Chapter 1

Introduction and Scope of the Thesis
**Introduction**

1. **Bioactivity/bioaffinity profiling of mixtures in drug discovery**

1.1. **Introduction**

Bioactivity describes the property of a compound or a set of compounds to modulate biochemical and physiological functions of living organisms. Identification of bioactive compounds in mixtures, such as natural extracts or metabolic mixtures, is an important challenge in different scientific and application fields that are often highly intertwined. Most of the times, bioactive compounds are present in very complex mixtures. These mixtures can at the same time contain multiple bioactives with different mechanisms of action or with synergistic and/or antagonistic effects. Bioactivity profiling of a mixture of compounds means bioactivity assessment of mixture constituents and identification of possibly present bioactives. Bioactivity profiling is generally directed at finding bioactives for a particular target, most often a drug target. Commonly, it is used in different screening programs that aim at the identification of bioactive compounds from different sample sources. For example, in attempt to discover new lead compounds, pharmaceutical industry as well as academic groups interested in drug discovery is dealing with analysis of various bioactive mixtures such as combinatorial synthetic libraries and natural products. The analysis of metabolic mixtures of drugs generated either in vivo or in vitro for potential bioactive metabolites represents another very important application of bioactivity profiling in drug discovery. Furthermore, bioactivity profiling of mixtures is important in toxicology, forensics, environmental and food analysis.

An important tool in discovery of new lead compounds is high-throughput screening (HTS). HTS is the process of testing large libraries of diverse chemical structures against disease targets to identify 'hits', i.e., bioactive compounds that modulate a particular biochemical pathway. HTS assays are performed in 96, 384, or 1536 well microtiter plates. Robotics, liquid handling devices, and sensitive detectors are brought together with advanced data processing and control software to optimize sample throughput. Based on ligand-target interactions, HTS is characterized by its simplicity, rapidness, low cost, and high efficiency, as well as a high information outcome. HTS has been developed to rapidly conduct millions of chemical, genetic, and/or pharmacological tests in order to identify potential drug candidates (1). HTS is an excellent tool for testing the bioactivity of large libraries of pure compounds. However, the identification of one or more bioactive components in a mixture requires a combination of analytical strategies for separation and chemical detection, often involving mass spectrometry (MS) and/or nuclear magnetic resonance spectroscopy (NMR), with biological assays for the bioactivity detection. Very often, bioactive mixtures are analyzed by their bioaffinity rather than their bioactivity. Bioaffinity represents the property of a compound to bind to the target protein and it is generally considered as the pre-condition for the compound to cause a biological effect as the bioactivity of a compound is the result of its binding to the target. Prior to
bioactivity assessment, extraction from the starting material and pre-concentration of the extract may be needed, for example in analysis of natural products of different origin with bioactivity ‘trapped’ in sometimes very complex matrixes. Finally, the correct attribution of bioactivity to a certain chemical structure has to be confirmed by additional tests and full characterization of the bioactive compound is required, both chemically and pharmacologically.

1.2. Bioassay-guided fractionation

Over the years, different strategies have been developed for bioactivity analysis of complex mixtures and identification of the bioactive components. The oldest and most frequently used approach is the (bio)assay-guided fractionation (BGF), which was introduced in drug discovery processes from natural extracts (originating from plant extracts). In parallel, similar approaches were developed and used under different names in other fields with interest in bioactivity profiling of complex mixtures (2). Maybe, the best known example in this regard is the effect-directed analysis (EDA) used for bioactivity assessment and identification of toxic pollutants present in environmental samples. Since the main focus of this thesis is on the strategies used in drug discovery, bioactivity profiling in other research fields will not be discussed further. Strategies used in these other fields were recently thoroughly reviewed by Jonker et al (3).

