URL: http://www.iupac.org/symposia/proceedings/phuket97/borris.html

Molecular Diversity, Biological Diversity and the Search for New Drugs

Robert P. Borris and Steven J. Gould

Department of Natural Products Drug Discovery, Merck Research Laboratories, P. Roy Vagelos Research and Development Center, P. O. Box 2000, Rahway, New Jersey, 07065 U.S.A.

Abstract: The success of the drug discovery process is often a function of the diversity of chemotypes examined. Natural products screening represents a potential source of organic chemicals of unparalleled diversity. To effectively realize this potential requires use of a selection strategy tailored to the needs of the individual screening program. Historically, plants and microorganisms have been extraordinarily rich sources of medicinally and agriculturally useful compounds. Interest in these sources of new bioactive molecules continues to the present time. In recent years, the evaluation of insects and marine invertebrates as sources of biologically active compounds has added to the array of new chemotypes. While the efficiencies of modern screening methods allow for the cost effective evaluation of vast numbers of samples, the cost of acquiring and processing natural products, particularly macroorganisms, is substantial relative to other sources of molecular diversity. A successful drug discovery program must then aim to evaluate the broadest diversity of relevant chemical classes in the minimum number of samples and least time. To achieve this goal, a careful assessment must be made of each source of molecular diversity, natural and man-made, and a proper balance of sources established. An evaluation of the ecological/environmental ramifications of collecting organisms from natural populations for screening operations and pre-development studies is an integral part of this assessment. While the screening of natural products, requiring a very small sample, is a practical endeavour, the realities of obtaining sufficient quantities of compound from natural populations of macroorganisms to support product development and eventual sale are often problematic. Alternative methods of producing interesting compounds, including synthesis, cell culture and agriculture, may be required in order to minimize the ecological impact of these discoveries. While our environmental concerns often focus on the direct impact of this research, they extend far beyond the scope of these studies, fostering the preservation and sustainable utilization of the natural habitats of our collaboratoring

*Invited lecture presented at the International Conference on Biodiversity and Bioresources: Conservation and Utilization, 23–27 November 1997, Phuket, Thailand. Other presentations are published in *Pure Appl. Chem.*, Vol. 70, No. 11, 1998.

source countries. This presentation will contrast the attributes of the molecular diversity obtainable from a variety of natural and man-made sources, with illustrative examples from recent experience. (Keywords: Biodiversity; natural products; drug discovery.)

Natural products research has been an integral part of drug discovery at Merck since the early days of antibiotic research in the 1940's. The primary focus of this effort has been on the discovery of new microbial metabolites with biological activity. Plants received some intermittent attention in the forties through seventies, but this source of novel molecules was not addressed in a systematic manner until the 1980's, when Merck, with collaborators in the People's Republic of China, embarked on a ten year evaluation of traditional Chinese medicines. In the middle of that decade, the scope of these studies was broadened to include plants from other regions of the globe through an ongoing collaboration with the New York Botanical Garden. Then, in 1991, Merck entered into the much publicized agreement with INBio to study the flora (and parts of the fauna) of Costa Rica in greater depth. In the 1990's, the scope of our studies was further broadened to include insects and marine invertebrates It is, then, clear that Merck scientists have been studying 'traditional medicines' and 'biodiversity' long before these buzzwords became fashionable.

The mission of a major pharmaceutical corporation is to discover and develop new drugs for the treatment of disease, *and* to manufacture and market these drugs in such a way as to maximize the financial return on the investments of our stockholders. Corporations such as Merck invest heavily in basic research, both in their own laboratories and in those of academic institutions, dealing with the fundamental understanding of wellness and disease, in order to keep our drug discovery efforts on the cutting edge of science. This investment enables our scientists to select and develop the best and most appropriate targets for the diseases we are trying to treat. The main thrust of the research at a large pharmaceutical corporation is directed toward the discovery and development of new drugs. Natural products research is actually a small part of the overall discovery process.

The screening of natural products is one of the earliest steps in drug discovery -- Lead Identification. A lead compound, also frequently refered to as a chemical template, is a compound with many of the characteristics of a desired new drug which will be used as a model for chemical modification, but which lacks either the potency or specificity expected of a product candidate. Lead compounds must be available in sufficient quantities to support the early stages of development such as biological characterization and toxicity studies while a total synthesis of the product candidate is completed. It is much easier to find a lead compound than to find a product candidate. Medicinal chemists are now *routinely* expected to be able to improve the potency and/or specificity of a lead compound, perhaps by several orders of magnitude, and thus take a lead compound and turn it into a drug.

