



Full wwPDB X-ray Structure Validation Report ⓘ

Jun 7, 2023 – 04:34 pm BST

PDB ID : 8CBS
Title : HIV-1 Integrase Catalytic Core Domain and C-Terminal Domain in Complex with Allosteric Integrase Inhibitor MUT871
Authors : Singer, M.R.; Pye, V.E.; Yu, Z.; Cherepanov, P.
Deposited on : 2023-01-25
Resolution : 1.70 Å (reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.33
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) : 2.33

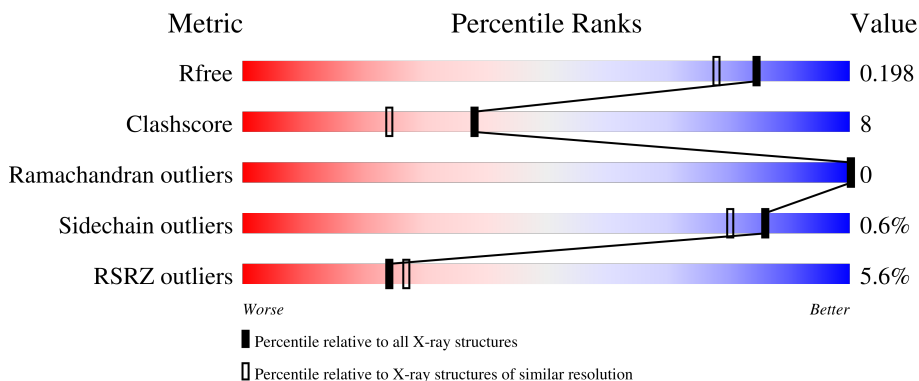
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.70 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 | 4298 (1.70-1.70) |
| Clashscore | 141614 | 4695 (1.70-1.70) |
| Ramachandran outliers | 138981 | 4610 (1.70-1.70) |
| Sidechain outliers | 138945 | 4610 (1.70-1.70) |
| RSRZ outliers | 127900 | 4222 (1.70-1.70) |

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 $\leq 5\%$. 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 | 233 | 3% 19% 18% 77% |
| 1 | B | 233 | % 52% 8% 39% |
| 1 | C | 233 | 4% 18% 78% |
| 1 | D | 233 | % 53% 8% 38% |

2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 3490 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 Integrase.

| Mol | Chain | Residues | Atoms | | | | | ZeroOcc | AltConf | Trace |
|-----|-------|----------|-------|-----|-----|-----|---|---------|---------|-------|
| | | | Total | C | N | O | S | | | |
| 1 | A | 53 | 449 | 287 | 85 | 77 | 0 | 2 | 0 | |
| 1 | B | 141 | 1125 | 723 | 195 | 202 | 5 | 0 | 5 | 0 |
| 1 | C | 52 | 439 | 280 | 82 | 77 | 0 | 2 | 0 | |
| 1 | D | 144 | 1132 | 726 | 196 | 205 | 5 | 0 | 3 | 0 |

There are 12 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|---------------------|------------|
| A | 219 | SER | - | expression tag | UNP P12497 |
| A | 243 | GLU | TRP | engineered mutation | UNP P12497 |
| A | 424 | LYS | PHE | engineered mutation | UNP P12497 |
| B | -20 | SER | - | expression tag | UNP P12497 |
| B | 4 | GLU | TRP | engineered mutation | UNP P12497 |
| B | 185 | LYS | PHE | engineered mutation | UNP P12497 |
| C | 219 | SER | - | expression tag | UNP P12497 |
| C | 243 | GLU | TRP | engineered mutation | UNP P12497 |
| C | 424 | LYS | PHE | engineered mutation | UNP P12497 |
| D | -20 | SER | - | expression tag | UNP P12497 |
| D | 4 | GLU | TRP | engineered mutation | UNP P12497 |
| D | 185 | LYS | PHE | engineered mutation | UNP P12497 |

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).

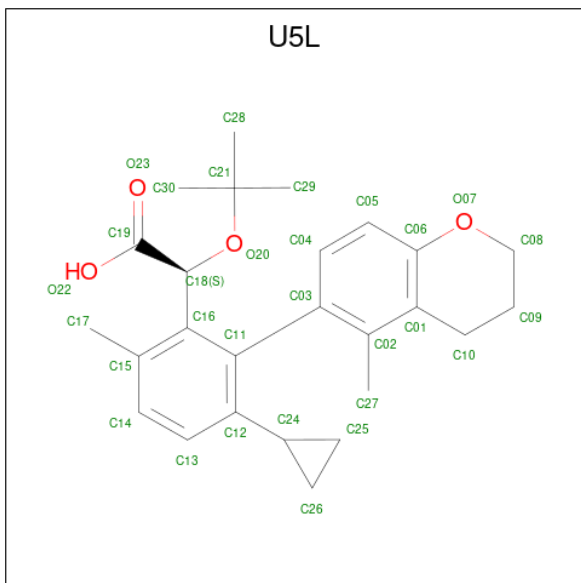


| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|-----|-------|----------|--------------------|---------|---------|
| 2 | A | 1 | Total C O 4 2 2 | 0 | 0 |
| 2 | A | 1 | Total C O 4 2 2 | 0 | 0 |
| 2 | B | 1 | Total C O 4 2 2 | 0 | 0 |
| 2 | B | 1 | Total C O 4 2 2 | 0 | 0 |
| 2 | B | 1 | Total C O 4 2 2 | 0 | 0 |
| 2 | B | 1 | Total C O 4 2 2 | 0 | 0 |
| 2 | C | 1 | Total C O 4 2 2 | 0 | 0 |
| 2 | C | 1 | Total C O 4 2 2 | 0 | 0 |
| 2 | D | 1 | Total C O 4 2 2 | 0 | 0 |
| 2 | D | 1 | Total C O 4 2 2 | 0 | 0 |
| 2 | D | 1 | Total C O 4 2 2 | 0 | 0 |
| 2 | D | 1 | Total C O 4 2 2 | 0 | 0 |
| 2 | D | 1 | Total C O 4 2 2 | 0 | 0 |
| 2 | D | 1 | Total C O 4 2 2 | 0 | 0 |

