May 2018

Off-Target Based Drug Repurposing Using Systems Pharmacology

Brent M. Kuenzi
University of South Florida, bkuenzi@mail.usf.edu

Follow this and additional works at: https://scholarcommons.usf.edu/etd
Part of the Bioinformatics Commons, Biology Commons, and the Pharmacology Commons

Scholar Commons Citation

This Dissertation is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.
Off-Target Based Drug Repurposing Using Systems Pharmacology

by

Brent M. Kuenzi

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy
Department of Cell Biology, Microbiology and Molecular Biology
College of Arts and Sciences
University of South Florida

Major professor: Uwe Rix, Ph.D.
    Eric B. Haura, M.D.
    John M. Koomen, Ph.D.
    Teresita Munoz-Antonia, Ph.D.
    Juan Del Valle, Ph.D.

Date of Approval:
    May 29, 2018

Keywords: tivantinib, ceritinib, AML, lung cancer, FAK1, RSK1/2, IGF1R, GSK3

Copyright © 2018, Brent M. Kuenzi
Dedication

I dedicate this work,

To everyone in my huge, crazy family especially my parents, Julie Newbury and Daniel Kuenzi, and my grandparents, Barbara and Norbert Kuenzi. A special dedication to my grandmother Rita Gurrlt whom recently passed away after a year long battle against Myelodysplastic Syndrome as well as my oldest sister, Tara Gilboy, who is living with a benign meningioma. Your strength and poise in the face of adversity makes me proud and your life is an inspiration to all of us in the family.

To my beloved fiancé, Melissa Hoffman. You are my best friend, most trusted colleague, role model, confidante and my biggest source of motivation and happiness for the past three years. From our first awkward interactions in the proteomics core to planning our future together, I have always looked up to you. Thank you for believing in me every step of the way. I could not have done this without you.

To my two sons, Kameryn and Logan, you have completely changed my life and I love you for it. You have made me stronger, better and more fulfilled than I could have ever imagined. I love you to the moon and back.
Acknowledgments

First, I owe an immense amount of gratitude to my mentor, Dr. Uwe Rix, for his incredible guidance, advice, and training. His drive, understanding, and respect for others creates an environment that I looked forward to being a part of every day. His patience and openness to new approaches and technologies have allowed me to grow and advance as a scientist and a person. I owe much of my success to you. I am also grateful to my committee members: Dr. Eric Haura, Dr. John Koomen, Dr. Juan Del Valle, and Dr. Teresita Antonio-Munoz. Your suggestions and assistance have helped advance not only my project, but also my education and career. I would also like to thank Dr. Christine Lovly for generously offering her time and reworking her schedule to serve as the external committee chair. Next, I would like to acknowledge each member of my lab family, both past and present: Lily Rix, Vinayak Palve, Silvia Kovalenko, Ou Deng, Claire Knezevic, Gabriela Wright, Natalia Sumi, and Anna Bryant. I’ve enjoyed not just working with you but also the friendships that we have built over the years. I would like to thank all the members of the weekly pY meeting group. This meeting has helped me progress in my research immensely and I am forever grateful for all the advice I have gotten from all of you over the years. I would especially like to thank Dr. John Koomen for taking the time to teach me about proteomics and to give me career advice. I wouldn’t be where I am today without your guidance. I would also like to especially thank Dr. Paul Stewart for all the advice in bioinformatics, programming, statistics, and vegetarian food. You’ve been pivotal in bolstering my success and I am very grateful. I would like to also thank all the members of the Haura lab both past and present: Matt Smith, Anu Majumder, Shikha Mahajan, Martial Bouchueng-Djidjou, Guolin Zhang, Jae Young Kim, Jianong Li, Adam Borne, Paul Stewart, and Eric Haura. You’ve all been amazing colleagues and friends. I’m
thankful for all the coffees, beers, chats, and advice over the years. I absolutely must thank Theresa Alvarez, Evelyn Rivera-Miranda, Emily Barsch, and Laurie Burns for helping schedule meetings, finding lost orders and helping to keep our lab afloat. I would also like to thank Theresa Alvarez for assisting with all my grant submissions. I would like thank my fellow Cancer Biology students and the Cancer Biology Student Organization for your support over the years. I want to especially thank Dr. Ankita Jhuraney. You’ve been an amazing friend and helped to build my scientific confidence and drive to push on.

The Moffitt Core Facilities were essential to this work and I owe them a lot of credit for the success of this work. I would especially like to acknowledge Bin Fang and Victoria Izumi of the Proteomics Core for routinely going above and beyond to teach me about all things proteomics. I would also like to thank Dr. Harshani Lawrence and Yunting Luo for their guidance and help with chemical synthesis. Finally, I am forever grateful to the Cancer Biology PhD Program and the opportunities it has provided me. Thank you to our program director, Dr. Ken Wright, who has been a fantastic advisor and Cathy Gaffney who was surely one the hardest working people that I have met in keeping all the graduate students alive. Thank you also to Janet Opel and Kimberly Hefner for your amazing work planning my defense. Without all of your assistance, this day would not have been possible. Thank You!
# Table of Contents

List of tables .................................................................................................................... iv  
List of figures ..................................................................................................................... v  
Abstract .............................................................................................................................. viii  

Chapter one: Background .............................................................................................. 1  
  Non-small cell lung cancer (NSCLC) ........................................................................ 1  
  Subtypes and current treatments .............................................................................. 1  
  Acute Myeloid Leukemia (AML) ............................................................................ 3  
  Subtypes and current treatments .............................................................................. 3  

Chapter two: Emerging strategies in cancer treatments .............................................. 6  
  Network Medicine .................................................................................................... 6  
  Polypharmacology .................................................................................................. 10  
  Drug repurposing .................................................................................................... 15  
  Integrated Functional Proteomics ....................................................................... 16  

Chapter three: Automated Processing of SAINT Templated Layouts ....................... 20  
  Underlying technologies ....................................................................................... 20  
  Affinity proteomics ............................................................................................... 20  
  Galaxy .................................................................................................................... 21  
  Implementation ...................................................................................................... 22  
  Discussion ............................................................................................................... 28  

Chapter four: Materials and methods ........................................................................ 30  
  Cell lines and transfections .................................................................................. 30  
  Cell culture ............................................................................................................ 30  
  Compounds ........................................................................................................... 31  
  RNA interference ................................................................................................. 31  
  Site-directed mutagenesis .................................................................................... 32  
  Chemistry.............................................................................................................. 32  
    c-(-)-tivantinib synthesis .................................................................................... 32  
    c-ceritinib synthesis ........................................................................................... 34  
  Antibodies and western blotting ......................................................................... 35  
  Cell viability and growth assays ........................................................................... 36  
    Cell viability ....................................................................................................... 36  
    Colony formation ............................................................................................... 37  
    Clonogenic assay .............................................................................................. 37  
  Immunoprecipitation assays ............................................................................... 37  
  Proteomics ............................................................................................................ 38  
    Drug affinity chromatography ......................................................................... 38  
    Phosphoproteomics ......................................................................................... 38  

i
List of tables

Table 1. Clinical characteristics of AML patients at time of diagnosis........................................78
List of Figures

Figure 1. FDA approved drugs for lung cancer treatment by year .............................................1
Figure 2. Pathological and molecular subtypes of lung cancer and currently available therapies for NSCLC ..................................................................................................................3
Figure 3. Genetic and treatment landscape of acute myeloid leukemia .............................................4
Figure 4. Disease-network modules ..................................................................................................8
Figure 5. Utilizing artificially intelligent network modules to predict disease phenotypes ..........10
Figure 6. Targeting complex networks with polypharmacology .........................................................11
Figure 7. Current Approaches to polypharmacology research ..............................................................13
Figure 8. APOSTL tools and features ...............................................................................................23
Figure 9. Example APOSTL quality control analysis .......................................................................26
Figure 10. Example APOSTL analysis ...............................................................................................27
Figure 11. Compound Synthesis and cellular evaluation of c-(-)-tivantinib .......................................33
Figure 12. c-ceritinib synthesis ........................................................................................................35
Figure 13. Network community optimization ....................................................................................48
Figure 14. Biological functions of GSK3 ..........................................................................................53
Figure 15. Cellular activity of tivantinib and various c-MET inhibitors in lung cancer cell lines .................................................................................................................................56
Figure 16. Determination of tivantinib’s kinase target interaction profile in A549 cells by chemical proteomics ....................................................................................................................59
Figure 17. Characterization of tivantinib’s kinase target interaction profile in H1648 cells by chemical proteomics ....................................................................................................................60
Figure 18. Functional analysis of GSK3α and GSK3β inhibition .........................................................61
Figure 19. Characterization of GSK3α and GSK3β inhibition .............................................................62
Figure 20. GSK3α is an actionable target in acute myeloid leukemia........................................................................66
Figure 21. Tivantinib has activity in acute myeloid leukemia cell lines.................................................................68
Figure 22. Tivantinib target profiling in acute myeloid leukemia cell lines .........................................................70
Figure 23. Analysis of cellular response following tivantinib treatment ..........................................................73
Figure 24. Identification of tivantinib and ABT-199 as a synergistic drug combination in AML cells ..........................................................75
Figure 25. Response of patient derived AML cells to tivantinib and ABT-199 ...............................................77
Figure 26. RSK isoform sequence conservation and mechanism of activation ..................................................86
Figure 27. Biological functions of RSK...........................................................................................................87
Figure 28. Structure and function of FAK1 .......................................................................................................91
Figure 29. Structure and function of IGF1R ....................................................................................................94
Figure 30. Ceritinib has beneficial off-target activity in ALK-negative NSCLC cells ........................................99
Figure 31. Cellular effects of ceritinib in NSCLC cell lines ........................................................................101
Figure 32. Integrated functional proteomics ................................................................................................102
Figure 33. Validation of c-ceritinib probe ........................................................................................................103
Figure 34. Chemical proteomic characterization of ceritinib ........................................................................104
Figure 35. Ceritinib target validation ............................................................................................................106
Figure 36. Phosphoproteomic characterization of H650 cells following ceritinib treatment ..................107
Figure 37. Ceritinib effector network ............................................................................................................107
Figure 38. Subnetwork analysis of ceritinib mechanism of action .............................................................108
Figure 39. Ceritinib inhibits cell viability through inhibition of IGF1R, FAK1, RSK1 and RSK2 ..................110
Figure 40. YB1 is a critical downstream node ...............................................................................................112
Figure 41. Ceritinib strongly synergizes with paclitaxel ..............................................................................113
Figure 42. Ceritinib elicits synergy through inhibition of RSK1/2, FAK1 and IGF1R .................................115
Figure 43. FAK1 autophosphorylation may be predictive of synergistic response to ceritinib and paclitaxel ..........................................................................................................................117
Figure 44. FAK1 autophosphorylation characterization across cell lines, patient tumors, and PDX tumors .................................................................118

Figure 45. FAK1 autophosphorylation biomarker validation .................................................................................................119
Abstract

The goal of this study was to identify novel drug repurposing opportunities in cancer by utilizing the off-target profiles of clinically relevant kinase inhibitors. This was based on the observation that the global target profiles of compounds are largely ignored and that many compounds have activity that cannot be explained by their cognate target alone. Additionally, by utilizing clinically relevant compounds, any results would hold a high potential for eventual clinical implementation.

We utilized a systems pharmacology approach utilizing cell viability-based drug screening to identify compounds with beneficial off-target activity and then using chemical and phosphoproteomics in order to elucidate the mechanisms of action of these compounds. We found that tivantinib has off-target activity in NSCLC cells through inhibition of GSK3. Based on tivantinib’s ability to inhibit GSK3α, we hypothesized that tivantinib would therefore have activity in acute myeloid leukemia (AML). We found that tivantinib had potent activity in AML through inhibition of GSK3. We also identified a highly synergistic combination with ABT-199 by drug synergy screening which was effective in HL60 cells and patient derived AML cells. We also found that the anaplastic lymphoma kinase (ALK) inhibitor, ceritinib, had activity across several ALK-negative lung cancer cell lines. We utilized integrated functional proteomics to identify the new targets and network-wide signaling effects. Combining pharmacological inhibitors and RNA interference revealed a polypharmacology mechanism involving the noncanonical targets IGF1R, FAK1, RSK1 and RSK2. Mutating the downstream signaling hub YB1 protected cells from ceritinib. Consistent with YB1 signaling being known to cause taxol resistance, combination of ceritinib with paclitaxel displayed strong synergy, particularly in cells expressing high FAK autophosphorylation, which we show to be prevalent in lung cancer. Together, we present a
systems chemical biology platform for elucidating multikinase inhibitor mechanisms, synergistic drug combinations, mechanistic biomarker candidates and identifying novel drug repurposing opportunities.
Chapter one: Background

Non-small Cell Lung Cancer (NSCLC)

Subtypes and current treatments

With over 1.8 million newly diagnosed cases and ~1.6 million deaths each year, lung cancer remains the most common cause of cancer-related mortality worldwide (1). Cigarette smoking has been shown to be the leading cause of lung cancer however there are additional carcinogens significantly associated with lung cancer development including asbestos and air pollution (2–4).

In broad terms, lung cancer is subdivided in two major subtypes, small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). SCLC is a highly aggressive disease and with a 5-year survival rate of 10%, it is associated with very poor prognosis. The poor outlook of a SCLC diagnosis is exemplified by the lack of novel therapies for the treatment of SCLC. There have not been significant advances in SCLC treatment since the introduction of the first platinum-based chemotherapies (5,6). Conversely, there have been major advances in the molecular pathology classification, diagnosis and treatment of NSCLC over the last two decades with a deep

Figure 1. FDA approved drugs for lung cancer treatment by year. Displayed are therapies approved for lung cancer treatment from January 1978 – September 2017. Shading intensity scales with the number of approved therapies.
characterization of the underlying genetics and a number of genomics-based treatment strategies being approved (Figure 1). As a result, it is now accepted that NSCLC is a collection of different pathologies all with their own molecular signatures (Figure 2a) (7–11). NSCLC in general can be subdivided into squamous cell carcinoma, adenocarcinoma, and large cell carcinoma. These histological subdivisions alone allow for better selection of treatments for NSCLC lacking driver mutations, such as to avoid treatment associated bleeding from bevacizumab in squamous cell carcinoma or the observed sensitivity of adenocarcinoma to pemetrexed (12–14).

Of the histological subtypes of NSCLC, adenocarcinoma is the most widely characterized with a number of additional genetic subtypes being identified. As a result, there have been major advances in genome-based treatment of NSCLC with multiple FDA-approved therapies for the major, targetable genomic alterations such as the EGFR inhibitors gefitinib and erlotinib for EGFR-driven NSCLC and the ALK inhibitors crizotinib and ceritinib for patients harboring ALK-rearrangements (Figure 2b) (15,16). Despite these advances, the majority of lung cancer patients don’t have actionable driver mutations. Therefore chemotherapeutics such as cisplatin, pemetrexed and docetaxel are some of the most common treatments for lung cancer (15). In addition to small molecule inhibitors, novel immunotherapeutic strategies have been under intense investigation and have led to a number of therapies being approved for NSCLC treatment (Figure 2c) (17). While these therapies are promising, only a subset of patients respond to these therapies. Therefore, there remains a great unmet need for additional therapeutic options for NSCLC treatment. In this dissertation we aimed to identify novel drug treatment strategies for NSCLC.
Acute Myeloid Leukemia (AML)

Subtypes and current treatments

Acute myeloid leukemia (AML) is a highly aggressive disease with 5-year survival rates ranging from 5-15% in older patients to 30% in young adults (18). Patient outcomes for AML have changed very little in the last 10 years with only a small percentage of patients exhibiting durable long-term responses. Current therapies consist of high dose chemotherapy and stem cell transplantation.
However, these therapies are still associated with significant amounts of morbidity (19,20). There have been recent advances in the genetics of AML which have lead to extensive genetic subtyping of patients (Figure 3a) (21). Patients can be stratified by both the cytogenetic risk category as well as by putative driver mutations such as NRAS, FLT3-ITD, NPM1, and TP53. Genetic alterations are very common in AML and it has been reported that in a group of patients at least one genetic mutation was found in over 97% of these patients (22). These mutations can be associated with specific chromosomal rearrangements such as t(8;21) and inv(16) being associated with KIT mutations (22). For this reason, low, intermediate and high risk cytogenetic subtypes of AML have overlapping but not identical genetic profiles. However, many patients have genetic mutations without large chromosomal abnormalities lending to further genetic complexity.

Figure 3. Genetic and treatment landscape of acute myeloid leukemia. (a) Pie charts of the prevalence of the most common genetic subtypes of AML stratified by patient cytogenetic risk group. Mutational data obtained from Ohgami RS., et al. (2015) Mod Pathol. (b) Key developments in AML treatment from January 1970 – January 2018. (c) Targeted therapy matrix of commonly altered genes in AML and FDA approved therapies (AML) with the ability to inhibit these gene products.
Despite these advances in characterizing and subtyping AML, the treatment landscape for AML has changed very little in almost 50 years (Figure 3b). The majority of patients still receive induction therapy (7+3 cytarabine + daunorubicin) or a bone marrow transplant (BMT). With the exception of gemtuzumab-ozogamicin (which was withdrawn from the market from 2010-2017) there were no new therapies approved for AML until 2017. There were however major recent advances in targeted therapy for AML with the FDA approval of 3 therapies in 2017 for patients with specific mutations (e.g. midostaurin or crenolanib for FLT3-ITD, enasidenib for IDH2) (Figure 3c). While these therapies are promising and many patients are benefiting, the vast majority of genetic alterations in AML do not have therapeutic options available and therefore these patients still receive induction therapy or BMT. Therefore, there remains a major unmet need for the identification of new therapeutic options for AML. In this dissertation we will be characterizing a novel drug repurposing opportunity for the treatment of AML.
Network medicine

Since the onset of the post-genomics era, it has become increasingly clear that cancer is often not a consequence of an alteration in a single gene, but rather is a complex interplay of perturbations in intra- and intercellular networks. Most cellular functions are exerted through intersections with other cellular components both in the cell as well as across cells, resulting in a highly complex network of hundreds of thousands of edges (or more) connecting proteins, RNA, metabolites, etc. The implication of this complexity is that the impact of a particular perturbation is not restricted to that gene/protein but rather travels along the edges of the network altering the activity of genes/proteins that are not otherwise altered. As a result, the genes/proteins responsible for driving a cancer are not always apparent and therefore understanding the network context of a particular cancer is of paramount importance (23,24).

Consequently, a number of approaches to systematically interrogate both the molecular complexity (nodes) of the disease as well as the inter-molecular relationships (edges) that lead to a particular phenotype have been developed. In particular, advances in network theory (24–30) have indicated that biological networks are not random but rather are governed by organizing principles that allow researchers to address and predict certain properties of genes/proteins within the network. For instance, most networks display a certain degree of clustering of local regions of highly interlinked nodes in the network. These topological modules help to reveal the global structure of the network. Additionally, given the ‘local hypothesis’ of biological networks, proteins involved in similar cellular functions or biochemical processes tend to interact with one another (31,32). This aggregation of functionally similar nodes is known as a functional module. Similarly,
proteins/genes implicated in the phenotype of a particular disease will more often interact with each other, which is known as a disease module. Given a particular network, these module types are likely to display a large degree of overlap as they all represent locally dense regions within the larger network. Therefore, the key assumption to be made when interpreting biological networks is that topological modules represent clusters of genes/proteins with closely related functions and thus a topological module is also representative of a particular functional module. Therefore, a disease is the breakdown of the function or regulation of a functional module (Figure 4). As a result, key insights into the underlying molecular processes governing the disease can be interrogated by analyzing the topological and therefore functional structure of the network. Analysis and identification of these topological modules can be computationally challenging but there have been a number of methods developed in order to efficiently identify these modules (See chapter four for an example) (33–35).

In addition to modular clustering, a key characteristic of a biological network is the centrality or the influence of individual nodes within a network. Given a random network, the degree (# of edges) of nodes within the network are roughly equal; however, biological networks often have a degree distribution in which some nodes have a higher degree. These nodes are therefore more likely to be influential than other nodes with lower degree which tend to segregate to the periphery of the network. These ‘hubs’ often coordinate specific cellular processes and may serve as attractive drug targets (36). Alternatively, in directed networks (such as many gene/protein regulatory networks), bottlenecks tend to be essential genes/proteins. Bottlenecks are nodes with a relatively high number of shortest paths between nodes being funneled through them. They therefore disproportionately control information flow through the network and may represent effective drug targets. As a result, various computational approaches have been developed to identify these nodes such as eigenvector centrality for hubs (See chapter four for details) and betweenness centrality for bottlenecks (37,38).
The underlying concepts described above are highly adaptable and can be quite powerful in identifying key nodes and edges within a network. For instance, one can study individual cellular processes and key nodes regulating that process by analyzing modules individually for nodes with high eigenvector centrality or more specifically study the crosstalk between cellular processes/pathways by looking at the undirected betweenness across the larger network. Combining these two approaches may even allow for prediction of drug resistance mechanisms based on the hubs and paths identified through the combination of random walk closeness centrality (39) and heat diffusion centrality (40) using the initial drug targets as initiator nodes where hubs with the highest 'heat' and closeness centrality may be more likely to be associated with bypass signaling mechanisms. These underlying concepts even transcend biological networks and are used by some of today’s largest tech companies in network models to make

Figure 4. Disease-network modules. Topological modules represent a collection of locally dense communities of interactions and are a property of the network itself. Functional modules represent a statistically significant collection of genes / proteins involved in a particular cellular function based on node functional characteristics. Functional modularity relies on the hypothesis that nodes involved in similar cellular processes (highlighted in blue) tend to interact with each other more often. Disease modules represent the specific collection of genes / proteins involved in the disease phenotype (highlighted in red). Note that each module type shares overlap and therefore one may make the assumption that topological modules correspond in part to functional modules and a disease phenotype may be viewed as an aberration in or dysregulation of a particular functional module. Conceptually adapted from Barabási, Gulbahce and Loscalzo. Nature Reviews Genetics (2011).
predictions of websites you may be interested in such as Google’s in-house centrality algorithm, PageRank (41), or Facebook using network modularity to make predictions of which members in your extended social network are most similar in order to make friend suggestions (42). Therefore with careful selection of various centrality metrics combined with the knowledge gained by analyzing the modular structure of a network, one can ask any number of questions for further investigation.

In order to potentially apply network models in an unbiased fashion to predict characteristics of a patient tumor, such as cancer aggressiveness or response to a particular therapy, it is important to more accurately model the entire system using hierarchical network models. While the above described 1-dimensional or ‘flat’ network models are useful for identifying novel drug targets and gaining insights into the underlying signaling pathways, they do not fully capture the hierarchical structure of the cell or of an organism as a whole. Hierarchical network models, which account for this type of information, have been under development for years using both computational approaches to predict gene function as well as careful curation of the literature. The most widely known and used hierarchical model is the Gene Ontology project which aims to accurately map the hierarchical structure of the cell through annotation of the biological processes, molecular functions and cellular compartments of all genes in the genome (43,44). Recently, there have been major advances in the application of these hierarchical models to predict simple characteristics of model systems, such as the growth rate of yeast based on their genetic characteristics (45–49). This is done by mapping genes to these curated hierarchical models and collapsing the leaves and branches of the network into an “ontotype” which contains functional, localization and molecular information of the altered signaling networks. This is done across the entire network to create a simplified tree with hundreds of ontotype branches. The structure of this network-tree allows for easy assigning of ontotypes as features for a random forest model which retains the hierarchical structure of the network while allowing for artificially intelligent prediction of an organism’s phenotype (45). This work has paved the way for potential
future clinical application of network models using information such as the patients’ exome sequences to predict response to any number of cancer therapies (Figure 5). While application of this methodology is still in its infancy and there are a number of regulatory hurdles to overcome, it represents a very promising future clinical application of network models.

Figure 5. Utilizing artificially intelligent hierarchical network models to predict disease phenotypes. A potential future clinical application of network models is to utilize hierarchical network models to predict patient response. Briefly, a patient’s exome mutations would be mapped to a hierarchical network and leaves/branches of this network would be collapsed into ontotypes. These ontotypes are then used as features to a random forest model which retains the structure of the network while allowing for prediction of any number of patient phenotypes such as response to therapy.

Polypharmacology

The current paradigm of drug discovery, that is developing drugs with high potency and selectivity for a single target, while proving successful in many indications such as BCR-ABL driven CML and EML4-ALK rearranged NSCLC (50–54), has certain limitations. This paradigm is based on a
A direct cause-effect relationship between an oncogenic driver and a cellular phenotype. However, the majority of cancers don’t have strong oncogenic drivers that can be targeted pharmacologically (55–57). For example, around 75% of NSCLC tumors do not harbor actionable oncogenic drivers and thus a targeted therapy does not exist for these patients (58,59). Therefore, the “one disease, one target, one drug” concept may be too simplistic to effectively treat such a multifactorial disease (60). Alternative strategies have emerged to target these cancers, such as the recent developments in immunotherapy (17) and a larger push toward drug combination strategies. However, these therapies are proving to also be effective only in a subset of patients (61–65). Thus novel strategies need to be employed that are not only well-tolerated for patients but are also able to overcome the vast complexity of the disease.

