Organic & Biomolecular Chemistry

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/obc

Design, characterization and cellular uptake studies of fluorescence-labeled prototypic cathepsin inhibitors †

Franziska Kohl,[a] Janina Schmitz,[a,b] Norbert Furtmann,[a,c] Anna-Christina Schulz-Fincke,[a] Matthias D. Mertens,[a] Jim Küppers,[a] Marcel Benkhoff,[a] Edda Tobiasch,[b] Ulrike Bartz,[b] Jürgen Bajorath,[c] Marit Stirnberg,[a] and Michael Gütschow*[a]

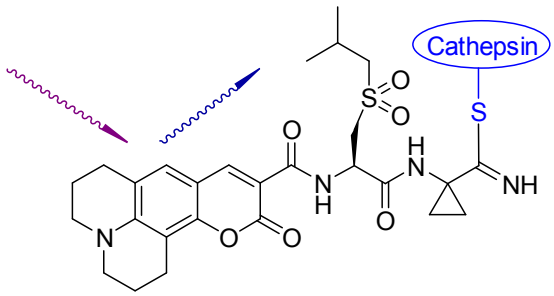
[a] *Pharmaceutical Institute, Pharmaceutical Chemistry I, University of Bonn, An der Immenburg 4, D-53121 Bonn, Germany. E-mail: guetschow@uni-bonn.de; Fax +49 228 732567 : Tel: +49 228 732317.*

[b] *Department of Natural Sciences, Bonn-Rhine-Sieg University of Applied Sciences, von-Liebig-Strasse 20, D-53359 Rheinbach, Germany.*

[c] *Department of Life Science Informatics, B-IT, LIMES Program Unit Chemical Biology and Medicinal Chemistry, University of Bonn, Dahlmannstr. 2, D-53113 Bonn, Germany.*

†Electronic Supplementary Information (ESI) available: Inhibition assays for cathepsins L, S, K, and B, molecular docking, UV and fluorescence spectra, experimental HPLC procedure for log $D_{7.4}$ estimation, as well as ^{13}C NMR and ^{1}H NMR spectra. See DOI:

Table of contents entry



Dipeptide nitriles bearing the coumarin 343 fluorophore were developed as prototypic cysteine cathepsin inhibitors and their cellular uptake was analyzed.

Abstract

Besides their extracellular activity crucial for several pathophysiological conditions, human cysteine cathepsins, in particular cathepsins K and S, represent important intracellular targets for drug development. In the present study, a prototypic dipeptide nitrile inhibitor structure was equipped with a coumarin moiety to function as a fluorescent reporter group. In a second inhibitor, a PEG linker was introduced between the dipeptide nitrile and the fluorophore. These tool compounds **6** and **7** were characterized by kinetic investigations as covalent reversible inhibitors of human cathepsins L, S, K and B. Probe **6** showed a pronounced inhibitory activity against cathepsins K and S, which was corroborated by modeling of inhibition modes. Probe **7** was highly potent (K_i = 93 nM) and selective for cathepsin S. To examine the ability of both probes to enter living cells, human embryonic kidney 293 cells were targeted. At a concentration of 10 μM, cellular uptake of probe **6** was demonstrated by fluorescence measurement after an incubation time of 30 min and 3 h, respectively. The probe's concentration in cell lysates was ascertained on the basis of the emission at 492 nm upon excitation at 450 nm, and the results were expressed as concentrations of probe **6** relative to the protein concentration originating from the lysate. After incubation of 10 μM of probe **6** for 3 h, the cellular uptake was confirmed by fluorescence microscopy. HPLC was used to assess the probes' lipophilicity, and the obtained log $D_{7.4}$ value of 2.65 for probe **6** was in agreement with the demonstrated cellular uptake.

Introduction

The cathepsins comprise a family of lysosomal proteolytic enzymes. Eleven human cathepsins are cysteine proteases of the papain-like subfamily C1A and represent the best characterized group of cathepsins. Cysteine cathepsins are processed to active enzymes during maturation. The final activation of the proenzymes proceeds in the late endosomes either autocatalytically or catalyzed by other lysosomal proteases. The majority of cysteine cathepsins act as endopeptidases, mainly to degrade proteins that have entered the lysosomal system. Nontraditional roles for cathepsins in the extracellular space as well as in the cytosol and nucleus have also been uncovered.[1,2] The cleavage of the peptide bond, catalyzed by cysteine cathepsins, occurs as a two-step acyl transfer. It is initiated by the nucleophilic attack of the active site cysteine leading to the release of the first product. In the subsequent step, the thiol ester bond of the acyl enzyme is hydrolyzed and the second product is formed.[2,3]

Cysteine cathepsin-catalyzed proteolytic cleavage plays a critical role in normal cellular functions as well as pathophysiological events such as osteoporosis, rheumatoid arthritis and osteoarthritis, inflammation, cancer, neurological disorders, multiple sclerosis, pancreatitis, diabetes, and lysosomal storage diseases.[1-4] For example, cathepsin K, the primary enzyme involved in osteoclastic bone resorption, is generally accepted as an important target for the treatment of osteoporosis.[5,6] This cathepsin, when secreted by osteoclasts, is capable to degrade several components of the bone matrix, including type I collagen, the main constituent of the organic bone matrix, as well as osteopontin and osteonectin. However, degradation of collagen mediated by cathepsin K also takes place in the lysosomal vesicles within the osteoclasts,[7,8] and lysosomal collagen degradation was shown to be inhibited by E64d, a membrane-permeable cysteine protease inhibitor.[9] Cathepsins L and B have been demonstrated to play an essential role in the development and progression of cancer. They are involved in the degradation of extracellular matrix proteins, a process promoting tumor angiogenesis, invasion and metastasis of tumor cells.[10.11] Cathepsin

S is crucial in the MHC class II antigen presentation pathway. It represents the major degrading enzyme of the MHC class II invariant chain necessary for loading of antigenic peptides and subsequent antigen presentation by specialized antigen-presenting cell types, such as B cells, macrophages and dendritic cells.[12] Due to the involvement in pathophysiological processes, cysteine cathepsins are well established as targets for drug development. For cathepsins K and S in particular, inhibition of the intracellular activity is desired. Thus, low-molecular weight inhibitors should be capable of entering cells to address their target cathepsins. In this study, two fluorescently labeled prototypic dipeptide nitrile inhibitors for cysteine cathepsins have been designed, their inhibitory profile characterized and their cellular uptake investigated.

