



# wwPDB X-ray Structure Validation Summary Report

Sep 24, 2014 – 11:20 PM EDT

PDB ID : 4TOX  
Title : Crystal structure of the E. coli ribosome bound to flopristin. This file contains the 50S subunit of the second 70S ribosome with flopristin bound.  
Authors : Noeske, J.; Huang, J.; Olivier, N.B.; Giacobbe, R.A.; Zambrowski, M.; Cate, J.H.D.  
Deposited on : 2014-06-06  
Resolution : 2.90 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>

---

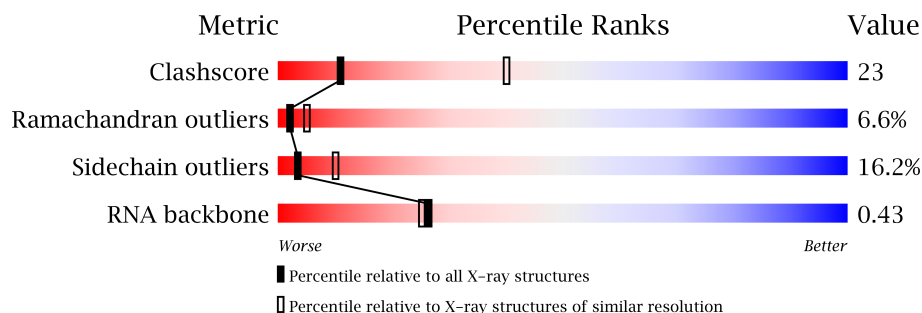
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.90 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	1326 (2.90-2.90)
Ramachandran outliers	78287	1290 (2.90-2.90)
Sidechain outliers	78261	1292 (2.90-2.90)
RNA backbone	1838	1055 (3.40-2.40)

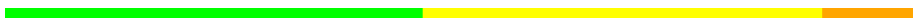
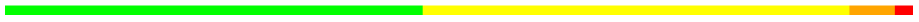


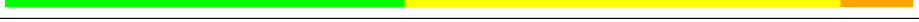



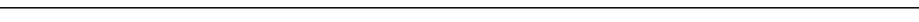

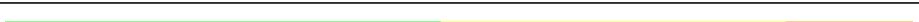



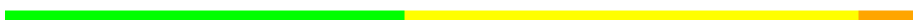
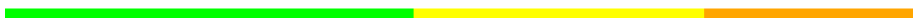

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	2903	
2	B	119	
3	C	271	
4	D	209	
5	E	201	
6	F	177	
7	G	176	
8	H	149	
9	I	141	
10	J	142	
11	K	122	
12	L	143	
13	M	136	
14	N	120	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
15	O	116	
16	P	114	
17	Q	117	
18	R	103	
19	S	110	
20	T	93	
21	U	102	
22	V	94	
23	W	76	
24	X	77	
25	Y	63	
26	Z	58	
27	0	56	
28	1	50	
29	2	46	
30	3	64	
31	4	38	

## 2 Entry composition

There are 35 unique types of molecules in this entry. The entry contains 91704 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	2897	Total	C	N	O	P	0	0	0
			62195	27745	11446	20107	2897			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	118	Total	C	N	O	P	0	0	0
			2529	1126	464	821	118			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	271	Total	C	N	O	S	0	0	0
			2083	1288	423	365	7			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	209	Total	C	N	O	S	0	0	0
			1565	979	288	294	4			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	201	Total	C	N	O	S	0	0	0
			1552	974	283	290	5			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	177	Total	C	N	O	S	0	0	0
			1411	899	249	257	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	176	Total	C	N	O	S	0	0	0
			1323	832	243	246	2			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	149	Total	C	N	O	S	0	0	0
			1110	699	197	213	1			

- Molecule 9 is a protein called 50S ribosomal protein L11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	141	Total	C	N	O	S	0	0	0
			1032	651	179	196	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	142	Total	C	N	O	S	0	0	0
			1129	714	212	199	4			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	122	Total	C	N	O	S	0	0	0
			939	587	180	166	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	143	Total	C	N	O	S	0	0	0
			1045	649	206	189	1			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	136	Total	C	N	O	S	0	0	0
			1074	686	205	177	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	N	120	Total	C	N	O	S	0	0	0
			961	593	196	167	5			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	O	116	Total	C	N	O		0	0	0
			892	552	178	162				

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
16	P	114	Total	C	N	O	S	0	0	0
			917	574	179	163	1			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
17	Q	117	Total	C	N	O		0	0	0
			947	604	192	151				

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
18	R	103	Total	C	N	O	S	0	0	0
			816	516	153	145	2			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
19	S	110	Total	C	N	O	S	0	0	0
			857	532	166	156	3			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
20	T	93	Total	C	N	O	S	0	0	0
			739	466	139	132	2			

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
21	U	102	Total	C	N	O			
			780	492	146	142	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
22	V	94	Total	C	N	O	S			
			753	479	137	134	3	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
23	W	75	Total	C	N	O	S			
			569	353	113	102	1	0	0	0

- Molecule 24 is a protein called 50S ribosomal protein L28.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
24	X	77	Total	C	N	O	S			
			625	388	129	106	2	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
25	Y	63	Total	C	N	O	S			
			509	313	99	95	2	0	0	0

- Molecule 26 is a protein called 50S ribosomal protein L30.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
26	Z	58	Total	C	N	O	S			
			449	281	87	79	2	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
27	0	56	Total	C	N	O	S			
			444	269	94	80	1	0	0	0

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
28	1	50	Total	C	N	O	0	0	0
			410	263	75	72			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
29	2	46	Total	C	N	O	S	0	0	0
			377	228	90	57	2			

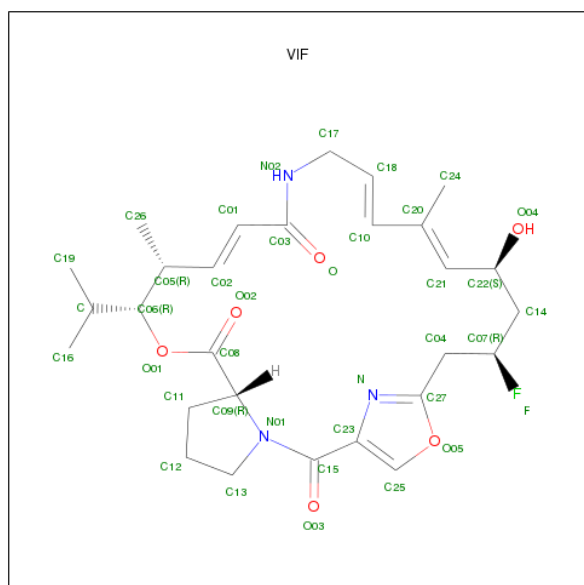
- Molecule 30 is a protein called 50S ribosomal protein L35.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
30	3	64	Total	C	N	O	S	0	0	0
			504	323	105	74	2			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
31	4	38	Total	C	N	O	S	0	0	0
			302	185	65	48	4			

