nature biotechnology

Article

https://doi.org/10.1038/s41587-024-02143-0

A small-molecule TNIK inhibitor targets fibrosis in preclinical and clinical models

Received: 26 June 2023

Accepted: 16 January 2024

Published online: 08 March 2024

Check for updates

Feng Ren[1,2], Alex Aliper[2,3], Jian Chen[4], Heng Zhao[1], Sujata Rao[5], Christoph Kuppe[6,7], Ivan V. Ozerov[3], Man Zhang[1], Klaus Witte[3], Chris Kruse[3], Vladimir Aladinskiy[2], Yan Ivanenkov[3], Daniil Polykovskiy[8], Yanyun Fu[1], Eugene Babin[2], Junwen Qiao[1], Xing Liang[1], Zhenzhen Mou[1], Hui Wang[1], Frank W. Pun[3], Pedro Torres Ayuso[9], Alexander Veviorskiy[2], Dandan Song[4], Sang Liu[1], Bei Zhang[1], Vladimir Naumov[3], Xiaoqiang Ding[10], Andrey Kukharenko[3], Evgeny Izumchenko[11] & Alex Zhavoronkov[2,3,5,8]

Idiopathic pulmonary fibrosis (IPF) is an aggressive interstitial lung disease with a high mortality rate. Putative drug targets in IPF have failed to translate into effective therapies at the clinical level. We identify TRAF2- and NCK-interacting kinase (TNIK) as an anti-fibrotic target using a predictive artificial intelligence (AI) approach. Using AI-driven methodology, we generated INS018_055, a small-molecule TNIK inhibitor, which exhibits desirable drug-like properties and anti-fibrotic activity across different organs in vivo through oral, inhaled or topical administration. INS018_055 possesses anti-inflammatory effects in addition to its anti-fibrotic profile, validated in multiple in vivo studies. Its safety and tolerability as well as pharmacokinetics were validated in a randomized, double-blinded, placebo-controlled phase I clinical trial (NCT05154240) involving 78 healthy participants. A separate phase I trial in China, CTR20221542, also demonstrated comparable safety and pharmacokinetic profiles. This work was completed in roughly 18 months from target discovery to preclinical candidate nomination and demonstrates the capabilities of our generative AI-driven drug-discovery pipeline.

Therapeutic target identification is a crucial step of the drug-discovery pipeline for all disease processes, including fibrosis. Erroneous targets selected at the early stage of drug development may result in a costly drug-discovery program, often ending in failure during phase II trials years later[1,2]. As a result, large pharmaceutical companies may behave as though they are risk averse, reducing their willingness to invest in potentially valuable targets and therapeutic strategies[3]. Although development of precise data-driven approaches for drug target discovery has a critical impact on the success of clinical trials, this task still has recognized limitations such as data complexity and batch effects[4,5]. More recently, AI-driven approaches have demonstrated efficacy in discovering target candidates[6] in the contexts of embryonic–fetal transition[7] and muscle aging[8]. Advanced pathway analysis and AI algorithms applied to multiomics data can identify targets

[1]Insilico Medicine Shanghai Ltd., Shanghai, China. [2]Insilico Medicine AI Limited, Abu Dhabi, UAE. [3]Insilico Medicine Hong Kong Ltd., Hong Kong Science and Technology Park, Hong Kong SAR, China. [4]Department of Clinical Pharmacology, Affiliated Xiaoshan Hospital, Hangzhou Normal University, Hangzhou, China. [5]Insilico Medicine US Inc., New York, NY, USA. [6]Institute of Experimental Medicine and Systems Biology, RWTH Aachen University, Aachen, Germany. [7]Department of Nephrology, University Clinic RWTH Aachen, Aachen, Germany. [8]Insilico Medicine Canada Inc, Montreal, Quebec, Canada. [9]Department of Cancer and Cellular Biology, Lewis Katz School of Medicine, Temple University, PA, USA. [10]Division of Nephrology, Zhongshan Hospital Shanghai Medical College, Fudan University, Shanghai, China. [11]Section of Hematology and Oncology, Department of Medicine, University of Chicago, Chicago, IL, USA. e-mail: alex@insilico.com

and biomarkers even when prior evidence is sparse[9,10]. The success of multiomic target-discovery systems has been demonstrated in cancer and age-associated diseases[11–15]. Fibrosis occurs in the final stage of chronic organ deficiencies, such as pulmonary, kidney or liver diseases and is characterized by excessive proliferation of matrix-producing cells that arise from dysregulated chronic inflammation triggered by infective, chemical, autoimmune or radioactive tissue damage[16,17]. IPF is a gradually advancing and generally lethal disease, characterized histologically by fibroblast proliferation and substantial extracellular matrix deposition[18,19]. Myofibroblasts play a crucial role in the fibrotic process, with transforming growth factor (TGF)-β being a major contributor to myofibroblast differentiation[20]. IPF is most prevalent in patients over 60 years of age[21], and its estimated incidence in the USA was reported to be 6.8 per 100,000 person-years[21]. Given the increasing incidence, the cost of treating patients with IPF poses a substantial burden on healthcare systems and represents a growing public health problem worldwide[21–25].

In untreated patients, IPF has a highly inconsistent clinical course, with a median survival of 2 to 3 years[26]. Less than 30% of patients with IPF benefit from administration of corticosteroids, with adverse effects reducing patients' quality of life[27]. Nevertheless, most of these patients succumb to respiratory failure or progressive decline in lung function[28]. Current targeted treatment options for IPF are limited to nintedanib and pirfenidone[29]. Following their FDA approval in 2014, nintedanib and pirfenidone are recommended by the American Thoracic Society, the European Respiratory Society, the Japanese Respiratory Society and the Latin American Thoracic Association for treating IPF[30]. Of these drugs, pirfenidone suppresses TGF-β expression[31], whereas nintedanib exerts its efficacy in IPF by inhibiting fibroblast growth factor receptor (FGFR), platelet-derived growth factor receptor (PDGFR) and vascular endothelial growth factor (VEGFR)[32], that is, growth factor receptors in the fibrotic pathway of tyrosine kinase structure, with FGFR being most relevant[33].

As with IPF, renal fibrosis also correlates with organ failure[34]. Renal fibrosis is histomorphologically identified by tubulointerstitial fibrosis and is the final manifestation of chronic kidney disease (CKD)[35]. The incidence of CKD is increasing worldwide and continues to grow within aged populations[36]. Thus, developing an effective anti-fibrotic medicine that attenuates renal fibrosis remains a large, unmet clinical need. Like IPF, renal fibrosis develops following prolonged exposure to ischemic conditions or inflammatory insults[37]. Increased TGF-β signaling within stromal cell types in the tubule–interstitial space appears to be central to the initiation of fibrosis[38,39]. No specific fibrosis inhibitor has passed clinical trials for the treatment of CKD.

Given the scarcity of anti-fibrotic therapies for IPF or CKD, we used our published generative AI pipeline[14,15,40–42] to address this clinical need. The process of drug development, from target identification to preclinical testing and ultimately to clinical validation, can take over a decade. In this study, our generative AI platform has identified TNIK inhibition as a potent anti-fibrotic strategy and assisted in developing a highly specific TNIK inhibitor, INS018_055. We have synthesized this compound and demonstrated its selective, anti-fibrotic activity in multiple murine and rat models of fibrosis. We report phase I clinical trial data highlighting the safety and tolerability of our small-molecule inhibitor. This comprehensive approach was completed in roughly 18 months from target discovery to preclinical candidate nomination and demonstrates the capabilities of our generative AI-driven drug-discovery pipeline.

Results

AI-discovered anti-fibrotic targets

Here we used PandaOmics, a commercially available target-discovery platform using multiple AI engines including generative pretrained transformers[14], for target identification in fibrosis (Fig. 1a). We included a list of multiomics datasets derived from tissue samples of patients with IPF as a basis for the computational pipeline (Fig. 1a, step 1). This pipeline combines several distinct but complementary computational approaches, all of which are similar in terms of output (a ranked list of genes, ordered by most-promising target) but different in terms of the input data. In addition to multiomics datasets, we included methods relying on biological network analysis and text data from scientific literature (Fig. 1a, step 2) with node degree bias control. These data types facilitate target hypotheses inferred from omics into the context of prior evidence, which include clinical trials, publications and grant applications. The text data are temporal in nature, allowing PandaOmics to not only capture trends but even project them into the future (Fig. 1a, step 3).

Unlike other target-discovery approaches such as Open Targets[43] or NewDrugTargets[44], our method was designed and validated to search targets for multiple diseases and aging[15]. Thus, we developed a time machine approach, in which we trained the computational models using data published before a certain time point and validated the model outputs by their ability to predict those targets that came into the focus of the pharmaceutical industry after this time point. Several distinct AI-driven target-discovery philosophies, including random walk on heterogeneous graphs and negative matrix factorization, were applied (Supplementary Video 1).

Our platform generated a ranked list of targets after applying a combination of score composition and filters, including disease-agnostic properties like protein family, accessibility by small molecules or therapeutic antibodies, novelty and crystal structure availability, among others (Extended Data Fig. 1). To generate the target hypotheses for fibrosis, we applied the 'protein and receptor kinase' scenario to fibrosis-related datasets. It includes scores reflecting the network component of the disease mechanisms combined with the transcriptional factor enrichment-based causal inference (Supplementary Information 1). Protein class (protein and receptor kinases only), novelty and small-molecule druggability filters were enabled in the downstream processing steps (Fig. 1a, step 4).

TNIK was identified as the number 1 target among the five top candidates using the 'protein and receptor kinase' approach with relatively high values of network neighbors, causal inference, pathways, interactome community, expression, heterogeneous graph walk and matrix factorization scores (Fig. 1a, step 5, Fig. 1b). TNIK has been associated with fibrosis-driving pathways including WNT[45,46], TGF-β[46,47], Hippo (yes-associated protein (YAP)–transcriptional coactivator with PDZ-binding motif (TAZ))[46,48], JNK[46,49,50] and nuclear factor (NF)-κB[46,50] signaling. However, TNIK has not been studied as a therapeutic target in IPF and was thus selected by the AI algorithm. Whereas tyrosine kinase inhibitors like nintedanib, imatinib and nilotinib have been tested in IPF, the role of serine–threonine kinases (such as TNIK) in IPF remains largely uninterrogated (Fig. 1c). As fibrosis can affect any organ and is responsible for up to 45% of all deaths in the industrialized world[51], we hypothesized that inhibition of TNIK may provide therapeutic benefit in multiple pathologies in which fibrosis may be implicated. In agreement with this hypothesis, TNIK was independently predicted as a disease target associated with multiple hallmarks of aging[15] and as a regulator of lipid metabolism in aged and/or high-fat diet-fed mice[52]. These findings support the hypotheses of our AI-driven pipeline (Supplementary Information 2).

We next performed a transparency analysis of the PandaOmics scores. The interactome community transparency revealed that TNIK inhibition is connected to multiple biological processes known to be important for fibrosis progression such as focal adhesion signaling, myofibroblast differentiation and mesenchymal cell migration (Extended Data Fig. 2a). Causal inference transparency revealed that TNIK is tightly connected with genes associated with IPF, including *TGFB1*, *FGR*, *FLT1*, *KDR* and others (Extended Data Fig. 2b). Finally, using an AI-powered de novo pathway-reconstruction tool, we showed that TNIK activates previously described pathways that are associated

https://doi.org/10.1038/s41587-024-02143-0

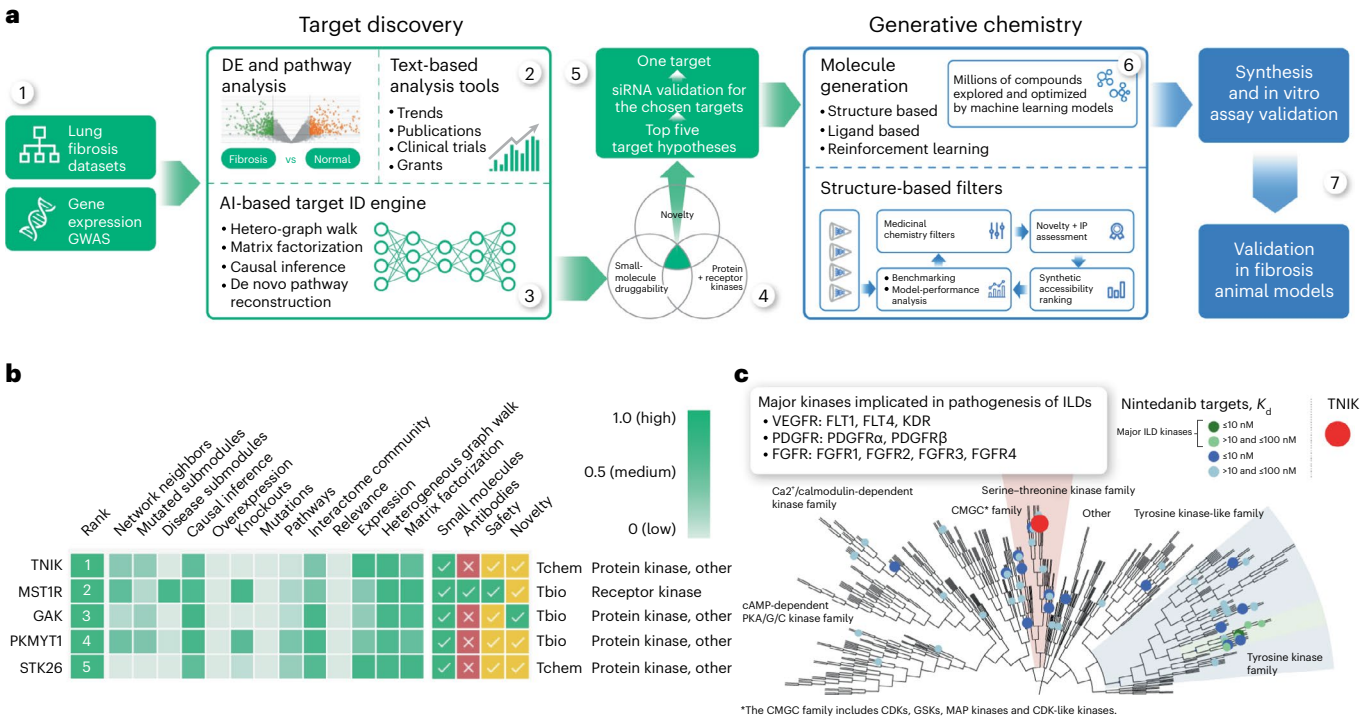


Fig. 1 | AI-augmented pipeline for target discovery. a, The PandaOmics target-discovery platform was applied to lung and kidney fibrosis datasets to generate target hypotheses, followed by the Chemistry42 platform application to generate small-molecule leads targeting TNIK. DE, differential gene; GWAS, genome-wide association study; hetero, heterogeneous; siRNA, small interfering RNA; IP, intellectual property. **b**, TNIK was scored the number 1 candidate using protein and receptor kinase PandaOmics settings based on relatively high values of network neighbors, mutated submodules, causal inference, pathways, interactome community, expression, heterogeneous graph walk and matrix factorization scores. GAK, cyclin G-associated kinase; MST1R, macrophage-stimulating 1 receptor; PKMYT1, protein kinase, membrane-associated tyrosine–threonine 1; STK26, serine–threonine kinase 26; Tchem, genes whose products can be targeted with small molecules better than the following bioactivity cutoff values: 30 nM for kinases, 100 nM for GPCRs and nuclear receptors, 10 μM for ion channels, and 1 μM for other target classes; Tbio, genes annotated with a Gene Ontology Molecular Function or Biological Process with an Experimental Evidence code, or targets with confirmed OMIM phenotype(s), or do not satisfy the Tdark criteria. **c**, TNIK is a member of the serine–threonine kinase STE20 family. This family does not contain any major targets of anti-fibrotic medications including nintedanib, the most prominent kinase inhibitor used for IPF treatment. This illustrates the relative novelty of the target. cAMP, cyclic AMP; FLT, FMS-related receptor tyrosine kinase 1; GSK, glycogen synthase kinase; KDR, kinase insert domain receptor; MAP, mitogen-activated protein; PKA, protein kinase A; ILD, interstial lung disease; K_d, dissociation constant.

with IPF. These gene sets are regulated by downstream transcriptional factors including TCF–LEF, SMAD, NF-κB and TEAD families (Extended Data Fig. 2c).

Validation of TNIK association with IPF using the single-cell gene expression dataset from unaffected lung and patients with IPF[53] demonstrates that *TNIK* expression in cytotoxic T cells, myofibroblasts and club cells is higher in fibrotic tissue than in unaffected controls (Extended Data Fig. 2d–f). This finding confirms TNIK's potential role in regulating the function of key cells involved in IPF development. We used this single-cell dataset to simulate *TNIK* knockout in IPF myofibroblasts using the scTenifoldKnk approach[54] followed by the Molecular Complex Detection (MCODE) algorithm[55] and performed pathway and biological process enrichment analysis. This analysis revealed that inhibiting TNIK primarily activates Hippo signaling and, consequently, downregulates YAP–TAZ (Extended Data Fig. 2g). Thus, TNIK is an attractive target for lung fibrosis, supported by both the unique omic-driven analysis and other AI-based approaches.

Generative AI-designed TNIK inhibitors

To identify TNIK inhibitors, we exploited available crystal structures of the TNIK kinase domain[56] and the Chemistry42 structure-based drug-design AI workflow (Fig. 1a, step 6)[40]. As using ATP competitors is a well-established strategy for targeting kinases[57], we selected the ATP-binding site as a pocket for compound generation. The AI-driven platform was configured to produce small-molecule structures capable of forming hydrogen bonds with the Cys108-NH of the TNIK hinge region. Targeting less-conserved adjacent allosteric pockets (such as a hydrophobic back cavity close to the gatekeeper residue) in addition to the active site can achieve better selectivity of the lead compounds[58]. As such, an additional hydrophobic pharmacophore point was applied to prioritize structures bearing hydrophobic functions to deeply occupy the back cavity formed by Met105, Leu73, Leu103, Ala52 and Val104.

The compounds, selected based on synthetic accessibility, novelty and medicinal chemistry properties, were synthesized and tested using a radiometric enzymatic assay (Fig. 1a, step 7). The first rounds of screening revealed a series of TNIK inhibitors with nanomolar-level binding affinity. However, in vitro absorption, distribution, metabolism and excretion profiling of the primary lead compounds revealed their high clearance in human and mice liver microsomes, cytochrome p450(CYP) inhibition at a half maximal inhibitory concentration (IC_{50}) less than 10 μM and kinetic solubility less than 2 μM (Supplementary Information 3). The lead optimization stage, which prioritized an improvement in the absorption, distribution, metabolism and excretion profile of this class of compounds, resulted in INS018_055 (WO2022179528A1) (Supplementary Information 3 and 4).

The carboxyl oxygen of INS018_055 forms a hydrogen bridge with Cys108-NH within the hinge region (Fig. 2a). Hydrogen bonding between the amide NH and the nitrogen in the proximal imidazole stabilizes a planar conformation. Moreover, the NH of this imidazole can form a hydrogen bond with the side chain of the gatekeeper Met105, as the distance of ~3.2 Å between the NH and the sulfur is suitable for bond formation. The second imidazole projects the *p*-fluorophenyl

https://doi.org/10.1038/s41587-024-02143-0

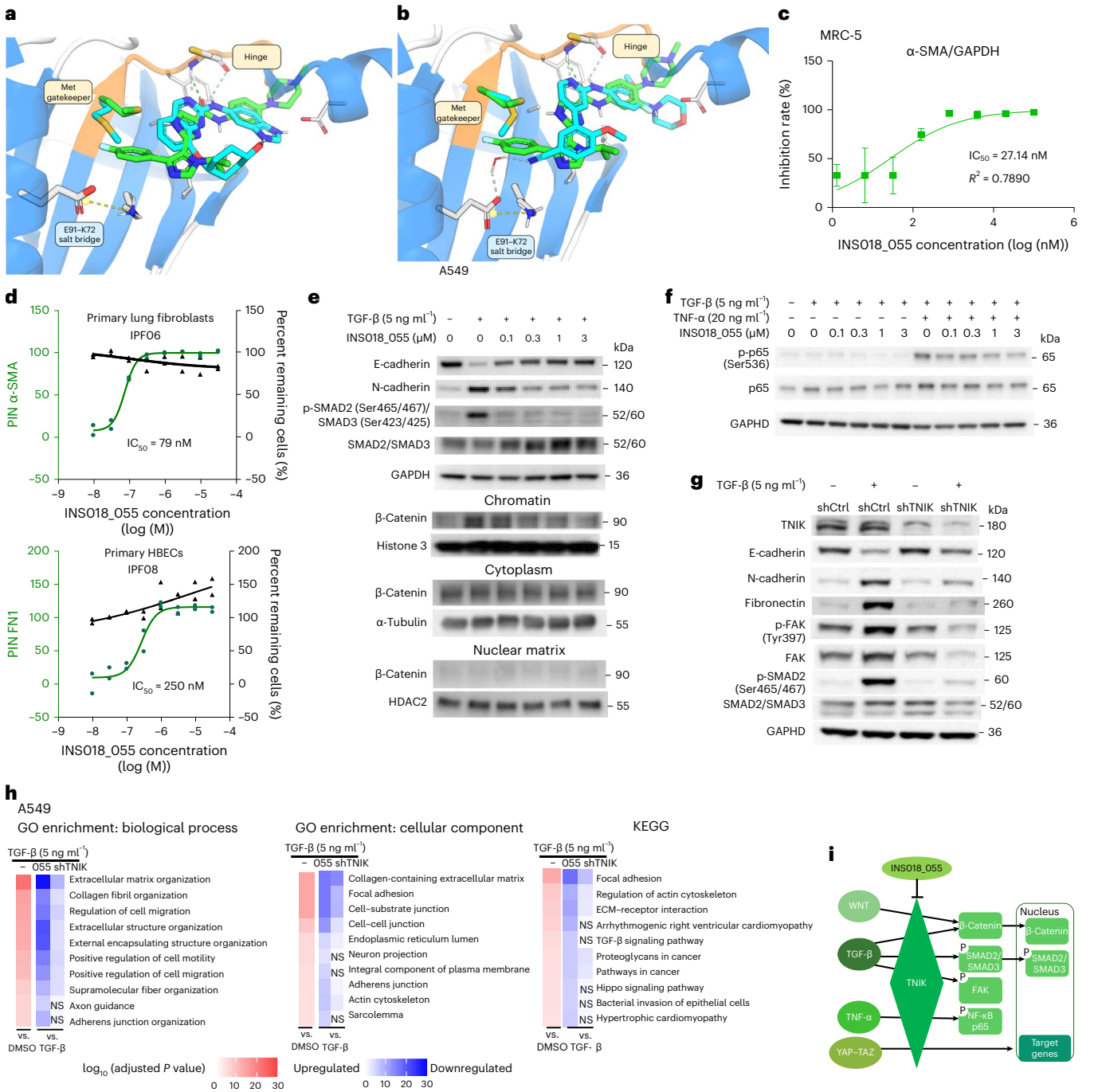


Fig. 2 | Architectural superposition of TNIK inhibitor structures with the predicted INS018_055-binding mode and effects on TGF-β-induced EMT and FMT cellular programs. a, Crystal structure of the NCB-0846 (cyan)-bound TNIK kinase domain (PDB 5D7A) aligned with the predicted binding mode of INS018_055 (green). **b**, Crystal structure of the compound 9 (cyan)-bound TNIK kinase domain (PDB 5AX9) aligned with the predicted binding mode of INS018_055 (green). Differences in Met gatekeeper orientations between inhibitors bound to the TNIK kinase domain are depicted, with Met side chain color corresponding to ligand color (cyan, green). The hinge region is shaded orange. Key ligand interactions are marked with dashed lines in the ligand color (cyan, green). The conservative E91–K72 salt bridge is shown as a yellow dashed line. Ligands and key pocket residues are represented as sticks. **c**, Inhibitory effect of INS018_055 on TGF-β-induced α-SMA protein expression in MRC-5 cells. $n = 3$, mean ± s.d. GAPDH, glyceraldehyde 3-phosphate dehydrogenase. **d**, Top, representative INS018_055-inhibitory effect (green) on FMT in primary human lung fibroblasts by measuring α-SMA. Bottom, INS018_055-inhibitory effect (green) on EMT in primary human bronchial epithelial cells (HBECs) by measuring fibronectin. Percent remaining cells (black) reflects nuclear count, which is a measurement of cell percentage without nuclear loss ($n = 2$ experimental replicates). PIN, percentage inhibition. **e**, Representative western blots showing changes in E-cadherin, N-cadherin, SMAD2/SMAD3, phospho (p)-SMAD2/SMAD3 and β-catenin in different A549 cell fractions following TGF-β stimulation and treatment with INS018_055. HDAC2, histone deacetylase 2. **f**, Representative western blot showing the inhibitory effect of INS018_055 on phosphorylated and total NF-κB p65 in TGF-β- and TNF-α-stimulated A549 cells ($n = 3$ biological replicates). **g**, Representative western blots showing changes in E-cadherin, N-cadherin, fibronectin, phospho-FAK and phospho-SMAD2 induced by treatment of A549 cells with TNIK shRNA (shTNIK) (shTNIK-4). shCtrl, control shRNA. **h**, Significantly upregulated pathways induced by TGF-β treatment and restoration by INS018_055 (055) or shTNIK-1. Gene ontology (GO) enrichment (left, middle) and KEGG analysis (right). $n = 3$ in each condition. Enrichment analysis was performed using the gseapy.enrichr Python package. *P* values were computed using Fisher's exact test (one-tailed hypergeometric test). Adjusted *P* values (*q* values) were calculated using the Benjamini–Hochberg method for correction for multiple-hypothesis testing. ECM, extracellular matrix; NS, not significant. **i**, Scheme showing TNIK function in regulating TGF-β, WNT, YAP–TAZ and TNF-α pathways as identified by in vitro perturbation experiments.

into the back pocket and the isopropyl toward Asn158 and Gln157. The *p*-fluorophenyl is not only properly accommodated within the back cavity but is also coplanar to the C-S-C of Met105 (the distance from the S to the closest C of the ring is 3.8 Å).

The X-ray structures of two early TNIK inhibitors, NCB-0846 and 4-methoxy-3-(2-(3-methoxy-4-morpholinophenylamino)pyridin-4-yl) benzonitrile (compound 9), in complex with the target were reported previously[56]. These compounds bind to the ATP-binding site and interact with the backbone of Cys108 at the hinge region via two hydrogen bonds (Fig. 2a,b). Crystal structures reveal that the side chain of the gatekeeper Met105 covers back pocket entry, such that these compounds do not contain functions capable of shifting this residue. Unlike NCB-0846, the cyano group of compound 9 interacts with the backbone of Phe172 through a water molecule close to the cavity and contacts Met105. By contrast, INS018_055 is predicted to occupy the back pocket deeply, forming a unique interaction among known TNIK inhibitors. To evaluate the binding affinity of INS018_055, we performed a surface plasmon resonance assay. INS018_055 showed potent affinity (K_d value of 4.32 nM) relative to two other known TNIK inhibitors (Extended Data Fig. 3a–c).

Selectivity of INS018_055

We next used a KinaseProfiler panel to assess the kinase selectivity of 10 µM INS018_055. The activity of protein kinases (excluding ATM and DNA-PK) was evaluated in a radiometric format, whereas lipid kinases as well as ATM, DNA-PK and ATR-ATRIP (ataxia telangiectasia-mutated and Rad3-related; ATR-interacting protein) were evaluated by homogeneous time-resolved fluorescence. Subsequently, 42 top-hit kinases were further evaluated to determine the INS018_055 IC_{50}. Our dose-dependent study revealed that TNIK was the most inhibited target with an IC_{50} of 31 nM (Supplementary Information 5). Notably, most of the kinases that achieved an IC_{50} at nanomolar (<1 µM) concentrations of INS018_055 have been reported to regulate fibrosis-driving processes, such as ALK4, TGFBR1 and DDR1 (Supplementary Information 5). These results indicate that our lead compound is selective at inhibiting TNIK at nanomolar concentrations and demonstrates a broader affinity toward targeting fibrosis-related kinases.

A published selectivity assay of NCB-0846 using a 50-kinase panel revealed that eight kinases achieved an IC_{50} < 100 nM[56]. INS018_055 inhibited six of these kinases with IC_{50} values at substantially higher concentrations (Supplementary Information 5). Although direct comparison of the data generated by different studies is difficult (even though both assays use K_m values for ATP concentration), these observations are suggestive of superior INS018_055 selectivity (at least among the reported kinases) associated with its different binding mode (particularly the occupation of the TNIK back pocket).

INS018_055 ameliorates TGF-β-induced EMT and FMT signaling

Epithelial-to-mesenchymal transition (EMT) and fibroblast-to-myofibroblast transition (FMT) cellular programs drive fibrosis[59] in a TGF-β-directed process[60–62]. To validate the predicted anti-fibrotic role of TNIK, we treated a lung fibroblast cell line (MRC-5) with INS018_055, which showed a dose-dependent reduction in TGF-β-induced expression of α-smooth muscle actin (α-SMA) protein (IC_{50} = 27.14 nM) (Fig. 2c). A much higher concentration of INS018_055 (50% cytotoxic concentration (CC_{50}) =84.3 µM) was required to suppress MRC-5 cell viability (Extended Data Fig. 3d). We also tested the effect of INS018_055 on TGF-β-stimulated FMT by measuring α-SMA levels in primary lung fibroblasts from three patients with IPF and three healthy donors (Fig. 2d). Sequentially diluted INS018_055 led to complete and concentration-dependent inhibition of TGF-β-mediated α-SMA expression in the fibroblasts of three donors with IPF, with IC_{50} values of 50 nM, 79 nM and 63 nM (Fig. 2d and Supplementary Information 6). In three healthy donors' fibroblasts, TGF-β-mediated α-SMA expression was reduced by INS018_055, albeit at higher concentrations (IC_{50} of 79 nM, 200 nM and 320 nM) than in donors with IPF (Supplementary Information 6), suggesting a greater potency of INS018_055 in IPF. Importantly, treatment of donor cell lines with INS018_055 did not induce substantial cytotoxicity as measured by nuclear frequency (Fig. 2d). Finally, representative images of the EMT and FMT assays carried out with donor IPF and healthy fibroblasts align perfectly with the automated quantitation of fibrosis-associated proteins (α-SMA and fibronectin 1 (FN1)) and cell density (Extended Data Fig. 4).

The effect of INS018_055 on EMT was assessed by changes in FN1 in human primary bronchial epithelial cells from patients with IPF and healthy donors. The IC_{50} values of INS018_055 in donors with IPF were 250 nM, 320 nM and 400 nM, whereas IC_{50} values for nintedanib were substantially higher (1,600 nM, 6,300 nM and 7,900 nM, respectively) (Supplementary Information 6). The IC_{50} values of INS018_055 were 63 nM, 63 nM and 500 nM in cells from healthy individuals (Supplementary Information 6), comparable to the IC_{50} value of the reference compound SB5252334. No loss of nuclei was observed in either donors with IPF or healthy donors, indicating low cytotoxicity.

