Abstract:
Cathepsins, key members of the papain-like family of cysteine proteases, are crucial for proteolysis processes
within human cells, including osteolysis, immunomodulation and apoptosis. Recent research has highlighted the
significant role of cathepsins, particularly the L, S, K, and B subtypes, in pancreatic cancer. This has driven the

development of novel cathepsin inhibitors as potential treatments to inhibit tumor progression, migration and
invasion. Targeting cathepsin S (CatS) has shown promise in reducing tumor progression and enhancing the
efficacy of chemotherapeutic agents in preclinical models. Building on our previous work where we employed
ethyl p-aminocinnamate ester derivatives for covalent inhibition of cysteine proteases, herein we have designed
and synthesized three new derivatives basing on an isosteric replacement (H–F) at the level of cinnamate moiety.
These derivatives emerged as potent covalent inhibitors of CatS (1.8–2.6 μM) with 2F showing also weak inhibition activity against CatL (20 %) and CatB (29 %). In vitro assays of 2F against pancreatic cancer cell lines BXPC3 and CAPAN1 revealed significant antiproliferative activity, with IC50 = 5.79 μM and 20.75 μM, respectively. These findings underscore the potential of α-fluorocinnamate-based cysteine protease inhibitors as promising candidates for further development in targeting CatS and CatL with the aim to reduce pancreatic
cancer cell proliferation.

Introduction:

Cathepsins are primary members of the papain-like family of cysteine proteases and play a crucial role in proteolysis processes within human cells.1 Under normal physiological conditions, these proteases are responsible for, among other regulatory functions, osteolysis, immunomodulation and apoptosis.2 In recent years, a growing body of research has revealed the close connection between metastatic cancers and cathepsins, notably the L, S, K and B subtypes. Consequently, the development of novel cathepsin inhibitors has emerged as a viable approach for inhibiting tumor progression, migration and invasion.3 Specifically, cathepsin S (CatS) is a peculiar cysteine protease with limited tissue distribution (mainly in antigen-presenting cells, e.g. pancreatic β-cells expressing MHC-II) and that functions at neutral pH.4 In addition, recent studies have shown that CatS plays a significant role in pancreatic cancer by promoting tumor growth and metastasis through degradation of extracellular matrix components.5 Pancreatic cancer is a highly

aggressive malignancy characterized by a ferocious nature, rapid progression and often late diagnosis, making it one of the deadliest forms of cancer.6 By the time it is diagnosed, pancreatic cancer has usually advanced to a later stage, making it refractory to treatment.7 High levels of extracellular CatS have also been associated with increased angiogenesis and immune evasion in pancreatic tumors.8 Nonetheless, targeting CatS has been shown to reduce tumor progression and enhance the efficacy of chemotherapeutic agents in preclinical models.9,10 In one of our previous studies, we proposed ethyl p- aminocinnamate ester as an electrophilic warhead to trap the catalytic thiol of cysteine, to be implemented in the development of new cysteine protease inhibitors, specifically for the covalent inhibition of SARS-CoV2 Mpro. 11 In the current work, a simple isosteric replacement H–F to the cinnamate moiety has led to the synthesis of three new derivatives (namely 1F, 2F and 3F) that proved to be potent covalent inhibitors of CatS (1.8–2.6 μM) with 2F showing also weak inhibition activity against CatL and CatB. The plausible interaction of these molecules within the active sites of CatS was investigated using a covalent docking approach, which demonstrated their strong potential to favorably interact within the binding pocket. The target compounds were tested in vitro against two pancreatic cancer cell lines, BXPC3 and CAPAN1, showing antiproliferative activity with 2F being the best candidate for further development in the future (IC50 = 5.94 μM). These findings highlight how α-fluorocinnamate-based cysteine protease inhibitors can serve as excellent starting points for the development of potent CatS inhibitors with a demonstrated correlation in their ability to reduce pancreatic cancer cell proliferation.

Results and discussion

Design:

In our precedent work, we identified novel pseudopeptide SARSCoV-2 Mpro inhibitors based on the ethyl cinnamate fragment, introduced as a C-terminal electrophilic warhead (Fig. 1).11 The mechanism of action of the cinnamic ester inhibitors relies on the covalent inhibition of Mpro by trapping the thiol group of Cys145 in a Michael-type reactivity. Consequently, our attention moved to the modification of the key ethyl cinnamate moiety, with the aim to introduce some functionalization able to enhance the reactivity toward thiol nucleophiles and therefore

inhibition activity against cysteine proteases. For this purpose, we decided to introduce an α-fluoro-substitution onto the cinnamic scaffold (Fig. 1). From a chemical perspective, it is well known that α,β-unsaturated carbonyl compounds are susceptible to nucleophilic attack at the β-carbon, due to the presence of the carbonyl electron withdrawing group (EWG) that acts as an anion-stabilizing and alkeneactivating group. Considering the reactivity of vinylogous carbonyl species, we supposed that the introduction of an additional EWG in α-position could enhance the electrophilicity of the β-position, in order to be more reactive against cysteine. Furthermore, the introduction of a fluorine atom could potentially improve the pharmacokinetic profile of the target compounds.12 For all these reasons, we decided to explore novel α-fluoro-substituted cinnamate esters as potential SARS-CoV-2 Mpro inhibitors. We decided to synthesize three fluorinated analogs of our reported Mpro inhibitors starting from 1 and 2 as reference compounds (Fig. 1), chosen from our laboratory's library. 1, which has been shown to block 40 % of cysteine protease enzymatic activity at 20 μM, bears a N-carbobenzyloxy-phenylalanine fragment connected to the ethyl p-aminocinnamate electrophile warhead and was originally the prototype of our previous work. 2, bearing a N-terminal 1H-indol-3- acetamide functional group, proved to be the most promising Mpro inhibitor of our library (~60 % inhibition of Mpro at 20 μM; IC50 = 12.4 ± 0.7 μM) and showed also potent antiviral activity against hCoV-OC43 (EC50 = 9.14 ± 0.70 μM). Following the H–F isosteric replacement principle, H in α position of the cinnamic moiety was replaced with F in the attempt to enhance the electrophilic behavior of the β position against the thiol group of Cys145 (Fig. 1). 1F and 2F are fluorinated analogues of 1 and 2, respectively, while 3F was designed directly attaching the 1H-indole-2-carboxylic acid to the central phenylalanine residue, which is common to all structures (Fig. 1). 2.2. Synthesis 1F, 2F, and 3F were synthesized using a convergent synthetic approach. Since the key intermediate ethyl p-amino α-fluorocinnamate (5) was not commercially available, the main efforts were focused on obtaining such fragment (Scheme 1). A selective electrophilic monofluorination was carried out on the α-position of the precursor ethyl benzoylacetate using Selectfluor™ (1-chloromethyl-4-fluoro-1,4-diazoniabicyclo[2.2.2]octane bis(tetrafluoroborate) or F-TEDA)

