G1131	C1064	A1001	U937	C866	G798
G1887	U1727	A1641	A1427	G1506	G1428	A1354	G1283	G1206	C1135	U1065	G1002	G938	A866	A800
A1888	G1728	G1642	U1576	C1507	C1432	A1359	U1284	A1210	G1136	U1066	G1003	A941	G867	A801
G1889	U1729	G1643	U1577	A1507	C1432	A1360	G1285	U1211	G1137	U1067	C1006	G942	U871	U803





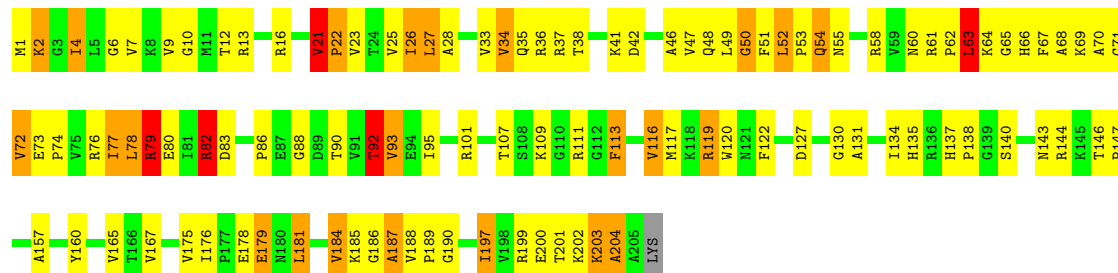
• Molecule 3: 50S ribosomal protein L2

Chain D:



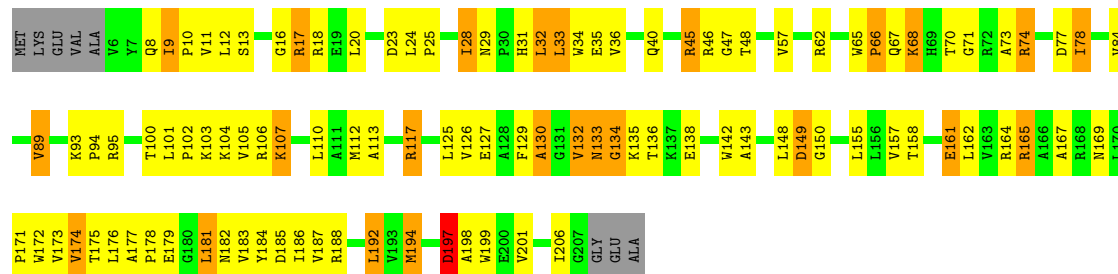
• Molecule 4: 50S ribosomal protein L3

Chain E:



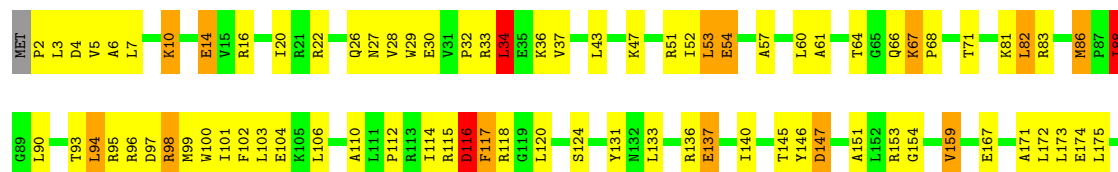
• Molecule 5: 50S ribosomal protein L4

Chain F:



• Molecule 6: 50S ribosomal protein L5

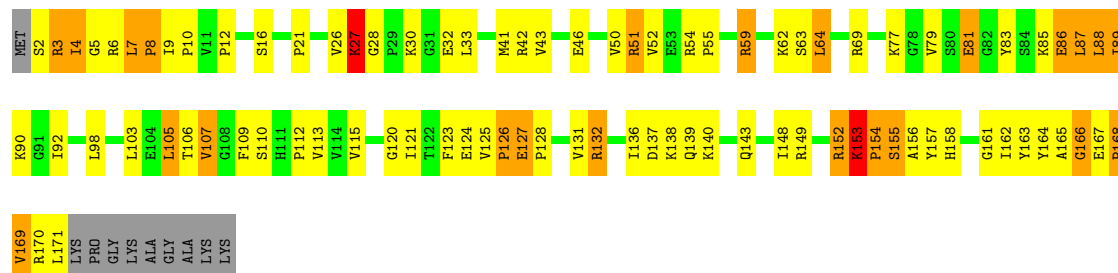
Chain G:





