Supporting Information

A Missing Enzyme in Thiamin Thiazole Biosynthesis: Identification of TenI as a Thiazole Tautomerase‡

Amrita B. Hazra, Ying Han[#], Abhishek Chatterjee[#], Yang Zhang[#], Rung-Yi Lai, Steven E. Ealick[#] and Tadhg P. Begley**

Department of Chemistry, Texas A&M University, College Station, TX 77843, [#]Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853

AUTHOR EMAIL ADDRESS: begley@chem.tamu.edu, see3@cornell.edu

RECEIVED DATE (to be automatically inserted after your manuscript is accepted if required according to the journal that you are submitting your paper to)

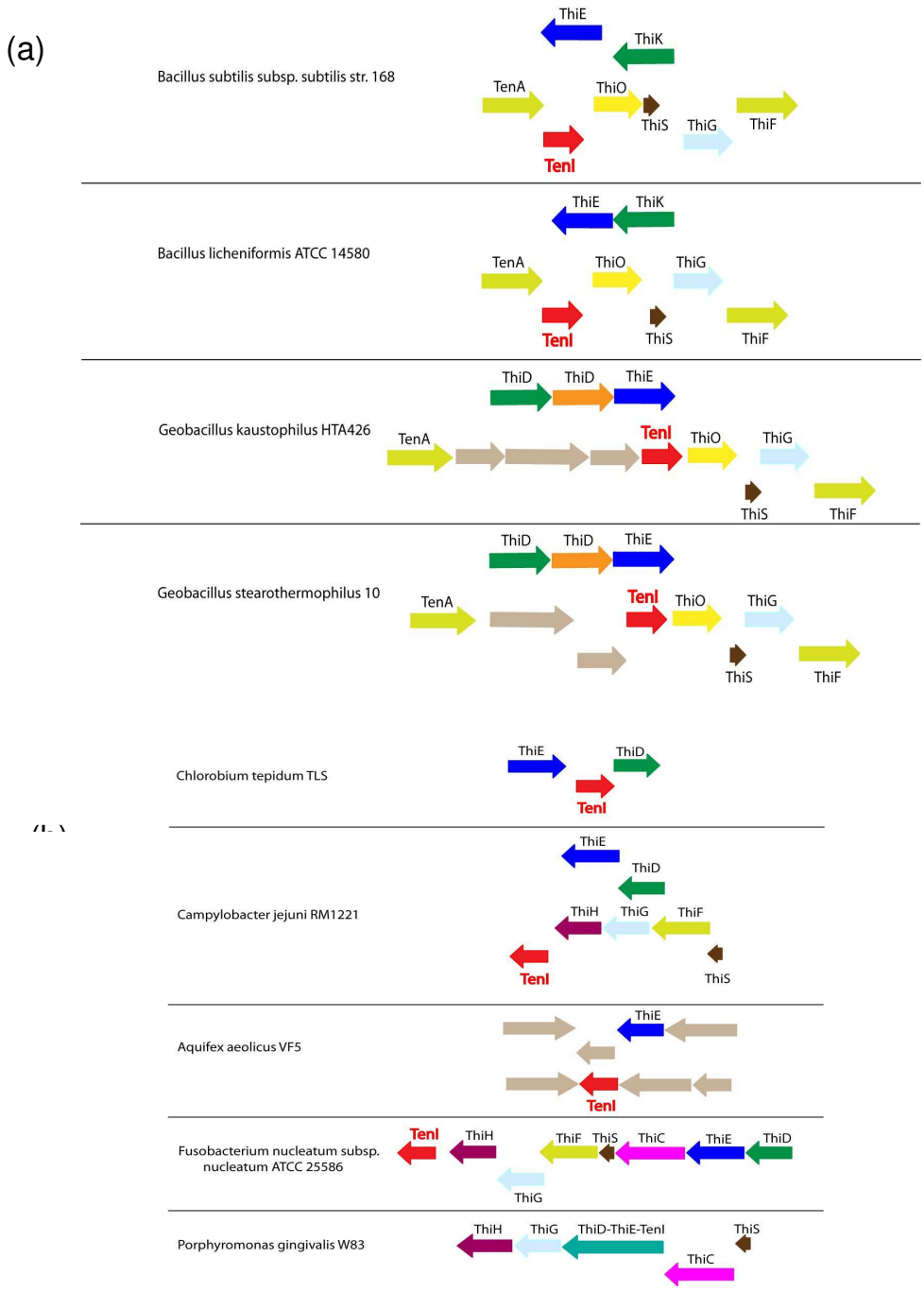

Figure 1: (a) Gene neighborhood of TenI showing clustering with TenA (thiaminase II) and other thiamin biosynthetic genes. (b) In some bacteria lacking TenA, TenI clusters with thiamin biosynthesis genes and sometimes even with ThiE (thiamin phosphate synthase), with which it has a high sequence similarity (http://theseed.uchicago.edu).

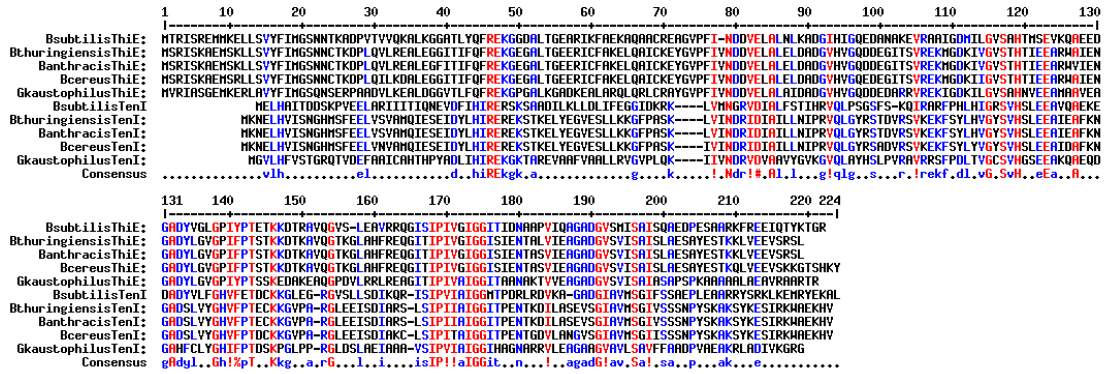

Figure 2: Primary sequence alignment of ThiE and TenI from five different bacterial species showing regions that are highly conserved between the two proteins.

