

## ARTICLE OPEN



# Wet-dry-wet drug screen leads to the synthesis of TS1, a novel compound reversing lung fibrosis through inhibition of myofibroblast differentiation

Nadja Anneliese Ruth Ring<sup>1,10</sup>, Maria Concetta Volpe<sup>1,2,10</sup>, Tomaž Stepišnik<sup>3</sup>, Maria Grazia Mamolo<sup>4</sup>, Panče Panov<sup>5</sup>, Dragi Kocev<sup>3</sup>, Simone Vodret<sup>1</sup>, Sara Fortuna<sup>4</sup>, Antonella Calabretti<sup>4</sup>, Michael Rehman<sup>1,5</sup>, Andrea Colliva<sup>1</sup>, Pietro Marchesan<sup>1</sup>, Luca Camparini<sup>6</sup>, Thomas Marcuzzo<sup>2</sup>, Rossana Bussani<sup>2</sup>, Sara Scarabellotto<sup>4</sup>, Marco Confalonieri<sup>2</sup>, Tho X. Pham<sup>7</sup>, Giovanni Ligresti<sup>7</sup>, Nunzia Caporarello<sup>8</sup>, Francesco S. Loffredo<sup>6,9</sup>, Daniele Zampieri<sup>4</sup>, Sašo Džeroski<sup>3</sup> and Serena Zacchigna<sup>1,2</sup>✉

© The Author(s) 2021

Therapies halting the progression of fibrosis are ineffective and limited. Activated myofibroblasts are emerging as important targets in the progression of fibrotic diseases. Previously, we performed a high-throughput screen on lung fibroblasts and subsequently demonstrated that the inhibition of myofibroblast activation is able to prevent lung fibrosis in bleomycin-treated mice. High-throughput screens are an ideal method of repurposing drugs, yet they contain an intrinsic limitation, which is the size of the library itself. Here, we exploited the data from our “wet” screen and used “dry” machine learning analysis to virtually screen millions of compounds, identifying novel anti-fibrotic hits which target myofibroblast differentiation, many of which were structurally related to dopamine. We synthesized and validated several compounds *ex vivo* (“wet”) and confirmed that both dopamine and its derivative TS1 are powerful inhibitors of myofibroblast activation. We further used RNAi-mediated knock-down and demonstrated that both molecules act through the dopamine receptor 3 and exert their anti-fibrotic effect by inhibiting the canonical transforming growth factor  $\beta$  pathway. Furthermore, molecular modelling confirmed the capability of TS1 to bind both human and mouse dopamine receptor 3. The anti-fibrotic effect on human cells was confirmed using primary fibroblasts from idiopathic pulmonary fibrosis patients. Finally, TS1 prevented and reversed disease progression in a murine model of lung fibrosis. Both our interdisciplinary approach and our novel compound TS1 are promising tools for understanding and combating lung fibrosis.

*Cell Death and Disease* (2022)13:2; <https://doi.org/10.1038/s41419-021-04439-4>

## INTRODUCTION

Fibrosis-related mortality is on the rise. Mortality data from the world health organization between 2001 and 2013 shows a general trend of increasing mortality for lung fibrosis in the EU [1]. This can be explained through several factors, including an aging population and insufficient treatment options to prevent or cure these disorders, due to the large diversity in both aetiology and symptoms [2].

Inflammatory cells are activated in response to tissue damage, leading to the proliferation of fibroblasts and their differentiation into activated myofibroblasts, as an attempt to repair the injury. In a later phase, inflammation decreases, and a scar is formed [3]. A major inducer of myofibroblast activation is transforming growth factor  $\beta$  (TGF $\beta$ ) [4], which increases the secretion of extracellular matrix components, and the expression of the contractile protein alpha-smooth muscle actin

( $\alpha$ SMA), currently considered to be the most reliable marker of this cell type [5].

While many therapeutic approaches focus on the inflammatory component of fibrosis, the use of corticosteroids as anti-inflammatory therapy has been shown to be ineffective and fraught with side effects [6, 7]. Pirfenidone and nintedanib are two drugs recently approved for the treatment of lung fibrosis, and while their exact mechanism of action has not been described, part of their anti-fibrotic activity is through suppression of inflammation [8]. Aside from inflammation, the key role of activated myofibroblasts in the development of fibrotic tissue is emerging as a crucial factor [5]. One therapeutic approach targeting these cells is the development of Bcl-2 homology domain mimetics, which target the anti-apoptotic Bcl2 protein and induce the apoptosis of myofibroblasts, bringing with it obvious risks and side effects [9].

<sup>1</sup>Cardiovascular Biology Laboratory, International Centre for Genetic Engineering and Biotechnology (ICGEB), Trieste, Italy. <sup>2</sup>Department of Medical, Surgical and Health Sciences, University of Trieste, Trieste, Italy. <sup>3</sup>Jožef Stefan Institute, Ljubljana, Slovenia; Jožef Stefan International Postgraduate School, Ljubljana, Slovenia. <sup>4</sup>Department of Chemical and Pharmaceutical Sciences, University of Trieste, Trieste, Italy. <sup>5</sup>Yale School of Medicine Section of Nephrology, New Haven, CT, USA. <sup>6</sup>Molecular Cardiology Unit, International Centre for Genetic Engineering and Biotechnology (ICGEB), Trieste, Italy. <sup>7</sup>Arthritis Center, Department of Medicine, Boston University School of Medicine, Boston, MA, USA. <sup>8</sup>Department of Physiology and Biomedical Engineering, Mayo Clinic, Rochester, MN, USA. <sup>9</sup>Department of Translational Medical Sciences, University of Campania “Luigi Vanvitelli”, Naples, Italy. <sup>10</sup>These authors contributed equally: Nadja Anneliese Ruth Ring, Maria Concetta Volpe. ✉email: [Serena.Zacchigna@icgeb.org](mailto:Serena.Zacchigna@icgeb.org)  
Edited by Professor Boris Zhivotovskiy

Received: 6 July 2021 Revised: 18 November 2021 Accepted: 29 November 2021

Published online: 17 December 2021

While myofibroblasts have been the target of therapies, few approaches have attempted to interfere with fibroblast differentiation into activated myofibroblasts [10]. In a previous study of our laboratory, we performed a high-throughput screen using a library of FDA-approved chemical compounds and observed their effect on ex vivo lung fibroblasts. Using the top hits of this screen we demonstrated that the inhibition of myofibroblast activation, shown by the accumulation of the contractile protein  $\alpha$ SMA is able to prevent the development of lung fibrosis in a murine model of lung fibrosis [11].

Here, we apply an interdisciplinary wet-dry-wet approach. We exploited data generated in our previously performed “wet” high-throughput screen, and used “dry” machine learning to virtually screen the ChEMBL database. ChEMBL is a publicly available database comprising a large collection of drug-like small molecules, including information about their structure, properties, and bioactivity [12, 13]. This enabled us to overcome the intrinsic limit of the compound library size in the original screen and extrapolate potential hits among millions of compounds, thereby increasing both coverage and value of the screen. Finally, we validated the predicted anti-fibrotic compounds in both primary cells and a mouse model of lung fibrosis (“wet”). Using this wet-dry-wet approach, we selected and synthesized several compounds and identified T51 as the most promising anti-fibrotic agent.

## RESULTS

### Machine learning analysis of a previous drug screen identifies novel anti-fibrotic candidates

We recently performed a high content high throughput screen (HC-HTS), using a library of 640 FDA-approved compounds [11]. The screen used primary fibroblasts from  $\alpha$ SMA-RFP/COLL-EGFP mice [14], which express red fluorescent protein (RFP) under the alpha smooth muscle actin ( $\alpha$ SMA) promoter and enhanced green fluorescent protein (EGFP) under the collagen  $\alpha$ 1(I) promoter. These cells differentiate into myofibroblasts in culture, either spontaneously or upon exposure to transforming growth factor beta (TGF $\beta$ ). This screen led to the identification of haloperidol as an inhibitor of myofibroblast activation and a potent anti-fibrotic agent in vivo.

To maximize the quality and the coverage of our “wet” screen in primary fibroblasts, we applied a machine learning approach to its results and learned a model to predict the anti-fibrotic effect of compounds not included in the original screen. Machine learning algorithms have the ability to learn and improve with experience presented in the form of data. Data consist of examples that typically reside in a table, where each row represents an example: in our case, each row represents a chemical compound. Each column from the table represents a specific property or descriptor of the compounds, in this case, extended connectivity fingerprints (ECFP). ECFPs are often used and typically prove to be well suited for predicting biological properties of molecules. They describe compounds with binary vectors, and each vector component corresponds to a set of substructures. If that component has a value of 1, this means that at least one of those substructures is present in the molecule. The dataset used to learn the models is usually called the training set, and the output is a predictive algorithm that can then be applied to the so-called testing set.

In this study, we used our existing “wet” screening data, i.e., experimental measurements of average RFP/EGFP intensity, as our training set. The predictive model has the form of a bagging ensemble of predictive clustering trees (Kocov, Vens et al. 2013), which simultaneously predict the effect of a compound on both RFP and EGFP intensity. Performance of the model was measured by mean absolute error, with performance on unseen compounds estimated via 10-fold cross-validation. The mean absolute errors of the model were estimated to be 0.0836 for EGFP and 0.0871 for RFP.

To discover novel compounds able to lower  $\alpha$ SMA expression, and thus, presumably, myofibroblast activation, we used the learned model to virtually screen our testing set: the ChEMBL database (containing over 1.7 million compounds in total). For each compound, its ECFP was obtained and fed to the predictive model to obtain a prediction of the compound effect on RFP/EGFP intensity. We performed a subsequent filtering step to filter the weak predictions and make a priority list for further testing and examination.

