HHS Public Access

Author manuscript

Nat Chem Biol. Author manuscript; available in PMC 2023 April 01.

Published in final edited form as:

Nat Chem Biol. 2022 October ; 18(10): 1065–1075. doi:10.1038/s41589-022-01048-w.

Targeting colorectal cancer with small-molecule inhibitors of ALDH1B1

Zhiping Feng[1], **Marisa E. Hom**[1,10], **Thomas E. Bearrood**[1], **Zachary C. Rosenthal**[1], **Daniel Fernández**[2,3], **Alison E. Ondrus**[1,11], **Yuchao Gu**[4,5,6], **Aaron K. McCormick**[4], **Madeline G. Tomaske**[4], **Cody R. Marshall**[1], **Toni Kline**[7], **Che-Hong Chen**[1], **Daria Mochly-Rosen**[1], **Calvin J. Kuo**[4], **James K. Chen**[1,8,9,*]

[1]Department of Chemical and Systems Biology, Stanford University, Stanford, CA, USA

[2]Macromolecular Structure Knowledge Center, Stanford University, Stanford, CA, USA

[3]Stanford ChEM-H, Stanford University, CA, USA

[4]Department of Medicine, Stanford University, Stanford, CA, USA

[5]Department of Biochemistry, Stanford University, Stanford, CA, USA

[6]Stanford Cancer Institute, Stanford University, Stanford, CA, USA

[7]SPARK at Stanford, Stanford University, Stanford, CA, USA

[8]Department of Developmental Biology, Stanford University, Stanford, CA, USA

[9]Department of Chemistry, Stanford University, Stanford, CA, USA

[10]Present address: Department of Otolaryngology, Vanderbilt University Medical Center, Nashville, TN, USA

[11]Present address: Department of Chemistry, University of Illinois at Chicago, Chicago, IL, USA

Abstract

Aldehyde dehydrogenases (ALDHs) are promising cancer drug targets, as certain isoforms are required for the survival of stem-like tumor cells. We have discovered the first selective inhibitors of ALDH1B1, a mitochondrial enzyme that promotes colorectal and pancreatic

*Corresponding: jameschen@stanford.edu.

AUTHOR CONTRIBUTIONS

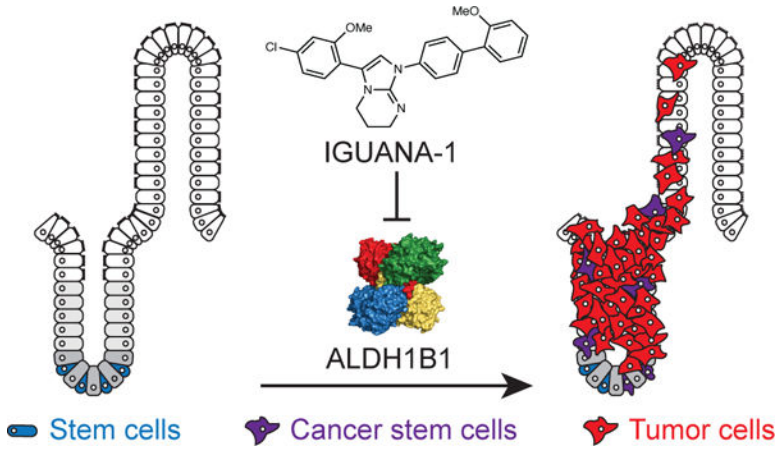
Z.F., M.E.H., T.E.B., Z.C.R., D.F., A.E.O., A.K.M., M.G.T., C.J.K., and J.K.C. designed the experiments and analyzed the data. A.E.O., T.K., and J.K.C. designed compounds and synthetic routes, and A.E.O. prepared compounds. A.E.O. conducted photoaffinity crosslinking, click chemistry, two-dimensional gel electrophoresis, and mass spectrometry analysis, and C.H.C. assisted with the two-dimensional gel electrophoresis. Z.F., M.E.H., T.E.B., Z.C.R., and C.H.C. expressed and purified ALDH proteins. Z.F., M.E.H., T.E.B., Z.C.R., and C.R.M. performed the enzyme kinetics assays and compound activity profiling. Z.F., Z.C.R., and D.F. conducted the X-ray crystallography studies of ALDH1B1 and ALDH1B1-inhibitor complexes. Z.F. and C.R.M. performed computational modeling of ALDH1-inhibitor complexes. Z.F., T.E.B., and Z.C.R. conducted spheroid and adherent culture growth assays. Z.F., M.E.H., and T.E.B. performed the ALDEFLUOR assays. Z.F. and M.E.H. conducted the Seahorse, HEK-293T viability, and Shh-LIGHT2 signaling assays. T.E.B. prepared the ALDH1B1 mutants. Z.F. and Y.G. generated the *ALDH1B1*$^{-/-}$ and *ALDH1B1* rescue clones, and Z.F. conducted related growth assays. Z.F. generated the 5-FU resistant SW480 cells and conducted related growth assays. A.K.M. and M.G.T. conducted the organoid experiments. Z.F. and Y.G. prepared total RNA samples for RNA-seq analysis, and Z.F. and J.K.C. analyzed the transcriptomic data. Z.F conducted cellular thermal shift assay and prepared TPP samples. T.E.B analyzed the TPP data. Z.F. and J.K.C. wrote the manuscript. A.E.O, D.M.R, C.J.K, and J.K.C. acquired funding for the project.

COMPETING FINANCIAL INTERESTS

J.K.C., Z.F., M.E.H., T.K., C.R.M., and A.E.O. have filed a PCT application related to the compounds described in this study.

cancer. We describe bicyclic imidazoliums and guanidines that target the ALDH1B1 active site with comparable molecular interactions and potencies. Both pharmacophores abrogate ALDH1B1 function in cells; however, the guanidines circumvent an off-target mitochondrial toxicity exhibited by the imidazoliums. Our lead isoform-selective guanidinyl antagonists of ALDHs (IGUANAs) exhibit proteome-wide target specificity, and they selectively block the growth of colon cancer spheroids and organoids. Finally, we have used genetic and chemical perturbations to elucidate the ALDH1B1-dependent transcriptome, which includes genes that regulate mitochondrial metabolism and ribosomal function. Our findings support an essential role for ALDH1B1 in colorectal cancer, provide molecular probes for studying ALDH1B1 functions, and yield leads for developing ALDH1B1-targeting therapies.

Graphical Abstract



INTRODUCTION

Cellular heterogeneity is a hallmark of tumor biology and a significant therapeutic challenge. Although most cancer drugs target the mechanisms that underlie rapid cell growth, multiple cell populations with distinct and potentially mutable phenotypic traits contribute to tumor progression. In particular, slow-cycling multipotent cells can persist after anti-proliferative therapies, eventually leading to tumor relapse, drug resistance, and metastasis. Drugs that selectively eliminate these stem-like populations therefore could effect more durable clinical responses.

One unique feature of stem and progenitor cells is their high levels of aldehyde dehydrogenase (ALDH) activity. ALDHs oxidize aldehydes through a step-wise, NAD(P)$^{+}$-dependent reaction cycle, and the human ALDH family consists of 19 isozymes with varying substrates and subcellular localizations. ALDH-expressing multipotent cells have been identified in a wide range of tissues, including the bone marrow, brain, breast, liver, pancreas, prostate, skeletal muscle, and skin[1]. Cells with high ALDH activities have also been observed in several cancers, and these stem-like populations exhibit greater tumorigenicity, chemo-radioresistance, and metastatic potential[2]. For example, high levels of ALDH1A3 have been reported in breast cancer[3], glioma[4], melanoma[5], and non-small cell lung cancer[6] stem cells. ALDH1A1 has been similarly linked to ovarian cancer stem cells[7],

and ALDH1B1 has been implicated in the tumor-initiating cells of colorectal cancer[8] and pancreatic ductal adenocarcinoma (PDAC)[9].

These ALDH isoforms are not only biomarkers for stem-like cell populations but also essential for their functions. For example, lentiviral *ALDH1A1* shRNAs suppress the clonogenic potential of ovarian cancer cells[10], and knocking down *ALDH1A3* impairs the tumorigenicity of lung cancer cells[6]. The oncogenic activities of ALDH1 isozymes have spurred the pursuit of small molecules that can inhibit these targets. Isoform-specific antagonists are a major challenge, and only a few selective inhibitors have been reported, including compounds that preferentially target ALDH1A1[11] or ALDH1A3[12,13].

Here we report the first specific small-molecule inhibitors of ALDH1B1, a mitochondrial ALDH isoform that can oxidize a broad range of substrates, including short- and medium-chain aliphatic aldehydes, retinaldehyde, and lipid peroxidation products[14]. ALDH1B1 is expressed at the base of colonic crypts where intestinal stem cells reside and in adult pancreas progenitors[9,15]. Although *Aldh1b1* knockout mice are viable[16], this isoenzyme plays important roles in cancer. ALDH1B1 is markedly upregulated in colorectal adenocarcinoma[15], and *ALDH1B1* silencing suppresses human colon cancer spheroid and xenograft growth[8]. ALDH1B1 is also highly expressed in certain human pancreatic cancer cells[17] and required for KRAS-driven PDAC in genetically engineered murine models[9].

We have discovered bicyclic imidazoliums and guanidines that selectively block ALDH1B1 activity with nanomolar potencies, established the molecular basis for ALDH1B1-inhibitor binding, and demonstrated the biochemical and functional specificity of guanidine antagonists in cells. We further show that these guanidine derivatives suppress the growth of colorectal cancer spheroids and organoids, with minimal effects on adherent cultures or non-cancerous cells. Finally, by comparing the transcriptional consequences of genetic and pharmacological ALDH1B1 deficiency, we have identified several ALDH1B1-dependent genes that regulate mitochondrial metabolism and ribosomal function. Our studies further establish ALDH1B1 as a promising cancer target, provide tool compounds for investigating ALDH1B1 functions, and take important first steps toward ALDH1B1-directed therapies.

RESULTS

Bicyclic imidazoliums can modulate ALDH enzymes

Our discovery of new ALDH-targeting pharmacophores arose fortuitously from a high-throughput screen for Hedgehog (Hh) pathway antagonists. These previous studies established bicyclic imidazolium **1** as a suppressor of Gli transcription factor activity and developed more potent analogs such as **2** (Fig. 1a and Supplementary Data Set 1)[18]. To identify potential protein targets through photoaffinity labeling, we synthesized aryl azide **3**, which also contains a propargyl group for click-chemistry tagging (Fig. 1a and Supplementary Data Set 1). Since the imidazoliums can suppress Hh signaling in NIH-3T3 fibroblasts[18], we treated live cultures of these cells with **3**, using **2** as a competitive inhibitor. We then exposed the fibroblasts to 254-nm light, lysed the cells, and labeled the photocrosslinked proteins with biotin-azide (Fig. 1b). Gel electrophoresis and far-western blotting revealed a specifically labeled protein with an apparent molecular weight of

about 60 kDa (Fig. 1c). As a complementary approach, we determined the subcellular localization of **3** by fixing the photoaffinity-labeled fibroblasts and tagging the propargyl group with tetramethylrhodamine (TAMRA)-azide. Fluorescence imaging revealed that **3** strongly localized to mitochondria (Extended Data Fig. 1a).

Guided by these results, we exploited the limited size of the mitochondrial proteome to identify the 60-kDa imidazolium-binding protein. We photoaffinity labeled isolated murine mitochondria with **3**, lysed the organelles, and tagged the probe-modified proteins with a fluorescent azide. The reaction mixture was then resolved by two-dimensional gel electrophoresis, revealing fluorescent protein spots in the 60-kDa range (Fig. 1d and Extended Data Fig. 1b). Mass spectrometry-based sequencing of these resolved spots identified ALDH2 as the most abundant candidate target (Fig. 1e and Supplementary Data Set 2), which we confirmed by western blotting (Fig. 1f).

To determine the effects of imidazoliums on ALDH2 and other ALDH family members, we tested **1**-**3** on ten human isoforms in enzyme kinetic assays. Surprisingly, only the photoaffinity probe suppressed ALDH2 activity at a 50-μM dose (Fig. 1g), but each compound was able to suppress or enhance the activity of other ALDH family members. **2** preferentially targeted members of the ALDH1 subfamily, and it potently and selectively inhibited ALDH1B1 (IC50 = 57 nM; Fig. 1h), the isoform most homologous to ALDH2 (72% sequence identity). Since ALDH1B1 is not expressed at detectable levels in NIH-3T3 cells (Supplementary Fig. 1), this isoenzyme is unlikely to be the relevant target for imidazolium-mediated Hh pathway inhibition. However, **2** provided us with a starting point for developing ALDH1B1-specific antagonists, and we focused our subsequent efforts toward this goal.

Structural basis of ALDH1B1 inhibition by imidazoliums

To understand the mechanism of ALDH1B1 inhibition, we conducted additional enzyme kinetic assays with **2** and varying concentrations of acetaldehyde or NAD^+. The inhibitor acted in a non-competitive manner with respect to the aldehyde substrate (Fig. 2a) and exhibited an uncompetitive relationship with NAD^+ (Fig. 2b). These findings are consistent with a sequential ordered mechanism of inhibition, in which NAD^+ is a prerequisite for ALDH1B1-imidazolium binding. The inability of acetaldehyde to competitively suppress imidazolium action also suggested that the inhibitor either targets an allosteric site or prevents ALDH1B1 from cycling through its aldehyde-binding state.

We next pursued structural studies of ALDH1B1. We co-expressed the enzyme with GroEL/GroES chaperones, which enhanced ALDH1B1 folding and solubility (Supplementary Fig. 2a–b). We then employed the vapor-diffusion method to grow crystals of ALDH1B1 bound to NAD^+ in the absence or presence of **2** (Supplementary Fig. 2c). Both enzyme complexes formed crystals in the same trigonal space group, with two polypeptide chains per asymmetric unit and one NAD^+ and one inhibitor (when included) per polypeptide chain. Like other ALDH1 isoforms and ALDH2, ALDH1B1 forms a tetramer composed of two dimers. Refinement of the crystallographic data provided us with structures of ALDH1B1 and its inhibitor-bound form with resolutions of 2.67 Å and 2.12 Å, respectively (Supplementary Table 1 and Supplementary Fig. 3a–b). Inhibitor binding did not coincide

with major changes in ALDH1B1 conformation, as the root-mean-square deviation (RMSD) of the two superimposed structures was 0.28 Å (Supplementary Fig. 4a). NAD^+ also adopted comparable conformations in the free and inhibitor-bound enzymes (Supplementary Fig. 4b–c).

In the ALDH1B1/NAD^+/**2** crystal structure, the inhibitor interacted with the active site through all three substituents on the imidazolium ring (Fig. 2d–e). The N1 2-methoxyphenyl group engaged the aldehyde-binding pocket that includes several conserved landmarks within the ALDH family[19,20]: the catalytic cysteine (Cys302; residue numbering is based on the mature mitochondrial protein), the oxyanion hole (Asn169), and the aromatic box (Phe170, Trp177, and Phe465). The aryl moiety also interacted with other hydrophobic and charged residues in this region (Asp121, Glu124, Val173, Met174, Phe296, Cys301, Cys303, and Val459). The C4 1,1'-biphenyl substituent interacted with an extended cleft (Glu288, Gln289, Glu292, Phe296, and Asn457), and the tetrahydroazepine ring occupied a pocket that bridges the two sites (Leu120, Glu124, Ile458, and Thr460). Neither the 2-methoxyphenyl-imidazolium nor 1,1'-biphenyl-imidazolium systems were coplanar, and the tetrahydroazepine ring adopted a boat-like conformation. The electron density map also revealed a close contact between the imidazolium ring and the Asn457 backbone carbonyl (3.3 Å) that likely reflects an n-to-π^* interaction (Fig. 2e).

To gain insights into the molecular basis of inhibitor specificity, we aligned the ALDH1B1/NAD^+/**2** structure with those of ALDH1A1, ALDH1A2, and ALDH1A3 (Supplementary Fig. 5). These overlays reveal potential steric clashes between the imidazolium substituents and several active site residues in the ALDH1A paralogs. For example, His293 and Val460 in ALDH1A1 are predicted to diminish protein interactions with the 1,1'-biphenyl and the tetrahydroazepine moieties, respectively. The biphenyl system could be occluded by Gln289 in ALDH1A2, and multiple residues in this isoform (Val120, Asp121, and Leu173) might abrogate binding of the 2-methoxyphenyl group. Our models also indicate that a subset of the analogous residues in ALDH1A3 (Ile132 and Leu185) could sterically clash with the 2-methoxyphenyl substituent. These observations are consistent with our enzyme kinetic studies, and they suggest that the ALDH1B1 selectivity of **2** stems from multiple repulsive interactions with the other ALDH1 isoforms.

SAR profile of imidazolium-based inhibitors

Since the ALDH1B1/NAD^+/**2** structure revealed protein contacts with each of the imidazolium substituents, we examined how chemical modifications at these sites would influence inhibitor potency and isoform selectivity. We synthesized 60 additional imidazolium derivatives (**4**-**63**; Supplementary Data Set 1) and assessed their activities against ALDH1B1, ALDH1A1, ALDH1A2, and ALDH1A3 at a 1-μM compound dose. This focused library exhibited a broad range of potencies against ALDH1B1, with the most active analogs suppressing enzyme activity by at least 16-fold and several exhibiting high isoform selectivity (Fig. 3a–b).

The structure-activity relationship (SAR) profiles of the imidazoliums revealed certain trends. The ALDH1B1 active site could accommodate both 2- and 4-substituted phenyl groups at the N1 position (e.g., **4** and **6**) and various cyclic systems fused to the imidazolium

(e.g., **7** and **9**). Substituents on the C4 1,1'-biphenyl system also altered isoform specificity. For example, adding a 2'-methoxy group or replacing the distal ring with pyridin-4-yl group increased inhibitor selectivity for ALDH1B1 (**8** and **13**, respectively), and functional groups at the 4' position could promote ALDH1A3 inhibition (**10** and **46**). To further examine the isoform selectivity of **8** and **13**, we profiled their activities against the broader panel of ten ALDH family members. Both derivatives had enhanced specificity for ALDH1B1 compared to **2** (Fig. 3c), and **2** and **8** had comparable potencies (Fig. 3d).

Cellular activities of imidazolium-based ALDH1B1 inhibitors

Having developed ALDH1B1-selective imidazoliums with nanomolar activities *in vitro*, we next evaluated their ability to suppress ALDH1B1 activity in cells. We first tested selected analogs in an ALDEFLUOR assay, which uses a fluorescent aldehyde to quantify cellular ALDH activity. For these experiments, we transiently expressed ALDH1B1 in *ALDH1A3*$^{-/-}$ A375 cells, a melanoma line with minimal endogenous ALDH expression[21]. The exogenous ALDH1B1 protein localized to the mitochondria in these cells (Extended Data Fig. 2a), and its ALDEFLUOR signal could be readily detected and quantified by flow cytometry. This cellular ALDH1B1 activity was abrogated by the pan-ALDH1 antagonist *N*,*N*-diethylaminobenzaldehyde (DEAB) and ALDH1B1-selective imidazoliums (Extended Data Fig. 2b) but not by an inactive analog (Extended Data. Fig. 2c–e).

We next assessed the effects of the ALDH1B1 antagonists on colon cancer cells. Public databases indicate that *ALDH1B1* is highly upregulated in colorectal adenocarcinoma relative to other cancer types and normal colorectal tissues (Supplementary Fig. 6)[22,23], and as described above, *ALDH1B1* silencing can suppress the growth of colon cancer spheroids and xenografts[8]. Since *ALDH1B1* shRNAs could have off-target effects[24], we re-examined ALDH1B1 function in colon cancer cells using CRISPR/Cas9 gene editing. We first used lentiviral vectors to constitutively express Cas9 and three individual *ALDH1B1* guide RNAs (gRNAs) in SW480 cells. As determined by TIDE (Tracking of Indels by Decomposition) analysis[25], two of the gRNAs (gRNA-1 and gRNA-2) could efficiently induce loss of wild-type *ALDH1B1* alleles (Supplementary Fig. 7).

