Identification of Small Molecule Inhibitors of 3-Phosphoglycerate Dehydrogenase to Target Serine Biosynthesis in Cancers

Identification of Small Molecule Inhibitors of 3-Phosphoglycerate Dehydrogenase to Target Serine Biosynthesis in Cancers

A dissertation presented

by

Edouard Mullarky

to

The Division of Medical Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biological and Biomedical Sciences

Harvard University

Cambridge, Massachusetts

January 2016

Dissertation Advisor: Dr. Lewis C. Cantley Edouard Mullarky

Identification of Small Molecule Inhibitors of 3-Phosphoglycerate Dehydrogenase to Target Serine Biosynthesis in Cancers

Abstract

Cancer cells are known to reprogram their metabolism in order to promote growth and proliferation. The amino acid serine is utilized in a plethora of anabolic reactions and supports the synthesis of all three major macromolecular classes: proteins, lipids, and nucleic acids. Serine can either be synthesized *de novo* via the phosphoserine pathway or imported from the extracellular space via amino acid transporters. The gene encoding the enzyme 3-phosphoglycerate dehydrogenase (PHGDH), which catalyzes the first committed step of the phosphoserine pathway, is focally amplified in human cancers suggesting that it is pro-tumorigenic. Cancer cell lines that harbor *PHGDH* amplifications, or over express *PHGDH* independently of amplification, are uniquely sensitive to genetic ablation of the pathway. In contrast, cancer cell lines that express little PHGDH, and instead rely on serine import, are resistant to genetic ablation of the pathway. Given these observations, we speculated that PHGDH might be a clinically interesting target in oncology and sought to develop small molecule inhibitors of PHGDH in order to provide tool compounds with which to study the

biology of PHGDH and evaluate the efficacy of inhibiting serine synthesis in cancers.

In order to identify inhibitors of PHGDH an *in vitro* enzymatic assay was developed and libraries of drug-like small molecules were screened. Hit compounds were validated in biochemical assays to determine potency and selectivity for PHGDH. Selected compounds were tested on cells for their ability to inhibit *de novo* serine synthesis and one lead, CBR-5884, was identified. CBR-5884 was selectively toxic to PHGDH amplified or overexpressing cancer cells but had no effect on cells that express little PHGDH. Mechanistically, CBR-5884 was found to be a non-competitive inhibitor that showed a time dependent onset of inhibition and disrupted the oligomerization state of PHGDH. These results provide a proof-of-concept for targeting PHGDH and suggest that inhibiting PHGDH in cancers addicted to serine synthesis is a potentially viable targeted therapy option.

Table of Contents

Acknowledgments

I am grateful to my thesis advisor, Dr. Lewis C. Cantley, for his mentoring, critical feedback, and support over the course of my thesis work. He has always been tremendously generous, patient, and encouraging. In addition, I would like to thank the members of my thesis advisory committee, Drs. Joan Brugge, Sheila Thomas, and Bruce Zetter, for their kindness and valuable suggestions throughout my thesis work.

I thank Dr. Luke L. Lairson and members of his lab, particularly Drs. Natasha Lucki, Reza Beheshti Zavareh, and Justin Anglin, for working with me to develop inhibitors of PHGDH and thoughtful discussions.

I thank members of Cantley lab for helpful discussions and willingness to help. I thank Yuxiang Zheng and Jared Johnson for discussing many aspects of biochemistry and metabolism and their technical assistance. I thank Gina DeNicola and Costas Lyssiotis for mentoring and support particularly at the early stages of my graduate student career. I thank Hyeseok Shim for being a good friend, and for her support and advice throughout graduate school both during classes and in the laboratory.

Finally, I would like to dedicate this thesis to my parents and family for their unstinting love and support. I would not have succeeded without them.

List of Tables

List of Figures

List of Abbreviations

Akt	Protein kinase B
3-PG	D-3-phosphoglycerate
CH_2-THF	5,10-methylenetetrahydrofolate
GSH	Glutathione
IDH	Isocitrate dehydrogenase
KRAS	V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
LDH	Lactate dehydrogenase
NRF2	Nuclear factor (erythroid-derived 2)-like 2
NSCLC	Non-small cell lung cancer
PI3K	Phosphoinositide 3-kinase
PHGDH	3-Phosphoglycerate dehydrogenase
PSAT1	Phosphoserine amino transferase
pSER	Phosphoserine
PSPH	Phosphoserine phosphatase
pPYR	phosphohydroxypyruvate
SER	Serine
SHMT	Serine hydroxymethyltransferase
TCA	Tricarboxylic acid
THF	Tetrahydrofolate
TNBC	Triple negative breast cancer

Chapter 1: Introduction

A Brief Overview of Cancer Cell Metabolism

Since Warburg's pioneering work, cancer cells have been known to reprogram their metabolism in order to promote growth and proliferation [1]. Indeed, metabolic alterations have now been accepted as a hallmark of cancer and many oncogenes including PI3K, Akt, and KRAS reprogram cellular metabolism [2,3,4,5,6]. Although it is now known that defective mitochondria are generally not the cause of aerobic glycolysis, the benefits of enhanced glycolysis remain an active area of research. In part, enhanced glycolysis allows glycolytic intermediates to be funneled into anabolic pathways [7]. In order to divide, a cell needs to duplicate its protein, lipid, and nucleic acid content. Nucleotides, amino acids, lipids, and cellular reducing power in the form of NADPH – critical for reductive biosynthesis and maintaining redox homeostasis – can all be produced by shunting flux away from glycolysis [8]. In addition, cancer cells are avid consumers of glutamine, the most abundant amino acid in plasma [2,9]. As is the case for glucose, glutamine consumption seems to drive anabolic metabolism and is activated by oncogenes such as cMyc [2,10,11]. Glutamine is a nitrogen donor for nucleotide and hexosamine synthesis. Via glutaminolysis, glutamine is a major contributor to anaplerotic flux for the TCA cycle and provides nonessential amino acids [9]. In addition, it is now apparent that despite enhanced glycolysis, many cancer cells still rely on oxidative phosphorylation for ATP production [7,12,13].

Thus, one of the fundamental adaptations of a cancerous cell is to rewire metabolism to promote biosynthesis. Much of the work on cancer metabolism stems from the belief that the idiosyncrasies of tumor metabolism will provide an Achilles' heel for the development of new targeted therapies [9,14,15].

Cellular Synthesis and Catabolism of Serine

Serine is a non-essential amino acid in that it can be synthesized *de novo* from glucose or from glycine, and dietary intake is therefore unnecessary. However, as noted by Eagle in the late 1950s, serine is better characterized as a conditionally essential amino acid, akin to glutamine, in that *de novo* serine synthesis can be insufficient to sustain cell proliferation in culture when serine is omitted from the media [16]. Serine synthesis from glucose proceeds via the phosphoserine pathway in three sequential enzymatic reactions. The first committed step is catalyzed by the enzyme 3-phosphoglycerate dehydrogenase (PHGDH), which oxidizes the glycolytic intermediate 3-phosphoglycerate (3-PG) to 3-phosphohydroxy pyruvate (pPYR) in an NAD^+ dependent mechanism. Subsequently, phosphoserine aminotransferase (PSAT1) transaminates pPYR to yield phosphoserine (pSER) using glutamate as the nitrogen donor. Finally, phosphoserine phosphatase (PSPH) hydrolyzes the pSER phosphate to yield serine (Figure 1.1) [17]. The equilibria of PHGDH and PSAT1 lie towards 3-PG and pPYR, respectively, thus ensuring that when PSPH is feedback inhibited by its product serine, the pathway as a whole is effectively inhibited [18,19]. Serine can be

synthesized from glycine via serine hydroxymethyltransferases (SHMTs) consuming 5,10-methylenetetrahydrofolate (CH_2-THF). Generally, the cytosolic isozyme, SHMT1, favors serine synthesis while the mitochondrial isozyme, SHMT2, catabolizes serine to produce glycine and CH_2-THF [20]. In addition to *de novo* synthesis, serine can be imported from the extracellular space via amino acid transporters such as the commonly expressed ASC system members, ASCT1 (SLC1A4) and ASCT2 (SLC1A5), which mediate the symport of alanine, serine or cysteine with sodium [21].

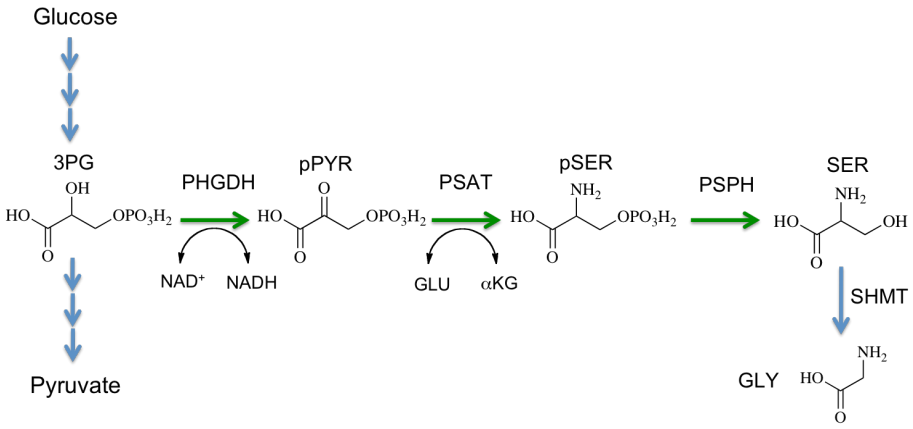


Figure 1.1 The Phosphoserine Pathway. The phosphoserine pathway enables serine synthesis from the glycolytic intermediate 3-phosphoglyerate. Phosphoglycerate dehydrogenase (PHGDH) oxidizes the glycolytic intermediate 3-phosphoglycerate (3-PG) to 3-phosphohydroxypyruvate (pPYR) using NAD^+; phosphoserine amino transferase (PSAT1) transaminates pPYR to phosphoserine (pSER) using glutamate as a nitrogen donor; phosphoserine phosphatase (PSPH) dephosphorylates pSER to yield serine. Serine can be converted to glycine (GLY) via the action of serine hydroxymethyltransferases (SHMTs).

In addition to the SHMT2 reaction, serine can be catabolized via the non-phosphorylated serine pathway or via serine dehydratase. Initially, it was thought that there were two pathways for serine synthesis from glucose – one based on phosphorylated intermediates and the other on unphosphorylated intermediates (Figure 1.2) [18]. The apparent dilemma was resolved when it was understood that the non-phosphorylated pathway was one of catabolism [22]. Akin to the phosphoserine pathway, the non-phosphorylated pathway proceeds in three steps. Initially, serine-pyruvate transaminase (SPT) transfers the serine amino group to pyruvate to yield alanine and hydroxypyruvate. Hydroxypyruvate is then reduced by glycerate dehydrogenase (GDH) to glycerate, which is then phosphorylated by glycerate kinase (GK) to give the glycolytic metabolite 2-phosphoglycerate (Figure 1.2) [18]. Thus, the phosphorylated and non-phosphorylated pathways are antiparallel; the former drawing flux away from glycolysis to produce serine, the latter returning serine derived carbon to glycolysis. Interestingly, SPT also has transaminase activity for glyoxylate and alanine, producing glycine and pyruvate. Germline loss-of-function mutations in SPT lead to hyperoxalurias due to the excessive production of oxalate from glyoxylate [23]. Serine catabolism via serine dehydratase (SDH) proceeds in a single step to yield pyruvate and ammonia (Figure 1.2). While the phosphoserine pathway enzymes and the SHMTs are widely disrupted across tissues, both the non-phosphorylated and SDH reactions occur predominantly in the gluconeogenic tissues of the kidneys and liver [18,23].

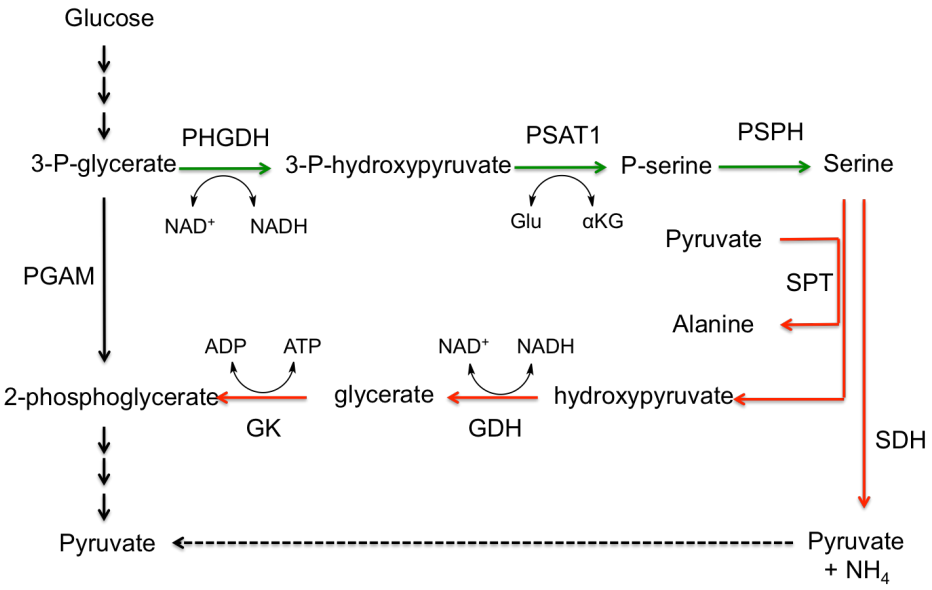


Figure 1.2. Serine synthesis and catabolism. Serine synthesis from glucose via the phosphoserine pathway (green arrows). Serine catabolism shown with red arrows can proceed either through (i) the non-phosphorylated pathway or (ii) serine dehydratase (SDH). The non-phosphorylated pathway consists of three sequential reactions: serine pyruvate transaminase (SPT), glycerate dehydrogenase (GDH), glycerate kinase (GK).

Anabolic Roles of Serine in Cellular Metabolism

Serine is an abundant component of proteins and is required for the synthesis of lipid head groups including sphingolipids and phosphatidylserine, a major component of cellular membranes (Figure 1.3) [24-26]. Indeed, serine deficiency in PHGDH knockout MEFs has been shown to cause abnormal sphingolipid profiles and the accumulation of potentially toxic deoxysphingolipids resulting from the incorporation of alanine instead of serine into lipid head groups [27]. Whereas other metabolites such as betaine (trimethylglycine) can donate one-

carbon units to the folate pool, serine is the major provider of one-carbon units to the folate cycle [28]. The hydroxymethyl group of the serine side chain enters the folate pool via the tetrahydrofolate (THF) dependent action of serine hydroxymethyltransferases (SHMT) yielding glycine and methylenetetrahydrofolate (CH_2-THF) (Figure 1.3). Importantly, the production of glycine and folate one-carbon units enables serine to contribute to nucleotide production. Glycine's nitrogen and two carbon atoms are incorporated directly into purine rings (Figure 1.3). In addition, serine derived folate one carbon units are incorporated into the purine ring and used to methylate dUMP to produce dTMP. Mechanistically, mitochondrial SHMT2 catabolizes serine to generate CH_2-THF. The methylene moiety of CH_2-THF is then sequentially oxidized to methenyl-THF, formyl-THF, and free formate by the mitochondrial folate pathway enzymes (MTHFD2 or MTHFD2L and MTHFD1L)[20]. Eventually, the formate effluxes from the mitochondria to the cytosol where it is a substrate for formate-THF synthetase and is thus reincorporated into the folate pool as formyl-THF. It is the serine derived cytosolic formyl-THF that is directly consumed to make purines; formyl-THF units are incorporated into a purine ring at positions 2 and 8 [28,29]. Thymidylate synthase (TS) consumes CH_2-THF directly to generate dTMP. The above discussion illustrates how serine contributes to the synthesis of all three major macromolecular classes, proteins, lipids, and nucleic acids, and is thus critical for biomass accumulation.

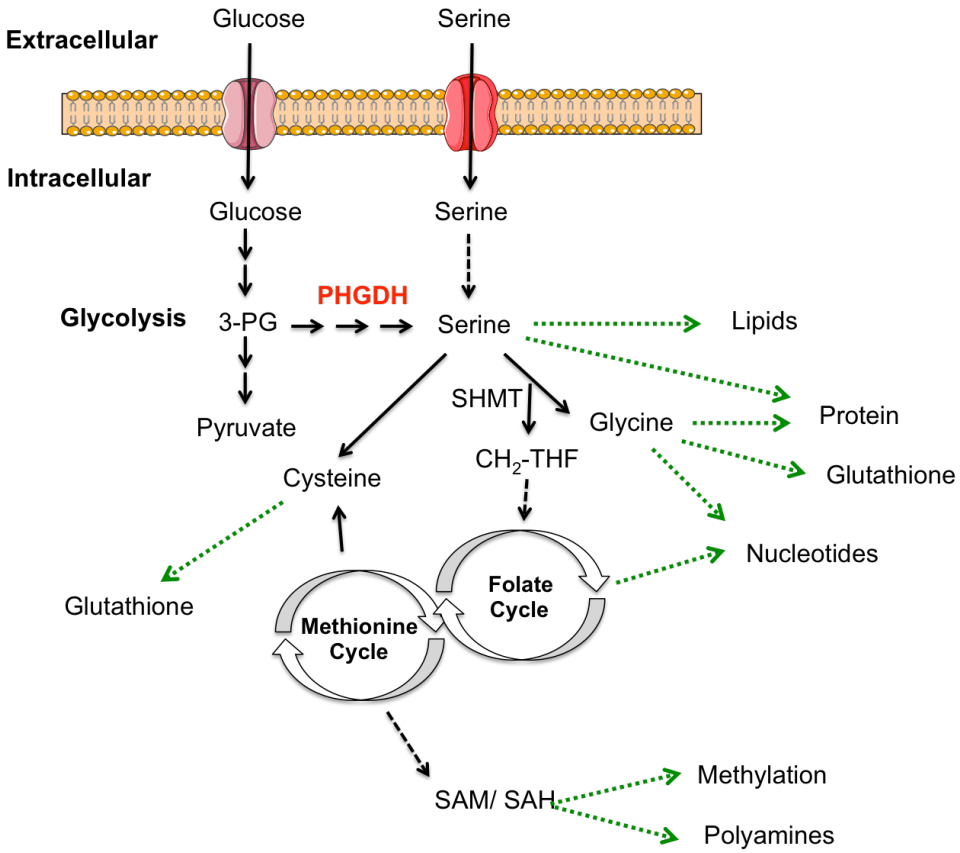


Figure 1.3. Anabolic roles of serine. Serine is either imported from the extracellular space or synthesized *de novo*. Serine is required for protein synthesis and is incorporated into the head groups of certain lipids, including phosphatidylserine and sphingosine. Importantly, serine can be converted to glycine by the action of SHMTs concomitantly charging the folate pool with a one-carbon unit in the form of CH_2-THF. Glycine and one-carbon units are required for nucleotide synthesis. Folate one carbon units can regenerate SAM from homocysteine. SAM is the major methyl donor for methyl transferase reactions and is used to produce polyamines. Via the transsulfuration pathway, serine can generate cysteine; both cysteine and glycine are required for glutathione synthesis.

Beyond supporting biomass production, serine is used to generate a plethora of small-molecule metabolites critical for cellular homeostasis and molecular biology processes. The methionine cycle enables the synthesis of S-adenosylmethionine (SAM), which functions as the universal methyl donor for

cellular methylation reactions including those of DNA, mRNA caps, histones, and other proteins [30,31]. While SAM can be synthesized *de novo* from the conjugation of methionine and the adenosine moiety of ATP via SAM synthetase, the salvage pathway relies on methylation of homocysteine to regenerate methionine. Importantly, the vitamin B_{12} dependent enzyme homocysteine methyltransferase consumes cytosolic methyl-THF to regenerate methionine and liberate THF [30,31]. Thus, serine derived folate one carbon units serve to regenerate SAM for cellular methylation reactions. In addition, SAM can be used for polyamine synthesis. Serine derived formyl-THF is also needed to produce formyl-methionine, which is used as the initiation codon in mitochondrial protein synthesis [32]. Serine can also be use to produce cysteine via the transsulfuration pathway; serine combines with homocysteine to yield cystathionine that is then hydrolyzed to cysteine and homoserine by cystathionase. Glutathione (GSH), which protects cells against oxidative and xenobiotic stress, is an enzymatically-synthesized tripeptide of glutamate, cysteine and glycine. Thus, up to two moles of serine can be consumed per mole of GSH produced [33]. Finally, mitochondrial glycine derived from serine can contribute to heme biosynthesis. Aminolevulinic acid synthase (ALAS), the first enzyme of the heme biosynthesis pathway, catalyzes the condensation reaction of glycine and succinyl-CoA to yield 5-aminolevulinic acid (ALA) [34].

Serine Synthesis in Cancer, a Historical Perspective

Historically, the importance of serine synthesis in cancer was studied by investigating the activities of serine synthesis and catabolism enzymes in homogenates of neoplastic tissues and by tracing the incorporation of radioactive serine, glycine, or glucose into nucleotides [35]. The studies revealed that in various tumor types ranging from rat sarcomas to human colon carcinomas, phosphoserine pathway activity and SHMT activity were increased compared to normal control tissues. In contrast, the activities of serine catabolism enzymes, SDH and SPT, were suppressed [35-37]. Importantly, it was noted that in proliferating hepatoma cells PHGDH and SHMT activity followed the same temporal regulation. Moreover, activity mirrored the incorporation of ^{14}C-serine or ^{14}C-gluocse into nucleotide bases [38,39]. Similar results were obtained with the study of lymphocytes following mitogenic stimulation, indicating that this particular metabolic flow is common among proliferating cells [39]. Combined, these observations led Snell to propose a general model in which the enhanced capacity for serine synthesis in cancer cells is utilized to generate glycine and one-carbon units for purine and thymidine synthesis [40]. Much of the recent work in the field of serine cancer metabolism corroborates Snell's hypothesis [41,42 43].

Serine Synthesis in Cancer, a Modern Perspective

The renewed interest in the role of serine synthesis in cancer emerged following the publication of two studies that sought to identify metabolic pathways important for tumorigenesis. Despite different functional screening methodologies, both studies converged on PHGDH [44,45]. The first study traced the metabolism of uniformly carbon-13 labeled glucose (U-^{13}C-glucose) using two-dimensional nuclear magnetic resonance (NMR)-based metabolomics in 293T cells. The metabolism of glycine from glucose was found to be highly enriched; ^{13}C-glycine enrichment was nearly commensurate with that of ^{13}C-lactate, a hallmark of the Warburg effect [45]. The second study used a pooled shRNA library against metabolic enzymes and transporters to screen for genes necessary for growth of human MCF10DCIS.com cells injected into mouse mammary fat pads [44]. Both studies then combined their analyses with published genomic somatic copy-number alterations (SCNA) from 3,131 cancer samples consisting of tissue specimens and cancer cell lines [46]. *PHGDH* was found to be amplified in a peak region of chromosome 1p12 in 16% of the samples [44,45]. Furthermore, some of the amplifications were focal and contained no known oncogene. Amplification was particularly prevalent in melanomas and breast cancers. Breast cancer and melanoma cell lines harboring *PHGDH* amplifications, or overexpressing PHGDH independent of amplification, were shown to be highly sensitive to PHGDH knockdown, while lines that expressed little to no PHGDH were insensitive [44,45]. These results suggested that PHGDH

overexpression leads to a dependency on, or addiction to, *de novo* serine synthesis. In breast cancer, PHGDH expression has been shown to correlate with both the triple negative and basal subtypes, increased metastatic potential, and shorter patient survival suggesting that targeting *PHGDH* in amplified cancers may be therapeutically advantageous [44,45,47].

The pro-tumorigenic role of serine synthesis is not limited to melanomas and breast cancer. Indeed, profiling of *de novo* serine synthesis in a large panel of non-small cell lung cancer (NSCLC) cell lines showed that high pathway activity was independent of amplification and instead regulated by the transcription factor NRF2, an oncogene frequently mutated in NSCLC [41]. While serine synthesis activity in NSCLC lines was heterogeneous, lines with high pathway activity were sensitive to PHGDH knockdown. Furthermore, expression of serine synthesis genes and SHMT2 correlated with poor patient survival. *De novo* serine synthesis was critical for NSCLC cells to maintain glutathione, purine nucleotide, and dTMP levels consistent with the known anabolic functions of serine [41].

In addition to NRF2, the histone H3 lysine methyltransferase G9A and the transcription factors c-Myc and ATF4 have been shown to positively regulate serine biosynthesis gene expression suggesting additional mechanisms for amplification independent overexpression in cancer [41,48-50]. Notably, serine biosynthesis genes have been shown to be regulated by ATF4 in a number of different cell types suggesting that it is a general mechanism for inducing serine biosynthesis consistent with the established role of ATF4 in the amino acid

depletion response [50-52] [41]. Interestingly, serine is involved in a regulatory feedback loop with pyruvate kinase M2 (PKM2). Serine binds to PKM2 allosterically to activate it and reduce the funneling of 3-PG into serine synthesis [53,54]. Conversely, excessive PKM2 activation can induce serine auxotrophy and ATF4 induction suggesting that pyruvate kinase M1 expressing cells may have a more limited serine synthesis capability [50,53,55].

NADPH powers reductive biosynthesis and cellular antioxidant defenses, and is therefore critically important for cellular proliferation and maintenance of redox homeostasis[,5,33,56]. Traditionally, the oxidative pentose phosphate pathway is thought to be the major provider of NADPH with malic enzyme sometimes being important [11,57]. In NSCLC cell lines, however, PHGDH knockdown was shown to deplete cellular NADPH levels specifically in PHGDH overexpressing cells suggesting an additional means by which PHGDH can promote tumor growth [41]. Although the exact mechanism by which serine enables NADPH production is not yet clear, recent evidence suggests that oxidation of serine derived methylene units via the folate cycle produces significant amounts of NAPDH [58,59,60]. Interestingly, using ectopic mutant IDH expression as a compartment specific NADPH reporter suggested that the oxidation of serine derived CH_2-THF via the folate cycle contributes to mitochondrial NADPH production rather than cytosolic NADPH production [59]. Serine flux to mitochondrial folate metabolism has been suggested to be important in mitigating mitochondrial oxidative stress under hypoxia since HIF1 induces expression of SHMT2 in a Myc dependent manner to prevent accumulation of reactive oxygen

species and maintain GSH levels [60]. Thus, the recent realization that NADPH is also a downstream product of serine expands the scope of the anabolic processes dependent on serine and suggests another way in which PHGDH overexpression can be pro-tumorigenic. Indeed, the importance of ROS and NADPH metabolism for tumorigenesis is well established [56,61].

While the anabolic functions of serine suggest that increased serine synthesis might enable tumorigenesis by passively supporting biomass production, some evidence suggests that PHGDH may be playing a more active role in tumorigenesis. MCF-10A cells, an immortalized but non-transformed human mammary epithelial cell line, grown in Matrigel reconstituted basement membrane form polarized, growth-arrested, hollow acinus-like spheroids that mimic *in vivo* glandular architecture. Provocatively, overexpressing PHGDH in MCF-10A cells disrupted acini polarity and resulted in a luminal filling phenotype indicating a failure of extracellular matrix detached cells to undergo apoptosis, which is suggestive of transformation [45]. In addition, overexpressing PSAT1 and PSPH in 3T3 cells allowed 3T3 cells to form tumors when injected into mice [62]. There is precedence for metabolic enzymes playing an active role in tumorigenesis. The TCA cycle enzymes fumarate hydratase (FH) and succinate dehydrogenase (SDH) follow a classical familial tumor suppressor paradigm with loss of heterozygosity (LOH) [63]. Typical of an oncogene, only a handful of IDH1/2 mutations are known and confer the same neomorphic activity [64]. Mechanistically, FH, SDH, and IDH mutations are thought to induce epigenetic deregulation conducive to transformation [63,65]. How PHGDH could have

pleiotropic effects sufficient to induce transformation is still difficult to imagine. Thus, whether *PHGDH* is a bona-fide oncogene remains to be determined, but these results combined with the occurrence of focal *PHGDH* amplifications and oncogene induced PHGDH overexpression suggests that, at a minimum, serine synthesis can play an important role in supporting tumorigenesis.

While it would be tempting to conclude that enhanced serine synthesis is a general hallmark of cancer akin to the Warburg effect, this is likely not the case. Nearly a third of the cell lines profiled in the large NSCLC panel had low serine biosynthetic activity [41]. Given that many cancer cell lines require serine import, and that the vast majority of the NCI-60 panel consume more serine from the media than they release, it is more likely that serine import is a common feature across cancer cell lines in culture and that only a subset of lines are addicted to *de novo* serine synthesis [43 66 44,47].

Potential Mechanisms for Addiction to *De Novo* Serine Synthesis

The mechanism through which PHGDH overexpressing cells become addicted to serine synthesis remains unclear. It is surprising that PHGDH overexpressing cells are sensitive to PHGDH knockdown given the presence of serine in tissue culture media [17]. There are at least three models that could account for the failure of extra-cellular serine to rescue PHGDH knockdown. The first model resolves to a kinetic issue where the rate of serine uptake via amino acid transporters is insufficient to compensate for loss of serine biosynthetic

activity. However, this seems unlikely because in a *PHGDH* amplified breast cancer line, MDA-MB-468, less than 20% of intracellular serine is derived from *de novo* synthesis [67].

The second model is that cancer cells become addicted to the *de novo* serine synthesis pathway not because it provides serine, but because flux through the phosphoserine pathway provides another metabolite. For example, it has been suggested that 50% of the α-ketoglutarate pool in MDA-MB-468 cells comes from the PSAT1 reaction and that this is an important anaplerotic flux for the TCA cycle [44]. This is surprising, however, because α-ketoglutarate is usually thought to be produced by transaminases and glutaminolysis [9 68]. Moreover, our attempts to rescue PHGDH knockdown with cell permeable α-ketoglutarate derivatives have been unsuccessful (unpublished). Interestingly, as an error reaction, PHGDH has been reported to produce the oncometabolite D-2-hydroxyglutarate from the NADH dependent reduction of α-ketoglutarate [69]. Again, however, this seems an unlikely mechanism because in *PHGDH* amplified MDA-MB-468 cells, PHGDH dependent D-2-hydroxyglutarate production is only on the order of 40μM whereas millimolar levels of D-2-hydroxyglutarate are observed in patients and cell lines with IDH1/2 mutations [64,67,70]. Moreover, in the *PHGDH* amplified cell line Hcc70, PHGDH knockdown inhibits growth but does not reduce basal D-2-hydroxyglutarate levels [67].

A third model for the necessity of *de novo* serine synthesis despite the presence of serine in tissue culture media is that the extra-cellular and intra-cellular serine pools are somehow distinct. Mechanistically, this could be

achieved via the formation of protein complexes that funnel *de novo* serine synthesis flux into specific downstream pathways or specific organelles (eg. mitochondria versus cytoplasm). Both complex mediated funneling and compartment specific metabolism are well established. For example, the purinosome is a complex that groups all six enzymes needed to synthesize inosine monophosphate (IMP), the purine precursor to AMP and GMP, from phosphoribosyl pyrophosphate (PRPP) [71]. Interestingly, a nuclear complex for TMP synthesis composed of thymidylate synthase, SHMT1 or SHMT2α, and dihydrofolate reductase has been reported [72,73]. While the association of *de novo* serine synthesis with the thymidylate synthase complex would fit well with the importance of serine synthesis for nucleotide production, nuclear localization of PHGDH has not been observed; furthermore, one would presume that cytosolic serine, imported from the extra-cellular space or synthesized from glucose, could simply diffuse to the nucleus [74].

While the mechanism for PHGDH addiction remains unclear there is an increasing body of work that links serine to the production of NADPH via the folate cycle suggesting that avenues of research pursuing the contribution of *de novo* serine synthesis to NADPH and folate metabolism may be a fruitful path of inquiry [58,60,75]. Indeed, beyond the ability of PHGDH overexpressing cells to grow in serine free media, the only phenotype that has been shown to distinguish PHGDH overexpressing cells from non-overexpressing cells is a decrease in NADPH following PHGDH knockdown in the former [41].

Chapter 2: **Identification of a Small Molecule Inhibitor PHGDH**

Abstract:

Cancer cells reprogram their metabolism to promote growth and proliferation. The genetic evidence pointing to the importance of the amino acid serine in tumorigenesis is striking. The gene encoding the enzyme 3-phosphoglycerate dehydrogenase (PHGDH), which catalyzes the first committed step of serine biosynthesis, is overexpressed in tumors and cancer cell lines via focal amplification and NRF2-mediated upregulation. PHGDH overexpressing cells are exquisitely sensitive to genetic ablation of the pathway. Here, we report the discovery of a novel and selective small molecule inhibitor of PHGDH, CBR-5884, identified by screening a library of 800,000 drug-like compounds. CBR-5884 inhibited *de novo* serine synthesis in cancer cells and was selectively toxic to cancer cell lines with high serine biosynthetic activity. Biochemical characterization of the inhibitor revealed that it was a non-competitive inhibitor that showed a time dependent onset of inhibition and disrupted the oligomerization state of PHGDH. The identification of a small molecule inhibitor of PHGDH not only enables thorough preclinical evaluation of PHGDH as a target in cancers, but also provides a tool with which to study serine metabolism.

Introduction:

Serine is required for a plethora of anabolic processes including the synthesis of lipids, protein and nucleic acids [24-26]. At the cellular level, serine can be imported from the extracellular space via amino acid transporters [21,76]. Alternatively, serine can be synthesized from glucose via the phosphoserine pathway [18]. *De novo* synthesis proceeds from the glycolytic intermediate 3-phosphoglycerate (3-PG) via three sequential enzymatic reactions (Figure 1.1), the first of which is catalyzed by the NAD^+ dependent enzyme 3-phosphoglycerate dehydrogenase (PHGDH) [17]. For decades, it has been known that cancer cells have enhanced serine synthesis, which contributes to nucleotide synthesis [38,40]. Recently, focal amplifications of the gene encoding *PHGDH* have been reported, particularly in breast cancers and melanomas [44-46]. Additionally, *KEAP1* and *NRF2* mutant non-small cell lung cancers (NSCLC) overexpress *PHGDH* [41]. Proliferation of *PHGDH* amplified cancer cell lines, and other lines that overexpress PHGDH without amplification, is inhibited by PHGDH knockdown. In contrast, lines that express little PHGDH are resistant to shRNA mediated ablation of the pathway presumably because serine import suffices [44,45]. A detailed mechanistic understanding of why some cancer cells are addicted to serine synthesis despite the availability of extracellular serine for import remains unclear. Interestingly, in triple negative breast cancer (TNBC) and NSCLC, *PHGDH* amplification and overexpression are associated with more

aggressive disease [41,44,45,47]. Thus, PHGDH inhibitors as a targeted therapy for these tumor types represent an exciting clinical opportunity.

The studies herein detail our efforts in identifying small molecule inhibitors of PHGDH. We reasoned that a PHGDH inhibitor would have the benefits of not only providing a tool compound with which to study the biology of serine synthesis, but also enable thorough preclinical evaluation of PHGDH as a target in cancers. We developed an *in vitro* assay for PHGDH amenable to high throughput screening and screened a library of 800,000 small molecules for inhibitors of PHGDH. A cell-based assay for serine synthesis was used to identify a lead, CBR-5884, that was active in cells. CBR-5884 selectively inhibited the proliferation of melanoma and breast cancer lines that have a high propensity for serine synthesis, but had no effect on lines that rely on extra-cellular serine uptake. Mechanistically, CBR-5884 was found to be a noncompetitive inhibitor, showed a time dependent onset of inhibition, and disrupted the oligomerization state of PHGDH.

Results:

Developing an in vitro assay for PHGDH

In order to screen for inhibitors of PHGDH an *in vitro* assay for PHGDH needed to be developed. Initially, recombinantly expressed PHGDH and the downstream enzyme PSAT1 were incubated with their substrates NAD^+, 3PG, and glutamate, and PHGDH activity was monitored by following the intrinsic fluorescence of the NADH produced (Figure 2.1 and 2.2). PSAT1 was included in the assay to prevent product inhibition of PHGDH. Given that the thermodynamic equilibrium for the PHGDH reaction lies heavily towards 3-PG [18], high concentrations of glutamate (30mM) were included in the assay to drive the PHGDH reaction forward.

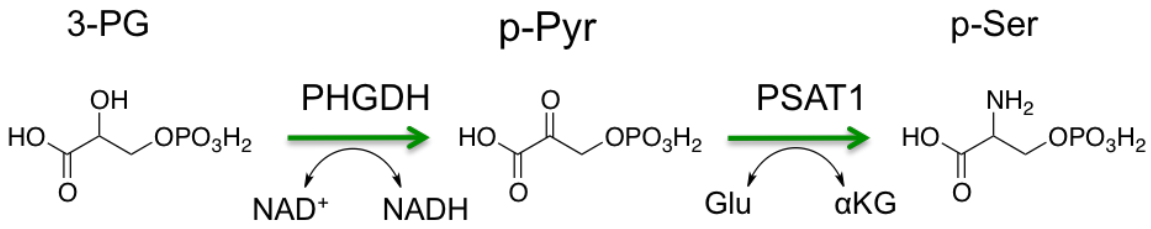


Figure 2.1. Schematic for the *in vitro* PHGDH enzymatic assay.
Recombinantly expressed PHGDH and PSAT1 are incubated with their substrates NAD^+ and glutamate. The reaction is initiated by the addition of 3-PG and NADH fluorescence (340/460nm Ex/Em) is monitored over time. 3-PG: 3-phosphoglycerate; NAD^+: nicotinamide adenine dinucleotide; αKG: alpha-ketoglutarate; p-Pyr: phosphohydroxypyruvate; p-Ser: phosphoserine.

