



# wwPDB X-ray Structure Validation Summary Report

Sep 15, 2014 – 07:51 PM EDT

PDB ID : 1VY0  
Title : Crystal Structure of Unmodified tRNA Proline (CGG) Bound to Codon CCG on the Ribosome  
Authors : Maehigashi, T.; Dunkle, J.A.; Dunham, C.M.  
Deposited on : 2014-03-25  
Resolution : 3.68 Å(reported)

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

---

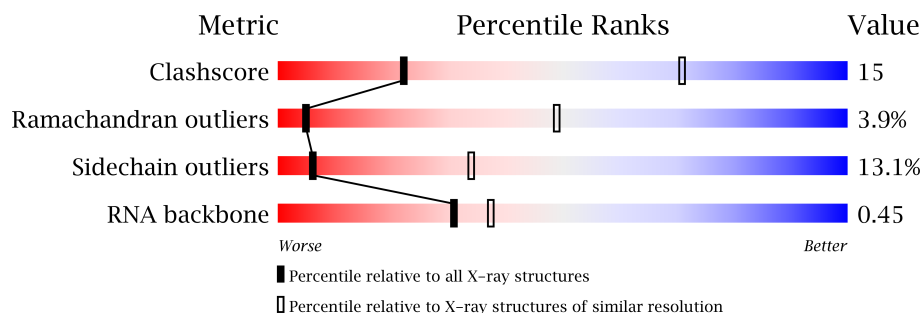
The following versions of software and data (see [references](#)) were used in the production of this report:

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

# 1 Overall quality at a glance

The reported resolution of this entry is 3.68 Å.

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



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	79885	1006 (3.92-3.44)
Ramachandran outliers	78287	1218 (3.96-3.40)
Sidechain outliers	78261	1216 (3.96-3.40)
RNA backbone	1838	1008 (4.52-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	1522	
2	B	256	
3	C	239	
4	D	209	
5	E	162	
6	F	101	
7	G	156	
8	H	138	
9	I	128	
10	J	105	
11	K	129	
12	L	132	
13	M	126	
14	N	61	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
15	O	89	<div><div></div></div>
16	P	88	<div><div></div></div>
17	Q	105	<div><div></div></div>
18	R	88	<div><div></div></div>
19	S	93	<div><div></div></div>
20	T	106	<div><div></div></div>
21	U	27	<div><div></div></div>
22	V	77	<div><div></div></div>
23	X	25	<div><div></div></div>
24	Y	17	<div><div></div></div>

## 2 Entry composition

There are 27 unique types of molecules in this entry. The entry contains 53592 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	1500	Total	C	N	O	P	0	0	0
			32247	14353	5981	10414	1499			

- Molecule 2 is a protein called 30S ribosomal protein S2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	237	Total	C	N	O	S	0	0	0
			1924	1228	344	347	5			

- Molecule 3 is a protein called 30S ribosomal protein S3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	205	Total	C	N	O	S	0	0	0
			1605	1011	313	280	1			

- Molecule 4 is a protein called 30S ribosomal protein S4.

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

- Molecule 5 is a protein called 30S ribosomal protein S5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	151	Total	C	N	O	S	0	0	0
			1155	729	218	204	4			

- Molecule 6 is a protein called 30S ribosomal protein S6.

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

- Molecule 7 is a protein called 30S ribosomal protein S7.

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

- Molecule 8 is a protein called 30S ribosomal protein S8.

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

- Molecule 9 is a protein called 30S ribosomal protein S9.

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

- Molecule 10 is a protein called 30S ribosomal protein S10.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	99	Total	C	N	O	S	0	0	0
			801	504	157	139	1			

- Molecule 11 is a protein called 30S ribosomal protein S11.

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

- Molecule 12 is a protein called 30S ribosomal protein S12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	125	Total	C	N	O	S	0	0	0
			975	614	196	164	1			

- Molecule 13 is a protein called 30S ribosomal protein S13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	121	Total	C	N	O	S	0	0	0
			964	597	199	166	2			

- Molecule 14 is a protein called 30S ribosomal protein S14 type Z.

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

- Molecule 15 is a protein called 30S ribosomal protein S15.

