# WORLDWIDE wwPDB PROTEIN DATA BANK

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

Sep 8, 2022 – 04:14 pm BST

| | | |
|---|---|---|
| PDB ID | : | 7ZM0 |
| Title | : | Structure of UCHL1 in complex with GK13S inhibitor |
| Authors | : | Grethe, C.; Gersch, M. |
| Deposited on | : | 2022-04-18 |
| Resolution | : | 2.24 Å(reported) |

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

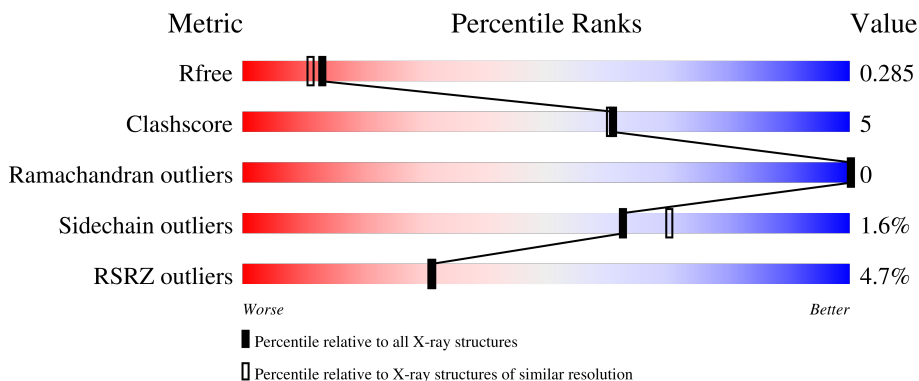| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Mogul | : | 1.8.4, CSD as541be (2020) |
| Xtriage (Phenix) | : | 1.13 |
| EDS | : | 2.30 |
| buster-report | : | 1.1.7 (2018) |
| Percentile statistics | : | 20191225.v01 (using entries in the PDB archive December 25th 2019) |
| Refmac | : | 5.8.0267 |
| CCP4 | : | 7.1.010 (Gargrove) |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | 2.30 |

# 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 2.24 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.
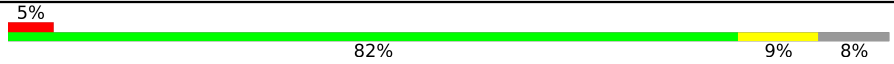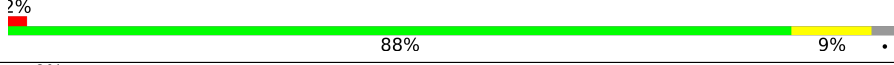
| Metric | Percentile Ranks | Value |
|--------|------------------|-------|
| Rfree | | 0.285 |
| Clashscore | | 5 |
| Ramachandran outliers | | 0 |
| Sidechain outliers | | 1.6% |
| RSRZ outliers | | 4.7% |

*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 | 2391 (2.26-2.22) |
| Clashscore | 141614 | 2539 (2.26-2.22) |
| Ramachandran outliers | 138981 | 2489 (2.26-2.22) |
| Sidechain outliers | 138945 | 2490 (2.26-2.22) |
| RSRZ outliers | 127900 | 2353 (2.26-2.22) |

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 | 227 | 84%    10%   6% |
| 1 | B | 227 | 4%    86%    11%   • |
| 1 | C | 227 | 7%    76%    14%   10% |
| 1 | D | 227 | 3%    82%    13%   • |
| 1 | E | 227 | %    87%    9%   • |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|------------------|
| 1 | F | 227 | 5% 82% 9% 8% |
| 1 | G | 227 | 3% 84% 10% 7% |
| 1 | H | 227 | 2% 88% 9% . |
| 1 | I | 227 | 9% 73% 17% 9% |
| 1 | J | 227 | 10% 78% 10% 12% |

# 2   Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 16195 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 Ubiquitin carboxyl-terminal hydrolase isozyme L1.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|---|---|---|---|---------|---------|-------|
| 1 | A | 213 | Total | C | N | O | S | 0 | 0 | 0 |
|   |   |   | 1615 | 1023 | 280 | 301 | 11 |   |   |   |
| 1 | B | 219 | Total | C | N | O | S | 0 | 0 | 0 |
|   |   |   | 1599 | 1012 | 279 | 297 | 11 |   |   |   |
| 1 | C | 204 | Total | C | N | O | S | 0 | 0 | 0 |
|   |   |   | 1487 | 947 | 257 | 272 | 11 |   |   |   |
| 1 | D | 217 | Total | C | N | O | S | 0 | 0 | 0 |
|   |   |   | 1607 | 1020 | 278 | 298 | 11 |   |   |   |
| 1 | E | 218 | Total | C | N | O | S | 0 | 0 | 0 |
|   |   |   | 1603 | 1013 | 280 | 299 | 11 |   |   |   |
| 1 | F | 208 | Total | C | N | O | S | 0 | 0 | 0 |
|   |   |   | 1529 | 976 | 260 | 282 | 11 |   |   |   |
| 1 | G | 212 | Total | C | N | O | S | 0 | 0 | 0 |
|   |   |   | 1556 | 985 | 269 | 291 | 11 |   |   |   |
| 1 | H | 220 | Total | C | N | O | S | 0 | 0 | 0 |
|   |   |   | 1656 | 1048 | 291 | 305 | 12 |   |   |   |
| 1 | I | 206 | Total | C | N | O | S | 0 | 0 | 0 |
|   |   |   | 1466 | 929 | 251 | 275 | 11 |   |   |   |
| 1 | J | 200 | Total | C | N | O | S | 0 | 0 | 0 |
|   |   |   | 1456 | 922 | 251 | 273 | 10 |   |   |   |

There are 40 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------|-----------|
| A | -3 | GLY | - | expression tag | UNP P09936 |
| A | -2 | PRO | - | expression tag | UNP P09936 |
| A | -1 | GLY | - | expression tag | UNP P09936 |
| A | 0 | SER | - | expression tag | UNP P09936 |
| B | -3 | GLY | - | expression tag | UNP P09936 |
| B | -2 | PRO | - | expression tag | UNP P09936 |
| B | -1 | GLY | - | expression tag | UNP P09936 |
| B | 0 | SER | - | expression tag | UNP P09936 |
| C | -3 | GLY | - | expression tag | UNP P09936 |

*Continued on next page...*

*Continued from previous page...*

| Chain | Residue | Modelled | Actual | Comment | Reference |
|:---:|:---:|:---:|:---:|:---:|:---:|
| C | -2 | PRO | - | expression tag | UNP P09936 |
| C | -1 | GLY | - | expression tag | UNP P09936 |
| C | 0 | SER | - | expression tag | UNP P09936 |
| D | -3 | GLY | - | expression tag | UNP P09936 |
| D | -2 | PRO | - | expression tag | UNP P09936 |
| D | -1 | GLY | - | expression tag | UNP P09936 |
| D | 0 | SER | - | expression tag | UNP P09936 |
| E | -3 | GLY | - | expression tag | UNP P09936 |
| E | -2 | PRO | - | expression tag | UNP P09936 |
| E | -1 | GLY | - | expression tag | UNP P09936 |
| E | 0 | SER | - | expression tag | UNP P09936 |
| F | -3 | GLY | - | expression tag | UNP P09936 |
| F | -2 | PRO | - | expression tag | UNP P09936 |
| F | -1 | GLY | - | expression tag | UNP P09936 |
| F | 0 | SER | - | expression tag | UNP P09936 |
| G | -3 | GLY | - | expression tag | UNP P09936 |
| G | -2 | PRO | - | expression tag | UNP P09936 |
| G | -1 | GLY | - | expression tag | UNP P09936 |
| G | 0 | SER | - | expression tag | UNP P09936 |
| H | -3 | GLY | - | expression tag | UNP P09936 |
| H | -2 | PRO | - | expression tag | UNP P09936 |
| H | -1 | GLY | - | expression tag | UNP P09936 |
| H | 0 | SER | - | expression tag | UNP P09936 |
| I | -3 | GLY | - | expression tag | UNP P09936 |
| I | -2 | PRO | - | expression tag | UNP P09936 |
| I | -1 | GLY | - | expression tag | UNP P09936 |
| I | 0 | SER | - | expression tag | UNP P09936 |
| J | -3 | GLY | - | expression tag | UNP P09936 |
| J | -2 | PRO | - | expression tag | UNP P09936 |
| J | -1 | GLY | - | expression tag | UNP P09936 |
| J | 0 | SER | - | expression tag | UNP P09936 |

- Molecule 2 is (3S)-1-(iminomethyl)-N-[1-[4-(pent-4-ynylcarbamoyl)phenyl]imidazol-4-yl]pyrrolidine-3-carboxamide (three-letter code: JMF) (formula: $C_{21}H_{24}N_6O_2$).

