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

Dec 16, 2023 – 08:18 pm GMT

| | | |
|---|---|---|
| PDB ID | : | 4BXX |
| Title | : | Arrested RNA polymerase II-Bye1 complex |
| Authors | : | Kinkelin, K.; Wozniak, G.G.; Rothbart, S.B.; Lidschreiber, M.; Strahl, B.D.; Cramer, P. |
| Deposited on | : | 2013-07-16 |
| Resolution | : | 3.28 Å(reported) |

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

---

The following versions of software and data (see references (i)) were used in the production of this report:

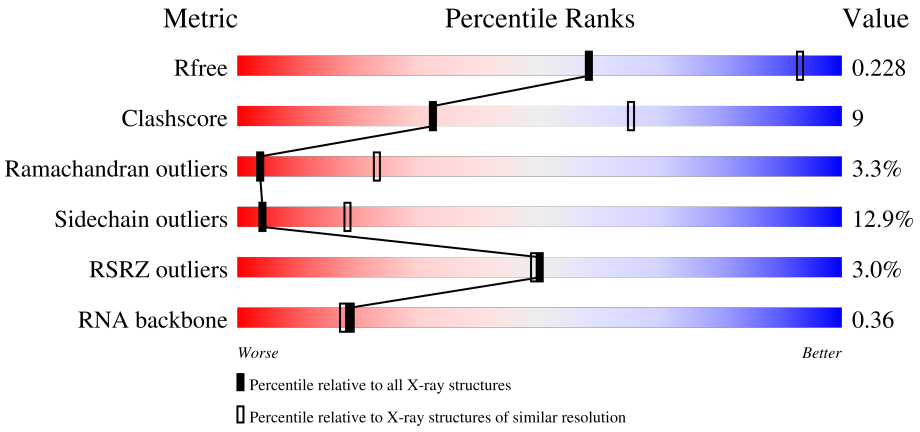
| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Mogul | : | 1.8.4, CSD as541be (2020) |
| Xtriage (Phenix) | : | 1.13 |
| EDS | : | 2.36 |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Refmac | : | 5.8.0158 |
| CCP4 | : | 7.0.044 (Gargrove) |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.36 |

# 1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.28 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.



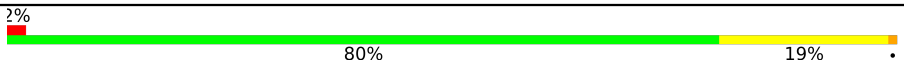
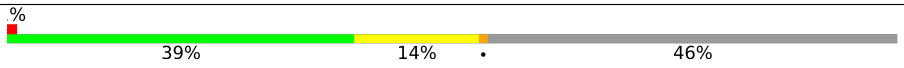
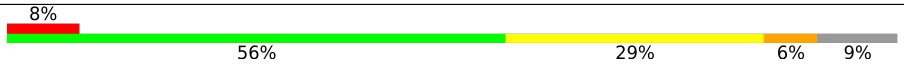
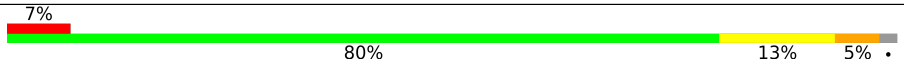
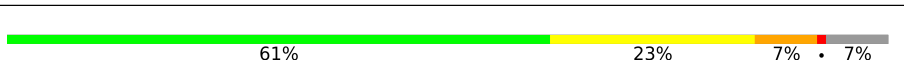
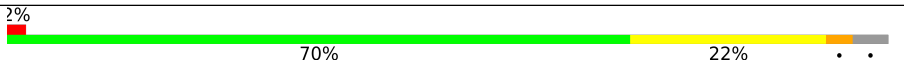
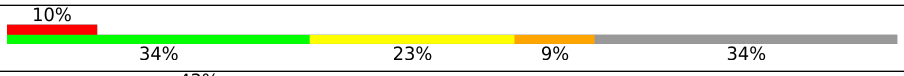
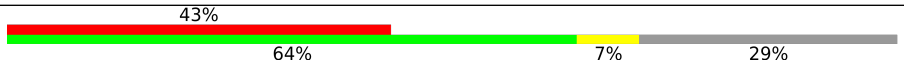
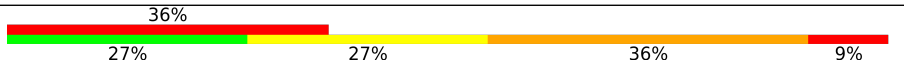
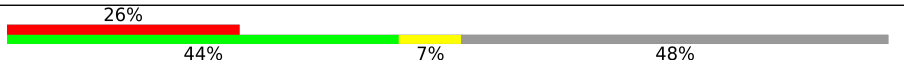

| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|---|---|---|
| $R_{free}$ | 130704 | 1177 (3.32-3.24) |
| Clashscore | 141614 | 1044 (3.30-3.26) |
| Ramachandran outliers | 138981 | 1026 (3.30-3.26) |
| Sidechain outliers | 138945 | 1025 (3.30-3.26) |
| RSRZ outliers | 127900 | 1141 (3.32-3.24) |
| RNA backbone | 3102 | 1091 (3.66-2.90) |

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 1 | A | 1733 | 57% 21% · 18% |
| 2 | B | 1224 | 62% 24% · 10% |
| 3 | C | 318 | 59% 20% · 16% |
| 4 | D | 221 | 59% 18% · 20% |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 5 | E | 215 |   |
| 6 | F | 155 |   |
| 7 | G | 171 |   |
| 8 | H | 146 |   |
| 9 | I | 122 |   |
| 10 | J | 70 |   |
| 11 | K | 120 |   |
| 12 | L | 70 |   |
| 13 | N | 14 |   |
| 14 | P | 11 |   |
| 15 | T | 27 |   |
| 16 | X | 146 |   |

## 2 Entry composition (i)

There are 18 unique types of molecules in this entry. The entry contains 32756 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 II SUBUNIT RPB1.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A | 1429 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 11249 | 7087 | 1964 | 2136 | 62 | | | |

- Molecule 2 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB2.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 | B | 1097 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 8720 | 5526 | 1523 | 1617 | 54 | | | |

- Molecule 3 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB3.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 3 | C | 266 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 2095 | 1317 | 348 | 417 | 13 | | | |

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB4.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 4 | D | 177 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1356 | 840 | 241 | 273 | 2 | | | |

- Molecule 5 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 1.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 5 | E | 214 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1752 | 1111 | 309 | 321 | 11 | | | |

- Molecule 6 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 2.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 6 | F | 84 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 679 | 434 | 115 | 127 | 3 | | | |

- Molecule 7 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB7.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 7 | G | 171 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1340 | 861 | 222 | 249 | 8 | | | |

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 3.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 8 | H | 133 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 1068 | 673 | 180 | 211 | 4 | | | |

