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

Feb 18, 2024 – 02:34 PM EST

| | | |
|---|---|---|
| PDB ID | : | 4E2I |
| Title | : | The Complex Structure of the SV40 Helicase Large T Antigen and p68 Subunit of DNA Polymerase Alpha-Primase |
| Authors | : | Zhou, B.; Arnett, D.R.; Yu, X.; Brewster, A.; Sowd, G.A.; Xie, C.L.; Vila, S.; Gai, D.; Fanning, E.; Chen, X.S. |
| Deposited on | : | 2012-03-08 |
| Resolution | : | 5.00 Å(reported) |

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

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

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

The following versions of software and data (see references (1)) were used in the production of this report:

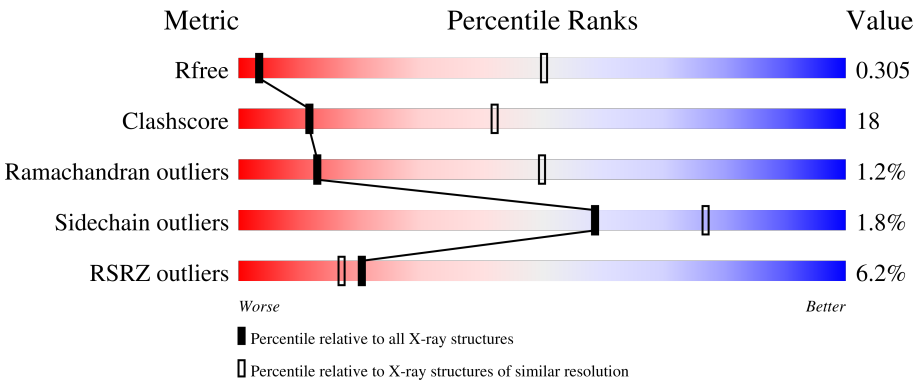
| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Xtriage (Phenix) | : | 1.13 |
| EDS | : | 2.36 |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Refmac | : | 5.8.0158 |
| CCP4 | : | 7.0.044 (Gargrove) |
| Ideal geometry (proteins) | : | Engh & 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 5.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.




| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|---|---|---|
| $R_{free}$ | 130704 | 1140 (6.20-3.80) |
| Clashscore | 141614 | 1000 (6.16-3.82) |
| Ramachandran outliers | 138981 | 1146 (6.20-3.80) |
| Sidechain outliers | 138945 | 1122 (6.20-3.80) |
| RSRZ outliers | 127900 | 1010 (6.22-3.72) |

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 | 362 | 6% 61% 38% |
| 1 | B | 362 | 2% 62% 36% |
| 1 | C | 362 | 2% 62% 37% |
| 1 | D | 362 | 2% 59% 39% |
| 1 | E | 362 | 2% 61% 37% |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 1 | F | 362 | 2% 60% 38% |
| 1 | G | 362 | 6% 61% 37% |
| 1 | H | 362 | 3% 63% 36% |
| 1 | I | 362 | 4% 59% 39% |
| 1 | J | 362 | 3% 60% 38% |
| 1 | K | 362 | 2% 60% 38% |
| 1 | L | 362 | 4% 60% 39% |
| 2 | 1 | 78 | 24% 54% 44% |
| 2 | 2 | 78 | 13% 55% 44% |
| 2 | 3 | 78 | 17% 54% 44% |
| 2 | 4 | 78 | 31% 56% 41% |
| 2 | 5 | 78 | 24% 55% 42% |
| 2 | 6 | 78 | 22% 55% 42% |
| 2 | 7 | 78 | 15% 55% 42% |
| 2 | 8 | 78 | 31% 55% 42% |
| 2 | 9 | 78 | 22% 54% 44% |
| 2 | U | 78 | 23% 53% 45% |
| 2 | W | 78 | 12% 55% 42% |

# 2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 41874 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 Large T antigen.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | | |
| 1 | A | 362 | 2933 | 1888 | 493 | 531 | 21 | 0 | 0 | 0 |
| 1 | B | 362 | 2933 | 1888 | 493 | 531 | 21 | 0 | 0 | 0 |
| 1 | C | 362 | 2933 | 1888 | 493 | 531 | 21 | 0 | 0 | 0 |
| 1 | D | 362 | 2933 | 1888 | 493 | 531 | 21 | 0 | 0 | 0 |
| 1 | E | 362 | 2933 | 1888 | 493 | 531 | 21 | 0 | 0 | 0 |
| 1 | F | 362 | 2933 | 1888 | 493 | 531 | 21 | 0 | 0 | 0 |
| 1 | G | 362 | 2933 | 1888 | 493 | 531 | 21 | 0 | 0 | 0 |
| 1 | H | 362 | 2933 | 1888 | 493 | 531 | 21 | 0 | 0 | 0 |
| 1 | I | 362 | 2933 | 1888 | 493 | 531 | 21 | 0 | 0 | 0 |
| 1 | J | 362 | 2933 | 1888 | 493 | 531 | 21 | 0 | 0 | 0 |
| 1 | K | 362 | 2933 | 1888 | 493 | 531 | 21 | 0 | 0 | 0 |
| 1 | L | 362 | 2933 | 1888 | 493 | 531 | 21 | 0 | 0 | 0 |

- Molecule 2 is a protein called DNA polymerase alpha subunit B.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | | |
| 2 | 2 | 78 | 606 | 380 | 99 | 122 | 5 | 0 | 0 | 0 |
| 2 | 3 | 78 | 606 | 380 | 99 | 122 | 5 | 0 | 0 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | S | | | |
| 2 | 6 | 78 | 606 | 380 | 99 | 122 | 5 | 0 | 0 | 0 |
| 2 | U | 78 | 606 | 380 | 99 | 122 | 5 | 0 | 0 | 0 |
| 2 | W | 78 | 606 | 380 | 99 | 122 | 5 | 0 | 0 | 0 |
| 2 | 5 | 78 | 606 | 380 | 99 | 122 | 5 | 0 | 0 | 0 |
| 2 | 7 | 78 | 606 | 380 | 99 | 122 | 5 | 0 | 0 | 0 |
| 2 | 9 | 78 | 606 | 380 | 99 | 122 | 5 | 0 | 0 | 0 |
| 2 | 1 | 78 | 606 | 380 | 99 | 122 | 5 | 0 | 0 | 0 |
| 2 | 4 | 78 | 606 | 380 | 99 | 122 | 5 | 0 | 0 | 0 |
| 2 | 8 | 78 | 606 | 380 | 99 | 122 | 5 | 0 | 0 | 0 |

