Dec 6, 2020 – 01:24 am GMT

PDB ID : 5ABB
EMDB ID : EMD-2446
Title : Visualization of a polytopic membrane protein during SecY-mediated membrane insertion
Authors : Bischoff, L.; Wickles, S.; Berninghausen, O.; vanderSluis, E.; Beckmann, R.
Deposited on : 2015-08-05
Resolution : 8.00 Å (reported)

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

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

EMDB validation analysis : 0.0.0.dev61
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.15.1
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 8.00 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>158937</td>
<td>4297</td>
</tr>
<tr>
<td>Ramachandran outliers</td>
<td>154571</td>
<td>4023</td>
</tr>
</tbody>
</table>

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 $\geq 3$, 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 EM map (all-atom inclusion $< 40\%$). The numeric value is given above the bar.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Length</th>
<th>Quality of chain</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>443</td>
<td><img src="image" alt="Graph" /></td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>116</td>
<td><img src="image" alt="Graph" /></td>
</tr>
<tr>
<td>3</td>
<td>Z</td>
<td>69</td>
<td><img src="image" alt="Graph" /></td>
</tr>
</tbody>
</table>
2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2512 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 PROTEIN TRANSLOCASE SUBUNIT SECY.

<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>443</td>
<td>Total</td>
<td>1772</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td>886</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N</td>
<td>443</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>O</td>
<td>443</td>
<td>0</td>
</tr>
</tbody>
</table>

- Molecule 2 is a protein called PROTEIN TRANSLOCASE SUBUNIT SECE.

<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>B</td>
<td>116</td>
<td>Total</td>
<td>464</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td>232</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N</td>
<td>116</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>O</td>
<td>116</td>
<td>0</td>
</tr>
</tbody>
</table>

- Molecule 3 is a protein called GREEN-LIGHT ABSORBING PROTEORHODOPSIN.

<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>Z</td>
<td>69</td>
<td>Total</td>
<td>276</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td>138</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N</td>
<td>69</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>O</td>
<td>69</td>
<td>0</td>
</tr>
</tbody>
</table>

There are 4 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>Z</td>
<td>84</td>
<td>THR</td>
<td>-</td>
<td>expression tag</td>
<td>UNP Q9F7P4</td>
</tr>
<tr>
<td>Z</td>
<td>85</td>
<td>ALA</td>
<td>-</td>
<td>expression tag</td>
<td>UNP Q9F7P4</td>
</tr>
<tr>
<td>Z</td>
<td>86</td>
<td>GLY</td>
<td>-</td>
<td>expression tag</td>
<td>UNP Q9F7P4</td>
</tr>
<tr>
<td>Z</td>
<td>87</td>
<td>ARG</td>
<td>-</td>
<td>expression tag</td>
<td>UNP Q9F7P4</td>
</tr>
</tbody>
</table>
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 atom inclusion in map 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 diamond above a residue indicates a poor fit to the EM map for this residue (all-atom inclusion < 40%). 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: PROTEIN TRANSLOCASE SUBUNIT SECY

Chain A:

- Molecule 2: PROTEIN TRANSLOCASE SUBUNIT SECE

Chain B:
Molecule 3: GREEN-LIGHT ABSORBING PROTEORHODOPSIN

Chain Z:
4 Experimental information

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>EM reconstruction method</td>
<td>SINGLE PARTICLE</td>
<td>Depositor</td>
</tr>
<tr>
<td>Imposed symmetry</td>
<td>POINT, C1</td>
<td>Depositor</td>
</tr>
<tr>
<td>Number of particles used</td>
<td>47471</td>
<td>Depositor</td>
</tr>
<tr>
<td>Resolution determination method</td>
<td>Not provided</td>
<td>Depositor</td>
</tr>
<tr>
<td>CTF correction method</td>
<td>Not provided</td>
<td>Depositor</td>
</tr>
<tr>
<td>Microscope</td>
<td>FEI TITAN KRIOS</td>
<td>Depositor</td>
</tr>
<tr>
<td>Voltage (kV)</td>
<td>200</td>
<td>Depositor</td>
</tr>
<tr>
<td>Electron dose ($e^-/\AA^2$)</td>
<td>20</td>
<td>Depositor</td>
</tr>
<tr>
<td>Minimum defocus (nm)</td>
<td>1000</td>
<td>Depositor</td>
</tr>
<tr>
<td>Maximum defocus (nm)</td>
<td>3500</td>
<td>Depositor</td>
</tr>
<tr>
<td>Magnification</td>
<td>148721</td>
<td>Depositor</td>
</tr>
<tr>
<td>Image detector</td>
<td>TVIPS TEMCAM-F416 (4k x 4k)</td>
<td>Depositor</td>
</tr>
<tr>
<td>Maximum map value</td>
<td>1.706</td>
<td>Depositor</td>
</tr>
<tr>
<td>Minimum map value</td>
<td>-0.587</td>
<td>Depositor</td>
</tr>
<tr>
<td>Average map value</td>
<td>0.013</td>
<td>Depositor</td>
</tr>
<tr>
<td>Map value standard deviation</td>
<td>0.142</td>
<td>Depositor</td>
</tr>
<tr>
<td>Recommended contour level</td>
<td>0.2</td>
<td>Depositor</td>
</tr>
<tr>
<td>Map size (Å)</td>
<td>385.9952, 385.9952, 385.9952</td>
<td>wwPDB</td>
</tr>
<tr>
<td>Map dimensions</td>
<td>368, 368, 368</td>
<td>wwPDB</td>
</tr>
<tr>
<td>Map angles (°)</td>
<td>90.0, 90.0, 90.0</td>
<td>wwPDB</td>
</tr>
<tr>
<td>Pixel spacing (Å)</td>
<td>1.0489, 1.0489, 1.0489</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>0.57</td>
<td>0/1771</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>0.57</td>
<td>0/463</td>
</tr>
<tr>
<td>3</td>
<td>Z</td>
<td>0.56</td>
<td>0/275</td>
</tr>
<tr>
<td>All</td>
<td>All</td>
<td>0.57</td>
<td>0/2509</td>
</tr>
</tbody>
</table>