Peptidomimetic inhibitors of cysteine cathepsins as well as activity-based probes can be designed by replacing the scissile peptide bond of a peptidic substrate by an electrophilic warhead structure, such as a halomethyl, diazomethyl and acyloxymethyl ketone, an epoxide or aziridine group, a vinyl sulfone or another Michael acceptor, as well as a β-lactam or a cyclopropenone structure. In particular, epoxide derivatives and acyloxymethyl ketones have been successfully devolved to activity-based probes for cysteine cathepsins.[13-17] The peptidic part of such covalent modifiers accounts for sufficient affinity to be accommodated in the active site of the target protease due to specific non-covalent interactions mainly with the S1-S4 binding pockets. The practically irreversible mode of inactivation results from the attack of the active site cysteine nucleophile at an electrophilic carbon of the warhead.[3,10-22] Other types of inhibitors without adjacent leaving groups, such as peptide aldehydes, ketones or nitriles inhibit cysteine proteases as they can sequester the active site cysteine via the reversible formation of thiohemiacetal, thiohemiketal or thioimidate intermediates, respectively. The peptide nitriles have attracted particular attention in the course of inhibitor development for cysteine proteases. Peptide nitriles have been developed against several human cathepsins and, by utilizing the interactions of the cathepsins' specificity pockets with

the corresponding residues of the inhibitors, strong selectivity has been achieved for the target cathepsin.[3,6,12,18,23-26]

Peptide nitriles represent promising drug candidates. For example, the cathepsin K inhibitor odanacatib (**1**, Fig. 1) is currently being developed for the treatment of osteoporosis. The basic compound balicatib (**2**), a further potent inhibitor of cathepsin K, possesses lysosomotropic properties leading to a reduced selectivity in cell-based assays. Both odanacatib and balicatib exhibit the characteristic structural features of low-molecular-weight inhibitors of cysteine cathepsins. In cathepsin K, the S2 pocket is hydrophobic and relatively small which is consistent with the preference for a leucine side chain or a cyclohexane moiety. The S3 pocket of cathepsin K is large and shallow. An extended P3 aryl substituent directly attached to the amide carbonyl (as in **1**, **2** and **5**) was shown to improve potency, e.g. when replacing the simple Cbz group.[6] Strong affinity was obtained with functionalized biaryl groups or with a piperazine directly attached to the phenyl group. During the development of odanacatib (**1**), metabolic liabilities have been minimized. Thus, one amide bond was replaced by the isosteric, non-basic F_3C-CH-NH motif providing an increase in potency and selectivity, and the 1,1-cyclopropane ring was introduced.[6,25] Recently, the activated cathepsin K pool in resorbing osteoclasts was localized using an analogous inhibitor of odanacatib in which the SO_2Me group was replaced by a linker-connected BODIPY fluorophore. The accessibility of cathepsin K in intracellular vesicles and in resorption lacunae was demonstrated with this inhibitor.[27]

The dipeptide nitrile **3** and the azadipeptide nitrile **4** were developed as highly potent inhibitors for cathepsin S. In cathepsin S, the S2 binding site is expanded as two glycine residues (133 and 160, numbering according to PDB ID 2FQ9)[28] replace alanine in cathepsin K (134 and 163, numbering according to PDB ID 1YK7).[29] Moreover, different side chain orientations of Phe205 in cathepsin S and the corresponding Leu209 in cathepsin K result in a deeper S2 pocket in case of cathepsin S. Accordingly, an isobutylsulfonylcysteine P2 building

block was incorporated pointing into the S2 pocket.[24,30] The S3 pocket in cathepsin S is reduced in size and an unsubstituted or *para*-fluoro-substituted phenyl residue in P3 turned out to be advantageous.[24,30,31] The cyclopropanecarbonitrile motif was introduced in both cathepsin K and S inhibitors, e.g. in **1** and **3**. In peptidomimetics with a central cyclopentane scaffold, this led to an increased activity in a cell-based antigen-presentation assay, compared to the corresponding aminoacetonitrile derivative.[26] The radioiodinated inactivator ^{125}I-BIL-DMK (**5**) was successfully utilized for probing active cysteine cathepsins in whole blood and in cell-based, enzyme-occupancy assays.[31,32] After cell lysis, protein separation by SDS-PAGE and quantification by autoradiography, the potency of reversible inhibitors that compete with the irreversible probe for the target enzyme's active site was determined. Moreover, the enzyme-occupancy assays provided information on the cellular permeability of the reversible inhibitors and their affinity towards the cathepsins in a native environment.[31]

Results and discussion

The preparation of the envisaged probe **6** is depicted in Fig. 2. The probe was designed to contain a suitable tetracyclic coumarin fluorophore which is connected via an amide bond to the isobutylsulfonylcysteine at P2 position. The cyclopropanamine moiety at P1 bears the nitrile warhead. Probe **6**, combining structural hallmarks of cysteine cathepsin inhibiting dipeptide nitriles, was expected to exhibit a pronounced affinity for cathepsins S and K. The synthesis of **6** started with alkylation of L-cysteine (**8**) using isobutyl bromide followed by Boc-protection of the amine moiety furnishing compound **9** in an excellent yield. After *S*-oxidation with $KMnO_4$ to afford the sulfone **10**, this was coupled in a mixed-anhydride procedure with 1-aminocyclopropanecarbonitrile. Acid-catalyzed deprotection of **11** yielded the methanesulfonate **12**. Coumarin 343 (**14**) was obtained by Knoevenagel reaction and subsequent heterocyclization from 8-hydroxyjulolidine-9-carboxaldehyde (**13**) and isopropylidene malonate, following a reported method to obtain coumarin-3-carboxylic

acids.[33] The preparation of probe **6** was accomplished by a HATU-promoted amide coupling of compounds **12** and **14** in moderate yield.

The second probe was designed by incorporating a polyethylene glycol (PEG) linker with two oxygen atoms (PEG2) between the dipeptide moiety and the coumarin fluorophore (Fig. 3). The amino alcohol **15** was Cbz-protected at the amino group followed by *O*-alkylation with *tert*-butyl bromoacetate. After deprotection and amide formation with coumarin 343 (**14**), the *tert*-butyl ester **19** was obtained. The ester group was cleaved and a HATU-promoted coupling with building block **12** yielded the probe **7**.