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|-----|-------|----------|-----------------|---------|---------|
| 3 | A | 1 | Total Cl 1 1 | 0 | 0 |

- Molecule 4 is (2 {S})-2-[3-cyclopropyl-6-methyl-2-(5-methyl-3,4-dihydro-2 {H}-chromen-6-yl)phenyl]-2-[(2-methylpropan-2-yl)oxy]ethanoic acid (three-letter code: U5L) (formula: C₂₆H₃₂O₄) (labeled as "Ligand of Interest" by depositor).



| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|-----|-------|----------|----------------------|---------|---------|
| 4 | B | 1 | Total C O 30 26 4 | 0 | 0 |
| 4 | D | 1 | Total C O 30 26 4 | 0 | 0 |

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|-----|-------|----------|-----------------|---------|---------|
| 5 | B | 1 | Total Mg 1 1 | 0 | 0 |
| 5 | D | 1 | Total Mg 1 1 | 0 | 0 |

- Molecule 6 is water.

| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|-----|-------|----------|------------------|---------|---------|
| 6 | A | 23 | Total O 23 23 | 0 | 0 |

Continued on next page...

Continued from previous page...

| Mol | Chain | Residues | Atoms | ZeroOcc | AltConf |
|------------|--------------|-----------------|------------------|----------------|----------------|
| 6 | B | 93 | Total O 93 93 | 0 | 0 |
| 6 | C | 22 | Total O 22 22 | 0 | 0 |
| 6 | D | 88 | Total O 88 88 | 0 | 0 |

ALA
ALA
CYS
TRP
TRP
GLY
ILE
LYS
GLN
GLU
PHE
GLY
ILE
PRO
TYR
ASN
PRO
GLN
SER
GLN
GLN
GLY
VAL
ILE
LYS
GLU
SER
MET
ASN
GLY
LEU
LYS
ILE
ILE
GLN
GLY
GLN
VAL
ARG
ASP
GLN
ALA
GLU
HIS
LEU
LEU
THR
PHE
VAL
GLN
MET
ALA
LYS
LYS
ARG

LYS
GLY
ILE
GLY
TRP
SER
ALA
GLY
GLU
ARG
GLY
VAL
ASP
ILE
ILE
ALA
THR
ASP
ILE
GLN
THR
LYS
GLU

● Molecule 1: Integrase



SER
ILE
GLN
ASN
PHE
ARG
VAL
TYR
TYR
ARG
ASP
MET
SER
ARG
ASP
PRO
VAL
ASP
TRP
LYS
GLY
PRO
ALA
GLN
LYS
LEU
LEU
GLU
LYS
GLY
GLU
GLY
ALA
VAL
VAL
ILE
GLN
ASP
ASN
SER
ASP
ILE
PRO
LYS
VAL
SER
GLN
PRO
ARG
ARG
LYS
ALA
LYS
ILE
ILE
ILE
ARG
ASP
TYR
GLY
LYS
MET
ALA
GLY
ASP

ASP
CYS
VAL
ALA
SER
ARG
GLN
ASP
GLU
ASP
MET
HIS
GLY
GLN
VAL
VAL
ASP
C56
V77
I84
E85
A86
E87
T93
R107
M120
W131
W132
G140
ILE
PRO
TYR
ASN
PRO
GLN
SER
GLN
G149
V150
I151
M154
R166
K173
R187
K188
GLY
GLY
ILE
G193

V194
E198
V201
A205
T206
D207
K211
GLU

4 Data and refinement statistics

| Property | Value | Source |
|---|---|------------------|
| Space group | P 1 21 1 | Depositor |
| Cell constants a, b, c, α , β , γ | 53.23Å 64.85Å 70.44Å 90.00° 101.50° 90.00° | Depositor |
| Resolution (Å) | 52.16 – 1.70 52.16 – 1.70 | Depositor EDS |
| % Data completeness (in resolution range) | 100.0 (52.16-1.70) 100.0 (52.16-1.70) | Depositor EDS |
| R_{merge} | (Not available) | Depositor |
| R_{sym} | (Not available) | Depositor |
| $\langle I/\sigma(I) \rangle$ ¹ | 1.03 (at 1.70Å) | Xtrriage |
| Refinement program | PHENIX 1.20.1_4487 | Depositor |
| R, R_{free} | 0.167 , 0.198 0.165 , 0.198 | Depositor DCC |
| R_{free} test set | 2574 reflections (4.97%) | wwPDB-VP |
| Wilson B-factor (Å ²) | 38.0 | Xtrriage |
| Anisotropy | 0.232 | Xtrriage |
| Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²) | 0.35 , 58.2 | EDS |
| L-test for twinning ² | $\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$ | Xtrriage |
| Estimated twinning fraction | No twinning to report. | Xtrriage |
| F_o, F_c correlation | 0.97 | EDS |
| Total number of atoms | 3490 | wwPDB-VP |
| Average B, all atoms (Å ²) | 51.0 | wwPDB-VP |

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.94% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: EDO, U5L, MG, CL

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 | 0.65 | 0/460 | 0.81 | 0/616 |
| 1 | B | 0.79 | 1/1156 (0.1%) | 0.75 | 0/1563 |
| 1 | C | 0.69 | 0/450 | 0.80 | 0/605 |
| 1 | D | 0.80 | 2/1162 (0.2%) | 0.79 | 0/1570 |
| All | All | 0.76 | 3/3228 (0.1%) | 0.78 | 0/4354 |

All (3) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|--------|-------|-------------|----------|
| 1 | B | 87 | GLU | CD-OE1 | 5.56 | 1.31 | 1.25 |
| 1 | D | 87 | GLU | CD-OE1 | 5.39 | 1.31 | 1.25 |
| 1 | D | 85 | GLU | CB-CG | -5.18 | 1.42 | 1.52 |

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

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 | 449 | 0 | 473 | 9 | 0 |
| 1 | B | 1125 | 0 | 1145 | 18 | 0 |
| 1 | C | 439 | 0 | 456 | 8 | 0 |
| 1 | D | 1132 | 0 | 1156 | 18 | 0 |

Continued on next page...

