wwPDB X-ray Structure Validation Summary Report (i)

Feb 3, 2024 – 10:21 PM EST

PDB ID	:	1N36
Title	:	Structure of the Thermus thermophilus 30S ribosomal subunit in the presence of crystallographically disordered codon and near-cognate transfer RNA anticodon stem-loop mismatched at the second codon position
Authors	:	Ogle, J.M.; Murphy IV, F.V.; Tarry, M.J.; Ramakrishnan, V.
Deposited on	:	2002-10-25
Resolution	:	3.65 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 (i) 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 (1)) were used in the production of this report:

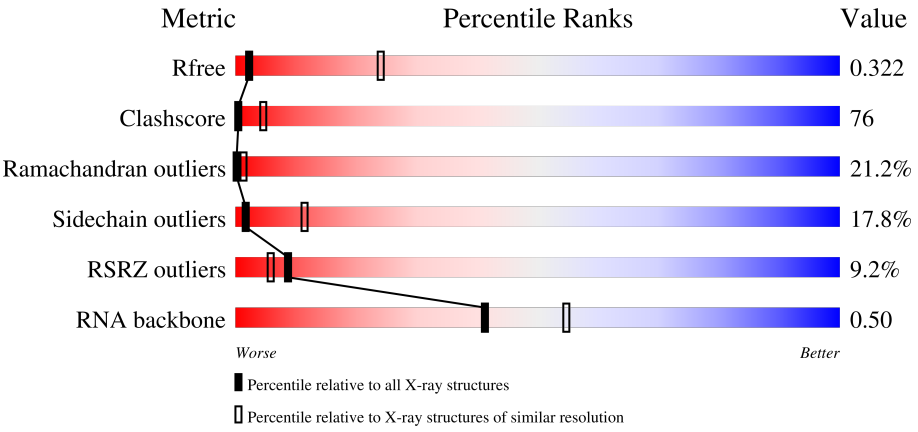
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.65 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.



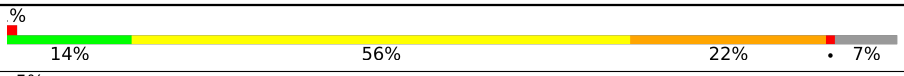
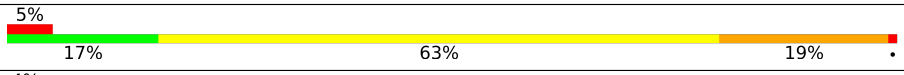
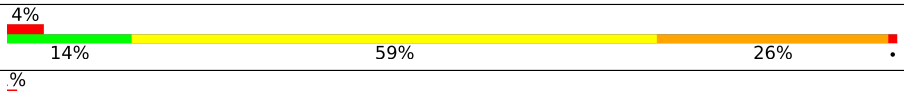
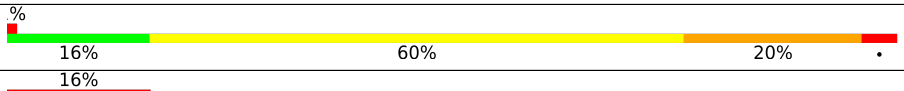
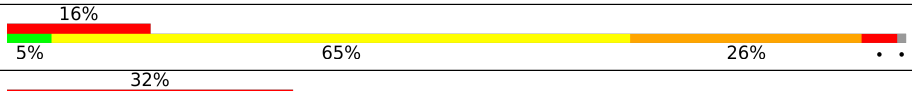
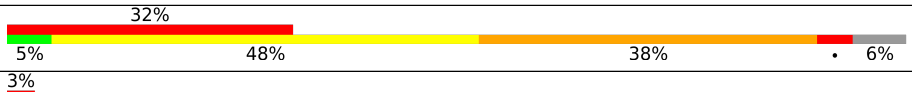
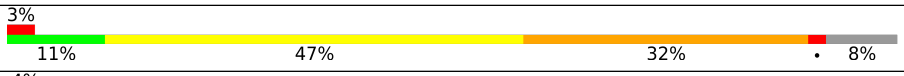
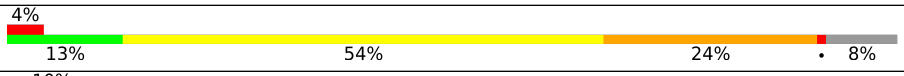
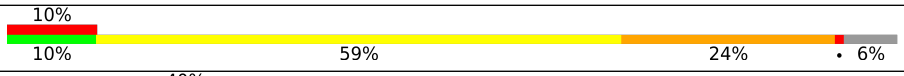
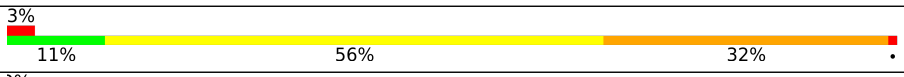
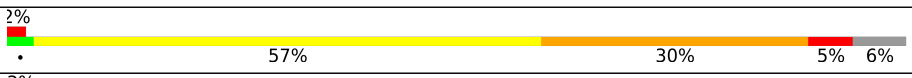
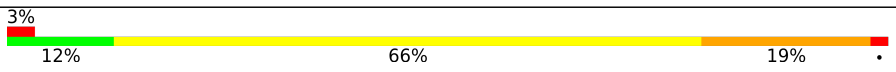
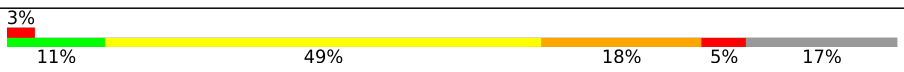
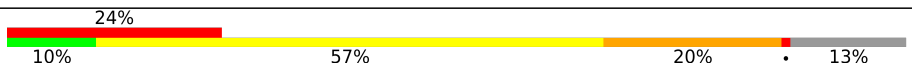
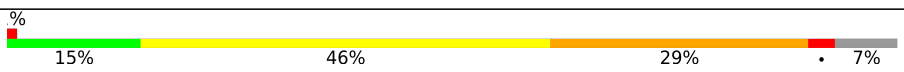
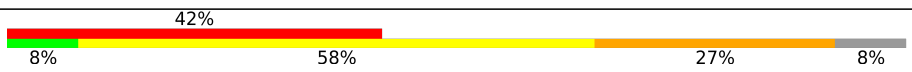
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1557 (3.82-3.50)
Clashscore	141614	1037 (3.80-3.52)
Ramachandran outliers	138981	1004 (3.80-3.52)
Sidechain outliers	138945	1002 (3.80-3.52)
RSRZ outliers	127900	1441 (3.82-3.50)
RNA backbone	3102	1024 (4.30-3.00)

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 >=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 <=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	1522	11% 9% 70% 18% ••
2	B	256	% 8% 55% 23% 5% 9%
3	C	239	8% 6% 49% 29% • 14%
4	D	208	3% 12% 57% 25% 6%

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	129	
12	L	135	
13	M	126	
14	N	60	
15	O	88	
16	P	88	
17	Q	104	
18	R	88	
19	S	92	
20	T	106	
21	V	26	

2 Entry composition (i)

There are 22 unique types of molecules in this entry. The entry contains 51680 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	1513	32508	14472	6016	10509	1511	22	0	0

