

A protein therapeutic modality founded on molecular regulation

Chapman M. Wright^a, R. Clay Wright^a, James R. Eshleman^b, and Marc Ostermeier^{a,1}

^aDepartment of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, MD 21218; and ^bDepartments of Pathology and Oncology, Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins University School of Medicine, Baltimore, MD 21231

Edited by David Baker, University of Washington, Seattle, WA, and approved August 19, 2011 (received for review February 18, 2011)

The exquisite specificity of proteins is a key feature driving their application to anticancer therapies. The therapeutic potential of another fundamental property of proteins, their ability to be regulated by molecular cues in their environment, is unknown. Here, we describe a synthetic biology strategy for designing protein therapeutics that autonomously activate a therapeutic function in response to a specific cancer marker of choice. We demonstrate this approach by creating a prodrug-activating enzyme that selectively kills human cancer cells that accumulate the marker hypoxia-inducible factor 1 α . This property arises primarily through increased cellular accumulation of the enzyme in the presence of the marker. Our strategy offers a platform for the development of inherently selective protein therapeutics for cancer and other diseases.

directed evolution | protein engineering | protein switch | enzyme/prodrug therapy

The attractiveness of proteins as therapeutics stems in part from the exquisite specificity by which they execute diverse functions—e.g., they catalyze exactly the right reaction or inhibit exactly the right cell receptor. Proteins possess another fundamental property central to life whose therapeutic potential has not been exploited—the ability to be regulated at the protein level by molecular signals. Prevailing approaches to cancer protein therapeutics focus on cancer marker recognition/modulation and downstream effects of that modulation (1). Such an approach is limiting because the therapeutic mechanism is restricted to those that naturally arise from modulation of the cancer marker. Furthermore, this approach precludes the use of cancer markers for which modulators cannot be found or for which modulation does not produce a therapeutic effect (2). In addition, many potential protein therapies lack the desired selective cancer cell targeting. The ability to link recognition of any cancer marker with activation of any desired therapeutic function would enormously expand the number of possible protein therapeutics. Here is where the therapeutic potential of regulation can be realized.

One approach to establishing unique regulatory relationships is to build proteins that function as switches—proteins whose cellular level of activity is modulated through interactions with an input signal such as a protein or small molecule. Our switch design strategy views all proteins as an extensive parts list from which switches can be built using domains with the prerequisite input and output functions of the desired switch. The design challenge for this approach is how to fuse the input (i.e., signal recognizing) and output domains (i.e., the function to be modulated) such that the input domain regulates the output domain's function. We have explored a directed evolution approach to this general design problem in which libraries of random insertions of one domain into the other are subjected to selections and screens designed to identify library members with switching behavior (3–5). Our approach extensively explores the geometric space of how two protein domains can be fused through insertion of one domain into another (3, 4). For example, we have identified switch proteins with maltose-activated β -lactamase activity from libraries of circularly permuted β -lactamase genes inserted into the gene encoding maltose-binding protein (MBP) (4, 5).

Many of these switches function as allosteric enzymes with maltose binding to the MBP domain inducing conformational changes in the β -lactamase domain that affect its catalytic activity (4, 6). For other switches, maltose binding does not affect the specific activity of the protein but instead increases the accumulation of the switch protein in the cell (7).

Here, we propose a synthetic biology strategy for designing protein therapeutics that link activation of a chosen therapeutic function to a specific cancer marker of choice. We demonstrate this strategy by creating a protein switch that renders human colon and breast cancer cells susceptible to the prodrug 5-fluorocytosine (5FC) in response to the cancer marker hypoxia-inducible factor 1 α (HIF-1 α).

Results and Discussion

A Protein Therapeutic Modality Based on Regulation. Our approach greatly expands the number of possible therapeutic strategies by providing a platform for how a cancer marker of choice can be used to trigger any therapeutic function of choice (Fig. 1). Central to this strategy is the development of a protein switch that couples recognition of the cancer marker to a therapeutic function. Such coupling could arise via newly established allosteric interactions (Fig. 1*B*) or through increased cellular accumulation in the presence of the marker (Fig. 1*C*). We sought to demonstrate this strategy by linking the unrelated functions of recognition of tumor-marker HIF-1 α and the enzymatic production of the chemotherapeutic 5-fluorouracil (5FU) from the prodrug 5FC by cytosine deaminase (Fig. 2*A*). HIF-1 α accumulates to a high level in many solid tumors, including breast, prostate, colorectal, and pancreatic cancers (8–10) but is virtually undetectable in normal, well-oxygenated tissues (11–13). Cancer cells that contain high levels of HIF-1 α are able to survive under extreme conditions, are resistant to therapy, and have a greater potential for metastasis (10, 14–18). Thus, the linking of HIF-1 α levels to 5FU production could in theory attack the most aggressive tumors, while having limited or no effect on well-oxygenated, normal cells. Such a strategy, like any strategy that utilizes an intracellular target, faces the challenge of efficient delivery of the therapeutic gene or protein to the cancer cell.

We constructed our switch using the CH1 domain from the human p300 protein as the HIF-1 α recognition input domain and yeast cytosine deaminase (yCD) as the prodrug activation output domain (Fig. 2*A*). The p300/CBP protein complex binds HIF-1 α in the cytoplasm and translocates it to the nucleus (13,

Author contributions: C.M.W., J.R.E., and M.O. designed research; C.M.W. and R.C.W. performed research; C.M.W., R.C.W., J.R.E., and M.O. analyzed data; and C.M.W. and M.O. wrote the paper.

Conflict of interest statement: M.O. and C.M.W. declare a conflict of interest in the form of a patent application.

This article is a PNAS Direct Submission.

¹To whom correspondence may be addressed at: Johns Hopkins University Department of Chemical and Biomolecular Engineering, 3400 North Charles Street, Baltimore, MD 21218. E-mail: oster@jhu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1102803108/-DCSupplemental.

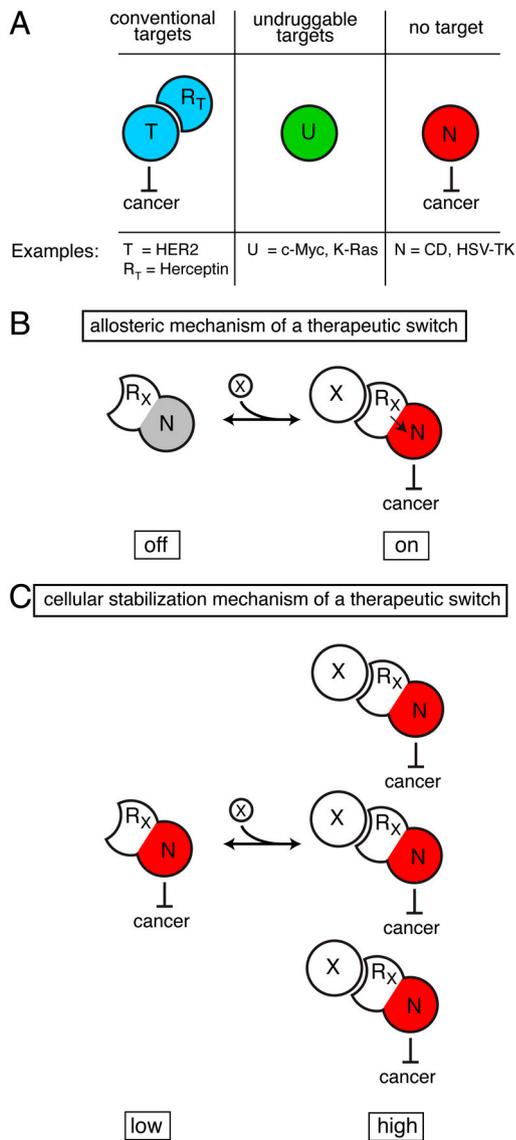


Fig. 1. A therapeutic modality based on regulation expands potential strategies. (A) Cancer marker T is a conventional protein target in which recognition and modulation of T by protein R_T causes a therapeutic effect. Cancer marker U is considered undruggable: Either a protein capable of recognizing and modulating U cannot be found or a therapeutic benefit for modulation of U does not exist. N is a protein with inherent therapeutic activity but is not targeted to cancer cells. (B) A therapeutic switch functioning by an allosteric mechanism couples R_X's recognition of cancer marker X to a therapeutic effect mediated by protein N. The switch is a fusion protein of R_X and N such that binding of X regulates N's therapeutic effect. (C) A therapeutic switch functioning by a cellular stabilization mechanism couples R_X's recognition of cancer marker X to cellular accumulation. The switch's therapeutic activity is higher in the presence of X, not because of change in specific activity of the switch, but because the switch accumulates to a higher level in the cell. These two switch mechanisms are not mutually exclusive and either can cause an otherwise nonspecific therapeutic effect to depend on the presence of any cancer maker, whether conventional or undruggable.

19–21). CH1 is a small 100 amino acid domain of the p300 protein that retains full HIF-1a affinity (19, 21). CH1 interacts with other cancer markers besides HIF-1a such as p53 (20) and HIF-2a (10)—offering other cancer-specific signals that could cause prodrug activation by our desired switch. yCD is an enzyme that converts 5FC to 5FU. We utilized the triple mutant of yCD (A23L/V108I/I140L) shown to have increased thermostability and full enzyme efficiency (22). 5FC is a nontoxic compound (LD₅₀ → 15 g/kg in rats), and 5FU is a well-established, highly active chemother-

apeutic used to treat multiple types of cancer, including breast, colorectal, and pancreatic cancers. Cytosine deaminase has been previously shown to activate the prodrug 5FC into the toxin 5FU in animal models (23) and the 5FC/CD combination has been used in gene-directed enzyme prodrug therapy (GDEPT) treatments in clinical trials (24).

Library Creation and Selection. We constructed a library in which DNA encoding the CH1 domain was randomly inserted into a plasmid encoding yCD (Fig. 2B, Fig. S1, and SI Text). Our yCD-CH1 library contained approximately 10 million members of which 25% had the DNA encoding the CH1 domain inserted somewhere in the yCD gene. We devised a two-tier genetic selection to isolate yCD-CH1 hybrids that exhibit low cellular activity in the absence of HIF-1a, but high cellular activity in the presence of HIF-1a (Fig. 2C). The library was transformed into GIA39 *Escherichia coli*—a uracil auxotroph that lacks a functional *E. coli* cytosine deaminase (25, 26). In the negative selection tier, library members deficient in 5FC to 5FU conversion in the absence of HIF-1a were selected. In the subsequent positive selection tier, library plasmids were transformed into GIA39 cells that harbored a plasmid encoding a fusion protein of GST and the C-TAD domain of HIF-1a under the control of the arabinose promoter. Freedman et al. demonstrated that this fusion (referred to here as “gstHIF-1a”) and the CH1 domain maintain the high affinity of the full-length proteins and could be copurified when they are expressed within the same *E. coli* cell (20). Our selections resulted in the identification of two recombinant genes that conferred to GIA39 cells a gstHIF-1a-dependent sensitivity to 5FC. These nearly identical proteins were named “Haps3” and “Haps59” for HIF-1a activated protein switch (Fig. 3A).

Characterization of Haps Proteins in Bacteria. Initial experiments suggested that *haps59* conferred to *E. coli* the greatest HIF-1a dependence on 5FC sensitivity. To quantify the effect, the plasmid encoding Haps59 was isolated and retransformed into fresh GIA39 cells harboring either the *gstHIF-1a* plasmid or an analogous negative control plasmid encoding only GST. These cells and appropriate control cells were challenged to grow on agar plates containing 5FC (Fig. 3B and Fig. S2). Only the combination of the presence of the *gstHIF-1a* gene and the addition of arabinose, which induces *gstHIF-1a* expression, increased the 5FC sensitivity of cells expressing Haps59. Qualitatively similar results were obtained in liquid media (Fig. S3). The results indicate that HIF-1a significantly increases the 5FC deaminase activity of cells expressing Haps59.