The combination of a number of factors has led to industry's current interest in natural products. In the 1950's and 1960's, when the role played by natural products in the drug discovery process was at its peak, the majority of bioassays targeting diseases other than bacterial and fungal infections involved whole animal models of disease. They were slow, cumbersome, expensive, often difficult to properly evaluate and not very sensitive. In short, the assays of the day were not well suited to the study of complex mixtures such as natural product extracts. Chemical methodology was also rather primitive. Chromatography meant gravity fed open

columns or thin layer plates. Countercurrent separations meant using the Craig apparatus. Spectroscopy meant IR and UV, with mass spectrometry and NMR beginning to come into routine use. Of course structure determination required degradation and synthesis. Grams of pure compound were required for complete identification and initial biological evaluation. Technology has changed. The current generation of mode of action assays is fast, selective, quantitative and readily amenable to automation. Fast and easy assays greatly facilitate the assay guided isolation of active constituents, which now routinely employs HPLC and HPCCC as frontline separation techniques. Once the active constituent is isolated its identification has been greatly simplified by modern spectroscopic methods. While IR and UV spectra are still useful, most structures are solved on the basis of high resolution mass spectrometry and high field, multinuclear, multidimensional NMR studies, routine in many laboratories. Chiroptical techniques such as circular dichroism have also simplified the determination of stereochemistry. Structure determination and initial biochemical investigations routinely require milligram quantities of compound rather than grams. The trend is obvious. The time cycle for natural product discovery has been substantially compressed while the sample requirements have been drastically reduced. This has enabled pharmaceutical researchers to study the chemical diversity of natural sources to a much greater extent without significantly contributing to deforestation or the extintion of marine invertebrates.

NATURAL PRODUCTS RESEARCH

A natural products research program may focus on one or more of the kinds of biomatter available to it. Microbiologists isolate and ferment pure cultures of microorganisms from environmental samples which had been collected in the biosphere. Botanists collect plant samples. Marine biologists collect invertebrates from the marine environment. Entomologists collect samples of insect populations. Whatever the source, it is necessary to separate the potential active compounds from the surrounding biological matrix via extraction and then to remove any known classes of general interfering substances via a judiciously selected processing scheme. Processing may include such steps as solvent partitioning, solid-phase extraction, ion exchange or even some chromatography. The resulting extracts or fractions, freed of interferences such as fatty acids, tannins, polysaccharides, inorganic salts and high molecular weight substances, can then be screened.

In a typical industrial screening program, a very large population of extracts, perhaps in the hundreds of thousands of samples per year is evaluated in a substantial number of assays covering a range of therapeutic areas. A hypothetical program of moderate size with fifty assays and a screening throughput of 100,000 samples/year would generate five million individual test results. If the assays in the panel are well conceived and constructed, the vast majority of samples will probably show no activity in any of the assays. A small percentage of samples will posess activity in many, if not all, of the assays. Such behavior suggests the presence of some general interference rather than a true activity. The remaining samples, having activity in one or more assays, are the samples of interest. While it is always tempting to quote ‘hit rates’ for an assay or assay panel, these figures are generally uninformative as they are a direct reflection of the criteria used to define “activity”. As a matter of practicality, if an assay in the hypothetical program described above had a 1% hit rate, it would generate 1000 “actives” each year, far more than

could be followed by even a large chemistry group. If the originally observed activity is reproduced in a second sample or regrowth, the extract having confirmed activity is subjected to some form of dereplication, the process of determining whether an active constituent has been encountered previously..

Dereplication is a formidable task, but it is essential to the success of any natural products program as it prevents the waste of resources required for the re-isolation of known compounds, allowing the group to focus on unique activities. Biological assays change frequently in large, broad-based screening programs, so one is confronted with the problem of rapidly recognizing not only the compounds known to be active in the current assays, but also those that have been isolated for earlier assays. Longterm screening programs, like the one at Merck, can generate thousands of compounds found to be active in one or more of the hundreds of assays used at various times in the screening panel. The use of HPLC-MS coupled with ultraviolet detection using a photodiode array detector greatly simplifies the task, especially when linked with computer databases and pattern matching algorithms. This is not to imply that that dereplication is or should be a strictly chemical undertaking. Indeed, information about the nature of the source organism and the spectrum of observed biological activities greatly facilitate this process.

