# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 4, 2024 – 04:59 pm GMT

|  |  |  |
|---|---|---|
| PDB ID | : | 5IYZ |
| Title | : | Tubulin-MMAE complex |
| Authors | : | Waight, A.B.; Bargsten, K.; Doronina, S.; Steinmetz, M.O.; Sussman, D.; Prota, A.E. |
| Deposited on | : | 2016-03-24 |
| Resolution | : | 1.80 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 ⓘ 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 ⓘ) 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 |
| buster-report | : | 1.1.7 (2018) |
| 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 ⓘ

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

The reported resolution of this entry is 1.80 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.
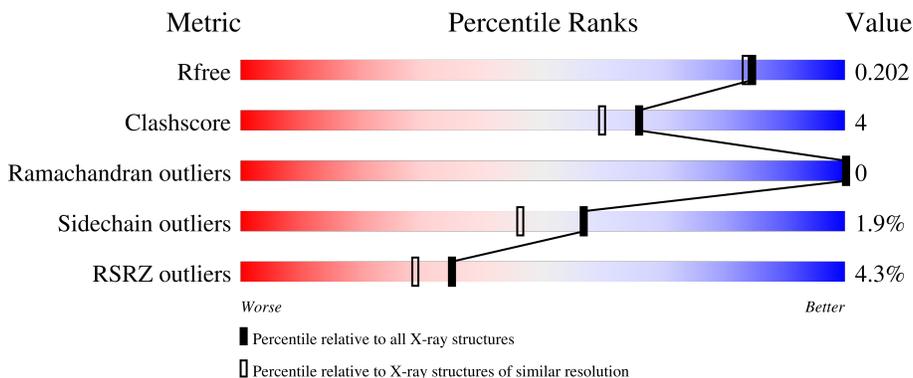


| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|---|---|---|
| $R_{free}$ | 130704 | 5950 (1.80-1.80) |
| Clashscore | 141614 | 6793 (1.80-1.80) |
| Ramachandran outliers | 138981 | 6697 (1.80-1.80) |
| Sidechain outliers | 138945 | 6696 (1.80-1.80) |
| RSRZ outliers | 127900 | 5850 (1.80-1.80) |

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 | 451 | %  91%  6%  • |
| 1 | C | 451 | 89%  8%  • |
| 2 | B | 445 | %  87%  10%  • |
| 2 | D | 445 | 3%  84%  10%  5% |
| 3 | E | 143 | 3%  74%  12%  14% |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|------------------|
| 4 | F | 384 | 18%       73%       13%   •   14% |

# 2　Entry composition ⓘ

There are 12 unique types of molecules in this entry. The entry contains 19140 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 Tubulin alpha-1B chain.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---|---------|---------|-------|
| 1 | A | 437 | Total 3467 | C 2206 | N 582 | O 654 | S 25 | 0 | 12 | 0 |
| 1 | C | 440 | Total 3482 | C 2208 | N 585 | O 664 | S 25 | 0 | 11 | 0 |

- Molecule 2 is a protein called Tubulin beta-2B chain.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---|---------|---------|-------|
| 2 | B | 430 | Total 3447 | C 2169 | N 586 | O 663 | S 29 | 3 | 14 | 0 |
| 2 | D | 424 | Total 3347 | C 2101 | N 570 | O 648 | S 28 | 0 | 3 | 0 |

- Molecule 3 is a protein called Stathmin-4.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---|---------|---------|-------|
| 3 | E | 123 | Total 1047 | C 646 | N 189 | O 207 | S 5 | 0 | 6 | 0 |

There are 2 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------|-----------|
| E | 3 | MET | - | initiating methionine | UNP P63043 |
| E | 4 | ALA | - | expression tag | UNP P63043 |

- Molecule 4 is a protein called TUBULIN-TYROSINE LIGASE.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---|---------|---------|-------|
| 4 | F | 331 | Total 2725 | C 1757 | N 461 | O 493 | S 14 | 0 | 3 | 0 |

There are 6 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------|-----------|
| F | 379 | HIS | - | expression tag | UNP E1BQ43 |
| F | 380 | HIS | - | expression tag | UNP E1BQ43 |
| F | 381 | HIS | - | expression tag | UNP E1BQ43 |
| F | 382 | HIS | - | expression tag | UNP E1BQ43 |
| F | 383 | HIS | - | expression tag | UNP E1BQ43 |
| F | 384 | HIS | - | expression tag | UNP E1BQ43 |

- Molecule 5 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).
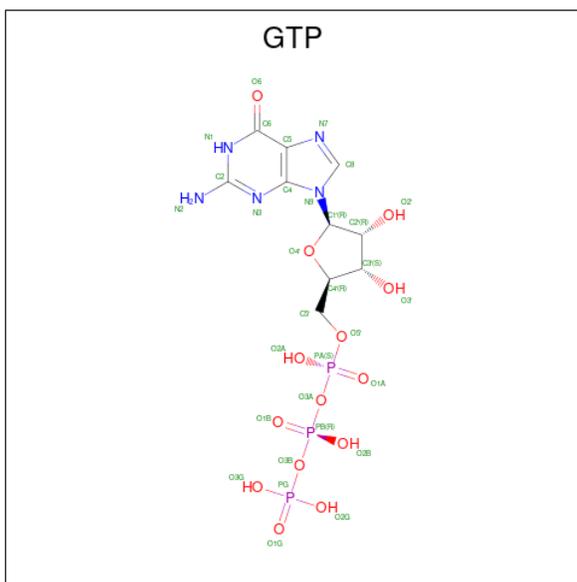


| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---|---|---|---------|---------|
| 5 | A | 1 | Total | C | N | O | P | 0 | 0 |
| | | | 32 | 10 | 5 | 14 | 3 | | |
| 5 | C | 1 | Total | C | N | O | P | 0 | 0 |
| | | | 32 | 10 | 5 | 14 | 3 | | |

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---------|---------|
| 6 | A | 1 | Total | Mg | 0 | 0 |
| | | | 1 | 1 | | |
| 6 | B | 1 | Total | Mg | 0 | 0 |
| | | | 1 | 1 | | |
| 6 | C | 1 | Total | Mg | 0 | 0 |
| | | | 1 | 1 | | |
| 6 | D | 1 | Total | Mg | 0 | 0 |
| | | | 1 | 1 | | |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| 6 | F | 1 | Total<br>1 | Mg<br>1 | 0 | 0 |

- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| 7 | A | 1 | Total<br>1 | Ca<br>1 | 0 | 0 |
| 7 | C | 1 | Total<br>1 | Ca<br>1 | 0 | 0 |

- Molecule 8 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: $C_{10}H_{15}N_5O_{11}P_2$).
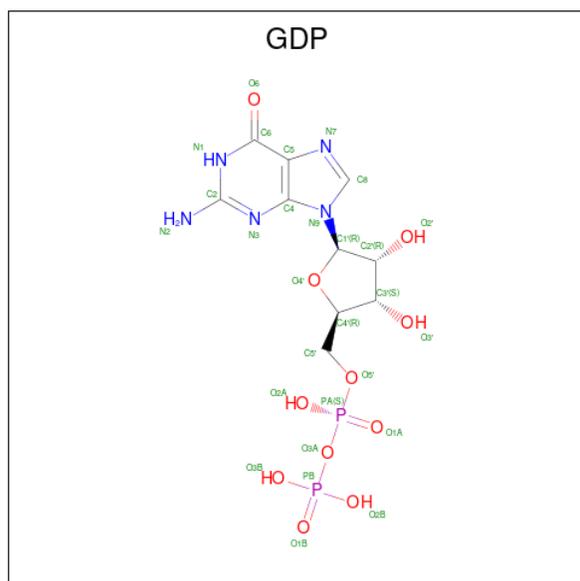


| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|---|---|
| 8 | B | 1 | Total<br>28 | C<br>10 | N<br>5 | O<br>11 | P<br>2 | 0 | 0 |
| 8 | D | 1 | Total<br>28 | C<br>10 | N<br>5 | O<br>11 | P<br>2 | 0 | 0 |

- Molecule 9 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$).

