



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 08:15 pm BST

PDB ID : 4A3K  
Title : RNA Polymerase II initial transcribing complex with a 7nt DNA-RNA hybrid  
Authors : Cheung, A.C.M.; Sainsbury, S.; Cramer, P.  
Deposited on : 2011-09-30  
Resolution : 3.50 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](mailto: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 ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) 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

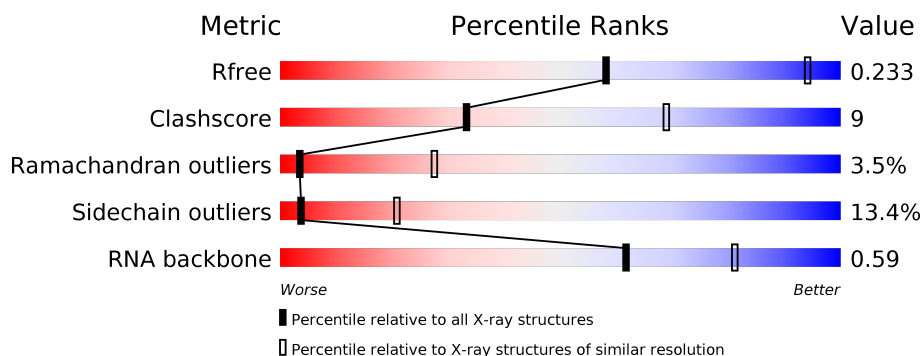
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.50 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.








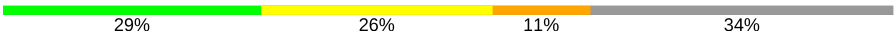

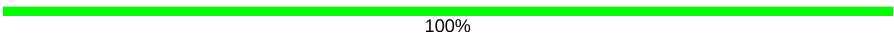

| Metric                | Whole archive<br>(#Entries) | Similar resolution<br>(#Entries, resolution range(Å)) |
|-----------------------|-----------------------------|---|
| $R_{free}$            | 130704                      | 1659 (3.60-3.40)                                      |
| Clashscore            | 141614                      | 1036 (3.58-3.42)                                      |
| Ramachandran outliers | 138981                      | 1005 (3.58-3.42)                                      |
| Sidechain outliers    | 138945                      | 1006 (3.58-3.42)                                      |
| RNA backbone          | 3102                        | 1002 (4.00-3.00)                                      |

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  $\geq 3$ , 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  $\leq 5\%$ .

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|------------------|
| 1   | A     | 1732   |                  |
| 2   | B     | 1224   |                  |
| 3   | C     | 318    |                  |
| 4   | D     | 221    |                  |
| 5   | E     | 215    |                  |
| 6   | F     | 155    |                  |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Length | Quality of chain   |
|-----|-------|--------|--|
| 7   | G     | 171    |  |
| 8   | H     | 146    |  |
| 9   | I     | 122    |  |
| 10  | J     | 70     |  |
| 11  | K     | 120    |  |
| 12  | L     | 70     |  |
| 13  | N     | 11     |  |
| 14  | P     | 7      |  |
| 15  | T     | 26     |  |

## 2 Entry composition

There are 17 unique types of molecules in this entry. The entry contains 31940 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     | 1422     | Total | C    | N    | O    | S  | 0       | 0       | 0     |
|     |       |          | 11174 | 7037 | 1954 | 2121 | 62 |         |         |       |

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

| Mol | Chain | Residues | Atoms |      |      |      |    | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|------|------|------|----|---------|---------|-------|
| 2   | B     | 1115     | Total | C    | N    | O    | S  | 0       | 0       | 0     |
|     |       |          | 8859  | 5609 | 1554 | 1641 | 55 |         |         |       |

- 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     | 178      | Total | C   | N   | O   | S | 0       | 0       | 0     |
|     |       |          | 1434  | 887 | 257 | 288 | 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 RPB7, 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(\*TP\*AP\*AP\*GP\*TP\*AP\*CP\*TP\*TP\*GP\*AP)-3'.

| Mol | Chain | Residues | Atoms |    |    |    |    | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|----|----|----|----|---------|---------|-------|
| 13  | N     | 10       | Total | C  | N  | O  | P  | 0       | 0       | 0     |
|     |       |          | 207   | 99 | 39 | 59 | 10 |         |         |       |

- Molecule 14 is a RNA chain called 5'-R(\*AP\*CP\*CP\*AP\*GP\*GP\*AP)-3'.

| Mol | Chain | Residues | Atoms |    |    |    |   | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|----|----|----|---|---------|---------|-------|
| 14  | P     | 7        | Total | C  | N  | O  | P | 0       | 0       | 0     |
|     |       |          | 152   | 68 | 31 | 46 | 7 |         |         |       |

- Molecule 15 is a DNA chain called 5'-D(\*AP\*GP\*CP\*TP\*CP\*AP\*AP\*GP\*TP\*AP\*CP\*TP\*TP\*DTP\*TP\*TP\*CP\*CP\*BRU\*GP\*GP\*TP\*CP\*AP\*TP\*T)-3'.

