# WORLDWIDE PDB PROTEIN DATA BANK

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

Aug 19, 2024 – 04:22 PM EDT

|  |  |  |
|---|---|---|
| PDB ID | : | 8VZV |
| Title | : | Human TDO (hTDO) in complex with LM10 |
| Authors | : | Ishigami, I.; Yeh, S.-R.; Lewis-Ballester, A. |
| Deposited on | : | 2024-02-12 |
| Resolution | : | 2.29 Å(reported) |

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

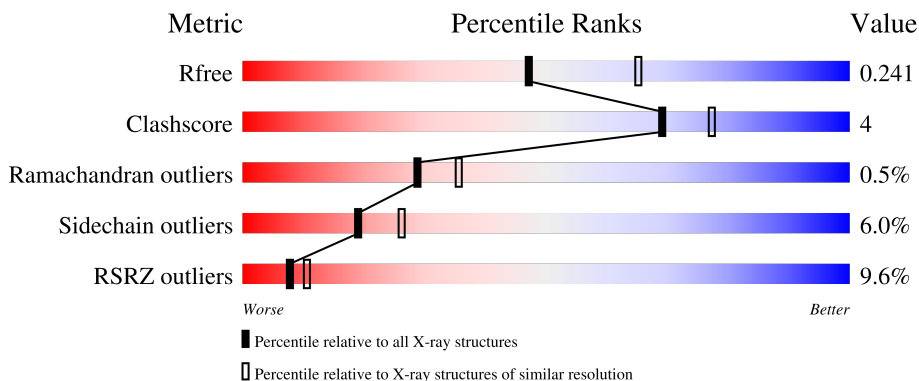|  |  |  |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Mogul | : | 1.8.5 (274361), CSD as541be (2020) |
| Xtriage (Phenix) | : | 1.13 |
| EDS | : | 2.37.1 |
| buster-report | : | 1.1.7 (2018) |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Refmac | : | 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) | : | 2.37.1 |

# 1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.29 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 | Percentile Ranks | Value |
|--------|------------------|-------|
| Rfree | | 0.241 |
| Clashscore | | 4 |
| Ramachandran outliers | | 0.5% |
| Sidechain outliers | | 6.0% |
| RSRZ outliers | | 9.6% |

*Worse* *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

| Metric | Whole archive (#Entries) | Similar resolution (#Entries, resolution range(Å)) |
|--------|--------------------------|-----------------------------------------------------|
| $R_{free}$ | 130704 | 5042 (2.30-2.30) |
| Clashscore | 141614 | 5643 (2.30-2.30) |
| Ramachandran outliers | 138981 | 5575 (2.30-2.30) |
| Sidechain outliers | 138945 | 5575 (2.30-2.30) |
| RSRZ outliers | 127900 | 4938 (2.30-2.30) |

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 | AAA | 380 | 8% / 80% 9% · 9% |
| 1 | BBB | 380 | 7% / 81% 9% · 9% |
| 1 | CCC | 380 | 12% / 73% 10% · 14% |
| 1 | DDD | 380 | 7% / 80% 8% · 10% |

WORLDWIDE PDB
PROTEIN DATA BANK

# 2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 11959 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 Tryptophan 2,3-dioxygenase.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | AAA | 344 | Total 2915 | C 1876 | N 507 | O 521 | S 11 | 0 | 0 | 0 |
| 1 | BBB | 344 | Total 2899 | C 1866 | N 503 | O 519 | S 11 | 0 | 0 | 0 |
| 1 | CCC | 326 | Total 2761 | C 1780 | N 482 | O 488 | S 11 | 0 | 0 | 0 |
| 1 | DDD | 343 | Total 2908 | C 1869 | N 509 | O 519 | S 11 | 0 | 0 | 0 |

There are 32 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|---|---|---|---|---|---|
| AAA | 17 | MET | - | initiating methionine | UNP P48775 |
| AAA | 390 | GLU | - | expression tag | UNP P48775 |
| AAA | 391 | HIS | - | expression tag | UNP P48775 |
| AAA | 392 | HIS | - | expression tag | UNP P48775 |
| AAA | 393 | HIS | - | expression tag | UNP P48775 |
| AAA | 394 | HIS | - | expression tag | UNP P48775 |
| AAA | 395 | HIS | - | expression tag | UNP P48775 |
| AAA | 396 | HIS | - | expression tag | UNP P48775 |
| BBB | 17 | MET | - | initiating methionine | UNP P48775 |
| BBB | 390 | GLU | - | expression tag | UNP P48775 |
| BBB | 391 | HIS | - | expression tag | UNP P48775 |
| BBB | 392 | HIS | - | expression tag | UNP P48775 |
| BBB | 393 | HIS | - | expression tag | UNP P48775 |
| BBB | 394 | HIS | - | expression tag | UNP P48775 |
| BBB | 395 | HIS | - | expression tag | UNP P48775 |
| BBB | 396 | HIS | - | expression tag | UNP P48775 |
| CCC | 17 | MET | - | initiating methionine | UNP P48775 |
| CCC | 390 | GLU | - | expression tag | UNP P48775 |
| CCC | 391 | HIS | - | expression tag | UNP P48775 |
| CCC | 392 | HIS | - | expression tag | UNP P48775 |
| CCC | 393 | HIS | - | expression tag | UNP P48775 |

*Continued on next page...*

*Continued from previous page...*

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------|-----------|
| CCC | 394 | HIS | - | expression tag | UNP P48775 |
| CCC | 395 | HIS | - | expression tag | UNP P48775 |
| CCC | 396 | HIS | - | expression tag | UNP P48775 |
| DDD | 17 | MET | - | initiating methionine | UNP P48775 |
| DDD | 390 | GLU | - | expression tag | UNP P48775 |
| DDD | 391 | HIS | - | expression tag | UNP P48775 |
| DDD | 392 | HIS | - | expression tag | UNP P48775 |
| DDD | 393 | HIS | - | expression tag | UNP P48775 |
| DDD | 394 | HIS | - | expression tag | UNP P48775 |
| DDD | 395 | HIS | - | expression tag | UNP P48775 |
| DDD | 396 | HIS | - | expression tag | UNP P48775 |

- Molecule 2 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$) (labeled as "Ligand of Interest" by depositor).
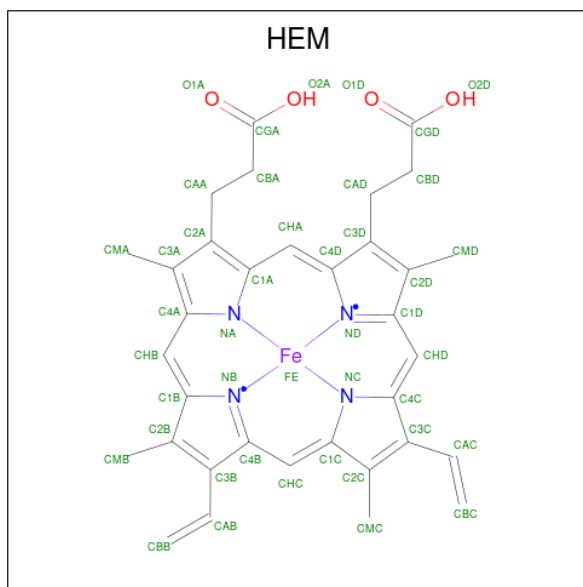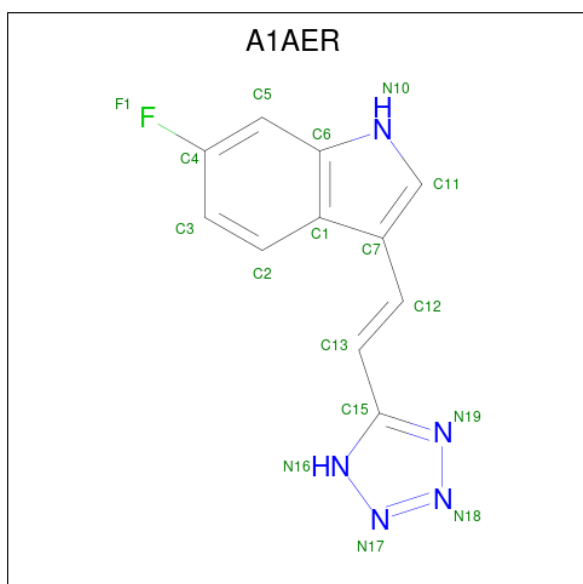


| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf |
|-----|-------|----------|-------|-----|-----|-----|-----|---------|---------|
| 2 | AAA | 1 | Total | C | Fe | N | O | 0 | 0 |
| | | | 43 | 34 | 1 | 4 | 4 | | |
| 2 | BBB | 1 | Total | C | Fe | N | O | 0 | 0 |
| | | | 43 | 34 | 1 | 4 | 4 | | |
| 2 | CCC | 1 | Total | C | Fe | N | O | 0 | 0 |
| | | | 43 | 34 | 1 | 4 | 4 | | |
| 2 | DDD | 1 | Total | C | Fe | N | O | 0 | 0 |
| | | | 43 | 34 | 1 | 4 | 4 | | |

- Molecule 3 is alpha-methyl-L-tryptophan (three-letter code: ZIQ) (formula: $C_{12}H_{14}N_2O_2$) (labeled as "Ligand of Interest" by depositor).
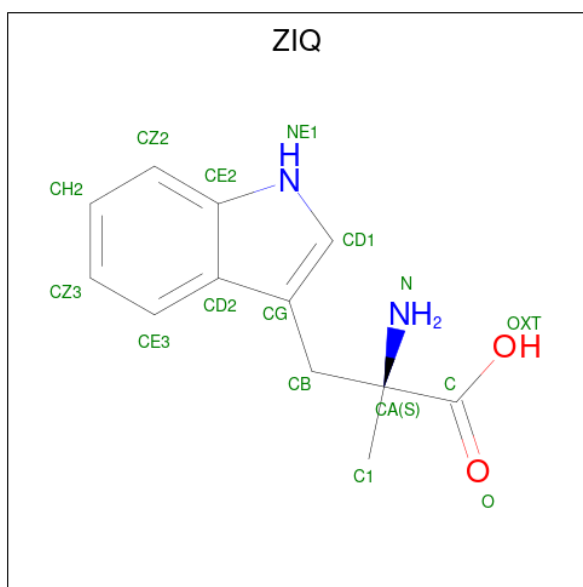
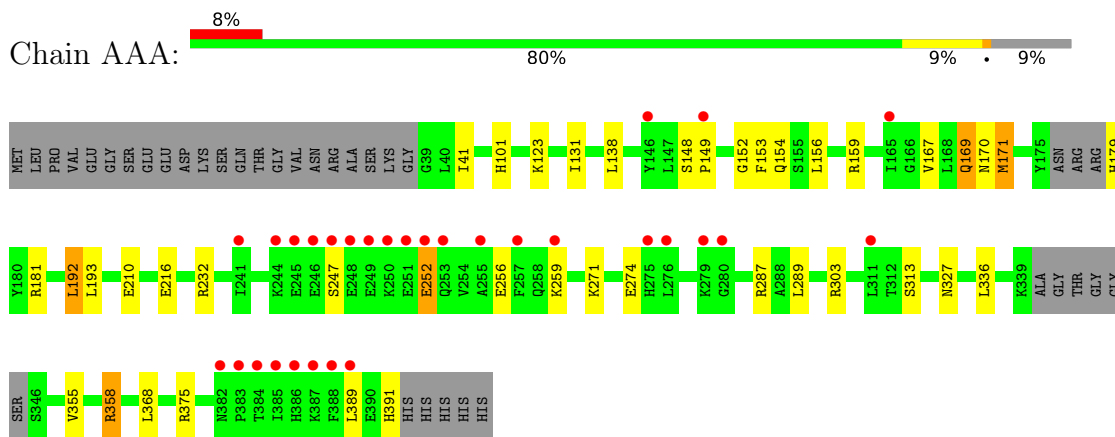| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|---|
| 3 | AAA | 1 | Total | C | N | O | 0 | 0 |
| | | | 16 | 12 | 2 | 2 | | |
| 3 | BBB | 1 | Total | C | N | O | 0 | 0 |
| | | | 16 | 12 | 2 | 2 | | |
| 3 | CCC | 1 | Total | C | N | O | 0 | 0 |
| | | | 16 | 12 | 2 | 2 | | |
| 3 | DDD | 1 | Total | C | N | O | 0 | 0 |
| | | | 16 | 12 | 2 | 2 | | |

- Molecule 4 is 6-fluoro-3-[(E)-2-(1H-tetrazol-5-yl)ethenyl]-1H-indole (three-letter code: A1AER) (formula: $C_{11}H_8FN_5$) (labeled as "Ligand of Interest" by depositor).

| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|---|
| 4 | AAA | 1 | Total 17 | C 11 | F 1 | N 5 | 0 | 0 |
| 4 | BBB | 1 | Total 17 | C 11 | F 1 | N 5 | 0 | 0 |
| 4 | DDD | 1 | Total 17 | C 11 | F 1 | N 5 | 0 | 0 |

• Molecule 5 is water.

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| 5 | AAA | 43 | Total 43 | O 43 | 0 | 0 |
| 5 | BBB | 60 | Total 60 | O 60 | 0 | 0 |
| 5 | CCC | 39 | Total 39 | O 39 | 0 | 0 |
| 5 | DDD | 47 | Total 47 | O 47 | 0 | 0 |

Page 7 wwPDB X-ray Structure Validation Summary Report 8VZV

# 3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.
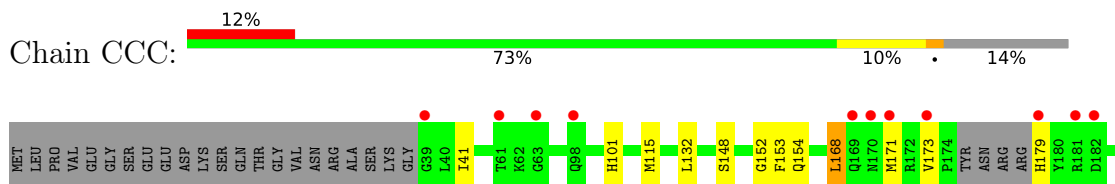
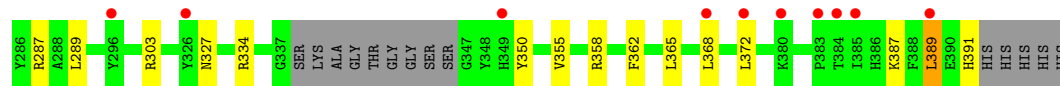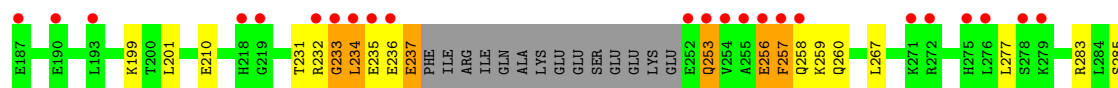- Molecule 1: Tryptophan 2,3-dioxygenase