to obtain the monofluorinated intermediate 3 in 72 % yield. Special attention was needed for the reaction on this 1,3-dicarbonyl scaffold, as Selectfluor™ can perform electrophilic fluorination at its acidic α position, potentially yielding both mono- and di-fluorinated derivatives. Among different attempts, the best result in terms of yield was obtained following the reaction conditions proposed by Tang et al., when the reaction is carried out by using 1.3 equiv of Selectfluor™ in a water:ACN 1:1 medium at room temperature overnight.13 3 was then reacted with p-nitrobenzaldehyde in the presence of Cs2CO3 to provide 4 in 62 % yield. This olefination-type reaction involved the stereoselective synthesis of α-fluoro-α,β-unsaturated esters via deacylation of α-fluoro-β-keto esters, following a route proposed by Qian et al. 14 The process was carried out in ACN, at 40 ◦C overnight, and furnished only the Z stereoisomer, in agreement with reported 1 H NMR data. The final step entailed the reduction of the p-nitrophenyl group using Fe powder in acidic medium in EtOH:water 1:1 mixture at 85 ◦C for 3 h, resulting in the key intermediate 5 in 78 % yield. 1F was obtained in moderate yield by directly coupling 5 with Ncarbobenzyloxy-L-phenylalanine through a conventional amide formation reaction using EDCI/DMAP as coupling reagents (Scheme 2). The synthesis of the indole-bearing compounds (i.e. 2F and 3F) started with an amide coupling reaction between the suitable indole-bearing carboxylic acids and L-phenylalanine methyl ester. As per 1F, the coupling was accomplished by EDCI/DMAP in DCM, providing intermediates 6a and 6b in good yields. Next, the latter were hydrolyzed using LiOH in a mixture 1:1 of 1,4-dioxane and water to yield the carboxylic acids 7a and 7b. Finally, another EDCI/DMAP-mediated coupling reaction between 5 and 7a or 7b was performed to produce the final products, 2F and 3F, in moderate yields. The structural identity and the purity of all the synthesized compounds was assessed by 1 H NMR, 13C NMR and HRMS. The characteristic signal indicating the presence of the C–F bond is represented by a doublet at approximately 6.85 ppm in the 1 H NMR spectra, with a peculiar JH–F = 35 Hz, confirming the Z geometry of the double bond. All compounds were >95 % pure by HPLC analysis (representative HPLC traces are included in the Supporting Information). 2.3. Biomimetic experiments A biomimetic experiment was performed to investigate the interaction of the final compounds with the

target cysteine protease.15 As an example, 2F was incubated with N-acetyl cysteine in phosphate buffer (pH = 7.4), containing 10 % ACN at 37 ◦C and the outcome was analyzed via HPLC-MS to measure the consumption of the starting material (Fig. 2). After just 1 h, the analysis revealed the appearance of a new peak (m/z = 677.4) corresponding to the Michael adduct formed between N-acetyl cysteine and 2F with a conversion rate of ~80 %. This result confirmed the reactivity of α-fluorocinnamate against thiol and highlights the increased reaction rate of the compound with N-acetyl cysteine compared to the derivatives missing F (1), which achieved a 70 % conversion only after 24 h of reaction. The adduct formation was also confirmed using NMR spectroscopy, following a literature procedure where α-fluorocinnamate was reacted with thiophenol. The resulting NMR signal patterns were then compared to verify the formation of the adduct. Key intermediate 5 was used to test the reactivity of α-fluorocinnamate warhead against thiol group of thiophenol. Monitoring the reaction by 19F NMR spectroscopy after 24 h, the disappearance of the α-fluorocinnamate signal (δ: − 130.5 ppm) was observed while a new peak referred to the Michael adduct was shown at − 191.7 ppm (Fig. S1 – panel A), values in agreement with the literature.16 1 H NMR spectra of the reaction mixture showed two new signals at 5.04 ppm (2 JH,F = 48.4 Hz, 3 JH,H = 4.3 Hz) and 4.49 ppm (3 JH,F = 26.5, 3 JH,H = 4.3 Hz), assigned to the α- and β-CH, respectively (Fig. S1 – panel B). 2.4. Enzyme inhibition assays Compounds 1F, 2F and 3F underwent a screening test for their inhibitory activity against the recombinant SARS-CoV-2 Mpro at 20 μM, and the results are shown in Table 1. As can be seen, all three compounds exhibited poor inhibitory activity against Mpro (19 %, 12 %, and 14 % for 1F, 2F, and 3F, respectively) in comparison to their non-fluorinated analogs.11 Therefore, H – F isosteric replacement in α-position of the electrophilic warhead in this case proved to be an unsuccessful strategy. All compounds were also evaluated against SARS-CoV-2 PLpro and showed no detectable inhibition at 20 μM, a similar outcome to that observed for parental derivatives.11 In light of these poor inhibitory results on viral proteases, we determined to screen these three compounds against other proteases belonging to the papain-like family, also based on the availability of these enzymes in our laboratories. Then,

