GENETICS

Phenotypic drug screening in a human fibrosis model identified a novel class of antifibrotic therapeutics

Michael Gerckens[1], Kenji Schorpp[2]†, Francesco Pelizza[3]†, Melanie Wögrath[1,4], Kora Reichau[5,6], Huilong Ma[5,6], Armando-Marco Dworsky[1,4], Arunima Sengupta[1], Mircea Gabriel Stoleriu[1,7,4], Katharina Heinzelmann[8,9], Juliane Merl-Pham[10], Martin Irmler[11], Hani N. Alsafadi[9,4,12], Eduard Trenkenschuh[13], Lenka Sarnova[14], Marketa Jirouskova[14], Wolfgang Frieß[13], Stefanie M. Hauck[10], Johannes Beckers[11,15,16], Nikolaus Kneidinger[4,17], Jürgen Behr[4,7,17], Anne Hilgendorff[1,4], Kamyar Hadian[2], Michael Lindner[4,7,18], Melanie Königshoff[8,9], Oliver Eickelberg[8], Martin Gregor[14], Oliver Plettenburg[5,6,19], Ali Önder Yildirim[1], Gerald Burgstaller[1,4]*

Fibrogenic processes instigate fatal chronic diseases leading to organ failure and death. Underlying biological processes involve induced massive deposition of extracellular matrix (ECM) by aberrant fibroblasts. We subjected diseased primary human lung fibroblasts to an advanced three-dimensional phenotypic high-content assay and screened a repurposing drug library of small molecules for inhibiting ECM deposition. Fibrotic Pattern Detection by Artificial Intelligence identified tranilast as an effective inhibitor. Structure-activity relationship studies confirmed *N*-(2-butoxyphenyl)-3-(phenyl)acrylamides (N23Ps) as a novel and highly potent compound class. N23Ps suppressed myofibroblast transdifferentiation, ECM deposition, cellular contractility, and altered cell shapes, thus advocating a unique mode of action. Mechanistically, transcriptomics identified SMURF2 as a potential therapeutic target network. Antifibrotic activity of N23Ps was verified by proteomics in a human ex vivo tissue fibrosis disease model, suppressing profibrotic markers SERPINE1 and CXCL8. Conclusively, N23Ps are a novel class of highly potent compounds inhibiting organ fibrosis in patients.

INTRODUCTION

Fibrotic diseases affect nearly every tissue in the body, account for more than 45% of all deaths in the industrialized world, and progressive forms of the disease rapidly lead to organ dysfunction, organ failure, and ultimately death (*1*–*3*). Because of its ubiquitous existence and high mortality, fibrosis, or "scarring," has become a high medical need for novel drug discovery strategies (*3*, *4*). However, effective antifibrotic therapeutics are missing from the clinic. The lack of antifibrotic therapies and its concomitant high medical need is best exemplified by idiopathic pulmonary fibrosis (IPF), which is a rapidly progressive and fatal fibrotic disorder. Patients with this common form of interstitial fibrotic lung disease face a median survival time of 3 to 5 years (*5*–*7*). Currently, only two approved antifibrotic drugs for IPF are on the market, pirfenidone and nintedanib; however, both substances partially slow down the rate in lung function decline but do not stop disease progression (*8*–*10*). Therefore, new therapeutic strategies and approaches are urgently required.

In fibrotic pathogenesis, repetitive and constant injury leads to a sustained and self-perpetuating activation of fibroblasts, leading to their transdifferentiation into synthetic and highly contractile α–smooth muscle actin (αSMA)–expressing myofibroblasts that massively deposit extracellular matrix (ECM), which stiffens the lung and destroys normal lung architecture (*3*, *6*, *11*, *12*). The matrisome of fibrotic ECM was shown to harbor a disease- and progression-specific signature of fibrillar collagens (types I, III, and V), proteoglycans, fibronectin, glycosaminoglycans, matrix-Gla protein, and microfibrillar-associated proteins (*11*, *13*–*16*).

Of all profibrotic signals reported, multifunctional transforming growth factor–β1 (TGFβ1) is the most intensively studied and a

[1]Institute of Lung Biology and Disease (ILBD) and Comprehensive Pneumology Center (CPC), Helmholtz Zentrum München, Member of the German Center for Lung Research (DZL), Munich, Germany. [2]Assay Development and Screening Platform, Institute of Molecular Toxicology and Pharmacology, Helmholtz Zentrum München-German Research Center for Environmental Health, Neuherberg, Germany. [3]Chemical and Process Engineering, Strathclyde University, Glasgow, Scotland, UK. [4]CPC-M bioArchive, Helmholtz Zentrum München, Comprehensive Pneumology Center Munich DZL/CPC-M, Munich, Germany. [5]Institute of Medicinal Chemistry, Helmholtz Zentrum München, German Research Center for Environmental Health (GmbH), Neuherberg, Germany. [6]Leibniz Universität Hannover, Institute of Organic Chemistry and Center for Biomolecular Drug Research (BMWZ), Hannover, Germany. [7]Asklepios Fachkliniken Munich-Gauting, Munich, Germany. [8]Division of Pulmonary, Allergy, and Critical Care Medicine, Department of Medicine, University of Pittsburgh, Pittsburgh, PA, USA. [9]Comprehensive Pneumology Center (CPC), Research Unit Lung Repair and Regeneration, Helmholtz Zentrum München, Member of the German Center for Lung Research (DZL), Munich, Germany. [10]Research Unit Protein Science, Helmholtz Zentrum München, German Research Center for Environmental Health GmbH, Neuherberg, Germany. [11]Institute of Experimental Genetics, Helmholtz Zentrum München, 85764 Neuherberg, Germany. [12]Wallenberg Center for Molecular Medicine (WCMM), Department of Experimental Medical Sciences, Lund University, Lund, Sweden. [13]Department of Pharmacy–Center for Drug Research, Pharmaceutical Technology and Biopharmaceutics, Ludwig-Maximillians University of Munich, Munich, Germany. [14]Laboratory of Integrative Biology, Institute of Molecular Genetics of the Czech Academy of Sciences, Prague, Czech Republic. [15]German Center for Diabetes Research (DZD e.V.), 85764 Neuherberg, Germany. [16]Chair of Experimental Genetics, Technische Universität München, 85354 Freising, Germany. [17]Department of Internal Medicine V, Ludwig-Maximillians University of Munich, Member of the German Center for Lung Research (DZL), Munich, Germany. [18]Paracelsus Medical Private University, Salzburg, Austria. [19]Institute for Lung Health (ILH), Justus Liebig University, Giessen, Germany.
*Corresponding author. Email: gerald.burgstaller@helmholtz-muenchen.de
†These authors contributed equally to this work.

central player in various fibrotic diseases capable of triggering transdifferentiation of fibroblasts into myofibroblasts (*17*–*21*). TGFβ1 binds to the TGFβ1 receptor, and downstream signaling occurs by posttranslational modifications of cytoplasmic members of the SMAD family, which act as transcription factors in the cell nucleus, regulating the expression of common profibrotic genes, including ECM proteins (*22*–*25*). Plasminogen activator inhibitor-1 (PAI-1) is an essential downstream target of the TGFβ1 pathway, suppressing the fibrinolytic system and is considered as a therapeutic target option for fibrosis (*26*). In addition, in IPF, profibrotic interleukin-8 (IL-8) was recently found to be secreted by a special fibrogenic mesenchymal progenitor cell population with autocrine effects on proliferation and motility and paracrine effects on macrophage recruitment (*27*).

However, target-based antifibrotic drug discovery faced failure for a substantial number of molecules to successfully translate into the clinics (*28*–*31*). In contrast, because of a much higher efficiency, pathway-unbiased phenotypic drug screening has gained momentum in first-in-class drug discovery (*32*), especially when patient-derived primary cells, physiological relevant stimuli, and a readout close to the clinical end point were applied (*33*).

In this study, we deployed an unprecedented phenotypic high-throughput/-content drug discovery approach using human patient material, and identified *N*-(2-butoxyphenyl)-3-(phenyl)acrylamides (N23Ps) as novel and highly potent [median inhibitory concentration (IC_{50}) < 1 μM] antifibrotic compounds mechanistically acting in the TGFβ1-based network of SMURF2/SMAD (de)ubiquitination. These data support further development of N23Ps as a therapeutic option for fibrosis.

RESULTS

ECM deposition in activated IPF patient–derived human fibroblasts

The aberrant and excessive deposition of ECM by pathologically activated fibroblasts and the concomitant increase in tissue stiffening are hallmarks of fibrogenesis, an irreversible and fatal process that progresses to organ dysfunction (*3*, *11*). The abnormal activation of fibroblasts and their transformation into highly synthetic and contractile myofibroblasts occurs by processes of repeated tissue (micro)injuries, subsequent inflammation, and reactivation of developmental signaling pathways such as TGFβ1, Wnt, Shh, Bmp, and, recently, IL-11 (*6*, *34*). However, during disease progression, TGFβ1 acts as a profibrotic master regulator inducing an activated physiological state in fibroblasts transforming them into highly contractile myofibroblasts that aberrantly deposit ECM-forming scar tissue (*6*, *11*, *21*).

Here, we developed and established a disease-relevant and innovative high-throughput phenotypic screening assay for drug discovery purposes that uses patient-derived primary human lung fibroblasts (phLFs). The amount of ECM deposited by these myofibroblasts was used as assay readout, which closely mimics the clinical end point of fibrotic diseases. phLFs were isolated by outgrowth from fibrotic lung tissue obtained from explanted lungs from five different patients with IPF (Fig. 1, A and B), expanded to 5×10^8 cells per patient until passage 5, and lastly frozen for subsequent screening purposes. Figure 1C lays out the workflow of the entire assay. However, immunolabeling of ECM proteins led to a substantial staining of intracellular nonfibrillar components, even without detergent permeabilization of phLFs, which led to false ECM quantification (fig. S1A). This was also exemplified by treatment with brefeldin A, an inhibitor of protein secretion (fig. S1, B and C). Conclusively, only immunolabeling of living cells before fixation led to exclusive staining of ECM components.

Next, automated high-resolution three-dimensional (3D) image acquisition was accomplished by confocal laser scanning microscopy and subsequent surface rendering allowed quantification of 3D ECM depositions. Automated cell count of cell nuclei served as normalization (Fig. 1D).

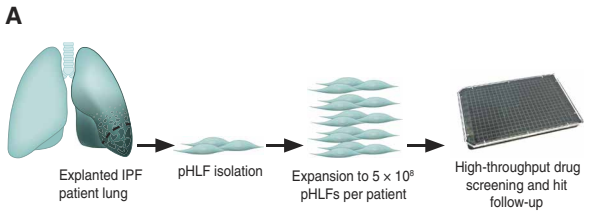
High-resolution 3D confocal imaging of immunolabeled phLFs revealed single ECM fibers appearing outside and on the surface of cells (Fig. 1E). Similarly, coimmunostaining of collagen I and collagen V disclosed an intricate extracellular fiber network, with partial colocalization of both collagen I and V subtypes, and areas of single-stained ECM fibers (Fig. 1F). In addition, 16-hour live 4D imaging of phLFs notably revealed active deposition and assembly of ECM fibers (Fig. 1G and movie S1). Together, by establishing a unique patient-derived human fibrosis model, we demonstrated the exclusive 3D quantification of ECM deposition in high spatial resolution (Fig. 1H).

A patient-derived human fibrosis model

Transdifferentiation to myofibroblasts is induced by treatment with profibrotic TGFβ1 mimicking the fundamental biological process leading to fibrotic diseases. Here, we applied active TGFβ1 (1 ng/ml) on patient-derived phLFs ($phLFs^{+TGFβ1}$) and subsequently quantified a significant increase in cytoskeletal αSMA expression by 2.3-fold compared to untreated phLFs hallmarking myofibroblasts (Fig. 2A).

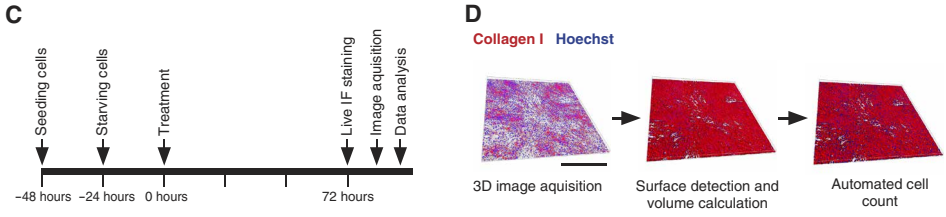
Next, for the selection of fibrosis-relevant ECM biomarkers for the quantitative assay readout, we bioinformatically compared a recently published human myofibroblast surface proteome containing 1126 proteins with 284 "core matrisome" proteins (Fig. 2B) (*35*, *36*). We obtained a pool of 17 deregulated ECM protein candidates (Fig. 2B) and selected collagen I and fibulin 1, two different ECM proteins marking collagen and elastic fibers, respectively (Fig. 2C). Furthermore, collagen I and collagen V were among the highest abundant transcripts especially expressed in myofibroblasts in the IPF cell atlas (*37*). Therefore, we included collagen V as a supplementary ECM deposition marker. The analysis of all three ECM proteins collagen I, fibulin 1, and collagen V after TGFβ1 treatment displayed a significant increase in deposition (Fig. 2D) but did not affect cell numbers of human fibroblasts (Fig. 2E). Similarly, protein analysis exhibited a significant increase in intracellular (3.4-fold) and secreted collagen I (3.2-fold) and collagen V (3.9- and 10.7-fold, respectively) after TGFβ1 stimulation (fig. S2, A and B). As proof of concept for ECM deposition inhibition in $phLFs^{+TGFβ1}$, we used ethyl-3,4-dihydroxy-benzoate (EDHB), which reportedly inhibits prolyl-4-hydroxylases and collagen synthesis (*38*, *39*). EDHB (50 μM) successfully inhibited ECM deposition in $phLFs^{+TGFβ1}$ (Fig. 2, F and G). Furthermore, intracellular protein expression of collagen I and collagen V was significantly inhibited in a dose-dependent manner, with mostly similar effects on fibulin 1 expression (fig. S2C). In line, secretion of collagen I was entirely abolished by EDHB treatment (fig. S2D).

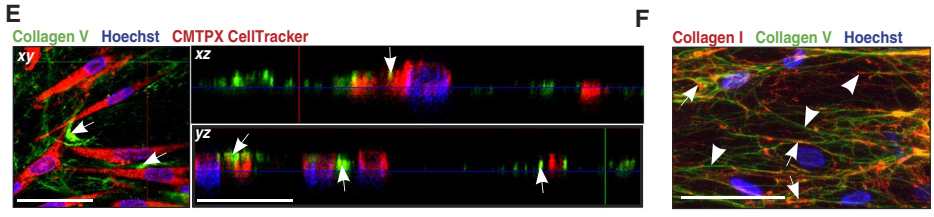
Notably, specific blocking of the TGFβ1 receptor using the inhibitor SB-431542 resulted in an inhibition of ECM deposition in phLFs (fig. S3, A to D). Recently, a pro–fibrotic cocktail (FC) containing TGFβ1, platelet-derived growth factor AB (PDGF-AB),

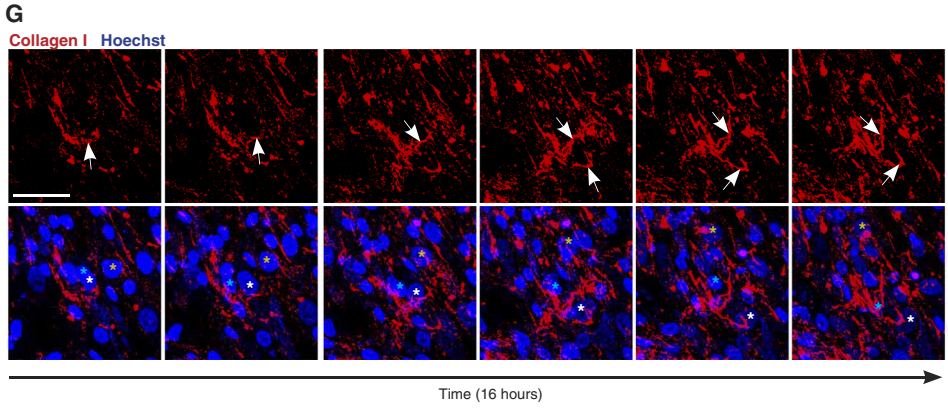


B

Pseudonym	Age*	Sex	Diagnosis
IPF#1	46 years	Male	IPF
IPF#2	63 years	Male	IPF
IPF#3	63 years	Female	IPF
IPF#4	55 years	Male	IPF
IPF#5	48 years	Female	IPF







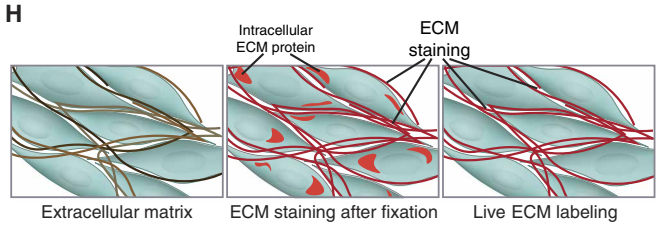


Fig. 1. 3D assessment of ECM deposition using IPF patient–derived phLFs. (**A**) phLFs are derived from explanted IPF lungs, expanded in cell culture, and used for high-throughput drug screening and hit validation. (**B**) Clinical data of patients from which the phLFs were derived. (**C**) Graphical representation of the actual workflow used in the ECM deposition assay. (**D**) Software-based volume rendering of confocal z-stacks of immunostained ECM (collagen I in red) is used for the quantification of ECM volume and automated cell count (Hoechst-stained cell nuclei in blue). Scale bar, 500 μm. (**E**) Orthoview of a confocal z-stack of phLFs (red) depositing collagen V (green) exclusively outside (indicated by white arrows) and on the surface of the cells. Cell nuclei are stained by Hoechst (blue). Scale bars, 50 and 25 μm. (**F**) Intricate 3D ECM network of collagen I (red)– and collagen V (green)–stained fibers, demonstrating areas of colocalizing fibers (white arrows) and single fibers (white arrowheads). Cell nuclei are stained by Hoechst (blue). The confocal z-stack is shown as a maximum intensity projection. Scale bar, 50 μm. (**G**) 4D confocal time-lapse imaging of phLFs for 16 hours showing various dynamic processes occurring during the assembly of single collagen I ECM fibers (red). White arrows indicate single ECM fibers. Colored asterisks indicate single cell-nuclei (Hoechst in blue) of distinct cells. Scale bar, 50 μm. (**H**) Illustration of ECM fibers surrounding fibroblasts (left), intra- and extracellular staining of matrisome proteins with conventional staining protocol after fixation (middle), and exclusive ECM labeling using live immunofluorescence (IF) staining before cell fixation (right).

