A chemical Paper-Chase

Sorbicillactone A

Sea sponges (poriferans) are a particularly productive source of natural substances. Rich in nutrients, they are a sitting target for fish because, although they are animals, they can't run away. They may have no sting or armor but they have effective chemical weapons; active substances that they can use against their natural predators. If such a poison is highly selective, for example attacking human tumor cells but not normal body cells, then this could be a promising candidate for an anti-cancer drug. But how can we identify such molecules? How can we harvest this treasure of active compounds and make this "Sea-bed Drugstore" openly available? After all, the more common sponges have been well investigated and the others are either rare or endangered and, for the most part, can't be farmed to increase their numbers. It is important to note that it is often not the sponge that "produces" the active compounds but co-existing micro-organisms – bacteria, single-cell algae and mould fungi. If one of these co-habitants is isolated and propagated on growth medium to examine the constituents, there is a good chance of discovering producers of brand new agents. However that still leaves the problem of how to approach the new agents? Follow us on one such "paper-chase" with the discovery of a new active substance called sorbicillactone A or "sorbi" for short.

A mould fungus network, as you might see on mouldy bread, but this is a brown-yellow colored one, nearly golden and swimming in a nutrient-rich solution (Fig. 1). Not a common mould fungus, as found everywhere, but pre-isolated from a Mediterranean sponge collected near Elba and separated from the many other microorganisms present, isolated and then propagated. Is this mould fungus, identified as ***Penicillium chrysogenum***, potentially a producer of completely new active agents?

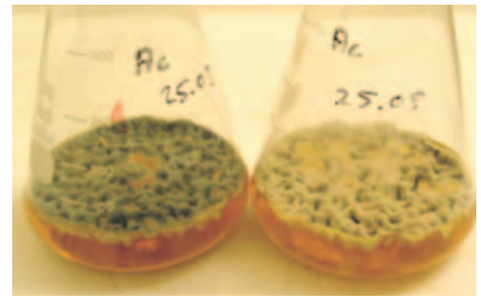


Fig. 1: ***P. chrysogenum*** grown in a nutrient-rich solution

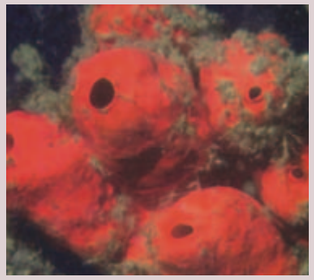
In the beginning came the extract

For this purpose you have to, initially, prepare an extract (Fig. 2). Firstly the fungus is sieved from the nutrient solution, carefully dried, milled into small pieces and then extracted in the same way that we all brew a cup of coffee or tea every day; extracting active compounds and flavours from that finely milled biological material. In the case of the mould, we not only use water in the extraction process but also different organic solvents.

Fig. 2: Extract of ***P. chrysogenum***

In contrast to tea or coffee, however, we don't want to end up with the whole crude extract mixture but rather the specific, pure active substances which show special biological activity or are, structurally, particularly new. But how do we get them out?

To do this, we have to get a picture of the range ("bouquet") of substances in the extract using, for example, high-performance liquid chromatography (***HPLC***). For this purpose, the substance mixture (here the dried extract) is dissolved in a small amount of solvent and then forced through a chromatographic column, a cylindrical metal tube filled with a special material. After applying the substance to the head of the column, it is transported through the column with the help of a pump and additional fresh solvent introduced behind



Sponges are a rich source of diverse, highly active natural products

HPLC (High-Performance Liquid Chromatography)

A further development of Column-Chromatography, HPLC, works under a pressure of approximately 100-1000 bar on particularly fine-grained adsorbent materials and exhibits a high capacity for separation. The time taken for a substance to elute from the column is called retention time and can be used together with spectroscopic information to identify compounds.

The development of HPLC instruments and chromatographic materials has allowed natural product extracts to be separated, nowadays, in a matter of minutes while reducing the quantity of sample material required.