| Mol | Chain | Residues | Atoms |    |     |    |     | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|----|-----|----|-----|---------|---------|-------|
| 15  | T     | 19       | Total | Br | C   | N  | O   | P       | 0       | 0     |
|     |       |          | 385   | 1  | 185 | 60 | 120 | 19      |         |       |

- Molecule 16 is ZINC ION (three-letter code: ZN) (formula: Zn).

| Mol | Chain | Residues | Atoms |    | ZeroOcc | AltConf |
|-----|-------|----------|-------|----|---------|---------|
| 16  | J     | 1        | Total | Zn | 0       | 0       |
|     |       |          | 1     | 1  |         |         |
| 16  | B     | 1        | Total | Zn | 0       | 0       |
|     |       |          | 1     | 1  |         |         |
| 16  | I     | 2        | Total | Zn | 0       | 0       |
|     |       |          | 2     | 2  |         |         |
| 16  | C     | 1        | Total | Zn | 0       | 0       |
|     |       |          | 1     | 1  |         |         |
| 16  | A     | 2        | Total | Zn | 0       | 0       |
|     |       |          | 2     | 2  |         |         |
| 16  | L     | 1        | Total | Zn | 0       | 0       |
|     |       |          | 1     | 1  |         |         |

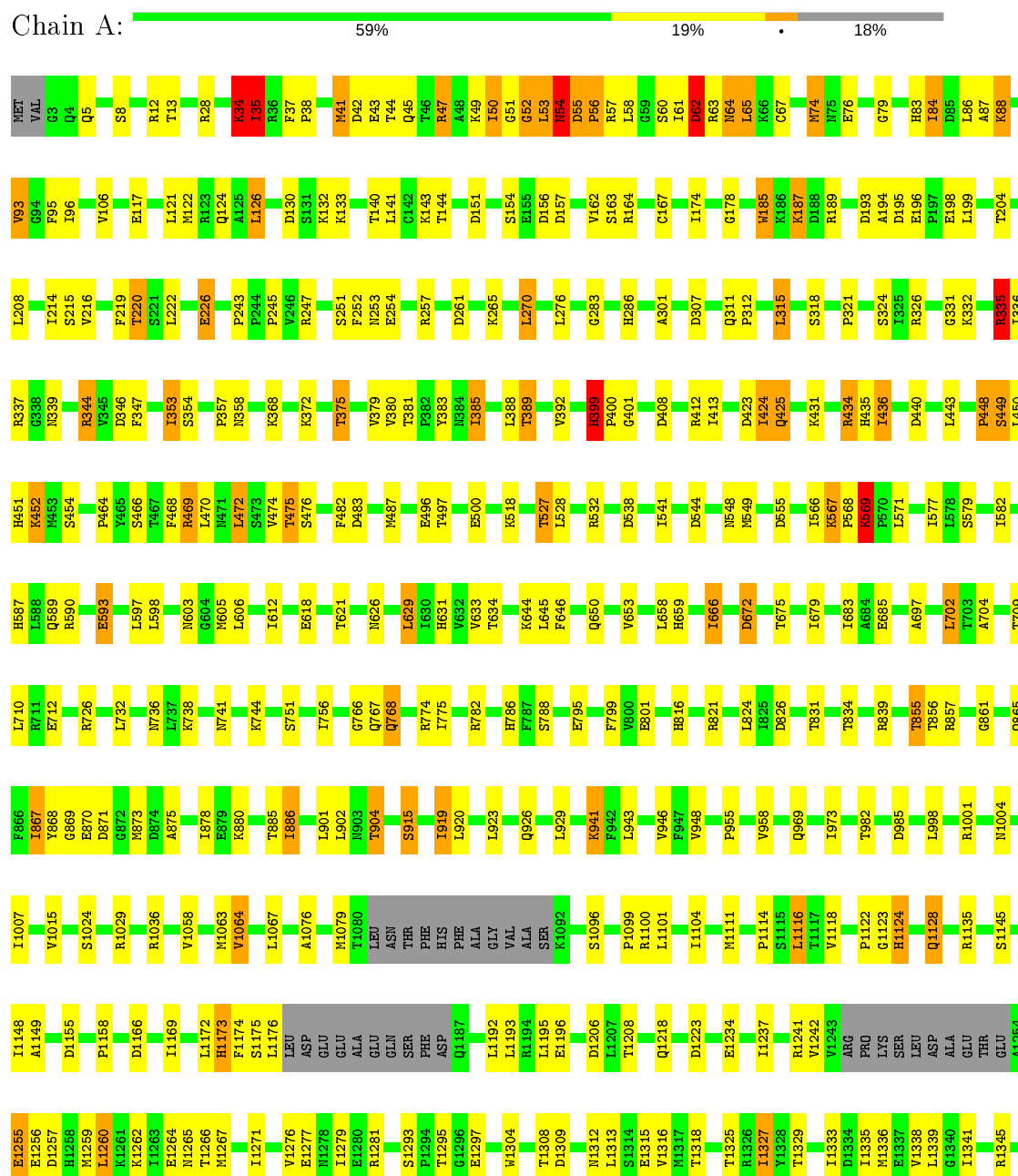
- Molecule 17 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

| Mol | Chain | Residues | Atoms |    | ZeroOcc | AltConf |
|-----|-------|----------|-------|----|---------|---------|
| 17  | A     | 1        | Total | Mg | 0       | 0       |
|     |       |          | 1     | 1  |         |         |

