



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 4, 2024 – 08:05 PM EST

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.; Schlunzen, 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](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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.36  
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) : 2.36

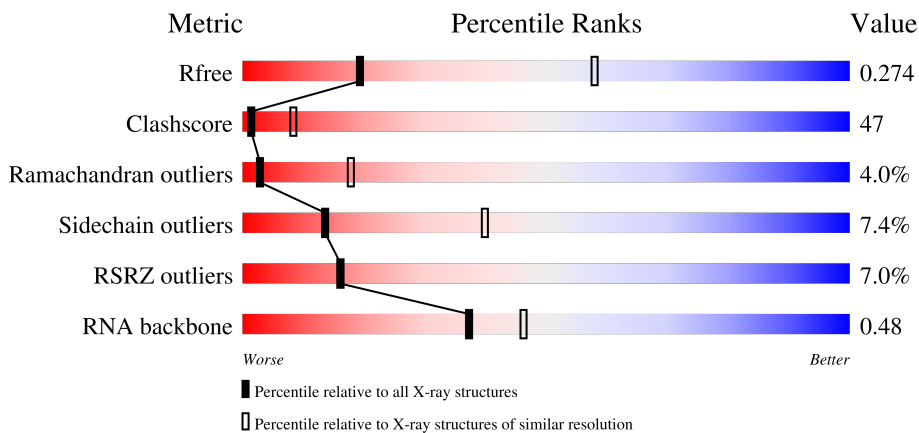
# 1 Overall quality at a glance

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 of 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	1520	 10% 60% 21% 8%
2	1	6	 17% 50% 33%
2	2	6	 17% 33% 33% 33%
3	B	227	 52% 37% 7%

*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 [i](#)

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
			Total	C	N	O	P			
1	A	1517	32594	14508	6027	10542	1517	0	0	0

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

- 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			

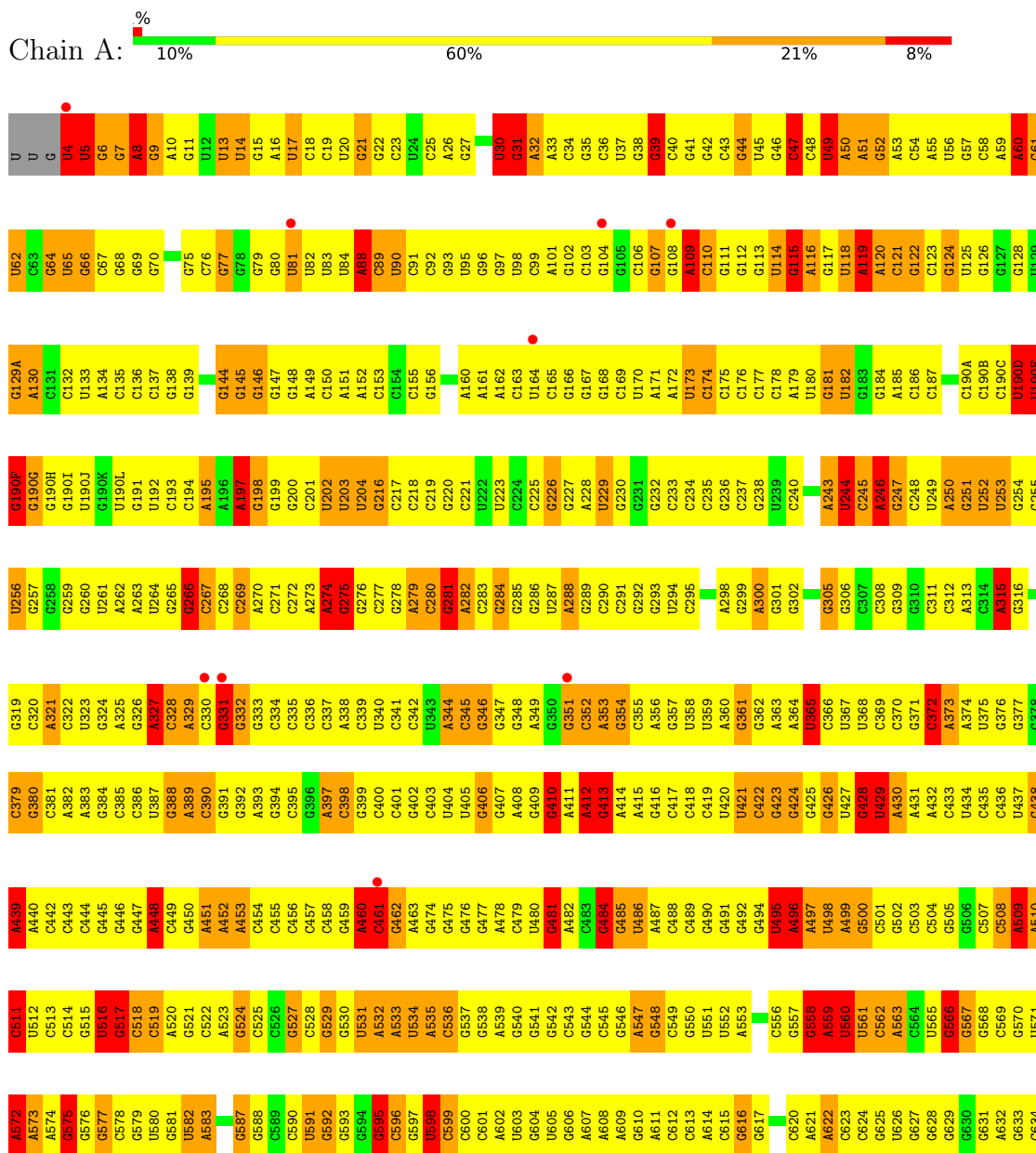
- Molecule 23 is ZINC ION (three-letter code: ZN) (formula: Zn).

