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

Nov 21, 2023 – 12:10 AM JST

| | | |
|---|---|---|
| PDB ID | : | 7DY1 |
| Title | : | Crystal Structure of Cyanobacterial Circadian Clock Protein KaiC |
| Authors | : | Furuike, Y.; Akiyama, S. |
| Deposited on | : | 2021-01-20 |
| Resolution | : | 2.20 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:

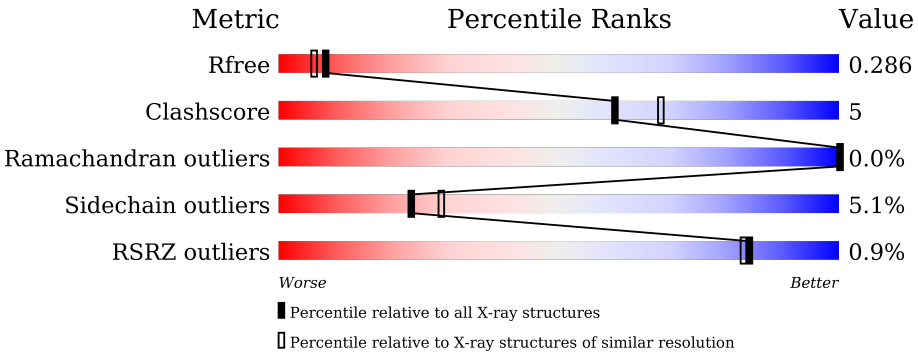
| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Mogul | : | 1.8.5 (274361), CSD as541be (2020) |
| Xtriage (Phenix) | : | 1.13 |
| EDS | : | 2.36 |
| buster-report | : | 1.1.7 (2018) |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Refmac | : | 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 (i)

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

The reported resolution of this entry is 2.20 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 | 4898 (2.20-2.20) |
| Clashscore | 141614 | 5594 (2.20-2.20) |
| Ramachandran outliers | 138981 | 5503 (2.20-2.20) |
| Sidechain outliers | 138945 | 5504 (2.20-2.20) |
| RSRZ outliers | 127900 | 4800 (2.20-2.20) |

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 | 518 | 76% 11% • 12% |
| 2 | B | 518 | 77% 10% • 13% |
| 2 | C | 518 | 75% 12% • 12% |
| 2 | D | 518 | 76% 11% • 12% |
| 2 | E | 518 | 75% 12% • 12% |
| 2 | F | 518 | 76% 11% 12% |

## 2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 21975 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.

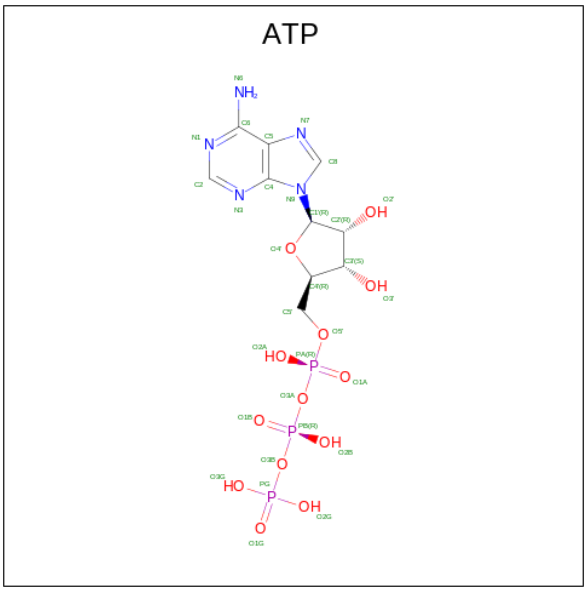
- Molecule 1 is a protein called Circadian clock protein kinase KaiC.

| Mol | Chain | Residues | Atoms | | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | P | S | | | |
| 1 | A | 454 | 3462 | 2197 | 596 | 658 | 1 | 10 | 0 | 1 | 0 |

- Molecule 2 is a protein called Circadian clock protein kinase KaiC.

| Mol | Chain | Residues | Atoms | | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | P | S | | | |
| 2 | B | 452 | 3413 | 2163 | 586 | 652 | 2 | 10 | 0 | 1 | 0 |
| | | | Total | C | N | O | P | S | | | |
| 2 | C | 456 | 3459 | 2187 | 600 | 660 | 2 | 10 | 0 | 0 | 0 |
| | | | Total | C | N | O | P | S | | | |
| 2 | D | 458 | 3491 | 2213 | 600 | 666 | 2 | 10 | 0 | 2 | 0 |
| | | | Total | C | N | O | P | S | | | |
| 2 | E | 455 | 3481 | 2205 | 598 | 666 | 2 | 10 | 0 | 1 | 0 |
| | | | Total | C | N | O | P | S | | | |
| 2 | F | 456 | 3448 | 2188 | 596 | 652 | 2 | 10 | 0 | 0 | 0 |

- Molecule 3 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$).




| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|---|---|
| 3 | A | 1 | Total | C | N | O | P | 0 | 0 |
| | | | 31 | 10 | 5 | 13 | 3 | | |
| 3 | A | 1 | Total | C | N | O | P | 0 | 0 |
| | | | 31 | 10 | 5 | 13 | 3 | | |
| 3 | B | 1 | Total | C | N | O | P | 0 | 0 |
| | | | 31 | 10 | 5 | 13 | 3 | | |
| 3 | B | 1 | Total | C | N | O | P | 0 | 0 |
| | | | 31 | 10 | 5 | 13 | 3 | | |
| 3 | C | 1 | Total | C | N | O | P | 0 | 0 |
| | | | 31 | 10 | 5 | 13 | 3 | | |
| 3 | C | 1 | Total | C | N | O | P | 0 | 0 |
| | | | 31 | 10 | 5 | 13 | 3 | | |
| 3 | D | 1 | Total | C | N | O | P | 0 | 0 |
| | | | 31 | 10 | 5 | 13 | 3 | | |
| 3 | D | 1 | Total | C | N | O | P | 0 | 0 |
| | | | 31 | 10 | 5 | 13 | 3 | | |
| 3 | E | 1 | Total | C | N | O | P | 0 | 0 |
| | | | 31 | 10 | 5 | 13 | 3 | | |
| 3 | E | 1 | Total | C | N | O | P | 0 | 0 |
| | | | 31 | 10 | 5 | 13 | 3 | | |
| 3 | F | 1 | Total | C | N | O | P | 0 | 0 |
| | | | 31 | 10 | 5 | 13 | 3 | | |
| 3 | F | 1 | Total | C | N | O | P | 0 | 0 |
| | | | 31 | 10 | 5 | 13 | 3 | | |

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|---|---|---|---|---|---|
| 4 | A | 2 | Total Mg<br>2 2 | 0 | 0 |
| 4 | B | 2 | Total Mg<br>2 2 | 0 | 0 |
| 4 | C | 2 | Total Mg<br>2 2 | 0 | 0 |
| 4 | D | 2 | Total Mg<br>2 2 | 0 | 0 |
| 4 | E | 2 | Total Mg<br>2 2 | 0 | 0 |
| 4 | F | 2 | Total Mg<br>2 2 | 0 | 0 |

- Molecule 5 is water.

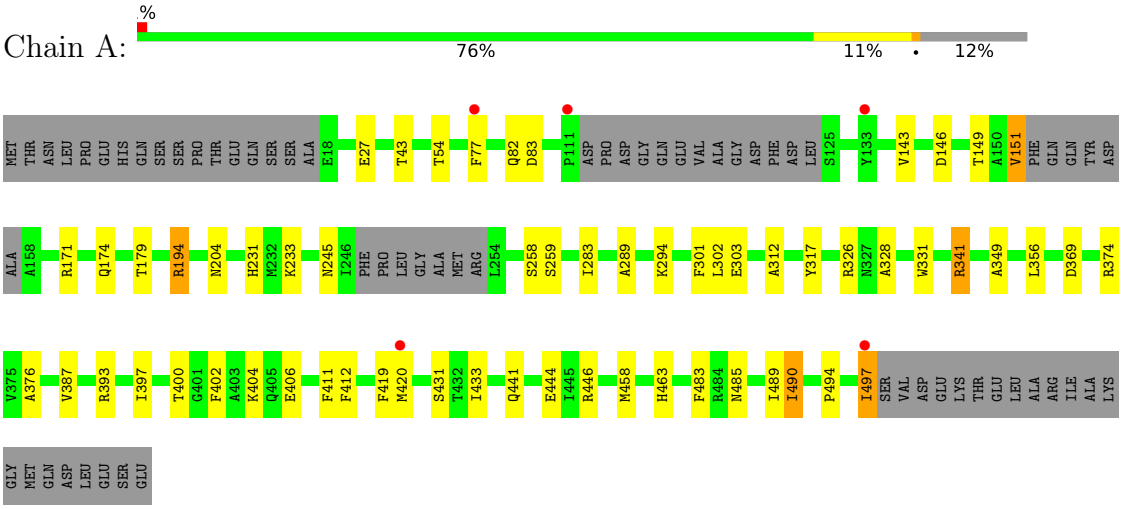
| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|---|---|---|---|---|---|
| 5 | A | 137 | Total O<br>137 137 | 0 | 0 |
| 5 | B | 137 | Total O<br>137 137 | 0 | 0 |
| 5 | C | 142 | Total O<br>142 142 | 0 | 0 |
| 5 | D | 161 | Total O<br>161 161 | 0 | 0 |
| 5 | E | 127 | Total O<br>127 127 | 0 | 0 |
| 5 | F | 133 | Total O<br>133 133 | 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.

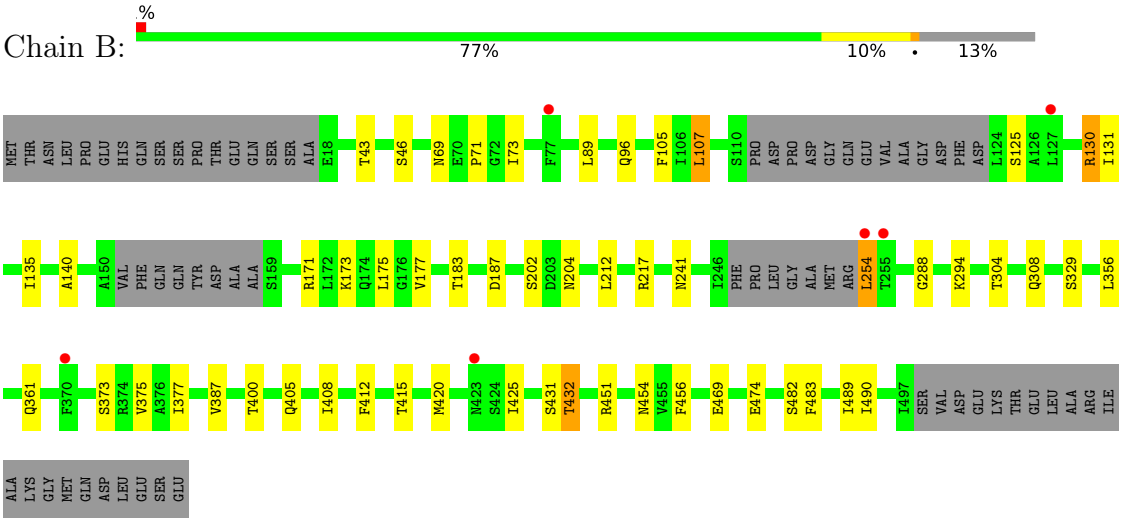
• Molecule 1: Circadian clock protein kinase KaiC

Chain A:



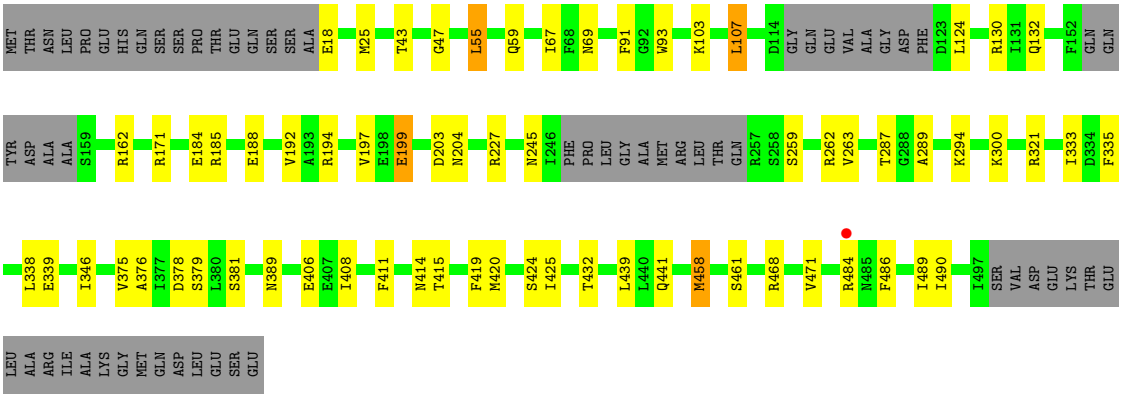

• Molecule 2: Circadian clock protein kinase KaiC

Chain B:



