The Danger of Annotation by Analogy: Most “thiI” Genes Play No Role in Thiamine Biosynthesis

Robert A. Bender*

Department of Molecular, Cellular, and Developmental Biology, The University of Michigan, Ann Arbor, Michigan 48109

In the book of Genesis, God gave Adam the power to name all the creatures that he encountered. From the moment that he named the first apple “delicious,” he ran into trouble. It may have looked delicious, but until he actually did the experiment, he had no idea of its real properties. More recently, bioinformatics programs have given us the power to name all the genes whose sequences we encounter. A paper by Martinez-Gomez et al. in this issue (2) shows that inappropriate nomenclature can still get us into trouble. Until we do the experiment, we may have no idea of the real properties of a gene product. Specifically, they show that only the rhodanese domain of the ThiI protein is required for a key thiolation reaction in the synthesis of thiamine. The other two domains (THUMP and AANH) are dispensable. And yet, nearly three-quarters of all genes annotated as thiI are allelic to the previously characterized E. coli gene, nuvA (3). ThiS/NuvA is the enzyme responsible for the thiolation of a uridine residue in tRNA. Three domains of ThiI are essential for the thiolation of tRNA: a THUMP domain that binds tRNA, an AANH domain that activates the uridine residue by adenylylation (9), and a rhodanese domain that transfers sulfur to the activated uridine residue (4). To accomplish this transfer of sulfur, a sulfur transferase (IscS) moves the sulfhydryl group from free cysteine to a specific cysteine residue in the rhodanese domain (Cys456), generating a persulfide at this site (1). The ThiS-persulfide then attacks the adenylylated uridine residue, releasing the adenylyl group and forming a disulfide bond between Cys456 and the uridine residue. Then a cysteine within the AANH domain (Cys344) attacks the disulfide, leaving a thioridine on the tRNA and a disulfide between Cys456 and Cys344 (4). Finally, the disulfide between Cys456 and Cys344 is rereduced and the cycle is ready to repeat. Thus, the AANH domain plays two roles, adenylylating the uridine residue and breaking the disulfide bond between Cys456 of the rhodanese domain and the tRNA.

In thiamine synthesis, the ThIFs protein complex is the ultimate sulfur donor to thiazole synthase. The thiolation of ThIFs requires activation of a C-terminal glycine residue by adenylylation (6) and an IscS-dependent formation of a persulfide on Thil (5, 6). Thus, it seemed logical to assume that the thiolation chemistry in thiazole synthesis could be predicted by analogy to thiolation of tRNA. The AANH domain of Thil would adenylylate (activate) the Thif protein. Thiolation of Cys456 of Thil by IscS would generate the required persulfide. The persulfide would attack and release the adenylyl group, forming a disulfide bond between the Cys456 of Thil and the terminal glycine of ThiS. Then transfer of the disulfide bond from ThiS to the Cys344 of the AANH domain of Thil would leave a thiolated ThiFS and a ThiI with a disulfide that could be rereduced to restart the cycle.

Unfortunately, this prediction by analogy turns out to be incorrect. It had already been shown that ThiS is adenylylated by Thil, not by the AANH domain of Thil (8). So the only role for the AANH domain appeared to be the ability of Cys344 to accept the disulfide bond from Cys456. However, Martinez-Gomez et al. have now shown that neither the THUMP nor the AANH domain of Thil is essential. The rhodanese domain	

* Mailing address: Department of MCBiology, The University of Michigan, Ann Arbor, MI 48109-1048. Phone: (734) 975-1776, Fax: (734) 647-0884. E-mail: rbender@umich.edu.

† Published ahead of print on 8 July 2011.
alone can carry out the thiolation reaction in vivo. The disulfide between ThiI and ThiS is resolved when a cysteine from ThiS attacks and forms a disulfide link between ThiF and ThiS (8). The Cys344 of ThiI plays no role at all. Thus, the AANH domain is not involved in either the adenylylation reaction or the disulfide transfer.

This shines a bright light on a problem in the annotation of thiI genes in general. Most of the genes annotated as thiI contain two of the three domains of the E. coli enzyme (THUMP and AANH), but they lack the rhodanese domain, the only domain actually involved in thiamine biosynthesis. So, the work of Martinez-Gomez et al. shows a need to revise the annotation of most of the genes currently annotated as thiI. Equally important, their work also suggests some novel chemistry at work in thiazole synthesis. When a genetic approach failed to find a suppressor that could replace ThiI in thiazole synthesis, they looked for physiological conditions that might bypass the ThiI activity and found that the addition of exogenous cysteine to the growth medium allowed thiI mutants to grow. This suppression of the thiI defect was seen even in double mutants with iscS, gdhA, and sufS but did not bypass the need for ThiFS. Thus, there must exist a novel way to mobilize sulfur from cysteine into thiazole (and perhaps elsewhere). This demonstrates that even in organisms whose physiology is as well studied as the enteric bacteria, there are still mysteries to be found.

The work of the Downs lab has long argued that there is more to physiology than the core pathways on our metabolic charts. The “leaky reactions” and “lack of specificity” seen in many enzymes are not errors to be overlooked; they are the stuff of evolution and of survival. Their demonstration that there is a novel, cysteine-dependent sulfur transferase pathway available in S. enterica opens up a whole path for investigation. The work of Martinez-Gomez et al. cautions us all to be careful of the article that we use in describing an enzyme. It is always safer to say “a thiolase” than “the thiolase.” The facts of life are seldom as simple as they seem. Just because it looks like a duck and quacks like a duck, we should not assume anything without experimental verification. After all, ducks and dinosaurs are more related than we might have expected.

REFERENCES

The views expressed in this Commentary do not necessarily reflect the views of the journal or of ASM.