wwPDB/EMDataBank EM Map/Model Validation Summary Report

Feb 27, 2018 – 05:02 pm GMT

PDB ID : 6BF7
EMDB ID: : EMD-7091
Title : Cryo-EM structure of human insulin degrading enzyme in complex with FAB H11-E heavy chain, FAB H11-E light chain
Authors : Liang, W.G.; Zhang, Z.; Bailey, L.J.; Kossiakoff, A.A.; Tan, Y.Z.; Wei, H.; Carragher, B.; Potter, S.C.; Tang, W.J.
Deposted on : 2017-10-26
Resolution : 6.50 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
with specific help available everywhere you see the symbol.

MolProbity : 4.02b-467
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30686
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 6.50 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.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Whole archive (#Entries)</th>
<th>EM structures (#Entries)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clashscore</td>
<td>136279</td>
<td>1886</td>
</tr>
<tr>
<td>Ramachandran outliers</td>
<td>132675</td>
<td>1663</td>
</tr>
<tr>
<td>Sidechain outliers</td>
<td>132484</td>
<td>1531</td>
</tr>
</tbody>
</table>

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%.
2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 21866 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 Insulin-degrading enzyme.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Residues</th>
<th>Atoms</th>
<th>AltConf</th>
<th>Trace</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>949</td>
<td>Total C N O S</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7730 4980 1299 1429 22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>B</td>
<td>940</td>
<td>Total C N O S</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7662 4939 1285 1417 21</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There are 22 discrepancies between the modelled and reference sequences:

<table>
<thead>
<tr>
<th>Chain</th>
<th>Residue</th>
<th>Modelled</th>
<th>Actual</th>
<th>Comment</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>110</td>
<td>LEU</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>A</td>
<td>171</td>
<td>SER</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>A</td>
<td>178</td>
<td>ALA</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>A</td>
<td>257</td>
<td>VAL</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>A</td>
<td>414</td>
<td>LEU</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>A</td>
<td>573</td>
<td>ASN</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>A</td>
<td>590</td>
<td>SER</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>A</td>
<td>789</td>
<td>SER</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>A</td>
<td>812</td>
<td>ALA</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>A</td>
<td>819</td>
<td>ALA</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>A</td>
<td>904</td>
<td>SER</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>B</td>
<td>110</td>
<td>LEU</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>B</td>
<td>171</td>
<td>SER</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>B</td>
<td>178</td>
<td>ALA</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>B</td>
<td>257</td>
<td>VAL</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>B</td>
<td>414</td>
<td>LEU</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>B</td>
<td>573</td>
<td>ASN</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>B</td>
<td>590</td>
<td>SER</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>B</td>
<td>789</td>
<td>SER</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>B</td>
<td>812</td>
<td>ALA</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>B</td>
<td>819</td>
<td>ALA</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
<tr>
<td>B</td>
<td>904</td>
<td>SER</td>
<td>CYS</td>
<td>engineered mutation</td>
<td>UNP P14735</td>
</tr>
</tbody>
</table>

- Molecule 2 is a protein called Fab H11-E heavy chain.
<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Residues</th>
<th>Atoms</th>
<th>AltConf</th>
<th>Trace</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>C</td>
<td>217</td>
<td>Total C N O S</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1614 1020 270 319 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>E</td>
<td>217</td>
<td>Total C N O S</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1614 1020 270 319 5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Molecule 3 is a protein called Fab H11-E light chain.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Residues</th>
<th>Atoms</th>
<th>AltConf</th>
<th>Trace</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>D</td>
<td>211</td>
<td>Total C N O S</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1623 1016 272 330 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>211</td>
<td>Total C N O S</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1623 1016 272 330 5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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: Insulin-degrading enzyme

Chain A:

Chain B:

- Molecule 1: Insulin-degrading enzyme
- Molecule 2: Fab H11-E heavy chain

Chain C:

- Molecule 2: Fab H11-E heavy chain

Chain E:

- Molecule 3: Fab H11-E light chain

Chain D:

- Molecule 3: Fab H11-E light chain

Chain F:
4 Experimental information

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reconstruction method</td>
<td>SINGLE PARTICLE</td>
<td>Depositor</td>
</tr>
<tr>
<td>Imposed symmetry</td>
<td>POINT, Not provided</td>
<td>Depositor</td>
</tr>
<tr>
<td>Number of particles used</td>
<td>16944</td>
<td>Depositor</td>
</tr>
<tr>
<td>Resolution determination method</td>
<td>FSC 0.143 CUT-OFF</td>
<td>Depositor</td>
</tr>
<tr>
<td>CTF correction method</td>
<td>PHASE FLIPPING AND AMPLITUDE CORRECTION</td>
<td>Depositor</td>
</tr>
<tr>
<td>Microscope</td>
<td>FEI TITAN KRIOS</td>
<td>Depositor</td>
</tr>
<tr>
<td>Voltage (kV)</td>
<td>300</td>
<td>Depositor</td>
</tr>
<tr>
<td>Electron dose ((e^-/\text{Å}^2))</td>
<td>7.9, 6.8</td>
<td>Depositor</td>
</tr>
<tr>
<td>Minimum defocus (nm)</td>
<td>940</td>
<td>Depositor</td>
</tr>
<tr>
<td>Maximum defocus (nm)</td>
<td>2200</td>
<td>Depositor</td>
</tr>
<tr>
<td>Magnification</td>
<td>46598</td>
<td>Depositor</td>
</tr>
<tr>
<td>Image detector</td>
<td>GATAN K2 SUMMIT (4k x 4k), GATAN K2 SUMMIT (4k x 4k)</td>
<td>Depositor</td>
</tr>
</tbody>
</table>
5 Model quality

5.1 Standard geometry

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).

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Bond lengths</th>
<th>Bond angles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>RMSZ</td>
<td>#</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>0.47</td>
<td>0/7924</td>
</tr>
<tr>
<td>1</td>
<td>B</td>
<td>0.45</td>
<td>0/7854</td>
</tr>
<tr>
<td>2</td>
<td>C</td>
<td>0.39</td>
<td>0/1653</td>
</tr>
<tr>
<td>2</td>
<td>E</td>
<td>0.39</td>
<td>0/1653</td>
</tr>
<tr>
<td>3</td>
<td>D</td>
<td>0.37</td>
<td>0/1658</td>
</tr>
<tr>
<td>3</td>
<td>E</td>
<td>0.39</td>
<td>0/1653</td>
</tr>
<tr>
<td>All</td>
<td>All</td>
<td>0.43</td>
<td>0/22400</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Non-H</th>
<th>H(model)</th>
<th>H(added)</th>
<th>Clashes</th>
<th>Symm-Clashes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>7730</td>
<td>0</td>
<td>7639</td>
<td>186</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>B</td>
<td>7662</td>
<td>0</td>
<td>7568</td>
<td>157</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>C</td>
<td>1614</td>
<td>0</td>
<td>1566</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>E</td>
<td>1614</td>
<td>0</td>
<td>1566</td>
<td>40</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>D</td>
<td>1623</td>
<td>0</td>
<td>1580</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>1623</td>
<td>0</td>
<td>1580</td>
<td>32</td>
<td>0</td>
</tr>
<tr>
<td>All</td>
<td>All</td>
<td>21866</td>
<td>0</td>
<td>21499</td>
<td>456</td>
<td>0</td>
</tr>
</tbody>
</table>

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 11.
The worst 5 of 456 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

<table>
<thead>
<tr>
<th>Atom-1</th>
<th>Atom-2</th>
<th>Interatomic distance (Å)</th>
<th>Clash overlap (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2:C:41:ARG:HB3</td>
<td>2:C:49:GLU:HB3</td>
<td>1.68</td>
<td>0.75</td>
</tr>
<tr>
<td>1:A:711:ARG:HH22</td>
<td>1:A:715:PHE:HB2</td>
<td>1.52</td>
<td>0.75</td>
</tr>
<tr>
<td>1:B:621:GLN:HB3</td>
<td>1:B:628:TYR:HB3</td>
<td>1.68</td>
<td>0.75</td>
</tr>
<tr>
<td>1:A:62:ARG:HG2</td>
<td>1:A:80:ASP:HB2</td>
<td>1.72</td>
<td>0.70</td>
</tr>
</tbody>
</table>

