Metabolomics in Drug Discovery: Cancer Cells Metabotyping to Predict the Mode of Action of Anticancer Agents

Dissertation

zur Erlangung des

Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

Der

Naturwissenschaftlichen Fakultät I

Biowissenschaften

der Martin-Luther-Universität

Halle-Wittenberg,

vorgelegt

Herr MSc. Mohamad Saoud

Halle (Saale), 2025

This thesis is based on the work done in the working groups of Prof. Ludger Wessjohann and Prof. Alain Tissier at the Leibniz Institute of Plant Biochemistry, Halle (Saale).

Thesis supervisor and editor: **Prof. Ludger A. Wessjohann**

Co- supervisor: Dr. Gerd U. Balcke and Dr. Robert Rennert

1st thesis reviewer: Prof. Alain Tissier (MLU)

2nd thesis reviewer: Prof. Ludger A. Wessjohann (IPB)

Part of the work reported here has been accepted for publication in *Advanced Science* 2024.

Advancing Anticancer Drug Discovery: Leveraging Metabolomics and Machine Learning for Mode of Action Prediction by Pattern Recognition

Mohamad Saoud, Jan Grau, Robert Rennert, Thomas Mueller, Mohammad Yousefi, Mehdi D. Davari, Bettina Hause, René Csuk, Luay Rashan, Ivo Grosse, Alain Tissier, Ludger A. Wessjohann,* and Gerd U. Balcke*

Adv. Sci. 2024, 2404085 DOI: 10.1002/advs.202404085

Datum der Disputation: 02.06.2025

Abstract:

In the development of new anticancer drugs, the identification of the mode of action (MoA) remains a significant challenge. This thesis demonstrates the integration of metabolomics into the drug discovery pipeline to predict the MoAs of novel anti-proliferative drug candidates, specifically for human prostate cancer cells (PC-3). By studying 38 drugs known to affect 16 key processes of cancer metabolism, we profiled low molecular weight intermediates of the central carbon and cellular energy metabolism (CCEM) using LC-MS/MS. These metabolic patterns revealed distinct MoAs, enabling the accurate prediction of MoAs for novel agents through machine learning algorithms. The methodology was further validated by transferring MoA predictions to two other cancer cell models, breast cancer and Ewing's sarcoma, confirming that correct MoA predictions across different cancer types are feasible, albeit with some reduction in prediction quality. Additionally, the metabolic profiling of treated cells provided deeper insights into intracellular processes, particularly drugs that induce various types of mitochondrial dysfunction. For instance, our findings suggest that pentacyclic triterpenoids inhibit oxidative phosphorylation and influence phospholipid biosynthesis, as corroborated by respiration parameters, lipidomics, and molecular docking studies. Leveraging biochemical insights from individual drug treatments, this research paves the way for new strategies in cancer treatment, including the optimization of combinatorial drug therapies, or enhancing the potential for targeted or synergistic approaches. This study exemplifies how incorporating metabolomics into the drug discovery process can substantially improve the characterization and development of new therapeutic agents.

Zusammenfassung:

Die Bestimmung des Wirkmechanismus (Mode of Action, MoA) ist eine schwierige Herausforderung bei der Entwicklung neuer Krebsmedikamente. Diese Dissertation demonstriert die Integration von Metabolomik in den Arzneimittelentwicklungsprozess, um die MoAs neuartiger antiproliferativer Wirkstoffkandidaten, insbesondere in menschlichen Prostatakrebszellen (PC-3), vorherzusagen. Durch die Untersuchung von 38 Wirkstoffen, die bekanntermaßen 16 Schlüsselprozesse des Stoffwechsels von Krebszellen beeinflussen, haben wir niedermolekulare Zwischenprodukte des zentralen Kohlenstoff- und Energiestoffwechsels (CCEM) mittels LC-MS/MS analysiert. Die durch die getesteten Substanzen veränderten metabolischen Muster offenbarten eindeutige MoAs und ermöglichten durch maschinelles Lernen eine präzise Vorhersage der MoAs für neue Wirkstoffe. Die Methodik wurde validiert, indem MoA-Vorhersagen auf zwei andere Krebszellmodelle, Brustkrebs und Ewing-Sarkom, übertragen wurden. Dies bestätigte, dass korrekte MoA-Vorhersagen über verschiedene Krebstypen hinweg machbar sind, wenn auch mit einer etwas geringeren Vorhersagequalität. Zudem lieferte das metabolische Profiling der behandelten Zellen tiefere Einblicke in intrazelluläre Prozesse, insbesondere bei Wirkstoffen, die verschiedene Arten von mitochondrialer Dysfunktion induzieren. Beispielsweise legen unsere Befunde nahe, dass pentazyklische Triterpenoide die oxidative Phosphorylierung hemmen und die Phospholipidbiosynthese beeinflussen, was durch Respirationsparameter, Lipidomik und molekulare Dockingstudien bestätigt wurde. Durch die Nutzung biochemischer Einblicke in einzelne Arzneimittelbehandlungen ebnet diese Forschung den Weg für neue Behandlungsstrategien, wie z.B. durch die Optimierung von kombinatorischen Arzneimitteltherapien, wodurch das Potenzial für gezielte Krebsbehandlungen verbessert wird. Diese Studie veranschaulicht, wie die Einbeziehung der Metabolomik in den Arzneimittelentwicklungsprozess die Charakterisierung und Entwicklung neuer therapeutischer Wirkstoffe erheblich verbessern kann

Acknowledgment:

I extend my deepest gratitude to my family for their unwavering support and patience throughout the course of my studies. Their unconditional love and sacrifices have been the cornerstone of my resilience and success. I am profoundly grateful for their belief in my abilities and the comforts they have provided me during this journey.

I am immensely grateful for the opportunity to collaborate with the esteemed team at the Leibniz Institute of Plant Biochemistry (IPB), especially Prof. Ludger Wessjohann, Prof. Alan Tissier, Dr. Gerd Balcke and Dr. Robert Rennert. Their insightful guidance and unwavering support have been instrumental in the advancement of my research. I deeply appreciate the enriching environment and resources provided by IPB, which have been crucial to my professional and personal growth.

I am grateful to Prof. Ivo Grosse and Dr. Jan Grau (Institute of Computer Science, Martin Luther University Halle-Wittenberg) for their assistance with the machine learning components of this project, and to Dr. Mahdi Davari and Mohammad Yousefi (Leibniz Institute of Plant Biochemistry, Dept. of Bioorganic Chemistry) for their critical contributions to the docking studies. Special thanks also to Dr. Thomas Müller (Medical Faculty, University Clinic for Internal Medicine IV (Hematology/Oncology), Martin Luther University Halle-Wittenberg) for his proficiency with the Seahorse assay. The expertise and support of each have been crucial to the success of this work. Their proficiency and insightful feedback have been pivotal in the successful application of advanced analytical techniques to my research. I am deeply appreciative of their commitment and invaluable contributions.

I am deeply thankful to my lab mates Martina Lerbs and Ibrahim Morgan for their invaluable support and encouragement during my research. Their shared knowledge and collaborative spirit have greatly enhanced both my work and my experience in the lab. I am grateful for the camaraderie and inspiration they have provided throughout this journey.

Table of Contents

Table of Figures:

List of Abbreviations:

1,3-BPG	1,3-Bisphosphoglycerate	**IMPDH**	Inosine 5'-monophosphate dehydrogenase
2DNP	2,4-Dinitrophenol	**IPC**	Ion pair chromatography
2-HG	2-Hydroxyglutarate	**IRIN**	Irinotecan
2-PGA	2-Phosphoglycerate	**ISOCIT**	Isocitrate
3-NP	3-Nitropropionic acid	**kNN**	k-Nearest neighbor classifier
3-PGA	3-Phosphoglyceric acid	**Lasso**	Least absolute shrinkage and selection operator
5-FU	5-Fluorouracil	**LC**	Liquid chromatography
6-AN	6-Aminonicotinamide	**LDH**	Lactate dehydrogenase
6-PG	6-phosphogluconate	**LOVA**	Lovastatin
6-PGDH	6-Phosphogluconate dehydrogenase	**LPC**	Monoacylglycerophosphocholines
6-PGL	6-Phosphogluconolactone	**LPE**	Monoacylglycerophosphoethanolamine
AAHR	Asiatic acid homopiperazinyl rhodamine	**LSCs**	Leukemic stem cells
ACC	Acetyl-CoA carboxylase	**m/z**	Mass-to-charge ratio
AcCoA	Acetyl CoA	**MAL**	Malate
ACLY	ATP citrate lyase	**MALO**	Malonic acid
ACN	Acetonitrile	**MASA**	Maslinic acid
ACT	Cis-aconitate	**MeA**	Mechanism of action
Adenylo-Suc	Adenylosuccinate	**METF**	Metformin
ADPPR	ADP-ribose 2'-phosphate	**MITO**	Mitoxantrone
ADPR	ADP-ribose	**ML**	Machine learning
α -KG	alpha-Ketoglutarate	**MoA**	Mode of action
AKT	Protein kinase B	**MPC**	Mitochondrial pyruvate carrier
ALPL	Alpelisib	**MRM**	Multiple reaction monitoring
AML	Acute myeloid leukemia	**MS or MS/MS**	Mass spectrometry
AMP	Adenosine monophosphate	**MTBE**	Methyl *tert*-butyl ether

AMYC	Antimycin A	**MTT**	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
APIG	Apigenin	**NAD**	Nicotinamide adenine dinucleotide
ASCL1	Achaete-Scute Homolog-1	**NADH**	Nicotinamide adenine dinucleotide hydrogen
ATCC	American Type Culture Collection	**NAMPT**	Nicotinamide phosphoribosyltransferase
ATOR	Atorvastatin	**NMR**	Nuclear magnetic resonance
ATOV	Atovaquone	**NQO1**	NAD(P)H dehydrogenase [quinone] 1
ATP	Adenosine triphosphate	**NSCLC**	Non-small cell lung cancer
AUC	Area under curve	**NTP**	Nucleoside triphosphate
AZID	Sodium azide	**OC**	Ovarian cancer
BETA	Betulinic acid	**OCR**	Oxygen consumption rate
BITN	Bithiniol	**oPLS-DA**	Orthogonal partial least squares discriminant analysis
BOWA	Boswellic acid	**OPP**	Oxidative pentose phosphate pathway
BRST	Breastin (patented *nerium oleander* cold water extract)	**ORID (OR)**	Oridonin
CA	Carbamoyl aspartate	**ORO**	Orotate
CAR	Acylcarnitines	**OXPHOS**	Oxidative phosphorylation
CCCP	Carbonyl cyanide chlorophenylhydrazone	**PBS**	Phosphate-buffered saline
CCEM	Central carbon and energy metabolism	**PC**	Phosphatitylcholines
CER	Ceramide	**PCA**	Principal component analysis
CIT	Citrate	**PDA**	Pancreatic ductal adenocarcinoma
CML	Chronic myeloid leukemia	**PDD**	Phenotypic drug discovery
CMPT	Camptothecin	**PDH**	Pyruvate dehydrogenase
CoQ	Coenzyme Q	**PE**	Diacylglycerophosphoethanolamines
CPLX I	Respiratory complex I	**PEA**	Phosphoethanolamine
CPLX II	Respiratory complex II	**PEP**	Phosphoenolypyruvate

CPLX III	Respiratory complex III	**PFA**	Paraformaldehyde
CPLX IV	Respiratory complex IV	**PG**	Phosphoglycerates
CPT-1	Choline phosphotransferase 1	**PI**	Phosphoinositides
CV	Crystal violet	**PKM2**	Pyruvate kinase M2
CYAN	Potassium cyanide	**PLB**	Phospholipid biosynthesis
DHAP	Dihydroxyacetone phosphate	**PLS-DA**	Partial least squares discriminant analysis
DHO	Dihydroorotate	**PPP**	Pentose phosphate pathway
DHODH	Dihydroorotate dehydrogenase	**PRFN (PR)**	Perifosine
DMSO	Dimethyl sulfoxide	**PTXL**	Paclitaxel
DNA	Deoxyribonucleic acid	**PYR**	Pyruvate
DOXO	Doxorubicin	**QQrB**	Cucurbitacin b
DT	Decision tree	**r.t.**	Retention times
ECAR	Extracellular acidification rate	**R5P**	5-phosphate
EDTA	Ethylenediaminetetraacetic acid	**RAPA**	Rapamycin
EGCG	Epigallocatechin gallate	**RNA**	Ribonucleic acid
EMOD	Emodin	**ROC**	Receiver operating characteristic
ER	Estrogen receptor	**ROTN**	Rotenone
ETC	Electron transport chain	**RP**	Reversed-phase
ETOP	Etoposide	**RT**	Room temperature
F6P	Fructose-6-phosphate	**Ru5P**	Ribulose-5-phosphate
FA	Free fatty acids	**S7P**	Sedoheptulose-7-phosphate
FAB	Fatty acid biosynthesis	**SAM**	S-adenosylmethionine
FADH2	Flavin adenine dinucleotide	**SCLC**	Small cell lung cancer
FASN	Fatty acid synthase	**SDH**	Succinate dehydrogenase
FBP	Fructose 1,6-bisphosphate	**SIT-DIMS**	Stable isotope tracer direct-infusion mass spectrometry
FCCP	Carbonyl cyanide 4-(trifluoromethoxy)phenylhydrazone	**SM**	Sphingomyelins

FCS	Fetal calf serum	**SUC**	Succinate
FDA	Food and Drug Administration	**SUC-CoA**	Succinyl-CoA
FLUV	Fluvastatin	**TBA**	Tetrabutylammonium
FUM	Fumarate	**TCA**	Tricarboxylic acid
G3P	Glyceraldehyde-3-phosphate	**TG**	Triacylglycerides
G6P	Glucose-6-phosphate	**TK**	Tyrosine kinase
G6PDH	Glucose 6-phosphate dehydrogenase	**TKIs**	Tyrosine kinase inhibitors
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	**TNBC**	Triple-negative breast cancer
GC	Gas ghromatography	**TOF**	Time of Flight
GDH	Glutamate dehydrogenase	**TopoI**	Topoisomerase I
Glc3P	Glycerol 3-phosphate	**TopoII**	Topoisomerase II
GLYA	Glycyrrhetinic acid	**TYMS**	Thymidylate synthase
GMP	Guanosine monophosphate	**UBE2C**	Ubiquitin-conjugating enzyme E2C
GPDi	Glucose-6-phosphate dehydrogenase inhibitor	**UbQ**	Ubiquinone-10
HER2	Human epidermal growth factor receptor 2	**Uncoupler**	Uncoupling of oxidative phosphorylation
HEXA	Hexachlorophorane	**UQH2**	Ubiquinol
HILIC	Hydrophilic interaction liquid chromatography	**VINC**	Vincristin
HMG-CoAr	Hydroxymethylglutaryl CoA reductase	**VIP**	Variable of importance
IC	Ion chromatography	**WRTN**	Wortmannin
IC50	Half maximal inhibitory concentration	**XMP**	Xanthosine monophosphate
IDH1/2	Isocitrate dehydrogenases 1/2	**Xu5P**	Xylulose-5-phosphate
IMP	Inosine monophosphate		

1 Introduction

1.1 Cancer and metabolism

1.1.1 Cancer definition and statistics

Cancer is fundamentally characterized by the abnormal growth of cells, caused by a multitude of genetic and environmental factors. These alterations disrupt the normal balance of cell proliferation, maintenance, and death, thereby favoring the growth of tumor cells. The primary distinction between malignant tumors and benign growths lies in their potential for invasion, metastasis, and significant impact on patients' health and mortality [1]. Various internal factors such as sporadic hereditary mutations, hormonal imbalances, and immune conditions, along with external influences such as lifestyle, exposure to chemical carcinogens, and radiation can play critical roles in cancer development. Interestingly, the majority of non-sporadic cancers are attributed to environmental and lifestyle factors, while only a small fraction of around ten percent directly caused by genetic defects [2].

Recent global statistics for 2022 highlight the persistent challenge of cancer with nearly 20 million new cases and nearly 9.7 million cancer-related deaths [3]. Incidence and mortality rates vary significantly across regions and cancer types, reflecting a complex interplay of genetic, environmental and socio-economic factors. Lung cancer emerged as the most frequently diagnosed cancer, accounting for almost 2.5 million new cases or about 12.4% of all cancer cases worldwide, followed by female breast cancer, colorectal, prostate, and stomach cancers. Notably, lung cancer also remained the leading cause of cancer related death with an estimated 1.8 million fatalities. In terms of gender-specific prevalence, breast cancer continued to be the most frequent cancer among women, while lung cancer was the most prevalent among men. Predictions indicate that cancer cases could rise to 35 million by 2050, representing a 77% increase from 2022 levels [3]. These statistics underscore the urgent need for effective cancer management strategies, encompassing prevention, early diagnosis, and advanced treatment modalities.

1.1.2 Cancer therapy approaches

The landscape of cancer therapies is diverse, including surgery, radiotherapy, chemotherapy, immunotherapy, and targeted therapy. Surgical interventions and radiotherapy are predominantly applied to localized tumors, often achieving significant cure rates [4]. Approximately 45% of newly diagnosed cancer cases may be treated effectively with radiotherapy, which accounts for about 40% of cancer cures. Radiotherapy, including external beam and brachytherapy, is used for different tumor types and is often combined with surgery and/or chemotherapy for more comprehensive treatment [5]. Conversely, advanced cancers with critical localization, infiltration into adjacent tissues, or suspected metastases require

systemic treatments such as chemotherapy, immunotherapy, and targeted therapies. These interventions aim to control disease spread and improve patient survival outcomes. Chemotherapy is used in primary, neoadjuvant and adjuvant settings. Primary chemotherapy is the first approach when no other options are available and aims to relieve symptoms and improve quality of life. Neoadjuvant chemotherapy is administered prior to the primary therapy to enhance its effectiveness. Adjuvant chemotherapy complements local treatment and is given after the primary therapy to reduce the risk of recurrence and improve overall survival [6]. Despite their benefits, these systemic therapies are frequently challenged by side effects, therapeutic resistance, and the need for personalized treatment strategies based on the molecular characteristics of the tumor. The evolution of cancer treatment reflects a growing understanding of the disease's complexity. Innovations in molecular biology and genetics have paved the way for targeted therapies and immunotherapies, offering hope for improved outcomes and personalized treatment plans. However, the need for curative therapies continues, as researchers strive to unravel the complexities of cancer biology and translate these insights into safer and more effective treatments.

The interconnection between cancer metabolism and therapeutic responses opens new horizons for cancer treatment. Understanding the metabolic adaptations of cancer cells provides a basis for the development of targeted therapies that can exploit the unique metabolic dependencies of tumors. The integration of metabolic tumor cell profile with personalized medicine approaches holds the potential to enhance the specificity and efficacy of cancer treatments, thereby improving the patients' outcomes.

1.1.3 Cancer metabolism

When Hanahan and Weinberg initially published their groundbreaking paper on the hallmarks of cancer in 2000 [7], they did not mention the significance of dysregulated cell metabolism. However, the accumulating body of evidences from numerous studies [8] along with a growing recognition of the Warburg effect's role in tumorigenesis [9], ultimately prompted them to incorporate dysregulated metabolism as one of the hallmarks of cancer in their revised paper in 2011 [10]. A century ago, the German biochemist Otto Warburg first identified the link between tumorigenesis and metabolic dysfunction. Warburg's findings showed that tumors have an increased glucose consumption compared to normal tissue. Interestingly, even when oxygen is abundantly available, tumors tend to prioritize the fermentation of glucose into lactate through the anaerobic glycolysis rather than its energetically more efficient oxidation within the mitochondria [11, 12]. Over the past two decades, the field of cancer metabolism has entered a dynamic and exciting new phase of research and discovery, thanks to the deeper understanding of the underlying processes that drive cell transformation. This progress is

further boosted by cutting-edge experimental methods, particularly the rapid advancements in OMICs technologies (please refer to section 1.3 for more information about OMICS approaches) [13-15]. Such discoveries improve our comprehension and characterization of cancer. Recent research, coinciding with the 100th anniversary of Otto Warburg's landmark report, highlights the multifaceted role of lactate. Lu et al. have shown that lactate, traditionally considered a metabolic byproduct, plays a pivotal role in promoting cell division, especially by influencing enzymes that regulate mitosis [16]. In particular, Lu and colleagues found that high lactate levels in cultured cancer cells lead to premature cell cycle progression and increased proliferation. Significantly, lactate accumulation may help tumor cells bypass chemotherapy-induced mitotic blockages, facilitating mitotic slippage [16]. This research also suggests that lactate, once considered just a consequence of rapid cell growth, indeed also actively drives cell cycle progression [16]. Additionally, lactate serves as a high-energy carbon source for anabolic synthesis and induces micro-environmental acidification [17].

A fundamental distinction can be made between normal proliferating cells and neoplastic (cancerous) cells. While uncontrolled growth of neoplastic cells is driven by oncogenes/ oncoproteins and is dependent on constant nutrient supply, normal cells have the ability to detect nutrient deprivation and respond by stopping proliferation. Therefore, while normal cells can enter the G0/G1 phase of the cell cycle as a protective response, nutrient deprivation could potentially lead to cell death in neoplastic cells [18-20]. Hence, cancer cells require elevated nutrient supply and energy-rich resources to sustain their high growth rate and proliferation. However, the finite availability of local resources often leads to nutrient depletion and accumulation of metabolic “waste” products. For adaptation, cancer cells undergo metabolic reprogramming, influenced by their origin, genetic alterations, and tissue environment [21, 22]. Some of these metabolic changes modulate their gene expression, protein function, or interactions with neighboring healthy cells [22]. Moreover, the impact of cancer cell metabolism on the systemic metabolism was explored by many researches, highlighting the interconnected metabolic wiring between tumors and distant tissues [13, 23, 24]. A notable study using isotope tracing and metabolomics (please refer to section 1.3 for more information on metabolomics) in adult zebrafish demonstrated how non-cancerous tissues are triggered to adjust their metabolism to accommodate tumor growth [25]. It was observed that the liver metabolizes alanine secreted by skin tumors promoting gluconeogenesis, resulting in a steady supply of circulating glucose that satisfies the increased glycolytic demands of the tumor. This research highlights the importance of metabolic interconnection between cancerous and healthy tissues and illustrates tumor’s capacity to manipulate host metabolism to its advantage [25].

Cancer is increasingly perceived as a metabolic disorder, marked by significant alterations in energy production and biosynthetic processes [26]. Research on the metabolic reprogramming of cancer cells, particularly focusing on the Warburg effect, has yielded several compelling hypotheses to explain proposed adaptation of enhanced glycolysis in cancer cells. Warburg initially hypothesized that tumors exhibit compromised respiration as a result of mitochondrial dysfunction [27]. However, contemporary research indicates that this is not the situation for the majority of cancers. Indeed, a significant body of evidence exists not only highlighting the retention of functional mitochondria in cancer cells but also underscoring the essential role of mitochondrial respiration for their growth, progression, and metastasis [15, 28, 29]. The shift towards enhanced glycolysis is also considered to be driven by tumor cells' need of faster adenosine triphosphate (ATP) production and the biosynthetic requirements of proliferating cells, even though it is less energy-efficient than oxidative phosphorylation (OXPHOS) [30]. It appears that cancer cells might derive more benefits from a quick availability of ATP and the precursors for macromolecule syntheses as provided by glycolysis [13, 31]. Many other hypotheses have been proposed to explain the anaerobic glycolysis in cancer cells, nevertheless, a key new discovery helped to shed light on the interpretation of these phenomena. The enhanced glycolytic activity is increasingly viewed in the context of redox balance, emphasizing the regeneration of oxidized electron carriers as a crucial aspect of cancer cell proliferation, rather than merely focusing on ATP generation [32, 33]. Luengo et al. propose that cancer cells increase glucose fermentation to maintain the nicotinamide adenine dinucleotide/nicotinamide adenine dinucleotide hydrogen (NAD^+/NADH) ratio, addressing a higher demand for NAD^+ than ATP during proliferation [34]. Pyruvate (PYR), the endpoint of glycolysis, can convert to lactate catalyzed by lactate dehydrogenase (LDH) or to acetyl-CoA via the activity of pyruvate dehydrogenase (PDH) in mitochondria. Nevertheless, a reduced NAD^+/NADH ratio inhibits PDH, driving cells to favor lactate production for NAD^+ regeneration. Luengo et al. discovered that enhancing PDH activity, thus directing PYR towards mitochondrial oxidation, impedes proliferation due to NAD^+ scarcity. Conversely, alternative NAD^+ regeneration pathways, like activation of NAD(P)H dehydrogenase [quinone] 1 (NQO1), can restore proliferation despite PDH's activity. This suggests the possibility of decoupling glycolysis from cell proliferation by modulating NAD^+ pathways. The Warburg effect appears to be more a result of NAD^+ requirements than ATP demands, highlighting a metabolic refinement in cancer cell adaptation. The electron carriers have a pivotal role in cellular metabolism orchestrating various biosynthetic processes crucial for cell growth and proliferation. The availability of $NAD(P)^+$ is not only crucial for glycolysis, but also required to generate precursors (NAD(P)H) for fatty acid biosynthesis. Zhaoqi Li and colleagues [35] further underscored this by demonstrating the essential role of environmental lipids in NAD^+ regeneration, especially under conditions of limited electron acceptors. Their findings highlight

that without external lipid sources, cells struggle to proliferate unless alternative NAD^+ regeneration mechanisms are employed, connecting lipid metabolism with the broader network of cellular energy and redox balance [35]. In this aspect, it is worth noting that nicotinamide phosphoribosyltransferase (NAMPT), a rate-limiting enzyme in the NAD salvage pathway, is overexpressed in a variety of cancers including breast, ovarian, and prostate cancer. Research has demonstrated that inhibiting NAMPT effectively depletes the NAD^+ and NADPH pools, which in turn alternate the reduced flux through glycolysis and may provide a strategic target for therapeutic intervention [36, 37].

Tumor cells intensively employ glycolysis, its branched pathways, and tricarboxylic acid (TCA) cycle metabolism (Figure 1) to produce fundamental building blocks necessary for biosynthesizing macromolecules such as nucleotides, lipids, and amino acids, which are indispensable for cellular proliferation [15]. The upregulation of key oncogenes like Myc and K-Ras, along with disruptions in signaling pathways such as PI3K, significantly contribute to the heightened glycolysis and TCA cycle activity which in turn fuel various biosynthetic pathways. These pathways include the pentose phosphate pathway (PPP), which is crucial for producing ribose and cytosolic NADPH, supporting nucleotide formation and maintaining antioxidant defenses, respectively. Furthermore, the engagement in one-carbon metabolism is vital for producing mitochondrial NADPH, nucleotides and facilitating methylation processes [38]. Concurrently, the intensified TCA cycle activity supplies crucial metabolites that are instrumental in the biosynthesis of nucleotides, lipids, amino acids and heme [39]. Interestingly, Sullivan et al. findings suggest that in certain proliferating cells, mitochondrial activity plays a more crucial role in the synthesis of macromolecular precursors than in the production of ATP [40]. An example is the role of oxaloacetate generated within the TCA cycle that can be converted to aspartate through a process that is pivotal for nucleotide production [41]. Nucleotide biosynthesis is essential for the proliferation of cancer cells, which constantly require the synthesis of purine and pyrimidine nucleotides for deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) synthesis to support the demands of rapidly replicating cells [42]. The nucleotides are synthesized through two main pathways: the energy-intensive *de novo* pathway, which builds nucleotides from small precursors through a series of enzymatic steps, and the salvage pathway, which recycles existing nucleosides and nucleobases into new nucleotides [43]. While the *de novo* pathway is typically dormant in non-proliferating cells, it is highly active in cancer cells, ensuring a steady supply of deoxy nucleoside triphosphate (d)NTPs necessary for the unbridled proliferation characteristic of cancer [43]. This heightened activity in cancer metabolism has been a target for therapeutic intervention since 1947 when Dr. Sidney Farber introduced aminopterin. As a folate analogue that impedes nucleotide biosynthesis, aminopterin demonstrated significant efficacy against acute lymphoblastic

leukemia [44-46]. This breakthrough paved the way for the development of antimetabolites, a novel drug category with the focus on exploiting cancer's metabolic vulnerabilities.

Another compelling study led by Prof. Gary Patti [32] suggested that the elevated rate of glycolysis observed in cancer cells does not function to support proliferation by actively suppressing glucose oxidation. Instead, it surpasses the maximal rate at which glucose can be oxidized, leading to fermentation. Thus, the metabolic phenotype of glucose fermentation is not inherently preferred by cancer cells but emerges as a secondary consequence of the saturation of the mitochondrial oxidative capacity.

In resting cells, nutrient import primarily fuels energy generation through carbon oxidation (Figure 1), liberating electrons that reduce cofactors such as $NAD(P)^+$, FAD, or ubiquinone-10 (UbQ) to NAD(P)H, flavin adenine dinucleotide ($FADH_2$), and ubiquinol-10, respectively. These reductions support biosynthesis and the mitochondrial electron transport chain (ETC), crucial for ATP production and for maintaining the redox balance [22]. Conversely, proliferating and cancer cells adapt their metabolism to prioritize the conversion of nutrients into building blocks while addressing energy needs [47]. The ETC's critical role involves recycling of NADH and $FADH_2$ back to NAD^+ and FAD, essential for the TCA cycle's functions [22]. A key ETC function is the interconversion of ubiquinol and ubiquinone (Figure 1), where the mitochondrial complexes I and II transfer electrons to ubiquinone converting it to ubiquinol. Subsequently, complex III reoxidizes ubiquinol to ubiquinone facilitating ATP synthesis and supporting the role of ubiquinone as an electron acceptor for enzymes like dihydroorotate dehydrogenase (DHODH), necessary for *de novo* pyrimidine biosynthesis [48-50]. Thus, efficient ubiquinone regeneration is critical for tumor growth *in vivo*, further emphasizing the ETC's importance. Additionally, the operational integrity of mitochondrial complex III is vital for cancer cell proliferation *in vitro*, underscoring the ETC's crucial role in maintaining the TCA cycle functionality and enabling DHODH activity, both essential for tumor development [51-53]. In this context, it is important to highlight that DHODH emerges as a promising therapeutic target in various cancers, particularly in prostate cancer [54]. Elevated DHODH transcript levels in prostate tumors have been linked to poor prognosis, and its inhibition has shown to suppress tumor progression. This underscores the therapeutic potential of DHODH inhibitors, either as standalone treatments or in combination with other chemotherapies, for advanced prostate cancer.

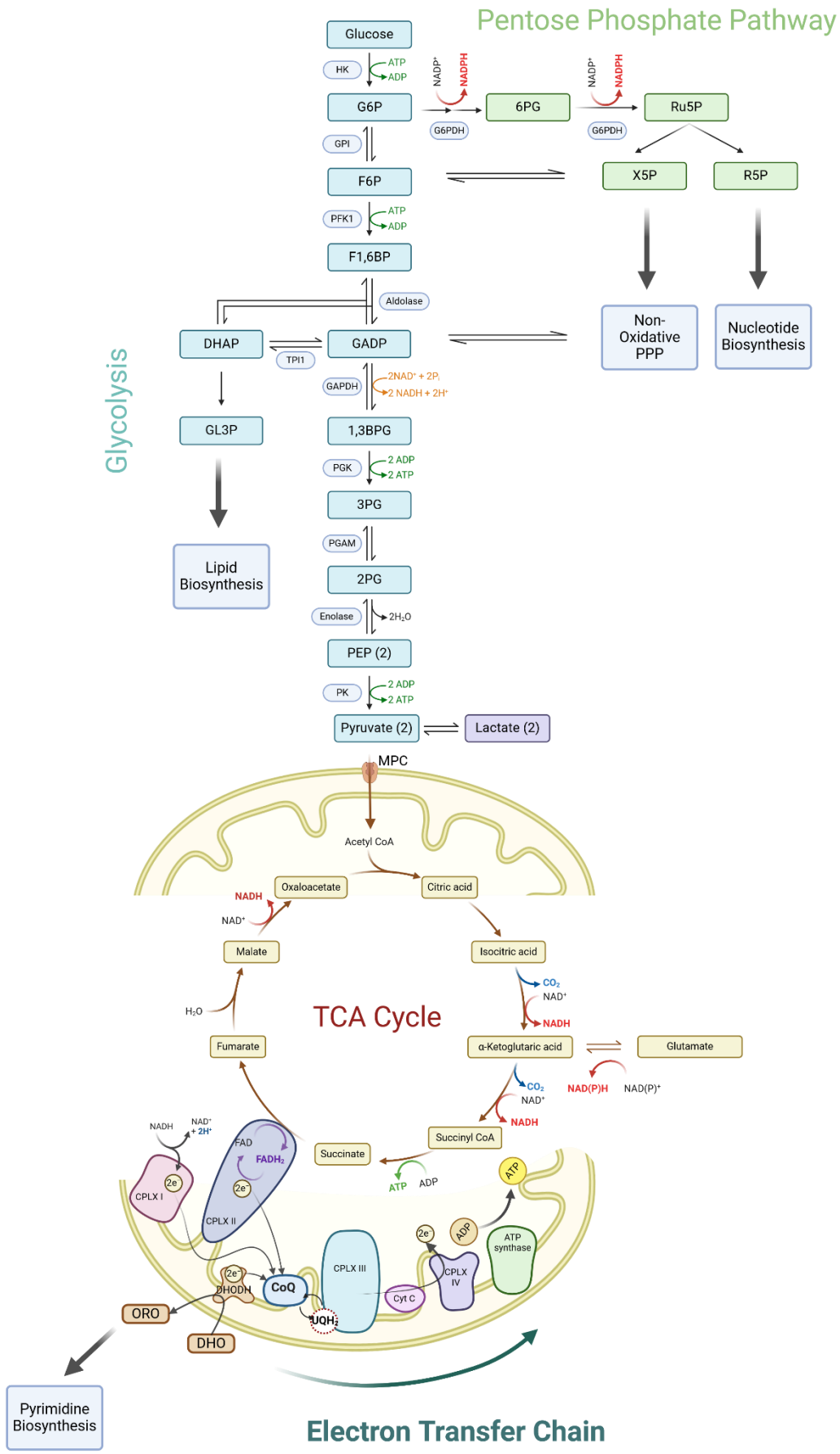

Pentose Phosphate Pathway
Glycolysis
Glucose
HK
ATP
ADP
G6P
GPI
F6P
PFK1
ATP
ADP
F1,6BP
Aldolase
DHAP
TPI1
GADP
GAPDH
2NAD+ + 2Pi
2 NADH + 2H+
1,3BPG
PGK
2 ADP
2 ATP
3PG
PGAM
2PG
Enolase
2H2O
PEP (2)
PK
2 ADP
2 ATP
Pyruvate (2)
Lactate (2)
NADP+
NADPH
6PG
G6PDH
Ru5P
X5P
R5P
Non-Oxidative PPP
Nucleotide Biosynthesis
GL3P
Lipid Biosynthesis
MPC
Acetyl CoA
Oxaloacetate
Citric acid
NADH
NAD+
Malate
Isocitric acid
H2O
CO2
TCA Cycle
Fumarate
α-Ketoglutaric acid
Glutamate
NAD(P)H
NAD(P)+
Succinyl CoA
Succinate
ATP
ADP
FAD
FADH2
2e-
CPLX I
CPLX II
CPLX III
CPLX IV
Cyt C
CoQ
UQH2
DHODH
ORO
DHO
ATP synthase
Pyrimidine Biosynthesis
Electron Transfer Chain

Figure 1: Main central carbon and energy metabolism (CCEM) pathways in (cancer) cells. Glucose enters cells via glucose transporters (GLUT) and is phosphorylated by hexokinases. The resulting glucose 6-phosphate (G6P) can be directed toward pyruvate generation or used for biosynthesis through pathways like the hexosamine biosynthesis pathway and the pentose phosphate pathway (PPP) [55]. PPP involves a series of intermediates, G6P initiates the pathway, and then it progresses through 6-phosphogluconolactone (6PGL) and 6-phosphogluconate (6PG). These intermediates are further metabolized into ribulose-5-phosphate (Ru5P) and xylulose-5-phosphate (Xu5P). Finally, PPP provides ribose 5-phosphate (R5P) for nucleotide synthesis and NADPH [56]. G6P subsequently transforms into fructose-6-phosphate (F6P) and then into fructose-1,6-bisphosphate (FBP) [57]. At this stage, dihydroxyacetone phosphate (DHAP) and glyceraldehyde-3-phosphate (G3P) are produced. G3P continues through the pathway to become 1,3-bisphosphoglycerate (1,3-BPG), followed by 3-phosphoglycerate (3-PG) and 2-phosphoglycerate (2-PG). Phosphoenolpyruvate (PEP) is the subsequent intermediate before reaching the final product, pyruvate (PYR). PYR has two fates: it can enter the tricarboxylic acid (TCA) cycle as acetyl-CoA via the mitochondrial pyruvate carrier (MPC) and pyruvate dehydrogenase (glucose oxidation) or it is reduced to lactate via pyruvate dehydrogenase (glucose fermentation) [28]. In the TCA cycle, citrate (CIT) formation initiates when acetyl-CoA combines with oxaloacetate. This results in the generation of citrate, which is then transformed into isocitrate. Isocitrate can undergo decarboxylation to produce alpha-ketoglutarate (α-KG), leading to the release of carbon dioxide. A-KG, in turn, undergoes decarboxylation to yield succinyl-CoA, generating NADH and releasing carbon dioxide. Succinyl-CoA is subsequently converted to succinate through a step that generates GTP or ATP via substrate-level phosphorylation. Lastly, succinate is transformed into fumarate, facilitated by the participation of the succinate dehydrogenase (SDH) complex (CPLX II), which functions in both the TCA cycle and the electron transport chain (ETC). Fumarate is transformed into malate, which is subsequently converted to oxaloacetate. The TCA cycle is reversible, and substrate oxidation reactions produce NADH and FADH2, donating electrons to complexes I and II. These electrons are passed through the ETC, ultimately reducing oxygen (O_2) to form water (H_2O). Complexes I, III, and IV pump protons across the inner mitochondrial membrane during this process, creating a proton gradient used by complex V (ATP synthase) to generate ATP from ADP. This mechanism is known as oxidative phosphorylation (OXPHOS) [48-50]. This graph was created with Biorender.com

The perception of cancer cell metabolism has evolved. It is currently seen not just as a cellular response to external factors but as a key determinant of cell fate and identity [33, 58]. This perspective is closely tied to the impact of specific metabolite availability on chromatin modification. In normal cells, methylation reactions rely on S-adenosylmethionine (SAM) as a methyl donor. Hence, a deficiency in methionine leads to reduced histone methylation. On the other hand, the removal of histone methylation involves a complex enzymatic process controlled by enzymes of the alpha-ketoglutarate (α-KG) dependent dioxygenase family, which utilize α-KG, an intermediate of the TCA cycle, as a co-substrate [59]. Consequently, a reduced availability of α-KG results in chromatin hypermethylation both *in vitro* and *in vivo* [58, 60]. Pathological accumulations of oncometabolites such as 2-hydroxyglutarate (2-HG), succinate, and fumarate are implicated in cancer progression, acting through well-defined mechanisms [61]. Particularly, 2-HG, produced by mutations in the TCA cycle enzymes isocitrate dehydrogenases 1/2 (IDH1/2), inhibits α-KG-dependent enzymes, leading to histone hypermethylation and disrupted cell differentiation. These mutations are associated with cancers like acute myeloid leukemia (AML) and have triggered the development of targeted therapies such as ivosidenib and enasidenib [62]. Despite therapeutic advances, the challenge of metabolic adaptability of tumors remains, with relapse often occurring as cancer cells adjust their metabolism to restore 2-HG levels [63]. Therefore, deciphering the convoluted relationship between metabolism and cellular function is central in understanding cancer biology and developing effective treatments.

The recognition that cancer cells have an altered regulation of central metabolic pathways led to the development of inhibitors with improved tumor specificity [64-68]. Metabolic inhibitors of cancer cell growth comprise important biosynthetic pathways (nucleotides, amino acids, lipids). However, there is a need for further research on cancer biology to identify effective screening methods and personalized therapeutic strategies that target metabolism in cancer cells [69].

1.2 Drug discovery approaches

In drug discovery, the primary objective is to identify an optimal agent characterized by selectivity, efficacy, bioavailability, safety, and cost-effectiveness. This agent should target "druggable" molecules that are accessible, critical to disease progression and modifiable with minimal side effects, and which are ideally specific disease-associated targets. A "target" is defined as a biological entity within the cell, as in case of DNA/RNA or a protein such as, transcription factor, ion channel, enzyme, or receptor, whose modulation due to binding or interaction with a pharmacological agent can lead to a therapeutic effect. Numerous methodologies for the evaluation of druggability have been introduced in scientific literature. These range from "guilt by association" approaches, where a protein is considered druggable if it belongs to a protein family that includes at least one member already targeted successfully by a drug, to techniques centered on binding site prediction [70].

Drug discovery strategies can be broadly classified into two categories: target-based (hypothesis-driven) and phenotypic (discovery-driven) (Figure 2). Target-based strategies rely on the identification of specific molecular targets that are implicated in the pathogenesis of a disease and the development of drugs that can modulate their pathogenic activity [71]. The advent of molecular medicine, supported by significant advancements in molecular biology techniques and the integration of advanced computational methodologies, has paved the way for the evolution of rational drug design, which in turn catalyzes the progression towards targeted-based drug discovery approaches [72-74]. It is worth noting that this approach requires screening campaigns specifically aimed at identifying modulators for the target under investigation. Phenotypic drug discovery (PDD) strategies are based on the inspection of phenotypic post-treatment changes in model organisms, guiding the selection of efficacious treatments. Phenotypic strategies are based on the systematic screening of chemical libraries to identify compounds that induce desirable phenotypic changes in relevant biological systems. Following the initial screening for hit compounds, a comprehensive elucidation of the mechanism of action (MeA) and mode of action (MoA) is undertaken [75]. MeA refers to the specific biochemical interaction between a compound and its molecular targets, while MoA refers to the physiological effects caused by the compound of study. On the other hand, the

interaction of ligands with their target biomolecules can be defined as target engagement (TE). Many studies showed that phenotypic strategies have been more successful in discovering novel first-in-class medicines, especially for complex diseases, because they can uncover unexpected MoA without preconceived hypotheses [75-77]. Moreover, a recent systematic review of 32,000 publications spanning 150 years of research demonstrated that target-based drug discovery is dominant, however, not very efficient. Only 9.4% of small molecule drugs are the result of target-based assays [78]. Therapeutic effects often rely on multiple mechanisms beyond the intended target. This suggests that target-based approaches may delay drug development since the fundamental challenge lies in the identification of a disease-relevant and -specific target, which, however, remains unknown for many diseases.

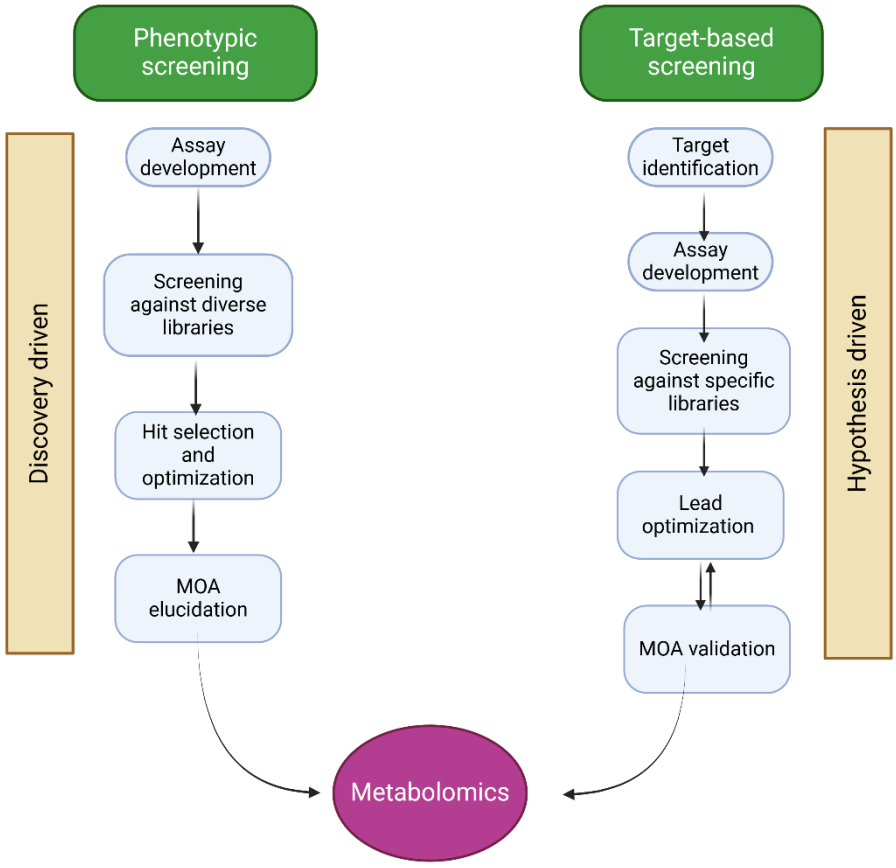


Figure 2: Two approaches commonly used in drug discovery research. Phenotypic screening (left) and target-based screening (right). This graph was created with Biorender.com

Phenotypic drug discovery (PDD) is supported by the high proportion of first-in-class medicines that originate from this approach [79]. Unlike target-based drug discovery that relies on a known causal link between a molecular target and a disease condition, PDD uses

chemical probes to explore disease-relevant biological systems without prior knowledge of a molecular target, the modulation of which could treat the disease [80]. This empirical, biology-first strategy provides distinctive molecules that can address the therapeutic biology of novel signaling pathways, molecular mechanisms, and drug targets. Phenotypic screening also enables the discovery of molecules that modulate multiple targets, also known as polypharmacology. In this case, the desired effect of a compound depends on a combination of targets (on-targets). However, these may not represent its complete target profile that may also include targets that are not essential for the desired activity (off-targets) [78].

Conventionally, polypharmacology has been perceived rather as a drawback of poorly optimized compounds because off-target effects that can be the reason of undesired therapeutic adverse effects. However, most, if not all, approved drugs interact with multiple targets at therapeutically relevant concentrations that may contribute to both the primarily desired therapeutical effect and undesired side effects. However, a simultaneous low potency modulation of several targets achieving an advanced efficacy “by synergy” has been proposed as a strategy to minimize side effects [75, 81].

The *Anticancer Agents and Targeting* research group at the Leibniz Institute of Plant Biochemistry (IPB) focusses on the identification, characterization and optimization of bioactive natural compounds and improved synthetic derivatives (based on over 25,000 compounds from plants or fungi and more than 5,000 plant extracts). With such a library of structurally diverse natural and synthetic compounds, phenotypic screening can be the method of choice to enhance the drug discovery research at the IPB. It is usually the case that screening campaigns result in many hit compounds with various activities and selectivities against a cancer model. Hence, besides the efforts to conduct screening studies, follow-up studies need to be carried out to identify the MoA for the selected promising hit compounds. Such studies can be laborious and time-consuming. Thus, a suitable methodology for the determination of underlying MoAs is needed to reduce the development times, costs and need of wet lab resources. Ideally, such a methodology would fasten the drug discovery process by identifying the MoA either by direct prediction or by further precise and targeted study guided by the outcome of the initial MoA prediction method.