To minimize possible off-target effects caused by constitutive CRISPR activity, we then transiently transfected plasmids encoding Cas9, an EGFP reporter, gRNA-1, and gRNA-2 into SW480 cells. Transfected cells were sorted by green fluorescence and clonally expanded, and multiple clones were characterized by DNA sequencing and western blot analysis (Extended Data Fig. 3a). A homozygously edited clone with complete loss of ALDH1B1 function (SW480 clone 2, Supplementary Fig. 8) and an unedited clone (SW480 clone 3) were selected for further study. Consistent with the shRNA phenotype[8], the *ALDH1B1*$^{-/-}$ lines were markedly less viable than the unedited cells when grown in Matrigel-based three-dimensional cultures (Extended Data Fig. 3b–c). This inhibitory effect was selective for spheroid growth, as adherent cultures did not exhibit the same degree of ALDH1B1 dependency (Extended Data Fig. 3d). A spheroid dependence on ALDH1B1 function is consistent with a role of this enzyme in stem cell maintenance, as colon cancer spheroids are enriched for and require these self-renewing populations[26]. We observed similar CRISPR phenotypes when the *ALDH1B1* gRNAs and *Cas9* plasmid were transiently

transfected into a second colon cancer-derived line, HCT116 cells (HCT116 clones 1 and 3; Extended Data Fig. 3e–h and Supplementary Fig. 8).

To ensure that these cellular phenotypes are not due to clonal variation, we examined the ability of exogenous *ALDH1B1* to rescue the *ALDH1B1*$^{-/-}$ SW480 clone. *ALDH1B1* was re-introduced by lentiviral transduction, and western blot analysis confirmed the restoration of ALDH1B1 expression to endogenous-like levels (Extended Data Fig. 4a). *EGFP* was similarly transduced into the knockout and unedited clones to provide comparison controls. Exogenous *ALDH1B1* but not *EGFP* restored the viability of *ALDH1B1*$^{-/-}$ SW480 cells under spheroid conditions, whereas both lentiviral constructs had minimal effects on adherent cultures (Extended Data Fig. 4b–d). These results firmly establish ALDH1B1 as a key promoter of colon cancer spheroid growth.

We consequently used the spheroid growth phenotype to evaluate the efficacy and specificity of our ALDH1B1-targeting inhibitors. SW480 spheroid formation was strongly suppressed by **2** and **8** at a 2-μM dose, whereas the inactive analog **63** did not impede the growth of these Matrigel-based cultures (Fig. 4a–b). However, the imidazoliums had comparable anti-proliferative activities against adherent SW480 cells (Fig. 4c). We hypothesized that this lack of selectivity was due to off-target cytotoxicity, which we surmised was related to the mitochondrial localization of the compounds. Other hydrophobic cations have been found to accumulate in mitochondria and disrupt oxidative phosphorylation[27]. Consistent with this idea, Seahorse assays revealed that both **2** and **8** suppress the oxygen consumption rate (OCR) of SW480 cells in an ALDH1B1-independent manner (Fig. 4d).

Isosteric guanidine-based ALDH1B1 inhibitors

Due to the mitochondrial toxicity of imidazoliums, we explored whether the constitutively cationic scaffold could be replaced with a bicyclic guanidine, which would become positively charged upon protonation (**64**-**97**; Supplementary Data Set 1). Several guanidines could inhibit ALDH1B1 with potencies and isoform selectivities that were comparable to those of their imidazolium counterparts (Fig. 5a). This new class of isoform-specific guanidinyl antagonists of ALDHs included **68**, **69**, **72**, and **80**, which we named IGUANA-1 through IGUANA-4. We next investigated how the IGUANAs interacted biochemically with ALDH1B1. Like the imidazoliums, they were non-competitive with aldehyde substrate and uncompetitive with NAD^+ (Extended Data Fig. 5a–b). We also solved the crystal structure of ALDH1B1 bound to IGUANA-4, the guanidine equivalent of **2**, at 2.30 Å resolution (Supplementary Table 1 and Supplementary Fig. 4c. Overlays of the ALDH1B1/NAD^+/IGUANA-4 and ALDH1B1/NAD^+/**2** structures confirmed that the two inhibitor classes adopt nearly identical binding modes (Extended Data Fig. 5c), including a close contact between the Asn457 backbone carbonyl and the guanidine π system (3.2 Å) in the ALDH1B1/NAD^+/IGUANA-4 complex (Extended Data Fig. 5d).

Like imidazoliums, the IGUANAs could effectively inhibit exogenous ALDH1B1 activity in *ALDH1A3*$^{-/-}$ A375 cells, as assessed by the ALDEFLUOR assay (Extended Data Fig. 6a). We also confirmed their ability to target endogenous ALDH1B1 in live SW480 cells by observing inhibitor-dependent stabilization of the enzyme in a cellular thermal shift assay (CETSA) (Extended Data Fig. 6b–c)[28]. The IGUANAs did not inhibit the OCR

of SW480 cells (Fig. 5b), and they could preferentially suppress the growth of SW480 spheroid cultures (Fig. 5c and Extended Data Fig. 7a–b) at least in part by inducing cell death (Extended Data Fig. 7c). The guanidines were also much less cytotoxic than the imidazoliums to HEK-293T cells, a non-cancerous line (Supplementary Fig. 9a), and they were considerably less potent against Hh signaling in NIH-3T3 cells (Supplementary Fig. 9b).

To more comprehensively assess the target specificity of IGUANAs, we examined their proteomic and genetic interactions in SW480 cells. We first conducted thermal proteome profiling (TPP)[29] to determine potential biochemical targets in this colon cancer line. We treated SW480 cells with DMSO or 2 μM IGUANA-1 for 1 hour and then heated the cultures to temperatures ranging from 37 and 68 °C. Proteins that resisted heat denaturation and remained soluble in the resulting cell lysates were labeled with isobaric tandem mass tags and quantified by mass spectrometry. Using this approach, we identified five proteins as the most likely candidates for direct targets ($\Delta T_m \geq 1.5$ °C and $P < 0.05$) (Fig. 5d and Supplementary Data Set 3). Confirming our CETSA results, ALDH1B1 exhibited the largest inhibitor-dependent increase in thermal stability in the TPP study. The TPP candidates with the next largest ΔT_m values, ARCN1 and NIPSAP2, were not validated by CETSA (Supplementary Fig. 10), indicating that they are false positives in the proteome-wide survey.

We next confirmed the functional specificity of the IGUANAs by determining whether ALDH1B1 is required for their ability to preferentially inhibit colon cancer spheroid growth. We compared the pharmacological phenotypes of *ALDH1B1*$^{-/-}$ SW480 cells that were lentivirally transduced with either *EGFP* or exogenous *ALDH1B1*. While the *ALDH1B1*-rescued line exhibited greater inhibitor sensitivity when cultured under spheroid rather than adherent conditions, the *EGFP*-expressing cells did not (Fig. 5e and Supplementary Fig. 11). We also generated an inhibitor-resistant variant of ALDH1B1, guided by the ALDH1B1/NAD^+/IGUANA-4 crystal structure and sequence alignments of the human ALDH1 subfamily and ALDH1B1 homologs across species (Supplementary Figs. 12–14). We selected seven residues in the inhibitor-binding site to mutate, and we evaluated the activity and inhibitor sensitivity of each ALDH1B1 construct in the ALDEFLUOR assay (Supplementary Fig. 15). We then tested selected mutants in enzyme kinetics assays (Supplementary Fig. 16). These experiments revealed three mutations (E124G, V173I, and N457G) that impart partial resistance to IGUANA-1, and by testing pairwise combinations of these residue changes, we obtained two ALDH1B1 variants (E124G/N457G and V173I/N457G) with markedly reduced inhibitor sensitivity (Supplementary Figs. 15–16). The E124G/N457G mutant could also restore the ability of *ALDH1B1*$^{-/-}$ SW480 cells to grow as spheroid cultures, albeit with lower efficacy than the wild-type enzyme (Supplementary Fig. 11). Spheroid cultures of E124G/N457G mutant-expressing cells were more resistant to IGUANA-1 than those transduced with wild-type *ALDH1B1* (Fig. 5e), approaching the inhibitor sensitivity of adherent SW480 cultures. Taken together, these proteomic and functional studies establish the IGUANAs as highly selective ALDH1B1 inhibitors and validate their mechanism of action in the colon cancer cells.

Therapeutic potential of ALDH1B1 inhibitors

Our genetic and pharmacological results suggest that ALDH1B1 antagonists could be effective treatments for advanced colorectal cancer. Since 5-FU is a current mainstay for advanced colorectal cancer[30], we compared how SW480 cells respond to IGUANAs and this cytotoxic chemotherapy. When used together, IGUANA-3 and 5-FU blocked SW480 spheroid growth with a combination index[31] of approximately 1 (Supplementary Fig. 17), corroborating their independent mechanisms of action. We also generated a 5-FU-resistant SW480 line by culturing the colon cancer cells in the presence of this drug for multiple passages. These cells were approximately 20-fold less sensitive to 5-FU than the parental line, but IGUANA-2 could suppress the spheroid growth of both lines with comparable potency (Fig. 6a).

To further evaluate ALDH1B1 inhibitor efficacy against colorectal cancer, we examined their effects on primary human organoid cultures representing normal colon, colon adenocarcinoma, or rectal adenocarcinoma (patient-derived organoids, PDOs[32]). We tested two human wild-type colon and four colorectal tumor PDO lines (Supplementary Table 2), which were cultured for three to five days in a submerged basement membrane extract type 2 format[33] and then treated with varying concentrations of IGUANA-1 for seven days. Although the colorectal adenocarcinoma PDOs carried varying oncogenic mutations in *APC*, *KRAS*, and/or *TP53*, the ALDH1B1 inhibitor was able to inhibit the growth of each line (Fig. 6b). The wild-type colon PDOs were less sensitive to IGUANA-1, perhaps reflecting the non-essential role for ALDH1B1 in normal colon physiology[16].

Identification of ALDH1B1-dependent genes

Our findings establish IGUANAs as ALDH1B1 inhibitors with the potency and specificity required for functional studies. Toward that goal, we used RNA sequencing (RNA-seq) to compare how genetic or chemical loss of ALDH1B1 function affects gene expression in colon cancer cells (Fig. 6c and Supplementary Data Set 4). One RNA-seq data set examined the transcriptomes of clonal ALDH1B1$^{-/-}$ SW480 cells that were lentivirally transduced with either *EGFP* or exogenous *ALDH1B1*. A second data set analyzed the transcriptomes of the parental SW480 cells after 24 hours of treatment with DMSO or 2 μM IGUANA-1.

To facilitate the identification of ALDH1B1-dependent genes, we focused on those with statistically significant transcriptional changes ($P_{adj} < 0.05$) in both data sets (Fig. 6d and Supplementary Data Set 4). Several transcripts were downregulated by both genetic and chemical loss of ALDH1B1 function, and Gene Ontology (GO) enrichment analyses revealed significant enrichment for genes associated with mitochondrial metabolism (*ATP5ME*, *ATP5PO*, *LDHB*, *NDUFB1*, *NDUFB8*, *NDUFC2*, *NDUF1A*, *NDFUA9*, *NDUFA11*, and *UQCR11*), and ribosomal biogenesis and function (*RPL15A*, *RPL21*, *RPL26*, *RPL29*, *RPL38*, and *TMA7*) (Supplementary Fig. 18). Both perturbations also reduced the expression of *KRT15* and *DCLK1*, which are genetic markers of colorectal cancer stem cells[34,35].

In addition to these conserved transcriptional changes, genetic disruption of *ALDH1B1* coincided with the upregulation of genes that were not induced by the ALDH1B1 inhibitor

(Fig. 6c and Supplementary Data Set 4), potentially reflecting cellular phenotypes caused by prolonged ALDH1B1 deficiency. In principle, these transcripts could encode oncogenes that are activated through compensatory mechanisms or tumor suppressors that are negatively regulated by ALDH1B1. Consistent with this idea, the *ALDH1B1*$^{-/-}$ cells upregulated multiple members of the pregnancy-specific glycoprotein and semaphorin families, some of which are known drivers of gastrointestinal malignancies[36–38]. Sustained loss of ALDH1B1 function also induced other colorectal cancer-associated factors that are either overexpressed and correlate with poor clinical prognoses (e.g., *AGPAT4*, *BST2*, *CACNA1F*, *CNTN1*, *IGF1*, *HSD17B2*, *ITIH3*, *ORL1*, and *PLOD2*) or function as tumor suppressors (e.g., *CEACAM1*, *EDN3*, *FILIP1L*, and *TREM2*). We note that neither genetic nor chemical loss of ALDH1B1 function significantly induced the expression of other ALDH family members (Supplementary Fig. 19). Together, these transcriptomic data sets support a model in which ALDH1B1 mobilizes metabolic and translational pathways in colon cancer cells that maintain stemness and promote tumor progression.

DISCUSSION

ALDH1 upregulation in human malignancies was first recognized as a direct mechanism of chemoresistance, exemplified by the ability of these enzymes to metabolize the active form of cyclophosphamide[39]. Subsequent studies have shown that ALDH1 isoforms are predominantly expressed in the stem-like populations that drive tumor initiation and progression, with high ALDH1 activity in tumors frequently correlating with poor clinical outcomes. ALDH1 function in cancer stem cells is not limited to drug detoxification, as the genetic suppression of individual isoforms has been shown to reduce tumorigenicity and metastasis. Although the precise roles of ALDH1 isoforms in these self-renewing populations remain unclear, their requirement for cancer stem cell maintenance has spurred the development of small-molecule ALDH1 antagonists.

To the best of our knowledge, IGUANAs represent the first selective inhibitors of ALDH1B1. While these compounds inhibit ALDH1B1 in a manner that is non-competitive with respect to aldehyde substrate, they engage the substrate-binding pocket rather than an allosteric site. We speculate that this kinetic behavior reflects an ordered, multi-step mechanism of ALDH1B1 action, as has been observed in other bisubstrate enzymes. In addition, ALDH enzymes composed of dimer pairs have been found to have one active subunit in each dimer per catalytic cycle, which may result from negative cooperativity among the tetramer subunits[40]. Our compounds therefore might block completion of the catalytic cycle and impede formation of new ALDH1B1/NAD^+ complexes that are capable of aldehyde binding.

Consistent with the isosteric nature of IGUANAs and their imidazolium counterparts, both pharmacophores bind to ALDH1B1 through equivalent molecular interactions. The ALDH1B1 antagonists adopt a pin-wheel-like conformation, with the imidazolium and guanidine scaffolds participating in a core n-to-π^* interaction with the Asn457 backbone carbonyl. Inhibitor potency and selectivity are enhanced by substituents that extend from and are non-coplanar with the central imidazolium or cyclic guanidine, engaging contiguous but distinct pockets in the active site. In comparison, structural studies of non-covalent

ALDH1A1 and ALDH1A3 antagonists reveal that these planar, multi-ringed structures derive their core binding energy through hydrophobic π-stacking interactions[12,13,41–43]. Our imidazolium and guanidine derivatives therefore represent an alternative mode for targeting the ALDH1 active site.

Despite these biochemical similarities, our studies uncover key differences between the cellular functions of imidazolium and guanidine derivatives. Although both pharmacophores can target ALDH1B1 in cells with comparable potencies, only the imidazoliums disrupt the electron transport chain in mitochondria. IGUANAs avoid these mitochondrial off-target effects, possibly due to their reversible cationic character, greater total polar surface area, and/or hydrogen-bonding capabilities. Accordingly, only the IGUANAs can selectively suppress colon cancer spheroid growth, recapitulating the *ALDH1B1*$^{-/-}$ phenotype. Moreover, our CETSA and TPP studies, rescue experiments with *ALDH1B1*$^{-/-}$ SW480 cells, and transcriptomic analyses demonstrate that IGUANAs have high biochemical and functional specificity in cells.

The potency and specificity of IGUANAs make them valuable tools for interrogating ALDH1B1 functions. *ALDH1B1* silencing in SW480 cells has been reported to downregulate Wnt, Notch, and phosphoinositide 3-kinase/Akt signaling[8]. However, the mechanisms by which ALDH1B1 promotes these oncogenic pathways are not yet known. By applying both CRISPR mutagenesis and our small-molecule probes to colon cancer cells, we have uncovered connections between ALDH1B1 activity, mitochondrial metabolism, ribosomal function, and stemness. While the epistatic relationships between the latter processes are yet to be determined, there is growing evidence that cancer stem cells prefer mitochondrial oxidative metabolism for energy production[44] and rely upon specialized ribosomes[45]. Moreover, our combined approach allows these immediate transcriptional changes to be discerned from others that occur over a longer timeframe. Unlike pharmacological ALDH1B1 blockade, CRISPR-mediated disruption of *ALDH1B1* resulted in the upregulation of numerous oncogenes and tumor suppressors associated with colorectal cancer. The expression levels of these factors can correlate with patient survival, and additional ALDH1B1-dependent genes from our study could have similar prognostic value. We anticipate that our small-molecule probes will be valuable tools for elucidating the ALDH1B1-dependent metabolome in cancer cells and investigating the relationships between these metabolites, oncogenic signaling, and stemness.

In addition to enabling these mechanistic investigations, IGUANAs provide a starting point for designing ALDH1B1-targeting drugs. Our findings and those by other laboratories establish an essential role for ALDH1B1 in colorectal cancer and PDAC. Anti-mitotic drugs remain the standard first-line treatment for these malignancies, and most patients with advanced cases succumb to the disease within five years[46]. Consistent with these clinical observations, 5-FU has been reported to promote the stemness of patient-derived colorectal cancer cells[47], and FOLFIRINOX (folinic acid, 5-FU, irinotecan, and oxaliplatin) has the same effect on some pancreatic cancer lines[48]. Small-molecule ALDH1B1 antagonists could provide a means for targeting the stem-like populations that sustain colorectal and pancreatic cancer growth, leading to single-agent or combination therapies with more durable clinical responses. The efficacy of ALDH1B1 inhibitors against 5-FU-resistant SW480 spheroids

and colorectal tumor PDOs lends support for this therapeutic paradigm. Moreover, the viability of *Aldh1b1*$^{-/-}$ mice predicts that ALDH1B1 blockade will have tolerable effects on normal physiology. Developing ALDH1B1 antagonists with drug-like properties will be an important next step toward this goal.

ONLINE METHODS

Chemical synthesis

Synthetic routes and structural characterization data for the bicyclic imidazoliums and guanidines are described in the Supplementary Information (Supplementary Note).

Cell lines

SW480 and HCT116 colorectal cancer cells were purchased from ATCC and maintained in RPMI-1640 containing 10% fetal bovine serum (FBS; MilliporeSigma), 100 U/mL penicillin, and 100 μg/mL streptomycin. *ALDH1A3*$^{-/-}$ A375 melanoma cells[21] were provided by E. Patton (University of Edinburgh) and maintained in DMEM containing 10% FBS, 5 μg/mL L-glutamine, 100 U/mL penicillin, 100 μg/mL streptomycin, and 0.8 μg/mL puromycin. NIH-3T3 cells were purchased from ATCC and maintained in DMEM containing 10% calf serum (CS), 1% sodium pyruvate, 100 U/mL penicillin, and 100 μg/mL streptomycin. HEK-293T cells were purchased from ATCC and maintained in DMEM containing 10% FBS, 2 mM L-glutamine, 100 U/mL penicillin, and 100 μg/mL streptomycin. Shh-LIGHT2 cells[49] were maintained in DMEM containing 10% CS, 1% sodium pyruvate, 100 U/mL penicillin, 100 μg/mL streptomycin, 150 μg/mL zeocin, and 400 μg/mL G418. All cells were cultured at 37 °C with 5% CO_2.

