



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 02:19 am BST

PDB ID : 2ZM6  
Title : Crystal structure of the Thermus thermophilus 30S ribosomal subunit  
Authors : Kaminishi, T.; Wang, H.; Kawazoe, M.; Ishii, R.; Schlutzen, F.; Hanawa-Suetsugu, K.; Wilson, D.N.; Nomura, M.; Takemoto, C.; Shirouzu, M.; Fucini, P.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)  
Deposited on : 2008-04-11  
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](mailto: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

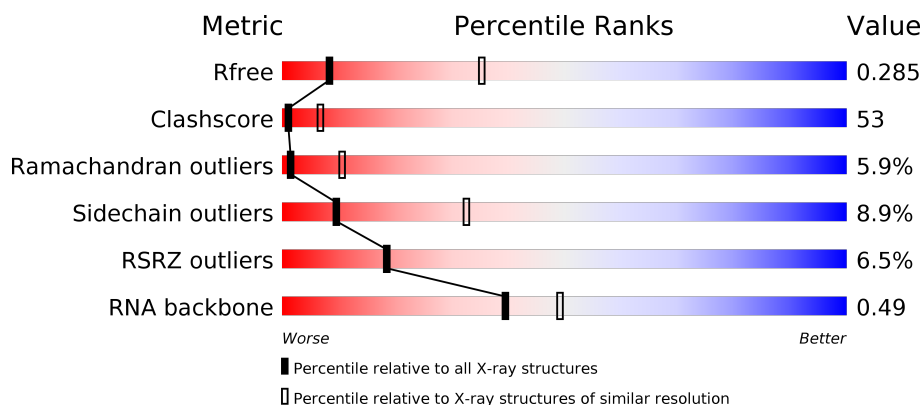
# 1 Overall quality at a glance

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  $\geq 3$ , 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	1509	<div> <div>9%</div> <div>59%</div> <div>21%</div> <div>10%</div> </div>
2	B	255	<div> <div>2%</div> <div>39%</div> <div>39%</div> <div>9%</div> <div>13%</div> </div>
3	C	238	<div> <div>6%</div> <div>33%</div> <div>42%</div> <div>11%</div> <div>13%</div> </div>
4	D	208	<div> <div>4%</div> <div>43%</div> <div>47%</div> <div>9%</div> </div>

*Continued on next page...*

Continued from previous page...

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

## 2 Entry composition

There are 22 unique types of molecules in this entry. The entry contains 51308 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	1506	Total	C	N	O	P	0	0	0
			32372	14408	5996	10462	1506			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	222	Total	C	N	O	S	0	0	0
			1810	1154	328	323	5			

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

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

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

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

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

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

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	153	Total	C	N	O	S	0	0	0
			1231	764	246	215	6			

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	125	Total	C	N	O	S	0	0	0
			993	629	195	169				

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

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

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

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

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	120	Total	C	N	O	S	0	0	0
			955	591	197	165	2			

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

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

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

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

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
17	Q	100	Total	C	N	O	S	0	0	0
			834	534	156	142	2			

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
18	R	68	Total	C	N	O	0	0	0
			559	357	109	93			

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
20	T	94	Total	C	N	O	S	0	0	0
			734	453	157	122	2			

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

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

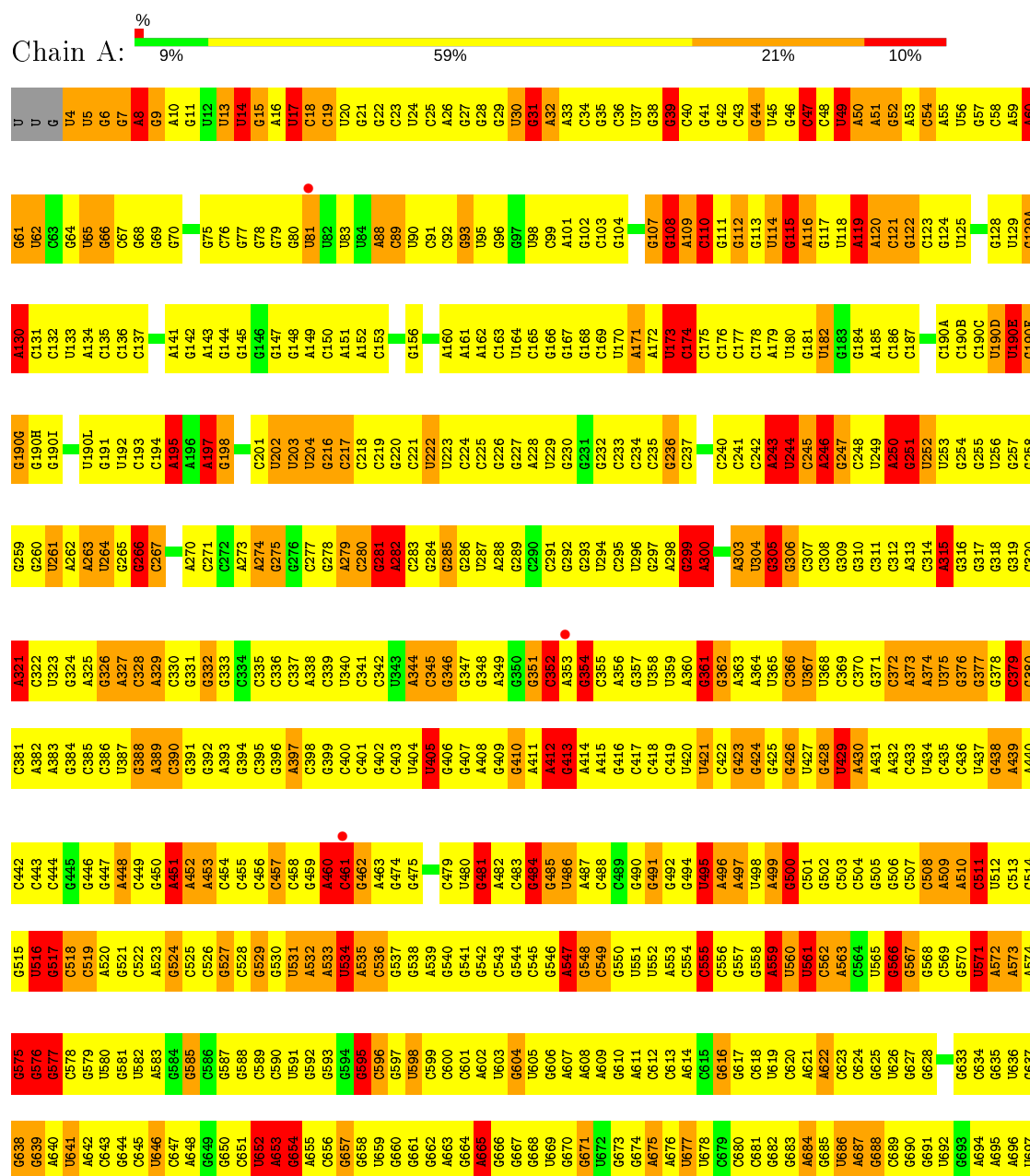
- Molecule 22 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
22	D	1	Total	Zn	0	0
			1	1		
22	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.