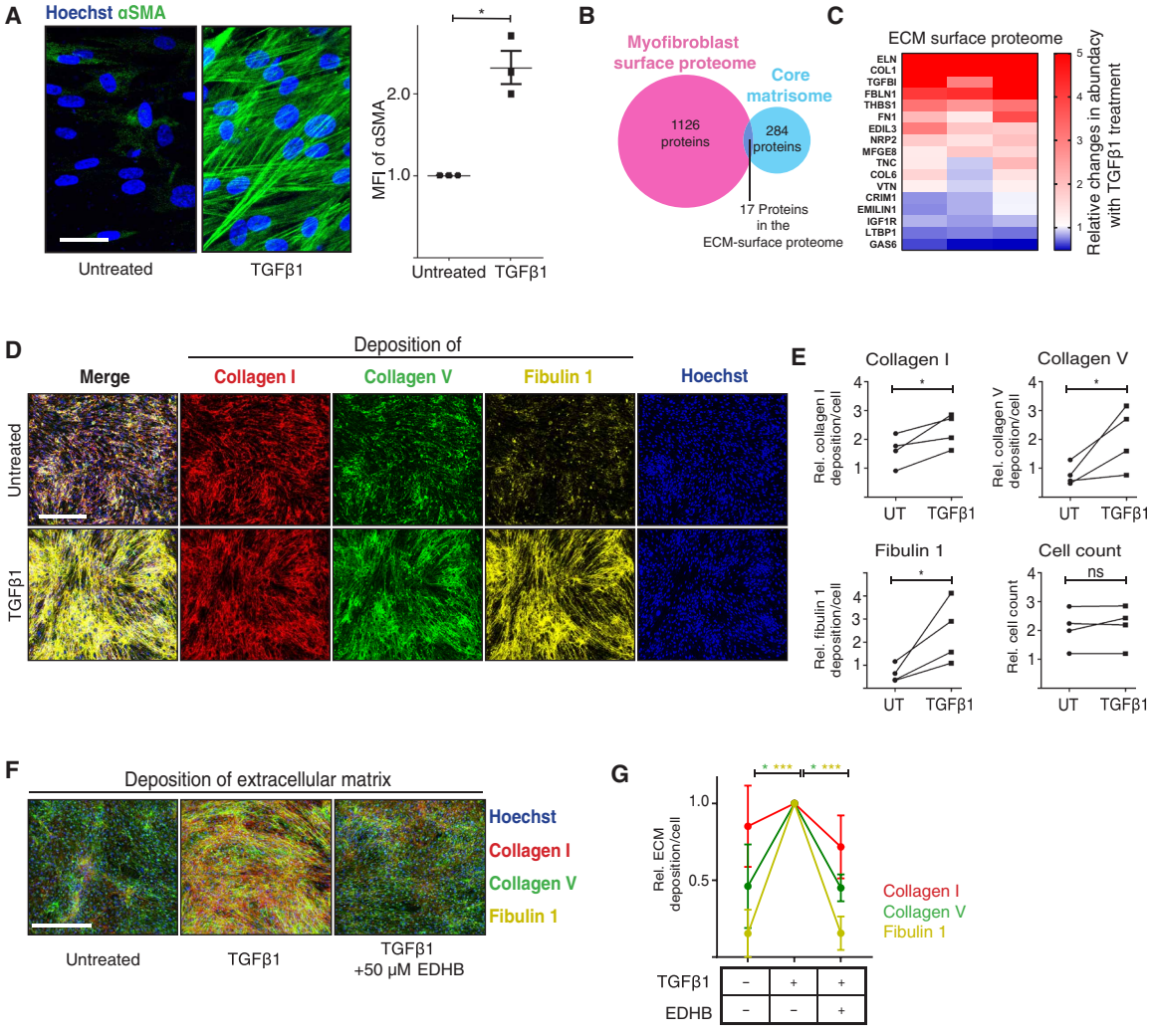


Fig. 2. 3D fibrosis disease model using IPF patient–derived phLFs. (**A**) phLFs treated with TGFβ1 (1 ng/ml) transdifferentiated to myofibroblasts that incorporated αSMA (green) into actin stress fibers. Cell nuclei are stained by Hoechst (blue). Scale bar, 200 μm. Quantification of the mean fluorescence intensity (MFI) of αSMA expression of three different patient phLFs after TGFβ1 treatment ($n = 3$). (**B**) Venn diagram showing an overlap of 17 ECM proteins between the myofibroblast surface proteome (pink) and a published core matrisome (blue). (**C**) Heatmap of protein expression levels of ECM proteins on the surface of myofibroblasts (phLFs$^{+TGFβ1}$), identifying collagen I and fibulin 1 among the highest up-regulated ECM proteins. Red and blue indicate high and low protein expression levels, respectively. (**D**) 3D confocal immunofluorescence microscopy of phLFs and phLFs$^{+TGFβ1}$. phLFs$^{+TGFβ1}$ showed increased ECM deposition of collagen I (red), collagen V (green), and fibulin 1 (yellow). Cell nuclei were stained by Hoechst (blue). The confocal z-stack is shown as a maximum intensity projection. Scale bar, 500 μm. (**E**) Software-based quantification of the deposited ECM volume displays a significant increase in the amount of deposited ECM in phLFs$^{+TGFβ1}$, whereas the amount of cells remains unchanged. $n = 4$ (four different patient phLFs). Statistics: One-way analysis of variance (ANOVA) with Bonferroni-correction. UT, untreated; ns, not significant. (**F**) 3D confocal immunofluorescence microscopy assessing the ECM deposition of phLFs, phLFs$^{+TGFβ1}$, and phLFs$^{+TGFβ1+EDHB}$. EDHB treatment inhibits the ECM deposition of collagen I (red), collagen V (green), and fibulin 1 (yellow). Cell nuclei were stained by Hoechst (blue). The confocal z-stack is shown as a maximum intensity projection. Scale bar, 500 μm. (**G**) Software-based quantification of the deposited ECM volume of data shown in (F). Data are presented as means ± SEM. Differences between groups were evaluated with paired *t* tests. *$P < 0.05$.

tumor necrosis factor–α (TNFα), and lysophosphatidic acid (LPA) was used to trigger a profibrotic response in cultured lung tissue slices (*40*). Now, by applying exactly this, FC to phLFs in our ECM deposition assay while simultaneously blocking the TGFβ1 receptor using the inhibitor SB-431542, no enhanced ECM deposition was observed (fig. S3, E to H). Therefore, we concluded that in our assay, only TGFβ1 was effective in enhancing ECM deposition. In summary, we developed an effective 3D fibrosis disease model based on IPF patient–derived human (myo)fibroblasts that is successfully interfered by pharmacological small molecules acting as inhibitors of ECM deposition.

Deep learning identifies antifibrotic compounds in a high-content screening of a U.S. Food and Drug Administration–approved drug library

To identify novel inhibitors of ECM deposition in our human fibrosis model, we used automatized high-content screening in combination with a novel deep learning hit-detection algorithm. Thus, we screened a U.S. Food and Drug Administration (FDA)–approved drug library of a unique collection of 1509 marketed drugs at a concentration of 100 μM. To efficiently interpret the huge amount of imaging data obtained, we developed a novel Fibrotic Pattern Detection by Artificial Intelligence (FANTAIL) algorithm. FANTAIL

deployed a trained deep convolutional neuronal network (CNN) (Fig. 3A) to discriminate fibrotic from nonfibrotic ECM patterns in 3D confocal microscopy images (Fig. 3B) and hyperparameter optimization (n = 3 and np = 128) led to a 98.83% sensitivity and 100% specificity (fig. S4, A to E). In addition, we trained the CNN to classify toxic compounds by applying training sets of phLFs$^{+TGF\beta1}$ plus 5% ethanol. False-positive hits due to cytotoxicity were clearly distinguished from nontoxic hits, and thus FANTAIL produced a hit rate of 2% (Fig. 3B). Subsequently, images of positive hits were clustered by a uniform manifold approximation and projection (UMAP), which removed immunofluorescence artifacts (Fig. 3C and fig. S4F). This resulted in 31 of 1509 FDA-approved drugs actively inhibiting ECM deposition in phLFs$^{+TGF\beta1}$, which were classified according to the manufacturer's molecular target class annotations (Fig. 3D). Next, to further emphasize the relevance of IPF-derived patient cells for screening outcomes, we rescreened the same FDA-approved drug library as before but now using the human fetal lung fibroblast cell line HFL1. Other cell lines such as CCL-151 lung fibroblasts were tested but were not functional in our assay (fig. S5, A to C). In addition, we detected substantial transcriptional differences between HFL1, CCL-151, and IPF-phLFs, especially after TGFβ1 stimulation (fig. S5, D to H). Now, by performing a rescreening of the FDA-approved drug library with HFL1 fibroblasts, FANTAIL produced 53 hits (3.4%) but altogether generated only two overlapping hit compounds between HFL1 and IPF-phLFs (fig. S5, I and J). Last, statistical testing by Monte Carlo simulation confirmed that the experimental outcome is based on true biological variations between the cells used, which altogether strongly advocates the usage of IPF patient–derived phLFs for antifibrotic drug development campaigns (fig. S5K). Conclusively, we provide here experimental evidence that the usage of a human fetal fibroblast cell line (HFL1) generates a hit profile substantially diverse from IPF-phLFs, which might be directly caused by biological dissimilarities between both cell types, particularly after TGFβ1 stimulation.

It was highly encouraging that in our list of 31 active hits derived from the screening campaign using IPF-phLFs, we identified several compounds that have been evaluated for treatment of IPF in preclinical (vorinostat) or even clinical trials (sildenafil and bosentan). The resulting 31 hits were clustered, and chemotherapeutics and anti-infectives, as well as compounds with assumed nonspecificity at the tested concentration, or those containing toxophoric groups were deprioritized. From the 16 remaining compounds, we selected *N*-(3′,4′-dimethoxycinnamoyl)-anthranilic acid (tranilast) for further profiling. Tranilast was chosen for further examination because of its good efficacy in inhibiting matrix deposition in our screening assay, along with its well-documented in vivo bioavailability, safety, and tolerance. In addition, tranilast was of interest because of its utilization as an antiallergic drug for treatment of bronchial asthma and its use in keloid and hypertrophic scar treatment. However, the exact mode of action (MoA) of this compound is still unknown (*41*, *42*). Next, we performed dose-response experiments to validate tranilast's inhibitory effect on ECM deposition. Expression of ECM markers (collagen I, collagen V, and fibulin 1) was significantly inhibited by tranilast treatment at concentrations of >150 μM (Fig. 3, E and F) that was not due to cytotoxic effects (fig. S6A). Concomitantly, we observed a significant dose-dependent reduction of cytosolic protein expression of αSMA, fibulin 1, collagen I, and collagen V after tranilast treatment (fig. S6, B to D). Furthermore, mRNA transcripts of αSMA/ACTA2 and collagen I, but not of collagen V and fibulin 1, were decreased by tranilast (fig. S6E), indicating a decoupled regulation of gene and protein expression. Thus, by integrating FANTAIL and high-content screening into our human fibrotic disease model, we successfully identified a panel of small molecules with antifibrotic activity displaying efficient inhibition of ECM deposition.

Structure-activity relationship studies find a novel series of highly potent antifibrotics

In our human fibroblast–based assay the inhibition by tranilast occurred only at very high concentrations (>150 μM). Therefore, we performed structure-activity relationship (SAR) studies of commercially available small-molecule tranilast derivatives along with newly designed compounds produced by medicinal chemistry efforts (fig. S7, A to C). By applying our ECM deposition assay using collagen V and fibulin 1 as readouts in dose-response relationship studies, we found a novel series of 2-butoxy–substituted derivatives (N23Ps), which were >100-fold more potent compared to tranilast in inhibiting ECM deposition (Fig. 4, A and B, and fig. S7, D and E). In particular, cmp**4**, cmp**12**, cmp**17**, and cmp**18**, all of which were characterized by a 2-butoxy substitution at R1, as well as distinct substitution patterns of methoxy, ethoxy, and butoxy groups at R6 and R7, were found to be active at 10 μM (fig. S7, A and B). Further dose response studies of cmp**4**, cmp**12**, cmp**17**, and cmp**18** demonstrated that specifically these modifications resulted in highly active compounds with substantial potency increases: >53-fold (collagen V) and >45-fold (fibulin 1) for cmp**4**, >10-fold (collagen V) and >10-fold (fibulin 1) for cmp**12**, >10-fold (collagen V) and >10-fold (fibulin 1) for cmp**17**, and >1000-fold (collagen V) and >500-fold (fibulin 1) for cmp**18** (Fig. 4B and fig. S7D).

Furthermore, testing of cmp**24** to cmp**35** (at 10 μM), which were all newly synthesized by medicinal chemistry efforts, produced two compounds with increased potency: first, a 2-*o*-benzyl substitution in cmp**35** at R1 and, second, a benzodioxol in cmp**31** (fig. S7C). In subsequent dose-response experiments, these substitutions resulted in a >10-fold (fibulin 1 and collagen V), as well as >10-fold (fibulin 1 and collagen V) potency increase for cmp**31** and cmp**35**, respectively (Fig. 4B and fig. S7E). Reduction of ECM deposition by cmp**4**, cmp**12**, cmp**17**, cmp**18**, cmp**31**, and cmp**35** was not only specific for lung fibroblasts but was also highly effective (inhibition at 50 nM) in TGFβ1-treated human dermal fibroblasts (fig. S7, F and G), which possibly advocates for a core fibrogenic target pathway.

