



wwPDB X-ray Structure Validation Summary Report

Mar 31, 2014 – 06:13 PM BST

PDB ID : 2HGI
Title : Crystal structure of the 70S *Thermus thermophilus* ribosome showing how the 16S 3'-end mimicks mRNA E and P codons. This entry 2HGI contains 30S ribosomal subunit. The 50S ribosomal subunit can be found in PDB entry 2HGJ.
Authors : Jenner, L.; Yusupova, G.; Rees, B.; Moras, D.; Yusupov, M.
Deposited on : 2006-06-27
Resolution : 5.00 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

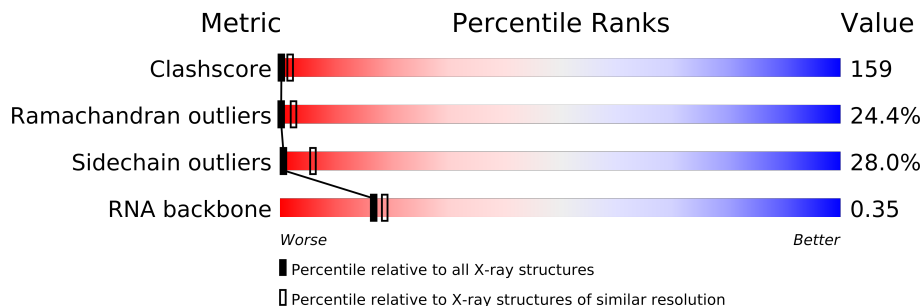
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : **NOT EXECUTED**
Percentile statistics : 21963
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable23004

1 Overall quality at a glance

The reported resolution of this entry is 5.00 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	1327 (6.50-3.50)
Ramachandran outliers	78287	1242 (6.50-3.50)
Sidechain outliers	78261	1221 (6.50-3.50)
RNA backbone	1838	1037 (7.00-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 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	1522	
2	C	77	
3	D	76	
4	E	256	
5	F	239	
6	G	209	
7	H	162	
8	I	101	
9	J	156	
10	K	138	
11	L	128	
12	M	105	
13	N	129	
14	O	132	
15	P	126	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
16	Q	61	
17	R	89	
18	S	88	
19	T	105	
20	U	88	
21	V	93	
22	W	106	
23	X	27	

2 Entry composition

There are 23 unique types of molecules in this entry. The entry contains 55032 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 16S rRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1515	Total	C	N	O	P	0	0	0
			32554	14490	6022	10527	1515			

- Molecule 2 is a RNA chain called tRNA fMET (unmodified bases).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	76	Total	C	N	O	P	0	0	0
			1624	723	295	530	76			

- Molecule 3 is a RNA chain called tRNA PHE (unmodified bases).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	76	Total	C	N	O	P	0	0	0
			1623	723	290	534	76			

- Molecule 4 is a protein called 30S ribosomal protein S2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	234	Total	C	N	O	S	0	0	0
			1900	1213	341	341	5			

- Molecule 5 is a protein called 30S ribosomal protein S3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	206	Total	C	N	O	S	0	0	0
			1612	1016	314	281	1			

- Molecule 6 is a protein called 30S ribosomal protein S4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	G	208	Total	C	N	O	S	0	0	0
			1703	1066	339	291	7			

- Molecule 7 is a protein called 30S ribosomal protein S5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	H	150	Total	C	N	O	S	0	0	0
			1146	724	217	201	4			

- Molecule 8 is a protein called 30S ribosomal protein S6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	I	101	Total	C	N	O	S	0	0	0
			843	531	155	154	3			

- Molecule 9 is a protein called 30S ribosomal protein S7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	J	155	Total	C	N	O	S	0	0	0
			1257	781	252	218	6			

- Molecule 10 is a protein called 30S ribosomal protein S8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	K	138	Total	C	N	O	S	0	0	0
			1116	705	215	193	3			

- Molecule 11 is a protein called 30S ribosomal protein S9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	L	127	Total	C	N	O	S	0	0	0
			1010	639	197	174				

- Molecule 12 is a protein called 30S ribosomal protein S10.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	M	98	Total	C	N	O	S	0	0	0
			794	499	156	138	1			

- Molecule 13 is a protein called 30S ribosomal protein S11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	N	119	Total	C	N	O	S	0	0	0
			885	549	168	165	3			

- Molecule 14 is a protein called 30S ribosomal protein S12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	O	124	Total	C	N	O	S	0	0	0
			970	611	195	163	1			

- Molecule 15 is a protein called 30S ribosomal protein S13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	P	125	Total	C	N	O	S	0	0	0
			997	617	207	171	2			

- Molecule 16 is a protein called 30S ribosomal protein S14.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
16	Q	60	Total	C	N	O	S	0	0	0
			492	312	104	72	4			

