



# wwPDB X-ray Structure Validation Summary Report

Sep 15, 2014 – 07:32 PM EDT

PDB ID : 1VY3  
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>

---

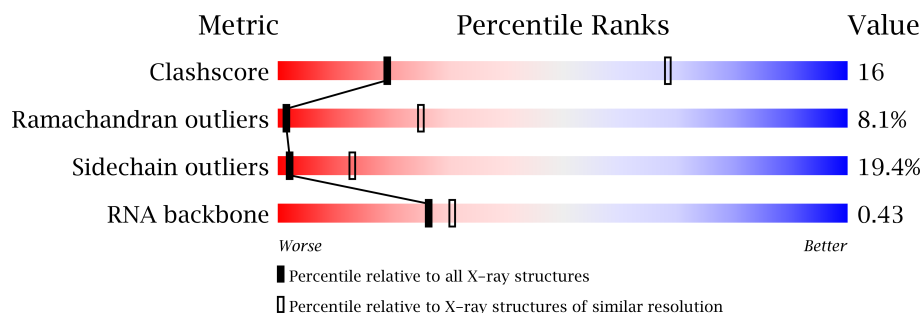
The following versions of software and data (see [references](#)) were used in the production of this report:

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 92287 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	2883	Total	C	N	O	P	0	0	0
			62091	27636	11613	19960	2882			

- 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	Q	1	Total	Mg	0	0
			1	1		
33	D	2	Total	Mg	0	0
			2	2		
33	B	3	Total	Mg	0	0
			3	3		
33	A	266	Total	Mg	0	0
			266	266		
33	5	1	Total	Mg	0	0
			1	1		

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
33	X	1	Total	Mg	0	0
			1	1		

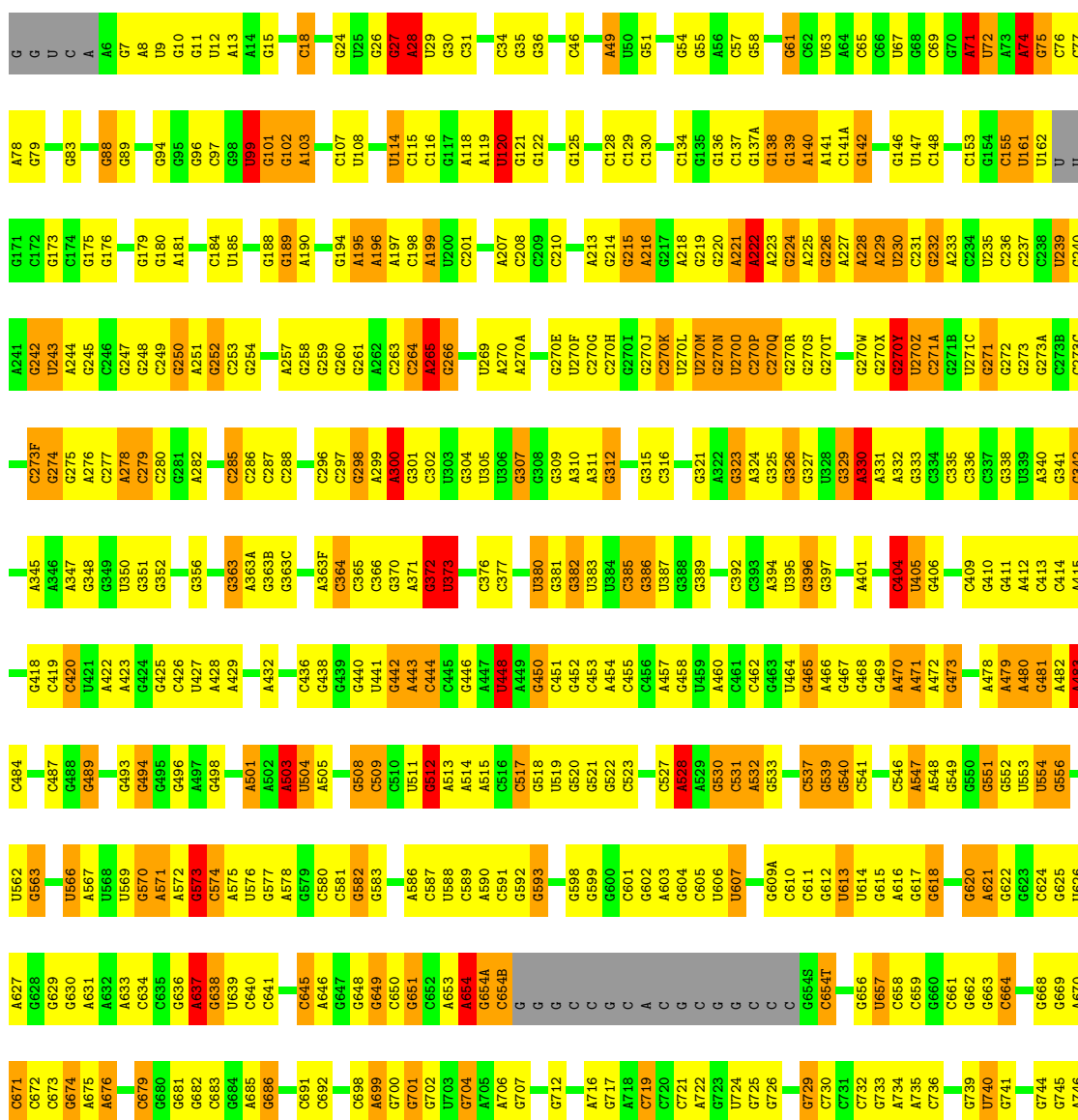
### 3 Residue-property plots

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.