- Molecule 32 is Flopristin (three-letter code: VIF) (formula:  $C_{28}H_{38}FN_3O_6$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
32	A	1	Total	C	F	N	O	0	0
			38	28	1	3	6		



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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
33	B	3	Total 3	Mg 3	0	0
33	A	166	Total 166	Mg 166	0	0
33	Q	1	Total 1	Mg 1	0	0
33	2	1	Total 1	Mg 1	0	0

- Molecule 34 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
34	4	1	Total 1	Zn 1	0	0

- Molecule 35 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
35	A	613	Total 613	O 613	0	0
35	B	13	Total 13	O 13	0	0
35	C	9	Total 9	O 9	0	0
35	D	4	Total 4	O 4	0	0
35	E	2	Total 2	O 2	0	0
35	J	1	Total 1	O 1	0	0
35	L	3	Total 3	O 3	0	0
35	N	1	Total 1	O 1	0	0
35	T	2	Total 2	O 2	0	0
35	V	1	Total 1	O 1	0	0
35	0	1	Total 1	O 1	0	0
35	2	3	Total 3	O 3	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
35	3	2	Total	O	0	0
			2	2		
35	4	1	Total	O	0	0
			1	1		

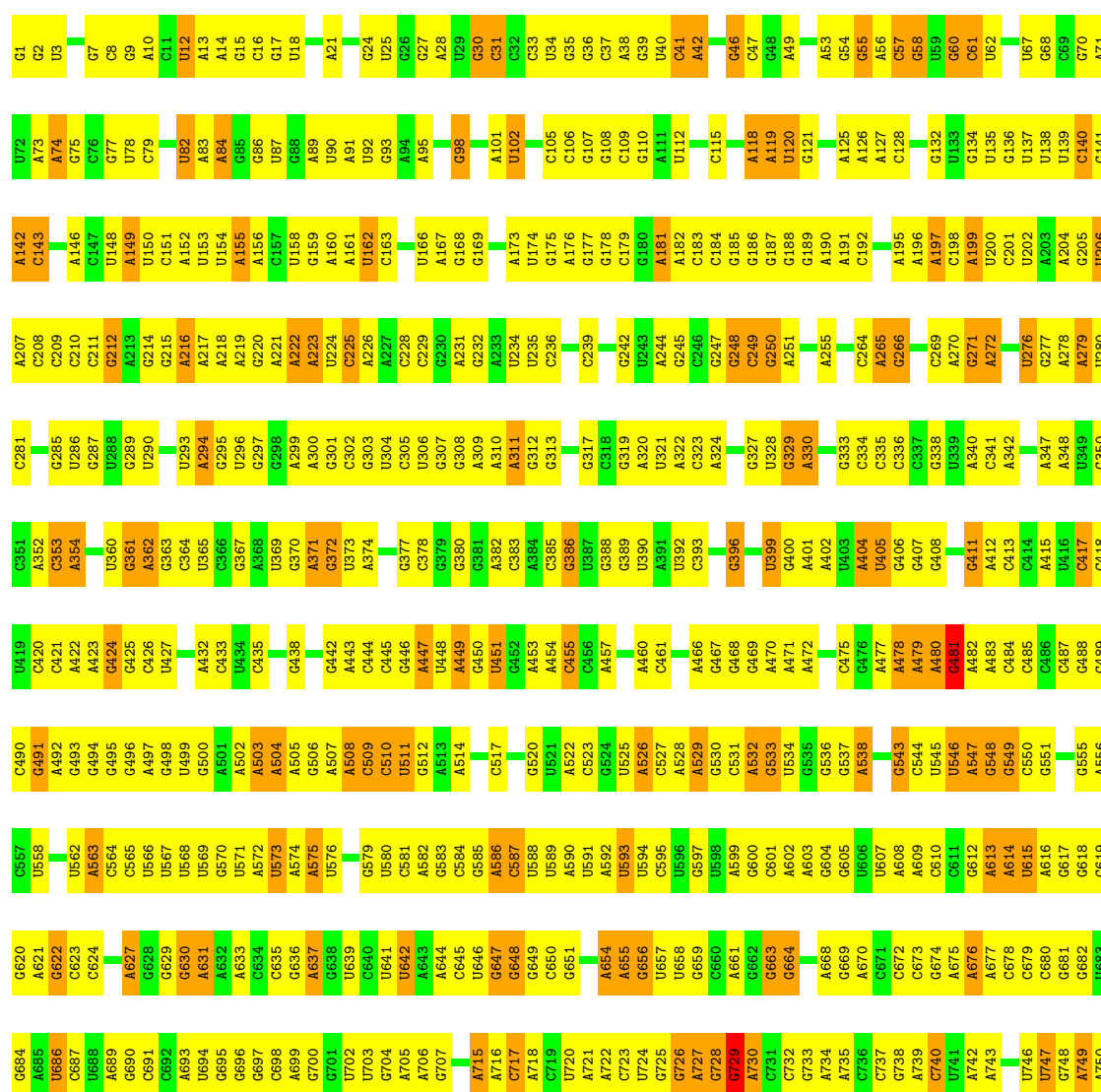
### 3 Residue-property plots

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.