A high sample throughput and low consumption of sample material is possible when connected to an autosampler. The sample throughput can be further increased by switching between several chromatography columns in parallel served by several pumps (parallel HPLC).

HPLC equipment

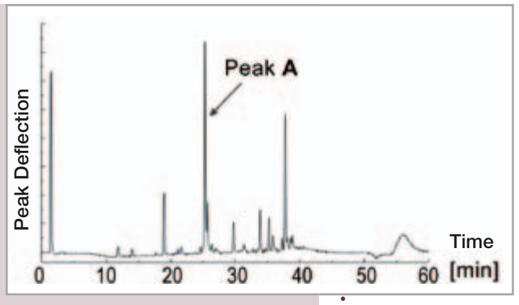


Fig. 3: Chromatogram

Mass Spectrometry
A method to identify the masses of atoms and molecules, using the deflection of the corresponding ions in an electrical or magnetic field.

it. The substance mixture passes through the entire column but not all components run at the same speed.

Considering their structures and their physico-chemical properties the components of the extract mixture elute from the column at different times depending on their individual affinity to the column adsorbent material. Each component/compound eluting from the column is registered with the help of a detector as a deflection (peak). The more peaks, the more components are in the mixture and the larger the peak, the greater the proportion of that component is present in the substance mixture (see Fig. 3).

THE QUESTION IS "WHICH PEAKS ARE OF INTEREST AND WHICH PEAKS ARE SIMPLY COMMON, WELL-KNOWN SUBSTANCES?"

Classically, you answer this question by using the same HPLC methodology "preparatively", i.e., with much larger volumes. Then, for each separately detected peak, the solvent is evaporated to produce pure substances in larger quantities, the structures of which can then be elucidated. However, that is far too time-consuming if you consider that many of the substances have already been discovered in other organisms (e.g., in other mould fungi) and so it is hardly worth the trouble. It would therefore be important to figure out, before isolating a substance in its pure form, whether a structure is still unknown and if so, what it looks like.

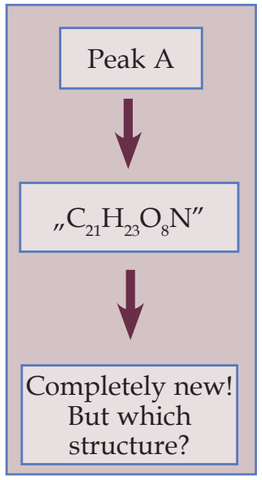
Fermentation of the fungus *Penicillium chrysogenum* as surface liquid cultures

COULD THIS BE RECOGNIZED DIRECTLY FROM THE PEAKS IN THE CHROMATOGRAM?

Yes, it could! Usually you must already have reference compounds which, under standard conditions required the same time to pass through the HPLC column (retention time). It is much more efficient, however, to couple the HPLC directly to spectroscopic methods i.e., analyze the peak as soon as it elutes from the column. This can already be done with minor quantities, as is the case for our new substance.

HPLC-MS coupling is an efficient procedure allowing a direct analysis of HPLC peaks with the help of *mass spectrometry* (MS). Using this hyphenated technique, one can measure the weight (molecular mass) of every molecule eluting from the column. These masses are inconceivably small but can, nevertheless, be measured accurately. Taking the weight of different atomic species in the molecule – e.g., carbon, hydrogen, oxygen and nitrogen, and the total weight of the molecule, it is then possible to make an initial determination about the atomic composition. This and much more information about over 200,000 well-known natural substances is collected in databases. In the case of our extract, this methodology very quickly supplied information that three of the six largest peaks of our fungus extract were well-known substances. Among them for example was roquefortine C, which is already well known from the Roquefort cheese fungus.

What was more interesting, by contrast, was a substance (Fig. 3: peak **A**) that appeared at a retention time of 25 minutes, to which the molecular formula $C_{21}H_{23}0_8N$ could be assigned very quickly (i.e., 21 carbon atoms, 23 hydrogen atoms, 8 oxygen atoms, and 1 nitrogen atom per molecule). This substance had obviously never been discovered in nature as no appropriate database entry existed. So, here we had one completely new substance!