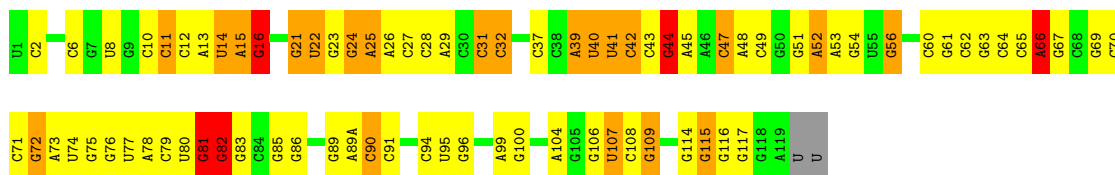
#### • Molecule 1: 23S rRNA

Chain A:



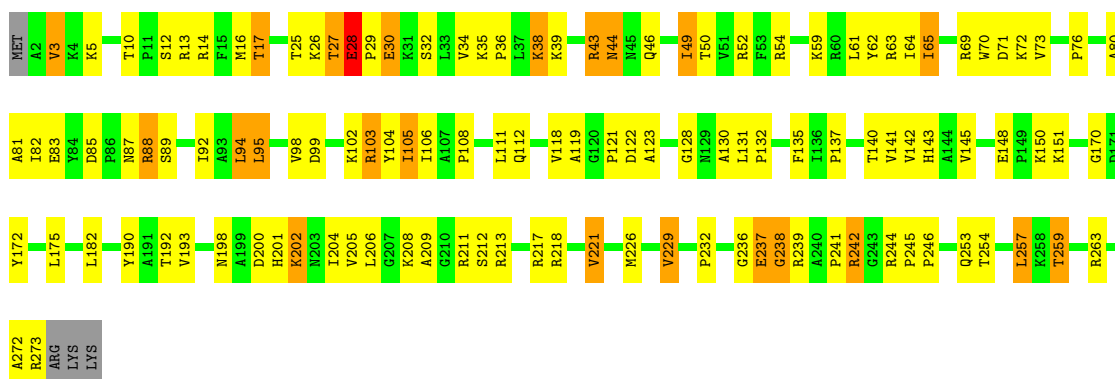
G1827	G1761	G1667	G1521	G1454	G1382	G1311	G1238	G1171	C1041	G968	U895	G818	U747
G1828	A1762	A1668	G1522	G1455	C1383	U1312	G1239	G1173	G1042	U969	A896	A819	G746
A1829	G1763	A1669	G1525	C1458	A1384	U1313	U1240	U1174	G1043	C970	C897	C825	A752
G1831	G1674	G1599	G1526	C1459	G1385	C1314	A1241	U1175	G1044	C971	C898	U826	C753
G1832	G1675	G1600	G1527	A1460	C1386	G1319	A1242	A1177	A1045	A972	A899	U827	C754
G1833	U1678	U1601	A1528	G1461	U1389	C1320	G1243	C1178	A1046	G974	A900	U828	
U1834	G1769	U1602	A1529	G1462	U1390	A1321	G1244	C1179	G1047	C974A	A901	A829	C758
G1835	G1770	G1530	U1391	C1464	U1391	G1324	A1247	C1180	A1048	G975	C902	G830	G759
C1836		G1531	U1392	G1465	U1392			C1181	A1050	C976	C904	C834	G760
C1837	A1773	C1532	U1393	G1466	U1394	G1327	G1250	A1182	G1051	G977		A835	A761
G1838	G1774	G1678	G1533	C1467	A1395	U1328	G1251	G1183	C1052	G978	U907	G762	G763
G1839	U1775	C1679	U1534	C1468		G1329	A1253	G1184	A1053	G979	A903	G764	A764
G1840	G1776	A1469		A1469	G1401	U1330			A1054	A980		C840	
U1841	U1777	A1536	G1537	A1470	C1402	C1331	A1254	G1187	G1055	A981	A910	A841	G768
G1842	U1778	G1613	U1538	G1471	C1403	G1332	U1255	U1188	G1056	C982	C912	G842	G769
C1843	U1779	G1615	G1539	G1472	U1405	G1333	G1256	A1189	A1057	A983	G914	G845	G770
C1844	A1780	G1616	U1540	G1473	U1406	C1334	C1257	G914	G1058	C985	C914	C846	G771
G1845	C1781	G1686	U1541	A1477	U1407	U1335	C1258	G1190	G1059	C986	C915	U847	C772
G1846	C1782	G1687	U1542	G1478	C1408	A1336	A1262	G1193	U1060	G987	C916	U848	
A1847	U1783	U1688	G1543	G1479	U1409	G1337	U1263	A1194	U1061	A988	G917	U849	A774
A1848	A1784	G1480	C1544	G1480	C1411	G1338	G1264	G1195	G1062	G989	A918		
A1853	A1785	G1482	A1545	G1483	A1412	G1339	A1265	U1198	G1063	A990		G852	G776
A1854	A1786	G1484	A1546	G1484	G1413	U1340	G1266	U1199	C1064	C991	C923	G854	A777
		G1485	C1546	G1485	U1415	U1341	U1267	C1200	U1065	C992		G855	G778
G1858		G1486		G1486	A1416	A1342	A1268	C1201	U1066	G993	G929	C856	
G1859	A1789	G1487	C1549	G1487	G1417	G1343	A1269	C1202	G1068	C994		C857	G780
G1863	G1790	G1697	C1550	G1488	C1417	G1344	C1270	G1203	A1069	A996	G932	U858	A781
U1864	A1791	A1698	C1551	G1488	G1418	G1271	G1271	A1204	G1070		G933	G859	A782
G1869	G1792	G1699	C1551	G1488	G1419	G1348	A1272	U1205	G1071		G934	U860	A783
C1870	C1793	A1700	A1584	U1489	U1420	A1349	U1273	G1206	C1072	A1001	C935	A861	A784
A1871	U1794	U1490		U1490	U1421		A1274	C1207	A1073	G1002	C936	A862	G785
A1872	G1795	C1493	G1558	C1493	G1422	U1352	A1275	G1208	C1076	G1003	U937	A863	G786
G1878	U1796	A1494	G1559	A1494	G1423	A1353		G1209	C1004	C938		G864	U787
C1879	U1797	A1495	G1560	A1495	G1424	A1354		A1210	C1005	G939	G940	C865	A788
C1880	G1798	A1496		A1496	G1425	G1355	A1278	U1211	U1077	A1009	A941	C866	A789
G1881	U1799	U1497	A1566	U1497	G1426		A1284	G1212	C1079	A1010	G942	C867	C790
C1882	G1801	C1498	A1567	C1498	A1427	G1358	G1285	A1213		A1011	U943	U868	C791
G1883	A1802	G1499	G1568	C1499	C1428	A1359	A1286	A1214	U1082	G1089	G952	G869	G792
A1884	C1803	G1500	A1569	A1500	G1429	A1360	A1287	G1215	U1083	A944	G945	A870	A793
	C1804	G1501		G1501	C1430	G1361	U1288	G1216	A1084	C1013	A945		G794
G1888	U1805	C1502	G1577	U1503	U1431	C1362	C1289	C1217	A1085	G1016	G946	G875	C795
U1806	C1806	U1504	U1578	G1504		C1363	C1290	G1218	A1086			C876	C796
G1807	G1734	C1505	A1579	C1505	A1434	G1364	C1291	G1219	G1087	G1016	G950		C797
		C1506	A1580	C1506	G1435		U1292	A1220	A1088	U1019	C951	G879	A800
	C1742	A1507	G1581	A1507	G1436	A1365	C1293	A1221	A1089	A1020	G952	G880	G801
	G1743	A1508	C1582	A1508	C1437		U1294	C1222	U1090	A1021	G953	G881	A802
		A1509	A1583	A1509	U1441	G1368		G1223	G1091	G1023	G956	G882	U803
		A1510	A1587	A1510	G1442	C1370	C1297	G1224	C1092	G1024		G883	A804
		A1511		A1511		G1371		C1225	G1093			C884	
		U1514		U1514	A1444A	U1372	U1300	G1228	U1094	G1025	A959	C885	
		C1445		C1445	G1446	A1373	A1301	G1230	U1095	U1026	A960	C886	U811
		U1516		U1516	G1447	G1374	A1302	G1231	A1096	A1027	C961	A887	C812
		G1517		G1517	G1448	C1375	G1303	G1233	U1097	A1028	G962	C888	C813
		U1518		U1518	G1449	C1377	C1306	G1235	G1098		U1033	C889	C814
		G1519		G1519	A1449A	G1378		G1236	G1099		C964	A890	C815
		U1520		U1520		A1379	G1309	G1237	C1100	G1039	C965	G892	C816
							G1310	A1237	C1102	C1040	C967	C894	C817





