**NMR properties of complex natural substances: density functional theory calculations as a valid complementary tool for structure elucidation**

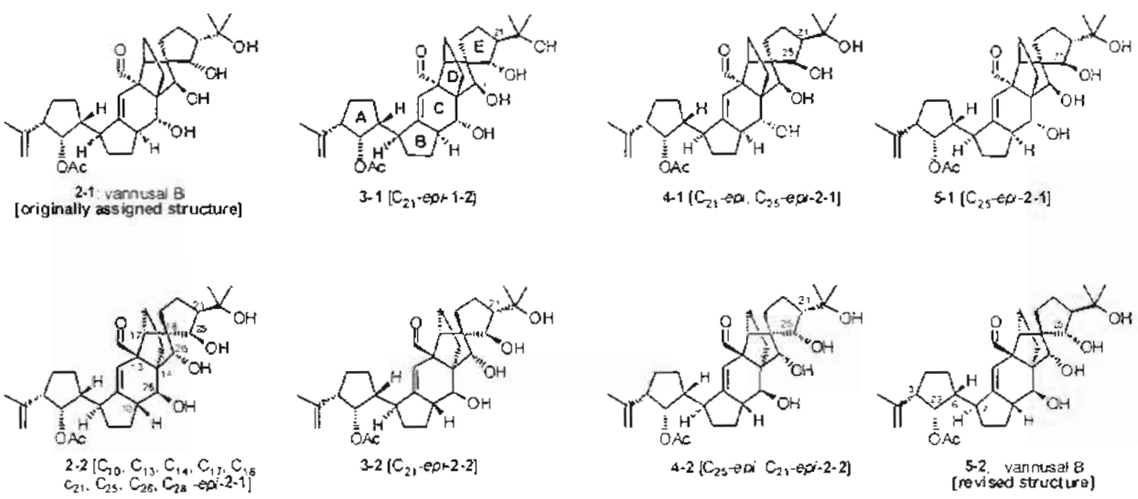
STM report by Dr. Giacomo Saielli
ITM-CNR, Padova Unit
Via Marzolo, 1 – 35131 Padova, Italy

## Introduction

During the three weeks spent at The Scripps Research Institute, La Jolla (San Diego) CA, USA, in the group of Prof. K.C. Nicolaou, G. Saielli conducted a careful analysis of the experimental NMR data – and subsequent comparison with the corresponding calculated NMR parameters from DFT calculations – collected by K.C. Nicolaou and his group for two class of compounds: vannusals and biyouyanagins. The comparison of calculated and experimental data highlighted some important issues concerning the original assignment of some proton and carbon resonances and showed the potential of the computational tools when applied to structural elucidation of complex natural substances. Furthermore, one of "unsolved misteries" reported in Nicolaou's 2005 review (Nicolaou&Snyder *Angew. Chem. Int. Ed.* **2005**, *44*, 1012), corylifolin, was reanalyzed in detail. A small number of putative structures has been proposed based on the comparison of calculated NMR shifts and couplings with the experimental $^{1}H$ and $^{13}C$ spectra. The details will be described in the next Section.

## Results

Vannusals. Vannusal A and vannusal B were isolated from isolated from *Euplotes vannus* in 1999 by Guella et al. (*Angew. Chem. Int. Ed.* **1999**, *38*, 1134).




Scheme 1: structures of various vannusals.

The quest for the total synthesis of Vannusal B (vannusal A being just the acetylated compound on C26), took about ten years to reach completion (Nicolaou et al. *J. Am. Chem. Soc.* **2010**, *132*, 7153), starting from the originally proposed structure, **2-1**, to the final revised structure, **5-2**, passing through six other vannusals synthesized in a step-by-step process aimed at gathering new intelligence at any new stage.

The eight vannusals can be divided in two groups: **n-1** compounds and **n-2** compounds, where the **n-2** isomers are epimers at C10, C13, C14, C17, C18, C21, C25, C26, C28 of the **n-1** isomers, that is most of the molecular structure except the "southwest" region (which, in fact, turned out to be correctly assigned by Guella). For each of the two groups we can identify four isomers, depending on the configuration of C21 and C25: R, R; R, S; S, R; S, S.

The path from the originally proposed structure to the revised structure, then, needed to: *i*) understanding that vannusal B belongs to the **n-2** group of compounds (as far as C10, C13, C14, C17, C18, C26, C28 configuration is concerned) and *ii*) understanding which one of the four isomers is the correct one concerning the configuration at C21 and C25.