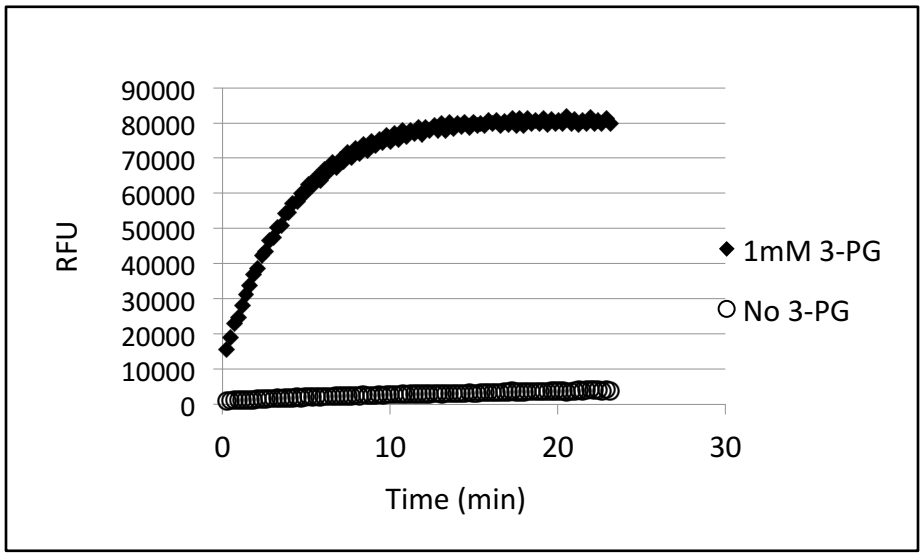


Figure 2.2. Progress curve for the PHGDH reaction. PHGDH activity is monitored by following the increase in NADH fluorescence (340/460nm Ex/Em) using the *in vitro* assay developed. 3-PG: 3-phosphoglycerate. RFU: relative fluorescence units.

An understanding of the Michaelis-Menten constants (K_m) of each substrate is important in order to design the screening assay conditions. Competitive inhibitors can only bind to the free enzyme (no substrate bound) and therefore the ability to identify them decreases with an increasing [S] / K_m ratio. In contrast, uncompetitive inhibitors bind to the enzyme substrate complex (ES). Thus, as the [S] / K_m ratio increases, the likelihood of an ES complex increases, which facilitates uncompetitive inhibitor binding. It has been shown that screening at the substrate K_m provides the best compromise for the assay to detect both competitive and non-competitive inhibitors [77]. The K_m constants for PHGDH were determined by titrating each substrate, NAD+ or 3-PG, one at a time, while

holding the remaining substrate at a fixed concentration and in excess such that the PHGDH reaction became pseudo first order with respect to the substrate being titrated. The K_m was determined to be 60± 10μM with respect to NAD^+ and 1.4± 0.6μM with respect to 3-PG (Figure 2.3). The K_m for NAD^+ was in good agreement with the literature value reported for chicken liver PHGDH (60μM) [78].

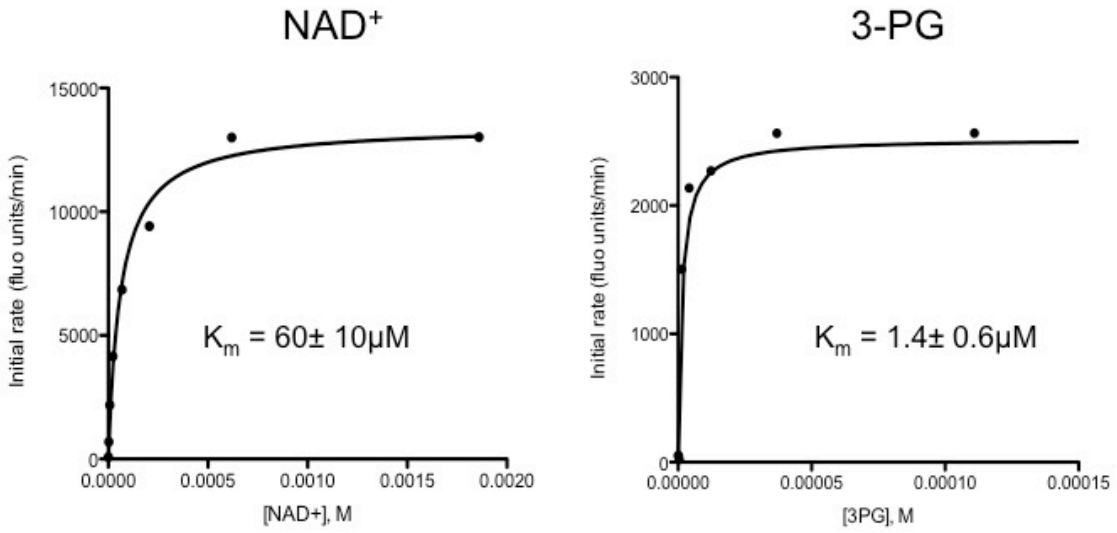


Figure 2.3. Michaelis-Menten plots for each PHGDH substrate. One at a time, each PHGDH substrate was titrated while holding the other substrate constant and in excess. The initial reaction rates were determined and plotted as a function of substrate concentration and the data were fit using non-linear regression to a Michaelis-Menten model. 3-PG: 3-phosphoglycerate; NAD^+: nicotinamide adenine dinucleotide. K_m values are the average of three replicates. Error is given as ± 1SD.

Inhibitors effectively decrease the fraction of active enzyme. Thus, in order to identify inhibitors, the initial rate of the enzymatic reaction needs to be linearly dependent on the PHGHD concentration otherwise binding of an inhibitor to PHGDH may not translate into a detectable decrease in signal. Moreover, given that the PHGDH reaction is coupled to the PSAT1 reaction, the rate of the PHGDH reaction may be limited by the rate of the downstream PSAT1 reaction. Running the screen under conditions in which the PSAT1 reaction is rate limiting would favor the identification of inhibitors targeting PSAT1 instead of PHGDH. Titrating PHGDH in the presence of NAD^+ at 1 K_m and 3-PG at 100 K_m in the presence of 1850nM PSAT1, showed that the reaction was linearly dependent on PHGDH from 20nM to 150nM PHGDH (Figure 2.4). The K_m for 3-PG (1.4± 0.6μM) was so low that it was impossible to efficiently detect NADH fluorescence when the assay was performed at the 3-PG K_m and a concentration200-fold the K_m had to be used (data not shown). Thus, 70nM PHGDH was selected as the optimal concentration for the given assay conditions.

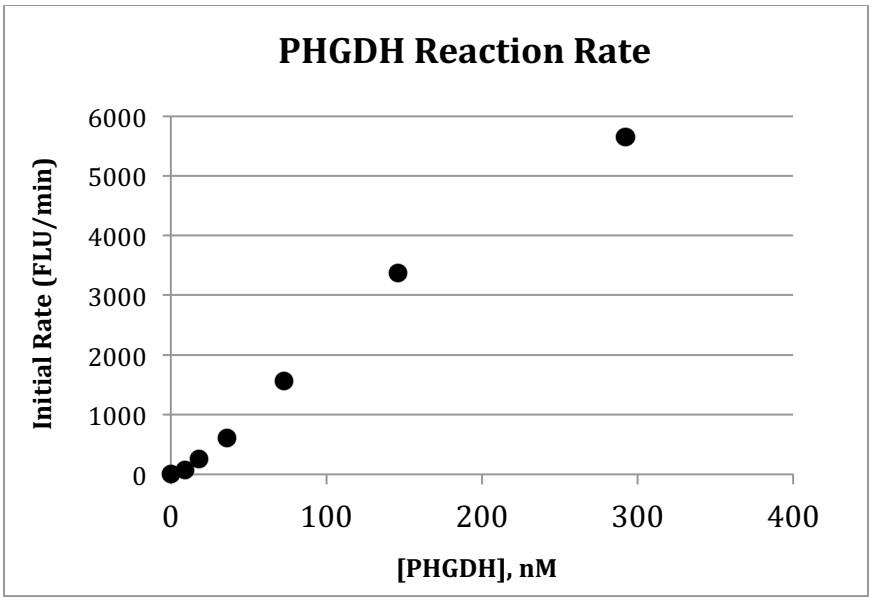


Figure 2.4. Linear dependence of the initial reaction rate on PHGDH concentration. Two-fold serial dilution of PHGDH in the presence of 1850nM PSAT1, 200 times the 3-PG K_m, and 1 times the NAD^+ K_m. Initial rates are plotted as a function of PHGDH concentration. The reaction rate is linearly dependent on PHGDH concentration from 18nM to 150nM. A representative example from 3 independent experiments is shown.

While assays measuring NADH fluorescence are amenable to high-throughput chemical screening, they are not ideal. A significant number of compounds in chemical libraries fluoresce at 460nm and would therefore interfere with the assay readout. In order to improve signal to noise ratio, since NADH is a relatively week fluorophore, and decrease the likelihood of compound interference, the reaction was coupled to diaphorase (Figure 2.5). Diaphorase can use the NADH produced upon PHGDH turnover to reduce resazurin to a potent fluorophore, resorufin. When the *in vitro* PHGDH assay was coupled to diaphorase, the progress curve showed a lag phase of approximately 10 minutes that was not seen in the PHGDH reaction progress curve suggesting that the

diaphorase reaction was slower than the PHGDH reaction (compare Figure 2.2 to Figure 2.6). Titrating resazurin in the diaphorase coupled PHGDH assay revealed that diaphorase was subject to substrate inhibition at high resazurin concentrations indicating the cause of the lag phase in the coupled reaction (Figure 2.7). At 0.1mM resazurin, the kinetics of the diaphorase coupled PHGDH reaction progress curved mirrored that of the PHGDH reaction monitored via NADH fluorescence. Thus, 0.1mM resazurin was selected for future assays. However, it was noted that even in the no diaphorase control, the resorufin signal increased overtime (Figure 2.7). This was problematic for the high throughput screen as it decreased the signal to noise ratio of the assay over time and would make it difficult to compare results across different compound screening assay plates. It was noted that when the diaphorase-coupled reaction was quenched with 50% methanol, the diaphorase independent reduction of resazurin still proceeded at the same rate indicating that the process was non-enzymatic (data not shown). Indeed, the dithiothreitol (DTT) included in the assay buffer was found to spontaneously reduce resazurin. Omitting DTT from the assay buffer was sufficient to abrogate the time dependent increase in background fluorescence (Supplementary Figure 2.1). Thus, DTT was removed from the assay buffer for the high throughput screen.

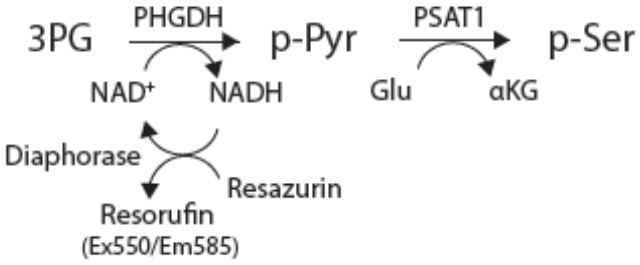


Figure 2.5. *In vitro* PHGDH assay coupled to diaphorase. Diaphorase couples the NADH produced upon PHGDH turnover to the reduction of resazurin to fluorescent resorufin concomitantly recycling the NAD^+. Resorufin fluorescence is a proxy for PHGDH activity.

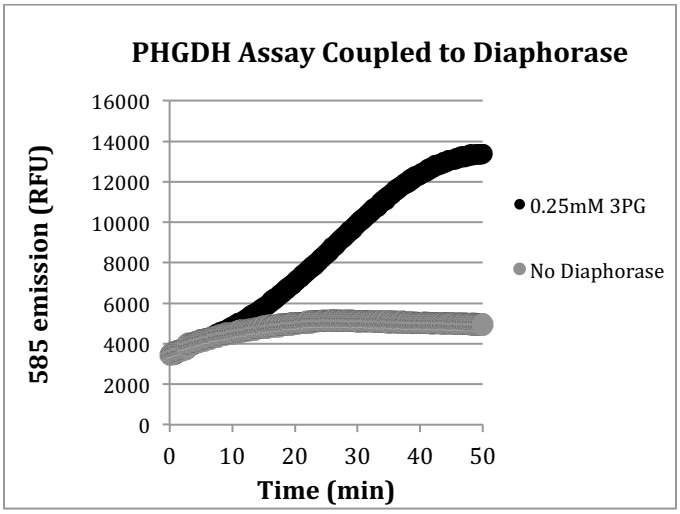


Figure 2.6. Progress curve for PHGDH reaction coupled to diaphorase. Coupling the PHGDH assay to diaphorase results in a lag phase of ~10min to the progress curve. Resorufin fluorescence (550/585nm Ex/Em) was monitored over time. A no diaphorase control was included to ensure that the signal was specific to the diaphorase reaction.

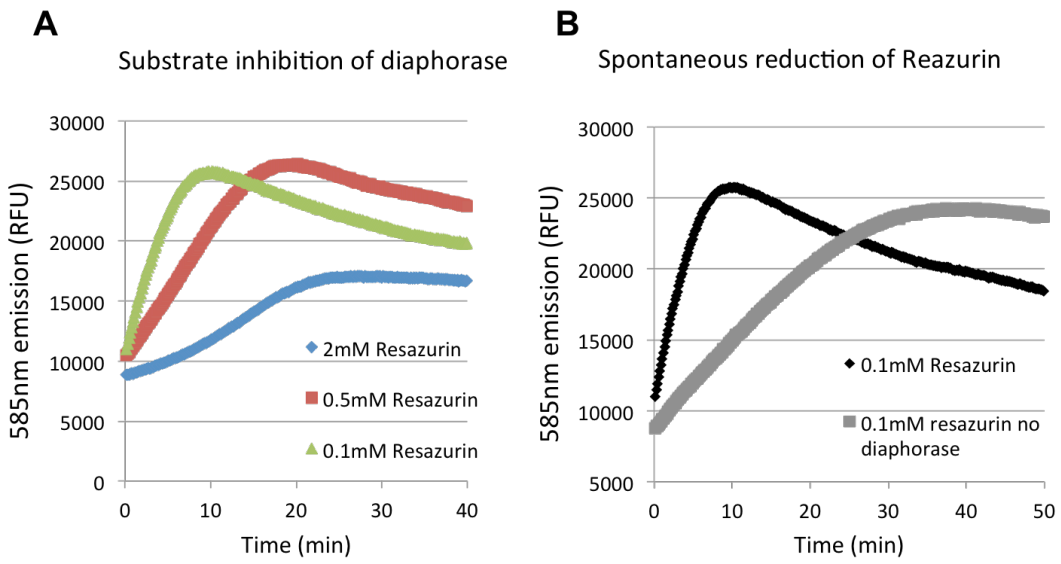


Figure 2.7. Substrate inhibition of diaphorase and spontaneous reduction of resazurin. (A) Titrating the resazurin concentration used for the diaphorase coupled PHGDH assay revealed that the lag phase was due to substrate mediated inhibition of diaphorase. (B) Omitting diaphorase from the coupled reactions reveals resazurin is spontaneously reduced leading to an increasing background signal over the course of the assay.

Given that NAD^+ dependent dehydrogenases, such as PHGDH, are a large family of enzymes, we were concerned that screening at the K_m of NAD^+ would identify compounds that bound competitively to the NAD^+ pocket and that such compounds would have significant off target activity on other dehydrogenases. Thus, after verifying that the reaction rate was still linearly dependent on the PHGDH concentration at increased NAD^+ concentrations, we decided to increase the NAD^+ concentration used in the screening assay to 8mM. Bromopyruvate (BrPYR) has been reported to be a nonspecific inhibitor of PHGDH and was thus selected as a positive control for inhibition in the high throughput screening assay [78]. A progress curve for the finalized diaphorase

coupled PHGDH assay miniaturized to 384-well plate format is presented (Figure 2.8). The assay was approximately linear over the first 30 minutes. The high throughput assay was to be performed in a single point format. Thus, we opted for stopping the reaction with the addition of 50% methanol following a 20-minute reaction corresponding to a signal to background ratio of 3 (Figure 2.8). Resorufin fluorescent signal was shown to be stable for at least five hours after stopping the reaction (data not shown). In a pilot experiment mimicking the screening protocol, the Z-factor for the screening assay was determined to be 0.66 indicating a high quality assay (Figure 2.8) [79]. Thus, we were confident the assay developed could be used for a high-throughput screening campaign. The final components of the optimized assay are presented in Table 1.

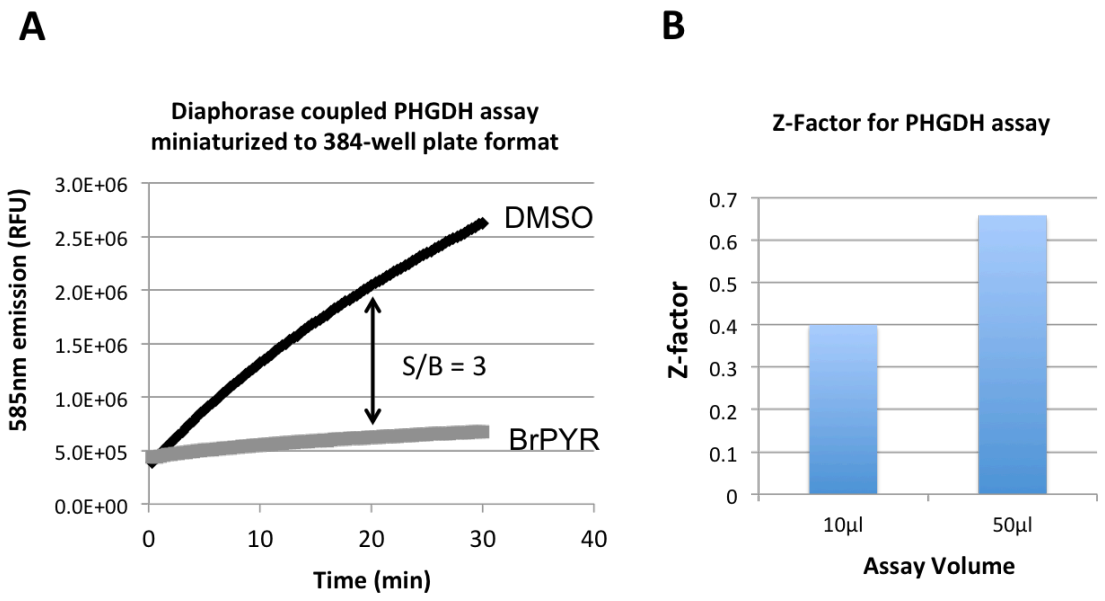


Figure 2.8. PHGDH assay miniaturization and Z-factor. (A) Progress curve for diaphorase coupled PHGDH assay in 384-well plate format with bromopyruvate (BrPYR) treatment (0.5mM) or vehicle control treatment (DMSO). At a 20 minute endpoint, the assay has a signal to background ratio (S/B) of three. (B) Z-factor for the 384 well plate format assay. Three assay plates of negative control wells (vehicle treated) and inhibitor positive control wells (bromopyruvate treated) were screened at a 10µL or 50µL final reaction volume using the automated high throughput screening protocol. The Z-factor for each assay volume was calculated from the mean and standard deviations of the positive and negative control wells.

Table 1. Summary of PHGDH assay conditions for screening.

Component	Amount per assay	unit
PHGDH	0.005	µg/µL
PSAT1	0.02	µg/µL
Diaphorase	0.001	U/ µL
3-phosphoglycerate	0.25	mM
Glutamic acid	30	mM
NAD+	8.3	mM
Resazurin	0.1	mM
Buffer	120 mM Tris, 2.4 mM EDTA	pH 8.5

Screening for small molecule inhibitors of PHGDH

The diaphorase coupled PHGDH assay was further miniaturized to 1536 well plate format with a Z-factor of >0.75 indicating a high quality assay. A library of 800,000 structurally diverse compounds with drug-like properties and known biological actives (kinase inhibitors, epigenetic modifiers, natural products) was screened in single point format at 13μM (Figure 2.9). Setting a hit threshold Z score of -3, corresponding to at least 50% PHGDH inhibition, gave a 0.5% hit rate yielding 3,906 hits. Putative hits were re-assayed in triplicate and counter screened against diaphorase to rule out false positives targeting diaphorase. The counter screen eliminated 3,498 compounds giving 408 PHGDH inhibitors (Figure 2.10).

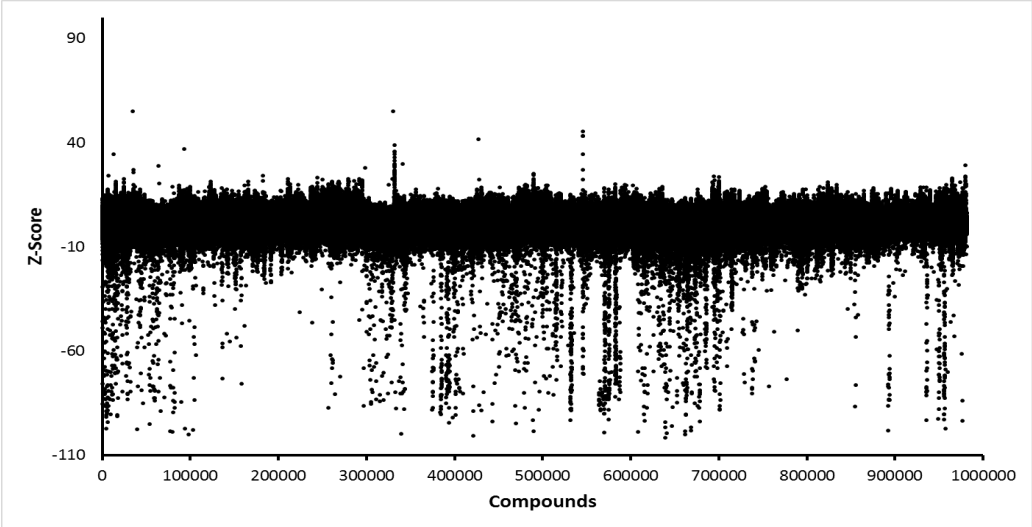


Figure 2.9. Z-score plot for a 800,000 compound library screened against PHGDH. Each point on the graph represents the Z-score of a single compound. A negative Z-score indicates inhibition and a positive one activation. Setting a threshold Z score of -3, corresponding to at least 50% PHGDH inhibition, gave a 0.5% hit rate yielding 3,906 hits.

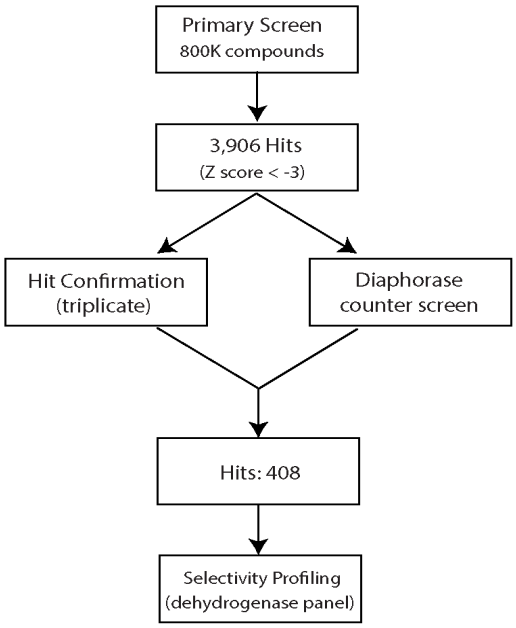


Figure 2.10. Screen triaging strategy. Setting a Z-score threshold of -3 gave 3,906 putative hits. After counter-screening against diaphorase to rule out false positives and confirming activity against PHGDH, 408 compounds remained. Selected compounds were profiled against a panel of metabolic $NAD(P)^+$ dehydrogenases to ascertain selectivity for PHGDH.

A triaging strategy based on hit potency and selectivity was designed (Figure 2.10). We reasoned that inhibitors specific to PHGDH would minimize general cellular toxicity compared to compounds that hit a variety of dehydrogenases. Thus, for commercially available hits (45 compounds) half maximal inhibitory concentrations (IC_{50}) were determined against a panel of $NAD(P)^+$ dependent dehydrogenases that included PHGDH, isocitrate dehydrogenase (IDH1), malate dehydrogenase (MDH1), and 3α-hydroxysteroid dehydrogenase (3α-HSD). Compounds at least 4-fold more selective for PHGDH were progressed for further analysis. Based on this triaging, seven of the most

potent PHGDH inhibitors were selected as lead compounds for evaluation in cell-based assays; selected structures are shown (Figure 2.11, Table 2). A number of these compounds are likely to target sulfhydryl groups and may therefore react with a PHGDH cysteine residue. For example, both CBR-5807 and CBR-6936 contain sulfhydryl reactive disulfide centers that are unstable. Interestingly, CBR-5807 (disulfiram) is an approved drug dosed in humans to treat alcoholism and known to inhibit aldehyde dehydrogenase by reacting with sulfhydryl groups [80].

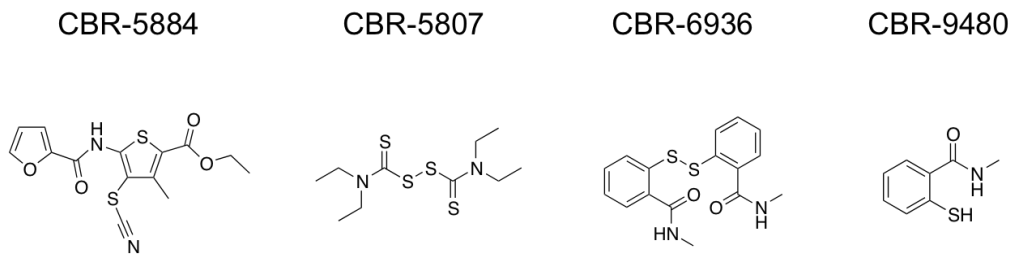


Figure 2.11. Sample structures of PHGDH inhibitors identified. CBR-5807 is an approved drug, disulfiram, that inhibits aldehyde dehydrogenase by targeting sulfhydryl groups.

Table 2. *In vitro* selectivity profiling of PHGDH inhibitors. For each compound, enzyme IC_{50} was normalized to that of PHGDH to indicate relative compound selectivity.

	PHGDH	3a-HSD	Diaphorase	IDH1	MDH1
CBR-5807	1	>14	>14	>14	>14
CBR-6903	1	>4	>4	>4	>4
CBR-6936	1	>125	>125	61	>125
CBR-5884	1	>14	>14	>14	>14
CBR-7406	1	11	>55	4	40
CBR-9480	1	>26	>26	>26	>26
CBR-7439	1	>46	>46	>46	>46

CBR-5884 inhibits serine synthesis in cells

We determined whether any of our seven leads inhibited serine synthesis in cancer cells. To do so, we turned to gas chromatography-mass-spectrometry (GCMS) with uniformly carbon-13 labeled glucose ($^{13}C_6$-glucose) tracing. Given the isotopic enrichment of serine, it is possible to decouple newly synthesized serine from extracellular serine or serine that was synthesized prior to tracer addition. Newly synthesized serine has a mass-shift of 3 (M+3) due to the incorporation of glucose-derived ^{13}C via 3-PG. We first investigated the kinetics of phosphoserine pathway labeling. While 3-PG labeling quickly saturated at ~90%, serine labeling plateaued around 6h with ~65% of the serine pool being ^{13}C labeled (Figure 2.12). Consistent with prior reports, both M+0 and M+3 phosphohydroxypyruvate was undetectable by GCMS possibly due to low abundance and/or poor stability (data not shown) [69]. The plateau phase in serine labeling likely reflects exchange between intra- and extra-cellular serine pools [81].

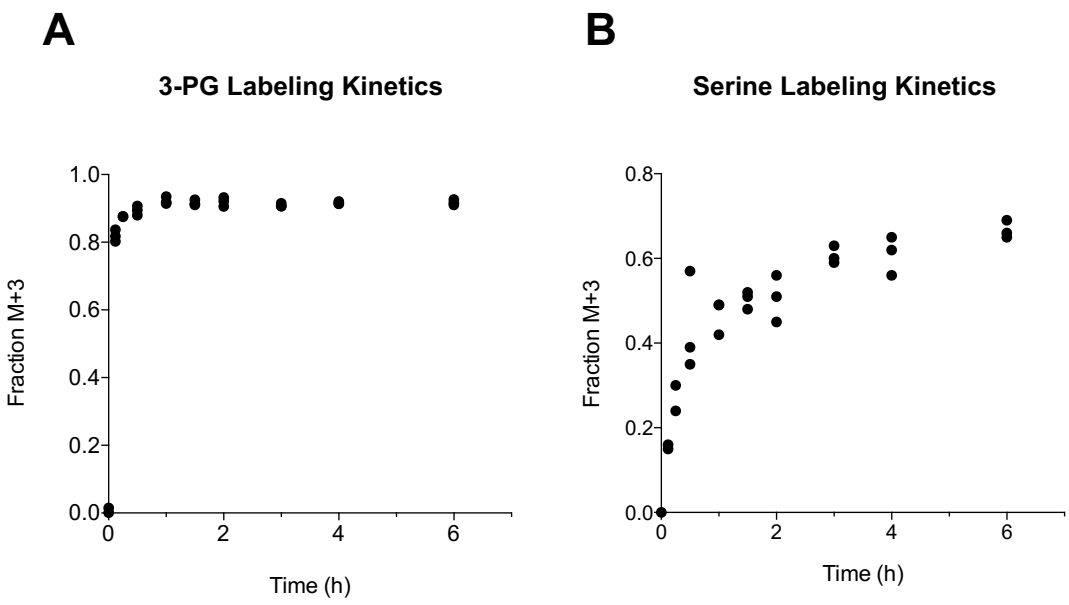


Figure 2.12. Kinetics of phosphoserine pathway labeling. $^{13}C_6$-glucose was added to Carney cells grown in serine free media at time t = 0 and cells were labeled for the indicated time points. Polar metabolites were extracted and analyzed by GCMS. (A) Glucose-derived 3-PG (M+3) and (B) glucose-derived serine (M+3) relative to total 3-PG or serine levels are plotted. Each point indicates an independent experiment.

With an understanding of serine labeling kinetics, we designed a $^{13}C_6$-glucose tracing assay to acutely interrogate the effects of compounds on serine synthesis (Figure 2.13). Assaying serine synthesis with a 3-hour compound treatment was preferred to longer treatments to guard against false positives that decrease serine labeling by an indirect effect, such as generally compromising cellular viability. Among our lead compounds, CBR-5884 was able to decrease *de novo* serine synthesis by 30%; the remaining compounds had little effect (Figure 2.13).

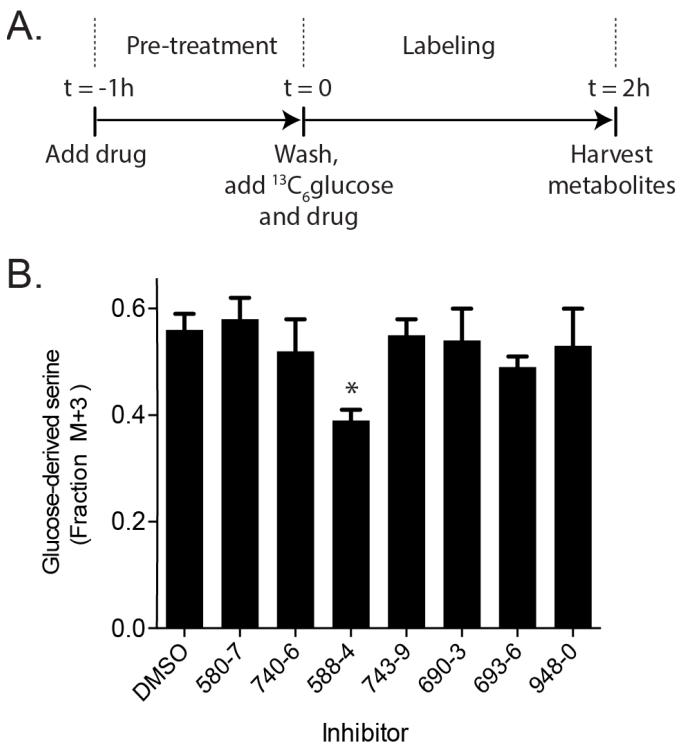


Figure 2.13. CBR-5884 inhibits serine synthesis in cells. (A) Acute inhibitor treatment assay schematic. Carney cells are pre-treated with drug at 30μM for 1h prior to initiating $^{13}C_6$-glucose labeling for 2h still in the presence of drug. Polar metabolites are harvested and analyzed by GC-MS. (B) The ability of the 7 lead PHGDH inhibitors to block serine synthesis was assayed as in (A). Only CBR-5884 decreased glucose derived serine labeling. M+3 serine levels relative to total serine levels are plotted. Asterisks indicate significant differences vs. DMSO treatment ($p < 0.05$, t-test, $n \geq 3$). Error is given as ± 1SD.

The dose at which CBR-5884 had an effect on serine labeling was consistent with the *in vitro* biochemical IC_{50} of 33 ±12μM for PHGDH (Figure 2.14). At such concentrations, CBR-5884 had no effect on two other NAD^+ dependent dehydrogenases, lactate dehydrogenase (LDH) and MDH1 (Figure 2.14, Figure 2.17). Importantly, under the acute treatment time period used in the labeling assays, CBR-5884 was not generally cytotoxic at concentrations up to 40μM as determined by two independent cellular viability assays (Figure 2.15).

Hence, decreases in serine labeling are a direct effect of CBR-5884 mediated PHGDH inhibition and not a consequence of general cellular toxicity.

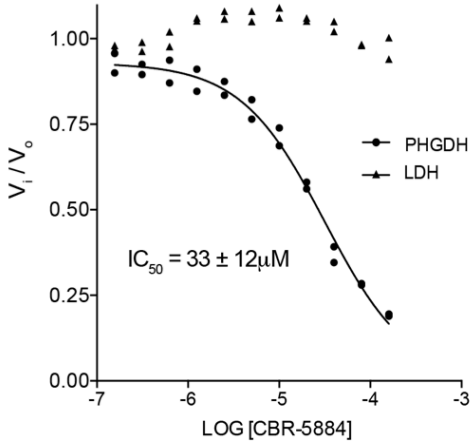


Figure 2.14. *In vitro* IC_{50} of CBR-5884 on PHGDH and LDH. (A) *In vitro* IC_{50} assays for PHGDH and lactate dehydrogenase, LDH. Initial rates of the enzymatic reaction (V_i) at the indicated CBR-5884 concentration normalized to that of the DMSO control (V_o) are plotted. The mean IC_{50} ± 1SD is given from three replicate experiments.

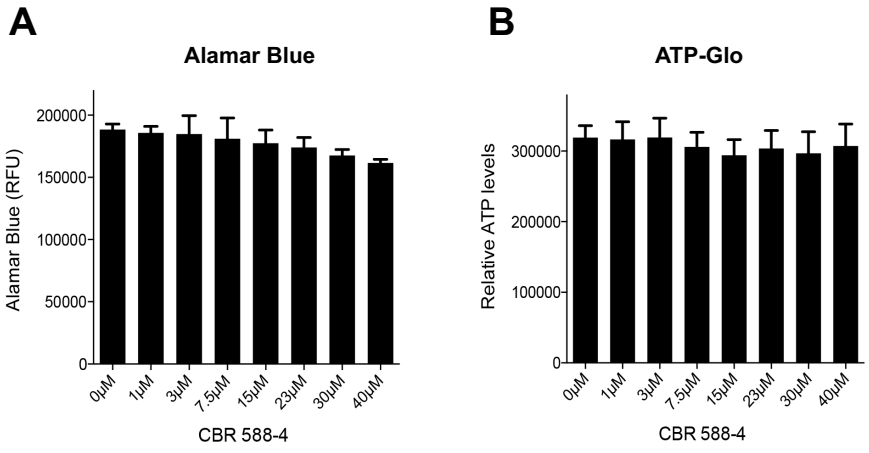


Figure 2.15. CBR-5884 is not generally toxic to cells. Carney cells were treated with a serial dilution of CBR-5884 for 3 hours to mimic the acute $^{13}C_6$-glucose labeling assay conditions and cellular reducing potential and ATP levels were determined using an (A) Alamar blue or (B) ATPGlo assay, respectively. Error is given as mean ± 1SD (n ≥ 3).

We resynthesized CBR-5884 in house and performed a dose response experiment for CBR-5884 employing the same acute treatment method as above. Serine labeling was significantly decreased at 30μM and trended towards a decrease at 15μM (Figure 2.16). Importantly, perturbations in labeling were specific to serine in that neither the PHGDH substrate, 3-PG, nor the end products of glycolysis, pyruvate and lactate, was affected (Figure 2.16). We further confirmed that glycolytic metabolites were unperturbed by CBR-5884 treatment using LC-MS/MS to interrogate a greater panel of metabolites (Figure 2.16). Thus, changes in serine labeling are a direct effect of CBR-5884 mediated PHGDH inhibition and not a consequence of changes in PHGDH substrate levels or general perturbations in glycolytic flux. The absence of an effect on lactate labeling was consistent with the *in vitro* data showing that CBR-5884 does not inhibit LDH under the drug concentrations used. In sum, the data argue that CBR-5884 is able to selectively inhibit serine synthesis in cells.