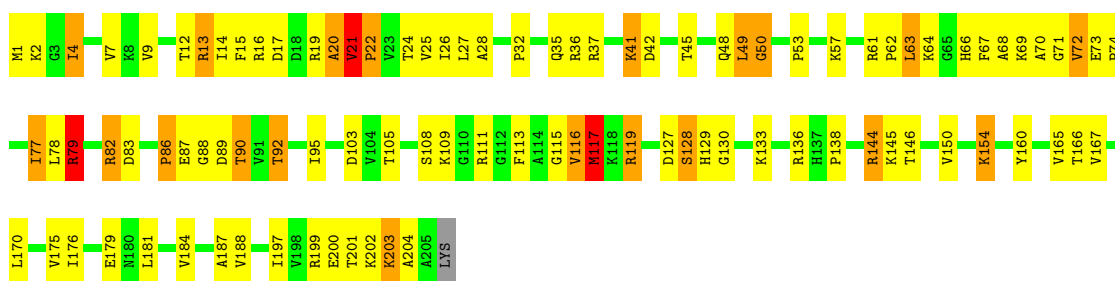
• Molecule 3: 50S ribosomal protein L2

Chain D:



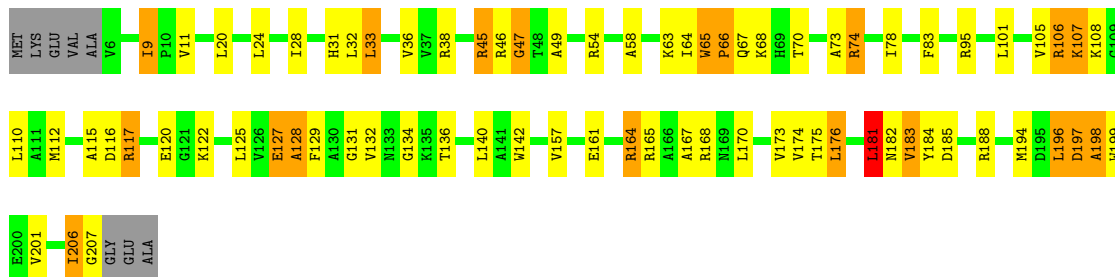
• Molecule 4: 50S ribosomal protein L3

Chain E:



• Molecule 5: 50S ribosomal protein L4

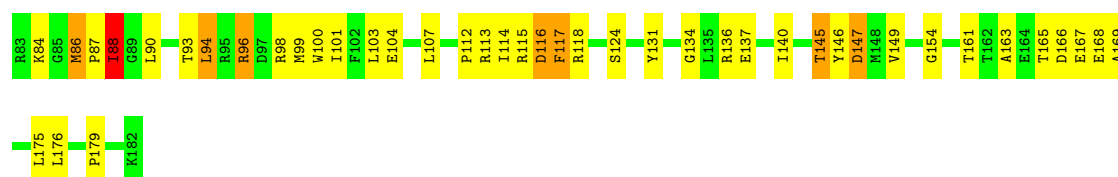
Chain F:



• Molecule 6: 50S ribosomal protein L5

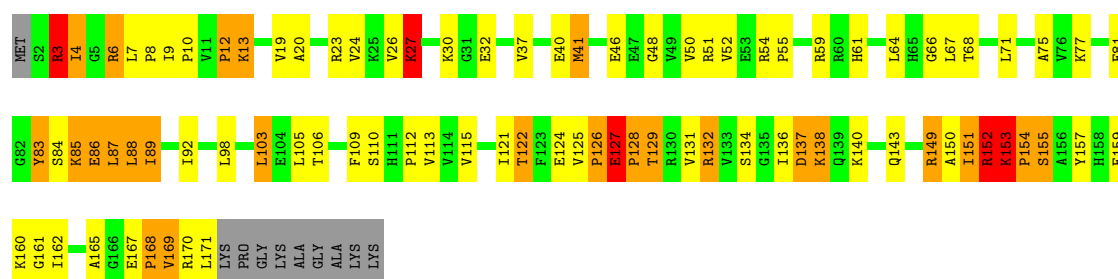
Chain G:





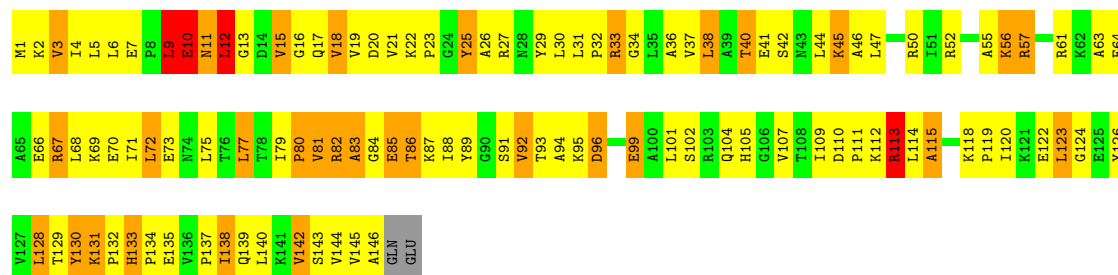
• Molecule 7: 50S ribosomal protein L6

Chain H:



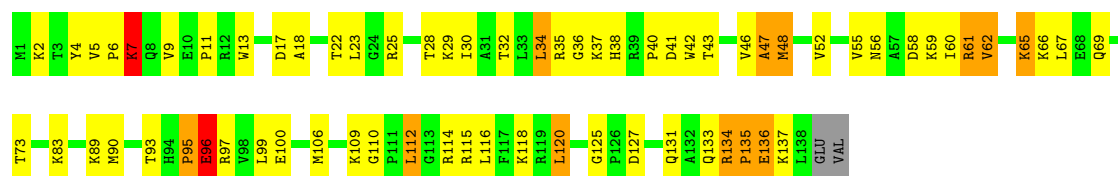
• Molecule 8: 50S ribosomal protein L9

Chain I:



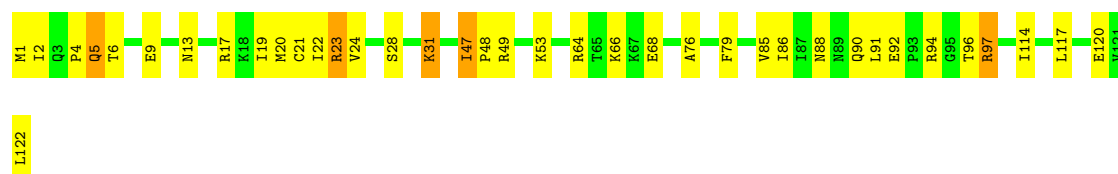
• Molecule 9: 50S ribosomal protein L13

Chain N:



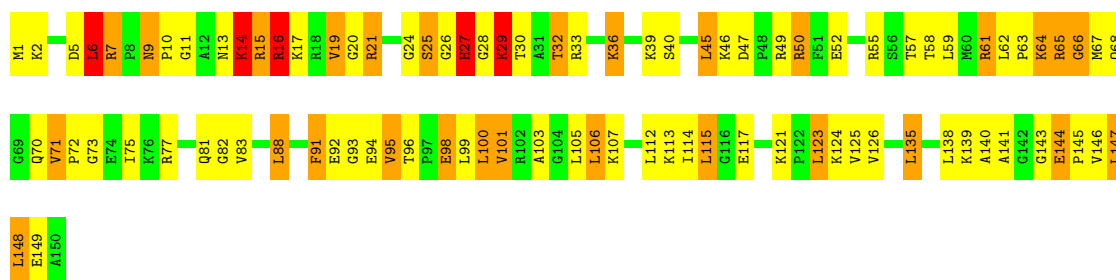
• Molecule 10: 50S ribosomal protein L14

