



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 10, 2023 – 09:39 AM EDT

PDB ID : 4KZX
Title : Rabbit 40S ribosomal subunit in complex with eIF1.
Authors : Lomakin, I.B.; Steitz, T.A.
Deposited on : 2013-05-30
Resolution : 7.81 Å(reported)

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

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.35.1
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.35.1

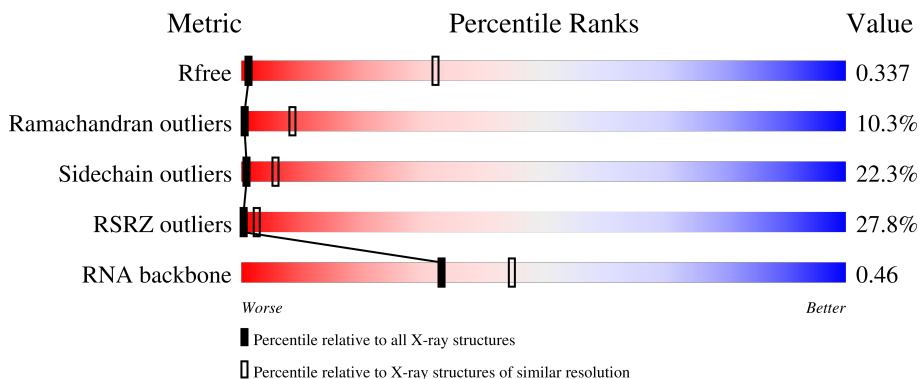
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 7.81 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1005 (11.50-3.90)
Ramachandran outliers	138981	1003 (11.50-3.90)
Sidechain outliers	138945	1003 (11.50-3.86)
RSRZ outliers	127900	1004 (9.50-3.80)
RNA backbone	3102	1079 (11.50-3.00)

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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	295	
2	B	264	
3	C	278	
4	D	243	
5	E	263	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	F	204	15% 67% 21% 6%
7	G	249	25% 65% 26% 5%
8	H	194	18% 61% 26% 7%
9	I	208	35% 69% 21% 8%
10	J	194	30% 58% 26% 8% 6%
11	K	165	44% 31% 15% 9% 5% 41%
12	L	158	18% 66% 26%
13	M	132	19% 62% 25% 7% 6%
14	N	151	26% 68% 29%
15	O	151	27% 60% 27% 10%
16	P	145	34% 50% 27% 8% 12%
17	Q	146	38% 68% 23% 5%
18	R	135	12% 65% 18% 7% 7%
19	S	152	51% 59% 21% 9% 10%
20	T	145	52% 70% 21% 5%
21	U	119	42% 49% 29% 5% 13%
22	V	83	2% 54% 30% 13%
23	W	130	51% 81% 16%
24	X	143	60% 68% 26% 6%
25	Y	133	12% 59% 26% 8% 5%
26	Z	125	40% 38% 16% 5% 40%
27	a	115	24% 63% 22% 7% 7%
28	b	84	42% 65% 25% 10%
29	c	69	3% 65% 20% 7% 7%
30	d	56	20% 66% 27% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
31	e	133	
32	f	156	
33	g	317	
34	i	1863	
35	l	113	

2 Entry composition [i](#)

There are 35 unique types of molecules in this entry. The entry contains 77211 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called 40S ribosomal protein SA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	208	1642	1045	289	300	8	0	0	0

- Molecule 2 is a protein called 40S ribosomal protein S3a.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	215	1741	1107	309	310	15	0	0	0

- Molecule 3 is a protein called 40S ribosomal protein S2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	226	1742	1127	300	306	9	0	0	0

- Molecule 4 is a protein called 40S ribosomal protein S3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	227	1764	1124	317	315	8	0	0	0

- Molecule 5 is a protein called 40S ribosomal protein S4X.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	263	2083	1329	385	359	10	0	0	0

- Molecule 6 is a protein called 40S ribosomal protein S5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	F	191	1509	943	286	273	7	0	0	0

- Molecule 7 is a protein called 40S ribosomal protein S6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
7	G	237	1923	1200	387	329	7	0	0	0

- Molecule 8 is a protein called 40S ribosomal protein S7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
8	H	190	1530	975	281	273	1	0	0	0

- Molecule 9 is a protein called 40S ribosomal protein S8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
9	I	206	1679	1054	329	291	5	0	0	0

- Molecule 10 is a protein called 40S Ribosomal protein S9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
10	J	182	1498	952	300	244	2	0	0	0

- Molecule 11 is a protein called 40S ribosomal protein S10.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
11	K	98	827	539	148	134	6	0	0	0

- Molecule 12 is a protein called 40S ribosomal protein S11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
12	L	158	1296	827	241	221	7	0	0	0

- Molecule 13 is a protein called 40S ribosomal protein S12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
13	M	124	950	594	169	179	8	0	0	0

- Molecule 14 is a protein called 40S ribosomal protein S13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
14	N	150	1208	773	229	205	1	0	0	0

- Molecule 15 is a protein called 40S ribosomal protein S14.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
15	O	136	1016	621	199	190	6	0	0	0

- Molecule 16 is a protein called 40S ribosomal protein S15.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
16	P	127	1060	673	201	179	7	0	0	0

- Molecule 17 is a protein called 40S ribosomal protein S16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
17	Q	141	1124	715	212	194	3	0	0	0

- Molecule 18 is a protein called 40S ribosomal protein S17.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
18	R	126	1019	639	188	187	5	0	0	0

- Molecule 19 is a protein called 40S ribosomal protein S18.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
19	S	137	1139	714	231	193	1	0	0	0

- Molecule 20 is a protein called 40S ribosomal protein S19.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
20	T	141	1112	701	213	195	3	0	0	0

- Molecule 21 is a protein called 40S ribosomal protein S20.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
21	U	104	822	514	156	148	4	0	0	0

- Molecule 22 is a protein called 40S ribosomal protein S21.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
22	V	82	619	378	117	119	5	0	0	0

- Molecule 23 is a protein called 40S ribosomal protein S15A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
23	W	129	1034	659	193	176	6	0	0	0

- Molecule 24 is a protein called 40S ribosomal protein S23.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
24	X	142	1106	698	220	184	4	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
X	1	MET	ALA	SEE REMARK 999	UNP G1SZ47

- Molecule 25 is a protein called 40S ribosomal protein S24.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
25	Y	126	1021	645	198	173	5	0	0	0

- Molecule 26 is a protein called 40S ribosomal protein S25.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
26	Z	75	598	382	111	104	1	0	0	0

- Molecule 27 is a protein called 40S ribosomal protein S26.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
27	a	107	844	527	173	138	6	0	0	0

- Molecule 28 is a protein called 40S ribosomal protein S27.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
28	b	84	659	413	122	116	8	0	0	0

- Molecule 29 is a protein called 40S ribosomal protein S28.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
29	c	64	506	308	102	94	2	0	0	0

- Molecule 30 is a protein called 40S ribosomal protein S29.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
30	d	53	445	278	90	72	5	0	0	0

- Molecule 31 is a protein called 40S ribosomal protein S30.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
31	e	59	473	293	104	75	1	0	0	0

- Molecule 32 is a protein called 40S ribosomal protein S27A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
32	f	71	581	367	109	98	7	0	0	0

- Molecule 33 is a protein called 40S ribosomal protein RACK1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
33	g	313	2436	1535	424	465	12	0	0	0

- Molecule 34 is a RNA chain called 18S ribosomal RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
34	i	1797	37514	16712	6633	12373	1796	0	0	0

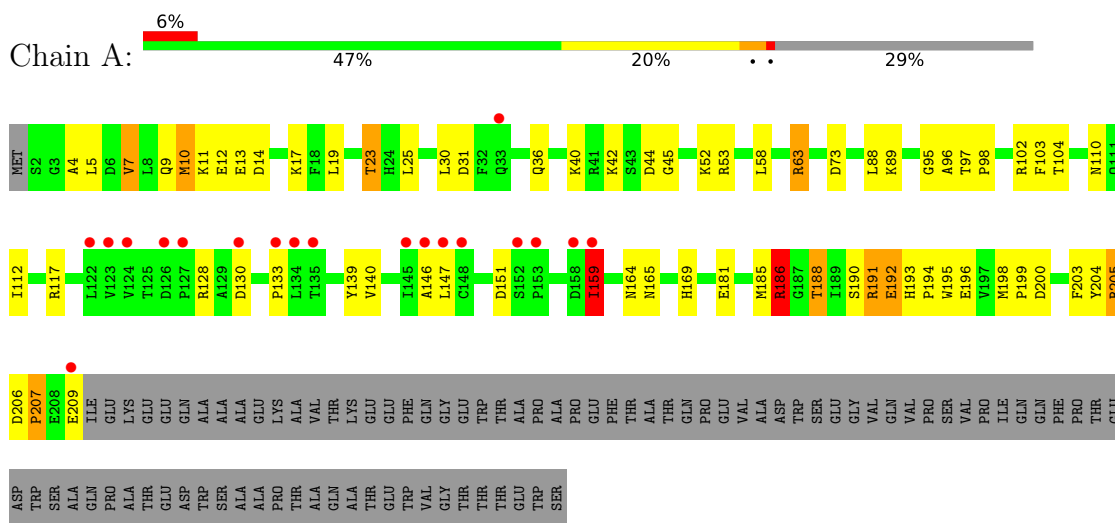
- Molecule 35 is a protein called Eukaryotic translation initiation factor 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
35	1	85	691	438	125	126	2	0	0	0