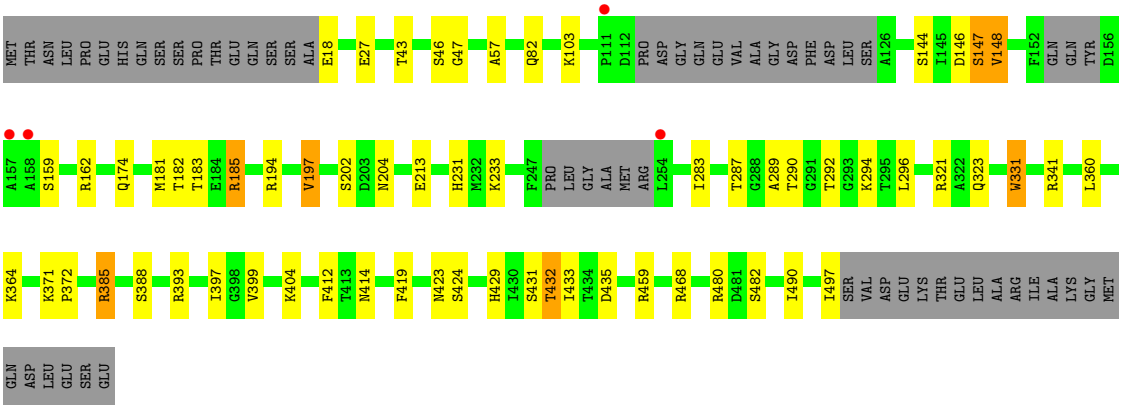

• Molecule 2: Circadian clock protein kinase KaiC

Chain C:



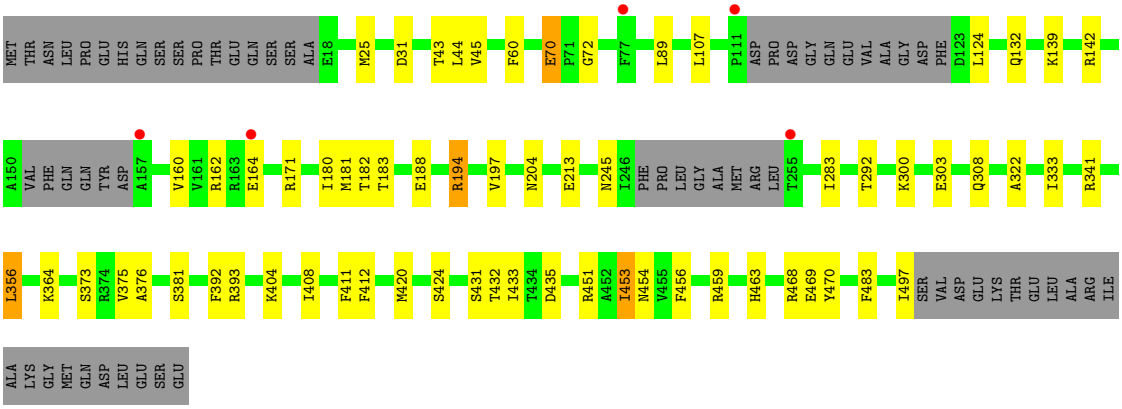

• Molecule 2: Circadian clock protein kinase KaiC

Chain D:

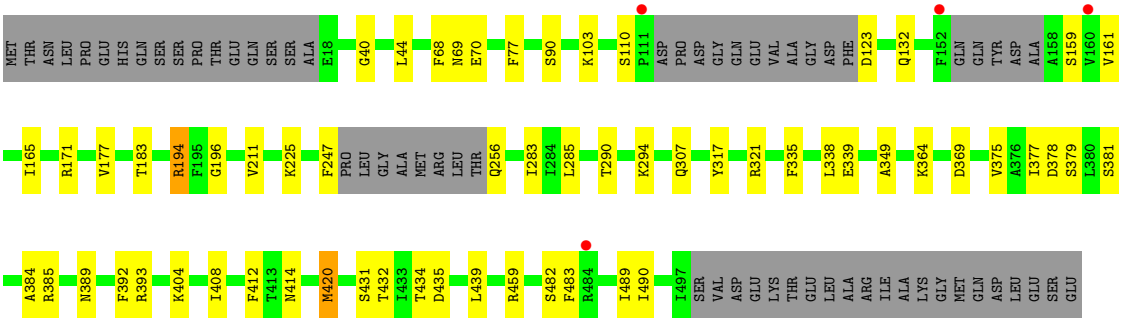

• Molecule 2: Circadian clock protein kinase KaiC

Chain E:


• Molecule 2: Circadian clock protein kinase KaiC

Chain F:


MET THR ASN LEU PRO GLU HIS GLN SER SER PRO THR GLU GLN SER SER ALA E18 G40 L44 F68 N69 E70 F77 S90 K103 S110 P111 ASP PRO ASP GLY GLN GLU VAL ALA GLY ASP PHE D123 Q132 F152 GLN GLN TYR ASP ALA A158 S159 V160 V161
I165 R171 V177 T183 R194 F195 G196 V211 K225 F247 PRO LEU GLY ALA MET ARG LEU THR Q256 I283 I284 L285 T290 K294 Q307 Y317 R321 F335 L338 E339 A349 K364 D369 V375 A376 I377 D378 S379 L380 S381
A384 R385 N389 F392 R393 K404 I408 F412 T413 N414 M420 S431 T432 I433 T434 D435 L439 R459 S482 F483 R484 I489 I490 I497 SER VAL ASP GLU LYS THR GLU LEU ALA ARG ILE ALA LYS GLY MET GLN ASP LEU GLU SER GLU

# 4 Data and refinement statistics (i)

| Property | Value | Source |
|---|---|---|
| Space group | P 21 21 21 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 131.10Å 136.49Å 190.77Å<br>90.00° 90.00° 90.00° | Depositor |
| Resolution (Å) | 47.28 – 2.20<br>47.28 – 2.20 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 99.7 (47.28-2.20)<br>99.7 (47.28-2.20) | Depositor<br>EDS |
| $R_{merge}$ | (Not available) | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 2.35 (at 2.20Å) | Xtriage |
| Refinement program | REFMAC 5.8.0253 | Depositor |
| R, $R_{free}$ | 0.233 , 0.284<br>0.238 , 0.286 | Depositor<br>DCC |
| $R_{free}$ test set | 8569 reflections (4.98%) | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 27.1 | Xtriage |
| Anisotropy | 0.331 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.34 , 39.2 | EDS |
| L-test for twinning[2] | $< \|L\| > = 0.48$, $< L^2 > = 0.31$ | Xtriage |
| Estimated twinning fraction | 0.000 for k,h,-l | Xtriage |
| $F_o$,$F_c$ correlation | 0.94 | EDS |
| Total number of atoms | 21975 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 29.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 44.42 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.5310e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: SEP, MG, TPO, ATP

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 | 0.75 | 0/3508 | 0.85 | 0/4738 |
| 2 | B | 0.74 | 0/3446 | 0.82 | 0/4659 |
| 2 | C | 0.74 | 0/3492 | 0.84 | 0/4716 |
| 2 | D | 0.76 | 0/3529 | 0.84 | 0/4769 |
| 2 | E | 0.73 | 0/3516 | 0.84 | 1/4747 (0.0%) |
| 2 | F | 0.74 | 0/3480 | 0.83 | 0/4703 |
| All | All | 0.74 | 0/20971 | 0.84 | 1/28332 (0.0%) |

There are no bond length outliers.

All (1) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 2 | E | 70 | GLU | CB-CA-C | 5.26 | 120.92 | 110.40 |

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 | 3462 | 0 | 3409 | 35 | 0 |
| 2 | B | 3413 | 0 | 3333 | 30 | 0 |
| 2 | C | 3459 | 0 | 3361 | 37 | 0 |
| 2 | D | 3491 | 0 | 3428 | 36 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|---|---|---|---|---|---|---|
| 2 | E | 3481 | 0 | 3424 | 31 | 0 |
| 2 | F | 3448 | 0 | 3367 | 36 | 0 |
| 3 | A | 62 | 0 | 24 | 0 | 0 |
| 3 | B | 62 | 0 | 24 | 0 | 0 |
| 3 | C | 62 | 0 | 24 | 1 | 0 |
| 3 | D | 62 | 0 | 24 | 2 | 0 |
| 3 | E | 62 | 0 | 24 | 1 | 0 |
| 3 | F | 62 | 0 | 24 | 5 | 0 |
| 4 | A | 2 | 0 | 0 | 0 | 0 |
| 4 | B | 2 | 0 | 0 | 0 | 0 |
| 4 | C | 2 | 0 | 0 | 0 | 0 |
| 4 | D | 2 | 0 | 0 | 0 | 0 |
| 4 | E | 2 | 0 | 0 | 0 | 0 |
| 4 | F | 2 | 0 | 0 | 0 | 0 |
| 5 | A | 137 | 0 | 0 | 1 | 0 |
| 5 | B | 137 | 0 | 0 | 3 | 0 |
| 5 | C | 142 | 0 | 0 | 4 | 0 |
| 5 | D | 161 | 0 | 0 | 4 | 0 |
| 5 | E | 127 | 0 | 0 | 0 | 0 |
| 5 | F | 133 | 0 | 0 | 3 | 0 |
| All | All | 21975 | 0 | 20466 | 191 | 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 5.

The worst 5 of 191 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:D:432:TPO:O3P | 5:D:701:HOH:O | 1.86 | 0.92 |
| 2:C:192:VAL:CB | 2:C:199:GLU:HG2 | 2.11 | 0.80 |
| 2:D:162:ARG:HG3 | 2:D:197:VAL:HG22 | 1.65 | 0.79 |
| 2:C:335:PHE:O | 2:C:339:GLU:HG3 | 1.88 | 0.73 |
| 2:E:160:VAL:O | 2:E:164:GLU:HG2 | 1.89 | 0.73 |

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 | 446/518 (86%) | 432 (97%) | 14 (3%) | 0 | 100 | 100 |
| 2 | B | 443/518 (86%) | 430 (97%) | 13 (3%) | 0 | 100 | 100 |
| 2 | C | 446/518 (86%) | 431 (97%) | 14 (3%) | 1 (0%) | 47 | 55 |
| 2 | D | 450/518 (87%) | 435 (97%) | 15 (3%) | 0 | 100 | 100 |
| 2 | E | 446/518 (86%) | 432 (97%) | 14 (3%) | 0 | 100 | 100 |
| 2 | F | 446/518 (86%) | 432 (97%) | 14 (3%) | 0 | 100 | 100 |
| All | All | 2677/3108 (86%) | 2592 (97%) | 84 (3%) | 1 (0%) | 100 | 100 |

All (1) Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 2 | C | 124 | LEU |

### 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 | 361/440 (82%) | 346 (96%) | 15 (4%) | 30 | 38 |
| 2 | B | 352/439 (80%) | 337 (96%) | 15 (4%) | 29 | 36 |
| 2 | C | 353/439 (80%) | 333 (94%) | 20 (6%) | 20 | 24 |
| 2 | D | 362/439 (82%) | 337 (93%) | 25 (7%) | 15 | 16 |
| 2 | E | 362/439 (82%) | 341 (94%) | 21 (6%) | 20 | 23 |
| 2 | F | 351/439 (80%) | 338 (96%) | 13 (4%) | 34 | 43 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed | Rotameric | Outliers | Percentiles |
|---|---|---|---|---|---|
| All | All | 2141/2635 (81%) | 2032 (95%) | 109 (5%) | 24 29 |

5 of 109 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 2 | D | 174 | GLN |
| 2 | D | 468 | ARG |
| 2 | F | 194 | ARG |
| 2 | D | 194 | ARG |
| 2 | D | 331 | TRP |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

| Mol | Chain | Res | Type |
|---|---|---|---|
| 2 | D | 414 | ASN |
| 2 | E | 463 | HIS |
| 2 | F | 414 | ASN |
| 2 | F | 245 | ASN |
| 2 | E | 454 | ASN |

### 5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

11 non-standard protein/DNA/RNA residues are modelled in this entry.