Copurification experiments indicated that Haps59 interacts with *gstHIF-1a* in *E. coli* (Fig. S4). Presumably this interaction must allosterically activate the cytosine deaminase activity of the switch and/or cause the increased accumulation of the switch in vivo. We observed a substantial increase in the accumulation of Haps59 when *gstHIF-1a* was coexpressed in vivo (Fig. 3C). Purified Haps59 exhibited a small increase in the 5FC deaminase activity in the presence of *gstHIF-1a*, but poor Haps59 stability complicated quantification of the effect (see SI Text). We conclude that increased accumulation of Haps59 in the presence of *gstHIF-1a* is likely to be the major mechanism by which *haps59* confers HIF-1a-dependent sensitivity to 5FC (i.e., the mechanism of Fig. 1C). The CH1 domain has a molten globule state in the absence of HIF-1a and transforms into a stable, structurally well-defined domain upon binding to HIF-1a (27). HIF-1a-induced stabilization of the CH1 domain of Haps59 might stabilize the entire fusion resulting in increased cellular accumulation. However, why Haps59 and not other yCD-CH1 fusions would have this property is a challenging problem. This result highlights both the difficulty in designing proteins that function in the complex cellular environment and the power of directed evolution to iden-

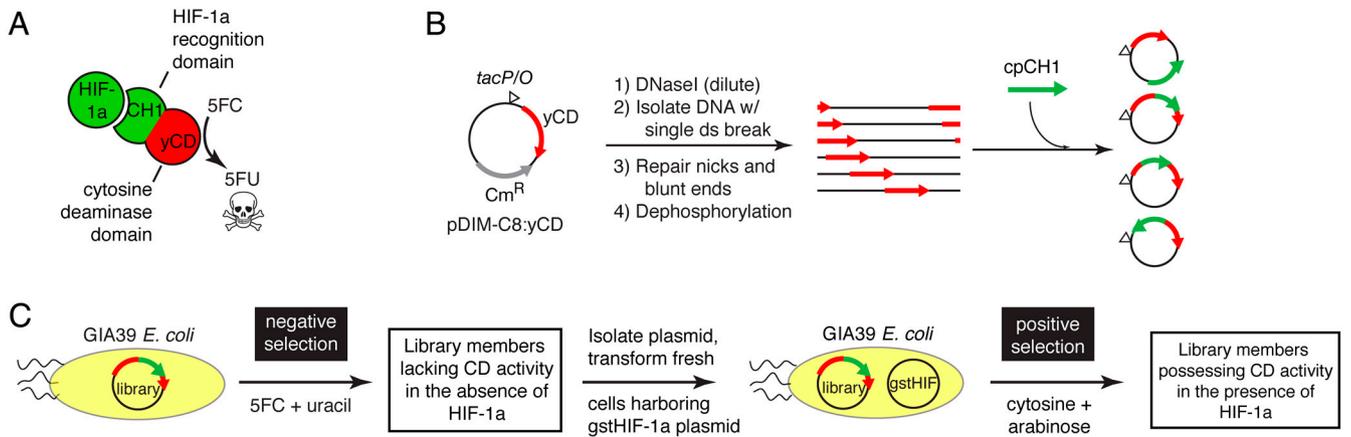


Fig. 2. Protein switch concept, creation, and isolation. (A) The desired switch is a fusion protein between a HIF-1a binding domain (CH1) and yeast cytosine deaminase (yCD) that increases the cellular conversion of 5FC to 5FU in the presence of HIF-1a. (B) Library creation. A plasmid containing the yCD gene is randomly linearized and ligated to a collection of circularly permuted CH1 genes (cpCH1; see Fig. S1), resulting in a library of random insertions of cpCH1 into the plasmid. (C) Two-tiered genetic selection for yCD-CH1 hybrids that possess cytosine deaminase activity only in the presence of HIF-1a.

tify proteins that provide the desired cellular phenotype when the appropriate genetic selection is applied.

Haps59 Increases the 5FC Sensitivity of Cancer Cells Expressing HIF-1a.

We next examined whether Haps59 would function in human cancer cells and confer sensitivity to 5FC selectively under conditions that cause HIF-1a to accumulate. Colorectal solid tumors accumulate high levels of HIF-1a (8–10). RKO is a colorectal cancer cell line that is known to accumulate high levels of HIF-1a in hypoxia (28) and is sensitive to 5FU in culture (29). 5FC was highly toxic to stable cell lines of RKO expressing yCD ($LC_{50} \approx 15 \mu\text{M}$) but nontoxic up to approximately 2 mM to stable cell lines constructed with the empty vector control (Fig. S5). We chose the CMV promoter to control expression of Haps59 and yCD because hypoxic conditions do not change the level of expression from this promoter (30). Initial experiments used the addition of Co^{2+} to the media to cause HIF-1a accumulation. Co^{2+} disrupts the degradation pathway of HIF-1a, allowing the protein to accumulate in the cytoplasm of the cell (31) (Fig. 4A). In the presence of Co^{2+} , Haps59-expressing RKO cell lines were 10-fold more sensitive to 5FC (Fig. 4B). Hypoxic growth condi-

tions (1% O_2) also caused a marked increase in the 5FC sensitivity of these cells (Fig. 4B). Neither hypoxic growth conditions nor the addition of Co^{2+} caused 5FC sensitivity to the empty vector control cells or altered the 5FC sensitivity of cells expressing yCD (Fig. 4B and Fig. S5). Additionally, neither the presence of Co^{2+} nor hypoxic conditions were inherently toxic to Haps59-expressing RKO cells (Fig. S6A).

Analogous experiments with MCF7 breast cancer cells illustrated the generality of the approach (Fig. 4D and E). Mirroring results with *E. coli* cells, Haps59 accumulated at a higher level in the presence of HIF-1a in both RKO and MCF7 cells (Fig. 4C and F), whereas yCD did not (Fig. S7A). This finding further supports the hypothesis that the increase in 5FC sensitivity results from a Haps59-HIF-1a interaction that causes increased accumulation of Haps59. This interaction was substantiated in RKO cells by coimmunoprecipitation experiments (Fig. S7B) and is the designed input signal to activate 5FU production. We confirmed the production of 5FU in Haps59-expressing RKO cells and found the production to be 2.5-fold higher in cells exposed to Co^{2+} (Fig. S6B). These results suggest that Haps59 functions as

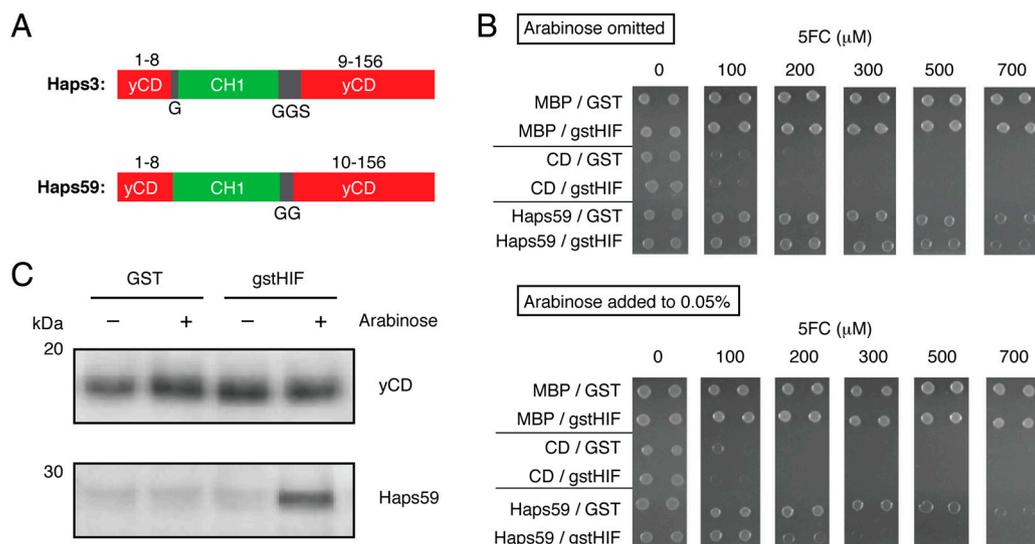


Fig. 3. Characterization of Haps59 in *E. coli*. (A) Switches contain a complete, noncircularly permuted CH1 domain (green) flanked by short linker sequences of the indicated sequences (black) inserted after amino acid 8 of yCD (red). The black numbers indicate the corresponding amino acid numbers in the wild-type yCD protein. (B) Growth of GIA39 cells expressing MBP, yCD, or Haps59 and either GST or gstHIF-1a on minimal media agar plates as a function of 5FC concentration. (Top) In the absence of arabinose and (Bottom) in the presence of arabinose to induce GST and gstHIF-1a expression. (C) Accumulation of yCD and Haps59 in GIA39 *E. coli* in the absence and presence of coexpressed GST or gstHIF-1a as detected by Western blot with anti-yCD antibodies.

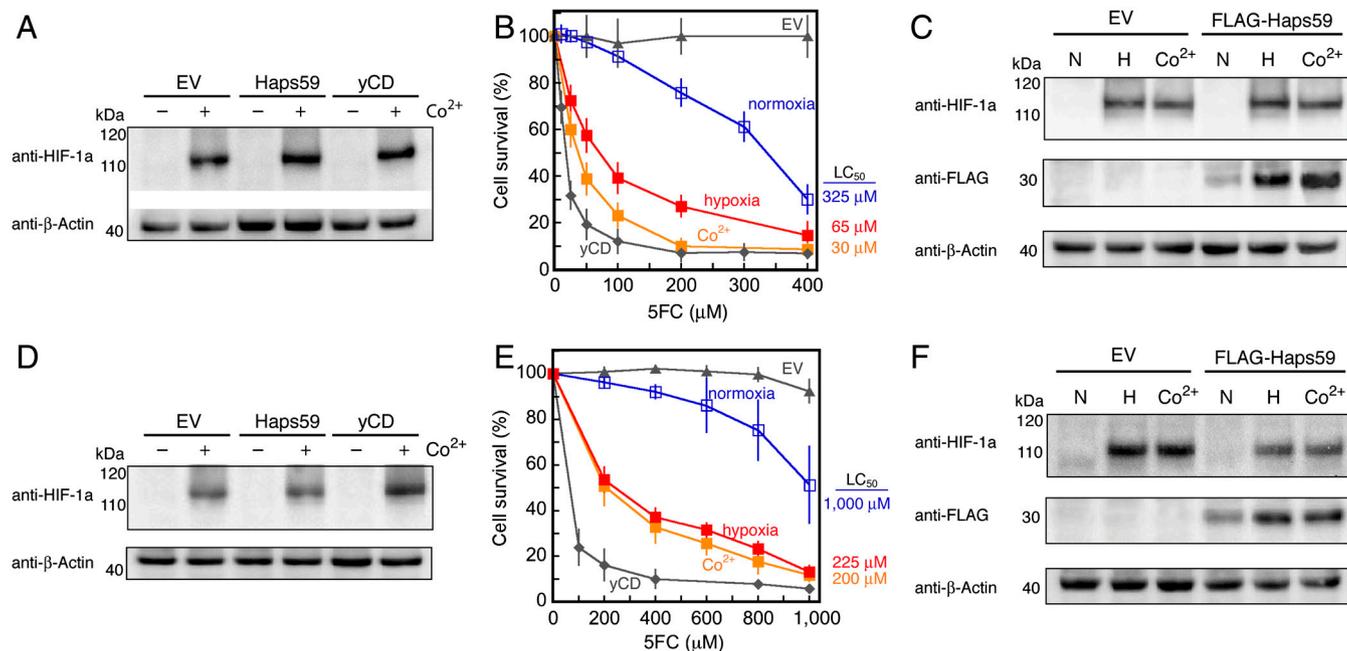


Fig. 4. Characterization of Haps59 in human cancer cells. (A–C) RKO colorectal cancer cells; (D–F) MCF7 breast cancer cells. (A and D) Co^{2+} causes accumulation of HIF-1a in cells expressing Haps59, yCD, or the empty vector control as detected by Western blot with anti-HIF-1a antibodies. (B and E) Both Co^{2+} (orange solid squares) and exposure to hypoxic growth conditions (red solid squares) increase the 5FC sensitivity of Haps59-expressing cells compared to Haps59-expressing cells not exposed to those conditions (open blue squares). The 5FC sensitivities of cells expressing yCD (solid diamonds) and EV control (solid triangle) are shown for comparison (see Fig. S5 for additional controls). Each point represents the mean from three different clones of each cell line, and each clone was tested in three separate experiments (error bars, SD, $N = 9$). (C and F) Haps59 accumulates at higher levels in cells when HIF-1a is present. In these experiments a FLAG epitope was appended to the N terminus of Haps59 (FLAG-Haps59) for switch detection by Western blot. Cells were cultured under normoxic (N), hypoxic (H), or normoxic conditions with the addition of 100 μM cobalt (Co^{2+}). Detection of β -actin served as a loading control.

designed—increasing 5FU production in cancer cells in response to HIF-1a.