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 [i](#)

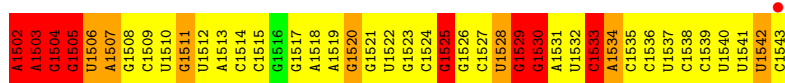
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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

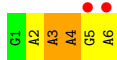
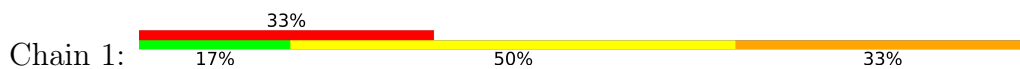




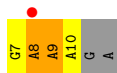
U1436	U1437	G1438	G1439	G1440	G1441	G1442	G1443	G1444	G1445	G1446	G1447	G1448	G1449	G1450	G1451	G1452	G1453	G1454	G1455	G1456	G1457	G1458	G1459	G1460	G1461	G1462	G1463	G1464	G1465	G1466	G1467	G1468	G1469	G1470	G1471	G1472	G1473	G1474	G1475	G1476	G1477	G1480	G1481	G1482	G1483	G1484	G1485	G1486	G1487	G1488	G1489	G1490	G1491	G1492	G1493	G1494	G1495	G1496	G1497	G1498	G1499	G1500	G1501																																																																																																																																																																																	
A1374	A1375	A1376	A1377	A1378	A1379	A1380	A1381	A1382	A1383	A1384	A1385	A1386	A1387	A1388	A1389	A1390	A1391	A1392	A1393	A1394	A1395	A1396	A1397	A1398	A1399	A1400	A1401	A1402	A1403	A1404	A1405	A1406	A1407	A1408	A1411	A1412	A1413	A1414	A1415	A1416	A1417	A1418	A1419	A1420	A1421	A1422	A1423	A1424	A1425	A1426	A1427	A1428	A1429	A1430	A1431	A1432	A1433	A1434	A1435																																																																																																																																																																																					
U1315	G1316	C1317	A1318	A1319	A1320	C1321	G1322	G1323	C1324	C1325	C1326	C1327	C1328	A1329	A1330	A1331	A1332	A1333	A1334	C1335	C1336	C1337	A1338	A1339	A1340	A1341	C1342	C1343	C1344	A1345	A1346	A1347	A1348	A1349	A1350	A1351	C1352	C1353	C1354	C1355	C1356	C1357	A1358	C1359	A1360	C1361	C1361A	C1362	A1363	A1364	C1365	C1366	C1367	C1368	C1369	C1370	C1371	C1372	C1373																																																																																																																																																																																					
G1193	U1194	C1195	U1196	U1197	G1198	U1199	C1200	A1201	C1202	C1203	A1204	U1205	G1206	G1207	G1210	U1211	U1212	U1213	C1214	G1215	G1216	C1217	C1218	U1219	G1220	C1221	G1222	C1223	A1224	A1225	C1226	A1227	C1228	A1229	C1230	G1231	G1232	G1233	C1234	U1235	U1236	C1237	A1238	A1239	U1240	G1241	C1242	C1243	C1244	A1245	G1246	U1247	U1248	C1249	A1250	A1251	U1252	C1253																																																																																																																																																																																						
G1130	G1131	G1134	U1135	U1136	C1137	G1138	A1139	C1140	G1141	G1142	G1143	C1144	C1145	A1146	G1147	U1148	C1149	U1150	A1151	A1152	C1153	G1154	U1155	G1156	A1157	C1158	U1159	G1160	C1161	C1162	A1163	G1164	C1165	A1166	A1167	A1168	A1169	C1170	C1171	G1172	G1173	G1174	G1175	A1176	G1177	A1178	A1179	A1180	G1181	A1182	A1183	G1184	C1185	U1186	U1187	A1188	C1189	A1190	A1191	C1192																																																																																																																																																																																				
G1068	C1069	G1072	U1073	G1074	C1075	G1076	U1077	U1078	A1080	G1081	G1082	U1083	G1084	U1085	U1086	G1087	G1088	G1089	U1090	G1091	A1092	A1093	G1094	U1095	C1096	C1097	G1098	G1099	C1100	A1101	C1102	C1103	G1104	A1105	G1106	C1107	G1108	A1109	A1110	A1111	G1112	C1113	C1114	G1117	A1118	C1119	G1120	U1121	U1122	A1123	G1124	U1125	U1126	G1127	C1128	C1129																																																																																																																																																																																								
