



wwPDB EM Validation Summary Report ⓘ

Mar 9, 2025 – 12:03 am GMT

PDB ID : 9HDQ
EMDB ID : EMD-52072
Title : The Human LINE-1 ORF2p target-primed reverse transcription complex with the fingers domain in an open conformation
Authors : Ghanim, G.E.; Hu, H.; Nguyen, T.H.D.
Deposited on : 2024-11-12
Resolution : 2.45 Å(reported)
Based on initial model : .

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

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

A user guide is available at

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

EMDB validation analysis : **FAILED**
MolProbity : 4.02b-467
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.41

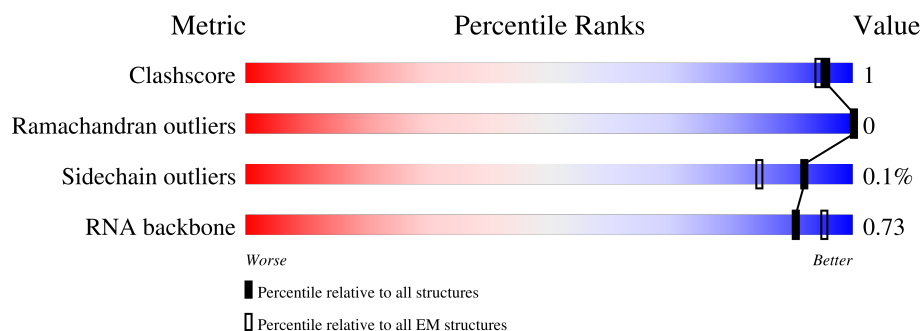
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 2.45 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) | EM structures (#Entries) |
|-----------------------|-----------------------------|-----------------------------|
| Clashscore | 210492 | 15764 |
| Ramachandran outliers | 207382 | 16835 |
| Sidechain outliers | 206894 | 16415 |
| RNA backbone | 6643 | 2191 |

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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 $\leq 5\%$.

| Mol | Chain | Length | Quality of chain |
|-----|-------|--------|--|
| 1 | A | 1797 | <div> <div>55%</div> <div>.</div> <div>43%</div> </div> |
| 2 | B | 37 | <div> <div>38%</div> <div>8%</div> <div>54%</div> </div> |
| 3 | C | 30 | <div> <div>33%</div> <div>67%</div> </div> |
| 4 | D | 30 | <div> <div>23%</div> <div>7%</div> <div>70%</div> </div> |
| 5 | E | 26 | <div> <div>100%</div> </div> |
| 6 | F | 19 | <div> <div>100%</div> </div> |
| 7 | G | 4 | <div> <div>100%</div> </div> |

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 10261 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 LINE-1 retrotransposable element ORF2 protein.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|------|------|------|----|---------|-------|
| | | | Total | C | N | O | S | | |
| 1 | A | 1025 | 8505 | 5501 | 1460 | 1505 | 39 | 0 | 0 |

There are 79 discrepancies between the modelled and reference sequences:

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|-----------------------|------------|
| A | -521 | MET | - | initiating methionine | UNP P0AEX9 |
| A | -520 | VAL | - | expression tag | UNP P0AEX9 |
| A | -519 | HIS | - | expression tag | UNP P0AEX9 |
| A | -518 | HIS | - | expression tag | UNP P0AEX9 |
| A | -517 | HIS | - | expression tag | UNP P0AEX9 |
| A | -516 | HIS | - | expression tag | UNP P0AEX9 |
| A | -515 | HIS | - | expression tag | UNP P0AEX9 |
| A | -514 | HIS | - | expression tag | UNP P0AEX9 |
| A | -513 | HIS | - | expression tag | UNP P0AEX9 |
| A | -512 | HIS | - | expression tag | UNP P0AEX9 |
| A | -511 | GLY | - | expression tag | UNP P0AEX9 |
| A | -510 | GLY | - | expression tag | UNP P0AEX9 |
| A | -509 | SER | - | expression tag | UNP P0AEX9 |
| A | -508 | ALA | - | expression tag | UNP P0AEX9 |
| A | -507 | TRP | - | expression tag | UNP P0AEX9 |
| A | -506 | SER | - | expression tag | UNP P0AEX9 |
| A | -505 | HIS | - | expression tag | UNP P0AEX9 |
| A | -504 | PRO | - | expression tag | UNP P0AEX9 |
| A | -503 | GLN | - | expression tag | UNP P0AEX9 |
| A | -502 | PHE | - | expression tag | UNP P0AEX9 |
| A | -501 | GLU | - | expression tag | UNP P0AEX9 |
| A | -500 | LYS | - | expression tag | UNP P0AEX9 |
| A | -499 | GLY | - | expression tag | UNP P0AEX9 |
| A | -498 | GLY | - | expression tag | UNP P0AEX9 |
| A | -497 | GLY | - | expression tag | UNP P0AEX9 |
| A | -496 | SER | - | expression tag | UNP P0AEX9 |
| A | -495 | GLY | - | expression tag | UNP P0AEX9 |
| A | -494 | GLY | - | expression tag | UNP P0AEX9 |

Continued on next page...