To appraise the requirements of probes **6** and **7** as prototypic cathepsin inhibitors suitable for exploring the cellular uptake, the relevant inhibition kinetics was determined. Both probes were evaluated on human cysteine cathepsins L, S, K and B using fluorogenic or chromogenic peptide substrates. As indicated by linear progress curves of the enzymatic assays, the inhibitors showed 'fast binding' behavior reflecting the rapid formation of the covalent and reversible enzyme-inhibitor thioimidate adduct. Progress curves were analyzed by linear regression and the rates were plotted versus inhibitor concentrations, as exemplarily shown in Fig. 4 for the inhibition of human cathepsin S by probe **6**. By applying the Cheng-Prusoff equation, K_i values (Table 1) were calculated from the obtained IC_{50} values. Probe **6** was found to be particularly potent toward human cathepsins S and K with K_i values of 0.87 μM and 0.26 μM, respectively, being one order of magnitude less active against cathepsins L and B.

Although the isobutylsulfonylcysteine group at P2 was reported to produce unfavorable interactions with Leu209 in cathepsin K,[30] in the case of probe **6** this effect appears to be overcompensated by a favorable interaction of the large P3 fluorophore with the S3 pocket of cathepsin K. On the other hand, probe **6** is a weaker inhibitor of cathepsin S when compared to dipeptide nitriles with the same P2 building block but aromatic N-capping groups of reduced size.[24,30] Covalent docking calculations were carried out to model the

binding of probe **6** to human cathepsins S and K. Fig. 5 shows the putative binding mode of probe **6** within the active site of cathepsin S. Through formation of a covalent thioimidate, the nitrile function of probe **6** is bound to the side chain of residue Cys25. The cyclopropanamine moiety in P1 addresses the S1 binding site. A possible hydrogen bonding interaction was observed between the carbonyl group of the isobutylsulfonylcysteine in P2 position and the backbone NH of residue Gly66. The P2 side chain perfectly fits into the lipophilic environment of the S2 pocket formed by residues Phe67, Met68, Gly133, Val157 and Phe205. However, the tetracyclic coumarin is too large to occupy the S3 pocket of cathepsin S and is orientated towards a solvent exposed area opposite of the S3 pocket. Hence, no beneficial interactions between the fluorescent moiety and the enzyme were predictable. A comparison of the co-crystallized nitrile inhibitor (PDB ID 2FQ9)[28] to the predicted binding mode of probe **6** within the active site of cathepsin S revealed a close spatial match of the backbone peptide structures and similar interaction patterns (see ESI†, Fig. S1).

For cathepsin K, the predicted binding conformation of probe **6** is shown in Fig. 6. Similar to cathepsin S, a thioimidate connection was formed, the cyclopropanamine moiety is directed into the S1 binding site and the P2 isobutylsulfonylcysteine carbonyl group is in hydrogen bonding distance to the backbone NH of residue Gly66. When compared to previous observations,[30] the isobutylsulfonylcysteine did not show any unfavorable contacts to residue Leu209 in our model and nicely fits into the lipophilic S2 pocket. In contrast to the putative binding mode of probe **6** in cathepsin S, the fluorescent moiety is able to bind to the S3 pocket of cathepsin K between residues Asn60, Asp61 and Tyr67, where it might form π-interactions with residue Tyr67. Again, probe **6** mimics the binding mode of the crystallographic (PDB ID 1YK7)[29] cyanamide inhibitor (see ESI†, Fig. S2).

Probe **7** showed the expected strong affinity for cathepsin S with a K_i value of 93 nM and a 10 to 750-fold selectivity for cathepsin S over the other cathepsins investigated. This finding reflects the combination of the P2 residue highly suitable for cathepsin S and the

absence of a bulky group directly attached to the *N*-terminus of the dipeptide. Instead, the PEG2 spacer allows for an orientation of the tetracyclic coumarin at a remote region. As a result, the unfavorable occupation of the coumarin in the S3 pocket is avoided. On the other hand, probe **7** was a weaker inhibitor of cathepsin K than **6**, presumably because the beneficial accommodation of the tetracyclic coumarin in the S3 subsite is prevented by the introduced linker.

The predicted binding modes of **6** within the active sites of cathepsin S and K provided a sound basis for rationalizing the differential inhibitory activities against both enzymes and also allowed conclusion for the binding of **7**, although the high flexibility of **7** complicated predictions of its binding modes. The pronounced activity of probe **6** towards cathepsin K can be explained by the larger S3 pocket of the enzyme that is able to accommodate the coumarin substructure. The different active site geometry around residues Gly59/Glu59 (cathepsin S/cathepsin K), Lys61/Asp61 and Phe67/Tyr67 (see ESI†, Fig. S3) is thought to give rise to the distinct binding conformations of both probes in the two enzymes (see ESI†, Fig. S4).

For the design of probes **6** and **7** coumarin 343 was chosen (**14**). Such rigidified 7-aminocoumarins have been widely used for *in vivo* and *in vitro* diagnostics and imaging.[34-40] The tetracyclic structure of coumarin 343, in which rotation of the amino group is constrained and the nitrogen lone pair can maximally interact with the aromatic system, accounts for a desired bathochromic shift in comparison to less rigid, 7-donor-substituted coumarins.[41,42] Coumarin 343 was expected to be a suitable building block also because of its sufficient fluorescence in aqueous media, biocompatibility and stability. However, coumarin 343 was not used for fluorescence labeling of protease inhibitors so far. As discussed above, the tetracyclic core of coumarin 343 as part of the structure of probes **6** and **7** influenced their binding modes to the active sites of cathepsins S and K.

Absorption and emission spectra of probe **6** were obtained spectrofluorimetrically

from solutions in CH_2Cl_2, MeOH, H_2O and phosphate buffered saline (PBS) (Fig. 7, see also ESI†, Fig. S5, S6). The spectra showed slight bathochromic shifts in polar solvents for both absorption maxima (443 nm in CH_2Cl_2, 441 nm in MeOH, 449 nm in H_2O, 452 nm in PBS) and emission maxima (474 nm in CH_2Cl_2, 484 nm in MeOH, 492 nm in H_2O and PBS), thus resulting in Stokes shifts between 31 nm and 43 nm. The linker-containing compound **7** exhibited similar spectral properties with absorption maxima (438 nm in CH_2Cl_2, 438 nm in MeOH, 449 nm in H_2O, 450 nm in PBS), emission maxima (472 nm in CH_2Cl_2, 482 nm in MeOH, 492 nm in H_2O, 492 nm in PBS) and Stokes shifts between 34 nm and 44 nm (see ESI†, Fig. S7, S8). For further investigations with both probes, an excitation wavelength of 450 nm and an emission wavelength of 492 nm were selected and fluorescence measurements were performed in PBS. This setting was used for the best possible detection of the probes and expected to avoid background fluorescence phenomena resulting from cell lysate components.