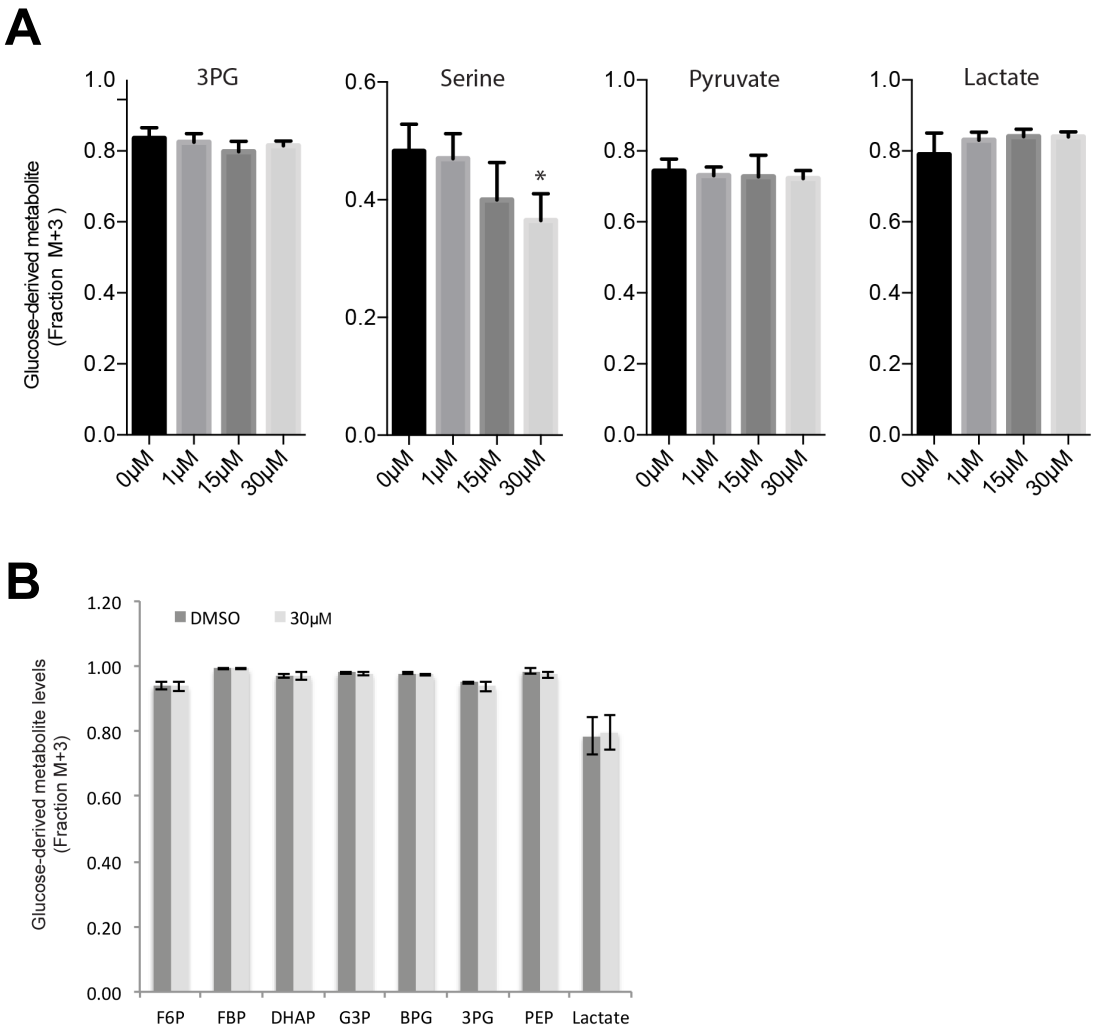


Figure 2.16. CBR-5884 inhibits serine synthesis in cells without perturbing glycolysis. (A) Dose response experiment on Carney cells using CBR-5884 synthesized in house following the same acute labeling assay as in Figure 2.13 but monitoring a panel of phosphoserine pathway and glycolytic metabolites. (B) Glycolytic metabolite levels for Carney cells pre-treated with DMSO or CBR-5884 (30μM) for 4 hours and then labeled for 2 hours with U-^{13}C-glucose under continued DMSO or drug treatment. Polar metabolites were extracted and analyzed by LC-MS/MS. The fractional labeling (M+3 metabolite level relative to total metabolite level) is plotted for each indicated metabolite Error is given as ± 1SD, n ≥ 3. F6P: fructose-6-phosphate; FBP: fructose-1,6-bisphosphate; DHAP: dihydroxy-acetone-phosphate; G3P: glyceraldehyde-3-phosphate; BPG: 1,3-bisphosphoglycerate or 2,3-bisphosphoglycerate (isotopomers); 3PG: 3-phosphoglycerate; PEP: phosphoenolpyruvate.

Given that CBR-5884 is an ethyl ester and therefore susceptible to intracellular esterases, we investigated whether the carboxylic acid derivative of the parent molecule was still active against PHGDH; were the acid less active, it would likely decrease the efficiency of targeting PHGDH *in situ*. Parent and acid derivatives were equally potent against and selective for PHGDH *in vitro* suggesting that intra-cellular de-esterification is unlikely to affect CBR-5884 activity (Figure 2.17).

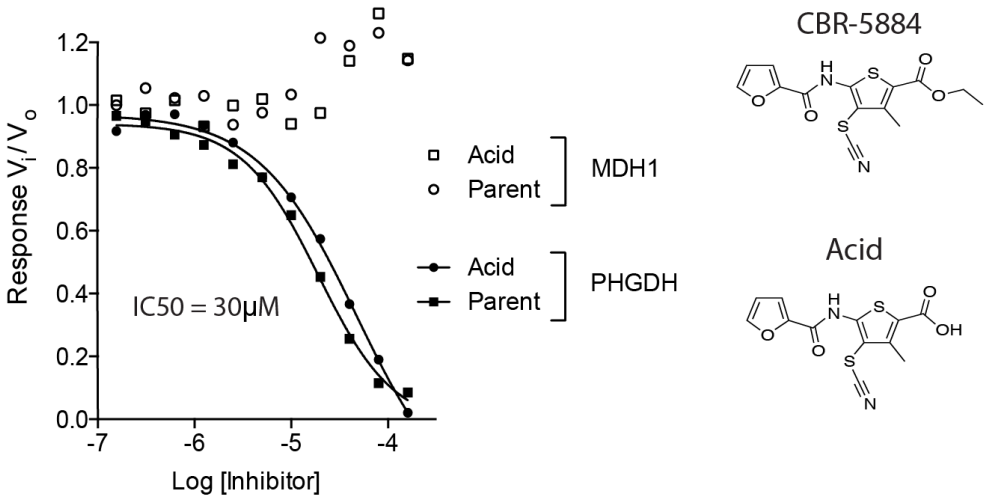


Figure 2.17. *In vitro* IC_{50} of CBR-5884 and acid derivative on PHGDH and MDH1. *In vitro* enzymatic assay for malate dehydrogenase 1 (MDH1) or PHGDH titrating CBR-5884 (parent) or the carboxylic acid derivative (acid). Structures of parent and acid are provided for comparison. Initial rates (V_i) normalized to the DMSO control rate (V_o) are plotted.

CBR-5884 selectively inhibits the proliferation of cancer cells with a high propensity for serine synthesis

We wanted to establish a system to test the ability of CBR-5884 to inhibit PHGDH-dependent cancer cell proliferation. Thus, we decided to first evaluate the ability of a panel of breast and melanoma cell lines to proliferate in serine replete or deplete media as a proxy for serine biosynthetic activity. Breast lines were selected based on *PHGDH* expression according to the Cancer Cell Line Encyclopedia (CCLE) data; no data were available for the melanoma lines (Figure 2.18)[82]. PHGDH levels were validated by western blotting (Figure 2.19). In the breast cancer cell lines, PHGDH expression levels appeared to be a good predictor of protein levels consistent with the report that in a large panel of non-small cell lung cancer cell lines PHGDH expression correlates strongly with serine biosynthetic activity [41]. Removing extracellular serine had no effect on the proliferation of the four high PHGDH expressing breast cancer lines (MDA-MB-468, MDA-MB-436, HCC70, Hs578T) or the PHGDH high expressing melanoma line WM266-3 (Figure 2.20). All four breast cancer lines cluster in the top quartile of the CCLE data set for *PHGDH* expression (Figure 2.18); MDA-MB-468, HCC70, and WM266-3 cells harbor *PHGDH* amplifications [44,45]. In contrast, serine depletion almost completely abrogated proliferation of low *PHGDH* expressing lines: MDA-MB-231, MCF-10A, and Gak (Figure 2.20). Interestingly, although Carney cells are sensitive to extracellular serine depletion, they can

adapt and resume proliferation as evidenced by increased PHGDH protein levels upon prolonged serine depletion (Figure 2.21).

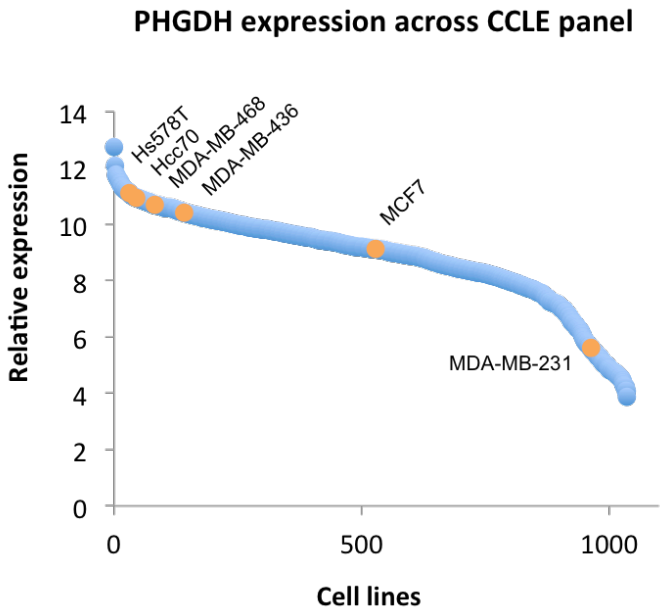


Figure 2.18. PHGDH expression across CCLE panel. Relative *PHGDH* expression levels across a panel of 1036 cell lines obtained from the CCLE data set. Selected breast cancer lines used for drug proliferation assays are highlighted in orange. Hcc70 and MDA-MB-468 cells are known to harbor PHGDH amplifications. No CCLE expression data were available for the melanoma lines.

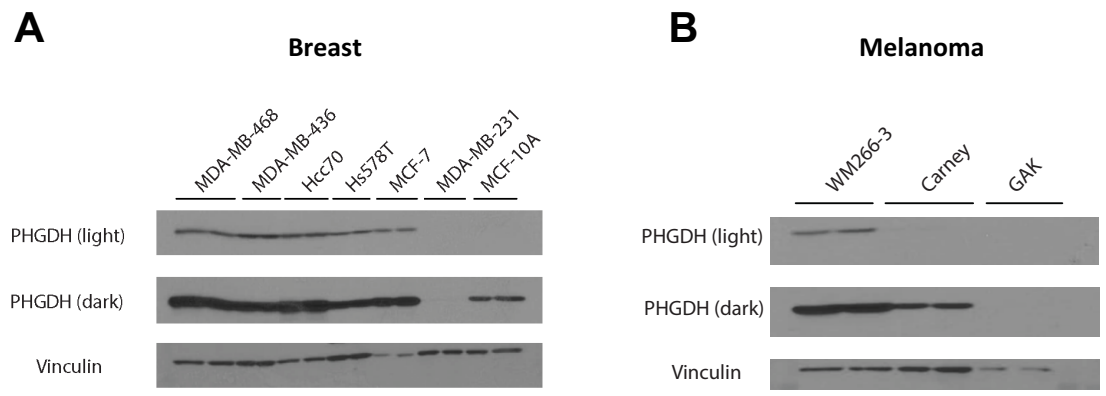


Figure 2.19. PHGDH protein levels across a panel of breast and melanoma lines. Western blot for PHGHD on panel of (A) breast and (B) melanoma lines grown in serine containing media; 2 lanes per cell line with each lane loaded with independent cell lysates. Hcc70, MDA-MB-468, and WM266-3 are *PHGDH* amplified.

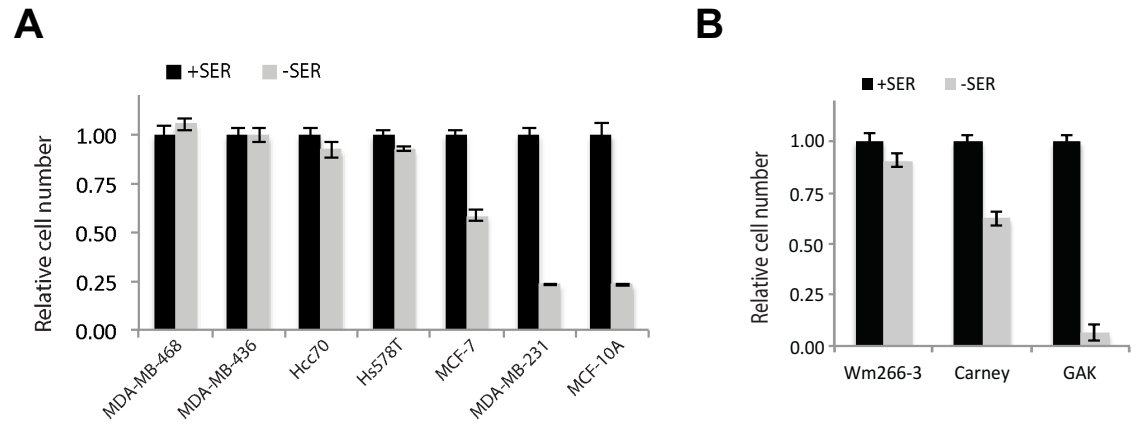


Figure 2.20. Proliferation of breast and melanoma lines in serine replete or serine deplete media. (A) Breast and (B) melanoma lines were grown in serine replete (+SER) or serine deplete (-SER) media for 3-5 days. Relative cell numbers were determined via crystal violet staining or a Cyquant assay.

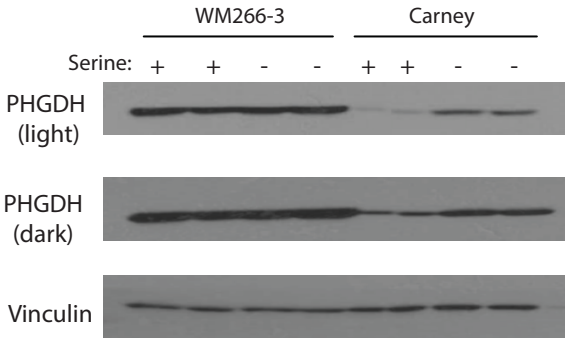


Figure 2.21. Carney cells adapt to serine deprivation by upregulating PHGDH protein levels. Western blot for WM266-3 and Carney cells grown in serine replete (+) or serine deplete (-) media; 2 lanes per cell line with each lane loaded with independent cell lysates.

Given that the ability to proliferate in the absence of extra-cellular serine is indicative of a high propensity for serine synthesis, we hypothesized that such lines should be sensitive to CBR-5884. Conversely, lines that cannot grow in serine free media have a low propensity for serine synthesis and should therefore be resistant to PHGDH inhibition. Treating the breast lines with CBR-5884 in serine replete media inhibited growth of the four lines that grew without extra-cellular serine in a dose dependent manner with growth inhibition ranging from 35% to 60% at 30μM CBR-5884. The inhibitor had no effect on the three lines sensitive to serine withdrawal indicating that the inhibitor was selectively toxic to cells with high serine synthesis activity (Figure 2.22). We next asked whether removing serine from the media, to enhance the reliance on *de novo* serine synthesis, could sensitize cells to PHGDH inhibition. Indeed, serine depletion increased the efficacy of CBR-5884 in lines already sensitive under serine replete conditions as evidenced by an 80% to 90% decrease in

proliferation with 30μM CBR-5884 (Figure 2.22). Moreover, MCF7 cells, which were of intermediate sensitivity to serine withdrawal, and insensitive to drug under serine replete conditions, became partially sensitive to the inhibitor under serine deplete conditions.

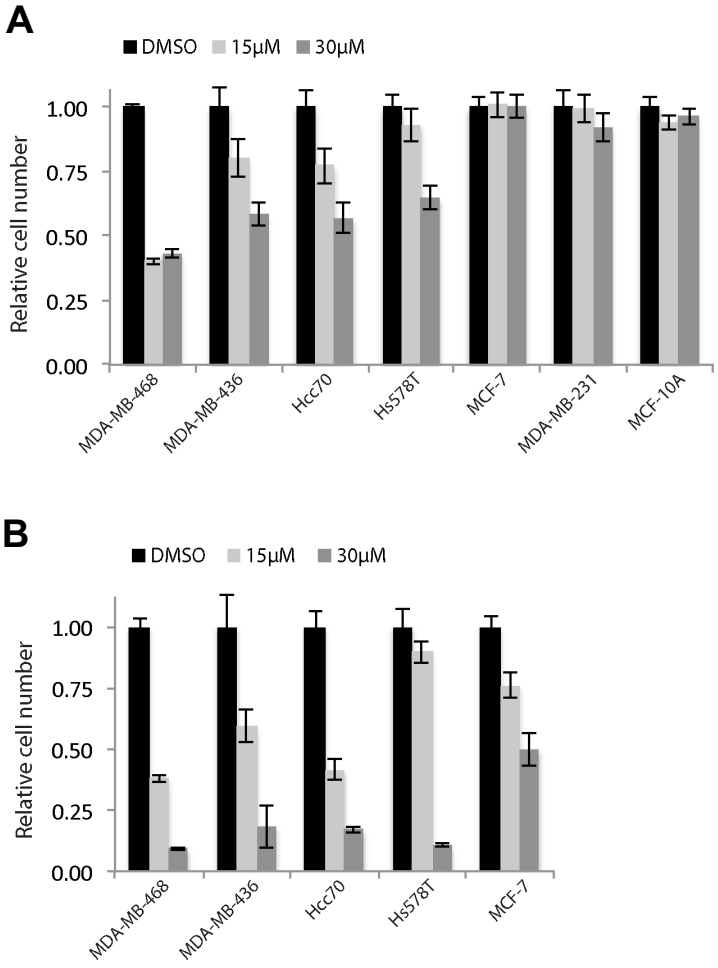


Figure 2.22. CBR-5884 selectively inhibits the proliferation of breast cancer lines that overexpress PHGDH. Proliferation assay for breast lines treated with CBR-5884 in (A) serine replete or (B) serine deplete media. MDA-MB-468 and Hcc70 are *PHGDH* amplified. MCF-10A is a non-transformed mammary epithelial cell line; other lines are cancer cell lines. MDA-MB-231 and MCF-10A lines were not included in –SER experiments because they are sensitive to serine withdrawal. Histograms depict mean ± standard error (n ≥ 3).

Importantly, under serine replete conditions, PHGDH knockdown phenocopied the effects of CBR-5884 treatment in that the drug-sensitive lines were also sensitive to PHGDH knockdown (Figure 2.23). Furthermore, as with the drug treatments, growing cells in serine free media enhanced the growth inhibitory effect of PHGDH knockdown although decreased proliferation under serine starved conditions seemed more pronounced with CBR-5884 treatment than with PHGDH knockdown (Figure 2.23). Similar trends were observed for the melanoma panel in terms of both the selectivity of CBR-5884 for cells with a high propensity for serine synthesis and the increased efficacy under serine deplete conditions (Supplementary Figure 2.2 and 2.3). Given the prevalence of esterases and their role in drug metabolism, and that charged molecules often have poor plasma membrane permeability, we were concerned that de-esterification of CBR-5884 *in vivo* might compromise activity on cells [83,84]. Indeed, the acid derivative of compound 588-4 was not effective on MDA-MB-468 cells and is therefore not a viable alternative to the parent compound (Supplementary Figure 2.4).

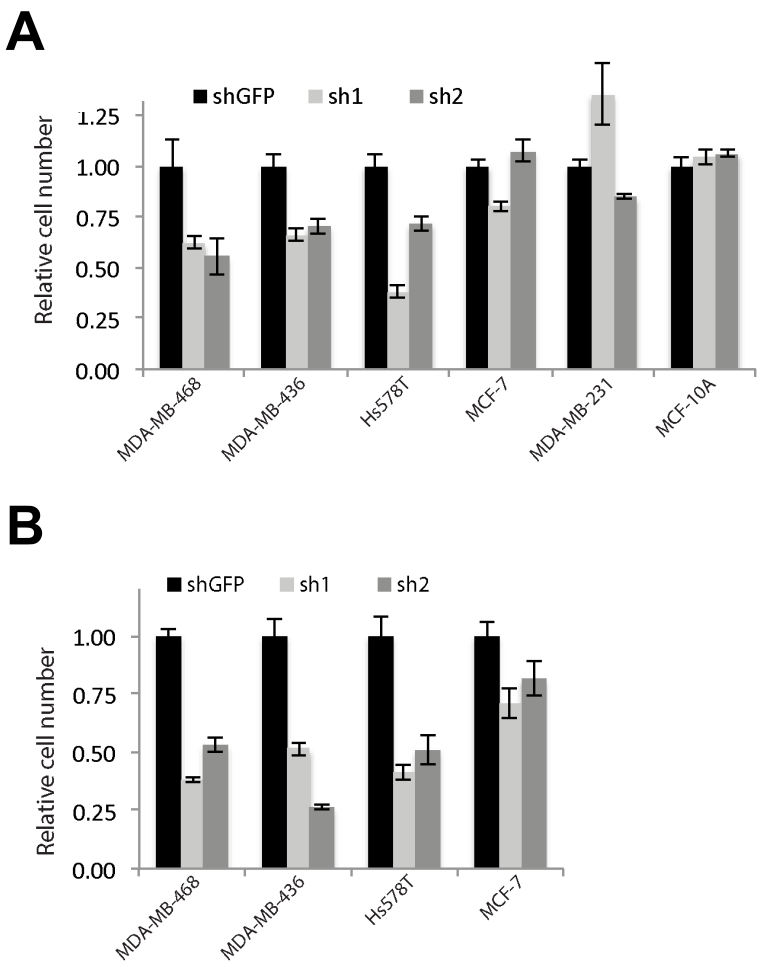


Figure 2.23. Knockdown of PHGDH in a panel of breast cancer lines. Proliferation assay for lines grown in (A) serine replete media (+SER) or (B) serine deplete media (-SER) with PHGDH knockdown (sh1 & sh2) or a nontargeting control (shGFP). MB-231 and MCF-10A lines were not included in –SER experiments because they are sensitive to serine withdrawal. Histograms depict mean ± standard error (n ≥ 3).

In order to more deeply characterize the metabolic changes following CBR-5884 treatment we turned to LCMS based targeted metabolomics. We knocked down PHGDH using two different hairpins and compared the resultant metabolic profile to that following CBR-5884 treatment. Interestingly, the metabolic signature following CBR-5884 treatment was similar to that following

either of the PHGDH knockdowns (sh1 and sh2) in that a majority of the significantly changed metabolites following CBR-5884 treatment overlapped with changes seen following PHGDH knockdown (Figure 2.24). Notably, the number of uniquely changed metabolites in the CBR-5884 treatment (20 metabolites) was less than the number of uniquely changed metabolites for sh1 (37 metabolites) and similar to the unique changes for sh2 (16 metabolites) suggesting that the specificity of CBR-5884 for PHGDH is comparable to that of sh1 and sh2 in terms of metabolic effects. Importantly, decreases in serine and purine nucleotides (IMP, AMP, ADP, GMP, GDP, dAMP, dGDP), known downstream products of serine metabolism, were among the 47 metabolites commonly altered across the CBR-5884 treatment and both knockdowns, indicating that CBR-5884 has on target activity similar to that of the knockdowns (Figure 2.24) [41,43]. Notably, glycerol-3-phosphate levels were increased in the CBR-5884 treatment whereas they were decreased in the two knockdown conditions, suggesting that the glycerol-3-phosphate dehydrogenase (GPDH) reaction that converts dihydroxyacetone phosphate (DHAP) to glycerol-3-phosphate might be uniquely perturbed following CBR-5884 treatment (Figure 2.24). Consistent with this, a decrease in DHAP was among the 20 metabolites that were uniquely perturbed by CBR-5884 treatment. Thus, we are currently investigating whether CBR-5884 has an off-target activity on GPDH.

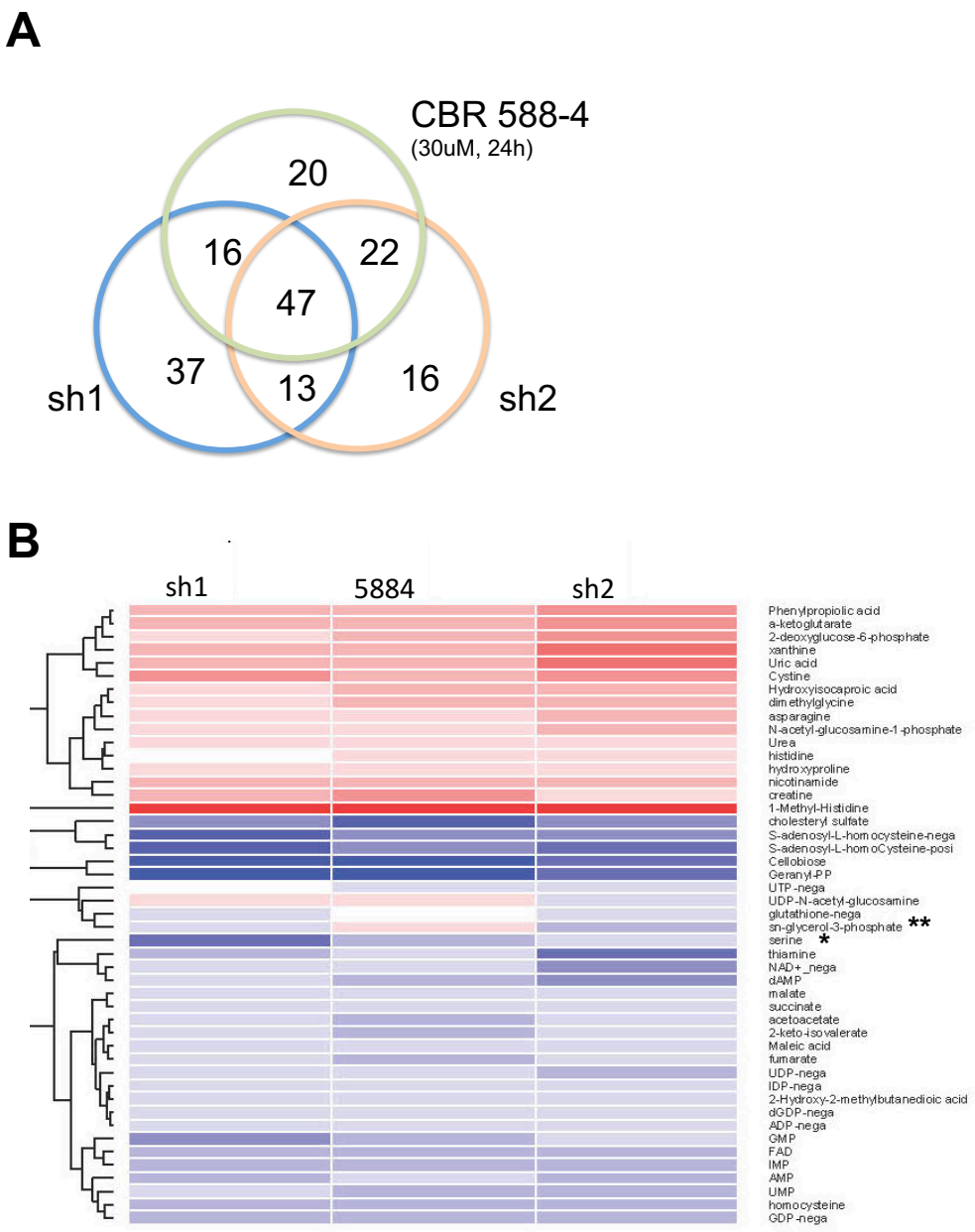


Figure 2.24. Comparison of metabolic changes following PHGDH knockdown versus CBR-5884 treatment. Polar metabolites were extracted from MDA-MB-468 cells following PHGDH knockdown (sh1 & sh2) or a nontargeting control (shGFP) knockdown and CBR-5884 treatment (24h, 30μM), and analyzed by targeted LCMS. Significantly altered metabolite levels for each group (sh1, sh2, or CBR-5884) were determined from comparison to shGFP control cells. (A) Venn diagram of all significantly changed metabolites. (B) Heat map depicting metabolic changes common to sh1, sh2 and CBR-5884 treatment. Blue: decrease; red: increase. * Serine. ** Glycerol-3-phosphate.

Analysis of CBR-5884 inhibition modality

We sought to more deeply characterize the mechanism by which CBR-5884 inhibits PHGDH. Inhibition constants (K_i) for CBR-5884 with respect to each substrate were determined. CBR-5884 inhibited PHGDH in a non-competitive mode with respect to both substrates, as evidenced by a decreasing V_{max} with increasing CBR-5884 concentration. The inhibition constants were 50 ±20μM and 50 ±3μM for 3PG and NAD^+, respectively (Figure 2.25).

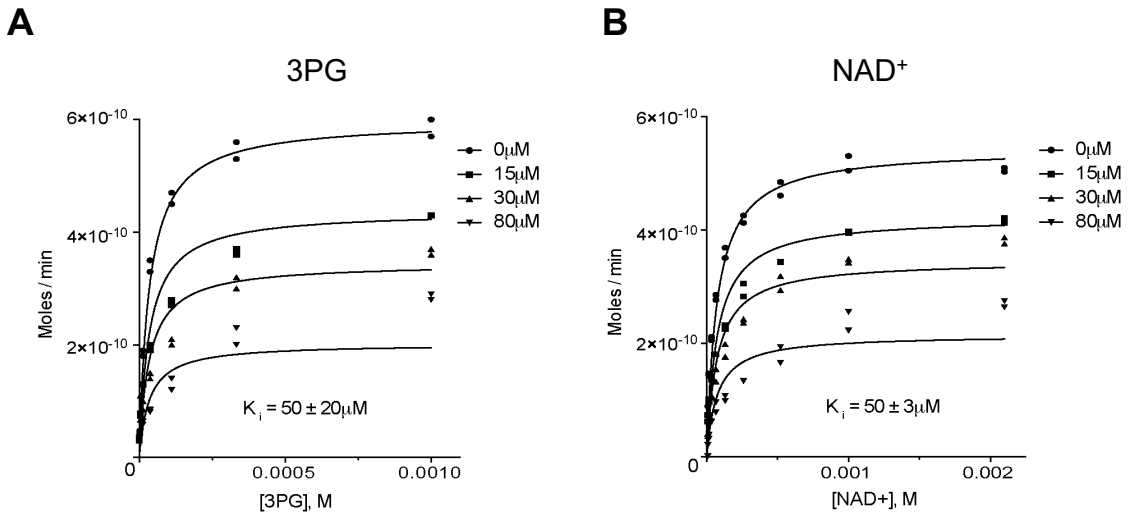


Figure 2.25. Mechanisms of CBR-5884 inhibition. Inhibition constants (K_i) were determined by titrating (A) 3PG or (B) NAD^+ while holding the other substrate constant at 4 different CBR-5884 concentrations and determining the initial reaction rate using a PHGDH assay. Plots were fit to a non-competitive model. Error is given as ± 1SD (n ≥ 3)

We assessed whether there was any time dependence to the onset of inhibition by varying the time period for which drug and PHGDH were pre-incubated before initiating the enzymatic reaction. CBR-5884 was progressively more potent with increasing pre-incubation time culminating in an IC_{50} of 7μM when drug and PHGDH were pre-incubated for 4 hours (Figure 2.26).

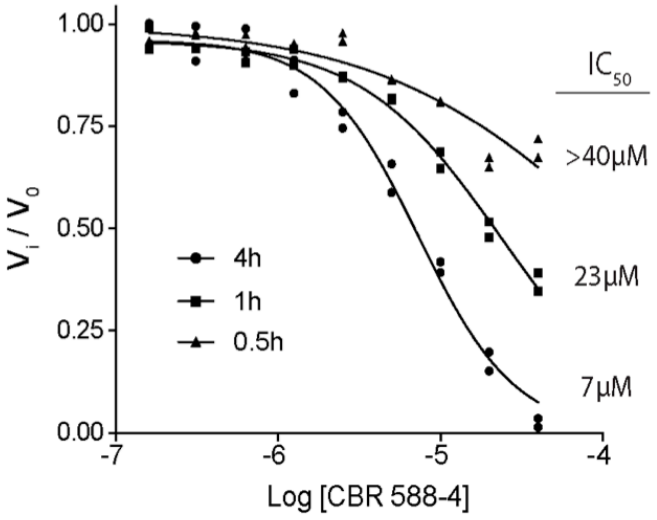


Figure 2.26. CBR-5884 is a time dependent inhibitor of PHGDH. Time dependent inhibition was measured by preincubating drug and PHGDH for 0.5h, 1h, or 4h as indicated prior to initiating the PHGDH reaction. Initial reaction rates (V_i) were determined and normalized to that of DMSO (V_o) and the IC_{50} determined. The observed IC_{50} following the different preincubation times are given.

Intrigued by the combination of a time dependent onset of inhibition and non-competitive inhibition, the latter suggesting that CBR-5884 might be binding to an allosteric pocket, we speculated that CBR-5884 could be affecting the PHGDH oligomerization state, where the time dependency of inhibition could

potentially stem from drug-induced conformational changes in PHGDH. To evaluate the PHGDH oligomerization state, we incubated PHGDH with drug and then cross-linked prior to SDS-PAGE. CBR-5884 shifted the PHGDH equilibrium from the tetrameric to the dimeric state (Figure 2.27). No such effect was observed with LDH, which is resistant to CBR-5884 mediated inhibition (Figure 2.27). In order to determine whether disruption of the PHGDH tetramer was necessary for inhibition, we investigated whether CBR-5884 could still inhibit a truncated form of PHGDH, which lacks the C-terminal domain responsible for tetramerization. In agreement with the crystal structure for $PHGDH^{3-314}$ (PDB: 2G76), the cross-linking assay confirmed that the truncated PHGDH protein was a constitutive dimer (Figure 2.27). However, CBR-5884 still inhibited $PHGDH^{3-314}$ with a similar potency (IC_{50} 20± 5μM) to that at which it inhibits the full-length protein suggesting that disruption of the tetramer is not essential for inhibition, at least *in vitro* (Figure 2.27). Together these results suggest that disruption of the tetramer might assist PHGDH inhibition, but is not necessary for inhibition. Finally, we investigated whether the time dependency of inhibition was due to the dissociation of the tetramer by varying the time for which PHGDH and drug were incubated prior to crosslinking (Figure 2.28). The tetramer to dimer ratio did not appear to be a function of drug and PHGDH preincubation time suggesting that the time dependency of inhibition does not stem from dissociation of the tetramer.

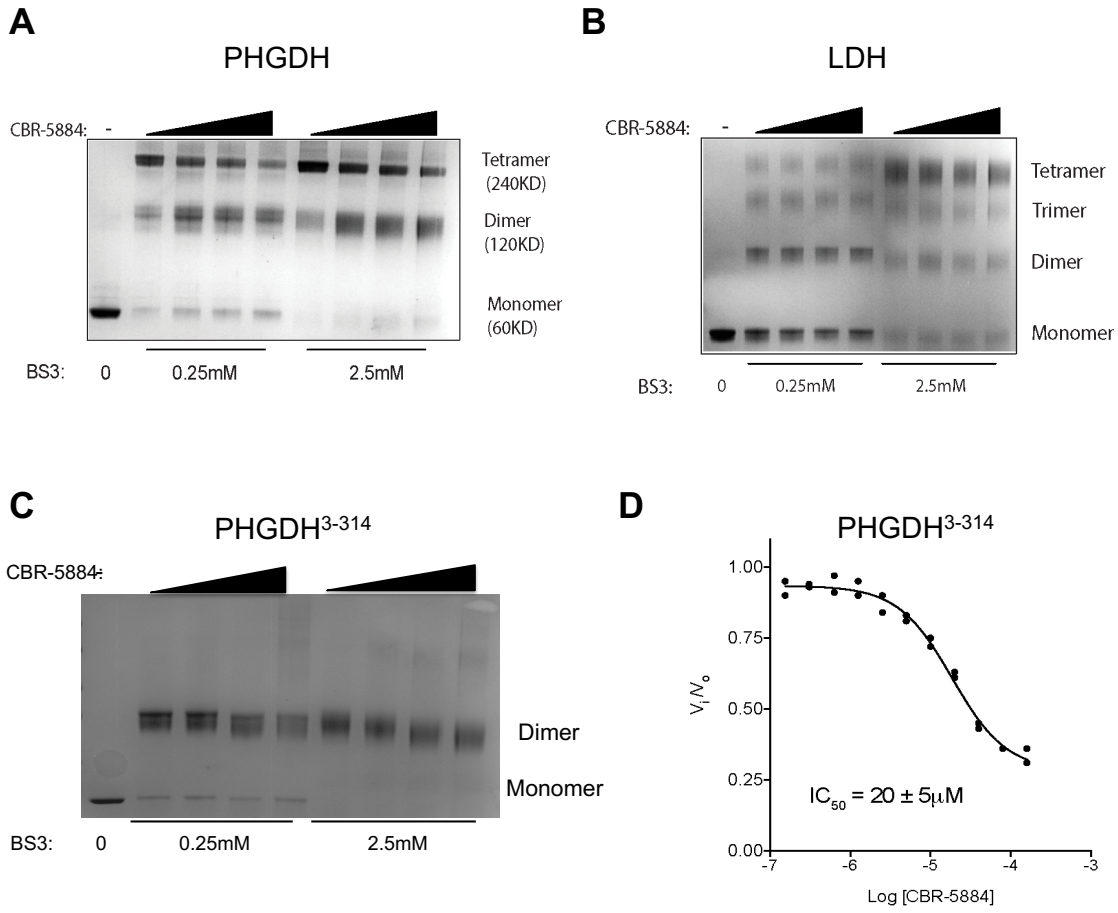


Figure 2.27. CBR-5884 disrupts the oligomerization state of PHGDH. (A) PHGDH or (B) LDH or (C) PHGDH$^{3-314}$ were preincubated with CBR-5884 (0, 50, 200, or 400μM) for 30min prior to crosslinking with BS3 (0.25 or 2.5mM) followed by SDS-PAGE and coomassie staining. Left-most lane had no BS3 indicating the monomeric species. Oligomerization state was inferred from reference to a molecular weight ladder. (D) *In vitro* enzymatic IC_{50} assay using a truncated form of PHGDH$^{3-314}$. Initial reaction rates (V_i) normalized to the DMSO control (V_o) are plotted. The IC_{50} for three independent experiments is given plus or minus SD.

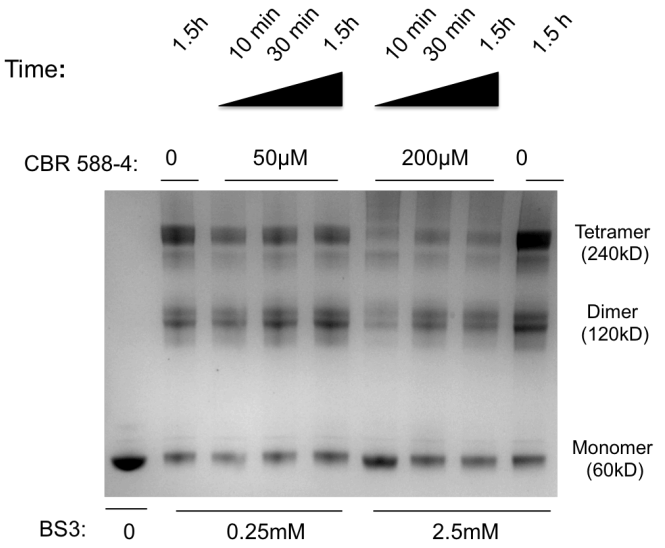


Figure 2.28. Disruption of the PHGDH tetramer is not time-dependent. PHGDH was preincubated with CBR-5884 (0, 50, 200μM) for 10min to 1.5h prior to crosslinking with BS3 (0.25mM or 2.5mM) followed by SDS-PAGE and coomassie staining.