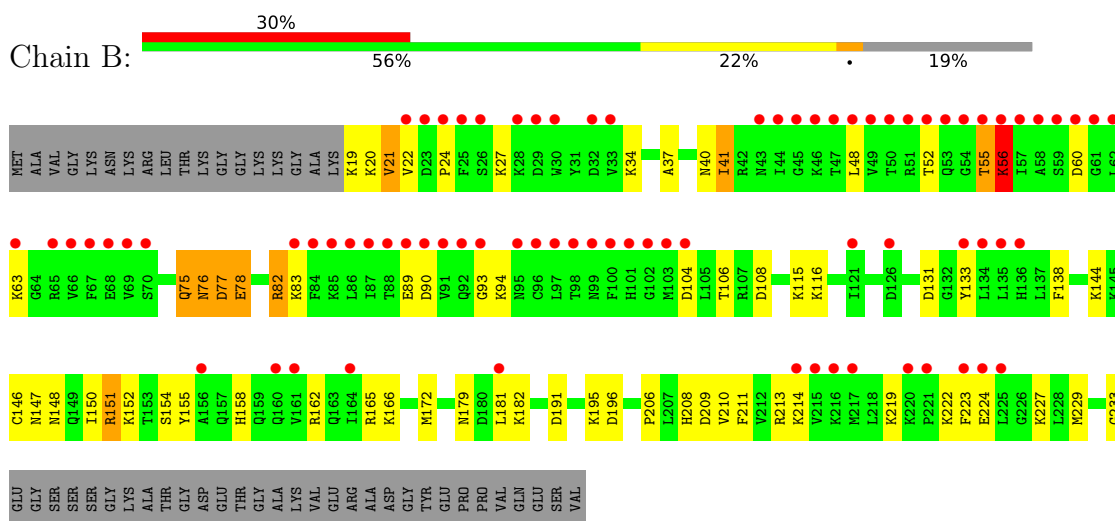
3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: 40S ribosomal protein SA