Haps59's Potential as a Cancer Therapeutic. HIF-1a is a molecular signature of cancer cells that can survive under extreme conditions, are resistant to therapy, and have a greater potential for metastasis (14–18). A treatment method that could specifically target these cells could significantly improve the effectiveness of cancer treatments when used in combination with traditional therapies. Haps59 is designed to establish a direct relationship between HIF-1a levels and the intracellular production of a chemotherapeutic drug. Haps59's therapeutic potential is derived from this unique regulatory property. Our establishment of this regulation results in a complex protein that autonomously “decides” whether its drug-producing capability should be activated.

Hypoxic regions within a solid tumor can vary from 10–80% (32), and aberrant HIF-1a accumulation as a result of both gain-of-function and loss-of-function mechanisms has been observed experimentally and clinically in normoxic conditions (10). Thus, a high percentage of solid tumor cells can potentially activate Haps59, and non-Haps59 expressing tumor cells would be susceptible to the strong 5FU bystander effect (33, 34). Haps59 exploits an intracellular cancer marker for this activation. Prevailing strategies for protein cancer therapeutics target more-readily accessible extracellular cancer markers such as membrane-bound cell receptors. However, it's been estimated that <10% of the human genome codes for cell surface proteins, limiting the available targets for therapy (2). There is a great need to develop effective targeting strategies, and this might be accomplished through intracellular cancer markers. Such strategies will face delivery challenges: Either the therapeutic protein or its corresponding gene must be efficiently delivered inside the cell. However, our placement of the selectivity at the protein level overcomes serious limitations of existing gene therapy approaches that require selective targeting of gene delivery to malignant cells. For example,

the combination of specific delivery of a cytosine deaminase gene to cancer cells followed by 5FC administration is an example of GDEPT (33, 34). One major limitation to existing CD/5FC GDEPT strategies is the poor transmission efficiency of the CD gene using current viral vectors because of the need for transmission specifically to cancer cells (35). Our strategy overcomes this limitation by moving the specificity from the transductional level to the protein level, allowing efficient means of gene delivery to be used regardless of cell-type specificity. In addition, our approach is complementary to both transcriptional (36) and transductional targeting and might be combined with these approaches to afford a double or triple layer of specificity: at the gene delivery level, at the transcription level, and at the protein level. Recent advances in protein delivery offer a potential route toward achieving efficient delivery without gene therapy (37, 38).

Nature chose proteins as a molecule of choice to carry out a wide array of specific, intricately regulated functions. Our increasing ability to emulate nature and design proteins capable of complex, carefully regulated functions offers an attractive path to achieving the effective and selective therapeutics we seek. By connecting disparate functions, protein switches expand the scope of potential therapeutic strategies in a manner that inherently encompasses specificity.

Methods

Library Creation and Selection. Creation of the yCD-CH1 hybrid library is described in detail in *SI Text*. For selections, yCD-CH1 library plasmid DNA was isolated from an aliquot of DH5 α *E. coli* cells, and 25 ng of the purified plasmid was used to transform GIA39 *E. coli* cells (Coli Genetic Stock Center# 5594) by electroporation. The transformed cells were plated on LB agar containing 50 $\mu\text{g}/\text{mL}$ chloramphenicol and then recovered using a sweep buffer (1 \times M9 salts containing 2% glucose and 15% glycerol) and stored in aliquots at -80°C . These cells were first plated on 24.5 \times 24.5 cm minimal media plates (1 \times nitrogen base, 1 \times yeast synthetic dropout without uracil, 2% glucose, and 20 g/L select agar) containing 75 $\mu\text{g}/\text{mL}$ 5FC, 1 $\mu\text{g}/\text{mL}$ uracil, 1 mM IPTG, and 50 $\mu\text{g}/\text{mL}$ chloramphenicol. Cells were plated at a

concentration of approximately 500,000 cfu/plate, with cfu defined on minimal media plates without 5FC but with 1 $\mu\text{g}/\text{mL}$ uracil. Colonies on these plates were recovered into sweep buffer and replated at the same cfu density on the same type of plate media except the 5FC concentration was lowered to 50 $\mu\text{g}/\text{mL}$. Library members surviving this second negative selection plate were recovered into sweep buffer, aliquoted, and stored at -80°C . Library plasmid DNA was isolated from these samples and used for the positive selection.

GIA39 cells harboring a *gstHIF-1a* plasmid were cotransformed with library plasmids from the negative selection and plated on LB media containing 100 $\mu\text{g}/\text{mL}$ ampicillin and 50 $\mu\text{g}/\text{mL}$ chloramphenicol. Cotransformed GIA39 cells were recovered from the plate using sweep buffer, aliquoted, and stored as described above. These cells were plated on 24.5×24.5 cm minimal media plates supplemented with 25 $\mu\text{g}/\text{mL}$ cytosine, 1 mM IPTG, 0.15% arabinose, 50 $\mu\text{g}/\text{mL}$ chloramphenicol, and 100 $\mu\text{g}/\text{mL}$ ampicillin at a density of approximately 500,000 cfu/plate. A total of 99 colonies formed on this selection plate, all of which were screened by colony PCR for *pDIM-C8* plasmids that contained the CH1 insert within the *yCD* gene. This was performed using primers that anneal outside the *yCD* gene followed by gel electrophoresis to observe the size of the PCR product. About 20% of these colonies harbored library members that contained a CH1 insert within the *yCD* gene, and these members were sequenced. Eight members of the 20 sequenced were in-frame, of which Haps3 appeared twice and Haps59 appeared three times. In-frame members were then individually replated on selection media to ensure their switching behavior in the presence of *gstHIF-1a*. Only Haps3 and Haps59 behaved as switches after replating.

5FC Sensitivity Assays with GIA39 Cells on Solid Media. Fresh GIA39 cell lines harboring a *pDIM-C8* plasmid for expression of MBP (negative control), *yCD* (positive control), or Haps59 were cotransformed with either the *gstHIF-1a* plasmid or an analogous negative control plasmid encoding only GST were created. These six cell lines were cultured to midlog phase (0.3 OD) in minimal media and then serially diluted with minimal media 3.3-fold in 96-well format. Minimal media consisted of 1 \times nitrogen base, 1 \times yeast synthetic dropout without uracil, 2% glucose, 5 $\mu\text{g}/\text{mL}$ uracil, 100 $\mu\text{g}/\text{mL}$ of ampicillin, and 50 $\mu\text{g}/\text{mL}$ chloramphenicol. One microliter of each cell line dilution was spotted on minimal media plates (OmniTray, 86×128 mm, Nunc) containing 1 mM IPTG, 20 g/L select agar, and different amounts of 5FC (0, 100, 200, 300, 500, and 700 μM). The media either omitted arabinose or contained 0.05% arabinose to induce GST and *gstHIF-1a*. The plates were incubated for 24–36 h at 37°C , and the results of these experiments can be seen in Fig. 3 and Fig. S2.

Haps59 Accumulation Studies in *E. coli*. Four GIA39 cell lines (expressing *yCD* + GST, *yCD* + *gstHIF-1a*, Haps59 + GST, or Haps59 + *gstHIF-1a*) were cultured in 25 mL of minimal media containing 1 \times nitrogen base, 1 \times yeast synthetic dropout without uracil, 2% glucose, 10 $\mu\text{g}/\text{mL}$ uracil, 50 $\mu\text{g}/\text{mL}$ chloramphenicol, and 100 $\mu\text{g}/\text{mL}$ ampicillin overnight at 37°C . These cultures were diluted into eight separate fresh minimal media flasks to compare the addition or omission of arabinose on *yCD* and Haps59 accumulation. These eight cultures were grown to an OD of 0.2, at which point IPTG (to 1 mM) was added to all cultures and 0.05% arabinose was added to half of the flasks to express GST or *gstHIF-1a*. The cultures were then incubated at 37°C for 12–16 h. An equal amount of cells were aliquoted based on final OD measurements for lysis. The bacterial cells (680 million cells per sample) were lysed using BugBuster™ (Novagen) following the manufacturer's protocol. Equal amounts of the cell lysates were separated on a 4–12% Bis-Tris NuPAGE gels (Invitrogen) and then transferred to PVDF membrane (BioRad). The Western protocol is described in *SI Text*.

5FC Sensitivity Experiments in RKO and MCF7 Cells Expressing *yCD* and Protein Switches. RKO cells stably expressing an empty vector (one clone), *yCD* (three clones), or Haps59 (three clones) were used to seed 96-well plates, 1,500 cells per well, in 100 μL of MEM media supplemented with 10% fetal bovine serum and 1% antibiotic and antimycotic (ABAM) (all from Gibco). MCF7 cells stably expressing an empty vector (one clone), *yCD* (three clones), or Haps59 (three clones) were used to seed 96-well plates, 3,500 cells per well, in 100 μL of DMEM media supplemented with 10% fetal bovine serum and 1% ABAM

(all from Gibco). After 24 h, 100 μL of MEM (RKO) or DMEM (MCF7) media containing 5FC (0–2 mM) was added to these cells. These experiments were carried out for 6 d with one media change after 3 d. After 6 d, the cells were lysed by rinsing twice with 100 μL of PBS and then incubated in 100 μL of 0.1% SDS in water. The cells were incubated at 37°C for 2 h in the SDS solution for complete lysis. To detect dsDNA, SYBR green (Invitrogen) was added to a final concentration of 0.075% to lysed cells and the fluorescence for each well was recorded at 520 nm after excitation at 485 nm. The percent survival was calculated using the equation $(A)_{\text{sample}} / (A)_{\text{control}} \times 100$, where $(A)_{\text{sample}}$ is the fluorescence of sample wells with 5FC and $(A)_{\text{control}}$ is the fluorescence of the control well lacking 5FC. The results displayed in Fig. 4 represent the mean with experiments performed on three separate days ($N = 6$ for empty vector (EV), $N = 9$ for *yCD*, and $N = 9$ for Haps59). Error bars are the standard deviation from the mean.

