universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

Metabolic effects of natural compounds on A549 lung carcinoma cells

verfasst von / submitted by

Sandra Panzalovic, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Magistra pharmaciae (Mag.pharm.)

Wien, 2023 / Vienna 2023

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 605
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Pharmazie
Betreut von / Supervisor:	Ass.-Prof. Karin Ortmayr, PhD

Table of Contents

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Dr. Karin Ortmayr, for their guidance and support throughout my master's program. Their expertise and patience helped me to complete this research and write this thesis.

I would also like to thank Univ.-Prof.Dr. Judith M. Rollinger for providing me with the opportunity to conduct my research at the Division of Pharmacognosy at the University of Vienna, and for all the resources and support they provided. I would also like to thank Univ. Prof. Mag. Dr. Sabine Glasl-Tazreiter and Univ.-Prof. Mag.pharm. Dr. Gottfried Reznicek for their insights and motivation and the whole Division of Pharmacognosy for serving on my thesis committee and providing helpful feedback and suggestions.

I am grateful to everyone who has supported me throughout this work. Without your help, guidance and patience, this thesis would not have been possible.

Abstract

In this thesis, the aim was the development of a general experimental approach to investigate the metabolic effects of natural compounds on human cells using metabolomics analysis. The concentration ranges of four natural compounds (solanine, helenin, taurine, and emodin) were optimized to investigate their effects on cell growth using A549 cells. The concentrations were carefully selected to avoid precipitation and cell death. Cell growth curves were generated for each concentration, and dose response curves were constructed to analyse the concentration-dependent effects. Another four natural compounds (genistein, berberine, agaric acid, and shikonin) were utilized to generate a deeper understanding of concentration range selection. Various effects on cell growth were discovered. Several natural compounds showed cell growth inhibiting activities (e.g., taurine, emodin) while other compounds showed little to no cell growth inhibiting effects (e.g., solanine, agaric acid) or led mainly to cell death (shikonin). The dose response curves revealed different response patterns for each compound, with solanine and agaric acid showing negative responses, and emodin exhibiting a slight sigmoidal pattern. The metabolomics analysis of the optimized concentration ranges of emodin and taurine revealed distinct metabolic changes compared to control samples. Principal Component Analysis (PCA) and volcano plots provided insights into the metabolic alterations induced by these compounds. One significant result was. Overall, the results demonstrated that the metabolic effects of different natural compounds differed from each other, they showed the improved consistency of the experiment with optimized concentration ranges and highlighted the potential of natural compounds as modulators of cell growth and metabolism.

Zusammenfassung

In dieser Arbeit lag der Hauptfokus auf der Entwicklung eines allgemeinen experimentellen Ansatzes zur Untersuchung der Stoffwechseleffekte von Naturstoffen auf menschliche Zellen mithilfe von Metabolomics-Analysen. Die Konzentrationsbereiche von vier Naturstoffen (Solanin, Helenin, Taurin und Emodin) wurden optimiert, um ihre Auswirkungen auf das Zellwachstum von A549-Zellen zu untersuchen. Die Konzentrationen wurden sorgfältig gewählt, um Ausfällungen und Zelltod zu vermeiden. Für jede Konzentration wurden Wachstumskurven erstellt und Dosis-Wirkungs-Kurven konstruiert, um die konzentrationsabhängigen Effekte zu analysieren. Vier weitere Naturstoffe (Genistein, Berberin, Agaricinsäure und Shikonin) wurden verwendet, um das Verständnis für die Auswahl des Konzentrationsbereichs zu vertiefen. Einige der Naturstoffe zeigten Zellwachstum-inhibierende Eigenschaften (z.B. Emodin, Taurin), während andere Naturstoffe wenig oder keine Inhibition vorwiesen (z.B. Agaricinsäure, Solanin) oder vorwiegend zu Zelltod führten (Shikonin). Die Dosis-Wirkungs-Kurven zeigten unterschiedliche Reaktionsmuster für jede Verbindung, wobei Solanin und Agaricinsäure negative Reaktionen zeigten und Emodin ein leicht sigmoidales Muster aufwies. Die Metabolomics-Analyse der optimierten Konzentrationsbereiche von Emodin und Taurin zeigte deutliche Stoffwechselveränderungen im Vergleich zu Kontrollproben. Die Hauptkomponentenanalyse (PCA) und Volcano-Plots lieferten Einblicke in die durch diese Verbindungen hervorgerufenen metabolischen Veränderungen. Insgesamt zeigten die Ergebnisse, dass sich die verschiedenen Naturstoffe in ihren metabolischen Effekten voneinander unterscheiden, sie demonstrierten eine verbesserte Konsistenz des Experiments mit optimierten Konzentrationsbereichen und verdeutlichten das Potenzial von Naturstoffen als Modulatoren von Zellwachstum und Stoffwechsel.

Aim of the Work

The aim of this thesis is to set up a general experimental approach that allows the research of metabolic effects of human cells under the influence of natural compounds by using metabolomics analysis. The anticipated objectives and aspirations behind this experimental setup were to achieve the capability to identify metabolic influences of natural compounds, leading to informative metabolic profiles that can be utilized to characterize the molecular effects and potentially identify targets associated with these compounds.

To make such an approach possible, it was of importance to investigate the optimal concentration ranges of the natural compounds, assess their impact on cell growth inhibition, and explore their effects on primary metabolism using metabolomics analysis.

The first step was to evaluate the effects of the compounds on cell growth and to identify any concentration-dependent patterns. The goal was to determine if there exists a general concentration range applicable to all compounds or if specific ranges are necessary for each compound.

In order to test the possibility to assess cell growth inhibition, the measurement of confluence was employed. By monitoring changes in confluence over time, the inhibitory effects of the compounds could be quantitatively evaluated. This approach allowed for the combined characterization of natural compound activity using growth inhibition- and metabolome profiling.

Furthermore, this thesis aimed to investigate whether the natural compounds have detectable effects on the primary metabolism through metabolomics. Here, we performed mass spectrometry-based metabolome profiling and compared the metabolic profiles of treated cells with control samples, expecting to identify significant changes that indicate effects of natural products on the metabolome of A549 cells. Additionally, the investigation aimed to determine if the effects on primary metabolism varied among the different compound groups. This analysis would provide insights into the potential modulation of primary metabolism by natural compounds, it would show if each compound class had characteristic metabolic effects and whether such effects differ depending on the specific compound class under investigation.

In summary, this work aimed to determine the concentration ranges suitable for the natural compounds, characterize cell growth inhibition using confluence measurements, and explore potential effects on primary metabolism through metabolomics analysis. By addressing these objectives, this work aimed to contribute to a deeper understanding of the cellular effects of these compounds and their potential as modulators of cell growth and metabolism.

Introduction

Natural compounds – A popular research field

The study of natural compounds has become increasingly popular in recent years due to several factors. Firstly, natural compounds have been used in traditional medicine for centuries and have a long history of safe and effective use. This has led to increased interest in natural compounds as potential sources of new drugs and therapies (Ganesan, 2008). Also, natural compounds are abundant in nature and can be found in a wide range of plant and animal sources, making them accessible and cost-effective (Newman and Cragg, 2012). Advances in technology have made it easier to isolate, identify, and study natural compounds, allowing researchers to explore their properties and potential applications (Li and Vederas, 2009).

In addition to their potential use in drug discovery, natural compounds have also been studied for their health benefits. There are many different substance classes that are commonly found in natural compounds, and they can have a variety of effects on the human body. Some of the most common substance classes include alkaloids, flavonoids, terpenoids, phenolic acids, and steroids (Saadeldeen *et al.*, 2020). Alkaloids are a diverse group of naturally occurring compounds that are found in many different plants. They can have a range of biological effects on the body, including pain relief, sedation, and stimulation of the central nervous system (Debnath *et al.*, 2018). Some well-known examples of alkaloids include caffeine, nicotine, and morphine (Bribi, 2018). Flavonoids are a large group of polyphenolic compounds that are found in many different plants. They have been shown to have antioxidant and anti-inflammatory properties, and they may also have a range of other health benefits, such as reducing the risk of cancer and cardiovascular disease (Heim, Tagliaferro and Bobilya, 2002). Terpenoids are a diverse group of compounds that are common across various plants, and they can have a range of biological effects on the body. Some terpenoids, such as carotenoids, have antioxidant properties, while others, such as menthol, can have analgesic and anti-inflammatory effects (Bohlmann and Keeling, 2008). Phenolic acids are a group of compounds that are prevalent in multiple plant species, and they have been shown to have a range of biological effects on the body. Some phenolic acids, such as ferulic acid, have antioxidant properties, while others, such as caffeic acid, can also have anti-inflammatory effects (Kumar and Goel, 2019). Steroids are a group of compounds that are found in various plants and animals, and they can have a range

of biological effects on the body. While some steroids are produced naturally in the human body, many others can be derived from plants and other natural resources. For example, some common plant-derived steroids include sitosterol, stigmasterol, and campesterol, which are found in a variety of foods such as nuts, seeds, and vegetable oils. These plant-derived steroids have been found to have a range of potential health benefits, including anti-inflammatory, anti-cancer, and cholesterol-lowering effects (Ras, Geleijnse and Trautwein, 2014).

Natural compounds have been researched for centuries for their medicinal properties which led to outstanding accomplishments concerning drug discovery. An emblematic example is the alkaloid taxol. Taxol is a natural compound derived from the bark of the Pacific yew tree. It was discovered in the 1960s and has since become one of the most important drugs in the treatment of cancer. Taxol works by binding to microtubules in cancer cells, preventing them from dividing and ultimately leading to cell death which is why it has been used to treat a wide range of cancers, including breast, ovarian, and lung cancer (Wall, 1998). Another accomplishment was the discovery of the natural compound artemisinin: Artemisinin is a sesquiterpene found in the sweet wormwood plant *Artemisia annua* that has anti-malarial properties. It was first isolated in the 1970s by Chinese scientists, and its use in treating malaria was confirmed in the 1980s. Artemisinin and its derivatives are now widely used in the treatment of malaria, particularly in combination with other anti-malarial drugs (McIntosh and Olliaro, 2000). Another drug - which is very prominent in the daily lives of many people - is aspirin. Aspirin is a synthetic compound derived from salicylic acid, which is found in willow bark. Salicylic acid has been used for its analgesic and anti-inflammatory properties for centuries, but its use was limited by its toxicity. In the late 19th century, researchers at Bayer discovered a way to modify salicylic acid to make it less toxic and at the same time more effective by adding an acetyl group (-COCH3) to the hydroxyl (-OH) functional group of salicylic acid. The resulting compound was acetylsalicylic acid, more commonly known as aspirin, which is nowadays one of the most widely used drugs in the world, with applications in pain relief, anti-inflammation, and cardiovascular disease prevention (Berk *et al.*, 2013). Another breakthrough that was especially relevant for society during World War II was the synthetic production of penicillin (Swann, 1983). Penicillin is a natural compound produced by the Penicillium fungi that has antibacterial properties. It was discovered by Alexander Fleming in 1928, and its use in treating bacterial infections was confirmed in the 1940s. Penicillin and its derivatives are now widely used in the treatment of bacterial infections, and they have saved

countless lives since their discovery (Ligon, 2004). These are just a few instances of the many natural compounds that have been studied for their therapeutic potential and which are highly in usage up to the present day. Meanwhile, the discovery and use of natural compounds in several treatments for various diseases continues to be an active area of research.

Metabolomics – Overview and History

Metabolomics emerged in the early 21st century as a field of study focused on the comprehensive analysis of all small molecule metabolites present in a biological sample. The term "metabolomics" was first coined in 1998 by Oliver Fiehn (Fiehn, 2002), a researcher at the University of California, Davis. However, the field of metabolomics can be traced back to the development of analytical techniques such as gas chromatography and mass spectrometry in the 1950s and 1960s, which enabled the identification and quantification of small molecule metabolites.

Since then, metabolomics has rapidly advanced, with the development of new analytical technologies, such as nuclear magnetic resonance spectroscopy (MRS), high-performance liquid chromatography (HPLC), and Fourier transform mass spectrometry (FT-MS) (Alarcon-Barrera *et al.*, 2022). The field has grown in scope and applications, including studies of metabolism in health and disease (Laíns *et al.*, 2019), drug discovery and development (Wishart, 2016), environmental monitoring (Skelton *et al.*, 2014), and systems biology (Rosato *et al.*, 2018).

Metabolomics, genomics, transcriptomics, and proteomics are all "omics" techniques that aim to study biological systems at a molecular level. While these techniques share some similarities, they differ in their focus and the types of molecules they analyse:

- Genomics is the study of an organism's entire genome, including all its genes and their functions.
- Transcriptomics focuses on the study of gene expression and the transcripts that are produced from the genome.
- Proteomics is the study of an organism's complete set of proteins, including their structures, functions, and interactions.

- Metabolomics, on the other hand, focuses on the study of all small molecule metabolites present in a biological system, which are substrates and products of enzymatic reactions in the metabolic network as well as the end products of cellular processes.

Metabolites are diverse and include compounds such as amino acids, lipids, sugars, and organic acids. The main goal of metabolomics is to understand the metabolic state of a biological system and how it responds to different stimuli or perturbations, while the other techniques provide information about the potential for gene expression, protein synthesis, or interactions (Horgan and Kenny, 2011). Overall, metabolomics complements the other 'omics techniques by providing a snapshot of the metabolic state of a biological system, which can provide important insights into the underlying physiological and biochemical processes (Sussulini, 2017).

Metabolomics – Analytical methods

There are two approaches of metabolomics analysis: targeted and untargeted metabolomics. Targeted metabolomics refers to the measurement and analysis of a pre-defined set of metabolites (also called "targeted metabolites") in a biological sample. This approach typically involves the use of specific analytical techniques, such as liquid chromatography-mass spectrometry (LC-MS) or gas chromatography-mass spectrometry (GC-MS) that are optimized for the detection and quantification of the target metabolites. The coupled mass spectrometers are not of one particular kind, various mass spectrometers can be used for coupling with chromatography (e.g. quadrupole MS, time of flight MS or tandem MS). Targeted metabolomics is often used to validate hypotheses about specific metabolic pathways or biomarkers, or to quantify known metabolites with high accuracy and precision (Ribbenstedt, Ziarrusta and Benskin, 2018).