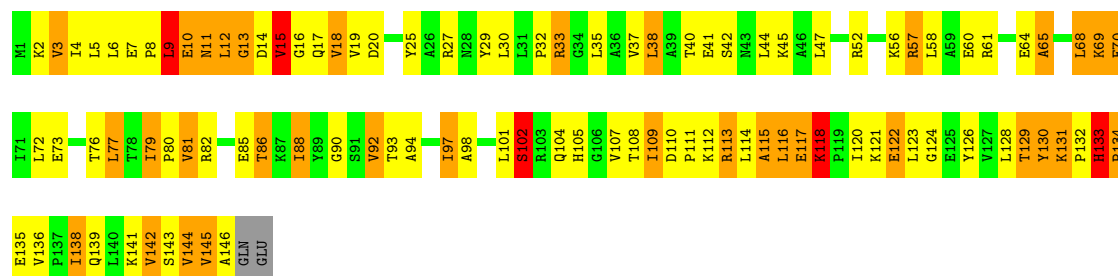
• Molecule 7: 50S ribosomal protein L6

Chain H:



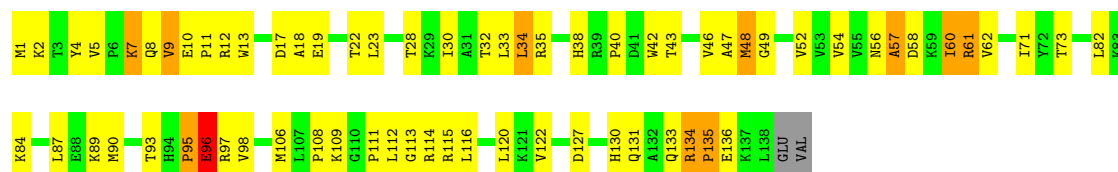
• Molecule 8: 50S ribosomal protein L9

Chain I:



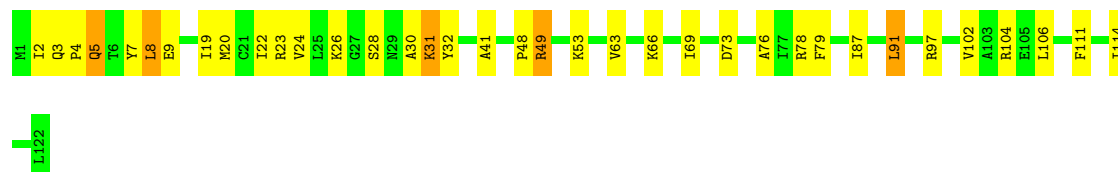
• Molecule 9: 50S ribosomal protein L13

Chain N:



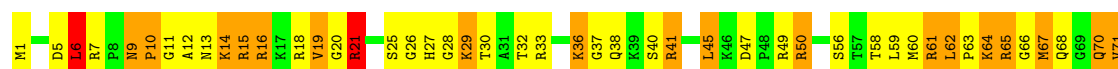
• Molecule 10: 50S ribosomal protein L14

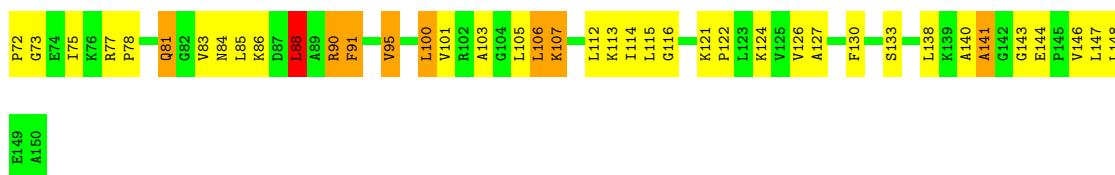
Chain O:



• Molecule 11: 50S ribosomal protein L15

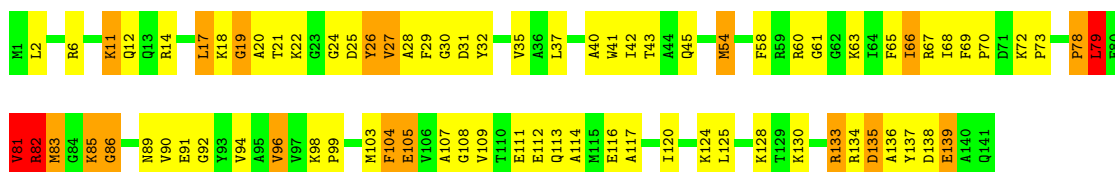
Chain P:





- Molecule 12: 50S ribosomal protein L16

Chain Q:



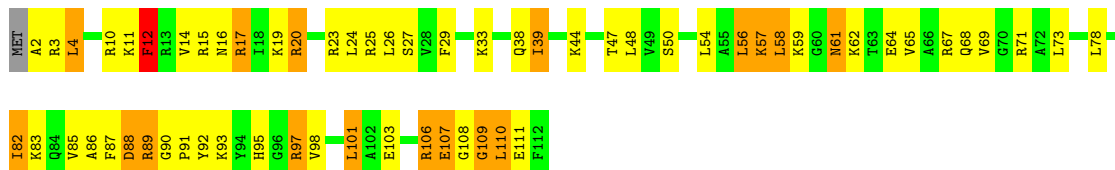
- Molecule 13: 50S ribosomal protein L17

Chain R:



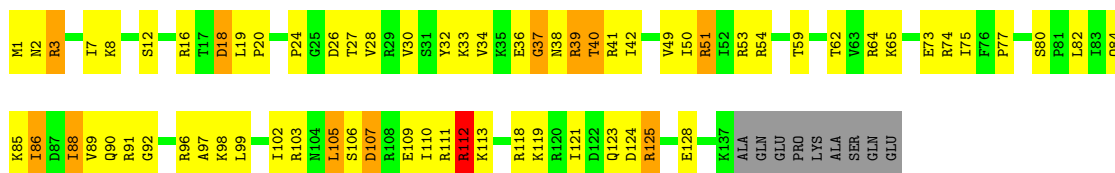
- Molecule 14: 50S ribosomal protein L18

Chain S:



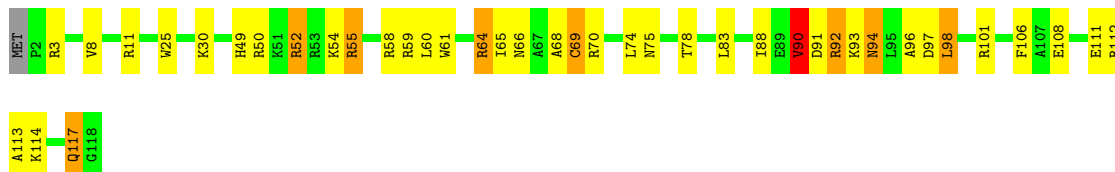
- Molecule 15: 50S ribosomal protein L19

Chain T:

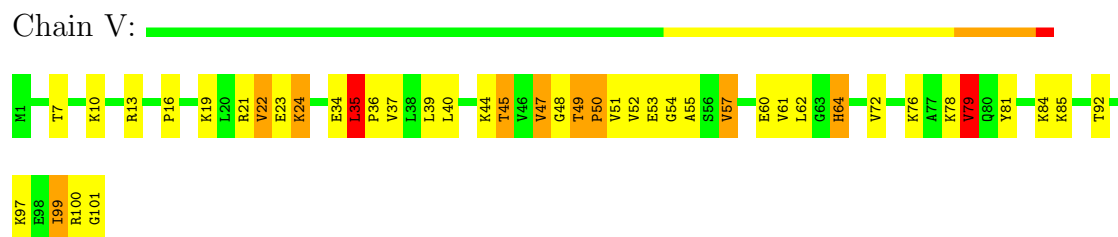


- Molecule 16: 50S ribosomal protein L20

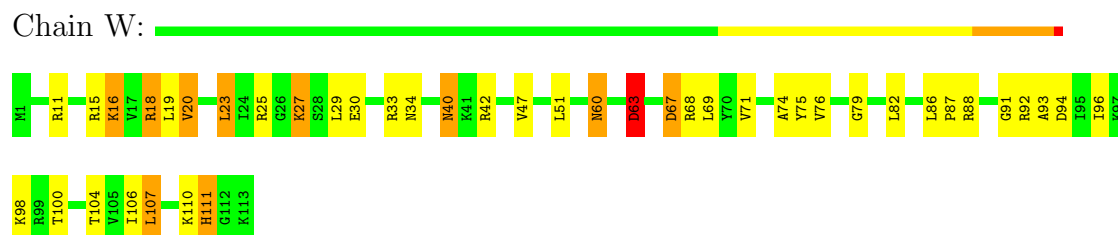
Chain U:



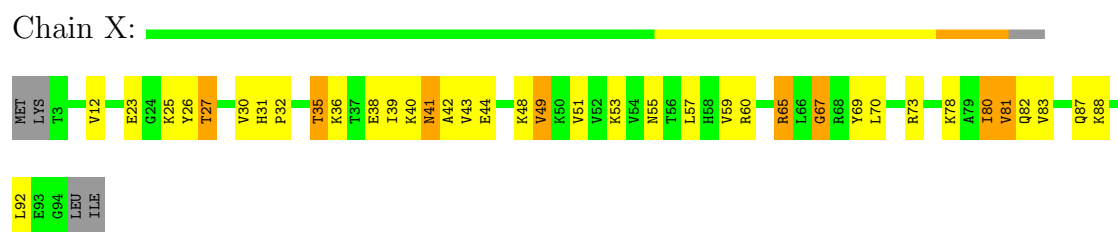
- Molecule 17: 50S ribosomal protein L21



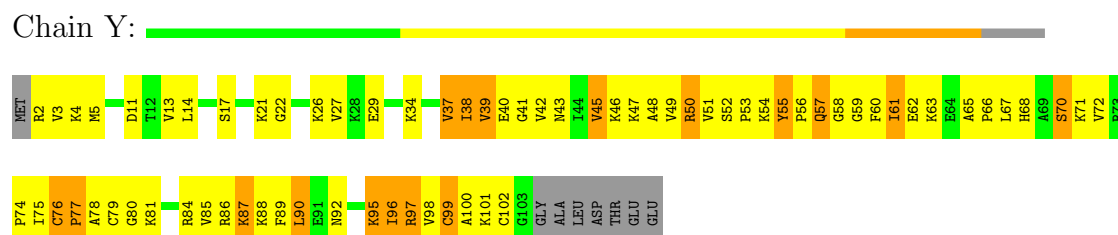
- Molecule 18: 50S ribosomal protein L22



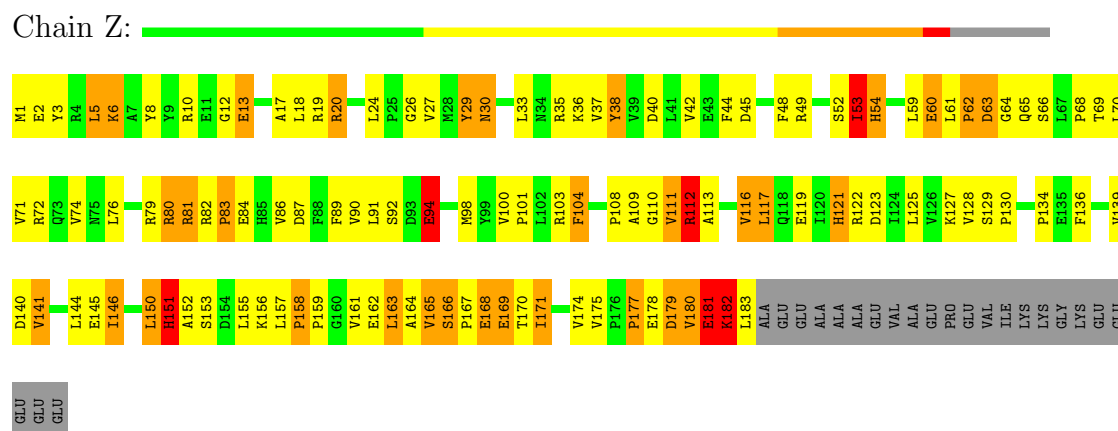
- Molecule 19: 50S ribosomal protein L23



