wwPDB X-ray Structure Validation Summary Report (i)

Nov 9, 2024 – 11:11 AM EST

PDB ID	:	4YLP
Title	:	E. coli Transcription Initiation Complex - 16-bp spacer and 5-nt RNA
Authors	:	Zuo, Y.; Steitz, T.A.
Deposited on	:	2015-03-05
Resolution	:	5.50 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:

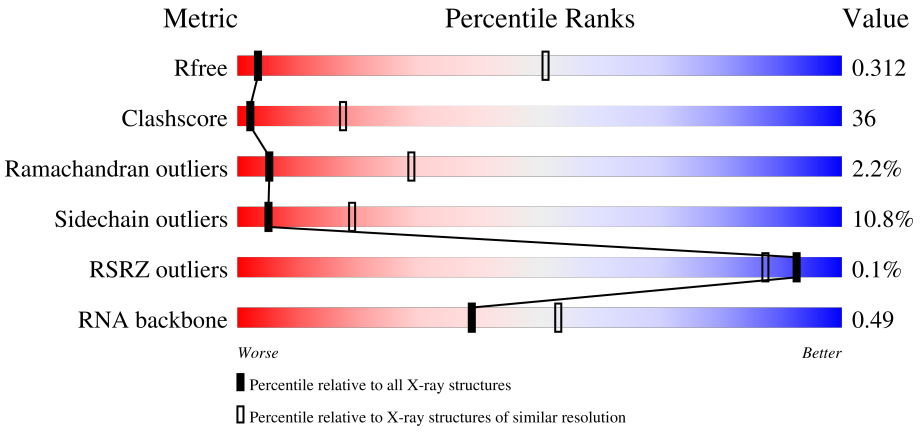
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 5.50 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}	164625	1029 (7.00-4.00)
Clashscore	180529	1069 (7.00-4.00)
Ramachandran outliers	177936	1010 (7.04-3.96)
Sidechain outliers	177891	1004 (7.04-3.94)
RSRZ outliers	164620	1023 (7.00-4.00)
RNA backbone	3690	1172 (7.80-3.00)

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	242	40% 44% 10% • 5%
1	B	242	39% 42% 13% 6%
1	G	242	49% 40% 5% • 5%
1	H	242	51% 36% 7% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	M	242	52% 39% 5%
1	N	242	52% 37% 5% 6%
2	C	1342	47% 48% 5%
2	I	1342	44% 49% 7%
2	O	1342	49% 45% 5%
3	D	1407	45% 44% 7%
3	J	1407	44% 44% 9%
3	P	1407	45% 44% 7%
4	E	90	59% 40%
4	K	90	54% 39% 7%
4	Q	90	57% 40%
5	F	628	41% 32% 6% 21%
5	L	628	45% 29% 21%
5	R	628	42% 30% 7% 21%
6	1	49	31% 69%
6	4	49	35% 63%
6	7	49	37% 63%
7	2	49	39% 61%
7	5	49	35% 65%
7	8	49	41% 59%
8	3	5	80% 20%
8	6	5	40% 40% 20%
8	9	5	40% 60%

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
9	ZN	D	1502	-	-	X	-
9	ZN	P	1501	-	-	X	-
9	ZN	P	1502	-	-	X	-

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 94668 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 protein called DNA-directed RNA polymerase subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	230	Total	C	N	O	S	0	0	0
			1787	1112	317	352	6			
1	B	228	Total	C	N	O	S	0	0	0
			1767	1100	312	349	6			
1	G	230	Total	C	N	O	S	0	0	0
			1787	1112	317	352	6			
1	H	228	Total	C	N	O	S	0	0	0
			1767	1100	312	349	6			
1	M	230	Total	C	N	O	S	0	0	0
			1787	1112	317	352	6			
1	N	228	Total	C	N	O	S	0	0	0
			1767	1100	312	349	6			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	ALA	-	expression tag	UNP A7ZSI4
A	-5	HIS	-	expression tag	UNP A7ZSI4
A	-4	HIS	-	expression tag	UNP A7ZSI4
A	-3	HIS	-	expression tag	UNP A7ZSI4
A	-2	HIS	-	expression tag	UNP A7ZSI4
A	-1	HIS	-	expression tag	UNP A7ZSI4
A	0	HIS	-	expression tag	UNP A7ZSI4
B	-6	ALA	-	expression tag	UNP A7ZSI4
B	-5	HIS	-	expression tag	UNP A7ZSI4
B	-4	HIS	-	expression tag	UNP A7ZSI4
B	-3	HIS	-	expression tag	UNP A7ZSI4
B	-2	HIS	-	expression tag	UNP A7ZSI4
B	-1	HIS	-	expression tag	UNP A7ZSI4
B	0	HIS	-	expression tag	UNP A7ZSI4
G	-6	ALA	-	expression tag	UNP A7ZSI4
G	-5	HIS	-	expression tag	UNP A7ZSI4
G	-4	HIS	-	expression tag	UNP A7ZSI4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	-3	HIS	-	expression tag	UNP A7ZSI4
G	-2	HIS	-	expression tag	UNP A7ZSI4
G	-1	HIS	-	expression tag	UNP A7ZSI4
G	0	HIS	-	expression tag	UNP A7ZSI4
H	-6	ALA	-	expression tag	UNP A7ZSI4
H	-5	HIS	-	expression tag	UNP A7ZSI4
H	-4	HIS	-	expression tag	UNP A7ZSI4
H	-3	HIS	-	expression tag	UNP A7ZSI4
H	-2	HIS	-	expression tag	UNP A7ZSI4
H	-1	HIS	-	expression tag	UNP A7ZSI4
H	0	HIS	-	expression tag	UNP A7ZSI4
M	-6	ALA	-	expression tag	UNP A7ZSI4
M	-5	HIS	-	expression tag	UNP A7ZSI4
M	-4	HIS	-	expression tag	UNP A7ZSI4
M	-3	HIS	-	expression tag	UNP A7ZSI4
M	-2	HIS	-	expression tag	UNP A7ZSI4
M	-1	HIS	-	expression tag	UNP A7ZSI4
M	0	HIS	-	expression tag	UNP A7ZSI4
N	-6	ALA	-	expression tag	UNP A7ZSI4
N	-5	HIS	-	expression tag	UNP A7ZSI4
N	-4	HIS	-	expression tag	UNP A7ZSI4
N	-3	HIS	-	expression tag	UNP A7ZSI4
N	-2	HIS	-	expression tag	UNP A7ZSI4
N	-1	HIS	-	expression tag	UNP A7ZSI4
N	0	HIS	-	expression tag	UNP A7ZSI4

- Molecule 2 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1341	Total	C	N	O	S	0	0	0
			10576	6636	1842	2055	43			
2	I	1341	Total	C	N	O	S	0	0	0
			10576	6636	1842	2055	43			
2	O	1341	Total	C	N	O	S	0	0	0
			10576	6636	1842	2055	43			

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1362	Total	C	N	O	S	0	0	0
			10568	6633	1887	1998	50			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	J	1362	Total	C	N	O	S	0	0	0
			10568	6633	1887	1998	50			
3	P	1362	Total	C	N	O	S	0	0	0
			10568	6633	1887	1998	50			

- Molecule 4 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	90	Total	C	N	O	S	0	0	0
			708	430	136	141	1			
4	K	90	Total	C	N	O	S	0	0	0
			708	430	136	141	1			
4	Q	90	Total	C	N	O	S	0	0	0
			708	430	136	141	1			

- Molecule 5 is a protein called RNA polymerase sigma factor RpoD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	497	Total	C	N	O	S	0	0	0
			4022	2512	719	768	23			
5	L	497	Total	C	N	O	S	0	0	0
			4022	2512	719	768	23			
5	R	497	Total	C	N	O	S	0	0	0
			4022	2512	719	768	23			

There are 45 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-14	MET	-	expression tag	UNP P00579
F	-13	ARG	-	expression tag	UNP P00579
F	-12	GLY	-	expression tag	UNP P00579
F	-11	SER	-	expression tag	UNP P00579
F	-10	HIS	-	expression tag	UNP P00579
F	-9	HIS	-	expression tag	UNP P00579
F	-8	HIS	-	expression tag	UNP P00579
F	-7	HIS	-	expression tag	UNP P00579
F	-6	HIS	-	expression tag	UNP P00579
F	-5	HIS	-	expression tag	UNP P00579
F	-4	THR	-	expression tag	UNP P00579
F	-3	ASP	-	expression tag	UNP P00579
F	-2	GLN	-	expression tag	UNP P00579
F	-1	PHE	-	expression tag	UNP P00579

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	0	THR	-	expression tag	UNP P00579
L	-14	MET	-	expression tag	UNP P00579
L	-13	ARG	-	expression tag	UNP P00579
L	-12	GLY	-	expression tag	UNP P00579
L	-11	SER	-	expression tag	UNP P00579
L	-10	HIS	-	expression tag	UNP P00579
L	-9	HIS	-	expression tag	UNP P00579
L	-8	HIS	-	expression tag	UNP P00579
L	-7	HIS	-	expression tag	UNP P00579
L	-6	HIS	-	expression tag	UNP P00579
L	-5	HIS	-	expression tag	UNP P00579
L	-4	THR	-	expression tag	UNP P00579
L	-3	ASP	-	expression tag	UNP P00579
L	-2	GLN	-	expression tag	UNP P00579
L	-1	PHE	-	expression tag	UNP P00579
L	0	THR	-	expression tag	UNP P00579
R	-14	MET	-	expression tag	UNP P00579
R	-13	ARG	-	expression tag	UNP P00579
R	-12	GLY	-	expression tag	UNP P00579
R	-11	SER	-	expression tag	UNP P00579
R	-10	HIS	-	expression tag	UNP P00579
R	-9	HIS	-	expression tag	UNP P00579
R	-8	HIS	-	expression tag	UNP P00579
R	-7	HIS	-	expression tag	UNP P00579
R	-6	HIS	-	expression tag	UNP P00579
R	-5	HIS	-	expression tag	UNP P00579
R	-4	THR	-	expression tag	UNP P00579
R	-3	ASP	-	expression tag	UNP P00579
R	-2	GLN	-	expression tag	UNP P00579
R	-1	PHE	-	expression tag	UNP P00579
R	0	THR	-	expression tag	UNP P00579

- Molecule 6 is a DNA chain called NT strand DNA (49-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	1	49	Total	C	N	O	P	0	0	0
			996	476	178	294	48			
6	4	49	Total	C	N	O	P	0	0	0
			996	476	178	294	48			
6	7	49	Total	C	N	O	P	0	0	0
			996	476	178	294	48			

- Molecule 7 is a DNA chain called T strand DNA (49-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	2	49	Total	C	N	O	P	0	0	0
			1012	481	191	292	48			
7	5	49	Total	C	N	O	P	0	0	0
			1012	481	191	292	48			
7	8	49	Total	C	N	O	P	0	0	0
			1012	481	191	292	48			

