POU5F1B **is a human-restricted ROCK inhibitor-sensitive oncoprotein that restructures membrane nanodomains to increase cell adhesion**

Laia Simó-Riudalbas[1,2*], Sandra Offner[1], Laurence Abrami[1], Eduard M Unterauer[3,4], Julien Duc[1], Evarist Planet[1], Ralf Jungmann[3,4], Gisou van der Goot[1], Didier Trono[1]

[1]School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

[2] Center for Integrative Genomics, Genopode Building, University of Lausanne, Lausanne, Switzerland.

[3]Max Planck Institute of Biochemistry, Planegg, Germany.

[4]Faculty of Physics and Center for NanoScience, Ludwig Maximilian University, Munich, Germany.

*: Corresponding author: laia.simoriudalbas@unil.ch

Evolution confers new species with distinctive biological features, which can translate in purely mechanistic speciation or in novel phenotypic traits. *POU5F1B* arose by retrotransposition of *POU5F1/OCT4* in the last common ancestor of modern Hominoidae. Its human product is endowed with oncogenic properties, notably promoting gastrointestinal cancer growth and metastasis. Here, we reveal that the oncogenic action of POU5F1B requires ubiquitination of lysine residues found only in the human protein. This post-translational modification is essential for POU5F1B to localize to the cytoplasm and become acylated by the DHHC17 palmitoyltransferase. This leads to POU5F1B association with detergent-resistant membrane subdomains, where it triggers the accumulation of integrins and signaling molecules, and to a stimulation of cell focal adhesion. Finally, we determined that POU5F1B stability is critically dependent on ROCK, a crucial regulator of cytoskeletal dynamics and cell motility.

The *POU5F1B* retrogene is transcribed by activation of upstream transposable element-embedded regulatory sequences in several human solid tumors including colon, stomach, breast, prostate and uterus[1]. Its expression is a negative prognostic factor in colorectal[1], gastric[2] and hepatocellular[3] malignancies and is associated with advanced histological grades of cervical cancer[4], where the *POU5F1B* locus is a papillomavirus integration hotspot[5]. Supporting its role as a cancer-promoting factor, the POU5F1B protein stimulates the proliferation of colorectal[1], gastric[2] and hepatocellular[3] carcinoma cells *in vitro* and increases the growth[1,2] and metastatic[1] potential of human gastrointestinal tumor cells in mouse xenotransplantation experiments. While its POU5F1/OCT4 relative is a DNA-binding transcription factor concentrated in the nucleus, POU5F1B accumulates in the cytoplasm where it associates with detergent-resistant membranes (DRM)[1]. The two proteins differ at 15 out of 359/360 positions, 5 of which are unique to human POU5F1B compared with its orthologues from other great apes (Fig. 1A). Amongst these, lysine residues at positions 135 and 182, replacing respectively the aspartate and threonine present in human POU5F1 and non-human POU5F1B, stood out as possible targets of ubiquitination, a post-translational modification that can influence the stability, activity or subcellular localization of proteins[6]. Confirming this hypothesis, we found that POU5F1B is ubiquitinated and that this modification is largely suppressed by the $K^{182}T$, $K^{182}A$, and $K^{135}E$ mutations (Fig. 1B). Furthermore, we observed that replacing either lysine by the corresponding amino acid of POU5F1/OCT4 redirected POU5F1B to the nucleus (Fig. 1C), and that it also abrogated both its DRM association, assessed using the gradient-based DRM extraction technique as biochemical readout of cholesterol-rich nanodomains[7,8], and its growth-promoting action on colorectal

cancer cells (Fig. 1D and 1E). As well, POU5F1B localized to the nucleus of cells treated with the ubiquitin activating enzyme inhibitor TAK-243[9] (Supp. Fig. 1A).

The membrane association of proteins is promoted by covalent linkage of fatty acids, amongst which S-acylation (commonly referred to as S-palmitoylation) plays a particular role in recruitment to detergent resistant membrane subdomains[10,11]. Using the acyl-resin assisted capture (Acylrac) method, whereby S-acylated proteins are trapped on thiol-reactive sepharose beads, we could demonstrate that POU5F1B is S-acylated, and that this modification is reduced by the $K^{135}E$ and $K^{182}T$ mutations (Fig. 2A). We evaluated the dynamics of POU5F1B S-acylation by monitoring radioactive 3H-palmitate incorporation using immunoprecipitation and autoradiographic analysis (Fig. 2B), and confirmed POU5F1B S-acylation. The SwissPalm (https://swisspalm.org/proteins/Q06416) predicted S-acylation on cysteine198 of POU5F1B with a high confidence score. We thus mutated this residue, and found that signal was indeed lost in the Acylrac assay (Fig. 2B).

Protein S-acylation is mediated by ZDHHC protein S-acyltransferases (ZDHHC-PATs), which are characterized by an aspartate-histidine-histidine-cysteine (ZDHHC) core domain. The human genome encodes 23 ZDHHC proteins, the substrate specificity of which is partly dictated by their subcellular localization[12]. Using siRNAs against each one of these 23 enzymes alone or in combinations, we identified ZDHHC-17 as the PAT responsible for POU5F1B palmitoylation (Fig. 2C). Furthermore, upon zDHHC-17 knockdown we not only observed that POU5F1B partially dissociated from DRMs (Fig. 2D), but also a concomitant decrease in the overall protein content of this fraction (Fig. 2E). In sum, this first set of experiments indicates that POU5F1B ubiquitination at residues present only in the human protein allows for its retention in the cytoplasm, where its ZDHHC-17-mediated palmitoylation fosters association with detergent-resistant membrane subdomains that become secondarily enriched in other proteins.

To explore further this phenomenon, we subjected extracts of POU5F1B-expressing LS1034 CRC cells and POU5F1B-knockdown derivatives to gradient fractionation followed by Coomassie staining (Fig. 3A). It confirmed that POU5F1B results in markedly increasing the protein content of DRMs (Fig. 3A). A similar effect was observed after overexpressing wild type POU5F1B in SW480 cells, but not its $K^{135}E$ mutant (Supp. Fig. 2A). Hierarchical clustering of mass spectrometry spectral counts from input and fractions F2 and F4-5-6 extracted from control and POU5F1B-knockdown LS1034 cells revealed that POU5F1B triggered a migration of proteins from soluble to detergent-resistant membrane fractions (Fig. 3B). Subjecting this protein subset to functional enrichment analysis with gene ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) using DAVID Knowledgebase v2023q3[13,14] (Fig. 3C) pointed to cadherin binding, focal adhesion, vesicle transport, endocytosis and adherens junction as enriched features.

We went on to examine how POU5F1B affects protein association with the cell surface. For this, we performed cell-surface biotinylation coupled to precipitation with streptavidin beads on control and POU5F1B-knockdown LS1034 cells. Based on stringent filtering criteria (fold-change in the same direction with two different shRNAs; high intensity candidates –over quantile 25%-; and p value<0.05), we found 142 proteins to be enriched and 145 depleted in the surface-associated proteome of control compared with POU5F1B-depleted LS1034 cells (Fig. 3D). In their overwhelming majority (274/287), these proteins displayed no difference in total levels, indicating that POU5F1B was acting on their intracellular distribution, not on their expression or stability. Proteins enriched in the surface-associated proteome of POU5F1B^{+} cells included the cell-ECM adhesion molecules integrins ITGB1, ITGA2, ITGA1, ITGAV, and ITGA9, SLC3A2, which modulates integrin-related signaling and is essential for integrin-dependent cell spreading and migration[15], the cell adhesion molecule EPCAM, the lipoprotein receptor LRP6, the receptor tyrosine kinase (RTK) MET and the signaling receptor NOTCH1, with GO/KEGG pointing to integrin complex, basal plasma membrane, and ECM-receptor interaction as features enhanced in this setting (Fig. 3D). Western blot analysis of total and active form of ITGB1 in the LS1034 knockdown and SW480 overexpression systems confirmed the enrichment of this cell-ECM adhesion molecule in DRMs of POU5F1B-expressing cells (Supp. Fig. 2B). By contrast, cadherin binding and motor proteins were singled out as depleted features in the surface-associated proteome of POU5F1B-knockdown cells, with dynein (DYNC1LI1, DYNLRB2, DYNLRB1), dynactin (DCTN2, DCTN1) and myosin (MYO1E, MYO1B, MYO6, MYL3), all cytoskeletal motor proteins important for vesicle trafficking[16], standing prominently amongst proteins redistributed away from the surface of these cells (Fig. 3D). The 13 proteins displaying changes in both total and surface-associated proteomes in the presence of POU5F1B included VPS4B and MYOE1, involved in endosomal and membrane trafficking (surface down, SILAC up), SLC3A2 (surface up, SILAC down), and MET (surface up, SILAC up) (Supp. Fig. 2C). Overall, these data suggests that POU5F1B induces cell membrane remodeling by reshaping the cell surface-associated proteome with concentration of key effectors of cell-cell and cell-ECM adhesion and signaling molecules notably in DRMs.

