



wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 01:56 PM BST

PDB ID : 2YKR
EMDB ID: : EMD-1884
Title : 30S ribosomal subunit with RsgA bound in the presence of GMPPNP
Authors : Guo, Q.; Yuan, Y.; Xu, Y.; Feng, B.; Liu, L.; Chen, K.; Lei, J.; Gao, N.
Deposited on : 2011-05-30
Resolution : 9.80 Å(reported)
Based on PDB ID : 3OFA, 2RCN

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
<http://wwpdb.org/validation/2016/EMValidationReportHelp>

MolProbity : 4.02b-467
Mogul : unknown
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : trunk27241

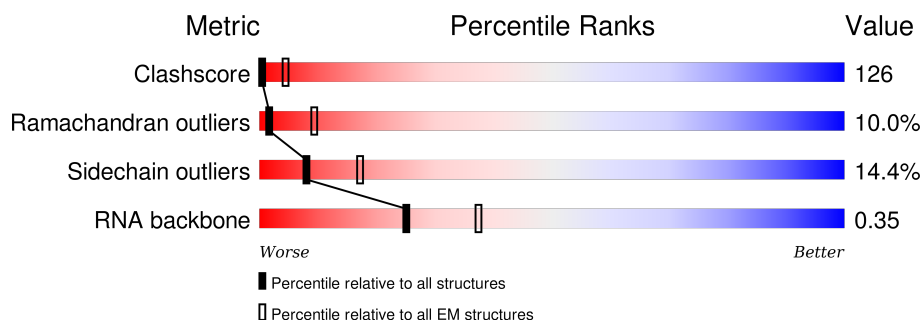
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 9.80 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.



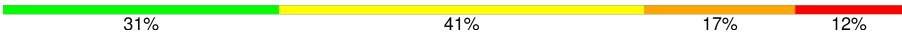
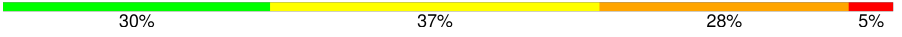

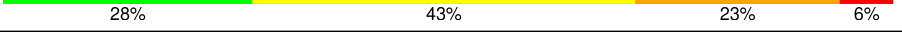

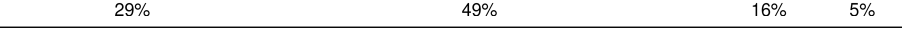
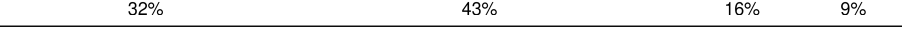
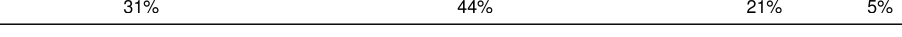

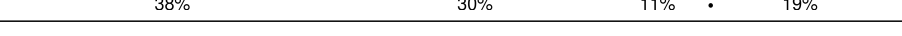
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686
RNA backbone	3027	244

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	1533	14% 86%
2	B	218	26% 44% 24% 6%
3	C	206	27% 43% 25% 5%
4	D	205	29% 48% 16% 7%
5	E	150	21% 51% 19% 9%
6	F	100	31% 40% 21% 8%
7	G	151	52% 32% 11% 5%
8	H	129	26% 39% 26% 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	I	127	
10	J	98	
11	K	117	
12	L	123	
13	M	114	
14	N	100	
15	O	88	
16	P	82	
17	Q	80	
18	R	55	
19	S	79	
20	T	85	
21	U	51	
22	W	350	

2 Entry composition

There are 22 unique types of molecules in this entry. The entry contains 53633 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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					AltConf	Trace
1	A	1533	Total	C	N	O	P	0	0
			32892	14671	6036	10653	1532		

- Molecule 2 is a protein called 30S RIBOSOMAL PROTEIN S2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	218	Total	C	N	O	S	0	0
			1705	1081	305	312	7		

- Molecule 3 is a protein called 30S RIBOSOMAL PROTEIN S3.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	206	Total	C	N	O	S	0	0
			1625	1028	305	289	3		

- Molecule 4 is a protein called 30S RIBOSOMAL PROTEIN S4.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	205	Total	C	N	O	S	0	0
			1643	1026	315	298	4		

- Molecule 5 is a protein called 30S RIBOSOMAL PROTEIN S5.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	150	Total	C	N	O	S	0	0
			1106	687	211	202	6		

- Molecule 6 is a protein called 30S RIBOSOMAL PROTEIN S6.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	100	Total	C	N	O	S	0	0
			818	515	148	149	6		

- Molecule 7 is a protein called 30S RIBOSOMAL PROTEIN S7.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	151	Total	C	N	O	S	0	0
			1182	735	227	216	4		

- Molecule 8 is a protein called 30S RIBOSOMAL PROTEIN S8.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	129	Total	C	N	O	S	0	0
			979	616	173	184	6		

- Molecule 9 is a protein called 30S RIBOSOMAL PROTEIN S9.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	I	127	Total	C	N	O	S	0	0
			1022	634	206	179	3		

- Molecule 10 is a protein called 30S RIBOSOMAL PROTEIN S10.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	98	Total	C	N	O	S	0	0
			787	493	150	143	1		

- Molecule 11 is a protein called 30S RIBOSOMAL PROTEIN S11.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	K	117	Total	C	N	O	S	0	0
			877	540	174	160	3		

- Molecule 12 is a protein called 30S RIBOSOMAL PROTEIN S12.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	123	Total	C	N	O	S	0	0
			955	590	196	165	4		

- Molecule 13 is a protein called 30S RIBOSOMAL PROTEIN S13.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	M	114	Total	C	N	O	S	0	0
			884	546	178	157	3		

- Molecule 14 is a protein called 30S RIBOSOMAL PROTEIN S14.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	N	97	Total	C	N	O	S	0	1
			775	483	161	128	3		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
N	39	ASP	GLU	CONFLICT	UNP B7M1M1

- Molecule 15 is a protein called 30S RIBOSOMAL PROTEIN S15.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	O	88	Total	C	N	O	S	0	0
			714	439	144	130	1		

- Molecule 16 is a protein called 30S RIBOSOMAL PROTEIN S16.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	P	82	Total	C	N	O	S	0	0
			649	406	128	114	1		

- Molecule 17 is a protein called 30S RIBOSOMAL PROTEIN S17.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	Q	80	Total	C	N	O	S	0	0
			649	411	121	114	3		

- Molecule 18 is a protein called 30S RIBOSOMAL PROTEIN S18.