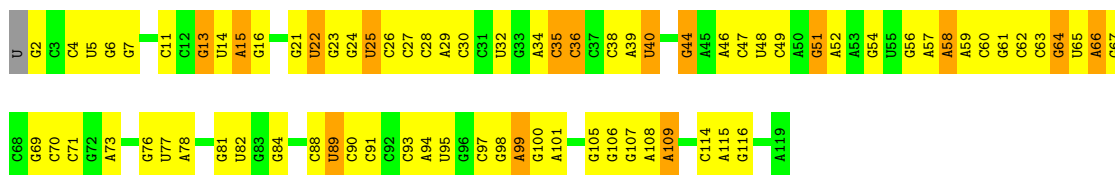
#### • Molecule 1: 23S rRNA

Chain A:



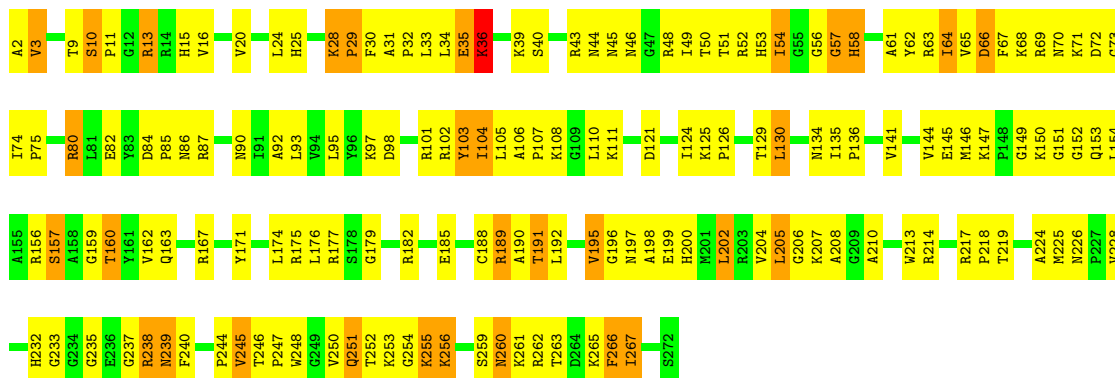
A1810	G1743	C1670	A1598	A1528	A1453	C1314	G1250	C1172	G1099	G1031	G966	A892	G818	A751
G1811	A1744	U1671	U1599	G1529	C1454	C1315	C1251	U1173	C1100	A1032	G969	C893	A819	A752
U1812	A1745	A1672	C1600	G1530	G1455	U1316	G1252	U1174	U1101	U1033	U969	U894	A820	A753
G1813	G1746	C1601	G1601	G1531	G1456	G1317	A1253	A1175	C1102	G1034	C1102	U895	A821	U754
G1814	U1747	U1674	U1602	U1532	U1457	U1318	A1254	U1176	A1103	U1035	G971	A896	G822	U755
A1815	C1675	C1533	C1533	C1533	U1458	C1319	U1255	G1177	C1104	U1035	A972	C897	U756	A756
C1816	A1603	C1534	U1534	U1534	C1462	C1320	G1256	C1178	U1105	G1041	A973	C898	G757	C758
G1817	C1604	A1535	C1605	A1535	C1463	C1321	C1257	U1179	G1106	U1045	G974	G902	U827	U761
U1818	U1677	C1536	C1606	G1536	C1464	U1322	U1258	U1180	G1107	A1046	A975	C903	U828	
A1819	A1678	G1537	C1607	G1537	G1464	G1324	U1259	U1181	U1108	A1047	G976	G904	A829	
U1820	G1754	U1538	A1608	G1538	G1465	U1325	C1259	U1182	C1109	G1047	G977	G905	G830	
A1821	A1655	G1681	A1609	U1539	C1465	U1326	U1260	U1183	G1110	A1050	A980	A905	G831	A764
G1822	G1756	C1682	A1610	G1540	U1468	U1327	A1262	U1184	A1111	U1050	A981	U906	U832	C765
G1823	A1757	U1683	G1613	C1541	A1469	A1328	A1264	G1186	G1112	G1051	A982	G907	U833	C766
	U1758		G1613	C1541	A1470	U1329	A1265	U1187	U1113	C1052	A983	C908	A833	U766
G1826	U1759	U1688	A1616	A1545	U1476	G1330	G1266	U1188	G1115	C1053	A984	A910	G836	G768
U1827	C1760	A1689	C1617	A1545	U1477	G1331	U1267	G1192	U1057	G1056	C985	A911	U769	U776
G1828	C1761	A1690	A1618	C1546	A1478	G1332	A1269	G1193	A1057	U1058	C986	C912	C837	G771
A1829	C1762	C1691	G1619	G1547	U1479	G1333	A1270	A1194	U1119	U1059	C987	U913	A844	C772
G1830	G1764	U1692	G1620	A1548	G1479	C1335	G1271	G1195	G1120	U1060	C988	G914	A845	G775
G1831	U1765	C1694	U1621	A1549	G1480	A1336	A1272	C1196	C1121	U1061	G989	C915	U846	G776
C1832	G1766	G1695	G1622	C1550	U1481	G1337	U1273	G1197	G1122	U1062	A990	C922	U847	G777
C1833	G1767				U1482	G1338	A1274	U1198	G1125	G1063	G993	A927	C848	G778
U1834					G1483	G1339	A1275	U1199	A1126	C1064	C994	A928	U850	U779
G1835	A1773	A1698	C1625	A1553	U1484	U1340	A1276	C1200	G1127	U1065	C995	U931	C851	G780
	C1774	G1702	A1626	U1554	U1491	G1341	G1277		U1128	U1066	A996	U932		A781
C1838	U1775	G1703	G1555	C1556	U1492	A1342	G1280	A1205	G1128	A1067	C997	U933	G856	A782
G1839	G1776	G1704	C1556	C1556	G1491	U1344	G1281	C1207	A1129	G1068	C998	U934	G857	A783
C1843	U1777	A1705	A1637	U1562	G1492	U1344	U1285	C1208	U1132	A1069	U999	A933	G858	G784
G1846	U1779	C1706	C1638	C1563	G1493	C1345	A1286	U1209	A1133	A1070	A1000	U934	G859	G785
A1847	U1780	G1707	C1639	U1563	C1494	C1350	A1287	G1210	A1134	G1071	A1001	C935	G864	C786
U1848	U1781	C1708	A1640	C1564	G1495	C1351	A1288	C1211	C1135	A1072	G1002	A936	C865	C787
	U1782	U1709	A1641	A1566	A1496	U1352	C1289	G1212	G1136	C1073	G1003	C937	A866	A788
G1854	A1783	G1710		U1567	U1497	A1353	C1290	G1221	G1137	G1074	U1004	G938	A867	A789
U1855	A1785	A1713	C1644	G1568	C1498	A1354	C1291	U1222	G1138	C1075	C1005	G939	C867	U790
U1856	U1786	G1714	C1646	A1569	C1499	G1355	C1292	U1223	G1139	C1076	C1006	G940	U868	C791
G1857		G1715	U1647	U1570	G1500	G1356	G1293	G1223	U1141	U1078	A1009	A941	G869	C792
A1858	U1789	U1716	U1648	A1501	G1501	C1357	U1294	U1224	A1142	C1079	A1010	A943	U871	A793
U1859			G1649	A1502	A1503	G1358	C1295	G1225	A1143	A1080	G1011	C944	C873	C795
	G1721		A1650	U1576	A1504	A1359	C1296	A1226	A1144	U1081	C1012	C945	G874	G796
G1862	G1792		G1651	C1577	A1504	G1360	C1297	G1229	U1082	U1083	C1013	C946	C875	G797
C1863	C1793	C1726	A1652	U1578	A1509	C1361	G1298	A1230	A1147	U1084	U1015	C947	G876	G799
U1864	A1794	C1727	G1653	C1581	G1510	C1362	C1299	U1231	U1148	A1085	G1016	C948	A878	
U1865	C1795	C1728	A1654	C1582	U1510	G1363	G1300	G1232	G1149	A1086	U1019	C950	G879	A802
A1866	U1796	U1729	A1655	U1583	G1514	C1364	G1301	G1233	C1152	A1087	G950	C951	G880	G805
G1867	G1797	C1730	C1656	A1584	A1515	A1365	A1302	U1234	C1153	A1088	A1020	C952	G881	C806
C1868	U1798	G1731		C1585	G1516	A1367	G1303	G1235	G1154	A1089	A1021	G953	G882	
G1869	G1799	C1732	G1659	A1586	G1519	G1368	C1304	G1236	A1156	A1090	G1022	G954	G883	
C1870	C1800	G1733	G1660	C1587	U1520	G1369	C1305	A1237	C1091	U1023	G1023	G956	U884	U810
A1871	A1801	G1734	G1661	U1587	G1521	C1370	C1306	G1238	C1092	G1024	G1024	C957	C885	U811
A1872	A1802	A1735	U1662	U1588	G1522	C1371	C1307	G1239	G1162	U1093	G1025	C957	U884	U812
G1873	C1803	U1736	G1663	U1589	A1523	U1372	A1308	U1240	G1167	U1094	G1026	C957	U884	U813
C1874	C1804	G1737	A1664	A1593	U1523	U1373	G1309	A1241	G1167	U1095	A1027	C959	U884	C814
G1875	A1805	G1738	A1665	U1594	G1524	G1374	G1310	A1247	C1168	A1096	A1028	A960	U884	C815
A1876	C1806	A1739	G1666	U1594	G1525	U1375	G1311	U1247	G1168	A1097	A1029	C961	U884	C816
	G1807	C1740	C1667	C1595	C1526	U1376	G1312	G1248	G1169	U1097	C1030			
U1880	A1808	C1741	A1668	C1596	C1527	G1377	U1313	U1249	G1171	A1098				
C1881	A1809	U1742	A1669	A1597	G1527	G1377	U1313	U1249	G1171	A1098				