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 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.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Analysed</th>
<th>Favoured</th>
<th>Allowed</th>
<th>Outliers</th>
<th>Percentiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>945/966 (98%)</td>
<td>913 (97%)</td>
<td>32 (3%)</td>
<td>0</td>
<td>100 100</td>
</tr>
<tr>
<td>1</td>
<td>B</td>
<td>936/966 (97%)</td>
<td>902 (96%)</td>
<td>34 (4%)</td>
<td>0</td>
<td>100 100</td>
</tr>
<tr>
<td>2</td>
<td>C</td>
<td>213/218 (98%)</td>
<td>208 (98%)</td>
<td>5 (2%)</td>
<td>0</td>
<td>100 100</td>
</tr>
<tr>
<td>2</td>
<td>E</td>
<td>213/218 (98%)</td>
<td>204 (96%)</td>
<td>9 (4%)</td>
<td>0</td>
<td>100 100</td>
</tr>
<tr>
<td>3</td>
<td>D</td>
<td>209/211 (99%)</td>
<td>205 (98%)</td>
<td>4 (2%)</td>
<td>0</td>
<td>100 100</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>209/211 (99%)</td>
<td>206 (99%)</td>
<td>3 (1%)</td>
<td>0</td>
<td>100 100</td>
</tr>
<tr>
<td>All</td>
<td>All</td>
<td>2725/2790 (98%)</td>
<td>2638 (97%)</td>
<td>87 (3%)</td>
<td>0</td>
<td>100 100</td>
</tr>
</tbody>
</table>

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 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.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Analysed</th>
<th>Rotameric</th>
<th>Outliers</th>
<th>Percentiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>835/861 (97%)</td>
<td>830 (99%)</td>
<td>5 (1%)</td>
<td>87 93</td>
</tr>
<tr>
<td>1</td>
<td>B</td>
<td>827/861 (96%)</td>
<td>821 (99%)</td>
<td>6 (1%)</td>
<td>85 93</td>
</tr>
<tr>
<td>2</td>
<td>C</td>
<td>179/181 (99%)</td>
<td>178 (99%)</td>
<td>1 (1%)</td>
<td>87 93</td>
</tr>
<tr>
<td>2</td>
<td>E</td>
<td>179/181 (99%)</td>
<td>178 (99%)</td>
<td>1 (1%)</td>
<td>87 93</td>
</tr>
<tr>
<td>3</td>
<td>D</td>
<td>187/187 (100%)</td>
<td>187 (100%)</td>
<td>0</td>
<td>100 100</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>187/187 (100%)</td>
<td>187 (100%)</td>
<td>0</td>
<td>100 100</td>
</tr>
<tr>
<td>All</td>
<td>All</td>
<td>2394/2458 (97%)</td>
<td>2381 (100%)</td>
<td>13 (0%)</td>
<td>90 94</td>
</tr>
</tbody>
</table>

5 of 13 residues with a non-rotameric sidechain are listed below:

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Res</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>B</td>
<td>49</td>
<td>ARG</td>
</tr>
<tr>
<td>1</td>
<td>B</td>
<td>229</td>
<td>ARG</td>
</tr>
<tr>
<td>1</td>
<td>B</td>
<td>896</td>
<td>LYS</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>896</td>
<td>LYS</td>
</tr>
<tr>
<td>1</td>
<td>B</td>
<td>838</td>
<td>ARG</td>
</tr>
</tbody>
</table>

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Res</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>B</td>
<td>681</td>
<td>HIS</td>
</tr>
<tr>
<td>1</td>
<td>B</td>
<td>730</td>
<td>HIS</td>
</tr>
<tr>
<td>1</td>
<td>B</td>
<td>857</td>
<td>HIS</td>
</tr>
<tr>
<td>1</td>
<td>B</td>
<td>300</td>
<td>GLN</td>
</tr>
<tr>
<td>1</td>
<td>B</td>
<td>332</td>
<td>HIS</td>
</tr>
</tbody>
</table>

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

There are no carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.