they were tested at 20 μM against CatL, CatB and CatS yielding interesting results. 1F exhibited 90 % inhibitory activity against CatS, while at the screening concentration 2F and 3F completely inhibited CatS enzyme activity. 2F also showed a weak inhibitory activity against CatL (29 % inhibition) and CatB (20 % inhibition), an outcome that might be exploited for the development of multi-target inhibitors since both proteases have been widely correlated with metastatic aggressiveness, such as in pancreatic cancer. These preliminary results clearly indicate that our novel α-fluorinated cinnamic electrophilic warhead was able to switch the inhibitory activity from SARS-CoV-2 Mpro to CatS, highlighting the potential for this moiety in the panorama of papain-like cysteine protease inhibitors in terms of selectivity. Continuous assays (30–0.5 μM) of 1F, 2F and 3F provided noteworthy IC50 values in the low-micromolar range (1F = 2.6 ± 0.5 μM; 2F = 2.0 ± 0.6 μM; 3F = 1.8 ± 0.2 μM). 2.5. Docking 1F, 2F and 3F underwent computational evaluation using a covalent docking technique to understand the preferred conformation and orientation if a covalent bond were formed between the β-carbon of the cinnamic ester and the thiol group of the catalytic cysteine (Fig. 3). First, we attempted to explain the low reactivity of our compounds towards Mpro, performing non-covalent docking of 1F, 2F and 3F. The binding of a covalent binder needs two steps: an initial non-covalent binding followed by a chemical reaction (in our case a Michael addition) leading to the final covalent bonding. Performing non-covalent docking of 1F, 2F and 3F on Mpro, we observed that in the non-covalent poses, the fluorocinnamate group was about 10 Å away from Cys145, a geometry that did not allow for a reaction (Fig. S2). Then, we performed non-covalent docking of 1F, 2F and 3F on the target cathepsins to test if the docked structures were compatible with the subsequent formation of a covalent bond. In this case, we observed (Fig. S3) that in the non-covalent poses the fluorocinnamate group was around 5 Å far from Cys147, a geometry compatible with the addition reaction. This evaluation was conducted specifically for CatS, while 2F was also docked in the active site of CatL to identify interactions that might favor the inhibition of this enzyme as well. This is particularly important because dual inhibitors CatS and CatL have recently been developed for the treatment of pancreatic cancer.17,18 In the case of CatS, the covalent docking revealed two main interactions

with Tyr193 and Tyr334 within the binding site. A T-shaped and a π-π stack interaction are established between the phenyl group of both 1F and 3F and the aromatic moiety of Tyr193 and Tyr334, respectively (Fig. 3 – panels A and B). 2F also shows similar interactions but in this case the interaction is maintained by the indole core (Fig. 3 – panel C). This difference in orientation can be attributed to the slightly different chemical structure of 2F, which includes an extra carbon between the indole moiety and the carbonyl group. Additionally, Fig. 3 – panel D shows the predicted binding mode of 2F to CatL, where the anchor point is the thiol group of Cys25, and the main interaction is a hydrogen bond with Met161. 2.6. Cytotoxicity assays 1F, 2F and 3F were evaluated for their cytotoxic activity against two pancreatic cancer cell lines, BXPC3 and CAPAN1. Their dose–response curves are included in Fig. 4. All the compounds were able to interfere with cell growth at μM concentrations. Among these, 2F demonstrated the highest antiproliferative activity in both cell lines. In BXPC3 cells, 2F exhibited an IC50 of 5.79 μM, while in CAPAN1 cells, the IC50 was slightly higher at 20.75 μM. On the other hand, 1F was the least active compound, displaying the highest IC50 values. Specifically, 1F showed an IC50 of 25.94 μM in BXPC3 cells and an even higher IC50 of 40.58 μM in CAPAN1 cells, suggesting lower activity against this cell line. 3F exhibited intermediate cytotoxicity compared to 2F and 1F across both cell lines. In BXPC3 cells, 3F had an IC50 of 14.05 μM, while in CAPAN1 cells, 3F showed an IC50 of 28.34 μM, again demonstrating moderate activity relative to the other two compounds. Rationalizing the data, the higher activity of 2F against both pancreatic cancer cell lines might be linked to its ability in vitro to inhibit both CatS and CatL. 2.7. ADME and drug-likeness properties The predicted drug-like properties of compounds 1F, 2F and 3F were evaluated using SwissADME (Absorption, Distribution, Metabolism, Excretion) online tool and they displayed favorable pharmacokinetic properties as shown in Table 2. Regarding physicochemical properties, all synthesized compounds showed moderate solubility in water except for 3F, according to the ESOL solubility. The lipophilicity (logP) of all compounds is predicted to be In conclusion, three new α-fluorocinnamate-based derivatives, 1F, 2F, and 3F have been successfully synthesized and evaluated as covalent inhibitors against a considerable panel of cysteine proteases of the