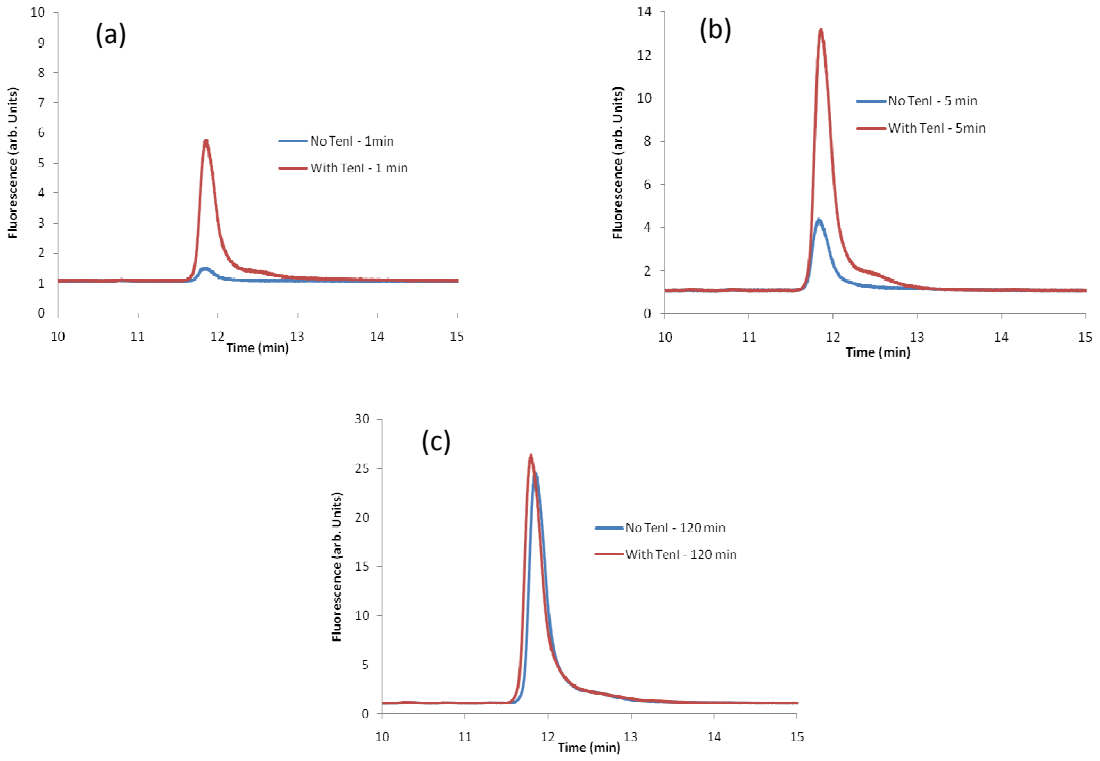

Figure 3: Time course for the thiazole reconstitution reaction in the presence and absence of TenI. (a), (b), (c) The red trace shows thiochrome formed in the presence of TenI and the blue trace shows the formation of thiochrome in the absence of TenI at different time points.

Preparation of cThz*-P 14 and cThz* 24 from adenylated cThz*-P 30: Overexpressed *S. cerevisiae* THI4p (thiazole synthase) from 4 L of culture (~200 mg in10 mL) was divided into twenty 500 µL aliquots and heat denatured (100 °C, 2minutes). The precipitated protein was removed by centrifugation and the supernatants were combined and filtered through a 10 kDa MW cut-off Microcon filter. Adenylated cThz*-P **30** was purified by HPLC using the following linear gradient at a flow rate of 3 mL/min: solvent A is water, solvent B is 100 mM KPi, pH 6.6, solvent C is methanol. 0 min: 100% B; 3 min: 10% A, 90%B; 17 min: 34% A, 60% B, 6% C; 21 min: 35% A, 25% B, 40% C; 23 min: 100%B and the collected fractions were pooled. A second HPLC purification, using a low concentration of ammonium acetate buffer, was performed on the pooled fractions using the following linear gradient at a flow rate of 3 mL/min: Solvent A is water, solvent B is 25 mM NH_4OAc, pH 6.6, solvent C is methanol. 0 min: 100% B; 2 min: 10% A, 90%B; 6 min: 15% A, 20% B, 65% C; 8 min: 15% A, 20% B, 65%; 11 min: 100% B; 14 min: 100%B. The collected fractions were then lyophilized to yield micromolar quantities of adenylated cThz*-P **26**. This was then treated with 1 unit nucleotide pyrophosphatase at pH 7.2 to yield **14** which was further treated with 1 unit calf intestinal phosphatase in phosphate buffer, pH 7.8 for 20 min. to yield **24** (Figure 4). The concentration of the cThz*-P was determined by comparing the amount of AMP released by nucleotide pyrophosphatase treatment with a standard quantity of AMP by HPLC analysis (Figure 5).

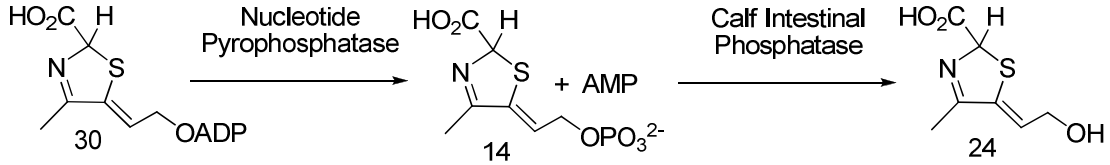

Figure 4: Conversion of adenylated cThz*-P **30** to cThz*-P **14** and cThz* **24**. **[**Chatterjee, A.; Jurgenson, C.T.; Schroeder, F.C.; Ealick, S.E.; Begley,T.P. *J. Am. Chem. Soc.* **2007**, 129, 2914-22]

