Evolution of Vitamin B₂ Biosynthesis

STRUCTURAL AND FUNCTIONAL SIMILARITY BETWEEN PYRIMIDINE DEAMINASES OF EUBACTERIAL AND PLANT ORIGIN

Markus Fischer‡, Werner Römisch, Sabine Saller, Boris Illarionov, Gerald Richter, Felix Rohdich, Wolfgang Eisenreich, and Adelbert Bacher

From the Lehrstuhl für Organische Chemie und Biochemie, Technische Universität München, Lichtenbergstrasse 4, Garching D-85747, Germany

Received for publication, April 21, 2004, and in revised form, June 16, 2004
Published, JBC Papers in Press, June 18, 2004, DOI 10.1074/jbc.M404406200

The Arabidopsis thaliana open reading frame At4g20960 predicts a protein whose N-terminal part is similar to the eubacterial 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate deaminase domain. A synthetic open reading frame specifying a pseudomature form of the plant enzyme directed the synthesis of a recombinant protein which was purified to apparent homogeneity and was shown by NMR spectroscopy to convert 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate into 5-amino-6-ribosylamino-2,4(1H,3H)-pyrimidinedione 5'-phosphate at a rate of 0.9 μmol mg⁻¹ min⁻¹. The substrate and product of the enzyme are both subject to spontaneous anomerization of the ribosyl side chain as shown by ¹³C NMR spectroscopy. The protein contains 1 eq of Zn²⁺/subunit. The deaminase activity could be assigned to the N-terminal section of the plant protein. The deaminase domains of plants and eubacteria share a high degree of similarity, in contrast to deaminases from fungi. These data show that the riboflavin biosynthesis in plants proceeds by the same reaction steps as in eubacteria, whereas fungi use a different pathway.

Flavocoenzymes derived from riboflavin (vitamin B₂; compound 6, Fig. 1) serve as essential redox cofactors in all cells and have also been shown to be involved in a variety of nonredox processes such as light sensing (1, 2), photorepair of DNA (3), circadian time keeping, and bioluminescence (4–7). Whereas plants and many microorganisms obtain riboflavin by biosynthetic intermediate 4, which is common to the bacterial and fungal pathway, is dephosphorylated by an unknown process affording 5-amino-6-ribosylamino-2,4(1H,3H)-pyrimidinedione which is condensed with 3,4-dihydroxy-2-butanone 4-phosphate biosynthesized from ribulose 5-phosphate (13–27). The resulting 6,7-dimethyl-8-ribityllumazine (5) undergoes a mechanistically unusual dismutation affording riboflavin (6) and the dephosphorylated biosynthetic intermediate 4 (19, 28–32). The biosynthesis of riboflavin in plants has been studied in some detail. A gene specifying a bifunctional GTP cyclohydrolase II/dihydroxybutanone phosphate synthase was cloned from tomato (33, 34). 6,7-Dimethyl-8-ribityllumazine synthase was cloned from spinach, and its structure has been determined by x-ray crystallography (35). Riboflavin synthase was partially purified from spinach (36), and a riboflavin synthase gene was later cloned from Arabidopsis thaliana (37).

The conversion of the GTP cyclohydrolase II product 1 into the substrate of lumazine synthase, 5-amino-6-ribosylamino-2,4(1H,3H)-pyrimidinedione, has not been studied in plants to the best of our knowledge. In this paper, we show that a protein specified by the hypothetical gene At4g20960 of A. thaliana catalyzes the deamination of 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate (1). Thus, riboflavin biosynthesis in plants proceeds by the same reaction steps as in eubacteria, whereas fungi use a different pathway.

EXPERIMENTAL PROCEDURES

Materials—The preparation of [2,1,2,3,4,5,13C₆]GTP will be described elsewhere. Restriction enzymes were obtained from New England Biolabs and Amersham Biosciences. T4 DNA ligase and reverse transcriptase (Superscript II) were from Invitrogen. Oligonucleotides were custom synthesized by MWG Biotech (Ebersberg, Germany) and by Interactiva (Ulm, Germany). Vent DNA polymerase was purchased from New England Biolabs. A plasmid mini preprep kit from PEQLab (Erlangen, Germany) was used for plasmid DNA isolation and purification. DNA fragments and PCR amplificates were purified with a gel extraction kit or Cycle Pure Kit from PEQLab. Casein hydrolysate and yeast extract were from Invitrogen, and isoprppyl β-D-thiogalactopyranoside was from Biomol (Hamburg, Germany). Thrombin was from Sigma. The isolation of recombinant Escherichia coli GTP cyclohydrolase II has been described previously (12).

Microorganisms—Bacterial strains and plasmids used in this study are summarized in Table I.

Restriction Enzyme Digestion of DNA—DNA was digested at 37 °C with restriction enzymes in reaction buffers supplied by the supplier. The treated DNA was analyzed by horizontal electrophoresis in 0.8–3% agarose gels.