- Molecule 1: Tryptophan 2,3-dioxygenase



- Molecule 1: Tryptophan 2,3-dioxygenase

E187 E190 L193 K199 T200 L201 E210 H218 G219 T231 R232 G233 L234 E235 E236 E237 PHE ILE ARG ILE GLN ALA LYS GLU GLU SER GLU GLU LYS GLU E252 Q253 V254 A255 E256 F257 Q258 K259 Q260 L267 K271 R272 H275 L276 L277 S278 K279 R283 L284 S285

Y286 R287 A288 L289 Y296 R303 Y326 N327 R334 G337 SER LYS ALA GLY THR GLY GLY SER SER G347 Y348 H349 Y350 V355 R358 F362 L365 L368 L372 K380 P383 T384 I385 H386 K387 F388 L389 E390 H391 HIS HIS HIS HIS

• Molecule 1: Tryptophan 2,3-dioxygenase

Chain DDD:

MET LEU PRO VAL GLU GLY SER GLU GLU ASP LYS SER GLN THR GLY VAL ASN ARG ALA SER LYS GLY G39 L40 I41 T61 H101 M115 D142 S148 P149 G152 F153 Q154 Q169 H179 D182 K185 L191 L192 L193 E210 H218 G219 F238

I241 Q242 A243 LYS GLU E246 S247 E248 E249 K250 E251 E252 Q253 V254 A255 E256 F257 Q258 K259 K271 R272 H275 R287 A288 L289 L293 P302 R303 R334 S338 LYS ALA GLY THR GLY GLY SER S346 R352 V355 L365 L368 P383 T384 I385 H386 K387 F388

L389 E390 HIS HIS HIS HIS HIS HIS

# 4 Data and refinement statistics (i)

| Property | Value | Source |
|---|---|---|
| Space group | P 21 21 2 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 143.43Å    152.92Å    87.59Å<br>90.00°     90.00°     90.00° | Depositor |
| Resolution (Å) | 30.00  –  2.29<br>29.91  –  2.29 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 97.3 (30.00-2.29)<br>97.3 (29.91-2.29) | Depositor<br>EDS |
| $R_{merge}$ | (Not available) | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $< I/\sigma(I) >$ [1] | 1.46 (at 2.29Å) | Xtriage |
| Refinement program | REFMAC 5.8.0267 | Depositor |
| R, $R_{free}$ | 0.202  ,  0.238<br>0.207  ,  0.241 | Depositor<br>DCC |
| $R_{free}$ test set | 4274 reflections (5.03%) | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 45.5 | Xtriage |
| Anisotropy | 0.050 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | 0.33 , 31.4 | EDS |
| L-test for twinning[2] | $< |L| > = 0.48$, $< L^2 > = 0.32$ | Xtriage |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| $F_o,F_c$ correlation | 0.95 | EDS |
| Total number of atoms | 11959 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 60.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.87% of the height of the origin peak. No significant pseudotranslation is detected.*

---

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

# 5   Model quality (i)

## 5.1   Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: HEM, A1AER, ZIQ

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 | AAA | 0.71 | 0/2983 | 0.81 | 1/4013 (0.0%) |
| 1 | BBB | 0.71 | 1/2965 (0.0%) | 0.83 | 2/3988 (0.1%) |
| 1 | CCC | 0.71 | 0/2826 | 0.80 | 0/3804 |
| 1 | DDD | 0.72 | 0/2975 | 0.81 | 1/4003 (0.0%) |
| All | All | 0.71 | 1/11749 (0.0%) | 0.81 | 4/15808 (0.0%) |

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

| Mol | Chain | #Chirality outliers | #Planarity outliers |
|-----|-------|---------------------|---------------------|
| 1 | DDD | 0 | 1 |

All (1) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|-------|---|-------------|----------|
| 1 | BBB | 300 | GLU | CD-OE2 | 5.48 | 1.31 | 1.25 |

All (4) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(°) | Ideal(°) |
|-----|-------|-----|------|-------|---|-------------|----------|
| 1 | BBB | 252 | GLU | CB-CA-C | 5.70 | 121.80 | 110.40 |
| 1 | BBB | 232 | ARG | NE-CZ-NH2 | 5.19 | 122.90 | 120.30 |
| 1 | AAA | 252 | GLU | CB-CA-C | 5.16 | 120.71 | 110.40 |
| 1 | DDD | 390 | GLU | CA-C-O | 5.10 | 130.81 | 120.10 |

There are no chirality outliers.

All (1) planarity outliers are listed below:

| Mol | Chain | Res | Type | Group |
|-----|-------|-----|------|-------|
| 1 | DDD | 242 | GLN | Peptide |

## 5.2   Too-close contacts ⓘ

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

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 1 | AAA | 2915 | 0 | 2897 | 20 | 0 |
| 1 | BBB | 2899 | 0 | 2888 | 16 | 0 |
| 1 | CCC | 2761 | 0 | 2746 | 29 | 0 |
| 1 | DDD | 2908 | 0 | 2890 | 21 | 0 |
| 2 | AAA | 43 | 0 | 30 | 5 | 0 |
| 2 | BBB | 43 | 0 | 30 | 4 | 0 |
| 2 | CCC | 43 | 0 | 30 | 3 | 0 |
| 2 | DDD | 43 | 0 | 30 | 3 | 0 |
| 3 | AAA | 16 | 0 | 0 | 0 | 0 |
| 3 | BBB | 16 | 0 | 0 | 0 | 0 |
| 3 | CCC | 16 | 0 | 0 | 0 | 0 |
| 3 | DDD | 16 | 0 | 0 | 0 | 0 |
| 4 | AAA | 17 | 0 | 0 | 3 | 0 |
| 4 | BBB | 17 | 0 | 0 | 1 | 0 |
| 4 | DDD | 17 | 0 | 0 | 1 | 0 |
| 5 | AAA | 43 | 0 | 0 | 2 | 0 |
| 5 | BBB | 60 | 0 | 0 | 2 | 0 |
| 5 | CCC | 39 | 0 | 0 | 2 | 0 |
| 5 | DDD | 47 | 0 | 0 | 2 | 0 |
| All | All | 11959 | 0 | 11541 | 87 | 0 |

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 4.

The worst 5 of 87 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

| Atom-1 | Atom-2 | Interatomic distance (Å) | Clash overlap (Å) |
|--------|--------|--------------------------|-------------------|
| 1:AAA:391:HIS:C | 1:DDD:390:GLU:C | 2.14 | 1.05 |
| 1:DDD:334:ARG:NH1 | 5:DDD:501:HOH:O | 2.15 | 0.79 |
| 4:BBB:403:A1AER:N10 | 5:BBB:501:HOH:O | 2.20 | 0.73 |
| 1:DDD:246:GLU:HB3 | 1:DDD:250:LYS:HD3 | 1.71 | 0.73 |

*Continued on next page...*

*Continued from previous page...*

| Atom-1 | Atom-2 | Interatomic distance (Å) | Clash overlap (Å) |
|--------|--------|--------------------------|-------------------|
| 1:DDD:169:GLN:HG3 | 1:DDD:179:HIS:NE2 | 2.05 | 0.72 |

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

| Mol | Chain | Analysed | Favoured | Allowed | Outliers | Percentiles | |
|-----|-------|----------|----------|---------|----------|-------------|---|
| 1 | AAA | 338/380 (89%) | 329 (97%) | 9 (3%) | 0 | 100 | 100 |
| 1 | BBB | 338/380 (89%) | 325 (96%) | 12 (4%) | 1 (0%) | 41 | 50 |
| 1 | CCC | 318/380 (84%) | 306 (96%) | 7 (2%) | 5 (2%) | 9 | 9 |
| 1 | DDD | 337/380 (89%) | 328 (97%) | 9 (3%) | 0 | 100 | 100 |
| All | All | 1331/1520 (88%) | 1288 (97%) | 37 (3%) | 6 (0%) | 29 | 35 |