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| | | | Total | Zn | | |
| 3 | A | 1 | 1 | 1 | 0 | 0 |
| 3 | B | 1 | 1 | 1 | 0 | 0 |
| 3 | C | 1 | 1 | 1 | 0 | 0 |
| 3 | D | 1 | 1 | 1 | 0 | 0 |
| 3 | E | 1 | 1 | 1 | 0 | 0 |
| 3 | F | 1 | 1 | 1 | 0 | 0 |
| 3 | G | 1 | 1 | 1 | 0 | 0 |
| 3 | H | 1 | 1 | 1 | 0 | 0 |
| 3 | I | 1 | 1 | 1 | 0 | 0 |
| 3 | J | 1 | 1 | 1 | 0 | 0 |

*Continued on next page...*

*Continued from previous page...*

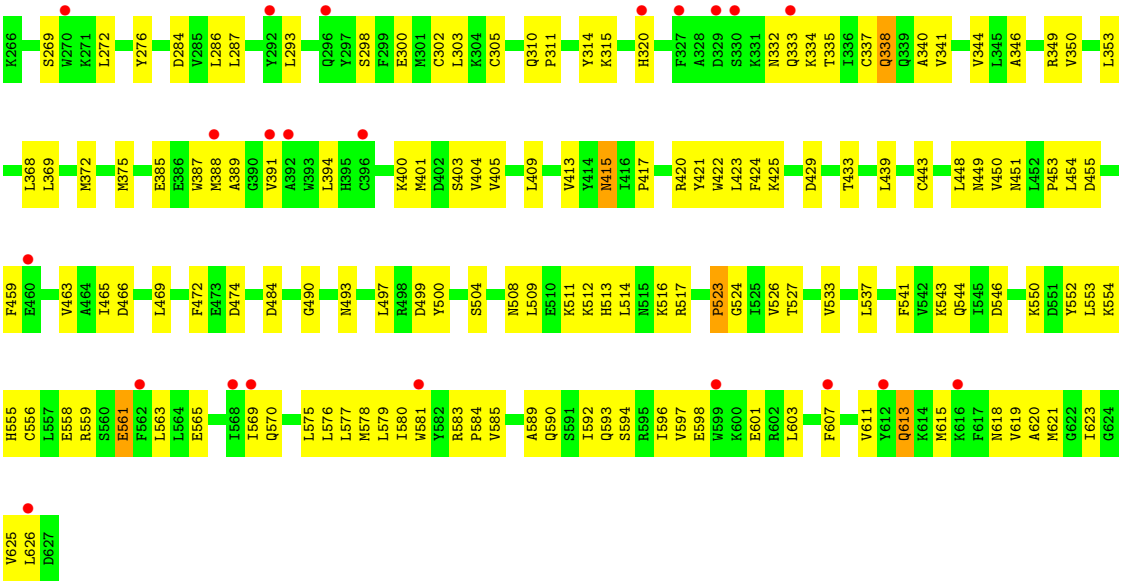
| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| | | | Total | Zn | | |
| 3 | K | 1 | 1 | 1 | 0 | 0 |
| | | | Total | Zn | | |
| 3 | L | 1 | 1 | 1 | 0 | 0 |

# 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.

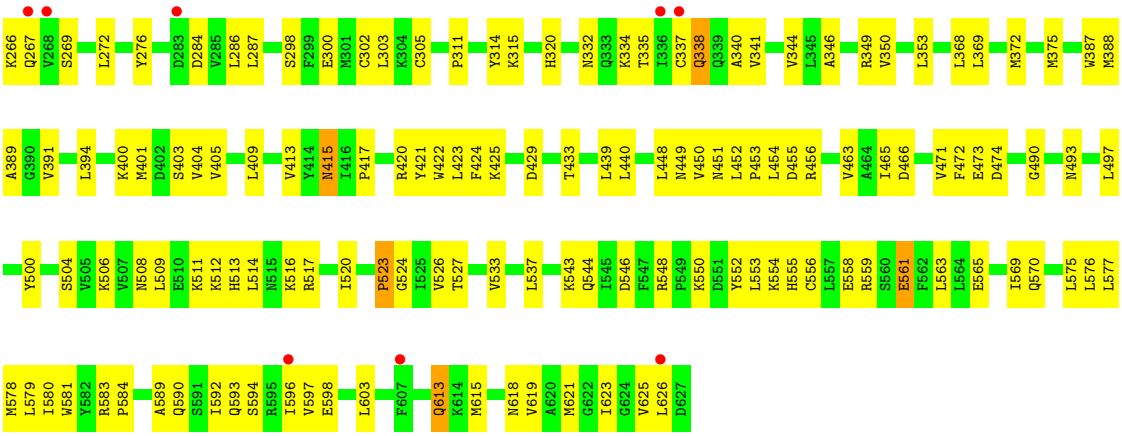
• Molecule 1: Large T antigen

Chain A: 6% 61% 38%




• Molecule 1: Large T antigen

Chain B: 2% 62% 36%



• Molecule 1: Large T antigen

Chain C: 2% 62% 37%

K266 S269 L272 Y276 D284 V285 L286 L287 S298 F299 E300 M301 C302 L303 K304 C305 P311 Y314 K315 E318 K319 H320 F327 N332 Q333 K334 T335 I336 C337 Q338 Q339 A340 V341 D342 T343 V344 L345 A346 R349 V350 L353 Q354 L368 L369 M372

M375 W387 M388 A389 G390 V391 L394 K400 M401 V404 V405 L409 V413 Y414 N415 I416 P417 K418 K419 R420 Y421 W422 L423 F424 K425 D429 T433 L439 L440 C443 L448 N449 V450 N451 L452 P453 L454 D455 V463 A464 I465 D466 L469 V470 V471 F472 E473

D474 G490 N493 L497 Y500 S504 V505 K506 V507 N508 L509 E510 K511 K512 H513 L514 N515 K516 R517 T518 Q519 I520 P523 G524 I525 V526 T527 V533 L537 F541 Q544 T545 D546 K550 D551 Y552 L553 K554 H555 C556 L557 E558 R559 S560 E561 F562 L563 L564 E565

I569 Q570 L575 L576 L577 M578 L579 I580 W581 Y582 R583 P584 A589 Q590 S591 I592 Q593 S594 R595 I596 V597 E598 E601 R602 L603 Q613 K614 M615 N618 V619 A620 M621 G622 I623 G624 V625 L626 D627

• Molecule 1: Large T antigen

Chain D: 2% 59% 39%

K266 S269 L272 Y276 D284 V285 L286 L287 Y292 L293 S298 F299 E300 M301 C302 L303 K304 C305 Q310 P311 Y314 K315 H320 N332 Q333 K334 T335 I336 C337 Q338 Q339 A340 V341 D342 T343 V344 L345 A346 R349 V350 L353 L368 L369 M372 M375