On the other hand, untargeted metabolomics involves the comprehensive analysis of all detectable metabolites in a biological sample, without any prior knowledge of their identities or concentrations. This approach allows for the identification of novel metabolites and metabolic pathways, as well as the characterization of metabolic changes associated with various physiological or pathological conditions (Pereira Braga and Adamec, 2019). Untargeted

metabolomics almost exclusively involves the use of high-resolution mass spectrometry. Analytical methods, such as LC-MS or GC-MS, coupled with data processing software are frequently used in to detect and quantify as many metabolites as possible in a sample. At this point it is of significance that the chosen methods should be able to cover a wide chemical spectrum of said metabolites by for instance considering the polarity (Ribbenstedt, Ziarrusta and Benskin, 2018).

Because the two approaches deliver different types of information on metabolites, the choice between targeted or untargeted is dependent on what exactly the research issue is.

Depending on whether targeted or nontargeted metabolomics is used, there are various advantages and disadvantages to each analytical approach.

Table 1 - Advantages and disadvantages of targeted and untargeted metabolomics analysis.

Advantages	
Targeted metabolomics	Untargeted metabolomics
+ Allows for the accurate and precise measurement of specific metabolites. + Can be used to validate hypotheses about specific metabolic pathways or biomarkers. + Typically has higher sensitivity and specificity than untargeted metabolomics. + Can be less time-consuming and expensive than untargeted metabolomics.	+ Allows for the comprehensive analysis of all detectable metabolites in a biological sample. + Can be used to identify novel metabolites and metabolic pathways. + Provides a global view of metabolic changes associated with various physiological or pathological conditions. + Can be used for discovery-driven research questions.
Disadvantages	
Targeted metabolomics	Untargeted metabolomics
– Limited to the measurement of pre-defined metabolites, which may not capture the full complexity of a metabolic system.	– May generate large amounts of data that can be challenging to analyse and interpret.

– May miss important metabolites that were not included in the target list. – May not be suitable for discovery-driven research questions.	– May have lower sensitivity and specificity than targeted metabolomics. – Can be more time-consuming and expensive than targeted metabolomics.

(Wishart, 2016) (Ribbenstedt, Ziarrusta and Benskin, 2018)

There are also differences between using LC-MS and GC-MS in metabolomics analysis. LC-MS is used for the analysis of polar and nonpolar metabolites, including amino acids, peptides, nucleotides, carbohydrates, lipids, and other small molecules. It is especially useful for analysing molecules with a higher molecular weight and greater structural complexity. LC-MS is also able to detect a wide range of molecules within a single sample, making it a useful tool for untargeted metabolomics analysis (Wishart, 2016). GC-MS, however, is used to analyse volatile and semi-volatile metabolites, including small organic acids, sugars, amino acids, and fatty acids (Fiehn, 2002). The analysation of non-volatile metabolites is possible as well if said metabolites can be changed into volatile ones through derivatization (Adams *et al.*, 1999). This is also the reason why GC-MS is used more frequently in targeted metabolomics analysis, where specific metabolites of interest are identified and quantified, as the preparation expense of the to be analysed samples would be too big in an untargeted approach (Fernie and Morgan, 2013). In summary, both LC-MS and GC-MS are important tools in metabolomics analysis, and the choice between the two will depend on the type of metabolites being analysed and the purpose of the study.

Natural compounds & metabolomics

Metabolomics is an important tool for drug discovery and development and has the potential to accelerate the identification of natural compounds with therapeutic potential. Natural compounds are a valuable source of new drugs and therapeutics, and metabolomics analysis can facilitate their identification, like Kim et al. showed in their work by screening anti-inflammatory compounds from the leaves of *Actinidia arguta* (Kim, Lee and Auh, 2019). By analysing changes in metabolite levels associated with disease or drug treatment, researchers can identify natural compounds with potential therapeutic effects (Alarcon-Barrera *et al.*,

2022). Another point would be the understanding of mechanisms of action (Wishart, 2008). By identifying metabolic pathways and cellular processes affected by natural compounds, researchers can gain a better understanding of how they, for instance, exert their anti-cancer effects. This knowledge can inform the development of more effective and targeted therapies which can be seen in a study by Wang et al. that investigated potential targets of colon cancer. Their analysis of the metabolic changes of eicosanoid metabolites produced by the cytochrome P450 monooxygenases showed an increase of said metabolites in plasma and colon of mice which suggested that the cytochrome P450 monooxygenase eicosanoid pathway could be a new target for treatment (Wang *et al.*, 2019). A further aspect is that metabolomics analysis has the potential to enable personalized medicine by identifying natural compounds that are effective in specific patient populations (Chowdhury *et al.*, 2021). By analysing metabolite profiles in patient samples, researchers can identify natural compounds that are more effective in patients with certain genetic or metabolic characteristics (B. Li *et al.*, 2017). For instance, in a study by Yerges-Armstrong et al., the metabolic mechanisms underlying aspirin resistance were investigated in high-risk patients. They found that the purine metabolism pathway was significantly affected by aspirin, and inosine levels were elevated in both good- and poor-responders after aspirin treatment. However, post-dose inosine levels were higher in poor-responders compared to good-responders, suggesting a potential correlation between inosine levels and aspirin response in these patients (Yerges-Armstrong *et al.*, 2013). Furthermore, natural compounds are a valuable source of drugs and therapeutics, and metabolomics analysis can facilitate the discovery of new compounds in a sustainable manner. By focusing on natural compounds, researchers can avoid the environmental and ethical concerns associated with synthetic drug development (Lahlou, 2013). Moreover, natural compounds have a great chemical diversity (Lautié *et al.*, 2020), they are more often readily available and cost-effective than synthetic compounds (Newman and Cragg, 2016), making them attractive candidates for drug discovery.

Cancer research is of paramount importance due to the increasing number of patients afflicted with this disease and the growing demand for novel therapeutic interventions. Metabolomics analysis has gained significant relevance in cancer research due to its potential in understanding the intricate relationship between cellular metabolism and proliferation, a fundamental process in cancer development. This connection between metabolism and cancer has spurred the investigation of natural compounds with anti-cancer properties using metabolomics

approaches. The integration of metabolomics in the study of anti-cancer agents provides valuable insights into their mechanisms of action and aids in the development of novel therapeutic strategies (Schmidt *et al.*, 2021). A well-known example of one natural compound that has been further analysed through metabolomics analysis is curcumin, a polyphenol found in turmeric that has shown to have anti-cancer properties. Metabolomics analysis has revealed that curcumin induces changes in metabolic pathways involved in cell growth, proliferation, and apoptosis in cancer cells. The main targets of curcumin laid in glutathione metabolism, glucose metabolism and lipid/phospholipid metabolism (Bayet-Robert and Morvan, 2013). These findings have been supported by *in vitro* and *in vivo* studies demonstrating the anti-cancer effects of curcumin in a variety of cancer types (Gupta, Patchva and Aggarwal, 2012). Another example is resveratrol, a polyphenol found in grapes, berries, and peanuts that has shown to have anti-cancer properties. Metabolomics analysis has showed that resveratrol induces changes in metabolic pathways, e.g. increasing the extracellular level of polyunsaturated fatty acids, interacting in the cellular biogenic amine metabolism (by increasing the syntheses of several biogenic amines and modified amino acids) and increasing the concentration of amino acids and therefore leading to cell growth inhibition of human breast cancer cell lines (Jäger *et al.*, 2011). Similar findings could be discovered in *in vitro* and *in vivo* studies demonstrating the anti-cancer effects of resveratrol in other cancer types (Singh *et al.*, 2017). Similarly, epigallocatechin gallate (EGCG), a polyphenol found in green tea has shown to have anti-cancer properties. Through metabolomics analysis it was discovered that EGCG induces changes in metabolic pathways, involving changes in the glycerophospholipid metabolism, starch and sucrose metabolism, amino sugar and nucleotide sugar metabolism, and glutathione metabolism (Zhang *et al.*, 2020).

Since metabolomics is – like the name suggests – the study of metabolic pathways, not only anti-cancer effects through metabolic pathways, but also other metabolic disorders are interesting research topics. For instance, in former scientific work several natural compounds have shown regulating effects in glucose metabolism and therefore improved insulin resistance by impacting the glucose pathway. This includes a recent review (Saadeldeen *et al.*, 2020) in which almost 100 natural compounds of different groups, including terpenoids, alkaloids, quinones, flavonoids, phenols, phenyl propanoids, steroids, and some other types of compounds, showed how metabolic pathways – primarily glycolysis – were influenced by different natural compounds which resulted in improving insulin resistance.

After reviewing the previous examples of research in metabolomics, it is notable how the various approaches and techniques utilized can vary significantly. It is critical to choose an appropriate approach that correlates with the research question being addressed. Despite this, it would be beneficial to have a standardized experimental approach that could be applied to the investigation of the metabolic effects of natural compounds on human cells using metabolomics analysis.

The main objective of this master's thesis is to develop and establish such an approach. The significance of this endeavour lies in the fact that the use of natural compounds for the treatment of various diseases has gained interest in recent years. However, the understanding of the metabolic pathways and mechanisms underlying the effects of these natural compounds on human cells remains limited. A standardized approach for investigating these metabolic effects would provide a reliable and consistent framework for research in this area.

Results & Discussion

Selection of natural compounds

The aim of this work is to establish a general framework for profiling the metabolic effects of natural compounds using metabolomics. Since an exhaustive screening of natural products is outside the scope of this work, 20 natural compounds were selected as test compounds to represent the chemical diversity among natural products during the establishment of the method.

To assess the chemical variability, and different compound classes among natural products, the in-house database of pure natural compounds ("Naturstoffdatenbank") available in the group of Prof. Judith Rollinger at the Division of Pharmacognosy was considered. A previous analysis, conducted using Diversity Picker tool in KNIME workbench and NP-Scout (Chen *et al.*, 2019; Stork *et al.*, 2020) to assess natural product-likeness, showed that to sample as chemically diverse compounds as possible, different compound classes should be selected in different proportions (see Table 2). Table 2 shows that heterocyclic and heteroaromatic compounds as well as alkaloids need to be used in higher numbers to adequately represent the chemical diversity in this compound class. Thus, to select a total of 20 compounds from the natural compound database, a literature search was conducted to identify representatives of each compound class that might evoke metabolic effects. The resulting compound selection is shown in Table 3.

Table 2- Proportions of different groups of natural compounds to enable a vast chemical variability in a selection of 20 natural compounds.

Groups of natural compounds	Proportions in [%]	Equivalent in numbers
Heterocyclic and heteroaromatic compounds	19%	4
Alkaloids	14%	3
Organic acids	6%	1
Naphtalene derivatives	5%	1
Phloroglucinols	5%	1
Phenols	4%	1
Steroids and derivatives	4%	1
Anthracene derivatives	3%	1
Flavonoids	3%	1

Lactams	2%	1
Macrolides	2%	1
Phenylpropanes and phenylpropanoids	2%	1
Glucosinolates	1%	1
Saponins and sapogenins	1%	1
Stilbenes	1%	1

Table 3 - List of selected natural compounds and their expected effect.

Natural compound	Group of natural compound	Known effects
Protoveratrin	Alkaloid	quantitative and qualitative changes in the metabolism of cerebral cortex tissue of guinea pigs (Wollenberger, 1955)
Harmin	Alkaloid	inductor of apoptosis and autophagy of gastric cancer cells (C. Li *et al.*, 2017)
Emodin	Anthracene derivative	effects in glucose metabolism, induces apoptosis in human cancer cells (Kim *et al.*, 2018) (Saadeldeen *et al.*, 2020)
Xanthohumol	Chalkon	effects on glucose and lipid metabolism of obese mice (Miranda *et al.*, 2016)
Lanatosid C	Cardiac glycoside	atherosclerosis-promoting effects on lipid metabolism (Shi *et al.*, 2016)
Liquiritigenin	Flavonoid	inhibiting activities on xanthin oxidase (Hsu *et al.*, 2021)
Solanine	Glycoalkaloid	showed inhibition of NSCLC by regulating glycolysis related pathways (Zou *et al.*, 2021)
Harpagosid	Monoterpene	anti-inflammatory activity through inhibition of NF-κB activation (Huang *et al.*, 2006)
Shikonin	Naphtalin derivative	anticancer activity on colon carcinoma cells, effects on multiple metabolic pathways (Chen *et al.*, 2020)
Taurine	Organic acid	cell growth inhibition on A549 cells, induces apoptosis, antioxidative, often found in mitochondria (Tu *et al.*, 2018)

Agaricic acid	Fatty acid	inducer of mitochondrial permeability transition (García *et al.*, 2005)
Gastrodin	Phenole	improve energy metabolism disorders and mitochondrial dysfunction in rat brains (Wu *et al.*, 2023)
Honokiol	Phenylpropane	regulates mitochondrial substrate utilization and cellular fatty acid metabolism in hearts of diabetic mice (Jayakumari *et al.*, 2021)
Hyperforin	Phloroglucinol	antidepressant, affects influx/release from internal stores of Ca(2+), Zn(2+), H(+) and Na(+), potent modulator of mitochondrial functions (Bouron and Lorrain, 2014)
α-Hederin	Saponine	inhibits cell growth by decreasing glycolysis pathways (Fang *et al.*, 2021)
Alantolactone / Helenin	Sesquiterpene	induces cell apoptosis through different pathways and through mitochondrial dysfunction (Liu *et al.*, 2019)
Berberine	Alkaloid	antioxidant and anti-inflammatory activities, regulates insulin resistance and lipid metabolism (Feng *et al.*, 2019)
Resibufogenin	Bufadienolid	TCM-drug, anticancer activities (Ning *et al.*, 2015)
Genistein	Flavonoid	anti-inflammatory, promotion of apoptosis, modulation of steroidal hormone receptors and metabolic pathways (Mukund *et al.*, 2017)

Test experiment with 5-Fluorouracil

In order to get familiarized with the working procedure, 5-Fluorouracil (5-FU) was chosen as a test compound. It is already known that 5-FU has cell growth inhibiting properties and it is commonly used in cytostatic therapies (Longley, Harkin and Johnston, 2003). 5-FU served as primary aid to generate test data which would help to optimize the analysation and interpretation of the growth data in future experiments with natural compounds. To that end, we measured cell growth of the A549 lung cancer cell line under the influence of different concentrations of 5-FU.

The growth rate data was acquired by working with 96 well plates and a well plate reader (Tekan Spark Cyto) using automated bright-field microscopy. A549 cells were seeded in 96-well plates, allowed to attach over-night before 5-FU was added in different concentrations between 0 and 20 µM in individual wells of the plate (Figure 10). The well plate was incubated in the plate reader at 37°C and 5% CO2 atmosphere for 72 hours, in which bright-field microscopy images were acquired in an automated manner every 2 hours to determine the cell confluence until the control samples (cells without 5-FU) in the reference wells reached approximately 80% of confluence. During this time the well plate reader measured the confluence of each well every two hours. This allowed to convert the obtained cell growth data into cell growth curves for each concentration (see Figure 1) by plotting the measured confluence against the time and to generate cell growth rates for each concentration of 5-FU. The evaluation of the data took place via Excel.