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

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

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

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

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

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

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

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

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

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

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

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

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	25	ASP	GLU	conflict	UNP Q5SHQ2
H	37	ARG	LYS	conflict	UNP Q5SHQ2
H	52	ASP	GLU	conflict	UNP Q5SHQ2
H	61	VAL	ILE	conflict	UNP Q5SHQ2
H	62	TYR	HIS	conflict	UNP Q5SHQ2
H	81	HIS	LYS	conflict	UNP Q5SHQ2
H	88	LYS	ARG	conflict	UNP Q5SHQ2
H	115	SER	PRO	conflict	UNP Q5SHQ2

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

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

- 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
			792	498	156	137	1			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	119	Total	C	N	O	S	0	0	0
			885	549	168	165	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	118	Total	C	N	O	S	0	0	0
			937	579	193	163	2			

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

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	104	Total	C	N	O	S	0	0	0
			857	547	161	147	2			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Q	50	LYS	ARG	conflict	UNP Q5SHP7
Q	53	LEU	VAL	conflict	UNP Q5SHP7
Q	62	SER	ALA	conflict	UNP Q5SHP7
Q	79	SER	GLU	conflict	UNP Q5SHP7
Q	82	MET	LEU	conflict	UNP Q5SHP7
Q	90	ILE	VAL	conflict	UNP Q5SHP7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
Q	96	GLN	ALA	conflict	UNP Q5SHP7

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

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

- 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	99	Total	C	N	O	S	0	0	0
			763	470	162	129	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).

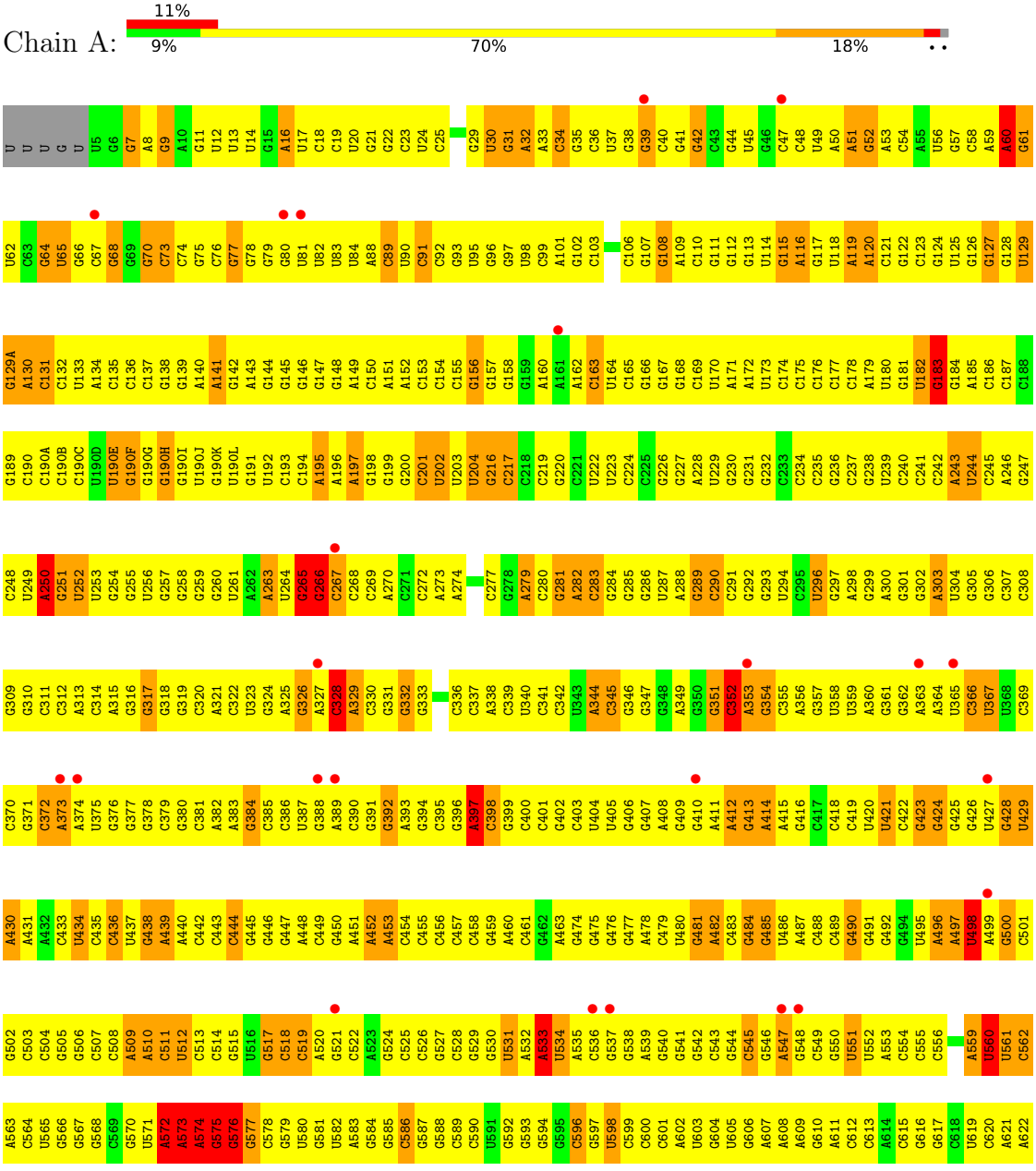
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
22	D	1	Total	Zn	0	0
			1	1		
22	N	1	Total	Zn	0	0
			1	1		