Chain O:



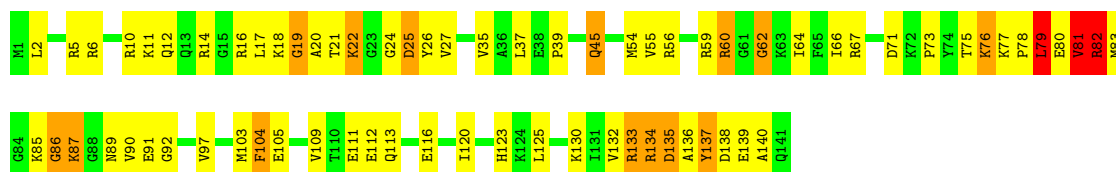
• Molecule 11: 50S ribosomal protein L15

Chain P:



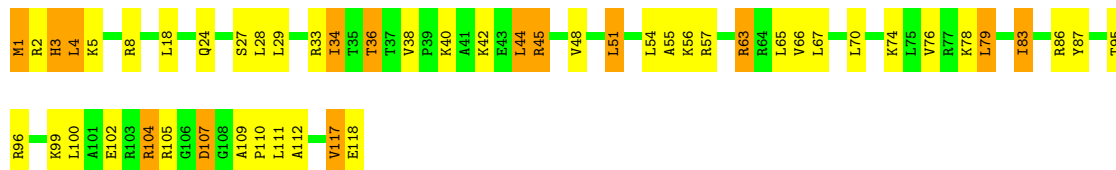
- Molecule 12: 50S ribosomal protein L16

Chain Q:



- Molecule 13: 50S ribosomal protein L17

Chain R:



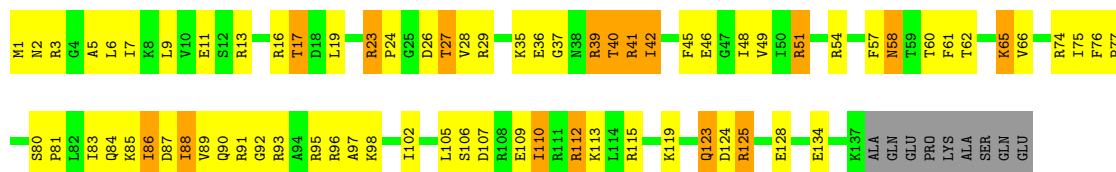
- Molecule 14: 50S ribosomal protein L18

Chain S:



- Molecule 15: 50S ribosomal protein L19

Chain T:



- Molecule 16: 50S ribosomal protein L20

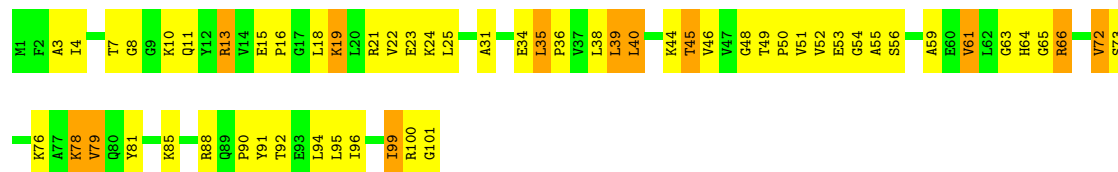
Chain U:





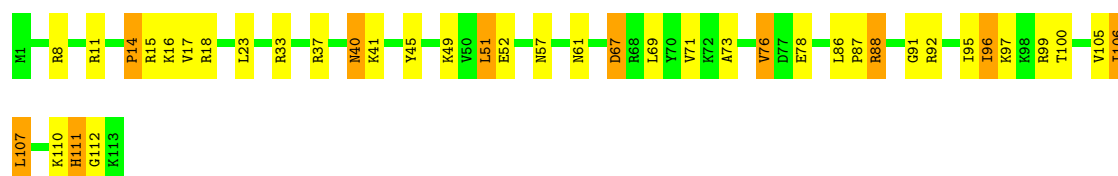
• Molecule 17: 50S ribosomal protein L21

Chain V:



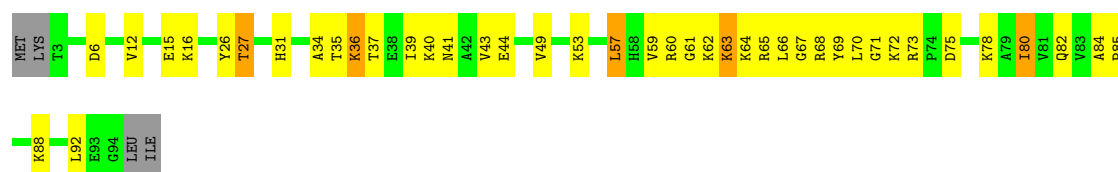
• Molecule 18: 50S ribosomal protein L22

Chain W:



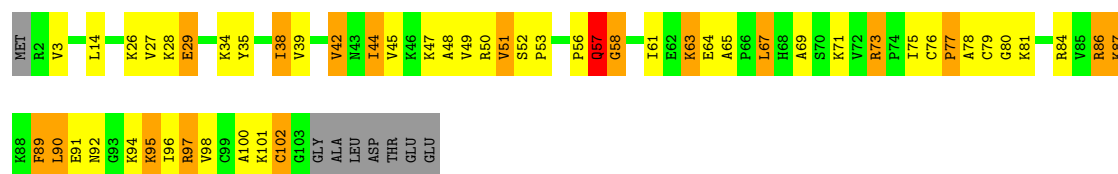
• Molecule 19: 50S ribosomal protein L23