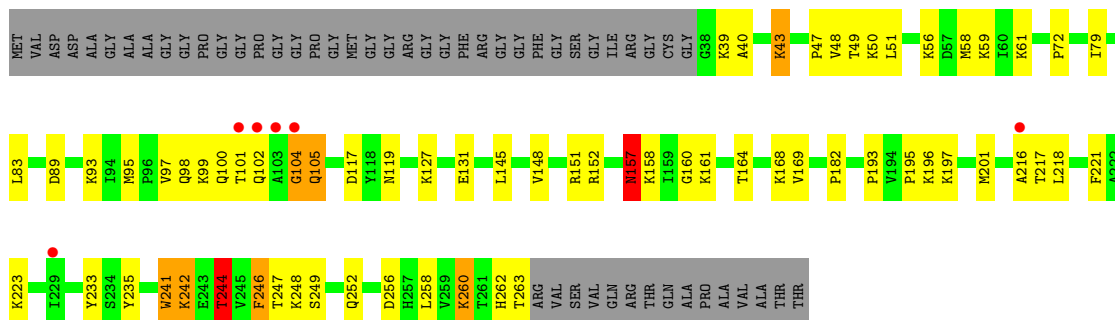


- Molecule 2: 40S ribosomal protein S3a

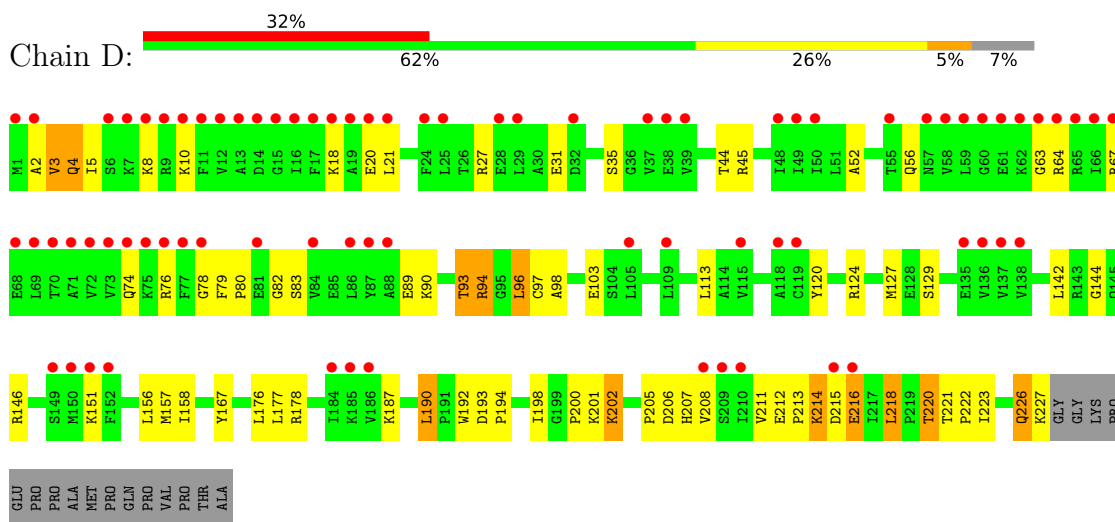


- Molecule 3: 40S ribosomal protein S2

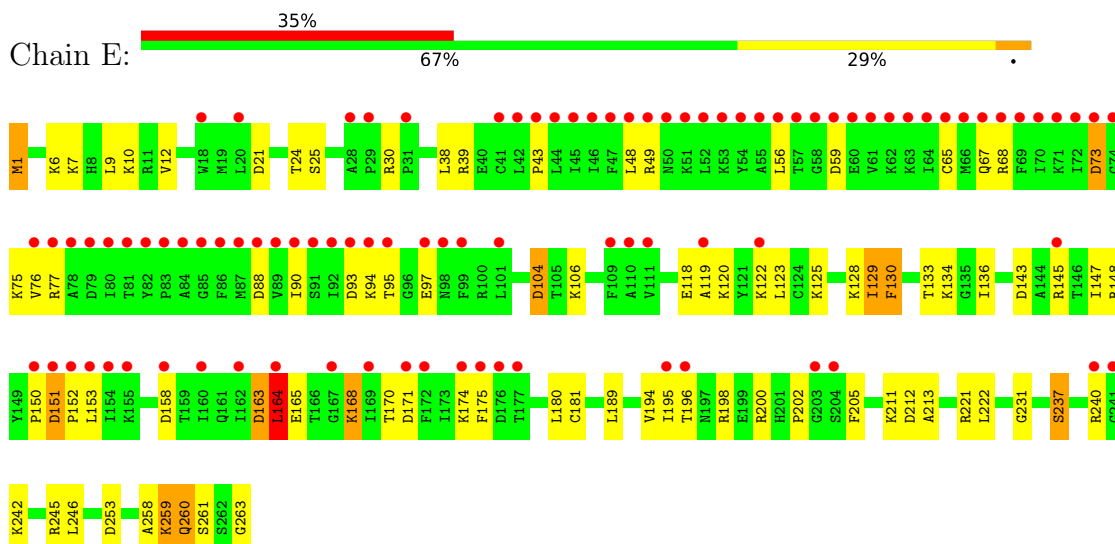




• Molecule 4: 40S ribosomal protein S3

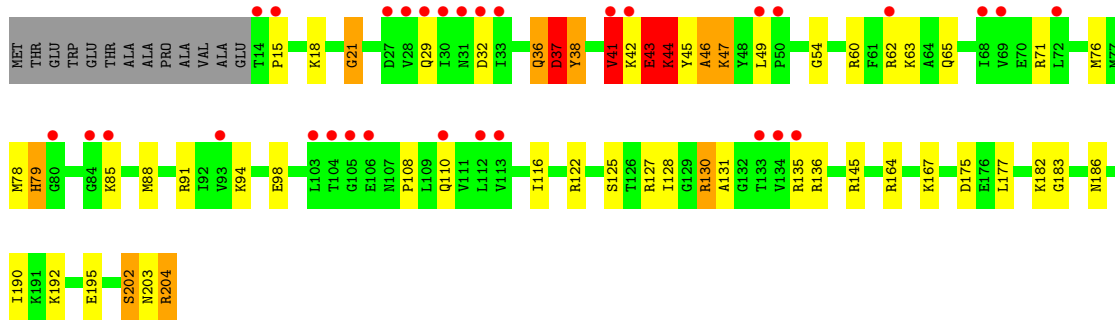


• Molecule 5: 40S ribosomal protein S4X

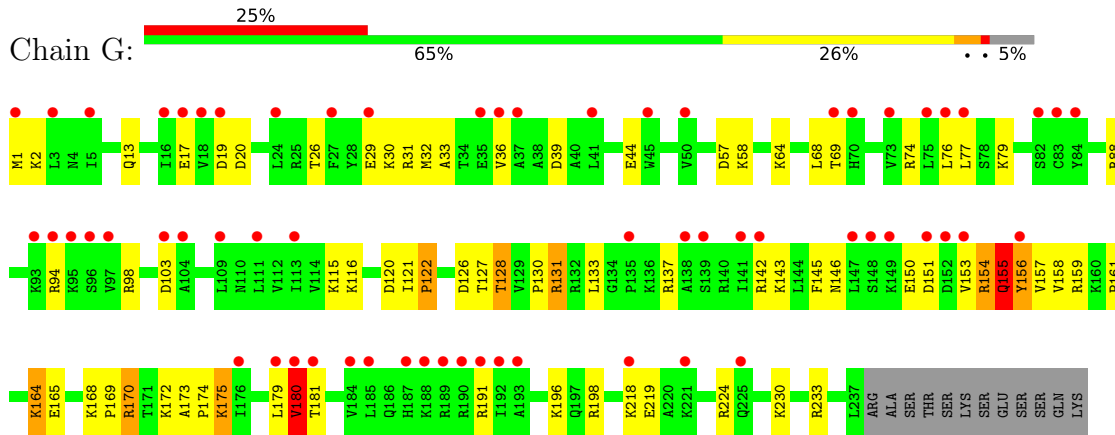


• Molecule 6: 40S ribosomal protein S5

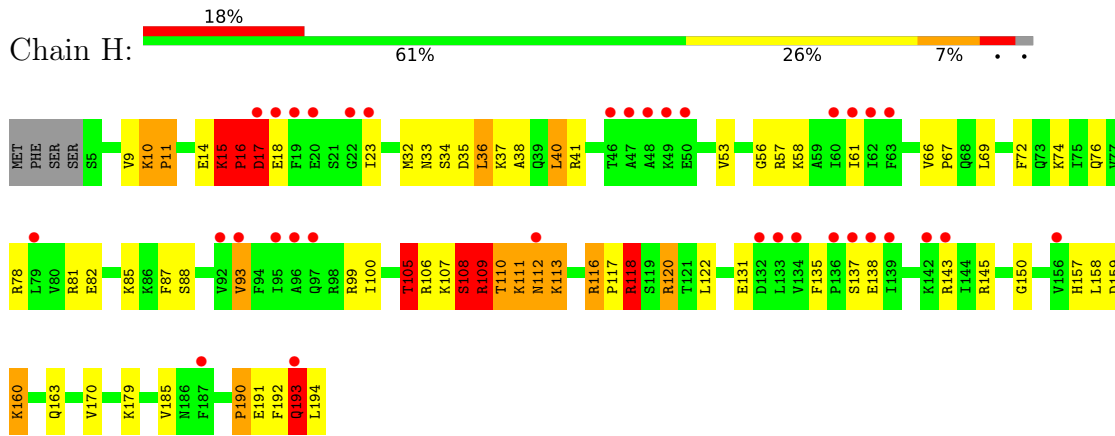




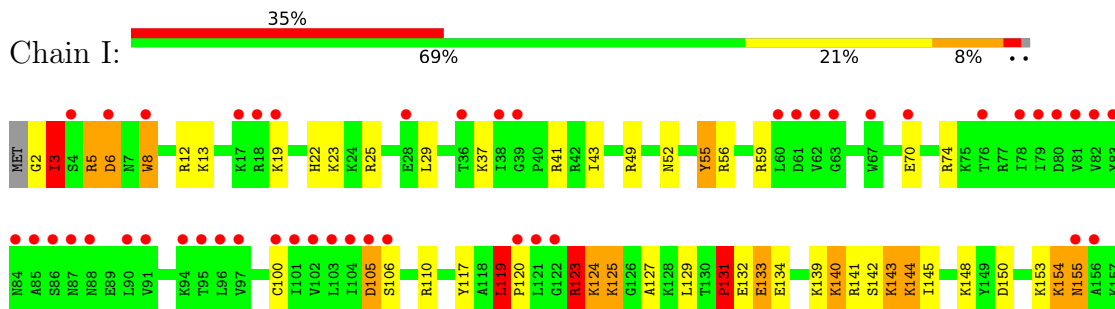
- Molecule 7: 40S ribosomal protein S6

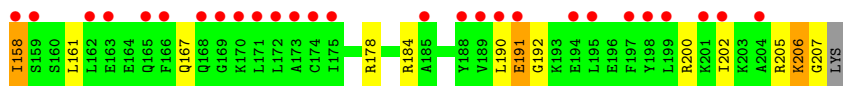


- Molecule 8: 40S ribosomal protein S7

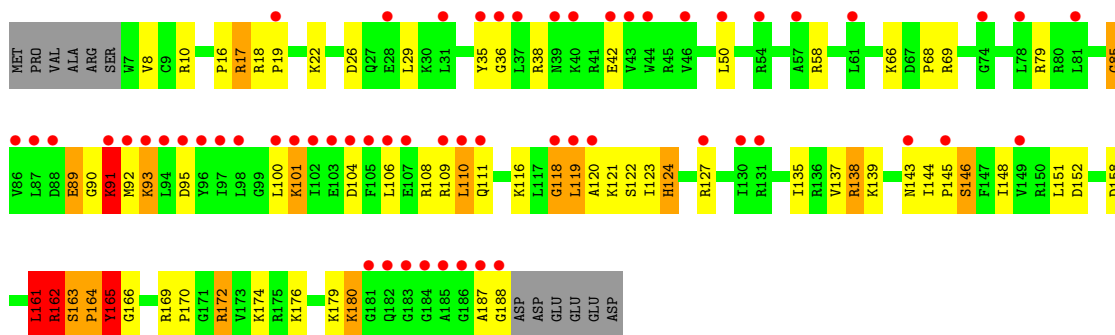


- Molecule 9: 40S ribosomal protein S8

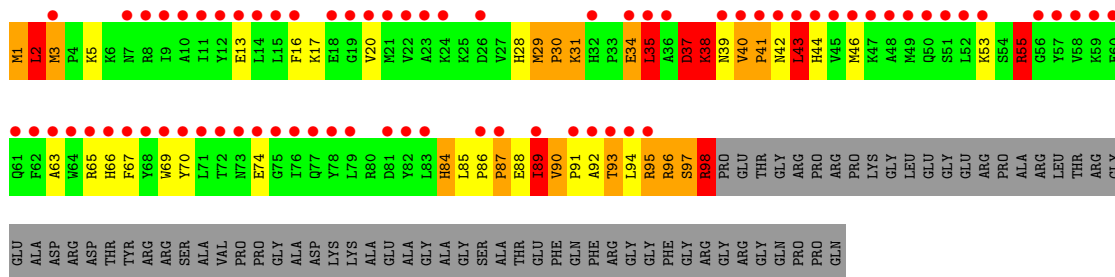
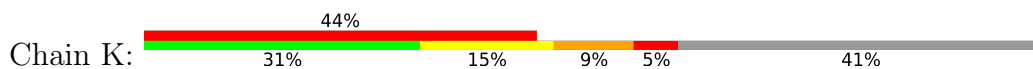




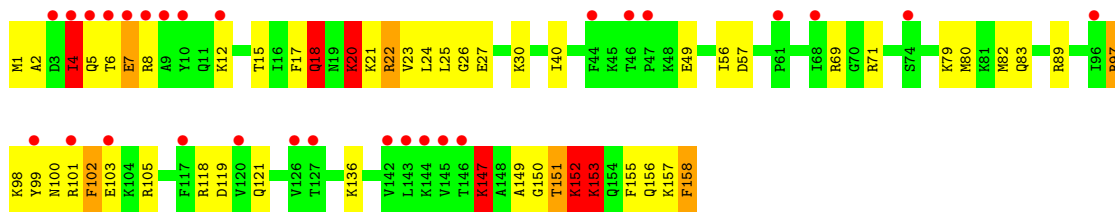
• Molecule 10: 40S Ribosomal protein S9



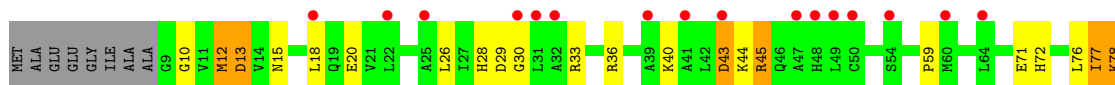
• Molecule 11: 40S ribosomal protein S10