To induce HIF-1a accumulation, MEM or DMEM media containing 75 μM CoCl_2 was used. The presence of Co^{2+} had little effect on the growth of the RKO cells, regardless of whether or not the cells had been transfected with a gene (Fig. S6). Alternatively, a hypoxic environment was induced by incubation of the cancer cells in a MIC-101 hypoxia chamber (Billups-Rothenberg). The chamber was purged with three 3-min purges of 1% O_2 gas containing 5% CO_2 and 94% N_2 over 2 h. A 3-min purge was then used every 24 h to maintain a hypoxic environment for the 6 d. Incubation at 1% O_2 tension also had no effect on cell growth compared to the parental cell lines grown in the same conditions (Fig. S6). After 6 d (with one media change after 3 d), the cells were lysed and the percent survival was calculated using the same protocol as described above. Hypoxic environment experiments were repeated on separate days ($N = 6$ for all), and error was calculated by standard deviation from the mean.

HIF-1a Dependent Accumulation of Haps59 in RKO and MCF7 Cancer Cells. *yCD* antibodies cross-reacted with many human proteins (Fig. S7A). To circumvent this problem, a FLAG™ tag (Sigma-Aldrich) was appended to the N terminus of Haps59 for protein switch accumulation experiments. RKO or MCF7 cells expressing an EV or Haps59 with the N-terminal FLAG™ tag (FLAG-Haps59) were incubated in normoxia (with and without 100 μM CoCl_2), or hypoxia (1% O_2) conditions for 12–18 h. After exposure, cell cultures were incubated at 4°C for 30 min and then treated with radioimmunoprecipitation assay buffer (Sigma) containing protease inhibitor cocktail (Sigma). Cells were scrapped from the flask and placed in a 1.5-mL tube and incubated on ice for 30 min. Cell debris was removed by centrifugation ($20,000 \times g$ for 30 min, 4°C) and the supernatant was transferred to a fresh 1.5-mL tube. Protein concentrations of the whole cell lysates were determined using the detergent compatible protein assay (BioRad). A total of 50 μg of each lysate were separated on 4–12% Bis-Tris NuPAGE gels (Invitrogen) and then transferred to PVDF membrane (BioRad). The membrane was blocked with 3% nonfat milk for 30 min.

Primary anti-FLAG™-HRP conjugated antibodies (Sigma) were diluted into the SignalBoost™ Immunoreaction Enhancer buffer (EMD Biosciences) according to the manufacturer's protocol and incubated at room temperature for 1 h. The membrane was then washed and ECL was visualized using a Universal Hood II and QuantityOne software (BioRad); the results can be seen in Fig. 4 C and F. After detection of FLAG™, the membrane was stripped and then reprobed with beta-actin-HRP conjugated antibodies (Abcam) to verify protein-loading levels. For HIF-1a detection, 50 μg of the same samples were separated on 4–12% Bis-Tris NuPAGE gels and transferred to PVDF membrane. The membrane was blocked with 3% nonfat milk for 30 min. Primary antibodies for HIF-1a were diluted 1:1,000 into blocking buffer and incubated at 4°C overnight. The membrane was then washed, followed by the addition of mouse-HRP conjugated secondary antibodies, and ECL was visualized as described above.

ACKNOWLEDGMENTS. The authors thank Sharon Gerech, Phillip Cole, and Denis Wirtz for comments on the manuscript and Gregg L. Semenza and Hirohiko Kamiyama for helpful discussions. This work was supported by National Institutes of Health Grants R01GM066972 (to M.O.) and R01CA130938 (to J.R.E.).

1. Leader B, Baca QJ, Golan DE (2008) Protein therapeutics: A summary and pharmacological classification. *Nat Rev Drug Discov* 7:21–39.
2. Verdine GL, Walensky LD (2007) The challenge of drugging undruggable targets in cancer: Lessons learned from targeting BCL-2 family members. *Clin Cancer Res* 13:7264–7270.
3. Guntas G, Ostermeier M (2004) Creation of an allosteric enzyme by domain insertion. *J Mol Biol* 336:263–273.

4. Guntas G, Mitchell SF, Ostermeier M (2004) A molecular switch created by in vitro recombination of nonhomologous genes. *Chem Biol* 11:1483–1487.
5. Guntas G, Mansell TJ, Kim JR, Ostermeier M (2005) Directed evolution of protein switches and their application to the creation of ligand-binding proteins. *Proc Natl Acad Sci USA* 102:11224–11229.
6. Wright CM, Majumdar A, Tolman JR, Ostermeier M (2010) NMR characterization of an engineered domain fusion between maltose binding protein and TEM1 beta-

- lactamase provides insight into its structure and allosteric mechanism. *Proteins* 78:1423–1430.
7. Sohka T, et al. (2009) An externally tunable bacterial band-pass filter. *Proc Natl Acad Sci USA* 106:10135–10140.
 8. Zhong H, et al. (1999) Overexpression of hypoxia-inducible factor 1 α in common human cancers and their metastases. *Cancer Res* 59:5830–5835.
 9. Mabeesh NJ, Amir S (2007) Hypoxia-inducible factor (HIF) in human tumorigenesis. *Histol Histopathol* 22:559–572.
 10. Semenza GL (2010) Defining the role of hypoxia-inducible factor 1 in cancer biology and therapeutics. *Oncogene* 29:625–634.
 11. Huang LE, Gu J, Schau M, Bunn HF (1998) Regulation of hypoxia-inducible factor 1 α is mediated by an O₂-dependent degradation domain via the ubiquitin-proteasome pathway. *Proc Natl Acad Sci USA* 95:7987–7992.
 12. Yu AY, et al. (1998) Temporal, spatial, and oxygen-regulated expression of hypoxia-inducible factor-1 in the lung. *Am J Physiol* 275:L818–826.
 13. Semenza GL (2004) Hydroxylation of HIF-1: Oxygen sensing at the molecular level. *Physiology (Bethesda)* 19:176–182.
 14. Sun HC, et al. (2007) Expression of hypoxia-inducible factor-1 α and associated proteins in pancreatic ductal adenocarcinoma and their impact on prognosis. *Int J Oncol* 30:1359–1367.
 15. Dales JP, et al. (2005) Overexpression of hypoxia-inducible factor HIF-1 α predicts early relapse in breast cancer: Retrospective study in a series of 745 patients. *Int J Cancer* 116:734–739.
 16. Akakura N, et al. (2001) Constitutive expression of hypoxia-inducible factor-1 α renders pancreatic cancer cells resistant to apoptosis induced by hypoxia and nutrient deprivation. *Cancer Res* 61:6548–6554.
 17. Khandrika L, et al. (2009) Hypoxia-associated p38 mitogen-activated protein kinase-mediated androgen receptor activation and increased HIF-1 α levels contribute to emergence of an aggressive phenotype in prostate cancer. *Oncogene* 28:1248–1260.
 18. Liao D, Corle C, Seagroves TN, Johnson RS (2007) Hypoxia-inducible factor-1 α is a key regulator of metastasis in a transgenic model of cancer initiation and progression. *Cancer Res* 67:563–572.
 19. Kung AL, Wang S, Klco JM, Kaelin WG, Livingston DM (2000) Suppression of tumor growth through disruption of hypoxia-inducible transcription. *Nat Med* 6:1335–1340.
 20. Freedman SJ, et al. (2002) Structural basis for recruitment of CBP/p300 by hypoxia-inducible factor-1 α . *Proc Natl Acad Sci USA* 99:5367–5372.
 21. Freedman SJ, et al. (2003) Structural basis for negative regulation of hypoxia-inducible factor-1 α by CITED2. *Nat Struct Biol* 10:504–512.
 22. Korkegian A, Black ME, Baker D, Stoddard BL (2005) Computational thermostabilization of an enzyme. *Science* 308:857–860.
 23. Kievit E, et al. (1999) Superiority of yeast over bacterial cytosine deaminase for enzyme/prodrug gene therapy in colon cancer xenografts. *Cancer Res* 59:1417–1421.
 24. Crystal RG, et al. (1997) Phase I study of direct administration of a replication deficient adenovirus vector containing the E. coli cytosine deaminase gene to metastatic colon carcinoma of the liver in association with the oral administration of the pro-drug 5-fluorocytosine. *Hum Gene Ther* 8:985–1001.
 25. Mahan SD, Ireton GC, Stoddard BL, Black ME (2004) Alanine-scanning mutagenesis reveals a cytosine deaminase mutant with altered substrate preference. *Biochemistry* 43:8957–8964.
 26. Mahan SD, Ireton GC, Knoeber C, Stoddard BL, Black ME (2004) Random mutagenesis and selection of *Escherichia coli* cytosine deaminase for cancer gene therapy. *Protein Eng Des Sel* 17:625–633.
 27. Dial R, Sun ZY, Freedman SJ (2003) Three conformational states of the p300 CH1 domain define its functional properties. *Biochemistry* 42:9937–9945.
 28. Dang DT, et al. (2006) Hypoxia-inducible factor-1 α promotes nonhypoxia-mediated proliferation in colon cancer cells and xenografts. *Cancer Res* 66:1684–1936.
 29. Thant AA, et al. (2008) Role of caspases in 5-FU and selenium-induced growth inhibition of colorectal cancer cells. *Anticancer Res* 28:3579–3592.
 30. Shibata T, Giaccia AJ, Brown JM (2000) Development of a hypoxia-responsive vector for tumor-specific gene therapy. *Gene Ther* 7:493–498.
 31. Epstein AC, et al. (2001) Celegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell* 107:43–54.
 32. Zhao D, Ran S, Constantinescu A, Hahn EW, Mason RP (2003) Tumor oxygen dynamics: Correlation of in vivo MRI with histological findings. *Neoplasia* 5:308–318.
 33. Greco O, Dachs GU (2001) Gene directed enzyme/prodrug therapy of cancer: Historical appraisal and future perspectives. *J Cell Physiol* 187:22–36.
 34. Russell PJ, Khatri A (2006) Novel gene-directed enzyme prodrug therapies against prostate cancer. *Expert Opin Investig Drugs* 15:947–961.
 35. Schepelmann S, Springer CJ (2006) Viral vectors for gene-directed enzyme prodrug therapy. *Curr Gene Ther* 6:647–670.
 36. Marignol L, et al. (2009) Hypoxia response element-driven cytosine deaminase/5-fluorocytosine gene therapy system: A highly effective approach to overcome the dynamics of tumour hypoxia and enhance the radiosensitivity of prostate cancer cells in vitro. *J Gene Med* 11:169–179.
 37. Yan M, et al. (2010) A novel intracellular protein delivery platform based on single-protein nanocapsules. *Nat Nanotechnol* 5:48–53.
 38. Cronican JJ, et al. (2010) Potent delivery of functional proteins into Mammalian cells in vitro and in vivo using a supercharged protein. *ACS Chem Biol* 5:747–752.

Supporting Information

Wright et al. 10.1073/pnas.1102803108

SI Text

General. All chemicals were purchased from Sigma-Aldrich unless otherwise noted. All cell culture materials were purchased from Gibco (Invitrogen), and cell lines were purchased from American Type Culture Collection (ATCC). For tissue culture, RKO cells were grown in MEM media, containing 10% FBS and 1% antibiotic and antimycotic (ABAM) (Gibco). GIA39 *Escherichia coli* cells (Coli Genetic Stock Center# 5594) were purchased from the *E. coli* genetic stock center. Antibodies for Western blots (yeast cytosine deaminase: Ab35251 and beta-Actin: Ab20272) were purchased from Abcam. The FLAG™ M2 (F3165) and FLAG™-HRP conjugated (A8592) antibodies were purchased from Sigma, and the HIF-1 α antibody (BDB 610958) was purchased from BD Biosciences (Fisher Scientific). All antibodies were used according to the manufacturers' instructions.