Chain X:



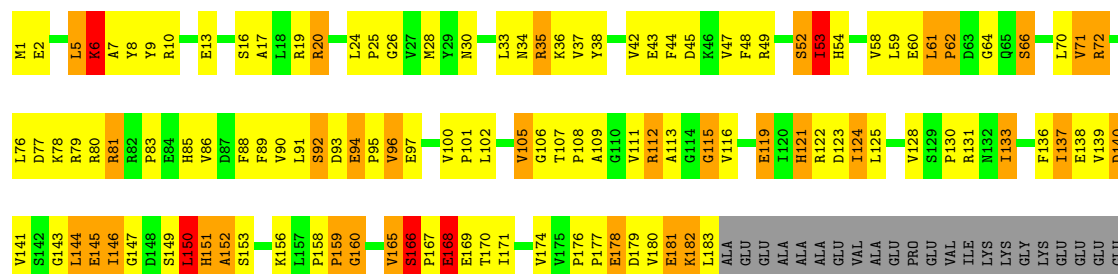
• Molecule 20: 50S ribosomal protein L24

Chain Y:



• Molecule 21: 50S ribosomal protein L25

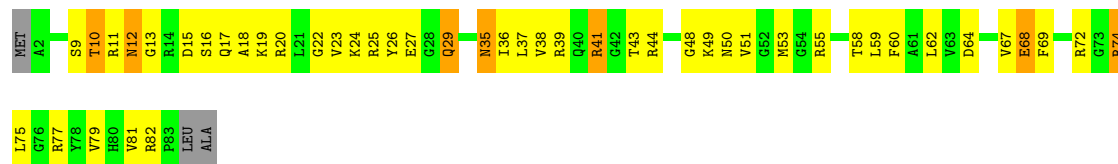
Chain Z:



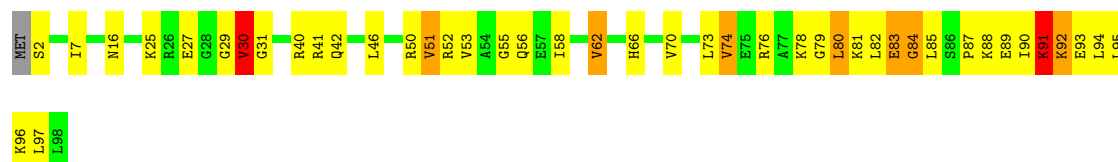


GLU

- Molecule 22: 50S ribosomal protein L27

Chain 0: 

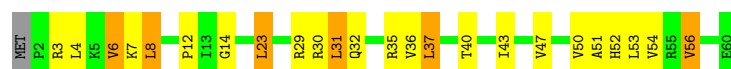
- Molecule 23: 50S ribosomal protein L28

Chain 1: 

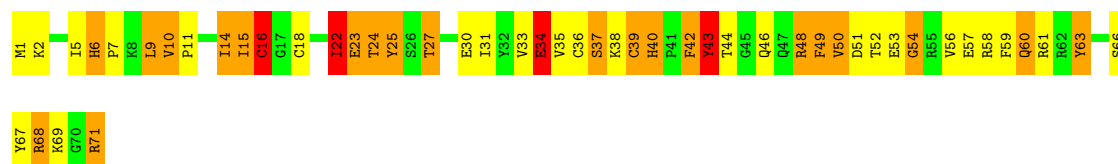
- Molecule 24: 50S ribosomal protein L29

Chain 2: 

- Molecule 25: 50S ribosomal protein L30

Chain 3: 

- Molecule 26: 50S ribosomal protein L31

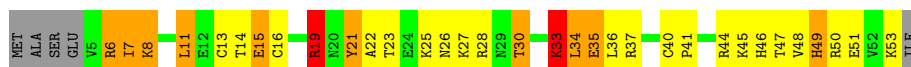
Chain 4: 

- Molecule 27: 50S ribosomal protein L32

Chain 5: 

- Molecule 28: 50S ribosomal protein L33

Chain 6: 



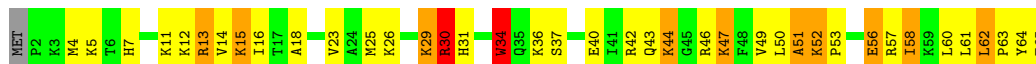
- Molecule 29: 50S ribosomal protein L34

Chain 7:



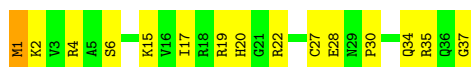
- Molecule 30: 50S ribosomal protein L35

Chain 8:



- Molecule 31: 50S ribosomal protein L36

Chain 9:



- Molecule 32: tRNA acceptor end mimic

Chain a:

There are no outlier residues recorded for this chain.