Mol	Chain	Residues	Atoms				AltConf	Trace
18	R	55	Total	C	N	O	0	0
			456	288	86	82		

- Molecule 19 is a protein called 30S RIBOSOMAL PROTEIN S19.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	S	79	Total	C	N	O	S	0	0
			638	408	120	108	2		

- Molecule 20 is a protein called 30S RIBOSOMAL PROTEIN S20.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	T	85	Total	C	N	O	S	0	0
			665	411	137	114	3		

- Molecule 21 is a protein called 30S RIBOSOMAL PROTEIN S21.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	U	51	Total	C	N	O	S	0	0
			426	265	86	74	1		

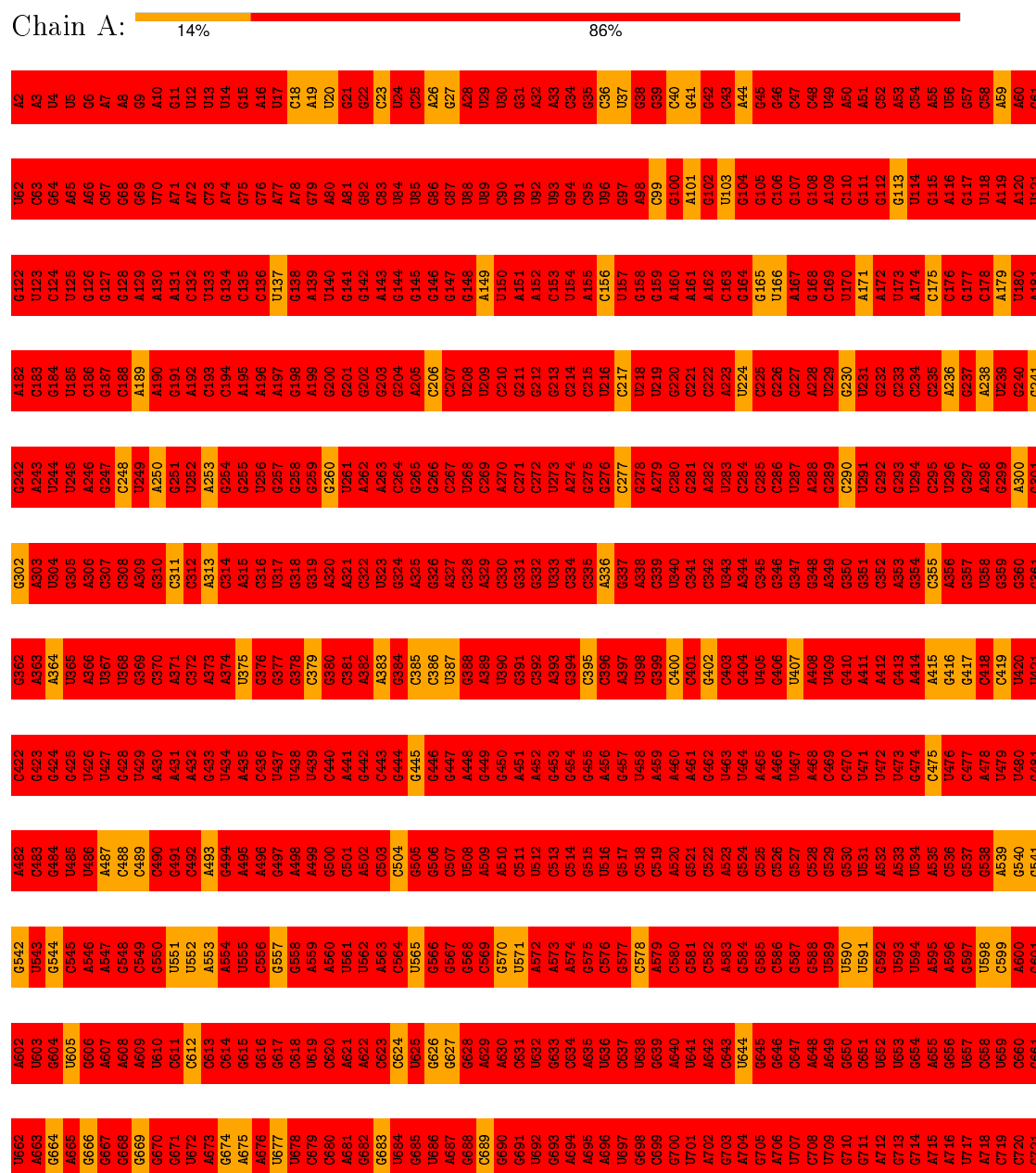
- Molecule 22 is a protein called PUTATIVE RIBOSOME BIOGENESIS GTPASE RSGA.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	W	282	Total	C	N	O	S	0	4
			2186	1378	388	410	10		

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. 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. 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: 16S rRNA