Preliminary structure-activity relationship around the CBR-5884 scaffold

In order to increase the drug-like nature of CBR-5884 and optimize potency, we performed a small structure-activity relationship (SAR) study around the CBR-5884 scaffold. We were concerned with the relatively high doses of CBR-5884 required for *in vitro* activity and with the presence of thiophene and furan moieties. Both furan and thiophene moieties are susceptible to cytochrome P450 (CYP) oxidation forming a 2,3-epoxide that can ablate inhibitory activity and result in further metabolism of the compound to potentially toxic products [85]. Attempts to replace the furan with various benzyl substituents and a single chlorothiophene substituent were unsuccessful in that a majority of the resultant

derivatives were completely inactive (**2-6**) or five-fold less active (**1**) in the biochemical PHGDH assay suggesting that modifications of the furan moiety are poorly tolerated (Figure 2.29). Modifications of the thiophene core resulted in a 5-10 fold loss in activity (**7-9**). Interestingly, removing the thiocyano group or substituting it with a cyano group decreased activity approximately 10-fold and 5-fold, respectively, suggesting the thiocyano group might be essential for activity (**8, 9**). While the acid derivative of the parent molecule was equally active the *tert*-butyl ester was inactive (**10, 11**).

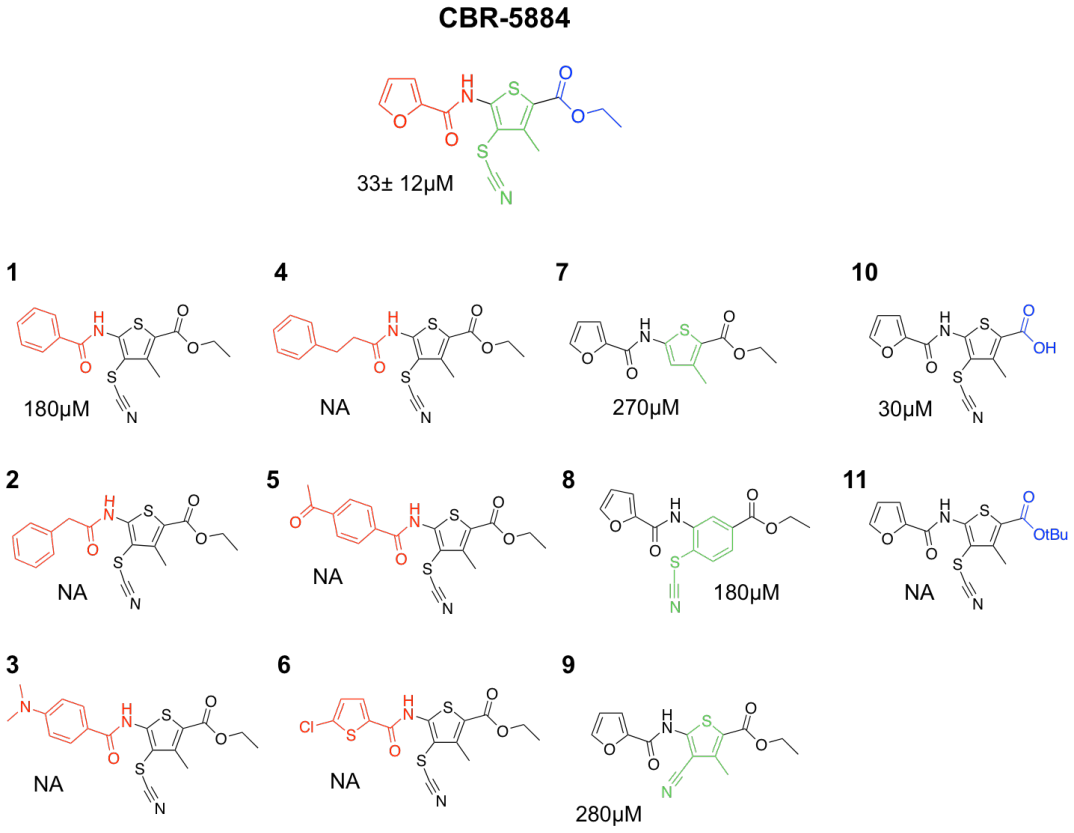


Figure 2.29. Preliminary SAR study around the CBR-5884 scaffold. The parent compound CBR-5884 is shown as well as the various derivatives. For each compound, the associated IC_{50} determined via an in vitro PHGDH assay is provided. NA: not any, meaning no detectable inhibition.

Discussion:

We have reported the discovery of a novel PHGDH inhibitor, CBR-5884, and shown that it inhibits serine synthesis in cells. Furthermore, CBR-5884 specifically inhibited the proliferation of melanoma and breast cancer lines with high levels of serine synthesis activity with little effect on lines reliant on serine import. Thus, CBR-5884 is selective for lines addicted to serine synthesis and phenocopies sensitivity to PHGDH knockdown. Finally, a biochemical analysis of CBR-5884 revealed that it was a non-competitive inhibitor that showed a time dependent onset of inhibition and disrupted the oligomerization state of PHGDH.

Recent work examining how malignant cells rewire their metabolism to support growth and proliferation has revealed a number of clinically interesting targets [11,86]. Perhaps the most promising is the discovery of gain-of-function mutations in the isocitrate dehydrogenase (IDH) enzymes that result in the production of the oncometabolite 2-hydroxyglutate [64,87]. The findings translated into chemical probes that yielded insights into the biology of IDH mutations [70,88] and led to clinical programs (eg. NCT02481154). The genetic evidence pointing to a role for PHGDH in cancer is similarly striking: PHGDH is one of few metabolic enzymes genetically deregulated in cancer [63,89]. Notably, elevated PHGDH expression correlates with clinical aggressiveness and poor prognosis in TNBC [45,47] and NSCLC [41]. There is a paucity of targeted therapies for these cancers and chemotherapies remain mainstay [90,91]. Hence, the clinical potential of PHGDH inhibitors as targeted agents for TNBC and NSCLC tumors addicted

to serine synthesis, as a single agent or in combination with standard of care, is an exciting perspective.

Beyond the preclinical applications, a PHGDH inhibitor provides a new tool to study *de novo* serine synthesis. For example, it remains unclear why a serine biosynthesis enzyme is critical for tumor growth when serine is available in the serum [44,45]. CBR-5884 provides a valuable tool complementary to genetic strategies to study this and other phenomena as small molecules provide greater temporal resolution and do not deplete the actual protein. The dynamic nature of metabolic flux means that the initial perturbation in flux following PHGDH knockdown may resolve to a new steady state before one can interrogate the system. With the greater temporal resolution of an inhibitor, these early and potentially important events can be studied. Moreover, CBR-5884 has already suggested an interesting feature of PHGDH biochemistry, namely that human PHGDH could be regulated by transitions between different oligomerization states. These findings are reminiscent of pyruvate kinase M2 (PKM2) regulation in that both an endogenous metabolite, fructose-1,6-bisphosphate, and pharmacological small molecule activators enhance PKM2 activity by stabilizing the tetrameric form [92,93]. Endogenously, 2-phosphoglycerate (2PG) has been reported to activate PHGDH [94]. Mechanistically, 2PG could be functioning by modulating the PHGDH oligomerization state. Thus, it would be interesting to determine whether 2PG modulates the PHGDH oligomerization state or, more generally, whether PHGDH activity is subject to such regulation *in situ*. Finally, it is possible that CBR-5884 is a covalent inhibitor of PHGDH. While PHGDH does

not rely on an active site cysteine, there are a number of cysteines that could potentially perturb enzymatic function were they modified (PDB: 2G76). In an attempt to determine whether changes in PHGDH oligomerization state are unique to CBR-5884, it would be interesting to determine whether disulfiram (CBR-5807), CBR-6936, or sulfhydryl blocking reagents have similar effects.

The identification of a selective novel small molecule inhibitor of PHGDH capable of modulating *de novo* serine synthesis in PHGDH-dependent cancer cells represents a significant step towards the goal of targeting cancer metabolism in oncology. However, future *in vivo* evaluation of the CBR-5884 chemical series will require medicinal chemistry-based optimization. Indeed, CBR-5884 was found to be unstable in mouse plasma (data not shown) and, as described above, replacement of the ethyl ester moiety with the corresponding negatively charged carboxylic acid resulted in a derivative that retains enzyme inhibitory activity but loses activity on cells. Thus, CBR-5884 is more likely to serve as a tool compound, and as a starting point for generating more drug-like molecules, than an actual drug. Efforts aimed at identifying cell permeable plasma-stable derivatives are ongoing. Furthermore, the cell-based potency of this series will likely need to be improved to enable *in vivo* evaluation at exposure levels that are not generally toxic. Medicinal chemistry approaches aimed at optimizing the potency and ADME properties of CBR-5884 are ongoing. Nevertheless, our results provide a proof-of-concept that small molecule inhibitors of PHGDH represent a viable new class of anti-cancer drugs.

Methods

Cells, transfections, and infections:

Breast and melanoma lines were passaged in RPMI supplemented with: 10% FBS, penicillin, streptomycin, normocin (InvivoGen). For lentivirus production and infection, Lenti-X 293T cells (Clontech) were transfected at 90% confluence with lipofectamine 2000 (Invitrogen) with the indicated pLKO.1 shRNAs and packaging plasmids pCMV-dR8.2 and pCMV-VSV-G. PHGDH, sh1: TRCN0000233029; sh2: TRCN0000221864; nontargeting control, shGFP: TRCN0000072181. Viral supernatants were collected after 2 and 3 days. Filtered (0.45μm) viral supernatant supplemented with 2μg/ml polybrene was added to target cells overnight, and media changed the next morning. Selection (2μg/ml puromycin) was initiated the following day for 3 days at which point the cells were either harvested for protein or seeded for proliferation assays. Breast cell lines were obtained from ATCC and melanoma lines were a gift from Dr. Haoqiang Ying (MD Anderson Cancer Center).

Immunoblots:

Media was removed, cells were washed with PBS, and proteins were isolated directly from intact cells via acid extraction using a 10% TCA solution (10mM Tris-HCL pH 8.0, 10% trichloroacetic acid, 25mM NH_4OAc, 1mM Na_2EDTA). Precipitated proteins were harvested and solubilized in a 0.1 M Tris-HCl, pH 11 solution containing 3% SDS and boiled for 5-10 minutes. Lysate protein content was quantified via the BCA method, run on an SDS-PAGE,

transferred to a nitrocellulose membrane, blocked (5% BSA), and probed with primary antibodies: αPHGDH (Sigma HPA021241, 1/10000, 5% BSA); αVinculin (Sigma V9264, 1/5000, 5% BSA).

Proliferation assays:

Cells were plated at a low density in 96 or 24 well plates in serine containing media. The following day, media was aspirated, cells were washed with PBS, and fresh serine replete or serine deplete media containing drug (15μM, 30μM) or vehicle (DMSO) was added. Serine replete or deplete media was made from serine/ glycine free DMEM (containing pyruvate) that had been supplemented with 10% dialyzed FBS and either serine (400μM final, serine replete) or PBS (serine deplete). Cells were grown for 3 to 5 days, adding fresh drug containing media everyday, before assaying relative cell numbers. For 96 well plates, relative cell numbers were determined using a Cyquant (Life Technologies) assay according to the manufacturer's instructions. For 24 well plates, crystal violet staining was performed as follows: cells were washed with PBS, fixed with 10% formalin (10min), washed twice with PBS and stored at 4^{o}C in PBS until the completion of the experiment after which PBS was removed and cells were stained with 0.1% crystal violet in 20% methanol for 15min. After staining, cells were washed thrice with water and air dried overnight. Cell-bound crystal violet was solubilized in 10% acetic acid and 595nm absorbance measured. Serine/glycine free DMEM was custom ordered from Gibco.

Carney serine labeling time course and acute drug treatments with $^{13}C_6$-glucose tracing:

Carney cells were acclimated to growth in MEM (Corning) supplemented with 10% dialyzed FBS, penicillin, streptomycin, and normocin (InvivoGen) by passaging for 2 weeks (~6 passages). Cells were plated at 9×10^5 cells per 6cm dish a day prior to the experiment. For the serine labeling time course: cells were washed with PBS and fresh glucose free MEM supplemented with $^{13}C_6$-glucose (3g/L, Cambridge Isotopes) and 10% dialyzed FBS was added. Cells were then incubated at 37^{o}C and harvested, over 9 time points covering a range of 0.12h to 6h, by quickly washing cells with cold PBS and flash freezing. For acute drug inhibition assays: media was replaced with fresh media containing CBR-5884 (1μM, 15μM, 30μM) or vehicle control (DMSO) for 1h. Media was then aspirated, cells were washed with PBS, and fresh glucose free MEM supplemented with $^{13}C_6$-glucose (3g/L, Cambridge Isotopes) and 10% dialyzed FBS containing drug or DMSO was added. After 2h, cells were quickly washed with cold PBS on ice and flash frozen. Polar metabolites were then extracted and analyzed as described in the GCMS methods. The seven initial lead PHGDH inhibitors were assayed as in the CBR-5884 dose response experiment except that only 30μM drug or DMSO control conditions were used. Glucose free MEM was custom ordered from Gibco.

Acute toxicity assay:

Carney cells acclimated to growth in MEM media supplemented with 10%

dialyzed FBS, penicillin, streptomycin, and normocin (InvivoGen) were plated in a 96 well plate at 6,000 cells/well. The next day, cells were treated with CBR-5884 from 1μM to 40μM for 3h. Drug containing media was then removed, fresh drug-free media added, and cell viability was determined via a CellTiter-Glo (Promega G7572) or Alamar Blue (Invitrogen DAL1025) assay according to the manufacturers' protocol. For Alamar Blue assays, cells were incubated with Alamar Blue containing media for 2h at 37°C following removal of drug before reading fluorescence.

GCMS metabolite analysis:

Polar metabolites were extracted with 2ml MeOH/H_2O (4:1) for 30min on dry ice, scraped, transferred to 2ml tubes, centrifuged (30min, 15000rpm), and the supernatants dried under vacuum. Samples were derivatized as previously described using methoxyamine hydrochloride and *N*-Methyl-*N-tert*-butyldimethylsilyltrifluoroacetamide (34). Samples were pulse spun to remove insoluble matter, the supernatant transferred to inserts set in brown glass vials and capped. Analysis was performed on an Agilent 6890 GC instrument. Samples were loaded onto a 30m DB-35MS capillary column using helium carrier gas and interfaced to an Agilent 5975B MS. Electron impact (EI) ionization was set at 70eV. Each sample was injected at 270°C at a flow rate of 1 ml/min. To mobilize metabolites, the GC oven temperature was held at 100°C for 3 min and increased to 300°C at 3.5°C/min for a total run time of approximately 1h. 1μl of each sample was injected in splitless mode. All analyses were operated in full

scan mode while recording a mass to charge ratio (Δm/z) spectra in the range of 100-650 m/z. Specific metabolite Δm/z analyzed are available upon request. For quantification of metabolites, unknown samples were fit to a standard curve of known metabolites assessed in the linear range of 0.006-0.33 nmoles on column. Standards were run in parallel to samples in order to maintain accuracy of retrofit quantitation. Fractional enrichment of ^{13}C in metabolites has been corrected for the natural abundance of ^{13}C and ^{15}N using METRAN and in house scripts written in Matlab (35, 36).

LC-MS/MS metabolite analysis:

Polar metabolites were extracted and dried as in the GCMS method. Samples were resuspended in 15µl of HPLC grade water. 5µl of each sample was injected and analyzed using a 5500 QTRAP triple quadrupole mass spectrometer (AB/Sciex) coupled to a Prominence UFLC system (Shimadzu) as reported previously [4,95].

PHGDH and PSAT1 purification:

pET28a PHGDH, pET28a PSAT1, or pNIC28-Bsa4 PHGDH$^{3-314}$, was transformed into BL21 *E. coli*. A single colony was grown to an OD_{600} 0.7 in 1L of Luria broth and protein expression was induced with IPTG (0.5mM). The culture was chilled on ice for 30min, cultured for 20h at room temp, pelleted (6000g, 20min), and flash frozen. Pellets were resuspended in 60ml of lysis buffer (50mM Tris pH 8.5, 10mM $MgCl_2$, 300mM NaCl, 10% glycerol, 5mM imidazole),

sonicated and cell debris pelleted by centrifugation (20,000g, 30min). The supernatant was collected and 2ml of Ni-agarose beads (Qiagen), pre-washed twice with wash buffer (50mM Tris pH 8.5, 10mM $MgCl_2$, 300mM NaCl, 10% glycerol, 30mM imidazole), and 60µl of 2-mercaptoethanol were added. Beads and lysate were incubated for 3h on a rocker at 4°C. Beads were batch-washed 4 times with wash buffer and transferred to a column. Bound proteins were eluted (elution buffer: 50mM Tris pH 8.5, 10mM $MgCl_2$, 250mM NaCl, 10% glycerol, 250mM imidazole) collecting 1ml fractions. Fraction protein content was measured via a Bradford assay. The most concentrated fractions were pooled and dialyzed overnight into 2L of dialysis buffer (50mM Tris pH 8.5, 10mM $MgCl_2$, 250mM NaCl, 20% glycerol, 0.15% 2-mercaptoethanol). The next morning, dialysis buffer was changed and dialysis continued for 4h. Protein purity was assessed via SDS-PAGE and coomassie staining.

IDH1 purification:

IDH1 gene was PCR amplified from SC322129 (IDH1 (NM_005896) Human cDNA) from Origene. The cDNA sequence was cloned into a pVB-CBD vector by inserting the IDH1 gene downstream of OmpA and CBD (cellulose binding domain). BL21-Gold (DE3) cells (Agilent) were transformed with the pVB-CBD-IDH1 plasmid and single colonies were generated. Protein expression was performed under IPTG (1mM) induction at 25°C for 24h. The supernatant containing the IDH1 protein was separated by centrifugation. The IDH1-CBD was captured on a Macroporous Bead Cellulose MT100 column (Iontosorb). TEV

protease (Sigma-Aldrich) digestion was performed at 37°C for 1 hour. Size exclusion FPLC was performed using a Superdex 200 10/300 GL (GE Health Care Life Sciences) column. Purified protein purity was more than 95% as determined by SDS-PAGE.

PHGDH assays:

PHGDH activity was measured in 96-well plates (100µL/well) at 28°C by monitoring NADH fluorescence (Ex340nM/ Em460nM) over time with a FLUOstar Omega (BMG labtech). PSAT1 was included to prevent product inhibition of PHGDH. Assays were performed in PHGDH assay buffer (50mM Tris pH 8.5 and 1mM EDTA). Substrates and enzyme concentrations were: 3-phosphoglycerate 240µM; NAD^+ 120µM; glutamate 30mM; 5.7ng/µl for full length PHGDH or 240ng/µl for $PHGDH^{3-314}$; PSAT1 80ng/µl except for K_i measurements where one of the PHGDH substrates was held constant at 3mM while the other substrate was titrated from 2mM to 8µM for NAD^+ or 1mM to 0.15µM for 3PG. An NADH standard curve in PHGDH assay buffer was included to quantify fluorescent signal. For IC_{50} assays, a 2-fold serial dilution of drug ranging from 160µM to 0.15µM was preincubated with enzyme for 30min before initiating the enzyme reaction with substrate mix. For time dependent IC_{50} assays, drug preincubation time was 4h, 1h, or 30min. For K_i measurements, drug (0, 15µM, 30µM, 80µM) and enzyme were preincubated for 30min. Initial rate plots were fit using Prism. Chemicals were purchased from Sigma.

LDH and MDH1 assays:

Enzyme activities were assayed using kits (LDH: Sigma MAK06; MDH1: Sigma MAK196-1KT) according to the manufacturer's instructions with commercially available recombinant enzyme (LDH: Sigma 59747; MDH1: Sigma SRP6103). Drug, titrated as for the PHGDH IC_{50} assays, and enzyme were pre-incubated for 30min prior to initiating reaction with substrate.

Cross-linking assays:

PHGDH (1.5μg) or LDH (2.2μg, Sigma 59747) were incubated with CBR-5884 (50μM, 200μM, 400μM) or vehicle control (DMSO) in 25mM HEPES, pH 7.3, and 1mM NAD^+ in 18μL total volume for 30min. BS3 (Pierce, PI-21585) cross-linker dissolved in PBS was added to a final concentration 0, 0.25mM, or 2.5mM and incubated for 30min on a rocker. The reaction was then quenched for 15min by adding 1M Tris pH 7.5 to a final concentration of 27mM. Cross-linked proteins were mixed with sample buffer, boiled for 5min, and run on SDS-PAGE. Gels were stained with colloidal coomassie stain (Bio-Rad 161-0803) for 24h and destained according to the manufacturer's instructions.

Primary PHGDH screen:

800,000 compounds were screened at a single dose (13.3μM) in 1536-plate format. Bromopyruvate (0.5mM) was used as inhibitor control. Assay set up: 2μL of assay buffer was added to each well of a 1536 well plate. Compounds were transferred and 2μL of reaction mixture 2 was added followed by a 3-min

incubation. Reactions were initiated by adding 2μL of reaction mixture 1. Plates were incubated (30min at 37°C) before fluorescence quantification (550/590nm; Ex/Em) using an Envision plate reader. Reaction mixture 1: 3-phosphoglycerate (0.25 mM), resazurin (0.1 mM), NAD^+ (8.3 mM), glutamic acid (30 mM), Tris-HCl pH 8.0 (30 mM), 0.6mM EDTA. Reaction mixture 2: PSAT1 (0.02 ug/uL), PHGDH (0.005 ug/uL), and diaphorase (0.001 U/uL) diluted in Tris-HCl pH 8.0 (30 mM) and EDTA (0.6mM) buffer. Assay buffer Tris-HCl pH 8.0 (30 mM) and EDTA (0.6mM).

Diaphorase counterscreen:

All 800,000 compounds were screened against diaphorase (Worthington Biochemical Corporation) as in the primary screen but drug and enzyme were incubated for 20min at 25°C before initiating the reaction. Modified reaction mixtures were as follows. Mixture 1: resazurin (0.1 mM), NADPH (0.1 mM), buffer (30 mM Tris-HCl, 0.6 mM EDTA pH 8.0). Mixture 2: diaphorase (0.0002 U/uL), buffer (30 mM Tris-HCl, 0.6 mM EDTA pH 8.0).

Screening hit selection criteria, confirmation, and selectivity profiling:

Results were analyzed using Genedata Screener software. Compounds with a robust Z-score < -3 in the PHGDH screening assay and robust Z-score > -2 in the diaphorase counter screen were selected as hits. Hit confirmation was done in triplicate (13.3μM) in 1536-plate format as described above. In order to assess selectivity of confirmed hits, hits were tested in 1536-plate format in an 8-

point dose-response experiment (2-fold serial dilution) against a panel of dehydrogenases including: PHGDH, isocitrate dehydrogenase (IDH1), malate dehydrogenase (MDH1), and 3α-hydroxysteroid dehydrogenase (3α-HSD). *PHGDH assay*: performed as in the primary screen. *IDH1 assay*: 2μL of assay buffer was added to each well. Compounds were transferred and 2μL of reaction mixture 2 was added followed by a 3-min incubation. Reactions were started by the addition of 2μL of reaction mixture 1. Plate was incubated for 45 min at 37°C before fluorescence quantification. *MDH1 assay*: 2μL of assay buffer was added to each well. Compounds were transferred and 2μL of reaction mixture 4 was added followed by a 3-min incubation. Reactions were started by the addition of 2μL of reaction mixture 3. Plate was incubated for 2.5h at 37°C before fluorescence quantification. *3α-HSD assay*: 2μL of assay buffer was added to each well. Compounds were transferred and 2uL of reaction mixture 6 was added followed by a 3-min incubation. Reactions were started by the addition of 2μL of reaction mixture 5. Plate was incubated for 30 min at 37°C before fluorescence quantification. In all cases fluorescence (550/590nm; Ex/Em) was quantified with an Envision plate reader and IC_{50} values were obtained using Genedata Screener software. Compounds that were not at least 4-fold selective for PHGDH were eliminated and the seven most potent PHGDH inhibitors were selected for further cell-based work. Reaction mixture 1: isocitrate (0.36 mM), resazurin (0.1 mM), NADP+ (0.033 mM), $MnCl_2$ (1 mM), buffer (30 mM Tris-HCl, 0.6 mM EDTA pH 8.0). Reaction mixture 2: IDH1 (0.25 μg/μL; purified in house), and diaphorase (0.005 U/μL) diluted in buffer (30 mM Tris-HCl, 0.6 mM EDTA pH

8.0). Reaction mixture 3: L-malic acid (2 mM), resazurin (0.1 mM), NAD+ (0.5 mM), buffer (30 mM Tris-HCl, 0.6 mM EDTA pH 8.0). Reaction mixture 4: porcine MDH1 (Sigma M2634, 0.00025 µg/µL), and diaphorase (0.005 U/µL), buffer (30 mM Tris-HCl, 0.6 mM EDTA, pH 8.0). Reaction mixture 5: cholic acid (0.05 mM), resazurin (0.1 mM), NAD+ (0.007 mM), buffer (20 mM sodium pyrophosphate pH 9.0, 4 µg/mL BSA). Reaction mixture 6: 3a-HSD (from Sigma; 0.025 µg/µL), and diaphorase (0.005 U/µL), buffer (20 mM sodium pyrophosphate pH 9.0, 4 µg/mL BSA). Chemicals were purchased from sigma and diaphorase from Worthington Biochemical Corporation

Chemical syntheses:

Materials:

2-amino-4-methyl-3-cyanato-5-thiophenecarboxylate ethyl ester was obtained from Life Chemicals. Cyanoacetic acid was obtained from AK Scientific. All remaining reagents and solvents were obtained from Sigma-Aldrich. All reagents were used as received without further purification.

Ethyl 5-(furan-2-carboxamido)-3-methyl-4-thiocyanatothiophene-2-carboxylate:

86.3 mg 2-amino-4-methyl-3-cyanato-5-thiophenecarboxylate ethyl ester (357 µmol, 1 eq.) was suspended in 1 mL DCM with a magnetic stirbar (Figure 2.30). 86.5µL of pyridine (1070 µmol, 3 eq.) was added to the mixture, followed

by drop-wise addition of 35.3 µL 2-furoyl chloride (357 µmol, 1 eq.). After 1 hour the reaction mixture was purified directly by normal-phase silica gel chromatography. Collected fractions were dried *in vacuo* to obtain 87.2 mg ethyl 5-(furan-2-carboxamido)-3-methyl-4-thiocyanatothiophene-2-carboxylate as a white solid, 73% yield. ^{1}H NMR (400 MHz $CDCl_3$) δ 9.36 (s, 1H), 7.67 (dd, *J* = 0.8, 1.7 Hz, 1H), 7.41 (dd, *J* = 0.8, 3.6 Hz, 1H), 6.67 (dd, *J* = 1.7, 3.6 Hz, 1H), 4.36 (q, *J* = 7.1 Hz, 2H), 1.40 (t, *J* = 7.1 Hz, 3H) ppm. LC-MS $(M+H)^+$ = 337.09, $(M-CN)^+$ = 310.09.

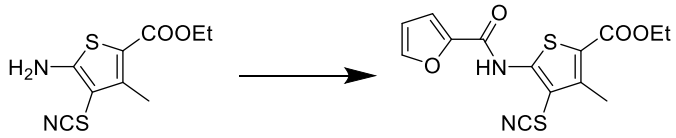


Figure 2.30. Synthesis of ethyl 5-(furan-2-carboxamido)-3-methyl-4-thiocyanatothiophene-2-carboxylate. The synthetic scheme to for CBR-5884 is depicted.

***Tert*-butyl 4-cyano-3-methylbut-3-enoate:**

Adapted from literature procedure (39). 5.00 g (36.1 mmol, 1 eq.) *tert*-butyl acetoacetate, 2.69 g (31.6 1 eq.) cyanoacetic acid, and 488 mg (6.33 mmol, 0.2 eq.) ammonium acetate were added to a solution of 0.9 mL acetic acid and 9.0 mL benzene (Figure 2.31). The reaction mixture was refluxed in a Dean-Stark trap for 24 hours. After cooling to room temperature, the reaction mixture was washed with saturated sodium bicarbonate, brine, then dried with sodium sulfate and condensed *in vacuo*. The crude product was vacuum-distilled at 65°C

to obtain 1.28g *tert*-butyl 4-cyano-3-methylbut-3-enoate as a mixture of E/Z isomers, 23% yield, clear oil. ^{1}H NMR (400 MHz $CDCl_3$) δ 5.30 (m, 2H), 3.36 (d, *J* = 1.5 Hz, 1H), 3.11 (d, *J* = 1.1 Hz, 1H), 2.15 (d, *J* = 1.1 Hz, 1H), 2.02 (d, *J* = 1.5 Hz, 1H), 1.48 (m, 18H) ppm.

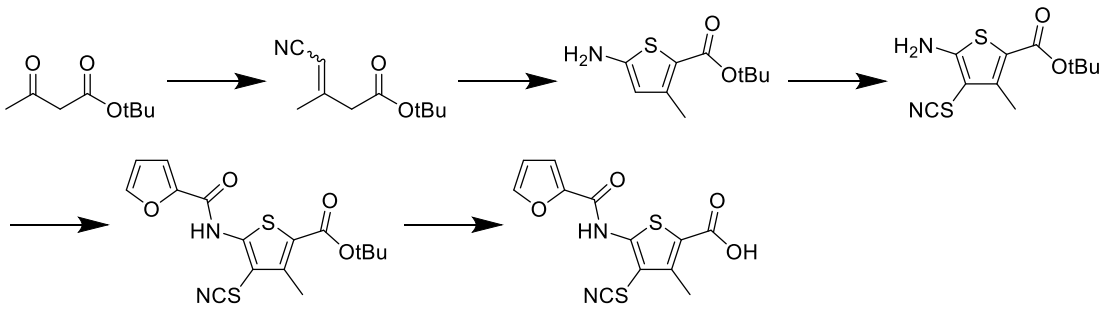


Figure 2.31. Synthetic scheme for 5-(furan-2-carboxamido)-3-methyl-4-thiocyanatothiophene-2-carboxylic acid. The synthetic scheme to obtain the acid derivative of CBR-5884 is provided.

***Tert*-butyl 5-amino-3-methylthiophene-2-carboxylate:**

Adapted from literature procedure (39). 1.28 g (7.09 mmol, 1 eq.) *tert*-butyl 4-cyano-3-methylbut-3-enoate and 227 mg (7.09 mmol, 1 eq.) sulfur flakes were suspended in 5.5 mL ethanol. 806 µL (7.80 mmol, 1.1 eq.) diethylamine was added drop-wise and the mixture was stirred for 4 hours, forming a dark orange/red solution (Figure 2.31). Reaction mixture was dried *in vacuo* and purified by flash chromatography. After drying collected fractions *in vacuo*, 814 mg *tert*-butyl 5-amino-3-methylthiophene-2-carboxylate was obtained as a

yellow-orange oil, 54% yield. ^{1}H NMR (400 MHz $CDCl_3$) δ 5.99 (s, 1H), 4.01 (br. s, 2H), 2.40 (s, 3H), 1.552 (s, 9H) ppm.

***Tert*-butyl 5-amino-3-methyl-4-thiocyanatothiophene-2-carboxylate:**

Adapted from literature procedure (40). 100 mg (0.469 mmol, 1 eq.) *tert*-butyl 5-amino-3-methylthiophene-2-carboxylate and 71.4 mg (0.938 mmol, 2 eq.) ammonium thiocyanate were dissolved in 0.5 mL methanol and cooled with an ice bath (Figure 2.31). A solution of 14.9 µL (0.291, 0.62 eq.) bromine in methanol was added drop-wise to the reaction mixture. The mixture was stirred on ice for 10 minutes; then the ice bath was removed and the mixture stirred for another hour while warming to room temperature. The reaction mixture was then dried *in vacuo* and purified by flash chromatography. The collected fractions were dried *in vacuo* to afford 32.0 mg *tert*-butyl 5-amino-3-methyl-4-thiocyanatothiophene-2-carboxylate as a white solid. ^{1}H NMR (400 MHz $CDCl_3$) δ 5.05 (br. s, 1H), 2.59 (2, 3H), 1.57 (s, 9H) ppm.

***Tert*-butyl 5-(furan-2-carboxamido)-3-methyl-4-thiocyanatothiophene-2-carboxylate:**

Prepared analogously as ethyl 5-(furan-2-carboxamido)-3-methyl-4-thiocyanatothiophene-2-carboxylate. From 32.0 mg (0.119 mmol) *tert*-butyl 5-amino-3-methyl-4-thiocyanatothiophene-2-carboxylate as starting material, 19.8 mg *tert*-butyl 5-(furan-2-carboxamido)-3-methyl-4-thiocyanatothiophene-2-carboxylate was obtained as a white solid, 46% yield (Figure 2.31). ^{1}H NMR (400

MHz $CDCl_3$) δ 9.35 (s 1H), 7.68 (dd, *J* = 0.8, 1.7 Hz, 1H), 7.42 (dd, *J* = 0.8, 3.6 Hz, 1H), 6.68 (dd, *J* = 1.8, 3.6 Hz, 1H), 2.72 (s, 3H), 1.60 (s, 9H) ppm.

5-(furan-2-carboxamido)-3-methyl-4-thiocyanatothiophene-2-carboxylic acid:

0.25 mL trifluoroacetic acid was added slowly to a solution of 5.3 mg (13.7 μmol) *tert*-butyl 5-(furan-2-carboxamido)-3-methyl-4-thiocyanatothiophene-2-carboxylate in 0.25 mL dichloromethane in an ice bath and stirred for 30 minutes, allowing the mixture to warm to room temperature over time (Figure 2.31). The reaction mixture was dried *in vacuo* with minimal heating applied (no greater than 30°C). The dried product was sonicated in diethyl ether to dissolve remaining starting material, centrifuged briefly to precipitate the product, then decanted to remove starting material in the ether layer. This ether treatment was repeated two more times. After drying the remaining solid *in vacuo*, 1.6 mg 5-(furan-2-carboxamido)-3-methyl-4-thiocyanatothiophene-2-carboxylic acid was obtained as a white solid, 38% yield. ^{1}H NMR (400 MHz $CDCl_3$) δ 7.89 (dd, *J* = 0.8, 1.8 Hz, 1H), 7.46 (dd, *J* = 0.8, 3.6 Hz, 1H), 6.75 (dd, *J* = 1.8, 3.6 Hz, 1H), 2.71 (s, 3H) ppm. LC-MS $(M+H)^+$ = 309.0, $(M-OH)^+$ = 291.0, $(M-CN)^+$ = 282.0.

Acknowledgements:

We thank U. Oppermann, MG. Vander Heiden, KR. Mattaini, M. Yuan, for technical assistance and reagents. We thank J. Johnson, Y. Zheng, H. Shim, BD.

Ngo, and other Cantley lab members for helpful discussions. Lewis C. Cantley was supported by NIH grants P01 CA117969 and P01CA120964. Conflicts of interest: Lewis C. Cantley owns equity in, receives compensation from, and serves on the Board of Directors and Scientific Advisory Board of Agios Pharmaceuticals. Agios Pharmaceuticals is identifying metabolic pathways of cancer cells and developing drugs to inhibit such enzymes in order to disrupt tumor cell growth and survival.

Attributions:

Edouard Mullarky, Luke L. Lairson, and Lewis C. Cantley designed the study. Edouard Mullarky and Luke L. Lairson developed the screening assay and Edouard Mullarky, Luke L. Lairson, and Natasha C. Lucki executed the screen. Edouard Mullarky, Natasha C. Lucki, Reza Beheshti Zavareh, Jenny C.Y. Wong performed enzymatic assays to assess selectivity and potency of inhibitors. Edouard Mullarky, Jenny C.Y. Wong, Ana P. Gomes performed cell based proliferation assays with inhibitor or knockdown. Edouard Mullarky, John M. Asara, Brandon N. Nicolay performed GC or LCMS assays. Luke L. Lairson and Justin Anglin designed and executed chemistry. Edouard Mullarky and Jenny C.Y. Wong performed protein cross-linking assays. Justin Anglin performed glutathione drug adduct experiments.

Chapter 3: Unresolved Questions, Discussion, and Future Directions

The work presented in this thesis has uncovered a new small molecule inhibitor of PHGDH. CBR-5884 will be a useful probe to further study the fundamental biology of serine synthesis both in cancer and non-transformed cells. Furthermore, CBR-5884 provides starting chemical matter from which more potent inhibitors of PHGDH can be developed.

In chapter 2, we described an *in vitro* enzymatic assay for PHGDH amenable to high throughput screening and its use to screen a library of 800,000 drug-like compounds for inhibitors of PHGDH. Following confirmation and counter screening, we were able to identify 408 bona fide PHGDH inhibitors. We triaged the commercially available subset of hits based on a combination of cell based and biochemical enzymatic assays. Thus, we were able to identify one compound, CBR-5884, that was selective for PHGDH in a panel of metabolic dehydrogenases and inhibited serine synthesis in cancer cells. Importantly, CBR-5884 was able to selectively inhibit the proliferation of breast cancer and melanoma cell-lines that over-express PHGDH while having minimal effects on lines that rely on import of extracellular serine. These results phenocopied those obtained with PHGDH knockdown in breast and melanoma cell lines, which is consistent with an on-target effect of CBR-5884. Biochemical characterization of CBR-5884 revealed that it was a non-competitive inhibitor that showed a time dependent onset of inhibition and disrupted the oligomerization state of PHGDH. Overall, the results suggest that targeting PHGDH in cancers addicted to serine

synthesis may be a promising therapeutic strategy. Nevertheless, there are some important limitations to CBR-5884 and further work is needed.

Relatively high concentrations of CBR-5884 were needed to inhibit PHGDH overexpressing cancer cell proliferation and the inhibition of serine labeling was relatively mild (~30%) at the maximal CBR-5884 dose used (30μM). Although we tried to investigate and rule out potential off-target effects, given the high concentrations of CBR-5884, it is possible that part of the anti-proliferative activity observed results from off-target effects. Using metabolomics to compare metabolic changes following drug treatment and PHGDH knockdown, we identified glycerol-3-phosphate as a metabolite uniquely perturbed following CBR-5884 treatment. Thus, we are currently investigating whether glycerol-3-phosphate dehydrogenase is inhibited by CBR-5884.