MES

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---|---|---|---------|---------|
| | | | Total | C | N | O | S | | |
| 9 | B | 1 | 12 | 6 | 1 | 4 | 1 | 0 | 0 |

- Molecule 10 is N-methyl-L-valyl-N-[(3R,4S,5S)-1-{(2S)-2-[(1R,2R)-3-{[(1S,2R)-1-hydroxy-1-phenylpropan-2-yl]amino}-1-methoxy-2-methyl-3-oxopropyl]pyrrolidin-1-yl}-3-methoxy-5-methyl-1-oxoheptan-4-yl]-N-methyl-L-valinamide (three-letter code: 4Q5) (formula: $C_{39}H_{67}N_5O_7$).



4Q5

| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---|---|---------|---------|
| | | | Total | C | N | O | | |
| 10 | B | 1 | 51 | 39 | 5 | 7 | 0 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---|---|---------|---------|
| 10 | D | 1 | Total<br>51 | C<br>39 | N<br>5 | O<br>7 | 0 | 0 |

- Molecule 11 is PHOSPHOMETHYLPHOSPHONIC ACID ADENYLATE ESTER (three-letter code: ACP) (formula: $C_{11}H_{18}N_5O_{12}P_3$).
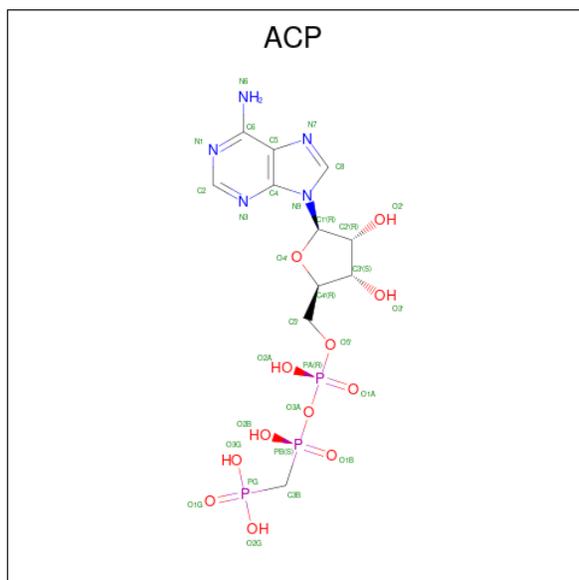


| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---|---|---|---------|---------|
| 11 | F | 1 | Total<br>31 | C<br>11 | N<br>5 | O<br>12 | P<br>3 | 0 | 0 |

- Molecule 12 is water.

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---------|---------|
| 12 | A | 259 | Total<br>259 | O<br>259 | 0 | 0 |
| 12 | B | 314 | Total<br>314 | O<br>314 | 0 | 0 |
| 12 | C | 434 | Total<br>434 | O<br>434 | 0 | 0 |
| 12 | D | 178 | Total<br>178 | O<br>178 | 0 | 0 |
| 12 | E | 67 | Total<br>67 | O<br>67 | 0 | 0 |
| 12 | F | 101 | Total<br>101 | O<br>101 | 0 | 0 |

# 3 Residue-property plots ⓘ

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.

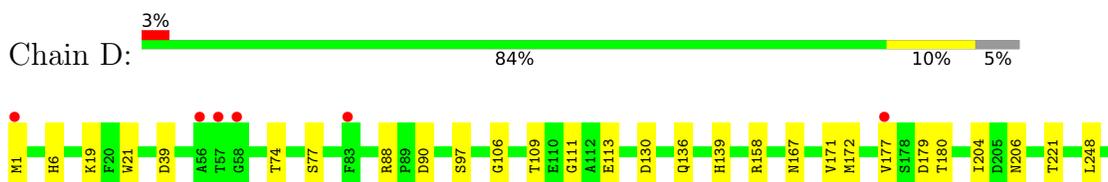- Molecule 1: Tubulin alpha-1B chain
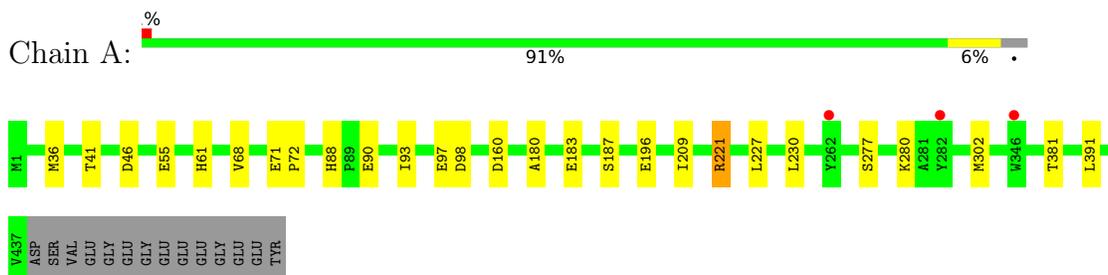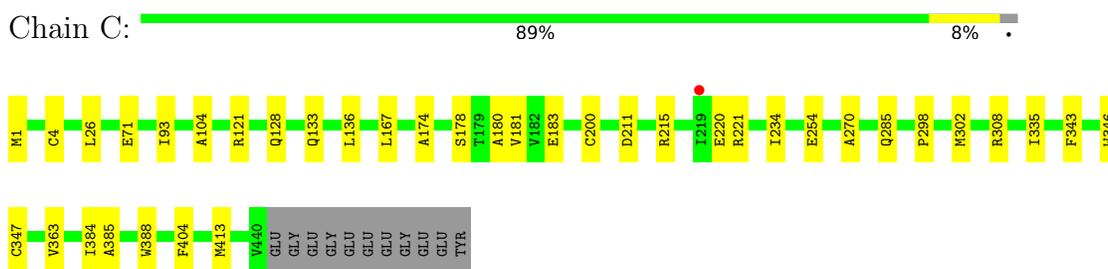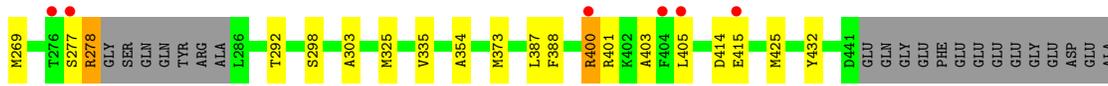
Chain A:



- Molecule 1: Tubulin alpha-1B chain

Chain C:



- Molecule 2: Tubulin beta-2B chain

Chain B:



- Molecule 2: Tubulin beta-2B chain

Chain D:

M269 T276 S277 R278 GLY SER GLN GLN TYR ARG ALA L286 T292 S298 A303 M325 V335 A354 M373 L387 F388 R400 R401 K402 A403 F404 L405 D414 E415 M425 Y432 D441 GLU GLN GLY GLU PHE GLU GLU GLU GLY GLU ASP GLU ALA

• Molecule 3: Stathmin-4

Chain E:

3%

74%　　　　12%　　　14%

MET ALA ASP M6 I9 E10 S28 PHE ASP GLY VAL PRO GLU PHE ASN ALA SER LEU PRO ARG ARG ARG D44 P45 S46 L47 E48 L72 K75 R76 E81 Q84 N90 S107 N108 L123 L139 K140 E141 A142 A143 SER ARG
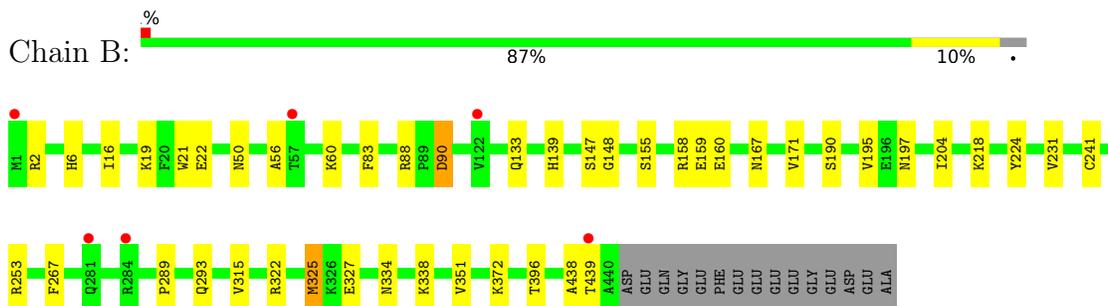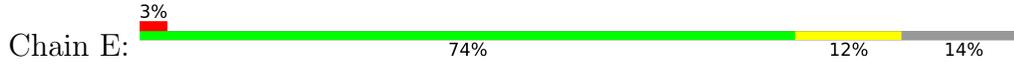
• Molecule 4: TUBULIN-TYROSINE LIGASE