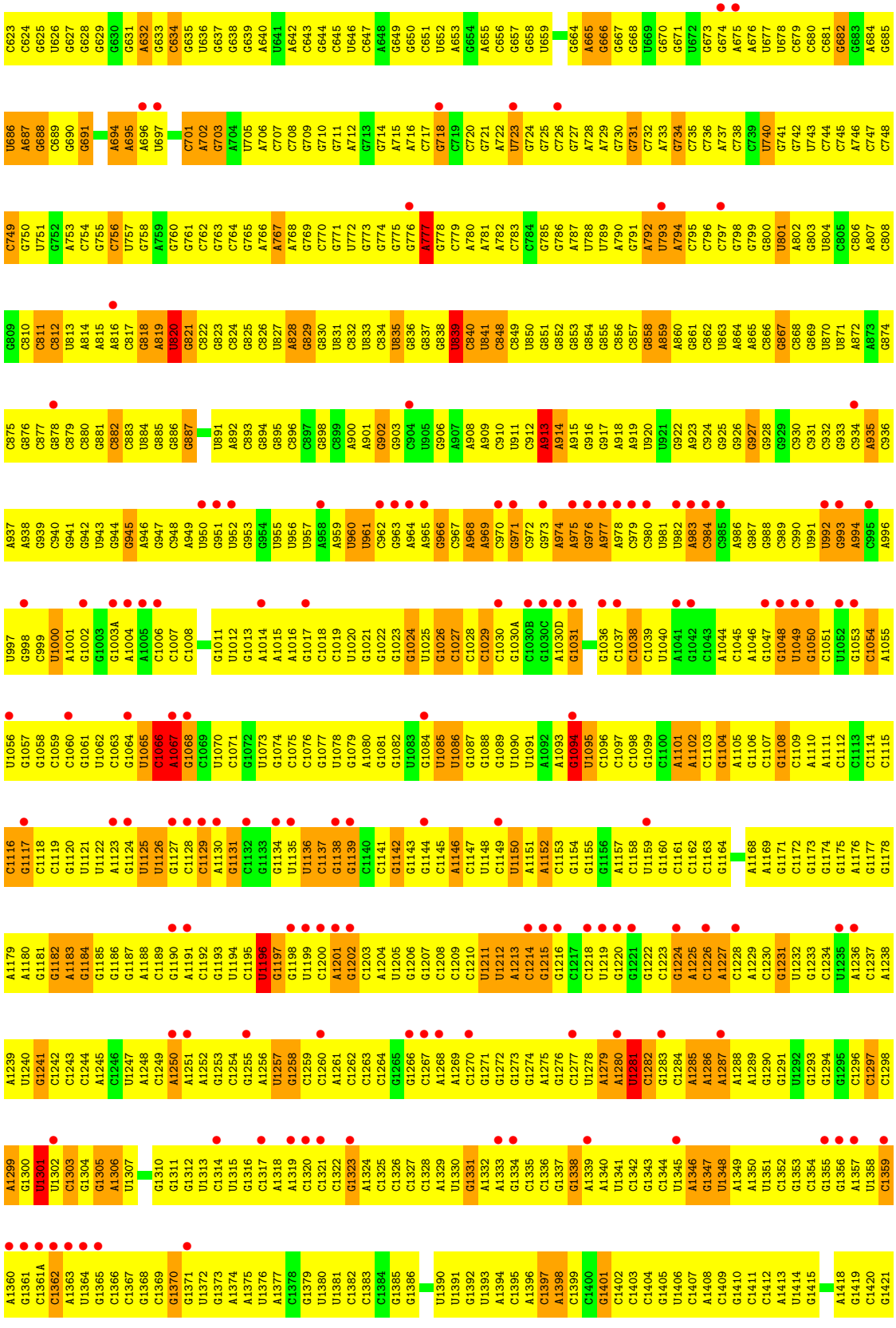
3 Residue-property plots (i)

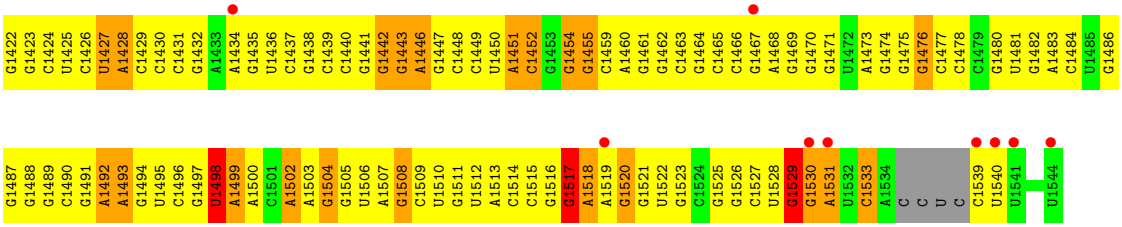
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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

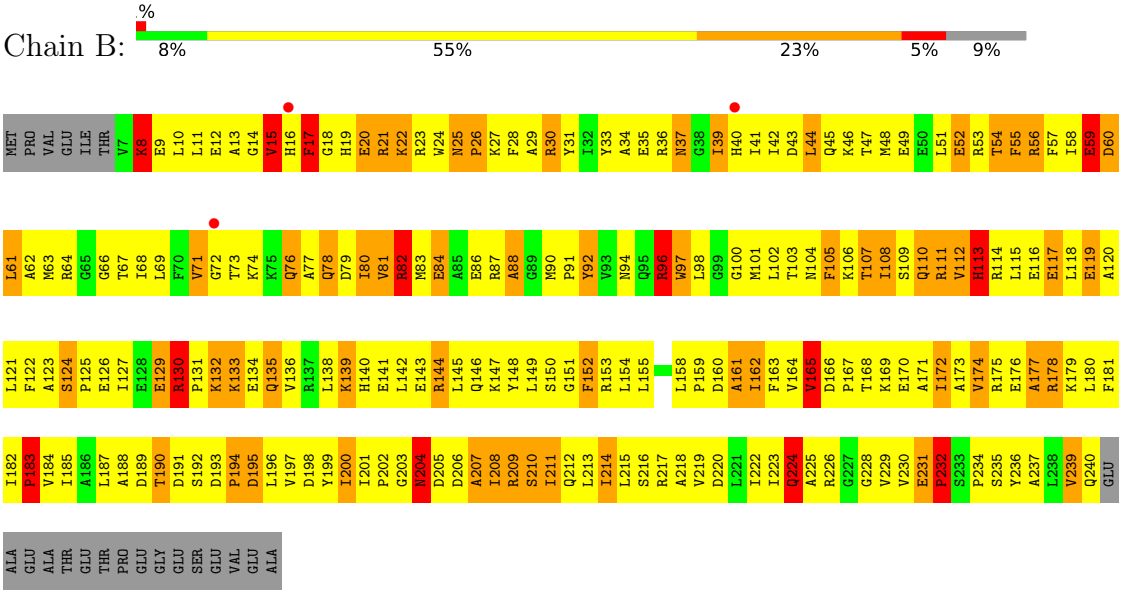
Chain A:



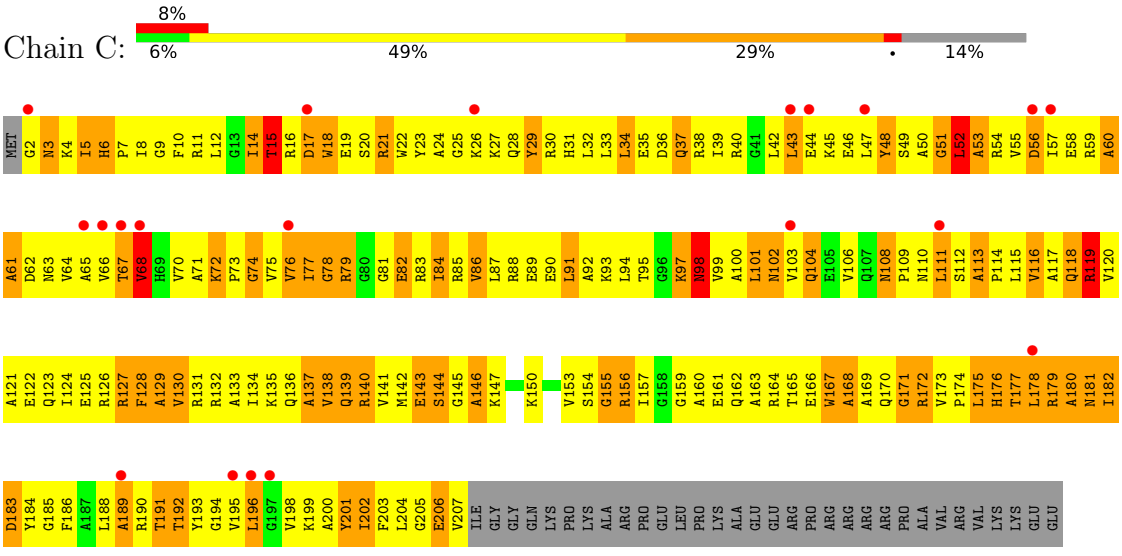

C623 C624 G625 U626 G627 G628 G629 G630 G631 A632 G633 C634 G635 U636 G637 G638 G639 A640 U641 A642 C643 G644 C645 U646 C647 A648 G649 G650 C651 U652 A653 G654 A655 C656 G657 G658 U659 G664 A665 G666 G667 G668 U669 G670 G671 U672 G673 G674 A675 A676 U677 U678 C679 C680 C681 G682 G683 A684 G685
U686 A687 G688 C689 G690 G691 A694 A695 A696 U697 C701 A702 G703 A704 U705 A706 C707 C708 G709 G710 G711 A712 G713 G714 A715 A716 C717 G718 C719 C720 G721 A722 U723 G724 G725 C726 G727 A728 A729 G730 G731 C732 A733 G734 C735 C736 A737 C738 C739 U740 G741 G742 U743 C744 C745 A746 C747 C748
C749 G750 U751 G752 A753 C754 G755 C756 U757 G758 A759 G760 G761 C762 G763 C764 G765 A766 A767 A768 G769 C770 G771 U772 G773 G774 G775 G776 A777 G778 C779 A780 A781 A782 C783 G784 G785 G786 A787 U788 U789 A790 G791 A792 U793 A794 C795 C796 C797 G798 G799 G800 U801 A802 G803 U804 C805 C806 A807 C808
G809 C810 C811 C812 U813 A814 A815 A816 C817 G818 A819 U820 G821 C822 G823 C824 G825 C826 U827 A828 G829 G830 U831 C832 U833 C834 U835 G836 G837 C838 U839 C840 U841 C848 C849 U850 G851 G852 G853 G854 G855 C856 C857 G858 A859 A860 C861 C862 U863 A864 A865 C866 G867 C868 G869 U870 U871 A872 A873 G874
C875 G876 C877 G878 C879 C880 G881 C882 C883 U884 G885 G886 G887 U891 A892 C893 C894 G895 C896 C897 G898 C899 A900 A901 G902 G903 C904 U905 G906 A907 A908 A909 C910 U911 C912 A913 A914 A915 G916 G917 A918 A919 U920 U921 G922 A923 C924 G925 G926 G927 G928 G929 C930 C931 C932 G933 C934 A935 C936
A937 A938 G939 C940 G941 G942 U943 G944 G945 A946 G947 C948 A949 U950 G951 U952 G953 C954 U955 U956 U957 A958 A959 U960 U961 C962 G963 A964 A965 G966 C967 A968 A969 C970 G971 C972 G973 A974 A975 G976 A977 A978 C979 C980 U981 U982 A983 C984 C985 A986 G987 G988 C989 C990 U991 U992 G993 A994 C995 A996
U997 G998 C999 U1000 A1001 G1002 G1003 G1003A A1004 A1005 C1006 C1007 C1008 G1011 U1012 G1013 A1014 A1015 A1016 G1017 C1018 C1019 U1020 G1021 G1022 G1023 G1024 U1025 G1026 C1027 C1028 C1029 C1030 G1030A C1030B G1030C A1030D G1031 G1036 C1037 C1038 C1039 C1040 A1041 G1042 C1043 A1044 C1045 A1046 C1047 G1048 U1049 G1050 C1051 U1052 G1053 C1054 A1055
U1056 G1057 G1058 C1059 C1060 G1061 U1062 C1063 G1064 U1065 C1066 A1067 G1068 C1069 U1070 C1071 G1072 U1073 G1074 C1075 C1076 G1077 U1078 G1079 A1080 G1081 G1082 U1083 G1084 U1085 U1086 G1087 G1088 G1089 U1090 U1091 A1092 A1093 G1094 U1095 C1096 C1097 C1098 G1099 C1100 A1101 A1102 C1103 G1104 A1105 G1106 C1107 G1108 C1109 A1110 A1111 C1112 C1113 C1114 C1115
C1116 G1117 C1118 C1119 G1120 U1121 U1122 A1123 G1124 U1125 U1126 G1127 C1128 C1129 A1130 G1131 C1132 C1133 G1134 U1135 U1136 C1137 G1138 G1139 C1140 C1141 G1142 G1143 G1144 C1145 A1146 C1147 U1148 C1149 U1150 A1151 A1152 C1153 G1154 G1155 G1156 A1157 C1158 U1159 G1160 C1161 C1162 C1163 G1164 A1168 A1169 G1171 C1172 G1173 G1174 G1175 A1176 G1177 G1178
A1179 A1180 G1181 G1182 A1183 G1184 G1185 G1186 G1187 A1188 C1189 G1190 A1191 C1192 G1193 U1194 C1195 U1196 G1197 G1198 U1199 C1200 A1201 G1202 C1203 A1204 U1205 G1206 G1207 C1208 C1209 C1210 U1211 U1212 A1213 C1214 G1215 G1216 C1217 C1218 U1219 G1220 G1221 G1222 C1223 G1224 A1225 C1226 A1227 C1228 A1229 C1230 G1231 U1232 G1233 C1234 U1235 A1236 C1237 A1238
A1239 U1240 G1241 C1242 C1243 C1244 A1245 C1246 U1247 A1248 C1249 A1250 A1251 A1252 G1253 C1254 G1255 A1256 U1257 G1258 C1259 C1260 A1261 C1262 C1263 C1264 C1265 G1266 C1267 A1268 A1269 C1270 G1271 G1272 G1273 G1274 A1275 G1276 C1277 U1278 A1279 A1280 U1281 C1282 G1283 C1284 A1285 A1286 A1287 A1288 A1289 G1290 G1291 U1292 G1293 G1294 G1295 C1296 C1297 C1298
A1299 G1300 U1301 U1302 C1303 G1304 G1305 A1306 U1307 G1310 G1311 G1312 U1313 C1314 U1315 G1316 C1317 A1318 A1319 C1320 C1321 C1322 G1323 A1324 C1325 C1326 C1327 C1328 A1329 U1330 G1331 A1332 A1333 G1334 C1335 C1336 U1337 G1338 A1339 A1340 U1341 C1342 G1343 C1344 U1345 A1346 G1347 U1348 A1349 A1350 U1351 C1352 G1353 C1354 G1355 G1356 A1357 U1358 C1359
A1360 G1361 C1361A C1362 A1363 U1364 G1365 C1366 C1367 G1368 C1369 G1370 G1371 U1372 G1373 A1374 A1375 U1376 A1377 C1378 G1379 U1380 U1381 C1382 C1383 C1384 G1385 G1386 U1390 U1391 G1392 U1393 A1394 C1395 A1396 C1397 A1398 C1399 C1400 G1401 C1402 C1403 C1404 G1405 U1406 C1407 A1408 C1409 G1410 C1411 C1412 A1413 U1414 G1415 A1418 G1419 C1420 G1421