- Molecule 8 is a RNA chain called RNA (5'-R(*(GTP))-R(P*AP*GP*UP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	3	5	Total	C	N	O	P	0	0	0
			117	48	20	42	7			
8	6	5	Total	C	N	O	P	0	0	0
			117	48	20	42	7			
8	9	5	Total	C	N	O	P	0	0	0
			117	48	20	42	7			

- Molecule 9 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	D	2	Total	Zn	0	0
			2	2		
9	J	2	Total	Zn	0	0
			2	2		
9	P	2	Total	Zn	0	0
			2	2		

- Molecule 10 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

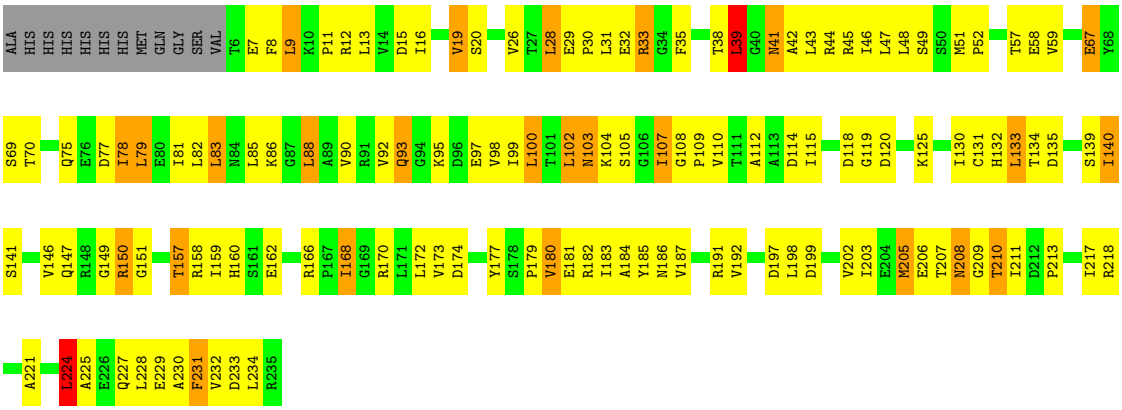
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	D	1	Total	Mg	0	0
			1	1		
10	J	1	Total	Mg	0	0
			1	1		
10	P	1	Total	Mg	0	0
			1	1		

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.

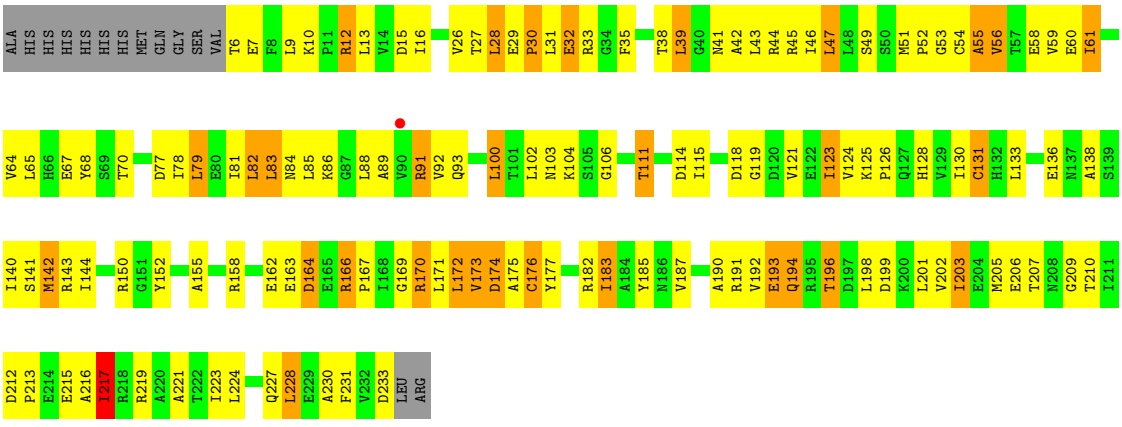
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain A:



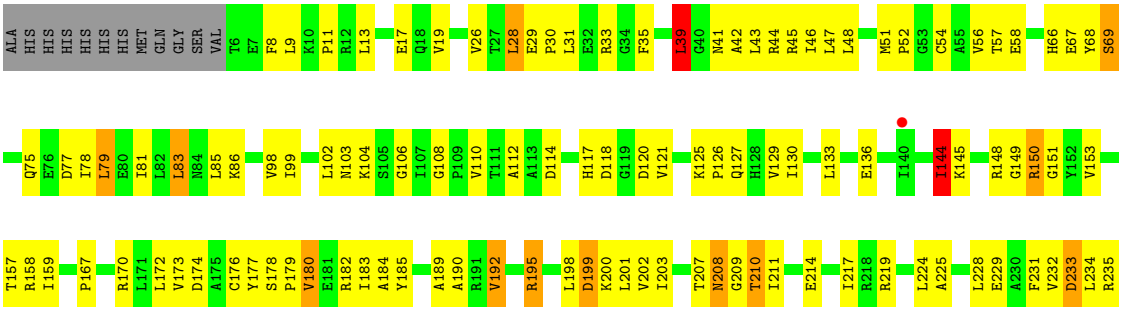
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain B:



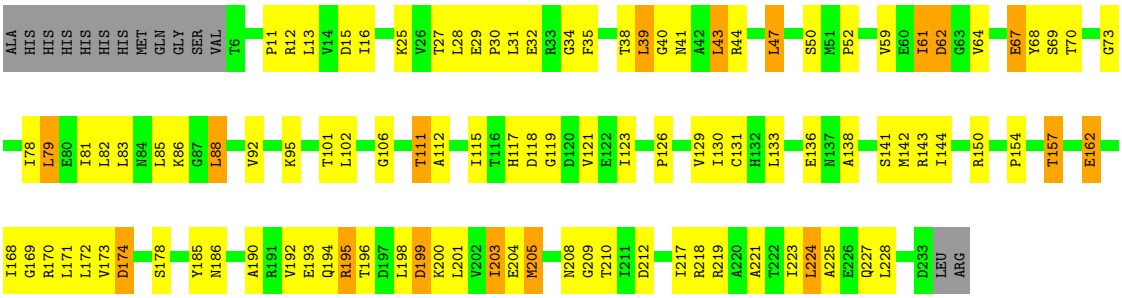
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain G:



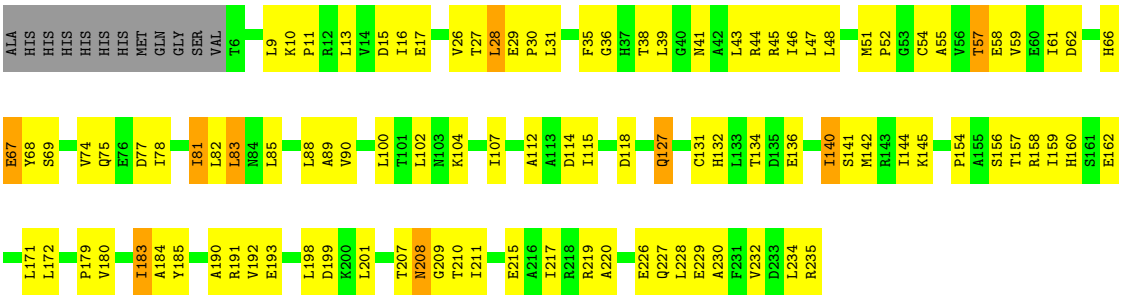
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain H:



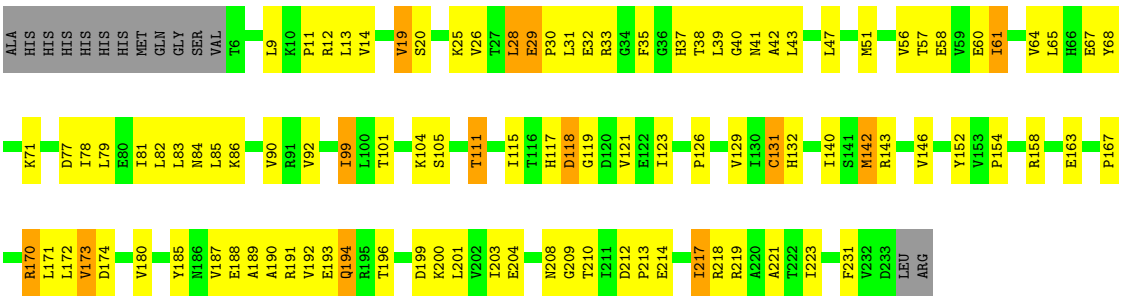
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain M:



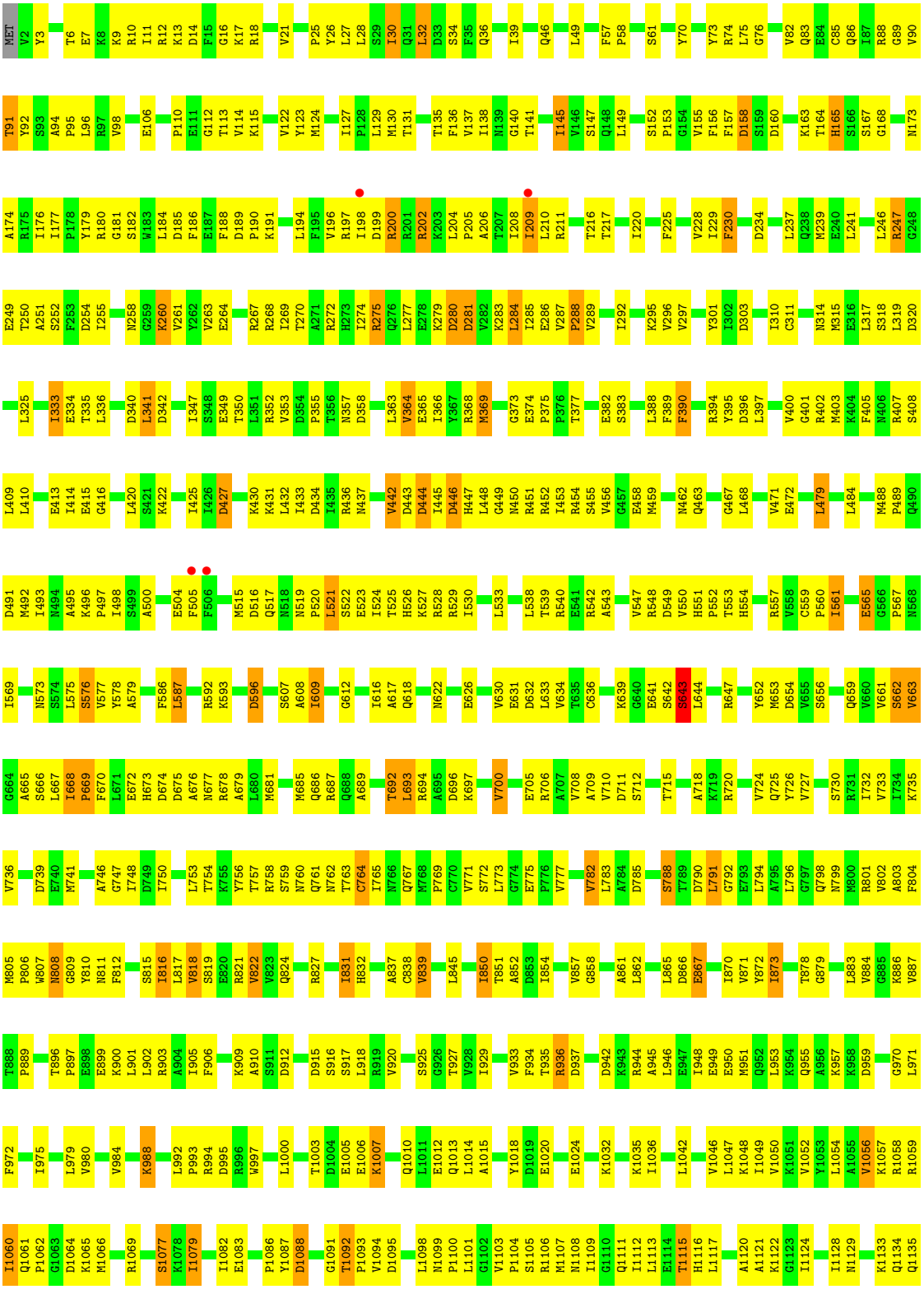
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain N:



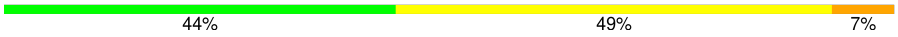
• Molecule 2: DNA-directed RNA polymerase subunit beta

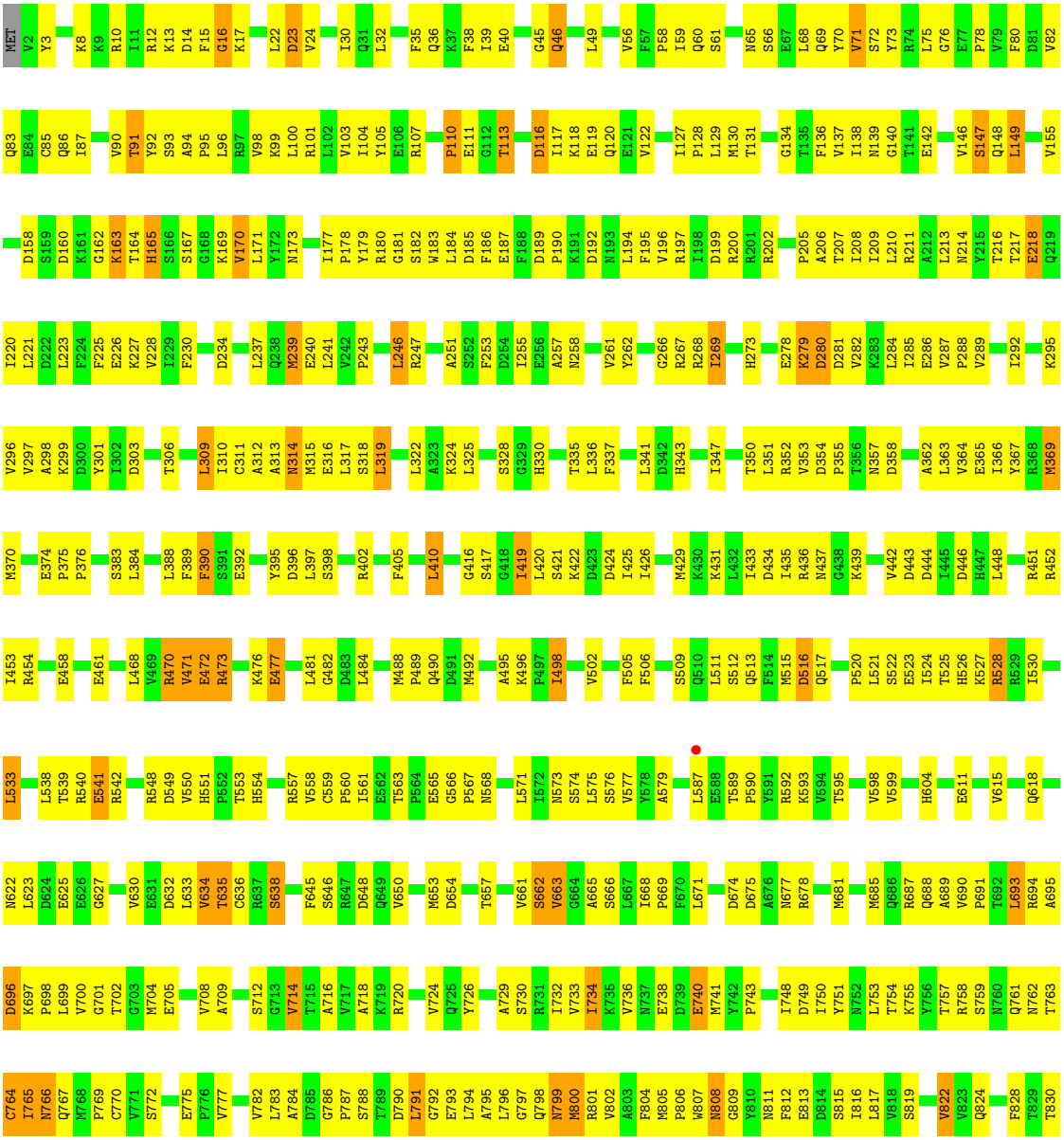
Chain C:

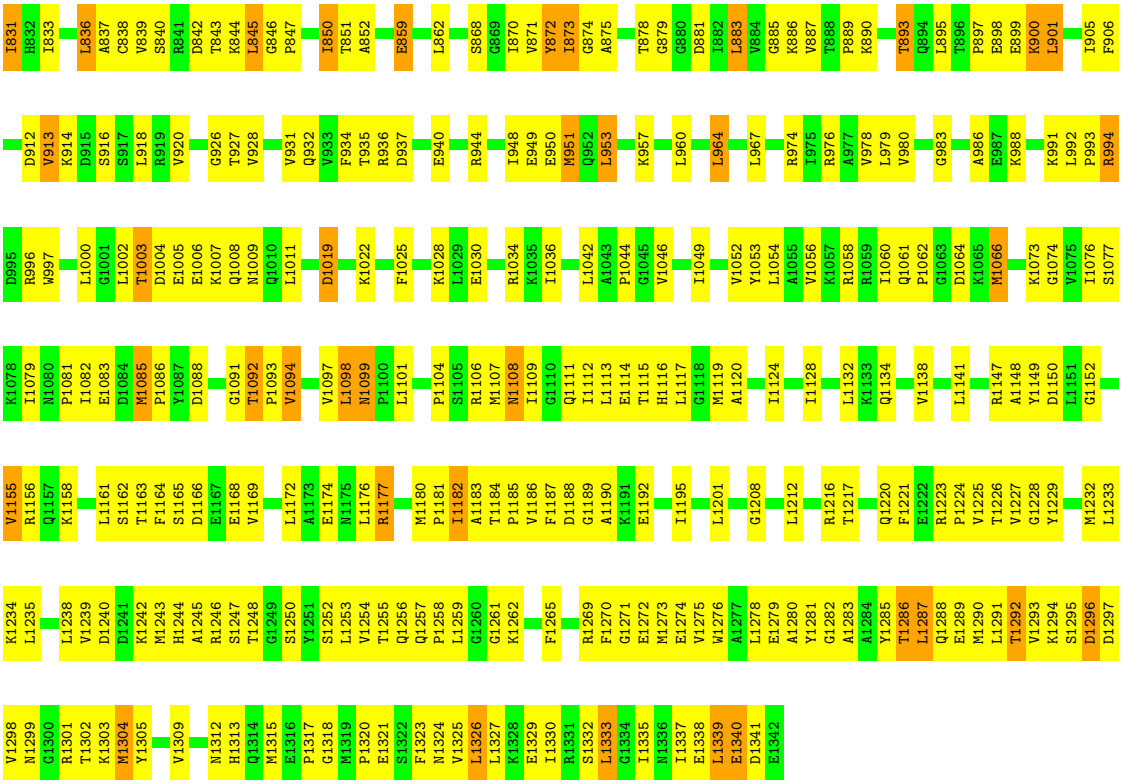


• Molecule 2: DNA-directed RNA polymerase subunit beta

Chain I:

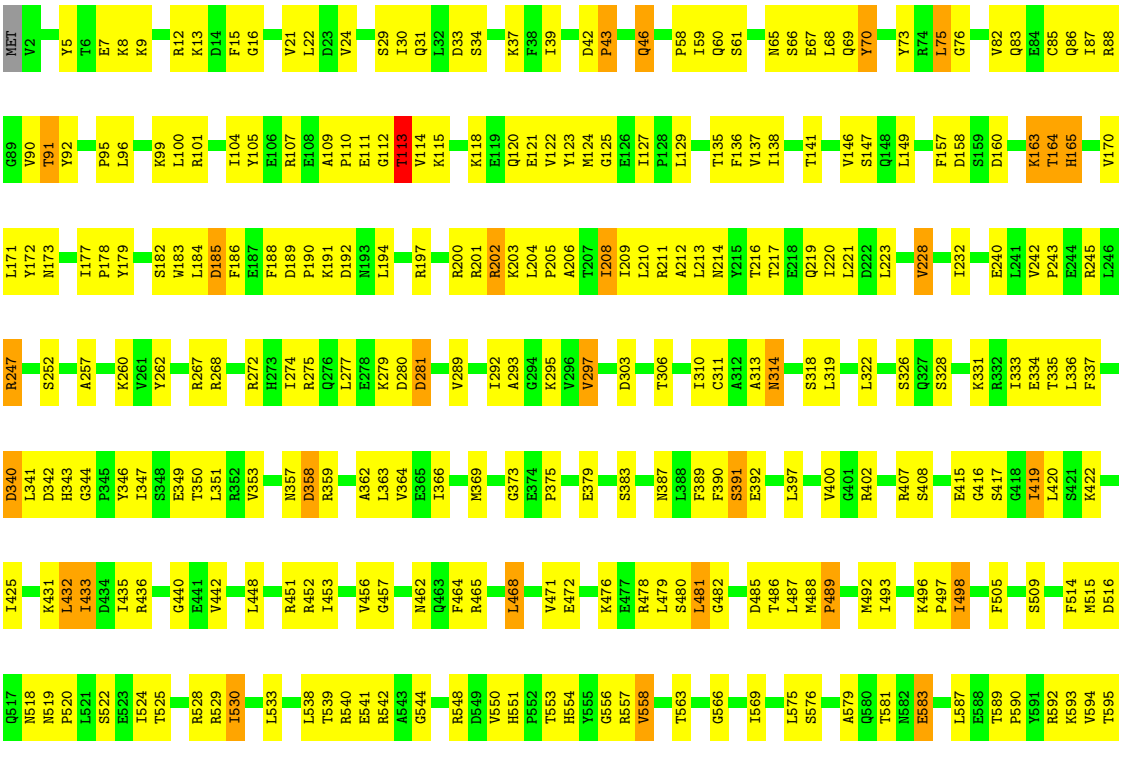


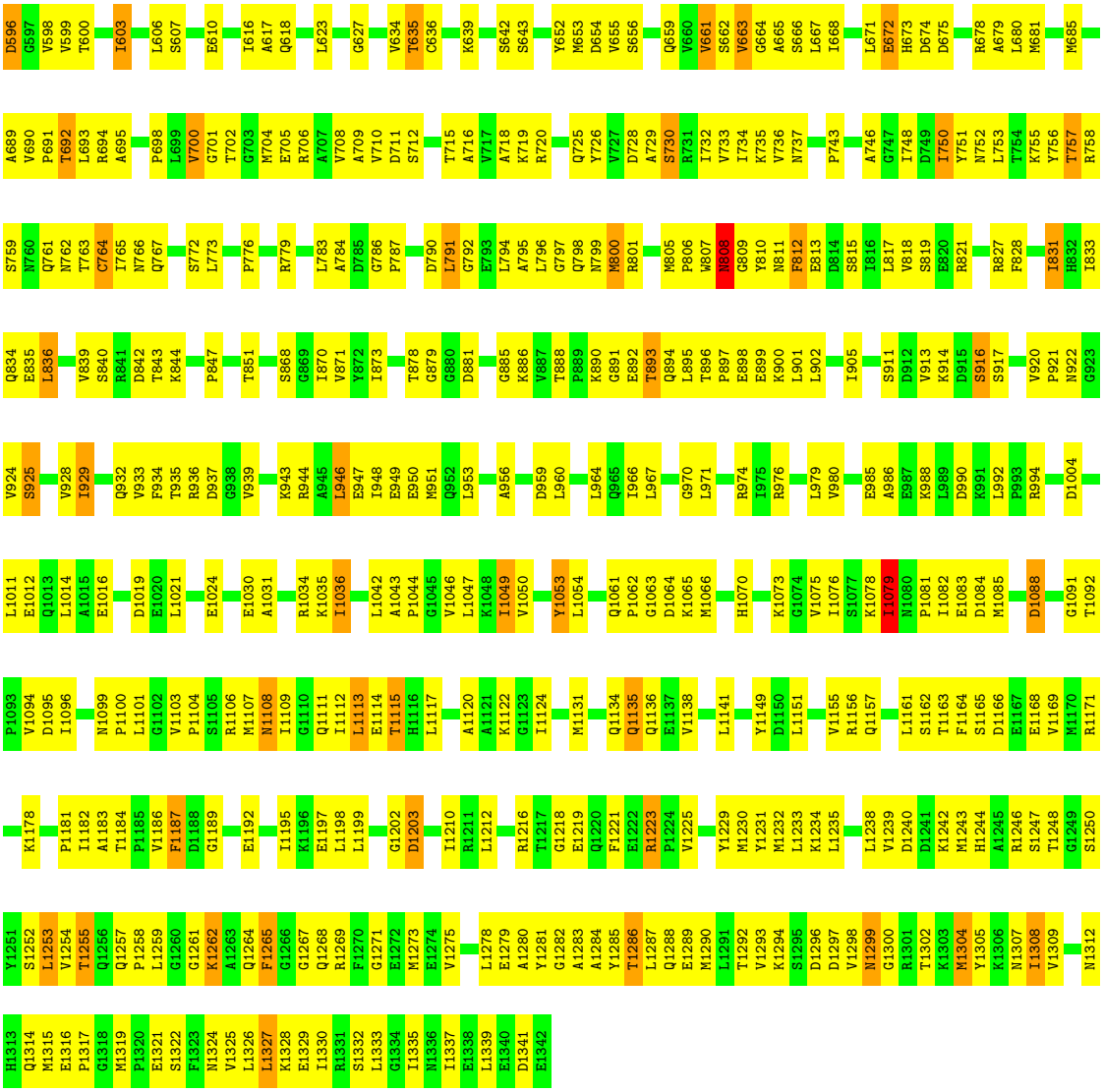




• Molecule 2: DNA-directed RNA polymerase subunit beta

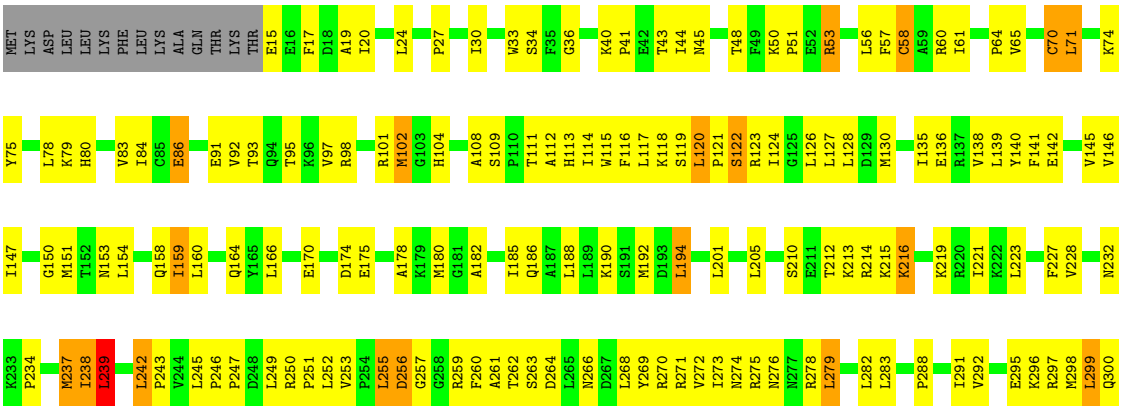
Chain O:

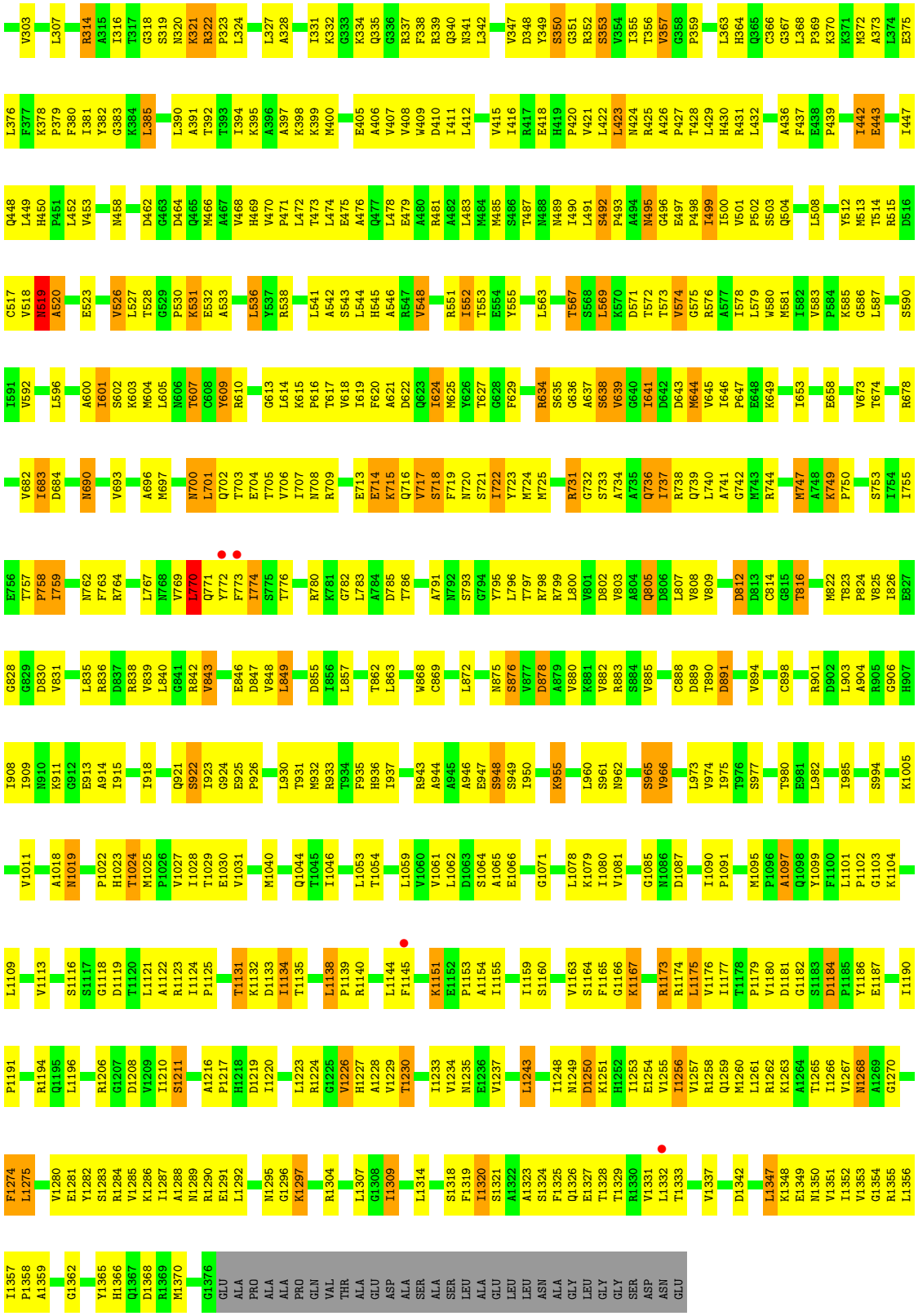




• Molecule 3: DNA-directed RNA polymerase subunit beta'

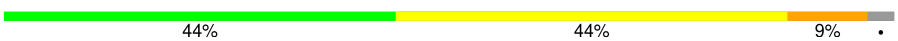
Chain D:

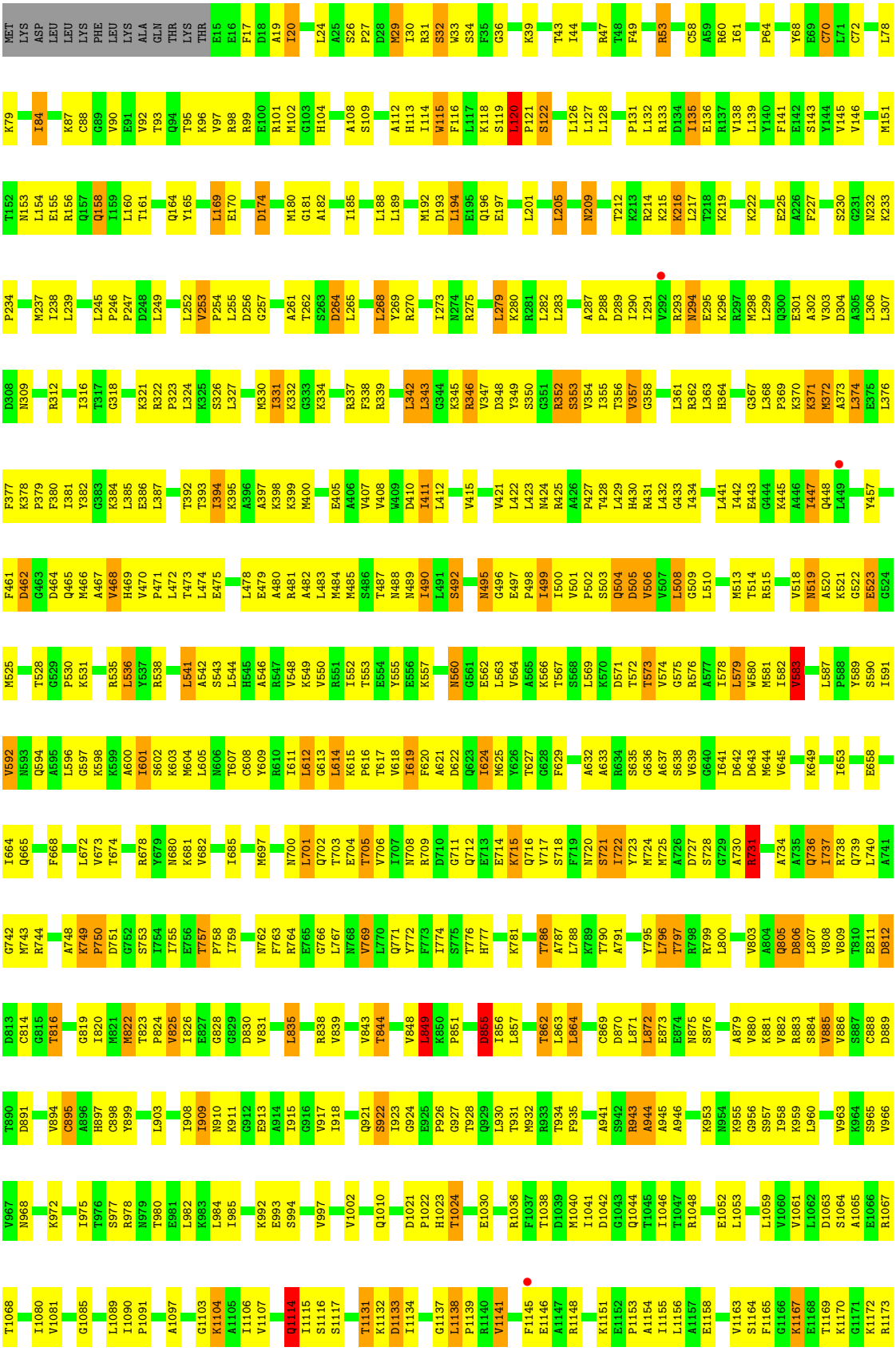


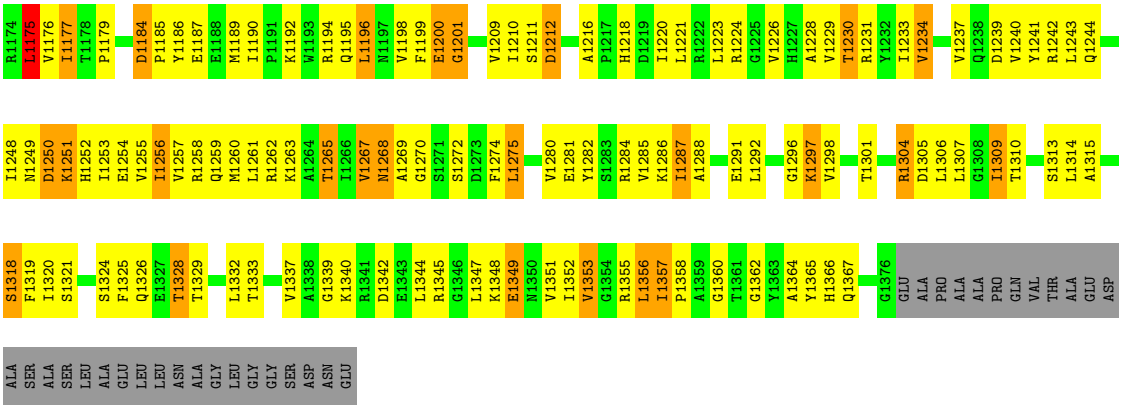


• Molecule 3: DNA-directed RNA polymerase subunit beta'

Chain J:

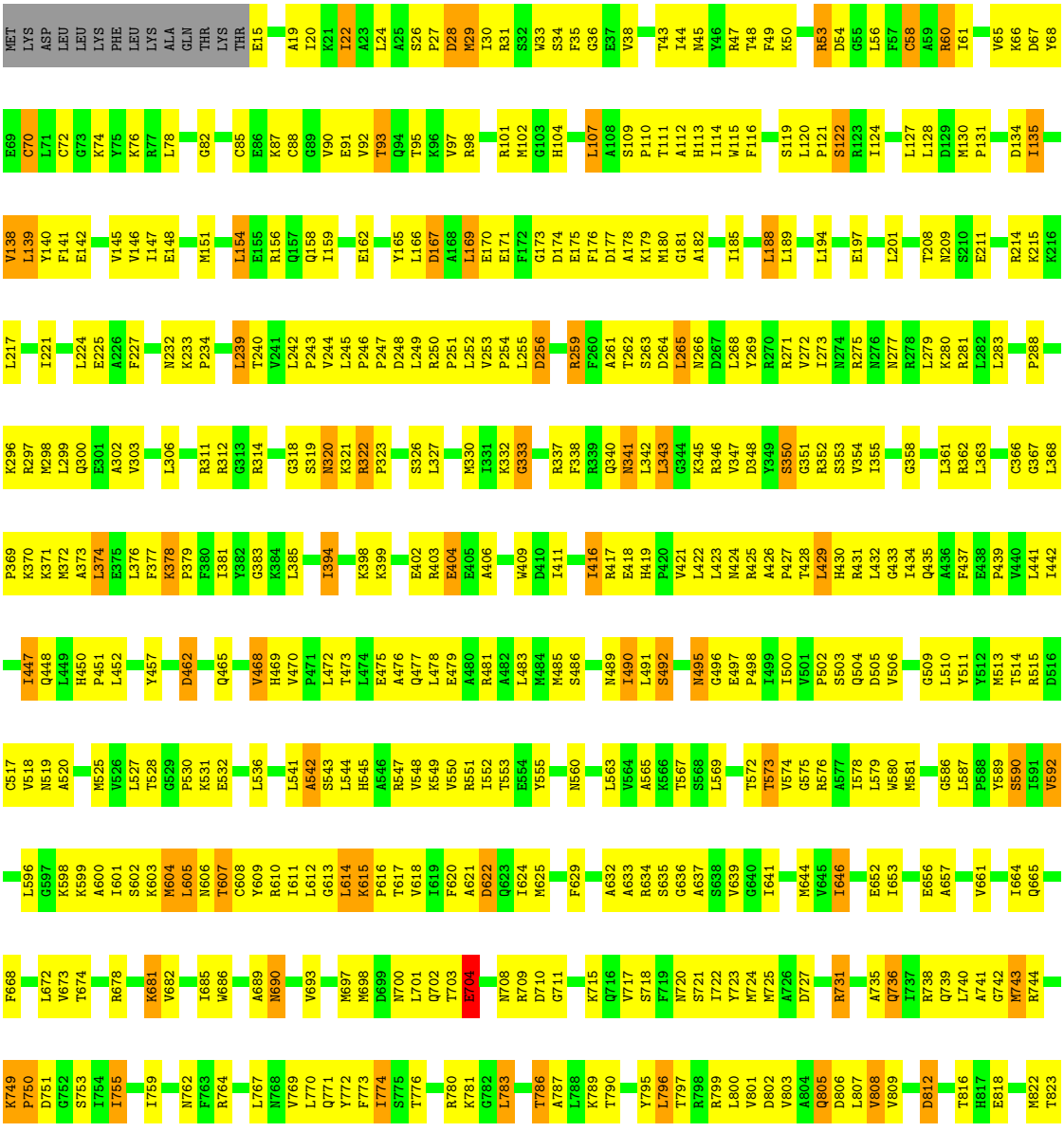


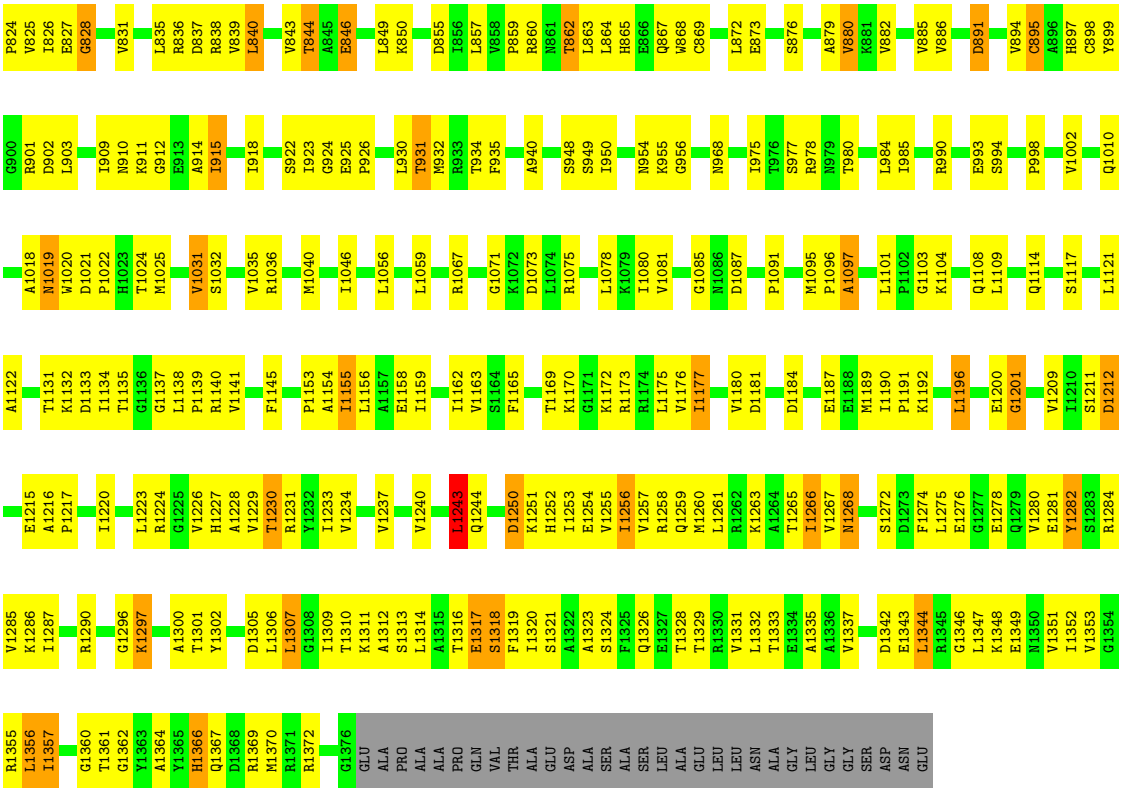

MET LYS ASP LEU LEU LYS PHE LEU LYS ALA GLN THR LYS THR
E15 E16 F17 D18 A19 I20 L24 A25 S26 P27 D28 M29 I30 R31 S32 W33 S34 F35 G36 K39 T43 I44 R47 T48 F49 R53 C58 A59 R60 I61 P64 Y68 E69 C70 L71 C72 L78
K79 I84 K87 C88 G89 V90 E91 V92 T93 Q94 T95 K96 V97 R98 R99 E100 R101 M102 G103 H104 A108 S109 A112 H113 I114 W115 F116 L117 K118 S119 L120 P121 S122 L126 L127 L128 P131 L132 R133 D134 I135 E136 R137 V138 L139 Y140 F141 E142 S143 Y144 V145 V146 M151
T152 N153 L154 E155 R156 Q157 Q158 I159 L160 T161 Q164 Y165 L169 E170 D174 M180 G181 A182 I185 L188 L189 M192 D193 L194 E195 Q196 E197 L201 L205 N209 T212 K213 R214 K215 K216 L217 T218 K219 K222 E225 A226 F227 S230 G231 N232 K233
P234 M237 I238 L239 L245 P246 P247 D248 L249 L252 V253 P264 L255 D256 G257 A261 T262 S263 D264 L265 L268 Y269 R270 I273 N274 R275 L279 K280 R281 L282 L283 A287 P288 D289 I290 I291 V292 R293 N294 E295 K296 R297 M298 L299 Q300 E301 A302 V303 D304 A305 L306 L307
D308 N309 R312 I316 T317 G318 K321 R322 P323 L324 K325 S326 L327 M330 I331 K332 G333 K334 R337 F338 R339 L342 L343 G344 K345 R346 V347 D348 Y349 S350 G351 R352 S353 V354 I355 T356 V357 G358 L361 R362 L363 H364 G367 L368 P369 K370 K371 M372 A373 L374 E375 L376
F377 K378 P379 F380 I381 Y382 G383 K384 L385 E386 L387 T392 T393 T394 K395 A396 A397 K398 K399 M400 E405 A406 V407 V408 W409 D410 I411 L412 V415 V421 L422 L423 N424 R425 A426 P427 T428 L429 H430 R431 L432 G433 I434 L441 I442 E443 G444 K445 A446 I447 Q448 L449 Y457
F461 D462 G463 D464 Q465 M466 A467 V468 H469 V470 P471 L472 T473 L474 E475 L478 E479 A480 R481 A482 L483 M484 M485 S486 T487 N488 N489 I490 L491 S492 N495 G496 E497 P498 I499 I500 V501 P502 S503 Q504 D505 V506 V507 L508 G509 L510 M513 T514 R515 V518 N519 A520 K521 G522 E523 G524
M525 T528 G529 P530 K531 R535 L536 Y537 R538 L541 A542 S543 L544 H545 A546 R547 V548 K549 V550 R551 I552 T553 E554 V555 E556 K557 N560 G561 E562 L563 V564 A565 K566 T567 S568 L569 K570 D571 T572 T573 V574 G575 R576 A577 I578 L579 W580 M581 I582 V583 L587 P588 Y589 S590 I591
V592 N593 Q594 A595 L596 G597 K598 K599 A600 I601 S602 K603 M604 L605 N606 T607 C608 Y609 R610 I611 L612 G613 L614 K615 P616 T617 V618 I619 F620 A621 D622 Q623 I624 M625 Y626 T627 G628 F629 A632 A633 R634 S635 G636 A637 S638 V639 G640 I641 D642 D643 M644 V645 K649 I653 E658
I664 Q665 F668 L672 V673 T674 R678 Y679 N680 K681 V682 I685 M697 N700 L701 Q702 T703 E704 T705 V706 T707 N708 R709 D710 G711 Q712 E713 E714 K715 Q716 V717 S718 F719 N720 S721 I722 Y723 M724 M725 A726 D727 S728 G729 A730 R731 A734 A735 Q736 I737 R738 Q739 L740 A741
G742 M743 R744 A748 K749 P750 D751 G752 S753 I754 I755 E756 T757 P758 I759 N762 F763 R764 E765 G766 L767 N768 V769 L770 Q771 Y772 F773 I774 S775 T776 H777 K781 T786 A787 L788 K789 T790 A791 Y795 L796 T797 R798 R799 L800 V803 A804 Q805 D806 L807 V808 V809 T810 E811 D812
D813 C814 G815 T816 G819 I820 M821 M822 T823 P824 V825 I826 E827 G828 Q829 D830 V831 L835 R838 V839 V843 T844 V848 I849 K850 P851 D855 I856 L857 T862 L863 L864 C869 D870 L871 L872 E873 E874 N875 S876 A879 V880 K881 R883 S884 V885 V886 S887 C888 D889
T890 D891 V894 C895 A896 H897 C898 Y899 L903 I908 I909 N910 K911 G912 E913 A914 I915 G916 V917 I918 Q921 S922 I923 G924 E925 P926 G927 T928 Q929 L930 T931 M932 R933 T934 F935 A941 S942 R943 A944 A945 A946 K953 N954 K955 G956 S957 I958 K959 L960 V963 K964 S965 V966
V967 N968 K972 I975 T976 S977 R978 N979 T980 E981 L982 K983 L984 I985 K992 E993 S994 V997 V1002 Q1010 D1021 P1022 H1023 T1024 E1030 R1036 F1037 T1038 D1039 M1040 I1041 D1042 G1043 Q1044 T1045 I1046 T1047 R1048 E1052 L1053 L1059 V1060 V1061 L1062 D1063 S1064 A1065 E1066 R1067
T1068 I1080 V1081 G1085 L1089 I1090 P1091 A1097 G1103 K1104 A1105 I1106 V1107 Q1114 I1115 S1116 S1117 T1131 K1132 D1133 I1134 G1137 L1138 P1139 R1140 V1141 F1145 E1146 A1147 R1148 K1151 E1152 P1153 A1154 I1155 L1156 A1157 E1158 V1163 S1164 F1165 G1166 K1167 E1168 T1169 K1170 G1171 K1172 R1173



• Molecule 3: DNA-directed RNA polymerase subunit beta'

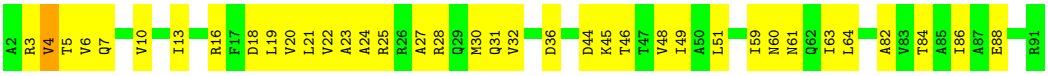
Chain P:





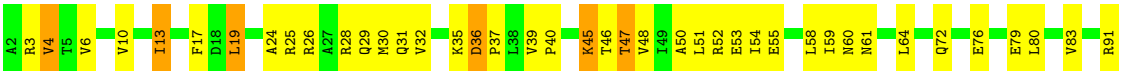
• Molecule 4: DNA-directed RNA polymerase subunit omega

Chain E:



• Molecule 4: DNA-directed RNA polymerase subunit omega

Chain K:

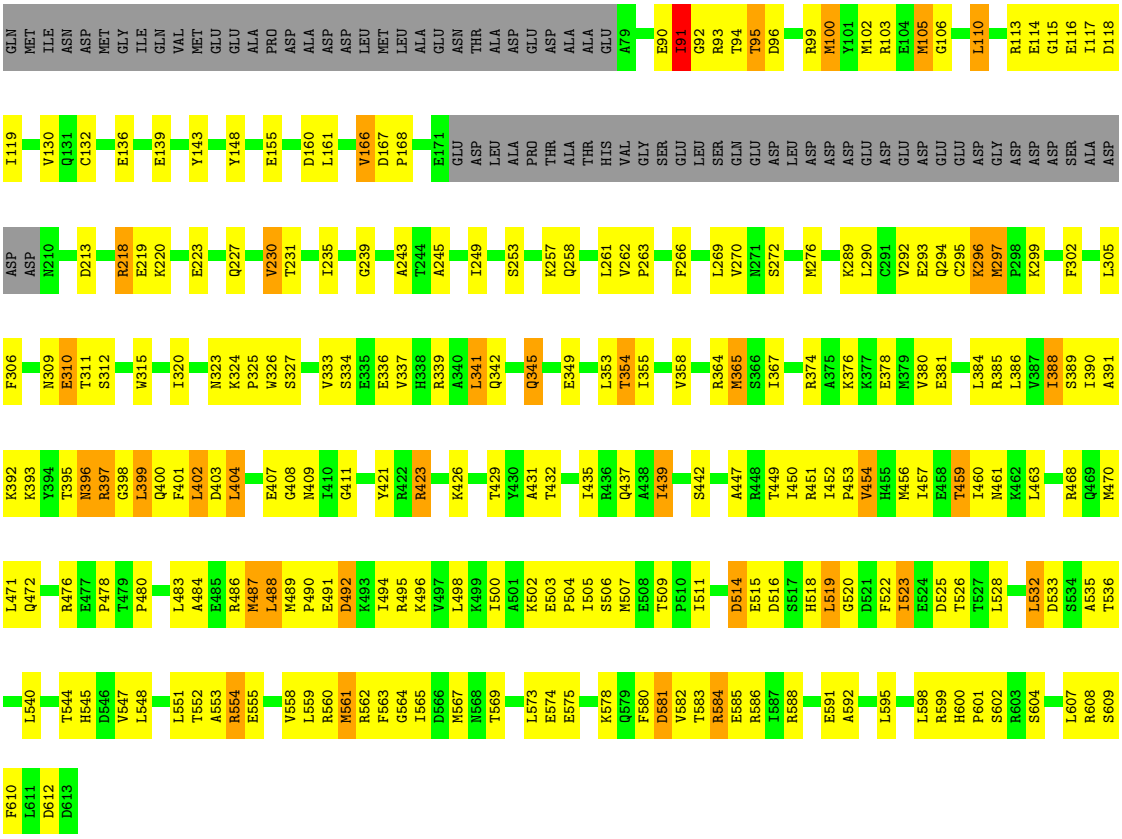


• Molecule 4: DNA-directed RNA polymerase subunit omega

Chain Q:

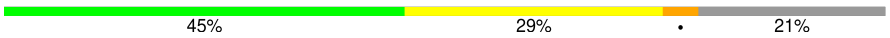
• Molecule 5: RNA polymerase sigma factor RpoD