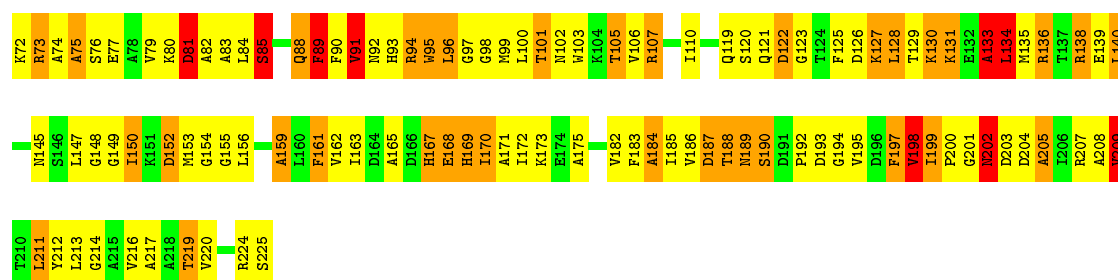


A1502	G1442	C1382	C1322	C1262	U1202	G1142	A1082	A1022	C962	G902	U842	A782	G722
A1503	C1443	C1383	G1323	C1263	C1203	G1143	U1083	U1023	G963	G903	U843	C782	U723
A1504	C1384	C1324	C1324	U1264	A1204	G1144	C1084	G1024	U964	U904	G844	A784	G724
G1505	U1445	G1385	A1325	C1265	U1205	U1145	U1085	U1025	U965	U905	A845	G785	G725
A1506	U1446	G1386	C1326	C1266	G1206	A1146	U1086	G1026	G966	A906	G846	G786	G726
A1507	A1447	G1387	U1327	C1267	G1207	C1147	G1087	C1027	G967	A907	G847	A787	G727
A1508	C1448	C1388	G1328	C1268	C1208	U1148	U1088	U1028	U968	A908	C848	U788	A728
C1509	C1449	C1389	A1329	A1269	C1209	C1149	U1089	C1029	U969	A909	G849	A789	A729
G1510	U1450	U1390	U1330	G1270	C1210	A1150	U1090	U1030	C970	C910	U850	A790	G730
G1511	U1451	G1391	G1331	A1271	U1211	A1151	U1091	C1031	G971	U911	G851	G791	G731
G1512	C1452	G1392	A1332	G1272	U1212	A1152	A1092	G1032	C972	C912	G852	A792	C732
A1513	G1453	U1393	A1333	C1273	A1213	G1153	A1093	G1033	G973	A913	C853	U793	G733
G1514	A1454	A1394	G1334	A1274	C1214	G1154	G1094	G1034	A974	A914	U854	A794	G734
G1515	G1455	C1395	U1335	A1275	G1215	A1155	U1095	A1035	A975	A915	U855	C795	C735
G1516	A1456	A1396	G1336	G1276	A1216	G1156	C1096	A1036	G976	U916	C856	C796	C736
G1517	G1457	C1397	G1337	C1277	C1217	A1157	C1097	C1037	A977	G917	C857	C797	C737
A1518	G1458	A1398	G1338	G1278	A1218	U1158	C1098	G1038	A978	A918	U858	U798	C738
A1519	G1459	A1399	A1339	G1279	C1219	U1159	G1099	G1039	C979	A919	G859	G799	C739
C1520	C1460	C1400	A1340	A1280	G1220	C1160	C1100	U1040	U880	U920	A860	G800	U740
C1521	G1461	U1341	U1341	C1281	G1221	C1161	A1101	G1041	U881	U921	G861	U801	G741
C1522	C1462	C1342	C1342	C1282	G1222	C1162	A1102	A1042	U882	G922	C862	A802	G742
G1523	C1463	G1343	G1343	U1283	C1223	A1163	C1103	G1043	A883	A923	U863	G803	A743
C1524	U1464	C1404	C1344	C1284	U1224	G1164	G1104	A1044	C984	C924	A864	U804	C744
G1525	A1465	G1405	U1345	A1285	A1225	U1165	A1105	C1045	C985	G925	A865	C805	G745
G1526	U1466	A1346	A1346	U1286	C1226	G1166	G1106	A1046	U986	G926	C866	C806	A746
U1527	C1467	G1347	G1347	A1287	A1227	A1167	C1107	G1047	C987	G927	G867	A807	A747
U1528	A1468	U1348	U1348	C1288	G1228	U1168	U988	G1048	G988	G928	C868	C808	G748
G1529	C1469	A1349	A1349	A1289	A1229	C1169	C1109	U1049	C989	G929	G869	G809	A749
G1530	A1410	A1350	A1350	G1290	C1230	A1170	C990	G1050	C930	C930	U870	C810	C750
A1531	U1471	U1351	U1351	G1291	G1231	A1171	U991	C1051	U991	C931	A871	C811	U751
U1532	U1472	C1352	C1352	G1292	U1232	C1172	C1112	U1052	U992	C932	A872	G812	G752
C1533	C1473	A1413	G1353	C1293	G1233	U1173	C1113	G1053	C993	C933	A873	U813	A753
A1534	U1474	U1354	U1354	G1294	C1234	G1174	C1114	A1054	A994	C934	A874	A814	C754
	G1475	G1355	G1355	U1295	U1235	G1175	U1115	A1055	C995	A935	U875	A815	G755
E43	A1476	G1416	G1356	C1296	A1236	A1176	U1116	U1056	A996	C936	C876	A816	C756
E44	U1477	A1357	A1357	G1297	C1237	G1177	A1117	G1057	U997	A937	A877	C817	U757
T45	U1478	A1418	U1358	U1298	A1238	G1178	U1118	G1058	C998	A938	A878	G818	C758
V46	C1479	C1359	C1359	A1299	A1239	A1179	C1119	C1059	C999	C939	C879	A819	A759
F47	U1480	A1360	A1360	G1300	U1240	A1180	C1120	U1060	A1000	C940	C880	U820	G760
U48	U1481	G1361	G1361	U1301	G1241	G1181	U1121	G1061	C1001	G941	G881	G821	G761
F49	G1482	A1362	A1362	C1302	G1242	G1182	U1122	U1062	G1002	G942	C882	U822	U762
A1483	G1483	A1363	C1363	C1303	C1243	U1183	C1123	C1063	G1003	U943	C883	C823	G763
A1484	U1424	U1364	U1364	G1304	G1244	G1184	G1124	A1064	A1004	G944	U884	G824	C764
U1485	U1425	G1365	G1365	G1305	C1245	G1185	U1125	U1065	A1005	G945	C885	A825	G765
G1486	G1426	C1366	C1366	A1306	A1246	G1186	U1126	C1066	G1006	A946	C886	C826	A766
G1487	C1427	C1367	C1367	U1307	U1247	G1187	G1127	A1067	U1007	G947	C887	U827	A767
G1488	A1428	A1368	A1368	U1308	A1248	A1188	C1128	G1068	U1008	C948	C888	U828	A768
G1489	A1429	C1369	C1369	G1309	C1249	U1189	C1129	C1069	U1009	A949	A889	G829	G769
U1490	A1430	G1370	G1370	G1310	A1250	G1190	A1130	U1070	U1010	U950	G890	G830	C770
G1491	A1431	G1371	A1371	A1311	A1251	A1191	C1131	C1071	C1011	G951	U891	A831	G771
A1492	G1432	U1372	U1372	G1312	A1252	C1192	C1132	G1072	A1012	U952	A892	G832	U772
A1493	G1373	G1373	G1373	C1313	G1253	G1193	G1133	U1073	C893	G953	C893	G833	G773
G1494	A1374	A1374	A1374	C1314	A1254	U1194	G1134	G1074	A1014	G954	G894	U834	G774
U1495	G1435	A1375	A1375	U1315	G1255	C1195	U1135	U1075	G1015	U955	G895	U835	G775
C1496	U1436	U1376	U1376	G1316	A1256	A1196	C1136	U1076	A1016	U956	C896	G836	G776
G1497	A1377	A1377	A1377	C1317	A1257	A1197	C1137	G1077	U897	U957	C897	U837	A777
U1498	C1378	C1378	C1378	A1318	G1258	G1198	G1138	U1078	G1018	A958	G898	G838	G778
A1499	U1379	U1379	U1379	A1319	C1259	U1199	G1139	G1079	A1019	A959	C899	C839	G779
A1500	U1440	C1380	C1380	C1320	G1260	U1199	G1139	A1080	G1020	U960	A900	C840	A780
C1501	A1441	U1381	U1381	U1321	A1261	A1201	C1141	A1081	A1021	U961	A901	C841	A781