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

| Mol | Type | Chain | Res | Link | Bond lengths | | | Bond angles | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Counts | RMSZ | $\#\|Z\| > 2$ | Counts | RMSZ | $\#\|Z\| > 2$ |
| 2 | TPO | B | 432 | 2 | 8,10,11 | 1.77 | 1 (12%) | 10,14,16 | 0.82 | 0 |
| 2 | TPO | F | 432 | 2 | 8,10,11 | 1.28 | 1 (12%) | 10,14,16 | 0.82 | 0 |

| Mol | Type | Chain | Res | Link | Bond lengths | | | Bond angles | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Counts | RMSZ | $\#\|Z\|>2$ | Counts | RMSZ | $\#\|Z\|>2$ |
| 2 | TPO | D | 432 | 2 | 8,10,11 | 1.07 | 1 (12%) | 10,14,16 | 0.99 | 1 (10%) |
| 2 | SEP | B | 431 | 2 | 8,9,10 | 0.68 | 0 | 8,12,14 | 1.31 | 1 (12%) |
| 2 | SEP | D | 431 | 2 | 8,9,10 | 0.82 | 0 | 8,12,14 | 1.13 | 1 (12%) |
| 2 | TPO | E | 432 | 2 | 8,10,11 | 1.64 | 1 (12%) | 10,14,16 | 0.86 | 0 |
| 2 | SEP | F | 431 | 2 | 8,9,10 | 0.85 | 0 | 8,12,14 | 1.02 | 1 (12%) |
| 2 | SEP | C | 431 | 2 | 8,9,10 | 0.66 | 0 | 8,12,14 | 0.79 | 0 |
| 1 | SEP | A | 431 | 1 | 8,9,10 | 0.78 | 0 | 8,12,14 | 0.92 | 0 |
| 2 | SEP | E | 431 | 2 | 8,9,10 | 0.73 | 0 | 8,12,14 | 1.00 | 1 (12%) |
| 2 | TPO | C | 432 | 2 | 8,10,11 | 1.62 | 1 (12%) | 10,14,16 | 0.89 | 0 |

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 |
|---|---|---|---|---|---|---|---|
| 2 | TPO | B | 432 | 2 | - | 1/9/11/13 | - |
| 2 | TPO | F | 432 | 2 | - | 4/9/11/13 | - |
| 2 | TPO | D | 432 | 2 | - | 0/9/11/13 | - |
| 2 | SEP | B | 431 | 2 | - | 2/5/8/10 | - |
| 2 | SEP | D | 431 | 2 | - | 1/5/8/10 | - |
| 2 | TPO | E | 432 | 2 | - | 0/9/11/13 | - |
| 2 | SEP | F | 431 | 2 | - | 0/5/8/10 | - |
| 2 | SEP | C | 431 | 2 | - | 2/5/8/10 | - |
| 1 | SEP | A | 431 | 1 | - | 0/5/8/10 | - |
| 2 | SEP | E | 431 | 2 | - | 2/5/8/10 | - |
| 2 | TPO | C | 432 | 2 | - | 0/9/11/13 | - |

All (5) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 2 | B | 432 | TPO | P-OG1 | 4.72 | 1.68 | 1.59 |
| 2 | E | 432 | TPO | P-OG1 | 4.34 | 1.67 | 1.59 |
| 2 | C | 432 | TPO | P-OG1 | 4.22 | 1.67 | 1.59 |
| 2 | F | 432 | TPO | P-OG1 | 3.24 | 1.65 | 1.59 |
| 2 | D | 432 | TPO | P-OG1 | 2.48 | 1.64 | 1.59 |

All (5) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 2 | B | 431 | SEP | OG-CB-CA | 2.83 | 110.90 | 108.14 |
| 2 | D | 431 | SEP | OG-CB-CA | 2.49 | 110.57 | 108.14 |
| 2 | D | 432 | TPO | O-C-CA | -2.19 | 119.03 | 124.78 |
| 2 | F | 431 | SEP | OG-CB-CA | 2.11 | 110.20 | 108.14 |
| 2 | E | 431 | SEP | OG-CB-CA | 2.02 | 110.11 | 108.14 |

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|---|---|---|---|---|
| 2 | C | 431 | SEP | CB-OG-P-O1P |
| 2 | F | 432 | TPO | N-CA-CB-OG1 |
| 2 | F | 432 | TPO | C-CA-CB-CG2 |
| 2 | F | 432 | TPO | N-CA-CB-CG2 |
| 2 | C | 431 | SEP | CB-OG-P-O3P |

There are no ring outliers.

3 monomers are involved in 6 short contacts:

| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|---|---|---|---|---|---|
| 2 | B | 432 | TPO | 2 | 0 |
| 2 | D | 432 | TPO | 3 | 0 |
| 1 | A | 431 | SEP | 1 | 0 |

## 5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry (i)

Of 24 ligands modelled in this entry, 12 are monoatomic - leaving 12 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 | $\#\|Z\|>2$ |
| 3 | ATP | C | 601 | 4 | 26,33,33 | 0.71 | 1 (3%) | 31,52,52 | 1.01 | 1 (3%) |
| 3 | ATP | F | 601 | 4 | 26,33,33 | 0.70 | 0 | 31,52,52 | 1.11 | 4 (12%) |
| 3 | ATP | D | 602 | 4 | 26,33,33 | 0.71 | 0 | 31,52,52 | 0.75 | 1 (3%) |
| 3 | ATP | A | 602 | 4 | 26,33,33 | 0.73 | 0 | 31,52,52 | 0.99 | 2 (6%) |
| 3 | ATP | E | 602 | 4 | 26,33,33 | 0.85 | 0 | 31,52,52 | 0.72 | 0 |
| 3 | ATP | D | 601 | 4 | 26,33,33 | 0.74 | 0 | 31,52,52 | 0.99 | 2 (6%) |
| 3 | ATP | C | 602 | 4 | 26,33,33 | 0.70 | 0 | 31,52,52 | 0.84 | 1 (3%) |
| 3 | ATP | E | 601 | 4 | 26,33,33 | 0.73 | 0 | 31,52,52 | 0.86 | 1 (3%) |
| 3 | ATP | B | 601 | 4 | 26,33,33 | 0.62 | 0 | 31,52,52 | 1.09 | 2 (6%) |
| 3 | ATP | F | 602 | 4 | 26,33,33 | 0.65 | 0 | 31,52,52 | 0.86 | 1 (3%) |
| 3 | ATP | B | 602 | 4 | 26,33,33 | 0.68 | 0 | 31,52,52 | 0.89 | 1 (3%) |
| 3 | ATP | A | 601 | 4 | 26,33,33 | 0.72 | 0 | 31,52,52 | 1.11 | 1 (3%) |

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 |
|---|---|---|---|---|---|---|---|
| 3 | ATP | C | 601 | 4 | - | 4/18/38/38 | 0/3/3/3 |
| 3 | ATP | F | 601 | 4 | - | 0/18/38/38 | 0/3/3/3 |
| 3 | ATP | D | 602 | 4 | - | 2/18/38/38 | 0/3/3/3 |
| 3 | ATP | A | 602 | 4 | - | 2/18/38/38 | 0/3/3/3 |
| 3 | ATP | E | 602 | 4 | - | 0/18/38/38 | 0/3/3/3 |
| 3 | ATP | D | 601 | 4 | - | 1/18/38/38 | 0/3/3/3 |
| 3 | ATP | C | 602 | 4 | - | 3/18/38/38 | 0/3/3/3 |
| 3 | ATP | E | 601 | 4 | - | 6/18/38/38 | 0/3/3/3 |
| 3 | ATP | B | 601 | 4 | - | 6/18/38/38 | 0/3/3/3 |
| 3 | ATP | F | 602 | 4 | - | 2/18/38/38 | 0/3/3/3 |
| 3 | ATP | B | 602 | 4 | - | 2/18/38/38 | 0/3/3/3 |
| 3 | ATP | A | 601 | 4 | - | 0/18/38/38 | 0/3/3/3 |

All (1) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 3 | C | 601 | ATP | C8-N7 | -2.06 | 1.31 | 1.34 |

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

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 3 | C | 601 | ATP | O3'-C3'-C4' | -2.74 | 103.12 | 111.05 |
| 3 | B | 602 | ATP | C5-C6-N6 | 2.63 | 124.35 | 120.35 |
| 3 | D | 601 | ATP | O3'-C3'-C4' | -2.46 | 103.92 | 111.05 |
| 3 | F | 601 | ATP | O3'-C3'-C4' | -2.44 | 104.01 | 111.05 |
| 3 | A | 602 | ATP | PA-O3A-PB | -2.43 | 124.50 | 132.83 |

There are no chirality outliers.

5 of 28 torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|---|---|---|---|---|
| 3 | B | 601 | ATP | C5'-O5'-PA-O1A |
| 3 | C | 601 | ATP | C5'-O5'-PA-O1A |
| 3 | C | 602 | ATP | PB-O3B-PG-O2G |
| 3 | E | 601 | ATP | C5'-O5'-PA-O1A |
| 3 | B | 601 | ATP | C3'-C4'-C5'-O5' |

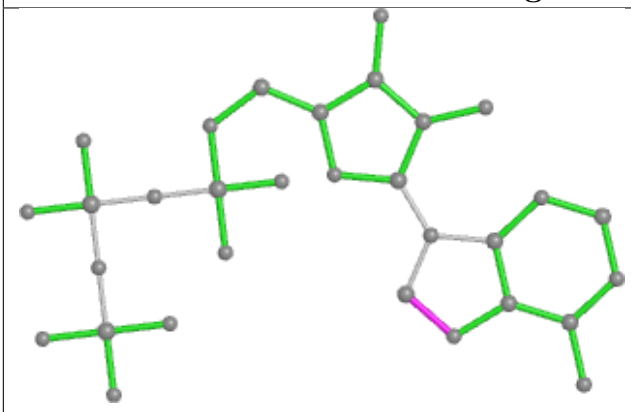
There are no ring outliers.

5 monomers are involved in 9 short contacts:

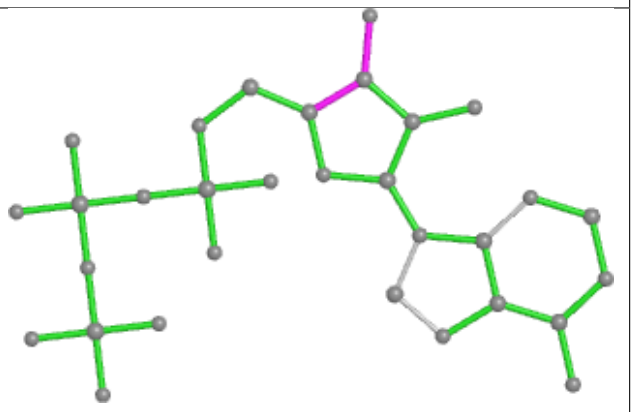
| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|---|---|---|---|---|---|
| 3 | C | 601 | ATP | 1 | 0 |
| 3 | F | 601 | ATP | 4 | 0 |
| 3 | D | 602 | ATP | 2 | 0 |
| 3 | E | 601 | ATP | 1 | 0 |
| 3 | F | 602 | ATP | 1 | 0 |

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.