Continued from previous page...

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|----------------|------------|
| A | -493 | GLY | - | expression tag | UNP P0AEX9 |
| A | -492 | SER | - | expression tag | UNP P0AEX9 |
| A | -491 | GLY | - | expression tag | UNP P0AEX9 |
| A | -490 | GLY | - | expression tag | UNP P0AEX9 |
| A | -489 | SER | - | expression tag | UNP P0AEX9 |
| A | -488 | ALA | - | expression tag | UNP P0AEX9 |
| A | -487 | TRP | - | expression tag | UNP P0AEX9 |
| A | -486 | SER | - | expression tag | UNP P0AEX9 |
| A | -485 | HIS | - | expression tag | UNP P0AEX9 |
| A | -484 | PRO | - | expression tag | UNP P0AEX9 |
| A | -483 | GLN | - | expression tag | UNP P0AEX9 |
| A | -482 | PHE | - | expression tag | UNP P0AEX9 |
| A | -481 | GLU | - | expression tag | UNP P0AEX9 |
| A | -480 | LYS | - | expression tag | UNP P0AEX9 |
| A | -479 | GLY | - | expression tag | UNP P0AEX9 |
| A | -478 | GLY | - | expression tag | UNP P0AEX9 |
| A | -477 | SER | - | expression tag | UNP P0AEX9 |
| A | -167 | VAL | ALA | conflict | UNP P0AEX9 |
| A | -112 | ASN | - | linker | UNP P0AEX9 |
| A | -111 | SER | - | linker | UNP P0AEX9 |
| A | -110 | SER | - | linker | UNP P0AEX9 |
| A | -109 | SER | - | linker | UNP P0AEX9 |
| A | -108 | ASN | - | linker | UNP P0AEX9 |
| A | -107 | ASN | - | linker | UNP P0AEX9 |
| A | -106 | ASN | - | linker | UNP P0AEX9 |
| A | -105 | ASN | - | linker | UNP P0AEX9 |
| A | -104 | ASN | - | linker | UNP P0AEX9 |
| A | -103 | ASN | - | linker | UNP P0AEX9 |
| A | -102 | ASN | - | linker | UNP P0AEX9 |
| A | -101 | ASN | - | linker | UNP P0AEX9 |
| A | -100 | ASN | - | linker | UNP P0AEX9 |
| A | -99 | ASN | - | linker | UNP P0AEX9 |
| A | -98 | LEU | - | linker | UNP P0AEX9 |
| A | -97 | GLY | - | linker | UNP P0AEX9 |
| A | -34 | THR | ARG | conflict | UNP Q12306 |
| A | -27 | GLU | ARG | conflict | UNP Q12306 |
| A | 5 | THR | ASN | conflict | UNP O00370 |
| A | 15 | ILE | VAL | conflict | UNP O00370 |
| A | 21 | ALA | PRO | conflict | UNP O00370 |
| A | 152 | THR | ILE | conflict | UNP O00370 |
| A | 175 | ALA | THR | conflict | UNP O00370 |
| A | 595 | MET | VAL | conflict | UNP O00370 |

Continued on next page...

Continued from previous page...

| Chain | Residue | Modelled | Actual | Comment | Reference |
|-------|---------|----------|--------|----------|------------|
| A | 625 | THR | MET | conflict | UNP O00370 |
| A | 627 | PHE | LEU | conflict | UNP O00370 |
| A | 744 | THR | ASN | conflict | UNP O00370 |
| A | 795 | GLU | ASP | conflict | UNP O00370 |
| A | 932 | PHE | LEU | conflict | UNP O00370 |
| A | 1141 | LEU | VAL | conflict | UNP O00370 |
| A | 1182 | ASN | LYS | conflict | UNP O00370 |
| A | 1183 | GLU | ASP | conflict | UNP O00370 |
| A | 1212 | LYS | ASN | conflict | UNP O00370 |

- Molecule 2 is a DNA chain called Target DNA strand 1.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|-----|----|-----|----|---------|-------|
| 2 | B | 17 | Total | C | N | O | P | 0 | 0 |
| | | | 345 | 170 | 43 | 115 | 17 | | |

- Molecule 3 is a DNA chain called Target DNA strand 2.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|----|----|----|----|---------|-------|
| 3 | C | 10 | Total | C | N | O | P | 0 | 0 |
| | | | 205 | 98 | 43 | 54 | 10 | | |

