



wwPDB X-ray Structure Validation Summary Report ⓘ

May 28, 2020 – 09:53 pm BST

PDB ID : 2E5L
Title : A snapshot of the 30S ribosomal subunit capturing mRNA via the Shine-Dalgarno interaction
Authors : Kaminishi, T.; Wilson, D.N.; Takemoto, C.; Harms, J.M.; Kawazoe, M.; Schlutzen, F.; Hanawa-Suetsugu, K.; Shirouzu, M.; Fucini, P.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on : 2006-12-21
Resolution : 3.30 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

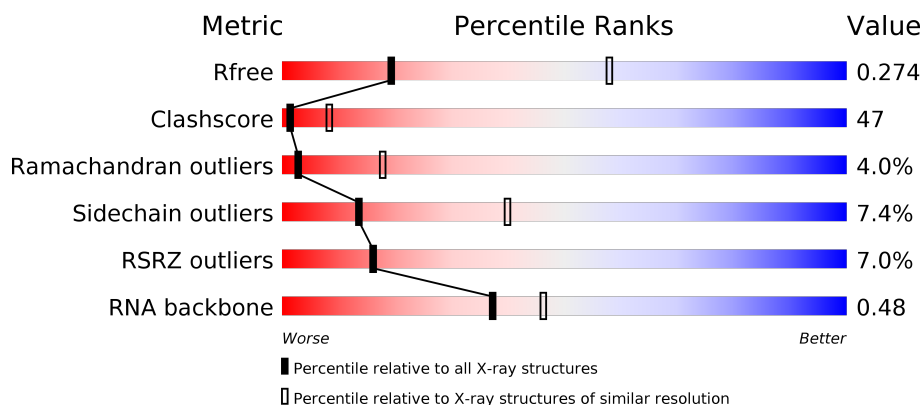
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 Å.

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



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)
RNA backbone	3102	1117 (3.70-2.90)

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 respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	1520	<div> <div></div> <div> <div>10%</div> <div>60%</div> <div>21%</div> <div>8%</div> </div> </div>
2	1	6	<div> <div>33%</div> <div>17%</div> <div>50%</div> <div>33%</div> </div>
2	2	6	<div> <div>17%</div> <div>33%</div> <div>33%</div> <div>33%</div> </div>
3	B	227	<div> <div></div> <div> <div>52%</div> <div>37%</div> <div>7%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	C	238	
5	D	208	
6	E	161	
7	F	101	
8	G	155	
9	H	138	
10	I	128	
11	J	104	
12	K	128	
13	L	131	
14	M	125	
15	N	60	
16	O	88	
17	P	88	
18	Q	104	
19	R	87	
20	S	92	
21	T	105	
22	V	26	

2 Entry composition

There are 23 unique types of molecules in this entry. The entry contains 51895 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a RNA chain called 16S ribosomal RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1517	Total	C	N	O	P	0	0	0
			32594	14508	6027	10542	1517			

- Molecule 2 is a RNA chain called 5'-R(*GP*AP*AP*AP*GP*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	1	6	Total	C	N	O	P	0	0	0
			131	60	30	36	5			
2	2	4	Total	C	N	O	P	0	0	0
			86	40	20	23	3			

- Molecule 3 is a protein called 30S ribosomal protein S2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	222	Total	C	N	O	S	0	0	0
			1811	1154	328	324	5			

- Molecule 4 is a protein called 30S ribosomal protein S3.

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

- Molecule 5 is a protein called 30S ribosomal protein S4.

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

- Molecule 6 is a protein called 30S ribosomal protein S5.

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

- Molecule 7 is a protein called 30S ribosomal protein S6.

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

- Molecule 8 is a protein called 30S ribosomal protein S7.

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

- Molecule 9 is a protein called 30S ribosomal protein S8.

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

- Molecule 10 is a protein called 30S ribosomal protein S9.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
10	I	127	Total	C	N	O	0	0	0
			1011	639	198	174			

- Molecule 11 is a protein called 30S ribosomal protein S10.

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

- Molecule 12 is a protein called 30S ribosomal protein S11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	K	115	Total	C	N	O	S	0	0	0
			853	531	160	159	3			

- Molecule 13 is a protein called 30S ribosomal protein S12.

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

- Molecule 14 is a protein called 30S ribosomal protein S13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	M	122	Total	C	N	O	S	0	0	0
			969	600	200	167	2			

- Molecule 15 is a protein called 30S ribosomal protein S14.

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

- Molecule 16 is a protein called 30S ribosomal protein S15.

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

- Molecule 17 is a protein called 30S ribosomal protein S16.

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

- Molecule 18 is a protein called 30S ribosomal protein S17.

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

- Molecule 19 is a protein called 30S ribosomal protein S18.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
19	R	73	Total	C	N	O	0	0	0
			597	380	118	99			

- Molecule 20 is a protein called 30S ribosomal protein S19.