JMF
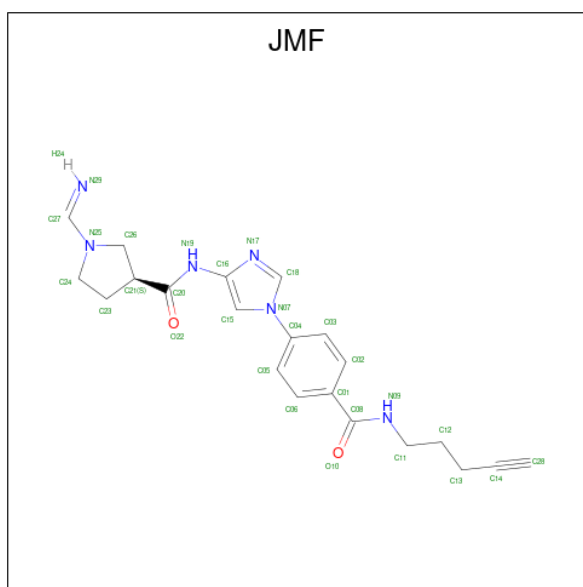


| Mol | Chain | Residues | Atoms | | | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|---|---|
| 2 | A | 1 | Total<br>29 | C<br>21 | N<br>6 | O<br>2 | 0 | 0 |
| 2 | B | 1 | Total<br>29 | C<br>21 | N<br>6 | O<br>2 | 0 | 0 |
| 2 | C | 1 | Total<br>29 | C<br>21 | N<br>6 | O<br>2 | 0 | 0 |
| 2 | D | 1 | Total<br>29 | C<br>21 | N<br>6 | O<br>2 | 0 | 0 |
| 2 | E | 1 | Total<br>29 | C<br>21 | N<br>6 | O<br>2 | 0 | 0 |
| 2 | F | 1 | Total<br>29 | C<br>21 | N<br>6 | O<br>2 | 0 | 0 |
| 2 | G | 1 | Total<br>29 | C<br>21 | N<br>6 | O<br>2 | 0 | 0 |
| 2 | H | 1 | Total<br>29 | C<br>21 | N<br>6 | O<br>2 | 0 | 0 |
| 2 | I | 1 | Total<br>29 | C<br>21 | N<br>6 | O<br>2 | 0 | 0 |
| 2 | J | 1 | Total<br>29 | C<br>21 | N<br>6 | O<br>2 | 0 | 0 |

- Molecule 3 is water.

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|---|---|---|---|---|---|---|
| 3 | A | 47 | Total<br>47 | O<br>47 | 0 | 0 |
| 3 | B | 24 | Total<br>24 | O<br>24 | 0 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Residues | Atoms | | ZeroOcc | AltConf |
|-----|-------|----------|-------|---|---------|---------|
| 3 | C | 24 | Total 24 | O 24 | 0 | 0 |
| 3 | D | 30 | Total 30 | O 30 | 0 | 0 |
| 3 | E | 23 | Total 23 | O 23 | 0 | 0 |
| 3 | F | 38 | Total 38 | O 38 | 0 | 0 |
| 3 | G | 32 | Total 32 | O 32 | 0 | 0 |
| 3 | H | 54 | Total 54 | O 54 | 0 | 0 |
| 3 | I | 30 | Total 30 | O 30 | 0 | 0 |
| 3 | J | 29 | Total 29 | O 29 | 0 | 0 |

# 3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.
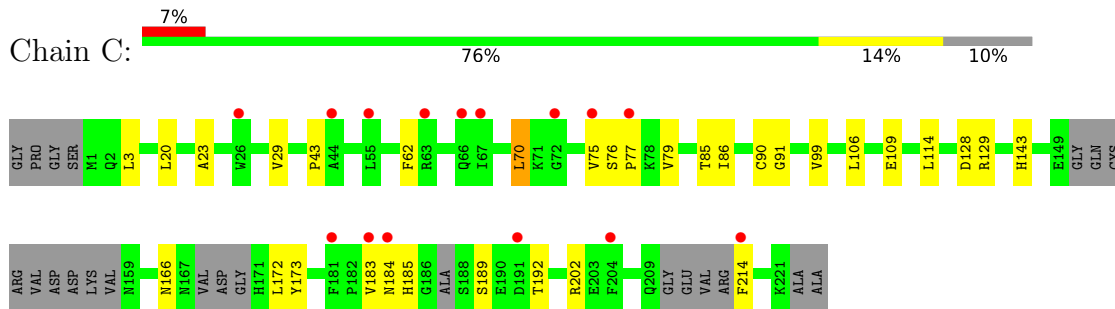
- Molecule 1: Ubiquitin carboxyl-terminal hydrolase isozyme L1
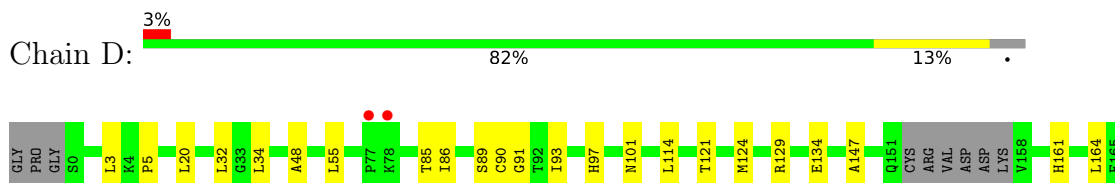
Chain A:

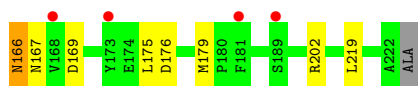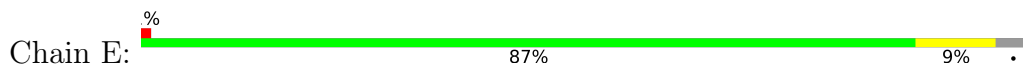- Molecule 1: Ubiquitin carboxyl-terminal hydrolase isozyme L1

Chain B:
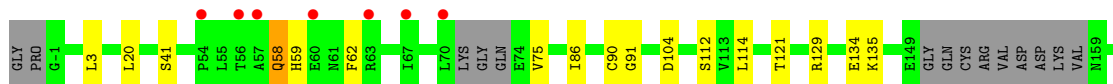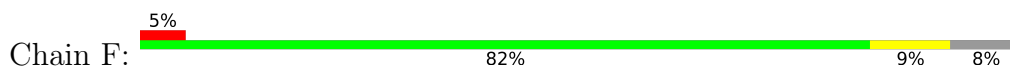
- Molecule 1: Ubiquitin carboxyl-terminal hydrolase isozyme L1

Chain C:

- Molecule 1: Ubiquitin carboxyl-terminal hydrolase isozyme L1

Chain D:

N166 N167 V168 D169 Y173 E174 L175 D176 M179 P180 F181 S189 R202 L219 A222 ALA

• Molecule 1: Ubiquitin carboxyl-terminal hydrolase isozyme L1

Chain E:

GLY PRO GLY S0 K4 A23 G24 R27 L39 G40 S41 E74 K83 Q84 T85 C90 G91 H97 E109 E127 Q148 Q151 CYS ARG VAL ASP LYS V158 L164 V168 H171 L175 A187 L194 R213 A223
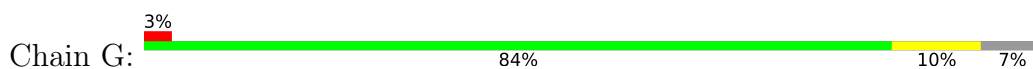
• Molecule 1: Ubiquitin carboxyl-terminal hydrolase isozyme L1

Chain F:

GLY PRO G-1 L3 L20 S41 P54 L55 T56 A57 Q58 H59 E60 N61 F62 R63 I67 L70 LYS GLY GLN E74 V75 I86 C90 G91 D104 S112 V113 L114 T121 R129 E134 K135 E149 GLY GLN CYS ARG VAL ASP LYS VAL N159

V168 Y173 D176 M179 P180 F181 G186 F204 E208 GLN GLY GLU VAL R213 A222 ALA
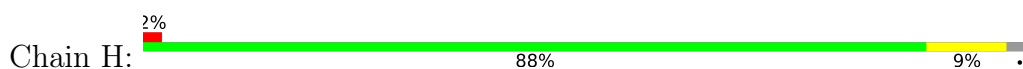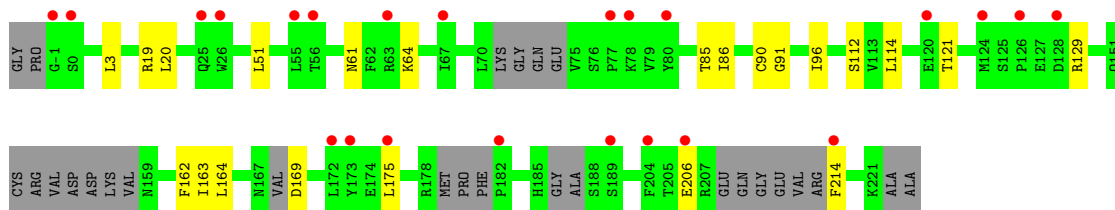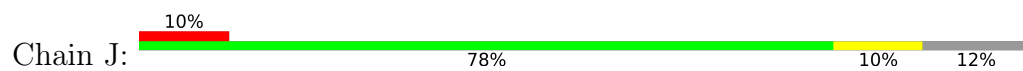
• Molecule 1: Ubiquitin carboxyl-terminal hydrolase isozyme L1

Chain G:

GLY PRO G-1 S0 M1 L20 S41 P45 A48 L55 Q73 E74 Y80 K83 C90 G91 T92 L114 T121 E122 K123 M124 S125 P126 R129 H143 G150 GLN CYS ARG VAL ASP ASP LYS VAL N159 L164 V168 L175 F181 A198

R202 Q209 GLY GLU VAL ARG F214 A222 ALA

• Molecule 1: Ubiquitin carboxyl-terminal hydrolase isozyme L1

Chain H:

GLY PRO G-1 L3 G24 Q25 W26 R27 E35 E36 E37 S41 G72 P77 I86 C90 G91 K105 T121 R129 Q151 C152 R153 VAL ASP ASP LYS V158 D169 E174 R178 V183 E203 R207 A222 ALA

• Molecule 1: Ubiquitin carboxyl-terminal hydrolase isozyme L1

Chain I:

GLY PRO GLY S0 M1 N9 M12 R19 L20 A23 S41 A46 C47 A48 L52 L55 E60 R63 I67 L70 LYS GLY GLN V75 S76 P77 K78 V79 Y80 T85 I86 C90 G91 T92 I96 Q103 D104 K105 L106 E109 S112

V113 L114 E120 T121 E122 K123 D128 E134 D144 A145 V146 A147 G150 GLN CYS ARG VAL ASP ASP LYS VAL N159 L164 F165 N166 N167 V168 ASP GLY HIS L172 Y173 E174 L175 P180 PHE P182 V183 ASN H185 G186 A197 T205 E206 R207 V212 R213 F214 S215

L219 C220 K221 A222 ALA

• Molecule 1: Ubiquitin carboxyl-terminal hydrolase isozyme L1

Chain J:

10%

78%

10%

12%

GLY PRO G-1 S0 L3 R19 L20 Q25 W26 L51 L55 T56 N61 F62 R63 K64 I67 L70 LYS GLY GLU V75 S76 P77 K78 V79 Y80 T85 I86 C90 G91 I96 S112 V113 L114 E120 T121 M124 S125 P126 E127 D128 R129 Q151

CYS ARG VAL ASP ASP LYS VAL N159 F162 I163 L164 N167 VAL D169 L172 Y173 E174 L175 R178 MET PRO PHE P182 H185 GLY ALA S188 S189 F204 T205 E206 R207 GLU GLN GLY GLU VAL ARG F214 K221 ALA ALA

# 4 Data and refinement statistics (i)

| Property | Value | Source |
|---|---|---|
| Space group | P 21 21 21 | Depositor |
| Cell constants<br>a, b, c, $\alpha$, $\beta$, $\gamma$ | 101.93Å 144.41Å 158.25Å<br>90.00° 90.00° 90.00° | Depositor |
| Resolution (Å) | 62.50 – 2.24<br>62.50 – 2.24 | Depositor<br>EDS |
| % Data completeness<br>(in resolution range) | 56.3 (62.50-2.24)<br>56.3 (62.50-2.24) | Depositor<br>EDS |
| $R_{merge}$ | 0.05 | Depositor |
| $R_{sym}$ | (Not available) | Depositor |
| $<I/\sigma(I)>$ [1] | 2.23 (at 2.25Å) | Xtriage |
| Refinement program | PHENIX 1.19.2_4158 | Depositor |
| R, $R_{free}$ | 0.240 , 0.288<br>0.240 , 0.285 | Depositor<br>DCC |
| $R_{free}$ test set | 1633 reflections (2.57%) | wwPDB-VP |
| Wilson B-factor (Å$^2$) | 46.3 | Xtriage |
| Anisotropy | 0.066 | Xtriage |
| Bulk solvent $k_{sol}$(e/Å$^3$), $B_{sol}$(Å$^2$) | (Not available) , (Not available) | EDS |
| L-test for twinning[2] | $<|L|> = 0.51$, $<L^2> = 0.34$ | Xtriage |
| Estimated twinning fraction | No twinning to report. | Xtriage |
| $F_o,F_c$ correlation | 0.92 | EDS |
| Total number of atoms | 16195 | wwPDB-VP |
| Average B, all atoms (Å$^2$) | 62.0 | wwPDB-VP |

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.23% of the height of the origin peak. No significant pseudotranslation is detected.*

---

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

# 5 Model quality ⓘ

## 5.1 Standard geometry ⓘ

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

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

| Mol | Chain | Bond lengths | | Bond angles | |
|---|---|---|---|---|---|
| | | RMSZ | #$|Z|$ >5 | RMSZ | #$|Z|$ >5 |
| 1 | A | 0.25 | 0/1645 | 0.42 | 0/2219 |
| 1 | B | 0.26 | 0/1630 | 0.43 | 0/2212 |
| 1 | C | 0.25 | 0/1513 | 0.42 | 0/2051 |
| 1 | D | 0.26 | 0/1637 | 0.44 | 0/2217 |
| 1 | E | 0.25 | 0/1633 | 0.43 | 0/2212 |
| 1 | F | 0.26 | 0/1557 | 0.43 | 0/2109 |
| 1 | G | 0.27 | 0/1584 | 0.43 | 0/2146 |
| 1 | H | 0.25 | 0/1686 | 0.44 | 0/2274 |
| 1 | I | 0.28 | 0/1489 | 0.43 | 0/2016 |
| 1 | J | 0.25 | 0/1479 | 0.42 | 0/2000 |
| All | All | 0.26 | 0/15853 | 0.43 | 0/21456 |