Estimation of Protein Concentration—Protein concentration was estimated by a dye binding assay (38) or photometrically (ε₅₀₀₅₀ for recombinant A. thaliana deaminase, 10,600 cm⁻¹ M⁻¹).

Transformation of Bacterial Cells—Ligation mixtures were trans-
TABLE I  

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Relevant characteristics</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli strains</td>
<td></td>
<td></td>
</tr>
<tr>
<td>XL1-blue</td>
<td>recA1 endA1 gyrA96 thi^{-1} hsdR17 supE44 relA1 lac [F9proAB lacI^Q ZDM15 Tn10 (Tet^R)]</td>
<td>(39)</td>
</tr>
<tr>
<td>M15[pREP4]</td>
<td>lac ara gal mtl recA^- uvr^- (pREP4; lacI Kan^-)</td>
<td>(61)</td>
</tr>
<tr>
<td>RR28 rib1</td>
<td>F^- thi pro lac gal ara mtl xyl supE44 endA hsd (r^- m^-) pheS recA ribD</td>
<td>(8)</td>
</tr>
<tr>
<td>Plasmids</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PNCO113</td>
<td>Plasmid cloning vector, Ap^R</td>
<td>(40)</td>
</tr>
<tr>
<td>PACYC184</td>
<td>Plasmid cloning vector, Cm^R</td>
<td>NEB^*</td>
</tr>
<tr>
<td>PNCO-G-Red</td>
<td>Plasmid carrying the gene coding for the reductase domain of B. subtilis, based on pNCO113</td>
<td>(11)</td>
</tr>
<tr>
<td>PNCO-MalE</td>
<td>pNCO113 containing the malE gene of E. coli</td>
<td>(41)</td>
</tr>
<tr>
<td>PNCO-ATTR2-wt</td>
<td>Plasmid carrying the wild type A. thaliana deaminase gene, based on pNCO113</td>
<td>This study</td>
</tr>
<tr>
<td>PACYC-BSRED</td>
<td>Plasmid carrying the reductase domain of B. subtilis, based on plasmid PACYC184</td>
<td>This study</td>
</tr>
<tr>
<td>PNCO-ATRib2-syn</td>
<td>pNCO113 carrying a synthetic gene coding for the A. thaliana deaminase</td>
<td>This study</td>
</tr>
<tr>
<td>PNCO-MalE-TB-ATR2-syn</td>
<td>Plasmid coding coding for a fusion protein consisting of MBP of E. coli and the deaminase of A. thaliana, based on pNCO-MalE</td>
<td>This study</td>
</tr>
<tr>
<td>PNCO-MalE-TB-ATR2-184</td>
<td>Plasmid coding for a fusion protein consisting of MBP of E. coli and a deaminase subdomain of A. thaliana (amino acids 64–184), based on pNCO-MalE-TB-ATR2-syn</td>
<td>This study</td>
</tr>
<tr>
<td>PNCO-MalE-TB-ATR2-250</td>
<td>Plasmid coding for a fusion protein consisting of MBP of E. coli and a deaminase subdomain of A. thaliana (amino acids 64–250), based on pNCO-MalE-TB-ATR2-syn</td>
<td>This study</td>
</tr>
</tbody>
</table>

*New England Biolabs.

Evolution of Vitamin B2 Biosynthesis

FIG. 1. Biosynthesis of riboflavin. Deaminase (A and D) and reductase (B and C) reaction steps are shown. The biosynthetic pathway proceeds via A and B in bacteria and via C and D in yeasts (11, 52). 1. 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate; 2. 5-amino-6-ribosylamino-2,4(1H,3H)-pyrimidinedione 5'-phosphate; 3. 2,5-diamino-6-ribitylaminol-4(3H)-pyrimidinone 5'-phosphate; 4. 5-amino-6-ribitylaminol-2,4(1H,3H)-pyrimidinedione 5'-phosphate; 5. 6,7-dimethyl-8-riboflavamine; 6. riboflavin.

Preparation of cDNA—A 20-μl reaction mixture containing 50 mM Tris-HCl, pH 8.3, 75 mM KCl, 3 mM MgCl2, 10 mM dithiothreitol, 0.5 mM dNTPs, 0.5 μg of oligo(dT)15, 2 μg of A. thaliana total RNA from leaves, and 200 units of reverse transcriptase was incubated at 37 °C for 15 min and subsequently at 48 °C for 30 min. The mixture was heated at 95 °C for 5 min.

cDNA Cloning—A segment of the hypothetical open reading frame (accession no. At4g20960) was amplified from A. thaliana cDNA using the oligonucleotides DN1 and DC1 (Table II) as primers whereby the first 189 bp coding for a putative targeting sequence (63 amino acids) were removed. The amplificate served as template in a second PCR using the oligonucleotides DC1 and DC2 as primers. A 1.3-kb DNA segment was purified by agarose gel electrophoresis, digested with the restriction endonucleases BamHI and PstI, and ligated into the pNCO113 vector, which had been treated with the same enzymes yielding the plasmid designated pNCO-ATTR2-wt.

