Industrial Crops & Products 124 (2018) 530–547

Contents lists available at ScienceDirect

Industrial Crops & Products

journal homepage: www.elsevier.com/locate/indcrop

Chemotypes and terpene synthase genes in *Thymus* genus: State of the art

Helena Trindade*, Luis Gaspar Pedro, Ana Cristina Figueiredo, José Gonçalves Barroso

Centro de Estudos do Ambiente e do Mar (CESAM Lisboa), Faculdade de Ciências da Universidade de Lisboa, Centro de Biotecnologia Vegetal (CBV), Edifício C2, Campo Grande, 1749-016 Lisboa, Portugal

ARTICLE INFO

Keywords:
Aromatic plants
Chemotypes
Essential oils
Terpene synthases
Thymus

ABSTRACT

Aromatic plants belonging to the Lamiaceae have been used in folk medicine as well as flavouring herbs. This family includes numerous species which are rich in essential oils, that have been studied from both the volatiles point of view and using molecular tools. *Thymus* is one important genus from this family, comprising about 215 species, whose essential oils have been used by the pharmaceutical industry given their biological activities, namely antibacterial, antifungal and anticancer activities. Essential oils are composed of a mixture of monoterpenes, among other constituents. Terpenoids show high structural variability and over 23.000 different metabolites are known. The occurrence of chemotypes in selected thymus species is listed and organized, and the concept of chemotype is addressed. A multiple approach, considering multivariate analysis of essential oil components and their metabolic relationship, is proposed for chemotype determination. Terpene synthase genes are described, with particular emphasis to the monoterpene synthases, its conserved motifs and features. The role of site directed mutagenesis and its contribution for understanding terpene synthases, revealing the role of some amino acids in the protein activity and final product synthesis is also explored. *T. caespititius*, a thyme species with carvacrol/thymol, α-terpineol chemotypes, among others, was chosen as an example to describe the state of the art on this topic. The high content of the phenolic monoterpenes makes this species interesting to explore as a source of compounds with antioxidant and antimicrobial properties. Two γ-terpinene synthases were identified in this thyme species, revealing the existence of two isogenes, one isoform directed to the plastids and the other probably cytosolic, indicating the complexity of terpene synthesis. Functional analysis suggests the existence of additional levels of regulation that might involve direct temperature effect, regulating protein activity and thereby affecting the final product. This opens new layers of regulation that will undoubtedly bring complexity to essential oil composition on aromatic plants. In this review, data available both on essential oils characterization and on monoterpene synthase genes for *Thymus* species is summarized, stressing the importance of integrating molecular and biochemical data on chemotype determination.

1. Introduction

Terpenoids are secondary metabolites not generally involved in plant growth and development, playing roles as attractants, repellents, toxins and antibiotics, or being precursors to such bioactive compounds (Gershenzon and Dudareva, 2007). Most of the terpenoids are considered part of the secondary metabolism, a term that has been proposed by the late 19th century to distinguish from the basic metabolism (Stahl, 1888 in Kliebenstein, 2004). In the true assumption, basic metabolism refers to all the processes, either anabolic or catabolic, required for cell maintenance and proliferation. Examples of such processes are respiration, nutrient assimilation, and growth/development. Secondary metabolites, on the other hand, refers to compounds produced in specialized cells, not required for the cells survival by itself, but thought to improve plant survival by contributing to a better fitness.

Emission of volatile compounds such as terpenoids allows plants to communicate, being important players in plant-insect, plant-pathogen, and plant-plant interactions (Dudareva et al., 2004; Paschold et al., 2006). They have been implicated in attraction of both pollinators and predators of phytophagous, in protection against photooxidative stress, in mediating thermotolerance, and in direct defense against microbes and insects (Figueiredo and Barroso, 2015; Tholl, 2006). They can thus act both in plant direct or indirect defense, in a complex response involving herbivores and their natural predators. The full extent of its ecological significance is not fully clear (Das et al., 2013) and is under the scope of chemical ecology, the science devoted for studying all these interactions. In nature, terpenes occur predominantly as hydrocarbons, alcohols and their glycosides, ethers, aldehydes, ketones, carboxylic acids and esters.

