wwPDB X-ray Structure Validation Summary Report (i)

May 29, 2020 – 08:11 am BST

PDB ID	:	4YLO
Title	:	E. coli Transcription Initiation Complex - 16-bp spacer and 4-nt RNA
Authors	:	Zuo, Y.; Steitz, T.A.
Deposited on	:	2015-03-05
Resolution	:	6.00 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 following versions of software and data (see references (1)) were used in the production of this report:

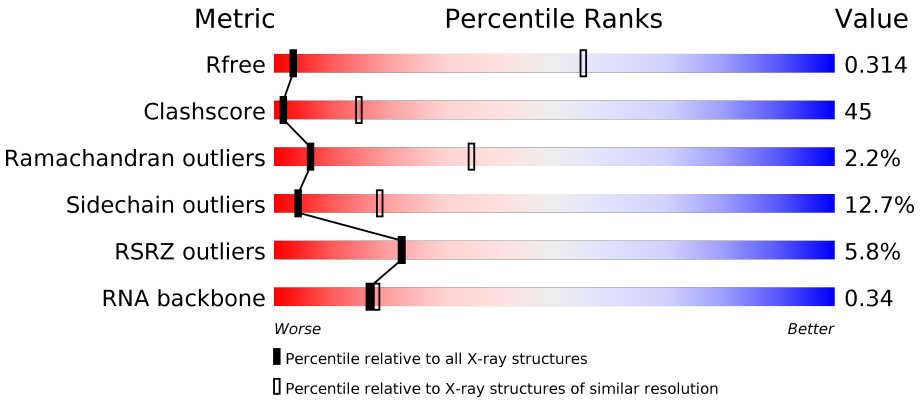
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 6.00 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	1000 (8.00-3.88)
Clashscore	141614	1049 (8.00-3.90)
Ramachandran outliers	138981	1016 (8.00-3.86)
Sidechain outliers	138945	1017 (8.00-3.82)
RSRZ outliers	127900	1015 (8.20-3.78)
RNA backbone	3102	1076 (8.70-3.00)

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 on 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	47% 40% 8% 5%
1	B	242	2% 41% 47% 6% 6%
1	G	242	7% 41% 48% 7% 5%
1	H	242	7% 45% 44% 5% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	M	242	5% 40% 48% 7% 5%
1	N	242	2% 40% 47% 7% 6%
2	C	1342	3% 45% 48% 6%
2	I	1342	2% 43% 48% 8%
2	O	1342	3% 43% 49% 8%
3	D	1407	4% 36% 51% 10%
3	J	1407	8% 40% 48% 9%
3	P	1407	7% 40% 48% 8%
4	E	90	9% 52% 42%
4	K	90	7% 56% 41%
4	Q	90	10% 52% 43%
5	F	628	13% 38% 35% 5% 21%
5	L	628	9% 37% 36% 6% 21%
5	R	628	9% 40% 33% 6% 21%
6	1	49	2% 24% 76%
6	4	49	4% 20% 78%
6	7	49	8% 31% 69%
7	2	49	4% 37% 63%
7	5	49	33% 67%
7	8	49	35% 65%
8	3	4	50% 50%
8	6	4	50% 50%
8	9	4	50% 50%

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
10	ZN	D	1502	-	-	X	-

2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 94608 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'-D(*(GTP))-R(P*AP*GP*U)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	3	4	Total	C	N	O	P	0	0	0
			97	39	17	35	6			
8	6	4	Total	C	N	O	P	0	0	0
			97	39	17	35	6			
8	9	4	Total	C	N	O	P	0	0	0
			97	39	17	35	6			

- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	P	1	Total	Mg	0	0
			1	1		
9	J	1	Total	Mg	0	0
			1	1		
9	C	1	Total	Mg	0	0
			1	1		

- Molecule 10 is ZINC ION (three-letter code: ZN) (formula: Zn).

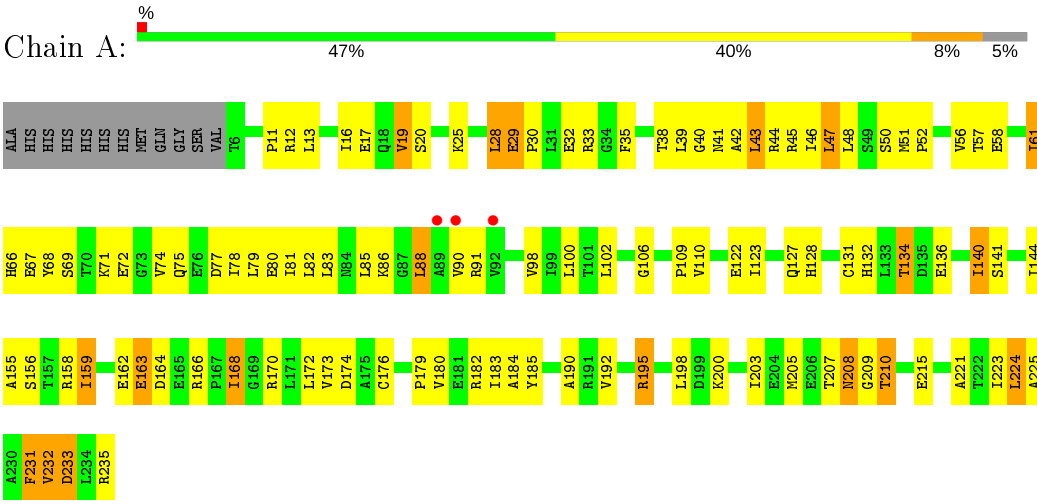
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	P	2	Total	Zn	0	0
			2	2		
10	J	2	Total	Zn	0	0
			2	2		
10	D	2	Total	Zn	0	0
			2	2		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

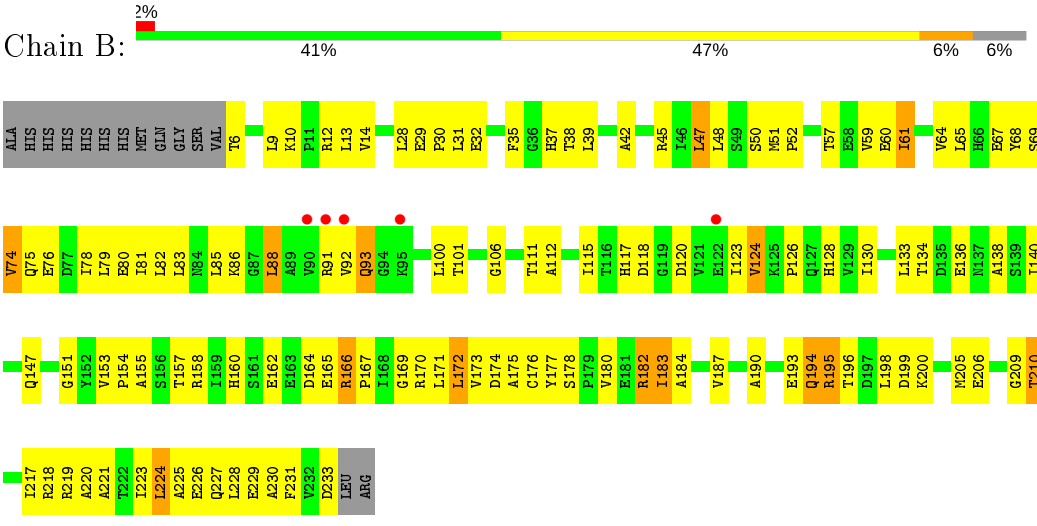
- Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain A:



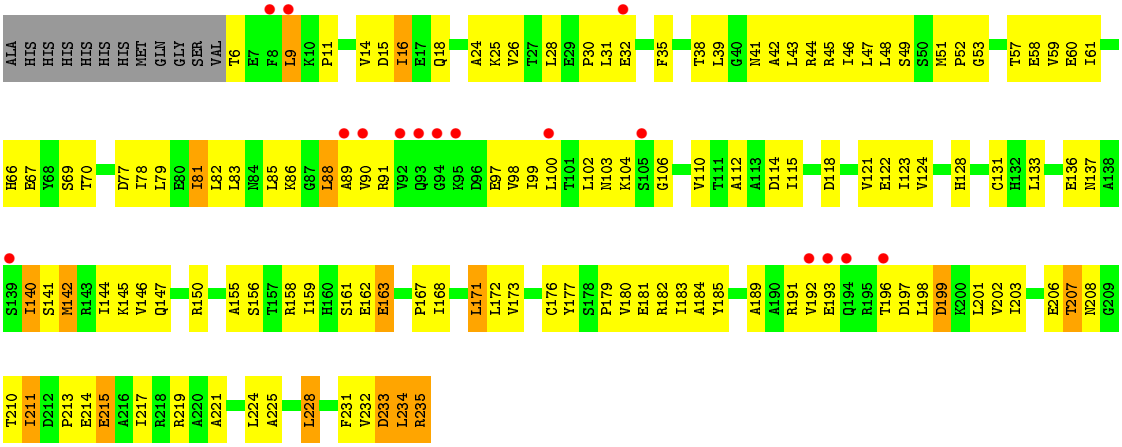
- Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain B:



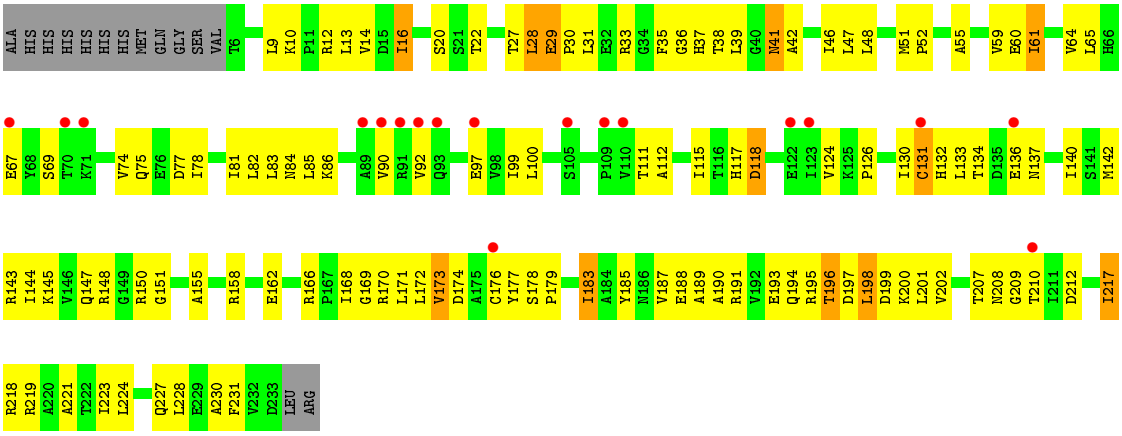
- Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain G:



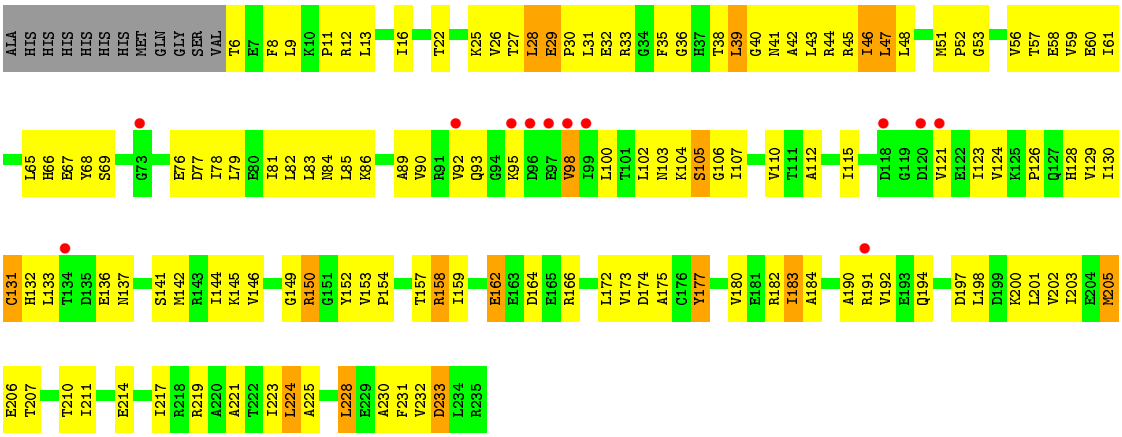
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain H:



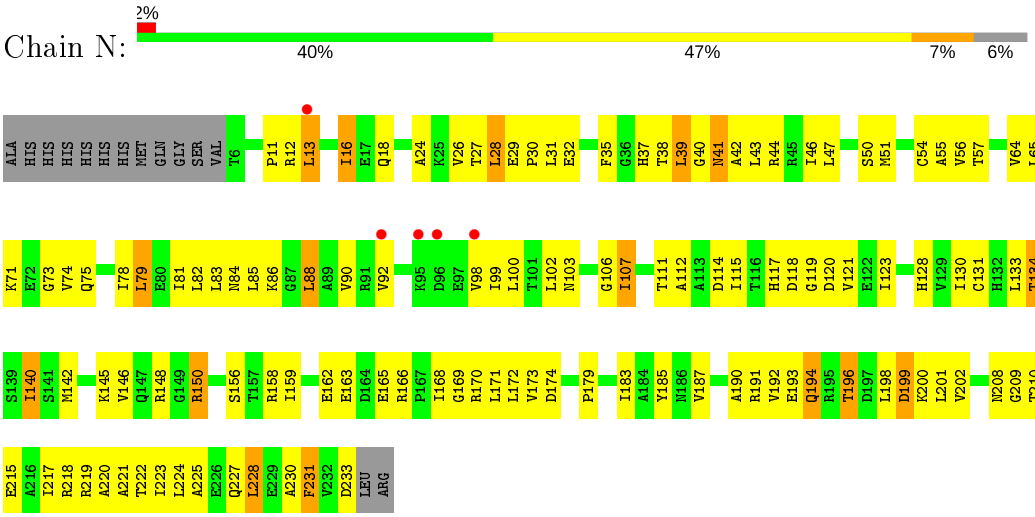
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain M:



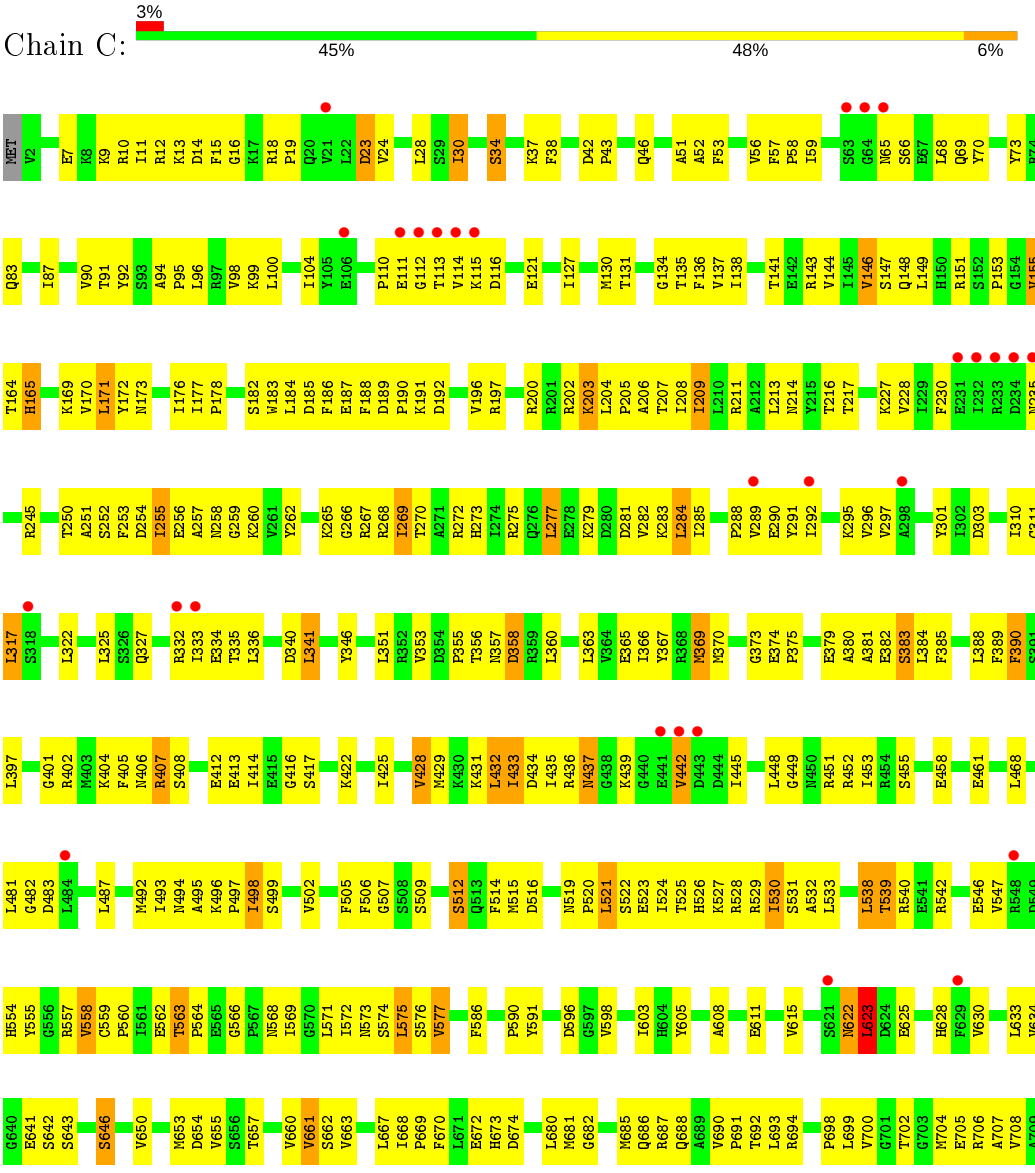
• Molecule 1: DNA-directed RNA polymerase subunit alpha

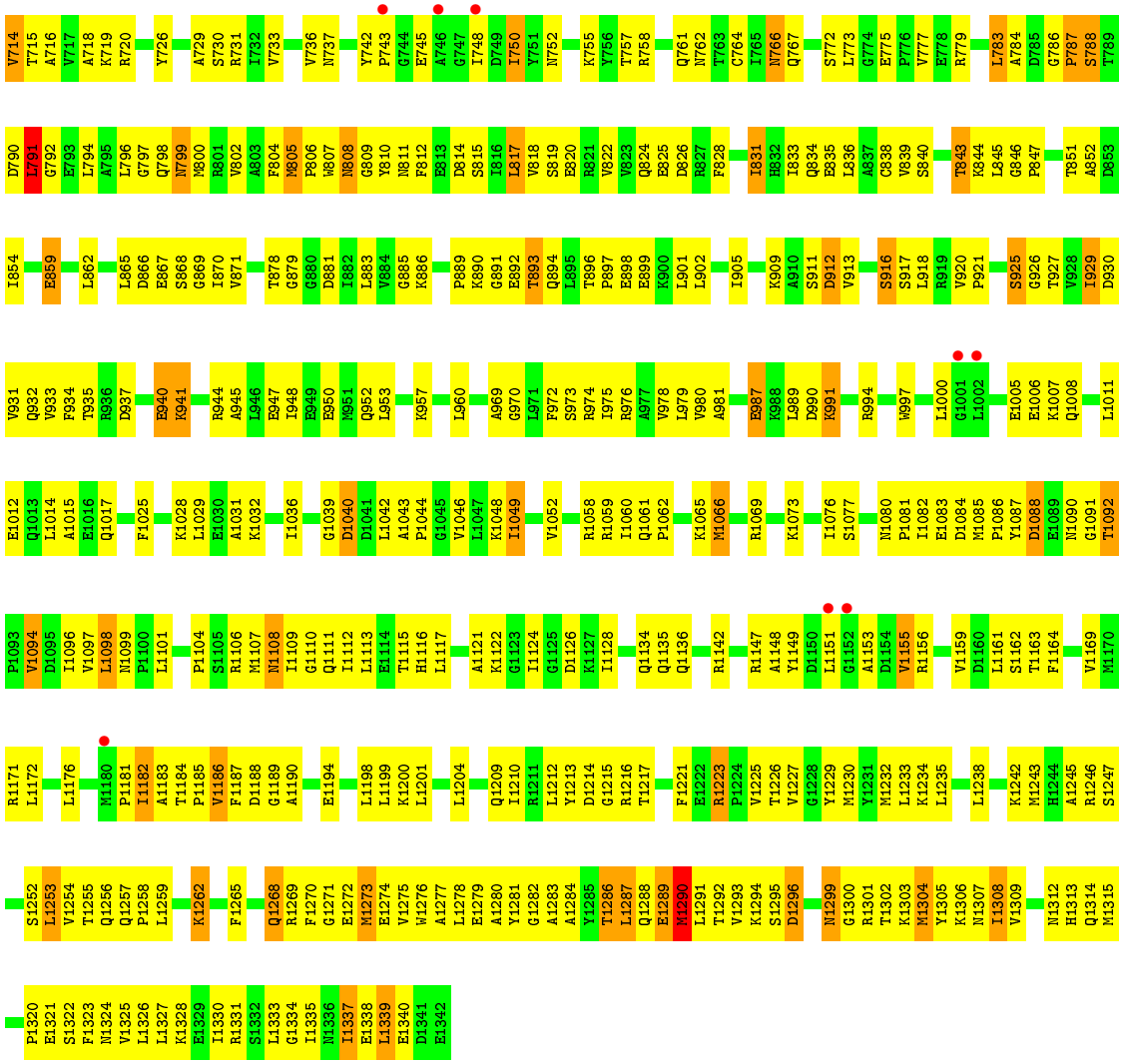
Chain N:



- Molecule 2: DNA-directed RNA polymerase subunit beta

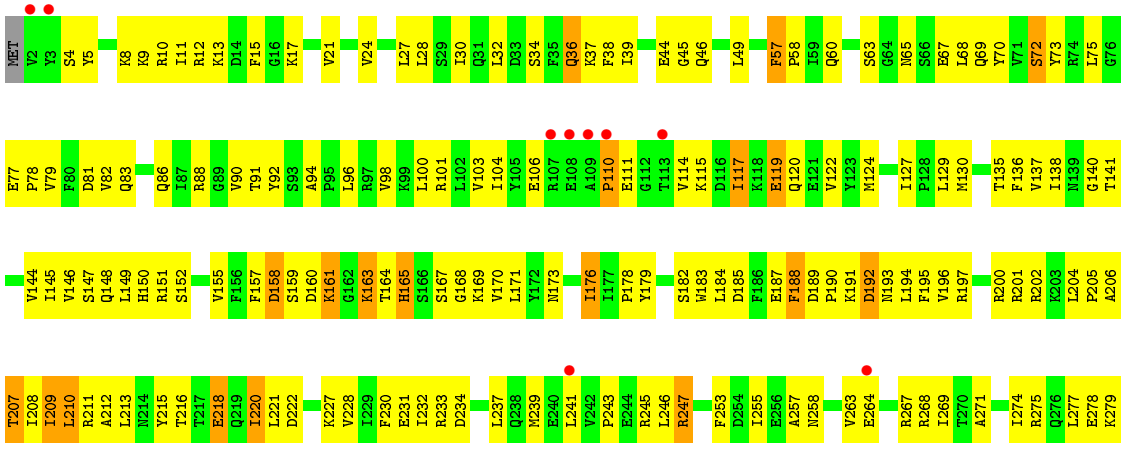
Chain C:

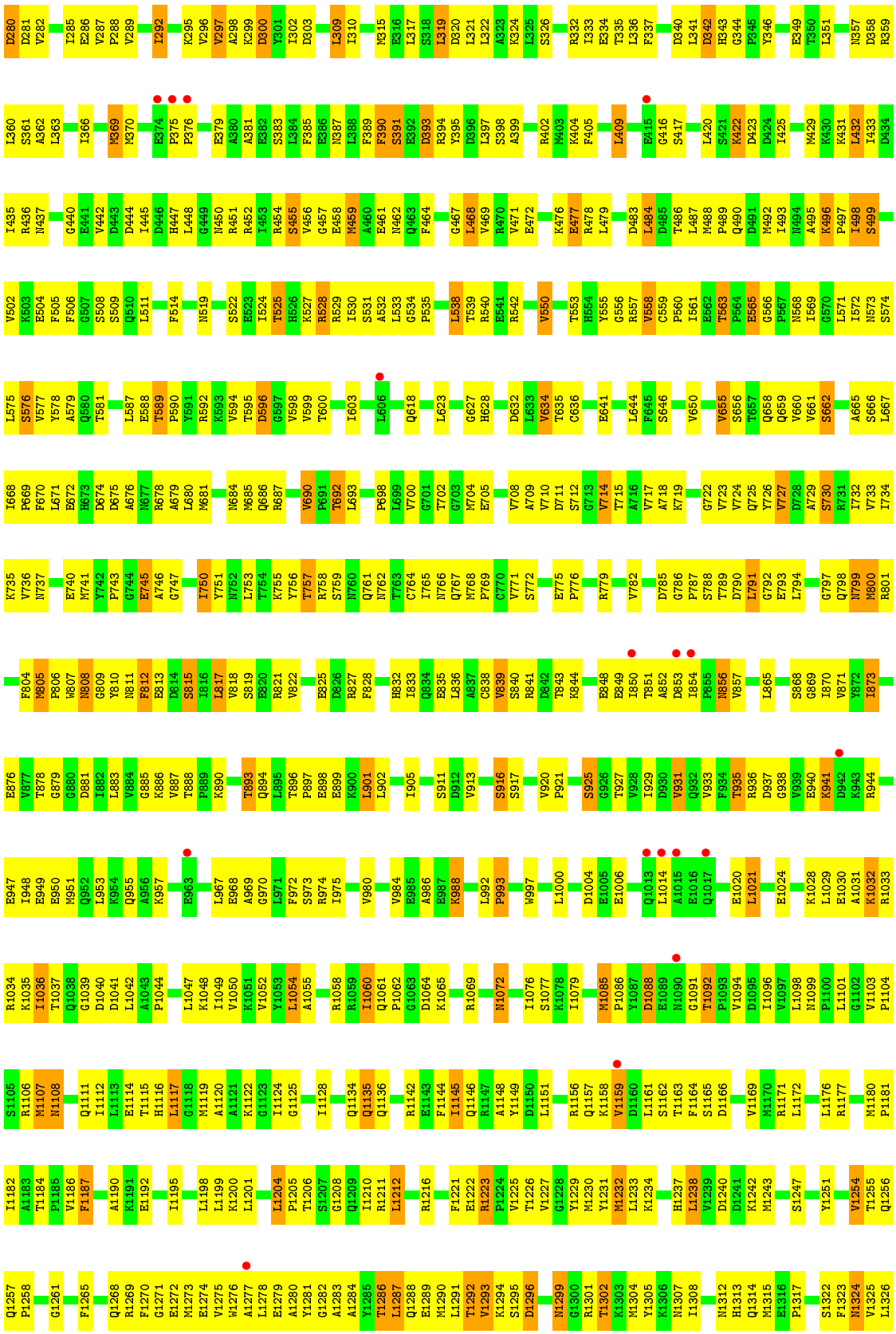




• Molecule 2: DNA-directed RNA polymerase subunit beta

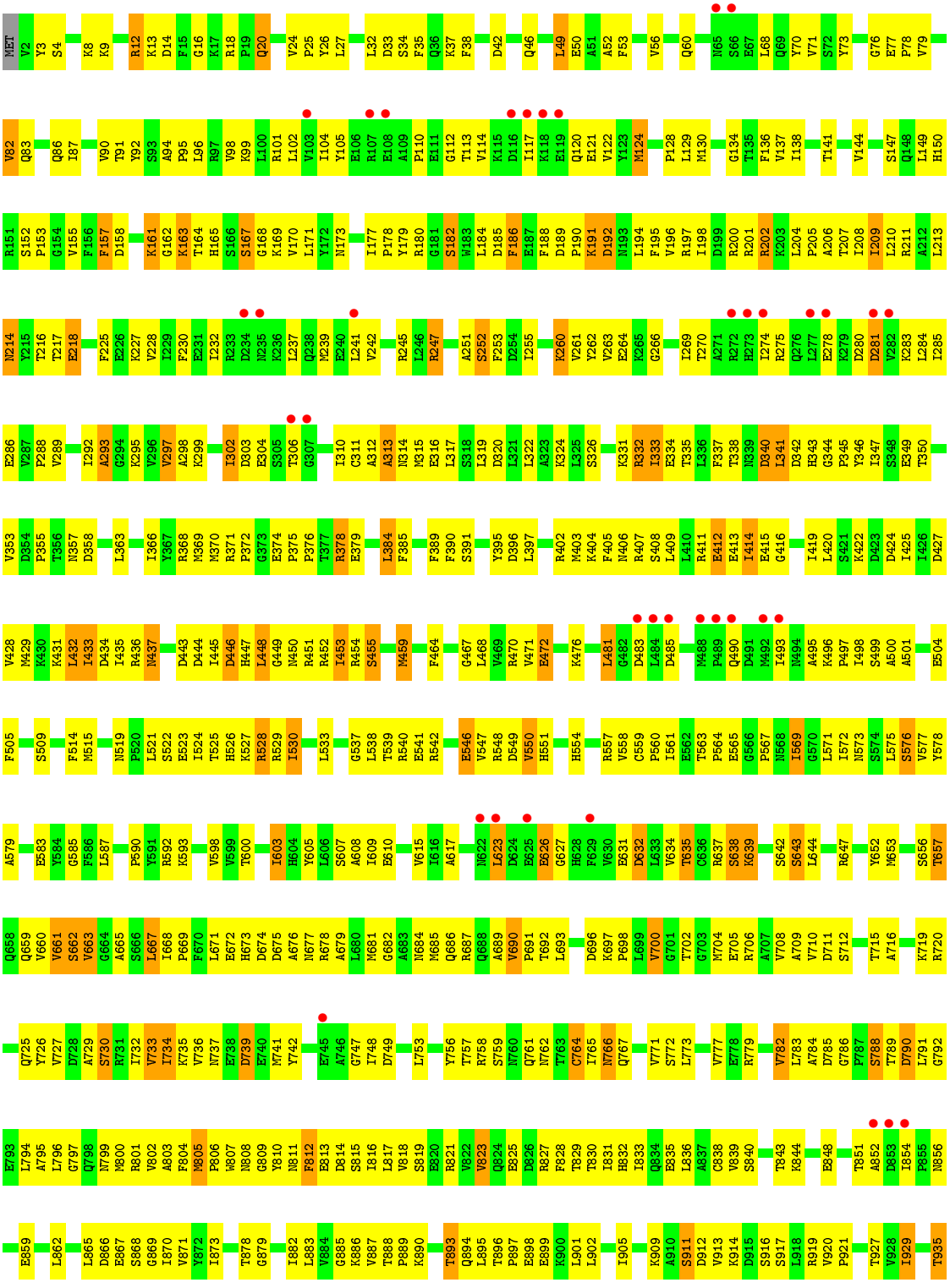
Chain I:

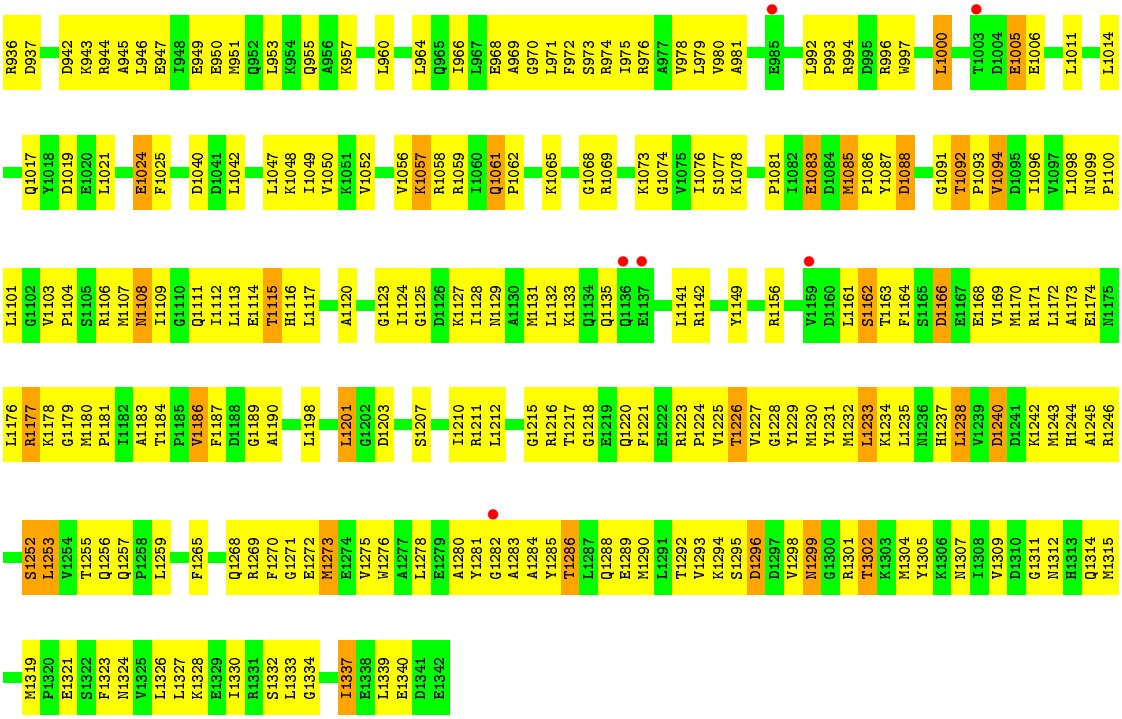




- Molecule 2: DNA-directed RNA polymerase subunit beta

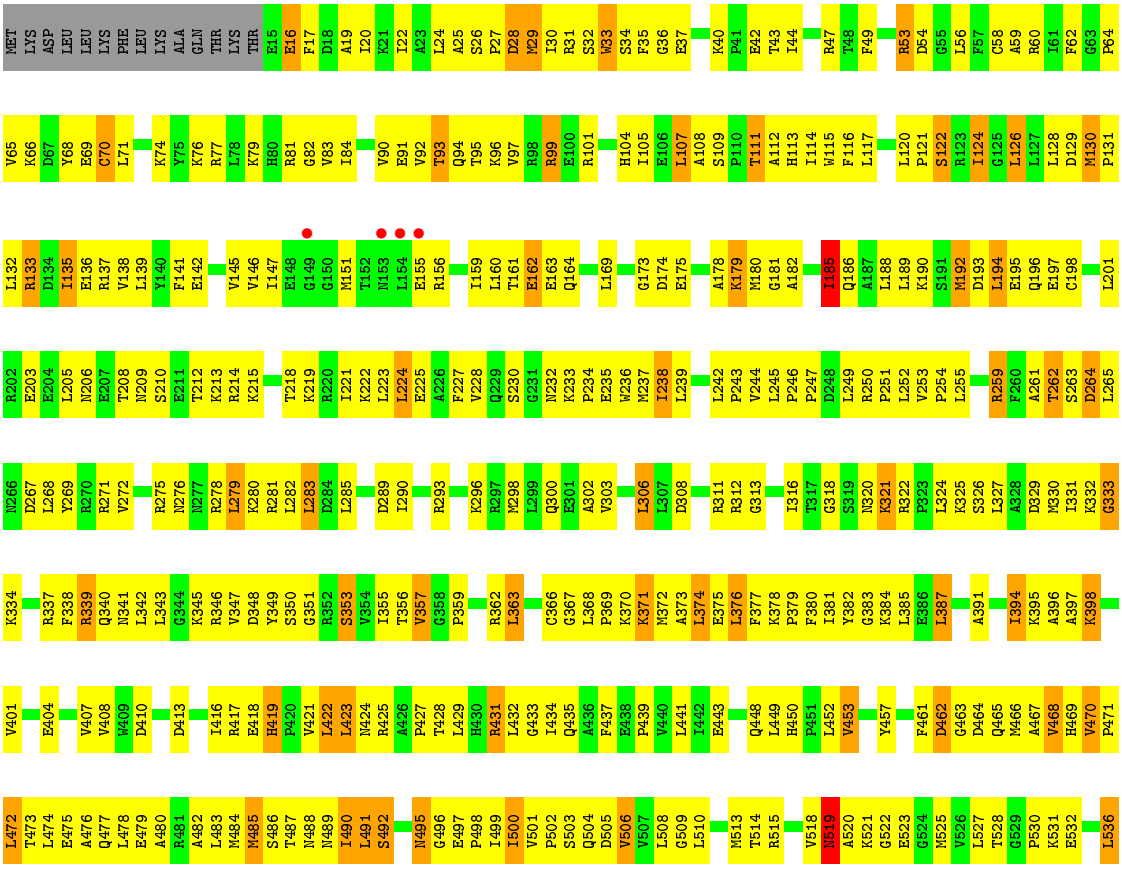
Chain O:

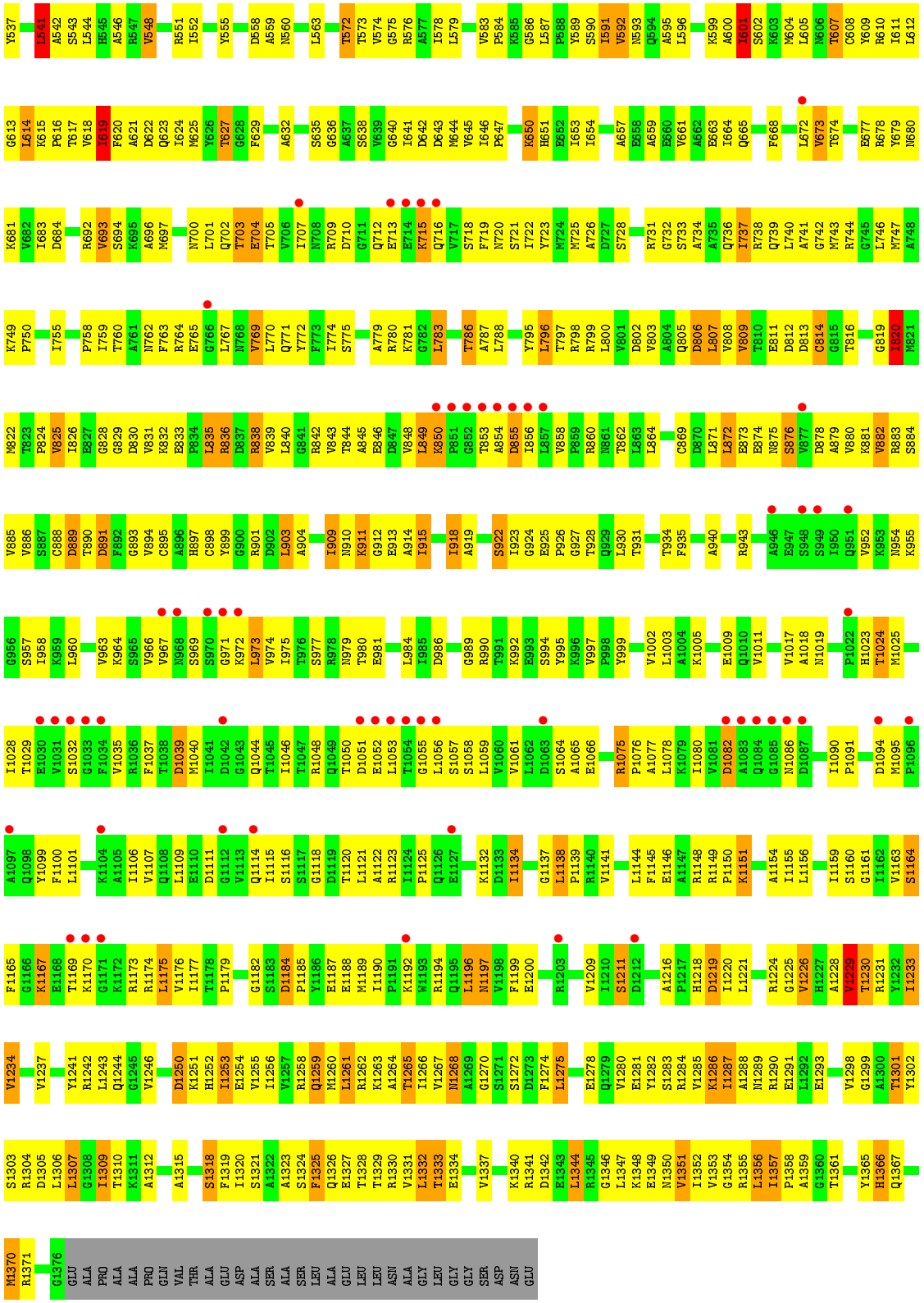




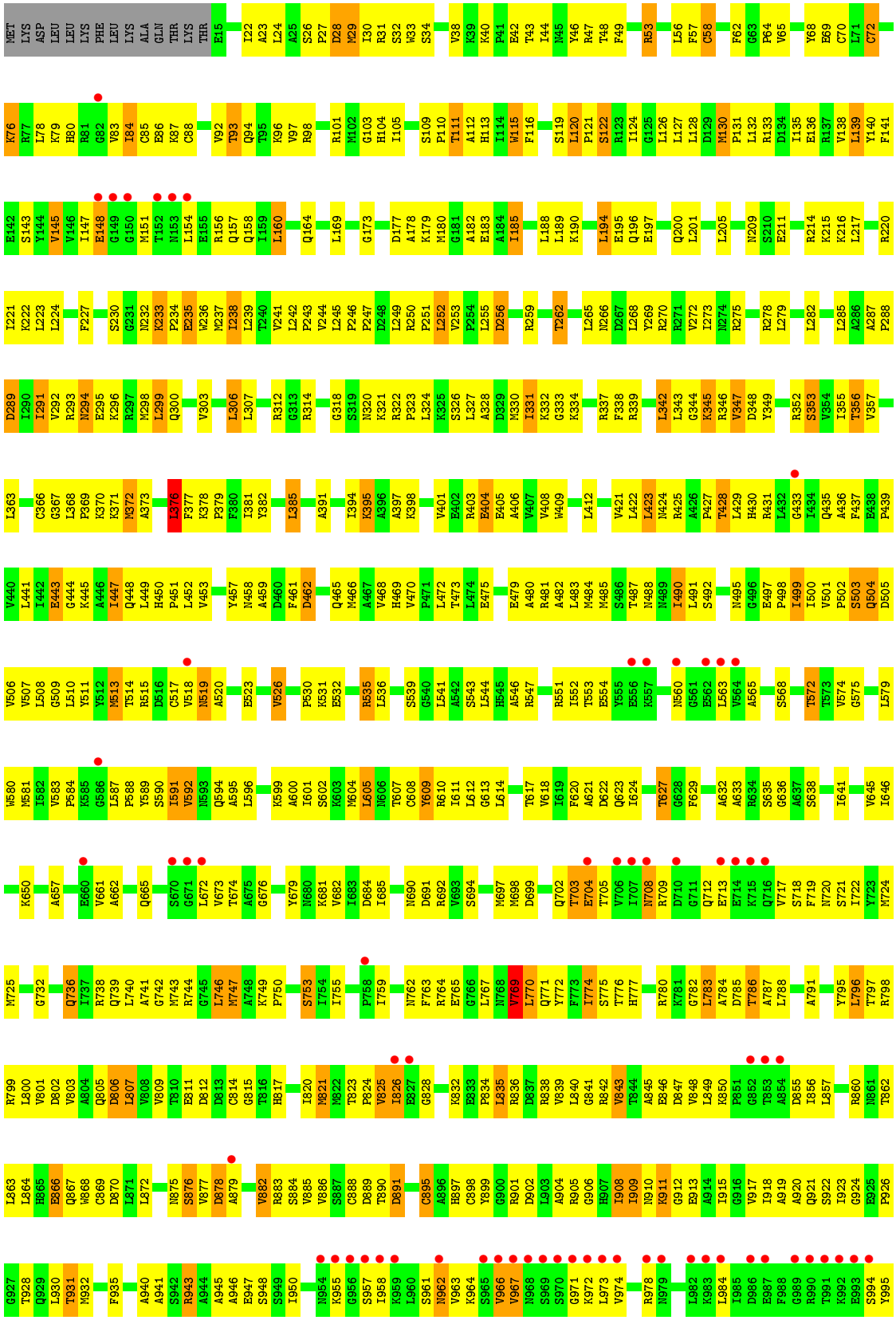
• Molecule 3: DNA-directed RNA polymerase subunit beta'