Gene Synthesis—The protein sequence coded by the pseudomatured open reading frame At4g20960 starting at methionine 64 (without targeting sequence) was retranslated into DNA sequence optimized for the expression in E. coli and divided into two segments (part A, bp 1–567; part B, bp 595–1125). Both segments were synthesized by a series of consecutive PCR amplifications and were then ligated. For both segments, synthetic starter DNA fragments (~100 bp) were prepared by PCR with the overlapping oligonucleotides AT-R2-1a and AT-R2-2a (segment A) and AT-R2-1b and AT-R2-2b (Fig. 2, segment B). In a sequence of seven PCR amplifications (segment A), respectively, six PCR amplifications (segment B), the oligonucleotides listed in Table II were used pairwise, starting with AT-R2-3a and AT-R2-4a (segment A) and AT-R2-3b and AT-R2-4b (segment B), for the elongation of each prior amplify. The final PCR product corresponding to segment A (609 bp) was digested with the restriction endonucleases EcoRI and PstI. The final PCR product corresponding to segment B (552 bp) was digested with PstI and HindIII. Both segments were combined and ligated into the plasmid pNCO113 (18, 38), which had been treated with the restriction enzymes EcoRI and HindIII, yielding the plasmid designated pNCO-ATR2-syn.

Construction of a Plasmid Expressing the Plant Deaminase Fused to the C Terminus of Maltose-binding Protein of E. coli—The synthetic gene coding for the plant deaminase was amplified by PCR using the oligonucleotides ATR2-NotI-Vo and ATR2-HindIII-Hi as primers (Table II) and the plasmid vector pNCO-ATR2-syn as template. The resulting 1119-bp fragment was purified and digested with NotI and HindIII and was then ligated into the plasmid pNCO-MalE (41), which had been treated with the same restriction enzymes yielding the plasmid designated pNCO-MalE-TB-ATR2-syn (Fig. 3).

Construction of Plasmids Expressing Subdomains of the Putative Deaminase of A. thaliana—The expression plasmid pNCO-MalE-TB-ATR2-syn was cleaved with the restriction endonucleases BamHI and HindIII, and a synthetic double-stranded oligonucleotide designated as ATR2-184 was ligated into the vector yielding the plasmid pNCO-MalE-TB-ATR2-184. To generate a further construct the plasmid pNCO-MalE-TB-ATR2-syn was digested with the restriction endonucleo-
Evolution of Vitamin B<sub>2</sub> Biosynthesis

<table>
<thead>
<tr>
<th>Designation</th>
<th>Sequence (5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DN1</td>
<td>gtaaaagcccaaagatgtaaatcgat</td>
</tr>
<tr>
<td>DC1</td>
<td>gactgtgatcctatcatcagaaacag</td>
</tr>
<tr>
<td>DN2</td>
<td>gaaacctgacagtctgattcccaattgg</td>
</tr>
<tr>
<td>DC2</td>
<td>gaaacctgacaatgctgacttcagccgtc</td>
</tr>
<tr>
<td>AT-R2-2a</td>
<td>gccaccttcatgttttcggggctgctcctccagccctgcaac</td>
</tr>
<tr>
<td>AT-R2-2b</td>
<td>cccacctcaaataaacacccgcaccccttcagggctggttc</td>
</tr>
<tr>
<td>AT-R2-3b</td>
<td>gattactccgccgattgatgtcttttgagccctttgatatgattt</td>
</tr>
<tr>
<td>AT-R2-4b</td>
<td>gaaacctgacttcagaagttctccgctgattgtagttctctccagcc</td>
</tr>
<tr>
<td>AT-R2-5b</td>
<td>catatcctgtaatccaccccaaggcttgagcctggccacggtgtagttccagcc</td>
</tr>
<tr>
<td>AT-R2-6b</td>
<td>tcgtgccacattccagcctcgatgtcctcctccagccaggt</td>
</tr>
<tr>
<td>AT-R2-7b</td>
<td>caggtgcctgttgctcgcacccactgctgcctcctccagccaggt</td>
</tr>
<tr>
<td>AT-R2-8b</td>
<td>tcgatacctgggctgacttcgagccctgtccgctccatgtactgatc</td>
</tr>
<tr>
<td>AT-R2-9b</td>
<td>ccagatcgacttcgacgctgcctccatgtactgacacttgac</td>
</tr>
<tr>
<td>AT-R2-10b</td>
<td>ccagatcgacttcgacgctgcctccatgtactgacacttgac</td>
</tr>
<tr>
<td>AT-R2-11b</td>
<td>caggtgcctgattcgcgtccctccagccaggt</td>
</tr>
<tr>
<td>AT-R2-12b</td>
<td>tataattactcgacagatgtgctaatccctgctggtggtgatctgac</td>
</tr>
<tr>
<td>ATR2-184-Fw</td>
<td>gattcccaacctatggtttttctcgggttagttgtaacagctgtaga</td>
</tr>
<tr>
<td>ATR2-184-Rv</td>
<td>isagcctactgttccttttgaacggtacttagaacctgtaga</td>
</tr>
<tr>
<td>ATR2-250-Fw</td>
<td>gcagaatttacagtctcgccacggtacttagaacctgtaga</td>
</tr>
<tr>
<td>ATR2-250-Rv</td>
<td>gacagcttacagctcgccacggtacttagaacctgtaga</td>
</tr>
<tr>
<td>ATR2Not1-Vo</td>
<td>atataaatgagctcctaatggtttttctcgggttagttgtaacagctgtaga</td>
</tr>
<tr>
<td>ATR2-HindIII-Hi</td>
<td>tattattactcgacagatgtgctaatccctgctggtggtgatctgac</td>
</tr>
</tbody>
</table>