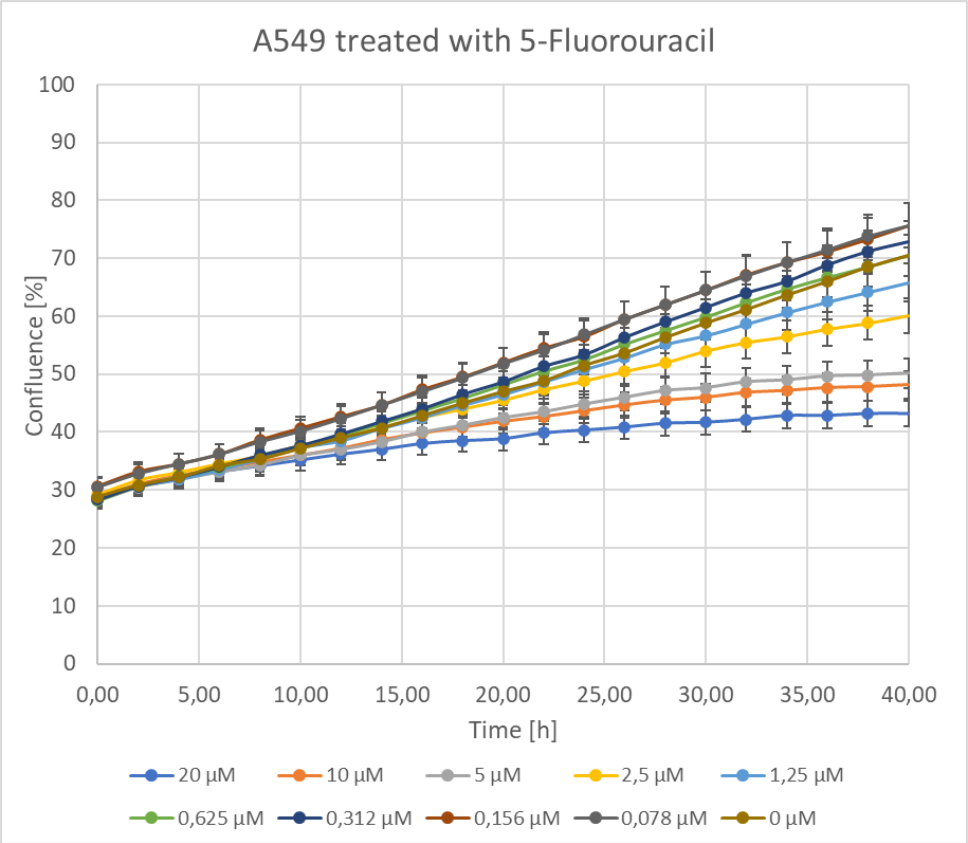

Figure 1 - Cell growth curve of different concentrations of 5-Fluorouracil in a 96 well plate with A549 cells. The different concentrations are depicted each by a different colour. The dots represent the confluences which were measured every two hours. The confluence is depicted in % on the y-axis, the time is depicted on the x-axis in hours. The standard deviation was calculated and added as error bars.

Consistent with previous literature reports (Wen *et al.*, 2010) and experimental data previously acquired in the lab, 5-FU inhibited the growth of A549 in a dose-dependent manner. This inhibition is measured by regarding the cell growth curves and the end confluence of each concentration as well as by calculating the growth rates of each concentration. As can be taken from Figure 1, the 0.156 µM concentration already shows a slight growth defect whereas 2.5 µM of 5-FU decrease the growth in relation to the control samples almost by half.

The growth curves were further analysed in order to certain parameters that can help the comparison of growth effects between different compounds. One possibility is to compare the growth rates, which can be calculated by subtracting the initial confluence from the end confluence in their natural logarithm and dividing the outcome by the time that has passed. The obtained growth rates then can be put in relation to the growth rates of the control samples from which response curves can be plotted. The same can be done by putting the end confluences of each different concentration in relation to the end confluence of the control sample. The difference between the two types of response curves is that the response based on the end confluences solely considers the endpoint values and leaves out the measured values in between, while the response curve based on the growth rates considers the entire growth curve by taking into account every measurement of confluence. However, in this particular case, both response curves show a very similar trend, as shown in Figure 2. Considering both curves could provide a more accurate analysis in case of unexpected measurement issues.

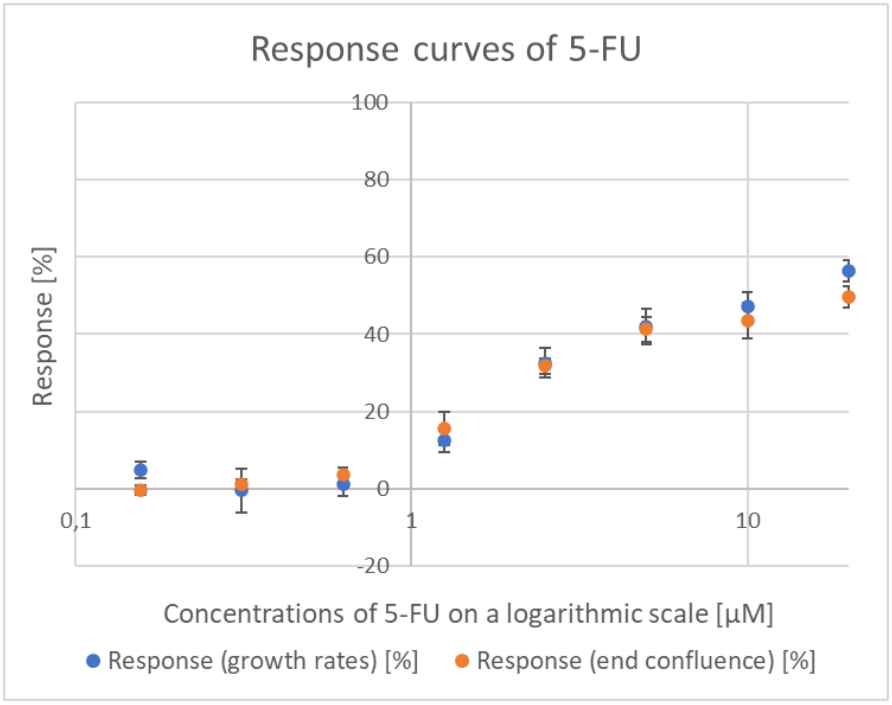

Figure 2 - Response curves of 5-Fluorouracil considering the growth rates (in blue) and the end confluence (in orange) of each concentration. The concentrations of 5-FU are depicted on the x-axis on a logarithmic scale, the response is depicted in % on the y-axis. The standard deviation was calculated and added as error bars.

Evaluation of the experimental approach for natural products

In order to investigate the applicability of the experimental approach used for 5-FU to similar experiments using natural compounds, several research questions needed to be addressed. Firstly, we asked whether different natural compounds could be tested in a similar concentration range, or whether each compound required testing in a specific concentration range. Secondly, it was important to determine if it was possible to characterize cell growth inhibition by measuring the confluence. Lastly, we sought to determine whether any detectable effects on primary metabolism could be observed through metabolomics analysis, and whether these effects varied between the different groups of compounds.

For the first experiment, we selected four natural compounds, namely solanine, helenin, emodin and taurine. After reviewing relevant literature reports (Zhang *et al.*, 2016; Tu *et al.*, 2018; Liu *et al.*, 2019; Xiao *et al.*, 2019) using these natural compounds in cell culture, it was evident that they all showed effects in different concentrations due to their diverse potency. This was a crucial indication that it was necessary to establish concentration ranges that were appropriate for each specific compound, taking into account their individual potency, in order to treat cells at sub-lethal doses and avoid excessive cell death.

Based on literature evidence (Zhang *et al.*, 2016; Kim *et al.*, 2018; Tu *et al.*, 2018; Liu *et al.*, 2019), it was decided to use the concentrations 5, 10, 20, 50 µM for solanine, 10, 25, 50, 100 µM for emodin, 5, 10, 20, 40 µM for helenin and 10, 20, 40, 100 mM for taurine. As solanine, emodin and taurine were solved in DMSO, the control samples were different concentrations thereof as well as reference wells without any additional substances (untreated cells). The exact schematic overview can be taken from Figure 11. After the well plate with seeded A549 cells reached a confluence of 20%, the different concentrations of natural compounds were applied and the well plate was put into the well plate reader for approximately 72 hours, where every two hours the confluence of each well was measured. The obtained cell growth data was again evaluated via Excel. The curves were visualized separately for each natural compound (see below in Figure 3).

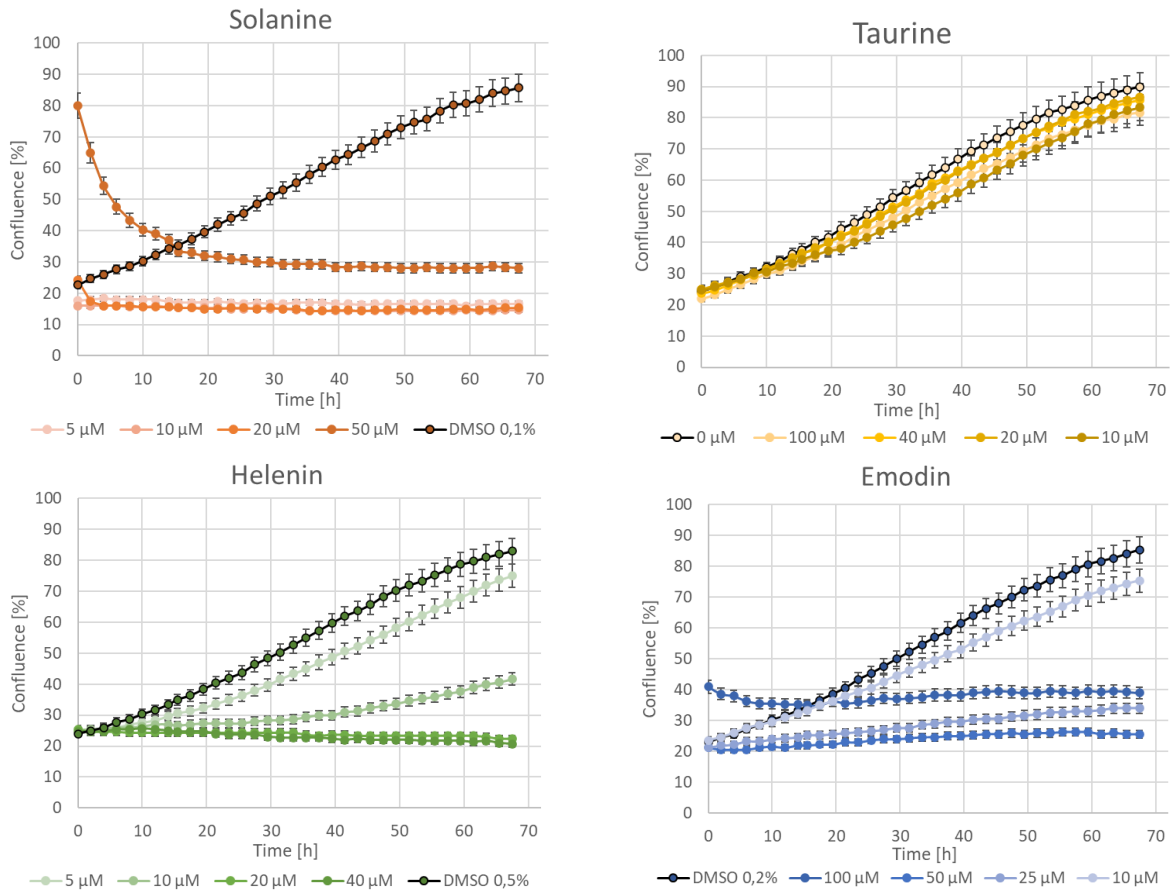

Figure 3 - Cell growth curves for the natural compounds Emodin, Helenin, Solanine and Taurine in different concentration ranges with DMSO and medium as references. The x-axis shows the time in hours, the y-axis shows the confluence in %. Each dot represents the confluence which was measured every two hours. The standard deviation was calculated and added as error bars.

Analysis of the cell growth data showed that the selected concentration ranges required optimizations to obtain a concentration dependent cell growth inhibition. Precipitation of the natural compounds and cell death were the two main issues. The 100 µM concentration of emodin was too high as the substance precipitated, interfering with image acquisition and causing an artificially high readout for cell confluence. Similar problems could be observed with high concentrations of solanine (50 µM). All the other concentrations of solanine led mostly to cell death. The 20 µM and 40 µM concentrations of helenin also led to cell death. The tested concentration range of taurine showed neither precipitation nor cell death but also no growth inhibition. To obtain more accurate results it will be necessary to lower the concentration ranges of emodin, solanine and helenin and increase the concentration range of taurine.

Optimization of concentration ranges

Considering the results from the previous experiment, the concentration ranges were adjusted to allow observing a broad range of growth phenotypes, from normal growth (low concentrations) to strong growth inhibition (high concentrations). With solanine it was decided to decrease the concentration range to under 1 µM as former research (Zou et al., 2021) also suggested working in a nM range. The selected concentrations were 125, 250, 500 and 750 nM. As concentrations over 10 µM of helenin led to cell death in the previous experiment, the Helenin concentrations were reduced to 4, 6, 8, 10 µM. For taurine, the concentration range was extended to 10, 50, 100 and 200 mM to test whether even higher concentrations may affect growth after all. For emodin, the lethal dose of 100 µM was removed and intermediate concentrations were selected instead: 10, 20, 40 and 50 µM. Control samples with only DMSO or only medium were included as a reference for untreated cells. The schematic overview can be seen in Figure 12.

A549 cells seeded in 96-well plates were exposed to the natural compounds at the optimized concentration ranges at a cell confluence of approximately 20% and subjected to time-lapse microscopy imaging (Tecan Spark Cyto) for approximately 48 hours. As can be seen from the resulting growth curves (Figure 4), no concentration showed precipitation, nor did we observe extensive cell death. Surprisingly, solanine did not show any significant cell growth inhibition in this experiment. Similarly, in contrast to the previous experiment, helenin showed only slight growth inhibition at 10 µM. A possible explanation could be that the concentrations were decreased too much, or that the stock solution of solanine did not stay stable since a light change in colour was evident. In contrast, taurine and emodin showed a concentration-dependent inhibition of cell growth.