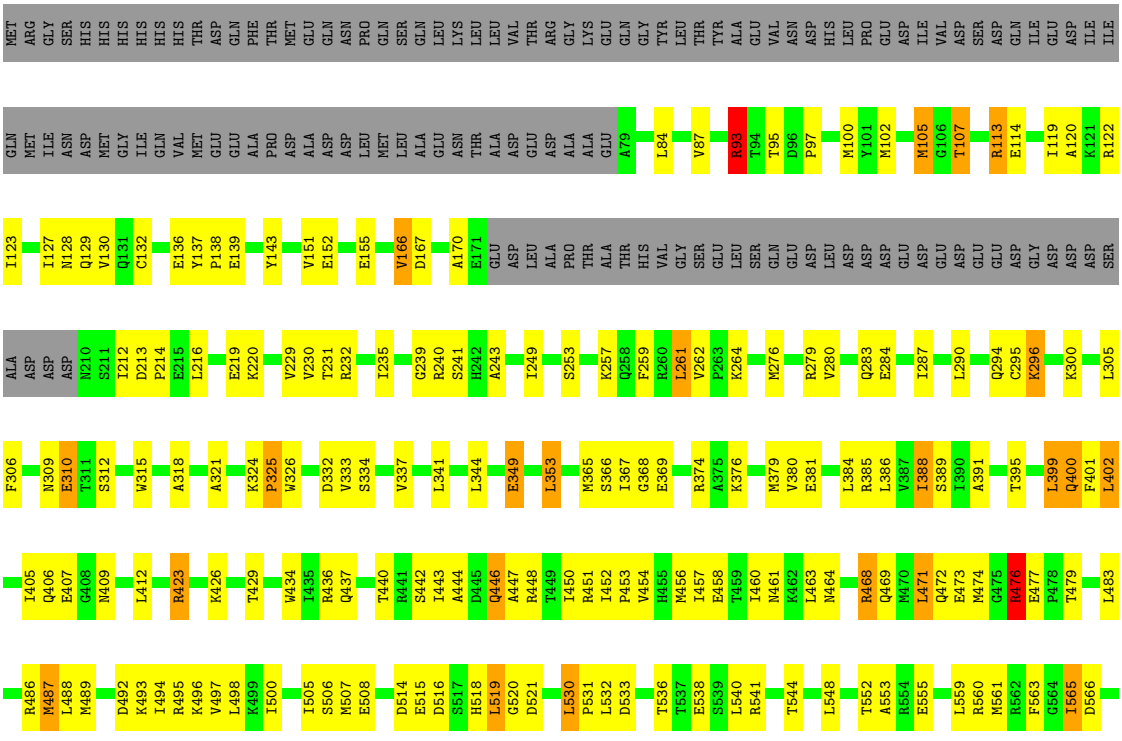
Chain F:

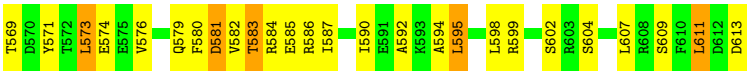


• Molecule 5: RNA polymerase sigma factor RpoD

Chain L:

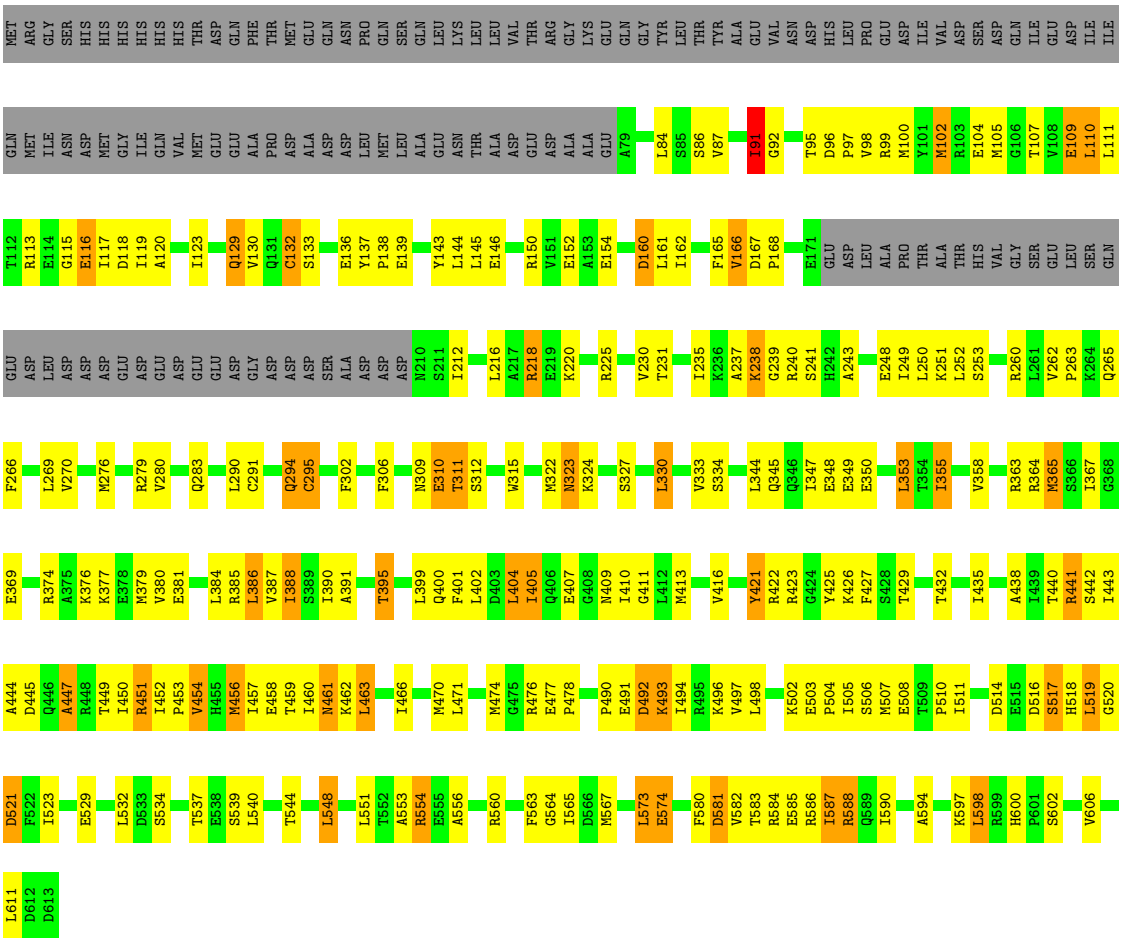






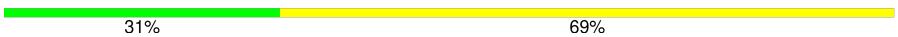
• Molecule 5: RNA polymerase sigma factor RpoD

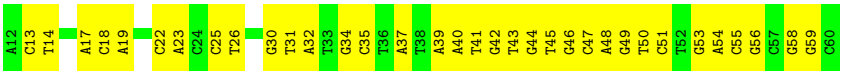
Chain R:



• Molecule 6: NT strand DNA (49-MER)

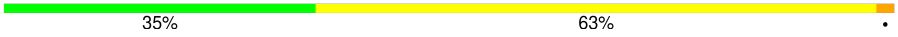
Chain 1:

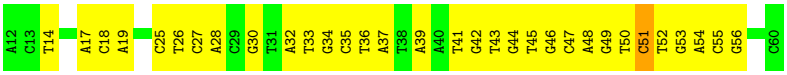




• Molecule 6: NT strand DNA (49-MER)

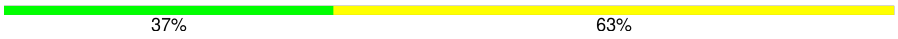
Chain 4:

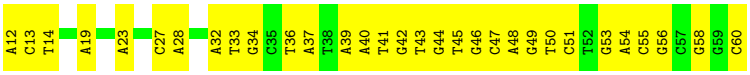




• Molecule 6: NT strand DNA (49-MER)

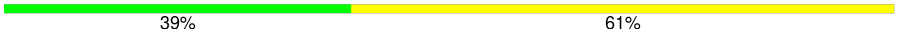
Chain 7:

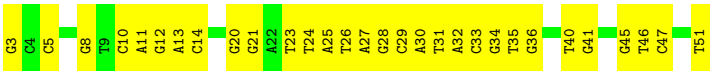




• Molecule 7: T strand DNA (49-MER)

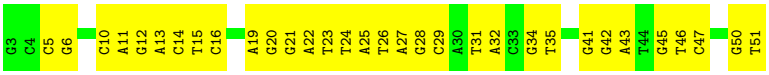
Chain 2:





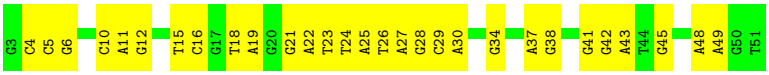
• Molecule 7: T strand DNA (49-MER)

Chain 5:



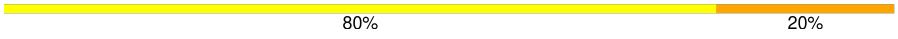
• Molecule 7: T strand DNA (49-MER)

Chain 8:



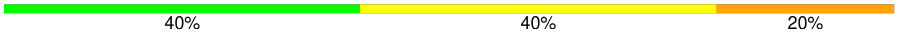
• Molecule 8: RNA (5'-R(*(GTP))-R(P*AP*GP*UP*C)-3')

Chain 3:



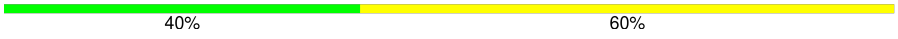
• Molecule 8: RNA (5'-R(*(GTP))-R(P*AP*GP*UP*C)-3')

Chain 6:



• Molecule 8: RNA (5'-R(*(GTP))-R(P*AP*GP*UP*C)-3')

Chain 9:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	237.67Å 204.99Å 248.84Å 90.00° 116.86° 90.00°	Depositor
Resolution (Å)	39.98 – 5.50 39.98 – 5.50	Depositor EDS
% Data completeness (in resolution range)	97.9 (39.98-5.50) 97.9 (39.98-5.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.07 (at 5.37Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.231 , 0.313 0.231 , 0.312	Depositor DCC
R_{free} test set	3384 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	324.1	Xtriage
Anisotropy	0.251	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.24 , 173.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.39$, $< L^2 > = 0.22$	Xtriage
Estimated twinning fraction	0.055 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	94668	wwPDB-VP
Average B, all atoms (Å^2)	198.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.45% 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 (i)

5.1 Standard geometry (i)

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

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	0.63	0/1809	0.91	5/2450 (0.2%)
1	B	0.58	0/1789	0.87	3/2425 (0.1%)
1	G	0.60	0/1809	0.87	2/2450 (0.1%)
1	H	0.59	0/1789	0.87	2/2425 (0.1%)
1	M	0.53	0/1809	0.76	1/2450 (0.0%)
1	N	0.55	0/1789	0.81	1/2425 (0.0%)
2	C	0.56	0/10745	0.78	5/14499 (0.0%)
2	I	0.58	1/10745 (0.0%)	0.78	5/14499 (0.0%)
2	O	0.53	0/10745	0.75	4/14499 (0.0%)
3	D	0.57	1/10729 (0.0%)	0.80	9/14487 (0.1%)
3	J	0.59	1/10729 (0.0%)	0.85	16/14487 (0.1%)
3	P	0.57	1/10729 (0.0%)	0.80	5/14487 (0.0%)
4	E	0.53	0/710	0.71	0/956
4	K	0.62	1/710 (0.1%)	0.82	0/956
4	Q	0.54	0/710	0.77	0/956
5	F	0.51	0/4076	0.73	1/5482 (0.0%)
5	L	0.53	0/4076	0.75	3/5482 (0.1%)
5	R	0.54	1/4076 (0.0%)	0.75	3/5482 (0.1%)
6	1	0.34	0/1114	0.68	0/1714
6	4	1.27	1/1114 (0.1%)	0.91	4/1714 (0.2%)
6	7	0.40	0/1115	0.66	0/1718
7	2	0.35	0/1136	0.67	0/1752
7	5	0.33	0/1136	0.68	0/1752
7	8	0.41	0/1137	0.66	0/1756
8	3	0.38	0/94	0.67	0/144
8	6	0.42	0/94	0.64	0/144
8	9	0.28	0/94	0.68	0/144
All	All	0.57	7/96608 (0.0%)	0.79	69/131735 (0.1%)