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

- Molecule 16 is a protein called 30S ribosomal protein S16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
16	P	84	Total	C	N	O	S	0	0	0
			705	446	140	118	1			

- Molecule 17 is a protein called 30S ribosomal protein S17.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
17	Q	100	Total	C	N	O	S	0	0	0
			834	534	155	143	2			

- Molecule 18 is a protein called 30S ribosomal protein S18.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
18	R	70	Total	C	N	O	0	0	0
			574	367	112	95			

- Molecule 19 is a protein called 30S ribosomal protein S19.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
19	S	84	Total	C	N	O	S	0	0	0
			674	430	126	116	2			

- Molecule 20 is a protein called 30S ribosomal protein S20.

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

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
21	U	25	Total	C	N	O	0	0	0
			217	134	52	31			

- Molecule 22 is a RNA chain called P-site tRNA fMET.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
22	V	77	Total	C	N	O	P	0	0	0
			1644	732	297	538	77			

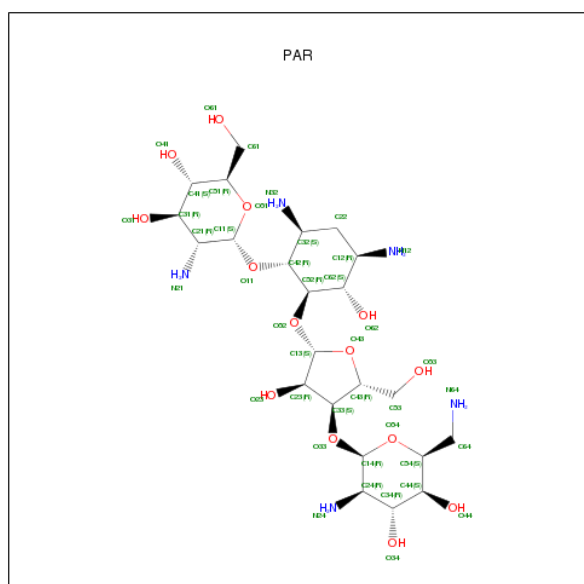
- Molecule 23 is a RNA chain called mRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
23	X	8	Total	C	N	O	P	0	0	0
			173	77	33	55	8			

- Molecule 24 is a RNA chain called A site ASL of tRNA-Proline CGG (unmodified).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
24	Y	8	Total	C	N	O	P	0	0	0
			174	77	33	56	8			

- Molecule 25 is PAROMOMYCIN (three-letter code: PAR) (formula:  $C_{23}H_{45}N_5O_{14}$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
25	A	1	Total	C	N	O	0	0
			42	23	5	14		

- Molecule 26 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
26	A	76	Total 76	Mg 76	0	0
26	V	1	Total 1	Mg 1	0	0
26	F	1	Total 1	Mg 1	0	0
26	M	1	Total 1	Mg 1	0	0

- Molecule 27 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
27	D	1	Total 1	Zn 1	0	0
27	N	1	Total 1	Zn 1	0	0