Cancer is a highly complex disease that relies on a dynamic network of genetic, protein, and metabolic interactions to drive carcinogenesis. This is particularly true for many so called ‘oncogene-negative’ cancers, which often display multiple mutations and epigenetic aberrations leading to deregulation of signaling and gene regulatory networks. These cancers lacking strong
oncogenes may therefore be more effectively treated using a network-based approach, wherein drugs target multiple pathways and cellular processes (25,66,67). Although inhibiting a single target may affect several cellular processes, efficient shutdown of oncogenic signaling and regulatory networks is often only achieved by directly engaging multiple proteins (Figure 6). This can be achieved by combining two or more targeted agents, a concept actively pursued in many preclinical and clinical studies such as blocking compensatory feedback activation in melanoma using a combination of BRAF and MEK inhibitors (68). Clinical translation of drug combinations however is often limited by issues with drug formulation, drug-drug interactions, increased toxicity, regulatory mechanisms as well as strategic business management (67,69). Alternatively, network-wide signaling effects can be produced with single compounds that innately inhibit multiple disease-related targets, a phenomenon referred to as polypharmacology (69,69–73). Therefore, a multi-targeted drug with a broader activity profile (i.e. targeting >1 protein with relatively equal potency) may be more effective in shutting down these complex networks. These compounds could be even further enhanced through the intelligent design of a combination therapy targeting other critical nodes within the network or by preemptively inhibiting known bypass signals (while being cognizant of the known limitations of combination therapy discussed above). Multi-targeted drugs should however not be confused with promiscuous drugs which target a very large amount of proteins leading to sometimes unpredictable adverse reactions and are generally unsafe for use.
There are a number of challenges in the design of polypharmacology drugs which limit the number of drugs designed for this purpose as well as the number of identified polypharmacology drugs. For example, rational design of a multi-targeted drug for specific, unrelated proteins poses a significant challenge when considering the structure-activity relationships of numerous targets. Therefore, a number of strategies have been developed to overcome the hurdle of designing multi-targeted drugs and utilizing polypharmacology (Figure 7 A-D). Early approaches involved large scale fragment-based high throughput screening (HTS) followed by validation and structure activity relationship (SAR)-based optimization (74) or through structure-based design focused on virtual screening across multiple crystal structures (75,76). While it is possible to design multi-targeted drugs in this manner, it requires substantial time and resources to identify suitable hit scaffolds and is very challenging to sufficiently optimize a compound following identification of a hit or lead molecule. Recently, advances in machine learning have allowed for automated design and optimization of hit compounds using Bayesian network models to solve the vast combinatoric problem of multiple target optimization (77,78). These models allow for prioritization of various compound features such as activity/potency against selected targets and even certain ADME properties such as blood-brain barrier permeability (79,80). With ligand-target prediction

Figure 7. Current approaches to polypharmacology research. (A) Fragment-based high throughput screening. (B) In silico design using machine learning and deep neural networks (C) Target identification and discovery (D) Structure based design.
success rates of up to 75%, automated methods are certainly appealing and quite promising. However, these machine-learning based approaches to polypharmacology design are still in early stage development and have yet to demonstrate their utility in developing clinically relevant compounds.

An alternative approach to rationally designing a multi-targeted drug is to identify a compound that is serendipitously multi-targeted and repurpose this compound into a new indication in which it may be effective. The majority of kinase inhibitors are type I inhibitors, in that they target the ATP binding pocket containing a conserved Phe of the Asp-Phe-Gly (DFG) motif of the hydrophobic pocket in the kinase’s active conformation (‘DFG-in’) (81,82). Type II inhibitors on the other hand occupy both the ATP binding pocket as well as an allosteric pocket that is exposed when the DFG motif moves out of the hydrophobic pocket in the kinase’s catalytically inactive conformation (‘DFG-out’) (83). Since this ATP binding pocket is in general highly conserved across different kinases (84), this leads to an inherent proclivity toward multi-targetedness for type I kinase inhibitors whereas type II inhibitors are in general more specific (85–87). However, the multi-targeted nature of type I inhibitors is not a detriment as these additional targets may lead to unexpected anti-proliferative activity in cancers lacking actionable driver mutations that may be more reliant on complex signaling networks. As discussed above, these targets may allow for better coverage across the cancer signaling network and therefore a greater chance of efficiently inhibiting critical survival signals (Figure 6). While we currently have insufficient knowledge of the underlying signaling networks that drive these cancers lacking actionable drivers and insufficient knowledge of the broader target profiles of the current repertoire of kinase inhibitors to make predictions of which drugs will be effective, cell-based drug screening approaches are quite effective in identifying these cases of unexpected, anticancer activity (88–90). With detailed knowledge of the mechanisms of action of these drugs, we can also learn about the underlying signaling networks and their inherent vulnerabilities.
**Drug Repurposing**

Each year dozens of new therapies are entering clinical trials for the treatment of a wide range of malignancies but the rate of FDA approval for oncology drugs entering phase I trials is less than 7%. Drugs that progress to phase III trials still have only an approximately 50% chance of success because many trials are suspended before completion. Of these suspended trials, most (>50%) are stopped due to a lack of efficacy in the selected patient population as compared to safety/toxicity issues, which is the least likely cause of suspension (9%) (91). This poor success rate is preceded by a 10 to 15-year timeline from target discovery to FDA approval and a development cost of 1–2 billion dollars. The high time and cost of drug development ultimately translates to a patient’s cost of treatment of $60,000 - 100,000 a year (92–96). This poor success rate and the high cost of new molecule development provides a large barrier to drug development and prevents investigators from readily initiating clinical trials.

In order to decrease the time to drug approval, reduce the development costs associated with these molecules and to increase the available pool of approved therapies, there has recently been a greater focus on drug repurposing, that is repositioning clinically advanced or FDA approved drugs from one indication into another. There have been a number of examples of drugs being successfully repurposed based on both their cognate / intended targets as well as based on their ‘off-target’ profiles. Most notably, the immunomodulatory drug thalidomide was originally developed to prevent morning sickness in pregnant women. This however resulted in a high incidence of birth defects due to off-target effects of thalidomide (97,98). While the exact mechanisms are still under investigation, it is likely that the teratogenetic effects are a result of oxidative stress and anti-angiogenic action of thalidomide (99). It was later discovered that based on these effects, thalidomide had anticancer activity in myeloma which led to the eventual approval of thalidomide for multiple myeloma (100,101). Similarly, metformin was originally approved as first line medication for type II diabetes. It has since been used in the treatment of polycystic ovarian syndrome, cardiovascular disease, as well as numerous different cancer types.
(102–108). The mechanisms of metformin’s efficacy remain incompletely understood but many of the beneficial effects of metformin can be attributed to activation of AMPK which has multiple roles in insulin signaling and glucose metabolism (109,110).

In addition to compounds with somewhat elusive mechanisms like thalidomide or metformin, some entire compound classes lend themselves more toward repurposing based on the nature of their target class. Targeted drugs, and kinase inhibitors in particular, have been shown to serendipitously display widely varying target profiles beyond their intended or “cognate” targets (69,111–114). While these “off-” or “non-canonical” targets are often either unknown or disregarded, they confer an inherent potential for anticancer activity. Notably, phenotypic screening approaches have found some kinase-targeted drugs to show antitumor activity in various subsets of cancer, which is unrelated to inhibition of their cognate targets (88–90). This is enabled by the conserved nature of the ATP binding pocket across different kinases (84), which in general leads to a more promiscuous target profile of kinase inhibitors. For example, SRC family member kinases are highly conserved and as a result the SRC inhibitor dasatinib potently targets the entire SRC family of kinases (115). This phenomenon in which a compound targets other proteins in addition to its intended target is not limited to kinase inhibitors as other compound classes also harbor multiple potent targets due to conservation of active site residues. For example, several other protein families and compounds designed to inhibit them have inherent multi-targetedness such as poly-ADP ribose polymerase (PARP) family members / PARP inhibitors (116,117) and histone deacetylase family members (HDAC) / HDAC inhibitors (118,119).

Integrated Functional Proteomics

Identifying the target(s) and mechanisms responsible for beneficial off-target activity is not always straightforward but is important when investigating a potential repurposing opportunity. The relationship between the observed activity and the responsible target(s) can sometimes be
inferred from analyzing multiple compounds of the same target class using clustering or machine learning and dimensionality reduction methods (90,120,121). These methods are incredibly powerful at identifying complex relationships in high dimensional datasets that aren’t otherwise apparent. While interesting for hypothesis generation, this method of target inference is limited to with which compound / classes a drug is co-clustering as well as our available knowledge of drug-target profiles and therefore should be critically evaluated and validated downstream of this analysis. Since the global target profiles of most drugs remain uninvestigated, the actual target(s) responsible for any observed activity is often incompletely understood. Therefore the underlying mechanism of action (MoA) is not always apparent, but in many cases is involving one or more non-canonical targets.

A number of different strategies have been developed to identify these targets including drug affinity chromatography, activity based-protein profiling, cellular thermal shift assays (CETSA) and Kinobeads (113,122,123). Each of these technologies share some similarities and have their own advantages and disadvantages, making the choice of which technique to use an important consideration when planning a study to investigate a drug target profile. Kinobeads for instance feature broadly selective kinase inhibitors covalently linked to sepharose beads allowing for affinity enrichment of much of the kinome from cells or tissue lysates. A drug target profile is determined by examining which kinases are competed away from the affinity matrix by LC-MS/MS. This has the advantage of scalability allowing for large scale analyses but the technology also relies heavily on competition away from the affinity matrix which is influenced by a number of factors that can result in false positive and negatives. CETSA, on the other hand, investigates the thermal stabilization of proteins upon binding to the drug of interest. The assay involves treating cells with a compound of interest, heating to denature proteins, and measuring delayed protein degradation as a result of ligand binding allowing for an estimation of binding free energies (123–125). This has the advantage of increased throughput allowing for large scale studies but is highly dependent on a protein stability and linear denaturation kinetics which many proteins and
kinases lack. Drug affinity chromatography is a precursor to the kinobeads technology and utilizes a chemically modified compound of interest covalently linked to sepharose beads. The proteins / kinases enriched are then compared directly to negative controls (compounds of different target class or competition) in order to determine the target profile of that compound. This has the advantage of enriching for targets in their biological context, many times enriching entire complexes, and doesn’t rely solely on competition-based target identification. Drug affinity chromatography does however require direct chemical modification of the compound of interest, which in some cases can be difficult, and could potentially alter the target profile of the compound (122). These concerns, while justified, are often not an issue but still need to be carefully considered when performing a drug affinity chromatography experiment.

In addition to global drug target identification, novel bioinformatic methods and tools need to be developed to facilitate the identification of these targets from large proteome-wide AP-MS datasets (see chapter three for an example). For example, the normalized spectral abundance factor (NSAF) was developed in order to analyze quantitative proteomics datasets by taking into account the protein size (amino acid length) when comparing the spectral counts of biological samples (126). This approach has been extensively used for affinity proteomics experiments as it is able to reduce the false discovery rate of interacting proteins by penalizing large proteins, which tend to have higher spectral counts. More recently, a semi-supervised Bayesian model, Significance Analysis of Interactomes (SAINT), was developed to identify true protein interactions with a ‘bait’ of interest in an affinity proteomics experiment. SAINT (and its optimized algorithm SAINTexpress) calculate the posterior probability of a true interaction for each protein identified by proteomics by comparing the probability distributions of true and false interactions based on the experimental and control inputs (127,128). This allows for the identification of high confidence interacting proteins and when used with drug affinity chromatography, the identification of drug targets. Finally, once protein interactions are identified it is important to further score these protein ‘hits’ based on common contaminants in affinity proteomics. To do this, the Contaminant
Repository for Affinity Proteomics (CRAPome) was developed which is a database containing the number of affinity proteomics experiments each protein in the proteome was identified in (129). This is a powerful resource in which to filter ‘hit’ proteins that may have been enriched not as a function of the experimental conditions but potentially as a nonspecific interaction with bead resins or other technical artifacts common to affinity proteomics. By combining these approaches and tools, one can very confidently identify true interacting partners with a drug of interest for subsequent validation.

While drug binding protein identification is critical to understanding a drug’s MoA, additional studies are almost always necessary in order to identify the biologically relevant targets of these compounds. Therefore, it is critical to interrogate the underlying signaling mechanisms responsible for the anticancer activity. Cellular signaling occurs through a number of diverse molecules and pathways but one of the most important signaling interactions is through phosphorylation. Kinase – substrate relationships are ubiquitous in biology and are involved in nearly every biochemical process in the cell, making them very important in diseases such as cancer (130). As a result, protein phosphorylation has become a very active area of research. One powerful method used to measure the global phosphorylation profile of cells is through the enrichment of phosphorylations using an antibody (phospho-tyrosine residues) or a positively charged metal ion resin to capture all negatively charged phosphorylated serine, threonine, and tyrosine residues followed by analysis of phosphorylated peptides by mass spectrometry (phosphoproteomics) (131). By analyzing the specific phosphorylation sites being altered by a drug treatment, one can gain critical insights into the pathways being altered that are driving the observed phenotype in these cells. In this dissertation, we will be applying the concepts of network medicine and polypharmacology in order to investigate off-target based drug repurposing opportunities using integrated functional proteomics.
Chapter three: Automated Processing of SAINT Templated Layouts

(Note to reader: Parts of this section have been published previously in Kuenzi et al. Journal of Proteome Research doi: 10.1021/acs.jproteome.6b00660 (132) and are being reproduced with permission from American Chemical Society copyright 2016 (Appendix B).

Underlying technologies

Affinity proteomics

Biological processes in the cell rely on a complex interplay of signaling complexes and molecular interactions such as posttranslational modifications, conformational control and protein-protein binding (133), one of the most common and important signaling mechanisms. Every type of signaling event relies on some form of protein-protein interaction in which a linear peptide motif containing a core determinant and surrounding residues on a substrate protein (or complex) allows for binding recognition by a second protein (134). These interactions range in strength and residence time (from seconds to days) and underlie all major cellular processes. Therefore, it is of critical importance to understand the full repertoire of binding interactions (interactome) of disease-associated proteins to identify biomarkers and design effective therapies.

The most widely used tool to study protein-protein interactions and other binding interactions such as between proteins and drugs / small molecules is affinity purification – mass spectrometry (AP-MS), which includes various techniques including immunoprecipitation, tandem affinity purification (TAP), and drug affinity chromatography. These techniques employ a common approach in which a “bait” (e.g. protein or drug of interest) is immobilized to a matrix and subsequently used to bind and pull down “prey” proteins (135–137). These interacting proteins
are then eluted and analyzed by mass spectrometry. AP-MS is being increasingly used for both small and large-scale analysis of interaction networks however computational tools for the analysis of AP-MS experiments have only begun to catch up to the advancements in experimental approaches (127).

With increased throughput of baits needed to fully map signaling pathways and the ever increasing quantity of data being generated by highly sensitive mass spectrometers, the analysis of quantitative AP-MS data is becoming progressively more computationally demanding. Several computational tools have been developed to overcome the computational challenges of AP-MS data (126–128,135,138). These tools employ various approaches in order to assign probabilities to bait-prey interactions using spectral counts or intensity distributions of negative control experiments as a comparator. The most widely used tools used for this purpose are the Significance Analysis of INTeractomes (SAINT) and the more recent, optimized version, SAINTexpress (127,128). While these algorithms can provide highly accurate predictions of true bait-prey interactions, they remain largely inaccessible to bench scientists as they are operated through the command line and require a large amount of data reformatting which can only be conceivably performed programmatically. Accordingly, many scientists are limited in the analyses they are able to perform to what can be easily achieved in Microsoft Excel. Therefore, approaches and tools need to be developed not just to statistically analyze AP-MS experiments, but to also make these analyses accessible to the bench scientists that are performing these experiments.

**Galaxy**

A number of bioinformatic workflow platforms have been developed to overcome the growing need to make bioinformatic tools accessible for proteomics and genomics research such as Taverna (139), Kepler (140), and Galaxy (141–144). All of these platforms have a similar design philosophy in which a user has access to a graphical user interface (GUI) to string together bioinformatic tools into workflows allowing for easy use and integration of multiple bioinformatic
tools. These tools are often developed in a scripting language such as perl, python or bash and implemented into the workflow GUI by a bioinformatician. These workflow management systems come with several advantages in addition to their ease of use in that they can be integrated into high performance computing (HPC) clusters, large-scale cloud storage systems, and they offer enhanced reproducibility of research through easy sharing of data and analysis workflows (145).

One of the most recent and widely used workflow management systems is the Galaxy platform. Galaxy was originally developed for the analysis of next-generation sequencing (NGS) data but has now expanded into all facets of data-intensive biology including proteomics (132,146,147), flow cytometry (148), epigenetics (149), and metabolomics (150,151). Coupled with simple tools for basic table manipulations, numerous analysis histories, well established tools such as BLAST (152), and common visualizations such as boxplots and heatmaps, Galaxy enables a user to perform nearly all biological analyses in their platform. Galaxy allows for simple implementation and integration of tools written in a wide variety of programming languages by requiring only the script itself and an XML file that describes usage of the script from the command line. This allows for rapid development and implementation of tools which can then be shared in the Galaxy Toolshed. It is this ease of use and flexibility that makes Galaxy an ideal solution for developing bioinformatic pipelines. Lastly, once a user has configured a Galaxy instance with the necessary tools, he/she can easily share a pre-configured Galaxy instance using services such as Amazon Web Services (AWS) or Docker (153,154).

Implementation

To help bridge the gap between bioinformatics and bench scientists we developed Automated Processing of SAINT Templated Layouts (APOSTL) (http://apostl.moffitt.org/), an interactive Galaxy pipeline for reproducible analysis of AP-MS data. APOSTL provides an intuitive user interface to identify novel interactions, interpret data and visualize information. APOSTL is the first tool available for affinity proteomics data analysis in Galaxy and has been developed in
collaboration with other proteomics bioinformaticians for extensive integration with other platforms such as the widely used Galaxy-P platform (https://usegalaxyp.org/), a proteomics specific Galaxy distribution (155). APOSTL provides a workflow to pre-process raw data formats into inter, prey and bait files for input into SAINTexpress (128), analyze interactions by the CRAPome (129), calculate NSAF values (126) for each protein, and integrate / analyze the results (Figure 8).

Figure 8. APOSTL tools and features

A major challenge in developing an AP-MS analysis platform lies with the sheer heterogeneity in proteomic workflows across institutions. Depending on the institute different quantification methods will be used, such as spectral counting or MS$^1$ intensity. This data can then be summarized in numerous platforms, all with their own data structures and formats, such as Scaffold (http://www.proteomesoftware.com/products/scaffold/) or PeptideShaker (147) for spectral counting data or MaxQuant (156–159) for MS$^1$-based quantification. Thus there constituted a need for a tool that supports the analysis of both spectral counting and MS$^1$-based
quantification that can handle disparate data types. Therefore a pre-processing platform was
developed that can accept a “Samples Report” from the free Scaffold Viewer, an “Experiment
Report” from the Scaffold Export tool in Galaxy, “Protein Report” files exported from SearchGUI /
PeptideShaker (146,147), and the open source data standard mzIdentML files (160) generated
directly from various search engines for spectral counting data. These data formats are then
processed into files specifying all possible interactions (inter), describing each prey protein (prey)
and experimental design (bait) files for input into SAINTexpress. For MS\textsuperscript{1} based quantification, a
“peptides.txt” file from MaxQuant is accepted as input. Since MS\textsuperscript{1} measurements are performed
at the peptide level and a single peptide can often be mapped to multiple proteins, a strategy to
aggregate peptides into proteins must be performed. Therefore, we assign peptides to all possible
proteins and then take the Tukey’s biweight (161) to obtain a more robust intensity estimation for
each individual protein before data reformatting.

Following reformatting inter, prey and bait files the data is analyzed by SAINTexpress and
the CRAPome. SAINTexpress has been integrated into Galaxy by wrapping the tool in python, in
which a system call is performed on the machine running Galaxy from python to run
SAINTexpress. The resulting “list.txt” file is then passed using python back to Galaxy and placed
into the active history. In parallel, the prey file from the SAINT preprocessing tool is used as a
protein list to query the CRAPome database (129). APOSTL utilizes a local version of the
CRAPome database, which allows for rapid querying. As a result, a customized query tool was
developed to perform the same searches and calculations performed on the CRAPome website
(http://crapome.org/). In addition, SAINTexpress outputs can be analyzed using clustered dotplots
from the Prohits-Viz platform (162,163).

Once the necessary result files are generated, data is passed into the “interactive analysis
environment” of APOSTL. This environment is a personalized web application generated
specifically for each individual user using the R based web framework, shiny. Upon importing data
into the APOSTL shiny server, APOSTL calculates the normalized spectral abundance factor
(NSAF) (126), the NSAF score and the CRAPomePCT. The NSAF score is an empirical fold change between the NSAF for test preys versus control preys normalized by the number of controls

\[
\text{NSAF score}_T = \ln \left( \frac{\left( \frac{SpC}{L} \right)_T}{\frac{1}{N} \sum_{i=1}^{N} \left( \frac{SpC}{L} \right)_i} + \epsilon \right) = \ln \left( \frac{\text{NSAF}_T + \epsilon}{\frac{\text{NSAF}_C}{n} + \epsilon} \right)
\]

where \( SpC \) is the spectral count (or intensity) of each test \( T \) or control \( C \) prey, \( L \) is the number of amino acids in prey \( i \), \( n \) is the number of control purifications, and the constant \( \epsilon \) is added to prevent division by 0. We define \( \epsilon \) as the inverse of the average spectral abundance factor (SAF) across all control purifications.

\[
\epsilon = \frac{1}{\text{mean}\left( \frac{SpC}{L} \right)_C}
\]

The NSAF score is designed as a less stringent alternative to the SaintScore to complement the analysis. To conveniently flag frequent background proteins, APOSTL utilizes the CRAPome database to calculate the probability of a true interaction based on the abundance of each protein in the CRAPome

\[
\text{CRAPomePCT} = 100 \left( 1 - \frac{f_i}{N} \right)
\]

where \( f \) is the frequency of identifications of prey \( i \) in the CRAPome and \( N \) is the total number of experiments annotated in the CRAPome database. We assign a cutoff for \( \text{CRAPomePCT} \) of 80% for visualization where all proteins below 80% probability are shaded with a fixed color.

Following calculations, data merging, and import in the shiny server, APOSTL offers a variety of interactive graphs using the R shiny framework and \textit{ggplot2} (164). Users can perform quality-control analyses such as correlations between replicates, boxplots of user specified proteins, density plots of all quantitative information (log2 fold-change, SaintScore, NSAF,
logOddsScore, and NSAFscore) (Figure 9a–c), and 2D histograms which cluster similar baits to one another. In the sidebar, users can specify cutoffs for SaintScore, log$_2$ fold-change and NSAFscore to filter their data.

**Figure 9. Example APOSTL quality control analysis.** Visualizations include (A) replicate correlations, (B) protein boxplots, and (C) density plots.

Additionally, a protein exclusion list can be appended to remove common contaminant proteins or the bait proteins themselves. Following filtering users can generate individual bubble graphs for all baits where each protein is represented as a circle. Axes and bubble scaling are customizable with the same options as specified in the density plots. Optionally, users can provide a CRAPome file from the CRAPome Query tool to assist with data filtering by allowing for scaled bubble color based on the CRAPomePCT. Identified interactions can be visualized in an interactive protein interaction network using visNetwork ([http://visjs.org/](http://visjs.org/)). Finally, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis and Gene Ontology analysis can be performed using ClusterProfiler (165) and visualized as bar graphs (Figure 10a–e).
Figure 10. Example APOSTL analysis. Visualizations include (A) protein interaction networks, (B) bubble graphs, (C) filtered interaction networks, (D) KEGG bar charts, and (E) Gene Ontology bar charts.
Discussion

We have developed APOSTL, a Galaxy integrated software suite and analysis pipeline aimed to make the analysis of AP-MS data more accessible and reproducible. One of the main advantages of APOSTL over other analysis methods is the ability to analyze AP-MS data at the peptide level using MS1 intensity quantification as well as spectral counting. MS1-based measurements are typically more accurate with a better linear dynamic range. Thus MS1-based quantification can provide greater accuracy in the low abundance range, which is usually lost with spectral counting (166). Additionally, while APOSTL was developed for the analysis of label-free AP-MS, it could be extended to AP-MS experiments coupled to stable isotope labeling (167–169) using customized Galaxy workflows to extract data at the peptide level through the development of additional tools.

APOSTL is suitable for the analysis of a wide variety of affinity proteomics applications, including tandem affinity purification mass spectrometry, drug affinity chromatography and proximity dependent biotin identification (BioID) (170) across human, murine or yeast datasets. All of these applications benefit from more reproducible research and easier replication of results using APOSTL's ability to export analysis parameters and Galaxy's built-in workflow features. An additional benefit of APOSTL is a large increase in analysis efficiency. Individual steps in a typical AP-MS analysis pipeline have been described to take anywhere from 1 day to 1 week (171). APOSTL accomplishes much of this within a few minutes by automating preprocessing steps and by providing an intuitive data analysis environment. Because individual APOSTL tools are integrated using Galaxy workflows, APOSTL is also highly adaptable with the potential to incorporate custom workflows, additional analysis tools, and new statistical algorithms.

Alternatively, data can be exported in each stage in the pipeline for offline examination or for analysis using other bioinformatics tools. The scalability of the APOSTL Galaxy platform also offers an institution level solution for an affinity proteomics data analysis platform that is accessible to multiple researchers. This allows for parallel analyses to occur across an institution,
between institutions, and also enables for group training workshops. Galaxy-based tools and workflows, such as APOSTL, are easily distributed through the Galaxy ToolShed enabling any institution to set up local instances of the framework.
Chapter four: Materials and methods

(Note to reader: Parts of this section have been published previously in Remsing Rix et al. ACS Chemical Biology doi: 10.1021/cb400660a or Kuenzi et al Nature Chemical Biology doi: 10.1038/nchembio.2489 (89) and are being reproduced with permission from American Chemical Society copyright 2016 or Nature Publishing Group copyright 2017 (Appendix B).

Cell lines and transfections

Cell culture

HL60 cells were kindly provided by Dr. G. Reuther (Moffitt Cancer Center, Tampa FL) and were cultured in IMDM (20% FBS). U937 and KG-1 cells were a kind gift from Dr. G. Superti-Furga (CeMM, Vienna, Austria) and were cultured in RPMI 1640 (10% FBS) and IMDM (20% FBS) respectively. H650, H1155, HOP62, H661, H2342, H23, A427, A549, H292, H2122, H157, H1299, Calu-6, H1395, H1437, H322, HCC4006, Calu-3, H2170, H3122, H460, H226 and HCC2935 cells were provided by the Moffitt Lung Cancer Center of Excellence Cell Line Core and were cultured in RPMI 1640 media containing 10% FBS. All cell lines tested negative for mycoplasma contamination and cell line authentication was done by short-tandem repeat (STR) analysis. The H157 cell line used in this study is reported in the International Cell Line Authentication Committee database of commonly misidentified cell lines. H157 was only used to demonstrate the range of sensitivity to ceritinib in a large, diverse set of lung cancer cell lines. The batch used in this study has been authenticated by STR analysis. Studies performed with frozen primary patient lung tissue were approved by the institutional scientific review committee (H. Lee Moffitt Cancer Center and Research Institute). Written informed consent was obtained before sample collection.
Compounds

Tivantinib (Moffitt Chemistry Core and ChemieTek), PHA-665752 (Pfizer) and 6-bromoindirubin-3'-oxime (BIO, Cayman Chemical) were dissolved in DMSO (10 mM) and LiCl and NaCl (Sigma-Aldrich) were dissolved in sterile diH$_2$O (10 M and 6 M, respectively). OSI-906, ceritinib, CEP-37440, crizotinib, cabozantinib, ABT-199, ribociclib (Chemietek), PF-04217903, PF-573228, BID1870 (Selleckchem), Compound C (Sigma), FMK (Axon Medchem) and SL0101 (Millipore) were dissolved in DMSO (10 mM) and diluted in RPMI 1640 + 10% FBS for use.