IS IT BRAND NEW? OR IS IT ULTIMATELY JUST A KNOWN STRUCTURE WITH SLIGHTLY MODIFIED MAIN FEATURES?

Which technique could be used to get a quick answer without going through the time-consuming process of isolating the compound in a pure form in sufficient quantities?

The puzzle begins: first details of the chemical structure

NMR spectroscopy is particularly suitable for this purpose. This technique examines the nuclear magnetic resonance (NMR) of a substance and the individual carbon and hydrogen atoms can be "seen". More importantly the special environment can be clarified; identifying which atom is joined to which atom. From this, the whole molecule reveals itself piece by piece, built up like a jig-saw puzzle. In such a way, the compound of interest represented by peak **A** consists of three structural parts (Fig. 4):

A1 consisted of a six-membered ring (more precisely a hexagon) of carbon atoms with double bonds in the ring, three oxygen functions, two methyl groups, and a branched chain;

A2 consisted of a chain with six carbon atoms, double bonds and a carbonyl group (a so-called sorbyl component);

A3 consisted of a fragment with four carbon atoms, almost identical to the well-known fumaric acid, which is often produced by mould fungi and also occurs as an acid in fruits.

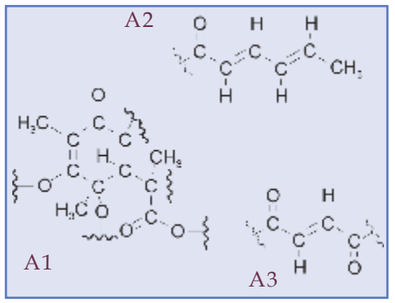


Fig. 4: Partial structures of the new substance represented by peak A

HOW ARE THESE FRAGMENTS CONNECTED?

Advanced NMR investigations, using the so-called HMBC method, have the ability to explore proton-carbon correlations across multiple bonds, thus providing information as to which locations on the 6-ring structure (**A1**) the two side structures (**A2** and **A3**) were connected, and suggested the total structure **A4** or, in a clearer notation, **A4'** (Fig. 5). Here each corner of the hexagon, each point in the zigzag chain, denotes a C atom.

600 MHz NMR for the measurement of the spectra of matter.

NMR Spectroscopy
A significant method for the structural elucidation of organic molecules. It utilizes the magnetic characteristics of certain atomic nuclei (see also glossary p. 107)

BUT WAS THIS THE STRUCTURE OF THE NATURAL SUBSTANCE?

Indeed, the structure was not complete. Again mass spectrometry was used clearly demonstrating that the actual structure of the new natural product had to be lighter than the postulated structure **A4**, namely be 18 mass units. 18 is exactly the mass of a molecule of water. Thus the actual natural product is characterized by the loss of an H_20 molecule, and the formation of a consecutive additional 5-membered ring with an integrated oxygen (O), a lactone ring. (Fig. 5, **A5**). This

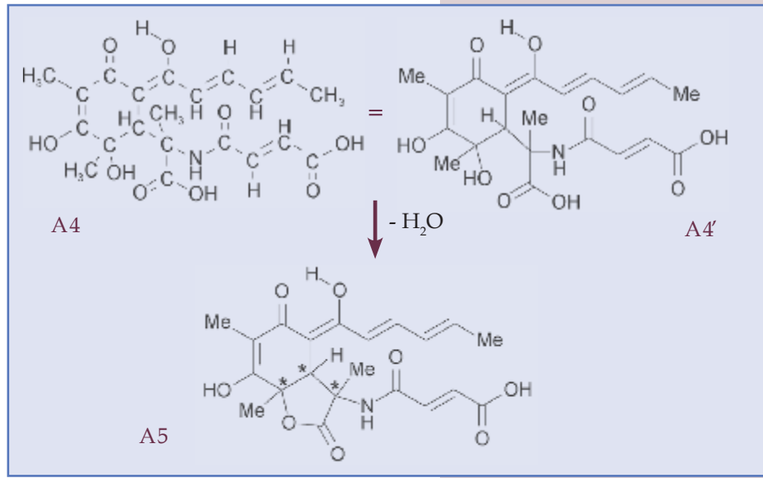


Fig. 5: Illustration of the established still flat structure of the new agent sorbicillactone A