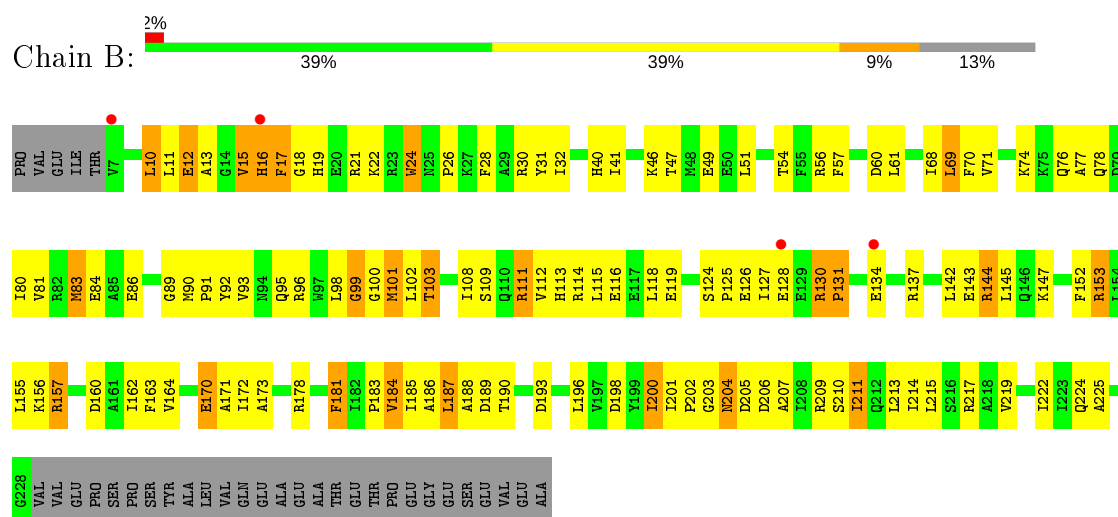
#### • Molecule 1: 16S ribosomal RNA



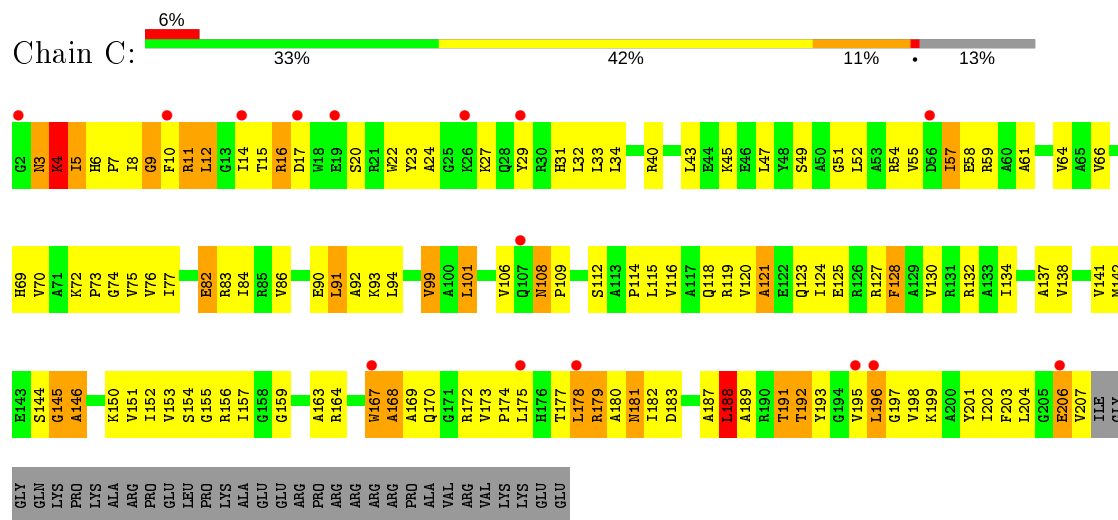


G1507	C1439	U1376	G1317	G1255	C1195	G1131	C1071	G1013	G951	U891	G821	A759	G698
G1508	C1440	A1377	A1318	A1256	U1196	C1132	G1072	A1014	U952	A892	C822	G760	C699
C1509	G1441	C1378	A1319	G1257	C1197	G1133	U1073	A1015	G953	C893	G823	G761	C701
U1510	G1442	G1379	C1320	G1258	G1198	G1134	G1074	A1016	G954	G894	C824	G762	G701
G1511	C1443	U1380	C1321	C1259	U1199	U1135	U1075	U955	U955	C895	G825	G763	A702
U1512	A1446	U1361	G1322	G1260	C1200	U1136	C1076	G1018	U956	C896	C826	G764	G703
A1513	C1447	C1362	G1323	A1261	A1201	C1137	G1077	C1019	U957	C897	U827	G765	A704
C1514	C1448	G1363	A1324	C1262	G1202	G1138	U1078	U1020	A958	G898	A828	A766	U705
C1515	G1449	G1365	C1325	C1263	C1203	G1139	G1079	G1021	A959	C899	G829	A767	A706
U1516	A1450	G1366	C1326	C1264	A1204	C1140	A1080	G1022	U960	A900	G830	A768	C707
A1517	G1451	G1367	C1327	G1265	U1205	C1141	G1081	G1023	U961	A901	G831	G769	C708
A1518	C1452	C1368	C1328	C1267	G1206	G1142	G1082	G1024	G962	G902	U835	G770	G709
A1519	G1453	C1369	A1329	A1268	G1207	G1143	U1083	U1025	G963	G903	G836	G771	G711
G1520	G1454	U1390	U1330	A1269	C1208	G1144	G1084	G1026	A964	C904	G837	G772	A712
G1521	G1455	U1391	G1331	C1270	C1209	C1145	U1085	C1027	A965	U905	G838	G773	G713
U1522	C1456	G1392	A1332	G1271	C1210	A1146	U1086	C1028	G966	G906	G839	G774	A714
G1523	A1460	U1393	A1333	G1272	U1211	C1147	G1087	G1029	A967	A907	U839	G775	G715
G1524	G1461	C1394	G1334	G1273	U1212	U1148	G1088	C1030	A968	A908	U840	A777	A716
G1525	G1462	C1395	C1335	G1274	A1213	C1149	U1089	G1030A	A969	A909	U841	G778	A717
G1526	C1463	A1396	C1336	C1275	G1214	U1150	U1090	C1030B	G970	G910	C848	G779	C717
C1527	G1464	C1397	C1337	C1277	A1151	A1152	U1091	G1030C	G971	U911	C849	G780	G718
C1465	G1465	A1398	G1338	U1278	C1216	G1153	C1098	A1030D	G972	C912	U850	A781	C719
C1466	C1466	C1399	A1339	A1279	C1217	G1154	G1099	G1031	G973	A913	U851	A782	C720
G1467	C1467	C1400	A1340	A1280	C1218	C1155	G1094	G1032	A974	A914	G852	C783	G721
G1470	G1470	G1401	U1341	U1281	U1219	C1156	U1095	G1033	A975	A915	G853	C784	G722
G1471	G1471	C1402	C1342	C1282	G1220	U1159	C1096	G1034	G976	G916	G854	G785	U723
G1472	C1472	C1403	G1343	G1283	G1221	G1160	C1097	C1037	A977	G917	G858	G786	G724
G1473	C1473	G1404	C1344	A1284	C1222	C1161	C1098	C1038	A978	A918	G859	U788	G725
G1474	G1474	G1405	U1345	A1285	C1223	G1162	G1099	C1039	C980	A919	U860	U789	C726