Photoaffinity crosslinking of imidazolium-binding proteins

To detect imidazolium-binding proteins, NIH-3T3 cells were seeded into a 15-cm plate (800,000 cells/plate) for 48–60 h to reach 85–90% confluence. The growth media was aspirated, and the cells were incubated with phenol red-free DMEM containing 0.5% CS, 7 μM **3**, and either 10 μM **2** or an equivalent amount of DMSO vehicle (0.01%, v/v) at 37 °C for 30 min. The cells were washed once with cold phosphate-buffered saline (PBS) and maintained in PBS on ice. The lid was removed from each plate, and the cells were exposed to 254-nm UV light using a photocrosslinker (UV Stratalinker 2400; Stratagene) for 20 s at a power of 250,000 μJ/cm^2. Cells were then dislodged by scraping in PBS, pelleted by centrifugation at 750 x *g* for 5 min at 4 °C, and washed with 2 × 10 mL fresh, ice-cold PBS. The cell pellet was dissolved in 0.1% SDS in PBS and sonicated at room temperature using a probe-tip sonicator. Total protein concentration was measured using the Pierce 660 Protein Assay Kit (Thermo Fisher Scientific). Biotin-azide (1.92 μL of a 5 mM solution in DMSO, Sigma) was added to 20 μg cell lysate in 75 μL 0.1% SDS in PBS in 1.5-mL microcentrifuge tubes. In a glass vial, 200 μL $CuSO_4$ (50 mM in H_2O) was added to a solution of 600 μL TBTA (1.7 mM in *t*-BuOH-DMSO (4:1)) and 200 μL TCEP (50 mM in H_2O, adjusted to pH 7–8 immediately before use). The resulting catalyst mixture was vortexed, a 9.6-μL aliquot was added to each tube, and tubes were incubated with shaking at 30 °C for 30 min. To quench the reaction and remove excess reagents, protein was precipitated with 1 mL acetone at −20 °C, tubes were incubated at −20 °C for 30 min, and protein was collected

by centrifugation at 20,000 x *g* for 10 min at 4 °C. The resulting supernatant was aspirated, and the pellet was washed two times with cold acetone in the same manner. After the last wash, residual acetone was removed by air-drying the pellet at room temperature for 5 min. The samples were then solubilized in SDS-PAGE loading buffer and resolved on 4–12% Criterion Bis-Tris protein gel (Bio-Rad) at 150 V for 1.5 h at room temperature.

The gel proteins were next transferred to an Immun-Blot Low Fluorescence PVDF membrane (Bio-Rad) using a Trans-Blot Turbo transfer system (Bio-Rad) at a current of 2.5 A for 7 min at room temperature. For far-western analysis of biotinylated target proteins, the membrane was incubated with blocking buffer (I-Block, Invitrogen) in 0.1% Tween-20 in PBS (PBST) for 1 h at room temperature and then treated with Pierce High Sensitivity Streptavidin-HRP (1:10,000 dilution; Thermo Fisher Scientific) in blocking buffer for 1.5 h at room temperature. The membrane was washed with 3 × 5 min PBST, and chemiluminescent signal was generated using Supersignal West Dura ECL reagent (Thermo Fisher Scientific) and detected using a ChemiDoc imaging system (Bio-Rad).

To determine the subcellular localization of the imidazolium-binding protein, NIH-3T3 cells were seeded in a 24-well plate (60,000 cells/well) containing poly-D-lysine-coated coverslips and cultured for 48–60 h to reach 85–90% confluence. The growth media was aspirated, and the cells were incubated with phenol red-free DMEM containing 0.5% CS, 7 μM **3**, and either 10 μM **2** or an equivalent amount of DMSO vehicle (0.01%, v/v) at 37 °C for 30 min. To assess mitochondrial morphology, cells were co-treated with 500 nM MitoTracker Deep Red FM (Invitrogen). The plate then cooled on ice, the lid was removed, and cells were exposed to 254-nm UV light as described above. Compound media was aspirated, and the cells were washed 3 × 5 min with PBS. After washing, cells were fixed with MeOH for 20 min at −20 °C. Fixed cells were washed 3 × 2 min with PBS, and TAMRA-azide (1.28 μL of a 5 mM solution in DMSO) in 200 μL PBS was added. Each well was then treated with 25.6 μL of the $CuSO_4$/TBTA/TCEP catalyst mixture described above, and the plate was covered in foil and rocked gently for 45 min at room temperature. The reaction mixture was aspirated, and the cells were washed 3 × 5 min with PBS, once with ddH_2O, and mounted on glass slides in ProLong Gold mounting media (Invitrogen). The coverslips were then imaged on a Zeiss LSM 700 confocal microscope.

To determine the identity of the imidazolium-binding protein, intact mitochondria were isolated from the liver of a 12-week-old female C57Bl/6 mouse as previously described[50]. The animal use was approved by the Stanford University Administrative Panel on Laboratory Animal Care (Protocol Number 14145). 100 μg of purified mitochondria were diluted in 75 μL PBS, 7 μM **3**, and either 10 μM **2** or an equivalent amount of DMSO vehicle (0.02%, v/v) were added. The mixture was rocked gently at 37 °C for 30 min and then transferred to one well of a 24-well plate. The lid of the plate was removed, and the mitochondria were exposed to 254-nm UV light as described above. The photocrosslinked mitochondria were transferred to a 1.5-mL microcentrifuge tube, homogenized in PBS containing 0.1% SDS, and the appropriate fluorophore-azide (1.92 μL of a 5 mM solution in DMSO) was added to 75 μL of this mixture. For in-gel fluorescence analysis, the competitor- and DMSO-treated samples were clicked to Cy3-azide and Cy5-azide (Click Chemistry Tools), respectively. For mass spectrometry analysis, the competitor- and DMSO-

treated samples were clicked to MegaStokes 673-azide (Sigma) and BODIPY-azide (BDT[51], provided by G. Ponhert, Friedrich Schiller University), respectively. Each tube was treated with a 9.6-μL aliquot of the $CuSO_4$/TBTA/TCEP catalyst mixture and incubated with shaking at 30 °C for 30 min.

For in-gel fluorescence analysis, competitor- and DMSO-treated samples were mixed in 1:1 ratio at the end of the reaction. For mass spectrometry analysis, the competitor- and DMSO-treated samples were mixed in a ratio of 10:1 at the end of the reaction. After mixing, protein was immediately precipitated with 1 mL acetone at −20 °C to quench the reaction and remove excess reagents. The acetone suspensions were incubated at −20 °C for 30 min, and protein was collected by centrifugation at 20,000 x *g* for 10 min at 4 °C. The resulting supernatant was aspirated, and the pellet was washed two times with cold acetone in the same manner. After the last wash, residual acetone was removed by air-drying the pellet at room temperature for 5 min. Protein samples were re-dissolved in 125 μL rehydration buffer (7 M urea, 2 M thiourea, 4% CHAPS), supplemented with 20 μL DTT (2 M), 5 μL Pharmalyte 3–10 IPG buffer (GE Healthcare), and trace bromothymol blue, and resolved on a pH 3–10 Immobiline Drystrip (GE Healthcare) using an Ettan IPGphor IEF System (GE Healthcare) for 50,000 Vh (8,000 V and 45 μA) for 14 h. The IPG strip was incubated with gentle agitation in equilibration buffer (6 M urea, 30% (w/v) glycerol, 2% (w/v) SDS, 0.05 M Tris-HCl, pH 8.8) supplemented with 10 mg/mL DTT, then in equilibration buffer supplemented with 40 mg/mL iodoacetamide for 15 min each. The IPG strip was rinsed with MES running buffer then mounted on a 4–12% Criterion Bis-Tris protein gel (Bio-Rad) and immobilized in agarose, and proteins were resolved at 150 V for 1.5 h at room temperature.

For in-gel fluorescence analysis, fluorescently labeled protein spots were detected using the ChemiDoc imaging system, and total protein was visualized with Sypro Ruby (Thermo Fisher Scientific). For mass spectrometry analysis, fluorescently labeled spots were visualized on a UV transilluminator, excised with a cut 25-μL pipette tip, and subjected to in-gel digestion for mass spectrometry-based sequencing. All mass spectrometry experiments were performed using an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific) with an Acquity M-Class UPLC system (Waters) for reverse phase separations. Separations were performed on an in-house pulled-and-packed fused silica chromatography column. The fused silica has an inner diameter of 100 μm, and was packed with a C18 ReproSil-Pur 1.8-μm stationary phase to a length of 25 cm. The UPLC system was set to a flow rate of 300 nL/min, where mobile phase A was 0.2% formic acid in water and mobile phase B was 0.2% formic acid in acetonitrile. Peptides were directly injected onto the chromatography column, with a gradient of 2–45% mobile phase B, followed by a high-B wash over a total 80 min. The mass spectrometer was operated in a data-dependent mode using CID fragmentation for MS/MS spectral generation.

The resulting mass spectra were analyzed using Byonic v2.0–25 (Protein Metrics) against the UniProt mouse database concatenated with a database of commonly observed protein contaminants. Precursor mass tolerances were set to 10 ppm with fragment tolerances set to 0.25 Da for CID fragmentation. Data were validated using the standard reverse-decoy technique at a 1% false discovery rate as described previously[52], and normalized spectral

abundance factors were calculated to account for differences in protein size. To confirm ALDH2 as the photocrosslinked and fluorescently tagged target, intact mitochondria were photocrosslinked with **3**, homogenized, clicked with biotin-azide, and resolved on a two-dimensional gel as described above. The gel proteins were transferred to a PVDF membrane, which was then probed with a goat polyclonal anti-ALDH2 antibody (1:1,000 dilution; Santa Cruz Biotechnology, sc-48837), an Alexa Fluor 488-conjugated donkey polyclonal anti-goat IgG antibody (1:1,000 dilution; Thermo Fisher Scientific, A11055), and Alexa Fluor 546-conjugated streptavidin (1:5,000 dilution; Thermo Fisher Scientific, S11223). Fluorescently stained proteins were then detected using the ChemiDoc imaging system.

ALDH1B1 constructs

Wild-type ALDH1B1 was subcloned into a pCMV6 vector (Origene, PS100001) via Gibson assembly, using HindIII- and MluI-digested pCMV6 plasmid, SW480 cell cDNA as a template, and the following primers: forward, 5'-ATCGCCGGCGCGCCAGATCTCAATGCTGCGCTTCCTGGCACCCCGG-3'; reverse, 5'-TTCTGCTCGAGCGGCCGCGTACGCGAGTTCTTCTGAGGAACCTT-3'.

For transient expression of mutant ALDH1B1 proteins in mammalian cells, human *ALDH1B1* variants were created using site-directed mutagenesis on the pCMV6-ALDH1B1 plasmid and the following primers: L120M forward, 5'-CCAAGAGTCTTACGCCATGGACTTGGATGAGGT-3'; L120M reverse, 5'-ACCTCATCCAAGTCCATGGCGTAAGACTCTTGG-3'; L120V forward, 5'-CCAAGAGTCTTACGCCGTGGACTTGGATGAGGT-3'; L120V reverse, 5'-ACCTCATCCAAGTCCACGGCGTAAGACTCTTGG-3'; E124G forward, 5'-CTTACGCCTTGGACTTGGATGGGGTCATCAAGGTGTATCGG-3'; E124G reverse, 5'-CCGATACACCTTGATGACCCCATCCAAGTCCAAGGCGTAAG-3'; E124L forward, 5'-GTCTTACGCCTTGGACTTGGATCTAGTCATCAAGGTGTATCGGTAC-3'; E124L reverse, 5'-GTACCGATACACCTTGATGACTAGATCCAAGTCCAAGGCGTAAGAC-3'; V173I forward, 5'-GTGGAACTTCCCCTTGATCATGCAGGGTTGGAA-3'; V173I reverse, 5'-TTCCAACCCTGCATGATCAAGGGGAAGTTCCAC-3'; Q289F forward, 5'-CATGGAGCATGCCGTGGAGTTTTGCCACGAAGCCCTG-3'; Q289F reverse, 5'-CAGGGCTTCGTGGCAAAACTCCACGGCATGCTCCATG-3'; E292H forward, 5'-GCCGTGGAGCAGTGCCACCATGCCCTGTTCTTCAACATGG-3'; E292H reverse, 5'-CCATGTTGAAGAACAGGGCATGGTGGCACTGCTCCACGGC-3'; N457G forward, 5'-GTGGGTAAACACCTACGGCATCGTCACCTGCCACACGCC-3'; N457G reverse, 5'-GGCGTGTGGCAGGTGACGATGCCGTAGGTGTTTACCCAC-3'.

For bacterial expression of the mutant ALDH1B1 proteins, each construct was subcloned via Gibson assembly into a pET-15b expression vector containing an N-terminal hexahistidine tag (Novagen), the corresponding pCMV6-ALDH1B1 plasmid as a template, and the following primers (which were designed to remove the mitochondria-targeting sequence of ALDH1B1): pET-15b forward, 5'-GTCACCATCAAGGTTCCTCAGAAGAACTCGTGAGGATCCGGCTGCTAACAAAGCCCGAAA-3'; pET-15b reverse , 5'-GAATGGGGCTTGGGAGGGCTGCTGCCGAGGACATATGGCTGCCGCGCGGCACCA

GGC-3’; ALDH1B1 forward, 5’-GCCTGGTGCCGCGCGGCAGCCATATGTCCTCGGCAGCAGCCCTCCCAAGCCCCATTC-3’; ALDH1B1 reverse, 5’-TTTCGGGCTTTGTTAGCAGCCGGATCCTCACGAGTTCTTCTGAGGAACCTTGATGGTGAC-3’.

For lentiviral expression of *EGFP* and human wild-type *ALDH1B1*, their cDNAs were separately cloned into the pCDH-CMV-MCS-EF1-Puro vector (System Biosciences) digested with EcoRI and PacI via Gibson assembly. pCDH-CMV-GFP-EF1-Puro was generated using pcDNA3-EGFP (Addgene, plasmid 13031) as a template the following primers: forward, 5’-TCCATAGAAGATTCTAGAGCTAGCGGTGAGCAAGGGCGATGAGCTGTTC-3’; reverse, 5’-GCGGCCGCGGATCCGATTTTCACTTGTACAGCTCGTCCATGCCG-3’. pCDH-CMV-ALDH1B1-EF1-Puro was generated using pCMV6-ALDH1B1 as a template and the following primers: forward, 5’-TCCATAGAAGATTCTAGAGCTAGCGATGCTGCGCTTCCTGGCACCC-3’; reverse, 5’-GCGGCCGCGGATCCGATTTTTAATTACGAGTTCTTCTGAGGAACC-3’.

For lentiviral expression of the mutant ALDH1B1 proteins, the pCDH-CMV-GFP-EF1-Puro vector was digested with EcoRI-HI and PacI to remove the GFP cDNA. Each ALDH1B1 mutant sequence was then subcloned into the resulting pCDH derivative via Gibson Assembly using the corresponding pCMV6-ALDH1B1 plasmid as a template and the following primers (which were designed to remove the mitochondria-targeting sequence of ALDH1B1): forward, 5’-TCCATAGAAGATTCTAGAGCTAGCGATGCTGCGCTTCCTGGCACCC-3’; reverse, 5’-GCGGCCGCGGATCCGATTTTTAATTACGAGTTCTTCTGAGGAACC-3’.

All constructs were confirmed using Sanger sequencing.

ALDH expression and purification

Cloning, expression, and purification of human recombinant ALDH1A1, ALDH1A3, ALDH2, ALDH3A1, ALDH3A2, ALDH4A1, ALDH5A1 and ALDH7A1 isozymes were conducted as described previously[53]. Human ALDH1A2 (Origene, SC109995) and ALDH1B1 (T. Hurley, Indiana University) plasmids were cloned into the pET-15b expression vector containing an N-terminal hexahistidine tag. The mitochondria-targeting sequence of ALDH1B1 was also removed. For protein expression without chaperones, the pET-15b-based plasmids were transformed into competent BL21 *E. coli* cells, which were subsequently grown in 1 L of Luria-Bertani medium containing 100 μg/mL ampicillin at 37 °C. Once the OD_{600} value reached 0.6–0.8, the culture was transferred to 16 °C and 0.5 mM isopropyl β-d-1-thiogalactopyranoside was introduced to induce protein expression overnight. The bacteria were then pelleted by centrifugation (3,000 x *g*, 4 °C for 30 min) and harvested. 40 mL BugBuster Master Mix buffer (MilliporeSigma) containing a protease inhibitor cocktail (cOmplete; Roche) was added, and the cells were lysed by mechanically rolling the suspension at room temperature for 25 min. The lysate was centrifuged (16,000 x *g*, 4 °C for 30 min), and the soluble fraction was collected and mixed with 2 mL of Ni-NTA Superflow resin (Qiagen). After the mixture was allowed to rock at 4 °C for 1 h, it was used

to create a 10-mL gravity column, which was sequentially treated with 40 mL of wash buffer 1 (20 mM sodium phosphate, 500 mM NaCl, 1 mM DTT, 20 mM imidazole, 5% glycerol, pH 7.4) and 25 mL of wash buffer 2 (20 mM sodium phosphate, 500 mM NaCl, 1 mM DTT, 40 mM imidazole, 5% glycerol, pH 7.4). The hexahistidine-tagged protein was then recovered by treated the column with 5 mL elution buffer (20 mM sodium phosphate, 500 mM NaCl, 1 mM DTT, 400 mM imidazole, 5% glycerol, pH 7.4) and collecting 500-μL fractions. The eluent was loaded onto a 15-mL 30-kDa Amicon Ultra centrifugation unit (MilliporeSigma), and the protein sample was repeatedly concentrated via centrifugation (5,000 x *g*, 22 °C for 20 min) and diluted with storage buffer (20 mM sodium phosphate, 500 mM NaCl, 1 mM DTT, 5% glycerol, pH 7.4) until less than 1 mM imidazole remained. The protein sample was then diluted with an equal volume of glycerol and the resulting concentration was determined (600nm Protein Assay Reagent; Thermo Fisher Scientific). Each samples was analyzed by SDS-PAGE and stored at −80 °C.

For wild-type and mutant ALDH1B1 protein co-expression with chaperones, pET-15b-ALDH1B1 and GroEL/GroES (Addgene, pBB541) plasmids were co-transformed into competent BL21 *E. coli* cells, which were subsequently grown in 1 L of Luria-Bertani medium containing 100 μg/mL ampicillin and 50 μg/mL spectinomycin at 37 °C. The hexahistidine-tagged ALDH1B1 protein was then expressed and purified as described above.

Enzyme kinetics assays

To evaluate compound activities against ALDH isoforms, reaction mixtures consisting of enzyme assay buffer (100 mM sodium phosphate, 1 mM $MgCl_2$, pH 8.0) and 0.1% DMSO, 1 mM β-mercaptoethanol, 1 mM NAD^+, and the bacterially expressed protein were prepared in white opaque 96-well assay plates (100 μL/well, flat-bottom and non-treated; Corning Costar). 5 μg/mL ALDH enzyme was used for single-dose (1, 10, 50 μM) compound test, and 1 μg/mL ALDH enzyme was used for IC50s determination. Acetaldehyde was then added to each well to achieve an initial substrate concentration of 1 mM, and the resulting enzymatic activity was measured based on NADH fluorescence. The fluorescence signal was read over the course of 10 min on a SpectraMax M2e microplate reader (340-nm excitation and 460-nm emission; Molecular Devices) operated by SoftMax Pro software. Steady-state kinetic measurements of ALDH1B1 inhibitors were also conducted with the general protocol described above, using 5 μg/mL ALDH1B1 and varying concentrations of acetaldehyde, NAD+, and the small-molecule antagonist.

To determine the inhibitor resistance of mutant ALDH isoforms, reaction mixtures consisting of enzyme assay buffer (100 mM sodium phosphate, 1 mM $MgCl_2$, pH 8.0), 5% glycerol, 0.1% DMSO, 1 mM β-mercaptoethanol, 1 mM propionaldehyde, and 2 μg/mL of the bacterially expressed ALDH1B1 were prepared in white opaque 96-well assay plates (100 μL/well, flat-bottom and non-treated; Corning Costar). Each mixture also contained varying concentrations of IGUANA-1. NAD^+ was then added to each well to achieve an initial concentration of 1 mM, and the resulting enzymatic activity was measured by monitoring NADH fluorescence for 30 min as described above.

All compound activities were normalized to a DMSO control, and dose-response inhibition curves were fitted with four-parameter non-linear regression using Prism (GraphPad) software.