G1009	G1010	G1011	U1012	G1013	A1014	A1015	A1016	G1017	C1018	C1019	U1020	G1021	G1022	G1023	G1024	U1025	A1026	C1027	C1028	C1029	U1030	G1030A	C1031	G1032	G1033	G1034	G1037	C1038	U1039	U1040	A1044	C1045	A1046	G1047	G1048	U1049	C1050	C1051	U1052	C1053	A1054	A1055	C1056	A1057	G1058	C1059	C1060	G1061	G1003A	G1004	A1005	C1006	C1007	C1008																																																																																																																																																																																										
C948	A949	U950	G951	U952	G953	G954	U955	U956	U957	A958	A959	U960	U961	C962	C963	A964	U965	C966	C967	A968	A969	C970	U971	C972	G973	A974	U975	G976	U977	C978	C979	U980	U981	A982	A983	C984	C985	A986	A987	C988	C989	C990	U991	U992	U993	U994	A995	A996	U997	G998	G1002	G1003	G1004	A1005	C1006	C1007	C1008																																																																																																																																																																																							
A888	A889	G890	U891	A892	G893	G894	U895	U896	C897	A898	C899	A900	A901	G902	G903	C904	U905	G906	A907	A908	A909	C910	U911	G912	A913	A914	A915	G916	G917	A918	A919	U920	U921	C922	A923	A924	A925	G926	G927	C928	G929	C930	U931	C932	G933	A934	A935	C936	A937	A938	G939	C940	G941	G942	G943	G944	A945	A946	C947	G948	G949																																																																																																																																																																																			
U820	C821	C822	G823	G824	C825	C826	U827	A828	G829	C830	A831	C832	U833	G836	G837	C838	U839	C840	U841	C842	C843	C844	C845	U846	C847	A848	C849	C850	C851	C852	C853	C854	C855	C856	C857	C858	C859	C860	C861	C862	C863	C864	C865	C866	C867	C868	C869	C870	C871	C872	C873	C874	C875	C876	C877	C878	C879	C880	C881	C882	C883	C884	C885	C886	C887																																																																																																																																																																															
A695	A696	U697	G698	A699	G700	C701	A702	G703	A704	U705	A706	C707	C708	G709	G710	G711	A712	G713	C714	A715	A716	C717	G718	U719	C720	G721	A722	C723	A724	G725	C726	U727	A728	G731	C732	A733	G734	C735	C736	A737	C738	U739	U801	G802	G803	U804	C805	C806	U807	C808	C811	C812	U813	A814	C815	U816	C817	A818	C819	C820	C821	C822	C823	C824	C825	C826	C827	C828	C829	C830	C831	C832	C833	C834	C835	C836	C837	C838	C839	C840	C841	C842	C843	C844	C845	C846	C847	C848	C849	C850	C851	C852	C853	C854	C855	C856	C857	C858	C859	C860	C861	C862	C863	C864	C865	C866	C867	C868	C869	C870	C871	C872	C873	C874	C875	C876	C877	C878	C879	C880	C881	C882	C883	C884	C885	C886	C887	C888	C889	C890	C891	C892	C893	C894	C895	C896	C897	C898	C899	C900	C901	C902	C903	C904	C905	C906	C907	C908	C909	C910	C911	C912	C913	C914	C915	C916	C917	C918	C919	C920	C921	C922	C923	C924	C925	C926	C927	C928	C929	C930	C931	C932	C933	C934	C935	C936	C937	C938	C939	C940	C941	C942	C943	C944	C945	C946	C947	C948	C949	C950	C951	C952	C953	C954	C955	C956	C957	C958	C959	C960	C961	C962	C963	C964	C965	C966	C967	C968	C969	C970	C971	C972	C973	C974	C975	C976	C977	C978	C979	C980	C981	C982	C983	C984	C985	C986	C987	C988	C989	C990	C991	C992	C993	C994	C995	C996	C997	C998	C999	C1000
U636	G637	G638	G639	A640	U641	A642	C643	G644	C645	A646	C647	A648	G649	C650	C651	U652	A653	A654	A655	A656	A657	A658	U659	G660	G661	G662	A663	G664	A665	A666	G667	G668	U669	G670	G671	U672	G673	G674	A675	A676	U677	U678	C679	C680	C681	G682	G683	A684	G685	U686	G687	G688	G689	G690	G691	U692	A693	G694	A695	A696	A697	A698	A699	A700	A701	A702	A703	A704	A705	A706	A707	A708	A709	A710	A711	A712	A713	A714	A715	A716	A717	A718	A719	A720	A721	A722	A723	A724	A725	A726	A727	A728	A729	A730	A731	A732	A733	A734	A735	A736	A737	A738	A739	A740	A741	A742	A743	A744	A745	A746	A747	A748	A749	A750	A751	A752	A753	A754	A755																																																																																																																									