C	C1466	C1378	U1313	C1243	G1175	A1105	C1030	A969	A901	C824
U	G1467	G1379	C1314	C1244	A1176	C1106	G1031	C970	G971	G825
U	A1468	U1315	U1315	A1245	G1177	C1107	A1032	G972	G903	C826
C	G1469	G1316	C1316	A1251	G1178	C1108	G1032A	G973	U905	U827
U	C1387	C1317	C1317	A1252	A1179	C1109	G1036	G974	G906	C829
U	C1388	C1318	C1318		A1180	A1110		A975		G829
	C1389	A1319	A1319		G1181	A1111	C1037		A909	G830
	C1390	C1320	C1320	G1255	G1182	C1112	C1038	G976	C910	U831
	U1391	C1321	C1321	A1256	A1183		U1039	A977	G911	C832
	C1392	C1322	C1322	U1257	G1184	G1118	C1040	A978	U911	U833
	C1478	G1393	G1323	G1258	G1185	C1119	A1041	C979	C912	
		A1394		C1259	G1186	G1120	G1042	C980	A913	G836
	G1488	C1395	C1326	C1260	G1187	U1121	C1043	U981	A914	
	G1489	A1396	C1327			U1122		U982		U841
	C1490	C1397	C1328	C1263	G1190	A1123	A1046	A983	G917	C842
	A1491	A1398		C1264	A1191	G1124		C984	A918	U843
	A1492	C1399	G1331	G1265	C1192	U1125	U1049	C985	A919	C848
	A1493	C1400	A1332	G1266	G1193	U1126	G1056	A986	U920	C849
	G1494	G1401	A1333	C1267	U1194	C1127	C1051	G987	U921	U850
	A1495	C1402	G1334	A1268	C1195	G1128	U1052	C988	G922	G851
		C1403	C1335	A1269	U1196	C1129	G1053	C989	A923	G852
	U1498	C1404	G1336	C1270	G1197	A1130	C1054	C990	C924	G853
	A1499	G1405	G1337	G1271	G1198	C1131	A1055	U991	G925	G854
	A1500		G1338	G1272	U1199	G1132	U1056	U992	G926	G855
	C1501	G1410	A1339	G1273	C1200	G1133	C1057	G993	G927	C856
	A1502	C1411		G1274	A1201	G1134	C1058	A994	G928	
	A1503	C1412	C1342	A1275	G1202	U1135	C1059	C995		A859
	G1504	A1413	G1343	G1276	C1203	U1136	C1060	A996	C932	A860
	G1505		C1344		A1204	C1137	G1061	U997	C933	
	U1506	G1419	U1345	A1279	U1205	G1138		C998	C934	A864
			A1346	U1280	G1206	C1139	U1065	C998A	A935	A865
	C1509	G1423	C1347	U1281	G1207	C1140	C1066	U999	C936	
	U1510	C1424	U1348	C1282	C1208	C1141	A1067	A1000	A937	U870
	G1511	U1425	A1349			G1142	G1068	G1001	A938	U871
	U1512	C1426	A1350	A1285	U1212	G1143	C1069	G1002	C939	A872
	A1513	U1427	U1351	A1286	A1213	C1144	U1070	G1003	C940	A873
		A1428	C1352	A1287	C1214	C1145	C1071	A1004		G874
	C1514	C1429	G1353	A1288	G1215	A1146	G1072	A1005	U943	
			C1354	A1289	G1216	C1147	U1073	C1006	G944	C877
	G1516	A1434	G1355	G1290	C1217	U1150	G1077	G1009	G945	G878
	U1517	G1435	G1356	G1291	C1218	A1151	U1078	G1010	A946	C879
	A1518	U1436	A1357	U1292	U1219	C1152	G1079	G1011	C947	C880
	U1519	C1437	U1358	G1293	G1220	A1153	A1080	U1012	C948	
	G1520	G1438		G1294	G1221	C1153	G1081	G1013	A949	C882
	G1521				G1222	A1157	G1082	A1014	G953	C883
	U1522		C1362A	C1297	C1223	G1158	U1083	A1015	G954	U884
	G1523	G1442	A1363	A1298	G1224	C1159		A1016	U955	G885
	C1524	G1443	U1364	C1299	A1225	U1159	U1086		U956	G886
		A1446	G1365	A1299	G1226	G1160				
		C1447	C1366	U1301	A1227	C1161		U1020	G988	G887
	C1527	C1448	C1367	U1302		C1162	G1094	G1021	A989	A889
	U1528	C1449	G1368	C1303	G1233	C1163		G1022	U960	G890
	G1529	U1450	A1369	G1304	C1234		C1096	G1023	U961	U891
	A1530	A1451	G1370	C1305	C1234	A1167	C1097	G1024	C982	A892
	U	C1452	G1371	G1306		C1237	C1098	G1025	G963	C893
	A	G1453	U1372	A1306	U1307	A1169	U1099	G1026	A964	G894
	C	G1454	C1373	U1307		A1170	C1100	C1027	G965	C895
	C	G1455	A1374	A1374	U1239	G1171	A1101	C1028		
	C	C1456	A1375	G1310	U1240	C1172				
	U		C1376	G1311	G1241	G1173				C899
	C	C1458	U1377	C1312	C1242	C1174				A900