DRMs are raft-like membrane nanodomains typically enriched in sphingolipids such as sphingomyelin (SM) and cholesterol. To test whether the distribution of these lipids was altered by POU5F1B expression, we labeled non-permeabilized control and POU5F1B-overexpressing SW480 cells with OlyA-mCherry, derived from the oyster mushroom *Pleurotus ostreatus* pore-forming ostreolysin, which selectively recognizes membrane regions where cholesterol and SM are jointly concentrated[17], and with D4-mCherry, a perfringolysin O derivative that binds to accessible cholesterol[18] (Supp. Fig. 2D). Both stainings revealed significant differences in the lipid distribution of control and POU5F1B-overexpressing cells,

with the latter characterized by a salient submembrane punctate pattern with OlyA-mCherry and an extended opalescent framing of the cell periphery with D4-mCherry (Supp. Fig. 2D). Although the morphological bases of these characteristics remain to be defined, they support a model whereby POU5F1B-induced remodeling of plasma membranes extends to their lipid constituents.

Integrins are a major component of focal adhesion, and we found them to be amongst the most upregulated proteins at the surface of POU5F1B expressing cells and in DRMs (Fig. 3D). Accordingly, we found p-Y^{397}FAK, the form of the focal adhesion kinase that becomes autophosphorylated upon integrin clustering[19], to display a discrete focal staining in wild-type, POU5F1B-expressing LS1034 cells (Fig. 4A) and in POU5F1B lentivector-transduced SW480 cells (Supp. Fig. 3A), contrasting with a more diffuse pattern in their respective POU5F1B-low counterparts. Staining for vinculin, a component of focal adhesions involved in linking integrins to the actin cytoskeleton, revealed a similarly focal localization in POU5F1B-expressing compared with control cells (Fig. 4B, Supp. Fig. 3B). High-resolution light microscopy images of the cell surface confirmed enrichment for ITGB1 in POU5F1B-overexpressing SW480 cells compared to control cells (Fig. 4C), and further documented the accumulation of a POU5F1B-GFP fusion protein beneath the surface of overexpressing SW480 cells (Fig. 4D), colocalized with phospho-FAK(Y^{397}) (Fig. 4D, right and bottom). Taken together, these data indicate that POU5F1B stimulates the formation of sharp focal adhesions, an event crucial for cancer cell migration and metastasis[22].

POU5F1B is not known to carry out enzymatic activities, and as such would long have been considered a non-druggable target. However, approaches aimed at triggering the proteasomal degradation of pathogenic effectors have changed this perspective, as exemplified by the thalidomide-, lenalidomide-, or pomalidomide-induced degradation of IKZF1 and IKZF3[22–25] for the treatment of multiple myeloma[26]. To explore this avenue, we searched for direct and indirect POU5F1B degraders using a recently described gain-of-signal cell-based assay[27]. For this, we established through lentivector-mediated transgenesis SW480 cell derivatives expressing POU5F1B fused to deoxycytidine kinase (DCK), an enzyme that converts the non-natural nucleoside 2-bromovinyldeoxyuridine (BVdU) into a poison[28], together with green fluorescence protein (GFP). Wild type cells, cells expressing unfused POU5F1B or unfused DCK served as controls (Supp. Fig. 4A). We first confirmed that SW480 cells expressing POU5F1B-DCK or unfused DCK were more sensitive to BVdU than SW480 cells expressing POU5F1B alone or wild type cells (Supp. Fig. 4A). We then exposed POU5F1B-DCK cells in duplicate to a library of 5,758 known bioactive compounds, ~25% of them FDA approved, for 24 hours before adding 200 µM BVdU to the culture in order to kill all cells but those in which degradation of the POU5F1B-DCK fusion protein had been induced by the pre-treatment. This first step of our screen singled out 34 compounds with a Z-score greater than the mean of the

negative controls plus three standard deviations (Supp. Fig. 4B), which were then used to perform a counter-screen with dose response curves in both POU5F1B-DCK- and DCK-expressing SW480 cells. At the outset of this second step, only three drugs were found to protect POU5F1B-DCK- but not DCK-expressing cells from BVdU-induced death: the two Rho kinase (ROCK) inhibitors capivasertib (or AZD5363) and OXA-06, and laurocapram, which acts on the permeability of lipid membranes (Fig. 5A). Importantly, the three compounds exerted negligible cytotoxic effect on control SW480 cells (Fig. 5A).

Pulse-chase experiments with ^{35}SCys/Met revealed that the half-life of POU5F1B was 60 min in SW480 cells and increased to 160 min in the presence of the proteasome inhibitor MG132 (Fig. 5B). It indicates that POU5F1B is degraded in the proteasome, consistent with our finding that it can be immunoprecipitated with two E3 ubiquitin-protein ligases, UBR5 and RNF213[1], and that it is a ubiquitinated protein (Fig. 1B). Most importantly we observed that upon treatment with capivasertib, OXA-06 or Y-27632, another ROCK inhibitor, the half-life of POU5F1B dropped to 30 min, whereas laurocapram has no effect in this assay (Fig. 5B, Supp. Fig. 4C). Thus, ROCK inhibition accelerates POU5F1B proteasomal degradation. This prompted us to ask whether POU5F1B was phosphorylated by the ROCK1/ROCK2 serine/threonine kinases, which as POU5F1B are involved in regulating cytoskeletal dynamics and focal adhesions. Probing POU5F1B immunoprecipitates by Western blot with phosphoserine-specific antibodies confirmed that the retrogene product was subjected to this post-translational modification, the intensity of which was decreased upon ROCK inhibition (Fig. 5C). Importantly, POU5F1B-triggered membrane reorganization was reverted by 24h treatment with the OXA-06 ROCK inhibitor (Fig. 5D).

In sum, the present work demonstrates that the pro-oncogenic action of POU5F1B depends on a cascade of ubiquitination, palmitoylation and phosphorylation, where the first of these PTMs target residues found only in human POU5F1B and not in its parent POU5F1 nor in POU5F1B from other great apes. Accordingly, POU5F1B represents a human-unique cancer-promoting protein, to our knowledge the first described to date. Our biochemical and functional studies further demonstrate that POU5F1B acts through unprecedented mechanisms, prompting the redistribution of cell adhesion and signaling molecules in cholesterol- and sphingomyelin-rich membrane nanodomains, correlating with increase in focal adhesion. Interestingly, we previously noted that POU5F1B-expressing cells secrete tenascin C[1], an ECM constituent that serves as a ligand for integrins. It suggests a model whereby POU5F1B-expressing cells are like the constructors of a new train line, laying in front of them Tenascin C as new rails on which to roll with integrins as wheels of the locomotive. While the mechanism by which ROCK inhibitors destabilize POU5F1B remain to be defined, it is interesting to note that Rho-ROCK signaling plays in important role in the regulation of cytoskeletal dynamics

and focal adhesions, both of which are targeted by POU5F1B action. Finally, our data point to the potential of ROCK inhibitors for the treatment of POU5F1B-expressing tumors.

METHODS

Cell culture. LS1034 and SW480 cell lines were obtained from the American Type Culture Collection (ATCC) and maintained in RPMI 1640 (Gibco) and L15 medium (Sigma) respectively, being supplemented with 10% FCS (Bioconcept 2-01F36-I). For lentiviral vector production, 293T cells were cultured in DMEM supplemented with 10% FBS with 100 IU ml^{-1} penicillin, 100 ug ml^{-1} streptomycin, and 26 µg ml^{-1} glutamine (Corning 30-009-CI) at 37°C in a humidified atmosphere of 5% CO_2. All cells tested negative for mycoplasma.

