Subsystem: Thiamin biosynthesis

Dmitry Rodionov, Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow, Russia

Thiamin pyrophosphate (vitamin B1) is an essential cofactor for several important enzymes of the carbohydrate metabolism. Many microorganisms, as well as plants and fungi synthesize thiamin, but it is not produced by vertebrates. Thiamin monophosphate is formed by coupling of two independently synthesized moieties, hydroxymethylthiamin-PP (HMP-PP) and hydroxyethylthiazole-P (HET-P).

The HET moiety is biosynthesized in *Bacillus subtilis* and most other bacteria from DXP, Glycine, and cysteine in a complex oxidative condensation reaction [1]. This reaction requires five different proteins, ThiO, ThiG, ThiS, ThiF, and a cysteine desulfurase. Glycine oxidase ThiO catalyzes the oxidation of glycine to the corresponding glycine imine. Sulfur carrier protein adenylyl transferase ThiF catalyzes the adenylation of the carboxy-terminus of the sulfur carrier protein ThiS, and cysteine desulfurase catalyzes the transfer of sulfur from cysteine to the ThiS-acyl adenylate to give ThiS-thiocarboxylate. ThiG is the thiazole synthase and catalyzes formation of the thiazole from dehydroglycine, DXP, and ThiS-thiocarboxylate [2, 3, 4]. The thiazole moiety of thiamin in *E. coli* is derived from Tyrosine, cysteine, and DXP using another enzyme (ThiH) of yet unknown function instead of ThiO [5]. In contrast, the THI4 protein family has been shown to be involved in the thiazole synthesis in some eukaryotes [7]. The HET kinase ThiM is involved in the salvage of thiazole from the culture medium.

The conversion of 5-aminomidazole ribonucleotide (AIR) into HMP is a fascinating reaction of the thiamin biosynthetic pathway in bacteria and is probably the most complex unresolved rearrangement in primary metabolism. The thiC gene complements all HMP requiring mutants in *E.coli*, and *B.subtilis* [6]. In yeast and plants, the pyrimidine moiety of thiamin is synthesized using a distinct gene (THI5 in yeasts), and the initial substrates appear to be histidine and pyridoxol-P [8, 9]. HMP-P is phosphorylated by the bifunctional HMP kinase/HMP-P kinase ThiD.

The ThiE protein catalyzes the formation of thiamin-P via coupling of HMP-PP and HET-P moieties. The ThiN protein recently identified in some ThiE-lacking archaea and in *T. maritima* [10] could complement *E.coli thiE* mutant strain and thus presents a case of non-orthologous gene displacement of the thiamin-P synthase ThiE [11]. The thiamin-phosphate phosphatase activity is present in yeast but corresponding gene has not yet been characterized. Bacteria synthesize TPP via single phosphorylation of TP using the ThiL kinase, whereas eukaryotes use distinct pathway to form an active coenzyme TPP: hydrolysis of TP to free thiamin is followed by pyrophosphorylation [8].

Search for thiamin-specific regulatory elements (*THI* riboswitches) and analysis of operon structures identified a large number of new candidate thiamin-regulated genes, mostly transporters, in various bacterial genomes [10]. In particular, the thiamin transporter function was assigned to yuaJ in the *Bacillus/Clostridium* group and thiT in Archaea. Also previously unknown HMP and HET transporter functions were tentatively assigned to several genes including thiXYZ and ykoEDC (ABC-type HMP transporters), cytX (HMP permease), thiU and thiW (HET permeases). Identification of the predicted uptake systems for thiamin precursors, HET and HMP, allowed to reconstruct the thiamin pathways in organisms that are unable to synthesize HET and HMP *de novo* and are forced to uptake them via specific transport system [10].
Fig. 1. Thiamin biosynthesis and uptake; HET and HMP salvage pathways

**Committed (connecting) intermediates**

- **DXP**: 1-deoxy-D-xylulose 5-phosphate
- **Cys**: L-Cysteine
- **Tyr**: L-Tyrosine
- **Gly**: L-Glycine
- **Gly-I**: L-Glycine imine
- **AIR**: Aminoimidazole ribotide
- **HMP**: 4-Amino-5-hydroxymethyl-2-methylpyrimidine
- **HMP-P**: 4-Amino-2-methyl-5-phosphonomethylpyrimidine
- **HMP-PP**: 4-amino-2-methyl-5-diphosphonomethylpyrimidine
- **HET**: 5-(2-Hydroxyethyl)-4-methylthiazole
- **HET-P**: 4-Methyl-5-(2-phosphoethyl)-thiazole
- **Thi**: Thiamin
- **Thi-PP**: Thiamin diphosphate
- **Thi-P**: Thiamin monophosphate

