# wwPDB X-ray Structure Validation Summary Report (i)

May 29, 2020 – 07:54 am BST

| | | |
|---|---|---|
| PDB ID | : | 4YLN |
| Title | : | E. coli Transcription Initiation Complex - 17-bp spacer and 4-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 following versions of software and data (see references (1)) were used in the production of this report:

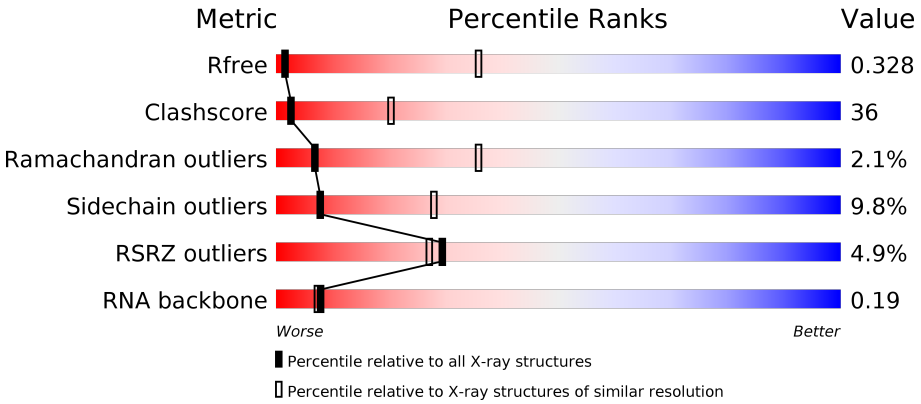
| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Mogul | : | 1.8.5 (274361), CSD as541be (2020) |
| Xtriage (Phenix) | : | 1.13 |
| EDS | : | 2.11 |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Refmac | : | 5.8.0158 |
| CCP4 | : | 7.0.044 (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 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.



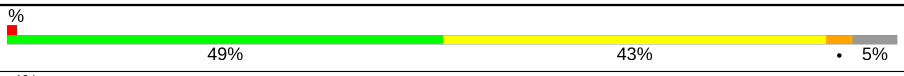
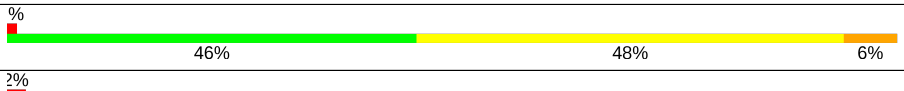
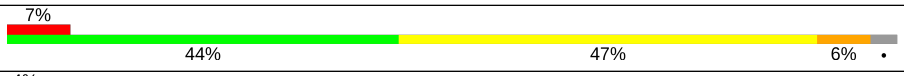
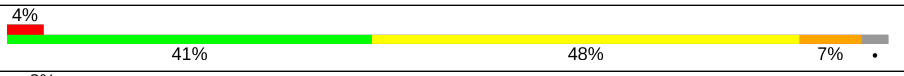
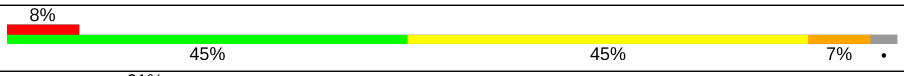
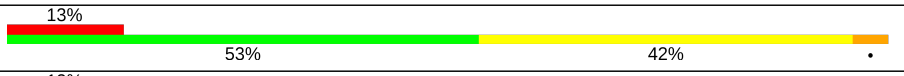
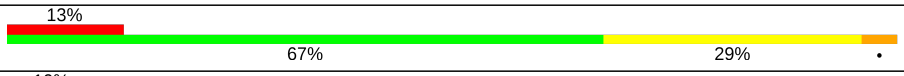
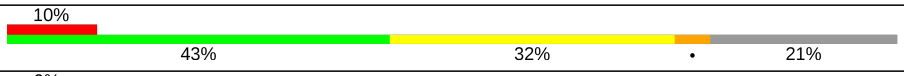
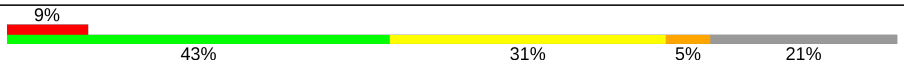
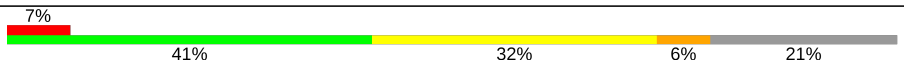
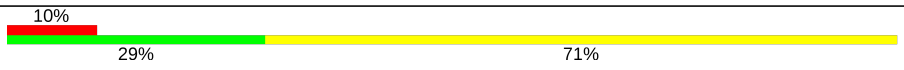
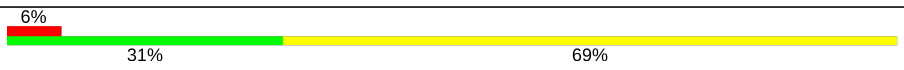
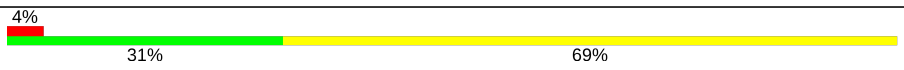
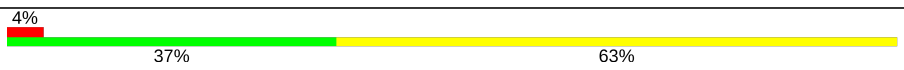
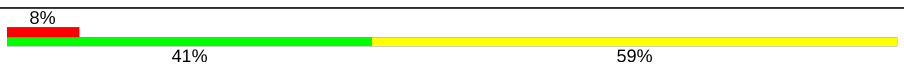
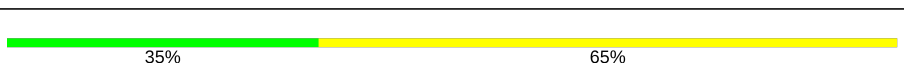
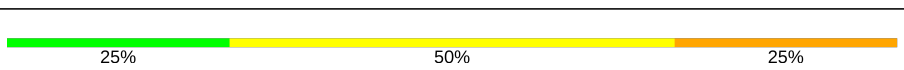
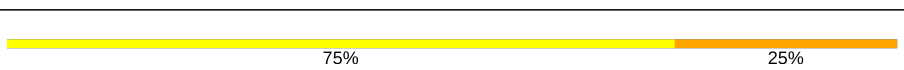

| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|---|---|---|
| $R_{free}$ | 130704 | 1019 (7.12-3.82) |
| Clashscore | 141614 | 1010 (7.10-3.90) |
| Ramachandran outliers | 138981 | 1014 (7.12-3.82) |
| Sidechain outliers | 138945 | 1191 (7.20-3.80) |
| RSRZ outliers | 127900 | 1023 (7.08-3.76) |
| RNA backbone | 3102 | 1074 (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 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 | 38% 51% 5% 5% |
| 1 | B | 242 | 40% 47% 7% 6% |
| 1 | G | 242 | 2% 48% 41% 6% 5% |
| 1 | H | 242 | 48% 38% 7% 6% |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 1 | M | 242 |   |
| 1 | N | 242 |   |
| 2 | C | 1342 |   |
| 2 | I | 1342 |   |
| 2 | O | 1342 |   |
| 3 | D | 1407 |   |
| 3 | J | 1407 |   |
| 3 | P | 1407 |   |
| 4 | E | 90 |   |
| 4 | K | 90 |   |
| 4 | Q | 90 |   |
| 5 | F | 628 |   |
| 5 | L | 628 |   |
| 5 | R | 628 |   |
| 6 | 1 | 49 |   |
| 6 | 4 | 49 |   |
| 6 | 7 | 49 |   |
| 7 | 2 | 49 |   |
| 7 | 5 | 49 |   |
| 7 | 8 | 49 |   |
| 8 | 3 | 4 |   |
| 8 | 6 | 4 |   |
| 8 | 9 | 4 |   |

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 | J | 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 ZINC ION (three-letter code: ZN) (formula: Zn).

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

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

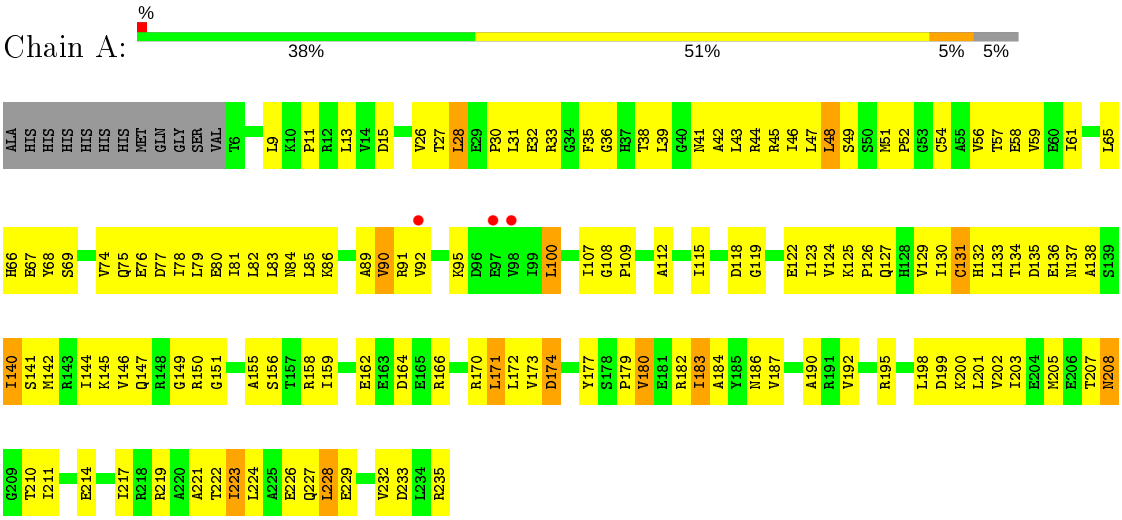
| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| 10 | P | 1 | Total | Mg | 0 | 0 |
| | | | 1 | 1 | | |
| 10 | D | 1 | Total | Mg | 0 | 0 |
| | | | 1 | 1 | | |
| 10 | 6 | 1 | Total | Mg | 0 | 0 |
| | | | 1 | 1 | | |

# 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.

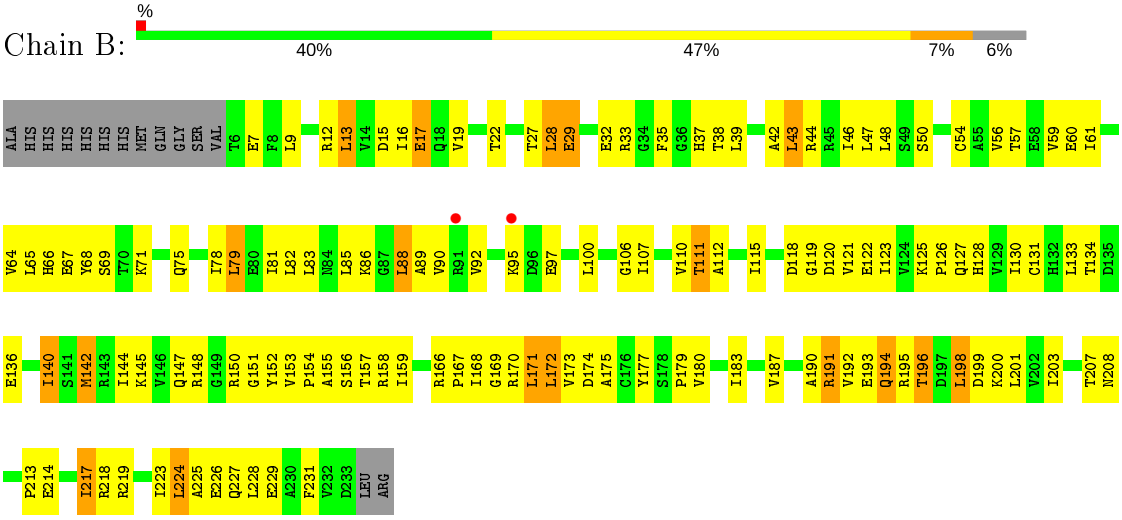
- Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain A:




- Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain B:

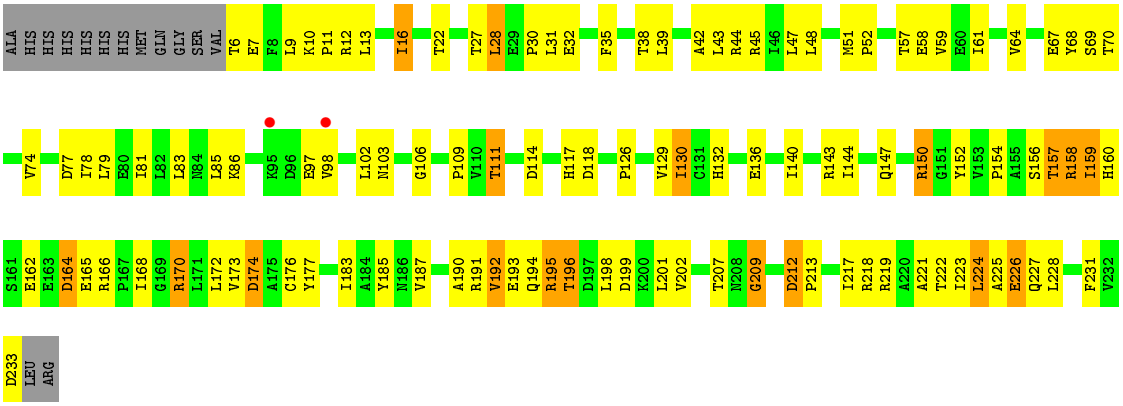



- Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain G:

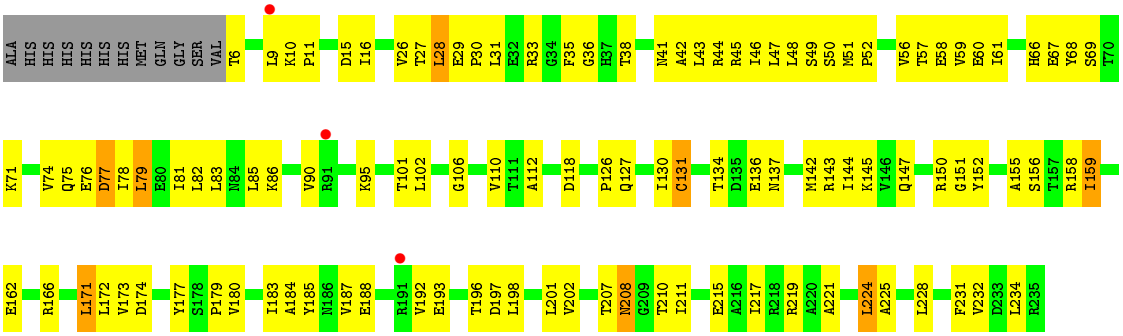
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain H:

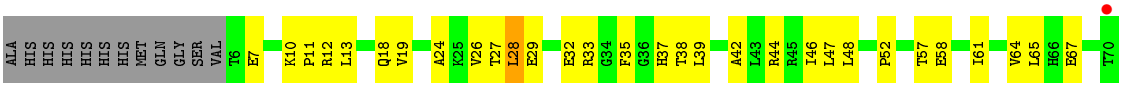
• Molecule 1: DNA-directed RNA polymerase subunit alpha

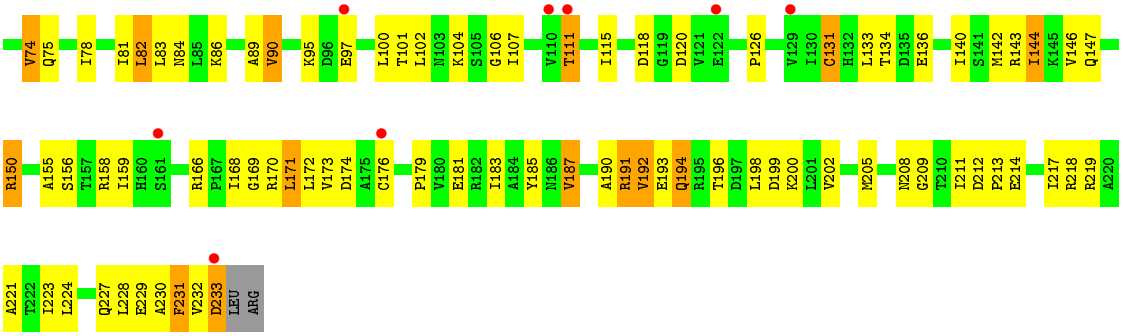
Chain M:

• Molecule 1: DNA-directed RNA polymerase subunit alpha

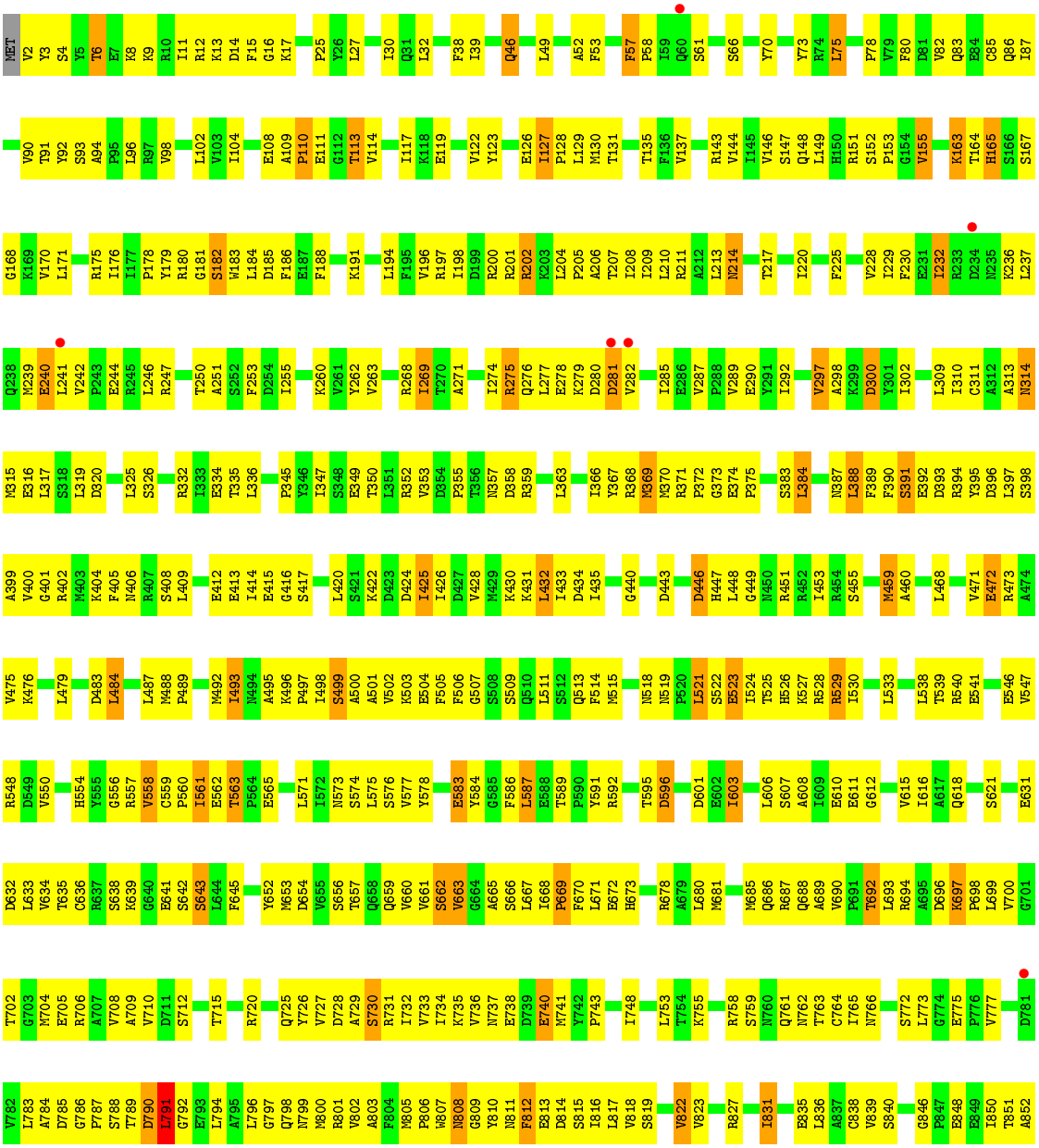
Chain N:

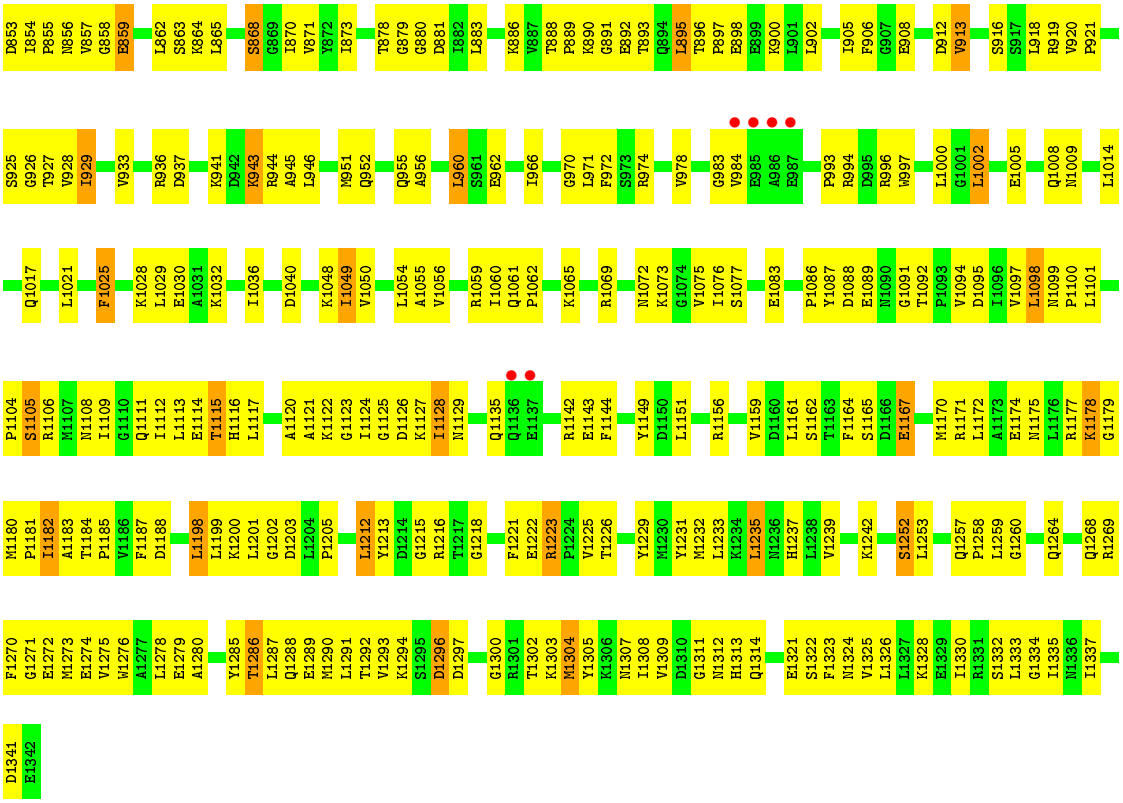






- Molecule 2: DNA-directed RNA polymerase subunit beta

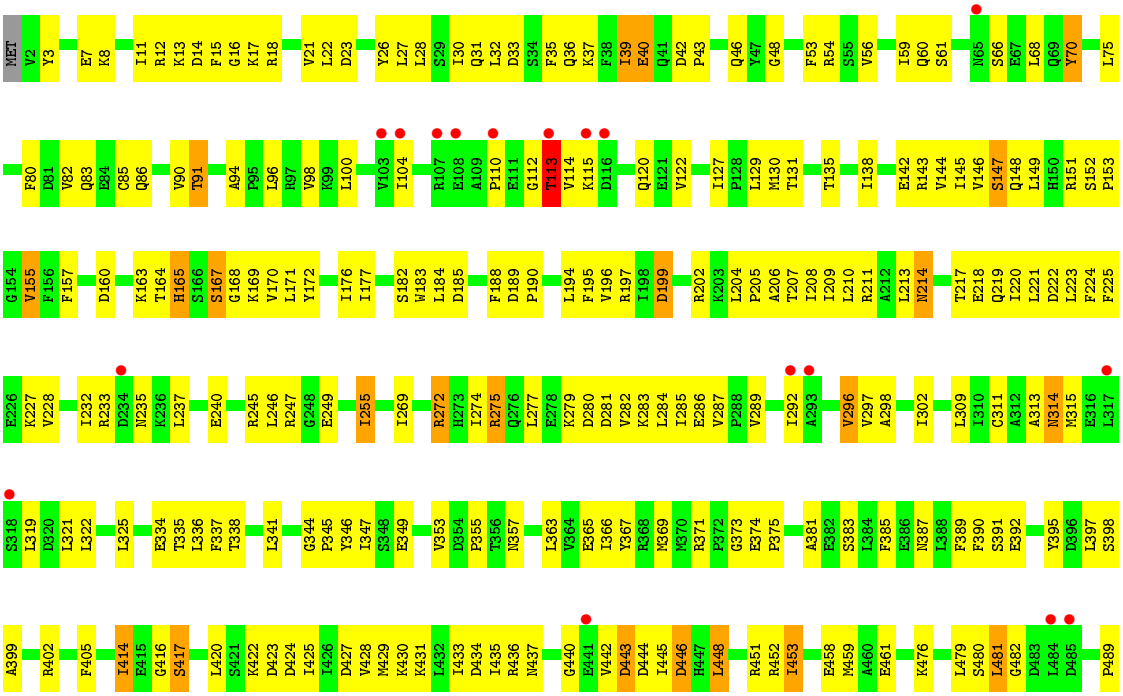
Chain C:

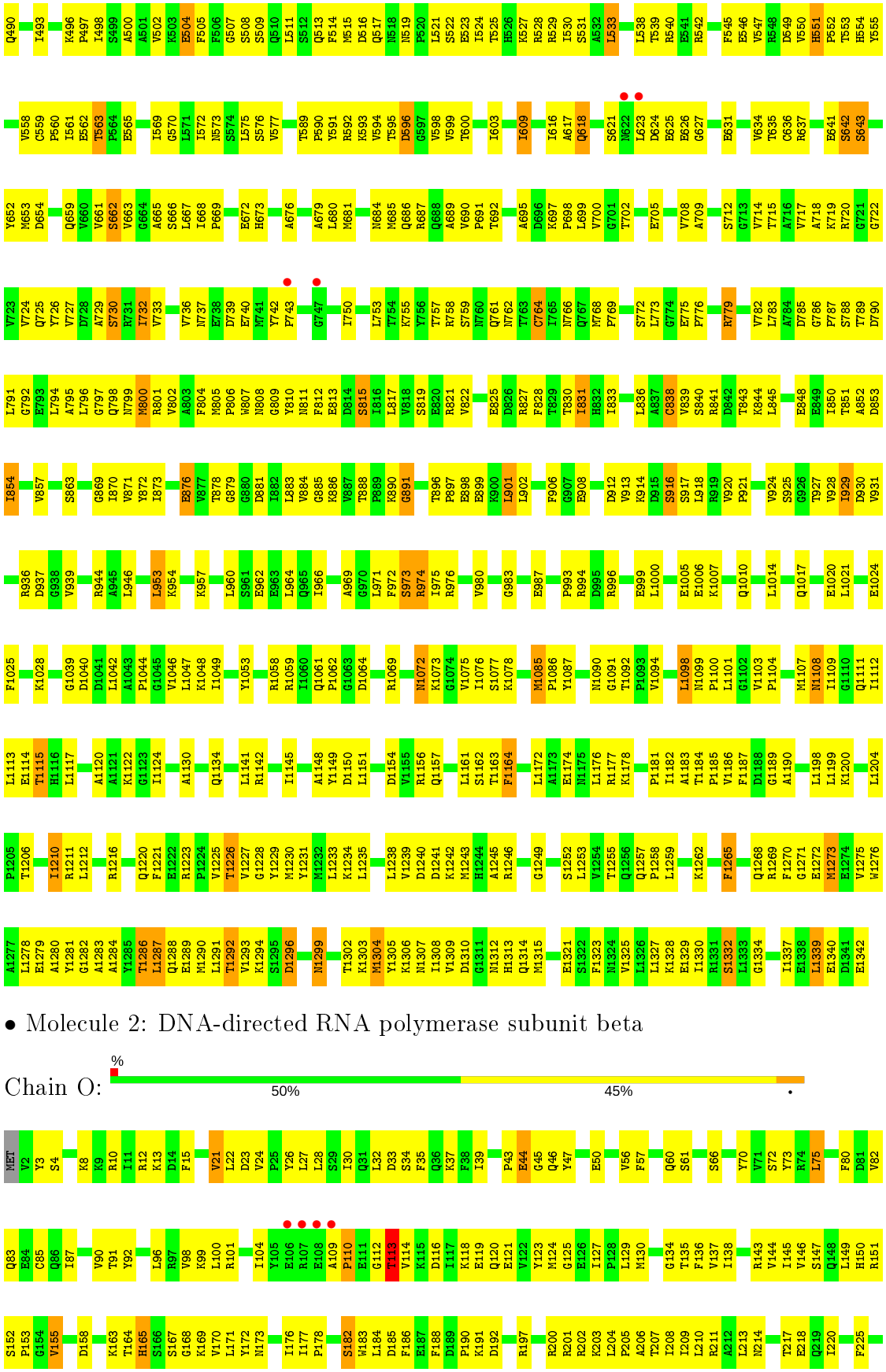





• Molecule 2: DNA-directed RNA polymerase subunit beta

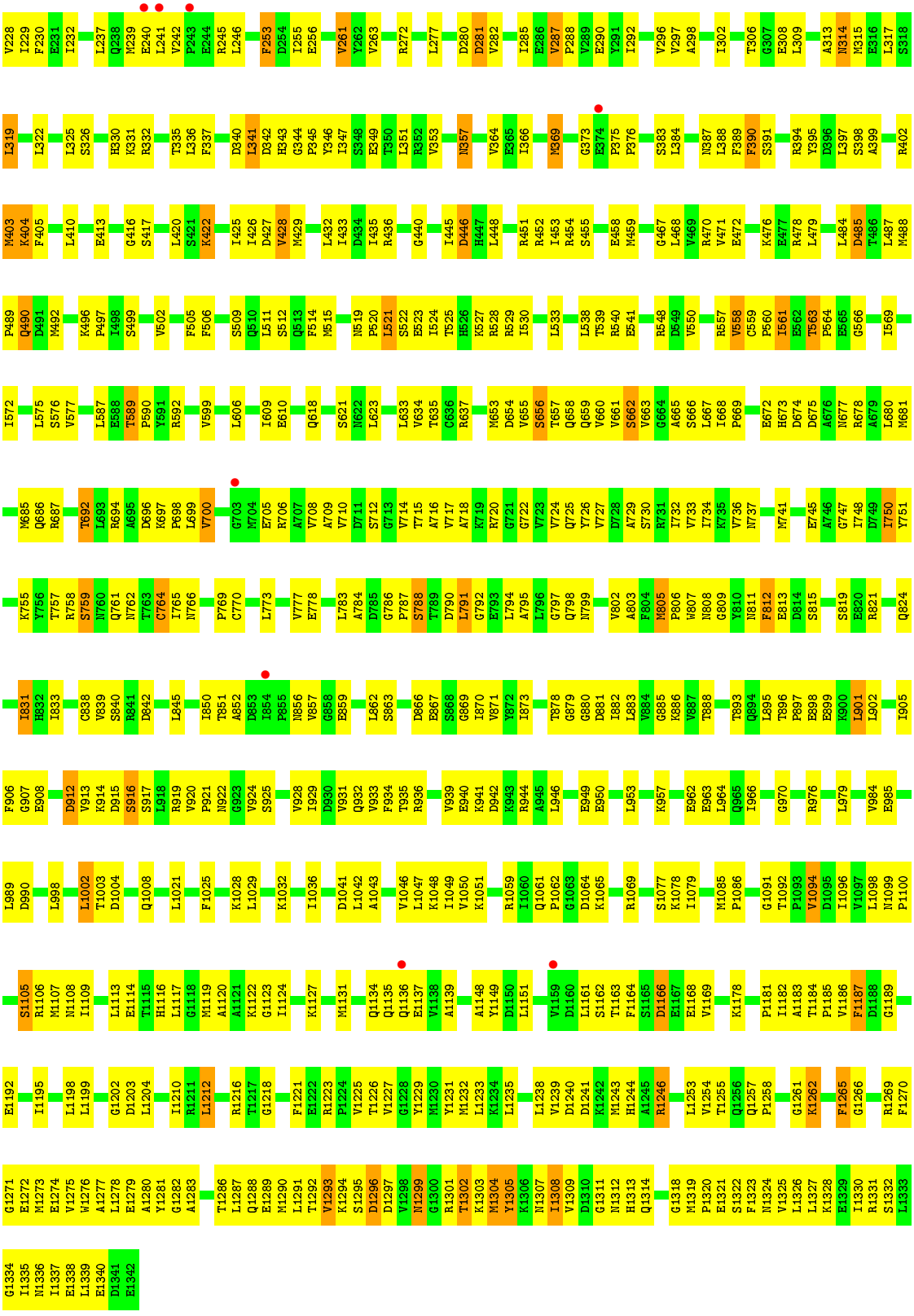
Chain I:



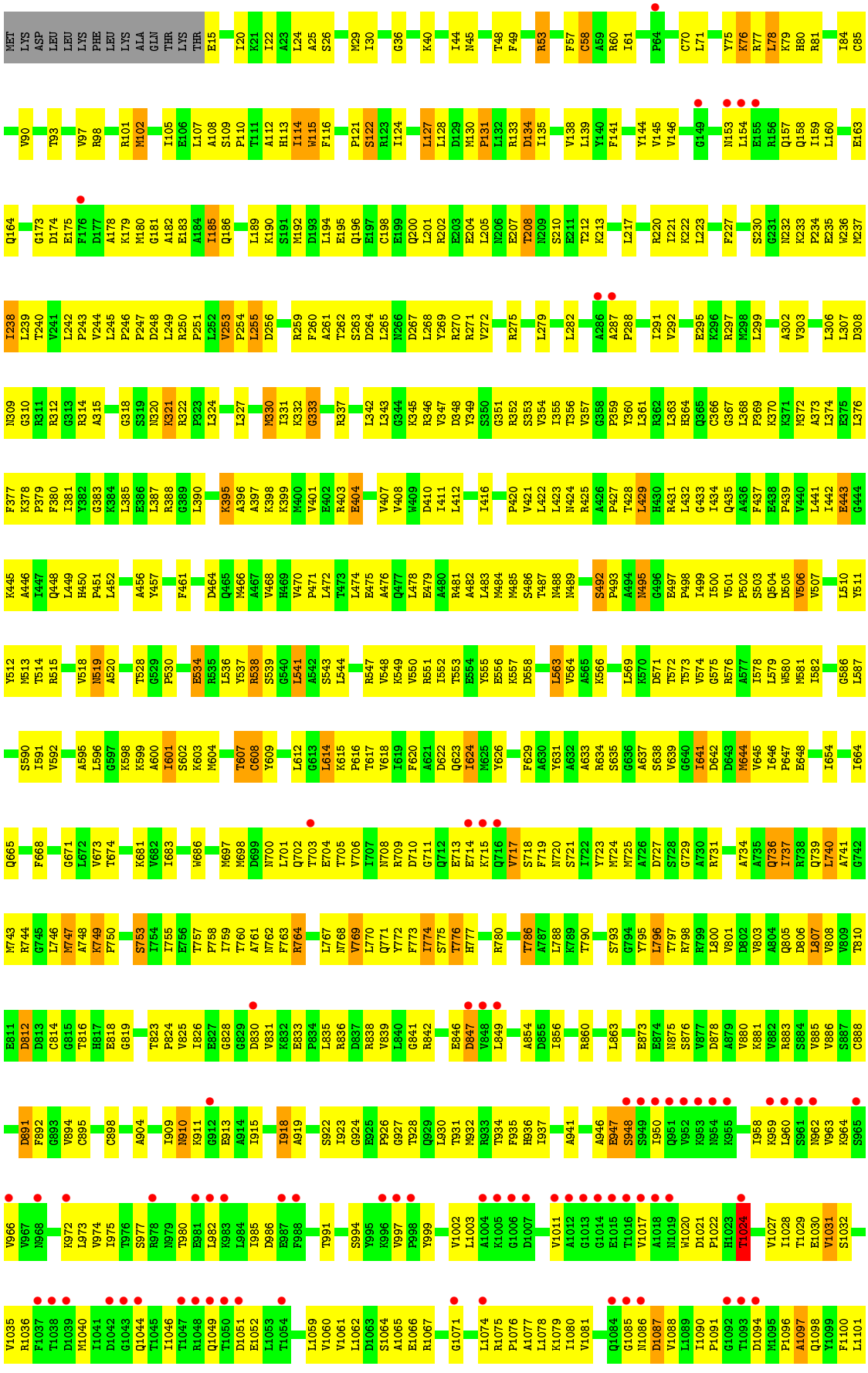
• Molecule 2: DNA-directed RNA polymerase subunit beta

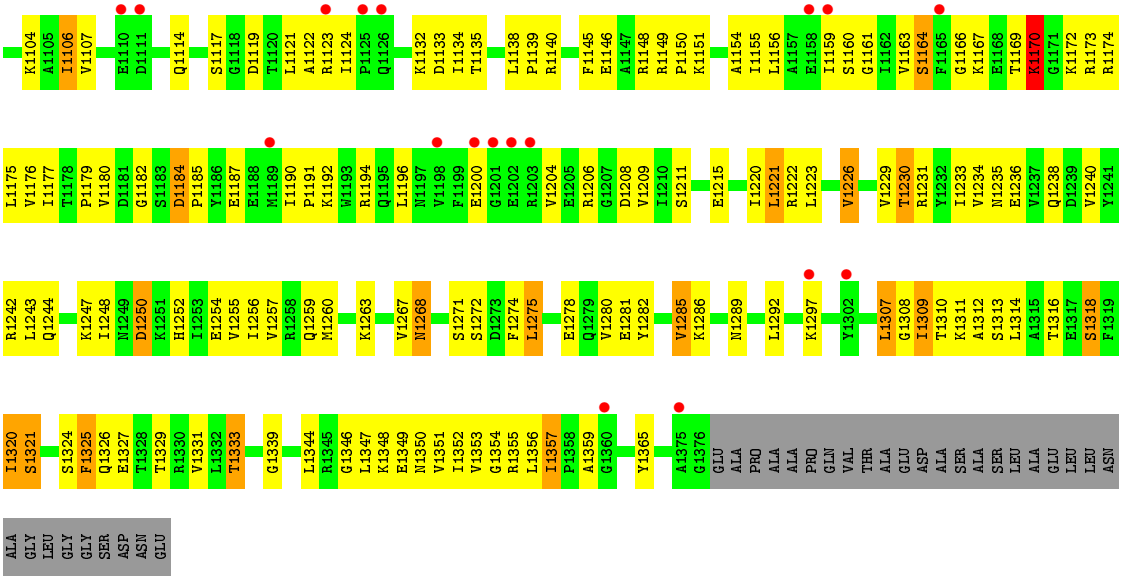
Chain O:



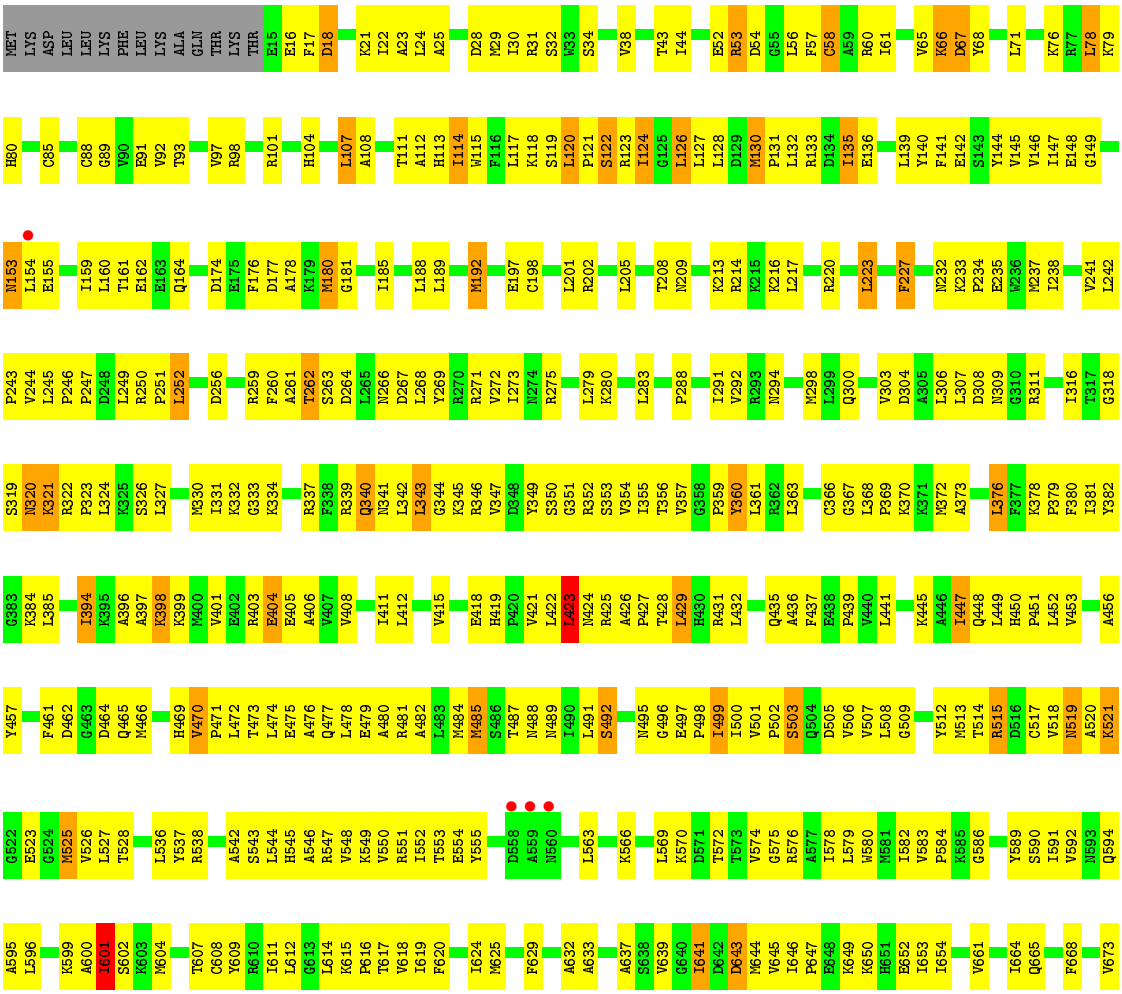
• Molecule 3: DNA-directed RNA polymerase subunit beta'

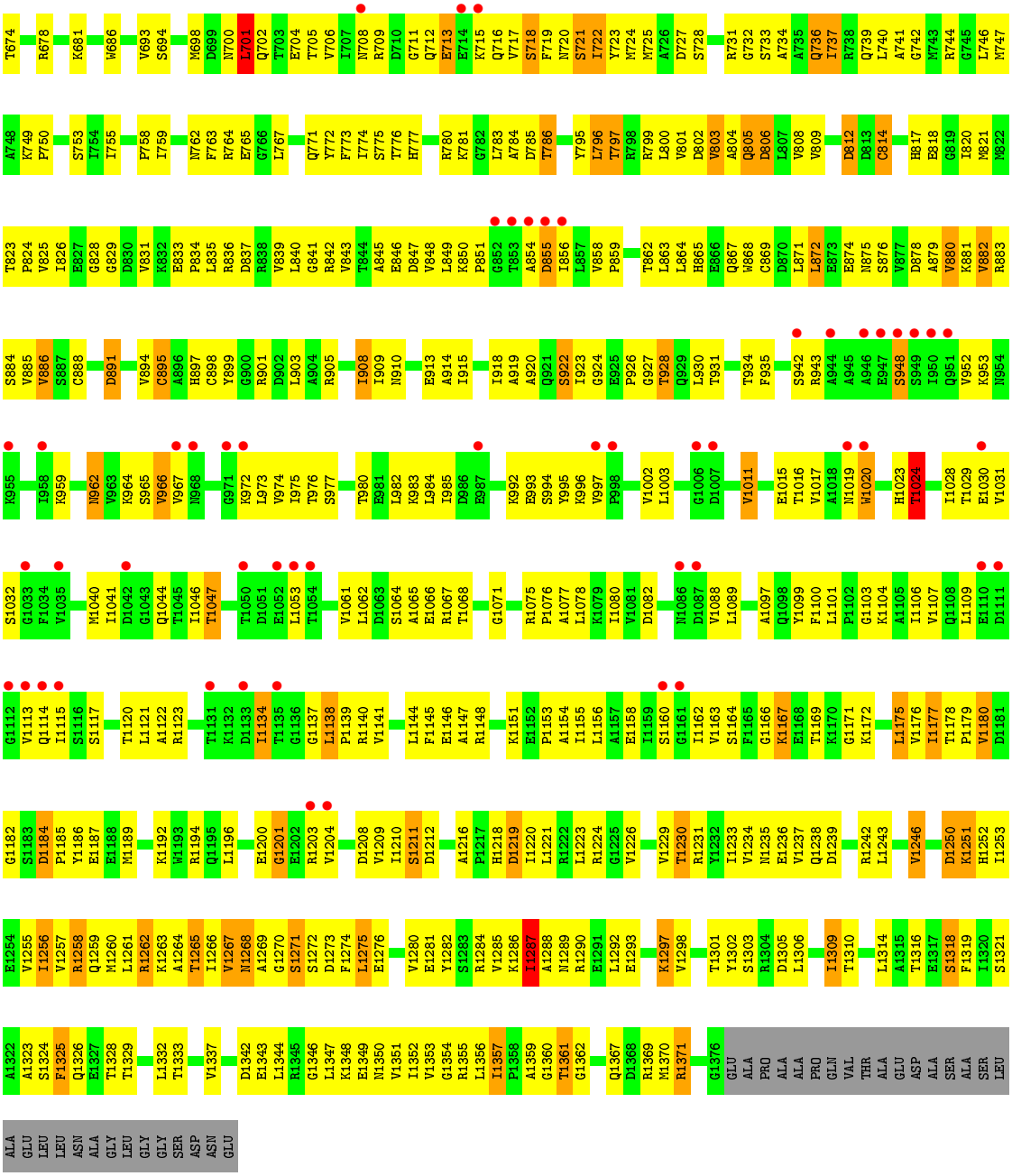
Chain D:


MET LYS ASP LEU LEU LYS PHE LEU LYS ALA GLN THR LYS THR E15 I20 K21 I22 A23 L24 A25 S26 M29 I30 G36 K40 I44 N45 T48 F49 R53 F57 C58 A59 R60 I61 P64 C70 L71 Y75 K76 R77 L78 K79 H80 R81 I84 C85
V90 T93 V97 R98 R101 M102 I105 E106 L107 A108 S109 P110 T111 A112 H113 I114 W115 F116 P121 S122 R123 I124 L127 L128 D129 M130 P131 L132 R133 D134 I135 V138 L139 Y140 F141 Y144 V145 V146 G149 N153 L154 E155 R156 Q157 Q158 I159 L160 E163
Q164 G173 D174 E175 F176 D177 A178 K179 M180 G181 A182 E183 A184 I185 Q186 L189 K190 S191 M192 D193 L194 E195 Q196 E197 C198 E199 Q200 L201 R202 E203 E204 L205 N206 E207 T208 N209 S210 E211 T212 K213 L217 R220 I221 K222 L223 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 F260 A261 T262 S263 D264 L265 N266 D267 L268 Y269 R270 R271 V272 R275 L279 L282 A286 A287 P288 I291 V292 E295 K296 R297 N298 L299 A302 V303 L306 L307 D308
N309 G310 R311 R312 G313 R314 A315 G318 S319 N320 K321 R322 P323 L324 I327 M330 I331 K332 G333 R337 I342 L343 G344 K345 R346 V347 D348 Y349 S350 G351 R352 S353 V354 I355 T356 V357 G358 P359 Y360 L361 R362 I363 H364 Q365 C366 G367 L368 P369 K370 K371 M372 A373 L374 E375 L376
F377 K378 P379 F380 I381 Y382 G383 K384 L385 E386 L387 R388 G389 L390 K395 A396 A397 K398 K399 N400 V401 E402 R403 E404 V407 V408 W409 D410 I411 L412 I416 P420 V421 L422 L423 N424 R425 A426 P427 T428 L429 H430 R431 L432 G433 I434 Q435 A436 F437 E438 P439 V440 L441 I442 E443 G444
K445 A446 Y447 Q448 L449 H450 P451 L452 A456 Y457 F461 D464 Q465 N466 A467 V468 H469 V470 P471 L472 T473 L474 E475 A476 Q477 L478 E479 A480 R481 A482 L483 M484 M485 S486 T487 N488 N489 S492 P493 A494 N495 G496 E497 P498 I499 I500 V501 P502 S503 Q504 D505 V506 V507 L510 Y511
Y512 M513 T514 R515 V518 N519 A520 T528 G529 P530 E534 R535 L536 Y537 R538 S539 G540 L541 A542 S543 L544 R547 V548 K549 V550 R551 I552 T553 E554 Y555 E556 K557 D558 L563 V564 A565 K566 L569 K570 D571 T572 T573 Y574 G575 R576 A577 I578 L579 W580 M581 I582 G586 L587
S590 I591 V592 A595 L596 G597 K598 K599 A600 I601 S602 K603 M604 T607 C608 Y609 L612 G613 L614 K615 P616 T617 V618 I619 F620 A621 D622 Q623 I624 K625 Y626 F629 A630 Y631 A632 A633 R634 S635 G636 A637 S638 D639 G640 I641 D642 D643 K644 V645 I646 P647 E648 I654 I664
Q665 F668 G671 L672 V673 T674 K681 V682 I683 W686 M697 M698 D699 N700 L701 Q702 T703 E704 T705 V706 I707 N708 R709 D710 G711 Q712 E713 E714 K715 Q716 V717 S718 F719 N720 S721 I722 Y723 N724 N725 A726 D727 S728 G729 A730 R731 A734 A735 Q736 I737 R738 Q739 L740 A741 G742
W743 R744 G745 L746 M747 A748 W749 P750 S753 I754 I755 E756 T757 P758 I759 T760 A761 N762 F763 R764 L767 N768 V769 L770 Q771 V772 F773 I774 S775 T776 H777 R780 T786 A787 L788 K789 T790 S793 G794 V795 L796 T797 R798 R799 L800 V801 D802 V803 A804 Q805 D806 L807 V808 V809 T810
E811 D812 D813 C814 G815 T816 H817 E818 G819 T823 P824 V825 I826 E827 G828 G829 D830 V831 K832 E833 P834 L835 R836 D837 R838 V839 L840 G841 R842 E846 D847 V848 L849 A854 D855 I856 R860 L863 E873 E874 N875 S876 V877 D878 A879 V880 K881 V882 R883 S884 V885 V886 S887 C888
D891 F892 G893 V894 C895 C898 A904 I909 N910 K911 G912 E913 A914 I915 I918 A919 S922 I923 G924 E925 P926 G927 T928 Q929 L930 T931 M932 R933 T934 F935 H936 I937 A941 A946 E947 S948 S949 I950 Q951 V952 K953 N954 K955 I958 K959 L960 S961 N962 V963 K964 S965
V966 V967 N968 K972 L973 V974 I975 T976 S977 R978 N979 T980 E981 L982 K983 I984 I985 D986 E987 F988 T991 S994 Y995 K996 V997 P998 Y999 V1002 L1003 A1004 K1005 G1006 D1007 V1011 A1012 G1013 G1014 E1015 T1016 V1017 A1018 N1019 W1020 D1021 P1022 H1023 T1024 V1027 I1028 T1029 E1030 V1031 S1032
V1035 R1036 F1037 T1038 D1039 M1040 I1041 D1042 G1043 Q1044 T1045 I1046 T1047 R1048 Q1049 T1050 D1051 E1052 L1053 T1054 L1059 V1060 V1061 L1062 D1063 S1064 A1065 E1066 R1067 G1071 L1074 R1075 P1076 A1077 L1078 K1079 I1080 V1081 Q1084 G1085 N1086 D1087 V1088 L1089 I1090 P1091 G1092 T1093 D1094 M1095 P1096 A1097 Q1098 Y1099 F1100 L1101



