Overall structure and dimerization of MvaS

Since MvaS could not be crystallized in the apo form, the AcAc-CoA complex is discussed here. MvaS consists of 13 α-helices and 14 β-sheets arranged in a larger “upper” and a smaller “lower” region. The upper region is characterized by the common thiolase fold arranged in a typical internally duplicated αβαβα-sandwich structure\textsuperscript{[1]}. The lower region is built by unstructured loops, 3 α-helices and 4 β-sheets, a feature unique among HMGCS enzymes (S3A in Supporting Information).\textsuperscript{[2]}

In order to determine the oligomerization state of MvaS, we conducted blue native page and PDBePISA analysis. Here, blue native page revealed a size of approximately 86 kDa, confirming that \textit{M. xanthus} MvaS also is a dimer. PISA showed a buried surface area of 6859 Å\textsuperscript{2} of 30069 Å\textsuperscript{2} total surface and a solvation free energy gain of -13.4 kcal mol\textsuperscript{-1} for a dimer generated by crystal symmetry, corroborating the experimental findings.
Scheme S1. Reaction mechanism of HMGCS. The catalytic triad comprising cysteine-histidine-glutamate is colored brown. Scheme modified according to Shafqat et al. [2]
Figure S1. Protein crystals of different MvaS complexes. A: AcCoA complex after 5 days. B: AcCoA complex after 14 days. C: AcAc-CoA complex after 5 days and D: HMG-CoA complex after 3 days.
**Figure S2.** Blue-native page to determine the oligomerisation state of MvaS in solution. 7.5 µg protein solution was loaded onto the gel. The marker contains albumin (66 kDa), lactat dehydrogenase (140 kDa), catalase (232 kDa), ferritin (440 kDa) and thyroglobulin (669 kDa).
Figure S3. Overall structure of MvaS and binding mode of acetoacetyl-CoA. A: Topology of the monomer with secondary structure assignment. The structure is coloured by secondary structure elements with α-helices in red, β-sheets in yellow and loop regions in green. B: Surface representation of AcAc-CoA bound MvaS in which the CoA binding tunnel is shown. The overall surface charge distribution was calculated using PDB2PQR and contoured at ± 10kT/e.\[^{[3]}^[4]\] C: The typical CoA binding mode is shown for acetoacetyl-CoA as a representative of all the three CoA-derivatives discussed here. Protein residues are colored according to their elements and AcAc-CoA is shown in the ball-and-stick representation.
Figure S4. Sequence alignment of MvaS (mxHMGS) with human HMGCS (hHMGS; Uniprot-ID: Q01581), plant HMGCS from Brassica juncea (pHMGS; Uniprot-ID: Q9M6U3) and Gram-positive bacterial HMGCS from Staphylococcus aureus (saHMGS; Uniprot-ID: Q9FD87). The secondary structure elements correspond to MvaS. * indicates residues involved in CoA binding and + highlights residues involved in catalysis. The two most divergent structural elements are highlighted by a blue (loop region) and a red box (helical subdomain).
Table S1. Data collection statistics for MvaS.

Values in parentheses are for the highest resolution shell. All data sets were collected from single crystals.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>complex with Ac-CoA</th>
<th>complex 2 with Ac-CoA</th>
<th>complex with AcAc-CoA</th>
<th>Complex with HMG-CoA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å) / beamline</td>
<td>1.5418 / inhouse</td>
<td>1.0332 / DESY, P11</td>
<td>1.0332 / DESY, P11</td>
<td>0.9184 / BESSY, BL14.1</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>29 – 2.20 (2.27 – 2.20)</td>
<td>42 – 1.50 (1.53 – 1.50)</td>
<td>47 – 1.50 (1.53 – 1.50)</td>
<td>46 – 1.48 (1.51 – 1.48)</td>
</tr>
<tr>
<td>Space group</td>
<td>P6122</td>
<td>P6122</td>
<td>P6122</td>
<td>P6122</td>
</tr>
<tr>
<td>Unit cell parameters (Å)</td>
<td>73.3, 73.3, 277.4</td>
<td>73.2, 73.2, 276.9</td>
<td>73.5, 73.5, 277.6</td>
<td>73.4, 73.4, 277.6</td>
</tr>
<tr>
<td>Mosaicity (°)†</td>
<td>0.16</td>
<td>0.04</td>
<td>0.12</td>
<td>0.05</td>
</tr>
<tr>
<td>Total No. of measured reflections</td>
<td>956150 (69768)</td>
<td>1143902 (52866)</td>
<td>1280707 (55634)</td>
<td>802665 (38588)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>23581 (1981)</td>
<td>71372 (3464)</td>
<td>72387 (3487)</td>
<td>74847 (3607)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>40.5 (35.2)</td>
<td>16.0 (15.3)</td>
<td>17.7 (16.0)</td>
<td>10.7 (10.7)</td>
</tr>
<tr>
<td>Mean I/σ(I)</td>
<td>12.4 (4.5)</td>
<td>17.3 (2.4)</td>
<td>18.0 (2.6)</td>
<td>12.3 (2.3)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>100 (99.9)</td>
<td>99.9 (99.5)</td>
<td>100 (100)</td>
<td>100 (99.9)</td>
</tr>
<tr>
<td>R&lt;sub&gt;meas&lt;/sub&gt; (%)‡</td>
<td>48.8 (126.7)</td>
<td>13.8 (1.398)</td>
<td>10.7 (105.7)</td>
<td>17.3 (122.1)</td>
</tr>
<tr>
<td>R&lt;sub&gt;pim&lt;/sub&gt; (%)‡</td>
<td>10.2 (28.8)</td>
<td>4.6 (48.7)</td>
<td>2.5 (26.0)</td>
<td>7.1 (50.1)</td>
</tr>
</tbody>
</table>

