Published in final edited form as:

Biochem J. 2015 February 15; 466(1): 137–145. doi:10.1042/BJ20141237.

A directed-overflow and damage-control *N*-glycosidase in riboflavin biosynthesis

Océane Frelin^{*}, Lili Huang[†], Ghulam Hasnain^{*}, James G. Jeffries^{‡,§}, Michael J. Ziemak^{*}, James R. Rocca^{||}, Bing Wang[¶], Jennifer Rice^{**}, Sanja Roje^{**}, Svetlana N. Yurgel^{**}, Jesse F. Gregory III[†], Arthur S. Edison[¶], Christopher S. Henry^{§,††}, Valérie deCrécy-Lagard^{‡‡}, and Andrew D. Hanson^{*,1}

^{*}Horticultural Sciences Department, University of Florida, Gainesville, FL 32611, U.S.A

[†]Food Science and Human Nutrition Department, University of Florida, Gainesville, FL 32611, U.S.A

[‡]Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL 60208, U.S.A

[§]Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL 60439, U.S.A

^{||}AMRIS Facility, McKnight Brain Institute, University of Florida, Gainesville, FL 32611, U.S.A

[¶]Department of Biochemistry and Molecular Biology, University of Florida, Gainesville, FL 32611, U.S.A

^{**}Institute of Biological Chemistry, Washington State University, Pullman, WA 99164, U.S.A

^{††}Computation Institute, University of Chicago, Chicago, IL 60637, U.S.A

^{‡‡}Department of Microbiology and Cell Science, University of Florida, Gainesville, FL 32611, U.S.A

Abstract

Plants and bacteria synthesize the essential human micronutrient riboflavin (vitamin B₂) via the same multistep pathway. The early intermediates of this pathway are notoriously reactive, and may be overproduced *in vivo* because riboflavin biosynthesis enzymes lack feedback controls. Here we demonstrate disposal of riboflavin intermediates by COG3236 (DUF1768), a protein of previously unknown function that is fused to two different riboflavin pathway enzymes in plants and bacteria (RIBR and RibA, respectively). We present cheminformatic, biochemical, genetic, and genomic evidence to show that: (i) plant and bacterial COG3236 proteins cleave the *N*-

¹To whom correspondence should be addressed: Andrew D. Hanson, Horticultural Sciences Department, University of Florida, Gainesville, FL 32611-0690, U.S.A., Tel: +1-352-273-4856, Fax: +1-352-392-5653, adha@ufl.edu.

AUTHOR CONTRIBUTION

Océane Frelin and Andrew Hanson conceived the study, analyzed data, and wrote the paper. Océane Frelin carried out experiments. Valérie de Crécy-Lagard provided genetic and genomic expertise, James Rocca, Océane Frelin, and Arthur Edison designed and executed NMR experiments, Bing Wang calculated NMR chemical shifts and assisted with analysis, James Jeffries and Christopher Henry made cheminformatic predictions, Lili Huang and Jesse Gregory conducted flavin analyses, Ghulam Hasnain prepared enzymes, Michael Ziemak performed cloning work, and Jennifer Rice, Sanja Roje, and Svetlana Yurgel planned and performed hybrid experiments.

glycosidic bond of the first two intermediates of riboflavin biosynthesis, yielding relatively innocuous products; (ii) certain COG3236 proteins are in a multienzyme riboflavin biosynthesis complex that gives them privileged access to riboflavin intermediates; and (iii) COG3236 action in *Arabidopsis thaliana* and *Escherichia coli* helps maintain flavin levels. COG3236 proteins thus illustrate two emerging principles in chemical biology: directed overflow metabolism, in which excess flux is diverted out of a pathway, and the pre-emption of damage from reactive metabolites.

Keywords

Arabidopsis thaliana; Maillard cascade; metabolite damage; *Vibrio vulnificus*; vitamin B₂; *Zea mays*

INTRODUCTION

There is more to metabolism than textbook pathways and controls such as feedback inhibition of enzyme activity. Several often-overlooked facets are captured by the emerging concepts of directed overflow metabolism [1] and of metabolite damage and its pre-emption or repair [2–4]. Directed overflow is the diversion of excess flux out of pathways when upstream feedback control is absent or ineffective; this limits the buildup of intermediates or end-products [1]. Metabolite damage – which is widespread and ceaseless – is the formation of unwanted side-products from normal metabolites via spontaneous reactions [5] or enzyme errors [6]. Damage pre-emption mechanisms destroy reactive metabolites or side-products before they do damage, and repair mechanisms recycle side-products to the metabolite they came from [2–4]. Directed overflow, damage pre-emption, and repair are all basically homeostatic in that they prevent buildup of normal metabolites or their corrupted forms [1,4].

Only a few cases of directed overflow or metabolite damage pre-emption and repair have been described [1,2]. There are, however, most likely many others – and combining chemical and biochemical logic with comparative genomics can help discover them [1,4]. Thus, directed overflow and pre-emption/repair are likely to operate, respectively, in pathways lacking feedback inhibition and in pathways with reactive intermediates. The riboflavin biosynthesis pathway of plants and bacteria (Figure 1A) fits both criteria. First, no feedback inhibition of its early enzymes has been reported for any organism, the only known regulatory mechanisms being transcriptional ones [7,8]. Second, the first two intermediates are highly reactive and have half-lives of a few hours [9–11]. It is consequently suggestive that the third enzyme (RIBR) of riboflavin biosynthesis in plants [12,13] and homologs of the first enzyme (RibA) in bacteria such as *Vibrio vulnificus* are fused to a conserved domain of unknown function (COG3236 or DUF1768) that is almost surely an enzyme [14] (Figure 1B). Furthermore, a mutation (*phs1*) that deletes the COG3236 domain of *Arabidopsis* RIBR is reported to reduce flavin content [12]. Because the domains of fusion proteins usually have related functions [15], these lines of evidence make the plant and bacterial COG3236 fusion domains candidates for directed overflow or damage pre-emption/repair enzymes in the riboflavin pathway. The same could apply to free-standing

COG3236 proteins [14], which occur in many organisms including *Escherichia coli* and *Nostoc punctiforme* (Figure 1B and Supplementary Figure S1A).

We therefore tested whether plant and bacterial COG3236 proteins are enzymes of directed overflow metabolism, damage pre-emption, or damage repair. We found that they hydrolyze the *N*-glycosidic bond in the first two riboflavin biosynthesis intermediates, that they likely have privileged access to these intermediates *in vivo*, and that they can divert flux out of the riboflavin pathway via relatively benign compounds. Besides documenting cross-kingdom cases of directed overflow and metabolite damage pre-emption, this work assigns a functional role to the COG3236 protein family and marshalls evidence for a multi-protein riboflavin biosynthesis complex.

EXPERIMENTAL

Bioinformatics

COG3236 sequences were from NCBI databases (<http://www.ncbi.nlm.nih.gov/>). Protein sequences were aligned using ClustalW (<http://www.genome.jp/tools/clustalw/>). Primers for gene expression analysis were designed using Primer Express 3.0 (Applied Biosystems) following manufacturer recommendations (optimal melting point, 60°C; optimal primer length, 20 bp; amplicon length, 60–80 bp). Mutagenesis primers were designed with PrimerX (<http://www.bioinformatics.org/primerx/index.htm>). Primers used in this study are listed in Supplementary Table S3.