RNA interference

siRNAs used were GSK3$\alpha$-1 (Qiagen, targeting 5'-AAGTGATTGGCAATGGCTCAT-3'), GSK3$\alpha$-2 (Cell Signaling, #6312), GSK3$\beta$-1 (Qiagen, targeting 5'-AAGTAATCCACCTCTGGCTAC-3'), GSK3$\alpha$/β dual-1 (Cell Signaling, #6301), GSK3$\alpha$/β dual-2 (Qiagen, targeting 5'-AAGAATCGAGAGCTCCAGATC-3'), MET (Santa Cruz, sc029397), GSK3$\beta$ SMARTpool (L-003010-00-0005), RSK1 SMARTpool (L-003025-00-0005), RSK2 SMARTpool (L-003026-00-0005), IGF1R SMARTpool (L-003012-00-0005), FAK1 SMARTpool (L-003164-00-0005), FER SMARTpool (L-003129-00-0005), PRKAA1 SMARTpool (L-005027-00-0005), CAMKK2 SMARTpool (L-004842-00-0005) and YB1 SMARTpool (L-010213-00-0005) and ON TARGET plus nontargeting (D-001810-10-20) (All Dharmacon). Dual knockdown of RSK1 and RSK2 was achieved by combining individual RSK1 and RSK2 siRNAs at a final concentration of 20 nM siRNA. Dual knockdown of GSK3$\alpha$ and GSK3$\beta$ was achieved either by combination of two siRNAs or through a single siRNA targeting both GSK3$\alpha$/β. Single knockdowns were supplemented with nontargeting siRNA (Dharmacon, D-001810-10-20) to a 20 nM final concentration of siRNA. Transfection of these siRNAs was conducted in a 6-well plate using Lipofectamine RNAiMAX Transfection Reagent (Invitrogen) according to the manufacturer's
instructions. Gene silencing was monitored by immunoblotting. Drug treatments were conducted 24 h after transfection at the indicated concentrations using complete RPMI as the diluent. Cells were counted in triplicate using Trypan blue and a hemocytometer, and analyzed in Excel or R.

Site-directed mutagenesis
pDESTmycYBX1 was a kind gift from T Tuschi (Addgene plasmid # 19878) (172). S102A and S102D mutants were created by Mutagenex and confirmed by sequencing. Plasmids were transiently transfected into H650 cells using Lipofectamine 2000 transfection reagent (Invitrogen) according to manufacturers’ instructions. Protein expression was monitored by immunoblotting. Drug treatments were conducted 24h after transfection at the indicated concentrations using RPMI 1640 (+10% FBS) as the diluent.

Chemistry
c-(-)-tivantinib synthesis
(-)-tivantinib, (+)-tivantinib and couplable analogs were synthesized by Dr. Harshani Lawrence and Yunting Luo in the Moffitt Cancer Center Chemistry Core (Appendix A) (Figure 11a). Cellular activity of (-)-tivantinib, (+)-tivantinib, c(-)-tivantinib, and c(+)-tivantinib against A549 cells was confirmed by cell viability assays to ensure that linker attachment did not significantly alter (-)-tivantinib target engagement (Figure 11b).
Figure 11. Compound synthesis and cellular evaluation of c-(-)-tivantinib. (a) Full synthetic route for preparation of (-)-tivantinib, (+)-tivantinib and their coupleable analogues. For a detailed description of the synthesis see Appendix A. (b) Enantiomeric differential inhibition of cellular viability of A549 by c-(-)-tivantinib and c-(+)-tivantinib compared with (-)-tivantinib and (+)-tivantinib.
**c-ceritinib synthesis**

The starting materials, ceritinib (15 mg, Chemietek, >99%), 3-(boc-amino) propyl bromide (3 eq.) and triethylamine (3 eq.) were dissolved in DMF (300 µL) and stirred under argon atmosphere at room temperature overnight. The reaction was monitored using HPLC–MS, and an additional 3-(boc-amino) propyl bromide (2 eq.) was added and stirred at room temperature overnight. HPLC–MS showed completion of the reaction with a minor double alkylated product (<10%). The crude reaction mixture was extracted using dichloromethane (5 mL) and 0.1 M NaHCO₃ (pH 8.5; 5 mL). The aqueous layer was washed with dichloromethane (3 mL × 3). The combined organic phase was washed with saturated NaCl (5 mL), dried (Na₂SO₄), filtered and concentrated to dryness to obtain a yellow, oily film. This product was further purified using dimercaptotriazine beads (20 EQ, 37 °C O/N) to remove excess 3-(boc-amino) propyl bromide and filtered through celite. Single and double alkylated products were then separated through silica (8% MeOH in DCM). LC–MS (ESI⁺) m/z: 715.3 [M+H]⁺. ¹H NMR (400 MHz, CD₃OD): δ 8.45 (d, J = 8.4 Hz, 1H), 8.15 (s, 1H), 7.92 (dd, J = 8.0, 1.5 Hz, 1H), 7.74 (s, 1H), 7.70 – 7.60 (m, 1H), 7.39 – 7.29 (m, 1H), 6.85 (s, 1H), 4.62 – 4.54 (m, 1H), 3.38 – 3.28 (m, 1H), 3.15 – 3.07 (m, 5H), 2.52 – 2.47 (m, 2H), 2.24 – 2.15 (m, 2H), 2.13 (s, 3H), 1.80 – 1.61 (m, 6H), 1.44 (s, 9H), 1.32 (d, J = 6.1 Hz, 6H), 1.25 (d, J = 6.8 Hz, 6H).

Purified boc-c-ceritinib was dissolved in dichloromethane/TFA (20%, 200 µL) and stirred for 15 min to perform boc-deprotection (Figure 12). The reaction was monitored using HPLC–MS, and showed completion of the reaction to obtain c-ceritinib (TFA salt). LC–MS (ESI⁺) m/z: 615.3 [M+H]⁺; HRMS (ESI⁺) m/z: calculated for C₃₁H₄₃ClN₆O₃S [M+H]⁺ 615.28786, found 615.28792 (0.10 p.p.m.). ¹H NMR (400 MHz, d₆-DMSO): δ 9.58 (s, 1H, disappears upon D₂O shake), 8.38 (d, J = 8.3 Hz, 1H), 8.33 (s, 1H, disappears upon D₂O shake), 8.28 (s, 1H), 7.89 (s, 2H, disappear upon D₂O shake), 7.83 (dd, J = 7.7, 1.1 Hz, 1H), 7.61 (dd, J = 7.8, 7.7 Hz, 1H), 7.48 (s, 1H), 7.37 (ddd, J = 8.3, 7.8, 1.1 Hz, 1H), 6.77 (s, 1H), 4.53 – 4.44 (m, 1H), 3.58 – 3.51 (m, 2H), 3.50 – 3.39 (m, 1H), 3.19 – 2.85 (m, 7H), 2.12 (s, 3H), 2.03 – 1.84 (m, 6H), 1.20 (d, J = 6.0 Hz, 6H), 1.13 (d,
J = 6.8 Hz, 6H). $^{13}$C NMR (100 MHz, d$_6$-DMSO): 159.38, 159.03, 158.68, 158.32, 157.46, 155.80, 154.17, 147.39, 138.29, 138.17, 135.26, 131.48, 127.16, 125.67, 124.75, 124.55, 117.82, 114.90, 111.60, 104.99, 104.82, 71.15, 55.15, 53.59, 52.76, 36.73, 34.82, 29.82, 22.24, 18.81, 15.27.

Figure 12. c-ceritinib synthesis. c-ceritinib was synthesized by adding an aminopropyl linker to terminal piperidine moiety.

Antibodies and western blotting

Cells were lysed using lysis buffer (0.20% NP40, 50 mM Tris pH 7.5, 5% Glycerol, 1.5 mM MgCl$_2$, 100 mM NaCl) containing Phosphatase Inhibitor Cocktail 2 (Sigma, P5726) and cOmplete
Protease Inhibitor Cocktail (Roche, 11873580001). Lysates were resolved by SDS-PAGE and incubated with primary antibodies. Antibodies used were against actin (A5441), CAMKK2 (HPA017389), ERK1/2 (M5670) (all Sigma) and pTyr279/Tyr216 GSK3 (Millipore, 05-413). Antibodies against β-catenin (sc-7199), Bak (sc-832), and MCL-1 (sc-819) were from Santa Cruz. Antibodies against GSK3α (#4337), GSK3β (#9315), pTyr1234/Tyr1235 MET (#3129), MET (#3127), p27Kip1 (#2552), pSer10 Histone H3 (#3377), Cleaved Caspase 3 (#9661), PARP-1 (#9542), BCL-XL (#2764), pSer473 AKT (#9271), AKT (#9272), ALK (#3633), pSer380 p90RSK (#12032), RSK1/2/3 (#9355), RSK1 (#9333), RSK2 (#5528), pSer235/236 RPS6 (#4858), RPS6 (#2217), pThr202/Tyr204 ERK1/2 (#4267), pTyr397 FAK1 (#8556), FAK1 (#13009), pTyr1131 IGF1R (#3021), IGF1R (#9750), AMPKα1 (#2795), pThr172 AMPKα1 (#2535), FER (#4268), pSer102 YB1 (#2900), and YB1 (#4202) were from Cell Signaling. Secondary antibodies were HRP-conjugated α-rabbit (NA934-1ML) or α-mouse (NA931-1ML, GE Healthcare).

Cell viability and growth assays

Cell viability

Cell viability assays were conducted according to the manufacturer’s specifications for CellTiter-Glo Luminescent Cell Viability Assay (Promega). Cells were seeded at 1,000 cells/well in a 384-well microtiter plate and treated after 24 h. Drugs were diluted in the respective culture medium at the indicated concentrations. Cells were treated for 72 h before the addition of CellTiter-Glo (Promega) reagent and read on an M5 Spectramax plate reader (Molecular Devices). Raw data for dose–response curves was analyzed using GraphPad Prism or preprocessed in python and analyzed in R using the drc package fitting a three parameter log-logistic regression function (173).
**Colony formation**

AML patient bone marrow mononuclear cells (BMNCs) were seeded into 6-well plates and treated overnight at the indicated concentrations of drug using IMDM (10% FBS) as the diluent. Treated cells were then collected and suspended in MethoCult™ H4034 Optimum methylcellulose medium (StemCell Technologies) containing additional drug, split into technical duplicates (200,000 cells/replicate) and plated in 30 mm cell culture dishes. A colony was defined as a cell cluster containing > 30 cells. Colonies were counted manually following 14 days of growth. Select samples were chosen for an additional readout after 19 days. Average colonies and standard deviation were calculated for each treatment. Samples for this project were archived and retrieved under both SRC and IRB approval for the Total Cancer Care® and Moffitt Cancer Center pilot protocol.

**Clonogenicity assay**

Cells were seeded in 6-well plates (~4000 cells/well) and treated with drugs after 24h. Plates were processed once control wells showed ~90% confluence. Cells were fixed with cold MeOH, stained with Crystal Violet, washed (dH2O) and imaged using a tabletop scanner.

**Immunoprecipitation assays**

Cells were seeded in 150 mm dishes (~4e6 cells / plate) and treated after 24h. Plates were harvested, lyzed and a total of 1 mg total protein was used for each immunoprecipitation. Antibodies were coupled to Protein A/G-Plus Agarose Beads (Santa Cruz) for 4h and incubated with 100 – 200 µL lysate (1:100) overnight. Beads were washed with lysis buffer followed by boiling elution and analyzed by immunoblotting. Antibodies used were FAK1 (#13009) and IGF1R (#9750, both Cell Signaling).
Proteomics

Drug affinity chromatography

Tethered inhibitor analogs or ampicillin were immobilized on NHS-activated Sepharose for Fast Flow resin (GE Healthcare) and blocked with ethanolamine overnight (ceritinib) or for 4h (tivantinib). Cell lines and patient samples were lyzed using 0.20% NP40, 50 mM Tris pH 7.5, 5% Glycerol, 1.5 mM MgCl2, 100 mM NaCl lysis buffer containing 25 mM NaF, 1 mM Na3VO4, 1 mM PMSF, 1 mM DTT, 30 µM TLCK, 30 µM TPCK, 1 µg/mL Leupeptin, 1 µg/mL Aprotinin, and 10 µg/mL Trypsin inhibitor. Total cell lysates containing 1 mg, 5 mg or 10 mg were then added to the affinity matrix for 2h or 6h. Competition experiments were performed via pretreatment of total cell lysate with 20 µM compound for 30min or 2h prior to affinity chromatography. Eluates were loaded on a Criterion™ XT Precast Gel (Bio-Rad), run briefly into the gel, fixed in 10% MeOH / 5% glacial acetic acid and excised as a single piece before performing an in-gel trypsin digestion.

Slices were washed in HPLC H2O and 50mM ammonium bicarbonate / 50% MeOH followed by reduction [2 mM triscarboxyethylphosphine (TCEP)/50 mM ammonium bicarbonate] and alkylation (20 mM iodoacetamine / 50 mM ammonium bicarbonate). Slices were then washed (50% MeOH / 50 mM ammonium bicarbonate), trypsin digested (20 ng/µL) overnight, and peptides extracted (50% – 100% acetonitrile / 0.1% TFA). Peptides were dried using vacuum centrifugation and resuspended in 2% acetonitrile / 0.1% formic acid before LC-MS/MS analysis.

Phosphoproteomics

For global phosphoproteomics (biological duplicate), cells were labeled using stable isotope labeling with amino acids in cell culture (SILAC; Invitrogen) (169) according to the manufacturer’s instructions. Briefly, H650 cells were grown in RPMI (+10% dialyzed FBS) containing [13C6]-, [15N4]-arginine and [13C6]-lysine or standard RPMI (+10% FBS) for 10 d. Incorporation of heavy label was confirmed by LC–MS/MS (>99%). For phosphotyrosine proteomics experiments
(biological and technical duplicates), cells were treated for 3h with 1.5 μM ceritinib and lyzed using 20 mM HEPES pH 8.0, 9 M urea, 1 mM Na3VO4, 2.5 mM Na4P2O7 and 1 mM β-glycerophosphate. Tyrosine phosphopeptides were enriched using the PTMScan Phospho-Tyrosine Mouse mAb (P-Tyr-100) Kit (Cell Signaling, #5636) according to the manufacturer's instructions. For SILAC-based experiments, samples were lyzed, digested and fractionated (SCX) into 12 concatenated fractions, and global phosphopeptide enrichment was done using PHOSSelect Iron Affinity Gel (Sigma) according to the manufacturer’s instructions.

**Liquid chromatography tandem mass spectrometry (LC-MS/MS)**

LC-MS/MS analysis of samples was performed on a linear ion trap (LTQ, Thermo Fisher) mass spectrometer coupled to a nanoflow liquid chromatograph (U3000, Dionex). Samples were first loaded onto a trap column (5 mm x 300 μm ID packed with C18 reversed-phase resin, 5 μm, 100 Å) and washed for 3 minutes at 8 μL/minute. The trapped peptides were eluted onto the analytical column (C18, 75 μm ID x 15 cm, Pepmap 100, Dionex). Peptides were eluted in a 60-minute gradient from 5% B to 45% B (solvent A: 2% acetonitrile + 0.1% formic acid; solvent B: 90% acetonitrile + 0.1% formic acid) with a flow rate of 300 nl/min. Five tandem mass spectra were collected in a data-dependent manner following each survey scan.

LC-MS/MS analysis of samples was performed on a hybrid LTQ-Orbitrap mass spectrometer (Thermo Fisher) coupled to a nanoflow liquid chromatograph (U3000, Dionex). Samples were first loaded onto a pre-column (5 mm x 300 μm ID packed with C18 reversed-phase resin, 5 μm, 100 Å) and washed for 8 minutes with aqueous 2% acetonitrile and 0.04% trifluoroacetic acid. The trapped peptides were eluted onto the analytical column, (C18, 75 μm ID x 15 cm, Pepmap 100, Dionex, Sunnyvale, CA). The 120-minute gradient was programmed as: 95% solvent A (2% acetonitrile + 0.1% formic acid) for 8 minutes, solvent B (90% acetonitrile + 0.1% formic acid) from 5% to 50% in 90 minutes, then solvent B from 50% to 90% B in 7 minutes.
and held at 90% for 5 minutes, followed by solvent B from 90% to 5% in 1 minute and re-equilibrate for 10 minutes. The flow rate on analytical column was 300 nl/min. Five tandem mass spectra were collected in a data-dependent manner following each survey scan. The MS scans were performed in Orbitrap to obtain accurate peptide mass measurement and the MS/MS scans were performed in linear ion trap using 60 second exclusion for previously sampled peptide peaks. For tivantinib experiments, data was searched against the UniProt human protein database (174) using the Sequest search engine (175). Carbamidomethylation of cysteine and methionine oxidation were selected as variable modifications. Three trypsin missed cleavages were allowed, the precursor mass tolerance was 1.2 Da. MS/MS mass tolerance was 1.0 Da. Search results were summarized in Scaffold 4.0. For ceritinib experiments, Mascot searches were performed against the UniProt human database (174) downloaded March 2015. Two trypsin missed cleavages were allowed, the precursor mass tolerance was 1.1 Da. MS/MS mass tolerance was 0.8 Da. Dynamic modifications included carbamidomethylation (Cys) and oxidation (Met). A minimum of two exclusive unique spectrum counts was required for protein identification. Mascot search results were summarized in Scaffold 4.4.5 and imported into R and Galaxy / APOSTL (132,141–144) for bioinformatic analysis.

LC-MS/MS analysis of samples on Q-Exacte was done using a nanoflow ultra high-performance liquid chromatograph (RSLC, Dionex) coupled to an electrospray bench top orbitrap mass spectrometer (Q-Exacte plus, Thermo). The sample was first loaded onto a pre-column (2 cm × 100 µm ID packed with C18 reversed-phase resin; 5 µm; 100 Å) and washed for 8 min with aqueous 2% acetonitrile and 0.04% trifluoroacetic acid. The trapped peptides were eluted onto the analytical column, (C18; 75 µm ID × 50 cm; 2 µm particle size; 100 Å pore size; Dionex). The gradient was programmed as follows: 95% solvent A (2% acetonitrile + 0.1% formic acid) for 8 min, solvent B (90% acetonitrile + 0.1% formic acid) from 5% to 38.5% in 90 min, then solvent B from 50% to 90% B in 7 min and held at 90% for 5 min, followed by solvent B from 90% to 5%
in 1 min and re-equilibration for 10 min. The flow rate on the analytical column was 300 nL/min. Sixteen tandem mass spectra were collected in a data-dependent manner following each survey scan. MS/MS scans were performed using 60s exclusion for previously sampled peptide peaks. For phosphoproteomics experiments, data were searched by MaxQuant v1.2.2.5 (156–159) using the UniProt human database (270–273) (downloaded 11/2014). Carbamidomethylation of cysteine, methionine oxidation and phosphorylation of serine, threonine and tyrosine were selected as variable modifications. For SILAC-based experiments, $^{[13]C_6}$-arginine and $^{[13]C_6}$-lysine were included as fixed modifications. Data was normalized using iterative rank order normalization (IRON)(176) and imported into R for bioinformatic analysis. Phosphopeptide identification and quantification by extracted ion chromatogram (XIC) was confirmed using Skyline (177). For tivantinib drug affinity chromatography experiments, data was searched against the UniProt human protein database (174) using the Mascot search engine. Carbamidomethylation of cysteine and methionine oxidation were selected as variable modifications. Two trypsin missed cleavages were allowed, the precursor mass tolerance was 1.2 Da. MS/MS mass tolerance was 1.0 Da. Search results were summarized in Scaffold 4.0 and imported into R and Galaxy / APOSTL (see Chapter three) (132,141–144) for bioinformatic analysis.

**Kinase assays**

*In vitro* kinase inhibition assays and IC$_{50}$ determinations were performed on the Reaction Biology Kinase Hotspot and Eurofins KinaseProfiler platforms using 10 µM ATP. Reaction Biology IC$_{50}$ values were determined using a 10 dose three-fold serial dilution series starting at 20 µM (Reaction Biology). Eurofins IC$_{50}$ values were determined using a nine dose three-fold serial dilution series starting at 20 µM.
Flow cytometry

Cell cycle
Cells were harvested following incubation with drug, fixed with 70% cold ethanol and stored at –20 °C until analyzed. Cells were washed with PBS, and cell cycle was determined by incubating the cells in a 1 µg/mL DAPI (4, 6-diamidino-2-phenylindole, Sigma)/0.1% Triton X-100/PBS solution and analyzed using a FACSCanto II benchtop analyzer (BD Biosciences). Raw data was processed using ModFit LT V3.2.1 (Verity Software House) and further analyzed in R.

Apoptosis
Treated cells were harvested and stained with Annexin V–APC (BD Biosciences) and 100 ng/mL DAPI according to the manufacturer’s instructions. Analyses were conducted using a FACSCanto II benchtop analyzer (BD Biosciences). Following gating, fcs files were imported into R for analysis. Plotting of scatter plots was done using Flowjo (Treestar, Inc.).

Differentiation
Treated cells were harvested and stained with PE Mouse Anti-Human CD11b/Mac-1 (BD Biosciences) to monitor cell differentiation. Analyses were conducted using a FACSCanto II benchtop analyzer (BD Biosciences). Following gating, fcs files were imported into R for analysis. Plotting of gated scatter plots was done using Flowjo (Treestar, Inc.).

Drug screening and synergy calculations
For off-target based drug screening, each cell line was seeded in 384-well microtiter plates and treated after 24h with 0.5µM or 2.5µM of each compound in biological duplicate. Raw data for the drug screen was pre-processed in excel and analyzed using Cluster3.0 (178). Unsupervised
hierarchical clustering was performed using Euclidean distance as the distance metric and complete linkage (furthest neighbor clustering) when merging clusters.

For drug combination screening, HL60 cells were seeded in 384-well microtiter plates and treated after 24h with 0.5 µM or 2.5 µM of each compound in biological duplicate. Raw data for the drug screen was pre-processed in Excel and analyzed using R. Hit compounds were selected by fold change (from DMSO) and by translational relevance. Drug combination effects were further evaluated using Bliss (179) or the combination index (CI) method described by Chou–Talalay (180,181) using CompuSyn software.

Microscopy

Fluorescent caspase activity imaging

Cells were seeded at 1000 cells/well in a 384-well microtiter plate. Drugs were diluted in media containing IncuCyte Caspase-3/7 Reagent (#4440), and cells were treated 24h after plating. Fluorescence and confluence were measured every 2h for 72h using an IncuCyte Live Cell Analysis System. Processed data was then imported into R for analysis.

Immunohistochemistry and tissue microarray analysis

Patient tissue microarray (TMA1) and PDX tissue microarray (TMA2) were constructed by US Biomax and Charles River Laboratories, respectively. Slides were deparaffinized and rehydrated in successive washes of a xylene/ethanol gradient followed by HIER (pressure cooker) in sodium citrate buffer (10 mM sodium citrate, 0.05% Tween 20, pH 6). Nonspecific binding was blocked by incubating with 1.5% BSA/PBST, and slides were incubated overnight in 0.5% BSA/PBST with anti-FAK1 pY397 (Invitrogen, 700255). Slides were washed twice with PBST, incubated with EnVision+ anti-rabbit (K4011, Dako) for 1 h and visualized by diaminobenzidine (DAB). Slides were counterstained with hematoxylin, rehydrated and hard-mounted. Each tumor was manually
scored for the percentage of cells staining negative, weakly, moderately, or strongly (0, 1, 2 or 3, respectively) by a board-certified pathologist (Dr. Theresa Boyle, MD, PhD). H-scores were calculated for each tumor as follows:

\[ H - \text{score} = 0(X_0) + 1(X_1) + 2(X_2) + 3(X_3) \]

where \( X \) indicates the percentage of cells in each staining intensity category from negative to strongly positive. Replicate tumors for array TMA2 were averaged (mean) before analysis. Tumor data for TMA1 and TMA2 were merged and analyzed in R. We assigned a cutoff of pFAK1 high as an H-score \( \geq 130 \), and defined moderate pFAK1 as an H-score of 100-130. H-score cutoffs were calculated based on previous literature (182) in which \( \sim25\% \) of lung adenocarcinoma (LUAD) had been defined as having high pFAK1 staining. Based on the observation that that our pFAK1 high population was significantly higher than all analyzed LUAD cell lines, we assigned cutoffs such that 25% LUAD had moderate pFAK1 staining.

**Bioinformatic and statistical analysis**

**Gene expression profiling data analysis**

GSK3\( \alpha \) and GSK3\( \beta \) expression levels across different human myeloid lineages were queried using the Bloodpool aggregation of hematopoietic expression profiles from numerous studies catalogued in the manually curated BloodSpot database (183), which provides gene expression profiles of a number of mouse/human hematopoietic cells (normal and AML).

**Gene silencing sensitivity profiling data analysis**

shGSK3\( \alpha \) and shGSK3\( \beta \) sensitivity data was downloaded from the Project DRIVE database (https://oncologynibr.shinyapps.io/drive/) and imported into python for analysis (184). The redundant siRNA activity (RSA) sensitivity was used as sensitivity measurement as RSA sensitivity is calculated using all shRNA reagents against a given gene to determine a score (185).
**Tivantinib sensitivity prediction**

In order to predict sensitivity of AML cell lines to tivantinib, we built a regularized linear regression model (elastic net) to select gene features that can predict a tivantinib response vector. Elastic net regularization is a machine learning algorithm that is specially suited for the case of many more input features (genes) than samples (cell lines). Candidate predictive features were selected from 18989 genes with normalized measures of gene expression in CCLE for cell lines that have tivantinib sensitivity data in CTRPv2 (n=297) (90). Data was split into training (0.75) and test sets (0.25). Let $X \in \mathbb{R}^{n \times p}$ be the matrix of predictive features, where $n$ is the number of cell lines included in the training set and $p$ is the number of features. Let $y \in \mathbb{R}^n$ be the vector of sensitivity values for the same cell line panel. The elastic net attempts to find the weighted ($\beta$) linear combination of columns of features (genes) that can best approximate tivantinib AUC ($y$) or by solving the following:

$$\arg\min_{\beta} \left\{ ||y - X\beta||_2^2 + \lambda \left( \alpha ||\beta||_2^2 + (1 - \alpha)||\beta||_1 \right) \right\}$$

where $\lambda$ and $\alpha$ are tunable parameters where $\lambda$ controls the overall penalty and $\alpha$ controls the mixing ratio of L1- and L2-norm. We optimized $\lambda$ and $\alpha$ for the model with a tuning grid of 1000 values of $\lambda$ from 10e-10 to 10e10 and 10 values of $\alpha$ from 0 to 1 using 10000 iterations of 10-fold cross validation. The values of $\lambda$ and $\alpha$ were chosen to be those that minimized the root mean square error for each fold. The trained model was then used to predict AML cell line sensitivity (n=34) to tivantinib (for which data does not exist in CTRPv2 and thus not used in model training). Statistically significant differences in tivantinib sensitivity between AML and non-AML cell lines was determined by Kolmogorov-Smirnov test.
Statistical analysis

Histograms and F-tests were used to test for normality and compare sample variances. Two-sided $t$-tests, Wilcoxon Rank sum tests, and Kolmogorov-Smirnov tests were performed where appropriate. A minimum of three replicates were used to establish significance.

