



wwPDB EM Map/Model Validation Report ⓘ

Oct 17, 2016 – 10:40 AM EDT

PDB ID : 5LMN
EMDB ID: : EMD-4073
Title : Structure of bacterial 30S-IF1-IF3-mRNA translation pre-initiation complex (state-1A)
Authors : Hussain, T.; Llacer, J.L.; Wimberly, B.T.; Ramakrishnan, V.
Deposited on : 2016-08-01
Resolution : 3.55 Å(reported)

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) : rb-20027939

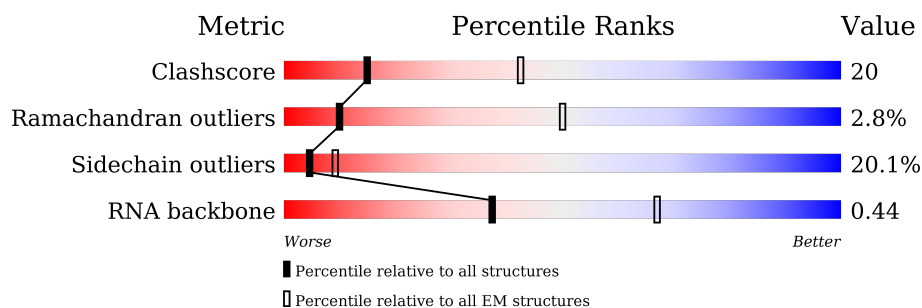
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.55 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	1522	23% 57% 17% ..
2	B	256	45% 32% 13% . 9%
3	C	239	55% 27% 5% 14%
4	D	209	56% 35% 9%
5	E	162	60% 27% 6% 7%
6	F	101	59% 38% ..
7	G	156	66% 31% .
8	H	138	56% 34% 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	I	128	
10	J	105	
11	K	129	
12	L	132	
13	M	126	
14	N	61	
15	O	89	
16	P	88	
17	Q	105	
18	R	88	
19	S	93	
20	T	106	
21	V	27	
22	W	72	
23	X	171	
24	Y	42	

2 Entry composition

There are 26 unique types of molecules in this entry. The entry contains 54149 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 ribosomal RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1514	Total	C	N	O	P	0	0
			32527	14481	6019	10515	1512		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	234	Total	C	N	O	S	0	0
			1900	1213	341	341	5		

- 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
			1612	1016	314	281	1		

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

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

- 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
			1146	724	217	201	4		

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

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

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

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

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

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

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

Mol	Chain	Residues	Atoms				AltConf	Trace
9	I	127	Total	C	N	O	0	0
			999	633	193	173		

- 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
			792	498	156	137	1		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
11	K	120	Total	C	N	O	S	0	0
			892	554	169	166	3		

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

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

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

Mol	Chain	Residues	Atoms					AltConf	Trace
13	M	117	Total	C	N	O	S	0	0
			933	577	192	162	2		

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

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

- 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
			734	459	147	126	2		

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

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

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

Mol	Chain	Residues	Atoms					AltConf	Trace
17	Q	99	Total	C	N	O	S	0	0
			823	528	151	142	2		

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

Mol	Chain	Residues	Atoms				AltConf	Trace
18	R	73	Total	C	N	O	0	0
			598	381	118	99		

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

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

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

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

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

Mol	Chain	Residues	Atoms				AltConf	Trace
21	V	24	Total	C	N	O	0	0
			208	128	50	30		

- Molecule 22 is a protein called Translation initiation factor IF-1.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	W	71	Total	C	N	O	S	0	0
			570	362	103	103	2		

- Molecule 23 is a protein called Translation initiation factor IF-3.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	X	168	Total	C	N	O	S	0	0
			1356	853	249	245	9		

- Molecule 24 is a RNA chain called mRNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	Y	20	Total	C	N	O	P	0	0
			439	196	89	134	20		

- Molecule 25 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
25	A	126	Total	Mg	0
			126	126	
25	N	1	Total	Mg	0
			1	1	