• Molecule 12: 40S ribosomal protein S11

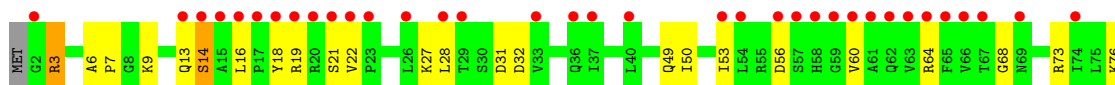


• Molecule 13: 40S ribosomal protein S12

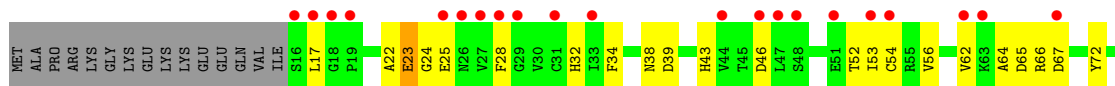




- Molecule 14: 40S ribosomal protein S13



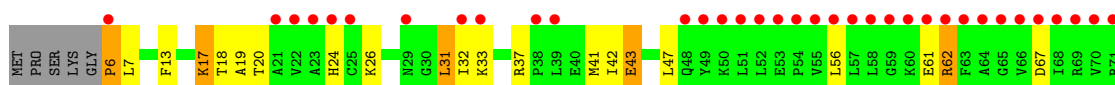
- Molecule 15: 40S ribosomal protein S14



- Molecule 16: 40S ribosomal protein S15

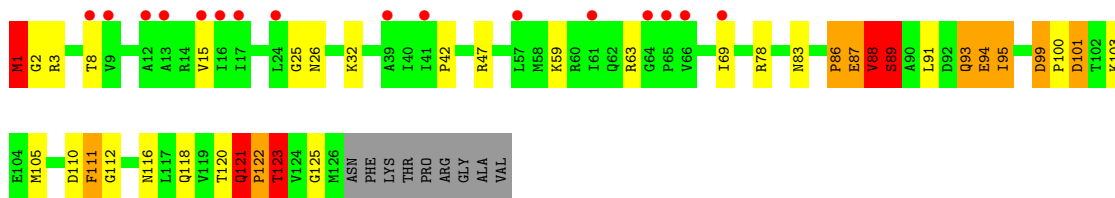


- Molecule 17: 40S ribosomal protein S16

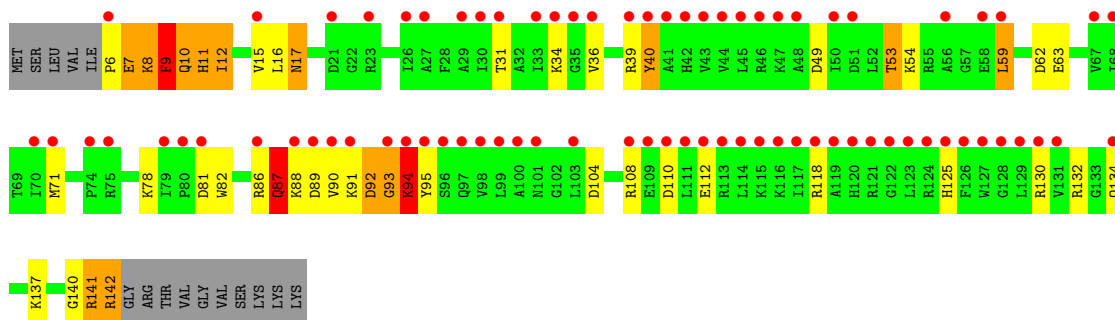


- Molecule 18: 40S ribosomal protein S17

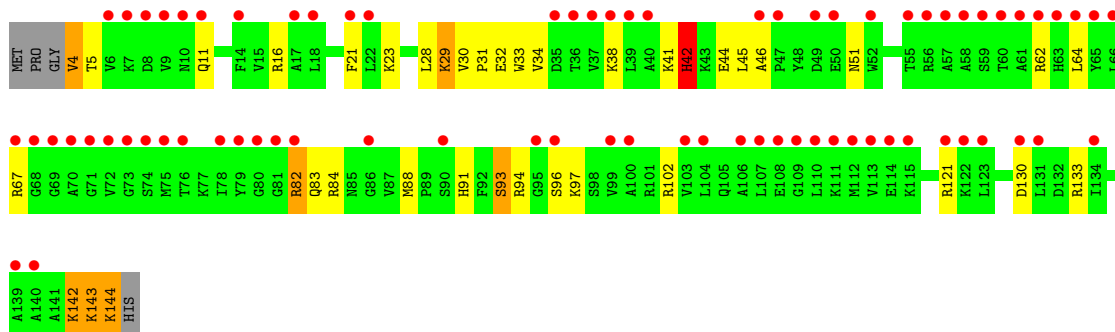




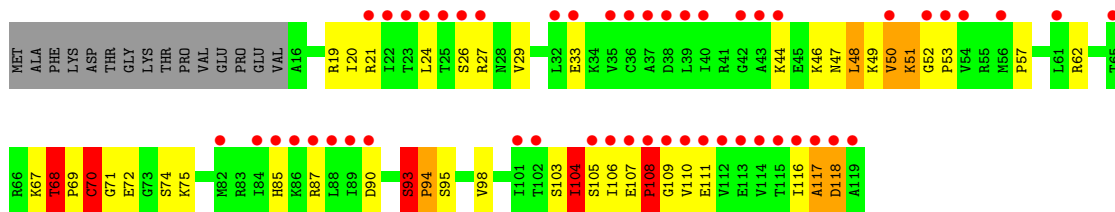
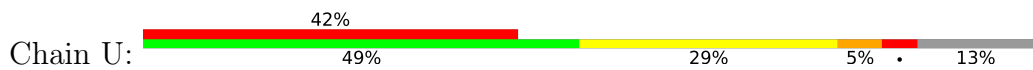
• Molecule 19: 40S ribosomal protein S18



• Molecule 20: 40S ribosomal protein S19



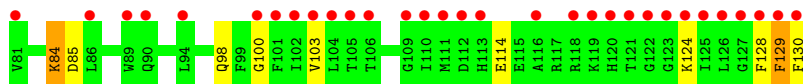
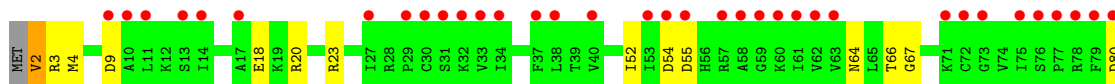
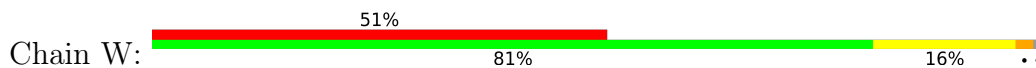
• Molecule 21: 40S ribosomal protein S20



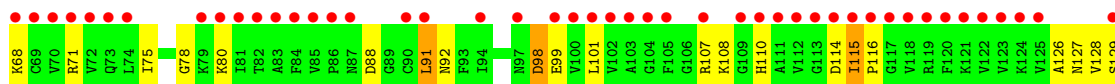
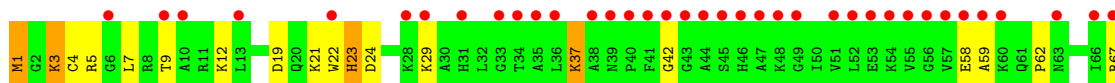
• Molecule 22: 40S ribosomal protein S21



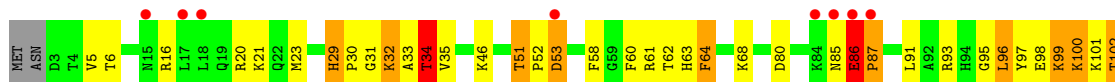
- Molecule 23: 40S ribosomal protein S15A



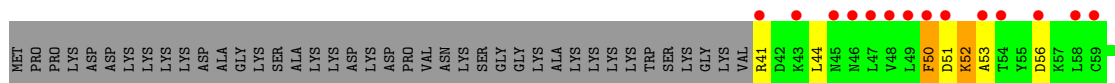
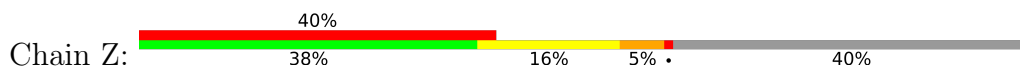
- Molecule 24: 40S ribosomal protein S23