Continued from previous page...

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 2 | A | 8 | 0 | 12 | 6 | 0 |
| 2 | B | 16 | 0 | 24 | 4 | 0 |
| 2 | C | 8 | 0 | 12 | 3 | 0 |
| 2 | D | 24 | 0 | 36 | 6 | 0 |
| 3 | A | 1 | 0 | 0 | 0 | 0 |
| 4 | B | 30 | 0 | 0 | 0 | 0 |
| 4 | D | 30 | 0 | 0 | 0 | 0 |
| 5 | B | 1 | 0 | 0 | 0 | 0 |
| 5 | D | 1 | 0 | 0 | 0 | 0 |
| 6 | A | 23 | 0 | 0 | 0 | 0 |
| 6 | B | 93 | 0 | 0 | 2 | 0 |
| 6 | C | 22 | 0 | 0 | 2 | 0 |
| 6 | D | 88 | 0 | 0 | 0 | 0 |
| All | All | 3490 | 0 | 3314 | 50 | 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 8.

All (50) 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:B:93:THR:HA | 2:B:601:EDO:H11 | 1.67 | 0.76 |
| 1:B:77[B]:VAL:HG22 | 1:B:84:ILE:HG22 | 1.70 | 0.72 |
| 1:C:249[A]:VAL:HG11 | 1:C:267:ILE:HD11 | 1.74 | 0.69 |
| 1:D:173:LYS:HE2 | 2:D:308:EDO:H22 | 1.77 | 0.67 |
| 1:D:187:ARG:NH1 | 1:D:193:GLY:O | 2.31 | 0.64 |
| 1:D:166:ARG:HB3 | 2:D:305:EDO:H21 | 1.79 | 0.63 |
| 1:C:271:TYR:CE2 | 2:C:501:EDO:H11 | 2.33 | 0.63 |
| 1:A:224:ARG:HG2 | 2:A:502:EDO:H21 | 1.81 | 0.63 |
| 1:B:77[B]:VAL:HG23 | 1:B:151:ILE:HD12 | 1.81 | 0.63 |
| 1:D:77[B]:VAL:HG22 | 1:D:84:ILE:HG22 | 1.82 | 0.61 |
| 1:B:185:LYS:HB2 | 2:C:501:EDO:H12 | 1.82 | 0.61 |
| 1:A:235:TRP:HD1 | 2:A:501:EDO:H21 | 1.66 | 0.60 |
| 1:C:272:GLY:HA3 | 1:D:132:TRP:O | 2.02 | 0.59 |
| 1:C:264:LYS:NZ | 6:C:602:HOH:O | 2.35 | 0.58 |
| 1:D:187:ARG:NH2 | 1:D:198:GLU:OE2 | 2.36 | 0.58 |
| 1:D:85:GLU:OE1 | 1:D:107:ARG:NH1 | 2.37 | 0.57 |
| 1:B:67[B]:HIS:HE1 | 1:B:92:GLU:OE1 | 1.91 | 0.53 |
| 1:D:93:THR:HA | 2:D:304:EDO:H22 | 1.92 | 0.50 |
| 1:B:77[B]:VAL:HG23 | 1:B:151:ILE:CD1 | 2.41 | 0.50 |
| 1:B:185:LYS:HA | 2:B:602:EDO:H21 | 1.93 | 0.50 |

Continued on next page...

Continued from previous page...

| Atom-1 | Atom-2 | Interatomic distance (Å) | Clash overlap (Å) |
|---------------------|--------------------|--------------------------|-------------------|
| 1:D:207:ASP:O | 1:D:211:LYS:HG2 | 2.11 | 0.49 |
| 1:B:77[B]:VAL:CG2 | 1:B:151:ILE:HD12 | 2.43 | 0.49 |
| 1:B:87:GLU:OE1 | 6:B:701:HOH:O | 2.20 | 0.48 |
| 1:D:173:LYS:CE | 2:D:308:EDO:H22 | 2.43 | 0.48 |
| 1:A:234:VAL:HA | 2:A:501:EDO:H22 | 1.95 | 0.47 |
| 1:B:85:GLU:OE1 | 1:B:107:ARG:NH1 | 2.47 | 0.47 |
| 1:A:235:TRP:H | 2:A:501:EDO:C2 | 2.28 | 0.47 |
| 1:B:62:GLN:HG2 | 1:B:114[B]:HIS:HB3 | 1.97 | 0.46 |
| 1:A:224:ARG:HE | 2:A:502:EDO:H21 | 1.80 | 0.46 |
| 1:B:201:VAL:HG12 | 1:D:205:ALA:HB2 | 1.97 | 0.46 |
| 1:C:240:LYS:HA | 6:C:615:HOH:O | 2.16 | 0.46 |
| 1:A:224:ARG:NH2 | 2:A:502:EDO:O1 | 2.47 | 0.45 |
| 1:D:166:ARG:HB3 | 2:D:305:EDO:C2 | 2.46 | 0.45 |
| 1:D:211:LYS:HD3 | 1:D:211:LYS:HA | 1.78 | 0.44 |
| 1:C:229:ASP:O | 1:C:231:ARG:N | 2.43 | 0.44 |
| 1:D:120:ASN:HA | 2:D:304:EDO:H11 | 1.99 | 0.44 |
| 1:C:224:ARG:HD3 | 1:D:131:TRP:CE2 | 2.53 | 0.44 |
| 2:B:601:EDO:C2 | 2:B:604:EDO:H11 | 2.49 | 0.43 |
| 1:B:62:GLN:HG2 | 1:B:114[A]:HIS:HB2 | 2.00 | 0.43 |
| 1:A:229:ASP:HB3 | 1:A:230:SER:H | 1.63 | 0.43 |
| 1:B:95:GLN:HG3 | 6:B:773:HOH:O | 2.18 | 0.43 |
| 1:D:77[B]:VAL:HG23 | 1:D:151:ILE:HD12 | 1.99 | 0.43 |
| 1:B:119:SER:HB2 | 2:B:604:EDO:H22 | 2.01 | 0.42 |
| 1:B:204[B]:ILE:HG22 | 1:D:201:VAL:HG11 | 2.01 | 0.42 |
| 1:C:271:TYR:HE2 | 2:C:501:EDO:H11 | 1.79 | 0.42 |
| 1:D:84:ILE:HG21 | 1:D:154:MET:HG2 | 2.00 | 0.42 |
| 1:A:249[A]:VAL:HG11 | 1:A:267:ILE:HD11 | 2.00 | 0.42 |
| 1:B:152:GLU:O | 1:B:156:LYS:HG3 | 2.18 | 0.42 |
| 1:A:245:GLY:O | 1:A:262:ARG:NH1 | 2.51 | 0.42 |
| 1:B:186:LYS:HE3 | 1:B:186:LYS:HB2 | 1.81 | 0.42 |