1.3 Application of OMICS approaches in drug discovery

Inspired by the term "genomics," several related OMICS terms have emerged over the past two decades including transcriptomics, proteomics, metabolomics, glycomics, and lipidomics [82]. OMICS technologies rely on high-throughput analytical methods, which share the survey of hundreds to thousands of features by a single analysis. Thereby, the type of analyte can be

of nucleotidic, peptidic, or small molecule nature. A second feature common to all OMICS approaches is the direct comparison of patterns between diverse treatment groups, which aims at the biological and molecular interpretation of regulatory patterns in comparative OMICS. The evolution of OMICS technologies and their implementation in biomedical and pharmaceutical research is shaping the post-genomic era. Metabolomics is the approach that most directly connects the gap between genotype and phenotype, which is often not resolved by genomic, transcriptomic or proteomic studies.

The term metabolomics was first introduced by Steven Oliver in a 1998's review article on functional genomics in yeast and is now widely accepted as a key approach of systems biology [83]. Metabolomics involves a high-throughput analysis of the metabolome, which includes all detectable small molecules (metabolites) within a mass range of 50 to approximately 1500 Dalton, exhibiting a broad spectrum of physiochemical properties and varying abundances [84, 85]. Metabolites are found inside or outside of cells, or even in specific cell compartments, biofluids, tissues, or entire organisms serving as pivotal intermediates or end products of metabolic pathways. They can be excreted and can be found in saliva, feces and urine. They encompass a diverse array of endogenous compounds, including amino acids, lipids, carbohydrates, nucleotides, amines and organic acids, and their profiles reflect the dynamic biochemical processes of the organism. Hence, the identification of metabolites and their concentrations directly reflects the biochemical phenotype [86]. Thus, metabolomics is an extremely powerful OMICS technique that is used in cancer research to effectively identify metabolites whose levels are affected by neoplastic progression [87]. Metabolomics has a wide range of applications, for instance in quantitative phenotyping to support biomarker discovery, drug development and discovery, clinical toxicology or nutritional investigations [64, 69]. The acquisition of metabolic data can be categorized into two main approaches: targeted and untargeted (Figure 3). The targeted approach involves a pre-selection of specific mass-to-charge (m/z) ratios and retention times (r.t.) for analysis, focusing on a pre-defined set of metabolites. In a way it is nothing else than conventional chemical analysis of defined constituents. This approach is hypothesis-driven, demanding prior knowledge about specific metabolic pathways and necessitating customized analytical protocols. However, the increasing availability of metabolite libraries has facilitated the development of comprehensive targeted methods [88]. Conversely, the untargeted approach employs advanced mass spectrometry (MS, MS/MS or MS^n) techniques, or other, mostly spectroscopic, techniques to broadly detect numerous metabolites in a single chromatographic run generating relevant spectral data for a quantification after the analysis (hypothesis-generative) [89]. This typically involves spectral database searches, for instance, by using METLIN or MS-DIAL, and comparisons to authentic standards. Targeted methods offer the advantage of straightforward

metabolite identification, while untargeted metabolomics analyses require additional steps of identification and often require pure chemical standards for verification and quantification, which can pose challenges in obtaining them [89].

Nonetheless, even with the continuous technological advancements, none of the existing analytical platforms can comprehensively analyze the entire metabolome. Achieving sufficient coverage of the metabolome is still a significant obstacle, and this can only be accomplished by employing a combination of various approaches and multiple techniques due to the wide spectrum of chemical properties and abundance of the metabolites [90-93].

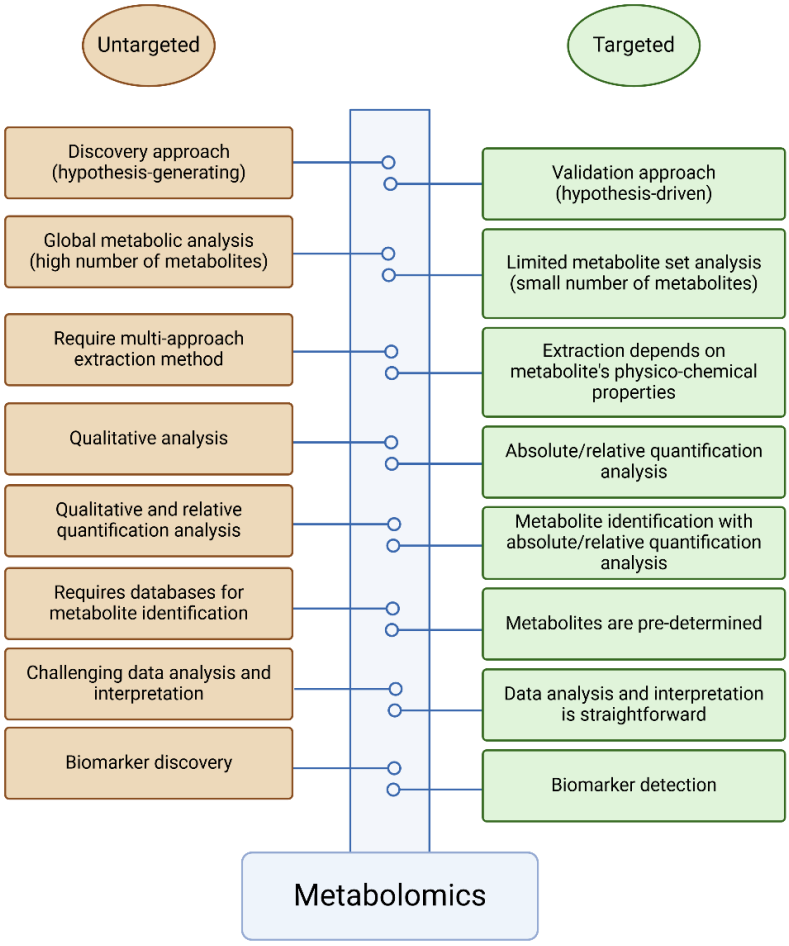


Figure 3: The main approaches for data acquisition in metabolomics: untargeted and targeted methodology. This graph was created with Biorender.com

In recent decades, significant technological advancements, primarily involving MS and nuclear magnetic resonance (NMR), have significantly improved the measurement of metabolites [94, 95]. In the branch of MS-based metabolomics, untargeted metabolomics often opts for high-resolution mass analyzers like the orbitrap and time of flight (TOF) devices due to their

exceptional mass accuracy. Conversely, targeted metabolomics typically leans towards high sensitivity instruments using single quadrupole or triple quadrupole [92].

While NMR spectroscopy is esteemed for its capacity to elucidate molecular structures, its sensitivity remains significantly lower than that of mass spectrometry, even with rapid advancements in the technology [96, 97]. In contrast, MS-based metabolomics has emerged as the predominant choice in metabolomics research due to its high sensitivity, extensive coverage, and adaptability for integration with various chromatographic techniques, such as liquid chromatography (LC) and gas chromatography (GC). This integration enhances the separation and quantification of metabolites, making MS-based metabolomics the preferred approach [95, 98, 99]. Typically, chromatographic separation precedes MS analysis and is accomplished using chromatographic columns packed with various materials. In the field of metabolomics, the most frequently employed LC columns are reversed-phase columns (RP) and hydrophilic interaction liquid chromatography (HILIC) columns [69]. RP columns with C18 are well-suited for metabolomics analyses due to their robustness and high separation power for semipolar compounds, including, e.g. phenolic acids, flavonoids, glycosylated steroids and other glycosylated substances, and alkaloids in plant research, while HILIC columns are dedicated to effectively separate water-soluble metabolites.

Another separation technique is ion chromatography (IC) which encompasses the separation of ions through ion exchange, exclusion, and pairing mechanisms. Ion pairing chromatography (IPC) separates hydrophilic or charged analytes using reversed phase, by masking the analyte's charge with amphiphilic counterions as ion-pairing reagents, such as tetrabutylammonium (TBA) hydroxide, contained in the mobile phase. These reagents, possessing charges opposite to the analytes, enable the formation of electrostatic bonds, pairing with them to create neutral, hydrophobic moieties. This process allows for the effective separation of ionic and polar metabolites, including organic acids, nucleotides, pharmaceuticals, vitamins, polyamines, sugar phosphates, and certain lipids and amino acids on C18 columns.

The combination of liquid chromatography and mass spectrometry is presently the most extensively employed approach for characterizing metabolic phenotypes through both untargeted and targeted analyses [69, 100]. While chromatographic separation of compounds before mass spectrometry injection markedly enhances the detection specificity for low abundance metabolites, the instrument running time for adequate chromatographic separation of complex biological samples constrains the number of samples that can be analyzed within a limited timeframe. On the other hand, to enhance sample throughput and accommodate

larger cohort studies, efforts are focused on developing methods for rapid profiling and quantification of metabolites derived from cell and tissue samples. Hence, the quickest and the lowest sample biomass required methods reported involve direct infusion MS, achieving sample duty cycles in just a few seconds with only few microliters required [101]. In this approach, no chromatographic separation is performed, which means that less abundant metabolites species can be overlooked [64]. Nevertheless, the development of the technology and data processing can reduce the drawbacks for the data acquisition, spectral annotation, and data interpretation [102].

In recent years, stable isotope tracing has emerged as a powerful tool for in-depth analyses of metabolic fluxes and pathways, offering precise atom-level tracking within complex metabolic networks [103]. Typically, this involves feeding a biological system with nutrients enriched with rare isotopes, such as ^{13}C, and utilizing mass spectrometry to detect isotopologues — molecules differentiated by isotopic content or isotopomers – isobaric molecules with different position of stable isotopes. The enrichment kinetics of stable isotopes in metabolites allow for the inference of biochemical pathway activities based on their distribution patterns and emergence rates post-tracer administration [104].

Typical metabolomics workflows include a series of structured steps: experimental design, collection and preparation of the biological samples, data collection and analysis, and ultimately the interpretation of the biological implication [105] (Figure 4). One of the most crucial steps within this framework throughout the extraction phase is a rapid quenching to arrest enzymatic activities, ensuring the stability of the metabolites' concentrations and their ratios within the sample, and in comparison, to controls [92]. This approach is essential for the accurate capture of a metabolic snapshot of the biological system, reflecting its state under specific conditions at the chosen time point, thereby providing a precise representation from a metabolomics perspective [95].

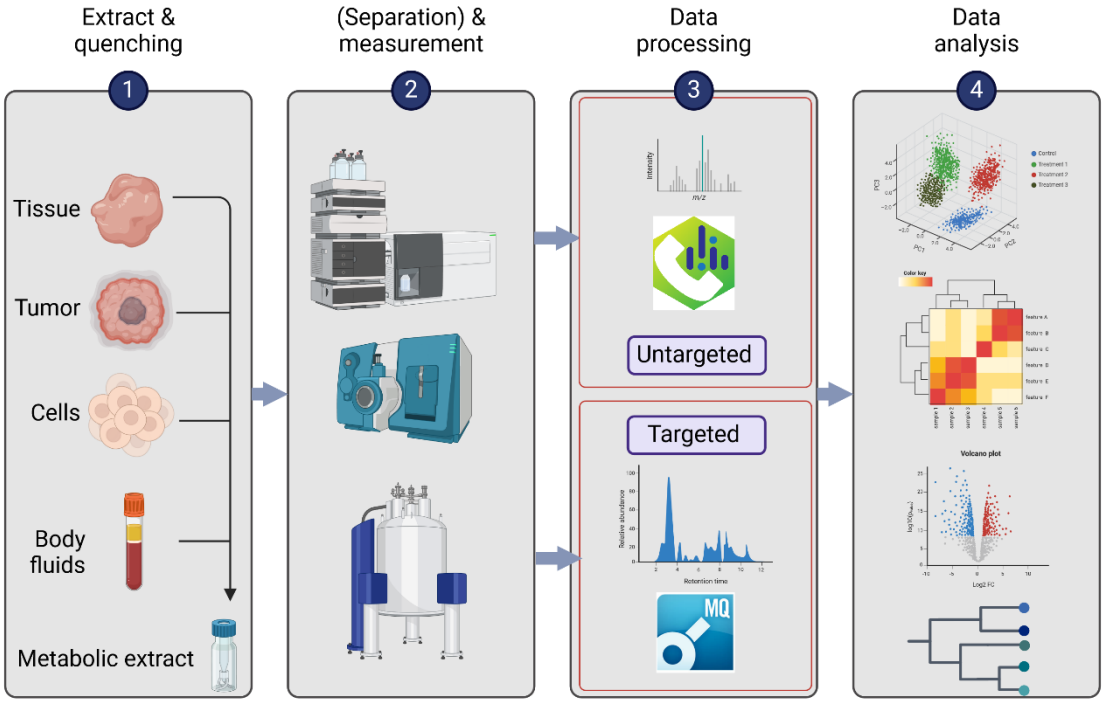


Figure 4: General workflow of metabolomics experiments. Starting with the extraction of the biological material (in vitro cells, tumor tissue, body fluids, etc.), and the detection of the metabolites with or without chromatographic separation (LC or GC). Followed by the data acquisition, depending on the type of method and amount of data produced. In the case of targeted metabolomics, a simple peak picking software is needed. However, in case of an untargeted approach, additional metabolite identification steps are needed based on metabolite and spectral libraries. Finally, a data analysis is required for data evaluation and interpretation. This graph was created with Biorender.com

Furthermore, continuous endeavors were dedicated to expand the coverage and enhance the sensitivity of targeted metabolomics. For instance, Balcke et al. [106, 107] established a LC-MS-based targeted metabolomics platform characterized by extensive coverage of metabolites of the central carbon and energy metabolism (CCEM), covering multiple vital pathways within the CCEM such as the glycolysis, TCA cycle, PPP, and the *de novo* biosynthesis of nucleotides, isoprenoids, lipids, and amino acids (Appendix table 1). This innovative platform allowed for the concurrent tracking of over 188 functional metabolites in a single analysis. By using this highly effective platform, multiple instances of metabolic reprogramming in cancer cells can be revealed.

1.4 Application of metabolomics in cancer research

1.4.1 Insights into the mechanisms of cancer development

Several diseases, including metabolic syndrome, cardiovascular diseases, and cancer are characterized by metabolic dysregulation during either initiation or developmental stages [105].

Notably, in the case of cancer, recent advances in metabolomics approaches have significantly enhanced our knowledge regarding the processes of tumor initiation, progression, metastasis and drug resistance [108-112]. A prominent example is the discovery of two subtypes of small cell lung cancer (SCLC) by Huang and colleagues [111]. Using a targeted LC-MS/MS approach, they detected across the two SCLC variants differences in the levels of purine nucleotides, such as inosine monophosphate (IMP), guanosine monophosphate (GMP), xanthosine monophosphate (XMP) and adenosine monophosphate (AMP). Accordingly, they determined the rates of nucleotide synthesis in both the achaete-scute homolog-1 (ASCL1) transcription factor subtypes, ASCL1-high and ASCL1-low. These different metabolic profiles were found to be under the control of inosine 5'-monophosphate dehydrogenase (IMPDH), hence being responsible for the different rates of cell proliferation between the two cell subtypes. Exemplified by this study, metabolomics not only increases the understanding of the pathways that carve distinctive metabolism in some cancers, but also helps suggest druggable metabolic targets for better selectivity and targeting. Therefore, an inhibition of IMPDH resulted in the selective reduction of ASCL1-low cells' growth, indicating that IMPDH holds promise as a therapeutic target for drug intervention [111].

Aberrant energy metabolism is a distinctive feature of cancer and can be utilized for the diagnosis and treatment as well as for overcoming drug resistances [22]. For instance, Patel et al. identified unique metabolic signatures in chronic myeloid leukemia (CML) cells, by using LC-MS-based metabolomics, distinguishing between those resistant to tyrosine kinase inhibitors (TKIs) and those sensitive to them [113]. The resistant CML cells exhibited enhanced glycolysis, whereas the sensitive cells showed enrichment in TCA cycle intermediates. This implies a metabolic shift from reliance on mitochondrial metabolism to glycolysis in TKI-resistant leukemic stem cells (LSCs). Indeed, targeting the rate-limiting enzyme in glycolysis, pyruvate kinase M2 (PKM2), led to the specific elimination of TKI-resistant LSCs, indicating that PKM2 could be a viable metabolic target in these persistent LSCs [113].

Another advanced application of metabolomics is the utilization in the clinic. Precision medicine aims to employ improved diagnostic testing to customize the individual's medical treatment based on their individual metabolic profiles. This aspect is exemplified by Xiao et al.'s study that categorized triple-negative breast cancer (TNBC) patients into three distinct subgroups based on LC-MS-based profiles assessed by non-targeted metabolomics. These included a subgroup with heightened ceramides and fatty acids, another with enhanced oxidation reactions and carbohydrate metabolism such as OXPHOS, purine and glucose metabolism, and a third subgroup showing minimal metabolic disturbances relative to normal tissue. The categorization of TNBC based on metabolomic profiles provides an opportunity to

customize individualized treatment strategies and could play a vital role in advancing precision medicine [114].

1.4.2 Insights into drug discovery, repurpose and development

Due to the significant expenses, lengthy timelines, and substantial failure rates associated with new drug discovery and development, the repurposing of existing drugs for the treatment of both common and rare diseases is gaining momentum. This approach involves the repurposing of compounds that have been de-risked, potentially resulting in reduced development costs and timelines [115]. Metabolomics, particularly in cancer research, contributes significantly to drug repurposing [8]. Following the previously mentioned work of Huang et al., mizoribine was repurposed, a medication initially approved as an immunosuppressant for organ transplantation and autoimmune disorders for the purpose of inhibiting IMPDH and consequently hampering the growth of ASCL1-low tumors [111, 116, 117]. Metformin (METF) has been extensively proposed as a potential anticancer agent and was investigated across various cancer types [118-121]. The metabolic patterns of tumor cells isolated from patients concurrently using METF were investigated. In their study, Liu et al. observed that METF accumulated within the tumors and affected numerous metabolic pathways, all related to mitochondrial functions. Interestingly, the metabolite profiles of tumor samples extracted from METF-treated patients closely resembled the profiles seen in a responsive animal tumor [122].

Furthermore, metabolomics can play a crucial role in the anticancer drug development process by analyzing the impact of drug candidates on cellular metabolism, which aids in refining these candidates to enhance their effectiveness [123]. Such analyses not only help in identifying specific inhibitors but also in predicting drug responses and potential resistance mechanisms arising from metabolic competition [124]. Moreover, metabolomics can be implemented in assessing drug toxicity, as exemplified by studies on sunitinib's effects on hepatic metabolism and bile acid homeostasis, providing insights that could mitigate adverse effects in cancer therapies [125].

Combining drugs can lead to synergistic effects, enhancing their effectiveness, reducing toxicity, and attenuating the development of drug resistance. Metabolomics has the potential to generate novel insights into the concept of combinatorial anticancer therapy. Lu et al. created a metabolomics-based drug screening platform based on stable isotope tracer direct-infusion mass spectrometry (SIT-DIMS) to assess the synergistic potential of drugs. By applying this platform, a new combination of drugs, namely the glutaminase inhibitor CB-839 and docetaxel, was discovered for treating prostate cancer, showcasing drug combination with

synergistic effects achieved with the help of metabolomics [126]. In another study, employing the same metabolomics-based phenotypic screening platform, Lu et al. successfully identified synergistic drug combinations for the treatment of AML, a highly aggressive hematologic malignancy associated with high mortality and relapse rates. They uncovered a novel synergy between the inhibitor IACS-010759, targeting complex I in the OXPHOS, and the FLT3 receptor tyrosine kinase inhibitor AC220 (quizartinib) specifically acting against AML. This discovery was subsequently validated and proven to significantly impede the growth of AML cells by causing a substantial disruption of their cell metabolism [127].

1.4.3 Metabolomics as a tool in MoA determination for anticancer agents

The emerging application of metabolomics in cancer research has provided the opportunity to capture the metabolic state of cancer cells as real-time image [128]. This gave access to more information supporting the researchers in understanding the metabolic processes that occur during the growth, proliferation, and death of cancer cells as well as upon treatment with any drug candidate. However, the application of metabolomics in drug discovery to predict and investigate the mode of cancer cell killing is still at its beginnings. Multiple studies have been conducted to predict the MoA in cancer drug discovery using proteomics and transcriptomics, however, metabolomics can be an even better tool in straightening the way of MoA investigation for new drug candidates [129]. The usefulness of metabolomics in MoA identification can be underlined by the two main drug discovery approaches [130]. In target-based screening, where the target is known, metabolomics helps investigate and confirm the anticipated changes in metabolite levels. Additionally, metabolomics can assist in MoA identification also in phenotypic screening, where the target of a drug candidate is often unknown. Metabolomics can be utilized to establish a model prediction technique based on similarity between the metabolic signature or “metabotype” of the investigated drug candidate and a reference compound with known molecular target in cancer cells. For that purpose, a comprehensive metabolomics “training study” needs to be prepared using a set of compounds with known MoAs (“reference compounds”), which can then be used in a “prediction study” for the elucidation of the MoA of newly discovered drug candidates. Hence, metabolomics can help delineate MoAs by identifying metabolic changes in cells upon drug exposure, which are reflected in complex metabotypes. Thereby, the prediction of the MoA for uncharacterized compounds is based on the hypothesis that drugs with similar targets will have similar effects on the metabolome [131]. Machine learning (ML) can then be employed to predict MoA(s) of a compound by using metabotype data of reference drugs with known MoA as training set (please refer to section 1.6 for further information about ML), as has been recently demonstrated for antimicrobial compounds screening study [131-133]. Demonstrating high-

throughput capabilities, CRISPRi and metabolomics were combined for the functional annotation of drug libraries in *E.coli* where drug-induced metabolic changes of 1,342 compounds were compared to metabotypes after knockout of 352 genes regulating essential biological processes [134]. When transferring this combined approach as proof of concept into a smaller side-project, A549 lung cancer cells were exposed to 14 anti-cancer drugs. Strong dUMP accumulation was observed only for those drugs targeting thymidylate synthase (TYMS), which was accompanied by a strong and selective metabolic similarity between genetically and chemically induced repression of TYMS. This work confirms the applicability of the computational and experimental frameworks introduced for the functional annotation of compound libraries to a wide range of biological systems, from bacteria to human cells. A significant finding of this study is that transcriptional downregulation of processes, which demonstrate target engagement, can yield phenotypes comparable to those achieved through chemoinhibition of the same targets.

1.5 The application of other OMICS techniques in MoA elucidation

Other OMICS techniques also play a significant role in drug discovery research, particularly in elucidating MoA and identifying therapeutic targets. Recent advances in MS-based proteomics have facilitated larger-scale investigations into proteomic alterations in cancer cells in response to drug treatments [135]. These investigations have supported the identification of potential biomarkers and drug targets [136]. Moreover, comprehensive proteomic studies investigating the MoAs of various drugs in the colon [137] and lung [138] cancer cell line have recently been published. Mitchell et al. introduced an advanced profiling methodology leveraging quantitative proteomics to unravel the MoAs of small molecules, thereby revealing the potential of new molecular targets for drug repurposing [137]. In their study, Mitchell et al. used proteomics to identify 914 proteins affected by 875 small molecules, covering nearly half of the targetable proteome. Interestingly, while some compounds had broad effects across the proteome, others acted as target specific, which may indicate better therapeutic prospects. Highlighting the utility of this approach in drug discovery, MoA elucidation and drug repurposing, the authors used a proteome fingerprint database to create networks of compounds and proteins. Compounds that share the same MoAs and proteins with similar functions were organized into communities, facilitating the characterization of the protein perturbation induced by compounds, uncovering new targets for specific compounds, and identifying off-target effects. This approach promises to enhance efficiency and safety of the drug discovery process [137].

Beyond metabolomics and proteomics, transcriptomics serves as a valuable tool in cancer research and drug discovery. Over the past decade, gene expression profiling has emerged

as a leading multidimensional assay technology, significantly advancing our understanding of cellular responses and molecular mechanisms. This technique measures the biological effects of a compound at the transcriptional level, providing a comprehensive representation of the biological state of a living system [139, 140]. Transcriptional changes following compound administration can now be assessed in high throughput, enabling the screening of numerous compounds across various cell lines at low costs. The use of transcriptomic data for characterizing the biological effects of small molecules gained popularity with the initiation of the Connectivity Map (CMap) project [141, 142]. CMap was created as the first reference sets compiling gene-expression profiles from human cells treated with bioactive molecules and utilizing pattern-matching software to analyze this data. Offering a cost-effective, high-throughput methodology, CMap has collected so far 3.02 million profiles from over 80,000 treatments. These profiles facilitate the identification of connections among small molecules sharing similar MoA and chemicals, physiological processes, diseases, and drugs [143].

1.6 Machine learning application in metabolomics and drug discovery

ML has gained increasing prominence as a valuable tool for the classification, regression, clustering and prediction of highly complex metabolomics data [144]. Hence, ML techniques serve as effective instruments for detecting patterns and constructing models in a large data environment as being characteristic for metabolomics [145], and this can be done primarily by using algorithms which rely on the capability to construct mathematical models using a set of sample data [146]. In standard practice, datasets used in the development of ML models are typically divided into two parts: the first part, called the training subset, typically comprises around 70% of the available data and is used by the ML algorithm to construct and predict the model. The second part, typically comprising around 30% of the data, called the testing subset (prediction set), is used impartially to evaluate the final model derived from the training phase in an unbiased manner [144].

To optimize model selection and hyper-parameters, a cross-validation phase is crucial, typically using about 20% of the training set as an auxiliary validation set. This step is essential in ML to ensure that sufficient data is available for the initial learning, where the algorithm learns from the training set through a function influenced by weights and biases. After training, the model's accuracy is tested on new data to verify its prediction accuracy. The model's predictive capability is assessed by how well it assigns inputs to statistically predefined classes. Another way to estimate the predictive performance of the ML model is the leave-one-out cross validation method. In this method, each data point in the dataset is sequentially left out to form the validation set, while the remaining data is used for training. This process repeats for each data point, allowing the model's accuracy to be assessed based on how well

it predicts each left-out observation. This technique ensures that each data point is used for both training and testing, offering a comprehensive evaluation of model performance. The leave-one-out cross-validation method is used in cases where dataset are scarce, to maximize the number of instances for training a model [147].

ML can be mainly categorized into two types: supervised and unsupervised learning. Supervised learning requires labeled data (training data) to enable the model to recognize and differentiate known patterns according to the labeling. The model then applies this knowledge to identify these patterns in unlabeled data (test data). Unsupervised learning, on the other hand, works without any need for labeled data and aims to uncover hidden underlying patterns in large data sets in an unbiased way. Typical supervised learning tools include neural networks, support vector machines, decision trees, random forests (RF) and hidden Markov models. Unsupervised methods include clustering algorithms such as hierarchical agglomerative clustering, k-means clustering, principal component analysis (PCA), self-organizing maps and non-negative matrix factorization [69].

Diving deeper, the decision tree (DT) methodology exemplifies supervised learning by employing a hierarchical structure to classify data, particularly demonstrating its capacity to handle non-linearly independent datasets. Its core objective is to develop a model that predicts a target variable's value based on inputs, thus simplifying complex decision-making processes [148]. Similarly, RF utilizes an ensemble of decision trees for tasks ranging from classification to regression, selecting the most common outcome across the diverse trees for classification and averaging predictions for regression. This ensemble approach effectively counters the overfitting pitfalls that single decision trees might fall into. Additionally, RF performs inherent feature selection, identifying and utilizing the most informative features for prediction, which enhances the model's interpretability and efficiency [149]. Thus, RF have gained attraction as a tool in diverse metabolomics studies especially for biomarker discovery [150-153]. Moreover, Lasso (least absolute shrinkage and selection operator) stands out by integrating variable selection and regularization. Thereby, this technique is not only boosting the interpretability and accuracy of models but is also tackling the issue of sparse coefficients in linear models by limiting the absolute sum of the coefficients to include only the most significant ones, effectively removing non-critical predictors [154]. These strategies exemplify the sophisticated and diverse approaches that machine learning offers for data classification and prediction, each granted with unique advantages and applicability [150, 155, 156].

In order to assess the performance of the prediction model used, another metric can be employed. The receiver operating characteristic (ROC) curve and the area under curve (AUC)

metrics evaluate the model's ability to differentiate classes, expressing performance in a single number. ROC and AUC indicate how well the model classifies outcomes, balancing sensitivity and specificity. An AUC of 1.0 represents a perfect model that perfectly distinguishes between two classes, while an AUC of 0.5 indicates a model that performs no better than random chance [157].

Recently, ML has been increasingly adopted in the field of metabolomics. For example, a study focused on the treatment of non-small cell lung cancer (NSCLC) with the drug pemetrexed, aiming to identify potential biomarkers and to establish models for the evaluation of drug efficacy. The researchers examined 323 subjects using metabolomics and ML. They found that certain metabolic markers, such as tryptophan, kynurenine, and xanthurenic acid, were highly effective in diagnosing NSCLC and evaluating the response to pemetrexed. The RF model achieved superior results with an AUC of 0.981 for diagnosing NSCLC and 0.954 for assessing pemetrexed efficacy [158]. Thus, this research promotes the use of metabolomics and ML as a promising approach with high potential to enable timely and accurate disease diagnosis, enhancing precision medicine practices, and supplementing current diagnostic tools.

Another study was done recently on ovarian cancer (OC) with the aim to utilize ML and metabolomics to develop a model for OC classification based on metabolic biomarkers. The authors analyzed existing metabolomics data related to OC and performed pathway and metabolite-set enrichment analyses. Then, they combined quantitative molecular descriptors with various ML algorithms to diagnose OC based on these metabolic data. The study identified five metabolic pathways linked to OC and created several classification models, with the most accurate achieving an 85.29% accuracy rate. This research suggests the potential for developing OC diagnostic models using metabolomics data and sheds light on OC-related biological pathways [159].

In their antimicrobial research, Zampieri et al. conducted a screening to identify antimicrobial drug candidates with potentially novel MoA to combat bacterial resistance against traditional antibiotics [131]. A reference database was constructed based on the metabotypes of 62 compounds with known MoAs in *Mycobacterium smegmatis*. The metabotypes were obtained using a high-throughput metabolomics profiling approach using the direct infusion method with high-resolution MS to annotate a panel of 446 metabolites. Utilizing a feature selection algorithm and statistical analysis, key discriminative metabolites were identified for MoA classification and for comparing metabolic profiles post-treatment. These metabolites were used as features in a random forest classifier, which was cross-validated to achieve an

accuracy of AUC = 0.84 on the ROC curve and could also distinguish between subclasses within the same MoA. Applied to 212 new anti-mycobacterial compounds from the GlaxoSmithKline Drug Discovery Initiative, the classifier indicated that over 70% of these compounds shared metabolomic profiles with known MoAs, suggesting limited novelty in their mechanisms. Only 8% displayed potentially novel MoAs.

2 Objectives and hypotheses

At the Leibniz Institute of Plant Biochemistry (IPB) longstanding efforts have been dedicated to extract, purify, identify, and characterize natural products from plant and fungal sources. Additionally, our team of organic chemists has synthesized numerous derivatives of those natural products, assembling a library of over 25,000 compounds and approximately 5,000 extracts. These initiatives aim to deepen our understanding of natural products, their biological effects and potential applications, amongst others particularly with respect to anticancer properties. The “Anticancer Screening and Targeting” research group at the department of Bioorganic Chemistry focuses on advancing cancer research and therapy development.

Our approach includes initial *in vitro* screening to identify promising candidates, which are then further characterized in detail to elucidate their molecular targets and to enlighten their modes of action (MoA). For that purpose, supporting the biochemical and cell-based assay efforts, a rapid and comprehensive methodology such as metabolomics and comparative metabotyping is highly desirable for generating and supporting hypotheses.

We propose that metabolomics can be an ideal tool for predicting the MoA of bioactive compounds. We hypothesize that compounds with the same MoA give a specific metabolic fingerprint (metabotype) and that a library of metabotypes of compounds with known MoAs can be used as a reference library for comparison with the metabotypes of novel anticancer agents with still unknown MoA. Accordingly, our hypothesis suggests that based on the similarity of metabotypes, each compound can then be assigned to the corresponding MoA. To obtain the metabotypes, we hypothesize that targeted LC-MS-based metabolomics measuring a set of metabolites covering the CCEM should be suitable for this methodology.

One of the primary objectives of this study is to establish and refine a workflow that is reproducible, straightforward and suitable for mammalian cell cultures. This includes optimizing protocols for metabolic quenching, extraction, measurement, and data normalization designed for targeted metabolomics. Such methodology can produce reliable metabotypes, which in turn provide a comprehensive understanding of the CCEM.

Our objectives include the assembling of a collection of 30 – 40 reference compounds, mostly approved drugs targeting multiple pathways known to impact the CCEM in cancer cells. Using the PC-3 prostate cancer cell line, we aim to study the metabolic profiles induced by these reference compounds by determining their IC_{50} values and conducting metabolomics experiments with experimental conditions and parameters optimized through preliminary studies. Finally, we measure the relative abundance of detectable CCEM metabolites.

Following the data generation, we intend to establish a collaboration to implement a machine learning (ML) approach in order to boost the predictive power of the methodology. Additionally, we aim to test the applicability of the methodology by investigating the MoA of novel cancer agents, for example, the *Nerium oleander* extract Breastin and the synthetic compound asiatic acid homopiperazinyl rhodamine (AAHR), using this approach.

Furthermore, we plan to test the transferability of this methodology to other cancer cell types with different genetic backgrounds. This will be done by comparing the metabotypes of these cell lines to the PC-3 after treatment with similar compounds and under the same condition.

Finally, this method will be integrated into the drug discovery process at IPB and made available for external collaborations, serving as a public research platform for MoA discovery of promising anticancer agents.

3 Materials and Methods

3.1 Cell culture

3.1.1 Cell culture maintenance

Cell handling and assay techniques were described in Dube et al. [160]. The prostate cancer cell line PC-3, the estrogen receptor positive (ER^+) and human epidermal growth factor receptor 2 negative ($HER2^-$) breast cancer cell line MCF-7 (both purchased from ATCC, Manassas, VA, USA) and the Ewing's sarcoma cell line MHH-ES1 (purchased from DSMZ, Braunschweig, Germany) were maintained in RPMI1640 basal cell culture medium supplemented with 10% (v/v) fetal calf serum (FCS), 1% (v/v) L-glutamine and 1% (v/v) penicillin/streptomycin. Cells were pre-cultured in T-75 cell culture flasks in a humidified atmosphere with 5% CO_2 at 37 °C. The cells were routinely sub-cultured when reaching 85% confluency and used with a maximum of six passages per batch. Adherent cells were washed with phosphate buffered saline (PBS) and detached by using trypsin/ethylenediaminetetraacetic acid (EDTA) (0.05% in PBS).

3.1.2 *In vitro* cell viability assays for IC_{50} determination

Cells were seeded in 96-well plates with a density of 6,000 cells/100 µl medium per well and were left to attach overnight. Cells were exposed to the compounds of interest, which were diluted in dimethyl sulfoxide (DMSO), or cell culture medium based on their solubility. The treatment was administered in cell culture medium with maximum of 0.5% DMSO (v/v), using dilution series of eight concentrations. The tested concentration ranges were compound-specific and narrowed down to the range around IC_{50} specifically determined for each compound in preliminary cell viability assays. In parallel, cells were treated with 0.5% DMSO (negative control) and 100 µM digitonin (positive control, for data normalization, set equal to 0% cell viability). Each cell viability assay was performed in two biological replicates each with technical quatruplicates. After 48 h incubation time, cells were washed once with PBS before incubating with MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) solution (0.5 mg/ml) for 1 h at 37 °C and 5% CO_2. Afterwards, MTT solution was discarded and DMSO was added to dissolve the formed formazan dye whose absorbance was measured at 570 nm in addition to a reference/background signal at 670 nm by using a SpectraMax M5 plate reader (Molecular Devices, San Jose, CA, USA). To address the potential interference of treatment with the MTT assay mechanism, which is based on changes in cellular metabolic activity, an additional viability assay was conducted. Crystal violet-based cell viability assay (CV) utilizes a dye that binds directly to DNA rather than being dependent on metabolic activity, thereby providing orthogonal measurement of cell viability under treatment conditions. Thus, CV was conducted. After a single PBS wash, the cells were fixed with 4% (v/v) paraformaldehyde (PFA) for 20 min at room temperature (RT). Then, the cells were dried for 10 min and stained

with 10% (w/v) crystal violet for another 20 min. The excess stain was removed by aqua bidest. wash and the cells were dried overnight at RT. Subsequently, acetic acid (33% v/v in aqua bidest.) was added to dissolve the stain, and the absorbance was measured at 570 nm and 670 nm as described before. The cell viability was calculated as a percentage in relation to untreated control cells. For data analyses and IC_{50} calculations, SigmaPlot 14.0 and Microsoft Excel 2013 were used. Mean values were calculated by using the four-parametric logistic function.

3.2 Cancer cell metabolomics

3.2.1 Cell treatment for metabolomics experiments

PC-3 cells were seeded as hexuplicates in T-25 cell culture flasks with a density of 0.5 million cells/5 ml. After cell attachment overnight, cells were treated with the compound of interest by applying IC_{50} concentrations as determined in the preliminary viability assays. Treatments and vehicle controls had the same final DMSO concentration. After 48 h under standard growth conditions (37 °C, 5% CO_2, in a humidified atmosphere) control flasks without test compound reached 85% confluence. Then the cell culture medium of the T-25 flask was discarded and replaced by pre-warmed PBS (37 °C) for a washing step that removed detached cells. Within a few seconds, the washing solution was discarded, and the remaining cells were quenched with 1.5 ml cold acidic ethanolic solution (10% HCl of in ethanol (v/v) pH 1.4, -80 °C). The T-25 flasks were sealed with watertight lids and then immersed in an ultrasound water bath, with the flasks weighted with lead weights and centered in the ultrasonic field. Cells were quantitatively detached by 5 min ultra-sonication (Bandelin Sonorex RK 106, 480 Watt, Berlin, Germany). Additionally, packed dry ice was placed in the water bath to maintain a low temperature close to 4 °C. Afterwards, the cell suspension was transferred from each T-25 flask into pre-chilled 2 ml Eppendorf tubes kept on dry ice. These where subsequently placed in a sample concentrator at 4 °C (TurboVap® LV, Biotage, Uppsala, Sweden), where the volume was reduced to 50 µl in a N_2 stream. After volume reduction, two consecutive centrifugation steps were performed at 10,600 rcf for 5 min at 1 °C, with another transfer of the supernatant to a fresh Eppendorf tube after the first centrifugation. Finally, the supernatant after the second centrifugation was transferred to a LC-MS vial with insert, and samples were frozen and stored at -80 °C until LC-MS measurement.

3.2.2 Metabolomics Assays

The separation of hydrophilic metabolites was performed by ion-pairing chromatography on a Nucleoshell RP18 column (2.1 x 150 mm, particle size 2.1 µm, Macherey & Nagel GmbH, Düren, Germany) using a Waters ACQUITY UPLC System equipped with an ACQUITY Binary

Solvent Manager and ACQUITY Sample Manager (5 µL injection volume; Waters GmbH, Eschborn, Germany). Eluents A and B were aqueous 10 mmol/L tributyl amine (adjusted to pH 6.2 with glacial acetic acid) and acetonitrile, respectively. Elution was performed isocratically for 2 min with 2% (v/v) eluent B in A, from 2 to 18 min with a linear gradient up to 36% (v/v) B, and from 18-21 min up to 95% (v/v) B, and isocratically from 21 min to 22.5 min with 95% (v/v) B, from 22.51 to 26 min again down to 2% (v/v) B in A. The flow rate was set to 400 µl/min, and the column temperature was maintained at 40 °C.

Mass spectrometric analyses of small molecules were performed by targeted MS/MS via MRM by using a QTRAP 6500 (AB Sciex GmbH, Darmstadt, Germany) operating in negative ionization mode and controlled by Analyst 1.7.1 software (AB Sciex GmbH, Darmstadt, Germany) (Appendix table 1 detailed the MRM metabolite list). The source operation parameters were the following: ion spray voltage, -4500 V; nebulizing gas, 60 psi; source temperature, 450 °C; drying gas, 70 psi; curtain gas, 35 psi.

Ideally, the sequence for sample injection should be randomized to reduce the risk of analytical errors. Additionally, it was advisable to include QC standards and pure water injections at specific intervals, such as every three sample injections, to monitor any shifts in retention time (r.t.) and assess sample carry-over throughout the batch. The performance check results were verified by ensuring that no signal was detected in the blank sample. In some cases, signals close to the background were considered normal for certain compounds due to their abundance or possible presence as contaminants, such as succinate. Subsequently, the QC mix results were examined, and it should be confirmed that the tested compounds elute at the expected r.t. with the anticipated intensities and signal-to-noise (S/N) ratio. Following the final sample injection, it was recommended to inject a blank and a QC mix for performance verification. To maintain column performance, the analytical sequence incorporated column washing steps using 10 µl injection volume of both methanol and magic mix (equal volume of acetonitrile (ACN), methanol, 2-propanol and water) in an alternating way for a total of four runs with 6 min gradual method after every 12 injections. In this method, a solvent mixture consisting of methanol, ACN, and 2-propanol in a 2:1:1 ratio serves as solvent A2, while solvent B2 contains 0.3 ammonium formate. By alternating between organic and aqueous solvents within the system, we ensure the durability and longevity of the LC pump system, reducing constant exposure to the same solvent.

Peak integration was performed using the MultiQuant software version 3.0.3 (Sciex, Toronto, Canada). In order to account for different cell numbers in the treated samples and vehicle controls, individual CCEM peak areas were normalized for each sample to the total peak area.

Finally, each normalized metabolite intensity in an individual sample was divided by the mean of the likewise normalized signal area of all vehicle control samples of the respective experimental set. All area ratio data were logarithmized to the basis of 2. The statistical analyses and graphical presentations were performed by MetaboAnalyst 5.0 using the log2-normalized data and range scaling [161].

3.3 Lipidomics Assay

PC-3 cells were quenched with 1.5 ml cold MeOH (-80 °C) and dispersed by ultrasonication as described above. Cell debris and solution were then transferred together into a new tube and were dried in a nitrogen stream. After addition of 1 steel bead (3 mm), 3 steel beads (1 mm) and 200 mg of glass beads (0.75-1 mm), 700 µl methyl tert-butyl ether (MTBE) and 200 µl of aqua bidest were added, and cryo bead milling was performed for 3 × 20 s (4.0 m s-1; MP24, Biomedicals Inc., California, USA). Following phase separation by centrifugation (2 min, 10,000 × g, 4 °C) the upper phase was collected and stored on ice. The lower phase was re-extracted with another 700 µl MTBE and, after centrifugation, the organic layers were combined and dried in a nitrogen stream. For analysis, the dry residue was dissolved in 500 µl MeCN/2-propanol/water (60/35/5 v/v/v).

Lipid separation on UPLC was performed according to Witting et al. [162] with the following modifications. Eluent A consisted of 60% MeCN and 40% water, and eluent B consisted of 90% 2-propanol and 10% MeCN, both with 10 mM ammonium formiate and 0.1% formic acid. Separation was carried out on a Nucleoshell RP18 (Macherey & Nagel, Düren, Germany) with the dimensions 2 mm × 150 mm × 2.7 µm. The gradient on a Waters Acquity was: 0-1.5 min 32% (v/v) B in A, 1.5- 18 min linear gradient to 98% (v/v) B, 18-20 min 98% (v/v) B, 20-24 min 32% (v/v) B in A, with 10 µL injection volume and 40 °C oven temperature.

Separated lipids were ionized by electrospray ionization and analyzed in positive ionization mode by HR QToF-MS/MS (Zeno7600; Sciex, Toronto, Canada) using these source parameters: ion source gas 1, 60 psi; ion source gas 2, 70 psi; curtain gas, 35 units; CAD gas, 7 units; temperature 450 °C, ion spray voltage floating, -5500 V. MS1 and CID-MS/MS spectra were scanned from 65 – 1500 Da using 100 ms accumulation time for MS1 and 20 ms each for a maximum of 40 data-dependent MS/MS experiments. For MS1 survey scans and MS/MS the declustering potential (DP) was 80 V with a spread of 50 V, whereas in MS1 the collision energy was fixed to 10 V. MS/MS scans were acquired with a collision energy of 35 V and a spread of 25 V. Mass accuracy was recalibrated after every five samples using internal standards provided for ESI by Sciex. For lipidomics data analysis, raw data were processed by MS-Dial version 4.9 using the lipidomics standard library embedded in the program [163].

Peak height tables with classified lipids were exported from MS-Dial, normalized to cell number and analyzed by using GraphPad Prism 9.1.5 software.

3.3.1 ^{13}C labeling lipidomics

PC-3 cells were seeded in T-25 cell culture flasks as described above. After 24 h of incubation, cell culture medium was exchanged against RPMI 1640 medium without glucose (Gibco, Darmstadt, Germany), supplemented with 10% heat-inactivated dialyzed FBS (Gibco, Darmstadt, Germany) and uniformly labeled D-glucose (U-$^{13}C6$, 99%) (Cambridge Isotope Laboratories, Inc., Massachusetts, USA) at a final concentration of 2 g L^{-1}. Simultaneously, samples were treated with BETA at the IC_{50} concentration (Table 1), whereas in the control flasks no BETA was added. Isotopolog analysis was performed using Lipidomics extraction 48 h post treatment, following the protocol described above. The samples were analyzed using UPLC-ToF in positive mode ESI with the chromatographic gradient as described. Isotopologs of the three most abundant PC (m/z;r.t.: 760.5836; 12.65 min, m/z;r.t.: 788.6158; 13.59 min, and m/z;r.t.: 732.5521; 11.78 min) were area integrated based on their theoretically expected [M+H] ions within a mass tolerance of 10 ppm. The data analysis was done using Sciex OS. The percentual ^{13}C incorporation was calculated as described in Balcke et al. [164] and visualized by GraphPad Prism 10.

3.3.2 CoQ assay

Coenzyme Q10 (CoQ) was analyzed by UPLC-QToF with positive mode ESI. The gradient was the same as described in the lipidomics section. CoQ was determined by [M+H] of 863.692 with a mass tolerance of 2 ppm. The retention time and the MS/MS spectrum were compared to an authentic standard (Merck). Peak area values of pentuplicate treatments were normalized to the cell number.

3.3.3 Cell counts

PC-3 cells of technical triplicates in each control/treatment group were detached by trypsination for 3 min in 0.05% trypsin/EDTA at 37 °C. After trypsination stop with FCS-containing culture medium, the resuspended cells were stained with Trypan blue, and cell counting was performed using a Neubauer counting chamber (Marienfeld, Lauda-Königshofen, Germany).

3.4 Localization of AAHR in PC-3 cells

PC-3 cells were seeded in a 10-well chamber slide with 1,000 cells per well and allowed to grow at standard growth conditions for 24 h. Cells were treated with 0.1 µM AAHR (test

compound) for 4 h followed by washing twice with PBS. Staining with 0.1 µM MitoTracker™ Green FM was done for 15 min. After washing with PBS twice, fresh medium was added. Fluorescence was recorded using a confocal microscope LSM900 (Carl Zeiss, Jena, Germany) with ex/em 561nm/576-700 nm for AAHR and ex/em 488 nm/490-540 nm for MitoTracker™ Green. Co-localization of AAHR and MitoTracker™ Green was analyzed using the Zen Blue image analysis software (Zeiss).

3.5 Seahorse analysis

To investigate the compound-induced alterations of the cellular energy metabolism, the Seahorse cell analysis technology platform from Agilent was utilized, applying the Seahorse XF Cell Mito Stress Test Kit (Agilent Technologies, Santa Clara, CA, USA). The analyses were performed on a Seahorse XF 96 - Extracellular Flux Analyzer according to instructions of the manufacturer. PC-3 cells were seeded in specific 96-well plates at day 0 using a density of 12,000 cells per well. Subsequently, starting at day 1, the cells were treated with 20 nM and 100 nM of AAHR for 24 h (a duration not yet causing apparent signs of cytotoxicity), or were left untreated, and finally were measured at day 2. Another sample of untreated cells was used to assess mitochondrial function after acute compound treatment. For this purpose, one of the standard chemicals at a time (sample) was replaced by AAHR.