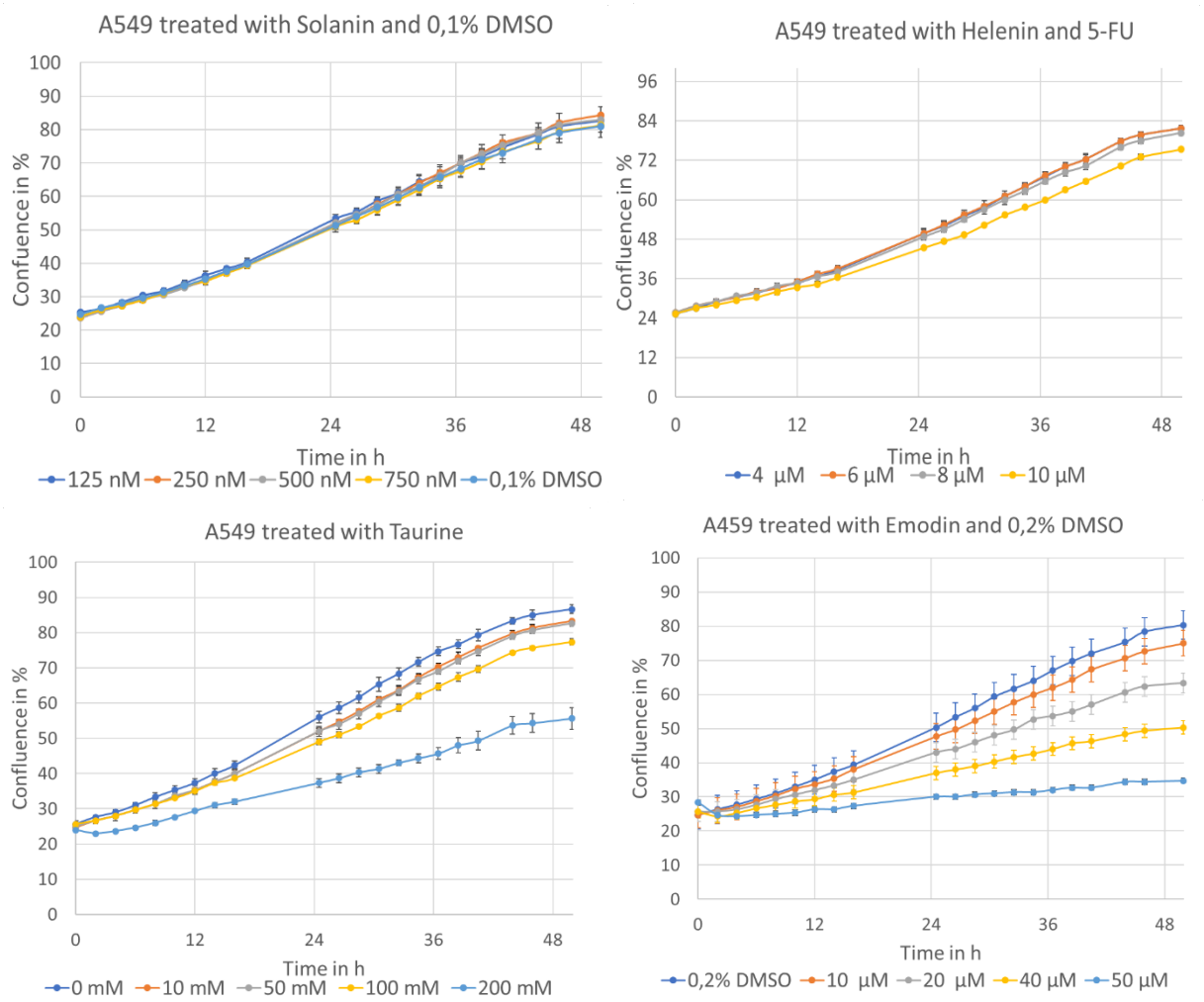

Figure 4 - Cell growth curves for the natural compounds Emodin, Helenin, Solanine and Taurine in different concentration ranges with DMSO and medium as references. The x-axis shows the time in hours, the y-axis shows the confluence in %. Each dot represents the confluence which was measured every two hours. The standard deviation was calculated and added as error bars.

To characterize the relationship between compound concentration and effects on growth, we constructed dose-response-curves from the collected growth data (see Figure 5). To that end, the first step was to determine the growth rates of the cells for each concentration and for each compound. Then each growth rate could be placed in relation to the growth rate of the reference wells which contained no additional substances, only medium or a matching concentration of DMSO. As an alternative to the evaluation of growth rates, we also evaluated the use of the cell confluence value at the end of the treatment as a readout for the response.

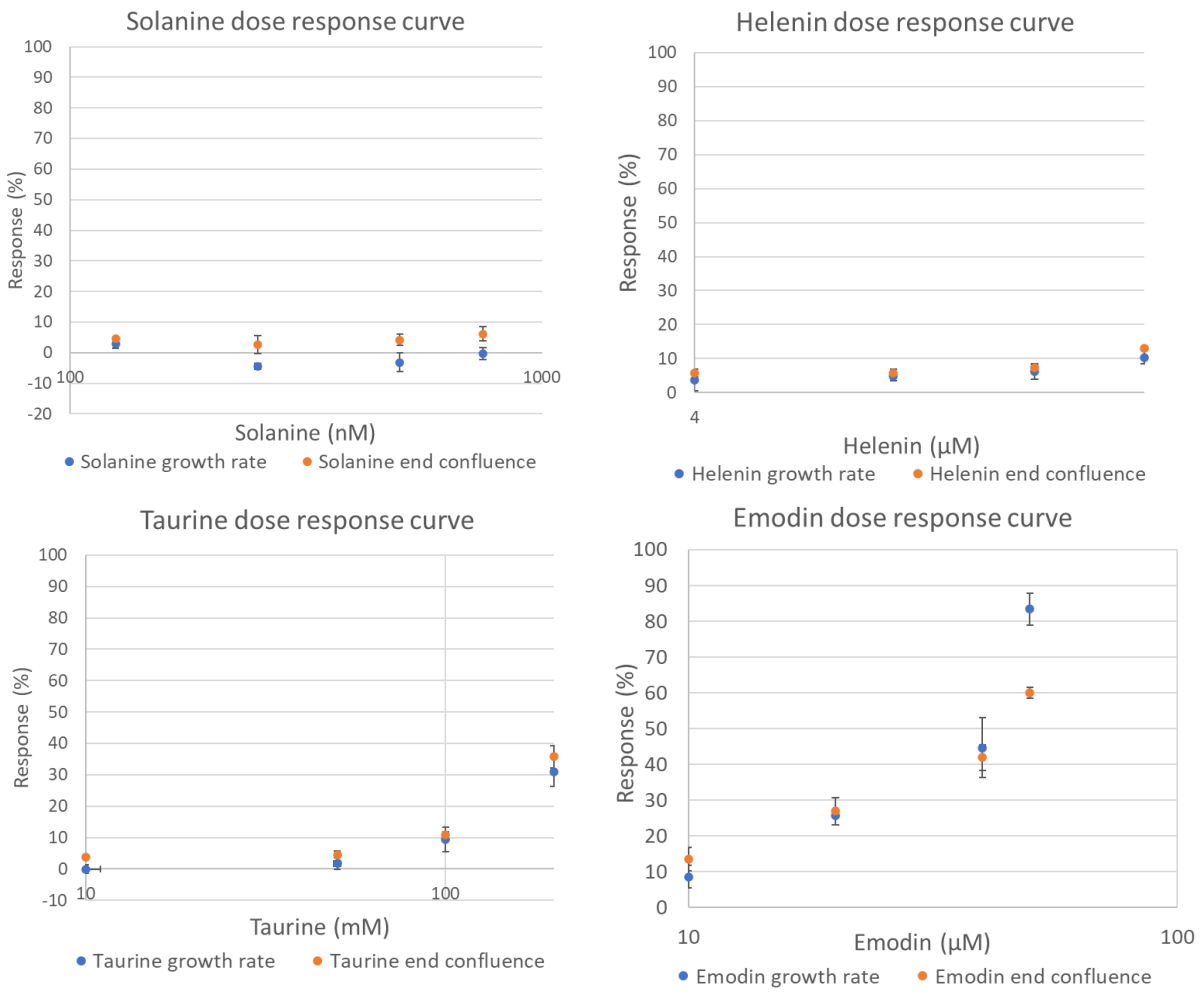

Figure 5 - Dose response curves for the natural compounds solanine, helenin, taurine and emodin. The y-axis shows the response in % which is the growth rate of each concentration in relation to the growth rate of the reference (blue) and the end confluence of each concentration in relation to the end confluence of the reference (orange). The x-axis shows the concentration of the natural compounds on a logarithmic scale. The standard deviation was calculated and added as error bars.

In this case the response for each concentration was determined by calculating the ratio of each growth rate and the growth rate of the control samples (only medium) or by calculating the ratio of each end confluence value and the end confluence value of the control samples and by multiplying the obtained ratios with 100 to turn them into a percentage value. Then, the results were subtracted from 100. A more detailed explanation hereof can be taken from the Materials & Methods chapter.

Since in this work only a few concentrations per natural compound were tested, the dose response curves do not have the typical sigmoidal form which would be expected of standard dose response curves that are commonly generated by testing wide concentration ranges and several concentrations with minimal effects, some close to the concentration with the

maximum effect (Emax). A slight sigmoidal course can be seen with emodin as it also showed a relatively precise concentration dependant growth inhibition. The curve of solanine falls even in the negative number range for the simple reason that the growth rates of some concentrations of solanine were smaller than the growth rate of the reference wells.

In this instance the dose response curves have almost identical shapes, regardless of whether the growth rate or the end confluence is used as the response value, with emodin being the exception. The difference between the two response curves arises because, in the case of the growth rate, the maximum growth rate is determined for each concentration, which tends to occur during the middle phase of cell growth (optimized growth rate). Whereas, when the end confluence values of the untreated samples are examined, it can be observed that the growth curve of the untreated sample flattens out towards the end which represents the stationary phase of growth, in which cell growth slows down when reaching 80% confluence. As a result, the cell growth effect appears to be reduced, indicating that the end confluence of the control sample is actually the issue leading to lower values and is no longer reliable. This could potentially suggest a slight advantage for the method involving growth rates.

In order to further test the applicability of the experimental approach used for 5-FU, another four natural compounds were chosen: Genistein, berberine, agaric acid and shikonin. Based on literature evidence, it was decided to use the concentrations 10, 15, 30, 50 µM for genistein (Yang *et al.*, 2016), 5, 10, 15, 20 µM for berberine (Yin *et al.*, 2008), 2.5, 5, 10, 20 µM for agaric acid (García *et al.*, 2005) and 10, 20, 30, 50 µM for shikonin (Sha *et al.*, 2021).

A549 cells seeded in 96-well plates were exposed to the natural compounds at the optimized concentration ranges at a cell confluence of approximately 20% and subjected to time-lapse microscopy imaging (Tecan Spark Cyto) for 72 hours. The schematic overview of the well plate can be taken from Figure 13. The obtained cell growth data from the well plate reader was evaluated in Excel. Again, cell growth curves for each concentration were compiled and visualized separately for each natural compound (Figure 6).

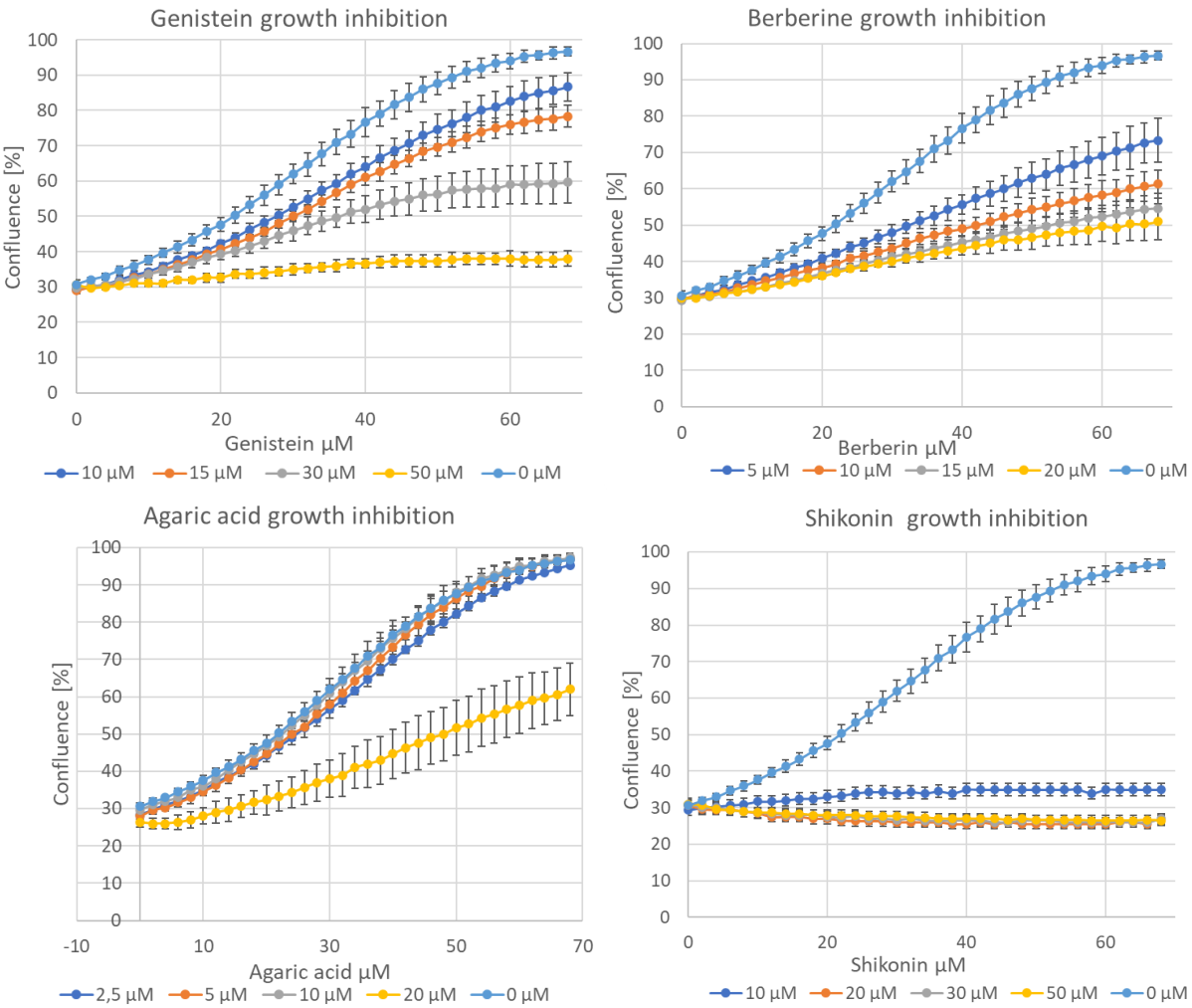

Figure 6 - Growth rates of the natural compounds genistein, berberine, agaric acid and shikonin in different concentration ranges with medium as reference. The x-axis shows the time in hours, the y-axis shows the confluence in %. Each dot represents the confluence which was measured every two hours. The standard deviation was calculated and added as error bars.