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 | 53/233 (23%) | 49 (92%) | 4 (8%) | 0 | 100 | 100 |
| 1 | B | 140/233 (60%) | 139 (99%) | 1 (1%) | 0 | 100 | 100 |
| 1 | C | 52/233 (22%) | 48 (92%) | 4 (8%) | 0 | 100 | 100 |
| 1 | D | 141/233 (60%) | 139 (99%) | 2 (1%) | 0 | 100 | 100 |
| All | All | 386/932 (41%) | 375 (97%) | 11 (3%) | 0 | 100 | 100 |

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains [i](#)

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

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

| Mol | Chain | Analysed | Rotameric | Outliers | Percentiles | |
|-----|-------|---------------|------------|----------|-------------|-----|
| 1 | A | 48/193 (25%) | 47 (98%) | 1 (2%) | 53 | 36 |
| 1 | B | 120/193 (62%) | 120 (100%) | 0 | 100 | 100 |
| 1 | C | 47/193 (24%) | 46 (98%) | 1 (2%) | 53 | 36 |
| 1 | D | 121/193 (63%) | 121 (100%) | 0 | 100 | 100 |
| All | All | 336/772 (44%) | 334 (99%) | 2 (1%) | 86 | 80 |

All (2) residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | A | 229 | ASP |
| 1 | C | 271 | TYR |

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

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 19 ligands modelled in this entry, 3 are monoatomic - leaving 16 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 |
| 2 | EDO | D | 305 | - | 3,3,3 | 0.71 | 0 | 2,2,2 | 0.72 | 0 |
| 2 | EDO | D | 302 | - | 3,3,3 | 0.63 | 0 | 2,2,2 | 0.38 | 0 |
| 2 | EDO | A | 502 | - | 3,3,3 | 0.37 | 0 | 2,2,2 | 0.57 | 0 |
| 2 | EDO | B | 604 | - | 3,3,3 | 0.43 | 0 | 2,2,2 | 0.58 | 0 |
| 2 | EDO | C | 502 | - | 3,3,3 | 0.56 | 0 | 2,2,2 | 0.22 | 0 |
| 2 | EDO | D | 306 | - | 3,3,3 | 0.62 | 0 | 2,2,2 | 0.11 | 0 |
| 4 | U5L | B | 605 | - | 32,33,33 | 2.05 | 9 (28%) | 39,50,50 | 1.56 | 6 (15%) |
| 2 | EDO | D | 308 | - | 3,3,3 | 0.52 | 0 | 2,2,2 | 0.11 | 0 |
| 2 | EDO | D | 303 | - | 3,3,3 | 0.78 | 0 | 2,2,2 | 0.55 | 0 |
| 2 | EDO | C | 501 | - | 3,3,3 | 0.53 | 0 | 2,2,2 | 0.19 | 0 |
| 2 | EDO | D | 304 | - | 3,3,3 | 0.40 | 0 | 2,2,2 | 0.86 | 0 |
| 4 | U5L | D | 301 | - | 32,33,33 | 1.88 | 7 (21%) | 39,50,50 | 1.51 | 6 (15%) |
| 2 | EDO | B | 603 | - | 3,3,3 | 0.75 | 0 | 2,2,2 | 0.66 | 0 |
| 2 | EDO | B | 602 | - | 3,3,3 | 0.43 | 0 | 2,2,2 | 0.32 | 0 |
| 2 | EDO | A | 501 | - | 3,3,3 | 0.66 | 0 | 2,2,2 | 0.30 | 0 |
| 2 | EDO | B | 601 | - | 3,3,3 | 0.59 | 0 | 2,2,2 | 0.47 | 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 | EDO | D | 305 | - | - | 1/1/1/1 | - |
| 2 | EDO | D | 302 | - | - | 0/1/1/1 | - |
| 2 | EDO | A | 502 | - | - | 0/1/1/1 | - |
| 2 | EDO | B | 604 | - | - | 1/1/1/1 | - |
| 2 | EDO | C | 502 | - | - | 0/1/1/1 | - |
| 2 | EDO | D | 306 | - | - | 1/1/1/1 | - |
| 4 | U5L | B | 605 | - | - | 3/20/30/30 | 0/4/4/4 |
| 2 | EDO | D | 308 | - | - | 0/1/1/1 | - |
| 2 | EDO | D | 303 | - | - | 0/1/1/1 | - |
| 2 | EDO | C | 501 | - | - | 0/1/1/1 | - |
| 2 | EDO | D | 304 | - | - | 0/1/1/1 | - |
| 4 | U5L | D | 301 | - | - | 2/20/30/30 | 0/4/4/4 |
| 2 | EDO | B | 603 | - | - | 1/1/1/1 | - |
| 2 | EDO | B | 602 | - | - | 0/1/1/1 | - |
| 2 | EDO | A | 501 | - | - | 1/1/1/1 | - |
| 2 | EDO | B | 601 | - | - | 1/1/1/1 | - |