### 3 Residue-property plots

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. 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. 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



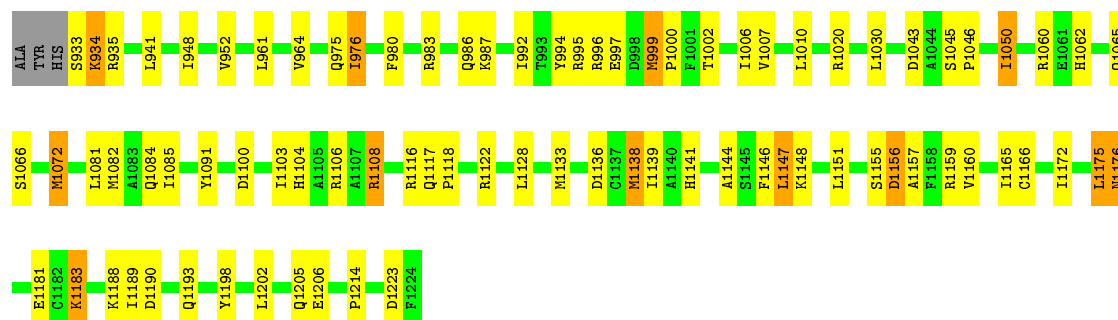
|       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |       |     |     |     |     |     |     |     |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-----|-----|-----|-----|-----|-----|-----|
| L1348 | L1349 | K1350 | E1351 | V1352 | V1355 | R1366 | L1370 | L1371 | V1372 | D1373 | V1374 | M1375 | T1376 | S1383 | R1386 | H1387 | R1391 | C1400 | S1401 | F1402 | E1403 | E1404 | T1405 | V1406 | E1407 | I1408 | A1414 | V1424 | S1425 | E1426 | V1427 | V1428 | M1433 | A1434 | P1435 | I1436 | G1437 | T1438 | G1439 | A1440 | F1441 | D1442 | V1443 | M1444 | I1445 | V1451 | M1454 |     |     |     |     |     |     |     |
| P1455 | GLU   | GLN   | LYS   | ILE   | THR   | GLU   | ILE   | GLU   | ASP   | GLY   | GLN   | ASP   | THR   | PRO   | TYR   | ASN   | GLU   | GLY   | LEU   | VAL   | THR   | PRO   | GLY   | THR   | PRO   | GLY   | LEU   | VAL   | ASN   | GLY   | PHE   | SER   | PRO   | LEU   | VAL   | ASP   | SER   | GLY   | SER   | ASN   | PRO   | ASP   | GLY   | ASP   | ALA   | GLY   | PHE   | THR | ALA | TYR | GLY | GLY | ALA |     |
| ASP   | THR   | GLY   | PRO   | ALA   | THR   | SER   | PRO   | PHE   | THR   | GLY   | GLU   | ALA   | THR   | PRO   | PHE   | GLY   | VAL   | SER   | PRO   | THR   | GLY   | THR   | PRO   | ASP   | VAL   | ASN   | GLY   | LEU   | THR   | THR   | PRO   | ALA   | THR   | SER   | PRO   | THR   | ASP   | THR   | PRO   | ASP   | THR   | ALA   | THR   | PRO   | THR   | PRO   | THR   | THR | THR | THR | THR | THR |     |     |
| PRO   | SER   | TYR   | SER   | PRO   | PRO   | SER   | PRO   | PRO   | TYR   | SER   | PRO   | PRO   | TYR   | PRO   | THR   | SER   | TYR   | PRO   | ALA   | PRO   | THR   | THR   | THR   | THR   | SER   | PRO   | GLY   | GLN   | THR   | THR   | PRO   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR | THR | THR | THR | THR | THR | THR |
| PRO   | THR   | SER   | SER   | THR   | THR   | SER   | PRO   | THR   | SER   | THR   | THR   | THR   | THR   | THR   | THR   | SER   | PRO   | ALA   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR   | THR | THR | THR | THR | THR | THR |     |
| ASN   | TYR   | SER   | PRO   | THR   | PRO   | SER   | PRO   | TYR   | THR   | PRO   | GLY   | THR   | SER   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO   | PRO | PRO | PRO | PRO | PRO | PRO |     |