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 9 | I | 119 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 971 | 596 | 179 | 186 | 10 | | | |

- Molecule 10 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 10 | J | 65 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 532 | 339 | 93 | 94 | 6 | | | |

- Molecule 11 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB11.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 11 | K | 115 | Total | C | N | O | S | 0 | 0 | 1 |
| | | | 920 | 590 | 157 | 171 | 2 | | | |

- Molecule 12 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 12 | L | 46 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 363 | 224 | 72 | 63 | 4 | | | |

- Molecule 13 is a DNA chain called 5'-D(*GP*AP*GP*GP*TP*AP*AP*GP*CP*TP*AP*

GP*CP*TP)-3'.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 13 | N | 10 | Total | C | N | O | P | 0 | 0 | 0 |
| | | | 209 | 99 | 42 | 58 | 10 | | | |

- Molecule 14 is a RNA chain called 5'-D(*CP*CP*CP*CP*CP*CP*CP*CP*CP*CP*CP)-3'.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 14 | P | 11 | Total | C | N | O | P | 0 | 0 | 0 |
| | | | 220 | 99 | 33 | 77 | 11 | | | |

- Molecule 15 is a DNA chain called 5'-D(*AP*GP*CP*TP*AP*GP*CP*TP*TP*AP*CP*C P*TP*GP *GP*TP*GP* BRUP*TP*GP*CP*TP*CP*TP*AP*AP*DC)-3'.

| Mol | Chain | Residues | Atoms | | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 15 | T | 14 | Total | Br | C | N | O | P | 0 | 0 | 0 |
| | | | 287 | 1 | 136 | 49 | 87 | 14 | | | |

- Molecule 16 is a protein called TRANSCRIPTION FACTOR BYE1.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 16 | X | 120 | Total | C | N | O | S | 0 | 0 | 0 |
| | | | 986 | 634 | 164 | 185 | 3 | | | |

- Molecule 17 is ZINC ION (three-letter code: ZN) (formula: Zn).

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| 17 | A | 2 | Total | Zn | 0 | 0 |
| | | | 2 | 2 | | |
| 17 | B | 1 | Total | Zn | 0 | 0 |
| | | | 1 | 1 | | |
| 17 | C | 1 | Total | Zn | 0 | 0 |
| | | | 1 | 1 | | |
| 17 | I | 2 | Total | Zn | 0 | 0 |
| | | | 2 | 2 | | |
| 17 | J | 1 | Total | Zn | 0 | 0 |
| | | | 1 | 1 | | |
| 17 | L | 1 | Total | Zn | 0 | 0 |
| | | | 1 | 1 | | |

- Molecule 18 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

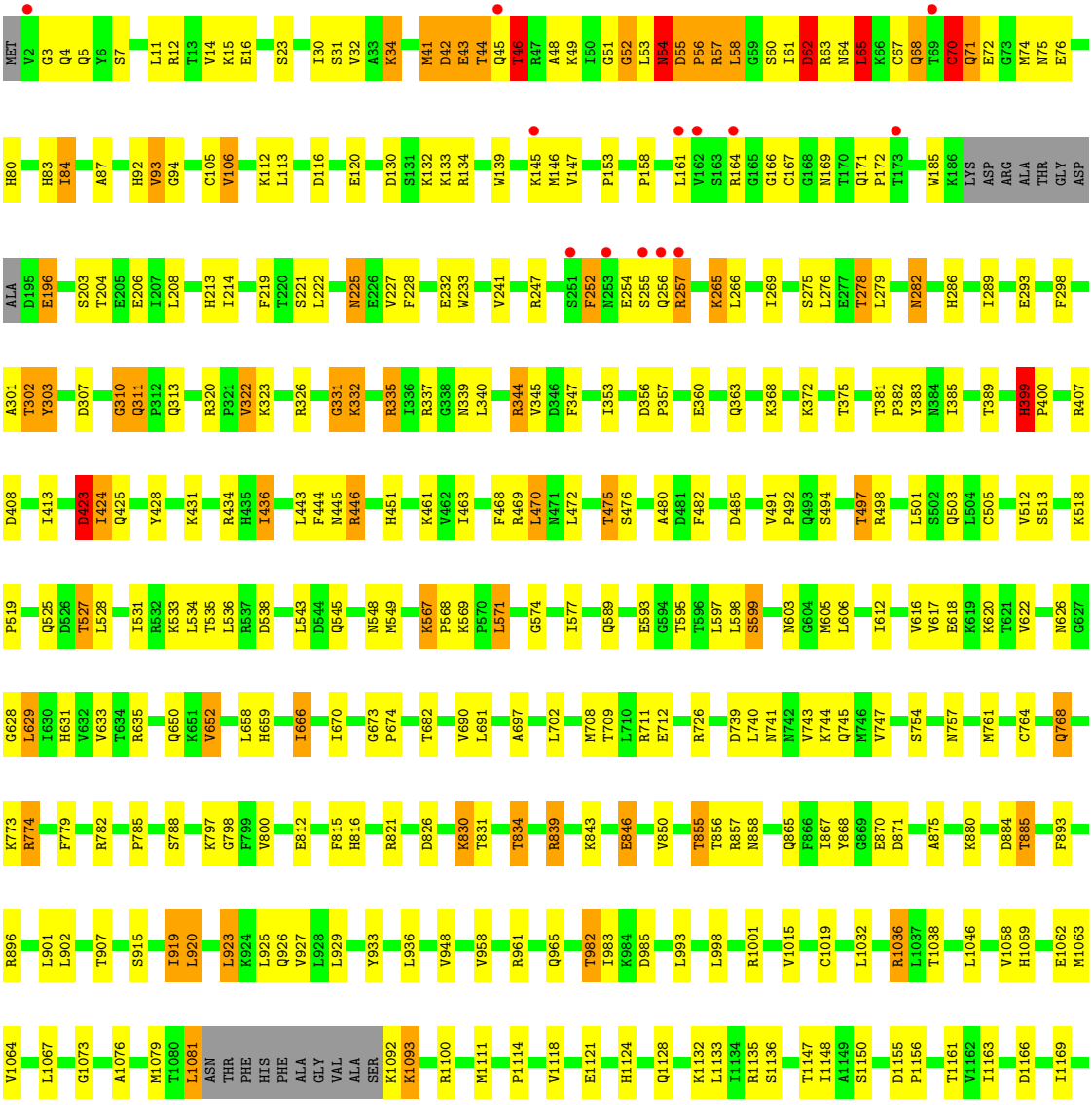
| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| 18 | A | 1 | Total | Mg | 0 | 0 |
| | | | 1 | 1 | | |