If an extract survives dereplication, it becomes the focus of an intense interdisciplinary collaboration aimed at isolating and identifying the active constituent in the minimum amount of time. The microbiologist manipulates media and growth conditions to optimize production of the desired compound. The chemists isolate a small amount of the compound guided at each step of the separation with bioassay data. The assay biologist or biochemists evaluate samples from the chemists and microbiologists and further characterize the observed activity. Using state of the art spectroscopic and chemical techniques, the natural products chemist determines the molecular structure of the active compound. By the time the active molecule has been identified, the microbiologists often have increased its titer in fermentations by a factor of ten or more, and the chemists have refined the steps in the isolation to allow direct implementation in a large-scale isolation in the pilot plant. It is not unusual for related minor metabolites to show up and be characterized in this process; when it occurs bioassays of these compounds can immediately add information on structure-activity relationships.

The situation is somewhat different when a macroorganism is involved. While the intersection between the natural product chemist and assay biochemist remains the same, the microbiologist has been replaced by a field botanist or marine biologist. These scientists become involved only if the initial supply of extract is not adequate to support the initial isolation and identification of the active molecule. Resupplying the extract of a macroorganism is a much more complicated task than resupplying a fermentation broth. For a microorganism it is a relatively straightforward task to produce hundreds or thousands of liters of a fermentation broth to support an isolation in a short time period. For a macroorganism, in order to have a reasonable level of confidence that the activity will be reproduced, one must collect the organism in the same geographic location and at the same stage in their life cycle as the original. While this may occasionally be possible in the space of a few weeks, more often resupply takes months or even years to accomplish successfully.

Once an active compound has been isolated and identified, it must be fully characterized in a biological and biochemical sense. While it is important to know that the compound possesses the desired activity, it is equally important to determine with what *other* molecular systems the molecule interacts. Potential lead compounds are evaluated against all molecular targets available

in the laboratory to determine the level of specificity for the desired target and what potential side effects or toxicity might be predicted. Occasionally, a new compound will interact with entire families of related targets, e.g. G-protein linked receptors or serine proteases. These promiscuous agents are usually of little interest.

Other discovery activities are often intimately involved in the biological characterization of a new compound. These may include the isolation of related minor constituents from the extract as noted above, microbial transformation of the active compound, and partial synthesis or preparation of chemical derivatives. For microbial products, directed biosynthesis experiments may be performed in which the microbiologist feeds the microbe an 'unnatural' precursor in an attempt to cause the biosynthesis of related molecules. These activities are aimed at the discovery of related compounds to help generate an initial structure activity relationship for the molecular class. The synthetic chemists will refine the SAR to select the best compound in the class for total synthesis. With this biochemical knowledge in hand, the compound progresses into pharmacologic evaluation to establish whether the compound has the predicted activity in an intact animal. The transition from *in vitro* to *in vivo* studies often requires a further refinement in the structure activity relationship.

When SAR studies have led to the identification of the best compound in the series, that compound is subjected to a rigorous safety assessment program which involves several species of animal and high doses of compound. While it may be possible to produce microbial products in sufficient quantity to support safety studies, it is assumed that these large quantities of macroorganism metabolites, often in excess of one kilogram of pure compound, will be produced by chemical synthesis. Any significant toxicity in any species of test animal can result in the termination of active interest in a compound, even if only observed at very high dosage levels. If some toxicity is observed, one or more related compounds, the backup candidates, may be similarly evaluated to determine whether the observed toxicity is due to an idiosyncrasy of the initial safety assessment candidate, or whether it is a more general phenomenon occurring throughout the class of compounds. Safety assessment data may necessitate some additional refinement in the structure activity relationships as the synthetic chemists attempt to reduce or eliminate unwanted biological activities. The most active compound from the original discovery process is usually not the one to be developed into a product. The success of a modern drug discovery program depends on the synthetic medicinal chemist's ability to subtly manipulate the chemical and biological properties of a lead molecule. If the observed toxicity of a safety assessment candidate cannot be reduced or eliminated by the medicinal chemists, it is likely that development of the entire class of compound will be abandoned.

Eventually, an active compound derived from the original lead structure will pass through safety assessment without observed toxicity. An Investigational New Drug Application (INDA) will then be filed and, when approved, the compound will make the transition from preclinical studies in animals to clinical trials in humans. Safety and efficacy must be adequately observed and documented in Phase I-III trials for a New Drug Application (NDA) to be filed with the regulatory authorities (e.g., the U.S. Food and Drug Administration). A necessarily detailed and critical evaluation of the NDA follows, and more clinical trials may be required by the regulatory authority to clarify relevant issues in the application before the drug will receive approval for marketing in a particular country. In recent years, the regulatory authorities have begun to require that additional clinical studies be performed after regulatory approval (Phase IV trials). While the entire process of clinical trials and regulatory approval can be summarized in a single

paragraph, this is a long and costly process, estimated (in 1993) to take 10 years to complete at an average pretax cost of $359 million (ref. 1). Only upon completion of this process, and after the massive investment in research and development, can a product be brought to market and begin generating revenue.