Calibration lines were generated for the quantification of probes **6** and **7**. By using different photomultiplier tube (PMT) values, linearity was verified over the entire concentration range of **6** from 0.1 nM to 1000 nM and of **7** from 1 nM to 1000 nM (data not shown). For each curve, ten calibration standards were generated in duplicates. The samples were analyzed fluorometrically and peak heights were plotted against the concentration of the probe. Linear regression analysis was applied to calculate the slope and linear correlation. The relevant calibration line for probe **6** ranging from 1 nM to 10 nM is shown in Fig. 8.

Prior to investigating the cellular uptake of probes **6** and **7** by human embryonic kidney (HEK) 293 cells, the possible influence of the compounds on cell viability was determined. To analyze the cytotoxicity of **6** and **7**, the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay was used, which is based on the reduction of MTT mainly catalyzed by an *N*-ethylmaleimide-sensitive flavin oxidase.[43-44] HEK cells were cultivated for 48 h and treated with **6** or **7** in different concentrations for 3 h. After removal of

the test compound, a 12 mM MTT solution was added and the formation of formazan was measured photometrically after 4 h. The statistical analysis of the MTT assay showed that **6** had only weak cytotoxic effects on HEK cells up to a concentration of 4 µM (Fig. 9). The cell viability decreased to ca. 80% when concentrations of 8 µM, 10 µM and 15 µM were applied. Similarly, compound **7** was not cytotoxic up to 2 µM and reduced the cell viability to ca. 85% at concentrations of 4 µM - 15 µM (see ESI†, Fig. S9).

Fluorescence of cell lysates can be used to assess the uptake of probes and molecular beacons.[45] Several fluorescent compounds have been designed as tools for monitoring the cellular entry of e.g. calixarenes,[46,47] cryptophycins,[48] α-galactosyl ceramide,[49] σ_2 receptor ligands,[50] bile acids,[51] protein kinase inhibitor peptides,[52] lipoproteins,[53] folate conjugates,[54] hepatitis C virus NS5A inhibitors,[55] or vitamin B_{12}.[56] 7-Nitro-2,1,3-benzoxadiazole (NBD) and dansyl served as particularly used fluorophores,[46,49-51,54,57] but also fluorescein,[48,52] indocarbocyanines,[53] rhenium (I) complexes,[56] or BODIPYs have been applied.[55] In the course of this study, two coumarin-labeled compounds were used to assess cellular uptake. Excitation and emission wavelengths of 450 nm and 492 nm, respectively, and a PMT value of 690 V were adjusted to measure the fluorescence intensities of probes **6** and **7**.

HEK cells were treated with either probe **6** or **7** for 10 s, 30 min and 3 h. Two different concentrations of each probe, *i.e.* 1 µM and 10 µM, were employed. To generate cell lysates, the cells were thoroughly washed, mechanically homogenized and the homogenates were fractionated by centrifugation. The protein concentration of each lysate was determined by a modified Bradford method which is based on an absorbance shift of Coomassie Blue.[58] In order to use an equal protein concentration in each fluorescence measurement, the lysates were diluted with PBS to a fixed protein concentration prior to the fluorescence determination. Thus, the protein concentration was used in this study as a relative benchmark for the cellular uptake.[53] This procedure was followed because the portion of cell-based proteins was unknown in the solutions to be analyzed. In pre-screenings it turned out not to be

necessary to include samples derived from untreated cells, because no fluorescence was detected in the corresponding lysates (data not shown). Instead, a very short incubation time of 10 s was chosen to prove the efficiency of the washing steps. It was assumed that the fluorescence quantified in lysates after 10 s incubation time resulted from the probe's association to the cell surface. Thus, it had to be decided whether the lysates prepared after 30 min or 3 h of incubation contained a significantly increased fluorescence compared to those obtained after an incubation time of 10 s. If so, a cellular uptake of the corresponding probe could be concluded.

Incubation of HEK cells with 1 µM of probe **6**, resulted in an only low fluorescence within the cell lysates (Table 2). The values were beyond the calibration line (Fig. 8) and, thus, concentrations of **6** were not determinable and only a limit (< 0.013 pmol/µg protein) could be defined for the relative amount of probe **6** (Table 2). However, incubation of HEK cells with 10 µM of **6** resulted in a quantifiable fluorescence of the lysates. The fluorescence was significantly higher after an incubation time of 30 min than 10 s. Moreover, the fluorescence intensity increased with prolonged incubation time (Fig. 10). These findings confirmed that the fluorescence readout reflected the lysate concentration of **6**. After correction using the value of the 10 s measurement, relative amounts of 0.016 pmol/µg protein and 0.065 pmol/µg protein were determined for the cellular uptake of probe **6** by HEK cells. Since the MTT test revealed that 10 µM of probe **6** only had a weak influence on HEK cell viability even after an incubation time of 3 h (Fig. 9). Hence, a considerable effect of cytotoxicity on the results of the cellular uptake of **6** was not assumed. Experiments for the permeability of probe **7** with concentration of 1 µM or 10 µM produced no significant cellular uptake (Table 2).

The cellular uptake of probe **6** was confirmed by fluorescence microscopy. The results are shown in Fig. 11. HEK cells were treated with **6** and LysoTracker, a red-fluorescent dye that stains acidic compartments in living cells, in particular in lysosomes. In accordance with

the aforementioned experimental setting, incubation with probe **6** was performed either for 10 s or 3 h at a concentration of 10 μM. Cells were fixed and nuclei were stained with DAPI, a blue-fluorescent dye that binds to double-stranded DNA. The fluorescence double-staining with LysoTracker and DAPI is shown in Figs. 11A and 11E. Probe **6** gave a significant green fluorescence only after an incubation time of 3 h (Fig. 11F).