Complementation of *E. coli* Ptet-*ribA* strain

The *E. coli* Ptet-*ribA* strain, in which the native *ribA* promoter is replaced by the inducible Ptet promoter [16], was obtained from B. El Yacoubi (University of Florida). This strain is auxotrophic for riboflavin in absence of the inducer anhydrotetracycline. The *E. coli ribA* gene and the gene encoding the *V. vulnificus* COG3236-RibA fusion protein were PCR-amplified from pET30a-RibA [13] and genomic DNA, respectively. The forward primers contained a KpnI site and a Shine-Dalgarno sequence before the start codon; the reverse primers contained an SphI site after the stop codon. Amplicons were cloned into KpnI/SphI-digested pBAD33 and sequences were verified. Both constructs were introduced into the Ptet-*ribA* strain by electroporation. Complementation tests were made on LB medium containing 50 µg/ml kanamycin, 100 µg/ml carbenicillin, 0.2% (w/v) arabinose, plus or minus 1 µg/ml anhydrotetracycline. Three independent transformants were used for each construct.

Expression and purification of COG3236 proteins

The *E. coli* and *N. punctiforme* COG3236 genes, the *V. vulnificus ribA* gene, and its COG3236 domain were PCR-amplified from genomic DNA. The COG3236 domains of the *Arabidopsis* and maize RIBR genes were amplified from pGEM-4Z-RIBR constructs [13]. Amplicons were cloned into pET28b, adding an N-terminal His-tag, and the resulting constructs were introduced into *E. coli* strain BL21-CodonPlus (DE3)-RIPL (Stratagene). Cultures (300 ml) were grown at 37°C in LB medium containing 50 µg/ml kanamycin, 50 µg/ml spectinomycin, and 30 µg/ml chloramphenicol. When A₆₀₀ reached 0.6–0.8, IPTG

(final concentration 1 mM) was added and incubation was continued for 4 h at 22°C. Subsequent steps were at 4°C. Cells were harvested by centrifugation (15 min at 4,000 *g*), resuspended in 5 ml of 50 mM HEPES-NaOH, pH 8.0, 300 mM NaCl, 10% (v/v) glycerol (Buffer A) containing 10 mM imidazole, and sonicated. Extracts were cleared by centrifugation (10 min at 10,000 *g*) and 2 ml of Ni²⁺-nitrilotriacetic acid agarose 50% slurry (Qiagen) was added to the supernatant. The mixture was slowly agitated for 1.5 h, then poured into a column. After washing the column with 60 ml of Buffer A containing 20 mM imidazole, proteins were eluted with 2.5 ml of Buffer A containing 250 mM imidazole. Proteins were desalted on PD-10 columns (Pierce) equilibrated with 50 mM HEPES-NaOH, pH 8.0, 300 mM betaine monohydrate, 10% (v/v) glycerol, 2 mM β-mercaptoethanol, and concentrated to 10–20 µg/µl using Amicon Ultra-4 10K units (Millipore). Proteins were estimated by the dye-binding method using bovine serum albumin as standard [17]. Purified proteins were aliquoted, flash-frozen in liquid nitrogen, and stored at –80°C. *E. coli* RibA and RibD, and maize ZmRIBR, were expressed and purified as described [13].

Site-directed mutagenesis of *E. coli* COG3236

Site-directed mutagenesis was performed on the pET28 *E. coli* GOG3236 construct using a Stratagene QuikChange® kit according to the manufacturer's recommendations. Residues were chosen for mutagenesis (to alanine) based on conservation (Supplementary Figure S1A) and position in the 3D-structure (PDB: 2b3w). Mutated constructs were sequence-verified.

Synthesis of riboflavin intermediates

Intermediate 1 was synthesized in 50-µl reaction mixtures containing 100 mM Tris-HCl, pH 8.0, 5 mM MgCl₂, 5 mM DTT, 5 mM GTP, and 100 µg *E. coli* RibA protein. After 1 h incubation at 22°C, RibA protein was removed using an Amicon Ultra-0.5 ml 10K unit (Millipore). The concentration of intermediate 1, which ranged from 2.8 to 3.2 mM, was estimated from HPLC traces by subtracting from the theoretical maximum (5 mM) the concentrations of unreacted GTP, GDP (a contaminant of GTP), and GMP (a side-product of RibA [18]). Intermediate 2 was synthesized as for intermediate 1 except that 150 µg of *E. coli* RibD protein was added as well as RibA. The concentration of intermediate 2, which ranged from 2.4 to 2.7 mM, was estimated from the amount of ammonia produced by the action of RibD by coupled assay (see below). Intermediates 1 and 2 spontaneously anomerize [9,18]; the anomers of intermediate 1 were not resolved by the HPLC conditions used, and those of intermediate 2 were partially resolved. Intermediate 3 was synthesized as described [13]; its concentration was estimated at 1.5 mM. Intermediates 1, 2, and 3 were not further purified and were freshly prepared before use in assays. The time course of degradation of intermediates 1 and 2 was run at 22°C in the dark for 24 h; samples were taken at various times, flash-frozen, and stored at –80°C until analysis.

COG3236 assay and product separation by HPLC

In a 50-µl reaction mixture, 40 µl of solution containing intermediate 1, 2, or 3 was mixed with 10 µl of COG3236 protein solution. Controls contained just the buffer used to desalt purified COG3236 proteins. After 1 h incubation at 22°C, 50 µl of 100 mM Tris-HCl, pH

8.0 was added and protein was removed using an Amicon Ultra-0.5 ml 10K unit. To 50 μ l of the resulting filtrate, 50 μ l of 100 mM Tris-HCl, pH 8.0 was added, and 25 μ l of the mixture was taken for ion-pair HPLC analysis using UV detection (at 293 nm for intermediate 1 or 284 nm for intermediate 2). To determine kinetic constants, the reaction volume was reduced to 25 μ l (20 μ l of substrate plus 5 μ l of enzyme) and each reaction included 100 μ g of bovine serum albumin to stabilize the enzyme. The amounts of COG3236 proteins and incubation times were such as to give linear rates of product formation. Kinetic constants were determined by non-linear regression using GraphPad Prism 6.0. Separations were made on a Discovery C18 column (250 $\times$ 4.6 mm, 5 μ m particle size) or a Spherisorb ODS-2 RP-18 column (250 $\times$ 4.6 mm, 5 μ m particle size) eluted isocratically with 1 mM tetrabutylammonium bisulfate, 100 mM KH₂PO₄, pH 6.5, 3% (v/v) methanol, and 1% (v/v) acetonitrile, at a flow rate of 1 ml/min. Derivatization of reaction mixtures with diacetyl and HPLC analysis of diacetyl derivatives was as described [13,19].