In drug discovery, the BGF approach has been successfully used for identification of many lead compounds from natural products, such as extracts from plants and animal venoms. The traditional BGF is a multistep approach which most often consists of four steps after optional sample preparation: 1) initial separation of the mixture of interest followed low-resolution fractionation (microfractionation), 2) bioassaying of collected fractions and identification of the bioactive fraction(s), 3) an additional (preferably orthogonal) separation step, followed by fractionation and bioactivity assessment performed on the bioactive fraction to isolate the pure compound for chemical analysis, and 4) (chemical analysis and) structure elucidation of the bioactive compound. In many cases, after two rounds of fractionation, the fractions are still too complex for bioactive compound identification and further rounds of fractionations are needed. The sample preparation that precedes BGF can be simple dilution of a lyophilized sample with optional centrifugation, as is often done in case of animal venoms. In case of plant material, the sample preparation step usually includes grinding and sawing of starting material followed by subsequent extraction of the bioactive sample using different organic solvents. The choice of the initial separation highly depends on the nature of the sample analyzed, but most often it is performed using liquid chromatography (LC) based on reversed-phase (RPLC), size-exclusion (SEC) or ion-exchange (IEX) separation mechanisms. SEC and IEX are especially useful for the initial separation of mixtures mainly containing peptides and proteins such as animal venoms (4,5). Other approaches include gel chromatography followed by LC (6) or preparative TLC (7). Subsequent fractionation of the separated mixture, also known as microfractionation, is performed manually in low resolution (long time intervals for
fraction collection), generally resulting in ≥1 mL fractions. The bioactivity testing of the fractions collected is usually performed after freeze-drying and re-constitution in a bioassay-compatible solvent, and it may include multiple up-concentration or dilution pipetting steps. From our point-of-view (see below), this approach is to be considered as off-line bioactivity testing since the bioassays are performed separately from the separation and detection. A variety of in vivo and in vitro bioassays are employed for the bioactivity assessment. Traditionally, bioassays included testing of the fractions on isolated animal organs, tissues or testing on whole animals. With the development of cell culturing techniques and the possibility of isolation and/or overexpression of specific drug targets, such as different enzymes and receptors, the choice and the possibilities of bioassays have drastically increased. Many successful applications of various bioassays using the BGF approach have been reported since (4,7−12). When one or more bioactive fractions are found, efforts are put in chemical analysis of the bioactive compound. Depending on the complexity of the bioactive fractions, an orthogonal separation can be performed with another round of microfractionation and bioassaying. In some cases, only one step of separation, fractionation and bioactivity testing is needed. After sufficiently pure samples have been obtained, the chemical analysis of the bioactive compound is performed using high-resolution MS and NMR. MS is widely used because of its high sensitivity and the amount of information it provides. In the case of small molecules, UV-VIS data (absorbance of ultraviolet or visible light) obtained with photo-diode array (PDA) detectors can be useful in dereplication (identification of already known compounds) and assigning of an unknown bioactive compound to a certain chemical class (13). However, NMR is most often needed for the final confirmation of the structure. This requires relatively high amounts of pure compounds. In case of bioactive peptides and proteins, MS-based proteomics approaches are widely used for amino-acid sequencing, and for the characterization of post-translational modifications and the positions of cysteine bridges. However, MS-based proteomics approaches do not provide full structure characterization. There may be issues with distinguishing between the isomeric amino acids leucine and isoleucine. For characterization of the 3D structures of proteins, other techniques like X-ray crystallography, electron microscopy, and NMR are needed. Huge efforts put in isolation of sufficient material for NMR prolong the time needed for identification of the bioactive. Finally, full chemical and pharmacological characterization of the compound is performed. Compounds obtained from natural products such as plants and microorganisms using the BGF approach have often reached the market in their original form. However, next to that, semi-synthetic approaches, pharmacophore identification and structure-activity relationship (SAR) studies on these compounds resulted in many new drugs with improved properties. This is for instance the case for various betalactam (14,15) and aminoglycoside(16) antibiotics and statins (17).
1.3. **Pre-column and on-column approaches for bioaffinity profiling of mixtures**

Even though highly successful and still widely used, the BGF approach is very laborious and time-consuming. Therefore, many academic groups have put significant effort in the development of alternative approaches for analysis of bioactive mixtures (18). Part of these advancements were due to advances in separation sciences, in hyphenation of separation technologies to MS, and especially in hyphenation of bioassays to separation techniques and MS detection (19). Approaches based on protein-ligand affinity separation will be briefly discussed here first (see the review of Jonker et al. (18) for an elaborate discussion).

Different techniques based on ultrafiltration, size-exclusion separation, and centrifugation can be used to analyze affinity of compounds binding to a solubilized target. Continuous ultrafiltration (20,21) is a technique for bioaffinity screening of mixtures by injecting the target protein into an ultrafiltration chamber, featuring a molecular-weight cut-off membrane, after which a bioactive mixture can be pumped through the chamber. After washing, a disruption step or sufficient elution time for the bound ligands to dissociate from the target protein allows processing of the eluted ligands towards MS for identification. This technique has been applied in a HTS-format for screening of a compound library for affinity towards a *Streptococcal* enzyme (22). Furthermore, on-line coupling of ultrafiltration with LC–MS was developed to screen natural extracts and other mixtures for bioaffinity towards the relevant targets (23–25). Pulsed ultrafiltration is a further development and improvement of this ultrafiltration approach (26). In pulsed ultrafiltration, continuous infusion is replaced by injection of a small amount of sample into the ultrafiltration unit. Molecules that do not bind to the target protein are flushed away through the molecular-weight cut-off membrane using a continuous buffer flow. Again, sufficient elution time or the use of a dissociation buffer allow migration of the ligands towards the MS for their subsequent identification. Pulsed ultrafiltration has been applied in a metabolic stability study (27) and in the screening for inhibitors of the retinoid X receptor (28). In the field of natural extracts screening, the use of pulsed ultrafiltration has been described for affinity selection of cyclooxygenase inhibitors from medicinal plants (29). SEC coupled to MS can be applied as an affinity selection method to screen binding of multiple ligands to a receptor (30,31). This technique is based on the incubation of mixtures of ligands with a target protein followed by a rapid SEC purification of the ligand-target complexes. A subsequent disruption step, commonly in continuous-flow format, precedes either direct analysis of the released ligands by MS, or a final SPE–(LC)–MS analysis. The SEC-based affinity trapping approach was developed in an automated HTS-format by Neogenesis and commercialized as the Automated Ligand Identification System (ALIS) (32). In ALIS, one uses a small on-line SEC affinity purification coupled to MS to screen the ligands binding to the relevant target protein. The method has been successfully used as an automated SEC–LC–MS setup to screen for ligands of lipid phosphatase (33) and the muscarinic M2 acetylcholine receptor (34). Similar methodologies were reported to study ligand binding to G protein-coupled receptors (GPCRs), an important class of drug targets (35). Another similar approach utilizing in-solution ligand binding to a
target protein is based on His-tagged target proteins with high-affinity for immobilized metal affinity columns (IMAC). Protein-ligand trapping on an IMAC column is followed by column washing with a disruption eluent allowing for release of the ligands from their target protein and subsequent identification (36). Alternatively, magnetic beads with IMAC properties have been used for ligand fishing in an off-line format (37,38). An online format of this technique has also been developed. Here, the process of washing and target protein-ligand disruption occurs on-line in an analytical chromatography setup by switching the eluents applied over the beads. Beads are trapped and finally released from the on-line system by use of a retractable magnet (39,40).