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

- Molecule 21 is a protein called 30S ribosomal protein S20.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
21	T	99	Total	C	N	O	S	0	0	0
			762	469	162	129	2			

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

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

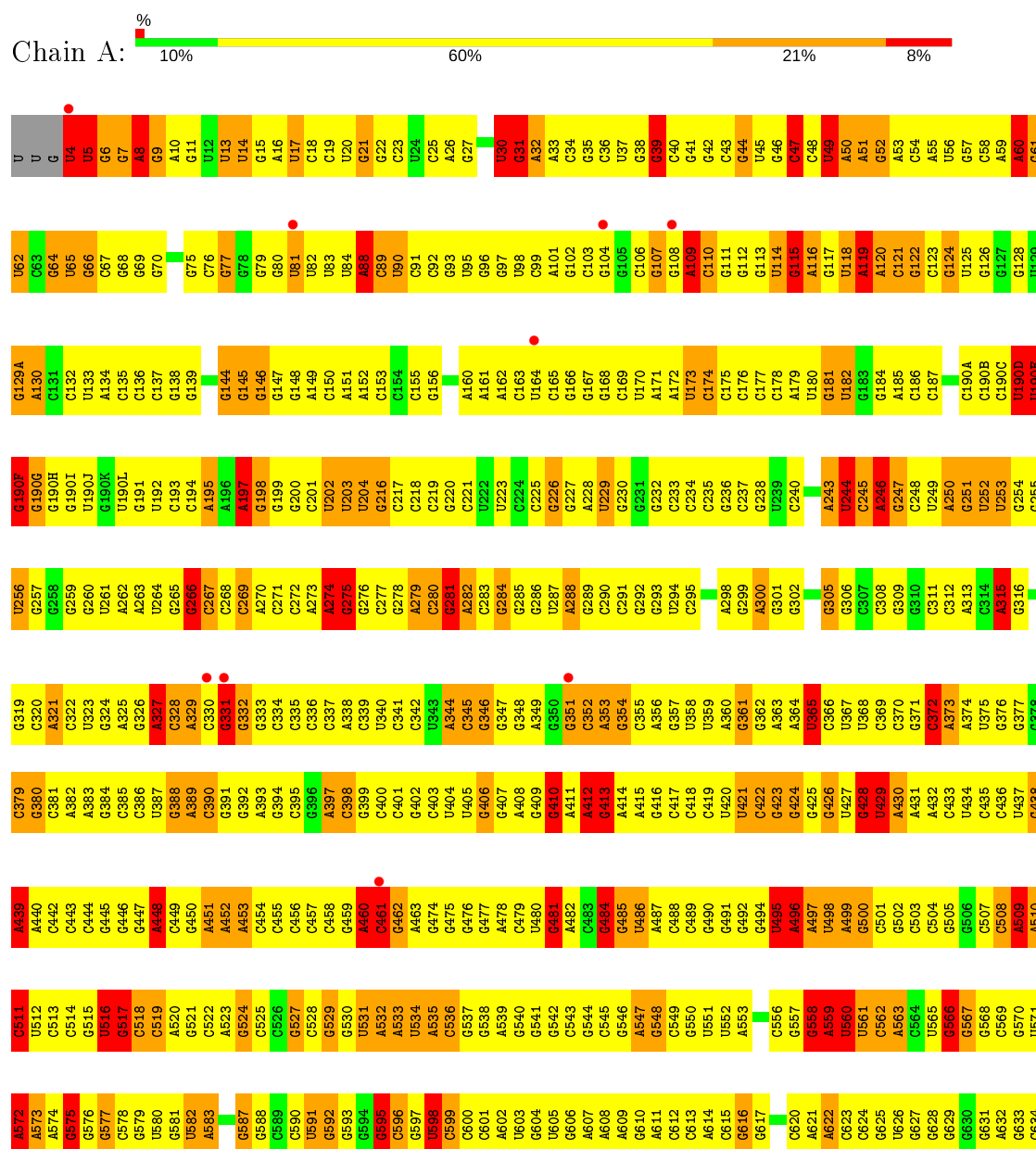
- Molecule 23 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

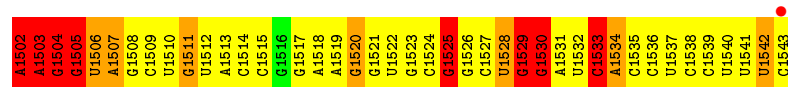
3 Residue-property plots

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

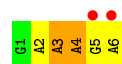
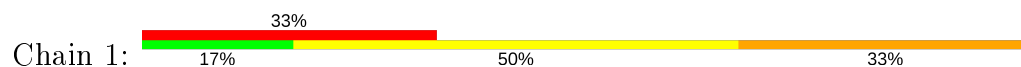
• Molecule 1: 16S ribosomal RNA