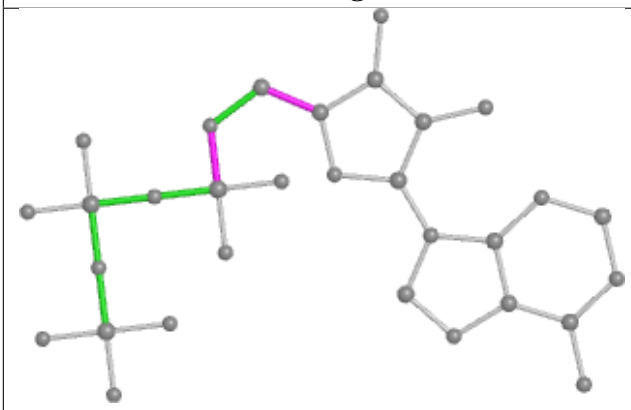
**Ligand ATP C 601**



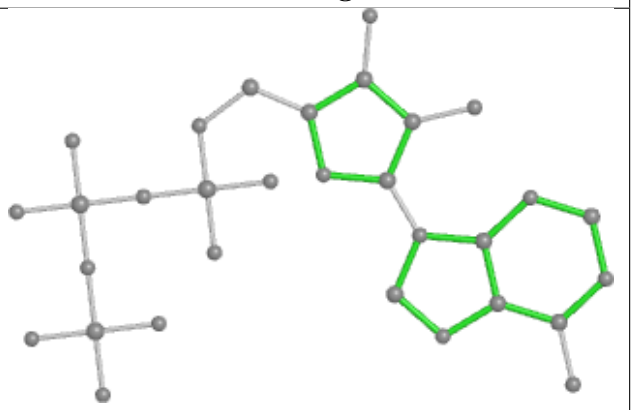
Bond lengths



Bond angles

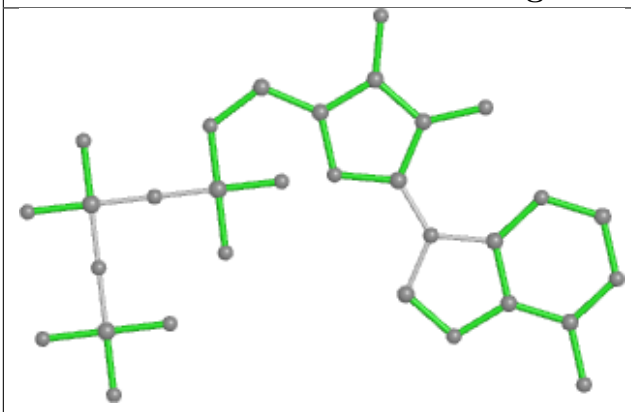


Torsions

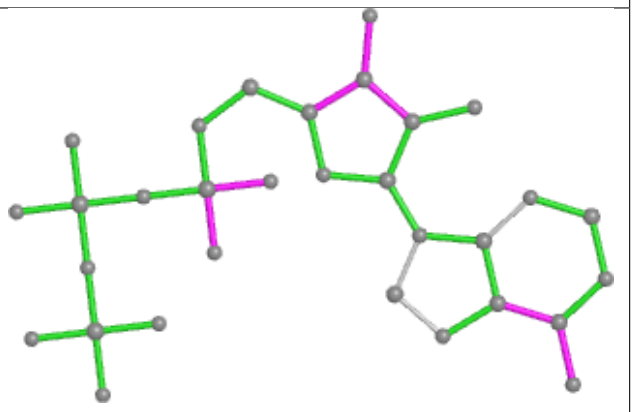


Rings

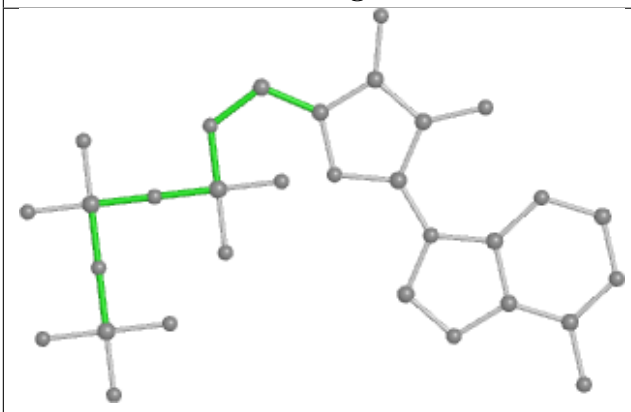
**Ligand ATP F 601**



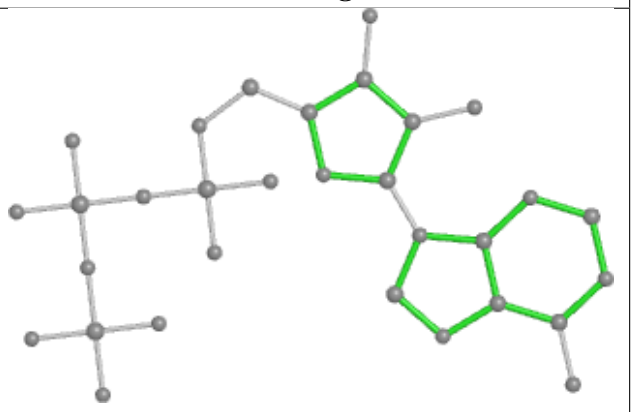
Bond lengths



Bond angles

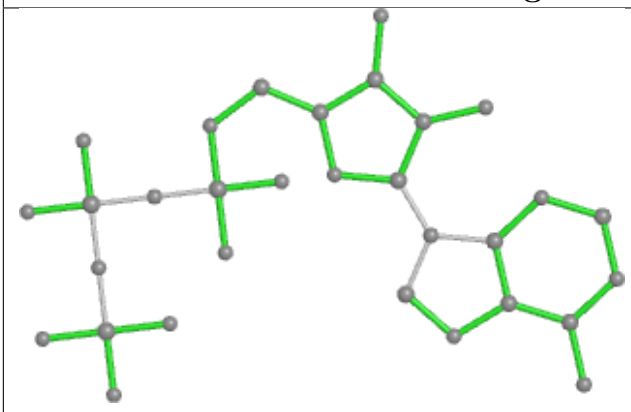


Torsions

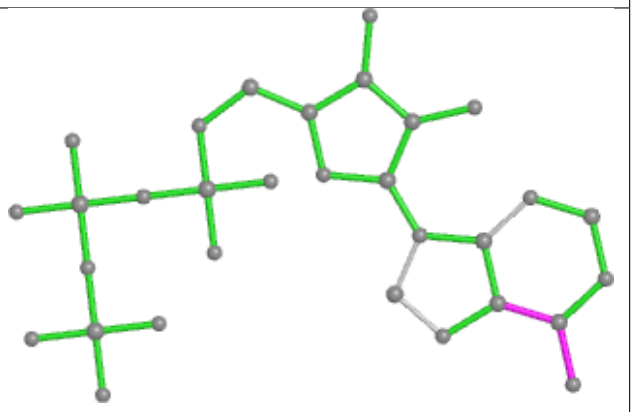


Rings

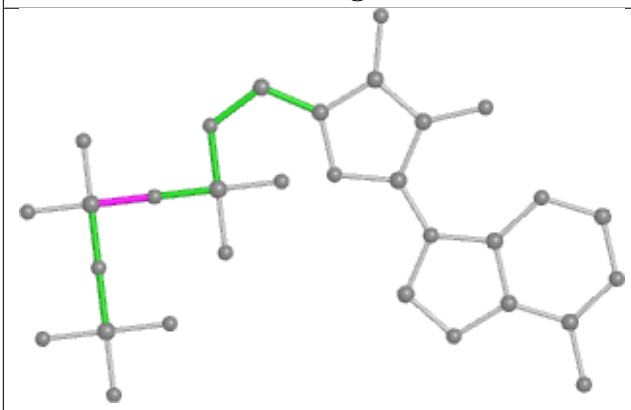
**Ligand ATP D 602**



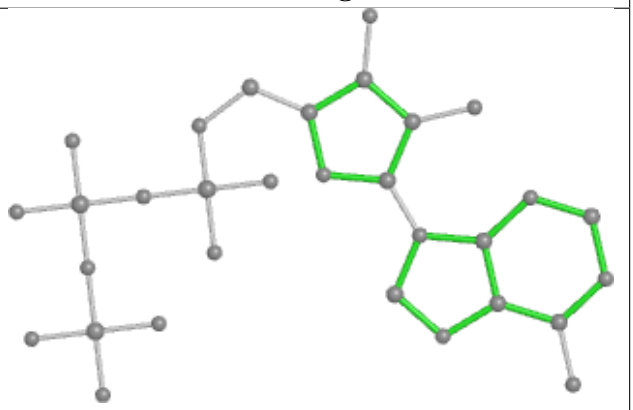
**Bond lengths**



**Bond angles**

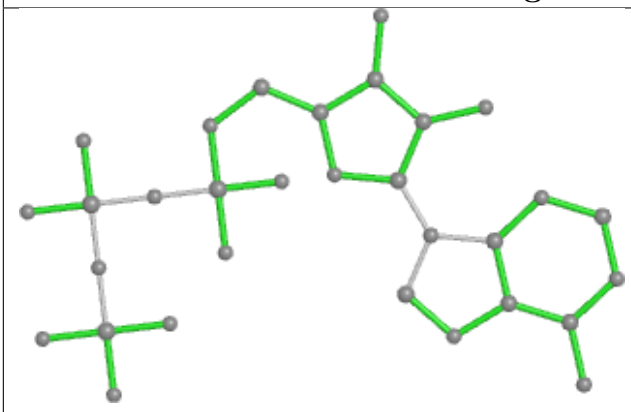


**Torsions**

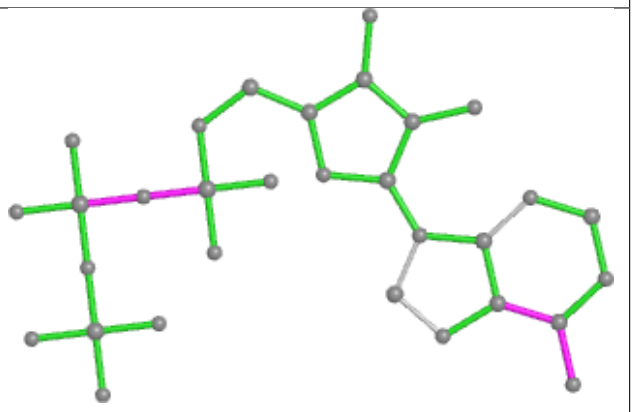


**Rings**

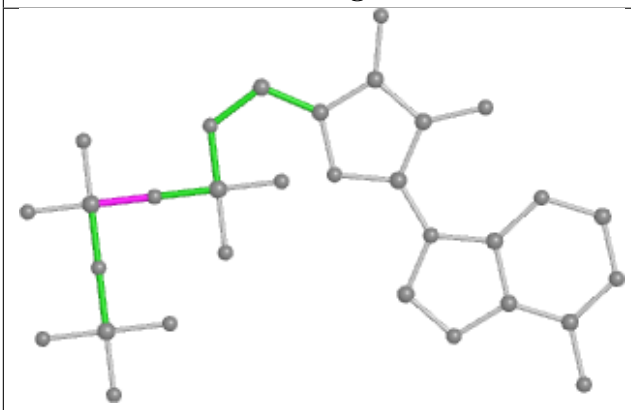
**Ligand ATP A 602**



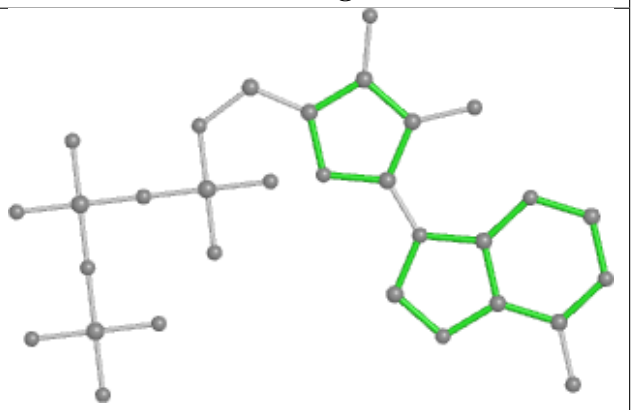
**Bond lengths**



**Bond angles**

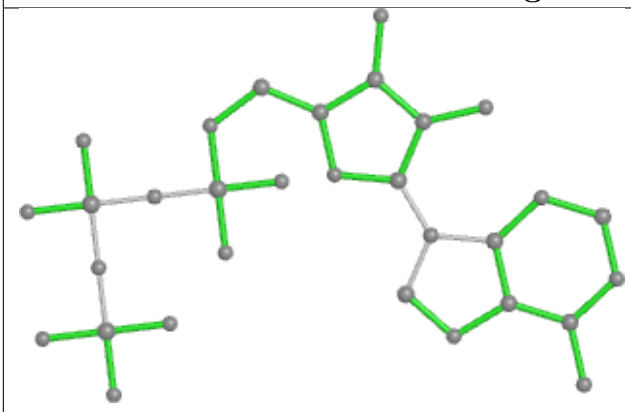


**Torsions**

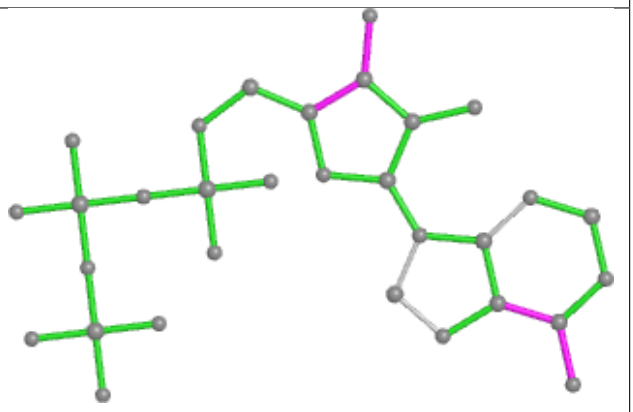


**Rings**

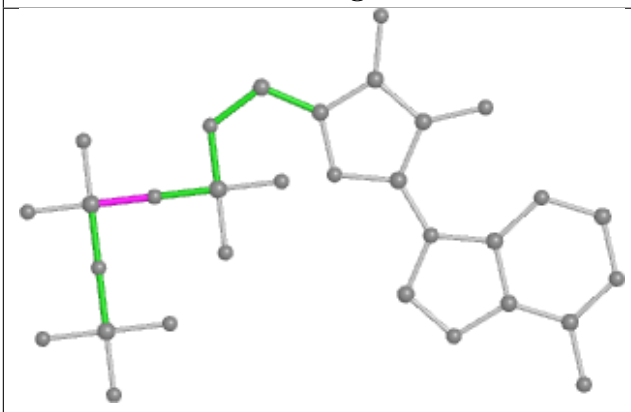
**Ligand ATP D 601**



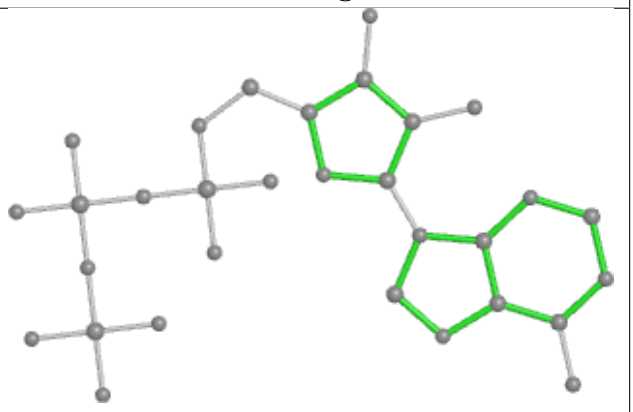
Bond lengths



Bond angles

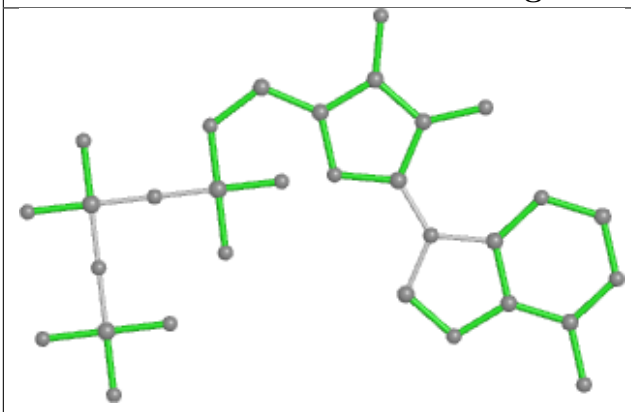


Torsions

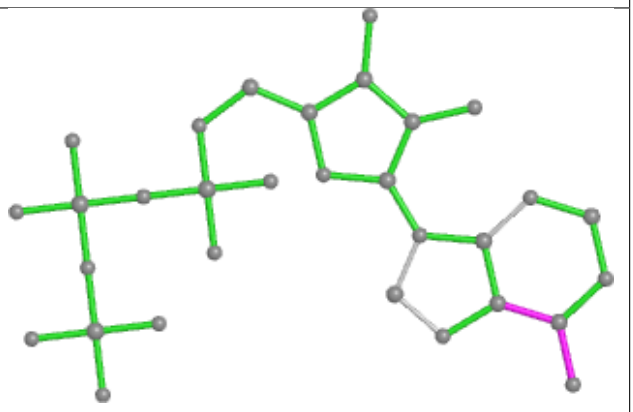


Rings

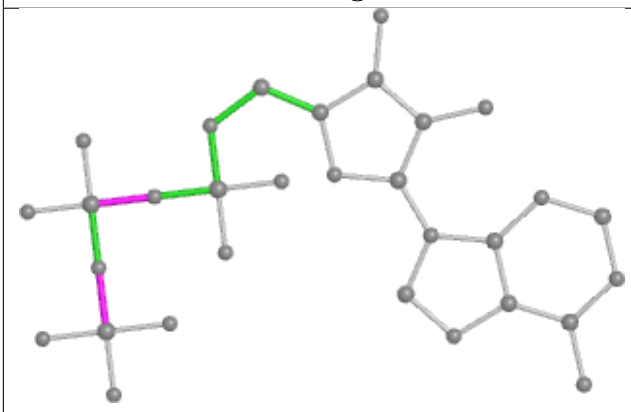
**Ligand ATP C 602**



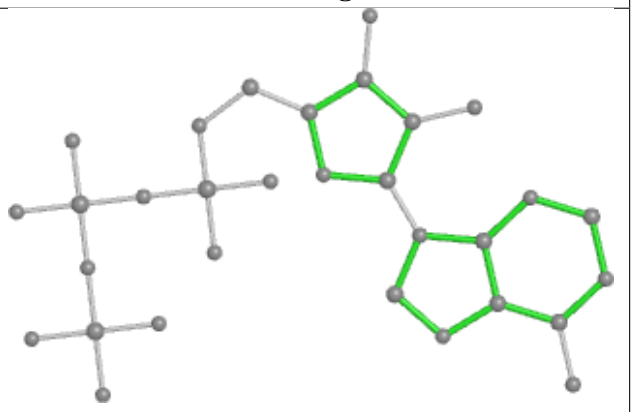
**Bond lengths**



**Bond angles**

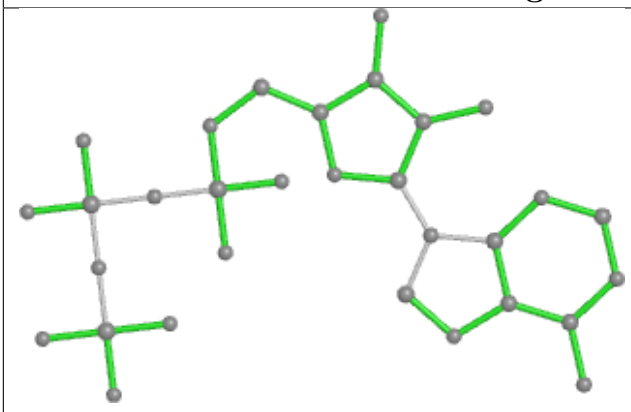


