wwPDB X-ray Structure Validation Summary Report (i)

Sep 10, 2023 – 05:12 PM EDT

PDB ID	:	4JK2
Title	:	X-ray crystal structure of Escherichia coli sigma70 holoenzyme in complex with guanosine pentaphosphate (pppGpp)
Authors	:	Murakami, K.S.
Deposited on	:	2013-03-09
Resolution	:	4.20 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 (i)) were used in the production of this report:

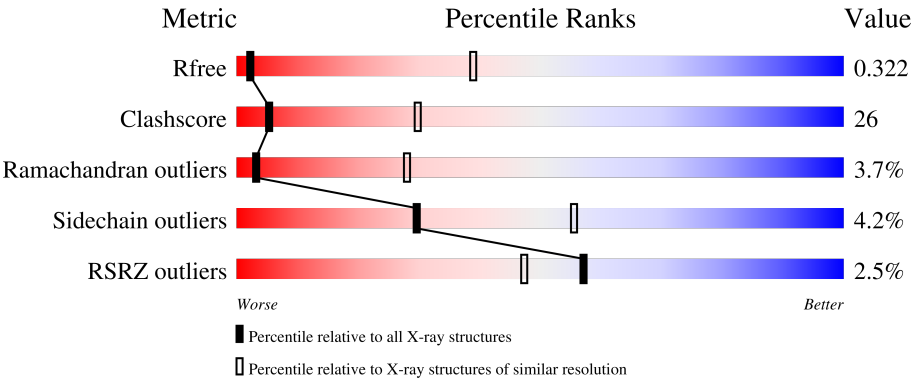
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 4.20 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.



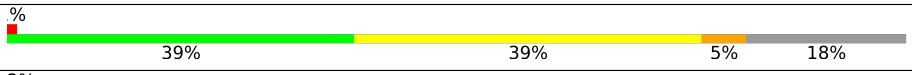
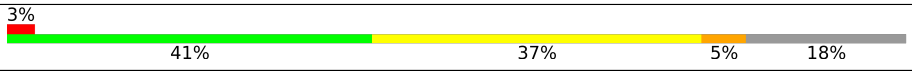
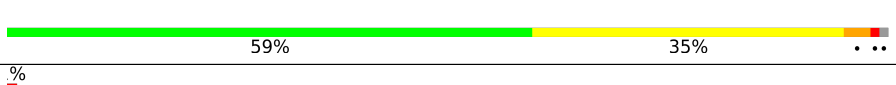
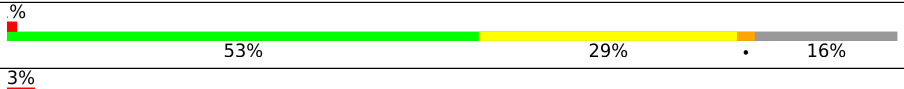
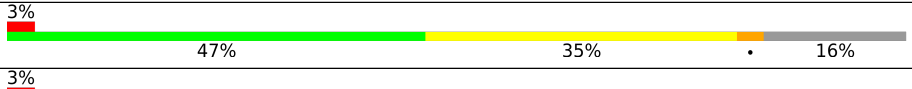
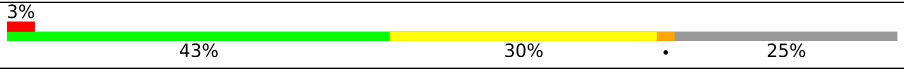
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1005 (4.62-3.78)
Clashscore	141614	1044 (4.60-3.80)
Ramachandran outliers	138981	1000 (4.60-3.80)
Sidechain outliers	138945	1007 (4.62-3.78)
RSRZ outliers	127900	1063 (4.70-3.70)

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	329	59% 36%
1	B	329	38% 27% 33%
1	F	329	46% 22% 30%
1	G	329	40% 24% 34%
2	C	1342	53% 41% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	1342	
3	D	1407	
3	I	1407	
4	E	91	
4	J	91	
5	X	613	
5	Y	613	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
7	0O2	D	1503	-	-	X	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 56129 atoms, of which 10 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 protein called Escherichia coli RNA polymerase alpha subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	323	Total	C	N	O	S	0	0	0
			2514	1571	443	492	8			
1	B	221	Total	C	N	O	S	0	0	0
			1706	1065	300	335	6			
1	F	229	Total	C	N	O	S	0	0	0
			1775	1106	313	350	6			
1	G	217	Total	C	N	O	S	0	0	0
			1671	1045	293	327	6			

- Molecule 2 is a protein called Escherichia coli RNA polymerase beta subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1335	Total	C	N	O	S	0	0	0
			10523	6601	1836	2043	43			
2	H	1335	Total	C	N	O	S	0	0	0
			10523	6601	1836	2043	43			

- Molecule 3 is a protein called Escherichia coli RNA polymerase beta' subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1160	Total	C	N	O	S	0	0	0
			9060	5695	1621	1697	47			
3	I	1160	Total	C	N	O	S	0	0	0
			9060	5695	1621	1697	47			

- Molecule 4 is a protein called Escherichia coli RNA polymerase omega subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	90	Total	C	N	O	S	0	0	0
			708	430	136	141	1			
4	J	76	Total	C	N	O	S	0	0	0
			605	368	115	121	1			

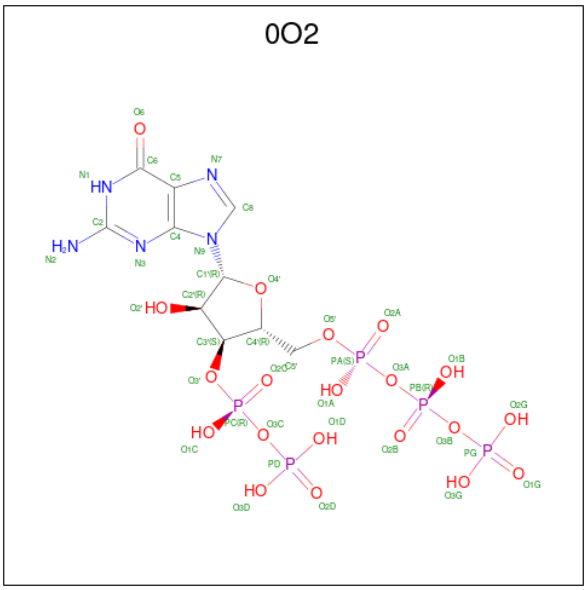
- Molecule 5 is a protein called Escherichia coli RNA polymerase sigma70 subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	X	517	Total	C	N	O	S	0	0	0
			4198	2621	745	806	26			
5	Y	458	Total	C	N	O	S	0	0	0
			3732	2335	671	703	23			

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	2	Total	Zn	0	0
			2	2		
6	I	2	Total	Zn	0	0
			2	2		

- Molecule 7 is guanosine 5'-(tetrahydrogen triphosphate) 3'-(trihydrogen diphosphate) (three-letter code: 0O2) (formula: $C_{10}H_{18}N_5O_{20}P_5$).



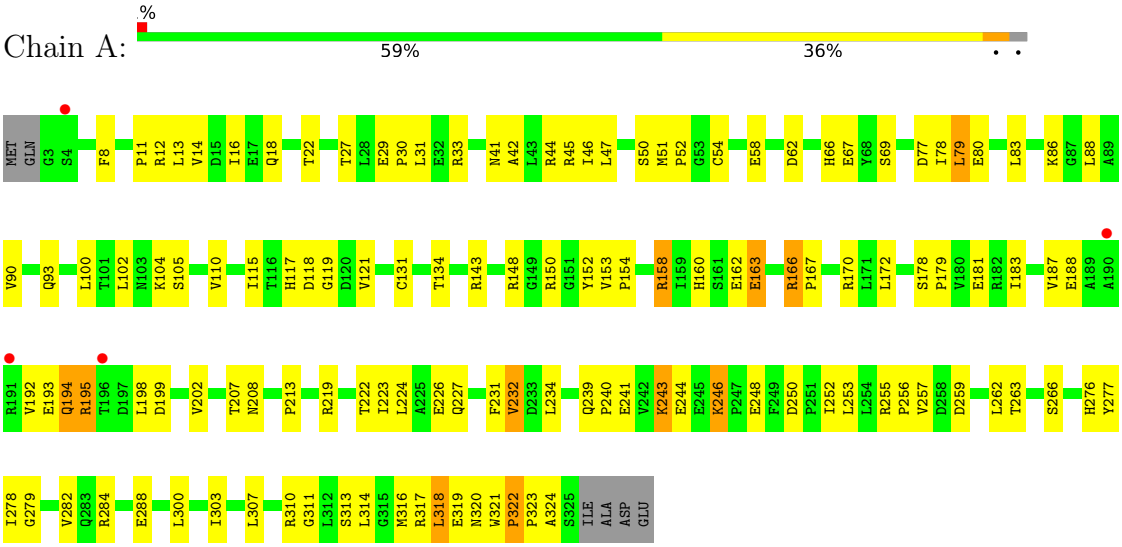
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
7	D	1	Total	C	H	N	O	P	0	0
			50	10	10	5	20	5		

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.