U1374	A1374	U1315	G1254	G1193	A1130	G1068	G1009	C948	G888	U820	C756	A695	G635
A1375	A1375	G1316	G1255	U1194	G1131	C1069	G1010	A949	A889	G821	U757	A696	U636
U1376	U1376	C1317	A1256	C1195			G1011	U950	G890	G822	U758	G697	U637
A1377	A1377	A1318	U1257	U1196	G1134	G1072	G1012	G951	U891	C823	A759	G698	G638
C1378	C1378	A1319	G1258	G1197	U1135	U1073	G1013	U952	U892	A824	G760	G699	G639
G1440	G1379	A1320	C1259	G1198	U1136	G1074	A1014	G953	C893	G825	G761	G700	A640
G1441	G1441	C1320	G1260	G1199	C1137	C1075	A1015	G954	G894	C826	C762	G701	U641
G1442	U1380	C1321	A1261	C1200	G1138	G1076	A1016	U955	G895	U827		A702	A642
G1443	U1381	G1322	C1262	A1201	G1139	G1077	G1017	U956	C896	A828	G765	G703	C643
A1446	C1382	C1323	C1263	G1202	C1140	U1078	C1018	U957	G897	G829	A766	A704	G644
G1447		C1324	C1264	C1203		G1079	G1019	G958	G898	G830	A767	U705	C645
G1448	G1385	C1325	G1265	G1204	G1143	A1080	U1021	A959	C899	U831	A768	A706	U646
U1450	G1386	C1326	G1266	U1205	G1144	G1081	G1022	U960	A900	U832	A769	C707	C647
A1451	G1387	C1327	C1267	G1206	C1145	G1082	G1023	U961	A901	U833	C770	C708	A648
C1452	C1388	C1328	A1268	G1207	A1146	U1083	G1024	G962	G902		G771	G709	G649
G1453	C1389	U1329	C1269		C1147	U1084	G1025	G963	G903	G837	U772	A710	G650
G1454	U1390	C1330	G1270	C1207	U1148	U1085	U1026	A964	C904	G838	G773	A712	U652
G1455	G1391	G1331	C1271	C1210	U1149	U1086	G1027	A965	U905	U839	G774	G713	A653
G1456	U1392	A1332	G1272	U1211	C1149	U1087	C1028	G966	G906	U840	G775	G714	G654
G1459	U1393	A1333	G1273	U1212	U1150	G1088	G1029	C967	A907	U841	A777	A715	A655
A1460	A1394	C1334	G1274	A1213	A1151	G1089	C1030	A968	A908	U842	G778	A716	C656
G1461	C1395	G1335		G1214	A1152	U1090	G1030A	A969	A909	C848	C779	C717	G657
G1462	C1396	C1336		G1215	C1153	U1091	G1030B	C970	A910	C849	C780	G718	G658
G1463	C1397	G1337	U1278	G1216	G1154	A1092	G1030C	G971	U911	U850	A781	C719	U659
G1464	A1398	C1338	A1279	C1217	G1155	A1093	A1030D	A972	C912	G851	A782	C720	G660
G1465	G1399	A1339	A1280	U1219	A1156	G1094	G1037	C973	A913	G852			G661
G1466	C1400	A1340	U1281	G1220	A1157	U1095	G1038	C979	A914	G853	G783	A722	G662
G1467	G1401	U1341	C1282	G1221	C1158	C1096	C1039	C980	A915	G854	G784	U723	A663
G1469	A1402	C1342	G1283	G1222	U1159	C1097	U1040	U981	G916	G855	G785	G724	G664
G1470	C1403	G1343	C1284	C1223	G1160	C1098	G1034	A977	G917	C857	A786	G725	A665
G1471	G1405	U1345	A1285	G1224	C1161	G1099		A978	A918	U857	A787	G726	G666
U1472	U1406	A1346	A1286	A1225	C1162	C1100		C979	A919	A859	U788	C727	G667
A1473	A1407	C1347	C1287	G1226	G1163	A1101	C1037	C980	U920	A860	U789	A728	G668
G1475	A1408	U1348	A1288	C1227	C1164	A1102	C1039	U982	U921	G861	G791		G669
G1476		A1349	A1289	C1228	C1165	C1103	U1040	A983	G922	U863	A792	G731	G670
G1477	C1411	A1350	G1290	A1229	A1167	G1104		C984	C923	U864	A793	C732	G671
	A1412	U1351	G1291	C1230	A1168	A1105	C1045	G985	G925	A865	A794	A733	U672
	A1413	C1352	U1292	G1231	A1169	C1106	C1046	A986	G926	C866	C795	G734	G673
	U1414	G1353	G1293	U1232	G1171	C1107	A1047	G987	G927	C867	C796	C735	G674
	G1415	C1354	G1294	G1233	C1172	G1108	G1048	G988	G928	C868	C797	C736	A675
	G1416	G1355	G1295	C1234	G1173	C1109	U1049	C989	G929	G869		A737	A676
	G1417	G1356	U1235	U1235	G1174	A1110	G1050	C990	C930	U870	U801	C738	U677
	A1418	A1357	C1297	C1236	G1175	A1111	U1051	U991	C931	U871	A802	C739	U678
	G1419	U1358	C1298	G1237	A1176	G1112	U1052	U992	C932	A872	G803	U740	C679
	C1420	C1359	A1299	A1238	G1177	C1113	G1053	G993	G933	A873	U804	G741	C680
	G1421	A1360	G1300	U1239	G1178		A1054	A994	C934	G874	C805	G742	C681
	G1422	G1361	U1301	U1240	A1179	C1108	A1055	C995	A935	C875	C806	U743	G682
	G1423	C1361A	U1302	G1241	A1180	C1117	U1056	A996	C936	G876	A807	C744	G683
	C1424	C1362	C1303	G1242	G1181	C1118	G1057	U997	C937	C877	C808	C745	A684
	U1425	A1363	G1304	C1243	G1182	C1119	G1058	G998	A938	G878		A746	G685
	A1426	U1364	G1305	C1244	A1183	G1120					C811	C747	U686
	U1427	G1365	A1306	A1245	G1184	U1121	C1059		G939	C879	C812	C748	A687
	A1428	C1366	U1307	C1246	G1185	U1122	C1060	G1002	C940	G880	C813	C749	G688
	C1429	C1367	U1308	C1247	G1186	A1123	G1061	G1003	G941	C881	U813	G750	C689
	C1430	G1368	G1309	U1247	G1187	G1124	U1062	G1003A	G942	C882	A814	G751	G690
	C1369	C1369	G1310	A1248	A1188	U1125	C1063	A1004	U943	U883	A815	U751	G691
	G1431	C1369	G1310	C1249	C1189	U1126	G1064	A1005	G944	U884	A816	G752	G692
	G1432	G1370	G1311	A1250	C1189	G1127	U1065	C1007	G945	G885	C817	A753	G693
	A1433	U1371	G1312	A1251	G1190	G1128	C1066	C1007	G946	G886	C818	C754	U694
	U1499	U1372	U1313	U1372	A1191	C1129	A1067	C1008	G947	G887	A819		A694
C1501		G1373	C1314	G1253	C1192								