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts ⓘ

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

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|---|---|---|---|---|---|---|
| 1 | A | 1615 | 0 | 1568 | 15 | 0 |
| 1 | B | 1599 | 0 | 1498 | 13 | 0 |
| 1 | C | 1487 | 0 | 1383 | 18 | 0 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 1 | D | 1607 | 0 | 1529 | 18 | 0 |
| 1 | E | 1603 | 0 | 1500 | 13 | 0 |
| 1 | F | 1529 | 0 | 1450 | 12 | 0 |
| 1 | G | 1556 | 0 | 1467 | 14 | 0 |
| 1 | H | 1656 | 0 | 1610 | 11 | 0 |
| 1 | I | 1466 | 0 | 1327 | 22 | 0 |
| 1 | J | 1456 | 0 | 1337 | 12 | 0 |
| 2 | A | 29 | 0 | 0 | 0 | 0 |
| 2 | B | 29 | 0 | 0 | 0 | 0 |
| 2 | C | 29 | 0 | 0 | 0 | 0 |
| 2 | D | 29 | 0 | 0 | 0 | 0 |
| 2 | E | 29 | 0 | 0 | 0 | 0 |
| 2 | F | 29 | 0 | 0 | 0 | 0 |
| 2 | G | 29 | 0 | 0 | 0 | 0 |
| 2 | H | 29 | 0 | 0 | 0 | 0 |
| 2 | I | 29 | 0 | 0 | 0 | 0 |
| 2 | J | 29 | 0 | 0 | 0 | 0 |
| 3 | A | 47 | 0 | 0 | 1 | 0 |
| 3 | B | 24 | 0 | 0 | 0 | 0 |
| 3 | C | 24 | 0 | 0 | 2 | 0 |
| 3 | D | 30 | 0 | 0 | 0 | 0 |
| 3 | E | 23 | 0 | 0 | 1 | 0 |
| 3 | F | 38 | 0 | 0 | 3 | 0 |
| 3 | G | 32 | 0 | 0 | 1 | 0 |
| 3 | H | 54 | 0 | 0 | 1 | 0 |
| 3 | I | 30 | 0 | 0 | 2 | 0 |
| 3 | J | 29 | 0 | 0 | 1 | 0 |
| All | All | 16195 | 0 | 14669 | 138 | 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 138 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:F:208:GLU:CB | 3:F:435:HOH:O | 2.23 | 0.85 |
| 1:J:121:THR:HB | 1:J:129:ARG:HG2 | 1.66 | 0.75 |
| 1:D:121:THR:HB | 1:D:129:ARG:HG2 | 1.75 | 0.69 |
| 1:A:121:THR:HB | 1:A:129:ARG:HG2 | 1.73 | 0.69 |
| 1:B:32:LEU:CB | 3:F:428:HOH:O | 2.44 | 0.65 |

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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

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

| Mol | Chain | Analysed | Favoured | Allowed | Outliers | Percentiles | |
|---|---|---|---|---|---|---|---|
| 1 | A | 207/227 (91%) | 206 (100%) | 1 (0%) | 0 | 100 | 100 |
| 1 | B | 215/227 (95%) | 208 (97%) | 7 (3%) | 0 | 100 | 100 |
| 1 | C | 194/227 (86%) | 189 (97%) | 5 (3%) | 0 | 100 | 100 |
| 1 | D | 213/227 (94%) | 210 (99%) | 3 (1%) | 0 | 100 | 100 |
| 1 | E | 214/227 (94%) | 210 (98%) | 4 (2%) | 0 | 100 | 100 |
| 1 | F | 200/227 (88%) | 198 (99%) | 2 (1%) | 0 | 100 | 100 |
| 1 | G | 206/227 (91%) | 202 (98%) | 4 (2%) | 0 | 100 | 100 |
| 1 | H | 216/227 (95%) | 211 (98%) | 5 (2%) | 0 | 100 | 100 |
| 1 | I | 194/227 (86%) | 192 (99%) | 2 (1%) | 0 | 100 | 100 |
| 1 | J | 186/227 (82%) | 183 (98%) | 3 (2%) | 0 | 100 | 100 |
| All | All | 2045/2270 (90%) | 2009 (98%) | 36 (2%) | 0 | 100 | 100 |

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

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

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

| Mol | Chain | Analysed | Rotameric | Outliers | Percentiles | |
|---|---|---|---|---|---|---|
| 1 | A | 168/190 (88%) | 167 (99%) | 1 (1%) | 86 | 90 |
| 1 | B | 157/190 (83%) | 155 (99%) | 2 (1%) | 69 | 76 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Analysed | Rotameric | Outliers | Percentiles | |
|-----|-------|----------|-----------|----------|-------------|---|
| 1 | C | 144/190 (76%) | 140 (97%) | 4 (3%) | 43 | 49 |
| 1 | D | 160/190 (84%) | 158 (99%) | 2 (1%) | 69 | 76 |
| 1 | E | 156/190 (82%) | 155 (99%) | 1 (1%) | 86 | 90 |
| 1 | F | 152/190 (80%) | 149 (98%) | 3 (2%) | 55 | 62 |
| 1 | G | 154/190 (81%) | 153 (99%) | 1 (1%) | 86 | 90 |
| 1 | H | 169/190 (89%) | 167 (99%) | 2 (1%) | 71 | 78 |
| 1 | I | 136/190 (72%) | 130 (96%) | 6 (4%) | 28 | 30 |
| 1 | J | 141/190 (74%) | 138 (98%) | 3 (2%) | 53 | 60 |
| All | All | 1537/1900 (81%) | 1512 (98%) | 25 (2%) | 62 | 70 |

5 of 25 residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | H | 151 | GLN |
| 1 | I | 103 | GLN |
| 1 | J | 206 | GLU |
| 1 | I | 70 | LEU |
| 1 | I | 105 | LYS |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | G | 148 | GLN |
| 1 | I | 102 | ASN |
| 1 | I | 166 | ASN |
| 1 | J | 84 | GLN |

#### 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 (i)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry (i)

10 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 | JMF | G | 301 | 1 | 28,31,31 | 5.45 | 19 (67%) | 34,41,41 | 2.34 | 6 (17%) |
| 2 | JMF | E | 301 | 1 | 28,31,31 | 5.47 | 19 (67%) | 34,41,41 | 2.05 | 6 (17%) |
| 2 | JMF | A | 301 | 1 | 28,31,31 | 5.53 | 19 (67%) | 34,41,41 | 1.80 | 6 (17%) |
| 2 | JMF | B | 301 | 1 | 28,31,31 | 5.46 | 19 (67%) | 34,41,41 | 1.93 | 5 (14%) |
| 2 | JMF | D | 301 | 1 | 28,31,31 | 5.45 | 19 (67%) | 34,41,41 | 2.20 | 6 (17%) |
| 2 | JMF | H | 301 | 1 | 28,31,31 | 5.45 | 19 (67%) | 34,41,41 | 2.12 | 5 (14%) |
| 2 | JMF | J | 301 | 1 | 28,31,31 | 5.49 | 19 (67%) | 34,41,41 | 2.06 | 5 (14%) |
| 2 | JMF | C | 301 | 1 | 28,31,31 | 5.47 | 19 (67%) | 34,41,41 | 2.12 | 6 (17%) |
| 2 | JMF | F | 301 | 1 | 28,31,31 | 5.49 | 19 (67%) | 34,41,41 | 2.12 | 5 (14%) |
| 2 | JMF | I | 301 | 1 | 28,31,31 | 5.45 | 19 (67%) | 34,41,41 | 2.36 | 6 (17%) |

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 | JMF | G | 301 | 1 | - | 6/19/33/33 | 0/3/3/3 |
| 2 | JMF | E | 301 | 1 | - | 5/19/33/33 | 0/3/3/3 |
| 2 | JMF | A | 301 | 1 | - | 5/19/33/33 | 0/3/3/3 |
| 2 | JMF | B | 301 | 1 | - | 5/19/33/33 | 0/3/3/3 |
| 2 | JMF | D | 301 | 1 | - | 9/19/33/33 | 0/3/3/3 |
| 2 | JMF | H | 301 | 1 | - | 5/19/33/33 | 0/3/3/3 |