The MS binding-assay approach (41) is a technique based on a conventional radioligand binding assay in which the radioactivity detection of bound ligands is replaced by MS detection. In short, ligands are incubated with the target protein in the presence of the radiolabeled ligand. The bound fraction is separated from the unbound one using vacuum filtration. The fraction with bound ligand is subjected to a disruption buffer, to release the bound ligands that are subsequently detected using MS. MS binding assays were developed for analysis of inhibitors of γ-aminobutyric acid (GABA) transporters (42).

Other approaches include frontal and zonal affinity chromatography (43), which are based on chromatographic affinity columns with an immobilized target protein. In frontal affinity analysis, a mixture of interest is continuously infused onto the column, while in the zonal affinity analysis the mixture of interest is injected as a plug onto the column. In both cases, an MS is used as detector. The affinity of the ligands for the immobilized drug target on the column determines the elution order: a longer elution time implies a higher affinity. Affinity chromatography techniques are applied to study binding affinities of the compounds (44–46) as well as in the screening of mixtures of compounds (47). Alternatively, capillary electrophoresis can be applied as separation technique to determine affinity interactions of proteins and ligands. Affinity capillary electrophoresis (ACE) has been used to study affinity and purity of llama-derived antibodies, also referred to as nanobodies (48). Immunoaffinity capillary electrophoresis (IA-CE) has been applied to study the affinity of peptides in complex matrices (49). The extensive review of application of ACE and IA-CE can be found in (50).

1.4. Post-column bioactivity profiling of mixtures
As discussed above, in the originally developed BGF approach, the bioactivity profiling of the mixture is performed in an off-line post-column fashion, i.e., the bioactivity/bioaffinity of the collected fractions is assessed after a separation step that is decoupled from the chemical analysis. On the other hand, the described alternative approaches for bioaffinity analysis are based on either pre-column bioassays followed by separation, or on-column bioaffinity analysis and separation. In these cases, the bioassays are coupled to the chemical detection, e.g., MS identification, either in an off-line or an on-line mode. In the last two decades, our group has made considerable efforts in the development of advanced
strategies for post-column bioactivity profiling of mixtures (19). These strategies are based on on-line (and at-line) post-column screening of complex mixtures and comprise parallel bioactivity/bioaffinity assessment and chemical analysis, which in turn allows for more efficient bioactivity profiling of mixtures. An important advantage of post-column analytics over the pre-column and on-column bioassay strategies is the possibility to also implement bioactivity analysis rather than only bioaffinity assessment. This implies that bioactivity analysis using functional cellular bioassays (51) and bioassays based on enzyme activity (52,53) are within the scope.

The concepts of on-line post-column screening (54,55) and advances in this field (19) have been reviewed in detail. Here, the discussion is focused at the on-line bioaffinity/bioactivity analysis with parallel MS detection. The basic principle of current on-line screening approaches lays in coupling of RPLC separation to continuous-flow bioassays (mostly fluorescence-based assays) with parallel MS detection. Parallel bioassaying and MS detection are enabled by a direct post-column split of the LC effluent in an approximately 1:9 ratio in most of the methods developed. In this case, the smaller fraction (10%) is subjected to the bioassay and larger fraction is directed to the MS. Typically, two types of bioassays are used: bioassays based on the enzymatic conversion of a fluorogenic substrate into a fluorescence product (56,57) and bioassays based on fluorescence enhancement upon binding of a tracer ligand to its target receptor or binding protein (58,59). The mixing of the LC effluent and the bioassay reagents is achieved by using two superloops followed by a continuous-flow incubation in a reaction coil after each superloop. A superloop represents a syringe through which the corresponding solution is delivered using an external LC pump. The first superloop and reaction coil allow interaction of the separated compounds with the target molecule (an enzyme or a receptor), while the second superloop introduces the fluorescent ligand that interacts with the ligand-target mixture. The flow rates and tubing dimensions determine the incubation times in the reaction coils and are limited to a few minutes in order to avoid severe post-column band broadening. In some cases, additional off-line NMR analysis was incorporated for full chemical characterization of the bioactive compounds (60). A schematic overview of a typical on-line post-column setup is shown in Fig. 1.