This algorithm selected 616 candidates, which were predicted to reduce  $\alpha$ SMA intensity by at least 25% compared to controls (Fig. 1A). After thresholding, compounds present in the original screen were removed from the candidate list. The remaining compounds were divided into two groups. Group 1 contained compounds that have already reached clinical trial stage, while novel compounds were included in Group 2. Group 1 was composed of 27 compounds, of which 24 were corticosteroids. Based on their potent anti-inflammatory activity, corticosteroids have been extensively considered for use as anti-fibrotic drugs [15–18]. However, they have invariably failed, exhibiting severe side effects [6, 7, 19] and therefore we did not consider them further. The three remaining compounds in group 1 were isotretinoin, alitretinoin, and fluvastatin, which belong to well-known classes of drugs and ranked relatively high in our original HTS, namely retinoids and statins. On the one hand, this renders them of little interest for our purpose, while on the other it confirms the power of our virtual screen in identifying novel therapeutic candidates.

Thus, we focused on Group 2, made up of 589 novel compounds. We used simplified molecular-input line-entry system (SMILES) in the online tool ChemMine to group them according to their structural similarity to known drugs [20]. The majority ( $n = 545$ ) contained structural components of FDA-approved drugs which were contained in our original screen compound library. Specifically, 327 compounds were related to the corticosteroid dexamethasone, 219 were related to either amiodarone, celecoxib, chlorambucil, ciprofloxacin, indomethacin, losartan, melatonin, or paclitaxel. Interestingly, the remaining 44 compounds were related to drugs unrepresented in the original HTS, including acyclovir, prasterone, progesterone, thiamine ion, topiramate, zanamivir, and dopamine. In particular, structural components of dopamine were contained in 43% of these compounds. In total, three candidates were chosen for validation in vitro, dopamine, topiramate, and zanamivir, as they have never been specifically considered in relation to fibrotic disorders.

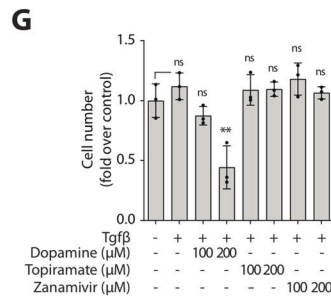
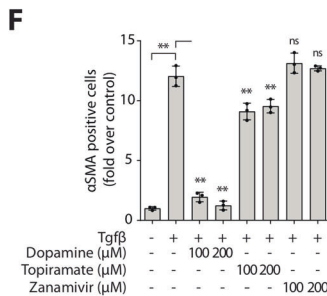
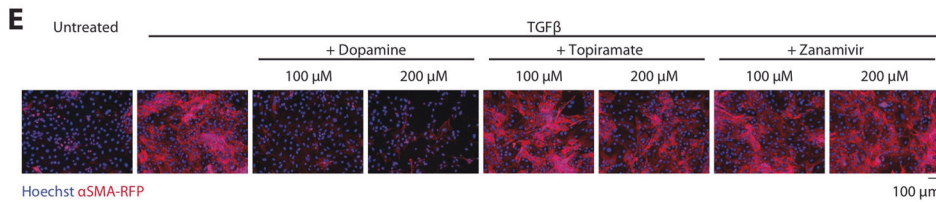
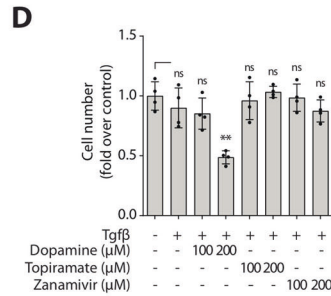
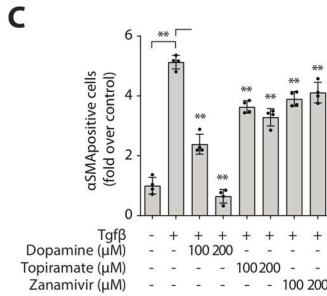
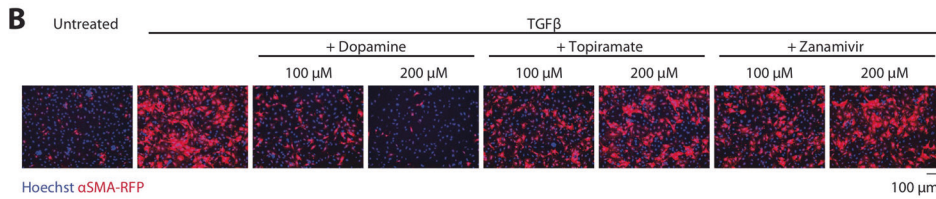
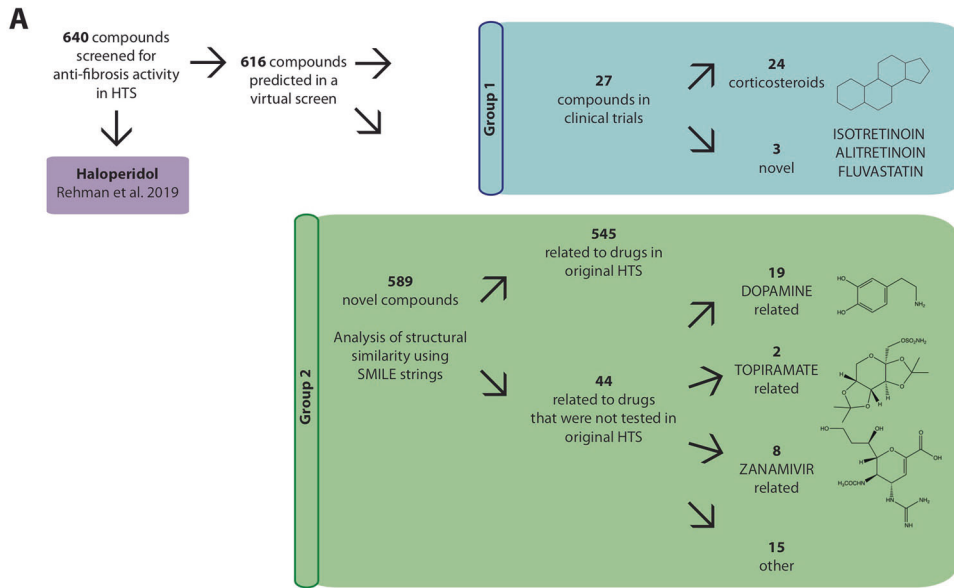
### Dopamine inhibits myofibroblast activation

To validate the efficacy of the selected drugs in reducing  $\alpha$ SMA expression, primary murine cardiac fibroblasts were isolated from  $\alpha$ SMA-RFP mice, as in our original screen. Cells were stimulated for 48 h with TGF $\beta$  and treated with dopamine, topiramate, or zanamivir [21, 22]. At 100  $\mu$ M dopamine was clearly the most effective one in reducing  $\alpha$ SMA expression (Fig. 1B, C). The effect was more evident at 200  $\mu$ M, but this dose resulted in reduced cell viability (Fig. 1D).

Given the current need for new therapeutics for lung fibrosis, we tested whether a similar effect was exerted in primary lung fibroblasts from  $\alpha$ SMA-RFP mice and confirmed the results (Fig. 1E–G). In this case, the potent anti-fibrotic effect of dopamine was even more evident at 100  $\mu$ M. These results confirm the ability of our virtual screen to identify interesting leads and justify further in-depth investigation on the mechanism of action of dopamine.

### Dopamine reduces myofibroblast activation through the dopamine receptor 3

In order to elucidate the receptor through which dopamine is able to counteract fibroblast differentiation into myofibroblasts, we tested the expression levels of the various dopamine



receptors (Drd) in primary fibroblasts from mouse heart and lung, using cerebellum as a positive control. We found that among the Drd family members, Drd3 and Drd4 were primarily expressed by both types of fibroblasts (Fig. 2a). To assess their functional relevance in mediating the anti-fibrotic effect of

dopamine, we silenced either receptor using specific siRNAs. The knock down of Drd3 (efficacy shown in Supplementary Fig. 1) significantly inhibited the effect of dopamine (Fig. 2b, c), pointing to this receptor as the major mediator of dopamine activity.

**Fig. 1 Virtual screen identifies dopamine as a powerful anti-fibrotic agent in primary murine fibroblasts.** **A** Flowchart showing the data processing of the machine learning predictions. In a previous publication, 600 compounds were screened, identifying haloperidol as a potent anti-fibrotic drug (purple). In the virtual screen, 616 compounds are predicted to decrease  $\alpha$ SMA expression by 25% and are divided into two groups. Group 1 (blue) contains all predicted compounds which are clinically approved. Group 2 (green) contains unapproved, novel, predicted compounds. These were further analyzed to identify structurally related, clinically approved, drugs. Of these, 545 compounds were structurally related to drugs present in the original “wet” screen, while the remaining 44 drugs were related to dopamine (19/44), topiramate (2/44), zanamivir (8/44), and others (15/44). **B** Murine heart fibroblasts were treated ex vivo with TGF $\beta$  and dopamine, topiramate, and zanamivir. Representative images are shown (Hoechst in blue,  $\alpha$ SMA-RFP in red). **C** Quantification of the number of  $\alpha$ SMA positive cells (shown as fold over control). **D** Quantification of the total number of cells per condition (shown as fold over control). **E** Murine lung fibroblasts were treated ex vivo with TGF $\beta$  and dopamine, topiramate, and zanamivir. Representative images are shown (Hoechst in blue,  $\alpha$ SMA-RFP in red). **F** Quantification of the number of  $\alpha$ SMA positive cells (shown as fold over control). **G** Quantification of the total number of cells per condition (shown as fold over control). All data are shown as mean  $\pm$  SD. Statistical significance was determined using a one-way ANOVA followed by Dunnett’s multiple comparison test, \* $P < 0.05$ , \*\* $P < 0.01$ ,  $n = 3$ .