*Continued on next page...*

*Continued from previous page...*

| Mol | Type | Chain | Res | Link | Chirals | Torsions | Rings |
|-----|------|-------|-----|------|---------|----------|-------|
| 2 | JMF | J | 301 | 1 | - | 10/19/33/33 | 0/3/3/3 |
| 2 | JMF | C | 301 | 1 | - | 5/19/33/33 | 0/3/3/3 |
| 2 | JMF | F | 301 | 1 | - | 6/19/33/33 | 0/3/3/3 |
| 2 | JMF | I | 301 | 1 | - | 6/19/33/33 | 0/3/3/3 |

The worst 5 of 190 bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|-------|---|-------------|----------|
| 2 | A | 301 | JMF | C06-C05 | 10.52 | 1.57 | 1.38 |
| 2 | J | 301 | JMF | C06-C05 | 10.51 | 1.57 | 1.38 |
| 2 | F | 301 | JMF | C06-C05 | 10.46 | 1.57 | 1.38 |
| 2 | C | 301 | JMF | C06-C05 | 10.46 | 1.57 | 1.38 |
| 2 | H | 301 | JMF | C06-C05 | 10.43 | 1.57 | 1.38 |

The worst 5 of 56 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(°) | Ideal(°) |
|-----|-------|-----|------|-------|---|-------------|----------|
| 2 | I | 301 | JMF | C24-N25-C27 | 9.62 | 130.47 | 123.90 |
| 2 | G | 301 | JMF | C24-N25-C27 | 9.04 | 130.08 | 123.90 |
| 2 | D | 301 | JMF | C24-N25-C27 | 8.31 | 129.57 | 123.90 |
| 2 | C | 301 | JMF | C24-N25-C27 | 7.89 | 129.29 | 123.90 |
| 2 | H | 301 | JMF | C24-N25-C27 | 7.83 | 129.24 | 123.90 |

There are no chirality outliers.

5 of 62 torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|-----|-------|-----|------|-------|
| 2 | D | 301 | JMF | C03-C04-N07-C15 |
| 2 | D | 301 | JMF | C03-C04-N07-C18 |
| 2 | D | 301 | JMF | C05-C04-N07-C15 |
| 2 | D | 301 | JMF | C05-C04-N07-C18 |
| 2 | J | 301 | JMF | C03-C04-N07-C15 |

There are no ring outliers.

No monomer is involved in short contacts.

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.
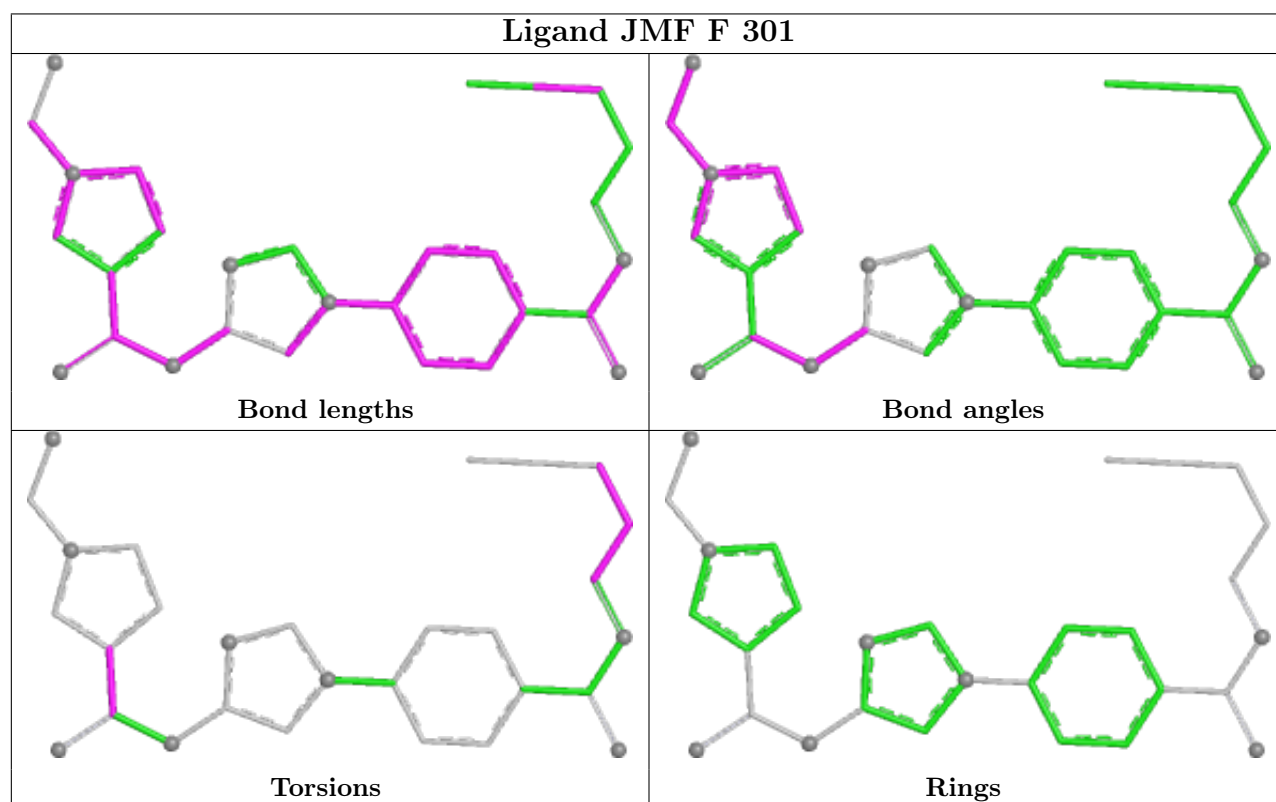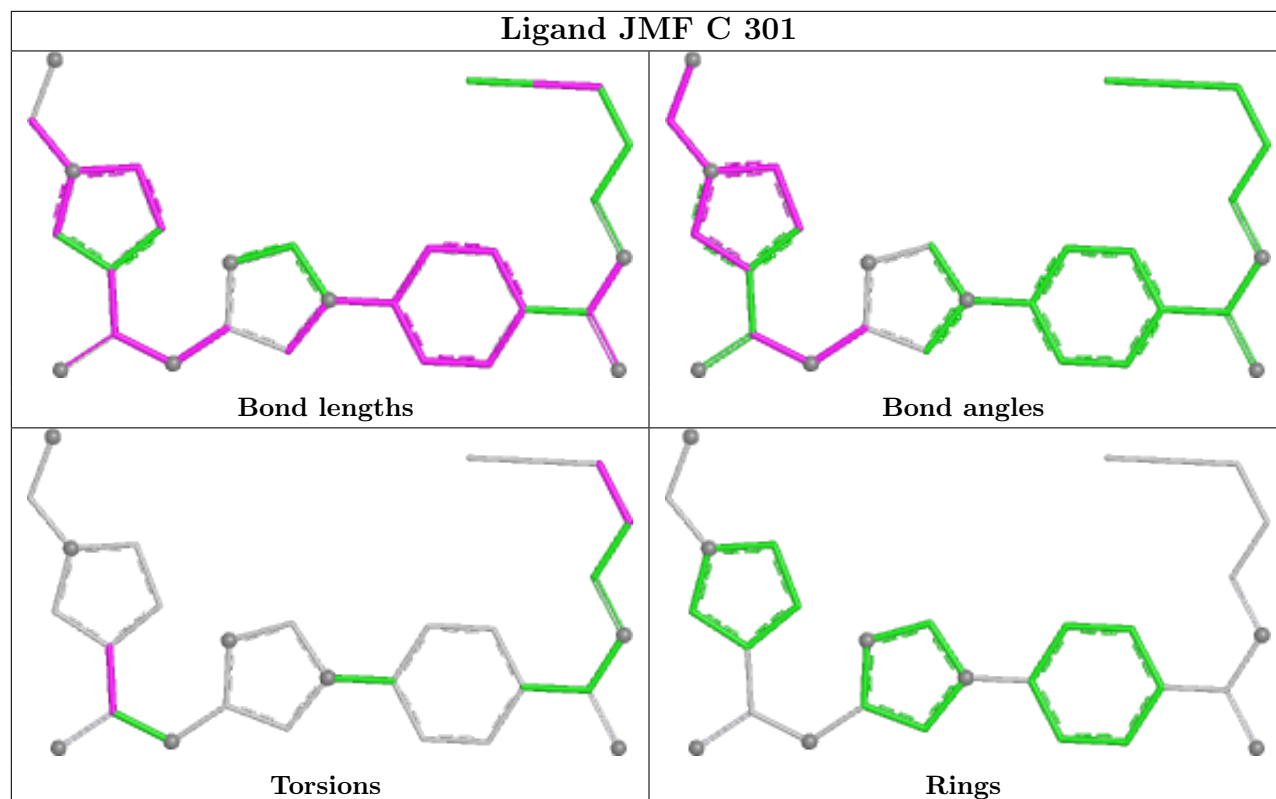
Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD 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 JMF G 301**