- Molecule 17 is a protein called 30S ribosomal protein S15.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
17	R	88	Total	C	N	O	S	0	0	0
			734	459	147	126	2			

- Molecule 18 is a protein called 30S ribosomal protein S16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
18	S	83	Total	C	N	O	S	0	0	0
			700	443	139	117	1			

- Molecule 19 is a protein called 30S ribosomal protein S17.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
19	T	104	Total	C	N	O	S	0	0	0
			857	547	161	147	2			

- Molecule 20 is a protein called 30S ribosomal protein S18.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
20	U	73	Total	C	N	O	S	0	0	0
			597	380	118	99				

- Molecule 21 is a protein called 30S ribosomal protein S19.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
21	V	80	Total	C	N	O	S	0	0	0
			647	414	119	112	2			

- Molecule 22 is a protein called 30S ribosomal protein S20.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
22	W	99	Total	C	N	O	S	0	0	0
			763	470	162	129	2			

- Molecule 23 is a protein called 30S ribosomal protein Thx.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
23	X	24	Total	C	N	O	0	0	0
			208	128	50	30			

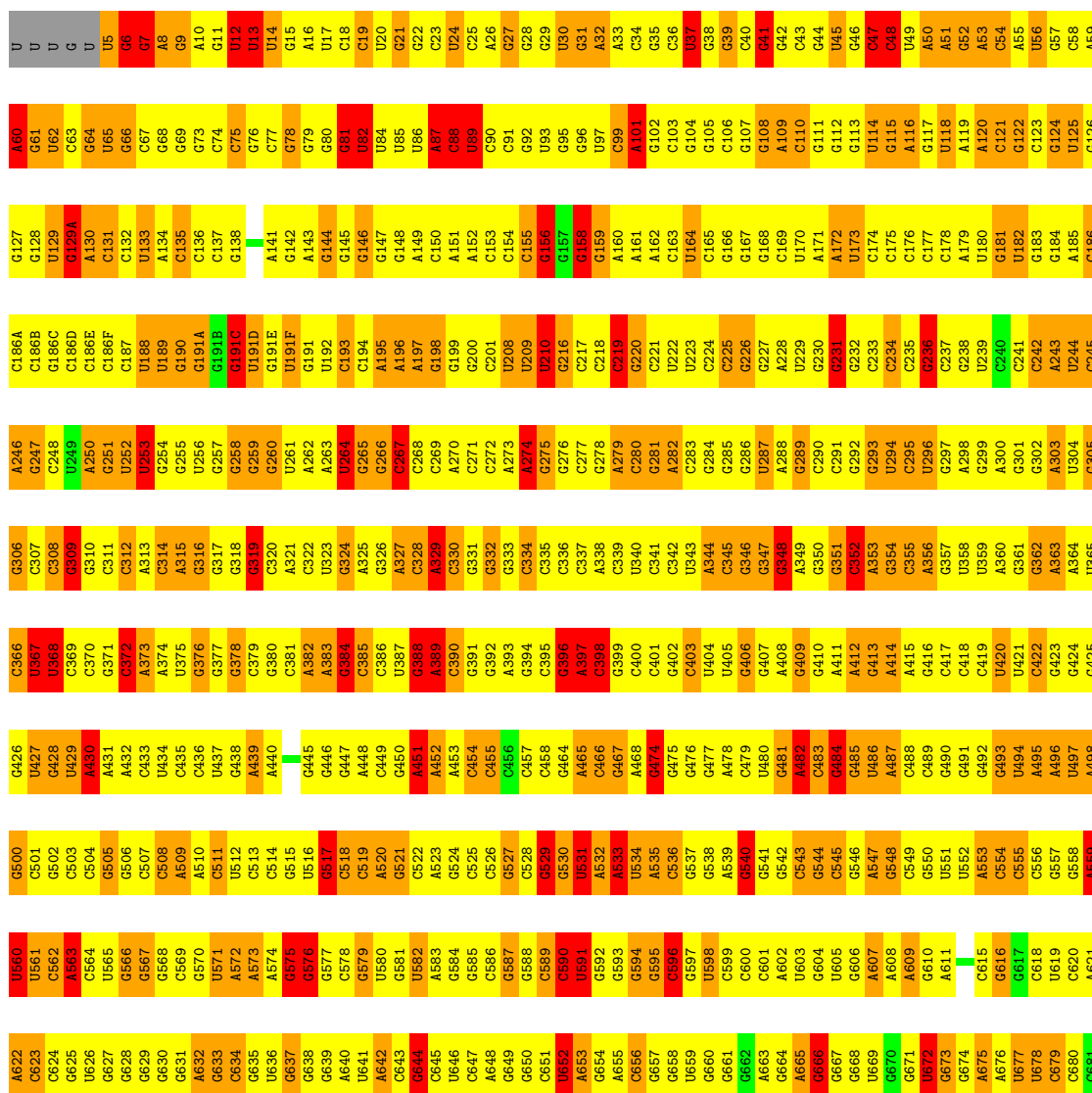
3 Residue-property plots

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: 16S rRNA

Chain A: 