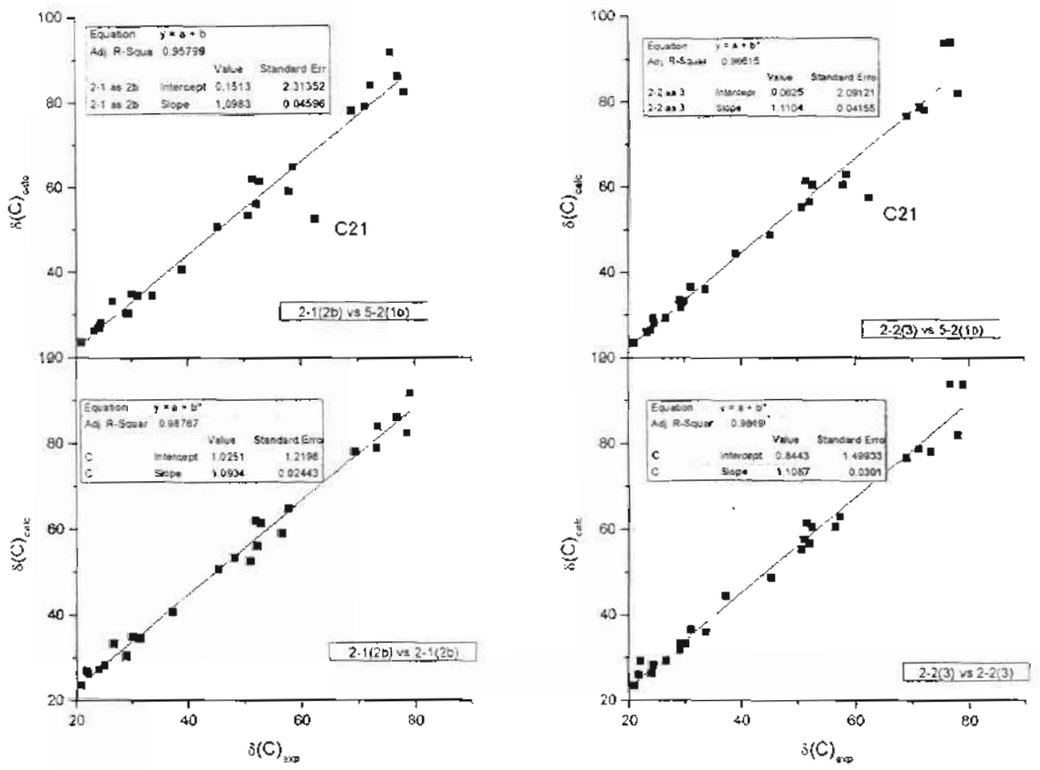



Figure 1: Correlation between calculated and experimental δ(C13) for **2-1** and **2-2**. (top) Experimental values of vannusal B; bottom experimental values of **2-1** and **2-2**, respectively.

