# wwPDB/EMDataBank EM Map/Model Validation Summary Report (i)

Dec 19, 2016 – 06:51 PM EST

| | | |
|---|---|---|
| PDB ID | : | 5MD8 |
| EMDB ID: | : | EMD-3477 |
| Title | : | The structure of the mature HIV-1 CA hexameric lattice with curvature parameters: tilt=17, twist=12 |
| Authors | : | Mattei, S.; Glass, B.; Hagen, W.J.H.; Kraeusslich, H.-G.; Briggs, J.A.G. |
| Deposited on | : | 2016-11-10 |
| Resolution | : | 8.60 Å(reported) |

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
http://wwpdb.org/validation/2016/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

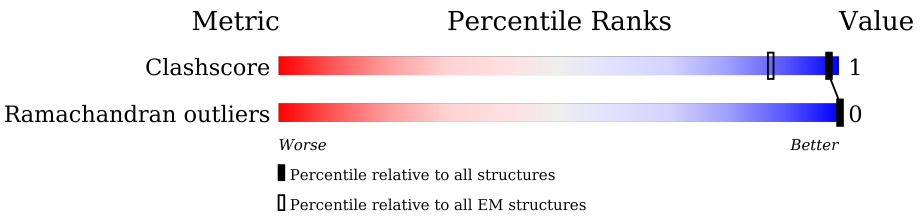
| | | |
|---|---|---|
| MolProbity | : | 4.02b-467 |
| Mogul | : | unknown |
| Percentile statistics | : | 20151230.v01 (using entries in the PDB archive December 30th 2015) |
| Ideal geometry (proteins) | : | Engh & Huber (2001) |
| Ideal geometry (DNA, RNA) | : | Parkinson et. al. (1996) |
| Validation Pipeline (wwPDB-VP) | : | rb-20028442 |

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
*ELECTRON MICROSCOPY*

The reported resolution of this entry is 8.60 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 | 114402 | 924 |
| Ramachandran outliers | 111179 | 726 |

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. 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%

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 1 | A | 74 | 99% . |
| 1 | B | 74 | 99% . |
| 1 | C | 74 | 99% . |
| 1 | D | 74 | 97% . |
| 1 | E | 74 | 97% . |
| 1 | F | 74 | 99% . |
| 1 | G | 74 | 99% . |
| 1 | W | 74 | 99% . |
| 2 | H | 147 | 98% . |
| 2 | I | 147 | 97% . |
| 2 | J | 147 | 97% . |

*Continued on next page...*

*Continued from previous page...*

| Mol | Chain | Length | Quality of chain |
|---|---|---|---|
| 2 | a | 147 | 98% |
| 2 | b | 147 | 98% |
| 2 | f | 147 | 98% |

# 2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5888 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 Gag protein.

| Mol | Chain | Residues | Atoms | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | | |
| 1 | A | 74 | 295 | 148 | 74 | 73 | 0 | 0 |
| 1 | B | 74 | 295 | 148 | 74 | 73 | 0 | 0 |
| 1 | F | 74 | 295 | 148 | 74 | 73 | 0 | 0 |
| 1 | W | 74 | 295 | 148 | 74 | 73 | 0 | 0 |
| 1 | C | 74 | 295 | 148 | 74 | 73 | 0 | 0 |
| 1 | D | 74 | 295 | 148 | 74 | 73 | 0 | 0 |
| 1 | E | 74 | 295 | 148 | 74 | 73 | 0 | 0 |
| 1 | G | 74 | 295 | 148 | 74 | 73 | 0 | 0 |

- Molecule 2 is a protein called Gag protein.

| Mol | Chain | Residues | Atoms | | | | AltConf | Trace |
|---|---|---|---|---|---|---|---|---|
| | | | Total | C | N | O | | |
| 2 | a | 147 | 588 | 294 | 147 | 147 | 0 | 0 |
| 2 | b | 147 | 588 | 294 | 147 | 147 | 0 | 0 |
| 2 | f | 147 | 588 | 294 | 147 | 147 | 0 | 0 |
| 2 | H | 147 | 588 | 294 | 147 | 147 | 0 | 0 |
| 2 | I | 147 | 588 | 294 | 147 | 147 | 0 | 0 |
| 2 | J | 147 | 588 | 294 | 147 | 147 | 0 | 0 |