The Seahorse XF Cell Mito Stress Test is based on the measurement of the oxygen consumption rate (OCR) and the extracellular acidification rate (ECAR) of viable cells during sequential manipulation with specific chemicals. We performed it to obtain information about mitochondrial function, and to quantify cellular bioenergy levels (ATP) upon compound treatment.

To further explore the precise mechanism of inhibition of mitochondrial respiration, the Mito Stress Test was modified to analyze the impact of acute AAHR treatment on OCR and ECAR. The standard assay comprises a consecutive treatment with oligomycin (complex V / ATP synthase inhibitor), carbonyl cyanide 4-(trifluoromethoxy)phenylhydrazone (FCCP, uncoupling agent) and rotenone/antimycin A (complex I / III inhibitors) leading to a reduction of OCR (increase of ECAR), strong increase of OCR to a maximum, and complete decrease of OCR to a minimum, respectively.

3.6 Docking Studies

Molecular docking was performed in YASARA software [165] version 20.12.24 by using the AutoDockLGA algorithm [166], and AMBER14 force field [167]. The cryo-EM structure of human CEPT1 complexed with CDP-choline (PDB ID: 8GYW) [168], and the AlphaFold [169] model for

CPT1 which was obtained from UniProt (accession number Q8WUD6) were utilized for molecular docking studies. The AlphaFill [170] server was used to assign the Mg^{2+} ion in CPT1, while the ion is already included in the CEPT1 cryo-EM structure. The binding site was defined according to the location of the ligand (CDP-choline) in the cryo-EM structure of CEPT1. Ligands and protein structures were energy-minimized, and a simulation box (10 Å) was defined around the Mg^{2+} ion. The validation of the docking protocol was performed by re-docking the ligand in the cryo-EM (CDP-choline) of CEPT1, using the same parameters. The results from 100 runs were clustered using a RMSD (Root Mean Square Deviation) of 2.5 Å, and MOE (Molecular Operating Environment) 2020 software [171] was employed to investigate the interactions between the ligands and the proteins.

3.7 Classification methods and prediction analysis

3.7.1 Normalization of CCEM profiles

CCEM profiles were normalized per metabolite by subtracting the mean across all drugs/replicates and dividing the result by the range between the maximum and minimum values across all replicates, i.e.

$$n_{m,r} = \frac{x_{m,r} - \bar{x}_m}{\max_s x_{m,s} - \min_s x_{m,s}}$$

where $x_{m,r}$: the original value of metabolite m in experiment (or replicate) r

$\bar{x}_m$: the mean value of metabolite m across all replicates/drugs

Calculated as: $x_{m,r}\ \bar{x}_m = \frac{1}{R}\sum_{r=1}^{R} x_{m,r}$

and R is the total number of replicates (or experiments) for metabolite m

$\max_s x_{m,s} - \min_s x_{m,s}$: These represent the maximum and minimum values of metabolite m across all replicates s. The difference between them is the range of metabolite m

The same coefficients were also used for normalizing the data in the prediction set to avoid re-scaling and subsequent re-training of models when processing prediction sets. For predictions in other cell models, a cell line-specific mean value was determined to account for different base profiles.

3.7.2 Correlation analysis of metabolites

For all pairs of metabolites, Pearson correlation was computed between their profiles across all drugs/replicates. Visualization as a heatmap and corresponding clustering of rows/columns was performed using the *pheatmap* R package (v. 1.0.12) [172].

3.7.3 Hierarchical clustering of CCEM profiles

Normalized CCEM profiles of all drugs/replicates were clustered using Euclidean distance and the *hclust* R function with method “ward.D”. The resulting cluster trees were visualized as dendrograms. Here, leaves are annotated with the MoA of the corresponding drug. Inner nodes of the cluster tree, under which all leaves belong to the same MoA, are annotated accordingly. To generate a more condensed view on the clustering result, each sub-tree in which all leaves belong to the same drug was aggregated into a single leaf. For all drugs, this results in a single leaf representing all replicates of that drug, except for PRFN/ORID. The latter has CCEM profiles so similar that replicates of both drugs appear intermixed. Heatmaps of selected profiles were generated using MetaboAnalyst 5.0. Peak integrated raw data were normalized to the sum of all signals as to compensate for the different cell numbers. Log(2)-fold changes between six replicate treatments per drug and six controls, both sampled after 48 h, were rank scaled.

3.7.4 Random forest – training

Random Forests were trained using the *randomForest* function of the corresponding R package (v. 4.7-1.1) [173] using 1,000 trees. Input data of the random forest method were the CCEM profiles of all replicates of all drugs in the training set. True class labels during training were the corresponding drugs, i.e., the random forest was trained to distinguish individual drugs, which in turn had been assigned a specific MoA.

3.7.5 Random forest – prediction

Each normalized CCEM profile of each replicate for each drug in the prediction set was considered independently, and probabilities for individual drugs were predicted using the trained random forest model using the prediction method with parameter “type=’prob’”. Returned probabilities were averaged across all replicate CCEM profiles of each drug, and then averaged across all drugs with a common MoA in the training set to yield the probability of a specific MoA. The two-level averaging ensures that all drugs have the same influence on the prediction result, irrespective of the number of replicate measurements.

3.7.6 Lasso-based approach – training & prediction

For the Lasso-based approach, the regression coefficients themselves were used to predict the MoA of a new drug. Here, the CCEM profile of an individual replicate of the drug in the prediction set were considered as the response variable and the CCEM profiles of all replicates of all drugs in the training set were considered as the input matrix of a Lasso regression using the *glmnet* function from the corresponding R package (v. 4.1-4) [174].

Regression coefficients were limited to non-negative values by setting the parameter for the lower boundary of regression coefficients to zero. The regularization parameter was set to a fixed value within the range of auto-computed values (lambda=0.025) and no intercept value was considered by setting “intercept=FALSE”. Regression coefficients were obtained using the coef method. Returned coefficients were averaged across all replicate CCEM profiles of each drug, and then averaged across all drugs with a common MoA to yield the prediction score of a specific MoA. For visualization purposes, prediction scores were normalized to a sum of one across all MoAs since each drug is represented by 5-6 replicates in the training set.

3.7.7 k-nearest neighbor – training & prediction

Predictions of the k-nearest neighbor classifier were obtained from the function *knn* of the class R package [175] using the CCEM profiles of each replicate of each drug in the training set and corresponding MoA as reference. The number of nearest neighbors was set to “k=5” to account for the number of replicates per drug.

3.7.8 Leave-one-out cross validation

For evaluating the prediction performance of each approach on the training set, a leave-one-out cross validation over the drugs in the training set was performed. To avoid data leakage from the training partition to the test partition within individual iterations of the cross validation, all replicates of a drug were excluded from the training partition and used as the test partition in a leave-one-out cross validation. Specifically, all replicate CCEM profiles of one drug with known MoA were considered as prediction partition, and the replicate CCEM profiles of all remaining drugs were used as the corresponding training partition. For each replicate of the drug in the prediction partition, predictions were obtained from the three alternative methods and those were compared against its true MoA. R source code of the evaluation procedure is provided as Appendix data file 1.

4 Results

4.1 Method establishment

To develop a novel and effective protocol for a cancer cell metabolomics experiment, it was essential to address gaps in existing methodologies and tailor the protocol to our specific experimental requirements. Unlike pre-existing protocols, that did not sufficiently meet the specific requirements of our study, this protocol was thoroughly developed to optimize the detection of low-abundance metabolites in PC-3 prostate cancer cells. In order to develop a functional protocol for a cancer cell metabolomics experiment, several aspects had to be considered. These include cell type, cell culture system, quenching procedure, metabolites extraction, storage conditions, normalization, and measurement method. In this study, the PC-3 cell line was used as the primary model. PC-3 cells grow as adherent monolayers *in vitro*, meaning they attach to the surface of the culture vessel. To ensure the detectability of low-abundance metabolites, it is essential to use an appropriate cell count before extraction. Therefore, T-25 flasks were preferred for culturing the PC-3 cells for the metabolomics experiment instead of cell culture plates, which are commonly used for cell-based assays. This approach enhances the ability to detect low-abundance metabolites in treated samples, where 50% of the cells are expected to survive compared to the untreated control, as the IC_{50} concentrations were used for the treatments. Optimized in preliminary trials, seeding 0.5 million cells results in 85% confluency 72 h after seeding in untreated T-25 flasks. Full confluence, which halts further cell proliferation in the culture, should be avoided as it can alter cancer metabolism and compromise data quality.

At the onset of the metabolomics experiment, after discarding the culture media, cells were washed with PBS to remove dead cells, debris, media residues and treatment $agent_{(s)}$. A suitable quenching protocol for adherent cells must address rapid arrest of all cellular enzymatic activities, ease of handling, choice of organic solvent, temperature, and subsequent extraction steps. We selected a quenching procedure using 10% HCl in ethanol (v/v) (pH 1.4, -70 °C), followed by 5 min ultrasonication. This protocol was optimized to ensure that all enzymatic activities are arrested by the very low temperature, the protein denaturation in the acidified ethanolic solvent, and that the stability of the extracted metabolites (especially for the nucleotides and the CoA esters) is maintained at a final pH 2.4, as recently proven by Balcke & Saoud et al. (unpublished data, not shown). Ultrasonication for 4 minutes at low temperature of -70 °C detaches the cells, disrupts the cell membrane, and releases intracellular metabolites into the solvent. This step eliminates the need for additional trypsinization or scraping, which could slow down the process and increase variability between replicates. We have proven that 5 minutes of ultrasonication is sufficient for cell detachment, as illustrated in the Figure 5.

Figure 5: Detachment of PC-3 cells after 5 minutes of ultrasonication.

Determining the required number of sample replicates before initiating the sample preparation protocol was crucial. Preliminary optimization trials indicated that hexareplicates were optimal for capturing inter-sample differences, which were more likely to be influenced by growth conditions in the control and treated cells. When cells are removed from the controlled environment of the cell culture incubator, metabolic alterations may occur due to changes in temperature (lower than the standard growth condition at 37 °C) and CO_2 levels (lower than 5%) [176, 177], resulting in variations in metabolite levels between replicates. It is therefore critical to expedite sample handling during washing, quenching and extraction. Once the quenching and ultrasonication step is completed, the samples need to be concentrated at a minimal volume to increase the metabolites' concentration, especially in case of low abundant metabolites. Solvent evaporation using a nitrogen (N_2) stream allows the removal of the organic solvent from the samples, concentrating the metabolites into the remaining aqueous phase. During the concentration step, all samples should be subjected to a homogeneous flow of N_2 to ensure a uniform final volume in all samples. Additionally, to minimize metabolic interconversion, it is essential to maintain all samples at temperatures below 4 °C throughout the whole extraction procedure. After extraction, samples were stored at -80 °C until analysis. Before LC-MS analysis, two consecutive centrifugation steps were performed to remove cell debris, and the clear supernatant was collected and used for LC-MS analysis.

The current LC-MS/MS method allows the accurate identification of 120 molecules using the multiple reaction monitoring (MRM) acquisition mode. This method employs ion-pairing separation and a targeted metabolomics approach developed with standards for the desired metabolites. These metabolites are involved in various cellular metabolic processes, including glycolysis, the tricarboxylic acid (TCA) cycle, the pentose phosphate pathway (PPP), nucleotide biosynthesis, amino acid and lipid biosynthesis, and cofactor. Detailed information on the chromatography and mass spectrometry parameters are provided in section 3.2.2.

Following the measurement of the metabolites, subsequent data analysis was conducted. However, a normalization step is necessary due to both biological and technical reasons. During method development, various normalization techniques were explored, including normalization based on the total protein signal, DNA content and cell count. However, these normalization methods significantly increased the complexity and susceptibility to errors in the entire metabolomics procedure. The added complexity of the method, the uncertainties regarding the accuracy of protein-based normalization (especially after protein denaturation in the quenching step), and the need for a substantial number of parallel samples for cell count estimation raised significant concerns. Cell counting could not be conducted with the sample dedicated to be analyzed in LC-MS/MS before or after the ultrasonication step due to limitations such as the inability to introduce staining and the undesirability of trypsinization, since both could compromise the preservation of the metabolic pattern. Next, normalization based on the DNA content of the cells was also considered. A study by Muschet et al. proposed a fast normalization method in cell culture metabolomics based on DNA content using fluorescence-based DNA quantification [178]. The published procedure involves extraction of the cellular DNA from metabolomics samples and staining with Hoechst 33342 dye. To assess the applicability of this method, several trials were conducted using PC-3 cells with control and treatment samples. A calibration curve, illustrated in Figure 6, was established using DNA extracted from metabolomics samples of PC-3 cells seeded at serial dilutions, the staining was done using Hoechst dye and followed by measuring the fluorescence. This approach was employed to quantitatively assess the correlation between cell count and the efficiency of DNA extraction from metabolomics samples.

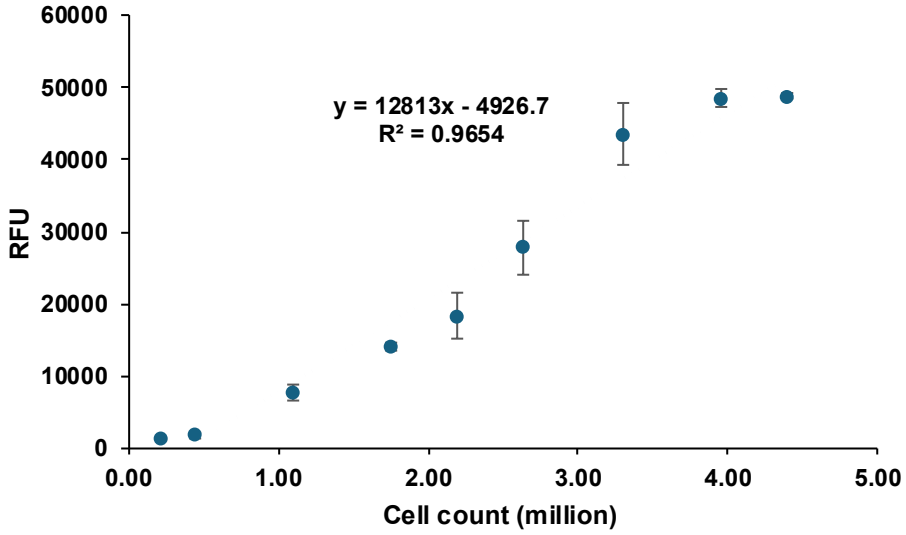


Figure 6: Calibration curve representing the correlation between the extracted DNA from PC-3 cells and the cell count. DNA was extracted as described by Muschet et al. stained with Hoechst 33342 dye and quantified fluorometrically [178].

However, several pitfalls were observed during the method application, including poor reproducibility within and between experiments, increased complexity of the metabolic profiling methodology, and incomplete DNA extraction. More specifically, when comparing the DNA content using the described method, and directly trypsinized cells with the same cell count, the recovery was inconsistent (data not shown). Consequently, there was a need for a simpler and more straightforward normalization procedure.

Hence, an integral normalization step was considered to ensure the homogeneity of the metabolites' intensities. For that purpose, the metabolites' intensities were normalized based on the sum of the total metabolites' peak areas in each sample. This approach assumes that the overall quantity of all metabolites within each sample remains relatively proportional to the cell count, regardless of whether the cell model was drug-treated or not. In other words, the total profile is assumed to be directly proportional to the total concentration of the sample [179]. The method normalizes intracellular metabolites levels by dividing their specific peak area by the total area of all signals in a particular sample. However, this normalization approach can be impacted by two major weaknesses: (1) the validity of the underlying assumption that total metabolite content is proportional to cell count, and (2) potential inaccuracy caused by very large (or small) signals in the samples. To assess the validity and accuracy of the normalization method, the total peak area was regressed against different cell numbers, which were independently counted. The test was conducted for untreated and antimycin A-treated PC-3 cells, in order to demonstrate that the assumption of linearity between cell number and total peak area is independent of the treatment (Figure 7).

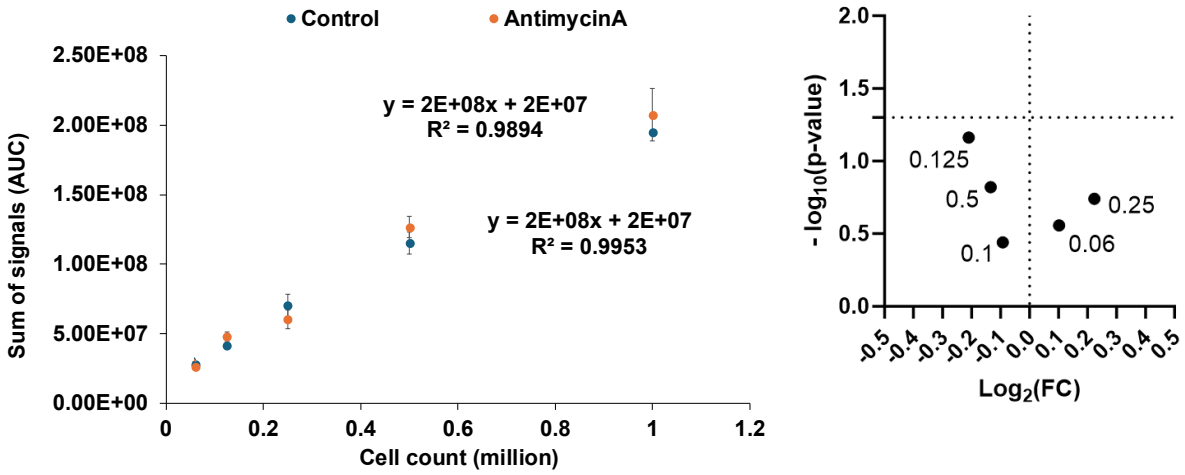


Figure 7: Validation of the total peak area normalization method. Left: Linear regression illustrating that the sum of all integrated metabolites peak areas correlates well with the cell counts, independent of the treatment. PC-3 cells were seeded in serial cell counts (1, 0.5, 0.25, 0.125, 0.06 million cells) in T-25 flasks and treated with a fixed concentration of antimycin A (a complex III inhibitor) for 4 h. After LC-MS/MS analysis, the total sum of the detectable metabolites' intensities for both control and treatment were calculated and blotted against the cell count. Right: The significance of the differences between control and treatment was calculated using multiple t-test and expressed as p-value. To evaluate the significance, log_{10}(p-value) was plotted against the average log_2 fold change of sum of peak areas of each treatment point and the respective control. The dotted line on the y axis represents the p-value= 0.05. Each data point represents the two tailed t test result calculated for the respective cell count in both treatment and control.

The cells were treated for 4 hours, which is sufficient time to elicit distinguishable metabolic alterations in the cells without affecting cell viability and, hence, cell count. In Figure 8-A a principal component analysis (PCA) is presented, illustrating the disparities between control and treatment samples, predominantly along Principal Component 1 (PC1; 74.3%). This component accounts for the majority of the variance, highlighting the distinct metabolic profiles induced by treatment. (Figure 8-B,C) showcases volcano plots representing the metabotypes of samples originating from varying cell counts after a 4 h treatment with antimycin A. These plots reveal significant metabolic shifts, characterized, for instance, by the accumulation of dihydroorotate (DHO) and carbamoylaspartate (CA), alongside a depletion of orotate (ORO) and intermediates of the pyrimidine biosynthesis pathway. These metabolic alterations, indicative of antimycin A's biological effect, are discussed in detail in section 4.3, underscoring the treatment's impact on cellular metabolism. One of the main outcomes of validating the total peak area normalization method (Figure 7) was the demonstration a verifiable correlation between cell number and sum signal intensity (R^2 > 0.989 for both control and treatment). Additionally, the strong similarity between the slopes in both trend lines for control and treatment groups further supports this correlation. Importantly, the comparison of the sum of all peaks areas between treated and control groups for each cell count, using statistical two tailed distribution t-tests, revealed no significant differences (Figure 7). All data points showed

no significance with a *p*-value higher than 0.05, supporting the hypothesis that metabolic changes can be effectively normalized by means of the total signal intensity (integral normalization), correlating accurately with actual cell counts. Moreover, the proposed weaknesses of this normalization approach have been addressed by providing evidence on the consistency of the sum of all signals regardless of metabolic changes due to treatment effects. This holds true even when the treatment results in major changes in single metabolites, as demonstrated by antimycin A treatment, which resulted in the accumulation of CA and DHO with more than 115- and 36-fold changes, respectively.

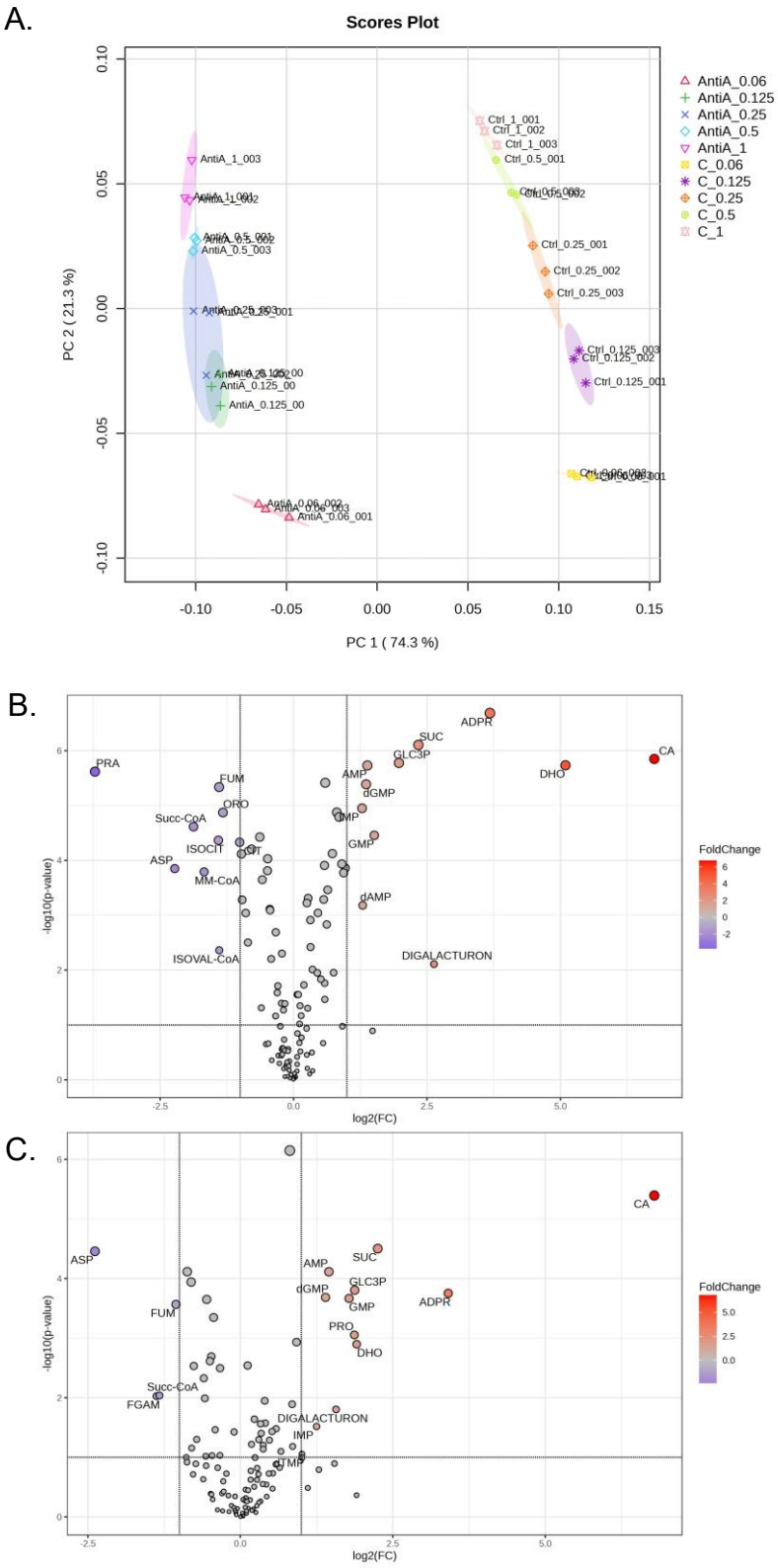


Figure 8: Impact of antimycin A (CPLX III inhibitor) on the metabolism of PC-3 cells with different cell count. Cells were treated for 4 hours, followed by metabolomics extraction and measurement. (A) Score plot of a PCA analysis of the metabotypes of control and antimycin A-treated PC-3 cells in all tested cell counts. PCA was done after range scaling. Volcano plots demonstrating the metabolites' intensity alteration after antimycin A treatment of the PC-3 cells at higher seeding count (B) and lower seeding count (C), in comparison to the untreated control. Log_2 aids in calculating fold change, and up-regulated versus down-regulated metabolites: Additionally, statistical significance can be evaluated with log_{10}(p-value). These graphs were created by using MetaboAnalyst 5.0.

To minimize disparities between control and treated samples that might be caused by variations in cell count due to cell death, an additional data processing step was implemented. This step involved the normalization of each data point (signal intensity) to a reference experimental value, specifically the median signal intensities (peak area) of the control samples within each individual experiment. In other words, all samples were normalized to the average total signal intensity of all respective control samples, followed by a base 2 logarithm transformation.

By applying this approach, we established a more convincing normalization method for the integration into our metabolomics workflow. This method is straightforward, easy to handle, and more appropriate as a standard procedure for detecting outliers and removing unreliably detected metabolites.

After data normalization, the analysis will proceed with pattern identification and the assessment of similarities across the obtained metabolic profiles. This process includes the implementation of machine learning techniques, which are discussed in detail below (4.7). Figure 9, provides a summary of the general workflow for the developed metabolomics methodology.

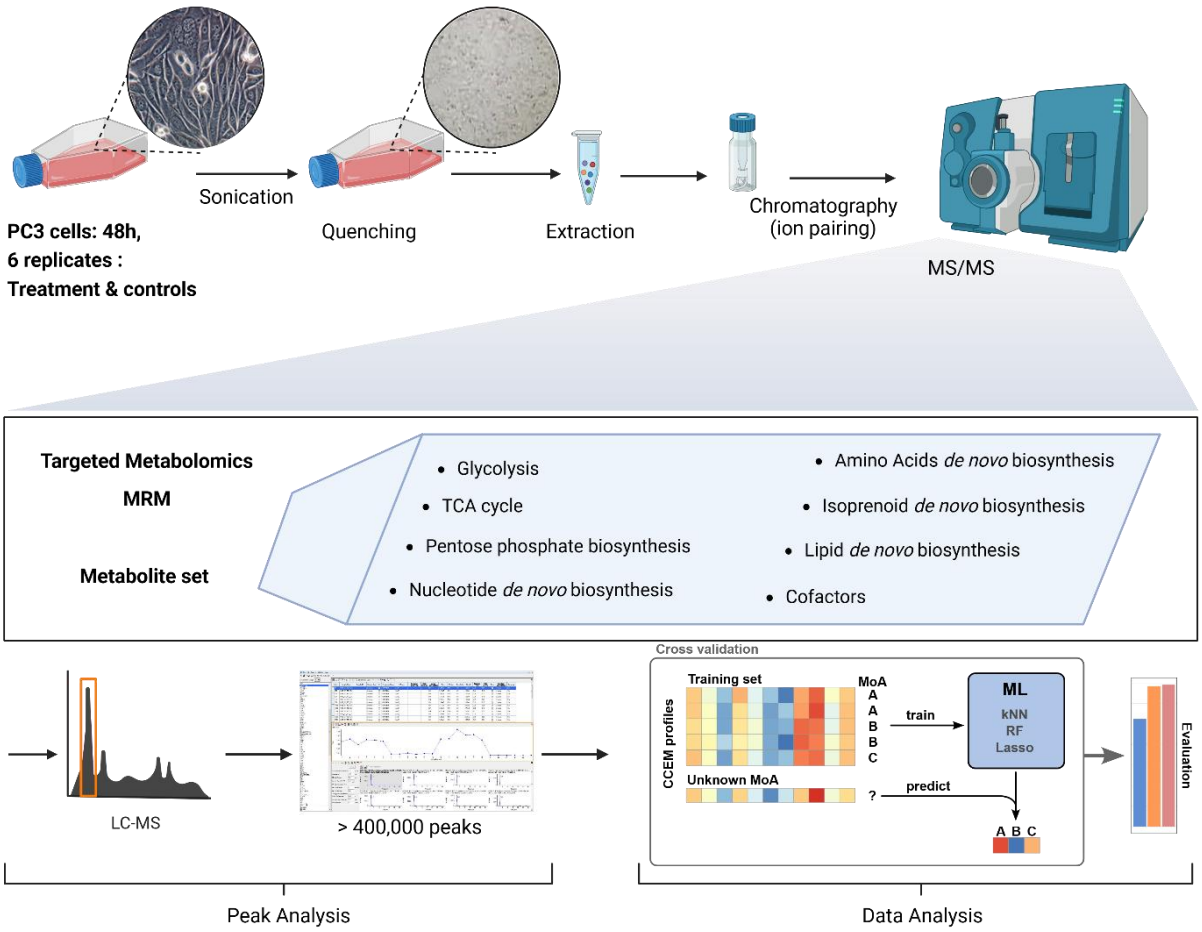


Figure 9: Analysis workflow for MoA prediction. The workflow starts with metabolite extraction from the untreated control cells and 48 h compound treated cells (six replicates for each group). The extraction starts with the quenching and sonication followed by extraction of the metabolites under cold and acidic conditions. The extracts contain hydrophilic metabolites, which are then separated by using ion pairing liquid chromatography and measured with MRM-based MS/MS analysis targeting central carbon and energy metabolism (CCEM). Post analysis, metabolite signals are collected by peak area integration. Finally, after normalization and scaling steps ML methods were used to further investigate the generated data. This graph was created with Biorender.com

4.2 Training of metabolic patterns enable the establishment of MoA prediction method

To establish a reference framework for metabolic responses in the CCEM, we utilized IC_{50} treatments for 48 h of the PC-3 prostate cancer cell model with 38 reference compounds with known MoA (Table 1). We assembled the reference compounds library based on the drugs' diverse and confirmed drug targets that impede PC-3 cell growth, sufficient knowledge on the MeA/MoA, and inclusion of at least two drugs for each MoA [65-68]. Our reference compounds not only comprised US Food and Drug Administration (FDA) approved chemotherapeutics, but also included additional experimental toxins to allow more specificity of MoA prediction. The list includes mitochondrial inhibitors, inhibitors of the NAD salvage, inhibitors of the

NADPH supply, inhibitors of the biosyntheses of amino acids, fatty acids and isoprenoids, and drugs having rather indirect effects on the cellular metabolism, such as inhibitors of PI3K/AKT/mTOR signaling, cytoskeleton and DNA replication. To ensure the comparability and reproducibility for all data, we decided to utilize the individual IC_{50} drug concentrations as predetermined by ourselves for all treatments.

*Table 1: IC_{50} values determined in preliminary cell viability assays (mostly by using MTT assay or, if indicated with #, by using CV assay), MoA and structural information for training and prediction compounds. Without further indication; PC-3 prostate cancer cell line; * : MCF-7 breast cancer cell line; ** : MHH-ES1 Ewing' sarcoma.*

Training/ Prediction	Training Compound	Isomeric SMILES	Abb. (drug)	MoA	Abb. (MOA)	IC_{50}	Pathway
TRAINING	3-Nitropropionic acid (CAS Nr- 504-88-1, Cayman chemical Item No. 14684)	C(C[N+](=O)[O-])C(=O)O	3-NP	Complex II Inhibitor [180, 181]	CPLX II	7.02 mM	Mitochondrial respiration
	Malonic acid (CAS Nr. 141-82-2, Sigma-Aldrich Item No. 8.00387)	C(C(=O)O)C(=O)O	MALO			7.13 mM	
	Metformin (CAS Nr. 1115-70-4, Cayman Chemical Item Nr. Cay13118-1)	CN(C(=N)N=C(N)N)C	METF	Complex I Inhibitor [182, 183]	CPLX I	16.68 mM	
	Rotenone (CAS Nr. 83-79-4, Merck Item Nr. R8875)	CC(=C)[C@H]1CC2=C(O1)C=CC3=C2O[C@@H]4COC5=CC(=C(C=C5[C@@H]4C3=O)OC)OC	ROTN			1.0 µM	
	Antimycin A (CAS Nr. 1397-94-0, Sigma-Aldrich Item Nr. A8674)	CCCCCC[C@@H]1[C@H]([C@@H](OC(=O)[C@H]([C@H](OC1=O)C)NC(=O)C2=C(C(=CC=C2)NC=O)O)C)OC(=O)CC(C)C	AMYC	Complex III Inhibitor [183, 184]	CPLX III	35.6 µM	
	Atovaquone (CAS Nr. 95233-18-4, MCE Item Nr. HY-13832)	OC1=C([C@H]2CC[C@@H](CC2)C2=CC=C(Cl)C=C2)C(=O)C2=CC=CC=C2C1=O	ATOV			8.9 µM	
	Potassium cyanide (CAS Nr. 151-50-8, Sigma-Aldrich Item Nr. 207810)	[C-]#N.[K+]	CYAN	Complex IV inhibitor [185]	CPLX IV	11.6 mM	
	Sodium azide (CAS Nr. 26628-22-8, Sigma-Aldrich Item Nr. S2002)	[N-]=[N+]=[N-].[Na+]	AZID			9.27 mM	
	2,4-Dinitrophenol (CAS Nr. 51-28-5, Sigma-Aldrich Item Nr. D198501)	C1=CC(=C(C=C1[N+](=O)[O-])[N+](=O)[O-])O	2DNP	Uncoupler [186, 187]	Uncoupler	0.46 mM	
	Emodin (CAS Nr. 518-82-1, TCI Item Nr. E0500)	CC1=CC2=C(C(=C1)O)C(=O)C3=C(C2=O)C=C(C=C3O)O	EMOD			37.0 µM	
	Carbonyl cyanide m-chlorophenyl	C1=CC(=CC(=C1)Cl)NN=C(C#N)C#N	CCCP			4.7 µM	

Training/ Prediction	Training Compound	Isomeric SMILES	Abb. (drug)	MoA	Abb. (MOA)	IC_{50}	Pathway
	hydrazone (CAS Nr. 555-60-2, Sigma-Aldrich Item Nr. C2759)						
	Lovastatin (CAS Nr. 75330-75-5, MCE Item Nr. HY-N0504)	CC[C@H](C)C(=O)O[C@H]1C[C@H](C=C2[C@H]1[C@H]([C@H](C=C2)C)CC[C@@H]3C[C@H](CC(=O)O3)O)C	LOVA	HMG-CoA reductase inhibitor [188]	HMG-CoAr	1.2 µM 12.18 µM* 8.02 µM**	Mevalonate Pathway
	Atorvastatin (CAS Nr. 134523-00-5, MCE Item Nr. HY-B0589)	CC(C)C1=C(C(=C(N1CC[C@H](C[C@H](CC(=O)O)O)O)C2=CC=C(C=C2)F)C3=CC=CC=C3)C(=O)NC4=CC=CC=C4	ATOR			3.5 µM 24.64 µM* 11.09 µM**	
	Fluvastatin (CAS Nr. 93957-54-1, MCE Item Nr. HY-14664)	CC(C)N1C2=CC=CC=C2C(=C1/C=C/[C@H](C[C@H](CC(=O)O)O)O)C3=CC=C(C=C3)F	FLUV			1.0 µM	
	Paclitaxel (CAS Nr. 33069-62-4, abcr Item Nr. AB251902)	CC1=C2[C@H](C(=O)[C@@]3([C@H](C[C@@H]4[C@]([C@H]3[C@@H]([C@@](C2(C)C)(C[C@@H]1OC(=O)[C@@H]([C@H](C5=CC=CC=C5)NC(=O)C6=CC=CC=C6)O)O)OC(=O)C7=CC=CC=C7)(CO4)OC(=O)C)O)C)OC(=O)C	PTXL	Antimicrotubule [189, 190]	Antimicrotubule	1.6 nM	Cytoskeleton
	Vincristin (CAS Nr. 2068-78-2, abcr Item Nr. AB348020)	CC[C@@]1(C[C@@H]2C[C@@](C3=C(CCN(C2)C1)C4=CC=CC=C4N3)(C5=C(C=C6C(=C5)[C@]78CCN9[C@H]7[C@@](C=CC9)([C@H]([C@@]([C@@H]8N6C=O)(C(=O)OC)O)OC(=O)C)CC)OC)C(=O)OC)O	VINC			0.9 nM	
	Doxorubicin (CAS Nr. 25316-40-9, TCI Item Nr. D4193)	C[C@H]1[C@H]([C@H](C[C@@H](O1)O[C@H]2C[C@@](CC3=C2C(=C4C(=C3O)C(=O)C5=C(C4=O)C(=CC=C5)OC)O)(C(=O)CO)O)N)O	DOXO	Topoisomerase II inhibitor [191]	TopoII	1.2 µM	DANN replication
	Mitoxantrone (CAS Nr. 65271-80-9, MCE Item Nr. HY-13502)	OCCNCCNC1=C2C(=O)C3=C(O)C=CC(O)=C3C(=O)C2=C(NCCNCCO)C=C1	MITO			1.6 µM	
	Etoposide (CAS Nr. 33419-42-0, TCI Item Nr. E0675)	C[C@@H]1OC[C@@H]2[C@@H](O1)[C@@H]([C@H]([C@@H](O2)O[C@H]3[C@H]4COC(=O)[C@@H]4[C@@H](C5=CC6=C(C=C35)OCO6)C7=CC(	ETOP			48.5 µM	

Training/ Prediction	Training Compound	Isomeric SMILES	Abb. (drug)	MoA	Abb. (MOA)	IC_{50}	Pathway
		=C(C(=C7)OC)O)OC)O)O					
	Irinotecan (CAS Nr. 97682-44-5, MCE Item Nr. HY-16562)	CCC1=C2CN3C(=CC4=C(C3=O)COC(=O)[C@@]4(CC)O)C2=NC5=C1C=C(C=C5)OC(=O)N6CCC(CC6)N7CCCCC7	IRIN	Topoisomerase I inhibitor [192]	TopoI	67.4 µM	
	Camptothecin (CAS Nr. 7689-03-4, Cayman Chemica Item Nr. Cay11694)	CC[C@@]1(C2=C(COC1=O)C(=O)N3CC4=CC5=CC=CC=C5N=C4C3=C2)O	CMPT			0.8 µM	
	Daporinad (CAS Nr. 658084-64-1, Selleckchem Item Nr. S2799)	C1CN(CCC1CCCCNC(=O)/C=C/C2=CN=CC=C2)C(=O)C3=CC=CC=C3	FK866	NAMPT Inhibitor [193]	NAMPT	0.3 nM #	NAD+ salvage
	GMX1778 (CAS Nr. 200484-11-3, Selleckchem Item Nr. S8117)	C1=CC(=CC=C1OCCCCCCN=C(NC#N)NC2=CC=NC=C2)Cl	GMX			0.1 nM #	
	GNE617 (CAS Nr. 1362154-70-8, Cayman Chemica Item Nr. Cay22463)	C1=CC(=CC=C1CNC(=O)C2=CN3C=CN=C3C=C2)S(=O)(=O)C4=CC(=CC(=C4)F)F	GNE			20.0 µM #	
	Hexachlorophene (CAS Nr.70-30-4, Cayman Chemical Item Nr. Cay23948)	CLC1CC(CL)C(C(C1O)CC1C(O)C(CL)CC(C1CL)CL)CL	HEXA	GDH Inhibitor [194]	GDH	10.0 µM 20.72 µM* 10.51 µM**	Glutaminolysis
	Bithionol (CAS Nr. 97-18-7, MCE Item Nr. HY-17592)	CLC1CC(SC2CC(CL)CC(C2O)CL)C(C(C1)CL)O	BITN			31.3 µM 22.25 µM* 20.87 µM**	
	Rapamycin (CAS Nr. 53123-88-9, MCE Item Nr. HY-10219)	C[C@@H]1CC[C@H]2C[C@@H](/C(=C/C=C/C=C/[C@H](C[C@H](C(=O)[C@@H]([C@@H](/C(=C/[C@H](C(=O)C[C@H](OC(=O)[C@@H]3CCCCN3C(=O)C(=O)[C@@]1(O2)O)[C@H](C)C[C@@H]4CC[C@H]([C@@H](C4)OC)O)C)/C)O)OC)C)/C)OC	RAPA	PI3K/mTOR Inhibitor [195, 196]	mTOR	0.05 µM	Signaling/ Protein phosphorylation
	Wortmannin (CAS Nr. 19545-26-7, MCE Item Nr. HY-10197)	CC(=O)O[C@@H]1C[C@]2([C@@H](CCC2=O)C3=C1[C@]4([C@H](OC(=O)C5=COC(=C54)C3=O)COC)C)C	WRTN			4.1 µM	
	Alpelisib (CAS Nr. 1217486-61-7, MCE Item Nr. HY-15244)	CC1=C(SC(=N1)NC(=O)N2CCC[C@H]2C(=O)N)C3=CC(=NC=C3)C(C)(C)C(F)(F)F	ALPL			26.1 µM	

Training/ Prediction	Training Compound	Isomeric SMILES	Abb. (drug)	MoA	Abb. (MOA)	IC_{50}	Pathway
	Perifosine (CAS Nr. 157716-52-4, MCE Item Nr. HY-50909)	CCCCCCCCCCCCCCCCCOP(=O)(OC1CC[N+](CC1)(C)C)[O-]	PRFN	AKT Inhibitor [197, 198]	AKT	2.6 µM	
	Oridonin (CAS Nr. 28957-04-2, MCE Item Nr. HY-N0004)	CC1(CC[C@@H]([C@]23[C@@H]1[C@@H]([C@]([C@]45[C@H]2CC[C@H]([C@H]4O)C(=C)C5=O)(OC3)O)O)O)C	ORID			15.7 µM	
	6-Aminonicotinamide (CAS Nr. 329-89-5, MCE Item Nr. HY-W010342)	NC(=O)C1CCC(N)NC1	6-AN	Inhibitor of oxidative PPP [199, 200]	OPP	180.0 µM#	NADPH biosynthesis
	glucose-6-phosphate dehydrogenase (G6PD) inhibitor-1 (CAS Nr. 2457232-14-1, Cayman Chemical Item Nr. Cay31484)	C1CCC(=O)C2=CN=C(N=C2C1)NC3=CSC(=C3)C#N	GPDi			59.0 µM	
	Maslinic acid (CAS Nr. 4373-41-5, IPB in-house compound library)	C[C@@]12CC[C@@H]3[C@@]([C@H]1CC=C4[C@]2(CC[C@@]5([C@H]4CC(CC5)(C)C)C(=O)O)C)(C[C@H]([C@@H](C3(C)C)O)O)C	MASA	Putative CEPT1 Inhibitor [201]	PLB	57.6 µM	Induce apoptosis
	Betulinic acid (CAS Nr. 472-15-1, IPB in-house compound library)	CC(=C)[C@@H]1CC[C@]2([C@H]1[C@H]3CC[C@@H]4[C@]5(CC[C@@H](C([C@@H]5CC[C@]4([C@@]3(CC2)C)C)(C)C)O)C(=O)O	BETA			19.8 µM	
	11-Keto-beta-boswellic acid (CAS Nr. 17019-92-0, IPB in-house compound library)	C[C@@H]1CC[C@@]2(CC[C@@]3(C(=CC(=O)[C@H]4[C@]3(CC[C@@H]5[C@@]4(CC[C@H]([C@]5(C)C(=O)O)O)C)C)[C@@H]2[C@H]1C)C)C	BOWA			76.4 µM	
	Epigallocatechin gallate (CAS Nr. 989-51-5, IPB in-house compound library)	C1[C@H]([C@H](OC2=CC(=CC(=C21)O)O)C3=CC(=C(C(=C3)O)O)O)OC(=O)C4=CC(=C(C(=C4)O)O)O	EGCG	FASN Inhibitor [202, 203]	FAB	61.4 µM	Lipid biosynthesis
	Apigenin (CAS Nr. 520-36-5, IPB in-house compound library)	OC1CCC(CC1)C1CC(=O)C2C(O1)CC(CC2O)O	APIG			22.8 µM	
Neg. Control	5-Flurouracil (CAS Nr. 51-21-8, Cayman Chemica Item Nr. Cay14416)	Fc1c[nH]c(=O)[nH]c1=O.	5-FU	Thymidylate synthase inhibitor [64]	TS	100 µM	Nucleotide biosynthesis

Training/ Prediction	Training Compound	Isomeric SMILES	Abb. (drug)	MoA	Abb. (MOA)	IC_{50}	P
PREDICTION	Breastin (IPB in-house compound library)		BRST	Unknown		1.5 µg/ml	
	AAHR (IPB in-house compound library)	C[C@@]12CC[C@]([C@@](COC(C)=O)(C)[C@H]3OC(C)=O)([H])[C@@](C[C@H]3OC(C)=O)(C)[C@@]1([H])CC=C4[C@@]2(C)CC[C@](CC5)(C(N6CCCN(C(C7=CC=CC=C7C(C8=C(C=C(N(CC)CC)C=C8)O9)=C(C=C/%10)C9=CC%10=[N+](CC)\CC)=O)CC6)=O)[C@@]4([H])[C@@H](C)[C@@H]5C.[Cl-]	AAHR		285 nM	
	Cucurbitacin B (CAS Nr. 6199-67-3, IPB in-house compound library)	CC(=O)OC(C)(C)/C=C/C(=O)[C@@](C)([C@H]1[C@@H](C[C@@]2([C@@]1(CC(=O)[C@@]3([C@H]2CC=C4[C@H]3C[C@@H](C(=O)C4(C)C)O)C)C)C)O)O	QQrB			30 nM	
	Glycyrrhetinic acid (CAS Nr. 471-53-4, IPB in-house compound library)	C[C@]12CC[C@](C[C@H]1C3=CC(=O)[C@@H]4[C@]5(CC[C@@H](C([C@@H]5CC[C@]4([C@@]3(CC2)C)C)(C)C)O)C)(C)C(=O)O	GLYA			73.1 µM	

For the vast majority of MoA training groups (95%), hierarchical clustering of the metabolic profiles revealed co-segregation of those training compounds addressing the same target (Figure 10). Interestingly, the dendrogram identifies two main clusters: cluster I harboring compounds affecting mitochondrial electron transport processes, such as direct inhibition of the oxidative phosphorylation (OXPHOS) or glutamate and NAD metabolism; and cluster II comprising MoA patterns of other metabolic processes, some of which are known to proceed in the cytosol (e.g. hydroxymethylglutaryl-CoA reduction (HMG-CoAr; EC 1.1.1.88), or fatty acid biosynthesis (FAB) by fatty acid synthase (FASN, EC 2.3.1.85) (cluster II) [204, 205]. Other inhibitors assigned to this clade are known to impair microtubule formation or degradation, to inhibit topoisomerases, or to modulate PI3K/AKT/mTOR signaling.