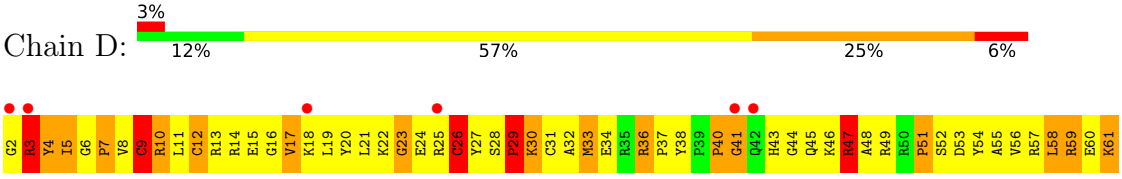
• Molecule 2: 30S RIBOSOMAL PROTEIN S2

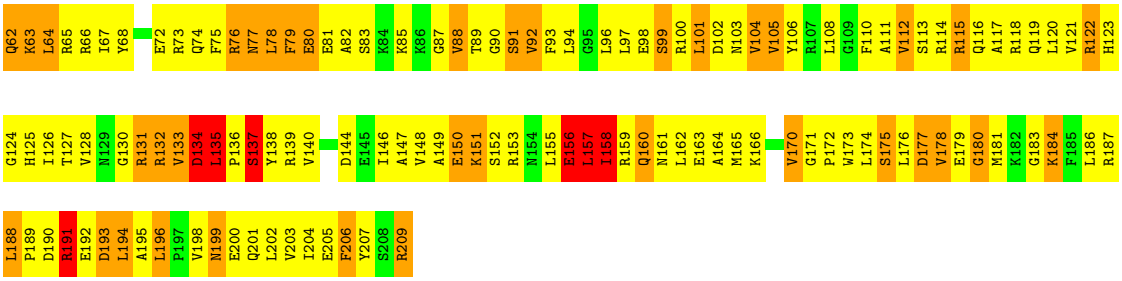


• Molecule 3: 30S RIBOSOMAL PROTEIN S3



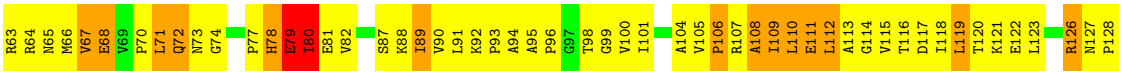
• Molecule 4: 30S RIBOSOMAL PROTEIN S4





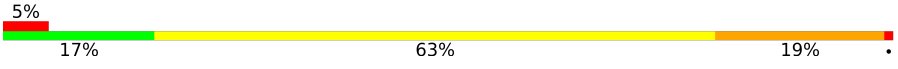
• Molecule 5: 30S RIBOSOMAL PROTEIN S5

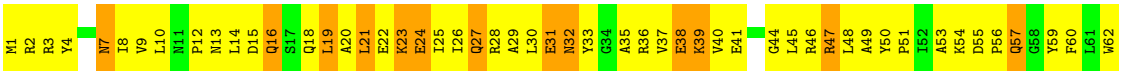
Chain E:

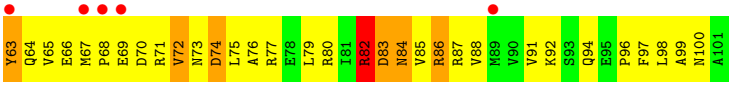


• Molecule 6: 30S RIBOSOMAL PROTEIN S6

Chain F:

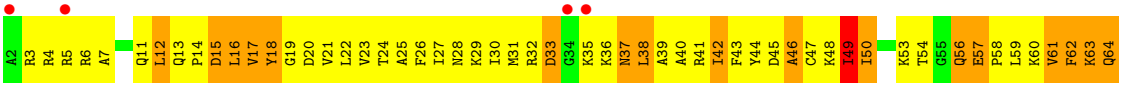






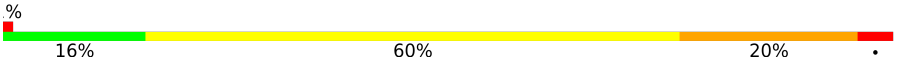
• Molecule 7: 30S RIBOSOMAL PROTEIN S7

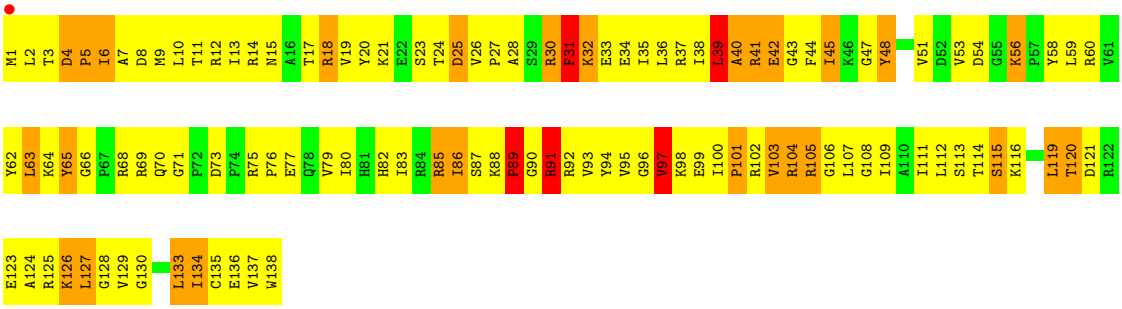
Chain G:

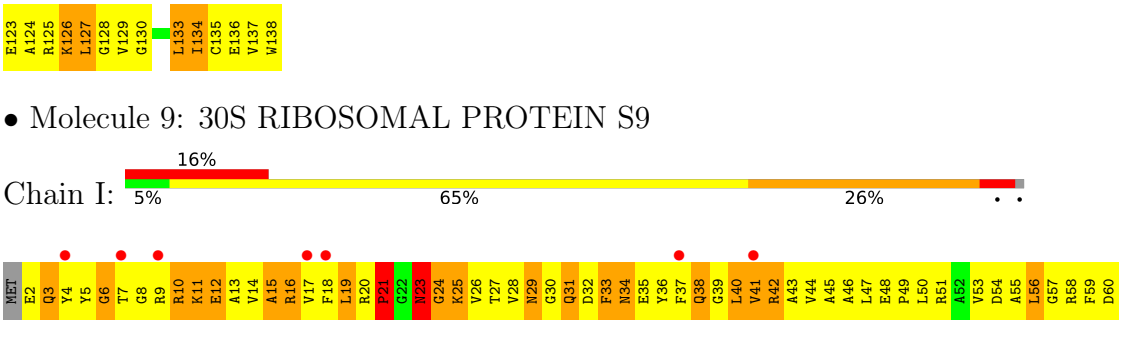


• Molecule 8: 30S RIBOSOMAL PROTEIN S8

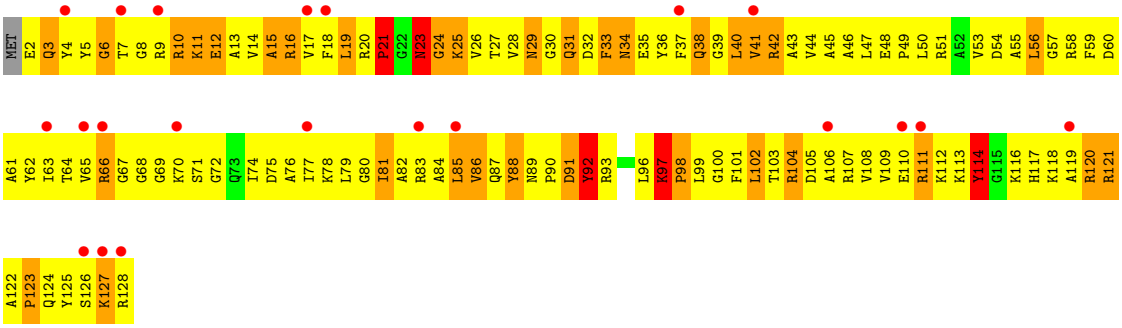
Chain H:



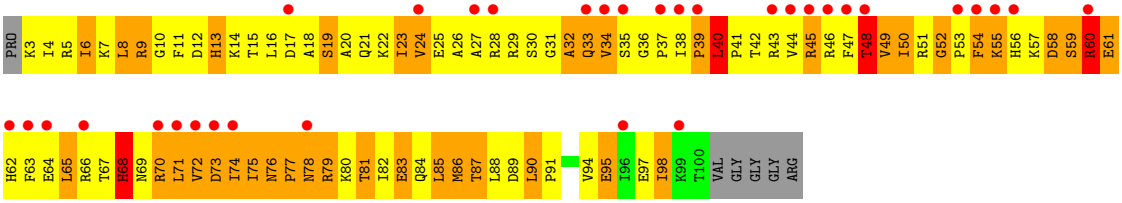




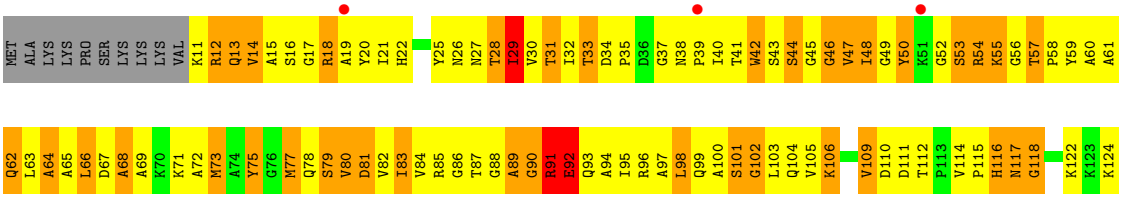
• Molecule 9: 30S RIBOSOMAL PROTEIN S9

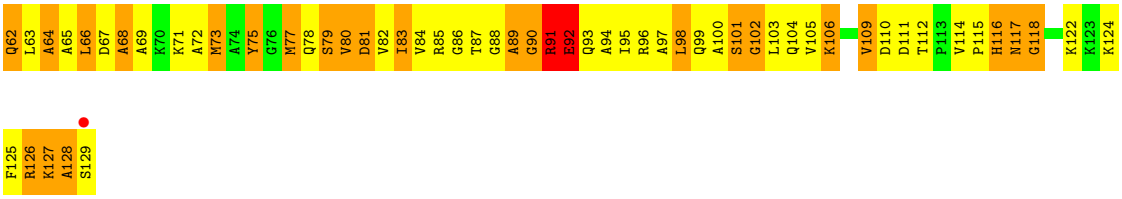


• Molecule 10: 30S RIBOSOMAL PROTEIN S10



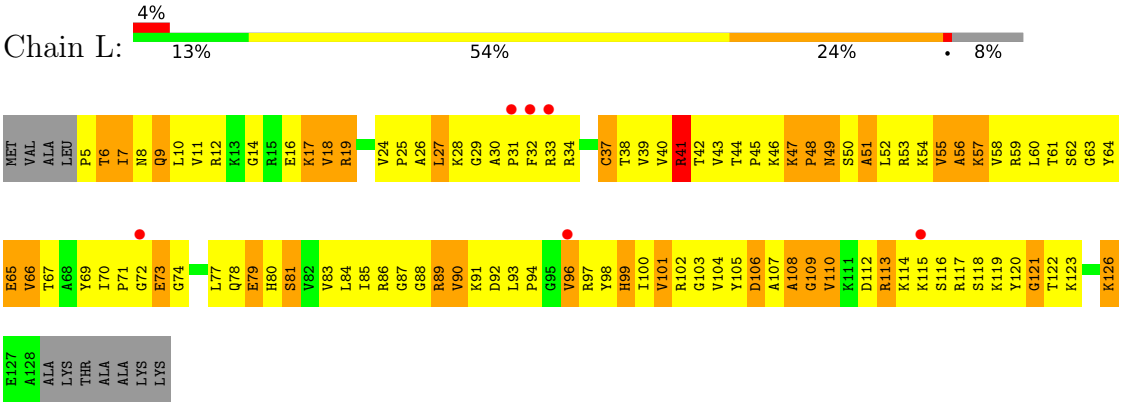
• Molecule 11: 30S RIBOSOMAL PROTEIN S11





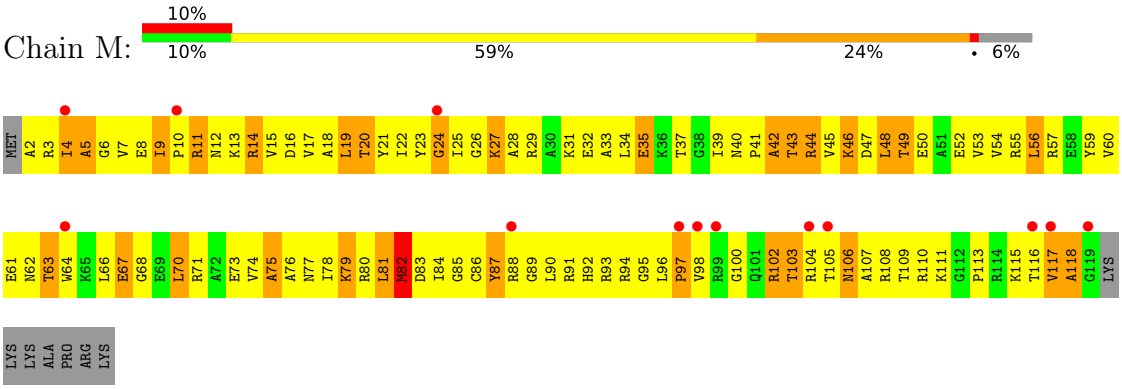
• Molecule 12: 30S RIBOSOMAL PROTEIN S12

Chain L:



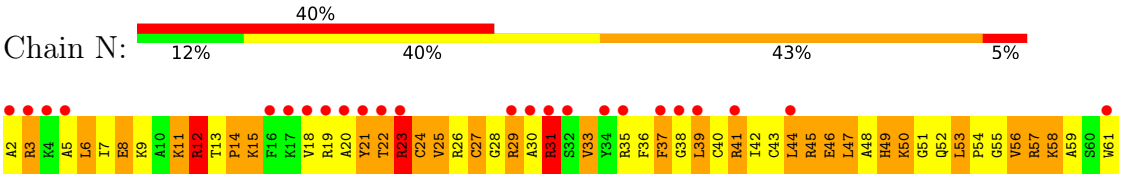
• Molecule 13: 30S RIBOSOMAL PROTEIN S13

Chain M:



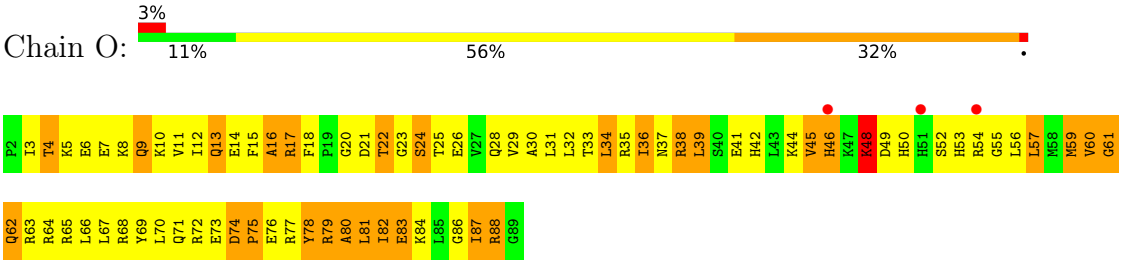
• Molecule 14: 30S RIBOSOMAL PROTEIN S14

Chain N:



• Molecule 15: 30S RIBOSOMAL PROTEIN S15

Chain O:

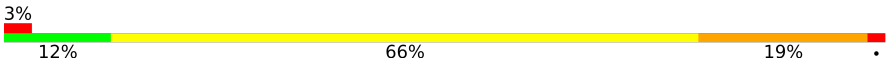


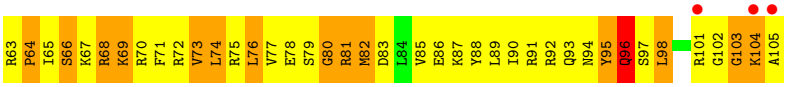
• Molecule 16: 30S RIBOSOMAL PROTEIN S16

Chain P:

• Molecule 17: 30S RIBOSOMAL PROTEIN S17

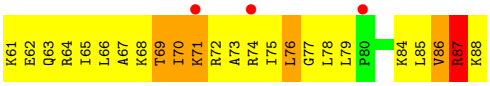
Chain Q:





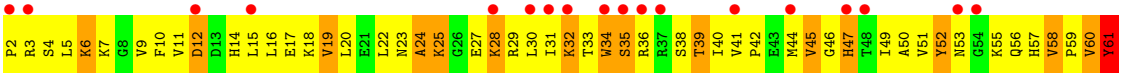
• Molecule 18: 30S RIBOSOMAL PROTEIN S18

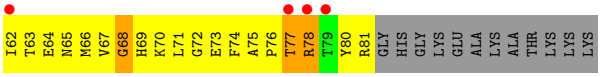
Chain R:



• Molecule 19: 30S RIBOSOMAL PROTEIN S19

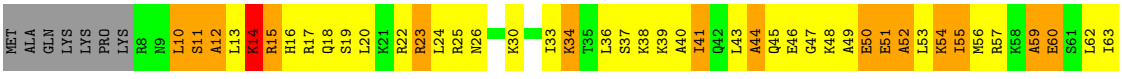
Chain S:

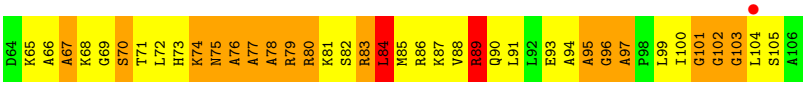




• Molecule 20: 30S RIBOSOMAL PROTEIN S20

Chain T:





• Molecule 21: 30S RIBOSOMAL PROTEIN THX

Chain V:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	402.84Å 402.84Å 174.28Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	141.42 – 3.65 148.66 – 3.64	Depositor EDS
% Data completeness (in resolution range)	92.6 (141.42-3.65) 89.6 (148.66-3.64)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.14	Depositor
$< I/\sigma(I) >$ [1]	2.44 (at 3.67Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.260 , 0.324 0.259 , 0.322	Depositor DCC
R_{free} test set	7116 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å^2)	107.1	Xtriage
Anisotropy	0.426	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.22 , 121.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.85	EDS
Total number of atoms	51680	wwPDB-VP
Average B, all atoms (Å^2)	93.0	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.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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	0.64	4/36387 (0.0%)	0.78	27/56789 (0.0%)
2	B	0.47	0/1935	0.79	0/2609
3	C	0.46	0/1636	0.77	0/2205
4	D	0.49	0/1733	0.70	0/2318
5	E	0.61	0/1162	0.88	2/1564 (0.1%)
6	F	0.42	0/856	0.72	0/1154
7	G	0.44	0/1276	0.77	1/1709 (0.1%)
8	H	0.65	0/1136	0.87	1/1527 (0.1%)
9	I	0.45	0/1029	0.73	0/1378
10	J	0.47	0/805	0.86	0/1082
11	K	0.49	0/900	0.81	0/1213
12	L	0.51	0/986	0.87	0/1320
13	M	0.40	0/947	0.73	0/1270
14	N	0.46	0/501	0.75	1/664 (0.2%)
15	O	0.51	0/745	0.74	0/992
16	P	0.58	0/716	0.88	1/963 (0.1%)
17	Q	0.58	0/870	0.90	2/1159 (0.2%)
18	R	0.45	0/603	0.75	0/799
19	S	0.47	0/661	0.82	0/890
20	T	0.48	0/765	0.79	0/1007
21	V	0.48	0/212	0.66	0/277
All	All	0.59	4/55861 (0.0%)	0.78	35/82889 (0.0%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	1	63
8	H	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
All	All	1	64

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	858	G	C5-C6	-6.65	1.35	1.42
1	A	1508	G	C5-C6	-5.12	1.37	1.42
1	A	574	A	C5-C6	-5.04	1.36	1.41
1	A	821	G	C5-C6	-5.03	1.37	1.42

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	290	C	N1-C1'-C2'	-8.80	102.33	112.00
1	A	1498	U	C2'-C3'-O3'	8.73	128.71	109.50
1	A	575	G	C2'-C3'-O3'	7.82	126.70	109.50
1	A	60	A	C2'-C3'-O3'	7.67	126.37	109.50
1	A	1454	G	N9-C1'-C2'	-7.34	103.92	112.00

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	1498	U	C3'

5 of 64 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	127	G	Sidechain
1	A	129	U	Sidechain
1	A	156	G	Sidechain
1	A	183	G	Sidechain
1	A	77	G	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	32508	0	16414	2582	0
2	B	1900	0	1951	415	0
3	C	1612	0	1677	503	0
4	D	1703	0	1764	410	0
5	E	1146	0	1207	238	0
6	F	843	0	857	150	0
7	G	1257	0	1296	278	0
8	H	1116	0	1177	221	0
9	I	1011	0	1043	277	0
10	J	792	0	835	283	0
11	K	885	0	904	167	0
12	L	970	0	1057	204	0
13	M	937	0	995	225	0
14	N	492	0	532	165	0
15	O	734	0	771	150	0
16	P	700	0	720	199	0
17	Q	857	0	930	177	0
18	R	597	0	668	137	0
19	S	647	0	673	182	0
20	T	763	0	861	206	0
21	V	208	0	221	52	0
22	D	1	0	0	0	0
22	N	1	0	0	0	0
All	All	51680	0	36553	6732	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 76.