# 3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Gag protein

Chain A: 99%

- Molecule 1: Gag protein

Chain B: 99%

- Molecule 1: Gag protein

Chain F: 99%

- Molecule 1: Gag protein

Chain W: 99%

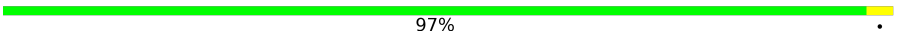
- Molecule 1: Gag protein

Chain C: 99%

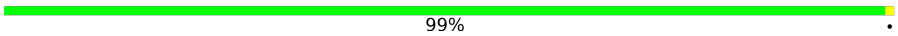
- Molecule 1: Gag protein

Chain D: 97%

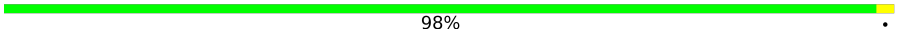
- Molecule 1: Gag protein

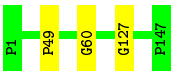
Chain E: 

- Molecule 1: Gag protein

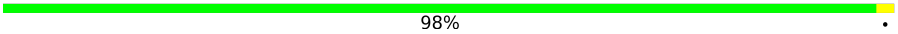
Chain G: 

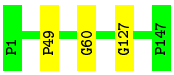
- Molecule 2: Gag protein

Chain a: 

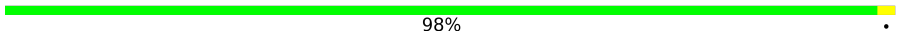




- Molecule 2: Gag protein

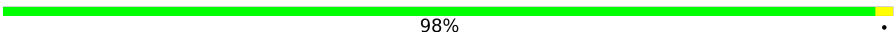
Chain b: 

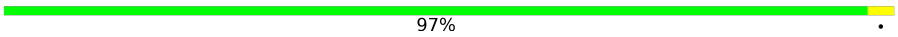

- Molecule 2: Gag protein

Chain f: 

- Molecule 2: Gag protein

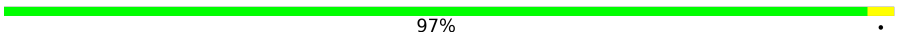
Chain H: 

- Molecule 2: Gag protein

Chain I: 

- Molecule 2: Gag protein

Chain J:



# 4 Experimental information (i)

| Property | Value | Source |
|---|---|---|
| Reconstruction method | SUBTOMOGRAM AVERAGING | Depositor |
| Imposed symmetry | POINT, Not provided | Depositor |
| Number of subtomograms used | 9273 | Depositor |
| Resolution determination method | FSC 0.143 CUT-OFF | Depositor |
| CTF correction method | Not provided | Depositor |
| Microscope | FEI TITAN KRIOS | Depositor |
| Voltage (kV) | 300 | Depositor |
| Electron dose ($e^-/Å^2$) | Not provided | Depositor |
| Minimum defocus (nm) | 2000 | Depositor |
| Maximum defocus (nm) | 6500 | Depositor |
| Magnification | 81000 | Depositor |
| Image detector | Not provided | Depositor |

# 5 Model quality (i)

## 5.1 Standard geometry (i)

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 | #$\|Z\|$ >2 | RMSZ | #$\|Z\|$ >2 |
| 1 | A | 1.35 | 0/294 | 1.55 | 1/366 (0.3%) |
| 1 | B | 1.35 | 0/294 | 1.54 | 1/366 (0.3%) |
| 1 | C | 1.35 | 0/294 | 1.54 | 1/366 (0.3%) |
| 1 | D | 1.35 | 0/294 | 1.54 | 1/366 (0.3%) |
| 1 | E | 1.35 | 0/294 | 1.54 | 1/366 (0.3%) |
| 1 | F | 1.35 | 0/294 | 1.54 | 1/366 (0.3%) |
| 1 | G | 1.36 | 0/294 | 1.55 | 1/366 (0.3%) |
| 1 | W | 1.36 | 0/294 | 1.55 | 1/366 (0.3%) |
| 2 | H | 1.49 | 3/587 (0.5%) | 1.48 | 0/732 |
| 2 | I | 1.49 | 3/587 (0.5%) | 1.48 | 0/732 |
| 2 | J | 1.49 | 3/587 (0.5%) | 1.49 | 0/732 |
| 2 | a | 1.49 | 3/587 (0.5%) | 1.48 | 0/732 |
| 2 | b | 1.49 | 3/587 (0.5%) | 1.48 | 0/732 |
| 2 | f | 1.49 | 3/587 (0.5%) | 1.48 | 0/732 |
| All | All | 1.43 | 18/5874 (0.3%) | 1.51 | 8/7320 (0.1%) |