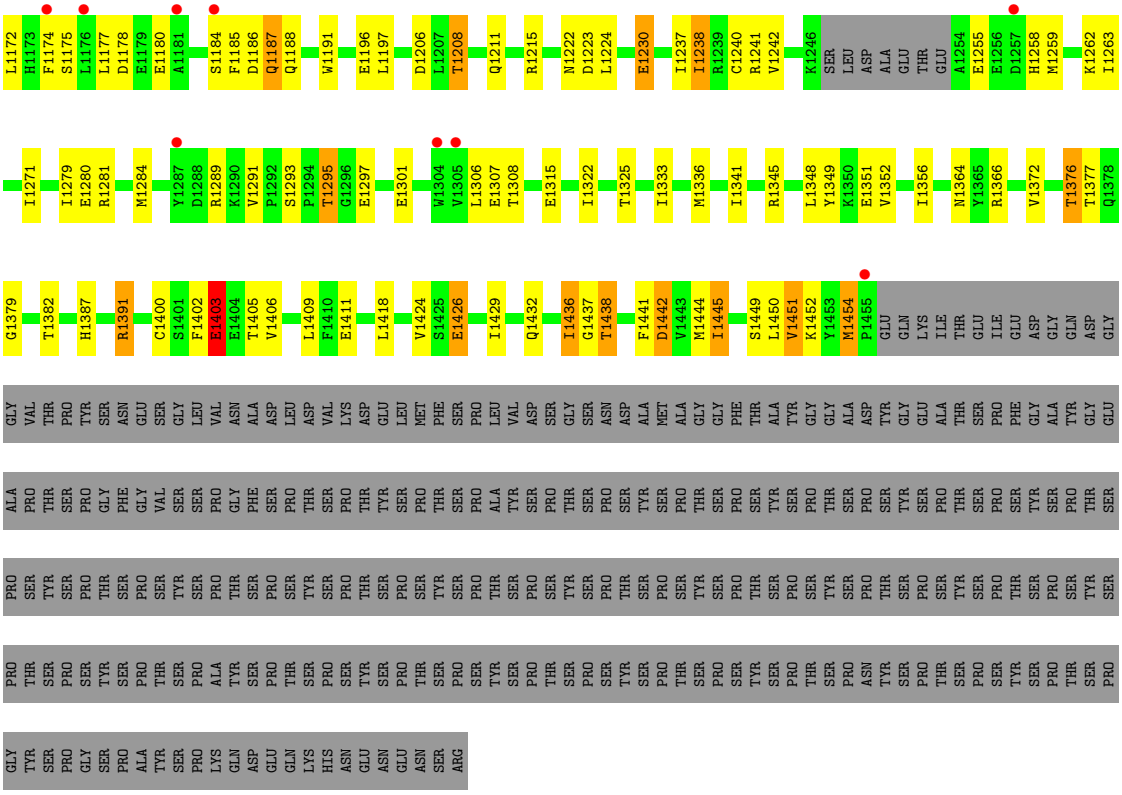
# 3 Residue-property plots (i)

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

• Molecule 1: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB1

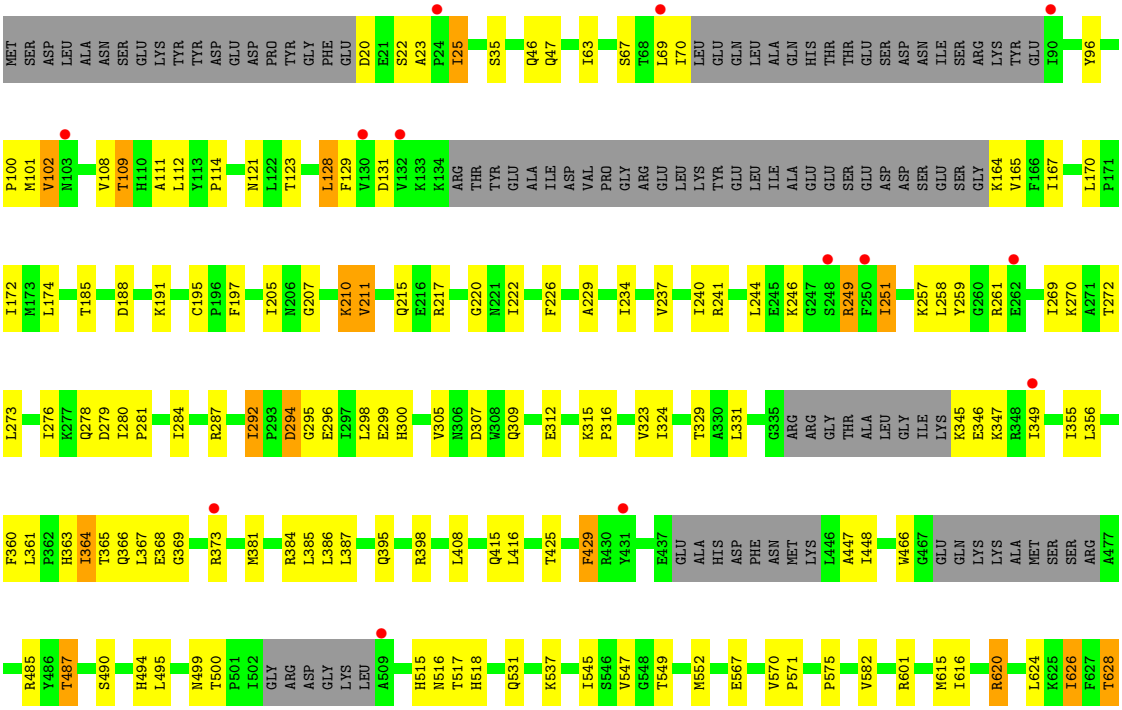
Chain A:

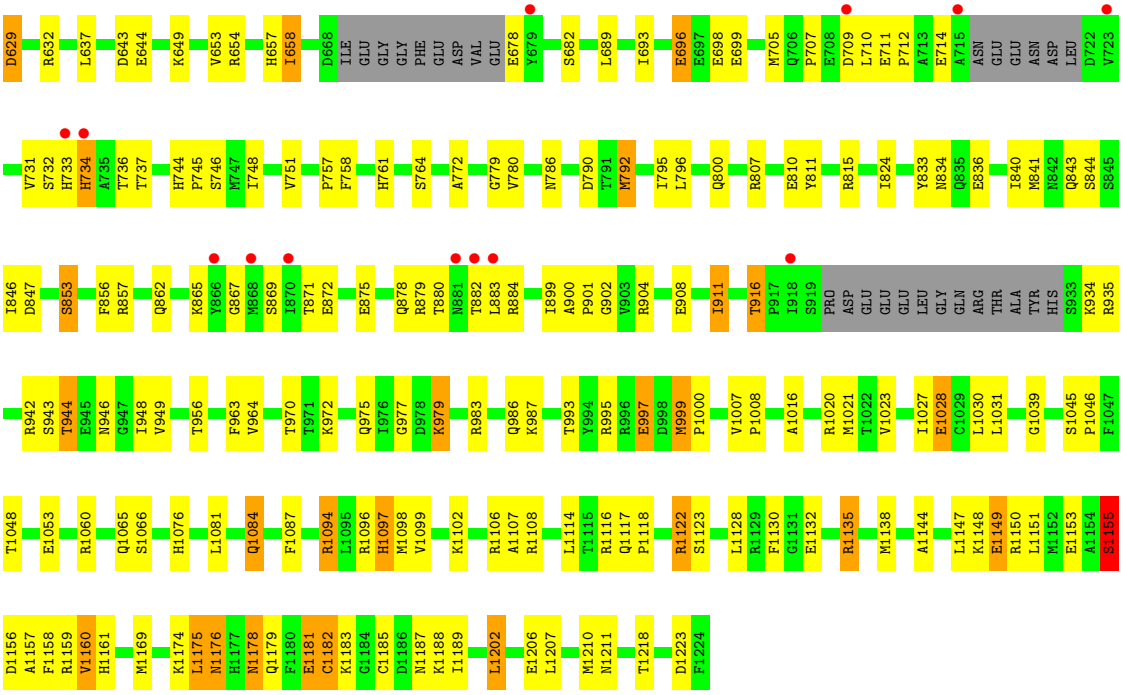





• Molecule 2: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB2

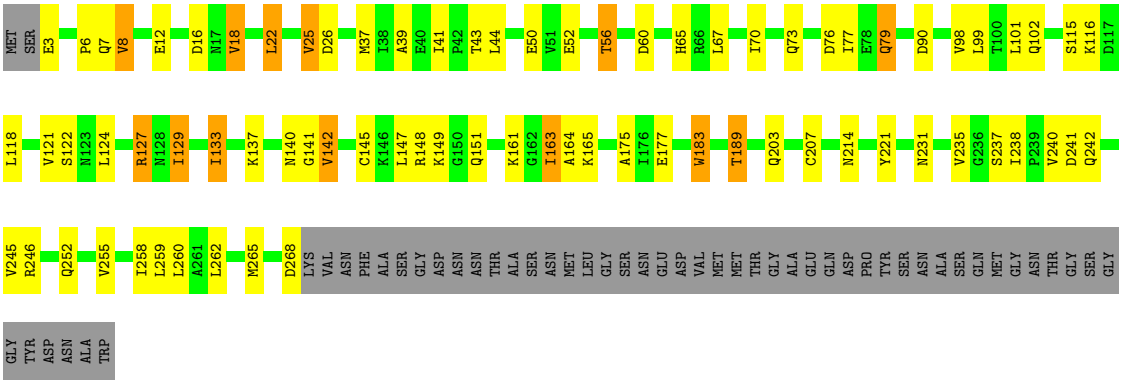
Chain B:



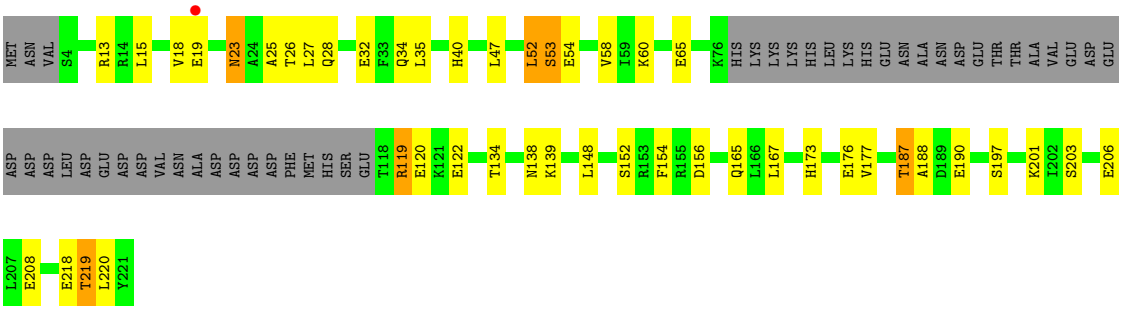
• Molecule 3: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB3

Chain C:

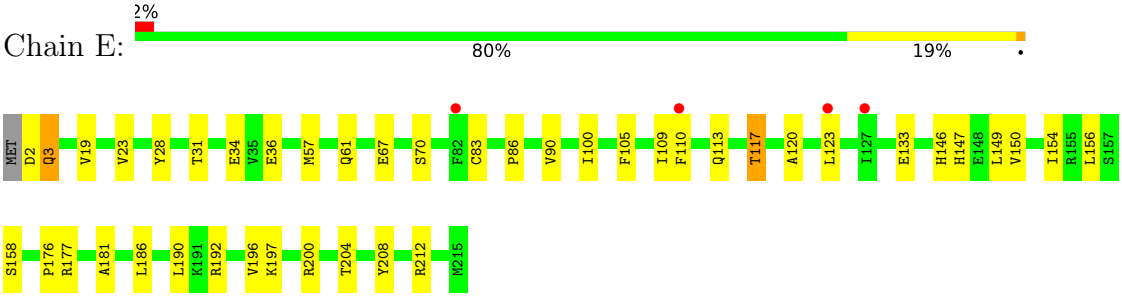
• Molecule 4: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB4

Chain D:

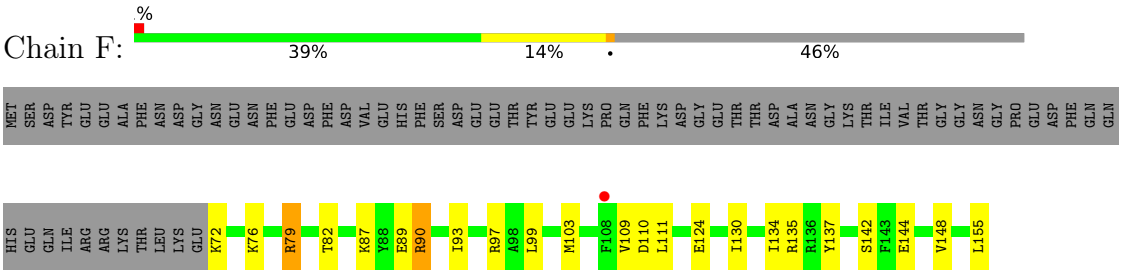
• Molecule 5: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 1

Chain E:



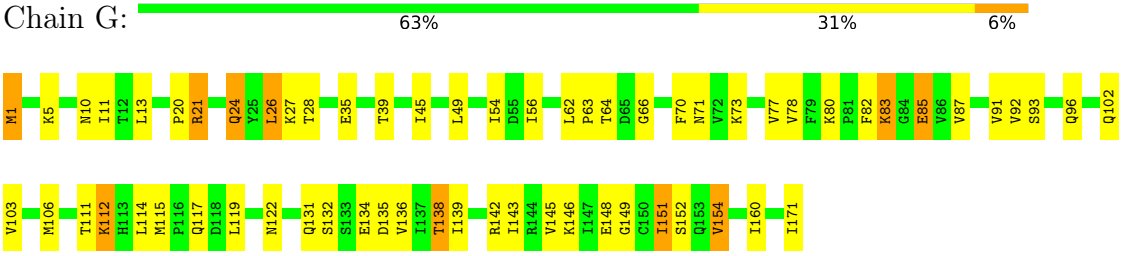

• Molecule 6: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 2

Chain F:



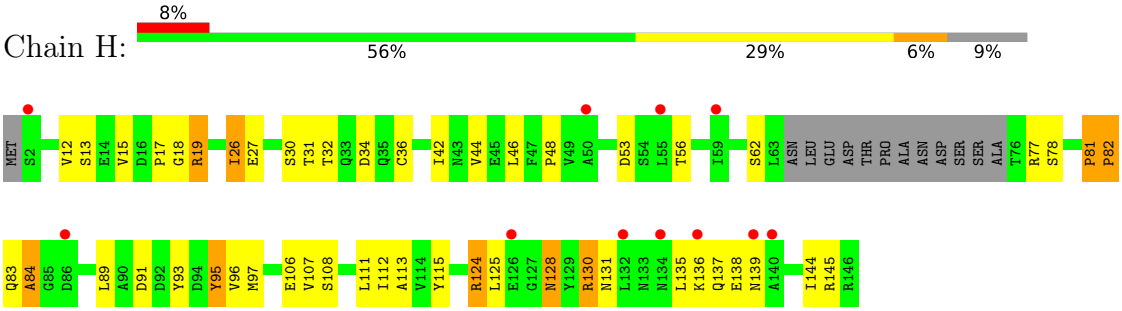

• Molecule 7: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB7

Chain G:



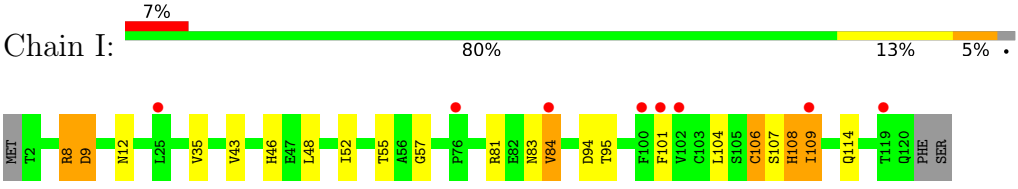

• Molecule 8: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 3

Chain H:



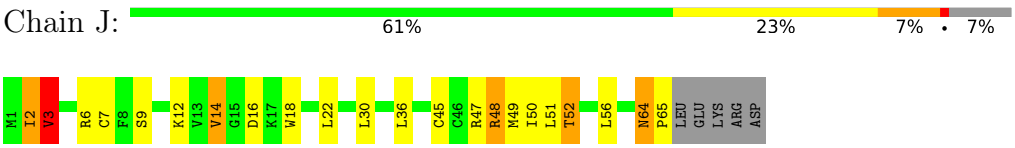

• Molecule 9: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9

Chain I:



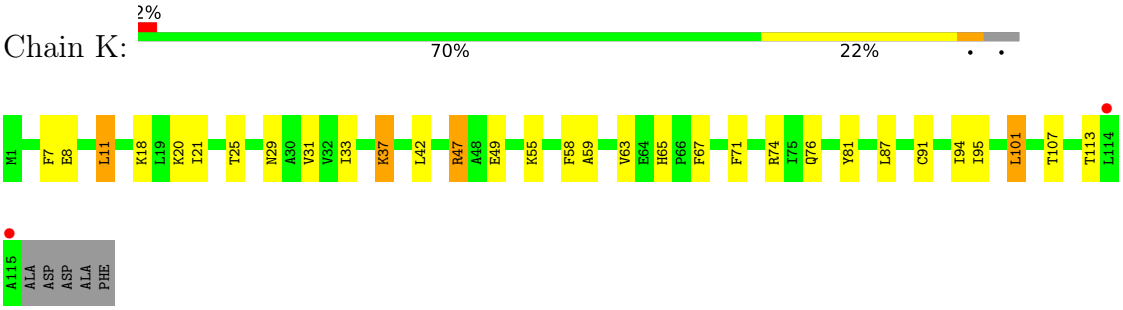

• Molecule 10: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5

Chain J:



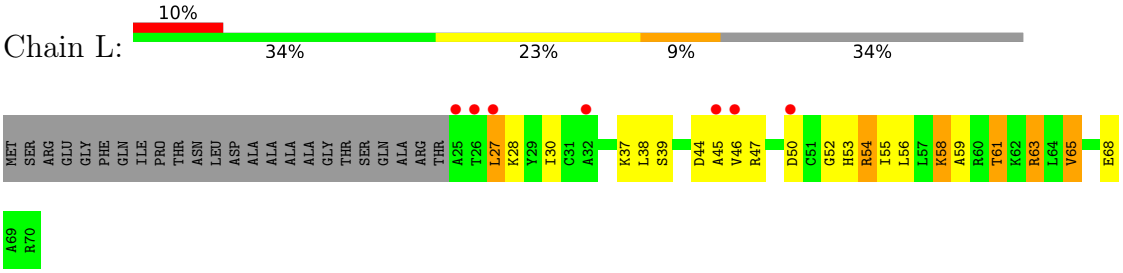

• Molecule 11: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB11

Chain K:



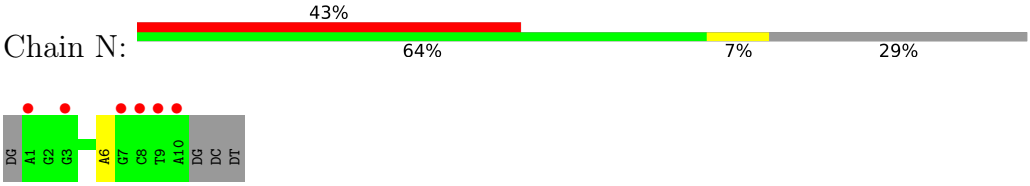

• Molecule 12: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4

Chain L:



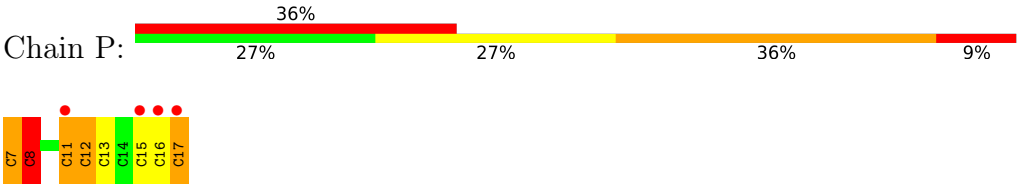

• Molecule 13: 5'-D(*GP*AP*GP*GP*TP*AP*AP*GP*CP*TP*AP*GP*CP*TP)-3'