Complementation Analysis of ribD Mutation in E. coli—The expression plasmids pACYC-BSRED in combination with pNCO-ATR2-wt were transformed into E. coli Rib1 cells (containing a mutation in the ribD gene) (11). 15 mg liter<sup>-1</sup> chloramphenicol and 170 mg liter<sup>-1</sup> ampicillin were added to secure the maintenance of both plasmids in the host strain. Aliquots were plated on LB medium containing the appropriate antibiotics. The appearance of transformants was checked after incubation at 30 °C for 24 h.

Partial Proteolysis—The combined fractions from the amylase affinity chromatography were supplemented with 2.5 mM CaCl<sub>2</sub> and 1 unit of thrombin/mg of protein. The mixture was incubated at room temperature for 16 h, concentrated to a volume of 5 ml, and placed on a column of Superdex 75 26/60 (Amersham Biosciences) which was then developed with buffer B containing 50 mM Tris-HCl, pH 7.4, 100 mM KCl, and 1 mM MgCl<sub>2</sub>. Fractions were combined and concentrated by ultrafiltration.

Native Molecular Mass—Native molecular mass was estimated using a Pharmacia FPLC system equipped with a Superose 6 HR (30 × 1.6 cm) column (Amersham Biosciences). The elution buffer contained 50 mM Tris-HCl, pH 7.4, 100 mM KCl, and 1 mM MgCl<sub>2</sub>. The column was calibrated using the following standard proteins: RNase A (13.7 kDa), chymotrypsinogen A (25.0 kDa), ovalbumin (43.0 kDa), bovine serum albumin (67.0 kDa), aldolase (158 kDa), catalase (232 kDa), and apoferritin (440 kDa).

DNA Sequencing—Sequencing was performed by the Sanger dideoxy chain termination method (42). Plasmid DNA was isolated from 5-ml cultures of XL1-Blue strains grown overnight in LB medium containing 170 mg/liter ampicillin using a plasmid mini-preparation kit from PEQLab. Custom sequencing was performed by MWG Biotech.

SDS-PAGE—SDS-PAGE was performed by procedures published previously (43).

Protein Sequencing—Sequence determination was performed by the automated Edman method using a 471 A protein sequencer (PerkinElmer Life Sciences).

Atomic Absorption Spectroscopy—Purified protein (4.8 mg/ml) was dialyzed against 50 mM Tris-HCl, pH 8.0. The solution was analyzed in a graphite furnace atomic absorption spectrophotograph Varian spectrAA-100 (Varian Deutschland GmbH, Darmstadt, Germany). The dialysis buffer was used as reference, and 0.1 mM zine sulfate was used as standard.
Evolution of Vitamin B\textsubscript{2} Biosynthesis

Fig. 2. Strategy for preparation of the synthetic gene coding for \textit{A. thaliana} 2,5-diamino-3-riboylamino-4(3H)-pyrimidinone 5′-phosphate deaminase with improved codon usage for the expression in \textit{E. coli}. Changed codons are marked by black boxes. New recognition sites for singular restriction endonucleases are drawn in italics. Gray boxes indicate sequence parts that contribute to the restriction site but are also present in the wild type gene.

Electrospray Mass Spectroscopy—Experiments were performed as described elsewhere (44).

NMR Spectroscopy—Samples were dissolved in 0.5 ml of 50 mM Tris, pH 7.8, containing 10 mM MgCl\textsubscript{2}, 5 mM dithiothreitol, and 10% D\textsubscript{2}O. \textsuperscript{1}H and \textsuperscript{13}C spectra were acquired with a DRX 500 spectrometer from Bruker Instruments, Karlsruhe, Germany, at transmitter frequencies of 500.13 and 125.76 MHz, respectively. The standard temperature was 17 °C, and 37 °C for kinetic measurements. Two-dimensional HMOC,\textsuperscript{1} TOCSY, and INADEQUATE spectra were measured using standard Bruker software (XWINNMR 3.0). The mixing time was 60 ms in the \textsuperscript{13}C TOCSY experiment. Composite pulse decoupling was used for \textsuperscript{13}C NMR measurements. 3-(Trimethylsilyl)-1-propanesulfonate served as an external standard for \textsuperscript{1}H and \textsuperscript{13}C NMR measurements.