Both genistein and berberine showed a concentration dependant cell growth inhibition (Figure 6), while only the highest concentration of agaric acid (20 µM) inhibited cell growth approximately by a third compared to the control samples with only medium. Cells treated with shikonin showed no increase in cell confluence, indicating cell death at all tested concentrations. Future experiments should test whether a dose-dependent growth inhibition can be observed in a lower concentration range (e.g. below 10 µM).

To generate dose response curves, similar to above growth rates and end confluence values were determined for each compound and concentration separately. Then, the ratio between each growth rate of the concentrations and the growth rate of the control samples which consisted only of medium could be determined, multiplied by 100 to turn it into a percentage value and subtracted from 100. A more detailed explanation hereof can be taken from the

Materials & Methods chapter. Additionally, the same steps were followed to determine the response of the end confluence values (Figure 7).

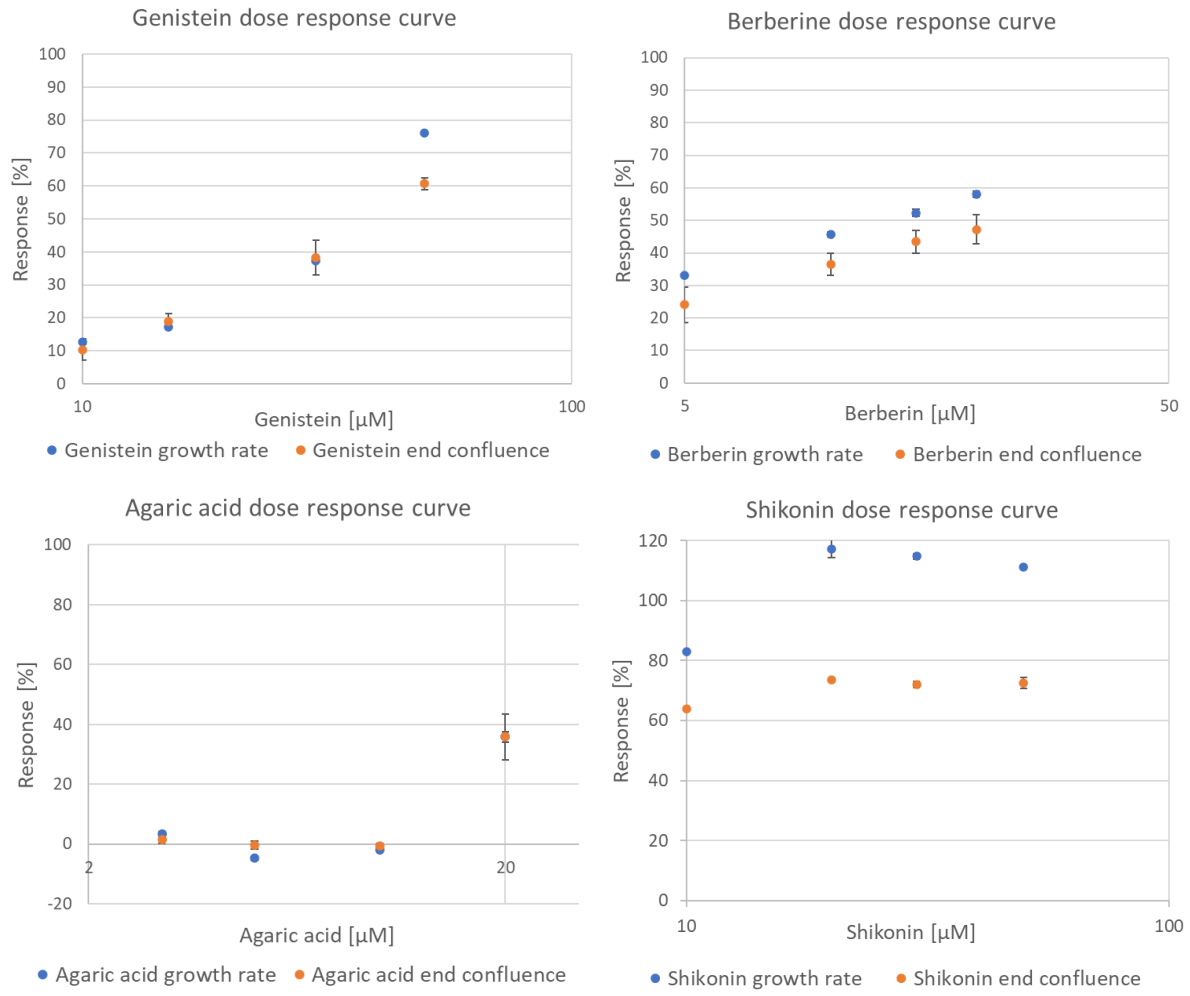

Figure 7 - Dose response curves for the natural compounds genistein, berberine, agaric acid and shikonin. The y-axis shows the response in % which is the growth rate of each concentration in relation to the growth rate of the reference (blue) and the end confluence of each concentration in relation to the end confluence of the reference (orange). The x-axis shows the concentration of the natural compounds on a logarithmic scale. The standard deviation was calculated and added as error bars.

Similar to the dose response curves of solanine, helenin, emodin and taurine, the course of the growth rate curves and the end confluence curves are almost identical for most of the tested compounds, regardless of whether the growth rate or the end confluence is used as the response value (Figure 7). Since agaric acid barely showed any growth inhibition, the dose response curves are flat, similar to the response curves of solanine. Agaric acid also shows a negative response because some of the growth rates of the different concentrations where lower compared to the control samples. Genistein shows a comparable course to emodin, where again a slight difference between the response curves is depicted. Similar to before, the

difference in response curves arises from the fact that the growth rate approach considers the optimized growth phase of the cells, where maximum growth rates are observed for each concentration, while the approach with the end confluence values does not consider that the untreated samples' growth curve flattens towards the end due to the stationary growth phase, leading to reduced apparent growth effects and unreliable end confluence values of the control sample. Shikonin led mostly to cell death which is why the response is high, even on the lowest concentration values. In this case the growth rate approach shows a response of over 100% for several concentrations (20, 30, 50 µM), mostly because the calculated growth rates have lower values compared to the growth rate of the control sample. Regarding the end confluence as the response value, the end confluences of all concentrations also have lower values than the end confluence of the untreated sample, but the response curve stays below 80%.

Overall, the experimental data demonstrates that natural compounds exhibit inhibitory effects on A549 cell growth across various concentration ranges. The utilization of bright-field microscopy imaging proved to be a valuable technique for quantifying the growth effects of these compounds. By monitoring cell growth after exposure to the compounds, it was possible to quantify their effects and establish a connection between molecular changes, such as metabolic alterations, and the inhibitory activity of the natural compounds.

Analysis of metabolic effects of natural compounds

This work aimed at establishing an experimental framework for characterizing the metabolic effects of natural compounds, i.e. to detect relative changes in metabolite abundances in response to the natural compounds. The analysis was performed using a non-targeted metabolomics approach which means there was no focus on a specific group of metabolites (or metabolic pathway), but the whole metabolome was analysed (Ribbenstedt, Ziarrusta and Benskin, 2018).

Samples for metabolome analysis were generated from cells treated with solanine, emodin, helenin and taurine by organic solvent extraction. The natural compounds were applied to A549 cells, cell extracts were generated after 42 hours of incubation, the cells were washed and extracted by applying pre-cooled extraction solvent (organic solvents) which led to a breakdown of the cells. The detailed procedure is described in the Materials & Methods

chapter. The metabolomics analysis of these samples was performed using flow-injection time of flight mass spectrometry in the lab of Prof. Dr. Uwe Sauer at ETH Zurich, Switzerland, by negative mode ionization. The method determines the accurate mass of ions generated from the injected samples using high-resolution mass spectrometry. To annotate the measured ions to known metabolites, putative MS1-based annotation was performed by calculating for each ion the difference between the measured mass and the exact masses of metabolites listed in the Human Metabolome Database (Wishart *et al.*, 2013) and the genome-scale reconstruction of human cell metabolism (Brunk *et al.*, 2018). Only metabolite annotations with a mass difference below 3 mDa were considered. Additionally, the measured relative abundance of individual metabolites was normalized by the amount of cells in the sample following a previously established procedure (Dubuis, Ortmayr and Zampieri, 2018; Ortmayr, Dubuis and Zampieri, 2019). The obtained data was analysed using multivariate statistical methods, including differential analysis by means of a t-test, and principal component analysis.

Principal Component Analysis (PCA) is a statistical technique used to reduce the dimensionality of a data set by transforming the original variables into a new set of orthogonal variables called principal components (PCs). In other words, the method aims to represent the complexity of the data with a small number of parameters (i.e. the principal components). The first PC accounts for the most variation in the data, the second PC for the second-most variation (Ringnér, 2008). Together, the first two PCs can capture a significant amount of the variability in the data, allowing researchers to visualize patterns and relationships among the samples that

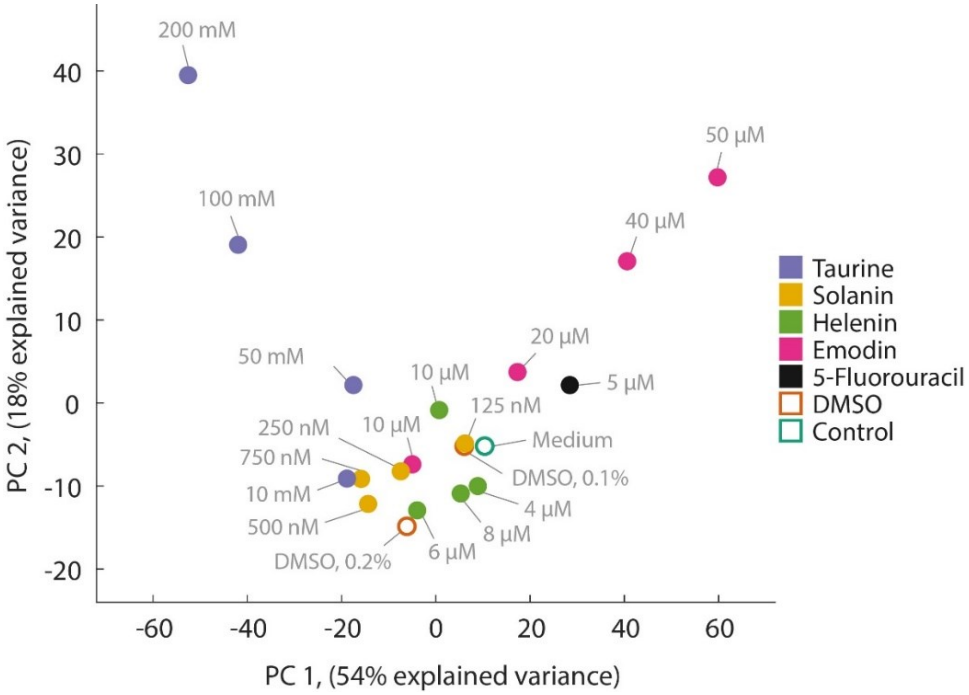

Figure 8 – PCA plot: Here emodin can be seen in pink, taurine in blue, solanine in orange and helenin in green. Each dot shows a different concentration. There are also dots for 5-FU and the references, DMSO and medium. If dots are in close approximation, it suggests a similar metabolic profile. Emodin and taurine show the greatest distance from the control samples.

may not be apparent from a simple inspection of the raw data. By plotting the scores of the first two or three PCs, one can quickly visualize similarities and differences between samples. In addition, PCA can identify which metabolites contribute most to the variation between samples, allowing researchers to focus on those metabolites for further investigation (Nyamundanda, Brennan and Gormley, 2010). Generally speaking, the Principal Component Analysis is an analysation method for big and complex data sets which enables the researcher a simpler visualisation by creating similarity trends.

In a PCA plot, each sample is shown as a dot, and samples that are more similar to each other in terms of their metabolite profiles will be located closer together on the plot (Figure 8). Figure 8 shows the PCA plot for the metabolic effects of taurine, solanine, helenin, emodin, 5-FU and the control samples which contain DMSO or only medium. In the plot's centre, multiple dots closely cluster together, forming a group around the control sample dots. This cluster signifies concentrations of natural compounds that induced metabolomic changes similar to those observed in the control samples, indicating no significant deviations compared to the controls. Conversely, deviations from this cluster, as observed for emodin and taurine, suggest distinct metabolic changes compared to the control samples. Furthermore, it is noteworthy that the distance from the control samples increases with higher concentrations of emodin and taurine suggesting stronger metabolic effects and/or increasing magnitude of metabolic changes. This observation reinforces the association between the dose of natural compounds and metabolic alterations, as previously indicated in the cell growth inhibition plots (Figure 4) and the dose-response curves (Figure 5), where a correlation between the effective dose of growth inhibition and metabolomic changes was expected.

To further investigate the metabolic changes of each compound and each concentration individually, the results of the differential analysis were visualized in form of volcano plots (Figure 9). In volcano plots, the x-axis typically represents a measure of the effect size or magnitude of the metabolic change, such as fold change or in this case $\log_2$ fold change. This indicates how much the metabolite levels have changed between different conditions or concentrations (Kumar, Hoque and Sugimoto, 2018). The y-axis, on the other hand, represents a statistical measure of the significance of the change (Berker, Muti and Cheng, 2023). This is represented as the negative logarithm (base 10) of the p-value which was adjusted for multiple hypothesis testing using the Benjamini-Hochberg procedure (BHFDR). The p-value indicates the

likelihood of observing the observed metabolic change due to random chance alone. A smaller p-value suggests a more significant and reliable change.