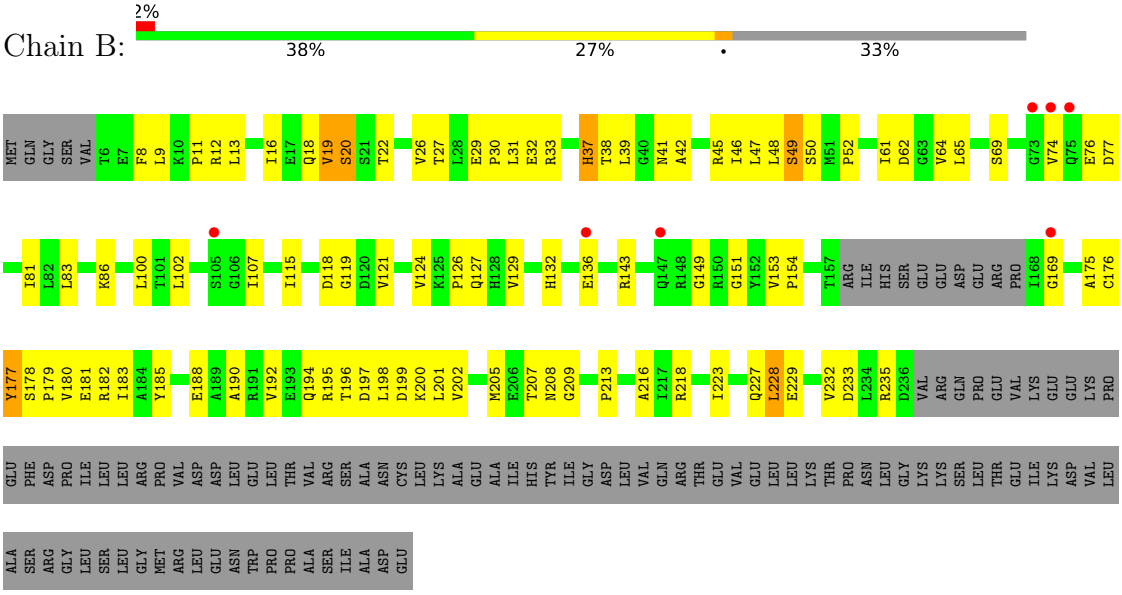
• Molecule 1: Escherichia coli RNA polymerase alpha subunit

Chain A:



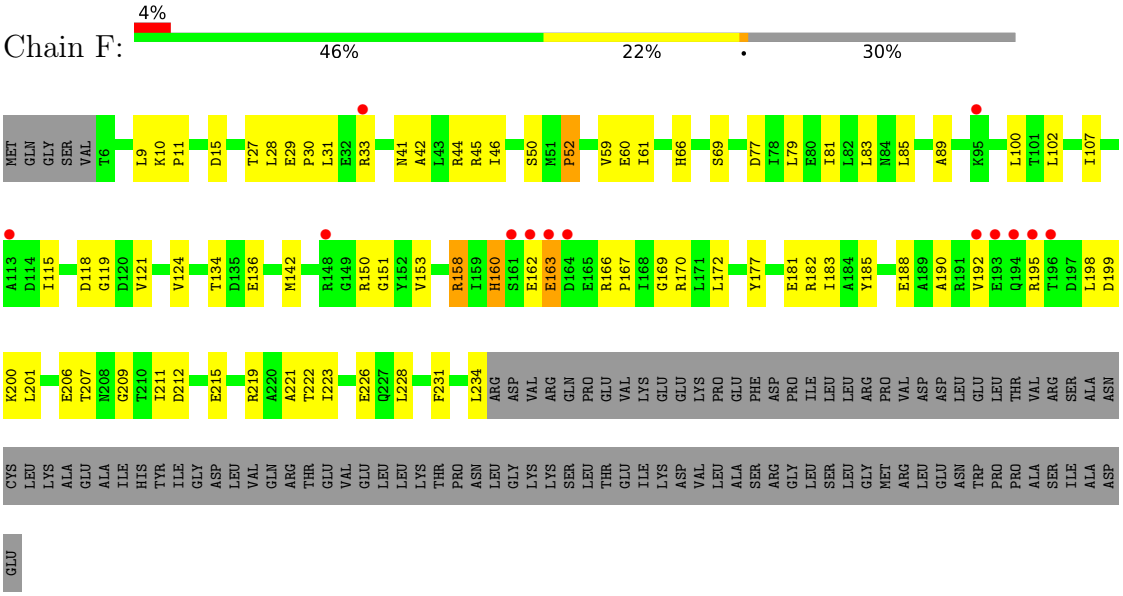
• Molecule 1: Escherichia coli RNA polymerase alpha subunit

Chain B:



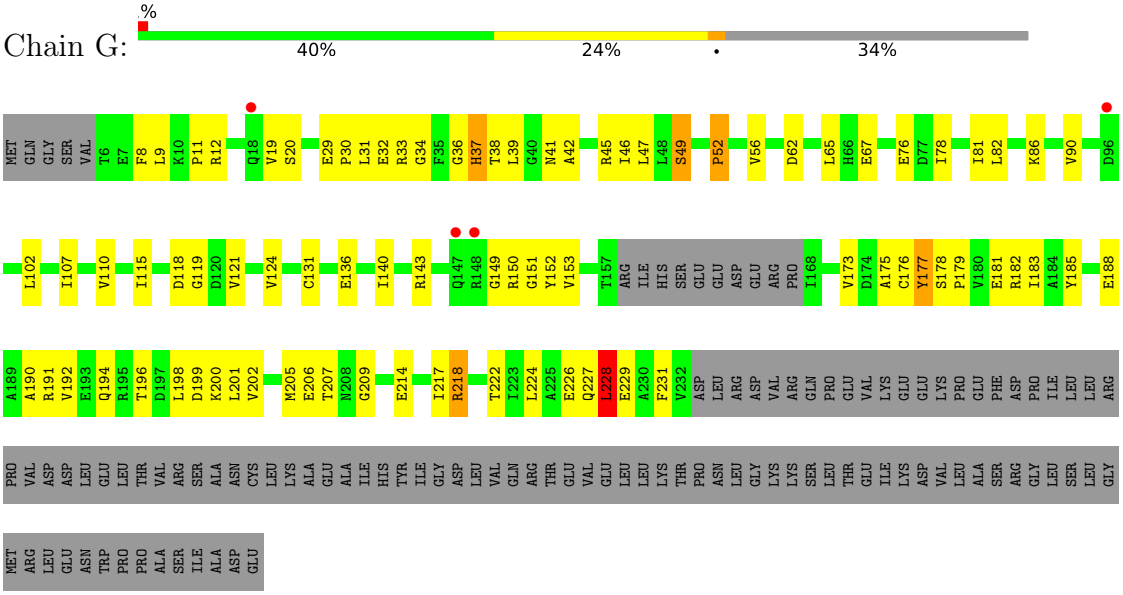
• Molecule 1: Escherichia coli RNA polymerase alpha subunit

Chain F:



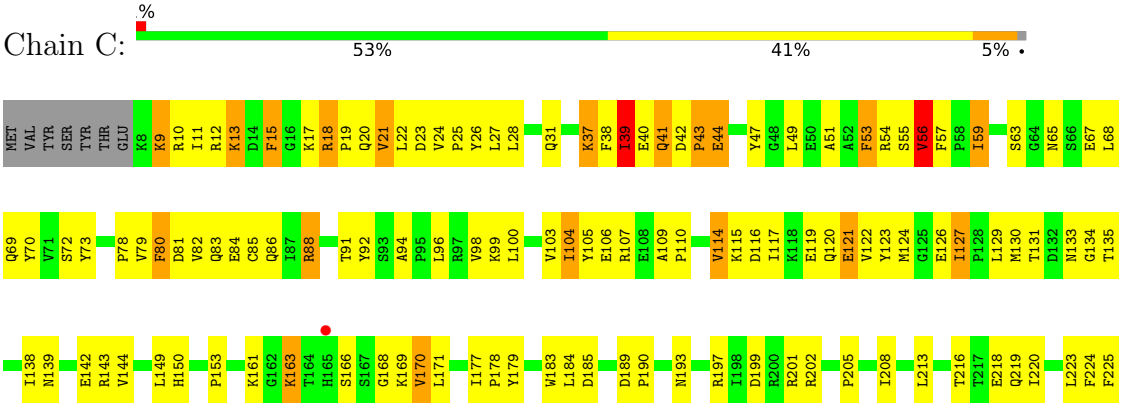
• Molecule 1: Escherichia coli RNA polymerase alpha subunit