An additional limitation to CBR-5884 is that it is unstable in mouse plasma (data not shown) and that the acid derivative is unable to inhibit cancer cell proliferation presumably because it cannot cross the plasma membrane. The poor pharmacokinetic properties of CBR-5884 will make it difficult to study the effect of PHGDH inhibition in murine cancer models. However, both the limited potency and poor pharmacokinetic properties could be ameliorated via medicinal chemistry efforts. To this end we are continuing to expand the SAR study around the CBR-5884 scaffold. Finally, an additional potential caveat is that CBR-5884 may well be a covalent inhibitor of PHGDH. A covalent inhibition mechanism would be consistent with the noncompetitive mode of inhibition, the importance of

the thiocyano group in the SAR study, and the time dependent inhibition observed.

We speculated that the thiocyano group could enable CBR-5884 to form an adduct with PHGDH cysteine sulfhydryl groups. In order to evaluate this possibility, the *tert*-butyl derivative of CBR-5884 (**11**) was incubated with 5mM glutathione (1:1 molar ratio) and analyzed by LCMS. Following overnight incubation, approximately 80% of compound **11** had been converted to a more polar compound with mass attributable to the glutathione adduct with loss of cyanide (Supplementary Figure 3.1). To confirm that the species was indeed the predicted adduct, an excess of dithiothreitol was added. This resulted in near complete recovery of a compound consistent with the free thiol by mass spectrometry (Supplementary Figure 3.1). These results suggest that CBR-5884 might be covalently reacting with PHGDH cysteine sulfhydryl groups and that CBR-5884 may be inactivated by glutathione conjugation in cells. It should be noted, however, that the rate of adduct formation is likely to be greatly overestimated in the above assay compared to the intracellular process given that the reaction is bimolecular and both compound **11** and glutathione were used at high concentrations (5mM).

In order to further investigate the potential covalent mechanism of inhibition, we examined the crystal structure of the truncated protein, $PHGDH^{3-314}$ (PDB: 2G76). Although there are no cysteine residues directly within the active site that can directly account for loss of PHGDH activity, there are a number of cysteines (C109, C232, C279, C293) that could potentially perturb enzymatic

function were they modified given their close proximity to the active site (Figure 3.1). PHGDH has been reported to rely on histidine and arginine residues for catalysis [96,97]. The histidine residue serves as a catalytic base to deprotonate the 3-PG hydroxyl group and the arginine residue coordinates the 3-PG carboxylic acid (Figure 3.1). Modification of C279 by CBR-5884 could potentially displace the catalytic histidine residue (H281) thereby inhibiting catalysis. Similarly, C232 modification could perturb R234 coordination of 3-PG. Phosphorylation of PHGDH threonine 77 by PKC zeta has been suggested to inhibit PHGDH activity, supporting the notion that modification of the cysteine residues could perturb enzymatic function (Figure 3.1) [98].

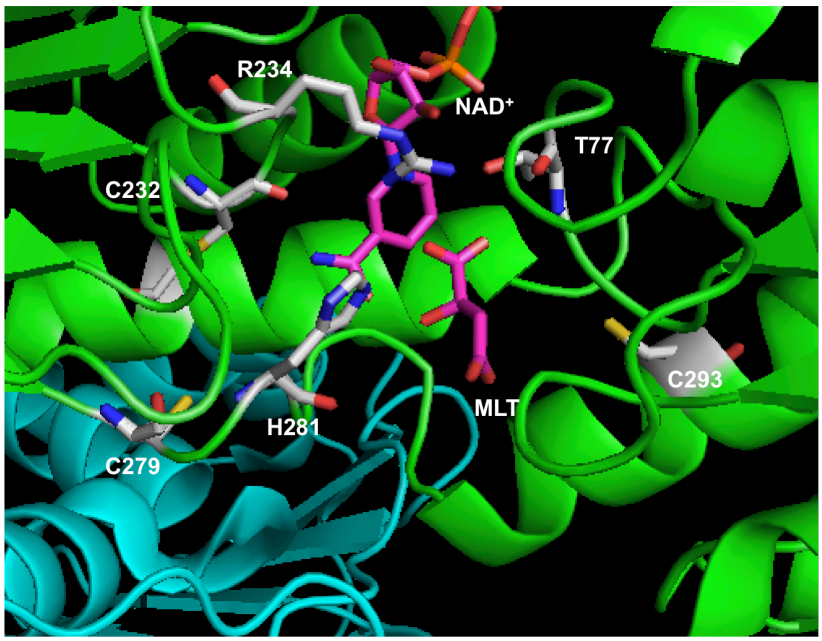


Figure 3.1: Structure of the PHGDH$^{3-314}$ active site. PHGDH catalyzes the oxidation of the 3-PG hydroxyl group to a ketone. The NAD^+ cofactor and 3-PG substrate analog malate (MLT) are shown in purple. PHGDH residues discussed are shown in grey. H281 catalytic base that abstracts the hydroxyl proton enabling ketone formation with a hydride transfer to NAD^+. R234 coordinates the 3-PG carboxylic acid. Cysteine residues that could potentially be modified by CBR-5887 to inhibit PHGDH activity are shown. The putative AGC kinase site, T77 putative, that is phosphorylated *in vitro* by PKC zeta inhibiting PHGDH activity is shown [98]. PHGDH$^{3-314}$ structure, PDB: 2G76.

Covalent modification of the cysteines proximal to the PHGDH active site could account for the time dependent inhibition of PHGDH and for why both the full length and truncated protein are inhibited. However, it would not account for the disruption of the tetrameric form of the full-length protein. Inspection of the PHGDH primary amino acid sequence indicates that there are cysteine residues in the full-length protein not present in the truncated protein. Were these cysteine

residues located at critical regions of the dimer-dimer interfaces, they could perhaps account for disruption of the tetramer.

In order to determine whether CBR-5887 is functioning though a covalent mechanism targeting cysteine residues, three parallel approaches are being followed. First, site directed mutagenesis is being pursued to determine whether point mutation of any of the cysteine residues can rescue from CBR-5884 inhibition and/or rescue disruption of the tetrameric form of PHGDH. Second, we are attempting to map the cysteine residues potentially modified by CBR-5884 via an LCMS based proteomic approach. Third, we are investigating whether thiol blocking reagents such as iodoacetamide and N-ethylmaleimide (NEM) can also inhibit PHGDH and disrupt the PHGDH oligomerization state.

It is important to determine whether CBR-5884 is a covalent modifier because this will determine how best to optimize the molecule. Covalent modifiers are traditionally undesirable in drug discovery because of potential toxicity issues (eg. hapten formation). Moreover, covalent inactivation is often considered a non-drug like mechanism of action in that the drug may be promiscuous and react with other proteins using the same covalent mechanism [99]. Nevertheless, covalent drugs are widely and safely used in the clinic (eg. beta-lactam antibiotics like penicillin) and recently there has been renewed interest in developing covalent kinase inhibitors in oncology [100]. In order to optimize selectivity, the current covalent drug design strategy focuses on developing probes that form a strong initial non-covalent complex with the enzyme, with the covalent warhead only reacting after the probe is bound and

with a specific residue that has the right geometric orientation with respect to the bound drug [99,100].

Given the design strategy of covalent probes discussed, one potential concern is that the CBR-5884 thiocyano group is non-specific in its mechanism of covalent inhibition. The dehydrogenase enzyme panel used to profile CBR-5884 selectivity (Table 2) may have lacked enzymes that contain an active site or otherwise important cysteine susceptible to CBR-5884 modification and may therefore have failed to identify the promiscuity of CBR-5884. In order to assess whether the presence of cysteine residues in close proximity to the active site was a unique feature of PHGDH, we examined the crystal structure of LDH (PDB: 9LDT). Analogous to PHGDH, the LDH structure revealed a catalytic histidine residue and arginine residues that coordinate the substrate carboxylic acid [101-103]. Importantly, as with PHGDH, LDH cysteine residues (C35 & C165) were found in close proximity to the active site. Given that CBR-5884 did not inhibit LDH it suggests that CBR-5884 might harbor some specificity for PHGDH. Moreover, given that the CBR-5884 derivative in which the thiocyano group is replaced with a cyano group (**9**) still inhibits PHGDH, it would suggest that CBR-5884 harbors some specificity for PHGDH, at least to the extent that formation of a relatively weak non-covalent complex is indicative of specificity.

In order to further validate that the mechanism of CBR-5884 inhibition is specific despite its cysteine targeting mechanism, we will profile the activity of CBR-5884 against other enzymes, such as GAPDH, that are known to rely on a catalytic cysteine residue for activity [33]. It should be noted, however, that

metabolic profiling of cells treated with CBR-5884 did not reveal perturbations of glycolytic flux downstream of GAPDH (Figure 2.16) as would be anticipated were the GAPDH catalytic cysteine targeted [6]. From a medicinal chemistry perspective, given that compound **9** only binds weakly to PHGDH, it would be important to improve the binding affinity of the non-covalent CBR-5884 scaffold and investigate whether the thiocyano moiety can be replaced with other covalent warheads such as an acrylamide moiety. For covalent kinase inhibitors, covalent warheads are typically only introduced after submicromolar non-covalent binding affinity has been achieved [104]. While CBR-5884 treatment only induces modest inhibition of serine labeling in cells, improved non-covalent binding of the CBR-5884 scaffold and modifications of the covalent warhead may lead to more selective and potent inhibition of PHGDH in cells.

Given that in pursing the results from the PHGDH inhibitor screen we focused on hits that were commercially available, we were only able to pursue 45 of the 408 PHGDH inhibitors identified. Thus, we are planning to synthesize some of the top compounds from the hit list in order to evaluate them in the cell and biochemical based assays developed. Top hits will be selected for resynthesis based on observed potency, selectivity, drug-like properties, and number of occurrences of the identified scaffold in the hit list. Using this approach we may be able to identify more potent and drug like scaffolds than that of CBR-5884.

Discovering that CBR-5884 induced dissociation of the tetrameric form of PHGDH to favor the dimeric form led us to ponder whether PHGDH activity could

be regulated by transitions between the tetrameric and dimeric forms in cells and whether disruption of the PHGDH tetramer was part of the inhibitory mechanism of CBR-5884. Precedence for metabolic enzymes being regulated in such a manner has been established for pyruvate kinase M2: an endogenous metabolite (fructose-1,6-bisphosphate), and pharmacological small molecule activators enhance PKM2 activity by stabilizing the tetrameric form over the less active dimeric form [92,93]. PHGDH is a tetramer formed from a dimer of dimers. The primary dimer is formed from interactions between the nucleotide binding domain of two monomers; tetramers then form by dimers interacting through their C-terminal domains (CTD). In *E.coli*, PHGDH is allosterically inhibited by serine via a V_{max} effect and the mechanism of inhibition involves serine binding to the interface of two CTDs [105]. While serine based feedback inhibition is not conserved in vertebrates, these observations suggest that it is possible that human PHGDH may also be regulated via modulation of the CTD and that this could lead to the disruption of the tetramer [78]. Consistent with the notion that disruption of the tetrameric form of PHGDH could be inhibitory, we noted that truncated $PHGDH^{3-314}$, which lacks the CTD and is therefore a constitutive dimer, had a specific activity ~5-fold lower than that of the full-length protein (Supplementary Figure 3.2). Furthermore, we noted that a significant fraction of patients with inherited disorders in serine synthesis are known to have missense mutations in the PHGDH CTD [106]. The mechanism by which CTD mutations disrupt PHGDH activity is still controversial but decreases in V_{max} have been reported [106-108]. These observations are consistent with a model in which PHDGH

activity may be regulated by transitions between dimer and tetrameric states. Having purified the various PHGDH CTD point mutants we are currently investigating whether their kinetic parameters and tetramer to dimer ratios differ from the wildtype protein. Preliminary experiments suggest that the mutants are somewhat less active and favor the dimeric form (data not shown), but further work is needed to validate our model. Finally, we also plan to investigate whether the endogenous activator of PHGDH, 2-phospohglycerate, could be functioning by modulating the PHGDH oligomerization state [94].

Appendix A: Identification of a Small Molecule Inhibitor of 3-Phosphoglycerate Dehydrogenase to Target Serine Biosynthesis in Cancer.

A copy of the manuscript accepted for publication in the *Proceedings of the National Academy of Sciences* is provided below.

Identification of a Small Molecule Inhibitor of 3-Phosphoglycerate Dehydrogenase to Target Serine Biosynthesis in Cancers.

Edouard Mullarky[a,b,f], Natasha C. Lucki[o], Reza Beheshti Zavareh[o], Justin Anglin[o], Ana P. Gomes[a,c], Brandon N. Nicolay[g], Jenny C.Y. Wong[a,b], Stefan Christen[j,k], Hidenori Takahashi[h,i,#], Pradeep K. Singh[d,e], John Blenis[a,c], J. David Warren[d,e], Sarah-Maria Fendt[j,k], John M. Asara[i], Gina M. DeNicola[a,b], Costas A. Lyssiotis[l,m,1,2], Luke L. Lairson[n,o,1,2], Lewis C. Cantley[a,b,1, 2]

[a]Meyer Cancer Center, [b]Department of Medicine, [c]Department of Pharmacology, [d]Department of Biochemistry, [e]Milstein Chemistry Core Facility, Weill Cornell Medical College, New York, NY 10065;

[f]Biological and Biomedical Sciences Graduate Program, Harvard Medical School, Boston, MA 02115;

[g]Massachusetts General Hospital Cancer Center, Harvard Medical School, Charlestown, MA 02129;

[h]Department of Systems Biology, [l]Division of Signal Transduction, Beth Israel Deaconess Medical Center and Department of Medicine, Harvard Medical School, Boston, MA 02115;

[j]Laboratory of Cellular Metabolism and Metabolic Regulation, Department of Oncology, KU Leuven, 3000 Leuven, Belgium;

[k]Laboratory of Cellular Metabolism and Metabolic Regulation, Vesalius Research Center, VIB Leuven, 3000 Leuven, Belgium;

[l] Department of Molecular and Integrative Physiology, [m]Department of Internal Medicine Division of Gastroenterology, University of Michigan, Ann Arbor, MI 48109;

[n]Department of Chemistry, The Scripps Research Institute, 10550, North Torrey Pines, La Jolla, California 92037.

[o]The California Institute for Biomedical Research, 11119 North Torrey Pines, La Jolla, California 92037.

[#] Current Address: Frontier Research Laboratories, Daiichi Sankyo Co., Shinagawa-ku, Tokyo 140-8710, Japan.

[1] Co-corresponding authors

[2] To whom correspondence should be addressed. E-mail: clyssiot@med.umich.edu, llairson@scripps.edu, LCantley@med.cornell.edu

Abstract

Cancer cells reprogram their metabolism to promote growth and proliferation. The genetic evidence pointing to the importance of the amino acid serine in tumorigenesis is striking. The gene encoding the enzyme 3-phosphoglycerate dehydrogenase (PHGDH), which catalyzes the first committed step of serine biosynthesis, is overexpressed in tumors and cancer cell lines via focal amplification and NRF2-mediated upregulation. PHGDH overexpressing cells are exquisitely sensitive to genetic ablation of the pathway. Here, we report the discovery of a novel and selective small molecule inhibitor of PHGDH, CBR-5884, identified by screening a library of 800,000 drug-like compounds. CBR-5884 inhibited *de novo* serine synthesis in cancer cells and was selectively toxic to cancer cell lines with high serine biosynthetic activity. Biochemical characterization of the inhibitor revealed that it was a non-competitive inhibitor that showed a time dependent onset of inhibition and disrupted the oligomerization state of PHGDH. The identification of a small molecule inhibitor of PHGDH not only enables thorough preclinical evaluation of PHGDH as a target in cancers, but also provides a tool with which to study serine metabolism.

Significance statement

Serine supports a number of anabolic processes including protein, lipid, and nucleic acid synthesis. Cells can either import serine or synthesize it *de*

novo. Recently, overexpression of *PHGDH*, the gene encoding the first committed step of serine synthesis, via focal amplification and NRF2-mediated upregulation, has been identified in human cancers. Cancer cell lines that overexpress *PHGDH* via amplification, or transcriptional upregulation, are uniquely sensitive to PHGDH knockdown while lines that express little PHGDH are insensitive suggesting that PHGDH may be a clinically interesting target. Here, we report a novel and specific small molecule inhibitor of PHGDH, which enables preclinical evaluation of PHGDH as a target in cancer and provides a tool to study the biology of *de novo* serine synthesis.

Introduction

Serine is required for a plethora of anabolic processes. Serine is an abundant component of proteins and is required for the synthesis of lipids including sphingolipids and phosphatidylserine, a major component of cellular membranes [1-3]. Alternatively, serine hydroxymethyltransferases (SHMTs) convert serine to glycine concomitantly charging the folate pool with “one-carbon” units [4,5]. Both glycine and folate one-carbon units are used to make nucleotides. Thus, serine serves numerous critically important roles in cellular metabolism.

At the cellular level, serine can be imported from the extracellular space via amino acid transporters [6,7]. Alternatively, serine can be synthesized from glucose via the phosphoserine pathway [8]. *De novo* synthesis proceeds from the glycolytic intermediate 3-phosphoglycerate (3-PG) via three sequential enzymatic

reactions (Figure A.1), the first of which is catalyzed by the NAD^+ dependent enzyme 3-phosphoglycerate dehydrogenase (PHGDH) [9]. For decades, it has been known that cancer cells have enhanced serine synthesis, which contributes to nucleotide synthesis [10,11]. Recently, focal amplifications of the gene encoding *PHGDH* have been reported, particularly in breast cancers and melanomas [12-14]. Additionally, *KEAP1* and *NRF2* mutant non-small cell lung cancers (NSCLC) overexpress *PHGDH* [15]. Proliferation of *PHGDH* amplified cancer cell lines, and other lines that overexpress PHGDH without amplification, is inhibited by PHGDH knockdown. In contrast, lines that express little PHGDH are resistant to shRNA mediated ablation of the pathway presumably because serine import suffices [13,14]. A detailed mechanistic understanding of why some cancer cells are addicted to serine synthesis despite the availability of extracellular serine for import remains unclear. Interestingly, in triple negative breast cancer (TNBC) and NSCLC, *PHGDH* amplification and overexpression are associated with more aggressive disease [13-16]. Thus, PHGDH inhibitors as a targeted therapy for these tumor types represents an exciting clinical opportunity.

The studies herein detail our efforts in identifying small molecule inhibitors of PHGDH. We reasoned that a PHGDH inhibitor would have the benefits of not only providing a tool compound with which to study the biology of serine synthesis, but also enable thorough preclinical evaluation of PHGDH as a target in cancers. We screened a library of 800,000 small molecules using an *in vitro* PHGDH assay. A cell-based assay for serine synthesis was used to identify a lead, CBR-5884, that was active in cells. CBR-5884 selectively inhibited the

proliferation of melanoma and breast cancer lines that have a high propensity for serine synthesis, but had no effect on lines that rely on extra-cellular serine uptake. Mechanistically, CBR-5884 was found to be a noncompetitive inhibitor, showed a time dependent onset of inhibition, and disrupted the oligomerization state of PHGDH.

Results

Screening for small molecule inhibitors of PHGDH.

An *in vitro* enzymatic assay for PHGDH activity amenable to high-throughput screening (HTS) was developed by coupling the production of NADH, upon 3PG oxidation, to the reduction of resazurin to resorufin using diaphorase as the coupling enzyme (Figure A.1). Thus, resorufin fluorescence served as a proxy for PHGDH activity. The assay was miniaturized to 1536 well format with a Z-factor of >0.75 indicating a high quality assay [17]. A library of 800,000 small molecules was screened in single point format at 13μM (Figure A.1). Setting a threshold Z score of -3, corresponding to at least 50% PHGDH inhibition, gave a 0.5% hit rate yielding 3,906 hits. Putative hits were re-assayed in triplicate and counter screened against diaphorase to rule out false positives targeting diaphorase. The counter screen eliminated 3,498 compounds giving 408 PHGDH inhibitors (Figure A.1).

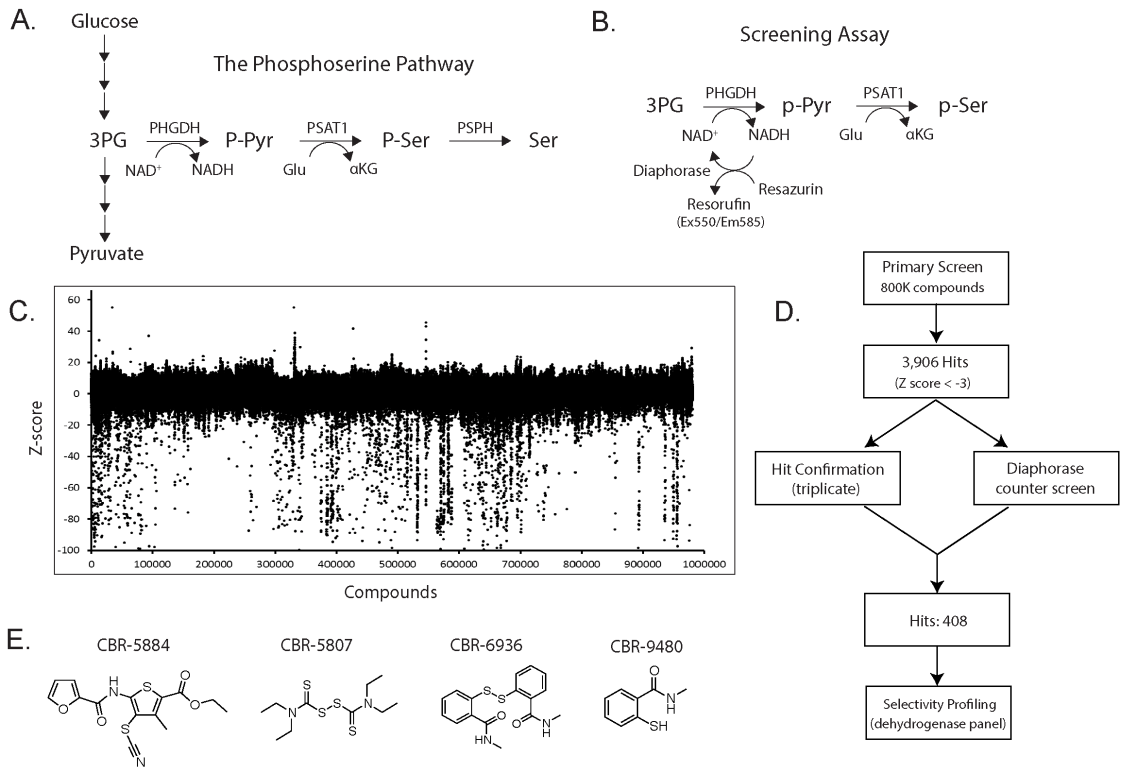


Figure A.1. Screening for inhibitors of PHGDH. (A) Serine synthesis from glucose via the phosphoserine pathway: phosphoglycerate dehydrogenase (PHGDH) oxidizes the glycolytic intermediate 3-phosphoglycerate (3-PG) to 3-phosphohydroxypyruvate (p-Pyr) using NAD^+; phosphoserine amino transferase (PSAT1) transaminates p-Pyr to phosphoserine (p-Ser) using glutamate as a nitrogen donor; phosphoserine phosphatase (PSPH) dephosphorylates p-Ser to yield serine. (B) *In vitro* PHGDH assay. Diaphorase couples the NADH produced upon PHGDH turnover to the reduction of resazurin to fluorescent resorufin. Resorufin fluorescence is a proxy for PHGDH activity. PSAT1 is included to prevent product feedback inhibition of PHGDH by p-Pyr. (C) Z-score plot for the 800,000 compound library screened using above PHGDH assay. Each point represents a single compound. A negative score indicates inhibition. (D) Screen triaging strategy. Setting a Z score threshold of -3 gave 3,906 putative hits. After counter-screening against diaphorase to rule out false positives and confirming activity against PHGDH 408 compounds remained. Selected compounds were profiled against a panel of metabolic $NAD(P)^+$ dehydrogenases to ascertain selectivity for PHGDH. (E) Structures of representative PHGDH inhibitors evaluated in cell-based assays.

A triaging strategy based on hit potency and selectivity was designed. We reasoned that inhibitors specific to PHGDH would minimize general cellular toxicity compared to compounds that hit a variety of dehydrogenases. Thus, half maximal inhibitory concentrations (IC_{50}) were determined for a panel of $NAD(P)^+$ dependent dehydrogenases that included PHGDH, isocitrate dehydrogenase (IDH1), malate dehydrogenase (MDH1), and 3α-hydroxysteroid dehydrogenase (3α-HSD). Compounds at least 4-fold more selective for PHGDH were progressed for further analysis. Based on this triaging, seven of the most potent PHGDH inhibitors were selected as lead compounds for evaluation in cell-based assays; selected structures are shown (Figure A.1). A number of these compounds are likely to target sulfhydryl groups and may therefore react with a PHGDH cysteine residue. For example, both CBR-5807 and CBR-6936 contain sulfhydryl reactive disulfide centers. Interestingly, CBR-5807 (Disulfiram) is an approved drug dosed in humans to treat alcoholism and known to inhibit aldehyde dehydrogenase by reacting with sulfhydryl groups [18].

CBR-5884 inhibits serine synthesis in cells.

We determined whether any of our seven leads inhibited serine synthesis in cancer cells. To do so, we turned to gas chromatography mass-spectrometry (GCMS) with uniformly carbon-13 labeled glucose ($^{13}C_6$-glucose) tracing. Given the isotopic enrichment of serine, it is possible to decouple newly synthesized serine from extracellular serine or serine that was synthesized prior to tracer

addition. Newly synthesized serine has a mass-shift of 3 (M+3) due to the incorporation of glucose-derived ^{13}C via 3-PG. We first investigated the kinetics of serine labeling. Serine labeling plateaued around 6h with ~65% of the serine pool being ^{13}C labeled. The plateau phase likely reflects exchange between intra- and extra-cellular serine pools [19].

With an understanding of serine labeling kinetics, we designed a $^{13}C_6$-glucose tracing assay to acutely interrogate the effects of compounds on serine synthesis (Figure A.2). Assaying serine synthesis with a 3-hour compound treatment was preferred to longer treatments to guard against false positives that decrease serine labeling by an indirect effect, such as generally compromising cellular viability. Among our lead compounds, CBR-5884 was able to decrease *de novo* serine synthesis by 30%; the remaining compounds had little effect (Figure A.2). The dose at which CBR-5884 had an effect on serine labeling was consistent with the *in vitro* biochemical IC_{50} of 33 ±12μM for PHGDH (Figure A.2). At such concentrations, CBR-5884 had no effect on two other NAD^+ dependent dehydrogenases, lactate dehydrogenase (LDH) and MDH1 (Figure A.2). Importantly, under the acute treatment time period used in the labeling assays, CBR-5884 was not generally cytotoxic at concentrations up to 40μM as determined by two independent cellular viability assays. Hence, decreases in serine labeling are a direct effect of CBR-5884 mediated PHGDH inhibition.

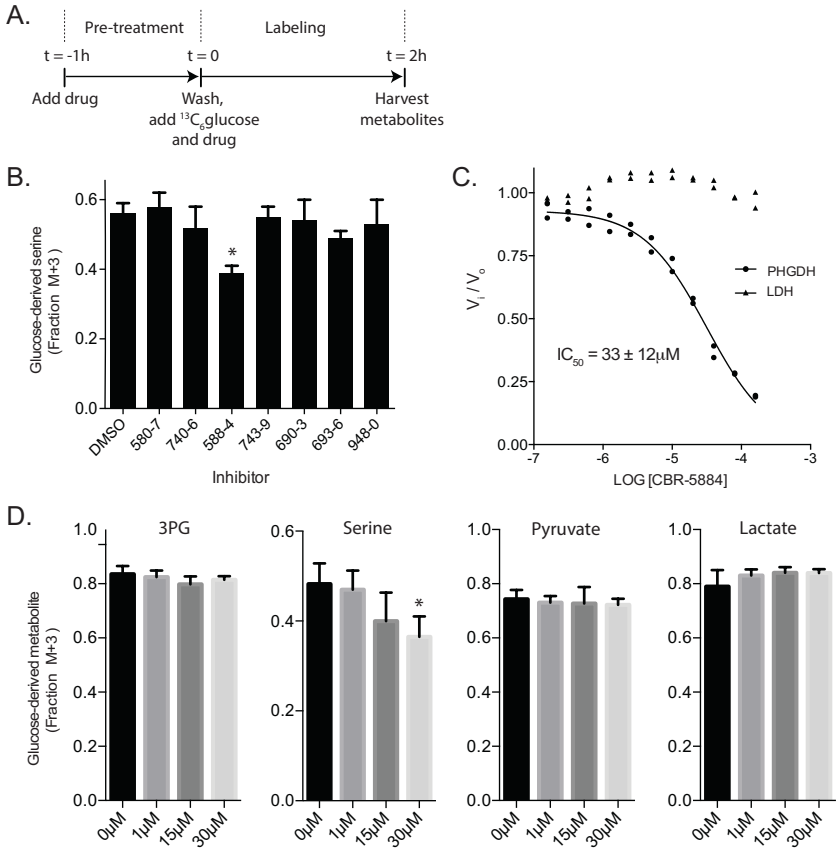


Figure A.2. CBR-5884 inhibits serine synthesis in cells. (A) Acute inhibitor treatment assay schematic. Carney cells are pre-treated with drug at 30μM for 1h prior to initiating $^{13}C_6$-glucose labeling for 2h still in the presence of drug. Polar metabolites are harvested and analyzed by GC-MS. (B) The ability of the 7 lead PHGDH inhibitors to block serine synthesis was assayed as in (A). $^{13}C_6$-glucose derived serine (M+3 serine) relative to total serine levels are plotted. (C) *In vitro* IC_{50} assays for PHGDH and lactate dehydrogenase, LDH. Initial rates of the enzymatic reaction (V_i) at the indicated CBR-5884 concentration normalized to that of the DMSO control (V_o) are plotted. (D) CBR-5884 dose response experiment as in (A) but monitoring a panel of phosphoserine pathway and glycolytic metabolites. Y-axes indicate the fraction of the indicated metabolite derived from glucose (M+3 metabolite level relative to total metabolite level). Asterisks indicate significant differences vs. DMSO treatment ($p < 0.05$, t-test, $n \geq 3$). Error is given as ± 1SD.

We resynthesized CBR-5884 in house and performed a dose response experiment for CBR-5884 employing the same acute treatment method as above. Serine labeling was significantly decreased at 30μM and trended towards a decrease at 15μM (Figure A.2). Importantly, perturbations in labeling were specific to serine in that neither the PHGDH substrate, 3-PG, nor the end products of glycolysis, pyruvate and lactate were affected (Figure A.2). We further confirmed that glycolytic metabolites were unperturbed by CBR-5884 treatment using LC-MS/MS to interrogate a greater panel of metabolites. Thus, changes in serine labeling are a direct effect of CBR-5884 mediated PHGDH inhibition and not a consequence of changes in PHGDH substrate levels or general perturbations in glycolytic flux (Figure A.2). The absence of an effect on lactate labeling was consistent with the *in vitro* data showing that CBR-5884 does not inhibit LDH under the drug concentrations used. In sum, the data argue that CBR-5884 is able to selectively inhibit serine synthesis in cells.

Given that CBR-5884 is an ethyl ester and therefore susceptible to intracellular esterases, we investigated whether the carboxylic acid derivative of the parent molecule was still active against PHGDH; were the acid less active, it would likely decrease the efficiency of targeting PHGDH *in situ*. Parent and acid derivatives were equally potent against and selective for PHGDH *in vitro* suggesting that intra-cellular de-esterification is unlikely to affect CBR-5884 activity.

CBR-5884 selectively inhibits the proliferation of cancer cells with a high propensity for serine synthesis.

We established a system to test the ability of CBR-5884 to inhibit PHGDH-dependent cancer cell proliferation. We first evaluated the ability of a panel of breast and melanoma cell lines to proliferate in serine replete or deplete media as a proxy for serine biosynthetic activity. Breast lines were selected based on *PHGDH* expression according to the Cancer Cell Line Encyclopedia (CCLE) data and validated by blotting for PHGDH (Figure A.3) [20]. Removing extracellular serine had no effect on proliferation of high *PHGDH* expressing lines: MDA-MB-468, MDA-MB-436, HCC70, and Hs578T (Figure A.3). All four lines cluster in the top quartile of the CCLE data set for *PHGDH* expression; MDA-MB-468 and HCC70 cells harbor *PHGDH* amplifications [14]. In contrast, serine depletion almost completely abrogated proliferation of low *PHGDH* expressing lines: MDA-MB-231 and MCF10A (Figure A.3). In melanoma cells, PHGDH protein levels were similarly commensurate with the ability to proliferate in serine free media. Interestingly, although Carney cells are sensitive to extracellular serine depletion, they can adapt and proliferate as evidenced by increased PHGDH protein levels upon serine depletion.

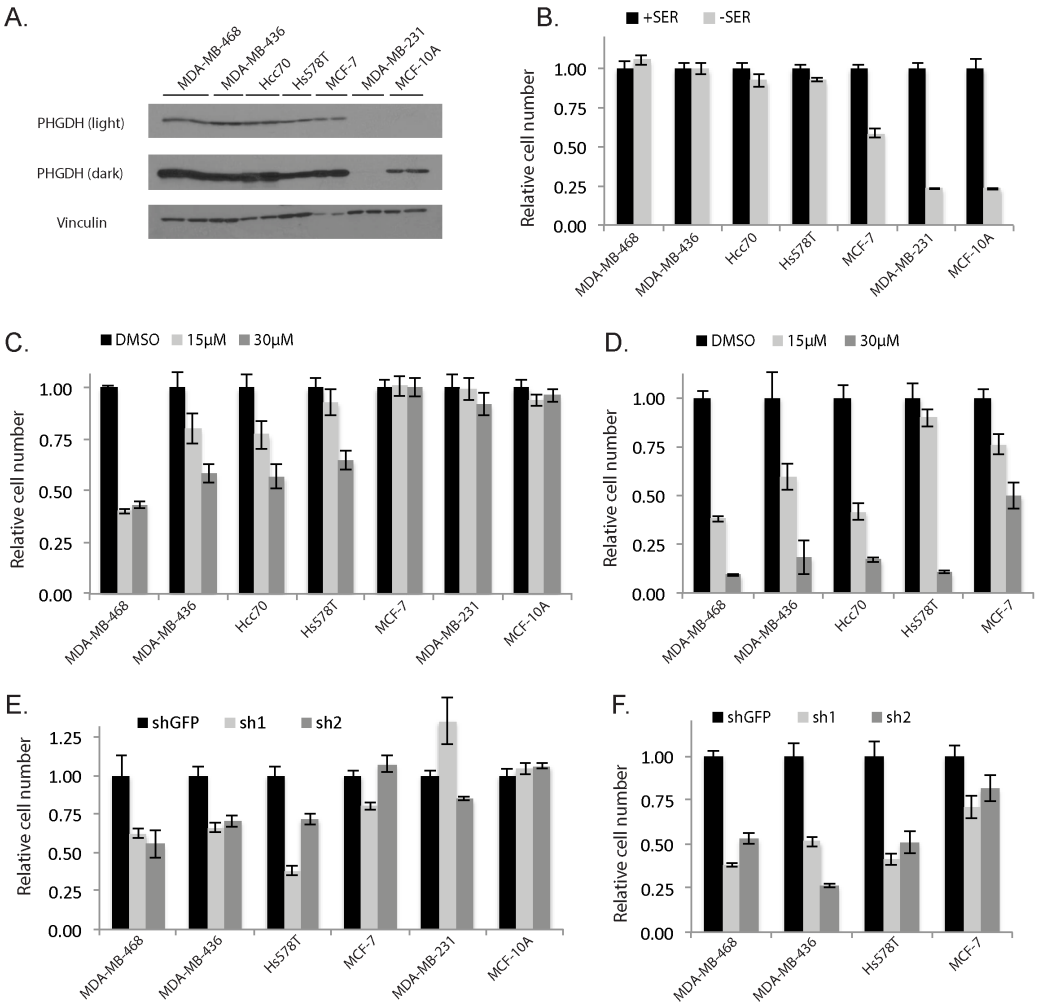


Figure A.3. CBR-5884 selectively inhibits the proliferation of breast cancer lines with a high propensity for serine synthesis. (A) Western blot for lines grown in +SER media; 2 lanes per cell line with each lane loaded with independent cell lysates. (B) Proliferation assay for breast lines grown in either serine replete (+SER) or deplete (-SER) media. Proliferation assay for lines treated with CBR-5884 in (C) +SER or (D) -SER media. Proliferation assay for lines grown in (E) +SER or (F) -SER media with PHGDH knockdown (sh1 & sh2) or a nontargeting control (shGFP). MDA-MB-468 and HCC70 are *PHGDH* amplified. MCF-10A cells are non-transformed mammary epithelial cells, other lines are cancer cell lines. MDA-MB-231 and MCF-10A lines were not included in –SER experiments in (D) and (F) because they are sensitive to serine withdrawal. Histograms depict mean ± standard error (n ≥ 3).

Given that the ability to proliferate in the absence of extra-cellular serine is indicative of a high propensity for serine synthesis, we hypothesized that such lines should be sensitive to CBR-5884. Conversely, lines that cannot grow in serine free media have a low propensity for serine synthesis and should therefore be resistant to PHGDH inhibition. Treating the breast lines with CBR-5884 in serine replete media inhibited growth of the four lines that grew without extra-cellular serine in a dose dependent manner with growth inhibition ranging from 35% to 60% at 30μM CBR-5884. The inhibitor had no effect on the three lines sensitive to serine withdrawal indicating that the inhibitor was selectively toxic to cells with high serine synthesis activity (Figure A.3). We next asked whether removing serine from the media, to enhance the reliance on *de novo* serine synthesis, could sensitize cells to PHGDH inhibition. Indeed, serine depletion increased the efficacy of CBR-5884 in lines already sensitive under serine replete conditions as evidenced by an 80% to 90% decrease in proliferation with 30μM CBR-5884 (Figure A.3). Moreover, MCF7 cells, which were of intermediate sensitivity to serine withdrawal, and insensitive to drug under serine replete conditions, became partially sensitive to the inhibitor under serine deplete conditions (Figure A.3). Importantly, under serine replete conditions, PHGDH knockdown phenocopied the effects of CBR-5884 treatment in that the drug-sensitive lines were also sensitive to PHGDH knockdown (Figure A.3). Furthermore, as with the drug treatments, growing cells in serine free media enhanced the growth inhibitory effect of PHGDH knockdown (Figure A.3). Similar trends were observed for the melanoma panel in terms of both the selectivity of

CBR-5884 for cells with a high propensity for serine synthesis and the increased efficacy under serine deplete conditions. Finally, the acid derivative of compound 588-4 was not effective on MDA-MB-468 cells likely owing to poor membrane permeability and is therefore not a viable alternative to the parent compound.