In fact, when we used Western blotting to confirm the protein expression of Drd3 and Drd4 we found that while both are found at detectable levels in brain and cerebellum extracts, only the former can be detected in primary murine lung fibroblasts (Supplementary Fig. 2). This strengthens the evidence that dopamine acts through Drd3, as it is expressed on both RNA and protein level.

We then tested our compounds using the bleomycin model of lung fibrosis applied to the COLL-EGFP transgenic mice. Animals were treated with dopamine through intra-tracheal delivery two days after bleomycin administration and sacrificed at two weeks. As shown in Fig. 2d, e, bleomycin injection resulted in strong up-regulation of collagen expression by lung fibroblasts, together with expression of  $\alpha$ SMA in the same cells. Dopamine treatment significantly blunted this response, reducing the levels of both fibrotic markers, especially  $\alpha$ SMA, in lung tissue. However, reduced myofibroblast differentiation in long term dopamine-treated lungs did not significantly alter the development of fibrosis, as assessed by Masson Trichrome staining and quantified using the Ashcroft score (Fig. 2f, g). This was not improved even when dopamine was administered repeatedly, at 7, 14, and 21 days following bleomycin administration (Fig. 2h, i).

Collectively, these results indicate that dopamine effectively reduces myofibroblast differentiation ex vivo through Drd3. However, it fails to significantly inhibit bleomycin-induced lung fibrosis in vivo, possibly because of residual collagen expression and/or its quick degradation [23].

### A novel, dopamine-like derivative inhibits the activation of lung myofibroblasts via Drd3

To overcome the limitations of dopamine as a potential therapeutic, we aimed at synthesizing related compounds, which could be equally effective in reducing myofibroblast differentiation, but more stable and capable of reducing fibrosis in vivo. We considered the 19 dopamine-related compounds identified by our machine learning approach and chose the least sterically active and therefore expected to be the most stable one (CHEMBL145929). Based on its structure, we designed six novel derivatives, named TS1-6, shown in Fig. 3A. While dopamine is a primary amine and is thus easily oxidized, the novel derivatives are all secondary amines. This renders them more stable and suitable for in vivo administration.

We comparatively evaluated the efficacy of TS1-6 to reduce the expression of both  $\alpha$ SMA and collagen in primary lung fibroblasts stimulated with TGF $\beta$ . As shown by the images in Fig. 3B and relative quantifications in Fig. 3c–e, TS1 was the most effective on both parameters and not toxic at 100  $\mu$ M. We then produced TS1 as a hydrochloride salt, which can be dissolved in water and can therefore be easily administered in vivo. To confirm efficacy and define the optimal dose for TS1-HCl in vivo administration, we tested multiple doses in the 100–600  $\mu$ M range. The 300  $\mu$ M concentration significantly reduced both  $\alpha$ SMA and collagen

protein expression without overt toxicity (Fig. 3f–i) and was therefore chosen for subsequent experiments.

### TS1 is a substrate of both murine and human Drd3

We confirmed the increased stability of TS1 compared to dopamine by HPLC analysis. While dopamine was rapidly degraded within the first 48 h, TS1 was totally stable in solution (Fig. 4A) and remained intact even after 30 days (not shown).

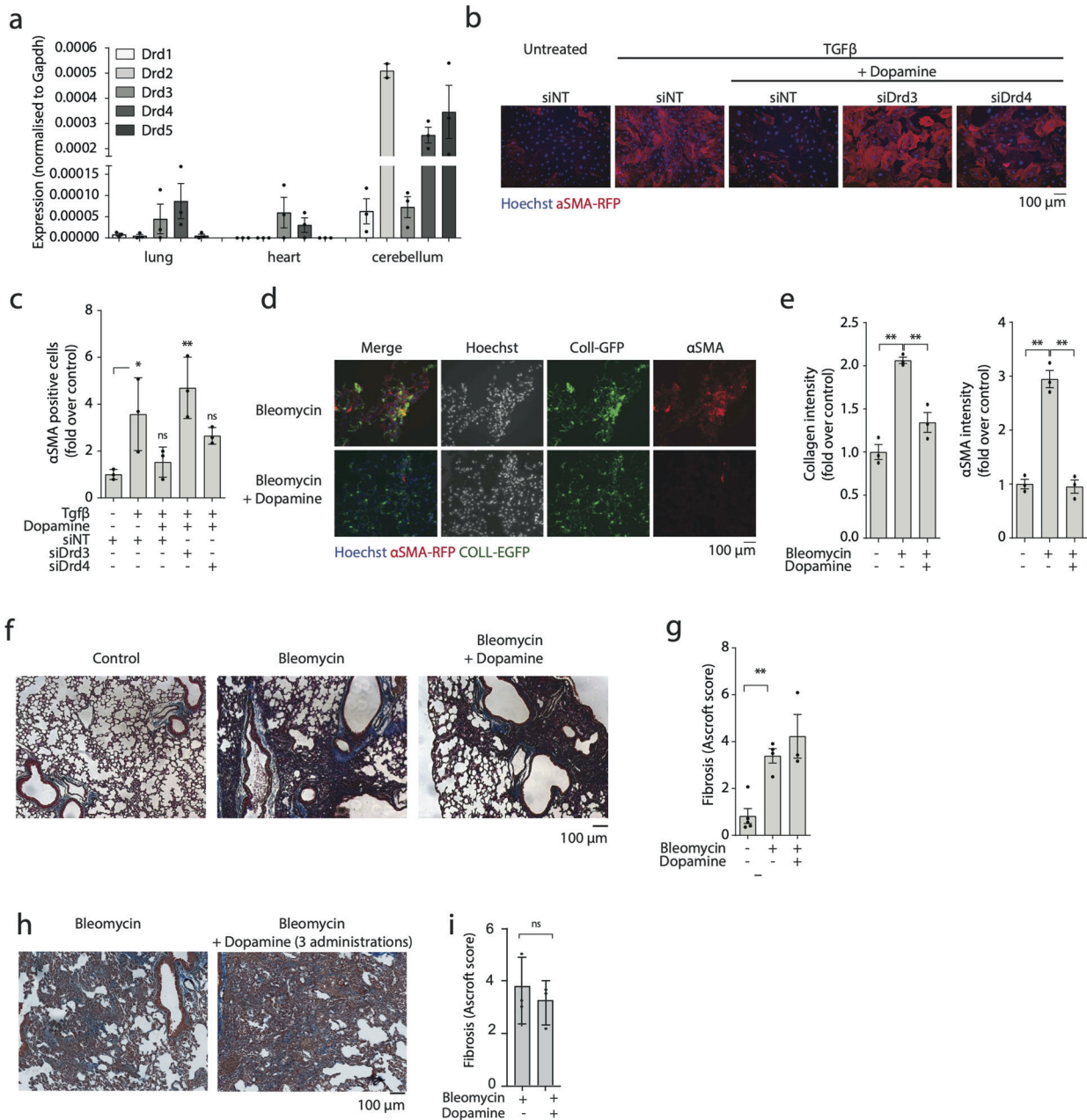
To confirm whether TS1 acts through the same receptor as dopamine, we silenced the two dopamine receptors expressed by lung fibroblasts (Drd3 and Drd4). While both non-targeting and anti-Drd4 siRNAs did not affect the ability of TS1 to block myofibroblast differentiation in the presence of TGF $\beta$ , the down-regulation of Drd3 completely abolished TS1 activity (Fig. 4B, C). Similar to dopamine, TS1 also acts on the canonical TGF $\beta$  pathway, as detected by a luciferase reporter of Smad 2/3 activation (Fig. 4D).

To further validate TS1 as a substrate of Drd3 through docking experiments, we first built a 3D model for the mouse receptor. The murine extracellular domain, made up of 272 residues, has never been crystallized. However, the structure of the human receptor is known [24] and differs from the murine protein by only 11 amino acids. Thus, we could model the latter with reasonable accuracy by homology (Fig. 4E, F).

TS1 was docked to the murine receptor (Supplementary Fig. 3) with a docking box covering the whole space enclosed by its helices (Supplementary Fig. 4A, B). Human Drd3 was originally crystallized with its substrate eticlopride, which was used to benchmark the method (Supplementary Fig. 4C, D). Interestingly, TS1 occupies the same binding site as eticlopride, and appears tightly wrapped by 12 residues (highlighted in Fig. 4E, F). These residues form 11 hydrophobic contacts, due to their aromatic side chains, and a hydrogen bond (2.8 Å) between Ser192 and the TS1 nitrogen (Fig. 4E, F). The position assumed by TS1 in the Drd3 channel shows that the molecule does not interact with any of the residues which differ from the human receptor (in yellow in Fig. 4E, F), suggesting that our compound could be active also in human cells.

Further docking experiments confirmed TS1 as a substrate of the human Drd3. Specifically, we identified two conformations of particular interest, suggesting that TS1 could bind more strongly to the human Drd3 than to the murine isoform. The conformation which was sampled most frequently overlapped with the one identified for the murine model (RMSD 0.3 Å). The second conformation showed more favourable interaction energies but was less populated (148 vs 892 conformers out of 2000) (Supplementary Fig. 5). While the static model predicts a strong binding of TS1 to both the murine and human receptor, it is difficult to predict the efficiency of entry into the receptor. Specifically, there is a variation in four amino acids between the murine and human receptor in a flexible loop that could influence the accessibility of the binding site.