Sequence Alignment and Phylogenetic Analysis—Amino acid sequences of \textit{E. coli} (accession no. P25539) and \textit{B. subtilis} (accession no. P17618) riboflavin type deaminase/reductase genes were used to query nonredundant NCBI (www.ncbi.nlm.nih.gov/) and plant cDNA/expressed sequence tag data bases (www.plantgdb.org/cgi-bin/PlantGDB/blast) using the BLAST algorithm (45). Protein sequences were aligned using ClustalW (clustalw.genome.ad.jp/) and edited by hand to remove positions with gaps. Phylogenetic analysis was carried out based on an alignment using 54 translated eubacterial and 9 plant genes and the program package TreeTop (www.genebee.msu.su/services/tree_top.html). The length of the alignment has been adapted to the deaminase domain of \textit{A. thaliana} starting from position 64 to position 250, which can fold independently, yielding enzymatically active protein.

RESULTS

The product of a putative open reading frame of \textit{A. thaliana} (accession no. At4g20960 (46)) shows similarity to the ribG gene specifying the deaminase domain of the bifunctional pyrimidine deaminase/reductase of \textit{B. subtilis} (Fig. 4) (11). A hypothetical plastid targeting sequence of about 63 amino acid residues precedes the plant deaminase domain. Sequence analysis with the software package ChloroP (47, 48) suggested that the gene comprises an N-terminal plastid targeting sequence; this agrees well with the earlier finding that several other enzymes of the riboflavin pathway of plants carry typical plastid targeting sequences (33, 34, 37, 49). The \textit{in silico} analysis with software SignalP suggested a cutting site between the amino acid residues 63 and 64. The mature form should thus begin with the N-terminal sequence MRREEDEVDDSFY.

Based on these findings, a DNA segment extending from bp 190 to 1280 (deaminase domain without the putative plastid targeting sequence) was amplified from \textit{A. thaliana} cDNA and was cloned into the pNCO113 plasmid vector. The nucleotide sequence of the cloned cDNA segment showed no differences by comparison with the previously published genomic sequence data and confirmed the predicted intron location in the chromosomal gene located on \textit{A. thaliana} chromosome 4 from 11211957 to 11213651 (predicted exon topology, 1–500, 597–895, 973–1695) (46). A recombinant \textit{E. coli} strain harboring the plant gene segment on an expression plasmid under the control of a T5 promoter and lac operator produced only small amounts of the predicted protein.

Because the wild type plant gene under study comprises numerous codons that are known to be poorly transcribed in \textit{E. coli}, the yield of expressed recombinant enzyme was very low. To overcome these problems we synthesized a synthetic gene whose sequence was optimized for expression in \textit{E. coli} (Fig. 2). The aim was to produce sufficient amounts of the plant protein required for a detailed functional characterization and, ultimately, crystallization studies. For that purpose, 87 codons were replaced. Moreover, 20 new unique restriction sites were implemented (Fig. 2). The optimized sequence has been deposited in GenBank (accession no. AY456384).

This optimized heterologous system showed a very high expression level. Unfortunately, a strong tendency of the protein to accumulate in inclusion bodies limited the final recovery level. To minimize the aggregation behavior, a fusion protein consisting of the maltose-binding protein of \textit{E. coli} fused to the N terminus of the plant deaminase was constructed. Between both domains a short linker sequence including a thrombin cleavage site was introduced (Fig. 3). A recombinant \textit{E. coli} strain harboring a plasmid with that sequence under the control of a T5 promoter and a lac operator produced large amounts of a soluble recombinant protein with an apparent mass of 83 kDa as judged by SDS-PAGE (Fig. 5).

\textsuperscript{1}The abbreviations used are: HMOC, heteronuclear multiple quantum coherence; MBP, maltose-binding protein; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total correlation spectroscopy; INADEQUATE, incredible natural abundance double quantum transfer experiment.
The recombinant fusion protein was purified by affinity chromatography on amylose affinity matrix. The maltose binding domain was then cleaved off by thrombin digestion, and the pseudomature deaminase domain was isolated in apparent purity form by gel permeation chromatography (Fig. 5).

Partial Edman degradation afforded the expected N-terminal sequence GRREEDVEDDSFY. Mass spectrometry afforded a relative mass of 39.650 Da in good agreement with the calculated mass of 39.655 Da. Whereas the engineered protein is believed to start with Glycine because of the introduction of the thrombin cleavage site, the wild type protein carries Met in the corresponding position.

In the pseudomature deaminase domain of Arabidopsis thaliana one zinc ion/subunit was found by atomic absorption spectroscopy.

