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

Dec 5, 2022 – 02:22 pm GMT

| | | |
|---|---|---|
| PDB ID | : | 8A0L |
| Title | : | Tubulin-CW1-complex |
| Authors | : | Prota, A.E.; Diaz, J.F.; Steinmetz, M.O.; Oliva, M.A. |
| Deposited on | : | 2022-05-28 |
| Resolution | : | 2.00 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 (i)) were used in the production of this report:

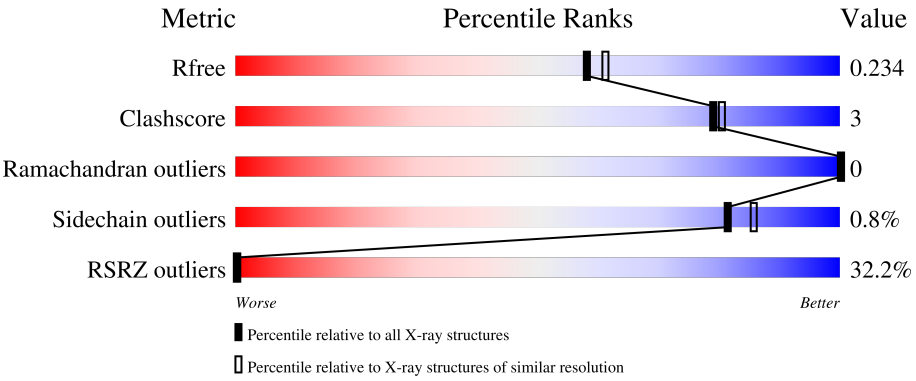
| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Mogul | : | 1.8.4, CSD as541be (2020) |
| Xtriage (Phenix) | : | 1.13 |
| EDS | : | 2.31.3 |
| buster-report | : | 1.1.7 (2018) |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Refmac | : | 5.8.0267 |
| CCP4 | : | 7.1.010 (Gargrove) |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.31.3 |

# 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 2.00 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 | 8085 (2.00-2.00) |
| Clashscore | 141614 | 9178 (2.00-2.00) |
| Ramachandran outliers | 138981 | 9054 (2.00-2.00) |
| Sidechain outliers | 138945 | 9053 (2.00-2.00) |
| RSRZ outliers | 127900 | 7900 (2.00-2.00) |

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 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 | 26% 91% 6% · |
| 1 | C | 451 | 13% 92% 6% · |
| 2 | B | 445 | 14% 88% 8% · |
| 2 | D | 445 | 34% 86% 10% · |
| 3 | E | 143 | 24% 79% 7% 14% |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 4 | F | 384 |   |

# 2 Entry composition (i)

There are 13 unique types of molecules in this entry. The entry contains 18367 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 | C | N | O | S | 0 | 5 | 0 |
| | | | 3437 | 2180 | 582 | 652 | 23 | | | |
| 1 | C | 440 | Total | C | N | O | S | 0 | 7 | 0 |
| | | | 3465 | 2196 | 585 | 660 | 24 | | | |

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

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 | B | 427 | Total | C | N | O | S | 0 | 9 | 0 |
| | | | 3398 | 2137 | 580 | 653 | 28 | | | |
| 2 | D | 427 | Total | C | N | O | S | 0 | 8 | 0 |
| | | | 3394 | 2136 | 576 | 652 | 30 | | | |

- Molecule 3 is a protein called Stathmin-4.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 3 | E | 123 | Total | C | N | O | S | 0 | 2 | 0 |
| | | | 1026 | 632 | 185 | 204 | 5 | | | |

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 |

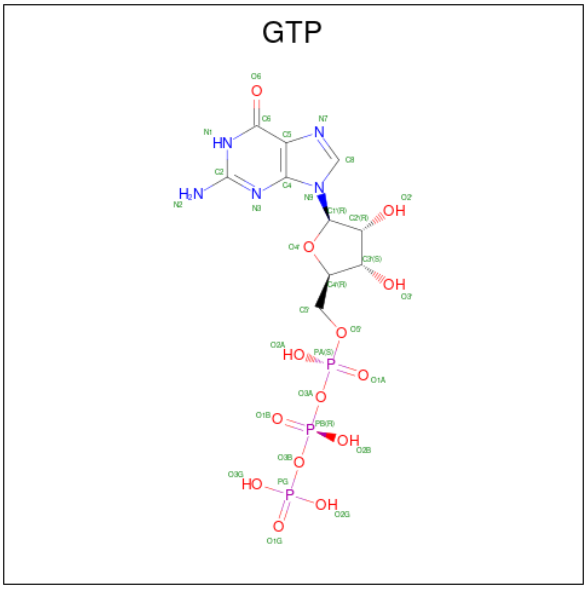
- Molecule 4 is a protein called Tubulin beta-2B chain.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 4 | F | 343 | Total | C | N | O | S | 0 | 2 | 0 |
| | | | 2814 | 1805 | 479 | 516 | 14 | | | |

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 | | |

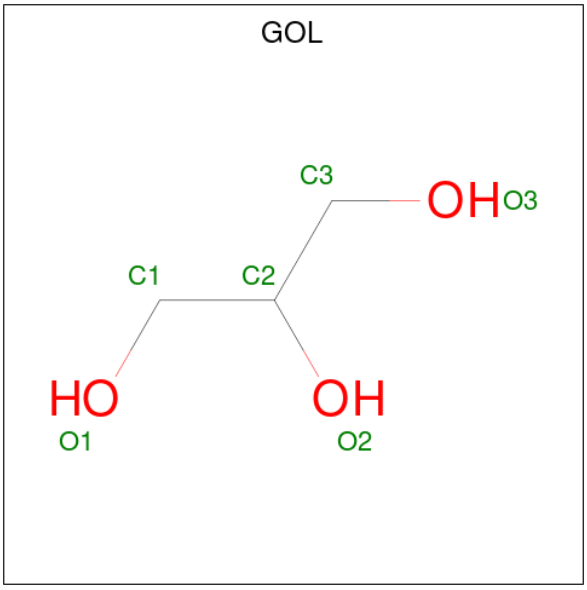
- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

| 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 | | |

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

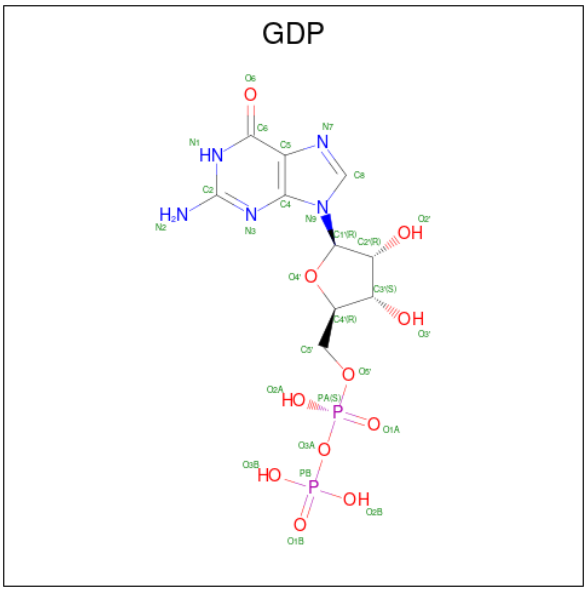
| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|---|---|---|---|---|---|
| 7 | A | 2 | Total Ca<br>2 2 | 0 | 0 |
| 7 | B | 1 | Total Ca<br>1 1 | 0 | 0 |
| 7 | C | 1 | Total Ca<br>1 1 | 0 | 0 |

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



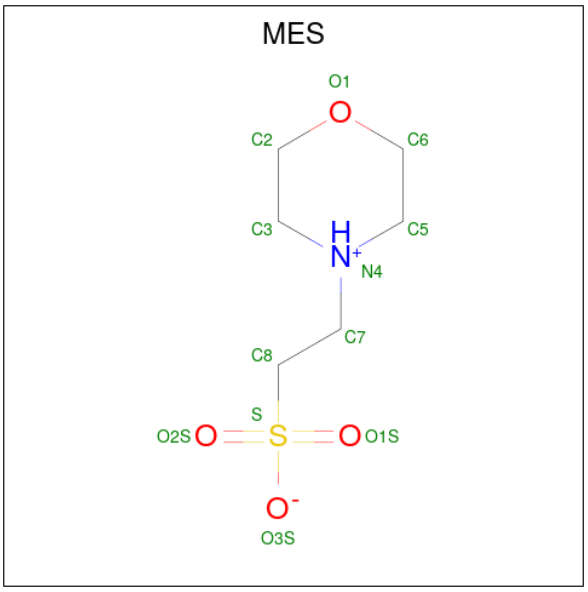

| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|---|---|---|---|---|---|
| 8 | A | 1 | Total C O<br>6 3 3 | 0 | 0 |
| 8 | A | 1 | Total C O<br>6 3 3 | 0 | 0 |

- Molecule 9 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: $C_{10}H_{15}N_5O_{11}P_2$).