5 of 6 Ramachandran outliers are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | BBB | 340 | ALA |
| 1 | CCC | 257 | PHE |
| 1 | CCC | 233 | GLY |
| 1 | CCC | 234 | LEU |
| 1 | CCC | 236 | GLU |

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

| Mol | Chain | Analysed | Rotameric | Outliers | Percentiles | |
|-----|-------|----------|-----------|----------|-------------|---|
| 1 | AAA | 320/348 (92%) | 297 (93%) | 23 (7%) | 14 | 18 |
| 1 | BBB | 317/348 (91%) | 300 (95%) | 17 (5%) | 22 | 30 |
| 1 | CCC | 303/348 (87%) | 287 (95%) | 16 (5%) | 22 | 31 |
| 1 | DDD | 319/348 (92%) | 299 (94%) | 20 (6%) | 18 | 24 |
| All | All | 1259/1392 (90%) | 1183 (94%) | 76 (6%) | 19 | 26 |

5 of 76 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | DDD | 148 | SER |
| 1 | DDD | 271 | LYS |
| 1 | DDD | 153 | PHE |
| 1 | DDD | 242 | GLN |
| 1 | DDD | 352 | ARG |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3   RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5   Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 5.6   Ligand geometry ⓘ

11 ligands 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 | HEM | AAA | 401 | 5,1 | 41,50,50 | 1.50 | 7 (17%) | 45,82,82 | 2.00 | 16 (35%) |
| 4 | A1AER | BBB | 403 | - | 18,19,19 | 2.35 | 8 (44%) | 19,26,26 | 3.07 | 8 (42%) |
| 3 | ZIQ | BBB | 402 | - | 13,17,17 | 2.78 | 6 (46%) | 14,25,25 | 5.36 | 7 (50%) |
| 2 | HEM | CCC | 401 | 5,1 | 41,50,50 | 1.46 | 7 (17%) | 45,82,82 | 1.85 | 11 (24%) |
| 3 | ZIQ | DDD | 402 | - | 13,17,17 | 2.89 | 5 (38%) | 14,25,25 | 5.82 | 8 (57%) |
| 4 | A1AER | DDD | 403 | - | 18,19,19 | 2.08 | 6 (33%) | 19,26,26 | 3.19 | 6 (31%) |
| 3 | ZIQ | AAA | 402 | - | 13,17,17 | 3.11 | 2 (15%) | 14,25,25 | 5.85 | 8 (57%) |
| 2 | HEM | DDD | 401 | 5,1 | 41,50,50 | 1.35 | 7 (17%) | 45,82,82 | 1.88 | 15 (33%) |
| 3 | ZIQ | CCC | 402 | - | 13,17,17 | 2.99 | 4 (30%) | 14,25,25 | 5.60 | 6 (42%) |
| 4 | A1AER | AAA | 403 | - | 18,19,19 | 2.41 | 5 (27%) | 19,26,26 | 3.34 | 7 (36%) |
| 2 | HEM | BBB | 401 | 5,1 | 41,50,50 | 1.50 | 8 (19%) | 45,82,82 | 2.08 | 17 (37%) |

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 | HEM | AAA | 401 | 5,1 | - | 3/12/54/54 | - |
| 4 | A1AER | BBB | 403 | - | - | 0/3/5/5 | 0/3/3/3 |
| 3 | ZIQ | BBB | 402 | - | - | 2/9/11/11 | 0/2/2/2 |
| 2 | HEM | CCC | 401 | 5,1 | - | 2/12/54/54 | - |
| 3 | ZIQ | DDD | 402 | - | - | 4/9/11/11 | 0/2/2/2 |
| 4 | A1AER | DDD | 403 | - | - | 0/3/5/5 | 0/3/3/3 |
| 3 | ZIQ | AAA | 402 | - | - | 4/9/11/11 | 0/2/2/2 |
| 2 | HEM | DDD | 401 | 5,1 | - | 2/12/54/54 | - |
| 3 | ZIQ | CCC | 402 | - | - | 0/9/11/11 | 0/2/2/2 |
| 4 | A1AER | AAA | 403 | - | - | 0/3/5/5 | 0/3/3/3 |
| 2 | HEM | BBB | 401 | 5,1 | - | 2/12/54/54 | - |

The worst 5 of 65 bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|-------|------|------------|----------|
| 3 | AAA | 402 | ZIQ | CE3-CD2 | -9.05 | 1.23 | 1.42 |
| 3 | CCC | 402 | ZIQ | CE3-CD2 | -8.15 | 1.25 | 1.42 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|-------|---|-------------|----------|
| 3 | DDD | 402 | ZIQ | CE3-CD2 | -7.39 | 1.27 | 1.42 |
| 3 | BBB | 402 | ZIQ | CE3-CD2 | -7.30 | 1.27 | 1.42 |
| 4 | AAA | 403 | A1AER | N18-N17 | 5.49 | 1.42 | 1.32 |

The worst 5 of 109 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|-----|-------|-----|------|-------|---|----------------|-------------|
| 3 | AAA | 402 | ZIQ | CZ3-CE3-CD2 | 17.51 | 145.18 | 120.89 |
| 3 | DDD | 402 | ZIQ | CZ3-CE3-CD2 | 17.34 | 144.94 | 120.89 |
| 3 | CCC | 402 | ZIQ | CZ3-CE3-CD2 | 17.02 | 144.50 | 120.89 |
| 3 | BBB | 402 | ZIQ | CZ3-CE3-CD2 | 16.07 | 143.18 | 120.89 |
| 3 | AAA | 402 | ZIQ | CH2-CZ3-CE3 | -9.19 | 107.55 | 120.44 |

There are no chirality outliers.

5 of 19 torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|-----|-------|-----|------|-------|
| 3 | AAA | 402 | ZIQ | OXT-C-CA-C1 |
| 3 | DDD | 402 | ZIQ | O-C-CA-C1 |
| 3 | DDD | 402 | ZIQ | OXT-C-CA-C1 |
| 3 | AAA | 402 | ZIQ | O-C-CA-C1 |
| 3 | DDD | 402 | ZIQ | O-C-CA-CB |

There are no ring outliers.

7 monomers are involved in 20 short contacts:

| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|-----|-------|-----|------|---------|--------------|
| 2 | AAA | 401 | HEM | 5 | 0 |
| 4 | BBB | 403 | A1AER | 1 | 0 |
| 2 | CCC | 401 | HEM | 3 | 0 |
| 4 | DDD | 403 | A1AER | 1 | 0 |
| 2 | DDD | 401 | HEM | 3 | 0 |
| 4 | AAA | 403 | A1AER | 3 | 0 |
| 2 | BBB | 401 | HEM | 4 | 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 HEM AAA 401**

Bond lengths

Bond angles

Torsions

Rings

## Ligand A1AER BBB 403



Bond lengths



Bond angles



Torsions



Rings

## Ligand ZIQ BBB 402



Bond lengths



Bond angles



Torsions



Rings

## Ligand HEM CCC 401


Bond lengths


Bond angles


Torsions


Rings

## Ligand ZIQ DDD 402



Bond lengths



Bond angles



Torsions



Rings

## Ligand A1AER DDD 403



Bond lengths



Bond angles



Torsions



Rings

## Ligand ZIQ AAA 402


Bond lengths


Bond angles


Torsions


Rings

## Ligand HEM DDD 401


Bond lengths


Bond angles


Torsions


Rings

**Ligand ZIQ CCC 402**



Bond lengths

Bond angles

Torsions

Rings

**Ligand A1AER AAA 403**



Bond lengths

Bond angles

Torsions

Rings

**Ligand HEM BBB 401**



Bond lengths



Bond angles



Torsions



Rings

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

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 | AAA | 344/380 (90%) | 0.36 | 30 (8%) | 10 | 14 | 30, 53, 117, 141 | 0 |
| 1 | BBB | 344/380 (90%) | 0.16 | 28 (8%) | 12 | 16 | 29, 46, 112, 136 | 0 |
| 1 | CCC | 326/380 (85%) | 0.45 | 44 (13%) | 3 | 4 | 31, 61, 118, 152 | 0 |
| 1 | DDD | 343/380 (90%) | 0.22 | 28 (8%) | 11 | 15 | 31, 53, 104, 147 | 0 |
| All | All | 1357/1520 (89%) | 0.29 | 130 (9%) | 8 | 10 | 29, 53, 115, 152 | 0 |