Bond lengths

Bond angles

Torsions

Rings

## Ligand JMF E 301



Bond lengths



Bond angles



Torsions



Rings

## Ligand JMF A 301



Bond lengths



Bond angles



Torsions



Rings

## Ligand JMF B 301



Bond lengths



Bond angles



Torsions



Rings

## Ligand JMF D 301



Bond lengths



Bond angles



Torsions



Rings

## Ligand JMF H 301



Bond lengths



Bond angles



Torsions



Rings

## Ligand JMF J 301



Bond lengths



Bond angles



Torsions



Rings

## Ligand JMF C 301


Bond lengths


Bond angles


Torsions


Rings

## Ligand JMF F 301


Bond lengths


Bond angles


Torsions


Rings

**Ligand JMF I 301**



Bond lengths

Bond angles

Torsions

Rings

## 5.7  Other polymers (i)

There are no such residues in this entry.

## 5.8  Polymer linkage issues (i)

There are no chain breaks in this entry.

# 6 Fit of model and data ⓘ

## 6.1 Protein, DNA and RNA chains ⓘ

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

| Mol | Chain | Analysed | <RSRZ> | #RSRZ>2 | | | OWAB(Å²) | Q<0.9 |
|-----|-------|----------|--------|---------|-----|-----|----------|-------|
| 1 | A | 213/227 (93%) | -0.05 | 0 | 100 | 100 | 25, 47, 89, 100 | 0 |
| 1 | B | 219/227 (96%) | 0.16 | 9 (4%) | 37 | 37 | 27, 56, 89, 102 | 0 |
| 1 | C | 204/227 (89%) | 0.49 | 15 (7%) | 14 | 14 | 27, 68, 111, 124 | 0 |
| 1 | D | 217/227 (95%) | 0.16 | 6 (2%) | 53 | 53 | 24, 59, 85, 101 | 0 |
| 1 | E | 218/227 (96%) | 0.23 | 3 (1%) | 75 | 76 | 32, 64, 96, 115 | 0 |
| 1 | F | 208/227 (91%) | 0.30 | 12 (5%) | 23 | 22 | 30, 66, 95, 105 | 0 |
| 1 | G | 212/227 (93%) | 0.31 | 6 (2%) | 53 | 53 | 35, 68, 97, 102 | 0 |
| 1 | H | 220/227 (96%) | -0.03 | 4 (1%) | 68 | 69 | 18, 44, 82, 99 | 0 |
| 1 | I | 206/227 (90%) | 0.58 | 21 (10%) | 6 | 6 | 35, 78, 112, 128 | 0 |
| 1 | J | 200/227 (88%) | 0.66 | 23 (11%) | 4 | 4 | 32, 75, 104, 121 | 0 |
| All | All | 2117/2270 (93%) | 0.28 | 99 (4%) | 31 | 31 | 18, 63, 100, 128 | 0 |

The worst 5 of 99 RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|-----|-------|-----|------|------|
| 1 | I | 67 | ILE | 4.8 |
| 1 | F | 67 | ILE | 4.7 |
| 1 | I | 76 | SER | 4.5 |
| 1 | C | 67 | ILE | 4.5 |
| 1 | J | 25 | GLN | 4.4 |

## 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 (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($\text{Å}^2$) | Q<0.9 |
|-----|------|-------|-----|-------|------|-----|------------|-------|
| 2 | JMF | G | 301 | 29/29 | 0.88 | 0.20 | 48,70,98,101 | 0 |
| 2 | JMF | D | 301 | 29/29 | 0.89 | 0.23 | 43,63,90,98 | 0 |
| 2 | JMF | C | 301 | 29/29 | 0.89 | 0.27 | 48,72,108,114 | 0 |
| 2 | JMF | J | 301 | 29/29 | 0.89 | 0.26 | 41,77,106,114 | 0 |
| 2 | JMF | E | 301 | 29/29 | 0.91 | 0.21 | 39,62,88,96 | 0 |
| 2 | JMF | F | 301 | 29/29 | 0.91 | 0.29 | 43,64,109,112 | 0 |
| 2 | JMF | A | 301 | 29/29 | 0.93 | 0.17 | 37,59,87,91 | 0 |
| 2 | JMF | I | 301 | 29/29 | 0.93 | 0.22 | 45,65,100,115 | 0 |
| 2 | JMF | B | 301 | 29/29 | 0.93 | 0.20 | 37,60,83,87 | 0 |
| 2 | JMF | H | 301 | 29/29 | 0.95 | 0.16 | 25,47,88,97 | 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.
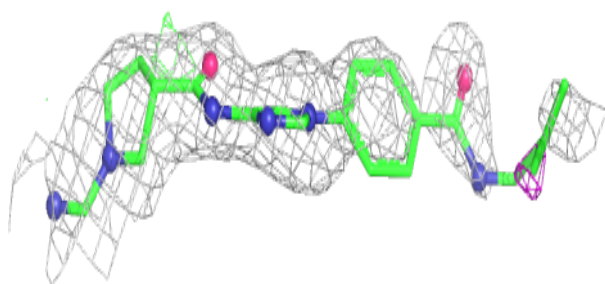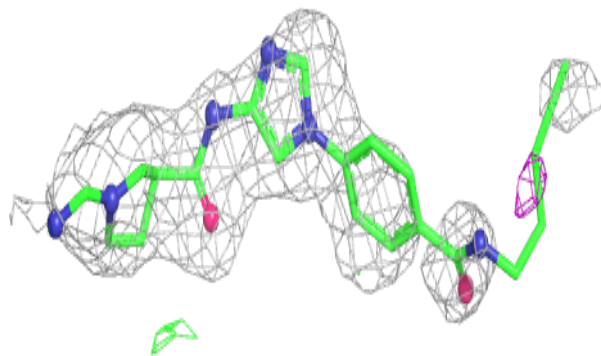
**Electron density around JMF G 301:**