† Mosaicity values reported by XDS.[5]
§ Values in parentheses refer to the highest resolution shell.
R<sub>meas</sub> = \( \frac{\Sigma_{hkl} (N/(N - 1))^{1/2} \Sigma_i | I_i(hkl) - \langle I(hkl) \rangle |}{\Sigma_{hkl} \Sigma_i I_i(hkl)} \), where N is the number of observations of the reflection with index hkl and I<sub>i</sub> is the intensity of its <sub>i</sub>th observation.
R<sub>pim</sub> = \( \frac{\Sigma_{hkl} (1/(N - 1))^{1/2} \Sigma_i | I_i(hkl) - \langle I(hkl) \rangle |}{\Sigma_{hkl} \Sigma_i I_i(hkl)} \).[6]
Table S2. Refinement statistics for MvaS.

Values in parentheses are for the highest resolution shell.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>complex with Ac-CoA</th>
<th>complex 2 with Ac-CoA</th>
<th>complex with AcAc-CoA</th>
<th>Complex with HMG-CoA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution range (Å)</td>
<td>28 – 2.20 (2.28 – 2.20)</td>
<td>37 – 1.50 (1.55 – 1.50)</td>
<td>46 – 1.50 (1.55 – 1.50)</td>
<td>46 – 1.48 (1.53 – 1.48)</td>
</tr>
<tr>
<td>$R_{	ext{cryst}}$ (%)</td>
<td>17.6 (20)</td>
<td>12.5 (15.9)</td>
<td>14.0 (18)</td>
<td>12.8 (17.3)</td>
</tr>
<tr>
<td>$R_{	ext{free}}$ (%)</td>
<td>22.0 (26)</td>
<td>15.7 (22.0)</td>
<td>16.9 (24)</td>
<td>15.7 (21.5)</td>
</tr>
<tr>
<td>No. of non-H atoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>3025</td>
<td>3102</td>
<td>3163</td>
<td>3136</td>
</tr>
<tr>
<td>Ion</td>
<td>5</td>
<td>12</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Ligand</td>
<td>54</td>
<td>118</td>
<td>79</td>
<td>82</td>
</tr>
<tr>
<td>Water</td>
<td>329</td>
<td>359</td>
<td>418</td>
<td>510</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bonds (Å)</td>
<td>0.002</td>
<td>0.010</td>
<td>0.008</td>
<td>0.008</td>
</tr>
<tr>
<td>Angles (°)</td>
<td>0.54</td>
<td>1.12</td>
<td>1.02</td>
<td>1.05</td>
</tr>
<tr>
<td>Average B factors (Å²)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>12.3</td>
<td>15</td>
<td>17</td>
<td>11</td>
</tr>
<tr>
<td>Ion</td>
<td>31.7</td>
<td>47</td>
<td>37</td>
<td>31</td>
</tr>
<tr>
<td>Ligand</td>
<td>18.0</td>
<td>24</td>
<td>24</td>
<td>17</td>
</tr>
<tr>
<td>Water</td>
<td>17.2</td>
<td>26</td>
<td>28</td>
<td>25</td>
</tr>
<tr>
<td>Ramachandran plot</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Favoured regions (%)</td>
<td>97.1</td>
<td>98.1</td>
<td>97.9</td>
<td>97.4</td>
</tr>
<tr>
<td>Outliers (%)</td>
<td>0.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>MolProbity score*</td>
<td>1.00</td>
<td>1.10</td>
<td>1.03</td>
<td>1.21</td>
</tr>
<tr>
<td>PDB entry code</td>
<td>5HWP</td>
<td>5HWR</td>
<td>5HWQ</td>
<td>5HWO</td>
</tr>
</tbody>
</table>

*As reported by MolProbity at http://molprobity.biochem.duke.edu/*
References