The worst 5 of 130 RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|-----|-------|-----|------|------|
| 1 | AAA | 385 | ILE | 5.8 |
| 1 | BBB | 384 | THR | 5.7 |
| 1 | DDD | 248 | GLU | 5.6 |
| 1 | CCC | 385 | ILE | 5.4 |
| 1 | CCC | 252 | GLU | 5.2 |

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

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

## 6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 6.4 Ligands ⓘ

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

WORLDWIDE
wwPDB
PROTEIN DATA BANK

| Mol | Type | Chain | Res | Atoms | RSCC | RSR | B-factors($Å^2$) | Q<0.9 |
|---|---|---|---|---|---|---|---|---|
| 4 | A1AER | AAA | 403 | 17/17 | 0.81 | 0.41 | 94,103,122,123 | 0 |
| 4 | A1AER | BBB | 403 | 17/17 | 0.84 | 0.25 | 79,90,123,124 | 0 |
| 4 | A1AER | DDD | 403 | 17/17 | 0.92 | 0.22 | 80,90,100,117 | 0 |
| 3 | ZIQ | CCC | 402 | 16/16 | 0.96 | 0.10 | 45,51,62,68 | 0 |
| 2 | HEM | CCC | 401 | 43/43 | 0.96 | 0.12 | 47,61,71,85 | 0 |
| 3 | ZIQ | DDD | 402 | 16/16 | 0.97 | 0.10 | 42,47,54,57 | 0 |
| 3 | ZIQ | AAA | 402 | 16/16 | 0.97 | 0.13 | 32,37,53,64 | 0 |
| 3 | ZIQ | BBB | 402 | 16/16 | 0.97 | 0.08 | 33,36,38,41 | 0 |
| 2 | HEM | AAA | 401 | 43/43 | 0.97 | 0.13 | 40,51,64,74 | 0 |
| 2 | HEM | DDD | 401 | 43/43 | 0.98 | 0.11 | 36,44,66,81 | 0 |
| 2 | HEM | BBB | 401 | 43/43 | 0.98 | 0.10 | 37,45,61,63 | 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.
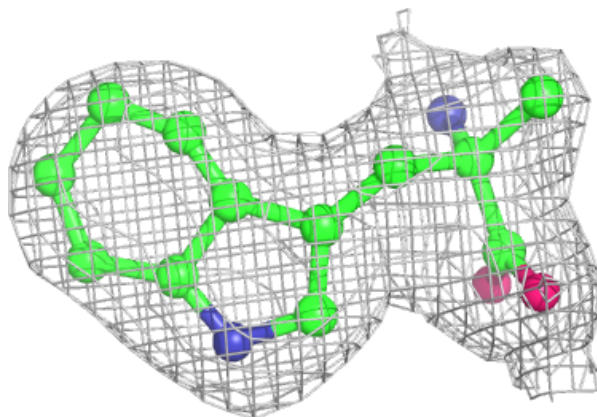
**Electron density around A1AER AAA 403:**