Lipophilicity represents a significant physicochemical property for the prediction of cellular permeation, which is known to primarily depend on the experimental distribution coefficient (log *D*) at pH 7.4.[59] In this study, the distribution coefficient of probes **6** and **7** (ESI†, Fig. S12, S13) was determined by an HPLC-based method.[40,60,61] The calibration line was generated by means of five drugs with known log $D_{7.4}$ values.[60] Their chromatograms were obtained by UV detection, whereas those of analytes **6** and **7** were monitored with the fluorescence channel. Thus, it was also possible to spike the mixture of the calibration compounds with the corresponding probe and simultaneously analyze the resulting chromatogram (ESI†, Fig. S12, S13). The obtained log $D_{7.4}$ value of 2.65, i.e. higher than 2.0,[59] was in agreement with the observed cellular uptake of **6**. However, the presence of the PEG linker in probe **7** decreased the lipophilicity as indicated by the experimentally obtained log $D_{7.4}$ value of 2.30. This might be one reason for the hampered cellular permeation of **7**, which, accordingly, could not be verified by the cellular uptake experiments. The introduction of a fluorescence label can cause changed uptake properties and a subcellular localization in comparison with an analogous, non-fluorescent bioactive compound.[55] However, such differences are expected to be less substantial when incorporating a linker-free fluorophore of rather small size, as was the case for the cell-permeable compound **6**.

Conclusion

Prototypic inhibitors of cysteine cathepsins were equipped with a suitable fluorescence tag and characterized as covalent-reversible inhibitors of human cathepsins. Probe **6** exhibited a

pronounced inhibition for cathepsins K and S, whose intracellular activity is of particular interest for drug development. Probe **7** was shown to be a potent and selective cathepsin S inhibitor. The incorporation of a PEG linker accounted for these different activity profiles. The probes were examined for internalization by HEK cells. At a concentration of 1 μM in cases of **6** and **7**, and a concentration of 10 μM of **7**, permeation into the cells could not be shown. However, when given **6** at a concentration of 10 μM, a time-dependent uptake pattern was demonstrated and confirmed by fluorescence microscopy. The analysis of the lysates after a short incubation time (10 s) was performed to account for the unspecific binding of the fluorescent molecule to the surface of the cells. This technique seems to be applicable for investigating the cellular uptake of pharmacological tool compounds by a common fluorescence microscopy.

Experimental

General methods and materials

Thin-layer chromatography was carried out on Merck aluminum sheets, silica gel 60 F254. Detection was performed with UV light at 254 nm. Preparative column chromatography was performed on Merck silica gel 60 (70-230 mesh). Melting points were determined on a Büchi 510 oil bath apparatus and were uncorrected. Mass spectra were recorded on an API 2000 mass spectrometer (electron spray ion source, Applied Biosystems, Darmstadt, Germany) coupled with an Agilent 1100 HPLC system using a Phenomenex Luna HPLC C_{18} column (50 × 2.00 mm, particle size 3 μm). The purity of the tested compounds was determined by HPLC-DAD obtained on an LC-MS instrument (HPLC Agilent 1100). ^{1}H NMR (500 MHz) and ^{13}C NMR (125 MHz) spectra were recorded on a Bruker Avance DRX 500 and ^{1}H NMR (600 MHz) and ^{13}C NMR (151 MHz) spectra on a Bruker Avance III 600. Chemical shifts δ are given in ppm referring to the signal center using the solvent peaks for reference: DMSO-d_6 2.49 / 39.7 ppm.

(*R*)-*N*-(*tert*-Butyloxycarbonyl)-*S*-(isobutyl)cysteine (9)[24]

A solution of L-cysteine (**8**, 4.85 g, 40 mmol) in a 1:1 EtOH/2N NaOH mixture (80 mL) was treated with isobutyl bromide (6.03 g, 44 mmol) and tetrabutylammonium iodide (0.45 g, 1.21 mmol). After stirring for 3 d at rt, $(Boc)_2O$ (9.60 g, 44 mmol) was added, and the mixture was stirred for 1 d at rt. EtOH was evaporated and the pH of the aqueous residue was brought to ~ 1 with concd HCl. It was extracted with ethyl acetate (3 × 30 mL), washed with 10% $KHSO_4$ (1 × 30 mL) and brine (1 × 30 mL). It was dried (Na_2SO_4) and evaporated to obtain **9** as an oily yellow product (10.5 g, 38 mmol, 95%).

(*R*)-N-(*tert*-Butyloxycarbonyl)-*S*-(isobutyl)cysteinylsulfone-1-aminocyclopropanecarbonitrile (11)

A solution of $KMnO_4$ (9.48 g, 60 mmol) in H_2O (100 mL) was slowly added to a solution of compound **9** (8.32 g, 30 mmol) in AcOH (60 mL), the mixture was stirred at rt for 2.5 h, and a saturated aqueous $KHSO_3$ solution was added until the reaction mixture became colorless. After evaporation, the aqueous suspension was extracted with ethyl acetate (3 × 80 mL). The combined organic layers were washed with H_2O (1 × 25 mL), brine (1 × 25 mL) and dried (Na_2SO_4). The solvent was evaporated to obtain **10** as a colorless oil (8.54 g, 27.6 mmol, 92%) which was used in the next step without further purification. A solution of compound **10** (3.71 g, 12 mmol) in dry THF (28 mL), cooled to -25 °C, was treated with *N*-methylmorpholine (1.21 g, 12 mmol) and isobutyl chloroformate (1.65 g, 12 mmol). A solution of 1-aminocyclopropanecarbonitrile hydrochloride (1.42 g, 12 mmol) in H_2O (2 mL) was treated with 2N NaOH (6.0 mL) and added to the reaction mixture. It was stirred for 2 h, the solvent was evaporated, and the aqueous suspension was extracted with ethyl acetate (3 × 30 mL). The combined organic layers were washed with 10 % $KHSO_4$ (1 × 30 mL), sat. $NaHCO_3$ (1 × 30 mL), H_2O (1 × 30 mL) and brine (1 × 30 mL). After drying (Na_2SO_4), it was