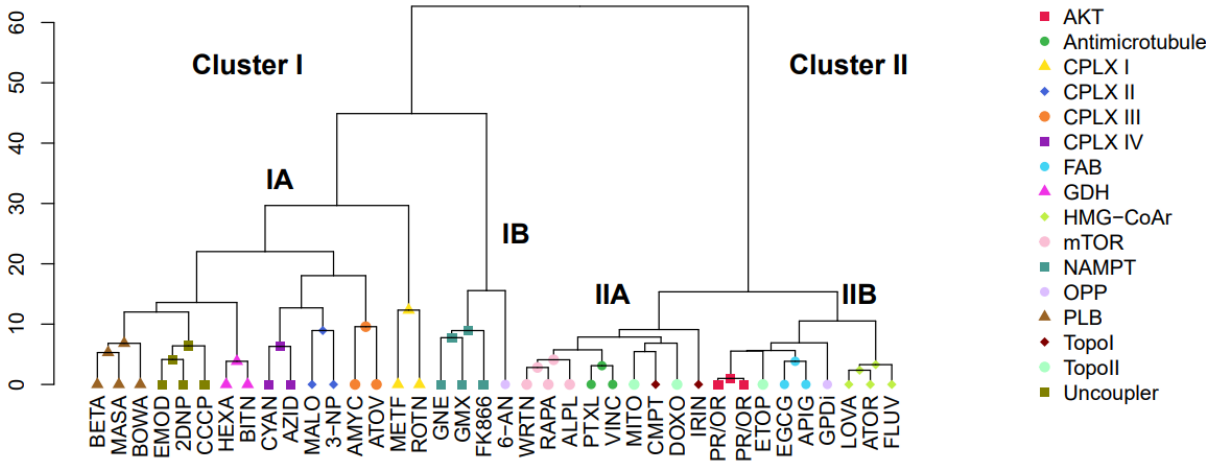


Figure 10: Hierarchical cluster analysis of metabolic patterns induced by 38 reference compounds inhibiting different molecular targets modulating the metabolism of human prostate cancer cells (PC-3). Aggregated leaves of hexuplicate experimental data are presented. Metabolic patterns of AKT inhibitors were so similar that individual replicates of the inhibitors PR and OR co-clustered with each other. For this reason, we defined a mixed group "PR/OR".

Compounds: BETA – betulinic acid, MASA – maslinic acid, BOWA – boswellic acid, EMOD – emodin, 2DNP – 2,4-dinitrophenol, CCCP – carbonyl cyanide chlorophenylhydrazone, HEXA – hexachlorophene, BITN – bithionol, CYAN – potassium cyanide, AZID – sodium azide, MALO – malonic acid, 3-NP – 3-nitropropionic acid, AMYC – antimycin A, ATOV – atovaquone , METF – metformin, ROTN – rotenone, GNE – GNE-617, GMX – GMX1778, FK866 – FK866, 6-AN - 6-aminonicotinamide, WRTN - wortmannin, RAPA – rapamycin, ALPL – alpelisib, PTXL – paclitaxel, VINC – vincristin, MITO – mitoxantrone, CMPT – camptothecin, DOXO – doxorubicin, IRIN – irinotecan, PRFN (PR) – perifosine, ORID (OR) – oridonin, ETOP – etoposid, EGCG – epigallocatechin gallate, APIG – apigenin, GPDi – glucose-6-phosphate dehydrogenase inhibitor, LOVA – lovastatin, ATOR – atorvastatin, FLUV – fluvastatin.

MoA: AKT – protein kinase B (AKT), Antimicrotubule, CPLX I – complex I, CPLX II – complex II, CPLX III – complex III, CPLX IV – complex IV, FAB – fatty acid biosynthesis, GDH - glutamate dehydrogenase, HMG-CoAr – HMG-CoA reductase, mTOR – PI3K/mTOR signaling, NAMPT – nicotinamide phosphoribosyltransferase, OPP – oxidative pentose phosphate pathway, PLB – phospholipid biosynthesis, TopoI – topoisomerase I, TopoII – topoisomerase II, Uncoupler – uncoupling of oxidative phosphorylation.

Only a few reference compounds targeting the same MoA did not co-cluster in our analysis. For example, the inhibitors of the oxidative pentose phosphate pathway (OPP), 6-aminonicotinamide (6-AN) and G6PDi, both causing accumulation of 6-phosphogluconate (6-PG), the substrate of 6-phosphogluconate dehydrogenase (6-PGDH, EC 1.1.1.44). However, their drug-induced metabolites' patterns in PC-3 cells showed different clustering behavior. 6-AN's pattern co-segregated with nicotinamide phosphoribosyltransferase (NAMPT, EC 2.4.2.12) inhibitors, while G6PDi's pattern showed higher similarity with protein kinase B (AKT, EC 2.7.11.1) and fatty acid biosynthesis (FAB) inhibitors (Figure 10). Another interesting finding was that etoposide (ETOP) did not cluster with four other topoisomerase inhibitors tested, suggesting unique metabolic interactions resulting in a distinct metabotype. Through this analysis, we obtained a clustered map of metabolic fingerprints (characteristic metabotypes) of established and proven reference drugs with known MoAs/MeAs, indicating the high quality of the procedure. This map can now be used to identify previously unknown

effects of these reference compounds and, particularly, to characterize the MoAs/MeAs of new, yet uncharacterized, natural products.

4.3 Inhibition of mitochondrial functions produces distinctive metabotypes

Although cancer cells produce ATP primarily by enhancing glycolytic flux, they still require functional mitochondria [206-208], due to the essential role of OXPHOS and its connection to seven ubiquinone (CoQ)-dependent mitochondrial dehydrogenases crucial for CCEM [209]. Consequently, targeting different stages of OXPHOS has specific effects on various metabolic processes linked to these CoQ-dependent mitochondrial dehydrogenases. This highlights the complexity of metabolic dependencies in cancer cells.

Initially, we observed that OXPHOS inhibitors lead to specific metabotypes that are categorized into cluster I (Figure 10). Surprisingly, these metabolic patterns further subdivide into distinct sub-clusters depending on which complex (CPLX) of the respiratory chain is inhibited. In the subsequent results section, a more detailed description of the observed metabolic patterns for several metabolites within certain metabolic pathways, in relation to the specific MoA is provided. Metabolites involved in glycolysis, PPP, the TCA cycle, the glycerol phosphate shuttle, and nucleotide biosynthesis can distinguish the resulting patterns following OXPHOS inhibition. The novelty of these data lies in the comprehensive analysis of these metabolites, particularly when comparing the effects of various OXPHOS inhibitors. These findings underscore the unique metabolic signatures associated with the inhibition of different OXPHOS complexes. Furthermore, these observations may improve our understanding of the potential of strategically combining these inhibitors for enhancing cytotoxic effects. The selection of the promising combinations can be based on the respective individual MoA-associated metabotype present in cancer cells.

TCA-cycle: When complexes I to IV (CPLX I-IV) were inhibited, a consistent pattern of depletion was noted in certain TCA cycle-related metabolites, such as cis-aconitate (ACT), isocitrate (ISOCIT), and succinyl-CoA (SUC-CoA) (Figure 11). Contrarily, this pattern was not observed with uncouplers. Additionally, inhibition of CPLX II-IV resulted in the accumulation of succinate (SUC) and a strong depletion of fumarate (FUM) and malate (MAL). However, these effects were not observed with CPLX I inhibition (Figure 12).

The glycerol phosphate shuttle and glycolysis exhibited distinct metabolic responses depending on the inhibited OXPHOS complex. Both CPLX inhibitors and uncouplers generally induced the accumulation of glycerol 3-phosphate (Glc3P) in the cytosol (Figure 11). Synthesis

of Glc3P could potentially rescue CPLX I deficiency by providing reducing equivalents to CoQ via the Glc3P shuttle [210]. However, with CPLX I inhibition using rotenone (ROTN) and metformin (METF), there was a consistent depletion of glycolytic and PPP intermediates. That indicated for a suppressed sugar metabolism in PC-3 cells and, as a result, limited substrate availability for the Glc3P shuttle (Figure 11). CPLX I inhibition resulted in up to eightfold depletion of dihydroxyacetone phosphate (DHAP) (Figure 13). In contrast, inhibiting CPLX II-IV led to the accumulation of DHAP, Glc3P and fructose 1,6-bisphosphate (FBP), while lower glycolysis metabolites were depleted (Figure 11). This differential pattern suggests that the glycerol phosphate shuttle responds differently to severe energy depletion depending on the inhibited complex. Despite these metabolic changes, ATP levels did not decrease significantly after OXPHOS inhibition. It is worth noting that CPLX I contributes to 40% of ATP synthesis through the proton motive force, as reported by Wang et al. [68]. Although ATP levels did not decrease significantly after inhibiting OXPHOS, the depletion of phosphocreatine was much more pronounced after CPLX I inhibition (up to a 196-fold decline) compared to CPLX II-IV inhibition (5 to 9-fold decline) (Figure 14).

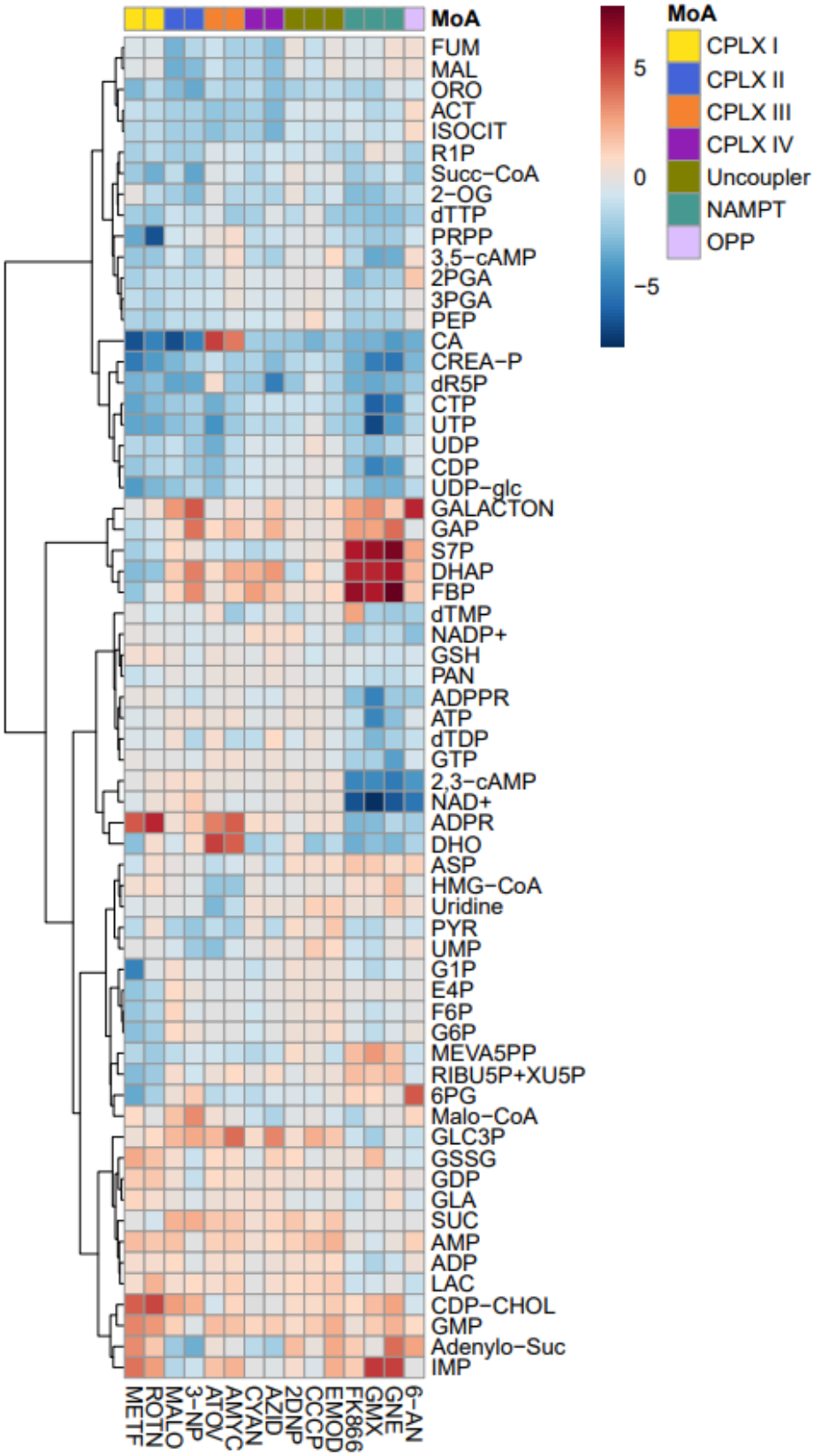


Figure 11: Relative abundance of selected CCEM intermediates after inhibition of the OXPHOS complexes I-IV, application of uncouplers, inhibition of nicotinamide phosphoribosyltransferase (NAMPT) and the oxidative pentose phosphate pathway. Data depict average log_2-fold changes of cell number-normalized peak areas obtained after 48 h of drug treatment (n = 6) relative to a vehicle control (n = 6). Legends of MoA and compound labels are given in Figure 10 and Table 1.

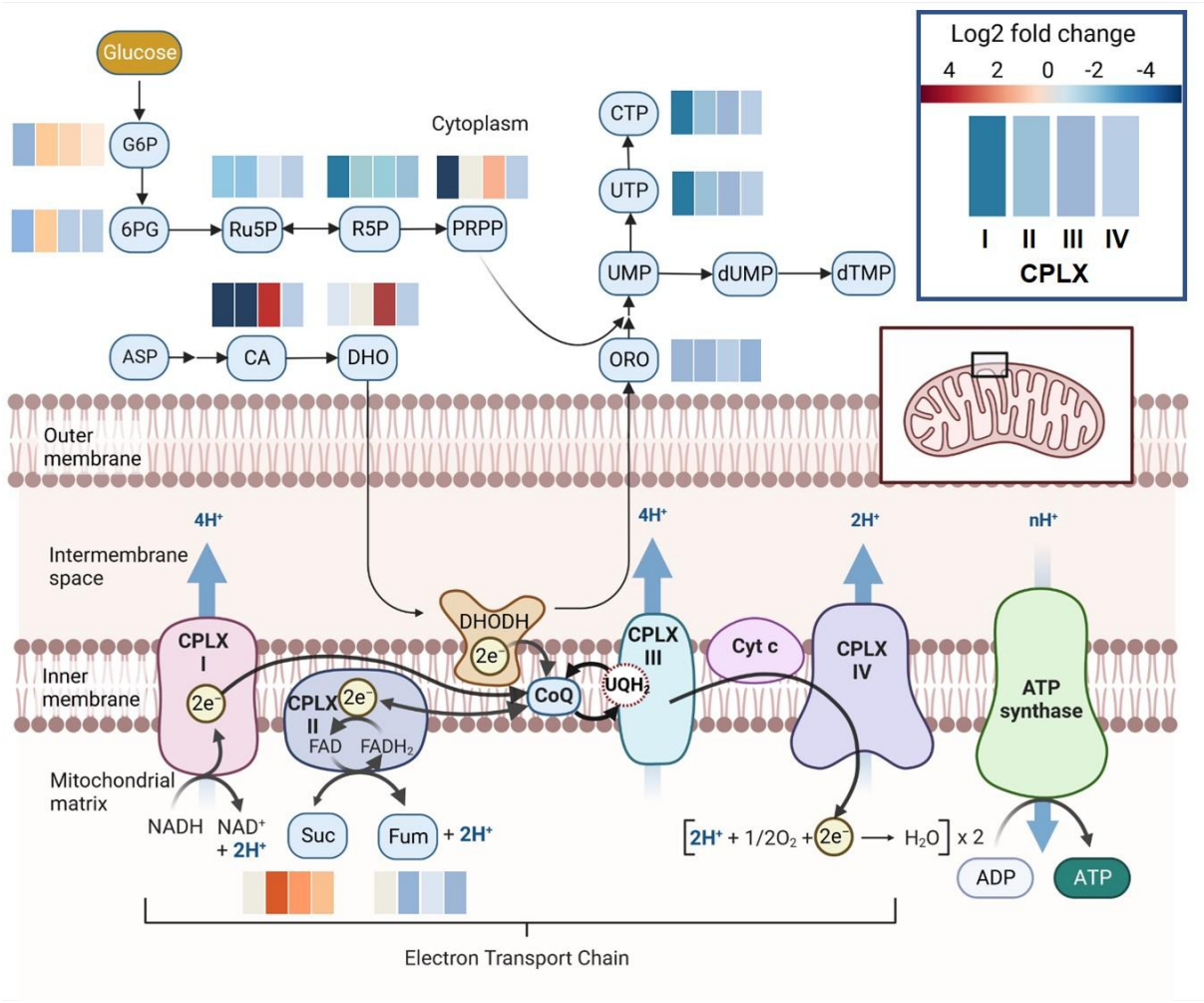


Figure 12: Relative abundance of selected intermediates from glycolysis, PPP, TCA cycle and pyrimidine BS in response to the CPLX I-IV inhibition. The heatmap displays the log_2 fold changes of these metabolites compared to untreated controls across treatments with inhibitors of CPLX I, II, III, and IV, respectively (from left to right). The mitochondrial coenzyme Q junction links pyrimidine nucleotide biosynthesis to the mitochondrial ETC. In order to fuel ATP formation, OXPHOS generates a proton gradient across the inner mitochondrial membrane through the ETC, which involves four protein complexes (CPLX I-CPLX IV). Both CPLX I and CPLX II transfer electrons to ubiquinone (coenzyme Q, CoQ) in the inner mitochondrial membrane resulting in ubiquinol (UQH_2) reconstitution. By contrast, CPLX III is responsible for the reoxidation of UQH_2 to UbQ, which is important not only for the function of CPLX I and CPLX II. CPLX IV then transfers the electrons to molecular oxygen. Metabolite' abbreviations are given in Appendix table 1. This graph was created with Biorender.com

Pentose phosphate pathway PPP: OXPHOS inhibition result in the alteration of PPP intermediates, contributing to suppressed sugar metabolism (Figure 11). While in CPLX I there is a strong depletion in the PPP intermediates, other OXPHOS inhibitors led to a moderate and strong accumulation in these metabolites.

Pyrimidine biosynthesis: Pyrimidine BS intermediates levels were reduced when OXPHOS was inhibited at CPLX I-IV but not with uncouplers (Figure 11). Pyrimidine BS in mammals begins with CAD, a large trifunctional enzyme composed of carbamoyl phosphate synthetase (CPS2), aspartate transcarbamoylase, and dihydroorotase. Different modes of OXPHOS inhibition impact CAD-associated metabolites in distinct ways (Figure 12). For instance,

inhibition of CPLX I and CPLX II leads to a significant decrease in carbamoylaspartate (CA) by 33- to 132-fold, while CPLX III inhibition results in 13-45 times higher CA level than in control cells. Similarly, dihydroorotate (DHO), the product of the third step catalyzed by CAD, accumulates strongly (20-36 fold) only after CPLX III inhibition. The subsequent biosynthetic enzyme, dihydroorotate dehydrogenase (DHODH, EC 1.3.1.14), located in the mitochondrial membrane, oxidizes DHO to orotate (ORO) in a CoQ-dependent manner. Despite the contrasting CA and DHO patterns, OXPHOS inhibition generally causes a decrease in ORO levels (Figure 12). Thus, CPLX III inhibition presents a distinct metabotype, characterized by the accumulation of upstream intermediates DHO and CA, alongside the depletion of ORO and downstream pyrimidine biosynthesis intermediates, due to the interaction with DHODH. Based on these observations, a combinatorial treatment with CPLX III and uncouplers was suggested and was proven to have synergistic effect in PC-3 cells using sub toxic concentration of (antimycin A or atovaquone) in one hand and (CCCP or 2-DNP) on the other hand (unpublished data, data not shown).

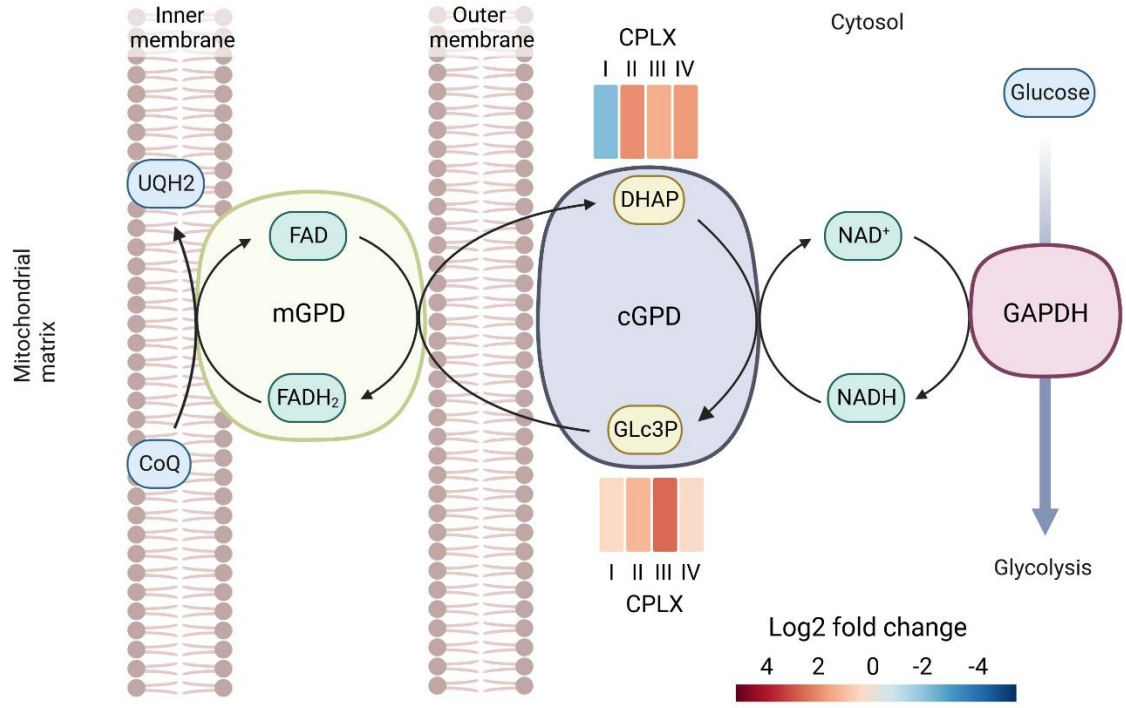


Figure 13: The glycerol phosphate shuttle antagonizes mitochondrial dysfunction in cancer cells [210]. By using cytosolic NADH, e.g. as a side-product of glycolysis, to reduce DHAP to glycerol 3-phosphate (Glc3P) cytosolic NAD is regenerated, which supports ongoing glycolytic flow. Glc3P can pass electrons to mitochondrial glycerol 3-phosphate dehydrogenase (EC 1.1.5.3), transferring reducing power to FADH2. FADH2 finally reduces CoQ at the inner mitochondrial membrane, which also regenerates DHAP. This graph was created with Biorender.com

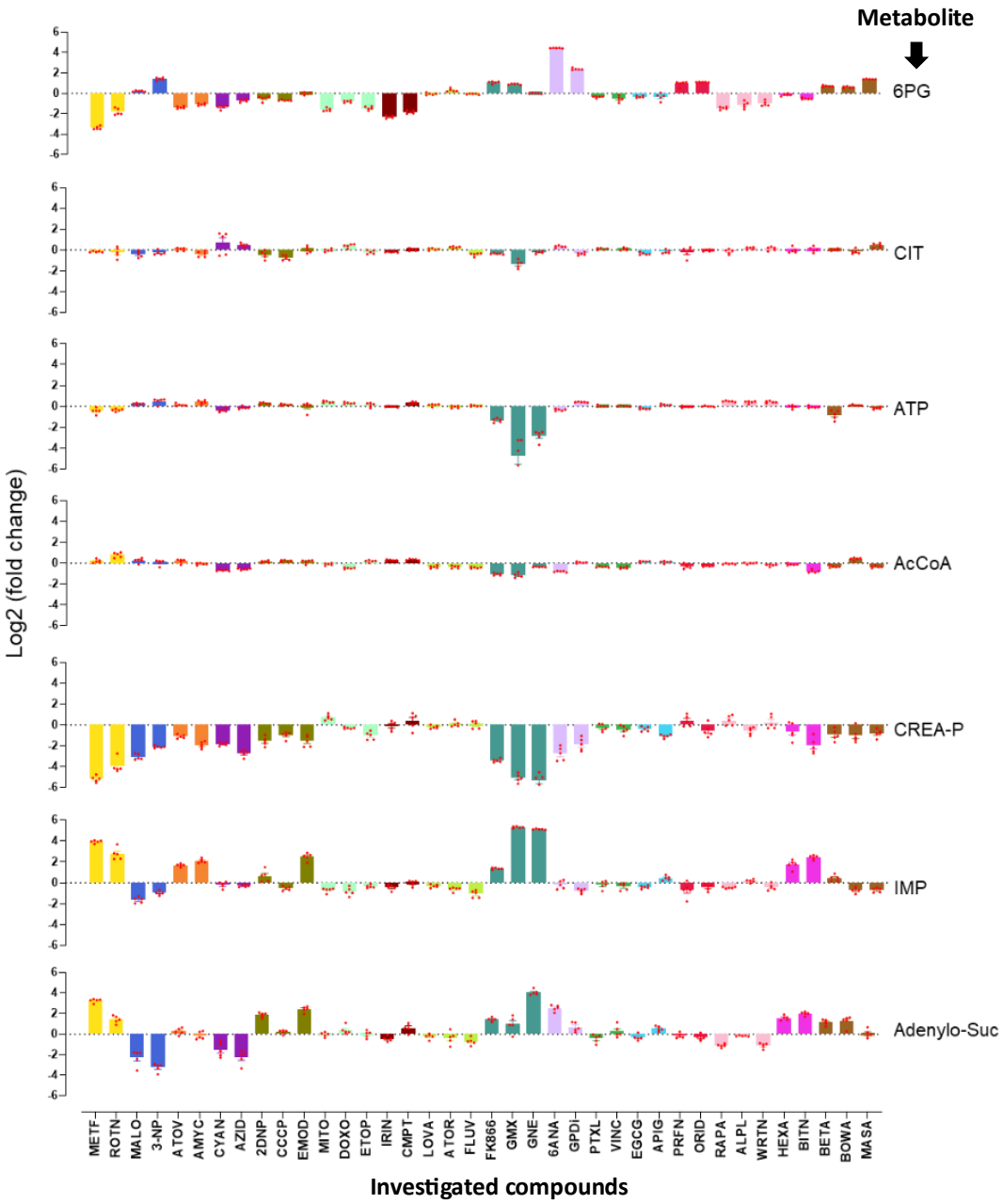


Figure 14: Relative abundance of selected metabolites across all training data. Bars represent the average log_2-fold change of normalized peak areas of n = 6 treatments and controls after 48 h incubation.

Compounds: BETA – betulinic acid, MASA – maslinic acid, BOWA – boswellic acid, EMOD – emodin, 2DNP – 2,4-dinitrophenol, CCCP – carbonyl cyanide chlorophenylhydrazone, HEXA – hexachlorophene, BITN - bithionol, CYAN – potassium cyanide, AZID – sodium azide, MALO – malonic acid, 3-NP – 3-nitropropionic acid, AMYC – antimycin A, ATOV – atovaquone , METF – metformin, ROTN – rotenone, GNE – GNE-617, GMX – GMX1778, FK866 – FK866, 6ANA – 6-aminonicotinamide, WRTN – wortmannin, RAPA – rapamycin, ALPL – alpelisib, PTXL – paclitaxel, VINC – vincristin, MITO – mitoxantrone, CMPT – camptothecin, DOXO – doxorubicin, IRIN – irinotecan, PRFN (PR) – perifosine, ORID (OR) – oridonin, ETOP – etoposid, EGCG – epigallocatechin gallate, APIG – apigenin, GPDi – glucose-6-phosphate dehydrogenase inhibitor, LOVA – lovastatin, ATOR – atorvastatin, FLUV – fluvastatin.

MoA: AKT – protein kinase B (AKT), Antimicrotubule, CPLX I – complex I, CPLX II – complex II, CPLX III – complex III, CPLX IV – complex IV, FAB – fatty acid biosynthesis, GDH – glutamate dehydrogenase, HMG-CoAr – HMG-CoA reductase, mTOR – PI3K/mTOR signaling, NAMPT – nicotinamide phosphoribosyltransferase, OPP – oxidative pentose phosphate pathway, PLB – phospholipid biosynthesis, TopoI – topoisomerase I, TopoII – topoisomerase II, uncoupler – uncoupling of oxidative phosphorylation.

Moreover, in cancer cells, glutamine/glutamate catabolism is a major source to fuel the TCA cycle with carbon and to provide energy via OXPHOS. Glutamate dehydrogenase (GDH) is a mitochondrial enzyme, which catalyzes the conversion of glutamate to the TCA cycle intermediate 2-oxoglutarate and ammonium and vice versa. Consequently, co-segregation of metabolic patterns after GDH inhibition as compared with the inhibition of other mitochondrial functions in sub-cluster IA is plausible (Figure 10).

Metabolic patterns of treatments with pentacyclic triterpenes likewise emerge in cluster I. Several pentacyclic triterpene acids have been extensively studied for their anti-cancer properties, including induction of apoptosis and inhibition of cell proliferation [211]. They target various genes and signaling pathways such as Bcl-2, NF-kB, and PI3K/Akt/mTOR, as reviewed recently [212]. Indication that such triterpenoids also affect phospholipid biosynthesis (PLB) has been provided by micelle models mimicking mitochondrial membranes, where membrane rigidification and increased membrane permeabilization were induced by the treatment with triterpenoids in dose-dependent manner [213]. A recent study investigating the MeA of ursolic acid on human prostate cancer (PCa) models found an accumulation of intermediates of PLB [214]. Nevertheless, the precise influence of pentacyclic triterpenoids on cancer cell metabolism remains so far only partially understood.

To further investigate these effects at the metabolic level we assessed the metabotypes of 11-keto-β-boswellic acid (BOWA), maslinic acid (MASA) and betulinic acid (BETA) in PC-3 cells. Our data suggests that pentacyclic triterpenoids affect mitochondrial metabolism, exhibiting metabolic patterns akin to uncouplers and GDH inhibition (Figure 10). Despite sharing some characteristics of OXPHOS inhibition in their metabolic profiles, they differ from conventional OXPHOS inhibitors, which is further investigated in section 4.8.2.

4.4 Inhibition of NAD salvage leads to severe ATP starvation

Despite significant cell growth inhibition observed in these experiments, we consistently detected stable intracellular levels of three key metabolites within CCEM. Across all investigated MoAs, a balanced state was maintained for ATP, acetyl-CoA (AcCoA), and citrate (CIT), with the exception of ATP levels in the context of NAMPT inhibition (Figure 14). In cancer cells, NAD is mainly synthesized from nicotinamide via NAD salvage [215]. NAMPT is the rate-limiting enzyme in the NAD salvage and is therefore a popular therapeutic target, especially in the combinatorial therapy [36]. As expected, NAMPT inhibition results in a metabolic pattern, which is characterized by strong depletion of NAD and NADP but also of their cleavage products ADP-ribose (ADPR) and ADP-ribose 2'-phosphate (ADPPR) (Figure 11). NAD(P) depletion, however, has several consequences leading to a specific metabolic pattern (Figure

10, sub-cluster IB), one of them, as mentioned, is a strong depletion in ATP (Figure 11 and Figure 14). Intracellular ATP homeostasis is normally maintained under moderate energy deprivation by activation of creatine kinase (EC 2.7.3.2) and adenylate kinase (EC 2.7.4.3). Consistently, for MoAs affecting the energy metabolism, we observed reduced levels of energy-storing phosphocreatine (CREA-P), and changes in the levels of inosine monophosphate (IMP) and adenylosuccinate (Adenylo-Suc), the latter two being involved in the purine nucleotide cycle that balances the equilibrium of adenylate kinase (Figure 14 and Figure 15).

As reported previously in Tan et al., limited availability of NAD also impairs the glycolytic ATP generation in cancer cells since NAD is a required coenzyme of the enzyme glyceraldehyde-3-phosphate dehydrogenase (GAPDH) which fosters the diversion of the carbon flow into upper glycolysis and PPP [37, 216]. Accordingly, the metabolic patterns of three NAMPT inhibitors consistently show very strong accumulation of FBP and triose phosphates, accompanied by a depletion of intermediates of the lower glycolysis such as 3-phosphoglyceric acid (3-PGA), phosphoenolypyruvate (PEP) and pyruvate (PYR) (Figure 11). Lacking NAD(P) also affects intermediate levels of the non-oxidative PPP, resulting in increased levels of pentose phosphates and sedoheptulose-7-phosphate (S7P), whereas (6-PG) levels of the OPP remain unaffected (Figure 11 and Figure 14). This discriminates NAMPT inhibitors from OXPHOS inhibitors, which do not accumulate pentose or heptulose phosphates, and from OPP inhibitors, which strongly accumulate 6-PG.

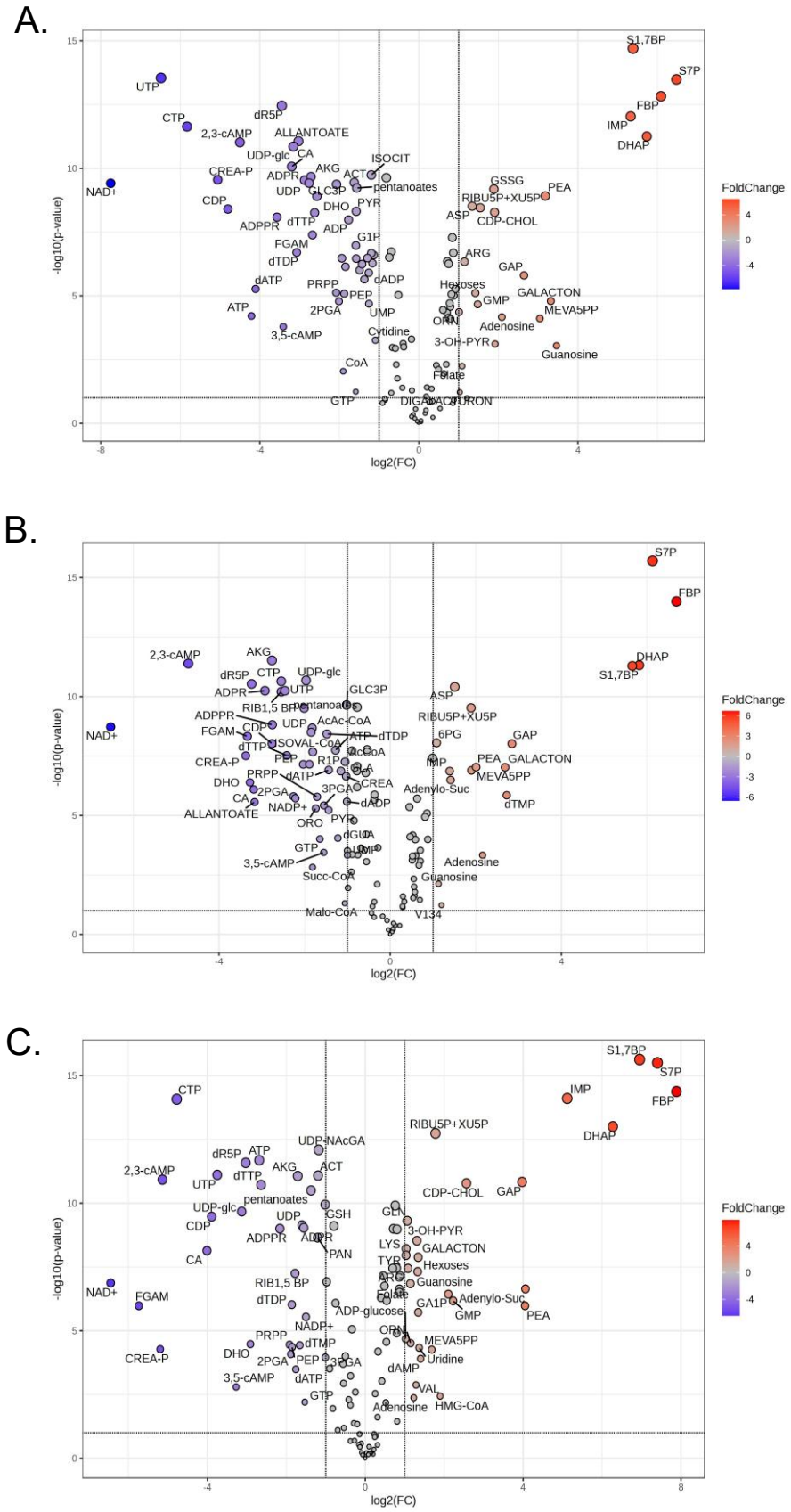


Figure 15: Volcano plots depicting relative metabolite abundances represented by log2(Fold Change) (average of n = 6 replicates) after 48 h treatment with (A) GMX1778, (B) FK866 and (C) GNE-617 versus non-treated vehicle controls. Accumulation (red), depletion (blue) and non-significant changes (grey).

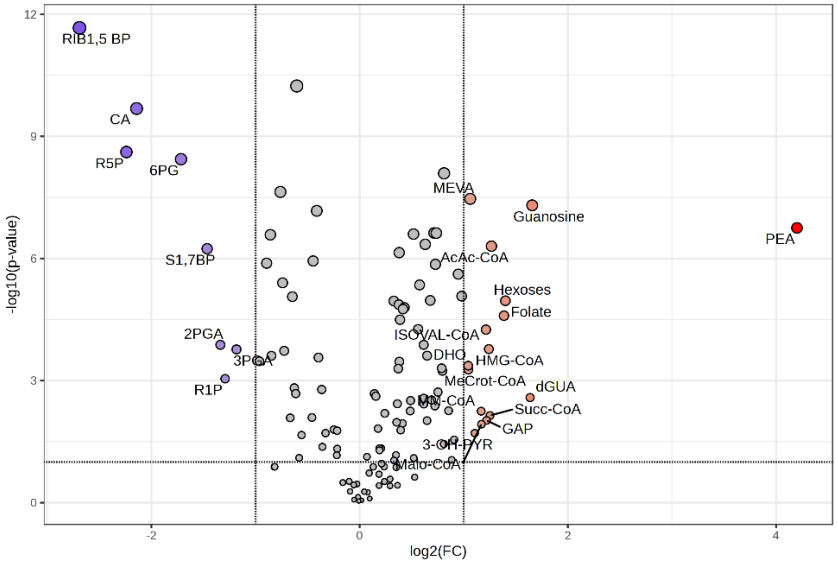
However, the metabolic patterns of both OPP inhibitors tested diverged, with only 6-AN sharing similarities with the metabolic patterns seen after NAMPT inhibition. Notably, the second OPP inhibitor tested, G6PDi, was previously described to inhibit only glucose 6-phosphate dehydrogenase (G6PDH, EC 1.1.1.49), whereas 6-AN is known to inhibit also the second line of NADPH-producing enzymes in the OPP, 6-PGDH [200, 217]. Until today, it has not been described whether 6-AN acts as nicotinamide analogue in NAD salvage via NAMPT converting 6-AN to 6-amino-NAD. However, it is known that 6-amino-NAD(P) are potent inhibitors of metabolic processes requiring NAD(P), such as glycolysis and OPP [218]. Therefore, strong commonalities between NAMPT inhibitors and 6-AN indicate that 6-amino-NAD might act as a NAMPT inhibitor as well, highlighting the need for further investigation.

In summary, inhibitors of NAD biosynthesis and OXPHOS inhibitors share great similarities in their metabolic patterns. Since both have a strong influence on cellular energy metabolism, their patterns are also grouped together in a common cluster (cluster I, Figure 10). Nevertheless, it is still possible to distinguish their metabotypes based on the developed methodology.

4.5 Metabolite patterns specific for non-mitochondrial targets

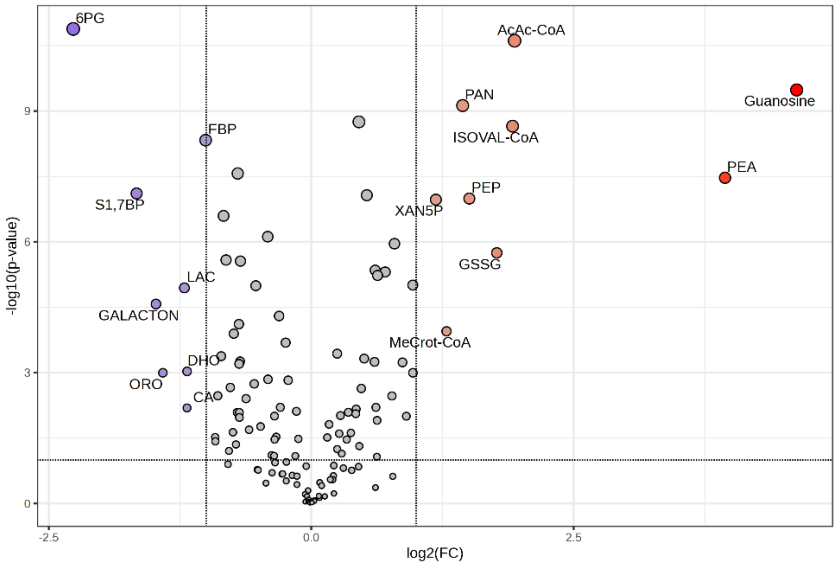
In Figure 10, the similarity cluster II comprises targets, i.e. MoAs that are not directly associated with mitochondrial targets, for instance, microtubules, FAB or hydroxymethylglutaryl CoA reductase (HMG-CoAr). As observed for cluster I, compounds in cluster II are grouped according to their MoAs, with patterns that reflect specific influence on the CCEM (Figure 16).

A.



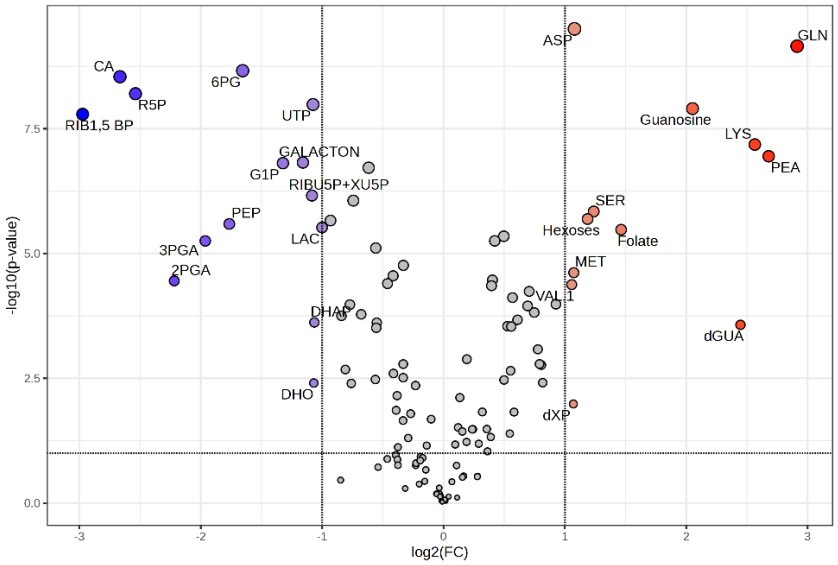
Camptothecin

B.



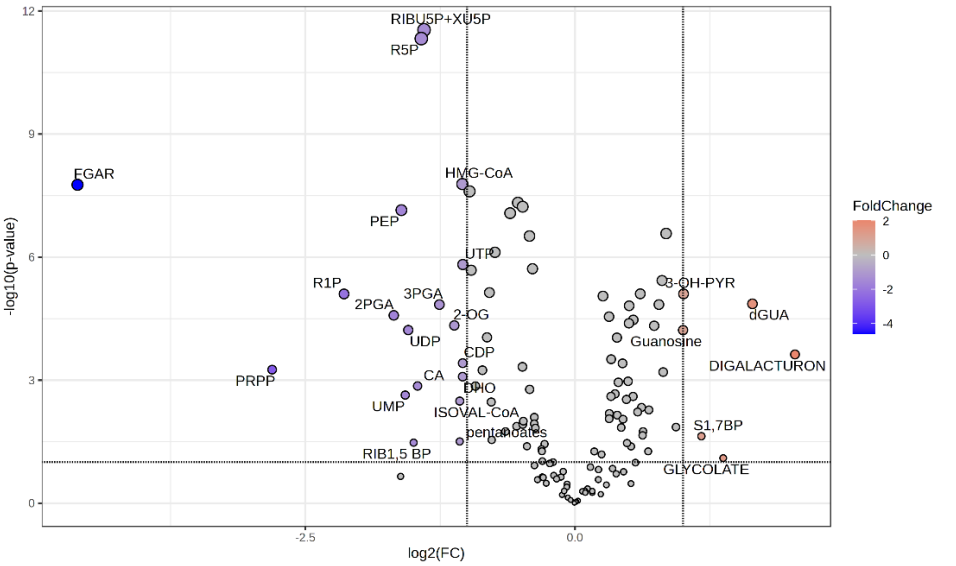
Irinotecan

C.



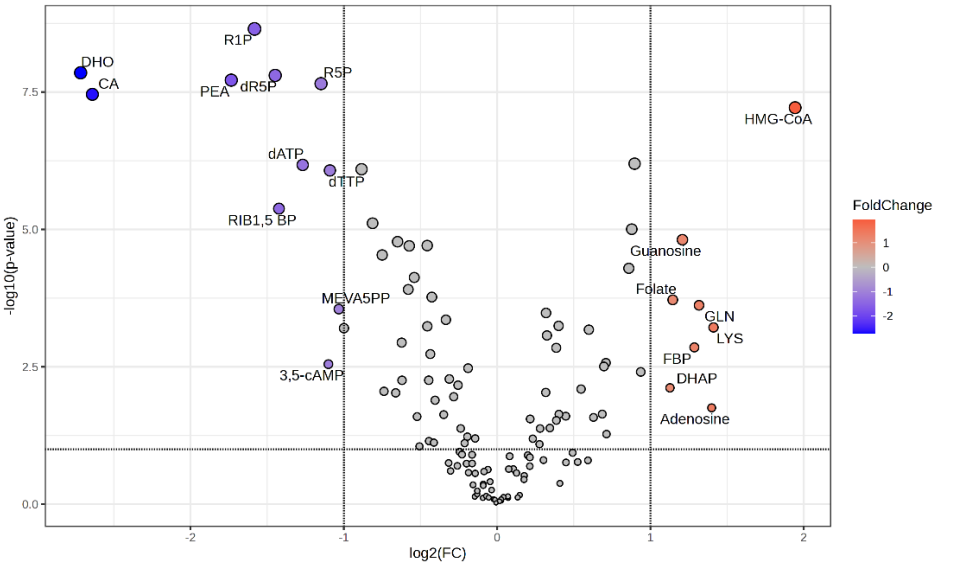
Mitoxantrone

D.



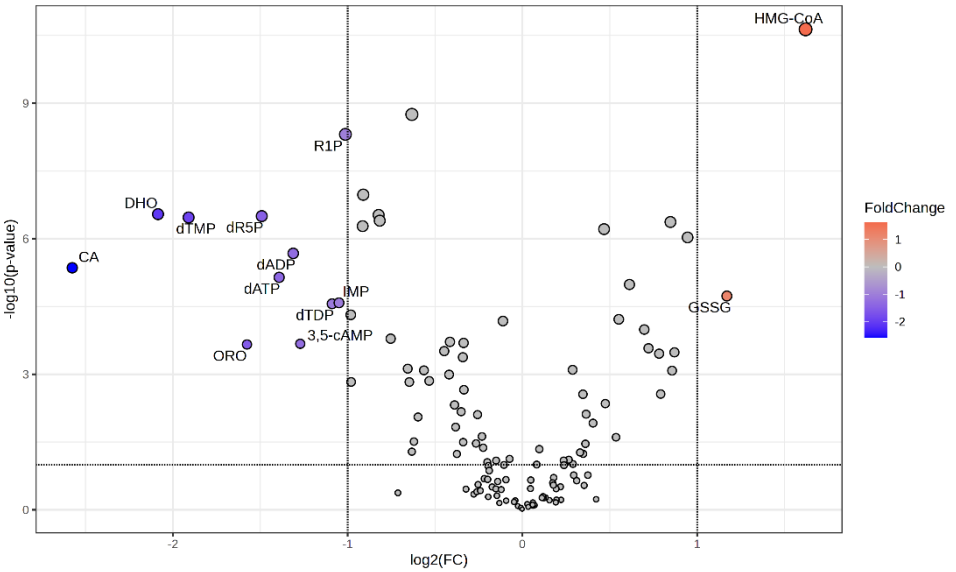
Doxorubicin

E.