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

Chain B:  63% 24% 9%

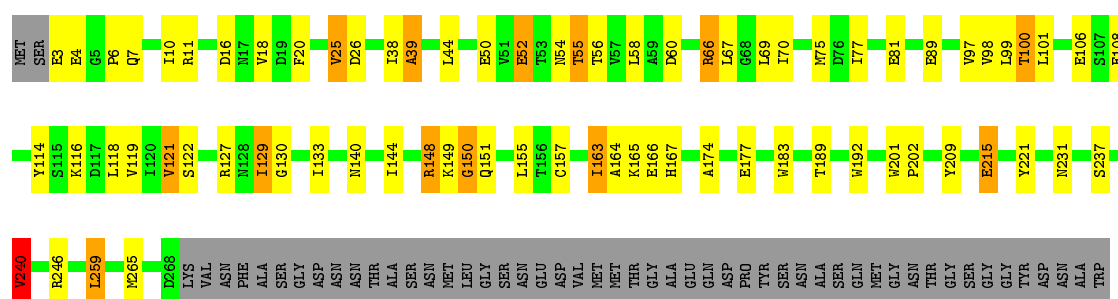
|      |      |      |      |      |      |      |      |      |     |
|------|------|------|------|------|------|------|------|------|-----|
| M842 | S732 | E644 | LYS  | L416 | E312 | I204 | V108 | SER  | MET |
| Q843 | H733 | K649 | LEU  | F417 | M313 | I205 | T109 | ASP  | ASP |
| S844 | H734 | E650 | K510 | K418 | L314 | M206 | H110 | LEU  |     |
| S845 | E651 | L651 | P511 | K423 | K315 | G207 | A111 | ALA  |     |
| I846 | F738 | K652 | C533 | L424 | P316 | V211 | L112 | ASN  |     |
| D847 | P745 | V653 | R594 | T425 | V323 | E216 | Y113 | SER  |     |
| R848 | S746 | R654 | K537 | F428 | I334 | R217 | P114 | GLU  |     |
| F856 | M747 | I658 | M542 | F429 | G335 | R217 | R118 | LYS  |     |
| I748 | I748 |      |      |      | R336 | F226 |      | TYR  |     |
| K865 | V751 | M662 | I545 | T435 | G337 |      | M121 | ASP  |     |
| Y866 |      | A663 | T546 | V436 | G338 | A229 |      | GLU  |     |
| G867 | S754 | T664 | V547 | E437 | T339 |      | L128 | ASP  |     |
| M868 | I755 | E665 | V647 | GLU  | A340 | P231 |      | PRO  |     |
| S869 | I756 |      |      | ALA  | L341 |      | K134 | TYR  |     |
| I870 | I756 |      |      | HIS  | G342 | A238 | ARG  | GLY  |     |
| T871 | F757 | D668 | M552 | ASP  | L343 |      | THR  | GLY  |     |
| E872 | F758 | ILE  | L566 | PHE  | K344 | S242 | TYR  | GLU  |     |
| T873 | P759 | GLY  | V570 | ASN  | K345 | G247 | GLY  | D20  |     |
| K876 | H761 | GLY  |      | MET  | E346 |      | ALA  | E21  |     |
| P877 |      | PHE  | LYS  | LYS  | K347 |      | ILE  | I25  |     |
| Q878 | R766 | GLU  | S574 | L446 | R348 | I251 | ASP  | ASP  |     |
| R879 |      | ASP  | P575 | A447 | I349 |      | VAL  | D29  |     |
|      | A772 | VAL  | D576 | I448 | T359 | V256 | PRO  | D29  |     |
| T882 |      | GLU  | A577 | M449 | Y351 | K257 | GLY  |      |     |
| L883 | G779 | E678 | T578 | A450 |      | L258 | ARG  | V33  |     |
| R884 |      |      | R579 | K451 | L356 |      | GLU  | I34  |     |
|      | N786 | S682 |      | T452 | Q357 | R261 | LEU  | S35  |     |
| T889 | D790 | L689 | M583 | I453 | K358 | E282 | LYS  | R39  |     |
| V890 | T791 | V690 | G584 | T454 | E359 |      | TYR  | Q46  |     |
| D891 | M792 | E931 | V586 |      | F360 | I269 | GLU  |      |     |
|      |      |      |      |      |      |      | LEU  |      |     |
| I899 |      | V692 | L661 |      | P362 | T272 | ILE  | I63  |     |
| A900 | L796 | I693 | V589 | G464 | R363 | L273 | ALA  |      |     |
| P901 | Y797 | D694 | H590 | M465 | I364 | P274 | GLU  |      |     |
| G902 |      | A695 | R591 | M466 | T365 | Y275 | GLU  | I70  |     |
| V903 | K801 | E996 |      |      |      | I276 | SER  |      |     |
| R904 |      | E997 | R595 | Q469 | E368 | K277 | GLU  | LEU  |     |
|      | T806 |      |      | K470 | G369 | Q278 | GLN  | GLU  |     |
| G907 | R807 | I703 | K606 |      | F370 | D279 | ASP  | GLN  |     |
| E908 | A808 | A704 |      | M473 |      | I280 | SER  | ALA  |     |
|      | M809 | M705 | I616 | S474 | R373 |      | GLU  | GLN  |     |
| I911 | E810 | Q706 | R620 | S475 |      | V283 | HIS  | ALA  |     |
|      | Y811 | P707 |      | R476 |      |      | SER  | LEU  |     |
| K914 | L812 | E708 | L624 |      | L385 | L289 | GLY  | THR  |     |
| T915 |      | D709 |      | V479 | L386 | G290 | K164 | THR  |     |
|      | R815 | L710 | L624 | S480 | L387 | I291 |      | GLU  |     |
| I918 | S831 | E711 | T628 | Q481 |      |      | I167 | SER  |     |
| S919 | G832 | A715 | D629 | L482 | K393 | D294 | G168 | ASP  |     |
| PRO  | R833 |      |      | L483 | Q395 | G295 | R169 | ASN  |     |
| GLU  | Y833 | ASN  | R632 | M484 | E296 | E296 | L170 | ILE  |     |
| GLY  | M834 | GLU  |      | R485 | I297 | I297 | M178 | ARG  |     |
| GLU  | Q835 | GLU  | L637 | Y486 | R398 | L298 |      | LYS  |     |
| LEU  | E836 | ASN  | F638 |      | G402 | E299 | L181 | TYR  |     |
| GLY  | D837 | ASP  | I639 |      |      | H300 | T185 | GLU  |     |
| GLY  | S838 | LEU  | V640 | ARG  | LA08 | D307 | E186 |      |     |
| GLN  | M839 | GLY  | D722 | E641 |      |      |      | I90  |     |
| ARG  |      |      |      |      | LA12 |      | S187 |      |     |
| THR  | V841 |      | E642 | ASP  |      | I211 |      | M101 |     |
|      |      |      | E643 | GLY  |      |      |      | H402 |     |
|      |      |      |      |      |      |      |      |      |     |





