Atta-ur-Rahman (Ed.) Studies in Natural Products Chemistry, Vol. 33

STRUCTURE-ACTIVITY RELATIONSHIPS OF SESQUITERPENE LACTONES

THOMAS J. SCHMIDT[1]

Institut für Pharmazeutische Biologie und Phytochemie der Westfälischen Wilhelms-Universität Münster, Hittorfstraße 56, D-48149 Münster, Germany

ABSTRACT: Sesquiterpene lactones (STLs) are one of the largest biogenetically homogenous groups of natural products known. Currently, the Dictionary of Natural Products holds a total of over 11000 entries on sesquiterpenes, of which almost 5000 contain at least one lactone group. One reason for the large number of individual compounds and the associated structural diversity of this class of natural products is most certainly related to their enormously broad spectrum of biological activities, rendering them a highly valuable defense tool for those organisms synthesizing them. It is thus no surprise that the plant family from which most STLs have been isolated, i.e. the Compositae, is among the largest and ecologically most diverse plant families on our planet.

Many of the biological activities of the most common types of STLs are known to be mediated by a simple general chemical mechanism, i.e. alkylation of biological macromolecules in terms of Michael type additions. These effects have been surveyed in numerous comprehensive reviews. Moreover, some very specific receptor-mediated activities of less frequent types of STLs are known. Quite surprisingly, despite the plethora of reports on STL bioactivity, only very few systematic studies on structure-activity relationships have been carried out. Detailed studies of this kind, however, would be highly desirable with respect to several aspects of medicinal/pharmaceutical, agrochemical and ecological interest. Most importantly, many STL-containing plants have been used in traditional medicines of all cultures for many centuries and continue to be utilized also in modern phytotherapy. Therapeutic use of STLs as *pure chemicals* is - in spite of their broad utilization in form of plants or crude extracts- restricted to very few examples, which is mainly due to a lack of knowledge on the structural requirements for *selectivity* with respect to a desired biological activity. It may, however, be conceived, that STLs could play a valuable role as lead structures to new therapeutic agents, if more information, especially in the form of quantitative structure-activity relationships (QSAR), existed. The present chapter attempts to present an overview on the current knowledge on structure-activity relationships of STLs of different types and different activities. Selected examples of current research in the author's laboratory on structure-reactivity- and structure-activity relationships of STLs are presented.

---

[1]Dedicated to my teacher and mentor, Prof. em. Dr. Günter Willuhn, on the occasion of his 70th birthday.

INTRODUCTION

Approximately a century after the advent of modern pharmacology, the search for bioactive compounds in living organisms is still one of the most important approaches in finding new therapeutic drugs and agrochemicals. Many, if not most, of the new therapeutic entities approved over the last two decades represent compounds originating directly or indirectly from natural sources. Low-molecular weight natural products, frequently termed "secondary metabolites", (which does not imply that they are of secondary importance) are the major part of the chemical communication systems that serve organisms from the different domains of life to cope with each other and with their environment. Co-evolution over many millions of years has led to development of highly efficient chemical mechanisms allowing e.g. plants to defend themselves against all kinds of predating microorganisms, fungi and animals (plants can't run away when they are attacked). Hence, there is a very intimate relationship between the pharmacological potential of such natural products that can be used for the purposes of man and the potential as ecological chemicals developed for the "purpose" of evolutionary success.

The way in which nature has developed such compounds may be compared to the way in which new drugs are developed by chemists: It is a process of optimisation in which new structures are generated on the basis of existent bioactive molecules and "tested" with respect to an improvement of biological performance. The major difference is that natural products are not optimised for the purpose of man but for the goal of maximum ecological success. This optimisation process has continued over millions of years and will certainly not come to an end point before living organisms stop to co-exist on this planet. However, it has (already) led to a high degree of perfection in many instances.

The class of natural products under focus of this work, the sesquiterpene lactones (STLs) represent one prime example, as a very large group of secondary metabolites of high diversity, with respect to chemical structure as well as biological activity. The first section of this contribution ("Structural Diversity of Sesquiterpene Lactones") therefore focuses on a general description and discussion of STL structure in terms of chemical and biological diversity.

The majority of STLs exert their biological activity by a common chemical mechanism of action based on interference with the function of cellular macromolecules through formation of covalent bonds between electrophilic partial structures of the STLs and nucleophilic centres of biological targets (Micheal type addition leading to alkylation). Thus, an important part of this chapter will focus on the mechanism of action, on structure-reactivity relationships and structure-activity relationships of such alkylant STLs (section "Alkylant Sesquiterpene Lactones"). One group of alkylant STLs with a highly specific activity towards malaria parasites, the artemisininoids, are treated separately, since the chemical reaction by which they react with biomolecules differs from that of the majority of STLs. These compounds, their mechanism of action and SAR are introduced under "Artemisininoids". Apart from the large group of alkylant STLs, some smaller groups of STLs exist whose biological activity is not related to alkylation of biomolecules but to non-covalent interaction with particular receptors. Here, thapsigargin and similar guaianolides, specific inhibitors of intracellular Ca^{2+}-ATPases, are discussed under "Thapsigargin and Analogues". Finally, common structure

activity relationships of two different types of sesquiterpene lactones, the *seco*-prezizaanes and the picrotoxanes, which possess convulsant, insecticidal and antiparasitic activity by antagonizing GABA-coupled chloride channels in the central nervous system of various organisms will be discussed under "GABA-Antagonistic Sesquiterpene Lactones".

STRUCTURAL DIVERSITY OF SESQUITERPENE LACTONES

Chemical and biological diversity

Biological diversity (i.e. variability of the geno- and phenotype) within a group of organisms is related to their capability of adaptation to different environments.

Chemical diversity can similarly be defined as structural and physicochemical variability within a group of related chemicals, e.g. chemicals derived from a common set of reactions or -in terms of natural products- derived from one common basic biogenetic pathway. It is a general rule that nature (very much like synthetic chemists) is parsimonious, i.e. diversity is usually created by as few and as simple changes as possible. In terms of chemical diversity, it can be expected that as few as possible simple reaction steps are taken leading to as much as possible structural variation. A diverse group of natural products can be considered as a set of simple variations on a common theme where only slight changes were introduced to create new chemical entities. As all changes in the phenotype, the changes that lead to chemical diversity are due to changes in genotype which are the consequence of random mutation/crossover events. The conservation of such changes follows the rules of natural selection. Hence, only such new phenotypes (in our context: chemotypes) are maintained during evolution which are of ecological advantage for their bearers. Thus, if a great number of different chemical entities are found in one particular biosynthetic group of natural products, this may straightforwardly be interpreted as a tendency of this particular group to lead to ecological and evolutionary success. As can be seen, the chemical diversity of natural products must therefore be related to biological diversity, ecological success and -in case of low-molecular weight secondary metabolites- to biological activity.

The sesquiterpene lactones

The sesquiterpene lactones (STLs) are one such group of natural products displaying great diversity. They have been subject to a vast number of phytochemical studies in the past four decades, mainly due to the fact that many of them show a broad variety of conspicuous biological activities (for reviews see e.g. [1-3]). As a major result of this effort of numerous research groups (it is due to mention the outstanding contributions of the late Ferdinand Bohlmann in this respect) it may be stated that STLs form one of the largest biogenetically homogenous groups of secondary natural products presently known. Currently, the Dictionary of Natural Products (DNP) holds a total of almost 5000 different sesquiterpene structures containing at least one lactone group [4].

The distribution of STLs within the plant kingdom shows a strong focus within the Asteraceae (Compositae) family which can thus be considered the major "pool" of STLs' structural diversity [4, 5]. Asteraceae are among the largest family of higher plants and considered the most highly developed group of dicotyledonous angiosperms (class Rosopsida) [6]. Members of this family are distributed world-wide and populate virtually

any habitat in which vascular plants thrive. It may hence quite straightforwardly be assumed that the structural and biological diversity of STLs is related to the biological diversity of these plants and their ecological and evolutionary success. Since this success based on biological activity must be intimately related to any consideration of structure-activity relationships (SAR), the first part of this chapter is dedicated to a general discussion of the structure and properties of STLs.

The approximately 5000 STL structures currently known represent almost 50% of all DNP entries on sesquiterpenes (approx. 11000). In comparison with the other major classes of terpenoids, where lactones are much less frequent (13%, 24% and 14% of the mono-, di- and triterpenoids, respectively [4]), this high percentage appears quite conspicuous.

A total number of 4861 structures can be considered as STLs *sensu stricto*, i.e. compounds containing a lactone group in the sesquiterpenoid core. Compounds representing e.g. esters of a hydroxycoumarin with a sesquiterpene acid were not counted. Quite interestingly, the majority (n = 4217 or 87 %) of all sesquiterpene lactones is contributed by only 7 of the over 100 sesquiterpene types distinguished by the DNP, which are derived from the carbon skeleta shown in Fig. (**1**). Biogenetically, all these STLs (i.e. the eudesmanolides, elemanolides, eremophilanolides, guaianolides, xanthanolides and pseudoguaianolides) are derived from the germacrane skeleton so that

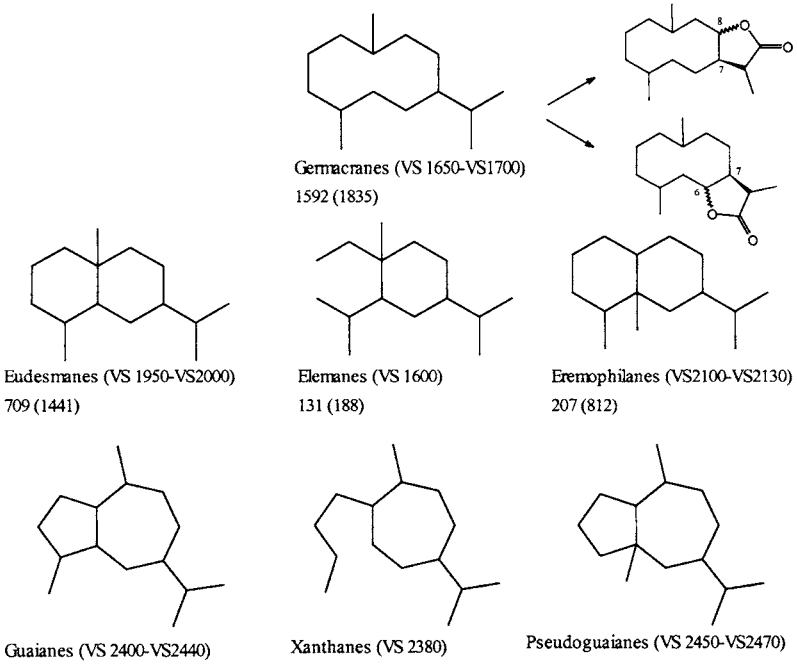


Fig. (1). Major sesquiterpenoid skeletal types representing the structural basis for 87% of all STLs [4]. The numbers indicate the number of lactones, those in brackets the total number of entries in each of the classes (VS1600-VS2470). Common modes of lactonization as found among all these types are shown for the germacrane type. (Note that C-6 and C-8 are actually equivalent in this further unsubstituted representation of the germacrane skeleton which, however, does not occur in nature).

this large group deserves most attention in a general discussion of STL structural diversity. Typically, the compounds from these groups occur in two different modes of lactonization involving the isopropyl side chain including C-7 and C-6 or C-8 leading to a γ-lactone with one pending carbon as the most typical structure element (see Fig. (**1**)). In almost all STLs of these types isolated from higher plants, the configuration at C-7 is as shown. However, lactonization in both directions introduces a chiral centre at C-6 or C-8 and examples for each epimer exist, so that 28 subtypes result from the shown skeleta without further substitution. Introduction of further substituents at any position produces further chiral centres, so that the possible variations are countless.

By means of modern computer technology, it is now possible to extract and handle structural information from large databases and thus, for this article, all STL structures were extracted from the Dictionary of Natural Products (rel. 10:2, 2002) and compiled into a database. This database contains the 2D structure of each of the 4861 STL based on the SMILES [7] formula (available as SDF file from the author on request). On grounds of the information inherent in these molecular structures it is now possible for the first time to characterise the basic chemical properties for the whole class of natural products and thereby to obtain new insights into their chemical diversity and the structural features underlying their biological activity.

Most STLs are "drug-like" molecules

One of the most important issues in the process of identification of new potential drug molecules is assessment of their physicochemical properties with respect to (expected) pharmacokinetics. Many new bioactive compounds that show promising effects *in-vitro* have to be abandoned at later stages of development because of poor pharmacokinetic characteristics, most often because of insufficient gastrointestinal absorption after p.o. administration. It is therefore desirable to assess such properties at an earlier stage, optimally even before biological tests in order to have only such compounds included in pharmacodynamic testing whose absorption, distribution, metabolism and elimination (ADME) properties are at least estimated to be appropriate. To this end, estimation of such drug characteristics by theoretical modelling techniques is a field of highly active research [8].

In the case of natural products with the major purpose of chemical defense, it can be expected that they have been optimized in the course of evolution not only with respect to pharmacodynamic but also to such pharmacokinetic aspects, i.e. many if not all compounds of such a group should show physicochemical properties leading to good bioavailability.

It was therefore of interest to quantify such properties for the STL database. A wide variety of parameters can be calculated in relation to ADME characteristics, the most well-known (and simplest) of which is probably Lipinsky's "Rule of Five" [9], predicting that poor absorption is likely to occur whenever two or more of the following rules 1.-4. are invoked:

1. Hydrogen-bond donors (NH or OH groups) > 5
2. Hydrogen-bond acceptors (N or O atoms) > 2 x 5 = 10
3. Molecular weight > 500
4. (calculated) octanol/water partition coefficient log P > 5.0
5. (as always, there are exceptions from this rule, e.g. caused by active transport of the compound via specific carriers).

It is quite noteworthy that within the STL database, only an insignificant fraction of the molecules appear as poor drug candidates from this point of view: Only 115 of the 4861 compounds (2.4 %) meet two or more of these points, i.e. conflict with Lipinski's rule. In fact, only 211 STL (4.3 %) violate any *one* of the mentioned criteria. A total of 189 compounds possess a molecular weight > 500, the number of H-bond acceptors is >10 in 126 molecules, while a number of H-bond donors >5 occurs only in 15 and a log P > 5 only in 28 cases. Thus, based on these simple qualitative criteria, the class of STLs as a whole appears to be a very "drug-like" family of natural products with respect to their expected absorption properties.

In order to investigate the major determinants of chemical diversity of STL, 145 different structural and molecular properties (descriptors) were calculated for the database using the program MOE [10]. The resulting data table was submitted to principal component analysis (PCA) [11, 12]. PCA is the basic method of chemometrics, and widely used as a mathematical tool to investigate the diversity of large sets of chemical data [11, 12]. Basically, the information content of a data table containing a large number of properties or features (here: descriptor variables) measured or calculated for a set of objects/samples (here: compounds) represents a multidimensional space with as many dimensions as variables considered. Each object thus has a definite position within this space which is determined by its values for each variable. The individual variables thus determine the distance between the samples in feature space and hence the diversity of the data set. In cases where the number of variables is large (typically larger than the number of samples), it is very difficult to assess the distance ((dis)similarity) between the objects and to interpret quantitatively the influence of the variables on the objects' positions. In such cases, PCA can be used to reduce the dimensionality of the complex feature space by combining the information content in so called latent variables or principal components (PCs) which form the dimensions in a new feature space, i.e. principal component space. The PCs have the advantage of being orthogonal (i.e. non-correlated with each other, in contrast with the original variables which may be mutually correlated so that many of them may contain similar information) and thus, in most cases, few PCs contain the major part of the information of the original data table. It becomes then possible to determine the distance between features (similarity/dissimilarity) in a feature space of low dimensionality (e.g. in a 3D plot of each compounds values (scores) in the first three PCs which often explain the major part of the variance in a data set), and to directly assess the original variables' influence on the object's positions (e.g.in a corresponding 3D plot of each feature's influence (loadings) in the same PCs).

The first two principal components from the PCA for 4861 STLs with respect to the mentioned 145 features (Fig. (**2**) **A**) were found to be related to descriptors of molecular size (PC1, 50% variance explained) and to lipophilicity (PC2, 10% variance explained). It turned out that a very similar picture as obtained by plotting the scores in PC2 vs. those of PC1 resulted by plotting two of the original variables, logP(o/w) vs. the van der Waals volume (vdw_vol; Fig (**2**) **B**). The 4861 compounds form four major clusters with respect to molecular size, while logP(o/w) follows a relatively sharp normal distribution with mean ± sd = 1.5 ± 1.08 over the entire database. Histograms showing the numbers of compounds within certain intervals of vdw_vol and logP(o/w) are shown in Fig. (**2**) **C**.

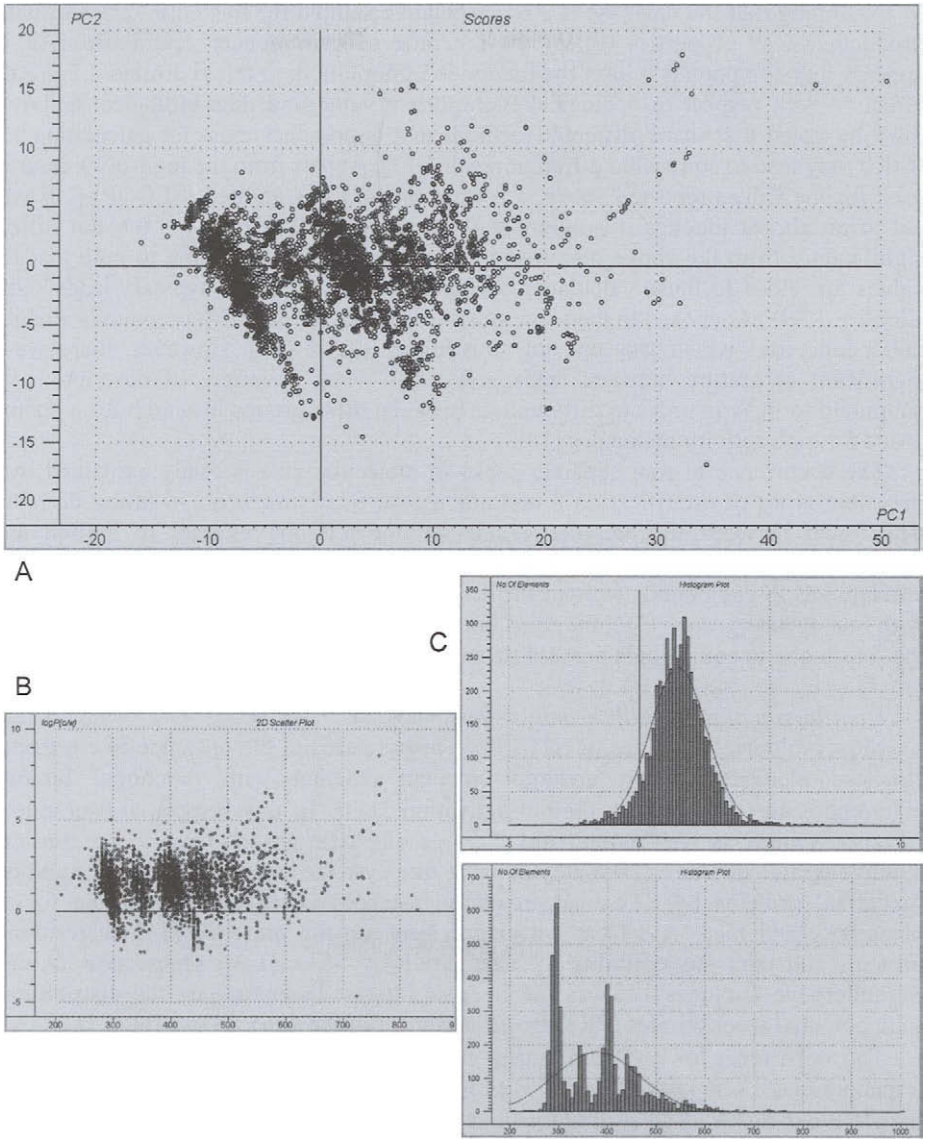


Fig. (2). Molecular size and lipophilicity are the main sources of diversity of STL among 145 2D descriptors of molecular structure and properties.
A: Result of a principal component analysis (PCA) over all desciptors: Scores for PC2 plotted vs those of PC1. PC1 accounts for 50 % of the variance, PC2 for further 10 %.
B: Scatter plot of logP(o/w) vs. van der Waals volume yielding an almost identical picture due to the fact that size and lipophilicty are the major determinants of PC1 and PC2.
C: Histogram plots of logP(o/w) (left) and van der Waals volume (right) showing the frequency of compounds within certain intervals of the two descriptors.

Assuming that the database is a representative sample for this entire class of natural products whose properties reflect the outcome of evolutionary optimization, it thus appears that compounds within the mentioned lipophilicity interval around 1.5 are most effective with respect to ecological/evolutionary value and thus biological activity. It must be noted that many different methods and approaches exist for calculating logP, which may lead to somewhat different results [13]. Appart from the logP(o/w) descriptor used for the values reported above, a second method was applied (SLOGP) [10] which led to an almost identical histogram with mean SLOGP = 1.84 ± 1.09, not differing significantly from the above mentioned values. It is quite interesting to note that these values are close to those calculated by Oprea for a set of 176 typical "leads" (mean ClogP = 1.79) as opposed to the value obtained with 532 "drugs" (compounds already in therapeutic use) which was one log unit higher (2.62) [14]. However, there was no significant separation between leads and drugs when a variety of descriptors were submitted to PCA in order to differentiate between these groups in search for a chemical space for high activity molecules [14].

The occurrence of four separate peaks in molecular size is easily explained by the prevalent types of esterification found among the STL structures. A strong correlation was found between the van der Waals volume and the number of carbon atoms (correlation coefficient $r^2 = 0.96$). The largest peak in the vdw_vol histogram roughly corresponds to compounds without further side chains (C_{15}), the next one to compounds with one acetate group (C_{17}), the third one to esters with a C_5-carboxylic acid (C_{20}) and the fourth one to compounds possessing an acetate and a C_5-ester group (C_{22}).

Distribution of potentially reactive sites (PRS)

Most biological activities of STLs have been related to the presence of electrophilic structure elements, which undergo covalent reaction with functional biological macromolecules resulting in their deactivation [1-3]. In this respect, α,β-unsaturated carbonyl groups as well as epoxide and free aldehyde groups have to be considered reactive partial structures. The alkylation of free cysteine residues in enzymes and other functional proteins by STLs has in many instances been held responsible for STL bioactivity and there is a clear correlation between the presence of such residues in proteins and their susceptibility to inactivation by STLs [1-3], see section "Alkylant Sesquiterpene Lactones". It was therefore of interest to investigate the distribution of such potential reactive sites (PRS) in the structures of the 4861 STLs (Table 1).

The occurrence of α,β-unsaturated carbonyl structures in comparison with other terpenoid classes, is rather high. A very high percentage, 63 % of all sesquiterpenes in the DNP and 81 % of all sesquiterpenoid *lactones*, contain at least one α,β-unsaturated carbonyl structure element, while these numbers are lower in case of the mono-, di- and triterpenes (10/75%, 14/59% and 7/51%, respectively (numbers based on DNP [4]). However, it becomes evident from these figures that lactone groups tend to be associated with α,β-unsaturated carbonyls throughout all classes of terpenoids.

Among the 3950 STLs possessing *at least one* α,β-unsaturated carbonyl group, the numbers of compounds with 1 to 5 α,β-unsaturated carbonyl structures are 2150, 1354, 388, 51 and 7, respectively, accounting for an average number of 1.28 α,β-unsaturated reactive centers per molecule. 3748 compounds are α,β-unsaturated carboxyl derivatives (lactones and esters), the conjugated lactone group occurring 3501 and a conjugated

ester group 1469 times. The most frequent α,β-unsaturated lactone moiety is the α-methylene-γ-butyrolactone ring (2724 entries = 56 % of all STLs, of which 2598 are unsubstituted at the exocyclic carbon). Further frequent α,β-unsaturated lactone types are γ- lactones with a conjugated double bond inside the ring (747). Conjugated δ-lactones occur less frequently (1.5 %). Among the α,β-unsaturated ketones (688), cyclic ones are most frequent (652) and among these, the cyclopentenone and cyclodecenone structures are the prevalent types (both over 300 compounds).

Within the large group of 4217 germacrane-derived STLs (i.e. germacranolides, eudesmanolides, elemanolides, eremophilanolides, guaianolides (including 4,5-seco guaianolides = xanthanolides) and pseudoguaianolides, see Fig. 1), 85% contain at least one and 41% more than one α,β-unsaturated carbonyl structure. Over all other groups of STLs, these figures are considerably lower (54 and 8%, respectively).

Another frequent PRS is the epoxy group, occurring at least once in 921 compounds (19%), while chlorohydrins, likely to have resulted from acid hydrolysis of the former, occur in 65 structures (the number of 1,2-diols and derivatives thereof is much higher but it is not possible to differentiate between such that result from hydrolysis of epoxides and which do not). Finally, a free aldehyde group, as another potential alkylant structure element occurs in 240 compounds (209 with α,β-unsaturated aldehyde group).

Overall, these figures show the tendency of STL structure towards accumulation of alkylant structure elements, which has possibly been increased in the course of evolution of STL-producing plants. The observation that about 45 % of all compounds with an α,β-unsaturated carbonyl possess more than one such PRS can be explained as a likely consequence of the fact that different electrophilic centers such as e.g. α,β-unsaturated lactone and -ketone structures possess different reactivity. Variations in the chemical environment of nucleophilic centers in biological target molecules cause different reactivities (see "Alkylant Sesquiterpene Lactones") so that a compound possessing more than one reactive centre is capable of reacting with a larger variety of different targets. It will thus become a more efficacious tool with respect to its function in chemical defence. Consequently, it was of interest to study the occurrence of combinations of such reactive centers, and to find out whether certain combinations are more frequent than others. Assuming that high numbers of occurrence are related with high ecological value, this might reveal whether particular combinations render STLs more efficient defense tools than others. To this end, 16 descriptors encoding the presence of the various reactive structure elements were taken into account. Besides the various types of α,β-unsaturated carbonyl structures, epoxy-, aldehyde- and peroxy groups were considered as potentially reactive structure elements (see Fig. (**3**)).

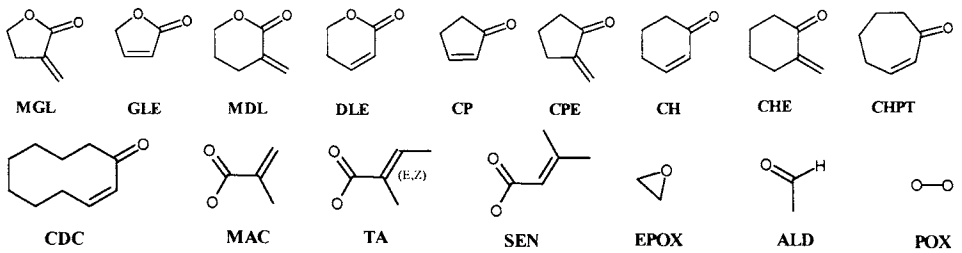


Fig. (3). Potential reactive structure elements (PRS) taken into account in the statistical analysis.

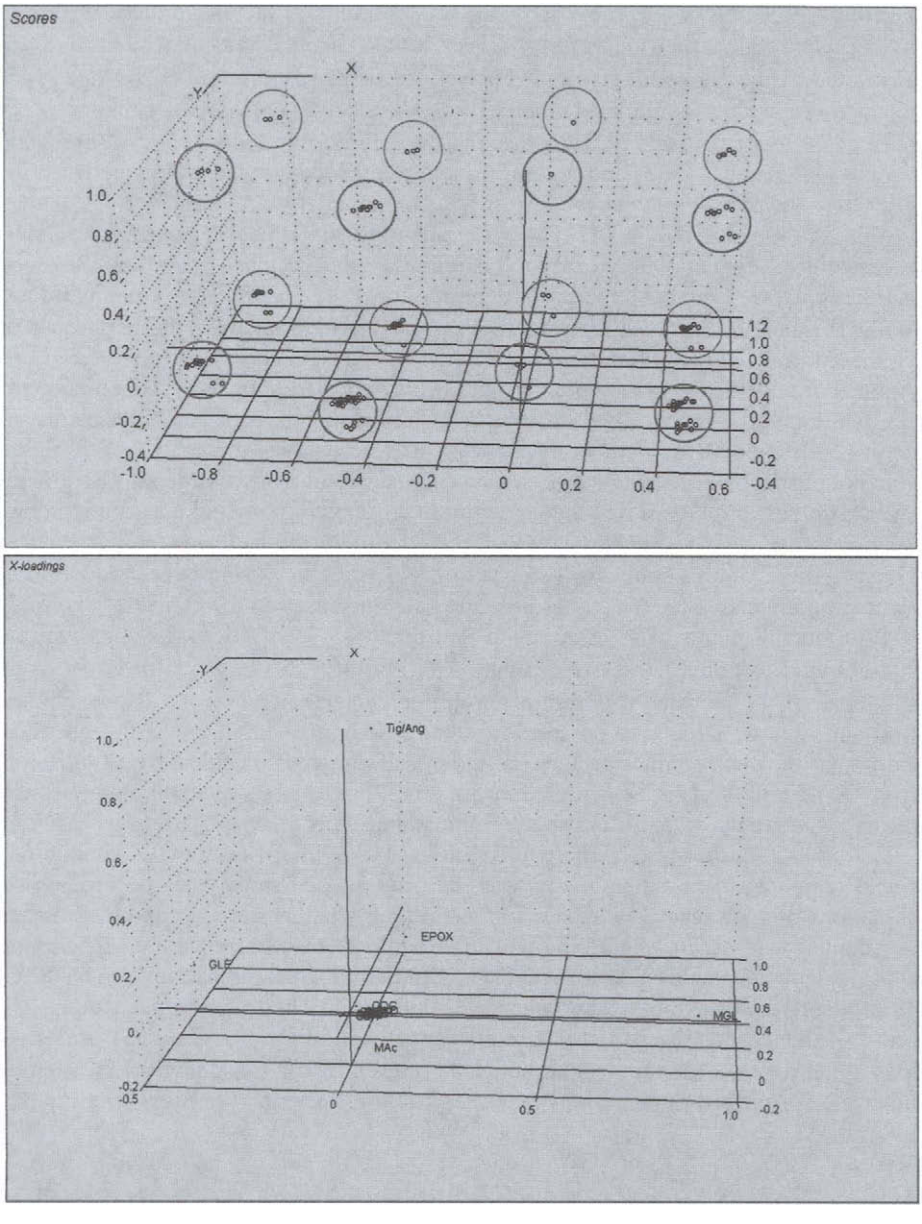


Fig. (4). Principal component analysis (PCA) with respect to the occurrence of various potential reactive sites in 4861 STL. Scores (**A**) and loadings (**B**) of the first three PCs are plotted in X, Y and Z direction. The scores show the clustering of the compounds into 16 major groups while the loadings show the influence of the variables on the position of these clusters in feature space and the distinction between the clusters.

A data table was created in which a value of 1 was assigned to the respective data column in case of presence of a particular structure element, while absence of that element was assigned zero. The resulting data matrix (4861 x 16) was then submitted to principal component analysis (PCA) to identify the major sources of variance with respect to the occurrence of potentially reactive structure elements.

The first principal component PC1 (see Fig. (**4**)) explains 32 % of the variance and receives significant contributions from variables MGL and GLE. PC1 divides the database into four groups: compounds with an MGL and no GLE (n= 2681), such with both, MGL and GLE (43), such without either MGL or GLE (1433) and such possessing a GLE and no MGL (704). Both variables contribute to the PC in opposite directions which mirrors the fact that both structure elements are commonly formed by carbon atoms of the isoprop(en)yl side chain of germacrane-derived STL and are thus in most cases mutually exclusive.

The second most important principal component PC2 (17 % variance explained) divides each of these groups into compounds possessing an epoxy group and such without this feature. PC3 (13 % variance explained) contains information on the presence/absence of a tiglate or angelate group (TA). Thus, the major part of the variance (62 %) is explained by only four of the mentioned alkylant centers and divides the database into 16 separate clusters, as shown in Fig. (**4**) and Table 1.

Table 1. Major combinations of potential reactive structure elements (PRS) among 4861 STLs.
Compare clusters in 3D scores plot in Fig. (**4**).

Cluster #	MGL	GLE	EPOX	TA	n
1	+	-	+	+	97
2	+	-	-	+	324
3	+	-	+	-	433
4	+	-	-	-	1827*
5	+	+	+	+	1
6	+	+	-	+	2
7	+	+	+	-	14
8	+	+	-	-	26
9	-	-	+	+	19
10	-	-	-	+	124
11	-	-	+	-	189
12	-	-	-	-	1101**
13	-	+	+	+	30
14	-	+	-	+	73
15	-	+	+	-	138
16	-	+	-	-	463

* Within this group, 1300 are true monofunctional MGL, while 527 represent combinations of MGL with other PRS taken into account (see text).

** Of these compounds, 826 are devoid of any PRS, while 275 represent combinations of the other PRS taken into account (see Fig. (**3**)).

The most frequent combination of the MGL is with an epoxy group (n=545). Combination of an MGL with a tiglic/angelic ester occurs 424 times. In the majority of MGL-containing compounds (n=1827) this PRS is not combined with either GLE, EPOX, or TA. Of these structures, 1300 do not contain a combination with any other of

the PRS taken into consideration, i.e. they are true monofunctional alkylants with electrophilic reactivity based solely on MGL. The remaining 527 compounds are bifunctional (n=455), trifunctional (n=70) and even tetrafunctional (n=2) combinations with those PRS not contributing significantly to PC1 and PC2. The major combinations of MGL in this group are with a methacrylic ester group (MAC n = 176), with a cyclopentenone (CP, n = 111) and with an aldehyde function (ALD, n = 116).

The considerable number of monofunctional MGL-compounds (28% of all STL in the database) probably reflects the capability of the MGL to render the compounds sufficiently reactive to yield quite an effective chemical protection. It should be mentioned at this point that the MGL structure can be considered a primitive character in terms of evolution, since it occurs already in STLs isolated from liverworts, e.g. the frullanolides [4]. It may thus be speculated that compounds possessing this structure element have represented the starting point in the evolution of the biosynthetic pathways leading to more complex alkylants.

The major group of "oligofunctional" STLs combines the MGL with an epoxy group. This may reflect the fact that MGL is a soft electrophile whose reactivity is directed towards soft nucleophiles such as SH groups, while the epoxide is a very hard electrophile which reacts preferably with hard nucleophiles such as OH- and NH groups. (see section "Alkylant Sesquiterpene Lactones".) It may therefore be assumed that the combination of two PRS with fundamentally different reactivity renders these compounds more versatile with respect to their biological targets and thus maximises their biological efficacy.