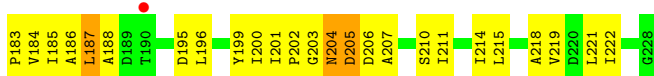
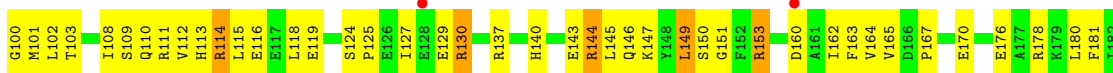
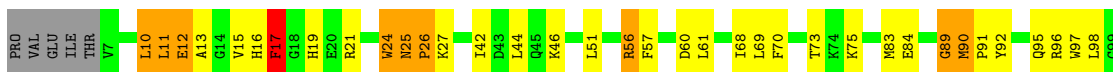
- Molecule 2: 5'-R(\*GP\*AP\*AP\*AP\*GP\*A)-3'



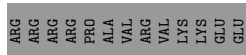
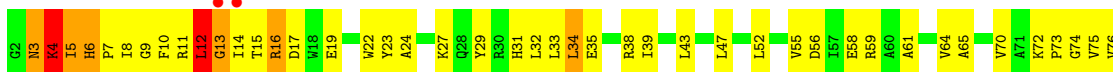
- Molecule 2: 5'-R(\*GP\*AP\*AP\*AP\*GP\*A)-3'