From left to right the crucial extract from the fungus and the different purity steps resulting from the crude extract during the purification process

Chiral Substance
A substance with two non-identical forms whose spatial structure varies only in that they are a mirror image of one-another.

was, obviously, a completely new structure, which, thus, deserved a new name. Due to the presence of a sorbyl side-chain, the similarity with the natural product sorbicillin (similarly extracted from the fungus) and the presence of a lactone ring, the substance indicated by peak **A** was now given the name sorbicillactone A.

However, is the **A5** structure now totally complete? Hardly, because the molecule really can't be as flat as depicted in Fig. 5. In particular the three carbon atoms highlighted with a * have their four partners arranged in four different directions in space, corresponding to a tetrahedron. This raises a question: in which directions in space are the two methyl groups (Me) and the hydrogen (H) located at the 6-5-ring system pointing? Upwards or downwards? Or verbalized in the language of the chemist:

WHAT IS THE STEREOSTRUCTURE OF SORBICILLACTONE A?

NMR spectroscopy also provides an initial answer to this question, again without the need to isolate the substance in pure form. The spatial interaction of the three identified atoms or atom groups can be shown because they communicate with each other (represented by the double-headed arrows in Fig. 6) and this can only be explained by them being on the same side of the ring system. So we know that they are all oriented either above (**A6**) — or all below (mirror-**A6**) the surface of this page.

Picture or mirror image; what is the precise structure?

Is it really important? Both possibilities; "all three above" or "all three below" , would act like a picture and mirror image, both would have completely the same dimensions and would have identical angles and distances between the atoms. So, is this "making a mountain out of a molehill" just academical? Certainly not, because, with active substances the "image or mirror-image" issue can make an enormous difference. You only need to remember the terrible example of thalidomide (only one form shows the growth-impairing characteristics) or in many more harmless examples of active substances such as carvone, where one form smells like caraway seeds, the mirror image (the enantiomer) smells like mint. But how do we work out with our fungus product whether it is the "image" or the "mirror image"? This is where circular dichroism spectroscopy (CD spectroscopy) is of great assistance in the process. It measures the interaction of a ***chiral substance*** with polarized light. The opposites, the image and the mirror image, behave completely differently (in the same way that a left hand feels different in left-hand glove as opposed to a right-hand glove). The image and mirror image molecules supply mirror-image CD spectra. Now, how can one interpret these spectra, or better still:

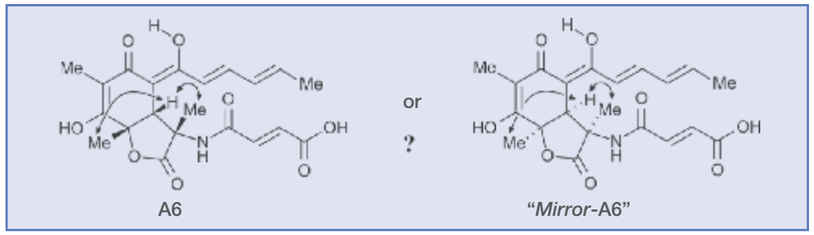


Fig. 6: Illustration of the chiral forms of sorbicillactone A

"WHICH IS WHICH"?

Here, theoretical chemistry can help. Using sophisticated calculations, the CD spectrum of a given structure or its mirror image can be predicted. These CD curves can then be compared with the experimental ones, and the 3D structure can be assigned exactly.

In this case, and without having to isolate the substance, it was discovered that the three groups (Me, H, Me) all are located on the upper side of the molecule. Consequently, **A6** is the exact structure of sorbicillactone A (Fig. 7).