X-ray crystallography

Stock solutions of the ALDH1B1 protein were buffer-exchanged with the crystallization buffer (100 mM ACES, 10 mM NAD^+, 10 μM DTT, pH 7.0) and concentrated to 20 mg/mL using a 30-kDa MWCO Amicon centrifugal filter (MilliporeSigma) at 4 °C. ALDH1B1 inhibitor was subsequently added to a final concentration of 1 mM, and the protein solution (usually 0.15–0.3 μL) was then mixed with an equal volume of various crystallization buffers from commercial sources. Crystals were grown using sitting drop vapor diffusion in a 6 °C or 12 °C incubator from crystallization experiments set-up using an Oryx8 Nanodrop dispensing robot (Douglas Instruments). Within one to two weeks, crystals were harvested and cryocooled under liquid N_2 stream. In general, crystals harvested even from a single crystallization condition showed a wide variation in X-ray diffracting power, and therefore a large number were screened for initial data quality assessment. The best candidates were selected and stored for further data collection, and in some cases, X-ray diffracting raster sampling was used to locate the best diffracting region on the crystal. Then, a full data set collection was performed. Data collections to a minimum Bragg spacing of 2.12 Å (ALDH1B1/NAD^+/**2**), 2.30 Å (ALDH1B1/NAD^+/IGUANA-4), and of 2.68 Å (ALDH1B1/NAD^+) were performed at 100 K using Stanford Synchrotron Radiation Lightsource (SSRL) beamlines BL9–2, BL12–1 and BL12–2 (SLAC National Accelerator Laboratory).

All ALDH1B1 diffraction-quality crystals were obtained from crystallization or co-crystallization (in case of ligand added to the protein solution) from a precipitant solution composed of 10% polyethyleneglycol 4000 and 20% glycerol with different combinations of buffers/additives: 0.1 M Tris-Bicine, pH 8.5, and a mixture of alcohols (0.02 M of each 1,6-hexanediol, 1-butanol, 1,2-propanediol, 2-propanol, 1,4-butanediol, and 1,3-propanediol) for ALDH1B1/NAD^+; or 0.1 M Tris-Bicine, pH 8.5, and a mixture of ethylene-glycols (0.03 M of each di-, tri-, tetra-, and penta-ethylene glycol) for ALDH1B1/NAD^+/**2** and ALDH1B1/NAD^+/IGUANA-4. Structures were solved by the molecular replacement method using the polypeptide chain of apo ALDH2 (PDB ID: 3N80) as the search model, and the search was carried over the enantiomorphic pair to fix handedness. Each of the crystals belonged to the trigonal space group P 3_2 2 1 and contained two polypeptide chains per asymmetry unit. The full polypeptide chain, residues Leu8-Ser500, was unambiguously traced in the electron density maps. Extra electron density was evident after structure solution and accounted for one NAD^+ cofactor molecule and one inhibitor molecule per polypeptide chain. The NAD^+ cofactor, **2**, and IGUANA-4 were modeled at full occupancy in both polypeptide chains A and B. Extra density accounting for the organic ligands in the crystallization buffer were detected and included in the final model. Solvent water molecules were first assigned based on their hydrogen bonding properties; in later stages of refinement, further water molecules were automatically added. Refinement progressed to convergence and reached an excellent agreement between the model and the experimental data. Supplementary Table 1 presents data collection, refinement, and structure quality check parameters. Data was integrated,

reduced, scaled, and analyzed with different computing modules within the CCP4 suite[54]. Graphic renderings were prepared with PyMOL software[55].

Generation of *ALDH1B1* knockout cell lines

To target the *ALDH1B1* locus by CRISPR/Cas9 mutagenesis, the following gRNAs were evaluated (PAM sequences in bold): gRNA-1, 5'-CGGTGGTA**GGG**TTGACCGTCGGG-3'; gRNA 2, 5'-GGAGCGGGATCGAGTCTACT**TGG**-3'; gRNA-3, 5'-CGAAACGCTCTCCGCCACAG**AGG**-3'. Each gRNA was cloned into the lentiCRISPRv2 vector (Addgene, plasmid 98290) by BsmBI digestion and DNA ligation, and lentivirus was generated using procedures as previously described[56]. Briefly, the lentiCRISPRv2-derived gRNA plasmids and a 3rd generation lentiviral system (Addgene, plasmids 12251, 12253, and 12259) were co-transfected into HEK-293T cells. The resulting lentivirus was collected from the media and used to infect SW480 cells. Stable populations were selected by puromycin treatment (2.0 μg/mL), and genomic DNA PCR and DNA TIDE sequencing analysis was performed to assess the frequency of targeted mutations. The PCR primers used for these studies were: gRNA-1 forward, 5'-CTTAGCCTCCAGGGCAGGAC-3'; gRNA-1 reverse, 5'-GGAAAGCCTGCCTCCTTGAT-3'; gRNA-2 forward, 5'-CCCAGACATCCCCTACAACC-3'; gRNA-2 reverse, 5'-GCCATACCCCGTGATGATGT-3'; gRNA-3 forward, 5'-GGCGATTCCAACCTCAAGAGA-3'; gRNA-3 reverse: 5'-TTAAGCCCATCCTCACCCAG-3'. The reverse primers were also used for sequencing.

To generate *ALDH1B1*$^{-/-}$ colon cancer lines, gRNA-1 and gRNA-2 were separately cloned into the pX458 vector (Addgene, plasmid 48138) by BbsI digestion and DNA ligation. The resulting constructs were co-transfected into SW480 and HCT116 cells using FuGENE HD transfection reagent (Promega) according to the manufacturer's guidelines. The cells were then cultured for 48 h, and the cells were sorted using an Aria II cell sorter (BD Biosciences). The cells were gated using FSC and SSC to exclude debris, FSC and trigger pulse width to identify single cells, and propidium iodide staining and FSC to select live cells (Supplementary Fig. 20). Single cells from the top 5% EGFP-positive population were then sorted into 96-well plates. The single clones were cultured and characterized by genomic DNA PCR and sequencing, using the following primers: forward, 5'-CTTAGCCTCCAGGGCAGGAC-3; reverse, 5'-GCCATACCCCGTGATGATGT-3'.

To conduct rescue experiments with *ALDH1B1*$^{-/-}$ cells, pCDH-CMV-EF1-Puro vectors encoding EGFP, wild-type ALDH1B1, or ALDH1B1 mutants were separately packaged in HEK-293T cells using the 3rd generation lentiviral system described above, and the resulting lentiviruses were used to infect *ALDH1B1*$^{+/+}$ (clone 3) and/or *ALDH1B1*$^{-/-}$ (clone 2) SW480 cells. Stable lines were then generated by puromycin selection (2.0 μg/mL).

Immunoblot analyses

To analyze lysates derived from SW480 and HCT116 cells, adherent cultures were grown to 70–80% confluence, detached from the plates with trypsin, and pelleted by centrifugation at 480 x *g* for 5 min at 4 °C. The cell pellets were washed once with PBS and lysed in cold RIPA buffer (Thermo Fisher Scientific) containing with protease (cOmplete; Roche)

and phosphatase (PhosSTOP; Roche) inhibitor. Lysates were kept on ice for 15 min with intermittent vortexing, and then clarified by centrifugation at 16,000 x *g* for 15 min at 4 °C. The resulting supernatant was analyzed to determine total protein levels (600nm Protein Assay Reagent; Thermo Fisher Scientific), and the final protein concentration in each sample was adjusted to 1.0 μg/μL, mixed with 6x SDS-SAGE loading buffer (240 mM Tris-HCl, 6% SDS, 0.3 M DTT, 30% glycerol, 0018% bromophenol blue, pH 6.8). The samples were then resolved on 4–12% Criterion Bis-Tris protein gels (Bio-Rad) at 130 V for 1.5 h at room temperature.

The gel proteins were next transferred to PVDF membranes using the Trans-Blot Turbo Transfer System (Bio-Rad), and the membranes were blocked with 5% non-fat dry milk in 1x Tris-buffered saline containing 0.1% Tween 20 (TBST) for 1 h at room temperature. The blots were then probed with the primary antibody in the same blocking buffer overnight at 4 °C. The membranes were washed 4 × 5 min with TBST and then incubated with horseradish peroxidase (HRP)-conjugated secondary antibody in blocking buffer for 2 h at room temperature. After the membranes were washed 4 × 5 min TBST, the secondary antibodies were detected by chemiluminescence using SuperSignal West Dura Extended Duration Substrate (Thermo Fisher Scientific) and a Gel Doc XR imaging system (Bio-Rad). stripping buffer (Thermo Fisher Scientific, 21059) was used to remove antibodies when the membranes needed to be re-probed with another primary antibody. Antibodies used for these experiments include: mouse monoclonal anti-ALDH1B1 (1:500 dilution; Santa Cruz Biotechnology, sc-393583); rabbit polyclonal ARCN1 antibody(1:500 dilution, Proteintech, 23843–1-AP); rabbit polyclonal NIPSNAP2 anbitody (1:500 dilution, Abcam, ab153833); rabbit polyclonal anti-KPNB1 antibody (1:1,000 dilution; Santa Cruz Biotechnology, sc-11367); HRP-conjugated sheep polyclonal anti-mouse IgG (1:1,000; GE Healthcare, NA931–1ML); HRP-conjugated donkey polyclonal anti-rabbit IgG (1:1000 dilution; GE Healthcare, NA934–1ML).

Matrigel spheroid cultures

To prepare Matrigel-coated culture plates, 96-well and 24-well plates were placed on ice and coated with 25 μL or 300 μL, respectively, of LDEV-free Matrigel matrix (Corning) that had been thawed overnight at 4 °C. The plates were then placed in a 37 °C incubator for 30 min to allow the Matrigel to solidify. SW480 or HCT116 cells were suspended in spheroid culture medium (RPMI-1640 containing 20 ng/mL recombinant human epidermal growth factor (Abcam, ab9697), 20 ng/mL recombinant human fibroblast growth factor (BioVision, 4037), 5 μg/mL heparin (Sigma), 1x N-2 supplement (Thermo Fisher Scientific), 100 U/mL penicillin, and 100 μg/mL streptomycin) and then seeded into the 96-well and 24-well Matrigel-coated plates at densities of 1,000 or 5,000 cells/well and total well volumes of 200 μL or 1 mL, respectively. Compounds were diluted in spheroid culture medium and added to the wells, and the spheroids were grown at 37 °C with 5% CO_2 for 7 days.

To quantify spheroid growth at the 7-day endpoint, unstained spheroids were imaged with phase-contrast illumination using a Keyence BZ-X710 microscope. Some cultures were also stained with crystal violet to enhance brightfield imaging. Briefly, the Matrigel spheroid cultures were treated with PBS containing 0.5% crystal violet and 5% methanol for 15

min at 37 °C and then gently washed 4 × 5 min with PBS. Spheroid sizes were quantified with ImageJ (Fiji) software using masks created with the default “Threshold” function and foreground/background adjustments to optimize masking fidelity. The full field of view was then selected for “Analyze Particles” analysis, providing numbered list of each spheroid and its area in the micrograph. In the few cases that two adjacent spheroids were treated as one object by the image processing software, their measurements were manually corrected using the “Freehand Selections” tool. To exclude background and debris from dead cells, spheroids with areas less than 500 μm^2 were excluded from the image analyses.

Colon cancer cell viability assays

SW480 or HCT116 cells were seeded into regular or Matrigel-coated 96-well plates at a density of 1000 cells/well. The cells were cultured in the corresponding adherent or spheroid culture medium (200 μL/well) containing the designated compounds or an equivalent amount of DMSO vehicle. Each culture was maintained for 7 days, and then cell viabilities were measured using a CellTiter-Glo 3D Cell Viability Assay kit (Promega). Briefly, 100 μL of the culture medium was removed from each well and replaced with 100 μL 3D Cell Assay reagent. The cells in each well were next homogenized by vigorous pipetting, and the plate was incubated for 10 min at room temperature to stabilize the chemiluminescent signal. A 150-μL aliquot from each well was then transferred into a white opaque 96-well assay plate (flat-bottom and non-treated; Corning Costar) and luminescence intensities were measured with a Veritas microplate luminometer (Turner BioSystems). All compound activities were normalized to a DMSO control, and dose-response inhibition curves were fitted with four-parameter non-linear regression using Prism software.

Immunofluorescence microscopy

ALDH1A3$^{-/-}$ A375 cells were seeded into a 6-well plate at a density of 6×10^5 cells/well and cultured for 24 h. The cells were then transfected with either pCMV6 or pCMV6-ALDH1B1 using FuGENE HD (Promega) at a 3:1 DNA:FuGENE HD ratio according to the manufacturer’s protocols. The cells were cultured for another 24 h, removed from the plate by trypsinization, and re-seeded at a density of 1.2×10^5 cells/well into a 24-well plate containing poly-D-lysine coated glass coverslips. After another 24 h of cell growth, the culture media was removed, and the cells were incubated with fresh media containing 500 nM MitoTracker Deep Red (Invitrogen, M22426) for 30 min at 37 °C and subsequently fixed with PBS containing 4% PFA for 10 min at room temperature. After permeabilization with PBS containing 0.3% Triton X-100, the cells were incubated with anti-ALDH1B1 antibody (1:100 dilution, Santa Cruz Biotechnology, sc-393583) overnight at 4 °C. Cells were then incubated with Alexa Fluor 488-conjugated goat polyclonal anti-mouse IgG (Invitrogen, A32723) and DAPI (Thermo Fisher Scientific, P36931) for 1 h at room temperature in the dark. Coverslips were imaged on a Zeiss LSM 800 confocal Microscope using ZenBlue software.

ALDEFLUOR assays

ALDH1A3$^{-/-}$ A375 cells were seeded into a 6-well plate at a density of 5×10^5 cells/well and cultured for 24 h. The cells were then transfected with either pCMV6, pCMV6-GFP (Addgene, plasmid 11153), or pCMV6-ALDH1B1. After 24 h, the media was replaced

and transfection efficiencies were estimated by the EGFP-positive cells in the pCMV6-GFP-transfected well. The cells were cultured for another 24 h and their ALDH activities were quantified using the ALDEFLUOR assay. Briefly, the cells were isolated from the plate with trypsin and resuspended in ALDEFLUOR assay buffer (STEMCELL Technologies) at a concentration of 4×10^5 cells/mL. Each ALDEFLUOR assay used 0.5 mL of this cell suspension, and the cells were treated with individual compounds for 15 min at 37 °C and 5% CO_2. The cell suspensions were next treated with 2.5 μL ALDEFLUOR reagent, incubated for an additional 30 min at 37 °C and 5% CO_2, collected by centrifugation (450 x *g*) for 5 min at room temperature, resuspended in 200 μL assay buffer, and analyzed with an LSR II flow cytometer (BD Biosciences). The cells were gated using FSC-A and SSC-A to exclude debris, FSC-A and FSC-H to identify single cells, and DAPI and SSC-A to select live cells (Supplementary Fig. 21). Cellular ALDH1B1 activities were then assessed using the ALDEFLUOR signal (Ex: 488 nm, Em: 515/20 or 525/50 nm) and SSC-A.

For mutant ALDH1B1 analyses, the *ALDH1A3*$^{-/-}$ A375 cells were seeded and cultured in 6-well plates as described above and then transfected with wild-type or mutant pCMV6-ALDH1B1 using FuGENE HD (Promega) at a 3:1 DNA:FuGENE HD ratio according to the manufacturer's protocols. The cells were cultured for another 48 h, and their ALDH activities were quantified using the ALDEFLUOR assay. The cells were trypsinized from the plates, resuspended in 2.0 mL ALDEFLUOR assay buffer at a concentration of 1×10^6 cells/mL, and then treated with 10 μL ALDEFLUOR reagent. Within 10 s, 500-μL aliquots of this cell suspension was transferred to 1.5-mL centrifuge tubes containing 5 μL of either DMSO, 1.5 mM DEAB in EtOH, 100 μM IGUANA-1 in DMSO, or 10 μM IGUANA-1 in DMSO. The tubes were then vortexed and placed in a 37 °C and 5% CO_2 incubator for 30 min. The cells were pelleted by centrifugation (1000 x *g*) for 5 min at room temperature, resuspended in 200 μL assay buffer, and analyzed with an LSR II (BD Biosciences) or FACS Symphony A5 (BD Biosciences) flow cytometer using the protocols described above. For each transfected ALDH1B1 construct, the DEAB-treated cells were used to establish gates for quantifying populations with exogenous ALDH activity, and these boundaries were applied to the other three samples. The gates assumed a normal distribution of the bulk, untransfected cell population, as some ALDH1B1 variants were not fully inhibited by DEAB.

Mitochondrial respiration assays

SW480 cells were seeded into a XF96 microplate (Agilent) at a density of 2×10^4 cells/well and cultured for 24 h. The cells were then incubated for 1 h at 37 °C in a non-CO_2 incubator, and the plate was then loaded into a 96-well Seahorse XF analyzer (Agilent) and sequentially injected with the indicated compounds, 1 μM oligomycin A, 1 μM FCCP, or 0.5 μM rotenone/antimycin according to the Agilent Mito Stress protocol. The oxygen consumption rates (OCRs, pmol/min) for each well were measured every 5 min for a 15-min period. Basal OCR values were normalized as 100% for data analyses.

Cellular thermal shift assays

CETSA was conducted as previously described[57]. Briefly, SW480 cells were cultured in six 15-cm dishes and were then harvested by trypsinization upon ~ 80% confluency.

Trypsinized cells were combined and then equally split into two 50-mL falcon tubes. Cell pellets were obtained by centrifugation at 480 x *g* for 5 min at room temperature and were resuspended with 20 mL culture medium with either 10 μM IGUANA-1 or DMSO. The live SW480 cells were cultured in the incubator for one hour, and cell pellets were collected by centrifugation at 480 x *g* for 5 min at room temperature. The cells were washed with 10 mL PBS, then pelleted again by centrifugation, and resuspended in 1 mL PBS containing protease inhibitors (cOmplete; Roche). The resulting cell suspension was distributed among eight PCR tubes as 100-μL aliquots, and the tubes were then heated to eight different temperatures (45.0, 46.4, 48.8, 52.6, 57.1, 60.9, 63.4 and 65.0 °C) in a Bio-Rad PCR machine. Heated cells were then incubated at room temperature for 3 min and snap-frozen in liquid nitrogen. Cell lysis was obtained by freezing-thawing the cells twice using liquid nitrogen and incubation at room temperature. The resulting cell lysates were briefly vortexed and transferred to a 1.5 mL Eppendorf tube for centrifugation at 20,000 g for 20 min at 4°C. A 90-μL aliquot of each supernatant was carefully transferred into another 1.5 mL Eppendorf tube, and the samples were analyzed by immunoblotting as described above. A TGX Stain-Free precast gel (Bio-Rad) was also used to detect total protein in each lane.

Thermal proteome profiling

TPP studies were conducted as previously described[29]. Briefly, soluble proteins were isolated SW480 cells as described for the CETSA experiments, except six 15-cm dishes were used, the cells were treated with 2 μM IGUANA-1 or DMSO, and ten different temperatures were evaluated (37.3, 40.9, 45.1, 48.2, 50.5, 53.4, 55.9, 59,5, 64.0 and 67.9 °C). The final 90-μL aliquot of each supernatant was mixed with 4x volume of pre-chilled (−80 °C) acetone and incubated at −80 °C overnight. Samples were centrifuged at 9,400 x *g* for 10 min to separate the supernatant acetone from the precipitated protein pellet. Air-dried protein pellets were reconstituted in 100 mM triethylammonium bicarbonate buffer, sonicated, and vortexed to solubilize proteins. The samples were then reduced with 10 mM DTT at 55 °C for 30 min followed by alkylation with 30 mM acrylamide for 30 min at room temperature. Digestion was performed using 1 μg of Trypsin/LysC protease (Promega) with overnight incubation at 37 °C. 10-μL aliquots of each sample were used for peptide quantification with a Pierce Quantitative Fluorometric Peptide Assay kit (Thermo Fisher Scientific).

Following peptide quantification, 10-μg aliquots of each peptide sample were labeled with a TMT10plex Isobaric Label Reagent set (Thermo Fisher Scientific). Post-labelling, the samples were quenched with 5% hydroxylamine, combined into one series for a given group, and de-salted with C18 Monospin Reverse Phase Columns (GL Sciences). The desalted samples were then dried and reconstituted in 2% aqueous acetonitrile prior to analysis.

Mass spectrometry experiments were performed using an Orbitrap Eclipse Tribrid mass spectrometer (Thermo Fisher Scientific) attached to an Acquity M-Class UPLC system (Waters). The UPLC system was set to a flow rate of 300 nL/min, where mobile phase A was 0.2% formic acid in water and mobile phase B was 0.2% formic acid in acetonitrile. The analytical column was prepared in-house with an I.D. of 100 μm pulled to a nanospray

emitter using a P2000 laser puller (Sutter Instrument). The column was packed with NanoLCMS Solutions 1.9-μm C18 stationary phase to a length of approximately 25 cm. Labeled peptides were directly injected into the column with a gradient of 2–45% mobile phase B, followed by a high-B wash over a total of 180 min. The mass spectrometer was operated in a data-dependent mode using MS3 HCD and CID fragmentation for spectral generation.