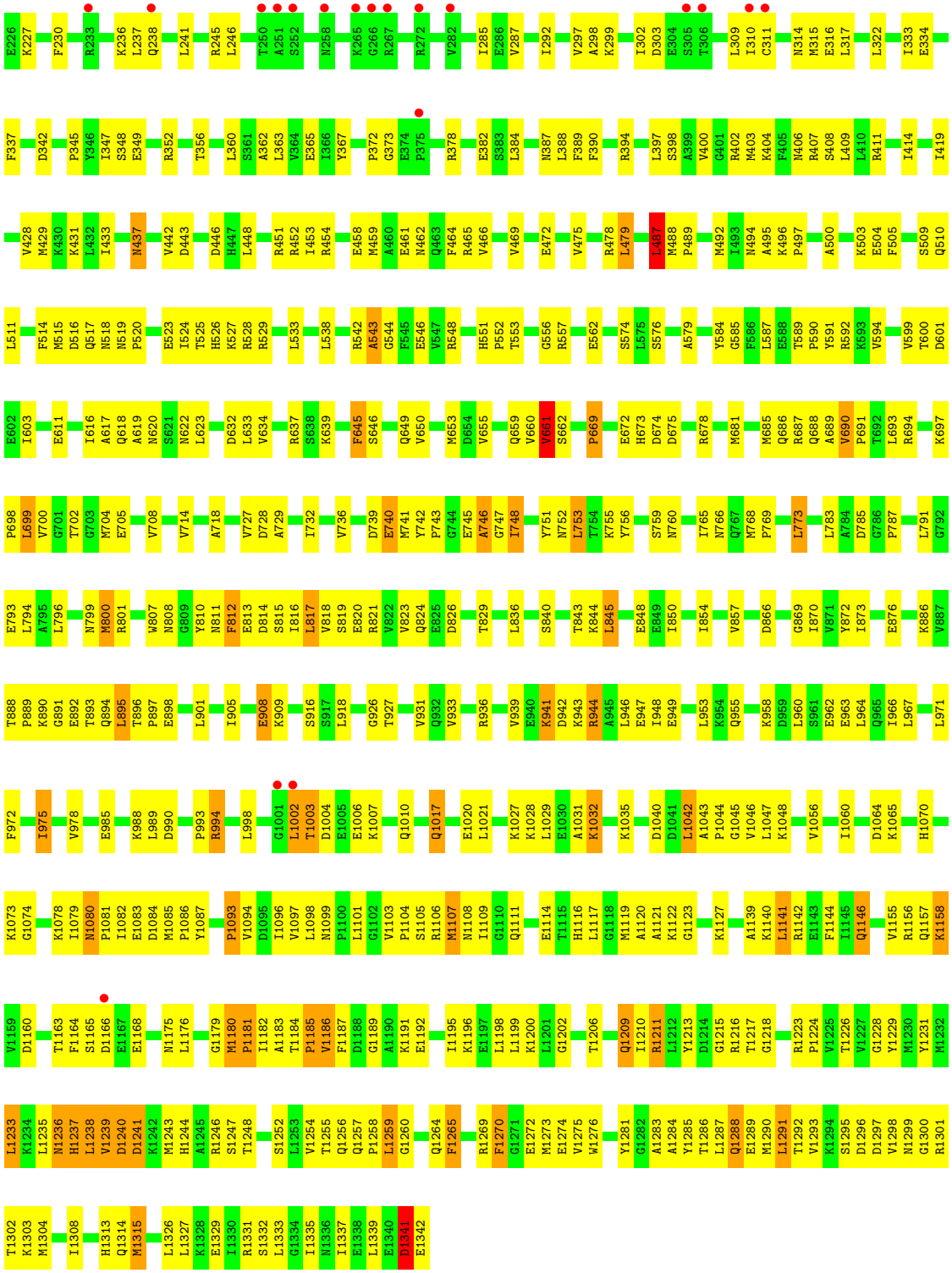
Chain G:



• Molecule 2: Escherichia coli RNA polymerase beta subunit

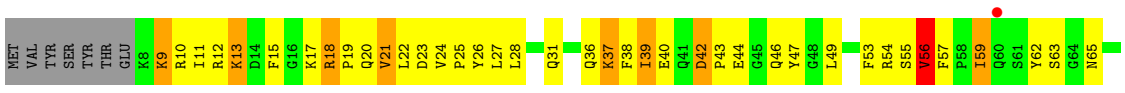
Chain C:

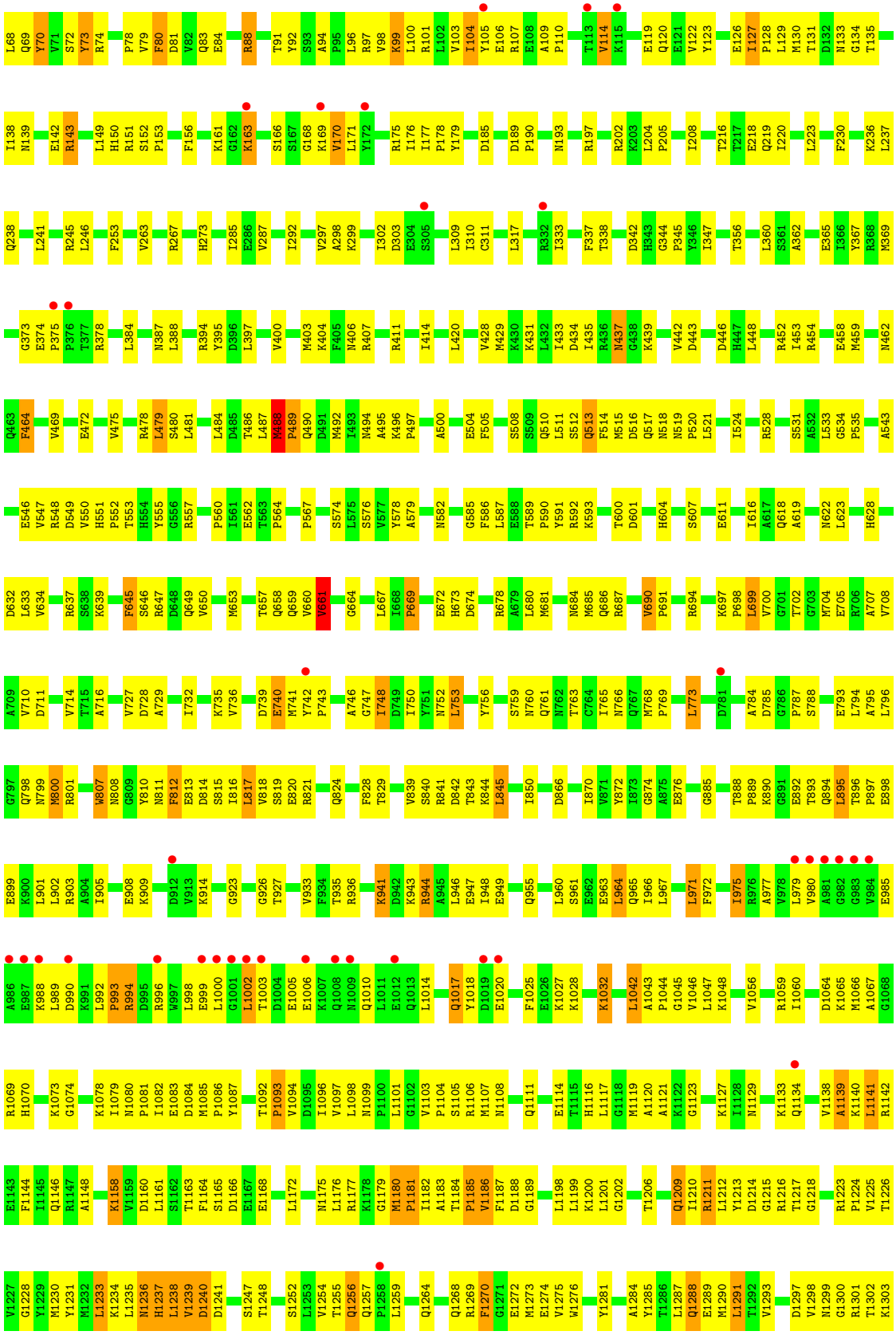


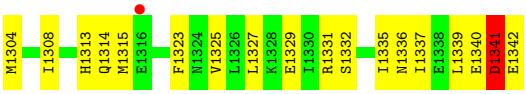


• Molecule 2: Escherichia coli RNA polymerase beta subunit

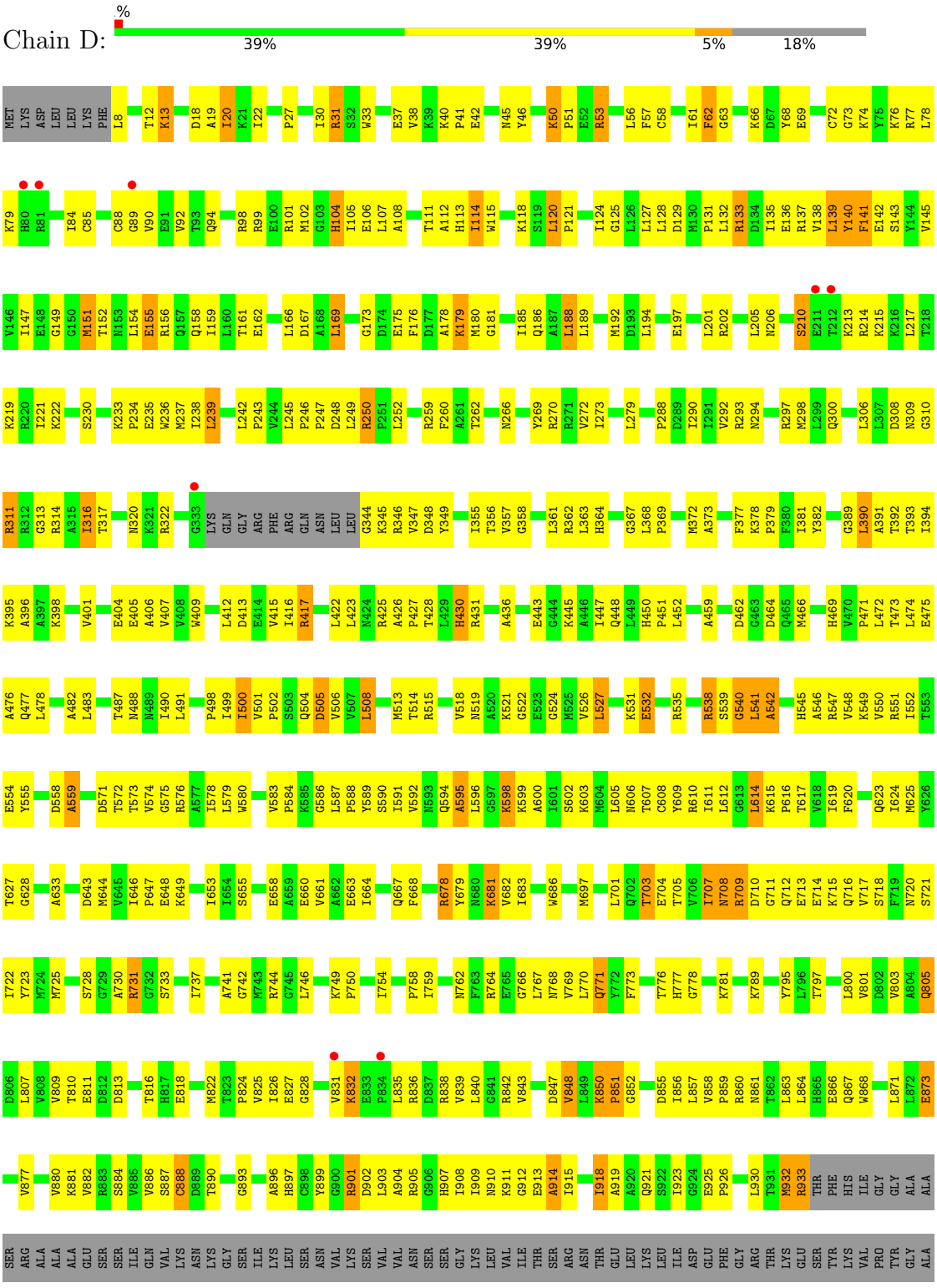
Chain H:

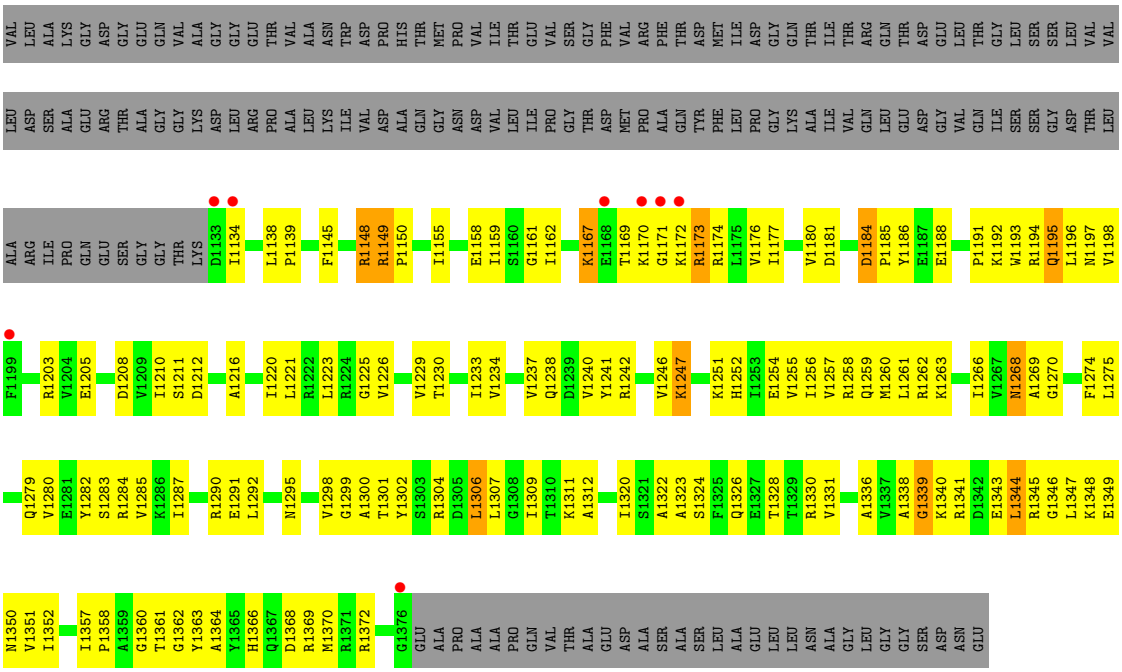



L68 Q69 Y70 V71 S72 Y73 R74 P78 V79 F80 D81 V82 Q83 E84 R88 T91 Y92 S93 A94 P95 L96 R97 V98 K99 L100 R101 L102 V103 I104 Y105 E106 R107 E108 A109 P110 T113 V114 K115 E119 Q120 E121 V122 Y123 E126 I127 P128 L129 M130 T131 D132 N133 G134 T135
I138 N139 E142 R143 L149 H150 R151 S152 P153 F156 K161 G162 K163 S166 S167 G168 K169 V170 L171 Y172 R175 I176 I177 P178 Y179 D185 D189 P190 N193 R197 R202 K203 L204 P205 I208 T216 T217 E218 Q219 I220 L223 F230 K236 L237
Q238 L241 R245 L246 F253 V263 R267 H273 I285 E286 V287 I292 V297 A298 K299 I302 D303 E304 S306 L309 I310 C311 L317 R332 I333 F337 T338 D342 H343 G344 P345 Y346 I347 T356 T360 S361 A362 E365 I366 Y367 R368 M369
G373 E374 P375 P376 T377 R378 L384 N387 L388 R394 Y395 D396 L397 V400 M403 K404 F405 N406 R407 R411 I414 L420 V428 M429 K430 K431 L432 I433 D434 I435 R436 N437 G438 K439 V442 D443 D446 H447 L448 R452 I453 R454 E458 M459 N462
Q463 F464 V469 E472 V475 R478 L479 S480 L481 L484 D485 T486 L487 M488 P489 Q490 D491 M492 I493 N494 A495 K496 P497 A500 E504 F505 S508 S509 Q510 L511 S512 Q513 F514 M515 D516 Q517 N518 N519 P520 L521 I524 R528 S531 A532 L533 G534 P535 A543
E546 V547 R548 D549 V550 H551 P552 T553 H554 Y555 G556 R557 P560 I561 E562 T563 P564 P567 S574 L575 S576 V577 Y578 A579 N582 G585 F586 L587 E588 T589 P590 Y591 R592 K593 T600 D601 H604 S607 E611 I616 A617 Q618 A619 N622 L623 H628
D632 L633 V634 R637 S638 K639 F645 S646 R647 D648 Q649 V650 M653 T657 Q658 Q659 V660 V661 G664 L667 T668 P669 E672 H673 D674 R678 A679 L680 M681 N684 H685 Q686 R687 V690 P691 R694 K697 P698 L699 V700 G701 T702 G703 M704 E705 R706 A707 V708
A709 V710 D711 V714 T715 A716 V727 D728 A729 I732 K735 V736 D739 E740 M741 Y742 P743 A746 G747 T748 D749 I750 Y751 N752 L753 Y756 S759 N760 Q761 N762 T763 C764 I765 N766 Q767 M768 P769 L773 D781 A784 D785 G786 P787 S788 E793 L794 A795 L796
G797 Q798 N799 M800 R801 W807 N808 G809 Y810 N811 F812 E813 D814 S815 I816 L817 V818 S819 E820 R821 Q824 F828 T829 V839 S840 R841 D842 T843 K844 L845 I850 D866 I870 V871 Y872 I873 G874 A875 E876 G885 T888 P889 K890 G891 E892 T893 Q894 L895 T896 P897 E898
E899 K900 L901 L902 R903 A904 I905 E908 K909 D912 V913 K914 G923 G926 T927 V933 F934 T935 R936 K941 D942 K943 R944 A945 L946 E947 I948 E949 Q955 L960 S961 E962 E963 L964 Q965 I966 L967 L971 F972 I975 R976 A977 V978 L979 V980 A981 G982 G983 V984 E985
A986 E987 K988 L989 D990 K991 L992 P993 R994 D995 R996 W997 L998 E999 L1000 G1001 L1002 T1003 D1004 E1005 E1006 K1007 Q1008 N1009 Q1010 L1011 E1012 Q1013 L1014 Q1017 Y1018 D1019 E1020 F1025 E1026 K1027 K1028 K1032 L1042 A1043 P1044 G1045 V1046 L1047 K1048 V1056 R1059 I1060 D1064 K1065 M1066 A1067 G1068
R1069 H1070 K1073 G1074 K1078 I1079 N1080 P1081 I1082 E1083 D1084 M1085 P1086 Y1087 T1092 P1093 V1094 D1095 I1096 V1097 L1098 N1099 P1100 L1101 G1102 V1103 P1104 S1105 R1106 M1107 N1108 Q1111 E1114 T1115 H1116 L1117 G1118 M1119 A1120 A1121 K1122 G1123 K1127 I1128 N1129 K1133 Q1134 V1138 A1139 K1140 L1141 R1142
E1143 F1144 I1145 Q1146 R1147 A1148 K1158 V1159 D1160 L1161 S1162 T1163 F1164 S1165 D1166 E1167 E1168 L1172 N1175 L1176 R1177 K1178 G1179 M1180 P1181 I1182 A1183 T1184 P1185 V1186 F1187 D1188 G1189 L1198 L1199 K1200 L1201 G1202 T1206 Q1209 I1210 R1211 L1212 Y1213 D1214 G1215 R1216 T1217 G1218 R1223 P1224 V1225 T1226
V1227 G1228 Y1229 M1230 Y1231 M1232 L1233 K1234 L1235 N1236 H1237 L1238 V1239 D1240 D1241 S1247 T1248 S1252 L1253 V1254 T1255 Q1256 Q1257 P1258 L1259 Q1264 Q1268 R1269 F1270 G1271 E1272 M1273 E1274 V1275 W1276 Y1281 A1284 Y1285 T1286 L1287 Q1288 E1289 M1290 L1291 T1292 V1293 D1297 V1298 N1299 G1300 R1301 T1302 K1303

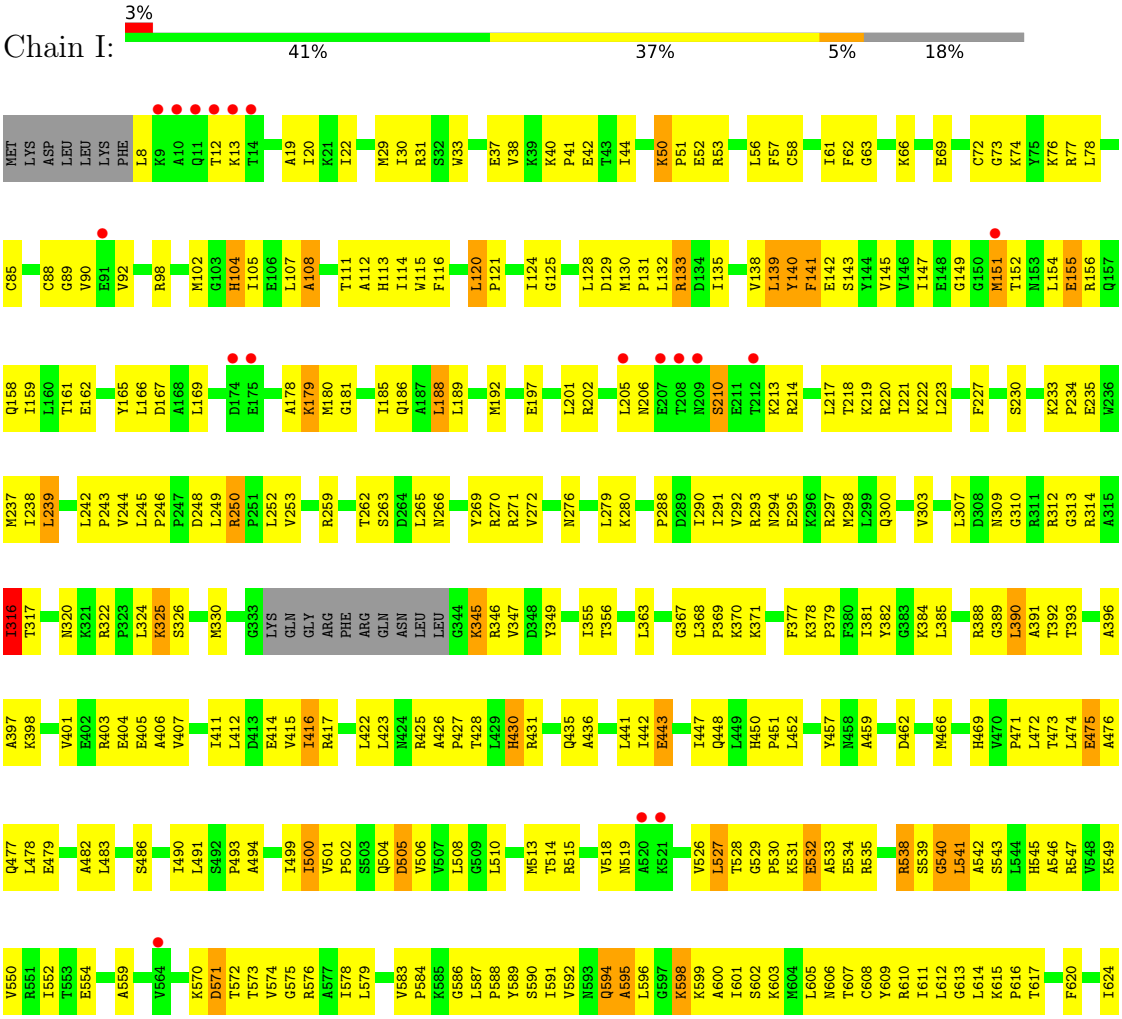


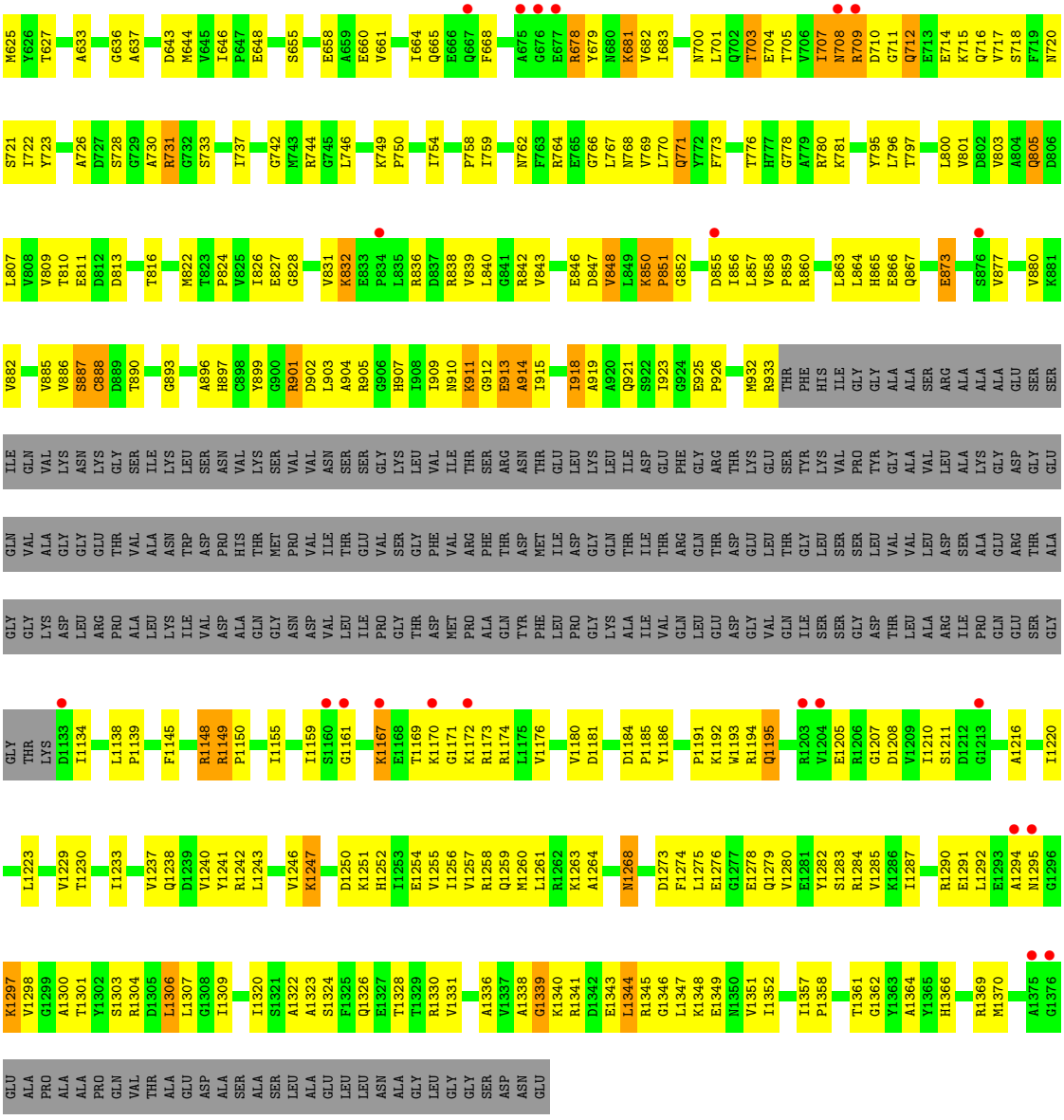
• Molecule 3: Escherichia coli RNA polymerase beta' subunit