G1475	C1475	U1406	A1346	A1286	G1224	C1163	C1100	U1040	G981	U920	A860	G790	G727
G1476	C1476	C1407	G1347	A1287	A1225	G1164	A1101	U1041	U982	U921	G861	A791	A728
G1477	C1477	A1408	U1348	A1288	C1226	C1165	A1102	A1041	U983	G922	C862	G792	G731
C1479	C1479	G1409	A1349	A1289	A1227	G1166	C1103	A1044	A983	A923	U863	A792	C732
G1480	U1480	C1411	C1350	G1290	C1228	A1167	G1104	C924	C984	G924	A864	U793	A733
U1481	G1481	C1412	U1351	G1291	A1229	A1168	A1105	C1045	C985	G925	A865	A794	C734
G1482	C1482	A1413	C1352	U1292	C1230	A1169	G1106	A1046	A986	G926	C866	C795	G734
A1483	G1483	U1414	G1353	G1293	G1231	G1171	C1107	G1047	G987	G927	G867	C796	C735
C1484	U1484	G1415	C1354	G1294	U1232	C1172	G1108	G1048	G988	G928	C868	C797	C736
U1485	G1485	G1416	G1355	G1295	C1233	G1173	C1109	U1049	C989	G929	G869	G798	A737
G1486	G1486	G1417	C1356	C1296	C1234	G1174	A1110	C1050	C990	C930	U870	G799	C738
G1487	C1487	A1418	A1357	C1297	U1235	A1175	A1111	U1051	U991	C931	U871	G800	C739
G1488	U1488	G1419	U1358	C1298	A1236	G1176	C1112	U1052	U992	G932	A872	U801	U740
G1489	C1489	C1420	C1359	A1299	C1237	G1177	G1113	G1053	G993	G933	A873	A802	G741
G1490	C1490	G1421	A1360	G1300	A1238	G1178	C1114	C1054	A994	C934	G874	G803	G742
G1491	C1491	G1422	G1361	U1301	A1239	A1179	C1115	A1055	C995	A935	C875	U804	U743
A1492	A1492	G1423	C1361A	U1302	U1240	A1180	G1116	U1056	A996	C936	G876	C805	C744
A1493	G1493	C1424	C1362	C1303	G1241	G1181	C1117	G1057	U997	A937	C877	C806	C745
G1494	C1494	U1425	A1363	G1304	C1242	G1182	C1118	G1058	G998	A938	G878	A807	A746
U1495	U1495	C1426	U1364	G1305	G1243	A1183	C1119	C1059	G999	G939	C879	C808	G747
C1496	G1496	U1427	G1365	A1306	C1244	G1184	U1120	U1060	G999	C940	C880	G748	C748
G1497	G1497	C1427	C1366	U1307	A1245	G1185	U1121	G1061	G1003	G941	G881	C811	C749
U1498	C1498	C1430	C1367	U1308	C1246	G1186	U1122	G1062	G1004	G942	G882	C812	G750
A1499	C1499	C1431	G1368	G1309	U1247	G1187	A1123	C1063	A1005	U943	C883	U813	U751
A1500	C1500	G1432	C1369	G1310	A1248	G1188	G1124	G1064	C1006	G944	U884	A814	G752
C1501	C1501	A1433	G1370	G1311	C1249	A1189	U1125	U1065	C1007	G945	G885	A815	A753
A1502	C1502	G1434	G1371	G1312	U1250	G1190	U1126	C1066	C1008	A946	G886	A816	C754
A1503	C1503	G1435	U1372	U1313	A1251	A1191	G1127	A1067	G1009	G947	G887	C817	G755
G1504	G1504	U1436	C1373	C1314	A1252	C1192	C1128	G1068	G1010	G948	G888	G818	C756
G1505	C1505	C1437	A1374	U1315	G1253	C1193	C1129	G1069	G1011	A949	A889	A819	U757
U1506	U1506	G1438	A1375	G1316	C1254	U1194	A1130	U1070	U1012	U950	G890	U820	G758