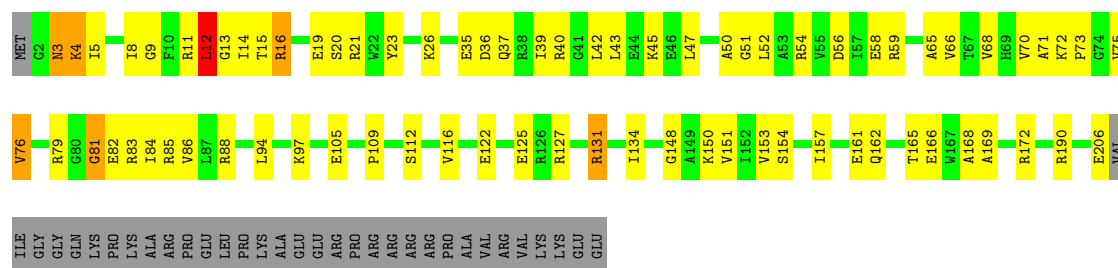
• Molecule 2: 30S ribosomal protein S2

Chain B:

ALA	THR	GLU	THR	PRO	GLU	GLY	GLU	SER	GLU	VAL	GLU	ALA	MET
	THR	GLU	THR	PRO	GLU	GLY	GLU	SER	GLU	VAL	GLU	ALA	

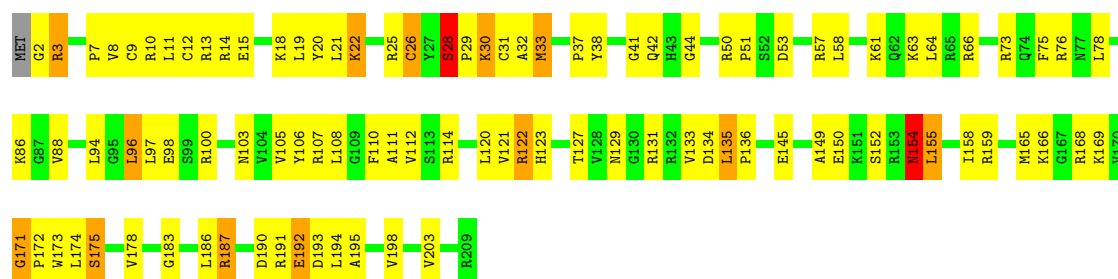
• Molecule 3: 30S ribosomal protein S3

## Chain C:



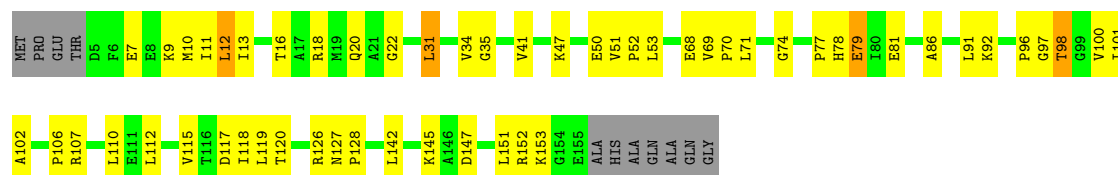
## • Molecule 4: 30S ribosomal protein S4

## Chain D:



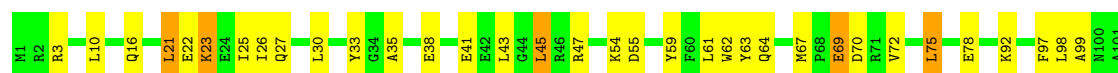
## • Molecule 5: 30S ribosomal protein S5

## Chain E:



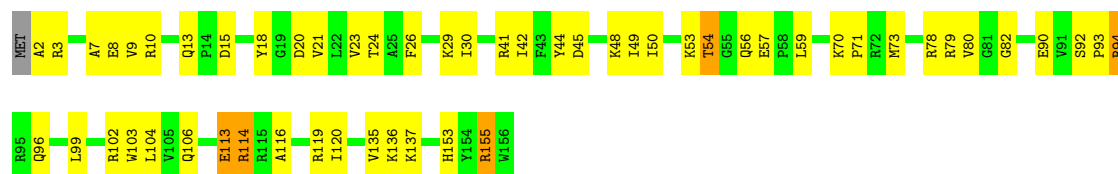
## • Molecule 6: 30S ribosomal protein S6