All (16) bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|-----|-------|-----|------|---------|-------|-------------|----------|
| 4 | D | 301 | U5L | C10-C01 | -7.06 | 1.40 | 1.51 |
| 4 | B | 605 | U5L | C10-C01 | -6.91 | 1.40 | 1.51 |
| 4 | B | 605 | U5L | C17-C15 | 4.16 | 1.59 | 1.51 |
| 4 | B | 605 | U5L | C05-C04 | 3.31 | 1.44 | 1.38 |
| 4 | D | 301 | U5L | C27-C02 | 3.02 | 1.57 | 1.51 |
| 4 | D | 301 | U5L | C03-C11 | 2.97 | 1.54 | 1.50 |
| 4 | B | 605 | U5L | O23-C19 | 2.79 | 1.30 | 1.22 |
| 4 | B | 605 | U5L | C03-C11 | 2.76 | 1.53 | 1.50 |
| 4 | B | 605 | U5L | C14-C13 | 2.59 | 1.43 | 1.38 |
| 4 | B | 605 | U5L | C11-C16 | 2.35 | 1.44 | 1.40 |
| 4 | D | 301 | U5L | O23-C19 | 2.30 | 1.29 | 1.22 |
| 4 | B | 605 | U5L | C28-C21 | 2.20 | 1.57 | 1.51 |
| 4 | D | 301 | U5L | C04-C03 | 2.12 | 1.43 | 1.40 |
| 4 | D | 301 | U5L | C16-C18 | 2.10 | 1.55 | 1.52 |
| 4 | D | 301 | U5L | C13-C12 | -2.01 | 1.37 | 1.39 |
| 4 | B | 605 | U5L | C16-C18 | 2.01 | 1.55 | 1.52 |

All (12) bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(°) | Ideal(°) |
|-----|-------|-----|------|-------------|-------|-------------|----------|
| 4 | D | 301 | U5L | C25-C24-C12 | 3.94 | 132.30 | 120.68 |
| 4 | B | 605 | U5L | C13-C12-C11 | 3.82 | 121.49 | 118.49 |
| 4 | B | 605 | U5L | C13-C12-C24 | -3.60 | 114.10 | 120.20 |

Continued on next page...

Continued from previous page...

| Mol | Chain | Res | Type | Atoms | Z | Observed(°) | Ideal(°) |
|-----|-------|-----|------|-------------|-------|-------------|----------|
| 4 | D | 301 | U5L | C11-C03-C02 | 3.59 | 126.16 | 122.00 |
| 4 | B | 605 | U5L | C10-C01-C02 | -3.44 | 114.90 | 121.31 |
| 4 | D | 301 | U5L | O22-C19-O23 | -3.34 | 116.52 | 124.09 |
| 4 | B | 605 | U5L | O07-C06-C01 | -3.03 | 119.74 | 122.72 |
| 4 | D | 301 | U5L | O20-C18-C19 | -2.76 | 101.70 | 109.38 |
| 4 | B | 605 | U5L | O22-C19-O23 | -2.74 | 117.87 | 124.09 |
| 4 | D | 301 | U5L | C17-C15-C14 | -2.28 | 115.85 | 120.31 |
| 4 | D | 301 | U5L | C13-C14-C15 | -2.21 | 118.61 | 121.97 |
| 4 | B | 605 | U5L | C25-C24-C12 | 2.08 | 126.81 | 120.68 |

There are no chirality outliers.

All (11) torsion outliers are listed below:

| Mol | Chain | Res | Type | Atoms |
|-----|-------|-----|------|-----------------|
| 4 | B | 605 | U5L | C16-C18-C19-O22 |
| 4 | D | 301 | U5L | C16-C18-C19-O22 |
| 4 | B | 605 | U5L | C16-C18-C19-O23 |
| 4 | D | 301 | U5L | C16-C18-C19-O23 |
| 2 | B | 601 | EDO | O1-C1-C2-O2 |
| 2 | B | 604 | EDO | O1-C1-C2-O2 |
| 2 | D | 305 | EDO | O1-C1-C2-O2 |
| 2 | A | 501 | EDO | O1-C1-C2-O2 |
| 4 | B | 605 | U5L | C15-C16-C18-O20 |
| 2 | D | 306 | EDO | O1-C1-C2-O2 |
| 2 | B | 603 | EDO | O1-C1-C2-O2 |