evaporated. The residue was recrystallized from ethyl acetate to give **11** as a white solid (1.08 g, 2.89 mmol, 24% from **10**); mp 190-192°C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 1.01-1.02 (m, 6H, CH(CH_3)$_2$), 1.11-1.16 (m, 2H, CC*H*HC*H*H), 1.38 (s, 9H, C(CH$_3$)$_3$), 1.44-1.46 (m, 2H, CCH*H*CH*H*), 2.18 (sept, ^{3}J = 5.6 Hz, 1H, C*H*(CH$_3$)$_2$), 2.99 (dd, 1H, ^{2}J = 11.5 Hz, ^{3}J = 5.6 Hz, (CH$_3$)$_2$CHC*H*HS), 3.04 (dd, 1H, ^{2}J = 11.6 Hz, ^{3}J = 5.4 Hz, (CH$_3$)$_2$CHCH*H*S), 3.34 (dd, ^{2}J = 12.2 Hz, ^{3}J = 7.4 Hz, 1H, NHCHC*H*HS), 3.43 (dd, ^{2}J = 12.0 Hz, ^{3}J = 3.0 Hz, 1H, NHCHCH*H*S), 4.34-4.38 (m, 1H, NHC*H*CO), 7.27 (d, ^{3}J = 6.9 Hz, 1H, N*H*CHCO), 9.00 (s, 1H, N*H*CCN); ^{13}C NMR (125 MHz, DMSO-d_6) δ 15.78, 15.91, 20.11, 22.52, 22.62, 22.99, 28.24, 49.40, 54.46, 60.15, 78.99, 120.77, 155.00, 170.84; LC-MS(ESI) (90% H_2O to 100% MeOH in 20 min, then 100% MeOH over 10 min, DAD 220-400 nm), 100% purity, t_r = 12.78, *m/z* = 374.3 ([M + H]$^+$), 391.5 ([M + NH$_4$]$^+$).

(*R*)-*S*-(Isobutyl)cysteinylsulfone-1-aminocyclopropanecarbonitrile methanesulfonate (12)

Compound **11** (0.93 g, 2.5 mmol) was dissolved in dry THF (16 mL). Under ice-cooling, dry methanesulfonic acid (1.44 g, 15 mmol) was added, and the resulting reaction mixture was stirred for 5 d at rt. The precipitated product was filtered off, washed with petroleum ether, and dried to obtain **12** as a white solid (0.79 g, 2.13 mmol, 85%); mp > 195 °C, decomp; ^{1}H NMR (500 MHz, DMSO-d_6) δ 1.03-1.05 (m, 6H, CH(CH_3)$_2$), 1.20-1.22 (m, 2H, CC*H*HC*H*H), 1.53–1.55 (m, 2H, CCH*H*CH*H*), 2.22 (sept, ^{3}J = 6.7 Hz, 1H, C*H*(CH$_3$)$_2$), 2.32 (s, 3H, $CH_3SO_3^-$), 3.15 (dd, 1H, ^{2}J = 14.0 Hz, ^{3}J = 6.5 Hz, (CH$_3$)$_2$CHC*H*HS), 3.18 (dd, 1H, ^{2}J = 13.9 Hz, ^{3}J = 6.6 Hz, (CH$_3$)$_2$CHCH*H*S), 3.49 (dd, ^{2}J = 14.7 Hz, ^{3}J = 7.1 Hz, 1H, NHCHC*H*HS), 3.66 (dd, ^{2}J = 14.5 Hz, ^{3}J = 6.0 Hz, 1H, NHCHCH*H*S), 4.18 (t, ^{3}J = 6.5 Hz, 1H, $^{+}$H$_3$NC*H*CO), 8.46 (bs, 3H, NH_3^+), 9.52 (s, 1H, N*H*CCN); ^{13}C NMR (125 MHz, DMSO-d_6) δ 15.69, 15.72, 19.99, 22.42, 23.02, 47.17, 52.98, 60.06, 120.19, 167.53, the $CH_3SO_3^-$ signal is obscured by the DMSO signal; LC-MS(ESI) (90% H_2O to 100% MeOH in 20 min, then 100% MeOH over 10 min, DAD 220-400 nm), 84% purity, t_r = 6.02, *m/z* = 274.3 ([M + H]$^+$).

2,3,6,7-Tetrahydro-11-oxo-1*H*,5*H*,11*H*-[1]benzopyrano-[6,7,8-ij]quinolizine-10-carboxylic acid, coumarin 343 (14)

8-Hydroxyjulolidine-9-carboxaldehyde (**13**, 2.17 g, 10.0 mmol) was dissolved in absolute EtOH (10 mL). Isopropylidene malonate (1.44 g, 10.0 mmol) and piperidinium acetate (61 mg, 0.42 mmol) were added to the solution and stirred for 20 min at rt and refluxed for 2 h. The reaction mixture was allowed to cool down to rt and chilled in an ice bath for 30 min. The product was filtered off and washed with EtOH (10 mL) to yield **14** as a yellow solid (1.19 g, 4.17 mmol, 42%); mp 249-251 °C, lit.[62] mp 253 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 1.84-1.90 (m, 4H, CH_2), 2.69-2.73 (m, 4H, CH_2), 3.32-3.36 (m, 4H, CH_2), 7.22 (s, 1H, 8-H), 8.43 (s, 1H, 9-H), the CO_2H signal could not be detected; ^{13}C NMR (125 MHz, DMSO-d_6) δ 19.64, 20.62, 26.89, 49.26, 49.78, 104.90, 105.51, 107.41, 119.66, 127.57, 148.81, 149.19, 152.77, 160.74, 164.76, the $NCH_2CH_2CH_2$ signal could not be detected; LC-MS (ESI) (90% H_2O to 100% MeOH in 10 min, then 100% MeOH over 10 min, DAD 220-450 nm), 100% purity, *m/z* = 286.1 ($[M + H]^+$).

(*R*)-*N*-[(2,3,6,7-Tetrahydro-11-oxo-1*H*,5*H*,11*H*-[1]benzopyrano-[6,7,8-ij]quinolizin-10-yl)carbonyl]-*S*-(isobutyl)cysteinylsulfone-1-aminocyclopropanecarbonitrile (6)

A mixture of coumarin 343 (**14**, 28 mg, 0.10 mmol), compound **12** (37 mg, 0.10 mmol) and DIPEA (39 mg, 0.30 mmol) was dissolved in CH_2Cl_2 (5 mL). HATU (38 mg, 0.10 mmol) was added and the solution was stirred at rt for 2 h. The reaction mixture was diluted with CH_2Cl_2 (20 mL) and washed with 10% citric acid (3 × 30 mL), saturated $NaHCO_3$ solution (3 × 30 mL) and brine (30 mL). The organic phase was dried over Na_2SO_4, filtered and evaporated in vacuo. The residue was purified by column chromatography using CH_2Cl_2/MeOH (19:1) as eluent. The product **6** was obtained after lyophilization as yellow solid (13 mg, 0.024 mmol, 24%); mp >220 °C decomp; ^{1}H NMR (500 MHz, DMSO-d_6) δ 0.98 (d, 3H, *J* = 6.9 Hz,