E385 E386 W387 M388 A389 G390 V391 L394 H395 C396 K400 M401 D402 S403 V404 V405 L409 V413 Y414 N415 R420 Y421 W422 L423 F424 K425 D429 K432 T433 A437 A438 L439 C443 G444 G445 K446 A447 L448 N449 V450 N451 L452 P453 L454 D455 V463 A464 I465

L469 F472 E473 D474 G490 N493 L494 D495 N496 L497 R498 D499 Y500 S504 V505 K506 V507 N508 L509 E510 K511 K512 H513 L514 N515 K516 R517 I520 P523 G524 I525 V526 T527 M528 N529 S532 V533 P534 K535 T536 L537 K543 Q544 T545 D546 F547 R548 P549 K550 D551 Y552 L553

K554 H555 C556 L557 E558 R559 S560 E561 F562 L563 L564 E565 I569 Q570 L575 L576 L577 M578 L579 I580 W581 Y582 R583 P584 A589 Q590 S591 I592 Q593 S594 R595 I596 V597 E598 L603 F607 Q613 K614 M615 N618 V619 A620 M621 G622 I623 G624 V625 L626 D627

• Molecule 1: Large T antigen

Chain E: 2% 61% 37%

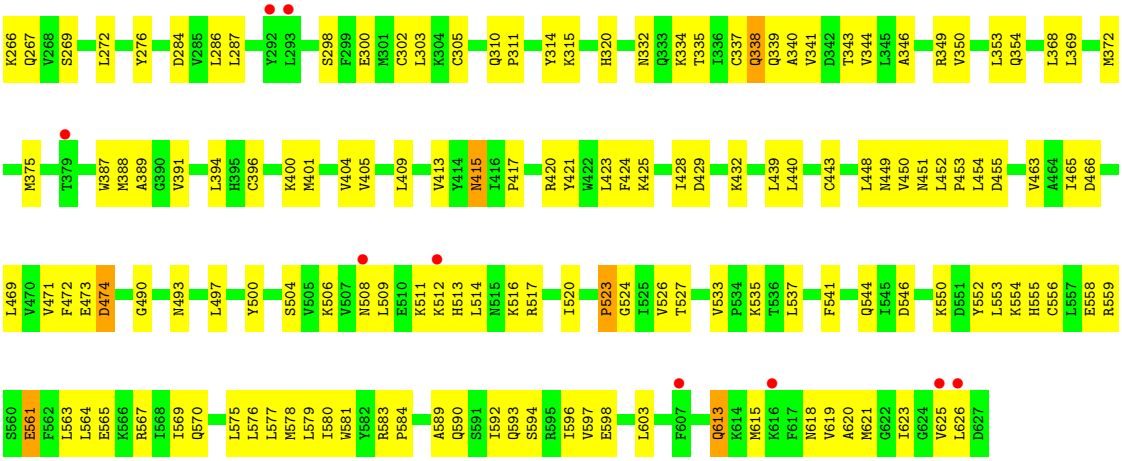
K266 S269 L272 Y276 D284 V285 L286 L287 Y292 L293 S298 F299 E300 M301 C302 L303 K304 C305 P311 Y314 K315 H320 N332 Q333 K334 T335 I336 C337 Q338 Q339 A340 V341 V344 L345 A346 R349 V350 L353 Q354 L368 L369 M372 M375

W387 M388 A389 G390 V391 L394 K400 M401 V404 V405 L409 V413 Y414 N415 I416 P417 K418 K419 R420 Y421 W422 L423 F424 K425 D429 K432 T433 L439 L440 C443 A447 L448 N449 V450 N451 L452 P453 L454 D455 V463 A464 I465 L469 V470 V471 F472 E473 D474

D484 G490 N493 L494 D495 N496 L497 R498 D499 Y500 N508 L509 E510 K511 K512 H513 L514 M515 K516 R517 P523 G524 I525 V526 T527 M528 N529 E530 V533 T536 L537 R540 F541 V542 K543 Q544 T545 D546 F547 R548 P549 K550 D551 Y552 L553 K554 H555 C556 L557 E558 R559 S560

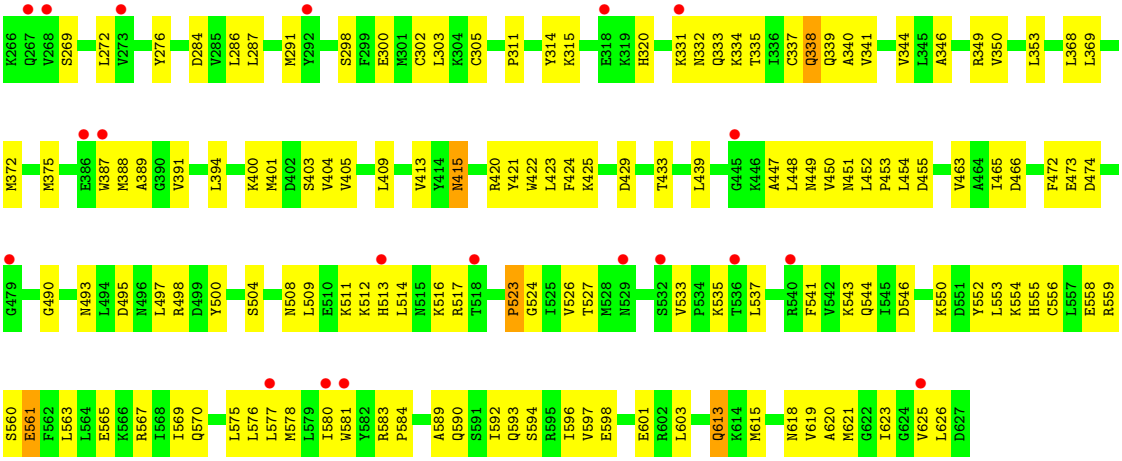
• Molecule 1: Large T antigen

Chain F:



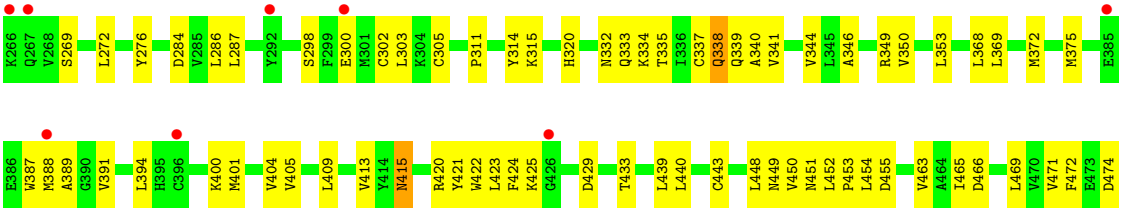

• Molecule 1: Large T antigen

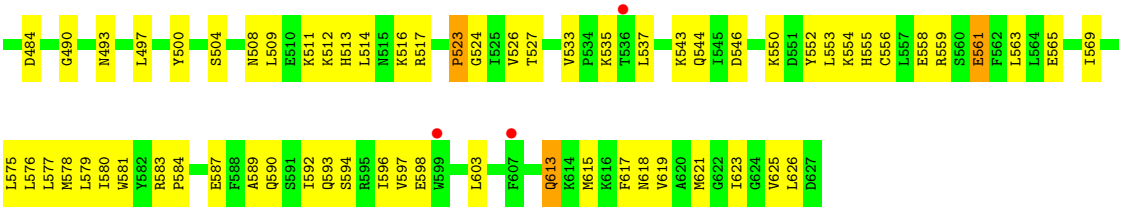
Chain G:




• Molecule 1: Large T antigen

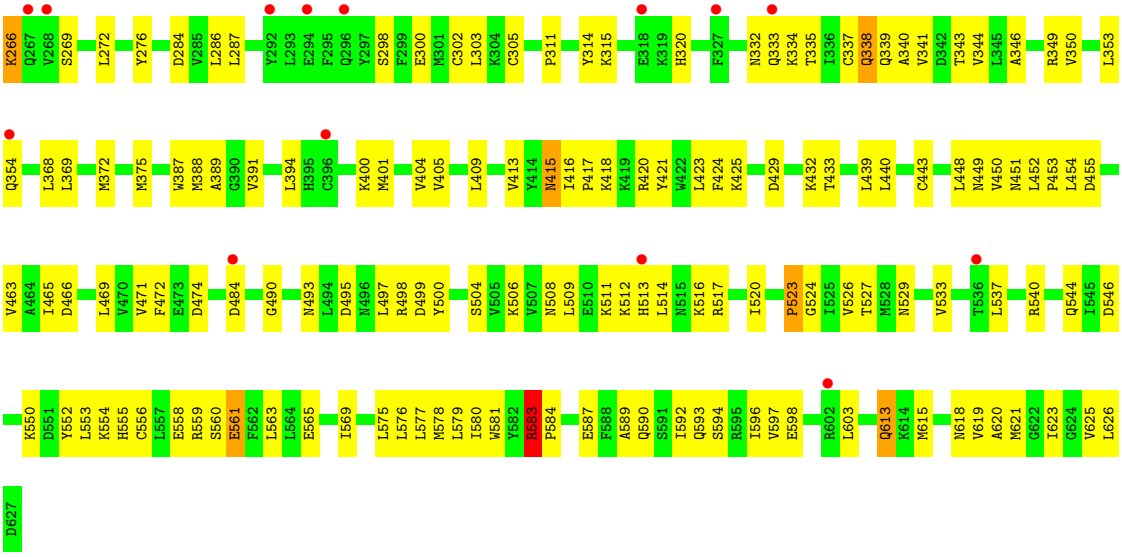
Chain H:





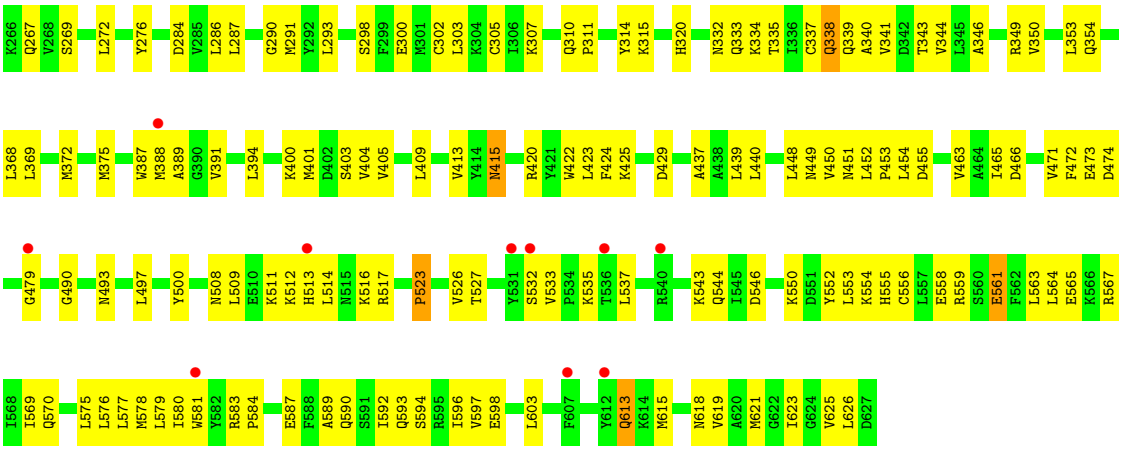

• Molecule 1: Large T antigen

Chain I:



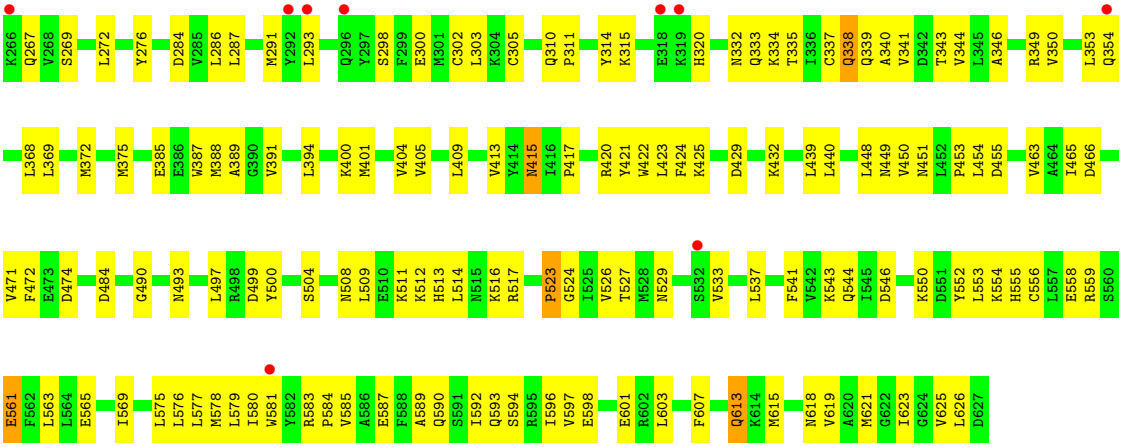

• Molecule 1: Large T antigen

Chain J:



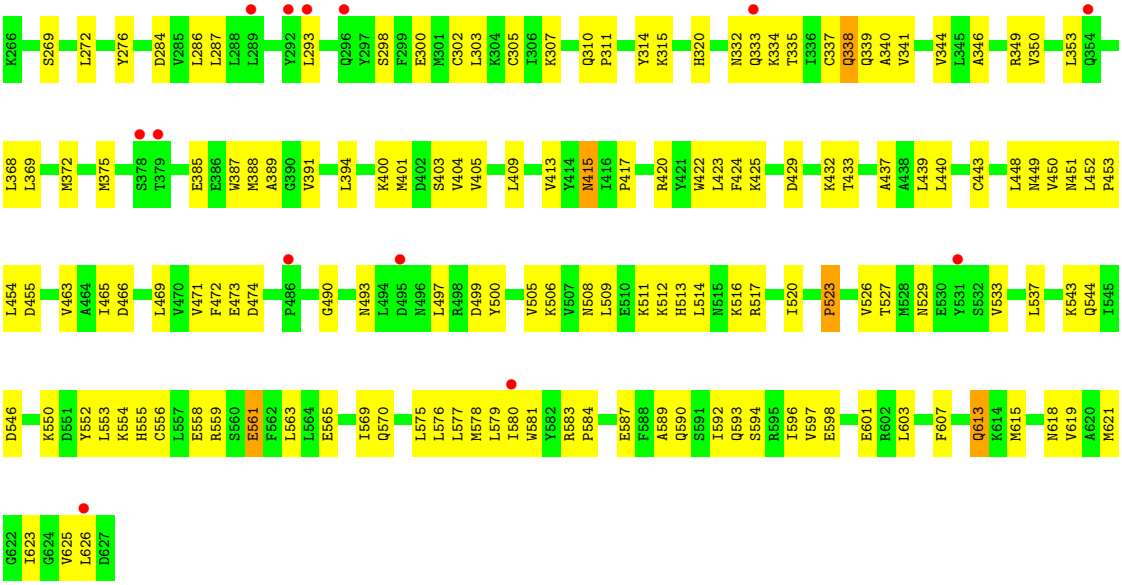

• Molecule 1: Large T antigen

Chain K:



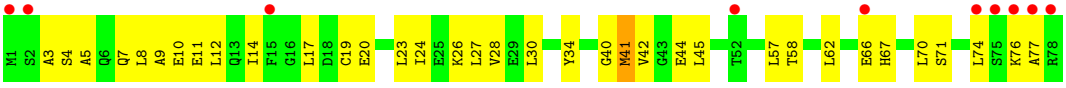

- Molecule 1: Large T antigen

Chain L:



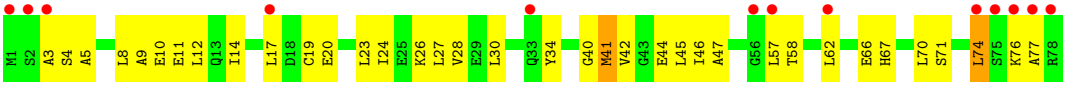

- Molecule 2: DNA polymerase alpha subunit B

Chain 2:



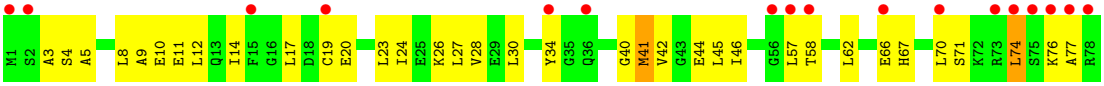

- Molecule 2: DNA polymerase alpha subunit B

Chain 3:



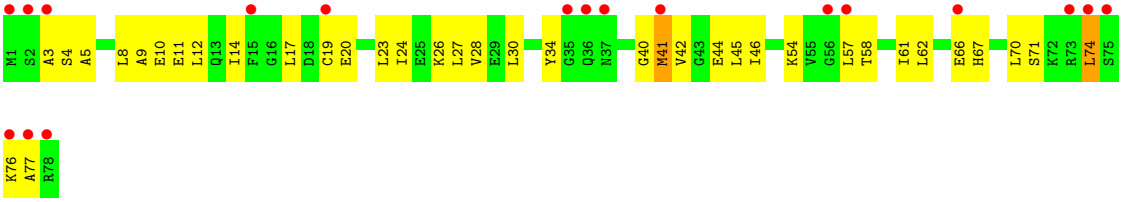
• Molecule 2: DNA polymerase alpha subunit B

Chain 6:



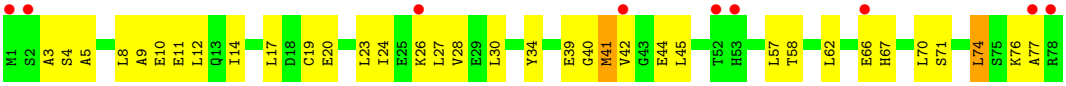

• Molecule 2: DNA polymerase alpha subunit B

Chain U:



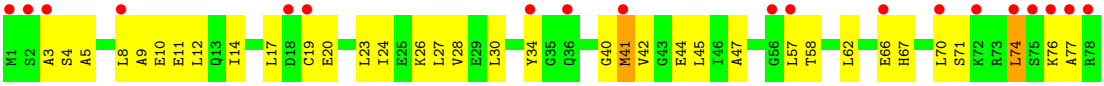

• Molecule 2: DNA polymerase alpha subunit B

Chain W:



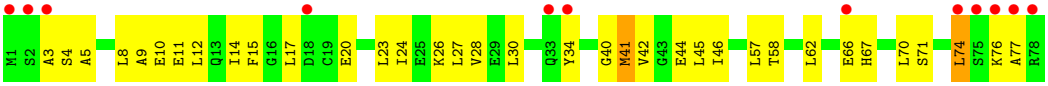

• Molecule 2: DNA polymerase alpha subunit B

Chain 5:



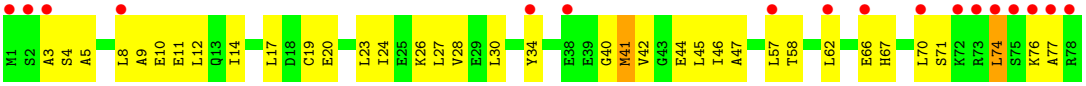

• Molecule 2: DNA polymerase alpha subunit B

Chain 7:



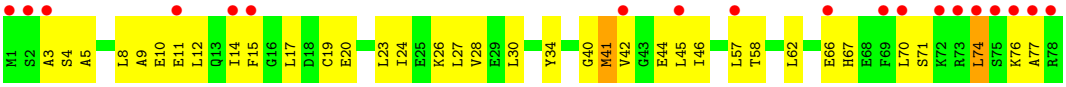

• Molecule 2: DNA polymerase alpha subunit B

Chain 9:



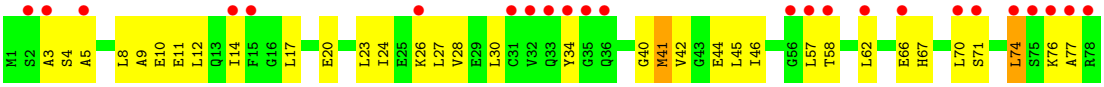

• Molecule 2: DNA polymerase alpha subunit B

Chain 1:



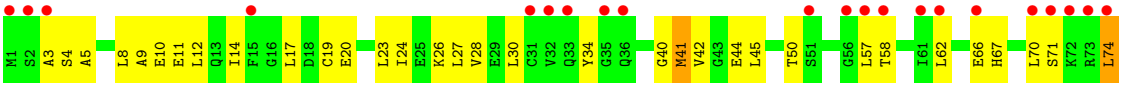

• Molecule 2: DNA polymerase alpha subunit B

Chain 4:




• Molecule 2: DNA polymerase alpha subunit B

Chain 8:



# 4 Data and refinement statistics (i)

| Property | Value | Source |
|---|---|---|
| Space group | P 41 21 2 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 249.10Å 249.10Å 387.03Å<br>90.00° 90.00° 90.00° | Depositor |
| Resolution (Å) | 50.00 – 5.00<br>49.91 – 5.01 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | (Not available) (50.00-5.00)<br>74.4 (49.91-5.01) | Depositor<br>EDS |
| $R_{merge}$ | (Not available) | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 2.40 (at 5.10Å) | Xtriage |
| Refinement program | CNS | Depositor |
| R, $R_{free}$ | 0.305 , 0.314<br>0.299 , 0.305 | Depositor<br>DCC |
| $R_{free}$ test set | 2006 reflections (5.06%) | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 280.1 | Xtriage |
| Anisotropy | 0.181 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.34 , 273.6 | EDS |
| L-test for twinning[2] | $< \|L\| > = 0.47$, $< L^2 > = 0.30$ | Xtriage |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| $F_o$,$F_c$ correlation | 0.89 | EDS |
| Total number of atoms | 41874 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 299.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.94% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

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

# 5 Model quality (i)

## 5.1 Standard geometry (i)

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

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 | 0.25 | 0/2992 | 0.43 | 0/4030 |
| 1 | B | 0.25 | 0/2992 | 0.43 | 0/4030 |
| 1 | C | 0.26 | 0/2992 | 0.63 | 3/4030 (0.1%) |
| 1 | D | 0.25 | 0/2992 | 0.43 | 0/4030 |
| 1 | E | 0.25 | 0/2992 | 0.43 | 0/4030 |
| 1 | F | 0.25 | 0/2992 | 0.43 | 0/4030 |
| 1 | G | 0.25 | 0/2992 | 0.43 | 0/4030 |
| 1 | H | 0.25 | 0/2992 | 0.43 | 0/4030 |
| 1 | I | 0.26 | 0/2992 | 0.63 | 3/4030 (0.1%) |
| 1 | J | 0.25 | 0/2992 | 0.43 | 0/4030 |
| 1 | K | 0.25 | 0/2992 | 0.43 | 0/4030 |
| 1 | L | 0.24 | 0/2992 | 0.43 | 0/4030 |
| 2 | 1 | 0.25 | 0/612 | 0.43 | 0/820 |
| 2 | 2 | 0.26 | 0/612 | 0.44 | 0/820 |
| 2 | 3 | 0.25 | 0/612 | 0.43 | 0/820 |
| 2 | 4 | 0.25 | 0/612 | 0.44 | 0/820 |
| 2 | 5 | 0.26 | 0/612 | 0.43 | 0/820 |
| 2 | 6 | 0.26 | 0/612 | 0.44 | 0/820 |
| 2 | 7 | 0.25 | 0/612 | 0.43 | 0/820 |
| 2 | 8 | 0.25 | 0/612 | 0.43 | 0/820 |
| 2 | 9 | 0.26 | 0/612 | 0.43 | 0/820 |
| 2 | U | 0.25 | 0/612 | 0.43 | 0/820 |
| 2 | W | 0.26 | 0/612 | 0.44 | 0/820 |
| All | All | 0.25 | 0/42636 | 0.46 | 6/57380 (0.0%) |

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 1 | I | 583 | ARG | NE-CZ-NH1 | -20.00 | 110.30 | 120.30 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 1 | C | 583 | ARG | NE-CZ-NH1 | -19.96 | 110.32 | 120.30 |
| 1 | I | 583 | ARG | NE-CZ-NH2 | 19.44 | 130.02 | 120.30 |
| 1 | C | 583 | ARG | NE-CZ-NH2 | 19.40 | 130.00 | 120.30 |
| 1 | C | 583 | ARG | CD-NE-CZ | 8.62 | 135.67 | 123.60 |

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 | 2933 | 0 | 2984 | 105 | 0 |
| 1 | B | 2933 | 0 | 2984 | 101 | 0 |
| 1 | C | 2933 | 0 | 2984 | 103 | 0 |
| 1 | D | 2933 | 0 | 2984 | 112 | 0 |
| 1 | E | 2933 | 0 | 2984 | 105 | 0 |
| 1 | F | 2933 | 0 | 2984 | 116 | 0 |
| 1 | G | 2933 | 0 | 2984 | 112 | 0 |
| 1 | H | 2933 | 0 | 2984 | 97 | 0 |
| 1 | I | 2933 | 0 | 2984 | 116 | 0 |
| 1 | J | 2933 | 0 | 2984 | 112 | 0 |
| 1 | K | 2933 | 0 | 2984 | 111 | 0 |
| 1 | L | 2933 | 0 | 2984 | 125 | 0 |
| 2 | 1 | 606 | 0 | 602 | 29 | 0 |
| 2 | 2 | 606 | 0 | 602 | 30 | 0 |
| 2 | 3 | 606 | 0 | 602 | 28 | 0 |
| 2 | 4 | 606 | 0 | 602 | 30 | 0 |
| 2 | 5 | 606 | 0 | 602 | 27 | 0 |
| 2 | 6 | 606 | 0 | 602 | 34 | 0 |
| 2 | 7 | 606 | 0 | 602 | 28 | 0 |
| 2 | 8 | 606 | 0 | 602 | 27 | 0 |
| 2 | 9 | 606 | 0 | 602 | 29 | 0 |
| 2 | U | 606 | 0 | 602 | 30 | 0 |
| 2 | W | 606 | 0 | 602 | 28 | 0 |
| 3 | A | 1 | 0 | 0 | 0 | 0 |
| 3 | B | 1 | 0 | 0 | 0 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|---|---|---|---|---|---|---|
| 3 | C | 1 | 0 | 0 | 0 | 0 |
| 3 | D | 1 | 0 | 0 | 0 | 0 |
| 3 | E | 1 | 0 | 0 | 0 | 0 |
| 3 | F | 1 | 0 | 0 | 0 | 0 |
| 3 | G | 1 | 0 | 0 | 0 | 0 |
| 3 | H | 1 | 0 | 0 | 0 | 0 |
| 3 | I | 1 | 0 | 0 | 0 | 0 |
| 3 | J | 1 | 0 | 0 | 0 | 0 |
| 3 | K | 1 | 0 | 0 | 0 | 0 |
| 3 | L | 1 | 0 | 0 | 0 | 0 |
| All | All | 41874 | 0 | 42430 | 1479 | 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 18.