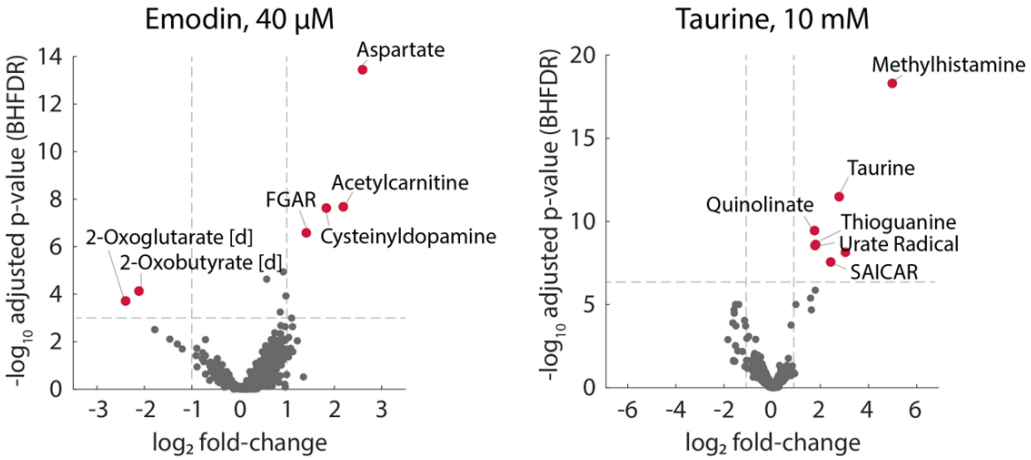

Figure 9 - Volcano plots of emodin at 40 µM and taurine at 10 mM. The x-axis shows the change of metabolite abundances (increase or decrease), the y-axis shows the adjusted p-value which describes the statistical significance of the change.

In the context of depicting the metabolic changes of different concentrations of natural compounds on a volcano plot, each data point on the plot represents a specific metabolite at a specific concentration. The position of the point on the plot is determined by both the magnitude of the metabolic change (log_2 fold change) along the x-axis and the statistical significance of the change (negative log p-value) along the y-axis. Generally, points that are located further away from the centre of the plot (towards the left or right side) and higher up on the y-axis indicate metabolites with larger changes that are statistically significant. These points are often considered as potential candidates for further investigation, as they represent metabolites that are most strongly associated with the concentration of the natural compound being studied. In Figure 9 volcano plots for emodin at a 40 µM concentration and taurine at a 10 mM concentration are depicted. The red highlighted data points represent the metabolites which deviate from the metabolic footprint of the control samples.

Emodin is an anthraquinone which is found in many plants, for instance *Rheum rhabarbarum*. In former research it showed that it affects the glucose metabolism (Xiao et al., 2019) and it also led to changes in cholesterol metabolism (Kim et al., 2018). In Figure 9, emodin shows the highest log_2 fold change at the highest significance with the increase of the metabolite aspartate. Aspartate is a versatile molecule that participates in various aspects of cell metabolism, including protein synthesis, energy production (citrate cycle), nitrogen

metabolism, and nucleotide synthesis. Its presence and availability are crucial for these metabolic processes to proceed efficiently (Furuchi and Homma, 2005). If increased, it might lead to alterations of these various metabolic activities.

Taurine is an organic acid and is also an endogenous component. It can be found in many mammals as metabolic breakdown product of the amino acid cysteine. Taurine already showed in former research growth inhibition of A549 cells (Tu et al., 2018), it has antioxidative activities and is found in mitochondria in higher concentrations which is why it might induce changes in oxidative phosphorylation. The metabolic profile of taurine treatment showed the increase of methylhistamine as the most significant effect. In former research (Nguyen and Cho, 2021) methylhistamine already showed to have tumour progression inhibiting activities in A549 cells among other cancer cell lines. What is also notable is the increase in taurine itself, which indicates that the natural compound was absorbed by the A549 cells successfully.

Overall, the experimental setup employed in this thesis enabled the generation of metabolomics data, allowing for a comprehensive analysis of metabolic changes induced by natural compounds. Notably, the observed growth defects at specific compound concentrations coincided with detectable metabolic effects. Importantly, the metabolic effects exhibited distinct patterns among different compounds, as evident from the principal component analysis as taurine and emodin displayed contrasting metabolic profiles, suggesting different modes of action. These findings hold promise for linking metabolic changes to compound activity and mode of action, paving the way for further investigations in this field.

Conclusion

The aim of this master's thesis was to develop a general experimental approach to investigate the metabolic effects of natural compounds on human cells using metabolomics analysis. This involves answering several questions, including the appropriate concentration range for the compounds, the ability to measure cell growth inhibition, and the detectability of effects on primary metabolism through metabolomics.

Regarding the concentration range, it is important to note that the optimal concentration for each compound will differ and depend on the compound's potency. At an optimal concentration, the compound should induce a significant growth defect but not cell death to allow capturing specific metabolic effects related to the compound's activity, and not unspecific consequences of cell death. Therefore, it is necessary to determine the concentration range for each natural compound individually, as factors such as the compounds solubility, stability, and cytotoxicity can all impact the range. Here, we performed several experiments that can serve as a blueprint for how optimal concentrations can be determined by initially testing various concentrations to establish a concentration-response relationship. Once the optimal concentration range has been determined, the subsequent metabolomics analysis can provide valuable insights into the metabolic effects of the natural compounds on the cells.

Regarding the measurement of cell growth inhibition, it is possible to assess cell growth using various techniques such as bright-field microscopy to determine cell confluence. Such utilization of growth monitoring proved to be a valuable tool for determining growth effects and establishing dose-response relationships. By continuously monitoring and quantifying cell growth over time, it was possible to accurately assess the impact of natural compounds on cellular proliferation. This approach enabled the characterization of the concentration-dependent effects of the natural compounds, providing valuable insights into their efficacy and potency. The obtained dose-response relationships contribute to a better understanding of the compound's growth inhibition activity and enable the identification of optimal concentrations for further studies or potential therapeutic applications.

Finally, regarding the detectability of effects on primary metabolism through metabolomics, the variability observed in the metabolic responses elicited by different natural compounds adds an intriguing aspect to this research. The further analysis of the change in metabolites can provide insight into the impact of natural compounds on cellular metabolism.

In summary, we successfully adapted a general experimental approach for investigating the metabolic effects of natural compounds on human cells using metabolomics analysis. Metabolome profiling is a promising avenue for advancing our understanding of the impact of natural compounds on cellular metabolism and relating metabolic effects to compound activity.

Based on the results of this thesis, it seems that the development of this experimental approach is nearing completion, as no major technical issues are currently outstanding. However, it should be noted that the frameworks' validity and applicability have been tested on a limited selection of natural compounds. Moving forward, this experimental setup would enable researchers to commence screening various compounds and explore their specific metabolic effects. It allows for a comprehensive evaluation of the potential of natural compounds and aids in identifying compounds with distinct metabolic profiles. Further research in this field could have the potential to lead to the development of novel therapeutic interventions for a variety of diseases.

Materials and Methods

Working materials

The RPMI medium 1640 (cat. No. 11875093) and PBS (cat. No. 10010) were purchased from Gibco; Thermo Fisher Scientific, Inc. (Waltham, MA, USA). The penicillin/streptomycin antibiotics (10kU-10 mg/mL P4333-100ML), the fetal bovine serum (FBS) full (F2442-500ML), 5-Fluorouracil (F6627-1G) and Trypsin-EDTA (0.25%, cat. no. T4049-100ML) were purchased from Sigma-Aldrich. 5 mL P/S antibiotics were added to the RPMI medium 1640 in order to prevent bacterial contamination, 25 mL FBS full were added as it contains various proteins which enable a more proficient cell growth. The T75 (cat. No. 734-2314P) and T25 cell culture flasks (cat. No. 734-2311P) as well as the 96-well plates (cat. No. 734-2327) were purchased from VWR International GmbH. The cell line stock of A549 lung carcinoma cells was provided by the lab of Dr. Mattia Zampieri at ETH (Eidgenössische Technische Hochschule) in Zurich, Switzerland, who received the cell line as part of a project with the National Cancer Institute (NCI).

Thawing cell line A549 from cryo-preserved stock

In the beginning, we prepared the workspace by cleaning the laminar air flow hood thoroughly with 70% ethanol. In the meantime, basic medium RPMI 1640 (500 mL) was warmed up in the water bath. 9 mL of the prepared medium were transferred to a 15 mL tube and heated at 37°C in the water bath. The frozen cell line stocks were collected from the liquid nitrogen tank and immediately thawed in the water bath.

Once the cells were thawed, they were put under the laminar flow hood and pipetted slowly into the prepared 15 mL tube. By doing that, the DMSO was diluted to avoid potential toxic effects on the cells. DMSO is used for freezing cell line stocks in order to prevent water to form rough crystals that could damage the cells. Cells were centrifuged for about five minutes at 400g. The medium above the cell pellet was then carefully aspirated, and then 5mL fresh medium were again added and in order to resuspend the pellet by pipetting.

The cell suspension was then transferred into a previously prepared T25 flask (cell culture flask with 25 cm^2 culture surface) and put in the incubator (5% CO_2 at 37°C) over night.

The day after, the cells were checked under the microscope whether they had attached. Cells were subsequently passaged, i.e. transferred to a new flask at lower cell density, whenever the confluence reached approximately 80-90% (determined by visual inspection).

Passaging adherent A549 cells

The medium in the T25 flask was aspirated carefully while tilting the flask 90°. Then warm PBS (37°C) was pipetted into the flask and the flask was swayed gently. The PBS served to thoroughly remove any traces of medium and was again aspirated. Then trypsin was added, and the flask was put in the incubator for about five minutes. Trypsin is a protease enzyme which digests proteins of the extracellular matrix. By adding trypsin, the cells detach from the surface of the bottom of the flask and this way can be transferred. The volume of PBS and trypsin depends on the flask size, i.e. 1 mL of trypsin is added in a T25 flask, 2 mL in a T75 flask; 5 mL PBS are added in a T25 flask, 8 mL in a T75 flask.

After checking under the microscope if the cells were detached, warm serum-containing medium was added to the cells. To inactivate trypsin, the threefold amount of volume of medium in relation to trypsin was added and the cell suspension was pipetted a few times up and down. Then the whole cell suspension was transferred to the previously prepared 15 mL tube. Cells were centrifuged for five minutes at 400g to generate a cell pellet.

The medium from the tube was aspirated without disrupting the cell pellet. Taking the split ratio into account, fresh medium was added and resuspended a few times to avoid cell clusters. Then the cell suspension was transferred to the prepared T75 flask. Cells were then returned into the CO_2-incubator.

At this point, passaging took place every two to three days, usually when approximately 80% of confluence were reached, by the methods mentioned above.

Cell Seeding in 96-well plates

The medium in the T75 flask was aspirated carefully while tilting the flask 90°. Then PBS was pipetted into the flask and the flask was swayed gently. This was again aspirated and served to thoroughly remove any trails of old medium. Then trypsin was added, and the flask was put in the incubator for five minutes. After the cells detached, medium was added. The cell suspension was resuspended a few times and transferred to a previously prepared 15 ml tube and put to centrifugate at 400g for five minutes.

After centrifugation, the medium above the cell pellet was aspirated and the pellet was resuspended in 2 mL of medium. To determine the cell concentration in the suspension, cells were counted using an automated cell counter (cat. No. 734-2675, VWR). To that end, 12 µL of the cell suspension were mixed with 12 µL of trypan blue reagent. Trypan blue is a dye used for determining cell viability since it only permeates and dyes dead and perforated cells blue. 10 µL of the cell suspension-trypan blue mixture were respectively pipetted on each side of a cell counter slide and put in a cell counter.

In order to fill 60 plates of the 96 well plate (the outer wells would be filled with PBS) including an error factor of 20%, a 10.8 mL cell suspension would be needed but due to exercise purposes it was decided to prepare a 12 mL cell suspension.

The optimal seeding cell density which was previously determined amounts to 2.1×10^5 cells/mL, 150 µL cell suspension per well. In consideration of the dilution factor, a dilution of the present cell suspension with medium was prepared. The 96-well plate was put under the hood together with a pipette tray in which the cell suspension was pipetted. Following, a multi pipette was used to pipette the cell suspension into the inner wells, the outer wells were filled with PBS. Finally, the 96-well plate was put in the incubator to allow the cells to attach and grow.

The remaining cells in the flask were passaged. Passaging took place by the methods mentioned above.

Test experiment with 5-Fluorouracil

Cells were seeded (cell density 2.1x10^5 cells/mL, 150 µL cell suspension **per well**) approximately 24 hours prior to experiment start. The 96-well plate was inspected under the microscope and the confluences were measured in a well plate reader (Tecan Spark Cyto) to verify the intended starting cell confluence between 20-25%. A 5 mM 5-FU stock solution was prepared which was diluted in two steps to achieve a concentration of 200 µM. 20 µL of the diluted stock solution were spiked directly to the cells with 480 µL medium which resulted in a concentration of 20 µM. Then, 100 µL of the 20 µM diluted solution were pipetted with 100µL medium to the next column of wells which resulted in a concentration of 10 µM. This step was repeated as a dilution series as shown below (Figure 10). The well plate was put in the well plate reader for incubation and confluence measuring.

	1	2	3	4	5	6	7	8	9	10	11	12
A	-	-	-	-	-	-	-	-	-	-	-	-
B	-	20	10	5	2,5	1,25	0,625	0,313	0,156	0,078	0	-
C	-	20	10	5	2,5	1,25	0,625	0,313	0,156	0,078	0	-
D	-	20	10	5	2,5	1,25	0,625	0,313	0,156	0,078	0	-
E	-	20	10	5	2,5	1,25	0,625	0,313	0,156	0,078	0	-
F	-	20	10	5	2,5	1,25	0,625	0,313	0,156	0,078	0	-
G	-	20	10	5	2,5	1,25	0,625	0,313	0,156	0,078	0	-
H	-	-	-	-	-	-	-	-	-	-	-	-

*Figure 10 – Schematic overview of a 96-well plate with 5-FU concentrations in **µM** in each well of a 96 well plate. The outer wells which are marked with '-' are filled with PBS.*

Test experiment with natural compounds

Cell seeding occurred in the same order of steps as mentioned above with the difference in cell suspension concentration. Due to the pre-experiment with 5-FU it was decided to start with a lower concentration of 1.6x10^4 cells/mL.

For this experiment, four natural compounds were chosen since four compounds could be divided rationally on a 96 well plate, with the compounds being taurine, solanine, emodin and

helenin. All natural compounds were provided by the group of Prof. Dr. Rollinger, Division of Pharmacognosy of the University of Vienna.