Chain N:



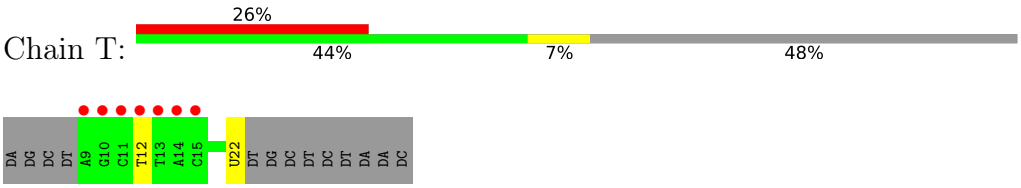

• Molecule 14: 5'-D(*CP*CP*CP*CP*CP*CP*CP*CP*CP*CP*CP)-3'

Chain P:



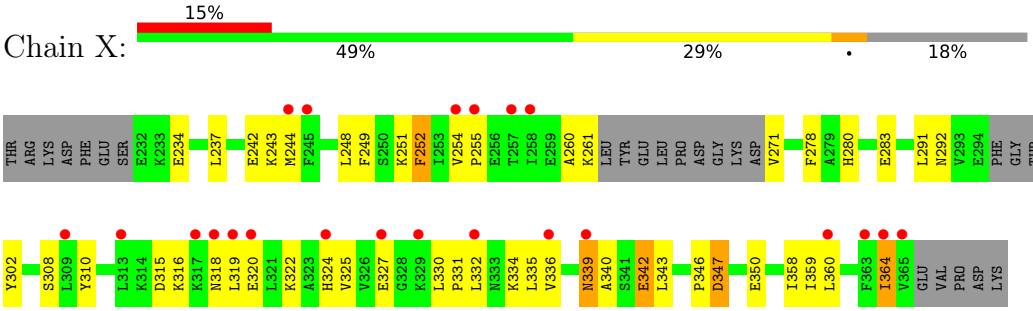

• Molecule 15: 5'-D(*AP*GP*CP*TP*AP*GP*CP*TP*TP*AP*CP*CP*TP*GP *GP*TP*GP* BRUP*TP*GP*CP*TP*CP*TP*AP*AP*DC)-3'

Chain T:




• Molecule 16: TRANSCRIPTION FACTOR BYE1

Chain X:


15%
49%
29%
18%
THR ARG LYS ASP PHE GLU SER E232 K233 E234 L237 E242 K243 M244 F245 L248 F249 S250 K251 F252 I253 V254 P255 E256 T257 I258 E259 A260 K261 LEU TYR GLU LEU PRO ASP GLY LYS ASP V271 F278 A279 H280 E283 L291 N292 V293 E294 PHE GLY THR LEU ASP K300 I301
Y302 S308 L309 Y310 L313 K314 D315 K316 K317 N318 L319 E320 L321 K322 A323 H324 V325 V326 E327 G328 K329 L330 P331 L332 N333 K334 L335 V336 N339 A340 S341 E342 L343 P346 D347 E350 I358 I359 L360 F363 I364 V365 GLU VAL PRO ASP LYS

# 4 Data and refinement statistics (i)

| Property | Value | Source |
|---|---|---|
| Space group | C 2 2 21 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 222.92Å 392.67Å 281.04Å<br>90.00° 90.00° 90.00° | Depositor |
| Resolution (Å) | 49.08 – 3.28<br>49.08 – 3.28 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 100.0 (49.08-3.28)<br>100.0 (49.08-3.28) | Depositor<br>EDS |
| $R_{merge}$ | 0.01 | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 2.00 (at 3.25Å) | Xtriage |
| Refinement program | BUSTER 2.11.5 | Depositor |
| R, $R_{free}$ | 0.180 , 0.208<br>0.199 , 0.228 | Depositor<br>DCC |
| $R_{free}$ test set | 3698 reflections (1.98%) | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 98.3 | Xtriage |
| Anisotropy | 0.446 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.32 , 100.8 | EDS |
| L-test for twinning[2] | $< \|L\| > = 0.48$, $< L^2 > = 0.30$ | Xtriage |
| Estimated twinning fraction | 0.012 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l<br>0.018 for 1/2*h+1/2*k,3/2*h-1/2*k,-l | Xtriage |
| $F_o$,$F_c$ correlation | 0.94 | EDS |
| Total number of atoms | 32756 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 120.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.80% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

# 5 Model quality (i)

## 5.1 Standard geometry (i)

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

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 | 0.54 | 0/11450 | 0.82 | 10/15483 (0.1%) |
| 2 | B | 0.51 | 0/8889 | 0.77 | 3/11987 (0.0%) |
| 3 | C | 0.50 | 0/2133 | 0.79 | 1/2891 (0.0%) |
| 4 | D | 0.53 | 0/1365 | 0.84 | 2/1837 (0.1%) |
| 5 | E | 0.45 | 0/1788 | 0.67 | 0/2406 |
| 6 | F | 0.60 | 0/691 | 0.78 | 0/933 |
| 7 | G | 0.53 | 0/1368 | 0.82 | 1/1844 (0.1%) |
| 8 | H | 0.49 | 0/1086 | 0.76 | 0/1470 |
| 9 | I | 0.45 | 0/989 | 0.72 | 0/1331 |
| 10 | J | 0.54 | 0/541 | 0.82 | 0/727 |
| 11 | K | 0.48 | 0/938 | 0.70 | 0/1267 |
| 12 | L | 0.56 | 0/365 | 0.90 | 0/485 |
| 13 | N | 0.99 | 0/235 | 0.89 | 0/361 |
| 14 | P | 1.59 | 5/241 (2.1%) | 1.04 | 0/370 |
| 15 | T | 1.27 | 0/298 | 0.99 | 0/458 |
| 16 | X | 0.50 | 0/999 | 0.75 | 0/1336 |
| All | All | 0.55 | 5/33376 (0.0%) | 0.79 | 17/45186 (0.0%) |

All (5) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 14 | P | 17 | C | C1'-N1 | 5.93 | 1.57 | 1.48 |
| 14 | P | 15 | C | C1'-N1 | 5.59 | 1.57 | 1.48 |
| 14 | P | 11 | C | C3'-O3' | 5.20 | 1.49 | 1.42 |
| 14 | P | 8 | C | C1'-N1 | 5.16 | 1.56 | 1.48 |
| 14 | P | 7 | C | C1'-N1 | 5.16 | 1.56 | 1.48 |

The worst 5 of 17 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 1 | A | 3 | GLY | C-N-CA | 8.18 | 142.14 | 121.70 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 4 | D | 25 | ALA | C-N-CA | 7.72 | 141.00 | 121.70 |
| 1 | A | 56 | PRO | C-N-CA | 6.95 | 139.07 | 121.70 |
| 1 | A | 399 | HIS | N-CA-CB | 6.93 | 123.07 | 110.60 |
| 1 | A | 54 | ASN | C-N-CA | 6.55 | 138.09 | 121.70 |

There are no chirality outliers.