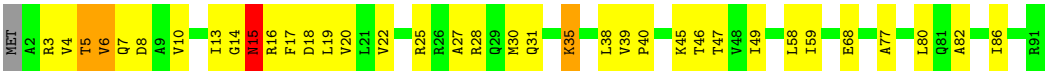
• Molecule 3: Escherichia coli RNA polymerase beta' subunit





• Molecule 4: Escherichia coli RNA polymerase omega subunit

Chain E:

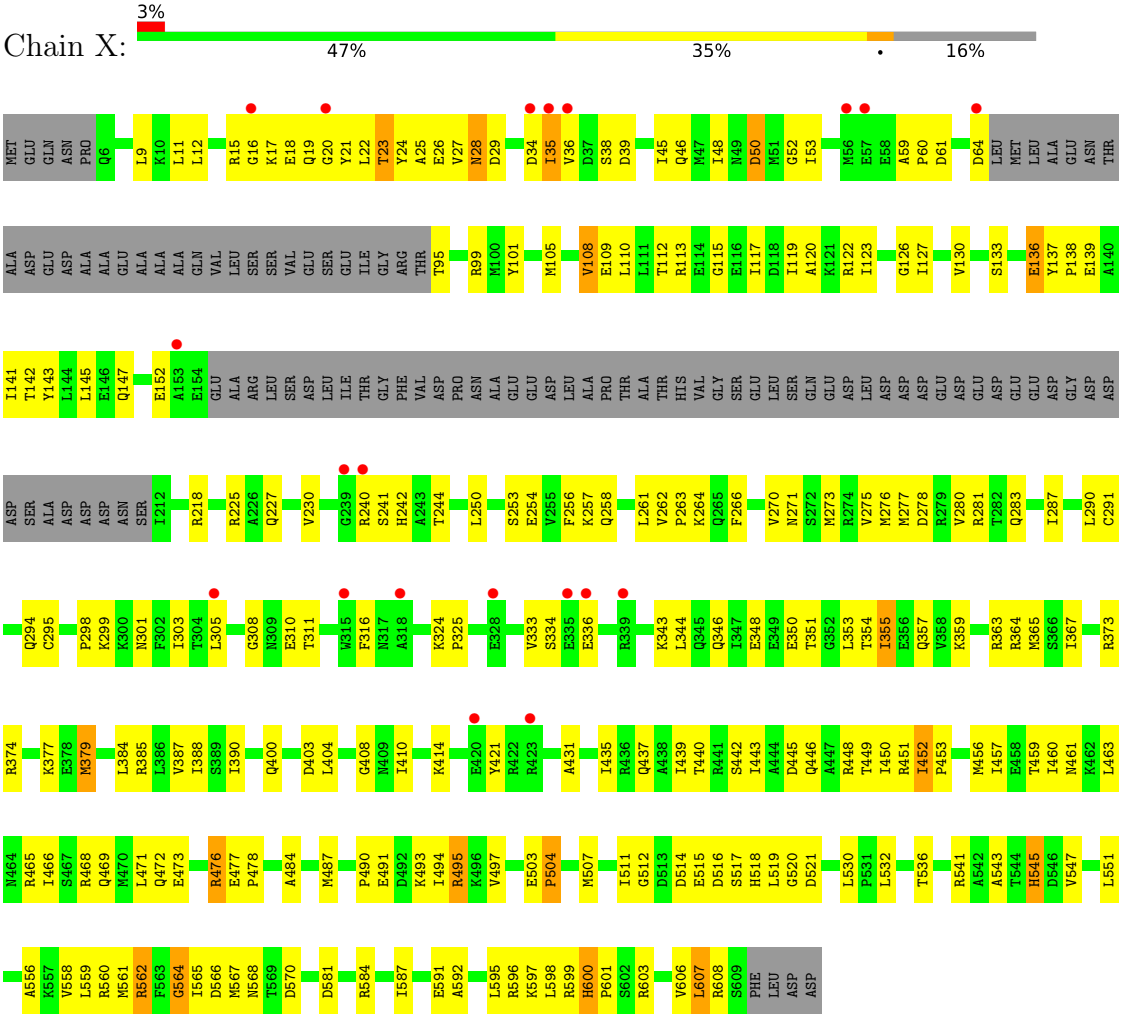


• Molecule 4: Escherichia coli RNA polymerase omega subunit

Chain J:

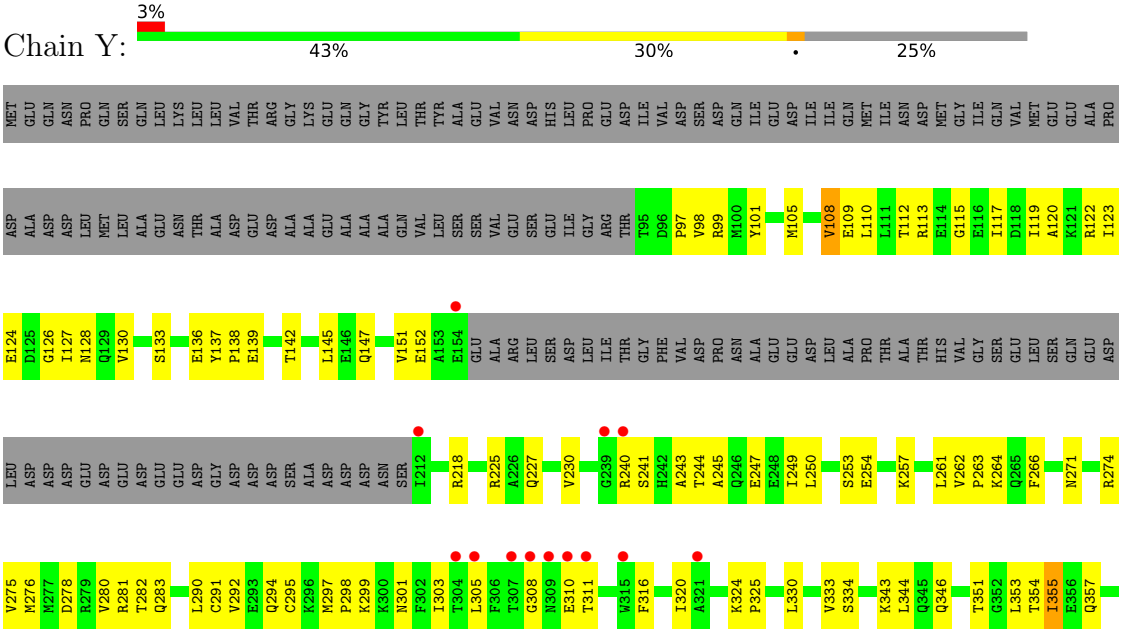
• Molecule 5: Escherichia coli RNA polymerase sigma70 subunit