It can be expected that the basic reactivity of MGL and α,β-unsaturated ester groups (TA, MAC) is quite similar. In both cases the conjugated enone is in a very similar chemical environment where it is partially deactivated by electron density from the sp^3 oxygen bound to the carbonyl, so that in comparison to α,β-unsaturated ketones, both are weaker electrophiles. The fact that MGL is combined with α,β-unsaturated ester groups in so many instances in spite of this similarity might be explained by the following hypotheses: (a) in order to introduce a second PRS it is biosynthetically more simple to attach such an ester group to an OH group of an STL molecule than to modify the STL skeleton itself. (b) The ester group attached to an STL may rotate more or less freely and is thus more flexible to reach nucleophilic centers of biomolecules in a larger variety of spatial arrangements. Moreover, in contrast to the MGL group, the geometry of the enone system (i.e. relative position of the carbonyl oxygen and the reactive β-carbon) in the ester groups is not confined to one arrangement which might also enhance versatility with respect to the number of possible targets (compare discussion of our findings on the reactivity of MGL towards free cysteine in section "Alkylant Sesquiterpene Lactones").

In these considerations on STL structure, one important aspect was not taken into account so far, namely, that of explicit stereochemistry. The STL database consists, as mentioned above, of 2D representations in the form of the SMILES string. It is at present not possible to include structural features based on molecular geometry, such as shape, charge distribution, etc. in our investigation on the molecular properties of STLs as a whole, since this would necessitate the generation of 3D structures with the correct stereochemistry for each of the 4861 molecules, which would only be possible by

manually re-modelling each compound. It is easy to imagine, however, that a very important part of structural diversity among these molecules results from their high variability in stereochemistry. Thus, 807 structures in the database represent duplicates, i.e stereoisomers whose 2D structure is identical with that of one of the remaining 4054 compounds. Apart from the possibility to generate diversity by changing the stereochemistry, i.e. by generating a different compound through inversion of one or more chiral centres, some very interesting examples exist for STLs in which a certain stereochemical arrangement introduces structural diversity in one particular molecule. Thus, e.g. in case of 1,10-*trans*-4,5-*trans*-germacradiene-7(12)8-cis-olides, the co-existence of up to four different conformers has been reported which can all be observed in the NMR spectra at room temperature [15, 16]. Thus, it can be expected with respect to molecular recognition (i.e. binding to different target structures), that all four geometries play a role which will certainly widen the efficacy of such compounds as defence tools. On the other hand, the fact that all the conformers can be detected as separate species in the NMR spectra at room temperature (i.e. slow exchange), is related to considerable activation barriers between these different forms of the compound. This might hinder a molecule adopting one conformation to bind to a target which requires just one of the other possible forms to fit. It might thus be even more advantageuous to have full flexibility at ambient temperature. An example is found in the 10-α-methyl pseudoguaianolide (=helenanolide) series. Here it could be shown in the case of helenalin (**HEL**, structure see Fig. (**4**), next section) and its ester derivatives that they exist in an equilibrium of two different conformations exchanging at room temperature with very high rates between 10^5 and 10^6 sec^{-1}, which we could prove by low-temperature NMR studies [17]. Further investigations on the dependence of these compounds' conformation on their substitution pattern have shown that such fast exchange equilibria occur only if both C=C double bonds are present, i.e. if the molecule has two PRS. Compounds of the 11,13-dihydrohelenalin- and 2,3-dihydrohelenalin series adopt only one conformation [18]. Thus, helenalin combines two advantages with respect to efficacy as a chemical defence tool, namely, alkylant bifunctionality and conformational flexibility. Implications with respect to the high bioactivity of **HEL** in comparison with less flexible compounds, e.g. its 6,8-diastereomer mexicanin I, have been discussed [17]. In fact, **HEL** is among the most active compounds whenever tested in comparison with others (see e.g. [3] and literature cited with respect to biological activity of **HEL**, next section).

It appears an intriguing "invention" by nature to synthesize one compound and end up with three or four "different" defense tools at the same time. Combination of the differential reactivity of several individual PRS with high flexibility thus seems to represent the current "state of the art" with respect to evolutionary development of STLs.

This is parsimony which would not be in the sense of a synthetic chemist synthesizing a new drug molecule: It is not directed towards specificity but serves rather the opposite: to obtain a broad-band chemical mace as efficacious as possible against as many targets as possible. This section has shown the enormous versatility of STL structure and the major determinants of their structural diversity. In this light on the outstanding success of STLs as secondary natural products in evolution the next section will focus in more detail on reactivity as the basis of this success, and on particular structure-activity relationships associated therewith.

ALKYLANT SESQUITERPENE LACTONES

Evidence pointing towards a general alkylant mechanism underlying the majority of biological effects of STLs dates back to the pioneering work of Kupchan and others in the early 1970s [19-21]. A large number of reports have since dealt with or at least taken into account this simple mechanism of (re)action underlying the most prominent activities of STL [3].

Structure-reactivity relationships

As mentioned in the previous section, α,β-unsaturated carbonyl structures readily react with free sulfhydryl groups of low-molecular weight thiols and most probably also of proteins, since most proteins whose activity is affected by STL possess free cysteine groups (for an overview see [3]).

Nucleophiles other than SH groups (amino groups in basic amino acids and in nucleic acids, as well as OH groups) were generally found to be much less susceptible to alkylation by STL which is in agreement with the classification of the SH group and enone as soft nucleophile and electrophile, respectively. According to Pearson's HSAB concept, soft nucleophiles and electrophiles should react more readily with each other than with reactands classified as hard electrophiles and nucleophiles, e.g. non-conjugated carbonyl groups and amino or hydroxyl groups, respectively [22, 23]. Several studies have sown that reactivity towards free amino groups of amino acids and nucleosides is low and does not play a role for the biological activity of most STLs (e.g. [24, 25]). As an exception, compounds possessing (masked) aldehyde groups, such as hymenoxon, react with DNA [26, 27]. Reactivity towards alcoholic OH groups is extremely low in comparison with SH groups. Own results on the stability of ethanolic *Arnica* tincture have shown that the reaction half-live of the cyclopentenone group of 11α,13-dihydrohelenalin esters such as **DHAC** (Fig. (**4**)) is in the order of several years (>3 y in 70 % EtOH at room temperature), when dissolved in ethanolic solution [28]. In order to obtain significant amounts of 2-ethoxy- and 2-isopropyloxy derivatives of helenanolides (structures **12** and **13** in Fig (**9**)), aluminium chloride had to be used as catalyst as reported in [28]. It appears noteworthy, that under these conditions, only the cyclopentenone showed addition of the alcohol while the methylene lactone structure of helenalin (**HEL**) did not yield any adduct [29]. Thus, reaction with OH-nucleophiles is irrelevant with respect to biological activity of STLs. In studies on the reactivity towards biologically relevant nucleophiles our main focus was hence on thiol compounds.

The physiologically important tripeptide glutathione (gly-cys-γ-glu, GSH) has frequently been reported to undergo Michael type alkylation by STL [24, 25, 30-33].

Our own work has shown that great differences exist between different alkylant centers of helenanolide type STLs (Fig. (**4**)) with respect to their reactivity towards GSH [31, 34, 35]. Thus, helenalin (**HEL**) in the reaction with the cysteine-SH group of GSH displayed high regio- and stereoselectivity. It reacted exclusively via the cyclopentenone group (CP) and yielded a single adduct **HEL-SG** when exposed to equimolar concentration of GSH. The methylene-γ-lactone group (MGL) only underwent reaction at higher GSH concentration and at a much lower reaction rate [31] leading to the bis-GSH adduct **HEL-SG$_2$**.

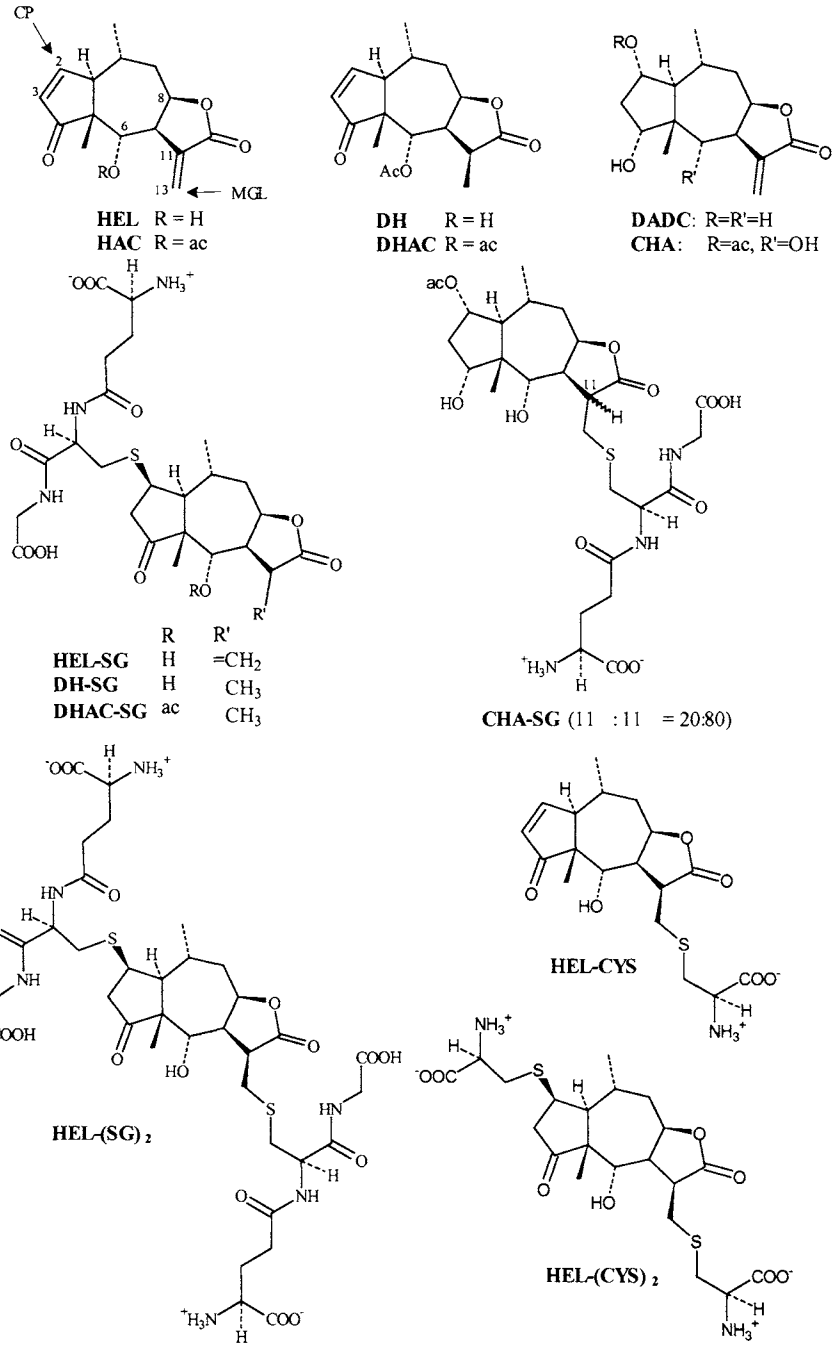
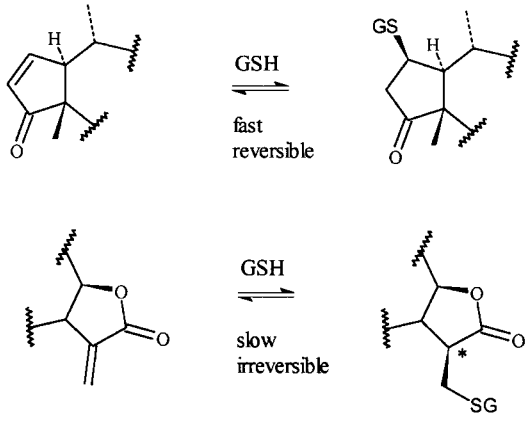


Fig. (4). Structures of representative helenanolide STLs and their thiol adducts

Studies on the role of GSH addition under physiological conditions were conducted since it was an open question how the low micromolar concentrations generally required to invoke the biological effects of STLs should inactivate proteins inside cells in the presence of the very high intracellular concentrations of GSH (0.5 – 10 mM [36]). We determined the reaction kinetics between GSH and STL possessing either a MGL or a CP group at various pH values and relative concentrations [34]. At physiological pH, the reaction for the CP-compound (**DHAC**) was found to occur so fast that the STL should be deactivated quite rapidly after entering into a cell. However, we found that adduct formation at the CP moiety is fully reversible at this pH. When incubated at equimolar concentrations, GSH and **DHAC** quite rapidly reached an equilibrium with approximately equal amounts of the free STL and the adduct. Equilibration to the same ratio was found when the isolated GSH-adduct of **DHAC** was incubated in buffer. Thus, even at high excess GSH a certain fraction of cyclopentenone-containing molecules will be present in the unreacted form and thus be able to react with other, more vital targets. Conversely, the reaction of the MGL-compound (**DADC**) was slower by a factor of ca 15 [34] but this reaction turned out to be practically irreversible [35]. Thus, STLs with MGL groups are more efficiently deactivated by GSH addition but due to the slow reaction kinetics they have more time to react with sensitive proteins after penetrating into a cell.

In agreement with these chemical results, which are summarized in Fig. (**5**), the biological activity of STL-GSH adducts was found to be similar to those of the unreacted STL in case of helenalin (reversible reaction at CP) [34, 35] while the adduct of chamissonolide **CHA** (irreversible addition at MGL) was found biologically inactive [35]. These findings are of high importance with respect to the explanation of certain differences in the toxicity profiles of helenanolide STLs [35] and differential effects of CP and MGL-containing STLs on gene expression profiles [37], which will be discussed below.



[a,b] reaction half-lives ($t_{1/2}$) at c(**HEL**) = 100 μM, c(GSH) = 500 μM, pH 7.0: CP: 6 min; MGL: 90 min [34].

Fig. (5). Differential reactivity of CP- and MGL-containing helenanolides towards glutathione (GSH).

It is highly noteworthy that completely different results from those with GSH were obtained with free cysteine (cys). In case of the free amino acid, the MGL of **HEL** reacted at an extremely high rate (about 100 times faster than at the CP group) yielding exclusively a C-13 monoadduct (**HEL-CYS**) when equimolar concentrations of the reactands were applied. Only at excess cys, formation of the bis-adduct **HEL-CYS$_2$** was observable, and the reaction at the CP group proceeded at a similar rate as with GSH [31]. Molecular orbital calculations indicated that a special mechanism underlies the reaction at the MGL (see Fig. (**6**)), in which the methylene lactone's reactive site is activated by an interaction of the carbonyl oxygen with the amino acid's protonated amino group which leads to electron withdrawal from the conjugated system and thus to higher reactivity of the β-carbon. The geometry of cys and the MGL group allows simultaneous occurrence of this interaction and approach of the sulphur atom to the β-carbon which explains the high reactivity [31] (see Fig. (**6**)). This would not be possible at the CP group due to the *trans* arrangement of the C=C and C=O double bonds. Experiments with N-acetylcysteine (cysNAc) and the methylene lactone **DADC** confirmed this proposed mechanism. CysNAc due to the N-acetylation is not able to activate the MGL group as described and, consistently, after 20 h incubation at equimolar concentration no adduct formation of **DADC** was observed with cysNAc while free cys led to complete reaction within minutes.

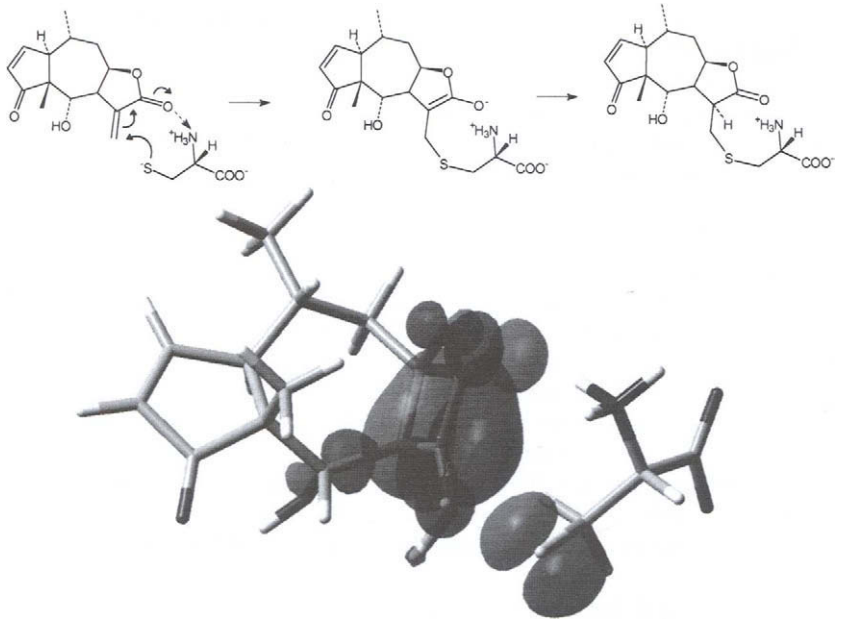


Fig. (6). Reaction mechanism observed during reaction of **HEL** and L-cysteine. The 3D-model shows the transition structure obtained by energy minimisation of the cysteine thiolate anion and a molecule of helenalin using the semiempirical AM1 hamiltonian. The surface representation of the highest occupied molecular orbital (HOMO) shows that this structure corresponds to the middle formula (above) in which the major part of the π-electron density is transferred to the 11,12 bond. Note the proximity of the NH_3^+ to the lactone carbonyl oxygen.

This finding readily explains the frequent observation by others, that free cys can be used to completely prevent the major biological effects of STL [19, 38-40]. With respect to structure-activity relationships, the result of Kupchan et al., who reported that no correlation exists between the reactivity of STLs with free cys and their cytotoxicity [20] may also be explained by our observation that cysteine addition obviously represents a special case and is dramatically influenced by the right fit of the two reactands. Consequently, it can be expected that reactivity towards a single thiol compound cannot serve as a model for a particular bioactivity of STL since cysteine groups occur in many different chemical environments in proteins. Stabilizing effects leading to enhancement of a particular biological acitivity of a particular STL will be greatly dependent on the environment within the protein and the capability of the STL molecule to form stabilizing (non-covalent) interactions with this environment.

A further interesting new finding related to the general biological activity of STL emerged from the work on GSH addition when we investigated the possible involvement of glutathione-S-transferases (GST) in the formation of STL-GSH adducts. GST catalyze the reaction of various electrophiles with GSH in order to detoxify such agents quickly and with high efficiency. It was thus an important question whether STL would also be conjugated with GSH more quickly in the presence of GST. No effect, whatsoever, was found on the rate of reaction between GSH and either the CP or the MGL group of **HEL**. Quite strikingly, however, the isolated mono- and bis-GSH adducts of **HEL** were found to be inhibitors of the enzyme preparation. The monoadduct **HEL-SG** and the adduct of **DHAC**, **DHAC-SG**, showed IC_{50} values of 15.6 and 1.1 μM, respectively [41]. The latter can thus be classified as potent GST inhibitor [42]. Since we had previously demonstrated the reversibility of GSH addition to the CP moiety, it was important to investigate whether the activity was due to the adducts or to some free STL released from them. In a modified experiment under conditions where no spontaneous adduct formation would occur in the solution, it was shown that the free STL do not inhibit GST and that the inhibition is essentially due to the adducts [41]. The inhibitory activity is hence dependent on the presence of the GSH-moiety which can be expected to mediate binding of the adducts to the GSH binding site of the enzyme.

It was thus shown for the first time, that GSH adducts of STL possess their own biological activity and may thus be involved in the overall effects observed after administration of STL. They should thus not be considered mere deactivation products. Although the adducts did not show any inhibitiory activity in experiments with glutathione reductase and glyoxalase [41], it can be expected that further enzymes involved in GSH metabolism will be affected by such adducts in a similar manner. STL-GSH adducts might thus even be considered as "Trojan horses" that can bring the STL molecules close to active centers of certain enzymes with GSH binding sites.

More recently, and hitherto unreported, we studied the reaction of **HEL** towards another eminently important physiological thiol, coenzyme A (CoASH). Under essentially the same conditions as reported in our NMR study on the reaction with GSH [31], 17 mM **HEL** was exposed to a twofold molar excess of CoASH in D_2O and the reaction monitored by ^{1}H-NMR. It was found that **HEL** reacts exclusively at C-2 of the cyclopentenone moiety. After 24h, 44% of the **HEL** had reacted. The ^{1}H-NMR spectrum obtained after completion of the reaction (approx. 10 days) is depicted in Fig. (**7**) and the

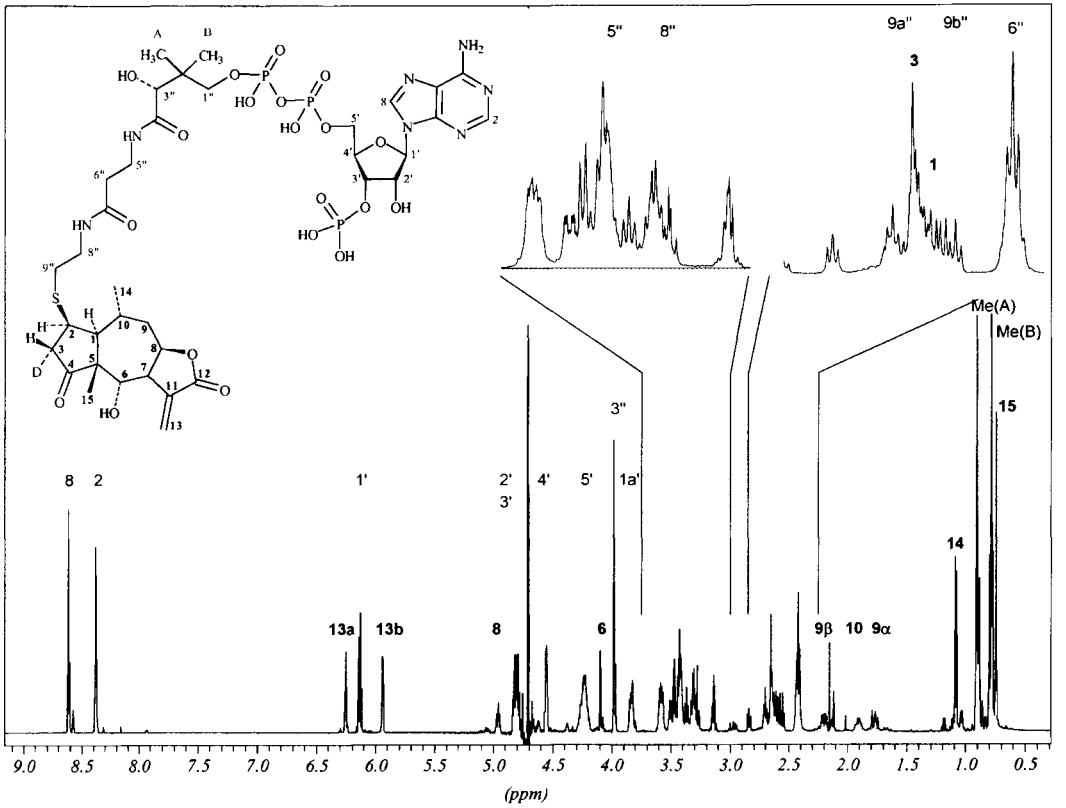


Fig. (7). Structure and ^{1}H-NMR spectrum of the covalent adduct of helenalin and coenzyme A formed in unbuffered D_2O solution. Signal assignment is based on ^{1}H/^{1}H-COSY spectra in comparison with unreacted compounds and changes of signals observed during the reaction.

data are presented in Table 2. It is obvious that one single adduct has formed at C-2 of **HEL**. In spite of the excess amount of CoASH, at best traces (if any at all) of a C-2,C-13 bis-adduct could be detected, so that this reaction in analgous manner as that with GSH is highly regioselective. The stereochemistry at C-2 of the adduct could be determined from the $^1H/^1H$-coupling constants in analogy to the GSH-adduct [31]. The newly formed signal of H-3 appears as a singlet so that it should be in the same position (H-3α) as in case of the GSH adducts, leading to a small coupling constant (<1 Hz) with H-2 which must hence be α-oriented so that the CoA-moiety is in the β-position [31]. Thus, the reaction also shows the same stereoselectivity as in case of GSH addition, i.e. attack of the nucleophile proceeded from the β-side of **HEL**.

Considering the crucial role that CoASH plays in primary cell metabolism, it appears straightforward to assume involvement of CoASH-addition in the biological activity of such compounds as helenalin, which will have to be shown in further studies. In a similar way as demonstrated for the GSH-adducts [41], CoASH adducts could influence CoA-dependent enzymes and thus contribute to the many known effects of STLs on processes of energy storage and -recovery, such as inhibition of oxidative

Table 2. ^1H NMR chemical shift values of the sesquiterpene lactone moiety in C-2-adducts of helenalin (HEL).

	HEL-SG[1]			HEL-COA[2]			HEL-PAP[2]		
proton	δ(ppm)	mult	J(Hz)	δ(ppm)	mult	J(Hz)	δ(ppm)	mult	J(Hz)
1	2.65	dd	6.5, 11.2	2.60	dd	6.3, 11.1	2.20	n.d.	n.d.
2	3.54	d	6.5	3.47	d	6.3	3.53	n.d.	n.d.
3	2.73	s*		2.65	s*		2.43	s**	
6	4.11	d	2.3	4.09	d	2.4	4.12	n.d.	n.d.
7	3.51	dddd	3 x ≈ 2, 8	3.50	dddd	3 x ≈ 2, 8.2	3.51	n.d.	n.d.
8	4.97	ddd	2, 6, 8	4.96	ddd	1.4, 6.6, 8.2	4.96	br dd	≈ 8
9α	1.78	ddd	1.6, 10.7, 15.7	1.76	dd	1.7, 10.6, 15.6	n.d.	n.d.	n.d.
9β	2.23	ddd	2.3, 6.4, 15.6	2.20	dd	2.1, 6.5, 15.5	2.22	n.d.	n.d.
10	1.97	m		1.90	m		2.08	n.d.	n.d.
13a	6.26	d	2.5	6.25	d	2.5	6.25	br s	
13b	5.94	d	2.3	5.94	d	2.4	5.93	br s	
14	1.08	d (3H)	6.4	1.08	d	6.5	1.09	d	6.9
15	0.80	s		0.74	s		n.d.	n.d.	n.d.

[1]according to [31] [2]previously unpublished. Assignments based on $^1H/^1H$-COSY spectra. *H-3(β) appears as singlet due to reaction in D_2O, so that position 3(α)is occupied by a deuteron and $^3J_{2(\alpha),3(\beta)} < 1$ Hz, compare [31]. ** assignment tentative, not based on COSY signal; shift value corresponds to a singlet observed in ^1H-difference spectrum (spectrum of reaction mixture – spectrum of pure papain).

phosphorylation, acetyl-CoA-synthetase, fatty acid synthetase and many others (review see [3]). Such effects are likely to be related to the recent observation that STLs disrupt mitochondrial membrane potential which is important in connection with their pro-apoptotic activity [43-45].

This further example for the preferential alkylation of a low-molecular weight thiol via the cyclopentenone structure of **HEL** indicates that the reaction will generally proceed in this way unless other stabilizing factors such as observed with free cysteine will enhance the susceptibility of the methylene lactone group.

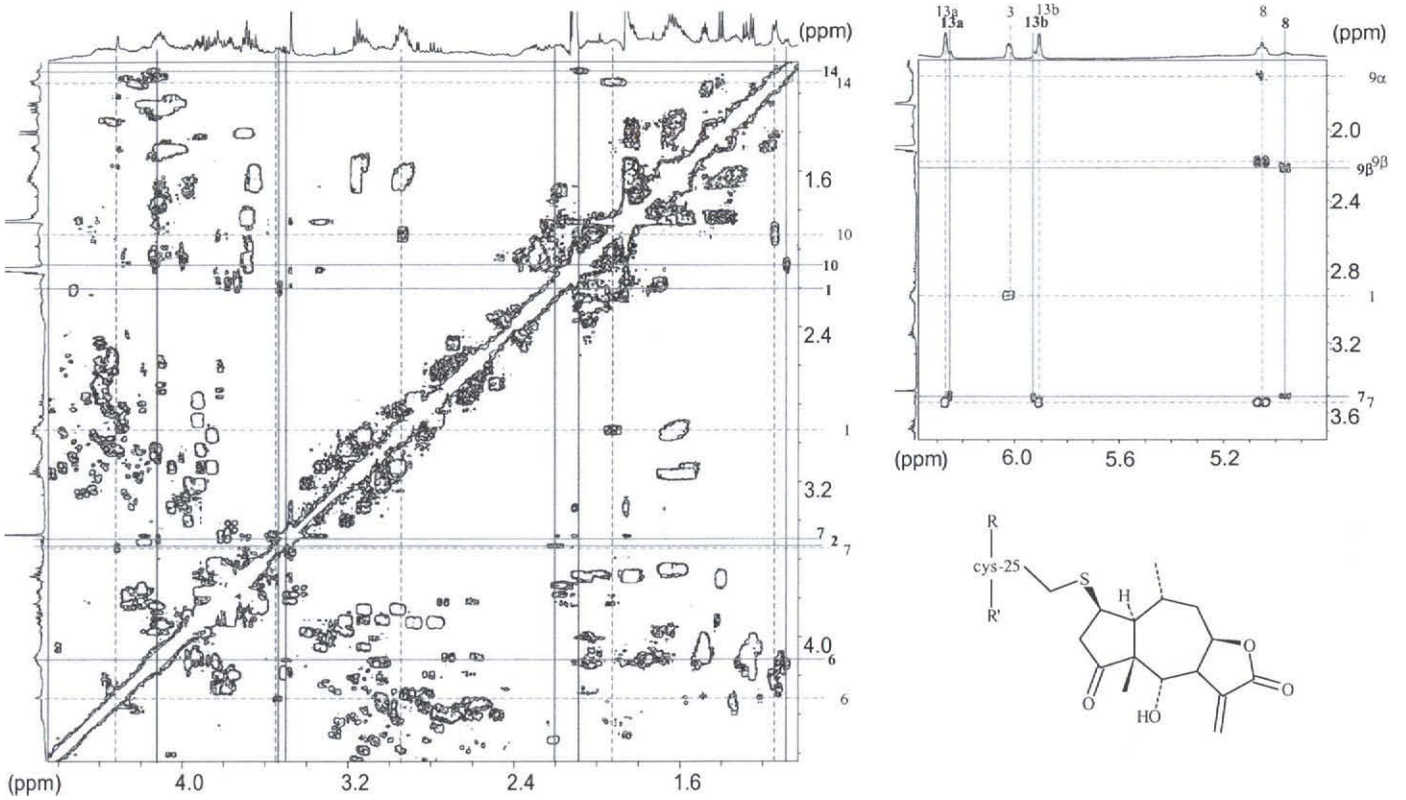


Fig. (8). $^{1}H/^{1}H$-COSY spectrum of **HEL** after reaction with papain (PAP). Dashed lines: signals of unreacted **HEL**; continuous lines: signals of **HEL-PAP**

Further evidence in support of this hypothesis could recently be obtained by investigations on the reaction of **HEL** with a full-grown protein, the cysteine protease papain (PAP). It has been reported in the literature that helenalin inhibits cysteine proteases of the cathepsin type [46]. Papain was chosen as a model protein due to its availability in high amounts at low cost. PAP possesses only one free cysteine residue (cys-25) in its catalytically active centre and its known 3D structure [47] (PDB code 1PPN) indicated that there should be enough space in the vicinity of the free SH of cys-25 to accommodate a **HEL** molecule.

The reaction was performed with 5 mM **HEL** and approximately equal concentration of PAP in D_2O and proton spectra recorded in daily intervals. Although the reaction proceeded quite sluggishly, after several days, diagnostic changes in the spectral signals were observed that were in agreement with reaction at C-2 of the cyclopentenone moiety. Fig. (**8**) shows the COSY spectrum obtained on day 18 after start of the reaction. Although only about 10 % of **HEL** have reacted, all cross signals attributable to the STL moiety of the newly formed **HEL**-adduct could be localized along with those of unreacted **HEL**, in comparison with a COSY spectrum of pure PAP obtained under the same conditions (not shown). Since the shift values of the new signals showed great similarity with those observed with GSH and CoASH (see Table 2), it became clear that the reaction had occurred at C-2 of the cyclopentenone group, i.e. that the it displayed the same regioselectivity as observed with the low-molecular weight thiols; Moreover, the signal of H-3 once more appeared as a singlet so that the nucleophile should be attached to C-2 in the β-postion as in case of the former adducts [31].

Since PAP possesses only one free cysteine residue, it is most probable that the reaction occurred with the free SH of this amino acid. It must not remain unmentioned, however, that assignment of the cysteine protons was not unambiguous due to signal crowding in the region of the spectrum where the α-proton of cys appears.

In agreement with the very low rate of reaction, **HEL** was unable to inhibit the proteolytic activity of PAP in a standard photometric assay. However, it is an important finding that the reaction can be observed and that the characteristics with respect to regio- and stereoselectivity are the same as observed with low molecular weight thiols. To our knowledge, this is the first instance in which reaction of an STL with the SH group of a protein was directly observed.

A full account on the mentioned results for the reaction of **HEL** with CoASH and papain is currently in preparation. Further investigations on the reactivity and involvement of this reaction in the biological activity of alkylant STLs and of synthetic analogues are in progress in our group.

Cytotoxicity

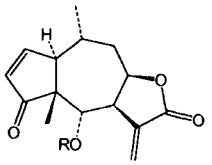
The eminent importance of alkylant potency/reactivity also manifests itself in the - still relatively few- systematic investigations on QSAR of such STLs.

The first study attempting a quantitative description of structural/physicochemical features on cytotoxic activity by Kupchan and coworkers [20] showed that a significant correlation of activity (cytotoxicity against KB cells) with lipophilicity (log P o/w) was only observable when subsets containing exclusively mono- or bifunctional alkylants were treated separately which indicated that alkylant potency was of higher importance for activity than other physicochemical properties. Interestingly, no correlation was

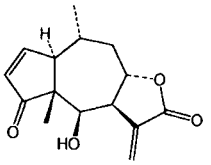
found in this study between reactivity of the STLs with free cysteine and their cytotoxicity, which could later on be explained by our finding that the reaction with the free amino acid represents a special case and is probably not representative for the behaviour towards physiologically important thiols (see above).