- Molecule 2: 5'-R(*GP*AP*AP*AP*GP*A)-3'



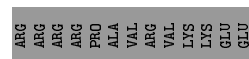
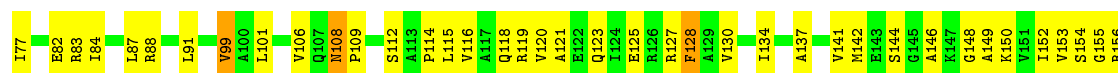
- Molecule 2: 5'-R(*GP*AP*AP*AP*GP*A)-3'



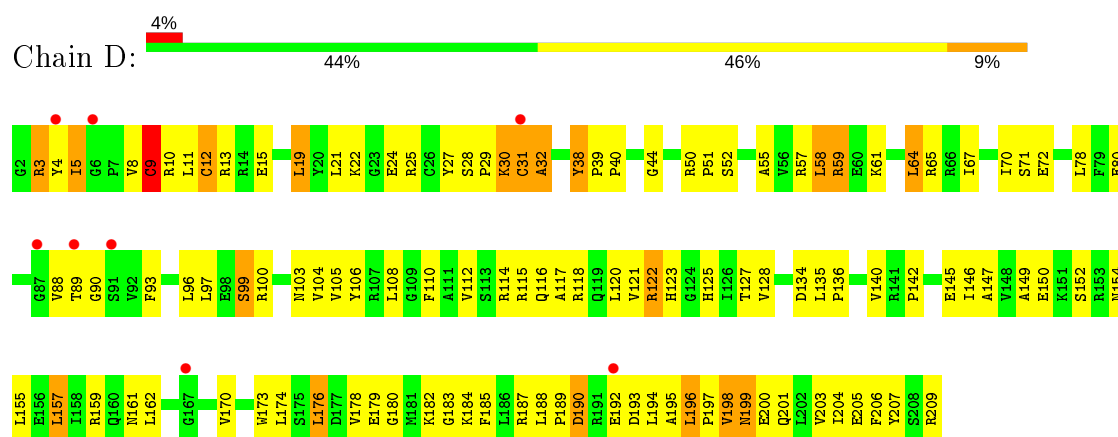
- Molecule 3: 30S ribosomal protein S2



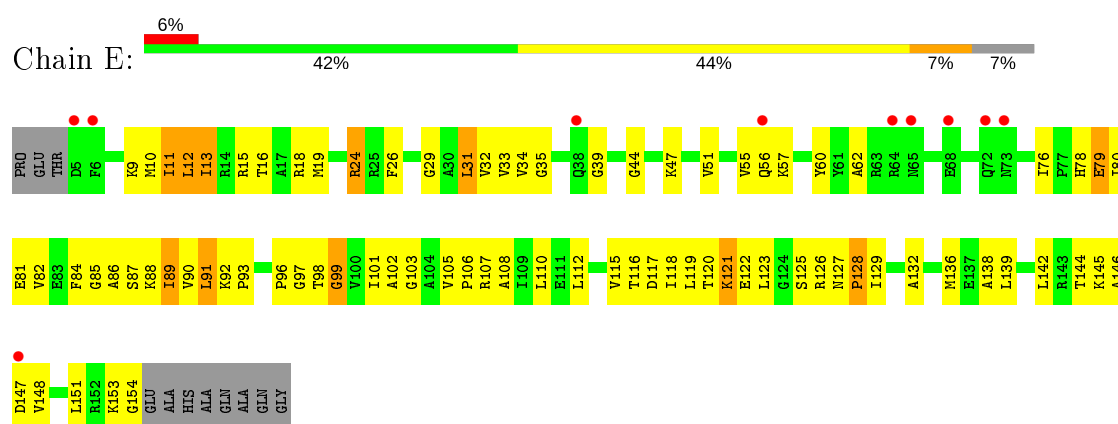
- Molecule 4: 30S ribosomal protein S3



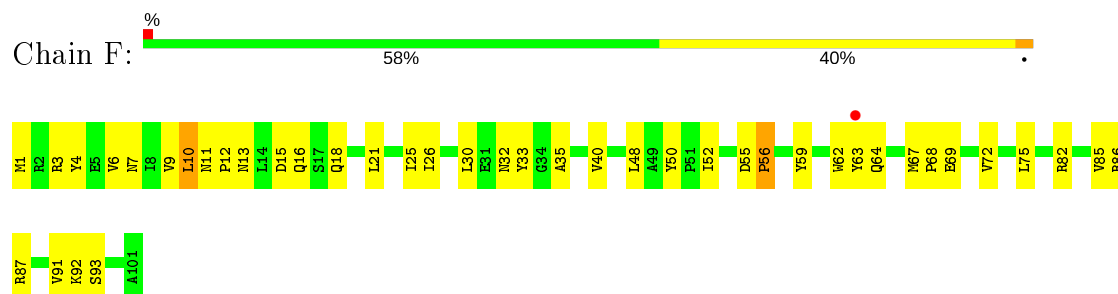
- Molecule 5: 30S ribosomal protein S4



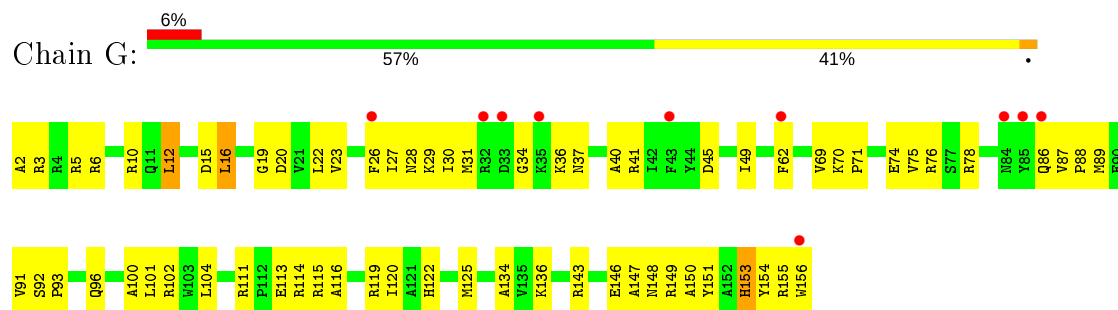