- Molecule 20: 50S ribosomal protein L24



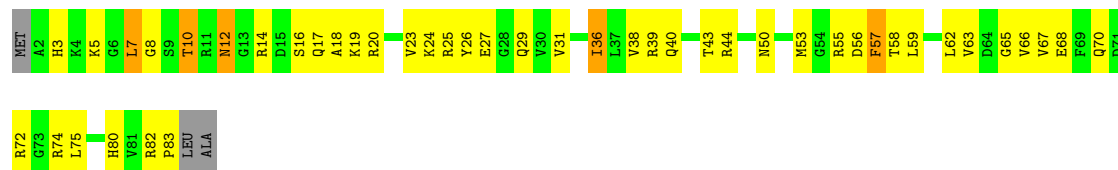
- Molecule 21: 50S ribosomal protein L25





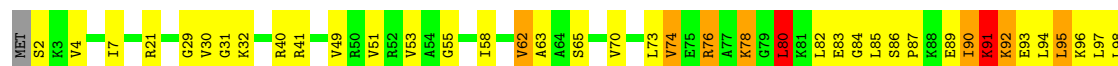
- Molecule 22: 50S ribosomal protein L27

Chain 0:



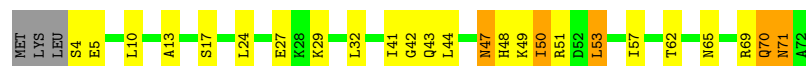
- Molecule 23: 50S ribosomal protein L28

Chain 1:



- Molecule 24: 50S ribosomal protein L29

Chain 2:



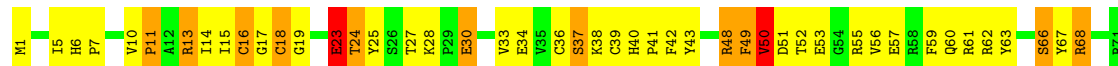
- Molecule 25: 50S ribosomal protein L30

Chain 3:



- Molecule 26: 50S ribosomal protein L31

Chain 4:



- Molecule 27: 50S ribosomal protein L32

Chain 5:



- Molecule 28: 50S ribosomal protein L33

Chain 6:



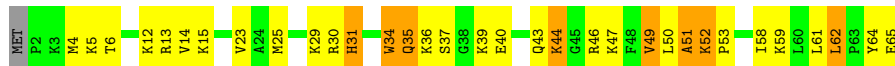
- Molecule 29: 50S ribosomal protein L34

Chain 7:



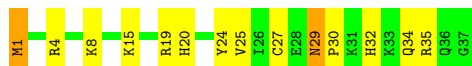
- Molecule 30: 50S ribosomal protein L35

Chain 8:



- Molecule 31: 50S ribosomal protein L36

Chain 9:



- Molecule 32: tRNA acceptor end mimic

Chain a:



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	209.82Å 447.39Å 619.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	187.58 – 3.68	Depositor
% Data completeness (in resolution range)	99.1 (187.58-3.68)	Depositor
$R_{merge}$	0.34	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	PHENIX (phenix.refine: 1.8.3_1479)	Depositor
R, $R_{free}$	0.212 , 0.272	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	92250	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	70.0	wwPDB-VP

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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