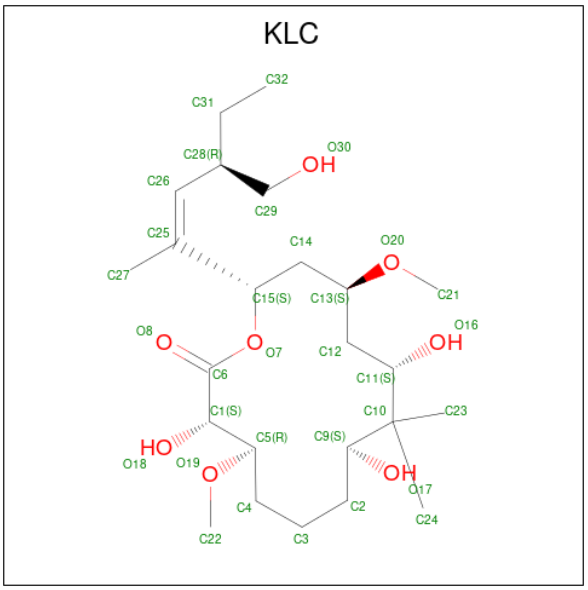

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|---|---|
| 9 | B | 1 | Total | C | N | O | P | 0 | 0 |
| | | | 28 | 10 | 5 | 11 | 2 | | |
| 9 | D | 1 | Total | C | N | O | P | 0 | 0 |
| | | | 28 | 10 | 5 | 11 | 2 | | |

- Molecule 10 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$).



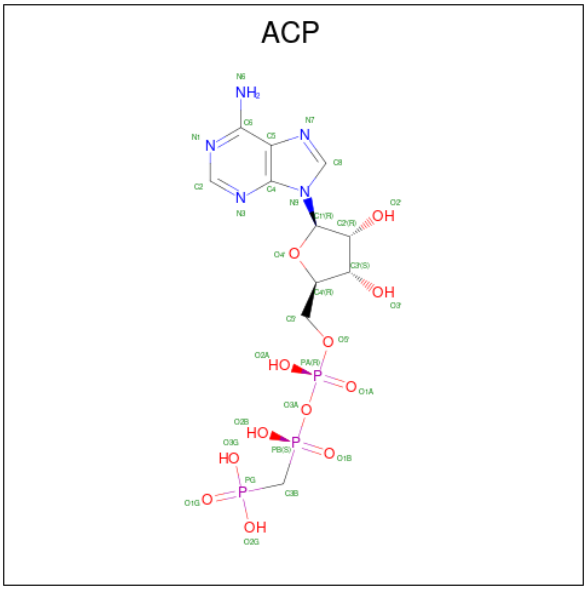

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|---|---|
| 10 | B | 1 | Total | C | N | O | S | 0 | 0 |
| | | | 12 | 6 | 1 | 4 | 1 | | |

- Molecule 11 is (3 {S},4 {R},8 {S},10 {S},12 {S},14 {S})-14-[( {Z},4 {R})-4-(hydroxymethyl)hex-2-en-2-yl]-4,12-dimethoxy-9,9-dimethyl-3,8,10-tris(oxidanyl)-1-oxacyclotetradecan-2-one (three-letter code: KLC) (formula: $C_{24}H_{44}O_8$) (labeled as "Ligand of Interest" by depositor).




| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|---|---|---|---|---|---|
| 11 | B | 1 | Total C O<br>32 24 8 | 0 | 0 |

- Molecule 12 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 |
|---|---|---|---|---|---|---|---|---|---|
| 12 | F | 1 | Total | C | N | O | P | 0 | 0 |
| | | | 31 | 11 | 5 | 12 | 3 | | |

- Molecule 13 is water.

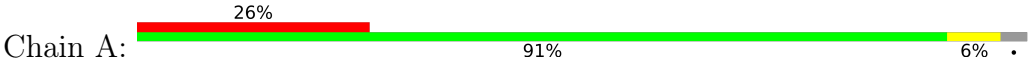
| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| 13 | A | 138 | Total | O | 0 | 0 |
| | | | 138 | 138 | | |
| 13 | B | 129 | Total | O | 0 | 0 |
| | | | 129 | 129 | | |
| 13 | C | 233 | Total | O | 0 | 0 |
| | | | 233 | 233 | | |
| 13 | D | 68 | Total | O | 0 | 0 |
| | | | 68 | 68 | | |
| 13 | E | 32 | Total | O | 0 | 0 |
| | | | 32 | 32 | | |
| 13 | F | 18 | Total | O | 0 | 0 |
| | | | 18 | 18 | | |

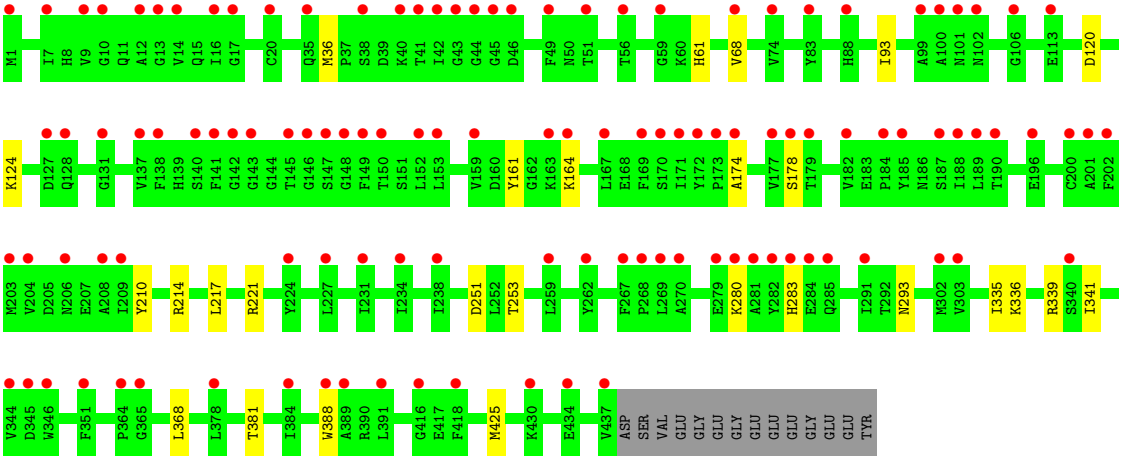
# 3 Residue-property plots (i)

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: Tubulin alpha-1B chain

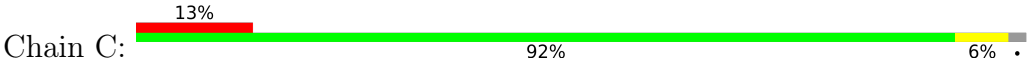
Chain A:

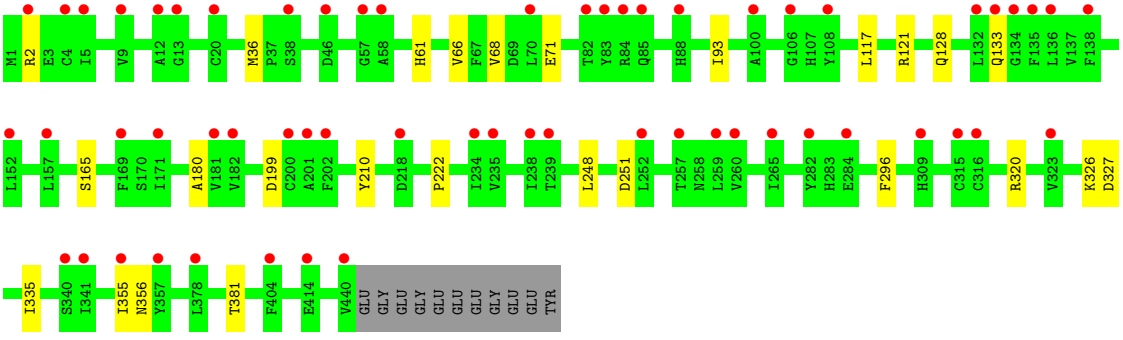



• Molecule 1: Tubulin alpha-1B chain

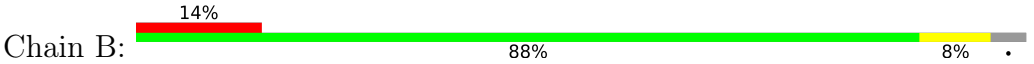
Chain C:

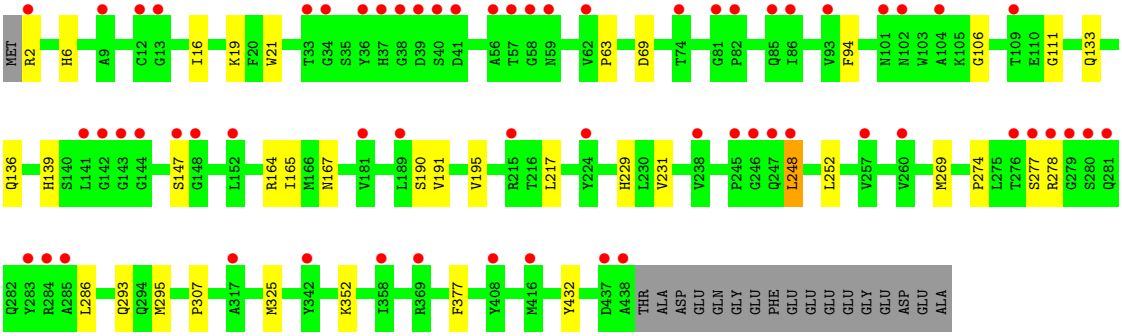


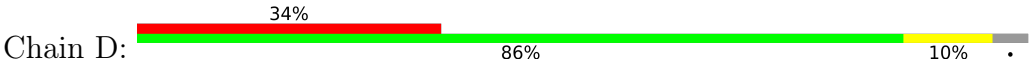

• Molecule 2: Tubulin beta-2B chain

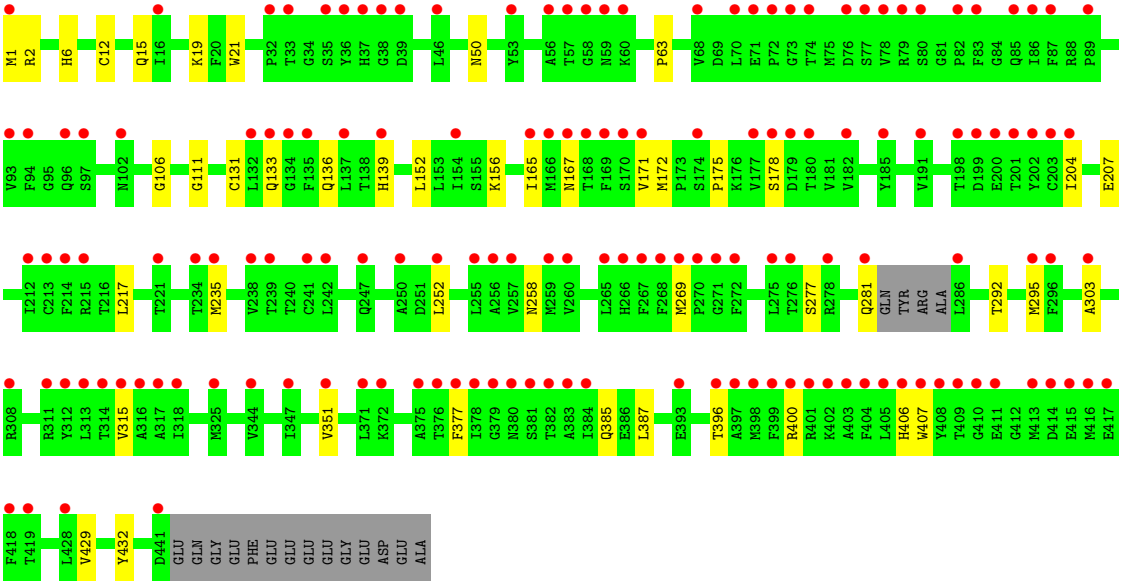
Chain B:



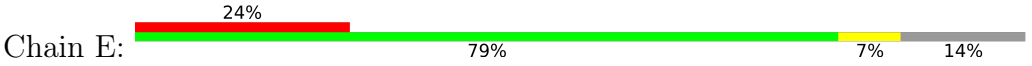


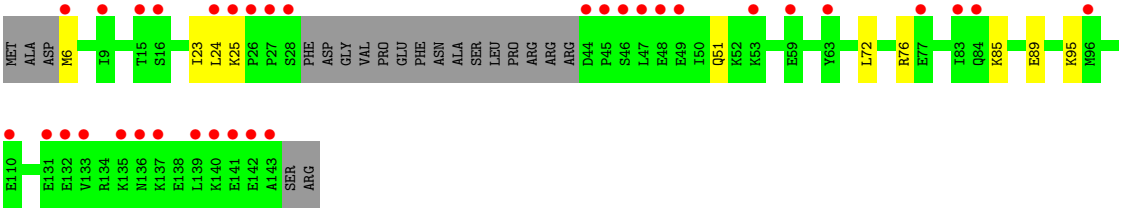

• Molecule 2: Tubulin beta-2B chain



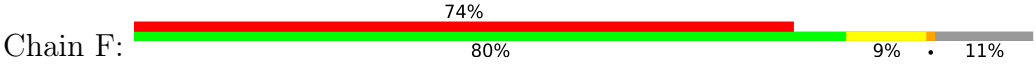



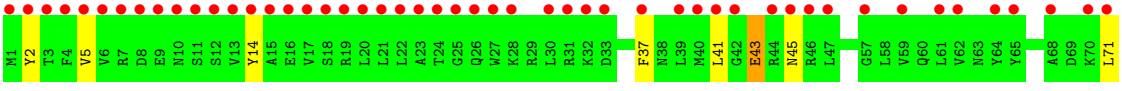

• Molecule 3: Stathmin-4

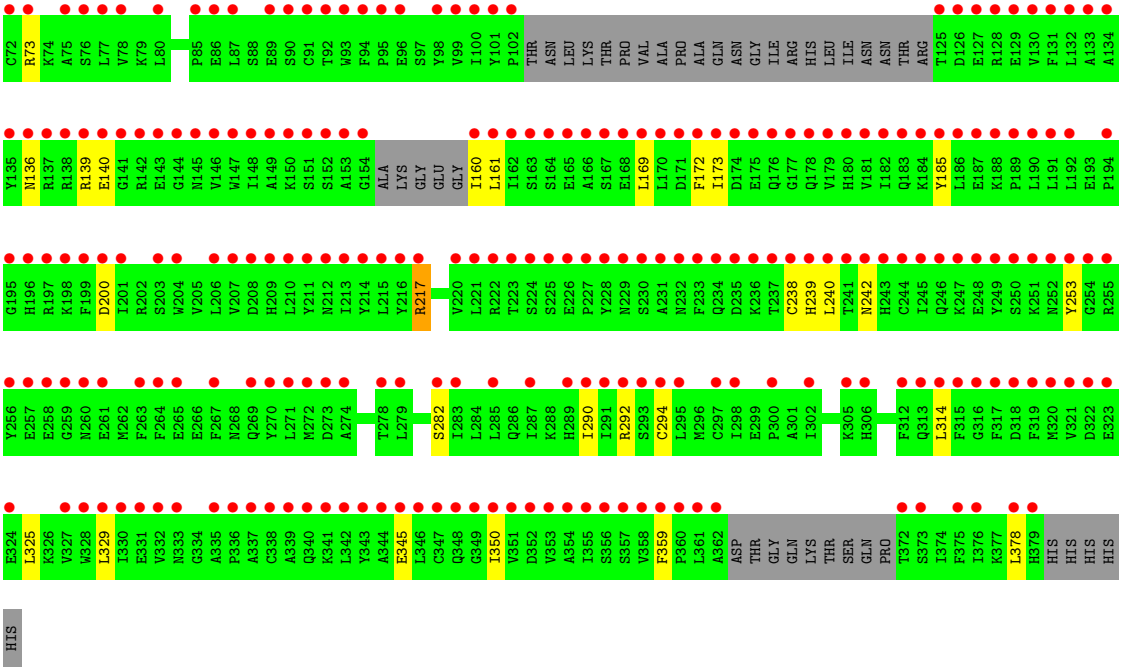



• Molecule 4: Tubulin beta-2B chain




C72 R73 K74 A75 S76 L77 V78 K79 L80 P85 E86 L87 S88 E89 S90 C91 T92 W93 F94 P95 E96 S97 Y98 V99 I100 Y101 P102 THR ASN LEU LYS THR PRO VAL ALA PRO ALA GLN ASN GLY ILE ARG HIS LEU ILE ASN ASN THR ARG T125 D126 E127 R128 E129 V130 F131 L132 A133 A134
Y135 N136 R137 R138 R139 E140 G141 R142 E143 G144 N145 V146 W147 I148 A149 K150 S151 S152 A153 G154 ALA LYS GLY GLU GLY I160 L161 I162 S163 S164 E165 A166 S167 E168 L169 L170 D171 F172 I173 D174 E175 Q176 G177 Q178 V179 H180 V181 I182 Q183 K184 Y185 L186 E187 K188 P189 L190 L191 L192 E193 P194
G195 H196 R197 K198 F199 D200 I201 R202 S203 W204 V205 L206 V207 D208 H209 L210 Y211 N212 I213 Y214 L215 Y216 R217 V220 L221 R222 T223 S224 S225 E226 P227 Y228 N229 S230 A231 N232 F233 Q234 D235 K236 T237 C238 H239 L240 T241 N242 H243 C244 I245 Q246 K247 E248 Y249 S250 K251 N252 Y253 G254 R255
Y256 E257 E258 G259 N260 E261 M262 F263 F264 E265 E266 F267 N268 Q269 Y270 L271 M272 D273 A274 T278 L279 S282 I283 L284 L285 Q286 I287 K288 H289 I290 I291 R292 S293 C294 L295 M296 C297 I298 E299 P300 A301 I302 K305 H306 F312 Q313 L314 F315 G316 F317 D318 F319 M320 V321 D322 E323
E324 L325 K326 V327 W328 L329 I330 E331 V332 N333 G334 A335 P336 A337 C338 A339 Q340 K341 L342 Y343 A344 E345 L346 C347 Q348 G349 I350 V351 D352 V353 A354 I355 S356 S357 V358 F359 P360 L361 A362 ASP THR GLY GLN LYS THR SER GLN PRO T372 S373 I374 F375 I376 K377 L378 H379 HIS 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, $\alpha$, $\beta$, $\gamma$ | 104.63Å 156.56Å 179.32Å<br>90.00° 90.00° 90.00° | Depositor |
| Resolution (Å) | 47.51 – 2.00<br>49.62 – 2.00 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 99.8 (47.51-2.00)<br>100.0 (49.62-2.00) | Depositor<br>EDS |
| $R_{merge}$ | 0.13 | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 1.01 (at 2.00Å) | Xtriage |
| Refinement program | PHENIX 1.16_3549 | Depositor |
| R, $R_{free}$ | 0.188 , 0.226<br>0.197 , 0.234 | Depositor<br>DCC |
| $R_{free}$ test set | 9918 reflections (5.00%) | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 44.3 | Xtriage |
| Anisotropy | 0.138 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | (Not available) , (Not available) | EDS |
| L-test for twinning[2] | $< \|L\| > = 0.50$, $< L^2 > = 0.33$ | Xtriage |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| $F_o$,$F_c$ correlation | 0.97 | EDS |
| Total number of atoms | 18367 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 66.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.42% 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: MG, ACP, GTP, GDP, KLC, CA, MES, GOL

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 | 0.30 | 0/3530 | 0.48 | 0/4793 |
| 1 | C | 0.34 | 0/3564 | 0.51 | 0/4841 |
| 2 | B | 0.32 | 0/3500 | 0.50 | 0/4740 |
| 2 | D | 0.29 | 0/3493 | 0.45 | 0/4730 |
| 3 | E | 0.29 | 0/1037 | 0.40 | 0/1376 |
| 4 | F | 0.26 | 0/2883 | 0.43 | 0/3894 |
| All | All | 0.30 | 0/18007 | 0.47 | 0/24374 |