The worst 5 of 18 bond length outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed(Å) | Ideal(Å) |
|---|---|---|---|---|---|---|---|
| 2 | f | 60 | GLY | CA-C | -6.63 | 1.41 | 1.51 |
| 2 | a | 60 | GLY | CA-C | -6.63 | 1.41 | 1.51 |
| 2 | I | 60 | GLY | CA-C | -6.57 | 1.41 | 1.51 |
| 2 | J | 60 | GLY | CA-C | -6.55 | 1.41 | 1.51 |
| 2 | H | 60 | GLY | CA-C | -6.51 | 1.41 | 1.51 |

The worst 5 of 8 bond angle outliers are listed below:

| Mol | Chain | Res | Type | Atoms | Z | Observed($^o$) | Ideal($^o$) |
|---|---|---|---|---|---|---|---|
| 1 | A | 218 | CYS | O-C-N | -5.78 | 113.46 | 122.70 |
| 1 | C | 218 | CYS | O-C-N | -5.78 | 113.46 | 122.70 |
| 1 | B | 218 | CYS | O-C-N | -5.76 | 113.49 | 122.70 |
| 1 | D | 218 | CYS | O-C-N | -5.76 | 113.49 | 122.70 |
| 1 | W | 218 | CYS | O-C-N | -5.70 | 113.58 | 122.70 |

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 | 295 | 0 | 77 | 0 | 0 |
| 1 | B | 295 | 0 | 77 | 0 | 0 |
| 1 | C | 295 | 0 | 77 | 0 | 0 |
| 1 | D | 295 | 0 | 77 | 1 | 0 |
| 1 | E | 295 | 0 | 77 | 1 | 0 |
| 1 | F | 295 | 0 | 77 | 0 | 0 |
| 1 | G | 295 | 0 | 77 | 0 | 0 |
| 1 | W | 295 | 0 | 77 | 0 | 0 |
| 2 | H | 588 | 0 | 156 | 0 | 0 |
| 2 | I | 588 | 0 | 156 | 1 | 0 |
| 2 | J | 588 | 0 | 156 | 1 | 0 |
| 2 | a | 588 | 0 | 156 | 0 | 0 |
| 2 | b | 588 | 0 | 156 | 0 | 0 |
| 2 | f | 588 | 0 | 156 | 0 | 0 |
| All | All | 5888 | 0 | 1552 | 2 | 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.

All (2) 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:D:148:THR:N | 2:I:147:PRO:C | 1.76 | 1.35 |
| 1:E:148:THR:N | 2:J:147:PRO:C | 2.00 | 1.15 |

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 | 72/74 (97%) | 70 (97%) | 2 (3%) | 0 | 100 | 100 |
| 1 | B | 72/74 (97%) | 70 (97%) | 2 (3%) | 0 | 100 | 100 |
| 1 | C | 72/74 (97%) | 70 (97%) | 2 (3%) | 0 | 100 | 100 |
| 1 | D | 72/74 (97%) | 70 (97%) | 2 (3%) | 0 | 100 | 100 |
| 1 | E | 72/74 (97%) | 70 (97%) | 2 (3%) | 0 | 100 | 100 |
| 1 | F | 72/74 (97%) | 70 (97%) | 2 (3%) | 0 | 100 | 100 |
| 1 | G | 72/74 (97%) | 70 (97%) | 2 (3%) | 0 | 100 | 100 |
| 1 | W | 72/74 (97%) | 70 (97%) | 2 (3%) | 0 | 100 | 100 |
| 2 | H | 145/147 (99%) | 141 (97%) | 4 (3%) | 0 | 100 | 100 |
| 2 | I | 145/147 (99%) | 142 (98%) | 3 (2%) | 0 | 100 | 100 |
| 2 | J | 145/147 (99%) | 142 (98%) | 3 (2%) | 0 | 100 | 100 |
| 2 | a | 145/147 (99%) | 142 (98%) | 3 (2%) | 0 | 100 | 100 |
| 2 | b | 145/147 (99%) | 142 (98%) | 3 (2%) | 0 | 100 | 100 |
| 2 | f | 145/147 (99%) | 142 (98%) | 3 (2%) | 0 | 100 | 100 |
| All | All | 1446/1474 (98%) | 1411 (98%) | 35 (2%) | 0 | 100 | 100 |

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains (i)

There are no protein residues with a non-rotameric sidechain to report in this entry.

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry (i)

There are no ligands in this entry.

## 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.