We next sought to compare these observations to a known TNIK inhibitor. The molecule KY-05009 was reported to attenuate TGF-β-induced EMT in the human lung adenocarcinoma cell line A549 (ref. 46). Treatment of these epithelial cells with TGF-β induced a fibroblast-like spindle-shape phenotype, which was further enhanced by cotreatment with tumor necrosis factor (TNF)-α (Extended Data Fig. 5a). INS018_055 reversed the TGF-β-induced morphology change that was paralleled by downregulation of N-cadherin and phospho-SMAD2/SMAD3 transcription factor complex levels as well as upregulation of E-cadherin expression in a dose-dependent manner (Fig. 2e and Extended Data Fig. 5b). Furthermore, INS018_055 suppressed the level of phosphorylated p65 protein induced by TGF-β and TNF-α cotreatment (Fig. 2f and Extended Data Fig. 5b). The WNT–β-catenin signaling pathway is involved in wound healing, regulates fibrogenesis and is known to contribute to EMT elicited by TGF-β stimulation[63]. TGF-β-induced relocalization of β-catenin to chromatin was reduced by INS018_055 treatment in a dose-dependent manner, with no significant effect on cytoplasm or nuclear matrix protein levels (Fig. 2e and Extended Data Fig. 5b). These observations suggest that INS018_055 may suppress β-catenin activation by inhibiting its DNA-binding activity. Short hairpin RNA (shRNA)-mediated depletion of TNIK attenuated TGF-β-induced EMT, as evidenced by lower expression of fibronectin, N-cadherin, phospho-SMAD2, phospho-focal adhesion kinase (FAK) and total FAK relative to cells treated with control shRNA. Furthermore, E-cadherin expression was increased following TNIK depletion as compared to the control group (Fig. 2g and Extended Data Fig. 6a–c). As expected, bulk RNA-seq of A549 cells treated with TGF-β revealed an upregulation of transcriptional processes associated with extracellular matrix organization, cell–cell junctions, focal adhesions and collagen fibril organization, while treatment with INS018_055 significantly reverted these transcriptional changes (Fig. 2h). Notably, TNIK knockdown also resulted in transcriptional changes akin to those induced by INS018_055 treatment. Furthermore, Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis revealed a significant enhancement of Hippo signaling in cells treated with TGF-β (Fig. 2h and Extended Data Fig. 6d). Transcripts of Hippo pathway target genes, including those encoding known regulators of fibrosis such as PAI-1, SMAD7, CTGF[64], GLI2 (ref. 65) and PUMA[66], were downregulated in cells pretreated with INS018_055. These data indicate that INS018_055 effectively inhibits key fibrotic pathways such as WNT–β-catenin, TGF-β–SMAD2, TNF-α–NF-κB and YAP–TAZ signaling (Fig. 2i).

INS018_055 attenuates murine bleomycin-induced lung fibrosis

To evaluate the anti-fibrotic activity of INS018_055 in vivo, we used the murine bleomycin-induced lung fibrosis model[67,68]. Intratracheal bleomycin administration impaired pulmonary function,

https://doi.org/10.1038/s41587-024-02143-0

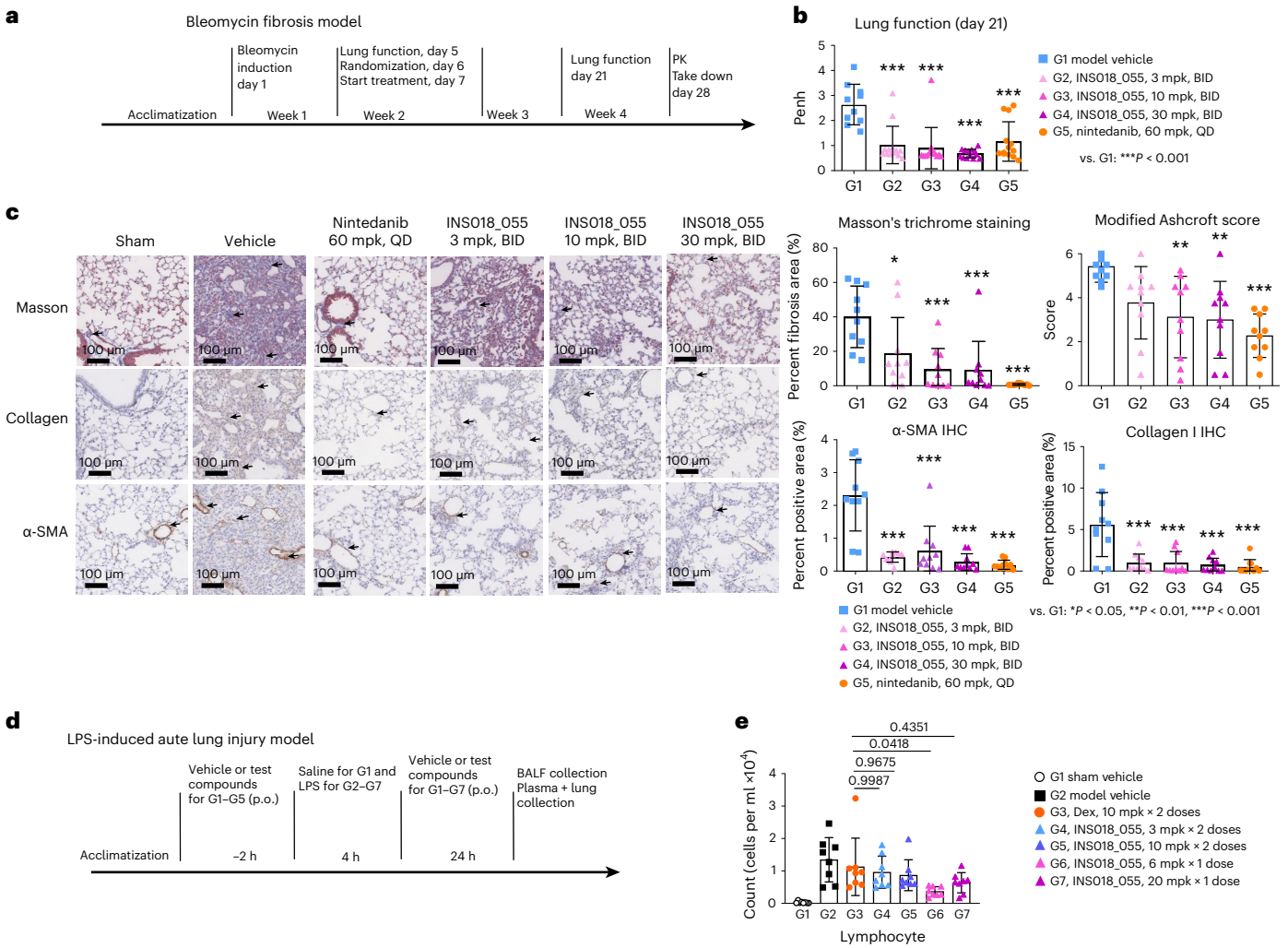


Fig. 3 | In vivo effects of INS018_055 treatment in mouse models of lung diseases. a, Study design of the bleomycin-induced lung fibrosis model in C57BL/6 male mice ($n = 10$ per group). **b**, Lung function on day 21 measured by Penh (mean ± s.d.; group (G)1, $n = 10$; groups 2–5, $n = 13$) (group 1 (vehicle) compared to groups 2–5, $P < 0.0001$; group 5 (nintedanib) compared to groups 2, 3 and 4, with $P = 0.9986$, 0.9426 and 0.464, respectively). **c**, Representative measurements of mice treated with INS018_055 (3, 10 or 30 mg per kg, BID) and nintedanib (60 mg per kg, QD) from **a** showing Masson's trichrome staining, modified Ashcroft scores and immunohistochemistry (IHC) of collagen I and α-SMA. $n = 10$ per group (mean ± s.d.) (Masson's trichrome staining: group 1 (vehicle) compared to groups 2, 3, 4 and 5, with $P = 0.0233$, 0.0004, 0.0004 and <0.0001, respectively; group 5 compared to groups 2, 3 and 4, with $P = 0.0799$, 0.8126 and 0.8289, respectively. Modified Ashcroft: group 1 compared to groups 2, 3, 4 and 5 with $P = 0.1021$, 0.0071, 0.004 and 0.0001, respectively; group 5 compared to groups 2, 3 and 4 with $P = 0.1718$, 0.791 and 0.8933, respectively. α-SMA: group 1 compared to groups 2–5 with $P \leq 0.0001$; group 5 compared to groups 2, 3 and 4 with $P = 0.9689$, 0.6125 and >0.9999, respectively. Collagen I: group 1 compared to groups 2–5 with $P \leq 0.0001$; group 5 compared to groups 2, 3 and 4, with $P = 0.9985$, 0.9986 and >0.9999, respectively). Ordinary one-way ANOVA and post hoc Šídák's multiple-comparison test were used to assess statistical significance. **d**, Study of INS018_055 in the LPS-induced acute lung injury model in C57BL/6 male mice. $n = 8$ per group. **e**, Lymphocyte cell counts from **d**. $n = 8$ per group (mean ± s.d.) (group 2 (vehicle) compared to groups 3, 4, 5, 6 and 7, $P = 0.9898$, 0.7465, 0.4720, 0.0035 and 0.0698, respectively). Dex, dexamethasone. **f**, Measurements of IL-6, IL-7, TNF-α, IL-1β and IL-4 in BALF by enzyme-linked immunosorbent assay (ELISA). $n = 8$ per group (mean ± s.d.) (group 2 (vehicle) compared for IL-6, IL-7 and IL-1β to groups 3–7 with $P < 0.0001$; for TNF-α, group 3, $P = 0.0023$ and groups 4–7, $P < 0.0001$; for IL-4, group 3, $P = 0.004$ and groups 4–7, $P < 0.0001$). Ordinary one-way ANOVA and post hoc Šídák's multiple-comparison test was used to assess statistical significance (exact P values are provided except for $^{***}P < 0.001$).

as measured by a significant increase in the respiratory parameter 'enhanced pause' (Penh) on day 21 (Fig. 3a). Two weeks of treatment with INS018_055 led to a significant, dose-dependent reduction in Penh, in line with the level achieved by nintedanib treatment (Fig. 3b). INS018_055-treated mice exhibited over 50% reduction in fibrotic area at 3 mg per kg twice daily (BID) and more than 75% reduction at 10 and

30 mg per kg BID dosages relative to vehicle-treated animals. Treatment with INS018_055 and nintedanib significantly decreased modified Ashcroft scores (a semi-quantitative grading for the severity of lung fibrosis) and lung fibrotic area relative to vehicle-treated mice. Additionally, INS018_055- and nintedanib-treated animals showed a marked reduction in α-SMA- and collagen I-positive regions (Fig. 3c). Furthermore, bleomycin administration increased lung inflammation as measured by perivascular and parabronchial infiltration of inflammatory cells, which was attenuated in a dose-dependent manner following INS018_055 or nintedanib administration (Extended Data Fig. 7a). Bronchoalveolar lavage fluid (BALF) showed reduced myeloid cell counts and pro-inflammatory cytokines in INS018_055-treated animals (Extended Data Fig. 7b).

To test whether INS018_055 synergizes with the therapeutic benefit of pirfenidone treatment in mice, we measured the in vivo effect of INS018_055 and pirfenidone combination therapy, both at suboptimal doses, in our bleomycin-induced lung fibrosis model (Extended Data Fig. 8a). Two weeks of combined INS018_055–pirfenidone treatment led to a significant improvement in lung function, as measured by a reduction in Penh (Extended Data Fig. 8b). A small but statistically insignificant improvement ($P = 0.0984$) was also observed for the subtherapeutic dose of INS018_055 alone. Treatment with the INS018_055–pirfenidone combination (3 mg per kilogram body weight (mpk) by mouth (p.o.), BID + 60 mpk, p.o., BID) significantly decreased the modified Ashcroft score, fibrotic lung area (Masson's trichrome staining) and α-SMA abundance as compared with vehicle treatment of mice, while a decreasing trend of collagen I was also observed (Extended Data Fig. 8c,d). Importantly, mice treated with the combined therapies exhibited the strongest anti-fibrotic phenotype across all parameters described. To test the effect of the INS018_055 combination with a therapeutic-level dose of pirfenidone, mice were treated with INS018_055 (3 mpk, p.o., BID), pirfenidone at the maximum therapeutic dose (200 mpk, p.o., BID) or a combination treatment of INS018_055 and pirfenidone (1 mpk + 200 mpk, 3 mpk + 200 mpk, 10 mpk + 200 mpk; p.o., BID) (Extended Data Fig. 9). A complete prevention of clinical symptoms was observed in mice treated with INS018_055 and pirfenidone (10 mpk + 200 mpk, BID).

INS018_055 reduces murine lipopolysaccharide-induced lung inflammation

As we observed that INS018_055 reduced inflammation in the bleomycin model (Extended Data Fig. 7), we hypothesized that it might elicit a broad anti-inflammatory response. To test this hypothesis, we measured the anti-inflammatory activity of INS018_055 in a lipopolysaccharide (LPS)-induced acute lung injury model (Fig. 3d). We tested two administration schemes: (1) two INS018_055 groups were dosed twice, where one dose was given before and the other after LPS challenge, (2) mice in the other two INS018_055 groups were treated with a single dose after LPS challenge. Intratracheal LPS increased BALF lymphocyte numbers (Fig. 3e), which were reduced by a single dose of INS018_055 administered 4 h after LPS challenge (Fig. 3d). Dexamethasone, a glucocorticoid used to treat acute respiratory distress syndrome[69], was used as a positive control and was administered once before and once after LPS challenge. INS018_055 at both dosing levels markedly reduced LPS-induced release of interleukin (IL)-1β, IL-6, IL-7 and TNF-α and almost completely restored IL-4 concentrations in BALF (Fig. 3f). Compared to the dexamethasone group, INS018_055-treated groups, particularly groups with 10 mg per kg (two doses) or a single dose of 20 mg per kg (after LPS stimulation), were considerably better at attenuating lung inflammation as measured by IL-1β, IL-6, IL-7 and TNF-α levels. Finally, INS018_055-treated animals had higher levels of the anti-inflammatory cytokine IL-4 than dexamethasone-treated animals. In total, these data argue in favor of INS018_055's more potent anti-inflammatory effect in this setting.

INS018_055 inhalation exclusively targets lung fibrosis

To achieve a therapeutic concentration in the lung while minimizing systemic exposure, we formulated an aerosolized INS018_055 intervention that was tested in a rat bleomycin-induced lung fibrosis model (Fig. 4a). One week after bleomycin induction, rats were exposed to INS018_055 aerosols generated by nebulization solution (0.1, 0.3, 1 or 6 mg ml^{-1}) for 30 min per day for 21 d. These values are equivalent to 0.04, 0.136, 0.485 and 2.575 mg per kg of actually delivered doses, respectively. Control animals were given NaCl solution by inhalation, while rats orally treated with pirfenidone (350 mg per kg) served as a positive control.

Plasma and lung concentrations of INS018_055 were measured in three animals immediately following administration, demonstrating approximately 50-fold higher levels in the lung than in the blood (Supplementary Information 7). Bleomycin-induced loss of lung function was significantly improved in rats treated with aerosolized INS018_055 solution across all doses and groups. Lung function was assessed by restoration of forced vital capacity (FVC, ml), pulmonary compliance (ml $cm^{-1}H_2O$) and recovery from increased pulmonary resistance (cm H_2O) (Fig. 4b). INS018_055 treatment (1 and 6 mg ml^{-1}) substantially inhibited lung fibrosis and inflammation (Fig. 4c).

TNIK inhibition mitigates kidney fibrosis

Given the lack of anti-fibrotic drugs in other diseases in which fibrosis drives organ pathology, we tested whether INS018_055 could provide therapeutic benefit in a murine kidney fibrosis model. The CC_{50} of INS018_055 in a proximal tubular kidney cell line, HK-2, was 37.37 μM (Extended Data Fig. 3d), which is 300-fold higher than its IC_{50} for α-SMA inhibition (0.104 μM), indicative of anti-fibrotic function with limited cytotoxicity (Fig. 5a). To further evaluate the effect of INS018_055 in vivo, we employed the unilateral ureteral obstruction (UUO) model[70] which mimics renal interstitial fibrosis found in humans. The vehicle-treated group showed a significant increase in kidney hydroxyproline content and Sirius red staining, indicating higher tissue fibrosis than in naive mice (Fig. 5b,c). By contrast, animals treated with INS018_055 (3, 10 and 30 mg per kg, BID) and the activin receptor-like kinase 5 (ALK5) inhibitor SB525334 (ref. 71) exhibited reduced fibrosis compared to vehicle-treated mice. Furthermore, both compounds significantly decreased hydroxyproline content (Fig. 5c) and suppressed collagen 1 staining (Fig. 5d).

Effects of INS018_055 treatment on dermal fibroblasts

To assess the anti-fibrotic effect of INS018_055 in other tissues, we used a normal human dermal fibroblast (NHDF) cell line that was previously used to measure the production of extracellular matrix proteins such as fibronectin and collagen in the context of skin fibrosis[72]. INS018_055 treatment of TGF-β-stimulated NHDF cells resulted in a concentration-dependent inhibitory effect on procollagen I, α-SMA and fibronectin expression and release with IC_{50} values of ~232 nM, 25 nM and ~135 nM, respectively (Extended Data Fig. 10a). We next sought to test the effect of topical administration of INS018_055 in a model of bleomycin-induced skin fibrosis. While bleomycin administration significantly increased collagen and hydroxyproline levels, animals treated topically with INS018_055 at doses of 0.05%, 0.15% and 0.45% showed a significant reduction in collagen and hydroxyproline levels compared to control rats (Extended Data Fig. 10b).

Phase 0 and I clinical trial testing of INS018_055

We carried out a phase 0 single-micro-dose study in healthy participants (ACTRN12621001541897) to evaluate INS018_055 pharmacokinetics (PK), in which we observed that an intravenous administration of 100 μg was well tolerated (Supplementary Information 8). A randomized, double-blind, placebo-controlled phase I clinical trial (NCT05154240) was then initiated to evaluate safety and tolerability (primary objective) as well as PK (secondary objective) of INS018_055 in 78 healthy

https://doi.org/10.1038/s41587-024-02143-0

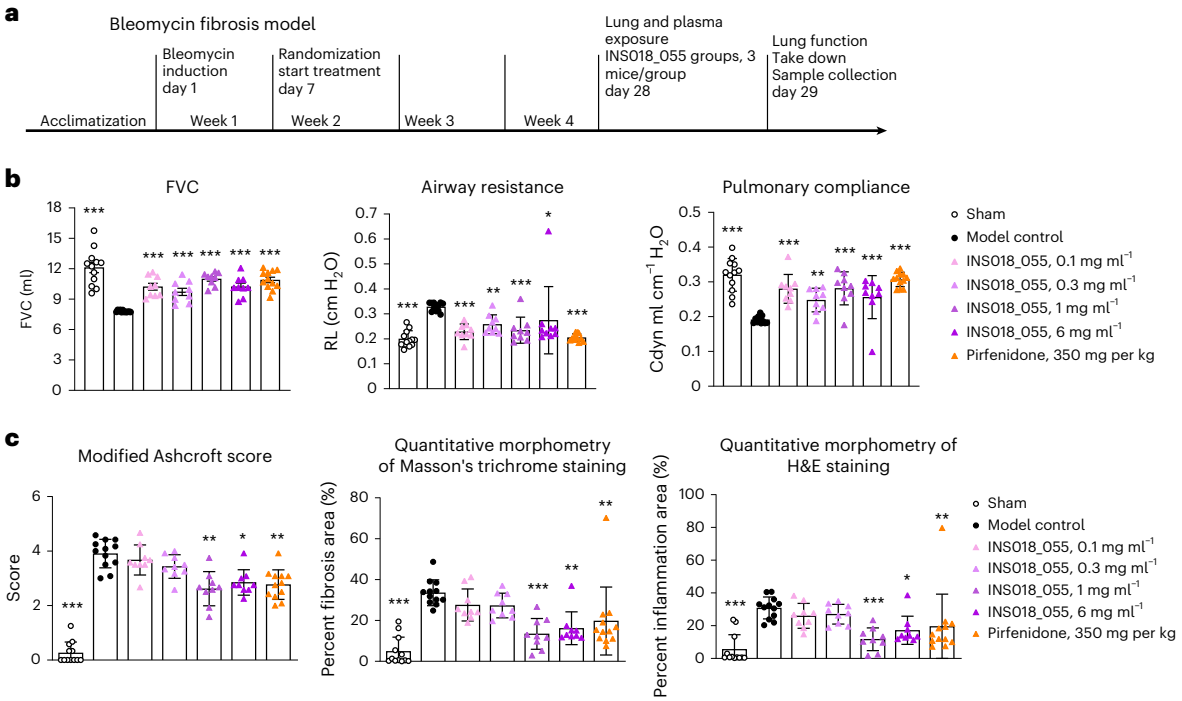


Fig. 4 | In vivo effects of INS018_055 treatment by inhalation in rat models of lung fibrosis. a, Study design of the bleomycin-induced lung fibrosis model in male Sprague Dawley rats ($n = 12$ for INS018_055 groups with three animals killed on day 28 for exposure analysis for plasma and lung, $n = 12$ for other groups). **b**, Lung function on day 29 measured by FVC, airway resistance (RL) and pulmonary compliance (Cdyn) (mean ± s.d.). Statistical analysis was performed using uncorrected two-sided Fisher's least-significant difference test as post hoc analysis after ANOVA analysis. Compared with the model control group, P values of the sham group, INS018_055 groups (0.1 mg ml^{-1}, 0.3 mg ml^{-1}, 1 mg ml^{-1}, 6 mg ml^{-1}) and the pirfenidone group (350 mg per kg) are all <0.001 for FVC; <0.001, <0.001, 0.0062, <0.001, 0.0336 and <0.001, respectively, for airway resistance; <0.001, <0.001, 0.002, <0.001, <0.001 and <0.001, respectively, for pulmonary compliance; $n = 9$ for INS018_055 groups, $n = 12$ for other groups (exact P values are provided except for ***$P < 0.001$). **c**, Quantitation for INS018_055 (0.1, 0.3, 1 or 6 mg ml^{-1}, QD, inhalation) and pirfenidone (350 mg per kg, QD, oral) groups in the bleomycin-induced lung fibrosis model shown as modified Ashcroft score, Masson's trichrome staining and hematoxylin and eosin (H&E) staining (mean ± s.d.). Statistical analysis was performed using Dunn's multiple-comparison tests after the Kruskal–Wallis test (compared with the model control group, P values of the sham group, the INS018_055 (1 mg ml^{-1}) group, the INS018_055 (6 mg ml^{-1}) group and the pirfenidone (350 mg per kg) group are <0.0001, 0.0050, 0.0223 and 0.0066 for modified Ashcroft score, respectively; <0.0001, 0.0007, 0.0040 and 0.0028 for Masson's trichrome staining, respectively; and <0.0001, 0.0008, 0.0250 and 0.0065 for H&E staining, respectively; $n = 9$ for INS018_055 groups, $n = 12$ for other groups (exact P values are provided except for ***$P < 0.001$).

volunteers. Part A of the trial was a single-ascending-dose-administration (SAD) study, with 40 participants assigned randomly at a ratio of 3:1 to receive INS018_055 or placebo at sequential doses on day 1 (10, 30, 60, 90, 120 mg). To evaluate safety and tolerability in the context of dietary effects, safety and PK evaluations were repeated with a standard high-fat meal in the cohort that received 90 mg. Part B was a multiple-ascending-dose-administration (MAD) study, with 24 healthy participants assigned randomly at a ratio of 3:1 to receive INS015_055 at sequential doses (30, 60, 120 mg) or placebo daily for 7 d.

PK profile for SAD and MAD trials

For SAD, in fasted individuals, profiles of plasma concentration versus time for INS018_055 showed a rapid absorption phase, reaching peak concentrations with median time taken to reach the maximum concentration (T_{max}) values ranging from 1.00 to 1.53 h after the dose, followed by declines in a multiphasic manner with geometric mean elimination half-life ($t_{1/2}$) values ranging from 7.42 to 9.74 h (Fig. 6a and Supplementary Information 9). INS018_055 levels were still detectable up to the last sampling point at 72 h across all doses except the cohort receiving 30 mg, whereas INS018_055 was detectable up to 48 h after the dose. Overall, the INS018_055 geometric mean peak (C_{max}) and total exposures (area under the curve (AUC)$_{0-t}$) and total exposure across time (AUC from zero to infinity (AUC_{0-inf})) increased proportionally (Supplementary Information 9) with dose except for C_{max} at the 90-mg level, which was similar to that at the 60-mg dose. The fed scenario in the cohort receiving 90 mg showed lower INS018_055 plasma absorption rates (median T_{max} of 3.00 h) than the fasted scenario (median T_{max} of 1.01 h). Total geometric C_{max} and mean AUC values were lower under fed conditions (169 ng ml^{-1} for C_{max}, 1,390 and 1,400 h × ng ml^{-1} for AUC_{0-t} and AUC_{0-inf}) than under fasted conditions (270 ng ml^{-1} for C_{max}, 1,620 and 1,630 h × ng ml^{-1} for AUC_{0-t} and AUC_{0-inf}).

For MAD, plasma exposure was measured in intensive time points after the first dose and the final dose as well as once (before the dose) in days between (Fig. 6a and Supplementary Information 9). INS018_055 C_{max}, exposure (AUC_{0-tau}) and average concentration (C_{av}) values increased with the escalation in dose level. Profiles of plasma concentration versus time for INS018_055 after multiple oral administration once a day (QD) for 7 d at dose levels of 30, 60 and 120 mg under fasted conditions were characterized by a rapid absorption phase, with peak concentrations reached at a median T_{max} ranging from 1.00 to 2.00 h after the dose, followed by overall decline in a multiphasic manner, with $t_{1/2}$ values consistent across dose levels with values ranging from 9.36 to 11.9 h after the last dose. Mean INS018_055 plasma concentrations increased with escalating doses of INS018_055. Plasma INS018_055 was detectable up to the last sampling time point at 72 h after administration at all dose levels except for the cohort receiving 30 mg, in which plasma concentrations were detectable up to 48 h. Accumulation of INS018_055 was not observed. The mean ratio of C_{max} and AUC_{0-tau} on day 7 to day 1 ranged from 0.739 to 0.905 and 0.996 to 1.05, respectively, showing a stable exposure throughout the study.

https://doi.org/10.1038/s41587-024-02143-0

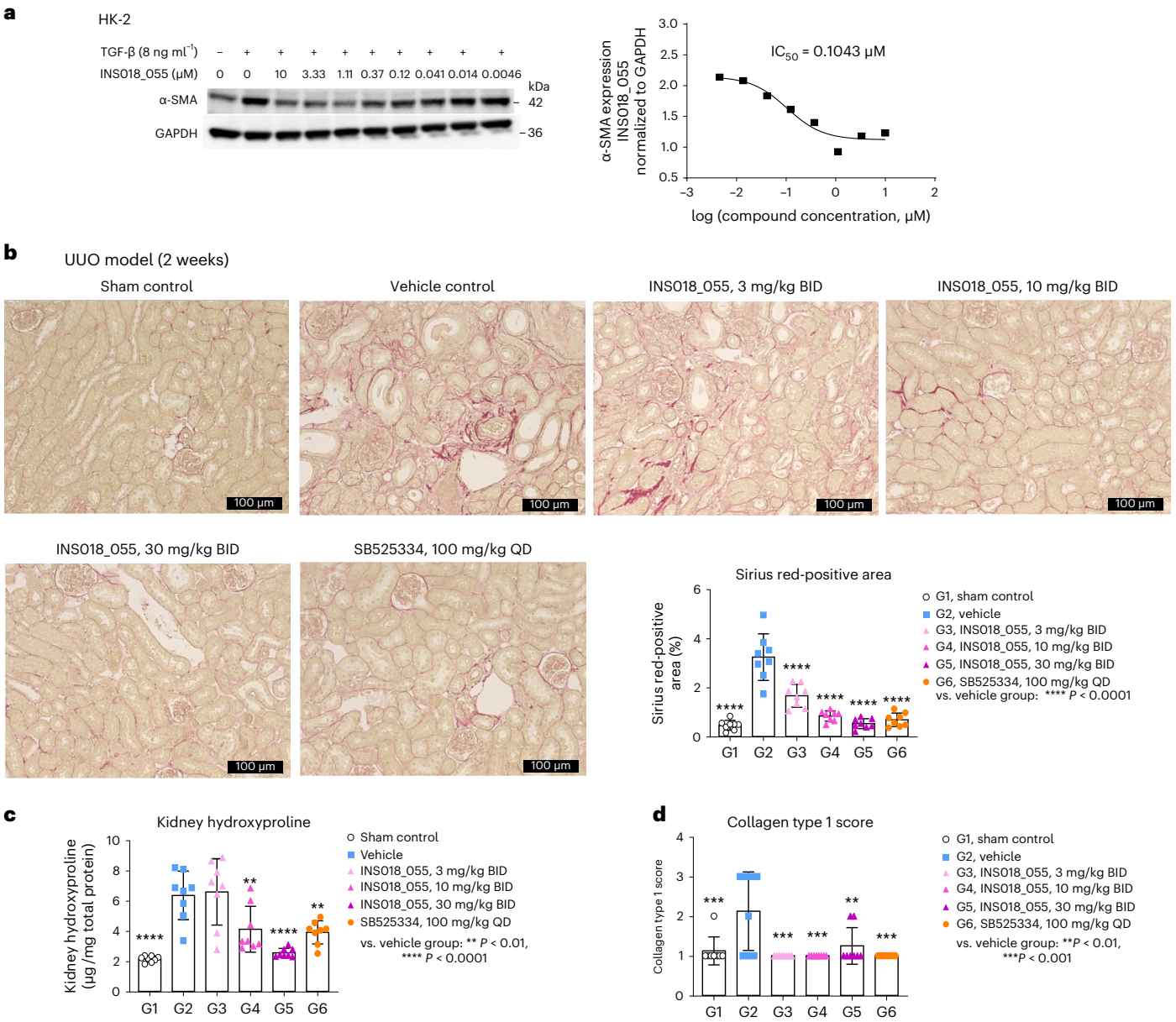


Fig. 5 | In vitro and in vivo studies on the effect of INS018_055 on kidney cells and the mouse model of kidney fibrosis. a, Left, representative western blot showing protein level changes of α-SMA with TGF-β and INS018_055 treatment. Right, graphical representation and IC_{50} determination of the effect of α-SMA expression in HK-2 cells. This represents a single independent experiment: no statistical conclusions were made. **b**, Representative images of Sirius red staining and quantitation of positive area (bottom right). **c**, Quantification of hydroxyproline content in the kidney. **d**, Scoring of IHC staining of collagen type 1 (for **b**–**d**, $n = 8$, mean ± s.d.; statistical analysis was performed using two-sided Bonferroni multiple-comparison test (using Prism 6 software) to compare between the vehicle group and other treatment groups for **b**–**d**). *P* values < 0.05 were considered statistically significant. Compared with the vehicle group, *P* values of the sham control group, INS018_055 groups (3, 10 and 30 mg per kg, BID) and the SB525334 group (100 mg per kg, QD) are all <0.0001 for Sirius red-positive area; <0.0001, >0.9999, 0.0074, <0.0001 and 0.003, respectively, for kidney hydroxyproline content; and <0.0006, 0.0001, 0.0001, 0.0028 and 0.0001, respectively, for collagen type 1 score (exact *P* values are provided except for ***$P < 0.001$).