The first QSAR study taking into account quantitatively the effect of different alkylant structure elements on STL cytotoxicity was based on our data obtained with 20 helenanolide type STLs versus a murine Ehrlich ascites tumour cell line [48]. Two simple indicator variables encoding the presence/absence of cyclopentenone and methylene lactone structures (CP and MGL, 1 in case of presence, 0 in case of absence) showed a high degree of correlation with cytotoxicity, yielding a correlation coefficient $r^2 = 0.89$. An even stronger correlation was found for the water-accessible surface area of the reactive β-carbon atoms which in combination with two further descriptors, namely, the number of H-bond acceptors and the distance between C-12 and C-15 as a conformational descriptor, led to statistically highly significant correlations with activity [49]. The four mentioned descriptors yielded a linear regression model with correlation coefficients of $r^2 = 0.94$ and q^2 (after leave-one-out cross validation, LOO) = 0.90. Consistently, re-anaylsis of Kupchan's data set [20] showed that among these compounds -which all possess a methylene lactone group-, additional presence of a cyclopentenone was of high importance for activity and the number of H-bond acceptors was of similar modulatory importance as in our helenanolide data set [3].

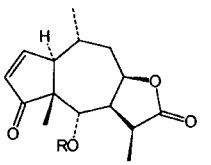
In a more recent study, we could extend these findings to a larger and structurally more diverse set of STLs and their cytotoxicity against KB cervix carcinoma cells [50]. It was found once more that the simple indicator variables CP and MGL (presence/absence (1/0) of cyclopentenone and/or methylene lactone) were sufficient to account for 70% of the variablity within the biological data. Using a set of 14 descriptors termed partial charge-based *fr*actional *a*ccessible *s*urface *a*reas (Q_frASAs) for 37 of our STLs and principal component regression [11] as well as genetic algorithm-driven variable selection combined with PLS [51] statistics (GA-PLS [52]), we arrived at QSAR models of very good statistical quality. The Q_frASA descriptors encode fractions of the water-accessible surface area attributable to atoms within 14 different partial charge intervals and were devised by us in order to include the surface charge distribution over the entire molecule in the structure-activity correlation. The Q_frASA variables contributing most significantly to the cytotoxicity QSAR models were those corresponding to atoms of the reactive alkylant partial structures of each molecule. Thus, in a similar but more sophisticated way as in our earlier QSAR [49], the alkylant structure elements were identified as major determinants of cytotoxicity. Most notably, Kupchan's data for cytotoxicity aganst the KB cell line [20] when re-analysed using the same method led to a virtually identical model so that we found it justified to combine both data sets and construct a QSAR for a total of 62 different and structurally very diverse STLs (Fig. (**9**)). The final QSAR thus obtained (Fig. (**10**) and (**11**)) showed high internal predictivity and explanatory power and at the same time also showed good capacity to predict the activity of untested STLs. It is possible to visualize the influence of particular Q_frASA descriptors on activity by mapping the PLS coefficient of each respective atom on the molecular surface. Some examples are depicted in Fig (**11**). It must be noted that Q_frASA descriptors are a global representation of the molecules'



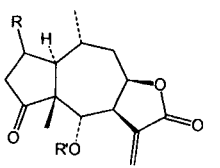
#	R	pIC_{50}
1	H	6.19
2	ac	5.99
3	*i*-but	6.08
4	*i*-val	5.61



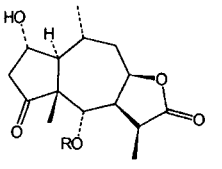
#	pIC_{50}
5	5.90



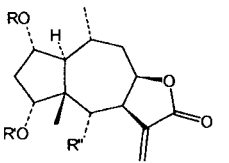
#	R	pIC_{50}
6	H	5.12
7	ac	5.20
8	i-val	5.49



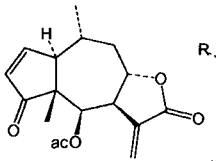
#	R	R‘	pIC_{50}
9	αOH	tig	5.25
10	αOH	ang	5.21
11	αOH	sen	4.91
12	βOet	H	4.88
13	βO*i*-pr	H	4.98



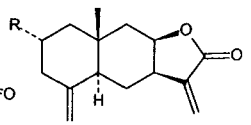
#	R	R‘	pIC_{50}
14	H	tig	3.95
15	H	sen	3.96



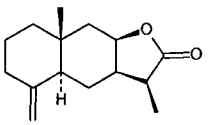
#	R	R‘	R“	pIC_{50}
16	H	H	H	4.70
17	ac	H	H	5.06
18	ac	H	OH	5.64
19	ac	H	Oac	5.83
20	ac	ac	H	5.38
21	H	H	Otig	4.69
22	H	tig	OH	5.00



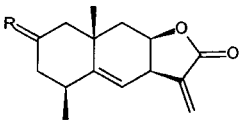
#	pIC_{50}
23	5.94



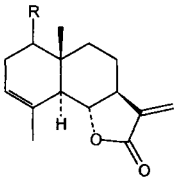
#	R	pIC_{50}
24	H	5.37
25	OH	5.75
26	Oac	5.42



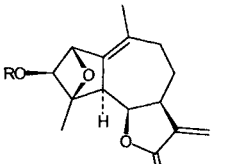
#	pIC_{50}
27	< 3



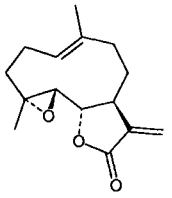
#	R	pIC_{50}
28	H_2	5.45
29	O	5.42



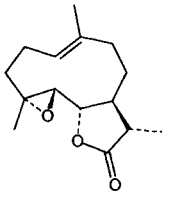
#	R	pIC_{50}
30	αOH	5.21
31	βOH	5.09



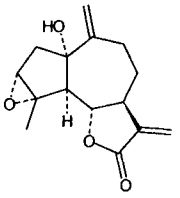
#	R	O	pIC_{50}
38	H		4.77
39	2-mebu		5.56



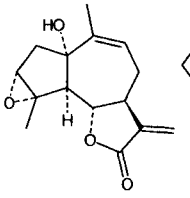
#	pIC_{50}
32	5.62



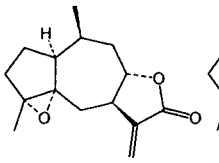
#	pIC_{50}
33	< 3



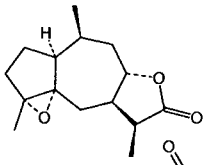
#	pIC_{50}
34	5.12



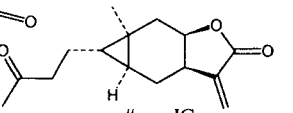
#	pIC_{50}
35	5.19



#	pIC_{50}
36	4.97

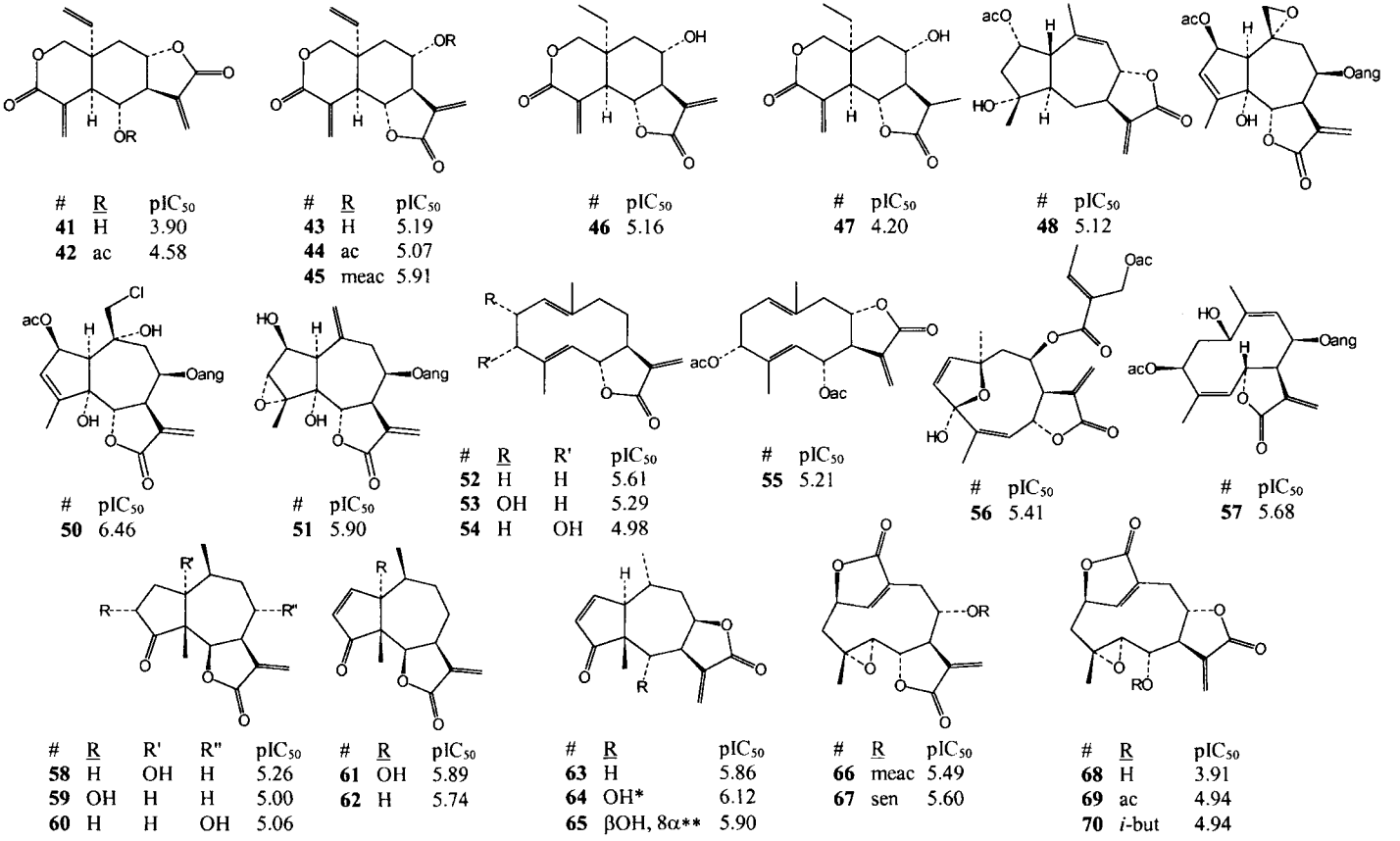


#	pIC_{50}
37	< 3



#	pIC_{50}
40	4.83

Fig. (9) (continued)



#	R	pIC_{50}
41	H	3.90
42	ac	4.58

#	R	pIC_{50}
43	H	5.19
44	ac	5.07
45	meac	5.91

#	pIC_{50}
46	5.16

#	pIC_{50}
47	4.20

#	pIC_{50}
48	5.12

#	pIC_{50}
50	6.46

#	pIC_{50}
51	5.90

#	R	R'	pIC_{50}
52	H	H	5.61
53	OH	H	5.29
54	H	OH	4.98

#	pIC_{50}
55	5.21

#	pIC_{50}
56	5.41

#	pIC_{50}
57	5.68

#	R	R'	R"	pIC_{50}
58	H	OH	H	5.26
59	OH	H	H	5.00
60	H	H	OH	5.06

#	R	pIC_{50}
61	OH	5.89
62	H	5.74

#	R	pIC_{50}
63	H	5.86
64	OH*	6.12
65	βOH, 8α**	5.90

#	R	pIC_{50}
66	meac	5.49
67	sen	5.60

#	R	pIC_{50}
68	H	3.91
69	ac	4.94
70	*i*-but	4.94

Fig. (9). Structures of STLs investigated in QSAR study on cytotoxicity vs. KB carcinoma cells [50]. Biological data are expressed as pIC_{50} = -log IC_{50} (M); for compounds **1-40** they represent own measurements [50], those of **41-70** were taken from Literature [20]. Note that **64** and **65** are identical with **1** and **5**.
ac : acetate; nmeac: methacrylate; i-but: isobutyrate; i-val: isovalerate; tig: tiglate; ang: angelate; sen: senecionate; i-pr: isopropyl; et: ethyl.

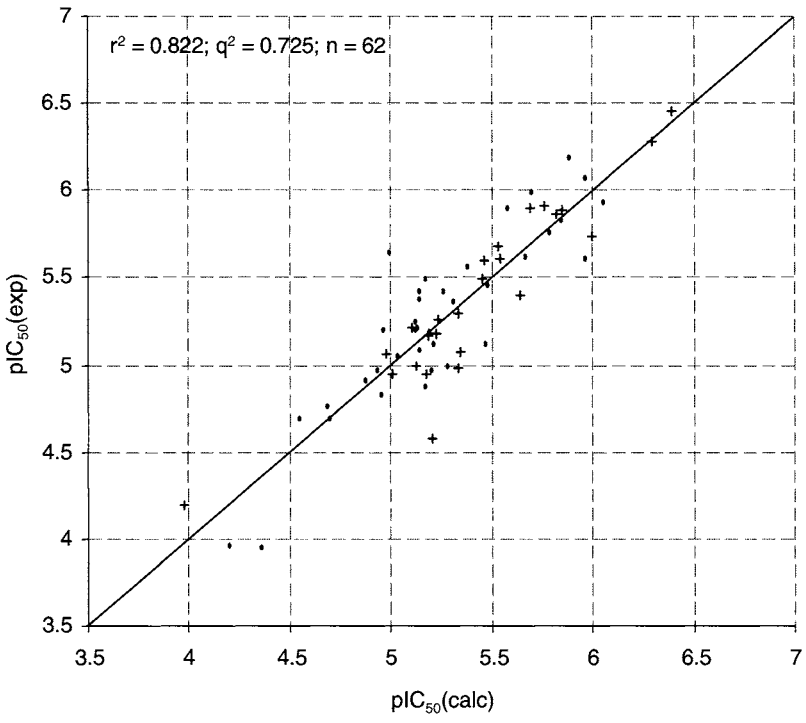


Fig. (10). Experimental vs. calculated pIC_{50} values from a QSAR model for the cytotoxic activity of 62 STLs derived exclusively from Q_frASA descriptors by GA-PLS. The GA-PLS algorithm identified 11 of the 14 descriptors as relevant, the number of latent variables constructed was 5. Data for compounds shown as dots represent own measurements (n=37), those shown as + were taken from literature [20] (n=25).

potential to interact with any kind of target and were thus found useful also for QSARs with respect to other types of compounds and biological activities [50]. Two further new applications to STLs will be presented below.

Recently discovered molecular targets

With respect to particular macromolecular targets, a vast number of reports from earlier years have shown the susceptibility of a plethora of enzymes and other functional proteins to inhibition by STLs. An overview of such proteins whose physiological functions are reportedly affected by STLs can be found in [3].

More recent findings on potential molecular targets include enzymes of leukotriene biosynthesis, namely, 5-lipoxygenase as well as leukotriene C_4-synthase [53]; both enzymes were inhibited in a dose- and time-dependent manner by **HEL** and **DHAC**, the former showing the stronger effect. It was found that the inhibitory activity was attenuated in a concentration-dependent manner by GSH. The corresponding C-2 GSH-adducts (**HEL-SG** and **DHAC-SG**) also showed some (although weak) activity on LTC_4 synthase, but this could be observed only if the trombocytes used in the assay were

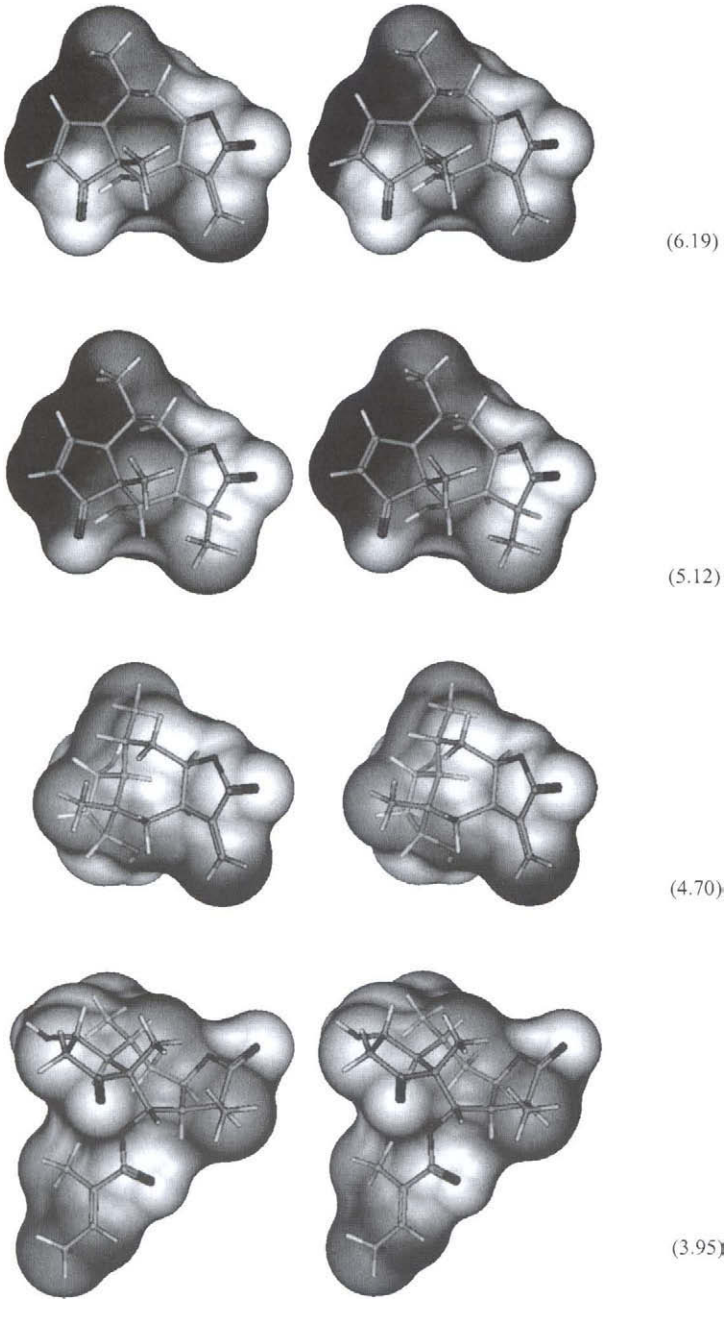


Fig. (11). Influence of Q_frASA descriptors mapped on the molecular surface (stereoscopic view) of four helenanolides (structures **1**, **6**, **16**, **14** in Fig. (**9**)); Only positive influence on KB cytotoxicity is shown. Darker shade corresponds to stronger enhancement of activity. According to [50], modified. Numbers in brackets: pIC_{50} values.

destroyed before activity was measured. The effect of the native STLs was strong enough to expect a contribution to their overall anti-inflammatory activity (IC_{50} of **HEL** on LTC_4 synthase after 60 min pre-incubation was 12 μM) [53].

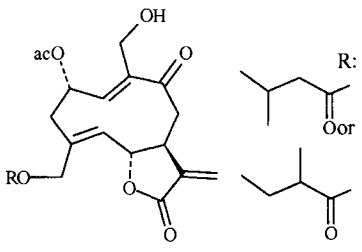
As an interesting new target with respect to a specific anti-tumoral activity, aromatase was found to be inhibited by several STLs [54]. Aromatase is the key enzyme in the biosynthesis of estrogens and inhibitors are used successfully to treat hormone-dependent breast cancer. Quite noteworthy, it was recently found that aromatase inhibition–unlike most other effects of STLs- is not dependent on the presence of potentially alkylant substructures. The inhibitory effect of a guaianolide (10-epi-8-deoxy derivative of structure **25** in Fig. (**12**)) was even increased when the exocyclic methylene double bond was hydrogenated to the corresponding 11β,13-dihydro derivative [54]. Although no experimental 3D structure of aromatase exists, the authors provided a very plausible explanation for their results based on a theoretical model, in which the STL was superposed with the structure of the natural substrate, testosterone, in the catalytic centre of the enzyme. The finding that a non-alkylant STL inhibits such an important target is especially noteworthy since the compound can be expected to possess drastically diminished non-specific toxicity in comparison with the native STL and may thus be a valuable lead towards new therapeutic drugs.

Nuclear transcription factors

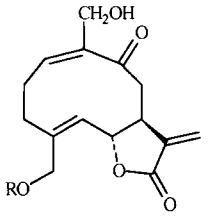
As a very important group of macromolecular targets, factors of nuclear transcription, namely, NF-κB and NFAT, have been found susceptible to inhibition by STLs in a number of studies in recent years. Results from two independent research groups have shown that the function of nuclear transcription factor NF-κB is inhibited by low concentrations of a variety of STLs [55-60]. This factor is responsible for the activation of gene transcription in response to a number of exogenous signals (for reviews see [61-64]), and the implications of inhibition by natural products on inflammatory processes in relation to their use as antiinflammatory agents have been discussed in detail [62].

The term NF-κB is applied to a family of highly homologous proteins occuring ubiquitously in homo- and heterodimeric form in mammalian cells. The most abundant heterodimer consisting of two subunits of M_r 50 and 65 kDa (p50 and p65) is known to function as a genetic switch for many different genes of the immediate response type which encode important cytokines such as interleukins 1, 2, 6 and 8 as well as inducible enzymes such as cyclooxygenase II (COX II) and inducible nitric oxide synthase (iNOS). Certain exogenous stimuli such as tumor necrosis factor α (TNF-α), phorbol esters, bacterial lipopolysaccharides and others elicit a signal cascade that results in activation of NF-κB. Inactivation of NF-κB by STL, originally proposed to be a consequence of impaired degradation of its endogenous inhibitor IκB [55, 56], has more recently been shown to result from direct action on the p65 subunit [57, 59, 60].

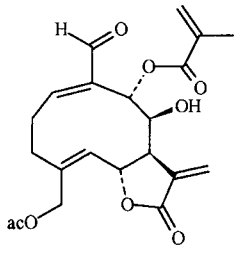
We conducted a study on SAR with respect to the inhibitory activity of 28 STLs (Fig. (**12**)) on NF-κB DNA binding, in which it was shown that the level of activity is roughly correlated with the number of potential alkylant centres [59]. Based on the known X-ray structure of the p65 subunit [65] we proposed a mechanism of action which involves alkylation of two cysteine residues, C38 and C120, which are located in the L1 and E' regions in the vicinity of the DNA-binding domain. The distance between



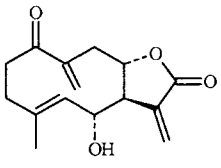
1 (12.5)



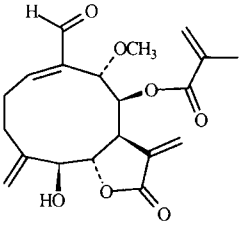
2 (12.5)



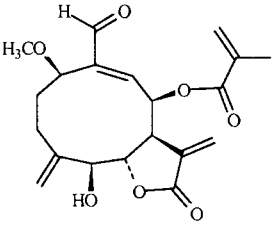
3 (5)



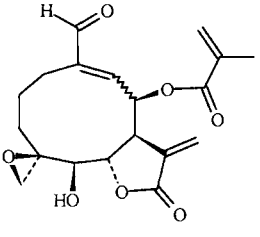
4 (50)



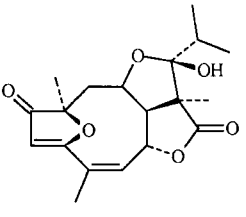
5 (50)



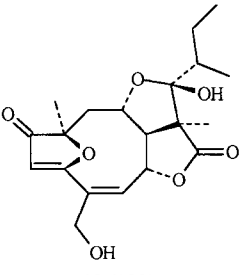
6 (20)



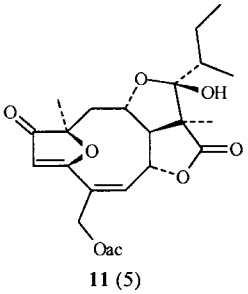
7 =**E** (5) **8** = **Z** (10)



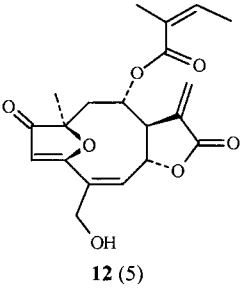
9 (>300)



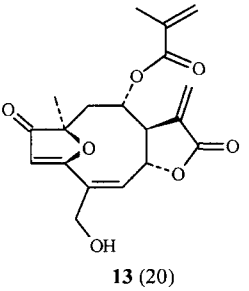
10 (100)



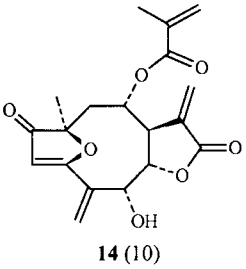
11 (5)



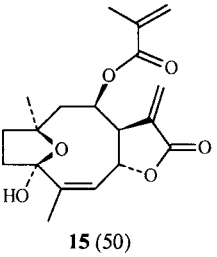
12 (5)



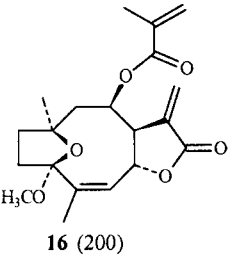
13 (20)



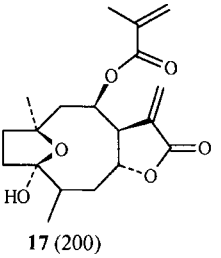
14 (10)



15 (50)



16 (200)



17 (200)

Fig. (12) (continued)

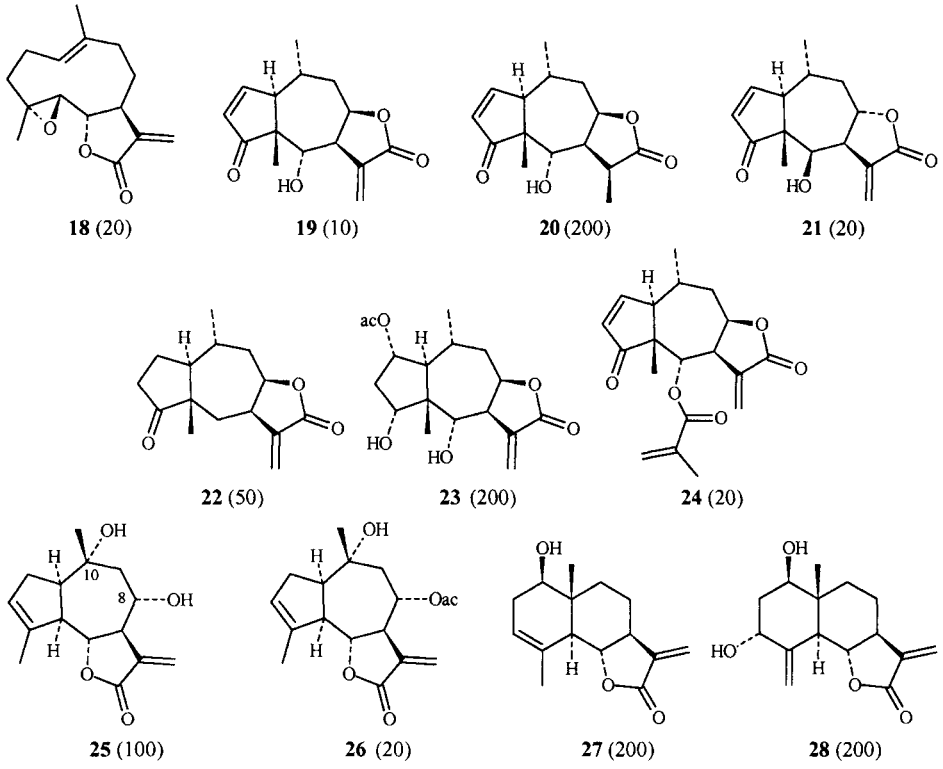


Fig. (12). Structures of STLs in SAR study on NF-κB inhibition [59]. Numbers in brackets: concentration (μM) required for 100% inhibition of NF-κB DNA binding.

the two free SH groups is ca 8 Å, which makes simultaneous (or subsequent) alkylation by a single bifunctional STL molecule possible. Between the two cysteines, a pocket or cleft is observed in the protein structure which is large enough to accommodate an STL molecule. A tyrosine (T36) whose side chain extends into this gap is known to be directly involved in DNA binding [65]. The respective part of the X-ray structure and the model showing cross-linkage of the two cys-residues by a molecule of helenalin is shown in Fig. (**13**). We could show by molecular modelling that all potentially bifunctional alkylants in the studied series fit into this cleft and can reach the two cysteines with their reactive partial structures. In the energy minimised structures of the bifunctional STL-protein adducts, several amino acid residues which reportedly contribute to DNA binding show displacement (which is most prominent in case of the mentioned T36), and we concluded that the factor, altered in this way, should no longer be able to correctly recognize and bind to the κB-DNA motif [65]. The model was capable of explaining in a qualitative manner the observed levels of activity for bifunctional and monofunctional alkylant STLs. However, the behaviour of some bifunctional compounds with relatively low activity and of some monofunctionals with unexpectedly high activity remained unclear [59]. Further studies were therefore conducted to find factors influencing activity apart from number of reactive structure

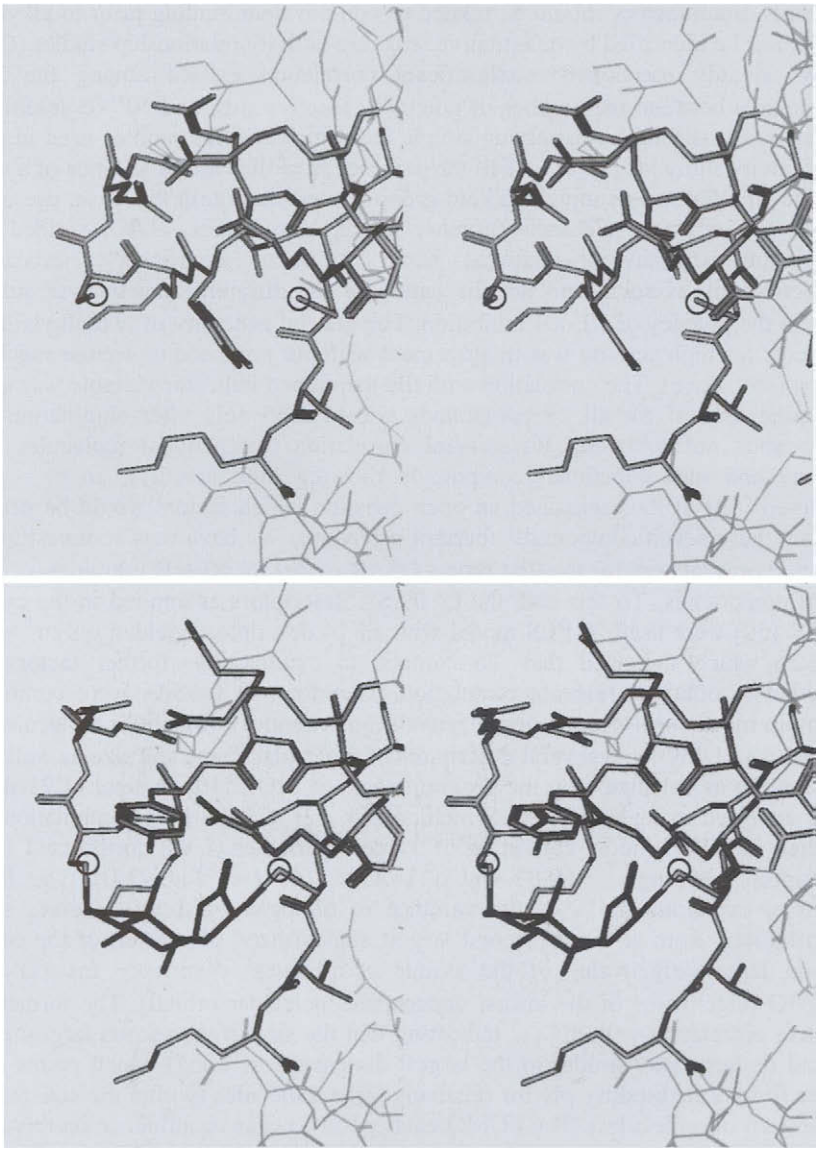


Fig. (13). Stereoscopic views of molecular models showing the proposed cross linkage of L1-C 38 and E'-C120 of NF-κB-p65 by helenalin (**HEL**), according to [59].
Top: Part of the DNA binding domain from the X-ray structure of NF-κB-p65 homodimer subunit A [65] (PDB-code: 1RAM).
Bottom: Energy minimized structure after cross linkage of cysteine residues by **HEL**.
Line representation: Protein portion fixed during energy minimization.
Stick representation: Protein portion submitted to energy minimization.
Circles: S-atoms of C38 and C120.

elements. Such factors should be related to non-covalent binding prior to alkylation and might best be identified by quantitative structure-activity relationship studies (QSAR).

As already mentioned, a significant correlation existed among the 28 tested compounds between the number of potetially reactive sites and NF-κB inhibition. This correlation was found by applying simple indicator variables such as used in our initial cytotoxicity study [49] (1 in case of the presence, 0 in the case of absence of a methylene lactone structure or a conjugated keto group); Interestingly, in this case, use of the sum of alkylant structure elements (number of alkylant centers, NAC) yielded an equal correlation with even somewhat better statistical significance, indicating that bifunctionality as such, and not the nature of the different centers, was sufficient to explain the potency of NF-κB inhibition. The general necessity of two alkylant structure elements for high activity was in agreement with our proposed molecular mechanism of action (see above). The correlation with the mentioned indicator variable was significant but quite low (r^2 for all 28 compounds = 0.44) and only after elimination of seven compounds not following the general correlation (bifunctional molecules with low activity and monofunctional compounds showing high activity), an r^2 = 0.86 was observed [59]. It thus remained an open question which factors would be necessary to explain these seven compounds' aberrant behaviour. We have now re-investigated these data and present here for the first time a QSAR model for NF-κB inhibition of the whole set of compounds. To this end, the Q_frASA descriptors as applied in the cytotoxicity QSAR [50] were used. A PLS model with all 14 descriptors yielded only r^2 = 0.51 and q^2=0.26 which indicated that –in contrast to cytotoxicity- further factors must be included to obtain a relevant correlation. Therefore, Q_frASAs were combined with quantum mechanical descriptors (eigenvalues (=energies) of frontier molecular orbitals: εHOMO, εLUMO) and several descriptors of molecular shape and size as well as charge distribution as calculated by the program package MOE [10]. A total of 21 descriptors were analysed using the GA-PLS method [50, 52] and the best combination obtained resulted in a PLS model consisting of 3 latent variables (LVs) constructed from eight descriptors yielding r^2 = 0.83 and q^2(LOO)= 0.67 (see Fig. (**14**)). The first latent variable, explaining 54% of the variance in biological activity, receives significant contributions from st_dim2 (second largest standardized dimension of the molecule = second largest eigenvalue of the atomic coordinates' covariance matrix) and from εLUMO (eigenvalue of the lowest unoccupied molecular orbital). The former shows a positive correlation with pIC_{100}, indicating that the size in the second largest dimension should be large (i.e. similar to the largest dimension, st_dim1) which points towards a better fit into the binding site for relatively large molecules (within the size range of the present set of molecules). The LUMO energy has a negative influence on this LV which is in agreement with the proposed mechanism of action, since a LUMO of lower energy will yield a more favourable interaction with the HOMO of a sulfhydryl group and thus will facilitate the formation of a covalent bond.