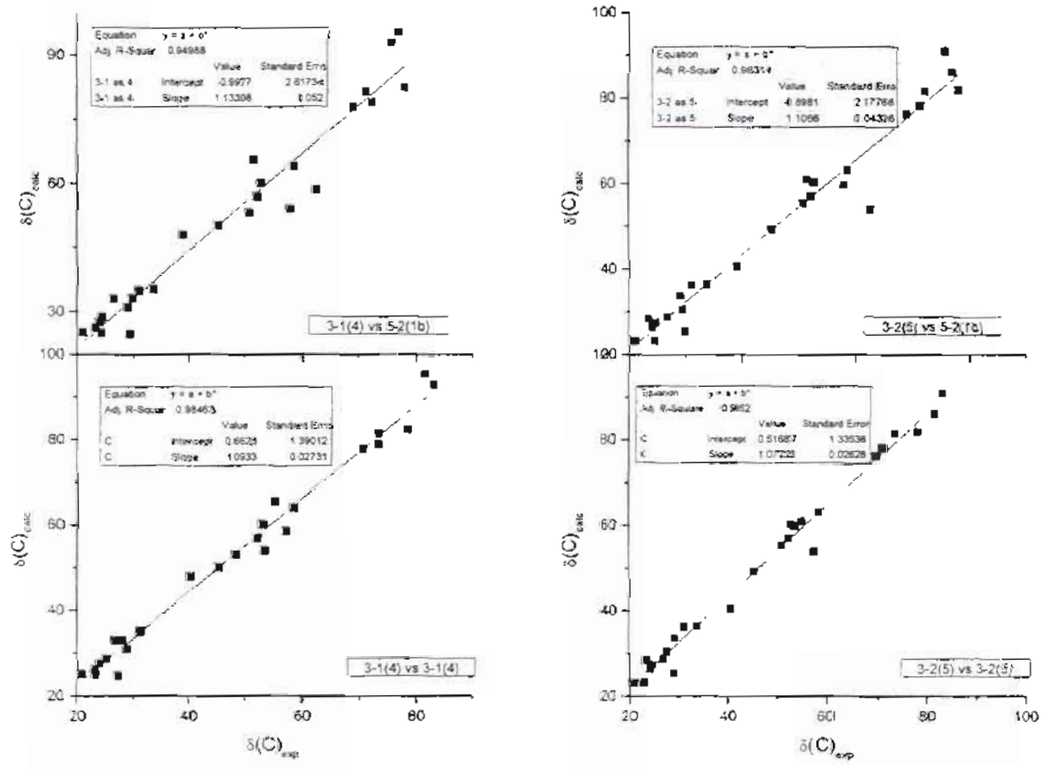



Figure 2: Correlation between calculated and experimental δ(C13) for **3-1** and **3-2**. (top) Experimental values of vannusal B; bottom experimental values of **3-1** and **3-2**, respectively.

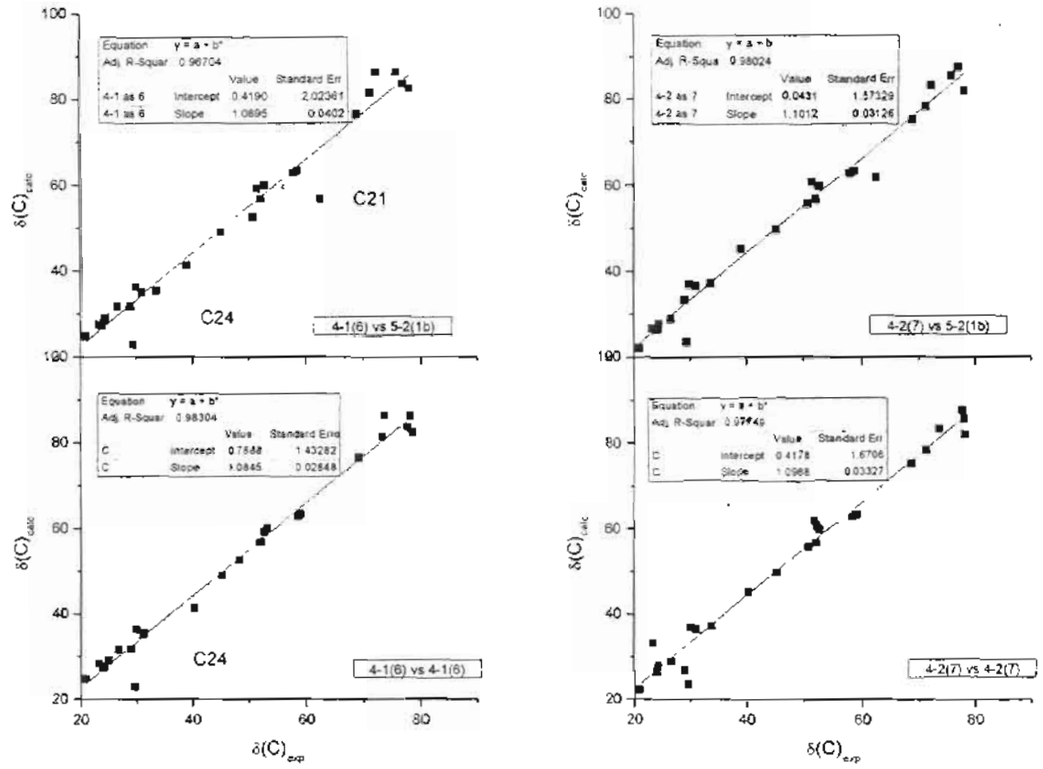



**Figure 3**: Correlation between calculated and experimental δ(C13) for **4-1** and **4-2**. (top) Experimental values of vannusal B; bottom experimental values of **4-1** and **4-2**, respectively.