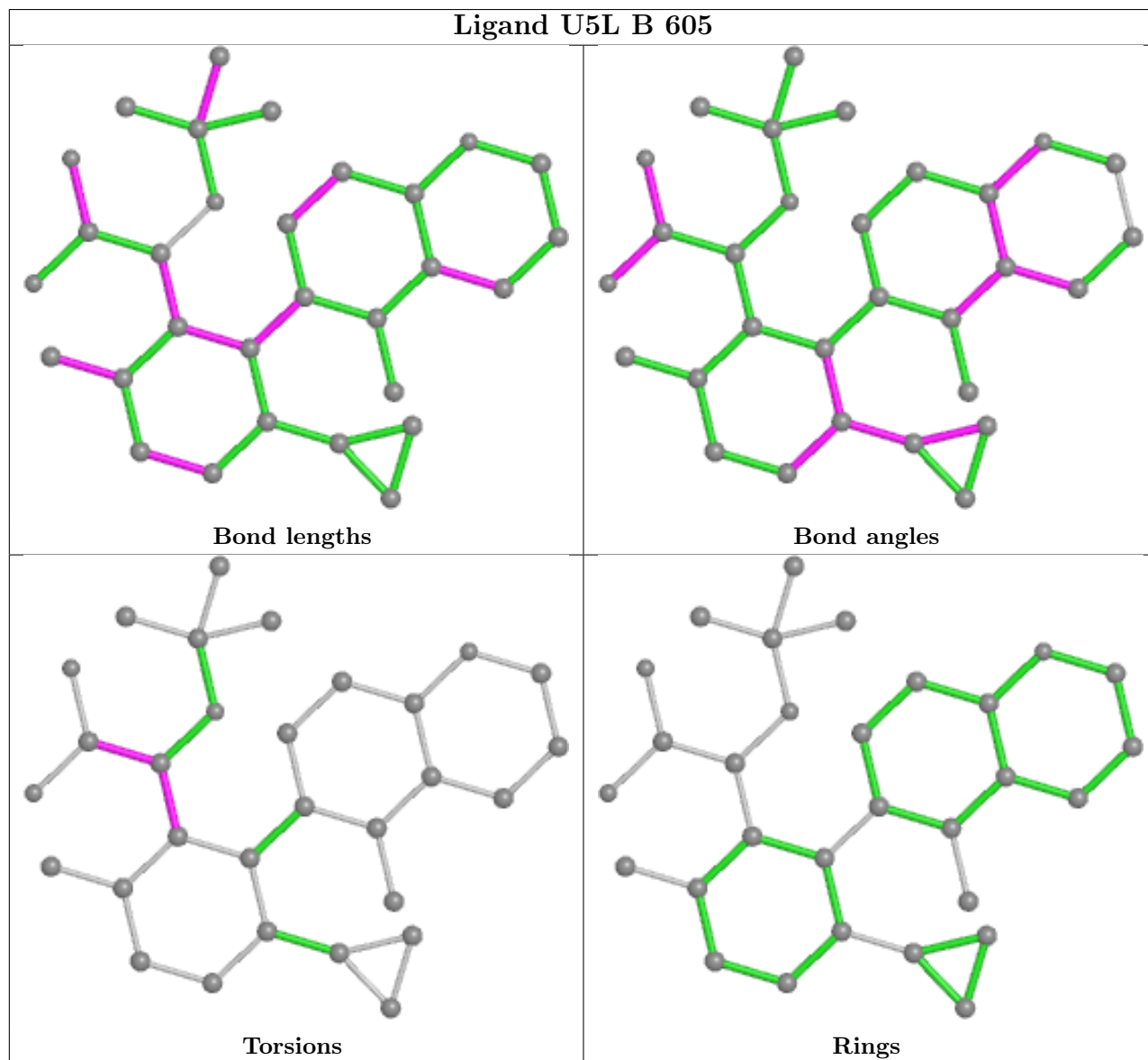
There are no ring outliers.

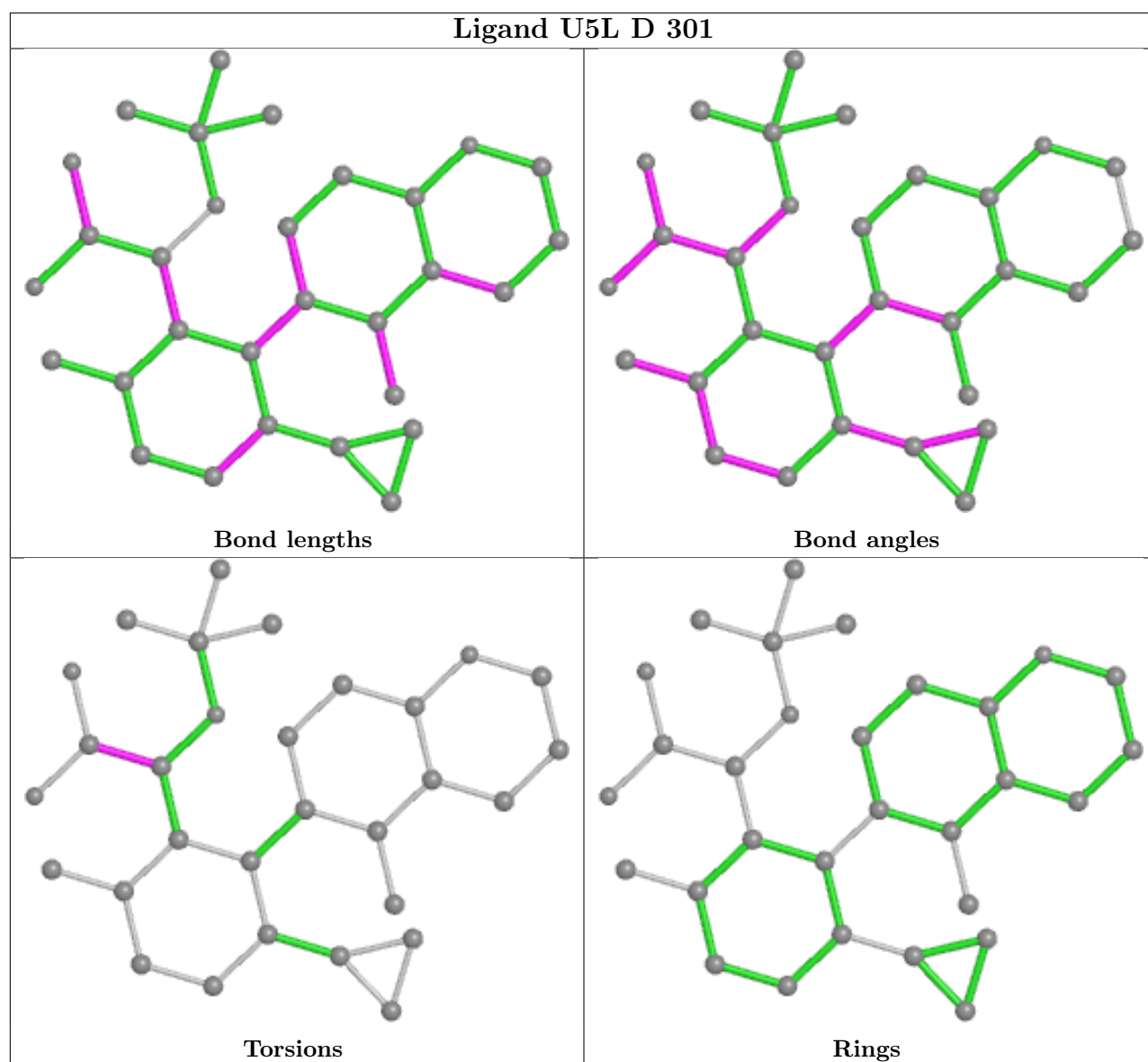
9 monomers are involved in 19 short contacts:

| Mol | Chain | Res | Type | Clashes | Symm-Clashes |
|-----|-------|-----|------|---------|--------------|
| 2 | D | 305 | EDO | 2 | 0 |
| 2 | A | 502 | EDO | 3 | 0 |
| 2 | B | 604 | EDO | 2 | 0 |
| 2 | D | 308 | EDO | 2 | 0 |
| 2 | C | 501 | EDO | 3 | 0 |
| 2 | D | 304 | EDO | 2 | 0 |
| 2 | B | 602 | EDO | 1 | 0 |
| 2 | A | 501 | EDO | 3 | 0 |
| 2 | B | 601 | EDO | 2 | 0 |

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths,

bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 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.