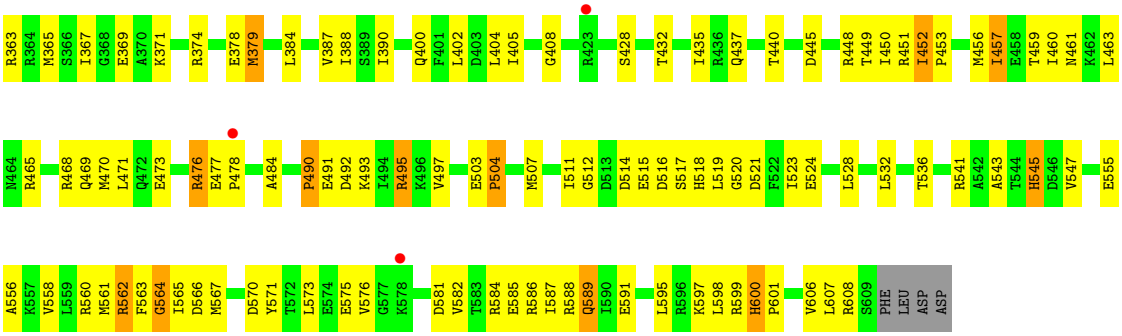
Chain X:



• Molecule 5: Escherichia coli RNA polymerase sigma70 subunit

Chain Y:




R363 R364 M365 S366 I367 G368 E369 A370 K371 R374 E378 M379 L384 V387 I388 S389 I390 Q400 F401 L402 D403 L404 I405 G408 R423 S428 T432 I435 R436 Q437 T440 D445 R448 T449 I450 R451 I452 P453 M456 I457 E458 T459 I460 N461 K462 L463
M464 R465 R468 Q469 M470 L471 Q472 E473 R476 E477 P478 A484 P490 E491 D492 K493 I494 R495 K496 V497 E503 P504 M507 I511 G512 D513 D514 E515 D516 S517 H518 L519 G520 D521 F522 I523 E524 L528 L532 T536 R541 A542 A543 T544 H545 D546 V547 E555
A556 K557 V558 L559 R560 M561 R562 F563 G564 I565 D566 M567 D570 Y571 T572 L573 E574 E575 V576 G577 K578 D581 V582 T583 R584 E585 R586 I587 R588 Q589 T590 E591 L595 R596 K597 L598 R599 H600 P601 V606 L607 R608 S609 PHE LEU ASP ASP

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	185.32Å 205.41Å 309.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.94 – 4.20 30.04 – 4.20	Depositor EDS
% Data completeness (in resolution range)	80.4 (29.94-4.20) 70.4 (30.04-4.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.36 (at 4.26Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, R_{free}	0.244 , 0.322 0.244 , 0.322	Depositor DCC
R_{free} test set	3506 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	155.9	Xtriage
Anisotropy	0.174	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 57.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	56129	wwPDB-VP
Average B, all atoms (Å^2)	76.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.28% 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, 0O2

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	0.19	0/2548	0.36	0/3454
1	B	0.19	0/1725	0.39	0/2337
1	F	0.19	0/1797	0.38	0/2436
1	G	0.19	0/1690	0.37	0/2290
2	C	0.20	0/10690	0.38	0/14423
2	H	0.20	0/10690	0.37	0/14423
3	D	0.20	0/9198	0.38	0/12413
3	I	0.20	0/9198	0.38	0/12413
4	E	0.19	0/710	0.38	0/956
4	J	0.19	0/607	0.37	0/817
5	X	0.19	0/4253	0.36	0/5719
5	Y	0.20	0/3783	0.36	0/5083
All	All	0.20	0/56889	0.38	0/76764

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	2514	0	2566	110	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1706	0	1738	91	0
1	F	1775	0	1800	79	0
1	G	1671	0	1706	91	0
2	C	10523	0	10546	600	0
2	H	10523	0	10546	574	0
3	D	9060	0	9257	658	0
3	I	9060	0	9257	591	0
4	E	708	0	719	52	0
4	J	605	0	612	33	0
5	X	4198	0	4250	197	0
5	Y	3732	0	3809	157	0
6	D	2	0	0	0	0
6	I	2	0	0	0	0
7	D	40	10	16	9	0
All	All	56119	10	56822	2973	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 26.

The worst 5 of 2973 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)
3:I:1173:ARG:HA	3:I:1174:ARG:HB2	1.23	1.20
3:D:1173:ARG:HA	3:D:1174:ARG:HB2	1.24	1.18
2:H:488:MET:HB2	2:H:490:GLN:H	1.07	1.11
3:I:850:LYS:HD2	3:I:851:PRO:HD2	1.30	1.09
2:H:1073:LYS:HD3	3:I:462:ASP:HB3	1.28	1.08

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	
1	A	321/329 (98%)	266 (83%)	40 (12%)	15 (5%)	2	24
1	B	217/329 (66%)	187 (86%)	22 (10%)	8 (4%)	3	28
1	F	227/329 (69%)	193 (85%)	28 (12%)	6 (3%)	5	35
1	G	213/329 (65%)	186 (87%)	22 (10%)	5 (2%)	6	37
2	C	1333/1342 (99%)	1073 (80%)	208 (16%)	52 (4%)	3	27
2	H	1333/1342 (99%)	1078 (81%)	206 (16%)	49 (4%)	3	28
3	D	1154/1407 (82%)	926 (80%)	182 (16%)	46 (4%)	3	26
3	I	1154/1407 (82%)	925 (80%)	184 (16%)	45 (4%)	3	27
4	E	88/91 (97%)	77 (88%)	6 (7%)	5 (6%)	1	21
4	J	74/91 (81%)	64 (86%)	6 (8%)	4 (5%)	2	22
5	X	511/613 (83%)	450 (88%)	46 (9%)	15 (3%)	4	32
5	Y	454/613 (74%)	409 (90%)	34 (8%)	11 (2%)	6	36
All	All	7079/8222 (86%)	5834 (82%)	984 (14%)	261 (4%)	3	28

5 of 261 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	319	GLU
1	B	20	SER
1	B	52	PRO
2	C	21	VAL
2	C	39	ILE

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	
1	A	281/286 (98%)	273 (97%)	8 (3%)	43	65
1	B	189/286 (66%)	186 (98%)	3 (2%)	62	79
1	F	197/286 (69%)	194 (98%)	3 (2%)	65	80
1	G	185/286 (65%)	182 (98%)	3 (2%)	62	79

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	C	1150/1157 (99%)	1094 (95%)	56 (5%)	25	52
2	H	1150/1157 (99%)	1097 (95%)	53 (5%)	27	54
3	D	971/1168 (83%)	921 (95%)	50 (5%)	24	51
3	I	971/1168 (83%)	918 (94%)	53 (6%)	21	49
4	E	74/75 (99%)	72 (97%)	2 (3%)	44	66
4	J	65/75 (87%)	65 (100%)	0	100	100
5	X	460/540 (85%)	447 (97%)	13 (3%)	43	65
5	Y	407/540 (75%)	392 (96%)	15 (4%)	34	59
All	All	6100/7024 (87%)	5841 (96%)	259 (4%)	30	55

5 of 259 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	I	867	GLN
3	I	1148	ARG
3	D	681	LYS
3	D	614	LEU
3	I	1369	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 91 such sidechains are listed below:

Mol	Chain	Res	Type
2	H	510	GLN
2	H	1264	GLN
2	H	517	GLN
2	H	1017	GLN
3	I	300	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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 5 ligands modelled in this entry, 4 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
7	0O2	D	1503	-	32,42,42	2.51	12 (37%)	44,68,68	2.05	12 (27%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	0O2	D	1503	-	-	7/29/49/49	0/3/3/3

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	D	1503	0O2	O6-C6	6.55	1.36	1.23
7	D	1503	0O2	C5-C6	-6.05	1.35	1.47
7	D	1503	0O2	O2'-C2'	-4.42	1.32	1.43
7	D	1503	0O2	C2'-C1'	-4.35	1.47	1.53
7	D	1503	0O2	C2-N2	3.74	1.43	1.34

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	D	1503	0O2	PA-O3A-PB	-5.74	113.14	132.83
7	D	1503	0O2	PC-O3C-PD	-5.41	114.28	132.83
7	D	1503	0O2	C5-C6-N1	3.93	120.89	113.95
7	D	1503	0O2	C2-N1-C6	-3.86	117.99	125.10

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	D	1503	0O2	O6-C6-C5	-3.63	117.29	124.37

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	D	1503	0O2	C3'-O3'-PC-O1C
7	D	1503	0O2	C2'-C3'-O3'-PC
7	D	1503	0O2	C3'-O3'-PC-O3C
7	D	1503	0O2	C3'-O3'-PC-O2C
7	D	1503	0O2	PD-O3C-PC-O2C

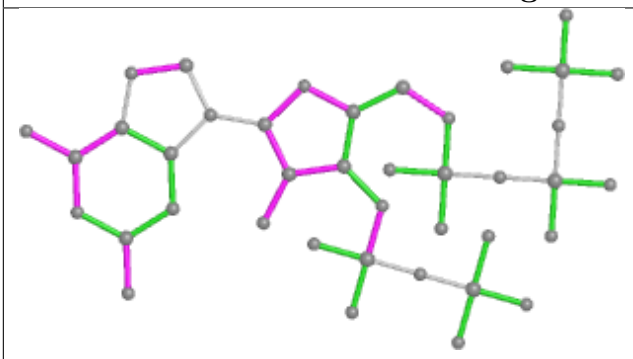
There are no ring outliers.