• Molecule 2: 30S RIBOSOMAL PROTEIN S2

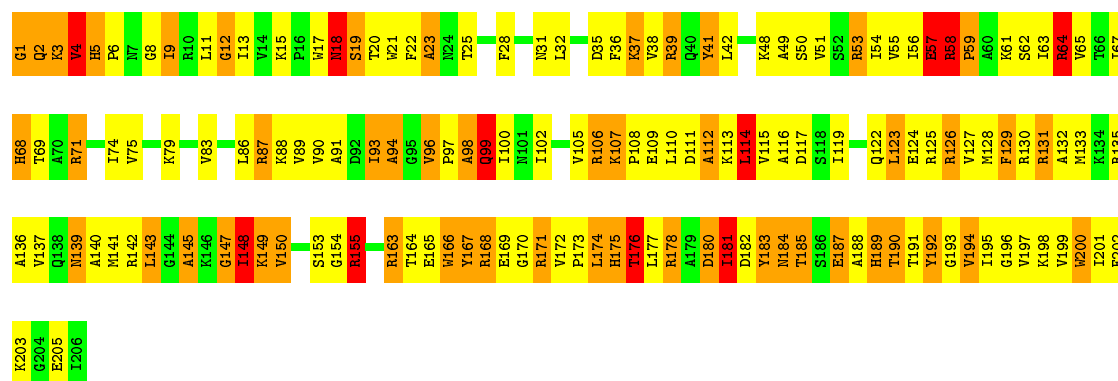
Chain B:  26% 44% 24% 6%

H8	A11	G12	V13	H14	F15	G16	H17	Q18	T19	R20	G21	Y21	N22	N23	P24	K25	N26	K27	F28	F29	R30	F31	G32	A33	R34	R35	R36	V37	A38	T39	L40	H41	L42	E43	E44	T45	V46	F47	U48	F49	J50	E51	A52	L53	L56	I59	A60	S61	G64	R65	T66	L67	F68	W69	G70	T71
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----