Bacterial strains for flavin analysis

The deletion of the *E. coli* COG3236 gene (*ybiA*) from the KEIO collection [20] was transferred to *E. coli* strain MG1655 by P1 transduction. The deletion was verified by PCR. The *E. coli ybiA* gene was cloned into the SacI/EcoRI restriction sites of pBlueScript II SK (+) for overexpression. The plasmid was introduced to the wild-type MG1655 strain by electroporation. Three independent colonies of wild-type and the deletant strain harboring the vector alone, and three independent colonies of the wild-type strain overexpressing COG3236 (*YbiA*) were cultivated in M9-glucose medium containing 1 mM IPTG and 100 μ g/ml ampicillin. Cells were harvested when A₆₀₀ reached 1.0. Cell pellets were frozen in liquid nitrogen and stored at -80°C until extraction for flavin analysis. Protein content was quantified from a 1-ml culture aliquot using the Thermo Scientific Pierce BCA protein assay kit.

Overexpression of COG3236 proteins in *Arabidopsis*

The COG3236 domains from the *V. vulnificus*, *Arabidopsis*, and maize were PCR-amplified as above and fused by PCR to the pea Rubisco small subunit targeting peptide amplified from pSSU13-precipTatC (obtained from X. Ma, University of Florida). Amplicons were digested with XbaI-BamHI and ligated to the corresponding sites of pFMV-nos [21] between the FMV-35S promoter and the nos terminator. The generated pFMV-SSU-COG constructs were then digested by NotI and ligated to the NotI-digested binary vector pHK1001 [21] and introduced into *Agrobacterium tumefaciens* (ABI strain) by electroporation. *Arabidopsis* Col-0 was transformed using the floral dip method [22]. Transgenic T1 lines were selected on solid Murashige and Skoog medium containing 1% sucrose and 0.6% agar supplemented with 50 μ g/ml kanamycin. Independent primary transformants were then transferred to potting soil. Plants were grown in a 12 h light-12 h dark cycle at 22°C at a photon flux density [12] of 80 μ mol m⁻² s⁻¹. Independent T1 lines were screened by PCR for the presence of the construct using the pFMV primers. Leaves from transgenic lines with or without bleaching symptoms were harvested, ground in liquid nitrogen, and stored at -80°C until analysis.

Flavin analysis

Flavins were extracted from 30–50 mg of leaf tissue in 300 μ l of methanol:methylene chloride (9:10, v/v). The suspension was vortexed for 10 s, 158 μ l of 10 mM KH_2PO_4 , pH 6.5 were added, and the mixture was vortexed again. After centrifugation (14,000 g, 15 min, 4°C) the aqueous phase was transferred to new tubes, filtered through an Amicon Ultra-0.5 mL 10K unit, and a 50- μ l sample was taken for HPLC analysis using fluorescence detection (excitation 452 nm; emission 516 nm). The same protocol was used for *E. coli* cultures except that, for a cell pellet from a 10-ml culture at $A_{600} = 1$, 1 ml of methanol:methylene chloride and 526 μ l of KH_2PO_4 were used for extraction and 10 μ l samples were taken for HPLC analysis. Separation was performed on a Microsorb-MV C18 column (150 x 4.6 mm, 5 μ m particle size) eluted isocratically at a flow rate of 1 ml/min with 75% (v/v) 10 mM KH_2PO_4 , pH 6.5, 25% (v/v) methanol. Recoveries of riboflavin, FMN, and FAD were close to 90%. Methods for bacterial two-hybrid analysis, enzymatic determination of ammonia, *in vitro* chloroplast import assays, qRT-PCR analysis, predicting COG3236 action using BNICE, NMR analysis, and chemical shift calculations are described in the Supplementary Materials and Methods section.

RESULTS

Bacterial COG3236-RibA fusions have GTP cyclohydrolase II activity

The RIBR fusion partner of the plant COG3236 protein is proven [13] to be the reductase that mediates step three in riboflavin biosynthesis, but there is no evidence yet that the RibA-like domain of the bacterial fusion proteins mediates step one, i.e. that it has GTP cyclohydrolase II activity (EC 3.5.4.25). Because the genomes of the bacteria with COG3236-RibA fusions encode other RibA family proteins [23] it seemed possible that these are functional GTP cyclohydrolases and the fusion protein is not. We therefore tested the *V. vulnificus* COG3236-*ribA* fusion gene for the ability to functionally complement an *E. coli ribA* mutant. The fusion gene was as effective as *E. coli ribA*, used as a positive control (Supplementary Figure S1B), confirming that the fusion protein is an active GTP cyclohydrolase II.

COG3236 proteins destroy the first two intermediates of riboflavin biosynthesis *in vitro*

The fusion of COG3236 to the first or third enzyme of riboflavin biosynthesis, and the reactivity of the early pathway intermediates, led us to test plant and bacterial COG3236 proteins for the capacity to transform these intermediates to other products. To this end, the COG3236 domains of the *V. vulnificus*, *Arabidopsis*, and maize fusion proteins and the free-standing *E. coli* and *N. punctiforme* COG3236 proteins were expressed in *E. coli* and purified (Supplementary Figure S1C), and the first three intermediates of riboflavin biosynthesis (henceforth: intermediates 1, 2, and 3) were made enzymatically *in situ* from GTP. The COG3236 proteins were then added to fresh preparations of each intermediate, and product formation was monitored by HPLC with UV detection. All five COG3236 proteins gave rise to a new HPLC peak when mixed with intermediate 1 or intermediate 2 (Figures 2A and 2B). The new peaks formed from intermediates 1 and 2 differed in retention time and had absorption spectra distinct from each other and from the parent intermediate

(Supplementary Figure S2). In contrast, none of the proteins had detectable activity against intermediate 3, or against GTP, nucleoside monophosphates, or ADP-ribose.

To exclude the possibility that the observed activity came from contaminating proteins, each of four likely active site residues [14] in the *E. coli* protein was changed to alanine (Supplementary Figure S1A) and the mutant proteins (Supplementary Figure S1D) were tested for activity against intermediate 1. Three mutations abolished activity and the fourth drastically cut it (Figure 2C) showing that the activity resides in the COG3236 protein itself.

Cheminformatic prediction of the products of COG3236 action

To guide investigation of the products of COG3236 action, we applied the Biochemical Network Integrated Computational Explorer (BNICE) [24,25] to predict enzymatic transformations that intermediates 1 and 2 could undergo in the absence of reactants except water and oxygen. We applied seven general BNICE one-step transformations on intermediates 1 and 2, which generated eight potential bond-cleavage reactions (Supplementary Figure S3): dephosphorylation, hydrolysis of the 5-amino group of the pyrimidine ring, ring-opening at any one of four positions, hydrolysis of the C₆-N bond, and hydrolysis of the *N*-glycosidic bond. The first seven possibilities can be provisionally excluded because the functional groups involved occur in intermediate 3, which COG3236 does not attack. This leaves *N*-glycosidic bond cleavage, which fits with comparative genomic evidence [14] that proteins from a sister family to COG3236 mediate de-ADP-ribosylation, entailing hydrolysis of an *N*-glycosidic bond.