The worst 5 of 1479 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:G:286:LEU:HD12 | 1:L:349:ARG:HH21 | 1.16 | 1.07 |
| 1:J:349:ARG:HH21 | 1:K:286:LEU:HD12 | 1.30 | 0.97 |
| 1:D:349:ARG:HH21 | 1:E:286:LEU:HD12 | 1.30 | 0.96 |
| 1:H:349:ARG:HH21 | 1:I:286:LEU:HD12 | 1.35 | 0.92 |
| 1:K:349:ARG:HH21 | 1:L:286:LEU:HD12 | 1.38 | 0.89 |

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 | 360/362 (99%) | 323 (90%) | 33 (9%) | 4 (1%) | 14 52 |
| 1 | B | 360/362 (99%) | 322 (89%) | 34 (9%) | 4 (1%) | 14 52 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed | Favoured | Allowed | Outliers | Percentiles | |
|---|---|---|---|---|---|---|---|
| 1 | C | 360/362 (99%) | 323 (90%) | 33 (9%) | 4 (1%) | 14 | 52 |
| 1 | D | 360/362 (99%) | 323 (90%) | 33 (9%) | 4 (1%) | 14 | 52 |
| 1 | E | 360/362 (99%) | 323 (90%) | 33 (9%) | 4 (1%) | 14 | 52 |
| 1 | F | 360/362 (99%) | 322 (89%) | 33 (9%) | 5 (1%) | 11 | 46 |
| 1 | G | 360/362 (99%) | 323 (90%) | 33 (9%) | 4 (1%) | 14 | 52 |
| 1 | H | 360/362 (99%) | 323 (90%) | 33 (9%) | 4 (1%) | 14 | 52 |
| 1 | I | 360/362 (99%) | 323 (90%) | 33 (9%) | 4 (1%) | 14 | 52 |
| 1 | J | 360/362 (99%) | 322 (89%) | 33 (9%) | 5 (1%) | 11 | 46 |
| 1 | K | 360/362 (99%) | 322 (89%) | 33 (9%) | 5 (1%) | 11 | 46 |
| 1 | L | 360/362 (99%) | 323 (90%) | 33 (9%) | 4 (1%) | 14 | 52 |
| 2 | 1 | 76/78 (97%) | 68 (90%) | 7 (9%) | 1 (1%) | 12 | 48 |
| 2 | 2 | 76/78 (97%) | 68 (90%) | 7 (9%) | 1 (1%) | 12 | 48 |
| 2 | 3 | 76/78 (97%) | 68 (90%) | 7 (9%) | 1 (1%) | 12 | 48 |
| 2 | 4 | 76/78 (97%) | 66 (87%) | 9 (12%) | 1 (1%) | 12 | 48 |
| 2 | 5 | 76/78 (97%) | 68 (90%) | 7 (9%) | 1 (1%) | 12 | 48 |
| 2 | 6 | 76/78 (97%) | 68 (90%) | 7 (9%) | 1 (1%) | 12 | 48 |
| 2 | 7 | 76/78 (97%) | 68 (90%) | 7 (9%) | 1 (1%) | 12 | 48 |
| 2 | 8 | 76/78 (97%) | 67 (88%) | 8 (10%) | 1 (1%) | 12 | 48 |
| 2 | 9 | 76/78 (97%) | 68 (90%) | 7 (9%) | 1 (1%) | 12 | 48 |
| 2 | U | 76/78 (97%) | 68 (90%) | 7 (9%) | 1 (1%) | 12 | 48 |
| 2 | W | 76/78 (97%) | 68 (90%) | 7 (9%) | 1 (1%) | 12 | 48 |
| All | All | 5156/5202 (99%) | 4617 (90%) | 477 (9%) | 62 (1%) | 13 | 50 |

5 of 62 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 474 | ASP |
| 1 | B | 474 | ASP |
| 1 | C | 474 | ASP |
| 1 | D | 474 | ASP |
| 1 | E | 474 | ASP |

### 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 | 323/323 (100%) | 318 (98%) | 5 (2%) | 65 | 80 |
| 1 | B | 323/323 (100%) | 318 (98%) | 5 (2%) | 65 | 80 |
| 1 | C | 323/323 (100%) | 317 (98%) | 6 (2%) | 57 | 75 |
| 1 | D | 323/323 (100%) | 318 (98%) | 5 (2%) | 65 | 80 |
| 1 | E | 323/323 (100%) | 318 (98%) | 5 (2%) | 65 | 80 |
| 1 | F | 323/323 (100%) | 318 (98%) | 5 (2%) | 65 | 80 |
| 1 | G | 323/323 (100%) | 318 (98%) | 5 (2%) | 65 | 80 |
| 1 | H | 323/323 (100%) | 318 (98%) | 5 (2%) | 65 | 80 |
| 1 | I | 323/323 (100%) | 316 (98%) | 7 (2%) | 52 | 71 |
| 1 | J | 323/323 (100%) | 318 (98%) | 5 (2%) | 65 | 80 |
| 1 | K | 323/323 (100%) | 318 (98%) | 5 (2%) | 65 | 80 |
| 1 | L | 323/323 (100%) | 318 (98%) | 5 (2%) | 65 | 80 |
| 2 | 1 | 67/67 (100%) | 65 (97%) | 2 (3%) | 41 | 63 |
| 2 | 2 | 67/67 (100%) | 66 (98%) | 1 (2%) | 65 | 80 |
| 2 | 3 | 67/67 (100%) | 65 (97%) | 2 (3%) | 41 | 63 |
| 2 | 4 | 67/67 (100%) | 65 (97%) | 2 (3%) | 41 | 63 |
| 2 | 5 | 67/67 (100%) | 65 (97%) | 2 (3%) | 41 | 63 |
| 2 | 6 | 67/67 (100%) | 65 (97%) | 2 (3%) | 41 | 63 |
| 2 | 7 | 67/67 (100%) | 65 (97%) | 2 (3%) | 41 | 63 |
| 2 | 8 | 67/67 (100%) | 65 (97%) | 2 (3%) | 41 | 63 |
| 2 | 9 | 67/67 (100%) | 65 (97%) | 2 (3%) | 41 | 63 |
| 2 | U | 67/67 (100%) | 65 (97%) | 2 (3%) | 41 | 63 |
| 2 | W | 67/67 (100%) | 65 (97%) | 2 (3%) | 41 | 63 |
| All | All | 4613/4613 (100%) | 4529 (98%) | 84 (2%) | 59 | 77 |