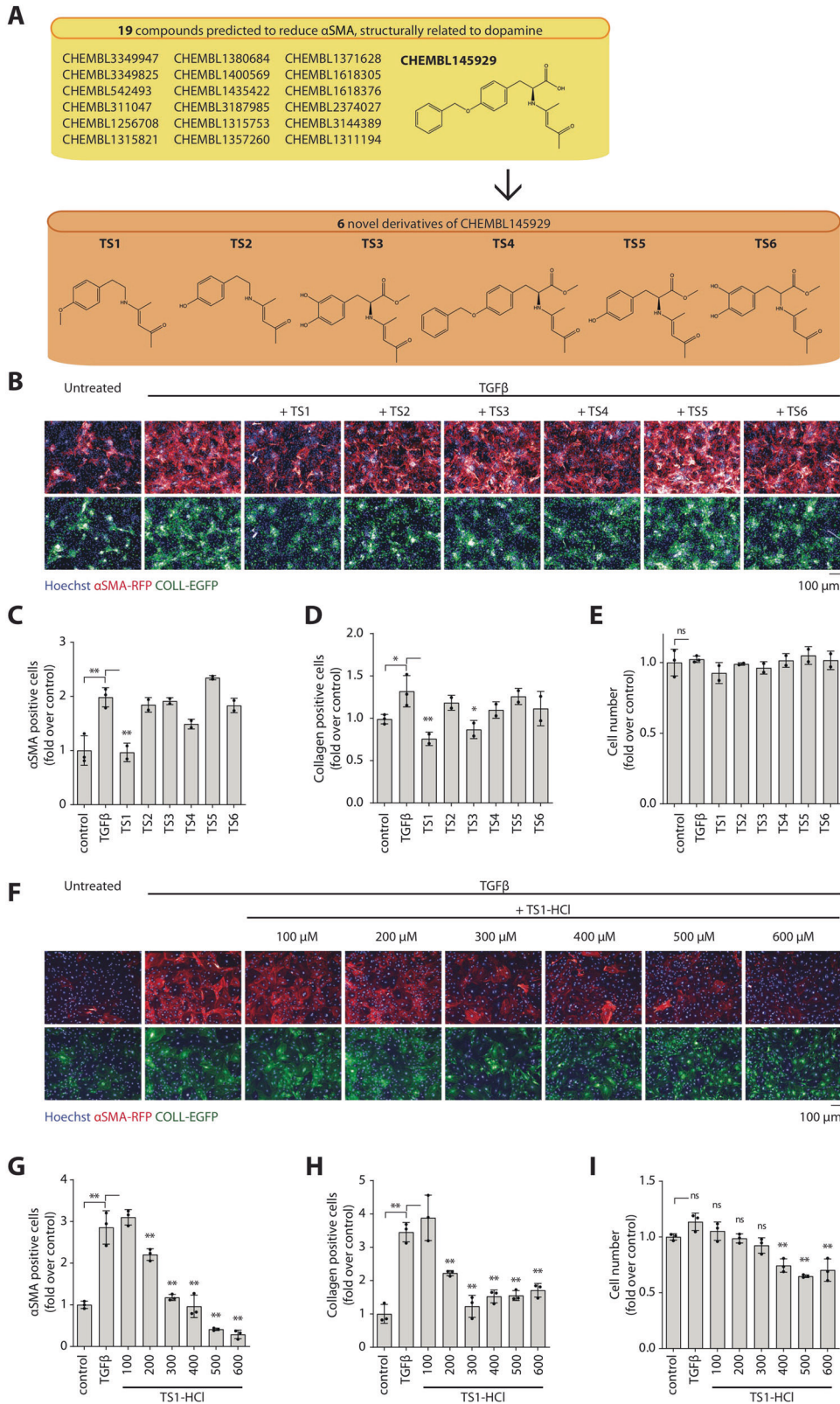


**Fig. 2 Dopamine inhibits lung myfibroblast differentiation by acting through Drd3 but is unable to prevent bleomycin-induced lung fibrosis in vivo.** **a** RT-PCR analysis shows the abundance of each dopamine receptor in fibroblasts from the murine lung and heart, using cerebellum lysate as a positive control. Results are normalized to *Gapdh* expression. **b** Murine lung fibroblasts were treated ex vivo with non-targeting siRNA (siNT) or pools of two siRNAs targeting either Drd3 or Drd4. After 48 h, cells were treated with TGF $\beta$  and dopamine for an additional 48 h prior to fixation. Representative images are shown (Hoechst in blue,  $\alpha$ SMA-RFP in red). **c** Quantification of the number of  $\alpha$ SMA positive cells (shown as fold over control). **d** COLL-EGFP mice were treated with bleomycin and dopamine by intratracheal injection and lungs were harvested after 14 days. Representative images are shown, with nuclei shown in blue, COLL-EGFP in green, and  $\alpha$ SMA in red. **e** Quantification of collagen and  $\alpha$ SMA intensity of lungs shown in panel (d), normalized over untreated control lungs. **f** Mice were treated with bleomycin, either alone or in combination with dopamine, by intratracheal injection. Lungs were harvested after 30 days and stained with Masson-Trichrome to visualize collagen fibres in blue. **g** Quantification of lung fibrosis using the Ashcroft score. **h** Mice were treated with bleomycin by intratracheal injection. Dopamine was administered at 7, 14, and 21 days after bleomycin injection. Lungs were harvested after 30 days and stained with Masson-Trichrome to visualise collagen fibres in blue. **i** Quantification of lung fibrosis using the Ashcroft score. Data in panels **a**, **e**, **g**, and **i** are shown as mean  $\pm$  S.E.M., while data in panel **c** is shown as  $\pm$  SD. Statistical significance was determined using a one-way ANOVA followed by Dunnett's multiple comparison test, \* $P < 0.05$ , \*\* $P < 0.01$ .

Supported by docking results, we tested the activity of TS1 on human adult lung fibroblasts and confirmed that it significantly inhibits TGF $\beta$ -induced myfibroblast activation in vitro, similar to dopamine (Fig. 4G, H).

#### TS1 reduces myfibroblast activation and fibrosis in vivo

Finally, we tested the capacity of TS1 to effectively suppress the formation of lung fibrosis in vivo. To visualize the effect of TS1 on myfibroblast differentiation, we injected TS1 intra-tracheally two



days after the administration of bleomycin in COLL-EGFP  $\alpha$ SMA-RFP mice. As expected, bleomycin strongly increased the expression of both collagen and  $\alpha$ SMA by lung fibroblasts. Both genes were markedly reduced in mice treated with TS1 (Fig. 5a). While the effect on  $\alpha$ SMA was comparable to that of dopamine,

TS1 was by far more potent in reducing the expression of collagen, as shown by the details in Fig. 5a and the whole organ section in Fig. 5b. Quantification of both collagen and  $\alpha$ SMA protein expression both in the presence and in the absence of TS1 is reported in Fig. 5c. Reduced myofibroblast differentiation by TS1

**Fig. 3 TS1 reduces activation and collagen expression of murine lung fibroblasts.** **A** Flowchart showing the 19 compounds structurally related to dopamine which were predicted by machine learning to reduce  $\alpha$ SMA expression (yellow). CHEMBL145929 was selected as the representative molecule from which six novel derivatives were synthesized, TS1-TS6 (orange). **B** Murine lung fibroblasts from  $\alpha$ SMA-RFP, COLL-EGFP mice were treated in vitro with TGF $\beta$  and lipophilic TS1-6. Representative images are shown (Hoechst in blue, COLL-EGFP in green,  $\alpha$ SMA-RFP in red). **C–E** Quantification of the number of  $\alpha$ SMA positive cells, collagen positive cells, and total cell number (shown as fold over control). **F** Murine lung fibroblasts from COLL-EGFP/ $\alpha$ SMA-RFP mice were treated ex vivo with TGF $\beta$  and the hydrophilic TS1-HCl. Representative images are shown (Hoechst in blue, COLL-EGFP in green,  $\alpha$ SMA-RFP in red). **G–I** Quantification of the number of  $\alpha$ SMA positive cells, collagen positive cells, and total cell number (shown as fold over control). All data are shown as mean  $\pm$  SD. Statistical significance was determined using a one-way ANOVA followed by Dunnett's multiple comparison test, \* $P < 0.05$ , \*\* $P < 0.01$ ,  $n = 3$ .

resulted in a significant decrease in lung fibrosis at 30 days, as quantified in Fig. 5d and shown by whole organ sections in Fig. 5e. Finally, we also assessed the value of TS1 as a therapeutic drug, by administering it in lungs with established fibrosis. For this purpose, we treated mice with bleomycin, and then we administered TS1 two weeks following bleomycin delivery. This treatment significantly reduced the amount of fibrosis at 21 days (Supplementary Fig. 6) and at 30 days (Fig. 5f, g) as quantified using the Ashcroft score. Intriguingly, the levels of bleomycin-induced inflammation and fibrosis were more pronounced at 21 days than at 30 days, allowing TS1 to induce a more significant reduction at this earlier time point. Moreover, as the Ashcroft score reveals, TS1 induces a continued improvement in the reduction of fibrosis between 21 and 30 days.

To further explore the therapeutic potential of TS1 we assessed its capacity to inhibit the activation of myofibroblasts harvested from the lungs of IPF patients. As shown in Supplementary Fig. 7, TS1 reduced the expression of both collagen and  $\alpha$ SMA upon cell stimulation with TGF $\beta$ . More impressive and significant results were observed in the absence of exposure to TGF $\beta$  (Fig. 5h), consistent with chronic activation of myofibroblasts in IPF lungs.