Earlier, we could show by NMR analysis that the product of GTP cyclohydrolase II reaction mixtures, the signal at 151.2 ppm, resonates at 151.2 ppm. Whereas the engineered protein is believed to start with Glycine because of the introduction of the thrombin cleavage site, the wild type protein carries Met in the corresponding position. In the pseudomature deaminase domain of Arabidopsis thaliana one zinc ion/subunit was found by atomic absorption spectroscopy.

Earlier, we could show by NMR analysis that the product of GTP cyclohydrolase II reaction mixtures, the signal at 151.2 ppm, resonates at 151.2 ppm. Whereas the engineered protein is believed to start with Glycine because of the introduction of the thrombin cleavage site, the wild type protein carries Met in the corresponding position.

In the pseudomature deaminase domain of Arabidopsis thaliana one zinc ion/subunit was found by atomic absorption spectroscopy.

Earlier, we could show by NMR analysis that the product of GTP cyclohydrolase II reaction mixtures, the signal at 151.2 ppm, resonates at 151.2 ppm. Whereas the engineered protein is believed to start with Glycine because of the introduction of the thrombin cleavage site, the wild type protein carries Met in the corresponding position.

In the pseudomature deaminase domain of Arabidopsis thaliana one zinc ion/subunit was found by atomic absorption spectroscopy.

Earlier, we could show by NMR analysis that the product of GTP cyclohydrolase II reaction mixtures, the signal at 151.2 ppm, resonates at 151.2 ppm. Whereas the engineered protein is believed to start with Glycine because of the introduction of the thrombin cleavage site, the wild type protein carries Met in the corresponding position.

In the pseudomature deaminase domain of Arabidopsis thaliana one zinc ion/subunit was found by atomic absorption spectroscopy.

Earlier, we could show by NMR analysis that the product of GTP cyclohydrolase II reaction mixtures, the signal at 151.2 ppm, resonates at 151.2 ppm. Whereas the engineered protein is believed to start with Glycine because of the introduction of the thrombin cleavage site, the wild type protein carries Met in the corresponding position.

In the pseudomature deaminase domain of Arabidopsis thaliana one zinc ion/subunit was found by atomic absorption spectroscopy.
Evolution of Vitamin B<sub>2</sub> Biosynthesis

signals at 156.1 and 155.7 ppm (deaminase product 2, Figs. 1 and 6). Notably, the signal at 151.2 ppm disappears more rapidly than that at 156.1 ppm, and the signal at 156.1 ppm appears more rapidly compared with that at 151.7 ppm. The signals of the enzyme product at 156.1 and 155.7 ppm appear broader than those of the substrate at 151.2 and 151.6 ppm.

The signals of the A. thaliana deaminase product obtained from [2,1′,2′,3′,4′,5′,13C]<sub>6</sub>1 show are shown in Fig. 7A. The complex multiplet patterns arise by 13C-13C coupling. Notably, the 13C NMR signals observed in the spectrum of a conversion of [2,1′,2′,3′,4′,5′,13C]<sub>6</sub>1 with the bifunctional deaminase/reductase from E. coli are apparently identical (Fig. 7B). This indicates that the A. thaliana deaminase product is identical with that obtained from the E. coli enzyme, namely 5-aminoo-6-ribosylamino-2,4(1H,3H)-pyrimidinedione 5′-phosphate (11). This conclusion is confirmed by a detailed NMR analysis including two-dimensional 1H-13C and 13C-13C correlation experiments (Table III). Two-dimensional INADEQUATE measurements afforded the 13C-13C connectivities of carbon atoms 1′ with 2′ and of 4′ with 5′ of both anomers (Fig. 8). Two-dimensional 13C-13C TOCSY spectroscopy confirmed these correlations and afforded additional 13C correlations conducive to the complete assignments of both spin systems (Fig. 9 and Table III). On the basis of kinetic arguments (for details, see above), the C-2 signals at 155.7 and 156.1 ppm can be assigned to the α and β anomers, respectively.

The 1H-13C connectivities in the carbon moieties of the glycosides 2α and 2β can be assigned on the basis of the 1H-13C correlation spectra. The HMQC spectrum (Fig. 10) shows cross-signals for pairs of directly connected 1H-13C pairs. The HMQC-TOSCY spectrum identifies pairs of hydrogen atoms on vicinal carbon atoms. The chemical shifts of both respective anomers are summarized in Table III.