## Chain F:



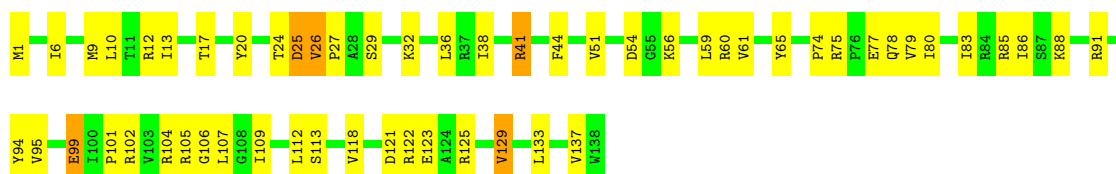
## • Molecule 7: 30S ribosomal protein S7

## Chain G:



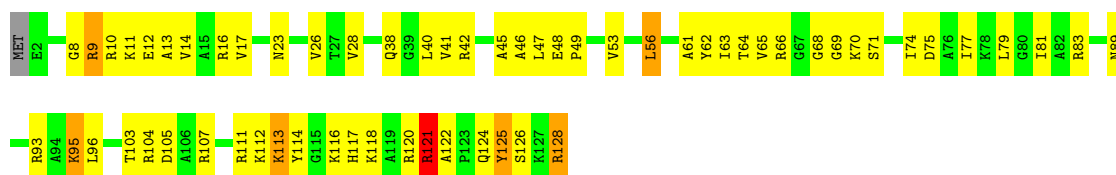
## • Molecule 8: 30S ribosomal protein S8

## Chain H:



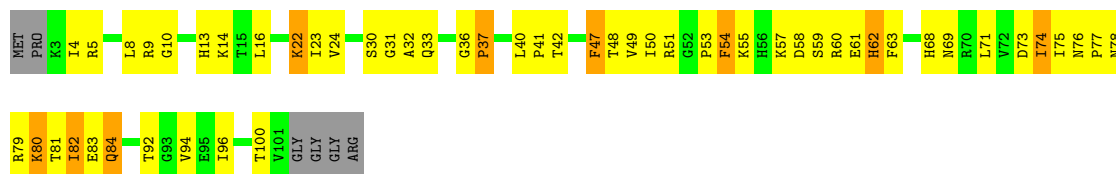
• Molecule 9: 30S ribosomal protein S9

Chain I:



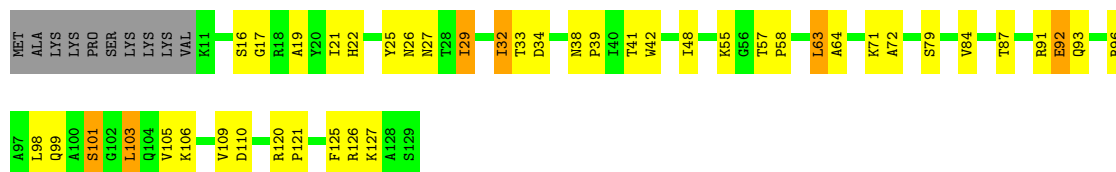
• Molecule 10: 30S ribosomal protein S10

Chain J:



• Molecule 11: 30S ribosomal protein S11

Chain K:



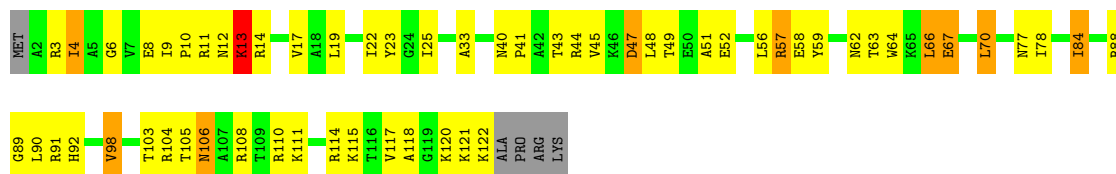
• Molecule 12: 30S ribosomal protein S12

Chain L:



• Molecule 13: 30S ribosomal protein S13

Chain M:



- Molecule 14: 30S ribosomal protein S14 type Z

Chain N:



- Molecule 15: 30S ribosomal protein S15

Chain O:



- Molecule 16: 30S ribosomal protein S16

Chain P:



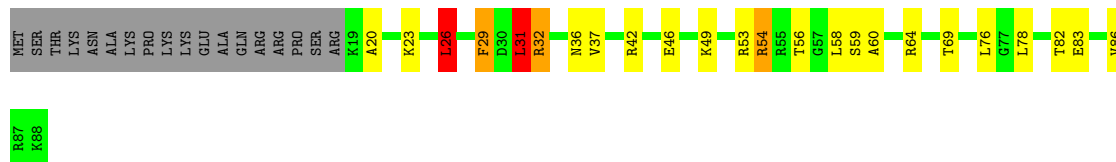
- Molecule 17: 30S ribosomal protein S17

Chain Q:



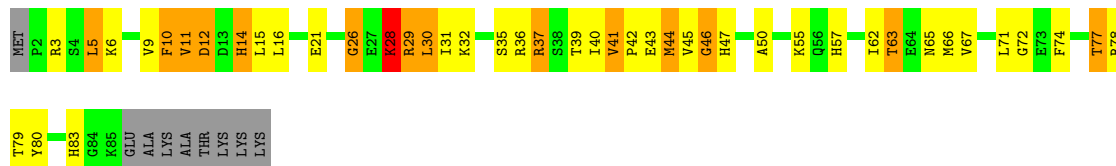
- Molecule 18: 30S ribosomal protein S18

Chain R:



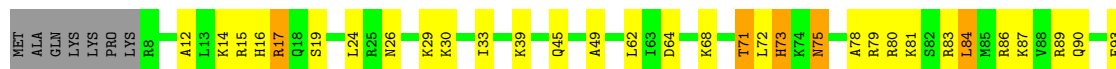
- Molecule 19: 30S ribosomal protein S19

Chain S:



- Molecule 20: 30S ribosomal protein S20

Chain T:





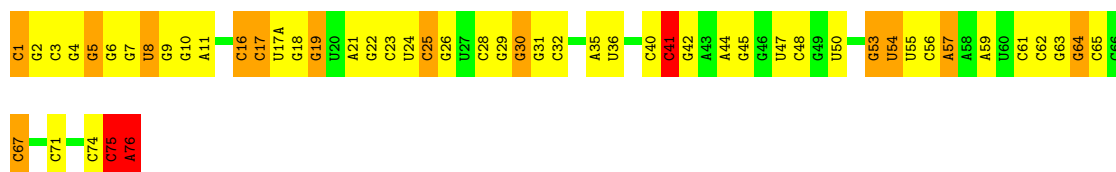
- Molecule 21: 30S ribosomal protein Thx

Chain U:



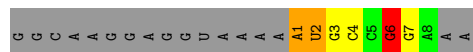
- Molecule 22: P-site tRNA fMET

Chain V:



- Molecule 23: mRNA

Chain X: 



- Molecule 24: A site ASL of tRNA-Proline CGG (unmodified)

Chain Y: 



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	209.82Å 447.39Å 619.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	187.58 – 3.68	Depositor
% Data completeness (in resolution range)	99.1 (187.58-3.68)	Depositor
$R_{merge}$	0.34	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	PHENIX (phenix.refine: 1.8.3_1479)	Depositor
R, $R_{free}$	0.212 , 0.272	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	53592	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	85.0	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, MG, PAR

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	0.59	3/36098 (0.0%)	1.21	155/56341 (0.3%)
2	B	0.31	0/1959	0.52	0/2642
3	C	0.31	0/1629	0.53	0/2195
4	D	0.38	0/1733	0.58	1/2318 (0.0%)
5	E	0.35	0/1171	0.56	0/1576
6	F	0.38	0/856	0.54	0/1154
7	G	0.33	0/1276	0.50	0/1709
8	H	0.33	0/1136	0.55	0/1527
9	I	0.31	0/1029	0.55	0/1379
10	J	0.33	0/814	0.54	0/1095
11	K	0.36	0/900	0.57	0/1213
12	L	0.37	0/991	0.61	0/1327
13	M	0.32	0/974	0.59	0/1303
14	N	0.41	0/501	0.60	0/664
15	O	0.35	0/745	0.54	0/992
16	P	0.36	0/721	0.57	0/970
17	Q	0.35	0/847	0.54	0/1131
18	R	0.35	0/579	0.64	1/768 (0.1%)
19	S	0.33	0/689	0.61	0/926
20	T	0.36	0/765	0.64	0/1007
21	U	0.31	0/221	0.54	0/288
22	V	0.85	1/1836 (0.1%)	1.45	24/2859 (0.8%)
23	X	0.95	0/193	1.91	7/299 (2.3%)
24	Y	0.68	0/194	1.53	0/301
All	All	0.54	4/57857 (0.0%)	1.07	188/85984 (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
12	L	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
22	V	1	C	OP3-P	-10.81	1.48	1.61
1	A	1227	A	N9-C4	-6.86	1.33	1.37
1	A	1434	A	N9-C4	-5.66	1.34	1.37
1	A	32	A	N9-C4	5.42	1.41	1.37