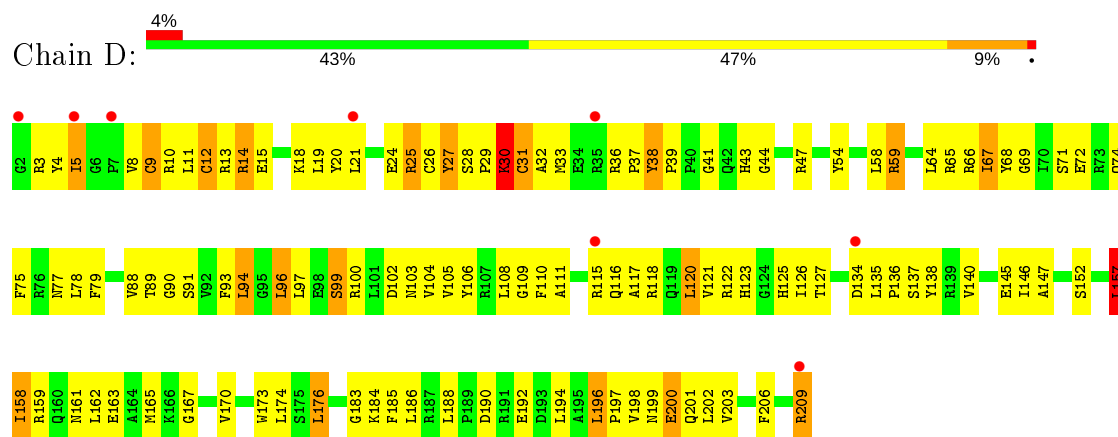
• Molecule 2: 30S ribosomal protein S2



- Molecule 3: 30S ribosomal protein S3

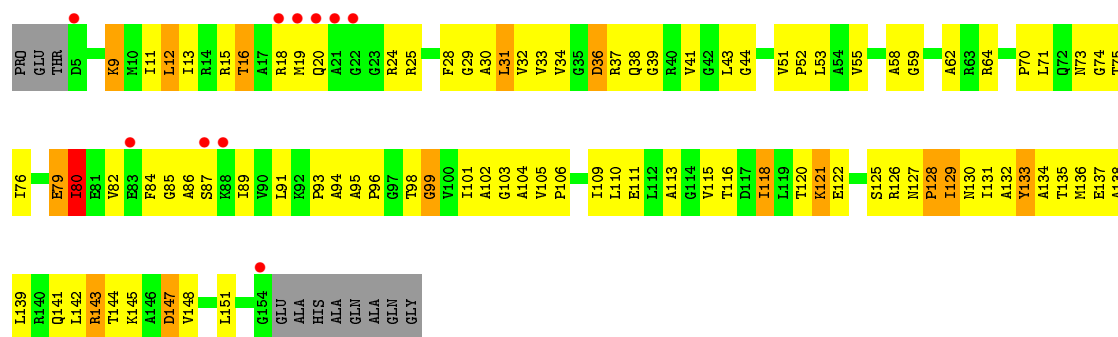


- Molecule 4: 30S ribosomal protein S4



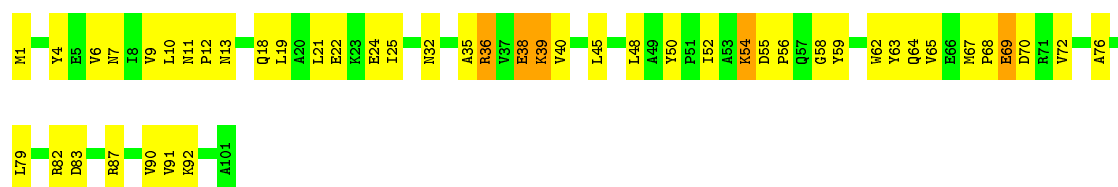
- Molecule 5: 30S ribosomal protein S5





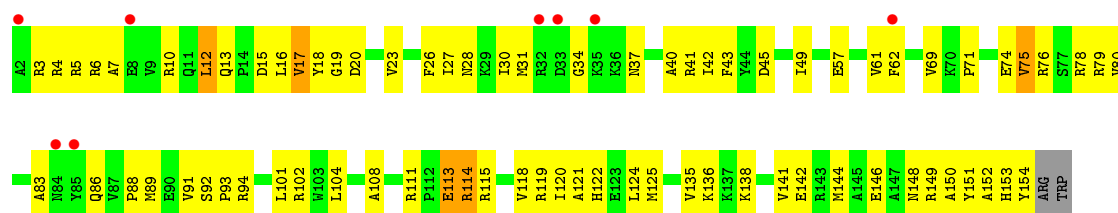
• Molecule 6: 30S ribosomal protein S6

Chain F: 53% 42% 5%



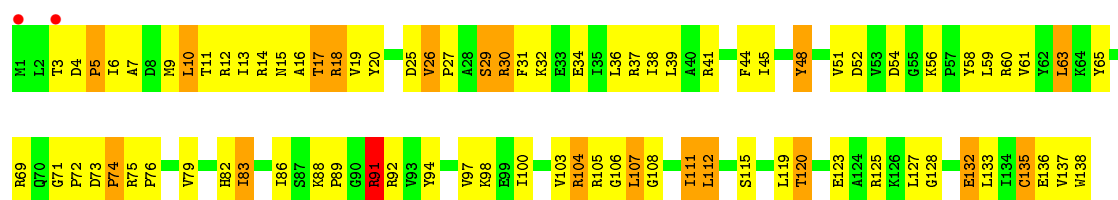