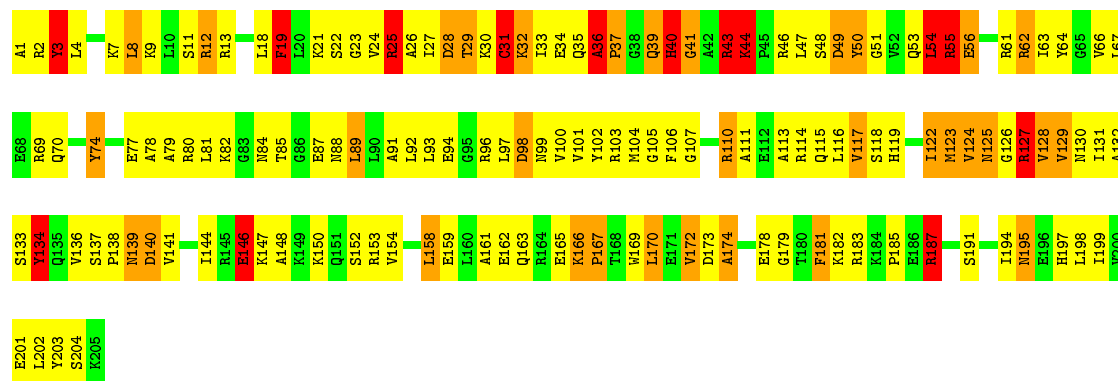
• Molecule 3: 30S RIBOSOMAL PROTEIN S3

Chain C: 27% 43% 25% 5%



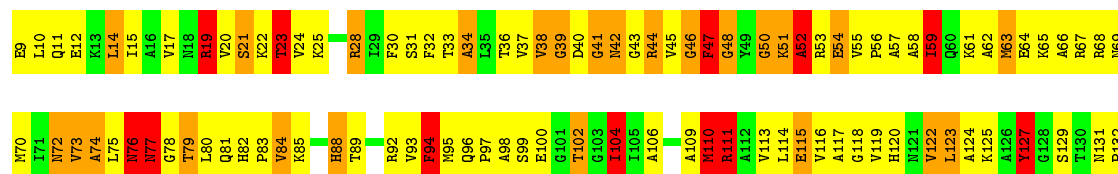
• Molecule 4: 30S RIBOSOMAL PROTEIN S4

Chain D: 29% 48% 16% 7%



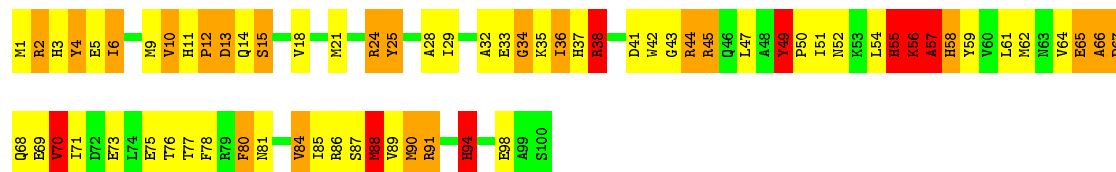
• Molecule 5: 30S RIBOSOMAL PROTEIN S5

Chain E: 21% 51% 19% 9%

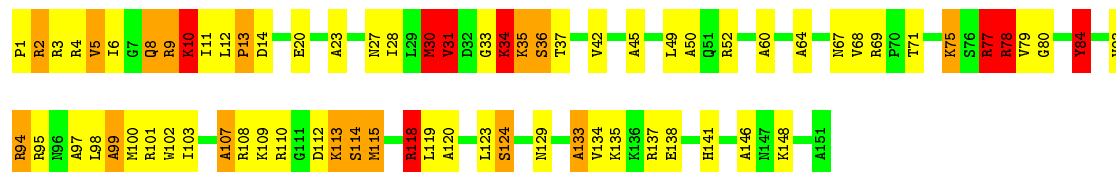




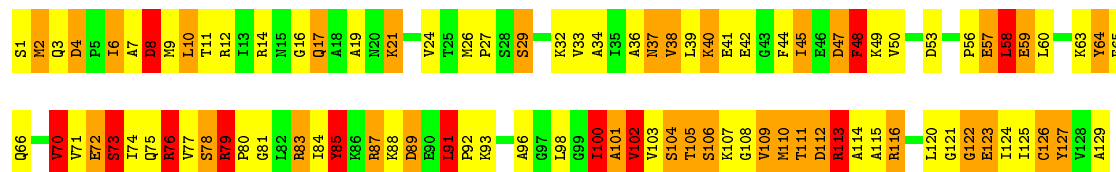
• Molecule 6: 30S RIBOSOMAL PROTEIN S6



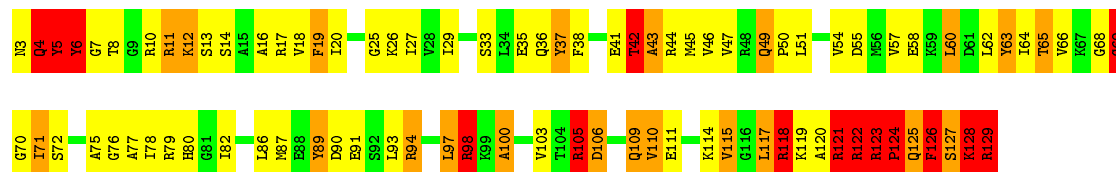
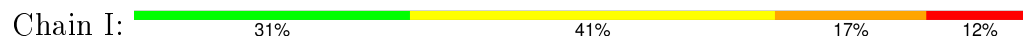
• Molecule 7: 30S RIBOSOMAL PROTEIN S7



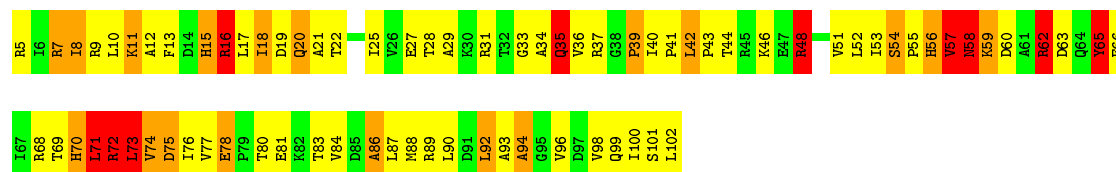
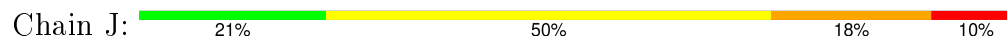
• Molecule 8: 30S RIBOSOMAL PROTEIN S8



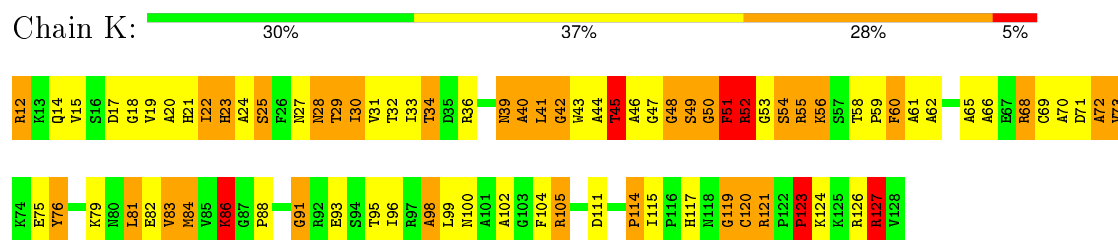
• Molecule 9: 30S RIBOSOMAL PROTEIN S9



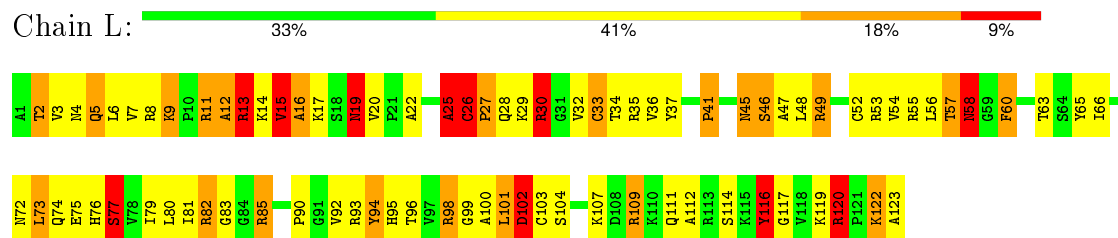
• Molecule 10: 30S RIBOSOMAL PROTEIN S10