Overexpression and shRNA vectors. For the POU5F1B-expressing vector, genomic DNA from DLD1 cells (ATCC® CCL-221™) was PCR amplified with the POU5F1B primers 5'-CACCATGGCGGGACACCTGGCTTCGGATTTC-3', 5'-GTTTGAATGCATGGGAGAGCCCAG-3', cloned into a pENTR TOPO donor vector (Thermo Fisher), that was recombined with the doxycycline-inducible lentiviral destination vector pSin-TRE-3xHA-puro. For the DCK-based positive selection assay, we kindly received from William Kaelin's laboratory the pLX304-gate-DCK-V5-IRES-GFP plasmid, in which we Gateway cloned *POU5F1B*. Short hairpin RNA (shRNA) were designed with the help of the i-Score Designer website, and adapted to the miR-E shRNA structure and cloned into the LT3GEPIR pRRL backbone following Fellmann C et. al. instructions[29]: shRNA3 5'-TGCTGTTGACAGTGAGCG**CACAGGTGATTATGATTTAAAG**TAGTGAAGCCACAGATGTA**CTTTAAATCATAATCACCTGTG**TGCCTACTGCCTCGGA-3'; shRNA5 5'-TGCTGTTGACAGTGAGCG**CACATTCAGTCAACATTTAATG**TAGTGAAGCCACAGATGTA**CATTAAATGTTGACTGAATGTG**TGCCTACTGCCTCGGA-3') with a scrambled shRNA serving as negative control 5'-TGCTGTTGACAGTGAGCG**CAACAAGATGAAGAGCACCAAG**TAGTGAAGCCACAGATGTA**CTTGGTGCTCTTCATCTTGTTG**TGCCTACTGCCTCGGA-3'.

Lentivirus production and stable gene expression. Lentiviral particles produced as described at http://tronolab.epfl.ch were used to transduce SW480 and LS1034 cells in 6-well plates at multiplicity of infection 1, before selection in 1 µg ml^{-1} puromycin for 7 days. POU5F1B, GFP and lacZ overexpression in SW480 cells was verified, after three days of 500 ng ml^{-1} doxycycline activation, by western blot with HRP-conjugated anti-HA antibody (clone 3F10 Roche 12013819001, 1:1,000); and *POU5F1B* downregulation upon shRNA transduction in LS1034 was analyzed by quantitative real-time PCR using P3 pair of primers. Constitutive expression of POU5F1B-DCK and DCK proteins was verified by western blot with anti-V5 antibody.

Immunofluorescence. Cell lines were plated on glass coverslips into 24-well plates and cultured for 3 days up to 70% confluence in 500 ng ml^{-1} doxycycline-containing medium. Growth medium was refreshed (1 ml) prior a 15 min fixation with 1 ml 8% paraformaldehyde

(PFA) added dropwise, to achieve a final concentration of 4% PFA. Cells were washed three times with PBS, permeabilized with 0.05% saponin PBS for 20 min, and blocked with 1% BSA 0.05% saponin PBS for 30 min before incubation with anti-HA (bioLegend, Covance catalog #MMS-101P, 1:1,000) and anti-lamin-B1 (abcam ab16048, 1:1,000) antibodies in 1% BSA 0.05% saponin PBS 2 h at room temperature (RT) under agitation. Phospho-FAK Tyr397 (Thermofisher 700255, 1:200) and anti-vinculin clone 7F9 (sigma MAB3574, 1:200) antibodies were incubated overnight at 4 degrees in blocking solution without saponin. Samples were washed three times with PBS and incubated with Alexa 647-conjugated (A647) anti-mouse antibody (1:1,000) and A568 anti-rabbit antibody (1:1,000) in 1% BSA PBS for 40 min at RT. Three final washes were performed before mounting the slides in Vectashield with DAPI (Vector Laboratories). Images were acquired on a ZEISS LSM 700 confocal microscopy and analyzed with Fiji software.

Isolation of detergent-resistant membranes (DRMs). Approximately 1 × 10^7 cells were resuspended in 0.5 ml cold TNE buffer (25 mMTris-HCl, pH 7.5, 150 mM NaCl, 5 mM EDTA, and 1% Triton X-100; Surfact-Amps, ThermoFisher) with a tablet of protease inhibitors (Roche). Membranes were solubilized in a rotating wheel at 4 °C for 30 min. DRMs were isolated using an Optiprep™ gradient [30]: the cell lysate was adjusted to 40% Optiprep™, loaded at the bottom of a TLS.55 Beckman tube, overlaid with 600 µl of 30% Optiprep™ and 600 µl of TNE, and centrifuged for 1.5 h at 55,000 rpm at 4 °C. Six fractions of 400 µl were collected from top to bottom. DRMs were found in fraction 2. Equal volumes from each fraction were analyzed by SDS-PAGE and western blot analysis using HRP-conjugated anti-HA, caveolin1 (Santa Cruz sc-894, 1:500) and transferrin receptor (ThermoFisher 13-6800, 1:1,000) antibodies.

***In vitro* proliferation assays.** Cell proliferation was determined with PrestoBlue™ reagent (A13262, Invitrogen) for 6 days.

Acyl-RAC capture assay. Protein S-palmitoylation was assessed by the Acyl-RAC assay as previously described (Werno and Chamberlain, 2015) with some modifications. Cells or supernatants were lysed in 400 ml buffer (0.5% Triton-X100, 25 mM HEPES, 25 mM NaCl, 1 mM EDTA, pH 7.4,and protease inhibitor cocktail). Cell lysis were incubated 30 min at RT with 10mM TCEP. Then, 200 ml of blocking buffer (100 mMHEPES, 1 mM EDTA, 87.5 mM SDS, and 1.5% [v/v] methyl methanethiosulfonate (MMTS)) was added to the lysates and incubated for 4 h at 40ºC to block free the SH groups with MMTS. Proteins were acetone precipitated and resuspended in buffer (100 mMHEPES, 1 mM EDTA, 35 mM SDS). For treatment with hydroxylamine (NH2OH) and capture by Thiopropyl Sepharose beads, 2 M of hydroxylamine was added together with the beads (previously activated for 15 min with water) to a final concentration of 0.5 M of hydroxylamine and 10% (w/v) beads. As a negative control, 2 M Tris was used instead of hydroxylamine. These samples were then incubated overnight at room

temperature on a rotating wheel. After washes, the proteins were eluted from the beads by incubation in 40 ml SDS sample buffer with ß-mercaptoethanol for 5 min at 95ºC. Finally, samples were separated by SDS-PAGE andanalyzed by Western blot. A fraction of the cell lysate (total cell extract) was saved as the input.

Radiolabeling 3 H-palmitic acid incorporation. To follow S-acylation, lentiviral vector transduced and 96h doxycycline activated cells were incubated 1 hour in L15 medium without serum, followed by 2 hours at 37ºC in L15 with 200 mCI /ml 3 H palmitic acid, washed with L15 and incubated different 2 hours at 37ºC with L15 prior overnight anti-HA immunoprecipitation. Beads were incubated 5 min at 90ºC in reducing sample buffer prior to SDS-PAGE. Immunoprecipitates were split into two, run on 4-20% gels and analyzed either by autoradiography (3 H-palmitate) after fixation (25% isopropanol, 65% H2O, 10% acetic acid), incubation 30 min in enhancer Amplify NAMP100, and drying; or western blot.

Transfection of siRNA. SW480 cells were transfected with siRNA (Qiagen, Supplementary Table S1) (15 pmol/ 9.6cm2 plate) using Lipofectamine RNAiMAX transfection reagent. Cells were harvested 72h post transfection and whole cell protein lysates were analyzed by western blot.

LC-MS/MS analysis of DRM fractions. Input and DRM fraction 2 (F2) and equal volume mixes of fractions 4, 5, and 6 (F456) from shRNA-scramble, shRNA3 and shRNA5-transduced LS1034 cells were analyzed by Liquid Chromatography with tandem mass spectrometry (LC-MS-MS). Each sample was digested by Filter Aided Sample Preparation (FASP)[31] with minor modifications. Dithiothreitol (DTT) was replaced by Tris (2-carboxyethyl)phosphine (TCEP) as reducing agent and Iodoacetamide by Chloracetamide as alkylating agent. A combined proteolytic digestion was performed using Endoproteinase Lys-C and Trypsin. Peptides were desalted on SDB-RPS StageTips[32] and dried down by vacuum centrifugation. Samples were then fractionated into 12 fractions using an Agilent OFFGEL 3100 system. Resulting fractions were desalted on SDB-RPS StageTips and dried by vacuum centrifugation. For LC MS/MS analysis, peptides were resuspended and separated by reversed-phase chromatography on a Dionex Ultimate 3000 RSLC nanoUPLC system in-line connected with an Orbitrap Lumos Fusion Mass-Spectrometer. Database search was performed using MaxQuant 1.6.10.43[33] against concatenated database consisting of the UniProt human database (Uniprot release 2019_06, 74468 sequences) and common fetal bovine serum protein[34]. Carbamidomethylation was set as fixed modification, whereas oxidation (M), phosphorylation (S,T,Y), Gln to pyro-Glu and acetylation (Protein N-term) were considered as variable modifications. Raw spectral counts extracted with MaxQuant were subjected to log2 transformation, and any missing values were imputed using the 'min' function available in the DEP package[35]. Proteins that exhibited fewer than 10 log2-transformed total spectral counts across all samples were excluded for further analysis. To identify proteins undergoing

significant changes between conditions of interest, a statistical linear model with interactions was applied using the limma R package[36]. Proteins that demonstrated notably different behavior between the 'scramble' and 'sh' conditions in F2 compared to their F456 counterparts, and exhibited a foldChange of less than 0.5 log2 in all 'inputs', were selected for further analysis.