• Molecule 7: 30S ribosomal protein S7

Chain G: 5% 50% 46% ..



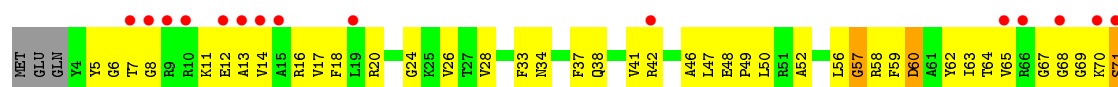
• Molecule 8: 30S ribosomal protein S8

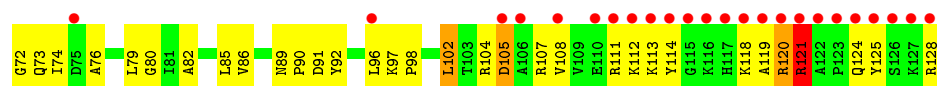
Chain H: 40% 46% 13% ..



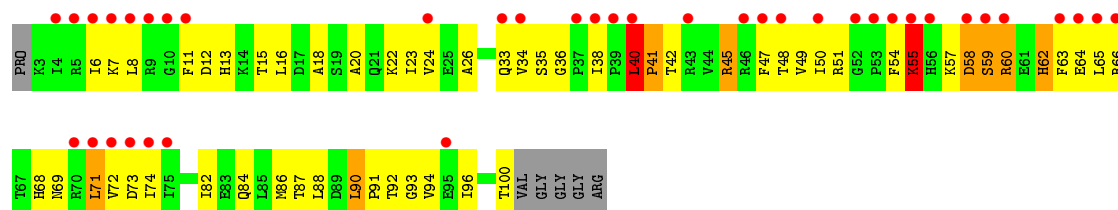
• Molecule 9: 30S ribosomal protein S9

Chain I: 30% 41% 52% 5% ..

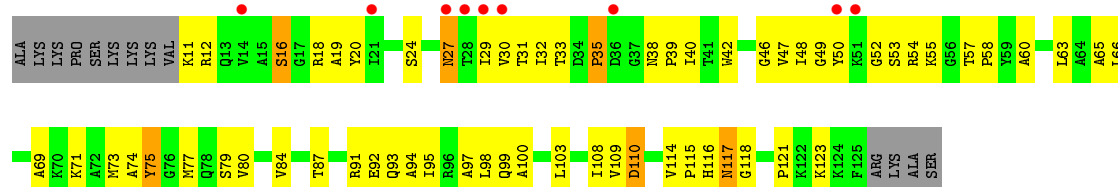




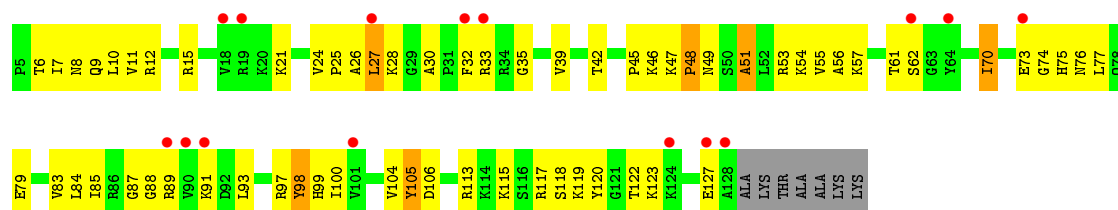
- Molecule 10: 30S ribosomal protein S10