**Torsions**

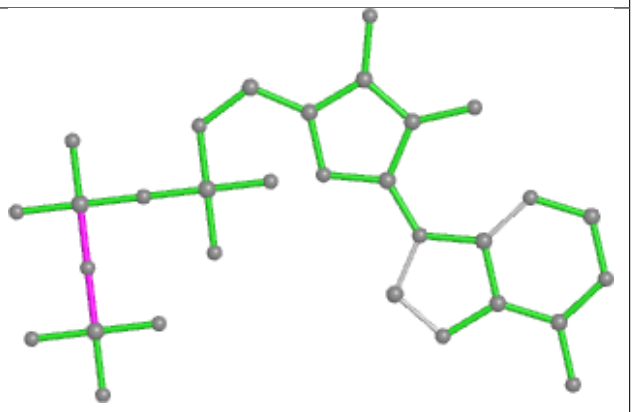


**Rings**

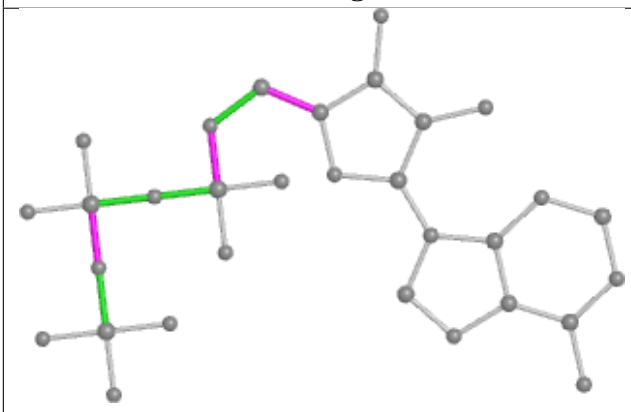
**Ligand ATP E 601**



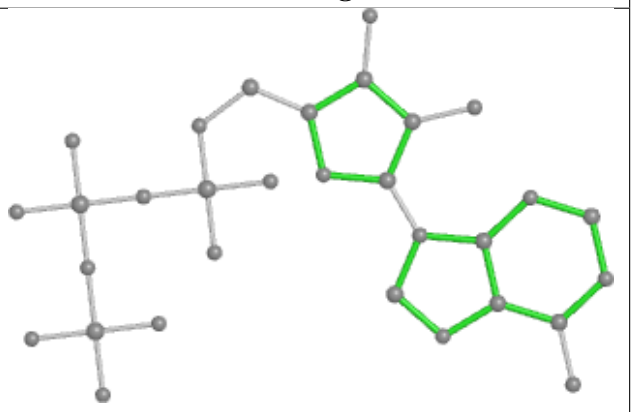
Bond lengths



Bond angles

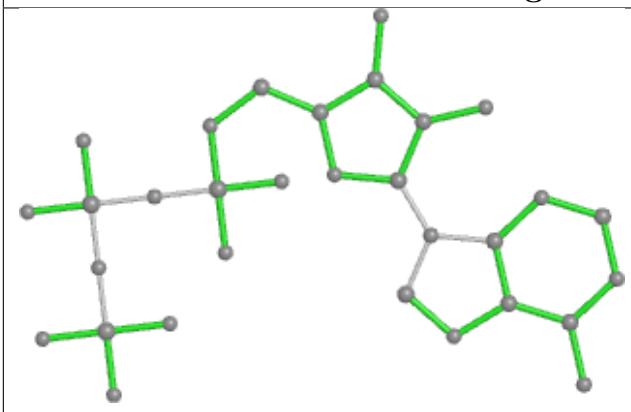


Torsions

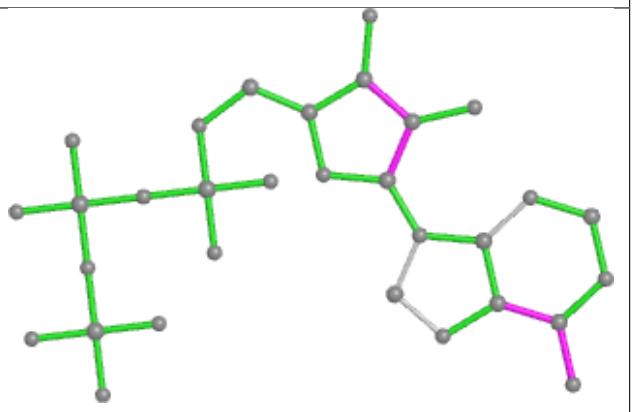


Rings

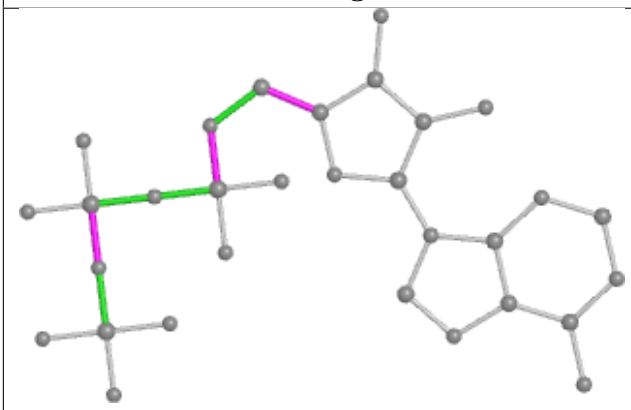
**Ligand ATP B 601**



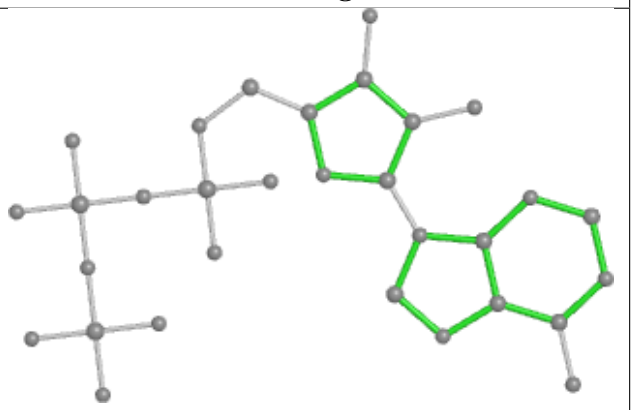
Bond lengths



Bond angles

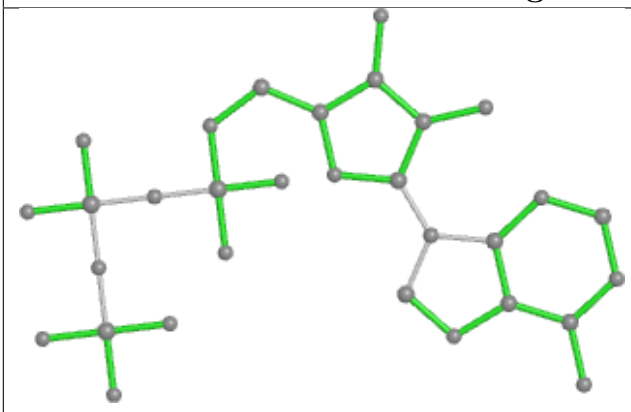


Torsions

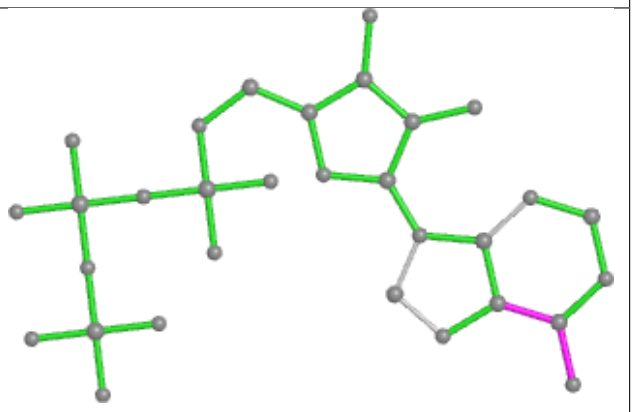


Rings

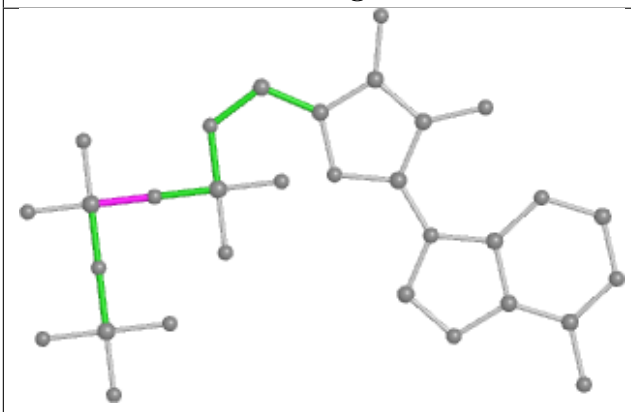
**Ligand ATP F 602**



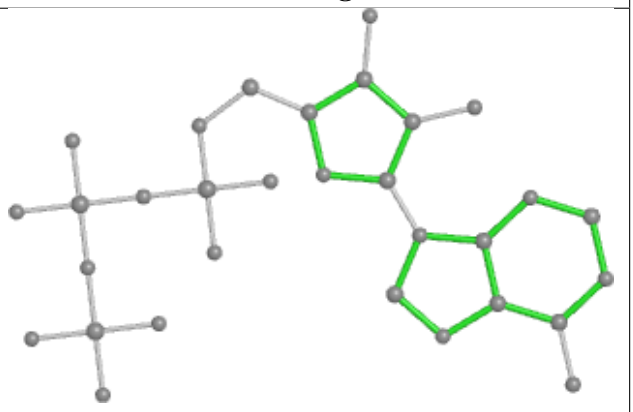
Bond lengths



Bond angles

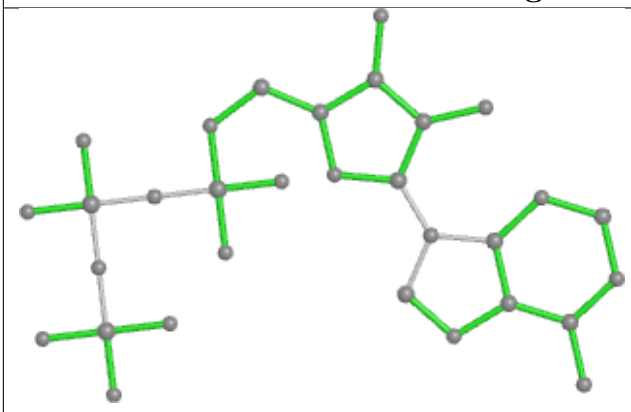


Torsions

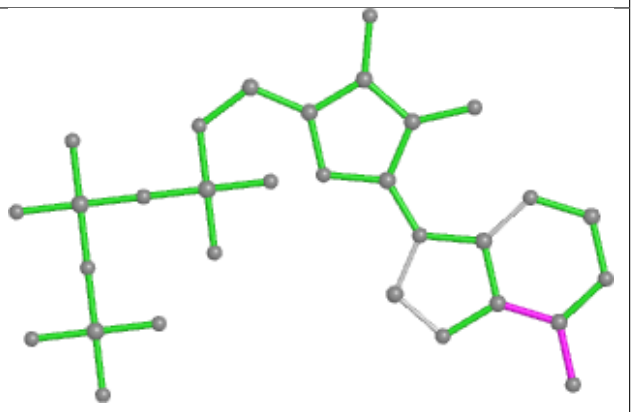


Rings

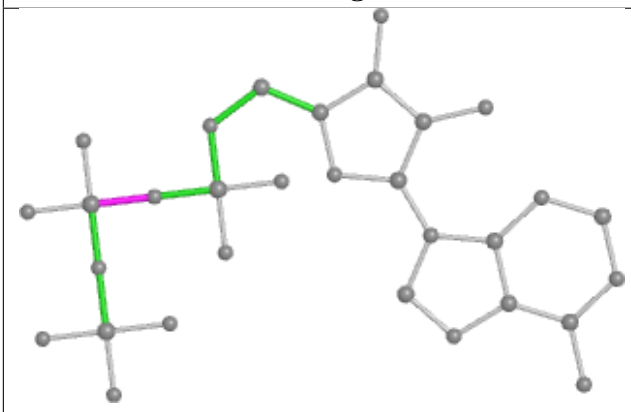
**Ligand ATP B 602**



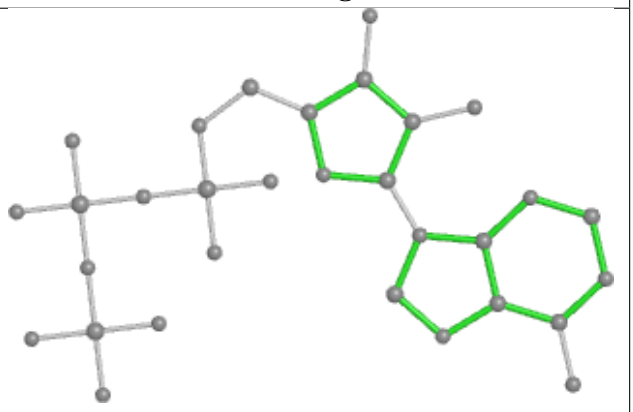
**Bond lengths**



**Bond angles**

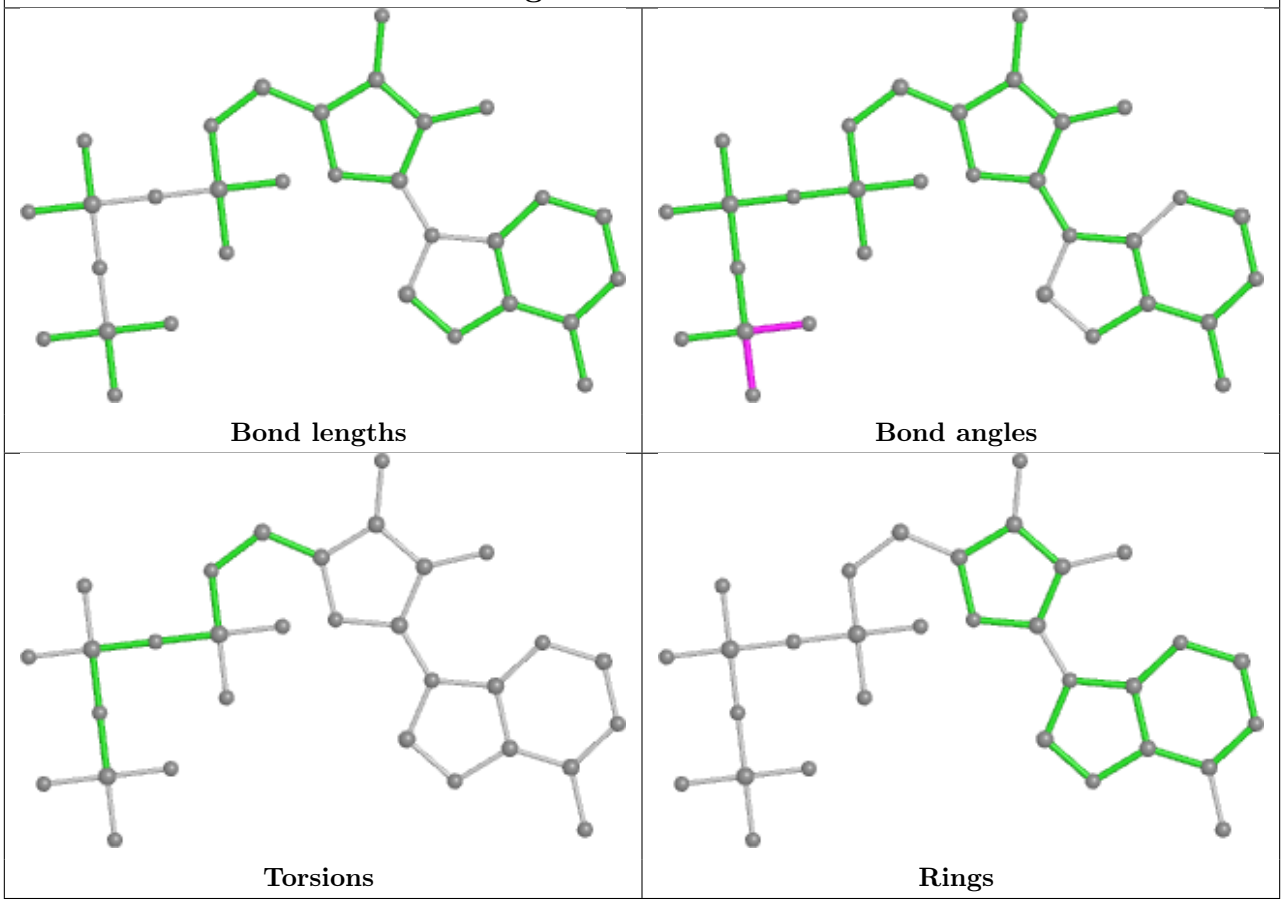


**Torsions**



**Rings**

**Ligand ATP A 601**




## 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 | 453/518 (87%) | -0.19 | 5 (1%) | 80 | 79 | 15, 27, 46, 68 | 0 |
| 2 | B | 450/518 (86%) | -0.16 | 6 (1%) | 77 | 75 | 17, 30, 48, 66 | 0 |
| 2 | C | 454/518 (87%) | -0.14 | 1 (0%) | 95 | 94 | 13, 29, 48, 87 | 0 |
| 2 | D | 456/518 (88%) | -0.15 | 4 (0%) | 84 | 83 | 14, 26, 45, 78 | 0 |
| 2 | E | 453/518 (87%) | -0.16 | 5 (1%) | 80 | 79 | 14, 29, 48, 74 | 0 |