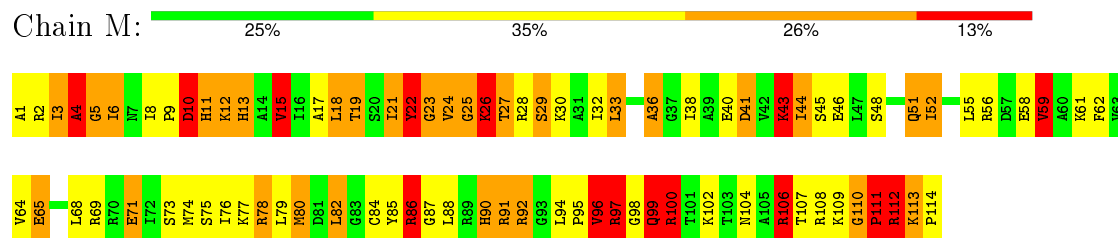
- Molecule 11: 30S RIBOSOMAL PROTEIN S11



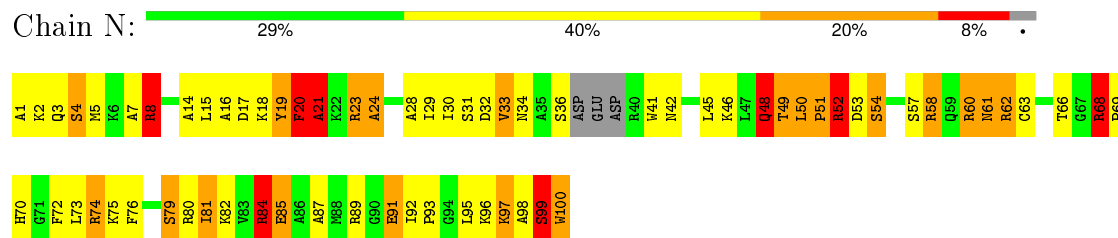
- Molecule 12: 30S RIBOSOMAL PROTEIN S12



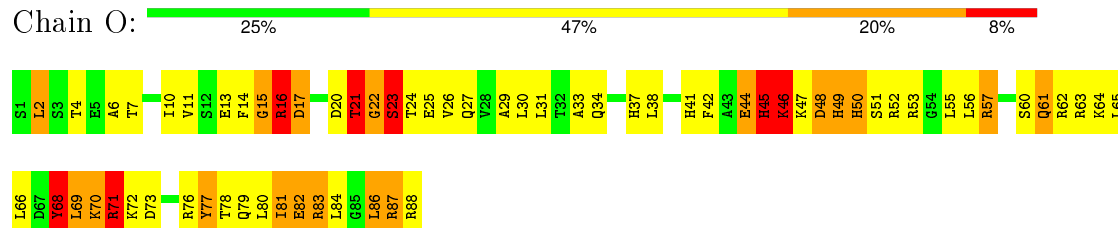
- Molecule 13: 30S RIBOSOMAL PROTEIN S13



- Molecule 14: 30S RIBOSOMAL PROTEIN S14

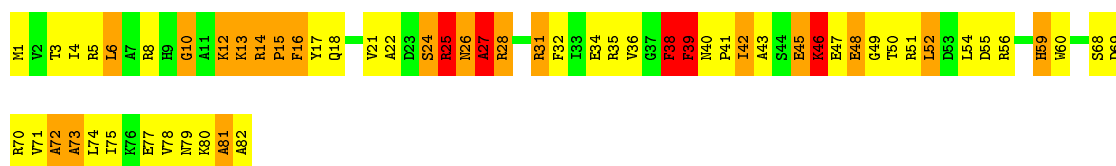


- Molecule 15: 30S RIBOSOMAL PROTEIN S15

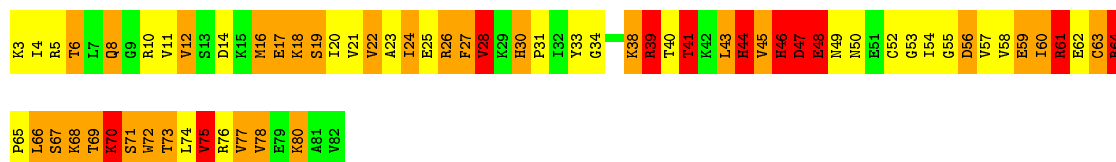


- Molecule 16: 30S RIBOSOMAL PROTEIN S16





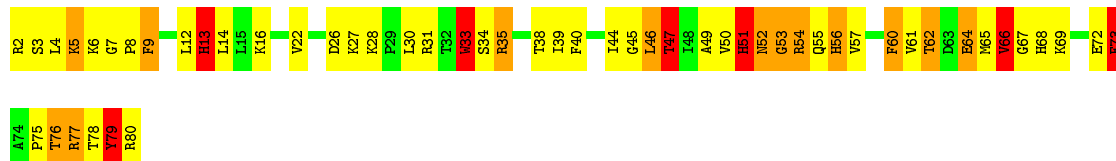
- Molecule 17: 30S RIBOSOMAL PROTEIN S17



- Molecule 18: 30S RIBOSOMAL PROTEIN S18



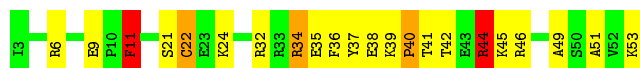
- Molecule 19: 30S RIBOSOMAL PROTEIN S19