Moreover, the ratio of total accessible surface area due to atoms with positive partial charge over such with negative partial charge (rASA = ASA+/ASA-) shows a negative influence in LV 1. The absolute values of rASA for the 28 molecules range between ca 1.1 and 1.9. Hence values close to the lower boundary of this range have a positive influence on activity, i.e. a more or less equal size of positive and negative areas is

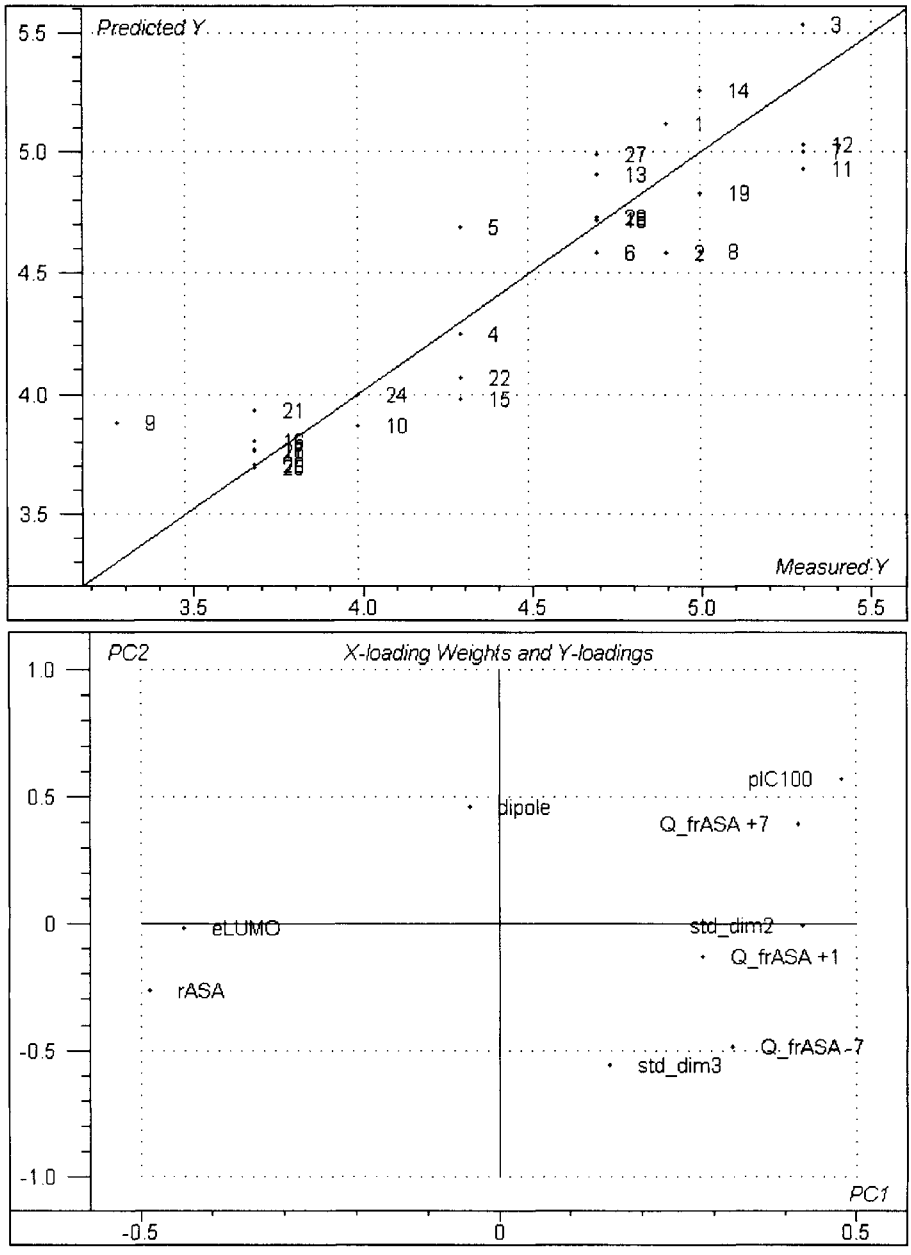


Fig. (14). A: Experimental vs. calculated pIC_{100} values from a QSAR model for inhibitory effect of 28 STLs on NF-κB activation (biological data from [59], structures see Fig. (**12**)). The model was generated by GA-PLS analysis (number of latent variables: 3) from the 8 descriptors shown in the loading weights plot (**B**). This plot illustrates the impact of each descriptor on the first two latent variables (PC1 and PC2) explaining 54 % and 27 % of the variance in the Y data (pIC_{100}), respectively.

desirable for active molecules. Three of the Q_frASA descriptors were selected by the GA-PLS method, all of which show high positive loadings in LV1. Q_frASA +7 (accessible surface of atoms with partial charge (Q) > 0.3 e) is related to carbonyl carbons of ester and lactone groups and Q_frASA -7 (due to atoms with Q < -0.3 e) represents ester carbonyl oxygens and oxygen atoms of free hydroxy groups. The fact that both of these descriptors show about the same loading weight in LV1 indicates that the number of ester/lactone carbonyl groups is important, probably because of their H-bond accepting capacity. Q_frASA +1 (attributable to atoms with Q between 0 and +0.05 e) is due to oxygen-substituted sp3 carbon atoms and points into the same direction. Quite interestingly, Q_frASA -7 shows a strong negative effect on the second latent variable (LV2 explaining further 27 % of the variance), while Q_frASA +7 is of positive impact also in this respect. This finding can be explained by the fact that Q_frASA -7 along with its relation to ester carbonyl oxygens contains information on the number of free OH groups. When Q_frASA -7 was replaced by two desciptors encoding the number of ester groups (nEst) and the number of free OH groups (nOH), the resulting PLS model was of identical statistical quality ($r^2 = 0.83$ $q^2 = 0.67$) and nEst showed strong positive impact on latent variable 1 while nOH did not contribute to this variable but showed a negative loading in LV2. A deleterious effect of free OH on activity might indicate a negative influence of hydrophilicity or solvation. Although e.g. log P was found not to contribute to our previous SAR (see above [59]), it is possible that the influence of this relatively weak effect was not significant in this coarse analysis.

Major contributions to the second latent variable arise from the molecular dipole moment (dipole) and from st_dim3, i.e. the smallest standardized dimension. The dipole moment shows a positive loading in this latent variable pointing towards the importance of charge separation and thus of electrostatic interactions. The finding that st_dim3 should be rather small while at the same time st_dim2 should be large (see above) points towards a preference of the binding site for molecules showing a more or less flattened shape. This is in agreement with our qualitative SAR model [59] which indicates that (a) in order to cross-link the two important cys residues of NF-kB-p65, an STL has to be large enough to span the distance of approx. 8 A, and (b) in order to effectively hinder DNA complexation, it would be advantageous that the STL molecule after binding extends into the open space which would normally be occupied by atoms of the DNA. These two requirements reflect the positive contributions of the first two standardized dimensions. Finally, the molecule should not be too large in the dimension perpendicular to those important for (a) and (b), in order to fit into the gap between C38 and C120, which is in agreement with the necessity for a small third standardized dimension.

It is noteworthy, that also compounds which did not follow the simple correlation with the number of alkylant centers, i.e. monofunctional alkylants with high activity as well as bifunctional ones with relatively low activity are quite well represented by this more sophisticated model, indicating that factors of shape (i.e. fit into the binding site), as well as charge distribution (i.e. electrostatic interaction) and H-bond acceptance modulate the impact of alkylant potency.

It must be noted, that the pIC_{100} values used here are relatively coarse estimates of the biological activity under study and that the investigated data set -although quite diverse in structural terms- is quite small. Although the model discussed above allows quite interesting insights into the structural requirements for NF-κB inhibition, more

detailed QSAR studies are desirable. Such further investigations currently await the development of assay methods that allow for measurement of high-quality IC_{50} data for an even wider variety of compounds which will be available in the foreseeable future [66].

The proposed mechanism of NF-κB inhibition [59], based entirely on theoretical considerations, was more recently corroborated by experimental data obtained in a mutation study. As the main result, a single-point mutant of p65 in which C38 was replaced by a serine (C38S) was completely insensitive to several bi- and monofunctional STLs so that the crucial role of direct interaction with the wild type protein (wt p65) at this position was unambiguously proven [60]. Although C120 was unmodified in this mutant, apparently the presence of this residue alone is not sufficient to render the protein susceptible to inactivation by STLs. This result is readily explained by the fact that C38 is the much more exposed of the two cysteines (compare Fig. (**13**), Top). It can be assumed that the second alkylation at C120 by a bifunctional alkylant will occur much more easily if the STL has already been bound to C38 and thereby fixed in the vicinity of C120. Thus, if C38 is replaced by a residue that does not react with STLs, the concentrations required to reach the more buried C120 must be drastically increased. Although the role of C120 could not be completely explained (some contradictory results were obtained with a C120→A and a C38,120→S mutant), the results provided unambiguous proof for the basic assumption of our theoretical model, i.e. a direct interaction of STLs with C38, and do not contradict the hypothesis that alkylation at this position is followed by reaction with C120.

In addition to its importance for inflammatory responses, NF-κB also plays a role in the regulation of cell homeostasis and apoptosis [62-64] and it was thus straightforward to assume a link between the known cytotoxic activity of STLs and their NF-κB inhibitory activity [3]. It was demonstrated that helenalin indeed induces apoptosis and that it does so via a CD95 death receptor independent pathway and requires the activation of caspases [43]. Further STLs for which quite detailed studies on pro-apoptotic activity and mechanism exist are the germacronolides parthenolide and costunolide (structures **32** and **52** in Fig. (**9**)) [44, 45].

In a more recent study, the effects of **HEL** and two related helenanolides (dihydrohelenalin acetate **DHAC**, and chamissonolide **CHA**) on gene transcription profiles and pro-apoptotic activity were studied [37]. As in so many other studies, the bifunctional alkylant **HEL** showed the strongest effects in all investigated assays, while **DHAC** and **CHA** as monofunctional alkylants were generally less active. A very interesting dependence of the observed effects on the different reactive structure elements present in these molecules was found. Apart from significant down-regulation of the mRNA levels for several pro-inflammatory cytokines (IL-2, IL-6, IFN-γ, TNF-α, iNOS, GM-CSF) observed with all three STLs, **HEL** and **DHAC** (both with a cyclopentenone structure) but not **CHA** (methylene lactone structure), also led to a pronounced down-regulation of house keeping genes (GAP-DH, β-actin), transcription factors (NF-ATc, I-κBα) as well as cell homeostasis factors (bcl-2, cyclin-D1). Furthermore, differences in the time-dependence of cytotoxic effects were observed, the cyclopentenones leading to a more rapid onset of cytotoxicity than **CHA**, which only showed some cytotoxicity at high concentrations and long incubation time (72 h). Most

strikingly, **HEL** and **DHAC** induced caspases-3 and -8 and led the cells into apoptosis as shown by flow cytometric measurements, while **CHA** did not induce apoptosis under identical conditions, but rather showed an anti-apoptotic activity. Thus, cytotoxic or pro-apoptotic activity correlates with the observed down regulation of house-keeping and cell homeostasis factors which depends on the presence of a cyclopentenone moiety. The down regulation of pro-inflammatory genes, on the other hand, appears to be independent of the cyclopentenone group since also **CHA** shows these effects. It is noteworthy, that the observed effects on gene transcription did not correlate well with the capacity of the compounds to inhibit NF-κB p65 DNA binding in the same cell system. Only **HEL** displayed a significant effect under the chosen conditions (complete inhibition at <20 μM) while in case of **DHAC** only a weak effect, in case of **CHA** no inhibition at all was observed at 20 μM. Thus, direct NF-κB inhibition is apparently not the only effect to explain the anti-inflammatory activity of these compounds, which is in line with reports that helenanolide type STLs also inhibit DNA binding of NFATc [67].

It was concluded from the presented results that anti-inflammatory and cytotoxic/pro-apoptotic effects of these STLs are based on different mechanisms of action mediated by the distinct reactive partial structures [37]. This conclusion being based on only three compounds so far, it would be highly desirable to conduct analogous experiments with a wider variety of STLs in order to find out whether the findings of this study indicate a general trend. At present, it appears straightforward to assume a link between these differential activities of cyclopentenone- and methylene lactone-containing STLs with the conspicuous differences in their reactivity towards thiol compounds mentioned earlier. Search for direct links between reactivity and biological activity thus continues to be one of the major topics of research in our group.

Inhibition of 5HT release from blood platelets

Although research at the transcription level has dominated pharmacological research on STLs in the last fes years, it should not be forgotten that evidence from earlier studies has indicated involvement of many downstream effects in their anti-inflammatory activity. One of the most prominent effects in this respect is related to anti-secretive activity due to membrane-stabilization. It was shown by Hall and coworkers that STLs inhibit the activity and release of lysosomal enzymes [46]; furthermore, some STLs were demonstrated to inhibit mast-cell degranulation [68, 69] and, as found more recently, the release of human neutrophil elastase [70].

Perhaps the most prominent example with respect to anti-secretive effects of STLs is related to the well-known anti migraine activity of feverfew (*Tanacetum parthenium*, Asteraceae). The main STL from this plant is parthenolide (**PAR**), structure **1** in Fig. (**15**), and it has been shown that **PAR** is a potent inhibitor of serotonin (5HT)-release from thrombocytes. Although it is not finally clarified to which extent this effect is involved along with the numerous other reported activities of **PAR** in the clinically proven efficacy of feverfew [71, 72], it should at least contribute to the overall effect.

A structure-activity study on 54 STLs including **PAR** was published by Marles et al. who investigated the influence of these compounds 5HT release from bovine thrombocytes [73]. As the most prominent results, compounds lacking an α,β-unsaturated carbonyl moiety were found completely inactive. In addition to various observations of more qualitative nature, the authors also attempted QSAR analyses with

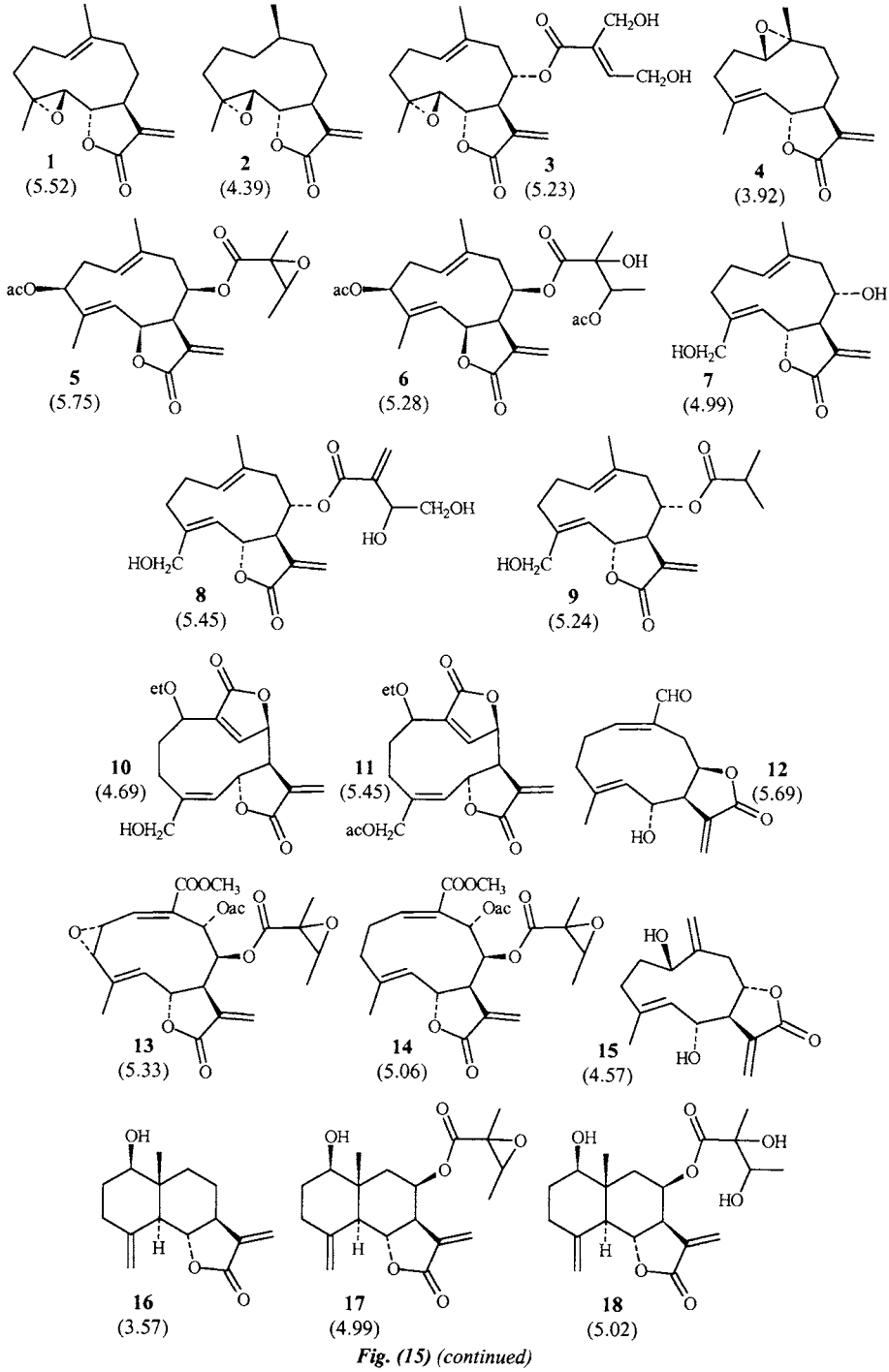


Fig. (15) (continued)

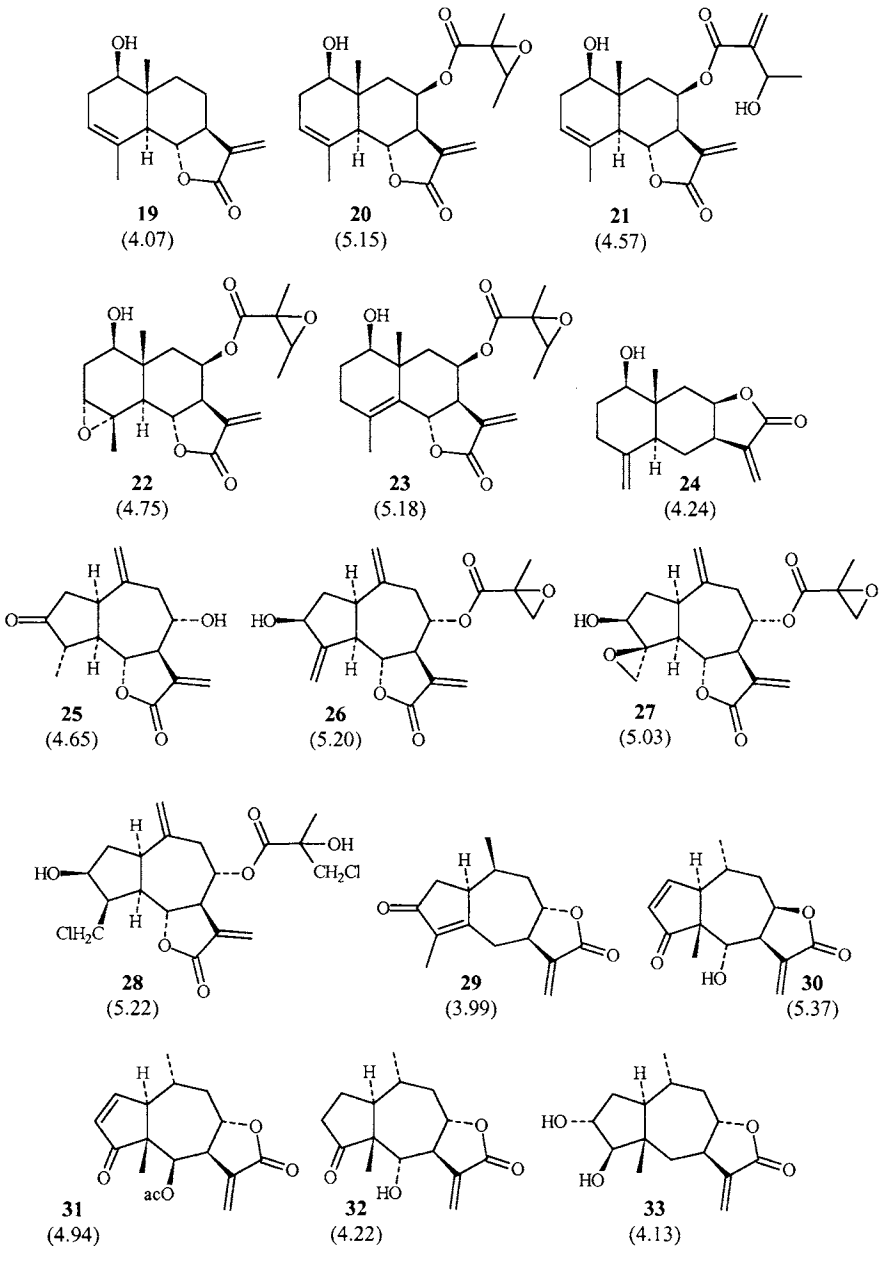


Fig. (15) (continued)

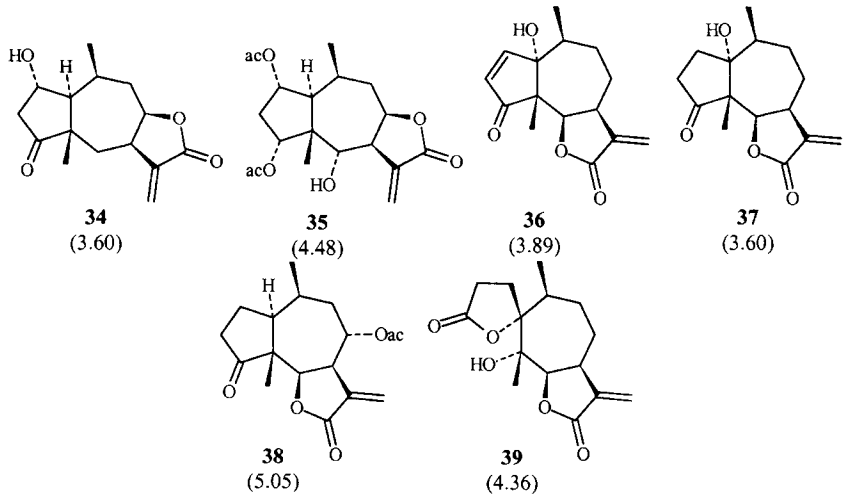


Fig. (15). Structures of STLs in SAR study on 5HT release inhibition, data from [73]. Numbers in brackets: pIC_{50} values.

this set of molecules. No correlation was found between pIC_{50} and calculated log P (CLOGP) values. A statistically significant correlation of activity with calculated molar refractivity (CMR) was observed for a subset of 8 eudesmanolides while this did not apply to the total data set consisting besides these compounds of germacranolides, guaianolides and pseudoguaianolides. It appears somewhat questionable how a putative receptor should distinguish one skeletal class of STLs from the others and then discriminate among these compounds by global steric bulk. The relevance of this correlation was already questioned by the authors [73]. Finally, an attempt was made to derive a CoMFA model for the subset of germacranolides (n=18, 16 with measurable IC_{50} value). Unfortunately, however, neither statistical details nor any information on the protocol used or on the final model are reported so that it is impossible to assess the validity of the model or of the conclusions drawn from it [73].

As part of our studies into the applicability of Q_frASA descriptors (see above) we have now re-investigated a subset of 39 active compounds (i.e. such compounds with reported measurable IC_{50} values (n=42) with the exception of three compounds not possessing an α-methylene-γ-lactone ring; see Fig (**15**)). In the same way as reported above for the NF-κB data, it was found that the Q_frASA descriptors alone did not lead to very satisfactory correlations. However, when quantum mechanical descriptors as well as such encoding molecular shape were included in the analysis (using the same descriptors as for the NF-κB data, see above), it was possible to derive PLS correlations of some statistical quality. The best model obtained by GA-PLS consisted of 5 Q_frASA descriptors along with stdim_1 and εLUMO which were combined in 3 latent variables and yielded $r^2 = 0.67$ and $q^2 = 0.52$. It was found that two compounds, **HEL** and **PAR**, did not follow the general trend of the model. When these two outliers were omitted, the r^2 and q^2 values were dramatically improved to 0.85 and 0.78, respectively (see Fig. (**16**) A). In the loading weights plot (Fig. (**16**) B) it can be seen that the first latent variable (67 % variance explained) is once more determined by the εLUMO (negative

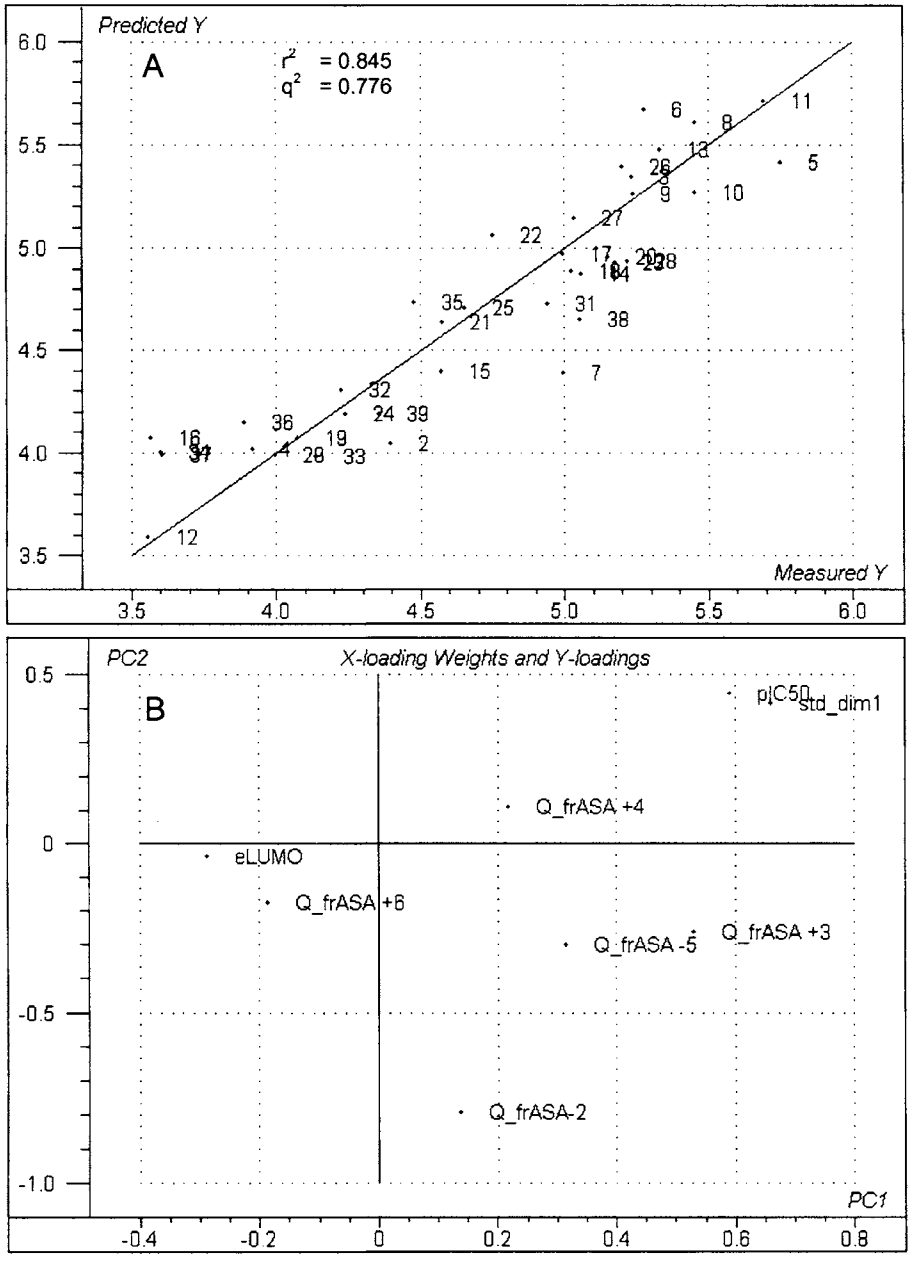


Fig. (16). **A**: Experimental vs. calculated pIC_{50} values from a QSAR model for inhibitory activity of 37 STLs on 5HT release from bovine platelets (biological data from [73]; structures see Fig (**15**)). The model was generated by GA-PLS analysis (number of latent variables: 3) from the 7 descriptors shown in the loading weights plot (**B**). This plot illustrates the impact of each X-variable (descriptor) on the first two latent variables (PC1 and PC2) explaining 67 % and 16 % of the variance in the Y data (pIC_{50}), respectively.

regression coefficient as above and thus in agreement with alkylation as basic mechanism of action) as well as a shape descriptor, st_dim1, i.e. the largest standardized dimension of the molecule. St_dim1 shows the most positive loadings in both, LV1 and LV2, so that a large maximum extension of a molecule corresponds with high activity. It must be noted that st_dim1 could not be replaced by any descriptor of global molecular size such as mass, volume etc., each of which yielded models of much inferior quality. Hence it became clear that the influence of st_dim1 must be related to its impact on molecular shape, i.e. st_dim1 must be large with respect to st_dim2 and st_dim3: an extended overall shape is important and correlates with high activity. The Q_frASA descriptors dominate the second and third latent variable and are thus of relatively minor impact on overall activity (16 and 2 % variance explained, respectively). With respect to LV 2, Q_frASA -2 (atoms with partial charge between -0.05 and -0.1 e) shows the strongest negative influence. This descriptor corresponds to the accessible surface area of various olefinic carbon atoms, some quaternary sp^3 carbons and others, whose influence cannot readily be interpreted and whose role must at present –in spite of its statistical relevance- remain unclear.

In conclusion, it could be shown that a common structure-activity correlation of good statistical quality is possible for this diverse set of STLs, with a clear dependence of activity on alkylant potency, molecular shape, as well as charge distribution. It appears somewhat dissatisfactory that **PAR** as one of the most active compounds had to be excluded from the data set and we are currently investigating the reason for its aberrant behaviour. Since it may be assumed that all of these compounds exert their effect on platelet serotonin release by interacting with a common target structure, we are attempting to apply the *Quasar* approach to this data set (an example for this method can be found in the final section on "GABA-antagonistic sesquiterpene lactones") and to construct a binding site model which may allow for QSAR predictions and design of new 5HT-release inhibitors [M. Ak, current thesis].

Allergic contact dermatitis (ACD)

STLs have long been known as the cause of allergic contact dermatitis (ACD) which occurs in most cases in response to contact with asteraceous plants [74]. The type of immune response involved is delayed type hypersensitivity and also termed "type IV allergy". As most other biological effects of STLs, their sensitizing potential is dependent on the presence of reactive partial structures. It is known that covalent adducts formed between STLs (playing the role of haptens) and proteins of the skin are the responsible allergens. In view of the information on occurrence of such structure elements presented in the first section of this chapter, the previous estimate of about 700 potential sensitizers among the STLs [74] should be updated. It can be estimated that between 3000 and 4000 of the almost 5000 reported STLs are potential allergens (only 830 compounds of 4861 STLs in the database do not possess any of the substructures considered as reactive). A large body of evidence exists showing that the α-methylene-γ-lactone group is the main requisite of STL structure leading to the formation of such antigens. However, the presence of an MGL alone is not sufficient, since some MGL-containing STLs without sensitizing potential are known. On the other hand, it has been demonstrated that the presence of an MGL is not an essential requirement to provoke an

immunological response since other potentially reactive sites, such as α,β-unsaturated ketone- and/or -ester groups as well as epoxide structures are also capable of sensitization [75]. An important issue is the question of cross-sensititvity between different STLs, i.e. what are the structural features to determine whether or not a person sensitized to one STL will show a reaction to another compound of this group (and thus, other plants).

With respect to structure-cross-sensitivity relationships, only little detailed information is available [74]. It appears quite logical that the degree of cross-sensitivity is primarily dependent on the similarity between molecules. When the cross-reactions in individuals sensitized to a very hydrophobic STL of small size, costunolide (COS), were tested, it appeared that the likelihood of reaction to other STLs decreased with increasing hydrophilicity and oxygenation as well as substitution with acetate groups (especially in the vicinity of the MGL) [74]. Whatsoever, no general "code" of cross sensitivity exists except that it appears dependent on molecular similarity. The two following examples show that stereochemistry is extremely influential in this respect. It is quite easy to imagine that molecular recognition systems such as immune receptors will be able to distinguish between enantiomeric forms of a molecule. The phenomenon termed enantiomer-specific ACD [74, 76] depends on the size and complexity as well as charge distribution of the chiral part of the molecule. In case of a simple optically active methylene lactone derivative, 3-methylene-5-methyl-tetrahydrofuran-2-one (**MMT**), no specificity was found with respect to cross-sensitivity of guinea pigs towards the (+) and (-) forms, respectively. Frullanolide is a eudesmanolide from liverworts where it occurs in one species in the form of the (-) enantiomer, (**-FRU**), while another contains the dextrorotatory form, (**+FRU**) (Fig. (**17**)) [76]. The reactive chiral substructure of **MMT** also occurs in **FRU**. In this case, strong specificity exists in guinea pigs sensitized against one enantiomer which do not show cross reaction towards the other optical isomer [74, 76]. In Fig. (**17**) **A**, 3D models of the both forms are shown, in which the reactive methylene lactone structures were aligned in identical positions (arrows). The difference in molecular shape becomes evident and it is easy to imagine that an immune receptor will only recognize one form. On the other hand, the allergen of tulip species with a molecular shape approximately the same as in **MMT**, tulipalin (**+TUL**, **-TUL**), displays complete enantiomeric specificity in guinea pigs [76]. Here, it becomes evident that the orientation of hydrophilic partial structures, influencing H-bonding- and electrostatic interactions, is of importance for cross-sensitivity.