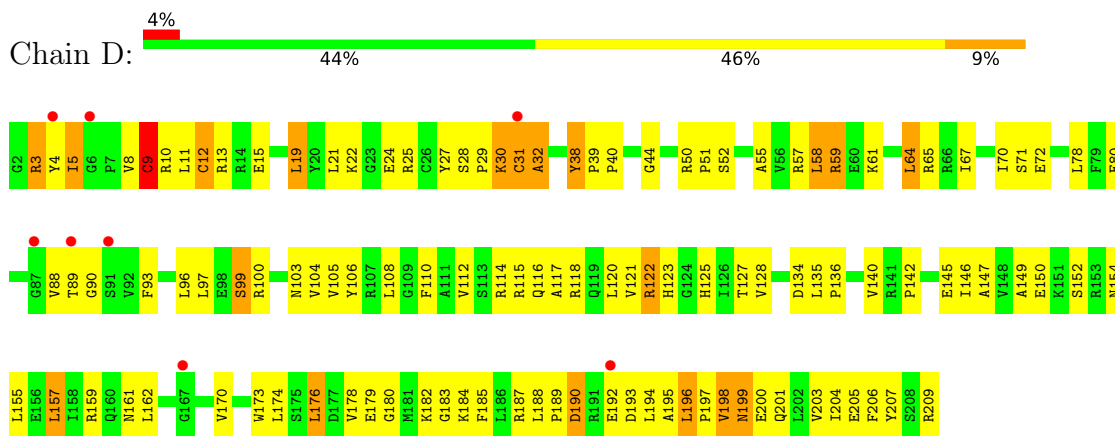
- Molecule 3: 30S ribosomal protein S2



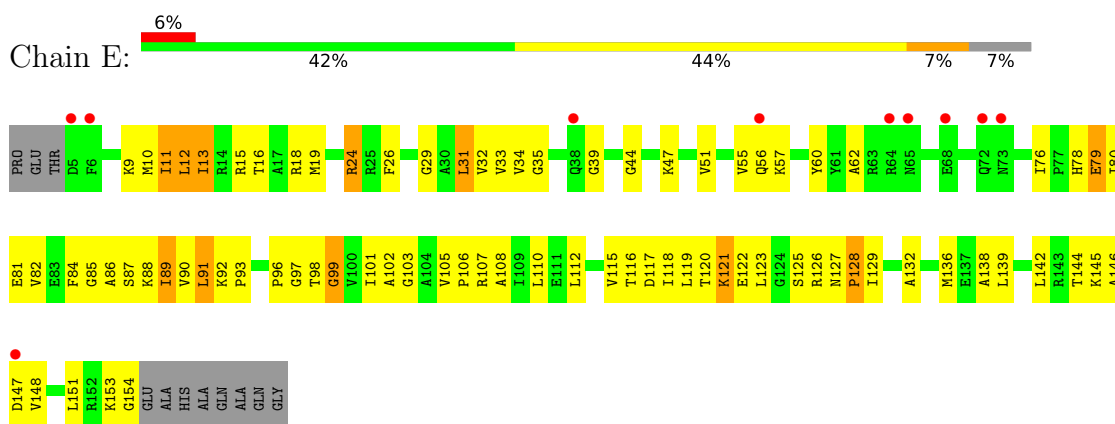
- Molecule 4: 30S ribosomal protein S3



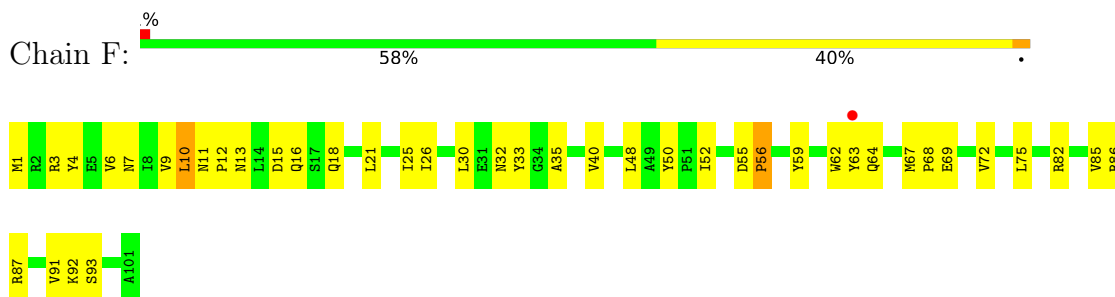
- Molecule 5: 30S ribosomal protein S4



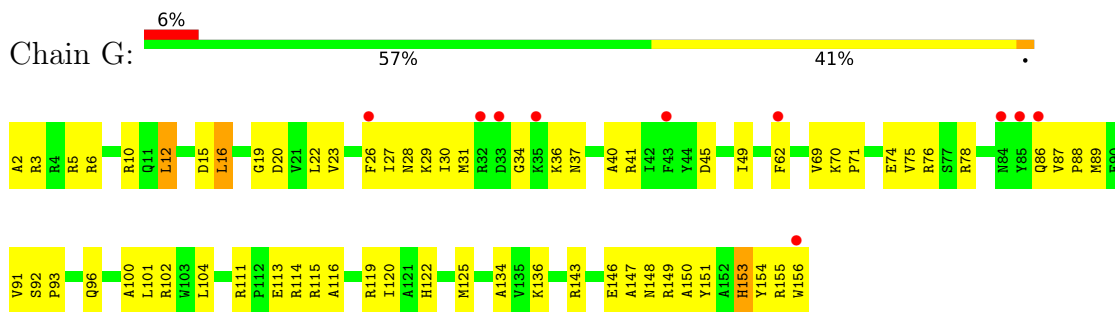
- Molecule 6: 30S ribosomal protein S5



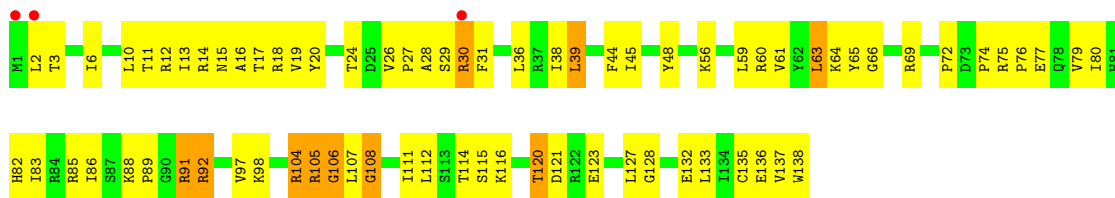
- Molecule 7: 30S ribosomal protein S6



