Plant secondary metabolites are one of the most important sources of therapeutic drugs and in fact many drugs currently in use are derived from plants or lead compounds of plant origin. Terpenoids are the most numerous and chemically diverse class of plant metabolites with over 25,000 compounds identified so far. This huge chemical diversity is however hardly exploited for the development of new drugs. This is due to several reasons such as poor availability of the source plant material, too low concentrations in the plant material and difficulties to obtain pure compounds. The TERPMED project is aimed at providing solutions to overcome these difficulties for two classes of plant terpenoids showing promise as potential drugs for treating cancer and central nervous system disorders: sesquiterpene lactones and phenolic diterpenes. Within these two classes of compounds, the project focuses primarily on parthenolide and carnosic acid. Parthenolide is the principal bioactive sesquiterpene lactone and the presumed active ingredient of *Tanacetum parthenium* (feverfew) extracts, which have been approved in Europe as an herbal drug sold without prescription for the treatment of migraine. More recently, parthenolide has been the subject of several studies suggesting its use as a novel treatment for cancer. Regarding carnosic acid, this compound and its derivative carnosol are the most abundant phenolic diterpenes in the leaves of rosemary (*Rosmarinus officinalis*) and Salvia plant species, and are also relevant for human health. For example, rosemary extracts are already commercialized as food preservatives due to their anti-oxidant properties. Moreover, some recent studies indicate that carnosic acid and carnosol have potential therapeutic properties that could make these compounds useful for the treatment and/or prevention of diseases such as some neurodegenerative disorders and cancer.

In this context, the overall scientific objective of the TERPMED project is to gain understanding of the metabolism of the above said compounds and other closely related compounds that can be identified as potentially interesting during project development, and make use of the knowledge generated for designing metabolic engineering strategies aimed at establishing efficient plant production platforms for the most promising bioactive molecules.

To achieve this goal, a range of analytical methods based on HPLC-PDA, GC-MS and LC-MS have been developed and optimized to specifically detect and quantify these compounds in samples from the plant species targeted in the project: *Tanacetum parthenium*, *Inula britannica*, *Rosmarinus officinalis* and *Salvia fruticosa*. An HPLC-PDA-MS method coupled to an on-line antioxidant detection system has been set up to specifically detect PDs with antioxidant activity. Unbiased metabolomic approaches based on accurate mass-LC-MS methods have also been developed and optimized for large-scale metabolite profiling of tissue samples. By using such untargeted approaches followed by multivariate analysis techniques, a comprehensive and global view of the metabolome of plant extracts has been obtained to establish differences between

species, genotypes, tissues and organs from the collection of plants generated in the project.

The above analytical methods have allowed to establish that parthenolide is primarily stored in the ovary trichomes of *T. parthenium* disc florets, Inula sesquiterpene lactones accumulate preferentially in the trichomes of *I. britannica* flower corolla, carnosic acid mainly accumulates in leaf trichomes of *R. officinalis* and *S. fruticosa*, and carnosol is primarily found in other leaf cells of these two species. High quality RNA preparations obtained from ovaries and ovary trichomes of *T. parthenium* disc florets, ovaries of *I. britannica* flowers, and leaf trichomes of *R. officinalis* and *S. fruticosa* have been subjected to high-throughput RNA sequencing (RNA-seq) and the resulting sequence data have been assembled, annotated and made available via an interactive EST database that is accessible through the TERPMED website (http://www.terpmed.eu).

One of the key objectives of the TERPMED is the elucidation of the biosynthetic pathways leading to parthenolide and to carnosic acid. The four *T. parthenium* enzymes responsible for converting farnesyl diphosphate into parthenolide have been cloned and functionally validated in yeast and the whole parthenolide pathway has been successfully reconstituted in leaves of *N. benthamiana* plants. The first two enzymes of the carnosic acid pathway from geranylgeranyl diphosphate have also been cloned and functionally validated *in vitro* by enzyme activity assays and *in vivo* by expression in both yeast and *N. benthamiana*. These terpene synthases catalyze the two-step synthesis of miltiradiene, which is the diterpene backbone precursor of carnosic acid and carnosol. The miltiradiene pathway has been stably reconstructed in leaf trichomes of *N. tabacum*. Identification of the cytochrome P450s mediating successive oxidation reactions leading to carnosic acid and carnosol is in progress and some strong candidates have already been identified.

Purification procedures have been developed to obtain highly purified preparations of parthenolide and carnosic acid at the gram scale, whereas a range of minor SLs and PDs have been purified in yields ranging from μg to mg. All these compounds have been tested for their ability to activate the Nrf2-ARE antioxidant pathway. Despite all SLs with an exocyclic methylene group activate this pathway, they are not promising lead compounds for further development of drugs targeting the Nrf2-ARE pathway due to their neurotoxicity. On the contrary, carnosol is a promising lead for development of drugs to treat neurodegenerative diseases since it activates the Nrf2-ARE pathway with a low level of toxicity. Production of carnosol through chemical oxidation of carnosic acid has already proven feasible.

A generic metabolite database has been developed to store and make available information about compounds identified in the project. This database is linked to the sequence database through an integrated database for which a pathway representation is

used. Information for the biosynthetic pathways can be visualized in pathway maps representing the knowledge on reaction networks obtained in the project. For compounds in the pathways that are present in the compound database, a link from the pathway representation to the compound database is provided. For selected enzymes of the pathways links to the protein sequence are provided.

In summary, the TERPMED project has laid the foundation for improved and sustainable production of known biologically active molecules and has also enabled to identify new bioactive molecules that can be considered as promising lead compounds for further development of drugs targeting cancer, neurological disorders and malaria. Moreover, the project has generated a pool of highly trained researchers in this interdisciplinary area. Therefore, the primary issues addressed within the framework of the TERPMED project have an obvious European dimension as they will support the European Union ambition of becoming one of the world's leading knowledge-based economies.