Another well known example ("diastereomeric ACD") shows that recognition is also different between diastereomeric forms of STLs. Parthenin (**PRT**) (Fig. (**17**))is known as the allergenic principle of the plant *Parthenium hysterophorus*, occurring in India where it is also known as the "scourge of India" due to the severeness and frequency of allergy it causes [74, 76]. The diastereomer hymenin (**HYM**), differing from **PRT** in the orientation of the OH group at C-1, occurs in the South American variety of the same species. Patients allergic to Indian *P. hysterophorus* reacted strongly to **PRT** as expected but did not cross-react to **HYM**. Fig. (**17**) **B** shows 3D representations of the two compounds (methylene lactone structure aligned). It becomes clear that the change in configuration at C-1 -which might appear quite subtle when looking at the structural formula- leads to a global change in molecular shape and charge distribution (the latter

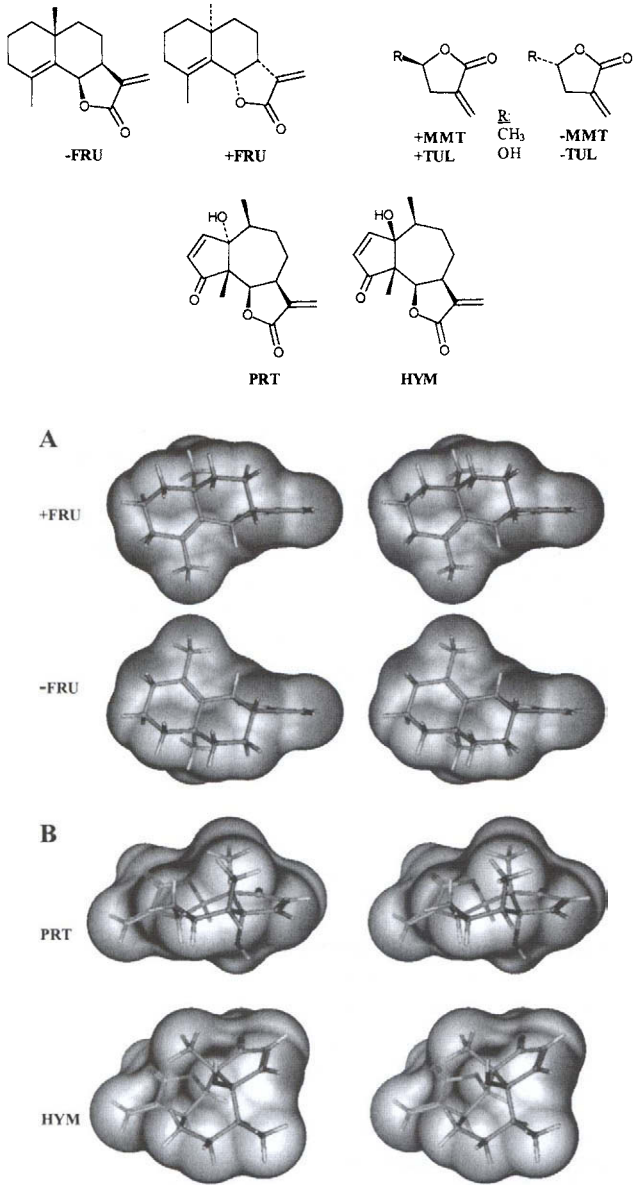


Fig. (17). Differences in stereochemistry leading to lack of cross-sensitivity in allergic contact dermatitis (see text).
A: (+)-Frullanolide (**+FRU**, top) and (-)-Frullanolide (**-FRU**, bottom). Naturally occurring enantiomers.
B: Parthenin (**PRT**) and Hymenin (**HYM**), two diastereomers differing only in the stereochemistry at one chiral centre. In both cases the exocyclic methylene lactone structure was superposed.

not being visible in the grey shade picture). It is obvious that distinction between these two completely different molecules should not represent a problem to immune receptors.

In spite of such interesting examples, it still remains an open question how molecular similarity is defined with respect to immunological response, since it is also known that cross-reactivity exists in some instances between STLs and mono- as well as diterpenoids [74]. It appears therefore a promising topic for future studies to investigate this issue in more detail (i.e. with large sets of compounds and applying quantitative structure-activity relationship methods) in order to elucidate the general rules by which specificity in ACD cross-sensitivity is determined.

Anti-microbial activity

Plants accumulate STLs as a means of chemical defence. Thus, one of the primary ecological benefits of these compounds is certainly the fact that they are toxic to all kinds of microorganisms. Once more it is their alkylant potency that is most obviously responsible. The work on antimicrobial activity of earlier years has been summarized in various reviews, e.g. [1-3]. More recent work has indicated that some STLs may have potential as leads for new antiprotozoal drugs. Apart from the prominent antimalarial artemisinin which will be treated separately in the following section, several more "common" STLs have been demonstrated to possess antiplasmodial [77-79], antileishmanial [80] and antitrypanosomal activity [81-84]. Unfortunately, in most cases such activities are accompanied by high non-specific cytotoxicity so that it is always necessary to compare the antiprotozoal activity with cytotoxic effects determined under the same conditions.

Thus, e.g., **HEL** (**1** in Fig. (**18**)) has recently been reported to possess antiplasmodial activity in a similar concentration as the potent antimalarial STL artemisinin (see following section) [78]. The reported IC_{50} value of 0.23 μM against *Plasmodium falciparum* asexual blood forms was somewhat lower but still quite close to those usually observed in cytotoxicity assays [46, 50, 78]. In addition, activities for derivatives **2**, **3** and **4**, Fig. (**18**) were reported and compared with that of the germacranolide neurolenin B (**14**, Fig. (**18**)), which had previously been reported to possess antimalarial activity [77]. Results on seven further compounds investigated in that same study [85] which were not included in [78] are reported together with the structures in Fig. (**18**).
Moreover, **HEL** and some further STLs were recently shown to possess potent anti-trypanosomal activity [81]. As in many other studies, **HEL** was shown to be the most potent derivative tested showing an IC_{50} of 0.051 μM against *Trypanosoma brucei* trypomastigotes (responsible for the east african sleeping sickness), which was lower than its cytotoxicity against mammalian L6 cell by factor 19.5 and indicating some selective toxicity against these parasites, which would, however, still have to be improved in potential **HEL**-derivatives that might be used in therapy. In a similar way as previously found for other biological activities, the 6,8-diastereomer of **HEL**, mexicanin I (**6** in Fig **(18)**) was less active (IC_{50} = 0.318 μM), once more showing the importance of stereochemistry for the biological effects of these compounds. Whatsoever, alkylant bifunctionality was found to be even more influential, since four monofunctional STLs tested in this study yielded IC_{50}s between 1.4 and 23.6 μM [81]. Interestingly, and in agreement with very early results of Lee et al. on antibacterial and antifungal activity [86], a cyclopentenone structure appeared to confer more effective antitrypanosomal

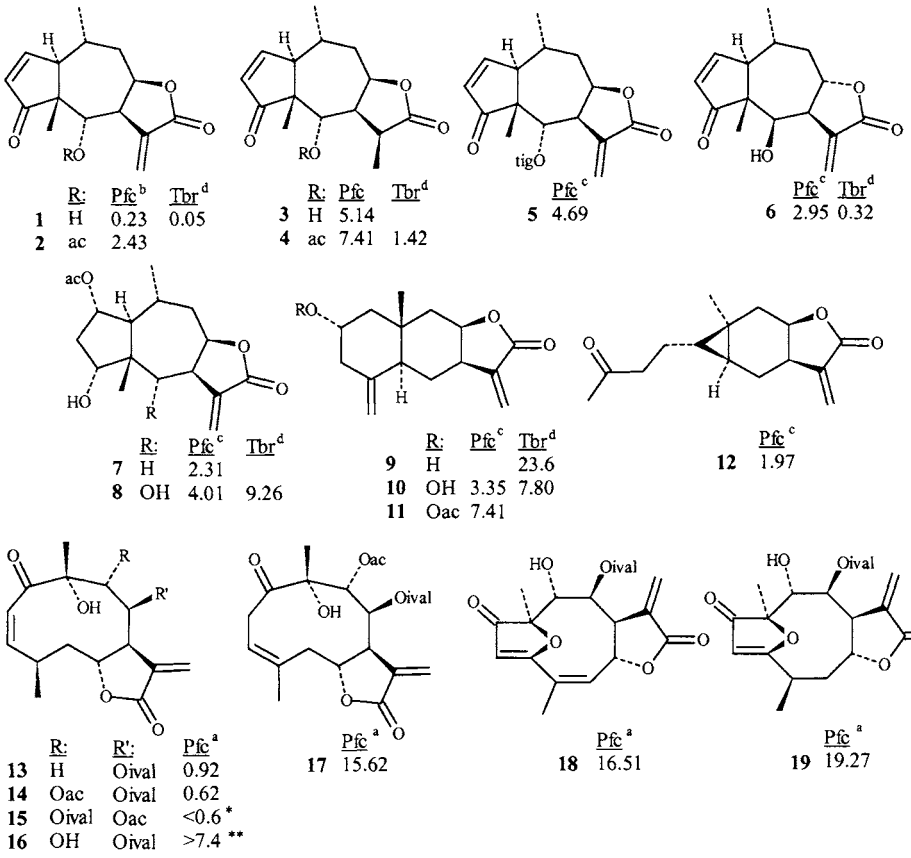


Fig. (18). Structures and antiprotozoal activity (IC_{50}, μM) of some STLs. Pfc: *Plasmodium falciparum* asexual blood forms; Tbr: *Trypansoma brucei rhodesiense* trypomastigotes; Data from: [a][77] [b][78], [c]previously unpublished [85], [d][81]. * mixture (**15**:**16**) 3:2: 0.62; ** mixture (**15**:**16**) 1:3: 7.35.

activity than the α-methylene-γ-lactone group, since dihydrohelenalin acetate **4** was more potent than three compounds with methylene lactone groups. It is also noteworthy, that selectivity towards the trypanosomes was only observed for the bifunctional alkylants while the monofunctionals displayed cytotoxicity in the same range or even higher than antitrypanosomal activity. It was anticipated that the activity against trypanosomes could involve interference with the metabolism of the special thiol metabolism of these protozoa in which trypanothione replaces glutathione. Initial experiments in which the potential inhibitory activity of **HEL** on trypanothione reductase was investigated yielded a negative result [87]. Mechanistic studies and structure-activity studies with a larger number of compounds are still in progress.

As stated above, it is of crucial importance to distinguish specific antiprotozoal effects from non-specific cytotoxicity. It is certainly possible to do so for individual compounds, e.g. by calculating the ratio of IC_{50} values for antiprotozoal and for cytotoxic activity as was done e.g. in [77, 78]. On the other hand, it has been demonstrated that cytotoxicity data for STLs vs. different cell lines do not necessarily

correlate well with each other so that it may be difficult to extrapolate from one cell system to all cells that might be damaged if an STL is administered (see comments on this issue and literature cited in [3]). It may, however, be assumed that a high degree of correlation between two series of activities for the same compounds should indicate that the effects are related with respect to the underlying mechanism of action, and vice versa for a low degree of correlation. It was therefore of interest to investigate activity-activity relationships between the anti-protozoal and cytotoxic activities of these STLs. Table 3 shows the correlation coefficients between these data for the compounds in Fig. (**18**).

Table 3. Correlation matrix for the antiprotozoal and cytotoxic activity of the STLs in Fig (18).
Correlation coefficients (*r*) between pIC_{50} values. In brackets: number of observations (compounds) underlying these values.

	P.falc.	*T. br. r.*	*T. cr.*	L6	KB	GLC4	COLO320
P.falc.							
T. br. r.	0.76 (5)						
T. cr.	0.70 (5)	0.96 (6)					
L6	0.51 (5)	0.47 (6)	0.55 (6)				
KB	0.44 (11)	0.68 (6)	0.61 (6)	0.49 (6)			
GLC4	0.63 (9)	*	*	*	0.96 (4)		
COLO320	0.61 (9)	*	*	*	0.97 (4)	0.99 (9)	

P. falc.: Asexual blood forms of *Plasmodium falciparum* [77, 78, 85]
T. br. r.: Trypomastigotes of *Trypanosoma brucei rhodesiense* [81]
T. cr.: Trypomastigotes of *Trypanosoma cruzi* [81].
L6: Uninfected rat skeletal myoblasts used for the *T. cruzi* assay [81].
KB: Human cervix carcinoma cell line, 72 h incubation [50].
GLC4: Human lung cancer cell line, 72 h incubation [77, 78, 88].
COLO320: Human colon cancer cell line , 72 h incubation [77, 78, 88].
* Correlation coefficients not determined due to lack of coherent data.

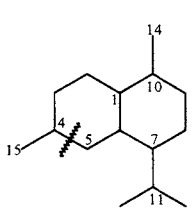
Several conclusions can be drawn from these data. There is a strong correlation between the antitrypanosomal effects ($r > 0.9$), which, on the other hand do not correlate well with the cytotoxic activities against primary (L6) cells (both $r < 0.6$). The same holds true for the antiplasmodial data. The correlation of each of the antiprotozoal effects with the three cancer cell lines is somewhat higher but still quite moderate ($r < 0.7$ in all cases), confirming the postulated existence of a certain selectivity [77, 78, 81]. Moreover, some (although not very strong) correlation is observed between the antitrypanosomal and antiplasmodial effects ($r > 0.7$) so that the mechanisms underlying the antiprotozoal effects are more related to each other than to any of the mammalian cells. The degree of correlation between the three human tumour cell lines, on the other hand is very high ($r > 0.95$ in all cases), indicating that the antitumoral activity is exerted by the same mechanism for each compound. Quite interestingly, there is no significant correlation between the cervix carcinoma cells (KB) and the primary cell culture (L6; $r < 0.5$). Unfortunately, the L6 data could not be correlated with the other two cancer cell lines since coherent data were only available for two of the 19 compounds. However, the absence of correlation between L6 and KB cell cytotoxicity seems to indicate that different mechanisms underly the toxic effects against transformed (cancer) and untransformed (primary) cells. In view of these apparent selective effects, it appears worth-while to establish more detailed (Q)SAR for these activities which we are currently attempting.

ARTEMISININOIDS

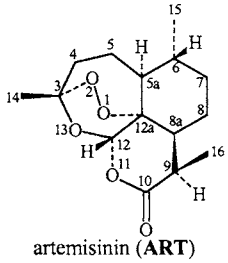
The STL Artemisinin or Quinghaosu (**ART**), see Fig. (**19**) was isolated by Chinese researchers from *Artemisia annua* (Sweet or Annual Wormwood, Asteraceae, chin. "Quinghao") in 1972 in the course of a program directed at the evaluation of traditionally used medicinal drugs. It was identified as the active principle in this plant which had been utilized in traditional chinese medicine for the treatment of febrile illnesses for many centuries [89]. The unusual sesquiterpenoid structure of the compound, also termed arteannuin, was solved in 1979 [90]. Artemisinin and derivatives were found to be highly effective antimalarial agents and are now successfully used for the treatment of severe forms of malaria caused by *Plasmodium falciparum*, especially of strains resistant against other antimalarials [91-94]. The structure of **ART** is derived from the cadinane skeleton which has been modified by scission between C-4 and C-5 to yield a *seco*-cadinane (note that the numbering scheme commonly applied to **ART** and derivatives is different from that used for cadinane-type sesquiterpenes). The most prominent feature of the structure is the 1,2,4-trioxane ring system of which the *endo*-peroxide moiety was demonstrated to be essential for its antiparasitic activity [91-93].

Mechanism of action

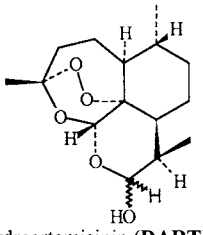
The mechanism of action as far as currently known consists of several steps, the first of which is selective uptake of **ART** by infected erythrocytes which is higher by a factor of ca 100 in comparison with normal healthy red blood cells. Inside the infected cell, **ART** interacts with hemin, a soluble iron-porphyrin material formed from hemoglobin as a digestion product by the parasites. Hemin is toxic to the parasites which therefore transform it enzymically into an insoluble nontoxic polymer, hemozoin. Redox-activation of the endoperoxide group occurs by reaction with the Fe-atom of hemin (and other iron containing porphyrins as well as with free iron) leading to formation of reactive radical intermediates which are capable of undergoing covalent addition (alkylation) to porphyrins and -possibly more important- to specific proteins of the parasites ("malaria specific proteins"), thus ultimately leading to cell damage and parasite death. The convincing lines of evidence leading to this general mechanism have been reviewed in great detail by Robert & Meunier [95] and by Meshnick [92, 93]. As an important part of this evidence, several studies have shown the necessity of redox-activation leading to the formation of carbon-centered radical species which are highly reactive alkylants. It has been shown that **ART** after activation directly alkylates porphyrins such as heme, as well as proteins. The reaction with proteins seems to be favoured since after reaction of radiolabelled **ART** with hemoglobin, the major part of the radioactivity is associated with the protein part; moreover formation of covalent adducts with albumin was proven with **ART** and **DART**, and was found to proceed more quickly in the presence of heme. Most imporatantly, after incubation of radiolabelled aretemisininoids with infected erythrocytes, alkylation of six non-abundant proteins ("Malaria specific proteins") was found, while no adduct formation with proteins is found when uninfected blood cells are used [92, 93]. Moreover, several lines of evidence suggest that reaction with heme –although detectable in infected erythrocytes- does not lead to antiplasmodial effect but is rather a side-reaction accompanying the more important alkylation of proteins [92, 93]. The reaction with heme has been investigated



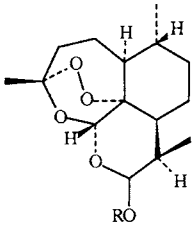
cadinane



artemisinin (**ART**)



dihydroartemisinin (**DART**)

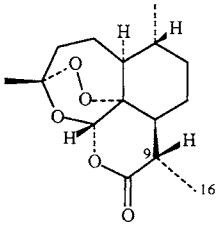


R

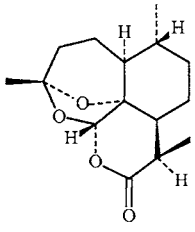
( )-Me artemether

( )-Et arteether

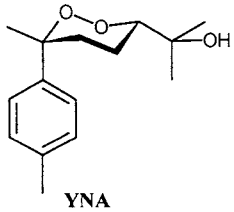
( )-$CO(CH_2)_2COONa$ artesunate



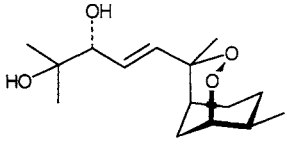
9-epi -**ART**



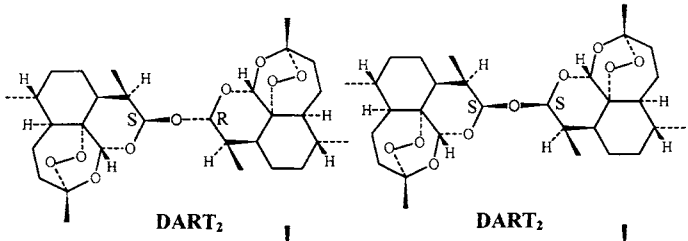
DOA



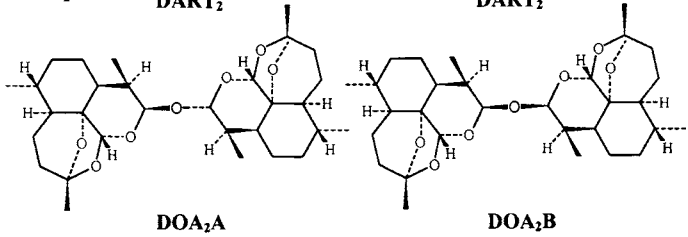
YNA



YNC



$DART_2$ **$DART_2$**



DOA_2A **DOA_2B**

Fig. (19). Stuctures of the artemisininoids mentioned in the text

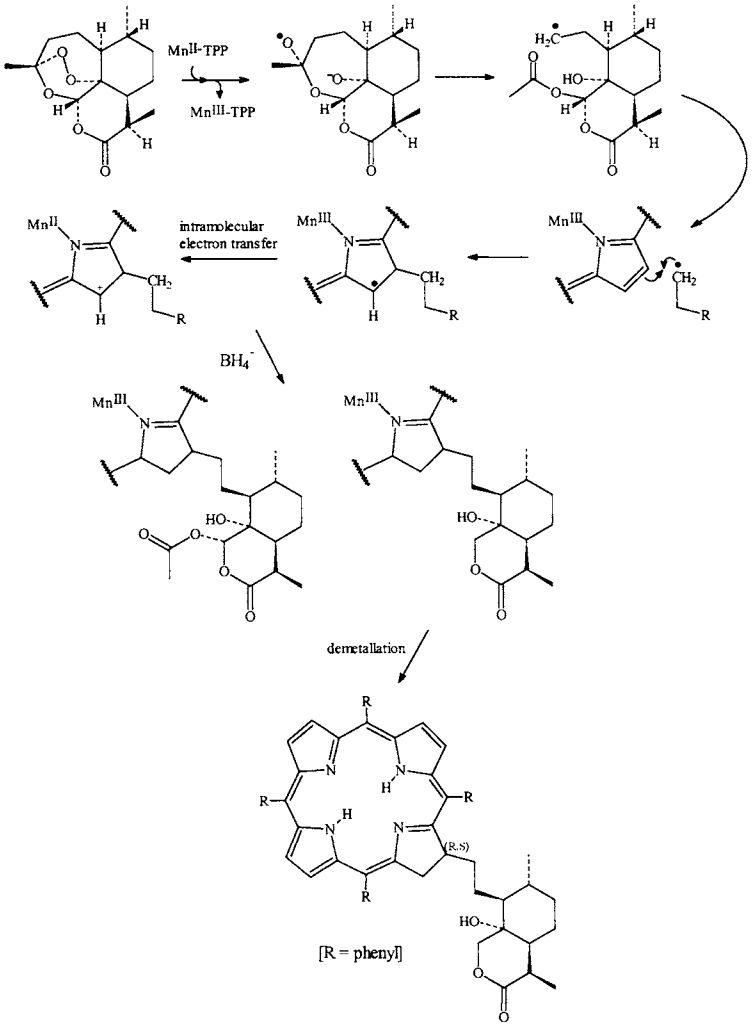


Fig. (20). Mechanism of radical formation and alkylation of a heme surrogate (Mn^{III}-tetraphenyl-porphyrine) by **ART** as shown by Robert & Meunier [96].

in great detail and it was proven unambiguously by NMR of the resulting adduct that activation by a manganese(II)-containing heme analog (tetraphenylporphyrin-Mn(II), TPP-Mn(II)) results in the formation of a carbon-centered radical at C-4 which subsequently alkylates the porphyrin ring system as shown in Fig. (**20**) [96]. Convincing evidence has been presented for the same mechanism to occur in the case of an Fe(II) central atom [96].

In order to alkylate proteins, the reactive species should either be able to leave the porphyrin complex or the activation reaction is to take place in the immediate vicinity of such alkylable structures. It has not been ultimately clarified how such interaction with proteins occurs. In a very recent study, the mechanism involving formation of C-

centered radicals has been seriously questioned [97], especially based on the finding that some **ART** derivatives without an oxygen function at C-10 are completely unaffected by treatment with ferrous heme but show high anti-malarial activity. A further piece of evidence contradicting the hypothesis of heme-activation is that artemisininoids are readily metabolized by CYP450 enzymes, during which ART is hydroxylated at the periphery of the molecule *syn* to the peroxide bridge which does not interfere with the enzyme's heme group [97]. The authors propose an interaction of **ART** with some specific target which then may be accompanied or followed by activation of the peroxide and alkylation of this particular target [97]. One such malaria specific protein, and possibly the most important one with respect to plasmodicidal activity, has quite recently been identified as a membrane Ca^{2+}-ATPase of the parasite. PfATP6, is a plasmodial orthologue of sarco-endoplasmic-reticulum Ca^{2+}-ATPases (SERCAs) [98]. The inhibitory profile and potency of **ART** on this important calcium pump were very similar to those of the potent SERCA inhibitory STL thapsigargin (see following section) but unlike the latter, and in support of the activation hypothesis, **ART** required the presence of iron to inhibit the enzyme [98].

In any case, it is clear that artemisininoids act by an alkylant mechanism and that this mechanism differs from that shown by other alkylant STLs as discussed above. It is highly noteworthy that selective toxicity occurs towards various states in the life-cycle of *Plasmodium* since heme-associated toxification by the described mechanism should occur also in uninfected cells. This selectivity is likely to be -at least in part- due to selective sequestration in infected erythrocytes as mentioned above. Evidence has been presented that –apart from targeting malaria-specific proteins- processes consuming metabolic energy, i.e. active transport, is involved in this selective uptake [92] but the transport mechanism is not known in detail.

Structure-activity relationships

A great number of structure-activity studies on artemisinin and related analogs exist in the literature, most of which were conducted in order to optimize the naturally occurring lead structure with respect to better bioavailability or pharmaceutical properties. It was found that the major structural feature to retain activity is the endoperoxide bridge (see above), while the lactone group is not necessary for activity. Reduction of the lactone carbonyl to dihydroartemisinin (**DART**) leads to better solubility as well as increased anti-malarial activity. In fact, ether- and ester derivatives of **DART**, such as artemether, arteether and artesunate, are the **ART**-analogs of choice currently used in mararia chemotherapy [92, 94]. A shortcoming of such lactol derivatives is, however, their lower stability as well as the occurrence -although in relatively few cases- of neurotoxic side effects [92, 93]. In many studies it has also been shown that the complex ring system of **ART** is not essential for activity. Other sesquiterpenoids, the bisabolane derivatives yingzhaosu A and -C (**YNA, YNC**) from *Artabotrys uncinatus* (Annonaceae), show similarly high activity as **ART** in spite of lacking the lactone structure and showing a much less complex cyclisation pattern [99, 100]. A large number of fully synthetic compounds with endoperoxide structures, mostly mimicking the trioxane system of **ART**, have been tested and recent advances in the synthesis and biological evaluation of such derivatives have been summarized by Bez et

al. [101]. In this contribution dedicated to structure-activity of STLs, we shall confine the discussion to (Q)SAR studies for such derivatives directly derived from **ART**.

The overall picture emerging from these studies is that unhindered approach of the peroxide group to a target molecule (with presumably planar geometry and containing an Fe(II) central atom) must be possible to warrant high activity. This would be in line with the "activation hypothesis" described above and requires a flat "α-side" of the molecule while molecular shape on the "β-side" seems to be of negligible influence. Thus, e.g. 9-epi-**ART** has been demonstrated to be a remarkably poor antimalarial which was attributed to sterical hindrance imposed by the C-16 methyl group when oriented towards the α-side and to effects on the mode of interaction with heme and subsequent radical formation associated with this geometrical change [102]. These results were in line with previous findings on synthetic C-5a-α-substituted trioxane analogues, which were inactive against plasmodia and also did not yield any porphyrin-drug adducts [103].

The methods applied in recent years by various groups to construct QSAR models for **ART** and analogues include docking calculations to heme [104, 105], CoMFA [106-109] and hologram QSAR [108] as well as the hypothetical active-site lattice (HASL) approach [107], self-organizing maps of the molecular electrostatic potential [110], quantum-similarity measures [111] and topological molecular connectivity descriptors [112].

Tonmunphean et al. conducted a QSAR study based on molecular docking in which binding energies of 30 artemisininoids to heme and some Fe--O interatomic distances obtained from the optimized complex geometries were correlated with antimalarial activity towards two different strains of *Plasmodium falciparum* [104]. Of the parameters investigated, total binding energy showed the highest degree of correlation with the biological data ($r^2 = 0.87$ and 0.66 for the two different strains), while the Fe--O distances only showed moderate correlation ($r^2 < 0.3$). Based on the docking energies, the activity of 5 further derivatives was calculated and the predictions were in very good agreement with the experimental activity [104]. These results proved on a quantitative basis the high importance of unhindered interactions with heme for good antimalarial activity.

Several studies on data sets of considerable size, mainly utilizing the CoMFA method [113], were conducted by the group of Avery [106-109].

In a recent study based on CoMFA and 2D fragment-based QSAR ("hologram QSAR", HQSAR [114]), 211 **ART** analogues were included [108]. The performance of an alignment of 3D structures obtained by standard *in-vacuo* energy-minimisation was compared with such representing a hypothetical "bioactive conformation" obtained by optimising the structures in a complex with hemin derived from docking calculations. Surprisingly, both methods yielded 3D QSARs of similarly high internal predictivity; the alignment based on "bioactive conformation" performed even somewhat worse in predicting the activities for an external test set. Although some differences are observed in the typical CoMFA contour maps obtained by the two different alignment rules, which are discussed in some detail by the authors, the biological relevance of these differences is not clear since both models are more or less equivalent with respect to internal statistical quality. In short, the maps for the CoMFA model based on alignment of the standard geometries once more show that increased steric bulk on the α-side of the molecule is associated with a loss of activity while an increase of activity may be

expected as a consequence of steric bulk around the C-3 and C-9 substituents. Favourable electrostatic interactions are associated with negative partial charge on the α-side (roughly around the peroxide bridge) while favourable interactions with positive partial charge (or unfavourable ones with negative charge) appear around the lactone carbonyl oxygen and adjacent regions. This latter finding possibly reflects the observation that 10-dihydro- and 10-deoxy-**ART** derivatives are more potent antimalarials than **ART** itself, indicating some detrimental effect of the C-10 carbonyl oxygen. With respect to these points it can be stated that they were in full agreement with earlier QSAR analyses from the same group, where CoMFA and another 3D-QSAR approach, the hypothetical active-site lattice method (HASL), had been used and led to similar results for a series of 154 compounds [106, 107].

Quite interestingly, it was shown in the more recent study [108] that a subset of 34 compounds whose biological data had been obtained with racemic mixtures led to very poor models when treated on its own. Consistently, the statistical parameters of models for the total data set were also improved when such racemates were eliminated from the data set, i.e. not considered in the model building process.

In addition to the 3D-QSAR method, the HQSAR approach was applied to the same data set in this study [108]. In this modelling technique, descriptors derived from molecular fingerprints (molecule fragments of variable size) are correlated via PLS with the biological activity [114]. It appears noteworthy, that HQSAR led to models of equal or even somewhat improved statistical quality as did the 3D QSAR approach, since HQSAR is independent of molecular alignment which is a major advantage over the CoMFA strategy whose major pitfall is that slight changes in alignment can lead to considerable differences between the resulting models [115]. Thus, the best HQSAR model in this study contained all 191 active molecules of the data set and showed even slightly better internal predictivity than the best CoMFA model obtained after elimination of the racemic compounds (q^2 = 0.76, n = 191 vs. 0.73, n = 157). This HQSAR model was constructed by including information on molecular chirality and the fact that such information was necessary to arrive at a high quality model pointed into the same direction as the poor results obtained with racemates in the CoMFA study.

In spite of some reports that the separated enantiomeric forms of artemisininoid analogues yielded equal in-vitro anti malarial potency, the authors hypothesized on grounds of these results that chirality might still play a role in the antimalarial potency of artemisininoids [108], but this issue still awaits further experimental data.

A further CoMFA study from the same group [109] addressed the structure-activity relationships of some 70 artemisininoids against *Leishmania donovani*, a protozoal parasite causing another severe tropical disease termed Kala-Azar (visceral Leishmaniasis [94]). The compounds showed in-vitro activity against *L. donovani* promastigotes in the micromolar range while the antimalarial IC_{50} values are in the nanomolar range. However, the results on the whole were quite similar to those of the above mentioned study for anti-malarial activity. In the same way as in case of *Plasmodium*, the peroxide group was an essential requirement for activity. Substitution at the C-9β position improved activity in both cases and 10-deoxo-derivatives showed better activity compared with the lactones. Considering the overall similarity of the

activity profiles and of the CoMFA models, the authors concluded that a common mechanism of action could be responsible for both antiparasitic activities.

Taking into account the likely mechanism of action, it is straightforward to assume that apart from a close fit to heme as the potential target, electronic factors are likely to be involved in the antimalarial activity of **ART** and analogues. Jefford et al. recognized the necessity to take such electronic factors into account and conducted a study with the aim to compare directly the electronic properties of **ART** and several related organic peroxides by creating self-organizing maps of their molecular electrostatic potentials (MEPs) [110]. This approach originally proposed and applied by Gasteiger et al. [116] reduces the dimensionality of a compound's MEP (which is a 4-dimensional object and thus difficult to visualize and compare) by projecting it onto a 2D plane without significant loss of information. This goal is obtained by feeding the 4D-information into a self-organizing neural network (Kohonen-neural network model, KNN [117]) which – after an iterative training process - results in an optimized 2D representation of the MEP.

This study resulted in the finding that a conspicuous difference exists between the MEPs of active and inactive compounds [110]. The active analogues **ART**, yingzhaosu A and a synthetic trioxane with high activity all share a high degree of delocalization with respect to their negative electrostatic potential leading to continuous ribbons across the KNN maps, while in case of inactive derivatives the negative regions are segregated into isolated patches by regions of positive potential.

Although the presented results do not offer the possibility to predict antimalarial potency in a quantitative manner, this method apparently allows an efficient discrimination between active and inactive analogues. Moreover, it takes into account aspects of charge delocalization which are difficult to capture by means of force-field based methods such as CoMFA. In view of the time and cost involved in such calculations, however, the authors propose this computationally more sophisticated method as a second selection step which could best be used after having sorted out the more active from the less active candidates by a less expensive 3D-QSAR modelling technique [110].

Finally, three further studies on QSAR of artemisininoids applying a variety of quantum-chemical and conventional molecular descriptors [105], molecular quantum-similarity measures (MQSM, [111]) and topological descriptors based on molecular connectivity [112] have led to models of quite satisfactory statistical performance. However, apart from showing the applicability of the respective QSAR approaches to this type of compounds both studies offer comparatively little new information with respect to structure-activity relationships.