| 2 | F | 454/518 (87%) | -0.13 | 4 (0%) | 84 | 83 | 13, 29, 47, 69 | 0 |
| All | All | 2720/3108 (87%) | -0.15 | 25 (0%) | 84 | 83 | 13, 29, 48, 87 | 0 |

The worst 5 of 25 RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|---|---|---|---|---|
| 2 | D | 111 | PRO | 4.3 |
| 2 | D | 157 | ALA | 4.0 |
| 1 | A | 111 | PRO | 3.9 |
| 2 | C | 484 | ARG | 3.5 |
| 2 | D | 254 | LEU | 3.3 |

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

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 |
|---|---|---|---|---|---|---|---|---|
| 2 | TPO | C | 432 | 11/12 | 0.85 | 0.14 | 28,31,43,48 | 4 |
| 2 | TPO | B | 432 | 11/12 | 0.86 | 0.15 | 25,29,35,41 | 4 |
| 2 | TPO | E | 432 | 11/12 | 0.86 | 0.16 | 24,29,37,43 | 4 |
| 2 | TPO | D | 432 | 11/12 | 0.87 | 0.19 | 23,27,43,46 | 4 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Type | Chain | Res | Atoms | RSCC | RSR | B-factors(Å$^2$) | Q<0.9 |
|---|---|---|---|---|---|---|---|---|
| 2 | TPO | F | 432 | 11/12 | 0.89 | 0.14 | 25,36,61,62 | 0 |
| 2 | SEP | F | 431 | 10/11 | 0.90 | 0.14 | 22,30,42,45 | 0 |
| 2 | SEP | C | 431 | 10/11 | 0.94 | 0.11 | 27,31,36,36 | 0 |
| 2 | SEP | B | 431 | 10/11 | 0.94 | 0.12 | 24,27,30,33 | 4 |
| 1 | SEP | A | 431 | 10/11 | 0.94 | 0.10 | 24,28,34,38 | 0 |
| 2 | SEP | E | 431 | 10/11 | 0.96 | 0.09 | 23,26,35,36 | 0 |
| 2 | SEP | D | 431 | 10/11 | 0.96 | 0.12 | 23,25,29,29 | 0 |