**Protein and Enzyme Details**

- **ThiG**: Thiazole biosynthesis protein ThiG
- **ThiH**: Thiazole biosynthesis protein ThiH
- **ThiI**: Glycine oxidase ThiI (EC 1.4.3.19)
- **ThiS**: Sulfur carrier protein ThiS
- **ThiF**: Sulfur carrier protein adenyltransferase ThiF
- **THI4**: Thiazole biosynthetic enzyme Thi4
- **ThiC**: Thiamine biosynthesis protein ThiC
- **THI5**: Hydroxymethylpyrimidine synthesis protein THI5
- **ThiM**: Hydroxethylthiazole kinase (EC 2.7.1.50)
- **ThiD**: Phosphonomethylpyrimidine kinase (EC 2.7.4.7)
- **ThiN**: Thiamin-phosphate synthase ThiN (EC 2.5.1.3)
- **ThiE**: Thiamin-phosphate pyrophosphorylase (EC 2.5.1.3)
- **TenA**: Thiaminase II (EC 3.5.99.2)
- **TPK**: Thiamine-monophosphate kinase (EC)
- **TPPK**: Thiamin pyrophosphokinase (EC 2.7.6.2)
- **TK**: Thiamine kinase (EC 2.7.1.89)
- **TP**: Thiamin-phosphate phosphatase (EC 3.1.3.-)
Fig. 2. Thiamin biosynthesis. Subsystem spreadsheet.

<table>
<thead>
<tr>
<th>Organism</th>
<th>Variant Code</th>
<th>HMP synthesis</th>
<th>HET synthesis</th>
<th>Thi-P synthesis</th>
<th>HMP salvage</th>
<th>HET salvage</th>
<th>Thiamin Uptake</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Shewanella oneidensis MR-1</em></td>
<td>1</td>
<td>2216-1</td>
<td>2212</td>
<td>2211-3</td>
<td>2215-11</td>
<td>2215</td>
<td>2079</td>
</tr>
<tr>
<td><em>Escherichia coli K12</em></td>
<td>11</td>
<td>3907-1</td>
<td>3903</td>
<td>3902-3</td>
<td>3906-11</td>
<td>2078</td>
<td>69-18, 65-19, 67-20</td>
</tr>
<tr>
<td><em>Mycobacterium tuberculosis CDC1551</em></td>
<td>2</td>
<td>438-1</td>
<td>431</td>
<td>429-4</td>
<td>426-11</td>
<td>437</td>
<td>5595-18, 5596-19, 5597-20</td>
</tr>
<tr>
<td><em>Sporosarcina억</em></td>
<td>3</td>
<td>2075-1</td>
<td>2073-4</td>
<td>2070-11</td>
<td>9516</td>
<td></td>
<td>5685-18, 5686-19, 5687-20</td>
</tr>
<tr>
<td><em>Thermotoga maritima MS38</em></td>
<td>3</td>
<td>781-1</td>
<td>780</td>
<td>783-9</td>
<td>478-14, 479-15, 477-16</td>
<td></td>
<td>227-18, 226-19, 225-20; 123-24</td>
</tr>
<tr>
<td><em>Pyrobaculum aerophilum str. IM2 [A]</em></td>
<td>33</td>
<td>229-1</td>
<td>122</td>
<td>1538-9, 1539-10</td>
<td>1770-9</td>
<td>1770</td>
<td>184</td>
</tr>
<tr>
<td><em>Neurospora erassa [E]</em></td>
<td>4</td>
<td>3970-36</td>
<td>3898</td>
<td>184-11</td>
<td>6888</td>
<td></td>
<td>184</td>
</tr>
<tr>
<td><em>Saccharomyces cerevisiae [E]</em></td>
<td>44</td>
<td>1690-38</td>
<td>3168-35</td>
<td>3168-35</td>
<td>2183</td>
<td>5452</td>
<td>3779-36, 5080-36, 5194-36</td>
</tr>
<tr>
<td><em>Methanococcus maripaludis S2 [A]</em></td>
<td>5</td>
<td>187-1</td>
<td>1352</td>
<td>1352-11</td>
<td>1352-11</td>
<td>1352-11</td>
<td>1352-21</td>
</tr>
<tr>
<td><em>Pyrococcus abyssi GE5 [A]</em></td>
<td>55</td>
<td>683-1</td>
<td>827</td>
<td>1140-9</td>
<td>1140-11</td>
<td>1140-11</td>
<td>1143-21</td>
</tr>
<tr>
<td><em>Archeaeoglobus fulgidus DSM 4304 [A]</em></td>
<td>6</td>
<td>2396-1</td>
<td>695</td>
<td>2069-11</td>
<td>2194</td>
<td>2069-11</td>
<td>2069-11</td>
</tr>
<tr>
<td><em>Haemophilus influenzae R2886</em></td>
<td>77</td>
<td>631-21</td>
<td>632-26</td>
<td>633-27</td>
<td>636</td>
<td>635-22</td>
<td>635-22</td>
</tr>
<tr>
<td><em>Bradyrhizobium japonicum USDA 110</em></td>
<td>8</td>
<td>6659-1</td>
<td>6657</td>
<td>6655-4</td>
<td>6658-11</td>
<td>5905</td>
<td>2637-14, 2836-15, 2836-16</td>
</tr>
<tr>
<td><em>Bacillus subtilis subsp. subtilis str. 168</em></td>
<td>88</td>
<td>870-1</td>
<td>1170</td>
<td>1168-4</td>
<td>1167-11, 3834-11, 3837-1</td>
<td></td>
<td>3102-17</td>
</tr>
<tr>
<td><em>Silicibacter pomeroyi DSS-3</em></td>
<td>9</td>
<td>48</td>
<td>46</td>
<td>49-11</td>
<td>45</td>
<td>51-14, 54-15</td>
<td>1864-17</td>
</tr>
<tr>
<td><em>Brucella melitensis 16M</em></td>
<td>99</td>
<td>1734</td>
<td>1732-4</td>
<td>1735-11</td>
<td>1731</td>
<td>1736-14, 1738-15</td>
<td>143-18, 142-19, 141-20</td>
</tr>
<tr>
<td><em>Treponema pallidum subsp. Pallidum</em></td>
<td>10</td>
<td>161</td>
<td>114</td>
<td>114</td>
<td></td>
<td></td>
<td>114</td>
</tr>
</tbody>
</table>