For the phase I trial conducted in China (CTR20221542), profiles of plasma concentration versus time for INS018_055 in the SAD dose-escalation phase (with dose levels of 30, 60 and 120 mg) and in MAD cohorts administered for 7 d with 30, 60 and 90 mg BID are presented in Fig. 6b. INS018_055 showed overall comparable PK profiles in the Chinese study as compared to the New Zealand study without noteworthy differences. The only notable difference was that the plasma exposure increase from 60 to 120 mg for the single dose was slightly greater than the expected dose proportionality in the Chinese trial. After administering 30-, 60- and 90-mg doses BID for 7 d, no significant plasma accumulation of INS018_055 was observed.

Across both SAD and MAD components of phase I trials of INS018_055, the severity of all treatment-emergent adverse events including fecal occult blood positivity, headache and phlebitis (Supplementary Information 9 and 10) was mild and resolved at the end of the studies. The following exceptions were reported: one individual from the active drug-treated group in the New Zealand MAD trial exhibited influenza-like illness and seven (five from the active drug-treated group and two from the placebo group) individuals in the MAD trial in China also experienced influenza-like illnesses that were not related to the study and coincided with high prevalence of coronavirus disease 2019 (COVID-19) in China during the time of the study. The results

https://doi.org/10.1038/s41587-024-02143-0

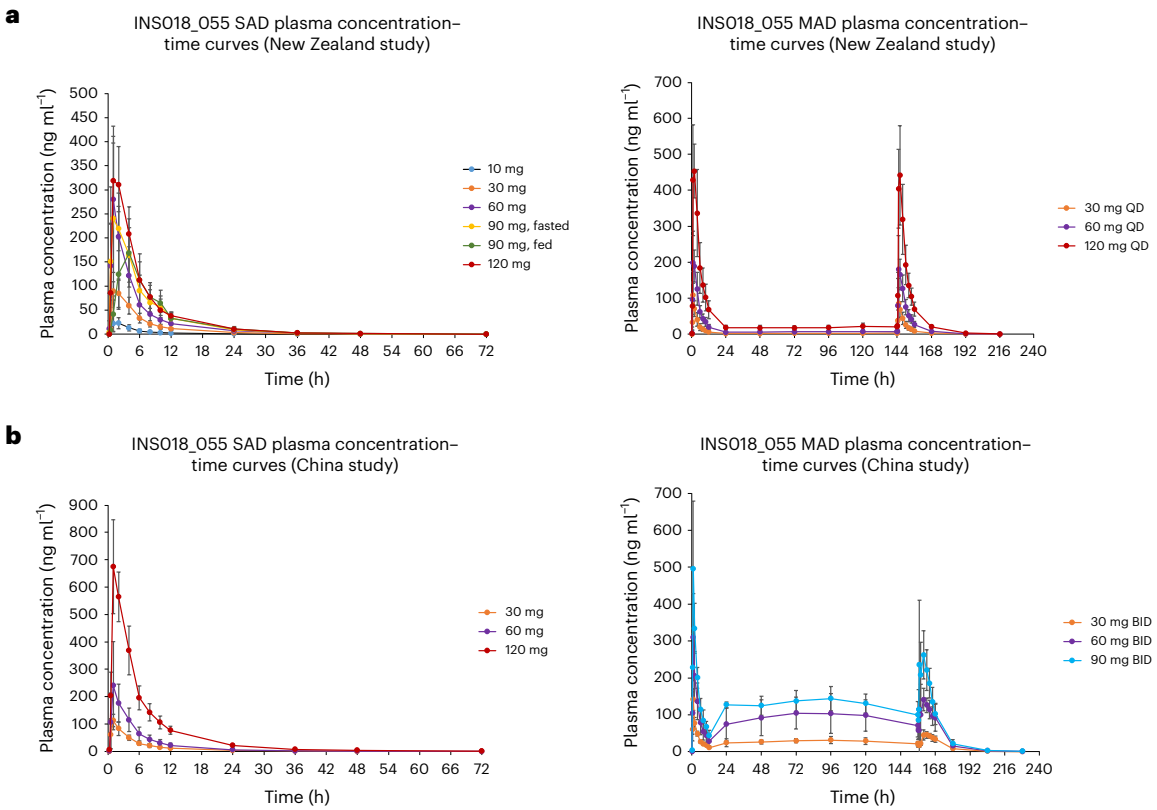


Fig. 6 | Pharmacokinetic analysis in the clinical phase I trial. a,**b**, Plasma concentrations of INS018_055 versus time in SAD and MAD of the phase I study performed in New Zealand (**a**) (mean ± s.d., $n = 6$ per group) and in SAD and MAD of the phase I study performed in China (**b**) (mean ± s.d., $n = 6$ per group).

of the two phase I studies indicate that INS018_055 is safe and well tolerated in healthy volunteers with good oral bioavailability and dose-proportional PK.

Discussion

The discovery of more effective anti-fibrotic therapies represents a major clinical challenge due to an incomplete understanding of fibrotic reprogramming in diseased tissue. Here, we navigate this limitation by harnessing the power of a generative AI platform to unbiasedly identify TNIK as an anti-fibrotic target. TNIK was also implicated in multiple hallmarks of aging, suggesting broader therapeutic potential. Using a comprehensive generative AI-driven drug-design technique, we designed a small-molecule inhibitor of TNIK, INS018_055. We demonstrate that TNIK inhibition effectively ameliorates fibrotic processes in vitro and in vivo in lung, kidney and skin fibrosis disease models. Furthermore, INS018_055 improved overall lung function in a murine model of lung fibrosis as compared to that of vehicle-treated mice. Target discovery to preclinical candidate nomination took only 18 months to accomplish, with INS018_055 evaluated in two clinical phase I trials (NCT05154240 and CTR20221542). To date, we have found that INS018_055 demonstrated safety and tolerability in healthy volunteers, providing a strong basis for further clinical studies in patients with IPF. Collectively, our preclinical and phase I clinical trial findings speak to the potential of AI-driven drug discovery in streamlining drug design in the setting of fibrosis and other contexts. Nevertheless, it is important to note that these results require further assessment in phase II and III clinical trials to validate these promising findings. At the time of this publication, two phase II trials using INS018_055 (NCT05975983 and NCT05938920) for the treatment of IPF are underway.

TNIK was identified by tasking the PandaOmics target-discovery platform[14] to search for new anti-fibrotic targets. PandaOmics was used to prioritize potential candidate targets based on the combination of omic- and text-based models (scores), which were specifically validated to identify targets using the time machine approach (as described in the Methods and Supplementary Information). TNIK scored as number 1 using the 'protein and receptor kinase' PandaOmics setting, which was based on the combination of bioinformatic and AI-driven philosophies applied to multi-model omics datasets followed by application of novelty, small-molecule druggability and protein class filters. This general methodology and pipeline can be used for drug discovery in the context of other diseases outside of fibrosis.

Our approach has provided mechanistic insight into the role of TNIK signaling during fibrosis. TNIK belongs to the germinal center kinases, which activate the profibrotic c-JUN pathway[73]. TNIK is also a RAP2 effector, which regulates cell spreading[74]. Furthermore, TNIK has been described as an essential mediator of the WNT signaling axis by direct interaction with the T-cell factor and lymphoid enhancer factor (TCF-LEF) LCF–L complex and TCF in gut epithelial cells[45]. After identifying TNIK as a target in IPF, we used Chemistry42 (refs. 40,41,75–77) to discover the lead inhibitory compounds targeting this kinase. Chemistry42 first applied 30 generative models in parallel to generate compound structures. Next, the model included feedback in the virtual screening approach to optimize the compounds further. During the generation process, the structures and associated scores produced by each generative model were dynamically shared with the other models to jointly explore the chemical space, optimize and finally generate favorable leads. INS018_055 showed the most desirable drug properties with strong selectivity for inhibiting TNIK.

While TNIK inhibition has been previously described in the context of cancer progression[46], our drug-discovery pipeline has circumvented many of the known limitations of this conventional approach. NCB-0846, an inhibitor with high potency[56], was developed and found to exhibit anti-cancer activity in multiple preclinical cancer models[56,78,79]. More recently, NCB-0846 was also reported to ameliorate liver

fibrosis[80], which is in line with orthogonal findings in our study. However, selectivity profiling has demonstrated that, at a concentration of 0.1 µM, NCB-0846 displayed inhibitory effects on additional kinase targets, including cyclin-dependent kinase (CDK)2 (>80% inhibition)[56]. One study found that NCB-0846 may impair cell viability at therapeutic doses despite its strong inhibitory effect on collagen trafficking and deposition in a murine model of liver fibrosis[80]. By contrast, INS018_055 inhibited cell viability at much higher concentrations, eliciting clear anti-fibrotic effects and ruling out non-specific cytotoxicity (Extended Data Fig. 3). Moreover, INS018_055 exhibits a favorable safety profile in nonclinical toxicology, with results from two phase I clinical trials further confirming that INS018_055 is generally well tolerated and has desirable PK profiles.

As IPF progression is driven by both fibrosis and inflammation[81,82], INS018_055 targets both main disease-driving processes in IPF evolution. The processes that initiate fibrosis are not yet completely understood, and, as a result, there are no established approaches to prevent its induction. Nevertheless, the prevalence of pulmonary fibrosis has dramatically increased due to the COVID-19 pandemic, reiterating the need for effective anti-fibrotic therapeutic interventions[83]. Given its potent anti-inflammatory and anti-fibrotic activity, INS018_055 may be used as a preventive agent for patients with a high risk of developing post-COVID-19 acute respiratory distress syndrome, including but not limited to older and immunocompromised individuals[84,85]. Moreover, those at high risk of developing organ failure due to progressive CKD may[86] also benefit from preventative therapy with INS018_055.

In summary, this generative AI-enabled study aimed at developing better compounds for fibrosis-related diseases provides evidence that generative AI platforms offer time-efficient solutions for generating target-specific drugs with potent anti-fibrotic activity. INS018_055 demonstrated good tolerability and favorable PK profiles in both phase I trials conducted in New Zealand and China, with a phase IIa trial in patients with IPF currently underway. We believe that this study underscores the strength of AI-enabled drug-discovery approaches, which will likely revolutionize drug discovery.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41587-024-02143-0.

References

1. Harrison, R. K. Phase II and phase III failures: 2013–2015. *Nat. Rev. Drug Discov.* **15**, 817–818 (2016).
2. Arrowsmith, J. & Miller, P. Trial watch: phase II and phase III attrition rates 2011–2012. *Nat. Rev. Drug Discov.* **12**, 569 (2013).
3. Krieger, J. L., Li, D. & Papanikolaou, D. *Missing Novelty in Drug Development* NBER Working Paper No. w24595 (National Bureau of Economic Research, 2018).
4. Oprea, T. I. et al. Unexplored therapeutic opportunities in the human genome. *Nat. Rev. Drug Discov.* **17**, 317–332 (2018).
5. Aliper, A. Prediction of clinical trials outcomes based on target choice and clinical trial design with multi-modal artificial intelligence. *Clin. Pharmacol. Ther.* **114**, 972–980 (2023).
6. Pun, F. W., Ozerov, I. V. & Zhavoronkov, A. AI-powered therapeutic target discovery. *Trends Pharmacol. Sci.* **44**, 561–572 (2023).
7. West, M. D. et al. Use of deep neural network ensembles to identify embryonic–fetal transition markers: repression of *COX7A1* in embryonic and cancer cells. *Oncotarget* **9**, 7796–7811 (2018).
8. Mamoshina, P. et al. Machine learning on human muscle transcriptomic data for biomarker discovery and tissue-specific drug target identification. *Front. Genet.* **9**, 242 (2018).
9. Ozerov, I. V. et al. In silico pathway activation network decomposition analysis (iPANDA) as a method for biomarker development. *Nat. Commun.* **7**, 13427 (2016).
10. Aliper, A. et al. Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data. *Mol. Pharm.* **13**, 2524–2530 (2016).
11. Broner, E. C. et al. Doublecortin-like kinase 1 (DCLK1) is a novel NOTCH pathway signaling regulator in head and neck squamous cell carcinoma. *Front. Oncol.* **11**, 677051 (2021).
12. Stamatas, G. N. et al. An analysis of gene expression data involving examination of signaling pathways activation reveals new insights into the mechanism of action of minoxidil topical foam in men with androgenetic alopecia. *Cell Cycle* **16**, 1578–1584 (2017).
13. Pasteuning-Vuhman, S. et al. New function of the myostatin/activin type I receptor (ALK4) as a mediator of muscle atrophy and muscle regeneration. *FASEB J.* **31**, 238–255 (2017).
14. Pun, F. W. et al. Identification of therapeutic targets for amyotrophic lateral sclerosis using PandaOmics — an AI-enabled biological target discovery platform. *Front. Aging Neurosci.* **14**, 914017 (2022).
15. Pun, F. W. et al. Hallmarks of aging-based dual-purpose disease and age-associated targets predicted using PandaOmics AI-powered discovery engine. *Aging* **14**, 2475–2506 (2022).
16. Duffield, J. S., Lupher, M., Thannickal, V. J. & Wynn, T. A. Host responses in tissue repair and fibrosis. *Annu. Rev. Pathol.* **8**, 241–276 (2013).
17. Humphreys, B. D. Mechanisms of renal fibrosis. *Annu. Rev. Physiol.* **80**, 309–326 (2018).
18. Raghu, G. Idiopathic pulmonary fibrosis: lessons from clinical trials over the past 25 years. *Eur. Respir. J.* **50**, 1701209 (2017).
19. Wynn, T. A. Integrating mechanisms of pulmonary fibrosis. *J. Exp. Med.* **208**, 1339–1350 (2011).
20. Vaughan, M. B., Howard, E. W. & Tomasek, J. J. Transforming growth factor-β1 promotes the morphological and functional differentiation of the myofibroblast. *Exp. Cell Res.* **257**, 180–189 (2000).
21. Raghu, G., Weycker, D., Edelsberg, J., Bradford, W. Z. & Oster, G. Incidence and prevalence of idiopathic pulmonary fibrosis. *Am. J. Respir. Crit. Care Med.* **174**, 810–816 (2006).
22. Maher, T. M. et al. Global incidence and prevalence of idiopathic pulmonary fibrosis. *Respir. Res.* **22**, 197 (2021).
23. Collard, H. R. et al. Health care utilization and costs of idiopathic pulmonary fibrosis in U.S. Medicare beneficiaries aged 65 years and older. *Ann. Am. Thorac. Soc.* **12**, 981–987 (2015).
24. Raghu, G., Chen, S. Y., Hou, Q., Yeh, W. S. & Collard, H. R. Incidence and prevalence of idiopathic pulmonary fibrosis in US adults 18–64 years old. *Eur. Respir. J.* **48**, 179–186 (2016).
25. Coultas, D. B., Zumwalt, R. E., Black, W. C. & Sobonya, R. E. The epidemiology of interstitial lung diseases. *Am. J. Respir. Crit. Care Med.* **150**, 967–972 (1994).
26. Ley, B., Collard, H. R. & King, T. E. Jr. Clinical course and prediction of survival in idiopathic pulmonary fibrosis. *Am. J. Respir. Crit. Care Med.* **183**, 431–440 (2011).
27. Collard, H. R. & King, T. E. Jr. Treatment of idiopathic pulmonary fibrosis: the rise and fall of corticosteroids. *Am. J. Med.* **110**, 326–328 (2001).
28. Lynch, J. P. 3rd, Fishbein, M. C., Saggar, R., Zisman, D. A. & Belperio, J. A. Idiopathic pulmonary fibrosis. *Expert Rev. Respir. Med.* **1**, 377–389 (2007).
29. Petnak, T., Lertjitbanjong, P., Thongprayoon, C. & Moua, T. Impact of antifibrotic therapy on mortality and acute exacerbation in idiopathic pulmonary fibrosis: a systematic review and meta-analysis. *Chest* **160**, 1751–1763 (2021).

30. Pleasants, R. & Tighe, R. M. Management of idiopathic pulmonary fibrosis. *Ann. Pharmacother.* **53**, 1238–1248 (2019).
31. Stahnke, T. et al. Suppression of TGF-β pathway by pirfenidone decreases extracellular matrix deposition in ocular fibroblasts in vitro. *PLoS ONE* **12**, e0172592 (2017).
32. Hilberg, F. et al. BIBF 1120: triple angiokinase inhibitor with sustained receptor blockade and good antitumor efficacy. *Cancer Res.* **68**, 4774–4782 (2008).
33. Wollin, L. et al. Mode of action of nintedanib in the treatment of idiopathic pulmonary fibrosis. *Eur. Respir. J.* **45**, 1434–1445 (2015).
34. Liu, Y. Cellular and molecular mechanisms of renal fibrosis. *Nat. Rev. Nephrol.* **7**, 684–696 (2011).
35. Vallianou, N. G., Mitesh, S., Gkogkou, A. & Geladari, E. Chronic kidney disease and cardiovascular disease: is there any relationship? *Curr. Cardiol. Rev.* **15**, 55–63 (2019).
36. Chen, T. K., Knicely, D. H. & Grams, M. E. Chronic kidney disease diagnosis and management: a review. *JAMA* **322**, 1294–1304 (2019).
37. Black, L. M., Lever, J. M. & Agarwal, A. Renal inflammation and fibrosis: a double-edged sword. *J. Histochem. Cytochem.* **67**, 663–681 (2019).
38. Choi, M. E., Ding, Y. & Kim, S. I. TGF-β signaling via TAK1 pathway: role in kidney fibrosis. *Semin. Nephrol.* **32**, 244–252 (2012).
39. Schnaper, H. W., Hayashida, T., Hubchak, S. C. & Poncelet, A. C. TGF-β signal transduction and mesangial cell fibrogenesis. *Am. J. Physiol. Renal Physiol.* **284**, F243–F252 (2003).
40. Ivanenkov, Y. A. et al. Chemistry42: an AI-driven platform for molecular design and optimization. *J. Chem. Inf. Model.* **63**, 695–701 (2023).
41. Ren, F. et al. AlphaFold accelerates artificial intelligence powered drug discovery: efficient discovery of a novel CDK20 small molecule inhibitor. *Chem. Sci.* **14**, 1443–1452 (2023).
42. Aliper, A. et al. Prediction of clinical trials outcomes based on target choice and clinical trial design with multi-modal artificial intelligence. *Clin. Pharmacol. Ther.* **114**, 972–980 (2023).
43. Ochoa, D. et al. Open Targets Platform: supporting systematic drug-target identification and prioritisation. *Nucleic Acids Res.* **49**, D1302–D1310 (2021).
44. Cannon, D. C. et al. TIN-X: target importance and novelty explorer. *Bioinformatics* **33**, 2601–2603 (2017).
45. Mahmoudi, T. et al. The kinase TNIK is an essential activator of Wnt target genes. *EMBO J.* **28**, 3329–3340 (2009).
46. Kim, J. et al. A novel aminothiazole KY-05009 with potential to inhibit Traf2- and Nck-interacting kinase (TNIK) attenuates TGF-β1-mediated epithelial-to-mesenchymal transition in human lung adenocarcinoma A549 cells. *PLoS ONE* **9**, e110180 (2014).
47. Kaneko, S. et al. Smad inhibition by the Ste20 kinase Misshapen. *Proc. Natl Acad. Sci. USA* **108**, 11127–11132 (2011).
48. Li, Q. et al. The Misshapen subfamily of Ste20 kinases regulate proliferation in the aging mammalian intestinal epithelium. *J. Cell. Physiol.* **234**, 21925–21936 (2019).
49. Larhammar, M., Huntwork-Rodriguez, S., Rudhard, Y., Sengupta-Ghosh, A. & Lewcock, J. W. The Ste20 family kinases MAP4K4, MINK1, and TNIK converge to regulate stress-induced JNK signaling in neurons. *J. Neurosci.* **37**, 11074–11084 (2017).
50. Shkoda, A. et al. The germinal center kinase TNIK is required for canonical NF-κB and JNK signaling in B-cells by the EBV oncoprotein LMP1 and the CD40 receptor. *PLoS Biol.* **10**, e1001376 (2012).
51. Henderson, N. C., Rieder, F. & Wynn, T. A. Fibrosis: from mechanisms to medicines. *Nature* **587**, 555–566 (2020).
52. Pham, T. C. P. et al. TNIK is a conserved regulator of glucose and lipid metabolism in obesity. *Sci. Adv.* **9**, eadf7119 (2023).
53. Adams, T. S. et al. Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. *Sci. Adv.* **6**, eaba1983 (2020).
54. Osorio, D. et al. scTenifoldKnk: an efficient virtual knockout tool for gene function predictions via single-cell gene regulatory network perturbation. *Patterns* **3**, 100434 (2022).
55. Bader, G. D. & Hogue, C. W. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* **4**, 2 (2003).
56. Masuda, M. et al. TNIK inhibition abrogates colorectal cancer stemness. *Nat. Commun.* **7**, 12586 (2016).
57. Ayala-Aguilera, C. C. et al. Small molecule kinase inhibitor drugs (1995–2021): medical indication, pharmacology, and synthesis. *J. Med. Chem.* **65**, 1047–1131 (2022).
58. Zuccotto, F., Ardini, E., Casale, E. & Angiolini, M. Through the 'gatekeeper door': exploiting the active kinase conformation. *J. Med. Chem.* **53**, 2681–2694 (2010).
59. Hinz, B. et al. The myofibroblast: one function, multiple origins. *Am. J. Pathol.* **170**, 1807–1816 (2007).
60. Roberts, A. B. et al. Transforming growth factor type β: rapid induction of fibrosis and angiogenesis in vivo and stimulation of collagen formation in vitro. *Proc. Natl Acad. Sci. USA* **83**, 4167–4171 (1986).
61. Miettinen, P. J., Ebner, R., Lopez, A. R. & Derynck, R. TGF-β induced transdifferentiation of mammary epithelial cells to mesenchymal cells: involvement of type I receptors. *J. Cell Biol.* **127**, 2021–2036 (1994).
62. Desmouliere, A., Geinoz, A., Gabbiani, F. & Gabbiani, G. Transforming growth factor-β 1 induces α-smooth muscle actin expression in granulation tissue myofibroblasts and in quiescent and growing cultured fibroblasts. *J. Cell Biol.* **122**, 103–111 (1993).
63. Lamouille, S., Xu, J. & Derynck, R. Molecular mechanisms of epithelial–mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* **15**, 178–196 (2014).
64. Wahab, N. A., Weston, B. S. & Mason, R. M. Modulation of the TGFβ/Smad signaling pathway in mesangial cells by CTGF/CCN2. *Exp. Cell Res.* **307**, 305–314 (2005).
65. Fabian, S. L. et al. Hedgehog–Gli pathway activation during kidney fibrosis. *Am. J. Pathol.* **180**, 1441–1453 (2012).
66. Wang, W. et al. p53/PUMA expression in human pulmonary fibroblasts mediates cell activation and migration in silicosis. *Sci. Rep.* **5**, 16900 (2015).
67. Shahzeidi, S., Jeffery, P. K., Laurent, G. J. & McAnulty, R. J. Increased type I procollagen mRNA transcripts in the lungs of mice during the development of bleomycin-induced fibrosis. *Eur. Respir. J.* **7**, 1938–1943 (1994).
68. Adamson, I. Y. & Bowden, D. H. The pathogenesis of bleomycin-induced pulmonary fibrosis in mice. *Am. J. Pathol.* **77**, 185–197 (1974).
69. Tomazini, B. M. et al. Effect of dexamethasone on days alive and ventilator-free in patients with moderate or severe acute respiratory distress syndrome and COVID-19: the CoDEX Randomized Clinical Trial. *JAMA* **324**, 1307–1316 (2020).
70. Chevalier, R. L., Forbes, M. S. & Thornhill, B. A. Ureteral obstruction as a model of renal interstitial fibrosis and obstructive nephropathy. *Kidney Int.* **75**, 1145–1152 (2009).
71. Denby, L. et al. MicroRNA-214 antagonism protects against renal fibrosis. *J. Am. Soc. Nephrol.* **25**, 65–80 (2014).
72. Montagnani, S. et al. Granulocyte macrophage colony stimulating factor (GM-CSF) biological actions on human dermal fibroblasts. *Eur. J. Histochem.* **45**, 219–228 (2001).
73. Fu, C. A. et al. TNIK, a novel member of the germinal center kinase family that activates the c-Jun N-terminal kinase pathway and regulates the cytoskeleton. *J. Biol. Chem.* **274**, 30729–30737 (1999).
74. Taira, K. et al. The Traf2- and Nck-interacting kinase as a putative effector of Rap2 to regulate actin cytoskeleton. *J. Biol. Chem.* **279**, 49488–49496 (2004).

75. Polykovskiy, D. et al. Entangled conditional adversarial autoencoder for de novo drug discovery. *Mol. Pharm.* **15**, 4398–4405 (2018).
76. Vanhaelen, Q., Lin, Y. C. & Zhavoronkov, A. The advent of generative chemistry. *ACS Med. Chem. Lett.* **11**, 1496–1505 (2020).
77. Zhavoronkov, A. et al. Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nat. Biotechnol.* **37**, 1038–1040 (2019).
78. Sekita, T. et al. Feasibility of targeting Traf2-and-Nck-interacting kinase in synovial sarcoma. *Cancers* **12**, 1258 (2020).
79. Torres-Ayuso, P. et al. TNIK is a therapeutic target in lung squamous cell carcinoma and regulates FAK activation through Merlin. *Cancer Discov.* **11**, 1411–1423 (2021).
80. Buchl, S. C. et al. Traf2 and NCK interacting kinase is a critical regulator of procollagen I trafficking and hepatic fibrogenesis in mice. *Hepatol. Commun.* **6**, 593–609 (2022).
81. Chanda, D. et al. Developmental pathways in the pathogenesis of lung fibrosis. *Mol. Aspects Med.* **65**, 56–69 (2019).
82. Martinez, F. J. et al. Idiopathic pulmonary fibrosis. *Nat. Rev. Dis. Primers* **3**, 17074 (2017).
83. Gallelli, L., Zhang, L., Wang, T. & Fu, F. Severe acute lung injury related to COVID-19 infection: a review and the possible role for escin. *J. Clin. Pharmacol.* **60**, 815–825 (2020).
84. Spagnolo, P. et al. Pulmonary fibrosis secondary to COVID-19: a call to arms? *Lancet Respir. Med.* **8**, 750–752 (2020).
85. Grillo, F., Barisione, E., Ball, L., Mastracci, L. & Fiocca, R. Lung fibrosis: an undervalued finding in COVID-19 pathological series. *Lancet Infect. Dis.* **21**, e72 (2021).
86. Jansen, J. et al. SARS-CoV-2 infects the human kidney and drives fibrosis in kidney organoids. *Cell Stem Cell* **29**, 217–231 (2022).

Methods

Ethics

Human tissue and clinical trials. All tissues used for isolation were obtained with informed consent and conformance to HIPAA regulations to protect the privacy of the donor's personally identifiable information. Humans who participated in any clinical trial in this study provided written consent for study participation. The entirety of this study adheres to the Declaration of Helsinki.

Phase 0 clinical trial in Australia. This trial was conducted at the CMAX Clinical Research Center, 18A North Terrace, Adelaide, SA 5000, Australia. This phase 0 clinical trial, also the first-in-human clinical trial of INS018_055, was conducted in Australia (ACTRN12621001541897). This trial was conducted in accordance with the ethical principles of good clinical practice, according to the International Council for Harmonisation (ICH) Harmonised Guideline E6(R2) Integrated Addendum to ICH E6(R1): Guideline for Good Clinical Practice ICH E6(R2), annotated with comments by the Australian Therapeutic Goods Administration (2018). This study was approved by the Bellberry Human Research Ethics Committee.

Phase I clinical trial in China. This trial was conducted at Zhejiang Xiaoshan Hospital, 728 Yucai North Road, Xiaoshan District, Hangzhou, Zhejiang 311202, China. Detailed information including facilities and inclusion and exclusion criteria can be found at the following link: http://www.chinadrugtrials.org.cn/clinicaltrials.prosearch.dhtml (registration number CTR20221542). This study was approved by the Zhejiang Xiaoshan Hospital clinical trial ethics committee. The number of the ethics committee approval letter is EC-2022102505.

Phase I clinical trial in New Zealand. This trial was conducted at the New Zealand Clinical Research Center, level 4, 264 Antigua Street, Central City, Christchurch, 8011, New Zealand (registration number NCT05154240). This study was approved by the Northern B Health and Disability Ethics Committee. The ethics reference number for this protocol is 2021 FULL 11770.

Animal studies. All animal studies were ethically and humanely conducted following institutional animal care and use committee (IACUC), IRB or relevant animal-handling ethics organizational guidelines by our partnering CROs. Murine LPS and bleomycin studies were carried out by HD Biosciences, which were approved under the following AUF protocol numbers: AUF 146 and AUF 117. These AUF protocols were approved by the IACUC at HD Biosciences, and the studies were conducted at an AAALAC-accredited facility. The bleomycin fibrosis study in rats that were administered INS018_055 as an inhalable agent was carried out by JOINN Laboratories. This study's IACUC-approved procedural number is S-ACU22-1144. JOINN Laboratories is fully accredited by the AAALAC, and the study adheres to regulations and rules laid out by the IACUC. The UUO kidney study was performed by SMC Laboratories. The animal care and use committee-approved protocol number is U32. All animals used in this study were housed and cared for in accordance with the Japanese Pharmacological Society Guidelines for Animal Use. The skin fibrosis model was carried out by TheraIndx LifeSciences. This study was performed using protocols approved by the Institutional Animals Ethics Committee of the test facility, which was designed under CPCSEA guidelines for animal care. The registration number for this protocol is 1852/PO/Rc/S/16/CPCSEA. For all aforementioned studies, animals were housed in groups (<5 per cage) in temperature-controlled facilities (20–26 °C) with 12-h light–12-h dark–light cycles. All animals had ad libitum access to drinking food and water.

ChatPaperGPT tool

To encourage a thorough, unbiased interrogation of our study, we developed an interactive generative AI tool, ChatPaperGPT, to navigate all contents of our study. We encourage all readers to critically examine our study, its accompanying state-of-the-art drug-discovery methodology and all associated data using our ChatPaperGPT tool (https://papers.insilicogpt.com/).

Antibodies

Antibodies used were anti-human-fibronectin (Invitrogen, MA5-11981), anti-human E-cadherin (BD Biosciences, 610182), anti-human N-cadherin (D4R1H) (Cell Signaling, 13116S), anti-human phospho-SMAD2 (Ser465/467)/SMAD3 (Ser423/425) (D27F4) (Cell Signaling, 8828S), anti-human phospho-SMAD2 (Ser465/467) (Cell Signaling, 3108S), anti-human SMAD2/SMAD3 (D7G7) (Cell Signaling, 12470S), anti-human phospho-FAK (Tyr397) (D20B1) (Cell Signaling, 8556), anti-human FAK (Cell Signaling, 3285), anti-human TNIK (Cell Signaling, 32712), anti-human NF-κB p65 (D14E12) (Cell Signaling, 8242S), anti-human phospho-NF-κB p65 (Ser536) (93H1) (Cell Signaling, 3033S), anti-human β-catenin (Cell Signaling, 9562), anti-human α-tubulin (Abcam, ab18251), anti-human HDAC2 (Abcam, ab12169), anti-human histone 3 (Abcam, ab176842), anti-α-SMA antibodies (Cell Signaling, 19245; Abcam, ab5694; Abcam, ab32575), anti-GAPDH (Absin, 016D), anti-collagen I antibody (Abcam, ab34710; LSL), goat anti-mouse IgG (Abcam, ab205719), goat anti-rabbit IgG (Abcam, ab205718) and goat polyclonal antibodies conjugated to HRP (DAKO, K4003; Vector Lab, PI-1000).