The worst 5 of 188 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
22	V	75	C	O5'-P-OP2	-13.28	93.75	105.70
1	A	1301	U	N3-C2-O2	-9.51	115.54	122.20
22	V	17	C	N1-C2-O2	9.47	124.58	118.90
1	A	1200	C	N1-C2-O2	8.49	123.99	118.90
1	A	1301	U	N1-C2-O2	8.40	128.68	122.80

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
12	L	47	LYS	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	32247	0	16277	763	0
2	B	1924	0	1975	64	0
3	C	1605	0	1668	45	0
4	D	1703	0	1763	63	0
5	E	1155	0	1213	36	0
6	F	843	0	857	20	0
7	G	1257	0	1296	41	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	H	1116	0	1177	40	0
9	I	1010	0	1037	43	0
10	J	801	0	849	50	0
11	K	885	0	904	29	0
12	L	975	0	1062	43	0
13	M	964	0	1034	31	0
14	N	492	0	529	27	0
15	O	734	0	771	23	0
16	P	705	0	725	21	0
17	Q	834	0	904	26	0
18	R	574	0	644	13	0
19	S	674	0	699	38	0
20	T	763	0	861	25	0
21	U	217	0	234	12	0
22	V	1644	0	836	26	0
23	X	173	0	88	3	0
24	Y	174	0	88	4	0
25	A	42	0	45	1	0
26	A	76	0	0	0	0
26	F	1	0	0	0	0
26	M	1	0	0	0	0
26	V	1	0	0	0	0
27	D	1	0	0	0	0
27	N	1	0	0	0	0
All	All	53592	0	37536	1299	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 15.

The worst 5 of 1299 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:1244:C:H42	1:A:1293:G:H1	1.10	0.98
4:D:9:CYS:SG	4:D:22:LYS:CE	2.52	0.97
4:D:9:CYS:SG	4:D:31:CYS:O	2.23	0.96
4:D:9:CYS:SG	4:D:22:LYS:HE3	2.07	0.95
1:A:559:A:H4'	1:A:560:U:H3'	1.49	0.93

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	
2	B	235/256 (92%)	174 (74%)	44 (19%)	17 (7%)	2	29
3	C	203/239 (85%)	163 (80%)	34 (17%)	6 (3%)	7	57
4	D	206/209 (99%)	176 (85%)	24 (12%)	6 (3%)	7	57
5	E	149/162 (92%)	136 (91%)	8 (5%)	5 (3%)	6	54
6	F	99/101 (98%)	95 (96%)	4 (4%)	0	100	100
7	G	153/156 (98%)	135 (88%)	16 (10%)	2 (1%)	18	74
8	H	136/138 (99%)	121 (89%)	14 (10%)	1 (1%)	30	84
9	I	125/128 (98%)	103 (82%)	17 (14%)	5 (4%)	5	48
10	J	97/105 (92%)	75 (77%)	19 (20%)	3 (3%)	7	56
11	K	117/129 (91%)	100 (86%)	14 (12%)	3 (3%)	8	60
12	L	123/132 (93%)	98 (80%)	18 (15%)	7 (6%)	3	37
13	M	119/126 (94%)	95 (80%)	15 (13%)	9 (8%)	2	27
14	N	58/61 (95%)	48 (83%)	6 (10%)	4 (7%)	2	31
15	O	86/89 (97%)	80 (93%)	5 (6%)	1 (1%)	19	76
16	P	82/88 (93%)	74 (90%)	7 (8%)	1 (1%)	19	76
17	Q	98/105 (93%)	91 (93%)	5 (5%)	2 (2%)	11	66
18	R	68/88 (77%)	56 (82%)	9 (13%)	3 (4%)	4	45
19	S	82/93 (88%)	56 (68%)	15 (18%)	11 (13%)	0	10
20	T	97/106 (92%)	76 (78%)	15 (16%)	6 (6%)	2	34
21	U	23/27 (85%)	19 (83%)	3 (13%)	1 (4%)	4	46
All	All	2356/2538 (93%)	1971 (84%)	292 (12%)	93 (4%)	5	49