The most conspicuous structure-activity relationships with respect to antiprotozoal activity of artemisininoids emerging from the mentioned studies -as well as from numerous further reports not directly aimed at (Q)SAR- are summarized in Fig. (**21**). The region on the α-side labelled A does not tolerate any changes, i.e. the peroxide must be present and no further substituents should point in this direction in order to allow interactions of this more or less planar region with the biological target. Regions B and C around the methyl groups at C-3 and C-9, respectively, can accommodate larger, more

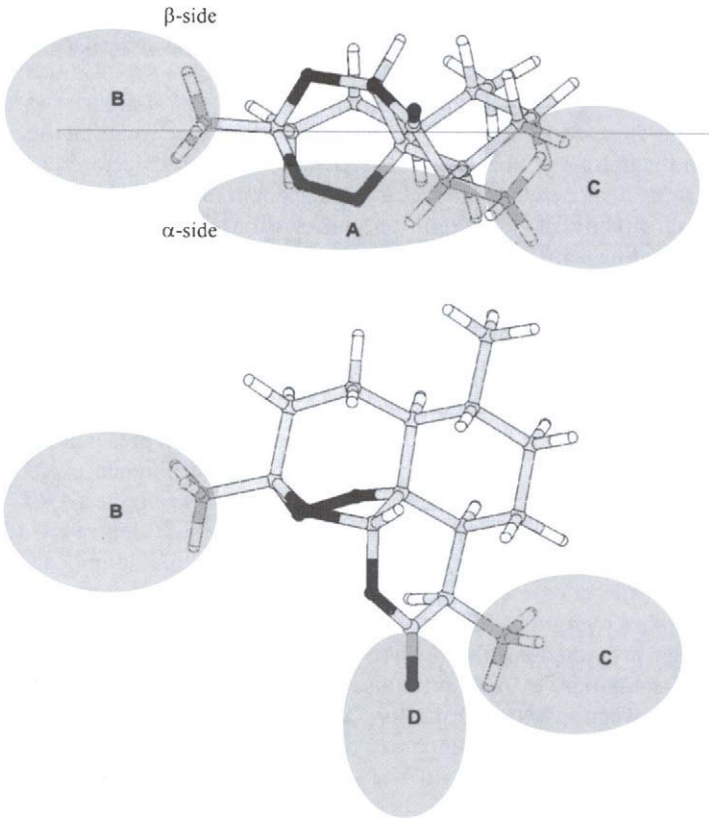


Fig.(21). Summary of the most important structure-activity relationships for antiprotozoal activity of artemisininoids as derived from various studies (for explanation of the shaded areas see text).
The 3D structure of **ART** is shown from two different angles. Above, the molecule is displayed in a "lateral" view, below it has been rotated by 90° around the X-axis (line) relative to the former, thus showing the molecule from the β-side.

hydrophobic residues whose presence leads to an increase in activity. In region C it appears to be essential that such substituents are introduced in the β-configuration at C-9, since otherwise they would be axially oriented and interfere with region A. In region D around the lactone carbonyl oxygen, reduction to a lactol or CH_2 as well as introduction of more lipophilic substituents leads to enhanced activity. From the CoMFA analyses it appears as though the lactone group has some detrimental effect due to electrostatic factors. With respect to steric bulk, this region can tolerate residues of considerable size and –in contrast to C-9- it is not necessary to have these substituents attached in the β-configuration at C-10 since a quasi-equatorial orientation is also obtained with α-substituents so that interference with region A will occur in neither case.

At the end of this section, some studies on the cytotoxic and anti-tumour potential of dimeric **ART** derivatives deserve to be mentioned. On one hand, cytotoxicity testing has been conducted because compounds for potential use against protozoal infections should possess a low degree of cytotoxic activity towards human cells in order to achieve a

maximum therapeutic index of drug safety. Thus, it was shown that some semi-synthetic dimers possess increased antimalarial activity over **ART** [118, 119]. On the other hand, some of these dimers were shown to possess interesting potential as new anti-tumour leads [118-122]. It was shown that **ART**, apart from its antiprotozoal activity also shows considerable growth-inhibitory activity on various tumour cell lines [120-122]. Thus, e.g. its IC_{50} value against a murine Ehrlich ascites cell line (EN2) was 0.98 μM when incubated with the cells for 72 h and cell viability determined by the MTT assay [121]. In the same way as described above, the presence of an endoperoxide group appeared important for activity, since deoxyartemisinin (**DOA**) was about 100 times less active. Quite interestingly, dimers of dihydroartemisinin (**DART**), structures see Fig. (**19**), linked via an acetal bridge between the lactol oxygens were demonstrated to show a pronounced dependence of their cytotoxic effect on the stereochemistry of the linkage: In the "non-symmetrical" dimer (10S, 10'R: **$DART_2A$**), activity was enhanced to an IC_{50} of only 0.11 μM while the "symmetrical dimer" (10S, 10'S: **$DART_2$-B**) showed a somewhat reduced activity in comparison with **ART** (IC_{50} = 2 μM). Most strikingly, a dimer of dihydro-deoxyartemisinin, analogous to **$DART_2A$** (**DOA_2A**) devoid of the peroxide group, although much weaker than the **$DART_2$** derivatives, still showed significant cytotoxicity (IC_{50} = 8.9 μM). The same dependence on stereochemistry as for the **$DART_2$** derivatives was observed, the "symmetrical dimer" **DOA_2B** being much less cytotoxic (IC_{50} = 99.8 μM). Hence it was concluded that the peroxide bridge is of less importance to these toxic effects on tumour cells than to antimalarial activity and that, probably, a different mechanism of action is involved. Mechanistic studies investigating the effect of **ART** and the **$DART_2$** derivatives on cell cycle resulted in the interesting finding, that the dimers caused a significant dose-dependent G_1-accumulation of the cells, while **ART** displayed a smaller (but still significant) accumulation of the cells in the S-phase. Thus it could be concluded that different mechanisms underlie the cytotoxic activity of **ART** and the dimeric acetals [121]. After testing dimers **$DART_2A$** and **$DART_2B$** in the NCI drug screening panel of 60 human tumour cell lines, some very promising activities were found for **$DART_2A$** with IC_{50} values in the nanomolar range [122]. A COMPARE analysis [123] over the NCI database did not result in any significant correlations of the activity profile with anti-tumour compounds with known mechanism of action so that the authors postulated a new mechanism of action for their **ART**-derivatives [122].

Considering the interesting activity differences related to stereochemistry and the presumably new mechanism of action, more detailed studies on structure-activity relationships of such dimeric atemisininoids would be of high interest but have apparently not been carried out up to present.

Finally it should not remain unmentioned, that several STLs of other structural types have been found which contain a cyclic peroxide structure. The STL database (see "Structural diversity of sesquiterpene lactones") contains 88 entries with peroxide moieties of which 24 contain this structure as part of a ring system. Of these, 21 are not derived from *seco*-cadinane but belong to the guaianolide and xanthanolide series. It remains to be shown whether these compounds possess anti-protozoal activity in a similar way as the artemisininoids.

THAPSIGARGIN AND ANALOGUES

A comparatively small group of structurally unique guaianolide type STLs known from *Thapisa* species (Apiaceae/Umbelliferae) are known to inhibit with high specificity and very high affinity (low nanomolar concentrations) intracellular calcium-pumps termed sarco-endoplasmatic reticulum Ca^{2+} ATPases (SERCA). Thapsigargin (**TG**) was identified as the active principle in the roots of *T. garganica* [124-127] which cause severe inflammation of the skin after contact and which had been used in traditional medicine as a counter-irritant for centuries [128, 129].

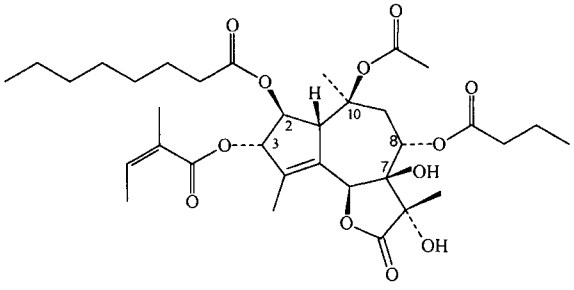


Fig. (22). Structure of thapsigargin (**TG**), a guaianolide from *Thapsia garganica*.

Thapsigargin was found to efficiently induce mast cell degranulation and histamine liberation [124, 125, 130, 131] as well as eicosanoid biosynthesis [132]. Lateron, it was demonstrated to be a potent inducer of apoptosis [133, 134], a promotor of skin tumours [135] and activator of nuclear factor κB [136, 137]. These effects are accompanied by drastic changes in intracellular Ca^{2+} homeostasis and the molecular mechanism underlying this action was elucidated to be a direct and specific inhibition of SERCAs [138-140]. This group of enzymes are located in the sarcoplasmatic reticulum (SR) membrane and are responsible for rapid re-uptake of Ca^{2+} after release from the SR in the course of a Ca^{2+} signal [139, 140]. Inhibition of SERCAs initially leads to strong elevation of cytosolic Ca^{2+} which -among other effects- can be held responsible for mast cell degranulation [130] as well as induction of NF-κB action and apoptosis [136, 137]. SERCAs belong to the superfamily of P-type ATPases and consist of a large cytosolic part containing the ATP-binding and phosphorylation sites coupled to a transmebranary part which consists of 10 membrane-spanning α-helices and is responsible for the binding of Ca^{2+} and its liberation into the SR [141-144]. The catalytic mechanism of the SERCAs according to an E1-E2 model is explained by changing the affinity of Ca^{2+} binding sites from high (E1) to low (E2) [144]. It was shown that **TG** at extremely low concentrations (sub-nanomolar) disrupts this cycle by effectively stabilizing the Ca-free E2 state of the enzyme thus forming a dead-end complex incapable of binding either Ca or ATP [141-143]. It was evidenced that the binding site for **TG** includes amino acid residues from one or several of the hydrophobic transmembrane helices involved in Ca^{2+}-binding and release [145] and that the M3 transmembrane segment is essential for **TG** sensitivity of SERCA [146].

Quite recently, the structures of SERCA from rabbit hind leg muscle in the E1Ca-state complexed with trinitrophenyl-AMP and two bound Ca^{2+}-ions as well as of Ca- and

ATP-free SERCA complexed with **TG** (E2(**TG**)) were solved by X-ray crystallography [144, 147]. The mechanism of SERCA as a Ca^{2+} pump could thus be elucidated. Moreover, the full structure of the Thapsigargin binding site is now known (see Fig. (**23**)) and the numerous observations made previously with respect to SAR of Thapsigargin and analogues [3] (see Table 4), can be rationalized on an atomistic scale.

The Thapsigargin-type guaianolides are unique not only with respect to their molecular mechanism but also concerning their structure. They differ from all other guaianolides known from higher plants by showing a C-7α(12)-6β lactone ring fusion in which C-11 is attached to C-7 in the α-orientation. The absolute stereochemistry of **TG** has been solved unambiguously [125] and the 7(R),11(S) *trans*-diol group has been hypothesized to originate from a 7,11-epoxide formed from an ordinary (C-7-β-substituted) slovanolide guaiane skeleton [148]. Furthermore, the **TG**-derivatives are highly functionalized and -as another difference to most STLs from higher plants- in several cases esterified with medium-chain fatty acids such as octanoic and hexanoic acid. **TG** and analogues thus are among the most lipophilic STLs known.

Structure-activity relationships

From Fig. (**23**) it becomes clear that the C-2-O-octanoate moiety is important for binding to the SERCA forming hydrophobic contacts over a considerable area of the binding site. The low activity of C-2 deoxy derivatives like trilobolide (structure **2** in Table 4) [149] is thereby easily explained. Moreover, the results from an SAR study showing a high degree of correlation of histamine-liberating potency (degranulation of rat peritoneal mast cells) in a series of 26 **TG** analogues with lipophilicity (log k', RP18 HPLC) [150] are in good agreement with this observation. For 15 out of the 26 compounds studied, the correlation coefficient r between log k' and -log EC_{50} was 0.95 (note from the author: r for all 20 compounds with measurable pEC_{50} is still as high as 0.86). The subset of 15 compounds was chosen mainly since all of them possess an unchanged 7,11-trans glycol moiety and the highly significant correlation with log k' indicated that in this series lipophilicity is indeed the major (if not sole) determinant of activity. This is in agreement with the observation that binding to the enzyme is strongly dominated by hydrophobic interactions [147]. Apart from the C-2 octanoate moiety, also the angelate side chain as well as the n-butanoate group at C-3α and C-8α, respectively, are engaged in binding, both fitting snugly into hydrophobic cavities of the enzyme. Here, it was shown that the mere presence of these side chains does not warrant high activity, since the C-8 epimer of **TG** shows 3000 times lower activity and the 2β,3β-dioctanoyl derivative is less active than **TG** by a factor of 400 while the 2β,3α-dioctanoate is still quite active (activity decreased by factor 11) [145]. The need for the correct orientation of the C-3- and C-8 residues is explained by the fact that they do not fit into the mentioned cavities in the opposite configuration. In case of the C-8-epimer, the drastic decrease can be explained by a steric clash of the bulky side chain as well as the inability to form the reported hydrogen bond between the ester carbonyl and the backbone NH of isoleucine I829 [147]. Moreover, the observation that replacement of the C-3 angeloyl group by a second octanoyl or a keto function leads only to 11- or 70-fold decrease in activity, while C-3-epimerization leads to 400-fold decrease respectively [145], is in the same line.

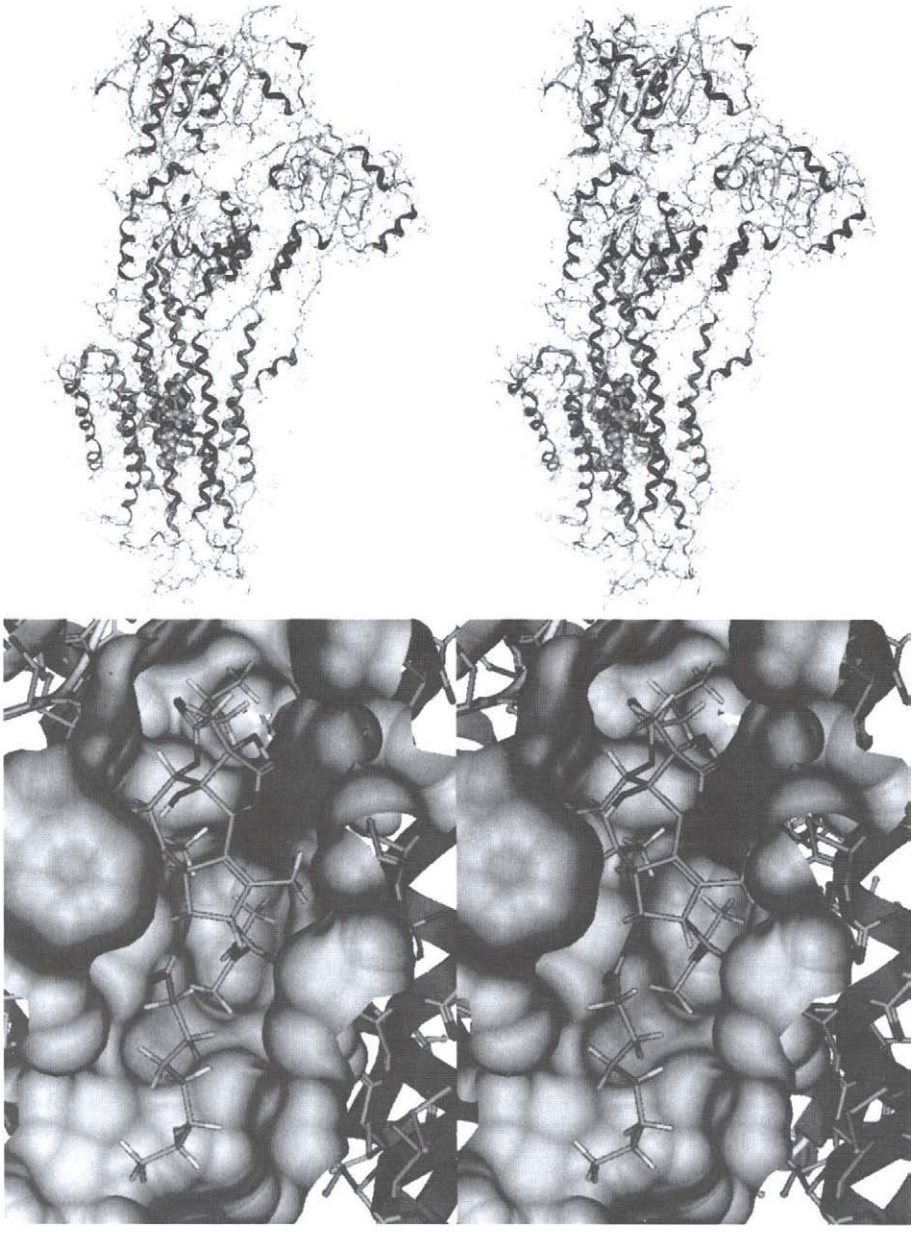
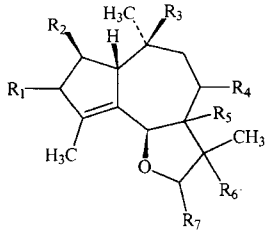


Fig. (23). Stereo representation of the X-ray structure of SERCA with a bound molecule of thapsigargin (**TG**) [147] (PDB code 1IWO). The lower picture shows **TG** in a surface representation of the binding site illustrating the perfect fit of the very potent inhibitor.

Table 4. Structures and biological activity of the Thapsigargin derivatives mentioned in the text.
Biological activity is in each case expressed as the ratio of the IC_{50} value of the analogue to that of thapsigargin (**TG**) from the same study. Note, however, that the activity values represent results from different testing systems. [145, 149-153].



	R_1	R_2	R_3	R_4	R_5	R_6	R_7	rel IC_{50}	refs.
1 (TG)	α-O-ang	O-oct	O-ac	α-O-nbut	β-OH	α-OH	=O	1.0	a
2	α-O-ang	H	O-ac	α-O-mebut	β-OH	α-OH	=O	63.6	[149]
3	β-O-oct	O-oct	O-ac	α-O-nbut	β-OH	α-OH	=O	438	[145]
4	α-O-ang	O-oct	O-ac	β-O-nbut	β-OH	α-OH	=O	3124	[145]
5	α-O-oct	O-oct	O-ac	α-O-nbut	β-OH	α-OH	=O	11	[145]
6	=O	O-oct	O-ac	α-O-nbut	β-OH	α-OH	=O	66	[145]
7	α-O-ang	O-oct	OH	α-O-nbut	β-OH	α-OH	=O	42	[145]
8	α-O-ang	O-hex	O-ac	α-O-nbut	β-OH	α-OH	α/β-OH	1.63	[151]
9	α-O-ang	O-hex	O-ac	α-O-nbut	β-OH	α-OH	α-O-ac	1.92	[151]
10	α-O-ang	O-hex	O-ac	α-O-nbut	β-OH	α-OH	β-O-et	1.17	[151]
11a	α-O-ang	O-hex	O-ac	α-O-nbut	β-OH	α-O-C(CH_3)$_2$-O-α		16.4	[151]
11b	α-O-ang	O-hex	O-ac	α-O-nbut	β-OH	α-O-CH(β-O-et)-O-α		2.51	[151]
11c	α-O-ang	O-hex	O-ac	α-O-nbut	β-OH	α-O-CH(α-O-et)-O-α		40.4	[151]
12	α-O-ang	O-oct	O-ac	α-O-nbut	β-O-ac	α-OH	=O	2.8	[145]
13	α-O-ang	O-oct	O-ac	α-O-nbut	β-OH	α-O-ac	=O	2.5	[145]
14	α-O-ang	O-oct	O-ac	α-O-nbut	β-O-ac	α-O-ac	=O	15	[145]
15	α-O-ang	O-oct	O-ac	α-O-nbut	β-O-β		=O	n.a.	[150]
16	α-O-ang	O-hex	O-ac	α-O-nbut	β-OH	α-O-α		5.4	[151]
17	α-O-ang	O-oct	O-ac	α-O-X*	β-OH	α-OH	=O	3.4	[153]

[a] IC_{50} = 0.17 nM: $^{45}CA^{2+}$ uptake inhibition in bovine cerebellar microsomes [145], 9.6 nM: Inhibition of Ca^{2+}-ATPase activity (phosphate release from ATP) in rabbit hind-leg muscle microsomes [151], 33 nM: Inhibition of histamine release from rat peritoneal mast cells [150], 33 nM: Specific binding to mouse brain microsomal fraction [151]. * X = –CO-(CH_2)$_{11}$-NH-CO-CH(CH_2(CH_3)$_2$)-NH_2; To form a prostate cancer-selective prodrug (substrate for PSA), the oligopeptide sequence QLKSSH is attached to the terminal valine residue [153].

An octanoyl group -although being more sterically demanding, can still fit into the tunnel like cavity normally holding the angeloyl residue if it is in the correct α-orientation, while in the β-orientation it will interfere with the C-2β-octanoyl residue so that both will probably not fit into the rather narrow cleft. A C-3-keto group cannot contribute to binding within the angeloyl-pocket and, consistently, the activity decreases by factor 70 [145].

The importance of the C-10β-O-acetyl group was also demonstrated since deacetylation at this position led to a 40-fold decrease of activity [145]. It can now be seen from the binding site structure, that the actetyl group fits into a cleft "behind" the

phenylalanine F834 side chain where it is able to contribute some hydrophobic interactions which would not be the case with a free hydroxy group.

It has been demonstrated that the lactone structure as such is not essential for activity in case of **TG**, since the corresponding lactols showed only slightly reduced activity in comparison with **TG** itself [151]. Indeed, the lactone carbonyl does not contribute to binding and -apart from making the molecule slightly more polar- the α- or β-oriented lactol OH groups will not cause any direct disturbances. The acetate of the α-lactol, however, was shown to be a somewhat weaker ligand, while the β-lactol ethylacetal showed no significant reduction of potency. In line with similar results obtained for 11,12α-configured acetonides and orthoformiates [151], this observation indicated that the α-side of **TG** is more intimately involved in receptor binding, which is confirmed by structure of **TG** bound to SERCA where it becomes evident that more bulky α-oriented groups will disrupt binding for steric reasons, while β-oriented residues will not lead to any unfavourable contacts with the binding site's surface. The finding that the C-7β-O and C-11α-O monoacetates were both equipotent and only about 3 times less active than **TG** [145] at first sight appears somewhat surprising since the latter substitution on the α-side might be expected to produce a steric clash for the reasons pointed out above for the C-12-lactol acetate. However, an acetate group at C-11α could orient towards the cavity that hosts the C-8-O-butanoyl residue. Also in case of the C-7β-acteate, the acetate group may be expected to lead only to minor disturbances of binding since it can adopt an orientation where it points towards the opening of the binding cleft where some less strong interaction with the Phe834 side chain may occur to explain the slight decrease in affinity. Consequently, the diacetate at both, the C-11 and C-7 positions was still quite active (less active than **TG** by a factor of 15 [145]) which, considering the additivity of the two relatively weak steric effects, is easily explained.

However, even in the light of the binding site structure, one of the previous observations cannot readily be explained. The absence of a measurable activation of mast cells by the 7β-11β epoxy derivative [150] would not be expected since it should fit into the SERCA's binding site without any major problems. In contrast, (and in agreement with the above-mentioned result for the α-lactol derivatives from the same study [151]) an 11α,12α-epoxide of Thapsigargicin was only 5 times less active in inhibiting SERCA in a microsomal preparation from rabbit longitudinal muscle than the natural compound [151] and thus behaved as would be expected in view of the binding site's structural requirements. However, since different assay systems were used, it is not clear whether these results can be directly compared. The former finding was obtained with whole cells and it is possible that metabolic inactivation or covalent reaction of the epoxide with cellular structures led to loss of activity in case of the 7,11-epoxide.

The activity of **TG** is of interest not only with respect to its wide use as a molecular tool to elucidate intracellular processes related to Ca-signalling [128, 129], but also with respect to its potential usefulness as a therapeutic agent. Apart from the traditional (and now obsolete) use as irritant, the strong pro-apoptotic activity would indicate a potential as anti-cancer agent [152, 153]. Moreover, the fact that modulation of Ca-levels can influence the pathophysiology of cystic fibrosis and related gene deficiencies has led to intense research concerning the usefulness of **TG** in treatment of such severe innate diseases [154, 155]. As is generally the case with highly bioactive natural products,

however, **TG**'s effects are not restricted to the desirable activity towards the target cells involved but displays a very low degree of selectivity. While it may be quite difficult (if not impossible) to taylor the structure into selective ligands for the SERCA of particular cell types, it appears more straightforward to obtain derivatives with selective action by masking the structure in such a way that it is selectively liberated in the tissue that is to be affected. In this respect, recent attempts to target androgen-independent prostate cancer with **TG**-related pro-drugs must be mentioned, which have led to the design of several compounds from which **TG**-derivatives equipotent to **TG** itself can be liberated selectively in prostate cancer tissue [153]. In these derivatives (e.g. structure **17** in table 4), 8-debutanoyl **TG** is esterified at C-8-OH with a long-chain aliphatic ω-aminocarboxylic acid linker whose amino group is attached to an oligopeptide which is known to be specifically cleaved off by a prostate cancer - specific protease known as "prostate specific antigen" (PSA). Such prodrugs turned out to be efficient in treatment of prostate cancer in mice [156] and are currently in further development.

In effect, it can be summarized that the various SAR studies on **TG** derivatives were quite useful in probing the SERCA's binding site before its actual structure was known. Whatsoever, none of the semi-synthetic derivatives was found more active than natural **TG**. Lead optimization with respect to potency would by all means have been a difficult goal in case of such a strong ligand. The unsurpassed potency of **TG** as a SERCA inhibitor may serve as an excellent example for the efficiency by which nature optimizes its drug leads with respect to a particular target. A further example, in which sesquiterpenoids from two quite different structural groups have been optimised to bind to a common target with high specificity and affinity will be introduced in the next section.

GABA-ANTAGONISTIC SESQUITERPENE LACTONES

Sesquiterpenoids from two different structural groups, possessing either the picrotoxane or the *seco*-prezizaane (*seco-allo*-cedrane [157]) skeleton have long been known as the toxic principles in a number of plant species. The former are found mainly in the Menispermaceae and Coriariaceae families while the latter are restricted to *Illicium* species (Illiciaceae) (review see [3]). A further group of terpenoids, the picrodendranes, toxic principles of *Picrodendron* species (Picrodendraceae), showing the same mechanism of action and a carbon skeleton structurally closely related to the picrotoxanes, will be included in this discussion although they are most likely of diterpenoid origin.

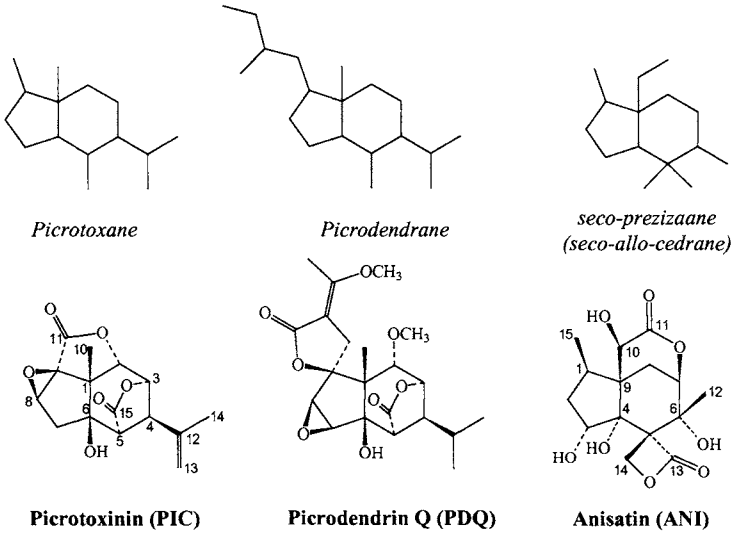


Fig. (24). Structures of the most prominent GABA-antagonistic terpenoids and their carbon skeleta.

Mechanism of action

Representatives of all three types of terpenoid lactones cause lethal convulsions by allosterically inhibiting the action of γ-aminobutyric acid (GABA) at GABA-receptor coupled chloride channels within the central nervous system. The prototype compounds from the two sesquiterpenoid groups, picrotoxinin (**PIC**) and anisatin (**ANI**) are highly potent neurotoxins both lethally active at LD_{50} of about 1 and 3 mg/kg (mouse, i.p.), respectively, and have been shown to possess the same mechanism of action [158-163]. The most potent picrodendrane derivative, picrodendrin Q (**PDQ**) shows even higher affinity to rat brain $GABA_A$ receptors to which it binds about 40 times more strongly than picrotoxinin [164, 165].

The mode by which the allosteric inhibition occurs in the case of picrotoxinin was formerly believed to result from a physical blockade of the Cl^--channel pore by the compound itself [166, 167] but it was more recently shown that binding of the compound to the channel protein rather leads to stabilization of its closed state [168-170]. A number of studies have provided good evidence that picrotoxinin and anisatin as

well as picrodendrin Q do not only share the general mechanism of action but actually do so by interacting with a common binding site at the channel protein. These findings, together with recent results where some of the *seco*-prezizaane type compounds isolated by us interestingly showed selective toxicity towards insects [162], have prompted us to establish common structure-activity relationships for all three groups of compounds which will be described below.

Structure-activity relationships

Studies on structure-activity relationships for the picrotoxane type convulsants were initiated by Jarboe et al. in 1968 [171], even before the exact mechanism of action was known. The convulsant and lethal activity of 27 natural and synthetic picrotoxanes was compared and this study led to the identification of the major structural determinants for *in-vivo* toxicity within this group, which may be summarized as follows: The C-5(15)-3 fused γ-lactone ring must be *cis*-oriented with respect to the carbocyclic ring junction (or *trans* with respect to the methyl group C-10 and the OH-group at C-6). The alkenyl side chain at C-4 must be oriented *trans* with respect to the mentioned lactone ring, a *cis*-oriented side chain leading to complete loss of activity. Hydrogenation of the $\Delta^{12,13}$ double bond still yields a highly toxic derivative while hydration to a tertiary alcohol leads to 30-50-fold reduction of activity and shift of the double bond to $\Delta^{4(12)}$ renders the compound inactive. The OH group at C-6 should be unsubstituted, since acetylation or formation of a 6,12-anhydro derivative leads to loss of activity. These findings, combined with each other, represented the first biophore model for this type of convulsant which could lateron be confirmed and extended to other convulsant terpenoids by the use of computer-assisted molecular modelling [163, 165, 172, 173], see below.

Three computer-assisted studies of the 1980's and early 90's [166, 174, 175] were aimed at refinement of the biophore and elucidation of the properties of the binding site for convulsants of the picrotoxane type together with different series of (mainly synthetic) compounds acting by the same mechanism. The investigations by Klunk et al. in 1983 [166] tried to establish a common SAR for picrotoxanes and synthetic γ-butyrolactone-type convulsants. Although the study arrived at very plausible results, both initial assumptions underlying the study, namely, a common binding site for the two groups of compounds under study as well as the mechanism of a physical channel blockade were later disproved [168-170, 176]. The second study [174] in which a very diverse set of $GABA_A$ antagonistic convulsants of various structural types was investigated, arrived at a common pharmacophore- or binding site model. It had, however, already been shown at that time that distinct binding sites for convulsant barbiturates and picrotoxanes exist [177]. Furthermore, anisatin **ANI**, also included in this work, was incorporated in the model in the form of the non-natural (+)-enantiomer, since the absolute stereochemistry had not been established at that time. It was only in the following year 1990 that the absolute stereochemistry of natural (-)-anisatin was established [178], and only then the way was open to compare the 3D structures of *seco*-prezizaane type convulsants with those of the picrotoxinin group (see below). A 1993 CoMFA study in search for a common QSAR including picrotoxinin and picrotin (12,13-dihydro-12-hydroxypicrotoxinin) together with 38 synthetic GABA-antagonistic insecticides of diverse chemical structure resulted in a 3D-QSAR of respectable

statistical quality [175]. Alignment of the picrotoxanes with the other compounds under study was based on superposition of the 9(11),2 lactone ring of the former with structure elements of the latter that were assumed to be important for activity. This lactone ring, however, is clearly not essential for GABA-antagonistic activity since, e.g. tutin and coriamyrtin (structures **15** and **17** in Fig (**27**)) as well as the picrodendrins (e.g. picrodendrin Q **PDQ**) are effective convulsants as well as insecticides and do not possess this structure element but all show the basic pharmacophore established by Jarboe as early as 1968 (see above). Unfortunately, none of the other highly potent picrotoxane derivatives was included in this work so that it appears doubtful whether the spatial regions identified as important for activity by this CoMFA analysis are of any relevance with respect to the binding of picrotoxanes at the Cl^--channel protein.

Several studies including CoMFA analyses for picrotoxanes/picrodendranes [165], in combination with synthetic insecticides of diverse structural groups such as cyclodienes, bicycloorthobenzoates and others [172, 173], have more recently been conducted by the group of Y. Ozoe. The results led to establishment of a simplified receptor model (Fig. (**25**)) for diverse types of compounds interacting with the GABA-sensitive Cl^- channel of

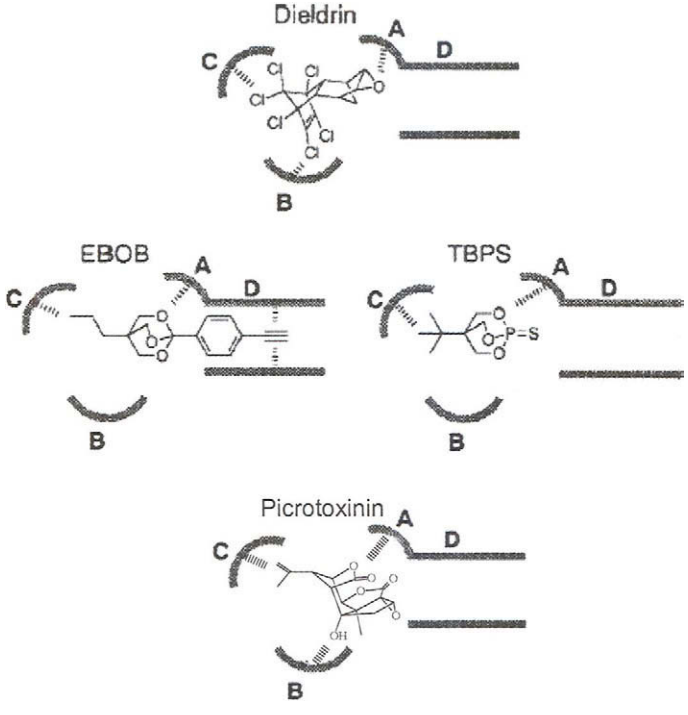


Fig. (25). Hypothetical binding site model for GABA-antagonistic insecticides established by Ozoe et al, reproduced with permission from the author [165]. The picture (bottom) showing picrotoxinin was added by the author of the present publication. Note that PIC was originally assumed to bind in an opposite mode, i.e. the epoxy group was assigned subsite A and the lactone structure subsite B. The picture presented here corresponds to the binding mode assumed on grounds of the current knowledge [163, 182].

insects [173]. It was concluded from the results of CoMFA studies for the separate classes of antagonists, that these structurally diverse compounds bind to an identical site but in different or overlapping orientations. This binding region was postulated to consist of four regions of interaction (Fig. (**25**), according to [173]). Sub-site **A** supposedly accepts an electronegative part of the ligands. On both sides of this region, there may be spaces that accomodate hydrophobic and electronegative parts of ligands (**C** and **D**). In addition, the existence of a further subsite **B** was inferred and it was concluded that compounds capable of interacting with two of the four sub-sites should act as antagonists; the electronegative interaction in site **A** was speculated to play the major role in determining the interactions. Transferring the requirements for convulsant activity derived by Jarboe (see above, [171]) to the insect, it is straightforward to assume that site **A** should interact with the picrotoxanes' lactone oxygens so that the isopropenyl group would extend into region **C**. However, it is known that insect and mammalian receptors are similar but show some differences in their pharmacological responses and this was corroborated by the mentioned studies which led to somehwhat different CoMFA results for housefly and rat receptors. In consequence, some hints with respect to insect-selectivity of some picrodendrane-type compounds could be derived from this information [173].