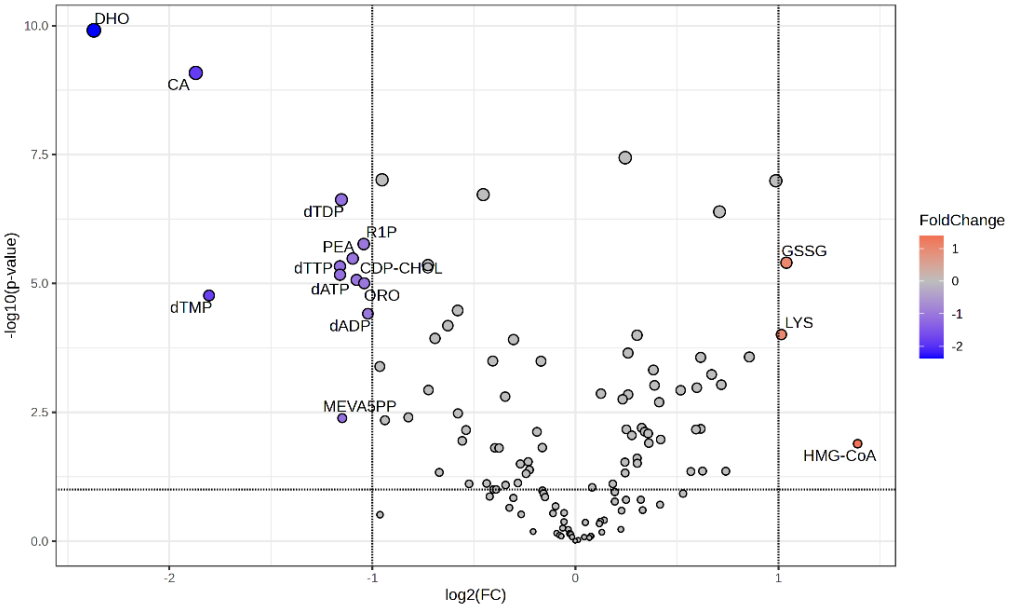
Lovastatin

F.



Fluvastatin

G.



H.

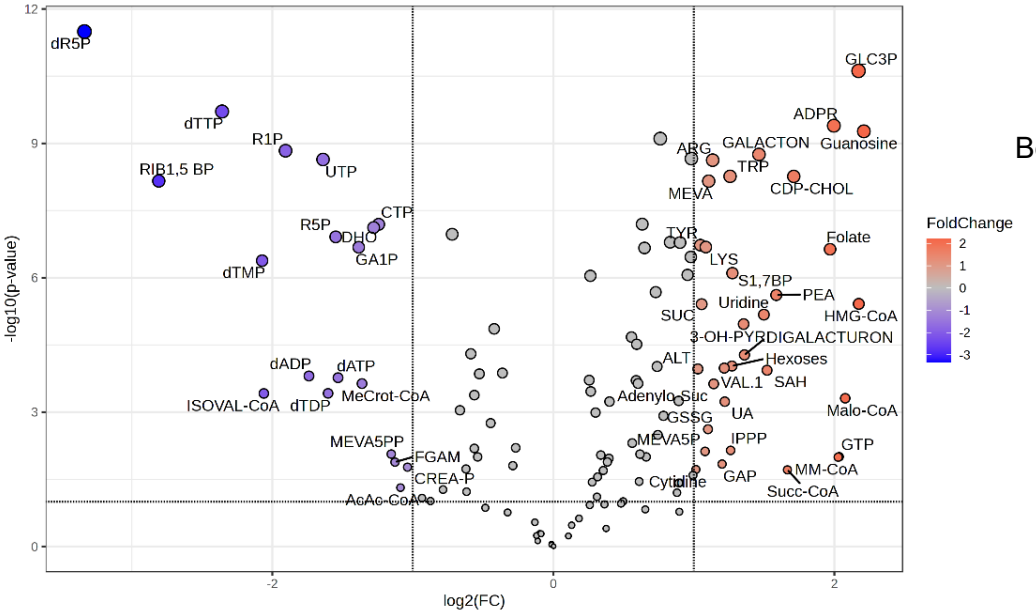


Figure 16: Volcano plots representing relative metabolite abundances (average of n = 6 replicates) after 48 h treatment with selected inhibitors versus non-treated vehicle controls. Accumulation (red), depletion (blue) and non-significant changes (grey) upon treatment with: (A) Camptothecin, (B) Irinotecan (Topoisomerase I inhibitors), (C) Mitoxantrone, (D) Doxorubicin (Topoisomerase II inhibitors), (E) Lovastatin (LOVA), (F) Fluvastatin (FLUV), (G) Atorvastatin (ATOR) (HMG-CoA reductase inhibitors) and (H) Betulinic acid (BETA) (putative PLB inhibitor). A legend for the metabolites' abbreviations is given in Appendix table 1.

Within cluster II, cluster analysis further unveiled two sub-clusters, IIA and IIB. Surprisingly, the metabolic patterns of mTOR inhibitors and AKT inhibitors, both targeting the PI3K/AKT/mTOR pathway, did not cluster in one and the same sub-cluster. Even though both AKT and mTOR inhibitors impact the same signaling pathway, their distinct patterns led us to employ orthogonal partial least square analysis (oPLS-DA) for differentiation. This analysis

effectively distinguishes mTOR inhibitors (rapamycin, wortmannin, alpelisib) from AKT inhibitors (oridonine, perifosine) (Figure 17).

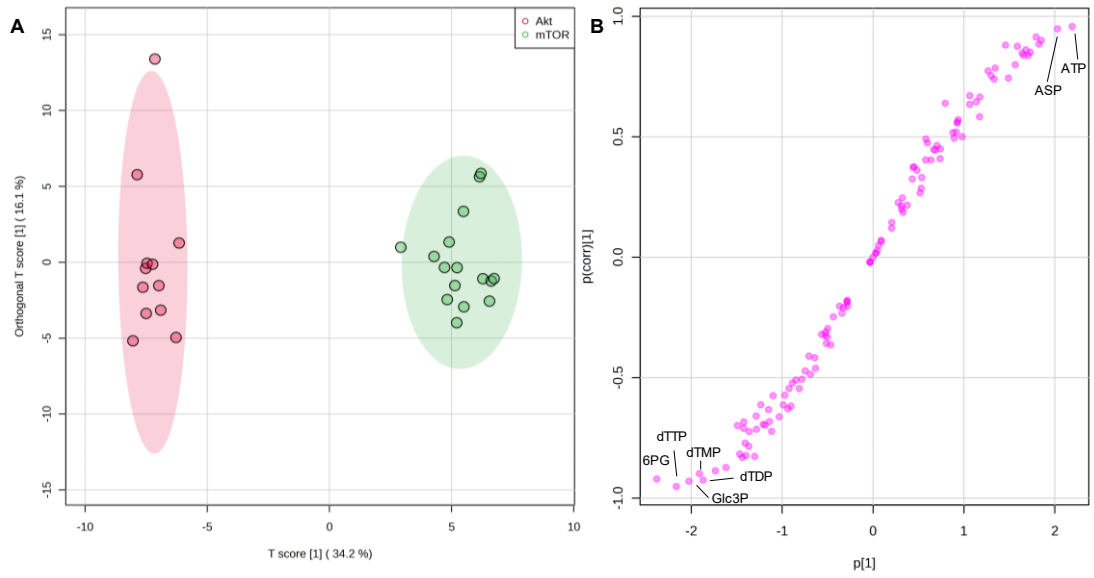


Figure 17: Orthogonal partial least squares (oPLS-DA) discrimination between all AKT inhibitors and all mTOR inhibitors: (A) scores, (B) loadings (importance of variables). For the analysis, all data achieved upon 48 h treatment with wortmannin, rapamycin and alpelisip were grouped as mTOR and compared against all data of perfosine and oridonin, grouped as AKT. Log_2-fold changes vs. non-treated controls were range-scaled.

Notably, the most significant differences between these MoAs arise from lower levels of 6-PG, Glc3P and thymidylates when applying mTOR inhibitors compared to AKT inhibitors. Conversely, mTOR inhibitors exhibit higher levels of ATP and aspartate compared to AKT inhibitors.

With the exception of ETOP, topoisomerase I and II inhibitors induce a specific metabotype in PC-3 cells, where several glycolytic intermediates (2-PGA, 3-PGA, FPB) and PPP intermediates (6-PG, R5P, Ru5P+XU5P and S7P) were depleted compared to the untreated control. Although Jaccard clustering was unable to further distinguish between the two topoisomerase subtypes, we observed an accumulation of several coenzyme A esters with the topoisomerase I inhibitors camptothecin (CMPT) and irinotecan (IRIN), which does not occur when using topoisomerase II inhibitors mitoxantrone (MITO) and doxorubicin (DOXO) (Figure 16, panels A-D).

Furthermore, three statins – lovastatin (LOVA), fluvastatin (FLUV), and atorvastatin (ATOR) – induce a specific metabotype in PC-3 cells with HMG-CoA accumulating and a reduction of mevalonate 5-pyrophosphate (MEVA5PP), which acts downstream of HMG-CoA reductase, as expected (Figure 16, panels E-G). This pattern comprises a variety of yet unknown metabolic consequences, primarily the modulation of the thymidylate synthesis activity (lower

dTMP and dTTP) accompanied by decreased levels of CA, DHO and ORO in the early pyrimidine biosynthesis. Further investigation of this link between mevalonate pathway inhibition and pyrimidine biosynthesis would be crucial, not least since the mevalonate pathway is the source of ubiquinone, an essential molecule for pyrimidine biosynthesis and the electron transport chain.

In summary, in cluster I, co-segregated metabolic patterns strongly suggest mitochondrial dysfunction, distinguishing between OXPHOS inhibition and NAD metabolism inhibition. These differ distinctly from MoA patterns in cluster II, which aid in identifying MoAs affecting non-mitochondrial targets or indirectly influencing metabolism.

4.6 Experimental validation of the MoA prediction approach and transferability to other cancer cell models

To assess the general applicability of the methodology also to other cancer cells beyond the PC-3 prostate cell line, we conducted similar metabotyping experiments with two additional cell lines: MCF-7 hormone receptor-positive breast cancer cells and MHH-ES-1 Ewing's sarcoma cells. Specifically, we examined the effects of four compounds: atorvastatin (ATOV) and lovastatin (LOVA), both inhibitors of HMG-CoA reductase (HMG-CoAr), as well as bithionol (BITN) and hexachlorophorane (HEXA), both inhibitors of glutamate dehydrogenase (GDH).

4.6.1 Metabolite profiling reveals consistent MoA-dependent clustering across cancer cell lines

Utilizing cluster analysis and partial least squares discriminant analysis (PLS-DA) to classify the MoAs, we found that the metabolite profiles after drug treatment exhibited significant similarities across the three cancer cell models. As illustrated in Figure 18, the PLS-DA analysis conducted on PC-3, MCF-7, and MHH-ES-1 cells treated with these compounds demonstrated that the segregation of MoAs by Principal Component 1 (PC1) accounted for 33% of the total variance between the two distinct clusters. Given that this analysis inherently prioritizes maximum clustering along PC1. Hence, this underscores that the primary difference in the metabolic patterns of the treated cells can be attributed to the type of treatment employed and is independent of the specific cell line used in the treatment.

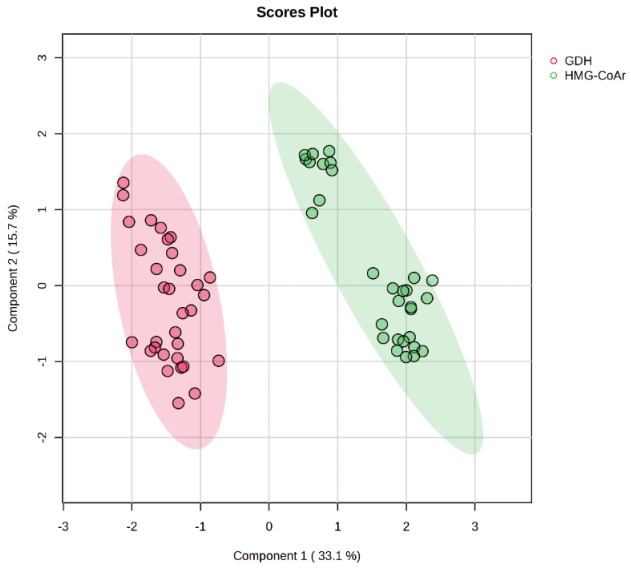


Figure 18: Partial Least Squares Discriminant Analysis (PLS-DA) for the metabotypes of PC-3 cells, MCF-7 and MHH-ES-1 cells. The three cell lines were treated with four compounds which target two pathways in the cells namely: glutamate dehydrogenase (GDH) using hexachlorophorane and bithiniol (red population). And HMG-CoA reductase (HMG-CoAr) inhibitors using lovastatin and the atorvastatin (green population).

Moreover, variable importance in the projection (VIP) metabolites contributing most strongly to the separation of both MoAs, showed consistent regulation independent of the cancer cell model (Figure 19 and Figure 20).

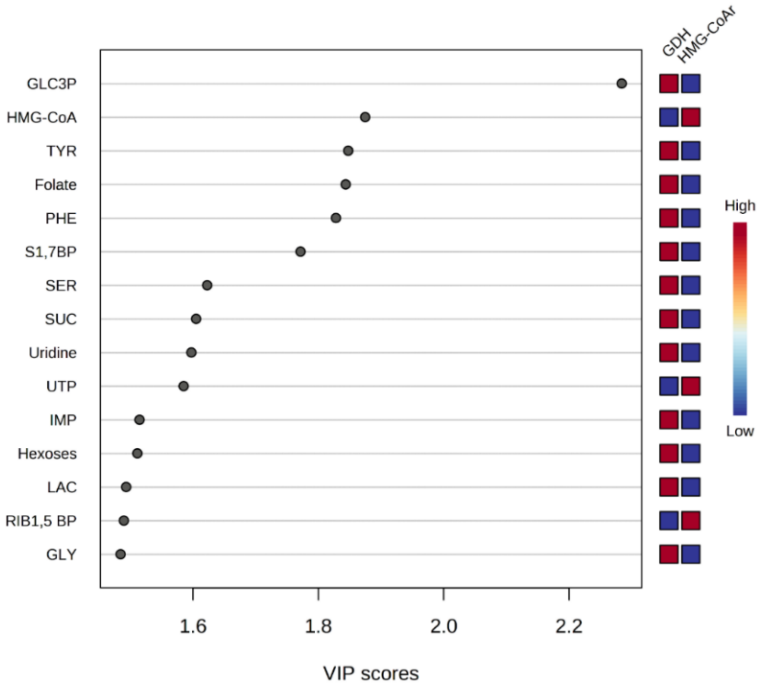


Figure 19: Variable of Importance (VIP) scores depicting the 15 most significant metabolites (VIP > 1.2) contributing to the group separations observed in the PLS-DA model (in Figure 18) differentiating HMG-CoAr inhibition and GDH inhibition across three cancer cell lines (PC-3, MCF-7 and MHH-ES-1). The Analysis was done using MetaboAnalyst 5.0 software.

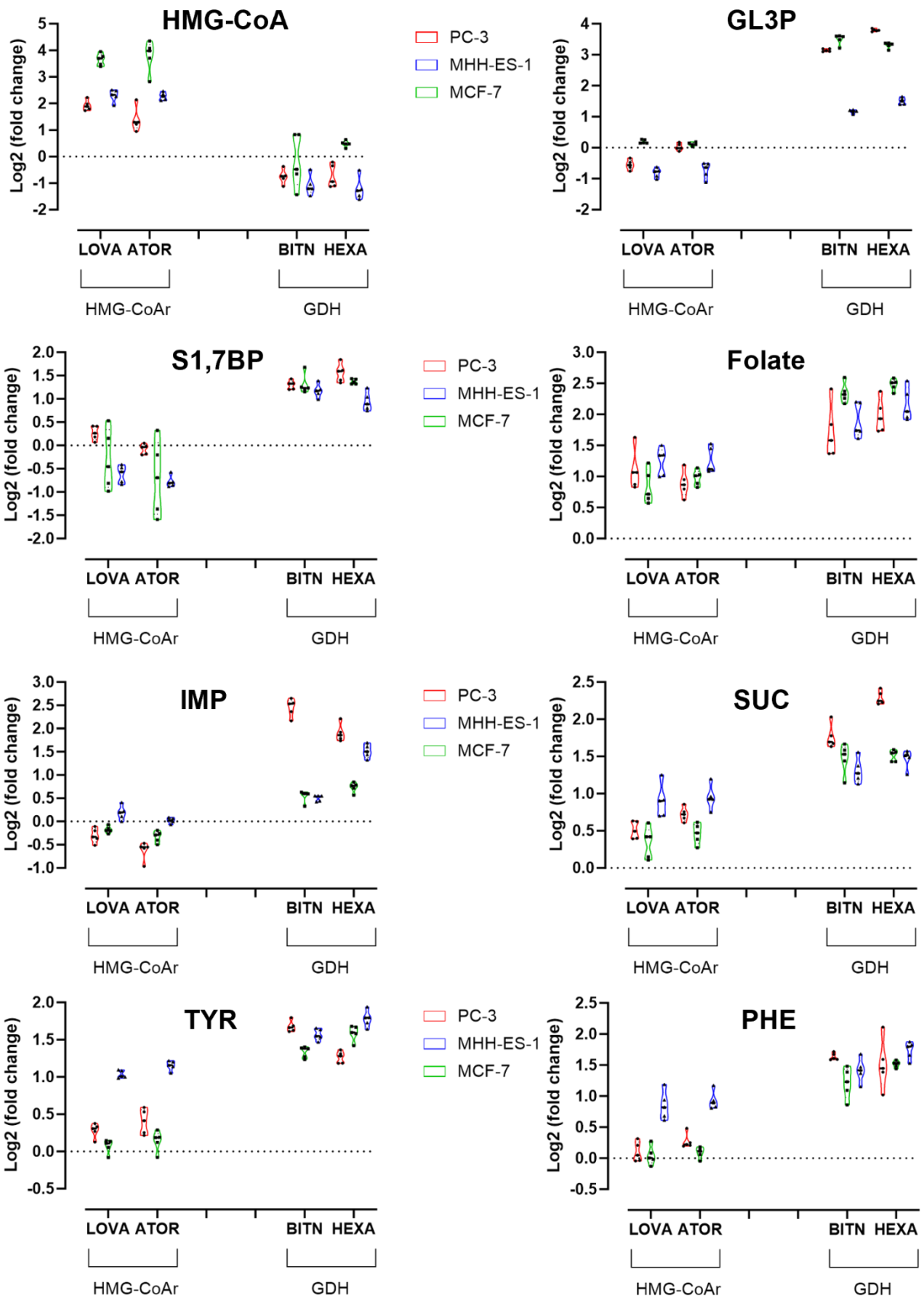


Figure 20: Selected VIP metabolites based on the PLS-DA analysis (in Figure 18) differentiating HMG-CoAr inhibition and GDH inhibition across three cancer cell models (PC-3, MCF-7 and MHH-ES-1).

In addition, Figure 21 presents a heat map showing the normalized data for the 70 most diverse metabolites in PC-3, MCF-7 and MHH-ES-1 cells upon treatment with the previously mentioned compounds. The heatmap shows that the metabolic profiles of the treated cells segregate into two main clusters corresponding to either HMG-CoA reductase inhibition or GDH inhibition, respectively. Furthermore, within each main cluster, sub-clusters attributed to the specific cell lines are discernible, highlighting significant inter-metabolic variations between cell lines both before and after treatment. In addition, significant changes in key pathways were observed in response to treatment, such as the marked accumulation of the metabolite HMG-CoA following inhibition of HMG-CoA reductase by ATOR and LOVA in all three cell lines.

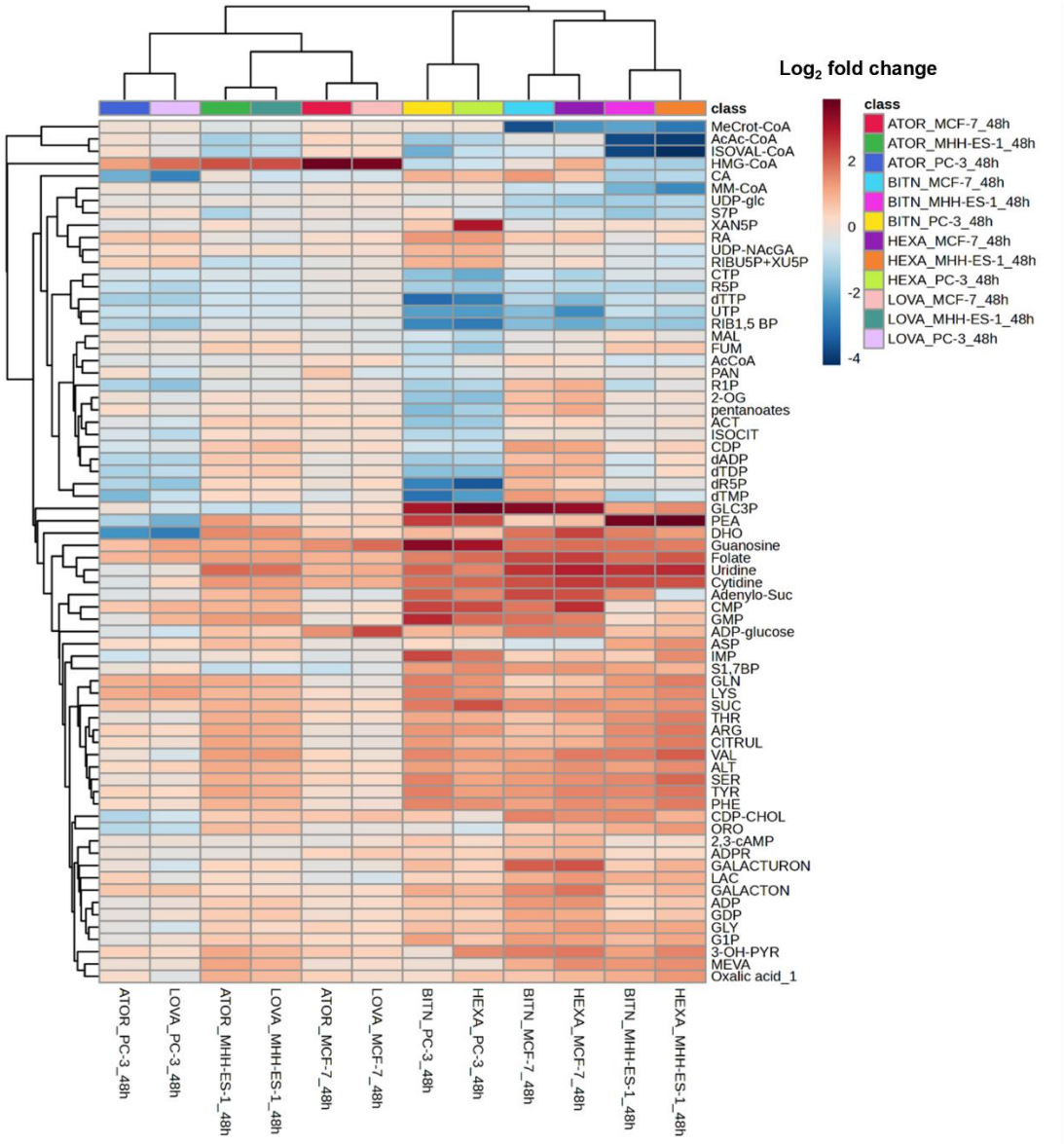


Figure 21: Heatmap illustrating the individual relative intensities of the 70 most varied metabolites in PC-3, MCF-7, and MHH-ES-1 cells. Cells were treated with four compounds targeting two specific pathways: glutamate dehydrogenase (GDH) inhibitors—hexachlorophorane (HEXA) and bithionol (BITN)—and HMG-CoA reductase (HMG-CoAr) inhibitors—lovastatin (LOVA) and atorvastatin (ATOR).

4.6.2 Employing ML models to assess the method transferability across cancer cell lines

We attempted to predict the MoAs GDH and HMG-CoAr in the additional cell lines (MCF-7 and MHH-ES1) by using PC-3 data as a training set, and employing two different machine learning approaches (ML): Lasso and Random Forest. More detailed explanation on the used ML approaches can be found in section 3.7 and 4.7.

The performance of these models showed significant variation, as illustrated in Figure 22, which presents a comparative analysis of the prediction outcomes. Panels A and C display the average prediction scores, representing the mean probability assigned to each MoA across all replicates. While this metric provides a general overview of the model's predictive performance, it may mask underlying variability or inconsistencies within individual replicates.

To provide a more detailed assessment, Panels B and D introduce the metric of the fraction of replicates assigned to each MoA. This approach involves identifying the most likely MoA for each replicate and subsequently calculating the proportion of replicates that were assigned to that specific MoA. The "fraction of replicates" metric delivers a clearer indication of the model's consistency and reliability across different samples, offering valuable insights into the stability of the predictions.

The Lasso-based approach, as depicted in Panels A and B, successfully predicted the correct MoA for three out of the four drugs evaluated. The fraction of replicates metric in Panel B underscores the consistency of these predictions, particularly for HEXA, ATOR, and LOVA, where the model demonstrated strong concordance across replicates. This consistency indicates that the Lasso model not only accurately identified the correct MoA but did so with high reliability across various samples. In contrast, the Random Forest classifier, as shown in Panels C and D, exhibited greater sensitivity to the variability inherent in different cell lines. Although the average prediction scores in Panel C provide a general sense of the model's performance, the fraction of replicates in Panel D reveals a more variable pattern, with a lower proportion of replicates consistently predicting the correct MoA. This suggests that the Random Forest model may be more susceptible to the influence of cell line-specific baseline profiles, leading to less reliable predictions across different cell lines.

In summary, the comparison of these two metrics — average prediction scores and the fraction of replicates — highlights the robustness of the Lasso method. While the average prediction scores provide a broad overview of the model's performance, the fraction of replicates metric offers a more precise evaluation of prediction reliability and consistency, thereby elucidating the strengths and limitations of each approach in this experimental context.

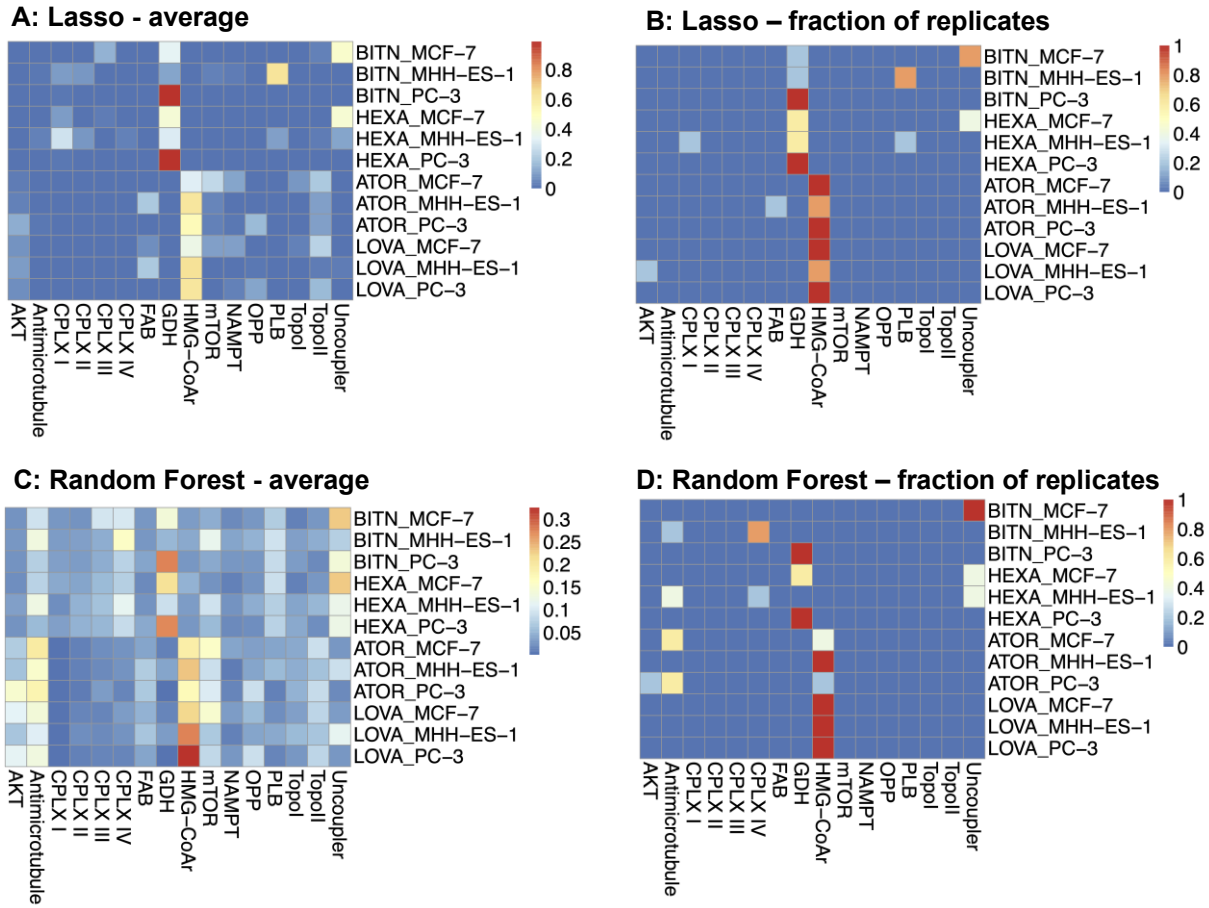


Figure 22: Prediction results across cell models for the Lasso (A-B) and Random Forest (C-D) methods. Panels A and C show average prediction, while Panels B and D show the fraction of replicates assigned to a specific MoA. While the Random Forest classifier widely fails to yield clear prediction scores and classifies several drugs incorrectly, the Lasso method only fails for BITN, but successfully predicts the correct MoA of HEXA, ATOR and LOVA for the majority of replicates in the MCF-7 and MHH-ES-1 cell models. Machine learning was done by Dr. Jan Grau and Prof. Ivo Grosse, Institute of Computer Science, Martin Luther University Halle-Wittenberg, Halle (Saale)

4.7 Evaluation of different machine learning (ML) approaches to predict metabolic patterns **(this chapter work was done in cooperation with Dr. Jan Grau and Prof. Ivo Grosse, MLU, Halle (Saale))**

We apply ML with the goal of predicting the MoA of new drugs based on their normalized CCEM profile. Since training data (i.e., CCEM profiles of drugs with known MoA) are of high quality but rather sparse (2-3 drugs per MoA), we needed to resort to methods that can accomplish classification under these constraints. Specifically, we employed Random Forests for classification as they provide reliable results with sparse training data. Additionally, Lasso regression was used to explain an unclassified CCEM profile based on a limited selection of classified profiles in the training set. Random Forests are specifically designed for classification problems and aim to avoid overfitting by repeated sampling from the variables considered in decision trees and bootstrapping on the training samples. The Lasso-based

approach fits coefficients for the CCEM profiles of all drugs in the training set to explain the CCEM profile of a test drug (response variable), while regularization yields coefficients different from zero only for those training CCEM profiles that contribute to the fit substantially. The MoAs of training CCEM profiles that obtain large coefficients in the fit and, hence, contribute the most to explaining the test CCEM profile are considered for predicting the MoA of the test drug. As different drugs of the training set may obtain coefficients different from zero, the Lasso-based approach is expected to handle intermediate/mixed MoAs more appropriately. To establish a baseline, we also incorporated a straightforward k-nearest neighbor classifier (Figure 23-A). Limitations of the quantity of high-quality training data available render a partitioning into dedicated training and test sets for evaluating the alternative approaches unreasonable. To establish a performance evaluation across all drugs and associated MoAs, nonetheless, we follow a leave-one-out cross validation strategy. Here, the data are partitioned such that all replicates of one drug (test partition) excluded from the training set, the respective method is trained on the CCEM profiles of the remaining drugs in the training set, and the MoA of the individual replicates of the left-out drug is predicted. By excluding all repliates of the test drug from the training set, we avoid data leakage in the same manner as for dedicated training and test sets. Initially, prediction accuracy (classification rates) was averaged over all drugs sharing the same MoA and then across all MoAs. We found that the baseline model (k-nearest neighbor) achieved a mean accuracy of just 0.655, while the Lasso-based approach and the random forest yielded substantially improved mean accuracy values of 0.854 and 0.864, respectively (Figure 23-A). Hence, we consider the Lasso-based and the random forest approaches similarly suited for predicting MoAs of novel drug candidates.

Subsequently, the predictions of both approaches were evaluated on the level of MoAs to achieve a more fine-grained picture of the prediction performance. We found notable differences as to the prediction accuracy of individual MoAs (Figure 23,B-C). Certain MoAs, such as CPLX I, II and IV, GDH, HMG-CoAr and NAMPT, can be well predicted using either approach. Other MoAs (OPP, TopI, TopoII, uncoupler) show reduced prediction performance by either approach, although to a varying degree. Finally, antimicrotubule, CPLX II and PLB inhibitors can be classified almost perfectly by only one of the two approaches (Figure 23,B-C). However, misclassification can be considered from two perspectives. First, replicate CCEM profiles of drugs from a certain MoA (e.g., TopoII) may be classified into multiple different MoAs. Second, some MoAs (e.g., AKT with Lasso, CPLX IV with random forest) co-classify with other MoAs, which might be explained by the lack of unique metabolic patterns that allow discrimination between these MoAs. Finally, not all effects of a training inhibitor

("reference drug") may be known yet for the cell line used, which could explain reduced predictability.

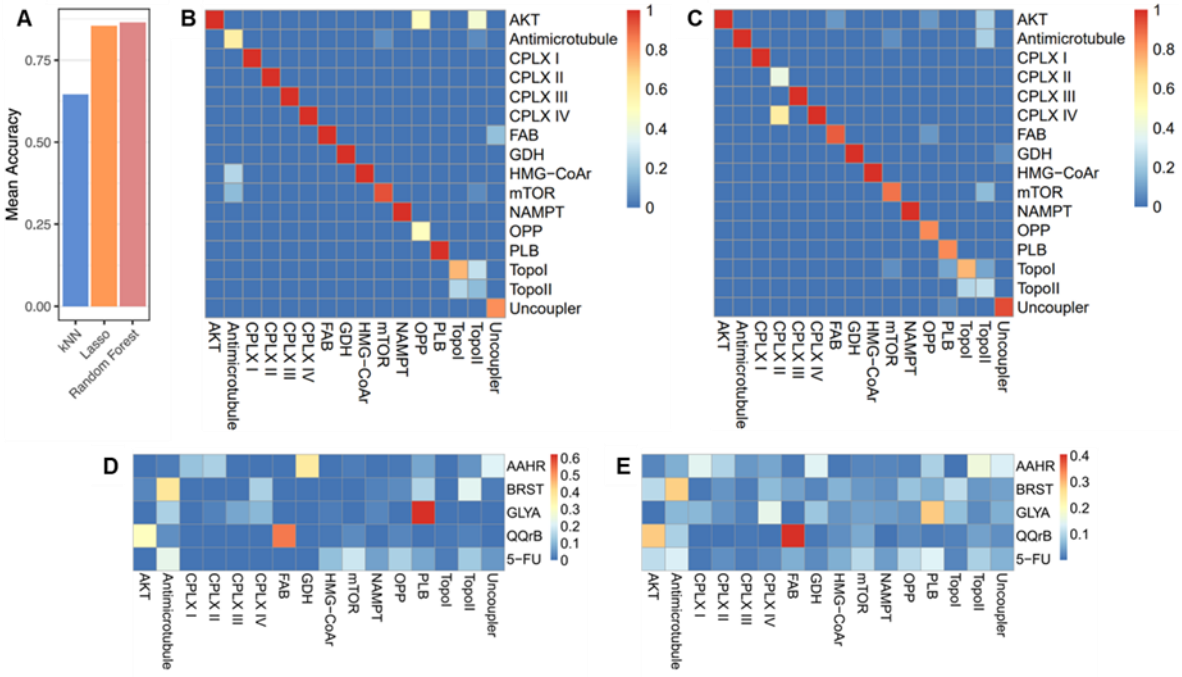


Figure 23: Performance evaluation and MoA prediction of machine learning approaches. (A) Mean accuracy of k-nearest neighbor classifier (kNN), Lasso-based approach (Lasso) and Random Forest classifier (Random Forest) in leave-one-out cross validation over the drugs in the training set. (B-C) Mean accuracy per MoA, where entries on the main diagonal correspond to correct classification in the training data and off-diagonal elements indicate predictions of alternative MoA (row), B-Lasso, C-Random Forest. (D-E) Normalized prediction scores for predicted MoAs of four cytotoxic compounds obtained by D-Lasso-, and E-Random Forest classifier. Machine learning was done by Dr. Jan Grau and Prof.Ivo Grosse, Institute of Computer Science, Martin Luther University Halle-Wittenberg, Halle (Saale)

MoA abbreviations are listed in Fig. 5, GLYA – glycyrrhetinic acid, AAHR – asiatic acid homopiperazinyl rhodamine B conjugate, BRST – Breastin (patented Nerium oleander cold water extract), QQrB – cucurbitacin B, 5-FU – 5-flurouracile.

4.8 ML predicts MoAs of novel cytotoxic compounds with anti-cancer potential

Next, both ML approaches were applied to predict the MoAs of one semi-synthetic and three plant-derived drug candidates that were previously phenotypically identified as potential anti-cancer agents by our groups (Table 1). The novel compound AAHR, composed of triterpene asiatic acid and rhodamine B, has demonstrated robust cytotoxicity, and the capacity to overcome drug resistance in human preclinical tumor models [219, 220]. In contrast to glycyrrhetinic acid (GLYA, a pure pentacyclic triterpenoid acid derived from licorice roots), AAHR exhibits a higher toxicity towards PC-3 prostate cancer cells (Table S1). Breastin (BRST) is a defined cold-water extract from leaves of *Nerium oleander*. Phytochemical characterization revealed several monoglycosidic cardenolides as major constituents [221]. Cucurbitacin B (QQrB) is a highly toxic tetracyclic triterpene compound derived from Cucurbitaceae plants (Table 1).

4.8.1 Example 1: The triterpene-rhodamine conjugate AAHR is a potent OXPHOS inhibitor

ML approaches suggest that AAHR inhibits multiple targets, of which CPLX I-III, uncoupler and GDH were indicative to the inhibition of mitochondrial processes (Figure 23D-E). Given the lack of existing literature on AAHR and its MoA, we conducted wet lab experiments to validate these predictions. Firstly, we employed confocal laser scanning microscopy (CLSM) to investigate the localization of AAHR within the cell. Imaging results revealed that AAHR localizes within the mitochondrial membrane in PC-3 cells, as indicated by a 86% fluorescence co-localization with MitoTracker™ Green (Figure 24).

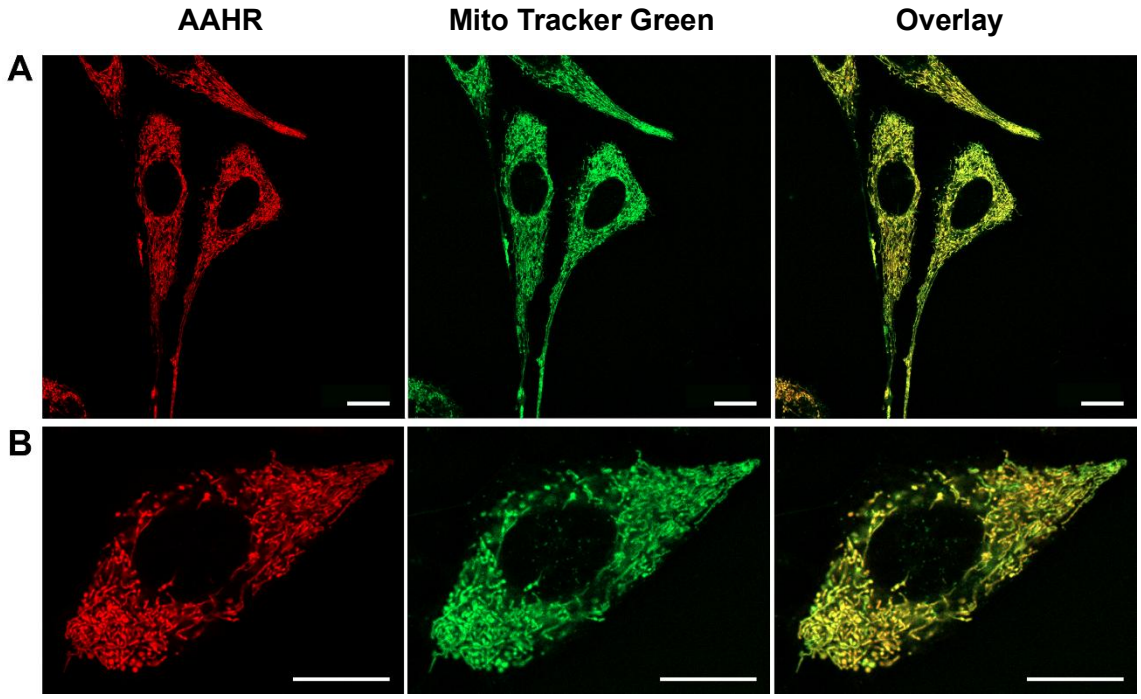


Figure 24: AAHR accumulation in the mitochondria of PC-3 cells.

PC-3 cells were treated with AAHR for 4 h followed by a counterstaining with MitoTracker™ Green and analyzed by confocal laser scanning microscopy. (A) Overview of few cells; (B) close-up of a single cell. Fluorescence of AAHR and MitoTracker™ Green is depicted in red and green, respectively. The overlay shows the co-localization of both AAHR and MitoTracker™ Green. Pearson's Correlation Coefficient was calculated using eight micrographs and revealed 0.8558±0.058. Scale bars represent 20 µm.

Following the imaging studies, we performed a series of biochemical assays, including lipid profiling, Seahorse cell analysis, and CoQ (coenzyme Q) measurements, to further elucidate the effects of AAHR on mitochondrial function. Lipid profiling demonstrated a significant accumulation of acylcarnitines. Seahorse analysis showed a marked reduction in mitochondrial respiration and ATP production, while CoQ measurements revealed a depletion of cellular CoQ levels (Figure 25A, Figure 27 and Figure 29).

To pinpoint the exact mechanism by which AAHR inhibits mitochondrial respiration, we modified the Seahorse assay by substituting traditional OXPHOS inhibitors with AAHR (Figure 25-B) (Seahorse assay was done by Dr. Thomas Müller, UKH, Halle). Analysis of the oxygen consumption rate (OCR) and extracellular acidification rate (ECAR) patterns indicated that AAHR does not function as an uncoupling agent or complex V inhibitor. Instead, the data suggests that AAHR inhibits mitochondrial respiration by targeting CPLX I-III and/or through an indirect mechanism affecting overall cellular respiration.

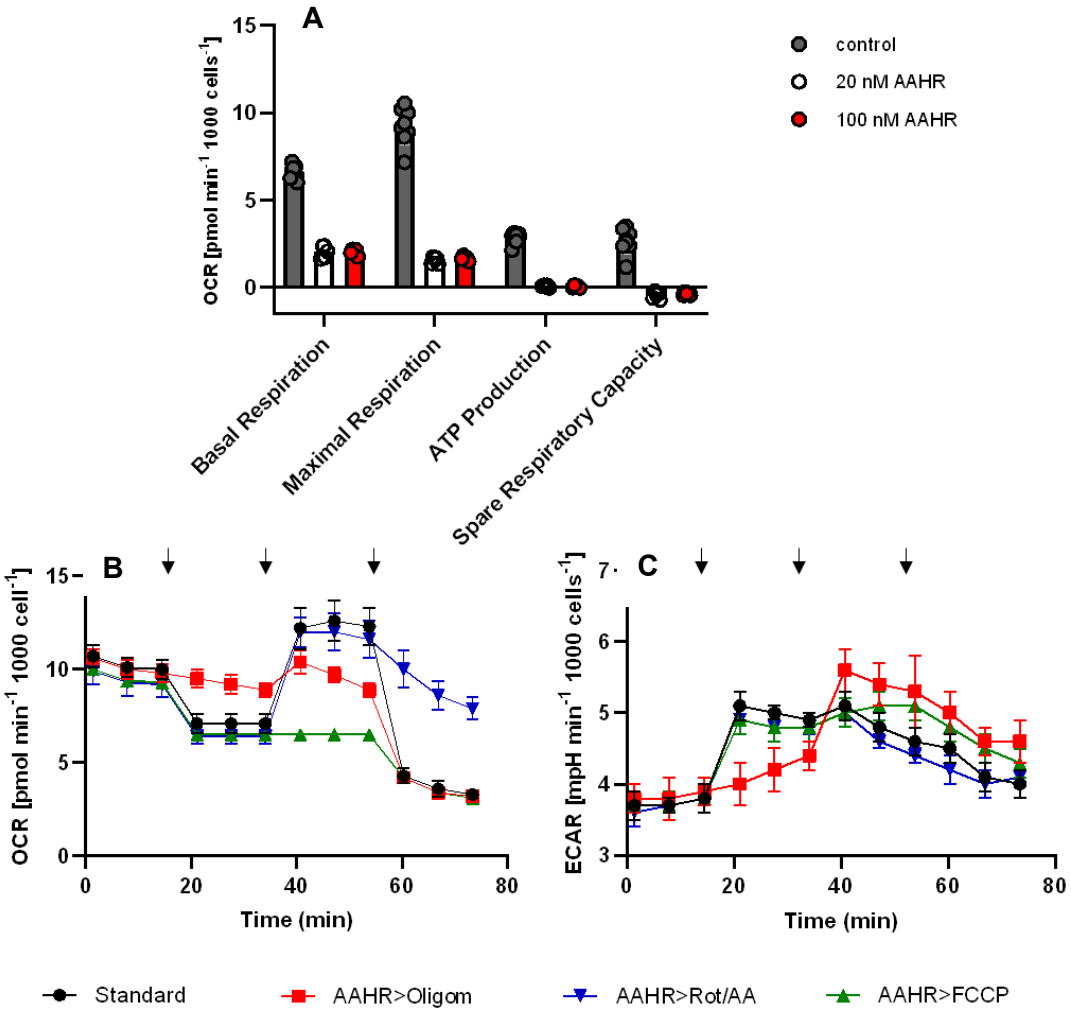


Figure 25: Analyses of mitochondrial function and ATP production in PC-3 cells by performing Seahorse assays.

(A) Parameters obtained from a Seahorse XF Cell Mito Stress Test following a 24 h treatment with AAHR. Basal respiration: oxygen consumption rate (OCR) used to investigate cellular ATP demand resulting from mitochondrial respiration at baseline; Maximal respiration: OCR attained by stimulating the respiratory chain to operate at maximum capacity, shows the maximum rate of respiration the cell can achieve. Spare respiratory capacity: indicates the capability of the cell to respond to an energetic demand and how closely the cell is to respiring to its theoretical maximum. ATP production: oxygen consumption linked cellular ATP production rate. Seahorse assay was done by Dr. Thomas Müller, Medical Faculty, University Clinic for Internal Medicine IV (Hematology/Oncology, Martin Luther University Halle-Wittenberg, Halle (Saale).