- Molecule 4 is a RNA chain called Template P(A)30 RNA.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|----|----|----|---|---------|-------|
| 4 | D | 9 | Total | C | N | O | P | 0 | 0 |
| | | | 198 | 90 | 45 | 54 | 9 | | |

- Molecule 5 is a DNA chain called Target DNA strand 3.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|-----|----|-----|----|---------|-------|
| 5 | E | 26 | Total | C | N | O | P | 0 | 0 |
| | | | 524 | 260 | 64 | 174 | 26 | | |

- Molecule 6 is a DNA chain called Target DNA strand 4.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|-----|----|----|----|---------|-------|
| 6 | F | 19 | Total | C | N | O | P | 0 | 0 |
| | | | 399 | 190 | 95 | 95 | 19 | | |

- Molecule 7 is a DNA chain called Unassigned Nucleic Acid.

| Mol | Chain | Residues | Atoms | | | | | AltConf | Trace |
|-----|-------|----------|-------|----|----|----|---|---------|-------|
| 7 | G | 4 | Total | C | N | O | P | 0 | 0 |
| | | | 84 | 40 | 20 | 20 | 4 | | |

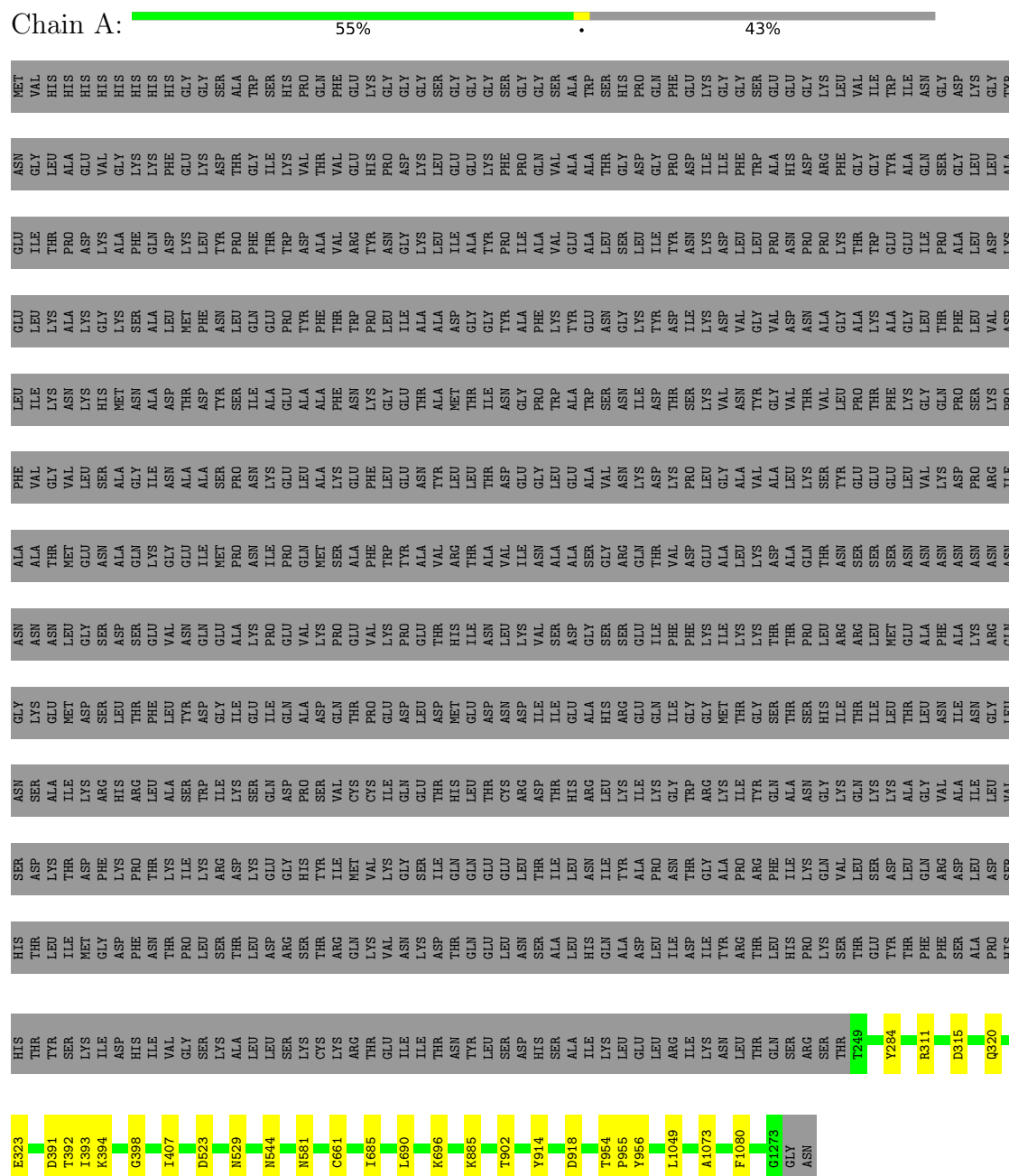
- Molecule 8 is ZINC ION (three-letter code: ZN) (formula: Zn).

| Mol | Chain | Residues | Atoms | | AltConf |
|-----|-------|----------|-------|----|---------|
| 8 | A | 1 | Total | Zn | 0 |
| | | | 1 | 1 | |

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. 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. 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: LINE-1 retrotransposable element ORF2 protein



There are no outlier residues recorded for this chain.