Drug-protein interaction identification

Following LC-MS/MS analysis, high confidence drug-protein interactions are identified using one of two methods. A filtering approach based on a series of scoring criteria is implemented to identify relevant drug targets. Criteria include: high fold change over negative controls (competition), high NSAF values, low standard deviation, and whether the protein is a kinase. Each protein is scored and the protein list is filtered and sorted for proteins with the highest score. Proteins with top scores are selected for follow up. Additionally, outputs from LC-MS/MS experiments are imported into Galaxy (141–144) and formatted for SAINTexpress (128) analysis by APOSTL (132). Proteins are then scored by the CRAPome (129), and analyzed using APOSTL’s interactive environment. Proteins with a high SaintScore, log$_2$ fold change, and NSAF values are prioritized for follow up by *in vitro* kinase assays. As drug affinity chromatography can pull down direct targets as well as binding partners, all kinases identified were included in subsequent network analysis where specified.

Phosphoproteomic analysis

For phosphoproteomics experiments, MS1 intensity was used as the measure of abundance. Following IRON normalization, data were filtered for PEP score $< 0.1$, contaminants were removed and rows with all 0 intensity values were excluded from analysis. Log$_2$ fold change values were calculated for both pY (label free) and pSTY (SILAC) data sets, and student $t$-tests
were performed for the pY data set. Cutoffs of fold change = 2 and P value = 0.1 for network analysis were assigned.

Network analysis
Proteins passing filtering from the chemical (kinases) and phosphoproteomics experiments were queried for known interactions using STRING (confidence > 0.9) (186). The resulting undirected network was imported into GEPHI for visualization and analysis (187). The community structure of the network was analyzed by calculating the modularity (see below) (33,34) to identify topological modules. Modules 1–4 were selected for KEGG pathway analysis, which was done using the ClusterProfiler R package (165). The proteins present in the adherens junction, insulin signaling, mTOR signaling and focal adhesion pathways were merged with the KRAS pathway (KRAS, RAF, MEK) and other key upstream / downstream nodes not identified by proteomics (PI3K, AKT, YB1, p70RSK, RHOA). The resulting proteins were queried by STRING to generate the subnetwork. Eigenvector centrality and optimal communities were calculated using the igraph R package (see below for details) (188). Hive plot was made using the HiveR package (189), and the adjacency matrix was made using ggplot2 (164).

Network community optimization
To identify the underlying community structure in medium to large scale networks we performed a fast-greedy community optimization using an agglomerative hierarchical clustering method (33,34) (Figure 13). At initiation, each node in the network belongs to its own community resulting in |V| communities. At each iteration, the algorithm merges pairs of communities and chooses the merge pair for which the resulting modularity is maximized. Modularity is defined as

\[ Q = \frac{1}{2|E|} \sum_{ij} \left[ A_{ij} - \frac{k_i k_j}{2|E|} \right] \delta_{c_i c_j} \]
where $k_i$ is the degree of node $i$, $A_{ij}$ is an element of the adjacency matrix, $|E|$ is the total number of edges in the network, $\delta_{c_ic_j}$ is the Kronecker delta, and $c_i$ is the label of the community to which node $i$ is assigned. Therefore the change in $Q$ upon joining two communities $c_i$ and $c_j$ is

$$\Delta Q_{c_ic_j} = 2 \left( \frac{|E_{c_ic_j}|}{2|E|} - \frac{|E_{c_i}||E_{c_j}|}{4|E|^2} \right)$$

where $|E_{c_ic_j}|$ is the number of edges from community $c_i$ to community $c_j$ and $|E_{c_i}|$ is the total degrees of nodes in community $c_i$. For each iteration, the algorithm stops once all nodes in the network are in a single community after $(|V| - 1)$ steps of merging. The partition with the largest modularity value is the result of the algorithm. For visualization, we then collapse all nodes belonging to individual communities into single nodes scaled on the total number of nodes within the community and assign edges between community $c_i$ and community $c_j$ scaled by $|E_{c_ic_j}|$.

Figure 13. Network community optimization. Circles represent network nodes, lines represent edges. Color indicates community assignment.
Functional module determination

Based on the assumption described in Chapter two, we then translated the topological modules determined from community optimization into functional modules using the most statistically overrepresented gene ontologies present within each module. To determine whether any GO terms annotate a specified list of genes in the module of interest at frequency greater than that would be expected by chance, we used ClusterProfiler (165) to calculate a p-value using the hypergeometric distribution.

Eigenvector centrality

To prioritize critical nodes with high influence in the network and therefore high importance, we calculated the eigenvector centrality for all nodes within the network. For a given network $G = (V,E)$ with $|V|$ nodes and $|E|$ edges, let $A = (a_{i,j})$ be the adjacency matrix. The eigenvector centrality of node $i$ can be defined as

$$x_i = \frac{1}{\lambda} \sum_{j \in M(i)} x_j = \frac{1}{\lambda} \sum_{j \in E} a_{i,j}x_j$$

where $M(i)$ is a set of neighbors of node $i$ and the eigenvalue $\lambda$ is a constant. The resulting eigenvector is then normalized to the total number of nodes $n$.

NetworKIN analysis

Phosphorylation information for altered pSTY peptides was formatted and input into NetworKIN (http://networkin.info/) (190). The output was filtered for kinase-substrate interactions with NetworKIN sore > 2. The resulting predicted-kinase substrate interactions were used to generate a data-dependent kinase-substrate subnetwork containing significantly altered pSTY peptides for the phosphoproteomics and known / newly identified ceritinib targets.
**ReKINect analysis**

Mutational data for the H650 cells was obtained from CCLE (191) and filtered for missense mutations and mutations that are known to yield errors in ReKINect. The corresponding mutations were mapped onto the protein reference sequence (UniProt) to generate a mutant FASTA file containing sequences of both the reference and mutated proteins. The mutant FASTA was analyzed using ReKINect (192) ([http://rekinect.info/home](http://rekinect.info/home)), and mutated proteins were queried for pathways (KEGG) using ClusterProfiler (165).
Chapter five: GSK3α and GSK3β are functionally relevant targets of tivantinib

(Note to reader: Parts of this section have been published previously in Remsing Rix et al. ACS Chemical Biology doi: 10.1021/cb400660a (193) and are being reproduced with permission from American Chemical Society copyright 2016 (Appendix B).

Glycogen Synthase Kinase 3 (GSK3)

Structure and biological functions of GSK3

Glycogen synthase kinase 3 (GSK3) was originally discovered as one of the kinases responsible for phosphorylating glycogen synthase (194). It was later discovered that GSK3 was actually two separate isoforms, a 51kDa protein (GSK3α) and a 47 kDa protein (GSK3β), encoded by their own genes (195). GSK3α and GSK3β are both serine/threonine kinases with ~98% homology in their kinase domains (195). Based on this similarity, much of the current research does not distinguish between GSK3α and GSK3β. However, even though they are structurally very similar, they still have some cellular functions that are unique to GSK3α and GSK3β. These GSK3α and GSK3β specific functions are not fully understood but are highlighted by the inability of GSK3α to rescue GSK3β-null mice from embryonic lethality as well as AML expressing a specific sensitivity toward knockdown of GSK3α (196,197). With many biological functions beyond their role in regulating glycogen synthesis, it will be rather difficult to fully elucidate these GSK3α and GSK3β specific functions. While this makes studying GSK3 biology difficult, the fact that GSK3 has many biological functions also makes it a desirable drug target. In this section we will highlight GSK3’s
role in the cancer–related biological functions of cell proliferation / survival, glycogen synthesis, microtubule regulation, and protein synthesis.

**Cell proliferation and survival.** One of the most well-described roles of GSK3 is its involvement in WNT signaling which has been associated particularly with colorectal cancer, but also plays an important role in many other malignancies (Figure 14a)(198). The GSK3 protein pool participating in WNT signaling exists in a multiprotein complex, known as the β-catenin destruction complex (199). This complex includes axin, β-catenin, adenomatous polyposis coli (APC) and GSK3 (200–202). Without WNT activation, active GSK3 phosphorylates each member of the complex (axin, APC, β-catenin) thereby promoting axin stabilization. Once stabilized, the interaction between APC and β-catenin is facilitated and promotes β-catenin ubiquitination and degradation. However once WNT binds to its receptor, “frizzled”, GSK3 binding to axin is disrupted by disheveled (DVL) and ‘frequently rearranged in advanced T-cell lymphomas’ (FRAT). GSK3 is then released from the destruction complex leading to β-catenin stabilization, accumulation in the cytoplasm and translocation of β-catenin to the nucleus. Here β-catenin binds to TCF/LEF transcription factors promoting expression of pro-survival and proliferation genes (203–205).

**Glycogen synthesis.** GSK3’s role in glycogen synthesis stems from its discovery close to 40 years ago when it was found to be a protein kinase capable of phosphorylating glycogen synthase (194). Glycogen synthase is a key enzyme regulating the conversion of glucose to glycogen for energy storage (206). GSK3 regulates this process by phosphorylating and inhibiting glycogen synthase in the absence of insulin (Figure 14b). Once global glucose stores are low, insulin is secreted, stimulating insulin receptor. Activated insulin receptor signals through the PI3K-AKT axis which phosphorylates GSK3α/β at S21/9, respectively, and inactivating them (207–210). Without its inhibitory phosphorylation, glycogen synthase is then able to promote glycogenesis. Interestingly, glycogen metabolism has recently gained a greater focus in the
context of cancer as it is upregulated in many tumor types, likely as a result of the Warburg effect (211,212).

**Microtubule regulation.** In addition to its canonical roles in WNT signaling and glycogenesis, GSK3 has been shown to localize to the microtubules and alter microtubule dynamics which is critical to GSK3’s involvement in Alzheimer’s disease (213–217). GSK3 does this primarily through tau regulation (Figure 14c). Tau is a microtubule associated protein that functions in the formation and maintenance of microtubules. Active GSK3 phosphorylates tau at

![Diagram of GSK3 functions](image)

**Figure 14. Biological functions of GSK3.** (a) cell proliferation and survival, (b) glycogen synthesis, (c) microtubule dynamics and (d) protein synthesis.
numerous residues which promotes tau dissociation from the microtubules (217–219). Once tau dissociates, microtubules become destabilized and depolymerization can occur. Interestingly, tau has been described as a predictive marker to the tubulin inhibitor, paclitaxel, in breast cancer in which patients harboring tumors expressing low levels of tau have microtubules that are more susceptible to destabilization by paclitaxel. These tumors were thus found to be more likely to respond to paclitaxel therapy (220).

**Protein synthesis.** Similar to the mechanism in which GSK3 regulates glycogen synthase, GSK3 also regulates the elongation initiation factor eIF2B and therefore protein synthesis (Figure 14d). In the absence of insulin, GSK3 places an inhibitory phosphorylation on eIF2B (208,221,222). In the presence of insulin, GSK3 is inhibited by phosphorylation of S21/9 by AKT resulting in activated eIF2B. Activated eIF2B acts as a guanine exchange factor (GEF) for its substrate eIF2 which binds tRNAs to the ribosome thus promoting the initiation of translation and protein synthesis (223). Recently, GSK3 has been shown to regulate protein synthesis through other proteins as well including inhibitory phosphorylation of 4EBP1, a protein that inhibits protein synthesis by sequestering the mRNA cap binding protein eIF4E away from the eIF4G complex (224).

**State of tivantinib in NSCLC**

The receptor tyrosine kinase c-MET is involved in cell migration and metastasis of various malignancies (225). c-MET amplification and activity provides an important mechanism by which cancer cells develop resistance to targeted drugs, such as EGFR inhibitors in non-small cell lung cancer (NSCLC) (226). Thus, c-MET is an attractive therapeutic target and several inhibitors are currently in clinical development. One of the most advanced c-MET inhibitors at the initiation of this study, tivantinib (ARQ197) (220), has progressed into phase III trials (62,228,229). In phase II studies, tivantinib displayed clinical activity in NSCLC patients in combination with the EGFR inhibitor erlotinib, particularly in patients with KRAS mutations (62). This was unexpected as the
primary rationale for testing tivantinib in NSCLC was to prevent emergence of resistance to erlotinib due to compensatory c-MET signaling in patients with EGFR mutations, which are mutually exclusive with KRAS mutations (230). Moreover, although described to be highly selective for c-MET, reportedly due to its unique ATP-independent binding mode (227,231), tivantinib showed anticancer activity in various cell lines across diverse tumor types, many of which are not driven by c-MET signaling (227). We therefore hypothesized that tivantinib inhibits a wider range of targets than appreciated and that some of these are functionally relevant for its activity. Further supporting this hypothesis, two recent studies suggest that tivantinib’s anticancer activity in different tumor types may be related to modulation of microtubule dynamics rather than c-MET inhibition (165–167).

**Tivantinib off-target activity**

_Tivantinib has cellular activity independent of c-MET inhibition_

To obtain a broader view of tivantinib’s activity in lung cancer, we screened a panel of 24 KRAS-mutant and wild-type lung cancer cell lines (Figure 15a). Tivantinib inhibited the viability of the majority of these whereas crizotinib (PF-02341066), PF-04217903 and cabozantinib (XL-184), which are much more potent c-MET kinase inhibitors than tivantinib, had no significant effects. There was no obvious link between tivantinib sensitivity and KRAS mutation status. Determination of the IC\(_{50}\) values for inhibition of cellular viability confirmed the differential activity of these compounds with tivantinib displaying an IC\(_{50}\) of about 500 nM for the most sensitive NSCLC cell lines. In comparison, the highly selective c-MET inhibitor PF-04217903 and the less selective crizotinib had no measurable or only weak cellular activity, respectively (Figure 15b). Confirming the functional integrity of these compounds, though, c-MET autophosphorylation in A549 cells was effectively inhibited by crizotinib, PF-04217903 and another widely used c-MET inhibitor, PHA-665752, whereas tivantinib showed essentially no effect (Figure 15c). Considering the
reported maximum plasma concentration of 5-7 $\mu$M from phase I clinical trials (235,236), tivantinib's activity against several of these cell lines was well within physiologically relevant concentrations. In summary, tivantinib displayed potent activity against a broad panel of lung cancer cell lines, which was unrelated to inhibition of c-MET kinase activity and KRAS mutation status.

**Tivantinib inhibits lung cancer viability through inhibition of GSK3**

Tivantinib's in vitro inhibition profile was originally determined against a panel of 230 kinases, based on which it was considered a highly selective c-MET inhibitor (227). In light of our data, however, we hypothesized that one or more of the remaining approximately 300 protein kinases...
in the human kinome could be previously unrecognized tivantinib targets responsible for the cellular activity in NSCLC cells. We therefore applied a mass spectrometry-based chemical proteomics strategy to characterize tivantinib’s target profile in NSCLC cells in a proteome-wide and unbiased fashion (122). To this end, we designed the tivantinib analogue c(-)-tivantinib (9, Figure 11) based on the reported co-crystal structure of tivantinib with c-MET, which suggests that the indole moiety is solvent accessible when binding to a kinase (231). According to our previous experience performing chemical proteomics with various kinase inhibitors, similar structure-activity relationships are likely maintained across the majority of targets (237,238). c(-)-Tivantinib (9) was synthesized by modifying the published synthetic route to tivantinib (8) and its enantiomer (+)-3S,4S-tivantinib (7), which has been reported to be inactive against c-MET (Figures 11a) (231,239). Interestingly, 7 also showed substantially weaker activity in (-)-tivantinib-sensitive cell lines making it an excellent control compound (Figure 11b). Separation of the racemic mixture resulting from reduction of the maleimide intermediate 6 by preparative chiral HPLC yielded the optically pure enantiomers c(-)-tivantinib (9) and c(+)-tivantinib (10), which retained differential cellular activity, albeit somewhat reduced possibly due to altered cell permeability (Figure 11b). Chemical proteomics with A549 total cell lysates using analogues 9 and 10 (the latter as control) identified several protein kinases not previously implicated as tivantinib targets (Figure 16a). Applying stringent filters for reproducibility between replicates and enrichment by 9 over 10, GSK3\(\alpha\) and GSK3\(\beta\) were found to be the highest confidence target candidates that interacted with tivantinib. Subsequent immunoblotting confirmed the interaction between c(-)-tivantinib (9) and GSK3\(\alpha\) and GSK3\(\beta\) and suggested selectivity of (-)-tivantinib over both (+)-tivantinib and blocked beads (no immobilized drug, ruling out non-specific binding to the matrix) for these kinases (Figure 16b). Furthermore, competition with the potent ATP-competitive pan-GSK3 inhibitor BIO suggested that tivantinib interacts directly with GSK3\(\alpha\) and GSK3\(\beta\) by binding to their ATP-binding pockets. In comparison, c-MET was only weakly detectable by LC-
MS/MS analysis in A549 cells and did not pass our stringent filtering criteria. Furthermore, while a more sensitive targeted proteomics analysis using multiple reaction monitoring (MRM) mass spectrometry with stable isotope-labeled standard (SIS) peptides also detected c-MET as a tivantinib binder, absolute quantification demonstrated that GSK3α and GSK3β were much more strongly enriched than c-MET (Figure 16c). All three kinases were selectively recovered by c-(−)-tivantinib (9) over c-(+)-tivantinib (10) (Figure 16d). Chemical proteomics using H1648 cells, which are known to overexpress c-MET, again prominently enriched for GSK3α and GSK3β, but now also for c-MET (Figure 17a). Consistent with in vitro kinase assays (Figure 17b), these results suggest that c-MET is only a weak tivantinib target. Instead, our data demonstrate that tivantinib is prominently binding to GSK3α and GSK3β.

Consistent with GSK3 inhibition, an increase in total β-catenin levels, which is degraded upon phosphorylation by GSK3 (240), was detected following tivantinib treatment of A549 cells (Figure 17c). To further evaluate the functional consequences of tivantinib binding to GSK3α and GSK3β, enzyme inhibition was determined by in vitro kinase assays. These experiments demonstrated that (−)-tivantinib potently inhibits GSK3α and GSK3β with IC₅₀ values in the upper nanomolar range, whereas (+)-tivantinib is a significantly weaker inhibitor of both kinases (Figure 18a). Thus, GSK3α and GSK3β are more potently inhibited than c-MET by (−)-tivantinib and the differential activity between the tivantinib enantiomers is supported by the difference observed in their cellular activities. Notably, in contrast to the vast majority of GSK3 inhibitors, which either affect GSK3α and GSK3β with equal potency or display selectivity for GSK3β (241), tivantinib inhibits GSK3α 2-3-fold more strongly than GSK3β. Although recently, there has been significant progress towards GSK3α-specific inhibitors (242,243), tivantinib constitutes to the best of our knowledge the first GSK3 inhibitor in clinical development that has selectivity for GSK3α. Furthermore, although other targets with different structure-activity relationships cannot be
completely ruled out, tivantinib appears to be much more selective for GSK3α and GSK3β on a kinome-wide level than established GSK3 inhibitors. This relative GSK3α and GSK3β specificity, as well as the kinome-wide target selectivity, may harbor novel therapeutic opportunities for tivantinib in malignant diseases, such as acute myeloid leukemia or pancreatic cancer, in which GSK3α has recently been described to be a promising new target (244, 245). Furthermore, GSK3 inhibitors may have utility in Alzheimer’s disease, diabetes and bipolar disorder (246, 247). Considering the narrow therapeutic index of the FDA-approved GSK3 inhibitor LiCl, the clinical

Figure 16. Determination of tivantinib’s kinase target interaction profile in A549 cells by chemical proteomics. (a) Phylogenetic tree of the human protein kinome displaying c-(-)-tivantinib targets. Node size indicates the spectral abundance factor. Node color corresponds to the ratio of the number of unique spectra identified by c-(-)-tivantinib to c-(+)-tivantinib. (b) Immunoblot analysis of GSK3α/β purification by c-(+)-tivantinib and c-(+)-tivantinib affinity chromatography. (c) Absolute quantification of GSK3α/β and c-MET peptides identified by LC-MS/MS analysis of c-(+)-tivantinib affinity purifications using LC-MRM. (d) Relative abundance ratios of GSK3α, GSK3β and c-MET peptides between c-(+)-tivantinib and c-(+)-tivantinib affinity purifications. The Human Kinome Map was adapted with permission from Cell Signaling Technology (www.cellsignal.com). TCL: total cell lysate; BB: blocked beads; Tiva: tivantinib.
safety and remarkable target selectivity of tivantinib may offer an interesting alternative. Both compounds showed antiproliferative activity in A549 cells (Figure 18b). Although initially it may seem high, the low millimolar activity of LiCl is in excellent agreement with GSK3-dependent growth inhibition in leukemia (244,248). Like tivantinib, BIO was able to induce apoptosis in A549 cells, albeit slightly delayed, as indicated by PARP-1 cleavage (Figure 18c). Curiously, we also noted decreased c-MET levels upon treatment with either compound. Regulation of c-MET levels may offer an explanation for why tivantinib has clinical activity in tumors driven by mutant or

![Figure 17. Characterization of tivantinib's kinase target interaction profile in H1648 cells by chemical proteomics.](image)

(a) Phylogenetic tree of the human protein kinome displaying targets identified by affinity chromatography experiments c(-)-tivantinib in H1648 cells. Node size indicates the spectral abundance factor. Node color corresponds to the ratio of the number of unique spectra identified by c(-)-tivantinib to c(+(+)-tivantinib enrichments. Displayed are all protein kinases that have been observed in both biological replicates of the (+)tivantinib affinity purifications with more than 2 unique spectra across both replicates. (b) Percent remaining MET kinase activity in in vitro kinase assay following tivantinib treatment at the indicated concentrations. (c) Immunoblot of β-catenin levels following treatment with tivantinib or 5-BIO at the indicated concentrations (µM).
overexpressed c-MET. Interestingly, siRNA-mediated knockdown of GSK3α and GSK3β suggest that these kinases have non-redundant functions in A549 cells (Figure 18d). Loss of GSK3β was primarily responsible for reducing c-MET levels, whereas GSK3α knockdown had a stronger effect on The combination of GSK3α- and GSK3β-specific siRNAs or using single siRNAs that cause simultaneous knockdown of both genes, however, was yet more effective in causing apoptosis (Figures 18d). Dual knockdown of GSK3α and GSK3β furthermore caused a significant reduction of viability in A549 cells compared to non-targeting siRNA (Figures 18e), which is consistent with the essential functions of GSK3α and GSK3β in many cells.

<table>
<thead>
<tr>
<th>GSK3α</th>
<th>IC50 [μM]</th>
<th>Millipore</th>
<th>Reaction Biology</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-) Tivantinib</td>
<td>659</td>
<td>112</td>
<td></td>
</tr>
<tr>
<td>(+) Tivantinib</td>
<td>1869</td>
<td>928</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GSK3β</th>
<th>IC50 [μM]</th>
<th>Millipore</th>
<th>Reaction Biology</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-) Tivantinib</td>
<td>1855</td>
<td>199</td>
<td></td>
</tr>
<tr>
<td>(+) Tivantinib</td>
<td>6944</td>
<td>1860</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 18. Functional analysis of GSK3α and GSK3β inhibition.** (a) IC50 values for inhibition of GSK3α and GSK3β in vitro kinase activity by (-)-tivantinib and (+)-tivantinib on Millipore and Reaction Biology assay platforms (mean values of independent experimental duplicate analyses). (b) Effect of tivantinib and the known pan-GSK3 inhibitors BIO and LiCl on cell viability of A549 NSCLC cells. (c) Effects of tivantinib and BIO (in μM) on A549 cells regarding PARP-1 cleavage and total c-MET levels. (d) Effects of individual or combined siRNA-mediated knockdowns of GSK3α and GSK3β in A549 cells on PARP-1 cleavage and c-MET levels 72 hours post-transfection. GSK3α1+β1: combination of gene-specific siRNAs GSK3α-1 and GSK3β-1; GSK3α/β-dual-1: single siRNA targeting both genes. NT: non-targeting siRNA. (e) Effects of individual or combined GSK3α and GSK3β knockdowns on cell number of A549 cells 96 hours post-transfection.
In the context of cancer, GSK3 is generally considered a tumor suppressor and thus, inhibition of GSK3 by tivantinib as a potential mechanism-of-action in NSCLC is initially counterintuitive. However, in addition to the aforementioned studies, several recent reports propose GSK3 to be a promising anticancer target, e.g. in glioblastoma multiforme and MLL-rearranged leukemia (248,249), suggesting that GSK3 function is context-dependent. Consistently, tivantinib also potently inhibited viability of MV-4-11 and RS4;11 MLL-rearranged leukemia cells, stabilized β-catenin and induced G1 cell cycle arrest, as has been described previously for GSK3 inhibitors in these cells (Figure 19a-b). In addition, tivantinib was able to cause an increased stabilization of β-catenin levels consistent with GSK3 inhibition (Figure 19c). In lung cancer, tivantinib induced G2/M arrest, which is consistent with previous reports (Figure 19d-e) (232,233).