Our own work on structure-activity relationships of convulsant sesquiterpene lactones was initiated after isolation of a variety of *seco*-prezizaane type compounds from *Illicium floridanum* and *I. parviflorum*, both endemic to the southern United States [157, 179-181]. A study on the *in-vivo* toxicity of several of these compounds [182] showed that anisatin, major constituent in leaves and fruits of *I. floridanum* (as in those of other previously studied toxic *Illicium* (=star anise) species, review see [3]), was apparently the only derivative in the tested series which was toxic to mammals, although further compounds, such as pseudoanisatin (**PSA**, structure **4** in Fig. (**27**)) were accumulated in the plant in similarly high amounts [179, 180]. In the light of reports that indicated an identical mechanism of action and -possibly- the existence of a common binding site for anisatin and picrotoxinin [156-161] the question arose which structural features in these compounds were similar enough to warrant an identical mode of molecular recognition at the channel protein.

Comparison of the structures of **PIC** and **ANI** shows that alignment of the -apparently similar- carbon skeleta does not lead to a satisfactory alignment with respect to potential interaction points which should -on grounds of the high degree of functionalization- in both cases be represented by hydrogen bond donor- and/or acceptor sites. Comparison of the distances and angles between all H-bond acceptor sites (oxygen atoms) in the 3D-structures indeed led to the finding that a triangle enclosing essentially identical distances and angles exists between the carbonyl and ring oxygens of the 5(15),3-lactone and the epoxy oxygen atom of **PIC** and between the δ-lactone carbonyl- and ring oxygens and the oxetane (β-lactone) ring oxygen of **ANI**. Fig. (**26**) shows the two 3D structures aligned with respect to these atoms (labelled **A**, **B** and **C**, respectively) and it becomes clear that the two molecules possess identical possibilities to accept H-bonds at these positions. Furthermore, it was found that the carbons in the lactone ring as well as the adjacent isopropenyl side chain of **PIC** in this situation are quite well aligned with the fragment consisting of C-15, 1, 9, 10, 8 and 7 in **ANI** so that similar

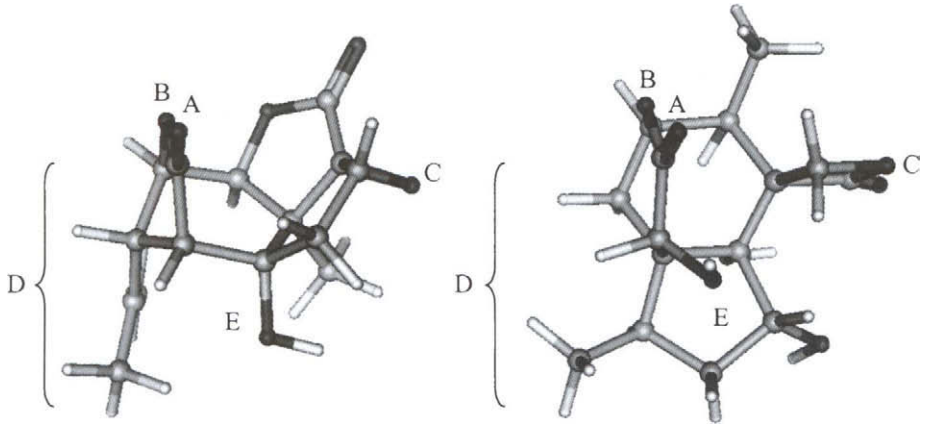


Fig. (26). Molecular models of **PIC** (left) and **ANI** (right) aligned with respect to the oxygen atoms labelled **A, B** and **C**. Further regions contributing to the biophore model described in the text are labelled **D** and **E**. Modified, according to [182].

hydrophobic interactions in these molecule parts can also be expected (region **D** in Fig. (**26**). The OH group at C-6 of **PIC** may then additionally serve as a H-bond donor interacting with an acceptor that can equally well be reached by the C-10-OH of **ANI** (**E** in Fig (**26**)). Thus, a common biophore was established for anisatin and picrotoxinin [182]. This common pattern was found in 3D-molecular models for a wide variety of highly active picrotoxanes and *seco*-prezizaanes while all of the less toxic or inactive compounds available showed deviations [182]. Most importantly, this first common SAR model for convulsant activity of the two types of sesquiterpenoids was in complete agreement with the picrotoxane biophore established by Jarboe et al. 30 years earlier [171] and with the findings of Ozoe et al. [165, 172, 173]. It must not remain unmentioned that there was one particular observation which could not readily be explained by our hypothesis. 2α-hydroxyneoanisatin (structure **3** in Fig (**27**)), in spite of its almost identical structure with anisatin, did not show any convulsant toxicity which could either be due to poor receptor affinity, or to pharmacokinetic reasons [182].

Since the model so far was based on *in-vivo* data obtained with whole mice, where the actual receptor affinities might to some extent have been obscured by effects of penetration through the blood-brain barrier, distribution, metabolism and elimination, it appeared desirable to confirm our biophore model by obtaining receptor binding data. Furthermore it was of interest to investigate the compounds' affinity to insect GABA receptors which are known to be similar to those of mammals but nevertheless show distinct pharmacological properties [183, 184]. To this end, 13 *seco*-prezizaane type compounds (structures see Fig. (**27**)) were subjected to binding studies towards rat (*Rattus norvegicus*) and housefly (*Musca domestica*) GABA receptors in a cooperation with the group of Y. Ozoe [162] who had already investigated SAR for picrotoxanes and picrodendranes (see above). The binding data obtained by displacement of the selective GABA antagonist [^{3}H]-4'-ethynyl-4-*n*-propylbicycloorthobenzoate (EBOB) from the rat receptor confirmed our previous model and added as a further piece of knowledge the

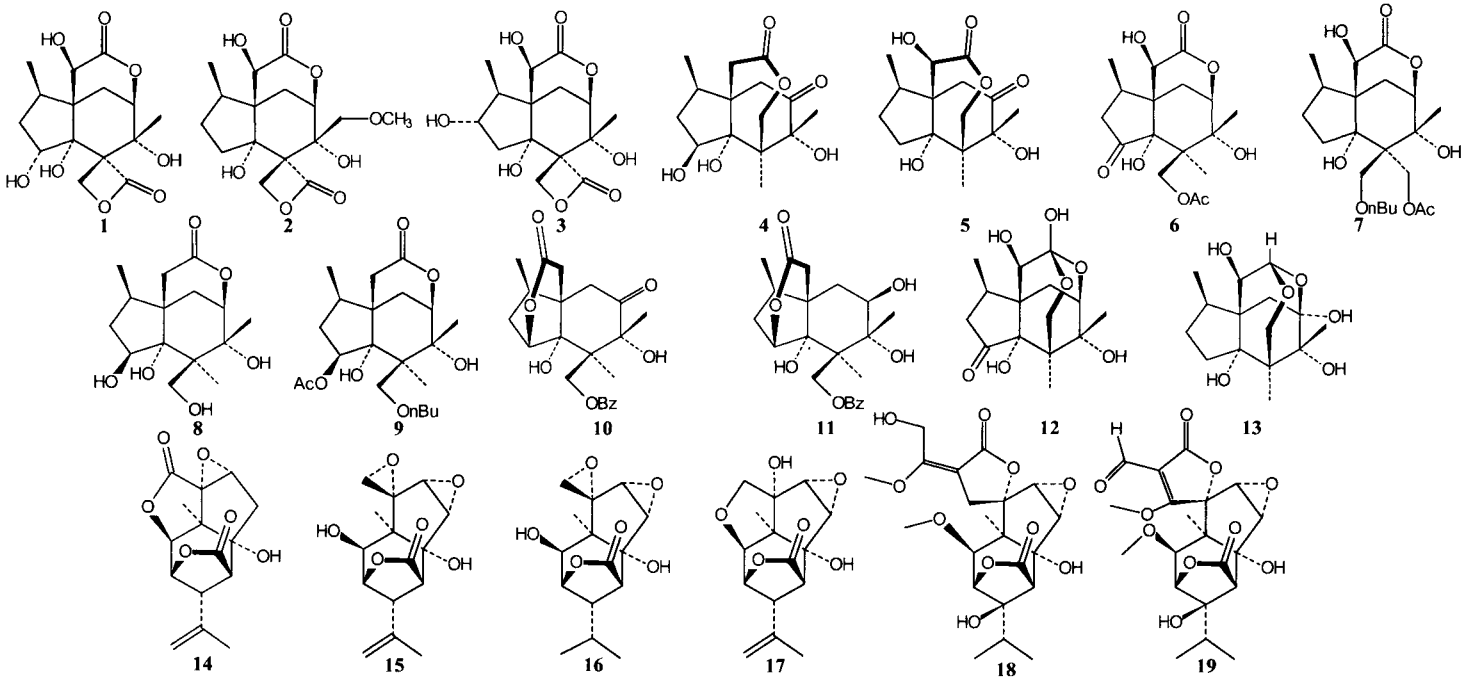


Fig. (27). (continued)

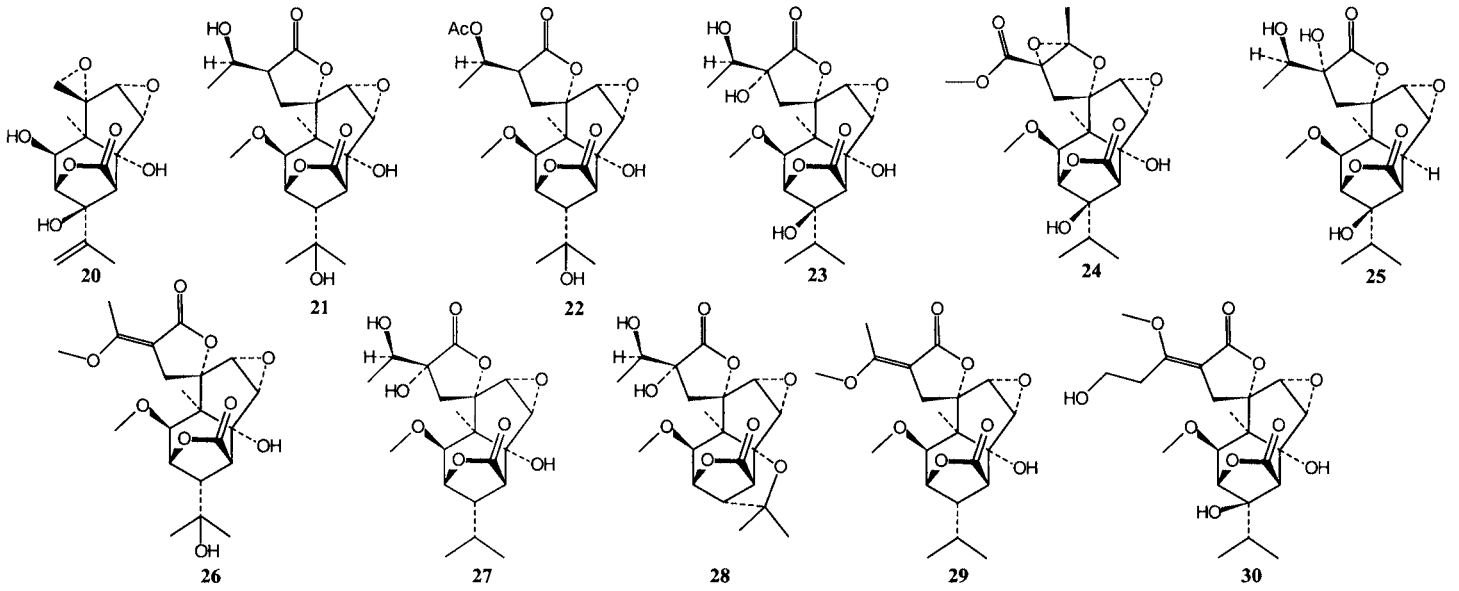


Fig. (27). Structures of the 30 convulsant/insecticidal terpenoids of the seco-prezizaane (**1-13**) and picrotoxane/picrodendrane types (**14-30**) whose binding data to rat and housefly GABA receptors were included in the *Quasar* receptor-surface modelling study [163].

finding that the lack of toxicity in case of 2-hydroxyneoanisatin (see above) was due to lack of receptor affinity and not primarily to ADME effects. As the major result of the binding experiments, however, some of the compounds showed selective affinity towards the insect receptors. Fig. (**28**) shows a plot of the percentage of [^{3}H]-EBOB displaced from the housefly vs. the rat receptor preparations, measured for the 13 compounds at 10 μM concentration. Most prominently, in contrast to the strong convulsants anisatin and veranisatin A (**VER**) which were found to bind equally strong to rat and housefly preparations, pseudoanisatin (**PSA**) was almost as potent as **VER** and **ANI** towards the housefly receptor while showing only negligible affinity to the mammalian one. Another derivative, parviflorolide (**5** in Fig (**27**)) showed almost equal potency to **PSA**. These compounds thus deserve attention with respect to their potential as leads to new selective insecticides [162].

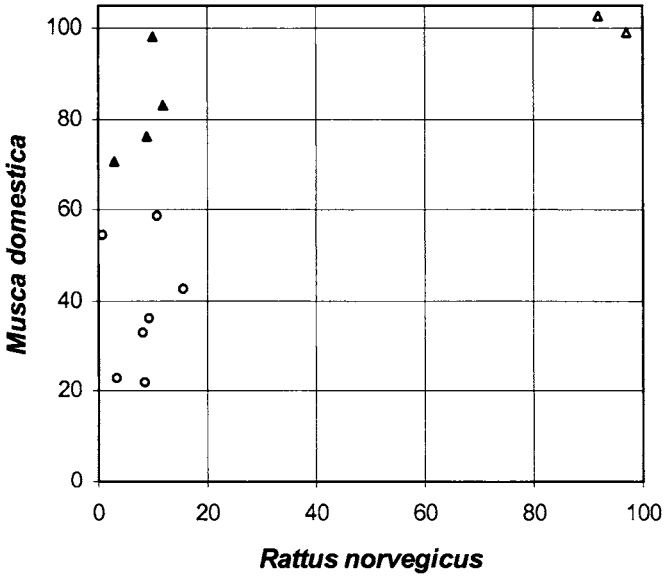


Fig. (28). Displacement of [^{3}H]-EBOB from the binding site at rat brain and housefly head GABA receptor-coupled Cl^- channels. Data were taken from [162]. Open circles: Compounds showing low affinity towards both receptors, open triangles: compounds with high affinity towards both receptors; filled triangles: compounds showing medium to high affinity towards the insect and negligible activity towards the mammalian receptor.

As a further result from this study, a CoMFA [113] model based on our previous biophore alignment could be generated for the housefly binding data in combination with data for a variety of picrotoxanes and picrodendranes obtained earlier. The resulting q^2 value, although quite low, was significant (0.45), and the fact that a model could be obtained with this diverse set of compounds further confirmed that the chosen alignment was reasonable [162]. A QSAR model for the rat receptor data was not attempted at that time since -in contrast with the fly data- only few active compounds (i.e. ligands for

which an IC_{50} could be determined) were present and most compounds were only very weakly active. Whatsoever, it became clear from the data that apparently the binding site at the mammalian receptor is more selective so that fewer compounds meet the requirements for tight binding. It was therefore of high interest to investigate further these differences in the two receptors' binding preferences and thereby to shed light on the reasons for the observed insect selectivity.

To this end, we recently generated models of the binding sites for convulsant terpenoids at both the housefly and the rat receptor [162] using the program *Quasar* [185, 186]. In this approach, a model of the receptor surface is obtained, based on the 3D structures of the aligned ligands and their experimental binding energies. A *Quasar* model consists of a pocket formed by virtual particles with atomistic properties such as hydrogen bond donating/accepting, salt bridge donating/accepting, hydrophobic etc.. Apart from interactions with such pseudo-atoms, also the induced fit of the receptor as well as multiple conformations of the ligands are used to model the binding site (For a detailed description of the *Quasar* approach see [185, 186].) The compounds are divided into a training set and a test set before the modelling process is initiated. The structures and binding data of the training set are used to construct the model. Internal cross validation is used to validate model quality in the course of model building. Finally, the resulting model is validated by predicting the binding energies for the test set compounds.

Thirty compounds (13 *seco*-prezizaanes and 17 picrotoxanes/ picrodendranes, see Fig. (**27**)) were included in the modelling study [163]. The test set consisted of 8 and 7 compounds for the housefly and rat data, respectively, and the binding energies were predicted by the final models with good quality (see Fig. (**29**)), indicating that these

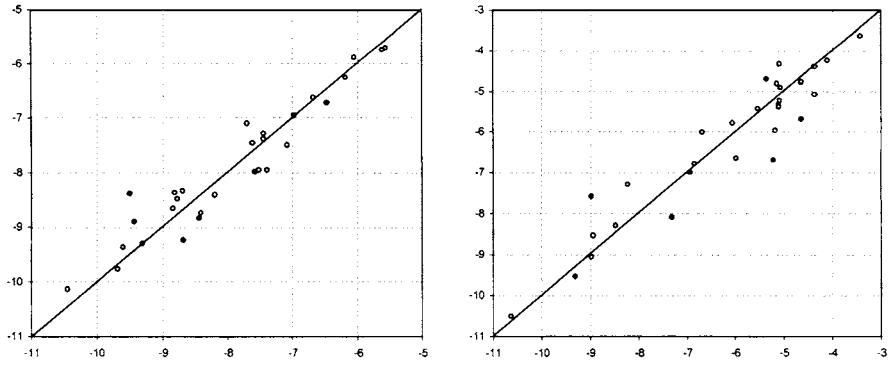


Fig. (29). Predicted vs. experimental binding energies (ΔG^0, kcal mol^{-1}) for the receptor surface models of the housefly (left) and rat (right) binding sites. Open circles represent compounds of the training set, filled dots those of the test set. For details see [163].

QSAR models can be used to predict the activity of untested compounds with some accuracy. The receptor surface models obtained are depicted in Fig. (**30**), with the structures of veranisatin A (**VER**, **2**, in Fig. (**27**), non-selective) and pseudoanisatin (**PSA**, insect-selective) inside the binding pocket. Both models show a common interaction region for the lactone carbonyl and ring oxygens consisting of salt-bridge accepting and/or hydrogen bond-donoating particles (SB+, red, DON, green; upper left corner) as well as a hydrophobic region (corresponding to region **D** in Fig. (**26**)) hosting

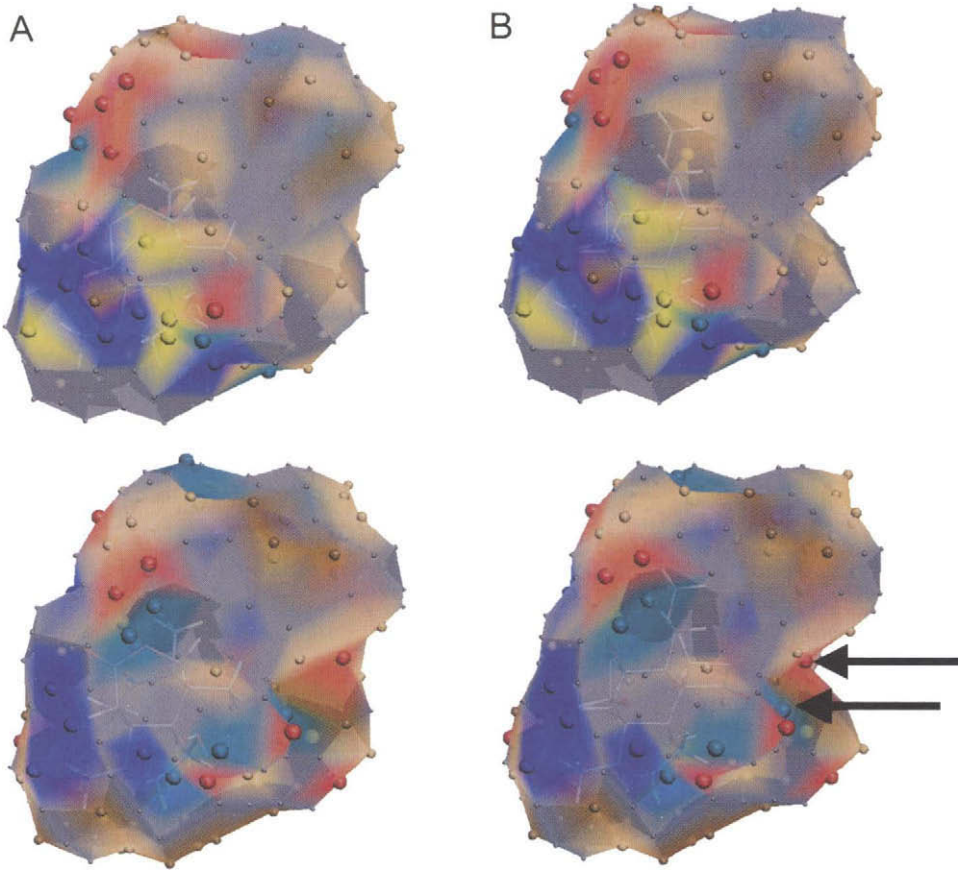


Fig.(30). Pseudoanisatin (**PSA**, insect-selective; **A**) and veranisatin A (**VER**, non-selective, **B**) shown inside the fly (top) and rat (bottom) receptor surface models (from [163]).
Receptor particles: Blue: Salt bridge negative, SB-; Red: Salt bridge positive, SB+, Yellow: Hydrogen bond acceptor, Acc; Green: Hydrogen bond donor (Don); Grey (small spheres): Hydrophobic, neutral; Yellow-brown (small spheres): Hydrophobic positive (Hy-); Red-brown (small spheres): Hydrophobic, negative (Hy-).
Note the interaction of the β-lactone ring oxygen of **VER** with the SB+ and Don particles on the right side of the rat receptor (arrows). Especially the SB+ particle is moved towards the lactone oxygen as compared to compound **PSA**.

the hydrophobic region around the C-15 methyl group (grey, brown, lower left, major part not visible from this angle; this region holds part of the isoprop(en)yl side chain of the picrotoxanes/picrodendranes).

The major difference between the receptor models is observed in the region marked by arrows, where the rat receptor shows an SB+ particle (red) and an H-bond-donating particle (green), while the fly model in this region consists exclusively of hydrophobic pseudo-atoms. These particles can interact with the β-lactone oxygen of **VER** and **ANI** (and corresponding structure elements in the picrotoxane/picrodendrane group, e.g. the epoxy group in **PIC**) while **PSA** does not have any structure elements with negative partial charge/ H-bond-accepting properties in this region. This difference seems to be the major point in explaining the inactivity of most of the *seco*-prezizaanes towards mammalian $GABA_A$-receptors. To explain why pseudoanisatin is highly potent towards the fly receptor, individual contributions of functional groups to the total binding energy were compared. The results of a functional group analysis performed by *Quasar* showed that the major discriminants explaining the difference in binding energy of **PSA** vs. the rat and fly receptors ($\delta\Delta G^0$=-3.6 kcal/mol) are the interactions of the ε-lactone and the C-7-oxo carbonyl groups, as well as the OH group at C-3β. The two carbonyl groups' interaction with the fly receptor is more favourable by -1.1 and -1.5 kcal/mol, respectively, that of the OH group by -1.1 kcal. Furthermore it was observed that the δ-lactone structure of **VER** interacted more strongly with the rat receptor, while the interactions of ε-lactone in **PSA** were more favourable in the fly model. We thus arrived at the conclusion that an insect-selective compound should not possess any electronegative structure elements in the region corresponding to **C** in Fig. (**26**) (equivalent to the oxetane oxygen in **ANI**, and the epoxide or spiro-γ-lactone oxygens in **PIC** and **PDQ**, respectively), that it should not have a δ-lactone structure corresponding to that of **ANI** and that it should possess an additional H-bond accepting moiety in the region occupied by the keto carbonyl of **PSA** [163].

These results in hands, the question arises whether the differences in such theoretical models based solely on the potential ligands' structure and binding data can also be used to obtain information on the real binding site in the channel proteins.

No three-dimensional structure for a GABA receptor-coupled chloride channel is currently available, but a number of amino acid sequences from different organisms are known. These channels belong to the superfamily of cys-loop neurotransmitter receptors that also includes nicotinic acetylcholine receptors (nAChR, whose 3D structure has now been established [187]), stychnine-sensitive glycine receptors, and serotonin receptors of subclass $5HT_3$ [183]. The general structure of such ion channels is constituted by pentamers of very similar protein subunits each of which possesses four transmembrane helical segments. One of these helices from each of the five monomers (M2) forms the lining of the channel pore [183, 187]. Among other GABA-sensitive chloride channel proteins, the full sequence of the *Rdl* protein from the fruit fly (*Drosophila melanogaster*) is known [188] and may serve as a template for the housefly protein (it shows 100% sequence identity with the corresponding *Musca domestica* protein in the area of interest). Most importantly, it has long been known that a single amino acid exchange (A302→S) confers the insects' resistance to the GABA-antagonistic insecticide dieldrin (Rdl for "*R*esistance to *diel*drin") as well as to **PIC** [183]. It can thus be expected that the binding site is located in the vicinity of this residue which is also

known to lie close to the cytoplasmic end of the M2 helix [183]. It is, however, not possible on grounds of the available experimental data to identify residues from the other membrane-spanning helices or from the loop regions connecting the helices on the cytoplasmic end are close to Ala 302 that may be involved in forming the binding site. Based on the recently published structure of the *Torpedo marmorata* (Electric Ray) nAChR, a homology model of the insect chloride channel was therefore generated which is depicted in Fig. (**31**) [Schmidt, T.J., in preparation].

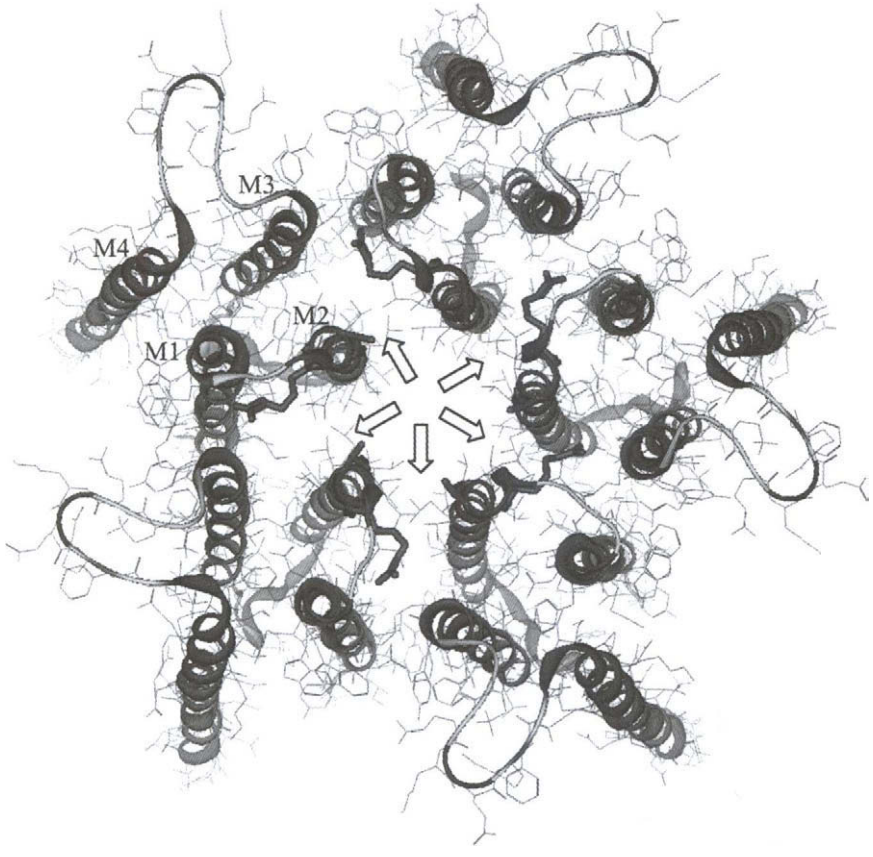


Fig. (31). Hypothetical 3D model of the membrane-spanning part of the insect GABA-gated chloride channel.
Topography of the five subunits viewed along the channel axis and from the cytosolic side. Helices are labelled M1-M4. Based on the presented evidence, the binding site for non-competitive antagonists is hypothesized to involve residues A302 and R300 (atoms represented in stick mode) from two neighboring subunits' M2 helices (arrowheads).

In spite of a relatively low degree of overall sequence identity (15-18% identical amino acids with the various subunits of nAChR), the sequence alignment led to correct assignment of the important M2 helix in *Rdl* to that of nAChR, so that construction of a

3D model seemed feasible. In this model, the important alanine residue A302 was correctly placed in its presumable position [183] close to the cytoplasmic end of helix M2. Amino acid residues possessing atoms within a sphere of radius 9Å (maximum diameter of a **PIC** molecule) from the methyl group of A302 were considered potential contributors to the binding site in question. The fragment thus obtained comprised residues 297-309 (TPARVALGVTTVL) and residues 297, 299-300, 302-303 and 306 (T-AR-AL--T---) from the M2 helix of the neighboring subunit. Involvement of arginine R300 indicates that the protonated side chain of this residue actually represents the salt-bridge-acceptor or hydrogen bond donor interacting with the lactone carbonyl and ring oxygens of the terpenoids. (**A** and **B** in Fig. (**26**), SB+ particles shown as red spheres in upper left corner of the pseudoreceptor models, Fig. (**30**)) It can be assumed that the opening mechanism of the GABA-gated Cl^- channel is similar to that of the nAChR sodium channel which has been described [187] to lead to a 15° rotation of the M2 helix in clockwise direction which lets the helix collapse back towards the inside of the cavity formed by the other three helices. Thereby, the residues blocking the channel pore move apart leaving an opening wide enough to allow a hydrated ion to pass [187]. The location of the mentioned amino acid residues in the 3D model (see Fig. (**31**)) would indicate that the binding site lies at the interface between the M2 helices of two neighboring subunits, possibly involving A302 of one subunit and R300 of the neighboring one (region marked by arrowheads in Fig. (**31**)). This region provides enough space to accomodate one of the antagonist molecules and it can be expected that this would effectively hinder the movement of the M2 helices required to open the channel.

Docking studies with the goal of modelling the binding site in an atomistic context are currently under way. Alignment of the known sequences [188] of mammalian $GABA_A$-receptor subunits may then reveal which residues are ultimately responsible for the selectivity of some of these compounds which will probably open the way to design of insecticides with even higher selectivity.

It may finally be mentioned that some of the non-toxic *seco*-prezizaane compounds have recently shown high activity against nematodes [189]. Many of these animals are pathogenic parasites affecting man and livestock [94]. GABA receptors are the known target of nematocidal compounds such as the avermectins, which are used to treat nematode infections [190]. One of the *seco*-prezizaanes, most interestingly of the group that shows significant binding neither to mammalian nor to insect GABA-receptors, displayed in-vitro nematocidal activity higher than that of abamectin [189]. Although the tested nematodes were not pathogenic but free-living nematodes isolated from soil, further studies on this compound's potential as a lead to new nematocidal agents appear well justified.

CONCLUDING REMARKS

As shown in the first section on structural diversity, STLs have been optimised during evolution towards the ultimate goal of high versatility as chemical defence tools. Due to their relatively non-specific mechanism of action, most of them are not suitable to serve as drugs. However, as shown in the section on alkylant STLs, studies on structure-reactivity- and structure-activity relationships yield valuable information with respect to potential selectivity of particular structures vs particular targets. Along with the observation that the vast majority of these compounds are quite "drug-like" with respect to their physicochemical properties, this indicates that they may well serve as lead structures for new anti tumour, anti-inflammatory and anti-protozoal drugs. A straightforward way to achieve higher specificity will certainly be to apply structural changes that direct these compounds into the target tissue of interest, or such that are released by metabolic reactions specific to the target cells, in a similar way as recently practised with thapsigargin.

Considering that only a small fraction of the 20000 Asteraceae species as the major pool of STL structures have been subject to phytochemical studies, and that many plants of other families may contain such compounds, the knowledge on their structural diversity as described in the first section will increase further in the future. At the same time, their biological activity and potential use is a field of very active research in many groups and it appears that a lot of interesting findings in this respect are yet to be made.

Apart from the majority of STLs acting by more or less unspecific alkylation of proteins, four groups of very specifically-acting STLs were introduced. Artemisinin derivatives are used in malaria therapy with outstanding success. Development and approval of new analogues requires understanding of the structure-activity relationships in this group which has to remain an active field of research. New aspects will certainly arise from the recent discovery of a likely malaria-specific target. It is a very exciting finding that this target is closely related to the target of thapsigargin, i.e. the mammalian SERCA protein, and new studies with respect to common structure-activity relationships for artemisininoids and thapsigargin derivatives will certainly emerge.

In the case of thapsigargin, which has long been used as a molecular tool to study intracellular calcium signalling, also new aspects can be expected with respect to tumour therapy and treatment of gene deficiencies whose consequences may be modulated by interference with intracellular calcium homeostasis.

Finally, in case of the GABA-antagonistic sesquiterpenoids treated in the last section, interesting new insights into the structure-activity- and especially structure-selectivity relationships were introduced. Recently discovered selective toxicity towards insects as well as helminths point towards the potential of such compounds as leads to new insecticides and antiparasitic drugs.

Sesquiterpene lactones thus certainly deserve to remain an active field of natural products research for many years.