The collected mass spectra were analyzed using Proteome Discoverer 2.2 (Thermo Fisher Scientific) against the UniProt human database, concatenated with common contaminant proteins (e.g., human keratins). Data were searched using Byonic v4.1.5 (Protein Metrics) as the peptide identification node. The precursor ion tolerance was set to 12 ppm. The fragment ion tolerance was set to 0.4 Da and 12 ppm for MS2 and MS3 scans, respectively. Trypsin was set as the enzyme, and up to two missed cleavages were allowed. Cysteine modified with propionamide was set as a fixed modification in the search. Variable modifications include TMT10plex on N-termini and lysine, oxidation on methionine, and deamidation on asparagine and glutamine. Data were validated using the standard reverse-decoy technique at a 1% false discovery rate.

Protein spectral counts for each experimental condition were normalized to account for the volume of the TMT labelling reaction and then normalized to the lowest temperature. The data set was further processed to exclude indistinguishable splicing isomers, common contaminant proteins (e.g., bovine serum albumin), and proteins that were not detected at the lowest temperature sample. Analysis was performed with the TPP R package described previously[29].

SYTOX Green staining

SW480 cells were seeded into an 8-well chamber slide precoated with 50 μL Matrigel (Corning) at a density of 5,000 cells/well. Spheroids were allowed to grow for 4 days in the spheroid culture medium and subsequently treated with either 2 μM IGUANA-1 or DMSO for 3 days. The spheroids were then incubated with 10 μM SYTOX Green (ThermoFisher, S7020) overnight to stain the nuclei of dead cells, washed briefly, with PBS and slides, and imaged on a Zeiss LSM 800 confocal microscope.

Cytotoxicity assays

HEK-293T cells were seeded into a 96-well plate at a density of 5×10^3 cells/well and treated with varying concentrations of selected compounds or an equivalent amount of DMSO vehicle. The cells were cultured for 4–5 days without a media change, and after the DMSO-treated cells approached 100% confluence, the cell viabilities were measured using a CellTiter-Glo 3D Cell Viability Assay kit (Promega) and a Veritas microplate luminometer (Turner BioSystems) as described above. All compound activities were normalized to a DMSO control, and dose-response inhibition curves were fitted with four-parameter non-linear regression using Prism software.

Hedgehog signaling assays

NIH-3T3 cells stably transfected with Gli-dependent firefly luciferase reporter (Shh-LIGHT2 cells[49]) were seeded into 96-well plates at a density of 2×10^4 cells/well and cultured for 24 h. The cells were then cultured in DMEM containing 0.5% CS, 100 U/mL penicillin, 100 μg/mL streptomycin, and 200 nM Smoothened agonist (SAG)[58] for 30 h. CellTiter AQueous One (20 μL/well; Promega) was added to the cells and incubated for 30 min at 37 °C, and cell viabilities were measured as the 490-nm absorbance of the culture media using a Spectramax M2e microplate reader (Molecular Devices). The cells were then washed with PBS and incubated with Bright-Glo reagent (100 μL/well; Promega) for 5 min at room temperature. The resulting chemiluminescence was measured on a Veritas microplate luminometer (Turner BioSystems) and normalized to the CellTiter signal. The normalized luciferase activities for each compound treatment regimen were normalized further to the DMSO control, and dose-responsive curves were were fitted with four-parameter non-linear regression using Prism software.

Generation of 5-FU-resistant SW480 cells

The parental SW480 cells were cultured in RPMI-1640 medium containing 10% FBS, 100 U/mL penicillin, 100 μg/mL streptomycin, and 1 μM 5-FU for two weeks. The 5-FU concentration in the culture medium was then increased to 5 μM and after another two weeks raised further to 10 μM. The resulting 5-FU-resistant SW480 cells were maintained in culture medium with 10 μM 5-FU.

Patient-derived organoid cultures

All human tissue procurement activities involved in development of the PDOs from wild-type colon or colorectal adenocarcinomas were approved and governed by Stanford University School of Medicine's Institutional Review Board (Protocol Number 28908). Individual informed consent for research was obtained in writing prior to tissue procurement from all patients who contributed to the development of the organoids used in this study. Resected material or core biopsy samples were stored in HyperThermosol FRS solution (StemCell Technologies) for no more than 24 hours prior to organoid generation. PDO culture lines from colorectal tumor samples were generated initially in an air-liquid interface culture system, as previously described[32]. Briefly, tissue samples were minced, washed, and re-suspended in Collagen Type I-A matrix (FUJIFILM Wako Laboratory Chemicals). Ongoing *in vitro* culture within the collagen matrix was maintained using WENR media. Upon PDO generation from source tissue, collagen matrix and organoids were digested using Collagenase Type IV (Worthington Biochemical) and TryPLE Express (Gibco) in sequence. Mutation profiling was performed by exome and/or targeted PCR sequencing and confirmed to match the original tumor. The resulting pool of partially-dissociated organoids were re-suspended in Cultrex Reduced Growth Factor Basement Membrane Matrix, Type 2 (BME-2) (R&D Systems) for continued culture in submerged BME-2 in WENR media[33]. Organoid lines from wild-type colon were generated in submerged BME-2 in WENR media as previously described[33].

To evaluate inhibitor responses in tumor and wild-type colon PDOs, near-confluent organoids were harvested from solidified BME-2 matrix, dissociated using TryPLE Express

supplemented with 10 μM Y-27632 dihydrochloride (Peprotech) in conjunction with gentle mechanical disruption at 37 °C, and diluted to 1,000 cells/μL with fresh BME-2 in black, clear-bottomed, 96-well plates (Corning). 10-μL aliquots of the matrix-cell suspension were plated as domes on the floor of each well and allowed to solidify in a 37 °C incubator for 10–15 min. 150 μL/well WENR media supplemented with 10 μM Y-27632 dihydrochloride (Peprotech) and 2.5 μM CHIR 99021 (R&D Systems) was then added, and organoids were allowed to grow unperturbed for 3 to 5 days. Following this initial growth period, varying concentrations of IGUANA-1 were applied in WENR media. After a 7-day incubation, all media was withdrawn from the wells and replaced with a 10% dilution of Alamar Blue Cell Viability Reagent (Invitrogen) in WENR media. After 4 h (tumor PDOs) or 6 h (wild-type PDOs) of incubation in a dark, 37 °C incubator, the effect of IGUANA-1 on the PDOs was quantified, using resazurin metabolism as a proxy for cell viability. The raw fluorescence values were measured using a microplate reader (TECAN; Ex: 530 nm, Em: 590 nm), and normalized to that of a DMSO control. Values for wells with collapsed or otherwise compromised matrix domes were excluded from subsequent analyses.

Transcriptome profiling

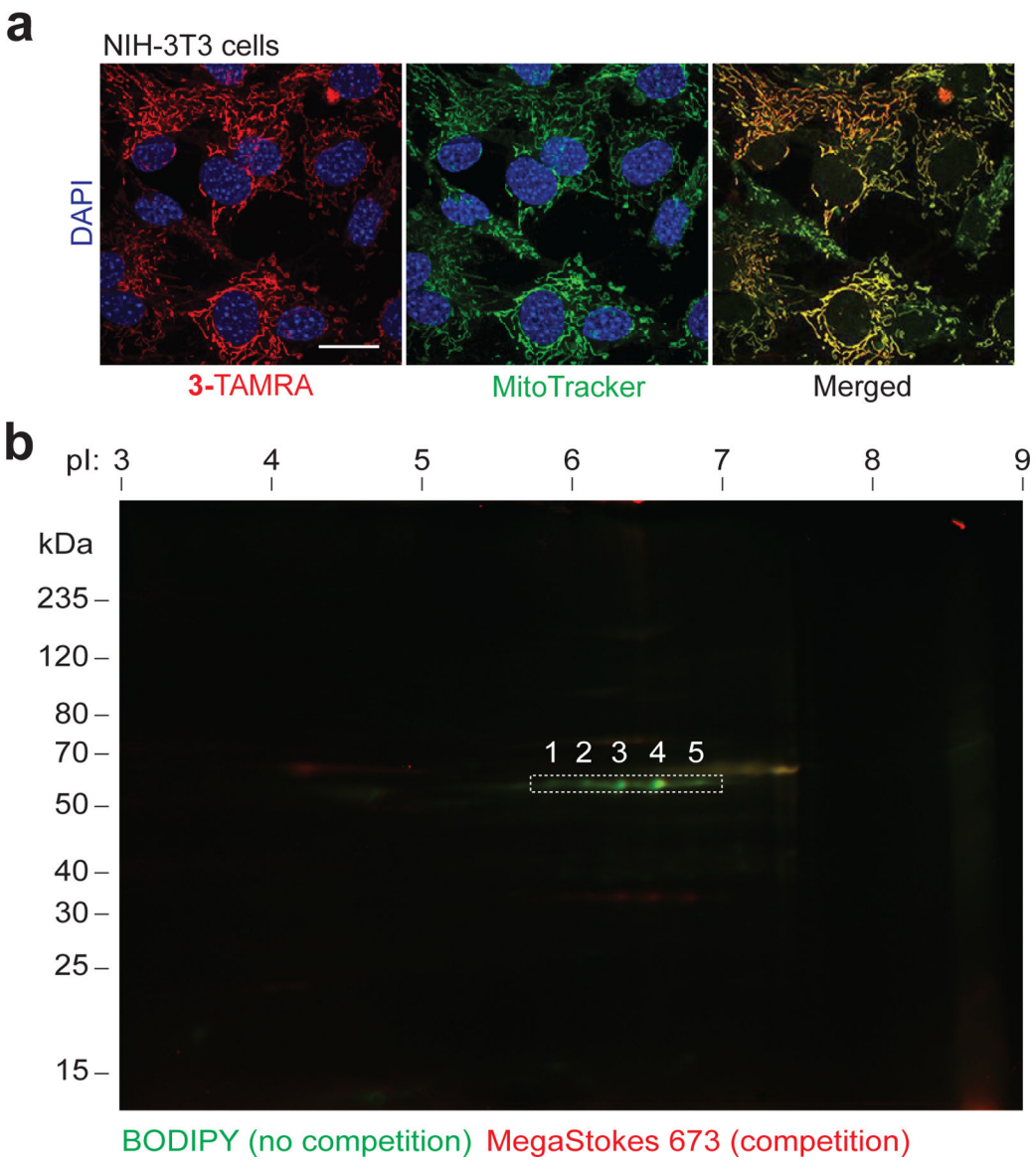
RNA-seq analysis of NIH-3T3 cells was conducted as described previously[59]. RNA-seq analyses of parental SW480 cells and an *ALDH1B1*$^{-/-}$ clone were conducted as three biological replicates as follows. To identify gene expression changes associated with chemical loss of ALDH1B1 function, parental SW480 cells were seeded into 6-well plates at density of 0.8×10^3 cells/well and cultured for 24 h. The cells were then treated with either 2 μM IGUANA-1 or an equivalent amount of DMSO vehicle for 24 h. To identify gene expression changes associated with genetic loss of ALDH1B1 function, *ALDH1B1*$^{-/-}$ SW480 cells (clone 2) that were lentivirally transduced with either *EGFP* or exogenous *ALDH1B1* were seeded into 6-well plates as described above and cultured for 48 h. Total mRNA was extracted from each sample using an RNeasy Mini kit (Qiagen) according to the manufacturer's protocols, and RNA libraries were constructed using NEBNext Ultra II RNA Library Prep Kit for Illumina (New England Biolabs). For each sample, approximately 40 million paired-end reads were obtained, and raw data of FASTQ were processed through fastp. Paired-end clean reads were aligned to the reference genome using the Spliced Transcripts Alignment to a Reference (STAR) software, and FeatureCounts was used to count the read numbers mapped of each gene. Differential expression analysis was performed using DESeq2 R package. The resulting *P* values were adjusted using the Benjamini and Hochberg's approach for controlling the False Discovery Rate[60].

Statistics and reproducibility

The experimental results shown in Fig. 1c–d, Fig. 2a–b, Extended Data Fig. 2b, Extended Data Fig. 3b, Extended Data Fig. 4a–b, Extended Data Fig. 5a–b, Extended Data Fig. 6b, Supplementary Fig. 1b, and Supplementary Fig. 2a–c were independently repeated at least three times with similar results. The results shown in Fig. 1f–h, Fig. 3c–d, Fig. 4a–d, Fig. 5c, Fig. 6a, Extended Data Fig. 2d, Extended Data Fig. 3f, Extended Data Fig. 6a, Extended Data Fig. 7a,c, Supplementary Fig. 10 and Supplementary Fig. 17a were independently repeated two times with similar results. Representative graphs, gels, and micrographs for these experiments are presented. Descriptions of the error bars and number of replicates are

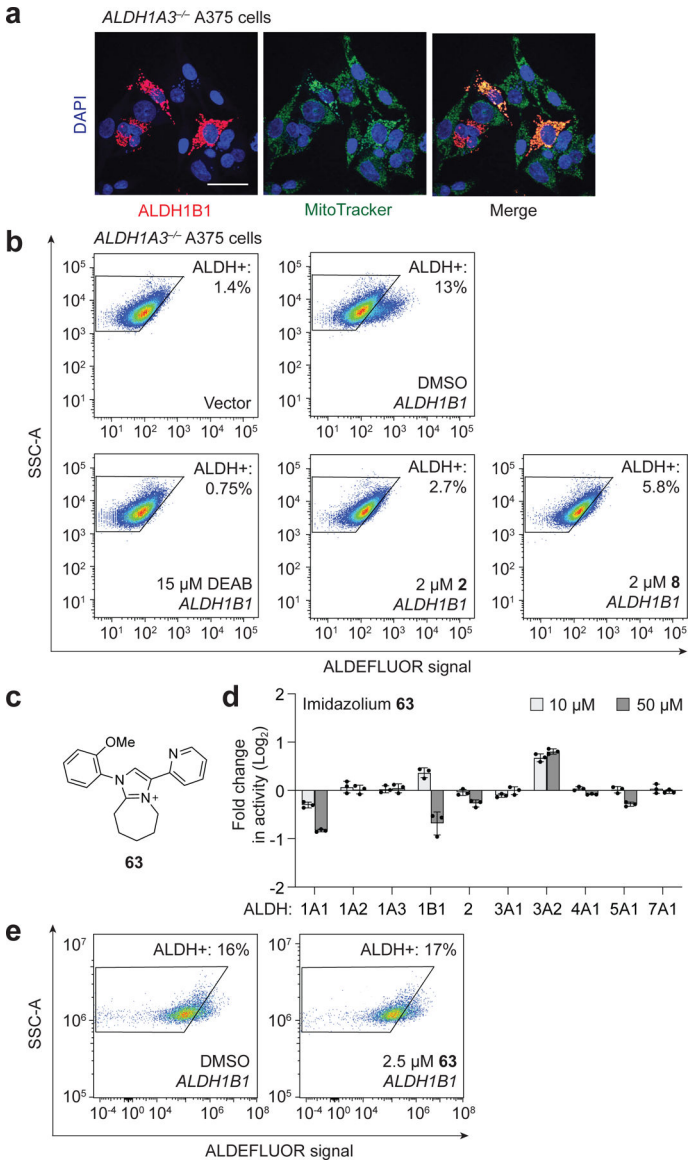
included in the corresponding figure legends. Student's t-test (two-tailed, unpaired) analysis in the Microsoft Excel were used to compare results and calculate *P* values. Statistical source data for the supplementary figures are provided in Supplementary Data Set 5.

Extended Data



Extended Data Fig. 1. Photoaffinity labeling of imidazolium-binding proteins

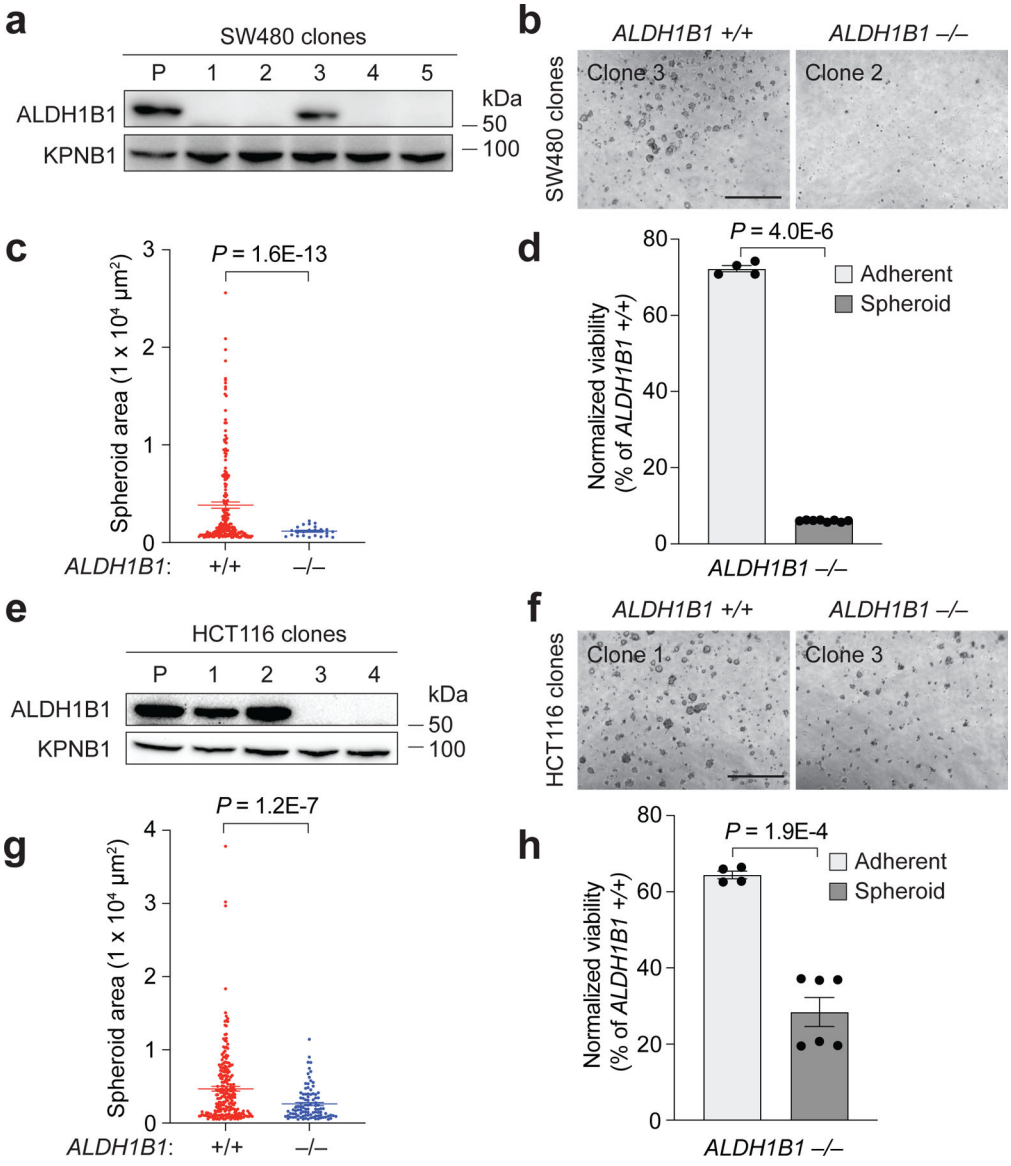
(a) Photoaffinity labeling and TAMRA-azide tagging of live NIH-3T3 cells, demonstrating the mitochondrial localization of the imidazolium target. Scale bar: 10 μm. (b) Mouse liver mitochondria were photocrosslinked with probe **3** in the absence or presence of competitor **2**. The mitochondria were then homogenized, reacted with BODIPY azide (no competitior) or MegaStrokes 673 azide (competitor), and resolved by two-dimensional gel electrophoresis. Protein spots that were specifically labeled by **3** (dashed box, spots 1–5) were then isolated for mass spectrometry-based sequencing.