Chain F:

18%

73%　　　　13% •　　14%

M1 R7 M10 E16 R19 L20 T24 G25 R31 M40 R44 L47 S84 L87 S88 E89 S90 C91 V99 I100 Y101 P102 T103 ASN LEU LYS THR PRO VAL ALA PRO ALA GLN ASN GLY ILE ARG HIS LEU ILE ASN ASN THR ARG THR D126 E127 R128

L132 A133 A134 Y135 N136 R137 R138 R139 E140 G141 ARG GLU G144 N145 V146 W147 I148 A149 K150 S151 SER ALA GLY ALA LYS GLY GLU G159 I160 L161 I162 S163 S164 E165 A166 S167 E168 L169 L170 D171 F172 I173 D174 E175 GLN GLY GLN V179 H180 V181 I182 Y185 L186 E187 K188 L192

D200 Y211 R217 V220 L221 R222 T223 S224 S225 E226 P227 A231 ASN PHE GLN ASP LYS T237 C238 H239 L240 T241 N242 H243 C244 I245 Q246 K247 E248 Y249 S250 K251 N252 Y253 G254 R255 Y256 E257 E258 G259 F263 N268 M272 E280 L284 R292 D318 F319 M320

E323 I330 E331 L361 A362 ASP THR GLY GLN LYS THR SER GLN PRO THR S373 H379 H380 HIS HIS HIS HIS

# 4 Data and refinement statistics (i)

| Property | Value | Source |
|---|---|---|
| Space group | P 21 21 21 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 104.51Å   156.61Å   182.42Å<br>90.00°   90.00°   90.00° | Depositor |
| Resolution (Å) | 71.96  –  1.80<br>71.96  –  1.80 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 99.9 (71.96-1.80)<br>99.9 (71.96-1.80) | Depositor<br>EDS |
| $R_{merge}$ | (Not available) | Depositor |
| $R_{sym}$ | 0.06 | Depositor |
| $< I/\sigma(I) >$ [1] | 1.00 (at 1.80Å) | Xtriage |
| Refinement program | PHENIX 1.9_1692 | Depositor |
| R, $R_{free}$ | 0.167  ,  0.202<br>0.168  ,  0.202 | Depositor<br>DCC |
| $R_{free}$ test set | 13782 reflections (5.00%) | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 37.2 | Xtriage |
| Anisotropy | 0.162 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.34 , 52.5 | EDS |
| L-test for twinning[2] | $< |L| > = 0.49, < L^2 > = 0.32$ | Xtriage |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| $F_o,F_c$ correlation | 0.97 | EDS |
| Total number of atoms | 19140 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 54.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.59% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

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

# 5 Model quality ⓘ

## 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 4Q5, MES, GDP, GTP, ACP, CA, MG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

| Mol | Chain | Bond lengths | | Bond angles | |
|-----|-------|------|---------|------|---------|
| | | RMSZ | #$|Z|$>5 | RMSZ | #$|Z|$>5 |
| 1 | A | 0.41 | 0/3582 | 0.54 | 0/4864 |
| 1 | C | 0.55 | 0/3590 | 0.60 | 0/4877 |
| 2 | B | 0.49 | 0/3559 | 0.55 | 0/4817 |
| 2 | D | 0.37 | 0/3426 | 0.50 | 0/4640 |
| 3 | E | 0.43 | 0/1071 | 0.46 | 0/1423 |
| 4 | F | 0.32 | 0/2794 | 0.49 | 0/3772 |
| All | All | 0.44 | 0/18022 | 0.54 | 0/24393 |

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## 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 | 3467 | 0 | 3426 | 17 | 0 |
| 1 | C | 3482 | 0 | 3416 | 22 | 0 |
| 2 | B | 3447 | 0 | 3364 | 34 | 0 |
| 2 | D | 3347 | 0 | 3229 | 31 | 0 |
| 3 | E | 1047 | 0 | 1067 | 10 | 0 |
| 4 | F | 2725 | 0 | 2712 | 31 | 0 |
| 5 | A | 32 | 0 | 12 | 0 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 5 | C | 32 | 0 | 12 | 0 | 0 |
| 6 | A | 1 | 0 | 0 | 0 | 0 |
| 6 | B | 1 | 0 | 0 | 0 | 0 |
| 6 | C | 1 | 0 | 0 | 0 | 0 |
| 6 | D | 1 | 0 | 0 | 0 | 0 |
| 6 | F | 1 | 0 | 0 | 0 | 0 |
| 7 | A | 1 | 0 | 0 | 0 | 0 |
| 7 | C | 1 | 0 | 0 | 0 | 0 |
| 8 | B | 28 | 0 | 12 | 0 | 0 |
| 8 | D | 28 | 0 | 12 | 0 | 0 |
| 9 | B | 12 | 0 | 12 | 4 | 0 |
| 10 | B | 51 | 0 | 0 | 0 | 0 |
| 10 | D | 51 | 0 | 0 | 0 | 0 |
| 11 | F | 31 | 0 | 14 | 3 | 0 |
| 12 | A | 259 | 0 | 0 | 2 | 0 |
| 12 | B | 314 | 0 | 0 | 10 | 0 |
| 12 | C | 434 | 0 | 0 | 5 | 0 |
| 12 | D | 178 | 0 | 0 | 8 | 0 |
| 12 | E | 67 | 0 | 0 | 3 | 0 |
| 12 | F | 101 | 0 | 0 | 3 | 0 |
| All | All | 19140 | 0 | 17288 | 140 | 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 4.

The worst 5 of 140 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) |
|--------|--------|--------------------------|-------------------|
| 4:F:320:MET:HG2 | 4:F:330:ILE:HD11 | 1.60 | 0.83 |
| 1:C:128:GLN:NE2 | 12:C:601:HOH:O | 2.11 | 0.83 |
| 4:F:40:MET:SD | 12:F:595:HOH:O | 2.40 | 0.78 |
| 4:F:200:ASP:OD2 | 4:F:241:THR:OG1 | 2.05 | 0.74 |
| 3:E:6:MET:N | 12:E:202:HOH:O | 2.21 | 0.74 |

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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 | 447/451 (99%) | 439 (98%) | 8 (2%) | 0 | 100 | 100 |
| 1 | C | 449/451 (100%) | 437 (97%) | 12 (3%) | 0 | 100 | 100 |
| 2 | B | 440/445 (99%) | 435 (99%) | 5 (1%) | 0 | 100 | 100 |
| 2 | D | 423/445 (95%) | 417 (99%) | 6 (1%) | 0 | 100 | 100 |
| 3 | E | 125/143 (87%) | 125 (100%) | 0 | 0 | 100 | 100 |
| 4 | F | 320/384 (83%) | 314 (98%) | 6 (2%) | 0 | 100 | 100 |
| All | All | 2204/2319 (95%) | 2167 (98%) | 37 (2%) | 0 | 100 | 100 |

There are no Ramachandran outliers to report.

### 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 | 380/379 (100%) | 375 (99%) | 5 (1%) | 69 | 62 |
| 1 | C | 382/379 (101%) | 378 (99%) | 4 (1%) | 76 | 71 |
| 2 | B | 385/383 (100%) | 380 (99%) | 5 (1%) | 69 | 62 |
| 2 | D | 370/383 (97%) | 362 (98%) | 8 (2%) | 52 | 39 |
| 3 | E | 116/127 (91%) | 112 (97%) | 4 (3%) | 37 | 22 |
| 4 | F | 301/342 (88%) | 292 (97%) | 9 (3%) | 41 | 27 |
| All | All | 1934/1993 (97%) | 1899 (98%) | 35 (2%) | 57 | 48 |

5 of 35 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 4 | F | 91 | CYS |
| 4 | F | 161 | LEU |
| 4 | F | 192 | LEU |
| 1 | C | 335 | ILE |
| 1 | C | 220 | GLU |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 2 | B | 334 | ASN |

### 5.3.3　RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5　Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 5.6　Ligand geometry ⓘ

Of 15 ligands modelled in this entry, 7 are monoatomic - leaving 8 for Mogul analysis.