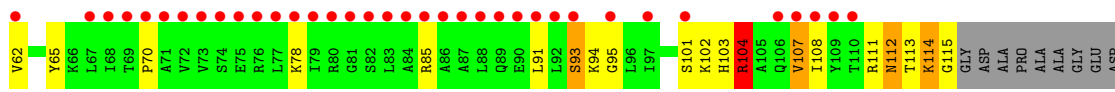


- Molecule 25: 40S ribosomal protein S24



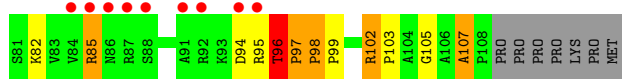
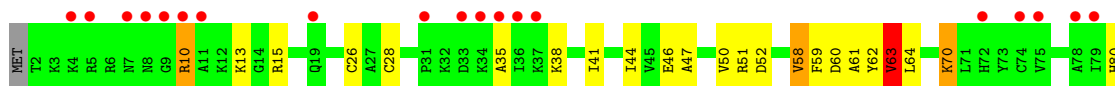
- Molecule 26: 40S ribosomal protein S25



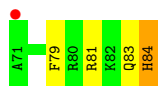
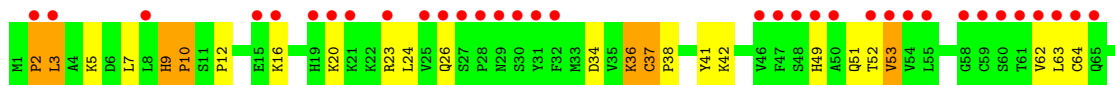
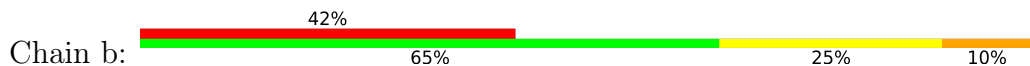


ALA

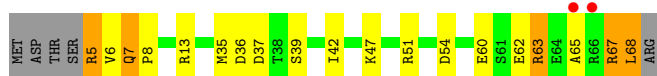
• Molecule 27: 40S ribosomal protein S26



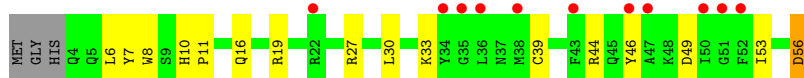
• Molecule 28: 40S ribosomal protein S27



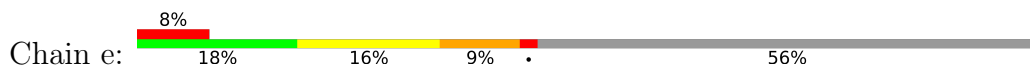
• Molecule 29: 40S ribosomal protein S28

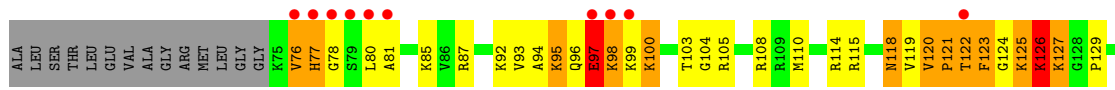


• Molecule 30: 40S ribosomal protein S29



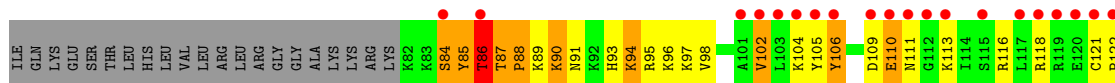
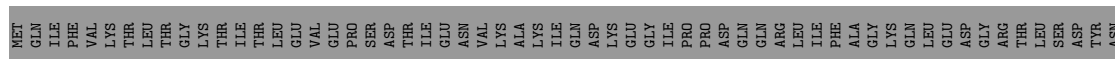
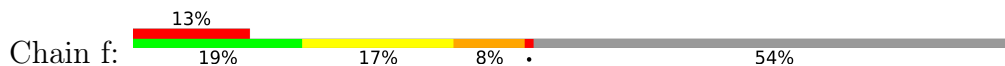
• Molecule 31: 40S ribosomal protein S30



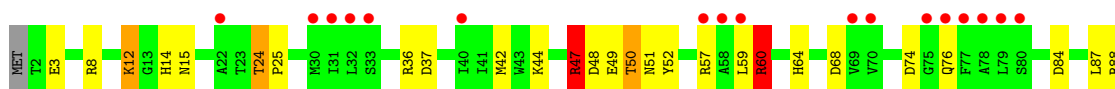
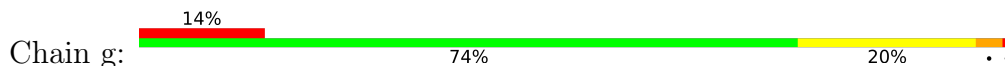


S133

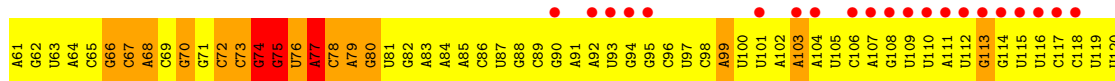
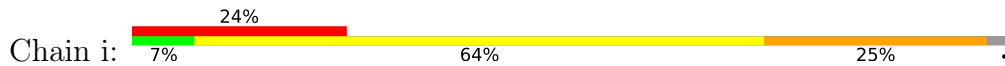
• Molecule 32: 40S ribosomal protein S27A