5 of 84 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | K | 613 | GLN |
| 2 | W | 74 | LEU |
| 1 | L | 338 | GLN |
| 2 | 3 | 74 | LEU |
| 2 | 7 | 74 | LEU |

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

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | J | 415 | ASN |
| 1 | L | 493 | ASN |
| 1 | J | 496 | ASN |
| 1 | K | 451 | ASN |
| 1 | L | 613 | GLN |

#### 5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

### 5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

### 5.6 Ligand geometry (i)

Of 12 ligands modelled in this entry, 12 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)

There are no chain breaks in this entry.

# 6 Fit of model and data (i)

## 6.1 Protein, DNA and RNA chains (i)

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 | 362/362 (100%) | 0.48 | 22 (6%) | 21 | 17 | 262, 300, 300, 300 | 0 |
| 1 | B | 362/362 (100%) | 0.30 | 8 (2%) | 62 | 52 | 260, 300, 300, 300 | 0 |
| 1 | C | 362/362 (100%) | 0.37 | 8 (2%) | 62 | 52 | 260, 300, 300, 300 | 0 |
| 1 | D | 362/362 (100%) | 0.29 | 9 (2%) | 57 | 48 | 257, 300, 300, 300 | 0 |
| 1 | E | 362/362 (100%) | 0.35 | 7 (1%) | 66 | 58 | 260, 300, 300, 300 | 0 |
| 1 | F | 362/362 (100%) | 0.38 | 9 (2%) | 57 | 48 | 258, 300, 300, 300 | 0 |
| 1 | G | 362/362 (100%) | 0.47 | 20 (5%) | 25 | 22 | 260, 300, 300, 300 | 0 |
| 1 | H | 362/362 (100%) | 0.31 | 11 (3%) | 50 | 40 | 260, 300, 300, 300 | 0 |
| 1 | I | 362/362 (100%) | 0.33 | 14 (3%) | 39 | 32 | 260, 300, 300, 300 | 0 |
| 1 | J | 362/362 (100%) | 0.34 | 10 (2%) | 53 | 43 | 261, 300, 300, 300 | 0 |
| 1 | K | 362/362 (100%) | 0.35 | 9 (2%) | 57 | 48 | 260, 300, 300, 300 | 0 |
| 1 | L | 362/362 (100%) | 0.32 | 13 (3%) | 42 | 35 | 262, 300, 300, 300 | 0 |
| 2 | 1 | 78/78 (100%) | 1.24 | 19 (24%) | 0 | 1 | 281, 300, 300, 300 | 0 |
| 2 | 2 | 78/78 (100%) | 0.73 | 10 (12%) | 3 | 5 | 278, 300, 300, 300 | 0 |
| 2 | 3 | 78/78 (100%) | 0.92 | 13 (16%) | 1 | 3 | 280, 300, 300, 300 | 0 |
| 2 | 4 | 78/78 (100%) | 1.38 | 24 (30%) | 0 | 1 | 281, 300, 300, 300 | 0 |
| 2 | 5 | 78/78 (100%) | 1.10 | 19 (24%) | 0 | 1 | 281, 300, 300, 300 | 0 |
| 2 | 6 | 78/78 (100%) | 1.18 | 17 (21%) | 0 | 1 | 281, 300, 300, 300 | 0 |
| 2 | 7 | 78/78 (100%) | 0.88 | 12 (15%) | 2 | 3 | 280, 300, 300, 300 | 0 |
| 2 | 8 | 78/78 (100%) | 1.40 | 24 (30%) | 0 | 1 | 280, 300, 300, 300 | 0 |
| 2 | 9 | 78/78 (100%) | 1.12 | 17 (21%) | 0 | 1 | 280, 300, 300, 300 | 0 |
| 2 | U | 78/78 (100%) | 1.28 | 18 (23%) | 0 | 1 | 280, 300, 300, 300 | 0 |
| 2 | W | 78/78 (100%) | 0.88 | 9 (11%) | 4 | 6 | 279, 300, 300, 300 | 0 |
| All | All | 5202/5202 (100%) | 0.48 | 322 (6%) | 20 | 17 | 257, 300, 300, 300 | 0 |

The worst 5 of 322 RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|---|---|---|---|---|
| 2 | 1 | 1 | MET | 8.4 |
| 2 | 1 | 78 | ARG | 7.2 |
| 2 | 1 | 2 | SER | 6.8 |
| 2 | 6 | 78 | ARG | 6.3 |
| 2 | 5 | 78 | ARG | 6.2 |

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

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

## 6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

## 6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, $95^{th}$ 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 |
|---|---|---|---|---|---|---|---|---|
| 3 | ZN | C | 700 | 1/1 | 0.45 | 0.17 | 300,300,300,300 | 0 |
| 3 | ZN | K | 700 | 1/1 | 0.52 | 0.15 | 300,300,300,300 | 0 |
| 3 | ZN | F | 700 | 1/1 | 0.65 | 0.09 | 300,300,300,300 | 0 |
| 3 | ZN | H | 700 | 1/1 | 0.70 | 0.10 | 300,300,300,300 | 0 |
| 3 | ZN | I | 700 | 1/1 | 0.75 | 0.12 | 300,300,300,300 | 0 |
| 3 | ZN | G | 700 | 1/1 | 0.76 | 0.14 | 300,300,300,300 | 0 |
| 3 | ZN | E | 700 | 1/1 | 0.79 | 0.16 | 300,300,300,300 | 0 |
| 3 | ZN | B | 700 | 1/1 | 0.83 | 0.17 | 300,300,300,300 | 0 |
| 3 | ZN | A | 700 | 1/1 | 0.91 | 0.08 | 300,300,300,300 | 0 |
| 3 | ZN | D | 700 | 1/1 | 0.92 | 0.05 | 245,245,245,245 | 0 |
| 3 | ZN | J | 700 | 1/1 | 0.94 | 0.10 | 300,300,300,300 | 0 |
| 3 | ZN | L | 700 | 1/1 | 0.97 | 0.04 | 300,300,300,300 | 0 |

## 6.5 Other polymers (i)

There are no such residues in this entry.