Plasmids and Genes. Yeast cytosine deaminase was cloned from yeast genomic DNA using the primers: 5'-ttataagatcctatggtgacagggggaatggcaag and 5'-ttataaactagctactaccataatcttcaaccaatc with the NcoI and SpeI restriction sites underlined, and inserted into the pDIM-C8 plasmid, which contains a *tac* promoter and confers chloramphenicol resistance (1–3). Three thermostabilizing mutations (A23L, V108I, and I140L) (4) were incorporated using QuikChange (Stratagene) following the manufacturer's protocol. This thermostable version of the gene is referred to as "yCD" throughout *SI Text* for simplicity and the plasmid is referred to as pDIM-yCD. The DNA encoding the C-TAD domain of HIF-1 α (amino acids 786–826) and the CH1 domain of the human p300 protein (amino acids 334–420) were ordered from IDT with *E. coli* codon optimization. The *gstHIF-1 α* plasmid (for *gstHIF-1 α* expression) was constructed using the GST fusion tag from the pGEX-6P-1 plasmid (GE Healthcare) and the *araC* gene and the arabinose promoter from the pBAD plasmid (Invitrogen) and conferred ampicillin resistance. The DNA encoding C-TAD domain of HIF-1 α was fused to the DNA encoding GST as described by Freedman et al. (5) to form "gstHIF-1 α ." The *gstHIF-1 α* fusion is inducible by the addition of arabinose to the media. DNA encoding human codon optimized yCD, Haps3, and Haps59 were purchased from GenScript and cloned into the pcDNA 3.1(+) plasmid with neomycin resistance (Invitrogen).

Library Creation. Purified pDIM-yCD plasmid was digested with dilute concentrations of DNaseI as described (1–3). Singly cut plasmids were isolated using gel electrophoresis and purified using Qiagen's gel extraction kit following the manufacturer's instructions. Isolated singly cut plasmids were repaired and blunted using T4 DNA polymerase (0.5 U/mg) and T4 DNA Ligase (150 U/mg) both from New England Biolabs. The repaired, linear DNA was isolated using gel electrophoresis and used in a ligation reaction with DNA encoding CH1 domain inserts. Three types of CH1 domain inserts were prepared. Two inserts are described as direct inserts and used as an unaltered CH1 gene with appended DNA that encoded peptide linkers. The appended linkers encoded a glycine on the N terminus and either a GGS peptide linker ("3-mer") or a GGGGS ("5-mer") peptide linker on the C terminus. The third CH1 domain insert was prepared using the circular permutation method shown in Fig. S1. The gene encoding the CH1 domain had a piece of DNA appended coding for a (GSGGG)₃ linker that joined together the N and C termini of the CH1 domain. The appended *CHI* gene was cyclized and digested with a nonspecific nuclease to create random circular permutations of this gene. To accomplish this, the CH1 DNA

was excised from its plasmid using BamHI sites located within the linker region and the ends of the gene ligated together under dilute DNA concentrations to favor intramolecular ligation over intermolecular ligation. A standard cyclization reaction diluted 5 μ g of the excised *CHI* gene into 500–600 μ L of 1 \times T4 DNA ligase buffer (New England Biolabs) to a DNA concentration of approximately 8–10 ng/ μ L. The dilute *CHI* genes were cyclized using T4 DNA ligase (20 units/ μ L) at room temperature (RT) for 1 h. Cyclized *CHI* genes were isolated using gel electrophoresis. S1 nuclease was added (2.5 U/ μ g) to purified, cyclized DNA to make a variety of single double-stranded breaks within the CH1 DNA. The singly cut DNA was isolated using gel electrophoresis and repaired as described above in the digestion of pDIM-yCD. To create the yCD-CH1 hybrid library, a ligation reaction was performed with the randomly linearized pDIM-yCD plasmid DNA and a 5-fold higher amount (molar basis) of a 1:1:1 mixture of the three types of CH1 inserts. A typical ligation reaction mixture included 500 ng of plasmid DNA, approximately 200 ng of CH1 inserts, and 5% PEG in 1 \times T4 ligase buffer. Ligated plasmids were electroporated into DH5 α *E. coli* cells and the transformation mixture plated on LB agar containing 50 μ g/mL chloramphenicol in a 24.5 \times 24.5 cm Bio-Assay dish (Nunc, Thermo Fisher Scientific). The number of transformants was 9.6×10^6 of which approximately 25% contained insert DNA, as estimated by gel electrophoresis of plasmid DNA isolated from the library.

5-fluorocytosine (5FC) Sensitivity Assays with GIA39 Cells in Liquid Media. GIA39 cells harboring a pDIM-C8 plasmid for expression of maltose-binding protein (MBP) (negative control), yCD (positive control), Haps3, or Haps59 were cotransformed with either the *gstHIF-1 α* plasmid or an analogous negative control plasmid encoding only GST. These eight GIA39 cell lines were used for 5FC sensitivity experiments. In 96-well format, 1-mL cultures of the eight GIA39 cell lines described above were grown in minimal media at 37 $^{\circ}$ C with shaking. Minimal media consisted of 1 \times nitrogen base, 1 \times yeast synthetic dropout without uracil, 2% glucose, 5FC (varied from 0 to 1 mM), 1.5 μ g/mL uracil, 1 mM IPTG, 100 μ g/mL of ampicillin, 50 μ g/mL chloramphenicol, and with or without 0.15% arabinose (to express *gstHIF-1 α* and GST). The bacteria cells were cultured for 40–48 h at 37 $^{\circ}$ C after which the OD at 600 nm was measured using a SPECTRAMax Plus 96-well plate reader (Molecular Devices) and analyzed using Softmax Pro software (Molecular Devices). The relative cell densities were calculated using the equation $(A)_{\text{sample}} / (A)_{\text{control}} \times 100$, in which $(A)_{\text{sample}}$ is the absorbance of a sample well containing 5FC and $(A)_{\text{control}}$ is the absorbance of the control well lacking 5FC. Experiments performed with GIA39 cell lines expressing yCD and Haps proteins were repeated on three separate days in duplicate ($N = 6$). Control experiments performed with GIA39 cells expressing MBP were repeated twice in duplicate on separate days ($N = 4$). Error was calculated by standard deviation from the mean.

Western Blot for Accumulated yCD and Haps59 in GIA39 Cells. The membrane containing transferred cell lysates was blocked with 3% nonfat milk for 30 min. To observe the expression of Haps59 and yCD, primary antibodies for yCD were diluted into blocking buffer according to the manufacturer's protocol and incubated at RT for 1 h. The membrane was then washed, followed by the addition of sheep-HRP conjugated secondary antibodies (Bethyl Laboratories) using the Snap ID protein detection system

(Millipore) following the manufacturer's protocol. ECL was visualized using a Universal Hood II and QuantityOne software (BioRad).

Purification of yCD, Haps3, and Haps59. After failed attempts to purify the protein switches using a His tag and Ni²⁺ affinity, the genes encoding yCD, Haps3, and Haps59 were cloned into the pGEX-6P-1 plasmid for fusion to the affinity tag GST for purification (GE Healthcare). Separate 500-mL cultures of DH5 α bacteria cells harboring each plasmid were grown in the presence of ampicillin at 37 °C until an OD of 0.4–0.6, at which point protein was induced by the addition of 1 mM IPTG for 3 h. Harvested cells were resuspended in lysis buffer (50 mM Tris buffer, pH 7.5, containing 150 mM NaCl, 0.1 mM EDTA) containing protease inhibitor cocktail (Sigma). Cells were lysed using a French Press and the insoluble material pelleted by centrifugation at 15,000 \times g for 45 min at 4 °C. Cleared lysate was loaded onto a 5 mL GSTrap FF column using an AKTA purifier FPLC (GE Healthcare). The GSTrap column was washed with four column volumes of lysis buffer to clear unbound proteins. The fusion protein was eluted using 50 mM Tris pH 8.0, containing 10 mM reduced glutathione. Alternatively, an in-column cleavage was performed on bound GST fusions using PreScission Protease (GE Healthcare) at 80 U/mg. Storage buffer (25 mM Tris buffer, pH 7.5, containing 50 mM NaCl) was used to elute the cleaved protein from the GSTrap column. Purity was judged to be >95% using SDS-PAGE and protein concentrations were determined using A₂₈₀ and the calculated extinction coefficient for each enzyme. Purification of the gstHIF-1a fusion was performed exactly the same as above, except the fusion was not cleaved with PreScission Protease. GST-HIF-1a was eluted from the GSTrap column in elution buffer (50 mM Tris buffer, pH 8.0, containing 10 mM reduced glutathione) and dialyzed into storage buffer (25 mM Tris buffer, pH 7.5, containing 50 mM NaCl and 10% glycerol) then frozen at –20 °C.

5FC Activity Assay. The cytosine deaminase activity assays for yCD and the Haps proteins were performed on purified samples that were never frozen and within two weeks of purification. The 5FC activity assay was based on similar assays reported by Mahan and coworkers (6, 7). To determine the cytosine deaminase activity, yCD or Haps3 or Haps59 were added to a final concentration of 100 nM in a reaction tube containing 50 mM Tris buffer, pH 7.5. The mixture was incubated at 37 °C for 5 min and then 5FC was added at various concentrations (50 μ M–1 mM). For experiments in the presence of 2.5 μ M purified gstHIF-1a, both the protein switch and gstHIF-1a were incubated at 37 °C for 5 min before the addition of 5FC. Aliquots (50 μ L) were removed from the reaction mixture at various time points (10, 20, 30, 60, 120, 180, and 240 s) and mixed with 0.1 N HCl. Measurements at 290 nm for 5FC and 255 nm for 5-fluorouracil (5FU) were recorded and inserted in the following equations to determine the concentrations of 5FC and 5FU: [5FC] = 0.119A₂₉₀ – 0.025A₂₅₅ and [5FU] = 0.105A₂₅₅ – 0.049A₂₉₀ as described elsewhere (6). Alternatively, CD activity was monitored using the decrease in absorbance at 235 nm, where 5FC absorbs. The production of 5FU was calculated using the decrease in absorbance at 235 nm over time and the extinction coefficient of 5FC. Kinetic parameters were determined using double reciprocal plots. Our calculated catalytic efficiency for triple mutant yCD at 37 °C was 4.98 \times 10⁴ E^{–1} M^{–1} s^{–1}, which is consistent with other reports (8, 9). Although both switches exhibited cytosine deaminase activity and produced absorbance spectra changes consistent with the production of 5FU, the measured catalytic activities were inconsistent from purification to purification and decreased markedly over the period of a day. In general, the two protein switches' catalytic activity in the presence of gstHIF-1a ranged from 30–85% of that measured for yCD. The rate of deamination

for both protein switches was always the same or higher (up to severalfold) when measured in the presence of purified gstHIF-1a. Analogous experiments with GST tags attached to the protein switches resulted in similar outcomes.

Copurification Experiments in *E. coli*. Copurification experiments were performed using GIA39 cells harboring the gstHIF-1a plasmid and either pDIM-Haps3 or pDIM-Haps59 (i.e., versions of the switches lacking the GST tag). A 500-mL culture was incubated at 37 °C until the OD reached 0.4–0.6. At this point, arabinose (0.15%) was added to the culture to express gstHIF-1a, and IPTG (1 mM) was added to express the protein switch. After 3 h at 37 °C, the cells were lysed and passed over a GSTrap column using an AKTA purifier FPLC (GE Healthcare). Unbound proteins were eluted by the passage of four column volumes of lysis buffer over the GSTrap column. Protein complexes were eluted using 50 mM Tris buffer, pH 8.0, containing 10 mM reduced glutathione and separated using SDS-PAGE. Because the switch proteins and gstHIF-1a have similar molecular weights (28 kDa and 31 kDa, respectively), an aliquot was treated with 20 U of PreScission Protease (GE Healthcare) to cleave off the GST tag of gstHIF-1a. The sample was subsequently passed over the GSTrap column again to remove the majority of free GST from the sample (Fig. S4).