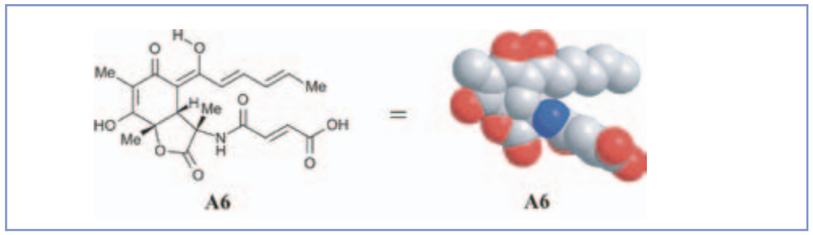


Fig. 7: Actual structure of sorbicillactone A

PERSPECTIVES

If we had a time machine and could ask both a 1960's Natural Product Chemist and a Natural Product Chemist of today to elucidate the structure of the same substance, would this be an unfair challenge? The 1960's chemist would have had no chance of solving this problem compared with his contemporary counterpart. Not because they were worse scientists back then but the analytical and spectroscopical methods mentioned above either did not yet exist or were in their infancy. The elucidation process took a great deal longer. Which poses the question "How long will structural elucidation take fifty years from now? Will it take only a matter of seconds and will be fully automated and computerized? Will the natural-product scientist of the future be simply a data manager?

Completely fiction? The quality and quantity of the spectroscopic data will certainly improve. The analytic devices will get ever faster in the future and will have higher sensitivity and improved data quality. Should it become a reality, comprehensive spectroscopic databanks will be generated and will be at the disposal of natural-product researchers from both, the industry and research sectors and database searches will have greater relevance. The identification of substances will be only a mouse-click away. This means that rapid identification of known substances and the determination of structures will become more of a routine.

The FCPC Rotor (Fast Centrifugal Partition Chromatograph) facilitates the preliminary purification of large samples of crude extract through Liquid-Liquid Chromatography

WHERE DO WE GO FROM HERE? IS THIS THE END OF THE STORY?

Certainly not! Because more recent research indicates that "sorbi" is an exciting new substance, perhaps even the first example of a new class of compounds; one that is well worth the time and effort to isolate in its pure form.

Natural-Substance Analysis

A core function of natural-product chemistry is the investigation of the composition of extracts and, in the case of new natural products, the elucidation of their three-dimensional structure. In principle the entire arsenal of analytical procedures of organic chemistry could be applied to achieve this. However natural substance extracts are often characterized by their high complexity. A good analysis must be able to capture the substance in its entirety and so it is essential that different powerful chromatographical and spectroscopical methods are used. It is a big challenge, using dereplication, to speedily and efficiently isolate the new, preferably bioactive, substance from the known substances in a sample with more than a hundred components. So, instead of the time-consuming process of isolating pure components directly from the extract, the elucidation is carried out through a combination of chromatography coupled with on-line analytical techniques (UV, MS, NMR, CD)

PERSPECTIVES

But how does it work with new and, so far, unknown substances, for which no reference data exists? Will structural elucidation also become a routine process? Will future development of software for the evaluation of spectroscopic data make this possible? Surely important data processing tools like this will be developed in the future and will relieve much of the work of the chemist? However, the chemist will never be replaced, since the structural elucidation of complex substances can bring chemists face to face with unexpected/surprising results for which software can not be designed. Furthermore, for elucidation, a deep knowledge on many natural products is needed. Theoretical structures can be programmed in silico (by the computer), which make no sense from a chemico-biological point of view.

Advances in hyphenated techniques such as HPLC-NMR or HPLC-MS will contribute greatly to increased efficiency during the structural elucidation process. The „Total Organic Analysis Device“ that preserves complete spectroscopic information; a complete data record resulting from only one HPLC separation is a not too distant vision.

Increasingly the arsenal of methods from other fields of activity, such as proteomics, are becoming available. Particularly methods developed for the analysis of low-molecular mass natural products will make the analysis of small quantities of substances possible. An interesting aspect is the development of microscopic methods, for example Atomic Force Microscopy now has resolution approaching atomic level. Using this technique it is becoming possible to represent the molecules coating a surface, giving us, effectively, a photo of the substance.