Analysis of CBR-5884 inhibition modality.

We sought to more deeply characterize the mechanism by which CBR-5884 inhibits PHGDH. Inhibition constants (K_i) for CBR-5884 with respect to each substrate were determined. CBR-5884 inhibited PHGDH in a non-competitive mode with respect to both substrates, as evidenced by a decreasing V_{max} with increasing CBR-5884 concentration. The inhibition constants were 50 ±20μM and 50 ±3μM for 3PG and NAD^+, respectively (Figure A.4). We assessed whether there was any time dependence to the onset of inhibition by varying the time period for which drug and PHGDH were pre-incubated before initiating the enzymatic reaction. CBR-5884 was progressively more potent with increasing pre-incubation time culminating in an IC_{50} of 7μM when drug and PHGDH were pre-incubated for 4 hours (Figure A.4). Intrigued by the combination of a time dependent onset of inhibition and non-competitive inhibition, the latter suggesting that CBR-5884 might be binding to an allosteric pocket, we speculated that CBR-5884 could be affecting the PHGDH oligomerization state, where the time dependency of inhibition could potentially stem from drug-induced conformational changes in PHGDH. To evaluate the PHGDH oligomerization state, we incubated

PHGDH with drug and then cross-linked prior to SDS-PAGE. CBR-5884 shifted the PHGDH equilibrium from the tetrameric to the dimeric state (Figure A.3). No such effect was observed with LDH which is resistant to CBR-5884 mediated inhibition. CBR-5884 still inhibited a truncated form of PHGDH, which lacks the C-terminal domain responsible for tetramerization and is therefore a constitutive dimer. Together these results suggest that disruption of the tetramer might assist PHGDH inhibition, but is not necessary for inhibition.

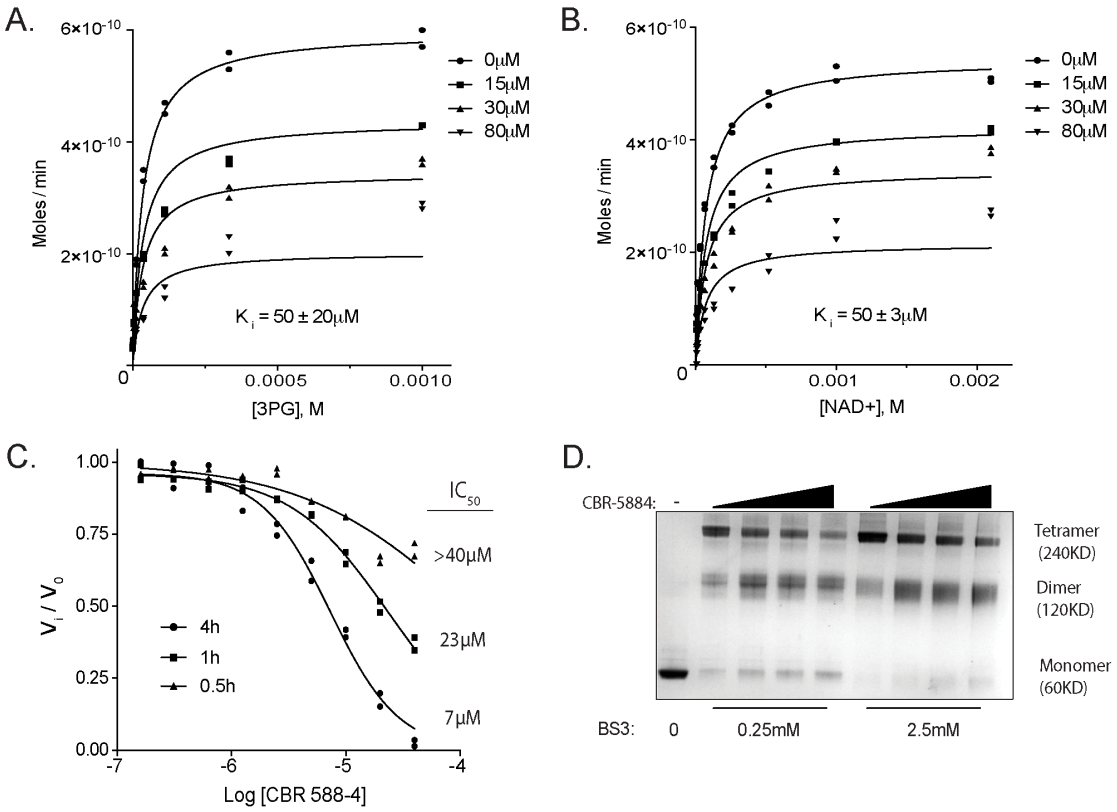


Figure A.4. Mechanisms of CBR-5884 inhibition. Inhibition constants (K_i) were determined by titrating (A) 3PG or (B) NAD^+ while holding the other substrate constant at 4 different CBR-5884 concentrations and determining the initial reaction rate using a PHGDH assay. Plots were fit to a non-competitive model. (C) Time dependent inhibition was measured by preincubating drug and PHGDH for 0.5h, 1h, or 4h as indicated prior to initiating the PHGDH reaction. Initial reaction rates (V_i) were determined and normalized to that of DMSO (V_o). (D) PHGDH was preincubated with CBR-5884 (0, 50, 200, or 400μM) prior to crosslinking with BS3 (0.25 or 2.5mM) followed by SDS-PAGE and coomassie staining. Left-most lane had no BS3 indicating the monomeric species. Oligomerization state was inferred from reference to a molecular weight ladder. Error is given as ± 1SD (n ≥ 3).

Discussion

We have reported the discovery of a novel PHGDH inhibitor, CBR-5884, and shown that it inhibits serine synthesis in cells. Furthermore, CBR-5884 specifically inhibited the proliferation of melanoma and breast cancer lines with high levels of serine synthesis activity with little effect on lines reliant on serine import. Thus, CBR-5884 is selective for lines addicted to serine synthesis and phenocopies sensitivity to PHGDH knockdown. Finally, a biochemical analysis of CBR-5884 revealed that it was a non-competitive inhibitor that showed a time dependent onset of inhibition and disrupted the oligomerization state of PHGDH.

Recent work examining how malignant cells rewire their metabolism to support growth and proliferation has revealed a number of clinically interesting targets [21,22]. Perhaps the most promising is the discovery of gain-of-function mutations in the isocitrate dehydrogenase (IDH) enzymes that result in the production of the oncometabolite 2-hydroxyglutarate [23,24]. The findings translated into chemical probes that yielded insights into the biology of IDH mutations [25,26] and led to clinical programs (eg. NCT02481154). The genetic evidence pointing to a role for PHGDH in cancer is similarly striking: PHGDH is one of few metabolic enzymes genetically deregulated in cancer [27,28]. Notably, elevated PHGDH expression correlates with clinical aggressiveness and poor prognosis in TNBC [13,16] and NSCLC [15]. There is a paucity of targeted therapies for these cancers and chemotherapies remain mainstay [29,30]. Thus, the clinical potential of PHGDH inhibitors as targeted agents for TNBC and NSCLC tumors addicted to

serine synthesis, as a single agent or in combination with standard of care, is an exciting perspective.

Beyond the preclinical applications, a PHGDH inhibitor provides a new tool to study *de novo* serine synthesis. For example, it remains unclear why a serine biosynthesis enzyme is critical for tumor growth when serine is available in the serum [13,14]. CBR-5884 provides a valuable tool complementary to genetic strategies to study this and other phenomena as small molecules provide greater temporal resolution and do not deplete the actual protein. Moreover, CBR-5884 has already suggested an interesting feature of PHGDH biochemistry, namely that human PHGDH could be regulated by transitions between different oligomerization states. These findings are reminiscent of pyruvate kinase M2 (PKM2) regulation in that both an endogenous metabolite, fructose-1,6-bisphosphate, and pharmacological small molecule activators enhance PKM2 activity by stabilizing the tetrameric form [31,32]. Endogenously, 2-phosphoglycerate (2PG) has been reported to activate PHGDH [33]. Mechanistically, 2PG could be functioning by modulating the PHGDH oligomerization state. Finally, it is possible that CBR-5884 is a covalent inhibitor of PHGDH. While PHGDH does not rely on an active site cysteine, there are a number of cysteines that could potentially perturb enzymatic function were they modified since they are in close proximity to the active site (PDB: 2G76). In an attempt to determine whether changes in PHGDH oligomerization state are unique to CBR-5884, it would be interesting to determine whether Disulfiram (CBR-5807), CBR-6936, or sulfhydryl blocking

reagents have similar effects. Mechanistic studies, including those described above, are ongoing.

The identification of a selective novel small molecule inhibitor of PHGDH capable of modulating *de novo* serine synthesis in PHGDH-dependent cancer cells represents a significant step towards the goal of targeting cancer metabolism in oncology. However, future *in vivo* evaluation of the CBR-5884 chemical series will require medicinal chemistry-based optimization. Indeed, CBR-5884 was found to be unstable in mouse plasma and, as described, replacement of the ethyl ester moiety with the corresponding negatively charged carboxylic acid resulted in a derivative that retains enzyme inhibitory activity but loses activity on cells. Thus, CBR-5884 is more likely to serve as a tool compound, and as a starting point for generating more drug-like molecules, than an actual drug. Efforts aimed at identifying cell permeable plasma-stable derivatives are ongoing. Furthermore, the cell-based potency of this series will likely need to be improved to enable *in vivo* evaluation at exposure levels that are not generally toxic. Medicinal chemistry approaches aimed at optimizing the potency and ADME properties of CBR-5884 are ongoing. Nevertheless, our results provide a proof-of-concept that small molecule inhibitors of PHGDH represent a viable new class of anti-cancer drugs.

Methods

Cells, transfections, and infections: Breast and melanoma lines were passaged in RPMI supplemented with: 10% FBS, penicillin, streptomycin, and normocin (InvivoGen). Lentivirus was produced from Lenti-X 293T cells (Clontech) transfected with packaging plasmids pCMV-dR8.2 and pCMV-VSV-G and indicated pLKO.1 shRNAs: PHGDH, TRCN0000233029 (sh1) and TRCN0000221864 (sh2); nontargeting control, TRCN0000072181 (shGFP). Experiments were performed in compliance with the Institutional Biosafety Committee. See SI Methods for details.

Immunoblots: Protein was extracted from cells via trichloroacetic acid precipitation and blotted for with primary antibodies: αPHGDH (Sigma HPA021241, 1/10000); αVinculin (Sigma V9264, 1/5000). See SI Methods for details.

Proliferation assays: Cells were plated at a low density in 96 or 24 well plates in serine containing media. The following day, media was aspirated, cells were washed with PBS, and fresh serine replete or deplete media containing drug (15μM, 30μM) or vehicle (DMSO) was added. Cells were grown for 3 to 5 days, with drug and media changed daily before assaying relative cell numbers. See SI Methods for details.

Acute drug treatments with $^{13}C_6$-glucose tracing: Carney cells acclimated to growth in MEM (Corning) were plated at 9×10^5 cells per 6cm dish the night before. The following morning, media was replaced with fresh media containing CBR-5884 (1μM, 15μM, 30μM,) or vehicle control (DMSO) for 1h. Media was then aspirated, cells were washed with PBS, and fresh glucose free MEM (Gibco) supplemented with $^{13}C_6$-glucose (3g/L, Cambridge Isotopes) and 10% dialyzed FBS containing drug or DMSO was added. After 2h, cells were quickly washed with cold PBS on ice and flash frozen. Polar metabolites were extracted as in the GCMS methods. See SI Methods for details.

Acute toxicity assay: Carney cells acclimated to growth in MEM media were plated in a 96 well plate at 6,000 cells/well. The next day, cells were treated with CBR-5884 from 1μM to 40μM for 3h. Drug containing media were then removed, fresh drug-free media added, and cell viability was determined via a CellTiter-Glo (Promega G7572) or Alamar Blue (Invitrogen DAL1025) assay according to the manufacturer's protocol. See SI Methods for details.

GCMS metabolite analysis: Polar metabolites were extracted with 2ml MeOH/H_2O (4:1) for 30min on dry ice, scraped, transferred to 2ml tubes, centrifuged (30min, 15000rpm), and the supernatants dried under vacuum. Samples were derivatized as previously described [34] and analyzed on an Agilent 6890 GC instrument. Metabolite quantification was inferred from a standard curve and fractional enrichment of ^{13}C in metabolites was corrected for the

natural abundance of ^{13}C and ^{15}N [35,36]. See SI Methods for details.

LC-MS/MS metabolite analysis: Polar metabolites were extracted and dried as in the GCMS method. Samples were resuspended in 15µl of HPLC grade water. 5µl of each sample was injected and analyzed using a 5500 QTRAP triple quadrupole mass spectrometer (AB/Sciex) coupled to a Prominence UFLC system (Shimadzu) as reported previously [37]

Protein purification: His_6-tagged pET28a PHGDH, pET28a PSAT1, and pNIC28-Bsa4 $PHGDH^{3-314}$ were purified via nickel agarose (Qiagen) from BL21 *E. coli* cultures. pVB-CBD IDH1 was purified via Macroporous Bead Cellulose capture, TEV protease (Sigma-Aldrich) digestion, and gel filtration chromatography from BL21 *E. coli* cultures. See SI Methods for details.

PHGDH assays: PHGDH activity was measured in 96-well plates by monitoring NADH fluorescence (Ex340nM/ Em460nM) over time. PSAT1 was included to prevent product inhibition of PHGDH. See SI Methods for details.

LDH and MDH1 assays: Enzyme activities were assayed using kits (LDH: Sigma MAK06; MDH1: Sigma MAK196-1KT) according to the manufacturer's instructions with commercially available recombinant enzyme (LDH: Sigma 59747; MDH1: Sigma SRP6103). Drug, titrated as for the PHGDH IC_{50} assays,

and enzyme were pre-incubated for 30min prior to initiating reaction with substrate

Cross-linking assays: PHGDH (1.5µg) or LDH (2.2µg, Sigma 59747) were incubated with CBR-5884 (50µM, 200µM, 400µM) or vehicle control (DMSO) in 25mM HEPES, pH 7.3, and 1mM NAD^+ in 18µL total volume for 30min prior to BS3 (Pierce) cross-linking and quenching. Samples were run on SDS-PAGE and colloidal coomassie stained (Bio-Rad). See SI Methods for details.

Primary PHGDH screen, diaphorase counter screen, and dehydrogenase panel selectivity profiling: 800,000 compounds were screened at a single dose (13.3µM) in 1536-plate format against PHGDH or diaphorase quantifying resorufin fluorescence (550/590nm; Ex/Em) with an Envision plate reader. Results were analyzed using Genedata Screener software. Compounds with a robust Z-score < -3 in the PHGDH screening assay and robust Z-score > -2 in the diaphorase counter screen were selected as hits. See SI Methods for detailed protocol, hit selection and confirmation, and selectivity profiling against dehydrogenase panel.

Chemical syntheses: CBR-5884, ethyl 5-(furan-2-carboxamido)-3-methyl-4-thiocyanatothiophene-2-carboxylate, and the acid derivative, 5-(furan-2-carboxamido)-3-methyl-4-thiocyanatothiophene-2-carboxylic acid, synthesis was adapted from the literature as described in SI methods [38,39].

Acknowledgements

We thank U. Oppermann, MG. Vander Heiden, KR. Mattaini, M. Yuan, for technical assistance and reagents. We thank J. Johnson, Y. Zheng, H. Shim, BD. Ngo, and other Cantley lab members for helpful discussions. L.C.C. was supported by NIH grants P01CA117969 and P01CA120964. C.A.L. was partially supported by a PanCAN-AACR Pathway to Leadership award and a Dale F. Frey award for Breakthrough Scientists from the Damon Runyon Cancer Research Foundation (DFS-09-14). G.M.D was supported by a PanCAN-AACR Pathway to Leadership award. S.M.F was supported by the Conquer Cancer Now Award, Concern Foundation. Conflicts of interest: L.C.C. owns equity in, receives compensation from, and serves on the Board of Directors and Scientific Advisory Board of Agios Pharmaceuticals. Agios Pharmaceuticals is identifying metabolic pathways of cancer cells and developing drugs to inhibit such enzymes in order to disrupt tumor cell growth and survival.

References

1. Kuge, O., Hasegawa, K., Saito, K. & Nishijima, M. Control of phosphatidylserine biosynthesis through phosphatidylserine-mediated inhibition of phosphatidylserine synthase I in Chinese hamster ovary cells. *Proc Natl Acad Sci* **95,** 4199–4203 (1998).

2. de Koning, T. J. *et al.* L-serine in disease and development. *Biochem. J.* **371,** 653–661 (2003).

3. Futerman, A. H. & Riezman, H. The ins and outs of sphingolipid synthesis.

Trends Cell Biol. **15,** 312–318 (2005).

4. Stover, P. & Schirch, V. Serine hydroxymethyltransferase catalyzes the hydrolysis of 5,10-methenyltetrahydrofolate to 5-formyltetrahydrofolate. *J Biol Chem* **265,** 14227–14233 (1990).

5. Tibbetts, A. S. & Appling, D. R. Compartmentalization of Mammalian Folate-Mediated One-Carbon Metabolism. *Annu. Rev. Nutr.* **30,** 57–81 (2010).

6. Palacín, M., Estévez, R., Bertran, J. & Zorzano, A. Molecular biology of mammalian plasma membrane amino acid transporters. *Physiol. Rev.* **78,** 969–1054 (1998).

7. Barker, G. A. & Ellory, J. C. The identification of neutral amino acid transport systems. *Exp. Physiol.* **75,** 3–26 (1990).

8. Snell, K. The duality of pathways for serine biosynthesis is a fallacy. *TIBS* 241–243 (1986).

9. Mullarky, E., Mattaini, K. R., Vander Heiden, M. G., Cantley, L. C. & Locasale, J. W. PHGDH amplification and altered glucose metabolism in human melanoma. *Pigment Cell Melanoma Res* **24,** 1112–1115 (2011).

10. Snell, K., Natsumeda, Y., Eble, J. N., Glover, J. L. & Weber, G. Enzymic imbalance in serine metabolism in human colon carcinoma and rat sarcoma. *Br. J. Cancer* **57,** 87–90 (1988).

11. Snell, K., Natsumeda, Y. & Weber, G. The modulation of serine metabolism in hepatoma 3924A during different phases of cellular proliferation in culture. *Biochemical Journal* **245,** 609–612 (1987).

12. Beroukhim, R. *et al.* The landscape of somatic copy-number alteration across human cancers. *Nature* **463,** 899–905 (2010).

13. Locasale, J. W. *et al.* Phosphoglycerate dehydrogenase diverts glycolytic flux and contributes to oncogenesis. *Nat. Genet.* **43,** 869–874 (2011).

14. Possemato, R. *et al.* Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature* **476,** 346–350 (2011).

15. DeNicola, G. M. *et al.* NRF2 regulates serine biosynthesis in non-small cell lung cancer. *Nat. Genet.* **47,** 1475–1481 (2015).

16. Pollari, S. *et al.* Enhanced serine production by bone metastatic breast cancer cells stimulates osteoclastogenesis. *Breast Cancer Res Treat* **125,**

421–430 (2011).

17. Zhang, J. H., Chung, T. D. & Oldenburg, K. R. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *Journal of biomolecular screening* **4,** 67–73 (1999).

18. Vallari, R. C. & Pietruszko, R. Human aldehyde dehydrogenase: mechanism of inhibition of disulfiram. *Science* **216,** 637–639 (1982).

19. Buescher, J. M. *et al.* A roadmap for interpreting 13C metabolite labeling patterns from cells. *Curr. Opin. Biotechnol.* **34,** 189–201 (2015).

20. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483,** 603–607 (2012).

21. Son, J. *et al.* Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway. *Nature* **496,** 101–105 (2014).

22. Tennant, D. A., Durán, R. V. & Gottlieb, E. Targeting metabolic transformation for cancer therapy. *Nat. Rev. Cancer* **10,** 267–277 (2010).

23. Dang, L. *et al.* Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* **462,** 739–744 (2009).

24. Parsons, D. W. *et al.* An integrated genomic analysis of human glioblastoma multiforme. *Science* **321,** 1807–1812 (2008).

25. Wang, F. *et al.* Targeted inhibition of mutant IDH2 in leukemia cells induces cellular differentiation. *Science* **340,** 622–626 (2013).

26. Rohle, D. *et al.* An inhibitor of mutant IDH1 delays growth and promotes differentiation of glioma cells. *Science* **340,** 626–630 (2013).

27. Mullen, A. R. & DeBerardinis, R. J. Genetically-defined metabolic reprogramming in cancer. *Trends Endocrinol. Metab.* **23,** 552–559 (2012).

28. Gottlieb, E. & Tomlinson, I. P. M. Mitochondrial tumour suppressors: a genetic and biochemical update. *Nat. Rev. Cancer* **5,** 857–866 (2005).

29. Foulkes, W. D., Smith, I. E. & Reis-Filho, J. S. Triple-negative breast cancer. *N. Engl. J. Med.* **363,** 1938–1948 (2010).

30. Boolell, V., Alamgeer, M., Watkins, D. N. & Ganju, V. The Evolution of Therapies in Non-Small Cell Lung Cancer. *Cancers (Basel)* **7,** 1815–1846 (2015).

31. Anastasiou, D. *et al.* Pyruvate kinase M2 activators promote tetramer formation and suppress tumorigenesis. *Nat. Chem. Biol.* **8,** 839–847 (2012).

32. Kung, C. *et al.* Small Molecule Activation of PKM2 in Cancer Cells Induces Serine Auxotrophy. *Chemistry & Biology* **19,** 1187–1198 (2012).

33. Hitosugi, T. *et al.* Phosphoglycerate mutase 1 coordinates glycolysis and biosynthesis to promote tumor growth. *Cancer Cell* **22,** 585–600 (2012).

34. Nicolay, B. N. *et al.* Loss of RBF1 changes glutamine catabolism. *Genes Dev.* **27,** 182–196 (2013).

35. Antoniewicz, M. R., Kelleher, J. K. & Stephanopoulos, G. Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. *Metab. Eng.* **8,** 324–337 (2006).

36. Antoniewicz, M. R., Kelleher, J. K. & Stephanopoulos, G. Elementary metabolite units (EMU): a novel framework for modeling isotopic distributions. *Metab. Eng.* **9,** 68–86 (2007).

37. Ying, H. *et al.* Oncogenic Kras maintains pancreatic tumors through regulation of anabolic glucose metabolism. *Cell* **149,** 656–670 (2012).

38. Guzi, T. *et al.* Methods for inhibiting protein kinases. 1–346

39. Ambrogi, V., Grandolini, G., Perioli, L. & Rossi, C. Convenient Synthesis of 2-Aminonaphthalene-1-thiol and 3-Aminoquinoline-4-thiol and Cyclocondensations to 1,4-Thiazino and 1,4-Thiazepino Derivatives. *Synthesis* 656–658 (1992). doi:10.1055/s-1992-26191

Appendix B: NRF2 Regulates Serine Biosynthesis in Non-small Cell Lung Cancer.

A copy of the manuscript published in *Nature Genetics* is provided below. See the indicated reference for bibliographic information [41].

NRF2 Regulates Serine Biosynthesis in Non-small Cell Lung Cancer

Gina M. DeNicola[1], Pei-Hsuan Chen[2,7], Edouard Mullarky[1,7], Jessica A. Sudderth[2], Zeping Hu[2], David Wu[1], Hao Tang[3], Yang Xie[3], John M. Asara[4], Kenneth E. Huffman[5], Ignacio I. Wistuba[6], John D. Minna[5], Ralph J. DeBerardinis[2] and Lewis C. Cantley[1]

[1]Department of Medicine, Weill Cornell Medical College, New York, NY 10065

[2]Children's Medical Center Research Institute, University of Texas - Southwestern Medical Center, Dallas, Texas 75390-8502

[3]Quantitative Biomedical Research Center, Department of Clinical Sciences, University of Texas - Southwestern Medical Center, Dallas, Texas 75390-8502

[4]Beth Israel Deaconess Medical Center, Division of Signal Transduction and Department of Medicine, Harvard Medical School, Boston, MA 02115

[5]Hamon Center for Therapeutic Oncology, University of Texas - Southwestern Medical Center, Dallas, Texas 75390-8502

[6]Department of Translational Molecular Pathology, The University of Texas, M.D. Anderson Cancer Center, Houston, TX 77030

[7]These authors contributed equally to this work.

*To whom correspondence should be addressed. E-mail: lcantley@med.cornell.edu

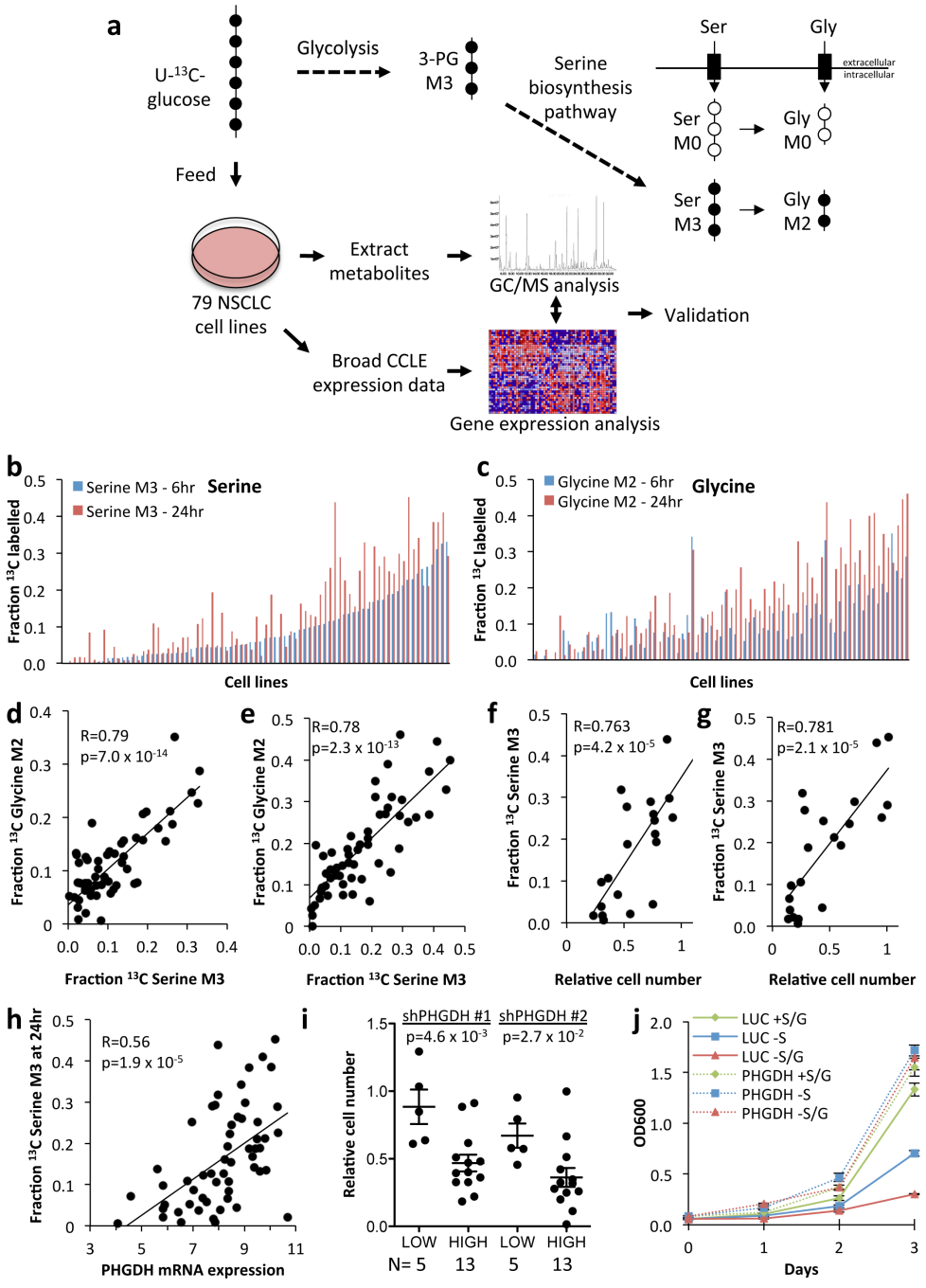
Tumours have high energetic and anabolic needs for rapid cell growth and proliferation[1], and the serine biosynthetic pathway was recently identified as an important source of metabolic intermediates for these processes[2,3]. We integrated metabolic tracing and transcriptional profiling of a large panel of non-small cell lung cancer (NSCLC) cell lines to characterize the activity and regulation of the serine/glycine biosynthetic pathway in NSCLC. Here, we show that the activity of this pathway is highly heterogeneous and is regulated by NRF2, a transcription factor frequently deregulated in NSCLC. We found that NRF2 controls the expression of the key serine/glycine biosynthesis enzymes PHGDH, PSAT1 and SHMT2 via ATF4 to support glutathione and nucleotide production. Moreover, we show that expression of these genes confers poor prognosis in human NSCLC. Thus, a substantial fraction of human NSCLC activates a NRF2-dependent transcriptional program that regulates serine and glycine metabolism and is linked to clinical aggressiveness.

Uniformly labelled ^{13}C-glucose ([U-^{13}C] glucose) is metabolized via the glycolytic intermediate 3-phosphoglycerate to serine M3 (three ^{13}C-labelled atoms) and, subsequently, to glycine M2 (two ^{13}C-labelled atoms) through the combined action of the serine biosynthetic pathway and serine hydroxymethyltransferase[4] (Figure B.1). To profile the activity of the serine/glycine biosynthesis pathway in NSCLC, we labelled a panel of 79 human NSCLC cell lines with ^{13}C-glucose and quantified serine and glycine labelling via gas chromatography/mass spectrometry (GC/MS). We determined 6 and 24 hours to be the optimal time points to detect labelling in serine and glycine

(Supplementary Figure 1). After 24 hours, the fractional abundance of serine M3 and glycine M2 ranged from 0–40% (Figure B.1), with significant correlation between the time points. Neither serine nor glycine labelling correlated with cellular doubling times. There was a significant correlation between serine M3 and glycine M2 labelling at both time points, indicating that the glycine produced from glucose was derived from serine (Figure B.1). As has been reported previously[3], *de novo* serine synthesis conferred the ability to grow in the absence of extracellular serine (Figure B.1). Thus, the serine biosynthesis pathway is not uniformly operant in NSCLC, and regulatory mechanisms exist for controlling the activation of this pathway in a subset of cell lines.

Figure B.1: **Serine biosynthesis activity in lung cancer.** (a) Synthesis of serine and glycine from glucose. Glycolytic metabolism of uniformly labelled ^{13}C-glucose (U-^{13}C-glucose) produces 3-phosphoglycerate (3-PG) labelled on all 3 carbons (M3). The serine biosynthetic pathway produces serine (M3, 3 carbons labelled), and subsequently glycine (M2, 2 carbons labelled) from 3-PG. Unlabelled (M0) serine and glycine are contributed from the cellular media. (b, c) Fraction of labelled intracellular serine (b) and glycine (c) at 6 and 24 hours in NSCLC cell lines. Cell lines are ordered according to serine M3 labelling at 6 hours. (d, e) Correlation between serine and glycine labelling at 6 (d) and 24 (e) hours. R=Pearson correlation coefficient. p-values were calculated by Student's t-distribution with n–2 degrees of freedom. (f–g) Serine 'high' cells (serine M3 Z-score > 0.5 at 24 hours) are resistant to serine (f) or serine/glycine (g) starvation for 3 days. Each data point represents a cell line. (h) Serine labelling at 24 hours correlates with *PHGDH* mRNA expression. mRNA expression data was obtained from the CCLE. Each data point represents a cell line. (i) Serine 'high' cell lines demonstrate sensitivity to PHGDH knockdown that is not rescued by serine. Relative cell number was quantified 5 days after PHGDH knockdown. Each data point represents a cell line. (j) H1975 cells expressing luciferase (LUC) or PHGDH were assayed as in (f–g). Results are the average of 3 biological replicates. Error bars represent SEM here and for all figures.

Figure B.1 (Continued)



To reveal a mechanism for increased serine biosynthesis in NSCLC, we correlated serine and glycine biosynthesis with gene expression using data from the Broad-Novartis Cancer Cell Line Encyclopedia[5]. We observed a significant correlation between serine and glycine labelling and the expression of PHGDH (Figure B.1), the first and rate-limiting step in serine biosynthesis[4]. “Serine-high” cells (serine M3 Z-score > 0.5 at 24 hours) were sensitive to PHGDH silencing (Figure B.1) and ectopic PHGDH rescued the proliferation of serine-low cell lines in serine-deficient media (Figure B.1). Previous studies have revealed a role for *PHGDH* copy number gain in increasing serine biosynthetic activity[2,3]. However, we did not see evidence for this mechanism in NSCLC cell lines. To investigate alternative mechanisms of PHGDH regulation, we performed gene set enrichment analysis (GSEA[6,7]) on the genes that positively correlated with serine and glycine biosynthesis. Interestingly, targets of the transcription factor nuclear factor erythroid-2–related factor 2 (NRF2, gene name *NFE2L2*) were the top hit, suggesting that NRF2 may be a regulator of PHGDH and the serine biosynthetic pathway.

Next, we examined NRF2 localization and found a significant correlation of nuclear NRF2 with serine biosynthesis (Figure B.2). Additionally, we ranked these cell lines according to expression of NRF2 target genes, which significantly correlated with NRF2 abundance. NRF2-high cell lines had significantly higher serine M3 labelling at 6 and 24 hours, and glycine M2 labelling at 24 hours (Figure B.2). Next, we silenced NRF2 and found that it regulates serine/glycine biosynthetic pathway gene expression (*PHGDH*, *PSAT1*, *PSPH*, *SHMT1*, and

SHMT2) (Figure B.2). Decreased *PHGDH* and *PSAT1* mRNA was accompanied by lower protein levels (Figure B.2). Furthermore, NRF2 or PHGDH knockdown decreased the production of ^{13}C-serine from ^{13}C-glucose (Figure B.2). These results demonstrate that transcriptional regulation of serine biosynthetic enzymes by NRF2 controls the production of serine from glucose.

ATF4 transcriptionally activates serine biosynthetic genes in response to serine starvation in NSCLC cells[8]. Interestingly, nuclear ATF4 expression correlated with both serine labelling and NRF2 protein expression. ATF4 has been reported as both a direct transcriptional target[9,10] and heterodimerization partner[11,12] of NRF2. In agreement with transcriptional regulation, we observed a marked reduction in *ATF4* mRNA expression, binding of RNA polymerase II to the *ATF4* promoter, and levels of newly synthesized *ATF4* mRNA upon NRF2 silencing (Figure B.3), with no effect on *ATF4* mRNA stability. Furthermore, ATF4 protein was decreased upon NRF2 knockdown (Figure B.3), but NRF2 did not regulate ATF4 translation[13]. Notably, ATF4 knockdown reduced the expression and activity of the serine pathway to an extent similar to that seen with NRF2 knockdown, while it did not affect NRF2 expression (Figure B.3). Ectopic expression of NRF2 in NRF2-depleted cells partially rescued NQO1, ATF4 and serine biosynthesis enzyme expression and induced expression of these genes in the serine-low cell line H1975, confirming the regulation of these genes by NRF2. Similarly, ectopic ATF4 expression rescued the effects of NRF2 silencing on serine biosynthesis enzyme expression (Figure B.3), and serine labelling from glucose at early time points (Figure B.3), although defects in serine production

were observed later. Furthermore, ectopic ATF4 rescued the growth of H1975 cells in serine deficient media (Figure B.3). We identified the ATF4 binding sites in the PHGDH, PSAT1 and SHMT2 promoters, and found that, while NRF2 itself did not bind to these sites, its silencing significantly decreased binding of ATF4 (Figure B.3). These results demonstrate that NRF2 regulates serine biosynthesis gene expression through ATF4.