Functional variants:

#1: HMP synthesis (ThiC-1), HET synthesis (ThiG, ThiH-3), thiamin synthase (ThiE-11);

#2: HMP synthesis (ThiC-1), HET synthesis (ThiG, ThiO-4), thiamin synthase (ThiE-11);

#3: HMP synthesis (ThiC-1), HET synthesis (ThiI4), thiamin synthase (ThiN-9);

#4: HMP synthesis (ThiS-35), HET synthesis (ThiI4), thiamin synthase (ThiN-9);

#5: HMP synthesis (ThiC-1), HET synthesis (ThiI4), thiamin synthases (ThiE-11 and ThiN-9);

#6: HMP synthesis (ThiC-1), HET synthesis (ThiI4), thiamin synthase (ThiE-9);

#7: HMP and HET salvage, thiamin synthase (ThiE-11);

#8: HMP synthesis (ThiC-1), HET synthesis (ThiG, ThiO-4), thiamin synthase (ThiE-11), HMP salvage;

#9: HMP salvage, HET synthesis (ThiG, ThiO-4), thiamin synthase (ThiE-11);

#10: only Thiamin Uptake;

*Variants #11,22,33,44,55,77,89,99 are the same as #1,2,3,4,5,7,8,9, respectively, but also include additional genes for Thiamin Uptake.*
Analysis of regulatory elements. Predicted thiamin-regulated transporter genes [10]

1. **YuaJ**: predicted Thiamin Transporter (possibly H\(^+\)-dependent)
   - Found only in bacteria from the *Bacillus/Clostridium* group
   - All these species lack the known thiamin ABC transporter ThiBPQ;
   - Occurs in genomes without thiamin biosynthetic pathway (*Streptococcus* spp.);
   - Has 6 predicted transmembrane segments;
   - Regulated by thiamin riboswitches;
   - *Bacillus cereus* is able to uptake thiamin in a proton-dependent manner [12].

2. **ThiXYZ and YkoEDC**: predicted ATP-dependent HMP Transporters
   - Found in some *Proteobacteria* and *Firmicutes*;
   - Not found in genomes without thiamin biosynthetic pathway;
   - Always co-occur with the *thiD* and *thiE* genes;
   - Present in *Pasteurellae, Brucella* and *Gram-positive cocci*, lacking HMP synthase *ThiC*;
   - Regulated by thiamin riboswitches;
   - The substrate-binding component *ThiY* is homologous to the HMP synthase *THI5* from *Yeast*
     and has an N-terminal signal peptide for possible membrane binding.
References.