Each group of organisms studied in a natural products discovery program possesses a characteristic spectrum of advantages and disadvantages. The microbial kingdom is incredibly diverse in both a taxonomic sense and a biochemical sense, and, importantly, this diversity is accessible even without traveling to distant locations. Because a wide range of microorganisms can be readily accessed, a high screening throughput can be both practical and profitable. Industrial microbiology is a well developed science, so one can be confident that observed activities will usually be reproduced, and that supply is unlikely to be a persistent problem. It is for these reasons that most industrial natural products programs focus on microbial products.

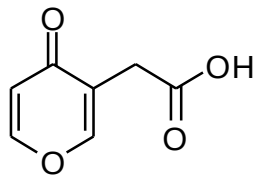
The discovery and development of new drugs from macroorganisms is not so straightforward. The logistics of acquiring a significant number of high quality samples for screening are daunting and expensive. It should be apparent that for a program based in the United States, the vast majority of plant and animal species are located overseas. Travel, the development of a local infrastructure in the source countries to process collections, and shipping samples and extracts back to the lab all add significantly to the relative cost of these samples. At Merck, we have tried to minimize the impact of these factors on our screening budget by establishing longterm collaborations, such as the ones with the New York Botanical Garden and the Instituto Nacional de Biodiversidad in Costa Rica. Obtaining additional quantities of active samples, as mentioned above, is not always easy. By carefully controlling as many of the collection variables as possible, we have been able to produce small quantities of active phytochemicals on a regular basis. However, in most cases, it is unlikely that we would be able to produce enough of an active compound from collections of wild populations to support a safety evaluation requiring a kilo or two of product. When compared with a microbial product, recollecting a plant or marine invertebrate in sufficient quantity to produce a few grams of compound is a very slow and expensive endeavor. For these and many other reasons our macroorganism screening programs have been based on the assumption that it may be impossible to obtain any resupply of the sample, and that any significant quantity of an active component would be produced synthetically.

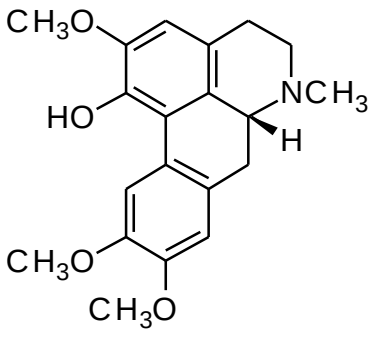
Despite the high costs and difficulties of obtaining samples of macroorganisms for screening and the dependence on synthesis for production of active compounds for development, macroorganisms remain an area of active interest in the pharmaceutical industry. The goal of this research is, of course, the discovery of novel active compounds for development. To do this effectively, it is necessary to maximize the diversity of chemotypes being evaluated in a screening program. We have learned from the collective experience of the scientific community that fungi, bacteria, actinomycetes, plants, insects, and marine invertebrates each produce a characteristic spectrum of secondary metabolites. While there is some overlap it is nevertheless necessary to evaluate each of these groups of organisms, and perhaps others, to see the full spectrum of nature's chemical abilities.

EXAMPLES

The following few examples have been drawn from recent experience in our laboratories. Although antimicrobial chemotherapeutics are a major class of targets in this program, they have been intentionally omitted from this discussion in order to highlight work in other therapeutic areas.

Interleukin-1β, a major determinant in the etiology of acute and chronic inflammatory disease, is synthesized as an inactive 31kDa precursor (Pre-IL-1β) which is cleaved to a 17.5 kDa mature, active protein. The enzyme catalyzing this cleavage, Interleukin-1β converting enzynme (ICE), is a unique cysteine-containing heterodimeric protease with a very high degree of specificity for Pre-IL-1β. It was felt that ICE may be a potentially important target for novel antiinflammatory agents, so an assay was developed to measure ICE activity *in vitro* and screening was commenced. After evaluating several thousand microbial fermentation extracts, a culture of a *Xylaria* species obtained from a bark sample collected in South Carolina was found to possess the desired activity. Isolation via ion exchange and reverse phase chromatography afforded xylaric acid (Fig. 1, **1**) as the active constituent. Xylaric acid was found to be a competitive, irreversible inhibitor of the enzyme with an IC_{50} of 33 μM. Screening of a library of a few thousand plant extracts resulted in the detection of an extract of the stems and leaves of *Croton urucurana* (H.B.K.)Baill. (Euphorbiaceae) collected in Paraguay. Isolation via solvent extraction and normal phase chromatography afforded the known aporphine alkaloid, thaliporphine (Fig. 1, **2**) as the active constituent. The biochemical activities of this compound were very similar to those of xylaric acid (ref. 2).