COG3236 has *N*-glycosidase activity against riboflavin biosynthesis intermediates

To test for *N*-glycosidase activity, we first used diacetyl derivatization. Diacetyl treatment of intermediates 1 or 2 yields fluorescent derivatives of their pyrimidine moieties and removes the phosphoribosyl group; the same derivatives are formed by treating the corresponding free pyrimidines [19]. When the products of COG3236 action on intermediates 1 or 2 were treated with diacetyl, fluorescent compounds were formed whose retention times were identical to the retention times of the products formed from intermediates 1 or 2 themselves and from standards of the corresponding free pyrimidines, 4-hydroxy-2,5,6-triaminopyrimidine (for intermediate 1) or 5,6-diaminouracil (for intermediate 2) (Figures 2D and 2E). Additionally, the underivatized products of COG3236 action on intermediates 1 or 2 had the same retention times and absorption spectra as underivatized 4-hydroxy-2,5,6-triaminopyrimidine or 5,6-diaminouracil, respectively (Supplementary Figure S2). These data indicate that COG3236 proteins cleave the *N*-glycosidic bond, giving a pyrimidine and ribose-5-phosphate.

N-Glycosidic bond cleavage was further investigated by ¹³C NMR. Intermediates 1 and 2 prepared from (U-¹³C₁₀, U-¹⁵N₅)-GTP were used as substrates for *E. coli* COG3236. Monitoring of both reaction mixtures clearly showed the disappearance of glycosidic bond resonances and the appearance of signals characteristic of the pyrimidine and ribose-5-phosphate moieties released by cleavage of this bond (Figures 3A and 3B). When authentic ribose-5-phosphate at natural abundance was spiked into the intermediate 1 reaction (Figure 3A, red overlay), all resonances from the spike corresponded to the center of the multiplets

from ^{13}C - ^{13}C scalar couplings observed in the ^{13}C -labeled enzymatic products, proving that ribose-5-phosphate is a reaction product. The pyrimidine resonances and those of other intermediates were assigned from *ab initio* quantum mechanical calculations of the species hypothesized as possible products. The assignments and calculations are given in Supplementary Table S1. These results confirm that COG3236 cleaves the *N*-glycosidic bond in intermediates 1 and 2.

COG3236 proteins prefer either intermediate 1 or intermediate 2

Kinetic constants were estimated for *E. coli* COG3236, for the isolated COG3236 domains of the *V. vulnificus* and maize fusion proteins, and for the native fusion proteins themselves to assess the effect of truncation (Table 1 and Supplementary Figure S4). Although the values are necessarily approximate because intermediates 1 and 2 cannot be prepared pure [9,11], two points emerge clearly. First, all five enzymes preferred one intermediate or the other. Second, the K_{cat} and K_{m} values for the preferred substrate (0.8 to 137 s^{-1} and 0.19 to 5.84 mM, respectively) are typical for *N*-glycosidases [26] and cofactor-metabolizing enzymes in general [27]. The preferred substrate, as judged from $K_{\text{cat}}/K_{\text{m}}$ ratios, differed between enzymes. The *V. vulnificus* COG3236 domain and native fusion protein showed a $\sim 10^3$ -fold preference for intermediate 1, whereas the maize COG3236 domain and native fusion preferred intermediate 2, although to different extents ($\sim 10^3$ -fold and ~ 10 -fold, respectively). The preference of *V. vulnificus* COG3236 for the earlier intermediate and of plant COG3236 for the later one matches their respective fusions to earlier and later enzymes in the riboflavin pathway (Figure 1B).

Riboflavin pathway enzymes can form a metabolon

That COG3236 domains fused to different riboflavin pathway enzymes prefer the intermediate nearest to them in the pathway suggests that the intermediates remain close to the enzymes, i.e. that they are channeled within a multi-enzyme complex. Mining of protein-protein interaction and genome databases uncovered strong evidence that such a complex exists in diverse organisms (Supplementary Figure S5A). Thus, in microbes, protein-protein interactions have been detected biochemically [28–32] or genetically [33] between each of the six riboflavin biosynthesis enzymes and at least one of the others. Second, sequenced genomes encode pairwise fusions between almost all of the riboflavin enzymes and triple fusions among several of them [34], which proves their potential to physically associate. That COG3236 is fused to two riboflavin enzymes indicates that it, too, can be integrated into the riboflavin metabolon.

Further support for a riboflavin biosynthesis complex came from a bacterial double-hybrid analysis with enzymes from *Sinorhizobium meliloti* (Supplementary Figure S5B). The bifunctional enzyme RibBA interacted with both RibH1 and RibH2 and, as expected [35], the multimeric lumazine syn-thases RibH1 and RibH2 interacted with self and with each other. Interactions between RibE and RibH1 or RibH2, and RibE with self, have been reported from other bacteria [35].

Spontaneous reactions of intermediates 1 and 2

Intermediates 1 and 2 are clearly labile [10,11] but the kinetics and products of their breakdown are not known – and this information is needed to interpret the *in-vivo* function of COG3236. Both intermediates are glycosylamines that could undergo Amadori rearrangement, followed by other reactions in the Maillard sequence leading to reactive and potentially toxic products [36–38] (Figure 4A). Also, C₆-N bond scission could yield 5-phosphoribosyl-amine and a pyrimidine [39] (Figure 4A). In physiological conditions, 5-phosphoribosylamine has a reactive amino group and hydrolyzes rapidly ($t_{1/2} = 38$ s) to ribose-5-phosphate and NH₃ [40,41]. The pyrimidine products (2,5-diamino-4,6-dihydroxypyrimidine from intermediate 1 and 5-aminobarbiturate from intermediate 2) are probably also reactive [42].

We therefore monitored the spontaneous decomposition of intermediates 1 and 2 by HPLC, by NH₃ release, and by NMR. The UV-absorbing HPLC peak of each intermediate disappeared, indicating breakdown of the pyrimidine ring. The rate of disappearance increased with time, the loss of intermediate 2 being faster (Figure 4B and Supplementary Figures S6A and S6B); such accelerating rates are not inconsistent with Maillard reactions, which can show a lag [43]. Both intermediates released about one mole of NH₃ per mole of reactant, the release rate being faster from intermediate 2 (Figure 4C). Release of NH₃ is consistent with 5-phosphoribosylamine formation [40]. The ¹³C NMR spectral changes were complex, as would be expected for mixtures of ring-opened pyrimidines and Maillard reaction products, but showed clearly that spontaneous degradation of intermediates 1 or 2 differs from the COG3236-catalyzed reactions (Figure 4D). The NMR data also supported breakdown via 5-phosphoribosylamine inasmuch as resonances of ribose-5-phosphate (the product of 5-phosphoribosylamine breakdown) increased over time in the spectra for both intermediates (Figure 4D). Collectively, these data confirm that both intermediates are unstable, that 2 is even more unstable than 1, and that they decompose in complex ways. Given the precedents [36–39], these ways could include Maillard reactions and 5-phosphoribosylamine formation, as well as opening of the pyrimidine ring. However, the data do not exclude ribose-5-phosphate formation from scission of the *N*-glycosidic bond, or NH₃ formation from pyrimidines released by scission of this bond or the C₆-N bond, or by the Maillard sequence (Supplementary Figure S6C).