![Fig. 1. Schematic overview of a typical on-line post-column screening platform.](image-url)
The on-line screening approach was successfully demonstrated in proof of principle studies as well as in screening of natural products (61) and metabolic mixtures (60). However, there are some important drawbacks and limitations of this approach. The main limitation of on-line screening is its restriction to bioassays with short incubation times. Thus, for example enzymatic assays with long incubation times and also elaborate radioligand binding affinity assays (that also comprise many pipetting steps including a filtration step) cannot be performed in an on-line approach. Obviously, cell-based assays such as gene-reporter assays requiring long incubation times cannot be used in this setup either. Other disadvantages include the necessity for modification and adaptation of the traditional microtiter plate based assays for on-line screening approaches, compatibility issues between LC mobile-phase composition and the solvent composition permitted in the bioassay, especially with respect to organic solvent content, and the need of (relatively) high concentrations of target receptor or enzyme to provide a robust system with an adequate dynamic range. Finally, suitable fluorescence enhancement ligands or substrates that are enzymatically converted into a fluorescent product are needed for the bioassay.

To address some of the drawbacks of the on-line screening approach, various modifications and alternative approaches have been developed. For example, the possibility of a direct MS bioassay readout has been explored for the analysis of mixtures towards targets for which fluorescence tracers or ligands are not available (62), but this approach did not find broad application due to issues with sensitivity, matrix effects, and ion-source contamination (19). For samples only available in small amounts, such as animal venoms, the on-line approach was miniaturized towards using a nanoLC separation with parallel MS detection and the bioassay being performed on a microfluidic chips (63). This allowed sample injections of ≤5 µg via a 10 to 500 nL injection loop in comparison to the 250 µg injections in the conventional on-line set-up using a 10 to 50 µL injection loop. The miniaturized on-line approach was applied to the screening of snake and cone snail venoms, and toad skin extracts for bioaffinity assessment towards nicotinic acetylcholine receptors (nAChR) by using its binding protein analogue (64–66). A schematic overview of a microfluidic on-line post-column screening setup is shown in Fig. 2.

![Fig. 2. A schematic overview of a microfluidic on-line post-column platform.](image-url)
Finally, the limitation regarding the choice of bioassays was tackled by introducing a post-column high-resolution fractionation onto microtiter plates prior to bioassaying. This approach was named at-line post-column bioaffinity profiling and, to some extent, was highly similar to the on-line approach. In this approach, the LC-effluent was split post-column: the larger fraction was guided to a modified autoinjector that served as the fractionation system, whereas the smaller fraction is directed to the MS. Prior to fractionation, the bioassay reagents were mixed with the LC effluent using two syringe pumps to deliver the reagents in a continuous-flow format. This approach allowed for coupling of the LC separation to more types of bioassays since there are no limitations in incubation time (67). Further development of the at-line fractionation approach was advanced by introducing the original BGF concept, i.e., the fractionation step was introduced directly after the flow split followed by a solvent-evaporation step (freeze-drying) and subsequent bioassaying. Compared to the original BGF approach, advancements are direct fractionation and solvent evaporation on microtiter well plates and a parallel post-column split to MS. In this way, potential bioassay interferences related to the presence of organic solvents used in the preceding RPLC separation are eliminated. These modifications resulted in a technique known as at-line nanofractionation and were for example applied in 96- and in 384-well plate formats to study metabolic mixtures of ligands towards the histamine receptor H4 (51,68). In these studies, both bioaffinity and bioactivity of metabolic mixtures were assessed using radioligand binding assays and a β-galactosidase reporter gene assay. Initially, at-line nanofractionation was developed as a decoupled approach. Namely, two separate chromatographic runs were performed for bioaffinity/bioactivity profiling and for MS detection. One LC–UV run was hyphenated to MS detection and the other was used in nanofractionation coupled to the bioassay. A schematic overview of the at-line nanofractionation approach is shown in Fig. 3.

Fig. 3. A schematic overview of an at-line nanofractionation approach.

In this thesis, further development of the at-line nanofractionation approach is described. The approach was used to study metabolic mixtures of ligands targeting chemokine receptors, a class of GPCRs comprising important drug targets in drug discovery for the treatment of various inflammatory and immunological diseases. Furthermore, the developed approach was advanced by introducing a flow split for parallel bioassaying and MS detection. It was applied to metabolic profiling of lead compounds binding to
the chemokine receptors CXCR1, CXCR2, and/or CXCR3. Additionally, the technique was applied to natural product screening with the focus on bioactive peptides, e.g., for screening of snake venoms for bioactive substances towards thrombin, factor Xa and the angiotensin-converting enzyme (ACE), which are drug targets for the pathologies of the cardiovascular system.