Ca. 10 g of pure sorbicillactone A

A NOVEL STRUCTURE – BUT DOES IT HAVE APPLICATIONS AS A NEW MEDICAL AGENT OR A NEW PESTICIDE?

In the discovery process you need the actual substance in your hands. So, it was initially isolated in small amounts (approx. 5 mg) and tested.

Sorbi has been shown to act against leukaemia cells, and the AIDS virus, HIV, while, at the same time, it was virtually non-toxic to healthy body cells. So, it's a promising new active substance with a completely unprecedented structure and with potential anti-tumor properties.

HOW DO YOU OBTAIN LARGE QUANTITIES OF THIS SUBSTANCE?

How nature makes this substance, i.e., how the "sorbi" fungus cells synthesize it, is quite well understood. Thus, the basic chemicals that the fungus uses are known: 6 molecules of the natural product acetic acid, 2 molecules of the amino acid methionine, 1 molecule of the amino acid alanine and 1 molecule related to fumaric acid. In the fungus these "simple" components were linked through several enzymes in the microorganisms to the very complex structure of sorbicillactone A.

After optimization of the growth conditions of the sponge fungus, production of sorbi succeeded at a larger scale in the laboratory. In addition nutrient solutions, cultivation duration etc. were analyzed to maximize the fungus' production of sorbicillactone A., The fungi were grown in large numbers under these optimized conditions in order to continually harvest the mycelia. The material was dried, pulverized, extracted, and eventually sorbicillactone A was isolated using a large-scale chromatographic column (six meters long!). This provided 50 g to 100 g of the pure agent ready for large pre-clinical studies. These investigations are still underway, and we wait with anticipation to see how things proceed with sorbicillactone A.

Gerhard Bringmann and Dirk Wunderlich

Additional Literature

Bringmann G, Lang G, GuIder TAM, Tsuruta H, Mühlbacher J, Maksimenka K, Steffens S, Schaumann K, Stöhr R, Wiese J, Imhoff JF, Perovic-Ottstadt S, Boreiko O, Müller WEG: The first sorbicillinoid alkaloids, the antileukemic sorbicillactones A and B, from a sponge-derived *Penicillium chrysogenum* strain (2005), Tetrahedron 61(30), 7252-7265

Bringmann G, Lang G, Mühlbacher J, Schaumann K, Steffens S, Rytik PG, Hentschel U, Morschhäuser J, Müller WEG: Sorbicillactone A: a structurally unprecedented bioactive novel-type alkaloid from a sponge-derived fungus. In: Sponges (Porifers) (2003), Müller WEG (Hrsg.), Springer-Verlag, Berlin, pp. 231-253

Bringmann G, Lang G: Novel marine natural products. Part 3. Full absolute stereostructures of natural products directly from crude extracts: the HPLC-MS/MS-NMR-CD „triad“. In: Sponges (Porifers) (2003), Müller WEG (Hrsg.), Springer-Verlag, Berlin, pp. 89-116

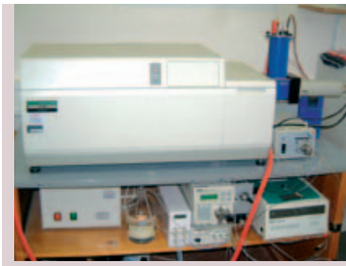
Bringmann G, Gulder TAM, Lang G, Schmitt S, Stöhr R, Wiese J, Nagel K, Imhoff JF: Large-scale production of the antileukemic marine natural product sorbicillactone A (2007), Marine Drugs 5, 23-30

Links on the Web

Homepage Prof. Bringmann:
www-organik.chemie.uni-wuerzburg.de/ak_bring

Bundesministerium für Bildung und Forschung (BMBF)
www.bmbf.de/press/1157.php

Bruker Homepage
www.bdal.de



LC-CD for the acquisition of CD spectra of compounds. This is used in tandem with computer-chemical methods to elucidate the exact configuration of chiral molecules.