As cmp**4** displayed an identical potency increase (98-fold) for collagen V and fibulin 1, we used it as an N23Ps prototype in all subsequent experiments. Next, to investigate cytotoxic activity and real-time inhibition of N23Ps, we applied confocal 4D live-cell imaging of phLFs$^{+TGF\beta1}$ together with cmp**4** (w_c = 10 μM) for a period of 48 hours (Fig. 4C and movie S2). Subsequent quantification unequivocally displayed a dynamic inhibition of ECM deposition by cmp**4** without showing cell death (Fig. 4D). Additional live-dead immunofluorescence staining of phLFs$^{+TGF\beta1+cmp4}$ confirmed fibroblast viability (Fig. 4E). We did, however, observe an outstanding morphology switch of phLFs$^{+TGF\beta1}$ compared to phLFs$^{+TGF\beta1+cmp4}$ (Fig. 4E). We analyzed this finding further in subconfluently cultured phLFs$^{+TGF\beta1}$ and phLFs$^{+TGF\beta1+cmp4}$ by assessing cell morphology parameters, demonstrating a significant switch from elongated to round cells, including an extensive rearrangement of the actin cytoskeleton (Fig. 4F). These substantial morphology changes, observed

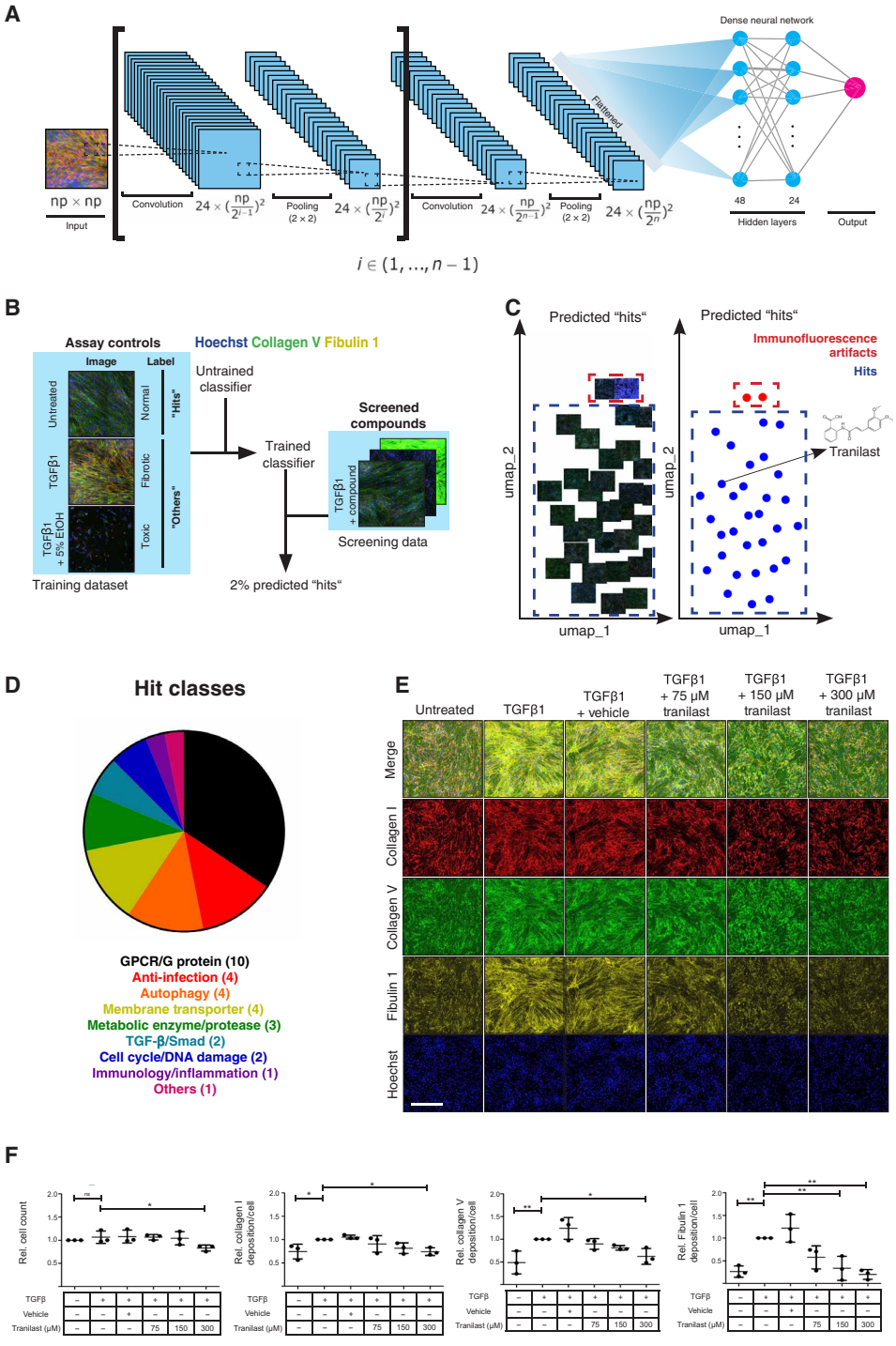
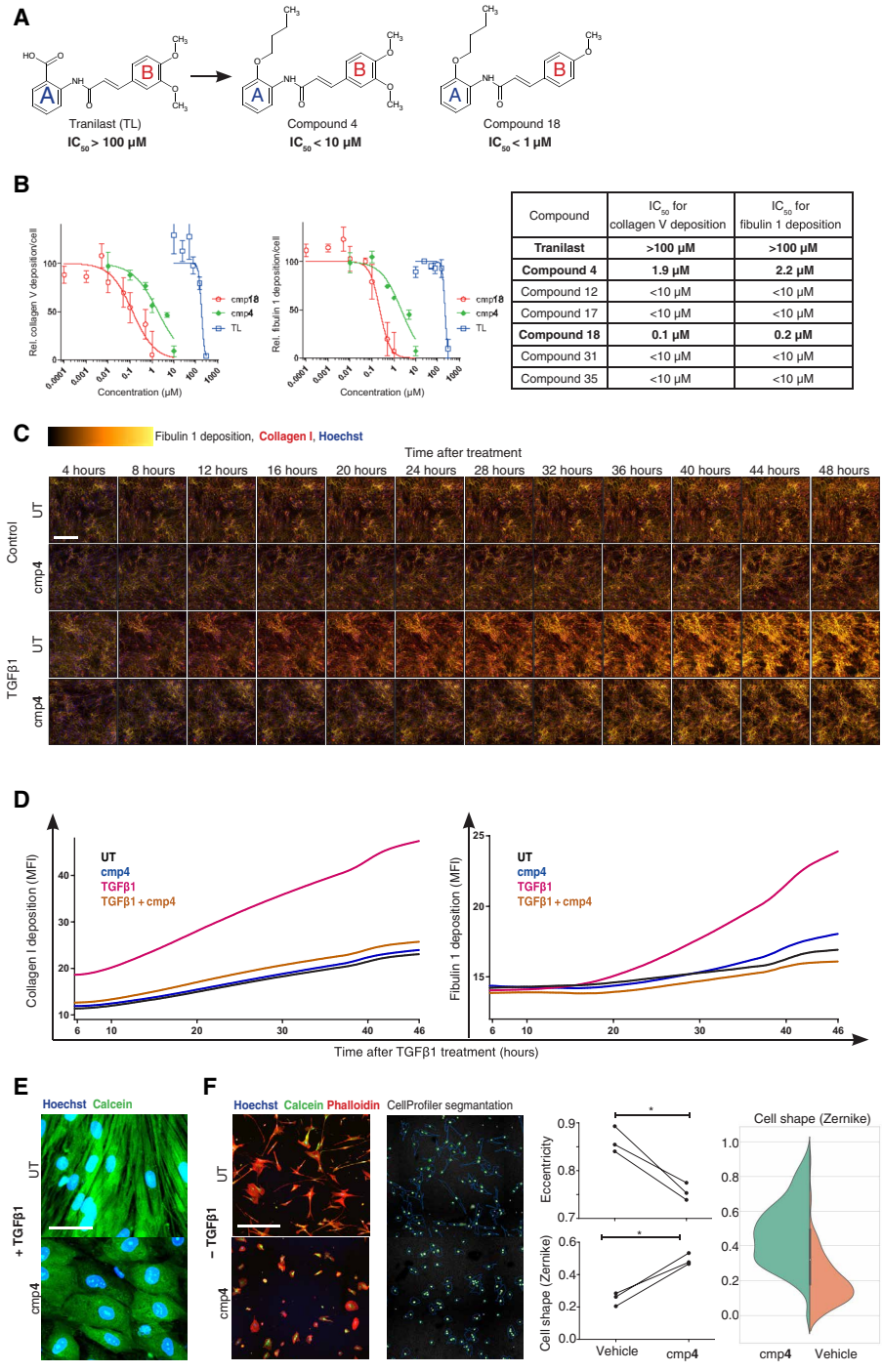


Fig. 3. FANTAIL using a CNN for hit identification within ECM deposition screening data of 1509 FDA-approved compounds. (**A**) Outline of the supervised multilayered deep CNN developed for detecting fibrotic and nonfibrotic patterns in images containing deposited ECM derived from 3D confocal microscopy of immunofluorescently labeled collagen I, collagen V, and fibulin 1. (**B**) The training dataset consisted of assay controls and additional samples treated with phLFs$^{+TGF\beta1}$ plus 5% ethanol (EtOH). The CNN network was exclusively trained to detect inhibitors of ECM deposition and false-positive hits due to cytotoxic effects. (**C**) UMAP clustering of predicted hits sorted out false-positive hits due to immunofluorescence artifacts. After this final filtering, *N*-(3′,4′-dimethoxycinnamoyl)-anthranilic acid (tranilast) was determined a promising candidate for repurposing in IPF. (**D**) Pie chart demonstrating the classification of detected hits into groups of similar molecular functions and biological processes. Tranilast is classified within the G protein–coupled receptor (GPCR) targeting molecules. (**E**) 3D confocal microscopy of immunofluorescently labeled collagen I (red), collagen V (green), and fibulin 1 (yellow) to validate the inhibitory effect of tranilast-treated phLFs$^{+TGF\beta1}$ on the deposition of ECM in a dose-dependent manner (untreated, vehicle, 75, 150, and 300 μM). Cell nuclei were stained by Hoechst (blue). The confocal z-stack is shown as a maximum intensity projection. Scale bar, 500 μm. (**F**) Quantification and statistical analysis of the inhibitory effects of tranilast on phLFs$^{+TGF\beta1}$ on ECM deposition of collagen I, collagen V, and fibulin 1 normalized to the number of cells. **P* < 0.05 and ***P* < 0.01. Statistics: One-way ANOVA with Bonferroni correction. All quantitative data represent means ± SEM. *n* = 3 (three different patient phLFs).



Compound	IC_{50} for collagen V deposition	IC_{50} for fibulin 1 deposition
Tranilast	**>100 µM**	**>100 µM**
Compound 4	**1.9 µM**	**2.2 µM**
Compound 12	<10 µM	<10 µM
Compound 17	<10 µM	<10 µM
Compound 18	**0.1 µM**	**0.2 µM**
Compound 31	<10 µM	<10 µM
Compound 35	<10 µM	<10 µM

Fig. 4. Dose-response relationships of N23Ps, live imaging of ECM deposition, and cell morphology switch. (**A**) Structural formula for tranilast and N23Ps (cmp**4** and cmp**18**) indicating 2-butoxy substitution at R1 (see fig. S7A) and IC_{50} values. (**B**) Dose-response curves of collagen V and fibulin 1 ECM deposition for N23Ps (cmp**4** and cmp**18**) compared to tranilast for determining IC_{50} values. $n = 3$ (three different patient phLFs). (**C**) Tables of IC_{50}s for active N23Ps compared to tranilast. (C) 3D confocal images displayed as maximum intensity projection of deposited ECM taken as frames from a live-cell experiment demonstrating dynamic ECM deposition in phLFs, $phLFs^{+TGF\beta1}$, and $phLFs^{+TGF\beta1+cmp4}$ during 48 hours. Immunostained collagen I and fibulin 1 depicted as one common signal (= ECM deposition) as a gold look-up table. Cell nuclei in blue (Hoechst). Scale bar, 500 µm. (**D**) Quantification by MFI of dynamic deposition of collagen I and fibulin 1 in phLFs, $phLFs^{+TGF\beta1}$, and $phLFs^{+TGF\beta1+cmp4}$ notably demonstrating inhibition of ECM deposition by cmp**4** over time. (**E**) Confocal images of calcein-stained confluent $phLFs^{+TGF\beta1}$ proving viability in untreated and cmp**4**-treated cells and identifying a morphology switch from elongated to round cells. Cell nuclei were stained by Hoechst (blue). Scale bar, 50 µm. (**F**) Confocal images of subconfluent TGFβ1-untreated phLFs stained for Hoechst (blue), calcein (green), and phalloidin (red) showing round cell morphologies after treatment with cmp**4** and extensive actin cytoskeletal rearrangements (red). Software-based segmentation of the cells in the images by CellProfiler allowed for the statistical analysis that exhibited significant changes in cell shape and eccentricity toward round cells after cmp**4** treatment. Scale bar, 500 µm. Data are presented as means ± SEM. Differences between groups were evaluated with paired *t* tests. $^*P < 0.05$. $n = 3$ (three different patient phLFs).

with N23Ps but not with tranilast, might strongly indicate an exclusive MoA for N23Ps. Together, in our SAR studies, replacement of 2-carboxy-phenyl at R1 with a 2-butoxy group in tranilast strongly increased its antifibrotic activity with a concomitant morphological phenotype switch, which strongly advocates for the discovery of a new therapeutic class of highly potent antifibrotics with a unique MoA.

N23Ps activate a unique antifibrotic regulatory network

To further interrogate a molecular mechanism for N23Ps, we applied microarray-based transcriptomics and in-depth software-based network analysis comparing bulk transcriptomes of transdifferentiated myofibroblasts (phLFs$^{+TGF\beta1}$) with those inhibited by cmp**4** (Fig. 5A). The entire microarray data of phLFs from three different patient samples were submitted to Gene Expression Omnibus (GSE141905). We found 2076 genes significantly deregulated in phLFs$^{+TGF\beta1}$ versus phLFs^{control}, 661 genes were significantly deregulated in cmp**4**-inhibited phLFs$^{+TGF\beta1\ +\ cmp4}$ versus phLFs$^{+TGF\beta1}$, and 1102 genes significantly deregulated in phLFs^{+cmp4} versus phLFs^{control} [all >2-fold, false discovery rate (FDR) < 10%] (Fig. 5B). The 20 highest and lowest deregulated genes from the analyses phLFs$^{+TGF\beta1+cmp4}$ versus phLFs$^{+TGF\beta1}$, phLFs$^{+TGF\beta1}$ versus phLFs^{control}, and phLFs^{+cmp4} versus phLFs^{control} are displayed as heatmaps (fig. S8, A to C). Among the differentially regulated genes (fold change > 2, FDR < 10%) of phLFs$^{+TGF\beta1}$ and phLFs$^{+TGF\beta1+cmp4}$, we found 362 overlapping genes (Fig. 5C). A majority (279) of these overlapping genes were counter-regulated by cmp**4** in phLFs$^{+TGF\beta1\ +\ cmp4}$, whereas the remaining 49 and 34 transcripts were either only up- or down-regulated in phLFs$^{+TGF\beta1}$ and phLFs$^{+TGF\beta1\ +\ cmp4}$, respectively (fig. S8D). A detailed gene set enrichment analysis (GSEA) of phLFs$^{+TGF\beta1+cmp4}$ revealed a negative enrichment for profibrotic functional gene signatures such as collagen formation, ECM organization, and smooth muscle contraction after cmp**4** treatment (Fig. 5D).

Usually, in omics analyses, hierarchical clustering dendograms and heatmaps of expression data are the gold standard to find coregulated and functionally related genes. However, in multidimensional datasets, the complexity is often too intricate to unravel interesting interactions. Therefore, we developed UMAP–regulation pattern clustering (UMAP-RPC) assay, a novel bioinformatic tool for facilitating pattern analysis in omics data (fig. S8, E to G). Here, clustering of genes depends on similarities in the expression patterns of each condition. To exemplify this, fig. S8F displays two patterns of transcript regulation across different conditions, one for matrix metalloproteinase-1 (*MMP1*), which is up-regulated in phLFs^{+cmp4} and phLFs$^{+TGF\beta1+cmp4}$, and a second condition for myosin heavy chain 1 (*MYH11*), which is up-regulated in phLFs$^{+TGF\beta1}$ only. After applying UMAP-RPC, we color-coded fold changes of transcript regulation for each gene (Fig. 5E). With that, we identified gene clusters that were clearly distinct in their expression patterns between the various conditions tested. Most genes in cluster A (red) were found up-regulated in phLFs$^{+TGF\beta1}$ only, whereas genes in cluster B (green) were up-regulated in both, phLFs$^{+TGF\beta1+cmp4}$ and phLFs^{+cmp4}, but not in phLFs$^{+TGF\beta1}$ (Fig. 5E). Next, both gene clusters were analyzed in the STRING database for functional interactions (*43*). Cluster A (red) included subnetworks involved in ECM organization (HSA-1474244) and the actin cytoskeleton (GO:0015629) (Fig. 5F), which mostly are myofibroblast differentiation markers (fig. S9A). Cluster B (green) included functional subnetworks of deubiquitination (HSA-5688426), laminin interactions (HSA-3000157), Rho guanosine triphosphatase (GTPase) effectors (HSA-195258), and ECM receptor interactions (hsa04512) (Fig. 5G and fig. S9B). Within cluster B, we identified two N23P-mediated up-regulated TGFβ1 pathway inhibitors *SMURF2* (3.19-fold, FDR < 10%) and *CDCP1* (17.41-fold, FDR < 10%) (*44*), which both act as potential druggable antifibrotic molecular targets (Fig. 5F and fig. S9B).