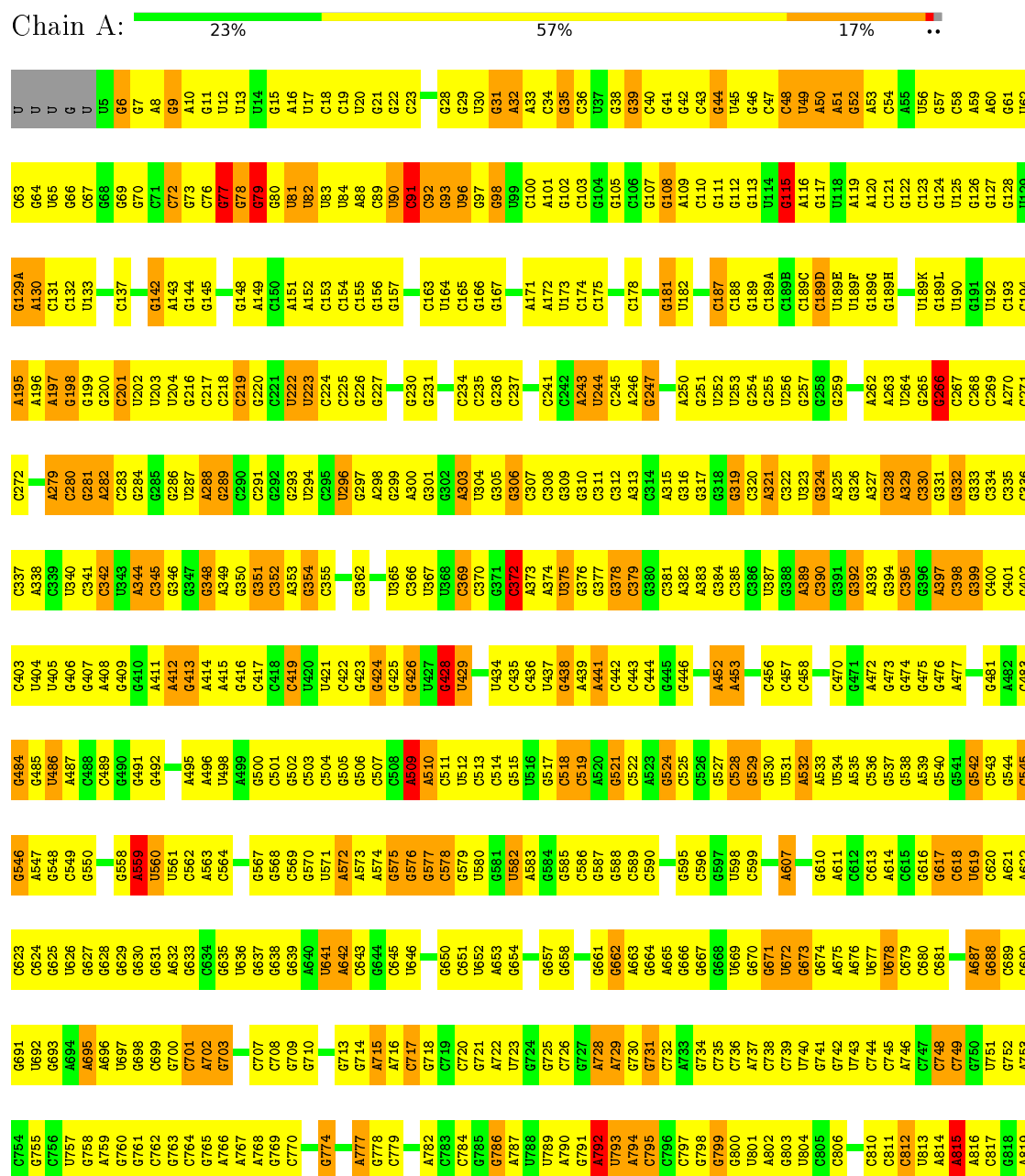
- Molecule 26 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
26	D	1	Total	Zn	0
			1	1	
26	N	1	Total	Zn	0
			1	1	

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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 ribosomal RNA

