Old substrates for new enzymes of terpenoid biosynthesis

Jörg Bohlmann[a,1] and Jonathan Gershenzon[b,1]

[a]Michael Smith Laboratories, University of British Columbia, 2185 East Mall, Vancouver, BC, Canada V6T 1Z4; and [b]Max Planck Institute for Chemical Ecology, Beutenberg Campus, Hans-Knoell-Strasse 8, 07745 Jena, Germany

Enzymologists have never had it so good. Thanks to advances in gene discovery, driven by high-throughput sequencing of genomes and transcriptomes, and improved heterologous expression of proteins, the functional characterization of new enzymes is generally straightforward: clone the gene, express the protein, assay for activity, and publish the results. However, without knowing the correct substrate to use in the assay such efforts can easily go astray. A paper by Schilmiller et al. (1) published in the latest issue of PNAS shows how one can elegantly avoid this pitfall by integrating genomic, genetic, enzymological, and metabolite-profiling approaches. In this important contribution, the authors report the discovery of a new substrate for enzymes of plant terpenoid biosynthesis.

Determining the correct substrate has historically been a major challenge for researchers of terpene biosynthetic enzymes in plants. Terpenes are an enormous class of plant metabolites with many diverse roles in growth, development, and resistance to environmental stresses (2). Terpenoids also have a myriad of applications as industrial biomaterials, including pharmaceuticals, fragrances and flavors, and insecticides (3), and some terpenoids may serve in the production of novel biofuels. The early steps of terpene biosynthesis in plants involve the formation and assembly of C_5 isopentenoid units that arise from two separate pathways. The C_5 units are polymerized into C_{10}, C_{15}, C_{20}, and larger diphosphate intermediates which then radiate out in a complex network of biosynthetic sequences forming thousands of products that vary with plant species, cell type, and organelle.

Monoterpenes are C_{10} terpene compounds that serve in plant defense, pollinator attraction, and other ecological roles. Their skeletons are formed by a family of enzymes known as monoterpene synthases (4, 5), but it has not always been easy to know which substrates these enzymes employ. Monoterpene synthases seem promiscuous in vitro, employing geranyl diphosphate (GPP), the ubiquitous C_{10} intermediate of the isoprenoid pathway in animals and bacteria, as well as neryl diphosphate (NPP), its Z-isomer (Fig. 1), and linalyl diphosphate, a tertiary isomer. For early researchers of monoterpene biosynthesis, NPP seemed a more
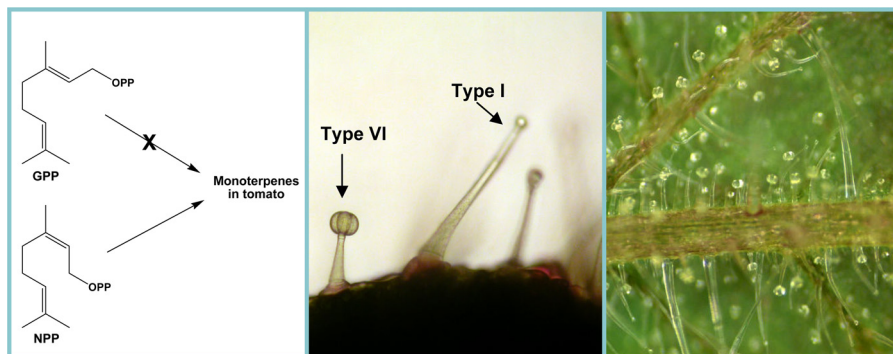

Fig. 1. Type I and type VI glandular trichomes on the stems and the leaf veins of the cultivated tomato (*Solanum lycopersicum*). The tiny type VI glandular hairs contain specialized cells for the formation of terpenoid volatile compounds. Using cell-type-specific ultrahigh-throughput transcriptome sequencing combined with metabolite profiling, genetic analysis, and biochemical characterization of heterologously expressed enzymes, Schilmiller et al. (1) discovered that neryl diphosphate (NPP), rather than the classical geranyl diphosphate (GPP), is the biologically relevant intermediate in the biosynthesis of tomato monoterpenes. Micrographs were taken by Anthony Schilmiller (Michigan State University, East Lansing).

likely candidate on chemical grounds than GPP as a substrate for monoterpene synthases, because GPP would first have to isomerize before cyclizing to form a 5- or 6-membered ring. In fact, it was once concluded from in vitro studies of crude preparations that NPP was a general substrate for monoterpene synthases (6, 7).

In the late 1970s, Rodney Croteau and his coworkers launched their landmark enzymological, mechanistic, structural, and molecular studies on monoterpene synthases. Using cell-free assays and recombinant proteins, they demonstrated that for many monoterpene synthases from plants as diverse as mints and conifers, GPP was probably the native substrate, because these enzymes could carry out the necessary isomerization before cyclization and kinetic parameters indicated that many monoterpene synthases actually worked more efficiently with GPP than NPP (6, 7).