5FC Experiments in Parental RKO and MCF7 Cells. RKO and MCF7 cells were acquired from the ATCC and genotyped before use. For controls, parental RKO cells were used to seed 96-well plates, 1,500 cells per well, in 100 μ L of MEM media supplemented with 10% fetal bovine serum and 1% ABAM (all from Gibco). Parental MCF7 cells were used to seed 96-well plates, 3,500 cells per well, in 100 μ L of DMEM media supplemented with 10% fetal bovine serum and 1% ABAM (all from Gibco). After 24 h, the media was changed with 100 μ L of MEM (RKO) or DMEM (MCF7) media containing either 5FC (0–20 mM) or 5FU (0–2 mM). Cells were incubated for 3 to 6 d and then lysed. The cells were lysed by rinsing twice with 100 μ L of PBS and then incubated in 100 μ L of 0.1% SDS in water. The cells were incubated at 37 °C for 2 h in the SDS solution for complete lysis. To detect dsDNA, SYBR® green (Invitrogen) was added to a final concentration of 0.075% to lysed cells and the fluorescence for each well was recorded at 520 nm after excitation at 485 nm. The percent survival was calculated using the equation (A)_{sample} / (A)_{control} \times 100, where (A)_{sample} is the fluorescence of the sample wells with 5FC and (A)_{control} is the absorbance of the control well lacking 5FC and 5FU. Results shown in Fig. S5 A and D are the mean of experiments performed on three different days. Error bars are the standard deviation from the mean.

Creation of RKO and MCF7 Stable Cell Lines. Genes encoding human codon optimized yCD, Haps3, and Haps59 were cloned into the pcDNA 3.1 (+) plasmid (Invitrogen) under the control of a CMV promoter and containing neomycin resistance. Transfections were performed using Lipofectamine 2000 following the manufacturer's protocol in 6-well plates. A bulk selection using G418 (Geneticin) (Gibco) at a concentration of 0.8 mg/mL was performed 24 h after the transfection. Cells that survived the bulk selection were separated into one cell per well in 96-well plates for further selection in the presence of Geneticin. Single cells that grew into colonies in the presence of Geneticin were examined for their 5FC sensitivity by the addition 800 μ M 5FC for yCD clones and 75 μ M Co²⁺ and 800 μ M 5FC for protein switch clones. Clones of RKO-yCD, RKO-Haps3, and RKO-Haps59 that were sensitive to 5FC containing media (i.e., functionally confirmed) were also confirmed genetically. Genomic DNA isolated from these clones was probed by PCR with primers that annealed outside of the multiple cloning site of the pcDNA3.1 (+) plasmid. Sequencing of the correct size PCR product con-

firmed the successful creation of the desired RKO cell lines. For the MCF7 cells, a gene encoding GFP was linked to the genes encoding yCD and Haps59 via a T2A peptide linker. The T2A linker is self-cleaving, yields two separate proteins (GFP + yCD or GFP + Haps59), with a short C-terminal peptide on GFP and only a single amino acid on the N terminus of the yCD or Haps59 proteins (10). The tandem genes (GFP-T2A-yCD or GFP-T2A-Haps59) were cloned into the pcDNA 3.1 (+) plasmid as well as the gene encoding GFP as an empty vector control. The same transfection protocol used for RKO cells was used for MCF7 cells. A bulk selection using G418 at a concentration of 0.6 mg/mL was performed 24 h after the transfection. Cells that survived the bulk selection were separated into one cell per well in 96-well plates using GFP as a reporter and further selection in the presence of Geneticin. Single cells that grew into colonies were characterized functionally and genetically as described above for RKO cells. Sequencing of MCF7-yCD and MCF7-Haps59 clones confirmed the successful creation of the stable cell lines.

5FU Production in RKO Cells. We confirmed that RKO cells expressing Haps59 generate 5FU and that HIF-1 α accumulation increased the amount of 5FU produced. We used an activity assay similar to that used by others to demonstrate the conversion of 5FC to 5FU by yCD in human cell lysates (11, 12). RKO cells expressing an empty vector, yCD or Haps59, were grown to confluency in four T-25 flasks in MEM media, with two flasks growing RKO-Haps59 cells. One RKO-Haps59 flask was exposed to 100 μ M Co²⁺ for >24 h. Cells in all four flasks were collected and resuspended in 400 μ L of PBS, pH 7.4 and then lysed using two freeze-thaw cycles. RKO-Haps59 cells that had been exposed to 100 μ M Co²⁺ were resuspended in PBS containing 100 μ M Co²⁺. After the lysis, cell debris was pelleted by centrifugation of samples for 5 min (20,000 \times g, 4°C) and protein concentrations were determined using a DC protein assay (BioRad). Cleared lysates (to a final concentration of 0.2 mg/mL) were added to a PBS solution containing 50 μ M 5FC and incubated at 37°C for 16–18 h.

The Analytical Pharmacology Core at Johns Hopkins University School of Medicine performed analysis of these samples. 5FU concentrations were quantified using an analytic assay based on reversed-phase HPLC with tandem mass spectrometric detection (13). The calibration curve and quality control samples were prepared in PBS, pH 7.4. Samples were quantified over the assay range of 50 to 5,000 ng/mL. Results can be seen in Fig. S6.

Western Blots for the Effect of Co²⁺ on HIF-1 α (Fig. 4 A and D) and Haps59 (Fig. S7A) Accumulation. Mammalian cell whole cell lysates were prepared after incubation with or without 150 μ M CoCl₂ for 4 h. After exposure, cell cultures were incubated at 4°C for 30 min and then treated with radioimmunoprecipitation assay (RIPA) buffer (Sigma) containing protease inhibitor cocktail (Sigma). Cells were scrapped from the flask and placed in a 1.5-mL tube and incubated on ice for 30 min. Cell debris was removed by centrifugation (20,000 \times g for 30 min) and the supernatant was transferred to a fresh 1.5-mL tube. Protein concentrations of the whole cell lysates were determined using the DC protein assay (BioRad). A total of 50 μ g of each lysate were separated on 4–12% Bis-Tris NuPAGE gels (Invitrogen) and then transferred to PVDF membrane (BioRad). The membrane was blocked with 3% nonfat milk for 30 min.

Primary antibodies for HIF-1 α were diluted into blocking buffer according to the manufacturer's protocol and incubated at 4°C overnight. The membrane was then washed, followed by the addition of mouse-HRP conjugated secondary antibodies (BioRad) using the Snap ID protein detection system (Millipore) following the manufacturer's protocol. ECL was visualized using a Universal Hood II and QuantityOne software (BioRad). After

detection of HIF-1 α , the membrane was stripped and then re-probed with beta-actin-HRP conjugated antibodies (Abcam) to verify protein-loading levels.

To observe the expression of Haps59 and yCD in these same RKO lysates, primary antibodies for yCD were diluted into blocking buffer according to the manufacturer's protocol and incubated at RT for 1 h. The membrane was then washed, followed by the addition of sheep-HRP conjugated secondary antibodies (Bethyl Laboratories) using the Snap ID protein detection system (Millipore) following the manufacturer's protocol. ECL was visualized using a Universal Hood II and QuantityOne software (BioRad). After initial protein screen, the membrane was stripped and then re-probed with beta-actin-HRP conjugated antibodies (Abcam) to verify protein-loading levels (Fig. S7).

Coimmunoprecipitation (coIP) Experiments. yCD antibodies were shown to cross-react with many mammalian proteins (Fig. S7A). To circumvent this problem, a FLAG tag (Sigma-Aldrich) was added to the N terminus of Haps59 for coIP experiments. CoIP experiments were performed on lysates of RKO cells transiently expressing an empty vector, yCD or FLAG-Haps59 (Haps59 with an N-terminal FLAG tag) after overnight incubation with 150 μ M Co²⁺. These cells were lysed using RIPA buffer (Sigma) containing protease inhibitor cocktail (Sigma), and then primary antibodies (anti-HIF-1 α or FLAG-M2) were added at a concentration of 1.5 μ g antibody per 100 μ L cell lysates. The mixture was incubated at 4°C overnight with light mixing. After incubation with the primary antibody, the coIP complexes were purified using PureProteome Protein G Magnetic Beads (Millipore) following the manufacturer's protocol. Protein samples were eluted in 70 μ L of denaturing sample buffer. An additional negative control was performed using the TATA-binding protein (TBP) antibody, which should have no affinity for HIF-1 α or FLAG-Haps59. The TBP antibody (Abcam) was added to the FLAG-Haps59 lysate, and this sample was treated exactly the same as the other samples.

For Western blot analysis, 20 μ L of each eluted complex sample was separated on 4–12% Bis-Tris NuPAGE gels (Invitrogen) and then transferred to PVDF membrane (BioRad). The membrane was blocked with 3% nonfat milk for 30 min. Primary antibodies against yCD or FLAG-M2 were diluted according to the manufacturer's protocol and incubated at room temperature for 1 h. Primary antibodies for HIF-1 α were diluted into blocking buffer according to the manufacturer's protocol and incubated at 4°C overnight. The membrane was then washed followed by the addition of mouse-HRP conjugated secondary antibody (FLAG-M2) or sheep-HRP conjugated secondary (yCD) using the Snap ID protein detection system (Millipore) following the manufacturer's protocol. ECL was visualized using a Universal Hood II and QuantityOne software (BioRad). To confirm the identity of the FLAG-Haps59 band, the membrane initially probed with yCD antibodies was stripped and then re-probed with anti-FLAG antibodies as described above.

Amino Acid Sequence for Haps Proteins. *Haps3*. MVTGGMASGD-PEKRKLIQQQLVLLHHAHKCORREQANGEVROCNLPH-CRTMKNVNLNMTHCQSGKSCQVAHCASSRQIISHWKN-CTRHDCPVCLPLKNAGGSKWDQKGMDIAYEEALLGYK-EGGVPIGGCLINNKDGSVLGRGHNMRFQGSATLHGE-ISTLENCGRLEGKVKYKDTTLYTTLSPCDMCTGAIIMYGI-PRCVIGENVNFKSKGEKYLQTRGHEVVVDDDERCKKL-MKQFIDERPDQWDFEDIGE

Haps59. MVTGGMASDPEKRKLIQQQLVLLHHAHKCORREQANGEVROCNLPHCRTMKNVNLNMTHCQSGKSCQVAHCASSRQIISHWKNCTRHDCPVCLPLKNAGGWDQKGM-DIAYEEALLGYKEGGVPIGGCLINNKDGSVLGRGHNM-RFQKGSATLHGEISTLENCGRLEGKVKYKDTTLYTTLSPCD-

MCTGAIIMYGIPRCVIGENVNFKSKGKYLQTR-
GHEVVVDDDERCKKLMKQFIDERPQDWFEDIGE

SI Acknowledgments This research was supported by the Analytical Pharmacology Core of the Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins (National Institutes of Health Grants P30 CA006973 and S10 RR026824). These experiments were made possible by Grant UL1 RR 025005 from the National

Center for Research Resources (NCRR), a component of the NIH, and NIH Roadmap for Medical Research—P30 CA006973 (funding from the NCI to W.G.N.); S10 RR026824 (funding from the NCRR to M.A.R.); UL1 RR025005 (funding from the NCRR to D.E.F.). The paper's contents are solely the responsibility of the authors and do not necessarily represent the official view of NCRR or NIH.