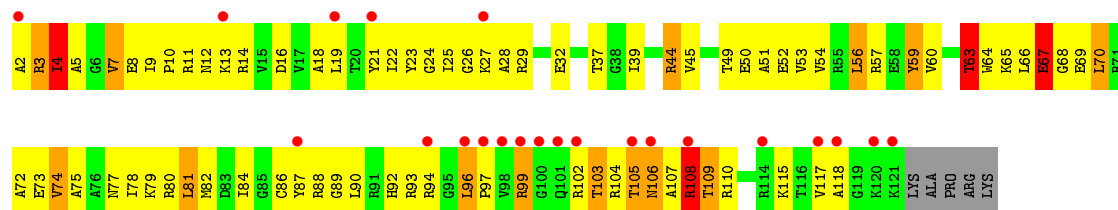
- Molecule 11: 30S ribosomal protein S11



- Molecule 12: 30S ribosomal protein S12

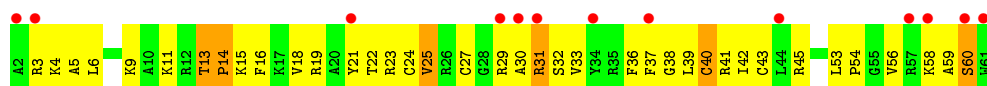


- Molecule 13: 30S ribosomal protein S13

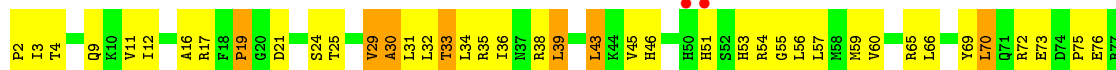


- Molecule 14: 30S ribosomal protein S14 type Z

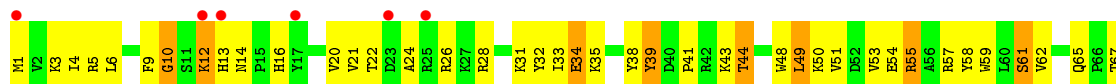
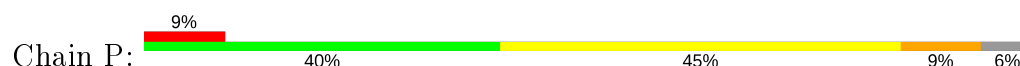




• Molecule 15: 30S ribosomal protein S15



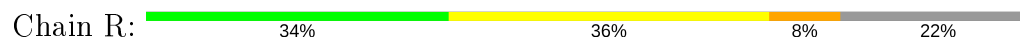
• Molecule 16: 30S ribosomal protein S16



• Molecule 17: 30S ribosomal protein S17

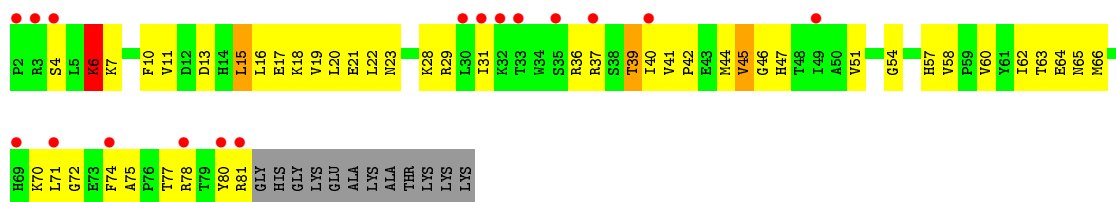


• Molecule 18: 30S ribosomal protein S18

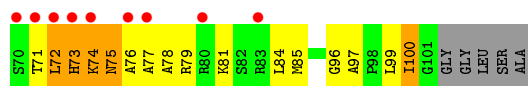
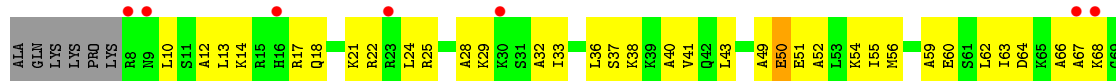