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 | 3437 | 0 | 3368 | 14 | 0 |
| 1 | C | 3465 | 0 | 3393 | 15 | 0 |
| 2 | B | 3398 | 0 | 3302 | 22 | 0 |
| 2 | D | 3394 | 0 | 3295 | 27 | 0 |
| 3 | E | 1026 | 0 | 1041 | 5 | 0 |
| 4 | F | 2814 | 0 | 2788 | 16 | 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 |
| 7 | A | 2 | 0 | 0 | 0 | 0 |
| 7 | B | 1 | 0 | 0 | 0 | 0 |
| 7 | C | 1 | 0 | 0 | 0 | 0 |
| 8 | A | 12 | 0 | 16 | 1 | 0 |
| 9 | B | 28 | 0 | 12 | 0 | 0 |
| 9 | D | 28 | 0 | 12 | 1 | 0 |
| 10 | B | 12 | 0 | 12 | 2 | 0 |
| 11 | B | 32 | 0 | 0 | 1 | 0 |
| 12 | F | 31 | 0 | 14 | 1 | 0 |
| 13 | A | 138 | 0 | 0 | 0 | 0 |
| 13 | B | 129 | 0 | 0 | 1 | 0 |
| 13 | C | 233 | 0 | 0 | 4 | 0 |
| 13 | D | 68 | 0 | 0 | 2 | 0 |
| 13 | E | 32 | 0 | 0 | 0 | 0 |
| 13 | F | 18 | 0 | 0 | 0 | 0 |
| All | All | 18367 | 0 | 17277 | 98 | 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 3.

The worst 5 of 98 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) |
|---|---|---|---|
| 2:B:147[A]:SER:HG | 2:B:190:SER:HG | 1.28 | 0.77 |
| 2:B:229:HIS:HB2 | 2:B:278:ARG:HE | 1.53 | 0.73 |
| 4:F:71:LEU:HD11 | 4:F:294:CYS:HB3 | 1.69 | 0.73 |
| 2:B:269:MET:HE1 | 2:B:307:PRO:HG3 | 1.73 | 0.70 |
| 2:D:2:ARG:HB3 | 2:D:133:GLN:HG3 | 1.76 | 0.67 |

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 | 440/451 (98%) | 436 (99%) | 4 (1%) | 0 | 100 | 100 |
| 1 | C | 445/451 (99%) | 435 (98%) | 10 (2%) | 0 | 100 | 100 |
| 2 | B | 434/445 (98%) | 427 (98%) | 7 (2%) | 0 | 100 | 100 |
| 2 | D | 431/445 (97%) | 424 (98%) | 7 (2%) | 0 | 100 | 100 |
| 3 | E | 121/143 (85%) | 120 (99%) | 1 (1%) | 0 | 100 | 100 |
| 4 | F | 337/384 (88%) | 324 (96%) | 13 (4%) | 0 | 100 | 100 |
| All | All | 2208/2319 (95%) | 2166 (98%) | 42 (2%) | 0 | 100 | 100 |

There are no Ramachandran outliers to report.

### 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 | 373/379 (98%) | 372 (100%) | 1 (0%) | 92 | 95 |
| 1 | C | 378/379 (100%) | 376 (100%) | 2 (0%) | 88 | 92 |
| 2 | B | 378/383 (99%) | 376 (100%) | 2 (0%) | 88 | 92 |
| 2 | D | 377/383 (98%) | 374 (99%) | 3 (1%) | 81 | 86 |
| 3 | E | 112/127 (88%) | 111 (99%) | 1 (1%) | 78 | 83 |
| 4 | F | 310/342 (91%) | 303 (98%) | 7 (2%) | 50 | 53 |
| All | All | 1928/1993 (97%) | 1912 (99%) | 16 (1%) | 81 | 86 |

5 of 16 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 4 | F | 238 | CYS |
| 4 | F | 217 | ARG |
| 3 | E | 51 | GLN |
| 4 | F | 200 | ASP |
| 2 | D | 207 | GLU |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 1 | A | 301 | GLN |
| 2 | D | 247 | 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 17 ligands modelled in this entry, 8 are monoatomic - leaving 9 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 | $\#\|Z\| > 2$ |
| 12 | ACP | F | 401 | - | 27,33,33 | 2.36 | 7 (25%) | 32,52,52 | 1.73 | 5 (15%) |
| 9 | GDP | D | 501 | 6 | 24,30,30 | 0.96 | 1 (4%) | 30,47,47 | 1.05 | 3 (10%) |

| Mol | Type | Chain | Res | Link | Bond lengths | | | Bond angles | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Counts | RMSZ | $\#\|Z\|>2$ | Counts | RMSZ | $\#\|Z\|>2$ |
| 8 | GOL | A | 506 | - | 5,5,5 | 0.79 | 0 | 5,5,5 | 1.14 | 0 |
| 8 | GOL | A | 505 | - | 5,5,5 | 0.99 | 0 | 5,5,5 | 0.84 | 0 |
| 9 | GDP | B | 501 | 6 | 24,30,30 | 0.95 | 1 (4%) | 30,47,47 | 1.20 | 5 (16%) |
| 11 | KLC | B | 505 | - | 31,32,32 | 0.85 | 0 | 31,44,44 | 1.72 | 4 (12%) |
| 5 | GTP | C | 501 | 6 | 26,34,34 | 1.14 | 2 (7%) | 32,54,54 | 1.31 | 5 (15%) |
| 10 | MES | B | 504 | - | 12,12,12 | 2.34 | 1 (8%) | 14,16,16 | 1.97 | 8 (57%) |
| 5 | GTP | A | 501 | 6 | 26,34,34 | 1.13 | 2 (7%) | 32,54,54 | 1.29 | 5 (15%) |

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 |
|---|---|---|---|---|---|---|---|
| 12 | ACP | F | 401 | - | - | 1/15/38/38 | 0/3/3/3 |
| 9 | GDP | D | 501 | 6 | - | 3/12/32/32 | 0/3/3/3 |
| 8 | GOL | A | 506 | - | - | 3/4/4/4 | - |
| 8 | GOL | A | 505 | - | - | 2/4/4/4 | - |
| 9 | GDP | B | 501 | 6 | - | 5/12/32/32 | 0/3/3/3 |
| 11 | KLC | B | 505 | - | - | 8/54/54/54 | 0/0/1/1 |
| 5 | GTP | C | 501 | 6 | - | 5/18/38/38 | 0/3/3/3 |
| 10 | MES | B | 504 | - | - | 2/6/14/14 | 0/1/1/1 |
| 5 | GTP | A | 501 | 6 | - | 7/18/38/38 | 0/3/3/3 |

The worst 5 of 14 bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 10 | B | 504 | MES | C8-S | -7.81 | 1.66 | 1.77 |
| 12 | F | 401 | ACP | PB-O3A | 7.38 | 1.66 | 1.58 |
| 12 | F | 401 | ACP | C4-N3 | 4.84 | 1.42 | 1.35 |
| 12 | F | 401 | ACP | PG-O3G | 4.17 | 1.64 | 1.54 |
| 12 | F | 401 | ACP | C2-N3 | 4.03 | 1.38 | 1.32 |

The worst 5 of 35 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 12 | F | 401 | ACP | N3-C2-N1 | -5.85 | 119.54 | 128.68 |
| 11 | B | 505 | KLC | C15-O7-C6 | 5.43 | 127.02 | 116.67 |
| 12 | F | 401 | ACP | C5-C6-N6 | 4.41 | 127.05 | 120.35 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 11 | B | 505 | KLC | O7-C15-C14 | 4.00 | 115.48 | 106.64 |
| 11 | B | 505 | KLC | O19-C5-C1 | 3.43 | 116.42 | 107.78 |

There are no chirality outliers.

5 of 36 torsion outliers are listed below:

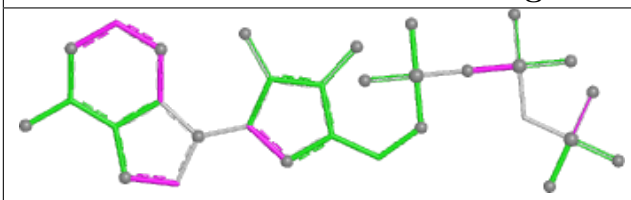
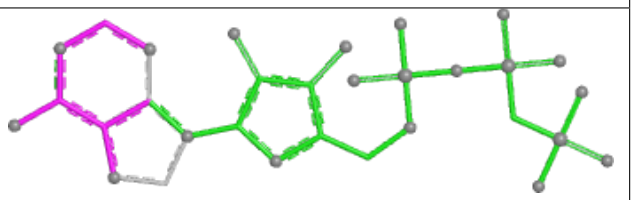
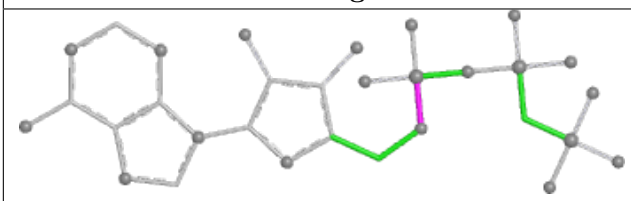
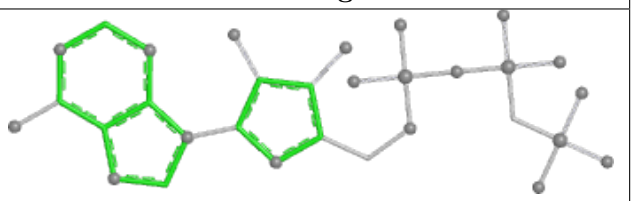
| Mol | Chain | Res | Type | Atoms |
|---|---|---|---|---|
| 5 | A | 501 | GTP | C5'-O5'-PA-O1A |
| 5 | A | 501 | GTP | C5'-O5'-PA-O2A |
| 5 | C | 501 | GTP | C5'-O5'-PA-O1A |
| 5 | C | 501 | GTP | C5'-O5'-PA-O2A |
| 8 | A | 505 | GOL | O1-C1-C2-O2 |