CH(CH_3)CH$_3$), 0.99 (d, 3H, J = 6.7 Hz, CH(CH$_3$)CH_3), 1.12-1.18 (m, 2H, CCHHCHH), 1.46-1.47 (m, 2H, CCHHCHH), 1.84-1.92 (m, 4H, N(CH$_2$CH_2CH$_2$)$_2$), 2.17 (sept, 1H, ^3J = 6.6 Hz, CH(CH$_3$)$_2$), 2.71 (t, 2H, ^3J = 6.3 Hz, NCH$_2$CH$_2$CH_2), 2.74 (t, 2H, ^3J = 6.0 Hz, NCH$_2$CH$_2$CH_2), 3.03-3.05 (m, 2H, (CH$_3$)$_2$CHCH_2SO$_2$), 3.32-3.35 (m, 4H, (NCH_2CH$_2$CH$_2$)$_2$), 3.57 (dd, 1H, ^3J = 5.0 Hz, ^2J = 14.7 Hz, NHCHCHH), 3.63 (dd, 1H, ^3J = 7.6 Hz, ^2J = 14.8 Hz, NHCHCHH), 4.89-4.93 (m, 1H, NHCH), 7.27 (s, 1H, 8-H), 8.52 (s, 1H, 9-H), 9.07 (s, 1H, NHCCN), 9.18 (d, 1H, ^3J = 7.9 Hz, NHCH); ^{13}C NMR (125 MHz, DMSO-d_6) δ 15.78, 15.82, 19.67, 19.70, 20.07, 20.61, 22.42, 22.50, 23.02, 29.91, 48.16, 49.18, 49.72, 54.23, 60.43, 104.79, 107.23, 107.45, 119.70, 120.69, 127.45, 147.91, 148.47, 152.37, 161.75, 162.74, 170.33; LC-MS (ESI) (60% H$_2$O to 100% MeOH in 10 min, then 100% MeOH over 10 min, DAD 220-450 nm), 96% purity, m/z = 541.3 ([M + H]$^+$).

Benzyl 2-(2-hydroxyethoxy)ethylcarbamate (16)

Following a similar literature procedure,[63] 2-(2-aminoethoxy)ethanol (**15**, 2.63 g, 25.0 mmol) was dissolved in a mixture of dry CH$_2$Cl$_2$ (23 mL) and triethylamine (2.56 g, 25.0 mmol) and cooled to 0 °C. A solution of benzyl chloroformate (4.26 g, 25.0 mmol) in dry CH$_2$Cl$_2$ (30 mL) was slowly added. The reaction mixture was first stirred for 2 h at 0 °C and then for 18 h at rt. The resulting mixture was washed with saturated NaHCO$_3$ solution (30 mL), whereas the aqueous layer was extracted with CH$_2$Cl$_2$ (3 × 50 mL). The combined organic layers were dried over Na$_2$SO$_4$, filtered and evaporated. The residue was purified by column chromatography using ethyl acetate to obtain **16** as a colorless liquid (5.82 g, 24 mmol, 97%); ^{1}H NMR (600 MHz, DMSO-d_6) δ 3.15 (q, 2H, ^3J = 5.9 Hz, NHCH_2CH$_2$O), 3.40 (t, 2H, ^3J = 5.6 Hz), 3.40 (t, 2H, ^3J = 6.0 Hz, CH_2OCH$_2$, CH$_2$OCH_2), 3.47 (q, 2H, ^3J = 5.3 Hz, OCH$_2$CH_2OH), 4.54 (t, 1H, ^3J = 5.5 Hz, OH), 5.00 (s, 2H, CH_2O(CO)NH), 7.21-7.25 (m, 1H, CONH), 7.28-7.32 (m, 1H, H$_{arom}$), 7.32-7.37 (m, 4H, H$_{arom}$); ^{13}C NMR (151 MHz, DMSO-d_6) δ 40.39, 60.32, 65.35, 69.21, 72.22, 127.84, 127.87, 128.45, 137.32, 156.28; LC-MS (ESI)

(90% H_2O to 100% MeOH in 10 min, then 100% MeOH over 10 min, DAD 200-400 nm), 94% purity, *m/z* = 240.2 ($[M+H]^+$), 257.2 ($[M+NH_4]^+$).

_N_-(Benzyloxycarbonyl)-2-(2-(2-aminoethoxy)ethoxy)acetate _tert_-butyl ester (17)[64]

Compound **16** (5.51 g, 23 mmol) was dissolved in dry THF (90 mL) and was cooled to 0 °C. A volume of 23 mL of a 1 M solution of potassium *tert*-butoxide in dry THF was added. The mixture was stirred for 30 min at 0 °C. Then, *tert*-butyl bromoacetate (5.38 g, 27.6 mmol) was added, it was stirred for 3 h in an ice bath and stirring was continued for 15 h at rt. Subsequently, H_2O (23 mL) was added to the mixture and evaporated. The residue was partitioned between ethyl acetate (1 × 100 mL) and H_2O (1 × 100 mL) and the aqueous layer was extracted with ethyl acetate (2 × 100 mL). The combined organic layers were dried over Na_2SO_4, filtered and evaporated. The residue was purified by column chromatography using ethyl acetate/petroleum ether (3:7) to obtain **17** as a colorless liquid (5.11 g, 63%); ^{1}H NMR (600 MHz, DMSO-d_6) δ 1.41 (s, 9H, C(CH_3)$_3$), 3.14 (q, 2H, ^3J = 5.9 Hz, NHCH_2CH$_2$O), 3.41 (t, 2H, ^3J = 6.0 Hz), 3.49-3.52 (m, 2H), 3.53-3.57 (m, 2H; NHCH$_2$CH_2O, OCH_2CH$_2$O, OCH$_2$CH_2O), 3.96 (s, 2H, OCH$_2$CO), 5.00 (s, 2H, CH_2O(CO)NH), 7.22 (t, 1H, ^3J = 5.7 Hz, CONH), 7.28-7.31 (m, 1H, H_{arom}), 7.31-7.37 (m, 4H, H_{arom}); ^{13}C NMR (151 MHz, DMSO-d_6) δ 27.89, 40.31, 65.34, 68.27, 69.23, 69.58, 69.98, 80.76, 127.82, 127.86, 128.45, 137.32, 156.28, 169.48; LC/MS (ESI) (60% H_2O to 100% MeOH in 10 min, then 100% MeOH over 10 min, DAD 200-400 nm), 76% purity; *m/z* = 354.3 ($[M+H]^+$), 371.4 ($[M+NH_4]^+$).

tert-Butyl 2-(2-(2-aminoethoxy)ethoxy)acetate (18)