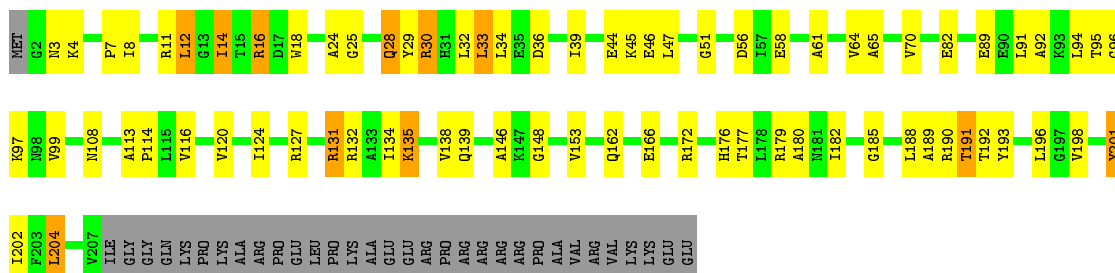


U820	C880	G954	A1080	G1143	G1272	C1336	C1399	G1470	A1534	U821	U891	G955	A1081	G1144	G1273	G1337	C1400	G1471	C1534
C824	C893	G1022	U1083	C1145	G1276	A1340	G1401	G1471	C1535	C825	U892	G1023	U1084	C1146	G1277	A1341	G1402	G1472	C1536
G826	G894	U960	U1085	U1147	U1278	G1343	G1403	G1473	C1537	C827	C895	G1024	U1086	C1148	U1279	G1344	G1404	G1473	C1538
U827	C896	C962	U1086	C1149	A1150	U1345	G1405	G1474	U1541	C828	C897	U1025	U1087	C1151	A1151	U1346	G1406	G1474	U1542
A828	C897	A965	U1088	U1151	G1281	A1346	G1409	G1475	C	A829	A900	C1028	U1089	U1152	A1152	U1347	G1410	G1475	U
G830	A900	G966	U1090	C1153	G1282	U1347	G1410	G1476	C	G831	A901	C1029	U1091	C1154	A1153	U1348	G1411	G1476	
C832	G902	A968	U1091	C1154	G1283	A1349	G1411	G1477		C833	G903	C1030	U1092	C1155	A1154	U1349	G1412	G1477	
C834	C904	A969	U1093	A1157	G1284	A1350	G1412	G1478		C835	C904	G1030	U1094	C1156	A1155	U1350	G1413	G1478	
U835	U905	G970	U1094	C1158	A1287	U1352	G1413	G1479		C836	U905	G1031	U1095	C1159	A1156	U1353	G1414	G1479	
G837	A908	G972	U1096	U1159	A1288	G1353	G1415	G1480		C838	A909	G1032	C1096	C1160	A1157	U1354	G1416	G1481	
U839	U910	A974	U1097	C1161	G1285	G1354	G1416	G1481		C839	A910	G1033	C1097	C1161	A1158	U1355	G1417	G1482	
C840	U911	A975	U1098	C1162	G1286	G1356	G1417	G1482		C841	U911	G1034	C1098	C1163	A1159	U1356	G1418	G1483	
U841	C912	G976	U1099	C1163	U1292	A1357	G1419	G1483		C842	U912	G1035	C1099	C1164	A1160	U1357	G1420	G1484	
C848	A913	A977	A1101	C1165	G1293	U1358	G1420	G1484		C849	A914	G1038	C1100	C1166	A1161	U1359	G1421	G1485	
U850	A915	G978	U1102	C1166	C1295	C1359	G1421	G1485		C851	A915	G1039	A983	C1167	A1162	U1360	G1422	G1486	
G851	G916	U981	C1103	A1163	C1296	A1360	G1422	G1486		C852	A916	U1040	C984	C1168	A1163	U1361	G1423	G1487	
C853	G917	U982	G1042	C1043	C1297	C1361	G1423	G1487		C854	A917	U1041	C985	C1169	A1164	U1362	G1424	G1488	
G855	C918	U983	G1043	C1044	C1298	C1362	G1424	G1488		C856	A918	U1042	C986	C1170	A1165	U1363	G1425	G1489	
U855	A919	U984	C1044	C1045	C1299	A1363	G1425	G1489		C857	A919	U1043	C987	C1171	A1166	U1364	G1426	G1490	
C858	A923	G979	U1045	C1046	C1300	U1364	G1426	G1490		C859	A920	U1044	C988	C1172	A1167	U1365	G1427	G1491	
A859	G924	C923	C1046	C1047	C1301	C1365	G1427	G1491		C860	A921	U1045	C989	C1173	A1168	U1366	G1428	G1492	
A860	G925	G926	U1051	C1052	C1302	C1366	G1428	G1492		C861	A922	U1046	C990	C1174	A1169	U1367	G1429	G1493	
G862	G927	G928	U1052	C1053	C1303	C1367	G1429	G1493		C863	A923	U1047	U991	C1175	A1170	U1368	G1430	G1494	
U863	G928	G929	C1054	A1055	C1304	C1368	G1430	G1494		C864	A924	U1048	U992	C1176	A1171	U1369	G1431	G1495	
A864	C930	C931	U1055	C1056	C1305	C1369	G1431	G1495		C865	A925	U1049	U993	C1177	A1172	U1370	G1432	G1496	
C866	C932	C933	U1056	C1057	C1306	C1370	G1432	G1496		C867	A926	U1050	U994	C1178	A1173	U1371	G1433	G1497	
C868	C934	A935	U1057	C1058	C1307	C1371	G1433	G1497		C869	A927	U1051	U995	C1179	A1174	U1372	G1434	G1498	
U870	C935	G1001A	U1058	C1059	C1308	C1372	G1434	G1498		C871	A928	U1052	U996	C1180	A1175	U1373	G1435	G1499	
U871	A938	G1002	C1060	C1061	C1309	C1373	G1435	G1499		C872	A929	U1053	U997	C1181	A1176	U1374	G1436	G1500	
A872	A939	G1003	C1063	C1064	C1310	C1374	G1436	G1500		C873	A930	U1054	U998	C1182	A1177	U1375	G1437	G1501	
U873	A940	G1004	C1064	C1065	C1311	C1375	G1437	G1501		C874	A931	U1055	U999	C1183	A1178	U1376	G1438	G1502	
C875	G941	C1006	C1066	C1067	C1312	C1376	G1438	G1502		C876	A932	U1056	U1000	C1184	A1179	U1377	G1439	G1503	
G876	G942	C1007	C1069	C1070	C1313	C1377	G1439	G1503		C877	A933	U1057	U1001	C1185	A1180	U1378	G1440	G1504	
C881	G945	G1010	C1071	C1072	C1314	C1378	G1440	G1504		C882	A934	U1058	U1002	C1186	A1181	U1379	G1441	G1505	