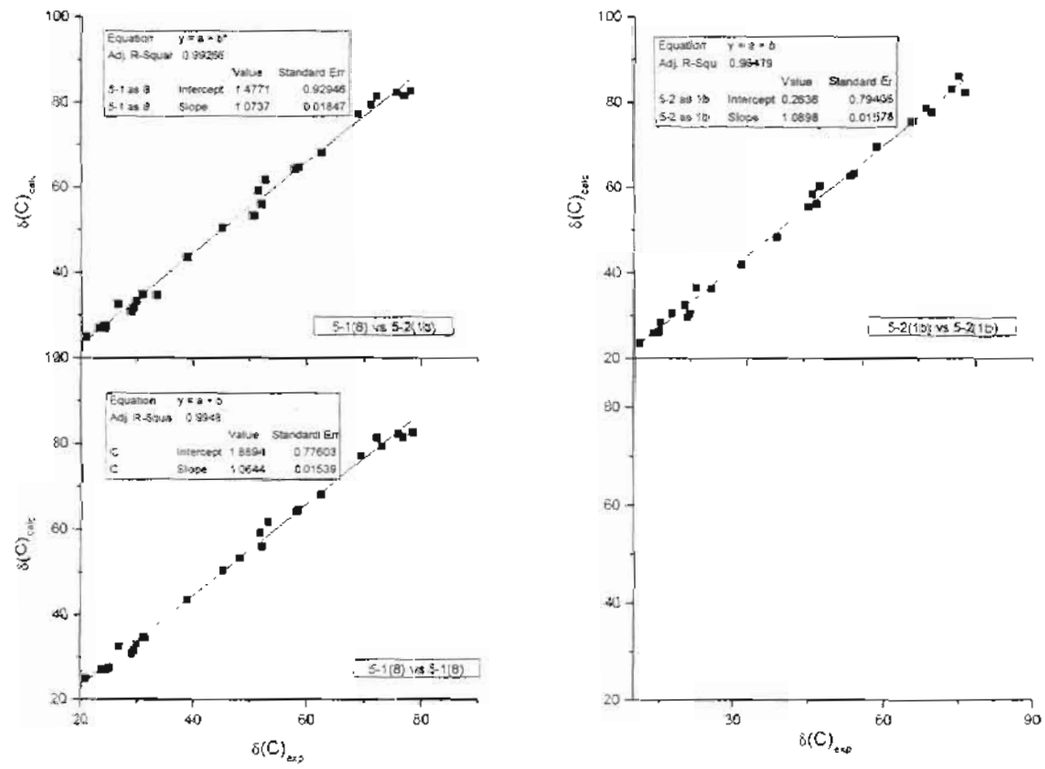



**Figure 4**: Correlation between calculated and experimental δ(C13) for **5-1** and vannusal B **5-2**. (top) Experimental values of vannusal B; bottom experimental values of **4-1** and vannusal B, respectively.

In Figures 1 to 4, we show the correlation of calculated carbon resonances for the eight vannusals with the experimental values of vannusal B, in the top panel. The corresponding bottom panel will show the

correlation of the calculated carbon resonances of the same compound but with the experimental values of that particular vannusal. This comparison highlights at once how far the structure is from vannusal B and what is the level of correlation when the correct experimental values of that structure are used.
The DFT protocol selected, based on the previous experience of our group, is based on the M06 functional, recently proposed by Truhlar and co-workers for the calculation of NMR properties, together with the pcS-2 basis set, developed by Jensen and co-workers specifically for the calculation of chemical shifts.
Calculations clearly show that the originally proposed structure, **2-1**, is quite unlikely the correct one, see Figure 1: C21 is largely in error and the general agreement is low ($R^2$=0.958). For comparison the correlation with the correct experimental values of **2-1** is much better. Compounds **3-1** and **3-2** are still far from the correct structure, see Figure 2. We note, however, that diastereomeric compounds of type **n-2** show a better correlation with the true vannusal B than compounds ot type **n-1**. Compounds **4-1**, and even more **4-2**, show a better correlation with the experimental values of vannusal B, suggesting that the molecular structure was close to the true natural substance, see Figure 3. We note that C24 and C23 could not be assigned in the experimental spectrum; unfortunately they belong to a flexible part of the molecule for which a careful population analysis of the various conformers is needed, but is beyond the scope of this investigation. Instead, C20 and C5 in **4-2** have been exchanged in the experimental assignment: comparison with the calculated values strongly support a reassignment for this pair. Finally, see Figure 4, a very good agreement is found with the correct structure **5-2**, corresponding to the true natural vannusal B. Also the almost enantiomeric structure **5-1** correlates well with the experimental values.
To summarise, Figure 5 shows the correlation coefficient $R^2$ for the various vannusals synthesized: it jumps above 0.99 only for the correct vannusal structure **5-2** ( and the almost enantiomeric structure **5-1**).