N23Ps block myofibroblast transdifferentiation and cellular contractility in a SMURF2-dependent manner

Next, in relationship to the functional interactions found in gene-cluster A of Fig. 5F, we explored N23Ps inhibitory effect on fibrosis-related biological processes such as fibroblast-to-myofibroblast transdifferentiation and cellular contractility. Myofibroblast transdifferentiation was significantly inhibited in N23P-treated phLFs$^{+TGF\beta1}$ (Fig. 6, A and B). Furthermore, N23Ps blocked the TGFβ1-induced contraction of phLFs in a 3D collagen contraction assay (Fig. 6, C and D). In addition, TGFβ1 prestimulation of phLFs 24 and 48 hours before N23P treatment prevented the inhibition of ECM deposition (fig. S10, A to D). In addition, N23Ps blocked ECM deposition of collagen V but not of fibulin 1 only in concurrent treatment with TGFβ1 (fig. S10, E to H). Still, the decline in fibulin 1 deposition without TGFβ1 treatment might be due to the phLFs' basal intrinsic TGFβ1 secretion and its autocrine effect. In summary, this implies a suppression in TGFβ1 induced transdifferentiation of phLFs into myofibroblasts by N23Ps, indicated by an inhibition of ECM deposition.

SMURF2, which we identified in gene cluster B of Fig. 5G, is a ubiquitin E3 ligase promoting the degradation of SMADs, and SMURFs are crucial intracellular signaling effectors in the TGFβ1 signaling pathway (*45*). Here, to gain mechanistic insight of N23P-related regulation in TGFβ1-dependent myofibroblast transdifferentiation, we performed a loss-of-function study using SMURF2 small interfering RNA (siRNA) knockdown. For that, we successfully depleted SMURF2 expression in IPF-phLFs by 70% (Fig. 6E). As shown previously, cmp**4** treatment significantly inhibited αSMA protein expression in TGFβ1-treated IPF-phLFs; however, this inhibition was prevented in SMURF2-depleted phLFs$^{+TGF\beta1+cmp4}$ (Fig. 6F), as well as phenotypically in myofibroblasts (Fig. 6G). In conclusion, N23Ps act via SMURF2-inhibited TGFβ1 signaling and thus block fibroblast-myofibroblast transdifferentiation with consequences for cell contractility and ECM deposition (Fig. 6H).

N23Ps inhibit a fibrotic signature in a human ex vivo fibrosis model

Last, we examined antifibrotic activity of N23Ps in a highly complex ex vivo human 3D tissue culture model (3D-LTC). Previously, we have demonstrated that living human precision cut lung slices (hPCLSs) treated with a pro-FC (TGFβ1, TNFα, PDGF-AB, and LPA) revealed a specific fibrotic signature (*40*, *46*). This human ex vivo model for IPF can successfully mimic injury and early fibrosis-like responses in human lung tissue. We treated the hPCLS from three different patients (n = 3) with FC, control cocktail, FC plus cmp**4**, or FC plus tranilast (Fig. 7A). Subsequent comprehensive proteome profiling of these samples by high-performance liquid chromatography–tandem mass spectrometry (LC-MS/MS) resulted in 135 significantly (P < 0.05) deregulated proteins in hPCLS^{+FC} of an overall 3362 detected proteins. The 20 significantly highest and lowest abundant proteins in hPCLS$^{+TGF\beta1}$ are depicted as a heatmap in fig. S11A.

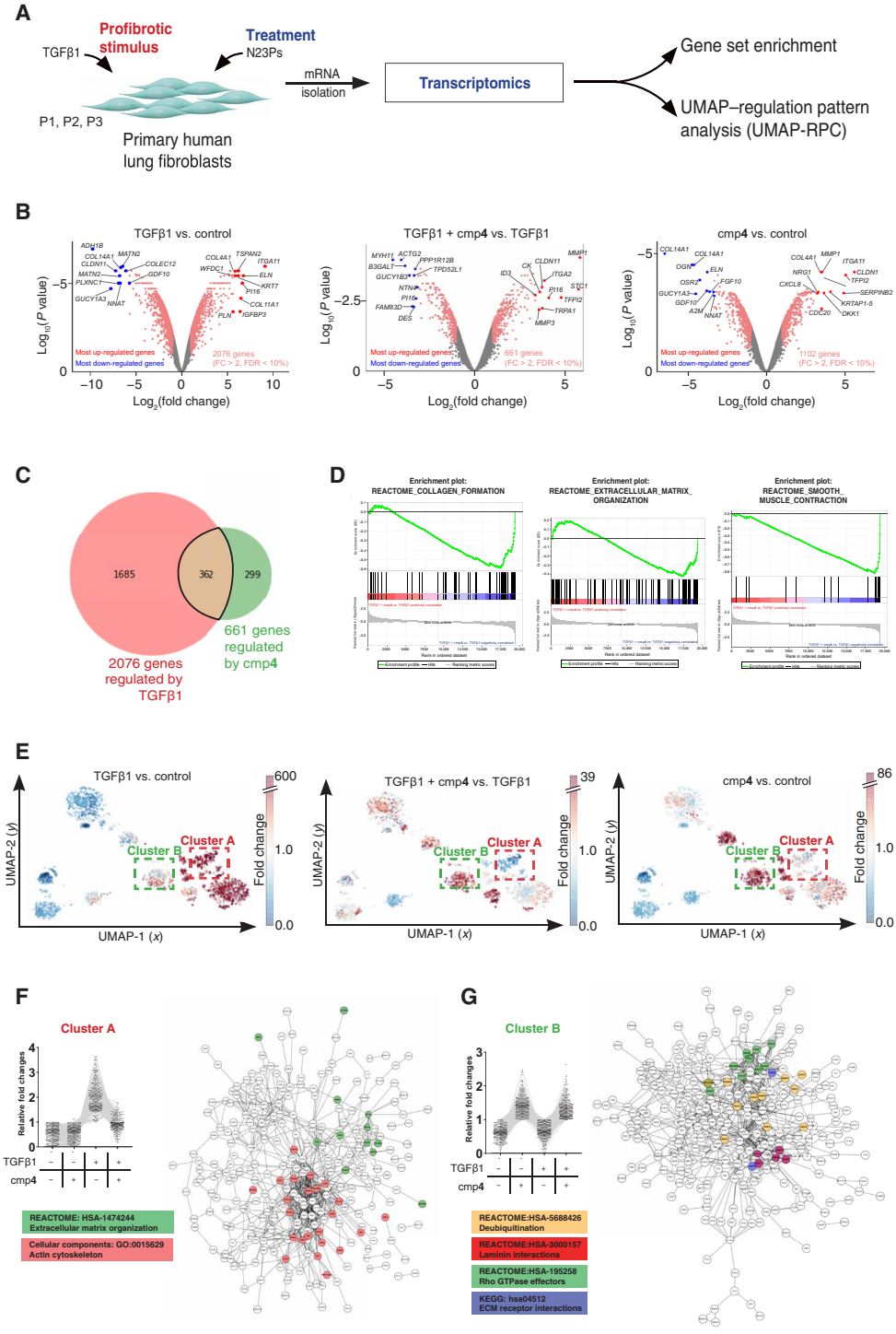


Fig. 5. Genome-wide transcriptomic analysis of N23Ps identifying a unique antifibrotic target network. (**A**) Experimental outline of transcriptional analysis of the human fibrosis model and treatment with ECM deposition inhibiting N23Ps. (**B**) Volcano plots depicting all significantly differentially expressed genes (>2-fold, FDR < 10%) in phLFs$^{+TGFβ1}$ phLFs$^{+TGFβ1+cmp4}$ and phLFs^{+cmp4} highlighting the 10 highest (red) and lowest (blue) abundant transcripts. FC, fold change. (**C**). Venn diagram showing 362 overlapping genes between 2076 deregulated genes in phLFs$^{+TGFβ1}$ and 661 deregulated genes in phLFs$^{+TGFβ1+cmp4}$. (**D**) GSEA of phLFs$^{+TGFβ1+cmp4}$ showing a negative enrichment for profibrotic gene signatures such as collagen formation, ECM organization, and smooth muscle contraction. (**E**) UMAP-RPC overlaying each gene in its cluster with color-coded transcript abundances as fold change. Up-regulated genes are depicted in red, and down-regulated genes are depicted in blue. Boxed cluster A (red) designates genes that were mostly found up-regulated in phLFs$^{+TGFβ1}$ only, and boxed cluster B (green) designates genes that were up-regulated in phLFs$^{+TGFβ1+cmp4}$ and phLFs^{+cmp4}. (**F**) On the basis of STRING DB analysis, cluster A included functional subnetworks of molecular components involved in the ECM organization (green) and the actin cytoskeleton (pink). (**G**) On the basis of STRING DB analysis cluster B included functional subnetworks of deubiquitination (yellow), laminin interactions (red), Rho guanosine triphosphatase (GTPase) effectors (green), and ECM receptor interactions (blue).

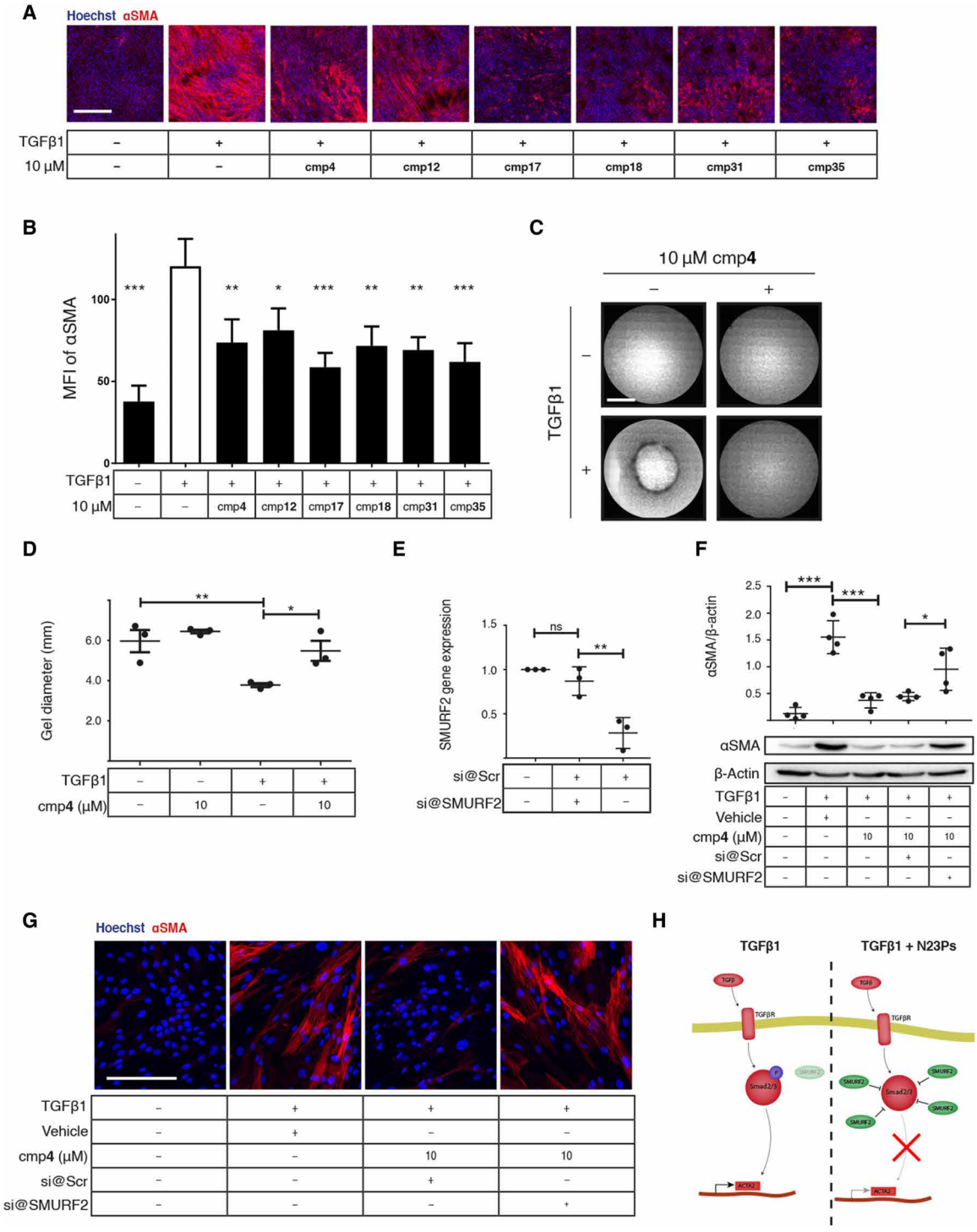


Fig. 6. Inhibition of myofibroblast transdifferentiation and contractility in a SMURF2-dependent manner. (**A**) Confocal microscopy images of phLFs concomitantly treated with TGFβ1 and active N23Ps and stained for αSMA (red) and Hoechst (blue) reducing αSMA-positive myofibroblasts. Scale bar, 500 μm. (**B**) Quantification of myofibroblasts by MFIs depicted in (A) exhibiting a significant inhibition of myofibroblast transdifferentiation. $n = 3$ (three different patient phLFs). (**C**) 3D collagen gel contractility assay (scale bar, 2000 μm) and its (**D**) quantification demonstrating a significant inhibition of cellular contractility by cmp**4**. $n = 3$ (three different patient phLFs). (**E**) Depletion of SMURF2 in IPF-phLFs by siRNA knockdown showing a significant reduction in gene expression of >70%. $n = 3$ (three different patient phLFs). (**F**) Protein expression analysis by Western blotting exhibiting a significant up-regulation of αSMA protein expression in IPF-phLFs$^{+TGFβ1+cmp4}$ depleted of SMURF2. $n = 4$ (four different patient phLFs). (**G**) Confocal microscopy images of IPF-phLFs$^{+TGFβ1+cmp4}$ depleted of SMURF2 and stained for αSMA (red) and Hoechst (blue). Scale bar, 200 μm. (**H**) Diagram depicting a possible MoA of N23Ps by a SMURF2-inhibited TGFβ1 signaling and prevention of fibroblast-myofibroblast transdifferentiation. $^*P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$. Statistics: One-way ANOVA with Bonferroni correction. All quantitative data represent means ± SEM.