The worst 5 of 6732 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)
4:D:158:ILE:H	4:D:158:ILE:CD1	1.57	1.15
1:A:243:A:H4'	1:A:244:U:H5'	1.22	1.12
4:D:176:LEU:HG	4:D:177:ASP:H	0.96	1.12
1:A:1250:A:H4'	9:I:68:GLY:HA2	1.31	1.12
1:A:1347:G:N2	1:A:1373:G:H2'	1.65	1.12

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	
2	B	232/256 (91%)	104 (45%)	70 (30%)	58 (25%)	0	0
3	C	204/239 (85%)	95 (47%)	59 (29%)	50 (24%)	0	0
4	D	206/208 (99%)	94 (46%)	62 (30%)	50 (24%)	0	0
5	E	148/161 (92%)	86 (58%)	38 (26%)	24 (16%)	0	3
6	F	99/101 (98%)	72 (73%)	15 (15%)	12 (12%)	0	5
7	G	153/155 (99%)	78 (51%)	45 (29%)	30 (20%)	0	1
8	H	136/138 (99%)	81 (60%)	32 (24%)	23 (17%)	0	2
9	I	125/128 (98%)	66 (53%)	38 (30%)	21 (17%)	0	2
10	J	96/104 (92%)	54 (56%)	19 (20%)	23 (24%)	0	0
11	K	117/129 (91%)	53 (45%)	33 (28%)	31 (26%)	0	0
12	L	122/135 (90%)	65 (53%)	28 (23%)	29 (24%)	0	0
13	M	116/126 (92%)	67 (58%)	28 (24%)	21 (18%)	0	1
14	N	58/60 (97%)	25 (43%)	19 (33%)	14 (24%)	0	0
15	O	86/88 (98%)	44 (51%)	23 (27%)	19 (22%)	0	1
16	P	81/88 (92%)	41 (51%)	18 (22%)	22 (27%)	0	0
17	Q	102/104 (98%)	66 (65%)	25 (24%)	11 (11%)	0	6
18	R	71/88 (81%)	38 (54%)	16 (22%)	17 (24%)	0	0
19	S	78/92 (85%)	43 (55%)	24 (31%)	11 (14%)	0	3
20	T	97/106 (92%)	31 (32%)	36 (37%)	30 (31%)	0	0
21	V	22/26 (85%)	14 (64%)	7 (32%)	1 (4%)	2	23
All	All	2349/2532 (93%)	1217 (52%)	635 (27%)	497 (21%)	0	1

5 of 497 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	15	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	17	PHE
2	B	39	ILE
2	B	78	GLN
2	B	82	ARG

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
2	B	202/220 (92%)	163 (81%)	39 (19%)	1 9
3	C	160/188 (85%)	127 (79%)	33 (21%)	1 7
4	D	180/180 (100%)	147 (82%)	33 (18%)	1 10
5	E	115/122 (94%)	98 (85%)	17 (15%)	3 18
6	F	90/90 (100%)	79 (88%)	11 (12%)	5 24
7	G	126/126 (100%)	105 (83%)	21 (17%)	2 14
8	H	119/119 (100%)	103 (87%)	16 (13%)	4 21
9	I	98/99 (99%)	71 (72%)	27 (28%)	0 3
10	J	87/91 (96%)	63 (72%)	24 (28%)	0 3
11	K	90/99 (91%)	73 (81%)	17 (19%)	1 9
12	L	104/111 (94%)	90 (86%)	14 (14%)	4 21
13	M	94/101 (93%)	80 (85%)	14 (15%)	3 17
14	N	49/49 (100%)	31 (63%)	18 (37%)	0 0
15	O	79/79 (100%)	68 (86%)	11 (14%)	3 20
16	P	72/74 (97%)	59 (82%)	13 (18%)	1 10
17	Q	96/96 (100%)	82 (85%)	14 (15%)	3 18
18	R	64/77 (83%)	55 (86%)	9 (14%)	3 19
19	S	71/79 (90%)	62 (87%)	9 (13%)	4 23
20	T	76/82 (93%)	68 (90%)	8 (10%)	7 30
21	V	19/21 (90%)	13 (68%)	6 (32%)	0 1

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1991/2103 (95%)	1637 (82%)	354 (18%)	2 11

5 of 354 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
11	K	91	ARG
15	O	22	THR
12	L	9	GLN
13	M	79	LYS
16	P	43	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 66 such sidechains are listed below:

Mol	Chain	Res	Type
15	O	50	HIS
17	Q	16	GLN
20	T	75	ASN
5	E	130	ASN
5	E	73	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	1511/1522 (99%)	253 (16%)	62 (4%)

5 of 253 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	8	A
1	A	9	G
1	A	31	G
1	A	32	A
1	A	39	G

5 of 62 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	559	A
1	A	1285	A
1	A	840	C

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1281	U
1	A	1397	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 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

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^{th} 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	1512/1522 (99%)	0.83	169 (11%) 5 3	12, 87, 171, 195	0
2	B	234/256 (91%)	-0.14	3 (1%) 77 65	21, 87, 149, 181	0
3	C	206/239 (86%)	0.38	20 (9%) 7 5	31, 115, 164, 182	0
4	D	208/208 (100%)	0.04	6 (2%) 51 37	15, 83, 140, 189	0
5	E	150/161 (93%)	0.01	1 (0%) 87 80	4, 55, 110, 132	0
6	F	101/101 (100%)	0.12	5 (4%) 28 20	32, 111, 151, 174	0
7	G	155/155 (100%)	-0.16	6 (3%) 39 27	27, 116, 156, 189	0
8	H	138/138 (100%)	-0.07	1 (0%) 87 80	0, 45, 102, 156	0
9	I	127/128 (99%)	0.66	21 (16%) 1 1	30, 120, 164, 186	0
10	J	98/104 (94%)	1.32	33 (33%) 0 0	43, 122, 168, 183	0
11	K	119/129 (92%)	0.17	4 (3%) 45 33	21, 87, 147, 195	0
12	L	124/135 (91%)	0.13	6 (4%) 30 21	2, 77, 127, 154	0
13	M	118/126 (93%)	0.47	13 (11%) 5 4	45, 118, 163, 195	0
14	N	60/60 (100%)	1.98	24 (40%) 0 0	50, 118, 171, 186	0
15	O	88/88 (100%)	-0.10	3 (3%) 45 33	23, 79, 143, 195	0
16	P	83/88 (94%)	0.06	2 (2%) 59 45	0, 60, 119, 138	0
17	Q	104/104 (100%)	-0.06	3 (2%) 51 37	8, 59, 132, 195	0
18	R	73/88 (82%)	-0.04	3 (4%) 37 26	3, 81, 147, 182	0
19	S	80/92 (86%)	1.33	22 (27%) 0 0	49, 124, 178, 181	0
20	T	99/106 (93%)	-0.33	1 (1%) 82 72	10, 66, 124, 141	0
21	V	24/26 (92%)	2.00	11 (45%) 0 0	84, 114, 145, 180	0
All	All	3901/4054 (96%)	0.46	357 (9%) 9 6	0, 90, 162, 195	0

The worst 5 of 357 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
14	N	4	LYS	10.9
19	S	3	ARG	8.4
17	Q	105	ALA	7.9
4	D	42	GLN	7.8
21	V	24	ARG	7.3

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^{th} 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^2)	Q<0.9
22	ZN	D	306	1/1	0.97	0.45	74,74,74,74	0
22	ZN	N	307	1/1	0.97	0.10	74,74,74,74	0

6.5 Other polymers (i)

There are no such residues in this entry.