(B) Patterns of OCR, and (C) ECAR (extracellular acidification rate) during the Seahorse XF Cell Mito Stress Test. Arrows indicate the sequential substance injection. In a standard assay, the order of injection is: 1. oligomycin, 2. FCCP, 3. rotenone/antimycin A. In the modified assays, one of the substances was replaced by AAHR, respectively.

4.8.2 Example 2: Pentacyclic triterpenes are potential modulators of PLB

Both ML models consistently predicted that the MoA of glycyrrhetinic acid (GLYA) involves modulation of phospholipid biosynthesis (PLB), as evidenced in (Figure 23D-E). This prediction is supported by the observed accumulation of CDP-choline and phosphoethanolamine (PEA), a pattern also seen with other pentacyclic triterpenes, such as betulinic acid (BETA) (Figure 16, panel H). While CDP-choline accumulation was not universally observed across all training compounds nor specifically with GLYA treatments, the overall metabolic signature indicative of PLB modulation remains characteristic.

The accumulation of CDP-choline typically occurs when enzymes such as choline phosphotransferase (CPT1, EC 2.7.8.1) or choline/ethanolamine phosphotransferase (CEPT1, EC 2.7.8.2) are inhibited in the Kennedy pathway. To explore this hypothesis further, we conducted in silico docking studies of pentacyclic triterpenes and the natural substrate CDP-choline on both CEPT1 and CPT1. The results suggest that pentacyclic triterpenoids could act as strong competitive inhibitors of these enzymes (Figure 26, Table 2). Docking study was done by MSc. Mohammad Yousefi and Dr. Mehdi Davari, IPB, Halle (Saale).

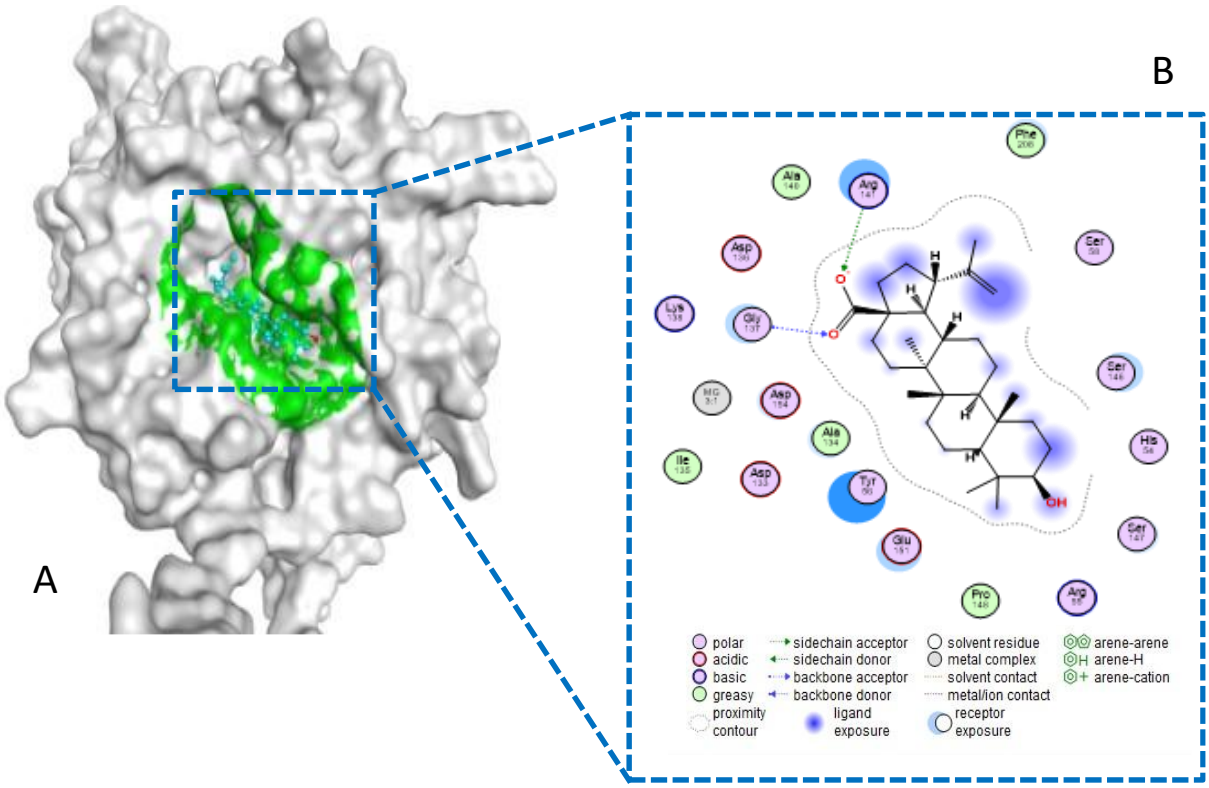


Figure 26: Molecular docking of pentacyclic triterpenes with CEPT1 and CPT1. (A) Molecular surface of CEPT1 with docked BETA as ligand (cyan). The binding site is represented in green. (B) 2D-interactions showing that the carboxyl group of acid-based pentacyclic diterpenes exhibits favorable hydrogen bond interactions, while the aromatic part has hydrophobic interactions with non-polar residues. Docking study was done by MSc. Mohammad Yousefi and Dr. Mehdi Davari, Leibniz Institute of Plant Biochemistry, Dept. of Bioorganic Chemistry, Halle (Saale).

Table 2: Binding affinity of triterpenoid ligands to the CEPT1 and CPT1 enzymes, in comparison to the native substrate CDP-choline. Binding energies were calculated based on the model shown in Figure. 26

	[kcal mol^{-1}]	
	CEPT1	**CPT1**
Betulinic acid	-8.33	-7.31
11-Keto-beta-boswellic acid	-7.3	-6.17
Glycyrrhetinic acid	-7.6	-7.68
Maslinic acid	-7.4	-6.13
Ursolic acid	-8.66	-7.44
CDP-Choline	-7.14	-8.33

Given this potential inhibitory effect, we hypothesized that treatment with pentacyclic triterpenoids, including AAHR, would lead to reduced cellular concentrations of intact phosphatidylcholines (PC). To investigate this, we performed lipidomics analysis. Contrary to our expectations of a general reduction in PC levels, we observed individually altered PC levels, indicating significant changes in the global membrane lipid composition following treatment with BETA and AAHR (Figure 27). Among the most affected lipid groups were PC, phosphatidylethanolamine (PE), and sphingomyelins (SM), which are derived from PC metabolism. Additionally, BETA and AAHR treatments led to the formation of lyso-phosphatidylethanolamines (LPE) and lyso-phosphatidylcholines (LPC), respectively. These alterations suggest membrane destabilization and the exposure of apoptosis-inducing "eat-me signals" [222].

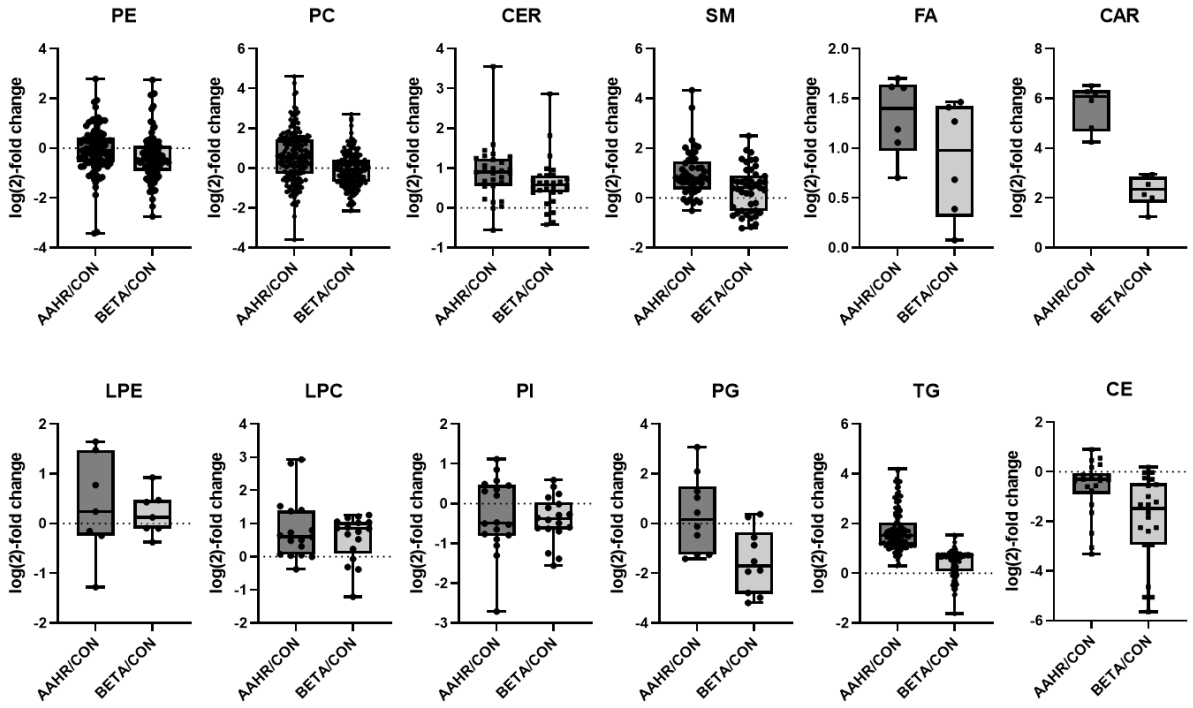


Figure 27: The ratio of lipid levels of PC-3 cells treated with AAHR (285 nM) or BETA (20 µM) for 48 h versus DMSO controls. Each dot represents an individual lipid. All values are normalized to cell number and represent the average of five independent samples per annotated lipid.

PE – diacylglycerophosphoethanolamines, PC – diacylglycerophosphocholines, CER – ceramide, SM – sphingomyolin, FA – free fatty acids, CAR – acylcarnitines, LPE – monoacylglycerophosphoethanolamines, LPC - monoacylglycerophosphocholines, PI – phosphoinositides, PG – phosphoglycerates, TG – triacylglycerides, CE – cholesterol esters.

To further corroborate that triterpenoic acids inhibit PLB at the enzymatic step of CPT1, we labeled PC3 by replacing glucose in the cell culture medium with $U^{13}C$-glucose and concomitantly added BETA in the treatments over 48 hours, whereas in controls no betulinic

acid was added. Fractional enrichment clearly demonstrates that less ^{13}C carbon is incorporated into the three most abundant PC in the presence of BETA (Figure 28).

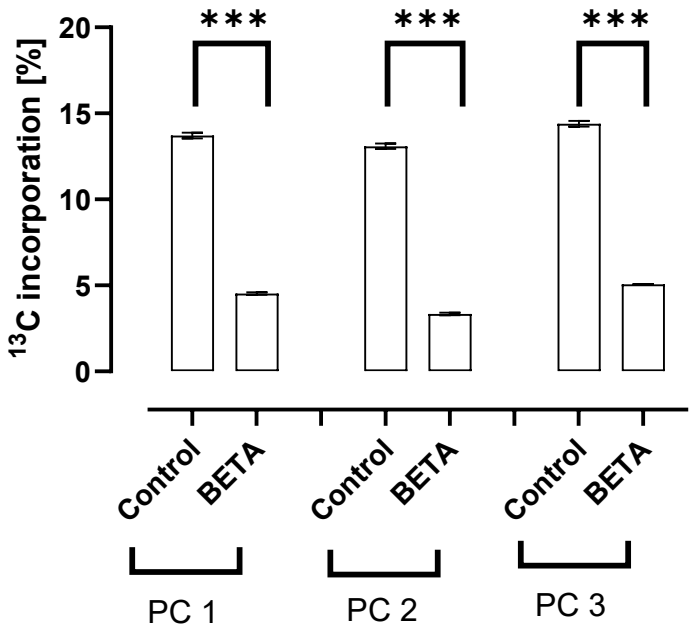


Figure 28: Comparison of the fractional 13C incorporation in the three most abundant phosphatidylcholines after 48 h treatment with betulinic acid versus untreated controls. PC1: PC 34:1(PC 16:0,18:1); PC2: PC 36:1(PC 18:0,18:1); PC3: PC 32:1(PC 16:0,16:1). Signal intensities are normalized to the cell count. Error bars represent S.D. Significance level 0.001 is based on pairwise t-test with homoscedastic data distribution (n=3).

Considering that mitochondrial membranes account for a significant proportion of cellular PC and PE, changes in the composition of these lipid groups could be directly related to mitochondrial dysfunction [223]. This is further supported by the observed profiles of free fatty acids (FA), acylcarnitines (CAR), and triacylglycerides (TG), all of which were elevated following treatment with BETA and AAHR. These compounds play crucial roles in mitochondrial energy metabolism, and their accumulation points to impaired OXPHOS [224]. A further indication of OXPHOS impairment is the marked decrease in coenzyme Q (CoQ) levels in PC-3 cells after BETA treatment (Figure 29).

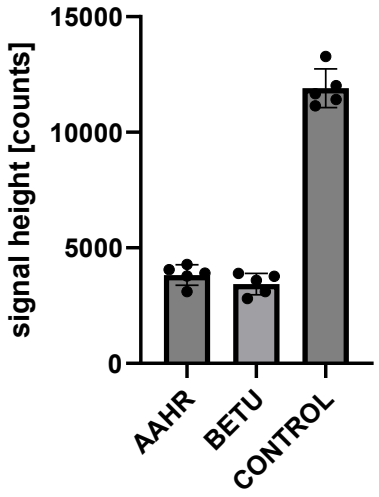


Figure 29: MS analyses of cellular coenzyme Q levels after 48 h treatment with AAHR (285 nM), BETA (20 µM) and DMSO (control).

However, it is important to note that the ML models also predict additional MoAs beyond PLB modulation for both GLYA and AAHR (Figure 23D-E). For AAHR, OXPHOS inhibition clearly predominates among ML-based model predictions, whereas both models predict OXPHOS inhibition as a side effect for pure pentacyclic triterpenoic acids (Figure 23D-E).

4.8.3 Example 3: Breastin impairs microtubule formation

BRST was isolated as cold extract from the leaves of *Nerium Oleander*. It contains defined amounts of cardenolide glycosides besides flavonoids and polysaccharides. The extraction and bioactivity evaluation including the anticancer properties of BRST were conducted by Rashan et al. [225]. In their work, Rashan conducted a COMPARE analysis involving 153 anticancer agents with well-established MoA across 74 tumor cell lines from the Oncotest panel. This analysis unveiled consistent correlations between BRST and drugs that inhibit mitosis and cause DNA damage. Consequently, they proceeded to investigate BRST's impact on microtubules using U2OS cells transfected with tubulin-GFP. Conducting confocal microscopy, they were able to confirm that BRST-induced disruption of microtubules exhibited a similar fashion to that of the tubulin-depolymerizing drug paclitaxel (Figure 30) [225]. Besides, they verified these findings by an *in vitro* tubulin polymerization assay and molecular docking of several ingredient compounds of the BRST extract. Hence, we applied BRST as prediction set with unknown MoA and we compared its metabotype on PC-3 cells to our training set (Figure 23D-E). Both ML approaches agree and assign the largest scores to the antimicrotubule MoA with further considerable scores for CPLX IV, PLB and TopoII in the case of Lasso and small but non-zero scores for multiple MoAs in the case of the Random Forest (Figure 23D-E), which is in agreement with the previous finding of Rashan and coworkers.

This finding validates our approach as a reliable method for MoA investigation in drug discovery, allowing for targeted research and reducing the need for extensive testing.

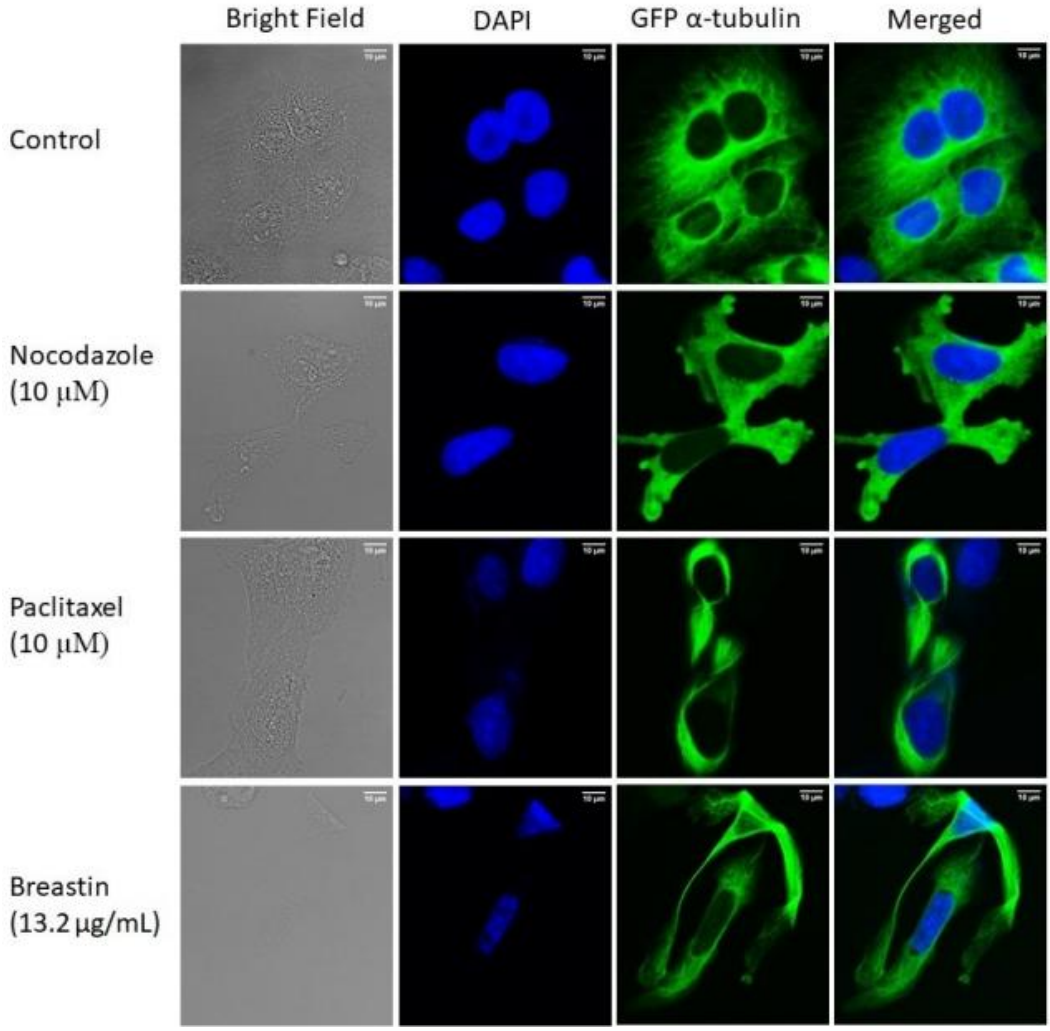


Figure 30: This figure, copied from Rashan et al. publication [225], shows the influence of BRST on microtubules in U2OS human osteosarcoma cells. U2OS expressing α-tubulin-GFP were subjected to a 24 h treatment with BRST to assess its impact on microtubules. As positive controls, paclitaxel and nocodazole were used. Control cells exhibited robust tubulin polymerization, forming a dense intracellular network. In contrast, cells treated with BRST displayed a stiff tubulin network resembling the effects seen in paclitaxel-treated cells. Nocodazole, on the other hand, caused extensive disintegration of the tubulin network. These results provide evidence that, similar to paclitaxel, BRST arrests tubulin polymerization.

4.8.4 Examples 4: Cucurbitacin B inhibits lipogenesis

Both ML models assigned cucurbitacin B (QQrB) with the highest score to the MoA of FAB inhibitors (Figure 23D-E). Cancer cells have various ways to maintain their fatty acid levels, including increased *de novo* FAB [226]. ATP citrate lyase (ACLY, EC 2.3.3.8), acetyl-CoA carboxylase (ACC, EC 6.4.1.2), and fatty acid synthase (FASN, EC 2.3.1.85) are the key enzymes responsible for lipogenesis via FAB and are targeted by anti-proliferative therapies [227]. FAB inhibitors such as epigallocatechin gallate (EGCG) and apigenin (APIG) impede

lipogenesis at the FASN stage [228, 203]. Similarly, QQrB inhibits ACLY, which produces a comparable metabolite profile to EGCG and APIG [229].

4.8.5 Example 5: The MoA of 5-fluorouracil (5-FU) can not be predicted

The method does face challenges, particularly in covering the full spectrum of potential MoAs for novel anticancer agents, raising questions about the applicability of the 16 MoAs considered here to new drugs. While CCEM-based method certainly cannot predict all possible MoAs based on the 16 MoAs included in our training set, it is designed to avoid generating misleading results when applied to compounds with MoAs not represented in the training data. This approach ensures that, in the absence of relevant training examples, the method refrains from making incorrect predictions. For example, 5-FU, a known TYMS inhibitor, exhibits a distinct response pattern characterized by thymidinylate depletion and elevated folate levels (Figure 31), yet this MoA was not included in the training process. By design, classification methods cannot predict classes that are absent from the training set. While the Lasso-based approach can interpolate between known MoAs within the training data, but nonetheless cannot predict MoAs beyond the training space. As a result, neither the Lasso nor the Random Forest methods accurately predicted the MoA of 5-FU. However, the low prediction scores observed for several MoAs (Figure 23 D/E), allow to identify such cases based on the prediction outcomes for further extension and improvement of our methodology.

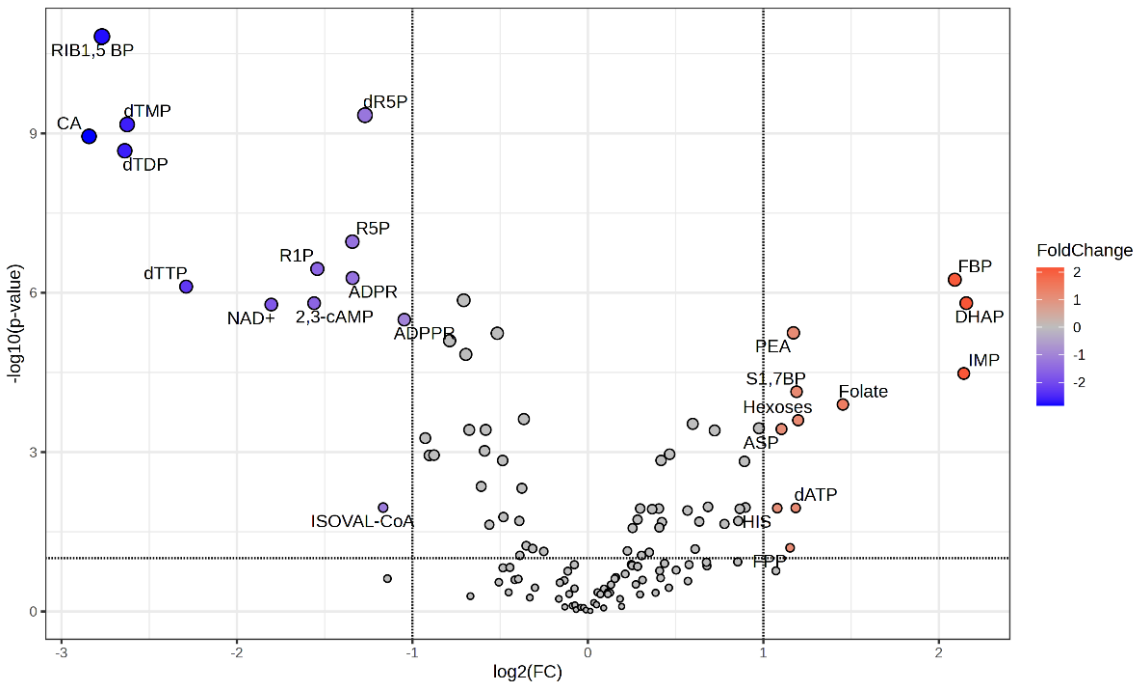


Figure 31: Volcano plots representing relative metabolite abundances (average of n = 6 replicates) after 48 h treatment with the TYMS inhibitor 5-fluorouracil (5-FU) versus non-treated vehicle controls, represented with accumulation (red), depletion (blue) and non-significant changes (grey) following the 48 h. A legend for the metabolites' abbreviations can be found in Appendix table 1.

5 Discussion

Metabolomics is the global analysis of small molecule metabolites that can provide critical information on the turntable of metabolic processes: the central carbon and energy metabolism (CCEM). Since the discovery of increased glycolytic flux by Warburg, it became clear that massive reprogramming of cellular energy metabolism is one of the hallmarks of cancer. Since this engages metabolic responses, metabolite profiling not only allowed to disclose cancer-specific phenotypes (metabotypes) but also promoted the discovery of oncometabolites such as 2-hydroxyglutarate or succinate, which accumulate in certain tumors [64].

More than 60% of chemotherapeutics are derived from natural products, which represent a vast repository of small molecules capable of targeting the energy metabolism of cancer cells [68]. The "Anticancer Agents and Targeting" research group within the Department of Bioorganic Chemistry at the Leibniz Institute of Plant Biochemistry, focuses on the identification, characterization, and synthesis of bioactive compounds from an extensive collection of natural and synthetic sources. Utilizing phenotypic screening as a cornerstone of our drug discovery process, we frequently identify promising compounds exhibiting a range of biological activities. To enhance the efficiency of our drug discovery pipeline, we are in pursuit of a rapid MoA detection method. Such a method would not only accelerate the process of drug development but also refine our approach by directing studies towards targeted investigations.

Understanding the MoA is crucial in drug discovery, as it reveals the biochemical interactions through which a drug exerts its effects. Although the FDA does not mandate MoA knowledge for approval, having this information enhances drug efficacy, safety, and the ability to predict potential side effects [230]. This understanding can also streamline the approval process, potentially reducing the need for more extensive clinical trials. One promising approach to address the challenge of identifying MoAs is through metabolomics, which offers insights into cellular responses to drugs by analyzing changes in metabolite intensities. Metabolic profiling plays a critical role in assessing cellular responses to cytotoxic agents, as alterations in metabolite intensity can reflect drug-induced modulation of specific targets. Moreover, metabolomics enables the investigation of effects on both upstream and downstream molecular targets, providing a comprehensive overview of the impact on related metabolic processes and can aid the prediction of potential side effects [131, 231].

However, several factors can influence the acquisition of a comprehensive metabolomic profile. The cellular metabolome exhibits significant diversity in terms of metabolite structure, molecular weight and polarity [232]. Therefore, no single metabolomics approach can collectively cover the entire metabolome [69]. Nevertheless, a targeted LC-MS metabolomics approach is

ideal for high-resolution analysis of a pre-defined set of metabolites encompassing key nodes in the metabolic network.

5.1 Evaluation of the established method: advantages and limitations

Cell culture metabolomics inherently faces numerous challenges, primarily due to the need to establish precise cell culture conditions and refine the extraction processes. One of the major challenges is the specification of factors such as cell age (passage number) and growth conditions, which are crucial in minimizing variations in cell proliferation rates that can significantly affect metabolic profiles. To address these variabilities, our protocol has incorporated stringent controls over these variables to ensure consistent and reliable results.

A critical component of our approach is the optimization of the quenching and extraction processes. Effective quenching is essential for halting cellular metabolic activities at precise time points, which is crucial for capturing the metabolic state of cells in response to treatment accurately. In our protocol, quenching involves the use of a 10% HCl in ethanol solution (v/v), - 70 °C, ensuring immediate arrest of enzymatic activities and preserving the integrity of labile metabolites such as nucleotides and CoA esters. Following quenching, ultrasonication is employed instead of traditional trypsination to detach cells. This method not only accelerates the process but also reduces variability between samples by eliminating additional mechanical or enzymatic steps that could alter the metabolic profile. The extraction procedure has been thoroughly developed to enhance the sensitivity of the metabolomics analysis. Our protocol demonstrates the capability to detect distinct metabolic profiles and accurately assign them to the correct MoA from as few as 60,000 cells. Despite its advantages, the method requires stringent control over experimental conditions to mitigate the effects of temperature fluctuations that could influence metabolic profiles. The use of nitrogen stream for solvent evaporation effectively concentrates metabolites but necessitates careful handling to ensure uniform sample volumes and preserve the quality of the data to be generated.

Moreover, with regards to the analysis method, we strategically selected ion-pairing chromatography (IPC) combined with targeted MS/MS operating in multiple reaction monitoring (MRM) mode. The CCEM metabolites are predominantly highly polar and carry anionic charges, making IPC an ideal choice. IPC enhances the retention and resolution of CCEM metabolites on the column, thus facilitating more efficient separation and identification. We employed the QTRAP 6500 system because of its high sensitivity and superior linear response, which are critical for detecting low-abundance metabolites and ensuring accurate quantitation across a broad range of concentrations. The MRM configuration can be adjusted to operate as scheduled MRM (sMRM), optimizing the dwell time per transition. This

optimization allows for relatively short dwell times (≤ 50 ms) typical for the QTRAP 6500, enabling rapid and efficient scanning without sacrificing the quality of data.

Our results in PC-3 prostate cancer cells clearly demonstrate that targeted UPLC-MS/MS analysis of as few as 120 metabolites of the CCEM allows reliable identification of one or even several co-occurring MoAs through which drugs affect cancer cells. Although the approach has a somewhat lower throughput than solely HR-MS-based rapid screenings [126, 132, 233], it avoids misclassification due to poorly annotated and artefactual signals and allows discrimination of important isobaric intermediates [234]. For example, ADP, which has a m/z [M-H] of 426.02 Da, shares this isobaric mass with commonly observed in-source fragments of ATP, CoA esters, NAD cofactors, ADPR, and is isobaric to the [M-H] of dGDP. Without adequate chromatographic separation, there is a risk that these signals could erroneously be grouped under ADP. Additionally, malate is known to produce a significant 115 Da in-source fragment, which coincidentally is the [M-H] of fumarate. Such overlaps highlight the critical need for effective chromatographic separation to ensure accurate metabolite identification and avoid analytical biases in metabolomics studies.

Furthermore, the protocol benefits from an integral normalization step based on the sum of total metabolite peak areas, which assumes that the overall signals of metabolites is proportional to the cell count, independent from treatment impact on cell metabolism. This approach has been rigorously tested and validated, showing a robust correlation between cell number and metabolite intensity, even under different treatment conditions. This normalization method simplifies the analytical process and enhances the reliability of comparing metabolic changes across samples to better understanding the metabolic responses induced by specific treatments.

5.2 Metabotyping of PC-3 cells treated with reference compounds

The premise underlying the metabolomics-driven prediction of drug's MoA for uncharacterized compounds is that drugs targeting similar pathways will produce comparable effects on the metabolome. However, variations in the uptake kinetics and target affinity among drugs with identical targets can influence the timing and intensity of their metabolic responses, as well as the minimal effective concentration needed to achieve the intended effects. Additionally, the specific MoA of a drug can either quickly halt cell growth or trigger cell death. Consequently, factors such as the duration of drug exposure, dosage, and the selection of appropriate comparable drugs are critical considerations in the design of comparative drug metabolic profiling studies. In our study, which aimed to establish a reference framework for metabolic responses in PC-3 cells, we selected 38 reference compounds, each possessing a known

MoA. This diverse set of compounds, selected for their robust MoA knowledge and influence on PC-3 cell growth, provided a comprehensive basis for examining metabolic alterations induced by various drug targets. We utilized a 48-hour exposure period of compounds at IC_{50} concentration on the cells. This approach ensures that all cell samples are expected to be in a comparable state prior to metabolomics analysis, both in terms of growth phase and growth inhibition. In addition, the use of IC_{50} concentrations was a good compromise between an application of lower concentrations, such as IC_{20}, which might not sufficiently capture subtle metabolic changes, and higher concentrations, such as IC_{80}, which could obscure results with patterns of apoptosis or necrosis induction while not providing the minimum required sample material. Hierarchical clustering of metabolic profiles revealed that about 95% of these compounds clearly segregated into two main clusters according to their impact either on mitochondrial functions or on cytosolic processes. This distinct clustering underscores the precision of metabolomic profiling in identifying specific drug actions.

Beside the pattern identification and metabotype clustering, drug-induced metabolic profiles contain important mechanistic information that may not be uncovered to this extent using high-throughput methods without chromatographic separation. We illustrated this for different types of oxidative phosphorylation (OXPHOS) inhibition in relation to regulatory metabolic effects on cellular pyrimidine biosynthesis. Moreover, antimycin A was recently found to stimulate fumarate (FUM) reduction by inhibiting complex III (CPLX III) in 143B osteosarcoma cells, leading to re-oxidation of ubiquinol to ubiquinone by CPLX II and succinate (SUC) accumulation [235]. In accordance with our study in PC-3 cells, SUC enrichment was also detected in CPLX III or fumarase mutant cells but not after CPLX I inhibition [236]. Stable isotope labeling of these CPLX III-compromised cells demonstrated that pyruvate carboxylase (EC 6.4.1.1) contributes to malate (MAL) and FUM formation, while SUC is produced from alpha-ketoglutarate (α-KG) and from FUM, respectively, via clockwise and counterclockwise fluxes of the TCA cycle. Thus, the different patterns observed for SUC, FUM, and MAL upon CPLX I compared with CPLX III inhibition can therefore be attributed to the fact that FUM reduction is possible only upon inhibition of CPLX III, while at the same time only a small fraction of FUM and MAL can be replenished by the reductive pathway under conditions of energy deficiency [236]. The inhibition of succinate dehydrogenase (SDH, EC 1.3.5.1) in CPLX II leads to depletion of FUM and MAL and accumulation of SUC, as SDH is responsible for the conversion of SUC to FUM.

The activity of DHODH depends on the availability of CoQ as well as on its reduction state, i.e. the ratio of ubiquinone to ubiquinol (UQH2). This ratio was assessed in isolated mitochondria of 143B osteosarcoma cells, comparing partial SDH inhibition with

pharmacological CPLX III inhibition. Remarkably elevated UQH2/CoQ ratios were exclusively observed with antimycin A-mediated CPLX III inhibition but not with CPLX II inhibition alone [235]. These findings align well with our observations, where CPLX III inhibition distinctly influences carbamoyl aspartate (CA) and dihydroorotate (DHO) patterns compared with upstream CPLX I and CII inhibitions (Figure 12).

A central question in cancer cell metabolomics research is whether the primary target modulation by a compound dominates the observed metabolic pattern or whether the metabolic changes are additionally significantly influenced by secondary or off-target interactions. Understanding these dynamics of complex interference when analyzing the relationship between the compound and its target through target engagement, can help to elucidate the drug-target interactions even when the MoA remains unclear [237]. The challenge lies when using OMICS techniques for determining a target of poly-pharmacological drugs is to identify those mainly affected. This often depends on the target location within the metabolic network and whether it is embedded in a key metabolic pathway. For instance, the signature of CPLX III / dihydroorotate dehydrogenase (DHODH) inhibition can be discerned by the accumulation of DHO and depletion of orotate ORO, indicating disruptions within a specific metabolic pathway. A broader analysis might involve examining both upstream and downstream metabolites, such as the accumulation of CA and depletion of pyrimidine intermediates, to reinforce the hypothesis of DHODH inhibition. Conversely, evaluating the impact of agents like microtubule inhibitors, DNA alkylating agents or topoisomerase inhibitors poses a different challenge, as their primary metabolic targets lie outside the main metabolic pathways. For drugs like doxorubicin or paclitaxel, which significantly impact the CCEM, understanding their specific targets might be less straightforward. The metabolic pattern of doxorubicin, for instance, might reflect its modulation of the cellular redox state and the impact of ROS on the ETC, leading to broader metabolic alterations [238]. Thus, while the metabolomic profile of these compounds is indicative of their main MoA, the precise target engagement details often remain elusive.

5.3 Machine learning application for MoA prediction of novel anticancer agents

Our study employed machine learning (ML) techniques such as Random Forests (RF) and Lasso regression (LASSO) to predict drug MoA from metabolomic profiles, addressing the challenge posed by sparse training data. The robustness of RF against sparse data is particularly notable, as it avoids overfitting through advanced sampling techniques and bootstrapping. Similarly, the use of LASSO, with its emphasis on regularization, allows for a

focused analysis by highlighting only those components of the data that make significant contributions to the model. Model-based approaches are complementary to unsupervised analyses and to different molecular signatures of drugs, in particular for the detection of short and long-term responses to drug treatment.

The necessity for a leave-one-out cross-validation strategy was dictated by the limited amount of high-quality training data available. This approach effectively prevented data leakage and allowed for a comprehensive evaluation of each model's performance. The significant improvement in accuracy demonstrated by both the LASSO and RF methods over the baseline k-nearest neighbor model highlights their suitability for this type of analysis.

Despite having a small training dataset, we could demonstrate the predictability of the MoA of new drug candidates. Overall, the RF approach resulted in more evenly distributed prediction scores across MoAs compared to the LASSO-based approach, which usually decides on one or more MoAs. This is due to the fact that each decision tree in the RF must decide on exactly one MoA, which can lead to different predictions if the MoA of the drug does not match exactly with the training drugs. In contrast, the LASSO approach aims to explain the observed CCEM profile of the drug by a small subset of CCEM profiles, leading to coefficients of many MoAs being pushed towards zero. Thus, the LASSO-based approach can better represent mixtures of MoAs because it considers additional MoAs as partial contribution(s) (different from zero) to improve the overall regression. In predictions across cell lines, the Lasso-based approach also proves to be more robust with regard to cell line-specific base profiles and predicts the correct MoA for three out of four tested drugs.

The restricted predictability of the TYMS inhibitor 5-FU (5-Fluorouracil) suggests that training on a narrow selection of drugs with limited diversity in MoAs will inevitably struggle to accurately align new metabolic patterns, unless the training data is expanded. However, the method's ability to avoid false positives in such cases is a significant strength, as it directs our attention to potential gaps in the training data and areas for methodological enhancement. Moreover, the low prediction scores for several MoAs provide a critical feedback mechanism, allowing us to identify and address these gaps systematically. This iterative process of testing and refining enhances the robustness of our approach, enabling it to better interpolate between known MoAs and potentially adapt to novel or unexpected MoAs in future studies.

Moreover, using untargeted metabolomics to collect insights on metabolic alterations after compound treatment offers an attractive approach by providing a rich dataset from thousands of signals, thereby avoiding the limitations of analyzing a restricted metabolite set as in a targeted approach. This method potentially enhances predictability through comprehensive

metabolic fingerprinting, which detects extensive analytes rather than focusing solely on CCEM. However, a major limitation is that the chemical structures of most MS features remain unidentified. Although this is not relevant for improved pattern recognition, it complicates the mining of pathways and mechanistic insights. Therefore, our approach, which focuses specifically on CCEM, holds significant advantages. By targeting well-defined metabolic pathways, we can more accurately infer MoA, effectively balancing the breadth of untargeted metabolomics with the depth of targeted investigations. This strategy simplifies the analytical process and enhances the interpretability and relevance of the findings to specific biochemical pathways.

Taken together, these results validate the approach, and its predictive power will further improve as the number of compounds and especially cell types used in the training sets increases.

5.4 Method transferability to other cancer cell lines

The method's applicability was evaluated by metabotyping two additional cancer cell lines, MCF-7 and MHH-ES-1, treated with compounds targeting two MoAs. Clustering analysis of the metabotypes consistently revealed two primary clusters, independent of the cell line, reflecting distinct metabolic responses to glutamate dehydrogenase (GDH) inhibitors bithionol and hexachlorophorane (BITN and HEXA) and HMG-CoA reductase inhibitors atorvastatin and lovastatin (LOVA and ATOR). This demonstrates that the method's specificity is driven by the targeted pathways, confirming its utility across various cellular contexts.

Furthermore, we employed the ML approaches using PC-3 metabotypes data as a training set to predict the MoA of the tested compounds in the additional cell lines. We found characteristic metabolic patterns associated with individual MoAs to be conserved, e.g. HMG-CoA accumulates in all cell models after HMG-CoA reductase inhibition. Nevertheless, and not surprisingly, correct MoA prediction based on PC-3 training data for one out four tested compounds (BETN) was associated with lower prediction reliability in the two other cancer cell models investigated. Although investigations into cancer cells have elucidated the fundamental principles of cell metabolism and describe how metabolites contribute to cell growth and identity, recent advancements emphasize the nuanced nature of metabolism. Rather than ‘‘one-size-fits-all’’, cells exhibit the ability to adopt diverse metabolic strategies based on factors such as lineage, developmental stage, or environmental context.

In our study, the selected cancer cell lines exhibit both similarities and differences in their metabolic processes, offering valuable insights. Similarities may signify universal metabolic

requirements or intrinsic demands of a specific cell type, while differences can unveil metabolic limitations in particular conditions or locations. For example, the MCF-7 breast cancer cell line, characterized by estrogen receptor positivity (ER^+) and HER2 negativity ($HER2^-$), consistently displays a mutation in the PIK3CA gene that is associated with heightened glycolytic activity [239, 240]. Conversely, the PC3 prostate cancer cells not only utilize glucose but also engage in long-chain fatty acid oxidation, primarily through increased activity of carnitine palmitoyltransferase-1 to meet their energetic needs [241].

Previous analyses of the metabolomes in up to 928 cancer cell lines revealed significant variations in their basal metabolite profiles, primarily stemming from distinct enzymatic regulations within the CCEM. These variations are, in part, a result of epigenetic effects or of genetic mutations specific to the respective cell lines. All three cancer cell models under exemplary investigation are known to carry specific gene mutations in various oncogenes affecting the CCEM. The impact of individual mutations or epigenetic alterations on drug response in these models remains unknown. Despite these genetic differences, similar metabotypes were observed after drug treatment across the three cell lines, which is promising as proof of concept for the method's robustness. Conducting comprehensive studies with additional cancer cell models, large drug libraries and machine learning will be crucial to understand the effects of individual mutations and drug treatments on metabolism. For a more general screening, it might be useful to select up to five cancer cell lines of relevance with the broadest metabolic diversities, so the ML- algorithms cover a large variety without the need to test many different cell lines.

5.5 Approach limitations

The study faces several challenges, including the coverage of all possible MoAs of novel anticancer agents questioning the applicability of the only 16 MoAs considered in this study to new drugs. The intent of this study was to serve as a showcase and proof of concept for the potential of metabolomics in preclinical drug discovery, acknowledging that the developed method as always has a limited coverage of all MoAs possible for a novel anticancer agent. Additionally, there is limited diversity in drugs with similar MoA in the number of drugs that could be examined in this work. However, the described methodology has been proven to be able to produce a MoA-representive metabolic pattern from as few as two compounds per MoA.

One significant challenge of our method is distinguishing between metabolic changes caused directly by drug-target interactions and those that are indirect adaptive responses. Additionally, there is a predominant focus on drugs targeting central metabolism, with the targeted

metabolomics approach concentrating mainly on this aspect. Nevertheless, the aim of this study is clarified as defining specific metabolic patterns within the central metabolism resulting from target engagement, whether inside or outside the central carbon metabolism. Moreover, focusing on metabolomics analyzing only CCEM may further restrict the method's applicability. Nonetheless, many anticancer drugs currently in use either directly target CCEM or exhibit specific effects on downstream CCEM patterns, even when they act on non-metabolic targets, as evidenced by MoAs such as AKT or antimicrotubular bioactives.

Issues related to data acquisition and normalization might be raised, particularly the lack of information on growth inhibition and potential biases from varying cell numbers and growth phases across treatments. These concerns were addressed by highlighting the constant use of compound-specific IC_{50} concentrations for data normalization, aiming for method simplicity and minimal variations, to keep the approach practical. For the same reason, absolute values were not discussed. The focus was on an equilibrated approach, providing a maximum of information with a minimum of effort. While the utilization of a single time point of 48 hours is presented, it may be considered a limitation, as prolonged exposure could potentially capture a broader range of cellular responses to the drug, emphasizing sustained effects rather than the initial direct drug impact and target engagement. Numerous examples of both metabolic and non-metabolic target inhibitions are presented, supporting the argument that the chosen 48-hour time point allows sufficient time for the refinement of the MoAs at the metabolic level across a broad spectrum of compounds. In addition, we tested a 4-hour exposure for the reference compounds (data are not shown), finding that while some drugs exhibited specific metabolic patterns already after 4 hours, others did not, and yet others showed different patterns after 48 hours compared to 4 hours. Consequently, we omitted the early time point and opted for a single 48-hour time point to be utilized in the metabotyping experiments.

As for the concerns about the effectiveness of such techniques with smaller datasets, the study defends the machine learning approach, citing examples such as Breastin, where it significantly contributed to confirm Breastin’s MoA as acting on microtubules, even though this MoA only indirectly affects CCEM, corroborating the validity of the approach even with the yet limited training set.

Thus, the present study aims to showcase the proof of concept of metabolomics and ML in preclinical drug discovery, considering the aforementioned limitations.

5.6 Summary and conclusions

Taken together, these results validate the approach, and its predictive power will further improve as the number of compounds and especially cell types used in the training sets increases. Additional cell lines will provide different metabolic signatures to be trained, and both similarities and differences of cellular responses will strengthen future prediction models.

Our approach creates new opportunities for metabolic research, for example, to identify drugs with overlapping or also with non- or poorly overlapping multiple MoAs, e.g., for the development of combination chemotherapies [242] based on model predictions. The example of the plant extract Breastin shows that multi-compound drugs and phytotherapeutics can be successfully applied and that ML correlations can provide useful hints to enlighten their MoAs, as proven by a causal study conducted in an independent lab [225].

Ultimately, our metabolomics methodology offers insights for transitioning from MoA to MeA. This is exemplified by cases where specific metabolite accumulations or deficiencies can indicate distinct targets, as demonstrated herein with NAMPT.

In conclusion, our methodology allows to correlate the metabolic impact of cytotoxic substances or even complex compound mixtures with pre-defined MoAs, especially if these directly or indirectly affect the central carbon and energy metabolism. ML-based pattern recognition not only enables the fine distinctiveness and prediction of MoAs, e.g. as shown by the differention of individual respiratory chain complexes. Mechanistic insights from metabolomics-based MoA studies imply valuable opportunities for metabolic research, for example, to identify drugs for the development of combination chemotherapies. Therefore, our approach can significantly improve drug discovery and development by directing MoA studies toward the identification of specific molecular targets or pathways, and subsequent medicinal chemistry improvements of the lead compound, which is not restricted to cancer research.

6 Future perspectives and outlook

Improve the predictive power and accuracy of our method. Firstly, increasing the number of compounds per MoA will substantially enhance the accuracy and confidence in assigning metabotypes of interest to their correct MoA in the prediction study. It is important to recognize that drugs that share the same primary target do not necessarily share the same secondary (or further off-target) targets. Therefore, incorporating a broader range of structurally diverse training compounds that modulate the same primary target will improve the likelihood of capturing a wider array of potential metabotypes arising from target interactions (primary and secondary). This expanded dataset can improve the predictive power and enhance the accuracy of the method. Second, by increasing the size of the training dataset by including additional MoA suggestions, the method can cover a wider range of potential MoAs for new drugs, thereby reducing misclassifications and increasing the possibility to categorize further MoA.

In parallel, one can expand the dataset by including more predictive compound data from IPB's in-house library, which can be used to investigate the MoAs of promising compounds, which was the primary goal of our method development. This expansion of data resources will provide valuable support for cancer research and the drug discovery process at IPB.