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 | #$|Z| > 2$ |
| 8 | GDP | B | 501 | 6 | 24,30,30 | 1.11 | 2 (8%) | 30,47,47 | 1.38 | 4 (13%) |
| 9 | MES | B | 503 | - | 12,12,12 | 2.30 | 1 (8%) | 14,16,16 | 2.52 | 7 (50%) |
| 10 | 4Q5 | D | 503 | - | 52,52,52 | 1.29 | 2 (3%) | 55,72,72 | 1.94 | 10 (18%) |

| Mol | Type | Chain | Res | Link | Bond lengths | | | Bond angles | | |
|-----|------|-------|-----|------|--------|------|---------|--------|------|---------|
| | | | | | Counts | RMSZ | #\|Z\| > 2 | Counts | RMSZ | #\|Z\| > 2 |
| 5 | GTP | C | 501 | 6 | 26,34,34 | 0.86 | 0 | 32,54,54 | 1.33 | 5 (15%) |
| 11 | ACP | F | 402 | 6 | 27,33,33 | 1.44 | 4 (14%) | 32,52,52 | 1.40 | 4 (12%) |
| 10 | 4Q5 | B | 504 | - | 52,52,52 | 1.47 | 5 (9%) | 55,72,72 | 1.76 | 9 (16%) |
| 5 | GTP | A | 501 | 6 | 26,34,34 | 0.94 | 1 (3%) | 32,54,54 | 1.33 | 3 (9%) |
| 8 | GDP | D | 501 | 6 | 24,30,30 | 0.88 | 1 (4%) | 30,47,47 | 1.16 | 3 (10%) |

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 |
|-----|------|-------|-----|------|---------|----------|-------|
| 8 | GDP | B | 501 | 6 | - | 5/12/32/32 | 0/3/3/3 |
| 9 | MES | B | 503 | - | - | 2/6/14/14 | 0/1/1/1 |
| 10 | 4Q5 | D | 503 | - | - | 2/76/86/86 | 0/2/2/2 |
| 5 | GTP | C | 501 | 6 | - | 5/18/38/38 | 0/3/3/3 |
| 11 | ACP | F | 402 | 6 | - | 4/15/38/38 | 0/3/3/3 |
| 10 | 4Q5 | B | 504 | - | - | 1/76/86/86 | 0/2/2/2 |
| 5 | GTP | A | 501 | 6 | - | 6/18/38/38 | 0/3/3/3 |
| 8 | GDP | D | 501 | 6 | - | 4/12/32/32 | 0/3/3/3 |

The worst 5 of 16 bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|-------|-----|-------------|----------|
| 10 | D | 503 | 4Q5 | CBN-CBM | -7.94 | 1.40 | 1.51 |
| 9 | B | 503 | MES | C8-S | -7.69 | 1.66 | 1.77 |
| 10 | B | 504 | 4Q5 | CBN-CBM | -6.35 | 1.42 | 1.51 |
| 10 | B | 504 | 4Q5 | CBH-CBD | 4.17 | 1.58 | 1.52 |
| 8 | B | 501 | GDP | O4'-C1' | 3.66 | 1.46 | 1.41 |

The worst 5 of 45 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|-----|-------|-----|------|-------|-----|----------------|-------------|
| 10 | B | 504 | 4Q5 | CAW-CAX-CAY | -6.25 | 102.76 | 112.42 |
| 9 | B | 503 | MES | C5-N4-C3 | 6.14 | 122.64 | 108.83 |
| 10 | D | 503 | 4Q5 | CBF-CBE-CBD | -5.99 | 94.09 | 104.16 |
| 10 | D | 503 | 4Q5 | CBG-NAZ-CBD | -5.74 | 102.72 | 111.74 |
| 10 | B | 504 | 4Q5 | CBG-NAZ-CBD | -5.25 | 103.48 | 111.74 |

There are no chirality outliers.

5 of 29 torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|-----|-------|-----|------|-------|
| 5 | A | 501 | GTP | PB-O3B-PG-O2G |
| 5 | A | 501 | GTP | PB-O3B-PG-O3G |
| 5 | A | 501 | GTP | C5'-O5'-PA-O1A |
| 5 | C | 501 | GTP | PB-O3B-PG-O3G |
| 5 | C | 501 | GTP | C5'-O5'-PA-O1A |

There are no ring outliers.
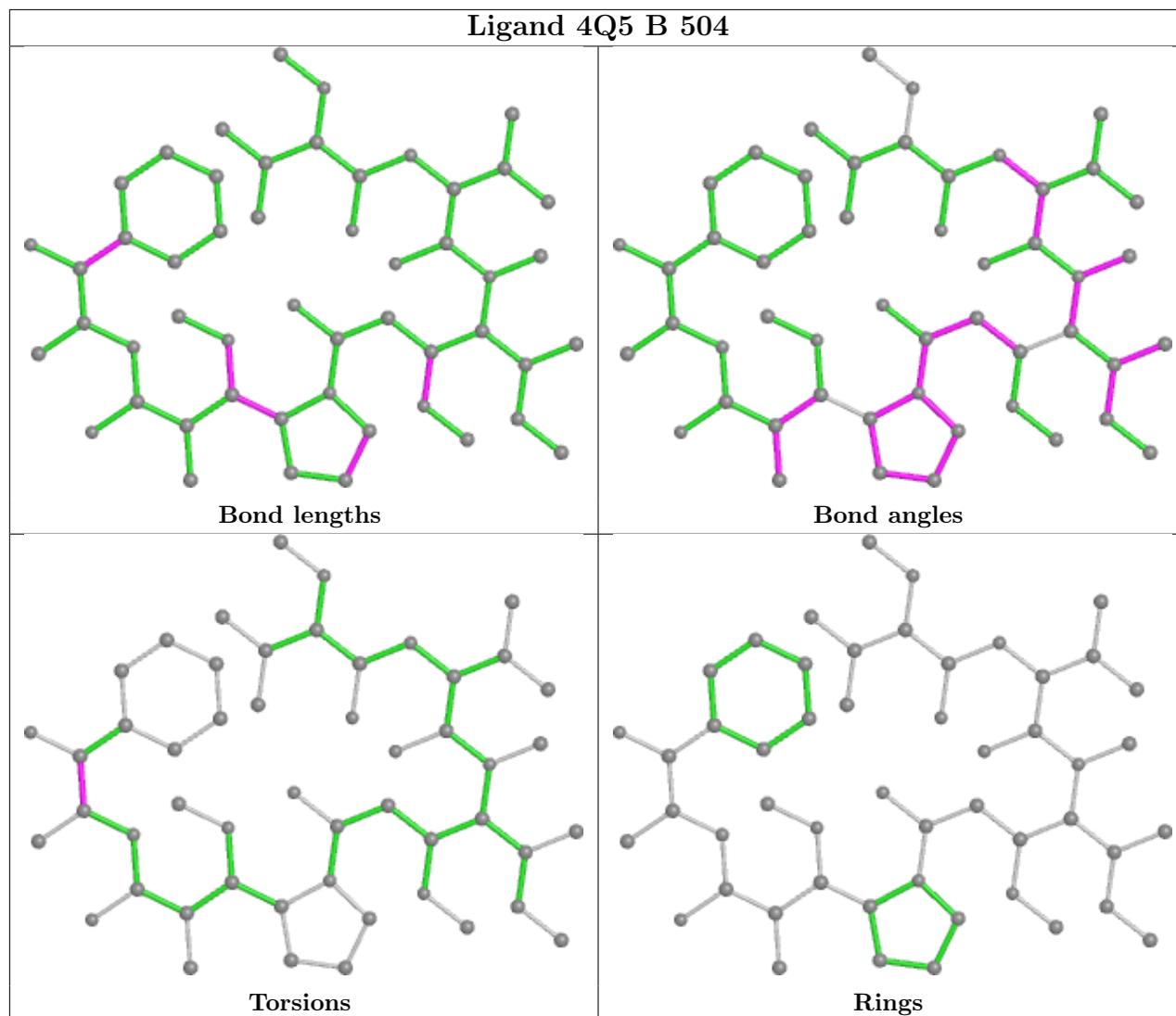
2 monomers are involved in 7 short contacts:

| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|-----|-------|-----|------|---------|--------------|
| 9 | B | 503 | MES | 4 | 0 |
| 11 | F | 402 | ACP | 3 | 0 |