$2mF_o$-$DF_c$ (at 0.7 rmsd) in gray
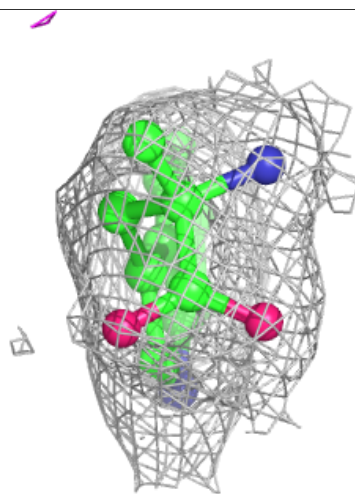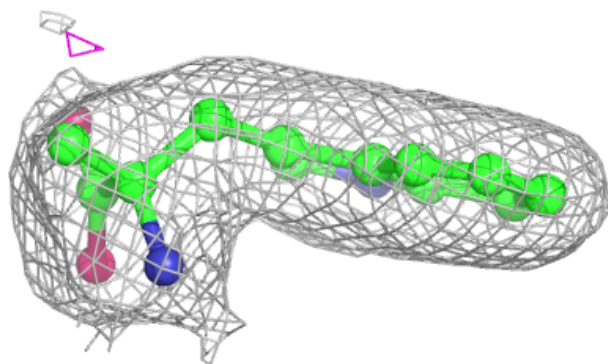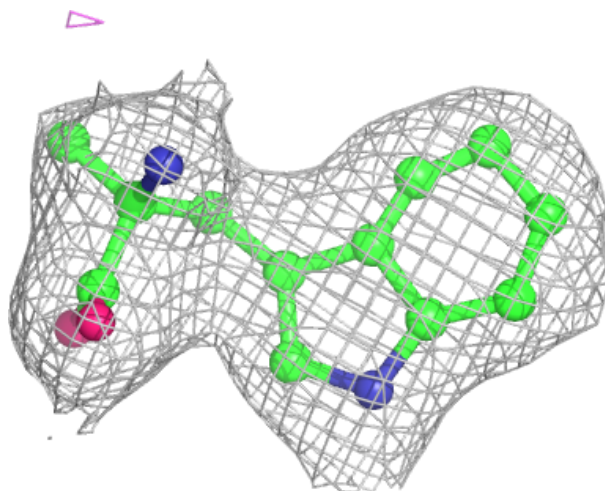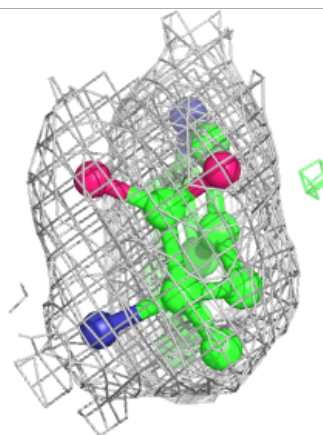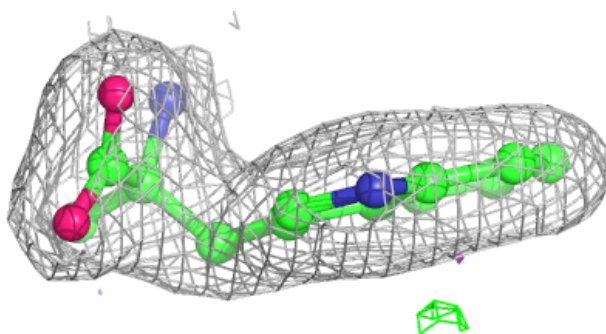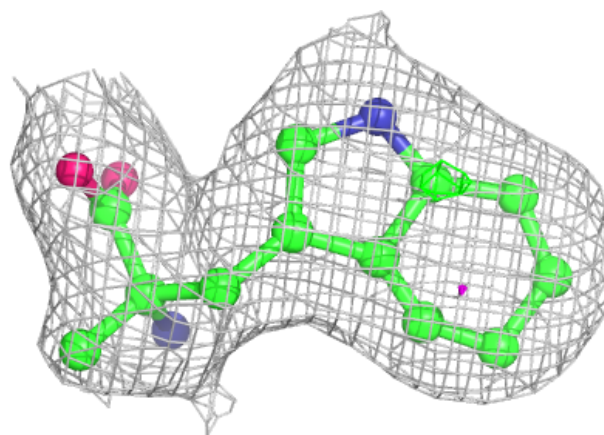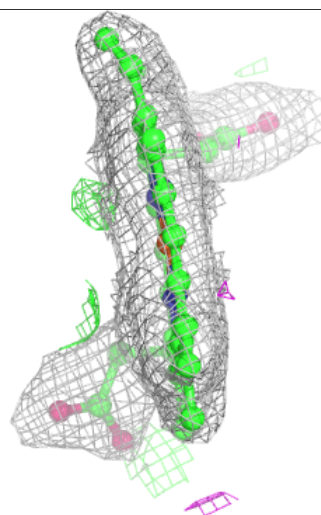$mF_o$-$DF_c$ (at 3 rmsd) in purple (negative) and green (positive)
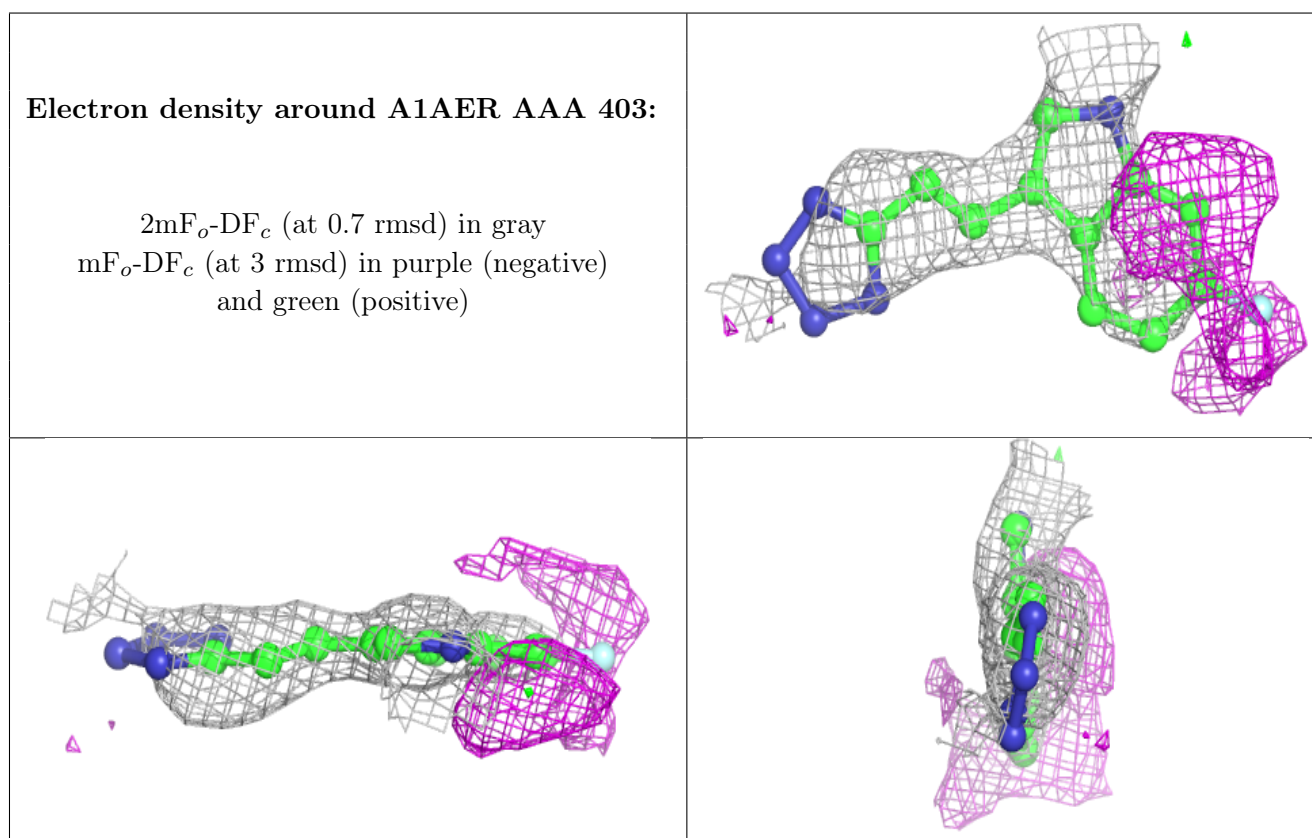
**Electron density around A1AER BBB 403:**

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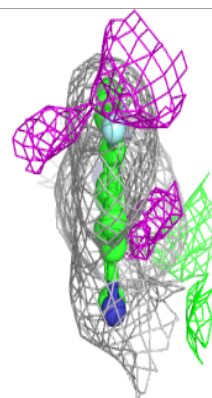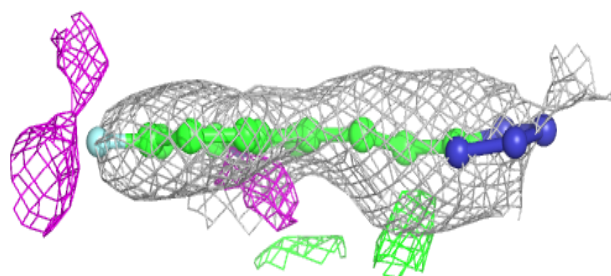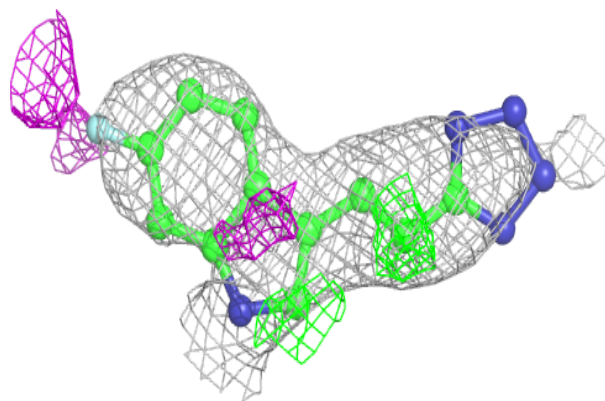