Taurine was solved in cell culture medium. Since solanine, emodin and helenin are not solvable in aqueous solutions, they had to be solved in DMSO. Because DMSO can have growth-inhibitory effects at some concentrations, control wells were included where cells were exposed to 0.5%, 0.2% and 0.1% DMSO, or only medium.

The concentrations of the compounds were chosen based on the experiences of former research from scientific papers (see main text). The ideal outcome would be a concentration dependant cell inhibition. A schematic overview of the well plate is shown below in Figure 11.

	1	2	3	4	5	6	7	8	9	10	11	12
A	-	-	-	-	-	-	-	-	-	-	-	-
B	-	5	10	20	50	0,1% DMSO	0,2% DMSO	100	50	25	10	-
C	-	5	10	20	50	0,1% DMSO	0,2% DMSO	100	50	25	10	-
D	-	5	10	20	50	0,1% DMSO	0,2% DMSO	100	50	25	10	-
E	-	5	10	20	40	0,5% DMSO	0	100	40	20	10	-
F	-	5	10	20	40	0,5% DMSO	0	100	40	20	10	-
G	-	5	10	20	40	0,5% DMSO	0	100	40	20	10	-
H	-	-	-	-	-	-	-	-	-	-	-	-

Figure 11 - Schematic overview of a 96-well plate with concentrations of solanine (red) in nM, emodin (blue) in µM, helenin (green) in µM, taurine (orange) in mM and references (gray). The outer wells with '-' are filled with PBS.

Experiments with natural compounds for metabolomics analysis

This time not one but five plates needed to be seeded: two plates with different cell concentrations ($7.10x10^3$, $9.47x10^3$, $1.18x10^4$, $1.42x10^4$ cells/ml) which served as reference for metabolome changes and the other three plates were to be exposed to natural compounds of which one would stay in the well plate reader for measuring cell growth, and the remaining two would undergo a cell extraction for metabolomics analysis.

Again, taurine, solanine, emodin and helenin were chosen. According to the results of the previous experiment, the concentrations were optimized in such a manner that a concentration dependant cell growth inhibition would be expected which hopefully would bring forth metabolic changes. The natural compound solutions were prepared the same way as mentioned in the test experiment. The reference wells with 0.1% and 0.2% DMSO were maintained, the 0.5% concentration was discarded since the concentrations of the natural compounds, which were solved in DMSO, were lowered after optimization. Additionally, a reference sample with 5 µM 5-Fluorouracil was added. The schematic overview of the well plates is shown below in Figure 12.

	1	2	3	4	5	6	7	8	9	10	11	12
A	-	-	-	-	-	-	-	-	-	-	-	-
B	-	125	250	500	750	0,1% DMSO	0,2% DMSO	10	20	40	50	-
C	-	125	250	500	750	0,1% DMSO	0,2% DMSO	10	20	40	50	-
D	-	125	250	500	750	0,1% DMSO	0,2% DMSO	10	20	40	50	-
E	-	4	6	8	10	5µM 5-FU	0	10	50	100	200	-
F	-	4	6	8	10	5µM 5-FU	0	10	50	100	200	-
G	-	4	6	8	10	5µM 5-FU	0	10	50	100	200	-
H	-	-	-	-	-	-	-	-	-	-	-	-

Figure 12 - Schematic overview of a 96 well plate with concentrations of solanine (red) in nM, emodin (blue) in µM, helenin (green) in µM, taurine (orange) in mM and references (gray). The outer wells with '-' are filled with PBS.

To gain further information another experiment was started. For this experiment, again four natural compounds were chosen since four compounds could be divided rationally on a 96-well plate. The selected compounds were genistein, berberine, agaric acid and shikonin. All compounds were solved in DMSO. The concentrations of the compounds were chosen based on the experiences of former research from scientific papers (see main text). The ideal outcome would be a concentration dependant cell inhibition. The schematic overview of the well plate can be taken from Figure 13.

	1	2	3	4	5	6	7	8	9	10	11	12
A	-	-	-	-	-	-	-	-	-	-	-	-
B	-	10	15	30	50	0	5	10	15	20	0,1% DMSO	-
C	-	10	15	30	50	0	5	10	15	20	0,1% DMSO	-
D	-	10	15	30	50	0	5	10	15	20	0,1% DMSO	-
E	-	3	5	10	20	0	10	20	30	50	0,2% DMSO	-
F	-	3	5	10	20	0	10	20	30	50	0,2% DMSO	-
G	-	3	5	10	20	0	10	20	30	50	0,2% DMSO	-
H	-	-	-	-	-	-	-	-	-	-	-	-

Figure 13 - Schematic overview of a 96 well plate with concentrations of genistein (red) in µM, berberin (blue) in µM, agaric acid (green) in µM, shikonin (orange) in µM and references (blue and orange). The outer wells with '-' are filled with PBS.

Metabolomics sampling

Before starting with the cell extraction, a wash and extraction solvent needed to be prepared. The wash solvent was a solution of nano pure water and ammonium carbonate at a concentration of 75 mM. The pH would be in an alkaline range and had to be brought to about 7.4 with acetic acid. The wash solvent had to be prepared freshly (should not stay for over an hour) and it had to be warmed up at 37°C in the water bath. The extraction solvent is a mixture of organic solvents and water, consisting of 40% acetonitrile, 40% methanol and 20% nano pure water. Right before the cell extraction, 25 µM phenyl hydrazine was added. The extraction solvent needed to be stored at -20°C.

One plate was brought from the incubator under the laminar air flow. A multichannel pipet was used to aspirate the medium in each well. This occurred with the plate turned slightly upwards and in a slow and careful manner as to prevent an unwished-for flushing away of cells. Then, in the same manner, 150 µL warm wash solvent was added into each well and again aspirated. The last step was the cell extraction in which quickly 100 µL of the pre-cooled extraction solvent was added to each well. The plate was sealed with aluminium-adhesive foil and immediately transferred to -20°C. For long term storage, the plate was transferred after approximately one hour in a -80°C freezer.

A replicate well plate was treated identically, except that in the last step of the cell extraction it was not induced with extraction solvent but with medium. It was put immediately in the well plate reader. This served to determine how many cells got lost during the washing process. Similarly, a plate holding different amounts of cells was extracted to generate samples with different cell concentrations for normalization of metabolite abundances.

Construction of dose-response curves

To establish dose-response curves, the first step was to determine the growth rates for each concentration. For this purpose, the formula $growth\ rate(\mu) = \frac{LN(final\ confluence) - LN(initial\ confluence)}{time\ difference}$ was used. This formula theoretically applies only during exponential growth. Therefore, a time window was determined during which the cells exhibited exponential growth characteristics. Using Excel, the logarithmic values of the confluences were plotted against the time and trend lines and their coherent linearity gradients were generated for different time windows. The time window with the highest linearity gradient value (above 0.99) was selected. The confluence values at the end and the beginning of this time window were then inserted into the formula.

Once the growth rates were calculated, the responses to each concentration could be determined. Initially, the formula $response(\%) = \frac{\mu\ (concentration)}{\mu\ (control\ sample)} * 100$ was used. However, working with this formula resulted in high values between 80-90 % and sometimes exceeding 100 % (when the growth rate was higher than that of the control samples, which was the case for some low concentrations). Visually, these values deviated from the norm of a typical dose-response curve. To address this, the formula was modified to subtract the resulting response values from 100 since the response of the control sample was considered 100 % when put into the formula mentioned above. This resulted in lower response values that increased with increasing concentration, indicating dose dependency and being visually more consistent. Therefore, the formula was amended to $response(\%) = 100 - \frac{\mu\ (concentration)}{\mu\ (control\ sample)} * 100.$

The obtained response values were plotted against the concentrations of the respective natural compounds on a graph in Excel, using a semi-logarithmic scale. Standard deviations were calculated and added as error bars.

The approach based on the end confluence values followed a similar procedure. In this case, the formula used was $response(\%) = 100 - \frac{end\ confluence(concentration)}{end\ confluence\ (control\ sample)} * 100$. The subsequent steps were equal to the steps of the growth rate approach.

References

Adams, M.A. *et al.* (1999) 'Simultaneous Determination by Capillary Gas Chromatography of Organic Acids, Sugars, and Sugar Alcohols in Plant Tissue Extracts as Their Trimethylsilyl Derivatives', *Analytical Biochemistry*, 266(1), pp. 77–84. Available at: https://doi.org/10.1006/abio.1998.2906.

Alarcon-Barrera, J.C. *et al.* (2022) 'Recent advances in metabolomics analysis for early drug development', *Drug Discovery Today*, 27(6), pp. 1763–1773. Available at: https://doi.org/10.1016/j.drudis.2022.02.018.

Bayet-Robert, M. and Morvan, D. (2013) 'Metabolomics Reveals Metabolic Targets and Biphasic Responses in Breast Cancer Cells Treated by Curcumin Alone and in Association with Docetaxel', *PLOS ONE*, 8(3), p. e57971. Available at: https://doi.org/10.1371/journal.pone.0057971.

Berk, M. *et al.* (2013) 'Aspirin: a review of its neurobiological properties and therapeutic potential for mental illness', *BMC medicine*, 11, p. 74. Available at: https://doi.org/10.1186/1741-7015-11-74.

Berker, Y., Muti, I.H. and Cheng, L.L. (2023) 'Visualizing metabolomics data with R', *NMR in Biomedicine*, 36(4), p. e4865. Available at: https://doi.org/10.1002/nbm.4865.

Bohlmann, J. and Keeling, C.I. (2008) 'Terpenoid biomaterials', *The Plant Journal*, 54(4), pp. 656–669. Available at: https://doi.org/10.1111/j.1365-313X.2008.03449.x.

Bouron, A. and Lorrain, E. (2014) '[Cellular and molecular effects of the antidepressant hyperforin on brain cells: Review of the literature]', *L'Encephale*, 40(2), pp. 108–113. Available at: https://doi.org/10.1016/j.encep.2013.03.004.

Bribi, N. (2018) 'Pharmacological activity of Alkaloids: A Review', 1. Available at: https://doi.org/10.63019/ajb.v1i2.467.

Brunk, E. *et al.* (2018) 'Recon3D enables a three-dimensional view of gene variation in human metabolism', *Nature Biotechnology*, 36(3), pp. 272–281. Available at: https://doi.org/10.1038/nbt.4072.

Chen, Y. *et al.* (2019) 'NP-Scout: Machine Learning Approach for the Quantification and Visualization of the Natural Product-Likeness of Small Molecules', *Biomolecules*, 9(2), p. 43. Available at: https://doi.org/10.3390/biom9020043.

Chen, Y. *et al.* (2020) 'Integration of Transcriptomics and Metabolomics Reveals the Antitumor Mechanism Underlying Shikonin in Colon Cancer', *Frontiers in Pharmacology*, 11. Available at: https://www.frontiersin.org/articles/10.3389/fphar.2020.544647 (Accessed: 16 March 2023).

Chowdhury, S. *et al.* (2021) 'The role of metabolomics in personalized medicine for diabetes', *Personalized Medicine*, 18(5), pp. 501–508. Available at: https://doi.org/10.2217/pme-2021-0083.

Debnath, B. *et al.* (2018) 'Role of plant alkaloids on human health: A review of biological activities', *Materials Today Chemistry*, 9, pp. 56–72. Available at: https://doi.org/10.1016/j.mtchem.2018.05.001.

Dubuis, S., Ortmayr, K. and Zampieri, M. (2018) 'A framework for large-scale metabolome drug profiling links coenzyme A metabolism to the toxicity of anti-cancer drug dichloroacetate', *Communications Biology*, 1(1), pp. 1–11. Available at: https://doi.org/10.1038/s42003-018-0111-x.

Fang, C. *et al.* (2021) 'α-Hederin inhibits the growth of lung cancer A549 cells in vitro and in vivo by decreasing SIRT6 dependent glycolysis', *Pharmaceutical Biology*, 59(1), pp. 11–20. Available at: https://doi.org/10.1080/13880209.2020.1862250.

Feng, X. *et al.* (2019) 'Berberine in Cardiovascular and Metabolic Diseases: From Mechanisms to Therapeutics', *Theranostics*, 9(7), pp. 1923–1951. Available at: https://doi.org/10.7150/thno.30787.

Fernie, A.R. and Morgan, J.A. (2013) 'Analysis of metabolic flux using dynamic labelling and metabolic modelling', *Plant, Cell & Environment*, 36(9), pp. 1738–1750. Available at: https://doi.org/10.1111/pce.12083.

Fiehn, O. (2002) 'Metabolomics – the link between genotypes and phenotypes', *Plant Molecular Biology*, 48(1), pp. 155–171. Available at: https://doi.org/10.1023/A:1013713905833.

Furuchi, T. and Homma, H. (2005) 'Free D-Aspartate in Mammals', *Biological and Pharmaceutical Bulletin*, 28(9), pp. 1566–1570. Available at: https://doi.org/10.1248/bpb.28.1566.

Ganesan, A. (2008) 'The impact of natural products upon modern drug discovery', *Current Opinion in Chemical Biology*, 12(3), pp. 306–317. Available at: https://doi.org/10.1016/j.cbpa.2008.03.016.

García, N. *et al.* (2005) 'Agaric acid induces mitochondrial permeability transition through its interaction with the adenine nucleotide translocase. Its dependence on membrane fluidity', *Mitochondrion*, 5(4), pp. 272–281. Available at: https://doi.org/10.1016/j.mito.2005.05.002.

Gupta, S.C., Patchva, S. and Aggarwal, B.B. (2012) 'Therapeutic Roles of Curcumin: Lessons Learned from Clinical Trials', *The AAPS Journal*, 15(1), pp. 195–218. Available at: https://doi.org/10.1208/s12248-012-9432-8.

Heim, K.E., Tagliaferro, A.R. and Bobilya, D.J. (2002) 'Flavonoid antioxidants: chemistry, metabolism and structure-activity relationships', *The Journal of Nutritional Biochemistry*, 13(10), pp. 572–584. Available at: https://doi.org/10.1016/S0955-2863(02)00208-5.

Horgan, R.P. and Kenny, L.C. (2011) ‘‘Omic’technologies: genomics, transcriptomics, proteomics and metabolomics’, *The Obstetrician & Gynaecologist*, 13(3), pp. 189–195.

Hsu, S.-J. *et al.* (2021) ‘Fast dereplication of xanthine oxidase-inhibiting compounds in alfalfa using comparative metabolomics’, *Food Research International*, 141, p. 110170. Available at: https://doi.org/10.1016/j.foodres.2021.110170.