**Figure 19. Characterization of GSK3α and GSK3β inhibition.** (a) Effect of tivantinib and the known pan-GSK3 inhibitors BIO and LiCl on cell viability of MV-4-11 and RS4;11 cells. (b) Cell cycle effects of tivantinib and LiCl treatment in MV-4-11 cells. (c) Immunoblot of β-catenin levels following treatment with tivantinib or 5-BIO in MV-4-11 cells at the indicated concentrations (µM). (d-e) Cell cycle effects of tivantinib treatment in A549 cells by (d) immunoblotting for pS10 Histone H3 and (e) DAPI DNA staining.
In summary, applying an unbiased, integrated chemical biology approach, we have identified GSK3α and GSK3β as novel targets of the clinical kinase inhibitor tivantinib in NSCLC cells. Importantly, tivantinib is a more potent inhibitor of GSK3α and GSK3β than of its intended target c-MET, which is only weakly inhibited. Furthermore, simultaneous loss of function of GSK3α and GSK3β causes apoptosis in lung cancer cells suggesting that inhibition of these kinases rather than c-MET plays an important role in tivantinib’s mechanism-of-action in NSCLC cells. Previous reports have suggested that tivantinib’s anticancer activity is due to disruption of microtubule dynamics(232,233). Considering the large number of substrates and pathways that are modulated by GSK3, many of which also affect microtubules such as Tau (240), it is possible that the observed effects are due to GSK3 inhibition. Further global and unbiased studies are required, however, to elucidate the exact downstream mechanism, by which tivantinib impairs cell viability. We anticipate that this knowledge will provide the basis for identification of relevant biomarkers that will greatly benefit tivantinib’s further clinical development, which is hampered by the erroneous assumption that tivantinib is a potent c-MET inhibitor. In addition, identification of GSK3α/β as tivantinib targets may pave the way for novel clinical studies with tivantinib, for instance in MLL-rearranged leukemia or in acute myeloid leukemia.

**Emerging therapeutic targets in AML**

Despite significant advances in targeted therapy development and a growing repertoire of drugs being tested in the treatment of acute myeloid leukemia (AML)(250), patient outcomes for AML have changed little in the last several decades. Only a small percentage of genetically defined AML patients exhibit durable long-term responses with current therapy. For instance, identification of the FLT3 internal tandem duplication mutation in 13-36% of AML (depending on the subgroup) (251) has led to the development of the FLT3 inhibitors quizartinib and midostaurin (252), the latter of which has recently acquired FDA approval. However, the 5-year overall survival rates of
the majority of AML cases ranges from 5-15% in older patients to 30% in young adults (18). This lack of improvement in patient survival rates over the past few decades is primarily attributed to the limited efficacy of currently available therapies in AML and the need for new targeted drugs. Although a number of promising drug candidates are being tested, such as the above mentioned FLT3 inhibitors, combination chemotherapy remains the standard of care (252). Thus, there persists a clear unmet need for new drugs for the treatment of AML.

Through the combination of chemical and RNAi screens, it has been suggested that GSK3α is a novel target in AML (282). In contrast to the more established role of GSK3α/β as a tumor suppressor pair, which inhibits Wnt signaling via β-catenin phosphorylation resulting in its degradation (253), it has been shown that GSK3α plays an important role in maintaining an undifferentiated leukemic state of AML blasts and therefore selective targeting of GSK3α, which avoids concomitant inhibition of GSK3β and subsequent β-catenin stabilization, could be a new therapeutic strategy for AML (244). Currently, the only FDA-approved GSK3 inhibitor is lithium chloride (LiCl), which is used for the treatment of epilepsy and bipolar disorder (254,255). However, given the narrow therapeutic index of LiCl, the lack of GSK3α specificity, and the limited kinome-wide selectivity (256,257), its utility as an AML therapy is questionable. There are a number of GSK3 inhibitors in development, but current compounds are either highly unselective featuring various off-targets in addition to GSK3α/β, lack isoform selectivity or have not yet advanced to clinical studies (258,259). In a previous study, we identified GSK3α/β as novel targets of tivantinib (ARQ197) in NSCLC (260), an advanced clinical drug candidate, which was initially thought to be a highly specific MET inhibitor (227). We observed that tivantinib, compared to other GSK3 inhibitors, has remarkable kinome-wide selectivity for GSK3α/β, as well as some specificity for GSK3α over GSK3β. Considering the identification of GSK3α as a potential oncogene, we hypothesized that tivantinib may be an effective, novel therapeutic option for AML. We therefore
aimed to next characterize tivantinib’s anticancer activity in AML cell lines and identify a potentially synergistic drug combination.

**Repurposing tivantinib in AML**

**GSK3α is an actionable target in AML**

GSK3α has been described to be a novel target in AML (244). Supporting this report, analysis of publically available expression levels of GSK3α and GSK3β using the BloodSpot database (which contains more than 2000 AML and normal samples assembled from six independent studies on AML) revealed that GSK3α is overexpressed across multiple AML subtypes as compared to normal hematopoietic lineages (183) (Figure. 20a). Interestingly, GSK3β expression in AML differs little from normal hematopoiesis (Figure. 20b). In order to evaluate AML sensitivity to GSK3α/β gene silencing, we analyzed the publically available shRNA screening data in Project DRIVE (184) which contains the cell viability following shRNA gene silencing of various genes across 384 cancer cell lines. Consistent with GSK3α being overexpressed in AML, we found that AML cell lines are significantly more sensitive to GSK3α silencing as compared to GSK3β gene silencing (Figure 20c). Furthermore, AML cell lines constitute the most sensitive population of hematopoietic cell lines to GSK3α gene silencing (Figure. 20d). Together this suggests that GSK3α is an actionable target in AML cell lines.
Tivantinib has anticancer activity in AML via GSK3 inhibition

Since we had previously identified GSK3α as a prominent tivantinib target (260), we wanted to determine tivantinib’s efficacy in AML. To the best of our knowledge, tivantinib has never been tested in AML cell lines, including the various large drug screening efforts such as the Cancer Therapeutic Response Portal v2 (CTRPv2) (90). Therefore to evaluate tivantinib’s efficacy across all the AML cell lines in the Cancer Cell Line Encyclopedia (CCLE) (191), we trained an elastic net regularized regression model to predict the area under the curve (AUC) sensitivity values of all the cell lines with tivantinib sensitivity data in CTRPv2 (which does not include tivantinib sensitivity information for AML cell lines) (Figure 21a) (90) using the gene expression profiles of

Figure 20. GSK3α is an actionable target in acute myeloid leukemia. (a–b) Hierarchical coexpression tree of (a) GSKα and (b) GSKβ gene expression across hematopoietic cell types. Size and color is associated with gene expression. Tree structure is based on genome wide gene expression profiles. (c) RSA sensitivity score of GSK3 shRNA sensitivity across AML cell lines. Wilcoxon-rank sum test was used to establish significance. (d) RSA sensitivity score of GSK3α shRNA sensitivity across hematopoietic cancer cell lines.
these cell lines (CCLE) (191) as features. Our model had good accordance between predicted and experimental values \( (r = 0.71, \text{ROC-AUC} = 0.83) \) which is comparable to similar models (261) (Figure 21b,c). Interestingly, many of the gene features selected through regularization are known to associate with GSK3 signaling (STRING) (186) such as the TCF7 cofactor MLLT11 which had the mostly highly weighted coefficient (262) (Figure 21d). We then applied this model to all AML cell lines in CCLE using their gene expression profiles, which predicted AML cell lines to be sensitive to tivantinib treatment. Interestingly, AML cell lines were predicted to be significantly more sensitive to tivantinib than non-AML cell lines (Figure 21e).

In order to validate the predicted AML sensitivity to tivantinib, we treated HL60, U937 and KG-1 AML cell lines with (-)-tivantinib, which is currently in advanced clinical development, its enantiomer (+)-tivantinib, which is a much weaker GSK3 inhibitor, the \textit{bona fide} pan-GSK3\(\alpha/\beta\) inhibitor LiCl and the c-MET inhibitor PF-04217903 as indicated. Intriguingly (-)-tivantinib, but not (+)-tivantinib (not tested in KG-1 cells), displayed nanomolar efficacy in HL60, U937 (Figure 21f, g), as well as KG-1 (Figure 21h) cell lines. This is in accordance with our previous results in MLL-rearranged acute lymphoblastic leukemia (ALL) cell lines, which are known to be sensitive to GSK3 inhibition(248,260). As expected, LiCl also showed strong activity while the potent and selective MET-inhibitor PF-04217903 was essentially inactive suggesting that GSK3, not MET, inhibition is responsible for tivantinib’s activity in AML cells (Figure 21f–h). Since tivantinib has previously been suggested to elicit anticancer activity in NSCLC through disruption of microtubule dynamics (232,233), we further evaluated the relative contribution that inhibition of MET, GSK3
Figure 21. Tivantinib has activity in acute myeloid leukemia cell lines. (a) Machine learning workflow to predict tivantinib efficacy in AML cell lines. (b) Correlation of predicted vs. actual area under the curve (AUC) values and (c) receiver operator characteristic (ROC) curve of prediction accuracy across all cell lines in the training and test sets. NRMSE = normalized root mean square error. (d) Lollipop plot of feature importance for top 15 genes selected during regularization. Inset is STRING network (medium confidence) of known associations of these
these genes with GSK3 signaling. (e) Empirical cumulative distribution function (ECDF) comparing the predicted AML AUC values to all the non-AML AUC values in CTRPv2. Statistical significance was determined using a Kolmogorov-Smirnov test. (f-h) Dose response curves and IC50 values for inhibition of viability by (-)-tivantinib, (+)-tivanitinib, LiCl, and PF-04217903 in (f) HL60 and (g) U937 cells. (h) Dose response curves and IC50 values for inhibition of viability by (-)-tivantinib and PF-04217903 in KG-1 cells. (i) Bar plot comparing spearman correlation of AML sensitivity predictions for (-)-tivantinib to paclitaxel, SGX253, and ML320 from elastic net regularized regression. Maximum possible spearman correlation is annotated. CCLE logo is provided courtesy of Ellen Gelfand. Copyright 2018 The Broad Institute of MIT & Harvard.

or microtubule polymerization plays in tivantinib’s mechanism of action in these cells. We trained additional elastic net regularized regression models to predict paclitaxel (microtubule inhibitor), SGX253 (MET inhibitor), and ML320 (highly selective GSK3 inhibitor) sensitivity across cell lines present in CTRPv2. We then applied these models to predict AML sensitivity, and performed pairwise comparisons (Spearman) of the model predictions. As expected, tivantinib’s sensitivity profile was uncorrelated with SGX253 further supporting that MET is not involved in tivantinib’s mechanism. Interestingly, tivantinib sensitivity was most highly correlated with ML320 suggesting the GSK3 inhibition is the primary mechanism in which tivantinib elicits activity in these cells whereas tivantinib was much more weakly correlated with paclitaxel (Figure 21i). In summary, this data demonstrates that tivantinib harbors potent anticancer activity in AML cell lines and this activity can likely be explained by GSK3 inhibition.
To confirm tivantinib’s ability to bind and inhibit GSK3α/β in these cells, we performed drug affinity chromatography using a couplable (−)-tivantinib analog as previously described (260). Pulldowns had high correlation between biological replicates (Figure 22a). We prioritized kinase targets with a SaintScore > 0.8, and an NSAFscore > 2x10^{-5} (Figure 22b). These criteria yielded two targets of tivantinib in these cells, GSK3α and GSK3β (Figure 22c). Tivantinib selectivity was confirmed by western blot, where (−)-c-tivantinib much more prominently enriched GSK3α/β as compared to (+)-c-tivantinib (Figure 22d). Importantly, the GSK3 inhibitor BIO was able to

Figure 22. Tivantinib target profiling in acute myeloid leukemia cell lines. (a) Correlation of total unique spectra between biological replicates of c(-)-tivantinib pulldowns in HL60 cells. (b) NSAF score vs SaintScore of c(-)-tivantinib pulldowns in HL60 cells. Red points indicate proteins included in analysis before kinase filtering. (c) Kinases enriched from drug affinity chromatography in HL60 cells passing SaintScore > 0.8 and NSAFscore > 2x10^{-5} cutoffs. Bubble size represents the number of total unique spectra. Bubble color represents probability of a specific interaction based on the CRAPome. (d) Western blot of eluates from drug affinity chromatography in HL60 and U937 cells. TCL = total cell lysate, BB = blocked beads, c(-)-Tiva / BIO = 5-BIO competition, Tiva = tivantinib.
compete away GSK3α/β suggesting a specific interaction. MET protein was unobserved by proteomics.

Treatment of HL60 cells with tivantinib decreased GSK3α/β phosphorylation on Tyr279/216, which as an autophosphorylation site is directly correlated with GSK3 kinase activity (Figure 23a) (263). Furthermore, upon treating with tivantinib or LiCl we observed an increase in total β-catenin levels, which is characteristic for GSK3 inhibitors (253). However, a larger and more prolonged increase in β-catenin was observed with LiCl than with tivantinib (Figure 23a). Since β-catenin stabilization requires inhibition of both GSK3α and GSK3β (264), this is consistent with LiCl strongly targeting both GSK3α/β isoforms and tivantinib being more selective for GSK3α, as we had shown previously by in vitro kinase assay (260).

To gain further insight into the downstream effects of GSK3 inhibition by tivantinib and LiCl, we investigated the cellular outcome following drug treatment via western blot and flow cytometry. Previous studies have suggested that tivantinib causes G2/M arrest through inhibition of microtubule polymerization, an observation which can likely also be explained by GSK3 inhibition (233). We therefore investigated the effects of tivantinib and LiCl on cell cycle arrest in AML cells. Tivantinib caused a pronounced and rapid increase in phosphorylation of histone H3 Ser10 (Figure 23b), which is indicative of cell cycle arrest. Detailed flow cytometry analysis showed a strong accumulation of cells in G2/M phase upon tivantinib treatment (Figure 23c). This was similarly prominent with LiCl suggesting the observed G2/M arrest is mediated through inhibition of GSK3α/β. Furthermore, after 24h of tivantinib treatment, we observed a strong and dose-dependent induction of apoptosis as assessed by PARP-1 and caspase 3 cleavage (Figure 23b). Consistent with previous reports (244), this was also apparent for LiCl although less pronounced than for tivantinib. We next assessed the timing and magnitude of the induction of apoptosis by Annexin V staining followed by flow cytometry. Similar increases in early apoptosis
were observed over time between tivantinib and LiCl; however, a much larger late apoptotic population was observed with tivantinib treatment (Figure 23d).

Given that GSK3\(\alpha\) silencing by RNAi has been described to induce cell differentiation (244), we stained HL60 cells with \(\alpha\)-CD11b to assess the ability of tivantinib and LiCl to differentiate AML cells by flow cytometry. Indeed, tivantinib treatment for 96 hours resulted in a significant increase of cell differentiation. Interestingly, LiCl caused a much stronger effect (Figure 23e), which is consistent with previous studies (244). Thus, tivantinib and LiCl have largely similar effects on AML cells as they both induce apoptosis, G2/M arrest, and differentiation; however tivantinib more potently induces apoptosis while LiCl has markedly larger effects on cell differentiation.
Figure 23. Analysis of cellular response following tivantinib treatment. (a) Effects of (-)-tivantinib (in µM), NaCl (20 mM) and the pan-GSK3 inhibitor LiCl (20 mM) on β-catenin and pGSK3α/β Y279/216 levels in HL60 cells. (b) Effects of tivantinib (in µM), NaCl (20 mM), and LiCl (20 mM) on PARP-1 and caspase 3 cleavage as well as pSer10 histone H3 levels after 4 and 24h. (c) Cell cycle analysis by DAPI DNA staining following treatment of HL60 cells with DMSO, NaCl (20 mM), LiCl (20 mM), or tivantinib (1 µM) for 24h. (d) Analysis of apoptotic populations by Annexin V staining following treatment of HL60 cells for 4, 12, 18, or 24h with DMSO, tivantinib (in µM), NaCl (20 mM), or LiCl (20 mM). (e) Cellular differentiation of HL60 cells following treatment with DMSO, tivantinib, NaCl or LiCl for 72 and 96h as assessed by CD11b staining. Asterisk denotes p<0.05 (*). Tiva = tivantinib.
Tivantinib synergizes with BCL-2 inhibition

Resistance against single drug therapy with targeted agents can often be delayed or suppressed by potent drug combinations. In the case of tivantinib, drug combinations may allow for a reduction of the tivantinib dose and thereby a less pronounced stabilization of β-catenin. In order to further amplify tivantinib’s anticancer activity in AML cells, we conducted a drug combination screen for inhibition of viability of HL60 cells using a collection of 240 clinically relevant targeted agents. The majority of these (90+%) are in clinical development so that identification of a drug that synergizes with tivantinib has the potential for clinical translation. The data is reproducible with good correlations between biological replicates (Figure 24a). One of the strongest hits from this screen for potential synergy with tivantinib in HL60 cells was the BCL-2 inhibitor navitoclax (ABT-263) (Figure 24a). This was interesting as cancer cell lines with activating mutations in β-catenin or increased β-catenin levels as the result of GSK3 inhibition have been shown to exhibit increased sensitivity to BCL-2 inhibitors (265). In addition to navitoclax, we identified its newer structural analogue ABT-199 as a potentially synergistic drug (Figure 24a)(266). We posited that since navitoclax has shown acute toxicity in patients, combination of tivantinib with ABT-199 may be a safer alternative while still retaining efficacy(267). Importantly, ABT-199 (FDA approved for CLL) has already been proven to be effective in AML cells (268,269) with a recent FDA designation as a breakthrough therapy and multiple AML specific clinical trials currently recruiting patients to test the safety and efficacy of ABT-199 alone or in combination with chemotherapy. We therefore selected ABT-199 for detailed synergy analysis. In addition to a clear shift of the dose response curve for the combination treatment, combination index (CI) values (181) along the curve suggested synergy (CI<0.7) to strong (CI<0.3) synergy within physiologically relevant concentration ranges (Figure 24b).
Figure 24. Identification of tivantinib and ABT-199 as a synergistic drug combination in AML cells. (a) Replicate correlations of cell viability following treatment with individual library compounds (2.5 μM) (left) and compounds in combination with tivantinib (0.25 μM) (middle) are displayed. Fold change corresponds to the ratio of inhibition of cell viability achieved by a drug combination with tivantinib (0.25 μM) compared to individual single library compounds (2.5 μM). Drugs passing fold change > 1.5 cutoff are highlighted in red. (b) Dose response curves of tivantinib, ABT-199, and combination ratio of 60:1 tivantinib to ABT-199 with individual CI values as determined by CompuSyn. Plasma concentration of tivantinib is annotated. (c) Effects of tivantinib and ABT-199 combination (in μM) on PARP-1 and caspase 3 cleavage as well as pSer10 histone H3 levels after 24h treatment. (d) Effects of tivantinib and ABT-199 combination on β-catenin stabilization and GSK3α/β pY279/216 levels. (e) Effects of tivantinib and ABT-199 combination on MCL-1, BCL-XL, and Bak. V = vehicle (DMSO). Tivantinib, ABT-199, and BIO concentrations are in μM. NaCl and LiCl concentrations are in mM.
As we had observed a strong induction of apoptosis with tivantinib treatment and as ABT-199 inhibits the anti-apoptotic protein BCL-2, we hypothesized that combining tivantinib and ABT-199 would also further increase apoptotic signaling. Upon combination treatment of tivantinib and ABT-199, we observed an increase in cleavage of caspase 3 and a complete cleavage of PARP-1 protein, in which no native 116 kDa PARP-1 remained, suggesting a large increase in apoptosis (Figure 24c). The pronounced G2/M arrest as indicated by pS10 histone H3 following tivantinib treatment was reversed by the drug combination. ABT-199 by itself did not affect GSK3 tyrosine phosphorylation or β-catenin levels, but the combination with tivantinib caused complete loss of GSK3 pY279/216 (Figure 24d). Importantly, addition of ABT-199 completely abrogated the increase of β-catenin that is observed with single agent tivantinib treatment (Figure 24d). Since MCL-1 and BCL-XL expression have been shown to cause resistance to ABT-199 (269,270), we hypothesized that the observed synergy with tivantinib is a result of altered MCL-1 and BCL-XL protein levels. Interestingly tivantinib single agent and, more pronouncedly, ABT-199 combination caused a loss of anti-apoptotic MCL-1 and BCL-XL protein levels while maintaining pro-apoptotic Bak levels (Figure 24e). In summary, these results suggest that the tivantinib and ABT-199 combination greatly increases the already strong apoptotic effects of tivantinib in AML cells by down regulating anti-apoptotic proteins while simultaneously suppressing activation of β-catenin.

In order to better evaluate the potential for clinical translation of our observations with tivantinib in AML, we next tested the efficacy of tivantinib in primary AML patient Bone Marrow Mononuclear Cells (BMNCs) as a single agent, as well as in combination with ABT-199. Using several different primary AML patient samples, tivantinib displayed a strong
ability to inhibit colony formation particularly at the clinically relevant concentration of 5 µM across all patients with only a few colonies remaining (Figure 25a). Single drug ABT-199 treatment showed slight variations in efficacy, but on average reduced colony formation to approximately 30-40 percent consistent with previous reports (Figure 25a) (269,270). The combination of tivantinib and ABT-199 exerted synergy in 4 of 7 patients with moderate (patients 1 & 2) to strong (patients 3 and 4) values (Figure 25a–b). Patients 5 and 7 were exquisitely sensitive to ABT-199 single agent treatment and therefore a Bliss value could not be accurately calculated (Figure 25a). Importantly, tivantinib efficacy did not show any obvious relationship with mutational status or karyotypes (Figure 25a; Table 1). Overall, these data suggest that tivantinib is highly effective in
inhibiting the colony forming capacity of primary AML patient samples as a single agent or in combination with ABT-199.

**Table 1.** Clinical characteristics of AML patients at time of diagnosis. WHO: World Health Organization; FAB: French-American British classification (M0-M7); WBC: white blood count (cells x 10^3 / dL); HgB at Dx: Hemaglobin at diagnosis (g/dL); Plt: platelet (cells x 10^3 / dL); BMBx: bone marrow biopsy.

<table>
<thead>
<tr>
<th></th>
<th>Patient 1</th>
<th>Patient 2</th>
<th>Patient 3</th>
<th>Patient 4</th>
<th>Patient 5</th>
<th>Patient 6</th>
<th>Patient 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>M</td>
<td>M</td>
<td>F</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Age</td>
<td>36</td>
<td>66</td>
<td>43</td>
<td>67</td>
<td>69</td>
<td>78</td>
<td>76</td>
</tr>
<tr>
<td>WHO</td>
<td>--</td>
<td>AML w/ recurrent abnorm.</td>
<td>AML NOS</td>
<td>AML MDS changes</td>
<td>AML w/ recurrent abnorm.</td>
<td>AML MDS changes</td>
<td>AML MDS changes</td>
</tr>
<tr>
<td>FAB</td>
<td>M5a</td>
<td>M4/5</td>
<td>M5</td>
<td>M4</td>
<td>M2</td>
<td>M6</td>
<td>M1</td>
</tr>
<tr>
<td>WBC</td>
<td>29.52</td>
<td>67.31</td>
<td>61.89</td>
<td>1.3</td>
<td>--</td>
<td>0.71</td>
<td>2.61</td>
</tr>
<tr>
<td>HgB at Dx</td>
<td>7.2</td>
<td>11.3</td>
<td>8.3</td>
<td>7.9</td>
<td>--</td>
<td>9.3</td>
<td>8.7</td>
</tr>
<tr>
<td>Plt</td>
<td>51</td>
<td>248</td>
<td>68</td>
<td>7</td>
<td>--</td>
<td>26</td>
<td>54</td>
</tr>
<tr>
<td>Blast % BMBx</td>
<td>100%</td>
<td>--</td>
<td>60-65%</td>
<td>39%</td>
<td>48%</td>
<td>15%</td>
<td>72%</td>
</tr>
<tr>
<td>Karyotype at Dx</td>
<td>Poor</td>
<td>Good</td>
<td>Good</td>
<td>Poor</td>
<td>Intermed.</td>
<td>Intermed.</td>
<td>Good</td>
</tr>
<tr>
<td>Complex?</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
</tbody>
</table>

**Discussion**

GSK3 plays a central role in a broad range of cellular processes, including glycogen metabolism, insulin signaling, apoptosis and microtubule function. Accordingly, it is under investigation as a potential target in Alzheimer’s disease and diabetes (253,271). In the context of cancer, GSK3 is best known in its function as a tumor suppressor, which is deactivated by AKT or Wnt signaling (253,271). However, most studies have focused on GSK3β, whereas significantly less is known about GSK3α. Moreover, it is increasingly appreciated that GSK3 signaling is context-dependent.
For instance, there have been several reports that describe tumor supporting roles of GSK3α in glioblastoma, pancreatic cancer, multiple myeloma, and MLL-rearranged leukemia (248,249,272,273).

GSK3α has been identified by Banerji et al. through a functional genomic screen as a promising target in AML where it maintains and promotes an undifferentiated leukemic state (244). However, targeting GSK3 in hematological malignancies does have theoretical challenges in that GSK3 is known to phosphorylate β-catenin thereby marking it for subsequent proteasomal degradation. Upon GSK3 inhibition, β-catenin accumulates, translocates to the nucleus and activates transcriptional pathways (253,274). Such increased β-catenin signaling has been implicated in a number of leukemogenic effects, such as self-renewal of leukemic stem cells, thus serving itself as an attractive target in AML (275). As β-catenin stabilization requires inhibition of both kinases (264) and most GSK3 inhibitors target GSK3α and GSK3β with similar potency, these compounds may possess some significant limitations. Accordingly, the nonspecific pan-GSK3 inhibitor LiCl, which is currently FDA approved for the treatment of epilepsy and bipolar disorder (254,255), has met limited success in clinical studies of AML (276–278). In addition to the relative selectivity for GSK3α over GSK3β, various other aspects, such as the binding mode, magnitude of inhibition, and kinome-wide target specificity influence the overall cellular outcome of inhibiting GSK3α. For this reason, not all GSK3 inhibitors should be treated equally.