papain-like family, demonstrating remarkable selectivity towards CatS with single digit micromolar values in the range 1.8–2.6 μM. 2F also exhibited weak inhibitory activity against CatL (29 % at 20 μM) and CatB (20 % at 20 μM). In silico investigations conducted on 2F revealed the most interesting interactions within the catalytic cavities of CatS and CatL, proving to be an interesting starting point for the development of dual inhibitors. Significantly, 2F showed antiproliferative effects against pancreatic cancer cell lines BXPC3 and CAPAN1, with IC50 values of 5.79 μM and 20.75 μM, respectively. These results highlight the promising potential of α-fluorocinnamate derivatives in the development of novel agents targeting CatS and CatL for the treatment of pancreatic cancer. Further studies will be conducted to thoroughly explore the mechanisms of action of this class of compounds, optimize their structure and evaluate their effectiveness. 4. Materials and methods 4.1. Chemistry Unless otherwise stated, reagents and solvents were purchased from Merck (Milan, Italy), Fluorochem (Hadfield, United Kingdom), TCI (Zwijndrecht, Belgium) or BLDPharm (Hamburg, Germany) and used without further purification. All reactions were carried out in oven-dried glass-ware, using dry solvents under nitrogen atmosphere and monitored by TLC on silica gel (Merck precoated 60 F254 plates), with detection by UV light (254 nm) or by permanganate or by HPLC. HPLC was performed on Agilent 1100 Series System using a Gemini 5 μM C18 110 Å LC Column 150 × 3 mm and with a gradient of water/ACN (+0.1 % HCOOH) ranging from 5 % ACN up to 100 % ACN in 30 min (flux of 1.0 mL/min and sample injection of 5 μL), choosing 220 or 254 nm as the wavelengths for the detection of compounds. Products were purified by flash column chromatography, using silica gel Merck 60 (230–400 mesh) as the stationary phase. Purity of the final tested compounds was assured to be >95 % as assessed by NMR and HPLC analysis. 1 H NMR and 13C NMR spectra were recorded at 298 K on a Brüker Avance Spectrometer (400 MHz), using commercially available deuterated solvents (Chloroform-d, MeOD, DMSO-d6). Chemical shifts are reported in parts per million (δ ppm), compared to TMS as an internal standard. Coupling constants (J) are given in hertz (Hz) and are quoted to the nearest 0.5 Hz. Peak multiplicities are described in the following way: s, singlet; bs, broad singlet; d, doublet; m, multiplet; br, broad. Highresolution mass spectra

(HRMS) were recorded using the Q-ToF Synapt G2-Si HDMS Acquity UPLC I-Class Photodiode Detector Array (PDA) (Waters). 4.2. Biomimetic experiments To a solution of 2F (10 mg, 0.02 mmol, 1.0 equiv) in 5 mL of PBS (pH = 7.4) with 10 % of ACN, N-acetylcysteine (3.6 mg, 0.022 mmol, 1.1 equiv) was added and the mixture was left stirring for 1 h at 37 ◦C. The reaction was checked via HPLC-MS and the adduct of 2F with N-acetylcysteine was observed, as reported in Fig. 2. Benzenethiol (97 µL, 0.7 mmol, 5.0 equiv) was added to a single-necked flask charged with 5 (30 mg, 0.14 mmol, 1.0 equiv) dissolved in ACN (2 mL). The mixture was left stirring overnight. Then, the reaction is concentrated under reduced pressure and the adduct was observed as reported in Fig. S1. 4.3. Enzymes The inhibitory activity of the compounds was evaluated by means of a Forster ¨ resonance energy transfer (FRET)-based enzymatic cleavage assay on a TECAN Infinite F2000 PRO plate reader (Agilent Technologies, Santa Clara, USA) using white flat-bottom 96-well microtiter plates (Greiner bio-one, Kremsmünster, Austria).19 Recombinant SARS-CoV-2 Mpro was expressed and purified as previously described,20 whereas the peptidic substrate DabcylKTSAVLQ↓SGFRKME-Edans (TFA salt) was obtained from a commercial source (Genescript, New Jersey, USA). The arrow indicates the cleavage position. The proteolytic activity of the SARS-CoV-2 Mpro was measured by monitoring the increasing fluorescence of SGFRKME-Edans upon hydrolytic shedding of the quencher Dabcyl-KTSAVLQ at 25 ◦C with a 335 nm excitation filter and a 493 nm emission filter. Each well contained 200 µL solution, composed of 185 µL of reaction buffer (20 mM Tris pH=7.5, 0.1 mM EDTA, 1 mM DTT and 200 mM NaCl), 5 µL of SARS-CoV-2 Mpro in enzyme buffer at a final concentration of 50 nM together with 5 µL of the fluorogenic substrate (final concentration 25 µM) and 10 µL of the compounds present at a final concentration of 20 µM (screening assay) The proteolytic activity of the other cysteine proteases, i.e. SARS-CoV-2 PLpro, CatL, CatB and CatS, were measured with the same method in an Infinite 200 PRO microplate reader (TECAN, Mannedorf, ¨ Switzerland). For recombinant SARS-CoV-2 PLpro, expressed and purified as previously described,20 the reaction buffer used was 20 mM Tris-HCl, 200 mM NaCl, 1 mM DTT, 0.1 mM EDTA (pH 7.5), and the screening assay (20 µM) was performed over 10 min. at 25 ◦C with a 360 nm excitation filter and a 460

nm emission filter. Human CatB and CatL were purchased from SIGMA-ALDRICH (Darmstadt, Germany) and screened in the same conditions using 50 mM Tris-HCl, 200 mM NaCl, 5 mM EDTA, 0.005 % Brij35 (pH 6.5) as a reaction buffer and 50 mM Tris-HCl, 200 mM NaCl, 5 mM EDTA, 2 mM DTT (pH 6.5) as an enzyme buffer. Assay buffer was mixed with CatB or CatL in enzyme buffer, then inhibitor in DMSO or DMSO (negative control) was added, followed by 100 µM (CatB) or 6.25 µM (CatL) substrate Z-Phe-Arg-AMC (BACHEM, Bubendorf, Switzerland). For both enzymes the fluorescence was recorded with a 365 nm excitation filter and a 460 nm emission filter. CatS (recombinant from E. coli, SIGMAALDRICH, Darmstadt, Germany) was incubated with enzyme buffer (35 mM KH2PO4, 35 mM CH3COONa, 2 mM DTT, 2 mM EDTA, pH 6.5) at 25 ◦C for ~30 min. Assay buffer (50 mM KH2PO4, 50 mM K2HPO4, 2.5 mM DTT, 2.5 mM EDTA, pH 6.5) was mixed with CatS (final concentration 15 nM) in enzyme buffer, followed by sequential addition of the inhibitor dissolved in DMSO or DMSO (negative control) and fluorogenic substrate, i.e. Z-Val-Val-Arg-AMC ((BACHEM, Bubendorf, Switzerland) with 10 µM final concentration. For the continuous assays, the final concentration of all three compounds was: 0–0.5–1–2.5–5–10–20–30 µM. The inhibitors and the substrate were dissolved and diluted in DMSO, leading to a final DMSO concentration of 7.5 % (v/v). As per the screening assay, the compounds and enzyme were incubated for ~30 min at 25 ◦C prior to substrate addition. The product released from the substrate hydrolysis was monitored in 30 s increments over a period of 10 min. The related KM value was determined in a separate experiment (34 µM). The IC50 value was calculated with GraFit (Version 6.0.12; Erithacus Software Limited, East Grinstead, West Sussex, UK) by fitting the relative enzymatic activities plotted against the respective inhibitor concentration to the four-parameter equation. 4.4. Docking The crystal structures of the CatS and CatL were obtained from the Protein Data Bank, PDB IDs: 4MZO (CatS) and 2XU3 (CatL). As input files for the three compounds the smiles were used. For both the preparation of the input files and the covalent docking calculation ICM-Pro 3.9–4 was used.21 Docking has been performed with biased probability Monte Carlo (BPMC) procedure and on the global optimization of the flexible ligand, and a weighted scoring function has been employed.22,23 For the preparation of the input files