- Molecule 6: 30S ribosomal protein S5



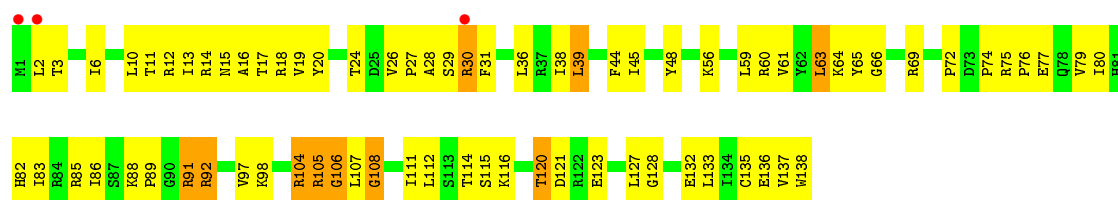
- Molecule 7: 30S ribosomal protein S6



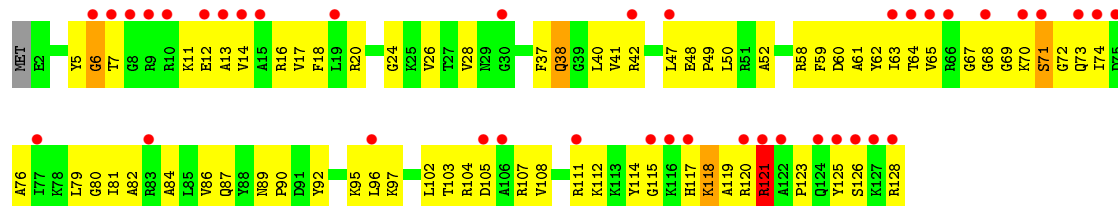
- Molecule 8: 30S ribosomal protein S7



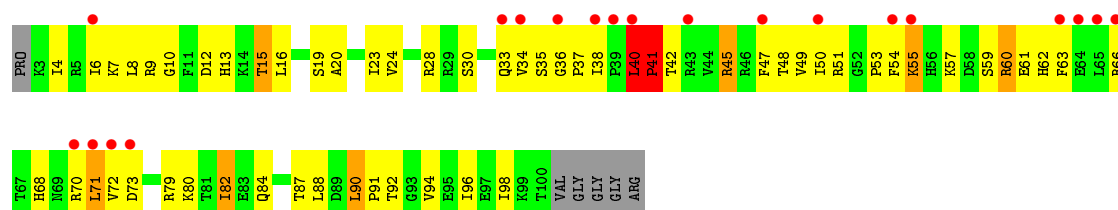
- Molecule 9: 30S ribosomal protein S8



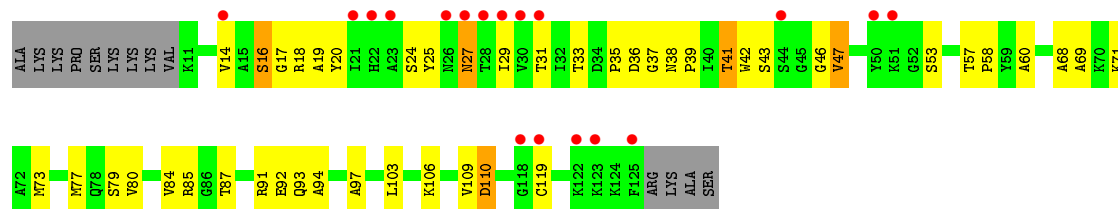
• Molecule 10: 30S ribosomal protein S9



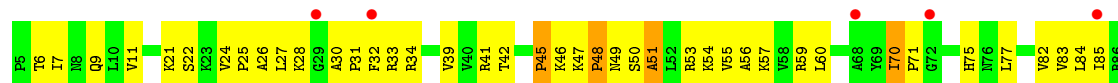
• Molecule 11: 30S ribosomal protein S10