Enzyme-linked immunosorbent assay kits

ELISA kits used were as follows: IL-1β (Sinobest Bio, YX-E01291M), IL-4 (Sinobest Bio, YX-E00064M), IL-6 (Sinobest Bio, YX-E00066M), TNF-α (Sinobest Bio, YX-E00104M), fibronectin (Takara, MK115), procollagen type I C peptide (Takara, MK101), rat hydroxyproline (KinesisDx, K11-0512), collagen (Abcam, ab222942).

Other reagents

Other reagents used were as follows: 20× TBS Tween-20 (Thermo, 228360), 4–20% Criterion TGX Precast Gels, 26 well, (Bio-Rad, 5675671095), 4% Paraformaldehyde Fix Solution (Beyotime, P0099), bovine pituitary extract (ScienCell, 0713-1100mg), BSA (Sigma, B2064), bleomycin sulfate (MCE, HY-17565), CellTiter-Glo Buffer (Promega, G756B), CellTiter-Glo Substrate (Promega, G755B), cOmplete, EDTA-free Protease Inhibitor Cocktail (Roche, 4693134693132001), Difco skim milk (BD, 23232100), dimethylsulfoxide (DMSO) (Sigma, D2650-1100mL), DMEM high-glucose culture medium (Gibco, 8128120212), DMEM (Gibco, 119611960051), Dulbecco's phosphate-buffered saline (DPBS) (Corning, 2103121031-CVC), eosin Y (Baso, BA4022), FBS (ExCell Bio, 12040), hematoxylin solution (Baso, BA4041), human recombinant epidermal growth factor (EGF) (Gibco, PHG0314), human TNF-α (Sigma, T0157), iBlot 2 NC Regular Stacks (Invitrogen, IB223001), keratinocyte serum-free medium (K-SFM) (Gibco, 10724011), Lipofectamine 3000 Transfection Kit (Thermo Fisher, L3000008), LPS from *Escherichia coli* O55:B5 (Sigma, L2880), MEM non-essential amino acids (100×) (Gibco, 11140050), methylcellulose (Sigma, M7140), MOPS SDS Running Buffer (1×) (Invitrogen, NP0001), nintedanib esylate (DC, DC8608), Novex ECL Chemiluminescent Substrate Reagent Kit (Invitrogen, WP220005), NuPAGE 4–12% Bis-Tris Gel, 15 well (Invitrogen, NP0330336BOX), NuPAGE LDS sample buffer (4×) (Invitrogen, NP0007), Opti-MEM (Gibco, 319831985062), PageRuler Plus Prestained Protein Ladder (Thermo Fisher, 226619), penicillin–streptomycin solution (10,000 U ml^{-1} penicillin, 10,000 μg ml^{-1} streptomycin) (HyClone, SV330010), Phosphatase Inhibitor Cocktail 2 (Sigma, P57265ML), Phosphatase Inhibitor Cocktail 3 (Sigma, P00445ML), PhosSTOP phosphatase inhibitor cocktail (Roche, 4906844906845001), Pierce western blot transfer buffer (10×) (Thermo Fisher, 335040), Pierce BCA Protein Assay Kit (Pierce, 223227), PMSF protease inhibitor (Dalian Meilun, MA0001-May06E), PMSF protease inhibitor (Beyotime, 105212110521210126), Recombinant Human

TGF-β 1 Protein (MCE, AV6120051, R&D, 240-B-010 and 240-b002), Restore Western Blot Stripping Buffer (Pierce, 221059), radioimmunoprecipitation assay (RIPA) buffer (Sigma, R0278), RIPA lysis buffer (Beyotime, 926192092619200827), RPMI 1640 medium (Gibco, 22400089), saline (Hualu, NMPN 37022750), Subcellular Protein Fractionation Kit for Cultured Cells (Thermo Fisher, 778840), SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Fisher, 334096), Thermo Scientific PageRuler Prestained Protein Ladder (Thermo, 226619), trypsin-EDTA (0.25%), phenol red (Gibco, 25200072).

Compounds

INS018_055 was provided by WuXi AppTec. SB525334, was purchased from TCI. Nintedanib esylate was purchased from DC (DC8608). Dexamethasone was purchased from Sigma (4902). Pirfenidone was purchased from DC (DC8792).

Synthesis of INS018_055

Synthesis, stability and characterization of INS018_055 are laid out in detail in Supplementary Information 4.

In vitro experiments

Cell lines. The human fetal lung fibroblast cell line MRC-5 was purchased from ATCC (CCL-171) (the cell viability assay was performed at WuXi AppTec) and Shanghai Cell Bank, Chinese Academy of Sciences (the α-SMA assay was performed at Boji). Human primary bronchial epithelial cells were derived from three donors with IPF (IPF05, IPF06 and IPF08) and three healthy donors (Br285, Br311 and 410955). Human primary lung fibroblasts were derived from three donors with IPF (IPF05, IPF06 and IPF08) and three healthy donors (FB218, 03HF67101 and FB2382). The human lung adenocarcinoma cell line A549 was purchased from ATCC (CCL-185). Human embryonic kidney 293T/17 cells were purchased from ATCC (CRL-11268). The human kidney cell line HK-2 was purchased from ATCC (CRL-2190). NHDFs were purchased from Bioalternatives (PF2).

Mycoplasma testing. All cell lines used were checked and cleared for mycoplasma at the following contract research organizations: WuXi AppTec (China), Charles Rivers (Netherlands), Guangzhou Boji Medical Biotechnological (China), Bioalternatives (France).

Cell culture conditions. All cell lines were cultured at 37 °C in a humidified atmosphere of 95% air and 5% CO_2. MRC-5 cells were cultured in MEM medium supplemented with 1% GlutaMAX, 10% FBS, 1% non-essential amino acids and 1% penicillin–streptomycin (100 U ml^{-1}). Human primary lung fibroblasts and human primary bronchial epithelial cells were cultured according to Charles River's internal protocol. A549 cells were cultured in RPMI 1640 medium supplemented with 10% FBS and antibiotics (penicillin and streptomycin, 100 U ml^{-1} each). The 293T/17 cells were cultured in DMEM medium supplemented with 10% FBS and 1% penicillin–streptomycin solution (100 U ml^{-1}). HK-2 cells were cultured in K-SFM containing 50% growth factor. Human primary lung fibroblasts, NHDF cells, were cultured in DMEM supplemented with 2 mM glutamine, 50 U ml^{-1} penicillin, 50 µg ml^{-1} streptomycin and 10% fetal calf serum.

Surface plasmon resonance assay for binding to His-tagged TNIK (9–315). Binding kinetics for INS018_055, NCB-0846 and KY-05009 were measured with the Biacore 8K system. His-tagged TNIK (9–315) (WuXi AppTec) was immobilized on an NTA sensor chip for 43 s at 30 µg ml^{-1} and 5 µg ml^{-1} (His capture and amine coupling). The three compounds INS018_055, NCB-0846 and KY-05009 were prepared by twofold dilutions from 1,000 nM to 2 nM. The association time and the dissociation time were 90 s and 180 s, respectively (flow rate, 30 µl per min; sample compartment temperature, 15 °C; analysis temperature, 15 °C). The running buffer contained 20 mM HEPES, pH 7.5, 150 mM NaCl, 1 mM $MgCl_2$, 0.05% Tween-20, 2% glycerol, 1 mM TCEP and 2% DMSO.

Enzyme selectivity profiling. INS018_055 was tested against selected kinases using Eurofins standard KinaseProfiler assays. A total of 430 kinases were tested. Protein kinases (with the exception of ATM(h) and DNA-PK(h)) were assayed in a radiometric format, whereas lipid kinases, ATM(h), ATR/ATRIP(h) and DNA-PK(h) were assayed using a homogeneous time-resolved fluorescence format. INS018_055 was prepared in a working stock 50× the final assay concentration in 100% DMSO. The required volume of the 50× stock of INS018_055 was added to the assay well, and then a reaction mix containing enzyme and substrate was added to the well. ATP at the selected concentration initiated the reaction. Data were processed by custom-built in-house analysis software at Eurofins. Results are presented as percentages of kinase activities remaining in comparison to the DMSO control. The following formula was used to calculate these values:

(Mean of sample counts − mean of blank counts)/(mean of control counts).

IC_{50} values were calculated using XLfit version 5.3 (ID Business Solutions). Sigmoidal dose–response (variable slope) curves were fit based on the mean result for each test concentration. Non-linear regression analysis was applied. When the top or bottom of the curve falls >10% outside of 100 or 0, either or both of these limits may be constrained at 100 and 0 when the QC criterion on R^2 is met.

Measurement of FMT and EMT on human primary cells by high-content analysis. For the measurement of FMT by α-SMA, on the first day, human lung-derived primary lung fibroblast cells were seeded. After 2 d, the medium of the cells was refreshed. On day 5, cells from donors with IPF or healthy donors were prepared in two sets and exposed to INS018_055. After 1 h, all cells were treated with 1.25 ng ml^{-1} TGF-β1. On the eighth day (72 h after triggering), cells were fixed with 4% formaldehyde, stained for α-SMA and with DAPI and then imaged and quantified via high-content analysis (HCA) (IN Cell Analyzer 2200, GE Healthcare).

For the measurement of EMT by FN1, on the first day, human primary bronchial epithelial cells were seeded. After 2 d, the medium of the cells was refreshed. On day 6, cells from donors with IPF or healthy donors were prepared in two sets and exposed to INS018_055. After 1 h, all cells were treated with 5 ng ml^{-1} TGF-β1. On the ninth day (72 h after triggering), cells were fixed with 4% formaldehyde, stained for FN1 and with DAPI and then imaged and quantified via HCA (IN Cell Analyzer 2200, GE Healthcare).

Analysis of α-SMA and FN1. Segmentation and quantification of α-SMA and FN1 immunoreactivity was carried out with an HCA algorithm, with density × area ($D \times A$) output. Data normalization of raw α-SMA or FN1 ($D \times A$) to percent inhibition values was performed on a plate-to-plate basis.

Percent inhibition = $(100 - (\mu_p - X_i)/(\mu_p - X_n)) \times 100$

μ_p is the average α-SMA or FN1 value of the positive control (TGF-β1 + 1 µM SB525334). μ_n is the average α-SMA or FN1 value of the vehicle control (TGF-β1 + 0.1% DMSO). X_i is the compound α-SMA or FN1 value. IC_{50} values for all compounds (if calculable, based on the point of inflection) were calculated with GraphPad Prism using non-linear fit of log (inhibitor) versus response (four parameter).

Analysis of percent remaining cells. DAPI fluorescence was applied for HCA-based quantification of the number of imaged cells on a plate-to-plate basis.

Percent remaining cells = $(X_i/\mu_n) \times 100\%$

μ_n is the average number of nuclei of the vehicle control (TGF-β1 + 0.1% DMSO). X_i is the compound number of nuclei.

INS018_055 treatment of A549 cells in the TGF-β- and TNF-α-induced EMT assay. A549 cells were cultured in RPMI 1640 medium supplemented with 10% FBS and antibiotics (penicillin and streptomycin, 100 U ml^{-1} each) at 37 °C in a humidified atmosphere of 95% air and 5% CO_2. A549 cells were seeded into six-well plates at 4×10^5 cells per well. After overnight culture, the medium was replaced with serum-free medium for 24 h of starvation before cellular induction with TGF-β alone at a final concentration of 5 ng ml^{-1} or with TGF-β and TNF-α at a final concentration of 20 ng ml^{-1} in combination with INS018_055 or DMSO for an additional 48 h, respectively. At the end of induction and compound treatment, whole-cell lysates were prepared using RIPA buffer (Sigma, R0278) with protease inhibitor or nuclear and cytoplasm fractions were extracted using the Subcellular Protein Fractionation Kit (Thermo Fisher Scientific).

Knockdown of *TNIK* by short hairpin RNA in A549 cells. TNIK short hairpin RNA sequences and vector information. TNIK shRNA[79] was cloned into the vector PLKO.1-puro by a CRO (GENEWIZ). Packaging plasmids (psPAX2, PCMV-VSV-G) were purchased from Addgene. Sequences were as follows: shRNA control (F, 5′-CCGGCAACAAGATGAAGAGCACCAACTCGAGTTGGTGCTCTTCATCTTGTTGTTTTT-3′; R, 5′-AATTAAAAACAACAAGATGAAGAGCACCAACTCGAGTTGGTGCTCTTCATCTTGTTG-3′), shTNIK-1 (F, 5′-CCGGGCTCCTAAACCGTATCATAAACTCGAGTTTATGATACGGTTTAGGAGCTTTTT-3′; R, 5′-AATTAAAAAGCTCCTAAACCGTATCATAAACTCGAGTTTATGATACGGTTTAGGAGC-3′), shTNIK-4 (F, 5′-CCGGGGGCAAGGCAAAGTCTATAATCTCGAGATTATAGACTTTGCCTTGCCCTTTTTG-3′; R, 5′-AATTCAAAAAGGGCAAGGCAAAGTCTATAATCTCGAGATTATAGACTTTGCCTTGCCC-3′).

Lentivirus packaging. Lentivirus were produced in 293T/17 cells transfected with PLKO.1-puro-shRNA (shRNA control, shTNIK-1 and shTNIK-4) using the Lipofectamine 3000 transfection kit. Next, 293T/17 cells were seeded in DMEM complete culture medium in a 10-cm tissue culture dish. Cells were incubated at 37 °C with 5% CO_2 overnight. Mixture I was composed of 5 μg PLKO.1-puro-shRNA, 7.5 μg psPAX2, 2.5 μg PCMV-VSV-G, 30 μl P3000 reagent and 1,500 μl Opti-MEM. Mixture II was composed of 45 μl Lipofectamine 3000 and 1,500 μl Opti-MEM.

After a 5-min incubation at room temperature, mixture I and mixture II were mixed and incubated for another 20 min at room temperature to allow DNA–Lipofectamine 3000 complexes to form. DNA–Lipofectamine 3000 complexes were added to the 293T/17 culture plate and incubated overnight. The next day, medium containing the DNA–Lipofectamine 3000 complexes was removed and replaced with fresh culture medium. Cells were incubated at 37 °C for another 72 h in a CO_2 incubator.

At 72 h after transfection, virus-containing supernatants were collected and filtered with a 0.45-μm filter. Viral stocks were aliquoted and stored at −80 °C.

Infection. RPMI 1640 medium supplement with 10% FBS and 1% penicillin–streptomycin was used for A549 culture medium. A549 cells were seeded at 2×10^6 cells per plate in culture medium in a 10-cm dish and incubated with 5% CO_2 at 37 °C overnight. The next day, the complete culture medium was replaced with fresh medium containing different virus ratios of 1/10 and 1/20. Virus at each ratio was used to treat cells in each dish. A total of 10 ml medium containing virus was added to each dish. Polybrene was added to each dish at a concentration of 6 μg ml^{-1}, and dishes were incubated overnight. The next day, the medium was replaced with fresh culture medium, and the cells were incubated for another 48 h.

TGF-β treatment. After 2 d of incubation, the infected A549 cells were split into six-well plates at a density of 4.0×10^5 cells per well in culture medium and cultured overnight. The next day, the medium was substituted with a serum-free medium for 24 h of starvation. A549 cells were treated in serum-free medium with or without 5 ng ml^{-1} TGF-β in duplicate for 48 h at 37 °C after starvation. At the end of induction, cells were lysed in RIPA buffer supplemented with protease inhibitor and Phosphatase Inhibitor Cocktails 2 and 3 (Sigma). The supernatant was stored at −80 °C after centrifugation. The total amount of protein in the lysate was determined using the BCA Protein Assay Kit. In experiment I, cell morphology was captured with a ×4 objective lens after induction and finally was magnified at ×40.

INS018_055 treatment of MRC-5 cells. MRC-5 cells were cultured in a monolayer with DMEM high-glucose medium supplemented with 10% FBS at 37 °C with 5% CO_2. Cells were seeded in six-well plates in three replicate wells per condition at a density of 2.5×10^5 cells per well and cultured for 24 h. The medium was replaced with serum-free medium, and cells were starved for 4 h. Afterward, the compound at eight concentrations (~100 μM–1.28 nM, fivefold gradient dilution) was added into each well. Vehicle (2% FBS–DMEM, containing 0.1% DMSO) was added into the blank control group and the model group, and pre-incubation was continued for 2 h. Next, TGF-β1 was added into each well of the administration group and the model group (final concentration in each well was 2 ng ml^{-1}), and the same volume of vehicle was added into each well of the blank control group for 72 h of incubation (the culture medium was changed once after 48 h of incubation). After incubation, the cells were collected and digested with 0.5% trypsin, protein was extracted and total protein content was determined by the BCA method.

INS018_055 treatment of HK-2 cells. INS018-055 stock was diluted to eight concentrations in a threefold dilution with DMSO from the final highest concentration at 10 μM. A total of 3.0×10^5 HK-2 cells per well were seeded into a six-well plate containing 2.0 ml culture medium at a density of 3.0×10^5 cells per well. The cells were cultured overnight at 37 °C. The next day, the complete medium was replaced with K-SFM medium supplemented with 0.025 mg ml^{-1} bovine pituitary extract, 2.5 ng ml^{-1} EGF and 1% penicillin–streptomycin for starvation. After 24 h of incubation, HK-2 cells were treated with compounds (eight concentrations, threefold dilution) for 30 min before being stimulated with 8 ng ml^{-1} TGF-β for an additional 48 h. TGF-β induction was used as the maximum induction positive control. The final DMSO concentration in the assay was 0.1%. Cells were cultured for 48 h at 37 °C. The culture medium was discarded at the end of induction. Cells were washed once using ice-cold DPBS. RIPA buffer was added to lyse the cells for 20 min. Cells were collected, and the supernatant was stored at −80 °C. The total amount of protein in the lysate was determined using the BCA Protein Assay Kit.

Protein extraction and western blot. For whole-cell lysates, cells were collected and lysed on ice for 20 min with RIPA buffer with protease and phosphatase inhibitors (cOmplete and PhosSTOP, Roche). The protein concentration of the cell lysates was determined using the BCA Protein Assay Kit. After denaturation, whole-cell lysates and cytoplasm and nuclear extraction fractions were loaded onto 4–12% gradient Bis-Tris gels (15 wells) and 4–20% Tris-glycine gels (26 wells), respectively. For western blot detection of target proteins, 15 μg of whole-cell lysates, 5–10 μg of the cytoplasm fraction, 1–2 μg of nuclear matrix and 2–4 μg of the chromatin fraction were loaded into each lane. For western blot detection of histone 3, 0.4–1 μg of the chromatin fraction was loaded. The gels were run for 0.5 h at 80 V before being run at 120 V for another 1 h. When electrophoresis was completed, the gels containing target proteins were transferred onto NC membrane using the iBlot 2 Gel Transfer Device for 13 min at 20 V. All membranes were blocked with 5% BSA in TBST buffer at room temperature for 1 h and then incubated with primary antibodies specific to the indicated targets in TBST buffer containing 5% BSA at 4 °C overnight. After the primary antibody,

membranes were washed three times with TBST buffer before incubation with HRP-conjugated secondary antibodies at room temperature for 1 h. Membranes were stripped with Restore Western Blot Stripping Buffer (Thermo Fisher) for GAPDH reference protein detection after exposure for the target proteins. Images of blots were acquired using Image Quant LAS 4000. Integrated intensity of bands from 16-bit blot images was used for quantitation with the software Image Studio Lite. Protein-quantification data of two independent experiments were calculated and analyzed using GraphPad.

α-SMA evaluation in normal human dermal fibroblasts. NHDF cells were seeded in a 96-well plate and cultured for 24 h. The medium was then replaced with an assay medium containing or lacking (control conditions) the test compound, and TGF-β (0.1 ng ml^{-1}) was added or not (unstimulated control). The cells were then incubated for 72 h. All experimental conditions were performed in five replicates. For in situ immunolabeling and image analysis, at the end of the incubation, fibroblasts were rinsed, fixed and permeabilized. The cells were then labeled using a primary antibody (anti-α-SMA), followed by the corresponding fluorescent secondary antibody (GAM Alexa 488). In parallel, cell nuclei were stained using Hoechst solution 33258 (bisbenzimide). Image acquisition was performed using a high-resolution imaging system, the IN Cell Analyzer 2200 (GE Healthcare) automated microscope (objective lens, ×20). Five pictures per well were taken. Labeling was quantified by measuring the fluorescence intensity of α-SMA signals normalized to the total number of nuclei identified (integration of numerical data was performed with the Developer Toolbox 1.5, GE Healthcare software).

Fibronectin and procollagen I evaluation in NHDFs. NHDF cells were seeded in 96-well plates and cultured for 24 h. The medium was then replaced with assay medium containing or not containing (control conditions) the test compounds, and TGF-β (10 ng ml^{-1}) was added or not (unstimulated control). The cells were then incubated for 72 h. Fibronectin and procollagen I contents were measured in the culture supernatants using specific ELISA kits (Takara, MK115 for fibronectin and MK101 for procollagen type I) according to the supplier's instructions.

Cell viability. MRC-5 cells were seeded at 8,000 cells per well in 10% FBS medium and cultured with concentrations diluted from 100 µM for 72 h at 37 °C with 5% CO_2. A549 cells were seeded at 5,000 cells per well in 10% FBS medium and cultured with concentrations diluted from 50 µM for 72 h at 37 °C with 5% CO_2. HK-2 cells were seeded at 8,000 cells per well in 10% FBS medium and cultured with concentrations diluted from 50 µM for 72 h at 37 °C with 5% CO_2.

CellTiter-Glo (Promega) was added to all conditions followed by 15 min of incubation at room temperature. Luminescent signal was detected with PerkinElmer EnVision. Data analysis was performed by calculating cell viability (%) with luminescent signal: cell viability (%) = (luminescent signal of sample − mean luminescent signal of medium)/(mean luminescent signal of cell control − mean luminescent signal of medium) × 100. CC_{50} values were determined according to the fitting curve of the emission ratio versus log (compound concentration) with GraphPad Prism.

Animal studies

Mice. For the bleomycin-induced lung fibrosis model and the LPS-induced acute lung injury model, C57BL/6 mice (male, 8 weeks old, ~23–26 g) were purchased from Jiangsu Gempharmatech. Animals were acclimatized for 1 week before the experiment. All in vivo experimental procedures were approved by the IACUC. All euthanasia was performed using CO_2 inhalation, and all efforts were made to minimize animal suffering.

For the UUO study, 7-week-old female C57BL/6 mice were obtained from Japan SLC. Animals were housed and fed a normal diet (CE-2, CLEA Japan) under controlled conditions. All animals used in the study were housed and cared for following the Japanese Pharmacological Society Guidelines for Animal Use.

Animal models

Preparation for dosing formulations for animal model studies. For bleomycin formulation for lung fibrosis, bleomycin was dissolved in saline and vortexed to obtain a clear solution. For bleomycin formulation for the scarring model, bleomycin was dissolved in PBS. For LPS formulation, LPS was dissolved in saline and vortexed to obtain a clear solution. INS018_055 (p.o.) treatment solutions were formulated in 0.5% methylcellulose, and formulations were prepared daily. INS018_055 (topical) was formulated in 15% DMSO, 41% Transcutol, 10% IPM, 8% glycerol, 8% oleic acid, 8% diisopropyl adipate, 3% Brij L4, 4% cetostearyl alcohol and 2% hydroxypropyl cellulose (wt/wt%). Nintedanib treatment solutions were formulated in 0.5% methylcellulose. Homogeneous SB525334 suspensions were prepared in 5% DMSO + 95% (0.5% HPMC and 0.1% Tween-80 in reverse-osmosis (RO) water), and formulations were freshly prepared before administration. Dexamethasone was formulated in 7.70 ml saline and kept on ice before use. Rapamycin was prepared in 0.2% sodium carboxymethyl cellulose and 0.25% polysorbate 80.

LPS-induced acute lung injury model. One day before the first experimental day, animals were grouped. All mice were anesthetized with Zoletil–xylazine (50 mpk and 10 mpk, i.p.). At 0 h, mice received LPS at a dose of 50 µl as 50 µg per mouse (groups 2–5, $n = 8$ per group) or 50 µl saline (group 1, $n = 8$ per group) by intratracheal administration. In the INS018_055 groups (groups 3 and 4), animals were treated with INS018_055 for 4 h after the LPS challenge. In the dexamethasone group (group 5), animals were treated with dexamethasone both 2 h before and 4 h after the LPS challenge. Twenty-four hours after the LPS challenge, animals were euthanized, and BALF samples were collected.

Bleomycin lung mouse model. Mice received bleomycin on day 1 at a dose of 0.66 mg per kg (equivalent to 1 U per kg), in a volume of 50 µl by intratracheal administration, or received saline (sham group). On day 7, dosing with the corresponding compounds was started. The vehicle group (group 1) was treated with 0.5% methylcellulose BID. Groups 2–4 were treated with INS018_055 at dosages of 3 mg per kg, 10 mg per kg and 30 mg per kg BID, respectively. Group 5 was treated with 60 mg per kg nintedanib QD. On days 5 and 21, lung function measurement was performed. On day 28, all animals were killed, and blood and tissue samples were collected along with BALF. The method of model development and study design of combination studies of INS018_055 on bleomycin-induced lung fibrosis mouse models were similar.

Lung function measures in the bleomycin-induced lung fibrosis mouse model. Lung function was measured according to the HDB standard protocol. Expiratory time, relaxation time, peak expiratory flow and peak inspiratory flow were collected and used to calculate Penh data.

Inhalation administration bleomycin-induced lung fibrosis rat model. On day 1 of the study, bleomycin hydrochloride was administered to 8–9-week-old Sprague Dawley rats at a final concentration of 1.5 mg ml^{-1} in sodium chloride solution (dosed volume at 1.0 ml per kg) by intratracheal atomization (once in the morning and once in the afternoon) to build the pulmonary fibrosis model. INS018_055 inhalation solutions were prepared at 0.1 mg ml^{-1}, 0.3 mg ml^{-1}, 1.0 mg ml^{-1} and 6.0 mg ml^{-1} in sodium citrate buffer.

Aerosol generation and environmental conditions. A German PARI compression atomizer (PARI TurboBOY) was set at an aerosol flow rate of 9–10 l min^{-1}. Aerosols from INS018_055 solutions were generated in single-chamber mode. The dilution flow of aerosol was set at 0 l min^{-1},

and the pumping-out flow from the chamber was set at 6 l min^{-1}. The fluid rate was set at 0.5 ml min^{-1} for test article replenishment. Before nebulization, aerosol-generation conditions were validated three times.

In this experiment, temperature was set at 20–26 °C, humidity between 30% and 80%, O_2 level ≥ 19% and CO_2 level ≤ 1%. And these parameters were recorded in real time during the experiment.

Aerosols of INS018_055 solutions were sampled and analyzed for concentrations and sizes of aerosol particles. For concentration assessment, a glass fiber filter membrane (φ = 37 mm) was used to measure aerosol concentrations during drug administration from approximately day 8 to day 28. Sampling was taken at 5 min (±1 min) after drug administration from a random exposure port. Sampling flow for the solution of 0.1 mg ml^{-1} INS018_055 was 1 l min^{-1} for 10 min. Sampling flow for solutions of 0.3 mg ml^{-1}, 1.0 mg ml^{-1} and 6.0 mg ml^{-1} INS018_055 was 1 l min^{-1} for 5 min. For particle size assessment, during administration on day 8, day 15, day 22 and day 28, an exposure port was selected randomly and connected with a next-generation impactor for sampling (MOC disc filter paper) for aerosol particle size measurement at a sampling flow rate of 15 l min^{-1} for 5 min. Aerosol particle size parameters were calculated with the system software: median mass aerodynamic diameter (MMAD), geometric standard deviation, fine particle fraction (percentage of fine particles with MMAD < 5 μm), etc. The actual delivered dose and parameters of aerosol particle size distribution are:

Index		0.1 mg ml^{-1}	0.3 mg ml^{-1}	1.0 mg ml^{-1}	6.0 mg ml^{-1}
Actual delivered dose (mg/kg)[a]		0.040 ± 0.008	0.136 ± 0.017	0.485 ± 0.081	2.575 ± 0.242
NGI	MMAD (μm)	2.747 ± 0.317	2.580 ± 0.077	2.435 ± 0.053	2.5175 ± 0.167
	GSD	1.314 ± 0.111	1.760 ± 0.036	1.856 ± 0.121	1.957 ± 0.066
	FPF (%)	99.935 ± 0.007	85.454 ± 4.675	80.866 ± 5.827	80.294 ± 4.187

[a]Actual delivered dose was calculated as follows:

$$\text{Dose (mg per kg)} = \frac{C(\mu g\ l^{-1}) \times RMV(l\ min^{-1}) \times D(min) \times DF}{BW\ (kg) \times 1{,}000}$$

. *C*, concentration (μg l^{-1}) in air inhaled; RMV, respiratory minute volume (l min^{-1}); RMV (l min^{-1}) = 0.608 × BW $(kg)^{0.852}$; *D*, duration of exposure (min); DF, deposition fraction, assumed as 100% for calculation of the actual delivered dose; BW, body weight (kg); FPF, fine particle fraction; GSD, geometric standard deviation; NGI, next-generation impactor.

Measurement of lung function in the lung fibrosis rat model. On day 29 and before euthanasia, animals were anesthetized with chloral hydrate (90 mg ml^{-1}, 5 ml per kg, i.p.). Necks were cut longitudinally at the middle to isolate trachea, and a catheter was intubated into the trachea toward the tail direction and fixed with a suture thread. The anesthetized animals were placed into the plethysmography chamber, to which the airway was connected. The AniRes2005 animal lung function-analysis system was used to assess parameters including FVC, airway resistance and pulmonary compliance. The ventilator respiration rate was set to 65 beats per minute; the respiration ratio was set to 20:10; the negative pressure controller was set to 30 cm H_2O. FVC detection was set to pressure-control mode, while pressure was set to 30 cm H_2O. The startup mode was selected to automatically detect at the end of breath, and the 'Start' button was clicked to passively inhale and then turned to exhale after reaching the set pressure value. Next, the 'Stop' button was clicked to complete an FVC detection. Each animal was evaluated for FVC at least five times. The final result was the mean value after the maximum and minimum data were excluded for analysis.

Unilateral urethral obstruction model in mice. The surgery took place on two separate days. Mice were divided into two slots based on their body weight before the day of the surgery. On day 0 (day of surgery), UUO surgery was performed after administration of three types of mixed anesthetic agents (medetomidine, midazolam, butorphanol). Mice were shaved at the incision site, the abdomen was cut open, and the left ureter was exteriorized and ligated with 4-0 silk sutures at two points. Finally, the peritoneum and the skin were closed with sutures.