1.5. Natural products in drug discovery

The use of natural products in the treatment of a large number of diseases goes deep in the past, probably as far as humans exist. The medicinal and/or toxic properties of primarily plants, but also animal extracts were known to old civilizations who documented their use in these purposes. The earliest records on medical use of plants date from around 2600 BCE and were made in ancient Mesopotamia. Other old documentation on natural medication encompass Egyptian “Ebers Papyrus” from around 1500 BCE, Chinese “Materia Medica” with first written records dating from around 1100 BCE, and Indian “Ayurveda” dating from around 1000 BCE. The grounds of the medical and pharmaceutical knowledge in the Western world were set by ancient Greeks and Romans, followed by the middle age Arab influences who brought knowledge of the West and East together (69). However, the isolation of active substances from natural products and understanding of their pharmacological activity has a more recent history, beginning in 1804 with the isolation of morphine from Opium crudum produced by cut seed pods of the poppy, Papaver somniferum (70). The isolation, purification, pharmacological characterization and precise dosing of many pharmaceutically active compounds from natural origin have since then followed, thereby setting up the grounds of modern medicine. Development of analytical methodologies played an important role here as the success rate of new chemical entities (NCE) discovered increased with the advances of analytical chemistry in separation and in identification techniques as well as advances in the equipment used. As mentioned in the section dedicated to the bioactivity profiling of mixtures, the BGF approach played a key role in identification of NCEs from natural products. By 1990, more than 80% of all drugs on the market were natural products, semi-synthetic analogues of natural products or compounds synthesized based on the pharmacophores of natural products (71). Furthermore, many naturally occurring compounds served as pharmacological tools to study different receptors and diseases. For example, five of the seven different sodium channels were characterized using venom-based compounds (72). Besides, the names for the nicotinic and muscarinic acetylcholine receptors were derived from the plant alkaloids selectively binding to these receptors. Even though the development of combinatorial chemistry and HTS caused profit-driven pharmaceutical industry to decrease its interest in natural products a few decades ago, they remain an important source of lead compounds. The interest in natural products as sources for NCEs is gaining interest again since the output from traditional pipelines is decreasing, especially the last decade. Moreover, the development of analytical techniques and screening methods together with the use of new genome mining approaches enabled faster dereplication and identification of new
bioactive compounds from natural products and consequently contributed to the revival of the interest in natural products as potential drugs. In the period between 1981 and 2010, 50% of NCEs approved in the category of small molecules were natural product related (73). In 2010 only, ten out of twenty new small molecular drugs were of natural origin (73). The role of plants, microorganisms and venoms as sources of active substances in drug discovery will be briefly discussed in the next paragraphs. For more extensive information on this topic, a number of reviews can be consulted (74–78).

1.6. **Plants as source of active substances**

Traditional medicine based on the use of plants has had a significant impact on modern medicine and drug discovery. In fact, the isolation of morphine at the very beginning of 19th century represents the first successful drug discovery effort in the history of medicine and pharmacy. This was followed by the isolation of a large number of other active substances from plants that were further studied and applied in treatment of the entire diapason of diseases. The original chemical structures of many of these compounds still represent important drugs such as morphine, digoxin, digitoxin and theophylline (79). Morphine is used for management of severe pain (80,81), while the cardiac glycosides digoxin and digitoxin, isolated from *Digitalis* species, are used as antiarrhythmic agents (82). Theophylline, isolated from *Teobroma cacao* plant, is a bronchodilator used to treat patients with asthma and chronic obstructive pulmonary disease (COPD) (83). There are also many examples of drugs and whole chemical classes of drugs that are based on the compounds found in plants. For example, another bronchodilator, tiotropium bromide that acts as the antimuscarinic agent, was developed based on atropine, an alkaloid found in *Atropa belladonna* (84,85). Papaverine, another active substance isolated from the poppy plant, served as a starting point for development of verapamil, an L-type calcium-channel blocking agent used in treatment of different pathologies related to the heart (79,86). Metformin and other bisguanidines that are used in the management of diabetes type II and metabolic syndrome were synthesized based on galegine, an active compound found in *Galega officinalis* (87). Local anesthetics were developed based on the pharmacophore of cocaine, which was isolated from the leaves of coca, *Erythroxylum coca* (88). Atracurium, a neuromuscular blocker used to control the muscle tonus of patients under general anesthesia, have a base in d-tubocurarine, an alkaloid found in various plants of Menispermaceae family (89). Besides, the role of ethnomedicine is based on the use of whole plants or their parts as herbal remedies and stays important in many cultures. They are used in primary health protection in less developed countries or as adjuvant or alternative therapy in developed countries. Moreover, plant-based drug discovery may benefit from current efforts put in the analysis and characterization of bioactive compounds from plants used in Traditional Chinese Medicine (TCM). Studies conducted so far on screening of TCM components for new leads show promising results and compounds with anticancer, anti-inflammatory, antiviral, anti-malarial and stroke prevention activity have been reported (90–92). Furthermore, clinical studies on the influence of TCM on the
overall condition of patients with cancer show that TCM as adjuvant to chemotherapy seems to improve the quality of life of these patients (93).