Flavin phenotypes of plant and bacterial COG3236 mutants and overexpressors

We first confirmed the flavin deficiency reported [12] for *Arabidopsis phs1* mutants, which lack the COG3236 domain of the RIBR fusion protein. In leaves grown in normal light, the FAD and FMN contents of the mutant were significantly lower (by 16% and 20%, respectively) than the wild type (Supplementary Figure S7A). We then compared the flavin contents of an *E. coli* COG3236 deletant strain with the wild type (Supplementary Table S2). As with *Arabidopsis*, the FAD and FMN contents were significantly reduced (both by 13%).

We also overexpressed various COG3236 proteins in *Arabidopsis*, and *E. coli* COG3236 in *E. coli*. In neither case were there marked effects on flavins. The *Arabidopsis* expression constructs contained the COG3236 domains of the *Arabidopsis*, maize, or *V. vulnificus*

fusion proteins (Figure 1B) preceded by a plastid-targeting peptide, the plant RIBR-COG3236 protein being plastidial [13]. *In vitro* assays confirmed chloroplast import of the engineered proteins (Supplementary Figure S7B). Analysis of transgenic plants showed normal levels of individual flavins and total flavins (Supplementary Figure S7C). Similarly, *E. coli* cells overexpressing COG3236 had flavin levels similar to vector-alone controls (Supplementary Table S2).

DISCUSSION

The COG3236 protein family has hitherto had only conjectural roles [14]. Here, we demonstrate that COG3236 proteins from plants and bacteria have *N*-glycosidase activity against intermediates 1 and 2 of the plant/bacterial riboflavin biosynthesis pathway. This activity was found in COG3236 proteins that occur naturally as fusions or in free-standing form. None of these proteins had detectable activity against ribonucleotides or ADP-ribose, which are structurally similar to intermediates 1 and 2. However, the *N*-glycosidic bond in intermediates 1 and 2 links the anomeric carbon to an amino substituent on the pyrimidine ring whereas in nucleotides the linkage is to the ring itself. The *N*-glycosidic bond in intermediates 1 and 2 is thus unlike that in nucleotides.

Our data suggest a model for the physiological function for COG3236 proteins, particularly those fused to riboflavin biosynthesis enzymes. This model combines directed overflow metabolism [1] and metabolite damage pre-emption [2] (Figure 5). It rests on four premises: (i) Absent feedback inhibition of GTP cyclohydrolase II, production of riboflavin intermediates can overshoot, leading to their accumulation. (ii) Accumulation of intermediates 1 and 2 is deleterious to riboflavin synthesis because they spontaneously degrade to reactive, injurious compounds. (iii) COG3236 action converts intermediates 1 and/or 2 into ribose-5-phosphate and low-toxicity pyrimidines, thus diverting a surplus of harmful intermediates into relatively harmless products and pre-empting the damage these intermediates would otherwise do. (iv) Riboflavin biosynthesis enzymes and their COG3236 partners form a complex that confines and channels the pathway intermediates so that COG3236 proteins have privileged access to these intermediates, and the riboflavin pathway takes the hardest damage hit if COG3236 action is blocked.

The support for these premises is as follows:

- i. Two lines of evidence argue that GTP cyclohydrolase II is not subject to feedback inhibition, potentially allowing riboflavin biosynthesis intermediates to accumulate. First, studies of microbes selected or engineered to overproduce riboflavin show that GTP cyclohydrolase II does not normally control flux in the pathway, and that GTP supply does [44–46]. Second, there are no reports of feedback inhibition of GTP cyclohydrolase II enzymes although they have been thoroughly studied [7].
- ii. Intermediates 1 and 2 are labile and are predisposed to form Maillard products, whose propensity to damage biomolecules is well established [36,37,43]. Intermediates 1 and 2 may also decompose to give the reactive nucleophile 5-phosphoribosylamine [39,40] and reactive pyrimidines [42].

- iii. Our data demonstrate that COG3236 proteins hydrolyze intermediates 1 and 2 to ribose-5-phosphate and a pyrimidine. That these products are less harmful than their parent intermediates can be strongly inferred. First, although ribose-5-phosphate is a glycosylating agent [37,47], plastids have a repair enzyme to reverse the damage it does [48] and in any case the riboflavin pathway is a negligible source of ribose-5-phosphate compared to the Calvin cycle. Second, the pyrimidines formed from intermediates 1 and 2 appear not to be toxic to microorganisms [49–51].
- iv. Our own data and public datasets show that plant and bacterial riboflavin biosynthesis enzymes can be fused or non-covalently bound to each other in various combinations, pointing to the existence of a multi-enzyme complex that can include COG3236, at least when it is fused to riboflavin enzymes.

The proposed model accounts for the flavin depletion observed in the *Arabidopsis phs1* mutant and the *E. coli* COG3236 knockout strain. Without COG3236, cells cannot deal with surplus intermediates 1 and 2, which can therefore build up within the riboflavin biosynthesis complex and give rise to harmful products. These products in turn attack and inactivate the enzymes of the complex, which are the nearest bystanders and experience the main force of the damage. Ablating COG3236 thus impacts the riboflavin pathway first and foremost.

The model can also explain why overexpression of COG3236 proteins in *Arabidopsis* chloroplasts or in *E. coli* cytosol has virtually no effect on flavin content. Being inside a complex, intermediates 1 and 2 are shielded from the overexpressed COG3236 proteins outside it. Were the intermediates not shielded in this way, COG3236 overexpression would be expected to hydrolyze them and hence to cause flavin deficiency.

The model is consistent with the observed K_m values for COG3236 enzymes. While typical for *N*-glycosidases, these values (0.19 to 5.84 mM) are well above those reported for the riboflavin biosynthesis enzymes RibA and RibD (5 to 63 μ M) [52–54] to which COG3236 domains are sometimes fused. If COG3236 enzymes deal with *excesses* of intermediates 1 and 2, they might be expected to reach high activity only at concentrations of these intermediates that are saturating for the biosynthetic enzymes. In this connection, it should be noted that the striking 10^3 -fold preference for intermediate 1 shown by the *V. vulnificus* enzyme obviates any need for activity against intermediate 2 because intermediate 2 cannot accumulate if intermediate 1 is removed. The converse is not true, and indeed while the native plant enzyme prefers intermediate 2 it also has substantial activity against intermediate 1, potentially allowing it to remove both.

The disposal of excess intermediates 1 and 2 is almost surely a major, if not the sole, function of the COG3236 proteins fused to riboflavin synthesis enzymes, and is most likely a function of the free-standing *E. coli* COG3236 protein, although not necessarily the only one. However, the COG3236 family is diverse and probably encompasses functions unrelated to riboflavin [14]. For instance, many bacterial genomes contain COG3236 genes that cluster on the chromosome with genes of ADP-ribose metabolism [14]. As ADP-ribose is a potent glycosylating agent [55], and the glycosylamines that launch the glycation process are

analogs of intermediates 1 and 2, it is conceivable that ADP-ribose-associated COG3236 proteins interrupt the glycation cascade by hydrolyzing its early intermediates.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank R.K. Sandwick for chemical advice on Maillard reactions, E. van Schaftingen and C. Goodman for critiquing the manuscript, B. El Yacoubi for the P_{tet}-*riba* strain, H.J. Klee for pHK1001, and X. Ma for pSSU13precipTatC.