There are no planarity outliers.

## 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 | 11249 | 0 | 11312 | 256 | 0 |
| 2 | B | 8720 | 0 | 8744 | 151 | 0 |
| 3 | C | 2095 | 0 | 2051 | 40 | 0 |
| 4 | D | 1356 | 0 | 1319 | 20 | 0 |
| 5 | E | 1752 | 0 | 1776 | 21 | 0 |
| 6 | F | 679 | 0 | 701 | 19 | 0 |
| 7 | G | 1340 | 0 | 1357 | 40 | 0 |
| 8 | H | 1068 | 0 | 1040 | 31 | 0 |
| 9 | I | 971 | 0 | 927 | 13 | 0 |
| 10 | J | 532 | 0 | 542 | 13 | 0 |
| 11 | K | 920 | 0 | 929 | 20 | 0 |
| 12 | L | 363 | 0 | 386 | 10 | 0 |
| 13 | N | 209 | 0 | 113 | 1 | 0 |
| 14 | P | 220 | 0 | 122 | 2 | 0 |
| 15 | T | 287 | 0 | 157 | 1 | 0 |
| 16 | X | 986 | 0 | 1009 | 18 | 0 |
| 17 | A | 2 | 0 | 0 | 0 | 0 |
| 17 | B | 1 | 0 | 0 | 0 | 0 |
| 17 | C | 1 | 0 | 0 | 0 | 0 |
| 17 | I | 2 | 0 | 0 | 0 | 0 |
| 17 | J | 1 | 0 | 0 | 0 | 0 |
| 17 | L | 1 | 0 | 0 | 0 | 0 |
| 18 | A | 1 | 0 | 0 | 0 | 0 |
| All | All | 32756 | 0 | 32485 | 571 | 0 |

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 9.

The worst 5 of 571 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) |
|---|---|---|---|
| 1:A:867:ILE:CG1 | 1:A:867:ILE:CD1 | 1.81 | 1.55 |
| 1:A:53:LEU:HD23 | 1:A:54:ASN:H | 1.06 | 1.10 |
| 1:A:567:LYS:HB3 | 8:H:96:VAL:H | 1.21 | 1.03 |
| 1:A:855:THR:HG21 | 1:A:857:ARG:HE | 1.22 | 1.02 |
| 6:F:93:ILE:HD11 | 6:F:134:ILE:HD11 | 1.39 | 1.01 |

There are no symmetry-related clashes.

## 5.3 Torsion angles (i)

### 5.3.1 Protein backbone (i)

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

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

| Mol | Chain | Analysed | Favoured | Allowed | Outliers | Percentiles |
|---|---|---|---|---|---|---|
| 1 | A | 1419/1733 (82%) | 1259 (89%) | 111 (8%) | 49 (4%) | 3 22 |
| 2 | B | 1075/1224 (88%) | 962 (90%) | 78 (7%) | 35 (3%) | 4 23 |
| 3 | C | 264/318 (83%) | 242 (92%) | 15 (6%) | 7 (3%) | 5 26 |
| 4 | D | 173/221 (78%) | 151 (87%) | 17 (10%) | 5 (3%) | 4 25 |
| 5 | E | 212/215 (99%) | 201 (95%) | 9 (4%) | 2 (1%) | 17 50 |
| 6 | F | 82/155 (53%) | 75 (92%) | 7 (8%) | 0 | 100 100 |
| 7 | G | 169/171 (99%) | 153 (90%) | 15 (9%) | 1 (1%) | 25 58 |
| 8 | H | 129/146 (88%) | 100 (78%) | 16 (12%) | 13 (10%) | 0 3 |
| 9 | I | 117/122 (96%) | 99 (85%) | 16 (14%) | 2 (2%) | 9 37 |
| 10 | J | 63/70 (90%) | 54 (86%) | 5 (8%) | 4 (6%) | 1 10 |
| 11 | K | 113/120 (94%) | 108 (96%) | 5 (4%) | 0 | 100 100 |
| 12 | L | 44/70 (63%) | 27 (61%) | 11 (25%) | 6 (14%) | 0 1 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed | Favoured | Allowed | Outliers | Percentiles |
|---|---|---|---|---|---|---|
| 16 | X | 114/146 (78%) | 99 (87%) | 9 (8%) | 6 (5%) | 2 12 |
| All | All | 3974/4711 (84%) | 3530 (89%) | 314 (8%) | 130 (3%) | 4 23 |

5 of 130 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 4 | GLN |
| 1 | A | 44 | THR |
| 1 | A | 54 | ASN |
| 1 | A | 57 | ARG |
| 1 | A | 70 | CYS |

### 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 | 1251/1520 (82%) | 1084 (87%) | 167 (13%) | 4 17 |
| 2 | B | 952/1061 (90%) | 835 (88%) | 117 (12%) | 4 20 |
| 3 | C | 234/274 (85%) | 204 (87%) | 30 (13%) | 4 18 |
| 4 | D | 140/200 (70%) | 123 (88%) | 17 (12%) | 5 21 |
| 5 | E | 196/197 (100%) | 183 (93%) | 13 (7%) | 16 46 |
| 6 | F | 74/137 (54%) | 68 (92%) | 6 (8%) | 11 36 |
| 7 | G | 152/152 (100%) | 130 (86%) | 22 (14%) | 3 15 |
| 8 | H | 117/128 (91%) | 99 (85%) | 18 (15%) | 2 12 |
| 9 | I | 113/116 (97%) | 102 (90%) | 11 (10%) | 8 29 |
| 10 | J | 60/65 (92%) | 50 (83%) | 10 (17%) | 2 10 |
| 11 | K | 99/102 (97%) | 88 (89%) | 11 (11%) | 6 24 |
| 12 | L | 40/57 (70%) | 28 (70%) | 12 (30%) | 0 1 |
| 16 | X | 110/136 (81%) | 88 (80%) | 22 (20%) | 1 5 |
| All | All | 3538/4145 (85%) | 3082 (87%) | 456 (13%) | 4 18 |

5 of 456 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 2 | B | 795 | ILE |
| 16 | X | 292 | ASN |
| 3 | C | 8 | VAL |
| 16 | X | 261 | LYS |
| 9 | I | 109 | ILE |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 79 such sidechains are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 3 | C | 242 | GLN |
| 10 | J | 53 | HIS |
| 4 | D | 28 | GLN |
| 7 | G | 14 | HIS |
| 12 | L | 66 | GLN |

### 5.3.3 RNA (i)

| Mol | Chain | Analysed | Backbone Outliers | Pucker Outliers |
|---|---|---|---|---|
| 14 | P | 10/11 (90%) | 5 (50%) | 2 (20%) |