C883	A946	G1011	C1072	C1073	C1315	C1379	G1441	G1505		C884	A935	U1059	U1003	C1187	A1182	U1380	G1442	G1506	
U884	A947	G1012	C1073	C1074	C1316	C1380	G1442	G1506		C885	A936	U1060	U1004	C1188	A1183	U1381	G1443	G1507	
G885	A948	U1013	C1074	C1075	C1317	C1381	G1443	G1507		C886	A937	U1061	U1005	C1189	A1184	U1382	G1444	G1508	
C888	U951	A1015	C1075	C1076	C1318	C1382	G1444	G1508		C889	A938	U1062	U1006	C1190	A1185	U1383	G1445	G1509	
U889	G953	A1016	C1076	C1077	C1319	C1383	G1445	G1509		C890	A939	U1063	U1007	C1191	A1186	U1384	G1446	G1510	
		G1017	C1078	C1079	C1320	C1384	G1446	G1510		C891	A940	U1064	U1008	C1192	A1187	U1385	G1447	G1511	
		C1141	C1079	C1142	C1321	C1385	G1447	G1511		C892	A941	U1065	U1009	C1193	A1188	U1386	G1448	G1512	
		G1142	C1143	C1144	C1322	C1386	G1448	G1512		C893	A942	U1066	U1010	C1194	A1189	U1387	G1449	G1513	
			C1145	C1146	C1323	C1387	G1449	G1513		C894	A943	U1067	U1011	C1195	A1190	U1388	G1450	G1514	
			C1146	C1147	C1324	C1388	G1450	G1514		C895	A944	U1068	U1012	C1196	A1191	U1389	G1451	G1515	
			C1147	C1148	C1325	C1389	G1451	G1515		C896	A945	U1069	U1013	C1197	A1192	U1390	G1452	G1516	
			C1148	C1149	C1326	C1390	G1452	G1516		C897	A946	U1070	U1014	C1198	A1193	U1391	G1453	G1517	
			C1149	C1150	C1327	C1391	G1453	G1517		C898	A947	U1071	U1015	C1199	A1194	U1392	G1454	G1518	
			C1150	C1151	C1328	C1392	G1454	G1518		C899	A948	U1072	U1016	C1200	A1195	U1393	G1455	G1519	
			C1151	C1152	C1329	C1393	G1455	G1519		C900	A949	U1073	U1017	C1201	A1196	U1394	G1456	G1520	
			C1152	C1153	C1330	C1394	G1456	G1520		C901	A950	U1074	U1018	C1202	A1197	U1395	G1457	G1521	
			C1153	C1154	C1331	C1395	G1457	G1521		C902	A951	U1075	U1019	C1203	A1198	U1396	G1458	G1522	
			C1154	C1155	C1332	C1396	G1458	G1522		C903	A952	U1076	U1020	C1204	A1199	U1397	G1459	G1523	
			C1155	C1156	C1333	C1397	G1459	G1523		C904	A953	U1077	U1021	C1205	A1200	U1398	G1460	G1524	
			C1156	C1157	C1334	C1398	G1460	G1524		C905	A954	U1078	U1022	C1206	A1201	U1399	G1461	G1525	
			C1157	C1158	C1335	C1399	G1461	G1525		C906	A955	U1079	U1023	C1207	A1202	U1400	G1462	G1526	
			C1158	C1159	C1336	C1400	G1462	G1526		C907	A956	U1080	U1024	C1208	A1203	U1401	G1463	G1527	
			C1159	C1160	C1337	C1401	G1463	G1527		C908	A957	U1081	U1025	C1209	A1204	U1402	G1464	G1528	
			C1160	C1161	C1338	C1402	G1464	G1528		C909	A958	U1082	U1026	C1210	A1205	U1403	G1465	G1529	
			C1161	C1162	C1339	C1403	G1465	G1529		C910	A959	U1083	U1027	C1211	A1206	U1404	G1466	G1530	
			C1162	C1163	C1340	C1404	G1466	G1530		C911	A960	U1084	U1028	C1212	A1207	U1405	G1467	G1531	
			C1163	C1164	C1341	C1405	G1467	G1531		C912	A961	U1085	U1029	C1213	A1208	U1406	G1468	G1532	
			C1164	C1165	C1342	C1406	G1468	G1532		C913	A962	U1086	U1030	C1214	A1209	U1407	G1469	G1533	
			C1165	C1166	C1343	C1407	G1469	G1533		C914	A963	U1087	U1031	C1215	A1210	U1408	G1470	G1534	
			C1166	C1167	C1344	C1408	G1470	G1534		C915	A964	U1088	U1032	C1216	A1211	U1409	G1471	G1535	
			C1167	C1168	C1345	C1409	G1471	G1535		C916	A965	U1089	U1033	C1217	A1212	U1410	G1472	G1536	
			C1168	C1169	C1346	C1410	G1472	G1536		C917	A966	U1090	U1034	C1218	A1213	U1411	G1473	G1537	
			C1169	C1170	C1347	C1411	G1473	G1537		C918	A967	U1091	U1035	C1219	A1214	U1412	G1474	G1538	
			C1170	C1171	C1348	C1412	G1474	G1538		C919	A968	U1092	U1036	C1220	A1215	U1413	G1475	G1539	
			C1171	C1172	C1349	C1413	G1475	G1539		C920	A969	U1093	U1037	C1221	A1216	U1414	G1476	G1540	
			C1172	C1173	C1350	C1414	G1476	G1540		C921	A970	U1094	U1038	C1222	A1217	U1415	G1477	G1541	
			C1173	C1174	C1351	C1415	G1477	G1541		C922	A971	U1095	U1039	C1223	A1218	U1416	G1478	G1542	
			C1174	C1175	C1352	C1416	G1478	G1542		C923	A972	U1096	U1040	C1224	A1219	U1417	G1479	G1543	
			C1175	C1176															