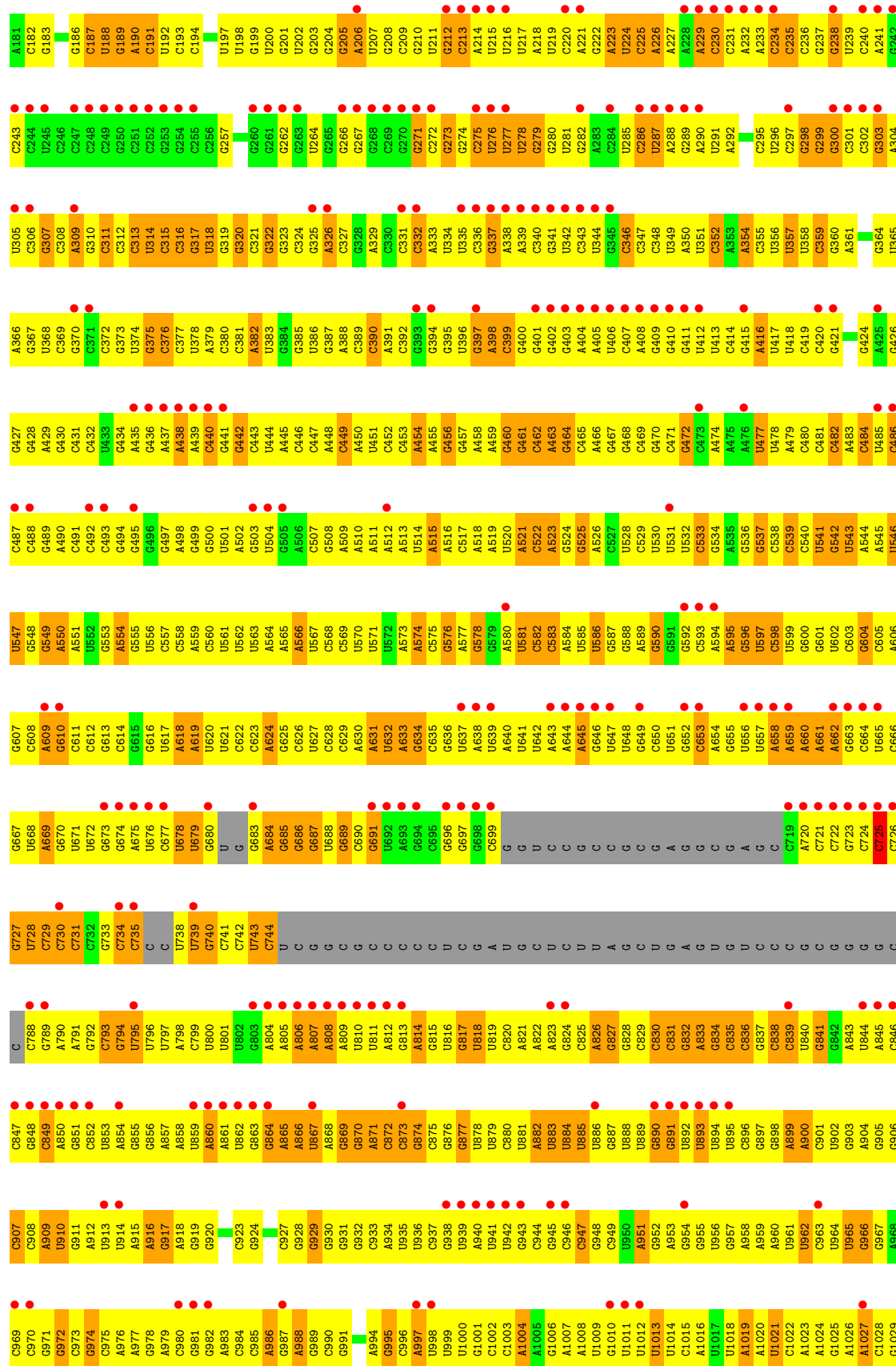


• Molecule 33: 40S ribosomal protein RACK1

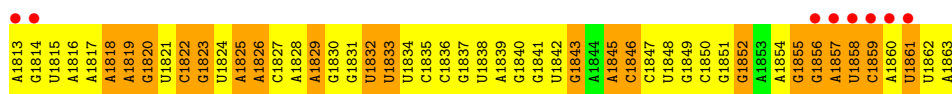


• Molecule 34: 18S ribosomal RNA

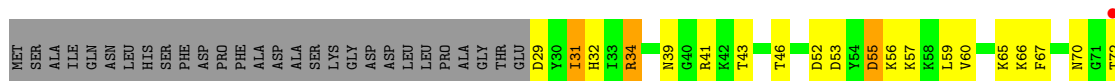




G1763	G1693	G1633	G1572	G1512	G1452	G1392	G1332	A1272	C1212	U1150	U1151	U1152	G1090	A1030
G1764	A1694	G1634	U1573	C1513	U1453	U1393	C1333	G1273	A1213	U1151	U1152	G1092	U1091	A1031
C1765	G1695	A1635	A1574	U1514	G1454	G1394	G1334	C1274	C1214	G1153	G1154	G1093	G1093	A1032
G1766	G1696	A1636	A1575	G1515	G1455	C1395	U1335	C1275	C1215	G1155	G1156	G1094	G1094	A1033
G1767	G1697	U1637	C1576	U1516	G1456	U1396	U1336	G1276	A1216	G1157	G1158	G1095	G1095	A1034
G1768	G1698	U1638	C1577	A1517	G1457	A1397	G1337	G1277	G1217	U1157	U1158	G1096	G1096	A1035
G1769	G1699	C1639	C1578	U1518	U1458	A1398	U1338	A1278	G1218	U1159	U1160	U1097	U1097	A1036
C1770	C1700	C1640	G1579	G1519	U1459	C1399	U1339	C1279	A1219	C1158	C1159	U1098	U1098	A1037
C1771	G1701	G1641	U1580	C1520	U1460	A1400	A1340	A1280	G1220	C1159	C1160	G1099	G1099	A1038
C1772	U1702	A1642	U1581	G1521	A1461	A1401	G1341	G1281	U1221	G1160	G1161	G1100	G1100	A1039
C1773	G1703	G1643	C1522	U1522	G1462	G1402	G1342	G1282	G1222	G1161	G1162	G1101	G1101	A1040
G1774	G1704	U1644	G1523	G1523	C1463	A1403	U1343	G1283	A1223	G1163	G1164	C1102	C1102	U1041
G1775	G1705	A1645	U1524	U1524	C1464	U1404	G1344	U1284	A1224	G1163	G1164	G1103	G1103	U1042
C1776	U1706	A1646	U1525	U1525	A1465	A1405	G1345	U1285	G1225	G1164	G1165	G1104	G1104	C1043
C1777	A1707	G1647	A1526	A1526	C1466	G1406	U1346	G1286	C1226	G1165	G1166	C1105	C1105	G1044
C1778	C1708	U1648	C1527	U1527	C1467	G1407	G1347	G1287	C1227	G1166	G1167	G1106	G1106	A1046
C1779	U1709	G1649	A1528	A1528	C1468	A1408	G1348	C1288	U1228	G1167	G1168	U1107	U1107	G1047
C1780	U1710	C1650	G1529	G1529	G1469	G1409	A1349	A1289	G1229	G1168	G1169	U1108	U1108	A1048
C1781	C1711	G1651	U1530	U1530	A1470	A1410	G1350	G1290	C1230	U1170	U1171	A1109	A1109	C1049
C1782	G1712	G1652	U1531	U1531	G1471	C1411	A1291	A1231	G1231	U1171	U1172	U1110	U1110	G1050
C1783	G1713	A1653	A1532	A1532	A1472	C1412	U1292	G1232	G1232	G1171	G1172	G1111	G1111	A1051
C1784	U1714	C1654	C1533	C1533	U1473	C1413	A1353	A1293	C1233	G1172	G1173	U1112	U1112	U1052
C1785	U1715	U1655	U1534	U1534	U1474	C1414	G1354	G1294	U1234	U1173	U1174	C1113	C1113	C1053
C1786	U1716	A1656	G1535	G1535	G1475	C1415	U1355	A1295	A1235	U1174	U1175	C1114	C1114	A1054
C1787	G1717	U1657	U1536	U1536	A1476	G1416	U1356	U1296	A1236	G1175	G1176	C1115	C1115	G1055
C1788	G1718	A1658	C1537	C1537	G1477	A1417	G1357	A1297	A1237	C1176	C1177	U1116	U1116	A1056
C1789	U1719	G1659	U1538	U1538	C1478	G1418	U1358	G1298	U1238	A1177	A1178	G1117	G1117	U1057
C1790	U1720	U1660	G1539	G1539	C1479	G1419	U1359	C1299	U1239	A1178	A1179	U1118	U1118	A1058
C1791	G1721	C1661	A1540	A1540	A1480	G1420	U1360	U1300	U1240	C1181	C1182	C1119	C1119	C1059
C1792	U1722	U1662	G1541	G1541	U1481	G1421	G1361	C1301	G1241	U1182	U1183	C1120	C1120	G1060
C1793	G1723	G1663	A1542	A1542	A1482	U1422	G1362	U1302	A1242	U1183	U1184	G1121	G1121	G1061
C1794	U1724	C1664	C1543	C1543	A1483	C1423	U1363	U1303	C1243	U1184	U1185	G1122	G1122	U1062
C1795	U1725	C1665	U1544	U1544	G1484	G1424	U1364	U1304	U1244	A1184	A1185	C1123	C1123	G1063
C1796	U1726	U1666	G1545	G1545	A1485	G1425	A1365	C1425	C1245	U1185	U1186	C1124	C1124	G1064
C1797	G1727	G1667	U1546	U1546	G1486	C1426	A1366	U1306	A1246	A1186	A1187	G1125	G1125	U1065
C1798	U1728	U1668	G1547	G1547	G1487	G1427	U1367	C1307	A1247	U1187	U1188	G1126	G1126	A1066
C1799	G1729	G1669	U1548	U1548	U1488	U1428	U1368	G1308	C1248	U1188	U1189	G1127	G1127	G1067
C1800	U1730	A1670	C1549	C1549	U1489	C1429	G1369	A1309	A1249	U1189	U1190	C1128	C1128	U1068
C1801	G1731	U1671	U1550	U1550	U1490	C1430	C1370	U1310	C1250	A1190	A1191	A1129	A1129	U1069
C1802	U1732	U1672	A1551	A1551	U1491	C1431	G1371	U1311	G1251	A1191	A1192	G1130	G1130	C1070
C1803	G1733	C1673	C1552	C1552	U1492	C1432	A1372	C1312	G1252	A1192	G1193	C1131	C1131	C1071
C1804	U1734	A1674	U1553	U1553	G1493	C1433	U1373	U1313	G1253	G1193	G1194	U1132	U1132	G1072
C1805	G1735	G1675	A1554	A1554	A1494	A1434	G1374	G1314	A1254	A1194	A1195	U1133	U1133	A1073
C1806	U1736	U1676	U1555	U1555	U1495	A1435	A1375	U1315	A1255	A1195	A1196	C1134	C1134	C1074
C1807	C1737	C1677	A1556	A1556	G1496	C1436	C1376	G1316	A1256	A1196	A1197	G1135	G1135	U1075
C1808	G1738	C1678	U1557	U1557	U1497	U1437	G1377	G1317	C1257	U1197	U1198	G1136	G1136	A1076
C1809	U1739	U1679	G1558	G1558	C1498	U1438	A1378	U1318	C1258	G1198	G1199	G1137	G1137	U1077
C1810	A1740	U1680	C1559	C1559	C1499	C1439	A1379	U1319	U1259	G1199	G1200	G1138	G1138	A1078
C1811	U1741	G1681	U1560	U1560	U1500	U1440	C1380	U1320	A1260	A1200	A1201	A1139	A1139	A1079
C1812	G1742	C1682	G1561	G1561	U1501	U1441	G1381	G1321	A1261	A1201	G1202	A1140	A1140	U1080
C1813	U1743	U1683	U1562	U1562	A1502	A1442	A1382	U1322	C1262	G1202	G1203	A1141	A1141	G1081
C1814	G1744	C1684	G1563	G1563	G1503	G1443	G1383	G1323	C1263	A1203	A1204	C1142	C1142	G1082
C1815	U1745	U1685	U1564	U1564	A1504	A1444	A1384	G1324	C1264	A1204	A1205	C1143	C1143	A1083
C1816	G1746	G1686	G1565	G1565	U1505	G1445	C1385	U1325	G1265	G1205	G1206	A1144	A1144	U1084
C1817	U1747	U1687	U1566	U1566	G1506	G1446	U1386	G1326	G1266	G1206	G1207	A1145	A1145	G1085
C1818	G1748	G1688	G1567	G1567	U1507	G1447	C1387	C1327	C1267	G1207	G1208	A1146	A1146	C1086
C1819	U1749	U1689	U1568	U1568	C1508	A1448	U1388	A1328	C1268	G1208	G1209	G1147	G1147	U1087
C1820	G1750	A1690	G1569	G1569	G1509	C1449	G1389	U1329	C1269	A1209	A1210	U1148	U1148	G1088
C1821	U1751	C1691	U1570	U1570	U1510	A1450	G1390	G1330	G1270	G1210	G1211	C1149	C1149	A1089
C1822	G1752	A1692	G1571	G1571	G1511	A1451	C1391	G1331	G1271	G1211	G1212			