the procedure present on the manual was followed (ICM 3.9–4 User Manual https://molsoft.com/gu i/introduction.html). All the water molecules were eliminated except waters 554, 555 and 558 for CatS; and waters 2138, 2293, 2360 and 2366 for CatL. For the covalent docking, the thoroughness was increased to 10.0, everything else was left on default. The figures were prepared with ICM's graphical user interface. 4.5. Cells The human pancreatic cancer cell line CAPAN1 and BXPC3 were obtained from ATCC, and their authentication has been carried out by the authors within the last 6 months. CAPAN1 cells were maintained in Iscove's Modified Dulbecco's Medium (Euroclone) supplemented with 20 % FBS and 1 % glutamine and BXPC3 cells in RPMI-1640 medium (Euroclone) with 10 % FBS and 1 % glutamine. All cell lines were maintained in a humidified 37 ◦C incubator with 5 % CO2. For cytotoxicity experiments, cells were seeded at 10.000 cells/mL and treated with compounds 1F, 2F and 3F in 96-well plates 48h after plating. After five days, cell viability was examined with the MTS assay (Promega) and absorbance was acquired using a plate reader (GloMax Discover, Promega). Drug–dose response curves were generated and IC50 (drug concentrations inhibiting 50 % of cell growth) calculated by using Prism7.05 (GraphPad Software, San Diego, CA, USA). 4.6. Spectral and characterization data Ethyl 2-fluoro-3-oxo-3-phenylpropanoate (3). Ethyl benzoylacetate (1.0 g, 5.14 mmol, 1.0 equiv) and Selectfluor™ (2.369 g, 6.68 mmol, 1.3 equiv) were dissolved in water/ACN 1:1 (20 mL) and the solution was left stirring at room temperature overnight. The reaction mixture was then concentrated under reduced pressure, diluted with ethyl acetate (12 mL) and washed with brine (3 × 10 mL). The organic layer was dried with anhydrous Na2SO4, filtered and concentrated under reduced pressure. The resulting crude was purified by column chromatography using n-hexane/ethyl acetate (95:5, v/v, then 8:2 v/v) as eluent to afford 3 as a clear liquid (781 mg, 72 % yield). 1 H NMR (400 MHz, CDCl3): δ 8.04 (dt, J = 8.5, 1.2 Hz, 2H, Ar H), 7.64 (dt, J = 7.4, 1.8 Hz, 1H, Ar H), 7.51 (d, J = 8.1 Hz, 2H, Ar H), 5.87 (d, J = 48.8 Hz, 1H, CHF), 4.30 (qd, J = 7.1, 1.9 Hz, 2H, CH2CH3), 1.26 (t, J = 7.1 Hz, 3H, CH2CH3). Spectral data match with those ones previously reported.13 Ethyl (Z)-2-fluoro-3-(4-nitrophenyl)acrylate (4). 3 (781 mg, 3.71 mmol, 1.0 equiv), p-nitro benzaldehyde (673 mg, 4.46 mmol, 1.2 equiv) and Cs2CO3 (2.421 g, 7.43 mmol, 2.0 equiv) were dissolved in ACN