Thus, TS1 reverses myofibroblast differentiation and fibrosis, both in preventive and curative approaches in mice, as well as in human IPF lung fibroblasts.

## DISCUSSION

Our work describes a unique wet-dry-wet approach for drug discovery using myofibroblast differentiation as a phenotypic readout. We started from a "wet" drug screen using a library of 640 compounds, an extensive library that nevertheless fails to represent the full scope of potentially therapeutic compounds. We then applied a "dry" machine learning strategy to extrapolate our wet results onto a database of over 1 million drug-like compounds to eventually design and synthesize a few novel candidates, one of which (TS1) potently inhibited lung fibrosis in a "wet" context. The power and efficiency of our approach is documented by the relatively low number of candidates needed to be produced and validated in vitro prior to moving to in vivo models.

Both academic institutions and pharma industry are investing high resources in the design and execution of virtual compound screens [25–27]. The ability to process large datasets using machine learning can be successfully applied in the prediction of druggable targets and is often combined with computational biology approaches that take the global properties of a compound into consideration [28]. High throughput data were used as a powerful data source for machine learning analysis in multiple contexts; for example, to improve the analysis of protein localization within cells and to search for novel antimicrobial compounds [29–31]. Our work provides further evidence that virtual screening can be a powerful tool to identify therapeutic compounds quickly and effectively.

Our virtual screening pointed to the selection of a number of molecules that were structurally related to dopamine, yet dopamine itself proved to be inefficient at reducing fibrosis

in vivo. This can be explained through the low stability of dopamine in vivo, both due to the instable chemical structure and the presence of degrading enzymes in circulation. TS1 is a much more stable compound, which has demonstrated success in vitro and in vivo in reducing both  $\alpha$ SMA and collagen expression levels, as well as counteracting the development of fibrosis both in a preventative and therapeutic protocol. Our docking and siRNA-mediated knock-down experiments have shown that this derivative acts through the Drd3 receptor, and we have further demonstrated a direct effect on inhibiting the pro-fibrotic canonical TGF $\beta$  pathway.

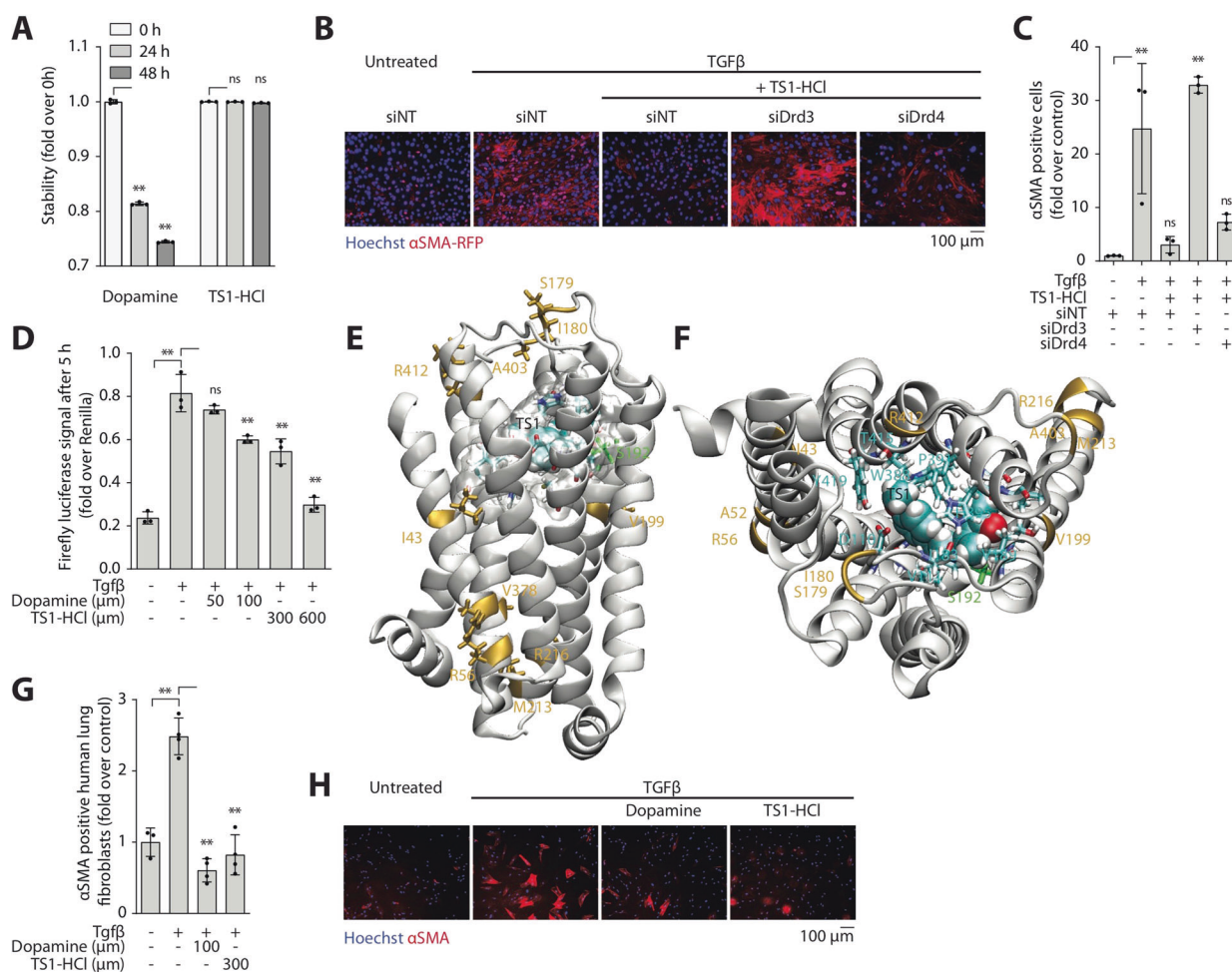
In the original "wet" screen, we identified haloperidol as a potent anti-fibrotic agent [11]. Interestingly, haloperidol is known to be a dopamine receptor antagonist, which could be seen as a contradiction to our findings. However, we have previously demonstrated that haloperidol inhibits myofibroblast differentiation by acting on sigma receptor 1, and moreover this antipsychotic drug has been shown to specifically inhibit Drd2, but not Drd3 or Drd4 [32]. In our work, we have shown Drd2 to be expressed below detectable levels in both lung and heart fibroblasts and therefore believe that TS1 is working through a different pathway than haloperidol.

While there are contradictory opinions on the role of dopamine agonists in fibrosis, there is strong evidence to support an anti-fibrotic effect of Drd3. One publication shows the anti-fibrotic effect of the Drd3 agonist pramipexole in a model of cardiac fibrosis induced by morphine, which activates Smad 2/3 signalling [33]. Additionally, more definitive evidence comes from Drd3 knock-out mice. Removal of this receptor in vivo results in increased cardiac fibrosis, mimicking the phenotype of aged animals [34].

Our results are complemented by a recent publication, showing that dopamine and various dopamine receptor agonists have an anti-fibrotic effect, by acting through the Drd1 receptor to inhibit the pro-fibrotic Yap/Taz pathway [35]. In this work, the main compound which is investigated is dihydroxydine, a moderately selective agonist able to bind both D1- and D2-type dopamine receptors. We have shown barely detectable expression levels of Drd1 in lung fibroblasts. This would possibly suggest a combined mechanism by which both Drd1 and Drd3 could contribute to the anti-fibrotic effect of dopamine receptor agonists, whereby the first is acting via the Yap/Taz pathway while the latter has an effect on canonical TGF $\beta$  signalling. A further interesting observation made by Haak et al. is a dramatic increase in the expression of Drd3 in lung epithelial cells following bleomycin treatment. This could further implicate Drd3 as an important pathway in fibrosis, indicating a compensatory role in mitigating the fibrotic response.

In conclusion, our work demonstrates the power of transdisciplinary approaches in the identification of novel anti-fibrotic drugs. By combining HTS and high content analysis, with machine learning and chemical synthesis, we were able to develop new powerful anti-fibrotic tools. Besides repurposing the known FDA-approved compound dopamine, we also identified and produced a novel derivative, TS1. Our data show that TS1 is highly stable and shows promising results both in vitro and in vivo in fighting





**Fig. 4** TS1 acts on the canonical TGFβ pathway via Drd3, inhibiting differentiation of both murine and human lung fibroblasts. **A** HPLC analysis shows compound stability of solubilised dopamine and TS1-HCl at room temperature over 48 h, data normalized over the initial time point (T0). **B** Murine lung fibroblasts were treated ex vivo with non-targeting siRNA (siNT) or pools of two siRNAs targeting either Drd3 or Drd4. After 48 h, cells were treated with TGFβ and TS1-HCl for an additional 48 h prior to fixation. Representative images are shown (Hoechst in blue, αSMA-RFP in red). **C** Quantification of the number of αSMA positive cells (shown as fold over control). **D** HEK293T cells stably expressing the TGFβ-sensitive (CAGA)<sub>12</sub>-luciferase reporter were treated with TGFβ and Dopamine or TS1-HCl to evaluate the effect of the compounds on the activation of the canonical TGFβ-Smad pathway. Data is the Firefly-luciferase intensity normalized over Renilla-luciferase intensity. Computational model of TS1 docked to murine Drd3: side (**E**) and top (**F**) view. Colour code: missense mutation with respect to the human receptor (yellow), residues interacting with TS1 with hydrogen bonds (green), and other interacting residues: carbons (cyan), nitrogens (blue), oxygens (red), hydrogens (white). TS1 atoms are represented with their van der Waal spheres (space-filling). **G** Human adult primary lung fibroblasts were treated ex vivo with TGFβ and TS1-HCl for 48 h prior to fixation and staining of αSMA. Quantification of the number of αSMA positive cells (shown as fold over control). **H** Representative images are shown (Hoechst in blue, αSMA in red). All data are shown as mean ± SD. Statistical significance was determined using a one-way ANOVA followed by Dunnett's multiple comparison test, \* $P < 0.05$ , \*\* $P < 0.01$ ,  $n = 3$ .

fibrosis through direct modulation of the canonical TGFβ signalling pathway.