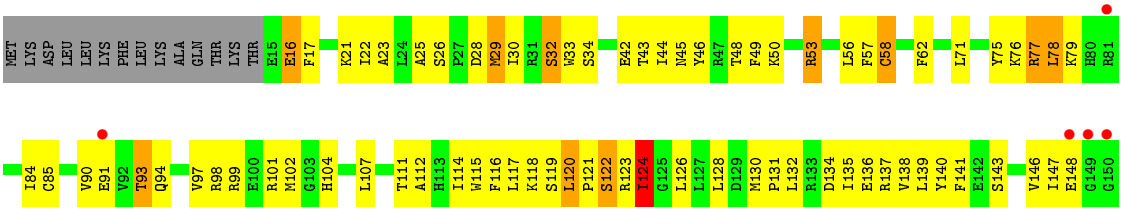
• Molecule 3: DNA-directed RNA polymerase subunit beta'

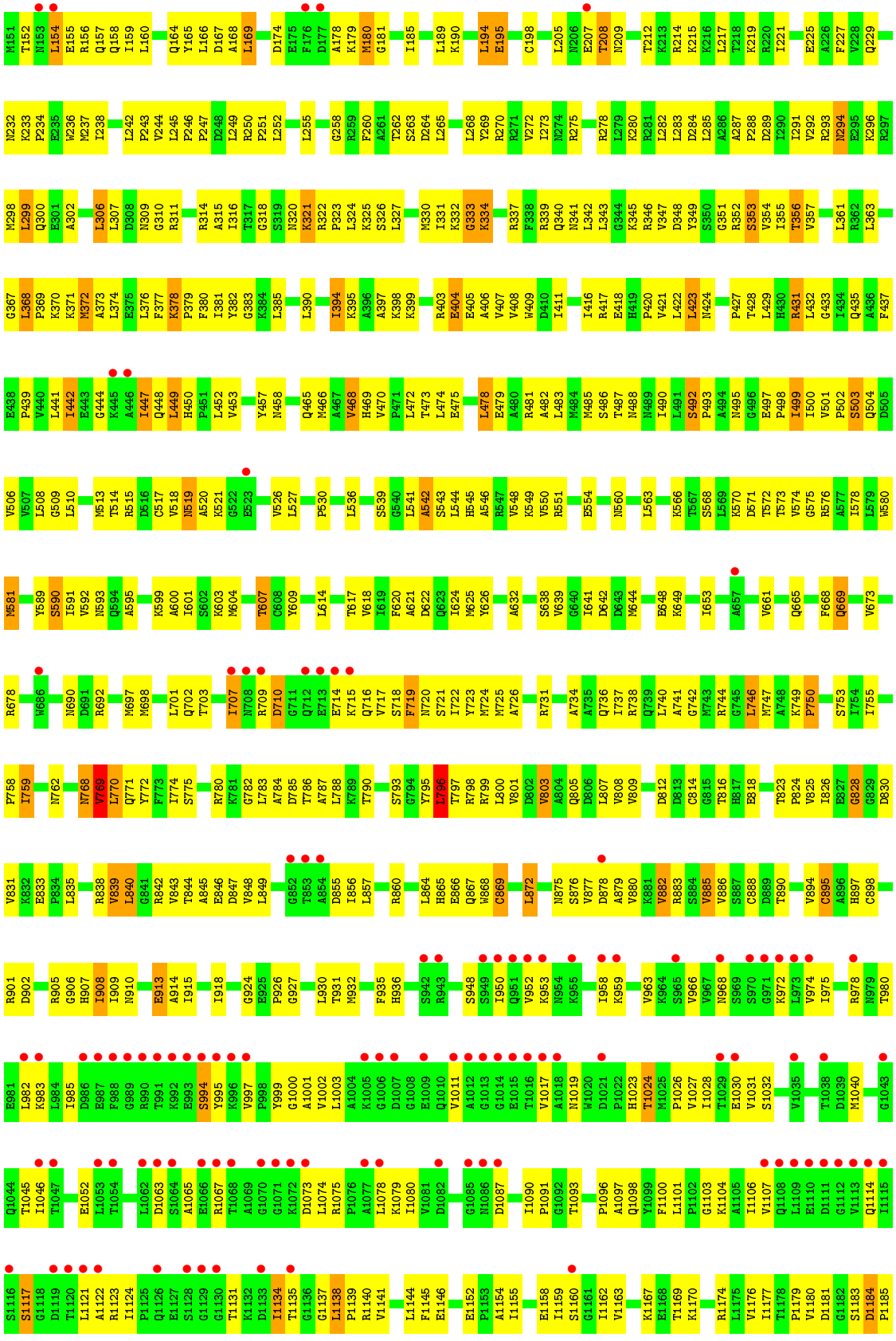



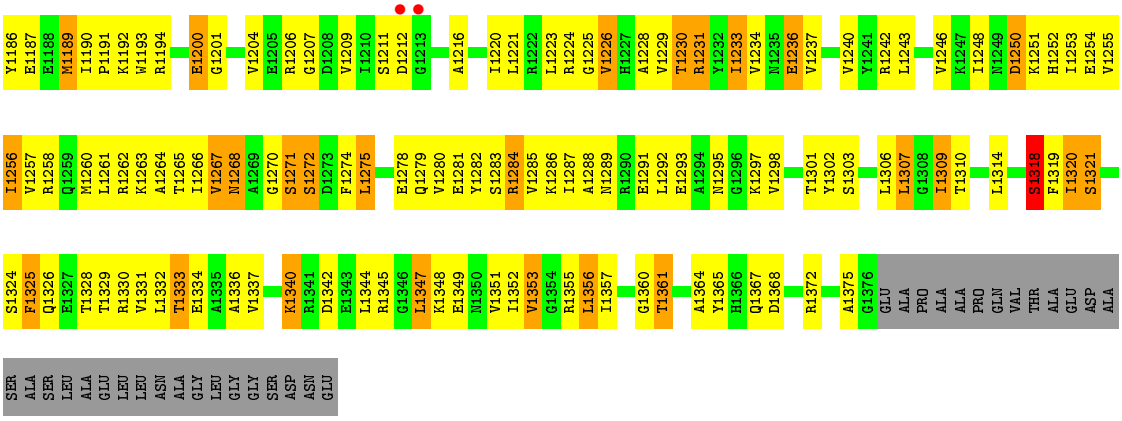


• Molecule 3: DNA-directed RNA polymerase subunit beta'

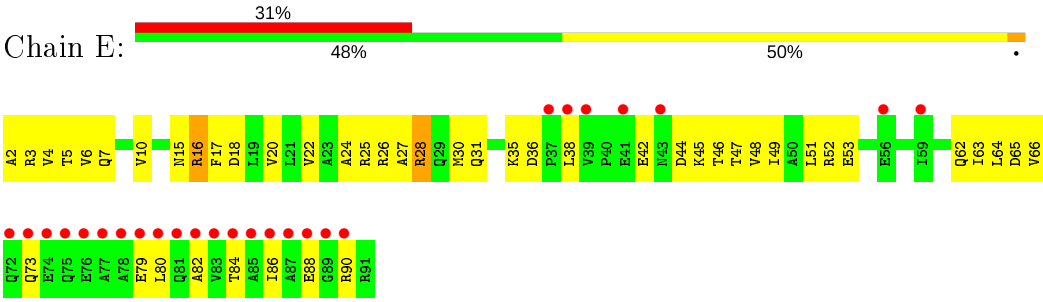
Chain P:

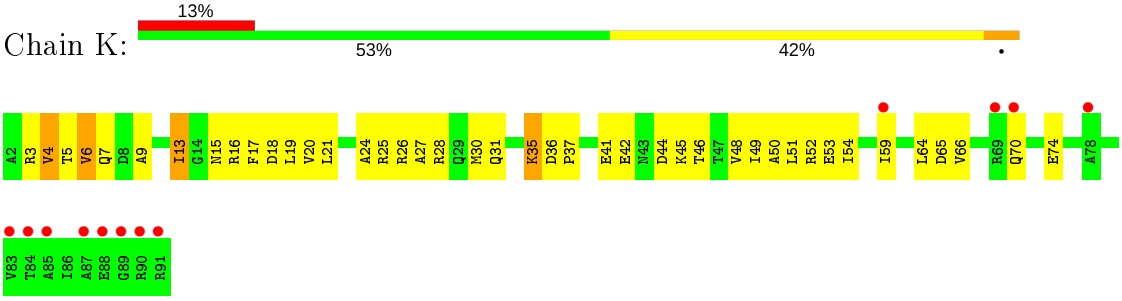

M151 T152 N153 L154 E155 R156 Q157 Q158 I159 L160 Q164 Y165 L166 D167 A168 L169 D174 E175 F176 D177 A178 K179 M180 G181 I185 L189 K190 L194 E195 C198 L205 N206 E207 T208 N209 T212 K213 R214 K215 K216 L217 T218 K219 R220 I221 E225 A226 F227 V228 Q229
N232 K233 P234 E235 W236 M237 I238 L242 P243 V244 L245 P246 P247 D248 L249 R250 P251 L252 L255 G258 R259 F260 A261 T262 S263 D264 L265 L268 Y269 R270 R271 V272 I273 N274 R275 R278 L279 K280 R281 L282 L283 D284 L285 A286 A287 P288 D289 I290 I291 V292 R293 N294 E295 K296 R297
M298 L299 Q300 E301 A302 L306 L307 D308 N309 G310 R311 R314 A315 I316 T317 G318 S319 N320 K321 R322 P323 L324 K325 S326 L327 M330 I331 K332 G333 K334 R337 F338 R339 Q340 N341 L342 L343 G344 K345 R346 V347 D348 Y349 S350 G351 R352 S353 V354 I355 T356 V357 L361 R362 L363
G367 L368 P369 K370 K371 M372 A373 L374 E375 L376 F377 K378 P379 F380 I381 Y382 G383 K384 L385 L390 I394 K395 A396 A397 K398 K399 R403 E404 E405 A406 V407 V408 W409 D410 I411 I416 R417 E418 H419 P420 V421 L422 L423 N424 P427 T428 L429 H430 R431 L432 G433 I434 Q435 A436 F437
E438 P439 V440 L441 I442 E443 G444 K445 A446 I447 Q448 L449 H450 P451 L452 V453 Y457 N458 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 P493 A494 N495 G496 E497 P498 I499 I500 V501 P502 S503 Q504 Q505 D506
V506 W507 L508 G509 L510 M513 T514 R515 D516 C517 V518 N519 A520 K521 G522 E523 V526 L527 P530 L536 S539 G540 L541 A542 S543 L544 H545 A546 R547 V548 K549 V550 R551 E554 N560 L563 K566 T567 S568 L569 K570 D571 T572 T573 V574 G575 R576 A577 I578 L579 W580
M581 Y589 S590 I591 V592 N593 Q594 A595 K599 A600 I601 S602 K603 M604 T607 C608 Y609 L614 T617 V618 I619 F620 A621 D622 G623 I624 M625 Y626 A632 S638 V639 G640 I641 D642 D643 M644 E648 K649 I653 A657 V661 Q665 F668 Q669 V673
R678 W686 N690 D691 R692 M697 M698 L701 Q702 T703 T707 N708 R709 D710 G711 Q712 E713 E714 K715 Q716 V717 S718 F719 N720 S721 I722 Y723 M724 M725 A726 R731 A734 A735 Q736 I737 R738 Q739 L740 A741 G742 M743 R744 G745 L746 M747 A748 K749 P750 S753 I754 I755
P758 I759 N762 M768 V769 L770 Q771 V772 F773 I774 S775 R780 K781 G782 L783 A784 D785 T786 A787 L788 K789 T790 S793 G794 Y795 L796 T797 R798 R799 L800 V801 D802 V803 A804 Q805 D806 L807 V808 Y809 D812 D813 C814 G815 T816 H817 E818 T823 P824 Y825 I826 E827 G828 G829 D830
V831 K832 E833 P834 L835 R838 V839 L840 G841 R842 V843 T844 A845 E846 D847 V848 L849 G852 T853 A854 D855 I856 L857 R860 L864 H865 E866 Q867 W868 C869 L872 N875 S876 V877 D878 A879 V880 K881 V882 R883 S884 V885 V886 S887 C888 D889 T890 V894 C895 A896 H897 C898
R901 D902 R905 G906 H907 I908 I909 N910 E913 A914 I915 I918 G924 E925 P926 G927 L930 T931 M932 F935 H936 S942 R943 S948 S949 I950 Q951 V952 K953 N954 K955 I958 K959 V963 K964 S965 V966 V967 N968 S969 S970 G971 K972 L973 V974 I975 R978 N979 T980
E981 L982 K983 L984 I985 D986 E987 F988 G989 R990 T991 K992 E993 S994 Y995 K996 V997 P998 Y999 G1000 A1001 V1002 L1003 A1004 K1005 G1006 D1007 G1008 E1009 Q1010 V1011 A1012 G1013 G1014 E1015 T1016 V1017 A1018 N1019 W1020 D1021 P1022 H1023 T1024 M1025 P1026 V1027 I1028 T1029 E1030 V1031 S1032 V1035 T1038 D1039 M1040 G1043
Q1044 T1045 I1046 T1047 E1052 L1053 T1054 L1062 D1063 S1064 A1065 E1066 R1067 T1068 A1069 A1070 G1071 K1072 D1073 L1074 R1075 P1076 A1077 L1078 K1079 I1080 V1081 D1082 G1085 N1086 D1087 I1090 P1091 G1092 T1093 P1096 A1097 Q1098 Y1099 F1100 L1101 P1102 G1103 K1104 A1105 I1106 V1107 Q1108 L1109 E1110 D1111 G1112 V1113 Q1114 I1115
S1116 S1117 G1118 D1119 T1120 L1121 A1122 R1123 I1124 P1125 Q1126 E1127 S1128 G1129 G1130 T1131 K1132 D1133 I1134 T1135 G1136 G1137 L1138 P1139 R1140 V1141 L1144 F1145 E1146 E1152 P1153 A1154 I1155 E1158 I1159 S1160 G1161 I1162 V1163 K1167 E1168 T1169 K1170 R1174 L1175 V1176 I1177 T1178 P1179 V1180 D1181 G1182 S1183 D1184 P1185