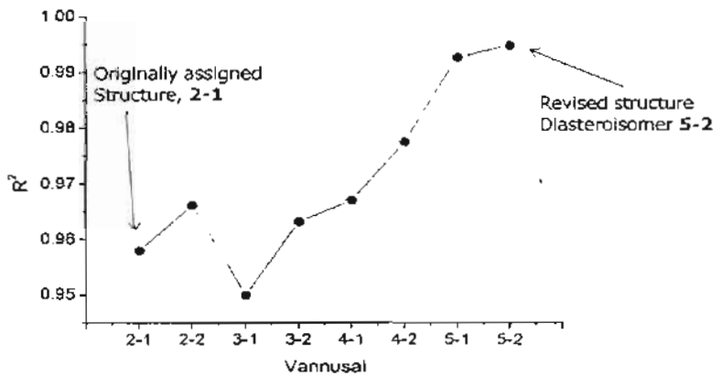
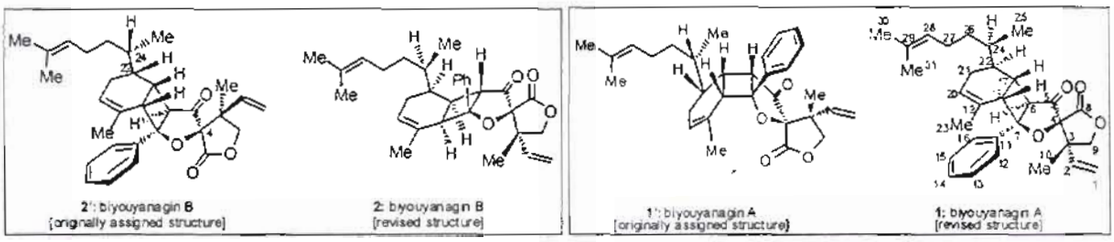



**Figure 5**: correlation coefficient, $R^2$, for the correlation of calculated δ(C13) with the natural vannusal B experimental values.

From the analysis of the above results we can conclude that DFT-based computational protocols are up to the task of distinguishing between stereo-isomers of natural substances with a complex architecture as the vannusals. Presented with the set of hypothetical structures of vannusal B the calculations clearly points to the correct structure and allow to discard that one originally proposed. In this sense they represent a very powerful complementary tool in structural elucidation.
Calculated proton chemical shifts appear more scattered, with respect to the experimental values, to allow for a meaningful comparison. This is largely due to the fact that experimental values refer to spectra obtained in methanol: vannusals have several hydroxyl groups so that solute-solvent interactions are significant and affect especially proton resonances.
However, coupling constants are much less affected by solvent effects: the key stereoisomerism at carbons C21 and C25 is nicely highlighted by calculated 3J(H21,H25) coupling constants for the four possible arrangements shown in Scheme 1, that is structures **2-n**, **3-n**, **4-n** and **5-n** (n = 1 or 2). The experimental value for vannusal B is J = 2.0 Hz. The calculated values for four small model systems of the "northeastern" region are 10.2, 6.4, 5.5 and 1.8 Hz, respectively strongly suggesting that **5-1** or **5-2** is the correct structure of the natural substance vannusal B.

Biyouyanagins. Biyouyanagins A and B have been isolated from *Hypericum Chinense* in 2005 (Tanaka et al. *Org. Lett.* **2005**, *7*, 2997) and in 2009 (Tanaka et al. *J. Nat. Prod.* **2009**, *72*, 1447), respectively. They have shown a significant activity as anti-HIV agents and lipopolysaccharide-induced cytokine production inhibitors. The originally proposed structures have been revised by K.C. Nicolaou et al. (K.C. Nicolaou et al. *Angew. Chem. Int. Ed.* **2007**, *46*, 4708; *Chem. Eur. J.* **2010** DOI: 10.1002/chem.201001474). The main difference between the originally proposed structures and the revised structures is in the stereoisomerism at the cyclobutane ring.




**Scheme 2**: originally proposed and revised structures of biyouyanagin A and B.

Comparison of calculated and experimental carbon chemical shifts for both biyouyanagins A and B revealed a very high quality correlations for the originally proposed and revised structures, indicating that both molecules should have a very similar $^{13}C$ spectrum, see Figure 6. Indeed biyouyanagin A and B have an almost indistinguishable $^{13}C$ spectrum.