Tivantinib was originally developed as an inhibitor of the receptor tyrosine kinase MET (227), but its target selectivity and mechanism of action has been a controversial subject. We and other groups have suggested that MET is actually not a significant target of tivantinib in many cancer cells (233,260,279). We have previously observed that tivantinib, although being indeed a weak MET inhibitor, much more prominently targets GSK3α and GSK3β in NSCLC cells and that inhibition of these targets can explain its potent anticancer activity in NSCLC (260). Importantly, we also noted remarkable kinome-wide specificity and some moderate selectivity of
tivantinib for GSK3α over GSK3β, which are unique features among clinical GSK3 inhibitors (259). It has also been suggested that tivantinib binds tubulin and inhibits microtubule dynamics resulting in anticancer activity (233, 279). While these observations are compelling, our results suggest that the sensitivity profile of tivantinib in AML cells more closely matches that of a GSK3 inhibitor. Furthermore, we show that pharmacological inhibition of GSK3 by LiCl largely mimics tivantinib’s effects in these cells with regard to viability, apoptosis, cell cycle arrest and differentiation. However, as GSK3α/β have well described roles in microtubule regulation through phosphorylation of Tau (MAPT) (218, 280) and other microtubule associated proteins, such as MAP2C (281), it would be difficult to precisely identify the contributions to tivantinib’s overall cellular effects that stem from targeting tubulin in addition to GSK3. Impairment of microtubule polymerization however may be translationally beneficial as it may result in synergistic anticancer activity in the context of dual GSK3 and BCL-2 inhibition as has been previously observed in breast cancer (282). However, additional studies are necessary to elucidate the complex interplay between these pathways in AML.

In light of the suggested role of GSK3α in AML, we investigated the potential for repurposing tivantinib for the treatment of AML, which to the best of our knowledge has not been reported. Consistent with previous reports (244), analysis of publically available datasets showed that GSK3α is overexpressed in AML and that knockdown of GSK3α has strong effects on the viability of AML cell lines. We also show that tivantinib interacts with and inhibits GSK3α/β in AML cells and that it potently kills these cells by inducing apoptosis. Interaction of tivantinib with its intended target c-MET, which was observed to a minor extent in NSCLC cells (260), was not detectable in AML cells. While tivantinib does target both GSK3α and GSK3β, its effects on β-catenin levels were somewhat less pronounced than with LiCl. This was consistent with our previous observation that tivantinib is more selectively targeting GSK3α (260). Also, β-catenin stabilization was more transient with tivantinib, whereas it is sustained for a longer period of time.
upon LiCl treatment. Tivantinib may therefore provide an important therapeutic advantage over pan-GSK3 inhibitors, such as LiCl. Banerji et al. have shown that LiCl readily causes differentiation of AML blasts at relatively low concentrations(244). Our results confirmed these observations and showed that tivantinib also induces differentiation. However, tivantinib was a much stronger inducer of apoptosis than of differentiation; and although LiCl also induces apoptosis, tivantinib is markedly more potent than LiCl in this regard, which might be due to additional effects of tivantinib on microtubules.

In addition to tivantinib exhibiting potent single agent activity in AML, we observed that the BCL-2 inhibitor ABT-199, which displays activity and is in clinical trials in AML(269), synergizes with tivantinib by further enhancing tivantinib’s already potent ability to inhibit cell viability and induce apoptosis. MCL-1 and BCL-XL expression have previously been associated with ABT-199 resistance (269,270), and it is noteworthy that we observed a dose-dependent decrease in MCL-1 and BCL-XL levels following tivantinib treatment, which was enhanced in combination with ABT-199. This downregulation is consistent with previous reports that have shown that GSK3 transcriptionally regulates BCL-XL expression and that GSK3 inhibition by BIO or SB-415286 leads to a reduction in BCL-XL levels (283,284). Interestingly, while BIO similarly downregulated BCL-XL in breast cancer cells, it also reduced MCL-1 expression, not via transcriptional control, but through a proteasome-dependent mechanism(283). This downregulation of MCL-1 and BCL-XL expression likely contributes to the synergy observed between tivantinib and ABT-199. By downregulating anti-apoptotic MCL-1 and BCL-XL that cause ABT-199 resistance, tivantinib is amplifying the relative apoptotic effect of ABT-199. This synergy is in excellent agreement with a previous study that described cancer cells with increased β-catenin levels, for instance as the consequence of GSK3 inhibition, to be particularly sensitive to inhibition of BCL-2 by the ABT-199 analogue navitoclax(265). Interestingly, in addition to modulation of anti-apoptotic proteins we observed that the tivantinib/ABT-199 combination completely abrogated β-catenin stabilization.
seen with tivantinib single drug treatment. This was apparent already after 4 hours and is therefore likely due to cross-talk between the GSK3 and BCL-2 pathways that is independent of altered transcription. This pronounced reduction of β-catenin persisted for 24 hours and could possibly help prevent some of the leukemogenic effects previously associated with β-catenin signaling in AML (275). Thus by modulating tivantinib's effects on β-catenin levels in conjunction with the amplification of apoptotic signaling, this suggests a superior therapeutic potential of this drug combination in AML.

In this context, it is important to note that tivantinib, as a single drug and even stronger in combination with ABT-199, showed potent anti-leukemic activity in AML patient-derived samples within clinically relevant concentrations. This appears to be independent of the mutational status of common prognostic genes although our sample size was too small to allow broader conclusions. A potential correlation with tivantinib sensitivity could be amplification of GSK3α expression levels as GSK3α is more highly expressed in several different subtypes of AML, including 11q23 MLL-rearranged leukemia, which has previously been shown to be sensitive to GSK3 inhibition (248,260). Interestingly, high expression of GSK3 and BCL-XL has previously been suggested to correlate with poor prognosis in AML (285). However, a more thorough investigation of GSK3α expression and signaling in AML is necessary to make detailed conclusions. Considering the safety profile of tivantinib and that the concentrations required for its activity in AML cells are well within the therapeutically achievable levels (235,236), repurposing tivantinib provides a tangible opportunity for clinical translation into AML.

In summary, repurposing the advanced clinical drug candidate tivantinib based on its off-target GSK3α identified it as a highly potent agent in AML cells. Combination with the BCL-2 inhibitor ABT-199, which is already under clinical investigation for AML, further enhanced tivantinib's potency and eliminated undesirable β-catenin activation. Together, these findings suggest that tivantinib, either as a single agent or in combination with ABT-199, represents a
novel and promising therapeutic option for AML, a disease, which is still in high need for new therapies.
Chapter six: Systems polypharmacology repurposing of ceritinib

(Note to reader: Parts of this section have been published previously in Kuenzi et al Nature Chemical Biology doi: 10.1038/nchembio.2489 (89) and are being reproduced with permission from Nature Publishing Group copyright 2017 (Appendix B).

p90 Ribosomal S6 Kinase (RSK)

Structure, conservation and activation

The p90 kDa ribosomal S6 kinase (RSK) was originally discovered as part of an effort to identify the kinase responsible for RPS6 phosphorylation in vitro. Two responsible kinases were biochemically purified which were named p90 RSKs. Interestingly, it was later found that two related kinases, p70 kDa ribosomal S6 kinase 1/2 (S6K1 and S6K2), were the predominant kinases responsible for RPS6 phosphorylation (286,287). Two additional RSK family members have been subsequently identified. Since most studies to date have not determined isoform specificity for RSK substrates, isoform specific functions of RSK remain largely unknown (288).

The human RSK family contains four different isoforms (each encoded by separate genes), namely RSK1 (RPS6KA1), RSK2 (RPS6KA3), RSK3 (RPS6KA2) and RSK4 (RPS6KA6) which are ~80% identical in sequence. The divergence in sequence resides largely in their N– and C– terminal sequences whereas the sequences surrounding critical residues, such as phosphorylation sites corresponding to RSK activation, are largely conserved among family members (Figure 26a). The most distinguishing feature of the RSK family is the presence of two non-identical protein kinase domains within the same protein, an N–terminal kinase domain (NTKD) and a C–terminal kinase domain (CTKD) (289,290). Interestingly, these domains belong
to different kinase families with the NTKD being a member of the AGC kinase family (also containing AKT and S6K) and the CTKD being a member of the CAMK kinase family (which also includes AMPK and DAPK) (UniProt). This difference between kinase domains corresponds to differential roles of the NTKD and CTKD. The only known function of the CTKD is to facilitate the activation of the NTKD of RSK which then can then phosphorylate downstream substrates (290–292). Differences among NTKD and CTKD also have allowed for identification of domain specific inhibitors of both the NTKD (BI-D1870 & SL-0101) and CTKD (FMK) which are widely used to study RSK function (85,293,294).

Full activation of RSK is a multistep process involving both RSK kinase domains as well as additional kinase binding partners (Figure 26b). All RSK isoforms contain a kinase interacting motif (KIM) (L-X-K/R-K/R-X-L) which facilitates mitogen-activated protein kinase 3/1 (ERK1/2) binding to RSK’s C–terminus. This domain is notably different from ERK’s typical consensus sequence; however, binding increases affinity of ERK for its substrates (295–297). Upon binding to the KIM, ERK is activated allowing it to phosphorylate the CTKD within the activation loop at S573. This phosphorylation is thought to lead to plasma membrane translocation of RSK and full activation of the CTKD (298–300). Once activated, the CTKD can autophosphorylate the hydrophobic motif of the interlinker region of RSK at S380 (292). This phosphorylation then serves as a docking site for 3-phosphoinositide-dependent protein kinase 1 (PDK1) which promotes PDK1 activity thus stimulating phosphorylation of the NTKD activation loop at S221 by PDK1 (298,301,302). Once phosphorylated, the NTKD becomes fully activated allowing for RSK autophosphorylation at S737 which facilitates ERK1/2 dissociation from the RSK C–terminus allowing for phosphorylation of downstream substrates by the NTKD (288,295). A number of downstream substrates for RSK have been identified including GSK3β (303,304), eIF4B (305), TSC2 (306,307), YB1 (308), and RPS6 (309). In general, RSK substrates follow a R/K-X-X-R-S/T or R-R-X-S/T motif with a preference of phosphorylation for serine residues over threonine (310).
Biological Processes and Biochemical Functions

Since the RSK family kinases play a role in many different biological processes and biochemical functions, we will focus on a few of the most cancer-relevant functions of RSK in this chapter.

**Protein synthesis.** Unsurprisingly, the first piece of evidence that RSK is involved in the regulation of protein synthesis stems from its discovery as an RPS6 kinase (311,312). RPS6 has well-described roles in protein translation and is a critical component of the ribosomal complex (313). It was later shown that RSK is able to associate with actively translating ribosomes and that phosphorylation of RPS6 at S235/236 by RSK promotes assembly of the cap binding complex. This phosphorylation is largely dependent on active RAS/MAPK signaling and correlates with increased translation (309,314,315). Interestingly, RSK has also been shown to affect protein synthesis earlier in the pathway by promoting active mTOR signaling (mTORC1), which is a well

---

**Figure 26. RSK isoform sequence conservation and mechanism of activation.** (a) conservation of amino acid sequence surrounding mechanistically relevant phosphorylation sites. Phosphorylation sites are bold/underlined. Non-conserved residues are highlighted in red. Amino acid position annotation is based on RSK1 sequence. (b) Mechanism of activation of RSK family members. 1. ERK binds to to KIM motif on RSK. 2. ERK phosphorylates the CTKD of RSK leading to its activation. 3. The CTKD phosphorylates S380 within the hydrophobic motif in the interlinker region. Optionally, ERK phosphorylates additional residues in RSK's interlinker region. 4. PDK1 binds to the phosphorylated hydrophobic motif. 5. PDK1 phosphorylates S221 of the NTKD leading to its activation. 6. The NTKD phosphorylates S737 in the KIM motif leading to ERK dissociation and full RSK activation. RSK can then phosphorylate substrates with its NTKD. Adapted from Romeo et al (2012) *Biochem J.*
known regulator of ribosome activity and mRNA translation (316). RSK does this through an inhibitory phosphorylation on tuberous sclerosis complex 2 (TSC2) S1798 (in conjunction with phosphorylation at S939 by AKT and S540/664 by ERK), disrupting its association with TSC1. The TSC1/2 complex normally inhibits the activator of mTORC1, Rheb, and therefore disruption of this complex allows for full mTOR activation by locking Rheb in its active GTP bound state (306,317–319). Active mTOR is then able to regulate a number of members of the translational machinery including p70 ribosomal protein S6 kinases, elongation initiation factors and 4E-BPs, which drives mRNA translation (316,320) (Figure 27a).

**Cell cycle progression and proliferation.** RSK has been directly implicated in regulating cell cycle progression and cancer proliferation through various mechanisms (Figure 27b). One of

![Figure 27. Biological functions of RSK.](image-url)
the first pieces of evidence that RSK was involved in cell cycle regulation was the observation that treatment of cancer cells with the RSK inhibitor SL-0101 inhibited proliferation. This was further supported by RNAi-based experiments knocking down RSK1 and RSK2 yielding similar results (293,321). Later mechanistic studies began to unravel the mechanisms in which RSK is able to regulate proliferation and the cell cycle at various stages. RSK has been shown to control G1 cell cycle by directly phosphorylating the serum response factor (SRF) transcription factor, activating it and thus promoting FOS transcription. RSK is also able to directly phosphorylate FOS at S362 to promote its activity (322–326). FOS then activates cyclin D to promote cell cycle entry (327). RSK is also able to control progression through G1 phase of the cell cycle through regulation of p27^{kip1}. RSK has been shown to directly phosphorylate p27^{kip1} at T198 which promotes its association with 14-3-3. Once bound to 14-3-3, p27^{kip1} is unable to bind and inhibit CDK2 and thus promoting cyclin E dependent G1 phase progression (328–331). In addition to G1 phase, RSK has been shown to have a distinct role in the regulation of the G2/M checkpoint of the cell cycle. M phase entry is in part controlled by CDC2 which is regulated by inhibitory phosphorylations on T14 and Y15 by WEE1 and MYT1. RSK has been shown to directly phosphorylate MYT1, inhibiting its activity and thereby promoting M phase entry through CDC2 and Cyclin B (332–338).

**Cell survival.** The Ras-MAPK-RSK axis has well described roles in regulating cell survival both by regulating pathways that elicit transcriptional control over pro-survival proteins and through direct phosphorylation of pro-apoptotic proteins (Figure 27c). RSK has been shown to promote CREB dependent transcription of the pro-survival genes Bcl-2, Bcl-XL and MCL-1. It does this by directly phosphorylating CREB at S133 thereby promoting its transcriptional activity (339,340). These pro-survival proteins regulate survival by controlling mitochondrial membrane permeability, preventing the release of cytochrome C and therefore apoptosis (341–343). RSK has also been shown to regulate cell survival through activation of the NF-κB pathway. RSK is able to regulate NF-κB activity by directly phosphorylating IκBα. IκBα is typically in complex with
NF-κB, preventing its dimerization and activation. Once phosphorylated by RSK, IκBα is degraded allowing NF-κB to dimerize similarly leading to transcription of pro-survival genes (326,344–346). Lastly, RSK has been shown to phosphorylate the pro-apoptotic protein Bad at S112 promoting its ubiquitination and degradation (339,347–349). A critical balance between pro-survival (e.g. Bcl-2, Bcl-XL, MCL1) and pro-apoptotic (e.g. Bad) proteins is maintained in which pro-survival proteins antagonize pro-apoptotic proteins preventing mitochondrial pore formation, release of cytochrome C and apoptotic caspase cascades (350–352). By increasing pools of pro-survival genes, RSK shifts this balance in favor of cell survival which helps to maintain tumorigenesis. In summary, RSK is able to promote cell survival through a number of mechanisms making it a critical regulator of this pathway.

**Focal Adhesion Kinase 1 (FAK1)**

**Structure and biological functions of FAK1**

The 125 kDa protein focal adhesion kinase 1 (FAK1) is a ubiquitously expressed protein tyrosine kinase involved in a large range of cellular functions. FAK1 contains an N-terminal FERM domain, a central protein kinase domain, three proline-rich regions (PRR) as well as a cleavable C-terminal focal-adhesion targeting (FAT) domain (Figure 28a). These domains contribute to FAK1’s kinase independent role as a scaffolding protein as well as serving to localize FAK1 for kinase dependent signaling. FAK1 has a number of key phosphorylation sites that serve to either activate the kinase or create docking sites for any of FAK1’s many binding partners. Autophosphorylation at Y397 for instance creates a Src-homology 2 (SH2) binding site, allowing for binding by SH2-containing proteins such as p120-RasGAP, phosphatidylinositol 3-kinase (PI3K), SHC and arguably most importantly SRC. While it is not known whether certain proteins differentially bind to pFAK1 Y397 in response to diverse stimuli or activation methods, it is likely that various proteins are recruited to FAK1 in order to elicit any of FAK1’s many downstream signals. In addition to
Y397, phosphorylation of Y576 / Y577 within the kinase domain by SRC is required for full activation of FAK1 kinase activity (353). This is supported by studies that introduced mutations (K \rightarrow E) within the activation loop of FAK1 at residues 578 and 581 which yielded a constitutively activated form of FAK1 suggesting that FAK1 resides in an autoinhibited state until activated by phosphorylation of Y576/577 (354). It was later shown that the FERM domain of FAK1 binds directly to the kinase domain thereby blocking access to the catalytic cleft (355).

FAK1 is able to be activated by a large number of external signals including cytokine receptors, G-protein coupled receptors (GPCR), integrin, and receptor tyrosine kinases (RTK) (Figure 28b) (356,357). The most well described method of FAK1 activation involves binding to integrin receptor clusters that are activated in response to extracellular matrix (ECM) binding. Upon binding, FAK1 will dimerize and catalyze its autophosphorylation at Y397 thus promoting SRC recruitment and full FAK1 activation (358). FAK1 is also able to be activated by cytokine receptors, G-protein coupled receptors, increases in intracellular pH, as well as activated receptor tyrosine kinases (359–365). Once activated FAK1 has roles in various cancer related cellular processes including cell cycle progression and cell survival (Figure 28c)(366,367), epithelial-to-mesenchymal transition (EMT) (368–370), cell motility and invasion (Figure 28d)(357,371). In this chapter I will briefly highlight FAK1’s role in cell cycle progression, cell survival, and cell motility / invasion.
Figure 28. Structure and function of FAK1. (a) Structure and key phosphorylation sites on FAK1. (b) Methods of activation of FAK1 leading to the role of FAK1 in (c) cell cycle progression and cell survival (d) epithelial to mesenchymal transition (EMT) and (e) cytoskeletal remodeling and invasion. Conceptually adapted from Sulzmaier et al. (2014) Nat. Rev. Cancer.
**Cell cycle progression and cell survival.** FAK1’s role in tumor growth and survival is becoming increasingly recognized. The best described role of FAK1 in survival involves its ability to promote PI3K-AKT signaling. Extracellular signals such as those received from integrins stimulate the FAK1-PI3K-AKT cascade to prevent anoikis, apoptosis as well as anchorage independent growth (372–374). This cascade promotes activity of transcription factors such as NF-κB and the oncoprotein YB1. FAK1-YB1 signaling for instance has been associated with resistance to taxane therapy in ovarian cancer through an AKT-dependent mechanism (375). Similar to the mechanisms described in the above “RSK: Biological Processes and Biochemical Functions” section, FAK1 can also promote cell survival through NF-κB activation and transcription of pro-survival genes such as BCL2 and BCL-XL (376,377).

**Cell motility and invasion.** FAK1 is canonically implicated in the formation and turnover of focal adhesions, which are critical elements of the cell motility machinery (378–381). Once recruited to a focal adhesion, FAK1 is activated in a RHOGEF dependent manner via p190RhoGEF (380,381). This complex associates with the adaptor protein paxillin (involved in focal adhesion maturation) stimulating its tyrosine phosphorylation (382). Active FAK1 is also able to recruit additional adaptor proteins, talin and cortactin, which help to localize FAK1 to focal adhesions and to link it to cytoskeletal components to promote mobility (383–385). Once phosphorylated by FAK1, cortactin promotes focal adhesion turnover through dynamic rearrangements in the actin cytoskeleton (357,385). Additionally, RHOGEF becomes activated once it is localized to the plasma membrane, and is therefore able to further promote actin polymerization through activation of RHOA (386–389). Interestingly FAK1 has additional, kinase-independent roles in regulating actin dynamics. Through binding to ARP2/3 by its FERM domain, FAK1 is able to enhance actin polymerization, directing polymerization toward cell protrusions thereby increasing motility (390–392). Lastly, FAK1 is able to further promote cell motility and invasion through the expression and activation of matrix metalloproteases (MMP) to facilitate
degradation of ECM components. FAK1 has been shown to increase the expression and surface presentation of MMPs through a number of downstream pathways including PI3K-AKT-MTOR signaling, SRC-p130CAS (393–396) as well as by promoting EMT through endophilin A2 binding and phosphorylation (369).

**Insulin-like growth factor 1 receptor (IGF1R)**

**Structure and biological functions**

The 95 kDa protein insulin-like growth factor 1 receptor (IGF1R) is a highly conserved receptor tyrosine kinase (RTK) in the insulin receptor family. The insulin receptor family also contains a highly homologous RTK, insulin receptor (IR), which shares 84 percent of its amino acid sequence with IGF1R. Both IGF1R and IR are dimeric receptors containing two extracellular α-subunits, two β-subunits as well as a small intracellular domain (397). The ligand-receptor interaction on the α-subunit results in phosphorylation of the tyrosine kinase domain of the β-subunit at Y1131, Y1135 and Y1136. Phosphorylation of Y1135 and Y1131 release the autoinhibitory loop which blocks the ATP binding pocket and thus tyrosine kinase activity whereas Y1136 phosphorylation locks IGF1R in its active conformation (Figure 29a) (398). Active IGF1R autophosphorylates additional tyrosine residues to form docking sites for SH2-domain containing proteins such as insulin receptor substrate 1 – 4 (IRS 1 – 4), SHC, or 14-3-3, which get phosphorylated by IGF1R upon binding (399). Recruitment and phosphorylation of these adaptor proteins allows for complex formation and downstream signaling through PI3K or RAS signaling pathways. As a result, IGF1R has been implicated in numerous cancers. For example, > 50% of breast tumors harbor activated IGF1R and 43% of stomach cancers display a molecular alteration (amplification, deletion, or mutation) in the IGF pathway (400). Therefore understanding IGF1R’s role in cancer development is of critical importance. In this chapter we will review IGF1R’s role in cancer cell growth / survival and invasion / metastasis.
Figure 29. Structure and function of IGF1R. (a) domain structure and method of activation of IGF1R. (b – c) role of IGF1R signaling in (b) cell proliferation, cell survival, and (c) cytoskeletal remodeling and invasion.
Cell growth and survival. IGF1R has well described roles in the growth and survival of cancer cells. It has been shown that IGF1R activation can protect cells from apoptosis induced by various agents including hypoxia, anticancer drugs, and osmotic stress and has proven itself as a critical determinant of apoptosis resistance in cancer cells (401–403). This apoptosis resistance is driven through the same pathways that IGF1R is able to promote cell proliferation, i.e. the PI3K and MAPK pathways. Upon activation, IRS1 is recruited to IGF1R thus promoting PI3K binding, an interaction which causes an increase in phosphoinositide 3,4,5-triphosphate (PIP₃) levels. PIP3 acts as a binding site for membrane localization and activation of PDK1 and AKT. AKT activation then leads to the expression of anti-apoptotic proteins Bcl-2, Bcl-XL, NF-κB and hypoxia inducible factors HIF-1α / HIF-2α (Figure 29b) (404–406). These pro-survival proteins regulate survival by controlling mitochondrial membrane permeability, preventing the release of cytochrome C from the mitochondria and therefore blocking the induction of apoptosis (341–343). In addition, the IGF1R – PI3K – AKT – MTOR signaling cascade leads to concomitant activation of S6K which has a dual role in cell survival. First S6K promotes cell growth by regulating protein synthesis components by phosphorylating RPS6 (as described in the p90 Ribosomal S6 Kinase : Biological Processes and Biochemical Functions S6K has been shown to phosphorylate the pro-apoptotic protein, BAD, at S136 thus inactivating it (407). Therefore, the IGF1R signaling pathway is able to promote cell proliferation and survival through a number of different critical signaling pathways.

Invasion and metastasis. In addition to its roles in proliferation and cell survival, IGF1R signaling has been implicated in driving the invasion and metastasis of cancer cells. Specifically, activated IRS-1 can alter the binding between E-cadherin and β-catenin (Figure 29c). Disruption of this binding leads to a dissociation of E-cadherin from actin filaments / cytoskeleton (connected through α-catenin) and a loss of E-cadherin expression. Loss of E-cadherin expression or function is a well understood mechanism to disrupt cell-cell junctions, thereby allowing tumor cells to
invade and metastasize (408–410). This is further supported by reports that IGF1R overexpression leads to tumor metastasis or that inhibition of IGF1R can inhibit tumor metastasis (411–413). It has also recently been shown that IGF1R regulates the expression of MMP2, disruption of which results in a less invasive cancer phenotype (414,415).

Targeted therapies in NSCLC

Targeted drugs have shown significant clinical success when directed against strong disease drivers, such as BCR-ABL in chronic myeloid leukemia (CML), EML4-ALK in non-small cell lung cancer (NSCLC) or JAK3 in rheumatoid arthritis, which are inhibited for instance by imatinib, crizotinib and tofacitinib, respectively (50,51,416). However, this target-driven approach is rarely effective in diseases with complex pathologies lacking dominant drivers or where these drivers are currently not druggable, such as most KRAS mutations. This is true particularly for many cancers as they often display broad landscapes of genetic mutations and epigenetic aberrations that lead to poorly understood deregulation of signaling and gene regulatory networks. Although inhibiting a single target may affect several cellular processes, efficient shutdown of deregulated oncogenic signaling networks is often only achieved by directly engaging multiple disease-associated proteins, which increases the likelihood of effective treatment by reducing signaling crosstalk or compensatory adaptation (35). Oncogene-negative cancers may therefore be more effectively targeted using a network-based approach that takes into account the effects of drugs on multiple pathways and cellular processes (23,66,67). This can be achieved by combining two or more targeted agents, a concept that is actively pursued in many preclinical and clinical studies and has been successfully implemented in melanoma using BRAF and MEK inhibitors, which block compensatory feedback activation (68). However, the clinical translation of drug combinations is often limited by issues with formulation and dosing, drug-drug interactions, increased toxicity, regulation or strategic business management, particularly when the individual drugs are developed by different companies (67,69). Alternatively, similar network-wide signaling
effects can be produced with single compounds that innately inhibit multiple disease-relevant targets, a phenomenon referred to as “polypharmacology” (67). The rational design of multi-targeted drugs for specific unrelated proteins poses a formidable challenge for drug discovery, though, as simultaneous potency optimization for two targets is difficult to accomplish and the risk for undesired inhibition of targets that elicit toxicity increases significantly (69). Thus, only few examples have been reported where this has been achieved in a rational manner, for instance in CML with dasatinib (BCR-ABL and SRC family kinases) and thyroid cancer (RET and VEGFR2) (54,74,417).