- Guntas G, Mitchell SF, Ostermeier M (2004) A molecular switch created by in vitro recombination of nonhomologous genes. *Chem Biol* 11:1483–1487.
- Guntas G, Mansell TJ, Kim JR, Ostermeier M (2005) Directed evolution of protein switches and their application to the creation of ligand-binding proteins. *Proc Natl Acad Sci USA* 102:11224–11229.
- Wright CM, Majumdar A, Tolman JR, Ostermeier M (2010) NMR characterization of an engineered domain fusion between maltose binding protein and TEM1 beta-lactamase provides insight into its structure and allosteric mechanism. *Proteins* 78:1423–1430.
- Korkegian A, Black ME, Baker D, Stoddard BL (2005). Computational thermostabilization of an enzyme. *Science* 308:857–860.
- Freedman SJ, et al. (2002) Structural basis for recruitment of CBP/p300 by hypoxia-inducible factor-1 alpha. *Proc Natl Acad Sci USA* 99:5367–5372.
- Mahan SD, Ireton GC, Stoddard BL, Black ME (2004) Alanine-scanning mutagenesis reveals a cytosine deaminase mutant with altered substrate preference. *Biochemistry* 43:8957–8964.
- Mahan SD, Ireton GC, Knoeber C, Stoddard BL, Black ME (2004) Random mutagenesis and selection of *Escherichia coli* cytosine deaminase for cancer gene therapy. *Protein Eng Des Sel* 17:625–633.
- Kievit E, et al. (1999) Superiority of yeast over bacterial cytosine deaminase for enzyme/prodrug gene therapy in colon cancer xenografts. *Cancer Res* 59:1417–1421.
- Stolworthy TS, et al. (2008). Yeast cytosine deaminase mutants with increased thermostability impart sensitivity to 5-fluorocytosine. *J Mol Biol* 377:854–869.
- de Felipe P, Martín V, Cortés ML, Ryan M, Izquierdo M. (1999) Use of the 2A sequence from foot-and-mouth disease virus in the generation of retroviral vectors for gene therapy. *Gene Ther* 6:198–208.
- Khatri A, et al. (2006) Combination of cytosine deaminase with uracil phosphoribosyl transferase leads to local and distant bystander effects against RM1 prostate cancer in mice. *J Gene Med* 8:1086–1096.
- Gopinath P, Ghosh SS (2008) Implication of functional activity for determining therapeutic efficacy of suicide genes in vitro. *Biotechnol Lett* 30:1913–1921.
- Xu Y, Grem JL (2003) Liquid chromatography-mass spectrometry method for the analysis of the anti-cancer agent capecitabine and its nucleoside metabolites in human plasma. *J Chromatogr B Analyt Technol Biomed Life Sci* 783:273–285.

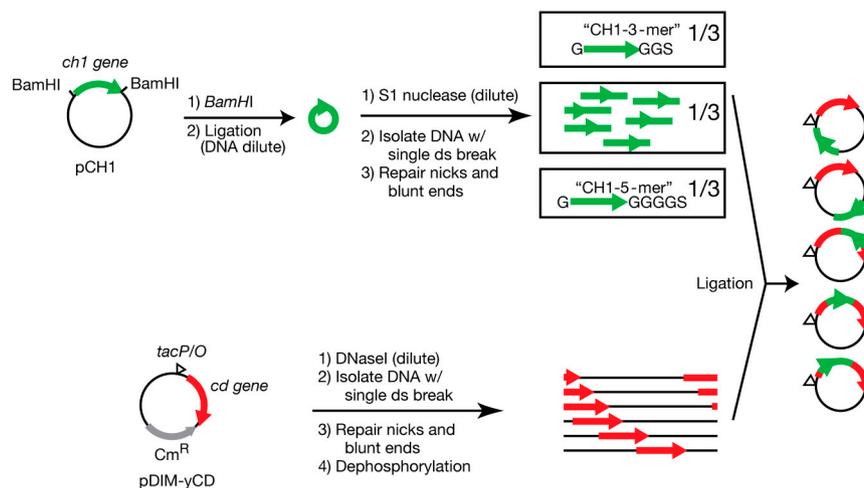


Fig. S1. Schematic showing domain insertion method used to create the yCD-CH1 hybrid library. The CH1 domain inserts (cpCH1, CH1-3-mer, and CH1-5-mer) were mixed in an equimolar ratio before they were used in the ligation mixture with singly cut pDIM-yCD plasmids.

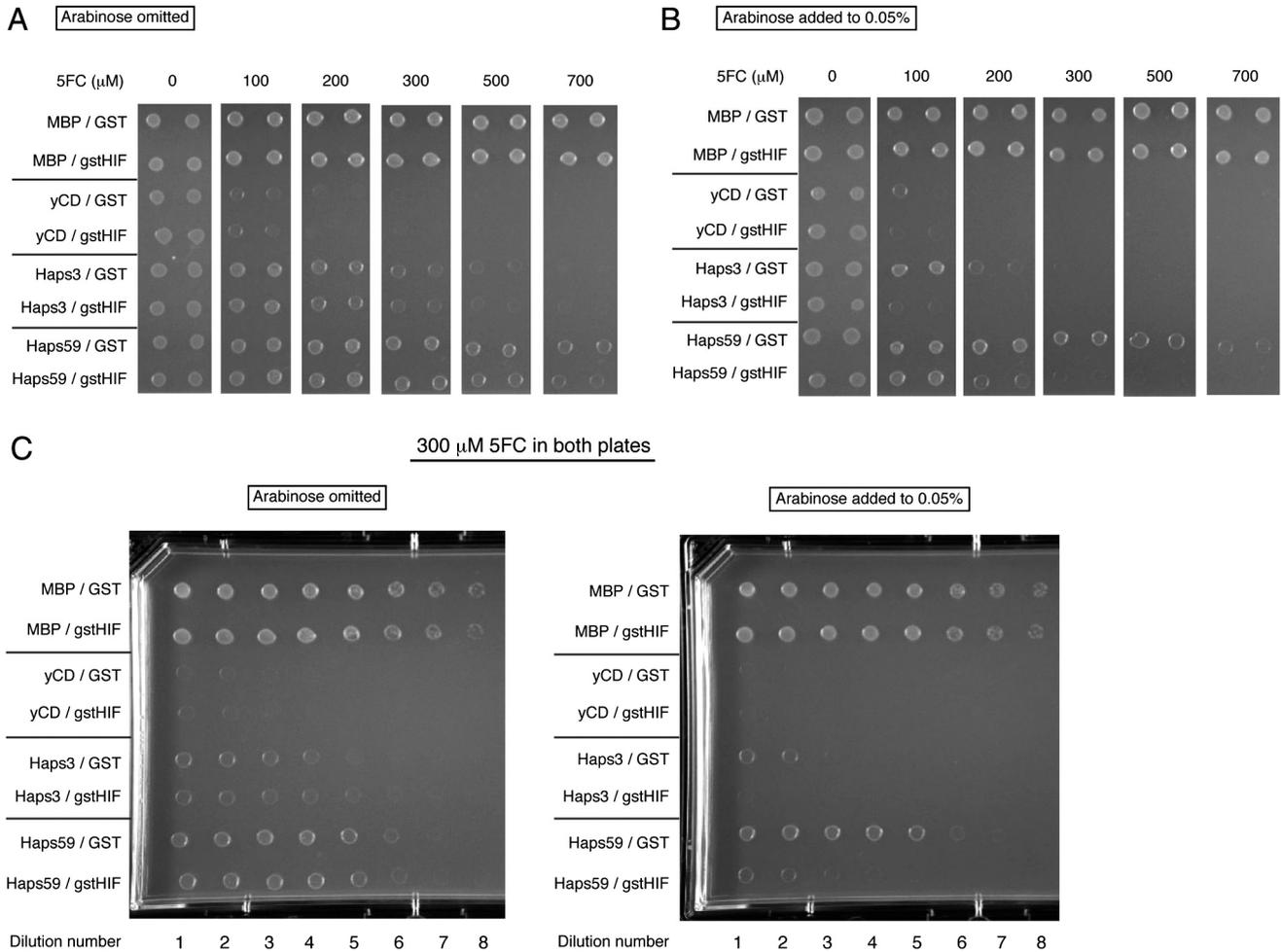


Fig. S2. Growth of GIA39 cells expressing MBP, yCD, or Haps59 and either GST or gstHIF-1a on minimal media agar plates as a function of 5FC concentration. *A* is in the absence of arabinose and *B* is in the presence of 0.05% arabinose to express GST or gstHIF-1a. (*C*) Example "full" plates showing all dilutions that exhibited growth. Both plates contained 300 μ M 5FC and either omitted arabinose (*Left*) or contained 0.05% arabinose (*Right*). Threefold less arabinose was used in these experiments compared to the liquid media experiments to lessen the burden of high expression from the arabinose promoter. Dilutions 3 and 4 from all plates were used to create Fig. 3*B* and *A* and *B* of this figure. The control lines (MBP and yCD) were unaffected by the expression of either GST or gstHIF-1a. 5FC toxicity in cells expressing Haps3 increased more with gstHIF-1a than GST, whereas only gstHIF-1a increased the 5FC toxicity for cells expressing Haps59.

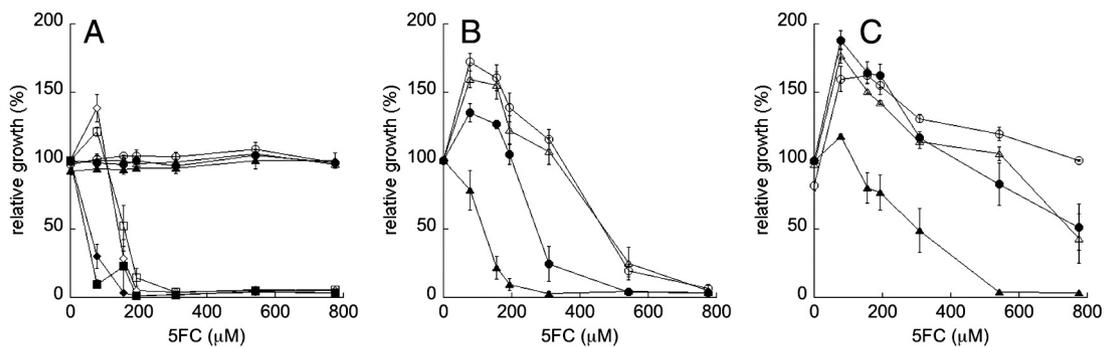


Fig. S3. Growth of GIA39 cells expressing MBP, yCD, or Haps59 and either GST or gstHIF-1a in minimal liquid media as a function of 5FC concentration. Media either omitted (open symbols) or contained (solid symbols) 0.15% arabinose to compare the effects of coexpressing GST or gstHIF-1a. (*A*) GIA39 cells expressing control proteins (MBP + GST, circles; MBP + gstHIF-1a, triangles; yCD + GST, diamonds; yCD + gstHIF-1a, squares). GIA39 cells expressing (*B*) Haps3 or (*C*) Haps59 with either GST (circles) or gstHIF-1a (triangles) demonstrate that gstHIF-1a increases the 5FC toxicity of cells expressing these switches. For all graphs, error bars, SD ($N = 6$). Experiments with yCD (*A*) revealed that both GST and gstHIF-1a induction by arabinose caused a small but equal increase in 5FC toxicity. This effect likely arises from the added burden of high expression from the arabinose promoter to cells that are coping with near lethal levels of 5FC. In contrast, gstHIF-1a expression increased the 5FC toxicity of cells expressing Haps3 (*B*) and Haps59 (*C*) to a much greater extent than did GST expression. We attribute the increase in cell density observed at sublethal 5FC concentrations for cells expressing yCD, Haps3, or Haps59 to a stress response that allows growth to higher densities in the minimal media because the increase is not observed when MBP is expressed.