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, 95th 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 | 53/233 (22%) | 0.54 | 6 (11%) 5 6 | 42, 61, 107, 141 | 0 |
| 1 | B | 141/233 (60%) | 0.05 | 3 (2%) 63 67 | 30, 40, 74, 96 | 0 |
| 1 | C | 52/233 (22%) | 0.80 | 10 (19%) 1 1 | 42, 61, 106, 114 | 0 |
| 1 | D | 144/233 (61%) | 0.07 | 3 (2%) 63 67 | 30, 42, 74, 99 | 0 |
| All | All | 390/932 (41%) | 0.22 | 22 (5%) 24 27 | 30, 45, 91, 141 | 0 |

All (22) RSRZ outliers are listed below:

| Mol | Chain | Res | Type | RSRZ |
|-----|-------|-----|------|------|
| 1 | A | 255 | SER | 5.5 |
| 1 | D | 56 | CYS | 5.4 |
| 1 | C | 255 | SER | 5.1 |
| 1 | A | 254 | ASN | 4.9 |
| 1 | C | 223 | PHE | 4.7 |
| 1 | B | 56 | CYS | 4.7 |
| 1 | C | 222 | ASN | 4.5 |
| 1 | A | 230 | SER | 4.2 |
| 1 | C | 273 | LYS | 4.0 |
| 1 | C | 271 | TYR | 3.6 |
| 1 | A | 246 | GLU | 3.5 |
| 1 | A | 231 | ARG | 3.4 |
| 1 | D | 211 | LYS | 3.3 |
| 1 | B | 194 | TYR | 3.1 |
| 1 | D | 194 | TYR | 3.0 |
| 1 | C | 270 | ASP | 2.7 |
| 1 | C | 254 | ASN | 2.7 |
| 1 | B | 58 | PRO | 2.4 |
| 1 | C | 272 | GLY | 2.3 |
| 1 | A | 253 | ASP | 2.3 |
| 1 | C | 231 | ARG | 2.1 |
| 1 | C | 268 | ILE | 2.1 |

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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](#)

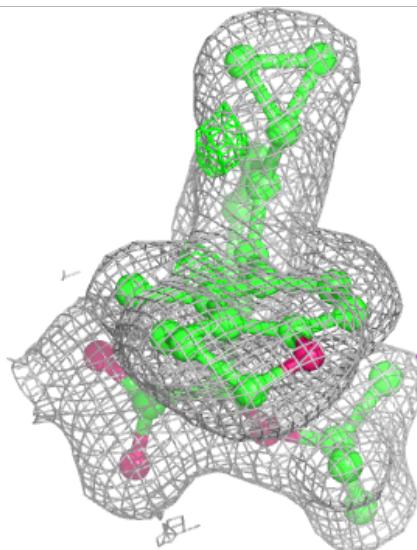
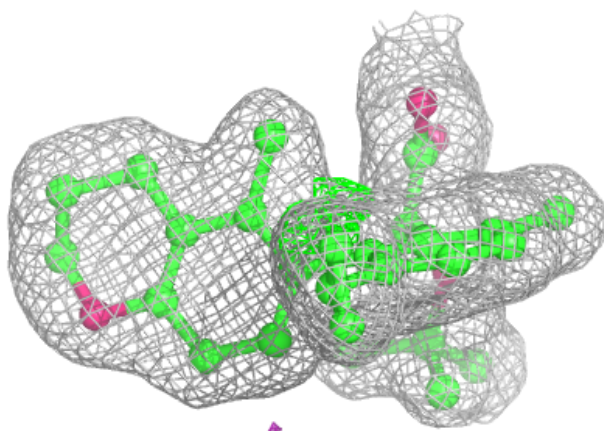
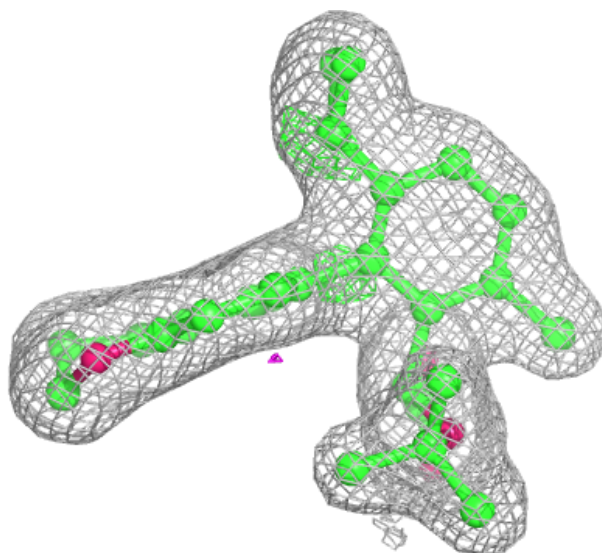
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