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

## Ligand GDP B 501



Bond lengths



Bond angles



Torsions



Rings

## Ligand 4Q5 D 503



Bond lengths



Bond angles



Torsions



Rings

## Ligand GTP C 501


Bond lengths


Bond angles


Torsions


Rings

## Ligand ACP F 402


Bond lengths


Bond angles


Torsions


Rings

## Ligand 4Q5 B 504



Bond lengths



Bond angles



Torsions



Rings

## Ligand GTP A 501



Bond lengths



Bond angles



Torsions



Rings

## Ligand GDP D 501



Bond lengths



Bond angles



Torsions



Rings
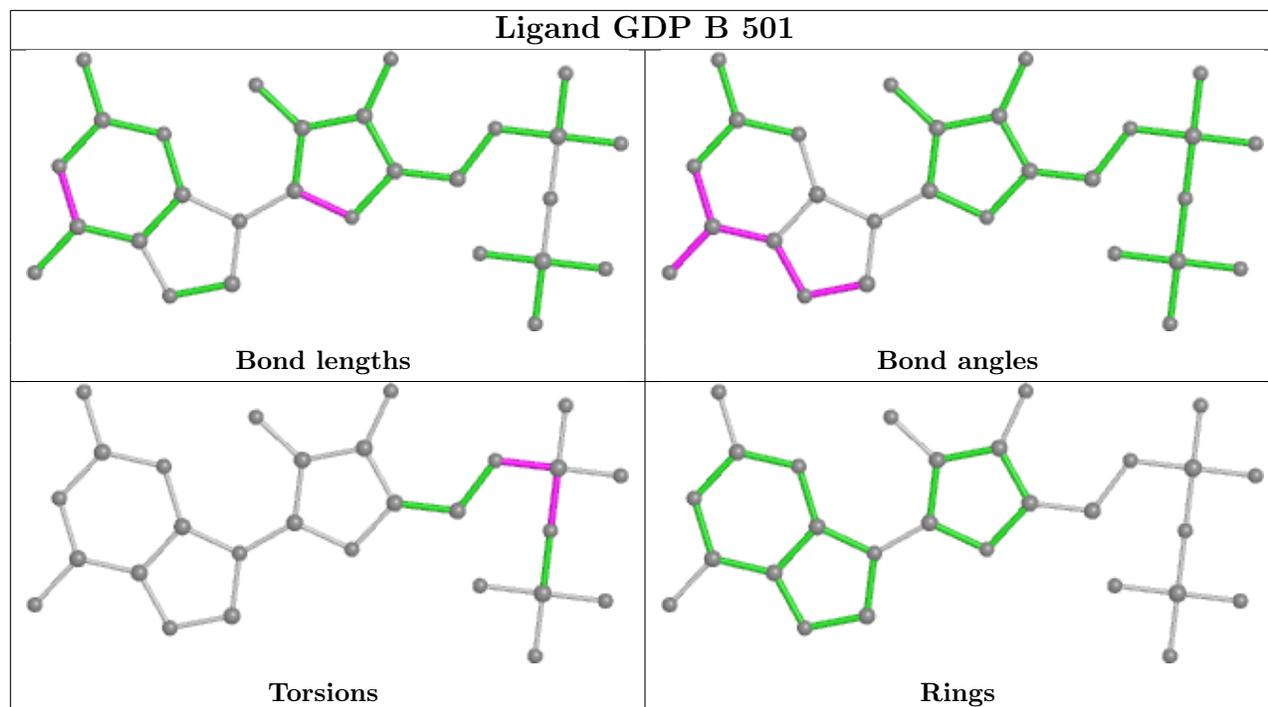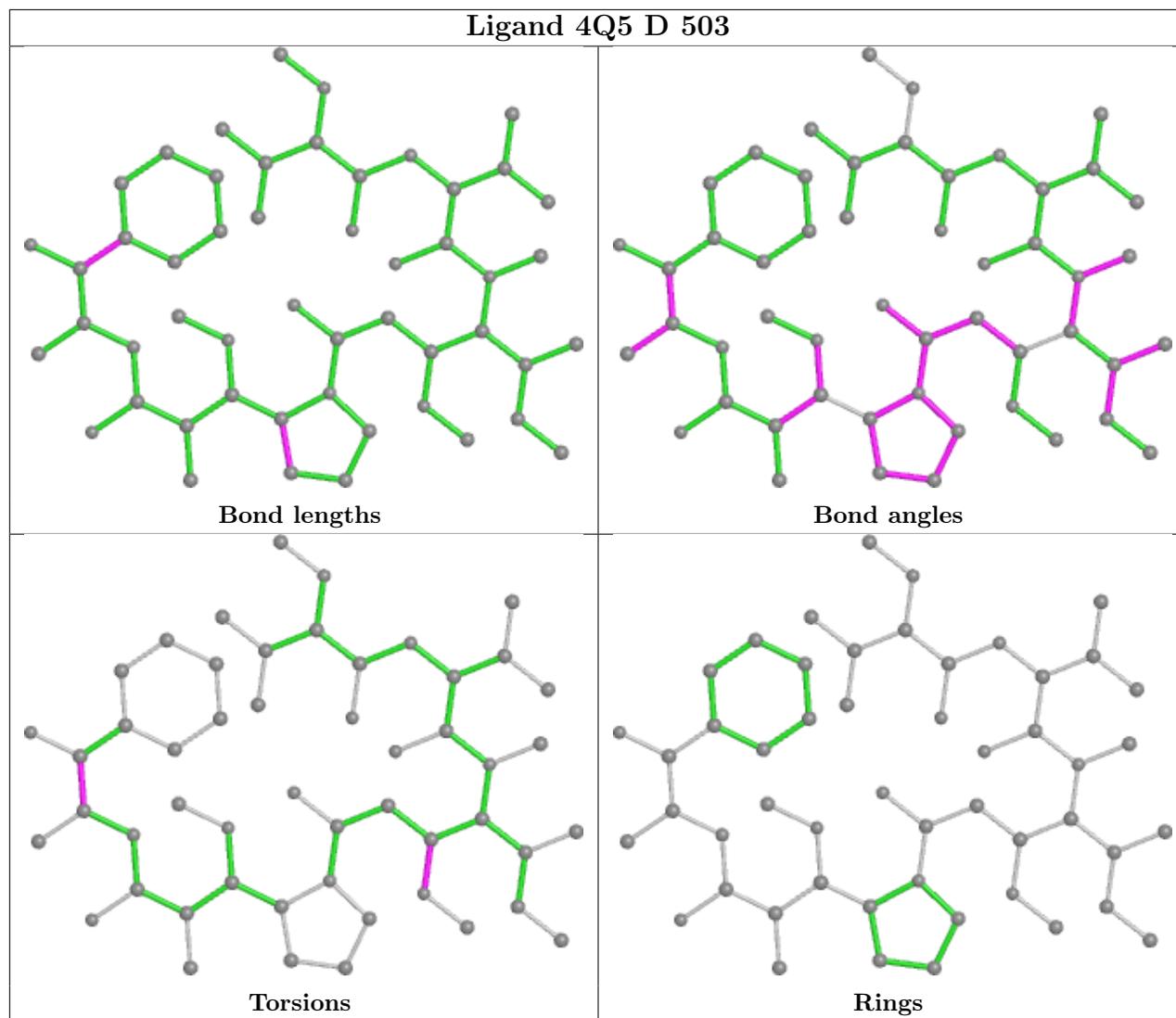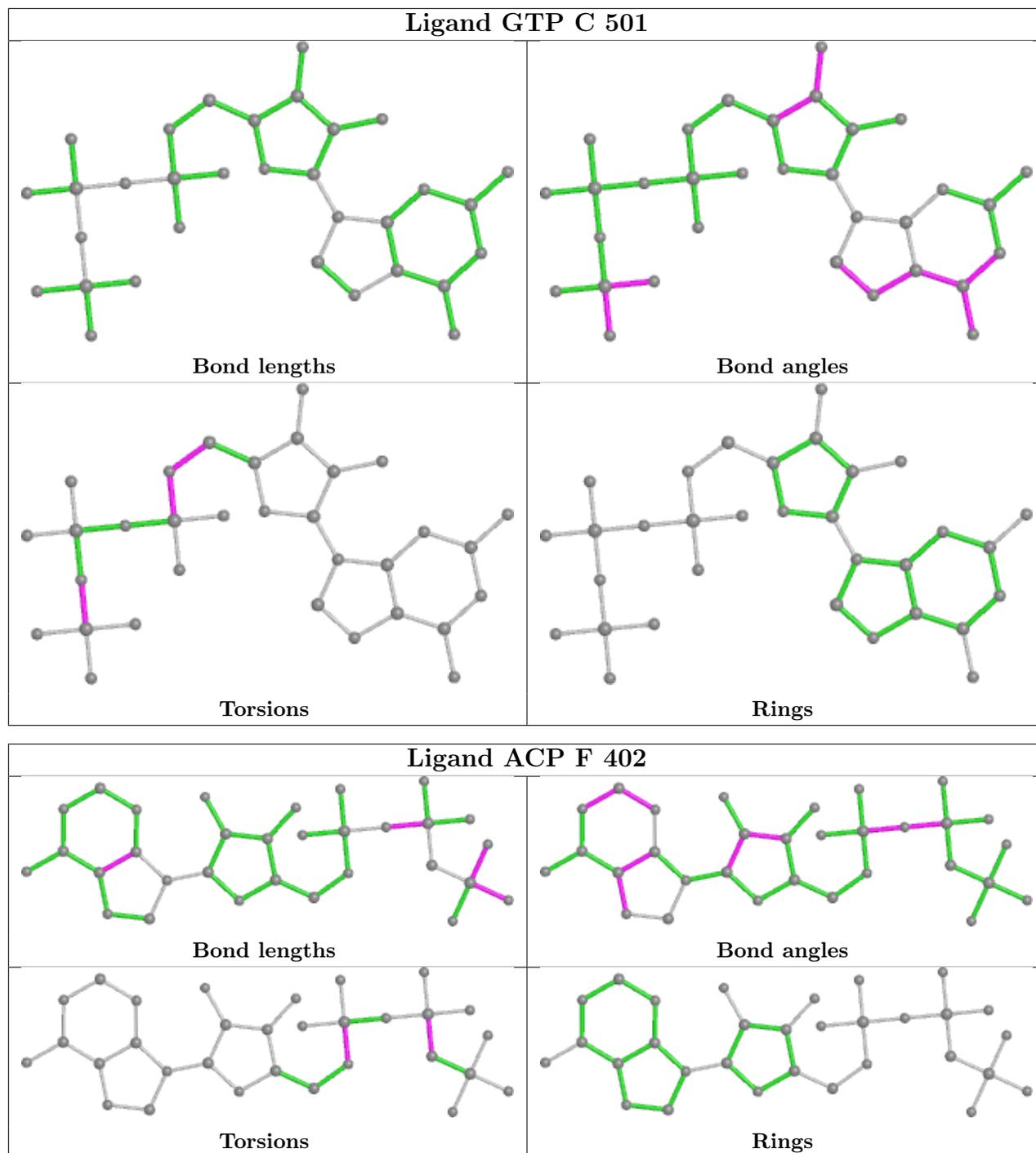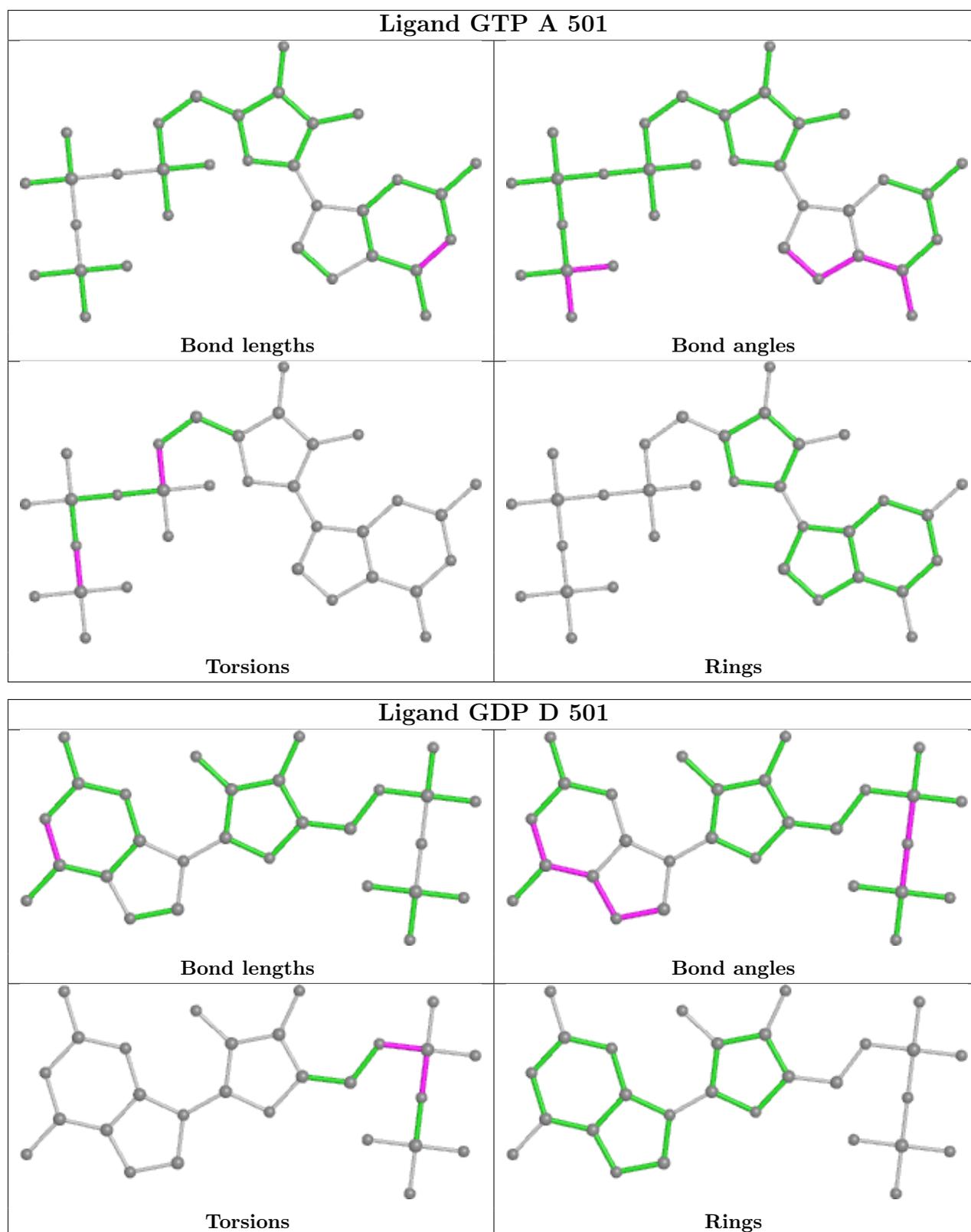
## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