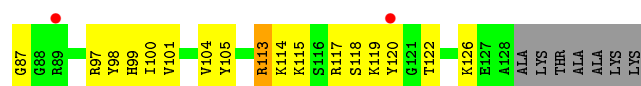


• Molecule 12: 30S ribosomal protein S11

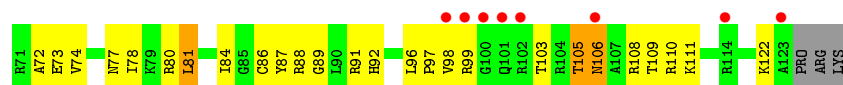


• Molecule 13: 30S ribosomal protein S12

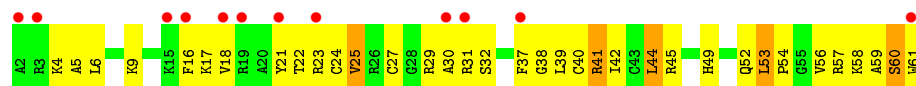




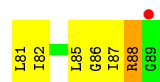
• Molecule 14: 30S ribosomal protein S13



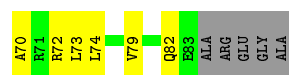
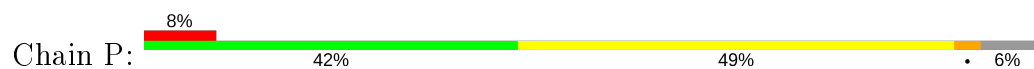
• Molecule 15: 30S ribosomal protein S14



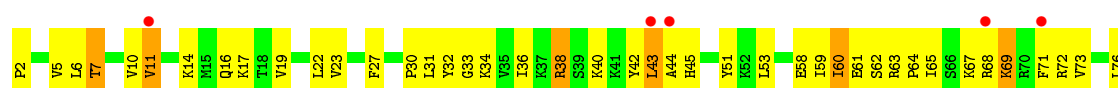
• Molecule 16: 30S ribosomal protein S15

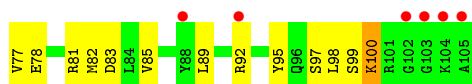


• Molecule 17: 30S ribosomal protein S16



• Molecule 18: 30S ribosomal protein S17





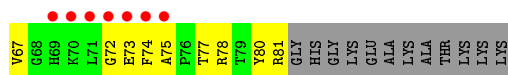
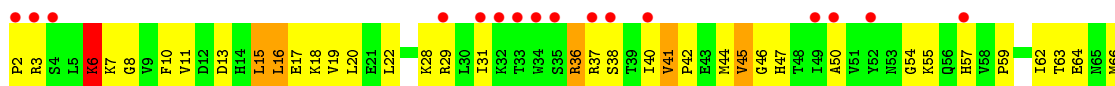
- Molecule 19: 30S ribosomal protein S18

Chain R: 44% 37% 1% 16%



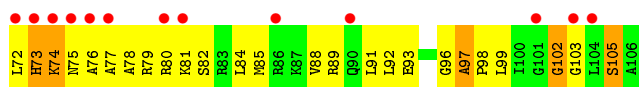
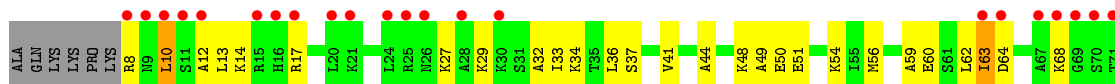
- Molecule 20: 30S ribosomal protein S19

Chain S: 25% 37% 43% 5% 13%



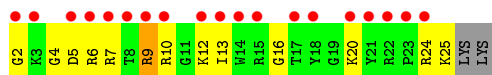
- Molecule 21: 30S ribosomal protein S20

Chain T: 33% 45% 43% 7% 6%



- Molecule 22: 30S ribosomal protein Thx

Chain V: 73% 42% 46% 8%



4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α , β , γ	411.79Å 411.79Å 173.16Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	148.83 – 3.30 148.83 – 3.30	Depositor EDS
% Data completeness (in resolution range)	97.0 (148.83-3.30) 97.1 (148.83-3.30)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.23 (at 3.33Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.259 , 0.301 0.231 , 0.274	Depositor DCC
R_{free} test set	10897 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å ²)	96.4	Xtriage
Anisotropy	0.310	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.18 , 82.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.38$, $\langle L^2 \rangle = 0.21$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.82	EDS
Total number of atoms	51895	wwPDB-VP
Average B, all atoms (Å ²)	104.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.54% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	0.97	19/36482 (0.1%)	0.99	145/56937 (0.3%)
2	1	0.51	0/148	0.76	0/230
2	2	0.42	0/97	0.73	0/150
3	B	0.67	0/1843	0.92	5/2479 (0.2%)
4	C	0.63	0/1636	0.89	2/2205 (0.1%)
5	D	0.80	3/1733 (0.2%)	0.96	8/2318 (0.3%)
6	E	0.82	0/1162	0.95	2/1564 (0.1%)
7	F	0.52	0/856	0.78	0/1154
8	G	0.50	0/1276	0.67	0/1709
9	H	0.76	0/1136	1.00	2/1527 (0.1%)
10	I	0.53	0/1029	0.78	0/1378
11	J	0.57	0/807	0.89	3/1085 (0.3%)
12	K	0.53	0/868	0.79	0/1173
13	L	0.62	0/986	0.85	0/1320
14	M	0.53	0/979	0.78	0/1310
15	N	0.66	0/501	0.93	1/664 (0.2%)
16	O	0.61	0/745	0.87	0/992
17	P	0.62	0/716	0.83	0/963
18	Q	0.74	0/870	0.92	1/1159 (0.1%)
19	R	0.59	0/603	0.86	0/799
20	S	0.51	0/661	0.82	0/890
21	T	0.49	0/764	0.73	0/1006
22	V	0.56	0/212	0.72	0/277
All	All	0.86	22/56110 (0.0%)	0.95	169/83289 (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	1	127