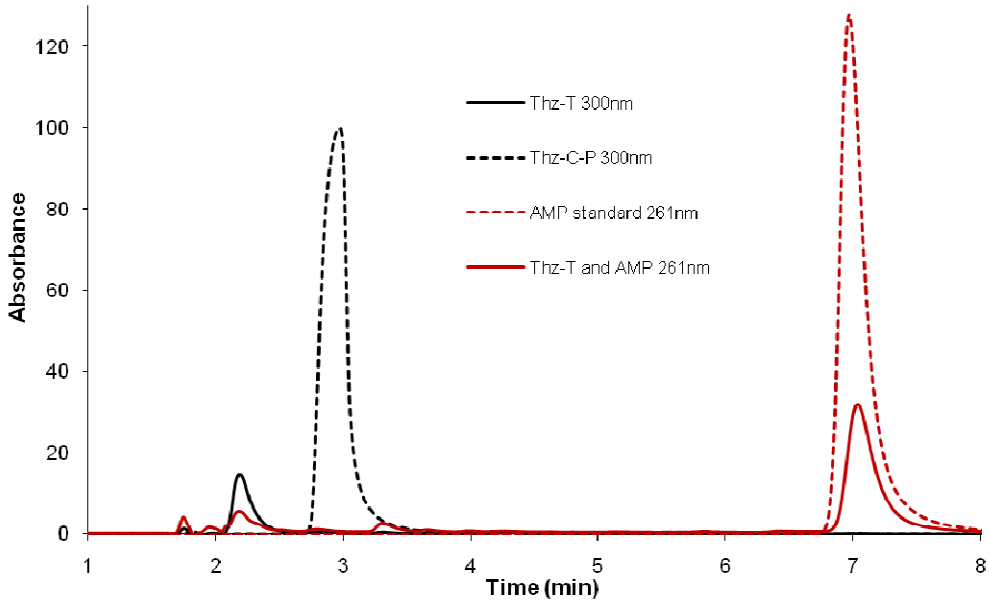

Figure 5: The concentration of cThz*-P is measured by comparing the amount of AMP released in the reaction shown in Figure 3 with a standard amount of AMP by HPLC analysis.

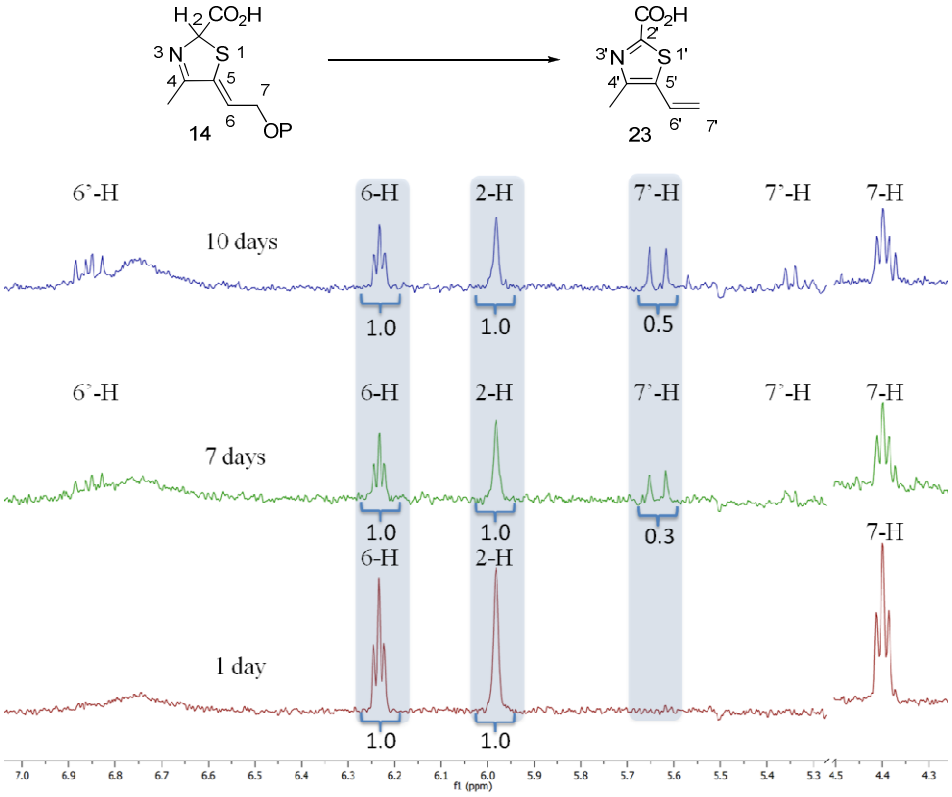

Figure 6: ^{1}H-1D-NMR analysis of the time course for cThz*-P degradation to form carboxyvinyl thiazole **23**.