The following chains have linkage breaks:

| Mol | Chain | Number of breaks |
|-----|-------|------------------|
| 2 | B | 1 |

All chain breaks are listed below:

| Model | Chain | Residue-1 | Atom-1 | Residue-2 | Atom-2 | Distance (Å) |
|-------|-------|-----------|--------|-----------|--------|--------------|
| 1 | B | 438:ALA | C | 439:THR | N | 2.74 |

# 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(Å²) | Q<0.9 |
|-----|-------|----------|--------|---------|----|----|----------|-------|
| 1 | A | 437/451 (96%) | -0.20 | 3 (0%) | 87 | 86 | 32, 49, 83, 104 | 0 |
| 1 | C | 440/451 (97%) | -0.12 | 1 (0%) | 95 | 93 | 26, 37, 61, 103 | 0 |
| 2 | B | 430/445 (96%) | -0.07 | 6 (1%) | 75 | 72 | 27, 41, 79, 132 | 2 (0%) |
| 2 | D | 424/445 (95%) | 0.00 | 12 (2%) | 53 | 47 | 33, 54, 86, 121 | 3 (0%) |
| 3 | E | 123/143 (86%) | 0.08 | 4 (3%) | 46 | 40 | 34, 61, 96, 132 | 0 |
| 4 | F | 331/384 (86%) | 0.80 | 68 (20%) | 1 | 0 | 39, 69, 125, 156 | 0 |
| All | All | 2185/2319 (94%) | 0.05 | 94 (4%) | 35 | 29 | 26, 49, 95, 156 | 5 (0%) |

The worst 5 of 94 RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|-----|-------|-----|------|------|
| 4 | F | 161 | LEU | 7.6 |
| 4 | F | 173 | ILE | 7.3 |
| 4 | F | 169 | LEU | 7.1 |
| 4 | F | 249 | TYR | 6.2 |
| 2 | D | 57 | THR | 6.0 |