The sequence of events following the addition of deaminase to 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinedione 5′-phosphate produced by the action of GTP cyclohydrolase II product is summarized in Reaction 1.

\[
\begin{align*}
1\alpha & \xrightarrow{k_1} 1\beta & 1\beta & \xrightarrow{k_3} 2\beta & 2\beta & \xrightarrow{k_5} 2\alpha & k_4
\end{align*}
\]

\text{REACTION 1}

This kinetic network can be described by the following set of rate equations.

\[
\begin{align*}
d([1\alpha]/dt) & = -[1\alpha]k_1 + [1\beta]k_2 & \text{(Eq. 1)}
d([1\beta]/dt) & = -[1\alpha]k_1 - [1\beta]k_2 - [1\beta]k_3 & \text{(Eq. 2)}
d([2\beta]/dt) & = -[2\beta]k_4 + [2\alpha]k_5 & \text{(Eq. 3)}
d([2\alpha]/dt) & = -[2\beta]k_4 - [2\alpha]k_5 & \text{(Eq. 4)}
\end{align*}
\]

The quantitative analysis of the signals in the time-resolved NMR spectra affords the concentrations of all reactants under study. Fitting these data to the set of rate equations yields the data shown in Fig. 11 and Table IV. Notably, the rate constants \(k_1\) and \(k_3\) characterizing the isomerization reactions are all similar, in the range of 0.11–0.20 min<sup>-1</sup>. The specific activity of the plant deaminase is 0.9 μmol mg<sup>-1</sup> min<sup>-1</sup>.

To obtain more evidence on the evolution of the riboflavin biosynthetic pathway of plants, we performed a comprehensive sequence data base search. A BLAST search with the A. thaliana deaminase domain studied in this paper as the search motif revealed significant sequence similarity with the N-terminal section of the bifunctional RibD protein of B. subtilis and of putative orthologs from a wide variety of eubacteria and from several plant species (Fig. 4).

An initial complementation assay showed that a chromosomal ribD mutant strain (Rib1) could not be reconstructed by the putative A. thaliana gene. More sophisticated complementation analysis showed that the phenotype of the mutant strain could be restored by plasmids directing the synthesis of the plant deaminase in combination with the eubacterial reductase domain. This means that the putative A. thaliana protein failed to catalyze the reduction of compound 2 (Fig. 1). In a cross-checking experiment the A. thaliana gene was coexpressed with the fungal type (Saccharomyces cerevisiae) reductase (RIB7) and deaminase (RIB2) gene. In both cases no complementation could be observed. This fact is remarkably well in line with the data base and the NMR analysis and means that the second step in the riboflavin pathway of A. thaliana is identical to the reaction found in eubacteria.

The N-terminal part of RibD proteins of eubacterial organisms is known to harbor the deaminase domain (11). Based on sequence alignments with functional deaminase domains of B. subtilis, C-terminal truncated plant proteins were constructed (Fig. 3). A subdomain consisting of amino acid residue 64–250 of the pseudomature plant enzyme was found to be sufficient for deaminase activity. These data suggest that the putative...
Fig. 7. 13C NMR signals of the product mixture obtained by treatment of [2,1'2,3',4',5',13C6]GTP with GTP cyclodihydrolase II and deaminase from A. thaliana (A) or bifunctional deaminase/reductase from E. coli (B).

Evolution of Vitamin B2 Biosynthesis

Remarkably, 1 of 17 completely sequenced archaebacterial species (Archaea) carries a putative gene with close neighborhood to cyanobacteria, plant pathogens, and soil bacteria.

open reading frame of A. thaliana directs the synthesis of a deaminase subdomain, which can fold independently.

Extended sequence analysis showed homology to a wide variety of eubacterial and plant sequences. The similarity between the cognate protein sequences is illustrated by the denrogram in Fig. 12. Notably, all plant deaminase sequences form a cluster in the phylogenetic tree. It is also worth noting that Gram-negative bacteria form a subcluster by themselves; the same is true for the group of Gram-positive eubacteria with only a few exceptions. The plant sequences are in close neighborhood to cyanobacteria, plant pathogens, and soil bacteria. Remarkably, 1 of 17 completely sequenced archaebacterial species (Pyrococcus furiosus) carries a putative gene with close similarity to the bifunctional RibD protein of B. subtilis and does not contain any putative orthologs of the Methanothermus jannaschii pyrimidine nucleotide reductase (50) which is believed to catalyze the fungal type reduction of the ribosyl side chain in Archaea. Notably, only one eubacterium (Buchnera aphidicola, subsp. Schizaphis graminum) carries separate genes for the deaminase (accession no. Q8K9A4) and the reductase (accession no. Q8K9A3), however, in a sequential order.

Pyrimidine deaminases of the riboflavin pathway show also sequence homology to yeast cytosine deaminase (Fig. 13). The enzyme has been found in prokaryotes and fungi, where it is a member of the pyrimidine salvage pathway. The x-ray structure of yeast cytosine deaminase has been determined in the
presence of an inhibitor at 1.14 Å resolution (51). Each active site contains a single catalytic zinc ion that is coordinated by a histidine, two cysteines, and a single bound water molecule that was found to act as a nucleophile in the deamination reaction. The coordination site for the zinc ion, which is known from the structure of the yeast cytosine deaminase, is strictly conserved over 54 bacterial and 9 plant pyrimidine deaminases of the riboflavin pathway analyzed in this study (Fig. 13). Some of the residues of the active site cavity of the cytosine deaminase are also present and highly conserved in riboflavin type deaminases, e.g., residues corresponding to His-62 (His-58 in the pseudomature A. thaliana deaminase), Glu-64 (Glu-60), Leu-88 (Leu-80), and Cys-91 (Cys-83). (Fig. 13). In addition, sequence similarity is also found to bacterial cytidine deaminase, e.g., of E. coli (accession no. P13652).