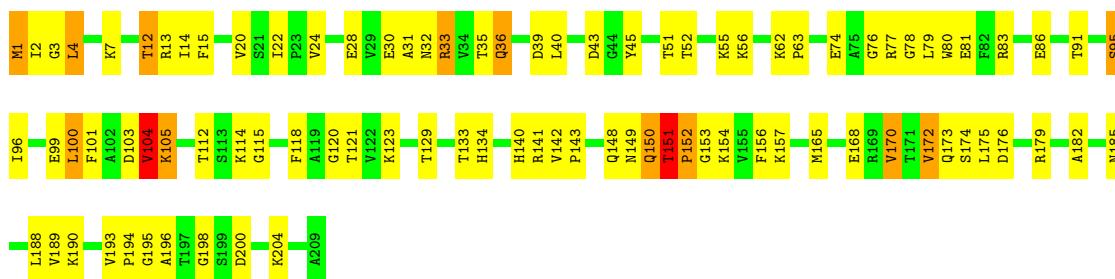
- Molecule 3: 50S ribosomal protein L2

Chain C:



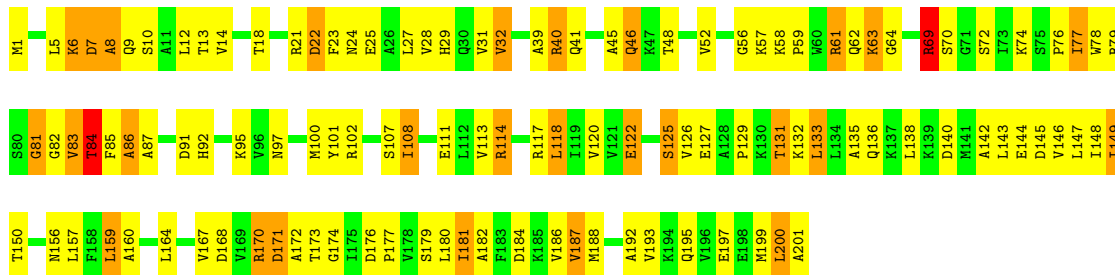
- Molecule 4: 50S ribosomal protein L3

Chain D:



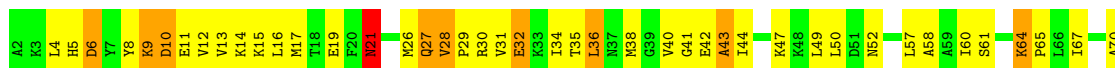
- Molecule 5: 50S ribosomal protein L4

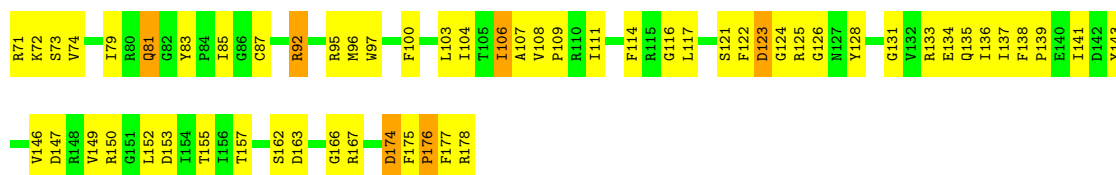
Chain E:



- Molecule 6: 50S ribosomal protein L5

Chain F:





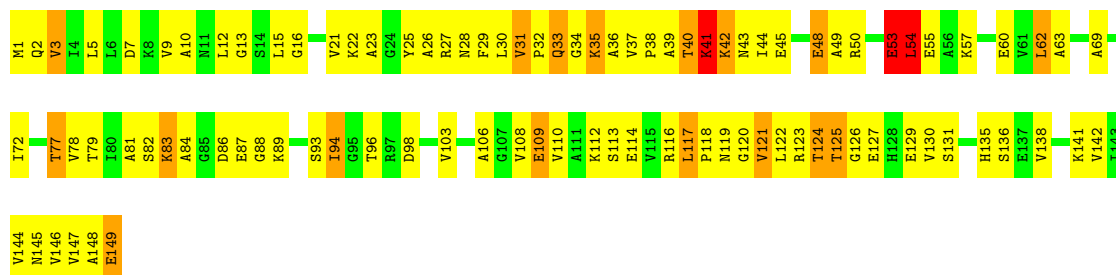
• Molecule 7: 50S ribosomal protein L6

Chain G:



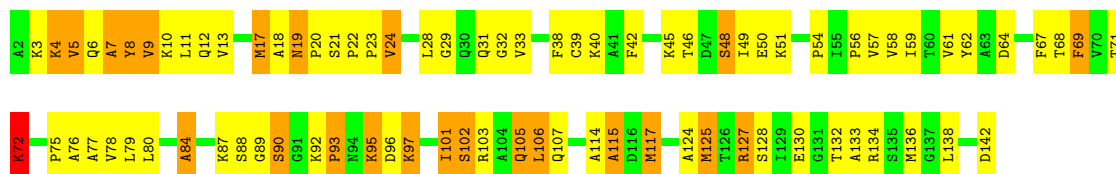
• Molecule 8: 50S ribosomal protein L9

Chain H:



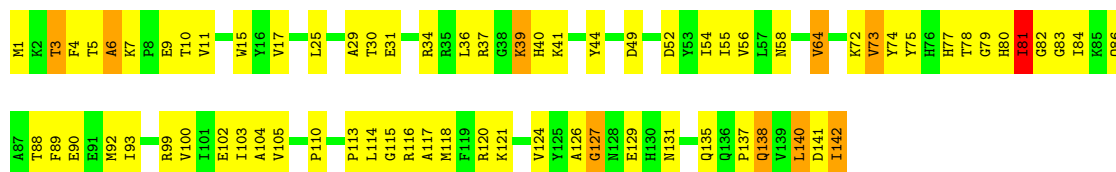
• Molecule 9: 50S ribosomal protein L11

Chain I:



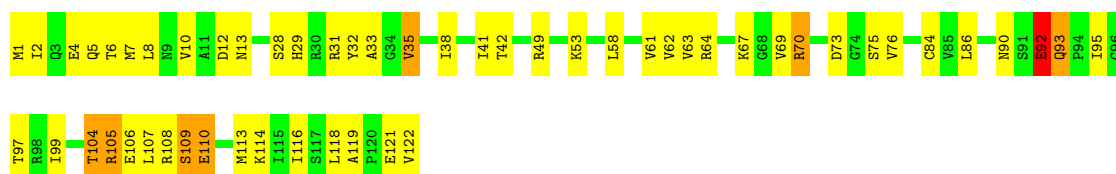
• Molecule 10: 50S ribosomal protein L13

Chain J:



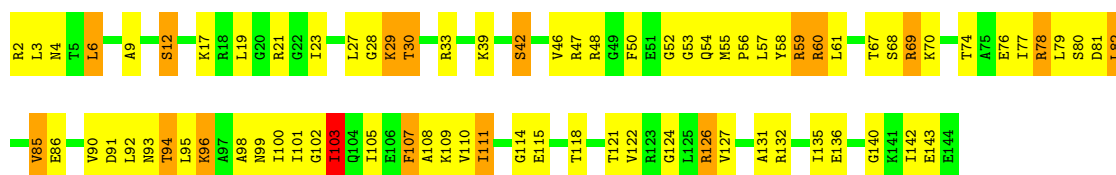
• Molecule 11: 50S ribosomal protein L14

Chain K:



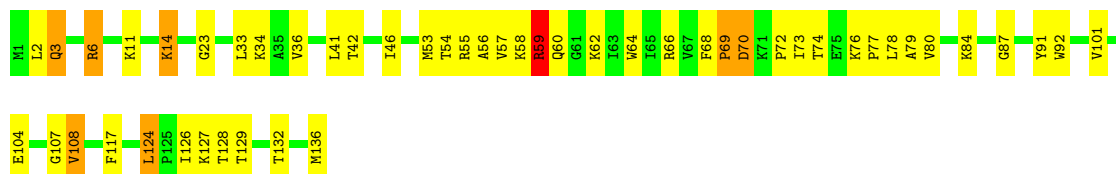
• Molecule 12: 50S ribosomal protein L15

Chain L:



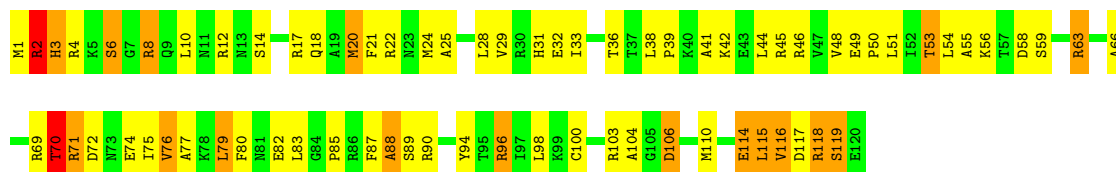
• Molecule 13: 50S ribosomal protein L16

Chain M:



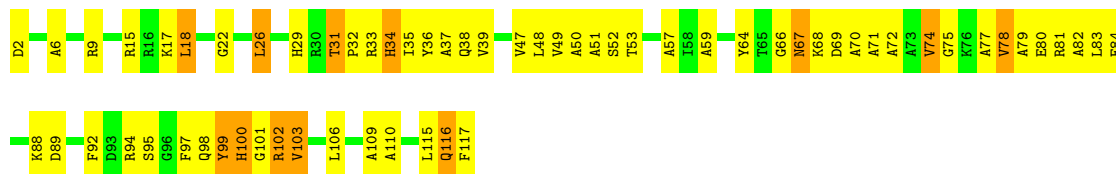
• Molecule 14: 50S ribosomal protein L17