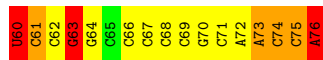


G1529	G1469	C1404	U1345	A1285	A1225	G1164	G1104	A1044	C989	G929	G869	A802	G742	G682
G1530	G1470	G1405	A1346	A1286	C1226	C1165	A1105	C1045	C990	C930	U870	G803	U743	G683
A1531	G1471	G1406	G1347	A1287	C1227	G1166	G1106	A1046	U991	C931	U871	U904	C744	A684
U1532	A1472	U1407	U1348	A1288	C1228	A1167	C1107	G1047	U992	C932	A872		C745	G685
C1533	A1473	A1408	A1349	A1289	C1229	A1169	G1108	G1048	G993	G933	A873	A807	U746	U686
A1534	G1474	C1409	A1350	G1290	C1230	A1170	C1109	U1049	A994	C934	G874	C808	A747	A687
C1535	G1475	G1410	U1351	G1291	G1231	G1171	A1110	G1050	C995	A935	C875	G803	C748	G688
U1536	G1476	C1411	C1352	U1292	U1232	C1172	A1111	C1051	A996	C936	C876	C810	C749	C689
C1537	G1477	G1412	G1353	G1293	G1233	C1173	C1112	U1052	U997	A937	C877	C811	G750	G690
C1538	C1478	A1413	C1354	G1294	C1234	G1174	C1113	G1053	G998	A938	G878	C812	U751	G691
C1539	C1479	U1414	G1355	G1295	U1235	G1175	C1114	C1054	C998A	G939	C879	U813	G752	U692
U1540	G1480	G1415	G1356	C1296	A1236	A1176	C1115	A1055	U999	C940	C880	A814	C753	G693
U1541	U1481	G1416	A1357	C1297	C1237	G1177	G1116	U1056	A1000	C941	G881	A815	C754	A694
U1542	G1482	G1417	U1358	C1298	A1238	G1178	C1117	G1057	G1001	G942	C882	A816	G755	A695
C	A1483	A1418	C1359	A1299	U1240	A1179	C1118	G1058	G1002	U943	C883	C817	C756	A696
U	C1484	G1419	A1360	G1300	U1241	A1180	C1119	C1059	U999	G944	U884	G818	U757	U697
	U1485	C1420	G1361	U1301	G1241	G1181	G1120	C1060	A885	A945	G885	A819	G758	G698
	G1486	G1421	C1362	U1302	C1242	G1182	U1121	G1061	A1005	A946	G886	U820	A759	G699
	G1487	G1422	C1363	A1303	C1243	A1183	U1122	U1062	C1006	G947	G887	G821	G760	G700
	G1488	G1423	A1363	G1304	C1244	A1184	A1123	C1063	C1007	C948	G888	C822	G761	G701
	G1489	C1424	U1364	G1305	C1245	G1185	G1124	C1064	A889	C949	A889	G823	C762	A702
	U1490	U1425	G1365	A1306	C1246	G1186	U1125	U1065	G1009	U950	G890	G824	G763	G703
	G1491	C1426	C1366	U1307	U1247	G1187	U1126	C1066	G1010	G951	U891	G825	C764	A704
	A1492	U1427	C1367	U1308	A1248	A1188	G1127	A1067	G1011	U952	A892	G826	G765	U705
	A1493	A1428	G1368	G1309	G1249	C1189	C1128	G1068	U1012	G953	C893	U827	A766	A706
	G1494	C1429	C1369	G1310	A1250	G1190	C1129	C1069	G1013	G954	C894	A828	A767	G707
	U1495	C1430	G1370	G1311	A1251	A1191	A1130	U1070	A1014	U955	G895	G829	A768	C708
	C1496	G1431	C1371	G1312	A1252	C1192	G1131	C1071	A1015	U956	C896	G830	G769	G709
	G1497	G1432	C1372	U1313	C1253	G1193	G1132	G1072	A1016	U957	C897	C832	G770	G710