All (5) RNA backbone outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 14 | P | 8 | C |
| 14 | P | 11 | C |
| 14 | P | 12 | C |
| 14 | P | 16 | C |
| 14 | P | 17 | C |

All (2) RNA pucker outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 14 | P | 11 | C |
| 14 | P | 16 | C |

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

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

| Mol | Type | Chain | Res | Link | Bond lengths | | | Bond angles | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Counts | RMSZ | $\#\|Z\| > 2$ | Counts | RMSZ | $\#\|Z\| > 2$ |
| 15 | BRU | T | 22 | 15 | 18,21,22 | 1.61 | 4 (22%) | 26,30,33 | 2.27 | 9 (34%) |

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

| Mol | Type | Chain | Res | Link | Chirals | Torsions | Rings |
|---|---|---|---|---|---|---|---|
| 15 | BRU | T | 22 | 15 | - | 2/7/21/22 | 0/2/2/2 |

All (4) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 15 | T | 22 | BRU | C2-N1 | 3.70 | 1.44 | 1.38 |
| 15 | T | 22 | BRU | C6-C5 | 3.23 | 1.40 | 1.34 |
| 15 | T | 22 | BRU | C4-C5 | 2.82 | 1.50 | 1.45 |
| 15 | T | 22 | BRU | C4-N3 | -2.20 | 1.34 | 1.38 |

The worst 5 of 9 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 15 | T | 22 | BRU | C5-C4-N3 | 5.17 | 119.28 | 113.34 |
| 15 | T | 22 | BRU | N3-C2-N1 | 4.93 | 121.44 | 114.89 |
| 15 | T | 22 | BRU | C4-N3-C2 | -4.51 | 121.52 | 127.35 |
| 15 | T | 22 | BRU | BR-C5-C4 | 3.71 | 122.33 | 118.03 |
| 15 | T | 22 | BRU | O4-C4-C5 | -3.52 | 121.43 | 125.84 |

There are no chirality outliers.

All (2) torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|---|---|---|---|---|
| 15 | T | 22 | BRU | O4'-C4'-C5'-O5' |
| 15 | T | 22 | BRU | C3'-C4'-C5'-O5' |

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates (i)

There are no monosaccharides 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 |
|---|---|---|
| 2 | B | 1 |
| 1 | A | 1 |

All chain breaks are listed below:

| Model | Chain | Residue-1 | Atom-1 | Residue-2 | Atom-2 | Distance (Å) |
|---|---|---|---|---|---|---|
| 1 | B | 934:LYS | C | 935:ARG | N | 4.65 |
| 1 | A | 1173:HIS | C | 1174:PHE | N | 3.16 |

# 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 | 1429/1733 (82%) | -0.00 | 22 (1%) 73 71 | 59, 101, 164, 229 | 0 |
| 2 | B | 1097/1224 (89%) | 0.12 | 26 (2%) 59 55 | 59, 114, 175, 208 | 0 |
| 3 | C | 266/318 (83%) | -0.15 | 0 100 100 | 77, 101, 141, 173 | 0 |
| 4 | D | 177/221 (80%) | -0.07 | 1 (0%) 89 90 | 80, 111, 165, 179 | 0 |
| 5 | E | 214/215 (99%) | 0.06 | 4 (1%) 66 64 | 76, 135, 181, 211 | 0 |
| 6 | F | 84/155 (54%) | -0.24 | 1 (1%) 79 78 | 63, 83, 110, 134 | 0 |
| 7 | G | 171/171 (100%) | 0.03 | 0 100 100 | 73, 100, 139, 166 | 0 |
| 8 | H | 133/146 (91%) | 0.59 | 11 (8%) 11 11 | 112, 148, 183, 209 | 0 |
| 9 | I | 119/122 (97%) | 0.35 | 8 (6%) 17 17 | 116, 148, 186, 203 | 0 |
| 10 | J | 65/70 (92%) | -0.11 | 0 100 100 | 82, 101, 138, 146 | 0 |
| 11 | K | 115/120 (95%) | 0.01 | 2 (1%) 70 67 | 70, 103, 141, 155 | 0 |
| 12 | L | 46/70 (65%) | 0.89 | 7 (15%) 2 2 | 86, 180, 196, 201 | 0 |
| 13 | N | 10/14 (71%) | 2.19 | 6 (60%) 0 0 | 239, 248, 270, 275 | 0 |
| 14 | P | 11/11 (100%) | 2.23 | 4 (36%) 0 0 | 204, 214, 222, 224 | 0 |
| 15 | T | 13/27 (48%) | 1.93 | 7 (53%) 0 0 | 165, 186, 260, 263 | 0 |
| 16 | X | 120/146 (82%) | 0.94 | 22 (18%) 1 1 | 172, 203, 219, 229 | 0 |
| All | All | 4070/4763 (85%) | 0.10 | 121 (2%) 50 49 | 59, 110, 188, 275 | 0 |

The worst 5 of 121 RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|---|---|---|---|---|
| 12 | L | 25 | ALA | 7.8 |
| 11 | K | 115 | ALA | 7.7 |
| 14 | P | 16 | C | 6.5 |
| 12 | L | 26 | THR | 6.1 |
| 1 | A | 69 | THR | 5.6 |

## 6.2 Non-standard residues in protein, DNA, RNA chains (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 |
|---|---|---|---|---|---|---|---|---|
| 15 | BRU | T | 22 | 20/21 | 0.63 | 0.32 | 191,208,212,215 | 0 |

## 6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

## 6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, $95^{th}$ percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

| Mol | Type | Chain | Res | Atoms | RSCC | RSR | B-factors(Å$^2$) | Q<0.9 |
|---|---|---|---|---|---|---|---|---|
| 18 | MG | A | 2458 | 1/1 | 0.92 | 0.10 | 77,77,77,77 | 0 |
| 17 | ZN | I | 1122 | 1/1 | 0.97 | 0.03 | 196,196,196,196 | 0 |
| 17 | ZN | L | 1071 | 1/1 | 0.98 | 0.06 | 207,207,207,207 | 0 |
| 17 | ZN | A | 2456 | 1/1 | 0.99 | 0.06 | 141,141,141,141 | 0 |
| 17 | ZN | J | 1066 | 1/1 | 0.99 | 0.23 | 90,90,90,90 | 0 |
| 17 | ZN | B | 2225 | 1/1 | 0.99 | 0.22 | 83,83,83,83 | 0 |
| 17 | ZN | I | 1121 | 1/1 | 0.99 | 0.15 | 124,124,124,124 | 0 |
| 17 | ZN | C | 1269 | 1/1 | 1.00 | 0.12 | 81,81,81,81 | 0 |
| 17 | ZN | A | 2457 | 1/1 | 1.00 | 0.17 | 79,79,79,79 | 0 |

## 6.5 Other polymers (i)

There are no such residues in this entry.