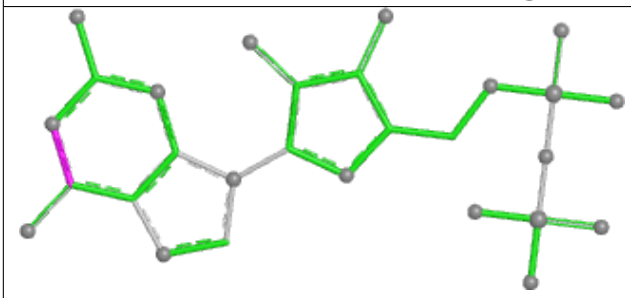
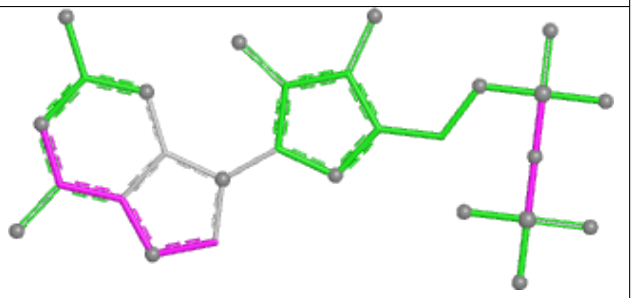
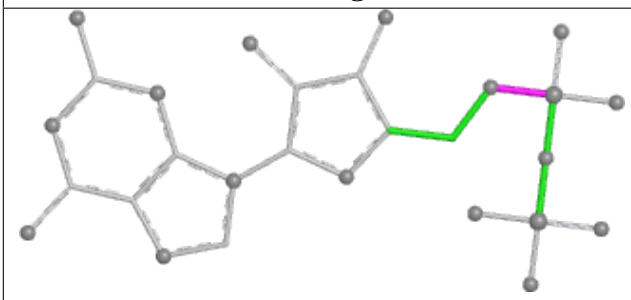
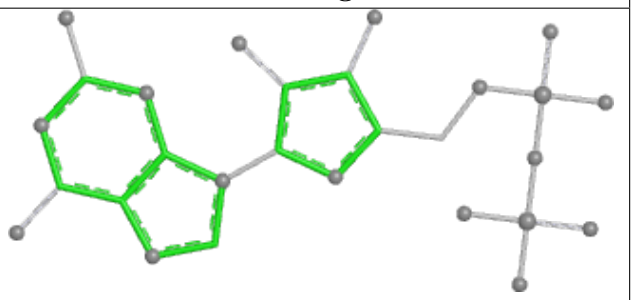
There are no ring outliers.

5 monomers are involved in 5 short contacts:

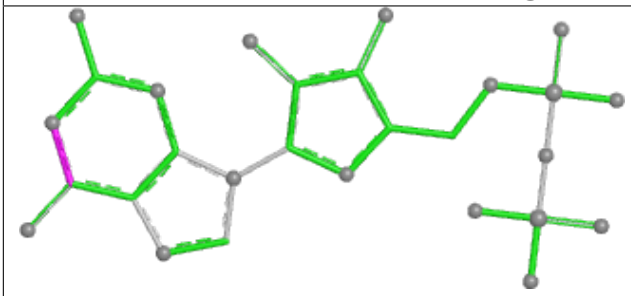
| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|---|---|---|---|---|---|
| 12 | F | 401 | ACP | 1 | 0 |
| 9 | D | 501 | GDP | 1 | 0 |
| 8 | A | 505 | GOL | 1 | 0 |
| 11 | B | 505 | KLC | 1 | 0 |
| 10 | B | 504 | MES | 2 | 0 |

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.

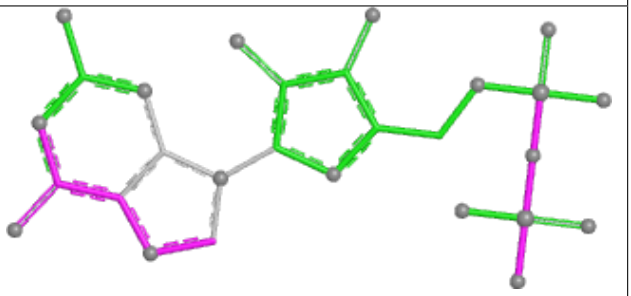
| Ligand ACP F 401 | |
|---|---|
|  Bond lengths |  Bond angles |
|  Torsions |  Rings |

| Ligand GDP D 501 | |
|---|---|
|  Bond lengths |  Bond angles |
|  Torsions |  Rings |