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

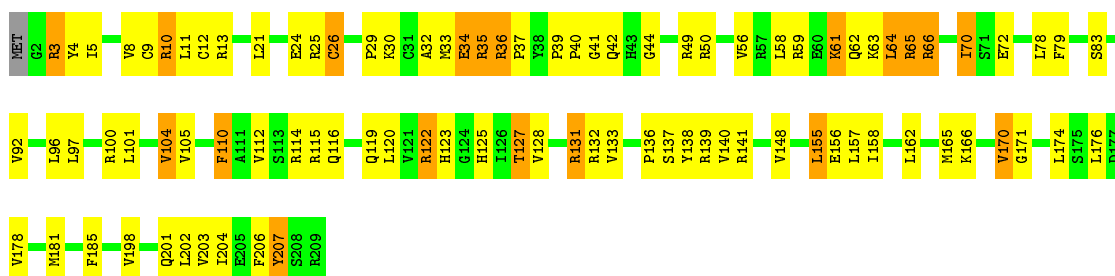
• Molecule 3: 30S ribosomal protein S3

Chain C:  55% 27% 5% 14%



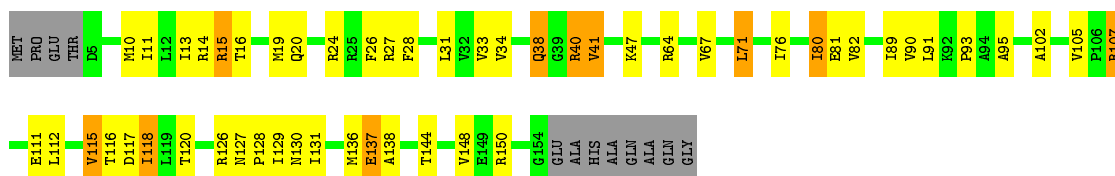
• Molecule 4: 30S ribosomal protein S4

Chain D:  56% 35% 9%



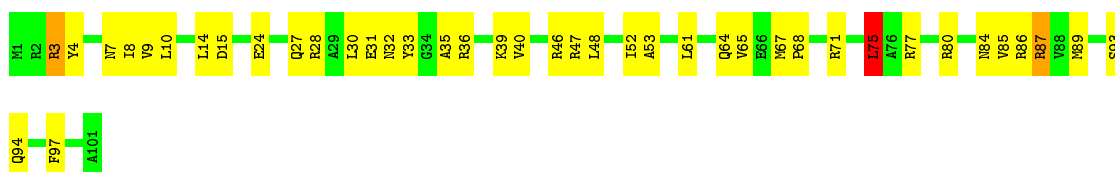
• Molecule 5: 30S ribosomal protein S5

Chain E:  60% 27% 6% 7%



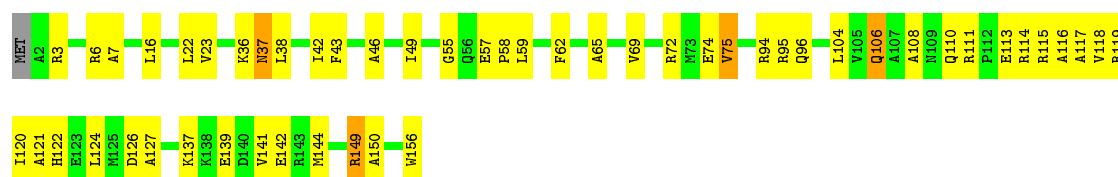
• Molecule 6: 30S ribosomal protein S6

Chain F:  59% 38% 2% 1%



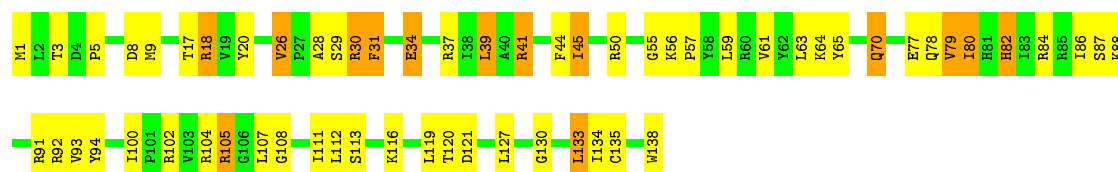
• Molecule 7: 30S ribosomal protein S7

Chain G:  66% 31% 2% 1%



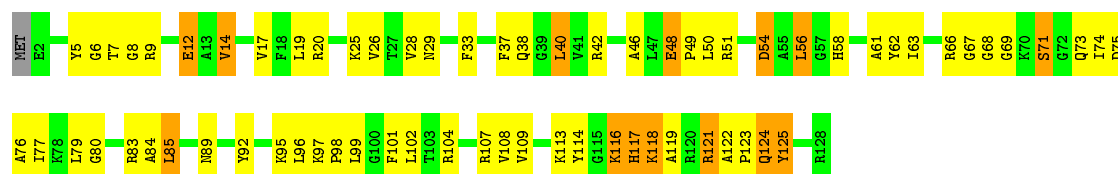
- Molecule 8: 30S ribosomal protein S8

Chain H: 56% 34% 10%



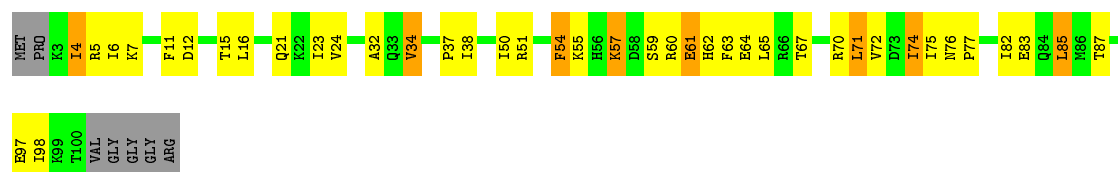
- Molecule 9: 30S ribosomal protein S9

Chain I: 45% 43% 11%



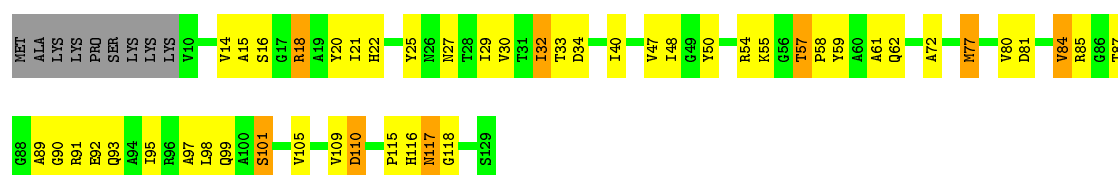