	U1498	A1433	G1373	C1314	C1254	U1194	C1133	G1073	G1017	A958	G898	G833	G771	G711
	A1499	A1434	A1374	U1315	G1255	C1195	G1134	G1074	C1018	A959	C899	U833	U772	A712
	A1500	G1435	A1375	G1316	A1256	U1196	U1135	C1075	C1019	U960	A900	C834	G773	G713
	C1501	U1436	U1376	C1317	U1257	G1197	U1136	C1076	U1020	U961	A901	U835	G774	G714
	A1502	C1437	A1377	A1318	G1258	G1198	C1137	G1077	G1021	C962	G902	G836	G775	A715
	C1503	G1438	C1378	A1319	C1259	U1199	G1138	U1078	G1022	A964	C904	G837	G776	A716
	G1504	C1439	G1379	C1320	A1261	C1200	G1139	G1079	G1023	A965	U905	U841	A777	G717
	U1505	U1440	U1380	G1321	C1262	A1201	C1140	A1080	G1024	G966	G906	C842	C778	C719
	U1506	G1441	U1381	C1322	C1263	G1202	C1141	G1081	U1025	C967	G907	C843	C779	C720
	A1507	G1442	C1382	G1323	C1264	A1204	G1142	G1082	C1027	C968	A908	U843	A780	G721
	G1508	G1443	C1383	A1324	G1265	U1205	G1143	U1083	C1028	A969	C909	C848	A781	G722
	C1509	A1446	G1384	C1325	G1266	G1206	G1144	G1084	C1028A	C970	C910	U850	A782	U723
	U1510	G1447	G1385	C1326	C1267	G1207	C1145	U1085	G1028B	G971	U911	G851	C783	G724
	G1511	C1448	G1386	C1327	C1268	G1208	C1146	U1086	C1029	C972	C912	G852	C784	G725
	U1512	U1449	G1387	C1328	A1268	C1209	C1147	G1087	G1029	G973	A913	G853	G785	G726
	A1513	U1450	C1388	A1329	A1269	C1210	U1148	G1088	C1030	G974	A914	G854	G786	C727
	C1514	A1451	C1389	U1330	G1270	C1211	C1149	U1089	G1031	A975	A915	G855	A787	A728
	U1515	C1452	U1390	G1331	G1271	U1211	U1150	U1090	A1032	G976	G916	C856	U788	G729
	G1516	G1453	U1391	A1332	G1272	U1212	A1151	U1091	G1032A	A977	G917	C857	A789	G730
	U1517	G1454	G1392	A1333	G1273	A1213	A1152	A1092	G1032B	A978	A918	G858	G791	G731
	A1518	G1455	U1393	G1334	G1274	C1214	C1153	A1093	G1033	C979	A919	A859	A792	C732
	U1519	C1456	A1394	C1335	A1275	G1215	G1154	G1094	G1034	C980	U920	A960	U793	A733
	G1520	A1460	C1395	C1336	G1276	G1216	U1155	U1095	A1035	U981	U921	G861	A794	G734
	U1521	G1461	A1396	G1337	C1277	C1217	G1156	C1096	G1036	U982	G922	C862	C795	C735
	C1522	G1462	C1397	G1338	U1278	C1218	A1157	C1097	G1037	A983	G923	U863	C796	G736
	G1523	C1463	A1398	A1339	A1279	U1219	C1158	C1098	G1038	C984	C924	A864	C797	A737
	U1524	G1464	C1399	A1340	U1280	G1220	U1159	G1099	C1039	C985	G925	A865	G798	C738
	G1525	C1465	C1400	U1341	U1281	G1221	C1160	C1100	U1040	A986	G926	C866	G799	C739
	U1526	G1466	G1401	C1342	C1282	G1222	C1161	A1101	A1041	G927	G927	G867	G800	U740
	C1527	G1467	C1402	G1343	C1283	C1223	C1162	A1102	C1042	G988	G928	C868	U801	G741
	U1528	A1468	C1403	C1344	C1284	G1224	C1163	C1103	C1043					