4 Experimental information

| Property | Value | Source |
|--------------------------------------|---|-----------|
| EM reconstruction method | SINGLE PARTICLE | Depositor |
| Imposed symmetry | POINT, Not provided | |
| Number of particles used | 183579 | Depositor |
| Resolution determination method | FSC 0.143 CUT-OFF | Depositor |
| CTF correction method | PHASE FLIPPING AND AMPLITUDE CORRECTION | Depositor |
| Microscope | TFS KRIOS | Depositor |
| Voltage (kV) | 300 | Depositor |
| Electron dose ($e^-/\text{\AA}^2$) | 60 | Depositor |
| Minimum defocus (nm) | 8000 | Depositor |
| Maximum defocus (nm) | 22000 | Depositor |
| Magnification | Not provided | |
| Image detector | TFS FALCON 4i (4k x 4k) | Depositor |

5 Model quality

5.1 Standard geometry

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

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

| Mol | Chain | Bond lengths | | Bond angles | |
|-----|-------|--------------|---------|-------------|---------|
| | | RMSZ | # Z >5 | RMSZ | # Z >5 |
| 1 | A | 0.30 | 0/8698 | 0.60 | 0/11724 |
| 2 | B | 0.61 | 0/381 | 1.09 | 0/587 |
| 3 | C | 0.63 | 0/231 | 0.99 | 0/353 |
| 4 | D | 0.52 | 0/224 | 1.04 | 0/347 |
| 5 | E | 0.65 | 0/579 | 1.10 | 0/892 |
| 6 | F | 0.62 | 0/455 | 1.01 | 0/699 |
| 7 | G | 0.74 | 0/95 | 1.04 | 0/144 |
| All | All | 0.38 | 0/10663 | 0.71 | 0/14746 |

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 | 8505 | 0 | 8793 | 15 | 0 |
| 2 | B | 345 | 0 | 202 | 2 | 0 |
| 3 | C | 205 | 0 | 112 | 0 | 0 |
| 4 | D | 198 | 0 | 100 | 1 | 0 |
| 5 | E | 524 | 0 | 309 | 0 | 0 |
| 6 | F | 399 | 0 | 210 | 0 | 0 |

Continued on next page...

Continued from previous page...

| Mol | Chain | Non-H | H(model) | H(added) | Clashes | Symm-Clashes |
|-----|-------|-------|----------|----------|---------|--------------|
| 7 | G | 84 | 0 | 45 | 0 | 0 |
| 8 | A | 1 | 0 | 0 | 0 | 0 |
| All | All | 10261 | 0 | 9771 | 18 | 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 1.

The worst 5 of 18 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:A:523:ASP:O | 1:A:529:ASN:ND2 | 2.36 | 0.56 |
| 2:B:19:DT:C2' | 2:B:20:DT:H71 | 2.42 | 0.49 |
| 1:A:320:GLN:HA | 1:A:323:GLU:OE2 | 2.13 | 0.49 |
| 1:A:914:TYR:O | 1:A:918:ASP:HB3 | 2.13 | 0.48 |
| 1:A:690:LEU:CD1 | 1:A:902:THR:HG23 | 2.44 | 0.48 |

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 PDB entries followed by that with respect to all EM entries.

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 | 1023/1797 (57%) | 1019 (100%) | 4 (0%) | 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 PDB entries followed by that with respect to all EM entries.

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 | 941/1593 (59%) | 940 (100%) | 1 (0%) | 92 | 96 |

All (1) residues with a non-rotameric sidechain are listed below:

| Mol | Chain | Res | Type |
|-----|-------|-----|------|
| 1 | A | 284 | TYR |

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

| Mol | Chain | Res | Type |
|-----|-------|------|------|
| 1 | A | 258 | ASN |
| 1 | A | 309 | GLN |
| 1 | A | 618 | ASN |
| 1 | A | 1055 | ASN |
| 1 | A | 1077 | ASN |

5.3.3 RNA [i](#)

| Mol | Chain | Analysed | Backbone Outliers | Pucker Outliers |
|-----|-------|------------|-------------------|-----------------|
| 4 | D | 8/30 (26%) | 0 | 0 |

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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.