## 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	92287	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	61.0	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	0.80	27/69543 (0.0%)	1.43	823/108563 (0.8%)
2	B	0.63	0/2878	1.28	17/4490 (0.4%)
3	D	0.58	0/2165	0.78	1/2919 (0.0%)
4	E	0.46	0/1601	0.75	2/2160 (0.1%)
5	F	0.48	0/1620	0.71	1/2194 (0.0%)
6	G	0.40	0/1499	0.60	0/2016
7	H	0.45	0/1332	0.73	0/1802
8	I	0.55	0/1151	0.80	0/1558
9	N	0.43	0/1131	0.64	0/1525
10	O	0.50	0/943	0.65	0/1269
11	P	0.49	0/1162	0.90	2/1544 (0.1%)
12	Q	0.57	0/1143	0.80	1/1527 (0.1%)
13	R	0.44	0/982	0.73	0/1312
14	S	0.40	0/892	0.75	1/1187 (0.1%)
15	T	0.44	0/1155	0.67	0/1542
16	U	0.50	0/982	0.68	0/1306
17	V	0.45	0/790	0.73	1/1057 (0.1%)
18	W	0.45	0/911	0.68	0/1220
19	X	0.49	0/739	0.66	0/993
20	Y	0.46	0/798	0.70	0/1064
21	Z	0.56	0/1493	0.79	1/2026 (0.0%)
22	0	0.74	1/657 (0.2%)	0.90	1/874 (0.1%)
23	1	0.46	0/770	0.69	0/1022
24	2	0.52	0/583	0.73	0/771
25	3	0.41	0/474	0.59	0/635
26	4	0.37	0/594	0.68	0/795
27	5	0.43	0/473	0.77	1/639 (0.2%)
28	6	0.37	0/431	0.67	0/575
29	7	0.57	0/438	0.71	0/575
30	8	0.58	0/525	0.82	0/691
31	9	0.32	0/310	0.48	0/407
32	a	0.92	0/40	1.50	0/60

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
All	All	0.72	28/100205 (0.0%)	1.29	852/150318 (0.6%)

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
5	F	0	1
7	H	0	2
14	S	0	1
21	Z	0	2
24	2	0	1
30	8	0	2
All	All	0	10

The worst 5 of 28 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	783	A	N9-C4	-8.12	1.32	1.37
1	A	2542	A	N9-C4	-7.43	1.33	1.37
1	A	1966	A	N9-C4	-7.42	1.33	1.37
1	A	1938	A	N9-C4	-6.62	1.33	1.37
1	A	2082	A	N9-C4	-6.52	1.33	1.37

The worst 5 of 852 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	774	A	C2-N3-C4	-12.25	104.47	110.60
1	A	1332	G	C6-C5-N7	-11.94	123.24	130.40
1	A	450	G	C5-C6-N1	-11.37	105.81	111.50
1	A	783	A	C2-N3-C4	-11.03	105.08	110.60
1	A	783	A	N1-C6-N6	10.53	124.92	118.60

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	E	21	VAL	Peptide
5	F	47	GLY	Peptide

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
7	H	127	GLU	Peptide
7	H	153	LYS	Peptide
14	S	109	GLY	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	62091	0	31301	1326	1
2	B	2573	0	1306	65	1
3	D	2115	0	2195	103	0
4	E	1568	0	1634	68	0
5	F	1585	0	1632	64	0
6	G	1474	0	1535	60	0
7	H	1307	0	1382	62	0
8	I	1136	0	1223	73	0
9	N	1104	0	1180	51	0
10	O	933	0	996	23	0
11	P	1145	0	1228	95	0
12	Q	1122	0	1179	49	0
13	R	968	0	1033	37	0
14	S	882	0	943	43	0
15	T	1141	0	1202	52	0
16	U	964	0	1022	54	0
17	V	779	0	852	43	0
18	W	900	0	964	26	0
19	X	725	0	778	24	0
20	Y	785	0	878	43	0
21	Z	1461	0	1493	104	0
22	0	648	0	672	39	0
23	1	763	0	848	36	0
24	2	581	0	629	23	0
25	3	469	0	518	15	0
26	4	581	0	574	41	0
27	5	459	0	480	31	0
28	6	424	0	450	29	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
29	7	430	0	480	19	0
30	8	517	0	582	43	0
31	9	307	0	338	17	0
32	a	74	0	51	0	0
33	5	1	0	0	0	0
33	A	266	0	0	0	0
33	B	3	0	0	0	0
33	D	2	0	0	0	0
33	P	2	0	0	0	0
33	Q	1	0	0	0	0
33	X	1	0	0	0	0
All	All	92287	0	61578	2425	1

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 2425 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:D:43:ARG:NH1	3:D:44:ASN:OD1	1.86	1.06
1:A:2701:C:H3'	1:A:2702:U:H5''	1.38	1.05
1:A:1138:G:H21	9:N:106:MET:HE3	1.22	1.04
1:A:498:G:N3	20:Y:47:LYS:NZ	2.07	1.01
1:A:571:A:H5'	1:A:2030:A:H62	1.26	1.00

All (1) 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)
1:A:1593:G:O2'	2:B:54:G:OP1[1_655]	2.14	0.06