FUNDING

This work was funded by US National Science Foundation (NSF) grants MCB-1153413 (to Andrew Hanson), MCB-1153357 (to Christopher Henry), and by the C.V. Griffin Sr. Foundation. NMR studies were supported by the NSF National High Magnetic Field Laboratory User Program in the Advanced Magnetic Resonance Imaging and Spectroscopy (AMRIS) Facility in the University of Florida McKnight Brain Institute. NMR data analysis and interpretation were supported by the Southeast Center for Integrated Metabolomics (NIH 1U24DK097209-01A1). Flavin analyses were supported by NSF grant IOS-1025398.

Abbreviations used

intermediate 1	2,5-diamino-6-ribosylamino-4(3 <i>H</i>)-pyrimidinone 5'-phosphate
intermediate 2	5-amino-6-ribosylamino-2,4(1 <i>H</i> ,3 <i>H</i>)-pyrimidinedione 5'-phosphate
intermediate 3	5-amino-6-ribitylamino-2,4(1 <i>H</i> ,3 <i>H</i>)-pyrimidinedione 5'-phosphate

References

1. Reaves ML, Young BD, Hosios AM, Xu YF, Rabinowitz JD. Pyrimidine homeostasis is accomplished by directed overflow metabolism. *Nature*. 2013; 500:237–241. [PubMed: 23903661]
2. Linster CL, Van Schaftingen E, Hanson AD. Metabolite damage and its repair or pre-emption. *Nat Chem Biol*. 2013; 9:72–80. [PubMed: 23334546]
3. Vinci CR, Clarke SG. Homocysteine methyltransferases Mht1 and Sam4 prevent the accumulation of age-damaged (*R,S*)-AdoMet in the yeast *Saccharomyces cerevisiae*. *J Biol Chem*. 2010; 285:20526–20531. [PubMed: 20421295]
4. Danchin A, Sekowska A. The logic of metabolism and its fuzzy consequences. *Environ Microbiol*. 2014; 16:19–28. [PubMed: 24387040]
5. Golubev AG. The other side of metabolism: a review. *Biochemistry (Mosc)*. 1996; 61:2018–2039.
6. Tawfik DS. Messy biology and the origins of evolutionary innovations. *Nat Chem Biol*. 2010; 6:692–696. [PubMed: 20852602]
7. Fischer M, Bacher A. Biosynthesis of flavocoenzymes. *Nat Prod Rep*. 2005; 22:324–350. [PubMed: 16010344]
8. Vorwieger A, Gryczka C, Czihal A, Douchkov D, Tiedemann J, Mock HP, Jakoby M, Weisshaar B, Saalbach I, Bäumlein H. Iron assimilation and transcription factor controlled synthesis of riboflavin in plants. *Planta*. 2007; 226:147–158. [PubMed: 17260143]
9. Fischer M, Römisch W, Saller S, Illarionov B, Richter G, Rohdich F, Eisenreich W, Bacher A. Evolution of vitamin B2 biosynthesis: structural and functional similarity between pyrimidine deaminases of eubacterial and plant origin. *J Biol Chem*. 2004; 279:36299–36308. [PubMed: 15208317]
10. Foor F, Brown GM. Purification and properties of guanosine triphosphate cyclohydrolase II from *Escherichia coli*. *J Biol Chem*. 1975; 250:3545–3551. [PubMed: 235552]

11. Magalhães ML, Argyrou A, Cahill SM, Blanchard JS. Kinetic and mechanistic analysis of the *Escherichia coli* *ribD*-encoded bifunctional deaminase-reductase involved in riboflavin biosynthesis. *Biochemistry*. 2008; 47:6499–6507. [PubMed: 18500821]
12. Ouyang M, Ma J, Zou M, Guo J, Wang L, Lu C, Zhang L. The photosensitive *phsI* mutant is impaired in the riboflavin biogenesis pathway. *J Plant Physiol*. 2010; 167:1466–1476. [PubMed: 20580123]
13. Hasnain G, Frelin O, Roje S, Ellens KW, Ali K, Guan JC, Garrett TJ, de Crécy-Lagard V, Gregory JF 3rd, McCarty DR, Hanson AD. Identification and characterization of the missing pyrimidine reductase in the plant riboflavin biosynthesis pathway. *Plant Physiol*. 2013; 161:48–56. [PubMed: 23150645]
14. de Souza RF, Aravind L. Identification of novel components of NAD-utilizing metabolic pathways and prediction of their biochemical functions. *Mol Biosyst*. 2012; 8:1661–1677. [PubMed: 22399070]
15. Suhre K. Inference of gene function based on gene fusion events: the Rosetta-stone method. *Methods Mol Biol*. 2007; 396:31–41. [PubMed: 18025684]
16. Da Re S, Le Quéré B, Ghigo JM, Beloin C. Tight modulation of *Escherichia coli* bacterial biofilm formation through controlled expression of adhesion factors. *Appl Environ Microbiol*. 2007; 73:3391–3403. [PubMed: 17384304]
17. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*. 1976; 72:248–254. [PubMed: 942051]
18. Ritz H, Schramek N, Bracher A, Herz S, Eisenreich W, Richter G, Bacher A. Biosynthesis of riboflavin: studies on the mechanism of GTP cyclohydrolase II. *J Biol Chem*. 2001; 276:2273–2277.
19. Richter G, Fischer M, Krieger C, Eberhardt S, Lüttgen H, Gerstenschläger I, Bacher A. Biosynthesis of riboflavin: characterization of the bifunctional deaminase-reductase of *Escherichia coli* and *Bacillus subtilis*. *J Bacteriol*. 1997; 179:2022–2028. [PubMed: 9068650]
20. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol*. 2006; 2:2006.0008.
21. Moussatche, P. PhD Thesis. University of Florida; 2004. The ethylene receptor multigene family: Insights on expression, localization and function in *Arabidopsis* and tomato.
22. Clough SJ, Bent AF. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J*. 1998; 16:735–743. [PubMed: 10069079]
23. Brutinel ED, Dean AM, Gralnick JA. Description of a riboflavin biosynthetic gene variant prevalent in the phylum Proteobacteria. *J Bacteriol*. 2013; 195:5479–5486. [PubMed: 24097946]
24. Hatzimanikatis V, Li C, Ionita JA, Henry CS, Jankowski MD, Broadbelt LJ. Exploring the diversity of complex metabolic networks. *Bioinformatics*. 2005; 21:1603–1609. [PubMed: 15613400]
25. Li C, Henry CS, Jankowski MD, Ionita JA, Hatzimanikatis V, Broadbelt LJ. Computational discovery of biochemical routes to specialty chemicals. *Chem Eng Sci*. 2004; 59:5051–5060.
26. Hansen MR, Dandanell G. Purification and characterization of RihC, a xanthosine-inosine-uridine-adenosine-preferring hydrolase from *Salmonella enterica* serovar. Typhimurium *Biochim Biophys Acta*. 2005; 1723:55–62.
27. Bar-Even A, Noor E, Savir Y, Liebermeister W, Davidi D, Tawfik DS, Milo R. The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry*. 2011; 50:4402–4410. [PubMed: 21506553]
28. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C, Jensen LJ. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res*. 2013; 41:D808–D815. [PubMed: 23203871]
29. Butland G, Peregrín-Alvarez JM, Li J, Yang W, Yang X, Canadien V, Starostine A, Richards D, Beattie B, Krogan N, Davey M, Parkinson J, Greenblatt J, Emili A. Interaction network containing conserved and essential protein complexes in *Escherichia coli*. *Nature*. 2005; 433:531–537. [PubMed: 15690043]