(15 mL) and the solution was left stirring at 40 ◦C overnight. The reaction mixture was then concentrated under reduced pressure, diluted with ethyl acetate (12 mL) and washed with brine (3 × 10 mL). The organic layer was dried with anhydrous Na2SO4, filtered and concentrated under reduced pressure. The resulting crude was purified by column chromatography using n-hexane/ethyl acetate (8:2 v/v) as eluent to afford 4 as a yellow amorphous solid (548 mg, 62 % yield). 1 H NMR (400 MHz, CDCl3): δ 8.26 (d, J = 8.9 Hz, 2H, Ar H), 7.79 (d, J = 8.8 Hz, 2H, Ar H), 6.97 (d, JH–F = 33.7 Hz, 1H, CH=CF), 4.38 (q, J = 7.1 Hz, 2H, CH2CH3), 1.40 (t, J = 7.1 Hz, 3H, CH2CH3). Spectral data match with those ones previously reported.14 Ethyl (Z)-3-(4-aminophenyl)-2-fluoroacrylate (5). 4 (548 mg, 2.29 mmol, 1.0 equiv), Fe powder (767 mg, 13.74 mmol, 6.0 equiv) and NH4Cl (735 mg, 13.74 mmol, 6.0 equiv) were suspended in water/EtOH 1:1 (16 mL) and the solution was left stirring at 80 ◦C for 3 h. The reaction mixture was then filtered on a short pad of celite, concentrated under reduced pressure, diluted with ethyl acetate (12 mL) and washed with brine (3 × 10 mL). The organic layer was dried with anhydrous Na2SO4, filtered, and concentrated under reduced pressure to afford 5 as an orange viscous liquid (373 mg, 78 % yield). 1 H NMR (400 MHz, CDCl3): δ 7.48 (d, J = 8.8 Hz, 2H, Ar H), 6.82 (d, JH–F = 36.2 Hz, 1H, CH—CF), 6.67 (d, J = 8.7 Hz, 2H, Ar H), 4.32 (q, J = 7.2 Hz, 2H, CH2CH3), 3.94 (s, 2H, NH2), 1.37 (t, J = 7.1 Hz, 3H, CH2CH3). 13C NMR (100 MHz, CDCl3): δ 162.0 (d, JC-F = 34.0 Hz, C—O), 148.3, 144.9 (d, JCF = 259.8 Hz, C-F), 132.1 (d, JC-F = 7.8 Hz), 120.9 (d, JC-F = 2.8 Hz), 118.2 (d, JC-F = 3.8 Hz), 114.7, 61.5, 14.2. 19F NMR (376 MHz, CDCl3): δ − 130.5. HRMS (ESI), m/z [M+H]+: calculated for C11H12FNNaO+ 232.0744; found 232.0731. Ethyl (S,Z)-3-(4-(2-(((benzyloxy)carbonyl)amino)-3-phenylpropanamido)phenyl)-2-fluoroacrylate (1F). Cbz-Phe-OH (68 mg, 0.23 mmol, 1.2 equiv), EDCI (112 mg, 0.59 mmol, 3.0 equiv), DMAP (36 mg, 0.29 mmol, 1.5 equiv) and 5 (41 mg, 0.19 mmol, 1.0 equiv) were dissolved in dry DCM (5 mL) and the solution was left stirring at room temperature overnight. The reaction mixture was then diluted with DCM (12 mL) and washed with saturated aq. solution of NH4Cl (3 × 10 mL), saturated aq. solution of NaHCO3 (3 × 10 mL) and brine (3 × 10 mL). The organic layer was dried with anhydrous Na2SO4, filtered and concentrated under reduced pressure. The resulting crude was purified

by column chromatography using n-hexane/ethyl acetate (6:4, v/v) as eluent to afford 1F as a white amorphous solid (27 mg, 29 % yield). 1 H NMR (400 MHz, CDCl3): δ 7.87 (s, 1H, NH), 7.56 (d, J = 8.8 Hz, 2H, Ar H), 7.41–7.15 (m, 12H, Ar H), 6.86 (d, JH–F = 35.3 Hz, 1H, CH—CF), 5.48 (s, 1H, NH), 5.10 (s, 2H, Ar-CH2-O), 4.54 (d, J = 7.2 Hz, 1H, α-CH), 4.34 (q, J = 7.1 Hz, 2H, CH2CH3), 3.12 (dt, J = 13.7, 7.7 Hz, 2H, α-CHCH2-Ar), 1.38 (t, J = 7.1 Hz, 3H, CH2CH3). 13C NMR (100 MHz, CDCl3): δ 169.3, 161.6 (d, JC–F = 34.0 Hz, C—O), 159.3, 146.8 (d, JC–F = 266.1 Hz, C-F), 138.5, 136.3, 136.0, 131.34 (d, JC–F = 7.9 Hz), 129.4, 129.1, 128.8, 128.5, 128.2, 127.4, 120.0, 117.0, 67.5, 62.0, 57.3, 38.6, 14.4. HRMS (ESI), m/z [M+H]+: calculated for C28H28FN2O5 + 491.1977; found 491.1998. HPLC rt: 22.6 min. Methyl (2-(1H-indol-3-yl)acetyl)-L-phenylalaninate (6a). 2-(1Hindol-3-yl)acetic acid (195 mg, 1.11 mmol, 1.2 equiv), EDCI (248 mg, 1.30 mmol, 1.4 equiv), DMAP (159 mg, 1.30 mmol, 1.4 equiv) and Lphenylalanine methyl ester (200 mg, 0.93 mmol, 1.0 equiv) were dissolved in dry DMF (10 mL) and the solution was left stirring at room temperature overnight. The reaction mixture was then diluted with ethyl acetate (12 mL) and washed with saturated aq. solution of NH4Cl (3 × 10 mL), saturated aq. solution of NaHCO3 (3 × 10 mL) and brine (3 × 10 mL). The organic layer was dried with anhydrous Na2SO4, filtered and concentrated under reduced pressure. The residue was treated with n-hexane and the precipitate was filtered and dried under vacuum to afford 6a as a pink amorphous solid (296 mg, 94 % yield). 1 H NMR (400 MHz, CDCl3): δ 8.22 (s, 1H, NH), 7.52 (d, J = 7.8 Hz, 1H, Ar H), 7.40 (d, J = 8.2 Hz, 1H, Ar H), 7.30–7.21 (m, 2H, Ar H), 7.14 (d, J = 7.7 Hz, 1H, Ar H), 7.10 (d, J = 7.4 Hz, 1H, Ar H), 7.08–6.96 (m, 3H, Ar H), 6.72 (d, J = 7.2 Hz, 2H, Ar H), 6.07 (d, J = 8.0 Hz, 1H, NH), 4.94–4.80 (m, 1H, α-CH), 3.72 (s, 2H, –CH2-Indole), 3.65 (s, 3H, OCH3), 3.04–2.88 (m, 2H, α-CH-CH2-Ar). 13C NMR (100 MHz, CDCl3) δ: 171.9, 171.2, 136.5, 135.7, 129.2, 128.5, 127.2, 127.0, 123.7, 122.8, 120.3, 118.9, 111.5, 108.9, 53.0, 52.3, 37.8, 33.4. HRMS (ESI), m/z [M+H]+: calculated for C20H21N2O3 + 337.1547; found 337.1508. Methyl (1H-indole-2-carbonyl)-L-phenylalaninate (6b). 1Hindole-2-carboxylic acid (269 mg, 1.67 mmol, 1.2 equiv), EDCI (397 mg, 2.08 mmol, 1.5 equiv), DMAP (254 mg, 2.08 mmol, 1.5 equiv) and Lphenylalanine methyl ester (300 mg, 1.39 mmol, 1.0 equiv) were dissolved in dry DMF (20 mL) and the solution was left stirring at room temperature overnight. The reaction