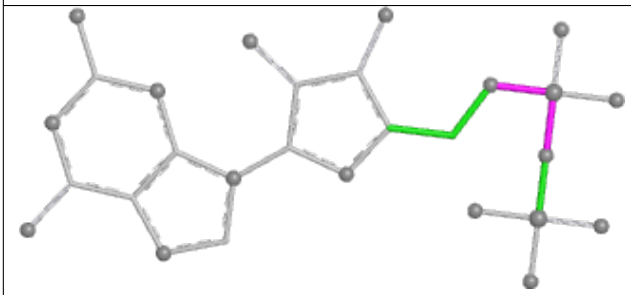
**Ligand GDP B 501**



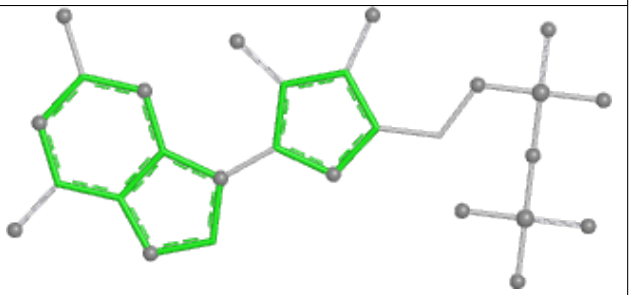
**Bond lengths**



**Bond angles**

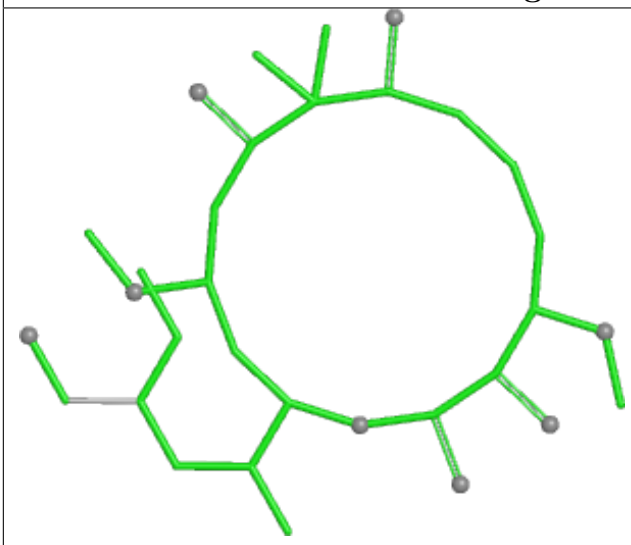


**Torsions**

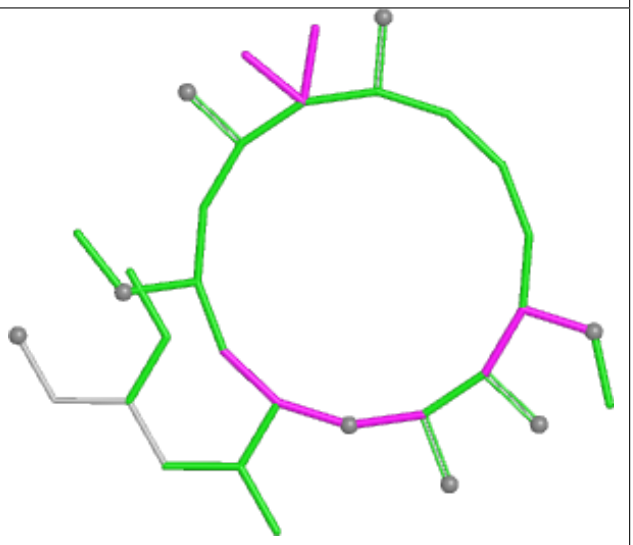


**Rings**

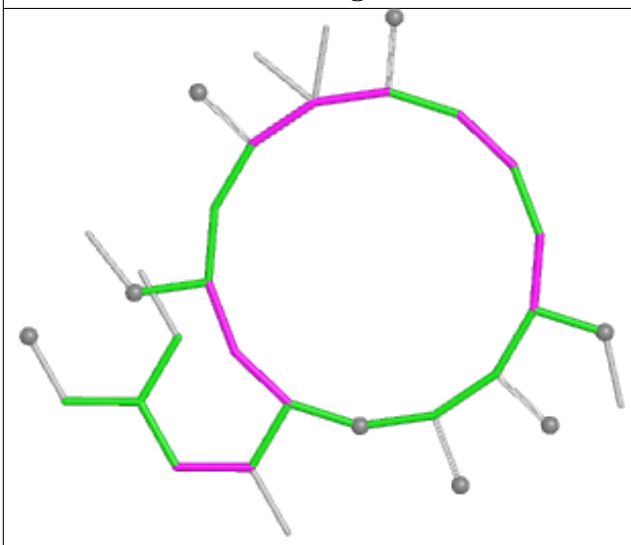
**Ligand KLC B 505**



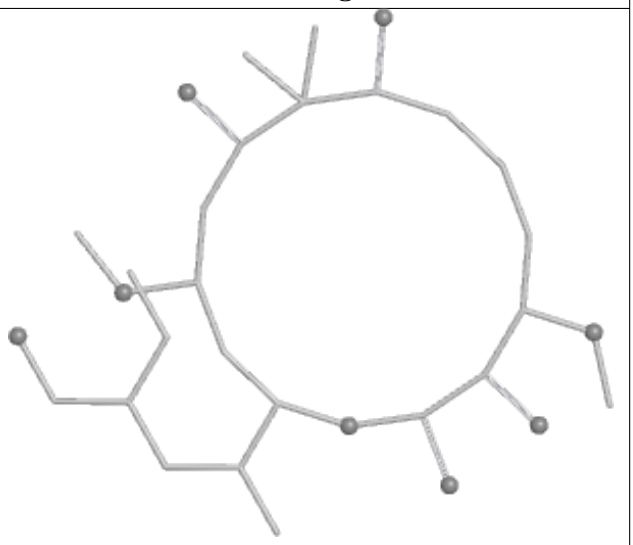
**Bond lengths**



**Bond angles**

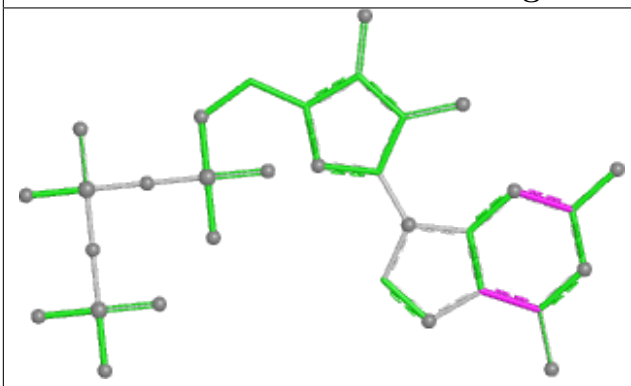


**Torsions**

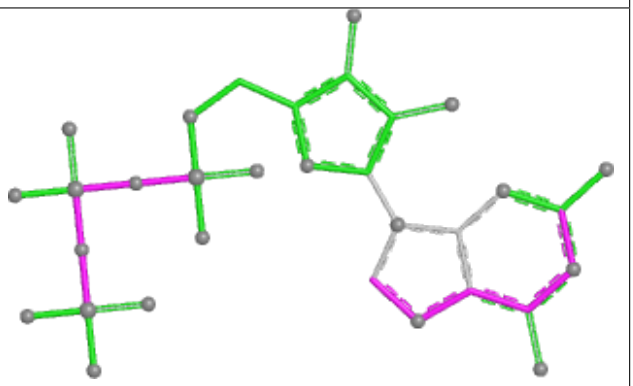


**Rings**

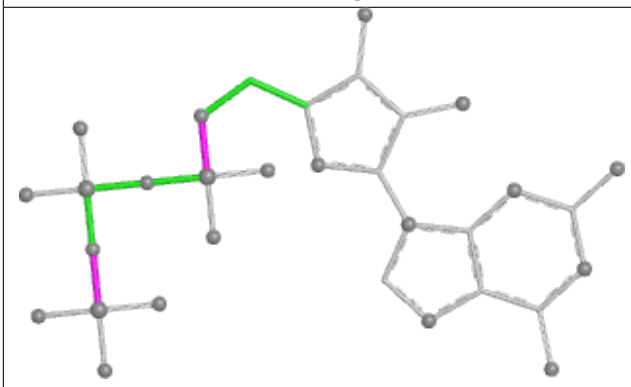
## Ligand GTP C 501



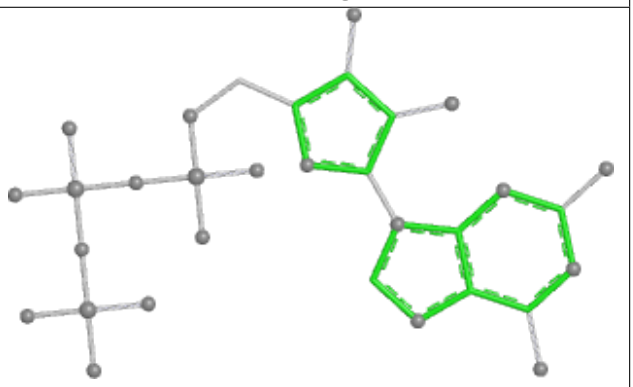
Bond lengths



Bond angles

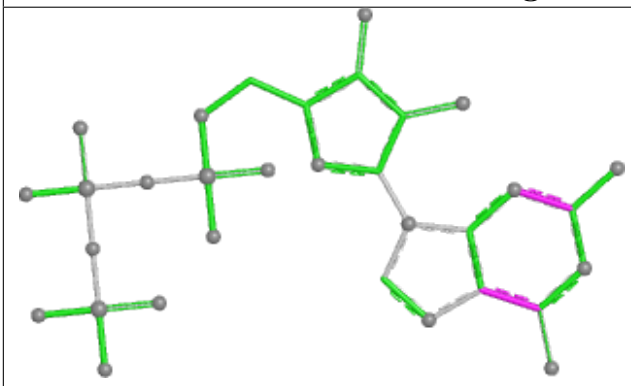


Torsions

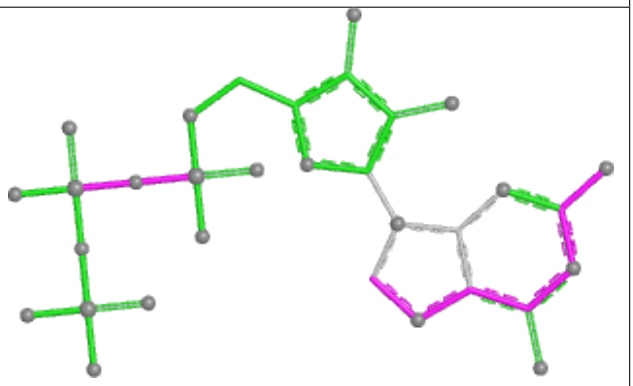


Rings

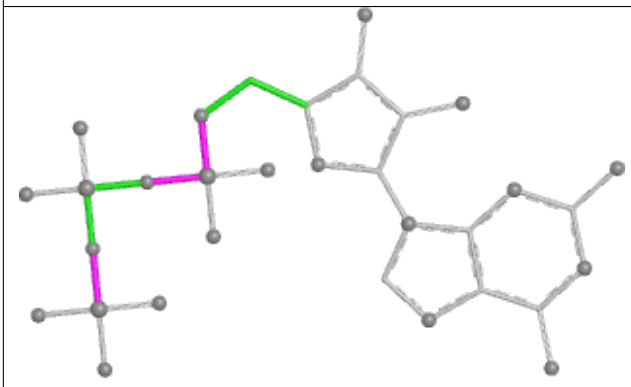
## Ligand GTP A 501



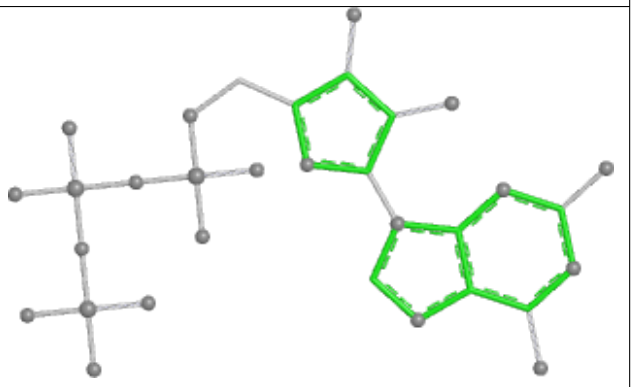
Bond lengths



Bond angles



Torsions



Rings

## 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

## 6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, $95^{th}$ percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

| Mol | Chain | Analysed | <RSRZ> | #RSRZ>2 | | | OWAB(Å$^2$) | Q<0.9 |
|---|---|---|---|---|---|---|---|---|
| 1 | A | 437/451 (96%) | 1.50 | 117 (26%) | 0 | 0 | 35, 54, 87, 163 | 0 |
| 1 | C | 440/451 (97%) | 1.08 | 59 (13%) | 3 | 2 | 31, 43, 70, 120 | 0 |
| 2 | B | 427/445 (95%) | 1.28 | 62 (14%) | 2 | 2 | 31, 50, 85, 151 | 0 |
| 2 | D | 427/445 (95%) | 1.78 | 153 (35%) | 0 | 0 | 38, 64, 98, 128 | 0 |
| 3 | E | 123/143 (86%) | 1.51 | 34 (27%) | 0 | 0 | 40, 65, 105, 145 | 0 |
| 4 | F | 343/384 (89%) | 5.19 | 283 (82%) | 0 | 0 | 58, 114, 170, 211 | 0 |
| All | All | 2197/2319 (94%) | 2.00 | 708 (32%) | 0 | 0 | 31, 58, 131, 211 | 0 |

The worst 5 of 708 RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|---|---|---|---|---|
| 4 | F | 17 | VAL | 17.9 |
| 4 | F | 130 | VAL | 17.8 |
| 4 | F | 362 | ALA | 15.9 |
| 4 | F | 20 | LEU | 15.8 |
| 4 | F | 250 | SER | 15.8 |