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

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

## 6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 6.4 Ligands ⓘ

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($\text{Å}^2$) | Q<0.9 |
|-----|------|-------|-----|-------|------|-----|-------------------------|-------|
| 6 | MG | B | 502 | 1/1 | 0.81 | 0.12 | 63,63,63,63 | 0 |
| 6 | MG | D | 502 | 1/1 | 0.81 | 0.13 | 64,64,64,64 | 0 |
| 10 | 4Q5 | D | 503 | 51/51 | 0.82 | 0.34 | 73,106,125,133 | 0 |
| 6 | MG | F | 401 | 1/1 | 0.86 | 0.11 | 78,78,78,78 | 0 |
| 9 | MES | B | 503 | 12/12 | 0.88 | 0.18 | 55,86,118,124 | 0 |
| 11 | ACP | F | 402 | 31/31 | 0.91 | 0.12 | 61,70,108,130 | 0 |
| 7 | CA | C | 503 | 1/1 | 0.93 | 0.31 | 86,86,86,86 | 0 |
| 7 | CA | A | 503 | 1/1 | 0.93 | 0.05 | 67,67,67,67 | 0 |
| 6 | MG | A | 502 | 1/1 | 0.96 | 0.15 | 37,37,37,37 | 0 |
| 10 | 4Q5 | B | 504 | 51/51 | 0.97 | 0.09 | 26,32,39,42 | 0 |
| 8 | GDP | D | 501 | 28/28 | 0.97 | 0.09 | 42,47,57,60 | 0 |
| 6 | MG | C | 502 | 1/1 | 0.97 | 0.14 | 32,32,32,32 | 0 |
| 8 | GDP | B | 501 | 28/28 | 0.99 | 0.12 | 25,28,30,31 | 0 |
| 5 | GTP | C | 501 | 32/32 | 0.99 | 0.12 | 25,30,32,33 | 0 |
| 5 | GTP | A | 501 | 32/32 | 0.99 | 0.09 | 30,33,38,39 | 0 |

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around 4Q5 D 503:**

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
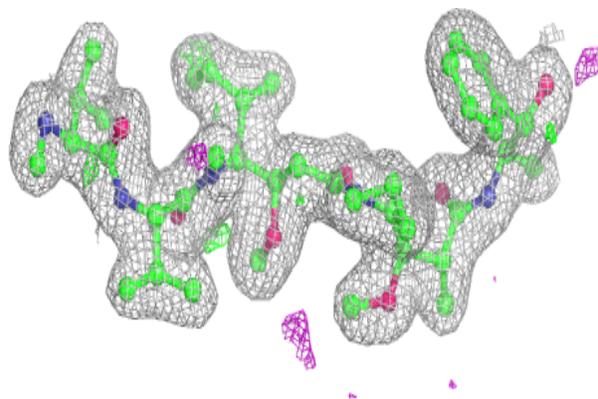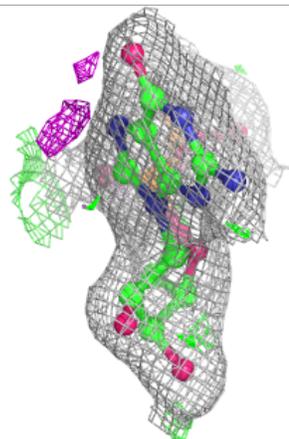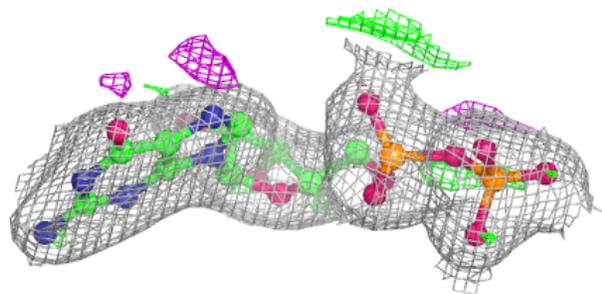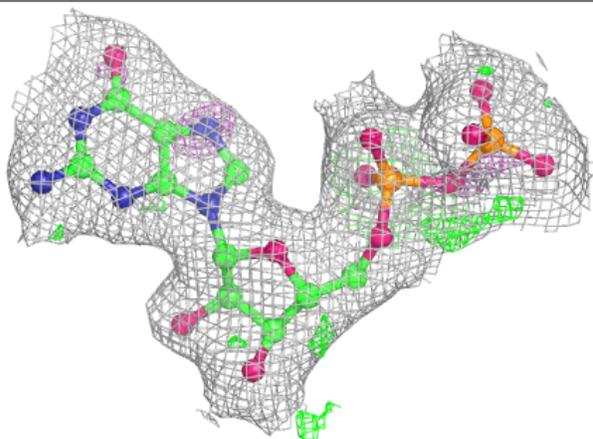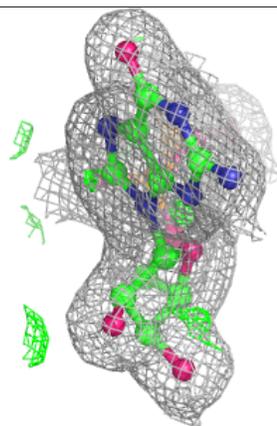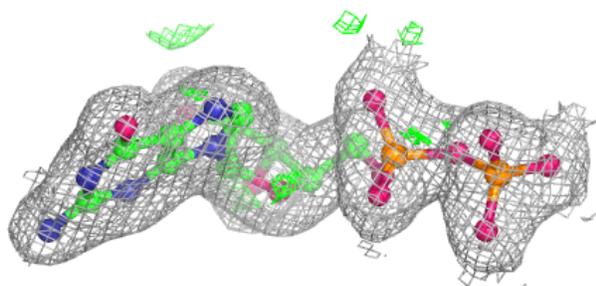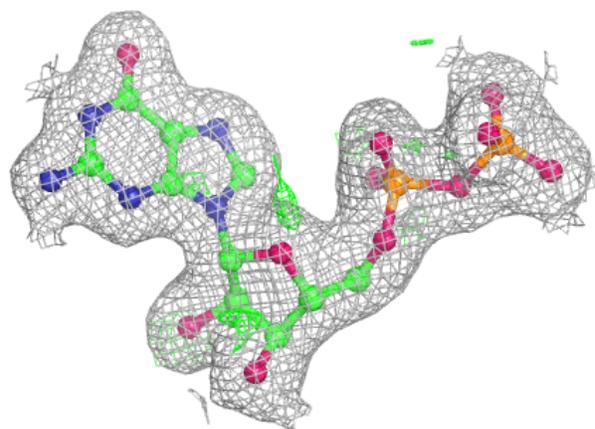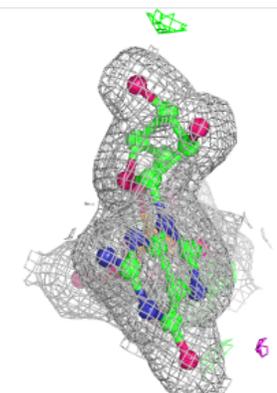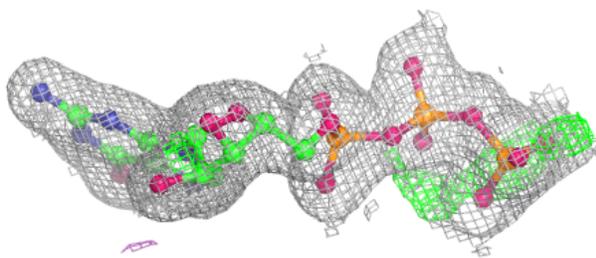mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)
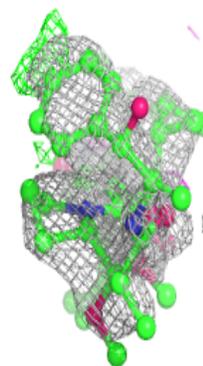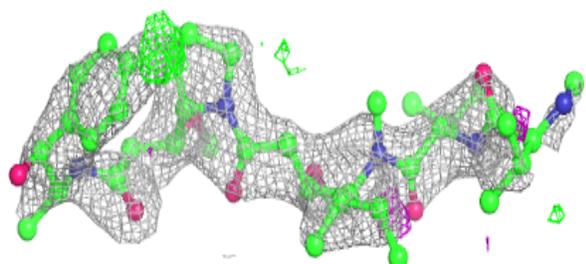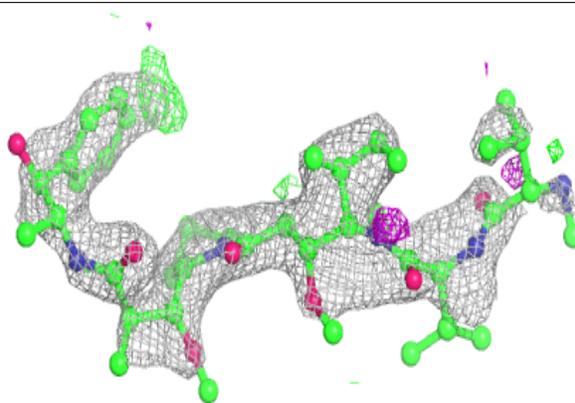