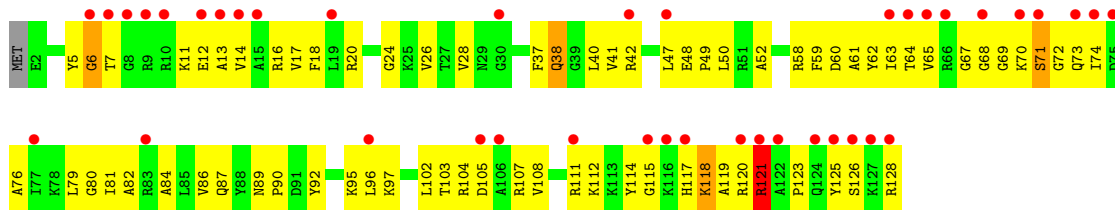
- Molecule 8: 30S ribosomal protein S7



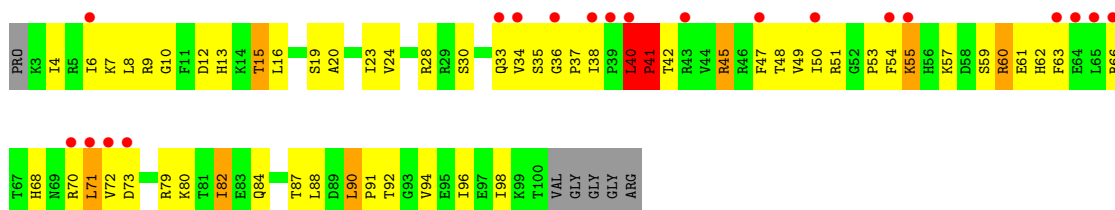
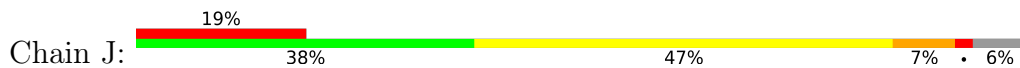
- Molecule 9: 30S ribosomal protein S8



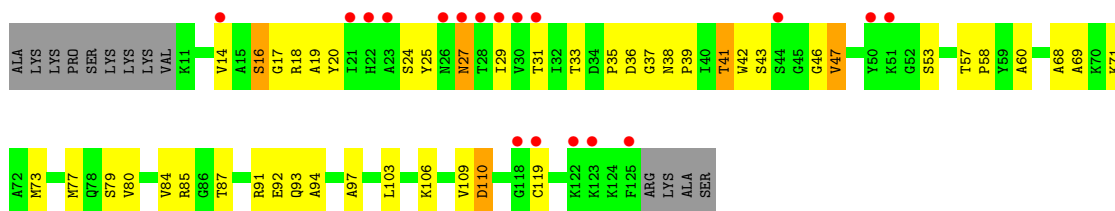
- Molecule 10: 30S ribosomal protein S9



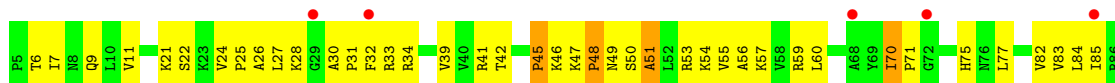
- Molecule 11: 30S ribosomal protein S10