Cell surface protein studies. Surface biotinylation assays were performed as previously described[37]. Briefly, cells were shifted to ice and incubated for 30 min with cold biotinylation solution (0.2 mg/ml sulfo-NHS-SS-biotin (Pierce) diluted in PBS -0.17-mg/mL final concentration-), then quenched three times with cold 100 mM NH4Cl, and lysed in IP buffer. The lysate was immunoprecipitated overnight at 4°C with prewashed streptavidin-coated beads (Sigma S1638), beads were washed with 1 mL 1% SDS, 150 mM NaCl, 1 mM EDTA in 50 mM HEPES pH 7.4 (×3) and 50 mM HEPES pH 8.0 (×2), and proteins were eluted from beads with 50 µL of 50 mM HEPES pH 8.0 containing 50-100 mM TCEP for 1 h. Peptides containing supernatants were digested overnight at 37 ºC in the thermoshaker with 1-2 µL Trypsin (20 µg /100 µL) in 50 mM HEPES pH 8. Samples were analyzed by LC-MS/MS as described above. Surface-associated proteome candidates were selected when the three experimental replicates showed p value<0.05 (outlier detection test as computed by MaxQuant), high intensity (above quantile 25%), and fold changes in the same direction for the two shRNAs.

DAVID enrichment analysis. For DRM and surface associated-proteome data sets, we obtained a list of differentially expressed candidates that were selected as detailed above. Resulting candidate lists were subjected to GO and KEGG functional annotation chart using the online bioinformatics resource DAVID[13,14].

High resolution microscopy imaging was performed as previously described[38]. Briefly, POU5F1B-, POU5F1B-GFP-, and lacZ-expressing SW480 cells were seeded on ibidi eight-well high glass-bottom chambers (no. 80807) at a density of 30,000 cm–2. Cells were fixed with 2.4% paraformaldehyde in PBS for 30 min at room temperature. After fixation, cells were washed three times with PBS. Blocking and permeabilization were performed with 0.05% saponin in blocking buffer for 30 min. After washing with PBS, cells were incubated with 100 nM anti-GFP nanobodies, anti-ITGB1 and anti-phospho FAK antibodies in blocking buffer without saponin for 60 min at room temperature. Unbound nanobodies were removed by washing three times with PBS. Fluorescence imaging was carried out as previously described (https://doi.org/10.1038/s41586-023-05925-9)[38].

Low-throughput BVdU treatment curves. SW480 cells stably transduced with bicistronic lentiviruses encoding (i) DCK, or POU5F1B-DCK and (ii) GFP, as well as POU5F1B-expressing and wild type SW480 control cells, were seeded into 96-well plates at 2000 cells per well in 100 µL of media. The next day, 1 M BVdU dissolved in DMSO was diluted into

media to prepare 6× stock solutions of BVdU at concentrations of 640 mM, 320 mM, 160 mM, 80 mM, 40 mM and 4 mM. For each stock solution, DMSO concentration was adjusted to a final concentration of 0.25%. 0.3 µL/well of the 6× stock solutions were dispensed in the 6 first rows of the 96-well dish to achieve final concentrations of 1600 µM, 800 µM, 400 µM, 200 µM, 100 µM and 10 µM respectively. Media with 0.25% DMSO was added to the seventh row as a control. Four days later, cell viability was assessed with PrestoBlue™ reagent (A13262, Invitrogen) following manufacturer's instructions.

High-throughput screening using the DCK*-based positive selection assay. High-throughput chemical library screening were performed as previously described[27]. Briefly, POU5F1B-DCK SW480 cells were seeded into 384-well plates (Corning, #3764) at a density of 750 cells per well in a volume of 30 µl of media. Plates already contained a library of 5,758 lyophilized known bioactive compounds, to which cells were exposed for 24 hours before adding BVdU at a final concentration of 200 uM into columns 2 to 24. After 4 days, the GFP fluorescence of each assay plates was quantified using an Acumen scanning laser cytometer. For each plate, the average and SD of the GFP fluorescence of wells in columns 3 to 22 were calculated. For each well on an assay plate, the GFP fluorescence was converted to a z score using the formula: $z(well) = [GFP (well) - \mu GFP (plate)] / \sigma GFP (plate)$, where μ GFP (plate) is the mean GFP fluorescence for that plate and σ GFP (plate) is the SD for that plate.

Radiolabeling 35S-cys/met incorporation. For metabolic labeling, transduced cells were washed with methionine /cysteine free medium, incubated 4hours pulse at 37°C with 50 mCi/ml 35S-methionine/cysteine, washed and further incubated for different times at 37°C in complete medium with a 10-fold excess of non-radioactive methionine and cysteine. Proteins were immunoprecipitated and analyzed by SDS-PAGE. Autoradiography and western blot were quantified using the Typhoon Imager (Image QuantTool, GE healthcare).

REFERENCES

1. Simó-Riudalbas, L. *et al.* Transposon-activated POU5F1B promotes colorectal cancer growth and metastasis. *Nat. Commun.* **13**, 4913 (2022).
2. Hayashi, H. *et al.* The OCT4 pseudogene POU5F1B is amplified and promotes an aggressive phenotype in gastric cancer. *Oncogene* **34**, 199–208 (2015).
3. Pan, Y. *et al.* POU5F1B promotes hepatocellular carcinoma proliferation by activating AKT. *Biomed. Pharmacother. Biomedecine Pharmacother.* **100**, 374–380 (2018).
4. Chen, L. *et al.* HPV-16 E2/E6 and POU5F1B as Biomarkers to Determine Cervical High-Grade Squamous Lesions and More. *J. Inflamm. Res.* **13**, 813–821 (2020).
5. Hu, Z. *et al.* Genome-wide profiling of HPV integration in cervical cancer identifies clustered genomic hot spots and a potential microhomology-mediated integration mechanism. *Nat. Genet.* **47**, 158–163 (2015).
6. Pickart, C. M. & Eddins, M. J. Ubiquitin: structures, functions, mechanisms. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **1695**, 55–72 (2004).
7. Anwar, M. U. *et al.* ER-Golgi-localized proteins TMED2 and TMED10 control the formation of plasma membrane lipid nanodomains. *Dev. Cell* **57**, 2334-2346.e8 (2022).
8. Levental, I., Levental, K. R. & Heberle, F. A. Lipid Rafts: Controversies Resolved, Mysteries Remain. *Trends Cell Biol.* **30**, 341–353 (2020).
9. Hyer, M. L. *et al.* A small-molecule inhibitor of the ubiquitin activating enzyme for cancer treatment. *Nat. Med.* **24**, 186–193 (2018).
10. Levental, I., Grzybek, M. & Simons, K. Greasing Their Way: Lipid Modifications Determine Protein Association with Membrane Rafts. *Biochemistry* **49**, 6305–6316 (2010).
11. Melkonian, K. A., Ostermeyer, A. G., Chen, J. Z., Roth, M. G. & Brown, D. A. Role of Lipid Modifications in Targeting Proteins to Detergent-resistant Membrane Rafts: MANY RAFT PROTEINS ARE ACYLATED, WHILE FEW ARE PRENYLATED *. *J. Biol. Chem.* **274**, 3910–3917 (1999).
12. Stix, R., Lee, C.-J., Faraldo-Gómez, J. D. & Banerjee, A. Structure and Mechanism of DHHC Protein Acyltransferases. *J. Mol. Biol.* **432**, 4983–4998 (2020).
13. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
14. Sherman, B. T. *et al.* DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res.* **50**, W216–W221 (2022).
15. Feral, C. C. *et al.* CD98hc (SLC3A2) mediates integrin signaling. *Proc. Natl. Acad. Sci.* **102**, 355–360 (2005).

16. Gross, S. P., Vershinin, M. & Shubeita, G. T. Cargo Transport: Two Motors Are Sometimes Better Than One. *Curr. Biol.* **17**, R478–R486 (2007).

17. Skočaj, M. *et al.* Tracking Cholesterol/Sphingomyelin-Rich Membrane Domains with the Ostreolysin A-mCherry Protein. *PLOS ONE* **9**, e92783 (2014).