Chain N:



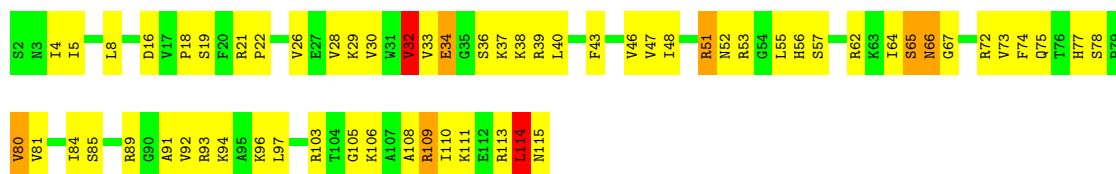
• Molecule 15: 50S ribosomal protein L18

Chain O:



• Molecule 16: 50S ribosomal protein L19

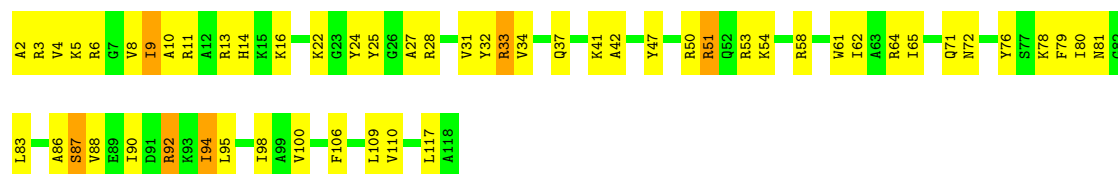
Chain P:





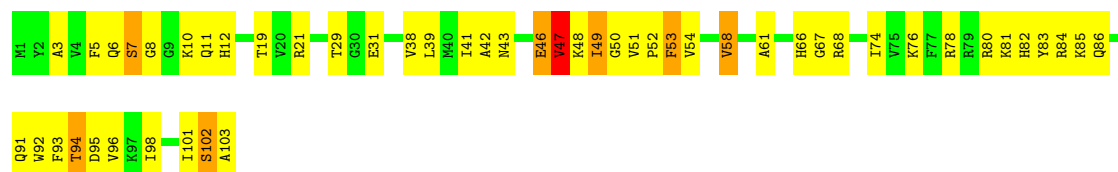
- Molecule 17: 50S ribosomal protein L20

Chain Q:



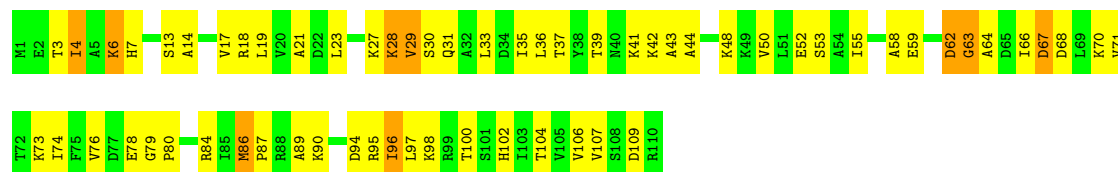
- Molecule 18: 50S ribosomal protein L21

Chain R:



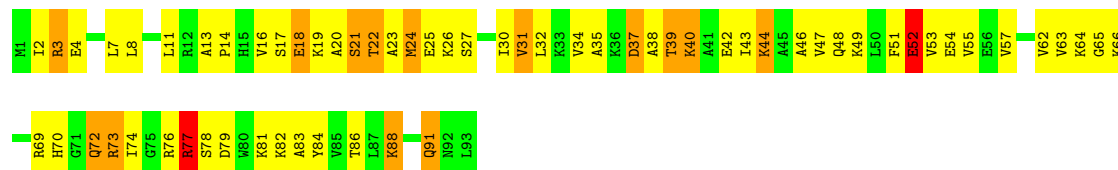
- Molecule 19: 50S ribosomal protein L22

Chain S:



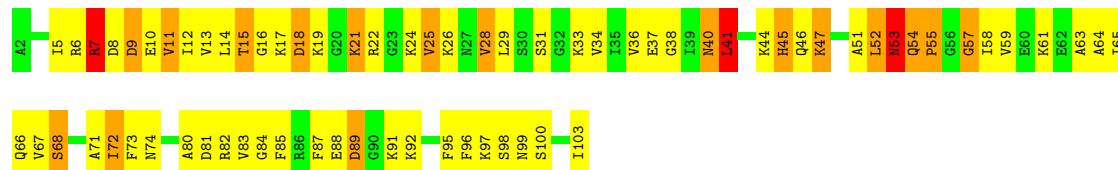
- Molecule 20: 50S ribosomal protein L23

Chain T:



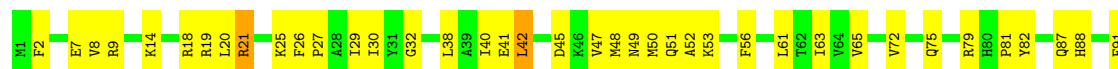
- Molecule 21: 50S ribosomal protein L24

Chain U:



- Molecule 22: 50S ribosomal protein L25

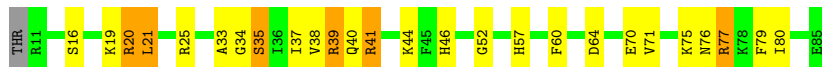
Chain V:





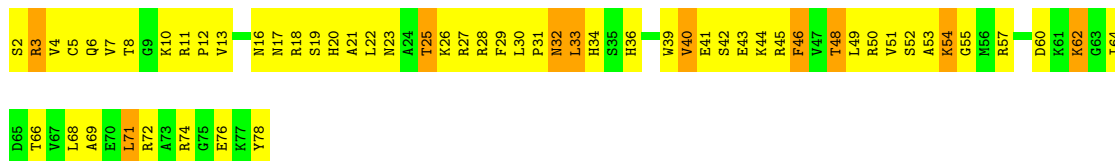
- Molecule 23: 50S ribosomal protein L27

Chain W:



- Molecule 24: 50S ribosomal protein L28

Chain X:



- Molecule 25: 50S ribosomal protein L29

Chain Y:



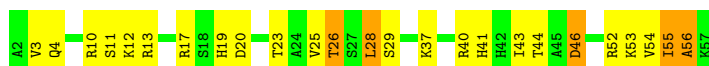
- Molecule 26: 50S ribosomal protein L30

Chain Z:



- Molecule 27: 50S ribosomal protein L32

Chain 0:



- Molecule 28: 50S ribosomal protein L33

Chain 1:



- Molecule 29: 50S ribosomal protein L34

Chain 2:



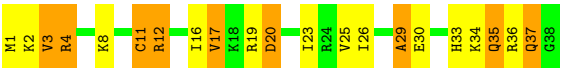
- Molecule 30: 50S ribosomal protein L35

Chain 3: 



- Molecule 31: 50S ribosomal protein L36

Chain 4: 



## 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, $\alpha$ , $\beta$ , $\gamma$	211.79Å 433.06Å 623.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	69.33 – 2.90	Depositor
% Data completeness (in resolution range)	87.4 (69.33-2.90)	Depositor
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, $R_{free}$	0.235 , 0.279	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	91704	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	84.0	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: ZN, VIF, MG

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	0.40	0/69659	0.90	13/108672 (0.0%)
2	B	0.35	0/2828	0.85	0/4410
3	C	0.34	0/2122	0.60	0/2852
4	D	0.33	0/1586	0.57	0/2134
5	E	0.35	0/1571	0.59	0/2113
6	F	0.32	0/1435	0.52	0/1926
7	G	0.32	0/1343	0.53	0/1816
8	H	0.35	0/1121	0.56	0/1515
9	I	0.37	0/1046	0.61	0/1410
10	J	0.32	0/1152	0.59	0/1551
11	K	0.34	0/948	0.56	0/1268
12	L	0.34	0/1054	0.61	0/1403
13	M	0.31	0/1093	0.54	0/1460
14	N	0.35	0/974	0.58	0/1301
15	O	0.30	0/902	0.51	0/1209
16	P	0.35	0/929	0.58	0/1242
17	Q	0.33	0/960	0.53	0/1278
18	R	0.34	0/829	0.58	0/1107
19	S	0.34	0/864	0.59	0/1156
20	T	0.35	0/745	0.60	0/994
21	U	0.37	0/788	0.59	0/1051
22	V	0.30	0/766	0.50	0/1025
23	W	0.31	0/576	0.49	0/762
24	X	0.34	0/635	0.60	0/848
25	Y	0.34	0/510	0.58	0/677
26	Z	0.31	0/453	0.55	0/605
27	0	0.35	0/450	0.61	0/599
28	1	0.34	0/417	0.56	0/554
29	2	0.36	0/380	0.59	0/498
30	3	0.31	0/513	0.54	0/676
31	4	0.46	0/303	0.64	0/397
All	All	0.38	0/98952	0.83	13/148509 (0.0%)

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	D	0	1

There are no bond length outliers.

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	783	A	N1-C6-N6	6.63	122.58	118.60
1	A	729	G	O4'-C1'-N9	5.87	112.89	108.20
1	A	2447	G	C4-N9-C1'	-5.83	118.92	126.50
1	A	481	G	O4'-C1'-N9	5.74	112.79	108.20
1	A	2425	A	P-O3'-C3'	5.58	126.39	119.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	D	151	THR	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	62195	0	31280	2083	0
2	B	2529	0	1281	64	0
3	C	2083	0	2154	140	0
4	D	1565	0	1616	85	0
5	E	1552	0	1619	108	0
6	F	1411	0	1444	53	0
7	G	1323	0	1371	62	0
8	H	1110	0	1148	93	0
9	I	1032	0	1085	76	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	J	1129	0	1162	55	0
11	K	939	0	1012	35	0
12	L	1045	0	1117	71	0
13	M	1074	0	1157	31	0
14	N	961	0	1000	69	0
15	O	892	0	923	48	0
16	P	917	0	962	49	0
17	Q	947	0	1019	56	0
18	R	816	0	839	52	0
19	S	857	0	922	43	0
20	T	739	0	807	54	0
21	U	780	0	831	71	0
22	V	753	0	780	24	0
23	W	569	0	581	20	0
24	X	625	0	652	67	0
25	Y	509	0	543	45	0
26	Z	449	0	488	16	0
27	0	444	0	458	26	0
28	1	410	0	440	19	0
29	2	377	0	418	28	0
30	3	504	0	572	31	0
31	4	302	0	342	21	0
32	A	38	0	0	4	0
33	2	1	0	0	0	0
33	A	166	0	0	0	0
33	B	3	0	0	0	0
33	Q	1	0	0	0	0
34	4	1	0	0	0	0
35	0	1	0	0	0	0
35	2	3	0	0	0	0
35	3	2	0	0	0	0
35	4	1	0	0	0	0
35	A	613	0	0	87	0
35	B	13	0	0	1	0
35	C	9	0	0	1	0
35	D	4	0	0	2	0
35	E	2	0	0	0	0
35	J	1	0	0	0	0
35	L	3	0	0	1	0
35	N	1	0	0	0	0
35	T	2	0	0	0	0
35	V	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	91704	0	60023	3361	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 23.

The worst 5 of 3361 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:2711:A:OP2	35:A:3548:HOH:O	1.68	1.11
1:A:192:C:OP1	35:A:3738:HOH:O	1.69	1.08
1:A:1395:A:OP2	35:A:3403:HOH:O	1.72	1.07
4:D:151:THR:O	4:D:153:GLY:N	1.88	1.06
1:A:756:A:N7	35:A:3301:HOH:O	1.90	1.04

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	C	269/271 (99%)	200 (74%)	49 (18%)	20 (7%)	2	4
4	D	207/209 (99%)	165 (80%)	33 (16%)	9 (4%)	4	15
5	E	199/201 (99%)	146 (73%)	40 (20%)	13 (6%)	2	5
6	F	175/177 (99%)	136 (78%)	26 (15%)	13 (7%)	2	4
7	G	174/176 (99%)	129 (74%)	30 (17%)	15 (9%)	1	3
8	H	147/149 (99%)	100 (68%)	32 (22%)	15 (10%)	1	2
9	I	139/141 (99%)	75 (54%)	49 (35%)	15 (11%)	1	2
10	J	140/142 (99%)	123 (88%)	12 (9%)	5 (4%)	5	22
11	K	120/122 (98%)	97 (81%)	16 (13%)	7 (6%)	3	7
12	L	141/143 (99%)	104 (74%)	27 (19%)	10 (7%)	2	4
13	M	134/136 (98%)	111 (83%)	17 (13%)	6 (4%)	4	14