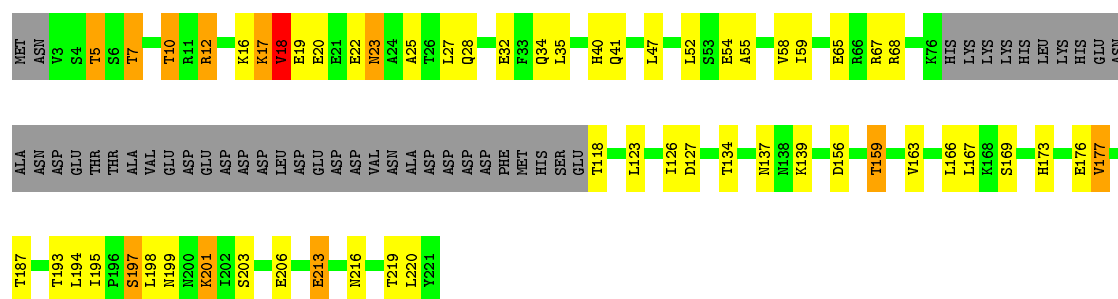
• Molecule 3: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB3

Chain C: 60% 19% 16%



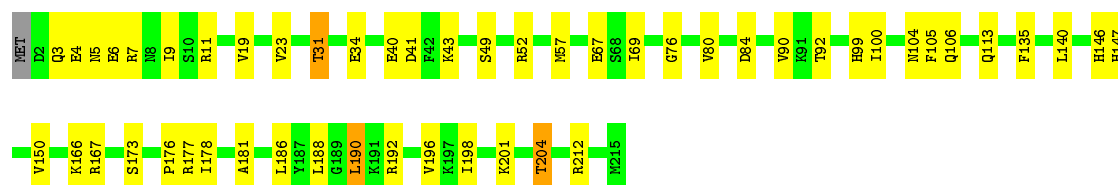
• Molecule 4: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB4

Chain D: 54% 21% 5% 19%



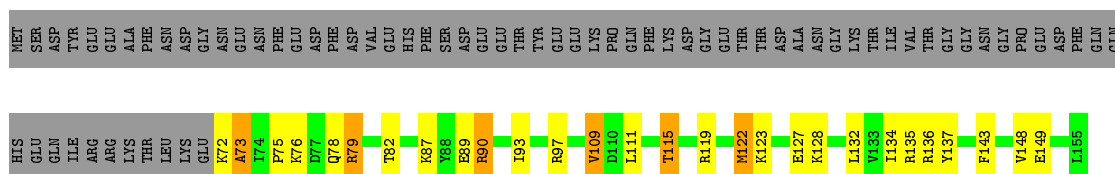
• Molecule 5: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 1

Chain E: 76% 22%



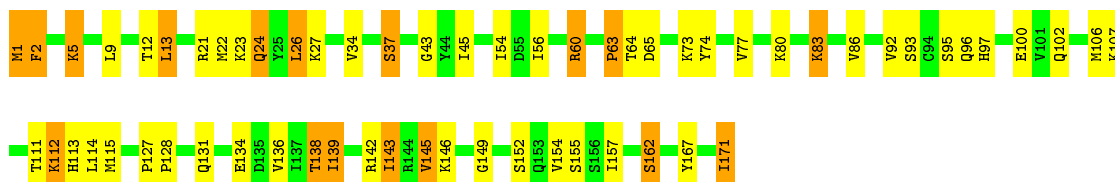
• Molecule 6: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 2

Chain F: 36% 14% 46%



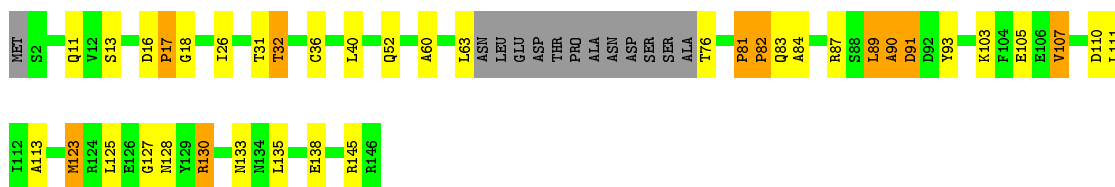
- Molecule 7: RPB7, DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB7