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	12	CYS	CB-SG	9.54	1.98	1.82
5	D	12	CYS	CA-CB	8.62	1.73	1.53
1	A	1108	G	C5-C6	7.41	1.49	1.42
1	A	660	G	C5-C6	-6.55	1.35	1.42
1	A	361	G	C5-C6	-6.41	1.35	1.42

The worst 5 of 169 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	12	CYS	CA-CB-SG	14.39	139.90	114.00
1	A	511	C	N1-C1'-C2'	12.07	129.70	114.00
1	A	934	C	N1-C1'-C2'	9.87	126.84	114.00
1	A	246	A	N9-C1'-C2'	9.69	126.60	114.00
1	A	1151	A	N9-C1'-C2'	9.47	126.31	114.00

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	511	C	C1'

5 of 127 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	30	U	Sidechain
1	A	39	G	Sidechain
1	A	47	C	Sidechain
1	A	49	U	Sidechain
1	A	60	A	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	32594	0	16454	3168	0
2	1	131	0	68	14	0
2	2	86	0	46	9	0
3	B	1811	0	1861	95	0
4	C	1612	0	1677	130	0
5	D	1703	0	1763	117	0
6	E	1146	0	1207	93	0
7	F	843	0	857	36	0
8	G	1257	0	1296	81	0
9	H	1116	0	1177	98	0
10	I	1011	0	1043	104	0
11	J	794	0	840	80	0
12	K	853	0	868	54	0
13	L	970	0	1057	75	0
14	M	969	0	1039	78	0
15	N	492	0	529	52	0
16	O	734	0	771	46	0
17	P	700	0	720	52	0
18	Q	857	0	930	53	0
19	R	597	0	668	43	0
20	S	647	0	673	61	0
21	T	762	0	859	48	0
22	V	208	0	221	14	0
23	D	1	0	0	0	0
23	N	1	0	0	0	0
All	All	51895	0	36624	4157	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 47.

The worst 5 of 4157 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:1027:C:C2'	1:A:1028:C:H5''	1.48	1.40
1:A:390:C:H4'	17:P:28:ARG:NH2	1.46	1.28
1:A:1027:C:H2'	1:A:1028:C:C5'	1.65	1.25
1:A:839:U:H5'	1:A:840:C:C5	1.71	1.24
1:A:243:A:H4'	1:A:244:U:C5'	1.65	1.24

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	
3	B	220/227 (97%)	169 (77%)	39 (18%)	12 (6%)	2	11
4	C	204/238 (86%)	149 (73%)	42 (21%)	13 (6%)	1	9
5	D	206/208 (99%)	165 (80%)	31 (15%)	10 (5%)	2	14
6	E	148/161 (92%)	113 (76%)	30 (20%)	5 (3%)	3	22
7	F	99/101 (98%)	83 (84%)	14 (14%)	2 (2%)	7	32
8	G	153/155 (99%)	129 (84%)	23 (15%)	1 (1%)	22	54
9	H	136/138 (99%)	113 (83%)	21 (15%)	2 (2%)	10	38
10	I	125/128 (98%)	94 (75%)	25 (20%)	6 (5%)	2	14
11	J	96/104 (92%)	75 (78%)	14 (15%)	7 (7%)	1	7
12	K	113/128 (88%)	88 (78%)	22 (20%)	3 (3%)	5	26
13	L	122/131 (93%)	96 (79%)	21 (17%)	5 (4%)	3	17
14	M	120/125 (96%)	89 (74%)	26 (22%)	5 (4%)	3	17
15	N	58/60 (97%)	45 (78%)	13 (22%)	0	100	100
16	O	86/88 (98%)	70 (81%)	14 (16%)	2 (2%)	6	29
17	P	81/88 (92%)	64 (79%)	16 (20%)	1 (1%)	13	42
18	Q	102/104 (98%)	86 (84%)	11 (11%)	5 (5%)	2	14
19	R	71/87 (82%)	57 (80%)	13 (18%)	1 (1%)	11	38
20	S	78/92 (85%)	63 (81%)	11 (14%)	4 (5%)	2	13
21	T	97/105 (92%)	72 (74%)	17 (18%)	8 (8%)	1	5
22	V	22/26 (85%)	19 (86%)	1 (4%)	2 (9%)	1	4
All	All	2337/2494 (94%)	1839 (79%)	404 (17%)	94 (4%)	3	18

5 of 94 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	B	12	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	B	21	ARG
3	B	24	TRP
3	B	130	ARG
4	C	4	LYS