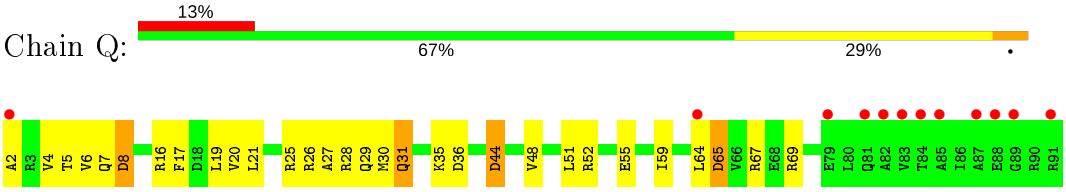

- Molecule 4: DNA-directed RNA polymerase subunit omega



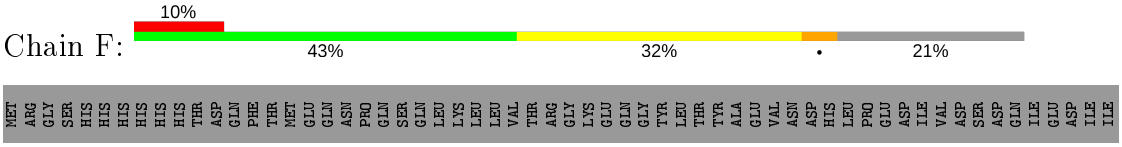

- Molecule 4: DNA-directed RNA polymerase subunit omega

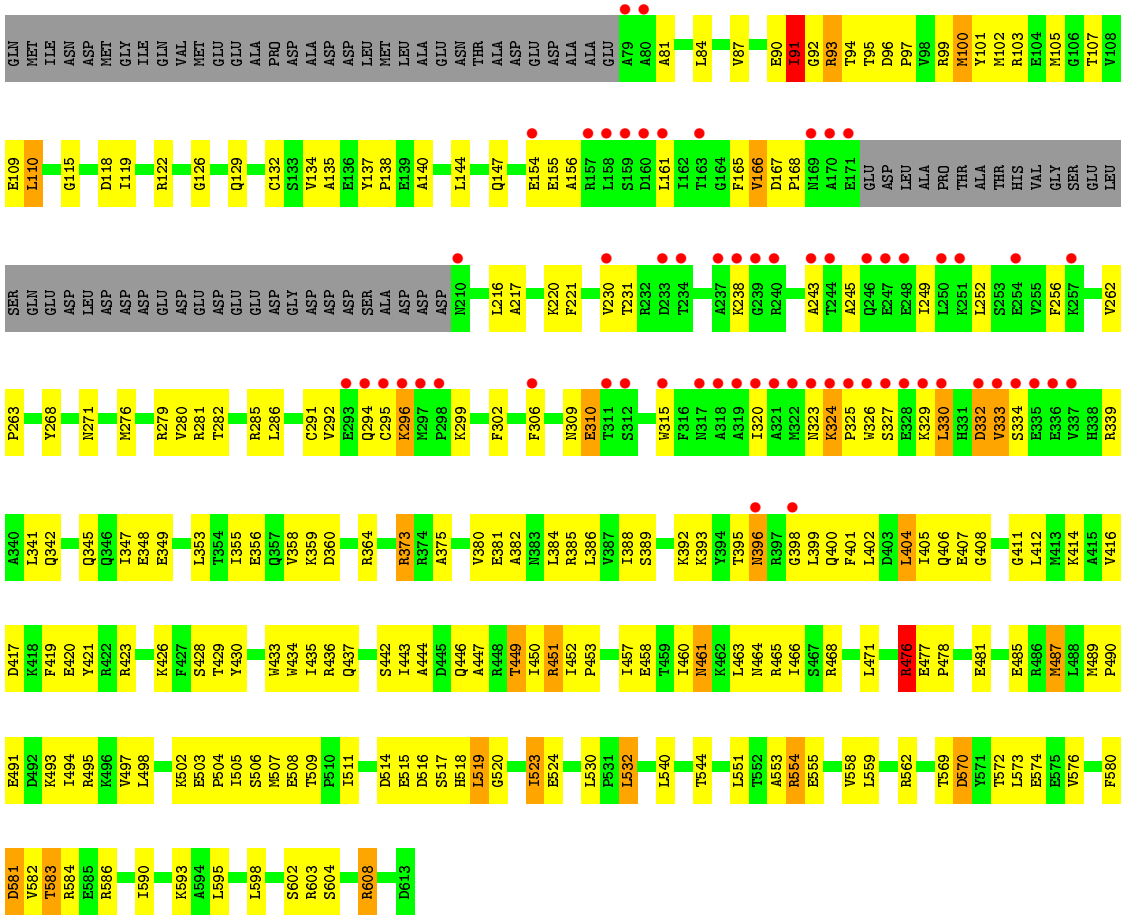



- Molecule 4: DNA-directed RNA polymerase subunit omega



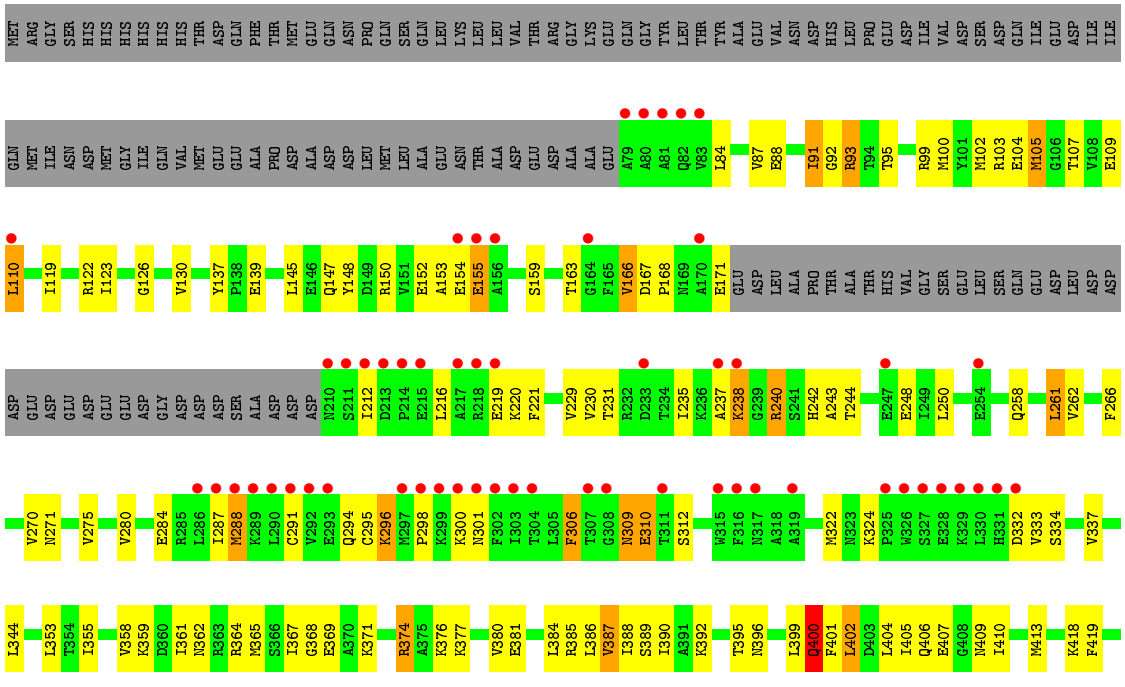

- Molecule 5: RNA polymerase sigma factor RpoD

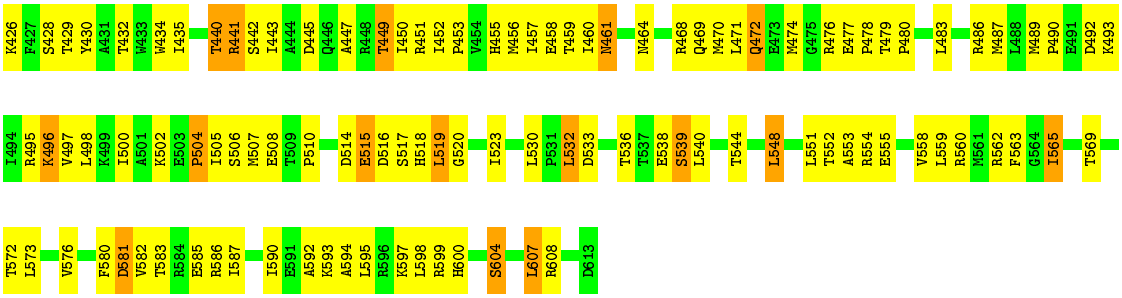


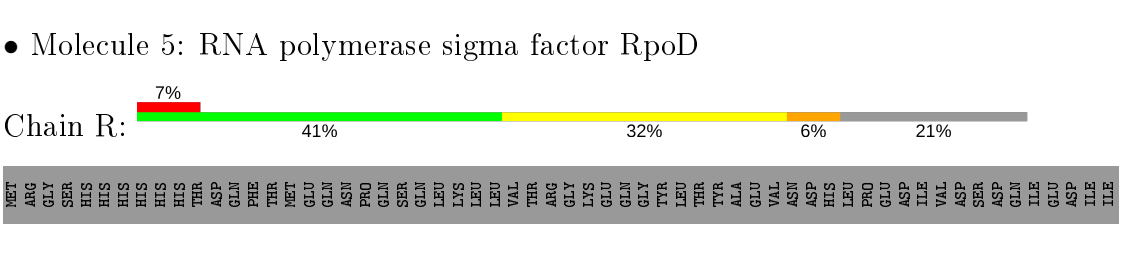


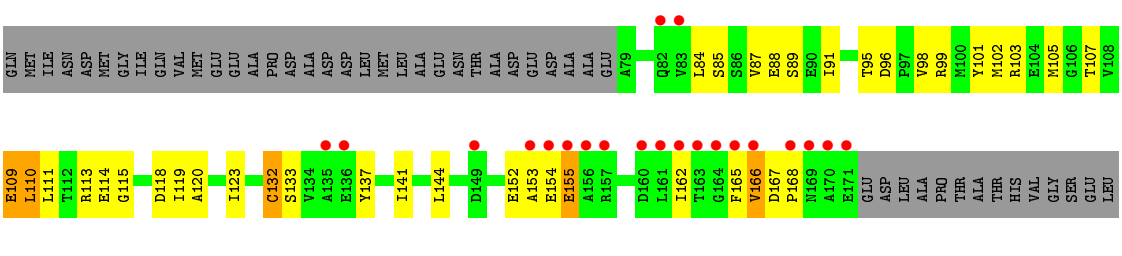
• Molecule 5: RNA polymerase sigma factor RpoD

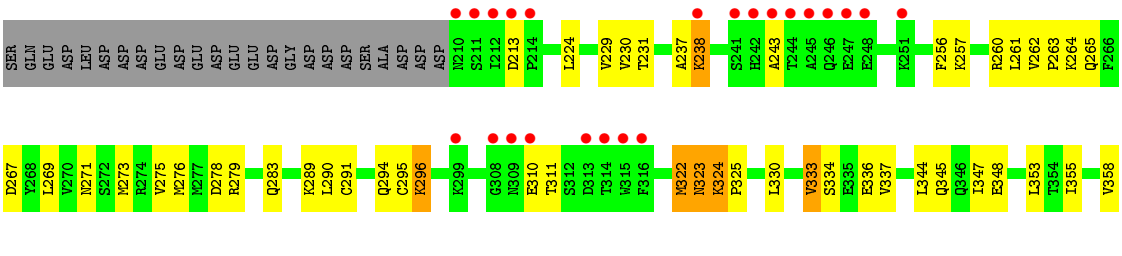
Chain L:

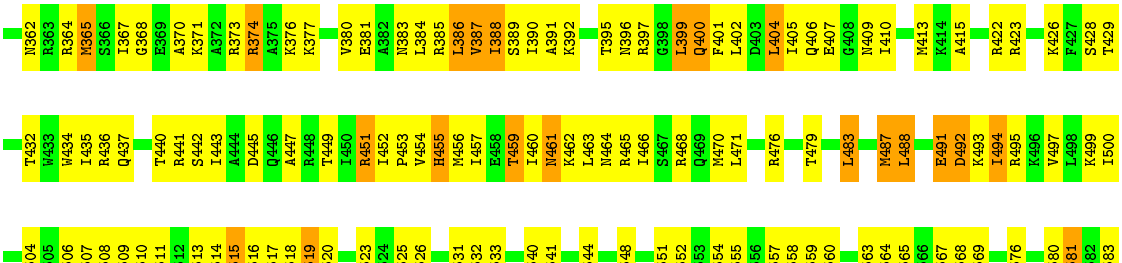



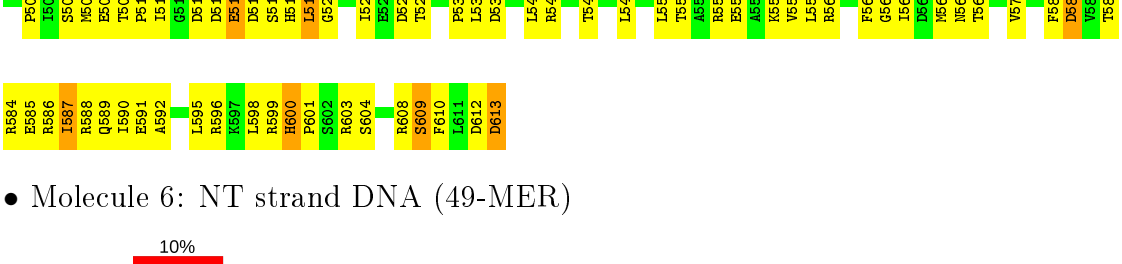


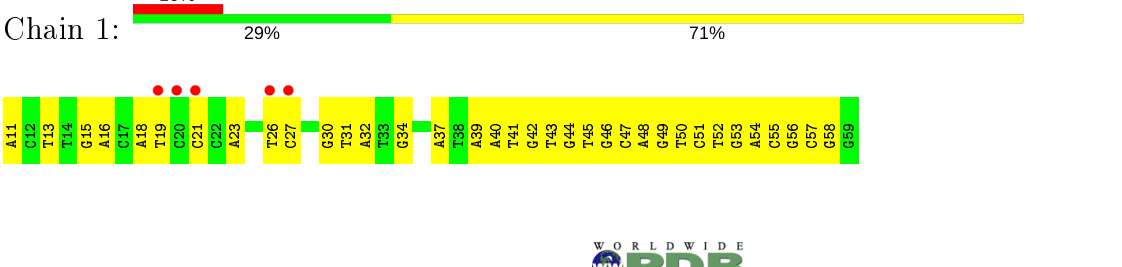














• Molecule 5: RNA polymerase sigma factor RpoD

Chain R:













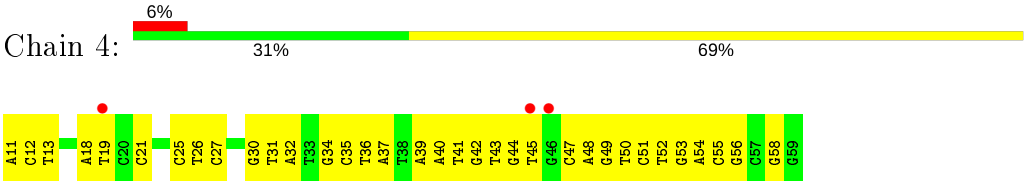



• Molecule 6: NT strand DNA (49-MER)

Chain 1:

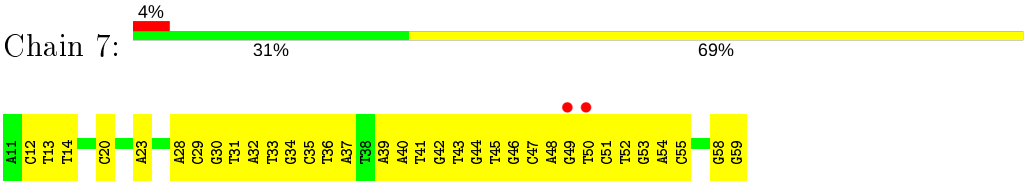
• Molecule 6: NT strand DNA (49-MER)

Chain 4:



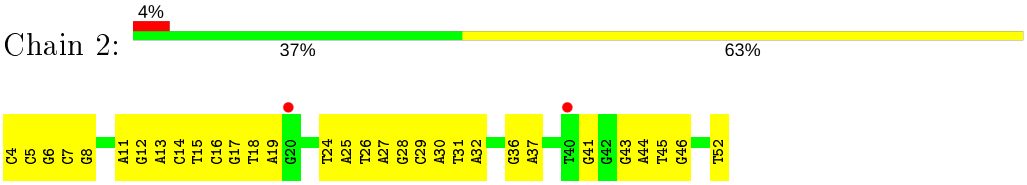

• Molecule 6: NT strand DNA (49-MER)

Chain 7:



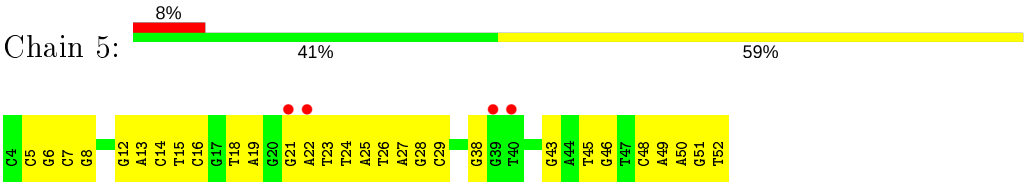

• Molecule 7: T strand DNA (49-MER)

Chain 2:



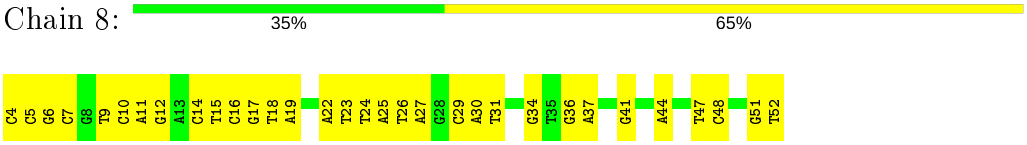

• Molecule 7: T strand DNA (49-MER)

Chain 5:



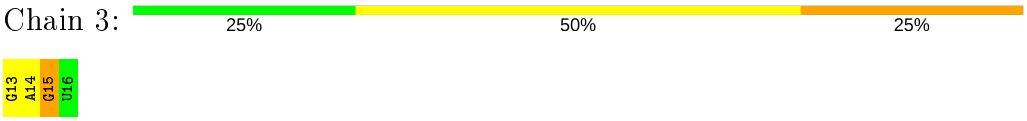

• Molecule 7: T strand DNA (49-MER)

Chain 8:




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

Chain 3:

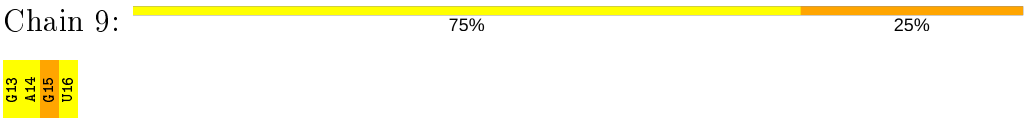



• Molecule 8: RNA (5'-D(*(GTP))-R(P*AP*GP*U)-3')

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<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 237.40Å 206.05Å 248.69Å<br>90.00° 116.55° 90.00° | Depositor |
| Resolution (Å) | 39.90 – 5.50<br>39.90 – 5.50 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 99.3 (39.90-5.50)<br>99.6 (39.90-5.50) | Depositor<br>EDS |
| $R_{merge}$ | (Not available) | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 1.20 (at 5.37Å) | Xtriage |
| Refinement program | REFMAC 5.8.0073 | Depositor |
| R, $R_{free}$ | 0.245 , 0.328<br>0.244 , 0.328 | Depositor<br>DCC |
| $R_{free}$ test set | 3459 reflections (4.98%) | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 268.1 | Xtriage |
| Anisotropy | 0.597 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.29 , 203.9 | EDS |
| L-test for twinning[2] | $< \|L\| > = 0.40$, $< L^2 > = 0.23$ | Xtriage |
| Estimated twinning fraction | 0.045 for h,-k,-h-l | Xtriage |
| $F_o$,$F_c$ correlation | 0.92 | EDS |
| Total number of atoms | 94608 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 219.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.48% 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 | 0.84 | 1/2450 (0.0%) |
| 1 | B | 0.54 | 0/1789 | 0.78 | 0/2425 |
| 1 | G | 0.56 | 0/1809 | 0.76 | 1/2450 (0.0%) |
| 1 | H | 0.53 | 0/1789 | 0.76 | 0/2425 |
| 1 | M | 0.53 | 0/1809 | 0.74 | 0/2450 |
| 1 | N | 0.54 | 0/1789 | 0.79 | 2/2425 (0.1%) |
| 2 | C | 0.54 | 0/10745 | 0.78 | 4/14499 (0.0%) |
| 2 | I | 0.54 | 3/10745 (0.0%) | 0.77 | 2/14499 (0.0%) |
| 2 | O | 0.53 | 0/10745 | 0.75 | 3/14499 (0.0%) |
| 3 | D | 0.54 | 0/10729 | 0.77 | 4/14487 (0.0%) |
| 3 | J | 0.58 | 2/10729 (0.0%) | 0.81 | 10/14487 (0.1%) |
| 3 | P | 0.55 | 1/10729 (0.0%) | 0.77 | 6/14487 (0.0%) |
| 4 | E | 0.54 | 1/710 (0.1%) | 0.72 | 0/956 |
| 4 | K | 0.53 | 0/710 | 0.73 | 0/956 |
| 4 | Q | 0.52 | 0/710 | 0.72 | 0/956 |
| 5 | F | 0.49 | 1/4076 (0.0%) | 0.69 | 0/5482 |
| 5 | L | 0.51 | 0/4076 | 0.72 | 0/5482 |
| 5 | R | 0.55 | 2/4076 (0.0%) | 0.74 | 1/5482 (0.0%) |
| 6 | 1 | 0.41 | 0/1115 | 0.69 | 0/1718 |
| 6 | 4 | 0.33 | 0/1112 | 0.66 | 0/1706 |
| 6 | 7 | 0.37 | 0/1114 | 0.67 | 0/1714 |
| 7 | 2 | 0.37 | 0/1134 | 0.67 | 0/1744 |
| 7 | 5 | 0.35 | 0/1134 | 0.65 | 0/1744 |
| 7 | 8 | 0.38 | 0/1136 | 0.64 | 0/1752 |
| 8 | 3 | 0.44 | 0/72 | 0.62 | 0/110 |
| 8 | 6 | 0.40 | 0/72 | 0.61 | 0/110 |
| 8 | 9 | 0.36 | 0/72 | 0.59 | 0/110 |
| All | All | 0.53 | 10/96535 (0.0%) | 0.76 | 34/131605 (0.0%) |

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected

by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

| Mol | Chain | #Chirality outliers | #Planarity outliers |
|---|---|---|---|
| 3 | D | 0 | 1 |

The worst 5 of 10 bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 3 | P | 1340 | LYS | CB-CG | 6.65 | 1.70 | 1.52 |
| 2 | I | 626 | GLU | CD-OE2 | 6.62 | 1.32 | 1.25 |
| 2 | I | 626 | GLU | CD-OE1 | 5.92 | 1.32 | 1.25 |
| 5 | R | 109 | GLU | CD-OE1 | 5.75 | 1.31 | 1.25 |
| 5 | F | 491 | GLU | CB-CG | 5.70 | 1.62 | 1.52 |

The worst 5 of 34 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 3 | J | 737 | ILE | CB-CA-C | -7.98 | 95.64 | 111.60 |
| 3 | J | 803 | VAL | CB-CA-C | -7.38 | 97.38 | 111.40 |
| 5 | R | 488 | LEU | CA-CB-CG | 7.32 | 132.12 | 115.30 |
| 3 | D | 737 | ILE | CB-CA-C | -7.15 | 97.30 | 111.60 |
| 2 | O | 57 | PHE | C-N-CD | -7.09 | 105.00 | 120.60 |

There are no chirality outliers.

All (1) planarity outliers are listed below:

| Mol | Chain | Res | Type | Group |
|---|---|---|---|---|
| 3 | D | 671 | GLY | 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 | 220 | 0 |
| 1 | B | 1767 | 0 | 1789 | 175 | 0 |
| 1 | G | 1787 | 0 | 1812 | 173 | 0 |
| 1 | H | 1767 | 0 | 1789 | 149 | 0 |
| 1 | M | 1787 | 0 | 1813 | 178 | 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 | 142 | 0 |
| 2 | C | 10576 | 0 | 10591 | 868 | 0 |
| 2 | I | 10576 | 0 | 10591 | 845 | 0 |
| 2 | O | 10576 | 0 | 10591 | 771 | 0 |
| 3 | D | 10568 | 0 | 10782 | 856 | 3 |
| 3 | J | 10568 | 0 | 10780 | 1069 | 2 |
| 3 | P | 10568 | 0 | 10780 | 901 | 0 |
| 4 | E | 708 | 0 | 719 | 42 | 0 |
| 4 | K | 708 | 0 | 719 | 48 | 0 |
| 4 | Q | 708 | 0 | 719 | 36 | 0 |
| 5 | F | 4022 | 0 | 4083 | 243 | 0 |
| 5 | L | 4022 | 0 | 4083 | 270 | 0 |
| 5 | R | 4022 | 0 | 4083 | 282 | 0 |
| 6 | 1 | 996 | 0 | 554 | 70 | 1 |
| 6 | 4 | 996 | 0 | 557 | 76 | 0 |
| 6 | 7 | 996 | 0 | 555 | 74 | 0 |
| 7 | 2 | 1012 | 0 | 556 | 62 | 0 |
| 7 | 5 | 1012 | 0 | 556 | 59 | 0 |
| 7 | 8 | 1012 | 0 | 554 | 64 | 0 |
| 8 | 3 | 97 | 0 | 44 | 7 | 0 |
| 8 | 6 | 97 | 0 | 44 | 8 | 0 |
| 8 | 9 | 97 | 0 | 44 | 4 | 0 |
| 9 | D | 2 | 0 | 0 | 0 | 0 |
| 9 | J | 2 | 0 | 0 | 2 | 0 |
| 9 | P | 2 | 0 | 0 | 0 | 0 |
| 10 | 6 | 1 | 0 | 0 | 0 | 0 |
| 10 | D | 1 | 0 | 0 | 0 | 0 |
| 10 | P | 1 | 0 | 0 | 0 | 0 |
| All | All | 94608 | 0 | 92790 | 6821 | 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 6821 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:D:608:CYS:SG | 3:D:617:THR:HG22 | 1.31 | 1.67 |
| 3:D:501:VAL:CG1 | 3:D:502:PRO:HD2 | 1.33 | 1.55 |
| 3:J:349:TYR:O | 3:J:470:VAL:HG23 | 1.24 | 1.30 |
| 3:D:645:VAL:CG2 | 3:D:701:LEU:HD13 | 1.59 | 1.30 |
| 5:L:573:LEU:HB2 | 7:5:46:DG:OP2 | 1.15 | 1.28 |