## 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%)	227 (84%)	34 (13%)	9 (3%)	6	54
4	E	203/206 (98%)	142 (70%)	41 (20%)	20 (10%)	1	18
5	F	200/210 (95%)	167 (84%)	25 (12%)	8 (4%)	5	48
6	G	179/182 (98%)	142 (79%)	25 (14%)	12 (7%)	2	32
7	H	168/180 (93%)	121 (72%)	23 (14%)	24 (14%)	0	9
8	I	144/148 (97%)	100 (69%)	23 (16%)	21 (15%)	0	8
9	N	136/140 (97%)	106 (78%)	16 (12%)	14 (10%)	1	17
10	O	120/122 (98%)	108 (90%)	10 (8%)	2 (2%)	14	69
11	P	148/150 (99%)	108 (73%)	23 (16%)	17 (12%)	1	13
12	Q	139/141 (99%)	98 (70%)	22 (16%)	19 (14%)	0	10
13	R	116/118 (98%)	99 (85%)	11 (10%)	6 (5%)	3	39
14	S	109/112 (97%)	78 (72%)	18 (16%)	13 (12%)	1	13
15	T	135/146 (92%)	108 (80%)	17 (13%)	10 (7%)	2	28
16	U	115/118 (98%)	101 (88%)	10 (9%)	4 (4%)	6	53
17	V	99/101 (98%)	79 (80%)	12 (12%)	8 (8%)	1	25
18	W	111/113 (98%)	100 (90%)	9 (8%)	2 (2%)	13	68
19	X	90/96 (94%)	82 (91%)	6 (7%)	2 (2%)	10	64
20	Y	100/110 (91%)	70 (70%)	18 (18%)	12 (12%)	1	13
21	Z	181/206 (88%)	113 (62%)	46 (25%)	22 (12%)	1	12
22	0	80/85 (94%)	66 (82%)	13 (16%)	1 (1%)	18	74
23	1	95/98 (97%)	72 (76%)	17 (18%)	6 (6%)	2	34
24	2	67/72 (93%)	55 (82%)	6 (9%)	6 (9%)	1	22
25	3	57/60 (95%)	52 (91%)	4 (7%)	1 (2%)	13	68
26	4	69/71 (97%)	35 (51%)	15 (22%)	19 (28%)	0	1
27	5	57/60 (95%)	46 (81%)	9 (16%)	2 (4%)	6	53
28	6	47/54 (87%)	22 (47%)	17 (36%)	8 (17%)	0	5
29	7	47/49 (96%)	43 (92%)	3 (6%)	1 (2%)	11	65
30	8	62/65 (95%)	48 (77%)	10 (16%)	4 (6%)	2	33
31	9	35/37 (95%)	31 (89%)	4 (11%)	0	100	100
All	All	3379/3526 (96%)	2619 (78%)	487 (14%)	273 (8%)	1	25

5 of 273 Ramachandran outliers are listed below:



Mol	Chain	Res	Type
3	D	26	LYS
3	D	28	GLU
3	D	122	ASP
3	D	123	ALA
4	E	2	LYS

### 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%)	181 (85%)	33 (15%)	4	26
4	E	165/166 (99%)	137 (83%)	28 (17%)	3	21
5	F	161/166 (97%)	137 (85%)	24 (15%)	4	28
6	G	155/156 (99%)	133 (86%)	22 (14%)	5	30
7	H	142/148 (96%)	117 (82%)	25 (18%)	3	18
8	I	122/124 (98%)	85 (70%)	37 (30%)	0	4
9	N	117/119 (98%)	96 (82%)	21 (18%)	2	17
10	O	100/100 (100%)	88 (88%)	12 (12%)	7	38
11	P	116/116 (100%)	82 (71%)	34 (29%)	0	4
12	Q	111/111 (100%)	92 (83%)	19 (17%)	3	20
13	R	101/101 (100%)	81 (80%)	20 (20%)	2	12
14	S	87/88 (99%)	68 (78%)	19 (22%)	1	9
15	T	120/127 (94%)	98 (82%)	22 (18%)	2	16
16	U	93/94 (99%)	77 (83%)	16 (17%)	3	20
17	V	82/82 (100%)	67 (82%)	15 (18%)	2	16
18	W	92/92 (100%)	76 (83%)	16 (17%)	3	19
19	X	74/78 (95%)	60 (81%)	14 (19%)	2	14
20	Y	85/91 (93%)	64 (75%)	21 (25%)	1	7
21	Z	162/179 (90%)	121 (75%)	41 (25%)	1	6
22	0	65/67 (97%)	53 (82%)	12 (18%)	2	15
23	1	82/83 (99%)	70 (85%)	12 (15%)	5	28

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
24	2	64/67 (96%)	47 (73%)	17 (27%)	1	6
25	3	51/52 (98%)	43 (84%)	8 (16%)	4	25
26	4	63/63 (100%)	43 (68%)	20 (32%)	0	3
27	5	51/52 (98%)	37 (72%)	14 (28%)	0	5
28	6	48/52 (92%)	38 (79%)	10 (21%)	2	10
29	7	42/42 (100%)	35 (83%)	7 (17%)	3	22
30	8	54/55 (98%)	41 (76%)	13 (24%)	1	7
31	9	34/34 (100%)	32 (94%)	2 (6%)	28	75
All	All	2853/2923 (98%)	2299 (81%)	554 (19%)	2	13

5 of 554 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
12	Q	103	MET
15	T	110	ILE
27	5	29	THR
13	R	28	LEU
14	S	27	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
15	T	58	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	2880/2916 (98%)	734 (25%)	57 (1%)
2	B	119/122 (97%)	32 (26%)	1 (0%)
32	a	1/3 (33%)	0	0
All	All	3000/3041 (98%)	766 (25%)	58 (1%)

5 of 766 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	9	U
1	A	13	A

*Continued on next page...*

*Continued from previous page...*

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

5 of 58 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	1022	G
1	A	1130	U
1	A	2712	U
1	A	1026	U
1	A	1078	U

## 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	1.48	4 (10%)	54,57,60	2.17	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 (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
32	a	76	PPU	P-OP1	4.82	1.52	1.46
32	a	76	PPU	C4-N9	-4.06	1.31	1.37

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
32	a	76	PPU	C5-C4	2.72	1.46	1.40
32	a	76	PPU	C2'-C3'	-2.41	1.49	1.53

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	-6.67	119.47	125.98
32	a	76	PPU	N3-C4-N9	5.13	134.20	125.39
32	a	76	PPU	C2-N1-C6	4.79	121.90	111.52
32	a	76	PPU	N1-C6-N6	4.54	121.83	117.04
32	a	76	PPU	N3-C2-N1	-4.32	125.09	128.89

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 276 ligands modelled in this entry, 276 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.