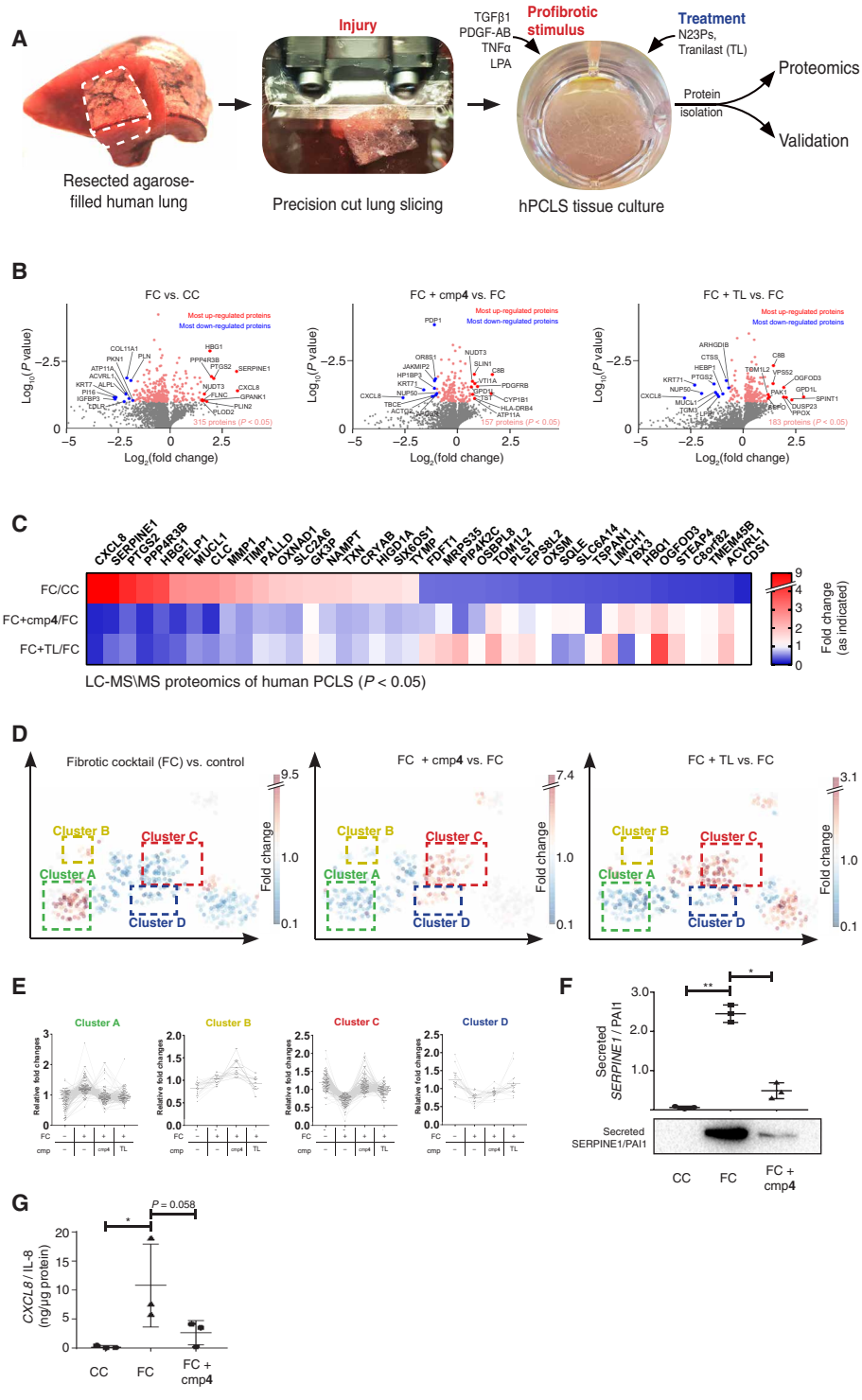


Fig. 7. A human ex vivo fibrosis model of PCLS confirms antifibrotic effects of N23Ps. (**A**) Human ex vivo fibrosis model derived from human lung resections. (**B**) Volcano plots depicting significantly ($P < 0.05$) differentially expressed proteins (pink) in FC-treated PCLS and their inhibition with cmp**4** and tranilast. The 10 highest and lowest abundant proteins are highlighted in red and blue, respectively. (**C**) Heatmap displaying protein abundancies as described in (B) and demonstrating protein deregulation as a consequence of cmp**4** and tranilast treatment. (**D**) UMAP-RPC clustering of 580 differentially expressed proteins ($P < 0.1$) according to their similarities in expression patterns of the various conditions tested, which is FC, control cocktail (CC), FC + cmp**4**, and FC + tranilast (TL). Colored boxes show clusters of proteins that display common abundancies between the different conditions tested. (**E**) Analysis of commonly regulated protein abundancy clusters found in (D). (**F**) Immunoblotting and quantification showing diminution of secreted profibrotic PAI-1/SERPINE1 by cmp**4** in living hPCLS. (**G**) ELISA data demonstrating diminution of profibrotic CXCL8/IL-8 by cmp**4** in living hPCLS. *$P < 0.05$ and **$P < 0.01$. Statistics: One-way ANOVA with Bonferroni correction. All quantitative data represent means ± SEM. $n = 3$ (three different patient PCLS).

Notably, a substantial number of proteins within this short list are related to tissue fibrosis (indicated by red asterisks in fig. S11A), and most of these proteins were highly abundant in hPCLS^{+FC} but of low abundance in hPCLS$^{+FC+cmp4}$ and hPCLS$^{+FC+tranilast}$ (Fig. 7C). Conversely, specific low abundant proteins in hPCLS^{+FC} were mostly enriched in hPCLS$^{+FC+cmp4}$ and hPCLS$^{+FC+tranilast}$ (Fig. 7C). Next, we applied UMAP-RPC (fig. S11B) on 580 proteins ($P < 0.1$) to generate clusters of similar protein abundances (Fig. 7D). We identified four unique protein clusters: Cluster A contained proteins that were highly abundant in hPCLS^{+FC}, but not in hPCLS$^{+FC+cmp4}$ or hPCLS$^{+FC+tranilast}$. Vice versa, cluster C enclosed proteins of low abundance in hPCLS^{+FC}, but of high abundance in hPCLS$^{+FC+cmp4}$ or hPCLS$^{+FC+tranilast}$. Cluster B and cluster D exhibited different protein abundance patterns for hPCLS$^{+FC+cmp4}$ and hPCLS$^{+FC+tranilast}$, which again could strongly advocate for a different MoA of N23Ps versus tranilast in human lung tissue. Analysis of functional protein interactions of cluster A by STRING database resulted in subnetworks of fibrosis-related proteins involved in ECM organization (HSA-1474244), actin cytoskeleton (GO:0015629), and IL signaling (HSA-449147) (fig. S11, C and D). Notably, these subnetworks were overlapping with those we found before in the transcriptomics data of phLFs (fig. S9A). Next, we experimentally validated single fibrosis-relevant proteins within the detected subnetworks in living hPCLS^{+FC} and hPCLS$^{+FC+cmp4}$ by immunoblotting and enzyme-linked immunosorbent assay (ELISA). Here, secreted PAI-1/SERPINE1 and cytoplasmic CXCL8/IL-8, which were the two highest abundant proteins found in hPCLS^{+FC} proteomics, were significantly expressed in hPCLS^{+FC}, but not in untreated controls as well as in hPCLS$^{+FC+cmp4}$ (Fig. 7, F and G). These findings clearly demonstrate an antifibrotic activity for novel N23Ps in the human ex vivo fibrosis model of IPF. Conclusively, our data strongly support a unique antifibrotic therapy route for human lung fibrosis by novel highly potent N23Ps, which mechanistically act through an inhibitory mode within the TGFβ1 network of protein (de)ubiquitination mediated by SMURF2 expression up-regulation.

DISCUSSION

In this study, we have used human lung tissue, in cellulo by applying human IPF-derived primary lung fibroblasts and ex vivo by hPCLS, to set up a highly relevant phenotypic screening and validation platform for chemical compounds, respectively. The high-throughput screening assay provides a robust and clinically relevant readout by quantification of ECM deposition. Hit identification was accomplished by an artificial intelligence (AI) model (FANTAIL) on the basis of deep learning CNNs analyzing 3D confocal microscopy images. In the screening phase, we identified tranilast, a cinnamoylanthranilate, as hit compound and subsequent SAR studies led to the discovery of a novel highly potent class of ECM deposition inhibitors, which were N23Ps. Mechanistically, we identified a regulatory subnetwork within the TGFβ1 signaling network of SMAD (de)ubiquitination specifically affected by N23Ps. To the best of our knowledge, neither a chemically similar compound nor compounds displaying similar biological effects are known. This pathway might act as a yet unidentified druggable target network in clinical antifibrotic therapy strategies. Last, to demonstrated human disease relevance, antifibrotic efficacy of N23Ps was successfully illustrated in a fibrosis ex vivo model based on human living PCLS.

The time-consuming and costly endeavor of drug discovery essentially discriminates two screening approaches: target- and phenotype-based projects. The prerequisite of target-based screening platforms is exact knowledge about a single disease-relevant molecular target through a given MoA. However, biology is highly complex and interconnects through intricate networks. Thus, many compounds found by target-based screening attempts lastly fail in preclinical and clinical trials (*33*). Besides, preclinical in vivo disease models are also often nonpredictive and produce misleading results, which cannot be translated into the clinics. For instance, the extensively used murine experimental lung fibrosis bleomycin model fails in recapitulating features of IPF. However, more advanced disease models of repetitive bleomycin injury, radiation, or virally induced models do exist (*47*, *48*). However, despite a vast number of antifibrotic compounds preclinically tested in vivo, only nintedanib and pirfenidon successfully proved clinical efficacy. Both drugs slow down disease progression but do not stop it (*49*, *50*). Therefore, innovative drug discovery platforms that rely on special key criterions might be more successful for translational efforts. Three of these criteria were reviewed and defined by Vincent *et al.* (*33*): (i) usage of primary patient-derived primary cells, (ii) disease-relevant stimuli that produce desired phenotypes, and (iii) assay readouts close to the clinical end point. Here, we used phLFs derived from lung explants from patients with IPF. The functionality and phenotypic characteristics of these phLFs cultured in vitro was intensively investigated in the past (*35*, *51*). phLFs showed consistent expression of fibroblastic surface markers until passage 7 but turned senescent from passage 10 on (*51*). This gives a reasonable window to expand phLFs between passages 1 to 5, which generates enough cells to lastly screen up to 30,000 compounds or even more. Reportedly, IPF-derived primary human fibroblasts are fundamentally different in their genomic profile and consequently have altered signaling pathways when compared to normal fibroblasts (*52*). However, these isolated phLFs were not subjected to long-term in vitro culturing, which per se still might introduce selective, functional, or even genetically modifications during long-term tissue culture (*53*). In our own observations, unchallenged IPF-derived fibroblasts did not primarily deposit more ECM when compared to normal fibroblasts. Therefore, we used a reliable trigger-able fibrotic system in vitro by applying profibrotic TGFβ1. Thus stimulated, the IPF-phLFs excessively produced and deposited massive insoluble ECM fibers, which we exploited as a robust assay readout. Deposited ECM plays a major role in the regulation of tissue architecture, mechanosensing, and cytokine release and in clinically relevant mechanics and diffusion capacity of the lung. Equally important, all molecular targets upstream of the measured end point can be chemically modulated, thus possibly finding novel ECM deposition inhibitors and molecular targets.

Moreover, we developed an end-to-end deep learning model based on the analysis of thousands of images of immunofluorescently stained ECM acquired by automated high-throughput microscopy. FANTAIL was set up as a convolutional neural network (CNN) and trained to identify qualitative and quantitative differences between treated and untreated phLFs, which means classifying "healthy" from "fibrotic" ECM, as well as detecting compound-related cytotoxic effects. We demonstrated feasibility of training an end-to-end deep learning CNN from a relatively small training set using aggressive data augmentation and fragmentation of original high-resolution 3D microscopy images into tiles. This reduces efforts and costs to produce large, high-quality training dataset as a major limitation of deep-learning CNN usage. Together, the role of deep learning in

drug discovery has been manifold and finds applications in compound activity predictions, generation of new chemical structures, chemical reaction predictions, calculating ligand-protein interactions, and biological imaging analysis (*54*). CNNs were applied to biological image analysis based on immunofluorescent mostly for feature extraction purposes, subtyping of cells, tracking of cells, colony counting, or even in a label-free fashion (*55*). Recently, in dermatology, a deep-learning model based on a CNN was successfully applied to image data of skin lesions for an automated diagnosis of skin cancer (*56*). To the best of our knowledge, so far, no deep-learning models have been developed that analyze complex ECM structures based on high-resolution microscopy immunofluorescent images as presented in this study. FANTAIL detects patterns inside the complex ECM images regardless of their orientation, splitting factor, distortion, or truncation. In contrast, shallow machine learning algorithms, or even non-AI, hit detection systems, often rely on manually engineered image features, which could lead to a loss of crucial information in the images, especially when dealing with images displaying a network of intricate and delicate ECM fibers. Moreover, pretrained AI hit detection systems can be optimized in hit accuracy by applying appropriate training sets. The aspect of accuracy in hit detection, especially specificity, becomes highly critical particularly when huge libraries will be tested.

FANTAIL identified tranilast in our screening data as an efficient antifibrotic hit compound. In a second step, on the basis of SAR studies, we found and developed novel derivatives of tranilast by introducing a 2-butoxy substitution at R1 (N23Ps), which resulted in a notably higher ECM deposition inhibition potency (>100-fold). Unexpectedly, N23Ps additionally affected fibroblasts causing a switch from elongated to round cell shapes. Contrarily, tranilast only exhibited decreasing matrix deposition but no cell morphology changes. This might hint at a possible unique MoA of N23Ps. The MoA of tranilast is still elusive, although inhibitory effects on cell proliferation are reported to be related with TGFβ1 signaling (*57*). However, it is reported that the carboxy function of tranilast is required for displaying optimal inhibitory effect in TGFβ1-stimulated collagen synthesis in cultured mesangial cells (*58*). By microarray-based transcriptomics and a unique subsequent UMAP-RPC embedding, we were able to identify unique signaling networks that were affected by N23Ps. One of these functional networks included the Gene Ontology (GO) term "deubiquitination" (REACTOME: HSA-5688426), including SMURF2 as one of the most up-regulated genes. SMURFs, or SMAD-ubiquitination-related factors, are E3 ubiquitin ligases that inactivate SMADs by ubiquitination and subsequent proteasomal degradation. As N23Ps induced the expression of SMURF2 and thus blocked TGFβ1 signaling, this specific route might highlight an innovative antifibrotic strategy for clinical therapy (Fig. 6H).

One potential limitation of the presented screening approach is that phLFs would become senescent starting at passage 10 (*51*), which possibly limits the screening of extremely large compound libraries. We have successfully expanded phLFs for the screening of libraries up to 30,000 compounds, and even a larger numbers of phLFs could be generated for bigger screens. Besides, the current workflow of the ECM deposition assay allows de facto only screening for TGFβ1 effects preventing transdifferentiation of phLFs into myofibroblasts. This strategy opens new avenues for the discovery of new molecular entities (NMEs) or molecular targets that, if effective, are capable of stopping disease progression not only in IPF and preventing acute disease exacerbations (*59*, *60*) but also for fibrosis in other organs such as skin and liver. Moreover, ECM resolution would be favorable, which might trigger regenerative processes and consequently confer symptomatic benefits for patients. However, such a screening approach would require an altered assay setup. In addition, comparing transcriptomics of the in vitro human cellular fibrosis model with proteomics from the ex vivo PCLS fibrosis model might produce a potential separation between the detected pathways due to different complexities found in both models. Moreover, technical issues arise, like applying agarose-filling for PCLS generation, which interestingly interferes with RNA extraction leading to low-quality RNA with consequences for downstream processes such as transcriptional analysis. However, in the future, our laboratory will apply advanced single-cell genomics derived from the ex vivo fibrosis model and from animal disease models, both of which will shed unprecedented insights into the MoA of antifibrotic compounds. Thus, N23Ps as a new and highly potent antifibrotic compound class might successfully translate and stop disease progression in fatal fibrotic lung diseases such as IPF and fibrosis affecting other organs.

MATERIALS AND METHODS

Study design

The purpose of this study was to identify novel antifibrotic drug candidates that inhibit ECM deposition of IPF-derived myofibroblasts. This objective was pursued through (i) the development and validation of a high-throughput/-content ECM deposition assay suitable for drug screening, (ii) the identification and confirmation of ECM deposition inhibitors, (iii) development of highly potent derivatives of hit compounds by SAR studies, (iv) target network identification, and (v) validation of SAR-optimized compounds in an human ex vivo injury and early fibrosis model. Research material included phLFs isolated from IPF patient lung explants, as well as normal healthy phLFs and ex vivo tissue culture derived from healthy tumor-free peritumoral tissue resections of patients with lung cancer undergoing lobectomy. High-throughput/-content screening of FDA libraries was performed with primary lung fibroblasts derived from one patient with IPF. For all in vitro validation and dose-response experiments, three to five independent biological replicates, which are different patients, were used. All experiments used up to four technical replicates. No outlier exclusion strategy was applied. Investigators were not blinded during data analysis. In experiments involving pharmacological treatments, vehicle controls were used at all times. For comparison of pharmacological effects, same concentrations of vehicle controls were used, but never higher that 1% (v/v) in the final culture medium. Specific pharmacological effects were only stated if (i) a significant effect compared to vehicle control and (ii) a dose-response relationship were observed, and (iii) a closely related analog, presumed to be inactive, could serve as negative control.

Human tissue and ethics statement

Human tissue has been obtained from the Comprehensive Pneumology Center cohort of the BioArchive CPC-M at the University Hospital Grosshadern of the Ludwig Maximilian University (Munich, Germany) and by the Asklepios Biobank of Lung Diseases (Gauting, Germany). Participants provided written informed consent to participate in this study, in accordance with approval by the local ethics committee of the LMU, Germany (Project 333-10, 454-12).

Primary cell culture

phLFs were isolated by outgrowth from human lung tissue derived from lung explants or tumor-free areas of lung resections as previously described (*61*, *62*). Cells were cultured in Dulbecco's modified Eagle's medium (DMEM) F-12 with 20% (v/v) special processed fetal bovine serum (FBS; PAN-Biotech) and penicillin (100 IU/ml) and streptomycin (100 μg/ml). Medium was changed every 2 to 3 days, and cells were passaged at 80 to 90% confluency in a ratio of 1:5 or 1:6. Cells were used for experiments until passage 7. For ECM deposition drug screening, 0.5 to 1 × 10^6 cells were expanded from passages 1 to 5, each time in a ratio of 1:6. More than 100 × 10^6 cells were trypsinized at passage 5 and cryopreserved in 90% (v/v) FBS and 10% (v/v) dimethyl sulfoxide (DMSO). Cells were frozen slowly using Mr. Frosty (Thermo Fisher Scientific) freezing containers. For reseeding, phLFs were thawed in a water bath at 37°C, and the cells were washed with culture medium, before plating. After reaching confluency in passage 6, cells were used for the ECM deposition assays. Primary human dermal fibroblast (catalog no. DF-F) were purchased from ZenBio Inc. and cultured according to the manufacturer's instructions.