Group 1 was sham-operated mice, and group 2 was treated with vehicle BID. Groups 3–5 were treated with INS018_055 at a dose of 3 mg per kg, 10 mg per kg and 30 mg per kg BID, respectively. Group 6 was treated with the ALK5 inhibitor SB525334 at a dose of 100 mg per kg QD. Treatments were applied from day 0 to day 14. And animals were killed on day 14, and blood and tissue samples were collected.

Bleomycin-induced scarring model in rats. Animals were anesthetized using isoflurane. Using a 1-ml syringe containing a 27-gauge needle, 100 μl bleomycin (1 mg ml^{-1} in PBS) solution was injected subcutaneously into two sites on the shaved dorsal regions, QD for 4 weeks. Rats in the naive control group were injected daily with an equivalent volume of sterile PBS. Treatments, vehicle (topical), rapamycin (1.5 mg per kg, i.p.) or INS018_055 (0.05%, 0.15% or 0.45%, topical) were administered 30 min before bleomycin administration, QD for 4 weeks. After 4 weeks of treatment, animals were killed with CO_2, and skin was collected.

Bronchoalveolar lavage fluid collection. Mice were first anesthetized with Zoletil–xylazine (50 mpk and 10 mpk, i.p.). Next, a 1.5–2-cm longitudinal incision was made on the ventral side of the neck to expose the trachea by blunt dissection, and curved forceps were placed underneath the trachea. A 10–12-cm piece of suture was then threaded underneath the trachea using forceps. The ends of the suture were then pulled cranially to 'stretch' the trachea toward the surgeon. A 20-G needle was then used to poke a hole in the trachea as close to the larynx as possible. A 22-G blunt stainless steel needle connected to PE tubing was placed in the hole and inserted approximately 5–8 mm into the trachea. The cannula was secured in the trachea by tying the suture around it with an overhand knot. A 1-ml syringe containing 0.4 ml saline was gently pushed forward and backward three times to recover BALF fluid. BALF was transferred into empty tubes and stored on ice. This process was repeated twice with 0.3 ml saline for a total volume of 1 ml.

BALF cytological analysis. Total cell numbers were counted with a standard hemocytometer by mixing 20 μl BALF with 20 μl trypan blue. Cytospins were prepared by cytocentrifugation using Cytospins at 106*g* for 5 min, and then smears of BALF cells were stained with Diff-Quick stain. Cell differentiation was performed by counting at least 400 cells using standard hemocytologic criteria to classify them as alveolar macrophages or monocytes, neutrophils, eosinophils or lymphocytes. The rest of the BALF fluid was centrifuged, and supernatants were collected and stored at −80 °C for cytokine, chemokine and total protein analysis.

BALF cytokine and chemokine analysis. From the collected BALF, concentrations of cytokines were measured from mice in groups 2–7 killed on day 28. Mouse BALF cytokines and chemokines were measured using the ELISA Analysis Kit (Sinobest Bio). Mouse BALF soluble collagen was measured using the Soluble Collagen kit (Biocolor). Mouse lung hydroxyproline content was measured using the Hydroxyproline Assay Kit (Njjcbio).

Kidney hydroxyproline content. Frozen kidney samples were processed by an alkaline acid hydrolysis method as follows. Kidney samples were dissolved in 2 N NaOH at 65 °C and autoclaved at 121 °C for 20 min. The lysed samples (400 μl) were acid hydrolyzed with 400 μl of 6 N HCl at 121 °C for 20 min and neutralized with 400 μl of 4 N NaOH containing 10 mg ml^{-1} activated carbon. Next, AC buffer (2.2 M acetic acid and 0.48 M citric acid, 400 μl) was added to the samples, followed by centrifugation to collect the supernatant. The prepared samples and standards (serial dilutions of *trans*-4-hydroxy-L-proline (Sigma-Aldrich)

starting at 16 μg ml^{-1}, each 400 μl) were mixed with 400 μl chloramine T solution (Nacalai Tesque) and incubated for 25 min at room temperature. The samples were then mixed with Ehrlich's solution (400 μl) and heated at 65 °C for 20 min to develop color. After samples were cooled on ice and centrifuged to remove precipitates, the optical density of each supernatant was measured at 560 nm. Hydroxyproline concentrations were calculated from the hydroxyproline standard curve. Kidney hydroxyproline contents were expressed as μg per mg protein, and the amount of protein was determined using the BCA Protein Assay Kit (Thermo Fisher Scientific).

Hematoxylin and eosin staining and quantification (lung). Left lung lobes were removed from each mouse and fixed in 10% NBF solution, embedded in paraffin and processed to obtain 4-μm sections for staining with H&E. H&E-stained slides were scanned using an Aperio ScanScope Model CS2 (Leica) at ×200 magnification. Images were then opened in the HALO Plus 5 workstation (version 2.3, Indica Labs) program. Using the pen annotation tool, the whole-lung section was selected as an annotation layer. The area of vessels was defined as background using the exclusion drawing tool. The total inflammation area was quantified by selecting the dark blue inflammation cells using the pen annotation tool manually. The percentage of H&E staining (inflammation area/total area of the lung) in the selected annotation was then calculated using the program.

Masson trichrome staining and quantification (lung). Masson's trichrome staining was conducted using a ready-to-use kit (Trichrome Stain (Masson) Kit, HT15, Sigma-Aldrich) as described by the manufacturer.

Lung sections were cut at a thickness of 4 μm, dried in an oven for 1 h and stained with M&T using our standard protocol. Briefly, sections were stained with Weigert's iron hematoxylin working solution for 10 min, stained in Biebrich scarlet–acid fuchsin solution for 10 min, differentiated in phosphomolybdic–phosphotungstic acid solution for 5 min or until the collagen was not red, transferred to aniline blue solution, stained for 1 min and then dedifferentiated in 1% acetic acid solution. Next, sections were dehydrated and coverslipped for subsequent image analysis. For image analysis of collagen deposition, Masson's trichrome-stained slides were scanned by using Aperio ScanScope Model CS2 (Leica) at ×200 magnification. Images were then opened with HALO. Using the pen tool, the whole left lung section was selected as an annotation layer. The area occupied by blue collagen fibers was measured using the pen tool. The percentage of fibrosis (positive areas) in the selected annotation was then calculated with the program. Fibrosis was expressed as a percentage per lung section.

Sirius red staining and quantification (kidney). At the end of the UUO study, the left kidney was fixed in Bouin's solution and embedded in paraffin. Kidney sections were stained using picro-Sirius red solution (Waldeck). To quantify the interstitial fibrosis area, brightfield images in the corticomedullary region were captured using a digital camera (DFC295) at 200-fold magnification, and the positive areas in five fields per section were measured using ImageJ software (National Institute of Health).

Immunohistochemistry staining and quantification of lung tissue. Sections (4 μm thick) were placed on slides, and, after overnight drying, paraffin was removed with xylene. Next, sections were placed in a graded ethanol series and immersed in distilled water. After heat-induced citrate antigen (pH 6.0) unmasking, sections were immersed in 3% hydrogen peroxide solution for 5 min. The sections were then incubated in blocking serum (Dako, X0909) for 15 min at room temperature, followed by using primary rabbit antibodies (anti-α-SMA and anti-collagen I) for 1 h. Next, secondary antibodies conjugated to HRP were added. For image analysis of fibrosis, stained sections were used and scanned with the Aperio CS2 Scanner machine. Images were then opened with HALO. Using the pen tool, the whole left lung section was selected as an annotation layer. The bronchus was excluded in the annotation layer. The area occupied by collagen fibers was measured using the 'Area Quantification v2.1.3' module. The percentage of positive areas in the selected annotation was then calculated using the program.

Immunohistochemistry staining and quantification of kidney tissue. Paraffin sections were deparaffinized and hydrophilized with xylene, 100–70% alcohol series and RO water, and then circles were drawn around the kidney sections. Endogenous peroxidase activity was blocked using 0.3% H_2O_2 for 5 min, and antigen retrieval was performed using antigen retrieval solution H (citrate buffer) at 121 °C for 10 min. After washing with PBS, kidney sections were treated with PBS with Tween-20 (PBST), followed by incubation with Block Ace (DS Pharma Biomedical) at room temperature for 10 min. The sections were incubated with the primary antibody (anti-collagen I) at 4 °C overnight. After washing with PBS, kidney sections were treated with PBST, followed by incubation with secondary antibody at room temperature for 30 min. After washing with PBS, kidney sections were fixed with 1% glutaraldehyde solution at room temperature for 1 min. After washing with RO water, kidney sections were treated with PBST and then colored using a chromogenic substrate (Simple Stain DAB, Nichirei Bioscience). After washing with RO water, kidney sections were immersed in an eightfold-diluted hematoxylin solution for 1 min and washed with RO water immediately. The stained sections were placed in running water for 15 min and sealed with Aquatex (Merck). For scoring of IHC analyses, brightfield images were captured using a digital camera (DFC295) at 200-fold magnification, and the score in one field per section was determined.

Evaluation of hydroxyproline and collagen content in skin. Hydroxyproline and collagen contents were measured in skin lysates using rat hydroxyproline (KinesisDx, K11-0512) and collagen (Abcam, ab222942) ELISA kits. Hydroxyproline and collagen concentrations were normalized to total protein content using the Bradford method.

Statistical analysis

For western blot figures, P values between groups were analyzed by Welch's t-test among three independent experiments. P values < 0.05 were considered statistically significant. For LPS and bleomycin mouse models, statistical analysis was performed using ordinary one-way-ANOVA, and post hoc Šídák's multiple-comparison test was performed to calculate statistical analysis between groups. The difference was considered significant when $P < 0.05$. For the inhalation study in the bleomycin rat model, in analyzing parameters of lung functions (FVC, airway resistance and pulmonary compliance), statistical analysis was performed using uncorrected Fisher's least-significant difference test as post hoc analysis after ANOVA analysis; in analyzing pathology results, statistical analysis was performed using Kruskal–Wallis test and Dunn's multiple-comparison test as post hoc analysis after ANOVA analysis. Differences were considered significant when $P < 0.05$. For the UUO model, statistical analyses were performed using the Bonferroni multiple-comparison test. P values < 0.05 were considered statistically significant. For the bleomycin-induced skin scarring experiment, statistical analysis was performed using one-way ANOVA, and then Dunnett's multiple-comparison test was performed. P values < 0.05 were considered statistically significant. For Fig. 2h, enrichment analysis was performed using the gseapy.enrichr Python package. P values were computed using Fisher's exact test. This is a binomial proportion test that assumes a binomial distribution and independence for probability of any gene belonging to any set. Adjusted P values (q values) were calculated using the Benjamini–Hochberg method for correction for multiple-hypothesis testing.

Single-cell RNA sequencing analysis

Single-cell RNA sequencing (scRNA-seq) data from 32 IPF and 28 control lungs were collected from the GEO database (GEO accession GSE136831)[53]. Data were preprocessed with a standard pipeline using the Scanpy package. Single-cell data was filtered by applying two conditions: including cells with a minimum of 200 genes and with genes expressed in more than three cells. Cells with mitochondrial fraction greater than 20% were excluded. Data were initially normalized with a scale factor of 10,000 and log transformed. PCA was run with the sc.pp.pca function and 'n_comps' = 50. Batch effects were corrected with 'sce.pp.harmony_integrate' function. A neighborhood graph was computed on the first 50 principal components derived after batch correction, and results were visualized using UMAP. Cell type annotation was obtained from the original article. Cluster gene signatures were identified by testing for differential expression of a subgroup against all other cells using a Wilcoxon rank-sum with the 'tl.rank_genes_groups' function. Differential expression between IPF and controls for each cell type was calculated using the 'tl.rank_genes_groups function' (method = 'Wilcoxon'), and *TNIK* expression was plotted (non zero values with *P* values, Benjamini–Hochberg adjustment and a threshold of 0.05) using the plotly package.

Simulated knockout profile of *TNIK*

scTenifoldKnk[54] is a method developed to perform virtual knockout experiments to predict gene functions. First, scTenifoldKnk, based on scRNA-seq data, constructs a denoised single-cell gene regulatory network (scGRN). The scGRN is copied, and then outward edges of the target gene in the adjacency matrix of the copied scGRN are zeroed out, thus creating a pseudo-knockout scGRN. Having two scGRNs (one initial net and a second pseudo-knockout net), target gene regulatory significance can be estimated with a manifold alignment procedure. Two scGRNs are mapped to the same low-dimensional space, and distance between gene projections shows the impact of gene knockout in the scGRN: larger perturbation of genes in low-dimensional space indicates the importance of the target gene in the scGRN. We used IPF myofibroblast scRNA-seq data derived from GSE136831 (ref. 53) in scTenifoldKnk and made a pseudo-knockout of *TNIK*. The list of perturbed genes was sorted by fold changes of distances between gene projections of two scGRNs; probabilities were assigned using χ^2 distribution with one degree of freedom. The most perturbed genes are supposed to have a tight connection with the target gene. Next, we applied the MCODE algorithm[47] for *TNIK* and the top 100 most perturbed genes and performed pathway and process enrichment analysis.

Chemistry42

The Molecular Sets platform (MOSES)[87] was used to train and benchmark the generative chemistry models in Chemistry42. The structure-based drug-design workflow implemented in the Chemistry42 platform was used to generate a library of virtual structures[40]. Pocket and Pharmacophore Reward modules were used for scoring produced designs and navigating the generative process. The TNIK ATP-binding site was selected as a target-binding pocket. The generated structures had to match a two-point pharmacophore hypothesis that consisted of the hydrogen bond acceptor forming an H-bond with the NH of Cys108 of the hinge region and hydrophobic center occupying the space near the gatekeeper Met105. In addition to the assessment by Pocket and Pharmacophore modules, the engine penalized structures that violated the predefined ranges of the physicochemical properties (logP, molecular weight, number of hydrogen bond donors, hydrogen bond acceptor, topological polar surface area, number of atoms, number of rotatable bonds, number of aromatic rings), medicinal chemistry filters and the synthetic accessibility score threshold.

Phase 0 study (Australia)

A phase 0 micro dosing clinical trial was conducted in Australia (ACTRN12621001541897) and can be found at https://www.anzctr.org.au/TrialSearch.aspx. Complete information on clinical trial registration, study protocol, data collection and outcomes is provided in the Reporting summary as well as in Supplementary Information 8. This includes statistical considerations, study design, patient-selection criteria, procedures, outcomes and PK analysis.

Phase I study (New Zealand)

The general design of this clinical trial (NCT05154240) can be found at https://www.clinicaltrials.gov. The randomized, double-blind, placebo-controlled study of INS018_055 was conducted from 21 February 2022 (first participant administered first dose) until 30 September 2022 (last participant contacted). Complete information on clinical trial registration, study protocol, data collection and outcomes is provided in the Reporting summary as well as in Supplementary Information 9. This includes statistical considerations, study design, patient-selection criteria, procedures, outcomes and PK analysis.

Phase I study (China)

Detailed information including facilities and inclusion and exclusion criteria can be found at the following link: http://www.chinadrugtrials.org.cn/clinicaltrials.prosearch.dhtml (registration number CTR20221542).

Complete information on clinical trial registration, study protocol, data collection and outcomes are provided in the Reporting summary as well as in Supplementary Information 10. This includes statistical considerations, study design, patient-selection criteria, procedures, outcomes and PK analysis.

Full clinical study protocols for all three trials are provided in the Supplementary Information.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Raw data from all experimental studies are available for public access through our repository, which can be found at https://insilico.com/repository/nbt-ins018-055-tnik. RNA sequencing data are available for download through the repository link along with all relevant source and raw data that were presented in this study. The molecular sets used for training are available through our open-access benchmarking platform MOSES[87]. For the current study, target ID scoring was calculated on 15 IPF datasets collected from GEO database. A list of used datasets is as follows: GSE93606, GSE38958, GSE28042 and GSE33566 derived from blood tissue and GSE101286, GSE72073, GSE150910, GSE92592, GSE52463, GSE83717, GSE21369, GSE15197, GSE99621, GSE138283 and GSE24206 derived from lung tissue. GSE136831 was used for scRNA-seq analysis (Single-cell RNA sequencing analysis). Other data that support the findings of this study, including clinical data, are available from the corresponding author upon reasonable request. Source data are provided with this paper.

Code availability

Code for benchmarking of the generative chemistry models is available in the provided reference[87]. PandaOmics and Chemistry42 are industry-grade commercial software platforms used by the industry since 2020. Both platforms are available at https://pandaomics.com and http://chemistry42.com, respectively. A detailed algorithmic workflow for running the platforms is described in the Methods and Supplementary Information. Demonstrations and trial access to the platforms are available from the corresponding author upon reasonable request.

References
87. Polykovskiy, D. et al. Molecular Sets (MOSES): a benchmarking platform for molecular generation models. *Front. Pharmacol.* **11**, 565644 (2020).

Acknowledgements
We thank J. Wang for her useful discussions about this study. We also thank the following contract research organizations for performing in vitro and in vivo experiments and clinical PK sample analysis in this study: HD Biosciences (China), WuXi AppTec (China), Charles River (Netherlands), SMC Laboratories (Japan), TheraIndx LifeSciences (India), Guangzhou Boji Medical Biotechnological (China), Bioalternatives (France), JOINN Laboratories (China), Eurofins (France), 360biolabs (Australia), Teddy Lab Services (China). We thank Nvidia for providing access to its latest GPUs throughout Insilico Medicine's history and the lifetime of this project. We acknowledge F. Pulous for preparing, assembling and developing the manuscript.

Author contributions
A.Z. and F.R. conceived the study, designed, performed and analyzed experiments, interpreted data, created figures and helped write the manuscript. A.A., H.Z., I.V.O, K.W., V.A., Y.I., Y.F., E.B., J.Q., X.L., Z.M., H.W. and A.K. intellectually contributed to experimental design for assays carried out in this study. A.A., H.Z., S.R., C. Kuppe, I.V.O, M.Z., K.W., C. Kruse, V.A., Y.I., D.P., Y.F., E.B., J.Q., X.L., Z.M., H.W., F.W.P., A.V., S.L., B.Z., V.N., X.D., A.K. and E.I. analyzed data included in this study. A.A., H.Z., C. Kuppe, I.V.O., M.Z., V.A., Z.M., P.T.A., A.V., S.L. and E.I. contributed to writing the manuscript. F.R., J.C., S.R., D.S., S.L., B.Z. and A.Z. contributed to design, data analysis and development of the clinical trials associated with this study.

Competing interests
For A.Z. and listed authors that are affiliated with Insilico Medicine: Insilico Medicine is a global clinical-stage commercial generative AI company with several hundred patents and patent applications and commercially available software. Insilico Medicine is a company developing an AI-based end-to-end integrated pipeline for drug discovery and development that is engaged in drug-discovery programs for aging, fibrosis and oncology. F.R., A.A., H.Z., S.R., I.V.O., M.Z., K.W., C. Kruse, V.A., Y.I., D.P., Y.F., E.B., J.Q., X.L., Z.M., H.W., F.W.P., A.V., S.L., B.Z., V.N., A.K. and A.Z. are affiliated with Insilico Medicine or were affiliated with Insilico Medicine when the studies were performed. No other conflicts are reported.

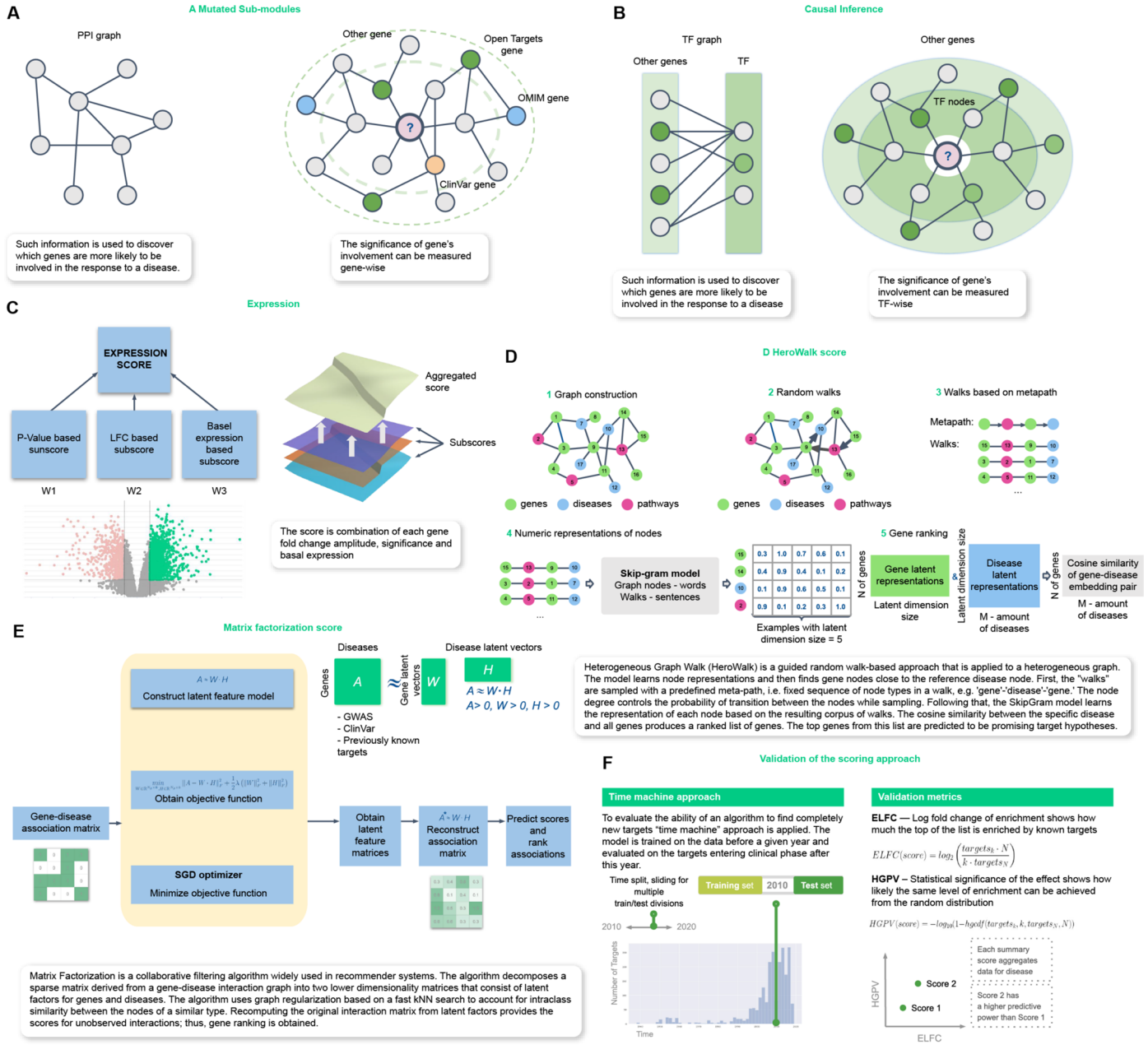
Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41587-024-02143-0.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41587-024-02143-0.

Correspondence and requests for materials should be addressed to Alex Zhavoronkov.

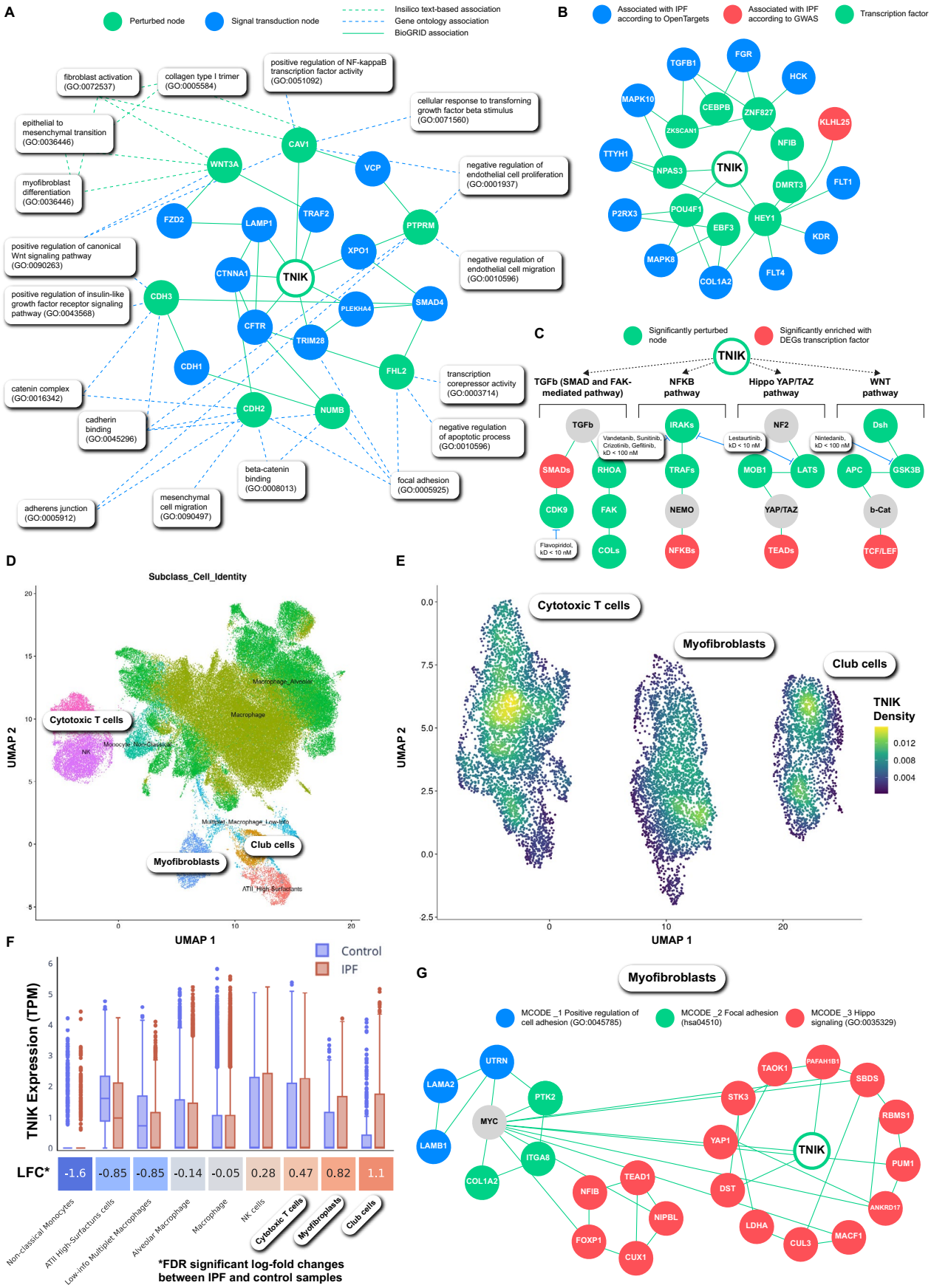
Peer review information *Nature Biotechnology* thanks the anonymous reviewers for their contribution to the peer review of this work.

https://doi.org/10.1038/s41587-024-02143-0



Extended Data Fig. 1 | Target ID Approach. (a) Mutated sub-modules score scheme. The significance of gene's involvement can be measured gene-wise. Colored nodes are genes associated with the disease of interest genes according to OMIM (green), ClinVar (orange), Open Targets (blue). **(b)** Causal inference score scheme. The significance of gene's involvement can be measured TF-wise. **(c)** Expression score scheme. The Expression score is combination of each gene fold change amplitude, significance, and basal expression. **(d)** HeroWalk score scheme. HeroWalk is a guided random walk-based approach that is applied to a heterogeneous graph. The model learns node representations and then finds gene nodes close to the reference disease node. First, the 'walks' are sampled with a predefined meta-path, that is fixed sequence of node types in a walk, for example 'gene'-'disease'-'gene.' The node degree controls the probability of transition between the nodes while sampling. Following that, the SkipGram model learns the representation of each node based on the resulting corpus of walks. The cosine similarity between the specific disease and all genes produces a ranked list of genes. The top genes from this list are predicted to be promising target hypotheses. **(e)** Matrix factorization scheme. Matrix Factorization is a collaborative filtering algorithm widely used in recommender systems. The algorithm decomposes a sparse matrix derived from a gene-disease interaction graph into two lower-dimensionality matrices that consist of latent factors for genes and diseases. The algorithm uses graph regularization based on a fast kNN search to account for intraclass similarity between similar nodes. Recomputing the original interaction matrix from latent factors provides the scores for unobserved interactions; thus, gene ranking is obtained. **(f)** Performance of the models are evaluated using two metrics ELFC and HGPV. Time Machine approach was applied to demonstrate the ability of the models to predict truly novel target hypotheses.

https://doi.org/10.1038/s41587-024-02143-0

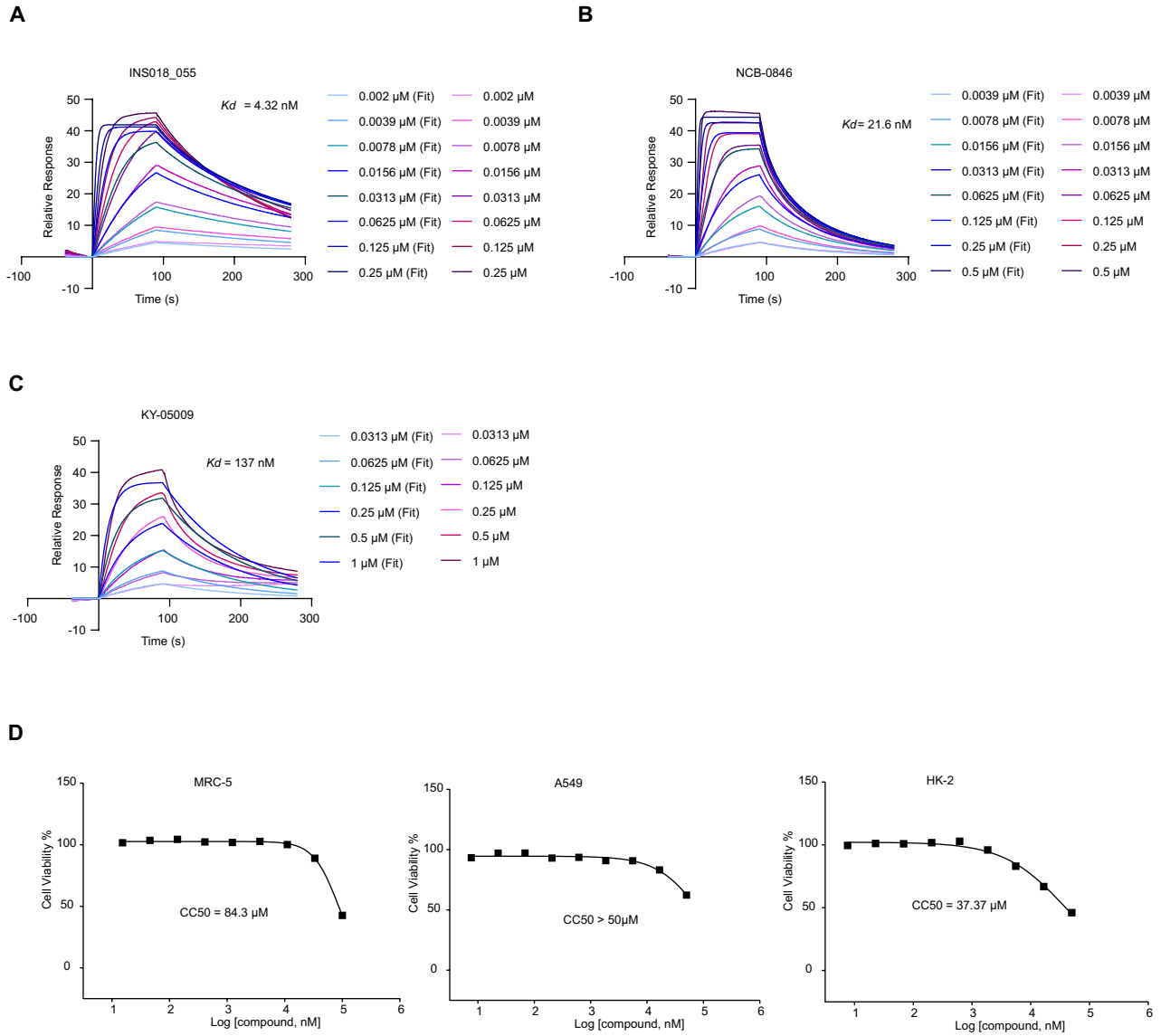


Extended Data Fig. 2 | See next page for caption.