*Continued on next page...*



Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
14	N	118/120 (98%)	91 (77%)	19 (16%)	8 (7%)	2	5
15	O	114/116 (98%)	97 (85%)	11 (10%)	6 (5%)	3	9
16	P	112/114 (98%)	89 (80%)	16 (14%)	7 (6%)	2	6
17	Q	115/117 (98%)	104 (90%)	10 (9%)	1 (1%)	25	66
18	R	101/103 (98%)	80 (79%)	14 (14%)	7 (7%)	2	4
19	S	108/110 (98%)	88 (82%)	13 (12%)	7 (6%)	2	5
20	T	91/93 (98%)	65 (71%)	14 (15%)	12 (13%)	0	1
21	U	100/102 (98%)	70 (70%)	19 (19%)	11 (11%)	1	2
22	V	92/94 (98%)	76 (83%)	14 (15%)	2 (2%)	10	37
23	W	73/76 (96%)	59 (81%)	11 (15%)	3 (4%)	4	17
24	X	75/77 (97%)	57 (76%)	14 (19%)	4 (5%)	3	9
25	Y	61/63 (97%)	42 (69%)	14 (23%)	5 (8%)	1	3
26	Z	56/58 (97%)	53 (95%)	1 (2%)	2 (4%)	5	22
27	0	54/56 (96%)	38 (70%)	12 (22%)	4 (7%)	2	4
28	1	48/50 (96%)	40 (83%)	5 (10%)	3 (6%)	2	6
29	2	44/46 (96%)	36 (82%)	6 (14%)	2 (4%)	4	14
30	3	62/64 (97%)	49 (79%)	10 (16%)	3 (5%)	4	12
31	4	36/38 (95%)	28 (78%)	4 (11%)	4 (11%)	1	2
All	All	3305/3364 (98%)	2548 (77%)	538 (16%)	219 (7%)	2	5

5 of 219 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	10	SER
3	C	29	PRO
3	C	35	GLU
3	C	36	LYS
3	C	58	HIS

### 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	C	216/216 (100%)	185 (86%)	31 (14%)	5	13
4	D	164/164 (100%)	148 (90%)	16 (10%)	12	34
5	E	165/165 (100%)	134 (81%)	31 (19%)	2	7
6	F	148/148 (100%)	120 (81%)	28 (19%)	2	7
7	G	137/137 (100%)	123 (90%)	14 (10%)	11	31
8	H	114/114 (100%)	88 (77%)	26 (23%)	1	3
9	I	109/109 (100%)	86 (79%)	23 (21%)	1	5
10	J	116/116 (100%)	100 (86%)	16 (14%)	5	14
11	K	103/103 (100%)	91 (88%)	12 (12%)	8	22
12	L	102/102 (100%)	78 (76%)	24 (24%)	1	3
13	M	109/109 (100%)	98 (90%)	11 (10%)	11	32
14	N	100/100 (100%)	82 (82%)	18 (18%)	2	7
15	O	86/86 (100%)	70 (81%)	16 (19%)	2	7
16	P	99/99 (100%)	82 (83%)	17 (17%)	3	8
17	Q	89/89 (100%)	74 (83%)	15 (17%)	3	9
18	R	84/84 (100%)	74 (88%)	10 (12%)	8	21
19	S	93/93 (100%)	80 (86%)	13 (14%)	5	14
20	T	80/80 (100%)	67 (84%)	13 (16%)	3	10
21	U	83/83 (100%)	63 (76%)	20 (24%)	1	3
22	V	78/78 (100%)	68 (87%)	10 (13%)	6	18
23	W	56/58 (97%)	51 (91%)	5 (9%)	14	40
24	X	67/67 (100%)	55 (82%)	12 (18%)	2	7
25	Y	55/55 (100%)	43 (78%)	12 (22%)	1	4
26	Z	48/48 (100%)	37 (77%)	11 (23%)	1	3
27	0	47/47 (100%)	41 (87%)	6 (13%)	6	18
28	1	45/45 (100%)	37 (82%)	8 (18%)	2	7
29	2	38/38 (100%)	31 (82%)	7 (18%)	2	7
30	3	51/51 (100%)	45 (88%)	6 (12%)	8	22
31	4	34/34 (100%)	26 (76%)	8 (24%)	1	3
All	All	2716/2718 (100%)	2277 (84%)	439 (16%)	3	10

5 of 439 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
12	L	12	SER
14	N	96	ARG
27	O	25	VAL
12	L	48	ARG
13	M	14	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
12	L	35	HIS
17	Q	37	GLN
25	Y	41	HIS
8	H	128	HIS
28	1	19	HIS

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	2895/2903 (99%)	637 (22%)	29 (1%)
2	B	117/119 (98%)	25 (21%)	0
All	All	3012/3022 (99%)	662 (21%)	29 (0%)

5 of 662 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	10	A
1	A	12	U
1	A	15	G
1	A	30	G
1	A	31	C

5 of 29 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	1738	G
1	A	2146	C
1	A	2585	U
1	A	2109	U
1	A	2162	G

## 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 ⓘ

Of 173 ligands modelled in this entry, 172 are monoatomic - leaving 1 for Mogul analysis.

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	VIF	A	3001	-	40,40,40	2.40	12 (30%)	53,55,55	2.59	21 (39%)

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	VIF	A	3001	-	-	0/46/58/58	0/1/3/3

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
32	A	3001	VIF	C15-N01	6.00	1.44	1.35
32	A	3001	VIF	F-C07	-5.94	1.25	1.40
32	A	3001	VIF	C03-N02	4.41	1.40	1.34
32	A	3001	VIF	C11-C09	-4.24	1.43	1.53
32	A	3001	VIF	O01-C06	-4.11	1.38	1.44

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
32	A	3001	VIF	C02-C01-C03	-9.85	103.77	121.88
32	A	3001	VIF	C17-N02-C03	-5.25	115.49	121.82
32	A	3001	VIF	C06-O01-C08	4.62	125.80	117.98
32	A	3001	VIF	C18-C10-C20	-4.54	119.09	125.94
32	A	3001	VIF	C23-N-C27	-4.52	104.88	107.50

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.