mixture was then diluted with ethyl acetate (12 mL) and washed with saturated aq. solution of NH4Cl (3 × 10 mL), saturated aq. solution of NaHCO3 (3 × 10 mL) and brine (3 × 10 mL). The organic layer was dried with anhydrous Na2SO4, filtered and concentrated under reduced pressure. The residue was treated with diethyl ether and the precipitate was filtered and dried under vacuum to afford 6b as a brown amorphous solid (244 mg, 54 % yield). 1 H NMR (400 MHz, CDCl3): δ 9.38 (s, 1H, NH), 7.64 (dd, J = 8.1, 1.2 Hz, 1H, Ar H), 7.44 (dd, J = 8.4, 1.2 Hz, 1H, Ar H), 7.34–7.27 (m, 3H, Ph H), 7.15 (dd, J = 7.4, 1.4 Hz, 3H, Ph H), 6.84 (dd, J = 2.1, 1.0 Hz, 1H, Ar H), 6.66 (d, J = 7.8 Hz, 1H, Ar H), 5.13 (dt, J = 7.6, 5.6 Hz, 1H, α-CH), 3.78 (s, 3H, OCH3), 3.36 – 3.19 (m, 2H, α-CH-CH2-Ar). 13C NMR (100 MHz, CDCl3): δ 172.1, 161.4, 136.8, 135.8, 130.1, 129.5, 128.8, 127.6, 127.4, 124.8, 122.1, 120.8, 112.2, 103.2, 53.5, 52.6, 38.2. HRMS (ESI), m/z [M+H]+: calculated for C19H19N2O3 + 323.1390; found 323.1411. (2-(1H-indol-3-yl)acetyl)-L-phenylalanine (7a). 6a (291 mg, 0.86 mmol, 1.0 equiv), LiOH (207 mg, 8.65 mmol, 10.0 equiv) were dissolved in water/1,4-dioxane 1:1 (12 mL, v/v) and the solution was left stirring at room temperature for 2 h. The reaction mixture was then concentrated under reduced pressure and treated with HCl 1 M. The precipitate was filtered and dried under vacuum to afford 7a as a white amorphous solid (170 mg, 61 % yield,). 1 H NMR (400 MHz, DMSO-d6): δ 10.82 (s, 1H, COOH), 8.13 (d, J = 8.1 Hz, 1H, Indole-NH), 7.41 (d, J = 7.8 Hz, 1H, Ar H), 7.32 (d, J = 8.1 Hz, 1H, Ar H), 7.26 – 7.11 (m, 5H, Ar H), 7.10 – 6.99 (m, 2H, Ar H), 6.92 (t, J = 7.5 Hz, 1H, Ar H), 4.44 (td, J = 9.0, 4.8 Hz, 1H, α-CH), 3.51 (s, 2H, CH2-Indole), 3.05 (dd, J = 13.7, 4.8 Hz, 1H, A part of an AB system), 2.89 (dd, J = 13.7, 9.2 Hz, 1H, B part of an AB system). 13C NMR (100 MHz, DMSO-d6): δ 173.1, 170.6, 137.6, 136.1, 129.1, 128.1, 127.2, 126.3, 123.8, 120.9, 118.7, 118.2, 111.2, 108.6, 53.5, 36.8, 32.3. HRMS (ESI), m/z [M+H]+: calculated for C19H18N2NaO3 + 345.1210; found 345.1278 (1H-indole-2-carbonyl)-L-phenylalanine (7b). 6b (240 mg, 0.74 mmol, 1.0 equiv), LiOH (177 mg, 7.4 mmol, 10.0 equiv) were dissolved in water/1,4-dioxane 1:1 (12 mL, v/v) and the solution was left stirring at room temperature for 2 h. The reaction mixture was then concentrated under reduced pressure and treated with HCl 1 M. The precipitate was filtered and dried under vacuum to afford 7b as a white amorphous solid (190 mg, 83 % yield). 1 H NMR (400 MHz, DMSO-d6): δ 10.11 (s, 1H, NH), 7.69