https://doi.org/10.1038/s41587-024-02143-0

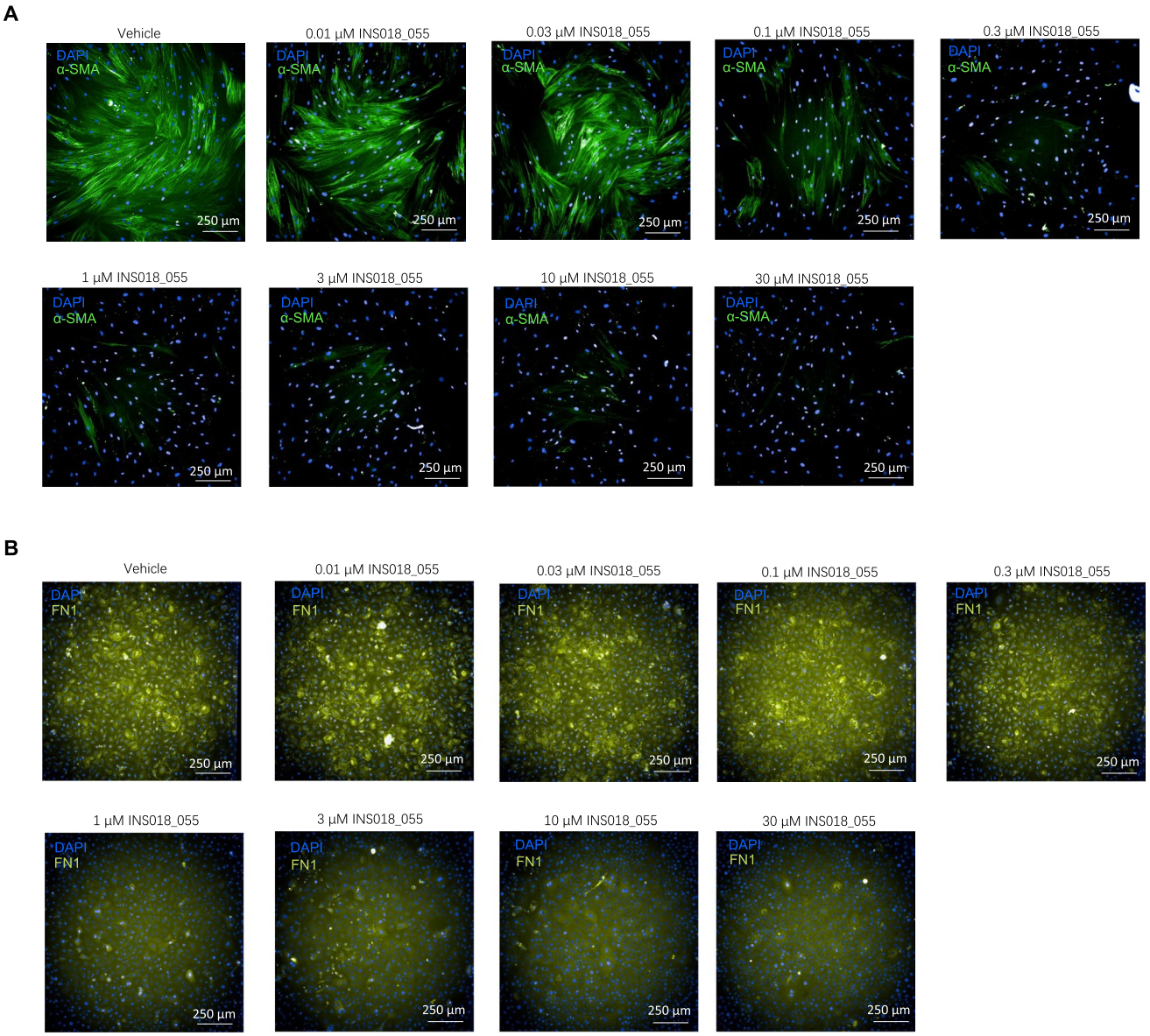
Extended Data Fig. 2 | PandaOmics scores transparency and in-silico validation of TNIK. (a) High values of the graph-based scores can be explained by the significantly perturbed genes observed in the TNIK interactome community neighborhood formed by proteins implicated in IPF and other fibrotic diseases. **(b)** Causal inference score relies on transcriptional factor inference. Majority of the genes regulated by the transcriptional factors are associated with IPF according to OpenTargets and GWAS studies. **(c)** TNIK regulates major signaling pathways known to be causal for the development of fibrotic conditions. **(d)** Clustering of cells derived from single cell RNA-seq data for IPF and healthy lung tissue. **(e)** Visualization of TNIK gene-weighted density in myofibroblasts, cytotoxic T cells and club cells using Nebulosa package. **(f)** Single cell gene expression profiles of IPF lung tissue are presented in boxplots for various unique cell types: Non-classical Monocytes (n = 1058, 1931), ATII High-Surfactuns cells (n = 2655, 496), Low-info Multiplet Macrophages (n = 1000, 2765), Alveolar Macrophage (n = 23905, 27403), Macrophage (n = 40747, 60483), NK cells (n = 2744, 4007), Cytotoxic T cells (n = 4061, 6296), Myofibroblasts (n = 204, 2886), and Club cells (n = 226, 1855) for the control and IPF groups, respectively. These profiles show high expression of TNIK in the key cell types responsible for disease progression: myofibroblasts, cytotoxic T cells, and club cells. The differential expression between IPF and controls for each unique cell type was calculated using the tl.rank_genes_groups function of scanpy package (method = 'Wilcoxon', two-sided), and the resulting log-fold changes (with p-values adjusted using the Benjamini-Hochberg method and threshold = 0.05, FDR) between the IPF and control groups for each cell type of TNIK were plotted on a heatmap. The center lines in the boxplots indicate the medians, while the box limits represent the 25th and 75th percentiles. (**g**) Simulated knockout profile of TNIK in myofibroblasts derived from IPF patients by scTenifoldKnk virtual knockout tool confirms the importance of Hippo and YAP/TAZ signaling in the TNIK-mediated regulation of the major IPF-related pathways.

https://doi.org/10.1038/s41587-024-02143-0



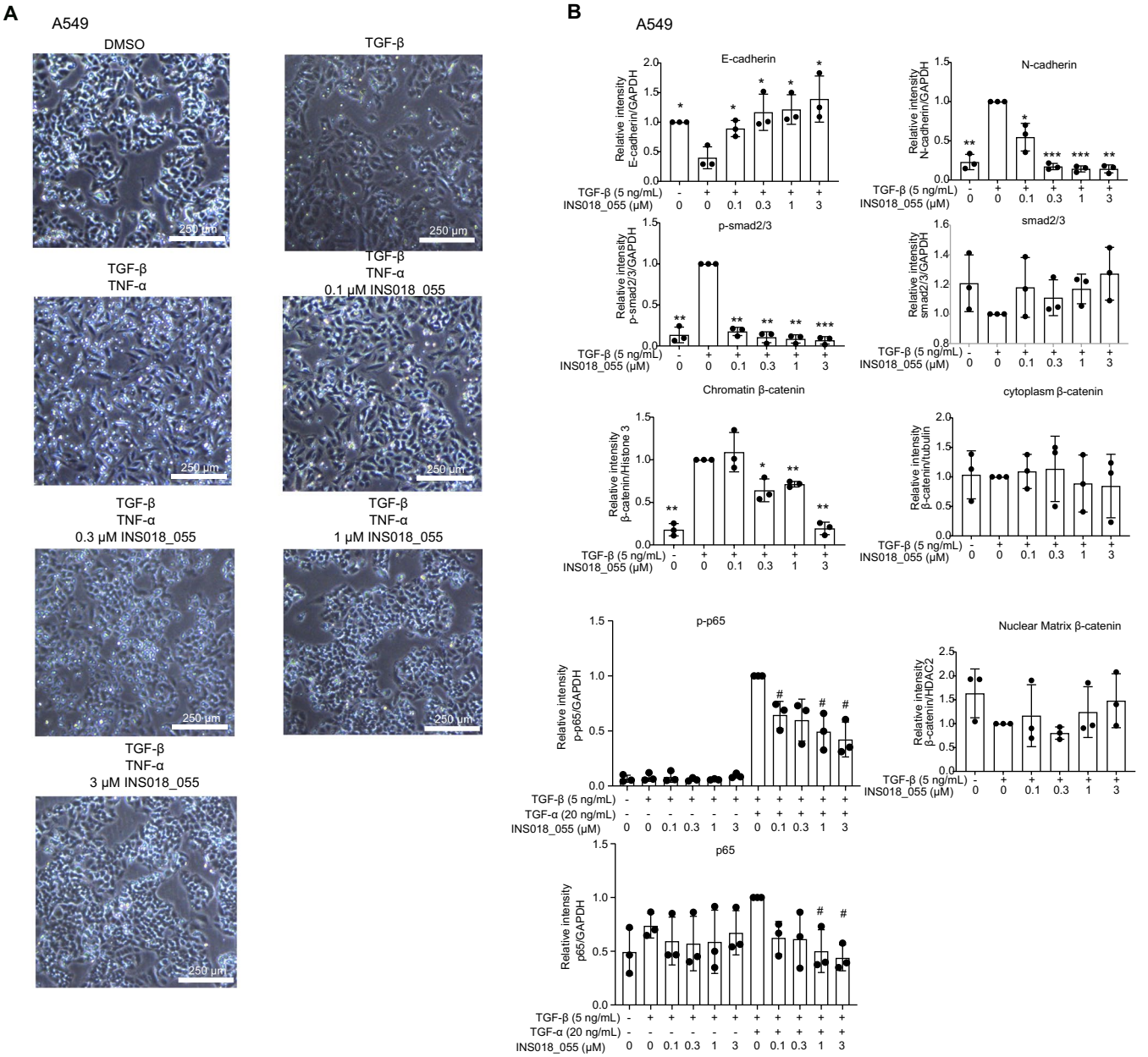
Extended Data Fig. 3 | INS018_055 target affinity and cell viability analysis. Surface plasmon resonance (SPR) assay to measure the binding kinetics of (**a**) INS018_055, (**b**) NCB-0846 and (**c**) KY-05009 to His tagged TNIK (9-315). (**d**) Cell viability analysis performed on MRC-5 cells **(left)**, A549 cells **(middle)** and HK-2 cells **(right)**. Cells were seeded at 8,000, 5,000 and 8,000 cells/well in 96-well plate, respectively and incubated with INS018_055 for 72 hours. Cell viability was measured by Cell Titer Glo. All experiments were performed in duplicate wells. CC_{50} was calculated with the data of cell viability (%). To determine the CC_{50} value, the data was fitted using an equation for a sigmoidal dose response (variable slope), as provided by GraphPad™ Prism software. The equation was shown as: Y = Bottom + (Top-Bottom)/(1+10^((LogIC_{50}-X)*HillSlope)).

https://doi.org/10.1038/s41587-024-02143-0



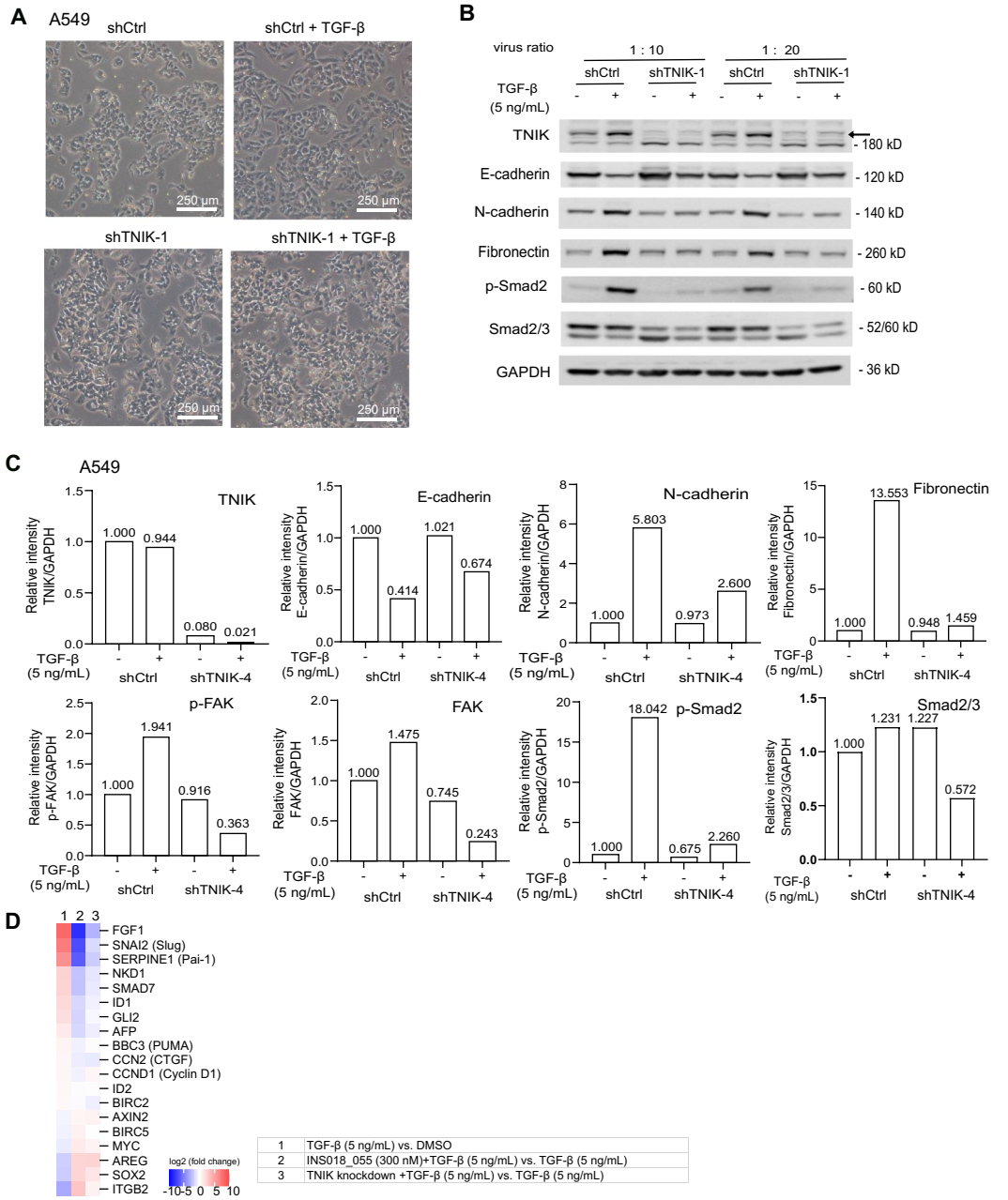
Extended Data Fig. 4 | IPF donor fibroblast and epithelial cell imaging for FMT and EMT assays. (**a**) Representative images of fibroblasts from a donor IPF patient stained with DAPI and alpha-smooth muscle actin (α-SMA) antibody. Cells were pre-treated with TGFβ to induce mesenchymal transition after which cells were treated with the indicated concentrations of INS018_055. Automated quantitation of α-SMA fluorescence intensity and cell density was made following INS018_055 treatment. (**b**) Representative images of fibroblasts from a donor IPF patient stained with DAPI and fibronectin (FN1) antibody. Cells were pre-treated with TGFβ to induce mesenchymal transition after which cells were treated with the indicated concentrations of INS018_055. Automated quantitation of FN1 fluorescence intensity and cell density was made following INS018_055 treatment. Fibroblasts and epithelial cells were isolated from 3 IPF donor tissue samples and 3 healthy donor samples. The experiment was carried out with all 6 independent samples in a single experiment by Charles River's high throughput imaging pipeline. Source data is provided.

https://doi.org/10.1038/s41587-024-02143-0



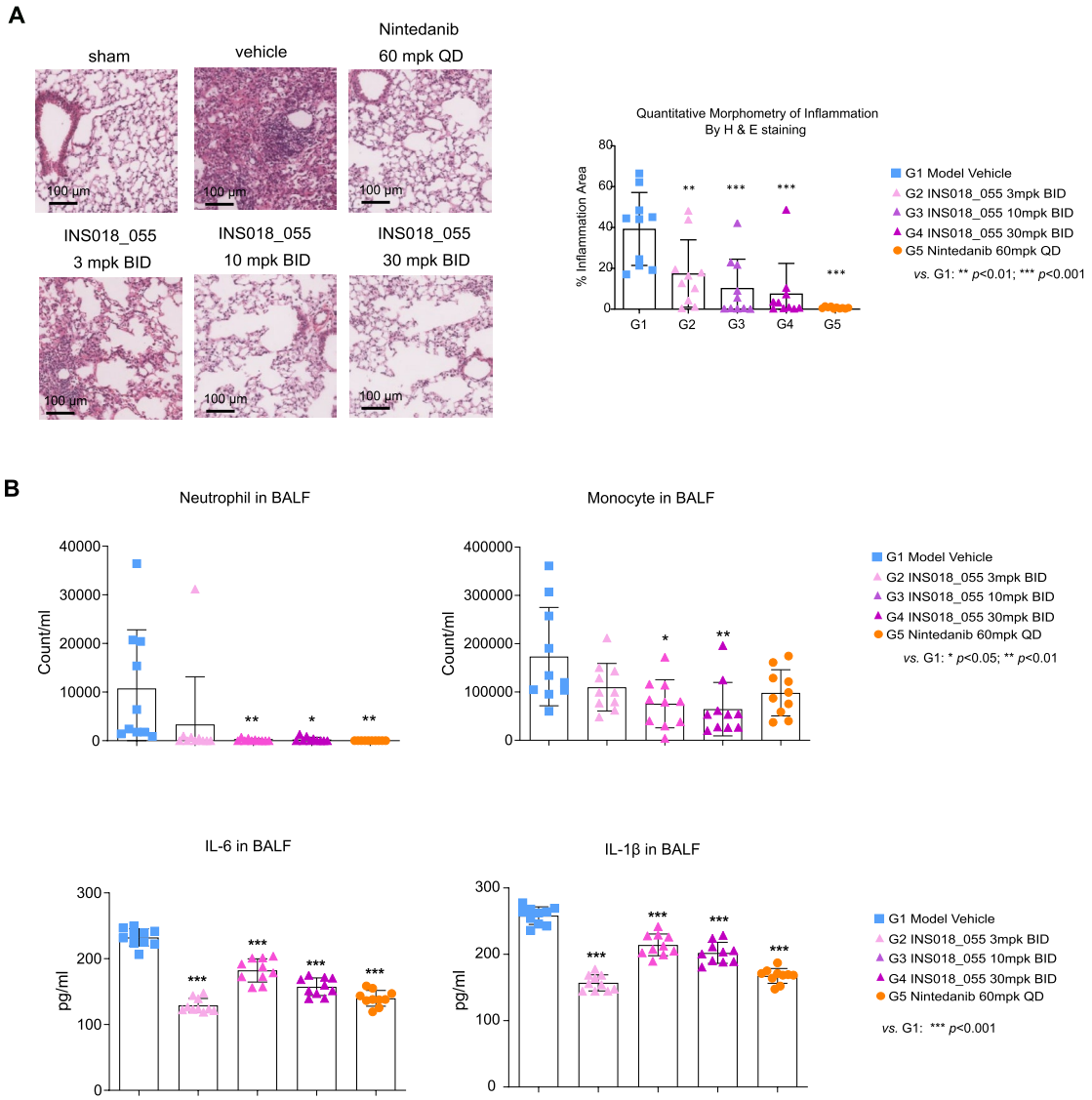
Extended Data Fig. 5 | Effect of IN18_055 on TGF-β-induced, or combination of TGF-β and TNF-α induced EMT/FMT cellular programs. (a) Representative pictures of morphology change of A549 cells. **(b)** Bar graphs (mean ± SD) of inhibition effect of INS018_055 on TGF-β induced changes of E-cadherin, N-cadherin, phospho-smad2, smad2/3 and beta-catenin levels in A549 cells; Bar graphs (mean ± SD) of inhibition effect of INS018_055 on TGF-β and TGF-α-induced phospho-NF-kb p65 and its total protein changes in A549 cells. Images of blots refers to Fig. 2e, f. Individual values are shown in dots. P values are analyzed by Welch's t-test (two-sided) among 3 independent experiments. p values < 0.05 were considered statistically significant (*: significant compared to group treated with TGF-β but without INS018_055; #: significant compared to group treated with TGF-β and TGF-α but without INS018_055). Compared to TGF-β control group, group (no TGF-β), groups with INS018_055 (0.1, 0.3, 1 and 3 μM) showed p values (E-cadherin) of 0.0302, 0.024, 0.0288, 0.0127 and 0.031; p values (N-cadherin) of 0.0054, 0.0469, 0.0008, 0.0007 and 0.0012; p values (p-smad2/3) of 0.004, 0.0012, 0.0018, 0.001 and 0.0008; p values (chromatin β-catenin) of 0.5714, 0.0436, 0.0043 and 0.0028. Compared to TGF-β and TGF-α control group, groups with INS018_055 (0.1, 0.3, 1 and 3 μM) showed p values (p-p65) of 0.0383, 0.0675, 0.0341 and 0.0242; p values (p65) of 0.0503, 0.1243, 0.0492 and 0.0151.

https://doi.org/10.1038/s41587-024-02143-0



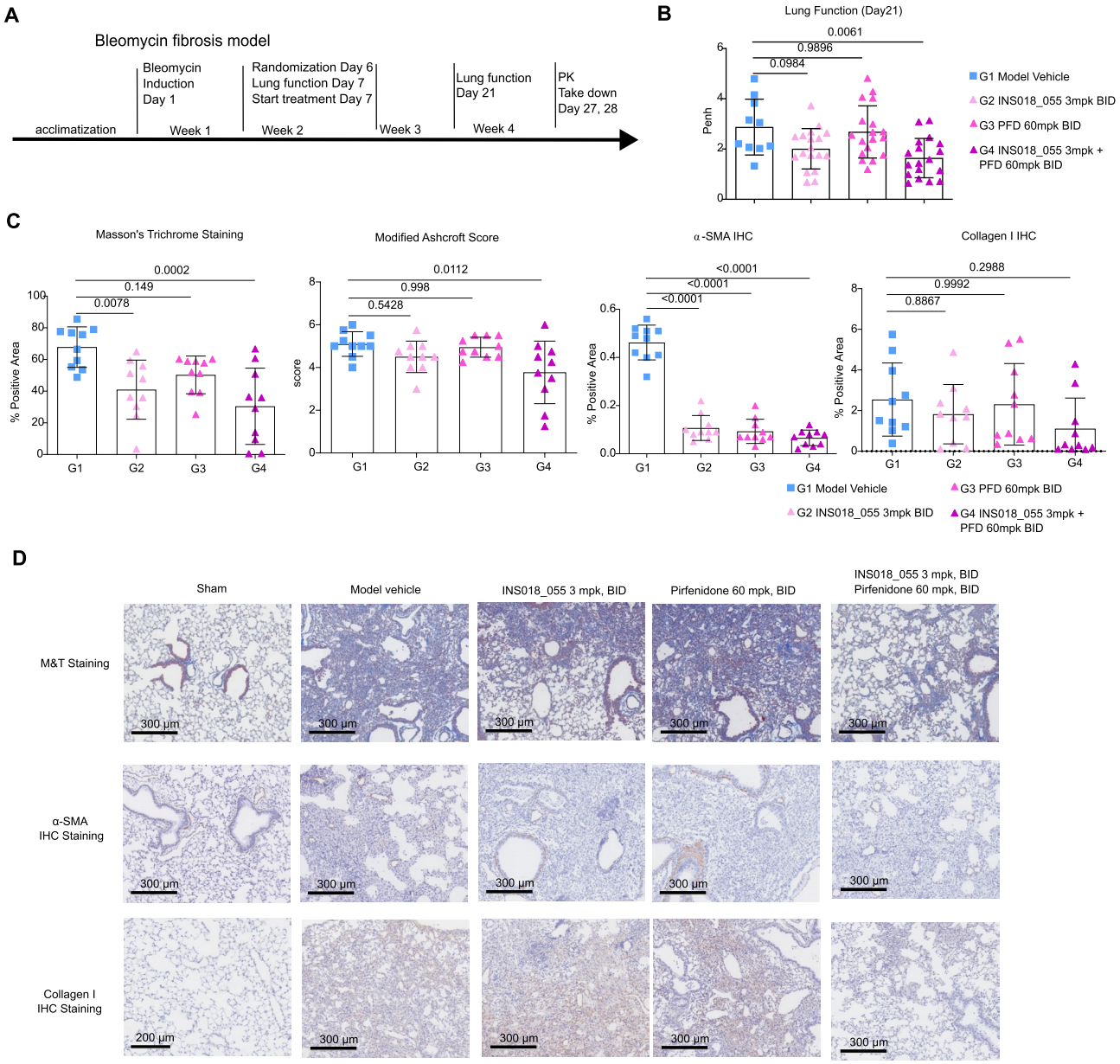
Extended Data Fig. 6 | Effect of TNIK deletion on TGF-β signaling in A549 cells. (a) Representative pictures of morphology change of A549 cells. Single independent experiment- no statistical comparisons were made. **(b)** Western Blot images of inhibition of TGF-β induced changes of E-cadherin, N-cadherin, fibronectin, and phospho-smad2 by shTNIK (shTNIK-1). Single independent experiment- no statistical comparisons were made. **(c)** Bar graphs of Fig. 2g, the inhibition of TGF-β induced changes of E-cadherin, N-cadherin, fibronectin, phospho-FAK and phospho-smad2 by shTNIK (shTNIK-4). **(d)** Heatmap of log2(fold changes) of gene expression change of Hippo signaling related proteins.

https://doi.org/10.1038/s41587-024-02143-0



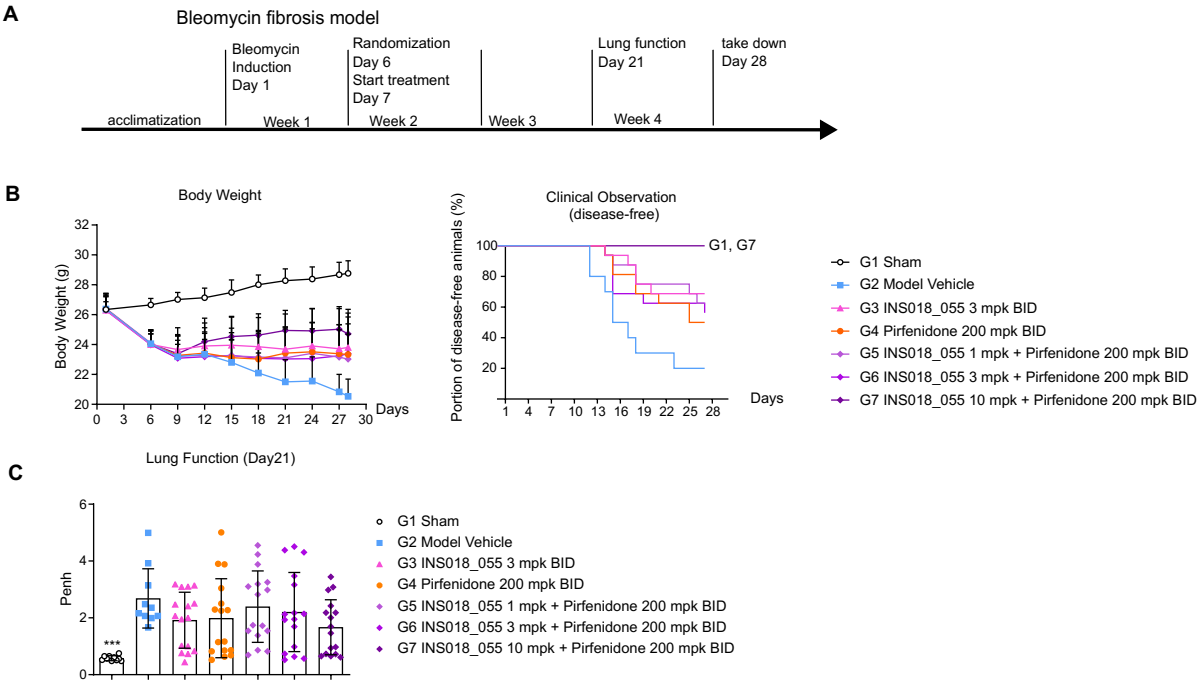
Extended Data Fig. 7 | INS018_055 inhibits inflammation in a bleomycin-induced lung (BLM) fibrosis mouse model. (a) Images and quantitation of inflammatory area under the treatment of INS018_055 (3, 10, or 30 mg/kg, BID) or nintedanib (60 mg/kg, BID) in bleomycin-induced lung fibrosis model by H&E staining. n = 10 per group. (mean ± SD) Individual values are shown in dots. Comparing G1 (model vehicle group) to G2, 3, 4, 5, yielded p = 0.0092, 0.0003, <0.0001, <0.0001, respectively; Comparing G5 (nintendanib) to G2, 3, 4: p = 0.0789, 0.6504, 0.9063, respectively) **(b)** cell counts of neutrophils, monocytes, and levels of IL-6, IL-1β in BALF of the animals in different treatment groups. n = 10 per group. (mean ± SD) Neutrophils: Comparing G1 (model vehicle) with G2, 3, 4, 5, yielded p = 0.1391, 0.0094, 0.0105, 0.0082, respectively; Comparing G5 (nintendanib) with G2, 3, 4 yielded p = 0.9073, >0.999, >0.999, respectively. Monocytes: Comparing G1 (model vehicle) with G2, 3, 4, 5, resulted in p = 0.2069, 0.01, 0.0031, and 0.0825, respectively; Comparing G5 (nintedanib) with G2, 3, 4, resulted in p = 0.9997, 0.9819, 0.8605, respectively; IL-6: Comparing G1 (model vehicle) with G2, 3, 4, 5, yielded p = <0.0001, <0.0001, <0.0001, <0.0001, respectively; Comparing G5 (nintedanib) with G2, 3, 4 resulted in p = 0.4435, <0.0001, 0.0403, respectively; IL-1β: Comparing G1 (model vehicle) with G2, 3, 4, 5, resulted in p = <0.0001, <0.0001, <0.0001, <0.0001, respectively; compared with G5: G2, 3, 4: p = 0.5115, <0.0001, <0.0001, respectively. All statistical analyses reported here are Ordinary One-way-anova with post-hoc Šídák's multiple comparisons testing. (asterisks were used instead of an exact value only for *** when p < 0.0001).