Targeted drugs, and kinase inhibitors in particular, have been shown to serendipitously display widely varying target profiles beyond their intended or “cognate” targets (69,111–114). While these “off-“ or “non-cannotical” targets are often either unknown or disregarded, they confer an inherent potential for polypharmacology applications. Notably, phenotypic screening approaches have found some kinase-targeted drugs to show antitumor activity in various subsets of cancer, which is unrelated to inhibition of their cognate targets and for which the underlying mechanism of action (MoA) therefore is not apparent, but is likely involving one or more non-canonical targets (88). Generating a detailed, systems-wide understanding of these beneficial off-target - and potentially polypharmacology - mechanisms can lead to novel drug repurposing opportunities that allow for the treatment of refractory cancers. Importantly, elucidation of these mechanisms can also be highly informative for understanding the wiring maps of complex oncogenic signaling networks, thus revealing new biological vulnerabilities as well as therapeutic opportunities with other drugs. Using a multi-tiered systems chemical biology approach, which integrates phenotypic screening with functional proteomics, we here describe the identification and mechanistic characterization of the polypharmacology activity of the FDA-approved, second-generation ALK inhibitor ceritinib in ALK-negative NSCLC. Furthermore, deep interrogation of ceritinib’s effector network enabled the subsequent development of a synergistic drug combination and identification of a predictive, mechanism-based biomarker candidate.
**Repurposing ceritinib in ALK-negative NSCLC**

A pharmacologic screen to identify beneficial off-target activity

We assembled a library of 240 compounds, which are mostly in clinical development or FDA-approved and span multiple target classes. In order to identify compounds with cellular activity unrelated to their cognate targets, this library contained multiple compounds per target that could serve as mutual controls. This library was screened for inhibition of cell viability against 20 NSCLC cell lines with various driver mutations. Unsupervised hierarchical clustering revealed that drugs of the same class largely cluster together (e.g. MAPK pathway inhibitors) (Figure 30a). Interestingly, the FDA-approved, second-generation ALK/IGF1R inhibitor ceritinib did not cluster with other ALK inhibitors (Figure 30a) suggesting that ceritinib may have a different anticancer activity profile. The screening data subset containing the ALK, IGF1R and EGFR inhibitors (to control for anti-EGFR activity of AP26113) was further analyzed using unsupervised hierarchical clustering to compare these inhibitors directly. As expected, all ALK inhibitors had potent activity against the ALK-rearranged H3122 cells. However, ceritinib displayed additional activity against
Figure 30. Ceritinib has beneficial off-target activity in ALK-negative NSCLC cells. (a) Dendrogram from unsupervised hierarchical clustering of the phenotypic drug screen of 240 compounds in 20 NSCLC cell lines. Cells were treated for 72 h at 0.5 µM and 2.5 µM of each compound in biological duplicate and viability was determined using CellTiterGlo. Colors highlight individual clusters. Left box (green) highlights ALK inhibitor cluster, right box (blue) highlights MAPK pathway inhibitor cluster. (b) Subset of drug screening data containing ALK, IGF1R and EGFR inhibitors at 2.5 µM. Points reflect the mean of two biological replicates. Individual clusters were manually chosen and numbered. Right box represents accompanying mutational data for these cell lines. The top box shows the in vitro IC₅₀ values for these drugs against ALK, EGFR and IGF1R. (c) Correlation of relative cell viability values for all cell lines for ceritinib and GSK1838705A. Cell lines highlighted in red represent cell lines displaying off-target activity with < 60% viability with ceritinib treatment and > 60% viability with GSK1838705A treatment. (d) Western blot of ALK across 13 cell lines (n = 2).
a number of mostly KRAS-mutant NSCLC lines whereas other ALK and IGF1R inhibitors had little to no effect in these cells (Figure 30b). To prioritize cell lines with the largest observed sensitivity unrelated to ALK inhibition, we compared ceritinib to GSK1838705A, which has similar potency against both ALK and IGF1R. This revealed specific ALK-independent activity of ceritinib in various cell lines, such as the KRAS-mutant H650, H1155 and H23 cells (Figure 30c). This was corroborated by lack of ALK expression in lung cancer cells in the absence of EML4-ALK chromosomal translocation, which in this cell line panel is only present in H3122 cells (Figure 30d). Since GSK1838705A is more potent for IGF1R than ceritinib, ceritinib's additional activity was also likely independent of IGF1R inhibition alone, which was further supported by the lack of cell sensitivity to other IGF1R inhibitors (Figure 30b).

**Cellular characterization of ceritinib activity**

To confirm the observed ceritinib activity we performed detailed dose-response curve analysis across cell lines that were sensitive to ceritinib but resistant to GSK1838705A. Ceritinib inhibited the viability of the most sensitive cells with an IC_{50} between 1-2 µM (E_{max} = 0% at 4 µM) (Figure 31a), which is within clinically achievable plasma concentrations ([ceri] = 1.4 - 2.3 µM).(420,421) This activity appeared to be elicited mostly through induction of G1 cell cycle arrest (Figure 31b-c). Ceritinib also showed substantial efficacy in longer-term clonogenic assays, in which it was able to completely eliminate H23 cells after 10 days (E_{max} = 0% at 2 µM) (Figure 31d). Together, these results suggest that ceritinib has ALK-independent antiproliferative activity in various NSCLC cell lines at concentrations likely to be clinically relevant.
Figure 31. Cellular effects of ceritinib in NSCLC cell lines. (a) Cell viability curve of ceritinib in H650, H23, H1155 and A549 cells. Concentrations are shown in µM (n = 3, SD). (b) Percent of H650 cells in G1, S and G2 phase following 24 h of ceritinib treatment at the indicated concentrations (n = 3, median ± SD). (c) Scatter plot and DNA histogram of raw flow cytometry data in (b). (d) Crystal violet stain of clonogenic assay for H23 cells following 10 days of ceritinib treatment at the indicated concentrations (µM) (n = 3).
Chemical proteomic characterization of ceritinib

To elucidate the mechanism of action (MoA) of ceritinib’s antiproliferative activity in these cells, we applied a systems approach combining both chemical and phosphoproteomics to gain a global view of ceritinib’s target profile as well as the network-wide phosphorylation changes following ceritinib treatment (Figure 32). In order to identify the targets of ceritinib responsible for the observed activity, we generated a coupleable ceritinib analog (c-ceritinib). Guided by a publically available co-crystal X-ray structure of ALK with ceritinib (PDB: 4MKC),(52) derivatization was achieved by modifying the solvent-exposed piperidine moiety of ceritinib to allow for immobilization while retaining activity for ALK (Figure 12). Activity was confirmed by in vitro kinase assays for ALK inhibition (Figure 33a) and by c-ceritinib’s ability to bind EML4-ALK in chemical proteomics experiments (Figure 33b). Pulldown assays were performed using lysates from H650,

![Figure 32. Integrated functional proteomics.](image)

Proteomics strategy to elucidate ceritinib’s MoA combining target identification by chemical proteomics and determination of the downstream signaling changes using both pY phosphoproteomics as well as global (pSTY) phosphoproteomics following SILAC labeling.
H23, and H3122 cells as well as patient-derived, primary lung cancer samples. Subsequent LC-MS/MS analysis identified, where expressed, ceritinib's cognate target ALK and its known off-target IGF1R. Interestingly, c-ceritinib also bound multiple other kinases including FAK1 (encoded by PTK2), RSK1/2 (encoded by RPS6KA3/1), ERK1/2, CAMKK2 and FER (Figure 34a), of which FAK1 and RSK1/2 were consistently most prominently enriched across cell lines (Figure 34b). Importantly, FAK1 and RSK1/2 can be competed away from the affinity matrix using free, unmodified ceritinib thereby confirming specificity of the interaction (Figure 34c). Ceritinib was also able to enrich for RSK1/2 and FAK1 in the lung cancer patient samples (Figure 34d–e). RSK1/2 have two functional kinase domains, with the C-terminal kinase domain (CTKD) being responsible for autophosphorylation of RSK1/2 at S380/S386 in the interdomain linker region, thus activating RSK1/2, and the N-terminal kinase domain (NTKD) being responsible for phosphorylating downstream substrates.(288,290–292) Cross-competition experiments with specific inhibitors of either the NTKD or the CTKD revealed that ceritinib is binding to the CTKD of RSK1/2 (Figure 34f).

Figure 33. Validation of c-ceritinib probe. (a) Percent ALK activity in in vitro kinase assays following treatment with ceritinib or c-ceritinib at the indicated concentrations. Performed in technical duplicate. (b) Western blot of eluates from drug affinity chromatography in H3122 cells showing c-ceritinib is able to bind to ALK (n = 1).
Figure 34. Chemical proteomic characterization of ceritinib. (a) Kinome tree representing kinases identified in chemical proteomics experiments with > 2 exclusive unique spectra. Circles consist of 3 sections representing identification in H650, H23, and H3122 cells, respectively. Kinase phylogenetic tree adapted courtesy of Cell Signaling Technology, Inc. (www.cellsignal.com). (d) Kinases identified in H650, H23 and H3122 chemical proteomics experiments. NSAF = normalized spectral abundance factor; CRAPomePCT: likelihood of specific interaction according to representation in CRAPome. Data is reflective of the mean of biological duplicates. (c) Western blot of eluates from c-ceritinib pulldowns in H650 cells ± 20 \(\mu\)M unmodified ceritinib (n = 2). Ceritinib is able to compete FAK1 and RSK1/2/3 from the affinity matrix. TCL = total cell lysate, ceri = ceritinib, amp = ampicillin.
Ceritinib inhibits multiple kinase targets including FAK1, RSK1/2 and FER

Several of the kinases identified by chemical proteomics were confirmed by in vitro kinase assays as new targets with IC₅₀ values ranging from low/mid nanomolar for FER (5 nM), FAK1 (ø 20 nM, 6.3 nM (Reaction Biology)/33 nM (Eurofins)) and CAMKK2 (ø 26 nM), to upper-nanomolar for RSK2 (ø 275 nM) and RSK1 (ø 584 nM), whereas AMPKα1 was only a weak target (2.5 µM) (Figure 35a). ERK1/2 and mTOR were not significantly inhibited at 1.5 µM suggesting that they are either very weak or indirect targets binding to a more potent ceritinib target (Figure 35b). Consistent with in vitro activity against RSK1/2, ceritinib treatment inhibited RSK1/2 S380/S386 autophosphorylation in H23 cells in a dose-dependent manner. Ceritinib also reduced AKT phosphorylation, indicative of IGF1R inhibition in these cells, as well as autophosphorylation of FAK1 with only minor effects on ERK phosphorylation (Figure 35c). In summary, chemical proteomics identified several new ceritinib targets with nanomolar IC₅₀ values, most prominently FAK1 and RSK1/2, which were functionally inhibited in cells.

RSK1/2, FAK1 and IGF1R represent central target nodes in a heavily interconnected ceritinib effector network

Next, we performed quantitative phosphoproteomic analysis of phosphotyrosine (pY) and phosphoserine/threonine/tyrosine (pSTY) phosphopeptides to identify the downstream signaling pathways altered by ceritinib treatment. Immunoprecipitation of pY phosphopeptides and subsequent label-free quantitative LC-MS/MS analysis resulted in detection
of 435 unique phosphopeptides. Global IMAC-enriched phosphoproteomics using SILAC-based quantification yielded 4433 unique pS, pT, and pY phosphopeptides. Individual replicates shared similar intensity distributions (Figure 36a) and had good reproducibility (Figure 36b). We then selected pY phosphopeptides with a Fold Change ≥ 2 or p-value ≤ 0.1, as well as pSTY phosphopeptides present in all replicates with a Fold Change ≥ 2. This analysis led to selection of 121 upregulated and 165 downregulated phosphopeptides (Figure 37a).

Proteins containing these altered phosphopeptides were merged with the kinases enriched in the chemical proteomics experiments and queried (confidence > 0.9) for protein interactions and associations using STRING. (186) Disconnected proteins were excluded from the analysis resulting in a ceritinib effector network of 139 unique protein nodes and 312 edges (interactions). In order to identify subnetworks, the community structure of this network was analyzed by calculating the modularity (modularity = 0.589 with resolution = 2.448) resulting in 9 distinct community modules (Figure 37b). These communities generally corresponded to
Figure 36. Phosphoproteomic characterization of H650 cells following ceritinib treatment. (a) Density plots of intensities across biological replicates. Technical replicates for pY experiments were averaged (mean) together for plotting. Individual replicates share similar intensity distributions for pY experiments (top) and pSTY experiments (bottom). (b) Replicate correlations for biological replicates for pY experiments (top) and pSTY experiments (bottom). Technical replicates for pY experiments were averaged (mean) together for plotting. Scale represents Log$_2$(Intensity).

Figure 37. Ceritinib effector network. (a) Waterfall plot of combined pY and pSTY phosphopeptides. Shaded boxes highlight phosphopeptides passing Fold Change cutoff of 2. (b) Reduced view of the resulting network from merging kinases identified in chemical proteomics with altered phosphoproteins. Edges were queried using STRING. Colors represent community modules. Edge and node size represents the number of connecting edges between modules and number of nodes within a module. Biological function annotations were assigned based on manual inspection.
distinct cellular processes including kinase signaling, DNA repair, and transcription/translation.

The largest module, module 1 (kinase signaling), contained all of ceritinib’s kinase targets with modules 3 and 4 connecting to module 1 directly through RSK1/2. In addition, module 2 was most interconnected with module 1 with 10 unique shared edges.

Figure 38. Subnetwork analysis of ceritinib mechanism of action. (a) KEGG pathway analysis of modules 1-4 from Figure 37b. Red line represents p = 0.05 following correction (Bonferroni). (b) Hive plot of proteins in adherens junction, insulin, mTOR, focal adhesion and KRAS pathways. Red edges represent edges connected to ceritinib targets. Node size, color and position on axes represent eigenvector centrality. (c) Adjacency matrix of network represented in (b). Box color represents proteins linking separate modules. Biological processes were queried from GeneGO. (d) Simplified topological pathway map of ceritinib-modulated network. Ceritinib targets are highlighted in red. Grey nodes were not observed, but manually added to complement signaling pathway connectivity. Red phosphosites were upregulated and blue phosphosites were downregulated following ceritinib treatment. WB = identified by Western blotting.
To gain further insight into the cellular pathways corresponding to these communities, modules 1-4 were chosen for KEGG pathway analysis. This analysis revealed a large number of perturbed pathways, most significantly adherens junction, insulin signaling, mTOR signaling and focal adhesion signaling (Figure 38a). Interestingly, these pathways displayed a large degree of overlap with multiple shared nodes. To pinpoint ceritinib’s most relevant target(s) and key downstream signals in these cells, the generated network was analyzed for critical nodes that may be indicative of inherent vulnerabilities. The most influential nodes are likely to be signaling hubs, which are central with a high edge degree, as well as bridges, which are involved in modular overlap and pathway crosstalk. We therefore merged the focal adhesion, adherens junction, mTOR and insulin signaling pathways, appended key nodes in the MAPK signaling pathway (KRAS, RAF, MEK) as well as upstream/downstream nodes in the network (unobserved by proteomics) (PI3K, AKT, YB1, p70RSK, RHOA) and analyzed the community structure of the resulting subnetwork. Interestingly, RSK1 seemed to be a key bridge protein among these pathways with the other ceritinib targets RSK2, FAK1 and IGF1R being very central to the network with high eigenvector centrality, which is indicative of its importance in the network (Figure 38b). (423) FER on the other hand showed a peripheral interaction with low eigenvector centrality suggesting lower functional relevance. When assessing the community modules of this subnetwork (33), three major communities and a large degree of crosstalk between them with proteins that belonged to multiple modules became apparent (Figure 38c). In particular, there was substantial crosstalk with pathways centered on adhesion signaling and cytoskeletal regulation. The targets and signaling nodes represented by these pathways were subsequently organized into a topological subnetwork for further functional interrogation (Figure 38d). Together, integrated analysis of the chemical and phosphoproteomics data sets suggests RSK1/2, FAK1 and IGF1R to be critical targets in a strongly interconnected ceritinib effector network.
Ceritinib inhibits viability through a polypharmacology mechanism

Since RSK1/2, IGF1R and FAK1 were central nodes in the ceritinib effector network, they were selected for functional validation to determine their relevance for ceritinib’s MoA. Interestingly, siRNA-mediated knockdown or pharmacological inhibition with selective inhibitors of RSK1/2 CTKD (using FMK; SL0101), FAK1 (using PF-573228) and IGF1R (using OSI-906) showed that inhibition of any of these kinases alone had little effect on cell viability across multiple cell lines (Figure 39a-d). However, combined targeting by simultaneous knockdown and pharmacological inhibition of RSK1/2 with either IGF1R or FAK1 resulted in a pronounced decrease of cell viability, which was consistent also upon reversing siRNA and probe molecules for target pairs (Figure 39a-c, upper vs. lower panels). This effect was conserved across both H650 and H1155 cells (Figure 39a-b), whereas in H23 cells co-targeting of RSK1/2 with IGF1R made a more significant contribution than the combination with FAK1 (Figure 39c).

![Figure 39](image)

**Figure 39.** Ceritinib inhibits cell viability through inhibition of IGF1R, FAK1, RSK1 and RSK2. (a-c) Relative cell counts following 96 h siRNA mediated knockdown of IGF1R or FAK1 and 72 h treatment with 20 μM FMK or SL0101, or siRNA mediated knockdown of RSK1/2 and treatment with 1.5 μM OSI-906 or PF-573228 in (a) H650, (b) H1155, (c) and H23 cells. Data is reflective of biological triplicates each performed in technical triplicate (median ± SD). Knockdown efficiencies were determined by immunoblotting (d). Asterisks indicate p-value cut-offs (* : 0.05; ** : 0.01; *** : 0.001) from Wilcoxon Rank Sum test. Asterisks without bars indicate comparison to NT. NT = non-targeting siRNA.
YB1 is a critical downstream node in the ceritinib effector network

Revisiting the ceritinib network model highlighted the multifunctional oncogenic transcription factor YB1 as central signaling effector nodes, on which several modulated pathways converged (Figure 38d). Immunoblot analysis demonstrated that ceritinib treatment inhibited the functionally relevant phosphorylation of YB1 at S102 (Figure 40a–d). Comparison with FMK consistently indicated RSK1/2 to be primarily responsible for YB1 phosphorylation (with contributions from FAK1 in H650 and H1155 cells). Similar to single inhibition of RSK1/2, knockdown of YBX1 (encoding for YB1) had little effect on cell viability across cell lines. However, co-inhibition of IGFR1 or FAK1 significantly reduced cell growth to a similar extent as observed with targeting of RSK1/2 directly suggesting that YB1 is a key downstream effector of RSK1/2 and that its inhibition is required, but not sufficient, for mediating ceritinib’s overall drug effect (Figure 40a–d). Consistently, overexpression of the constitutively active phosphomimetic YB1\textsuperscript{S102D} mutant conferred significant resistance of H650 cells to ceritinib while the inactive YB1\textsuperscript{S102A} mutant or wild-type YB1 provided no protection (Figure 40e–f). In summary, these results suggest that ceritinib inhibits viability of ALK-negative NSCLC cells through a polypharmacology mechanism that involves simultaneous inhibition of RSK1/2, IGFR1 and FAK1 and is largely dependent on the central downstream signaling effector YB1.
Ceritinib strongly synergizes with microtubule inhibition

Since ceritinib caused alterations in various pathways centered on regulation of the microtubules and cytoskeleton (Figure 38c) and considering that the key signaling node YB1 has been reported to confer resistance to microtubule-targeting drugs (375, 424), we hypothesized ceritinib to show synergy with the FDA-approved microtubule inhibitor paclitaxel. Indeed, employing the Bliss model of independence, the combination of ceritinib and paclitaxel displayed substantial synergy.
in H650 and H1155 cells across multiple concentrations (Figure 41a). This was corroborated by independent experiments analyzed using the Chou-Talalay combination index method, which revealed strong to very strong synergy as well as a large shift of the ceritinib dose response curve (Figure 41b). Addition of ceritinib also caused shifts in the paclitaxel dose-response curves in several cell lines, most pronounced again in H1155 cells (Figure 41c). Importantly, as indicated by cleaved caspase 3/7 and PARP1, the combination of ceritinib and paclitaxel in H1155 cells

**Figure 41. Ceritinib strongly synergizes with paclitaxel.** (a) Heatmap of cell viability (top) and deviation from Bliss (bottom) in H650 and H1155 cells following 72 h treatment with ceritinib and paclitaxel. (n = 3) (b) Dose response curve of ceritinib ± paclitaxel in H650 and H1155 cells (n = 3, SD). CI values were calculated using CompuSyn. Highlighted area reflects reported maximum range of ceritinib concentration in patient plasma (1.4 - 2.3 μM). (c) Cell viability curves of paclitaxel (blue) alone or with ceritinib (red) in H23 (0.5μM), H226, H661, H650, HCC2935, H1155 and H460 cells (all 2 μM ceritinib) following 72 h treatment. Individual curves are representative of biological triplicates performed in technical triplicate (mean ± SD). (d) (left) Relative apoptosis following treatment at the indicated concentrations of ceritinib and paclitaxel as determined by caspase 3/7 cleavage (n = 3). Data were recorded every 2 h for 72 h using an Incucyte Live Cell Analysis System. (right) Western blot of PARP1 and cleaved caspase 3 following 48 h of treatment (n = 3).
also caused a significant increase in apoptosis whereas single drug treatment showed no effect (Figure 41d).

To determine the underlying mechanism of this synergy, we combined paclitaxel with genetic knockdown of PTK2 (encoding FAK1), RPS6KA1/3 (encoding RSK1/2), IGF1R (encoding IGF1R) or YBX1 (encoding YB1). Knockdown of any of these genes moderately, but significantly enhanced the efficacy of paclitaxel suggesting they all contribute to varying degrees to the observed synergy, with the largest effect detected for PTK2 knockdown (Figure 42a). Conversely, genetic knockdown of FER (which was only on the periphery of the subnetwork), did not enhance paclitaxel activity (Figure 42a). Likewise, knockdown of CAMKK2 (not present in the subnetwork) or its downstream substrate PRKAA1 (encoding AMPKα1) did not synergize with paclitaxel either (Figure 42b). However, pharmacological inhibition of IGF1R, FAK1, and/or RSK1/2 markedly sensitized H1155 cells to paclitaxel (Figure 42c). Consistent with the gene silencing results, the FAK1/RSK/IGF1R inhibitor CEP-37440(425) strongly synergized with paclitaxel (Figure 42d). The dual CAMKK2/AMPK inhibitor compound C did not synergize with paclitaxel. In addition, combination of paclitaxel with the CDK4/6 inhibitor ribociclib, which has no activity against RSK1/2, IGF1R, or FAK1 (426), did not exhibit synergy (Figure 42e). Taken together, ceritinib displays pronounced synergy with paclitaxel in H1155 cells that leads to potent induction of apoptosis, and this synergy is dependent on inhibition of RSK1/2, IGF1R and FAK1.
FAK1 autophosphorylation correlates with pronounced synergy

Since the extent of synergy observed in H1155 cells was markedly more pronounced than in H650 or H23 cells, we sought to identify additional cell lines that responded similarly as well as a
potential predictive marker of exceptional sensitivity to this combination. To this end, a panel of 14 cell lines was analyzed for basal expression and phosphorylation of several key nodes in the ceritinib effector network. While for most of these signals there was no apparent correlation, FAK1 autophosphorylation at Y397 was notably much stronger in H1155 cells than in H650 or H23 (Figure 43a). Notably, H460 was the only other cell line in this panel that displayed a high amount of pFAK1. Indeed, these cells also responded strongly to the combination of ceritinib and paclitaxel with a drastic shift in the dose response curve and highly synergistic CI values in viability assays (Figure 43b), as well as a pronounced combination effect in clonogenic assays (Figure 43c). Since both the H1155 cells and H460 cells are large-cell lung carcinoma (LCLC) cell lines harboring KRAS mutations, we hypothesized that additional RAS mutant LCLC cell lines may respond similarly to this drug combination. Subsequent testing of the NRAS-mutant LCLC cell line H1299 showed that these cells also responded strongly to the combination of ceritinib and paclitaxel although resistant to ceritinib treatment alone (Figure 43d). Notably, H1299 cells also displayed a substantial degree of FAK1 autophosphorylation (Figure 43e). These findings suggest that high FAK1 pY397, possibly in the context of RAS mutations, may be predictive of exceptional synergy to the combination of ceritinib and paclitaxel.
In order to determine the prevalence of FAK1 autophosphorylation across various lung cancer subtypes, we analyzed two different lung cancer tissue microarrays (TMA) for pFAK1 Y397 by immunohistochemistry (IHC). TMA1 consisted of 60 human tumor samples (US Biomax) and TMA2 consisted of 69 patient-derived xenograft (PDX) tumors (Charles River Laboratories). The TMAs contained tumors of varying histologies, the majority being adenocarcinoma (LUAD) and squamous cell carcinoma (LUSQ), followed by a relatively high number of LCLC (Figure 44a). IHC analysis revealed a large range of pFAK1 staining [range = (0, 280), average = 96.1, median

Figure 43. FAK1 autophosphorylation may be predictive of synergistic response to ceritinib and paclitaxel. (a) Western blot of pFAK1, FAK1, pYB1, YB1, pIGF1R, IGF1R, pRSK, pRPS6, and RPS6 across 14 cell lines (n = 4). (b) Dose response curves of ceritinib ± paclitaxel in H460 cells (n = 3, SD). CI values were calculated using CompuSyn. Highlighted area reflects reported maximum range of ceritinib concentration in patient plasma (1.4 – 2.3 µM). (c) Crystal violet stain of clonogenic assay for H460 cells following 7 days of ceritinib and paclitaxel treatment at the indicated concentrations (µM) (n = 3). (d) Dose response curves of ceritinib ± paclitaxel in H1299 cells (n = 3, SD). (e) Western blot of pFAK1 in H1155, H460, H1299 and H650 cells (n = 3).
= 100] with most tumors having low (48%) to moderate (36%) staining (H-score < 130) and 16% showing strong pFAK1 staining (Figure 44b,d). Staining was of high quality with replicates of TMA2 showing good reproducibility (Figure 44c). Interestingly, 24% of LCLC tumors displayed high pFAK1 staining (Figure 44d). Adenocarcinoma in situ (AIS) tumors had the highest amount of pFAK1 staining with 50% of tumors having high staining with the caveat that the overall number of AIS cases was low (Figure 44d). Using this data, we prioritized 3 PDX models from TMA2 with high pFAK1 staining (one LCLC [PDX1 and PDX3] and one LUSQ [PDX2], Figure 45a) based on their FAK1 phosphorylation status for functional validation by ex vivo 3D clonogenic assays (which

**Figure 44.** FAK1 autophosphorylation characterization across cell lines, patient tumors and PDX tumors. (a) Composition of combined TMA1 and TMA2 tumors. (b) H-score distribution of combined TMA1 and TMA2. (c) Correlation of replicate tumors analyzed in TMA2. (d) Sankey diagram of pFAK1 staining across lung tumor histologies for TMA1 and TMA2. Chord area is proportional to the number of tumors. Percentage of each tumor type is represented on the left. LCLC = large cell lung carcinoma, SCLC = small cell lung carcinoma, LUAD = lung adenocarcinoma, LUSQ = lung squamous cell carcinoma, ADSQ = adenosquamous carcinoma, AIS = adenocarcinoma in situ, NOS = not otherwise specified.
were performed by Oncotest). Interestingly, all three models exhibited pronounced synergy between ceritinib and paclitaxel (Figure 45b-d). In summary, these results suggest that pFAK1 is a prevalent signal in lung cancer, and that further evaluation of pFAK1 as a potential mechanistic biomarker for synergy between ceritinib and paclitaxel is warranted.