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	0.72	9/69521 (0.0%)	1.34	555/108529 (0.5%)
2	B	0.58	0/2878	1.22	15/4490 (0.3%)
3	D	0.51	0/2165	0.70	0/2919
4	E	0.43	0/1601	0.73	3/2160 (0.1%)
5	F	0.42	0/1620	0.62	0/2194
6	G	0.31	0/1499	0.57	1/2016 (0.0%)
7	H	0.29	0/1332	0.58	0/1802
8	I	0.52	0/1151	0.79	1/1558 (0.1%)
9	N	0.41	0/1131	0.62	0/1525
10	O	0.41	0/943	0.62	1/1269 (0.1%)
11	P	0.44	0/1162	0.81	1/1544 (0.1%)
12	Q	0.47	0/1143	0.74	2/1527 (0.1%)
13	R	0.42	0/982	0.69	0/1312
14	S	0.36	0/892	0.65	0/1187
15	T	0.42	0/1155	0.63	0/1542
16	U	0.40	0/982	0.65	0/1306
17	V	0.38	0/790	0.61	1/1057 (0.1%)
18	W	0.49	0/911	0.67	0/1220
19	X	0.47	0/739	0.62	0/993
20	Y	0.44	0/798	0.68	0/1064
21	Z	0.58	1/1493 (0.1%)	0.77	0/2026
22	0	0.65	0/657	0.80	0/874
23	1	0.44	0/770	0.66	0/1022
24	2	0.39	0/583	0.65	0/771
25	3	0.35	0/474	0.57	0/635
26	4	0.33	0/594	0.68	0/795
27	5	0.44	0/473	0.73	0/639
28	6	0.35	0/431	0.69	0/575
29	7	0.49	0/438	0.68	0/575
30	8	0.55	0/525	0.79	0/691
31	9	0.26	0/310	0.45	0/407
32	a	0.74	0/40	1.58	1/60 (1.7%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
All	All	0.65	10/100183 (0.0%)	1.21	581/150284 (0.4%)

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
4	E	0	1
7	H	0	2
8	I	0	1
26	4	0	1
30	8	0	2
All	All	0	7

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1918	A	N9-C4	-6.88	1.33	1.37
1	A	471	A	N9-C4	-6.35	1.34	1.37
1	A	2589	A	N9-C4	-5.73	1.34	1.37
1	A	1677	A	N9-C4	-5.56	1.34	1.37
21	Z	54	HIS	CG-ND1	-5.27	1.27	1.38

The worst 5 of 581 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1931	U	N3-C2-O2	-12.05	113.77	122.20
1	A	28	A	C8-N9-C4	-10.75	101.50	105.80
1	A	140	A	C8-N9-C4	-9.94	101.82	105.80
1	A	1931	U	C5-C4-O4	9.85	131.81	125.90
1	A	774	A	C2-N3-C4	-9.80	105.70	110.60

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
26	4	38	LYS	Peptide
4	E	21	VAL	Peptide
7	H	127	GLU	Peptide
7	H	153	LYS	Peptide

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
8	I	134	PRO	Peptide

## 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	A	62071	0	31292	1367	0
2	B	2573	0	1306	57	0
3	D	2115	0	2195	106	0
4	E	1568	0	1634	67	0
5	F	1585	0	1632	76	0
6	G	1474	0	1535	54	0
7	H	1307	0	1382	67	0
8	I	1136	0	1223	84	0
9	N	1104	0	1180	38	0
10	O	933	0	996	21	0
11	P	1145	0	1227	79	0
12	Q	1122	0	1179	58	0
13	R	968	0	1033	48	0
14	S	882	0	943	45	0
15	T	1141	0	1202	45	0
16	U	964	0	1022	36	0
17	V	779	0	852	24	0
18	W	900	0	964	28	0
19	X	725	0	778	31	0
20	Y	785	0	878	52	0
21	Z	1461	0	1493	96	0
22	0	648	0	671	37	0
23	1	763	0	848	32	0
24	2	581	0	629	14	0
25	3	469	0	518	7	0
26	4	581	0	574	21	0
27	5	459	0	480	30	0
28	6	424	0	450	30	0
29	7	430	0	480	18	0
30	8	517	0	582	36	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
31	9	307	0	338	14	0
32	a	74	0	51	0	0
33	0	1	0	0	0	0
33	5	1	0	0	0	0
33	8	1	0	0	0	0
33	A	248	0	0	0	0
33	B	2	0	0	0	0
33	D	1	0	0	0	0
33	E	2	0	0	0	0
33	F	1	0	0	0	0
33	P	2	0	0	0	0
All	All	92250	0	61567	2407	0

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 16.