• Molecule 2: tRNA fMET (unmodified bases)

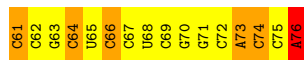
Chain C:





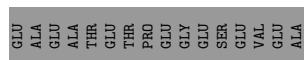
- Molecule 3: tRNA PHE (unmodified bases)

Chain D:



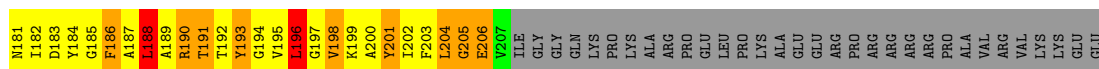
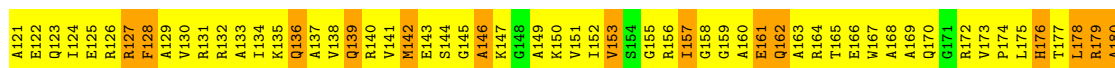
- Molecule 4: 30S ribosomal protein S2

Chain E:



- Molecule 5: 30S ribosomal protein S3

Chain F:



- Molecule 6: 30S ribosomal protein S4

Chain G:

V61 V62 L63 K64 V65 G66 P67 R68 R69 Q70 G71 P72 D73 P74 R75 P76 E77 Q78 V79 L80 H81 H82 H83 L84 R85 L86 S87 K88 P89 G90 G91 R91 R92 V93 V94 V95 G96 V97 V98 E99 I100 P101 R102 R103 R104 R105 G106 L107 G108 I109 A110 I111 I112 S113 T114 S115 K116 G117 V118 T120

D121 R122 E123 A124 R125 K126 L127 G128 V129 G130 G131 E132 L133 I134 C135 E136 V137 W138

• Molecule 11: 30S ribosomal protein S9

Chain L:

MET E2 Q3 Y4 Y5 G6 T7 G8 R9 R10 K11 E12 A13 V14 A15 R16 V17 F18 F19 R20 P21 P22 N23 G24 K25 V26 T27 V28 N29 G30 G31 D32 P33 N34 E35 Y36 Y37 Q38 Q39 L40 L41 R42 R43 A44 A45 A46 L47 E48 P49 L50 R51 A52 V53 D54 A55 L56 G57 H58 F59 D60

A61 Y62 T63 T64 Y65 R66 G67 G68 G69 R70 R71 S71 G72 G73 I74 D75 A76 I77 K78 L79 G80 G81 H81 R82 R83 R84 L85 V86 Q87 R88 N89 P90 P91 D91 Y92 R93 A94 K95 L96 K97 P98 L99 G100 F101 L102 T103 G104 D105 A106 R107 V108 V109 E110 R111 K112 K113 Y114 G115 K116 R117 K118 A119 R120

R121 A122 P123 Q124 Y125 S126 K127 R128

• Molecule 12: 30S ribosomal protein S10

Chain M:

MET PR0 K3 I4 R5 I6 K7 L8 R9 G10 F11 D12 H13 K14 T15 L16 D17 A18 S19 A20 Q21 K22 K23 I23 Q24 E25 A26 A27 L28 R29 S30 G31 A32 Q33 V34 E35 G36 P37 I38 P39 L40 P41 T42 R43 V44 R45 R46 F47 T48 V49 I50 R51 G52 P53 F54 K55 H56 K57 D58 S59 R60

B61 B62 F63 B64 L65 R66 T67 H68 G69 R70 L71 V72 D73 I74 I75 N76 P77 N78 R79 K80 R81 T81 L82 E83 Q84 L85 L86 R87 L88 D89 L90 P91 T92 G93 V94 E95 L96 E97 I98 R99 T100 VAL GLY GLY ARG

• Molecule 13: 30S ribosomal protein S11

Chain N:

MET ALA LYS LYS PRO SER LYS LYS LYS VAL R12 Q13 V14 A15 S16 G17 R18 A19 Y20 I21 H22 A23 Q24 R25 N26 N27 T28 I29 V30 T31 I32 T33 D34 P35 D36 G37 N38 P39 T40 W41 W42 S43 S44 G45 G46 V47 I48 G49 Y50 K51 G52 S53 R54 K55 G56 T57 P58 S59 A60

A61 Q62 L63 A64 A65 L66 D67 A68 A69 K70 K71 A72 M73 A74 Y75 G76 T77 Q78 S79 V80 D81 V82 I83 V84 R85 G86 T87 G88 A89 R90 R91 E92 Q93 A94 I95 R96 R97 L98 Q99 A100 S101 G102 L103 Q104 V105 K106 S107 I108 V109 D110 D111 T112 P113 V114 P115 H116 N117 G118 C119 R120

P121 K122 K123 K124 L125 R126 K127 A128 S129

• Molecule 14: 30S ribosomal protein S12

Chain O:

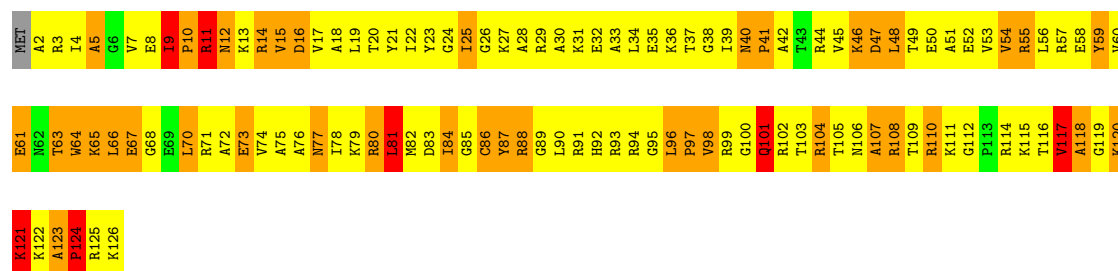
MET P5 T6 I7 N8 Q9 L10 V11 R12 R15 E16 K17 V18 R19 K20 K21 S22 K23 V24 P25 A26 L27 L28 G29 A30 P31 F32 R33 R34 G35 V36 G37 V38 V39 V40 R41 T42 V43 P44 P45 K46 K47 R48 M49 S50 A51 L52 R53 K54 V55 V56 S57 K58 K59 R60 L60 T61 S62 K63 G64