- Molecule 12: 30S ribosomal protein S11

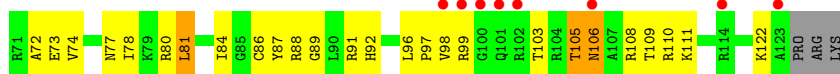


- Molecule 13: 30S ribosomal protein S12

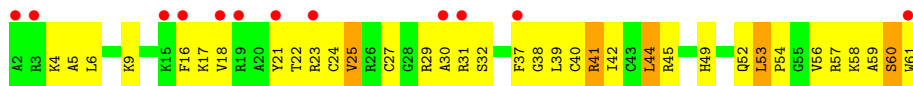




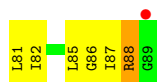
- Molecule 14: 30S ribosomal protein S13



- Molecule 15: 30S ribosomal protein S14



- Molecule 16: 30S ribosomal protein S15

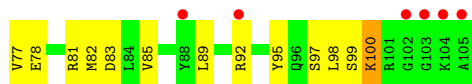


- Molecule 17: 30S ribosomal protein S16



- Molecule 18: 30S ribosomal protein S17

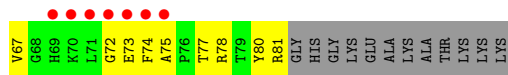
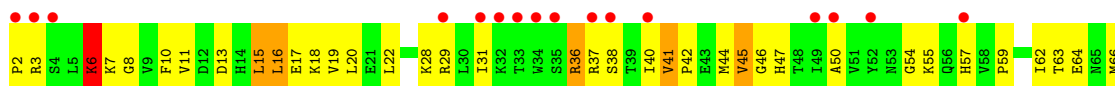




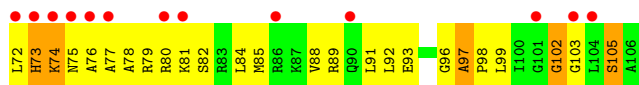
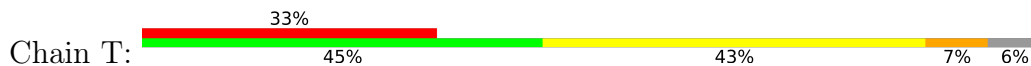
- Molecule 19: 30S ribosomal protein S18



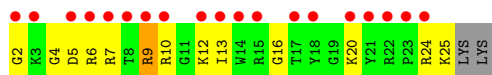
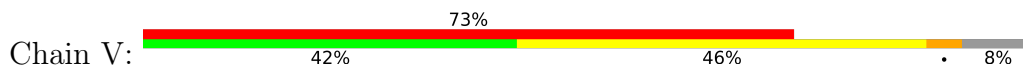
- Molecule 20: 30S ribosomal protein S19



- Molecule 21: 30S ribosomal protein S20



- Molecule 22: 30S ribosomal protein Thx



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	411.79Å 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
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.23 (at 3.33Å)	Xtrriage
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 <sup>2</sup> )	96.4	Xtrriage
Anisotropy	0.310	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.18 , 82.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.38$ , $\langle L^2 \rangle = 0.21$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.82	EDS
Total number of atoms	51895	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	104.0	wwPDB-VP

Xtrriage'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.*

<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 [i](#)

### 5.1 Standard geometry [i](#)

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 [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	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:243:A:H4'	1:A:244:U:C5'	1.65	1.24
1:A:839:U:H5'	1:A:840:C:C5	1.71	1.24

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all 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 [i](#)

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
15	N	60	SER
20	S	57	HIS
16	O	52	SER
18	Q	11	VAL
5	D	190	ASP

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

Mol	Chain	Res	Type
12	K	117	ASN
16	O	46	HIS
13	L	49	ASN
14	M	106	ASN
18	Q	16	GLN

### 5.3.3 RNA [i](#)

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	975	A
1	A	1214	C
1	A	993	G
1	A	1129	C
1	A	1257	U

#### 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 monosaccharides 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	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, 120, 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 monosaccharides 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(Å <sup>2</sup> )	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.