From MOS1 to HsMAR-Ra, form C-ter to PEC by structure modelling

Jeanne Cambefort1,2 and Corinne Augé-Gouillou2
(1) GILC UMR 7292 – CNRS – Tours
(2) IMT EA 6306 – Université F. Rabelais – Tours
(jeanne.cambefort, auge)@univ-tours.fr

Abstract
MOS1, SETMAR, and HsMAR-Ra are three closely related mariner transposases. The PDB contains two crystals of MOS1, the C-ter and the paired-end complex (PEC), and two SETMAR C-ter structures. No HsMAR-Ra structure is available, thus preventing the improvement of specific inhibitors [1] using docking approaches. Our goal is thus to model the PEC of HsMAR-Ra. We have drawn several C-ter structures of HsMAR-Ra by homology modelling, thanks to Modeller. Our models were molecularly analyzed with the SAVES software. Our models will be discriminated thanks to molecular dynamics. Once we will get a C-ter HsMAR-Ra, we will model the PEC of HsMAR-Ra.

Transposition cycle of Mariner [2]

Transposons are DNA fragments that can jump from one place in a genome to another. They possess two inverted terminal repeats (ITR) at their extremities and encode transposases. MOS1 and HsMAR-Ra belong to the transpose superfamily.

History of SETMAR [3]
40-58 million years ago, an Hsmar1 transposon was inserted downstream of a preexisting set gene. Hence, HsMAR-Ra and SETMAR are closely related proteins.

Sequential and structural alignments

Sequences were aligned with ClustalW

<table>
<thead>
<tr>
<th>Alignment</th>
<th>C-ter</th>
<th>N-ter</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOS1 – HsMAR-Ra</td>
<td>40.09%</td>
<td>28.21%</td>
</tr>
<tr>
<td>SETMAR – HsMAR-Ra</td>
<td>92.04%</td>
<td></td>
</tr>
</tbody>
</table>

- In the C-ter region, the sequence identity (92.04%) between SETMAR and HsMAR-Ra is high, so the models that will be built by homology modelling are likely to be reliable.
- With a sequence identity of 28.21% in the N-ter regions of MOS1 and HsMAR-Ra, modelling is more hazardous. But according to literature, we may obtain acceptable predictions even if the sequences share only 25% sequence identity if our models are carefully checked [8].

Tertiary structures

Interesting structures available in the PDB

<table>
<thead>
<tr>
<th>Structure</th>
<th>Resolution</th>
<th>Ion</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3HOT</td>
<td>3.25 Å</td>
<td>Mn</td>
<td>MOS1 full protein with DNA, dimer [4]</td>
</tr>
<tr>
<td>3HOS</td>
<td>3.5 Å</td>
<td>Mg</td>
<td>MOS1 full protein with DNA, dimer [4]</td>
</tr>
<tr>
<td>2F7T</td>
<td>2.2 Å</td>
<td>Mg</td>
<td>MOS1 catalytic domain, monomer [5]</td>
</tr>
<tr>
<td>3K9J</td>
<td>1.9 Å</td>
<td>Ca</td>
<td>SETMAR catalytic domain, dimer [6]</td>
</tr>
<tr>
<td>3K9K</td>
<td>2.55 Å</td>
<td></td>
<td>SETMAR catalytic domain, dimer [6]</td>
</tr>
</tbody>
</table>

Homology modelling of HsMAR-Ra C-ter with Modeller

Models were built with Modeller [10] in 3 different ways:
- Templates: 3HOT and 3K9J => 250 generated dimers
- Templates: 3HOT, 3HOS, 3K9J, and 3K9K => 250 generated dimers
- Templates: 3HOT, 3K9J, and 2F7T => 250 generated dimers

In the Modeller output, each structure is given with Molpdf and Dope scores. Then, in each case, the 10 models with minimum Molpdf scores and 10 models with minimum Dope scores were selected. The selected structures were evaluated in employing Verrity 3D [11] and Errat [12] tools. In this table, for each of 3 cases, the 3 best structures are presented. On the right side, the picture shows the 3 best structures (one in each case), they are aligned together with 3K9J.

<table>
<thead>
<tr>
<th>Model</th>
<th>Molpdf</th>
<th>Dope</th>
<th>Verrity 3D</th>
<th>Errat</th>
</tr>
</thead>
<tbody>
<tr>
<td>HsMAR-Ra.B106</td>
<td>29800.96</td>
<td>-48301.49</td>
<td>82.58</td>
<td>47.36</td>
</tr>
<tr>
<td>HsMAR-Ra.B191</td>
<td>29882.23</td>
<td>-48164.26</td>
<td>86.65</td>
<td>52.85</td>
</tr>
<tr>
<td>HsMAR-Ra.B223</td>
<td>30401.80</td>
<td>-46958.75</td>
<td>80.43</td>
<td>49.78</td>
</tr>
<tr>
<td>HsMAR-Ra.B040</td>
<td>35490.80</td>
<td>-48144.66</td>
<td>86.90</td>
<td>56.02</td>
</tr>
<tr>
<td>HsMAR-Ra.B042</td>
<td>36074.73</td>
<td>-47962.40</td>
<td>86.47</td>
<td>55.80</td>
</tr>
<tr>
<td>HsMAR-Ra.B203</td>
<td>37518.50</td>
<td>-48280.27</td>
<td>81.12</td>
<td>54.70</td>
</tr>
<tr>
<td>HsMAR-Ra.B012</td>
<td>26529.78</td>
<td>-47801.69</td>
<td>84.09</td>
<td>50.66</td>
</tr>
<tr>
<td>HsMAR-Ra.B091</td>
<td>26027.36</td>
<td>-48424.06</td>
<td>87.74</td>
<td>52.19</td>
</tr>
<tr>
<td>HsMAR-Ra.B122</td>
<td>26760.80</td>
<td>-48030.00</td>
<td>84.73</td>
<td>53.95</td>
</tr>
</tbody>
</table>

Since these alignments are very good, we may use homology modelling to model the C-ter region of HsMAR-Ra.

These models need to be confirmed by the results of molecular dynamics studies.

Bibliography

Conclusion and future works
Our results and some molecular dynamics will lead us to model the C-ter structure of HsMAR-Ra. Molecular dynamics will be performed with Xplor [13] software in using the Generalized Born [14] solvent model. Then we will model the full PEC structure of HsMAR-Ra by checking carefully the N-ter region whose secondary structure is well known. Now, we are seeking a collaboration to introduce the DNA in the PEC of HsMAR-Ra, in order to achieve our goal that is modelling the PEC of HsMAR-Ra.