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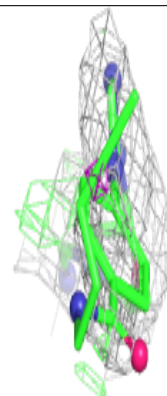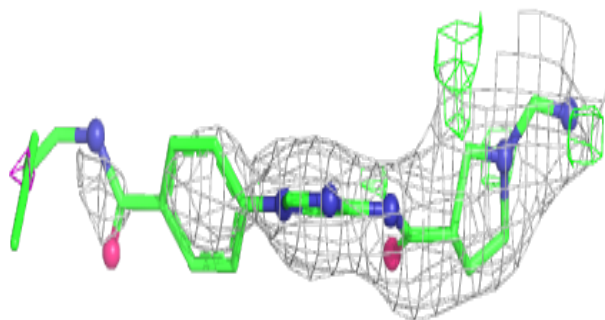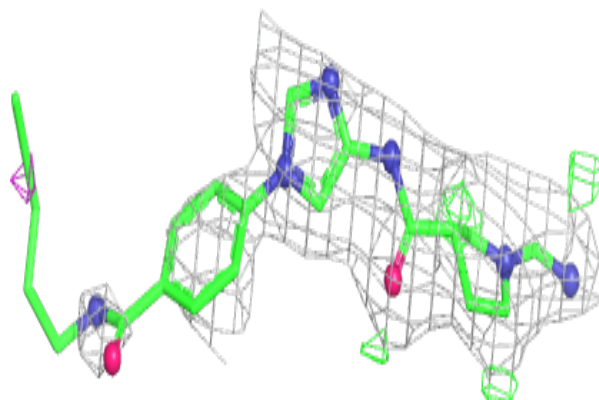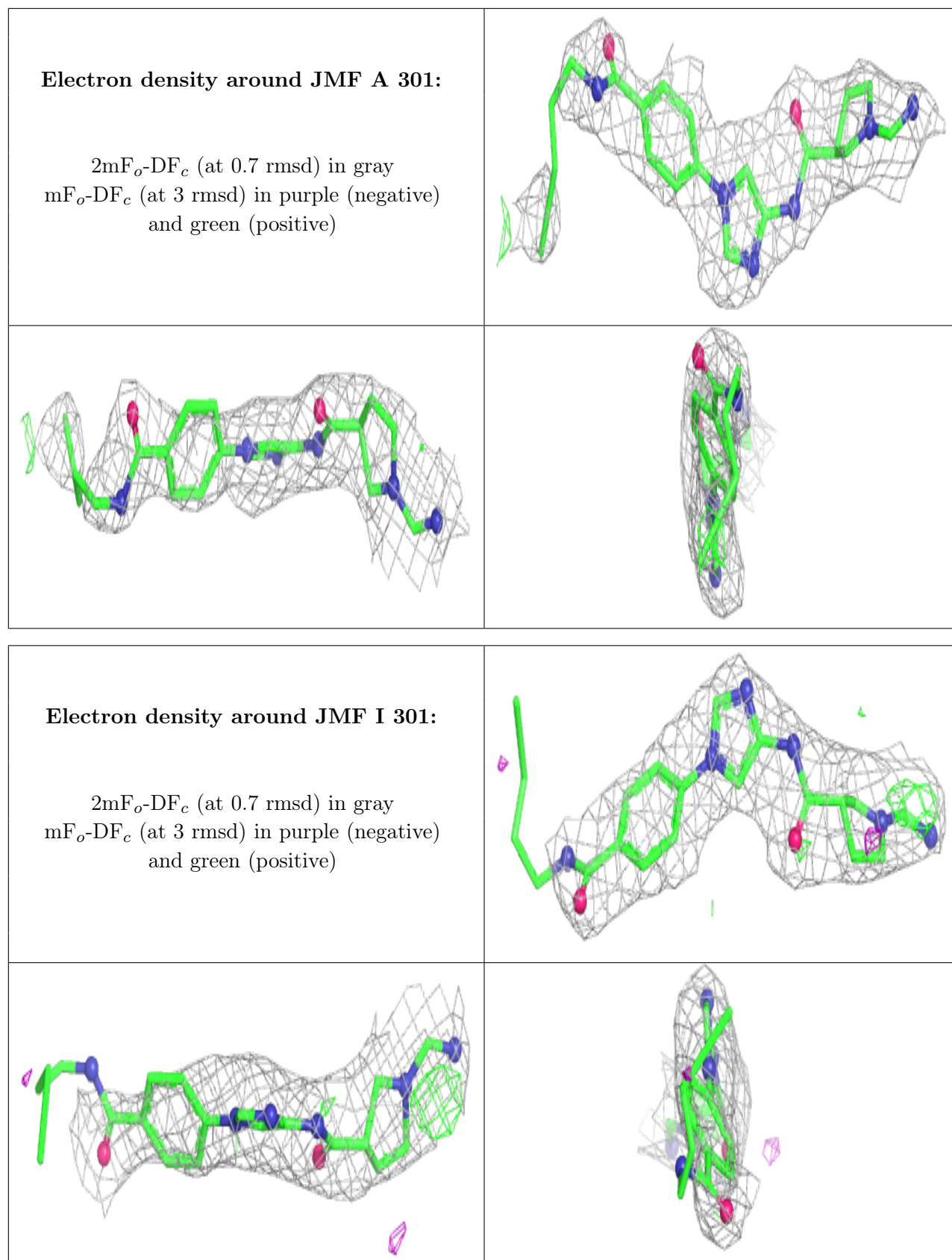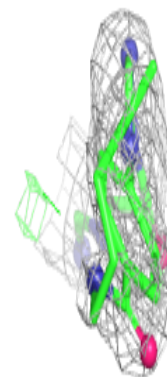$mF_o$-$DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)



**Electron density around JMF D 301:**

$2mF_o$-$DF_c$ (at 0.7 rmsd) in gray
$mF_o$-$DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
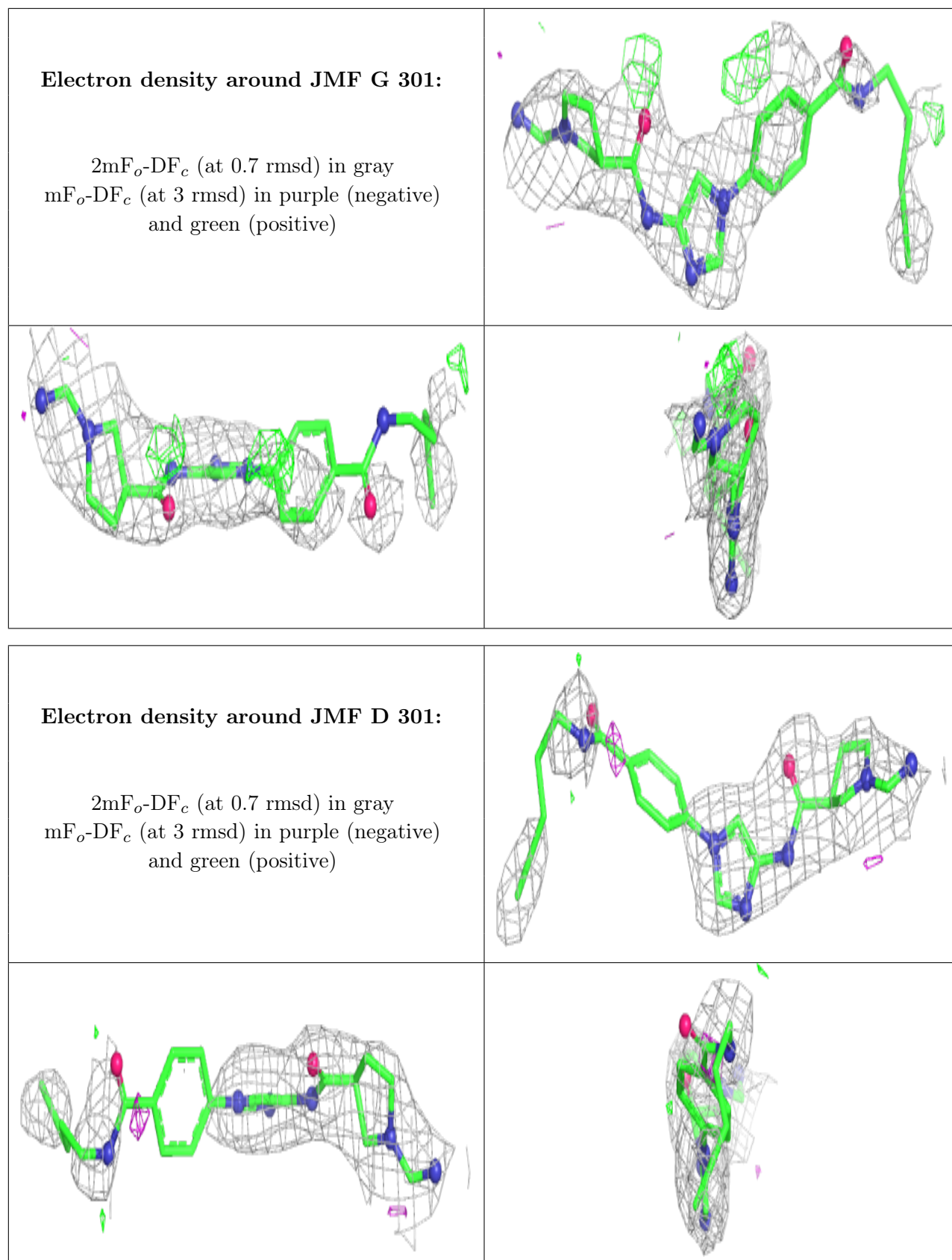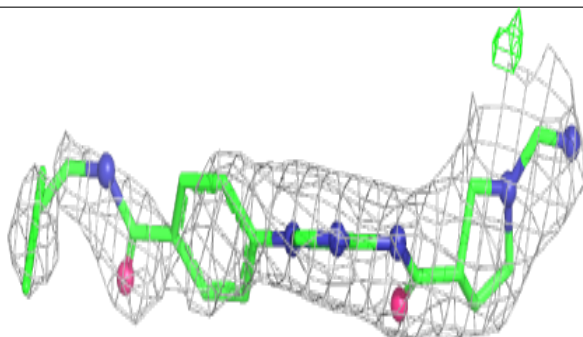
**Electron density around JMF C 301:**