**Electron density around A1AER DDD 403:**

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 ZIQ CCC 402:**

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
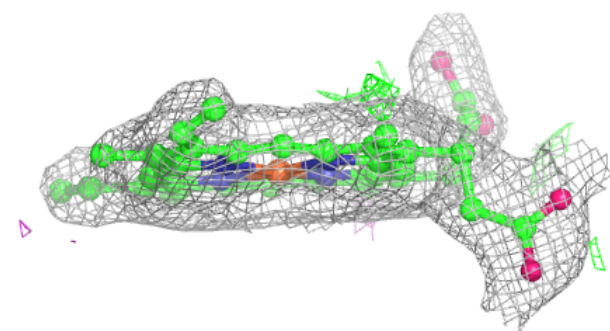$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
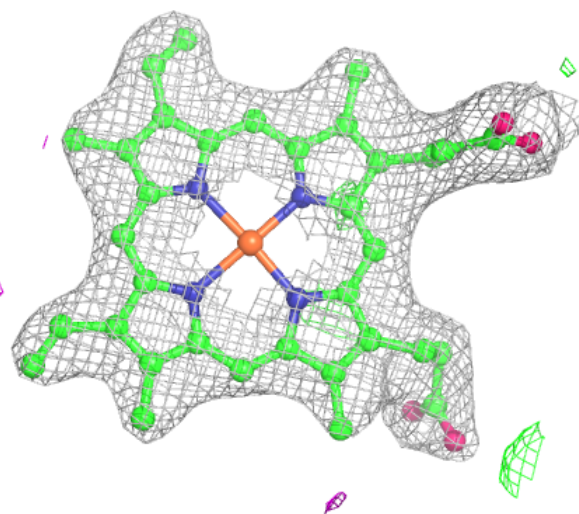and green (positive)
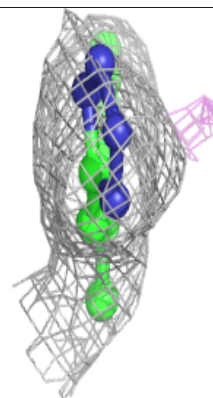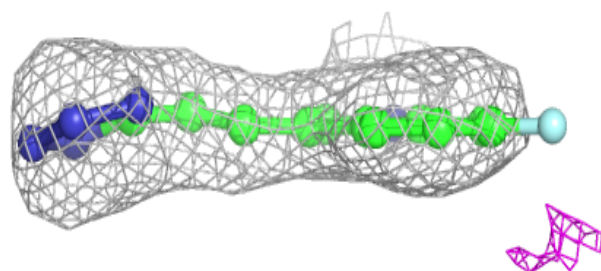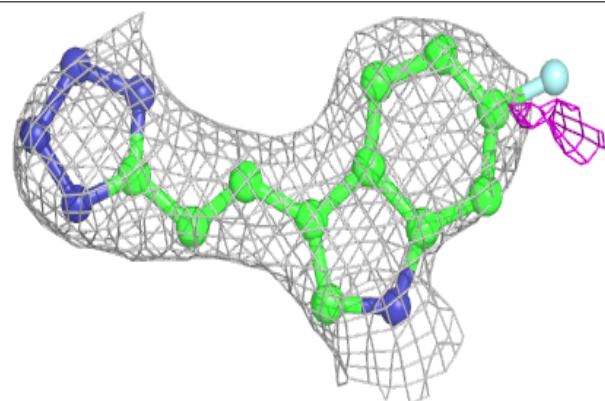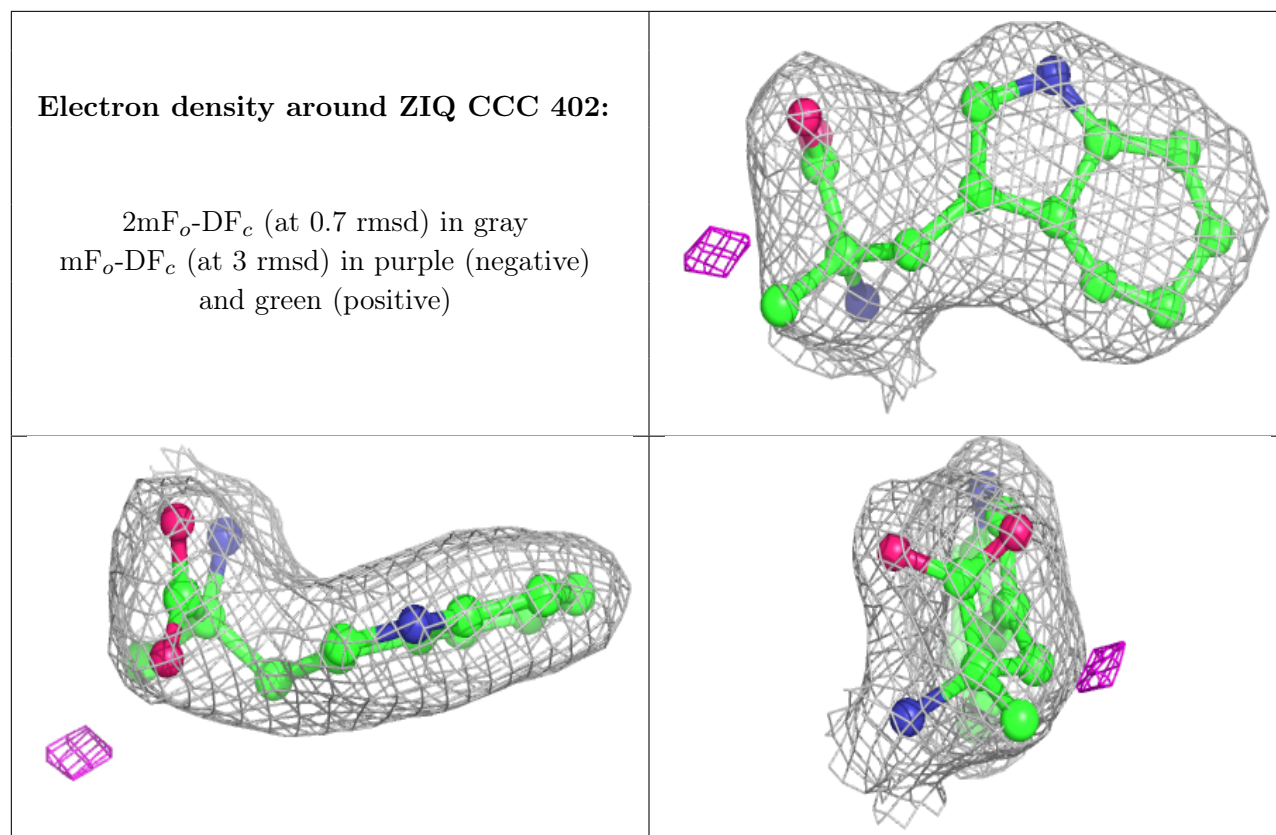
**Electron density around HEM CCC 401:**

$2mF_o$-$DF_c$ (at 0.7 rmsd) in gray
$mF_o$-$DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
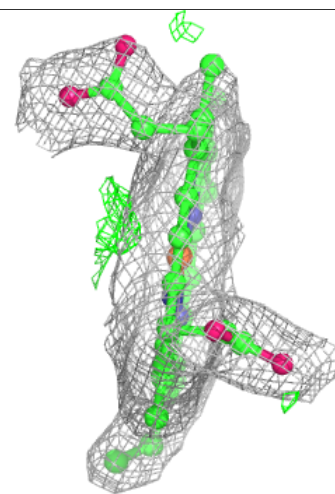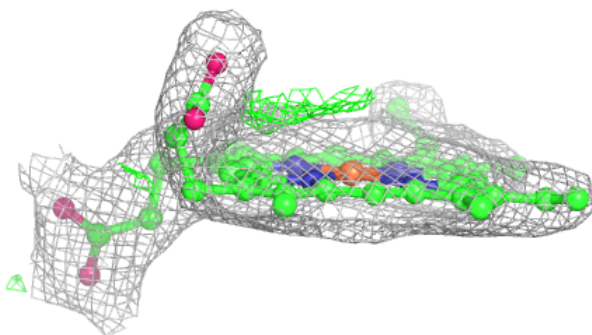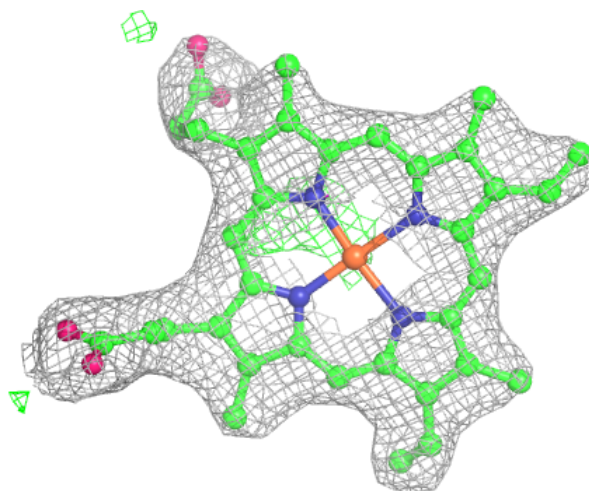
**Electron density around ZIQ DDD 402:**