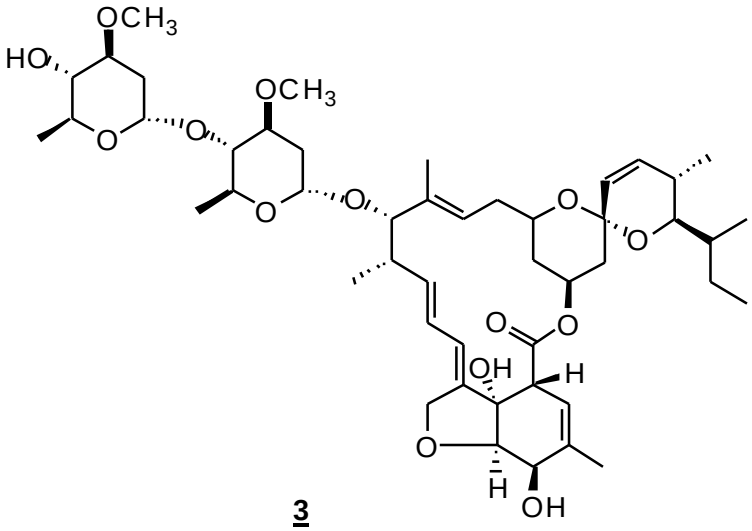


1 **2**

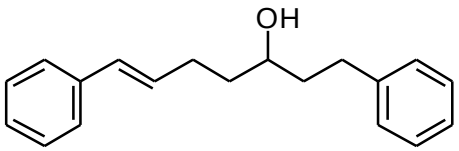
Figure 1. Interleukin 1β converting enzyme inhibitors: Xylaric acid (**1**) and thaliporphine (**2**)

The avermectins (ref. 3), such as avermectin B_{1a} (Fig. 2, **3**), a class of macrocyclic antiparasitic agents, have been the mainstay of Merck's animal health and agrochemicals businesses for a number of years. Products containing these compounds are widely used for the treatment of veterinary helminth infections, but also find application in human health for the treatment of onchocerciasis (river blindness). It is a common practice for companies to look for second and third generation drugs to augment, and at times supplant, their marketed products. The search for a second generation anthelminthic agent employed, as part of its assay panel, a motility assay

using the free living nematode *Caenorhabditis elegans* as a surrogate for the gastrointestinal parasitic nematodes. This assay was a very sensitive detector for the avermectins and the closely related milbemycins, and it was also employed in the search for novel anthelminthic chemotypes. Two extracts from a library containing several thousand plant extracts screened in this assay were found to have significant levels of activity. Both were nonpolar extracts of the roots of *Curcuma comosa* Roxb. (Zingiberaceae), a Thai medicinal plant. Isolation via solvent extraction and normal phase chromatography afforded a family of related diphenylheptanoids (e.g., Fig. 2, **4**) with EC_{95} values ranging from 700 ng/ml to over 100 µg/ml (ref. 4). Not only was this a novel class of nematocidal agents, but the initial isolation afforded a family of compounds which provided a substantial amount of information about their structure activity relationships.



3

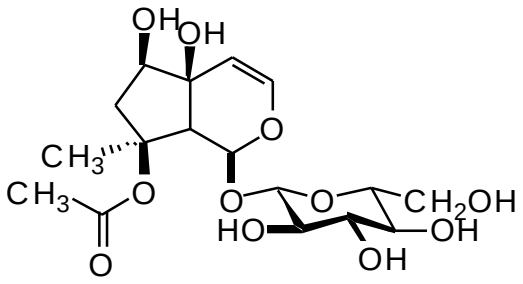


4

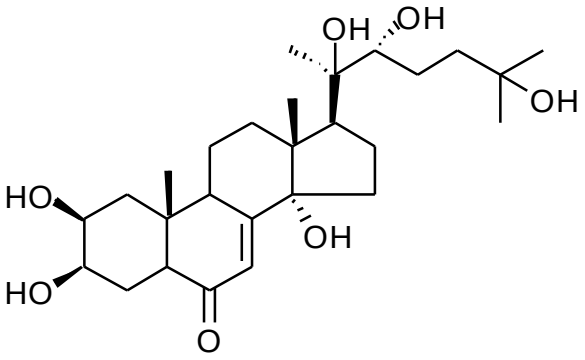
Figure 2. Anthelminthic agents: Avermectin B_{1a} (**3**) and a *Curcuma* diphenylheptanoid (**4**)