## MATERIALS AND METHODS

### Machine learning

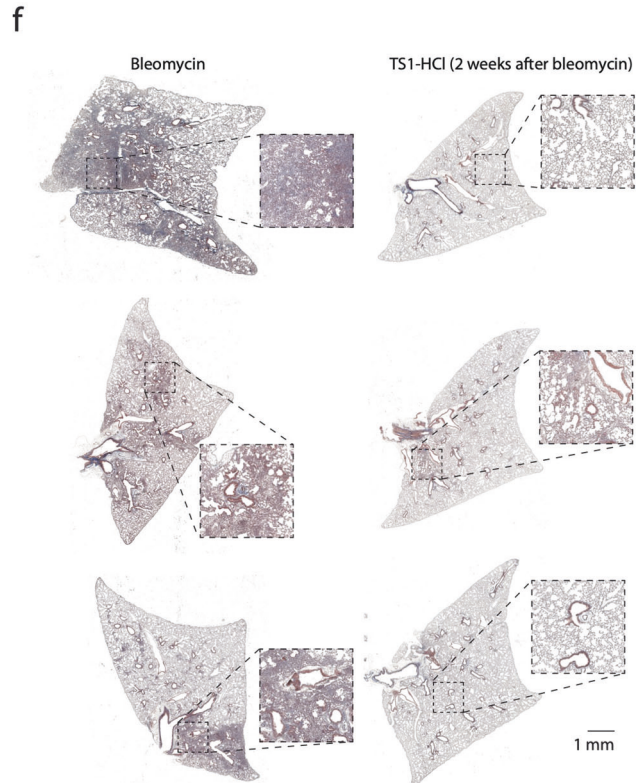
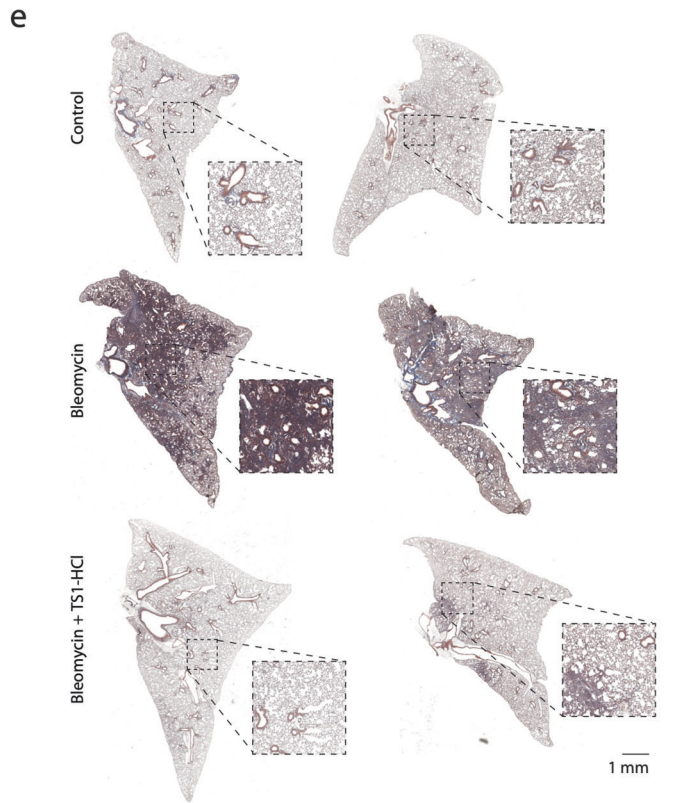
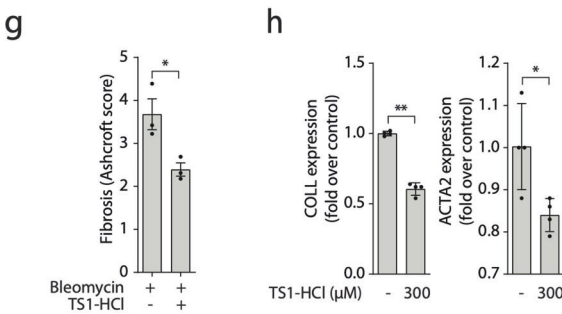
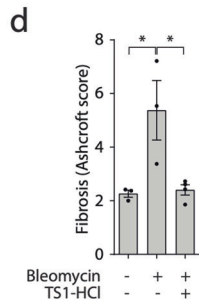
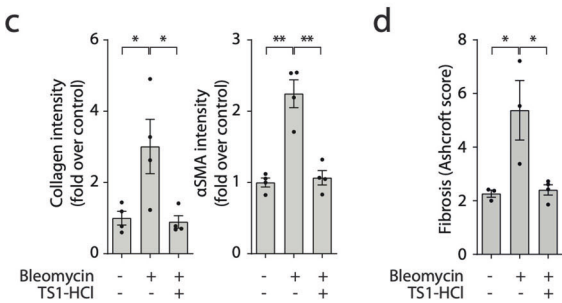
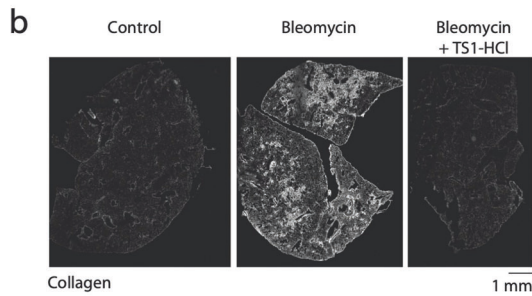
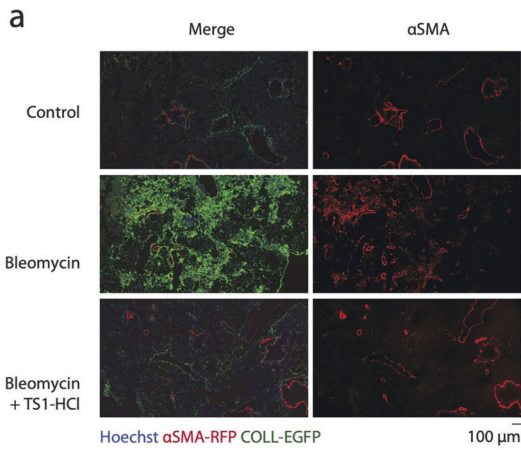
Starting from previously published data [11], we built a predictive model that takes compound descriptions as input and consequently predicts their effect on αSMA and collagen expression. The original screen was used as the training set. It included 640 compounds, identified with CAS numbers. In order to match CAS numbers with other identifiers, the Chemical Translation Service (<http://cts.fiehnlab.ucdavis.edu/>) and some manual checking were used. Each molecule was eventually described using Extended Connectivity Fingerprints (ECFP)—represented by a binary vector indicating whether a given compound contains a set of substructures. Bagging ensembles of predictive clustering trees (PCTs) were used as predictive models, as they yield state-of-the-art predictive

performance and can provide estimates of the reliability of their predictions. We used 10-fold cross-validation to estimate the predictive performance of the models on unseen compounds. Since the goal was to identify interesting candidates for wet-lab experiments, we also wanted to investigate the reliability of individual predictions. For this, we looked at the standard deviation of predictions of individual trees in the ensemble and found a clear correlation between lower standard deviation of the predictions of individual trees and lower prediction error of the ensemble prediction. We identified ChEMBL (version 24) as a suitable database of compounds to use as a testing set and apply the predictive model on, in search of interesting candidates.

### Cell culture and treatments

Primary murine fibroblasts were isolated from lung and heart of αSMA-RFP C57BL/6 or αSMA-RFP/COLL-EGFP C57BL/6 mice by tissue digestion and mechanical shearing. For this, tissue was minced and then digested for 1 h at 37 °C with a magnetic stirrer using collagenase, dipeptidase, and





accutase. Dissociated tissue was passed through a 70 μm cell strainer and centrifuged at 450 × g. The resulting cell pellet was resuspended in DMEM with 10% FBS, and fibroblasts were allowed to attach on plastic tissue culture plates for 2 h. Medium was changed, and cells were left to recover overnight in a humidified incubator at 37 °C before being immediately used for experiments. Human primary lung fibroblasts were purchased

from ATCC, while those from IPF lungs were obtained from Dr. Steven Huang (University of Michigan). Human recombinant TGFβ1 (Peprotech) was used for stimulation in combination with drug treatments, and cells were fixed after 48 h. Zanamivir, topiramate and dopamine were purchased from Sigma and resuspended in water or DMSO, as required. Human cells were stained with αSMA-Cy3 (Sigma-Aldrich, clone 1A4).