Extension of the method to assess the influence of anti-inflammatory modulators on CCEM. The application of this method extends beyond cancer research and drug discovery to include basic research aimed at understanding how, e.g., anti-inflammatory agents affect energy metabolism within cancer cells. This research may involve the use of both differentiated and undifferentiated human leukemia monocytic cell lines (THP-1) [243] in the presence of traditional anti-inflammatory agents and promising new compounds. The aim is to study their effect on inflammatory processes, either alone or in combination with an inflammatory inducer such as lipopolysaccharides (LPS).

Discovery of DHODH/CIII inhibitors by metabolomics screening. The application of this method goes beyond basic research and has been shown in previous studies to be suitable for screening campaigns [131]. In particular, the search for novel DHODH inhibitors is currently a prominent area of interest with potential applications in treating and controlling cancer and other diseases [244, 245]. DHODH has shown promise as a therapeutic target in diseases such as acute myeloid leukemia, multiple myeloma, and viral and bacterial infections [246-250]. As a result, the medicinal chemistry community is actively searching for new DHODH ligands. However, the use of this method for the screening and discovery of novel DHODH ligands requires further refinements in terms of increased HTS performance, including the adaptation of 96-well plates for whole cell utilization, preferably using PC-3 cells, and upgrading the

extraction protocol to align with plate-based procedures. This improvement will allow optimization for higher throughput. Further investment in bioinformatics is recommended to develop a ML tool to accelerate data analysis and enable faster CIII/DHODH inhibition decisions based on the resulting metabolic patterns and metabolite intensities.

Further investigation into the synergistic effects of CPLX III and uncoupler on PC-3 cells is considered a valuable future focus. It is intended to test additional, safer anticancer agents having combined or additive MoAs to minimize potential adverse effects. This strategy enhances the therapeutic potential. After confirming the efficacy of these potential treatments, progression to animal testing will be considered. This putative combinatorial therapy was identified through analyzing the metabolic profiles of the individual drugs, suggesting a promising avenue for enhanced therapeutic strategies.

Reducing the set of metabolites to a minimum number is still predictive. To increase the productivity of the method, it is possible to reduce the number of the relevant metabolites to a minimal but significant number. The effective metabolites, as defined by the classifier used in the developed machine learning model are referred to as VIP features. However, these VIP features represent the common metabolites that have the highest ranking in the shaping of the pattern for the overall metabotypes of the training set. It is important to distinguish these VIP features from those that are specific to individual MoAs. The aim is to find the minimum number of VIP metabolites that are sufficient to give the best cross-validation results and the highest accuracy in predicting MoA for newly discovered drugs. In addition to making the method more accessible to other researchers and transferable to other laboratories, this reduction will increase the applicability of the method, minimizing the time required for metabolic analyses and data processing.

Can a viable business model be developed? The transition from research to a business model involves translating scientific discoveries and innovations into commercially viable products or services. Commercial tool development should comprise the analysis of a smaller set of metabolites (~ 20 analytes) for MoA prediction, combined with an optimized extraction protocol and the integration of machine learning algorithms. The MoA finding tool can be the core of a commercialized product with an immense value in cancer research and drug discovery with likely sufficient interest from the researchers in academia and companies. The targeted customers are the cancer research labs and companies who have an interest in finding the MoA of a novel anticancer drug. The product could comprise a physical kit and an analytical service where the kit contains the standards of the most relevant metabolites. The kit can be distributed to the interested parties, to be mixed with the sample of choice, such as a metabolic

extract of cancer cells treated with the compound under investigation. Second, the logistics can be arranged and covered to run the analysis of the kit's contents in addition to the data analysis and processing including the ML. Afterwards, the results can be sent directly to the customer indicating the predicted MoA. This product can be relatively low cost, however, with high scientific value as it can save the researchers the funds for extensive MoA research.

7 References

(1) Ruddon, R. W. *Cancer biology*; Oxford University Press, 2007.

(2) Anand, P.; Kunnumakkara, A. B.; Sundaram, C.; Harikumar, K. B.; Tharakan, S. T.; Lai, O. S.; Sung, B.; Aggarwal, B. B. Cancer is a preventable disease that requires major lifestyle changes. *Pharmaceutical research* **2008**, *25* (9), 2097-2116. DOI: 10.1007/s11095-008-9661-9.

(3) Bray, F.; Laversanne, M.; Sung, H.; Ferlay, J.; Siegel, R. L.; Soerjomataram, I.; Jemal, A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians* **2024**, *74* (3), 229-263. DOI: 10.3322/caac.21834.

(4) Urruticoechea, A.; Alemany, R.; Balart, J.; Villanueva, A.; Vinals, F.; Capella, G. Recent advances in cancer therapy: an overview. *Current pharmaceutical design* **2010**, *16* (1), 3-10. DOI: 10.2174/138161210789941847.

(5) Delaney, G.; Jacob, S.; Featherstone, C.; Barton, M. The role of radiotherapy in cancer treatment: estimating optimal utilization from a review of evidence-based clinical guidelines. *Cancer: Interdisciplinary International Journal of the American Cancer Society* **2005**, *104* (6), 1129-1137. DOI: 10.1002/cncr.21324.

(6) Chu, E.; Sartorelli, A. Cancer chemotherapy. *Lange's Basic and Clinical Pharmacology* **2018**, 948-976.

(7) Hanahan, D.; Weinberg, R. A. The hallmarks of cancer. *Cell* **2000**, *100* (1), 57-70. DOI: 10.1016/s0092-8674(00)81683-9.

(8) Wishart, D. S. Emerging applications of metabolomics in drug discovery and precision medicine. *Nature reviews Drug discovery* **2016**, *15* (7), 473-484. DOI: 10.1038/nrd.2016.32.

(9) Kim, J. W.; Dang, C. V. Cancer's molecular sweet tooth and the Warburg effect. *Cancer research* **2006**, *66* (18), 8927-8930. DOI: 10.1158/0008-5472.CAN-06-1501.

(10) Hanahan, D.; Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **2011**, *144* (5), 646-674. DOI: 10.1016/j.cell.2011.02.013.

(11) Warburg, O.; Wind, F.; Negelein, E. The Metabolism of Tumors in the Body. *The Journal of general physiology* **1927**, *8* (6), 519-530. DOI: 10.1085/jgp.8.6.519.

(12) Weinhouse, S. On respiratory impairment in cancer cells. *Science* **1956**, *124* (3215), 267-269. DOI: 10.1126/science.124.3215.267.

(13) Vander Heiden, M. G.; Cantley, L. C.; Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **2009**, *324* (5930), 1029-1033. DOI: 10.1126/science.1160809.

(14) Liberti, M. V.; Locasale, J. W. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends in biochemical sciences* **2016**, *41* (3), 211-218. DOI: 10.1016/j.tibs.2015.12.001.

(15) DeBerardinis, R. J.; Chandel, N. S. We need to talk about the Warburg effect. *Nature metabolism* **2020**, *2* (2), 127-129. DOI: 10.1038/s42255-020-0172-2.

(16) Liu, W.; Wang, Y.; Bozi, L. H. M.; Fischer, P. D.; Jedrychowski, M. P.; Xiao, H.; Wu, T.; Darabedian, N.; He, X.; Mills, E. L.; et al. Lactate regulates cell cycle by remodelling the anaphase promoting complex. *Nature* **2023**, *616* (7958), 790-797. DOI: 10.1038/s41586-023-05939-3.

(17) Wang, Z. H.; Peng, W. B.; Zhang, P.; Yang, X. P.; Zhou, Q. Lactate in the tumour microenvironment: From immune modulation to therapy. *EBioMedicine* **2021**, *73*, 103627. DOI: 10.1016/j.ebiom.2021.103627.

(18) Castel, P.; Rauen, K. A.; McCormick, F. The duality of human oncoproteins: drivers of cancer and congenital disorders. *Nature Reviews Cancer* **2020**, *20* (7), 383-397. DOI: 10.1038/s41568-020-0256-z.

(19) Yuneva, M.; Zamboni, N.; Oefner, P.; Sachidanandam, R.; Lazebnik, Y. Deficiency in glutamine but not glucose induces MYC-dependent apoptosis in human cells. *The Journal of cell biology* **2007**, *178* (1), 93-105. DOI: 10.1083/jcb.200703099.

(20) Stine, Z. E.; Schug, Z. T.; Salvino, J. M.; Dang, C. V. Targeting cancer metabolism in the era of precision oncology. *Nature reviews Drug discovery* **2022**, *21* (2), 141-162. DOI: 10.1038/s41573-021-00339-6.

(21) Rabinowitz, J. D.; White, E. Autophagy and metabolism. *Science* **2010**, *330* (6009), 1344-1348. DOI: 10.1126/science.1193497.

(22) Pavlova, N. N.; Zhu, J.; Thompson, C. B. The hallmarks of cancer metabolism: Still emerging. *Cell metabolism* **2022**, *34* (3), 355-377. DOI: 10.1016/j.cmet.2022.01.007.

(23) DeBerardinis, R. J.; Cheng, T. Q's next: the diverse functions of glutamine in metabolism, cell biology and cancer. *Oncogene* **2010**, *29* (3), 313-324. DOI: 10.1038/onc.2009.358.

(24) Gilard, V.; Ferey, J.; Marguet, F.; Fontanilles, M.; Ducatez, F.; Pilon, C.; Lesueur, C.; Pereira, T.; Basset, C.; Schmitz-Afonso, I.; et al. Integrative Metabolomics Reveals Deep Tissue and Systemic Metabolic Remodeling in Glioblastoma. *Cancers (Basel)* **2021**, *13* (20). DOI: 10.3390/cancers13205157.

(25) Naser, F. J.; Jackstadt, M. M.; Fowle-Grider, R.; Spalding, J. L.; Cho, K.; Stancliffe, E.; Doonan, S. R.; Kramer, E. T.; Yao, L.; Krasnick, B.; et al. Isotope tracing in adult zebrafish reveals alanine cycling between melanoma and liver. *Cell metabolism* **2021**, *33* (7), 1493-1504 e1495. DOI: 10.1016/j.cmet.2021.04.014.

(26) Seyfried, T. N.; Flores, R. E.; Poff, A. M.; D'Agostino, D. P. Cancer as a metabolic disease: implications for novel therapeutics. *Carcinogenesis* **2014**, *35* (3), 515-527. DOI: 10.1093/carcin/bgt480.

(27) Warburg, O. On the origin of cancer cells. *Science* **1956**, *123* (3191), 309-314. DOI: 10.1126/science.123.3191.309.

(28) Cassim, S.; Vucetic, M.; Zdralevic, M.; Pouyssegur, J. Warburg and Beyond: The Power of Mitochondrial Metabolism to Collaborate or Replace Fermentative Glycolysis in Cancer. *Cancers (Basel)* **2020**, *12* (5). DOI: 10.3390/cancers12051119.

(29) Martins Pinto, M.; Paumard, P.; Bouchez, C.; Ransac, S.; Duvezin-Caubet, S.; Mazat, J. P.; Rigoulet, M.; Devin, A. The Warburg effect and mitochondrial oxidative phosphorylation: Friends or foes? *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **2023**, *1864* (1), 148931. DOI: 10.1016/j.bbabio.2022.148931.

(30) Pfeiffer, T.; Schuster, S.; Bonhoeffer, S. Cooperation and competition in the evolution of ATP-producing pathways. *Science* **2001**, *292* (5516), 504-507. DOI: 10.1126/science.1058079.

(31) Boroughs, L. K.; DeBerardinis, R. J. Metabolic pathways promoting cancer cell survival and growth. *Nature cell biology* **2015**, *17* (4), 351-359. DOI: 10.1038/ncb3124.

(32) Wang, Y.; Patti, G. J. The Warburg effect: a signature of mitochondrial overload. *Trends in cell biology* **2023**, *33* (12), 1014-1020. DOI: 10.1016/j.tcb.2023.03.013.

(33) Finley, L. W. S. What is cancer metabolism? *Cell* **2023**, *186* (8), 1670-1688. DOI: 10.1016/j.cell.2023.01.038.

(34) Luengo, A.; Li, Z.; Gui, D. Y.; Sullivan, L. B.; Zagorulya, M.; Do, B. T.; Ferreira, R.; Naamati, A.; Ali, A.; Lewis, C. A.; et al. Increased demand for NAD(+) relative to ATP drives aerobic glycolysis. *Molecular cell* **2021**, *81* (4), 691-707 e696. DOI: 10.1016/j.molcel.2020.12.012.

(35) Li, Z.; Ji, B. W.; Dixit, P. D.; Tchourine, K.; Lien, E. C.; Hosios, A. M.; Abbott, K. L.; Rutter, J. C.; Westermark, A. M.; Gorodetsky, E. F.; et al. Cancer cells depend on environmental lipids for

proliferation when electron acceptors are limited. *Nature metabolism* **2022**, *4* (6), 711-723. DOI: 10.1038/s42255-022-00588-8.

(36) Zhu, Y. M.; Xu, P.; Huang, X.; Shuai, W.; Liu, L.; Zhang, S.; Zhao, R.; Hu, X. Y.; Wang, G. From Rate-Limiting Enzyme to Therapeutic Target: The Promise of NAMPT in Neurodegenerative Diseases. *Frontiers in Pharmacology* **2022**, *13*. DOI: 10.3389/fphar.2022.920113.

(37) Tan, B.; Young, D. A.; Lu, Z. H.; Wang, T.; Meier, T. I.; Shepard, R. L.; Roth, K.; Zhai, Y.; Huss, K.; Kuo, M. S.; et al. Pharmacological Inhibition of Nicotinamide Phosphoribosyltransferase (NAMPT), an Enzyme Essential for NAD(+) Biosynthesis, in Human Cancer Cells METABOLIC BASIS AND POTENTIAL CLINICAL IMPLICATIONS. *Journal of Biological Chemistry* **2013**, *288* (5), 3500-3511. DOI: 10.1074/jbc.M112.394510.

(38) Vander Heiden, M. G.; DeBerardinis, R. J. Understanding the Intersections between Metabolism and Cancer Biology. *Cell* **2017**, *168* (4), 657-669. DOI: 10.1016/j.cell.2016.12.039.

(39) Zong, W. X.; Rabinowitz, J. D.; White, E. Mitochondria and Cancer. *Molecular cell* **2016**, *61* (5), 667-676. DOI: 10.1016/j.molcel.2016.02.011.

(40) Sullivan, L. B.; Gui, D. Y.; Hosios, A. M.; Bush, L. N.; Freinkman, E.; Vander Heiden, M. G. Supporting Aspartate Biosynthesis Is an Essential Function of Respiration in Proliferating Cells. *Cell* **2015**, *162* (3), 552-563. DOI: 10.1016/j.cell.2015.07.017.

(41) Birsoy, K.; Wang, T.; Chen, W. W.; Freinkman, E.; Abu-Remaileh, M.; Sabatini, D. M. An Essential Role of the Mitochondrial Electron Transport Chain in Cell Proliferation Is to Enable Aspartate Synthesis. *Cell* **2015**, *162* (3), 540-551. DOI: 10.1016/j.cell.2015.07.016.

(42) Villa, E.; Ali, E. S.; Sahu, U.; Ben-Sahra, I. Cancer Cells Tune the Signaling Pathways to Empower de Novo Synthesis of Nucleotides. *Cancers (Basel)* **2019**, *11* (5). DOI: 10.3390/cancers11050688.

(43) Chen, J.; Yang, S.; Li, Y.; Ziwen, X.; Zhang, P.; Song, Q.; Yao, Y.; Pei, H. De novo nucleotide biosynthetic pathway and cancer. *Genes & Diseases* **2023**, *10* (6), 2331-2338. DOI: 10.1016/j.gendis.2022.04.018.

(44) Farber, S. Some observations on the effect of folic acid antagonists on acute leukemia and other forms of incurable cancer. *Blood* **1949**, *4* (2), 160-167.

(45) Djerassi, I.; Farber, S.; Abir, E.; Neikirk, W. Continuous infusion of methotrexate in children with acute leukemia. *Cancer* **1967**, *20* (2), 233-242. DOI: 10.1002/1097-0142(1967)20:2<233::aid-cncr2820200209>3.0.co;2-8.

(46) Miller, D. R. A tribute to Sidney Farber-- the father of modern chemotherapy. *British journal of haematology* **2006**, *134* (1), 20-26. DOI: 10.1111/j.1365-2141.2006.06119.x.

(47) Thompson, C. B. Rethinking the regulation of cellular metabolism. *Cold Spring Harbor symposia on quantitative biology. Vol. 76. Cold Spring Harbor Laboratory Press* **2011**, *76*, 23-29. DOI: 10.1101/sqb.2012.76.010496.

(48) Sazanov, L. A. The mechanism of coupling between electron transfer and proton translocation in respiratory complex I. *Journal of bioenergetics and biomembranes* **2014**, *46* (4), 247-253. DOI: 10.1007/s10863-014-9554-z.

(49) Raimondi, V.; Ciccarese, F.; Ciminale, V. Oncogenic pathways and the electron transport chain: a dangeROS liaison. *British journal of cancer* **2020**, *122* (2), 168-181. DOI: 10.1038/s41416-019-0651-y.

(50) Zhao, R. Z.; Jiang, S.; Zhang, L.; Yu, Z. B. Mitochondrial electron transport chain, ROS generation and uncoupling (Review). *International journal of molecular medicine* **2019**, *44* (1), 3-15. DOI: 10.3892/ijmm.2019.4188.

(51) Vasan, K.; Werner, M.; Chandel, N. S. Mitochondrial Metabolism as a Target for Cancer Therapy. *Cell metabolism* **2020**, *32* (3), 341-352. DOI: 10.1016/j.cmet.2020.06.019.

(52) Meyers, R. M.; Bryan, J. G.; McFarland, J. M.; Weir, B. A.; Sizemore, A. E.; Xu, H.; Dharia, N. V.; Montgomery, P. G.; Cowley, G. S.; Pantel, S.; et al. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nature genetics* **2017**, *49* (12), 1779-1784. DOI: 10.1038/ng.3984.

(53) Dempster, J. M.; Rossen, J.; Kazachkova, M.; Pan, J.; Kugener, G.; Root, D. E.; Tsherniak, A. Extracting Biological Insights from the Project Achilles Genome-Scale CRISPR Screens in Cancer Cell Lines. *BioRxiv* **2019,** 720243. DOI: 10.1101/720243

(54) Guo, S.; Miao, M.; Wu, Y.; Pan, D.; Wu, Q.; Kang, Z.; Zeng, J.; Zhong, G.; Liu, C.; Wang, J. DHODH inhibition represents a therapeutic strategy and improves abiraterone treatment in castration-resistant prostate cancer. *Oncogene* **2024**, *43* (19), 1399-1410. DOI: 10.1038/s41388-024-03005-4.

(55) Patra, K. C.; Hay, N. The pentose phosphate pathway and cancer. *Trends in biochemical sciences* **2014**, *39* (8), 347-354. DOI: 10.1016/j.tibs.2014.06.005.

(56) Boros, L. G.; Lee, P. W.; Brandes, J. L.; Cascante, M.; Muscarella, P.; Schirmer, W. J.; Melvin, W. S.; Ellison, E. C. Nonoxidative pentose phosphate pathways and their direct role in ribose synthesis in tumors: is cancer a disease of cellular glucose metabolism? *Medical hypotheses* **1998**, *50* (1), 55-59. DOI: 10.1016/s0306-9877(98)90178-5.

(57) Arnold, P. K.; Finley, L. W. S. Regulation and function of the mammalian tricarboxylic acid cycle. *Journal of Biological Chemistry* **2023**, *299* (2), 102838. DOI: 10.1016/j.jbc.2022.102838.

(58) Baksh, S. C.; Finley, L. W. S. Metabolic Coordination of Cell Fate by alpha-Ketoglutarate-Dependent Dioxygenases. *Trends in cell biology* **2021**, *31* (1), 24-36. DOI: 10.1016/j.tcb.2020.09.010.

(59) Carey, B. W.; Finley, L. W.; Cross, J. R.; Allis, C. D.; Thompson, C. B. Intracellular alpha-ketoglutarate maintains the pluripotency of embryonic stem cells. *Nature* **2015**, *518* (7539), 413-416. DOI: 10.1038/nature13981.

(60) Reid, M. A.; Dai, Z.; Locasale, J. W. The impact of cellular metabolism on chromatin dynamics and epigenetics. *Nature cell biology* **2017**, *19* (11), 1298-1306. DOI: 10.1038/ncb3629.

(61) DeBerardinis, R. J.; Chandel, N. S. Fundamentals of cancer metabolism. *Science advances* **2016**, *2* (5), e1600200. DOI: 10.1126/sciadv.1600200.

(62) Pirozzi, C. J.; Yan, H. The implications of IDH mutations for cancer development and therapy. *Nature reviews Clinical oncology* **2021**, *18* (10), 645-661. DOI: 10.1038/s41571-021-00521-0.

(63) Intlekofer, A. M.; Shih, A. H.; Wang, B.; Nazir, A.; Rustenburg, A. S.; Albanese, S. K.; Patel, M.; Famulare, C.; Correa, F. M.; Takemoto, N.; et al. Acquired resistance to IDH inhibition through trans or cis dimer-interface mutations. *Nature* **2018**, *559* (7712), 125-129. DOI: 10.1038/s41586-018-0251-7.

(64) Schmidt, D. R.; Patel, R.; Kirsch, D. G.; Lewis, C. A.; Vander Heiden, M. G.; Locasale, J. W. Metabolomics in cancer research and emerging applications in clinical oncology. *CA: a cancer journal for clinicians* **2021**, *71* (4), 333-358. DOI: 10.3322/caac.21670.

(65) Farhadi, P.; Yarani, R.; Dokaneheifard, S.; Mansouri, K. The emerging role of targeting cancer metabolism for cancer therapy. *Tumor Biology* **2020**, *42* (10), 1010428320965284. DOI: 10.1177/1010428320965284.

(66) Luengo, A.; Gui, D. Y.; Vander Heiden, M. G. Targeting Metabolism for Cancer Therapy. *Cell chemical biology* **2017**, *24* (9), 1161-1180. DOI: 10.1016/j.chembiol.2017.08.028.

(67) Guerra, A. R.; Duarte, M. F.; Duarte, I. F. Targeting Tumor Metabolism with Plant-Derived Natural Products: Emerging Trends in Cancer Therapy. *Journal of agricultural and food chemistry* **2018**, *66* (41), 10663-10685. DOI: 10.1021/acs.jafc.8b04104.

(68) Wang, Q. Q.; Li, M. X.; Li, C.; Gu, X. X.; Zheng, M. Z.; Chen, L. X.; Li, H. Natural Products and Derivatives Targeting at Cancer Energy Metabolism: A Potential Treatment Strategy. *Current Medical Science* **2020**, *40* (2), 205-217. DOI: 10.1007/s11596-020-2165-5.

(69) Danzi, F.; Pacchiana, R.; Mafficini, A.; Scupoli, M. T.; Scarpa, A.; Donadelli, M.; Fiore, A. To metabolomics and beyond: a technological portfolio to investigate cancer metabolism. *Signal Transduction and Targeted Therapy* **2023**, *8* (1), 137. DOI: 10.1038/s41392-023-01380-0.

(70) Keller, T. H.; Pichota, A.; Yin, Z. A practical view of 'druggability'. *Current opinion in chemical biology* **2006**, *10* (4), 357-361. DOI: 10.1016/j.cbpa.2006.06.014.

(71) Swinney, D. C. Phenotypic vs. target-based drug discovery for first-in-class medicines. *Clinical Pharmacology & Therapeutics* **2013**, *93* (4), 299-301. DOI: 10.1038/clpt.2012.236.

(72) Belfield, G. P.; Delaney, S. J. The impact of molecular biology on drug discovery. *Biochemical Society Transactions* **2006**, *34* (Pt 2), 313-316. DOI: 10.1042/BST20060313.

(73) Yang, Y.; Wu, Z.; Yao, X.; Kang, Y.; Hou, T.; Hsieh, C. Y.; Liu, H. Exploring Low-Toxicity Chemical Space with Deep Learning for Molecular Generation. *Journal of Chemical Information and Modeling* **2022**, *62* (13), 3191-3199. DOI: 10.1021/acs.jcim.2c00671.

(74) Wang, Y.; Michael, S.; Yang, S. M.; Huang, R.; Cruz-Gutierrez, K.; Zhang, Y.; Zhao, J.; Xia, M.; Shinn, P.; Sun, H. Retro Drug Design: From Target Properties to Molecular Structures. *Journal of Chemical Information and Modeling* **2022**, *62* (11), 2659-2669. DOI: 10.1021/acs.jcim.2c00123.

(75) Vincent, F.; Nueda, A.; Lee, J.; Schenone, M.; Prunotto, M.; Mercola, M. Phenotypic drug discovery: recent successes, lessons learned and new directions. *Nature Reviews Drug Discovery* **2022**, *21* (12), 899-914. DOI: 10.1038/s41573-022-00472-w.

(76) Fang, Y. Biosensors: On the Origin of Label-Free Cell Phenotypic Profiles of Drug- Target Interactions. J Biochip Tissue Chip 3 2013, e126.

(77) Chen, Y.; Li, L.; Zhang, G. Q.; Xu, R. Phenome-driven disease genetics prediction toward drug discovery. *Bioinformatics* **2015**, *31* (12), i276-283. DOI: 10.1093/bioinformatics/btv245.

(78) Sadri, A. Is Target-Based Drug Discovery Efficient? Discovery and "Off-Target" Mechanisms of All Drugs. *Journal of medicinal chemistry* **2023**, *66* (18), 12651-12677. DOI: 10.1021/acs.jmedchem.2c01737.

(79) Swinney, D. C.; Anthony, J. How were new medicines discovered? *Nature reviews Drug discovery* **2011**, *10* (7), 507-519. DOI: 10.1038/nrd3480.

(80) Moffat, J. G.; Vincent, F.; Lee, J. A.; Eder, J.; Prunotto, M. Opportunities and challenges in phenotypic drug discovery: an industry perspective. *Nature reviews Drug discovery* **2017**, *16* (8), 531-543. DOI: 10.1038/nrd.2017.111.

(81) Lin, A.; Giuliano, C. J.; Palladino, A.; John, K. M.; Abramowicz, C.; Yuan, M. L.; Sausville, E. L.; Lukow, D. A.; Liu, L.; Chait, A. R.; et al. Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials. *Science translational medicine* **2019**, *11* (509). DOI: 10.1126/scitranslmed.aaw8412.

(82) Yan, S. K.; Liu, R. H.; Jin, H. Z.; Liu, X. R.; Ye, J.; Shan, L.; Zhang, W. D. "Omics" in pharmaceutical research: overview, applications, challenges, and future perspectives. *Chinese journal of natural medicines* **2015**, *13* (1), 3-21. DOI: 10.1016/S1875-5364(15)60002-4.

(83) Oliver, S. G.; Winson, M. K.; Kell, D. B.; Baganz, F. Systematic functional analysis of the yeast genome. *Trends in biotechnology* **1998**, *16* (9), 373-378. DOI: 10.1016/s0167-7799(98)01214-1.

(84) German, J. B.; Hammock, B. D.; Watkins, S. M. Metabolomics: building on a century of biochemistry to guide human health. *Metabolomics* **2005**, *1* (1), 3-9. DOI: 10.1007/s11306-005-1102-8.

(85) Petrick, L. M.; Shomron, N. AI/ML-driven advances in untargeted metabolomics and exposomics for biomedical applications. *Cell Reports Physical Science* **2022**, *3* (7). DOI: 10.1016/j.xcrp.2022.100978.

(86) Qiu, S.; Cai, Y.; Yao, H.; Lin, C.; Xie, Y.; Tang, S.; Zhang, A. Small molecule metabolites: discovery of biomarkers and therapeutic targets. *Signal Transduction and Targeted Therapy* **2023**, *8* (1), 132. DOI: 10.1038/s41392-023-01399-3.

(87) Han, J.; Li, Q.; Chen, Y.; Yang, Y. Recent Metabolomics Analysis in Tumor Metabolism Reprogramming. *Frontiers in Molecular Biosciences* **2021**, *8*, 763902. DOI: 10.3389/fmolb.2021.763902.

(88) Ribbenstedt, A.; Ziarrusta, H.; Benskin, J. P. Development, characterization and comparisons of targeted and non-targeted metabolomics methods. *PLoS One* **2018**, *13* (11), e0207082. DOI: 10.1371/journal.pone.0207082.

(89) Schrimpe-Rutledge, A. C.; Codreanu, S. G.; Sherrod, S. D.; McLean, J. A. Untargeted Metabolomics Strategies-Challenges and Emerging Directions. *Journal of the American Society for Mass Spectrometry* **2016**, *27* (12), 1897-1905. DOI: 10.1007/s13361-016-1469-y.

(90) Alarcon-Barrera, J. C.; Kostidis, S.; Ondo-Mendez, A.; Giera, M. Recent advances in metabolomics analysis for early drug development. *Drug discovery today* **2022**, *27* (6), 1763-1773. DOI: 10.1016/j.drudis.2022.02.018.

(91) Fernie, A. R.; Trethewey, R. N.; Krotzky, A. J.; Willmitzer, L. Metabolite profiling: from diagnostics to systems biology. *Nature reviews molecular cell biology* **2004**, *5* (9), 763-769. DOI: 10.1038/nrm1451.

(92) Lu, W.; Su, X.; Klein, M. S.; Lewis, I. A.; Fiehn, O.; Rabinowitz, J. D. Metabolite Measurement: Pitfalls to Avoid and Practices to Follow. *Annual review of biochemistry* **2017**, *86*, 277-304. DOI: 10.1146/annurev-biochem-061516-044952.

(93) Kang, Y. P.; Ward, N. P.; DeNicola, G. M. Recent advances in cancer metabolism: a technological perspective. *Experimental & molecular medicine* **2018**, *50* (4), 1-16. DOI: 10.1038/s12276-018-0027-z.

(94) Vignoli, A.; Ghini, V.; Meoni, G.; Licari, C.; Takis, P. G.; Tenori, L.; Turano, P.; Luchinat, C. High-Throughput Metabolomics by 1D NMR. *Angewandte Chemie International Edition* **2019**, *58* (4), 968-994. DOI: 10.1002/anie.201804736.

(95) Alseekh, S.; Aharoni, A.; Brotman, Y.; Contrepois, K.; D'Auria, J.; Ewald, J.; J, C. E.; Fraser, P. D.; Giavalisco, P.; Hall, R. D.; et al. Mass spectrometry-based metabolomics: a guide for annotation, quantification and best reporting practices. *Nature methods* **2021**, *18* (7), 747-756. DOI: 10.1038/s41592-021-01197-1.

(96) Markley, J. L.; Bruschweiler, R.; Edison, A. S.; Eghbalnia, H. R.; Powers, R.; Raftery, D.; Wishart, D. S. The future of NMR-based metabolomics. *Current opinion in biotechnology* **2017**, *43*, 34-40. DOI: 10.1016/j.copbio.2016.08.001.

(97) Emwas, A. H. The strengths and weaknesses of NMR spectroscopy and mass spectrometry with particular focus on metabolomics research. *Metabonomics: Methods and protocols* **2015**, *1277*, 161-193. DOI: 10.1007/978-1-4939-2377-9_13.

(98) Collins, S. L.; Koo, I.; Peters, J. M.; Smith, P. B.; Patterson, A. D. Current Challenges and Recent Developments in Mass Spectrometry-Based Metabolomics. *Annual review of analytical chemistry* **2021**, *14* (1), 467-487. DOI: 10.1146/annurev-anchem-091620-015205.

(99) Chen, C. J.; Lee, D. Y.; Yu, J.; Lin, Y. N.; Lin, T. M. Recent advances in LC-MS-based metabolomics for clinical biomarker discovery. *Mass spectrometry reviews* **2022**, e21785. DOI: 10.1002/mas.21785.

(100) Gika, H.; Virgiliou, C.; Theodoridis, G.; Plumb, R. S.; Wilson, I. D. Untargeted LC/MS-based metabolic phenotyping (metabonomics/metabolomics): The state of the art. *Journal of Chromatography B* **2019**, *1117*, 136-147. DOI: 10.1016/j.jchromb.2019.04.009.

(101) Marques, C.; Liu, L.; Duncan, K. D.; Lanekoff, I. A Direct Infusion Probe for Rapid Metabolomics of Low-Volume Samples. *Analytical Chemistry* **2022**, *94* (37), 12875-12883. DOI: 10.1021/acs.analchem.2c02918.

(102) Kirwan, J. A.; Weber, R. J.; Broadhurst, D. I.; Viant, M. R. Direct infusion mass spectrometry metabolomics dataset: a benchmark for data processing and quality control. *Scientific data* **2014**, *1*, 140012. DOI: 10.1038/sdata.2014.12.

(103) Higashi, R. M.; Fan, T. W.; Lorkiewicz, P. K.; Moseley, H. N.; Lane, A. N. Stable isotope-labeled tracers for metabolic pathway elucidation by GC-MS and FT-MS. *Mass spectrometry in metabolomics: Methods and protocols* **2014**, *1198*, 147-167. DOI: 10.1007/978-1-4939-1258-2_11.

(104) Schwaiger-Haber, M.; Stancliffe, E.; Anbukumar, D. S.; Sells, B.; Yi, J.; Cho, K.; Adkins-Travis, K.; Chheda, M. G.; Shriver, L. P.; Patti, G. J. Using mass spectrometry imaging to map fluxes quantitatively in the tumor ecosystem. *Nature communications* **2023**, *14* (1), 2876. DOI: 10.1038/s41467-023-38403-x.

(105) Pang, H.; Hu, Z. Metabolomics in drug research and development: The recent advances in technologies and applications. *Acta Pharmaceutica Sinica B* **2023**, *13* (8), 3238-3251. DOI: 10.1016/j.apsb.2023.05.021.

(106) Balcke, G. U.; Kolle, S. N.; Kamp, H.; Bethan, B.; Looser, R.; Wagner, S.; Landsiedel, R.; van Ravenzwaay, B. Linking energy metabolism to dysfunctions in mitochondrial respiration--a metabolomics in vitro approach. *Toxicology letters* **2011**, *203* (3), 200-209. DOI: 10.1016/j.toxlet.2011.03.013.

(107) Balcke, G. U.; Bennewitz, S.; Bergau, N.; Athmer, B.; Henning, A.; Majovsky, P.; Jimenez-Gomez, J. M.; Hoehenwarter, W.; Tissier, A. Multi-Omics of Tomato Glandular Trichomes Reveals Distinct Features of Central Carbon Metabolism Supporting High Productivity of Specialized Metabolites. *The Plant Cell* **2017**, *29* (5), 960-983. DOI: 10.1105/tpc.17.00060.

(108) Wang, Z.; Yip, L. Y.; Lee, J. H. J.; Wu, Z.; Chew, H. Y.; Chong, P. K. W.; Teo, C. C.; Ang, H. Y.; Peh, K. L. E.; Yuan, J.; et al. Methionine is a metabolic dependency of tumor-initiating cells. *Nature medicine* **2019**, *25* (5), 825-837. DOI: 10.1038/s41591-019-0423-5.

(109) Wu, J. Y.; Huang, T. W.; Hsieh, Y. T.; Wang, Y. F.; Yen, C. C.; Lee, G. L.; Yeh, C. C.; Peng, Y. J.; Kuo, Y. Y.; Wen, H. T.; et al. Cancer-Derived Succinate Promotes Macrophage Polarization and Cancer Metastasis via Succinate Receptor. *Molecular cell* **2020**, *77* (2), 213-227 e215. DOI: 10.1016/j.molcel.2019.10.023.

(110) Andrzejewski, S.; Klimcakova, E.; Johnson, R. M.; Tabaries, S.; Annis, M. G.; McGuirk, S.; Northey, J. J.; Chenard, V.; Sriram, U.; Papadopoli, D. J.; et al. PGC-1alpha Promotes Breast Cancer Metastasis and Confers Bioenergetic Flexibility against Metabolic Drugs. *Cell metabolism* **2017**, *26* (5), 778-787 e775. DOI: 10.1016/j.cmet.2017.09.006.

(111) Huang, F.; Ni, M.; Chalishazar, M. D.; Huffman, K. E.; Kim, J.; Cai, L.; Shi, X.; Cai, F.; Zacharias, L. G.; Ireland, A. S.; et al. Inosine Monophosphate Dehydrogenase Dependence in a Subset of Small Cell Lung Cancers. *Cell metabolism***2018**, *28* (3), 369-382 e365. DOI: 10.1016/j.cmet.2018.06.005.

(112) Obrist, F.; Michels, J.; Durand, S.; Chery, A.; Pol, J.; Levesque, S.; Joseph, A.; Astesana, V.; Pietrocola, F.; Wu, G. S.; et al. Metabolic vulnerability of cisplatin-resistant cancers. *The EMBO journal* **2018**, *37* (14). DOI: 10.15252/embj.201798597.

(113) Patel, S. B.; Nemkov, T.; Stefanoni, D.; Benavides, G. A.; Bassal, M. A.; Crown, B. L.; Matkins, V. R.; Camacho, V.; Kuznetsova, V.; Hoang, A. T.; et al. Metabolic alterations mediated by STAT3 promotes drug persistence in CML. *Leukemia* **2021**, *35* (12), 3371-3382. DOI: 10.1038/s41375-021-01315-0.

(114) Xiao, Y.; Ma, D.; Yang, Y. S.; Yang, F.; Ding, J. H.; Gong, Y.; Jiang, L.; Ge, L. P.; Wu, S. Y.; Yu, Q.; et al. Comprehensive metabolomics expands precision medicine for triple-negative breast cancer. *Cell research* **2022**, *32* (5), 477-490. DOI: 10.1038/s41422-022-00614-0.

(115) Pushpakom, S.; Iorio, F.; Eyers, P. A.; Escott, K. J.; Hopper, S.; Wells, A.; Doig, A.; Guilliams, T.; Latimer, J.; McNamee, C.; et al. Drug repurposing: progress, challenges and recommendations. *Nature reviews. Drug discovery* **2019**, *18* (1), 41-58. DOI: 10.1038/nrd.2018.168.

(116) Naffouje, R.; Grover, P.; Yu, H.; Sendilnathan, A.; Wolfe, K.; Majd, N.; Smith, E. P.; Takeuchi, K.; Senda, T.; Kofuji, S.; et al. Anti-Tumor Potential of IMP Dehydrogenase Inhibitors: A Century-Long Story. *Cancers (Basel)* **2019**, *11* (9). DOI: 10.3390/cancers11091346.

(117) Mullen, N. J.; Singh, P. K. Nucleotide metabolism: a pan-cancer metabolic dependency. *Nature Reviews Cancer* **2023**, *23* (5), 275-294. DOI: 10.1038/s41568-023-00557-7.

(118) Alghandour, R.; Ebrahim, M. A.; Elshal, A. M.; Ghobrial, F.; Elzaafarany, M.; MA, E. L. Repurposing metformin as anticancer drug: Randomized controlled trial in advanced prostate cancer (MANSMED). *Urologic Oncology: Seminars and Original Investigations. Vol. 39. No. 12.* **2021**, *39* (12), 831 e831-831 e810. DOI: 10.1016/j.urolonc.2021.05.020.

(119) Lord, S. R.; Harris, A. L. Is it still worth pursuing the repurposing of metformin as a cancer therapeutic? *British Journal of Cancer* **2023**, *128* (6), 958-966. DOI: 10.1038/s41416-023-02204-2.

(120) Pacal, L.; Kankova, K. Metformin in Oncology - How Far Is Its Repurposing as an Anticancer Drug? *Klinicka onkologie : casopis Ceske a Slovenske onkologicke spolecnosti vol. 33,2* **2020**, *33* (2), 107-113. DOI: 10.14735/amko2020107.

(121) Kim, H. W. Metabolomic Approaches to Investigate the Effect of Metformin: An Overview. *International journal of molecular sciences* **2021**, *22* (19). DOI: 10.3390/ijms221910275.

(122) Liu, X.; Romero, I. L.; Litchfield, L. M.; Lengyel, E.; Locasale, J. W. Metformin Targets Central Carbon Metabolism and Reveals Mitochondrial Requirements in Human Cancers. *Cell metabolism* **2016**, *24* (5), 728-739. DOI: 10.1016/j.cmet.2016.09.005.

(123) Liberti, M. V.; Dai, Z.; Wardell, S. E.; Baccile, J. A.; Liu, X.; Gao, X.; Baldi, R.; Mehrmohamadi, M.; Johnson, M. O.; Madhukar, N. S.; et al. A Predictive Model for Selective Targeting of the Warburg Effect through GAPDH Inhibition with a Natural Product. *Cell metabolism* **2017**, *26* (4), 648-659 e648. DOI: 10.1016/j.cmet.2017.08.017.

(124) Halbrook, C. J.; Pontious, C.; Kovalenko, I.; Lapienyte, L.; Dreyer, S.; Lee, H. J.; Thurston, G.; Zhang, Y.; Lazarus, J.; Sajjakulnukit, P.; et al. Macrophage-Released Pyrimidines Inhibit Gemcitabine Therapy in Pancreatic Cancer. *Cell metabolism* **2019**, *29* (6), 1390-1399 e1396. DOI: 10.1016/j.cmet.2019.02.001.

(125) Zhao, Q.; Zhang, T.; Xiao, X. R.; Huang, J. F.; Wang, Y.; Gonzalez, F. J.; Li, F. Impaired clearance of sunitinib leads to metabolic disorders and hepatotoxicity. *British journal of pharmacology* **2019**, *176* (13), 2162-2178. DOI: 10.1111/bph.14664.

(126) Lu, X.; Hackman, G. L.; Saha, A.; Rathore, A. S.; Collins, M.; Friedman, C.; Yi, S. S.; Matsuda, F.; DiGiovanni, J.; Lodi, A.; et al. Metabolomics-based phenotypic screens for evaluation of drug synergy via direct-infusion mass spectrometry. *iScience* **2022**, *25* (5), 104221. DOI: 10.1016/j.isci.2022.104221.

(127) Lu, X.; Han, L.; Busquets, J.; Collins, M.; Lodi, A.; Marszalek, J. R.; Konopleva, M.; Tiziani, S. The Combined Treatment With the FLT3-Inhibitor AC220 and the Complex I Inhibitor IACS-010759

Synergistically Depletes Wt- and FLT3-Mutated Acute Myeloid Leukemia Cells. *Frontiers in Oncology* **2021**, *11*, 686765. DOI: 10.3389/fonc.2021.686765.

(128) Gravel, S. P.; Avizonis, D.; St-Pierre, J. Metabolomics Analyses of Cancer Cells in Controlled Microenvironments. *The Tumor Microenvironment: Methods and Protocols* **2016**, *1458*, 273-290. DOI: 10.1007/978-1-4939-3801-8_20.

(129) Rabinowitz, J. D.; Purdy, J. G.; Vastag, L.; Shenk, T.; Koyuncu, E. Metabolomics in drug target discovery. *Cold Spring Harbor symposia on quantitative biology. Vol. 76. Cold Spring Harbor Laboratory Press* **2011**, *76*, 235-246. DOI: 10.1101/sqb.2011.76.010694.

(130) Puchades-Carrasco, L.; Pineda-Lucena, A. Metabolomics in pharmaceutical research and development. *Current opinion in biotechnology* **2015**, *35*, 73-77. DOI: 10.1016/j.copbio.2015.04.004.

(131) Zampieri, M.; Szappanos, B.; Buchieri, M. V.; Trauner, A.; Piazza, I.; Picotti, P.; Gagneux, S.; Borrell, S.; Gicquel, B.; Lelievre, J.; et al. High-throughput metabolomic analysis predicts mode of action of uncharacterized antimicrobial compounds. *Science translational medicine* **2018**, *10* (429). DOI: 10.1126/scitranslmed.aal3973.

(132) Campos, A. I.; Zampieri, M. Metabolomics-Driven Exploration of the Chemical Drug Space to Predict Combination Antimicrobial Therapies. *Molecular cell* **2019**, *74* (6), 1291-+. DOI: 10.1016/j.molcel.2019.04.001.

(133) Vincent, I. M.; Ehmann, D. E.; Mills, S. D.; Perros, M.; Barrett, M. P. Untargeted Metabolomics To Ascertain Antibiotic Modes of Action. *Antimicrobial agents and chemotherapy* **2016**, *60* (4), 2281-2291. DOI: 10.1128/Aac.02109-15.

(134) Anglada-Girotto, M.; Handschin, G.; Ortmayr, K.; Campos, A. I.; Gillet, L.; Manfredi, P.; Mulholland, C. V.; Berney, M.; Jenal, U.; Picotti, P.; et al. Combining CRISPRi and metabolomics for functional annotation of compound libraries. *Nature chemical biology* **2022**, *18* (5), 482-491. DOI: 10.1038/s41589-022-00970-3.

(135) Zhao, W.; Li, J.; Chen, M. M.; Luo, Y.; Ju, Z.; Nesser, N. K.; Johnson-Camacho, K.; Boniface, C. T.; Lawrence, Y.; Pande, N. T.; et al. Large-Scale Characterization of Drug Responses of Clinically Relevant Proteins in Cancer Cell Lines. *Cancer Cell* **2020**, *38* (6), 829-843 e824. DOI: 10.1016/j.ccell.2020.10.008.

(136) Liu, Z.; Jiang, S.; Hao, B.; Xie, S.; Liu, Y.; Huang, Y.; Xu, H.; Luo, C.; Huang, M.; Tan, M.; et al. A proteomic landscape of pharmacologic perturbations for functional relevance. *Journal of Pharmaceutical Analysis* **2024**, *14* (1), 128-139. DOI: 10.1016/j.jpha.2023.08.021.

(137) Mitchell, D. C.; Kuljanin, M.; Li, J.; Van Vranken, J. G.; Bulloch, N.; Schweppe, D. K.; Huttlin, E. L.; Gygi, S. P. A proteome-wide atlas of drug mechanism of action. *Nature biotechnology* **2023**, *41* (6), 845-857. DOI: 10.1038/s41587-022-01539-0.

(138) Ruprecht, B.; Di Bernardo, J.; Wang, Z.; Mo, X.; Ursu, O.; Christopher, M.; Fernandez, R. B.; Zheng, L.; Dill, B. D.; Wang, H.; et al. A mass spectrometry-based proteome map of drug action in lung cancer cell lines. *Nature Chemical Biology* **2020**, *16* (10), 1111-1119. DOI: 10.1038/s41589-020-0572-3.

(139) Bol, D.; Ebner, R. Gene expression profiling in the discovery, optimization and development of novel drugs: one universal screening platform. *Pharmacogenomics* **2006**, *7* (2), 227-235. DOI: 10.2217/14622416.7.2.227.

(140) Verbist, B.; Klambauer, G.; Vervoort, L.; Talloen, W.; Consortium, Q.; Shkedy, Z.; Thas, O.; Bender, A.; Gohlmann, H. W.; Hochreiter, S. Using transcriptomics to guide lead optimization in drug discovery projects: Lessons learned from the QSTAR project. *Drug Discovery Today* **2015**, *20* (5), 505-513. DOI: 10.1016/j.drudis.2014.12.014.

(141) Lamb, J.; Crawford, E. D.; Peck, D.; Modell, J. W.; Blat, I. C.; Wrobel, M. J.; Lerner, J.; Brunet, J. P.; Subramanian, A.; Ross, K. N.; et al. The Connectivity Map: using gene-expression signatures to connect

small molecules, genes, and disease. *Science* **2006**, *313* (5795), 1929-1935. DOI: 10.1126/science.1132939.