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
3	P	0	1

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	4	51	DC	O3'-P	40.58	2.09	1.61
2	I	638	SER	CB-OG	16.07	1.63	1.42
3	D	955	LYS	CE-NZ	10.97	1.76	1.49
4	K	91	ARG	C-O	7.42	1.37	1.23
3	P	681	LYS	CG-CD	5.15	1.70	1.52

The worst 5 of 69 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	4	51	DC	OP1-P-O3'	15.55	139.42	105.20
6	4	51	DC	P-O3'-C3'	15.39	138.17	119.70
6	4	51	DC	O3'-P-O5'	-10.32	84.38	104.00
3	J	120	LEU	C-N-CD	-9.82	99.00	120.60
1	N	29	GLU	C-N-CD	-9.03	100.74	120.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	P	1276	GLU	Peptide

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	1787	0	1813	209	0
1	B	1767	0	1789	217	0
1	G	1787	0	1813	166	0
1	H	1767	0	1789	160	0
1	M	1787	0	1813	134	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	N	1767	0	1789	116	0
2	C	10576	0	10591	815	0
2	I	10576	0	10591	916	0
2	O	10576	0	10591	739	0
3	D	10568	0	10781	927	1
3	J	10568	0	10780	1017	0
3	P	10568	0	10783	901	0
4	E	708	0	719	39	0
4	K	708	0	719	38	0
4	Q	708	0	719	47	0
5	F	4022	0	4083	280	0
5	L	4022	0	4083	220	0
5	R	4022	0	4083	298	0
6	1	996	0	555	65	1
6	4	996	0	556	71	0
6	7	996	0	554	60	1
7	2	1012	0	554	55	1
7	5	1012	0	554	53	0
7	8	1012	0	553	48	0
8	3	117	0	55	10	0
8	6	117	0	55	6	0
8	9	117	0	55	6	0
9	D	2	0	0	2	0
9	J	2	0	0	1	0
9	P	2	0	0	5	0
10	D	1	0	0	0	0
10	J	1	0	0	0	0
10	P	1	0	0	0	0
All	All	94668	0	92820	6810	3

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 36.

The worst 5 of 6810 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)
2:I:255:ILE:CG1	2:I:255:ILE:CD1	1.74	1.59
3:D:955:LYS:NZ	3:D:955:LYS:CE	1.76	1.48
3:P:514:THR:HG21	3:P:596:LEU:CD1	1.48	1.42
3:J:421:VAL:CG1	3:J:469:HIS:O	1.70	1.40
3:P:1095:MET:SD	3:P:1173:ARG:NH2	1.97	1.38

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:2:3:DG:O5'	7:2:51:DT:O3'[2_657]	1.64	0.56
3:D:1174:ARG:NH2	6:1:17:DA:OP1[2_657]	2.10	0.10
6:7:12:DA:O5'	6:7:60:DC:O3'[2_546]	2.13	0.07

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	228/242 (94%)	213 (93%)	11 (5%)	4 (2%)	7 34
1	B	226/242 (93%)	204 (90%)	14 (6%)	8 (4%)	3 20
1	G	228/242 (94%)	211 (92%)	14 (6%)	3 (1%)	10 42
1	H	226/242 (93%)	205 (91%)	17 (8%)	4 (2%)	7 34
1	M	228/242 (94%)	215 (94%)	12 (5%)	1 (0%)	30 68
1	N	226/242 (93%)	208 (92%)	12 (5%)	6 (3%)	4 25
2	C	1339/1342 (100%)	1220 (91%)	97 (7%)	22 (2%)	8 37
2	I	1339/1342 (100%)	1226 (92%)	88 (7%)	25 (2%)	6 32
2	O	1339/1342 (100%)	1235 (92%)	82 (6%)	22 (2%)	8 37
3	D	1360/1407 (97%)	1212 (89%)	120 (9%)	28 (2%)	5 29
3	J	1360/1407 (97%)	1212 (89%)	113 (8%)	35 (3%)	4 25
3	P	1360/1407 (97%)	1214 (89%)	111 (8%)	35 (3%)	4 25
4	E	88/90 (98%)	84 (96%)	4 (4%)	0	100 100
4	K	88/90 (98%)	84 (96%)	4 (4%)	0	100 100
4	Q	88/90 (98%)	84 (96%)	4 (4%)	0	100 100
5	F	493/628 (78%)	449 (91%)	30 (6%)	14 (3%)	4 24
5	L	493/628 (78%)	444 (90%)	30 (6%)	19 (4%)	2 19

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
5	R	493/628 (78%)	447 (91%)	30 (6%)	16 (3%)	3 21
All	All	11202/11853 (94%)	10167 (91%)	793 (7%)	242 (2%)	5 29

5 of 242 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	210	THR
1	B	209	GLY
2	C	165	HIS
2	C	808	ASN
2	C	812	PHE

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	198/208 (95%)	166 (84%)	32 (16%)	2 10
1	B	196/208 (94%)	163 (83%)	33 (17%)	1 9
1	G	198/208 (95%)	180 (91%)	18 (9%)	7 24
1	H	196/208 (94%)	171 (87%)	25 (13%)	3 14
1	M	198/208 (95%)	183 (92%)	15 (8%)	11 30
1	N	196/208 (94%)	179 (91%)	17 (9%)	8 25
2	C	1156/1157 (100%)	1027 (89%)	129 (11%)	5 17
2	I	1156/1157 (100%)	1038 (90%)	118 (10%)	6 20
2	O	1156/1157 (100%)	1044 (90%)	112 (10%)	6 22
3	D	1135/1168 (97%)	1009 (89%)	126 (11%)	5 18
3	J	1135/1168 (97%)	1003 (88%)	132 (12%)	4 17
3	P	1135/1168 (97%)	1014 (89%)	121 (11%)	5 19
4	E	74/74 (100%)	71 (96%)	3 (4%)	26 47
4	K	74/74 (100%)	65 (88%)	9 (12%)	4 16
4	Q	74/74 (100%)	68 (92%)	6 (8%)	9 28

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
5	F	439/554 (79%)	395 (90%)	44 (10%)	6 21
5	L	439/554 (79%)	401 (91%)	38 (9%)	8 25
5	R	439/554 (79%)	384 (88%)	55 (12%)	3 15
All	All	9594/10107 (95%)	8561 (89%)	1033 (11%)	5 19

5 of 1033 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	P	781	LYS
3	P	1159	ILE
3	P	774	ILE
1	H	43	LEU
1	G	173	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 143 such sidechains are listed below:

Mol	Chain	Res	Type
3	P	277	ASN
3	P	419	HIS
3	P	936	HIS
5	F	589	GLN
5	F	545	HIS

5.3.3 RNA

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
8	3	4/5 (80%)	0	1 (25%)
8	6	4/5 (80%)	0	1 (25%)
8	9	3/5 (60%)	0	0
All	All	11/15 (73%)	0	2 (18%)

There are no RNA backbone outliers to report.

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
8	3	13	GTP
8	6	13	GTP

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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 9 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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
6	4	2
6	1	1
7	2	1
7	5	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	4	46:DG	O3'	47:DC	P	5.33
1	1	46:DG	O3'	47:DC	P	4.95
1	2	12:DG	O3'	13:DA	P	2.74
1	5	11:DA	O3'	12:DG	P	2.33

Continued on next page...

Continued from previous page...

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	4	51:DC	O3'	52:DT	P	2.09

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	230/242 (95%)	-0.85	0 100 100	134, 152, 183, 205	0
1	B	228/242 (94%)	-0.83	1 (0%) 89 79	136, 167, 199, 236	0
1	G	230/242 (95%)	-0.81	1 (0%) 89 79	139, 162, 198, 240	0
1	H	228/242 (94%)	-0.75	0 100 100	141, 176, 208, 242	0
1	M	230/242 (95%)	-0.80	0 100 100	159, 179, 209, 245	0
1	N	228/242 (94%)	-0.75	0 100 100	169, 201, 249, 272	0
2	C	1341/1342 (99%)	-0.81	6 (0%) 89 79	107, 166, 250, 351	0
2	I	1341/1342 (99%)	-0.86	1 (0%) 92 87	98, 172, 227, 283	0
2	O	1341/1342 (99%)	-0.91	0 100 100	113, 174, 222, 263	0
3	D	1362/1407 (96%)	-0.84	4 (0%) 90 81	112, 184, 269, 324	0
3	J	1362/1407 (96%)	-0.80	3 (0%) 92 84	100, 172, 322, 386	0
3	P	1362/1407 (96%)	-0.86	0 100 100	117, 182, 291, 333	0
4	E	90/90 (100%)	-0.86	0 100 100	136, 169, 350, 413	0
4	K	90/90 (100%)	-0.80	0 100 100	112, 152, 324, 363	0
4	Q	90/90 (100%)	-0.92	0 100 100	128, 171, 328, 364	0
5	F	497/628 (79%)	-0.83	0 100 100	154, 271, 387, 434	0
5	L	497/628 (79%)	-0.76	0 100 100	138, 281, 365, 402	0
5	R	497/628 (79%)	-0.83	0 100 100	146, 261, 390, 426	0
6	1	49/49 (100%)	-0.52	0 100 100	205, 265, 288, 289	0
6	4	49/49 (100%)	-0.59	0 100 100	181, 228, 278, 302	0
6	7	49/49 (100%)	-0.51	0 100 100	184, 228, 266, 277	0
7	2	49/49 (100%)	-0.46	0 100 100	192, 268, 291, 312	0
7	5	49/49 (100%)	-0.47	0 100 100	163, 232, 279, 326	0
7	8	49/49 (100%)	-0.52	0 100 100	166, 227, 262, 322	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
8	3	4/5 (80%)	-1.19	0	100	100	230, 234, 236, 245	0
8	6	4/5 (80%)	-1.21	0	100	100	220, 221, 224, 239	0
8	9	4/5 (80%)	-1.02	0	100	100	215, 221, 224, 236	0
All	All	11550/12162 (94%)	-0.83	16 (0%)	92	87	98, 182, 331, 434	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	772	TYR	3.7
3	J	449	LEU	3.6
2	C	1278	LEU	3.6
2	C	506	PHE	3.5
2	C	1279	GLU	2.9

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, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
10	MG	P	1503	1/1	0.96	0.08	170,170,170,170	0
9	ZN	J	1501	1/1	0.99	0.02	211,211,211,211	0
9	ZN	P	1501	1/1	0.99	0.02	206,206,206,206	0
9	ZN	D	1501	1/1	0.99	0.02	220,220,220,220	0
9	ZN	D	1502	1/1	1.00	0.03	181,181,181,181	0
9	ZN	P	1502	1/1	1.00	0.02	158,158,158,158	0
10	MG	D	1503	1/1	1.00	0.01	141,141,141,141	0
10	MG	J	1503	1/1	1.00	0.02	145,145,145,145	0
9	ZN	J	1502	1/1	1.00	0.04	144,144,144,144	0

6.5 Other polymers (i)

There are no such residues in this entry.