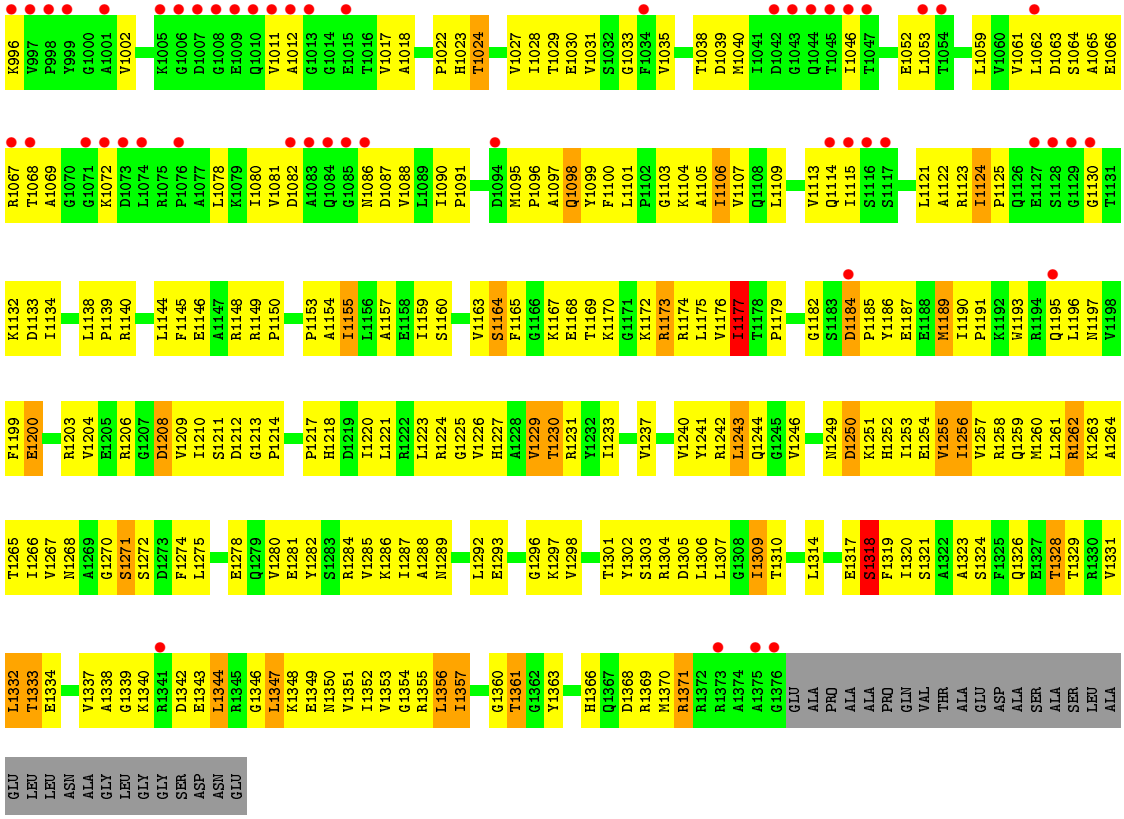
Chain D:



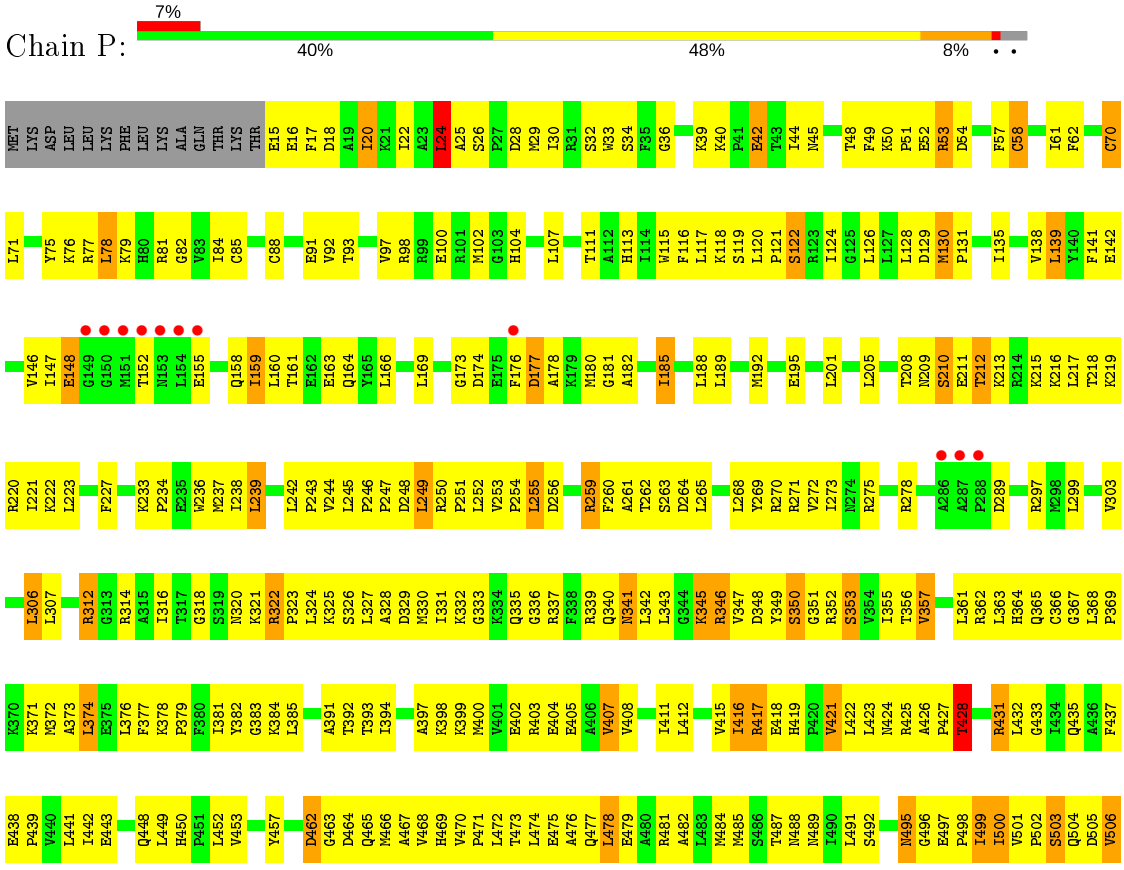


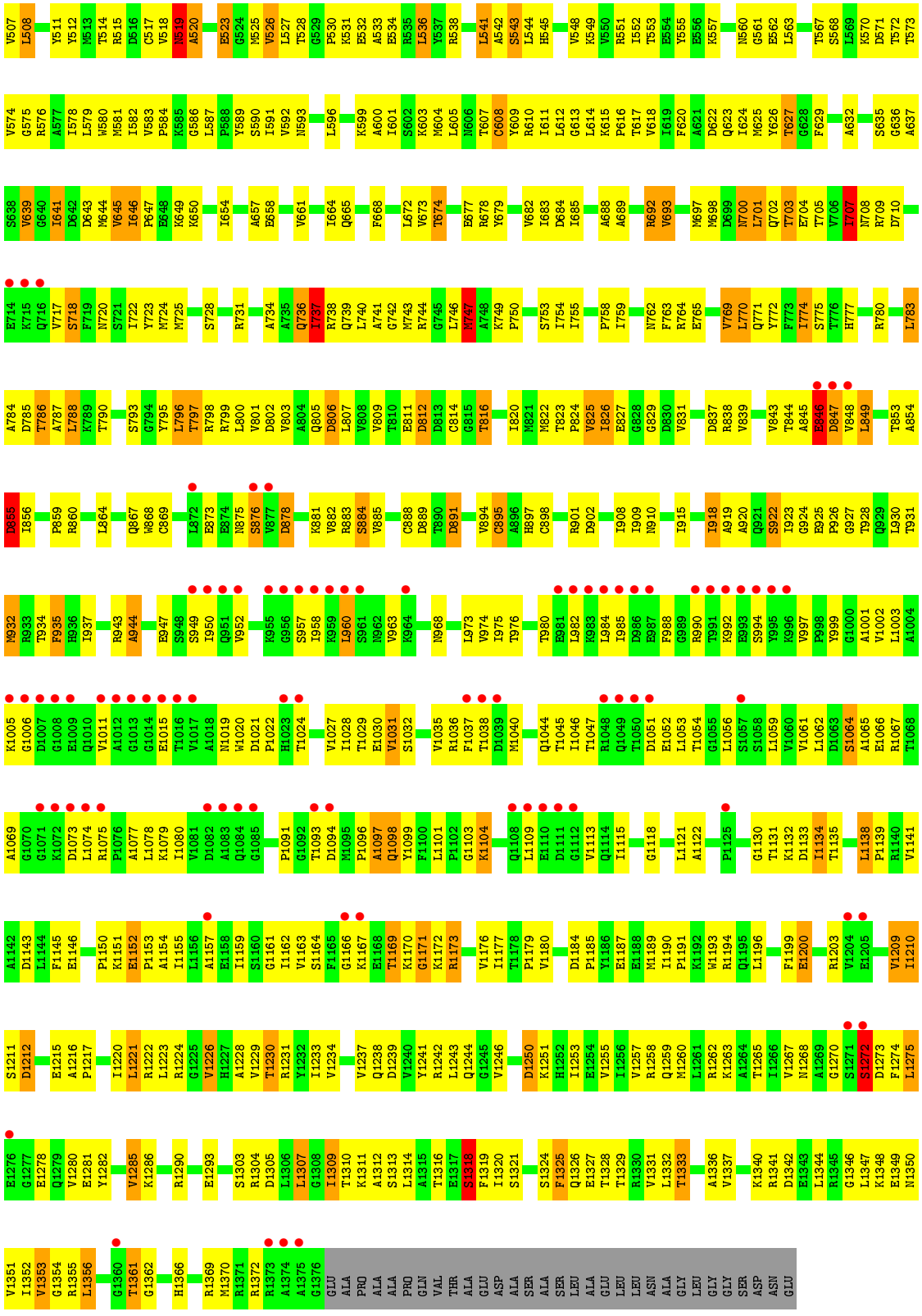
• Molecule 3: DNA-directed RNA polymerase subunit beta'


MET LYS ASP LEU LEU LYS PHE LEU LYS ALA GLN THR LYS THR
E15 I22 A23 L24 A25 S26 P27 D28 M29 I30 R31 S32 W33 S34 V38 K39 K40 P41 E42 T43 I44 N45 Y46 R47 T48 F49 R53 L56 F57 C58 F62 G63 P64 V65 Y68 E69 C70 L71 C72
K76 R77 L78 K79 H80 R81 G82 V83 I84 C85 E86 K87 C88 V92 T93 Q94 T95 K96 V97 R98 R101 M102 G103 H104 I105 S109 P110 T111 A112 H113 I114 W115 F116 S119 L120 P121 S122 R123 I124 G125 L126 L127 L128 D129 M130 P131 L132 R133 D134 I135 E136 R137 V138 L139 Y140 F141
E142 S143 Y144 V145 V146 I147 E148 G149 G150 M151 T152 N153 L154 E155 R156 Q157 Q158 I159 L160 Q164 L169 G173 D177 A178 K179 M180 G181 A182 E183 A184 I185 L188 L189 K190 L194 E195 Q196 E197 Q200 L201 L205 N209 S210 E211 R214 K215 K216 L217 R220
I221 K222 L223 L224 F227 S230 G231 N232 K233 P234 E235 W236 M237 I238 L239 T240 V241 L242 P243 V244 L245 P246 P247 D248 L249 R250 P251 L252 V253 P254 L255 D256 R259 T262 L265 N266 D267 I268 Y269 R270 R271 V272 I273 N274 R275 R278 L279 L282 L285 A286 A287 P288
D289 I290 I291 V292 R293 N294 E295 K296 R297 M298 L299 Q300 V303 L306 L307 R312 G313 R314 G318 S319 N320 K321 R322 P323 L324 K325 S326 L327 A328 D329 M330 I331 K332 G333 K334 R337 F338 R339 L342 L343 G344 K345 R346 V347 D348 Y349 R352 S353 V354 I355 T356 V357
I363 C366 G367 L368 P369 K370 K371 M372 A373 I376 F377 K378 P379 F380 I381 Y382 I385 A391 I394 N395 A396 A397 N398 V401 E402 R403 E404 E405 A406 V407 V408 W409 I412 V421 L422 L423 N424 R425 A426 P427 T428 L429 H430 R431 I432 G433 I434 Q435 A436 F437 E438 P439
V440 L441 Y442 E443 G444 K445 A446 I447 Q448 L449 H450 P451 L452 V453 Y457 N458 A459 D460 F461 D462 Q465 M466 A467 V468 H469 V470 P471 L472 T473 L474 E475 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 Y511 V512 N513 T514 R515 D516 C517 V518 N519 A520 E523 V526 P530 K531 E532 R535 L536 S539 G540 L541 A542 S543 L544 H545 A546 R547 R551 I552 T553 E554 Y555 E556 K557 N560 G561 E562 L563 V564 A565 S568 T572 T573 V574 G575 L579
W580 N581 I582 V583 P584 K585 G586 L587 P588 Y589 S590 I591 V592 N593 Q594 A595 L596 K599 A600 I601 S602 K603 W604 L605 N606 T607 C608 Y609 G610 I611 L612 G613 L614 T617 V618 I619 F620 A621 D622 Q623 I624 T627 G628 F629 A632 A633 E634 S635 G636 A637 S638 I641 V645 I646
K650 A657 E660 V661 A662 Q665 S670 G671 L672 V673 T674 A675 G676 Y679 N680 K681 V682 I683 D684 I685 N690 D691 R692 V693 S694 M697 M698 D699 Q702 T703 E704 T705 V706 I707 N708 R709 D710 G711 Q712 E713 E714 K715 Q716 V717 S718 F719 N720 S721 I722 Y723 M724
M725 G732 Q736 I737 R738 Q739 L740 A741 G742 M743 R744 G745 L746 M747 A748 K749 P750 S753 I754 I755 P758 I759 N762 F763 R764 E765 G766 L767 N768 V769 L770 Q771 Y772 F773 I774 S775 T776 H777 R780 K781 G782 L783 A784 D785 T786 A787 L788 A791 Y795 L796 T797 R798
R799 L800 V801 D802 V803 A804 Q805 D806 L807 V808 V809 T810 E811 D812 D813 C814 G815 T816 H817 I820 M821 M822 T823 P824 V825 I826 E827 G828 K832 E833 P834 L835 R836 D837 R838 V839 L840 G841 R842 V843 T844 A845 E846 D847 V848 L849 K850 P851 G852 T853 A854 D855 I856 L857 R860 N861 T862
L863 L864 H865 E866 Q867 W868 C869 D870 L871 L872 N875 S876 V877 D878 A879 V882 R883 S884 V885 V886 S887 C888 D889 T890 D891 C895 A896 H897 C898 Y899 G900 R901 D902 L903 A904 R905 G906 H907 I908 I909 N910 K911 G912 E913 A914 I915 G916 V917 I918 A919 A920 Q921 S922 I923 G924 E925 P926
G927 T928 Q929 L930 T931 M932 F935 A940 A941 S942 R943 A944 A945 A946 E947 S948 S949 I950 N954 K955 G956 S957 I958 K959 L960 S961 N962 V963 K964 S965 V966 V967 N968 S969 S970 G971 K972 L973 V974 R978 N979 L982 K983 L984 I985 D986 E987 F988 G989 R990 T991 K992 E993 S994 Y995