18. Ramachandran, R., Heuck, A. P., Tweten, R. K. & Johnson, A. E. Structural insights into the membrane-anchoring mechanism of a cholesterol-dependent cytolysin. *Nat. Struct. Biol.* **9**, 823–827 (2002).

19. Hanks, S. K. & Polte, T. R. Signaling through focal adhesion kinase. *BioEssays News Rev. Mol. Cell. Dev. Biol.* **19**, 137–145 (1997).

20. Seetharaman, S. & Etienne-Manneville, S. Cytoskeletal Crosstalk in Cell Migration. *Trends Cell Biol.* **30**, 720–735 (2020).

21. Massagué, J. & Ganesh, K. Metastasis-Initiating Cells and Ecosystems. *Cancer Discov.* **11**, 971–994 (2021).

22. Ito, T. *et al.* Identification of a Primary Target of Thalidomide Teratogenicity. *Science* **327**, 1345–1350 (2010).

23. Krönke, J. *et al.* Lenalidomide causes selective degradation of IKZF1 and IKZF3 in multiple myeloma cells. *Science* **343**, 301–305 (2014).

24. Lu, G. *et al.* The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins. *Science* **343**, 305–309 (2014).

25. Lopez-Girona, A. *et al.* Cereblon is a direct protein target for immunomodulatory and antiproliferative activities of lenalidomide and pomalidomide. *Leukemia* **26**, 2326–2335 (2012).

26. Zhu, Y. X. *et al.* Cereblon expression is required for the antimyeloma activity of lenalidomide and pomalidomide. *Blood* **118**, 4771–4779 (2011).

27. Koduri, V. *et al.* Targeting oncoproteins with a positive selection assay for protein degraders. *Sci. Adv.* **7**, eabd6263 (2021).

28. Neschadim, A. *et al.* Cell Fate Control Gene Therapy Based on Engineered Variants of Human Deoxycytidine Kinase. *Mol. Ther.* **20**, 1002–1013 (2012).

29. Fellmann, C. *et al.* An optimized microRNA backbone for effective single-copy RNAi. *Cell Rep.* **5**, 1704–1713 (2013).

30. Harder, T., Scheiffele, P., Verkade, P. & Simons, K. Lipid Domain Structure of the Plasma Membrane Revealed by Patching of Membrane Components. *J. Cell Biol.* **141**, 929–942 (1998).

31. Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359–362 (2009).

32. Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N. & Mann, M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat. Methods* **11**, 319–324 (2014).

33. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).

34. Shin, J. *et al.* Use of Composite Protein Database including Search Result Sequences for Mass Spectrometric Analysis of Cell Secretome. *PLOS ONE* **10**, e0121692 (2015).

35. Zhang, X. *et al.* Proteome-wide identification of ubiquitin interactions using UbIA-MS. *Nat. Protoc.* **13**, 530–550 (2018).

36. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

37. Sergeeva, O. A. & van der Goot, F. G. Anthrax toxin requires ZDHHC5-mediated palmitoylation of its surface-processing host enzymes. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 1279–1288 (2019).

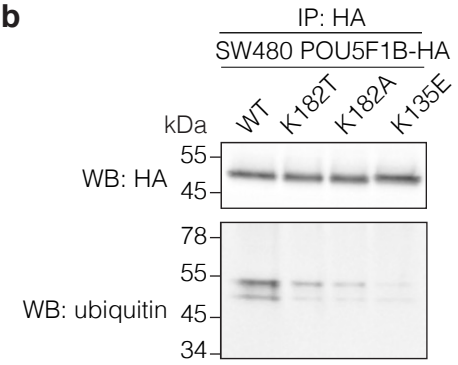
38. Reinhardt, S. C. M. *et al.* Ångström-resolution fluorescence microscopy. *Nature* **617**, 711–716 (2023).

Figure 1

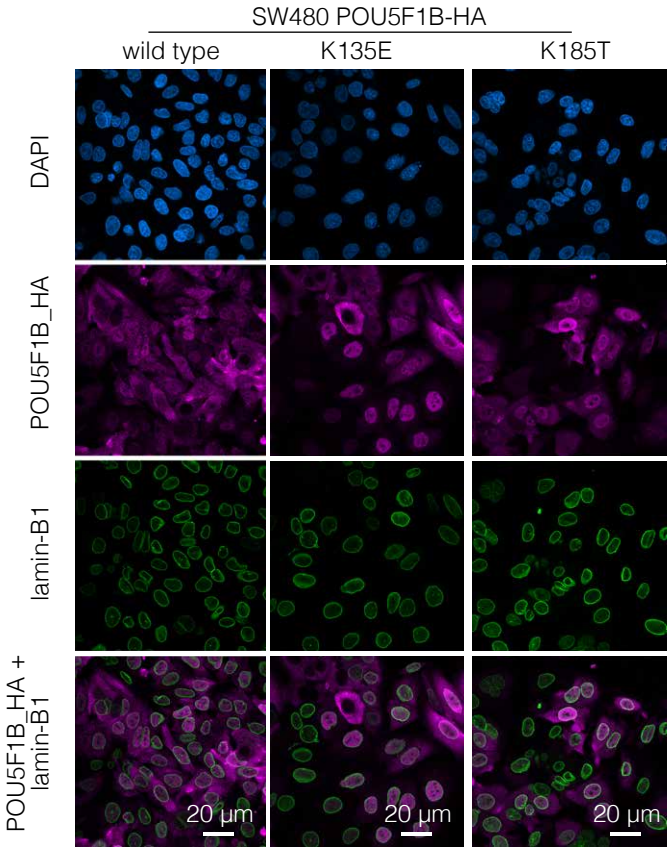
a

```
human_POU5F1        MAGHLASDFAFSPPPGGGGDGPGGPEPGWVDPRTWLSFQGPPGGPGIGPGVGPGSEVWGI  60
human_POU5F1B       MAGHLASDFAFSPPPGGGGDGPWGAEPGWVDPLTWLSFQGPPGGPGIGPGVGPGSEVWGI  60
gorilla_POU5F1B     MAGHLASDFAFSPPPGGGGDGPGGTEPGWVDPLTWLSFQGPPGGPGIGPGAGPGSEVWGI  60
orangutan_POU5F1B   MAGHLASDFAFSPPPGGGGDGPGGTEPGWVDPLTWLSFQGPPGGPGIGPGAGPGSEVWGI  60
chimpanzee_POU5F1B  MAGHLASDFAFSPPPGGEGDGPGGPEPGWVDPLTWLRFQGPPGGPGIGPGVGPGSEVWGI  60
                    ***************** **** * ******* *** *************.*********

human_POU5F1        PPCPPPYEFCGGMAYCGPQVGVGLVPQGGLETSQPEGEAGVGVESNSDGASPEPCTVTPG  120
human_POU5F1B       PPCPPPYELCGGMAYCGPQVGVGLVPQGGLETSQPESEAGVGVESNSNGASPEPCTVPPG  120
gorilla_POU5F1B     PPCPPPYEFCRGMAYCGPQVGVGLVPQGGLETSQPEGEAGVGVESNSNGASPEPCTVPPG  120
orangutan_POU5F1B   PPCPPPYEFCRGMAYCGPQVGVGLVPQGGLETSQPEGEAGVGVESNSNGASPEPCTVPPG  120
chimpanzee_POU5F1B  PPCPPPYEFCGGMAYCGPQVGAGLVPQGGLETSQPEGEAGVGVESNSDGASPEPCTVPPG  120
                    ********:* **********.**************.**********:********* **

human_POU5F1        AVKLEKEKLEQNPEESQDIKALQKELEQFAKLLKQKRITLGYTQADVGLTLGVLFGKVFS  180
human_POU5F1B       AVKLEKEKLEQNPEKSQDIKALQKELEQFAKLLKQKRITLGYTQADVGLILGVLFGKVFS  180
gorilla_POU5F1B     AVKLEKEKLEQNP-ESQDIKALQKELEQFAKLLKQKRITLGYTQADVGLILGVLFGKVFS  179
orangutan_POU5F1B   AVKLEKEKLEQNP-ESQDIKALQKELEQFAKLLKQKRITLGYTQADVGLILGVLFGKVFS  179
chimpanzee_POU5F1B  AVKLEKEKLEQNPEESQDIKALQKELEQFAKLLKQKRITLGYTQADVGLILGVLFGKVFS  180
                    ************* :********************************** **********

human_POU5F1        QTTICRFEALQLSFKNMCKLRPLLQKWVEEADNNENLQEICKAETLVQARKRKRTSIENR  240
human_POU5F1B       QKTICRFEALQLSFKNMCKLRPLLQKWVEEADNNENLQEICKAETLMQARKRKRTSIENR  240
gorilla_POU5F1B     QTTICCFEALQLSFKNMCKLRPLLQKWVEEADNNENLQEICKAESLMQARKRKRTSIENR  239
orangutan_POU5F1B   QTTICCFEALQLSFKNMCKLRPLLQKWVEEADNNENLQEICKAESLMQARKRKRTSIENR  239
chimpanzee_POU5F1B  QTTICRFEALQLSFKNMCKLRPLLQKWVEETDNNENLQEICKAETLMQARKRKRTSIENR  240
                    *.*** ************************:*************:*:*************

human_POU5F1        VRGNLENLFLQCPKPTLQQISHIAQQLGLEKDVVRVWFCNRRQKGKRSSSDYAQREDFEA  300
human_POU5F1B       VRGNLENLFLQCPKPTL-QISHIAQQLGLEKDVVRVWFCNRRQKGKRSSSDYAQREDFEA  299
gorilla_POU5F1B     VRGNLENLFLQCPKPTL-QISHIAQQLGLEKDVVRVWFCNRRQKGKRSSSDYAQREDFEA  298
orangutan_POU5F1B   VRGNLENLFLQCPKPTL-QISHIAQQLGLEKDVVRVWFCNRRQKGKRSSSDYAQREDFEA  298
chimpanzee_POU5F1B  VRGNLENLFLQCPKPTL-QISHIAQQLGLEKDVVRVWFCNRRQKGKRSSSDYEQREDFEA  299
                    ***************** ******************************** *******

human_POU5F1        AGSPFSGGPVSFPLAPGPHFGTPGYGSPHFTALYSSVPFPEGEAFPPVSVTTLGSPMHSN  360
human_POU5F1B       AGSPFSGGPVSFPPAPGPHFGTPGYGSPHFTALYSSVPFPEGEVFPPVSVITLGSPMHSN  359
gorilla_POU5F1B     AGSPFSGGPVSFPPAPGPHFGTPGYGSPHFTALYSSVPFPEGEVFPPVSVITLGSPMHSN  358
orangutan_POU5F1B   AGSPFSGGPVSFPPAPGPHFGTPGYGSPHFTALYSSVPFPEGEVFPPVSVITLGSPMHSN  358
chimpanzee_POU5F1B  AGSPFSGGPVSFPLAPGPHFGTPGYGSPHFTALYSSVPFPEGEVFPPVSVITLGSPMHLN  359
                    ************* ****************************.****** ******* *
```