Extended Data Fig. 2. Inhibition of cellular ALDH1B1 activity by imidazolium derivatives

(a) Immunofluorescence staining of *ALDH1A3*$^{-/-}$ A375 cells transiently transfected with ALDH1B1, demonstrating the mitochondrial localization of the exogenous protein. Scale bar: 40 μm. (b) Flow cytometry-based assays of ALDH1B1 activity and its pharmacological inhibition using *ALDH1A3*$^{-/-}$ A375 cells. The cells were transiently transfected with ALDH1B1 cDNA or a vector control, incubated with the designated compounds, and then treated with ALDEFLUOR reagent. The cells were gated by ALDEFLUOR signal intensity and side scatter area (SSC-A) to identify those with ALDH1B1 activity, and the percentage of cells outside of the negative control gate is shown for each condition. DMSO and the pan-ALDH inhibitor DEAB were used as negative and positive controls, respectively. (c) Chemical structure of **63**, which is inactive against ALDH1B1. (d) Activity of **63** against selected ALDH isoforms. Fold-change values are relative to a DMSO control and are the average of three biological replicates ± s.d. (e) Fluorescence-activated cell sorting

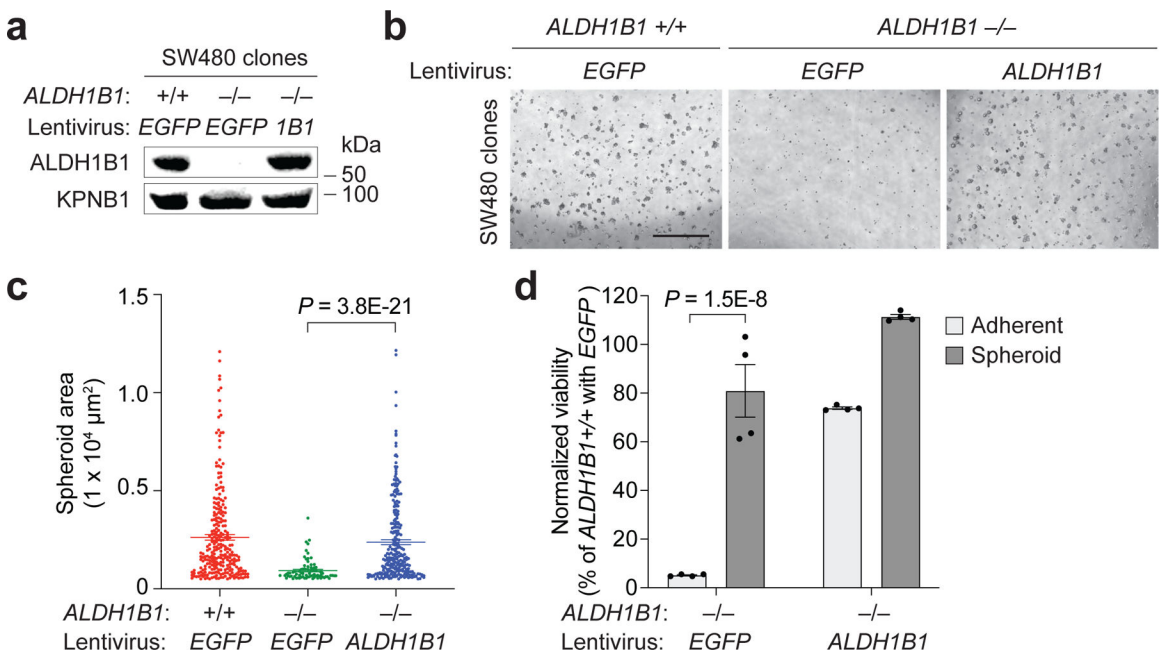
(FACS) plots of ALDH1B1-overexpressing *ALDH1A3*$^{-/-}$ A375 cells that were incubated with DMSO or **63** and then treated with ALDEFLUOR reagent.



Extended Data Fig. 3. ALDH1B1 promotes SW480 and HCT116 cell growth in spheroid culture

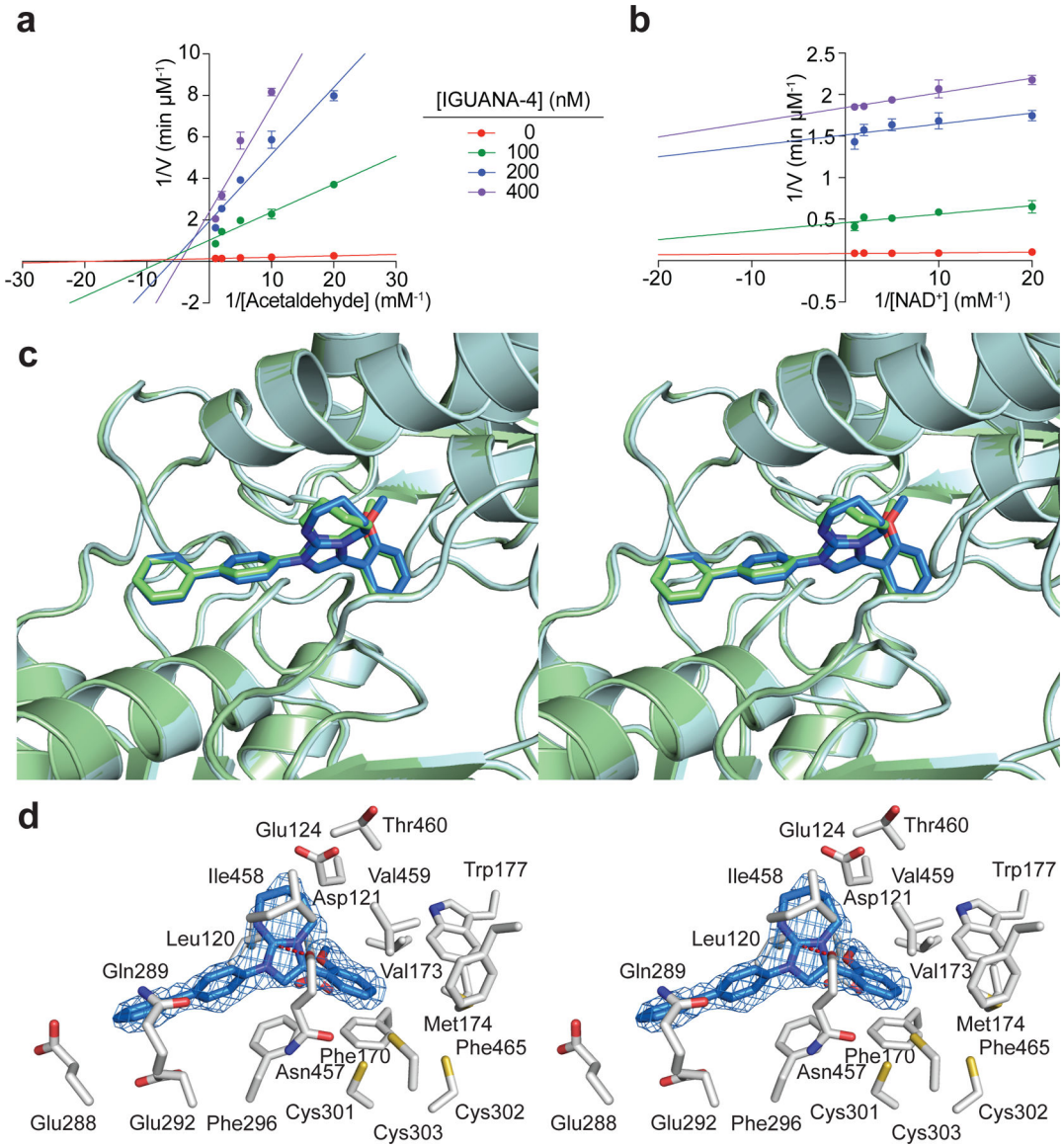
(a and e) Western blot detection of ALDH1B1 protein in individual SW480 and HCT116 cell clones that were transiently transfected with *Cas9* cDNA and *ALDH1B1* gRNA-1 and gRNA-2. Lysates from the parental lines (P) are also shown, and importin β1 (KPNB1) was used as a loading control. SW480 clone 3 and clone 2 were used as *ALDH1B1*$^{+/+}$ and *ALDH1B1*$^{-/-}$ clones for subsequent studies. (b and f) Phase-contrast micrographs of spheroid cultures derived from SW480 and HCT116 cells with differing *ALDH1B1* genotypes. (c and g) Quantification of spheroid sizes for the micrographs shown in (b) and (f). Each dot represents an individual spheroid with an area that is >500 µm2 in the images. Error bars represent the average spheroid size ± s.e.m. (d and h) Viability of the *ALDH1B1*$^{-/-}$ clone in either adherent or spheroid conditions, as determined by cellular ATP

levels and normalized to that of the as *ALDH1B1*$^{+/+}$ clone. Data are the average of four (d and h, adherent), eight (d, spheroid) and six (h, spheroid) biological replicates ± s.e.m. Scale bars: 1 mm.



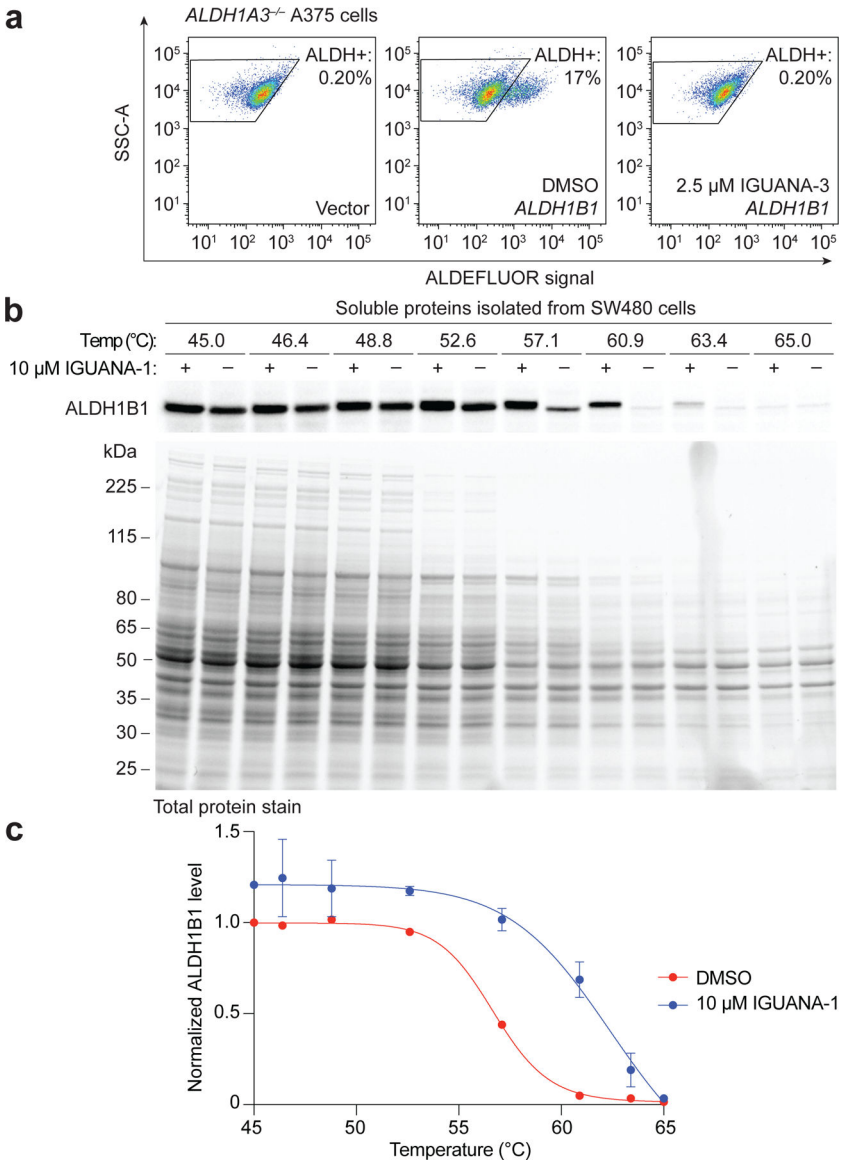
Extended Data Fig. 4. Exogenous *ALDH1B1* rescues the spheroid growth of ALDH1B1 knockout colon cancer cells

(a) Western blot analysis of SW480 clones with the indicated *ALDH1B1* genotypes and lentivirally transduced with either *EGFP* or exogenous *ALDH1B1*. Importin β1 (KNPB1) was used as gel-loading control. (b) Spheroid cultures of the SW480 clones described in (a). (c) Quantification of spheroid sizes for the micrographs shown in (b). Each dot represents an individual spheroid with an area that is >500 µm^2 in the image. Error bars represent the average spheroid size ± s.e.m. (d) Viability of the *ALDH1B1*$^{-/-}$ SW480 clone transduced with *EGFP* or *ALDH1B1* and cultured in either adherent or spheroid conditions, as determined by cellular ATP levels and normalized to that of the *ALDH1B1*$^{+/+}$ clone transduced with EGFP. Data are the average of four biological replicates ± s.e.m.



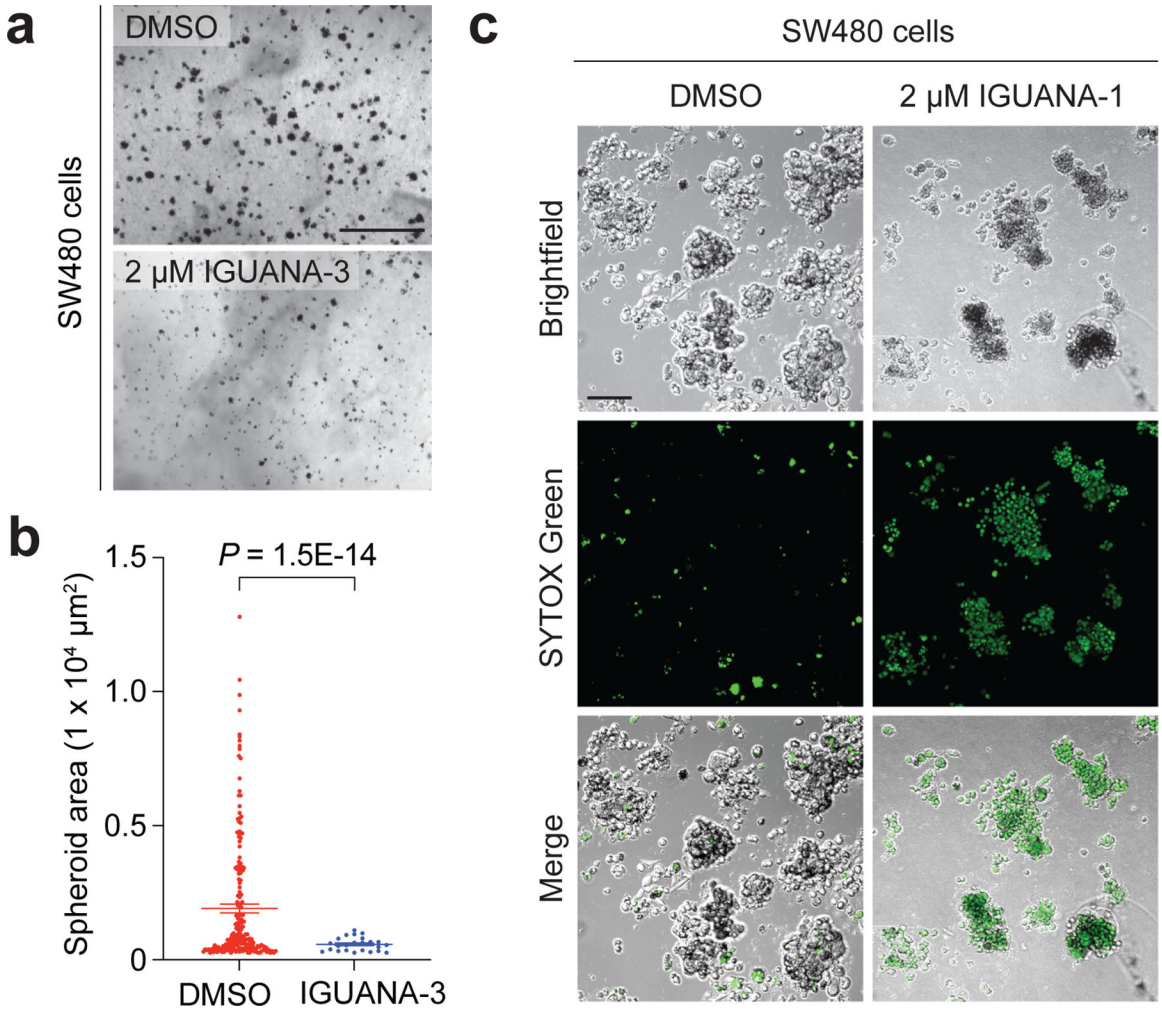
Extended Data Fig. 5. Molecular basis of ALDH1B1-IGUANA binding

(a-b) Lineweaver-Burk plots demonstrating that IGUANA-4 exhibits non-competitive inhibition with respect to acetaldehyde (a) and uncompetitive inhibition with respect to NAD^+ (b). Both enzyme kinetics assays used the same series of inhibitor concentrations, and the data are the average of three biological replicates ± s.e.m. (c) Wall-eyed stereoview of the ALDH1B1/NAD+/**2** (light green cartoon and green stick model) and ALDH1B1/NAD^+/IGUANA-4 (light blue cartoon and blue stick model) complexes shown as superimposed structures. Residues 143–154 and 490–500 and the NAD^+ cofactor were omitted for clarity. (d) Wall-eyed stereoview of the electron density map for IGUANA-4 (blue mesh) bound to ALDH1B1. The polder omit map was calculated with coefficients mF_o-DF_c and is contoured at 4σ. Residues in the inhibitor-binding site are shown, and the red dashed line indicates a potential n-to-π* interaction between the Asn457 backbone carbonyl and the guanidine (3.2 Å).



Extended Data Fig. 6. IGUANAs engage ALDH1B1 in live cells

(a) Flow cytometry-based ALDEFLUOR assays using *ALDH1A3*$^{-/-}$ A375 cells, demonstrating the ability of IGUANA-3 to inhibit cellular ALDH1B1 activity. The cells were transiently transfected with *ALDH1B1* cDNA or a vector control, incubated with IGUANA-3 or DMSO vehicle alone, and then treated with ALDEFLUOR reagent. The cells were gated by ALDEFLUOR signal intensity and side scatter area (SSC-A) to identify those with ALDH1B1 activity, and the percentage of cells outside of the negative control gate is shown for each condition. (b) Cellular thermal shift assay demonstrating that IGUANA-1 stabilizes endogenous ALDH1B1 in live SW480 cells. Western blot signals for ALDH1B1 and total protein staining of the soluble fraction are shown for each condition. (c) Corresponding melting curves of endogenous ALDH1B1 in the presence and absence of IGUANA-1. Data are the average of two biological replicates ± s.d., normalized to the DMSO condition at 45 °C.



Extended Data Fig. 7. IGUANAs suppress colon cancer spheroid growth

(a) Brightfield micrographs of SW480 spheroid cultures treated with IGUANA-3 and then stained with crystal violet. (b) Quantification of spheroid sizes for the micrographs shown in (b). Each dot represents an individual spheroid with an area that is >500 μm^2 in the micrograph. Error bars represent the average spheroid size ± s.e.m. (c) Brightfield, fluorescent, and merged micrographs of SW480 spheroids treated with IGUANA-1 or DMSO vehicle alone for 3 days and then stained overnight with the viability dye SYTOX Green. Scale bars: a, 1 mm; c, 100 μm.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We thank T. Hurley for the vector for bacterial ALDH1B1 expression, E. Patton for ALDH1A3$^{-/-}$ A375 cells, P. Beachy for Shh-LIGHT2 cells, G. Ponhert for BOPIDY azide, and S. Swick for assistance with compound characterization. This work was supported by the National Institutes of Health (R35 GM127030, R01 CA244334, and R01 GM113100 to J.K.C.; U01 CA217851 and U54 CA224081 to C.J.K.; R01 AA11147 to D.M.R.; and F32 CA183527 to A.E.O.) SPARK at Stanford (J.K.C.), Weston Havens Foundation (J.K.C.), the Emerson

Collective (J.K.C. and C.J.K.), the Human Cancer Models Initiative (C.J.K), and Stand Up to Cancer (C.J.K). Flow cytometry experiments were performed at the Stanford Shared FACS Facility supported by an NIH S10 Shared Instrument Grant (S10 OD026831). Use of the Stanford University Mass Spectrometry facility is supported in part by the National Institutes of Health (P30 CA124435) utilizing the Stanford Cancer Institute Proteomics/Mass Spectrometry Shared Resources. Use of the SSRL/SLAC National Accelerator Laboratory is supported by the U.S. Department of Energy, Office of Science and Office of Basic Energy Sciences under Contract No. DE-AC02–76SF00515. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and by the National Institutes of Health (P41 GM103393). NMR experiments included spectra acquired on a Bruker Avance NEO 500 MHz spectrometer supported by an NIH S10 Shared Instrument Grant (S10 OD028697).