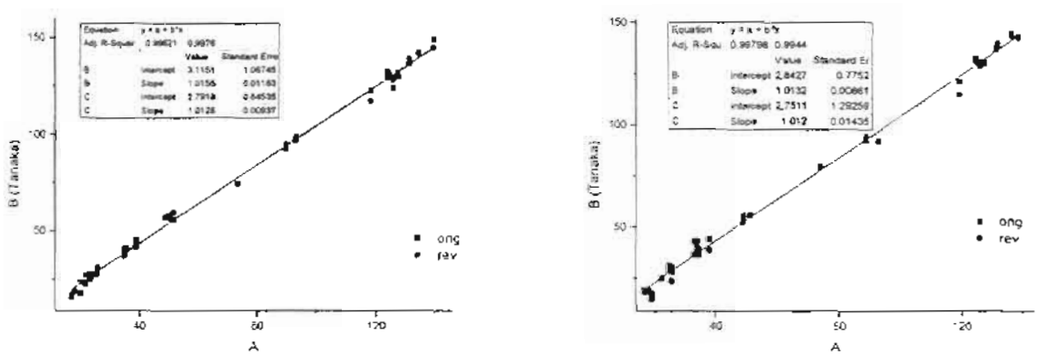
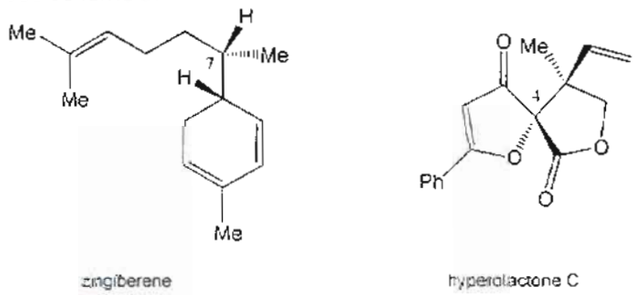



Figure 6: Correlation between calculated and experimental δ(C13) for (left) biyouyanagin A and (right) biyouyanagin B. Black squares: originally proposed structures. Red circles: revised structures

It appears, therefore, that both molecules (the originally proposed and the revised structure) are too similar to be clearly distinguished by comparison of calculated NMR parameters.
Nevertheless, the interaction with K.C. Nicolaou and his group revealed very fruitful: biyouyanagins are thought to be biosynthesized in a [2+2] photocycloaddition of two units, zingiberene and hyperolactone C, see Scheme 3.

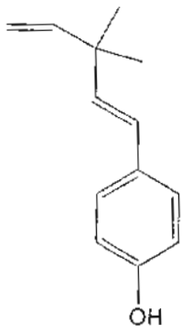



**Scheme 3**. Structure of the naturally occuring zingiberene and hyperolactone C.

Given the four diastereoisomers of zingiberenes, each with two C=C double bonds, and for each double bond two sides for approaching the enone, and the two epimeric hyperolactones at C4, there are a total of

32 possible biyouyanagins. Some might be strong anti-HIV agents, even stronger than the natural substances biyouyanagins A and B discovered since now. It is then, of interest to investigate by DFT-based computational protocols, the reaction mechanism, at least for the most stable biyouyanagins, of the [2+2] photocycloaddition exploring both the singlet and triplet PES. This cooperative work with K.C. Nicolaou's group which will be conducted in the next months.

Corylifolin. Corylifolin was isolated in 1998 (Snapka et al. *J. Nat. Prod.* **1998**, *61*, 362) from *Psoralea Corylifolia*, a plant which found large use in traditional Chinese medicine and indian ayurvedic medicine. Psoralen and Isopsoralen, also isolated from this plant, are active ingredients of several drugs of current use. Corylifolin was originally assigned the structure of Scheme 4. However, total synthesis of this compound revealed that the structure of corylifolin needed revision.




**Scheme 4**. (Left) Originally proposed structure of the natural substance corylifolin.

In particular, chemical shifts of the two methyl groups were not equivalent (1.09 and 1.17 ppm for $^{1}H$ and 23.9 and 25.0 ppm for $^{13}C$). Thus, some alternative structures have been considered and DFT calculations are currently in progress in order to assess their reliability as the correct corylifolin.

Padova, 03 Settembre 2010

Il Fruitore: Dott. Giacomo Saielli

Il Proponente: Prof. Alessandro Bagno