5 of 93 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	236	TYR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	C	12	LEU
3	C	190	ARG
4	D	28	SER
13	M	67	GLU

### 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	
2	B	205/220 (93%)	172 (84%)	33 (16%)	3	24
3	C	159/188 (85%)	145 (91%)	14 (9%)	14	58
4	D	180/181 (99%)	157 (87%)	23 (13%)	6	35
5	E	116/123 (94%)	104 (90%)	12 (10%)	10	48
6	F	90/90 (100%)	78 (87%)	12 (13%)	6	33
7	G	126/127 (99%)	114 (90%)	12 (10%)	12	53
8	H	119/119 (100%)	109 (92%)	10 (8%)	16	61
9	I	98/99 (99%)	81 (83%)	17 (17%)	3	19
10	J	89/92 (97%)	77 (86%)	12 (14%)	6	33
11	K	90/99 (91%)	81 (90%)	9 (10%)	11	50
12	L	104/109 (95%)	87 (84%)	17 (16%)	3	23
13	M	97/101 (96%)	73 (75%)	24 (25%)	1	7
14	N	49/50 (98%)	40 (82%)	9 (18%)	2	16
15	O	79/80 (99%)	72 (91%)	7 (9%)	14	57
16	P	72/74 (97%)	63 (88%)	9 (12%)	7	37
17	Q	95/97 (98%)	87 (92%)	8 (8%)	16	61
18	R	61/77 (79%)	50 (82%)	11 (18%)	2	17
19	S	73/80 (91%)	59 (81%)	14 (19%)	2	13
20	T	76/82 (93%)	67 (88%)	9 (12%)	8	39
21	U	20/22 (91%)	20 (100%)	0	100	100
All	All	1998/2110 (95%)	1736 (87%)	262 (13%)	6	34

5 of 262 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	I	11	LYS
11	K	29	ILE
19	S	21	GLU
9	I	64	THR
10	J	22	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
10	J	13	HIS
19	S	47	HIS
10	J	78	ASN
2	B	204	ASN
13	M	92	HIS

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1498/1522 (98%)	341 (22%)	49 (3%)
22	V	76/77 (98%)	22 (28%)	1 (1%)
23	X	7/25 (28%)	4 (57%)	1 (14%)
24	Y	7/17 (41%)	3 (42%)	0
All	All	1588/1641 (96%)	370 (23%)	51 (3%)

5 of 370 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	7	G
1	A	9	G
1	A	22	G
1	A	32	A
1	A	39	G

5 of 51 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	701	C
1	A	913	A
1	A	1503	A

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	752	G
1	A	934	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 ⓘ

Of 82 ligands modelled in this entry, 81 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$
25	PAR	A	1601	-	45,45,45	1.31	6 (13%)	67,67,67	1.39	8 (11%)

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
25	PAR	A	1601	-	-	0/18/94/94	0/4/4/4

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
25	A	1601	PAR	C52-C42	3.28	1.58	1.52
25	A	1601	PAR	C64-C54	2.89	1.59	1.51

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
25	A	1601	PAR	O54-C14	2.86	1.49	1.41
25	A	1601	PAR	O51-C11	2.35	1.47	1.41
25	A	1601	PAR	C11-C21	2.29	1.57	1.52

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
25	A	1601	PAR	O52-C13-C23	4.70	115.89	107.50
25	A	1601	PAR	C14-O54-C54	3.82	121.13	113.73
25	A	1601	PAR	O33-C14-C24	3.34	114.69	108.08
25	A	1601	PAR	O11-C42-C52	3.11	115.34	107.42
25	A	1601	PAR	O11-C42-C32	-3.03	101.71	108.97

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.