Chain G: 64% 26% 10%



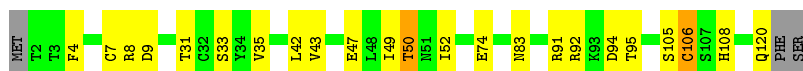
- Molecule 8: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 3

Chain H: 65% 19% 7% 9%



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

Chain I: 79% 17% 4%



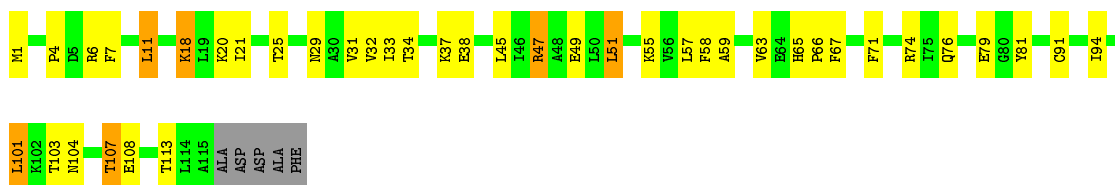
- Molecule 10: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5

Chain J: 53% 34% 6% 7%




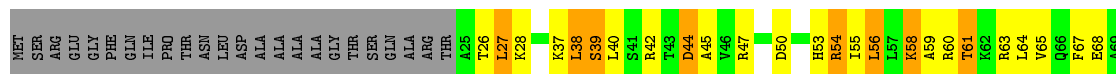
- Molecule 11: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB11

Chain K: 62% 29% 5% 4%



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

Chain L:  29% 26% 11% 34%



R70

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

Chain N:  64% 27% 9%



● Molecule 14: 5'-R(\*AP\*CP\*CP\*AP\*GP\*GP\*AP)-3'

Chain P:  100%

There are no outlier residues recorded for this chain.

● Molecule 15: 5'-D(\*AP\*GP\*CP\*TP\*CP\*AP\*AP\*GP\*TP\*AP\*CP\*TP\*TP\*DTP \*TP\*TP\*C  
P\*CP\*BRU\*GP\*GP\*TP\*CP\*AP\*TP\*T)-3'

Chain T:  54% 19% 27%



## 4 Data and refinement statistics

| Property  | Value   | Source           |
|---|---|------------------|
| Space group   | C 2 2 21  | Depositor        |
| Cell constants<br>a, b, c, $\alpha$ , $\beta$ , $\gamma$                | 221.91Å 391.36Å 283.92Å<br>90.00° 90.00° 90.00°                               | Depositor        |
| Resolution (Å)  | 48.95 – 3.50<br>49.11 – 3.37  | Depositor<br>EDS |
| % Data completeness<br>(in resolution range)                            | 100.0 (48.95-3.50)<br>98.2 (49.11-3.37)                                       | Depositor<br>EDS |
| $R_{merge}$   | (Not available)   | Depositor        |
| $R_{sym}$   | 0.12  | Depositor        |
| $\langle I/\sigma(I) \rangle$ <sup>1</sup>                              | 2.49 (at 3.40Å)   | Xtriage          |
| Refinement program  | BUSTER 2.11.2   | Depositor        |
| R, $R_{free}$   | 0.158 , 0.185<br>0.220 , 0.233  | Depositor<br>DCC |
| $R_{free}$ test set   | 3353 reflections (1.98%)  | wwPDB-VP         |
| Wilson B-factor (Å <sup>2</sup> )                                       | 97.8  | Xtriage          |
| Anisotropy  | 0.391   | Xtriage          |
| Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> ) | 0.31 , 91.2   | EDS              |
| L-test for twinning <sup>2</sup>  | $\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$                   | Xtriage          |
| Estimated twinning fraction   | 0.019 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l<br>0.028 for 1/2*h+1/2*k,3/2*h-1/2*k,-l | Xtriage          |
| $F_o, F_c$ correlation  | 0.92  | EDS              |
| Total number of atoms   | 31940   | wwPDB-VP         |
| Average B, all atoms (Å <sup>2</sup> )                                  | 112.0   | wwPDB-VP         |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.85% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: 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         | 1/11374 (0.0%) | 0.84        | 8/15383 (0.1%)  |
| 2   | B     | 0.51         | 0/9029         | 0.80        | 2/12171 (0.0%)  |
| 3   | C     | 0.50         | 0/2133         | 0.80        | 2/2891 (0.1%)   |
| 4   | D     | 0.52         | 0/1444         | 0.82        | 1/1935 (0.1%)   |
| 5   | E     | 0.47         | 0/1788         | 0.72        | 0/2406          |
| 6   | F     | 0.61         | 0/691          | 0.80        | 0/933           |
| 7   | G     | 0.50         | 0/1368         | 0.83        | 0/1844          |
| 8   | H     | 0.50         | 0/1086         | 0.80        | 0/1470          |
| 9   | I     | 0.46         | 0/989          | 0.80        | 0/1331          |
| 10  | J     | 0.56         | 0/541          | 0.83        | 1/727 (0.1%)    |
| 11  | K     | 0.48         | 0/938          | 0.73        | 0/1267          |
| 12  | L     | 0.59         | 0/365          | 1.04        | 0/485           |
| 13  | N     | 1.15         | 0/232          | 1.03        | 0/356           |
| 14  | P     | 1.04         | 0/170          | 0.80        | 0/263           |
| 15  | T     | 1.30         | 0/405          | 1.13        | 1/620 (0.2%)    |
| All | All   | 0.55         | 1/32553 (0.0%) | 0.82        | 15/44082 (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 |
|-----|-------|---------------------|---------------------|
| 1   | A     | 0                   | 2                   |