30. Hu P, Janga SC, Babu M, Díaz-Mejía JJ, Butland G, Yang W, Pogoutse O, Guo X, Phanse S, Wong P, et al. Global functional atlas of *Escherichia coli* encompassing previously uncharacterized proteins. *PLoS Biol.* 2009; 7:e96. [PubMed: 19402753]
31. Parrish JR, Yu J, Liu G, Hines JA, Chan JE, Mangiola BA, Zhang H, Pacifico S, Fotouhi F, DiRita VJ, Ideker T, Andrews P, Finley RL Jr. A proteome-wide protein interaction map for *Campylobacter jejuni*. *Genome Biol.* 2007; 8:R130. [PubMed: 17615063]
32. Kis K, Bacher A. Substrate channeling in the lumazine synthase/riboflavin synthase complex of *Bacillus subtilis*. *J Biol Chem.* 1995; 270:16788–16795. [PubMed: 7622491]
33. Spitzner A, Perzlmaier AF, Geillinger KE, Reihl P, Stolz J. The proline-dependent transcription factor Put3 regulates the expression of the riboflavin transporter MCH5 in *Saccharomyces cerevisiae*. *Genetics.* 2008; 180:2007–2017. [PubMed: 18940788]
34. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, et al. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* 2011; 39:D225–D229. [PubMed: 21109532]
35. Ladenstein RM, Fischer M, Bacher A. The lumazine synthase/riboflavin synthase complex: shapes and functions of a highly variable enzyme system. *FEBS J.* 2013; 280:2537–2563. [PubMed: 23551830]
36. Isbell HS, Frush HJ. Mutarotation, hydrolysis, and rearrangement of glycosylamines. *J Org Chem.* 1958; 23:1309–1319.
37. Munanairi A, O'Banion SK, Gamble R, Breuer E, Harris AW, Sandwick RK. The multiple Maillard reactions of ribose and deoxyribose sugars and sugar phosphates. *Carbohydr Res.* 2007; 342:2575–2592. [PubMed: 17850774]
38. Bacher A, Lingens F. Biosynthesis of riboflavin. Formation of 2,5-diamino-6-hydroxy-4-(1'-D-ribitylamino)pyrimidine in a riboflavin auxotroph. *J Biol Chem.* 1970; 245:4647–4652. [PubMed: 5456141]
39. Plaut GW. Studies on the nature of the enzymic conversion of 6,7-dimethyl-8-ribityllumazine to riboflavin. *J Biol Chem.* 1963; 238:2225–2243.
40. Schendel FJ, Cheng YS, Otvos JD, Wehrli S, Stubbe J. Characterization and chemical properties of phosphoribosylamine, an unstable intermediate in the *de novo* purine biosynthetic pathway. *Biochemistry.* 1988; 27:2614–2623. [PubMed: 2454658]
41. Inouye S. On the prediction of pK_a values of amino sugars. *Chem Pharm Bull.* 1968; 16:1134–1137. [PubMed: 5706827]
42. Elsner M, Gurgul-Convey E, Lenzen S. Relation between triketone structure, generation of reactive oxygen species, and selective toxicity of the diabetogenic agent alloxan. *Antioxid Redox Signal.* 2008; 10:691–699. [PubMed: 18177230]
43. Sandwick R, Johanson M, Breuer E. Maillard reactions of ribose 5-phosphate and amino acids. *Ann NY Acad Sci.* 2005; 1043:85–96. [PubMed: 16037226]
44. Burgess CM, Smid EJ, Rutten G, van Sinderen D. A general method for selection of riboflavin-overproducing food grade micro-organisms. *Microb Cell Fact.* 2006; 5:24. [PubMed: 16848883]
45. Stahmann KP, Revuelta JL, Seulberger H. Three biotechnical processes using *Ashbya gossypii*, *Candida famata*, or *Bacillus subtilis* compete with chemical riboflavin production. *Appl Microbiol Biotechnol.* 2000; 53:509–516. [PubMed: 10855708]
46. Hümbelin M, Griesser V, Keller T, Schurter W, Haiker M, Hohmann HP, Ritz H, Richter G, Bacher A, van Loon APGM. GTP cyclohydrolase II and 3,4-dihydroxy-2-butanone 4-phosphate synthase are the rate-limiting enzymes in riboflavin synthesis of an industrial *Bacillus subtilis* strain used for riboflavin production. *J Ind Microbiol Biotechnol.* 1999; 22:1–7.
47. Hildick-Smith GJ, Downey MC, Gretebeck LM, Gersten RA, Sandwick RK. Ribose 5-phosphate glycation reduces cytochrome c respiratory activity and membrane affinity. *Biochemistry.* 2011; 50:11047–11057. [PubMed: 22091532]
48. Fortpied J, Gemayel R, Stroobant V, van Schaftingen E. Plant ribulosa-mine/erythrosamine 3-kinase, a putative protein-repair enzyme. *Biochem J.* 2005; 388:795–802. [PubMed: 15705060]
49. Middelhoven WJ, Brons HJ, Breedveld MW, Suy IM, van der Plas HC. Gratuitous induction of xanthine oxidase by 5,6-diaminouracil in continuously grown cells of *Arthrobacter globiformis* M4. *Appl Microbiol Biotechnol.* 1989; 32:323–326.

50. Sadique J, Shanmugasundaram R, Shanmugasundaram ER. Formation of 4,5-diaminouracil in a riboflavineless mutant of *Aspergillus nidulans*. *Naturwissenschaften*. 1966; 53:282. [PubMed: 5987043]
51. Bacher A, Lingens F. Biosynthesis of riboflavin. Formation of 6-hydroxy-2,4,5-triamino-pyrimidine in Rib⁷ mutants of *Saccharomyces cerevisiae*. *J Biol Chem*. 1971; 246:7018–7022. [PubMed: 4942329]
52. Lehmann M, Degen S, Hohmann HP, Wyss M, Bacher A, Schramek N. Biosynthesis of riboflavin. Screening for an improved GTP cyclohydrolase II mutant. *FEBS J*. 2009; 276:4119–4129. [PubMed: 19583770]
53. Spoonamore JE, Bandarian V. Understanding functional divergence in proteins by studying intragenomic homologues. *Biochemistry*. 2008; 47:2592–2600. [PubMed: 18281960]
54. Burrows RB, Brown GM. Presence of *Escherichia coli* of a deaminase and a reductase involved in biosynthesis of riboflavin. *J Bacteriol*. 1978; 136:657–667. [PubMed: 30756]
55. Jacobson EL, Cervantes-Laurean D, Jacobson MK. ADP-ribose in glycation and glycoxidation reactions. *Adv Exp Med Biol*. 1997; 419:371–379. [PubMed: 9193679]
56. Ramaswamy V, Hooker JW, Withers RS, Nast RE, Brey WW, Edison AS. Development of a ¹³C-optimized 1.5-mm high temperature superconducting NMR probe. *J Magn Reson*. 2013; 235:58–65. [PubMed: 23969086]