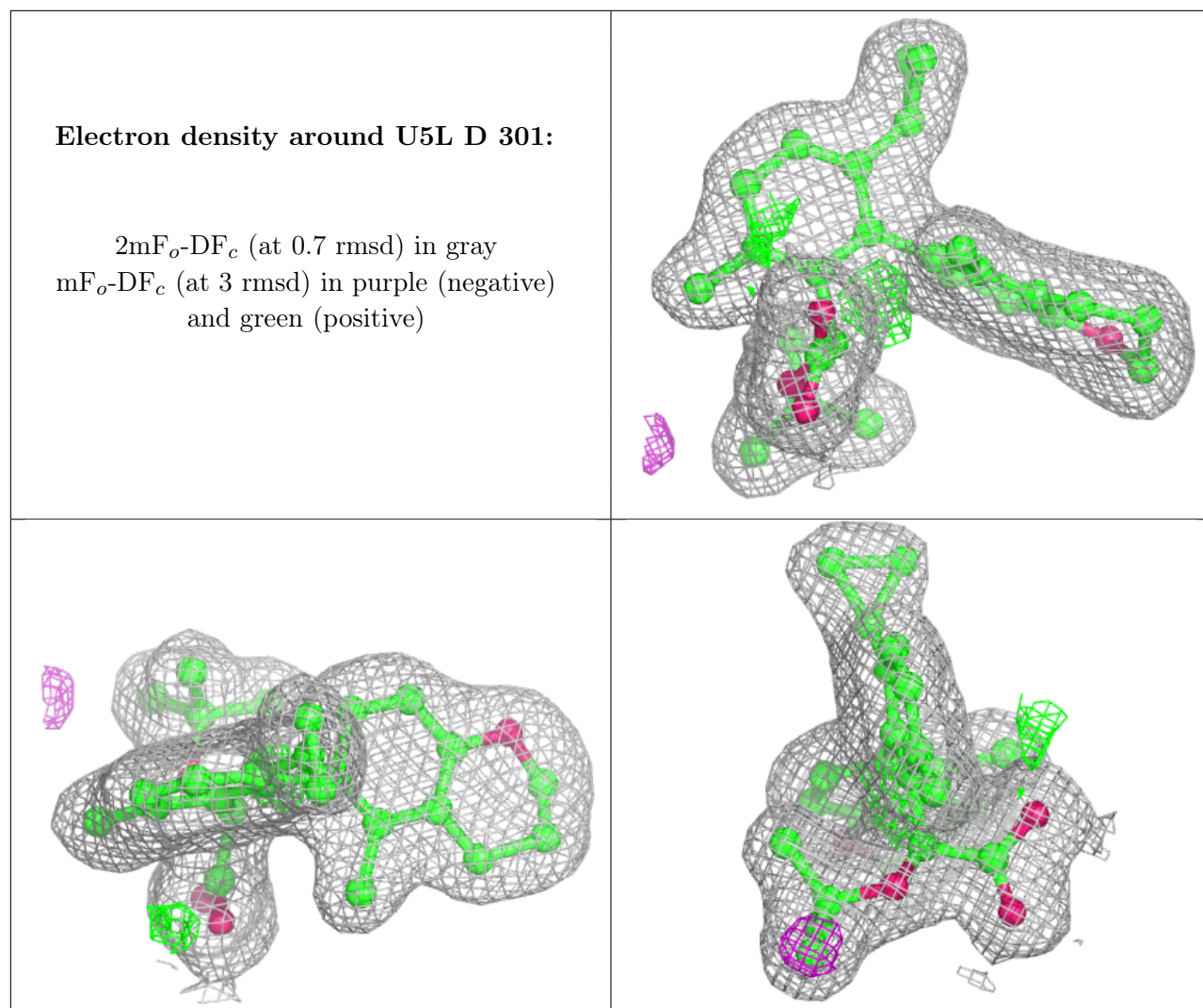
| Mol | Type | Chain | Res | Atoms | RSCC | RSR | B-factors(Å ²) | Q<0.9 |
|-----|------|-------|-----|-------|------|------|----------------------------|-------|
| 2 | EDO | D | 305 | 4/4 | 0.69 | 0.30 | 59,63,64,68 | 0 |
| 2 | EDO | B | 603 | 4/4 | 0.71 | 0.17 | 54,58,58,70 | 0 |
| 2 | EDO | D | 306 | 4/4 | 0.76 | 0.15 | 59,61,66,72 | 0 |
| 2 | EDO | C | 502 | 4/4 | 0.82 | 0.34 | 72,72,75,82 | 0 |
| 2 | EDO | B | 604 | 4/4 | 0.83 | 0.20 | 48,54,59,64 | 0 |
| 2 | EDO | B | 601 | 4/4 | 0.84 | 0.12 | 49,58,63,64 | 0 |
| 2 | EDO | A | 502 | 4/4 | 0.84 | 0.26 | 64,65,66,85 | 0 |
| 2 | EDO | A | 501 | 4/4 | 0.86 | 0.21 | 52,57,71,72 | 0 |
| 2 | EDO | D | 303 | 4/4 | 0.86 | 0.11 | 44,58,63,68 | 0 |
| 2 | EDO | C | 501 | 4/4 | 0.88 | 0.15 | 55,55,68,73 | 0 |
| 2 | EDO | D | 302 | 4/4 | 0.89 | 0.15 | 45,47,48,49 | 0 |
| 2 | EDO | D | 308 | 4/4 | 0.91 | 0.09 | 61,68,78,84 | 0 |
| 3 | CL | A | 503 | 1/1 | 0.92 | 0.05 | 65,65,65,65 | 0 |
| 2 | EDO | B | 602 | 4/4 | 0.94 | 0.33 | 59,68,71,72 | 0 |
| 2 | EDO | D | 304 | 4/4 | 0.95 | 0.17 | 53,60,65,67 | 0 |
| 4 | U5L | B | 605 | 30/30 | 0.97 | 0.09 | 31,34,38,41 | 0 |
| 4 | U5L | D | 301 | 30/30 | 0.97 | 0.10 | 32,35,39,42 | 0 |
| 5 | MG | D | 307 | 1/1 | 0.97 | 0.06 | 44,44,44,44 | 0 |
| 5 | MG | B | 606 | 1/1 | 0.99 | 0.05 | 36,36,36,36 | 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 U5L B 605:

$2mF_o-DF_c$ (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.