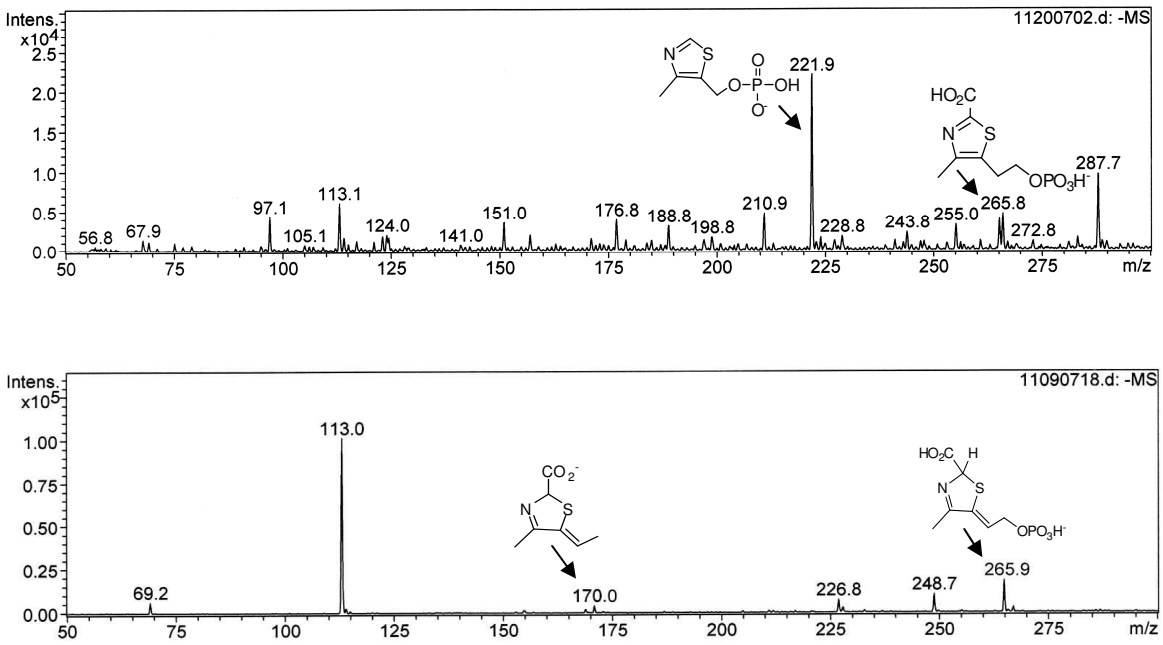

Figure 7: Negative mode ESI-MS of (a) cThz*-P and (b) cThz-P.

Table 1: TenI crystal structure parameters*

Data collection	
Space group	$C222_1$
Cell dimensions	
a, *b*, *c* (Å)	98.05, 105.42, 219.26
α, β, γ (°)	90, 90, 90
Resolution (Å)	50 – 2.23 (2.23 – 2.31)*
$I / \sigma I$	18.9 (4.8)
Completeness (%)	99.8 (100)
Redundancy	6.1 (6.4)
Refinement	
Resolution (Å)	50 – 2.23 (2.23 – 2.31)
No. reflections	55149
R_{work} / R_{free}	0.176 / 0.223
No. atoms	
Protein	6067
Ligand/ion	69
Water	302
B-factors	
Protein	24.3
Ligand/ion	32.2
Water	30.1
R.m.s deviations	
Bond lengths (Å)	0.011
Bond angles (°)	1.265

**Values in parentheses are for the highest resolution shell*