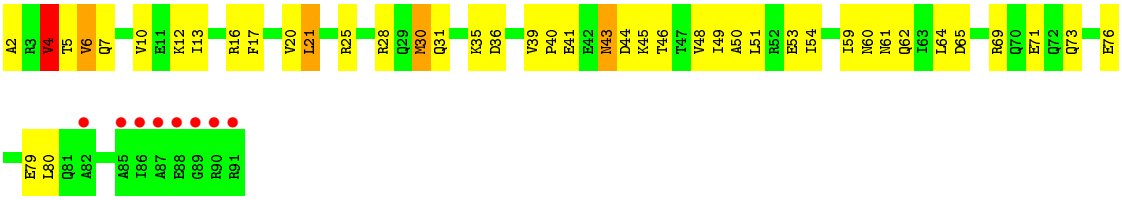


• Molecule 3: DNA-directed RNA polymerase subunit beta'



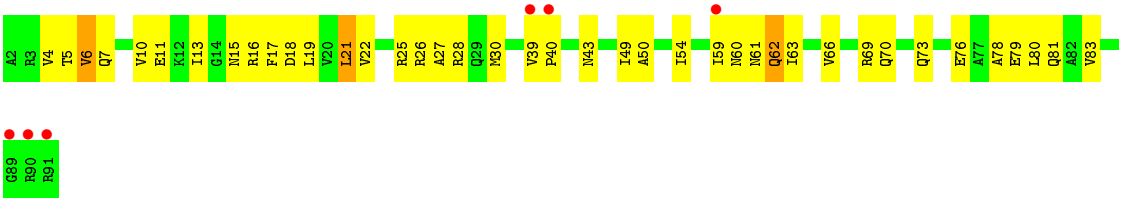


• Molecule 4: DNA-directed RNA polymerase subunit omega



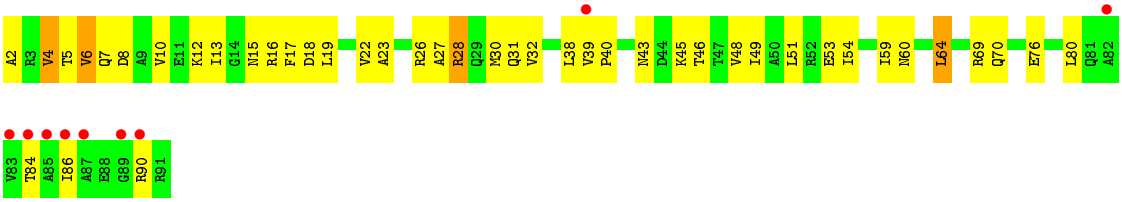
• Molecule 4: DNA-directed RNA polymerase subunit omega

Chain K:



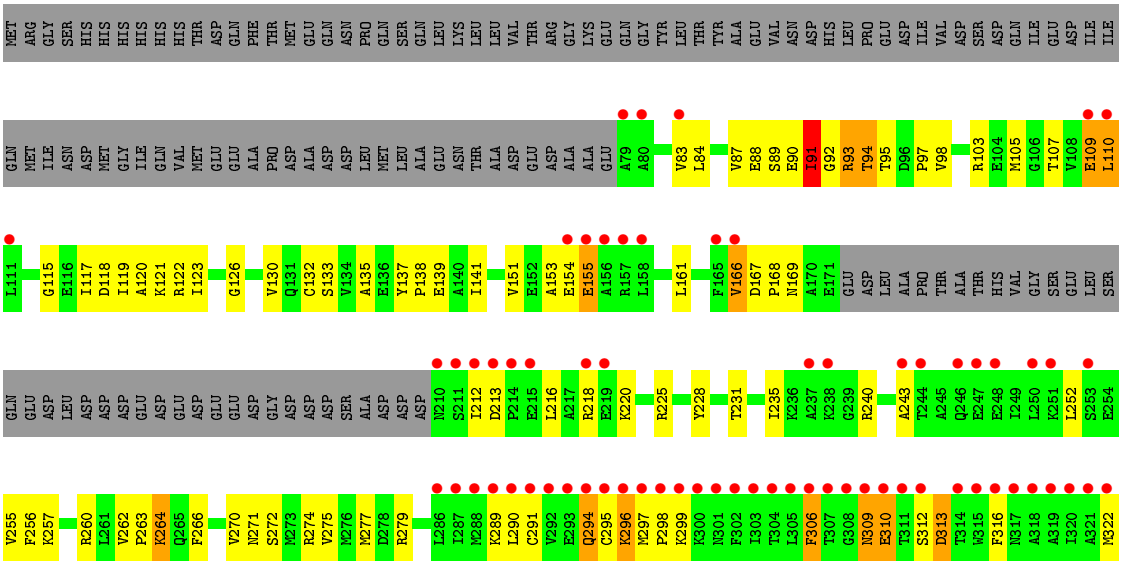
• Molecule 4: DNA-directed RNA polymerase subunit omega

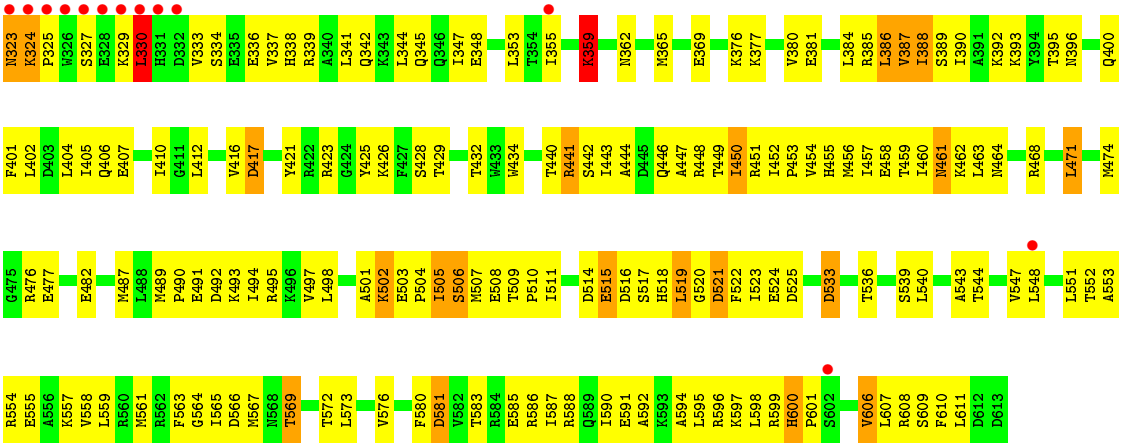
Chain Q:



• Molecule 5: RNA polymerase sigma factor RpoD

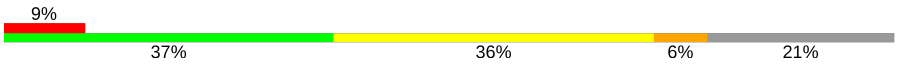
Chain F:

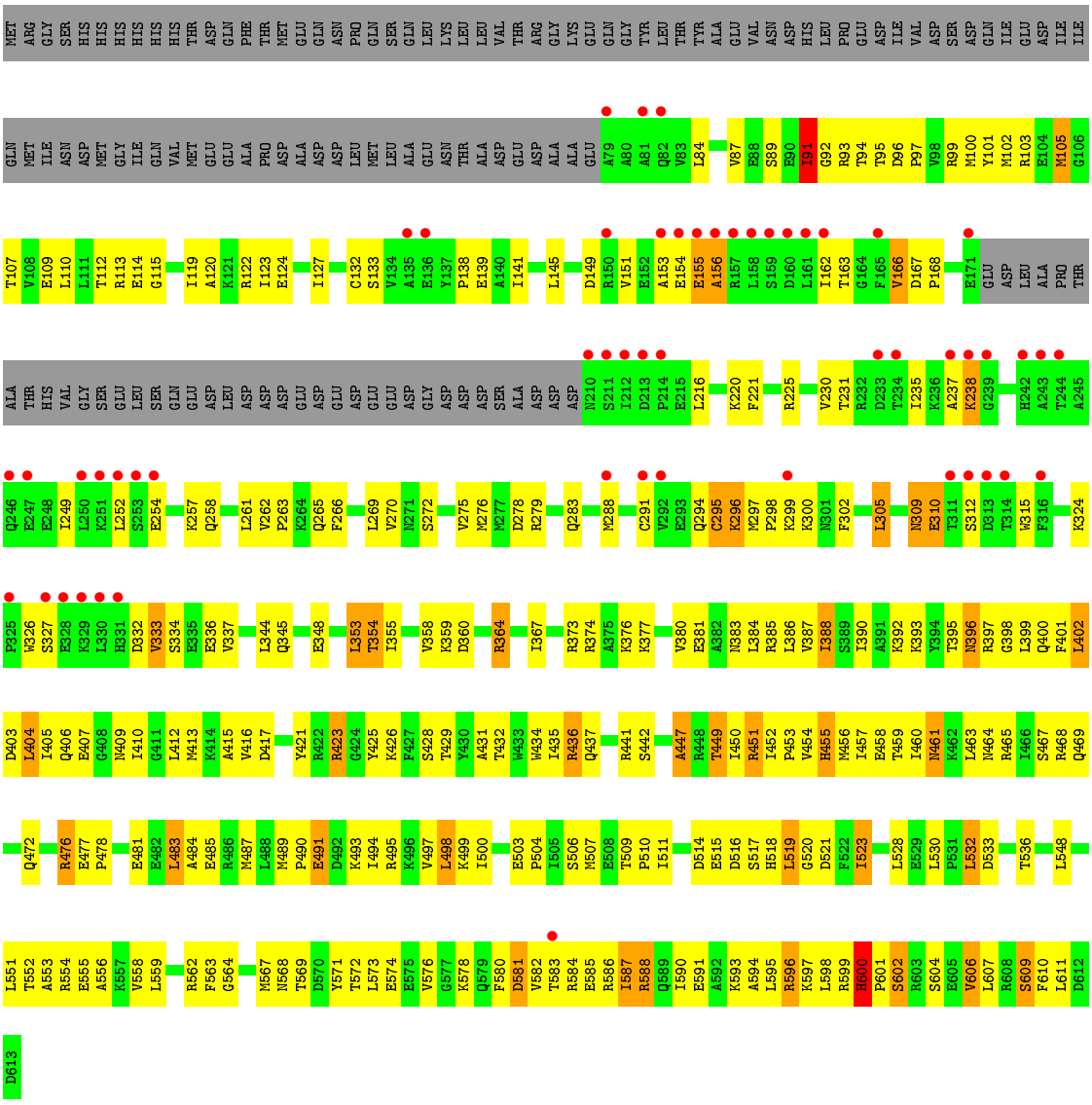




• Molecule 5: RNA polymerase sigma factor RpoD

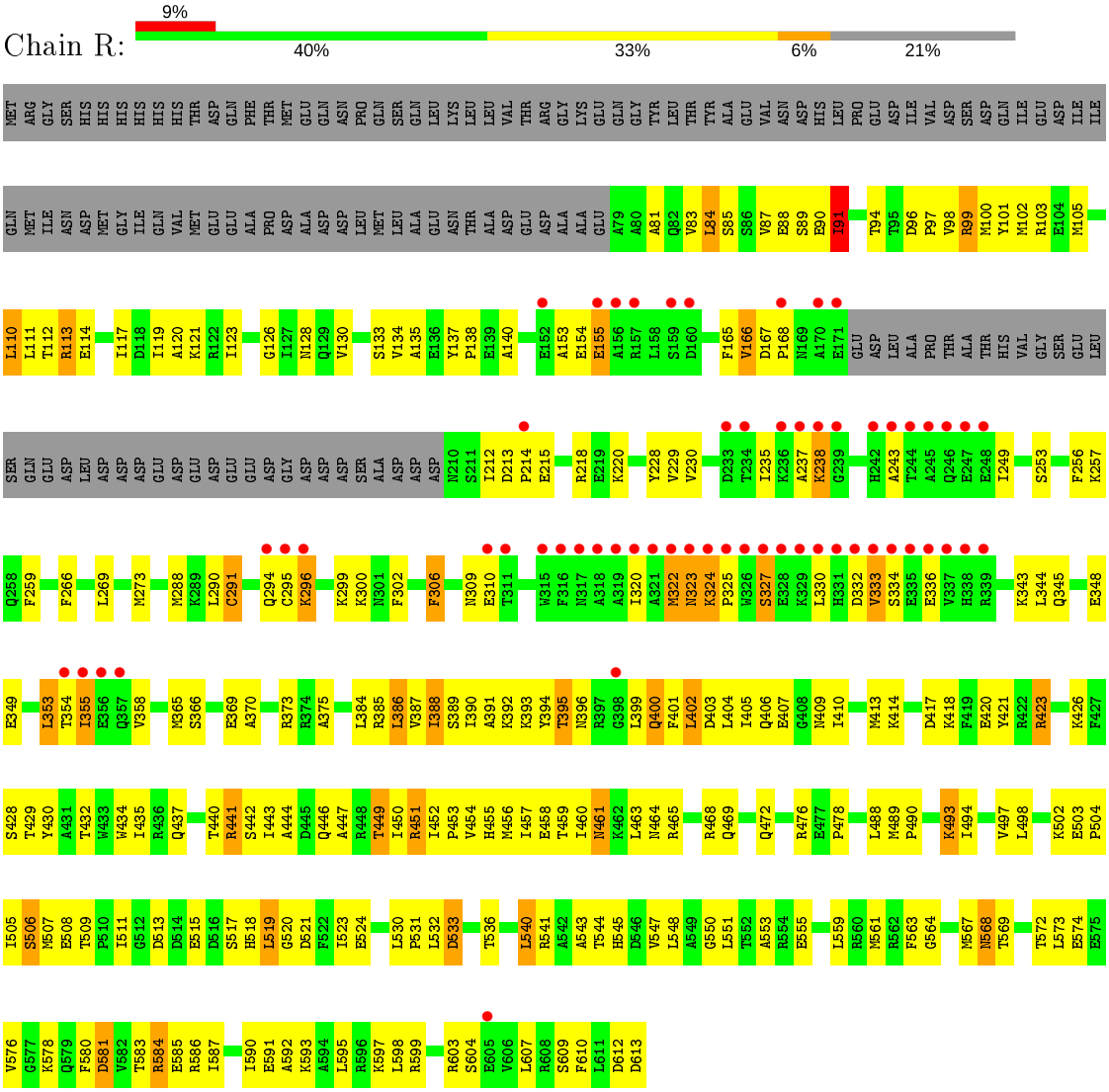
Chain L:





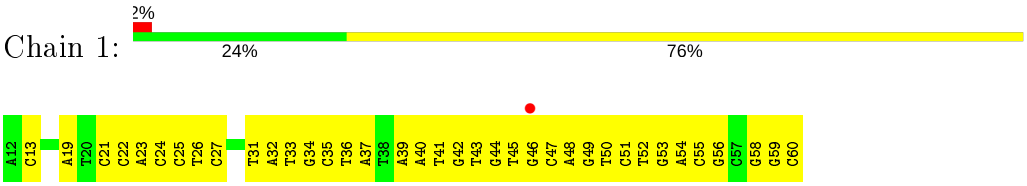
• Molecule 5: RNA polymerase sigma factor RpoD

Chain R:



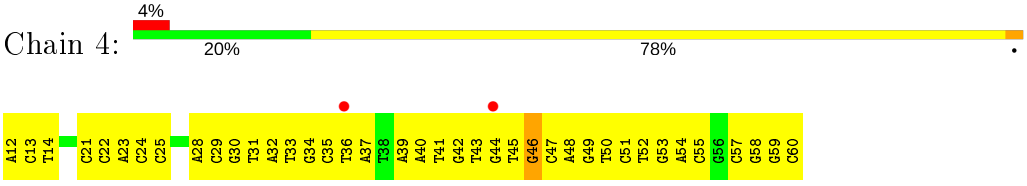
• Molecule 6: NT strand DNA (49-MER)

Chain 1:



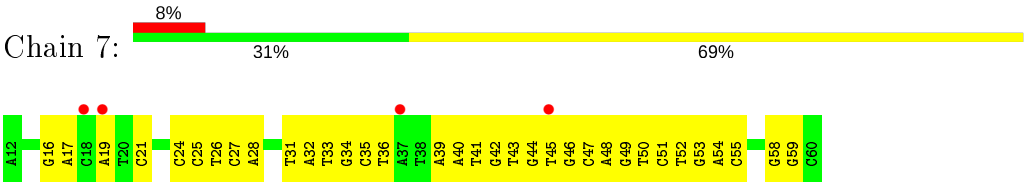
• Molecule 6: NT strand DNA (49-MER)

Chain 4:



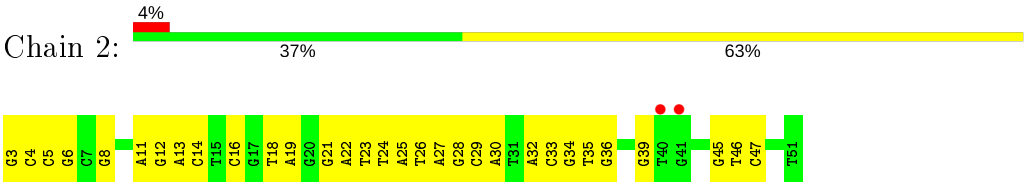
• Molecule 6: NT strand DNA (49-MER)

Chain 7:



• Molecule 7: T strand DNA (49-MER)

Chain 2:

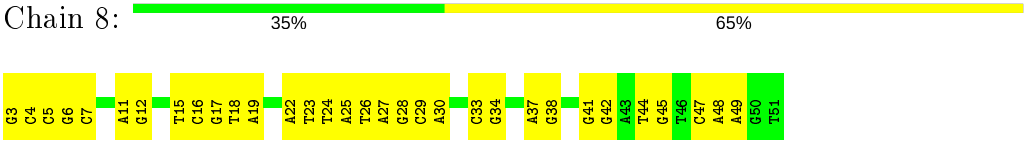


• Molecule 7: T strand DNA (49-MER)

Chain 5:

• Molecule 7: T strand DNA (49-MER)

Chain 8:



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

Chain 3:

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

Chain 6:

• Molecule 8: RNA (5'-D(*(GTP))-R(P*AP*GP*U)-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	240.89Å 208.17Å 256.32Å 90.00° 119.31° 90.00°	Depositor
Resolution (Å)	39.95 – 6.00 39.95 – 6.00	Depositor EDS
% Data completeness (in resolution range)	99.6 (39.95-6.00) 100.0 (39.95-6.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.44 (at 6.13Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.227 , 0.314 0.227 , 0.314	Depositor DCC
R_{free} test set	2938 reflections (5.31%)	wwPDB-VP
Wilson B-factor (Å^2)	343.5	Xtriage
Anisotropy	0.527	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 253.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.25$	Xtriage
Estimated twinning fraction	0.035 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	94608	wwPDB-VP
Average B, all atoms (Å^2)	238.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.84% 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: GTP, ZN, MG

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	0.59	0/1809	0.85	2/2450 (0.1%)
1	B	0.57	0/1789	0.84	1/2425 (0.0%)
1	G	0.60	1/1809 (0.1%)	0.83	1/2450 (0.0%)
1	H	0.57	0/1789	0.79	1/2425 (0.0%)
1	M	0.63	0/1809	0.87	2/2450 (0.1%)
1	N	0.60	0/1789	0.87	0/2425
2	C	0.58	3/10745 (0.0%)	0.79	11/14499 (0.1%)
2	I	0.61	3/10745 (0.0%)	0.81	3/14499 (0.0%)
2	O	0.61	4/10745 (0.0%)	0.81	4/14499 (0.0%)
3	D	0.66	3/10729 (0.0%)	0.91	20/14487 (0.1%)
3	J	0.63	2/10729 (0.0%)	0.85	15/14487 (0.1%)
3	P	0.59	4/10729 (0.0%)	0.80	9/14487 (0.1%)
4	E	0.62	0/710	0.89	1/956 (0.1%)
4	K	0.56	0/710	0.72	0/956
4	Q	0.55	0/710	0.74	1/956 (0.1%)
5	F	0.56	2/4076 (0.0%)	0.77	1/5482 (0.0%)
5	L	0.59	0/4076	0.78	2/5482 (0.0%)
5	R	0.55	0/4076	0.77	2/5482 (0.0%)
6	1	0.39	0/1112	0.67	0/1706
6	4	0.51	1/1114 (0.1%)	0.73	0/1714
6	7	0.39	0/1115	0.70	0/1718
7	2	0.38	0/1136	0.66	0/1752
7	5	0.41	0/1137	0.69	0/1756
7	8	0.36	0/1137	0.68	0/1756
8	3	0.33	0/72	0.58	0/110
8	6	0.50	0/72	0.71	0/110
8	9	0.50	0/72	0.65	0/110
All	All	0.59	23/96541 (0.0%)	0.81	76/131629 (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	J	0	1
3	P	0	1
All	All	0	2