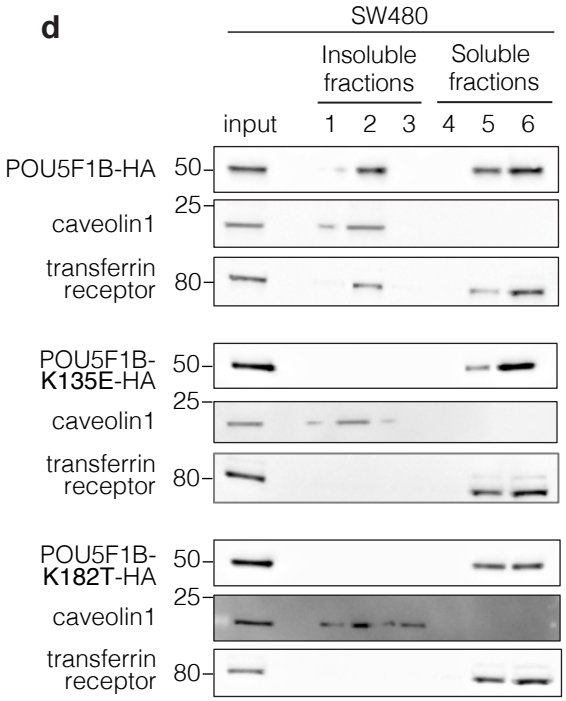
b



c



d



e

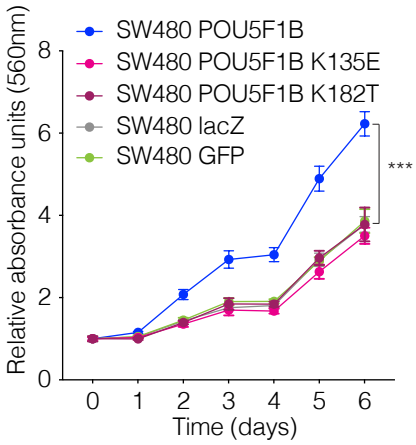


Figure 1. POU5F1B is a human specific oncogene. a, Alignment of human POU5F1 (360 aa), POU5F1B (359 aa), and gorilla, organgutan and chimpanzee POU5F1B (358, 358 and 359 aa respectively) protein sequences. POU5F1B human-specific amino acids are highlighted in pink, POU5F1B-specific amino acids in blue. **b,** Immunoprecipitation (IP) and western blot (WB) analysis of HA-tagged wild type (WT)-, K182T-, K182A-, and K135E-POU5F1B in SW480 cells. **c,** Representative immunofluorescence-confocal microscopy of POU5F1B-HA wild type-, K135E-, and K182T-overexpressing SW480 cells, with DAPI in blue, HA in pink and the nuclear membrane marker lamin-B1 in green. **d,** Isolation of detergent-resistant membranes (DRMs) from POU5F1B-, POU5F1B-K135E-, and POU5F1B-K182T-overexpressing SW480 cells. 1-3 correspond to insoluble fractions, 4-6 to soluble fractions, DRMs being traditionally found in fraction 2. Caveolin1 and transferrin receptor are used as controls for insoluble and soluble fractions, respectively (representative blot out of three independent experiments). **e,** 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay in POU5F1B-, POU5F1B K135E-, POU5F1B K182T-, lacZ-, and GFP-overexpressing SW480 cells.