- Molecule 10: 30S ribosomal protein S10

Chain J: 54% 31% 8% 7%



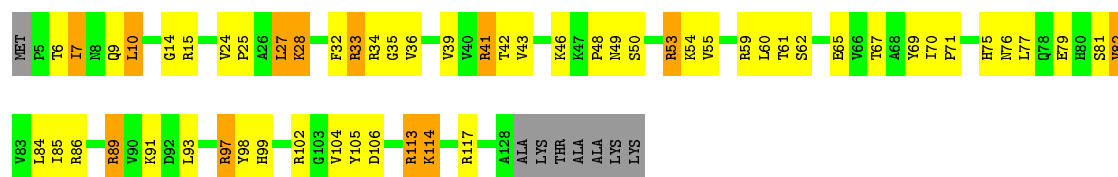
- Molecule 11: 30S ribosomal protein S11

Chain K: 55% 32% 6% 7%



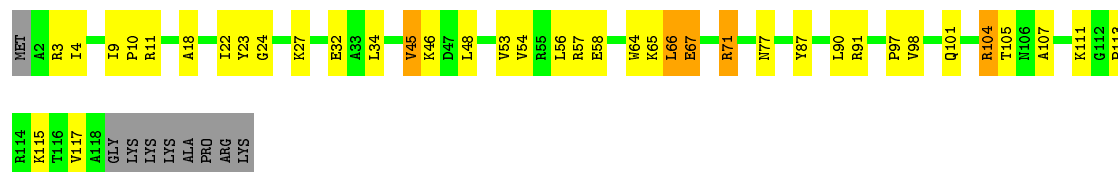
- Molecule 12: 30S ribosomal protein S12

Chain L: 51% 34% 9% 6%



- Molecule 13: 30S ribosomal protein S13

Chain M: 62% 27% 7%



- Molecule 14: 30S ribosomal protein S14 type Z

Chain N: 39% 48% 11%



- Molecule 15: 30S ribosomal protein S15

Chain O: 69% 25% 6%



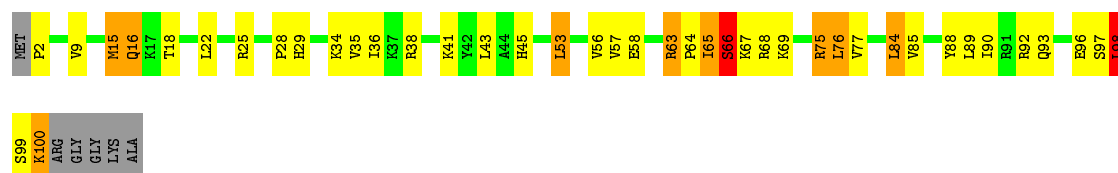
- Molecule 16: 30S ribosomal protein S16

Chain P: 64% 24% 6% 6%



- Molecule 17: 30S ribosomal protein S17

Chain Q: 54% 30% 9% 6%



- Molecule 18: 30S ribosomal protein S18

Chain R: 47% 31% 6% 17%



- [illegible]

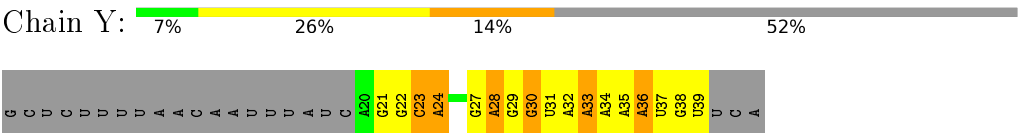
- [illegible]