https://doi.org/10.1038/s41587-024-02143-0



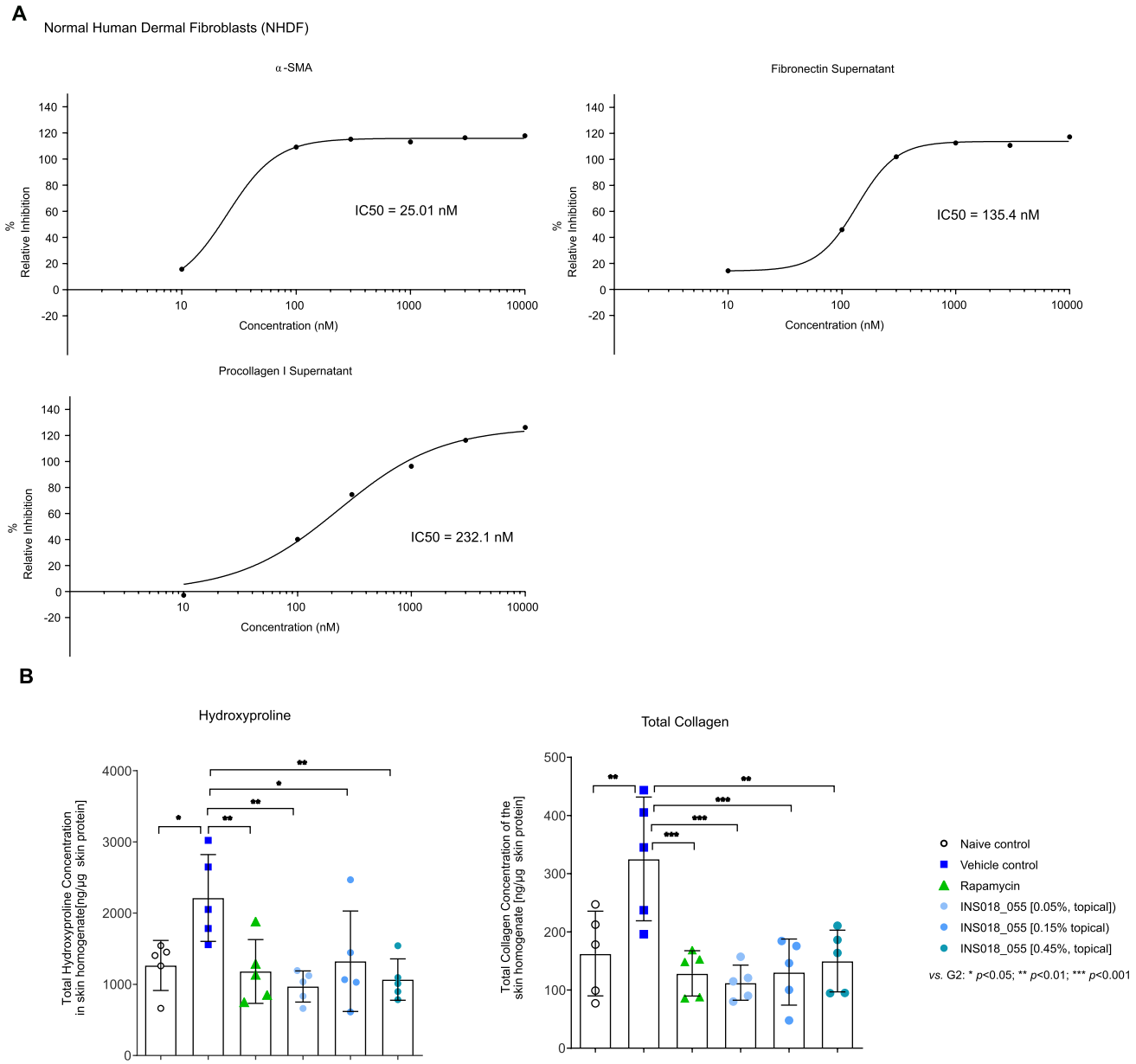
Extended Data Fig. 8 | *In vivo* study on the effect of combination therapy of INS018_055 and sub-optimal dose Pirfenidone on mouse model of bleomycin induced lung fibrosis. (a) Study design of bleomycin-induced lung fibrosis model in C57BL/6 male mice (n = 10 per group). **(b)** lung function on day 21 measured by Penh (mean ± SD) (n = 10 for G1, and n = 18 for G2-4, with data of animals for PK sampling included). Comparing G1 (model vehicle) with G2, G3, G4, yielded p value = 0.0984, 0.9896, and 0.0061, respectively; Comparing G4 (combo group) with G2, G3 yielded p values = 0.7383 and 0.0062, respectively. **(c)** Quantification of INS018_055 in bleomycin-induced lung fibrosis model Masson's trichrome staining for fibrosis area, Modified Ashcroft Score, and IHC staining of Collagen 1 and α-SMA (mean ± SD, n = 10). Individual values are shown in dots. MT staining: Comparing G1 (model vehicle) with G2, G3, G4, resulted in p value = 0.0078, 0.149, and 0.0002, respectively; Comparing G4 (combo group) with G2, G3 yielded p values = 0.6478 and 0.0774, respectively. Modified Ashcroft Score: Comparing G1 (model vehicle) with G2, G3, G yielded p value = 0.5428, 0.998, and 0.0112, respectively; Comparing G4 (combo group) with G2, G3 resulted in p values = 0.3414 and 0.0298, respectively. α-SMA: Comparing G1 (model vehicle) with G2, G3, G4 yielded p value = <0.0001, <0.0001, and <0.0001, respectively; Comparing G4 (combo group) with G2, G3 resulted in p values = 0.3935 and 0.7877, respectively. Collagen I: Comparing G1 (model vehicle) with G2, G3, G4 yielded p value = 0.8867, 0.9992, and 0.2988, respectively; Comparing G4 (combo group) with G2, G3 resulted in p values = 0.8884 and 0.4865, respectively. (**d**) Representative histopathology images including M&T Staining, α-SMA IHC Staining and Collagen I IHC Staining. All statistical analyses reported here are Ordinary One-way-anova with post-hoc Šídák's multiple comparisons testing. (asterisks were used instead of an exact value only for *** when p < 0.0001).

https://doi.org/10.1038/s41587-024-02143-0



Extended Data Fig. 9 | *In vivo* effects of INS018_055 treatment in combination with maximum therapeutic dose of Pirfenidone in mouse models of lung fibrosis. (a) study design of bleomycin-induced lung fibrosis model in C57BL/6 male mice. **(b)** Left, body weight change during study process; right: Disease-free curve based on clinical observations (Group,1,2, n = 10, Group 3-7, n = 16) (mean ± SD). **(c)** Lung function measured by Penh (Group,1,2, n = 10, Group 3-7, n = 16) (mean ± SD). Comparing G2 (model vehicle) with G1, p = 0.0005; G3-G7, p = 0.4629, 0.5745, 0.9888, 0.8805 and 0.1635, respectively. All statistical analyses reported here are Ordinary One-way-anova with post-hoc Šídák's multiple comparisons testing.

https://doi.org/10.1038/s41587-024-02143-0



Extended Data Fig. 10 | *In vitro* and *in vivo* study of INS018_055 on skin fibrosis. (a) IC_{50} values of inhibition effect of INS018_055 on levels of α-SMA expression, and levels of fibronectin and procollagen I in supernatant in cultured NHDF cells. Left: Cells were treated with INS018_055 and TGF-β (0.1 ng/ml) for 72 hours. The fluorescence intensity of α-SMA normalized to the total number of nuclei identified by staining with Hoechst 33258. Right and below: Cells were treated with INS018_055 and TGF-β (10 ng/ml) for 72 hours. Fibronectin and procollagen I contents were measured in the culture supernatants. Relative inhibition (%) = (Mean Stimulated control - Value)/ (Mean Stimulated control −Mean Non-stimulated control) X 100. All experimental conditions were performed in n = 5. (mean ± SEM). **(b)** INS018_055 inhibited collagen level in bleomycin induced skin-thickening rat model. Sprague Dawley rats were injected with 0.1 mL 1 mg/mL bleomycin subcutaneously into two sites on the shaved regions, once daily for 4 weeks. Treatments were administered with vehicle, Rapamycin (1.5 mg/kg, intraperitoneal) or INS018_055 (0.05%, 0.15% or 0.45%, topical), 30 minutes before administration of bleomycin, everyday for four weeks. Hydroxyproline and collagen content was measured (mean ± SD), (n = 5). The concentrations of hydroxyproline and collagen were normalized to total protein content using the Bradford method. Statistical analysis was performed using one-way ANOVA, then Dunnett's multiple comparison test was performed. For hydroxyproline, compared with vehicle control group (Group 2), p values of Group naive control, Group of rapamycin, Groups of INS018_055 [0.05%, 0.15% and 0.45%], are 0.0165, 0.0085, 0.0015, 0.0261, and 0.0033, respectively. For total collagen, p values of Group naive control, Group of rapamycin, Groups of INS018_055 [0.05%, 0.15% and 0.45%], are 0.0025, 0.0003, 0.0001, 0.0004, and 0.0012, respectively. p values < 0.05 were considered statistically significant.

nature portfolio

Corresponding author(s): Alex Zhavoronkov, PhD

Last updated by author(s): Dec 22, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	For target discovery part, all data used in this manuscript are public available, and described in methods and materials.
Data analysis	For target discovery part, the analytical softwares used in this manuscript (PandaOmics v2.0 and Chemistry42) are commercially available. We have provided supplementary information for both software in the manuscript files. For pre-clinical data, raw data were analyzed using Microsoft Excel software(Microsoft 365), GraphPad PRISM (Versions 6, 8) or SPSS software (Versions 25 and later). Versions used varies in different contract research organizations who performed the studies. For clinical phase 0 data (Australia), analyses were performed using Phoenix WinNonlin software (Version 8.3 or higher, Certara, USA). For clinical phase 1 data (New Zealand), all analyses were performed using SAS® Version 9.4. For clinical phase 1 data (China), analyses were performed using SAS version 8.3.1 or higher.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

PandaOmics v2.0 and Chemistry42 are commercially available:
https://insilico.com/pandaomics
https://insilico.com/chemistry42
We havea also provided full information on the drug discovery pipelines used in this mansucript in our supplementary information section.

Data Availability Statement:
Raw data from all experimental studies are available for public access through our repository which can be found at www.insilico.com/nbt-ins018-055-tnik/. The RNA sequencing data is available for download through this repository link along with all relevant source and raw data that were presented in this manuscript. The only two datasets which are not uploaded are the EMT/FMT full imaging set and the inhalation pathology imaging set as together these images are approximately 1 TB in size. The quantitation sheets are available in the repository and the images can be provided as a physical drive for any reasonable, public requests. This statement has been included in the manuscript as well.

Humans that participated in all of the clinical trials in this study provided written consent towards their study participation.

Clinical Datasets:
Phase 0 Study (Australia)
A phase 0 micro dosing clinical trial was conducted in Australia (ACTRN12621001541897). The complete information on clinical trial registration, study protocol, data collection and outcomes are provided in the reporting summary as well as in Supplementary information 9. This includes statistical considerations, study design, patient selection criteria, procedures, outcomes and PK analysis.

Phase I study (New Zealand):
The general design of clinical trial (NCT05154240) can be found in www.clinicaltrials.gov The randomized, double blind, placebo controlled study of INS018_055 was conducted from Feb 21, 2022 (first subject administered first dose) until Sept 30th, 2022 (last subject contacted). The complete information on clinical trial registration, study protocol, data collection and outcomes are provided in the reporting summary as well as in Supplementary information 10. This includes statistical considerations, study design, patient selection criteria, procedures, outcomes and PK analysis.

Phase I study (China):
Study design and Outcomes:
Detailed information including facilities, inclusion/exclusion criteria could be found in link http://www.chinadrugtrials.org.cn/clinicaltrials.prosearch.dhtml (registration number: CTR20221542).
The complete information on clinical trial registration, study protocol, data collection and outcomes are provided in the reporting summary as well as in Supplementary information 11. This includes statistical considerations, study design, patient selection criteria, procedures, outcomes and PK analysis.

Patient Samples for EMT and FMT Assays:
All tissues used for isolation are obtained under informed consent and conform to HIPAA regulations to protect the privacy of the donor's Personally Identifiable Information. The entirety of this study adheres to the Declarations of Helsinki. Quantitation of donor cell lines used are included in the repository link for this study provided above and identified imaging datasets for these studies can be provided in the form of a physical drive as the datasets are nearly 1 TB in size.

Publicly Available -Omics Datasets Utilized in this Study:
GSE93606, GSE38958, GSE28042, GSE33566 derived from blood tissue and GSE101286, GSE72073, GSE150910, GSE92592, GSE52463, GSE83717, GSE21369, GSE15197, GSE99621, GSE138283, GSE24206 derived from lung tissue. GSE136831 - was used for scRNA-seq analysis. These datasets were deposited into the GEO and were utilized in the contexts indicated following each accession number.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Clinical studies enrolled both male and female healthy volunteers.
Population characteristics	Clinical trial (phase 0, Australia): For phase 0 study, 4 female subjects (50%) and 4 male subjects (50%) were enrolled into the study. Five subjects were white (62.5%); two were Asian (25%), one was not specified for race (12.5%). All 8 (100%) subjects are non-hispanic/Latino. The mean age was 28.8 years (range 18 to 42years), and the mean BMI was 24.93 kg/m2 (range 19.7 to 29.0 kg/m2). Clinical trial (phase 1, New Zealand):

Part A (SAD):
For Part A, 22 female subjects (55.0%) and 18 male subjects (45.0%) were enrolled into the study. The majority of subjects were White (33 subjects [82.5%]) and not Hispanic or Latino (32 subjects [80.0%]). The mean age was 28.4 years (range 19.0 to 51.0 years), and the mean BMI was 24.95 kg/m2 (range 19.3 to 30.5 kg/m2).
Part B (MAD):
For Part B, 13 female subjects (54.2%) and 11 male subjects (45.8%) were enrolled into the study. The majority of subjects were White (19 subjects [79.2%]) and not Hispanic or Latino (21 subjects [87.5%]). The mean age was 25.8 years (range 18.0 to 50.0 years), and the mean BMI was 23.82 kg/m2 (range 18.7 to 30.1 kg/m2)
For Part C, 5 female subjects (35.7%) and 9 male subjects (64.3%) were enrolled into the study (total 14 subjects). The majority of subjects were White (12 subjects [85.7%]) and not Hispanic or Latino (10 subjects [71.4%]). The mean age was 29.1 years (range 19.0 to 48.0 years), and the mean BMI was 24.99 kg/m2 (range 19.9 to 31.1 kg/m2).
Clinical trial (phase 1, China):
Part A (SAD): 17 (70.8%) male and 7 (29.2%) female subjects were enrolled in the study; 23 (95.8%) subjects were of Han Chinese ethnicity; the mean age of the subjects in each dosage group (30 mg, 60 mg, 120 mg, and placebo) was 33.2 years old, 27.5 years old, 28.5 years old, 28.7 years old, and the mean values of weight, height, and body mass index (BMI) were 64.01 kg, 167.25 cm, and 22.79 kg/m2 , respectively, in all subjects.
Part B (MAD): 20 (83.3%) male and 4 (16.7%) female subjects were enrolled in the study; 22 (91.7%) subjects were of Han Chinese ethnicity; the mean age of subjects in each dosage group (30 mg, 60 mg, 90 mg, and placebo) was 31.8, 26.0, 25.3, and 27.8 years old, respectively; and the mean values of body weight, height, and body mass index (BMI) of all subjects were 63.06 kg, 167.79 cm, and 22.21 kg/m2 , respectively.

Recruitment

Healthy subjects were recruited according to pre-specified inclusion/exclusion criteria.
There is no known bias in selection of subjects and site.
Before entering the study, each trial participant was given a full explanation of the study to ensure the full understanding of the implications of participating the study before signing and giving a written informed consent which is in compliance with ICH E6(R2). The investigator submitted the informed consent form (ICF) to the IRB for review and approval before the initiation of the study.

Australia Phase 0:
Participants who met all of the following criteria at Screening were eligible to participate in the study:
1. Healthy female or male aged ≥ 18 and ≤ 55 years at Screening. Healthy was defined as no clinically relevant abnormalities identified by a detailed medical history, full physical examination, including blood pressure and pulse rate measurement, 12-lead electrocardiogram (ECG), and clinical laboratory tests.
2. BMI of 17.50 to 30.50 kg/m2; and a total body weight > 50 kg at Screening and Day -1.
3. Sufficient venous access for the purposes of the study.
4. Non-smoker. The participant must not have used any tobacco products within 2 months prior to Screening.
5. Females were non-pregnant and non-lactating, and agreed to use an acceptable, highly effective double contraception from Screening until study completion, including the Follow-up period. Double contraception was defined as a condom AND one other form of the following:

a. Established hormonal contraception (with approved long-acting implantable hormones, injectable hormones). Oral contraceptive pills [OCPs] could not be used as a second form of contraception by female participants, due to the unknown potential for OCP reduced effectiveness when administered in combination with INS018_055.
b. A vaginal ring or an intrauterine device (IUD) (including a hormonal IUD).
c. Documented evidence of surgical sterilisation at least 6 months prior to Screening (eg, tubal occlusion, hysterectomy, bilateral salpingectomy, or bilateral oophorectomy for women or vasectomy for men [with appropriate post-vasectomy documentation of the absence of sperm in semen] provided the male partner was a sole partner).

Women not of childbearing potential were post-menopausal for ≥ 12 months. Post-menopausal status was confirmed through testing of follicle-stimulating hormone (FSH) levels ≥ 40 IU/L at Screening for amenorrhoeic female participants.
Females who were abstinent from heterosexual intercourse were also eligible.
Periodic abstinence (eg, calendar, ovulation, symptothermal, post-ovulation methods) and withdrawal were not considered highly effective methods of birth control. Participant complete abstinence for the duration of the study and for 90 days after the last study treatment was acceptable.
Female participants who were in same-sex relationships were not required to use contraception.
Women of childbearing potential (WOCBP) had a negative pregnancy test at Screening and Day 1 and were willing to have additional pregnancy tests as required throughout the study.
Males were surgically sterile (> 30 days since vasectomy with no viable sperm), abstinent, or if engaged in sexual relations with a WOCBP, the participant and his partner were surgically sterile (eg, tubal occlusion, hysterectomy, bilateral salpingectomy, bilateral oophorectomy) or using an acceptable, highly effective contraceptive method from Screening until study completion, including the Follow-up period. Acceptable methods of contraception included the use of condoms and the use of an effective contraceptive for the female partner that included: OCPs, long-acting implantable hormones, injectable hormones, a vaginal ring, or an IUD. Participants with same-sex partners (abstinence from penile-vaginal intercourse) were eligible when this was their preferred and usual lifestyle.
Males could not donate sperm for at least 90 days after the last study treatment.
6. Willing and able to attend the trial visits and complete study assessments.
7. Willing to consume standard meals provided.
8. Able to read and understand study documents and follow Investigator and study personnel instructions during visits.
9. Able to sign the HREC approved ICF.

Exclusion Criteria

Participants who met any of the following criteria at Screening were not eligible to participate in the study:
1. Positive toxicology screening panel (urine test including qualitative identification of barbiturates, tetrahydrocannabinol [THC], amphetamines, benzodiazepines, opiates, and cocaine), or with a history of substance abuse or dependency or history

of recreational IV drug use over the last 5 years (by self-declaration).
2. Positive alcohol breath test at Screening or a history of regular alcohol consumption exceeding 14 drinks/week for women or 21 drinks/week for men (1 drink = 150 mL of wine or 360 mL of beer or 45 mL of hard liquor) within the 6 months prior to Screening.
3. Major surgery or significant trauma within 28 days (4 weeks) prior to Screening.
4. Blood pressure (BP) > 150 mmHg (systolic) or > 95 mmHg (diastolic) at Screening and Day -1, following at least 5 minutes of supine rest. If BP was > 140 mmHg (systolic) or > 90 mmHg (diastolic), the BP measurements was repeated 2 more times, at least 2 minutes apart, and the average of the 3 BP values was used to determine the participant's eligibility.
5. Heart rate < 45 beats per minute (bpm) or > 100 bpm at Screening and Day -1, following at least 5 minutes of supine rest. If heart rate was below 45 bpm or exceeds 100 bpm, the heart rate was repeated 2 more times, at least 2 minutes apart, and the average of the 3 heart rate values was used to determine the participant's eligibility.
6. 12-lead ECG demonstrating QTc > 450 msec for males or > 470 msec for females, or a QRS interval ≥ 120 msec at Screening and Day -1. If QTc exceeded 450 msec (males) or 470 msec (females), or QRS exceeded 120 msec, the ECG was repeated 2 more times, at least 2 minutes apart, and the average of the 3 QTc (or QRS) values was used to determine the participant's eligibility.
7. ANY of the following abnormalities in clinical laboratory tests at Screening, as assessed by the study-specific laboratory and confirmed by a single repeat, if deemed necessary: – Serum creatinine level above the upper limit of normal (ULN) or an estimated glomerular filtration rate value < 80 mL/min, based on the Cockcroft-Gault calculation, at Screening.
– Aspartate aminotransferase (AST) / serum glutamic oxaloacetic transaminase (SGOT) or alanine aminotransferase (ALT) / serum glutamic pyruvic transaminase (SGPT) > 1.5 × ULN.
– Fasting glucose > 5.4 mmol/L.
– Total bilirubin > 1.5 × ULN.

8. A white blood cell count < 4.0 × 109/L. Participants with borderline clinical laboratory values outside the reference range could be included in the study if the Investigator deemed that the values were not clinically significant.
9. Absolute neutrophil count of < 2 × 109/L.
10. Haematocrit below 0.4 for males and 0.35 for females.
11. Use of any IP or investigational medical device within 30 days prior to Screening, or 5 half-lives of the product (whichever was the longest) or participation in more than 4 investigational drug studies within 1 year prior to Screening.
12. Use of prescription or non-prescription drugs and dietary supplements within 7 days or 5 half-lives (whichever was longer) prior to INS018_055 administration, with the exception of paracetamol, which could be used at doses of ≤ 2 g/day, and contraceptives.
13. Blood donation (excluding plasma donations) of ≥ 500 mL or significant blood loss within 56 days prior to dosing.
14. History of sensitivity to heparin or heparin-induced thrombocytopenia.
15. Other severe acute or chronic medical or psychiatric condition including recent (within the past year) or active suicidal ideation or behaviour or laboratory abnormality, or any other abnormality that in the opinion of the Investigator could increase the risk associated with study participation or IP administration or could interfere with the interpretation of study results and make the participant inappropriate for entry into this study.
16. CRU staff members directly involved in the conduct of the study and their family members, CRU members otherwise supervised by the Investigator, or participants who were Sponsor employees including their family members directly involved in the conduct of the study.
17. Vaccination with live virus, attenuated live virus, or any live viral components within the 6 weeks prior to the first dose of study drug or was to receive these vaccines at any time during treatment or within 8 weeks following the end of study visit. Scheduled/intended to have a COVID-19 vaccine during the study (ie, from Screening through to Day 8).
18. History of any lymphoproliferative disorder (such as EBV related lymphoproliferative disorder, as reported in some participants on other immunosuppressive drugs), history of lymphoma, leukaemia, myeloproliferative disorders, multiple myeloma, or signs and symptoms suggestive of current lymphatic disease.
19. Clinically significant infection currently or within 6 months of first dose of study drug (those requiring hospitalisation or parenteral antimicrobial therapy or opportunistic infections), or a history of chronic or recurrent infectious disease.
20. Known infection with or test positive at Screening for HIV, hepatitis B or C viruses.
21. History of malignancy, except for non-melanoma skin cancer, excised more than 2 years ago, and cervical intraepithelial neoplasia that had been successfully cured more than 5 years prior to Screening.
22. Consumption of grapefruit or grapefruit juice or citrus fruits (ie, Seville oranges, pomelos, tangelos) within 7 days prior to the first dose of study medication until collection of the final pharmacokinetic blood sample.
23. History of severe allergic reactions (eg, anaphylaxis) or known sensitivity to any of the constituents of the test product.
24. Pregnant or lactating at Screening or planning to become pregnant (self or partner) at any time during the study, including the Follow-up period.
25. History of benign ethnic neutropenia.

Phase I Clinical Trial China:
Inclusion criteria:
Each subject must meet all of the following criteria to be enrolled in this study:
1. Subjects were male or female between 18 and 45 years of age, inclusive.
2. Subjects had a body mass index of 19 to 26 kg/m2 (inclusive) and weighed ≥ 50 kg (inclusive) for men and ≥ 45 kg (inclusive) for women at the time of screening.
3. Based on medical history, clinical laboratory findings, vital sign measurements, 12-lead ECG results, physical examination, chest radiograph results, and serum virological findings at screening, the investigators concluded that the subjects were in good general health.
4. Female subjects of childbearing potential must be non-pregnant and non- lactating and must be using one of the following methods of contraception throughout the treatment period until at least 28 days after the last dose of study drug, or have been surgically sterilized (i.e., hysterectomy, bilateral tubal ligation, or bilateral oophorectomy) or are postmenopausal (defined as 12 consecutive months of amenorrhea with documented plasma follicle stimulating hormone levels > 40 IU/mL). Female subjects must have a negative pregnancy test result at screening and prior to the first dose of study drug.
A highly effective method of contraception is one that has a contraceptive failure rate of less than 1% per year when used consistently. Examples are as follows:

a. Implanted contraceptives (e.g., Jadelle®)
b. IUDs containing copper or levonorgestrel (e.g., Mirena®)
c. Male sterilization, no sperm in ejaculation after vasectomy
d. Double barrier method: condom and occlusion cap (diaphragm or cervical cap/dome cap), barrier method with spermicide (foam/gel/film/cream/suppository) must be used as an add-on
e. Abstinence, defined as complete and continuous avoidance of all heterosexual sex (including during the entire period of risk associated with study treatment), was allowed without contraception only if this was the subject's preferred and daily lifestyle.
or an effective method with a contraceptive failure rate of less than 5% to 10% per year. Examples are as follows:
f. Injectable contraceptives (e.g., Depo Provera)
g. Oral contraceptives (combination hormonal contraceptives or progestogen-only "mini-pills")
h. Vaginal contraceptive ring (e.g., NuvaRing®)
Female subjects must also agree not to donate eggs from the time of administration until at least 28 days after the last dose of study drug.
Male subjects and their fertile female partners must agree to use one of the above methods of contraception for the entire treatment period until at least 28 days after the last dose of study drug. Male subjects must also agree not to donate sperm for the entire treatment period until at least 28 days after the last dose of study drug.
5. Subjects agree to comply with all protocol requirements.
6. Subjects were able to provide written informed consent.

Exclusion criteria:
Subjects who met any of the following criteria were excluded from the study:
1. Subject has current evidence or history of clinically significant hematologic, renal, endocrine, pulmonary, gastrointestinal, cardiovascular, hepatic, psychiatric, neurologic, or allergic disease (including drug allergy, but excluding asymptomatic seasonal allergy that was untreated at the time of administration).
2. Subjects have any condition that may affect drug absorption (e.g., gastrectomy).
3. Subjects had a history of cancer, except adequately treated basal cell or squamous cell carcinoma of the skin.
4. The subject rests for at least 5 minutes with a blood pressure (BP) > 140 mm Hg (systolic) or > 90 mm Hg (diastolic). At screening, if BP is > 140 mm Hg (systolic) or > 90 mm Hg (diastolic), BP should be measured 2 additional times and the average of the 3 BP values should be used to determine the subject's eligibility to participate.
5. At screening, the subject's 12-lead ECG shows a QT interval (QTc) corrected by the Bazett formula (QTc=QT/RR0.5) > 450 msec, or a QRS interval > 120 msec. The average of three QTc (or QRS interval) values from three standard 12-lead ECGs (repeated at intervals of no more than 5 minutes) should be used to determine the subject's eligibility for participation.
6. At the time of screening, the subject has any of the following abnormalities in clinical laboratory tests (if necessary, retest once for confirmation):
a. At screening, serum creatinine levels above the upper limit of normal (ULN) or creatinine clearance (Ccr) < 80 mL/min using the Cockcroft - Gault formula (Appendix 3) and no protein in the urine.
b. Aspartate aminotransferase or alanine aminotransferase values > 1.5 × ULN.
c. Fasting glucose > 110 mg/dL (6.1 mmol/L).
d. Total bilirubin > 1.5 × ULN.
e. Routine blood test values that are outside the normal reference range of local laboratory findings and are considered clinically significant by the investigator.
f. Positive fecal occult blood test result at screening or registration (day -1).
7. Subjects have a history of any disease that may have caused total bilirubin to be higher than ULN. Subjects whose clinical laboratory test values are not significantly outside the reference range may be enrolled in this study if the investigator does not consider the values to be clinically significant.
Note: In subjects with a history of Gilbert's syndrome, direct bilirubin may be measured, and if direct bilirubin < ULN, the subject is eligible for this study.
8. Subjects have any history of lymphoproliferative disease (such as Epstein Barr virus-associated lymphoproliferative disease as reported by some subjects receiving immunosuppressive drugs), lymphoma, leukemia, myeloproliferative disease, multiple myeloma, or signs and symptoms suggestive of current lymphatic disease.
9. Subjects have a history of relevant drug and/or food allergies (i.e., allergy to any study drug or excipient, or any severe food allergy that could result in inability to consume the standard diet of the clinical institution).
10. Subject has a current or clinically significant infection (e.g., an infection requiring hospitalization or parenteral antimicrobial therapy or the presence of an opportunistic infection within the past 6 months) or a history of chronic or recurrent infectious disease within the 6 months prior to the first dose of study drug.
11. Subject has other serious acute or chronic medical or psychiatric illness (including recent (within the past year) or active suicidal ideation or behavior or abnormal experimental results (which may increase the risk associated with study participation or experimental drug administration or may interfere with the interpretation of study results and, in the judgment of the investigator, may render the subject unsuitable for entry into this study).
12. Subjects have a history of symptomatic herpes zoster or herpes simplex, more than one episode of localized herpes zoster, or disseminated herpes zoster (single episode) present or within 12 weeks.
13. Subjects tested positive for hepatitis B surface antigen, hepatitis C virus antibody, or human immunodeficiency virus type 1 or 2 antibody at screening.
14. Subjects were pregnant or lactating females.
15. Subjects are men of childbearing potential who are unwilling or unable to use the contraceptive methods described in this protocol throughout the study period and for at least 28 days after the last dose of the experimental drug.
16. Subjects are unwilling or unable to comply with the lifestyle restrictions described in this protocol.
17. Subjects were smokers or had used nicotine or nicotine-containing products (e.g., snus, nicotine patches, nicotine chewing gum, simulated cigarettes, or inhalants) within 6 months prior to the first dose of the study drug.
18. Subjects tested positive for substance abuse or cotinine (indicating current active smoking) prior to the first dose of study drug.
19. Subjects have used any prescription or over-the-counter medication (except paracetamol [up to 2 g/day]), including herbal supplements, within 14 days prior to the first dose of study drug. Nutritional supplements are permitted provided that they are unlikely to interfere with the study results and that investigator consent has been obtained.
20. Subjects ingested grapefruit or grapefruit juice, limes or products containing limes (e.g., orange marmalade) or products

containing alcohol, caffeine or xanthines within 48 hours prior to the first dose of the study drug.
21. Subjects will be vaccinated with live virus, live attenuated virus, or any live viral component within 2 weeks prior to the first dose of study drug, or will receive these vaccines at any time during the study or within 8 weeks of study completion.
22. Subject tested positive for Severe Acute Respiratory Syndrome-associated Coronavirus 2 (SARS-CoV-2). Subject received the 2019 coronavirus disease (COVID-19) vaccine within 2 weeks prior to the first dose of study drug or is scheduled to receive the COVID-19 vaccine within 12 weeks of study drug administration, or tested positive for SARS-CoV-2 during screening or had COVID-19 symptoms within 4 weeks prior to Day -1.
23. Subjects who have undergone major trauma or major surgery within 4 weeks prior to screening or who are expected to require major surgery during the trial.
24. Subjects are at risk for bleeding: genetic predisposition to bleeding, a bleeding event within 12 months prior to screening start, or abnormal laboratory coagulation parameters.
25. Subjects have a first-degree relative with a genetic immunodeficiency.
26. Subjects were study site staff and their family members who were directly involved in the implementation of the study, study site staff who were otherwise supervised by the investigator, or sponsor employees (including their family members) who were directly involved in the implementation of the study.
27. Subject has a history of alcohol abuse or drug addiction or excessive alcohol consumption (regular alcohol intake > 21 units/week for male subjects and > 14 units/week for female subjects; 1 unit equals approximately ½ pint [200 mL] of beer, 1 small glass [100 mL] of wine, or 1 cup [25 mL] of spirits) within the past year or consumed alcohol 24 hours prior to the first dose of study drug.
28. Subjects engaged in strenuous activity or contact sports within 24 hours prior to dosing and during the study.
29. The subject donated > 450 mL of blood or blood products within 30 days prior to the first dose of the study drug.
30. Subjects received study drug in another pilot study within 30 days prior to dosing or 5 drug half-lives, whichever is longer.
31. Subjects received cytochrome P450 (CYP3A4 and CYP2C8) and P-gp inhibitors and/or inducers within 4 weeks prior to the first dose of INS018_055 or may have received CYP3A4 and CYP2C8 and P-gp inhibitors and/or inducers during the study.
32. Subjects were deemed by the investigator to be unsuitable for enrollment in the study.