Figure 2

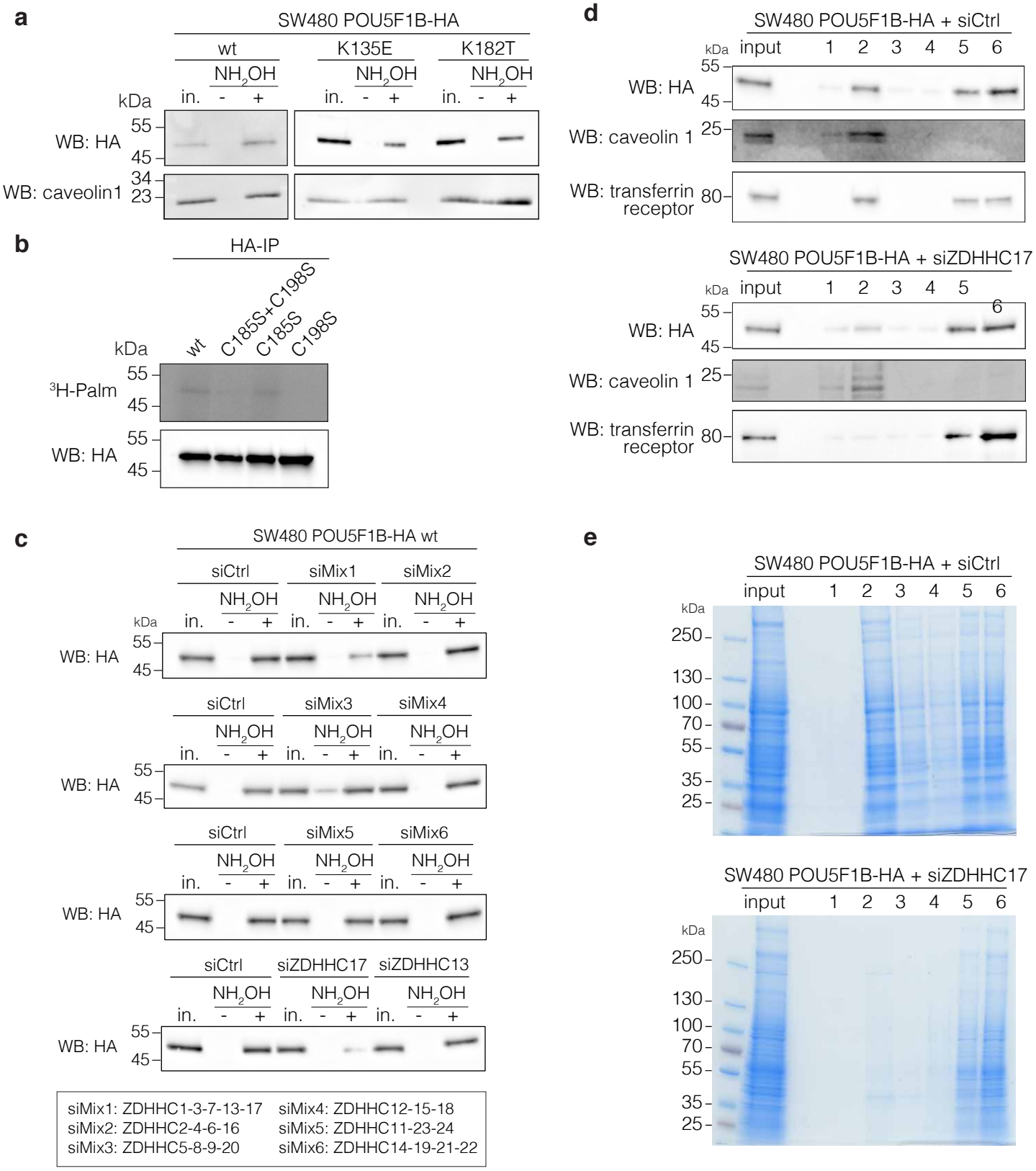


Figure 2. POU5F1B palmitoylation changes its subcellular localization. a, Acylrac capture assay in SW480 expressing POU5F1B-HA WT. Western blot (WB) of total cell extracts (in), S-acylated (plus hydroxylamine +NH_2OH) and control fractions (–NH_2OH). **b,** WB and autoradiography (^{3}H-Palm) of immunoprecitated HA-tagged wild type (wt)-, C185S-, C198S-POU5F1B in ^{3}H-palmitic acid-labeled SW480 cells. **c,** Same as in **a** in SW480 POU5F1B-HA wt cells knockdown for the indicated mixes or individual ZDHHC proteins. **d,** WB and **e,** commassie stain on isolated detergent-resistant membranes (DRMs) from siCtrl-, and siZDHHC17- POU5F1B-HA SW480 cells.

Figure 3

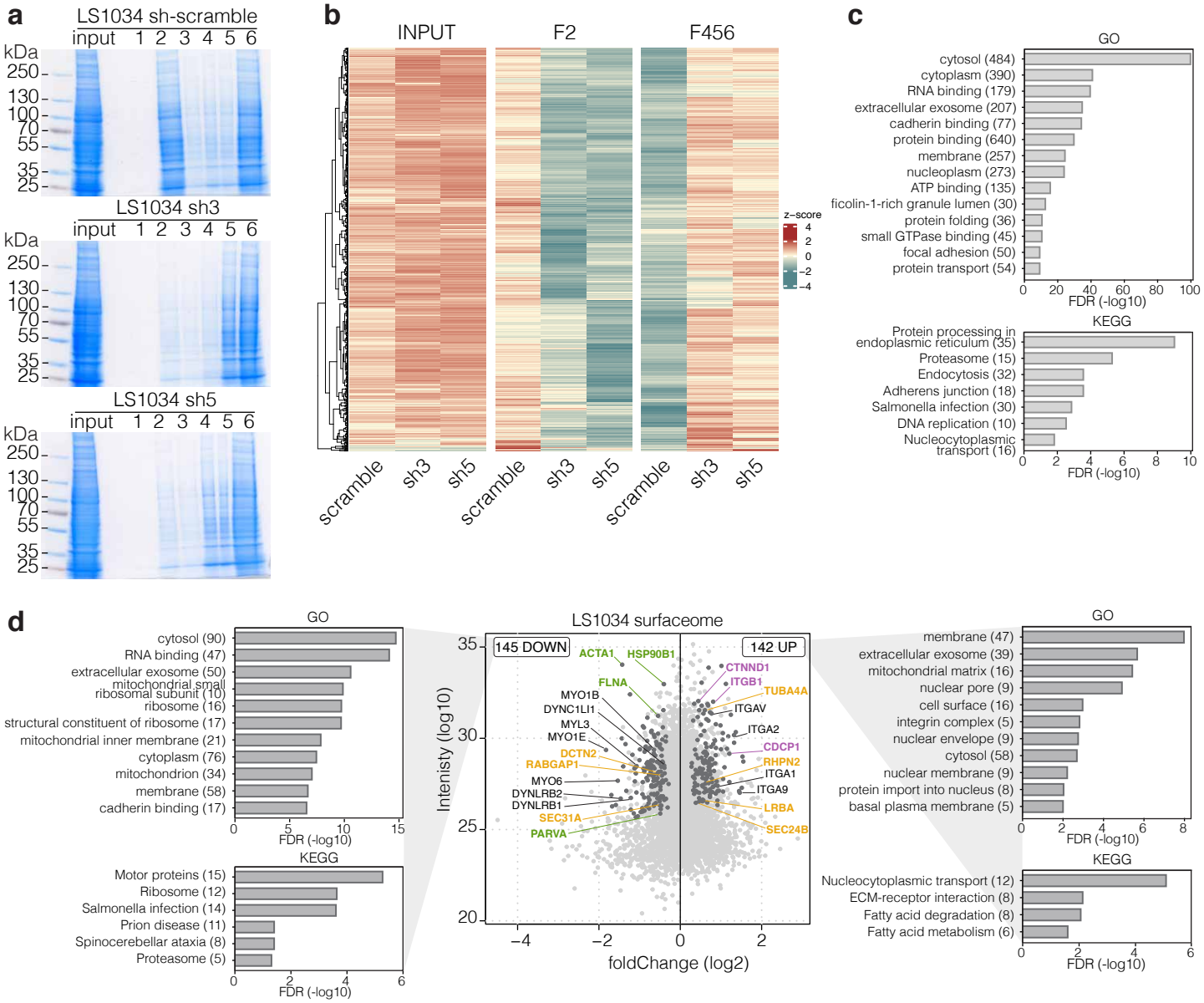


Figure 3. POU5F1B induces remodeling of cell membranes. a, Isolation of detergent-resistant membranes (DRMs) from sh-scramble, shRNA3 & shRNA5 LS1034 cells, fractions were run on SDS-PAGE and gels were stained with Coomassie. **b,** Heatmap of mass spectrometry spectral counts from input, fraction 2 (F2) and fractions 4, 5, and 6 (F456), extracted from sh-scramble, shRNA3 & shRNA5 LS1034 cells. **c,** Significantly enriched proteins in F2 sh-scramble not changing along the inputs (Sig.) were clustered in GO and KEGG annotations using the Functional Annotation Chart from DAVID bioinformatics resources 6.8. **d,** MA plot depicting relative abundance and average intensity of individual proteins identified by mass spectrometry in the surfaceome of sh-scramble vs. shRNA3 & shRNA5 LS1034 cells. Measurements were performed in independent triplicates, each dot represents a detected protein, with significantly changed ones ($P < 0.05$, outlier detection test as computed by MaxQuant) in dark grey and their numbers indicated in upper corners. Significantly enriched proteins were clustered as in **c**. Names of surface proteins also enriched in DRMs with cell adhesion (purple), vesicle trafficking (yellow), and cytoskeleton (green) functions are depicted.