(142) Subramanian, A.; Narayan, R.; Corsello, S. M.; Peck, D. D.; Natoli, T. E.; Lu, X.; Gould, J.; Davis, J. F.; Tubelli, A. A.; Asiedu, J. K.; et al. A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* **2017**, *171* (6), 1437-1452 e1417. DOI: 10.1016/j.cell.2017.10.049.

(143) Li, Y.; Fan, Z.; Rao, J.; Chen, Z.; Chu, Q.; Zheng, M.; Li, X. An overview of recent advances and challenges in predicting compound-protein interaction (CPI). *Medical Review* **2023**, *3* (6), 465-486. DOI: 10.1515/mr-2023-0030.

(144) Galal, A.; Talal, M.; Moustafa, A. Applications of machine learning in metabolomics: Disease modeling and classification. *Frontiers in genetics* **2022**, *13*, 1017340. DOI: 10.3389/fgene.2022.1017340.

(145) Dias-Audibert, F. L.; Navarro, L. C.; de Oliveira, D. N.; Delafiori, J.; Melo, C.; Guerreiro, T. M.; Rosa, F. T.; Petenuci, D. L.; Watanabe, M. A. E.; Velloso, L. A.; et al. Combining Machine Learning and Metabolomics to Identify Weight Gain Biomarkers. *Frontiers in bioengineering and biotechnology* **2020**, *8*, 6. DOI: 10.3389/fbioe.2020.00006.

(146) Dhall, D.; Kaur, R.; Juneja, M. Machine Learning: A Review of the Algorithms and Its Applications. Proceedings of ICRIC 2019: Recent innovations in computing (2020): 47-63.

(147) Geroldinger, A.; Lusa, L.; Nold, M.; Heinze, G. Leave-one-out cross-validation, penalization, and differential bias of some prediction model performance measures-a simulation study. *Diagnostic and Prognostic Research* **2023**, *7* (1), 9. DOI: 10.1186/s41512-023-00146-0.

(148) Ghosh, T.; Zhang, W.; Ghosh, D.; Kechris, K. Predictive Modeling for Metabolomics Data. *Computational methods and data analysis for metabolomics* **2020**, *2104*, 313-336. DOI: 10.1007/978-1-0716-0239-3_16.

(149) Ho, T. K. The random subspace method for constructing decision forests. *IEEE transactions on pattern analysis and machine intelligence* **1998**, *20* (8), 832-844.

(150) Liu, J. J.; Shen, W. B.; Qin, Q. R.; Li, J. W.; Li, X.; Liu, M. Y.; Hu, W. L.; Wu, Y. Y.; Huang, F. Prediction of positive pulmonary nodules based on machine learning algorithm combined with central carbon metabolism data. *Journal of Cancer Research and Clinical Oncology* **2024**, *150* (2), 33. DOI: 10.1007/s00432-024-05610-y.

(151) Acharjee, A.; Larkman, J.; Xu, Y.; Cardoso, V. R.; Gkoutos, G. V. A random forest based biomarker discovery and power analysis framework for diagnostics research. *BMC medical genomics* **2020**, *13* (1), 178. DOI: 10.1186/s12920-020-00826-6.

(152) Chardin, D.; Gille, C.; Pourcher, T.; Humbert, O.; Barlaud, M. Learning a confidence score and the latent space of a new supervised autoencoder for diagnosis and prognosis in clinical metabolomic studies. *BMC Bioinformatics* **2022**, *23* (1), 361. DOI: 10.1186/s12859-022-04900-x.

(153) Chen, T.; Cao, Y.; Zhang, Y.; Liu, J.; Bao, Y.; Wang, C.; Jia, W.; Zhao, A. Random forest in clinical metabolomics for phenotypic discrimination and biomarker selection. *Evidence-Based Complementary and Alternative Medicine* **2013**, *2013*, 298183. DOI: 10.1155/2013/298183.

(154) Tibshirani, R. Regression shrinkage and selection via the Lasso. *ournal of the Royal Statistical Society Series B: Statistical Methodology* **1996**, *58* (1), 267-288. DOI: 10.1111/j.2517-6161.1996.tb02080.x.

(155) McEligot, A. J.; Poynor, V.; Sharma, R.; Panangadan, A. Logistic LASSO Regression for Dietary Intakes and Breast Cancer. *Nutrients* **2020**, *12* (9). DOI: 10.3390/nu12092652.

(156) Chen, Y.; Wang, B.; Zhao, Y.; Shao, X.; Wang, M.; Ma, F.; Yang, L.; Nie, M.; Jin, P.; Yao, K.; et al. Metabolomic machine learning predictor for diagnosis and prognosis of gastric cancer. *Nature Communications* **2024**, *15* (1), 1657. DOI: 10.1038/s41467-024-46043-y.

(157) Hajian-Tilaki, K. Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. *Caspian journal of internal medicine* **2013**, *4* (2), 627-635.

(158) Sun, R.; Fei, F.; Wang, M.; Jiang, J.; Yang, G.; Yang, N.; Jin, D.; Xu, Z.; Cao, B.; Li, J. Integration of metabolomics and machine learning revealed tryptophan metabolites are sensitive biomarkers of pemetrexed efficacy in non-small cell lung cancer. *Cancer Medicine* **2023**. DOI: 10.1002/cam4.6446.

(159) Yao, J. Z.; Tsigelny, I. F.; Kesari, S.; Kouznetsova, V. L. Diagnostics of ovarian cancer via metabolite analysis and machine learning. *Integrative Biology* **2023**, *15*. DOI: 10.1093/intbio/zyad005.

(160) Dube, M.; Llanes, D.; Saoud, M.; Rennert, R.; Imming, P.; Haberli, C.; Keiser, J.; Arnold, N. Albatrellus confluens (Alb. & Schwein.) Kotl. & Pouz.: Natural Fungal Compounds and Synthetic Derivatives with In Vitro Anthelmintic Activities and Antiproliferative Effects against Two Human Cancer Cell Lines. *Molecules* **2022**, *27* (9). DOI: 10.3390/molecules27092950.

(161) Pang, Z.; Chong, J.; Zhou, G.; de Lima Morais, D. A.; Chang, L.; Barrette, M.; Gauthier, C.; Jacques, P. E.; Li, S.; Xia, J. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic acids research* **2021**, *49* (W1), W388-W396. DOI: 10.1093/nar/gkab382.

(162) Witting, M.; Maier, T. V.; Garvis, S.; Schmitt-Kopplin, P. Optimizing a ultrahigh pressure liquid chromatography-time of flight-mass spectrometry approach using a novel sub-2 mu m core-shell particle for in depth lipidomic profiling of Caenorhabditis elegans. *Journal of Chromatography A* **2014**, *1359*, 91-99. DOI: 10.1016/j.chroma.2014.07.021.

(163) Tsugawa, H.; Ikeda, K.; Takahashi, M.; Satoh, A.; Mori, Y.; Uchino, H.; Okahashi, N.; Yamada, Y.; Tada, I.; Bonini, P.; et al. A lipidome atlas in MS-DIAL 4. *Nature biotechnology* **2020**, *38* (10), 1159-+. DOI: 10.1038/s41587-020-0531-2.

(164) Balcke, G. U.; Vahabi, K.; Giese, J.; Finkemeier, I.; Tissier, A. Coordinated metabolic adaptation of Arabidopsis thaliana to high light. *The Plant Journal* **2024**. DOI: 10.1111/tpj.16992.

(165) Krieger, E.; Vriend, G. YASARA View-molecular graphics for all devices-from smartphones to workstations. *Bioinformatics* **2014**, *30* (20), 2981-2982. DOI: 10.1093/bioinformatics/btu426.

(166) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *Journal of computational chemistry* **2009**, *30* (16), 2785-2791. DOI: 10.1002/jcc.21256.

(167) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB Journal of chemical theory and computation **2015**, *11* (8), 3696-3713. DOI: 10.1021/acs.jctc.5b00255.

(168) Wang, Z. H.; Yang, M.; Yang, Y. F.; He, Y. L.; Qian, H. W. Structural basis for catalysis of human choline/ethanolamine phosphotransferase 1. *Nature Communications* **2023**, *14* (1). DOI: 10.1038/s41467-023-38290-2.

(169) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583-+. DOI: 10.1038/s41586-021-03819-2.

(170) Hekkelman, M. L.; de Vries, I.; Joosten, R. P.; Perrakis, A. AlphaFill: enriching AlphaFold models with ligands and cofactors. *Nature Methods* **2023**, *20* (2), 205-+. DOI: 10.1038/s41592-022-01685-y.

(171) (MOE), M. O. E. Chemical Computing Group ULC, 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2020. **2020.09**, 1-7.

(172) Kolde, R. pheatmap: Pretty Heatmaps. https://CRAN.R-project.org/package=pheatmap, 2019; Vol. R package version 1.0.12.

(173) Liaw, A.; Wiener, M. Classification and Regression by randomForest. *R News* **2002**, *2* (3), 18-22.

(174) Friedman, J.; Hastie, T.; Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of statistical software* **2010**, *33* (1), 1-22. DOI: DOI 10.18637/jss.v033.i01.

(175) Venables, W. N.; Ripley, B. D. *Modern Applied Statistics with S*; Springer, 2002.

(176) Soltani, M.; Zhao, Y.; Xia, Z.; Ganjalikhani Hakemi, M.; Bazhin, A. V. The Importance of Cellular Metabolic Pathways in Pathogenesis and Selective Treatments of Hematological Malignancies. *Frontiers in Oncology* **2021**, *11*, 767026. DOI: 10.3389/fonc.2021.767026.

(177) Zhu, J.; Thompson, C. B. Metabolic regulation of cell growth and proliferation. *Nature reviews Molecular cell biology* **2019**, *20* (7), 436-450. DOI: 10.1038/s41580-019-0123-5.

(178) Muschet, C.; Moller, G.; Prehn, C.; de Angelis, M. H.; Adamski, J.; Tokarz, J. Removing the bottlenecks of cell culture metabolomics: fast normalization procedure, correlation of metabolites to cell number, and impact of the cell harvesting method. *Metabolomics* **2016**, *12* (10), 151. DOI: 10.1007/s11306-016-1104-8.

(179) Sun, J.; Xia, Y. Pretreating and normalizing metabolomics data for statistical analysis. *Genes & Diseases* **2024**, *11* (3), 100979. DOI: 10.1016/j.gendis.2023.04.018 From NLM PubMed-not-MEDLINE.

(180) Kluckova, K.; Bezawork-Geleta, A.; Rohlena, J.; Dong, L. F.; Neuzil, J. Mitochondrial complex II, a novel target for anti-cancer agents. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **2013**, *1827* (5), 552-564. DOI: 10.1016/j.bbabio.2012.10.015.

(181) Wojtovich, A. P.; Smith, C. O.; Haynes, C. M.; Nehrke, K. W.; Brookes, P. S. Physiological consequences of complex II inhibition for aging, disease, and the mKATP channel. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **2013**, *1827* (5), 598-611. DOI: 10.1016/j.bbabio.2012.12.007.

(182) Fontaine, E. Metformin-Induced Mitochondrial Complex I Inhibition: Facts, Uncertainties, and Consequences. *Frontiers in endocrinology* **2018**, *9*, 753. DOI: 10.3389/fendo.2018.00753.

(183) Wang, L.; Duan, Q.; Wang, T.; Ahmed, M.; Zhang, N.; Li, Y.; Li, L.; Yao, X. Mitochondrial Respiratory Chain Inhibitors Involved in ROS Production Induced by Acute High Concentrations of Iodide and the Effects of SOD as a Protective Factor. *Oxidative medicine and cellular longevity* **2015**, *2015*, 217670. DOI: 10.1155/2015/217670.

(184) Owens, K. M.; Kulawiec, M.; Desouki, M. M.; Vanniarajan, A.; Singh, K. K. Impaired OXPHOS complex III in breast cancer. *PLoS One* **2011**, *6* (8), e23846. DOI: 10.1371/journal.pone.0023846.

(185) Hargreaves, I. P.; Duncan, A. J.; Wu, L.; Agrawal, A.; Land, J. M.; Heales, S. J. Inhibition of mitochondrial complex IV leads to secondary loss complex II-III activity: implications for the pathogenesis and treatment of mitochondrial encephalomyopathies. *Mitochondrion* **2007**, *7* (4), 284-287. DOI: 10.1016/j.mito.2007.02.001.

(186) Demine, S.; Renard, P.; Arnould, T. Mitochondrial Uncoupling: A Key Controller of Biological Processes in Physiology and Diseases. *Cells* **2019**, *8* (8). DOI: 10.3390/cells8080795.

(187) Terada, H. Uncouplers of oxidative phosphorylation. *Environmental health perspectives* **1990**, *87*, 213-218. DOI: 10.1289/ehp.9087213.

(188) Goulitquer, S.; Croyal, M.; Lalande, J.; Royer, A. L.; Guitton, Y.; Arzur, D.; Durand, S.; Le Jossic-Corcos, C.; Bouchereau, A.; Potin, P.; et al. Consequences of blunting the mevalonate pathway in cancer identified by a pluriomics approach. *Cell Death & Disease* **2018**, *9*. DOI: 10.1038/s41419-018-0761-0.

(189) Qi, C.; Wang, X.; Shen, Z.; Chen, S.; Yu, H.; Williams, N.; Wang, G. Anti-mitotic chemotherapeutics promote apoptosis through TL1A-activated death receptor 3 in cancer cells. *Cell research* **2018**, *28* (5), 544-555. DOI: 10.1038/s41422-018-0018-6.

(190) Weaver, B. A. How Taxol/paclitaxel kills cancer cells. *Molecular biology of the cell* **2014**, *25* (18), 2677-2681. DOI: 10.1091/mbc.E14-04-0916.

(191) Nitiss, J. L. Targeting DNA topoisomerase II in cancer chemotherapy. *Nature Reviews Cancer* **2009**, *9* (5), 338-350. DOI: 10.1038/nrc2607.

(192) Talukdar, A.; Kundu, B.; Sarkar, D.; Goon, S.; Mondal, M. A. Topoisomerase I inhibitors: Challenges, progress and the road ahead. *European Journal of Medicinal Chemistry* **2022**, *236*, 114304. DOI: 10.1016/j.ejmech.2022.114304.

(193) Navas, L. E.; Carnero, A. Nicotinamide Adenine Dinucleotide (NAD) Metabolism as a Relevant Target in Cancer. *Cells* **2022**, *11* (17). DOI: 10.3390/cells11172627.

(194) Bian, Y.; Hou, W.; Chen, X.; Fang, J.; Xu, N.; Ruan, B. H. Glutamate Dehydrogenase as a Promising Target for Hyperinsulinism Hyperammonemia Syndrome Therapy. *Current Medicinal Chemistry* **2022**, *29* (15), 2652-2672. DOI: 10.2174/0929867328666210825105342.

(195) Vasan, N.; Cantley, L. C. At a crossroads: how to translate the roles of PI3K in oncogenic and metabolic signalling into improvements in cancer therapy. *Nature Reviews Clinical Oncology* **2022**, *19* (7), 471-485. DOI: 10.1038/s41571-022-00633-1.

(196) Peng, Y.; Wang, Y.; Zhou, C.; Mei, W.; Zeng, C. PI3K/Akt/mTOR Pathway and Its Role in Cancer Therapeutics: Are We Making Headway? *Frontiers in oncology* **2022**, *12*, 819128. DOI: 10.3389/fonc.2022.819128.

(197) Gills, J. J.; Dennis, P. A. Perifosine: update on a novel Akt inhibitor. *Current oncology reports* **2009**, *11* (2), 102-110. DOI: 10.1007/s11912-009-0016-4.

(198) Song, M.; Liu, X.; Liu, K.; Zhao, R.; Huang, H.; Shi, Y.; Zhang, M.; Zhou, S.; Xie, H.; Chen, H.; et al. Targeting AKT with Oridonin Inhibits Growth of Esophageal Squamous Cell Carcinoma In Vitro and Patient-Derived Xenografts In Vivo. *Molecular cancer therapeutics* **2018**, *17* (7), 1540-1553. DOI: 10.1158/1535-7163.MCT-17-0823.

(199) Tyson, R. L.; Perron, J.; Sutherland, G. R. 6-Aminonicotinamide inhibition of the pentose phosphate pathway in rat neocortex. *Neuroreport* **2000**, *11* (9), 1845-1848. DOI: 10.1097/00001756-200006260-00009.

(200) Ghergurovich, J. M.; Garcia-Canaveras, J. C.; Wang, J. S.; Schmidt, E.; Zhang, Z. Y.; TeSlaa, T.; Patel, H.; Chen, L.; Britt, E. C.; Piqueras-Nebot, M.; et al. A small molecule G6PD inhibitor reveals immune dependence on pentose phosphate pathway. *Nature chemical biology* **2020**, *16* (7), 731-+. DOI: 10.1038/s41589-020-0533-x.

(201) Yadav, V. R.; Prasad, S.; Sung, B.; Kannappan, R.; Aggarwal, B. B. Targeting inflammatory pathways by triterpenoids for prevention and treatment of cancer. *Toxins (Basel)* **2010**, *2* (10), 2428-2466. DOI: 10.3390/toxins2102428.

(202) Pan, M. H.; Lin, C. C.; Lin, J. K.; Chen, W. J. Tea polyphenol (-)-epigallocatechin 3-gallate suppresses heregulin-beta1-induced fatty acid synthase expression in human breast cancer cells by inhibiting phosphatidylinositol 3-kinase/Akt and mitogen-activated protein kinase cascade signaling. *Journal of Agricultural and Food Chemistry* **2007**, *55* (13), 5030-5037. DOI: 10.1021/jf070316r.

(203) Brusselmans, K.; Vrolix, R.; Verhoeven, G.; Swinnen, J. V. Induction of cancer cell apoptosis by flavonoids is associated with their ability to inhibit fatty acid synthase activity. *Journal of Biological Chemistry* **2005**, *280* (7), 5636-5645. DOI: 10.1074/jbc.M408177200.

(204) Juarez, D.; Fruman, D. A. Targeting the Mevalonate Pathway in Cancer. *Trends in cancer* **2021**, *7* (6), 525-540. DOI: 10.1016/j.trecan.2020.11.008.

(205) Pavlova, N. N.; Thompson, C. B. The Emerging Hallmarks of Cancer Metabolism. *Cell metabolism* **2016**, *23* (1), 27-47. DOI: 10.1016/j.cmet.2015.12.006.

(206) Martinez-Reyes, I.; Cardona, L. R.; Kong, H.; Vasan, K.; McElroy, G. S.; Werner, M.; Kihshen, H.; Reczek, C. R.; Weinberg, S. E.; Gao, P.; et al. Mitochondrial ubiquinol oxidation is necessary for tumour growth. *Nature* **2020**, *585* (7824), 288-+. DOI: 10.1038/s41586-020-2475-6.

(207) Bartman, C. R.; Weilandt, D. R.; Shen, Y. H.; Lee, W. D.; Han, Y. J.; TeSlaa, T.; Jankowski, C. S. R.; Samarah, L.; Park, N. R.; da Silva-Diz, V.; et al. Slow TCA flux and ATP production in primary solid tumours but not metastases. *Nature* **2023**. DOI: 10.1038/s41586-022-05661-6.

(208) Bajzikova, M.; Kovarova, J.; Coelho, A. R.; Boukalova, S.; Oh, S.; Rohlenova, K.; Svec, D.; Hubackova, S.; Endaya, B.; Judasova, K.; et al. Reactivation of Dihydroorotate Dehydrogenase-Driven Pyrimidine Biosynthesis Restores Tumor Growth of Respiration-Deficient Cancer Cells. *Cell metabolism* **2019**, *29* (2), 399-416 e310. DOI: 10.1016/j.cmet.2018.10.014.

(209) Banerjee, R.; Purhonen, J.; Kallijarvi, J. The mitochondrial coenzyme Q junction and complex III: biochemistry and pathophysiology. *The FEBS journal* **2022**, *289* (22), 6936-6958. DOI: 10.1111/febs.16164.

(210) Liu, S. S.; Fu, S.; Wang, G. D.; Cao, Y.; Li, L. L.; Li, X. M.; Yang, J.; Li, N.; Shan, Y. B.; Cao, Y.; et al. Glycerol-3-phosphate biosynthesis regenerates cytosolic NAD(+) to alleviate mitochondrial disease. *Cell metabolism* **2021**, *33* (10), 1974-+. DOI: 10.1016/j.cmet.2021.06.013.

(211) Ghante, M. H.; Jamkhande, P. G. Role of Pentacyclic Triterpenoids in Chemoprevention and Anticancer Treatment: An Overview on Targets and Underling Mechanisms. *Journal of pharmacopuncture* **2019**, *22* (2), 55-67. DOI: 10.3831/KPI.201.22.007.

(212) Petrenko, M.; Guttler, A.; Pfluger, E.; Serbian, I.; Kahnt, M.; Eiselt, Y.; Kessler, J.; Funtan, A.; Paschke, R.; Csuk, R.; et al. MSBA-S - A pentacyclic sulfamate as a new option for radiotherapy of human breast cancer cells. *European Journal of Medicinal Chemistry* **2021**, *224*. DOI: 10.1016/j.ejmech.2021.113721.

(213) Verstraeten, S.; Catteau, L.; Boukricha, L.; Quetin-Leclercq, J.; Mingeot-Leclercq, M. P. Effect of Ursolic and Oleanolic Acids on Lipid Membranes: Studies on MRSA and Models of Membranes. *Antibiotics-Basel* **2021**, *10* (11). DOI: 10.3390/antibiotics10111381.

(214) Li, S.; Wu, R.; Wang, L.; Dina Kuo, H. C.; Sargsyan, D.; Zheng, X.; Wang, Y.; Su, X.; Kong, A. N. Triterpenoid ursolic acid drives metabolic rewiring and epigenetic reprogramming in treatment/prevention of human prostate cancer. *Molecular carcinogenesis* **2022**, *61* (1), 111-121. DOI: 10.1002/mc.23365.

(215) Kennedy, B. E.; Sharif, T.; Martell, E.; Dai, C.; Kim, Y.; Lee, P. W.; Gujar, S. A. NAD(+) salvage pathway in cancer metabolism and therapy. *Pharmacological research* **2016**, *114*, 274-283. DOI: 10.1016/j.phrs.2016.10.027.

(216) Tan, B.; Dong, S. C.; Shepard, R. L.; Kays, L.; Roth, K. D.; Geeganage, S.; Kuo, M. S.; Zhao, G. S. Inhibition of Nicotinamide Phosphoribosyltransferase (NAMPT), an Enzyme Essential for NAD(+) Biosynthesis, Leads to Altered Carbohydrate Metabolism in Cancer Cells. *Journal of Biological Chemistry* **2015**, *290* (25), 15812-15824. DOI: 10.1074/jbc.M114.632141.

(217) Kohler, E.; Barrach, H.; Neubert, D. Inhibition of NADP dependent oxidoreductases by the 6-aminonicotinamide analogue of NADP. *FEBS letters* **1970**, *6* (3), 225-228. DOI: 10.1016/0014-5793(70)80063-1.

(218) Street, J. C.; Alfieri, A. A.; Koutcher, J. A. Quantitation of metabolic and radiobiological effects of 6-aminonicotinamide in RIF-1 tumor cells in vitro. *Cancer research* **1997**, *57* (18), 3956-3962.

(219) Kraft, O.; Hartmann, A. K.; Brandt, S.; Hoenke, S.; Heise, N. V.; Csuk, R.; Mueller, T. Asiatic acid as a leading structure for derivatives combining sub-nanomolar cytotoxicity, high selectivity, and the ability to overcome drug resistance in human preclinical tumor models. *European Journal of Medicinal Chemistry* **2023**, *250*, 115189. DOI: 10.1016/j.ejmech.2023.115189.

(220) Wang, H.; Fang, B.; Peng, B.; Wang, L.; Xue, Y.; Bai, H.; Lu, S.; Voelcker, N. H.; Li, L.; Fu, L.; et al. Recent Advances in Chemical Biology of Mitochondria Targeting. *Frontiers in Chemistry***2021**, *9*, 683220. DOI: 10.3389/fchem.2021.683220.

(221) Rashan, L. J.; Franke, K.; Khine, M. M.; Kelter, G.; Fiebig, H. H.; Neumann, J.; Wessjohann, L. A. Characterization of the anticancer properties of monoglycosidic cardenolides isolated from Nerium oleander and Streptocaulon tomentosum. *Journal of ethnopharmacology* **2011**, *134* (3), 781-788. DOI: 10.1016/j.jep.2011.01.038.

(222) Lauber, K.; Bohn, E.; Krober, S. M.; Xiao, Y. J.; Blumenthal, S. G.; Lindemann, R. K.; Marini, P.; Wiedig, C.; Zobywalski, A.; Baksh, S.; et al. Apoptotic cells induce migration of phagocytes via caspase-3-mediated release of a lipid attraction signal. *Cell* **2003**, *113* (6), 717-730. DOI: 10.1016/S0092-8674(03)00422-7.

(223) Horvath, S. E.; Daum, G. Lipids of mitochondria. *Progress in lipid research* **2013**, *52* (4), 590-614. DOI: 10.1016/j.plipres.2013.07.002.

(224) Liepinsh, E.; Makrecka-Kuka, M.; Volska, K.; Kuka, J.; Makarova, E.; Antone, U.; Sevostjanovs, E.; Vilskersts, R.; Strods, A.; Tars, K.; et al. Long-chain acylcarnitines determine ischaemia/reperfusion-induced damage in heart mitochondria. *Biochemical Journal* **2016**, *473* (9), 1191-1202. DOI: 10.1042/BCJ20160164.

(225) Rashan, L. J.; Ozenver, N.; Boulos, J. C.; Dawood, M.; Roos, W. P.; Franke, K.; Papasotiriou, I.; Wessjohann, L. A.; Fiebig, H. H.; Efferth, T. Molecular Modes of Action of an Aqueous Nerium oleander Extract in Cancer Cells In Vitro and In Vivo. *Molecules* **2023**, *28* (4). DOI: 10.3390/molecules28041871.

(226) Batchuluun, B.; Pinkosky, S. L.; Steinberg, G. R. Lipogenesis inhibitors: therapeutic opportunities and challenges. *Nature Reviews Drug Discovery* **2022**, *21* (4), 283-305. DOI: 10.1038/s41573-021-00367-2.

(227) Montesdeoca, N.; Lopez, M.; Ariza, X.; Herrero, L.; Makowski, K. Inhibitors of lipogenic enzymes as a potential therapy against cancer. *The FASEB Journal* **2020**, *34* (9), 11355-11381. DOI: 10.1096/fj.202000705R.

(228) Wang, X.; Tian, W. Green tea epigallocatechin gallate: a natural inhibitor of fatty-acid synthase. *Biochemical and Biophysical Research Communications* **2001**, *288* (5), 1200-1206. DOI: 10.1006/bbrc.2001.5923.

(229) Gao, Y. J.; Islam, M. S.; Tian, J.; Lui, V. W. Y.; Xiao, D. Inactivation of ATP citrate lyase by Cucurbitacin B: A bioactive compound from cucumber, inhibits prostate cancer growth. *Cancer letters* **2014**, *349* (1), 15-25. DOI: 10.1016/j.canlet.2014.03.015.

(230) Trapotsi, M. A.; Hosseini-Gerami, L.; Bender, A. Computational analyses of mechanism of action (MoA): data, methods and integration. *RSC chemical biology* **2022**, *3* (2), 170-200. DOI: 10.1039/d1cb00069a.

(231) Belenky, P.; Ye, J. D.; Porter, C. B.; Cohen, N. R.; Lobritz, M. A.; Ferrante, T.; Jain, S.; Korry, B. J.; Schwarz, E. G.; Walker, G. C.; et al. Bactericidal Antibiotics Induce Toxic Metabolic Perturbations that Lead to Cellular Damage. *Cell reports* **2015**, *13* (5), 968-980. DOI: 10.1016/j.celrep.2015.09.059.

(232) Gielisch, I.; Meierhofer, D. Metabolome and proteome profiling of complex I deficiency induced by rotenone. *Journal of proteome research* **2015**, *14* (1), 224-235. DOI: 10.1021/pr500894v.

(233) Cherkaoui, S.; Durot, S.; Bradley, J.; Critchlow, S.; Dubuis, S.; Masiero, M. M.; Wegmann, R.; Snijder, B.; Othman, A.; Bendtsen, C.; et al. A functional analysis of 180 cancer cell lines reveals conserved intrinsic metabolic programs. *Molecular Systems Biology* **2022**, *18* (11), e11033. DOI: 10.15252/msb.202211033.

(234) Xu, Y. F.; Lu, W. Y.; Rabinowitz, J. D. Avoiding Misannotation of In-Source Fragmentation Products as Cellular Metabolites in Liquid Chromatography-Mass Spectrometry-Based Metabolomics. *Analytical chemistry* **2015**, *87* (4), 2273-2281. DOI: 10.1021/ac504118y.

(235) Spinelli, J. B.; Rosen, P. C.; Sprenger, H. G.; Puszynska, A. M.; Mann, J. L.; Roessler, J. M.; Cangelosi, A. L.; Henne, A.; Condon, K. J.; Zhang, T.; et al. Fumarate is a terminal electron acceptor in the mammalian electron transport chain. *Science* **2021**, *374* (6572), 1227-1237. DOI: 10.1126/science.abi7495.

(236) Mullen, A. R.; Hu, Z.; Shi, X.; Jiang, L.; Boroughs, L. K.; Kovacs, Z.; Boriack, R.; Rakheja, D.; Sullivan, L. B.; Linehan, W. M.; et al. Oxidation of alpha-ketoglutarate is required for reductive carboxylation in cancer cells with mitochondrial defects. *Cell reports* **2014**, *7* (5), 1679-1690. DOI: 10.1016/j.celrep.2014.04.037.

(237) Stefaniak, J.; Huber, K. V. M. Importance of Quantifying Drug-Target Engagement in Cells. *ACS Medicinal Chemistry Letters* **2020**, *11* (4), 403-406. DOI: 10.1021/acsmedchemlett.9b00570.

(238) Kciuk, M.; Gielecinska, A.; Mujwar, S.; Kolat, D.; Kaluzinska-Kolat, Z.; Celik, I.; Kontek, R. Doxorubicin-An Agent with Multiple Mechanisms of Anticancer Activity. *Cells* **2023**, *12* (4). DOI: 10.3390/cells12040659.

(239) Martino, F.; Lupi, M.; Giraudo, E.; Lanzetti, L. Breast cancers as ecosystems: a metabolic perspective. *Cellular and Molecular Life Sciences* **2023**, *80* (9), 244. DOI: 10.1007/s00018-023-04902-9.

(240) Magometschnigg, H.; Pinker, K.; Helbich, T.; Brandstetter, A.; Rudas, M.; Nakuz, T.; Baltzer, P.; Wadsak, W.; Hacker, M.; Weber, M.; et al. PIK3CA Mutational Status Is Associated with High Glycolytic Activity in ER+/HER2- Early Invasive Breast Cancer: a Molecular Imaging Study Using [(18)F]FDG PET/CT. *Molecular imaging and biology* **2019**, *21* (5), 991-1002. DOI: 10.1007/s11307-018-01308-z.

(241) Sadeghi, R. N.; Karami-Tehrani, F.; Salami, S. Targeting prostate cancer cell metabolism: impact of hexokinase and CPT-1 enzymes. *Tumor Biology***2015**, *36* (4), 2893-2905. DOI: 10.1007/s13277-014-2919-4.

(242) Krajnovic, T.; Kaluderovic, G. N.; Wessjohann, L. A.; Mijatovic, S.; Maksimovic-Ivanic, D. Versatile antitumor potential of isoxanthohumol: Enhancement of paclitaxel activity in vivo. *Pharmacological research* **2016**, *105*, 62-73. DOI: 10.1016/j.phrs.2016.01.011.

(243) Yoshida, L. S.; Kakegawa, T.; Yuda, Y.; Takano-Ohmuro, H. Shikonin changes the lipopolysaccharide-induced expression of inflammation-related genes in macrophages. *Journal of natural medicines* **2017**, *71* (4), 723-734. DOI: 10.1007/s11418-017-1106-5.

(244) Sabnis, R. W. Biaryl Compounds as Dihydroorotate Dehydrogenase Inhibitors for Treating Acute Myelogenous Leukemia (AML). *ACS Medicinal Chemistry Letters* **2022**, *13* (2), 158-159. DOI: 10.1021/acsmedchemlett.2c00017.

(245) Jiang, M.; Song, Y.; Liu, H.; Jin, Y.; Li, R.; Zhu, X. DHODH Inhibition Exerts Synergistic Therapeutic Effect with Cisplatin to Induce Ferroptosis in Cervical Cancer through Regulating mTOR Pathway. *Cancers (Basel)* **2023**, *15* (2). DOI: 10.3390/cancers15020546.

(246) Abohassan, M.; Al Shahrani, M.; Rajagopalan, P. Identification of second generation benzylidene chromanone analogues as novel, potent DHODH inhibitors in acute myeloid leukemia cells. *Journal of Biomolecular Structure and Dynamics* **2023**, *41* (13), 6168-6177. DOI: 10.1080/07391102.2022.2103031.

(247) So, J.; Lewis, A. C.; Smith, L. K.; Stanley, K.; Franich, R.; Yoannidis, D.; Pijpers, L.; Dominguez, P.; Hogg, S. J.; Vervoort, S. J.; et al. Inhibition of pyrimidine biosynthesis targets protein translation in acute myeloid leukemia. *EMBO Molecular Medicine* **2022**, *14* (7), e15203. DOI: 10.15252/emmm.202115203.

(248) Zheng, Y.; Li, S.; Song, K.; Ye, J.; Li, W.; Zhong, Y.; Feng, Z.; Liang, S.; Cai, Z.; Xu, K. A Broad Antiviral Strategy: Inhibitors of Human DHODH Pave the Way for Host-Targeting Antivirals against Emerging and Re-Emerging Viruses. *Viruses* **2022**, *14* (5). DOI: 10.3390/v14050928.

(249) Russo, T. A.; Umland, T. C.; Deng, X.; El Mazouni, F.; Kokkonda, S.; Olson, R.; Carlino-MacDonald, U.; Beanan, J.; Alvarado, C. L.; Tomchick, D. R.; et al. Repurposed dihydroorotate dehydrogenase inhibitors with efficacy against drug-resistant Acinetobacter baumannii. *Proceedings of the National Academy of Sciences* **2022**, *119* (51), e2213116119. DOI: 10.1073/pnas.2213116119.

(250) Peelen, E.; Muehler, A.; Vitt, D.; Kohlhof, H. IMU-838, a Small Molecule DHODH Inhibitor in Phase 2 Clinical Trial for Ulcerative Colitis, Shows Potent Anti-inflammatory Activity in Cell-Culture-Based and In Vivo Systems. *Journal of Crohn's and Colitis* **2022**, *16*, I172-I173.

8 Appendix

Appendix table 1: Analyzed CCEM metabolites

Name	Abbreviation	PubChem_ ID	Name	Abbreviation	PubChem_ ID
2',3'-cyclic AMP	2,3-cAMP	101812	D-glucuronic Acid	GLUCURON	94715
2-phosphoglycolic acid	2PG	529	glycine	GLY	750
glycerate 2-phosphate	2PGA	59	glycolic acid	GLYCOLATE	757
3',5'-cyclic AMP	3,5-cAMP	6076	glyoxylic acid	GLYOX	760
3-dehydroshikimate	3DeHy-SHIK	439774	guanosine-5'-monophosphate	GMP	135398631
3-hydroxypyruvic acid	3-OH-PYR	964	geranyl diphosphate	GPP	445995
3-phosphoglycerate	3PGA	439183	glutathione	GSH	124886
5-methyltetrahydrofolate	5Me-THF	135398561	glutathione disulfide	GSSG	65359
6-phosphogluconic acid	6PG	91493	guanosine triphosphate	GTP	135398633
8-hydroxycopalyl diphosphate	8OH-COPA-PP	56927957	guanosine	Guanosine	135398635
acetoacetyl-CoA	AcAc-CoA	92153	hexose_sum	Hexoses	
acetyl-CoA	AcCoA	444493	L-histidine	HIS	6274
aconitic acid	ACT	309	(2E)-4-hydroxy-3-methylbut-2-enyl diphosphate	HMBPP	21597501

Name	Abbreviation	PubChem_ID	Name	Abbreviation	PubChem_ID
adenosine	Adenosine	60961	hydroxymethylglutaryl-CoA	HMG-CoA	445127
adenylosuccinic acid	Adenylo-Suc	440122	inosine 5'-monophosphate	IMP	135398640
adenosine 5'-diphosphate	ADP	6022	D-myo-inositol trikisphosphates (sum)	IP3_sum	
ADP-glucose	ADP-glucose	16500	D-myo-inositol tetrakisphosphates (sum)	IP4_sum	25245205
ADP-ribose 2'-phosphate	ADPPR	439957	D-myo-inositol pentakisphosphates (sum)	IP5_sum	127296
ADP-D-ribose	ADPR	33576	phytic acid	IP6	890
AICA ribonucleotide	AICAR	65110	Isopentenyl pyrophosphate	IPPP	1195
5-aminoimidazole ribonucleotide	AIR	161500	isocitric acid	ISOCIT	1198
2-ketoglutaric acid	2-OG	51	isovaleryl-coenzyme A	ISOVAL-CoA	165435
L-alanine	ALA	5950	2-keto-3-deoxy-6-phosphogluconate	KDPG	3080745
allantoic acid	ALLANTOATE	203	lactic acid	LAC	612
allantoin	ALT	204	L-leucine+L-isoleucine (sum)	LEU+ILE	6106

Name	Abbreviation	PubChem_ ID	Name	Abbreviation	PubChem_ ID
adenosine 5'-monophosphate	AMP	6083	L-leucine+L-isoleucine (sum)	LEU+ILE	6106
L-arginine	ARG	6322	L-lysine	LYS	5962
ascorbic acid	ASC	54670067	malic acid	MAL	92824
L-asparagine	ASN	6267	malonyl-CoA	Malo-CoA	644066
L-aspartic acid	ASP	5960	2C-methyl-D-erythritol 2,4-cyclodiphosphate	MEcPP	126747
adenosine triphosphate	ATP	5957	3-methylcrotonyl CoA	MeCrot-CoA	99
ureidosuccinic acid	CA	279	methyl-D-erythritol Phosphate	MEP	443198
carboxyaminoimidazole ribotide	CAIR	165388	methionine	MET	876
3-ureidopropionic acid	CarbamoylAla	111	mevalonic acid	MEVA	439230
cytidine 5'-diphosphate	CDP	6132	mevalonate-P	MEVA5P	439400
CDP-choline	CDP-CHOL	13804	mevalonate-5PP	MEVA5PP	439418
4-diphosphocytidyl-2-C-methyl-D-erythritol	CDP-ME	46936756	methylmalonyl coa	MM-CoA	123909
4-CDP-2-C-methyl-D-erythritol 2-phosphate	CDP-MEP	443200	nicotinic acid	NA	938
3',5'-cyclic GMP	cGMP	135398570	nicotinate mononucleotide	NAAM	121991

Name	Abbreviation	PubChem_ ID	Name	Abbreviation	PubChem_ ID
chloride	chloride	312	NAD+	NAD+	5892
chorismic acid	CHOR	12039	NADH	NADH	439153
chorismic acid	CHOR-H2O	12039	NADP+	NADP+	5886
citric acid	CIT	311	NADPH	NADPH	5884
L-citrulline	CITRUL	9750	nicotinamide	nicotinamide	936
cytidine 5'-monophosphate	CMP	6131	nicotinamide mononucleotide	NMN	14180
coenzyme A	CoA	87642	L-ornithine	ORN	6262
creatine	CREA	586	orotic acid	ORO	967
phosphocreatine	CREA-P	9548602	orotidylic acid	ORO5P	160617
cytidine triphosphate	CTP	6176	oxaloacetic acid	Oxalic acid_1	970
L-cysteine	CYS	5862	oxaloacetic acid_2	Oxalic acid_2	970
L-cystine	CYSN	67678	D-pantothenic acid	PAN	6613
cytidine	Cytidine	6175	O-phosphoethanolamine	PEA	1015
2'-deoxyadenosine	dADO	13730	sum of valeric acid, isovaleric, anteisovaleric acid	pentanoates	
2'-deoxyadenosine-5'-diphosphate	dADP	188966	phosphoenolpyruvate	PEP	1005

Name	Abbreviation	PubChem_ID	Name	Abbreviation	PubChem_ID
2'-deoxyadenosine-5'-monophosphate	dAMP	12599	phenylalanine	PHE	6140
deoxyadenosine triphosphate	dATP	15993	phosphate	phosphate	1061
dehydroascorbic acid	DeHyASC	440667	phosphoribosylamine	PRA	150
2'-deoxyguanosine 5'-monophosphate	dGMP	135398597	L-proline	PRO	145742
deoxyguanosine	dGUA	135398592	PRPP	PRPP	7339
dihydroxyacetone phosphate	DHAP	668	pyruvic acid	PYR	1060
dihydroorotic acid	DHO	648	quinic acid	QUIN	6508
digalacturonic acid	DIGALACTURON	439694	ribose-1-phosphate	R1P	439236
disccharides_sum	Disccharides_sum		ribose phosphate	R5P	77982
dimethylallyl diphosphate	DMA-PP	647	D-ribonic acid	RA	5460677
deoxyribose 5-phosphate	dR5P	45934311	D-ribulose 1,5-Bisphosphate	RIB1,5 BP	123658
thymidine-5'-diphosphate	dTDP	164628	ribulose 5-phosphate+xylolose 5-phosphate (sum)	RIBU5P+XU5P	439184

Name	Abbreviation	PubChem_ID	Name	Abbreviation	PubChem_ID
thymidine 5'-monophosphate	dTMP	9700	sedoheptulose 1,7-bisphosphate	S1,7BP	164735
thymidine-5'-triphosphate	dTTP	64968	sedoheptulose 7-phosphate	S7P	165007
1-deoxy-D-xylulose 5-phosphate	dXP	443201	S-adenosylhomocysteine	SAH	439155
erythrose 4-phosphate	E4P	122357	SAICAR	SAICAR	160666
fructose 6-phosphate	F6P	69507	L-serine	SER	5951
flavin adenine dinucleotide	FAD	643975	shikimic acid	SHIK	8742
FAICAR	FAICAR	166760	succinic acid	SUC	1110
fructose-1,6-diphosphate	FBP	10267	succinyl-coenzyme A	Succ-CoA	92133
5-phosphoribosyl-N-formylglycineamidine	FGAM	154	sulfate	sulfate	1117
5'-phosphoribosyl-N-formylglycineamide	FGAR	151	L-threonine	THR	6288
flavin mononucleotide	FMN	643976	trigalacturonic acid	TRIGALACTURON	4472903
farnesyl diphosphate	FPP	445713	L-tryptophan	TRP	6305
fumaric acid	FUM	444972	L-tyrosine	TYR	6057
glucose-1-phosphate	G1P	65533	uric acid	UA	1175

Name	Abbreviation	PubChem_ID	Name	Abbreviation	PubChem_ID
glucose-6-phosphate	G6P	439284	uridine 5'-diphosphate	UDP	6031
glucosamine 1-phosphate	GA1P	188960	UDP-glucose	UDP-glc	8629
glucosamine 6-phosphate	GA6P	439217	UDP-N-acetyl-alpha-D-glucosamine	UDP-NAcGA	445675
gamma-aminobutyric acid	GABA	119	uridine 5'-monophosphate	UMP	6030
D-galactonic acid	GALACTON	128869	uridine	Uridine	6029
D-galacturonic Acid	GALACTURON	439215	uridine 5'-triphosphate	UTP	6133
glycerinaldehye 3-phosphate	GAP	729	L-valine	VAL	6287
5'-phosphoribosylglycineamide	GAR	90657462	L-valine	VAL	6287
guanosine diphosphate	GDP	135398619	xanthosine monophosphate	XAN5P	73323
geranylgeranyl pyrophosphate	GGPP	447277	tetrahydrofolic acid	THF	135444742
L-glutamic acid	GLA	33032	folic acid	Folate	135398658

Name	Abbreviation	PubChem_ ID	Name	Abbreviation	PubChem_ ID
glycerophosphoric acid	GLC3P	754	5-formyltetrahydrofolic acid	5-Formyl-THF	135403648
L-glutamine	GLN	5961			
gluconic acid	GLUCON	10690			

Curriculum Vitae (CV):

EXPERIENCE

Metabolomics Scientist 03/2025 - present

Metabolomics Core Technology Platform Heidelberg, Germany

- Project coordination & scientific collaboration (internal & external), LC-MS/GC-MS method development and optimization, instrument operation, maintenance & troubleshooting, data acquisition, processing & interpretation, quality control and assurance, documentation, SOP creation & implementation, Scientific reporting & publication writing, Training of users & knowledge transfer

Scientist at Anticancer Agents and Targeting Research Group 11/2024 - 03/2025

Leibniz Institute of Plant Biochemistry Halle, Germany

- Investigation, Analysis, Methodologys Establishment, Development & optimization, Visualization, Validation, Manuscript Writing, Supervision, Training, SOP implentation & creation, laboratory inventory

Associate Scientist at Anticancer Agents and Targeting Research Group 07/2018 - 10/2024

Leibniz Institute of Plant Biochemistry Halle, Germany

- Investigation, Analysis, Methodology, Visualization, Validation, Manuscript Writing, Supervision, Training.

Pharmacist - Community Pharmacist 08/2012 - 03/2015

Al-shefaa pharmacy Hama- Syria

- Pharmacist at Al-shefaa pharmacy
- Dispense prescription medications to patients
- Offer expertise in the safe use of medicines

EDUCATION & TRAINING

PhD study in the field of cancer cells metabolomics 07/2018 - 06-2025

Leibniz Institute of Plant Biochemistry, Martin-Luther University Halle-Wittenberg Halle, Germany

- PhD Project" Metabolomics in Drug Discovery: Cancer Cells Metabotyping to Predict the Mode of Action of Anticacner Agents"
- Research projects in the field of testing anticancer drugs from Natural or synthetic origin.

DAAD Researchers Exchange Programm 10/2019 - 11/2019

Institute of Biological Research Belgrade, Serbia

- "Siniša Stanković"- University of Belgrade. Belgrade, Serbia

Drug Development Course 10/2018

European Society for Animal Cell Technology Llafranc, Costa Brava, Spain

Master's degree course of Pharmaceutical Biotechnology 10/2015 - 06/2018

Leibniz Institute of Plant Biochemistry, Martin-Luther University Halle-Wittenberg Halle, Germany

- Master thesis "*In vitro* anticancer screening of sterol derivatives" IPB_ Anticancer Targeting&Screening, Halle (Saale)
- Research project on *In vitro* cross-linking of human recombinant tropoelastin_Faculty of Pharmacy, MLU, Halle (Saale)

Leadership of Syria course 11/2016 - 08/2017

University of Konstanz Konstanz, Germany

Bachelor's degree in Pharmacology and Pharmaceutical Chemistry 08/2009 - 08/2014

Al-Baath University Homs, Syria

- License in Pharmacology and Pharmaceutical Chemistry
- Major: Pharmacology
- Minor: Pharmaceutical chemistry

Declaration:

“I, Mohamad Saoud, hereby declare that I have completed this dissertation without unauthorized help of a second party and only with the assistance acknowledged herein. I have appropriately acknowledged and referenced all text passages that are derived from or based on the content of published or unpublished work of other authors”

I also declare that I have not made any previous unsuccessful attempts to attain a doctoral degree. Furthermore, I confirm that this current version of my thesis, or any other version, has not been previously submitted to any other faculty for evaluation.

I also declare that I have no previous convictions, and no preliminary investigations or legal proceedings have been initiated against me.

Mohamad Saoud Halle (Saale), 21.10.2024