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>#Chirality outliers</th>
<th>#Planarity outliers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>Z</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>All</td>
<td>All</td>
<td>0</td>
<td>42</td>
</tr>
</tbody>
</table>

There are no bond length outliers.

All (2) bond angle outliers are listed below:

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Res</th>
<th>Type</th>
<th>Atoms</th>
<th>Z</th>
<th>Observed(°)</th>
<th>Ideal(°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>300</td>
<td>TRP</td>
<td>C-N-CA</td>
<td>7.20</td>
<td>139.70</td>
<td>121.70</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>305</td>
<td>THR</td>
<td>N-CA-C</td>
<td>5.27</td>
<td>125.23</td>
<td>111.00</td>
</tr>
</tbody>
</table>

There are no chirality outliers.

All (42) planarity outliers are listed below:

<table>
<thead>
<tr>
<th>Mol</th>
<th>Chain</th>
<th>Res</th>
<th>Type</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>115</td>
<td>LYS</td>
<td>Mainchain</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>16</td>
<td>LEU</td>
<td>Mainchain</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>183</td>
<td>ILE</td>
<td>Mainchain,Peptide</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>204</td>
<td>ALA</td>
<td>Mainchain</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>218</td>
<td>LEU</td>
<td>Mainchain</td>
</tr>
</tbody>
</table>

Continued on next page...
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.
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 0.

There are no clashes within the asymmetric unit.

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>441/443 (100%)</td>
<td>367 (83%)</td>
<td>49 (11%)</td>
<td>25 (6%)</td>
<td>1 18</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>114/116 (98%)</td>
<td>88 (77%)</td>
<td>14 (12%)</td>
<td>12 (10%)</td>
<td>0 8</td>
</tr>
<tr>
<td>3</td>
<td>Z</td>
<td>67/69 (97%)</td>
<td>52 (78%)</td>
<td>11 (16%)</td>
<td>4 (6%)</td>
<td>1 17</td>
</tr>
<tr>
<td>All</td>
<td>All</td>
<td>622/628 (99%)</td>
<td>507 (82%)</td>
<td>74 (12%)</td>
<td>41 (7%)</td>
<td>2 16</td>
</tr>
</tbody>
</table>

All (41) Ramachandran outliers 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>A</td>
<td>60</td>
<td>ILE</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>82</td>
<td>ILE</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>257</td>
<td>VAL</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>305</td>
<td>THR</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>312</td>
<td>PRO</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>35</td>
<td>LEU</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>48</td>
<td>VAL</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>63</td>
<td>THR</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>81</td>
<td>GLY</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>217</td>
<td>PHE</td>
</tr>
<tr>
<td>1</td>
<td>A</td>
<td>307</td>
<td>SER</td>
</tr>
</tbody>
</table>
### 5.3.2 Protein sidechains

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

### 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 monosaccharides 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.
6 Map visualisation

This section contains visualisations of the EMDB entry EMD-2446. These allow visual inspection of the internal detail of the map and identification of artifacts.

No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

6.1.1 Primary map

The images above show the map projected in three orthogonal directions.

6.2 Central slices

6.2.1 Primary map

X Index: 184   Y Index: 184   Z Index: 184
The images above show central slices of the map in three orthogonal directions.

6.3 Largest variance slices

6.3.1 Primary map

The images above show the largest variance slices of the map in three orthogonal directions.

6.4 Orthogonal surface views

6.4.1 Primary map

The images above show the 3D surface view of the map at the recommended contour level 0.2. These images, in conjunction with the slice images, may facilitate assessment of whether an appropriate contour level has been provided.
6.5 Mask visualisation

This section was not generated. No masks/segmentation were deposited.
7 Map analysis

This section contains the results of statistical analysis of the map.

7.1 Map-value distribution

The map-value distribution is plotted in 128 intervals along the x-axis. The y-axis is logarithmic. A spike in this graph at zero usually indicates that the volume has been masked.
7.2 Volume estimate

The volume at the recommended contour level is 2877 nm$^3$; this corresponds to an approximate mass of 2598 kDa.

The volume estimate graph shows how the enclosed volume varies with the contour level. The recommended contour level is shown as a vertical line and the intersection between the line and the curve gives the volume of the enclosed surface at the given level.
7.3 Rotationally averaged power spectrum

*Reported resolution corresponds to spatial frequency of 0.125 Å⁻¹
8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.
9 Map-model fit

This section contains information regarding the fit between EMDB map EMD-2446 and PDB model 5ABB. Per-residue inclusion information can be found in section 3 on page 4.

9.1 Map-model overlay

The images above show the 3D surface view of the map at the recommended contour level 0.2 at 50% transparency in yellow overlaid with a ribbon representation of the model coloured in blue. These images allow for the visual assessment of the quality of fit between the atomic model and the map.
9.2 Atom inclusion

At the recommended contour level, 23% of all backbone atoms, 23% of all non-hydrogen atoms, are inside the map.