Huang, T.H.-W. *et al.* (2006) ‘Harpagoside suppresses lipopolysaccharide-induced iNOS and COX-2 expression through inhibition of NF-κB activation’, *Journal of Ethnopharmacology*, 104(1), pp. 149–155. Available at: https://doi.org/10.1016/j.jep.2005.08.055.

Jäger, W. *et al.* (2011) ‘Metabolomic Analysis of Resveratrol-Induced Effects in the Human Breast Cancer Cell Lines MCF-7 and MDA-MB-231’, *OMICS: A Journal of Integrative Biology*, 15(1–2), pp. 9–14. Available at: https://doi.org/10.1089/omi.2010.0114.

Jayakumari, N.R. *et al.* (2021) ‘Honokiol regulates mitochondrial substrate utilization and cellular fatty acid metabolism in diabetic mice heart’, *European Journal of Pharmacology*, 896, p. 173918. Available at: https://doi.org/10.1016/j.ejphar.2021.173918.

Kim, G.-D., Lee, J.Y. and Auh, J.-H. (2019) ‘Metabolomic Screening of Anti-Inflammatory Compounds from the Leaves of Actinidia arguta (Hardy Kiwi)’, *Foods*, 8(2), p. 47. Available at: https://doi.org/10.3390/foods8020047.

Kim, Y.-S. *et al.* (2018) ‘Emodin Sensitizes Hepatocellular Carcinoma Cells to the Anti-Cancer Effect of Sorafenib through Suppression of Cholesterol Metabolism’, *International Journal of Molecular Sciences*, 19(10), p. 3127. Available at: https://doi.org/10.3390/ijms19103127.

Kumar, N. and Goel, N. (2019) ‘Phenolic acids: Natural versatile molecules with promising therapeutic applications’, *Biotechnology Reports*, 24, p. e00370. Available at: https://doi.org/10.1016/j.btre.2019.e00370.

Kumar, N., Hoque, Md.A. and Sugimoto, M. (2018) ‘Robust volcano plot: identification of differential metabolites in the presence of outliers’, *BMC Bioinformatics*, 19(1), p. 128. Available at: https://doi.org/10.1186/s12859-018-2117-2.

Lahlou, M. (2013) ‘The Success of Natural Products in Drug Discovery’, 2013. Available at: https://doi.org/10.4236/pp.2013.43A003.

Laíns, I. *et al.* (2019) ‘Metabolomics in the study of retinal health and disease’, *Progress in Retinal and Eye Research*, 69, pp. 57–79. Available at: https://doi.org/10.1016/j.preteyeres.2018.11.002.

Lautié, E. *et al.* (2020) ‘Unraveling Plant Natural Chemical Diversity for Drug Discovery Purposes’, *Frontiers in Pharmacology*, 11. Available at: https://www.frontiersin.org/articles/10.3389/fphar.2020.00397 (Accessed: 22 June 2023).

Li, B. *et al.* (2017) ‘Novel Applications of Metabolomics in Personalized Medicine: A Mini-Review’, *Molecules*, 22(7), p. 1173. Available at: https://doi.org/10.3390/molecules22071173.

Li, C. *et al.* (2017) 'Anticancer activities of harmine by inducing a pro-death autophagy and apoptosis in human gastric cancer cells', *Phytomedicine*, 28, pp. 10–18. Available at: https://doi.org/10.1016/j.phymed.2017.02.008.

Li, J.W.-H. and Vederas, J.C. (2009) 'Drug Discovery and Natural Products: End of an Era or an Endless Frontier?', *Science*, 325(5937), pp. 161–165. Available at: https://doi.org/10.1126/science.1168243.

Ligon, B.L. (2004) 'Penicillin: its discovery and early development', *Seminars in Pediatric Infectious Diseases*, 15(1), pp. 52–57. Available at: https://doi.org/10.1053/j.spid.2004.02.001.

Liu, J. *et al.* (2019) 'Antitumor activity of alantolactone in lung cancer cell lines NCI-H1299 and Anip973', *Journal of Food Biochemistry*, 43(9), p. e12972. Available at: https://doi.org/10.1111/jfbc.12972.

Longley, D.B., Harkin, D.P. and Johnston, P.G. (2003) '5-Fluorouracil: mechanisms of action and clinical strategies', *Nature Reviews Cancer*, 3(5), pp. 330–338. Available at: https://doi.org/10.1038/nrc1074.

McIntosh, H.M. and Olliaro, P. (2000) 'Artemisinin derivatives for treating uncomplicated malaria', *The Cochrane Database of Systematic Reviews*, 1999(2), p. CD000256. Available at: https://doi.org/10.1002/14651858.CD000256.

Miranda, C.L. *et al.* (2016) 'Xanthohumol improves dysfunctional glucose and lipid metabolism in diet-induced obese C57BL/6J mice', *Archives of Biochemistry and Biophysics*, 599, pp. 22–30. Available at: https://doi.org/10.1016/j.abb.2016.03.008.

Mukund, V. *et al.* (2017) 'Genistein: Its role in metabolic diseases and cancer', *Critical Reviews in Oncology/Hematology*, 119, pp. 13–22. Available at: https://doi.org/10.1016/j.critrevonc.2017.09.004.

Newman, D.J. and Cragg, G.M. (2012) 'Natural Products As Sources of New Drugs over the 30 Years from 1981 to 2010', *Journal of Natural Products*, 75(3), pp. 311–335. Available at: https://doi.org/10.1021/np200906s.

Newman, D.J. and Cragg, G.M. (2016) 'Natural Products as Sources of New Drugs from 1981 to 2014', *Journal of Natural Products*, 79(3), pp. 629–661. Available at: https://doi.org/10.1021/acs.jnatprod.5b01055.

Nguyen, P.L. and Cho, J. (2021) 'Pathophysiological Roles of Histamine Receptors in Cancer Progression: Implications and Perspectives as Potential Molecular Targets', *Biomolecules*, 11(8), p. 1232. Available at: https://doi.org/10.3390/biom11081232.

Ning, J. *et al.* (2015) 'Characterization of Phase I Metabolism of Resibufogenin and Evaluation of the Metabolic Effects on Its Antitumor Activity and Toxicity', *Drug Metabolism and Disposition*, 43(3), pp. 299–308. Available at: https://doi.org/10.1124/dmd.114.060996.

Ortmayr, K., Dubuis, S. and Zampieri, M. (2019) 'Metabolic profiling of cancer cells reveals genome-wide crosstalk between transcriptional regulators and metabolism', *Nature Communications*, 10(1), p. 1841. Available at: https://doi.org/10.1038/s41467-019-09695-9.

Pereira Braga, C. and Adamec, J. (2019) 'Metabolome Analysis', in S. Ranganathan et al. (eds) *Encyclopedia of Bioinformatics and Computational Biology*. Oxford: Academic Press, pp. 463–475. Available at: https://doi.org/10.1016/B978-0-12-809633-8.20134-9.

Ras, R.T., Geleijnse, J.M. and Trautwein, E.A. (2014) 'LDL-cholesterol-lowering effect of plant sterols and stanols across different dose ranges: a meta-analysis of randomised controlled studies', *British Journal of Nutrition*, 112(2), pp. 214–219. Available at: https://doi.org/10.1017/S0007114514000750.

Ribbenstedt, A., Ziarrusta, H. and Benskin, J.P. (2018) 'Development, characterization and comparisons of targeted and non-targeted metabolomics methods', *PLOS ONE*, 13(11), p. e0207082. Available at: https://doi.org/10.1371/journal.pone.0207082.

Rosato, A. *et al.* (2018) 'From correlation to causation: analysis of metabolomics data using systems biology approaches', *Metabolomics*, 14(4), p. 37. Available at: https://doi.org/10.1007/s11306-018-1335-y.

Saadeldeen, F.S.A. *et al.* (2020) 'Natural products: Regulating glucose metabolism and improving insulin resistance', *Food Science and Human Wellness*, 9(3), pp. 214–228. Available at: https://doi.org/10.1016/j.fshw.2020.04.005.

Schmidt, D.R. *et al.* (2021) 'Metabolomics in cancer research and emerging applications in clinical oncology', *CA: A Cancer Journal for Clinicians*, 71(4), pp. 333–358. Available at: https://doi.org/10.3322/caac.21670.

Sha, L. *et al.* (2021) 'Shikonin inhibits the Warburg effect, cell proliferation, invasion and migration by downregulating PFKFB2 expression in lung cancer', *Molecular Medicine Reports*, 24(2), pp. 1–10. Available at: https://doi.org/10.3892/mmr.2021.12199.

Shi, H. *et al.* (2016) 'Lanatoside C promotes foam cell formation and atherosclerosis', *Scientific Reports*, 6(1), pp. 1–11.

Singh, S.K. *et al.* (2017) 'Resveratrol induces cell cycle arrest and apoptosis with docetaxel in prostate cancer cells via a p53/ p21WAF1/CIP1 and p27KIP1 pathway', *Oncotarget*, 8(10), pp. 17216–17228. Available at: https://doi.org/10.18632/oncotarget.15303.

Skelton, D.M. *et al.* (2014) 'Metabolomics for *in Situ* Environmental Monitoring of Surface Waters Impacted by Contaminants from Both Point and Nonpoint Sources', *Environmental Science & Technology*, p. 140115140646006. Available at: https://doi.org/10.1021/es404021f.

Stork, C. *et al.* (2020) 'NERDD: a web portal providing access to in silico tools for drug discovery', *Bioinformatics*, 36(4), pp. 1291–1292. Available at: https://doi.org/10.1093/bioinformatics/btz695.

Sussulini, A. (ed.) (2017) *Metabolomics: From Fundamentals to Clinical Applications*. Cham: Springer International Publishing (Advances in Experimental Medicine and Biology). Available at: https://doi.org/10.1007/978-3-319-47656-8.

Swann, J.P. (1983) 'The Search for Synthetic Penicillin during World War II', *The British Journal for the History of Science*, 16(2), pp. 154–190. Available at: https://doi.org/10.1017/S0007087400026789.

Tu, S. *et al.* (2018) 'Effect of taurine on cell proliferation and apoptosis human lung cancer A549 cells', *Oncology Letters*, 15(4), pp. 5473–5480. Available at: https://doi.org/10.3892/ol.2018.8036.

Wall, M.E. (1998) 'Camptothecin and taxol: Discovery to clinic', *Medicinal Research Reviews*, 18(5), pp. 299–314. Available at: https://doi.org/10.1002/(SICI)1098-1128(199809)18:5<299::AID-MED2>3.0.CO;2-O.

Wang, W. *et al.* (2019) 'Targeted Metabolomics Identifies the Cytochrome P450 Monooxygenase Eicosanoid Pathway as a Novel Therapeutic Target of Colon Tumorigenesis', *Cancer Research*, 79(8), pp. 1822–1830. Available at: https://doi.org/10.1158/0008-5472.CAN-18-3221.

Wen, W. *et al.* (2010) 'Suppression of Cyclin D1 by Hypoxia-Inducible Factor-1 via Direct Mechanism Inhibits the Proliferation and 5-Fluorouracil–Induced Apoptosis of A549 Cells', *Cancer Research*, 70(5), pp. 2010–2019. Available at: https://doi.org/10.1158/0008-5472.CAN-08-4910.

Wishart, D.S. (2008) 'Applications of Metabolomics in Drug Discovery and Development', *Drugs in R & D*, 9(5), pp. 307–322. Available at: https://doi.org/10.2165/00126839-200809050-00002.

Wishart, D.S. *et al.* (2013) 'HMDB 3.0--The Human Metabolome Database in 2013', *Nucleic Acids Research*, 41(Database issue), pp. D801-807. Available at: https://doi.org/10.1093/nar/gks1065.

Wishart, D.S. (2016) 'Emerging applications of metabolomics in drug discovery and precision medicine', *Nature Reviews Drug Discovery*, 15(7), pp. 473–484. Available at: https://doi.org/10.1038/nrd.2016.32.

Wollenberger, A. (1955) 'Action of protoveratrine on the metabolism of cerebral cortex. 1. Unstimulated cerebral-cortex tissue', *Biochemical Journal*, 61(1), pp. 68–77.

Wu, S. *et al.* (2023) 'Gastrodin and Gastrodigenin Improve Energy Metabolism Disorders and Mitochondrial Dysfunction to Antagonize Vascular Dementia', *Molecules*, 28(6), p. 2598. Available at: https://doi.org/10.3390/molecules28062598.

Xiao, D. *et al.* (2019) 'Emodin improves glucose metabolism by targeting microRNA-20b in insulin-resistant skeletal muscle', *Phytomedicine*, 59, p. 152758. Available at: https://doi.org/10.1016/j.phymed.2018.11.018.

Yang, Y. *et al.* (2016) 'Genistein inhibits A549 human lung cancer cell proliferation via miR-27a and MET signaling', *Oncology Letters*, 12(3), pp. 2189–2193. Available at: https://doi.org/10.3892/ol.2016.4817.

Yerges-Armstrong, L.M. *et al.* (2013) 'Purine Pathway Implicated in Mechanism of Resistance to Aspirin Therapy: Pharmacometabolomics-Informed-Pharmacogenomics', *Clinical pharmacology and therapeutics*, 94(4), pp. 525–532. Available at: https://doi.org/10.1038/clpt.2013.119.

Yin, J. *et al.* (2008) 'Berberine improves glucose metabolism through induction of glycolysis', *American Journal of Physiology-Endocrinology and Metabolism*, 294(1), pp. E148–E156. Available at: https://doi.org/10.1152/ajpendo.00211.2007.

Zhang, F. *et al.* (2016) 'Anticancer function of α-solanine in lung adenocarcinoma cells by inducing microRNA-138 expression', *Tumor Biology*, 37(5), pp. 6437–6446. Available at: https://doi.org/10.1007/s13277-015-4528-2.

Zhang, Z. *et al.* (2020) 'Integrated transcriptomic and metabolomic analyses to characterize the anti-cancer effects of (−)-epigallocatechin-3-gallate in human colon cancer cells', *Toxicology and Applied Pharmacology*, 401, p. 115100. Available at: https://doi.org/10.1016/j.taap.2020.115100.

Zou, T. *et al.* (2021) *Alpha-Solanine Anti-Tumor Effects in Non-Small Cell Lung Cancer Through Regulating the Glycolysis Pathway*. preprint. In Review. Available at: https://doi.org/10.21203/rs.3.rs-789534/v1.