• Molecule 19: 30S ribosomal protein S19

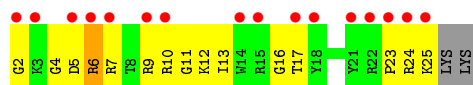




- Molecule 20: 30S ribosomal protein S20



- Molecule 21: 30S ribosomal protein Thx



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	411.50Å 411.50Å 172.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	184.03 – 3.30 184.03 – 3.30	Depositor EDS
% Data completeness (in resolution range)	97.7 (184.03-3.30) 97.7 (184.03-3.30)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.86 (at 3.33Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.292 , 0.323 0.252 , 0.285	Depositor DCC
$R_{free}$ test set	10942 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	95.8	Xtriage
Anisotropy	0.188	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.12 , 20.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.35$ , $\langle L^2 \rangle = 0.18$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.87	EDS
Total number of atoms	51308	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	109.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.66% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	29/36237 (0.1%)	0.99	122/56558 (0.2%)
2	B	0.68	0/1842	0.91	1/2479 (0.0%)
3	C	0.62	0/1636	0.88	3/2205 (0.1%)
4	D	0.72	2/1733 (0.1%)	1.01	7/2318 (0.3%)
5	E	0.81	0/1162	1.02	0/1564
6	F	0.57	0/856	0.83	0/1154
7	G	0.48	0/1248	0.71	0/1672
8	H	0.76	0/1136	1.02	3/1527 (0.2%)
9	I	0.56	0/1011	0.80	1/1354 (0.1%)
10	J	0.53	0/807	0.87	2/1085 (0.2%)
11	K	0.53	0/868	0.82	0/1173
12	L	0.59	0/986	0.89	1/1320 (0.1%)
13	M	0.54	0/965	0.88	3/1292 (0.2%)
14	N	0.58	0/501	0.98	1/664 (0.2%)
15	O	0.66	0/745	0.90	1/992 (0.1%)
16	P	0.66	0/716	0.88	0/963
17	Q	0.69	0/847	0.92	0/1131
18	R	0.56	0/564	0.89	0/748
19	S	0.53	0/661	0.92	1/890 (0.1%)
20	T	0.50	0/736	0.83	1/970 (0.1%)
21	V	0.60	0/212	0.77	0/277
All	All	0.87	31/55469 (0.1%)	0.96	147/82336 (0.2%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	160
4	D	0	1

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
5	E	0	1
8	H	0	1
17	Q	0	1
All	All	0	164

The worst 5 of 31 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1077	G	C5-C6	-12.45	1.29	1.42
1	A	1511	G	N3-C4	-8.17	1.29	1.35
1	A	1108	G	C5-C6	7.17	1.49	1.42
1	A	378	G	C5-C6	-6.89	1.35	1.42
1	A	1081	G	N3-C4	-6.69	1.30	1.35

The worst 5 of 147 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	246	A	N9-C1'-C2'	10.65	127.85	114.00
1	A	934	C	N1-C1'-C2'	9.98	126.98	114.00
14	N	40	CYS	CA-CB-SG	9.68	131.43	114.00
1	A	1336	C	N1-C1'-C2'	9.27	126.05	114.00
4	D	12	CYS	CA-CB-SG	9.23	130.62	114.00

There are no chirality outliers.

5 of 164 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	14	U	Sidechain
1	A	15	G	Sidechain
1	A	17	U	Sidechain
1	A	19	C	Sidechain
1	A	30	U	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	32372	0	16335	3399	0
2	B	1810	0	1861	119	0
3	C	1612	0	1677	169	0
4	D	1703	0	1763	134	0
5	E	1146	0	1207	112	0
6	F	843	0	857	42	0
7	G	1231	0	1273	81	0
8	H	1116	0	1177	113	0
9	I	993	0	1029	103	0
10	J	794	0	840	65	0
11	K	853	0	868	52	0
12	L	970	0	1057	77	0
13	M	955	0	1021	98	0
14	N	492	0	529	62	0
15	O	734	0	771	41	0
16	P	700	0	720	65	0
17	Q	834	0	906	70	0
18	R	559	0	624	49	0
19	S	647	0	673	57	0
20	T	734	0	832	55	0
21	V	208	0	221	15	0
22	D	1	0	0	0	0
22	N	1	0	0	0	0
All	All	51308	0	36241	4584	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 53.

The worst 5 of 4584 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''	1.52	1.36
1:A:1250:A:H2'	1:A:1251:A:C8	1.72	1.25
14:N:24:CYS:SG	14:N:39:LEU:HA	1.79	1.21
1:A:1027:C:H2'	1:A:1028:C:C5'	1.72	1.19
1:A:109:A:H2'	1:A:326:G:N2	1.58	1.18

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	220/255 (86%)	165 (75%)	41 (19%)	14 (6%)	1	9
3	C	204/238 (86%)	140 (69%)	47 (23%)	17 (8%)	1	5
4	D	206/208 (99%)	153 (74%)	42 (20%)	11 (5%)	2	12
5	E	148/161 (92%)	113 (76%)	24 (16%)	11 (7%)	1	7
6	F	99/101 (98%)	86 (87%)	9 (9%)	4 (4%)	3	18
7	G	151/155 (97%)	126 (83%)	21 (14%)	4 (3%)	5	27
8	H	136/138 (99%)	117 (86%)	13 (10%)	6 (4%)	2	16
9	I	123/128 (96%)	98 (80%)	21 (17%)	4 (3%)	4	22
10	J	96/104 (92%)	74 (77%)	14 (15%)	8 (8%)	1	5
11	K	113/128 (88%)	88 (78%)	17 (15%)	8 (7%)	1	7
12	L	122/131 (93%)	87 (71%)	26 (21%)	9 (7%)	1	7
13	M	118/125 (94%)	75 (64%)	31 (26%)	12 (10%)	0	3
14	N	58/60 (97%)	46 (79%)	12 (21%)	0	100	100
15	O	86/88 (98%)	62 (72%)	17 (20%)	7 (8%)	1	6
16	P	81/88 (92%)	58 (72%)	19 (24%)	4 (5%)	2	14
17	Q	98/104 (94%)	77 (79%)	16 (16%)	5 (5%)	2	13
18	R	66/87 (76%)	45 (68%)	17 (26%)	4 (6%)	1	10
19	S	78/92 (85%)	62 (80%)	15 (19%)	1 (1%)	12	40
20	T	92/105 (88%)	70 (76%)	16 (17%)	6 (6%)	1	9
21	V	22/26 (85%)	17 (77%)	3 (14%)	2 (9%)	1	4
All	All	2317/2522 (92%)	1759 (76%)	421 (18%)	137 (6%)	1	10