**Electron density around ACP F 402:**

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)
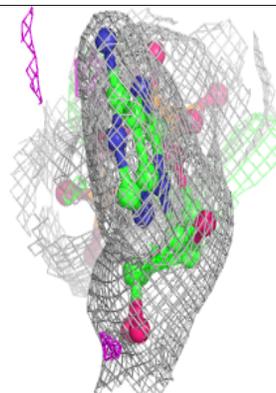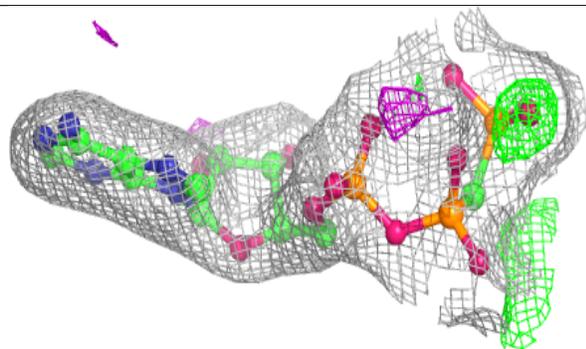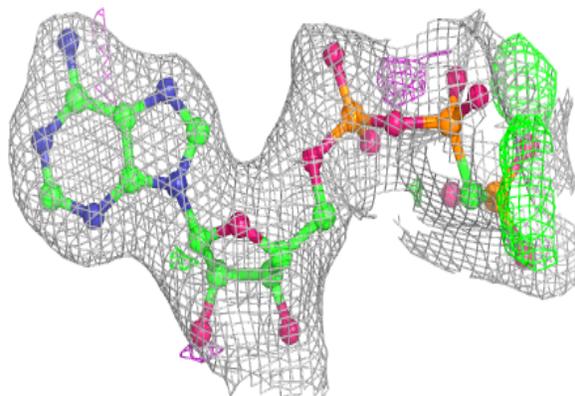
**Electron density around 4Q5 B 504:**

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
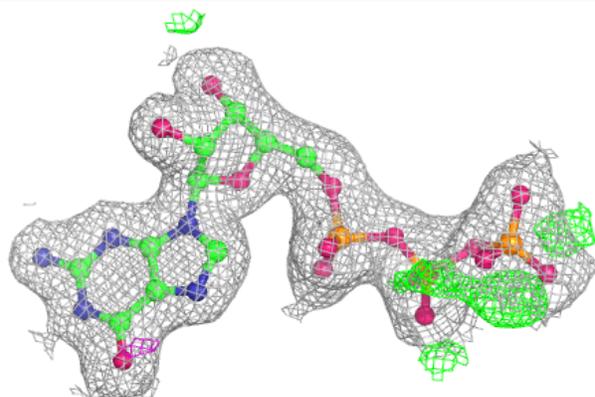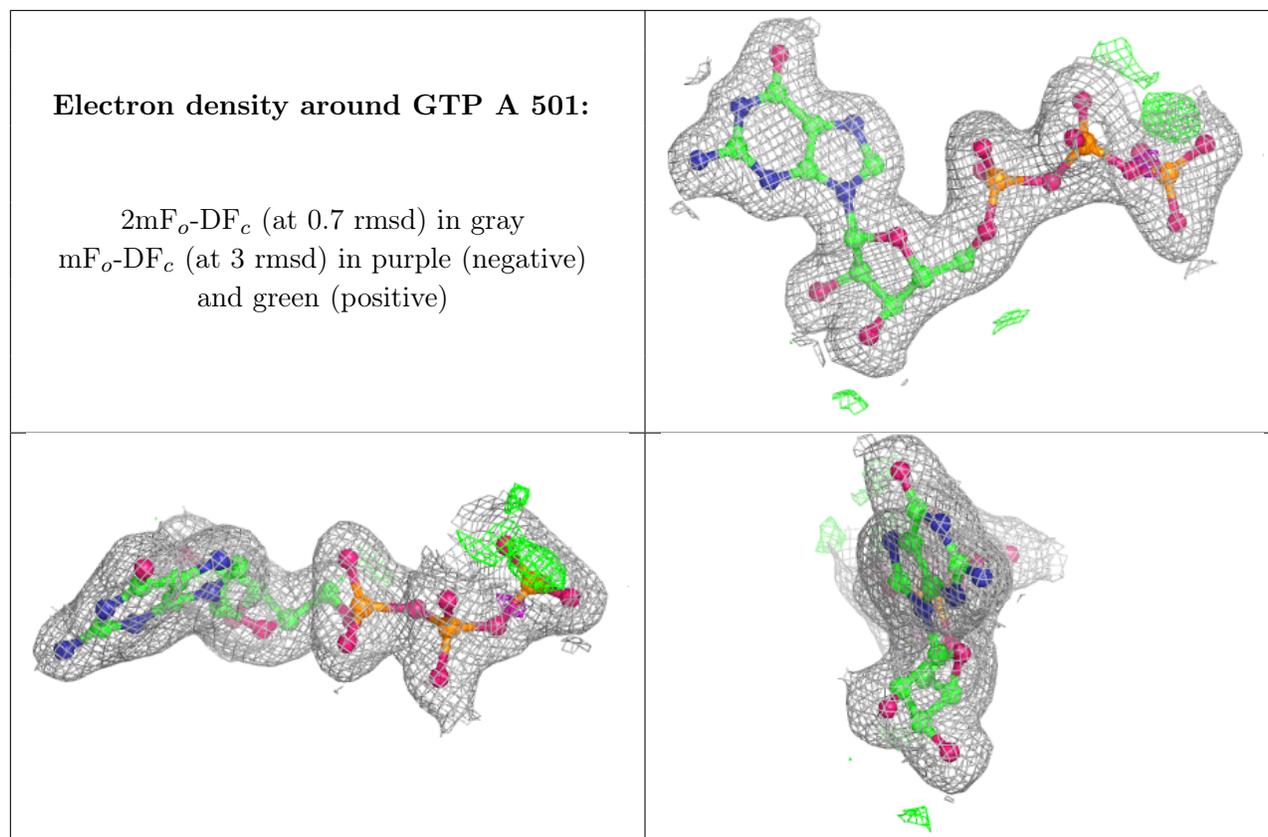and green (positive)
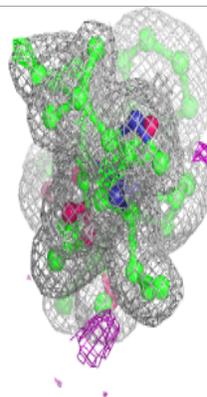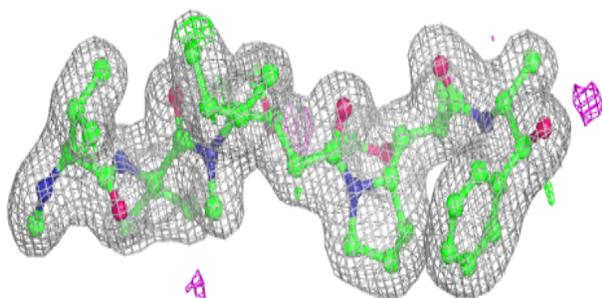
**Electron density around GDP D 501:**

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)

**Electron density around GDP B 501:**

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)







**Electron density around GTP C 501:**

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)

**Electron density around GTP A 501:**

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative) and green (positive)

## 6.5 Other polymers ⓘ

There are no such residues in this entry.