## 6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

## 6.4 Ligands (i)

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

| Mol | Type | Chain | Res | Atoms | RSCC | RSR | B-factors(Å$^2$) | Q<0.9 |
|---|---|---|---|---|---|---|---|---|
| 4 | MG | B | 603 | 1/1 | 0.91 | 0.14 | 30,30,30,30 | 0 |
| 4 | MG | C | 603 | 1/1 | 0.91 | 0.06 | 30,30,30,30 | 0 |
| 4 | MG | F | 604 | 1/1 | 0.91 | 0.04 | 24,24,24,24 | 0 |
| 4 | MG | F | 603 | 1/1 | 0.93 | 0.08 | 37,37,37,37 | 0 |
| 4 | MG | A | 604 | 1/1 | 0.95 | 0.09 | 19,19,19,19 | 0 |
| 3 | ATP | D | 601 | 31/31 | 0.96 | 0.11 | 15,22,26,27 | 0 |
| 3 | ATP | F | 601 | 31/31 | 0.96 | 0.11 | 17,23,29,31 | 0 |
| 3 | ATP | F | 602 | 31/31 | 0.96 | 0.10 | 19,23,27,32 | 0 |
| 3 | ATP | B | 601 | 31/31 | 0.96 | 0.10 | 17,23,28,29 | 0 |
| 3 | ATP | B | 602 | 31/31 | 0.97 | 0.11 | 15,22,26,29 | 0 |
| 3 | ATP | C | 601 | 31/31 | 0.97 | 0.10 | 18,23,31,35 | 0 |
| 3 | ATP | C | 602 | 31/31 | 0.97 | 0.10 | 16,21,32,37 | 0 |
| 3 | ATP | A | 601 | 31/31 | 0.97 | 0.12 | 16,23,29,34 | 0 |
| 4 | MG | B | 604 | 1/1 | 0.97 | 0.07 | 23,23,23,23 | 0 |
| 3 | ATP | D | 602 | 31/31 | 0.97 | 0.10 | 16,23,25,26 | 0 |
| 4 | MG | D | 604 | 1/1 | 0.97 | 0.11 | 17,17,17,17 | 0 |
| 4 | MG | E | 603 | 1/1 | 0.97 | 0.03 | 25,25,25,25 | 0 |
| 4 | MG | E | 604 | 1/1 | 0.97 | 0.07 | 26,26,26,26 | 0 |
| 3 | ATP | E | 601 | 31/31 | 0.97 | 0.11 | 17,25,31,33 | 0 |
| 3 | ATP | E | 602 | 31/31 | 0.97 | 0.11 | 16,22,26,26 | 0 |
| 4 | MG | A | 603 | 1/1 | 0.98 | 0.04 | 35,35,35,35 | 0 |

*Continued on next page...*

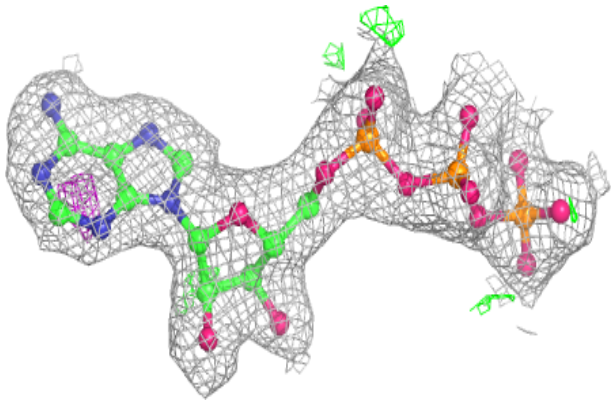
*Continued from previous page...*

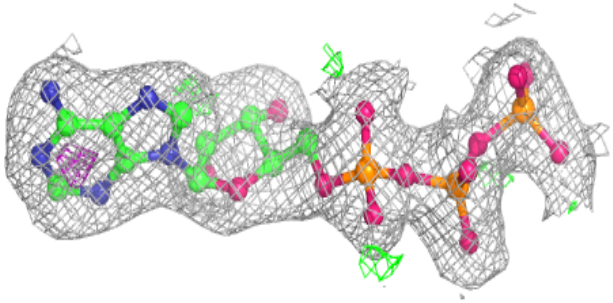
| Mol | Type | Chain | Res | Atoms | RSCC | RSR | B-factors(Å$^2$) | Q<0.9 |
|---|---|---|---|---|---|---|---|---|
| 4 | MG | C | 604 | 1/1 | 0.98 | 0.05 | 25,25,25,25 | 0 |
| 4 | MG | D | 603 | 1/1 | 0.98 | 0.07 | 17,17,17,17 | 0 |
| 3 | ATP | A | 602 | 31/31 | 0.98 | 0.10 | 16,20,25,27 | 0 |

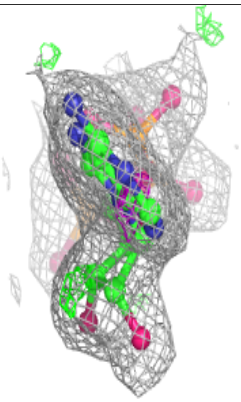
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around ATP D 601:**

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

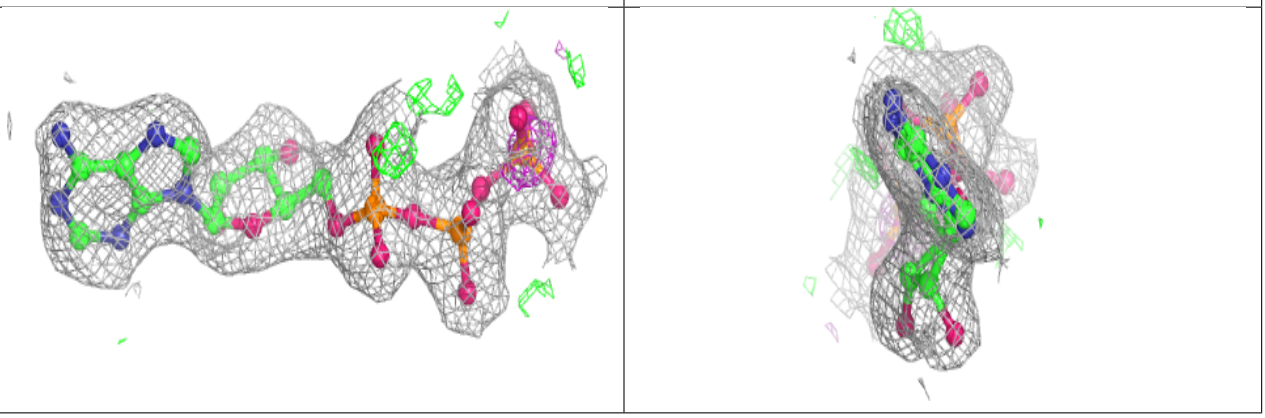






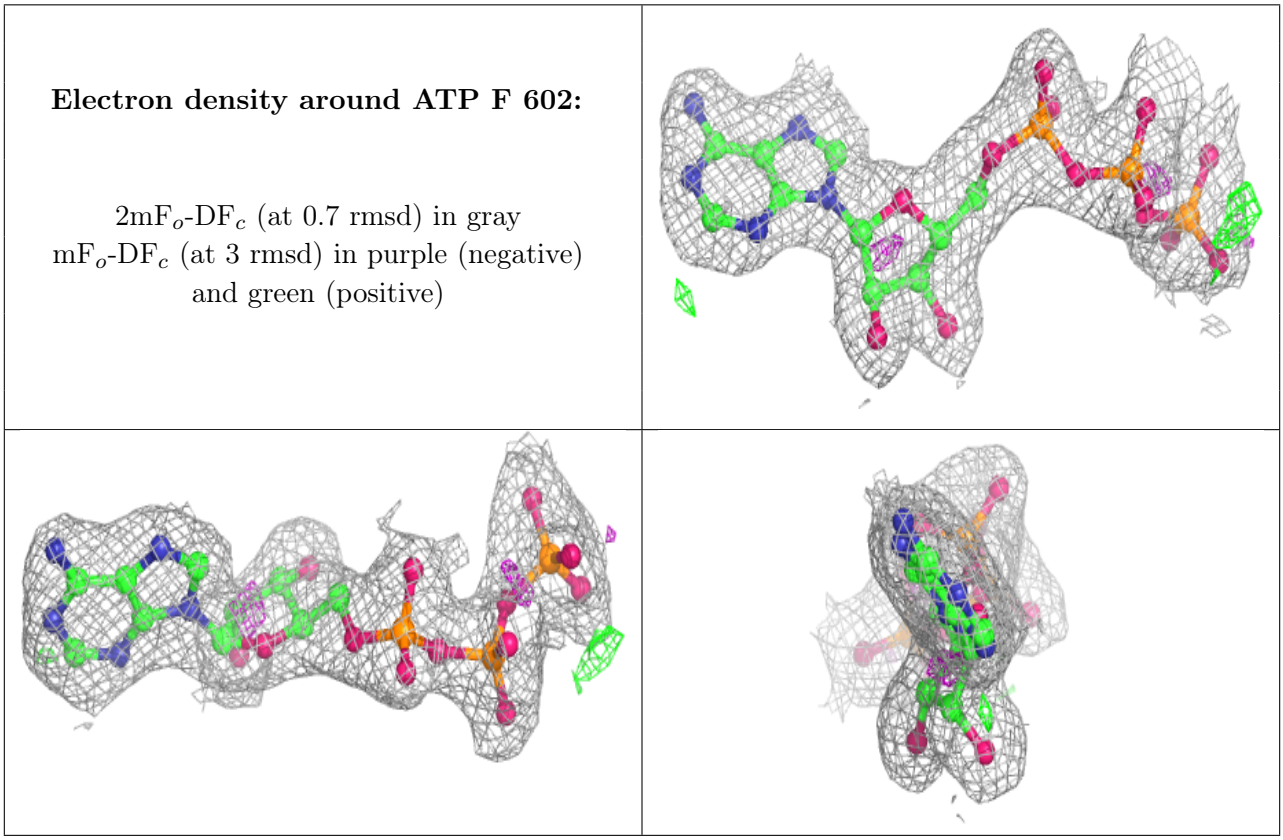
**Electron density around ATP F 601:**

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



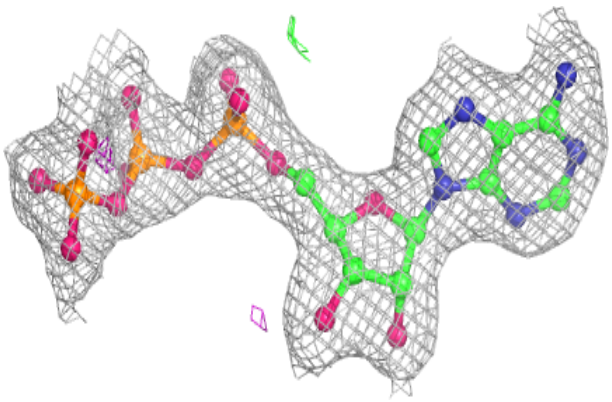
**Electron density around ATP F 602:**

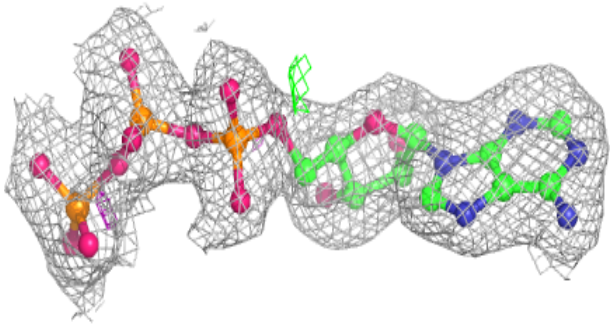
$2mF_o$-$DF_c$ (at 0.7 rmsd) in gray
$mF_o$-$DF_c$ (at 3 rmsd) in purple (negative) and green (positive)