1 monomer is involved in 9 short contacts:

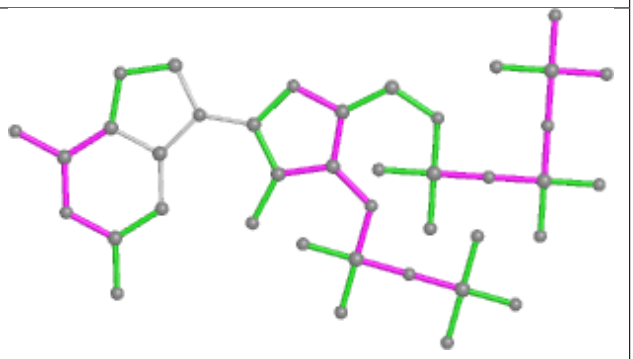
Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	D	1503	0O2	9	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

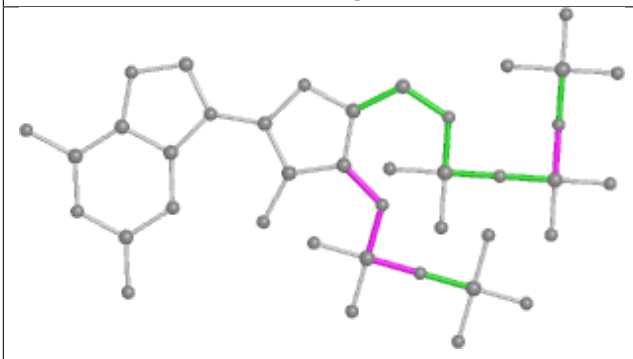
Ligand 0O2 D 1503



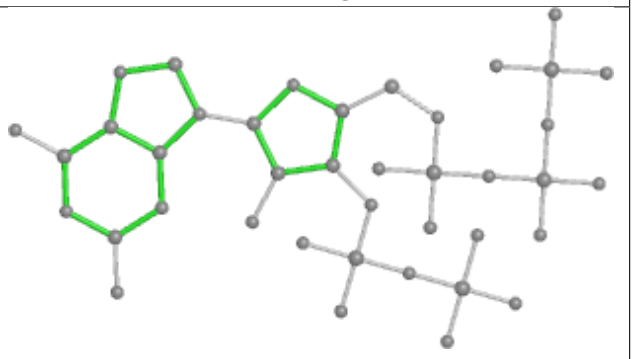
Bond lengths



Bond angles



Torsions



Rings

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	323/329 (98%)	-0.19	4 (1%) 79 70	0, 55, 172, 230	0
1	B	221/329 (67%)	-0.04	7 (3%) 47 37	0, 86, 193, 260	0
1	F	229/329 (69%)	-0.01	13 (5%) 23 20	2, 123, 212, 293	0
1	G	217/329 (65%)	0.04	4 (1%) 68 59	5, 113, 204, 271	0
2	C	1335/1342 (99%)	-0.29	20 (1%) 73 64	0, 38, 168, 304	0
2	H	1335/1342 (99%)	-0.13	39 (2%) 51 41	0, 78, 206, 346	0
3	D	1160/1407 (82%)	-0.21	16 (1%) 75 65	0, 28, 152, 297	0
3	I	1160/1407 (82%)	-0.10	40 (3%) 45 36	0, 54, 183, 316	0
4	E	90/91 (98%)	-0.22	0 100 100	0, 33, 116, 167	0
4	J	76/91 (83%)	0.08	1 (1%) 77 68	12, 83, 181, 230	0
5	X	517/613 (84%)	-0.11	20 (3%) 39 31	0, 98, 238, 341	0
5	Y	458/613 (74%)	-0.09	16 (3%) 44 35	1, 100, 216, 296	0
All	All	7121/8222 (86%)	-0.15	180 (2%) 57 47	0, 63, 198, 346	0

The worst 5 of 180 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	I	10	ALA	9.3
2	H	982	GLY	9.3
3	I	521	LYS	8.7
2	H	981	ALA	8.0
2	H	983	GLY	7.7

6.2 Non-standard residues in protein, DNA, RNA chains

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)

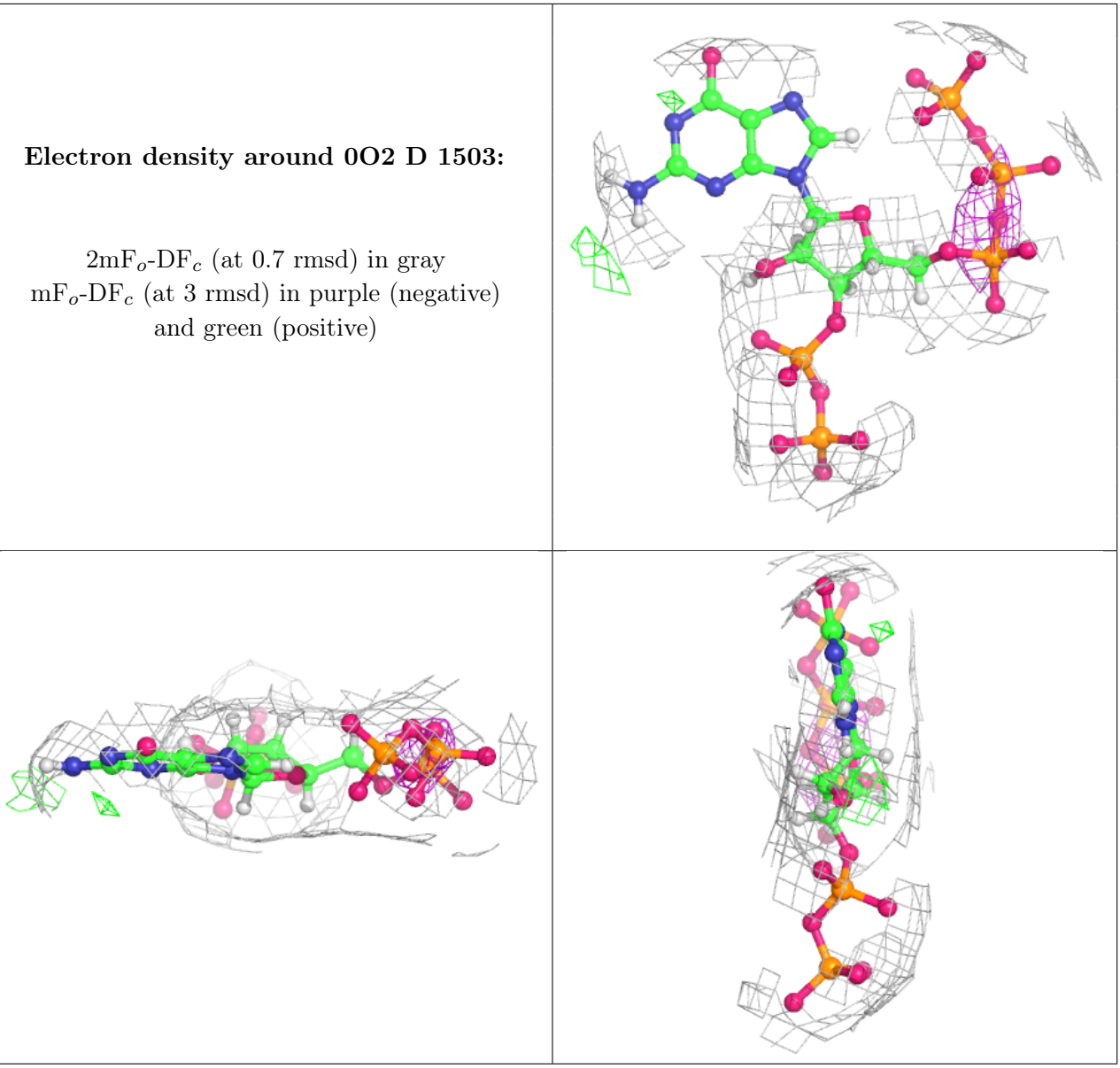
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
7	0O2	D	1503	40/40	0.90	0.17	20,20,20,20	0
6	ZN	I	1501	1/1	0.97	0.04	60,60,60,60	0
6	ZN	D	1502	1/1	0.97	0.18	8,8,8,8	0
6	ZN	D	1501	1/1	0.98	0.06	54,54,54,54	0
6	ZN	I	1502	1/1	0.99	0.17	49,49,49,49	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around 0O2 D 1503:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.