Data availability

Any data generated or analyzed during this study, associated protocols, and materials are available from the corresponding author on reasonable request. X-ray crystal structures have been validated and deposited with the Protein Data Bank with the following entries: 7MJC (ALDH1B1/NAD^+), 7MJD (ALDH1B1/NAD^+/**2**), and 7RAD (ALDH1B1/NAD^+/IGUANA-4). Raw and processed proteomics data for the TPP study are publicly available in the MassIVE respository (MassIVE ID: MSv000088824), which a member of the ProteomeXchange Consortium (ProteomeXchange ID: PXD031630). Raw and processed RNA-seq data for the identification of ALDH1B1-dependent genes are publicly available in the Gene Expression Omnibus database (GEO accession number: GSE165621).

REFERENCES

1. Vassalli G Aldehyde dehydrogenases: not just markers, but functional regulators of stem cells. Stem Cells Int 2019, 3904645, (2019). [PubMed: 30733805]
2. Xu X et al. Aldehyde dehydrogenases and cancer stem cells. Cancer Lett 369, 50–57, (2015). [PubMed: 26319899]
3. Marcato P et al. Aldehyde dehydrogenase activity of breast cancer stem cells is primarily due to isoform ALDH1A3 and its expression is predictive of metastasis. Stem Cells 29, 32–45, (2011). [PubMed: 21280157]
4. Mao P et al. Mesenchymal glioma stem cells are maintained by activated glycolytic metabolism involving aldehyde dehydrogenase 1A3. Proc. Natl. Acad. Sci. U. S. A 110, 8644–8649, (2013). [PubMed: 23650391]
5. Luo Y et al. ALDH1A isozymes are markers of human melanoma stem cells and potential therapeutic targets. Stem Cells 30, 2100–2113, (2012). [PubMed: 22887839]
6. Shao C et al. Essential role of aldehyde dehydrogenase 1A3 for the maintenance of non-small cell lung cancer stem cells is associated with the STAT3 pathway. Clin. Cancer Res 20, 4154–4166, (2014). [PubMed: 24907115]
7. Landen CN Jr. et al. Targeting aldehyde dehydrogenase cancer stem cells in ovarian cancer. Mol. Cancer Ther 9, 3186–3199, (2010). [PubMed: 20889728]
8. Singh S et al. ALDH1B1 is crucial for colon tumorigenesis by modulating Wnt/beta-catenin, Notch and PI3K/Akt signaling pathways. PLoS One 10, e0121648, (2015). [PubMed: 25950950]
9. Mameishvili E et al. Aldh1b1 expression defines progenitor cells in the adult pancreas and is required for Kras-induced pancreatic cancer. Proc. Natl. Acad. Sci. U. S. A 116, 20679–20688, (2019). [PubMed: 31548432]
10. Meng E et al. ALDH1A1 maintains ovarian cancer stem cell-like properties by altered regulation of cell cycle checkpoint and DNA repair network signaling. PLoS One 9, e107142, (2014). [PubMed: 25216266]
11. Li B, Yang K, Liang D, Jiang C & Ma Z Discovery and development of selective aldehyde dehydrogenase 1A1 (ALDH1A1) inhibitors. Eur. J. Med. Chem 209, 112940, (2021). [PubMed: 33328099]

12. Gelardi ELM et al. A selective competitive inhibitor of aldehyde dehydrogenase 1A3 hinders cancer cell growth, invasiveness and stemness in vitro. Cancers (Basel) 13, (2021).
13. Li J et al. A specific inhibitor of ALDH1A3 regulates retinoic acid biosynthesis in glioma stem cells. Commun. Biol 4, 1420, (2021). [PubMed: 34934174]
14. Stagos D et al. Aldehyde dehydrogenase 1B1: molecular cloning and characterization of a novel mitochondrial acetaldehyde-metabolizing enzyme. Drug Metab. Dispos 38, 1679–1687, (2010). [PubMed: 20616185]
15. Chen Y et al. Aldehyde dehydrogenase 1B1 (ALDH1B1) is a potential biomarker for human colon cancer. Biochem. Biophys. Res. Commun 405, 173–179, (2011). [PubMed: 21216231]
16. Singh S et al. ALDH1B1 links alcohol consumption and diabetes. Biochem. Biophys. Res. Commun 463, 768–773, (2015). [PubMed: 26086111]
17. Singh S et al. Aldehyde dehydrogenase 1B1 as a modulator of pancreatic adenocarcinoma. Pancreas 45, 117–122, (2016). [PubMed: 26566217]
18. Hom ME, Ondrus AE, Sakata-Kato T, Rack PG & Chen JK Bicyclic imidazolium inhibitors of Gli transcription factor activity. ChemMedChem 15, 1044–1049, (2020). [PubMed: 32268014]
19. Luo M & Tanner JJ Structural basis of substrate recognition by aldehyde dehydrogenase 7A1. Biochemistry 54, 5513–5522, (2015). [PubMed: 26260980]
20. Riveros-Rosas H, Gonzalez-Segura L, Julian-Sanchez A, Diaz-Sanchez AG & Munoz-Clares RA Structural determinants of substrate specificity in aldehyde dehydrogenases. Chem. Biol. Interact 202, 51–61, (2013). [PubMed: 23219887]
21. Sarvi S et al. ALDH1 bio-activates nifuroxazide to eradicate ALDH(High) melanoma-initiating cells. Cell Chem. Biol 25, 1456–1469 e1456, (2018). [PubMed: 30293938]
22. Cerami E et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2, 401–404, (2012). [PubMed: 22588877]
23. Tang Z et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 45, W98–W102, (2017). [PubMed: 28407145]
24. Putzbach W et al. Many si/shRNAs can kill cancer cells by targeting multiple survival genes through an off-target mechanism. Elife 6, (2017).
25. Brinkman EK, Chen T, Amendola M & van Steensel B Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Res 42, e168, (2014). [PubMed: 25300484]
26. Chen X et al. LGR5 is required for the maintenance of spheroid-derived colon cancer stem cells. Int. J. Mol. Med 34, 35–42, (2014). [PubMed: 24789370]
27. Reily C et al. Mitochondrially targeted compounds and their impact on cellular bioenergetics. Redox Biol 1, 86–93, (2013). [PubMed: 23667828]
28. Martinez Molina D et al. Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. Science 341, 84–87, (2013). [PubMed: 23828940]
29. Franken H et al. Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry. Nat. Protoc 10, 1567–1593, (2015). [PubMed: 26379230]
30. Vodenkova S et al. 5-fluorouracil and other fluoropyrimidines in colorectal cancer: Past, present and future. Pharmacol. Ther 206, 107447, (2020). [PubMed: 31756363]
31. Chou TC Drug combination studies and their synergy quantification using the Chou-Talalay method. Cancer Res 70, 440–446, (2010). [PubMed: 20068163]
32. Neal JT et al. Organoid modeling of the tumor immune microenvironment. Cell 175, 1972–1988 e1916, (2018). [PubMed: 30550791]
33. Lo YH et al. A CRISPR/Cas9-Eegineered ARID1A-deficient human gastric cancer organoid model reveals essential and nonessential modes of oncogenic transformation. Cancer Discov 11, 1562–1581, (2021). [PubMed: 33451982]
34. Giroux V et al. Mouse intestinal Krt15+ crypt cells Are radio-resistant and tumor initiating. Stem Cell Rep 10, 1947–1958, (2018).
35. Nakanishi Y et al. Dclk1 distinguishes between tumor and normal stem cells in the intestine. Nat. Genet 45, 98–103, (2013). [PubMed: 23202126]

36. Yang L et al. Pregnancy-specific glycoprotein 9 (PSG9), a driver for colorectal cancer, enhances angiogenesis via activation of SMAD4. Oncotarget 7, 61562–61574, (2016). [PubMed: 27528036]
37. Casazza A et al. Impeding macrophage entry into hypoxic tumor areas by Sema3A/Nrp1 signaling blockade inhibits angiogenesis and restores antitumor immunity. Cancer Cell 24, 695–709, (2013). [PubMed: 24332039]
38. De Vlaeminck Y et al. Targeting neuropilin-1 with nanobodies reduces colorectal carcinoma development. Cancers (Basel) 12, (2020).
39. Cox PJ, Phillips BJ & Thomas P The enzymatic basis of the selective action of cyclophosphamide. Cancer Res 35, 3755–3761, (1975). [PubMed: 172233]
40. Weiner H, Hu JH & Sanny CG Rate-limiting steps for the esterase and dehydrogenase reaction catalyzed by horse liver aldehyde dehydrogenase. J. Biol. Chem 251, 3853–3855, (1976). [PubMed: 945270]
41. Morgan CA & Hurley TD Characterization of two distinct structural classes of selective aldehyde dehydrogenase 1A1 inhibitors. J. Med. Chem 58, 1964–1975, (2015). [PubMed: 25634381]
42. Yang SM et al. Discovery of orally bioavailable, quinoline-based aldehyde dehydrogenase 1A1 (ALDH1A1) inhibitors with potent cellular activity. J. Med. Chem 61, 4883–4903, (2018). [PubMed: 29767973]
43. Huddle BC et al. Structure-based optimization of a novel class of aldehyde dehydrogenase 1A (ALDH1A) subfamily-selective inhibitors as potential adjuncts to ovarian cancer chemotherapy. J. Med. Chem 61, 8754–8773, (2018). [PubMed: 30221940]
44. Snyder V, Reed-Newman TC, Arnold L, Thomas SM & Anant S Cancer stem cell metabolism and potential therapeutic targets. Front. Oncol 8, 203, (2018). [PubMed: 29922594]
45. Bastide A & David A The ribosome, (slow) beating heart of cancer (stem) cell. Oncogenesis 7, 34, (2018). [PubMed: 29674660]
46. Siegel RL, Miller KD & Jemal A Cancer statistics, 2020. CA Cancer J. Clin 70, 7–30, (2020). [PubMed: 31912902]
47. Cho YH et al. 5-FU promotes stemness of colorectal cancer via p53-mediated WNT/beta-catenin pathway activation. Nat. Commun 11, 5321, (2020). [PubMed: 33087710]
48. Cong J et al. A novel chemoradiation targeting stem and nonstem pancreatic cancer cells by repurposing disulfiram. Cancer Lett 409, 9–19, (2017). [PubMed: 28864067]

METHODS-ONLY REFERENCES

49. Taipale J et al. Effects of oncogenic mutations in Smoothened and Patched can be reversed by cyclopamine. Nature 406, 1005–1009, (2000). [PubMed: 10984056]
50. Hoppel C, DiMarco JP & Tandler B Riboflavin and rat hepatic cell structure and function. Mitochondrial oxidative metabolism in deficiency states. J. Biol. Chem 254, 4164–4170, (1979). [PubMed: 571436]
51. Wolfram S et al. A small azide-modified thiazole-based reporter molecule for fluorescence and mass spectrometric detection. Beilstein J. Org. Chem 10, 2470–2479, (2014). [PubMed: 25383118]
52. Elias JE & Gygi SP Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat. Methods 4, 207–214, (2007). [PubMed: 17327847]
53. Chen CH, Cruz LA & Mochly-Rosen D Pharmacological recruitment of aldehyde dehydrogenase 3A1 (ALDH3A1) to assist ALDH2 in acetaldehyde and ethanol metabolism in vivo. Proc. Natl. Acad. Sci. U. S. A 112, 3074–3079, (2015). [PubMed: 25713355]
54. Winn MD et al. Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr 67, 235–242, (2011). [PubMed: 21460441]
55. DeLano WL Pymol: An open-source molecular graphics tool. CPP4 Newsletter on Protein Crystallography 40, 82–92, (2002).
56. Campeau E et al. A versatile viral system for expression and depletion of proteins in mammalian cells. PLoS One 4, e6529, (2009). [PubMed: 19657394]

57. Jafari R et al. The cellular thermal shift assay for evaluating drug target interactions in cells. Nat. Protoc 9, 2100–2122, (2014). [PubMed: 25101824]

58. Chen JK, Taipale J, Young KE, Maiti T & Beachy PA Small molecule modulation of Smoothened activity. Proc. Natl. Acad. Sci. U. S. A 99, 14071–14076, (2002). [PubMed: 12391318]

59. Tran U, Zhang GC, Eom R, Billingsley KL & Ondrus AE Small molecule intervention in a protein kinase C-Gli transcription factor axis. ACS Chem. Biol 15, 1321–1327, (2020). [PubMed: 32479053]

60. Benjamini Y & Hochberg Y Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc Series B. Stat. Methodol 57, 289–300, (1995).

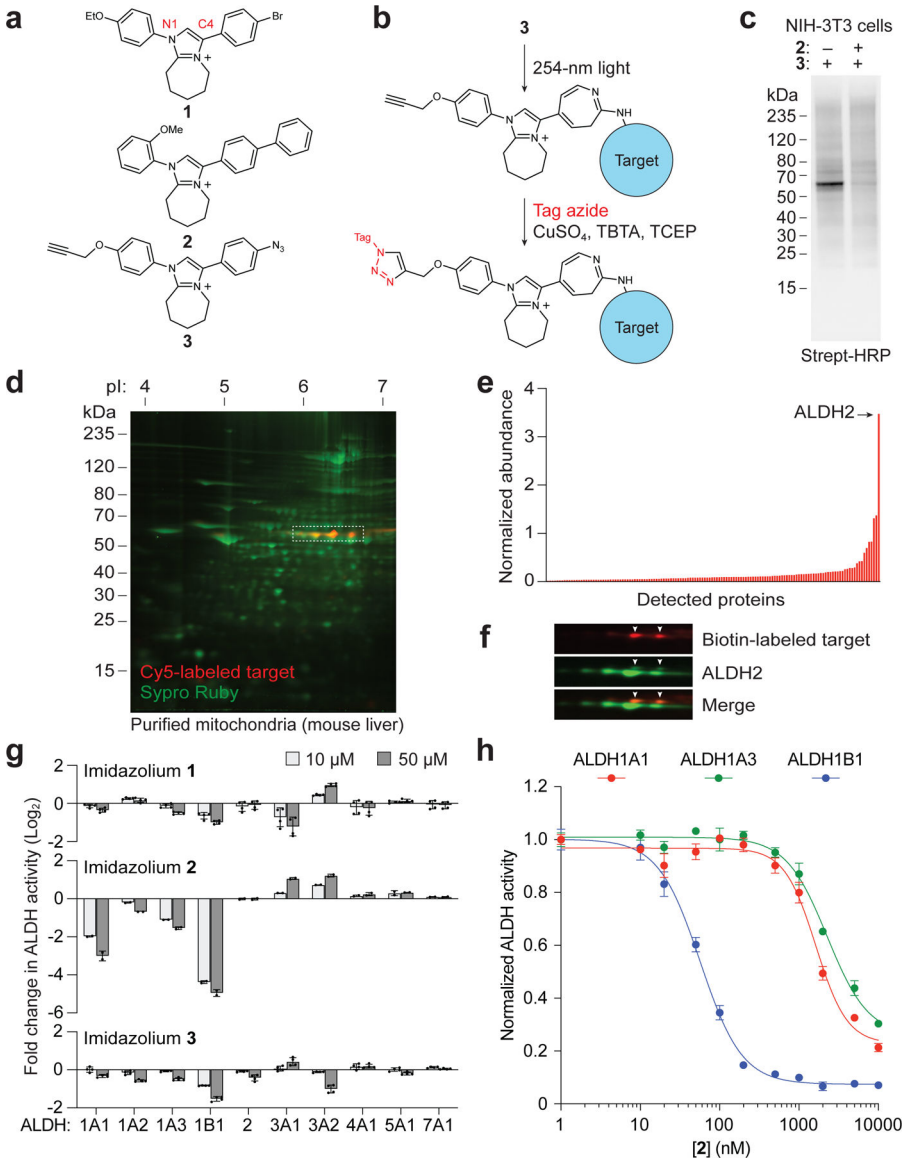


Figure 1. Identification of an ALDH1B1-selective imidazolium.

(a) Chemical structures of imidazoliums **1**-**3**. N1 and C4 positions on the imidazolium ring are labeled. (b) Reaction scheme for target photocrosslinking and click chemistry tagging. (c) Photoaffinity labeling of a 60-kDa imidazolium-binding protein in live NIH-3T3 cells. Biotin-azide was used to tag the photocrosslinked protein, which was then detected by far-western blot analysis. (d) Photoaffinity labeling and Cy5-azide tagging of the 60-kDa protein (dashed box) in intact, purified mitochondria from mouse liver. The mitochondrial lysate was resolved by two-dimensional gel electrophoresis, and Sypro Ruby was used to fluorescently stain the mitochondrial proteome. (e) Mass spectrometry sequencing of gel spots corresponding to the photoaffinity-labeled proteins, revealing ALDH2 as the candidate target with the highest normalized spectral abundance factor. (f) Confirmation of **3**-ALDH2 photocrosslinking by biotin tagging and western blot analysis. Arrowheads label proteins that are both biotinylated and stained by anti-ALDH2 antibody. (g) Activities of **1**-**3**

against selected ALDH isoforms in enzyme kinetics assays using acetaldehyde as substrate. Fold-change values are relative to a DMSO control and are the average of at least two biological replicates ± s.d. (h) Dose-response curves of **2** against ALDH1A1, ALDH1A3, and ALDH1B1. Data are the average of three biological replicates ± s.e.m.

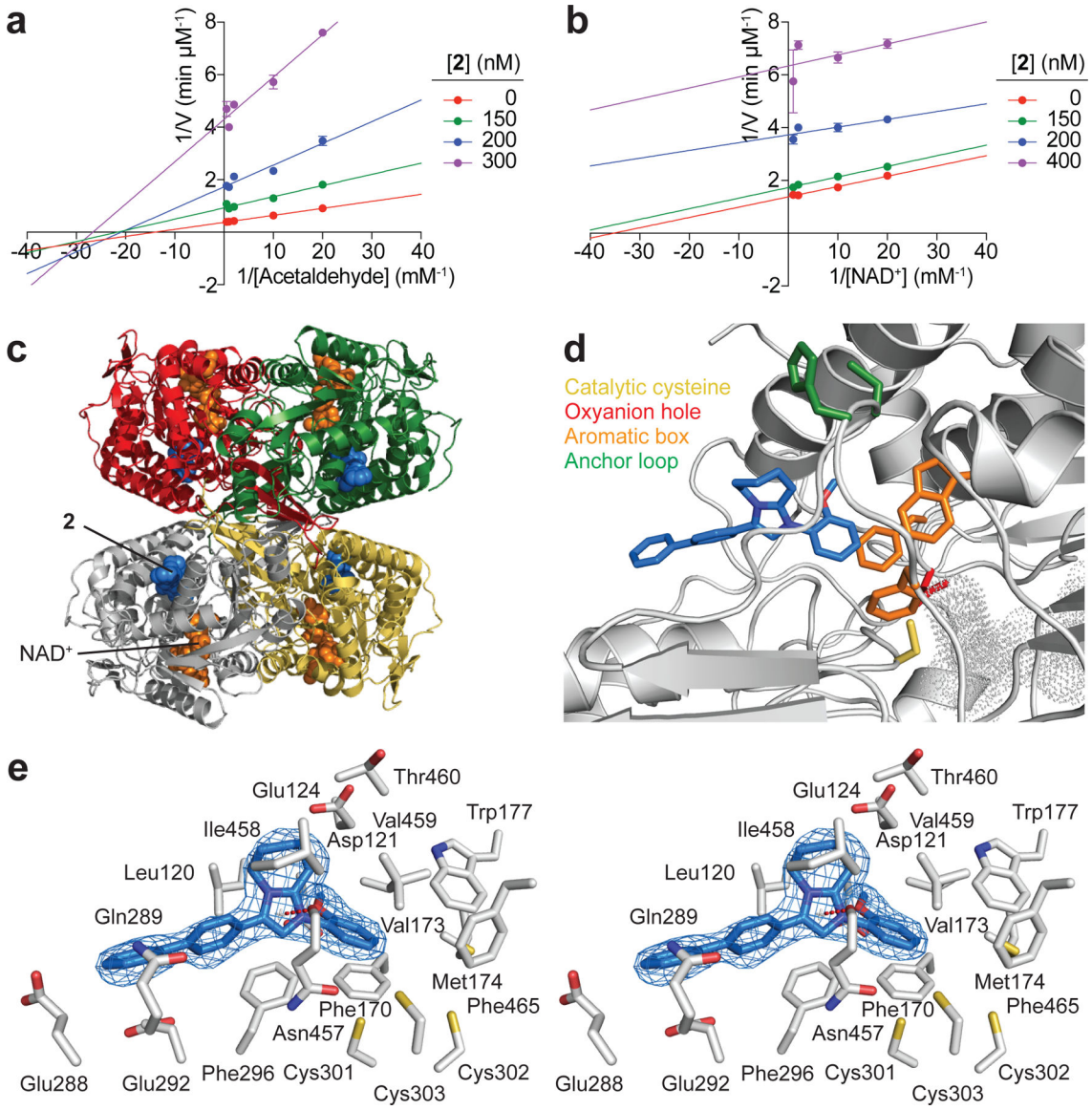


Figure 2. Molecular basis of ALDH1B1-imidazolium binding.