**Fig. 5 TS1 inhibits myofibroblast differentiation in vivo and reduces lung fibrosis in bleomycin-treated mice.** **a** COLL-EGFP/ $\alpha$ SMA-RFP mice were treated with bleomycin and two days later with TS1-HCl by intratracheal injection; lungs were harvested after 14 days. Representative images are shown, with nuclei shown in blue, COLL-EGFP in green, and  $\alpha$ SMA-RFP in red. **b** Representative whole organ images of treated mice, in which collagen expression is shown in white. **c** Quantification of collagen and  $\alpha$ SMA intensity of lungs shown in panels (a, b), normalized over untreated, control lungs. **d, e** Mice were treated with bleomycin, either alone or in combination with TS1-HCl by intratracheal injection. Lungs were harvested after 30 days and stained with Masson-Trichrome to visualise collagen fibres in blue. Lung fibrosis was quantified using the Ashcroft score. **f** Mice were treated with bleomycin, and after 2 weeks mice were treated with TS1-HCl by intratracheal injection. Lungs were harvested after a total of 30 days and stained with Masson-Trichrome to visualise collagen fibres in blue. **g** Quantification of lung fibrosis was performed using the Ashcroft score to quantify collagen fibres stained in blue. **h** Levels of expression of collagen (*col1a1*) and  $\alpha$ SMA (*acta2*) in human fibroblasts derived from IPF lungs, upon treatment with TS1-HCl at the indicated dose. Results are normalized to *Gapdh* expression. All data are shown as mean  $\pm$  S.E.M. Statistical significance was determined using a one-way ANOVA followed by Dunnett's multiple comparison test, \* $P < 0.05$ , \*\* $P < 0.01$ .

### Transfections

Silencing of *Drd3* and *Drd4* was performed in primary murine fibroblasts using pools of two siRNAs purchased from Sigma (siDrd3: SASI\_Mm01\_00104662, SASI\_Mm01\_00104663; siDrd4: SASI\_Mm01\_00044753, SASI\_Mm01\_00044754). Control cells were transfected with ON-TARGETplus Non-targeting Control siRNA #4 (Dharmacon). Cells were transfected using 50 nM of siRNA complexed with lipofectamine RNAiMAX (ThermoFisher).

### Immunohistochemistry

Cells were fixed with 4% PFA and nuclei were stained using Hoechst 3334. Images were acquired using an Operetta high content screening microscope (Perkin Elmer) with a 20 $\times$  objective. A minimum of 9 fields were acquired per well, and the number of  $\alpha$ SMA positive cells was quantified using the Columbus software (threshold set for TRITC channel cellular mean fluorescence intensity). For imaging of tissue sections from  $\alpha$ SMA-RFP/COLL-EGFP C57BL/6 mice, immunohistochemistry was performed on cryosections fixed with 4% PFA or paraffin embedded sections fixed with formalin, which were deparaffinised before processing. Nuclei were counterstained with Hoechst and tissue was imaged on a Nikon Eclipse Ti-E inverted fluorescence microscope. For visualization of fibrosis, tissue was stained using Masson-Trichrome (Bio Optica).

### qRT-PCR

RNA was prepared using Trizol (Invitrogen) and phenol chloroform extraction, and cDNA was prepared using the First strand cDNA synthesis kit (Thermo Scientific) following the manufacturer's instructions. Samples were run on a C1000 Thermal Cycler (Biorad) and analyzed using the proprietary software and the 2 $^{-\Delta\Delta C_T}$  and 2 $^{-\Delta\Delta C_T}$  methods.

Primers for dopamine receptors were taken from literature [36]: *mDrd1* fw CACGGATCCATCCTTAACCT, rev TGCCTCGGAGTCATCTTCCT; *mDrd2* fw ACCTGTCTCGGTACGATGATG, rev GCATGGCATAGTAGTTGTAGTGG; *mDrd3* fw CCAGTTCACATACGATGGC, rev CCCTGTGTGTGAAACCAA; *mDrd4* fw GGTGTGCGGACCCTACTCAG, rev GGCAGGACTCTCATTCCTT; *mDrd5* fw TGCTGTCCAATGAGACACCC, rev GATGGCGTAGTTCGGTTTACG. Human primers were the following: *col1a1* fw AAGGGACACAGAGTTTCAGTGG, rev CAGCACCAGTAGCACCATCATTTTC; *acta2* fw GTGAAGAAGAGGACAGCACTG, rev CCCATTCACCACATCACC. *Gapdh* was used as a house-keeping gene for both human and mouse experiments.

### Luciferase reporter assay

HEK293T cells stably expressing the TGF $\beta$ -sensitive (CAGA)12-luciferase reporter (pGL3 CAGA12 LC + PuroR-P2A-Renilla-T2A-EGFP) [37] were starved overnight in 0% FBS in low glucose DMEM, and then treated for 6 h with human recombinant TGF $\beta$ 1 (Peprotech), in combination with various concentrations of dopamine or TS1-HCl. Cells were lysed with Glo-lysis buffer (Promega) and assayed using the Dual-Glo luciferase kit (Promega) on the Wallac Envision plate reader (Perkin Elmer).

### Animal experiments

Ethical and experimental procedures were reviewed and approved by the ICGEB Animal Welfare board, meeting the requirements of the EU Directive 2010/63/EU, and by the Italian Ministry of Health (approval number 806/2018-PR). Bleomycin was injected intratracheally at 1 U/kg, dopamine was administered at 0.76  $\mu$ g/kg, TS1-HCl at 2 mg/kg. Mice were sacrificed at either 14 or 30 days post-treatment. The severity of fibrosis was determined on Masson-Trichrome stained lung sections using the Ashcroft score [38].

### Compound synthesis

Commercially available chemicals were of reagent grade and used as received. Reaction courses and product mixtures were routinely monitored by thin-layer chromatography (TLC) on silica gel pre-coated F<sub>254</sub> Merck plates. Melting points were determined using a Stuart SMP300 apparatus and uncorrected. Infrared spectra were recorded on a Jasco 4700 spectrophotometer in nujol mulls. Nuclear magnetic resonance spectra were registered on a Varian 400 MHz (400 for <sup>1</sup>H-NMR and 101 MHz for <sup>13</sup>C-NMR), shown in Supplementary Figure 8. Chemical shifts are reported as  $\delta$  (ppm) in CDCl<sub>3</sub> solution ( $\delta = 7.26$  ppm for <sup>1</sup>H-NMR and  $\delta = 77.2$  for <sup>13</sup>C-NMR) or DMSO-d<sub>6</sub> ( $\delta = 2.49$  ppm for <sup>1</sup>H-NMR and  $\delta = 39.52$  for <sup>13</sup>C-NMR); 20  $\mu$ L of D<sub>2</sub>O was added to assign NH and OH protons. Microanalyses (C, H, N) were carried out using an Elementar Vario ELIII apparatus and were in agreement with theoretical values  $\pm$  0.4%. ESI-MS spectra (LRMS) were recorded on a Bruker Daltonics Esquire 4000 spectrometer by infusion of a solution of the sample in MeOH (HPLC grade). Microwave oven synthesizer Anton Paar monowave 400.

### Compound stability analysis

HPLC analysis was performed using methanol and ammonium acetate (Sigma Aldrich). Analyses were performed on an HPLC Agilent 1260 Infinity II with a Supelco C18 Discovery 250  $\times$  4.6 mm, 5 mm particle size (Supelco) column with a column guard and an autosampler using a DAD detector at 210 nm. The flow was set to 1.00 mL/min with an injection volume of 20  $\mu$ L. The mobile phase was ammonium acetate 0.01 M and methanol (60:40) using isocratic condition at 25  $^{\circ}$ C.

### Homology modelling and docking

*Drd3* murine (UniProt P30728) and human (UniProt P35462), were aligned with BLAST [39]. Three-dimensional receptor models were built by homology model with SWISS-MODEL [40] via the ExPASy web server by employing PDB ID 3PBL [24] as a template (human dopamine D3 receptor, Lysozyme chimera, in complex with eticlopride). TS1 SMILES was converted into a protonated three-dimensional model with Obabel [41], and minimized with AM1 as implemented in MOPAC [42]. The Lamarckian Genetic Algorithm (LGA) was used for docking TS1 to both human and murine *Drd3* models. The docking box (50  $\times$  50  $\times$  50), spacing 0.375  $\text{Å}$ , was centred at (2.692, -16.966, 0.876) in the frame of the original template (PDB ID 3PBL) coordinates [24]. 2000 LGA runs with initial population size of 500, 10,000,000 maximum numbers of evaluations, and standard parameters were run as implemented in AutoDock [43]. Docked conformations were clustered with 2.0  $\text{Å}$  tolerance. For both murine and human models, the representative conformations of the larger cluster (comprising  $\approx$ 45% of the generated conformations) were selected as representative conformation for subsequent analysis. Specific ligand-protein interactions were derived through LigPlot+ [44].

### DATA AVAILABILITY

The data supporting the findings from this study are available within the manuscript and its supplementary information. Source data are provided with this paper.