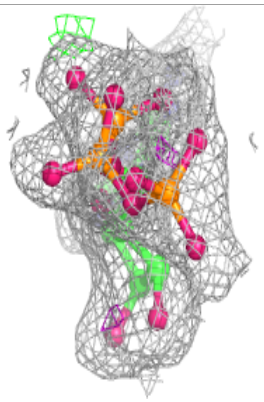


**Electron density around ATP B 601:**

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

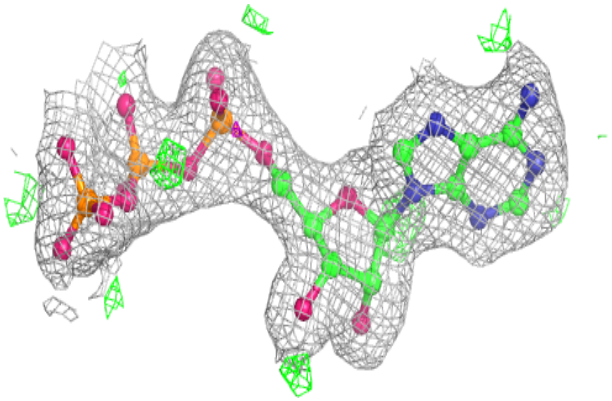


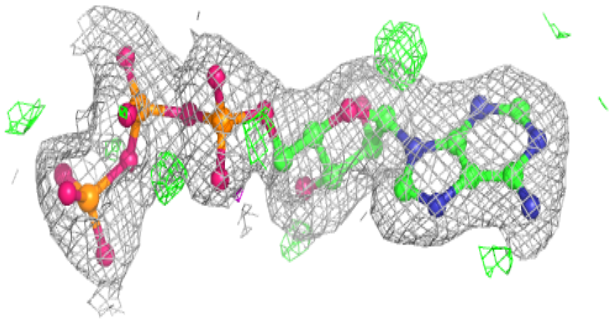


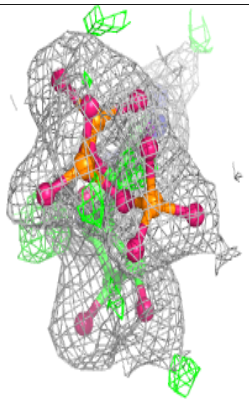


**Electron density around ATP B 602:**

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

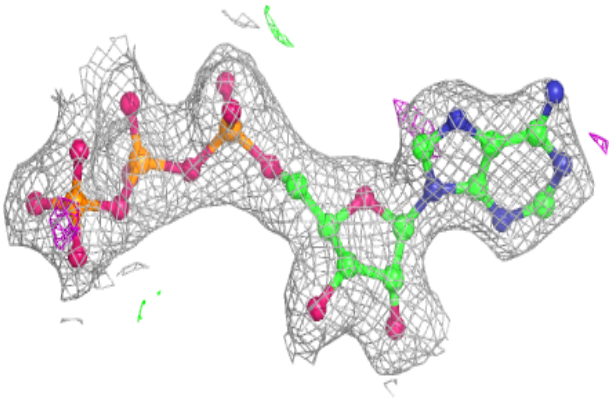


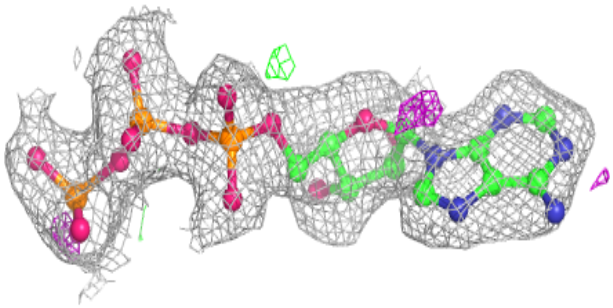


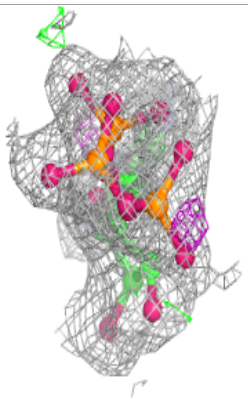


**Electron density around ATP C 601:**

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

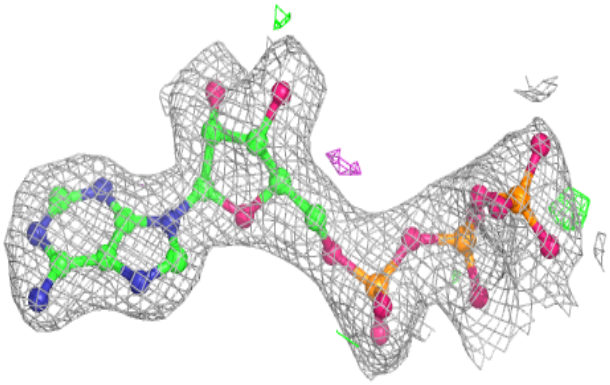


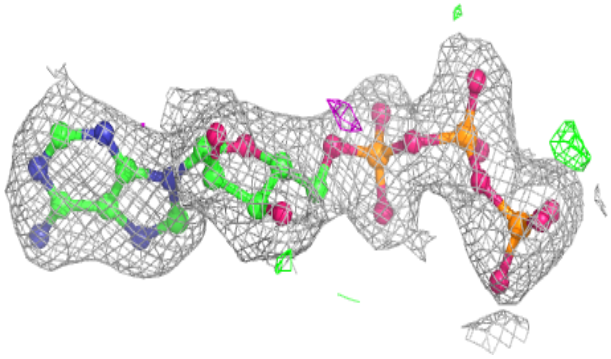


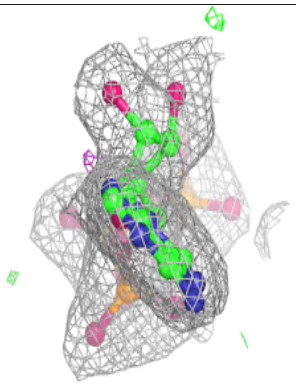


**Electron density around ATP C 602:**

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

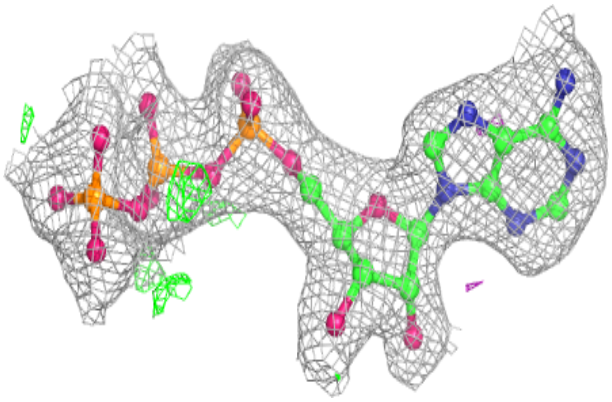


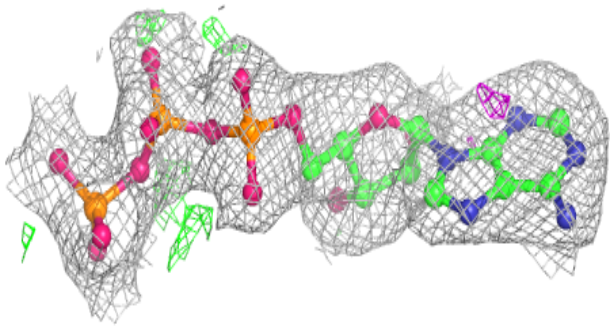


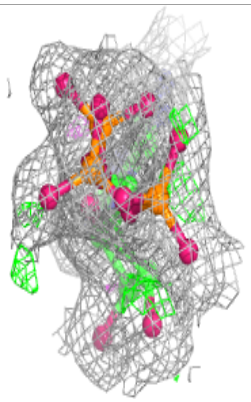


**Electron density around ATP A 601:**

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

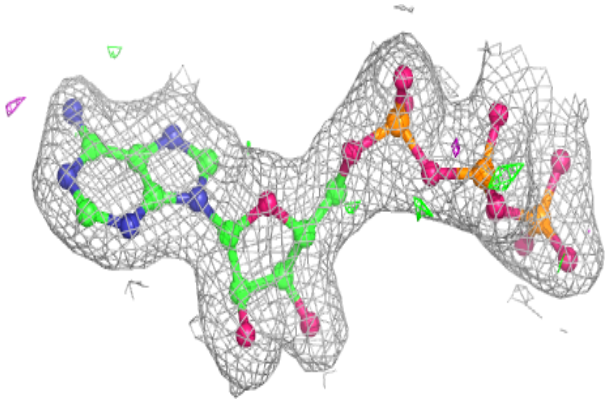


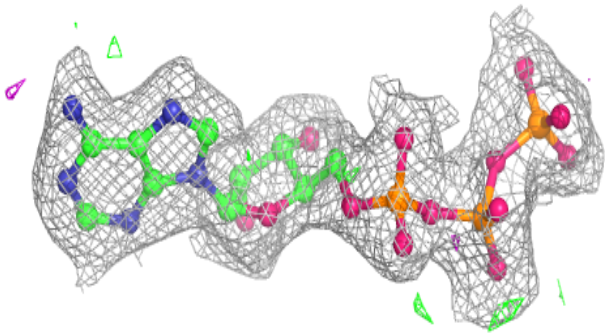


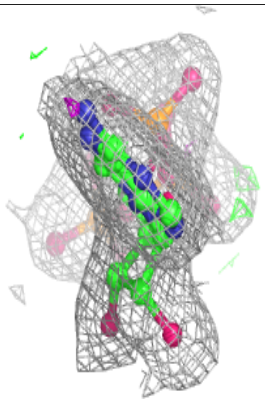


**Electron density around ATP D 602:**

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

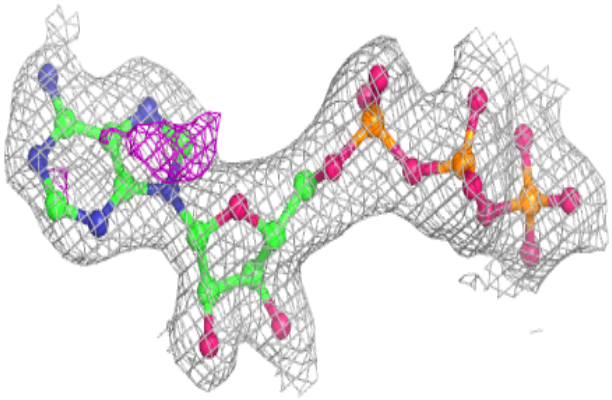


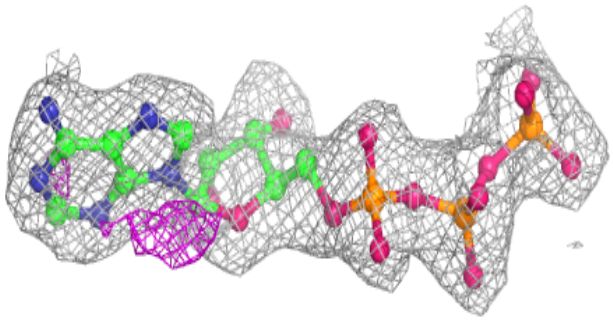


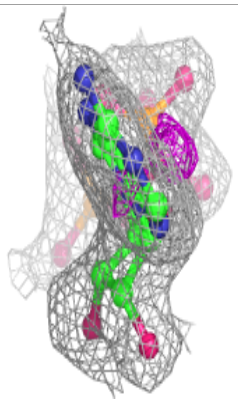


**Electron density around ATP E 601:**

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

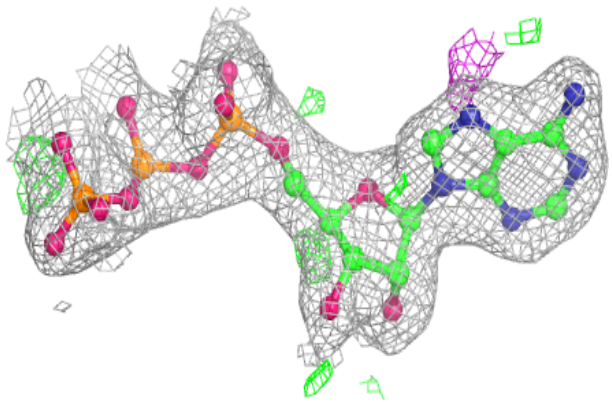


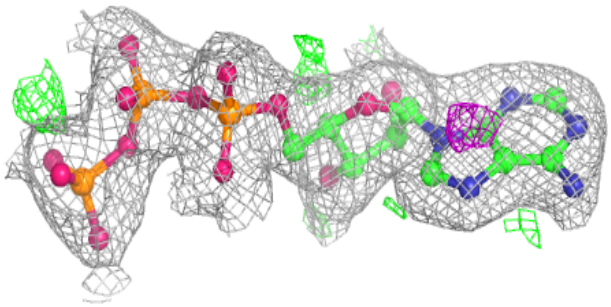


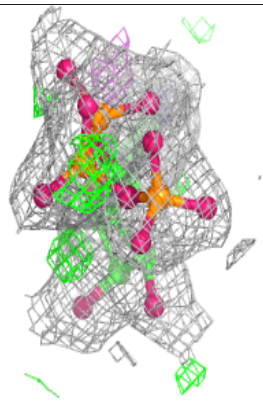


**Electron density around ATP A 602:**

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







## 6.5 Other polymers (i)

There are no such residues in this entry.