![Figure 45. FAK1 autophosphorylation biomarker validation.](image)

(a) Bee swarm diagram of H-scores across lung histologies for individual tumors in TMA1 and mean of tumors in TMA2 with median ± SD displayed. Medium and high pFAK1 cutoffs are annotated with dotted lines. (b-d) Heatmap of cell viability (left) and deviation from Bliss (right) in (b) PDX1, (c) PDX2 and (d) PDX3 tumors grown in 3D ex vivo culture following treatment with ceritinib and paclitaxel at the indicated concentrations (n = 2). LCLC = large cell lung carcinoma, SCLC = small cell lung carcinoma, LUAD = lung adenocarcinoma, LUSQ = lung squamous cell carcinoma, ADSQ = adenosquamous carcinoma, AIS = adenocarcinoma in situ, NOS = not otherwise specified.
Discussion

Network-based targeting strategies have the potential to reveal new therapeutic opportunities in diseases lacking strong and actionable drivers (23,67). Here, we present a systems chemical biology approach for elucidating polypharmacology mechanisms of multi-kinase inhibitors, which enables rational design of synergistic drug combinations and identification of mechanistic biomarker candidates. It furthermore facilitates the systems level understanding of complex signaling networks. This approach utilizes an integrated technology platform based on unbiased phenotypic drug screening, in this case for cancer cell viability, to identify cellular effects unrelated to a drug’s cognate target. Subsequent functional proteomics, comprised of chemical proteomics and phosphoproteomics, allows identification of a kinase inhibitor’s cell-specific targets and network-wide signaling effects, which are synthesized into a mechanistic hypothesis for functional validation. This platform is generalizable to other chemical proteomic technologies such as activity-based protein profiling (ABPP) or cellular thermal shift assays (CETSA). Importantly, this platform could also be adapted to other drug classes, such as HDAC inhibitors, and diseases other than cancer depending on the post-translational modification MS method, e.g. acetylation, or choice of screening assay/cell type, including primary cells. Furthermore, it is not limited to cases where polypharmacology is elicited entirely through non-canonical targets, as observed here for ceritinib, but can be applied similarly to drugs that act through mixed mechanisms involving cognate and other targets. The latter may in fact constitute a more common scenario. For example, it has been shown that concomitant inhibition of IGF1R with ALK in ALK-fusion positive lung cancer patients improves therapeutic efficacy, explaining some ceritinib’s enhanced efficacy in the disease (427).

Application of this approach identified the FDA-approved ALK-inhibitor ceritinib to harbor antiproliferative activity also in ALK-negative NSCLC cell lines. This activity, while less potent than in ALK-rearranged NSCLC (52), which is likely for serendipitous drug effects not optimized for the particular target(s), was within clinically relevant plasma concentrations (420,421). The underlying
MoA depended not on a single target, but constituted a complex case of polypharmacology involving simultaneous inhibition of multiple non-canonical targets, namely FAK1, IGF1R and RSK1/2. Inhibition of each kinase alone showed no significant effects, thus highlighting the relevance of targeting the signaling network at multiple nodes. In-depth analysis of the ceritinib-modulated network led to identification of YB1 as a key effector node for ceritinib single agent activity. Based on the network converging on microtubule regulation and the fact that YB1, either downstream of RSK1/2 or FAK1, confers resistance to microtubule targeting agents in prostate and ovarian cancers (375,424), this led to discovery of ceritinib synergy with paclitaxel, which was highly pronounced in a subset of cell lines. Although YB1 signaling contributed to the MoA of this drug combination, FAK1 targeting displayed somewhat stronger effects suggesting a role also of other, uncharacterized FAK1-regulated signals. Additional studies are required to determine the exact mechanism of this signaling interaction.

Consistent with a more prominent role of FAK1 for the ceritinib/paclitaxel synergy, high FAK1 autophosphorylation indicated exceptional sensitivity to this drug combination. Notably, cell lines lacking FAK1 autophosphorylation still responded to the combination albeit to lesser extent. Importantly, LCLC and AIS histology tumors showed a much higher pFAK1 prevalence than LUAD, which was previously reported to display high pFAK1 in approximately 25% of cases (LCLC and AIS were not present in that analysis)(182). Our findings of high pFAK in LCLC and AIS tumors suggest that the ceritinib/paclitaxel combination has a higher potential of pronounced effects in these subtypes, which compared to LUAD less commonly feature actionable kinase mutations and translocations and therefore constitute a higher unmet medical need. Dedicated clinical studies are required to determine the translational potential of this drug combination and associated biomarker candidate in ALK-negative lung cancers.

In summary, we present a systems chemical biology approach for elucidation of multi-kinase inhibitor polypharmacology mechanisms and their underlying network biology using an integrated phenotypic screening and functional proteomics platform. Using the FDA-approved,
second-generation ALK inhibitor ceritinib, we illustrate the potential of this strategy for the subsequent design of synergistic drug combinations and identification of a potential biomarker candidate that may enable repurposing of ceritinib, in combination with paclitaxel, for treatment of ALK-negative NSCLC.
Chapter seven: Future perspectives

In this dissertation, I have discussed an approach to identify and characterize novel drug repurposing opportunities based on the beneficial off-target effects of clinically advanced compounds. I have shown that by utilizing drug screening and chemical proteomics, we can identify and characterize the beneficial off-target activity of a compound. I also have shown that this approach can be extended with phosphoproteomics and network analysis to deeply characterize and elucidate complex polypharmacology mechanisms of action. By integrating these technologies and strategies, we have developed a pipeline in which to identify novel drug repurposing opportunities in which to treat cancers lacking strong oncogenic drivers.

The strategy described in this dissertation greatly enhances the identification of drug repurposing opportunities and may lead to an increase of new therapy options for difficult-to-treat patients in the future. We have identified two novel drug repurposing opportunities, both of which have potential clinical implications. Tivantinib, for example, has been extensively investigated clinically as a MET inhibitor (46 registered clinical trials); however many of the trials failed to meet primary endpoints, likely in part due to tivantinib’s inability to potently inhibit MET. ABT-199, which we identified to be effective in combination with tivantinib in AML, has also been extensively characterized in the clinic and has received FDA breakthrough therapy designation for multiple cancers, including AML (January 2016). Repurposing tivantinib for AML could represent a potential new effective therapy (and potential combination therapy), which warrants further investigation. In addition, the effective combination of ceritinib and microtubule disruption will be translated into a clinical trial at H. Lee Moffitt Cancer Center & Research Institute. We developed a clinical trial concept, which has been approved by Novartis, in order to test the combination of
ceritinib and the more clinically applicable taxane therapy, docetaxel. Patients will be eligible for the trial if they lack \textit{EGFR} activating mutations, lack an \textit{ALK} translocation, and if they had not responded to first line therapy (immunotherapy). This exemplifies the speed in which a repurposing approach can move a potential therapy from discovery to clinical implementation (~3 years compared to 10 – 15 years for standard drug discovery).

The main limitation of this approach is the lack of a scalable biomarker discovery platform for the drugs identified for potential repurposing. This approach could be enhanced through the implementation of a biomarker pipeline to identify predictors of single or combination therapy response. For example, with sufficient profiling, the area under the dose response curve or a binary classifier (sensitive / not sensitive, synergy / no synergy) may be able to be predicted using a gene expression based feature set (typically ≤ 7 genes) derived from an elastic net regularization optimized on the root mean square error or prediction accuracy, respectively (see Chapter 5 for an example implementation). The resulting model could then be translated to predict PDX or patient response to a compound of interest based on the intratumoral expression of the genes in this small feature set. While useful for biomarker candidate identification, extensive tuning would be necessary before clinical implementation could be seriously considered as it is difficult for an elastic net regularization to achieve sufficient accuracy for clinical implementation (ROC-AUC > 0.9). Alternatively, the network model used to prioritize ceritinib targets could be extended with significantly different genes between sensitive vs. resistant cell lines (based on p-values and top 1% signal-to-noise). Candidate biomarkers may then be able to be identified by analyzing the genes / proteins with the highest eigenvector and closeness centrality, the most mechanistically relevant being genes with a high target degree. Either of the described approaches could currently be implemented and enhance the biomarker discovery process by overcoming the antibody requirement of profiling network nodes by western blotting and allow for larger scale in profiling which could lead to even more robust biomarkers.
Future approaches however should take the concepts of network medicine and polypharmacology into consideration in order to build a comprehensive platform that is predictive of not only a tumor’s growth or responsiveness to therapy, but also the signaling events that underlie this response. By integrating the concepts of network medicine and polypharmacology with cutting edge artificial intelligence, it would be possible to model the mechanism of action of compounds in silico in order to make predictions of responsiveness to that compound based on the cellular context. To accomplish this, an effective approach could be to build an artificial neural network to predict tumor response to therapy by modeling the underlying network alterations and their interactions. This can be done by structuring the neural network to mirror exactly a hierarchical network model of the cell with banks of neurons capturing the complex network activation states of individual cellular / biological processes (428). The model would then be trained on genome / proteome-wide chemical-genetic interactions through which it will predict tumor growth. During training, the model would weight cellular processes and determine their signaling state which minimizes prediction error. By doing so, it would be modeling the pathway alterations responsible for the observed anticancer response (the mechanism of action of the drug). Therefore, a likely scenario could be that the most highly weighted process for tivantinib response in AML would be “GSK3 signaling.” It is also likely that modeling a drug with a polypharmacology mechanism of action, such as ceritinib in ALK negative NSCLC, would yield multiple highly weighted processes such as “activation of MAPK activity”, “focal adhesion”, and “insulin-like growth factor receptor signaling pathway”. With careful curation of compounds and enough data, this platform could be streamlined into a comprehensive drug repurposing platform to identify effective treatments for cancers lacking strong oncogenic drivers or even into a clinical service in which to make treatment decisions for individual patients with exquisite accuracy (>99%). Successful implementation of such a system could completely change personalized medicine and lead to more effective patient-therapy matching and an exponential increase in the number of FDA-approved therapies.
List of References


42. Ferrara E. A Large-Scale Community Structure Analysis In Facebook. EPJ Data Sci [Internet]. 2012 Dec;1(1). Available from: http://arxiv.org/abs/1106.2503


241. Kramer T, Schmidt B, Lo Monte F. Small-Molecule Inhibitors of GSK-3: Structural Insights and Their Application to Alzheimer’s Disease Models [Internet]. International Journal of


335. Palmer A, Gavin AC, Nebreda AR. A link between MAP kinase and p34(cdc2)/cyclin B during oocyte maturation: p90(rsk) phosphorylates and inactivates the p34(cdc2) inhibitory kinase Myt1. EMBO J. 1998 Sep 1;17(17):5037–47.


Costa P, Scales TME, Ivaska J, Parsons M. Integrin-Specific Control of Focal Adhesion Kinase and RhoA Regulates Membrane Protrusion and Invasion. PLOS ONE. 2013 Sep 9;8(9):e74659.


Appendix A:
c-(-)-tivantinib synthesis

(Note to reader: Parts of this section have been published previously in Remsing Rix et al. ACS Chemical Biology doi: 10.1021/cb400660a and are being reproduced with permission from American Chemical Society copyright 2016.)

For chemical structures and synthetic strategy, see Figure 11. The synthesis of intermediate 2 was carried out according to the literature-reported protocols (239,429,430) from the commercially available starting material tetrahydroquinoline (1). The intermediate 2 was achieved via 5 steps with 40% overall yield. Synthesis was performed by the Moffitt Chemistry Core Facility.

_Tert-butyl 3-(3-(2-amino-2-oxoethyl)-1H-indol-1-yl)propylcarbamate (4)_

The starting material, Indole-3-acetamide 3 (0.114 g, 1.30 mmol) in DMF (0.5 mL) was added drop wise, under argon atmosphere at 0 °C to a stirred suspension of sodium hydride (39 mg, 60% in mineral oil, washed with hexane prior to the reaction) in DMF (0.5 mL). The reaction mixture was stirred under inert conditions for 1 h, cooled to 0 °C, added tert-butyl 3-bromopropylcarbamate (233 mg, 1.96 mmol) in DMF (0.5 mL), and stirred at r.t. overnight (12-14 h). The reaction was monitored using HPLC-MS and TLC (4% MeOH in DCM, \(R_f = 0.4\), major spot and other minor impurities). The HPLC-MS showed completion of the reaction. The crude reaction mixture was diluted with ethyl acetate (15 mL) and the organic layer was washed with water (10 mL x 3). Combined organic phase was separated, dried (MgSO₄), filtered and concentrated to dryness to obtain a colorless foamy solid (250 mg, 116% yield). This product was
taken to the next stage without further purification. $^1$H NMR (400 MHz, DMSO-d$_6$) δ 7.53 (d, $J = 8.0$ Hz, 1H), 7.38 (d, $J = 8.0$ Hz, 1H), 7.32 (br s, 1H), 7.21 (s, 1H), 7.09 (t, $J = 8.0$ Hz, 1H), 7.00-6.95 (m, 2H), 6.84 (br s, 1H), 4.11 (t, $J = 6.8$ Hz, 2H), 3.43 (s, 2H), 2.91-2.87 (m, 2H), 1.85-1.77 (m, 2H), 1.36 (s, 9H); HPLC-MS (ESI+) m/z 332.2 (M + H)$^+$. 

3-(1-(3-Aminopropyl)-1H-indol-3-yl)-4-(5,6-dihydro-4H-pyrrolo[3,2,1-ij]quinolin-1-yl)-1H-pyrrole-2,5-dione (6)  
To a mixture of methyl 2-(5,6-dihydro-4H-pyrrolo[3,2,1-ij]quinolin-1-yl)-2-oxoacetate (2) (365 mg, 1.5 mmol) and 4 (546 mg, 1.65 mmol) in anhydrous THF (2.5 mL) was added a solution of potassium tert-butoxide (4.5 mL, 1M, THF) drop wise over 30 minutes. The reaction mixture was stirred at 0 ºC for 2 h, and monitored by tlc (tlc showed no starting material [Rf = 0.8, 60% ethyl acetate in hexane] left in the reaction mixture). Concentrated HCl (3.65 mL) was added drop wise at 0 ºC and the reaction mixture was allowed to warm to r.t. and stirred for an additional hour. The mixture was diluted with EtOAc (60 mL), washed with water (2 x 25 mL) and brine (25 mL). Combined organic phase was dried (Na$_2$SO$_4$), filtered and concentrated. Only 200 mg of crude product recovered from the organic phase (according to the $^1$H NMR majority of the isolated product was impurity). The aqueous layer was neutralized (sat. NaHCO$_3$, pH: 8-9), and extracted with EtOAc (2 x 25 mL). Combined EtOAc was dried (Na$_2$SO$_4$) and concentrated and only a small amount of required product (mainly impurity) was isolated. The sticky red solid that was precipitated in the aqueous layer was decanted and dissolved in MeOH (10 mL). Methanol was evaporated under high vacuum to afford the required product 6 as a red solid (512 mg, 81% yield). The NMR and HPLC-MS indicated isolation of the pure amine 6. $^1$H NMR (400 MHz, CD$_3$OD) δ7.79 (s, 1H), 7.63 (s, 1H), 7.43 (d, $J = 8.4$ Hz, 1H), 7.08 (apparent t, $J = 7.2$ Hz, 1H), 7.00 ($J = 8.0$ Hz, 1H), 6.74-6.70 (m, 2H), 6.50 (t, $J = 8.0$ Hz, 1H), 6.42 (d, $J = 8.0$ Hz, 1H), 4.32 (t, $J = 6.8$ Hz, 2H), 4.22 (t, $J = 5.6$ Hz, 2H), 2.94 (t, $J = 6.0$ Hz, 2H), 2.79 (apparent t, $J = 7.6$ Hz, 2H), 2.24-
2.19 (m, 2H), 2.15-2.08 (m, 2H); HPLC-MS (ESI+) m/z 425.3 (M + H)+; LC−MS (ESI+) m/z 447.2 (M + Na)+; HRMS (ESI + ve) m/z calculated for C_{26}H_{25}N_{4}O_{2} (M+H)+ 425.1972, found 425.1968.

Synthesis of (±)-trans-3-(1-(3-aminopropyl)-1H-indol-3-yl)-4-(5,6-dihydro-4H-pyrrolo[3,2,1-ij]quinolin-1-yl)pyrrolidine-2,5-dione (9 and 10 as a racemic mixture)

To a solution of 6 (164 mg, 0.386 mmol) in anhydrous MeOH (5 mL), magnesium turnings were added and heated to reflux under inert atmosphere for 40 min. The color of the reaction mixture was changed from red to pink. The reaction mixture was cooled to r.t. and methanol was evaporated. 1M HCl (30 mL) was added to dissolve the solid followed by sat. sodium bicarbonate (drop wise) to neutralize the mixture (pH = 8-9). The crude mixture was extracted with DCM (30 mL x 3) and the combined organic phase was dried (Na_{2}SO_{4}), filtered and concentrated to obtain the crude product (130 mg) as a pink solid (i.e. 9 and 10 as a racemic mixture). The crude product was purified using SiO\textsubscript{2} chromatography by gradient elution of DCM:MeOH (100% DCM to 75% DCM and Methanol). The SiO\textsubscript{2} column was pre washed/eluted with 1% triethylamine in DCM. [TLC; 20% MeOH in DCM, product \( R_f = 0.1 \)]. The product isolated from SiO\textsubscript{2} chromatography was re-crystallized using EtOAc: Hexane (1:1) to afford the desired pure product (±)-trans (3S,4S)-3-(1-(3-aminopropyl)-1H-indol-3-yl)-4-(5,6-dihydro-4H-pyrrolo[3,2,1-ij]quinolin-1-yl)pyrrolidine-2,5-dione (9 and 10 as a racemic mixture) as a pale pink solid (72 mg, 44%). \(^1\)H NMR (400 MHz, CD\textsubscript{3}OD) \( \delta \) 7.44 (d, \( J = 8.4 \) Hz, 1H), 7.33 (d, \( J = 8.0 \) Hz, 1H), 7.20-7.16 (m, 2H), 7.10-7.08 (m, 2H), 7.01 (t, \( J = 8.0 \) Hz, 1H), 6.92-6.85 (m, 2H), 4.45 (dd, \( J = 9.2, 6.4 \) Hz, 2H), 4.20 (t, \( J = 6.8 \) Hz, 2H), 4.12 (apparent t, \( J = 5.6 \) Hz, 2H), 2.96 (t, \( J = 6.0 \) Hz, 2H), 2.64 (t, \( J = 7.2 \) Hz, 2H), 2.22-2.16 (m, 2H), 2.01-1.96 (m, 2H); HPLC-MS (ESI+) m/z 427.2 (M + H)+; LC−MS (ESI+) m/z 427.2 (M + H)+; HRMS (ESI + ve) m/z calculated for C_{26}H_{27}N_{4}O_{2} (M+H)+ 427.2129, found 427.2124.

Separation of the (±)-trans-3-(1-(3-aminopropyl)-1H-indol-3-yl)-4-(5,6-dihydro-4H-pyrrolo[3,2,1-ij]quinolin-1-yl)pyrrolidine-2,5-dione using chiral analytical HPLC and chiral preparative HPLC
Analytical HPLC was performed using a JASCO HPLC system equipped with a PU-2089 Plus quaternary gradient pump and a UV-2075 Plus UV-VIS detector, and using a CHIRALPAK AD column (250 x 4.6 mm). The racemic mixture \((\pm)-\text{trans-3-(1-(3-aminopropyl)-1H-indol-3-yl)-4-(5,6-dihydro-4H-pyrrolo[3,2,1-ij]quinolin-1-yl)pyrrolidin-2,5-dione}\) was separated with method A: 80% IPA, 20% Hexane 0.5 mL/min., 60 min; analytical chromatography yielded faster eluting peak with \(t_R = 17.5\) min; and slower eluting peak with \(t_R = 40.9\) min.; (Chromatogram-1). The two peaks in chromatogram-1 appeared broad. The enantiomers were separated with better resolution with method B: 40% IPA 60% Hexane (0.1% TFA) 0.5 mL/min 60 min.; analytical chromatography yielded a faster eluting peak with \(t_R = 12.2\) min.; and a slower eluting peak with \(t_R = 32.0\) min.; (Chromatogram-2). The chromatogram-2 showed better resolution of peaks. However when the method B with 0.1% TFA was transferred to preparative chiral HPLC, the separated enantiomers had decomposed according to the \(^1\)H NMR and LCMS analysis. Therefore method B was further modified with 30% IPA and 70% Hexane, 120 min, 0.5 mL/min and analytical HPLC runs were performed without TFA (Method C). The Method C was adapted to preparative HPLC conditions as described below.

A mixture of \((\pm)\) trans 3-(1-(3-aminopropyl)-1H-indol-3-yl)-4-(5,6-dihydro-4H-pyrrolo[3,2,1-ij]quinolin-1-yl)pyrrolidin-2,5-dione (57 mg) was subjected to preparative chiral HPLC using a JASCO HPLC system equipped with a PU-2086 Plus quaternary gradient pump and a MD-2010 Plus multi-wavelength detector, and using a CHIRALPAK AD column (250 x 21 mm). The racemic mixture was separated using method C: 30% IPA 70% Hexane 10 mL/min, 120 min. The preparative chiral chromatography yielded a faster eluting peak \((t_R = 16.9\) min) of the trans isomer having a negative optical rotation \(([\alpha]_D^{25} = -103.7\) (c: 19.5 mg in 1.7 mL Methanol) assigned to \((-)\) \((3R,4R)-\text{trans-3-(1-(3-aminopropyl)-1H-indol-3-yl)-4-(5,6-dihydro-4H-pyrrolo[3,2,1-ij]quinolin-1-yl)pyrrolidin-2,5-dione}\) \((9)(81–83)\) and a slower eluting peak \((t_R = 53.3\) min.) of the trans isomer having a positive optical rotation \(([\alpha]_D^{25} + 89.6\) (c: 20.5 mg in 1.7 mL Methanol) assigned to \((+)\)
(3S,4S)-3-(1-(3-aminopropyl)-1H-indol-3-yl)-4-(5,6-dihydro-4H-pyrrolo[3,2,1-ij]quinolin1-yl)pyrrolidin-2,5-dione (10).\(^1\) Optical Rotations were conducted in Methanol at 25°C. Absolute stereochemical assignments were based upon relative retention time of related compounds.\(^1\) Total of 57 mg of racemic mixture (15 mg in 50-100 µl of DCM/injection) was used for preparative HPLC chromatography; 19.5 mg of the faster eluting peak (\(R,R\)-3) and 20.5 mg of the slower eluting peak (\(S,S\)-4) were isolated.
Appendix B:

Letters of approval

2/7/2018

Rightslink® by Copyright Clearance Center

Title: APOSTL: An Interactive Galaxy Pipeline for Reproducible Analysis of Affinity Proteomics Data
Author: Brent M. Kuenzi, Adam L. Borne, Jiannong Li, et al
Publication: Journal of Proteome Research
Publisher: American Chemical Society
Date: Dec 1, 2016
Copyright © 2016, American Chemical Society

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

BACK CLOSE WINDOW

Copyright © 2018 Copyright Clearance Center, Inc. All Rights Reserved. Privacy statement Terms and Conditions. Comments? We would like to hear from you. E-mail us at customerservice@copyright.com
GSK3 Alpha and Beta Are New Functionally Relevant Targets of Tivantinib in Lung Cancer Cells

Author: Lily L. Remsing Rix, Brent M. Kuenzi, Yunting Luo, et al
Publication: ACS Chemical Biology
Publisher: American Chemical Society
Date: Feb 1, 2014
Copyright © 2014, American Chemical Society

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: “Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society.” Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

BACK CLOSE WINDOW
Title: Polypharmacology-based ceritinib repurposing using integrated functional proteomics
Author: Brent M Kuenzi, Lily L Remsing Rix, Paul A Stewart, Bin Fang, Fumi Kinose et al.
Publication: Nature Chemical Biology
Publisher: Springer Nature
Date: Oct 9, 2017
Copyright © 2017, Springer Nature

Author Request

If you are the author of this content (or his/her designated agent) please read the following. If you are not the author of this content, please click the Back button and select no to the question "Are you the Author of this Springer Nature content?"

Ownership of copyright in original research articles remains with the Author, and provided that, when reproducing the contribution or extracts from it or from the Supplementary Information, the Author acknowledges first and reference publication in the Journal, the Author retains the following non-exclusive rights:

To reproduce the contribution in whole or in part in any printed volume (book or thesis) of which they are the author(s).

To reuse figures or tables created by the Author and contained in the Contribution in oral presentations and other works created by them.

To post a copy of the contribution as accepted for publication after peer review (in locked Word processing file, of a PDF version thereof) on the Author's own web site, or the Author's institutional repository, or the Author's funding body's archive, six months after publication of the printed or online edition of the Journal, provided that they also link to the contribution on the publisher's website.

Authors wishing to use the published version of their article for promotional use or on a web site must request in the normal way.

If you require further assistance please read Springer Nature's online author reuse guidelines.

For full paper portion: Authors of original research papers published by Springer Nature are encouraged to submit the author's version of the accepted, peer-reviewed manuscript to their relevant funding body's archive, for release six months after publication. In addition, authors are encouraged to archive their version of the manuscript in their institution's repositories (as well as their personal Web sites), also six months after original publication.

v1.0