### REFERENCES

1. Marshall DC, Saliccioli JD, Shea BS, Akuthota P. Trends in mortality from idiopathic pulmonary fibrosis in the European Union: an observational study of the WHO mortality database from 2001-2013. *Eur Respir J.* 2018;51:1701603

2. Rosenbloom J, Macarak E, Piera-Velazquez S, Jimenez SA. Human fibrotic diseases: current challenges in fibrosis research. *Methods Mol Biol.* 2017;1627:1–23.
3. Murtha LA, Schuliga MJ, Mabotwana NS, Hardy SA, Waters DW, Burgess JK, et al. The processes and mechanisms of cardiac and pulmonary fibrosis. *Front Physiol.* 2017;8:777.
4. Khalil H, Kanisicak O, Prasad V, Correll RN, Fu X, Schips T, et al. Fibroblast-specific TGF- $\beta$ -Smad2/3 signaling underlies cardiac fibrosis. *J Clin Invest.* 2017;127:3770–83.
5. Bochaton-Piallat ML, Gabbiani G, Hinz B. The myofibroblast in wound healing and fibrosis: answered and unanswered questions. *F1000Res.* 2016;5:F1000
6. de Salvi Guimaraes F, de Moraes WM, Bozi LH, Souza PR, Antonio EL, Bocalini DS, et al. Dexamethasone-induced cardiac deterioration is associated with both calcium handling abnormalities and calcineurin signaling pathway activation. *Mol Cell Biochem.* 2017;424:87–98.
7. Huang S, Frangogiannis NG. Anti-inflammatory therapies in myocardial infarction: failures, hopes, and challenges. *Br J Pharmacol.* 2018;175:1377–400.
8. Myllarniemi M, Kaarteenaho R. Pharmacological treatment of idiopathic pulmonary fibrosis—preclinical and clinical studies of pirfenidone, nintedanib, and N-acetylcysteine. *Eur Clin Respir J.* 2015;2:26385.
9. Lagares D, Santos A, Grasberger PE, Liu F, Probst CK, Rahimi RA, et al. Targeted apoptosis of myofibroblasts with the BH3 mimetic ABT-263 reverses established fibrosis. *Sci Transl Med.* 2017;9:eal3765.
10. Wei P, Xie Y, Abel PW, Huang Y, Ma Q, Li L, et al. Transforming growth factor (TGF)- $\beta$ 1-induced miR-133a inhibits myofibroblast differentiation and pulmonary fibrosis. *Cell Death Dis.* 2019;10:670.
11. Rehman M, Vodret S, Braga L, Guarnaccia C, Celsi F, Rossetti G, et al. High-throughput screening discovers antifibrotic properties of haloperidol by hindering myofibroblast activation. *JCI Insight.* 2019;4:e123987
12. Davies M, Nowotka M, Papadatos G, Dedman N, Gaulton A, Atkinson F, et al. ChEMBL web services: streamlining access to drug discovery data and utilities. *Nucleic Acids Res.* 2015;43:W612–20.
13. Gaulton A, Hersey A, Nowotka M, Bento AP, Chambers J, Mendez D, et al. The ChEMBL database in 2017. *Nucleic Acids Res.* 2017;45:D945–D54.
14. Magness ST, Bataller R, Yang L, Brenner DA. A dual reporter gene transgenic mouse demonstrates heterogeneity in hepatic fibrogenic cell populations. *Hepatology* 2004;40:1151–9.
15. Selman M. From anti-inflammatory drugs through antifibrotic agents to lung transplantation: a long road of research, clinical attempts, and failures in the treatment of idiopathic pulmonary fibrosis. *Chest* 2002;122:759–61.
16. Li X, Zhu L, Wang B, Yuan M, Zhu R. Drugs and targets in fibrosis. *Front Pharmacol.* 2017;8:855.
17. Raghu G, Rochweg B, Zhang Y, Garcia CA, Azuma A, Behr J, et al. An official ATS/ERS/JRS/ALAT clinical practice guideline: treatment of idiopathic pulmonary fibrosis. An update of the 2011 Clinical Practice Guideline. *Am J Respir Crit Care Med.* 2015;192:e3–19.
18. Swartz RD. Idiopathic retroperitoneal fibrosis: a review of the pathogenesis and approaches to treatment. *Am J Kidney Dis.* 2009;54:546–53.
19. Rafii R, Juarez MM, Albertson TE, Chan AL. A review of current and novel therapies for idiopathic pulmonary fibrosis. *J Thorac Dis.* 2013;5:48–73.
20. Backman TW, Cao Y, Girke T. ChemMine tools: an online service for analyzing and clustering small molecules. *Nucleic Acids Res.* 2011;39:W486–91.
21. Yamanaka T, Tsujimura K, Kondo T, Matsumura T. In vitro efficacies of oseltamivir carboxylate and zanamivir against equine influenza A viruses. *J Vet Med Sci.* 2006;68:405–8.
22. White HS, Brown SD, Woodhead JH, Skeen GA, Wolf HH. Topiramate enhances GABA-mediated chloride flux and GABA-evoked chloride currents in murine brain neurons and increases seizure threshold. *Epilepsy Res.* 1997;28:167–79.
23. Sanchez-Rivera AE, Corona-Avendano S, Alarcon-Angeles G, Rojas-Hernandez A, Ramirez-Silva MT, Romero-Romo MA. Spectrophotometric study on the stability of dopamine and the determination of its acidity constants. *Spectrochim Acta A Mol Biomol Spectrosc.* 2003;59:3193–203.
24. Chien EY, Liu W, Zhao Q, Katritch V, Han GW, Hanson MA, et al. Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science* 2010;330:1091–5.
25. Lyu J, Wang S, Balius TE, Singh I, Levit A, Moroz YS, et al. Ultra-large library docking for discovering new chemotypes. *Nature* 2019;566:224–9.
26. Walters WP. Virtual chemical libraries. *J Med Chem.* 2019;62:1116–24.
27. Zoffmann S, Verccruyse M, Benmansour F, Maunz A, Wolf L, Blum Marti R, et al. Machine learning-powered antibiotics phenotypic drug discovery. *Sci Rep.* 2019;9:5013.
28. Kandoi G, Acencio ML, Lemke N. Prediction of druggable proteins using machine learning and systems biology: A mini-review. *Front Physiol.* 2015;6:366.
29. Riniker S, Wang Y, Jenkins JL, Landrum GA. Using information from historical high-throughput screens to predict active compounds. *J Chem Inf Model.* 2014;54:1880–91.
30. Kraus OZ, Grys BT, Ba J, Chong Y, Frey BJ, Boone C, et al. Automated analysis of high-content microscopy data with deep learning. *Mol Syst Biol.* 2017;13:924.
31. Korbee CJ, Heemskerck MT, Kocev D, van Strijen E, Rabiee O, Franken K, et al. Combined chemical genetics and data-driven bioinformatics approach identifies receptor tyrosine kinase inhibitors as host-directed antimicrobials. *Nat Commun.* 2018;9:358.
32. Fan L, Tan L, Chen Z, Qi J, Nie F, Luo Z, et al. Haloperidol bound D2 dopamine receptor structure inspired the discovery of subtype selective ligands. *Nat Commun.* 2020;11:1074.
33. Gaweda G, Iyer RP, Shaver PR, Grilo GA, Dinkins ML, Stoffel HJ, et al. Dopamine receptor D3 agonist (Pramipexole) reduces morphine-induced cardiac fibrosis. *Biochem Biophys Res Commun.* 2020;529:1080–5.
34. Johnson TL, Tulis DA, Keeler BE, Virag JA, Lust RM, Clemens S. The dopamine D3 receptor knockout mouse mimics aging-related changes in autonomic function and cardiac fibrosis. *PLoS One.* 2013;8:e74116.
35. Haak AJ, Kostallari E, Sicard D, Ligresti G, Choi KM, Caporarello N, et al. Selective YAP/TAZ inhibition in fibroblasts via dopamine receptor D1 agonism reverses fibrosis. *Sci Transl Med.* 2019;11:eaa6296
36. Lu J, Li X, Wang Q, Pei G. Dopamine D2 receptor and beta-arrestin 2 mediate Amyloid-beta elevation induced by anti-Parkinson's disease drugs, levodopa, and piribedil, in neuronal cells. *PLoS One.* 2017;12:e0173240.
37. Drenth S, Itoh S, Vivien D, ten Dijke P, Huet S, Gauthier JM. Direct binding of Smad3 and Smad4 to critical TGF  $\beta$ -inducible elements in the promoter of human plasminogen activator inhibitor-type 1 gene. *EMBO J.* 1998;17:3091–100.
38. Ashcroft T, Simpson JM, Timbrell V. Simple method of estimating severity of pulmonary fibrosis on a numerical scale. *J Clin Pathol.* 1988;41:467–70.
39. Johnson M, Zaretskaya I, Raytselis Y, Merezuk Y, McGinnis S, Madden TL. NCBIBLAST: a better web interface. *Nucleic Acids Res.* 2008;36:W5–W9.
40. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 2018;46:W296–W303.
41. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: An open chemical toolbox. *J Cheminformatics.* 2011;3:33.
42. Stewart JJP. Special issue—MOPAC—A semiempirical molecular-orbital program. *J Comput Aided Mol Des.* 1990;4:1–45.
43. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem.* 2010;31:455–61.
44. Laskowski RA, Swindells MB. LigPlot+: multiple ligand–protein interaction diagrams for drug discovery. *J Chem Inf Model.* 2011;51:2778–86.

## ACKNOWLEDGEMENTS

This work was supported by project TRAIN (high Throughput screening and big data Analysis for INnovation) funded by the European Regional Development Fund and Interreg ITA-SLO 2014-2020, Fondazione CRTrieste, Italy, and Fondazione Roche, Italy. We are grateful to Barbara Bozigrav and Maristella Coglievina for excellent technical assistance.

## AUTHOR CONTRIBUTIONS

NARR and MCV performed most of the cell culture and in vivo experiments; TS, PP, DK, and SD designed machine learning algorithm and performed virtual screening; MGM, SS, and DZ synthesized new chemical compounds including T51; SV contributed to animal experiments; SF and AC performed docking analysis; MR, AC, and PM contributed to microscopy and data analysis; LC and FSL provided cells for TGF $\beta$  luciferase reporter assay; TM and RB contributed to whole lung section analysis; MC, TXP, GL, and NC contributed with human cells and critical discussion; SZ conceived the study, supervised activities and wrote the manuscript.

## COMPETING INTERESTS

The authors declare no conflict of interest.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41419-021-04439-4>.

**Correspondence** and requests for materials should be addressed to Serena Zacchigna.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.





**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021