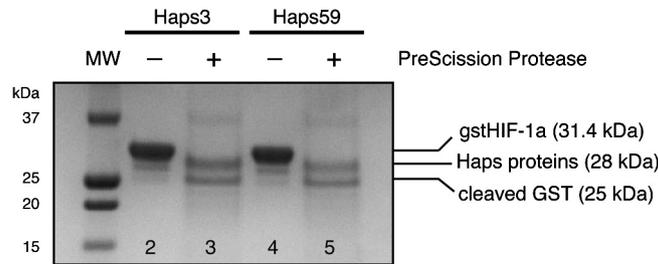


Fig. S4. Copurification of Haps3 and Haps59 with *gstHIF-1a* as visualized with Coomassie blue stain. GIA39 cells harboring the *gstHIF-1a* plasmid were used to express *gstHIF-1a* and either Haps3 or Haps59. The lysates were passed over a GSTrap column. Lanes 2 and 4 show the eluted proteins purified from cells expressing *gstHIF-1a* and either Haps3 (lane 2) or Haps59 (lane 4). In both of these lanes a heavy band corresponding to the molecular weight of *gstHIF-1a* (which is present in excess) and lighter, faster migrating band corresponding to the Haps protein was observed. To remove the majority of the *gstHIF-1a*, an aliquot of each purified protein mixture was treated with PreScission protease to cleave the GST tag off of *gstHIF-1a*, then passed over a GSTrap column to remove most of the cleaved GST protein. The flow through from this experiment was loaded in lane 3 (Haps3) and lane 5 (Haps59).

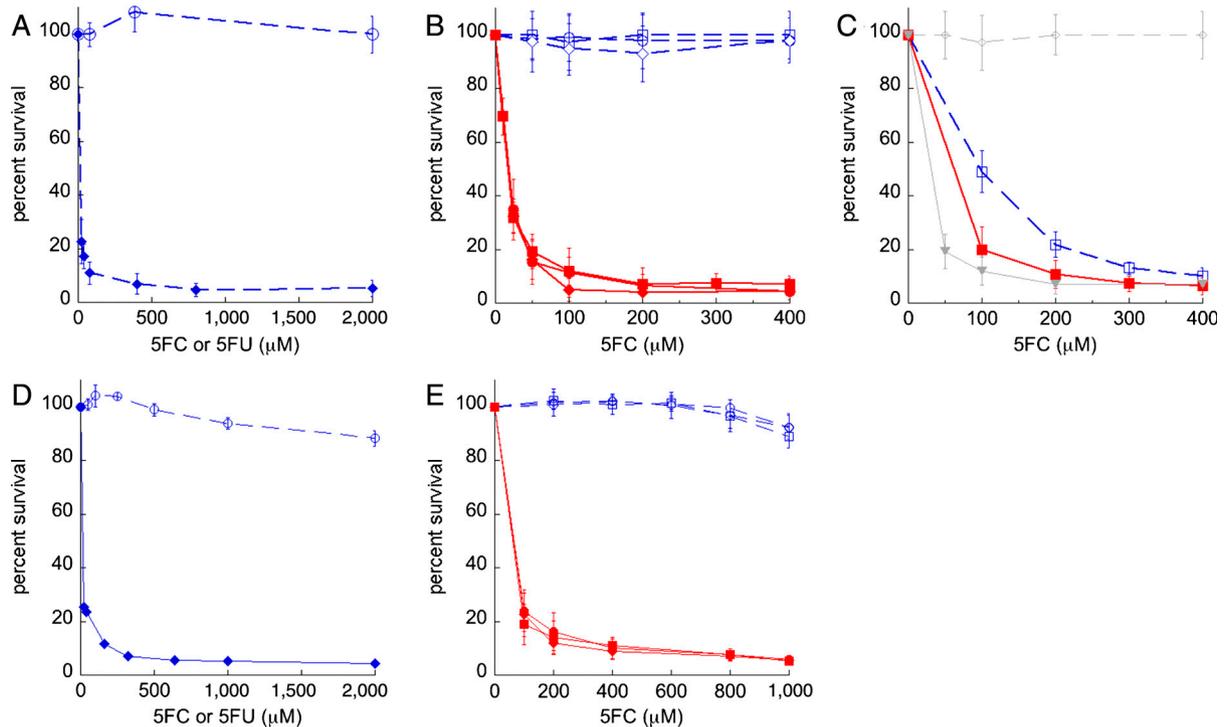


Fig. S5. 5FC and 5FU sensitivity experiments in RKO and MCF7 cells. (A) Parental RKO cells treated with 5FU (solid blue diamonds; \blacklozenge) or 5FC (blue circles; \circ). Each point represents the mean of six experiments performed on three separate days. Error bars, SD ($N = 6$). (B) Positive and negative control experiments in stable RKO cell lines. The 5FC sensitivity of RKO cells stably expressing *yCD* (red, solid symbols) or an empty vector control (blue, open symbols) that were incubated in normoxic conditions in the absence (squares) or presence (circles) of $100 \mu\text{M Co}^{2+}$ or in $1\% \text{O}_2$ (diamonds). Error bars, SD ($N = 6$). (C) Haps3-expressing RKO cells incubated in normoxic conditions show a small increase in sensitivity to 5FC in the presence of $100 \mu\text{M Co}^{2+}$ (solid red squares, \blacksquare) over cells in the absence Co^{2+} (open blue squares, \square). Error bars, SD ($N = 4$). Controls from RKO cells expressing *yCD* or EV (B) are shown in gray for comparison. (D) Parental MCF7 cells treated with 5FU (solid blue diamonds; \blacklozenge) or 5FC (blue circles; \circ). Each point represents the mean of nine experiments performed on three separate days. Error bars, SD ($N = 9$). (E) Positive and negative control experiments in stable MCF7 cell lines. The 5FC sensitivity of MCF7 cells stably expressing *yCD* (red, solid symbols) or an empty vector control (blue, open symbols) that were incubated in normoxic conditions in the absence (squares) or presence (circles) of $100 \mu\text{M Co}^{2+}$ or in $1\% \text{O}_2$ (diamonds). Error bars, SD ($N = 9$).

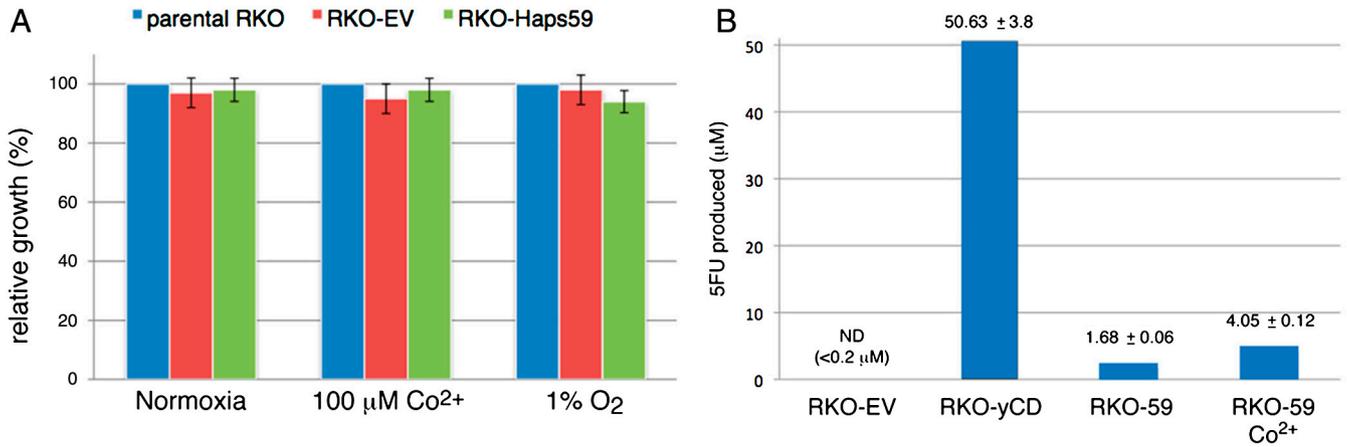


Fig. 56. (A) The relative growth of various RKO cells under experimental conditions. The presence of Co^{2+} or 1% O_2 did not have an effect on the growth of RKO cells expressing Haps59 or RKO cells expressing an EV when compared to the parental RKO cells. (B) 5FU production by RKO cells. Lysates from RKO cells expressing an empty vector (RKO-EV) or Haps59 (RKO-59; either with or without Co^{2+}) were added to a solution containing 5FC. 5FU production was measured by using an analytic assay based on reversed-phase HPLC with tandem mass spectrometric detection (9). Expression of yCD instead of Haps59 resulted in 10-fold more 5FU production, which corresponds to complete conversion 5FC. We speculate that the instability of Haps59, especially after cell lysis, could contribute to the large difference in production of 5FU compared to yCD lysates.

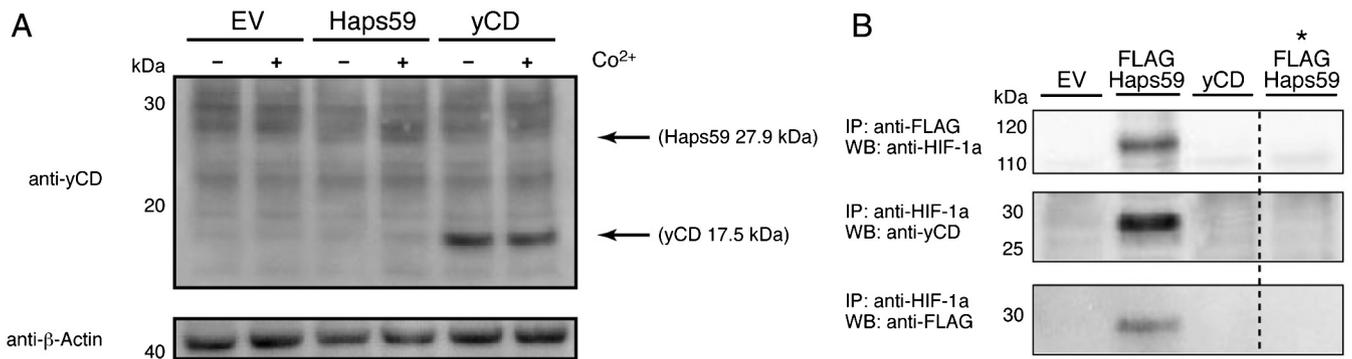


Fig. 57. (A) The effect of Co^{2+} on Haps59 and yCD accumulation in RKO cells stably expressing these proteins. Haps59 and yCD were detected by Western blot using anti-yCD antibodies. An EV was used as a negative control. The cross-reaction of the yCD antibodies with many human proteins complicated Haps59 detection, but the results suggest that Haps59 may accumulate to a higher level when the cells were grown under conditions that cause HIF-1a to accumulate. yCD accumulation was unaffected by HIF-1a accumulation. (B) Haps59 and HIF-1a interact in RKO cells. In these experiments, a FLAG epitope was appended to the N terminus of Haps59 (FLAG-Haps59) for switch immunoprecipitation and detection. Western blot (WB) analysis of coIP reactions with lysates of RKO cells expressing FLAG-Haps59, yCD, or an EV control show that HIF-1a is precipitated with anti-FLAG antibodies only in FLAG-Haps59-expressing cells (Top) and that FLAG-Haps59 is precipitated by anti-HIF-1a antibodies (Bottom). The far right lane (*) shows that coimmunoprecipitations of lysates of FLAG-Haps59-expressing cells using TATA-binding protein antibody [a negative control instead of the IP antibodies (Left)] did not result in detection of HIF-1a or FLAG-Haps59.