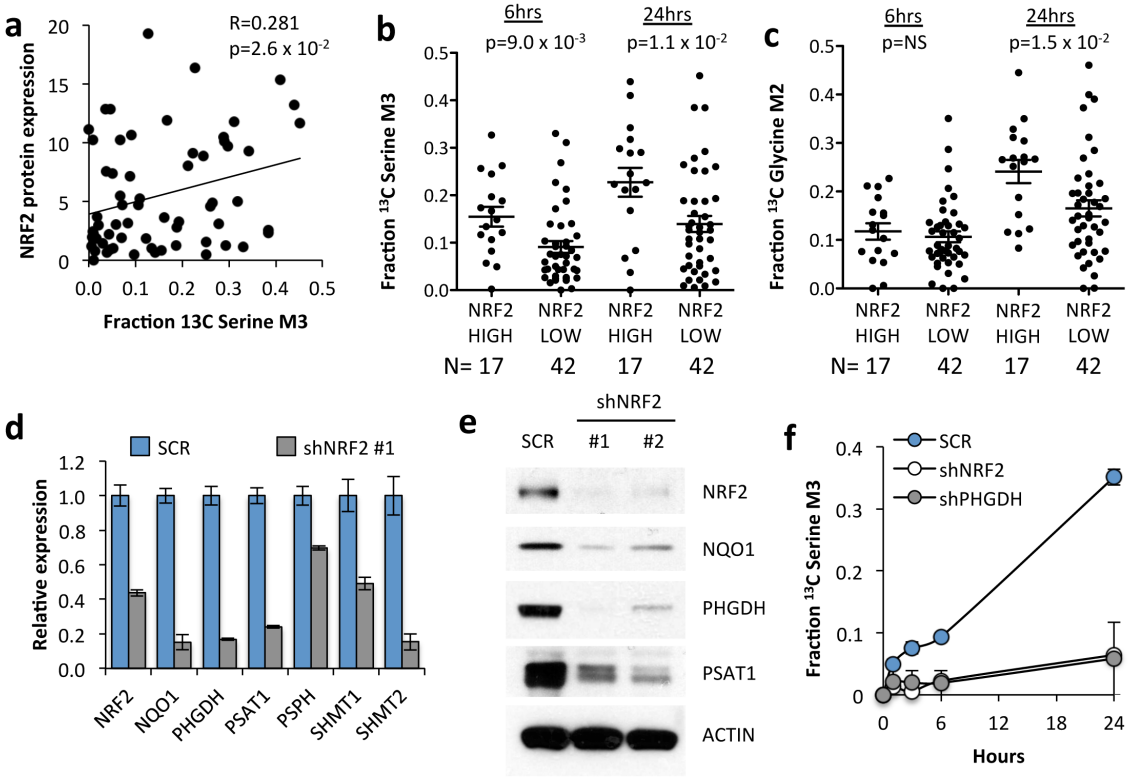
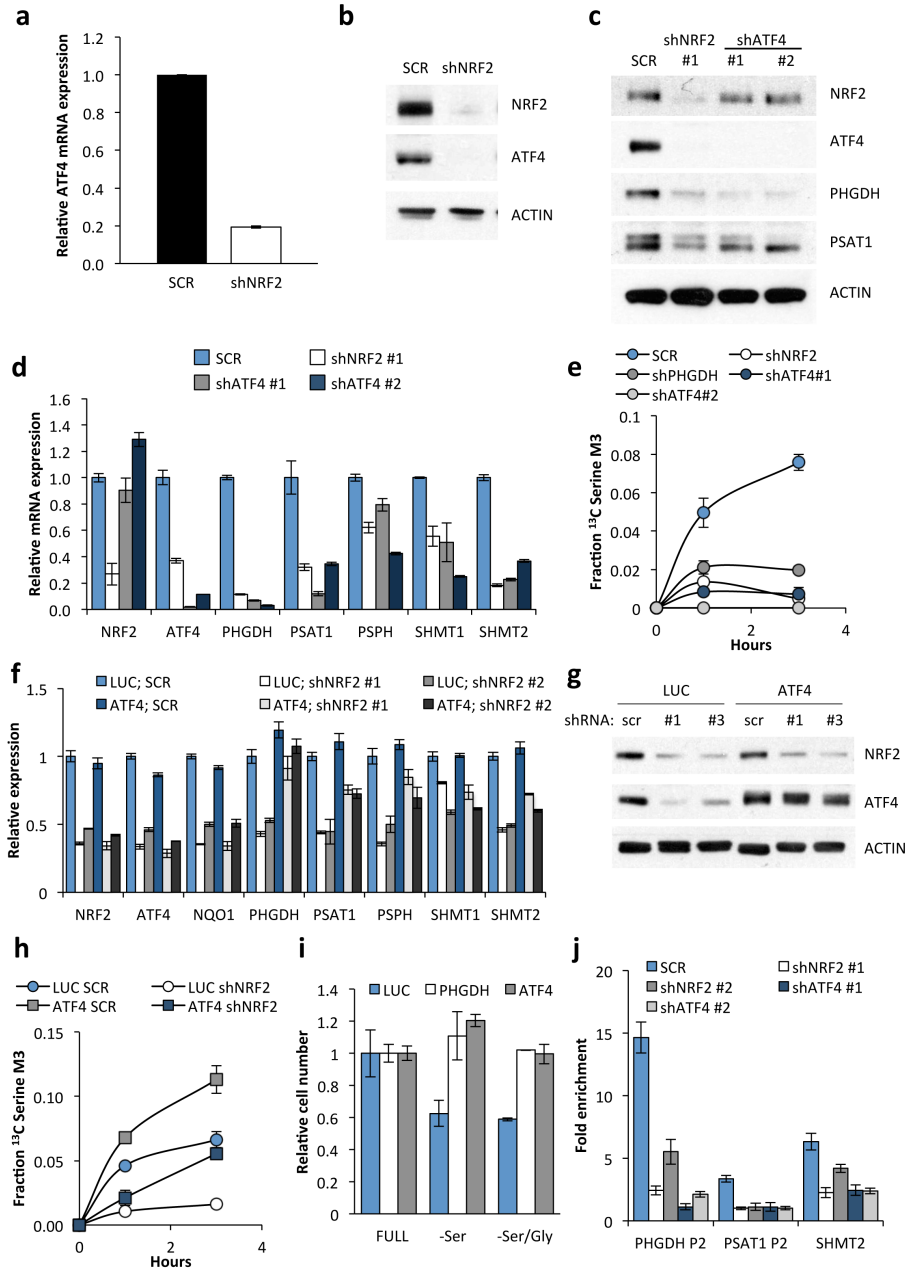


Figure B.2: NRF2 regulates serine biosynthesis. (a) Nuclear NRF2 protein expression correlates with ^{13}C-serine M3 labelling at 24 hours. Nuclear NRF2 protein expression data is found in Supplementary Figure 7. (b) Cell lines with high NRF2 activity have significantly higher ^{13}C serine M3 labelling at 6 and 24 hours. Cell lines were grouped according to the NRF2 score into “high” (>1.4) and “low” (<1.4). (c) NRF2 high cell lines have significantly higher ^{13}C glycine M2 labelling at 24 hours. (b,c) p-values were calculated using an unpaired, two-tailed student’s t-test. (d) mRNA expression in A549 cells expressing scramble shRNA (SCR) or NRF2 shRNA #1. Decreased *NQO1* expression confirmed that NRF2 activity was reduced upon knockdown. Results are the average of 3 technical replicates. (e) Western blot of serine pathway enzyme expression in lysates from A549 cells expressing scramble (SCR), or NRF2 shRNAs #1 or #2. (f) Cell lines were grown in the presence of U-^{13}C-glucose for the indicated time points, the metabolites extracted and the fractional ^{13}C-labelling on serine analysed by LC/MS. Results are the average of 3 biological replicates.

Figure B.3: NRF2 regulates the expression of serine/glycine biosynthesis genes through ATF4. (a) *ATF4* mRNA expression in A549 cells expressing scramble shRNA (SCR), or NRF2 shRNA #1. (b) Western blot of NRF2, ATF4 and ACTIN expression in cells from (a). (c) Western blot of NRF2, ATF4 and serine pathway enzyme expression in lysates from A549s expressing scramble (SCR), NRF2 shRNA #1, or ATF4 shRNAs #1 or #2. (d) mRNA expression in cells from (c). (e) ATF4 knockdown impairs serine biosynthesis. Cell lines from were grown in the presence of U-^{13}C-glucose for the indicated time points, the metabolites extracted and the fractional ^{13}C-labeling on serine analysed by LC/MS. (f) ATF4 rescues serine biosynthesis enzyme expression following NRF2 knockdown. A549 cells were infected with lentivirus encoding mATF4 prior to infection with scramble or NRF2-targeting lentivirus. (g) Western analysis of NRF2, ATF4, and ACTIN expression in the cells from (f). (h) ATF4 rescues the serine biosynthesis defect in shNRF2 A549 cells. Cells were assayed as in (e). (i) ATF4 rescues the growth of H1975 cells in serine deficient media. Cells expressing luciferase (LUC) or ATF4 were grown in the indicated media for 3 days and cell number normalized to cells grown in full media. (j) Chromatin immunoprecipitation of ATF4 to the PHGDH, PSAT1 and SHMT2 promoters. Samples were normalized to IgG control immunoprecipitations. Results are the average of 3 technical (a, d, f, j) or biological (e, h, i) replicates.

Figure B.3 (Continued)

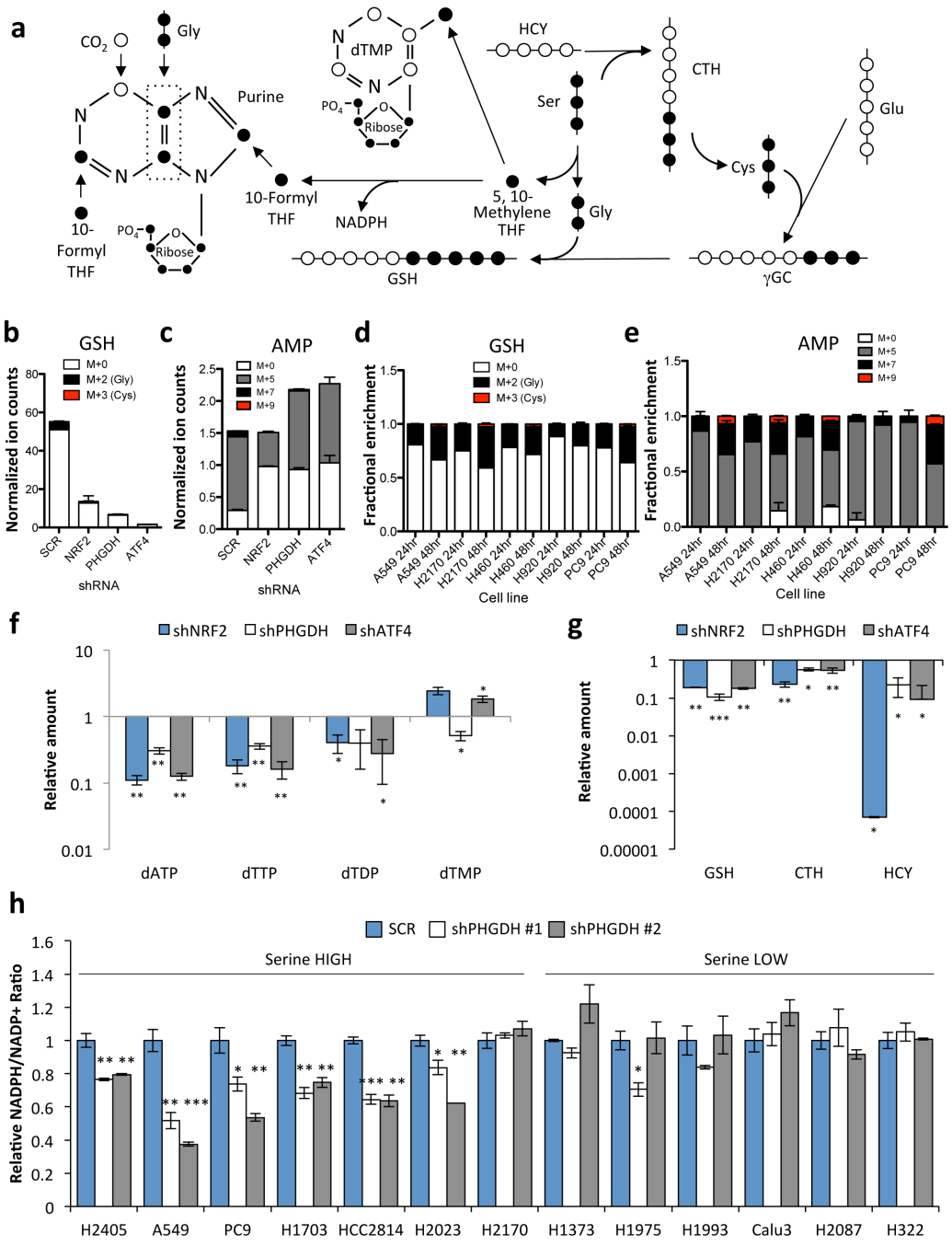


We examined how PHGDH-derived serine contributes to downstream metabolism (Figure B.4). PHGDH, NRF2 or ATF4 silencing decreased the incorporation of glucose-derived serine into cystathionine and glucose-derived glycine into glutathione (Figure B.4), without loss of cell viability. Metabolism of serine to glycine results in the production of one-carbon units via the folate cycle that are utilized for purine and thymidine synthesis (Figure B.4). We observed a decrease in the PHGDH-derived labelling (M+7) of purines including IMP, AMP, ADP, and inosine following PHGDH or ATF4 silencing, while NRF2 silencing decreased both the ribose (M+5) and PHGDH-derived (M+7) labelling of these purines (Figure B.4). Serine-high cell line labelling in glutathione and purines was significantly higher at 48 hours than 24 hours (Figure B.4). Furthermore, the majority of the ^{13}C-labelled serine and glycine had escaped the cell by 24 hours of ^{13}C-glucose labelling due to equilibration with unlabelled amino acids in the media, suggesting that the fraction labelling at 24 hours was underestimating the total contribution of PHGDH to these metabolite pools. In support of this notion, PHGDH, NRF2 or ATF4 silencing resulted in significant decreases in the total levels of purines and thymidine nucleotides (Figure B.4) as well as glutathione, cystathionine and homocysteine (Figure B.4). In contrast, we did not observe differences in the S-adenosyl methionine (SAM)/S-adenosyl homocysteine (SAH) ratio, which is also modulated by the folate cycle, or in histone or DNA methylation. Furthermore, we observed a significant decrease in the NADPH/NADP+ ratio in serine-high cells following PHGDH silencing that was not

observed in serine-low cells (Figure B.4). These results demonstrate that the serine biosynthesis pathway supports glutathione and nucleotide production in NSCLC.

Figure B.4: PHGDH-derived serine supports the transsulfuration and folate cycles. (a) Serine metabolism via the transsulfuration and folate cycles. ^{13}C-labelled carbons (l) unlabelled carbons ($_{o}$). Gly: Glycine, Ser: Serine, CTH: Cystathionine; HCY: Homocysteine, GSH: Glutathione, gGC: g-glutamyl cysteine, Glu: Glutamate. (b–e) A549 cells expressing scramble (SCR), NRF2 or PHGDH shRNAs were grown in the presence of U-^{13}C-glucose for 24 hours and metabolites were extracted and analysed by LC/MS. (b) Analysis of ^{13}C-labelling on the glycine component of glutathione. (c) Analysis of ^{13}C-labeling on the purine metabolite adenosine monophosphate (AMP). (d,e) Analysis of glutathione (d) and AMP (e) labelling in serine high cell lines. Cells were labelled with ^{13}C-glucose for 24 or 48 hours as indicated. (b,d) PHGDH-derived serine is incorporated into glutathione through the generation of glycine and cysteine. M+0 denotes no carbons labelled, M+2 denotes labelling on the glycine moiety, M+3 denotes labeling on the cysteine moiety, and M+5 denotes labeling on both glycine and cysteine. M+5 labeling was not observed. (c,e) M+5 labelling occurs following ribose-5-phosphate labelling via the pentose phosphate pathway. M+7 labelling is the result of ribose labelling plus either glycine or formyl-THF labelling in the purine ring. M+9 labelling is the result of ribose labelling plus either glycine and formyl-THF labelling in the purine ring. M+0 has no labelled carbons. (f–g) LC/MS analysis of total metabolite levels in the nucleotide (f) and transsulfuration (g) pathways. (h) NADPH/NADP+ ratios 4 days after PHGDH knockdown. Results are the average of 3 biological replicates.

Figure B.4 (Continued)

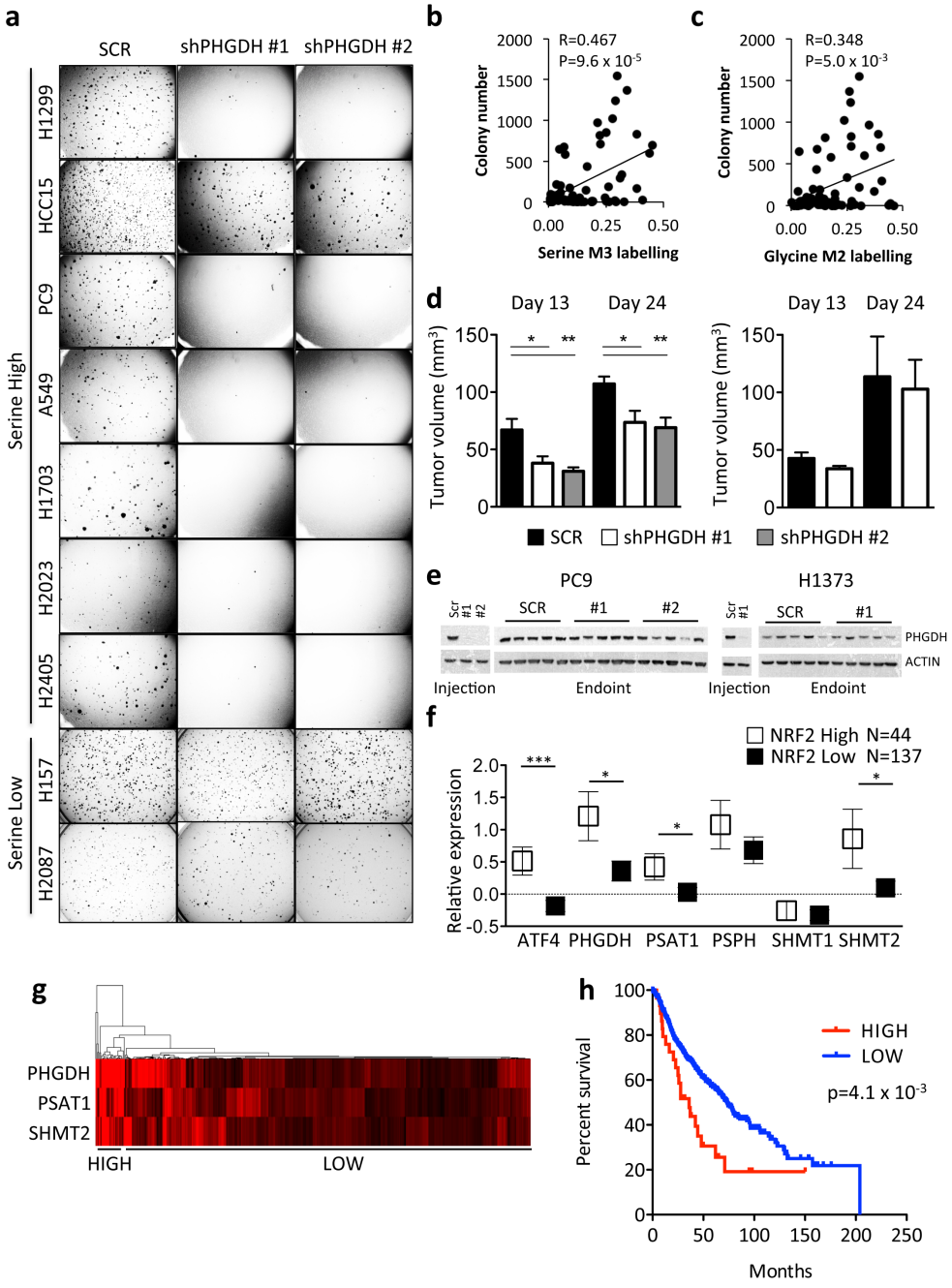


We next asked whether this pathway promotes tumourigenesis. PHGDH silencing significantly impaired the soft agar growth of serine-high, but not serine-low, cell lines (Figure B.5). Interestingly, we observed a significant correlation between serine and glycine labelling and colony number in soft agar (Figure B.5). Furthermore, PHGDH silencing impaired the xenograft growth of the serine-high cell line PC9 (Figure B.5) but not the serine-low cell line H1373 (Figure B.5). We observed that all tumours re-expressed PHGDH at endpoint (Figure B.5), suggesting that PHGDH was required for tumour formation. Next, we examined whether serine pathway gene expression correlated with overall survival of human patients. Human tumours with high NRF2 protein expression displayed significantly higher expression of *ATF4, PHGDH, PSAT1,* and *SHMT2* mRNA (Figure B.5), and *vice versa*. Furthermore, we found that high expression of *PHGDH*, *PSAT1* and *SHMT2* conferred a significantly poorer prognosis (Figure B.5) and was associated with higher tumour grade. These results demonstrate that in human NSCLC, NRF2 regulates the expression of serine biosynthetic enzymes, which correlates with poor prognosis.

We have demonstrated a striking heterogeneity in the activity of the serine biosynthetic pathway in NSCLC. Notably, intracellular amino acid labelling from ^{13}C-glucose is underestimated because of rapid exchange with ^{12}C-amino acids from the media, which is likely mediated by amino acid antiporters (see Supplementary Note). Caution should be used when interpreting data from ^{13}C-labelling experiments due to these exchange mechanisms.

Figure B.5: Activation of the serine biosynthesis pathway promotes tumourigenesis in NSCLC. (a) PHGDH knockdown impairs soft agar growth of serine high, but not serine low, cell lines. (b–c) Soft agar growth correlates with serine (b) and glycine (c) labelling at 24 hours. Each cell line was plated at 5,000 cells/well and the number of colonies counted after 14 days. (d) PHGDH knockdown impairs the xenograft growth of a serine high cell line (PC9, left) but not a serine low cell line (H1373, right). Results are the average of 5 tumors. (e) Western analysis of PHGDH expression of cell lines and xenografts from (d) upon injection and at endpoint. (f) Patients with high NRF2 protein expression (Z-score > 0.5) demonstrate elevated serine pathway gene expression in patient samples from TCGA lung adenocarcinoma cohort. Boxes represent mean values, error bars represent SEM. (g) Gene expression of PHGDH, PSAT1 and SHMT2 in the Director's Consortium lung adenocarcinoma dataset clusters patients into high and low expression cohorts. (h) Kaplan-Meier survival analysis of patients with high (n = 29, red) or low (n = 414, blue) expression of PHGDH, PSAT1 and SHMT2 based on the patient clustering from (g). Median survival is 36 (high) vs. 73.2 (low) months. The p-value was calculated using the Mantel-Cox test.

Figure B.5 (Continued)



Heterogeneity of metabolic or signalling pathways is a common phenomenon across tumours, cell lines, and even between cells of the same tumour. Here, by systematically analysing the serine/glycine biosynthesis pathway in a large, highly-annotated panel of NSCLC cell lines, we identified NRF2 as the molecular driver of this pathway. NRF2 is frequently deregulated in NSCLC through somatic mutations that disrupt the NRF2–KEAP1 interaction to constitutively activate NRF2[14–18]. Cancers with high NRF2 levels are associated with poor prognosis[19,20], resistance to therapy and rapid proliferation[18,21]. Importantly, NRF2 ablation in various tumour models results in elevated ROS and the suppression of tumour growth *in vivo*[21–23]. However, the precise mechanism by which NRF2 promotes tumourigenesis is unclear. Recent studies have revealed that, in addition to genes that promote ROS detoxification, NRF2 regulates genes involved in anabolic metabolism[24,25]. Here, we demonstrate that NRF2 regulates a serine biosynthesis metabolic program via ATF4 and PHGDH to supply the substrates for glutathione and nucleotide production (Figure B.6), which synergizes with the regulation of the pentose phosphate pathway to supply ribose for nucleotides[24,25]. NRF2 indirectly regulates ATF4 transcription via unknown mechanisms. Additionally, while ATF4 completely rescues PHGDH expression following NRF2 depletion, the partial rescue of other serine biosynthesis genes suggests that NRF2 regulates these genes combinatorially through ATF4 and additional factors. Furthermore, these findings suggest that multiple NRF2-regulated pathways co-ordinately contribute to tumourigenesis. Our work encourages the integration of metabolite tracing on large panels of

cancer cell lines in combination with gene expression analysis. This approach is a powerful tool for determining mechanisms responsible for differential regulation of metabolic pathways and may reveal additional links between the activity of metabolic pathways and genetic alterations in cancers.

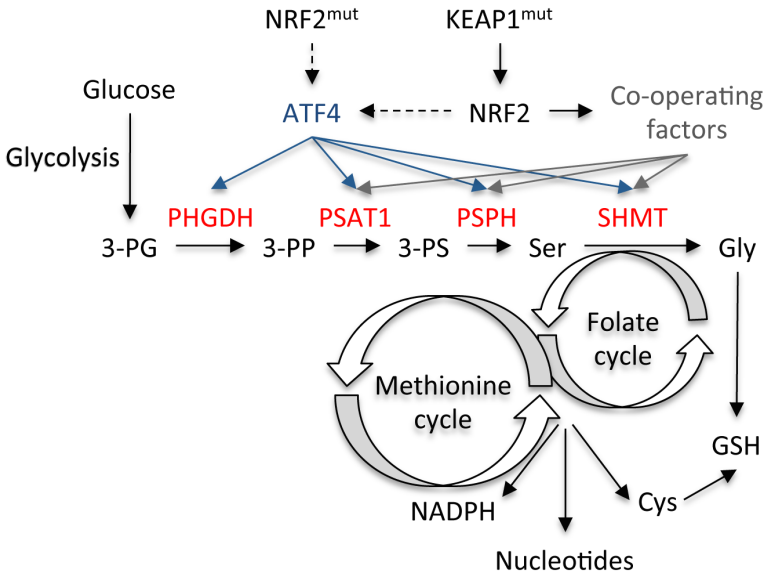


Figure B.6: Model of the regulation of serine/glycine biosynthesis by NRF2. An ATF4 transcriptional program, indirectly activated by NRF2/KEAP1 mutations, regulates the expression of serine/glycine biosynthesis enzymes. These enzymes produce serine and glycine from the glycolytic intermediate 3-PG and funnel the carbon into glutathione and nucleotides via the folate and transsulfuration cycles. NRF2/ATF4 regulated enzymes are shown in red. 3-PG: 3-phosphoglycerate; 3-PP: 3-phosphohydroxy pyruvate; 3-PS: 3-phosphoserine; Ser: Serine; Cys: Cysteine; GSH: Glutathione.

URLs

the TCGA Research Network: http://cancergenome.nih.gov/

Acknowledgements

We would like to thank George Poulogiannis for bioinformatics advice and Hansaa Abbasi, Claire Klimko and Min Yuan for technical support with mass spectrometry experiments. This work was supported by NIH grants P01 CA117969 and R01 GM041890 (L.C.C.), R01 CA157996-01 (R.J.D.), 5R01CA152301 (Y.X.), P50CA70907 (J.D.M, I.I.W., Y.X., K.E.H.), and CPRIT funding to J.D.M, Y.X., I.I.W., K.E.H. (RP110708, RP120732) and R.J.D (RP130272). P-H.C. was supported by a grant from the Welch Foundation to R.J.D. (I-1733). The mass spectrometry work was partially supported by NIH grants 5P30CA006516 and 5 P01CA120964 (J.M.A.). G.M.D. was the Malcolm AS Moore Hope Funds for Cancer Research Fellow and is supported by the PanCAN/AACR Pathway to Leadership grant.

Author Contributions

G.M.D., R.J.D., and L.C.C. designed the study. G.M.D. ad E.M. performed molecular biology experiments. G.M.D., P.H.C., E.M., J.A.S., Z.H., and J.M.A. performed metabolomics and isotope labelling and analysed the data. D.W. performed xenograft experiments. H.T. and Y.X. performed bioinformatics analysis. K.E.H., I.I.W. and J.D.M. contributed highly annotated lung cancer cell

lines. G.M.D., E.M., and L.C.C. wrote the manuscript. All authors commented on the manuscript.

Competing financial interests

L.C.C. owns equity in, receives compensation from, and serves on the Board of Directors and Scientific Advisory Board of Agios Pharmaceuticals. Agios Pharmaceuticals is identifying metabolic pathways of cancer cells and developing drugs to inhibit such enzymes in order to disrupt tumour cell growth and survival. R.J.D. is on the scientific advisory boards of Agios Pharmaceuticals and Peloton Therapeutics. Peloton Therapeutics is developing drugs to target altered molecular pathways in cancer, including altered metabolism.

References

1. Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324,** 1029–1033 (2009).

2. Locasale, J. W. *et al.* Phosphoglycerate dehydrogenase diverts glycolytic flux and contributes to oncogenesis. *Nat Genet* **43,** 869–874 (2011).

3. Possemato, R. *et al.* Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature* **476,** 346–350 (2011).

4. Mullarky, E., Mattaini, K. R., Vander Heiden, M. G., Cantley, L. C. & Locasale, J. W. PHGDH amplification and altered glucose metabolism in human melanoma. *Pigment Cell Melanoma Res* **24,** 1112–1115 (2011).

5. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483,** 603–607 (2012).

6. Mootha, V. K. *et al.* PGC-1alpha-responsive genes involved in oxidative

phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* **34,** 267–273 (2003).

7. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102,** 15545–15550 (2005).

8. Ye, J. *et al.* Pyruvate kinase M2 promotes de novo serine synthesis to sustain mTORC1 activity and cell proliferation. *Proc Natl Acad Sci U S A* **109,** 6904–6909 (2012).

9. Miyamoto, N. *et al.* Transcriptional regulation of activating transcription factor 4 under oxidative stress in retinal pigment epithelial ARPE-19/HPV-16 cells. *Invest. Ophthalmol. Vis. Sci.* **52,** 1226–1234 (2011).

10. Afonyushkin, T. *et al.* Oxidized phospholipids regulate expression of ATF4 and VEGF in endothelial cells via NRF2-dependent mechanism: novel point of convergence between electrophilic and unfolded protein stress pathways. *Arterioscler Thromb Vasc Biol* **30,** 1007–1013 (2010).

11. Ye, P. *et al.* Nrf2- and ATF4-dependent upregulation of xCT modulates the sensitivity of T24 bladder carcinoma cells to proteasome inhibition. *Mol Cell Biol* **34,** 3421–3434 (2014).

12. He, C. H. *et al.* Identification of activating transcription factor 4 (ATF4) as an Nrf2-interacting protein. Implication for heme oxygenase-1 gene regulation. *J Biol Chem* **276,** 20858–20865 (2001).

13. Harding, H. P. *et al.* Regulated translation initiation controls stress-induced gene expression in mammalian cells. *Mol Cell* **6,** 1099–1108 (2000).

14. Hayes, J. D. & McMahon, M. NRF2 and KEAP1 mutations: permanent activation of an adaptive response in cancer. *Trends Biochem Sci* **34,** 176–188 (2009).

15. Kim, Y. R. *et al.* Oncogenic NRF2 mutations in squamous cell carcinomas of oesophagus and skin. *J Pathol* **220,** 446–451 (2010).

16. Konstantinopoulos, P. A. *et al.* Keap1 mutations and Nrf2 pathway activation in epithelial ovarian cancer. *Cancer Res* **71,** 5081–5089 (2011).

17. Seng, S. *et al.* NRP/B mutations impair Nrf2-dependent NQO1 induction in human primary brain tumors. *Oncogene* **28,** 378–389 (2009).

18. Zhang, P. *et al.* Loss of Kelch-like ECH-associated protein 1 function in prostate cancer cells causes chemoresistance and radioresistance and promotes

tumor growth. *Mol Cancer Ther* **9,** 336–346 (2010).

19. Shibata, T. *et al.* Cancer related mutations in NRF2 impair its recognition by Keap1-Cul3 E3 ligase and promote malignancy. *Proc Natl Acad Sci U S A* **105,** 13568–13573 (2008).

20. Solis, L. M. *et al.* Nrf2 and Keap1 abnormalities in non-small cell lung carcinoma and association with clinicopathologic features. *Clin Cancer Res* **16,** 3743–3753 (2010).

21. Singh, A. *et al.* RNAi-mediated silencing of nuclear factor erythroid-2-related factor 2 gene expression in non-small cell lung cancer inhibits tumor growth and increases efficacy of chemotherapy. *Cancer Res* **68,** 7975–7984 (2008).

22. DeNicola, G. M. *et al.* Oncogene-induced Nrf2 transcription promotes ROS detoxification and tumorigenesis. *Nature* **475,** 106–109 (2011).

23. Ohta, T. *et al.* Loss of Keap1 function activates Nrf2 and provides advantages for lung cancer cell growth. *Cancer Res* **68,** 1303–1309 (2008).

24. Mitsuishi, Y. *et al.* Nrf2 redirects glucose and glutamine into anabolic pathways in metabolic reprogramming. *Cancer Cell* **22,** 66–79 (2012).

25. Singh, A. *et al.* Transcription factor NRF2 regulates miR-1 and miR-206 to drive tumorigenesis. *J Clin Invest* **123,** 2921–2934 (2013).

Online Methods

Animals. Male nude mice (CrTac:NCr-Foxn1nu) were obtained from Taconic labs and maintained under pathogen-free conditions. Experiments were performed according to IACUC guidelines. Mice were injected at 8 weeks of age with $2X10^6$ NSCLC cells on each flank. shRNAs were randomized so they were evenly distributed across mice. As the animal study was exploratory, no statistical test was used to determine adequate sample size. No mice were excluded from the analysis. The study was not blinded.

Cell culture. All NSCLC cell lines used in this study were obtained from the Hamon Cancer Center Collection (University of Texas Southwestern Medical Center). Cells were maintained in RPMI-1640 (Life Technologies Inc.) supplemented with 5% or 10% fetal calf serum (FCS) without antibiotics at 37°C in a humidified atmosphere containing 5% CO_2 and 95% air. All experiments were performed in media containing serine and glycine except where otherwise noted. All cell lines have been DNA fingerprinted using the PowerPlex 1.2 kit (Promega) and mycoplasma tested by the e-Myco (Boca Scientific) or MycoAlert kit (Lonza). Although NCI-H157 is listed in the database of commonly misidentified cell lines, it was originally derived by J.D.M. and fingerprinted before use.

Celltox green staining. Cells were incubated in RPMI + 10% FCS containing 1X Celltox green (Promega) and 5 uM Syto 17 (Life Technologies) for 20 minutes, washed in PBS, and imaged on an EVOS FL cell imaging system (Life Technologies).

Chromatin Immunoprecipitation. Cells ($5X10^6$) were fixed at 37°C in RPMI with 1% formaldehyde for 10 min, lysed in 1% SDS, 10 mM EDTA, 50 mM Tris-HCl pH 8.1 plus protease inhibitors and sonicated with a probe tip until DNA was an average size of 1 kilobase. Input was saved and lysate was diluted in immunoprecipitation buffer (1% Triton, 2 mM EDTA, 150 mM NaCl, 20 mM Tris-HCl pH 8.1) and mixed with beads (Dynal Protein A, Invitrogen) that were pre-bound overnight with antibodies to Nrf2 (H-300, Santa Cruz), ATF4 (11815, Cell Signaling), RNA pol II (pSer5, ab5131, abcam) or Rabbit IgG (sc-2027, Santa

Cruz). Chromatin was immunoprecipitated overnight, and beads were washed six times with RIPA buffer (50 mM HEPES pH 7.6, 1 mM EDTA, 0.7% Na deoxycholate, 1% NP-40, 0.5 M LiCl) and twice with TE. Beads were incubated with 1% SDS, 0.1 M $NaHCO_3$ for 30 min at room temperature, and then crosslinks were reversed on both the input and the immunoprecipitate by heating overnight in a 65°C water bath. DNA was purified with a QIAquick spin kit (Qiagen) and Q-PCR was performed in triplicate with a Fast Sybr green master mix on a Step One Real-Time PCR system (all Life Technologies). Primer sequences can be found in Supplementary Table 4.

DNA methylation. DNA was extracted from cells in lysis buffer (10 mM Tris, 100 mM NaCl, 10 mM EDTA, 0.5% SDS, 0.4 mg/ml Proteinase K, pH 8.0) by heating overnight at 65°C. DNA was purified by precipitating protein with NaCl, and precipitated with isopropanol. 2 ug DNA was denatured in 0.4 M NaOH, 10 mM EDTA at 95°C for 10min and neutralized with an equal volume of 2 M ammonium acetate, pH7.0. 200 ng of DNA was spotted on a nylon membrane (GE Healthcare), cross linked at maximum twice, and blocked in 5% milk for 1 hour. The membrane was incubated overnight with 5-meC antibody (clone 33D3; 1:1,000; Epigentek), followed by an HRP-secondary and chemiluminescence visualization. To ensure equal loading, the membrane was stained with 0.02% methylene blue in 0.3 M sodium acetate, pH 5.2. 5-meC intensity was normalized to methylene blue.

GC/MS metabolite tracing. All NSCLC cell lines were cultured under identical conditions to identify cell-autonomous differences in glucose

metabolism. Cells were seeded into 60-mm culture dishes and grown until 70–80% confluent. Each dish was then rinsed in warm phosphate-buffered saline and overlaid with 4 mL of RPMI-1640 containing 5% dialyzed fetal calf serum, 4 mM unlabelled glutamine, and 10 mM [U-^{13}C] glucose (Cambridge Isotope Laboratories). NB: RPMI-1640 contains unlabelled serine and glycine. After 6 hours or 24 hours, the medium was aspirated and cells were rapidly rinsed in cold normal saline solution. The cells were lysed in 0.5 mL of cold 50% methanol, with three rapid cycles of freeze-thawing between liquid nitrogen and a 37^{o}C water bath. The lysates were cleared of cellular debris by centrifugation, and metabolites in the supernatant were evaporated under blown air or by centrifugation under vacuum. Derivatization, mass spectrometry, and mathematical correction for natural abundance isotopes were performed according to published methods [26,27]. The following fragments were monitored, which represent derivatized species: serine – m/z 306 (M+0) – 309 (M+3); glycine, m/z 276 (M+0) – 278 (M+2). Every cell line was analysed in biological replicates where n≥3 and no two replicates were performed on the same day.

Lentivirus production and infection. Lenti-X 293T cells (Clontech) were transfected at 90% confluence with Lipofectamine 2000 (Invitrogen). Packaging plasmids pCMV-dR8.2 dvpr (addgene # 8455) and pCMV-VSV-G (addgene #8454) were used with the vectors found in Supplementary Table 5.

Viral supernatants were collected at 48 and 72 hours and added to target cells with 8mg/ml polybrene for 3 hours. Cells were selected in 1 mg/ml puromycin where appropriate.

LC/MS measurement of total and ^{13}C-labeled metabolites in A549 cells. Cells were plated the day before labelling at $2X10^6$ cells/10 cm dish and media was changed to glucose-free RPMI containing 10% dialyzed serum + 10 mM U-^{13}C-glucose (Cambridge Isotope Laboratories) for the indicated time points. Metabolites were extracted in ice-cold 80% methanol and analysed by targeted LC-MS/MS via selected reaction monitoring (SRM), as described[28]. The following precursor ions were monitored: serine – m/z 106 (M+0) and 109 (M+3); cystathionine – m/z 223 (M+0) and 226 (M+3); glutathione – m/z 308 (M+0), glutathione (G) – m/z 310 (M+2); homocysteine – m/z; 136; IMP – m/z 349 (M+0), 354 (M+5) and 356 (M+7); inosine – m/z 267 (M+0), 272 (M+5), 274 (M+7); AMP – m/z 348 (M+0), 353 (M+5) and 355 (M+7); dATP – m/z 490; dTTP – m/z 481; dTDP – m/z 401; dTMP – m/z 323 . The bracketed amino acid following glutathione indicates which constituent amino acid is ^{13}C labelled. Samples were analysed in triplicate. Data represents median-normalized values.