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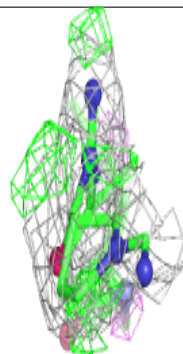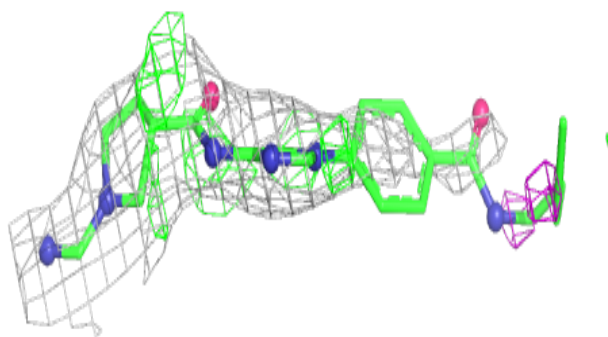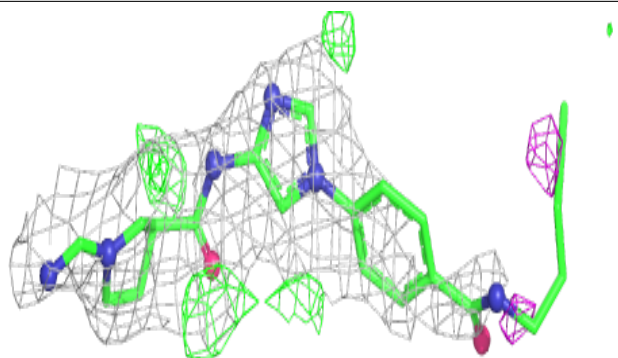


**Electron density around JMF J 301:**

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)
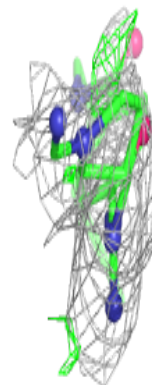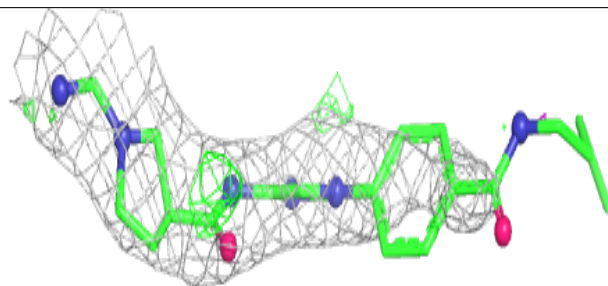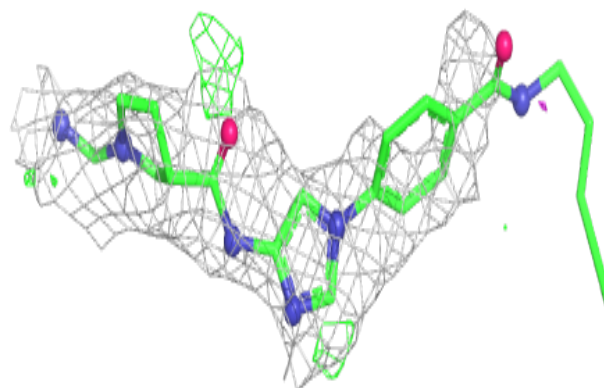
**Electron density around JMF E 301:**

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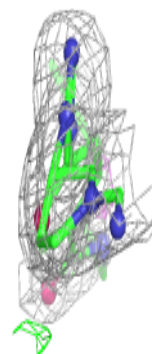


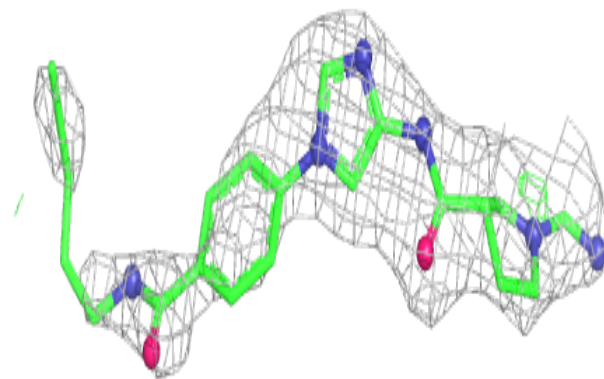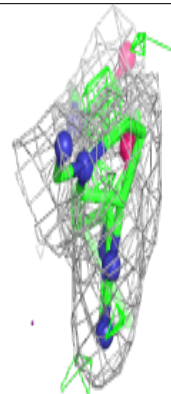**Electron density around JMF F 301:**

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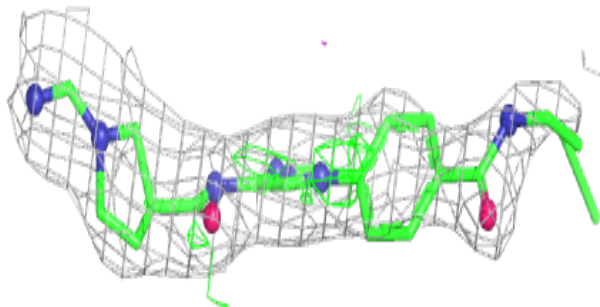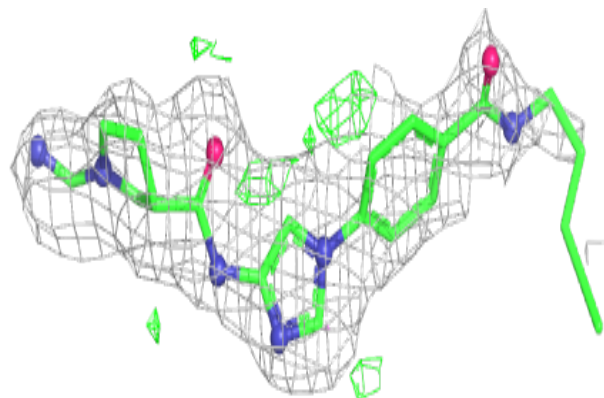and green (positive)

**Electron density around JMF A 301:**

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



**Electron density around JMF I 301:**

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

**Electron density around JMF B 301:**

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



**Electron density around JMF H 301:**

$2mF_o$-$DF_c$ (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.