1.7. **Microorganisms as source of active substances**

The serendipitous discovery of antibacterial properties of penicillin from a *Penicillium* fungi extract in 1928 by Alexander Fleming and its purification in 1940s initiated a revolution in the treatment of bacterial infections and opened an era of drugs isolated from microorganisms (94). In fact, microorganisms are the second most important source of active substances. Penicillin, which is still used as a drug, showed a relatively wide spectrum of antibacterial action, but poor bioavailability upon oral administration due to the sensitivity to stomach acid. The introduction of different functional groups using semi-synthetic approaches led to development of other penicillins that are stable for oral administration and have an even wider spectrum of action or are targeting specific types of bacteria (95). Interestingly, microorganisms yielded most of all known antibiotic chemical classes. Throughout the years, many more antibacterial agents from fungi and bacteria, such as tetracyclines (96), cephalosporins (95), aminoglycosides (97), and macrolides (98), were discovered and developed into drugs followed by their semi-synthetic analogues. Microorganisms also served as source for the development of some hypolipidemics, anticancer drugs, immunosuppressives, anthelminitics and antiparasitics (99). Many of these drugs made a significant impact in the treatment of the corresponding diseases, such as statins in the treatment of dyslipidemia or cyclosporine A (CSA) in the prevention of allograft rejection. CSA was isolated in 1970 from the *Tolypocladium inflatum* fungi and made a turn in the therapy of allograft rejection due to the highly specific mechanism of action, serving as a base for a new generation of immunosuppressives (100). CSA is a cyclic peptide that consists of 11 amino acids and probably the best-known small peptide from natural products that made it to the clinic. The cyclic form and modifications of the side chains contribute to its relative stability and good bioavailability upon oral administration. This is an important property considering the fact that many drugs based on linear peptides require intravenous application due to the sensitivity to stomach acid. Therefore, the study of the influence of peptide cyclization and/or side chain modifications on their specificity, bioactivity and oral bioavailability may lead to the development of peptide drugs with improved properties allowing further development and wider application of peptidergic drugs (101).

Furthermore, microorganisms in general continue to be an important source of active substances for drug discovery considering the large number of unexplored species as well as the improvement in analytical techniques and increasing amount of genomic information. Finally, the growing concern world-wide regarding the increases in resistance of bacteria and multi-resistant bacteria to existing antibiotics gave a great impulse the last decade on both academic and pharma research in the discovery of new classes of antibacterial agents.
1.8. **Venoms as source of active substances**

Venoms are complex mixtures of peptides and proteins affecting various physiological processes upon delivery in the soft tissue of a prey organism in a process called envenomation. The bioactive components found in venoms are characterized by high potency and high selectivity for their target making venoms a valuable source of lead compounds in drug discovery. The use of venoms in treatment of some medical conditions dates back a long time (102). Snake venoms have been used to treat rheumatism, gastrointestinal and others disorders, while spider venoms have been known to possess anti-asthmatic and anticancer activity. Moreover, animal venoms represent a source of important pharmacological tools in receptor and disease studies. Venom-based drug discovery started with the isolation of peptide from the venom of Brazilian viper, *Bothrops jararaca*, with bradykinin potentiating properties (103,104). This ultimately led to the development of captopril (105), the first marketed ACE inhibitor. ACE is an enzyme involved in the regulation of blood pressure and its inhibition has been successfully used in treatment of hypertension and heart failure. Since the discovery of captopril, the interest in bioactive compounds from venoms has increased resulting in six new drugs and many potential candidates (102). For example, two peptides found in snake venoms were developed into epifibatide and tirofiban, antiplatelet agents acting as glycoprotein IIb/IIIa inhibitors. Epifibatide (Integrilin) (106) is a cyclic heptapeptide derived from a snake desintegrin isolated from the venom of *Sistrurus miliarius barbouri*, the southeastern pygmy rattlesnake, while tirofiban (Aggrastat) (107) is a small molecule that was based on a peptide found in the venom of *Echis carinatus*, the saw-scaled vipers. Moreover, snake venoms are constantly being explored as sources of leads in drug discovery aimed at a wide range of syndromes and diseases (102,108−112). Peptides from snake venoms are also used as research tools (113). Alpha-bungarotoxin, a neurotoxin isolated from the venom of Taiwan banded krait, *Bungarus multicinctus*, is a competitive and irreversible antagonist of the nAChR that is used in studies of neuromuscular junctions, nAChRs and it contributed to understanding of multiple sclerosis disease (114). One particular source of venoms relevant in both medicine and as source of pharmacological tools are cone snails (115). Marine cone snails are predatory sea snails that immobilize their prey via a venomous harpoon sting. Venom of cone snails provides an excellent source of bioactive peptides for discovery of new lead compounds (116−118). The biggest success story is related to the development of ziconotide (Prialt) (119), an exact synthetic replicate of a ω-conotoxin originally found in the cone snail *Conus magus*. This highly selective blocker of voltage-sensitive calcium channels was approved for the treatment of chronic pain. Vetter and Lewis (118) highlighted other conopeptides acting on sodium ion channels and various receptors (i.e., acetylcholine, neurotensin, noradrenaline and 5-hydroxytryptamine receptors). However, little is known about their progress towards clinical development. A μ-conotoxin (XEP-018) from *Conus consors* venom was reported to be moving into preclinical development on account of its particularly long duration of action (120) but its current status is unknown. Another example of successful drug
development from animals venoms is exenetide, which was isolated from the venom of Gila monster, *Heloderma suspectum* (121). Exenetide is approved for treatment of type II diabetes based on its similarity to the glucagon-like peptide-1, which plays a role in the control of appetite and glucose levels in blood (122,123). Many other venomous species, such as spiders and sea anemones are currently being studied for the potential therapeutic use (72,124,125). The possibilities include bioactives that act through binding to calcium, potassium or chloride channels, inhibition of nAChRs, inhibition of thrombin or factor Xa, inhibition of noradrenaline transporter, blocking N-methyl-D-aspartate (NMDA) receptors and many others (126).