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
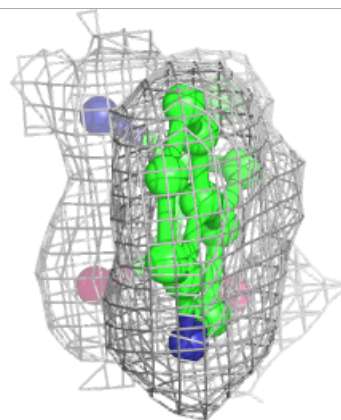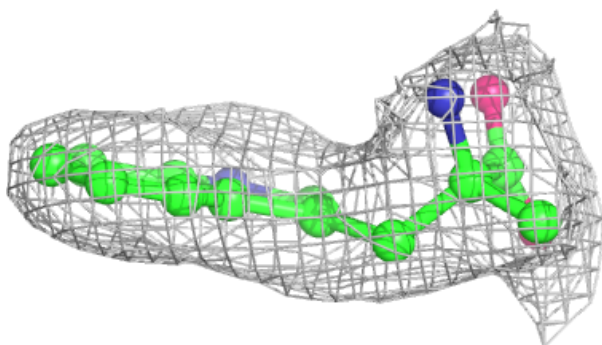
**Electron density around ZIQ AAA 402:**

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

**Electron density around ZIQ BBB 402:**

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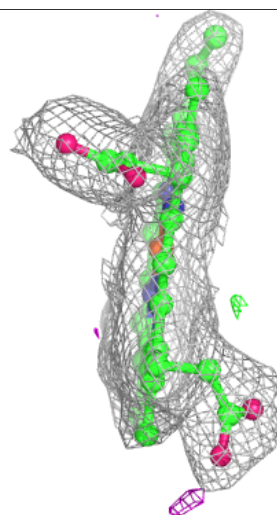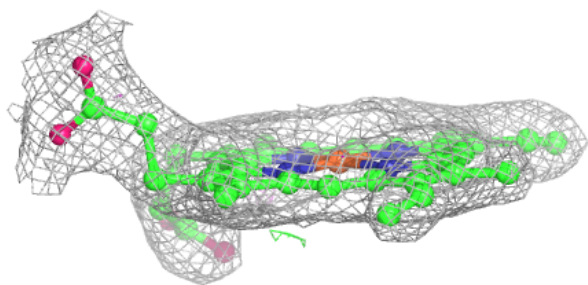$mF_o$-$DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)

**Electron density around HEM AAA 401:**

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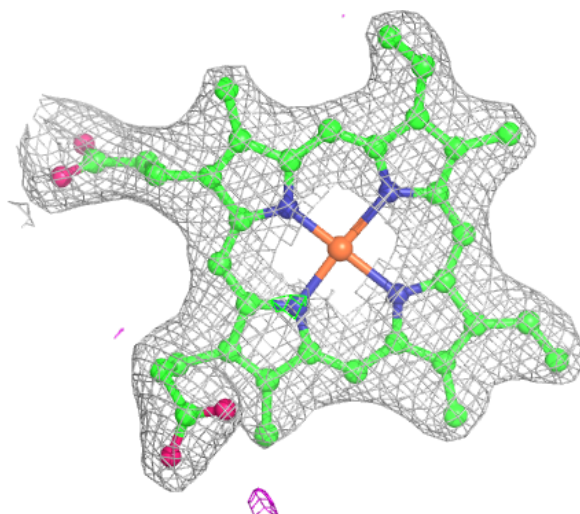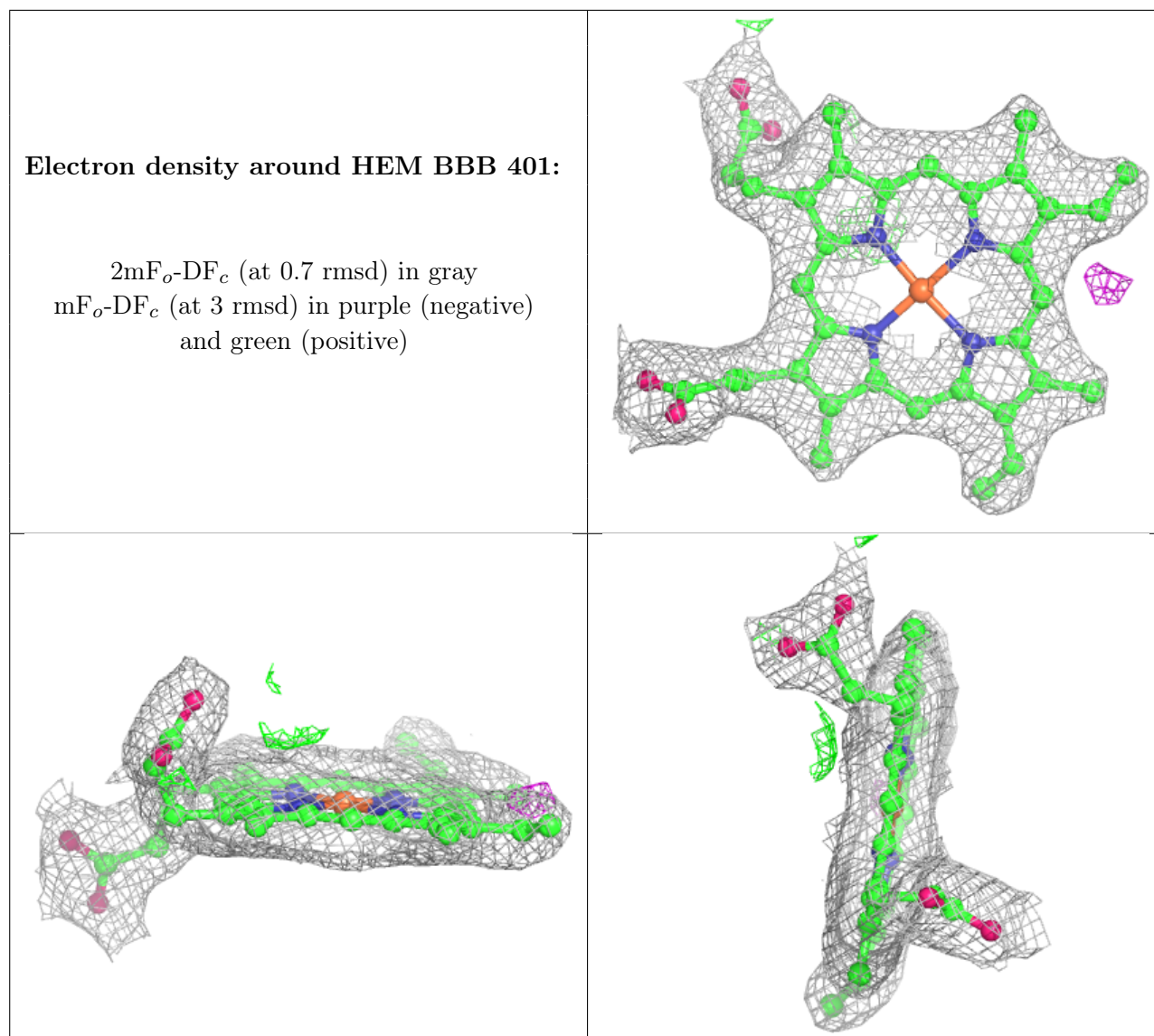and green (positive)

**Electron density around HEM DDD 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 HEM BBB 401:**

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.