Figure 4

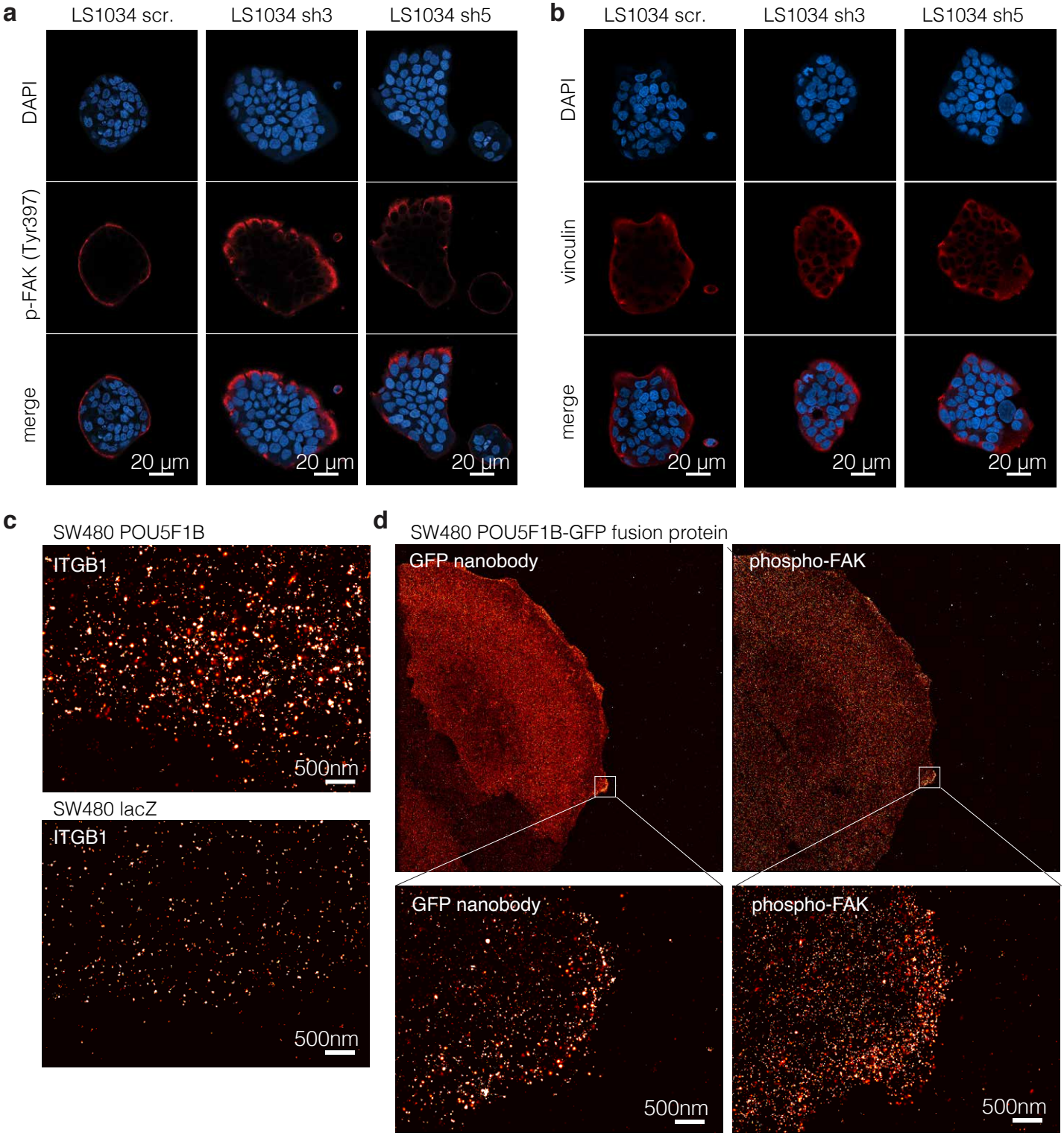


Figure 4. POU5F1B localizes at the cell surface where it sharpens focal adhesion assembly. a, Representative immunofluorescence-confocal microscopy of LS1034 cells transduced with two different POU5F1B-targeting shRNAs or a control scramble sequence, with DAPI in blue and phospho-FAK (Tyr397) in red. **b,** Same as in **a** with DAPI in blue and vinculin in red. **c,** Representative high-resolution microscopy image of POU5F1B and lacZ control overexpressing SW480 cells, immunostained with ITGB1 antibody. **d,** Representative high-resolution microscopy image of POU5F1B-GFP fusion overexpressing SW480 cells, immunostained with GFP nanobody (left), phospho-FAK (Tyr397) antibody (right), with zoom-in region and multiplexed image.

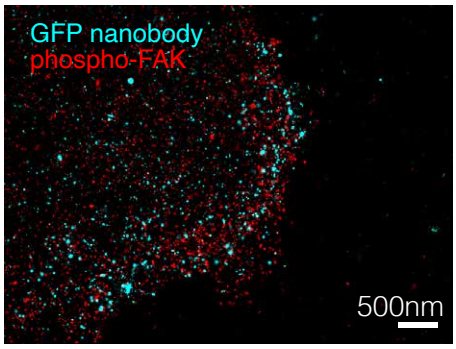


Figure 5

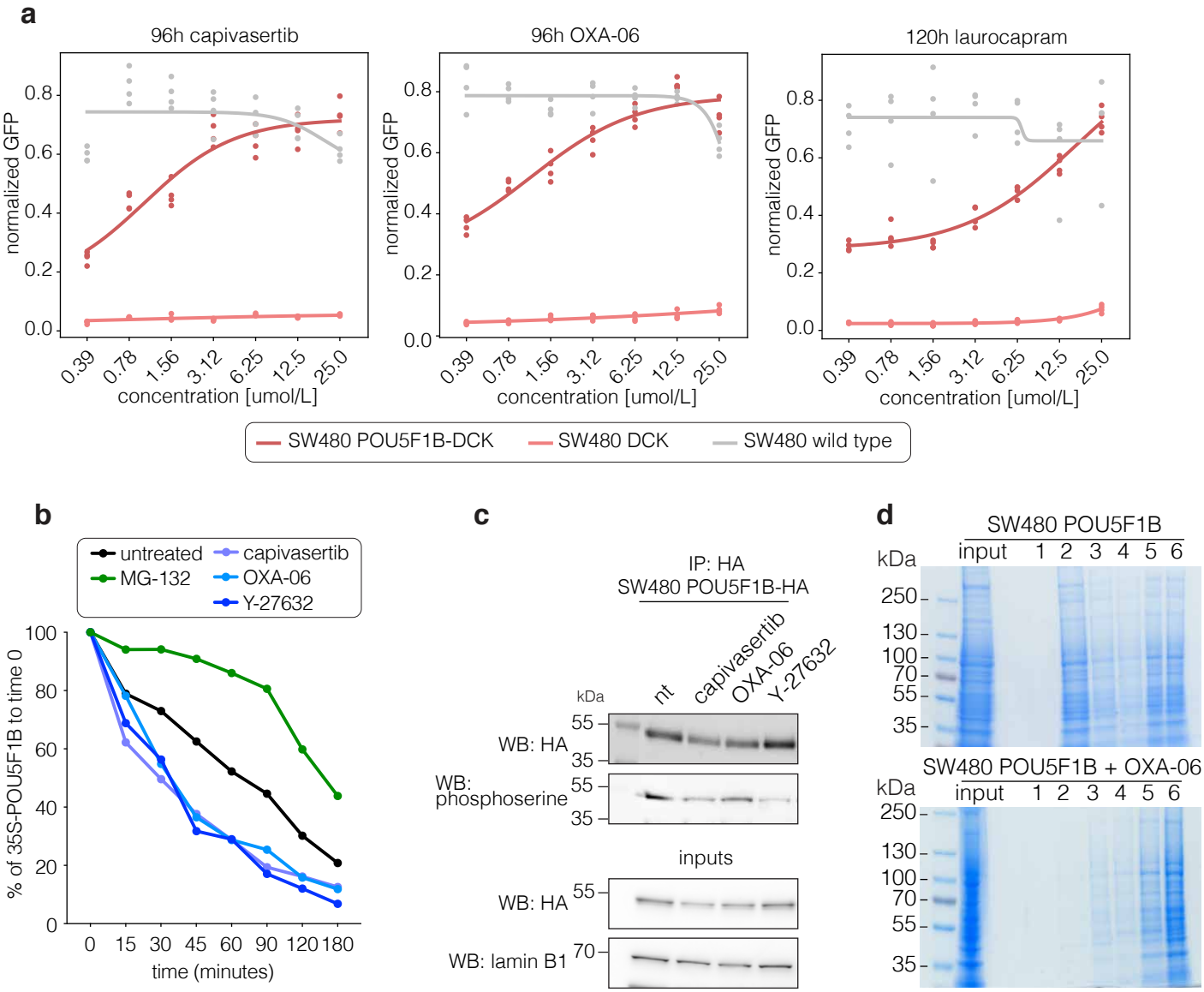


Figure 5. ROCK inhibition blocks POU5F1B function. a, Dose response curves in wild type, POU5F1B-DCK-, and DCK-expressing SW480 cells treated with increasing concentrations of capivasertib, OXA-06 and lauro-capram. **b,** Pulse-chase experiment with 35SCys/Met in POU5F1B-expressing SW480 cells untreated or treated with MG-132 (10uM), capivasertib (20uM), OXA-06 (20uM), and Y-27632 (10uM). Drugs were added at time 0 and POU5F1B-HA protein was IP-ed from harvested cell pellets at indicated times. **c,** anti-HA immunoprecipitation in SW480 POU5F1B-HA-overexpressing cells, untreated (nt) and treated 24h with capivasertib (20uM), OXA-06 (20uM), and Y-27632 (10uM), and blotted with anti-phosphoserine antibody. Inputs blotted with anti-HA and anti-laminB1 are depicted. **d,** Isolation of detergent-resistant membranes (DRMs) from untreated and OXA-06 (20uM, 24h) treated POU5F1B-expressing SW480 cells, fractions were run on SDS-PAGE and gels were stained with Coomassie.