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	
3	B	191/196 (97%)	170 (89%)	21 (11%)	6	24
4	C	160/187 (86%)	146 (91%)	14 (9%)	10	33
5	D	180/180 (100%)	163 (91%)	17 (9%)	8	30
6	E	115/122 (94%)	103 (90%)	12 (10%)	7	25
7	F	90/90 (100%)	87 (97%)	3 (3%)	38	66
8	G	126/126 (100%)	122 (97%)	4 (3%)	39	67
9	H	119/119 (100%)	110 (92%)	9 (8%)	13	39
10	I	98/99 (99%)	91 (93%)	7 (7%)	14	42
11	J	88/91 (97%)	82 (93%)	6 (7%)	16	44
12	K	87/98 (89%)	80 (92%)	7 (8%)	12	37
13	L	104/108 (96%)	100 (96%)	4 (4%)	33	62
14	M	97/100 (97%)	90 (93%)	7 (7%)	14	41
15	N	49/49 (100%)	44 (90%)	5 (10%)	7	27
16	O	79/79 (100%)	71 (90%)	8 (10%)	7	27
17	P	72/74 (97%)	67 (93%)	5 (7%)	15	44
18	Q	96/96 (100%)	89 (93%)	7 (7%)	14	41
19	R	64/76 (84%)	62 (97%)	2 (3%)	40	67
20	S	71/79 (90%)	65 (92%)	6 (8%)	10	35
21	T	76/81 (94%)	73 (96%)	3 (4%)	32	62
22	V	19/21 (90%)	19 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	1981/2071 (96%)	1834 (93%)	147 (7%)	13	40

5 of 147 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
7	F	86	ARG
10	I	60	ASP
18	Q	100	LYS
8	G	16	LEU
9	H	91	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
8	G	106	GLN
10	I	73	GLN
18	Q	16	GLN
9	H	15	ASN
9	H	82	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1517/1520 (99%)	331 (21%)	187 (12%)
2	1	5/6 (83%)	1 (20%)	1 (20%)
2	2	3/6 (50%)	2 (66%)	0
All	All	1525/1532 (99%)	334 (21%)	188 (12%)

5 of 334 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	5	U
1	A	6	G
1	A	7	G
1	A	8	A
1	A	9	G

5 of 188 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	703	G
1	A	884	U
1	A	1397	C
1	A	721	G
1	A	817	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 2 ligands modelled in this entry, 2 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](#)

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	1517/1520 (99%)	-0.02	21 (1%) 75 75	44, 89, 178, 199	0
2	1	6/6 (100%)	0.76	2 (33%) 0 0	199, 199, 199, 199	0
2	2	4/6 (66%)	2.20	1 (25%) 0 0	185, 193, 195, 198	0
3	B	222/227 (97%)	0.22	3 (1%) 75 75	46, 104, 169, 199	0
4	C	206/238 (86%)	0.10	4 (1%) 66 65	49, 107, 172, 198	0
5	D	208/208 (100%)	0.35	8 (3%) 40 37	32, 90, 156, 199	0
6	E	150/161 (93%)	0.58	10 (6%) 17 17	32, 72, 151, 195	0
7	F	101/101 (100%)	-0.13	1 (0%) 82 82	63, 116, 167, 182	0
8	G	155/155 (100%)	0.07	10 (6%) 18 18	70, 133, 184, 199	0
9	H	138/138 (100%)	0.20	3 (2%) 62 60	31, 72, 145, 181	0
10	I	127/128 (99%)	1.30	40 (31%) 0 0	55, 147, 191, 199	0
11	J	98/104 (94%)	1.01	20 (20%) 1 1	64, 138, 198, 199	0
12	K	115/128 (89%)	0.51	18 (15%) 2 2	59, 111, 172, 190	0
13	L	124/131 (94%)	0.42	7 (5%) 24 23	46, 104, 165, 199	0
14	M	122/125 (97%)	0.64	18 (14%) 2 2	71, 127, 180, 198	0
15	N	60/60 (100%)	1.09	12 (20%) 1 1	56, 89, 158, 190	0
16	O	88/88 (100%)	0.07	2 (2%) 60 59	45, 100, 167, 185	0
17	P	83/88 (94%)	0.66	7 (8%) 11 10	38, 91, 146, 185	0
18	Q	104/104 (100%)	0.86	11 (10%) 6 6	49, 90, 172, 199	0
19	R	73/87 (83%)	0.14	0 100 100	46, 103, 175, 199	0
20	S	80/92 (86%)	1.29	23 (28%) 0 0	74, 136, 187, 199	0
21	T	99/105 (94%)	1.59	35 (35%) 0 0	69, 122, 182, 199	0
22	V	24/26 (92%)	3.70	19 (79%) 0 0	72, 121, 168, 199	0
All	All	3904/4026 (96%)	0.32	275 (7%) 16 16	31, 99, 178, 199	0

The worst 5 of 275 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
18	Q	103	GLY	11.7
18	Q	104	LYS	11.4
21	T	73	HIS	11.3
20	S	3	ARG	11.1
20	S	2	PRO	9.0

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

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
23	ZN	D	210	1/1	0.97	0.34	85,85,85,85	0
23	ZN	N	62	1/1	0.98	0.10	87,87,87,87	0

6.5 Other polymers [i](#)

There are no such residues in this entry.