ECM deposition assay

phLFs were cultured in DMEM F-12 medium with 20% FBS and antibiotic supplement as mentioned above. Cells were seeded with 6000 cells per well in 384-well CellCarrier plates (PerkinElmer, catalog no. 6007550). Following overnight incubation, cells were starved in serum-reduced medium [1% FBS with 0.1 mM 2-phosphoascorbate (Sigma-Aldrich, catalog no. 49752)] for 24 hours. Afterward, cells were treated with TGFβ1 (1 ng/ml) or vehicle, and, additionally, small molecules or appropriate vehicle controls were added. After 72 hours of incubation, medium was changed for starving medium with Alexa Fluor 488 fluorophore–conjugated anti–collagen type 5 antibodies (1 μg/ml; Santa Cruz Biotechnology, Clone C-5, catalog no. sc-166155 AF488), Alexa Fluor 555 fluorophore–conjugated anti–collagen type 1 antibodies (0.66 μg/ml; Rockland, catalog no. 600-401-103-0.1), Alexa Fluor 637 fluorophore–conjugated anti–fibulin 1 antibodies (1 μg/ml; Santa Cruz Biotechnology, Clone C-5, catalog no. sc-25281 AF647), and Hoechst H33342 (1 μg/ml; Sigma-Aldrich). Fluorescence-conjugation of the collagen type 1 antibody (Ab) was performed using the Alexa Fluor 555 Protein Labeling Kit (Invitrogen, catalog no. A20174) according to the manufacturer's instructions. Labeling efficacy was controlled by photometrical means. Following 4 hours of incubation, cells were washed three times with phosphate-buffered saline (PBS) and fixed with paraformaldehyde. For automated liquid handling in 384-well plates, an INTEGRA Assist Plus (INTEGRA, Zizers, Switzerland) equipped with an INTEGRA Viaflo II pipette (INTEGRA, Zizers, Switzerland, catalog no. 4642), 125-μl GripTips pipette tips (INTEGRA, Zizers, Switzerland, catalog no. 6464), and sterile reagent reservoirs (INTEGRA, catalog no. 4311) were applied. All automated pipetting steps with the INTEGRA Assist Plus were performed at 9.5 μl/s to ensure proper integrity and attachment of the deposited ECM to the culturing surface within the wells of the 384-well plates. During cell seeding, the automated liquid handling was performed at 89.3 μl/s. Removal of liquids from the well plates was performed by manually inverting the plates. Following fixation, the automated imaging was achieved using a confocal laser scanning microscope (LSM 710, Zeiss) with automated focus detection for 3D image acquisition (1024 pixel by 1024 pixel by 9 pixel that equals a dimension of 1417 μm by 1417 μm by 16 μm). For postacquisition analysis, images were imported into IMARIS software (Bitplane) and volume detection or alternatively quantification of the mean fluorescence intensity (MFI), and Hoechst-stained cell nuclei were automatically counted using Imaris' spot detection algorithm.

hPCLS and FC treatment

PCLS were prepared as described before (*40*, *46*). Briefly, PCLS were prepared from tumor-free peritumor tissue. The lung tissue was inflated with 3% agarose solution and solidified at 4°C. Tissue blocks were cut in 500-μm-thick PCLS using a vibration microtome Hyrax V50 (Zeiss). PCLS were cultured in DMEM F-12 medium and treated with a pro-FC, as described before (*40*), or vehicle, as well as with small molecules or vehicles for 7 days. After culturing and treatments, supernatants were harvested. PCLS were washed in PBS, and protein was extracted as previously described (*63*). Briefly, PCLS were pooled in an Eppendorf tube and lysed in 500 μl of ice-cold radioimmunoprecipitation assay (RIPA) buffer [50 mM tris-Cl (pH 7.4), 150 mM NaCl, 1% NP-40, and 0.25% Na-deoxycholate] containing 1× Roche complete mini protease inhibitor cocktail (Roche, catalog no. 11697498001). After an incubation of 2 hours rotating at 4°C, the lung slices were removed from the lysates, and the protein content was measured.

Cytotoxicity assays

Viability/Cytotoxicity Assay Kit for Animal Live and Dead Cells was obtained from Biotium (catalog no. 3002). CellEvent Caspase 3/7 Green Detection Reagent was acquired from Invitrogen (catalog no. C10423). For MTT [3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide] assays, Thiazolyl Blue Tetrazolium was bought from Sigma-Aldrich (M5655-1G). All these kits and assays were used according to the manufacturer's instructions.

Antibodies for immunofluorescence and dyes

For immunofluorescence microscopy, the following antibodies were used: monoclonal mouse anti-collagen type 5 (1 mg/ml) from Sigma-Aldrich (catalog no. sc-166155), monoclonal mouse anti-collagen type 5 Alexa Fluor 488 conjugate from Sigma-Aldrich (catalog no. sc-166-155 AF488), polyclonal rabbit anti–collagen type 1 from Rockland (catalog no. 600-401-103-0.5), monoclonal mouse anti–fibulin 1 from Santa Cruz Biotechnology (catalog no. sc-25281), monoclonal mouse anti–fibulin 1 Alexa Fluor 647 conjugate (1 mg/ml) from Santa Cruz Biotechnology (catalog no. sc-25281 AF647), and polyclonal rabbit antifibronectin (1 mg/ml) from Santa Cruz Biotechnology (catalog no. sc-9068). Hoechst-33342 was obtained from Sigma-Aldrich (catalog no. B2261). The following secondary antibodies were used: Alexa Fluor 488 donkey anti-mouse Ab (Invitrogen, catalog no. A21202), Alexa Fluor 568 donkey anti-mouse Ab (Invitrogen, catalog no. A11004), and Alexa Fluor 568 donkey anti-mouse Ab (Invitrogen, catalog no. A11011). For immunofluorescence stainings of actin stress fibers, Alexa Fluor 568 Phalloidin (Invitrogen, A12380) was used. 4′,6-Diamidino-2-phenylindole (DAPI) was acquired from Sigma-Aldrich (catalog no. D9564).

Immunocytochemistry

For standard immunofluorescence staining, 5000 phLFs were seeded into 96-well imaging plates with a flat bottom (BD Biosciences, catalog no. 353376). After incubation, cells were fixed with either 4% paraformaldehyde for 30 min at 37°C or 100% methanol for 2 min at −20°C. If needed, then phLFs were permeabilized with

0.25% (v/v) Triton X-100 in PBS for 15 min. After washing with 100 µl of PBS, blocking was performed by incubation with 5% (w/v) bovine serum albumin (BSA) in PBS for 1 hour. Primary antibodies were diluted in 1% BSA (Sigma-Aldrich) in PBS, incubated for 16 hours at 4°C, and subsequently washed three times with PBS for 20 min each. Secondary antibodies were diluted in 1% BSA (Sigma-Aldrich) in PBS, incubated for 1 hour at room temperature, and subsequently washed three times with PBS for 20 min each. Paraformaldehyde (4%) in PBS (w/v) was prepared from paraformaldehyde from Sigma-Aldrich (catalog no. 15,812-7). BSA was obtained from Sigma-Aldrich (catalog no. A3059). Triton X-100 was obtained from AppliChem (catalog no. A1388).

Confocal 3D and 4D imaging

Confocal time-lapse microscopy was implemented on an LSM 710 system (Carl Zeiss) containing an inverted AxioObserver.Z1 stand equipped with phase-contrast and epi-illumination optics and operated by ZEN2009 software (Carl Zeiss). The following objectives were used for imaging: EC Plan-Neofluar 20×/0.8 NA (numerical aperture) (Carl Zeiss), LD C-Apochromat 40×/1.1 NA water objective lens (Carl Zeiss), and LCI PLN-NEOF DICIII 63×/1.30 NA water objective lens (Carl Zeiss). For 4D imaging, the cells were kept in an incubation chamber (Carl Zeiss) under standard cultivation conditions (37°C and 5% CO_2). Thickness of single confocal layers within the z-stacks was set according to optimized values suggested by the ZEN2009 software. The confocal datasets were either maximum intensity projected in the ZEN2009 software (Carl Zeiss) and/or imported into Imaris 9.0.0-9.3.1 software (Bitplane) for analysis.

Protein isolation, SDS–polyacrylamide gel electrophoresis, Western blotting, and ELISA

Cells were scraped off the plastic dish directly into 200 µl of ice-cold RIPA buffer containing 1× Roche complete mini protease inhibitor cocktail. After incubating the samples for 30 min on ice, insoluble material was removed by centrifugation at 14,000*g* for 15 min at 4°C, and the supernatant was further processed. Samples were mixed with 50 mM tris-HCl (pH 6.8), 100 mM dithiothreitol (DTT), 2% SDS, 1% bromophenol blue, and 10% glycerol, and proteins were separated using standard SDS–10% polyacrylamide gel electrophoresis. For immunoblotting, proteins were transferred to polyvinylidene difluoride (0.45 or 0.2 µm; Millipore, Billerica, MA, USA) membranes, which were blocked with 5% milk in 0.1% Tween 20 in tris-buffered saline and incubated with primary, followed by horseradish peroxidase (HRP)–conjugated secondary antibodies over night at 4°C and at room temperature for 1 hour, respectively. For immunoblotting, the following primary antibodies were used: monoclonal mouse anti–collagen type 5 (1 mg/ml) from Sigma-Aldrich (catalog no. sc-166155), polyclonal rabbit anti–collagen type 3 (1 mg/ml) from Rockland (catalog no. 600-401-105), polyclonal rabbit anti–collagen type 1 (1 mg/ml) from Rockland (catalog no. 600-401-103-0.5), monoclonal mouse anti–fibulin 1 (1 mg/ml) from Santa Cruz Biotechnology (catalog no. sc-25281), polyclonal rabbit anti-fibronectin (1 mg/ml) from Santa Cruz Biotechnology (catalog no. sc-9068), and monoclonal mouse anti–β-actin–peroxidase (1:10,000; Sigma-Aldrich, AC-15). Goat anti-rabbit and goat anti-mouse immunoglobulin G conjugated to HRP (1:10,000; Cell Signaling Technology) were applied as secondary antibodies. CXCL/IL-8 concentrations were determined using Human IL-8/CXCL8 DuoSet ELISA (DY208-05) according to the manufacturer's protocol.

mRNA isolation, cDNA synthesis, and quantitative reverse transcription polymerase chain reaction

RNA extraction from cultured phLFs was performed using the PeqGold RNA Kit (Peqlab) according to the manufacturer's instruction. The concentration of the isolated RNA was assessed spectrophotometrically at a wavelength of 260 nm (NanoDrop 1000). cDNA was synthesized with the GeneAMP Polymerase Chain Reaction (PCR) Kit (Applied Biosystems, Foster City, CA, USA) using random hexamers using 1 µg of isolated RNA for one 301 reaction. Denaturation was performed in an Eppendorf Mastercycler with the following settings: 302 303 lid = 45°C, 70°C for 10 min and 4°C for 5 min. Reverse transcription was performed in an Eppendorf Mastercycler with the following settings: lid = 105°C, 20°C for 10 min, 42°C for 60 min, and 99°C for 5 min. Quantitative reverse transcription PCR (qRT-PCR) reactions were performed in triplicates with SYBR Green I Master in a LightCycler 480II (Roche, Risch, Switzerland) with standard conditions: 95°C for 5 min followed by 45 cycles of 95°C for 5 s (denaturation), 59°C for 5 s (annealing), and 72°C for 20 s (elongation). Target genes were normalized to hypoxanthine-guanine phosphoribosyltransferase expression. All human primer sequences are documented in table S1.

Microarray and UMAP-RPC

Total RNA was isolated PEQGold Total RNA Kit (Peqlab) according to the manufacturer's instructions including genomic DNA elimination. The Agilent 2100 Bioanalyzer was used to assess RNA quality and RNA with an RIN (RNA integrity number) of >7 was used for microarray analysis. Total RNA (150 ng) was amplified using the WT PLUS Reagent Kit (Thermo Fisher Scientific Inc., Waltham, USA). Amplified cDNA was hybridized on Human ClariomS arrays (Thermo Fisher Scientific). Staining and scanning (GeneChip Scanner 3000 7G) was performed according to the manufacturer's instructions. Transcriptome Analysis Console (version 4.0.0.25, Thermo Fisher Scientific) was used for quality control and to obtain annotated normalized SST-RMA (Signal Space Transformation in conjunction with the regular robust multiple-array average normalization method) gene-level data. Statistical analyses were performed using the statistical programming environment R (R Development Core Team Ref1). Gene-wise testing for differential expression was performed using the paired limma *t* test and Benjamini-Hochberg multiple testing correction (FDR < 10%). To reduce background, gene sets were filtered using DABG (detection above background) $P < 0.05$ in at least one sample per pair and in at least two of three pairs per analysis. Heatmaps were generated using GraphPad Prism v7. The RPC was based on UMAP (*64*). Code is provided in code sections S3 and S4. mRNA abundancies from the microarray data were normalized (as seen as an example in fig. S11B), and abundancies of all four different conditions were summarized in a linear vector (fig. S11A) that was projected into a bidimensional space using UMAP (fig. S11C). Then, clusters of genes were extracted. Gene/protein interactions were visualized using the String Database (www.string-db.org).

SAR—Drug library and medicinal chemistry

The FDA-approved Drug Library (1509 compounds) was obtained from MedChemExpress (catalog no. HY-L022). Screening of the drug library was performed at 100 µM. Compounds 1 to 22 for the SAR studies were acquired from the following suppliers compiled in table S2. Synthesis of (E)-3-(3,4-dimethoxyphenyl)-*N*-(4-fluoro-2-propoxyphenyl)acrylamide (= compound 29): (*E*)-3,4-Dimethoxycinnamic acid (100 mg,

0.48 mmol, and 1.0 eq.) were dissolved in dichloromethane (1.5 ml). Oxyma (137 mg, 0.96 mmol, and 2.0 eq.), EDC (3-Dimethylaminopropyl)-ethyl-carbodiimide)-HCl (184 mg, 0.96 mmol, and 2.0 eq.), and sodium bicarbonate (40 mg, 0.48 mmol, and 1.0 eq.) were subsequently added under inert atmosphere. Stirring was continued for 10 min at 0°C. 2-Butoxy-4-fluoro aniline (81.3 mg, 0.48 mmol, and 1.0 eq.) was added, and the reaction was allowed to warm to room temperature and stirred overnight. The reaction mixture was poured on 5 ml of 1 M HCl and extracted with dichloromethane (3×). The combined organic layer was extracted with 1 M NaOH and brine and dried over sodium sulfate, and the solvent was removed in vacuo. After silica gel chromatography (petrol ether/ethyl acetate), the product was obtained as a solid (68.9 mg, 0.187 mmol, 39%). R_f = 0.3 PE/EE (petrol ether/ethyl acetate) = 3:1. ^{1}H nuclear magnetic resonance (NMR): (400 MHz, $CDCl_3$) δ = 8.52 to 8.42 parts per million (ppm) (m, 1H), 7.76 ppm (s, 1H), 7.66 ppm (d, J = 15.4 Hz, 1H), 7.14 ppm (dd, J = 8.3, 1.9 Hz, 1H), 7.06 ppm (d, J = 1.9 Hz, 1H), 6.87 ppm (dd, J = 8.2, 4.1 Hz, 1H), 6.70 to 6.60 ppm (m, 2H), 6.41 ppm (d, J = 15.4 Hz, 1H), 3.99 ppm (t, J = 6.7 Hz, 2H), 3.92 ppm (d, J = 7.1 Hz, 7H), 1.95 to 1.82 ppm (m, 2H), 1.08 ppm (t, J = 7.4 Hz, 3H). ^{13}C NMR: (100 MHz, $CDCl_3$) δ = 163.94, 150.93, 149.28, 142.06, 127.76, 124.30, 122.29, 118.98, 111.21, 110.07, 106.93, 106.72, 99.77, 99.50, 70.64, 56.08, 22.48, and 10.62 ppm. LC-MS: mass/charge ratio (m/z) calculated ($C_{20}H_{22}FNO_4$): 359.15, found 359.94 ($[M + H]^+$), t_R = 1.32 min. All noncommercial compounds were synthesized following the procedure given above. Analytical data for compounds 24 to 35 can be found in table S3. All chemicals were diluted in DMSO and added to the cell culture medium (v/v) in final concentrations up to 1% DMSO. Vehicle controls contained DMSO only. DMSO was obtained from Sigma-Aldrich (catalog no. D4540).