However, NPP is apparently not ready to go quietly into the dustbin of biochemical history. Schilmiller et al. (1) have now marshaled convincing evidence that NPP is indeed the native substrate of a monoterpene synthase from tomato, a plant that accumulates a bouquet of volatile monoterpenes in glandular hairs on the surfaces of its leaves and stems (Fig. 1). Not only was the product profile of this enzyme with NPP much more closely correlated with the natural monoterpene spectrum of this particular variety of tomato than the product profile with GPP, but the smok-

ing gun was the authors' discovery of a gene encoding an NPP synthase (to the best of our knowledge, the first enzyme of its type) whose expression in tomato is restricted to glandular trichomes and resides in the same region of the tomato genome as the monoterpene synthase, in a chromosome position responsible for controlling monoterpene formation.

Why nature has room for both GPP and its *cisoid* isomer, NPP, in its stable of C_{10} terpene intermediates is not clear, although the use of separate substrates may allow more opportunities for regulating the biosynthesis of monoterpenes, reduce competition for formation of other terpenes, or alter the product profile of the enzyme. Many monoterpene synthases have the unusual ability to form multiple products from a single substrate (4, 5), and the product profile of a given monoterpene synthase varies with the type of substrate used in vitro (1). In any case, the discovery of a NPP-using monoterpene synthase in tomato accompanied by an NPP synthase is a major addition to textbook knowledge of terpene biosynthesis. This report should prompt a reassessment of the substrates of other terpene synthases that have proved inactive with GPP, had product spectra that correlated

poorly to those found in the plant, or produced mainly acyclic monoterpenes. It should also stimulate the search for terpene synthases using other novel substrates. Remarkably, earlier this year a tomato terpene synthase was reported that makes sesquiterpenes (C_{15}) in glandular trichomes and also uses a *cisoid* substrate, in this case (*Z,Z*)-farnesyl diphosphate (*Z,Z*-FPP) rather than the much more common (*E,E*)-FPP (8).

Surprisingly, the two new terpene synthases described by Schilmiller et al. (1) and Sallaud et al. (8) have more in common than their unexpected preference for *cisoid* over *transoid* substrates. Although one enzyme uses a C_{10} substrate (NPP) (1) and the other C_{15} (*Z,Z*-FPP) (8), both are members of the same subfamily (TPS-e) of the large and apparently monophyletic group of plant terpene synthases. Members of the TPS-e subfamily share sequence features that are reminiscent of ancestral plant terpene synthases (4). Hence, the biologically relevant use of *cisoid* C_{10} and C_{15} substrates by two terpene synthases in tomato may not be a new trick of nascent genes, but may actually mark

the (re)-discovery of new substrates by old enzymes in an evolutionary sense.

Unlike these two newly described tomato terpene synthases that use *cisoid* C_{10} and C_{15} substrates for the formation of

Nature has room for both GPP and its *cisoid* isomer, NPP, in its stable of C_{10} terpene intermediates.

specialized mono- and sesquiterpenes, several previously characterized members of the TPS-e subfamily use the C_{20} copalyl diphosphate substrate for the formation of intermediates in the biosynthesis of gibberellins, plant diterpene hormones. Therefore, the discovery of new enzymes of the TPS-e group with an unexpected preference for *cisoid* C_{10} and C_{15} substrates reinforces an important notion for the annotation of genes that code for enzymes in specialized plant metabolism

(also known as plant secondary metabolism): neither substrate specificity nor product profiles can be predicted by sequence gazing (e.g., BLAST searches or other means of bioinformatics sequence comparisons) or through guilt by association with family members. Instead, rigorous biochemical testing in a biologically relevant context can readily establish enzyme function, as illustrated by Schilmiller et al. (1).

This work also serves as a reminder that for many enzymes promiscuity in substrate acceptance in vitro may be only an illusion. In vivo the enzyme may never be tempted by more than one substrate partner. However, even with all of the molecular, biochemical, and genetic resources at our disposal, there is no magic shortcut for determining the actual substrate in vivo, and an integration of approaches is often required, as demonstrated here in exemplary fashion (1).

1. Schilmiller AL, et al. (2009) Monoterpenes in the glandular trichomes of tomato are synthesized from a neryl diphosphate precursor rather than geranyl diphosphate. *Proc Natl Acad Sci USA* 106:10865–10870.
2. Gershenzon J, Dudareva N (2007) The function of terpene natural products in the natural world. *Nat Chem Biol* 3:408–414.
3. Bohlmann J, Keeling CI (2008) Terpenoid biomaterials. *Plant J* 54:656–669.
4. Bohlmann J, Meyer-Gauen G, Croteau R (1998) Plant terpenoid synthases: Molecular biology and phylogenetic analysis. *Proc Natl Acad Sci USA* 95:4126–4133.
5. Tholl D (2006) Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. *Curr Opin Plant Biol* 9:297–304.
6. Croteau R (1987) Biosynthesis and catabolism of monoterpenoids. *Chem Rev* 87:929–954.
7. Wise ML, Croteau R (1998) Monoterpene biosynthesis. *Comprehensive Natural Product Chemistry: Isoprenoid Biosynthesis*, ed Cane DE (Elsevier, London), pp 97–153.
8. Sallaud C, et al. (2009) A novel pathway for sesquiterpene biosynthesis from *Z,Z*-farnesyl pyrophosphate in the wild tomato *Solanum habrochaites. Plant Cell* 21:301–317.