Further exploration of the therapeutic potential of venoms will open new opportunities for drug discovery. The huge potential for finding new leads in venoms and developing them into drugs lays in the diversity of venomous animals and the compounds they contain. Moreover, the venom-based peptides have the favorable characteristics of being highly selective and potent towards respective drug targets, which makes them very interesting for discovery of new leads. Importantly, the high content of disulfide bridges of some venomous peptides, securing the three dimensional structure needed for very potent and specific protein interactions, provides stability against proteolytic degradation and adds to their chemical and thermal stability. Furthermore, the small size of venom peptides-derived drugs contributes to their low immunogenicity and opens the possibility for different administration routs (127).

1.9. **Bioactivity/bioaffinity profiling of metabolic mixtures in drug discovery**

Metabolic profiling and bioactivity/bioaffinity assessment of drug metabolites and of lead compounds is important for full pharmacological and toxicological characterization. The metabolic stability of a drug is important because it highly influences its bioavailability and consequently determines the formulation and route of administration. For example, highly lipophilic molecules (e.g., human estrogens and androgens) are characterized by the first-pass effect, i.e., extensive liver metabolism resulting in a poor oral bioavailability. Therefore, before more stable molecules than estradiol (e.g., ethinylestradiol) were synthesized to allow oral administration of estrogens and improve bioavailability (128), estrogens were applied as intramuscular injections. Besides, drug metabolism may result in the formation of one or more bioactive metabolites with changes in selectivity, bioaffinity and/or bioactivity. This may further lead to the development of serious side effects, e.g., toxicity due to a longer exposure to the drug and/or off-target side effects. Therefore, bioactivity profiling of the individual metabolites of a lead compound represents an important step in drug discovery. Moreover, the bioactivity profile of metabolites of a particular drug gives valuable information about the liable 'hotspots' (important functional groups) in the parent compound, which can in turn be used for structural improvements in the design of new drugs. In fact, there are examples where the bioactive metabolite of a known drug reached the market (129) (e.g., benzodiazepine oxazepam, being a metabolite of diazepam).
One of the important reasons for drug discontinuation is related to its metabolism (130). During the selection of candidates for the preclinical studies, all potential lead compounds are screened for their metabolic stability. This is mostly conducted using LC–MS analysis of the metabolic mixture generated using various methods (131). Full chemical and pharmacological characterization of the metabolites is usually only performed at a later stage in drug discovery/development, since the purification of all metabolites for MS and NMR based structure elucidation as well as full pharmacological characterization is an elaborate and costly process. Besides, the rationale behind late full characterization of drug metabolites is that in general more relevant information is obtained in the preclinical studies after exposing an experimental animal to a drug. That way, all metabolic pathways can be assessed and generally good prediction models of drug metabolism in humans can be made. However, the discovery of an unfavorable metabolic profile connected to the formation of bioactive metabolites at this stage will most probably result in discontinuation of development of the lead compound. This further means a significant loss of time and money invested in the particular lead compound. Therefore, the pharmaceutical industry could benefit from implementation of early stage bioactivity/bioaffinity profiling of metabolites in drug discovery to lower the risk that a lead candidate will fail to reach clinical studies and consequently cut the attrition costs. As previously discussed, the benefit of early stage metabolic and bioactivity/bioaffinity profiling is also reflected in valuable information provided to medicinal chemists allowing them to design structural improvements of promising lead compounds. The development of advanced analytical methodologies that integrate bioassays and chemical analysis is necessary to accomplish this. Many successful attempts have already been made in that respect, as briefly discussed in section 1.1. However, the at-line nanofractionation methodology has an important advantage compared to other approaches, which is the possibility of coupling the LC–MS analysis to most types of bioassays available. The application of the at-line nanofractionation methodology in bioaffinity profiling of metabolic mixtures towards chemokine receptors is demonstrated in this thesis.
Scope of the thesis

This thesis explores the possibility for implementation of the at-line nanofractionation methodology into drug discovery screening platforms. The methodology aims at bioactivity/bioaffinity profiling of complex samples; metabolic mixtures and natural products in particular.