* Corresponding author.
E-mail addresses: htrindade@fc.ul.pt (H. Trindade), lmpedro@fc.ul.pt (L.G. Pedro), acsf@fc.ul.pt (A.C. Figueiredo), jgbarroso@fc.ul.pt (J.G. Barroso).

https://doi.org/10.1016/j.indcrop.2018.08.021
Received 7 May 2018; Received in revised form 27 July 2018; Accepted 8 August 2018

Aside from their important ecological role, terpenes extracted from plants have been used by mankind for many different purposes. They include fragrances and flavours, pharmaceutical agents and insecticides which, taken altogether, possess an immense commercial value.

The study on plant terpene metabolism has undergone a huge development with Croteau's pioneer work starting on the 70s of last century (Gershenzon, 2006). He conducted several experiments leading to the formation of many cyclic terpenes in *in vitro* systems, from the intermediate geranyl diphosphate (GPP). A tremendous amount of enzymatic studies was performed and helped on the elucidation of the mechanism of action of terpene synthases, including the complex carbocationic mechanism of the cyclizing processes that are responsible for generating such an enormous variability on terpenes. Furthermore, Croteau and collaborators also studied other families of terpene-forming catalysts, such as prenyltransferases, cytochrome P450s, and the enzymes of menthol biosynthesis, accomplishing 350 publications (Gershenzon, 2006).

Terpene synthesis starts from a common C_5 precursor, the universal C_5 building block, isopentenyl diphosphate (IPP) and the isomer dimethylallyl diphosphate (DMAPP) (Tholl, 2006). The biosynthesis of terpenes follows the isoprene rule, based on the addition of isoprene units $(C_5)_n$, that build up the carbon skeleton of terpenes. They can be coupled in chain elongation reactions to yield increasingly longer polyisoprenoid diphosphates which can then be cyclized to generate single- or multi-ringed products (Christianson, 2008). Terpenes synthesis can proceed either by the mevalonate pathway (MVA), which is a cytosolic route or through the methylerythritol pathway (MEP), occurring in the plastids. On the cytosol, IPP is produced *via* the condensation of two acetyl-CoA moieties, through the MVA pathway. The plastidic MEP pathway, starting with glyceraldehyde 3-phosphate and pyruvate, gives rise to DMAPP and later to volatile monoterpenes, C_{10}, and diterpenes, C_{20}, while the MVA pathway provides precursors to volatile sesquiterpenes, C_{15} (Dudareva et al., 2004) and the triterpenes, C_{30}. The mevalonate-independent pathway has been discovered to operate in certain bacteria and in the plastids of both lower and higher plants (Davis and Croteau, 2000).

Terpene structural diversity is generated by the modification of a common scaffold structure, to which diversity is added. Since the differential modification of common backbone structures can change the biological activity, the new compounds show a potential new biological activity.

Why has this modular diversity become a trait selected under natural conditions? One possible explanation is that plants with these newly derived defenses have better chances of survival against insects and other pests, which on the other hand have evolved the ability to overcome existing plant defenses (Ehrlich and Raven, 1964). This results in a scenario where plant and herbivore insects display a coevolutionary trend. These new compounds can be synthesized *in vivo* through the structural modification of an existing toxic compound to which the pest had already a counter-defense, but still maintaining the compound's toxic activity.

1.1. Terpene synthases

Terpene synthases, also known as terpene cyclases, are a metal dependent family of enzymes and constitute key enzymes for terpene biosynthesis. Typically, the monoterpene synthases catalyse the cyclisation of geranyl diphosphate *via* an α-terpinyl cation intermediate, or instead use the linear geranyl cation intermediate through elimination and addition reactions, resulting in a diverse selection of monoterpene products (Fig. 1). The sesquiterpene synthases use farnesyl diphosphate instead, while for triterpene synthases the substrate is geranyl geranyl diphosphate. Synthesis of sesqui- and tri-terpenes have a different subcellular location from mono-, di-, and tetra-terpenes, as described.