FANTAIL—Inferential classification and detection model for the inhibition of ECM deposition

The KERAS high-level API (https://github.com/fchollet/keras/) with TensorFlow implementation was used to train CNN on a complex image detection and classification task. The CNN design (Fig. 3A) followed the most accepted guidelines, as in (*65*). The best convolutional process was reached with three convolutional layers convoluting the images with 24 filters per layer and pooling out data with a 2 by 2 pooling matrix in the convolutional layers. The specific image classification and detection task was based on the detection of interspersed fibrotic and cellular patterns with frequent image edge pattern interruptions. The dimensional orientation of the fibrotic patterns appeared randomly oriented with various shape, size, and clustering on images of large dimension (1024 by 1024 by RGB). Hence, the Rectified Linear Unit activation was used as the activation mode to detect pattern edges and the adadelta optimizer was chosen for an efficient CNN learning process. A dataset of image controls was used to train the classifier. This dataset consisted of 295 immunofluorescence images annotated as "toxic" (treated with 5% ethanol), 390 images annotated as fibrotic (treated with TGFβ1), and 390 images annotated as "normal" (untreated). Images with the annotation normal were labeled as "hits," while those with the annotations toxic and fibrotic were combined under the label "others." Images were randomly assigned to the training dataset (75%) and the validation dataset (25%). As 1024 by 1024–sized images would be an unusually large input for a CNN, we aimed to test the CNN efficiency of np × np large subsets of each image with np ∈ 128,256,512 pixels. Aggressive data augmentation was performed by fragmenting each $m \times m$ dimensional image (m = 1024) in $\left(\left(\frac{m}{\mathrm{np}} - 1\right) \times 4 + 1\right)^2$ tiles with a three-fourths tile overlap (fig. S3A) to increase redundancy and pattern fragmentation. For each image M with the dimensions $m \times m$, the np × np–sized tiles T were produced as follows

$$T = M[o_x, o_x + 1, o_x + 2, \ldots, o_x + \mathrm{np}; o_y, o_y + 1, o_y + 2, \ldots, o_y + \mathrm{np}]$$

$$o_x, o_y \in \bigcap_{k \in \left[0,1,\ldots,\left(\left(\frac{m}{\mathrm{np}}-1\right)\times 4\right)\right)} k \times \frac{n}{4}$$

Each data tile T was rotated by θ ∈ {0 °,90 °,180 °,270 °}, representing different spatial orientations of the ECM (fig. S3A). Hence, $4 \times \left(\left(\frac{m}{\mathrm{np}} - 1\right) \times 4 + 1\right)^2$ tiles were saved from each original image leading to a significant augmentation of data (100-fold for np = 512, 676-fold for np = 256, and 3364-fold for np = 128). The CNN as shown in Fig. 3A was trained on all fragments of the training set with fragment size np ∈ 128,256,512 and n = 3. Code is provided in code sections S1 and S2. Learning curves are shown in fig. S3 (C and D). Quickest and stable convergence was observed with 512 by 512 pixels imaging, expectedly considering the larger amount of information per fragment. Then, trained models were evaluated to predict the annotation of an original image of the validation set. Therefore, each original image was fragmented into $\left(\frac{m}{\mathrm{np}}\right)^2$ not overlapping np × np–sized tiles (fig. S3B), and the share of tiles classified as hits was calculated. To determine the accuracy of prediction for original images independent of a cutoff value for the share of tiles classified as hits, we used receiver-operating characteristic analysis (GraphPad Prism v7) and calculated the area under the curve (AUC) for each deep learning model. Different numbers of learning iterations (epochs) and different tile sizes of np × np were tested. In general, accuracy (given as AUC) increased with the number of epochs. After, 10,000 epochs models with np = 256 and np = 128 yielded a similarly high AUC = 0.999 (fig. S3E). Therefore, we chose model 7 with np = 128, n = 3 for further experiments.

UMAP for image clustering

The image clustering chosen was performed using the UMAP, a widely used manifold learning technique for dimension reduction. UMAP is constructed from a theoretical framework based in Riemannian geometry and algebraic topology (*64*). Each $m \times m$–dimensional image pixel matrix (m = 1024) is flattened as a linear vector (fig. S3G) and projected using UMAP on a 2D space (Fig. 3C). It was found in this study that the UMAP approach generated 2D data maps that logically respected the AI images classification performed in the present study putting images into a bidimensional graph that helped to graphically understand the concept of true positives hits and identify images with artifacts not recognized by the AI.

Quantification of αSMA content in phLFs

Six thousand cells per well phLFs were seeded in 384-well CellCarrier plates. Following overnight incubation, cells were starved in serum-reduced medium (1% FBS) for 24 hours. Afterward, cells were treated with TGFβ1 (1 ng/ml) and different compounds. After 48 hours, cells were fixed with 100% ice-cold methanol. Cells were stained for DAPI and αSMA Ab conjugated to Cy3 (Sigma-Aldrich, catalog no. C6198-2ML). For automated liquid handling in 384-well plates, an INTEGRA Assist Plus was used. Following fixation, the automated imaging was achieved using a confocal laser scanning microscope

(Zeiss, LSM 710) with automated focus detection for 3D image acquisition (ECM Deposition Assay). Images were analyzed by measuring the MFI of the αSMA signal in Zen Blue v2.5 (Zeiss).

Contractility assay of phLFs

In a 96-well imaging plate, 50 µl of 3D collagen gels were casted per well as described before (*62*), and 20,000 phLFs per well were seeded on top. Cells were treated with TGFβ1 (1 ng/ml) and/or cmp**4**. After 72 hours, cells were fixed with 4% paraformaldehyde. Collagen gels were imaged using an AxioImager2 (Zeiss), and the gel diameter was determined using Zen Blue v2.5 (Zeiss).

LC-MS/MS of PCLS

Each 10 µg of protein extract was digested using a modified FASP (filter aided sample preparation) protocol (*66*, *67*). Briefly, proteins were reduced and alkylated using DTT and iodoacetamide and diluted to 4 M urea before centrifugation on a 30-kDa filter device (Sartorius). After several washing steps using 8 M urea and 50 mM ammonium bicarbonate, proteins were digested on the filter by Lys-C and trypsin overnight. Generated peptides were eluted by centrifugation, acidified with trifluoroacetic acid and stored at −20°C. Samples were measured on a QExactive HF-X MS (Thermo Fisher Scientific) online coupled to an Ultimate 3000 nano-RSLC (Dionex). Tryptic peptides were automatically loaded on a trap column (300-µm inner diameter × 5 mm, Acclaim PepMap100 C18, 5 µm, 100 Å, LC Packings) before C18 reversed-phase chromatography on the analytical column (nanoEase MZ HSS T3 Column, 100 Å, 1.8 µm, 75 µm by 250 mm, Waters) at flow rate (250 nl/min) in a 95-min nonlinear acetonitrile gradient from 3 to 40% in 0.1% formic acid. Profile precursor spectra from 300 to 1500 m/z were recorded at 60,000 resolution with an automatic gain control (AGC) target of 3×10^6 and a maximum injection time of 30 ms. Subsequently, TOP15 fragment spectra of charges 2 to 7 were recorded at 15,000 resolution with an AGC target of 1×10^5, a maximum injection time of 50 ms, an isolation window of 1.6 m/z, a normalized collision energy of 28, and a dynamic exclusion of 30 s. Generated raw files were analyzed using Progenesis QI for proteomics (version 4.1, Nonlinear Dynamics, part of Waters) for label-free quantification as described (*68*, *69*). Features of charges 2 to 7 were used and all MS/MS spectra were exported as mgf file. Peptide search was performed using Mascot search engine (version 2.6.2) against the Swiss-Prot human protein database (20,237 sequences and 11,451,954 residues). Search settings were 10 ppm of precursor tolerance, 0.02-Da fragment tolerance, one missed cleavage allowed, carbamidomethyl on cysteine as fixed modification, deamidation of glutamine and asparagine allowed as variable modification, and oxidation of methionine. Applying the percolator algorithm (*70*) resulted in a peptide FDR of 0.46%. Search results were reimported in the Progenesis QI software. Proteins were quantified by summing up the abundances of all unique peptides per protein after normalization to identified glyceraldehyde-3-phosphate dehydrogenase and ACTB peptides. Resulting protein abundances were used for calculation of fold changes between conditions and repeated-measures analyses of variance (ANOVAs) within the Progenesis QI software. Proteomics expression data are provided as table S4.

Smurf2 siRNA-mediated silencing

phLFs were reverse transfected with 2 or 10 nM Silencer Predesigned Smurf2 siRNA (Ambion, catalog no. AM16708, Thermo Fisher Scientific, Carlsbad, USA) or 10 nM scrambled Silencer Negative control No. 1 siRNA (Ambion, AM4611, Thermo Fisher Scientific, Carlsbad, USA) in Lipofectamine RNAiMax transfection reagent (Thermo Fisher Scientific, Carlsbad, USA, 13778-150) as indicated followed by TGFβ1 treatment (1 ng/ml) for 48 hours if not indicated differently.

Statistical analysis

Analyses of interferential statistics were performed in GraphPad Prism v5 and v7. Unless otherwise indicated results indicate means ± SEM of minimum three biological replicates. Student's t test was used for formally distributed samples. $P < 0.05$ was considered as significant. One-way ANOVA with Bonferroni correction for multiple testing was used to compare multiple experimental groups or conditions. Data are represented as arithmetic means ± SEM. For quantitative SAR studies, dose-response curve modeling and IC_{50} determination GraphPad Prism v7 was used. MFI measurement data series from live imaging and learning curve data series for the deep learning models were smoothened as described in code section S5 with a degree of smoothening of 95 and 99%, respectively.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.abb3673

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. T. A. Wynn, T. R. Ramalingam, Mechanisms of fibrosis: Therapeutic translation for fibrotic disease. *Nat. Med.* **18**, 1028–1040 (2012).
2. T. R. Cox, J. T. Erler, Remodeling and homeostasis of the extracellular matrix: Implications for fibrotic diseases and cancer. *Dis. Model. Mech.* **4**, 165–178 (2011).
3. D. C. Rockey, P. D. Bell, J. A. Hill, Fibrosis—A common pathway to organ injury and failure. *N. Engl. J. Med.* **373**, 95–96 (2015).
4. S. L. Friedman, D. Sheppard, J. S. Duffield, S. Violette, Therapy for fibrotic diseases: Nearing the starting line. *Sci. Transl. Med.* **5**, 167sr1 (2013).
5. L. Richeldi, H. R. Collard, M. G. Jones, Idiopathic pulmonary fibrosis. *Lancet* **389**, 1941–1952 (2017).
6. I. E. Fernandez, O. Eickelberg, New cellular and molecular mechanisms of lung injury and fibrosis in idiopathic pulmonary fibrosis. *Lancet* **380**, 680–688 (2012).
7. G. Sgalla, A. Biffi, L. Richeldi, Idiopathic pulmonary fibrosis: Diagnosis, epidemiology and natural history. *Respirology* **21**, 427–437 (2016).
8. T. E. King Jr., W. Z. Bradford, S. Castro-Bernardini, E. A. Fagan, I. Glaspole, M. K. Glassberg, E. Gorina, P. M. Hopkins, D. Kardatzke, L. Lancaster, D. J. Lederer, S. D. Nathan, C. A. Pereira, S. A. Sahn, R. Sussman, J. J. Swigris, P. W. Noble, A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. *N. Engl. J. Med.* **370**, 2083–2092 (2014).
9. L. Richeldi, R. M. du Bois, G. Raghu, A. Azuma, K. K. Brown, U. Costabel, V. Cottin, K. R. Flaherty, D. M. Hansell, Y. Inoue, D. S. Kim, M. Kolb, A. G. Nicholson, P. W. Noble, M. Selman, H. Taniguchi, M. Brun, F. le Maulf, M. Girard, S. Stowasser, R. Schlenker-Herceg, B. Disse, H. R. Collard, Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N. Engl. J. Med.* **370**, 2071–2082 (2014).
10. K. R. Flaherty, A. U. Wells, V. Cottin, A. Devaraj, S. L. F. Walsh, Y. Inoue, L. Richeldi, M. Kolb, K. Tetzlaff, S. Stowasser, C. Coeck, E. Clerisme-Beaty, B. Rosenstock, M. Quaresma, T. Haeufel, R. G. Goeldner, R. Schlenker-Herceg, K. K. Brown, Nintedanib in progressive fibrosing interstitial lung diseases. *N. Engl. J. Med.* **381**, 1718–1727 (2019).
11. G. Burgstaller, B. Oehrle, M. Gerckens, E. S. White, H. B. Schiller, O. Eickelberg, The instructive extracellular matrix of the lung: Basic composition and alterations in chronic lung disease. *Eur. Respir. J.* **50**, 1601805 (2017).
12. M. Selman, C. Lopez-Otin, A. Pardo, Age-driven developmental drift in the pathogenesis of idiopathic pulmonary fibrosis. *Eur. Respir. J.* **48**, 538–552 (2016).
13. H. B. Schiller, I. E. Fernandez, G. Burgstaller, C. Schaab, R. A. Scheltema, T. Schwarzmayr, T. M. Strom, O. Eickelberg, M. Mann, Time- and compartment-resolved proteome profiling of the extracellular niche in lung injury and repair. *Mol. Syst. Biol.* **11**, 819 (2015).
14. A. Naba, K. R. Clauser, S. Hoersch, H. Liu, S. A. Carr, R. O. Hynes, The matrisome: In silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices. *Mol. Cell Proteomics* **11**, M111.014647 (2012).