## 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 (i)

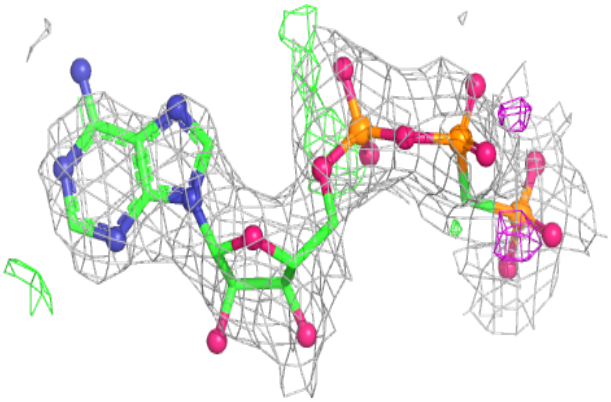
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, $95^{th}$ percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

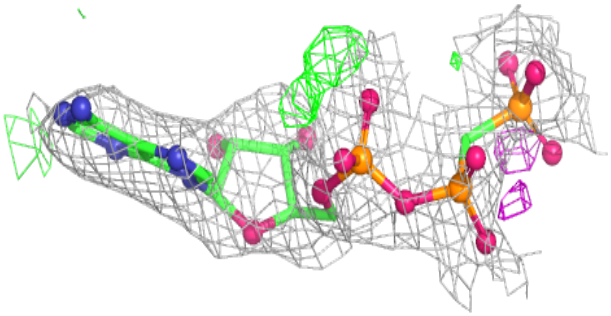
| Mol | Type | Chain | Res | Atoms | RSCC | RSR | B-factors(Å$^2$) | Q<0.9 |
|---|---|---|---|---|---|---|---|---|
| 7 | CA | B | 503 | 1/1 | 0.69 | 0.22 | 95,95,95,95 | 0 |
| 12 | ACP | F | 401 | 31/31 | 0.75 | 0.27 | 92,113,142,145 | 0 |
| 8 | GOL | A | 505 | 6/6 | 0.79 | 0.24 | 79,80,87,92 | 0 |
| 10 | MES | B | 504 | 12/12 | 0.85 | 0.18 | 64,74,90,98 | 0 |
| 8 | GOL | A | 506 | 6/6 | 0.88 | 0.17 | 58,67,78,86 | 0 |
| 11 | KLC | B | 505 | 32/32 | 0.89 | 0.20 | 48,62,70,73 | 0 |
| 6 | MG | D | 502 | 1/1 | 0.90 | 0.09 | 55,55,55,55 | 0 |
| 7 | CA | A | 504 | 1/1 | 0.90 | 0.18 | 99,99,99,99 | 0 |
| 9 | GDP | D | 501 | 28/28 | 0.92 | 0.15 | 49,58,67,73 | 0 |
| 6 | MG | A | 502 | 1/1 | 0.93 | 0.13 | 38,38,38,38 | 0 |
| 5 | GTP | A | 501 | 32/32 | 0.95 | 0.22 | 32,37,42,46 | 0 |
| 5 | GTP | C | 501 | 32/32 | 0.96 | 0.17 | 27,34,38,40 | 0 |
| 7 | CA | A | 503 | 1/1 | 0.96 | 0.08 | 70,70,70,70 | 0 |
| 6 | MG | C | 502 | 1/1 | 0.97 | 0.11 | 34,34,34,34 | 0 |
| 9 | GDP | B | 501 | 28/28 | 0.97 | 0.17 | 28,35,39,40 | 0 |
| 6 | MG | B | 502 | 1/1 | 0.97 | 0.21 | 32,32,32,32 | 0 |
| 7 | CA | C | 503 | 1/1 | 0.99 | 0.06 | 55,55,55,55 | 0 |

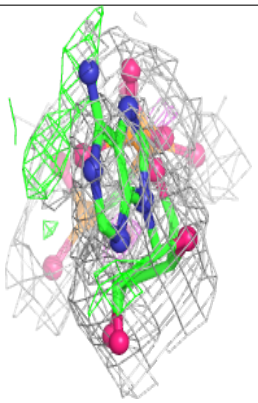
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around ACP F 401:**