The biosynthesis of natural compounds starting from a limited stock of simple metabolites, ends in final products, with an exquisite chemodiversity (Gershenzon and Dudareva, 2007). The structural features of these enzyme catalysts are a major cause of terpene diversity. The intermediates have several different metabolic fates, leading to the synthesis of structurally diverse products. It is not usually possible to predict the product profile of terpene synthases based on the primary structure alone. Therefore, the elucidation of enzyme structure-function relationships depends on three-dimensional (3D) structures and the position of amino-acid residues to the catalytic process.

To date, many terpene synthases (TPSs) have been characterized, some catalyzing the formation of a single terpene compound, although many enzymes synthesize complex product mixtures with high regio- and stereospecifity (Davis and Croteau, 2000; Degenhardt et al., 2009; Tholl, 2006). The tremendous range of possible variations in the carbocationic reactions (cyclizations, hydride shifts, rearrangements, and termination steps) catalyzed by the TPSs explains the wide range of possible products (Fischer et al., 2013; Roeder et al., 2007; Trapp and Croteau, 2001). Furthermore, enzymes such as cytochrome P450 (CYPs) monooxygenases and oxidoreductases are also involved in further modifications of the terpene skeletons, yielding the extremely high diversity of terpene compounds found in nature (Crocoll et al., 2010; Daviet and Schalk, 2010).

The simplest form of terpenoid chemistry is found for monoterpenes. Monoterpene synthases are commonly found in plants, and recently they were found also to occur in several bacteria, as has been demonstrated through genome mining efforts (Yamada et al., 2015). Monoterpenes biosynthesis starts with the synthesis of geranyl pyrophosphate (GPP), the precursor of all monoterpenes, which gives rise to the α-terpenyl cation. This highly unstable intermediate can be later converted to specific monoterpenes, such as sabinene, γ-terpinene, or α-terpineol, among others, through the action of certain monoterpene synthases (Fig. 1). The active-site cavity, with its specific amino acids surrounding it, is extremely important in defining which monoterpene will be produced. Nevertheless, distant structural features also play a vital role in modulating the catalytic specificity in terpene synthases, as has been proved through mutational experiments performed on residues located away from the active site (Greenhagen et al., 2006). As a result, a certain combination of active-site residues may not lead to the same final products if based in different scaffolds (Kampranis et al., 2007).

Terpene synthases are proteins composed of 550–850 amino acid residues and contain a C-terminal active site domain and an N-terminal domain. Monoterpene synthases, containing between 600 and 650 amino acid residues, are larger than sesquiterpene synthases due to the presence of 50–70 amino acids functioning as a signal peptide, absent in sesquiterpene synthases (Bohlmann et al., 1998; Martin et al., 2004). Diterpene synthases are approximately 210 amino acids longer than monoterpene synthases, due to the presence of an internal element. This additional element has been shown to be conserved in both sequence and position (Bohlmann et al., 1998).

Sesquiterpene synthesis starts from the linear precursor, farnesyl diphosphate (FPP), generating more than 300 distinct cyclic sesquiterpenes. Sesquiterpene synthase active site provides a shaped cavity for the binding and orientation of the flexible substrate. There is a conserved mechanism for carbocation initiation and in despite the existence of structural similarity, sesquiterpene synthases share only a minimal sequence identity of 6–15%. Final products have regiochemistry and stereochemistry, however the structural basis for this precise control is not understood.

Although it is generally accepted that GPP and FPP are respectively the precursors for monoterpenes and sesquiterpenes and these syntheses are compartmentally separated, bifunctional enzymes capable of efficient formation of both terpenes were discovered. It has been shown that such bifunctional enzymes depend on substrate availability and could be directed to different subcellular compartments, which extends the range of available substrates for enzyme utilization and increase the metabolite diversity (Gutensohn et al., 2012).