- Molecule 35: Eukaryotic translation initiation factor 1



4 Data and refinement statistics i

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	292.12Å 292.12Å 477.67Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	58.11 – 7.81 64.21 – 7.81	Depositor EDS
% Data completeness (in resolution range)	94.8 (58.11-7.81) 94.9 (64.21-7.81)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.96 (at 7.40Å)	Xtrriage
Refinement program	PHENIX 1.7.3_928	Depositor
R, R_{free}	0.347 , 0.347 0.334 , 0.337	Depositor DCC
R_{free} test set	1305 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å ²)	667.2	Xtrriage
Anisotropy	0.309	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.22 , 117.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.34$, $\langle L^2 \rangle = 0.17$	Xtrriage
Estimated twinning fraction	0.120 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.86	EDS
Total number of atoms	77211	wwPDB-VP
Average B, all atoms (Å ²)	216.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.66% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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

5 Model quality i

5.1 Standard geometry i

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	0.76	2/1679 (0.1%)	1.06	16/2283 (0.7%)
2	B	0.79	6/1769 (0.3%)	1.08	20/2367 (0.8%)
3	C	0.97	7/1778 (0.4%)	1.19	19/2399 (0.8%)
4	D	1.03	6/1792 (0.3%)	1.30	22/2412 (0.9%)
5	E	0.76	4/2125 (0.2%)	0.98	23/2856 (0.8%)
6	F	1.00	5/1531 (0.3%)	1.21	17/2059 (0.8%)
7	G	0.97	15/1946 (0.8%)	1.23	25/2590 (1.0%)
8	H	1.09	9/1553 (0.6%)	2.19	29/2079 (1.4%)
9	I	1.11	7/1708 (0.4%)	1.51	33/2278 (1.4%)
10	J	1.27	19/1522 (1.2%)	1.51	42/2031 (2.1%)
11	K	1.21	7/851 (0.8%)	1.78	32/1147 (2.8%)
12	L	1.10	6/1319 (0.5%)	1.40	17/1761 (1.0%)
13	M	1.00	3/960 (0.3%)	1.23	7/1287 (0.5%)
14	N	0.83	3/1232 (0.2%)	1.00	12/1656 (0.7%)
15	O	0.61	0/1029	1.05	12/1380 (0.9%)
16	P	0.75	1/1079 (0.1%)	1.43	32/1437 (2.2%)
17	Q	0.70	3/1142 (0.3%)	1.12	15/1528 (1.0%)
18	R	1.23	10/1031 (1.0%)	1.64	30/1383 (2.2%)
19	S	1.21	11/1157 (1.0%)	1.61	36/1548 (2.3%)
20	T	0.95	3/1132 (0.3%)	1.25	13/1517 (0.9%)
21	U	0.96	1/832 (0.1%)	1.59	28/1117 (2.5%)
22	V	0.75	1/626 (0.2%)	1.39	15/839 (1.8%)
23	W	0.85	4/1051 (0.4%)	0.86	9/1406 (0.6%)
24	X	1.00	8/1124 (0.7%)	1.24	21/1500 (1.4%)
25	Y	0.93	3/1038 (0.3%)	1.42	20/1380 (1.4%)
26	Z	1.04	5/604 (0.8%)	1.35	17/810 (2.1%)
27	a	0.89	4/860 (0.5%)	1.60	21/1156 (1.8%)
28	b	1.02	2/673 (0.3%)	1.36	12/902 (1.3%)
29	c	0.80	1/508 (0.2%)	1.18	8/680 (1.2%)
30	d	0.89	2/455 (0.4%)	0.79	3/603 (0.5%)
31	e	1.46	5/478 (1.0%)	1.43	11/628 (1.8%)
32	f	1.10	4/593 (0.7%)	1.49	15/786 (1.9%)
33	g	0.91	1/2493 (0.0%)	1.29	26/3394 (0.8%)
34	i	2.42	1848/41880 (4.4%)	2.22	2570/65161 (3.9%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
35	l	1.00	3/701 (0.4%)	1.11	4/936 (0.4%)
All	All	1.85	2019/82251 (2.5%)	1.87	3232/119296 (2.7%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	11
2	B	0	4
3	C	1	5
4	D	0	5
5	E	1	2
6	F	0	3
7	G	0	1
8	H	0	10
9	I	0	8
10	J	1	11
11	K	0	11
12	L	0	7
13	M	0	1
14	N	0	4
15	O	0	1
16	P	0	10
17	Q	0	4
18	R	1	5
19	S	1	10
20	T	1	6
21	U	0	8
22	V	0	9
23	W	0	2
24	X	0	4
25	Y	1	6
26	Z	0	6
27	a	0	2
28	b	0	3
31	e	0	5
32	f	0	6
33	g	0	13
34	i	6	0
All	All	13	183

The worst 5 of 2019 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
34	i	1683	C	O3'-P	41.12	2.10	1.61
34	i	1322	U	C2'-C1'	-25.29	1.25	1.53
34	i	66	G	C2'-C1'	-24.37	1.26	1.53
34	i	858	A	C2'-C1'	-23.82	1.27	1.53
34	i	652	G	C2'-C1'	-23.61	1.27	1.53

The worst 5 of 3232 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	H	109	ARG	NE-CZ-NH2	-53.46	93.57	120.30
34	i	1683	C	P-O3'-C3'	-46.88	63.44	119.70
8	H	109	ARG	NE-CZ-NH1	42.70	141.65	120.30
34	i	1774	G	P-O3'-C3'	38.10	165.42	119.70
34	i	1114	C	O4'-C1'-N1	35.28	136.42	108.20

5 of 13 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	C	157	ASN	CA
5	E	171	ASP	CA
10	J	138	ARG	CA
18	R	3	ARG	CA
19	S	92	ASP	CA

5 of 183 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	146	ALA	Mainchain
1	A	23	THR	Mainchain
1	A	4	ALA	Peptide
1	A	63	ARG	Sidechain
1	A	97	THR	Mainchain

5.2 Too-close contacts [i](#)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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	
1	A	206/295 (70%)	156 (76%)	23 (11%)	27 (13%)	0	5
2	B	213/264 (81%)	173 (81%)	25 (12%)	15 (7%)	1	14
3	C	224/278 (81%)	200 (89%)	13 (6%)	11 (5%)	2	20
4	D	225/243 (93%)	180 (80%)	23 (10%)	22 (10%)	0	10
5	E	261/263 (99%)	209 (80%)	28 (11%)	24 (9%)	1	11
6	F	189/204 (93%)	162 (86%)	15 (8%)	12 (6%)	1	17
7	G	235/249 (94%)	201 (86%)	19 (8%)	15 (6%)	1	16
8	H	188/194 (97%)	146 (78%)	11 (6%)	31 (16%)	0	3
9	I	204/208 (98%)	169 (83%)	12 (6%)	23 (11%)	0	7
10	J	180/194 (93%)	138 (77%)	18 (10%)	24 (13%)	0	5
11	K	96/165 (58%)	67 (70%)	11 (12%)	18 (19%)	0	2
12	L	156/158 (99%)	132 (85%)	10 (6%)	14 (9%)	1	11
13	M	122/132 (92%)	85 (70%)	16 (13%)	21 (17%)	0	3
14	N	148/151 (98%)	124 (84%)	18 (12%)	6 (4%)	3	23
15	O	134/151 (89%)	101 (75%)	14 (10%)	19 (14%)	0	4
16	P	125/145 (86%)	92 (74%)	16 (13%)	17 (14%)	0	4
17	Q	139/146 (95%)	109 (78%)	20 (14%)	10 (7%)	1	14
18	R	124/135 (92%)	96 (77%)	13 (10%)	15 (12%)	0	6
19	S	135/152 (89%)	106 (78%)	20 (15%)	9 (7%)	1	15
20	T	139/145 (96%)	119 (86%)	10 (7%)	10 (7%)	1	14
21	U	102/119 (86%)	76 (74%)	10 (10%)	16 (16%)	0	3
22	V	80/83 (96%)	55 (69%)	11 (14%)	14 (18%)	0	3
23	W	127/130 (98%)	110 (87%)	15 (12%)	2 (2%)	9	44
24	X	140/143 (98%)	121 (86%)	11 (8%)	8 (6%)	1	18
25	Y	124/133 (93%)	91 (73%)	15 (12%)	18 (14%)	0	4