5 of 137 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	10	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	12	GLU
2	B	21	ARG
2	B	24	TRP
2	B	95	GLN

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	191/219 (87%)	166 (87%)	25 (13%)	4	17
3	C	160/187 (86%)	141 (88%)	19 (12%)	5	21
4	D	180/180 (100%)	166 (92%)	14 (8%)	12	38
5	E	115/122 (94%)	103 (90%)	12 (10%)	7	25
6	F	90/90 (100%)	85 (94%)	5 (6%)	21	52
7	G	124/126 (98%)	120 (97%)	4 (3%)	39	67
8	H	119/119 (100%)	107 (90%)	12 (10%)	7	27
9	I	96/99 (97%)	90 (94%)	6 (6%)	18	47
10	J	88/91 (97%)	78 (89%)	10 (11%)	5	22
11	K	87/98 (89%)	82 (94%)	5 (6%)	20	51
12	L	104/108 (96%)	100 (96%)	4 (4%)	33	62
13	M	96/100 (96%)	83 (86%)	13 (14%)	4	16
14	N	49/49 (100%)	41 (84%)	8 (16%)	2	10
15	O	79/79 (100%)	69 (87%)	10 (13%)	4	19
16	P	72/74 (97%)	65 (90%)	7 (10%)	8	29
17	Q	95/96 (99%)	87 (92%)	8 (8%)	11	35
18	R	60/76 (79%)	57 (95%)	3 (5%)	24	55
19	S	71/79 (90%)	65 (92%)	6 (8%)	10	35
20	T	74/81 (91%)	70 (95%)	4 (5%)	22	53
21	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	1969/2094 (94%)	1794 (91%)	175 (9%)	9 32

5 of 175 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
8	H	17	THR
10	J	23	ILE
17	Q	100	LYS
8	H	26	VAL
8	H	112	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
7	G	106	GLN
10	J	56	HIS
19	S	14	HIS
7	G	122	HIS
8	H	82	HIS

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1506/1509 (99%)	332 (22%)	181 (12%)

5 of 332 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 181 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	703	G
1	A	883	C
1	A	1396	A

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	733	A
1	A	815	A

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 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 ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	1506/1509 (99%)	-0.33	9 (0%) 89 90	29, 94, 180, 218	0
2	B	222/255 (87%)	0.09	4 (1%) 68 67	43, 111, 195, 218	0
3	C	206/238 (86%)	0.24	15 (7%) 15 15	49, 123, 192, 215	0
4	D	208/208 (100%)	0.26	8 (3%) 40 37	25, 101, 171, 218	0
5	E	150/161 (93%)	0.30	10 (6%) 17 17	42, 85, 151, 185	0
6	F	101/101 (100%)	-0.05	0 100 100	53, 122, 186, 211	0
7	G	153/155 (98%)	0.12	8 (5%) 27 25	74, 136, 196, 218	0
8	H	138/138 (100%)	0.02	2 (1%) 75 75	30, 83, 156, 193	0
9	I	125/128 (97%)	1.36	39 (31%) 0 0	69, 144, 201, 218	0
10	J	98/104 (94%)	1.56	39 (39%) 0 0	61, 150, 210, 218	0
11	K	115/128 (89%)	0.19	9 (7%) 13 12	60, 118, 181, 209	0
12	L	124/131 (94%)	0.57	15 (12%) 4 3	41, 108, 169, 204	0
13	M	120/125 (96%)	0.69	22 (18%) 1 1	66, 133, 198, 218	0
14	N	60/60 (100%)	1.18	13 (21%) 0 1	56, 110, 174, 203	0
15	O	88/88 (100%)	0.16	2 (2%) 60 59	50, 103, 167, 203	0
16	P	83/88 (94%)	0.44	8 (9%) 8 8	36, 91, 152, 216	0
17	Q	100/104 (96%)	0.11	0 100 100	43, 95, 166, 211	0
18	R	68/87 (78%)	-0.03	0 100 100	49, 102, 184, 196	0
19	S	80/92 (86%)	0.78	17 (21%) 0 1	66, 143, 208, 218	0
20	T	94/105 (89%)	0.80	16 (17%) 1 1	64, 126, 190, 218	0
21	V	24/26 (92%)	3.16	16 (66%) 0 0	93, 129, 158, 188	0
All	All	3863/4031 (95%)	0.14	252 (6%) 18 18	25, 104, 188, 218	0

The worst 5 of 252 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
10	J	33	GLN	9.9
9	I	128	ARG	9.7
19	S	2	PRO	9.3
13	M	121	LYS	8.0
21	V	6	ARG	7.6

## 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
22	ZN	D	210	1/1	0.99	0.33	90,90,90,90	0
22	ZN	N	62	1/1	0.99	0.11	101,101,101,101	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.