- | | | | | | | | | | | | | | | | | | | | | | | |
|-----|----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|--|-----|-----|-----|--|-----|-----|-----|
| MET | G2 | K3 | G4 | D5 | R6 | R7 | T8 | R9 | R10 | G11 | K12 | I13 | W14 | R15 | | K20 | Y21 | R22 | | K25 | LYS | LYS |
|-----|----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|--|-----|-----|-----|--|-----|-----|-----|

-
- | Category | Value |
|----------|-------|
| MET | 1 |
| A1 | 1 |
| K2 | 1 |
| V12 | 1 |
| V13 | 1 |
| T14 | 1 |
| E15 | 1 |
| A16 | 1 |
| L17 | 1 |
| P18 | 1 |
| P19 | 1 |
| A20 | 1 |
| T21 | 1 |
| F22 | 1 |
| R23 | 1 |
| V24 | 1 |
| T32 | 1 |
| L33 | 1 |
| M40 | 1 |
| I45 | 1 |
| R46 | 1 |
| I47 | 1 |
| L48 | 1 |
| D51 | 1 |
| R52 | 1 |
| V53 | 1 |
| T58 | 1 |
| T63 | 1 |
| V68 | 1 |
| Y69 | 1 |
| R70 | 1 |
| K71 | 1 |

- [illegible]

- WORLDWIDE
PDB
PROTEIN DATA BANK
- EMDataBank**
Unified Data Resource for 3DEM



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	86892	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI POLARA 300	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	Not provided	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	78000	Depositor
Image detector	Not provided	Depositor

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

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	0.38	5/36403 (0.0%)	0.76	22/56800 (0.0%)
10	J	0.44	0/805	0.73	0/1082
11	K	0.43	0/907	0.71	0/1223
12	L	0.47	0/986	0.86	0/1320
13	M	0.43	0/943	0.76	0/1265
14	N	0.43	0/501	0.82	0/664
15	O	0.46	0/745	0.90	1/992 (0.1%)
16	P	0.55	0/716	0.85	1/963 (0.1%)
17	Q	0.46	0/836	0.82	0/1117
18	R	0.44	0/604	0.83	0/801
19	S	0.41	0/661	0.70	0/890
2	B	0.43	0/1935	0.77	2/2609 (0.1%)
20	T	0.46	0/765	0.90	0/1007
21	V	0.47	0/212	0.79	0/277
22	W	0.44	0/580	0.76	0/782
23	X	0.42	0/1375	0.77	0/1844
24	Y	0.33	0/494	0.79	0/770
3	C	0.45	0/1636	0.79	0/2205
4	D	0.45	0/1733	0.84	1/2318 (0.0%)
5	E	0.50	0/1162	0.86	0/1564
6	F	0.42	0/856	0.81	1/1154 (0.1%)
7	G	0.43	0/1276	0.79	0/1709
8	H	0.47	0/1136	0.89	2/1527 (0.1%)
9	I	0.43	0/1018	0.73	0/1368
All	All	0.41	5/58285 (0.0%)	0.78	30/86251 (0.0%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
8	H	0	1
All	All	0	2

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1459	C	O3'-P	-8.53	1.50	1.61
1	A	1458	G	O3'-P	7.51	1.70	1.61
1	A	999	C	O3'-P	-5.64	1.54	1.61
1	A	559	A	O3'-P	5.30	1.67	1.61
1	A	79	G	O3'-P	-5.17	1.54	1.61

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	93	G	O3'-P-O5'	-26.72	53.24	104.00
1	A	93	G	OP1-P-O3'	-22.17	56.43	105.20
1	A	93	G	P-O3'-C3'	-10.18	107.49	119.70
1	A	93	G	OP2-P-O3'	8.72	124.37	105.20
1	A	266	G	C2'-C3'-O3'	8.44	128.07	109.50

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1458	G	Sidechain
8	H	3	THR	Peptide

5.2 Too-close contacts [i](#)

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	32527	0	16428	1136	0
2	B	1900	0	1951	112	0
3	C	1612	0	1677	58	0
4	D	1703	0	1766	65	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	E	1146	0	1207	50	0
6	F	843	0	857	18	0
7	G	1257	0	1296	29	0
8	H	1116	0	1177	33	0
9	I	999	0	1015	79	0
10	J	792	0	829	37	0
11	K	892	0	913	38	0
12	L	970	0	1057	43	0
13	M	933	0	992	19	0
14	N	492	0	527	52	0
15	O	734	0	771	17	0
16	P	700	0	720	14	0
17	Q	823	0	891	29	0
18	R	598	0	670	34	0
19	S	647	0	673	27	0
20	T	763	0	861	31	0
21	V	208	0	221	11	0
22	W	570	0	599	37	0
23	X	1356	0	1401	40	0
24	Y	439	0	218	17	0
25	A	126	0	0	0	0
25	N	1	0	0	0	0
26	D	1	0	0	1	0
26	N	1	0	0	0	0
All	All	54149	0	38717	1806	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 20.