- Molecule 20: 30S RIBOSOMAL PROTEIN S20



- Molecule 21: 30S RIBOSOMAL PROTEIN S21



- Molecule 22: PUTATIVE RIBOSOME BIOGENESIS GTPASE RSGA



MET	M65	S216	L294
SER	I66	G217	G295
LYS	R67	V218	L296
ASN	R68	G219	C297
LYS	L73	K220	K298
LEU	Y144	S221	Y299
SER	L145	S222	R300
LYS	L146	L223	D301
GLY	A147	L224	C302
GLN	C148	N225	K303
GLN	E149	A226	H304
ARG	T150	Q231	D305
ARG	L151	K232	T306
VAL	R82	E233	D307
ASN	P83	I234	C310
ALA	P86	L235	I312
ASN	P86	L235	R313
HIS	ALA	E238	I321
GLN	ALA	VAL	R325
ARG	GLU	SER	F326
ARG	GLY	ASP	R331
LEU	VAL	ASN	I332
LYS	ASN	SER	Q338
THR	V93	GLY	VAL
SER	K94	LEU	LYS
LYS	G95	GLY	THR
GLU	I96	GLN	ARG
LYS	V97	HIS	LYS
PRO	E98	THR	ASN
ASP	A99	THR	PHE
ASP	V100	T281	SER
ASP	R103	G194	ASP
ASN	T104	T281	THR
L35	S105	R284	ARG
E38	S106	L285	LYS
P39	L107	Y256	ASN
P40	L108	H257	PHE
E41	R109	F258	SER
P110	P110	F259	ASP
G42	D111	H260	THR
I43	PHE	G261	ASP
V44	TYR	V264	ASP
I45	ASP	R271	
S46	GLY	E272	
R47	V116	F273	
G49	K117	L274	
M50	A120	L275	
H51	A121	W276	
A52	T126	I283	
D53	V127	T284	
V54	V128	Q285	
E55	I128	G286	
S56	A131	F287	
A57	I132	F290	
H62	L133	Y293	
R63	P134		
O64	E135		

4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	MAPS FROM EACH DEFOCUS GROUP	Depositor
Microscope	OTHER	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	20	Depositor
Minimum defocus (nm)	2000	Depositor
Maximum defocus (nm)	3850	Depositor
Magnification	59000	Depositor
Image detector	FEI EAGLE 4K 4K CCD	Depositor

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 >2	RMSZ	# Z >2
1	A	3.88	6344/36831 (17.2%)	4.27	11020/57458 (19.2%)
10	J	1.41	0/797	2.00	23/1077 (2.1%)
11	K	1.61	4/893 (0.4%)	1.82	24/1205 (2.0%)
12	L	1.42	1/969 (0.1%)	1.88	27/1300 (2.1%)
13	M	1.60	5/893 (0.6%)	1.97	30/1193 (2.5%)
14	N	1.33	0/786	1.93	19/1045 (1.8%)
15	O	1.42	2/722 (0.3%)	1.99	26/964 (2.7%)
16	P	1.42	0/659	2.00	27/884 (3.1%)
17	Q	1.69	4/658 (0.6%)	2.31	40/881 (4.5%)
18	R	1.35	0/463	1.99	14/621 (2.3%)
19	S	1.20	1/653 (0.2%)	1.69	14/877 (1.6%)
2	B	1.35	2/1736 (0.1%)	1.90	59/2338 (2.5%)
20	T	1.37	1/671 (0.1%)	1.96	18/888 (2.0%)
21	U	1.28	0/431	1.60	2/570 (0.4%)
22	W	1.30	3/2223 (0.1%)	1.58	40/3008 (1.3%)
3	C	1.48	2/1652 (0.1%)	1.88	48/2225 (2.2%)
4	D	1.45	4/1665 (0.2%)	1.82	40/2227 (1.8%)
5	E	1.57	8/1119 (0.7%)	1.93	38/1504 (2.5%)
6	F	1.37	1/836 (0.1%)	1.96	29/1128 (2.6%)
7	G	1.27	2/1196 (0.2%)	1.59	18/1602 (1.1%)
8	H	1.42	2/989 (0.2%)	1.96	36/1326 (2.7%)
9	I	1.45	1/1034 (0.1%)	1.93	36/1375 (2.6%)
All	All	3.21	6387/57876 (11.0%)	3.66	11628/85696 (13.6%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1238
10	J	0	6
11	K	0	10

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
12	L	0	7
13	M	0	13
14	N	0	9
15	O	0	9
16	P	0	5
17	Q	0	7
18	R	0	4
19	S	0	8
2	B	0	14
20	T	0	5
21	U	0	2
22	W	0	7
3	C	0	16
4	D	0	15
5	E	0	6
6	F	0	7
7	G	0	4
8	H	0	12
9	I	0	15
All	All	0	1419

The worst 5 of 6387 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1306	A	N7-C5	-22.89	1.25	1.39
1	A	466	A	N7-C5	-22.87	1.25	1.39
1	A	627	G	N7-C5	-22.74	1.25	1.39
1	A	373	A	N9-C4	-22.28	1.24	1.37
1	A	78	A	N9-C4	22.19	1.51	1.37

The worst 5 of 11628 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1399	C	P-O3'-C3'	47.06	176.18	119.70
1	A	1139	G	P-O3'-C3'	47.02	176.13	119.70
1	A	556	C	C6-N1-C2	-44.83	102.37	120.30
1	A	306	A	P-O3'-C3'	39.05	166.56	119.70
1	A	73	C	C6-N1-C2	-38.49	104.91	120.30

There are no chirality outliers.