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
26	Z	73/125 (58%)	52 (71%)	12 (16%)	9 (12%)	0	5
27	a	105/115 (91%)	72 (69%)	14 (13%)	19 (18%)	0	3
28	b	82/84 (98%)	57 (70%)	14 (17%)	11 (13%)	0	5
29	c	62/69 (90%)	44 (71%)	13 (21%)	5 (8%)	1	12
30	d	51/56 (91%)	46 (90%)	3 (6%)	2 (4%)	3	23
31	e	57/133 (43%)	37 (65%)	7 (12%)	13 (23%)	0	2
32	f	69/156 (44%)	38 (55%)	13 (19%)	18 (26%)	0	1
33	g	311/317 (98%)	271 (87%)	23 (7%)	17 (6%)	2	19
35	l	83/113 (74%)	49 (59%)	24 (29%)	10 (12%)	0	6
All	All	4909/5648 (87%)	3884 (79%)	520 (11%)	505 (10%)	0	8

5 of 505 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	9	GLN
1	A	31	ASP
1	A	45	GLY
1	A	103	PHE
1	A	164	ASN

5.3.2 Protein sidechains [i](#)

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	
1	A	174/244 (71%)	140 (80%)	34 (20%)	1	8
2	B	196/231 (85%)	155 (79%)	41 (21%)	1	6
3	C	187/215 (87%)	147 (79%)	40 (21%)	1	6
4	D	190/202 (94%)	144 (76%)	46 (24%)	0	4
5	E	225/225 (100%)	173 (77%)	52 (23%)	1	4
6	F	161/170 (95%)	117 (73%)	44 (27%)	0	3
7	G	207/218 (95%)	157 (76%)	50 (24%)	0	4

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
8	H	170/174 (98%)	124 (73%)	46 (27%)	0	3
9	I	177/179 (99%)	142 (80%)	35 (20%)	1	8
10	J	157/168 (94%)	128 (82%)	29 (18%)	1	9
11	K	89/136 (65%)	61 (68%)	28 (32%)	0	2
12	L	142/142 (100%)	105 (74%)	37 (26%)	0	3
13	M	101/108 (94%)	78 (77%)	23 (23%)	1	5
14	N	130/131 (99%)	103 (79%)	27 (21%)	1	6
15	O	106/119 (89%)	87 (82%)	19 (18%)	2	10
16	P	116/130 (89%)	84 (72%)	32 (28%)	0	3
17	Q	117/121 (97%)	89 (76%)	28 (24%)	0	4
18	R	114/121 (94%)	90 (79%)	24 (21%)	1	6
19	S	119/132 (90%)	95 (80%)	24 (20%)	1	7
20	T	113/116 (97%)	87 (77%)	26 (23%)	1	4
21	U	94/107 (88%)	74 (79%)	20 (21%)	1	6
22	V	67/68 (98%)	50 (75%)	17 (25%)	0	3
23	W	112/113 (99%)	98 (88%)	14 (12%)	4	19
24	X	114/115 (99%)	91 (80%)	23 (20%)	1	7
25	Y	108/115 (94%)	85 (79%)	23 (21%)	1	6
26	Z	66/103 (64%)	53 (80%)	13 (20%)	1	8
27	a	91/99 (92%)	76 (84%)	15 (16%)	2	12
28	b	76/76 (100%)	63 (83%)	13 (17%)	2	11
29	c	57/62 (92%)	46 (81%)	11 (19%)	1	8
30	d	47/49 (96%)	35 (74%)	12 (26%)	0	3
31	e	49/106 (46%)	26 (53%)	23 (47%)	0	0
32	f	64/140 (46%)	43 (67%)	21 (33%)	0	2
33	g	272/275 (99%)	224 (82%)	48 (18%)	2	11
35	l	74/96 (77%)	57 (77%)	17 (23%)	1	4
All	All	4282/4806 (89%)	3327 (78%)	955 (22%)	1	5

5 of 955 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
12	L	82	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
32	f	135	HIS
16	P	86	LEU
32	f	104	LYS
35	l	70	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 114 such sidechains are listed below:

Mol	Chain	Res	Type
14	N	62	GLN
35	l	48	GLN
19	S	11	HIS
33	g	237	ASN
30	d	26	ASN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
34	i	1721/1863 (92%)	498 (28%)	0

5 of 498 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
34	i	2	A
34	i	3	C
34	i	4	C
34	i	8	U
34	i	16	G

There are no RNA pucker outliers to report.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
34	i	11
10	J	3
4	D	2
19	S	2
31	e	1
9	I	1
21	U	1
3	C	1
7	G	1
18	R	1

The worst 5 of 24 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	i	326:A	O3'	327:C	P	8.42
1	i	309:A	O3'	310:G	P	7.21
1	i	304:A	O3'	305:U	P	5.77
1	i	209:C	O3'	210:G	P	5.60
1	i	1826:A	O3'	1827:C	P	5.04

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	208/295 (70%)	0.01	19 (9%) 9 11	173, 263, 322, 332	0
2	B	215/264 (81%)	1.79	78 (36%) 0 1	128, 196, 242, 252	0
3	C	226/278 (81%)	0.12	6 (2%) 54 48	82, 160, 268, 295	0
4	D	227/243 (93%)	1.58	78 (34%) 0 2	167, 217, 294, 337	0
5	E	263/263 (100%)	2.15	93 (35%) 0 2	117, 216, 286, 299	0
6	F	191/204 (93%)	0.91	31 (16%) 1 4	191, 263, 305, 317	0
7	G	237/249 (95%)	1.13	63 (26%) 0 2	164, 242, 325, 343	0
8	H	190/194 (97%)	0.75	34 (17%) 1 4	187, 326, 368, 385	0
9	I	206/208 (99%)	1.83	73 (35%) 0 2	89, 239, 276, 287	0
10	J	182/194 (93%)	1.34	58 (31%) 0 2	89, 166, 231, 276	0
11	K	98/165 (59%)	4.84	73 (74%) 0 0	222, 291, 316, 323	0
12	L	158/158 (100%)	0.71	28 (17%) 1 4	77, 176, 261, 270	0
13	M	124/132 (93%)	1.14	25 (20%) 1 3	298, 378, 406, 431	0
14	N	150/151 (99%)	1.33	39 (26%) 0 2	87, 150, 275, 297	0
15	O	136/151 (90%)	1.25	41 (30%) 0 2	92, 194, 256, 272	0
16	P	127/145 (87%)	1.69	49 (38%) 0 1	234, 305, 340, 361	0
17	Q	141/146 (96%)	2.31	55 (39%) 0 1	166, 287, 321, 331	0
18	R	126/135 (93%)	0.64	16 (12%) 3 7	174, 225, 322, 329	0
19	S	137/152 (90%)	3.00	77 (56%) 0 1	217, 311, 344, 357	0
20	T	141/145 (97%)	3.25	75 (53%) 0 1	238, 311, 341, 349	0
21	U	104/119 (87%)	2.89	50 (48%) 0 1	167, 266, 306, 317	0
22	V	82/83 (98%)	-0.22	2 (2%) 59 52	164, 218, 318, 328	0
23	W	129/130 (99%)	2.24	66 (51%) 0 1	107, 159, 204, 218	0
24	X	142/143 (99%)	3.83	86 (60%) 0 0	50, 90, 134, 155	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
25	Y	126/133 (94%)	0.67	16 (12%) 3 7	156, 205, 243, 261	0
26	Z	75/125 (60%)	2.90	50 (66%) 0 0	292, 320, 350, 360	0
27	a	107/115 (93%)	1.20	28 (26%) 0 2	87, 129, 234, 249	0
28	b	84/84 (100%)	1.75	35 (41%) 0 1	164, 223, 303, 334	0
29	c	64/69 (92%)	-0.06	2 (3%) 49 42	175, 224, 274, 287	0
30	d	53/56 (94%)	1.00	11 (20%) 1 3	183, 212, 290, 295	0
31	e	59/133 (44%)	0.59	10 (16%) 1 4	103, 158, 182, 189	0
32	f	71/156 (45%)	1.07	21 (29%) 0 2	213, 332, 401, 415	0
33	g	313/317 (98%)	0.57	43 (13%) 3 6	235, 304, 346, 364	0
34	i	1797/1863 (96%)	1.41	446 (24%) 0 2	49, 184, 371, 475	0
35	l	85/113 (75%)	0.08	3 (3%) 44 39	223, 235, 250, 253	0
All	All	6774/7511 (90%)	1.43	1880 (27%) 0 2	49, 228, 347, 475	0

The worst 5 of 1880 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
34	i	252	C	72.0
34	i	251	C	49.1
34	i	250	G	45.1
34	i	249	C	31.8
24	X	83	ALA	23.9

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.