All (1) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms   | Z    | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|---------|------|-------------|----------|
| 1   | A     | 867 | ILE  | CG1-CD1 | 5.19 | 1.86        | 1.50     |

The worst 5 of 15 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms   | Z    | Observed(°) | Ideal(°) |
|-----|-------|-----|------|---------|------|-------------|----------|
| 1   | A     | 34  | LYS  | C-N-CA  | 8.15 | 142.07      | 121.70   |
| 1   | A     | 399 | HIS  | N-CA-CB | 7.04 | 123.28      | 110.60   |
| 1   | A     | 54  | ASN  | C-N-CA  | 6.02 | 136.75      | 121.70   |
| 1   | A     | 35  | ILE  | N-CA-CB | 5.82 | 124.19      | 110.80   |
| 1   | A     | 56  | PRO  | C-N-CA  | 5.70 | 135.96      | 121.70   |

There are no chirality outliers.

All (2) planarity outliers are listed below:

| Mol | Chain | Res | Type | Group             |
|-----|-------|-----|------|-------------------|
| 1   | A     | 34  | LYS  | Mainchain,Peptide |

## 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 1   | A     | 11174 | 0        | 11233    | 201     | 1            |
| 2   | B     | 8859  | 0        | 8901     | 173     | 0            |
| 3   | C     | 2095  | 0        | 2051     | 46      | 0            |
| 4   | D     | 1434  | 0        | 1460     | 24      | 0            |
| 5   | E     | 1752  | 0        | 1776     | 24      | 0            |
| 6   | F     | 679   | 0        | 701      | 21      | 0            |
| 7   | G     | 1340  | 0        | 1357     | 43      | 1            |
| 8   | H     | 1068  | 0        | 1040     | 17      | 0            |
| 9   | I     | 971   | 0        | 927      | 6       | 1            |
| 10  | J     | 532   | 0        | 542      | 15      | 0            |
| 11  | K     | 920   | 0        | 929      | 28      | 0            |
| 12  | L     | 363   | 0        | 386      | 17      | 0            |
| 13  | N     | 207   | 0        | 114      | 4       | 0            |
| 14  | P     | 152   | 0        | 77       | 0       | 0            |
| 15  | T     | 385   | 0        | 216      | 5       | 0            |
| 16  | A     | 2     | 0        | 0        | 0       | 0            |
| 16  | B     | 1     | 0        | 0        | 0       | 0            |
| 16  | C     | 1     | 0        | 0        | 0       | 0            |
| 16  | I     | 2     | 0        | 0        | 0       | 0            |
| 16  | J     | 1     | 0        | 0        | 0       | 0            |
| 16  | L     | 1     | 0        | 0        | 0       | 0            |
| 17  | A     | 1     | 0        | 0        | 0       | 0            |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| All | All   | 31940 | 0        | 31710    | 547     | 2            |

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 547 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:CD1 | 1:A:867:ILE:CG1 | 1.86                     | 1.53              |
| 1:A:53:LEU:HD23 | 1:A:54:ASN:H    | 0.97                     | 1.11              |
| 4:D:40:HIS:HB3  | 7:G:73:LYS:HE3  | 1.34                     | 1.08              |
| 7:G:1:MET:HE2   | 7:G:2:PHE:H     | 1.19                     | 1.06              |
| 1:A:53:LEU:CD2  | 1:A:54:ASN:H    | 1.78                     | 0.97              |

All (2) 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) |
|---------------|-----------------------|--------------------------|-------------------|
| 1:A:904:THR:N | 9:I:33:SER:O[8_555]   | 2.18                     | 0.02              |
| 7:G:95:SER:OG | 7:G:97:HIS:CD2[3_654] | 2.18                     | 0.02              |

## 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     | 1414/1732 (82%) | 1252 (88%) | 112 (8%) | 50 (4%)  | <b>3</b> 27  |
| 2   | B     | 1095/1224 (90%) | 969 (88%)  | 82 (8%)  | 44 (4%)  | <b>3</b> 24  |
| 3   | C     | 264/318 (83%)   | 236 (89%)  | 24 (9%)  | 4 (2%)   | <b>10</b> 45 |
| 4   | D     | 174/221 (79%)   | 154 (88%)  | 14 (8%)  | 6 (3%)   | <b>3</b> 28  |
| 5   | E     | 212/215 (99%)   | 186 (88%)  | 24 (11%) | 2 (1%)   | <b>17</b> 56 |

*Continued on next page...*

Continued from previous page...