The worst 5 of 2407 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:2701:C:H3'	1:A:2702:U:H5''	1.27	1.07
1:A:1310:G:OP2	29:7:9:ARG:NH1	1.91	1.01
1:A:617:G:OP1	5:F:40:GLN:NE2	1.95	0.98
1:A:768:G:O2'	1:A:1379:A:N6	1.99	0.96
1:A:2392:A:H8	11:P:60:MET:HG2	1.25	0.96

There are no symmetry-related clashes.

## 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	270/276 (98%)	226 (84%)	32 (12%)	12 (4%)	4	45
4	E	203/206 (98%)	147 (72%)	36 (18%)	20 (10%)	1	18

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
5	F	200/210 (95%)	167 (84%)	20 (10%)	13 (6%)	2	33
6	G	179/182 (98%)	139 (78%)	26 (14%)	14 (8%)	1	26
7	H	168/180 (93%)	114 (68%)	33 (20%)	21 (12%)	1	12
8	I	144/148 (97%)	94 (65%)	31 (22%)	19 (13%)	0	11
9	N	136/140 (97%)	104 (76%)	20 (15%)	12 (9%)	1	22
10	O	120/122 (98%)	109 (91%)	9 (8%)	2 (2%)	14	69
11	P	148/150 (99%)	107 (72%)	27 (18%)	14 (10%)	1	20
12	Q	139/141 (99%)	99 (71%)	22 (16%)	18 (13%)	0	11
13	R	116/118 (98%)	106 (91%)	5 (4%)	5 (4%)	4	46
14	S	109/112 (97%)	76 (70%)	22 (20%)	11 (10%)	1	17
15	T	135/146 (92%)	106 (78%)	17 (13%)	12 (9%)	1	22
16	U	115/118 (98%)	102 (89%)	9 (8%)	4 (4%)	6	53
17	V	99/101 (98%)	82 (83%)	11 (11%)	6 (6%)	2	34
18	W	111/113 (98%)	99 (89%)	8 (7%)	4 (4%)	5	52
19	X	90/96 (94%)	77 (86%)	11 (12%)	2 (2%)	10	64
20	Y	100/110 (91%)	71 (71%)	13 (13%)	16 (16%)	0	6
21	Z	181/206 (88%)	118 (65%)	35 (19%)	28 (16%)	0	7
22	0	80/85 (94%)	61 (76%)	15 (19%)	4 (5%)	3	41
23	1	95/98 (97%)	75 (79%)	11 (12%)	9 (10%)	1	20
24	2	67/72 (93%)	53 (79%)	9 (13%)	5 (8%)	2	28
25	3	57/60 (95%)	52 (91%)	3 (5%)	2 (4%)	6	53
26	4	69/71 (97%)	35 (51%)	18 (26%)	16 (23%)	0	1
27	5	57/60 (95%)	44 (77%)	11 (19%)	2 (4%)	6	53
28	6	47/54 (87%)	23 (49%)	13 (28%)	11 (23%)	0	1
29	7	47/49 (96%)	45 (96%)	1 (2%)	1 (2%)	11	65
30	8	62/65 (95%)	51 (82%)	6 (10%)	5 (8%)	1	25
31	9	35/37 (95%)	35 (100%)	0	0	100	100
All	All	3379/3526 (96%)	2617 (77%)	474 (14%)	288 (8%)	1	24

5 of 288 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	26	LYS

Continued on next page...