Figure 1. Evidence connecting COG3236 with riboflavin biosynthesis

(A) The first three steps of riboflavin biosynthesis and their enzymes in bacteria (above arrows) and plants (below arrows). Steps two and three in bacteria are catalyzed by RibD, a bifunctional deaminase-reductase. RibA and RIBA1, GTP cyclohydrolase II; RibD and PYRD, pyrimidine deaminase; RibD and RIBR, pyrimidine reductase. Intermediates 1, 2, and 3 are 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate, 5-amino-6-ribosylamino-2,4(1H,3H)-pyrimidinedione 5'-phosphate, and 5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione 5'-phosphate, respectively; P, phosphate.

(B) Domain structures of fusions between COG3236 and RibA in γ -Proteobacteria, and between COG3236 and RIBR in plants. Two stand-alone bacterial COG3236 proteins are also shown. TP, plastid targeting peptide.

Figure 2. Chromatographic evidence that plant and bacterial COG3236 proteins cleave the pyrimidine moiety from riboflavin biosynthesis intermediates 1 and 2

(A) Action of *E. coli* (Ec), *N. punctiforme* (Np), *V. vulnificus* (Vv), *Arabidopsis* (At), and maize (Zm) COG3236 proteins on intermediate 1 (black arrow). Reactions (50 μ l) containing 3 mM intermediate 1 and 30 μ g of protein were incubated for 1 h at 22°C and analyzed by HPLC with UV detection. Control reactions contained no enzyme.

(B) Action of COG3236 proteins on intermediate 2 (grey arrow). Conditions were as in A except that substrate concentration was 2.4 mM and the amounts of protein (μ g) were: Vv, 5; Ec and Np, 1; Zm 0.25; At, 0.1. The shoulder on the intermediate 2 peak is due to partial resolution of the two anomers [9].

(C) Elimination or reduction of activity of *E. coli* COG3236 by site-directed mutagenesis of four conserved residues. A reaction with wild type (WT) protein is shown for comparison. Conditions were as in A except that 100 μ g of protein was used.

(D,E) Fluorometric HPLC analysis of the diacetyl derivatives of, respectively, intermediates 1 and 2, of the products formed therefrom by *E. coli* or *Arabidopsis* COG3236, and of authentic 4-hydroxy-2,5,6-triaminopyrimidine (HTP) or 5,6-diaminouracil (DAU). Control reactions contained no enzyme.

Figure 3. NMR analysis confirms that *E. coli* COG3236 cleaves the *N*-glycosidic bond of riboflavin intermediates 1 and 2

Aromatic (left) and ribose regions (right) are shown for reactions involving (A) intermediate 1 generated by RibA and (B) intermediate 2 generated by RibA plus RibD. The starting GTP is shown as reference at the bottom. Spectra were obtained using volumes of just 50 μ l by incubating labeled GTP with COG3236 in the 1.5-mm NMR tubes. An exceptionally sensitive custom ^{13}C -optimized NMR probe made from high temperature superconducting material [56] was utilized. The red overlay in the ribose region of 'Intermediate 1 + COG3236' is the ribose-5-phosphate spike of the same spectrum in black. Asterisks indicate the ribose-5-phosphate resonances that increased in the spike. Note that ribose-5-phosphate exists as a mixture of α and β anomers. Assignments of NMR spectra of unstable intermediates were made with calculated chemical shifts (Table S1).

Figure 4. Spontaneous breakdown chemistry of intermediates 1 and 2

(A) Hypothetical breakdown reactions of intermediates 1 and 2. 5AB, 5-aminobarbituric acid; AGEs, advanced glycation end products; DAU, diaminouracil; DHP, 2,5-diamino-4,6-dihydroxypyrimidine; HTP, 4-hydroxy-2,5,6-triaminopyrimidine; PRA, 5-phosphoribosylamine; R-5-P, ribose-5-phosphate.

(B) Kinetics of the disappearance of the UV-absorbing HPLC peaks of intermediates 1 and 2.

(C) Kinetics of NH_3 release from intermediates 1 (3 mM) and 2 (2.4 mM).

(D) Freshly prepared ^{13}C , ^{15}N -labeled intermediate 2 (2.4 mM) was incubated at 22°C in the NMR tube and ^{13}C NMR spectra were acquired during successive 4-h periods. The first spectrum (blue) and last spectrum collected at 20 h (red) are shown at the top. The spectra from Figure 3 of intermediate 2 (middle) and intermediate 2 + COG3236 (bottom) are shown for comparison. The left portion is the aromatic region and the right portion is the ribose region. The arrows above the degradation spectra indicate resonances that decrease ($\downarrow$) or increase ($\uparrow$) over time. Spectral changes for intermediate 1 were similar but less marked.

Figure 5. A model of COG3236 action in the riboflavin biosynthesis pathway

COG3236 proteins dispose of excess intermediates 1 and 2 by hydrolysis to ribose-5-phosphate and a pyrimidine. These are less damaging than compounds such as Maillard products and 5-phospho-ribosylamine (PRA) to which the intermediates could spontaneously break down. The pathway enzymes, and COG3236, form a multi-enzyme complex. HTP, 4-hydroxy-2,5,6-triaminopyrimidine; DAU, 5,6-diaminouracil.

Table 1

Kinetic constants of bacterial and plant COG3236 enzymes

Results are means $\pm$ S.E.M. of three replicates. The k_{cat} values for the full length proteins are based on the molecular mass of the COG3236 domain.

Enzyme	Intermediate 1			Intermediate 2		
	k_{cat} (s^{-1})	K_{m} (mM)	$k_{\text{cat}}/K_{\text{m}}$ ($\text{s}^{-1} \text{M}^{-1}$)	k_{cat} (s^{-1})	K_{m} (mM)	$k_{\text{cat}}/K_{\text{m}}$ ($\text{s}^{-1} \text{M}^{-1}$)
<i>E. coli</i> YbiA	0.006 ± 0.0002	4.04 ± 0.32	1	0.80 ± 0.1	5.84 ± 1.09	140
<i>V. vulnificus</i> COG domain	42.5 ± 2.48	0.45 ± 0.07	94,400	0.03 ± 0.002	0.49 ± 0.10	60
<i>V. vulnificus</i> full length	137 ± 6	1.09 ± 0.14	125,000	0.07 ± 0.003	0.51 ± 0.07	130
Maize COG domain	0.03 ± 0.001	$5.56 \pm 0.49^*$	5	4.74 ± 0.19	1.08 ± 0.11	4,390
Maize full length	$5.00^{\dagger} \pm 0.19$	0.94 ± 0.09	5,300	$9.39^{\ddagger} \pm 0.36$	0.19 ± 0.03	50,300

* Two replicates only.