| Mol | Chain | Analysed        | Favoured   | Allowed  | Outliers | Percentiles |     |
|-----|-------|-----------------|------------|----------|----------|-------------|-----|
| 6   | F     | 82/155 (53%)    | 77 (94%)   | 4 (5%)   | 1 (1%)   | 13          | 50  |
| 7   | G     | 169/171 (99%)   | 160 (95%)  | 6 (4%)   | 3 (2%)   | 8           | 41  |
| 8   | H     | 129/146 (88%)   | 101 (78%)  | 19 (15%) | 9 (7%)   | 1           | 12  |
| 9   | I     | 117/122 (96%)   | 96 (82%)   | 18 (15%) | 3 (3%)   | 5           | 33  |
| 10  | J     | 63/70 (90%)     | 55 (87%)   | 4 (6%)   | 4 (6%)   | 1           | 14  |
| 11  | K     | 113/120 (94%)   | 108 (96%)  | 5 (4%)   | 0        | 100         | 100 |
| 12  | L     | 44/70 (63%)     | 27 (61%)   | 9 (20%)  | 8 (18%)  | 0           | 2   |
| All | All   | 3876/4564 (85%) | 3421 (88%) | 321 (8%) | 134 (4%) | 3           | 27  |

5 of 134 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1   | A     | 47  | ARG  |
| 1   | A     | 57  | ARG  |
| 1   | A     | 58  | LEU  |
| 1   | A     | 74  | MET  |
| 1   | A     | 193 | ASP  |

### 5.3.2 Protein sidechains ⓘ

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     | 1240/1519 (82%) | 1070 (86%) | 170 (14%) | 3           | 20 |
| 2   | B     | 966/1061 (91%)  | 845 (88%)  | 121 (12%) | 4           | 23 |
| 3   | C     | 234/274 (85%)   | 207 (88%)  | 27 (12%)  | 5           | 26 |
| 4   | D     | 160/200 (80%)   | 131 (82%)  | 29 (18%)  | 1           | 9  |
| 5   | E     | 196/197 (100%)  | 176 (90%)  | 20 (10%)  | 7           | 32 |
| 6   | F     | 74/137 (54%)    | 64 (86%)   | 10 (14%)  | 4           | 21 |
| 7   | G     | 152/152 (100%)  | 128 (84%)  | 24 (16%)  | 2           | 15 |
| 8   | H     | 117/128 (91%)   | 100 (86%)  | 17 (14%)  | 3           | 18 |
| 9   | I     | 113/116 (97%)   | 100 (88%)  | 13 (12%)  | 5           | 26 |

Continued on next page...



*Continued from previous page...*

| Mol | Chain | Analysed        | Rotameric  | Outliers  | Percentiles |    |
|-----|-------|-----------------|------------|-----------|-------------|----|
| 10  | J     | 60/65 (92%)     | 51 (85%)   | 9 (15%)   | 3           | 17 |
| 11  | K     | 99/102 (97%)    | 85 (86%)   | 14 (14%)  | 3           | 19 |
| 12  | L     | 40/57 (70%)     | 32 (80%)   | 8 (20%)   | 1           | 7  |
| All | All   | 3451/4008 (86%) | 2989 (87%) | 462 (13%) | 4           | 21 |

5 of 462 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 2   | B     | 398 | ARG  |
| 2   | B     | 879 | ARG  |
| 9   | I     | 83  | ASN  |
| 2   | B     | 452 | THR  |
| 2   | B     | 628 | THR  |

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 42 such sidechains are listed below:

| Mol | Chain | Res  | Type |
|-----|-------|------|------|
| 2   | B     | 465  | ASN  |
| 2   | B     | 1084 | GLN  |
| 8   | H     | 137  | GLN  |
| 2   | B     | 572  | HIS  |
| 2   | B     | 975  | GLN  |

### 5.3.3 RNA ⓘ

| Mol | Chain | Analysed  | Backbone Outliers | Pucker Outliers |
|-----|-------|-----------|-------------------|-----------------|
| 14  | P     | 6/7 (85%) | 0                 | 0               |

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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

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  | 15,14 | 15,21,22     | 1.24 | 2 (13%)  | 17,30,33    | 3.13 | 3 (17%)  |

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 Torsions and Rings columns. '-' means no outliers of that kind were identified.

| Mol | Type | Chain | Res | Link  | Chirals | Torsions  | Rings   |
|-----|------|-------|-----|-------|---------|-----------|---------|
| 15  | BRU  | T     | 22  | 15,14 | -       | 0/4/21/22 | 0/2/2/2 |

All (2) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z     | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|-------|-------|-------------|----------|
| 15  | T     | 22  | BRU  | C4-C5 | 3.19  | 1.42        | 1.38     |
| 15  | T     | 22  | BRU  | C2-N3 | -2.55 | 1.33        | 1.38     |

All (3) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms      | Z     | Observed(°) | Ideal(°) |
|-----|-------|-----|------|------------|-------|-------------|----------|
| 15  | T     | 22  | BRU  | C4-N3-C2   | 11.36 | 124.73      | 115.14   |
| 15  | T     | 22  | BRU  | C5-C4-N3   | -4.26 | 118.54      | 123.64   |
| 15  | T     | 22  | BRU  | C2'-C1'-N1 | -4.10 | 104.82      | 114.27   |

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 9 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

| Mol | Chain | Number of breaks |
|-----|-------|------------------|
| 2   | B     | 2                |

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      | 5.38         |
| 1     | B     | 351:TYR   | C      | 352:ALA   | N      | 3.05         |

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.