B65 V66 T67 A68 L69 I70 P71 G72 E73 G74 H75 N76 L77 Q78 E79 H80 S81 V82 R83 L84 L85 R86 G87 G88 R89 V90 K91 L93 V96 R97 H99 I100 R101 R102 G103 V104 Y105 D106 A107 A108 G109 V110 K111 D112 R113 K114 K115 S116 R117 S118 K119 Y120 G121 T122 K123 K124 P125

K126 E127 A128 ALA THR ALA LYS LYS

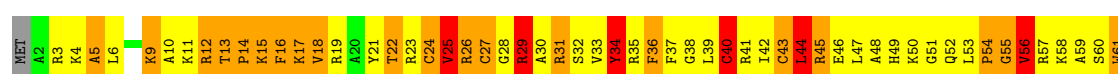
- Molecule 15: 30S ribosomal protein S13

Chain P:



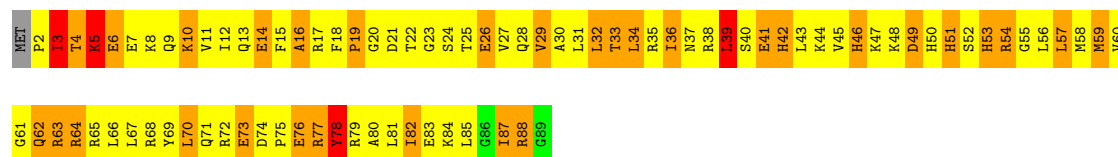
- Molecule 16: 30S ribosomal protein S14

Chain Q:



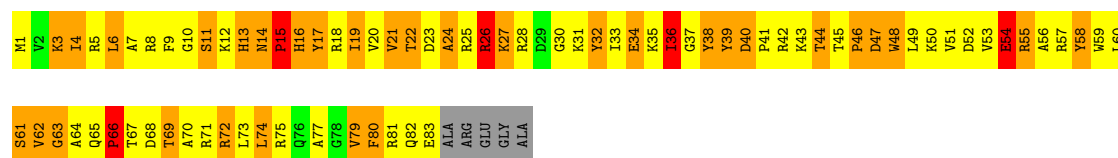
- Molecule 17: 30S ribosomal protein S15

Chain R:



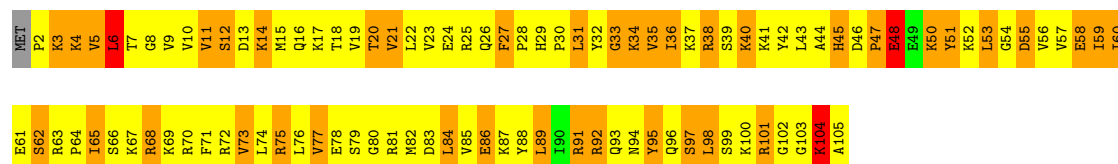
- Molecule 18: 30S ribosomal protein S16

Chain S:



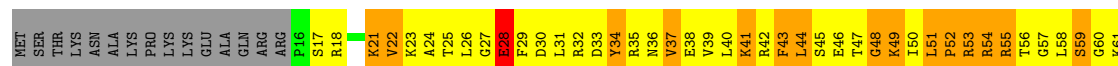
- Molecule 19: 30S ribosomal protein S17

Chain T:



- Molecule 20: 30S ribosomal protein S18

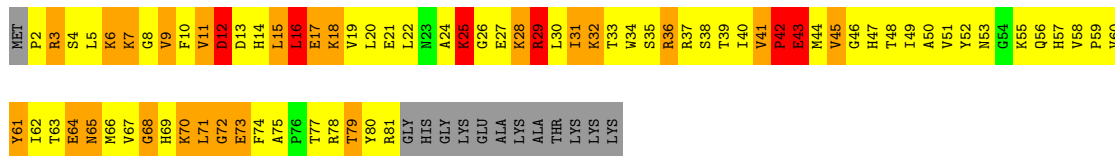
Chain U:





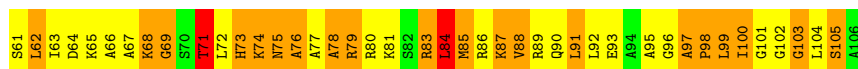
- Molecule 21: 30S ribosomal protein S19

Chain V:



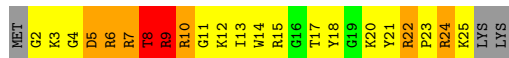
- Molecule 22: 30S ribosomal protein S20

Chain W:



- Molecule 23: 30S ribosomal protein Thx

Chain X:



4 Data and refinement statistics

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

Property	Value	Source
Space group	I 4 2 2	Depositor
Cell constants a, b, c, α , β , γ	509.52Å 509.52Å 804.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	300.00 – 5.00	Depositor
% Data completeness (in resolution range)	93.9 (300.00-5.00)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.09 (at 4.88Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.263 , 0.323	Depositor
Wilson B-factor (Å ²)	99.7	Xtriage
Anisotropy	0.066	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.55$, $\langle L^2 \rangle = 0.39$	Xtriage
Outliers	1 of 252174 reflections (0.000%)	Xtriage
Total number of atoms	55032	wwPDB-VP
Average B, all atoms (Å ²)	169.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.30% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# $ Z > 5$	RMSZ	# $ Z > 5$
1	A	0.94	48/36438 (0.1%)	1.12	150/56869 (0.3%)
2	C	1.00	1/1814 (0.1%)	0.92	3/2825 (0.1%)
3	D	0.80	1/1813 (0.1%)	0.94	3/2823 (0.1%)
4	E	0.65	0/1935	1.02	3/2609 (0.1%)
5	F	0.55	0/1636	0.98	4/2205 (0.2%)
6	G	0.68	1/1733 (0.1%)	1.07	5/2318 (0.2%)
7	H	0.68	0/1162	1.06	2/1564 (0.1%)
8	I	0.68	0/856	1.01	3/1154 (0.3%)
9	J	0.60	0/1276	0.90	0/1709
10	K	0.68	0/1136	1.10	5/1527 (0.3%)
11	L	0.54	0/1029	0.87	0/1379
12	M	0.49	0/807	0.88	0/1085
13	N	0.73	0/900	1.06	3/1213 (0.2%)
14	O	0.67	0/986	1.12	6/1320 (0.5%)
15	P	0.53	0/1008	0.90	2/1347 (0.1%)
16	Q	0.57	0/501	1.01	1/664 (0.2%)
17	R	0.71	0/745	1.00	1/992 (0.1%)
18	S	0.72	0/716	1.05	3/963 (0.3%)
19	T	0.73	0/870	1.07	2/1159 (0.2%)
20	U	0.64	0/603	1.02	3/799 (0.4%)
21	V	0.55	0/661	0.94	0/890
22	W	0.77	0/765	1.03	1/1007 (0.1%)
23	X	0.48	0/212	0.85	1/277 (0.4%)
All	All	0.85	51/59602 (0.1%)	1.07	201/88698 (0.2%)

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	217
2	C	0	13

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
3	D	0	7
5	F	0	1
10	K	0	1
13	N	0	1
All	All	0	240

The worst 5 of 51 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	173	U	O3'-P	19.93	1.85	1.61
1	A	397	A	C6-N6	17.27	1.47	1.33
1	A	37	U	N1-C2	15.22	1.52	1.38
1	A	37	U	C4-O4	-14.43	1.12	1.23
1	A	37	U	N3-C4	-12.52	1.27	1.38

The worst 5 of 201 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1064	G	N1-C2-N2	-70.70	52.57	116.20
1	A	1064	G	N3-C2-N2	58.69	160.98	119.90
1	A	1064	G	N1-C2-N3	-25.80	108.42	123.90
1	A	37	U	N3-C4-O4	-19.88	105.48	119.40
1	A	1064	G	C2-N3-C4	15.33	119.56	111.90

There are no chirality outliers.

5 of 240 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	12	U	Sidechain
1	A	21	G	Sidechain
1	A	24	U	Sidechain
1	A	6	G	Sidechain
1	A	7	G	Sidechain

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	32554	0	16429	6436	0
2	C	1624	0	826	326	0
3	D	1623	0	821	338	0
4	E	1900	0	1951	934	0
5	F	1612	0	1677	695	0
6	G	1703	0	1763	800	0
7	H	1146	0	1207	493	0
8	I	843	0	857	370	0
9	J	1257	0	1296	561	0
10	K	1116	0	1177	644	0
11	L	1010	0	1037	460	0
12	M	794	0	840	363	0
13	N	885	0	904	407	0
14	O	970	0	1057	440	0
15	P	997	0	1072	482	0
16	Q	492	0	529	255	0
17	R	734	0	771	297	0
18	S	700	0	720	308	0
19	T	857	0	930	415	0
20	U	597	0	668	331	0
21	V	647	0	673	300	0
22	W	763	0	861	356	0
23	X	208	0	221	83	0
All	All	55032	0	38287	14770	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 159.