(a-b) Lineweaver-Burk plots demonstrating that **2** exhibits non-competitive inhibition with respect to acetaldehyde (a) and uncompetitive inhibition with respect to NAD^+ (b). Data are the average of two biological replicates ± s.e.m. (c) Tetrameric structure of the ALDH1B1/NAD^+/**2** complex, with the monomers shown in a different colors. (d) Close-up view of molecular interactions in the inhibitor-binding site, with **2** is shown as a blue stick model and NAD^+ represented as dots. Conserved landmarks in the ALDH family are also shown as stick models, including the catalytic cysteine (Cys302; yellow), oxyanion hole (Asn169; red), aromatic box (Phe170, Trp177, and Phe465; orange), and aldehyde substrate anchor loop (Cys461 and His462; green). (e) Wall-eyed stereoview of the electron density map for **2** (blue mesh) bound to ALDH1B1. The polder omit map was calculated with coefficients mFo-DFc and is contoured at 4σ. Residues in the inhibitor-binding site are shown, and

the red dashed line indicates a potential n-to-π* interaction between the Asn457 backbone carbonyl and the imidazolium ring (3.3 Å).

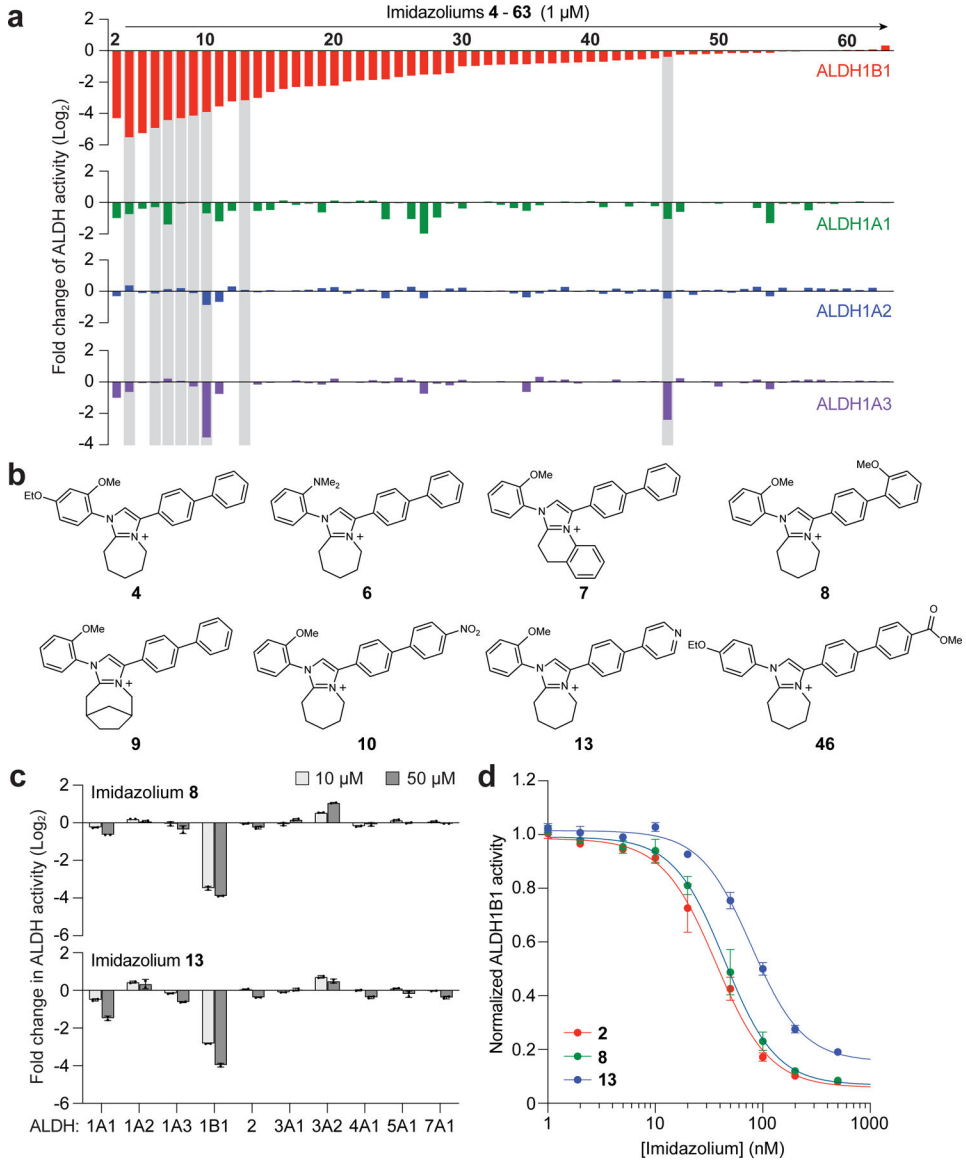


Figure 3. Structure-activity relationship analysis of the imidazolium pharmacophore.

(a-b) Compound activities against ALDH1 isoforms in enzyme kinetics assays using acetaldehyde as substrate. The compounds were tested at a 1-μM dose, and data are the average of at least two biological replicates. Chemical structures are shown for representative imidazoliums, which are also highlighted in the compound activity profiles with gray bars. (c) Profiling of **8** and **13** against a broader panel of ALDH isoforms. Data are the average of two biological replicates ± s.d. (d) Dose-response curves for **2**, **8**, and **13** in ALDH1B1 enzyme kinetics assays. Data are the average of three biological replicates ± s.e.m.

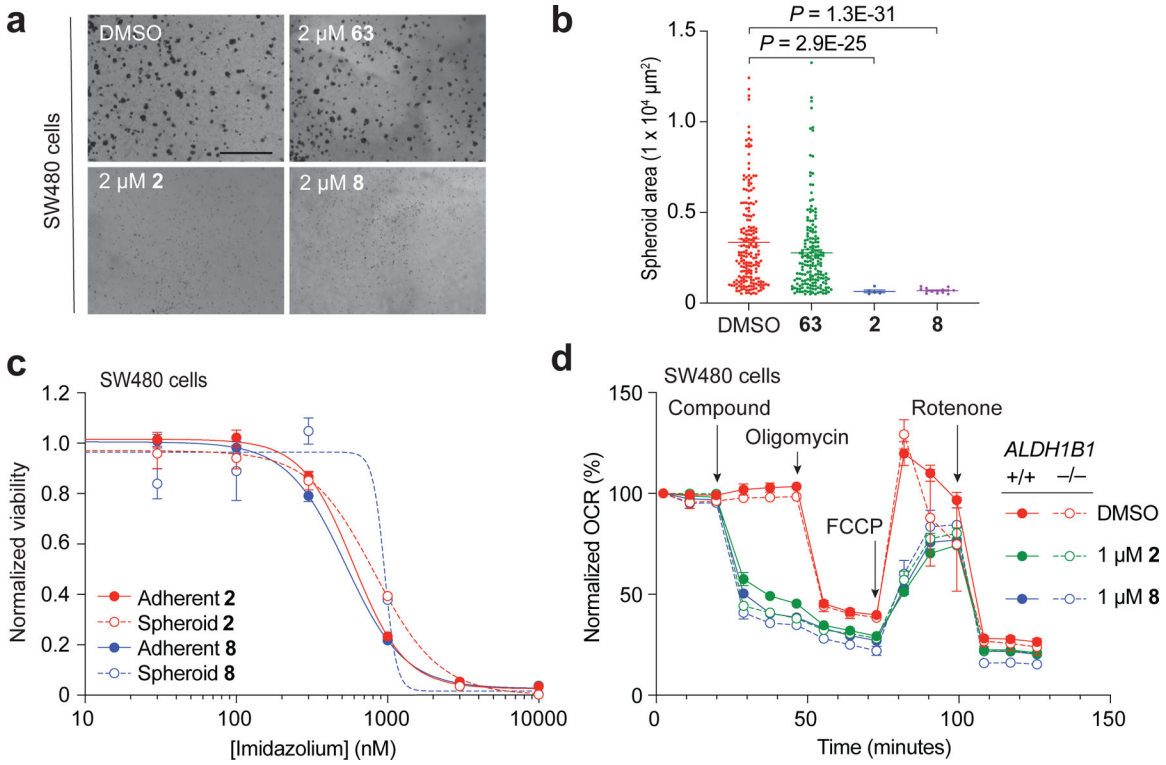


Figure 4. Imidazolium-based ALDH1B1 inhibitors exhibit off-target mitochondrial toxicity.

(a) Brightfield micrographs of SW480 spheroid cultures treated with the designated imidazolium derivatives and then stained with crystal violet. (b) Quantification of spheroid sizes in (a). Each dot represents an individual spheroid with an area that is >500 μm^2 in the micrographs. Error bars represent the average spheroid size ± s.e.m. (c) Dose-response curves for **2** and **8** on SW480 cells cultured in either adherent or spheroid conditions. Data are the average of three biological replicates ± s.e.m. (d) Seahorse assays evaluating the effects of imidazoliums on the oxygen consumption rates of *ALDH1B1*$^{+/+}$ and *ALDH1B1*$^{-/-}$ SW480 cells. Data are the average of three biological replicates ± s.e.m. Scale bars: 1 mm.

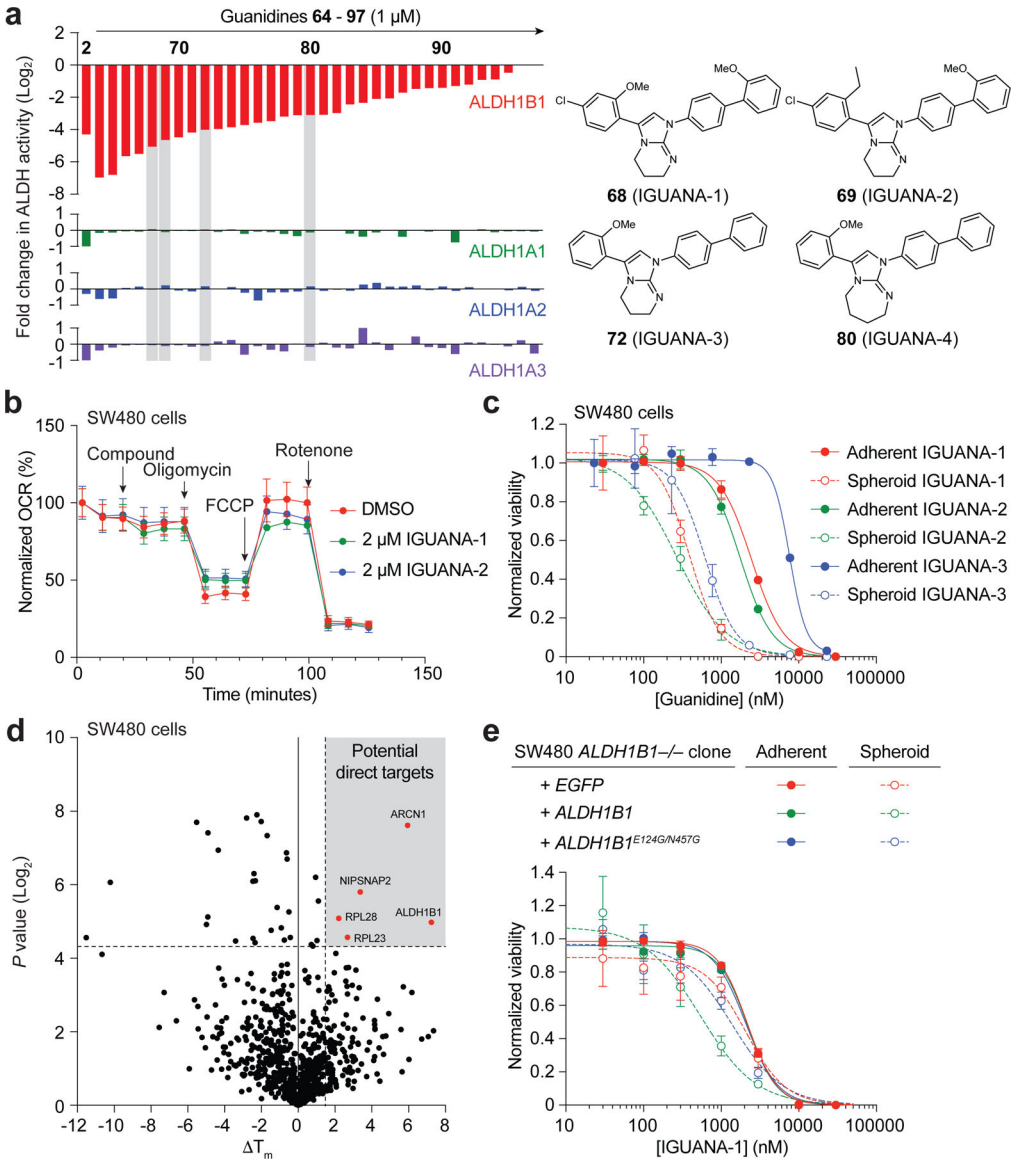


Figure 5. Development of guanidine-based ALDH1B1 antagonists.

(a) Compound activities against ALDH1 isoforms in enzyme kinetics assays, revealing several ALDH1B1-selective guanidines. The compounds were tested at a 1-μM dose, and data are the average of at least two biological replicates. Chemical structures are shown for representative guanidine, which are also highlighted in the compound activity profiles with gray bars. (b) Seahorse assays showing that guanidiums do not inhibit oxygen consumption rates of SW480 cells. Data are the average of three biological replicates ± s.e.m. (c) Dose-response curves for IGUANA-1, IGUANA-2, and IGUANA-3 on adherent or spheroid cultures of SW480 cells. Data are the average of three biological replicates ± s.e.m. (d) TPP analysis demonstrating the biochemical specificity of IGUANA-1in SW480 cells. Data are the average of two biological replicates, and a volcano plot with cutoffs of $\Delta T_m \geq 1.5$ °C and $P < 0.05$ is shown. Potential direct targets are colored red. (e) Dose-response curves for IGUANA-1on adherent or spheroid cultures of *ALDH1B1*$^{-/-}$ SW480 cells lentivirally

transduced with either *EGFP*, exogenous *ALDH1B1*, or an inhibitor-resistant *ALDH1B1* mutant. Data are the average of three biological replicates ± s.e.m.

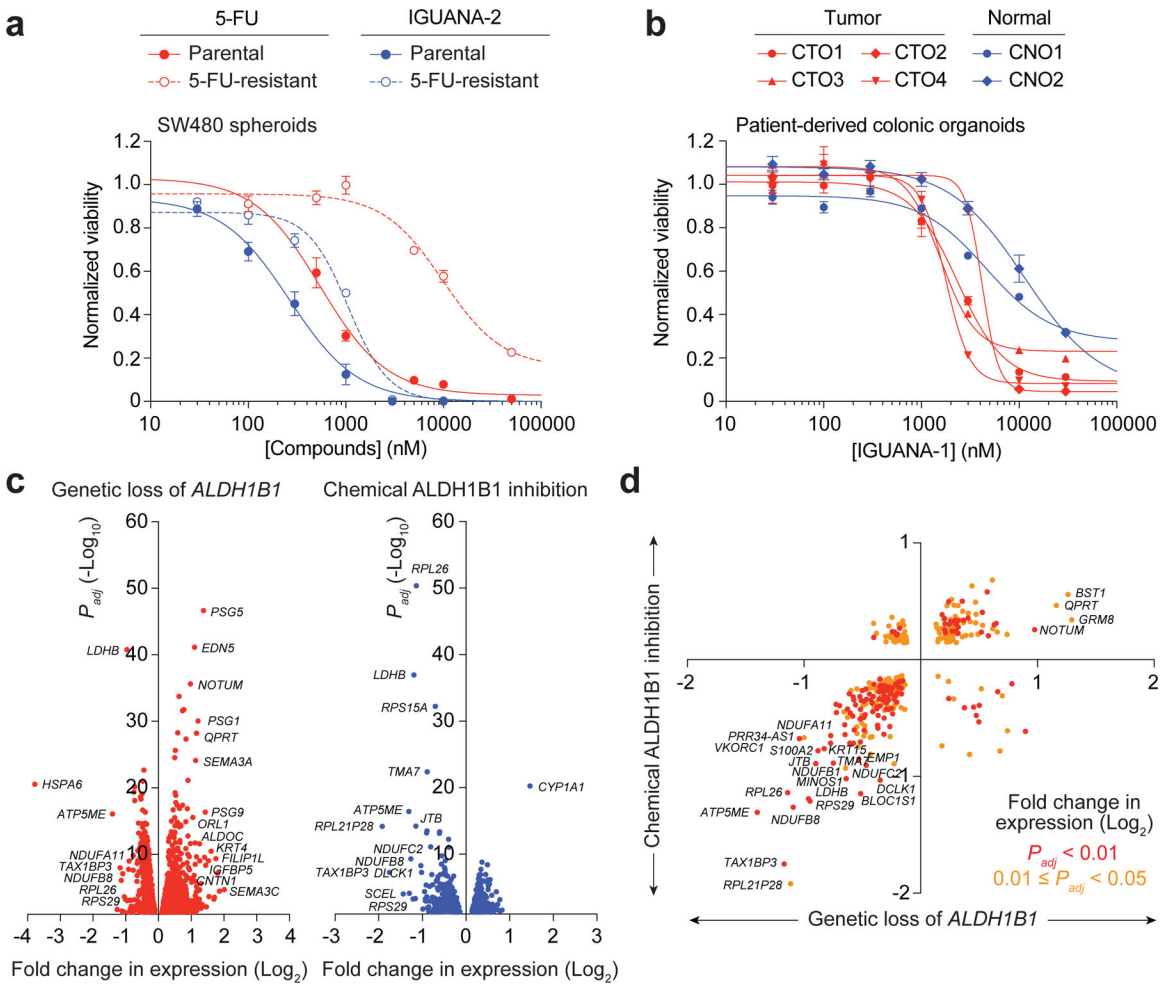


Figure 6. ALDH1B1 inhibitor activities in models of colorectal cancer.

(a) Activities of 5-FU and IGUANA-2 against parental and 5-FU-resistant SW480 cell lines. The cells were cultured as spheroids, and data are the average of three biological replicates ± s.e.m. (b) Activities of IGUANA-1 against patient-derived organoids from colorectal tumors (CTO1, CTO2, CTO3, and CTO4) and normal colonic tissues (CNO1 and CNO2). Data are the average of five (CNO1, CTO3), six (CTO2) and seven (CNO2, CTO1, CTO3) biological replicates ± s.e.m. (c) Volcano plots of the transcriptional perturbations caused by genetically or chemically induced loss of ALDH1B1 function. Gene expression changes with adjusted *P* values (P_{adj}) < 0.05 are graphed, and selected transcripts are annotated. *ALDH1B1* is excluded in the volcano plot for the genetic loss of *ALDH1B1* function since its P_{adj} value exceeds the range of the y-axis. (d) Comparison of transcriptional changes caused by the genetic and chemical perturbations. Transcripts with statistically significant changes in expression for both data sets are graphed, and those that are most dependent on ALDH1B1 activity are annotated.