*Continued from previous page...*

Mol	Chain	Res	Type
3	D	122	ASP
3	D	242	ARG
4	E	22	PRO
4	E	53	PRO

### 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	214/218 (98%)	174 (81%)	40 (19%)	2	15
4	E	165/166 (99%)	126 (76%)	39 (24%)	1	7
5	F	161/166 (97%)	132 (82%)	29 (18%)	2	17
6	G	155/156 (99%)	134 (86%)	21 (14%)	6	33
7	H	142/148 (96%)	120 (84%)	22 (16%)	4	25
8	I	122/124 (98%)	86 (70%)	36 (30%)	0	4
9	N	117/119 (98%)	97 (83%)	20 (17%)	3	20
10	O	100/100 (100%)	90 (90%)	10 (10%)	11	50
11	P	116/116 (100%)	85 (73%)	31 (27%)	1	5
12	Q	111/111 (100%)	95 (86%)	16 (14%)	5	29
13	R	101/101 (100%)	83 (82%)	18 (18%)	2	18
14	S	87/88 (99%)	69 (79%)	18 (21%)	2	10
15	T	120/127 (94%)	102 (85%)	18 (15%)	4	28
16	U	93/94 (99%)	78 (84%)	15 (16%)	3	24
17	V	82/82 (100%)	66 (80%)	16 (20%)	2	13
18	W	92/92 (100%)	73 (79%)	19 (21%)	2	10
19	X	74/78 (95%)	64 (86%)	10 (14%)	6	33
20	Y	85/91 (93%)	63 (74%)	22 (26%)	1	6
21	Z	162/179 (90%)	131 (81%)	31 (19%)	2	14
22	0	65/67 (97%)	56 (86%)	9 (14%)	5	31
23	1	82/83 (99%)	73 (89%)	9 (11%)	9	44

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
24	2	64/67 (96%)	57 (89%)	7 (11%)	9	45
25	3	51/52 (98%)	45 (88%)	6 (12%)	8	39
26	4	63/63 (100%)	45 (71%)	18 (29%)	0	4
27	5	51/52 (98%)	37 (72%)	14 (28%)	0	5
28	6	48/52 (92%)	35 (73%)	13 (27%)	1	5
29	7	42/42 (100%)	34 (81%)	8 (19%)	2	14
30	8	54/55 (98%)	44 (82%)	10 (18%)	2	15
31	9	34/34 (100%)	32 (94%)	2 (6%)	28	75
All	All	2853/2923 (98%)	2326 (82%)	527 (18%)	2	15

5 of 527 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
11	P	88	LEU
14	S	59	LYS
27	5	36	CYS
11	P	138	LEU
13	R	18	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
8	I	104	GLN
31	9	32	HIS
22	0	12	ASN
7	H	147	ASN
14	S	34	HIS

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	2879/2916 (98%)	747 (25%)	65 (2%)
2	B	119/122 (97%)	29 (24%)	2 (1%)
32	a	1/3 (33%)	0	0
All	All	2999/3041 (98%)	776 (25%)	67 (2%)

5 of 776 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	10	G
1	A	15	G
1	A	28	A
1	A	34	C
1	A	35	G

5 of 67 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	1045	A
1	A	1312	U
1	A	2726	U
1	A	1078	U
1	A	1178	C

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

1 non-standard protein/DNA/RNA residue is 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$
32	PPU	a	76	1,32	38,40,41	1.41	3 (7%)	54,57,60	2.40	18 (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
32	PPU	a	76	1,32	-	0/26/43/44	0/4/4/4

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
32	a	76	PPU	P-OP1	4.78	1.52	1.46

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
32	a	76	PPU	C4-N9	-3.43	1.32	1.37
32	a	76	PPU	C5-C4	3.01	1.47	1.40

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
32	a	76	PPU	C5-C4-N3	-7.10	119.05	125.98
32	a	76	PPU	C3'-N3'-C	-5.64	114.21	123.19
32	a	76	PPU	N3-C4-N9	5.47	134.77	125.39
32	a	76	PPU	N3-C2-N1	-4.87	124.61	128.89
32	a	76	PPU	N1-C6-N6	4.64	121.93	117.04

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 259 ligands modelled in this entry, 259 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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 ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.