The worst 5 of 14770 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
19:T:5:VAL:HG13	19:T:59:ILE:O	1.26	1.35
1:A:559:A:H4'	1:A:560:U:C5'	1.57	1.34
1:A:197:A:N6	1:A:221:C:H5'	1.42	1.30
6:G:31:CYS:SG	6:G:33:MET:SD	2.29	1.29
6:G:59:ARG:HA	6:G:59:ARG:CZ	1.64	1.26

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	
4	E	232/256 (91%)	119 (51%)	48 (21%)	65 (28%)	0	1
5	F	204/239 (85%)	117 (57%)	41 (20%)	46 (22%)	0	3
6	G	206/209 (99%)	99 (48%)	52 (25%)	55 (27%)	0	2
7	H	148/162 (91%)	89 (60%)	26 (18%)	33 (22%)	0	3
8	I	99/101 (98%)	62 (63%)	14 (14%)	23 (23%)	0	3
9	J	153/156 (98%)	65 (42%)	45 (29%)	43 (28%)	0	1
10	K	136/138 (99%)	71 (52%)	28 (21%)	37 (27%)	0	1
11	L	125/128 (98%)	57 (46%)	38 (30%)	30 (24%)	0	2
12	M	96/105 (91%)	53 (55%)	21 (22%)	22 (23%)	0	3
13	N	117/129 (91%)	63 (54%)	28 (24%)	26 (22%)	0	3
14	O	122/132 (92%)	60 (49%)	29 (24%)	33 (27%)	0	2
15	P	123/126 (98%)	64 (52%)	33 (27%)	26 (21%)	0	4
16	Q	58/61 (95%)	24 (41%)	16 (28%)	18 (31%)	0	1
17	R	86/89 (97%)	43 (50%)	26 (30%)	17 (20%)	0	4
18	S	81/88 (92%)	46 (57%)	22 (27%)	13 (16%)	0	8
19	T	102/105 (97%)	63 (62%)	20 (20%)	19 (19%)	0	5
20	U	71/88 (81%)	29 (41%)	23 (32%)	19 (27%)	0	2
21	V	78/93 (84%)	32 (41%)	24 (31%)	22 (28%)	0	1
22	W	97/106 (92%)	31 (32%)	43 (44%)	23 (24%)	0	3
23	X	22/27 (82%)	11 (50%)	5 (23%)	6 (27%)	0	1
All	All	2356/2538 (93%)	1198 (51%)	582 (25%)	576 (24%)	0	2

5 of 576 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	E	13	ALA
4	E	15	VAL
4	E	17	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	E	20	GLU
4	E	42	ILE

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	
4	E	202/220 (92%)	142 (70%)	60 (30%)	0	5
5	F	160/188 (85%)	122 (76%)	38 (24%)	1	9
6	G	180/181 (99%)	130 (72%)	50 (28%)	0	6
7	H	115/123 (94%)	70 (61%)	45 (39%)	0	1
8	I	90/90 (100%)	59 (66%)	31 (34%)	0	3
9	J	126/127 (99%)	97 (77%)	29 (23%)	1	10
10	K	119/119 (100%)	76 (64%)	43 (36%)	0	2
11	L	98/99 (99%)	74 (76%)	24 (24%)	1	9
12	M	88/92 (96%)	61 (69%)	27 (31%)	0	5
13	N	90/99 (91%)	70 (78%)	20 (22%)	1	11
14	O	104/109 (95%)	79 (76%)	25 (24%)	1	9
15	P	100/101 (99%)	73 (73%)	27 (27%)	1	6
16	Q	49/50 (98%)	36 (74%)	13 (26%)	1	7
17	R	79/80 (99%)	56 (71%)	23 (29%)	0	5
18	S	72/74 (97%)	45 (62%)	27 (38%)	0	1
19	T	96/97 (99%)	69 (72%)	27 (28%)	0	5
20	U	64/77 (83%)	50 (78%)	14 (22%)	1	11
21	V	71/80 (89%)	57 (80%)	14 (20%)	2	15
22	W	76/82 (93%)	57 (75%)	19 (25%)	1	8
23	X	19/22 (86%)	16 (84%)	3 (16%)	4	27
All	All	1998/2110 (95%)	1439 (72%)	559 (28%)	0	6

5 of 559 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	J	155	ARG
11	L	91	ASP
20	U	63	GLN
10	K	24	THR
10	K	98	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 83 such sidechains are listed below:

Mol	Chain	Res	Type
9	J	96	GLN
11	L	23	ASN
19	T	93	GLN
9	J	97	GLN
9	J	153	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1515/1522 (99%)	430 (28%)	157 (10%)
2	C	75/77 (97%)	29 (38%)	6 (8%)
3	D	75/76 (98%)	27 (36%)	4 (5%)
All	All	1665/1675 (99%)	486 (29%)	167 (10%)

5 of 486 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	6	G
1	A	8	A
1	A	9	G
1	A	13	U
1	A	14	U

5 of 167 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	703	G
1	A	872	A
1	A	1532	U
1	A	721	G
1	A	812	C

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

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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.