5 of 1419 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	2	A	Sidechain
1	A	3	A	Sidechain
1	A	4	U	Sidechain
1	A	5	U	Sidechain
1	A	6	G	Sidechain

5.2 Too-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 hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	32892	0	16464	9855	0
2	B	1705	0	1732	204	0
3	C	1625	0	1699	204	0
4	D	1643	0	1710	176	0
5	E	1106	0	1148	142	0
6	F	818	0	808	110	0
7	G	1182	0	1239	124	0
8	H	979	0	1034	139	0
9	I	1022	0	1070	132	0
10	J	787	0	828	130	0
11	K	877	0	887	165	0
12	L	955	0	1019	106	0
13	M	884	0	944	126	0
14	N	775	0	827	111	0
15	O	714	0	737	111	0
16	P	649	0	666	67	0
17	Q	649	0	690	143	0
18	R	456	0	478	57	0
19	S	638	0	665	75	0
20	T	665	0	714	109	0
21	U	426	0	449	19	0
22	W	2186	0	2180	107	0
All	All	53633	0	37988	11511	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 126.

The worst 5 of 11511 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:429:U:H5'	4:D:8:LEU:HD13	1.45	0.99
1:A:688:G:H5''	1:A:688:G:C8	1.99	0.98
1:A:82:G:H22	1:A:84:U:H3	1.11	0.95
1:A:450:G:H1	1:A:483:C:H42	1.13	0.95
1:A:1469:C:H5'	1:A:1469:C:C6	2.02	0.94

There are no symmetry-related clashes.

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 PDB entries followed by that with respect to all EM entries.

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	216/218 (99%)	149 (69%)	41 (19%)	26 (12%)	0	8
3	C	204/206 (99%)	158 (78%)	28 (14%)	18 (9%)	1	17
4	D	203/205 (99%)	139 (68%)	41 (20%)	23 (11%)	0	10
5	E	148/150 (99%)	104 (70%)	25 (17%)	19 (13%)	0	8
6	F	98/100 (98%)	75 (76%)	16 (16%)	7 (7%)	1	22
7	G	149/151 (99%)	118 (79%)	20 (13%)	11 (7%)	1	21
8	H	127/129 (98%)	90 (71%)	26 (20%)	11 (9%)	1	17
9	I	125/127 (98%)	83 (66%)	25 (20%)	17 (14%)	0	7
10	J	96/98 (98%)	74 (77%)	10 (10%)	12 (12%)	0	8
11	K	115/117 (98%)	81 (70%)	20 (17%)	14 (12%)	0	8
12	L	121/123 (98%)	85 (70%)	22 (18%)	14 (12%)	0	9
13	M	112/114 (98%)	82 (73%)	16 (14%)	14 (12%)	0	8
14	N	93/100 (93%)	61 (66%)	21 (23%)	11 (12%)	0	9
15	O	86/88 (98%)	65 (76%)	16 (19%)	5 (6%)	2	27
16	P	80/82 (98%)	58 (72%)	13 (16%)	9 (11%)	0	10
17	Q	78/80 (98%)	55 (70%)	16 (20%)	7 (9%)	1	17

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
18	R	53/55 (96%)	33 (62%)	12 (23%)	8 (15%)	0	5
19	S	77/79 (98%)	61 (79%)	12 (16%)	4 (5%)	2	30
20	T	83/85 (98%)	64 (77%)	14 (17%)	5 (6%)	2	26
21	U	49/51 (96%)	28 (57%)	11 (22%)	10 (20%)	0	3
22	W	274/350 (78%)	236 (86%)	25 (9%)	13 (5%)	3	32
All	All	2587/2708 (96%)	1899 (73%)	430 (17%)	258 (10%)	2	14

5 of 258 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	11	ALA
2	B	18	GLN
2	B	23	ASN
2	B	27	LYS
2	B	30	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 PDB entries followed by that with respect to all EM entries.

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	180/180 (100%)	156 (87%)	24 (13%)	5	28
3	C	170/170 (100%)	146 (86%)	24 (14%)	4	26
4	D	172/172 (100%)	147 (86%)	25 (14%)	4	24
5	E	113/113 (100%)	94 (83%)	19 (17%)	2	19
6	F	87/87 (100%)	74 (85%)	13 (15%)	4	23
7	G	124/124 (100%)	113 (91%)	11 (9%)	12	44
8	H	104/104 (100%)	84 (81%)	20 (19%)	2	13
9	I	105/105 (100%)	91 (87%)	14 (13%)	5	28
10	J	86/86 (100%)	71 (83%)	15 (17%)	2	17
11	K	90/90 (100%)	75 (83%)	15 (17%)	3	19
12	L	103/103 (100%)	88 (85%)	15 (15%)	4	24

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
13	M	92/92 (100%)	74 (80%)	18 (20%)	1	12
14	N	79/83 (95%)	67 (85%)	12 (15%)	3	22
15	O	76/76 (100%)	68 (90%)	8 (10%)	8	36
16	P	65/65 (100%)	57 (88%)	8 (12%)	6	30
17	Q	74/74 (100%)	52 (70%)	22 (30%)	0	3
18	R	48/48 (100%)	44 (92%)	4 (8%)	14	49
19	S	70/70 (100%)	58 (83%)	12 (17%)	2	17
20	T	65/65 (100%)	56 (86%)	9 (14%)	4	27
21	U	44/44 (100%)	40 (91%)	4 (9%)	12	43
22	W	238/302 (79%)	215 (90%)	23 (10%)	10	40
All	All	2185/2253 (97%)	1870 (86%)	315 (14%)	8	25

5 of 315 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	I	105	ARG
11	K	93	GLU
21	U	45	LYS
9	I	123	ARG
10	J	63	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 53 such sidechains are listed below:

Mol	Chain	Res	Type
10	J	35	GLN
11	K	117	HIS
22	W	193	HIS
10	J	56	HIS
11	K	21	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1532/1533 (99%)	479 (31%)	200 (13%)

5 of 479 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	3	A
1	A	4	U
1	A	5	U
1	A	6	G
1	A	7	A

5 of 200 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	531	U
1	A	754	C
1	A	1380	U
1	A	536	C
1	A	637	C

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 carbohydrates 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](#)

There are no chain breaks in this entry.