Phase I Clinical Trial New Zealand:
Inclusion criteria:
Each subject must meet all of the following criteria to be enrolled in this study:
1. The subject is a male or female 18 to 55 years of age, inclusive.
2. The subject has a body mass index 18 to 32 kg/m2, inclusive, and a total body weight ≥50 kg, inclusive, at screening.
3. The subject is considered by the investigator to be in good general health as determined by medical history, clinical laboratory test results, vital sign measurements, 12-lead ECG results, and physical examination findings at screening.
4. Female subjects of childbearing potential must be non-pregnant and non-lactating and must use one of the methods of contraception listed below for the duration of the treatment until at least 28 days after the last dose of the study drug, or be surgically sterile (ie, hysterectomy, bilateral tubal ligation, or bilateral oophorectomy) or postmenopausal (defined as amenorrhea 12 consecutive months and documented plasma follicle-stimulating hormone level >40 IU/mL). Female subjects must have a negative pregnancy test at screening and before the first dose of study drug.

Highly effective methods of contraception are those that result in a failure rate of less than 1% per year when used consistently. Examples are provided below:
a. Implant contraceptive (eg, Jadelle®)
b. Intrauterine device (IUD) containing either copper or levonorgestrel (eg, Mirena®)
c. Male sterilization with absence of sperm in the post-vasectomy ejaculate

OR an effective method that results in a failure rate of less than 5% to 10% per year. Examples are provided below:
d. Injectable contraceptive (eg, Depo Provera)
e. Oral contraceptive pill (combined hormonal contraceptive pill or progestogen-only 'mini-pill')

f. Vaginal contraceptive ring (eg, NuvaRing®)
Female subjects must also agree not to donate eggs, from dosing until at least 28 days after the last dose of study drug.
A male subject and his female partner who is of childbearing potential must agree to use one of the methods of contraception listed above for the duration of the treatment until at least 28 days after the last dose of the study drug. A male subject must also agree not to donate sperm, for the duration of the treatment until at least 28 days after the last dose of the study drug.
5. The subject agrees to comply with all protocol requirements.
6. The subject is able to provide written informed consent.
Inclusion criteria:

1. The subject has current evidence or history of clinically significant hematological, renal, endocrine, pulmonary, GI, cardiovascular, hepatic, psychiatric, neurologic, or allergic disease (including drug allergies, but excluding untreated, asymptomatic, seasonal allergies at time of dosing).
2. The subject has any condition possibly affecting drug absorption (eg, gastrectomy).
3. The subject has a history of cancer with the exception of adequately treated basal cell or squamous cell carcinoma of the skin.
4. The subject has supine blood pressure (BP) >140 mm Hg (systolic) or >90 mm Hg (diastolic), following at least 5 minutes of supine rest. If BP is >140 mm Hg (systolic) or >90 mm Hg (diastolic), the BP should be repeated 2 more times and the average of the 3 BP values should be used to determine the subject's eligibility at screening.
5. The subject has 12-lead ECG demonstrating corrected QT interval by Fridericia (QTcF) >450 msec, or a QRS interval >120 msec at screening. If QTcF exceeds 450 msec, or QRS interval exceeds 120 msec, the ECG should be repeated 2 more times and the average of the 3 QTcF (or QRS interval) values should be used to determine the subject's eligibility.
6. The subject has ANY of the following abnormalities in clinical laboratory tests at screening, as assessed by the study-specific laboratory and confirmed by a single repeat, if deemed necessary:
a. Serum creatinine (SCr) level above the upper limit of normal (ULN) or an estimated glomerular filtration rate (GFR) value <80 mL/min/1.73 m2 calculated with the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) formula and the absence of protein in urine, at screening.

b. Aspartate aminotransferase (AST) or alanine aminotransferase (ALT) values more than >1.5 × ULN.
c. Fasting glucose >110 mg/dL (6.1 mmol/L).
d. Total bilirubin >1.5 × ULN.
e. Hematological values outside the normal reference range for local laboratory results.
f. Positive fecal occult blood test at screening or at check-in (Day -1).
7. The subject has any medical history of disease that has the potential to cause a rise in total bilirubin over the ULN. Subjects with borderline clinical laboratory values outside the reference range may be included in the study if the investigator deems that the values are not clinically significant.
Note: Subjects with a history of Gilbert's syndrome may have a direct bilirubin measured and would be eligible for this study provided the direct bilirubin is <ULN.
8. The subject has a history of any lymphoproliferative disorder (such as Epstein Barr Virus-related lymphoproliferative disorder, as reported in some subjects on immunosuppressive drugs), history of lymphoma, leukemia, myeloproliferative disorders, multiple myeloma, or signs and symptoms suggestive of current lymphatic disease.
9. The subject has a history of relevant drug and/or food allergies (ie, allergy to any study drug or excipients, or any significant food allergy that could preclude a standard diet in the clinical unit).
10. The subject has a clinically significant infection currently or within 6 months of first dose of study drug (eg, those requiring hospitalization or parenteral antimicrobial therapy or opportunistic infections), or a history of chronic or recurrent infectious disease.
11. The subject has other severe acute or chronic medical or psychiatric condition including recent (within the past year) or active suicidal ideation or behavior or laboratory abnormality that may increase the risk associated with study participation or investigational product administration or may interfere with the interpretation of study results and, in the judgment of the investigator, would make the subject inappropriate for entry into this study.
12. The subject has or has had symptomatic herpes zoster or herpes simplex within 12 weeks, more than one episode of local herpes zoster, or a history (single episode) of disseminated zoster.
13. The subject has a positive test result for hepatitis B surface antigen, hepatitis C virus antibody, or human immunodeficiency virus (HIV) types 1 or 2 antibodies at screening.
14. The subject is a female who is pregnant or lactating.
15. The subject is a fertile male who is unwilling or unable to use a highly effective method of contraception as outlined in this protocol for the duration of the study and for at least 28 days after the last dose of investigational product.
16. The subject is unwilling or unable to comply with the lifestyle restrictions described in this protocol (Section 4.3.1).
17. The subject is a smoker or has used nicotine or nicotine-containing products (eg, snuff, nicotine patch, nicotine chewing gum, mock cigarettes, or inhalers) within 6 months before the first dose of study drug.
18. The subject has a positive test result for drugs of abuse or cotinine (indicating active current smoking) at screening or before the first dose of study drug.
19. The subject has used any prescription or over-the-counter medications (except paracetamol [up to 2 g per day]), including herbal supplements, within 14 days before the first dose of study drug. Nutritional supplements are allowed if unlikely to interfere with the study results and agreed by medical monitor and investigator.
20. The subject has consumed grapefruit or grapefruit juice, Seville orange or Seville orange-containing products (eg, marmalade), or alcohol-, caffeine-, or xanthine-containing products within 48 hours before the first dose of study drug.
21. The subject has used a known strong or moderate inhibitor or inducer of CYP1A2 within 4 weeks prior to Day 1 and through the last PK sampling point on Day 21 (only for Part C, DDI).
22. The subject will have vaccination with live virus, attenuated live virus, or any live viral components within the 2 weeks prior to the first dose of study drug or is to receive these vaccines at any time during treatment or within 8 weeks following completion of study treatment.
23. The subject has a positive test result for severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2). The subject has received the Coronavirus disease 2019 (COVID-19) vaccine within 2 weeks prior to the first dose of study drug or plans to receive a COVID-19 vaccine within 12 weeks after study drug dosing or has positive test for SARS-CoV-2 during screening or presence of COVID-19 symptoms within 4 weeks prior to Day -1.
24. The subject has undergone significant trauma or major surgery within 4 weeks of screening.

Ethics oversight

Phase 0 Clinical Trial Australia:
Institution: CMAX Clinical Research Center, 18A North Terrace, Adelaide SA 5000, Australia
A phase 0 clinical trial, also being the first-in-human clinical trial of INS018_055, was conducted in Australia (ACTRN12621001541897). This trial was conducted in accordance with the ethical principles of Good Clinical Practice, according to the International Council for Harmonisation (ICH) Harmonised Guideline E6(R2) Integrated Addendum to ICH E6(R1): Guideline for Good Clinical Practice ICH E6(R2), annotated with comments by the Australian Therapeutic Goods Administration (TGA; 2018). This study was approved by Bellberry Human Research Ethics Committee.

Phase I Clinical Trial China:
Institution: Zhejiang Xiaoshan Hospital, 728 Yucai N Rd, Xiaoshan District, Hangzhou, Zhejiang, China, 311202
Detailed information including facilities, inclusion/exclusion criteria could be found in link http://www.chinadrugtrials.org.cn/clinicaltrials.prosearch.dhtml (registration number: CTR20221542). This study was approved by Zhejiang Xiaoshan hospital Clinical trial ethics committee. The number of ethics committee approval letter is EC-2022102505.

Phase I Clinical Trial New Zealand:
Institution: New Zealand Clinical Research Center, Christchurch, New Zealand, Level 4, 264 Antigua Street, Central city , Christchurch, 8011 New Zealand
(registration number: NCT05154240).This study was approved by Northern B Health and Disability Ethics Committee. Ethics reference number for this protocol is 2021 FULL 11770.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Clinical (Australia): A total of 8 healthy adult subjects were enrolled. On day -1, baseline data was taken and safety were evaluated. On day 1, A single dose of 100 µg INS018_055 was given to each subject through IV injection. Blood was collected pre-dose and 5, 15, 30 minutes, then 1, 2, 4, 8 and 24 hours post-dose for PK evaluation. After dosing, individuals were monitored for safety until discharged in Day 2. The number of participants was selected to allow for evaluation of safety/tolerability and PK in this study and is consistent with standards of practice for Phase 0 studies. Clinical (New Zealand): 40 and 24 subjects were enrolled in Part A and Part B, respectively. Among 8 subjects per cohort, 6 subjects received treatment matching with 2 subjects received placebo (3:1). The determination of the subject number was not based on a formal statistical power calculation but based on clinical and practical considerations. Clinical (China): 24 and 24 subjects were enrolled in Part A and Part B, respectively. Among 8 subjects per cohort, 6 subjects received treatment matching with 2 subjects received placebo (3:1). The number of subjects was determined based on clinical and practical considerations, rather than on formal statistical certainty calculations. The total sample size of 56 subjects was considered sufficient for the purpose of the study based off of prior similar studies. The final enrollment is 48 since the trial is terminated due to COVID-9 outbreak. Animal study: For lung and kidney in vivo studies receiving oral dosing of INS018_055, the sample size in most groups are between 8-10. In some cases, when there were satellite groups receiving same treatment but for PK collection, some data (lung function, clinical observation, etc) were pooled together with efficacy groups. For the inhalation study, the sample sizes of each group are 12/group for vehicle and SOC group and 9/group for INS018_055 group with additional 3/group for PK collection. For skin fibrosis study, the sample size is 5/group. The determination on the sample sizes in in vivo study is based on the expected variations in the disease model and 3R principles (Replacement, Reduction and Refinement). In vitro: no special considerations.
Data exclusions	Clinical: There are no data exclusion except in below cases: In New Zealand phase 1 trial, one participant (101-066) discontinued treatment and was subsequently discontinued from the study due to a TEAE of influenza-like illness. Additionally in the New Zealand phase 1 trial, a few PK timepoints were referred to as 0 and not included in the study based on the pre established SAP that "BLQ value between 2 quantifiable concentrations will be set as missing." Animal Studies: No animals were excluded from this study unless noted in the supplementary information wherein mice were sacrificed early due to health concerns that are explicitly outlined in the methods section. This was less than 5 mice throughout the entire study which were sacrificed for humane reasons prior to endpoint data collection.
Replication	For patient number in clinical study, and sample size (animal number/group), the information could be found in manuscript, or the section of Life sciences study design-Sample size in this reporting summary file. Each clinical study and animal study were performed once. Two phase 1 trials were performed in different dosing regimen. Regarding preclinical animal studies, the efficacy of test compound was validated three times in bleomycin induced lung fibrosis models, though in different study designs. In vitro assays Assays on MRC-5, and A549, treated with INS018_055 were performed in 3 independent replicates. Assays on A549, knocked-by different shRNA targeting against TNIK were performed once, respectively: (Once by shTNIK-4 and once by shTNIK-1 with different virus dilution ratio). The results of these assays are consistent. Assay on HK-2 was performed once. EMT and FMT assays on primary lung cells were formed in 6 donors (3 healthy and 3 IPF patients), respectively. Assays performed on NHDF were performed in 5 replicates. For cell viability assays, A549 and HK2, once for each cell line. Cell viability on MRC-5 was performed twice. Data of the other assay is not shown in manuscript, the conclusion is similar, with CC50>100 µM.
Randomization	Clinical trials: For phase 0 study, there was no randomization and the study only has one cohort. For Phase I NZ and Phase I China trials, enrolled subjects were randomly assigned within each dose cohort to receive INS018_055 or a matched placebo. For all animal models utilized in this study, individual animals were randomly assigned to their respective groups.

Blinding

For all animal studies, the animals were randomized when grouping. The dosing and analysis were semi-blinded where different investigators experimentally handled the mice during model-induction (BLM, UUO, Skin fibrosis) while data analysis was performed by other investigators. This semi-blinded approach was required in order to assign animals to the correct grouping following disease-model induction since experimental animals that were not in the sham cohorts were visibly different to the sham mice but macroscopically similar across experimental groups.
In addition in UUO experiment, experimenters stated that during histological analysis, the tissue slides were coded a number for blind evaluation. Each number was generated using the RAND function of Excel software, sorted in ascending order and assigned the slides. The tissue slides were used for the following stains and evaluated by an experimenter.
Clinical trial (phase 0, Australia): Blinding was not used in phase 0 study, since it was a single cohort study.
Clinical trial (Phase 1, New Zealand): Part A and B of this clinical trial is a double-blind study design. The INS018_055 and matching placebo capsules were identical in appearance. The unblinded pharmacist was responsible for dispensing the study drug in capsules with identical appearance between INS018_055 and placebo, in a manner consistent with maintaining the blind. Reviews of unblinded safety data were performed to determine whether to move towards next dose cohort.
Blinding were maintained during dosing, sample collection and analysis.
Clinical trial (Phase1, China): INS018_055 has identical appearance to matched placebo capsules. The pharmacist was responsible for responsibility for dispensing the test drug in a manner that maintains blinding at all times.Blinding was maintained throughout the study period. If a subject becomes seriously ill or pregnant during the study period, blinding will be done only when it is known that administration of the test drug will affect the available treatment options for that subject. In the event of a medical emergency requiring identification of the individual subject receiving the test drug, the investigator should make every effort to contact the Medical Monitor within 24 hours prior to code unblinding to explain the need for code unblinding. The investigator is responsible for documenting the time, date, and reason for code unblinding, as well as the names of the participants.
(Detailed procedures could be found in protocols)

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Antibodies:
Western Blot: Primary antibodies: Antibody name Vendor Catalog # /Lot and dilution factor: Anti-human-fibronectin, Invitrogen, MA5-11981 (1:500); Anti-human-E-cadherin, BD Biosciences, 610182 (1:1000); Anti-human-N-cadherin (D4R1H) , Cell Signaling Technology (CST), 13116S (1:1000); Anti-human-phospho-smad2 (ser465/467)/smad3(ser423/425)(D27F4), CST, 8828S (1:1000); Anti-human-phospho-smad2 (ser465/467), CST, 3108S; Anti-human-smad2/3 (D7G7), CST, 12470S (1:1000); Anti-human-phospho-fak (tyr397) (D20B1) , CCST, 8556 (1:1000); Anti--human-FAK, CST, 3285 (1:1000); Anti--human-TNIK, CST, 32712 (1:1000); Anti-human-NF-κB p65 (D14E12), CST, 8242S (1:1000); Anti-human-Phospho-NF-κB p65 (Ser536) (93H1) (1:1000); CST, 3033S, Anti-human-beta-catenin, CST, 9562 (1:1000); Anti-human-Alpha-tubulin, Abcam, ab18251 (1:10000); Anti-human-HDAC2, Abcam, ab12169 (1:1000); Anti-human-Histone 3, Abcam, ab176842 (1:1000); anti-α-SMA antibodies (CST, #19245, 1:1000; #ab5694 and Santa Cruz, sc-56499, 1:500); GAPDH (CST, #2118S, 1:10000; Merk Millipore, MAB374, 1:2000).Secondary antibodies: goat Anti-Rabbit IgG(H+L), Abclonal, Lot 9300014001, 1: 8000; goat Anti-mouse IgG H&L, abcam, ab205719, 1:5000 or 20000; goat Anti-rabbit IgG H&L, abcam, ab205718, 1:20000)
Immunohistochemistry (IHC): Primary antibodies: anti-aSMA (Abcam, #ab5694, 1:400); Collagen I (Abcam, ab34710, 1:400; LSL co., LTD. cat#LB-1102, 1:2000). Secondary antibodies: goat anti-rabbit IgG HRP, DAKO, K4003, 10150251; Goat anti-rabbit IgG HRP, Vector Laboratories, Inc., Cat #PI-1000)
Immuno flurescence (IF): Anti-aSMA-Alexa Fluor 488 antibody (Abcam, Cat# ab184675, 1:600); anti-aSMA, Sigma-Aldrich, 1:100; anti-fibronectin 1 antibody (biohit healthcare, cat# 610001), 0.125 ug/mL in blocking buffer, DAPI (Invitrogen, Cat# D3571, 500 ng/mL)
Secondary antibodies: Donkey anti-mouse Alexa 546 (Invitrogen, cat# A10036).
ELISA kits: IL-1β, SINOBEST BIO, YX-E01291M; IL-4, SINOBEST BIO, YX-E00064M; IL-6, SINOBEST BIO, YX-E00066M; TNF-α, SINOBEST BIO, YX-E00104M; Fibronectin: Takara, cat# MK115 ; Procollagen type I C-peptide (PIP): Takara, cat# MK101 ; rat hydroxyproline (kinesisDx, K11-0512) ; collagen (Abcam, ab222942)

Validation

Validations of antibodies and assay kits are provided by manufacturer.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	Human fetal lung fibroblast cell line MRC-5 was purchased from ATCC (Cat#CCL-171) (cell viability assay performed in WuXi AppTec) and Shanghai Cell Bank, Chinese Academy of Sciences (α-SMA assay performed in BOJI) Human primary bronchial epithelial cells were derived from 3 IPF donors (IPF05, IPF06, and IPF08) and 3 healthy donors (Br285, Br311 and 410955). Human primary lung fibroblasts were derived from 3 IPF donors (IPF05, IPF06, IPF08) and 3 healthy donors (FB218, 03HF67101, FB2382). Human lung adenocarcinoma cell line A549 was purchased from ATCC (Cat#CCL-185). Human embryonic kidney 293T/17 was purchased from ATCC (Cat#CRL-11268). Human kidney cell line HK-2 was purchased from ATCC (Cat#CRL-2190). Normal human dermal fibroblasts (NHDF), Bioalternatives reference PF2
Authentication	No independent authentication of cell lines is available. The internal authentication of depends on the internal procedure of Contract research organizations. Assays using A549, 293T/17 and HK-2: WuXi AppTec (China); Assays using human primary bronchial epithelial cells and human primary lung fibroblasts: Charles Rivers (Netherlands); Assays using MRC-5: WuXi AppTec (China) and Guangzhou Boji Medical Biotechnological (China); Assays using NHDF: Bioalternatives (France)
Mycoplasma contamination	All cell lines used were checked and cleared for mycoplasma contimation at the following contract research organizations: WuXi AppTec (China), Charles Rivers (Netherlands), Guangzhou Boji Medical Biotechnological (China), Bioalternatives (France).
Commonly misidentified lines (See ICLAC register)	No

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	Studies performed on bleomycin-induced lung fibrosis mouse model: 7-8 weeks, C57BL/6 mice The study performed on bleomycin-induced lung fibrosis rat model: 8-9 weeks, Sprague-Dawley(SD) Rat Study performed on lipopolysaccharide (LPS)-induced acute lung injury mouse model: 8 weeks, C57BL/6 mice The study on bleomycin induced skin fibrosis rat model: 5-6 week, SD rats The study performed on unilateral ureteral obstruction(UUO) model: 7 weeks, C57BL/6 mice
Wild animals	No wild animals were used in this manuscript.
Reporting on sex	Studies performed on bleomycin-induced lung fibrosis mouse model: male The study performed on bleomycin-induced lung fibrosis rat model: male Study performed on LPS-induced acute lung injury mouse model: male The study on bleomycin induced skin fibrosis rat model: male The study performed on unilateral ureteral obstruction(UUO) model: female
Field-collected samples	No Field-collected samples were used in this manuscript.
Ethics oversight	All animal studies were ethically and humanely conducted following IACUC, IRB or relevant animal handling ethics organizational guidelines by our partnering CROs. Murine LPS and BLM studies were carried out by HD Biosciences (Shanghai) which were approved under the following AUF protocol numbers: AUF#146 and AUF#117. These AUF protocols were approved by the Institutional Animal Care and Use Committee at HD Biosciences and the studies were conducted at an AAALAC-accredited facilities. The BLM fibrosis study in rats that were administered INS018_055 as an inhalable agent was carried out by JOINN Laboratories (Suzhou). This study's IACUC approved procedural number is S-ACU22-1144. JOINN Laboratories is fully accredited by the AAALAC, and the study number provided adheres to the regulations and rules laid out by the Institutional Animal Care and Use Committee. The UUO Kidney study was performed by SMC Laboratories, Inc. (Japan). The Animal Care and Use Committee-approved protocol number is U32. All animals used in this study were housed and cared for in accordance with the Japanese Pharmacological Society Guidelines for Animal Use. The skin fibrosis model was carried out by TheraIndx LifeSciences PVT LTD (India). This study was performed using protocols approved by the Institutional Animals Ethics Committee (IAEC) of the test facility which was designed under CPCSEA guidelines for animal care. The registration number for this protocol is (No. 1852/PO/Rc/S/16/CPCSEA. For all the aforementioned studies, animals were group-housed (<5 per cage) in a temperature-controlled facilities (20-26 C) with 12h light/12h dark light cycles. All animals had ad libitum access to drinking food and water.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	Clinical Trial Phase 0 (Australia): ACTRN12621001541897 Clinical trial (New Zealand): NCT05154240

Clinical trial (China): CTR20221542

Study protocol

Trial information has also been provided for all 3 trials in Supplementary Information 9,10, and 11, respectively.

Phase 0 Study (Australia)
A phase 0 micro dosing clinical trial was conducted in Australia (ACTRN12621001541897). The trial can be found using the clinical trial number using this database https://www.anzctr.org.au/TrialSearch.aspx

Summary: A total of 8 healthy adult subjects were enrolled. On day -1, baseline data was taken and safety were evaluated. On day 1, A single dose of 100 µg INS018_055 was given to each subject through IV injection. Blood was collected pre-dose and 5, 15, 30 minutes, then 1, 2, 4, 8 and 24 hours post-dose for PK evaluation. After dosing, individuals were monitored for safety until discharged in Day 2.

Phase I study (New Zealand):
Full study protocol of clinical trial (NCT05154240) can be found in www.clinicaltrials.gov.

Summary:This is a phase I, randomized, double-blind, placebo-controlled, oral single and multiple ascending doses, parallel group and exploratory drug-drug interaction study to evaluate the safety, tolerability, pharmacokinetics, and interaction potential of INS018_055 in healthy subjects, conducted in one clinical site in New Zealand. Protocol and its amendments were approved by the IRB before implementation unless proceeding with the changes was in the subject's best interest. Planned enrollment was a total of 80 healthy male and female subjects (40, 24, 16 subjects in Part A, B and C, respectively), aged from 18 to 55 years, with body mass index (BMI) of from 18 to 32 kg/m2, and a total body weight ≥50 kg, considered by the investigator to be in good general health based at screening.

Phase I study (China):
Full study protocol can be found at http://www.chinadrugtrials.org.cn/clinicaltrials.prosearch.dhtml (registration number: CTR20221542).

Summary:

This is a phase I, randomized, 2-part (Part A and Part B), double-blind, placebo-controlled, dose-escalation study designed to evaluate the safety, tolerability, and PK characteristics of healthy subjects receiving oral administration of INS018_055.
The study will include a screening period, enrollment, treatment period, and end-of-study (EOS) visit.

Data collection

Clinical Trial (Austrialia): A phase 0 clinical trial, also being the first-in-human clinical trial of INS018_055, was conducted in Australia (ACTRN12621001541897). This trial was conducted in accordance with the ethical principles of Good Clinical Practice, according to the International Council for Harmonisation (ICH) Harmonised Guideline E6(R2) Integrated Addendum to ICH E6(R1): Guideline for Good Clinical Practice ICH E6(R2), annotated with comments by the Australian Therapeutic Goods Administration (TGA; 2018). This study was approved by Bellberry Human Research Ethics Committee and data collected at CMAX Clinical Research Center, 18A North Terrace, Adelaide SA 5000, Australia.

Clinical trial (New Zealand): This study was licensed as registration number: NCT05154240.This study was approved by Northern B Health and Disability Ethics Committee and the Ethics reference number for this protocol is 2021 FULL 11770. The phase 1 clinical trial was conducted atN ew Zealand Clinical Research Center, Christchurch, New Zealand, Level 4, 264 Antigua Street, Central city , Christchurch, 8011 New Zealand. The first subject first dose was in Feb 21st, 2022, and last subject contact was in Sep 30th, 2022.

Clinical trial (China): The phase 1 clinical trial was conducted at Zhejiang Xiaoshan Hospital, 728 Yucai N Rd, Xiaoshan District, Hangzhou, Zhejiang, China, 311202. The first subject signed ICF: in Jul 22nd, 2022, and last subject visit was in Jan 13th, 2023.

Information on the exact data collected are outlined in the relevant sections of this reporting summary and have been included in the manuscript as part of the supplementary information for the referenced clinical trials.

Outcomes

Phase 0 Clinical Trial (Australia):
Primary Objective:
The primary objective of the study was to determine the PK of plasma INS018_055 after a single IV microdose administered to healthy participants.
Primary endpoints:
The primary endpoints of this study were:
•Volume of distribution of INS018_055 in healthy participants.
•Elimination half-life (t1/2) of INS018_055.
•Clearance of INS018_055 in healthy participants.

Phase 1 Clinical trial (New Zealand):
•The primary objective of this study was to assess the safety and tolerability of single and multiple oral escalating doses of INS018_055 administered to healthy subjects. The secondary objectives of this study were as follows:

• To determine the pharmacokinetics (PK) of INS018_055 following single and multiple oral escalating doses in healthy subjects.
• To assess the effect of food on the PK of INS018_055 following an oral dose.
The exploratory objectives of this study were as follows:
• To determine the effect of single and multiple oral escalating doses of INS018_055 on the circulating cluster of differentiation (CD)4 + and CD8+ subpopulation of T cells (as a measure of pharmacological activity) in healthy subjects.
• To determine the effect of multiple oral escalating doses of INS018_055 on the levels of phospho-nuclear factor kappa B (NF-κB) p65 (Ser536), phospho-SMAD family member (Smad)2 (Ser465/467), interleukin (IL)-6, transforming growth factor beta (TGF-β), matrix metalloproteinase (MMP)-2, MMP-7, MMP-9, total NF-κB p65, and total Smad2 in healthy subjects. The results of these exploratory objective analyses will be presented separately from the clinical study report (CSR).
• To assess the interaction potential of INS018_055 with a cytochrome P450 (CYP)1A2 substrate (caffeine).
• To characterize the metabolite profiles of INS018_055 in plasma and urine following the highest dose level in the multiple ascending dose (MAD) study part (Cohort 8 in Part B). The results of these exploratory objective analyses will be presented separately from the CSR.

Phase 1 Clinical trial (China):
The primary objectives of this study are:
• The primary objective of this study is to assess the safety and tolerability of single and multiple oral ascending doses of INS018_055 administered to healthy subjects
The secondary objectives of this study were:
• To determine the pharmacokinetic (PK) profile of INS018_055 following single and multiple oral escalating doses in healthy subjects.
• Assessment of the effect of single and multiple dose escalation oral administration of INS018_055 on biomarkers (T cell CD4+ and CD8+ subpopulations, interleukin [IL]-6, transforming growth factor [TGF]-β, matrix metalloproteinases [MMP-2, MMP-9, MMP-7]) in healthy subjects (as a measure of pharmacological activity)