**TABLE IV**

Kinetic parameters of pyrimidine deaminase of A. thaliana obtained from $^{13}$C NMR spectroscopy

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Kinetic parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1\alpha \rightarrow 1\beta$</td>
<td>$k_1$</td>
<td>0.11 min$^{-1}$</td>
</tr>
<tr>
<td>$1\beta \rightarrow 1\alpha$</td>
<td>$k_2$</td>
<td>0.11 min$^{-1}$</td>
</tr>
<tr>
<td>$1\beta \rightarrow 2\beta$</td>
<td>$k_3$</td>
<td>1.34 min$^{-1}$ mg$^{-1}$</td>
</tr>
<tr>
<td>$2\beta \rightarrow 2\alpha$</td>
<td>$k_4$</td>
<td>0.20 min$^{-1}$</td>
</tr>
<tr>
<td>$2\alpha \rightarrow 2\beta$</td>
<td>$k_5$</td>
<td>0.18 min$^{-1}$</td>
</tr>
<tr>
<td>$1(\alpha\beta) \rightarrow 2(\alpha\beta)$</td>
<td>$V_{max}$</td>
<td>0.9 μmol min$^{-1}$ mg$^{-1}$</td>
</tr>
</tbody>
</table>

**FIG. 9.** Two-dimensional $^{13}$C$^{13}$C TOCSY spectrum of the deaminase products. The correlated signals of the two spin systems of $2\alpha$ and $2\beta$ are connected by dashed and solid lines, respectively.

**FIG. 10.** Two-dimensional $^1$H$^{13}$C correlation spectrum of the product mixture obtained by treatment of $[2,1,2',3',4',5',6',13C_2]$GTP with GTP cyclohydrolase II and deaminase from A. thaliana.

**FIG. 11.** Reaction dynamics of deaminase from A. thaliana. The time-resolved concentration was calculated from the integrals of the $^{13}$C-labeled C-2 carbon atoms of the reactants $1\alpha$ (○), $2\beta$ (■), $2\alpha$ (△), and $2\beta$ (△) (see Fig. 6).

**FIG. 12.** Phylogenetic tree of 2,5-diamino-6-ribosylamino-4(3H)-pyrimidinone 5'-phosphate deaminase domains.
Enzymes catalyzing the reduction of the ribose chain of 1 under formation of 3 have been partially purified from the flavogenic ascomycete Ashyha gossypii (53) and from Candida guillermondii (52). The enzyme catalyzing the reduction of 1 has been cloned from the yeast C. guillermondii (58). Genes involved in the riboflavin pathway of S. cerevisiae and A. gossypii have been claimed in patents (59, 60). Moreover, although reductase and deaminase occur as separate enzymes in S. cerevisiae and other fungi, bifunctional proteins with deaminase and reductase activity were found in the eubacteria, E. coli and B. subtilis (11). The deaminase as well as the reductase domain of the bifunctional enzyme of B. subtilis could be expressed separately as functional proteins (11).

The data reported in this paper show that the enzyme of A. thaliana deaminates the product of GTP cyclohydrolase II at a rate of 0.9 μmol mg⁻¹ min⁻¹. This rate is similar to that of the enzyme of E. coli (0.36 μmol mg⁻¹ min⁻¹) (15). The substrate and the product of deaminase are both subject to spontaneous epimerization. The equilibrium constant for each respective pair of anomers has a value close to 1, and the velocities of anomerization for the two anomer pairs are similar. Spontaneous hydrolysis of the N-glycosidic bond is slow compared with the epimerization reaction.

It has not been possible, up to now, to determine directly the absolute configuration of these anomers because NOESY spectroscopy gave insufficient results. It can be shown, however, on the basis of kinetic arguments, that the product of GTP cyclohydrolase II would anomerize its substrate, we therefore assume that all riboside type intermediates of riboflavin biosynthesis in plants have β stereochemistry.

These data show that plants use the same pathway as eubacteria (i.e. via 2) for the transformation of 1 into 4. On the other hand, fungi use a different pathway (via 3), and circumstantial evidence indicates that Archaea also use that pathway (50). But, to the best of our knowledge, nothing is known about the enzyme catalyzing the reduction of the ribosyl side chain in plants.

Acknowledgments—We thank Richard Feicht, Sebastian Schwamb, and Thomas Wojtulewicz for skillful help in protein preparation.

REFERENCES

Evolution of Vitamin B₂ Biosynthesis