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) |
|---|---|---|---|
| 3:D:1169:THR:OG1 | 6:1:16:DA:OP1[2_657] | 1.85 | 0.35 |
| 3:D:710:ASP:OD2 | 3:J:1282:TYR:OH[2_547] | 1.93 | 0.27 |
| 3:D:710:ASP:CA | 3:J:1302:TYR:OH[2_547] | 2.07 | 0.13 |

## 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%) | 214 (94%) | 11 (5%) | 3 (1%) | 12 47 |
| 1 | B | 226/242 (93%) | 204 (90%) | 17 (8%) | 5 (2%) | 6 35 |
| 1 | G | 228/242 (94%) | 209 (92%) | 16 (7%) | 3 (1%) | 12 47 |
| 1 | H | 226/242 (93%) | 207 (92%) | 13 (6%) | 6 (3%) | 5 31 |
| 1 | M | 228/242 (94%) | 214 (94%) | 14 (6%) | 0 | 100 100 |
| 1 | N | 226/242 (93%) | 209 (92%) | 14 (6%) | 3 (1%) | 12 47 |
| 2 | C | 1339/1342 (100%) | 1218 (91%) | 98 (7%) | 23 (2%) | 9 41 |
| 2 | I | 1339/1342 (100%) | 1214 (91%) | 105 (8%) | 20 (2%) | 10 45 |
| 2 | O | 1339/1342 (100%) | 1234 (92%) | 90 (7%) | 15 (1%) | 14 51 |
| 3 | D | 1360/1407 (97%) | 1220 (90%) | 109 (8%) | 31 (2%) | 6 34 |
| 3 | J | 1360/1407 (97%) | 1227 (90%) | 99 (7%) | 34 (2%) | 5 32 |
| 3 | P | 1360/1407 (97%) | 1226 (90%) | 99 (7%) | 35 (3%) | 5 31 |
| 4 | E | 88/90 (98%) | 83 (94%) | 5 (6%) | 0 | 100 100 |
| 4 | K | 88/90 (98%) | 84 (96%) | 3 (3%) | 1 (1%) | 14 51 |
| 4 | Q | 88/90 (98%) | 84 (96%) | 4 (4%) | 0 | 100 100 |
| 5 | F | 493/628 (78%) | 444 (90%) | 27 (6%) | 22 (4%) | 2 22 |
| 5 | L | 493/628 (78%) | 447 (91%) | 28 (6%) | 18 (4%) | 3 25 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed | Favoured | Allowed | Outliers | Percentiles |
|---|---|---|---|---|---|---|
| 5 | R | 493/628 (78%) | 449 (91%) | 30 (6%) | 14 (3%) | 5 30 |
| All | All | 11202/11853 (94%) | 10187 (91%) | 782 (7%) | 233 (2%) | 7 36 |

5 of 233 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 164 | ASP |
| 1 | B | 118 | ASP |
| 2 | C | 110 | PRO |
| 2 | C | 214 | ASN |
| 2 | C | 247 | ARG |

### 5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

| Mol | Chain | Analysed | Rotameric | Outliers | Percentiles |
|---|---|---|---|---|---|
| 1 | A | 198/208 (95%) | 181 (91%) | 17 (9%) | 10 34 |
| 1 | B | 196/208 (94%) | 172 (88%) | 24 (12%) | 5 21 |
| 1 | G | 198/208 (95%) | 178 (90%) | 20 (10%) | 7 27 |
| 1 | H | 196/208 (94%) | 174 (89%) | 22 (11%) | 6 23 |
| 1 | M | 198/208 (95%) | 178 (90%) | 20 (10%) | 7 27 |
| 1 | N | 196/208 (94%) | 176 (90%) | 20 (10%) | 7 26 |
| 2 | C | 1156/1157 (100%) | 1042 (90%) | 114 (10%) | 8 28 |
| 2 | I | 1156/1157 (100%) | 1052 (91%) | 104 (9%) | 9 32 |
| 2 | O | 1156/1157 (100%) | 1050 (91%) | 106 (9%) | 9 30 |
| 3 | D | 1135/1168 (97%) | 1026 (90%) | 109 (10%) | 8 29 |
| 3 | J | 1135/1168 (97%) | 1014 (89%) | 121 (11%) | 6 25 |
| 3 | P | 1135/1168 (97%) | 1017 (90%) | 118 (10%) | 7 26 |
| 4 | E | 74/74 (100%) | 70 (95%) | 4 (5%) | 22 49 |
| 4 | K | 74/74 (100%) | 67 (90%) | 7 (10%) | 8 29 |
| 4 | Q | 74/74 (100%) | 66 (89%) | 8 (11%) | 6 25 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed | Rotameric | Outliers | Percentiles |
|---|---|---|---|---|---|
| 5 | F | 439/554 (79%) | 406 (92%) | 33 (8%) | 13 39 |
| 5 | L | 439/554 (79%) | 394 (90%) | 45 (10%) | 7 26 |
| 5 | R | 439/554 (79%) | 393 (90%) | 46 (10%) | 7 26 |
| All | All | 9594/10107 (95%) | 8656 (90%) | 938 (10%) | 8 28 |

5 of 938 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 2 | I | 1085 | MET |
| 3 | J | 882 | VAL |
| 3 | P | 1250 | ASP |
| 2 | I | 1255 | THR |
| 3 | J | 321 | LYS |

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 148 such sidechains are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 3 | J | 419 | HIS |
| 3 | J | 979 | ASN |
| 3 | P | 1259 | GLN |
| 3 | J | 465 | GLN |
| 3 | J | 700 | ASN |

### 5.3.3 RNA

| Mol | Chain | Analysed | Backbone Outliers | Pucker Outliers |
|---|---|---|---|---|
| 8 | 3 | 2/4 (50%) | 1 (50%) | 0 |
| 8 | 6 | 3/4 (75%) | 1 (33%) | 1 (33%) |
| 8 | 9 | 3/4 (75%) | 1 (33%) | 1 (33%) |
| All | All | 8/12 (66%) | 3 (37%) | 2 (25%) |

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 (2) RNA pucker outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 8 | 6 | 13 | GTP |
| 8 | 9 | 13 | GTP |

## 5.4 Non-standard residues in protein, DNA, RNA chains

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

## 5.5 Carbohydrates

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

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

There are no such residues in this entry.

## 5.8 Polymer linkage issues

The following chains have linkage breaks:

| Mol | Chain | Number of breaks |
|---|---|---|
| 7 | 2 | 3 |
| 6 | 4 | 3 |
| 7 | 5 | 3 |
| 7 | 8 | 1 |
| 6 | 7 | 1 |

The worst 5 of 11 chain breaks are listed below:

| Model | Chain | Residue-1 | Atom-1 | Residue-2 | Atom-2 | Distance (Å) |
|---|---|---|---|---|---|---|
| 1 | 4 | 45:DT | O3' | 46:DG | P | 5.04 |
| 1 | 7 | 50:DT | O3' | 51:DC | P | 4.24 |
| 1 | 8 | 22:DA | O3' | 23:DT | P | 3.80 |
| 1 | 2 | 22:DA | O3' | 23:DT | P | 3.79 |
| 1 | 5 | 22:DA | O3' | 23:DT | P | 3.79 |

# 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%) | -0.13 | 3 (1%) 77 68 | 153, 175, 210, 235 | 0 |
| 1 | B | 228/242 (94%) | -0.18 | 2 (0%) 84 77 | 162, 194, 217, 238 | 0 |
| 1 | G | 230/242 (95%) | 0.05 | 6 (2%) 56 47 | 157, 185, 217, 248 | 0 |
| 1 | H | 228/242 (94%) | -0.14 | 2 (0%) 84 77 | 160, 191, 229, 261 | 0 |
| 1 | M | 230/242 (95%) | 0.05 | 3 (1%) 77 68 | 166, 200, 233, 252 | 0 |
| 1 | N | 228/242 (94%) | 0.25 | 9 (3%) 39 33 | 186, 233, 258, 273 | 0 |
| 2 | C | 1341/1342 (99%) | -0.10 | 12 (0%) 84 77 | 119, 186, 244, 277 | 0 |
| 2 | I | 1341/1342 (99%) | -0.10 | 21 (1%) 72 63 | 130, 195, 278, 377 | 0 |
| 2 | O | 1341/1342 (99%) | -0.10 | 12 (0%) 84 77 | 144, 183, 235, 270 | 0 |
| 3 | D | 1362/1407 (96%) | 0.16 | 94 (6%) 16 15 | 128, 214, 296, 349 | 0 |
| 3 | J | 1362/1407 (96%) | 0.08 | 56 (4%) 37 32 | 132, 194, 280, 314 | 0 |
| 3 | P | 1362/1407 (96%) | 0.29 | 119 (8%) 10 11 | 148, 208, 292, 330 | 0 |
| 4 | E | 90/90 (100%) | 1.11 | 28 (31%) 0 1 | 169, 206, 407, 461 | 0 |
| 4 | K | 90/90 (100%) | 0.42 | 12 (13%) 3 5 | 144, 199, 394, 442 | 0 |
| 4 | Q | 90/90 (100%) | 0.67 | 12 (13%) 3 5 | 167, 222, 416, 460 | 0 |
| 5 | F | 497/628 (79%) | 0.38 | 61 (12%) 4 7 | 182, 294, 404, 418 | 0 |
| 5 | L | 497/628 (79%) | 0.32 | 56 (11%) 5 8 | 169, 262, 400, 406 | 0 |
| 5 | R | 497/628 (79%) | 0.26 | 44 (8%) 9 11 | 172, 259, 413, 444 | 0 |
| 6 | 1 | 49/49 (100%) | 0.36 | 5 (10%) 6 9 | 201, 272, 311, 317 | 0 |
| 6 | 4 | 49/49 (100%) | 0.16 | 3 (6%) 21 19 | 209, 264, 308, 350 | 0 |
| 6 | 7 | 49/49 (100%) | 0.23 | 2 (4%) 37 32 | 211, 255, 278, 300 | 0 |
| 7 | 2 | 49/49 (100%) | 0.24 | 2 (4%) 37 32 | 215, 278, 312, 343 | 0 |
| 7 | 5 | 49/49 (100%) | 0.41 | 4 (8%) 11 12 | 198, 270, 339, 341 | 0 |
| 7 | 8 | 49/49 (100%) | 0.22 | 0 100 100 | 195, 260, 296, 335 | 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%) | 0.55 | 0 | 100 | 100 | 255, 255, 281, 321 | 0 |
| 8 | 6 | 3/4 (75%) | 0.38 | 0 | 100 | 100 | 263, 263, 272, 282 | 0 |
| 8 | 9 | 3/4 (75%) | 0.75 | 0 | 100 | 100 | 262, 262, 277, 295 | 0 |
| All | All | 11547/12159 (94%) | 0.09 | 568 (4%) | 29 | 27 | 119, 203, 358, 461 | 0 |

The worst 5 of 568 RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|---|---|---|---|---|
| 5 | L | 210 | ASN | 8.7 |
| 3 | P | 1068 | THR | 8.3 |
| 3 | P | 1006 | GLY | 8.1 |
| 5 | L | 211 | SER | 7.9 |
| 3 | D | 961 | SER | 6.8 |

## 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, 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 |
|---|---|---|---|---|---|---|---|---|
| 9 | ZN | J | 1501 | 1/1 | 0.88 | 0.15 | 200,200,200,200 | 0 |
| 10 | MG | 6 | 101 | 1/1 | 0.88 | 0.33 | 189,189,189,189 | 0 |
| 10 | MG | P | 1503 | 1/1 | 0.91 | 0.15 | 194,194,194,194 | 0 |
| 9 | ZN | D | 1502 | 1/1 | 0.94 | 0.11 | 212,212,212,212 | 0 |
| 9 | ZN | P | 1501 | 1/1 | 0.95 | 0.10 | 214,214,214,214 | 0 |
| 10 | MG | D | 1503 | 1/1 | 0.95 | 0.16 | 176,176,176,176 | 0 |
| 9 | ZN | D | 1501 | 1/1 | 0.97 | 0.09 | 228,228,228,228 | 0 |
| 9 | ZN | P | 1502 | 1/1 | 0.98 | 0.17 | 187,187,187,187 | 0 |
| 9 | ZN | J | 1502 | 1/1 | 0.98 | 0.12 | 174,174,174,174 | 0 |

## 6.5 Other polymers (i)

There are no such residues in this entry.