In the search for new drugs or agrochemicals, agents having a novel mechanism of action may represent safer alternatives to existing products. Many of the commonly used insecticides are neurotoxic compounds which are toxic not only to the target insects, but also to fish, birds and mammals. The maturation of insects involves molting, a complex biological event which is initiated by the steroid hormones, the ecdysteroids or ecdysones. Disruption of the molting process can be lethal to the insect. As ecdysteroid mediated molting occurs only in insects, modulation of the activity of these hormones at their receptors is an attractive target for the development of insect management agents that are nontoxic to other animal species. A transactivation assay was established in cultured Drosophila cMK7 cells to look for ecdysteroid receptor agonists. Screening a library of several thousand plant extracts afforded a single extract with significant levels of activity, a defatted methanol extract of *Ajuga reptans* L. (Lamiaceae) collected in New York. Isolation via solvent extraction and normal phase chromatography afforded an iridoid glycoside, 8-O-acetylharpagide (Fig. 3, **5**) as the active constituent (ref. 5). This compound activates the transcription of an ecdysteroid inducible promoter in the transactivation assay and competes with ponasterone A, a natural ecdysteroid, for binding at its receptor. Additionally, it causes differentiation of ecdysteroid responsive cells in culture in a manner analogous to that seen in the presence of 20-hydroxyecdysone (Fig. 3, **6**). Kubo (ref. 6)

has shown that 8-O-acetylharpagide has insect antifeedant properties. It is not yet clear whether these properties are the result of ecdysteroid mediated effects.



5



6

Figure 3. Ecdysone receptor agonists: 8-O-Acetylharpagide (**5**) and 20-Hydroxyecdysone (**6**)

CONCLUSION

The search for new biologically active natural products still holds great potential for the pharmaceutical and agrochemical industries. It is necessary to screen broadly both microorganisms and a wide range of macroorganisms to realize the vast potential of this resource, and, perhaps in the process, discover the next blockbuster drug.

REFERENCES

1. U. S. Congress, Office of Technology Assessment. *Pharmaceutical R & D Costs, Risks and Rewards*. Washington DC: U. S. Government Printing Office (OTA-H-522) (1993).
2. M. J. Salvatore, O. D. Hensens, D. L. Zink, J. Liesch, C. Dufresne, J. G. Ondeyka, T. M. Jurgens, R. P. Borris, S. Raghoobar, E. McCauley, L. Kong, S. E. Gartner, G. E. Koch, F. Pelaez, M. T. Diez, C. Cascales, I. Martin, J. D. Polishook, M. J. Balick, H. T. Beck, S. R. King, A. Hsu and R. B. Lingham. *J. Nat. Prod.* **57** (6), 755-760 (1994).
3. T. W. Miller, L. Chaiet, D. J. Cole, L. J. Cole, J. E. Flor, R. T. Goegelman, V. P. Gullo, H. Joshua, A. J. Kempf, W. R. Krellwitz, R. L. Monaghan, R. E. Ormond, K. E. Wilson, G. Albers-Schonberg and I. Putter. *Antimicr. Agents Chemother.* **15**, 368-371 (1979).
4. T. M. Jurgens, E. G. Frazier, J. M. Schaeffer, T. E. Jones, D. L. Zink, R. P. Borris, W. Nanakorn, H. T. Beck and M. J. Balick. . *J. Nat. Prod.* **57** (2), 230-235 (1994).
5. A. Elbrecht, Y. Chen, T. Jurgens, O. D. Hensens, D. L. Zink, H. T. Beck, M. J. Balick and R. P. Borris. . *Insect Biochem. Molec. Biol.* **26** (6), 519-523 (1996).
6. I. Kubo. In *Recent Advances in Phytochemistry: Phytochemical Potential of Tropical Plants, vol . 27* (K. R. Downum, ed.) pp 133-152. Plenum Press, New York (1993).