This chapter attempted an overview on the current state of knowledge on structure-activity relationships of this interesting group of plant constituents. Since it is certainly not possible to cover such a wide area exhaustively, and there are probably many studies in the field which were not cited, the author apologizes to all those who might feel their

work not duly mentioned on these pages. At the same time I shall be grateful to all who will send me their comments on this work which will help to inspire further research in this interesting field (schmidtt@uni-duesseldorf.de).

LIST OF ABBREVIATIONS

STL: Sesquiterpene lactone
PCA: Principal component analysis
DNP: Dictionary of Natural Products
PRS: Potentially reactive substructure
GSH: Glutathione
CP: Cyclopentenone
MGL: α-Methylene-γ-lactone
ADME: Absorption, distribution, metabolism and elimination
QSAR: Quantitative structure-activity relationships
r^2: Squared correlation coefficient
q^2: Squared correlation coefficient after cross validation
LOO: Leave-one-out cross validation
PLS: Partial least squares regression
GA-PLS: Genetic algorithm – partial least squares
Q_frASA: Partial charge (Q) based fractional accessible surface area descriptor
CoMFA: Comparative molecular field analysis
ACD: Allergic contact dermatitis
SERCA: Sarco-endoplasmatic reticulum Ca-ATPase
GABA: γ-aminobutryric acid
Quasar: Quasi-atomistic receptor surface modelling
nAChR: Nicotinic acetylcholine receptor

ACKNOWLEDGMENTS

Most gratefully, the author wishes to acknowledge the contributions of many collaborators to much of the presented work in which he himself was involved. It is due to mention the cooperations with I. Merfort and coworkers, who performed the experimental work on NF-κB inhibition, J. Heilmann and coworkers who conducted the KB cytotoxicity and gene expression studies, R. Brun who performed the anti-trypanosomal tests and E. Okuyama and Y. Ozoe who conducted in-vivo- and receptor binding studies, respectively, with the *seco*-prezizaanes. Big thanks are due to the NMR service at HHU. Furthermore, help of E. Müller in some of the reactivity studies is gratefully acknowledged.
Finally, thanks to my wife for being very patient and keeping my back clear all the time and to my kids for not complaining that many of their "evening lectures" had to be cancelled during the last few weeks of preparing this manuscript.

REFERENCES

[1] Ivie, G.W.; Witzel, D.A. In: *Handbook of Natural Toxins*; Keeler, R.F.; Tu, A.T., Eds.; Marcel Dekker: New York, **1983**; Vol. *1*, chapter 17, pp. 543-584.

[2] Picman, A.K. Biochem. Syst. Ecol. 1986, 14, 255-281.

[3] Schmidt, T.J.; *Current Org. Chem.* **1999**, *3*, 577-608.

[4] The Dictionary of Natural Products on CDROM, v. 10:2; Buckingham, J; Hodgson, A.M.; Walford, S. P., Eds.; Chapman & Hall/CRC: Boca Raton, **2002**.

[5] Seaman, F.C. *Bot. Rev.* **1982**, *48*, 121-595.

[6] Raven, P. H.; Evert, R. F.; Eichhorn, S. E. *Biology of Plants, Sixth Edition*; Worth Publishers: New York, **1999**.

[7] Weininger, D. In: *Handbook of Chemoinformatics*; Gasteiger, J., Ed.; Wiley-VCH: Weinheim, **2003;** Vol. *1*, section 3, pp. 80-102.

[8] Van de Waterbeemd, H.; Lennernäs, H.; Artursson, P., Eds.; *Drug Bioavailability – Estimation of Solubility, Permeability, Absorption and Bioavailability*; Wiley-VCH: Weinheim, 2003.

[9] Lipinski, C.; *J. Pharmacol. Toxicol. Methods*, **2000**, *44*, 235-249.

[10] MOE (the Molecular Operations Environment); available from Chemical Computing Group, 1010 Sherbrooke St. West, Suite 910, Montreal, Quebec, Canada.
For documentation on QSAR descriptors calculated by MOE see: *http://www.chemcomp.com/Journal_of_CCG/Features/descr.htm*.

[11] Varmuza, K. In: *Handbook of Chemoinformatics*; Gasteiger, J., Ed.; Wiley-VCH: Weinheim, **2003**, Vol. *3*, pp. 1098-1133.

[12] Livingstone, D. *Data Analysis for Chemists – Applications to QSAR and Chemical Product Design*; Oxford University Press: Oxford, **1995**.

[13] Mannhold, R. In: *Handbook of Chemoinformatics*; Gasteiger, J., Ed.; Wiley-VCH: Weinheim, **2003;** Vol. *3*, pp. 1300-1313.

[14] Oprea, T. I. In: *Handbook of Chemoinformatics*; Gasteiger, J., Ed.; Wiley-VCH: Weinheim, **2003;** Vol. *4*, pp. 1509-1531.

[15] Takeda, K. Tetrahedron, 1974, 30, 1525-1534.

[16] Tori, K.; Horibe, I.; Tamura, Y.; Kuriyama, H.; Takeda, K. *Tetrahedron Lett.*, **1976**, *5*, 387.

[17] Schmidt, T. J. *J. Mol. Struct.* **1996**, *385*, 99-112.

[18] Schmidt, T. J.; Fronczek, F. R.; Liu, Y.-H. *J. Mol. Struct.*, **1996**, *385*, 113-121.

[19] Kupchan, S.M.; Fessler, D.C.; Eakin, M.A.; Giacobbe, T.J. *Science*, **1970**, *168*, 376-377.

[20] Kupchan, S.M.; Eakin, M.A.; Thomas, A.M. *J. Med. Chem.*, **1971**, *14*, 1147-1152.

[21] Kupchan, S.M.; Giacobbe, T.J.; Krull, I.S. *Tetrahedron Lett.*, **1970**, *33*, 2859-2862.

[22] Pearson, R. G.; *J. Am. Chem. Soc.*, **1963**, *85*, 3533-3539.

[23] Pearson, R. G.; *Coord. Chem. Rev.*, **1990**, 100.

[24] Waddell, T.G.; Gebert, P.H.; Tait, D.L. *J. Pharm. Sci.*, **1983**, *72*, 1474-1476.

[25] Lee, K.-H.; Hall, I.H.; Mar, E.C.; Starnes, C.O.; ElGebaly, S.A.; Waddell,

T.G.; Hadgraft, R.I.; Ruffner, C.G.; Weidner, I. *Science*, **1977**, *196*, 533-535.
[26] Sylvia, V.L.; Joe, C.O.; Stipanovic, R.D.; Kim, H.L.; Busbee, D.L. *Toxicol. Lett.*, **1985**, *29*, 69-76.
[27] Sylvia, V.L.; Kim, H.L.; Norman, J.O.; Busbee, D.L. *Cell Biol. Toxicol.*, **1987**, *3*, 39-49.
[28] Schmidt, T.J.; Matthiesen, U.; Willuhn, G.; *Planta Med.*, **2000**, *66*, 678-681.
[29] Schmidt, T.J., unpublished.
[30] Picman, A.K.; Rodriguez, E.; Towers, G.H.N. *Chem.-Biol. Interactions*, **1979**, *28*, 83-89.
[31] Schmidt, T.J. *Bioorg. & Med. Chem.*, **1997**, *5*, 645-653.
[32] Hall, I.H.; Lee, K.-H.; Mar, E.C.; Starnes, C.O.; Waddell, T.G. *J. Med. Chem.*, **1977**, *20*, 333-337 = T10
[33] Page, J.D.; Chaney, S.G.; Hall, I.H.; Lee, K.-H.; Holbrook, D.J. *Biochim. Biophys. Act.*, **1987**, *926*, 186-194.
[34] Schmidt, T.J.; Lyß, G.; Pahl, H.L.; Merfort, I. *Bioorg. Med. Chem.* **1999**, *7*, 2849-2855.
[35] Heilmann, J.; Wasescha, M.R.; Schmidt, T.J. *Bioorg. Med. Chem.* **2001**, *9*, 2189-2194.
[36] Meister, A.; Anderson, M.E. Annu. Rev. Biochem. *Annu. Rev. Biochem.*, **1983**, *52*, 711-760.
[37] Gertsch, J.; Sticher, O.; Schmidt, T.J.; Heilmann, J. *Biochem. Pharmacol.* **2003**, *66*, 2141-2153.
[38] Dupuis, G.; Mitchell, J.C.; Towers, G.H.N. *Can. J. Biochem.*, **1974**, *52*, 575-581.
[39] Picman, A.K.; Elliott, R.H.; Towers, G.H.N. *Can. J. Zool.* **1981**, *59*, 285-292.
[40] Picman, J.; Picman, A.K. *Pharmazie*, **1990**, *45*, 57-59.
[41] Schmidt, T.J.; *Planta Med.*, **2000**, *66*, 106-109.
[42] Mannervik, B.; Danielson, U.H., *Crit. Rev. Biochem.* **1988**, *23*, 283-337.
[43] Dirsch, V.M.; Stuppner, H.; Vollmar, A.M; Cancer Res. 2001, 61, 5817-5823.
[44] Lee, M.-G.; Lee, K.-T.; Chi, S.-G.; Park, J.-H.; *Biol. Pharm. Bull.*, **2001**, *24*, 303-306.
[45] Wen, J.; You, K.-R.; Lee, S.-Y.; Song, C.-H.; Kim, D.-G.; *J. Biol. Chem.*, **2002**, *277*, 38954-38964.
[46] Hall, I.H.; Starnes, C.O. Jr.; Lee, K.H.; Waddell, T.G. *J. Pharm. Sci.*, **1980**, *69*, 537-543.
[47] Pickersgill, R.W.; Harris, G.; Garman, E.; *Acta Crystallogr. B*, **1992**, *B48*, 59-67.
[48] Beekman, A.C.; Woerdenbag, H.J.; van Uden, W.; Pras, N.; Konings, A.W.T.; Wikström, H.V.; Schmidt, T.J. *J. Nat. Prod.*, **1997**, *60*, 252-257.
[49] Schmidt, T.J.; *Pharm. Pharmacol. Lett.*; **1999**, *9*, 9-13.
[50] Schmidt, T.J.; Heilmann, J. *Quant. Struct.-Act. Relat. (QSAR)*, **2002**, *21*, 276-287.
[51] Eriksson, L.; Antti, H.; Holmes, E.; Johansson, E.; Lundstedt, T.; Shockcor, J.; Wold, S. In: *Handbook of Chemoinformatics*; Gasteiger, J., Ed.; Wiley-VCH: Weinheim, **2003**, Vol. *3*, pp. 1134-1166.
[52] Cho, S. J.; Cummins, D.; Bentley, J.; Andrews, C. W.; Tropsha, A., *"Back" to*

2D QSAR: Application of Genetic Algorithms and Partial Least Squares to Variable Selection of Topological Indices, http://mmlin1.pha.unc.edu/~jin/QSAR/GA_PLS/gapls.html.

[53] Tornhamre, S.; Schmidt, T. J.; Näsman-Glaser, B.; Ericsson, I.; Lindgren, J. A.; *Biochem. Pharmacol.*, **2001**, *62*, 903-922.

[54] Blanco, J. G.; Gil, R. R.; Bocco, J. L.; Meragelman, T. L.; Genti-Raimondo, S.; Flury, A.; *J. Pharmacol. Exp. Ther.*, **2001**, *297*, 1099-1105.

[55] Lyß, G.; Schmidt, T.J.; Merfort, I.; Pahl, H.L *Biol. Chem.*, **1997**, *378*, 951-961.

[56] Hehner, S.; P.; Heinrich, M.; Bork, P.M.; Vogt, M.; Ratter, F.; Lehmann, V.; Schulze-Osthoff, K.; Dröge, W.; Schmitz, M.l. *J. Biol. Chem.*, **1998**, *273*, 1288-1297.

[57] Lyß, G.; Knorre, A.; Schmidt, T.J.; Pahl, H.L.; Merfort, I. *J. Biol.Chem.*, **1998**, *273*, 33508-33516.

[58] Bork, P.M.; Schmitz, M.L.; Kuhnt, M.; Escher, C.; Heinrich, M. *FEBS Lett.*, **1997**, *402*, 85-90.

[59] Rüngeler, P.; Castro, V.; Mora, G.; Gören, N.; Vichnewski, W.; Pahl, H.L.; Merfort, I.; Schmidt, T.J. *Bioorg. Med. Chem*, **1999**, *7*, 2343-2352.

[60] Garcia-Pineres, A.J.; Castro, V.; Mora, G.; Schmidt, T.J.; Strunck, E.; Pahl, H.L.; Merfort, I.; *J. Biol. Chem.* **2001**, *276*, 39713-39720.

[61] Baeuerle, P.A.; Baltimore, D. Cell, 1996, 87, 13-20.

[62] Bremner, P.; Heinrich, M.; J. Pharm. Pharmacol., 2001, 54, 453-472.

[63] Baldwin, A. S. Jr. (ed). NF-κB in defense and disease, Perspective Series, in: *J. Clin. Invest.* **2001**, *107*, Nos. 1-3.

[64] Karin, M.; Lin, A.; *Nature Immunol.* **2002**, *3*, 221-227.

[65] Chen, Y.-Q.; Ghosh, S.; Ghosh, G.; *Nature Struct. Biol.* **1998**, *5*, 67-73.

[66] I. Merfort, personal communication.

[67] Klaas, C.A.; Wagner, G.; Laufer, S.; Sosa, S.; Della Loggia, R.; Bomme, U.; Pahl, H.L.; Merfort, I.; *Planta Med.* **2002**, *68*, 385-391.

[68] Wu, J.B.; Chun, Y.T.; Ebizuka, Y.; Sankawa, U. *Chem. Pharm. Bull.*, **1985**, *33*, 4091-4094.

[69] Wu, J.B.; Chun, Y.T.; Ebizuka, Y.; Sankawa, U. *Chem. Pharm. Bull.*, **1991**, *39*, 3272-3275.

[70] Siedle, B.; Gustavsson, L.; Johansson, S.; Murillo, R.; Castro, V.; Bohlin, L.; Merfort, I.; *Biochem. Pharmacol.*, **2003**, *65*; 897-903.

[71] Knight, D.W. *Natural Product Reports*, **1995**, 271-276.

[72] Heptinstall. S.; Awang, D. V. C. In: *Phytomedicines of Europe*; Lawson, L.; Bauer, R., Eds.; ACS Symposium Series, *691*; American Chemical Society: Washington, DC, **1998**; chapter 13, pp. 158-175.

[73] Marles, R.; Pazos-Sanou, L.; Compadre, C.M.; Pezzuto, J.M.; Bloszyk, E., Arnason, J.T. In: *Phytochemistry of Medicinal Plants*; Arnason, J.T., Mata, R.; Romeo, J.T., Eds.; Plenum Press: New York, **1995**; chapter 13, pp. 334-356.

[74] Warshaw, E.M.; Zug, K.A. *Contact Dermatitis*, **1996**, *7*, 1-23.

[75] Hausen, B.M.; Schmalle, H.W. *Contact Dermatitis*, **1985**, *13*, 329-332.

[76] Benezra, D.; Ducombs, G.; *Dermatosen*, **1987**, 35, 4-11.

[77] François, G.; Passreiter, C.M.; Woerdenbag, H.J.; Van Looveren, M. *Planta*

Med., **1996**, *62*, 127-129.
[78] François, G.; Passreiter, C.M., *Phytother. Res.*, **2004**, *18*, 184-186.
[79] Phillippson, J.D.; Wright, C. W.; *Planta Med.*, **1991**, *57 (Suppl. 1)*, S53-S59.
[80] Fuchino, H.; Koide, T.; Takahashi, M.; Sekita, S.; Satake, M.; Planta Med., **2001**, *67*, 647-653.
[81] Schmidt, T.J.; Brun, R.; Willuhn, G.; Khalid, S. A.; *Planta Med.* **2002**, *68*, 750-751.
[82] Hoet, S.; Opperdoes, F.; Brun, R.; Quentin-Leclercq, J.; *Nat. Prod. Rep.*, **2004**, *21*, 353-364.
[83] Fournet, A.; Munoz, V.; Roblot, F., Hocquemiller, R.; Cave, A.; Gantier, J.C.; *Phytother. Res.*, **1993**, *7*, 111-115.
[84] Grael, C.F.; Vichnewski, W.; Souza, G. E.; Lopes, J.L; Albuquerque, S.; Cunha, W. R.; *Phytother. Res.*, **2000**, *13*, 203-206.
[85] G. François, personal communication to G. Willuhn, C. M. Paßreiter and T. J. Schmidt.
[86] Lee, K. H.; Ibuka, T.; Wu, R.-Y.; Geissman, T. A. *Phytochemistry*, **1977**, *16*, 1177-1181.
[87] L. Krauth-Siegel, personal communication.
[88] Woerdenbag, H. J.; Merfort, I.; Paßreiter, C. M.; Schmidt, T. J.; Willuhn, G.; van Uden, W.; Pras, N.; Kampinga, H. H.; Konings, A. W. T.; *Planta Med.*, **1994**, *60*, 434-437.
[89] China Cooperative Research Group on Quinghaosu and Its Derivatives as Antimalarials. *J. Trad. Chin. Med.*, **1982**, *2*, 3-8.
[90] Liu, J.-M.; Ni, M.-Y.; Fan, J.-F.; *Acta Chim. Sin.*, **1979**, *37*, 129-143.
[91] Klayman, D. L. *Science*, **1985**, *228*, 1049-1055.
[92] Meshnick, S.R.; Taylor, T.E.; Kamchonwongpaisan, S.; *Mircobiol. Rev.*, **1996**, *60*, 301-315.
[93] Meshnick, S. R.; *Int. J. Parasitol.* **2002**, *32*, 1655-1660.
[94] Mehlhorn, H., Ed.; *Encyclopedic Reference of Parasitology*; Springer: Berlin, Heidelberg, New York, **2001**.
[95] Robert, A.; Meunier, B.; *Chem. Soc. Rev.*, **1998**, *27*, 273-279.
[96] Robert, A.; Meunier, B.; *J. Am. Chem. Soc.* **1997**, *119*, 5968-5969.
[97] Haynes, R. K.; Ho, W.-Y.; Chan, H.-W.; Fugmann, B.; Stetter, J.; Croft, S. L.; Vivas, L.; Peters, W.; Robinson, B. L.; *Angew. Chem.*, **2004**, *116*, 1405-1409.
[98] Eckstein-Ludwig, U.; Webb, R. J.; van Goethem, I. D. A.; East, J. M.; Lee, A. G.; Kimura, M.; O'Neill, P. M.; Bray, P. G.; Ward, S. A.; Krishna, S.; *Nature*, **2003**, *424*, 957-961.
[99] Liang, X.-T.; Yu, D.-Q., Wu, W.-L.; Deng, H.-C.; *Acta Chim. Sin.*, **1979**, *37*, 215-230.
[100] Hofheinz, W.; Bürgin, H.; Gocke, E.; Jaquet, C.; Masciadri, R.; Schmid, G.; Stohler, H.; Urwyler, H.; *Trop. Med. Parasitol.* **1994**, *45*, 261-265.
[101] Bez, G.; Kalita, B.; Sarmah, P.; Barua, N.C.; Dutta, D.K.; *Current Org. Chem.*, **2003**, *7*, 1231-1255.
[102] Jefford, C. W.; Burger, U.; Millasson-Schmidt, P.; Bernardinelli, G.; Robinson, B. L.; Peters, W.; *Helv. Chim. Act.*, **2000**, *83*, 1239-1246.
[103] Provot, O.; Camuzat-Denis, B.; Hamzaoui, M.; Moskowitz, H.; Mayrargue, J.;

Robert, A.; Cazelles, J.; Meunier, B.; Zouhiri, F.; Desmaele, D.; d'Angelo, J.; Mahuteau, J.; Gay, F.; Ciceron, L.; *Eur. J. Org. Chem.*, **1999**, 1935-1938.

[104] Tonmunphean, S.; Parasuk, V.; Kokpol, S.; *Quant. Struct. Act. Relat.*, **2001**, *19*, 475-483.

[105] Pinheiro, J. C.; Kiralj, R.; Ferreira, M. M. C.; Romero, O. A. S.; *QSAR Comb. Sci.*, **2003**, *22*, 830-842.

[106] Avery, M. A.; Gao, F.; Chong, W. K. M.; Mehrotra, S.; Milhaus, W. K.; *J. Med. Chem.* **1993**, *36*, 4264- 4275.

[107] Woolfrey, J. R.; Avery, M. A.; Doweyko, A. M.; *J. Comp.-Aided Mol. Design*, **1998**, *12*, 165-181.

[108] Avery, M. A.; Alvim-Gaston, M.; Rodrigues, C. R.; Barreiro, E. J., Cohen, F. E.; Sabnis, Y. A.; Woolfrey, J. R.; *J. Med. Chem.*, **2002**, *45*, 292-303.

[109] Avery, M. A.; Muraleedharan, K. M.; Desai, P. V.M Badyopadhyaya, A. K.; Furtado, M. M.; Tekwani, B. L.; *J. Med. Chem.*, **2003**, *46*, 4244-4258.

[110] Jefford, C. W.; Grigorov, M.; Weber, J.; Lüthi, H. P.; Tronchet, J. M.; *J. Chem. Inf. Comput. Sci.*, **2000**, *40*, 354-357.

[111] Girones, X.; Gallegos, A.; Carbo-Dorca, R.; *J. Chem. Inf. Comput. Sci.*, **2000**, *40*, 1400-1407.

[112] Gozalbes, R.; Galvez, J.; Moreno, A.M.; Garcia-Domenech, R.; *J. Pharm. Pharmacol.*, **1999**, *51*, 111-117.

[113] Cramer, R. D. III; Patterson, D. E.; Bunce, J. D.; *J. Am. Chem. Soc.*, **1988**, *110*, 5956-5967.

[114] Heritage, T. W.; Lowis, D. R. In: *Rational Drug Design. Novel Methodology and Practical Applications*; American Chemical Soicety: Washington DC, **1999**, pp 212-225.

[115] Kubinyi, H. In: *Handbook of Chemoinformatics*; Gasteiger, J., Ed.; Wiley-VCH: Weinheim, **2003**, Vol. *3*, pp. 1555-1574.

[116] Anzali, S.; Barnickel, G.; Krug, M.; Wagener, M.; Gasteiger, J.; In: *Computer-Assisted Lead Finding and Optimization- Current Tools for Medicinal Chemistry*; Van de Waterbeemd, H.; Testa, B.; Folkers, G., Eds.; Wiley-VCH: Weinheim, **1997**; pp. 93-106.

[117] Zupan, J. In: *Handbook of Chemoinformatics*; Gasteiger, J., Ed.; Wiley-VCH: Weinheim, **2003**, Vol. *3*, section 1.4, pp. 1167-1215.

[118] Posner, G. H.; Paik, I.-H.; Sur, S.; McRiner, A. J.; Borstnik, K.; Xie, S.; Shapiro, T. A.; *J. Med. Chem.*, **2003**, *46*, 1060-1065.

[119] Posner, G. H.; Ploypradith, P.; Parker, M. H.; O'Dowd, H.; Woom S.-H.; Northrop, J.; Krasavin, M.; Dolan, P.; Kensler, T. W.; *J. Med. Chem.*, **1999**, *42*, 4275-4280.

[120] Beekman, A. C.; Woerdenbag, H. J.; Kampinga, H. H., Konings, A. W. T.; *Phytotherapy Res.*, **1996**, *10*, 140-144.

[121] Beekman, A. C.; Barentsen, A. R. W.; Woerdenbag, H. J.; Van Uden, W.; Pras, N.; Konings, A. W. T.; El-Feraly, F. S.; Galal, A. M.; Wikström, H. V.; *J. Nat. Prod.*, **1997**, *60*, 325-330.

[122] Beekman, A. C.; Wierenga, P. K.; Woerdenbag, H. J.; Van Uden, W.; Pras, N.; Konings, A. W. T., El-Feraly, F. S.; Galal, A. M.; Wikström, H. V.; *Planta Med.*, **1998**, *64*, 615-619.

[123] Boyd, M. R.; Paull, K. D.; *Drug Develop. Res.*, **1995**, *34*, 91-109.
[124] Christensen, S.B.; Schaumburg, K.; *J. Org. Chem.*, **1987**, *48*, 396-399.
[125] Christensen, S.B.; Norup, E.; *Tetrahedron Lett.*, **1985**, *26*, 107-110.
[126] Rasmussen, U.; Christensen, S.B.; Sandberg, F.; *Planta Med.* **1981**, *43*, 336-341.
[127] Christensen, S.B.; Norup, E.; Rasmussen, U.; Madsen, J.Ø.; *Phytochemistry* **1984**, *23*, 1659-1663.
[128] Thastrup, O.; Dawson, A.P.; Scharff, O.; Foder, B.; Cullen, J.P.; Drøbak, B.K. Bjerrum, P.J. Christensen, S.B.; Hanley, M.R.; *Agents and Actions*, **1989**, *27*, 17-23.
[129] Treiman, M.; Caspersen, C.; Christensen, S.B.; *TiPS*, **1998**, *19*, 131-135.
[130] Rasmussen, U. Christensen, S.B.; Sandberg, F.; *Acta Pharm. Suec.*, **1978**, *15*, 133-140.
[131] Ali, H.; Christensen, S.B.; Foreman, J.C.; Pearce, F.L.; Piotrowski, W.; Thastrup, O.; *Br. J. Pharmac.*, **1985**, *85*, 705-712.
[132] Malcolm, K.C.; Fitzpatrick, F.A.; *J. Pharmacol. Exp. Ther.*, **1992**, *260*, 1244-1249.
[133] Tsukamoto, A.; Kaneko, Y.; *Cell Biol. Internat.*, **1993**, *17*, 969-970.
[134] Wei, H. Wei, W.; Bredesen, D.E. Perry, D.C.; *J. Neurochem.*, **1998**, *70*, 2305-2314.
[135] Hakii, H.; Fujiki, H.; Suganuma, M.; Nakayasu, M.; Tahira, T.; Sugimura, T.; Scheuer, P.J.; Christensen, S.B.; *J. Cancer Res.*, **1986**, *Clin. Oncol. 111*, 177-181.
[136] Pahl, H.L.; Baeuerle, P.A.; *Trends Cell Biol.*, **1997**, *7*, 51-55.
[137] Pahl, H.L.; Sester, M.; Burgert, H.-G.; Baeuerle, P.A.; *J. Cell Biol.*, **1996**, *132*, 511-522.
[138] Sagara, Y.; Inesi, G.; *J. Biol. Chem.*, **1991**, *266*, 13503-13506.
[139] Inesi, G.; Sagara, Y. *Arch. Biochem. Biophys.*, **1992**, *298*, 313-317.
[140] Inesi, G.; Sagara, Y.; *J. Membrane Biol.* **1994**, *141*, 1-6.
[141] Sagara, Y.; Wade, J.B.; Inesi, G.; *J. Biol. Chem.*, **1992**, *267*, 1286-1292
[142] Sagara, Y.; Fernandez-Belda, F.; de Meis, L.; Inesi, G.; *J. Biol. Chem.*, **1992** , *267*, 12606-12613.
[143] Witcome, M.; Henderson, I.; Lee, A.G.; East, J.M.; *Biochem. J.*, **1992**, *283*, 525-529.
[144] Toyoshima, C.; Nakasako, M.; Nomura, H.; Ogawa, H.; *Nature,* **2000**, *405*, 647-655.
[145] Christensen, S.B.; Andersen, A.; Poulsen, J.-C.J., Treiman, M.; *FEBS Lett.*, **1993**, *335*, 345-348.
[146] Nørregaard, A.; Vilsen, B.; Andersen, J.P.; *J. Biol. Chem.*, **1994**, *269*, 26589-26601.
[147] Toyoshima, C.; Nomura, H.; *Nature,* **2002**, *418*, 605-611.
[148] Holub, M.; Budesinsky, M.; *Phytochemistry*, **1986**, *25*, 2015-2026.
[149] Christensen, S.B.; Hergenhahn, M.; Roeser, H.; Hecker, E.; *Cancer Res. Clin.*

Oncol. **1992**, *118*, 344-348.

[150] Norup, E.; Smitt, U.W.; Christensen, S.B.; *Planta Med.*, **1985**, *251-255.*,
[151] Nielsen, S.F.; Thastrup, O.; Pedersen, R.; Olsen, C.E.; Christensen, S.B.; *J. Med. Chem.*, **1995**, *38*, 272-276.
[152] Christensen, S.B.; Andersen, A.; Kromann, H.; Treiman, M.; Tombal, B.; Denmeade, S.; Isaacs, J.T.; *Bioorg. Med. Chem.*, **1999**, *7*, 1273-1280.

[153] Jakobsen, C.M.; Denmeade, S.R.; Isaacs, J.T.; Gady, A.; Olsen, C.E.; Christensen, S.B.; *J. Med. Chem.*, **2001**, *44*, 4696-4703.

[154] Egan, M.E.; Glöckner-Pagel, J.; Ambrose, C.A.; Cahill, P.A.; Pappoe, L; Balamuth, N; Cho, E.; Canny, S.; Wagner, C.A.;Geibel, J.; Caplan, M.J; *Nature Medicine*, **2002**, *8*, 485-492.

[155] Delisle, B.P.; Anderson, C.L.; Balijepalli, R.C.; Anson, B.D.; Kamp,T.J.; January, C.T.; *J. Biol. Chem.*, **2003**, *278*, 35749–35754.

[156] Denmeade, S.R.; Jakobsen, C.M.; Janssen, S.; Khan, S.R.; Garrett, E.S.; Lilja, H.; Christensen, S.B.; Isaacs, J.T. *Journal of the National Cancer Institute* **2003**, *95*, 990-1000.

[157] Schmidt, T.J.; Müller, E.; Fronczek, F.R.; *J. Nat. Prod.*, **2001**, *64*, 411-414.
[158] Squires, R.F.; Casida, J.E.; Richardson, M.; Saederup, E. *Mol. Pharmacol.*, **1983**, *23*, 326-336.
[159] Kudo, Y.; Oka, J.-I.; Yamada, K. *Neurosci. Lett.*, **1981**, *25*, 83-88.
[160] Matsumoto, K.; Fukuda, H. *Neurosci. Lett.*, **1982**, *32*, 175-179.
[161] Matsumoto, K.; Fukuda, H. *Brain Res.*, **1983**, *270*, 103-108.
[162] Kuriyama, T.; Schmidt, T.J.; Okuyama, E.; Ozoe, Y.; *Bioorg. Med. Chem.*, **2002**, *10*, 1873-1881.
[163] Schmidt, T.J.; Gurrath, M.; Ozoe, Y.; *Bioorg. Med. Chem.*, **2004**, *12*, 4159-4167.
[164] Hosie, A.M.; Ozoe, Y.; Koike, K.; Ohmoto, T.; Nikaido, T.; Sattelle, D.B. *Br. J. Pharmac.*, **1996**, *119*, 1569-1576.
[165] Ozoe, Y.; Akamatsu, M.; Higata, T.; Ikeda, I.; Mochida, K.; Koike, K.; Ohmoto, T.; Nikaido, T. *Bioorg. Med. Chem.*, **1998**, *6*, 481-492.
[166] Klunk, W.E.; Kalman, B.L.; Ferrendelli, J.A.; Covey, D.F. *Mol. Pharmacol.*, **1983**, *23*, 511-518.
[167] Inoue, M.; Akaike, N. *Neurosci. Res.*, **1988**, *5*, 380-394.
[168] Newland, C.F.; Cull-Candy, S. G. *J. Physiol.*, **1992**, *447*, 191-213.
[169] Yoon, K.-W.; Covey, D.F.; Rothman, S.M. *J. Physiol.*, **1993**, *464*, 423-439.
[170] Zhang, H.-G.; ffrench-Constant, R.H.; Jackson, M.B. *J. Physiol.*, **1994**, *479*, 65-75.
[171] Jarboe, C.H.; Porter, L.A.; Buckler, R.T *J. Med. Chem.*, **1968**, *11*, 729-731.
[172] Ozoe, Y.; Akamatsu, M.; *Pest Manag. Sci.*, **2001**, *57*, 923-931.
[173] Akamatsu, M.; Ozoe, Y.; Ueno, T.; Fujita, T.; Mochida, K.; Nakamura, T.; Matsumura, F.; *Pestic. Sci.* **1997**, *49*, 319-332.
[174] Wong, M.G.; Andrews, P.R. *Eur. J. Med. Chem.*, **1989**, *24*, 323-334.
[175] Calder, J.A.; Wyatt, J.A.; Frenkel, D.A.; Casida, J.E. *J. CAMD*, **1993**, *7*, 45-

60.
[176] Holland, K.D.; Bouley, M.G.; Covey, D.F.; Ferrendelli, J.A. *Brain Res.*, **1993**, *615*, 170-174.
[177] Trifilletti, R.R.; Snowman, A.M.; Snyder, S.H. *Eur. J. Pharmacol.*, **1985**, *106*, 441-447.
[178] Niwa, H.; Nisiwaki, M.; Tsukada, I.; Ishigaki, T.; Ito, S.; Wakamatsu, K.; Mori, T.; Ikagawa, M.; Yamada, K. *J. Amer. Chem. Soc.*, **1990**, *112*, 9001-9003.
[179] Schmidt, T.J.; Schmidt, H.M.; Müller, E.; Peters, W.; Fronczek, F.R.; Truesdale, A.; Fischer, N.H. *J. Nat. Prod.*, **1998**, *61*, 230-236.
[180] Schmidt, T.J.; Peters, W.; Fronczek, F.R.; Fischer, N.H. *J. Nat. Prod.*, **1997**, *60*, 783-787.
[181] Schmidt, T.J. *J. Nat. Prod.*, **1999**, *62*, 684-687.
[182] Schmidt, T.J.; Okuyama, E.; Fronczek, F.R. *Bioorg. Med. Chem.*, **1999**, *7*, 2857-2865.
[183] Hosie, A. M.; Aronstein, K.; Sattelle, D. B.; ffrench-Constant, R. H.; *Trends Neurosci.* **1997**, *20*, 578-583.
[184] Mezler, M.; Müller, T.; Raming, K.; *Eur. J. Neurosci.* **2001**, *13*, 477-486.
[186] Vedani, A.; Dobler, M.; *J. Med. Chem.* **2002**, *45*, 2139-2149.
[187] *Quasar*-Manual, http://www.biograf.ch
[188] Miyazawa, A.; Fujiyoshi, Y.; Unwin, N.; *Nature*, **2003**, 423, 949-955.
[189] SwissProt and TrEMBL databases, *http://www. expasy.org*
[190] Y. Ozoe, personal communication on results of an ongoing cooperation.
[191] Campbell, W.C.; *Annu. Rev. Microbiol.*, **1991**, *45*, 445-474.