The worst 5 of 23 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	O	316	GLU	CD-OE2	13.14	1.40	1.25
3	J	943	ARG	CZ-NH1	11.93	1.48	1.33
3	D	431	ARG	CZ-NH1	11.55	1.48	1.33
2	I	565	GLU	CB-CG	10.38	1.71	1.52
3	P	1152	GLU	CD-OE2	-9.69	1.15	1.25

The worst 5 of 76 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	F	359	LYS	CG-CD-CE	-8.94	85.08	111.90
3	D	431	ARG	NE-CZ-NH2	-8.33	116.13	120.30
3	J	943	ARG	NE-CZ-NH2	-8.10	116.25	120.30
1	M	29	GLU	C-N-CD	-7.94	103.14	120.60
3	D	376	LEU	CA-CB-CG	-7.68	97.62	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	J	943	ARG	Sidechain
3	P	210	SER	Peptide

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	1787	0	1813	190	0
1	B	1767	0	1789	200	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	1787	0	1813	179	0
1	H	1767	0	1789	145	0
1	M	1787	0	1813	260	0
1	N	1767	0	1789	213	0
2	C	10576	0	10591	945	0
2	I	10576	0	10591	991	0
2	O	10576	0	10591	1002	1
3	D	10568	0	10781	1353	1
3	J	10568	0	10782	1175	0
3	P	10568	0	10780	1041	1
4	E	708	0	719	65	0
4	K	708	0	719	40	0
4	Q	708	0	719	46	0
5	F	4022	0	4083	368	1
5	L	4022	0	4083	367	0
5	R	4022	0	4083	348	0
6	1	996	0	557	71	0
6	4	996	0	555	103	1
6	7	996	0	554	69	0
7	2	1012	0	554	65	0
7	5	1012	0	553	62	0
7	8	1012	0	553	66	0
8	3	97	0	44	21	0
8	6	97	0	44	6	0
8	9	97	0	44	19	0
9	C	1	0	0	0	0
9	J	1	0	0	0	0
9	P	1	0	0	0	0
10	D	2	0	0	2	0
10	J	2	0	0	2	0
10	P	2	0	0	1	0
All	All	94608	0	92786	8368	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 45.

The worst 5 of 8368 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)
5:F:359:LYS:NZ	5:F:359:LYS:CE	1.67	1.54
3:D:484:MET:SD	3:D:484:MET:CE	2.03	1.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:P:139:LEU:CD2	3:P:185:ILE:HD11	1.49	1.43
3:J:367:GLY:O	3:J:447:ILE:CG2	1.68	1.38
1:G:25:LYS:NZ	1:G:202:VAL:HG11	1.34	1.37

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)
6:4:12:DA:O5'	6:4:60:DC:O5'[2_454]	1.86	0.34
5:F:482:GLU:OE2	2:O:275:ARG:NH2[2_455]	1.99	0.21
3:D:1282:TYR:OH	3:P:710:ASP:OD2[1_655]	2.08	0.12

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%)	210 (92%)	15 (7%)	3 (1%)	12 48
1	B	226/242 (93%)	207 (92%)	13 (6%)	6 (3%)	5 31
1	G	228/242 (94%)	211 (92%)	14 (6%)	3 (1%)	12 48
1	H	226/242 (93%)	207 (92%)	17 (8%)	2 (1%)	17 56
1	M	228/242 (94%)	209 (92%)	16 (7%)	3 (1%)	12 48
1	N	226/242 (93%)	207 (92%)	14 (6%)	5 (2%)	6 35
2	C	1339/1342 (100%)	1218 (91%)	98 (7%)	23 (2%)	9 42
2	I	1339/1342 (100%)	1210 (90%)	102 (8%)	27 (2%)	7 38
2	O	1339/1342 (100%)	1222 (91%)	87 (6%)	30 (2%)	6 35
3	D	1360/1407 (97%)	1210 (89%)	122 (9%)	28 (2%)	7 36
3	J	1360/1407 (97%)	1225 (90%)	110 (8%)	25 (2%)	8 40
3	P	1360/1407 (97%)	1208 (89%)	112 (8%)	40 (3%)	4 29

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
4	E	88/90 (98%)	82 (93%)	5 (6%)	1 (1%)	14 52
4	K	88/90 (98%)	83 (94%)	4 (4%)	1 (1%)	14 52
4	Q	88/90 (98%)	81 (92%)	6 (7%)	1 (1%)	14 52
5	F	493/628 (78%)	443 (90%)	34 (7%)	16 (3%)	4 26
5	L	493/628 (78%)	441 (90%)	36 (7%)	16 (3%)	4 26
5	R	493/628 (78%)	441 (90%)	36 (7%)	16 (3%)	4 26
All	All	11202/11853 (94%)	10115 (90%)	841 (8%)	246 (2%)	6 35

5 of 246 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	209	GLY
1	A	210	THR
2	C	113	THR
2	C	481	LEU
2	C	791	LEU

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%)	171 (86%)	27 (14%)	3 17
1	B	196/208 (94%)	172 (88%)	24 (12%)	5 20
1	G	198/208 (95%)	169 (85%)	29 (15%)	3 16
1	H	196/208 (94%)	171 (87%)	25 (13%)	4 18
1	M	198/208 (95%)	171 (86%)	27 (14%)	3 17
1	N	196/208 (94%)	167 (85%)	29 (15%)	3 15
2	C	1156/1157 (100%)	1016 (88%)	140 (12%)	5 20
2	I	1156/1157 (100%)	1013 (88%)	143 (12%)	4 19
2	O	1156/1157 (100%)	1019 (88%)	137 (12%)	5 20
3	D	1135/1168 (97%)	968 (85%)	167 (15%)	3 15

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
3	J	1135/1168 (97%)	986 (87%)	149 (13%)	4 18
3	P	1135/1168 (97%)	989 (87%)	146 (13%)	4 18
4	E	74/74 (100%)	63 (85%)	11 (15%)	3 15
4	K	74/74 (100%)	70 (95%)	4 (5%)	22 47
4	Q	74/74 (100%)	68 (92%)	6 (8%)	11 35
5	F	439/554 (79%)	388 (88%)	51 (12%)	5 21
5	L	439/554 (79%)	382 (87%)	57 (13%)	4 18
5	R	439/554 (79%)	393 (90%)	46 (10%)	7 24
All	All	9594/10107 (95%)	8376 (87%)	1218 (13%)	4 19

5 of 1218 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	I	727	VAL
3	J	568	SER
3	P	878	ASP
2	I	843	THR
3	J	29	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 152 such sidechains are listed below:

Mol	Chain	Res	Type
2	I	725	GLN
3	J	690	ASN
3	P	1098	GLN
2	I	856	ASN
3	J	209	ASN

5.3.3 RNA

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
8	3	3/4 (75%)	1 (33%)	1 (33%)
8	6	3/4 (75%)	1 (33%)	1 (33%)
8	9	3/4 (75%)	1 (33%)	1 (33%)
All	All	9/12 (75%)	3 (33%)	3 (33%)

All (3) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
8	3	15	G
8	6	15	G
8	9	15	G

All (3) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
8	3	13	GTP
8	6	13	GTP
8	9	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 carbohydrates 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	1	3
7	2	1
6	4	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	4	51:DC	O3'	52:DT	P	3.86
1	1	51:DC	O3'	52:DT	P	3.85
1	1	46:DG	O3'	47:DC	P	3.49
1	2	12:DG	O3'	13:DA	P	2.97
1	1	36:DT	O3'	37:DA	P	2.64

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%)	-0.02	3 (1%) 77 68	162, 207, 238, 299	0
1	B	228/242 (94%)	-0.15	5 (2%) 62 54	165, 211, 254, 284	0
1	G	230/242 (95%)	0.09	16 (6%) 16 16	175, 216, 259, 292	0
1	H	228/242 (94%)	0.39	18 (7%) 12 14	200, 264, 307, 325	0
1	M	230/242 (95%)	-0.01	12 (5%) 27 27	155, 186, 230, 311	0
1	N	228/242 (94%)	-0.18	5 (2%) 62 54	162, 205, 248, 280	0
2	C	1341/1342 (99%)	0.01	38 (2%) 53 46	126, 243, 333, 427	0
2	I	1341/1342 (99%)	-0.08	26 (1%) 66 59	142, 196, 276, 345	0
2	O	1341/1342 (99%)	-0.03	43 (3%) 47 41	126, 202, 287, 327	0
3	D	1362/1407 (96%)	0.06	62 (4%) 32 30	129, 204, 384, 437	0
3	J	1362/1407 (96%)	0.26	118 (8%) 10 12	145, 213, 344, 397	0
3	P	1362/1407 (96%)	0.18	97 (7%) 16 16	133, 223, 391, 441	0
4	E	90/90 (100%)	0.15	8 (8%) 9 12	138, 184, 398, 452	0
4	K	90/90 (100%)	0.24	6 (6%) 17 17	165, 213, 389, 436	0
4	Q	90/90 (100%)	0.30	9 (10%) 7 10	157, 202, 400, 463	0
5	F	497/628 (79%)	0.51	80 (16%) 1 4	176, 304, 513, 589	0
5	L	497/628 (79%)	0.23	54 (10%) 5 8	171, 277, 369, 399	0
5	R	497/628 (79%)	0.41	59 (11%) 4 8	170, 322, 414, 449	0
6	1	49/49 (100%)	-0.21	1 (2%) 65 58	206, 263, 311, 333	0
6	4	49/49 (100%)	-0.03	2 (4%) 37 34	50, 261, 289, 312	0
6	7	49/49 (100%)	0.17	4 (8%) 11 13	214, 295, 334, 336	0
7	2	49/49 (100%)	0.00	2 (4%) 37 34	186, 272, 334, 388	0
7	5	49/49 (100%)	-0.10	0 100 100	174, 257, 298, 371	0
7	8	49/49 (100%)	-0.14	0 100 100	229, 300, 347, 427	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
8	3	3/4 (75%)	0.07	0	100	100	310, 310, 316, 331	0
8	6	3/4 (75%)	0.58	0	100	100	242, 242, 251, 272	0
8	9	3/4 (75%)	0.69	0	100	100	257, 257, 296, 373	0
All	All	11547/12159 (94%)	0.10	668 (5%)	23	23	50, 219, 383, 589	0

The worst 5 of 668 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	R	326	TRP	12.3
5	R	325	PRO	10.7
3	D	1054	THR	9.2
3	P	153	ASN	8.5
3	P	959	LYS	8.4

6.2 Non-standard residues in protein, DNA, RNA chains (i)

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

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 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	ZN	J	1501	1/1	0.72	0.13	218,218,218,218	0
10	ZN	D	1501	1/1	0.83	0.11	215,215,215,215	0
10	ZN	P	1502	1/1	0.90	0.15	204,204,204,204	0
9	MG	C	1401	1/1	0.91	0.33	218,218,218,218	0
10	ZN	P	1501	1/1	0.96	0.08	252,252,252,252	0
10	ZN	J	1502	1/1	0.97	0.17	196,196,196,196	0
9	MG	J	1503	1/1	0.97	0.19	204,204,204,204	0
10	ZN	D	1502	1/1	0.97	0.21	172,172,172,172	0
9	MG	P	1503	1/1	0.98	0.15	187,187,187,187	0

6.5 Other polymers (i)

There are no such residues in this entry.