Luciferase assays. Luciferase assays were performed with the Dual-Glo Luciferase Assay System (Promega) according to the manufacturer's instructions.

NADPH/NADP+ ratios. NADP+ and NADPH were determined with the NADP/NADPH-glo assay kit (Promega) according to the manufacturer's protocol.

NRF2 score calculation. Cell lines were grouped into high and low groups based on the expression of five classic NRF2 target genes: *NQO1*, *GCLC*, *GLCM*, *SLC7A11*, and *AKR1C1*. The expression of each gene was normalized to the median across the cell lines, and these five genes were then averaged

together. NRF2 high vs. low was defined as the top quartile vs. bottom three quartiles. The top 20 overexpressed genes in the NRF2 high cell lines compared to the NRF2 low were used to calculate the NRF2 score: *AKR1C1*, *AKR1C2*, *SPP1*, *ALDH3A1*, *LOC344887*, *AKR1C3*, *OSGIN1*, *PGD*, *CYP4F11*, *AKR1B10*, *KIAA0319*, *SRXN1*, *NR0B1*, *SLC7A11*, *LOC100292680*, *ABCC2*, *CABYR*, *JAKMIP3*, *KYNU* and *PTGR1*. Gene expression values were normalized to the median and then these 20 genes were averaged together to get the NRF2 score. Individual gene expression values and NRF2 scores can be found in Supplementary Table 3, and NRF2 high vs. low clustering is found in Supplementary Figure 8.

Patient samples and survival analysis. The National Cancer Institute Director's Challenge Consortium study (Director's Consortium)[29] and the Cancer Genome Atlas (TCGA) lung adenocarcinoma data were used in this study to evaluate gene signatures' prognostic performance. The Director's Consortium data set collected 442 resected lung adenocarcinomas at four US institutions[29], while the TCGA Research Network data includes 203 patient samples for which gene expression and survival data are available. Unsupervised cluster analysis was used to group patients based on expression of PHGDH, PSAT1 and SHMT2, using average linkage clustering with the Spearman's rank correlation distance metric. Clustering was performed with Cluster 3.0. Heat map visualization was performed with JavaTreeView. Kaplan–Meier survival curves were used to determine survival rate as a function of time and survival differences were analysed by a log-rank Mantel-Cox test using Graph Pad Prism.

Proliferation assays. Cells were seeded at 500–10,000 cells/well in 96 well plates on day –1 and infected on day 0 with lentivirus. Alternatively, cells were switched into RPMI + 10% dialyzed FBS containing full amino acids, lacking serine, or lacking serine and glycine. Plates were fixed on the indicated days with 4% paraformaldehyde, stained with crystal violet, washed and dried. Crystal violet was solubilized in 10% acetic acid and the OD_{600} was measured.

Pulse-labelling with 4-thiouridine. Cells were labelled on 10 cm dishes at 70% confluence with 200 uM 4-thiouridine (4sU, Sigma Aldrich) for 60 minutes as described[30]. RNA was extracted with Trizol. 4-thiouridine containing mRNA molecules were biotinylated with biotin-HPDP (EZ-Link Biotin-HPDP, Pierce, Cat #21341), and purified with the uMACS streptavidin kit (Miltenyi Biotec). RNA was eluted in 100 mM DTT, cDNA was synthesized, and ATF4 levels quantified and normalized to b-actin.

Reagents. Serine and glycine were purchased from Sigma Aldrich. RPMI media lacking serine and glycine was custom prepared by Life Technologies.

Real-time PCR. RNA was isolated using an RNeasy kit (Qiagen). cDNA was synthesized using the Superscript VILO Master mix and analysed by quantitative PCR (q-PCR) using Fast Sybr green master mix on a Step One Real-Time PCR system (all Life Technologies). Target gene expression was normalized to actin expression, and shown relative to control samples. Primer sequences can be found in Supplementary Table 4.

siRNA transfection. 100,000 cells/well were reverse transfected in 800 ul of growth media in 12 well dishes. 2 ul DharmaFECT Duo was combined with

100 pmol siRNA in a final volume of 200 ul according to the manufacturer's instructions, which was then added to the cells. Cells were analysed after 2 days. Dharmacon ON-TARGET plus non-targeting siRNA (D-001810-10) and NFE2L2 (L-003755-00-0005) pools were used.

Soft agar assays. Soft agar assays were performed in triplicate in 6-well dishes. A 1 mL base layer of 0.8% agar in RPMI was plated and allowed to solidify, then 5,000 cells/well were plated in 0.4% agar on top. 1mL of RPMI was added the following day to each well, and changed as needed. NB: RPMI contains serine and glycine. Soft agar was stained with 0.01% crystal violet in 4%PFA/PBS solution and imaged in a ChemiDoc system (Bio-Rad). Colonies were quantified with Image J.

Statistical analysis. Data were analysed using a two-sided unpaired Student's t test and the Mantel-Cox test was used for survival analyses. For all statistical analyses GraphPad Prism 6 software was used, and values of $p < 0.05$ were considered statistically significant (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$). The mean ± standard error of the mean of at least three independent experiments performed in triplicates is reported. For all experiments similar variances between groups were observed. Normal distribution of samples was not determined.

Western blotting. Protein lysates were prepared using RIPA lysis buffer and separated on 4–12% NuPAGE gels (Invitrogen), transferred onto a nitrocellulose membrane (Millipore) and incubated with the following antibodies: ACTIN (ab6276) and NRF2 (EP1808Y) mAb (both Abcam), NQO1 (HPA007308) and PHGDH (HPA021241) (both Sigma), PSAT1 (PA5-22124, Pierce), and ATF4

(11815, Cell Signaling). Alternatively, nuclear extracts were prepared as described[31]. Histone extracts were prepared with the Histone Extraction Kit according to the manufacturer's instructions (Abcam, ab113476). Histone extracts were probed with the following antibodies: H3K4me3 (9727, Cell Signaling), H3K27me3 (07-449, Millipore), total H3 (4499, Cell Signaling).

Bibliography

1. Warburg, O. On respiratory impairment in cancer cells. *Science* **124,** 270–272 (1956).

2. Wise, D. R. *et al.* Myc regulates a transcriptional program that stimulates mitochondrial glutaminolysis and leads to glutamine addiction. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 18782–18787 (2008).

3. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144,** 646–674 (2011).

4. Ying, H. *et al.* Oncogenic Kras maintains pancreatic tumors through regulation of anabolic glucose metabolism. *Cell* **149,** 656–670 (2012).

5. Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. *Science* **324,** 1029–1033 (2009).

6. Yun, J. *et al.* Vitamin C selectively kills KRAS and BRAF mutant colorectal cancer cells by targeting GAPDH. *Science* **350,** 1391–1396 (2015).

7. Lunt, S. Y. & Vander Heiden, M. G. Aerobic glycolysis: meeting the metabolic requirements of cell proliferation. *Annu. Rev. Cell Dev. Biol.* **27,** 441–464 (2011).

8. Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324,** 1029–1033 (2009).

9. DeBerardinis, R. J. & Cheng, T. Q's next: the diverse functions of glutamine in metabolism, cell biology and cancer. *Oncogene* **29,** 313–324 (2010).

10. Yang, C. *et al.* Glioblastoma cells require glutamate dehydrogenase to survive impairments of glucose metabolism or Akt signaling. *Cancer Research* **69,** 7986–7993 (2009).

11. Son, J. *et al.* Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway. *Nature* **496,** 101–105 (2014).

12. Fan, J. *et al.* Glutamine-driven oxidative phosphorylation is a major ATP source in transformed mammalian cells in both normoxia and hypoxia.

Mol. Syst. Biol. **9,** 712–712 (2013).

13. Zu, X. L. & Guppy, M. Cancer metabolism: facts, fantasy, and fiction. *Bioch Biophy Res Com* **313,** 459–465 (2004).

14. Kroemer, G. & Pouyssegur, J. Tumor cell metabolism: cancer's Achilles' heel. *Cancer Cell* **13,** 472–482 (2008).

15. Vander Heiden, M. G. Targeting cancer metabolism: a therapeutic window opens. *Nat Rev Drug Discov* **10,** 671–684 (2011).

16. Eagle, H. *Amino acid metabolism in mammalian cell cultures*. (Science, 1959).

17. Mullarky, E., Mattaini, K. R., Vander Heiden, M. G., Cantley, L. C. & Locasale, J. W. PHGDH amplification and altered glucose metabolism in human melanoma. *Pigment Cell Melanoma Res* **24,** 1112–1115 (2011).

18. Snell, K. The duality of pathways for serine biosynthesis is a fallacy. *TIBS* 241–243 (1986).

19. Knox, W. E., Herzfeld, A. & Hudson, J. Phosphoserine phosphatase distribution in normal and neoplastic rat tissues. *Arch. Biochem. Biophys.* **132,** 397–403 (1969).

20. Tibbetts, A. S. & Appling, D. R. Compartmentalization of Mammalian folate-mediated one-carbon metabolism. *Annu. Rev. Nutr.* **30,** 57–81 (2010).

21. Palacín, M., Estévez, R., Bertran, J. & Zorzano, A. Molecular biology of mammalian plasma membrane amino acid transporters. *Physiol. Rev.* **78,** 969–1054 (1998).

22. Rowsell, E. V., Snell, K., Carnie, J. A. & Al-Tai, A. H. Liver-L-alanine-glyoxylate and L-serine-pyruvate aminotransferase activities: an apparent association with gluconeogenesis. *Biochemical Journal* **115,** 1071–1073 (1969).

23. Cochat, P. & Rumsby, G. Primary Hyperoxaluria. *N. Engl. J. Med.* **369,** 649–658 (2013).

24. Kuge, O., Hasegawa, K., Saito, K. & Nishijima, M. Control of phosphatidylserine biosynthesis through phosphatidylserine-mediated inhibition of phosphatidylserine synthase I in Chinese hamster ovary cells. *Proc Natl Acad Sci* **95,** 4199–4203 (1998).

25. de Koning, T. J. *et al.* L-serine in disease and development. *Biochem. J.* **371,** 653–661 (2003).

26. Futerman, A. H. & Riezman, H. The ins and outs of sphingolipid synthesis. *Trends Cell Biol.* **15,** 312–318 (2005).

27. Esaki, K. *et al.* L-Serine deficiency elicits intracellular accumulation of cytotoxic deoxy-sphingolipids and lipid body formation. *J Biol Chem* **290,** jbc.M114.603860–14609 (2015).

28. Tibbetts, A. S. & Appling, D. R. Compartmentalization of Mammalian Folate-Mediated One-Carbon Metabolism. *Annu. Rev. Nutr.* **30,** 57–81 (2010).

29. Fu, T.-F., Rife, J. P. & Schirch, V. The Role of Serine Hydroxymethyltransferase Isozymes in One-Carbon Metabolism in MCF-7 Cells as Determined by 13C NMR. *Arch. Biochem. Biophys.* **393,** 42–50 (2001).

30. Fabrega, C., Hausmann, S., Shen, V., Shuman, S. & Lima, C. D. Structure and mechanism of mRNA cap (guanine-N7) methyltransferase. *Molecular Cell* **13,** 77–89 (2004).

31. Loenen, W. A. M. S-adenosylmethionine: jack of all trades and master of everything? *Biochem. Soc. Trans.* **34,** 330–333 (2006).

32. Carp, H. Mitochondrial N-formylmethionyl proteins as chemoattractants for neutrophils. *J. Exp. Med.* **155,** 264–275 (1982).

33. Mullarky, E. & Cantley, L. C. Diverting Glycolysis to Combat Oxidative Stress. *Innovative Medicine* (2015). doi:10.1007/978-4-431-55651-0_1

34. Shoolingin-Jordan, P. M. *et al.* 5-Aminolevulinic acid synthase: mechanism, mutations and medicine. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1647,** 361–366 (2003).

35. Snell, K. Enzymes of serine metabolism in normal, developing and neoplastic rat tissues. *Adv. Enzyme Regul.* **22,** 325–400 (1984).

36. Snell, K., Natsumeda, Y., Eble, J. N., Glover, J. L. & Weber, G. Enzymic imbalance in serine metabolism in human colon carcinoma and rat sarcoma. *Br. J. Cancer* **57,** 87–90 (1988).

37. Snell, K. & Weber, G. Enzymic imbalance in serine metabolism in rat hepatomas. *Biochemical Journal* **233,** 617–620 (1986).

38. Snell, K., Natsumeda, Y. & Weber, G. The modulation of serine metabolism in hepatoma 3924A during different phases of cellular proliferation in culture. *Biochemical Journal* **245,** 609–612 (1987).

39. Snell, K. & Fell, D. A. Metabolic control analysis of mammalian serine metabolism. *Adv. Enzyme Regul.* **30,** 13–32 (1990).

40. Snell, K., Natsumeda, Y., Eble, J. N., Glover, J. L. & Weber, G. Enzymic imbalance in serine metabolism in human colon carcinoma and rat sarcoma. *Br. J. Cancer* **57,** 87–90 (1988).

41. DeNicola, G. M. *et al.* NRF2 regulates serine biosynthesis in non-small cell lung cancer. *Nat. Genet.* **47,** 1475–1481 (2015).

42. Labuschagne, C. F., van den Broek, N. J. F., Mackay, G. M., Vousden, K. H. & Maddocks, O. D. K. Serine, but Not Glycine, Supports One-Carbon Metabolism and Proliferation of Cancer Cells. *Cell Rep* **7,** 1248–1258 (2014).

43. Jain, M. *et al.* Metabolite Profiling Identifies a Key Role for Glycine in Rapid Cancer Cell Proliferation. *Science* **336,** 1040–1044 (2012).

44. Possemato, R. *et al.* Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature* **476,** 346–350 (2011).

45. Locasale, J. W. *et al.* Phosphoglycerate dehydrogenase diverts glycolytic flux and contributes to oncogenesis. *Nat. Genet.* **43,** 869–874 (2011).

46. Beroukhim, R. *et al.* The landscape of somatic copy-number alteration across human cancers. *Nature* **463,** 899–905 (2010).

47. Pollari, S. *et al.* Enhanced serine production by bone metastatic breast cancer cells stimulates osteoclastogenesis. *Breast Cancer Res Treat* **125,** 421–430 (2011).

48. Ding, J. *et al.* The Histone H3 Methyltransferase G9A Epigenetically Activates the Serine-Glycine Synthesis Pathwayto Sustain Cancer Cell Survival and Proliferation. *Cell Metabolism* **18,** 896–907 (2013).

49. Nilsson, L. M. *et al.* Mouse genetics suggests cell-context dependency for Myc-regulated metabolic enzymes during tumorigenesis. *PLoS Genet.* **8,** e1002573 (2012).

50. Kung, C. *et al.* Small molecule activation of PKM2 in cancer cells induces serine auxotrophy. *Chemistry & Biology* **19,** 1187–1198 (2012).

51. Singleton, D. C. & Harris, A. L. Targeting the ATF4 pathway in cancer therapy. *Expert Opin. Ther. Targets* **16,** 1189–1202 (2012).

52. Ye, J. *et al.* Pyruvate kinase M2 promotes de novo serine synthesis to sustain mTORC1 activity and cell proliferation. *Proc Natl Acad Sci* **109,** 6904–6909 (2012).

53. Chaneton, B. *et al.* Serine is a natural ligand and allosteric activator of pyruvate kinase M2. *Nature* **491,** 458–462 (2013).

54. Ye, J., Mancuso, A., Tong, X. & Ward, P. S. Pyruvate kinase M2 promotes de novo serine synthesis to sustain mTORC1 activity and cell proliferation. in (2012). doi:10.1073/pnas.1204176109/-/DCSupplemental

55. Lunt, S. Y. *et al.* Pyruvate Kinase Isoform Expression Alters Nucleotide Synthesis to Impact Cell Proliferation. *Molecular Cell* (2014). doi:10.1016/j.molcel.2014.10.027

56. Gorrini, C., Harris, I. S. & Mak, T. W. Modulation of oxidative stress as an anticancer strategy. *Nat Rev Drug Discov* **12,** 931–947 (2013).

57. Jiang, P., Du, W., Mancuso, A., Wellen, K. E. & Yang, X. Reciprocal regulation of p53 and malic enzymes modulates metabolism and senescence. *Nature* **493,** 689–693 (2013).

58. Fan, J. *et al.* Quantitative flux analysis reveals folate-dependent NADPH production. *Nature* **510,** 298–302 (2014).

59. Lewis, C. A. *et al.* Tracing compartmentalized NADPH metabolism in the cytosol and mitochondria of mammalian cells. *Molecular Cell* **55,** 253–263 (2014).

60. Ye, J. *et al.* Serine catabolism regulates mitochondrial redox control during hypoxia. *Cancer Discovery* **4,** 1406–1417 (2014).

61. DeNicola, G. M. *et al.* Oncogene-induced Nrf2 transcription promotes ROS detoxification and tumorigenesis. *Nature* **475,** 106–109 (2012).

62. Zhang, W. C. *et al.* Glycine decarboxylase activity drives non-small cell lung cancer tumor-initiating cells and tumorigenesis. *Cell* **148,** 259–272 (2012).

63. Mullen, A. R. & DeBerardinis, R. J. Genetically-defined metabolic reprogramming in cancer. *Trends Endocrinol. Metab.* **23,** 552–559 (2012).

64. Dang, L. *et al.* Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* **462,** 739–744 (2009).

65. King, A., Selak, M. A. & Gottlieb, E. Succinate dehydrogenase and fumarate hydratase: linking mitochondrial dysfunction and cancer. *Oncogene* **25,** 4675–4682 (2006).

66. Maddocks, O. D. K. *et al.* Serine starvation induces stress and p53-dependent metabolic remodelling in cancer cells. *Nature* **493,** 542–546 (2014).

67. Fan, J. *et al.* Human phosphoglycerate dehydrogenase produces the oncometabolite D-2-hydroxyglutarate. *ACS Chem. Biol.* **10,** 510–516 (2015).

68. Son, J. *et al.* Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway. *Nature* **496,** 101–105 (2014).

69. Fan, J. *et al.* Human Phosphoglycerate Dehydrogenase Produces the Oncometabolite d-2-Hydroxyglutarate. *ACS Chem. Biol.* (2014). doi:10.1021/cb500683c

70. Rohle, D. *et al.* An inhibitor of mutant IDH1 delays growth and promotes differentiation of glioma cells. *Science* **340,** 626–630 (2013).

71. An, S., Kumar, R., Sheets, E. D. & Benkovic, S. J. Reversible Compartmentalization of de Novo Purine Biosynthetic Complexes in Living Cells. *Science* **320,** 103–106 (2008).

72. Anderson, D. D., Quintero, C. M. & Stover, P. J. Identification of a de novo thymidylate biosynthesis pathway in mammalian mitochondria. *Proc Natl Acad Sci* **108,** 15163–15168 (2011).

73. Anderson, D. D., Woeller, C. F., Chiang, E.-P., Shane, B. & Stover, P. J. Serine hydroxymethyltransferase anchors de novo thymidylate synthesis pathway to nuclear lamina for DNA synthesis. *J. Biol. Chem.* **287,** 7051–7062 (2012).

74. Mattaini, K. R. *et al.* An epitope tag alters phosphoglycerate dehydrogenase structure and impairs ability to support cell proliferation. *Cancer Metab* **3,** 5 (2015).

75. Lewis, C. A. *et al.* Tracing compartmentalized NADPH metabolism in the cytosol and mitochondria of mammalian cells. *Molecular Cell* **55,** 253–263 (2014).

76. Barker, G. A. & Ellory, J. C. The identification of neutral amino acid transport systems. *Exp. Physiol.* **75,** 3–26 (1990).

77. Acker, M. G. & Auld, D. S. Considerations for the design and reporting of enzyme assays in high-throughput screening applications. *Perspectives in Science* (2014). doi:10.1016/j.pisc.2013.12.001

78. Walsh, D. A. & Sallach, H. J. Purification and Properties of Chicken Liver D-3-Phosphoglycerate Dehydrogenase*. *Biochemistry* **4,** 1076–1085 (1965).

79. Zhang, J. H., Chung, T. D. & Oldenburg, K. R. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *Journal of biomolecular screening* **4,** 67–73 (1999).

80. Vallari, R. C. & Pietruszko, R. Human aldehyde dehydrogenase: mechanism of inhibition of disulfiram. *Science* **216,** 637–639 (1982).

81. Buescher, J. M. *et al.* A roadmap for interpreting 13C metabolite labeling patterns from cells. *Curr. Opin. Biotechnol.* **34,** 189–201 (2015).

82. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483,** 603–607 (2012).

83. Li, B. *et al.* Butyrylcholinesterase, paraoxonase, and albumin esterase, but not carboxylesterase, are present in human plasma. *Biochem. Pharmacol.* **70,** 1673–1684 (2005).

84. Hosokawa, M. Structure and catalytic properties of carboxylesterase isozymes involved in metabolic activation of prodrugs. *Molecules* **13,** 412–431 (2008).

85. Smith, G. F. 1 Designing Drugs to Avoid Toxicity. *Progress in medicinal chemistry* (2011). doi:10.1016/B978-0-12-381290-2.00001-X

86. Tennant, D. A., Durán, R. V. & Gottlieb, E. Targeting metabolic transformation for cancer therapy. *Nat. Rev. Cancer* **10,** 267–277 (2010).

87. Parsons, D. W. *et al.* An integrated genomic analysis of human glioblastoma multiforme. *Science* **321,** 1807–1812 (2008).

88. Wang, F. *et al.* Targeted inhibition of mutant IDH2 in leukemia cells induces cellular differentiation. *Science* **340,** 622–626 (2013).

89. Gottlieb, E. & Tomlinson, I. P. M. Mitochondrial tumour suppressors: a genetic and biochemical update. *Nat. Rev. Cancer* **5,** 857–866 (2005).

90. Foulkes, W. D., Smith, I. E. & Reis-Filho, J. S. Triple-negative breast cancer. *N. Engl. J. Med.* **363,** 1938–1948 (2010).

91. Boolell, V., Alamgeer, M., Watkins, D. N. & Ganju, V. The Evolution of Therapies in Non-Small Cell Lung Cancer. *Cancers (Basel)* **7,** 1815–1846 (2015).

92. Anastasiou, D. *et al.* Pyruvate kinase M2 activators promote tetramer formation and suppress tumorigenesis. *Nat. Chem. Biol.* **8,** 839–847 (2012).

93. Kung, C. *et al.* Small Molecule Activation of PKM2 in Cancer Cells Induces Serine Auxotrophy. *Chemistry & Biology* **19,** 1187–1198 (2012).

94. Hitosugi, T. *et al.* Phosphoglycerate mutase 1 coordinates glycolysis and biosynthesis to promote tumor growth. *Cancer Cell* **22,** 585–600 (2012).

95. Yuan, M., Breitkopf, S. B., Yang, X. & Asara, J. M. A positive/negative ion-switching, targeted mass spectrometry-based metabolomics platform for bodily fluids, cells, and fresh and fixed tissue. *Nat Protoc* **7,** 872–881 (2012).

96. Grant, G. A. Contrasting catalytic and allosteric mechanisms for phosphoglycerate dehydrogenases. *Arch. Biochem. Biophys.* **519,** 175–185 (2012).

97. Prasad, P. V. & Hatefi, Y. Inactivation of D-(-)-. beta.-hydroxybutyrate dehydrogenase by modifiers of carboxyl and histidyl groups. *Biochemistry* **25,** 2459–2464 (1986).

98. Ma, L. *et al.* Control of Nutrient Stress-Induced Metabolic Reprogrammingby PKCz in Tumorigenesis. *Cell* **152,** 599–611 (2013).

99. Singh, J., Petter, R. C., Baillie, T. A. & Whitty, A. The resurgence of covalent drugs. *Nat Rev Drug Discov* **10,** 307–317 (2011).

100. Liu, Q. *et al.* Developing Irreversible Inhibitors of the Protein Kinase Cysteinome. *Chemistry & Biology* **20,** 146–159 (2013).

101. Lamzin, V. S., Dauter, Z. & Wilson, K. S. Dehydrogenation through the looking-glass. *Nat. Struct. Biol.* **1,** 281–282 (1994).

102. Hou, R., Chen, Z., Yi, X., Bian, J. & Xu, G. Catalytic reaction mechanism of L-lactate dehydrogenase: anab initio study. *Sc. China Ser. B-Chem.* **43,** 587–599 (2000).

103. Minarik, P., Tomaskova, N. & Kollarova, M. Malate dehydrogenases-structure and function. *General physiology and Biophysics* (2002).

104. Liu, Q. *et al.* Developing Irreversible Inhibitors of the Protein Kinase Cysteinome. *Chemistry & Biology* **20,** 146–159 (2013).

105. Grant, G. A., Schuller, D. J. & Banaszak, L. J. A model for the regulation of D-3-phosphoglycerate dehydrogenase, a Vmax-type allosteric enzyme. *Protein Sci.* **5,** 34–41 (1996).

106. Tabatabaie, L. *et al.* Novel mutations in 3-phosphoglycerate dehydrogenase (PHGDH) are distributed throughout the protein and result in altered enzyme kinetics. *Hum. Mutat.* **30,** 749–756 (2009).

107. Pind, S. *et al.* V490M, a Common Mutation in 3-Phosphoglycerate Dehydrogenase Deficiency, Causes Enzyme Deficiency by Decreasing the Yield of Mature Enzyme. *J Biol Chem* **277,** 7136–7143 (2002).

108. Klomp, L. W. *et al.* Molecular characterization of 3-phosphoglycerate dehydrogenase deficiency--a neurometabolic disorder associated with reduced L-serine biosynthesis. *Am. J. Hum. Genet.* **67,** 1389–1399 (2000).

Supplementary Materials

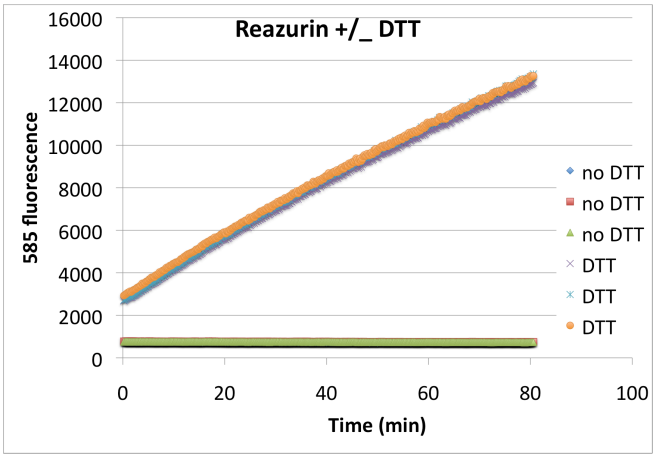


Figure S2.1. DTT spontaneously reduces resazurin. Resazruin was incubated in Tris (pH 8.5) buffer with or without DTT and the fluorescence of resorufin (550/585nm Ex/Em) measured over time measured.

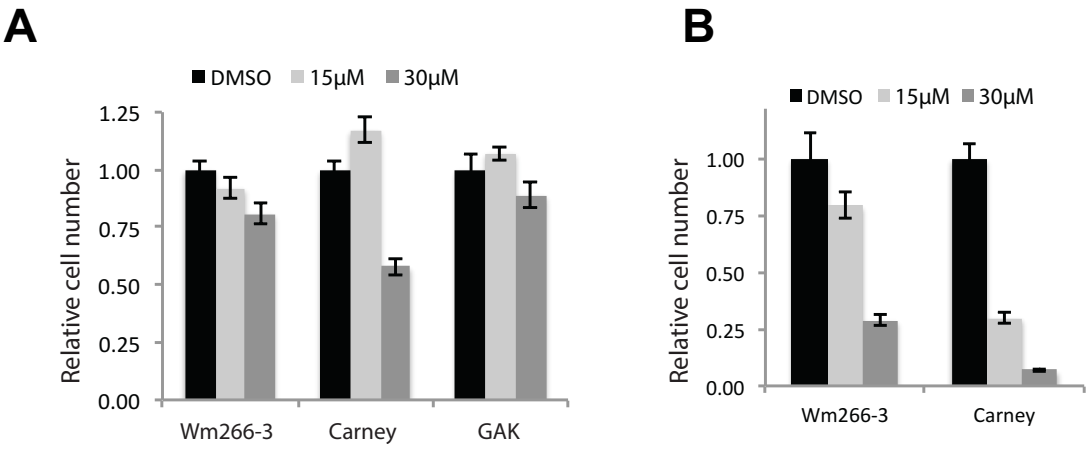


Figure S2.2. CBR-5884 selectively inhibits the proliferation of melanoma lines with a high propensity for serine synthesis. Proliferation assay for melanoma lines treated at the indicated CBR-5884 concentration in (A) serine replete (+SER) or (B) serine deplete (-SER) media. Wm266-3 cells are *PHGDH* amplified. GAK cells were not included in –SER experiments because they are sensitive to serine withdrawal. Histograms depict mean ± standard error (n ≥ 3).

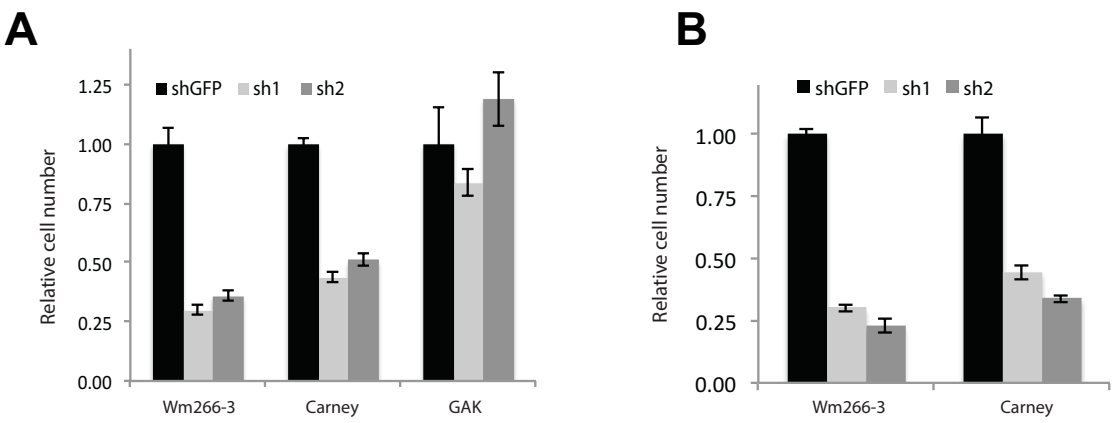


Figure S2.3. Knockdown of PHGDH in a panel of melanoma lines. Proliferation assay for lines grown in (A) serine replete media (+SER) or (B) serine deplete media (-SER) with PHGDH knockdown (sh1 & sh2) or a nontargeting control (shGFP). GAK cells were not included in –SER experiments because they are sensitive to serine withdrawal. Histograms depict mean ± standard error (n ≥ 3).

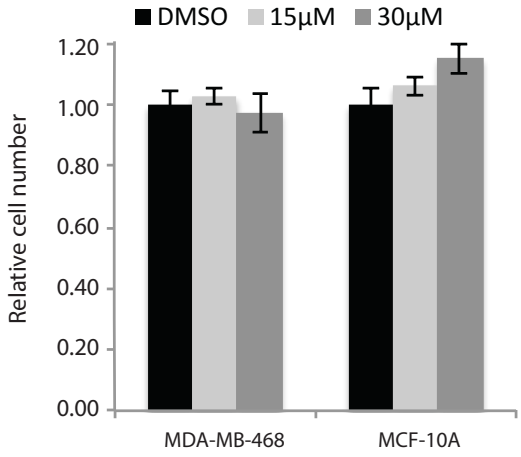


Figure S2.4. The acid derivative of CBR-5884 is not active on cells. MDA-MB-468 cells (PHGDH amplified) and MCF-10A cells were treated with the acid derivative of CBR-5884 for three days before assaying relative cell numbers. Despite CBR-5884 strongly inhibiting the proliferation of MDA-MB-468 cells, the acid derivative was unable to inhibit MDA-MB-468 proliferation.

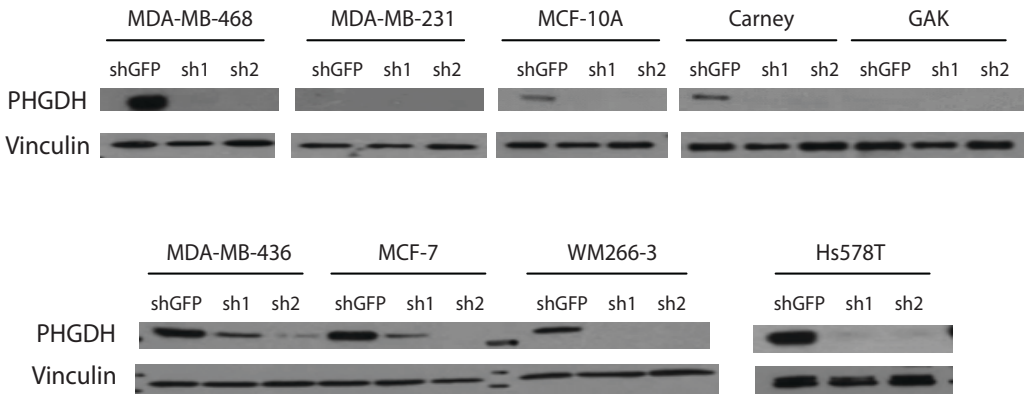
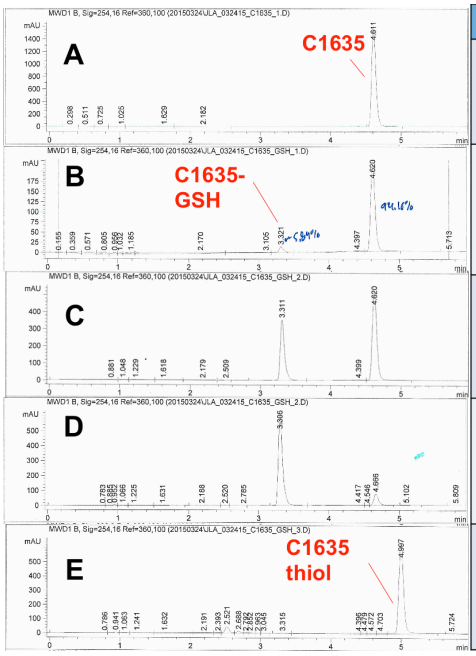


Figure S2.5. Western blot confirming PHGDH knockdown in breast and melanoma lines. On the day of seeding for the proliferation assays in serine replete or deplete media, the indicated lysates were harvested and PHGDH knockdown was confirmed by western blotting. PHGDH knockdown: sh1 and sh2. Nontargeting control: shGFP.



Conditions	% C1635	% C1635-GSH Disulfide
C1635 in ACN (Control)	100%	-
5 mM GSH, 5 mM C1635 **0.5 hr** *C1635 insoluble	94%	6%
5 mM GSH, 5 mM C1635 **1 hr** Added MeOH/ACN to solubilize C1635	62%	38%
~2 mM GSH, ~2 mM C1635 **10 hr**	17%	83%
~2 mM GSH, ~2 mM C1635 **10 hr** DTT added (excess)	0%	0% (~100% reduced to thiol)

Figure S3.1. Chromatograms from LCMS experiments to detect CBR-5884 and GSH adducts. The *tert*-butyl derivative of CBR-5884 at 5mM (labeled C1635 in figure) was incubated with 5mM glutathione (1:1 molar ratio) and analyzed by LCMS. CBR-5884 and glutathione were incubated for (A) 0h, (B) 0.5h, (C) 1h, and (D) 10h prior to LCMS analysis. To confirm that the species was indeed the predicted CBR-5884 and GSH adduct with loss of cyanide, an excess of (E) dithiothreitol was added which resulted in nearly complete recovery of a compound consistent with the free thiol by mass spectrometry. On the left, LC chromatograms are presented. Labeled peaks were identified via MS. C-1635, *tert*-butyl derivative of CBR-5884. C-1635-GSH, glutathione adduct. C-1635 thiol, free thiol resulting from cyanide loss.

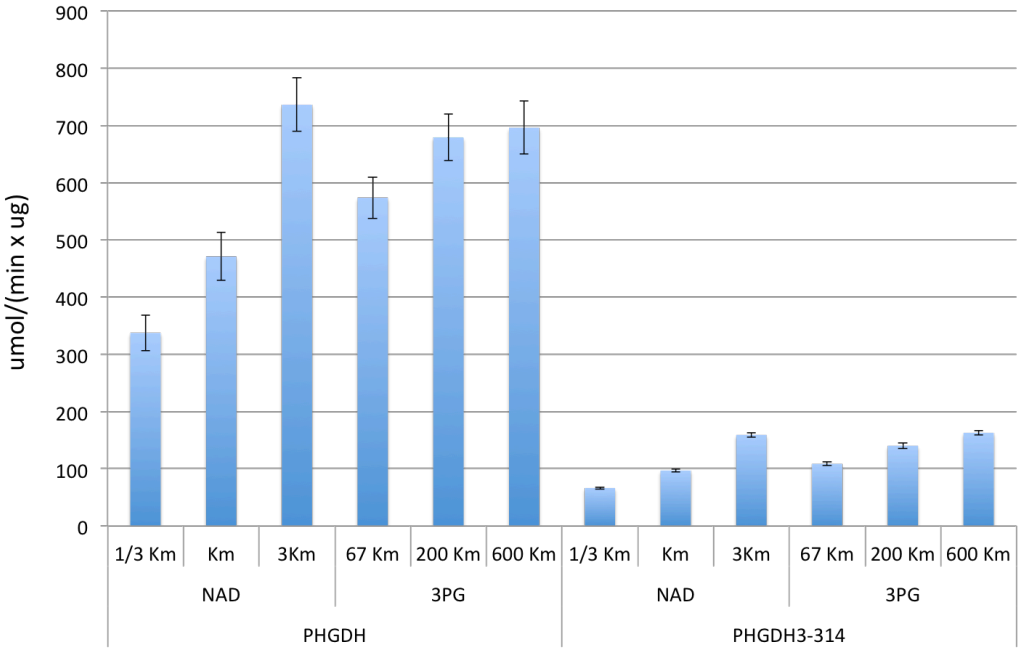


Figure S3.2. Specific activity of full length PHGDH versus PHGDH$^{3-314}$. The specific activity of full length PHGDH and truncated PHGDH$^{3-314}$ was determined from the initial reaction rate normalized to the protein input. Initial rates were determined at a variety of substrate combinations as indicated.