Parsley, sage, rosemary and thyme — four spices from an old song. Although the anonymous lyricist from hundreds of years ago wasn’t an organic chemist, it happens that each of these spices is related to a large class of hydrocarbon compounds called terpenes. The essential oils of violets and roses, peppermint, pine trees (hence “turpentine”), eucalyptus, oranges, and also frankincense and myrrh — all of them contain terpenes that are among thousands found in nature, many entering our lives as pungent fragrances and flavors.

“Most terpenes are plant-derived and have interesting, pleasant flavors or smells,” says University of California, Davis chemist Dean Tantillo, “but they have activities all over the map. Some have the ability to fight cancer.”

The terpene taxol, a widely used anti-cancer drug, effective because it inhibits cell division, exemplifies Tantillo’s interest in these compounds. Taxol comes from the bark of the Pacific Yew tree, and when chemists first isolated it in 1967, they faced huge obstacles in making it available as a pharmaceutical. Early efforts required 1,200 kilograms of bark to produce 10 grams of taxol. Estimates by the mid-1980s were that it would take 360,000 yew trees a year to meet the anticipated need for ovarian cancer alone, an amount impossible to sustain. Creative chemists eventually sidestepped this problem by developing ways to make taxol from available small molecules or a precursor chemical in yew-tree needles. The latter approach exploits nature’s ability to produce taxol’s complex core architecture, but the former involves a long sequence of chemical steps, and for most complex terpenes, chemists have been forced to follow this difficult approach.

In nature, explains Tantillo, the core of a complicated terpene is assembled by an enzyme catalyst in one step. “If you could harness the enzyme to do it for you, synthesizing a complex terpene would be much more efficient.” Tantillo’s research focuses on understanding these reactions, with the aim of unraveling a central mystery. Nature uses hundreds of different enzymes to produce hundreds of different terpenes, but all from only a handful of starting compounds. How does a given enzyme make an abundance of only one product when hundreds are possible?

“It’s rare that biology uses the same compound to make hundreds of different complex things so efficiently,” says Tantillo. “If we can understand why one enzyme is producing one terpene and a different enzyme is producing a different terpene, we should be able to apply that knowledge to rationally redesign the enzymes. The goal would be to take a readily available enzyme that makes one complicated organic molecule that perhaps has an interesting structure but no useful biological activity, understand how that enzyme works, and change it in ways that will cause it to produce a different molecule that is highly valued.”

With demanding quantum computations on Pople, PSC’s SGI Altix system, Tantillo has explored the detailed, atomic-level chemistry of terpenes. His recent papers from this work — *Nature Chemistry* (August 2009) and *JACS* (March 2010) — report unanticipated findings that alter the standard view of terpene-forming reactions and point toward new understanding of the mysterious chemistry at play in producing terpenes in nature.

Terpenes are derived biosynthetically from units of isoprene, a biological building block that consists of five carbons — four in a row and one on the side — and eight hydrogen atoms, $\text{C}_5\text{H}_8$. The basic formula of terpenes is to multiply that, $(\text{C}_5\text{H}_8)^n$, where $n$ is the number of linked isoprene units.

Much of Tantillo’s recent work has focused on “sesquiterpenes,” a large family of terpenes formed from three isoprenes, 15 carbon and 24 hydrogen atoms, rearranged through reactions into complex configurations of carbon rings, fused together or linked by straight hydrocarbon chains. (Tantillo has also examined **diterpenes**, 20 carbons and 32 hydrogens, that’s a precursor to taxol.) The starting compound for this work is farnesyl diphosphate (FPP), a straight chain of three isoprenes ending with a diphosphate group, from which nature constructs hundreds of different sesquiterpenes.

For his computational modeling of sesquiterpene reaction pathways, Tantillo relies on GAUSSIAN03.
The IRC plot maps out the topography of the entire pathway from reactant to product. The numbers on each structure are computed distances in Angstroms.

IRC PLOT

The results of an IRC calculation from a terpene reaction in which two events (lower left and right) combine into one process. The numbers on each structure are computed distances in Angstroms.