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

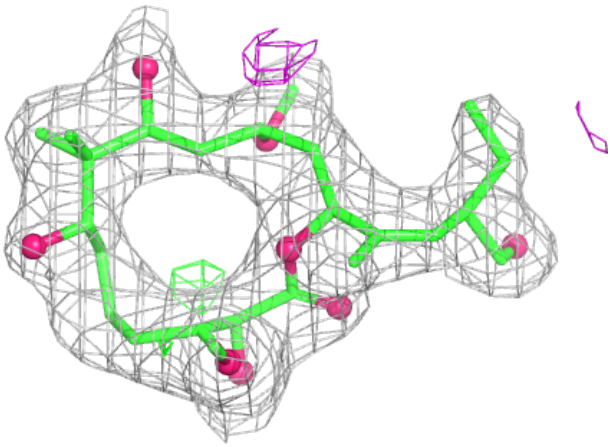


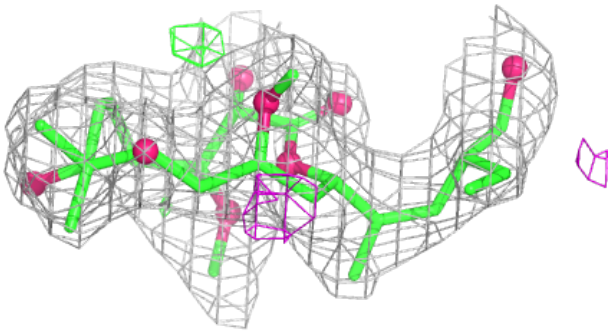


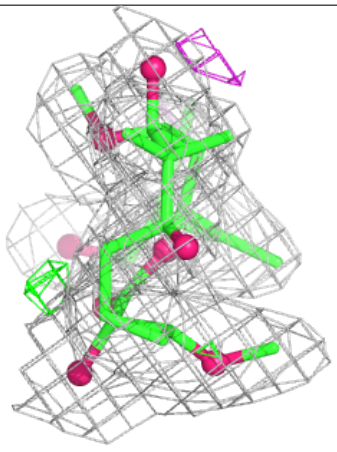


**Electron density around KLC B 505:**

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

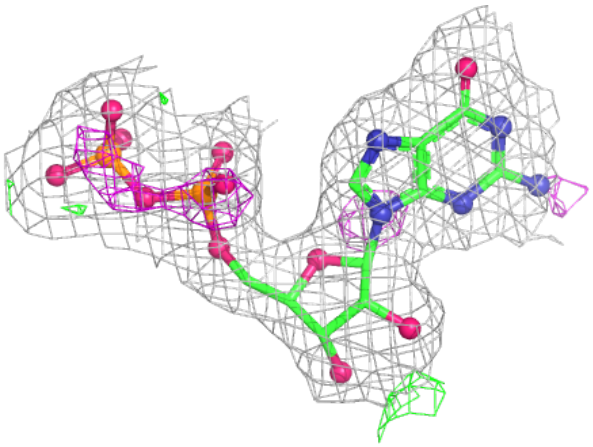


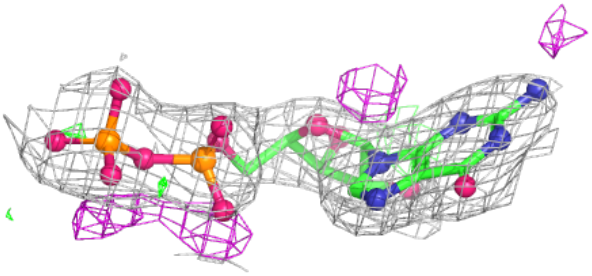


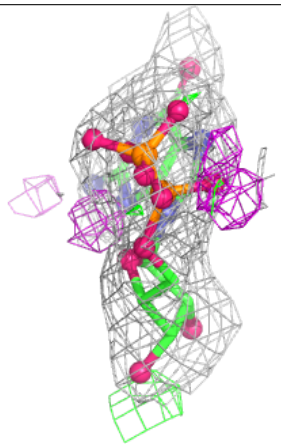


**Electron density around GDP D 501:**

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

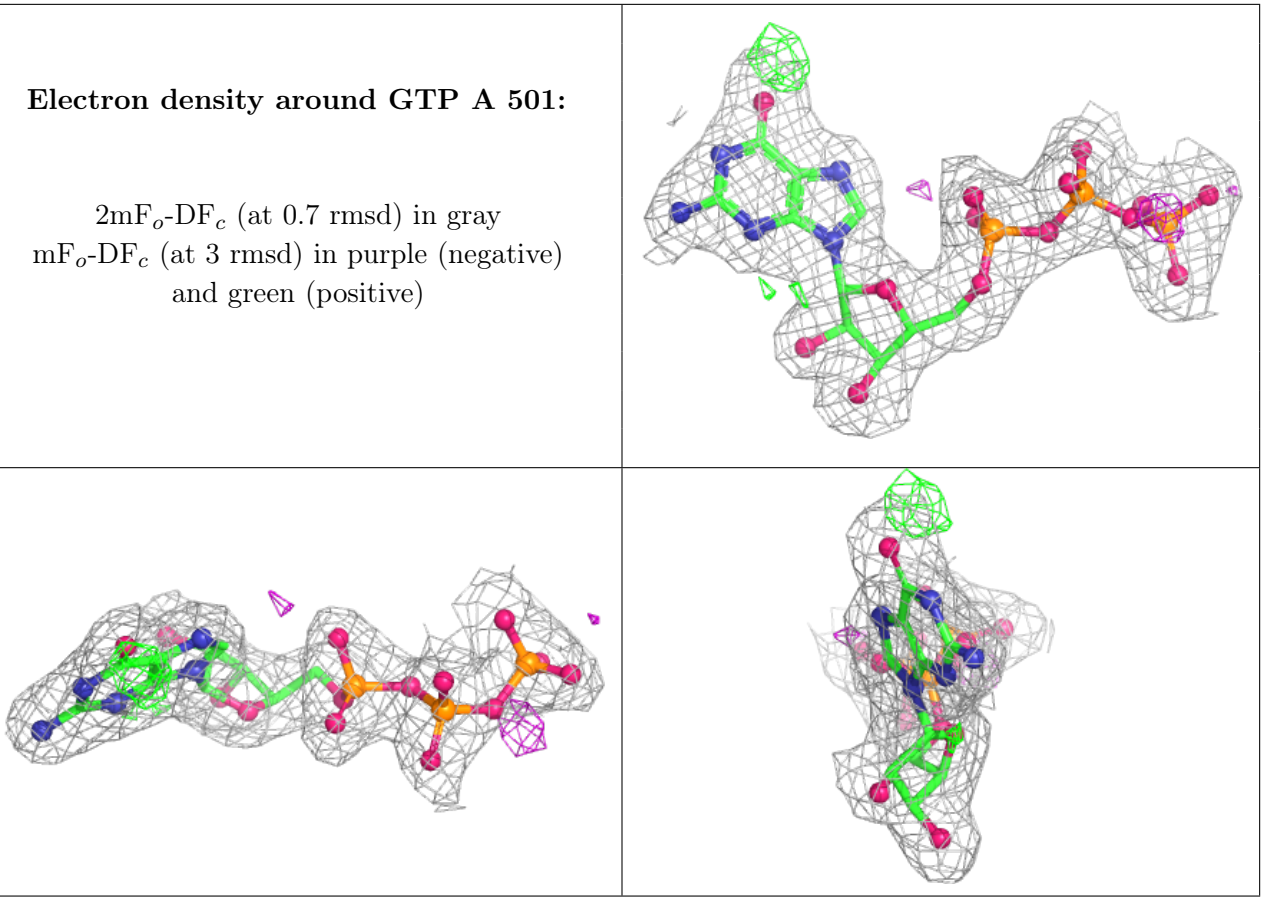






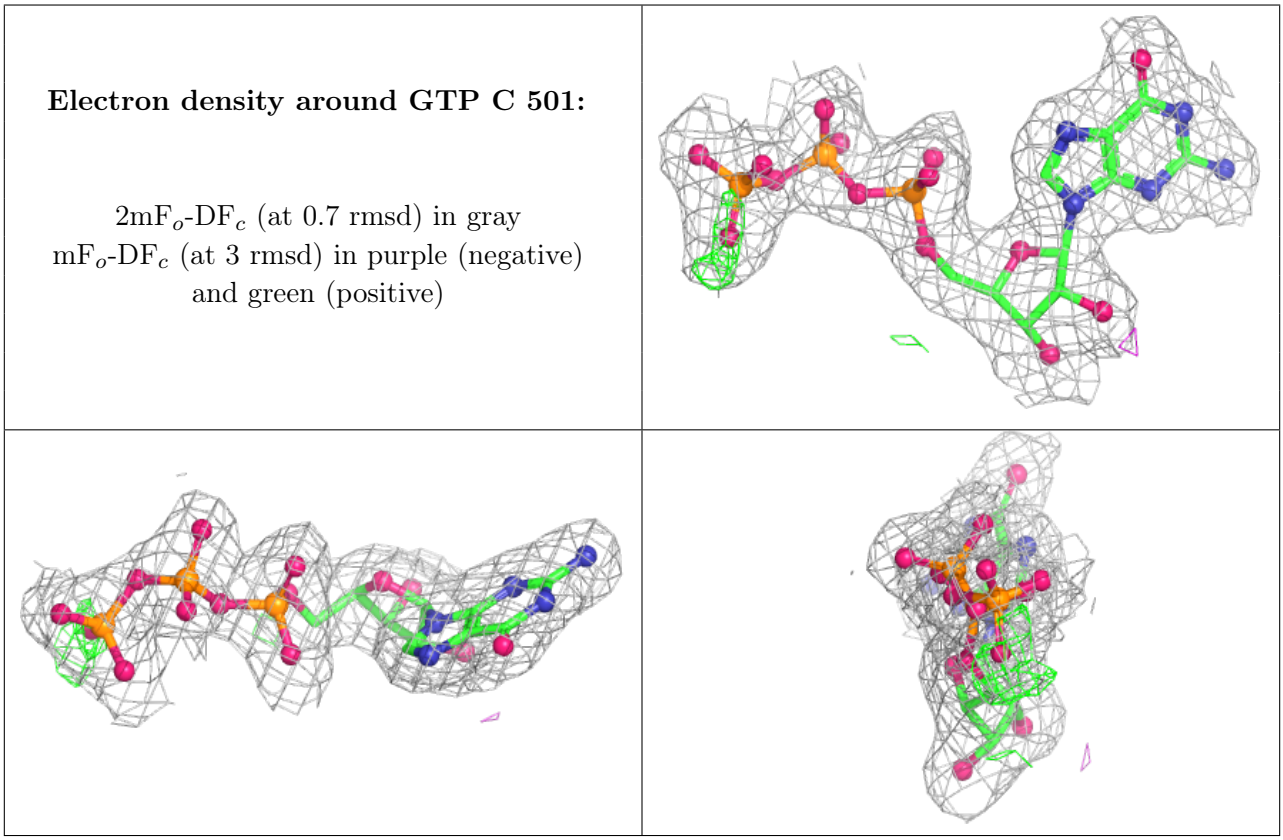
**Electron density around GTP A 501:**

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



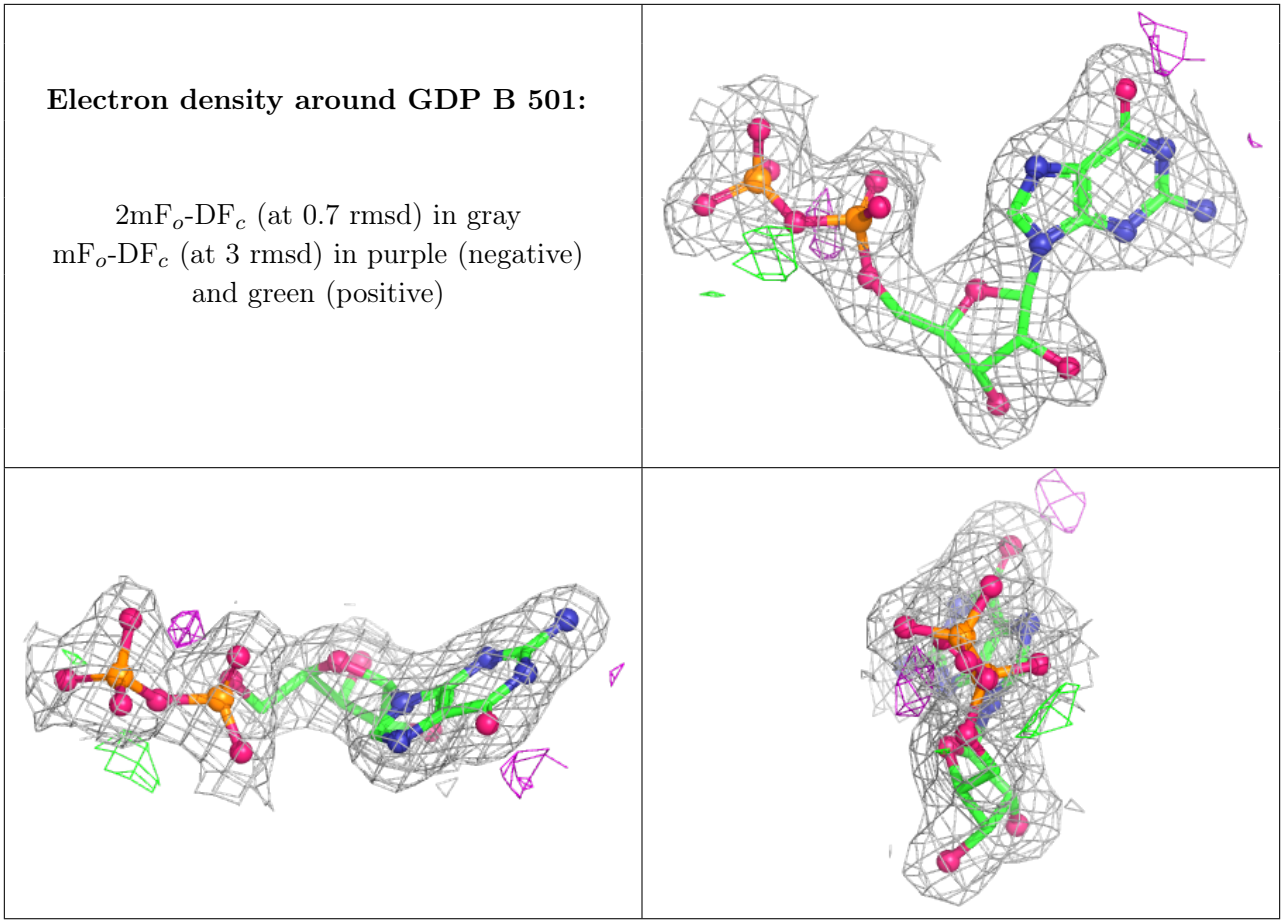
**Electron density around GTP C 501:**

$2mF_o$-$DF_c$ (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:**

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



## 6.5 Other polymers

There are no such residues in this entry.