Chapter 1 gives a general introduction on the importance of bioactivity/bioaffinity profiling of complex mixtures in drug discovery, and presents an overview of the currently available analytical strategies applied in this field including pre-column, on-column and post-column bioaffinity and bioactivity screening approaches. Finally, the role of natural products and of metabolic profiling in drug discovery is briefly summarized.

In Chapters 2 and 3, the at-line nanofractionation methodology was developed and demonstrated for metabolic profiling and parallel bioaffinity assessment of low-molecular weight (LMW) ligands towards the chemokine receptors CXCR1, CXCR2, and CXCR3. Chemokine receptors belong to the class of GPCRs and represent important drug targets for the treatment of immune-related diseases such as rheumatoid arthritis, asthma, COPD, cancer and allograft rejection. Currently, approximately 40 antagonists of chemokine receptors are commencing clinical studies, while many others are at the early stage of drug discovery. For metabolic profiling of LMW CXCR ligands, metabolic mixtures were generated using pig liver microsomes. The bioaffinity of these ligands and their metabolites was assessed using radioligand binding assays in 96-well plate format. Chapter 2 describes the development and optimization of a decoupled at-line nanofractionation approach, i.e., MS identification and bioaffinity profiling are performed after two separate RPLC runs. The method was further successfully applied to analysis of metabolic mixtures of VUF11211 and NBI-74330, which are LMW allosteric modulators of the CXCR3 receptor. This resulted in structure elucidation and bioactivity profiling of the metabolites formed. Moreover, the possibility to analyze chemokines as endogenous ligands was demonstrated. In Chapter 3, the method developed was extended to selectivity screening of metabolic mixtures towards CXCR2 versus the highly homologous CXCR1. Furthermore, a direct post-column flow split was introduced allowing parallel bioaffinity profiling and MS identification. The method was optimized for the analysis of both LMW ligands and chemokines, and was demonstrated to be applicable as fast analytical screening method in a small screening program on metabolic mixtures from a set of CXCR2 antagonists. Lastly, MK-7123, a dual high-affinity antagonist, was metabolically profiled and chemical structures of its metabolites were (partially) elucidated.

Chapters 4, 5 and 6 describe the integration of the at-line nanofractionation screening in natural products-based drug discovery. In Chapter 4, the at-line nanofractionation platform was developed in 384-well plate format for screening and identification of thrombin and factor Xa inhibitors in snake venoms. Both thrombin and factor Xa are enzymes involved in the coagulation cascade and valid drug targets for the treatment of abnormal coagulation in patients. In total, 39 different snake venoms were screened using
an RPLC separation coupled to fluorescence-based enzyme activity assays and parallel MS identification. Bioactivity profiles of the snakes screened were studied for the presence of inhibitors, but in most cases, only snake proteases were detected. In case inhibitors were detected, the bioactivity (i.e., a negative bioactivity peak) was assigned to the corresponding $m/z$-values obtained from the parallel MS measurements. Furthermore, the protease activity profiles (i.e., represented as positive bioactivity peaks) of the different snake venoms screened were compared in an attempt to understand the functional phylogeny.

In Chapter 5, the snake venom screening program was directed towards ACE inhibitors considering the known presence of peptides inhibiting ACE activity in snake venoms. As an advancement, the use of hydrophilic interaction liquid chromatography (HILIC) was investigated as a complementary tool to RPLC separation. All snakes were initially screened in RPLC mode and the most interesting candidates were subjected to re-screening with RPLC and an additional screening in HILIC mode. Since snake venoms comprise very complex samples, several constituents in a fraction may be responsible for the bioactivity. The additional use of the HILIC mode allowed narrowing down of the number of candidates responsible for the bioactivity per venom analyzed. Additionally, the bioactive fractions were, directly from the wells, subjected to nanoLC–MS/MS analysis, and peptide sequences were determined from the fragmentation patterns observed in the MS/MS spectra.

Since for many drug targets the bioactivity of bioactive compounds in mixtures can only be studied by using cellular bioassays, and since in many cases functional cellular responses are required, the implementation of such an assay format in the at-line nanofractionation screening approach was explored in Chapter 6. Therefore, at-line nanofractionation analytics was developed for screening of complex samples for bioactive compounds acting on the $\alpha_7$-nAChR, which is an important ligand-gated ion channel involved in diseases of central nervous system. For the bioactivity assessment, a calcium flux assay was optimized and used for the detection of agonists and for positive allosteric modulators of the $\alpha_7$-nAChR. For proof of principle purposes, the method was applied to screen an extract of a hallucinogen mushroom, both after RPLC and HILIC separation, resulting in identification of two bioactives.
References


Introduction and Scope of the thesis


80. Hansel, T. T., and Barnes, P. J. (2002) Tiotropium bromide: a novel once-daily anticholinergic bronchodilator for the treatment of COPD. *Drugs Today (Barc)* **38**, 585-600


C1

Introduction and Scope of the thesis