The worst 5 of 1806 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:563:A:N6	1:A:884:U:H3	1.13	1.44
23:X:91:ARG:N	23:X:94:ILE:HD11	1.33	1.41
5:E:15:ARG:CG	5:E:28:PHE:CE1	2.06	1.38
23:X:91:ARG:H	23:X:94:ILE:CD1	1.39	1.35
1:A:89:C:C2'	1:A:90:U:H5'	1.57	1.32

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 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	232/256 (91%)	194 (84%)	28 (12%)	10 (4%)	3	33
3	C	204/239 (85%)	176 (86%)	23 (11%)	5 (2%)	7	48
4	D	206/209 (99%)	180 (87%)	23 (11%)	3 (2%)	13	58
5	E	148/162 (91%)	135 (91%)	10 (7%)	3 (2%)	9	53
6	F	99/101 (98%)	91 (92%)	8 (8%)	0	100	100
7	G	153/156 (98%)	142 (93%)	8 (5%)	3 (2%)	9	53
8	H	136/138 (99%)	127 (93%)	7 (5%)	2 (2%)	13	58
9	I	125/128 (98%)	108 (86%)	13 (10%)	4 (3%)	5	42
10	J	96/105 (91%)	76 (79%)	16 (17%)	4 (4%)	3	34
11	K	118/129 (92%)	98 (83%)	16 (14%)	4 (3%)	5	41
12	L	122/132 (92%)	103 (84%)	15 (12%)	4 (3%)	5	41
13	M	115/126 (91%)	97 (84%)	15 (13%)	3 (3%)	7	46
14	N	58/61 (95%)	44 (76%)	10 (17%)	4 (7%)	1	20
15	O	86/89 (97%)	80 (93%)	5 (6%)	1 (1%)	16	63
16	P	81/88 (92%)	73 (90%)	6 (7%)	2 (2%)	7	48
17	Q	97/105 (92%)	82 (84%)	10 (10%)	5 (5%)	2	27
18	R	71/88 (81%)	61 (86%)	9 (13%)	1 (1%)	14	59
19	S	78/93 (84%)	66 (85%)	9 (12%)	3 (4%)	4	37
20	T	97/106 (92%)	84 (87%)	10 (10%)	3 (3%)	5	43
21	V	22/27 (82%)	20 (91%)	1 (4%)	1 (4%)	3	32
22	W	69/72 (96%)	55 (80%)	11 (16%)	3 (4%)	3	33
23	X	166/171 (97%)	145 (87%)	16 (10%)	5 (3%)	5	43
All	All	2579/2781 (93%)	2237 (87%)	269 (10%)	73 (3%)	10	45

5 of 73 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	9	GLU
2	B	29	ALA
3	C	192	THR
4	D	37	PRO
17	Q	66	SER

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	202/220 (92%)	142 (70%)	60 (30%)	0	3
3	C	160/188 (85%)	131 (82%)	29 (18%)	2	14
4	D	180/181 (99%)	146 (81%)	34 (19%)	2	12
5	E	115/123 (94%)	90 (78%)	25 (22%)	1	9
6	F	90/90 (100%)	72 (80%)	18 (20%)	1	11
7	G	126/127 (99%)	106 (84%)	20 (16%)	3	21
8	H	119/119 (100%)	90 (76%)	29 (24%)	1	6
9	I	96/99 (97%)	76 (79%)	20 (21%)	1	10
10	J	87/92 (95%)	75 (86%)	12 (14%)	4	28
11	K	91/99 (92%)	72 (79%)	19 (21%)	1	10
12	L	104/109 (95%)	79 (76%)	25 (24%)	1	6
13	M	94/101 (93%)	79 (84%)	15 (16%)	3	21
14	N	49/50 (98%)	41 (84%)	8 (16%)	3	20
15	O	79/80 (99%)	64 (81%)	15 (19%)	2	12
16	P	72/74 (97%)	59 (82%)	13 (18%)	2	14
17	Q	94/97 (97%)	76 (81%)	18 (19%)	2	12
18	R	64/77 (83%)	52 (81%)	12 (19%)	2	13
19	S	71/80 (89%)	59 (83%)	12 (17%)	2	18
20	T	76/82 (93%)	55 (72%)	21 (28%)	0	4
21	V	19/22 (86%)	15 (79%)	4 (21%)	1	10

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
22	W	62/63 (98%)	51 (82%)	11 (18%)	2	15
23	X	145/150 (97%)	123 (85%)	22 (15%)	3	23
All	All	2195/2323 (94%)	1753 (80%)	442 (20%)	4	11

5 of 442 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
8	H	59	LEU
11	K	16	SER
21	V	12	LYS
8	H	82	HIS
9	I	85	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
6	F	11	ASN
9	I	3	GLN
20	T	26	ASN
6	F	100	ASN
3	C	176	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1508/1522 (99%)	447 (29%)	99 (6%)
24	Y	19/42 (45%)	12 (63%)	2 (10%)
All	All	1527/1564 (97%)	459 (30%)	101 (6%)

5 of 459 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	6	G
1	A	7	G
1	A	8	A
1	A	9	G
1	A	13	U

5 of 101 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	748	C
1	A	992	U
1	A	1452	C
1	A	820	U
1	A	897	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](#)

Of 129 ligands modelled in this entry, 129 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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
1	A	4

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	84:U	O3'	88:A	P	5.75
1	A	1442(A):G	O3'	1442(B):A	P	5.10
1	A	841:U	O3'	848:C	P	4.14
1	A	204:U	O3'	216:G	P	3.78