15. M. L. Decaris, M. Gatmaitan, S. FlorCruz, F. Luo, K. Li, W. E. Holmes, M. K. Hellerstein, S. M. Turner, C. L. Emson, Proteomic analysis of altered extracellular matrix turnover in bleomycin-induced pulmonary fibrosis. *Mol. Cell. Proteomics* **13**, 1741–1752 (2014).
16. A. J. Booth, R. Hadley, A. M. Cornett, A. A. Dreffs, S. A. Matthes, J. L. Tsui, K. Weiss, J. C. Horowitz, V. F. Fiore, T. H. Barker, B. B. Moore, F. J. Martinez, L. E. Niklason, E. S. White, Acellular normal and fibrotic human lung matrices as a culture system for in vitro investigation. *Am. J. Respir. Crit. Care Med.* **186**, 866–876 (2012).
17. M. Sato, Y. Muragaki, S. Saika, A. B. Roberts, A. Ooshima, Targeted disruption of TGF-β1/Smad3 signaling protects against renal tubulointerstitial fibrosis induced by unilateral ureteral obstruction. *J. Clin. Invest.* **112**, 1486–1494 (2003).
18. W. A. Border, N. A. Noble, T. Yamamoto, J. R. Harper, Y. Yamaguchi, M. D. Pierschbacher, E. Ruoslahti, Natural inhibitor of transforming growth factor-β protects against scarring in experimental kidney disease. *Nature* **360**, 361–364 (1992).
19. D. E. Clouthier, S. A. Comerford, R. E. Hammer, Hepatic fibrosis, glomerulosclerosis, and a lipodystrophy-like syndrome in PEPCK-TGF-beta1 transgenic mice. *J. Clin. Invest.* **100**, 2697–2713 (1997).
20. P. J. Sime, Z. Xing, F. L. Graham, K. G. Csaky, J. Gauldie, Adenovector-mediated gene transfer of active transforming growth factor-beta1 induces prolonged severe fibrosis in rat lung. *J. Clin. Invest.* **100**, 768–776 (1997).
21. U. Bartram, C. P. Speer, The role of transforming growth factor β in lung development and disease. *Chest* **125**, 754–765 (2004).
22. E. Piek, C. H. Heldin, P. Ten Dijke, Specificity, diversity, and regulation in TGF-β superfamily signaling. *FASEB J.* **13**, 2105–2124 (1999).
23. F. Verrecchia, A. Mauviel, Transforming growth factor-β signaling through the Smad pathway: Role in extracellular matrix gene expression and regulation. *J. Invest. Dermatol.* **118**, 211–215 (2002).
24. K. L. Walton, K. E. Johnson, C. A. Harrison, Targeting TGF-β mediated SMAD signaling for the prevention of fibrosis. *Front. Pharmacol.* **8**, 461 (2017).
25. P. Pakshir, B. Hinz, The big five in fibrosis: Macrophages, myofibroblasts, matrix, mechanics, and miscommunication. *Matrix Biol.* **68-69**, 81–93 (2018).
26. R. Rabieian, M. Boshtam, M. Zareei, S. Kouhpayeh, A. Masoudifar, H. Mirzaei, Plasminogen activator inhibitor type-1 as a regulator of fibrosis. *J. Cell. Biochem.* **119**, 17–27 (2018).
27. L. Yang, J. Herrera, A. Gilbertsen, H. Xia, K. Smith, A. Benyumov, P. B. Bitterman, C. A. Henke, IL-8 mediates idiopathic pulmonary fibrosis mesenchymal progenitor cell fibrogenicity. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **314**, L127–L136 (2018).
28. G. Raghu, K. K. Brown, H. R. Collard, V. Cottin, K. F. Gibson, R. J. Kaner, D. J. Lederer, F. J. Martinez, P. W. Noble, J. W. Song, A. U. Wells, T. P. M. Whelan, W. Wuyts, E. Moreau, S. D. Patterson, V. Smith, S. Bayly, J. W. Chien, Q. Gong, J. J. Zhang, T. G. O'Riordan, Efficacy of simtuzumab versus placebo in patients with idiopathic pulmonary fibrosis: A randomised, double-blind, controlled, phase 2 trial. *Lancet Respir. Med.* **5**, 22–32 (2017).
29. J. M. Parker, I. N. Glaspole, L. H. Lancaster, T. J. Haddad, D. She, S. L. Roseti, J. P. Fiening, E. P. Grant, C. M. Kell, K. R. Flaherty, A phase 2 randomized controlled study of tralokinumab in subjects with idiopathic pulmonary fibrosis. *Am. J. Respir. Crit. Care Med.* **197**, 94–103 (2018).
30. T. E. King Jr., K. K. Brown, G. Raghu, R. M. du Bois, D. A. Lynch, F. Martinez, D. Valeyre, I. Leconte, A. Morganti, S. Roux, J. Behr, BUILD-3: A randomized, controlled trial of bosentan in idiopathic pulmonary fibrosis. *Am. J. Respir. Crit. Care Med.* **184**, 92–99 (2011).
31. I. Noth, K. J. Anstrom, S. B. Calvert, J. de Andrade, K. R. Flaherty, C. Glazer, R. J. Kaner, M. A. Olman; Idiopathic Pulmonary Fibrosis Clinical Research Network (IPFnet), A placebo-controlled randomized trial of warfarin in idiopathic pulmonary fibrosis. *Am. J. Respir. Crit. Care Med.* **186**, 88–95 (2012).
32. D. Haasen, U. Schopfer, C. Antczak, C. Guy, F. Fuchs, P. Selzer, How phenotypic screening influenced drug discovery: Lessons from five years of practice. *Assay Drug Dev. Technol.* **15**, 239–246 (2017).
33. F. Vincent, P. Loria, M. Pregel, R. Stanton, L. Kitching, K. Nocka, R. Doyonnas, C. Steppan, A. Gilbert, T. Schroeter, M.-C. Peakman, Developing predictive assays: The phenotypic screening "rule of 3". *Sci. Transl. Med.* **7**, 293ps15 (2015).
34. B. Ng, J. Dong, G. D'Agostino, S. Viswanathan, A. A. Widjaja, W. W. Lim, N. S. J. Ko, J. Tan, S. P. Chothani, B. Huang, C. Xie, C. J. Pua, A. M. Chacko, N. Guimarães-Camboa, S. M. Evans, A. J. Byrne, T. M. Maher, J. Liang, D. Jiang, P. W. Noble, S. Schafer, S. A. Cook, Interleukin-11 is a therapeutic target in idiopathic pulmonary fibrosis. *Sci. Transl. Med.* **11**, (2019).
35. K. Heinzelmann, N. Noskovičová, J. Merl-Pham, G. Preissler, H. Winter, M. Lindner, R. Hatz, S. M. Hauck, J. Behr, O. Eickelberg, Surface proteome analysis identifies platelet derived growth factor receptor-alpha as a critical mediator of transforming growth factor-beta-induced collagen secretion. *Int. J. Biochem. Cell Biol.* **74**, 44–59 (2016).
36. A. Naba, S. Hoersch, R. O. Hynes, Towards definition of an ECM parts list: An advance on GO categories. *Matrix Biol.* **31**, 371–372 (2012).
37. T. S. Adams, J. C. Schupp, S. Poli, E. A. Ayaub, N. Neumark, F. Ahangari, S. G. Chu, B. A. Raby, G. De Iuliis, M. Januszyk, Q. Duan, H. A. Arnett, A. Siddiqui, G. R. Washko, R. Homer, X. Yan, I. O. Rosas, N. Kaminski, Single cell RNA-seq reveals ectopic and aberrant lung resident cell populations in idiopathic pulmonary fibrosis. *Sci. Adv.* **8**, 759902 (2019).
38. J. D. Vasta, K. A. Andersen, K. M. Deck, C. P. Nizzi, R. S. Eisenstein, R. T. Raines, Selective inhibition of collagen prolyl 4-hydroxylase in human cells. *ACS Chem. Biol.* **11**, 193–199 (2016).
39. K. Majamaa, T. Sasaki, J. Uitto, Inhibition of prolyl hydroxylation during collagen biosynthesis in human skin fibroblast cultures by ethyl 3,4-dihydroxybenzoate. *J. Invest. Dermatol.* **89**, 405–409 (1987).
40. H. N. Alsafadi, C. A. Staab-Weijnitz, M. Lehmann, M. Lindner, B. Peschel, M. Königshoff, D. E. Wagner, An ex vivo model to induce early fibrosis-like changes in human precision-cut lung slices. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **312**, L896–L902 (2017).
41. S. Darakhshan, A. B. Pour, Tranilast: A review of its therapeutic applications. *Pharmacol. Res.* **91**, 15–28 (2015).
42. M. Isaji, M. Nakajoh, J. Naito, Selective inhibition of collagen accumulation by *N*-(3,4-dimethoxycinnamoyl)anthranilic acid (N-5') in granulation tissue. *Biochem. Pharmacol.* **36**, 469–474 (1987).
43. D. Szklarczyk, A. L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N. T. Doncheva, J. H. Morris, P. Bork, L. J. Jensen, C. von Mering, STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).
44. N. Noskovicova, K. Heinzelmann, G. Burgstaller, J. Behr, O. Eickelberg, Cub domain-containing protein 1 negatively regulates TGF-β signaling and myofibroblast differentiation. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **314**, L695–L707 (2018).
45. X. Lin, M. Liang, X. H. Feng, Smurf2 is a ubiquitin E3 ligase mediating proteasome-dependent degradation of Smad2 in transforming growth factor-β signaling. *J. Biol. Chem.* **275**, 36818–36822 (2000).
46. M. Gerckens, H. N. Alsafadi, D. E. Wagner, M. Lindner, G. Burgstaller, M. Königshoff, Generation of human 3D lung tissue cultures (3D-LTCs) for disease modeling. *J. Vis. Exp.* (2019).
47. A. L. Degryse, H. Tanjore, X. C. Xu, V. V. Polosukhin, B. R. Jones, F. B. McMahon, L. A. Gleaves, T. S. Blackwell, W. E. Lawson, Repetitive intratracheal bleomycin models several features of idiopathic pulmonary fibrosis. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **299**, L442–L452 (2010).
48. A. L. Degryse, W. E. Lawson, Progress toward improving animal models for idiopathic pulmonary fibrosis. *Am. J. Med. Sci.* **341**, 444–449 (2011).
49. L. Richeldi, Treatments for idiopathic pulmonary fibrosis. *N. Engl. J. Med.* **371**, 781–784 (2014).
50. P. Spagnolo, T. M. Maher, L. Richeldi, Idiopathic pulmonary fibrosis: Recent advances on pharmacological therapy. *Pharmacol. Ther.* **152**, 18–27 (2015).
51. K. Heinzelmann, M. Lehmann, M. Gerckens, N. Noskovičová, M. Frankenberger, M. Lindner, R. Hatz, J. Behr, A. Hilgendorff, M. Königshoff, O. Eickelberg, Cell-surface phenotyping identifies CD36 and CD97 as novel markers of fibroblast quiescence in lung fibrosis. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **315**, L682–L696 (2018).
52. M. C. Emblom-Callahan, M. K. Chhina, O. A. Shlobin, S. Ahmad, E. S. Reese, E. P. R. Iyer, D. N. Cox, R. Brenner, N. A. Burton, G. M. Grant, S. D. Nathan, Genomic phenotype of non-cultured pulmonary fibroblasts in idiopathic pulmonary fibrosis. *Genomics* **96**, 134–145 (2010).
53. L. R. Rodriguez, M. Emblom-Callahan, M. Chhina, S. Bui, B. Aljeburry, L. H. Tran, R. Novak, M. Lemma, S. D. Nathan, G. M. Grant, Global gene expression analysis in an in vitro fibroblast model of idiopathic pulmonary fibrosis reveals potential role for CXCL14/CXCR4. *Sci. Rep.* **8**, 3983 (2018).
54. H. Chen, O. Engkvist, Y. Wang, M. Olivecrona, T. Blaschke, The rise of deep learning in drug discovery. *Drug Discov. Today* **23**, 1241–1250 (2018).
55. C. L. Chen, A. Mahjoubfar, L. C. Tai, I. K. Blaby, A. Huang, K. R. Niazi, B. Jalali, Deep learning in label-free cell classification. *Sci. Rep.* **6**, 21471 (2016).
56. A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, S. Thrun, Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115–118 (2017).
57. M. Rogosnitzky, R. Danks, E. Kardash, Therapeutic potential of tranilast, an anti-allergy drug, in proliferative disorders. *Anticancer Res* **32**, 2471–2478 (2012).
58. S. C. Zammit, A. J. Cox, R. M. Gow, Y. Zhang, R. E. Gilbert, H. Krum, D. J. Kelly, S. J. Williams, Evaluation and optimization of antifibrotic activity of cinnamoyl anthranilates. *Bioorg. Med. Chem. Lett.* **19**, 7003–7006 (2009).
59. J. W. Song, S. B. Hong, C. M. Lim, Y. Koh, D. S. Kim, Acute exacerbation of idiopathic pulmonary fibrosis: Incidence, risk factors and outcome. *Eur. Respir. J.* **37**, 356–363 (2011).
60. M. M. Juarez, A. L. Chan, A. G. Norris, B. M. Morrissey, T. E. Albertson, Acute exacerbation of idiopathic pulmonary fibrosis-a review of current and novel pharmacotherapies. *J. Thorac. Dis.* **7**, 499–519 (2015).
61. C. A. Staab-Weijnitz, I. E. Fernandez, L. Knüppel, J. Maul, K. Heinzelmann, B. M. Juan-Guardela, E. Hennen, G. Preissler, H. Winter, C. Neurohr, R. Hatz, M. Lindner, J. Behr, N. Kaminski, O. Eickelberg, FK506-binding protein 10, a potential novel drug target for idiopathic pulmonary fibrosis. *Am. J. Respir. Crit. Care Med.* **192**, 455–467 (2015).
62. G. Burgstaller, B. Oehrle, I. Koch, M. Lindner, O. Eickelberg, Multiplex profiling of cellular invasion in 3D cell culture models. *PLoS One* **8**, e63121 (2013).

63. G. Burgstaller, A. Sengupta, S. Vierkotten, G. Preissler, M. Lindner, J. Behr, M. Königshoff, O. Eickelberg, Distinct niches within the extracellular matrix dictate fibroblast function in (cell free) 3D lung tissue cultures. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **314**, L708–L723 (2018).
64. L. McInnes, J. Healy, J. Melville, UMAP: Uniform manifold approximation and projection for dimension reduction. arXiv:1802.03426 [stat.ML] (9 February 2018).
65. C. Scheeder, F. Heigwer, M. Boutros, Machine learning and image-based profiling in drug discovery. *Curr. Opin. Syst. Biol.* **10**, 43–52 (2018).
66. J. R. Wisniewski, A. Zougman, N. Nagaraj, M. Mann, Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359–362 (2009).
67. A. Grosche, A. Hauser, M. F. Lepper, R. Mayo, C. von Toerne, J. Merl-Pham, S. M. Hauck, The proteome of native adult Müller glial cells from murine retina. *Mol. Cell. Proteomics* **15**, 462–480 (2016).
68. S. M. Hauck, J. Dietter, R. L. Kramer, F. Hofmaier, J. K. Zipplies, B. Amann, A. Feuchtinger, C. A. Deeg, M. Ueffing, Deciphering membrane-associated molecular processes in target tissue of autoimmune uveitis by label-free quantitative mass spectrometry. *Mol. Cell. Proteomics* **9**, 2292–2305 (2010).
69. J. Merl, M. Ueffing, S. M. Hauck, C. von Toerne, Direct comparison of MS-based label-free and SILAC quantitative proteome profiling strategies in primary retinal Müller cells. *Proteomics* **12**, 1902–1911 (2012).
70. M. Brosch, L. Yu, T. Hubbard, J. Choudhary, Accurate and sensitive peptide identification with Mascot Percolator. *J. Proteome Res.* **8**, 3176–3181 (2009).

Acknowledgments: We acknowledge the provision of human biomaterial and clinical data from the CPC-M bioArchive and its partners at the Asklepios Biobank Gauting, the Klinikum der Universität München, and the Ludwig-Maximilians-Universität München. We are also grateful to M. Neumann, C. Hollauer, and D. Kutschke from the Comprehensive Pneumology Center (CPC), and K. Richter (Institute of Experimental Genetics) for providing superb technical support. We thank the CPC's research school for providing financial support for the project and a fellowship to M.Ge. We are grateful for the support of the HMGU Innovation and Management team. We thank T. Conlon for editing the manuscript's text. **Funding:** This work was supported by the German Center of Lung Research (DZL), the Helmholtz Association, the LMU Munich by Lehre@LMU Förderung to M.G., and the Helmholtz Alliance "Aging and Metabolic Programming, AMPro" (to J.Bec.). H.M. acknowledges a CSC graduate fellowship. M.Gr. reports the following funding grants: MH CR (grant 17-31538A) and MEYS CR (grant LQ1604). M.J. acknowledges GACR funding (grant 18-02699S). **Author contributions:** G.B., A.Ö.Y., and O.E. conceived and initiated the study. G.B. designed the study. M.G., A.-M.D., K.H., A.S., M.W., and M.G.S. designed and performed experiments. M.J., L.S., and M.Gr. performed mouse in vivo experiments. W.F. and E.T. formulated compounds for mouse in vivo experiments. J.B., M.G.S., A.H., and M.L. performed surgical work and provided human tissue. M.G. and F.P. set up all methods for AI. K.S., M.G., and K.Had. performed the screening of the small-molecule library. M.I. and J.Bec. performed the microarray experiment and analysis. J.M.-P. and S.M.H. did the mass spectrometry experiments and analyzed the mass spec data. M.G., A.S., M.G.S., and H.N.A. created human lung slices and performed experiments on PCLS. K.R., H.M., and O.P. performed medicinal chemistry. G.B., M.G., A.Ö.Y., M.K., O.E., J.B., O.P., and A.H. analyzed and interpreted results. M.G. and G.B. prepared the figures. G.B. drafted and wrote the manuscript. G.B., M.G., M.K., O.E., O.P., A.H., M.I., and A.Ö.Y. edited and revised the manuscript. All authors read the manuscript and discussed the interpretation of the results. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data associated with this study are present in the paper or the Supplementary Materials. Microarray data have been deposited in the Gene Expression Omnibus (accession number GSE141905).

Submitted 19 February 2020
Accepted 8 November 2021
Published 22 December 2021
10.1126/sciadv.abb3673

Science Advances

Phenotypic drug screening in a human fibrosis model identified a novel class of antifibrotic therapeutics

Michael Gerckens, Kenji Schorpp, Francesco Pelizza, Melanie Wögrath, Kora Reichau, Huilong Ma, Armando-Marco Dworsky, Arunima Sengupta, Mircea Gabriel Stoleriu, Katharina Heinzelmann, Juliane Merl-Pham, Martin Irmler, Hani N. Alsafadi, Eduard Trenkenschuh, Lenka Sarnova, Marketa Jirouskova, Wolfgang Frieß, Stefanie M. Hauck, Johannes Beckers, Nikolaus Kneidinger, Jürgen Behr, Anne Hilgendorff, Kamyar Hadian, Michael Lindner, Melanie Königshoff, Oliver Eickelberg, Martin Gregor, Oliver Plettenburg, Ali Önder Yildirim, and Gerald Burgstaller

Sci. Adv. **7** (52), eabb3673. DOI: 10.1126/sciadv.abb3673

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.