(dd, J = 8.1, 1.2 Hz, 1H, Ar H), 7.50 (dd, J = 8.4, 1.0 Hz, 1H, Ar H), 7.35–7.26 (m, 3H, Ph H), 7.15 (dd, J = 7.4, 1.4 Hz, 3H, Ph H), 6.84 (s, 1H, Ar H), 6.66 (d, J = 7.8 Hz, 1H, Ar H), 5.18 (m, 1H, α-CH), 3.31 (m, 2H, α-CH-CH2-Ar). 13C NMR (100 MHz, DMSO-d6): δ 173.7, 166.4, 135.6, 135.8, 130.3, 129.0, 127.9, 127.6, 127.4, 123.9, 122.4, 120.0, 112.1, 103.2, 51.9, 38.2. HRMS (ESI), m/z [M+H]+: calculated for C18H16N2NaO3 + 331.1053; found 331.1112. Ethyl (S,Z)-3-(4-(2-(2-(1H-indol-3-yl)acetamido)-3-phenylpropanamido)phenyl)-2-fluoroacrylate (2F). 7a (129 mg, 0.40 mmol, 1.2 equiv), EDCI (192 mg, 1.00 mmol, 3.0 equiv), DMAP (61 mg, 0.50 mmol, 1.5 equiv) and 5 (70 mg, 0.33 mmol, 1.0 equiv) were dissolved in dry DCM (10 mL) and the solution was left stirring at room temperature overnight. The reaction mixture was then diluted with DCM (12 mL) and washed with saturated aq. solution of NH4Cl (3 × 10 mL), saturated aq. solution of NaHCO3 (3 × 10 mL) and brine (3 × 10 mL). The organic layer was dried with anhydrous Na2SO4, filtered and concentrated under reduced pressure. The resulting crude was treated with diethyl ether and the precipitate was collected by filtration to afford 2F as a brown amorphous solid (85 mg, 50 % yield). 1 H NMR (400 MHz, CDCl3): δ 8.59 (s, 1H, NH), 8.22 (s, 1H, NH), 7.53 (d, J = 8.4 Hz, 2H, Ar H), 7.46–7.31 (m, 4H, Ar H), 7.24 (d, J = 7.6 Hz, 1H, Ar H), 7.18–7.07 (m, 4H, Ar H), 7.01 (d, J = 2.3 Hz, 1H, Ar H), 6.98–6.92 (m, 2H, Ar H), 6.85 (d, JH–F=35.4 Hz, 1H, CH—CF), 6.31 (d, J = 7.8 Hz, 1H, NH), 4.84 (q, J = 7.3 Hz, 1H, α-CH), 4.35 (q, J = 7.1 Hz, 2H, CH2CH3), 3.73 (d, J = 2.4 Hz, 2H, CH2-Indole), 3.02 (dd, J = 14.0, 6.5 Hz, 1H, A part of an AB system), 2.95 (dd, J = 13.9, 7.4 Hz, 1H, B part of an AB system), 1.38 (t, J = 7.1 Hz, 3H, CH2CH3). 13C NMR (100 MHz, CDCl3): δ 172.6, 169.3, 161.7 (d, JC–F = 33.5 Hz, C—O), 146.6 (d, JC–F = 265.7 Hz), 138.9 (d, JC–F = 3.0 Hz), 136.5, 136.2, 131.3 (d, JC–F = 8.3 Hz), 129.3, 128.8, 127.2 (d, JC–F = 4.1 Hz), 127.1, 127.0, 123.8, 123.0, 120.5, 120.0, 118.6, 117.2 (d, JC–F = 5.1 Hz), 111.7, 108.4, 62.0, 55.3, 37.4, 33.4, 14.4. HRMS (ESI), m/z [M+H]+: calculated for C30H29FN3O4 + 514.2137; found 514.2111. HPLC rt: 20.4 min. Ethyl (S,Z)-3-(4-(2-(1H-indole-2-carboxamido)-3-phenylpropa namido)phenyl)-2-fluoroacrylate (3F). 7b (108 mg, 0.35 mmol, 1.2 equiv), EDCI (168 mg, 0.88 mmol, 3.0 equiv), DMAP (54 mg, 0.44 mmol, 1.5 equiv) and 5 (61 mg, 0.29 mmol, 1.0 equiv) were dissolved in dry DCM (10 mL) and the solution was left stirring at room temperature overnight. The reaction mixture was then

diluted with DCM (12 mL) and washed with saturated aq. solution of NH4Cl (3 × 10 mL), saturated aq. solution of NaHCO3 (3 × 10 mL) and brine (3 × 10 mL). The organic layer was dried with anhydrous Na2SO4, filtered and concentrated under reduced pressure. The resulting crude was purified by column chromatography using n-hexane/ethyl acetate (5:5, v/v) as eluent to afford 3F as a white solid (17 % yield, 25 mg). 1 H NMR (400 MHz, CDCl3): δ 9.10 (s, 1H, NH), 8.09 (s, 1H, NH), 7.65 (d, J = 8.0 Hz, 1H, Ar H), 7.57 (d, J = 8.8 Hz, 2H, Ar H), 7.41 (dd, J = 8.6, 5.1 Hz, 3H, Ar H), 7.30 (d, J = 4.2 Hz, 5H, Ar H), 7.16 (t, J = 8.0 Hz, 1H, Ar H), 7.01 (d, J = 7.0 Hz, 1H, Ar H), 6.92 (d, J = 2.3 Hz, 1H, Ar H), 6.85 (d, JH–F=35.3 Hz, 1H, CH—CF), 5.01 (q, J = 7.7 Hz, 1H, α-CH), 4.34 (q, J = 7.1 Hz, 2H, CH2CH3), 3.34 (dd, J = 13.7, 6.6 Hz, 1H, A part of an AB system), 3.24 (dd, J = 13.5, 8.1 Hz, 1H, B part of an AB system), 1.37 (t, J = 7.1 Hz, 3H, CH2CH3). 13C NMR (100 MHz, MeOD): δ 172.3, 163.9, 147.7 (d, JC–F = 263.0 Hz), 140.9 (d, JC–F = 3.2 Hz), 138.5, 138.2, 132.1 (d, JC–F = 8.3 Hz), 131.6, 130.6, 130.4, 129.5, 128.9, 128.9, 128.3, 127.9, 125.3, 122.9, 121.2, 118.1 (d, JC–F = 4.5 Hz), 113.0, 105.3, 62.9, 57.2, 39.2, 14.5. HRMS (ESI), m/z [M+H]+: calculated for C29H27FN3O4 + 500.1980; found 500.1957. HPLC rt: 21.8 min. CRediT authorship contribution statement Andrea Citarella: Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. Serena Petrella: Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. Davide Moi: Writing – review & editing, Writing – original draft, Data curation. Alessandro Dimasi: Writing – review & editing, Methodology, Data curation. Tommaso Braga: Investigation, Data curation. Lorenzo Ruberto: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis. Stefano Pieraccini: Writing – original draft, Supervision. Maurizio Sironi: Supervision. Nicola Micale: Writing – original draft, Supervision, Resources, Methodology, Investigation, Formal analysis. Tanja Schirmeister: Writing – original draft, Supervision, Resources, Methodology, Investigation, Formal analysis. Giovanna Damia: Writing – original draft, Visualization, Supervision, Resources, Project administration, Funding acquisition, Formal analysis. Valerio Fasano: Supervision. Alessandra Silvani: Supervision. Clelia Giannini: Supervision. Daniele Passarella: Writing – review & editing,

Supervision, Resources, Project administration, Investigation. Declaration of competing interest The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Andrea Citarella reports financial support was provided by University of Milan. Andrea Citarella reports a relationship with University of Milan that includes: employment. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.