Compound **17** (1.48 g, 4.19 mmol) was dissolved in dry MeOH (20 mL) containing Pd/C (10% Pd) and was hydrogenated under atmospheric pressure at rt for 24 h. The catalyst was filtered off and washed twice with MeOH. The removal of the organic solvent under vacuo resulted in a yellow liquid (**18**, 0.87 g, 95%) which was used without further purification in

the next step. ^{1}H NMR (600 MHz, DMSO-d_6) δ 1.41 (s, 9H, C(CH_3)$_3$), 2.65 (t, 2H, ^{3}J = 5.8 Hz, $NH_2CH_2CH_2O$), 3.36 (t, 2H, ^{3}J = 5.8 Hz), 3.49-3.52 (m, 2H), 3.54-3.58 (m, 2H; $NH_2CH_2CH_2O$, OCH_2CH_2O, OCH_2CH_2O), 3.97 (s, 2H, $OCH_2(CO)O$); ^{13}C NMR (151 MHz, DMSO-d_6) δ 28.24, 41.62, 68.62, 69.98, 70.34, 73.10, 81.10, 169.82.

***tert*-Butyl *N*-[(2,3,6,7-tetrahydro-11-oxo-1*H*,5*H*,11*H*-[1]benzopyrano-[6,7,8-ij]quinolizin-10-yl)carbonyl]-2-(2-(2-aminoethoxy)ethoxy)acetate (19)**

A mixture of coumarin 343 (**14**, 0.40 g, 1.40 mmol), compound **18** (0.46 g, 2.10 mmol) and DIPEA (0.36 g, 2.80 mmol) was dissolved in dry CH_2Cl_2 (15 mL). HATU (0.53 g, 1.40 mmol) was added and the solution was stirred at rt for 17 h. The solvent was evaporated and the crude product was suspended in ethyl acetate and washed with 10% $KHSO_4$ (3 × 40 mL), saturated $NaHCO_3$ solution (3 × 40 mL) and brine (40 mL). The organic phase was dried over Na_2SO_4, filtered and evaporated in vacuo. The residue was purified by column chromatography using ethyl acetate as eluent to obtain compound **19** as an orange oil (0.48 g, 0.99 mmol, 71 %); ^{1}H NMR (500 MHz, DMSO-d_6) δ 1.40 (s, 9H, C(CH_3)$_3$), 1.84-1.90 (m, 4H, N($CH_2CH_2CH_2$)$_2$), 2.68-2.73 (m, 4H, N($CH_2CH_2CH_2$)$_2$), 3.30-3.33 (m, 4H, N($CH_2CH_2CH_2$)$_2$), 3.45 (t, 2H, ^{3}J = 4.6 Hz), 3.53 (t, 2H, ^{3}J = 4.7 Hz), 3.55-3.56 (m, 2H), 3.58-3.59 (m, 2H, $NHCH_2CH_2O$, $NHCH_2CH_2O$, OCH_2CH_2O, OCH_2CH_2O), 3.98 (s, 2H, OCH_2CO), 7.23 (s, 1H, 8-H), 8.50 (s, 1H, 9-H), 8.77 (t, 1H, ^{3}J = 4.7 Hz, NH); ^{13}C NMR (125 MHz, DMSO-d_6) δ 19.72, 20.67, 26.93, 27.88, 38.92, 49.14, 49.67, 68.34, 69.20, 69.72, 70.01, 80.72, 104.75, 107.50, 107.96, 119.56, 127.28, 147.65, 148.13, 152.22, 161.97, 162.59, 169.47; LC-MS (ESI) (90% H_2O to 100% MeOH in 10 min, then 100% MeOH over 10 min, DAD 220-450 nm), 87% purity, *m/z* = 487.1 ($[M + H]^+$).

***(R)*-*N*-[*N*-[(2,3,6,7-Tetrahydro-11-oxo-1*H*,5*H*,11*H*-[1]benzopyrano-[6,7,8-ij]quinolizin-10-yl)carbonyl]-2-(2-(2-aminoethoxy)ethoxy)acetyl]-*S*-(isobutyl)cysteinylsulfone-1-**

aminocyclopropanecarbonitrile (7)

Compound **19** (0.44 g, 0.90 mmol) was dissolved in dry CH_2Cl_2 (10 mL). Under ice-cooling, trifluoroacetic acid (10 mL) was added, and the resulting reaction mixture was stirred for 2 h. The solvent was evaporated, and the residue was diluted with CH_2Cl_2 (3 × 10 mL) and evaporated to remove the excess of trifluoroacetic acid. The crude product **20** was used without further purification. A mixture of **20** (0.39 g, 0.90 mmol), methanesulfonate **12** (0.33 g, 0.90 mmol) and DIPEA (1.16 g, 9.00 mmol) was dissolved in CH_2Cl_2 (15 mL). HATU (0.34 g, 0.90 mmol) was added and the solution was stirred at rt for 15 h. The reaction mixture was diluted with CH_2Cl_2 (20 mL) and washed with 10% $KHSO_4$ (3 × 30 mL), saturated $NaHCO_3$ solution (3 × 30 mL) and brine (30 mL). The organic phase was dried over Na_2SO_4, filtered and evaporated in vacuo. The residue was purified by column chromatography, first using CH_2Cl_2/MeOH (19:1) as eluent, followed by second chromatography with ethyl acetate/MeOH (19:1) to obtain compound **7** as a yellow solid (0.10 g, 0.15 mmol, 16%); mp 140-142 °C; ^{1}H NMR (500 MHz, DMSO-d_6) δ 0.99 (d, 3H, J = 6.7 Hz, CH(CH_3)CH_3), 1.00 (d, 3H, J = 6.6 Hz, CH(CH_3)CH_3), 1.12-1.14 (m, 2H, CCHHCHH), 1.45-1.47 (m, 2H, CCHHCHH), 1.84-1.91 (m, 4H, N(CH_2CH_2CH_2)$_2$), 2.17 (sept, 1H, ^3J = 6.7 Hz, CH(CH_3)$_2$), 2.69-2.74 (m, 4H, N(CH_2CH_2CH_2)$_2$), 2.98-3.06 (m, 2H, (CH_3)$_2$CHCH_2SO$_2$), 3.31-3.34 (m, 4H, N(CH_2CH_2CH_2)$_2$), 3.45-3.56 (m, 6H), 3.58-3.64 (m, 4H, NHCH_2CH_2O, OCH_2CH_2O, OCH$_2$CH_2O, NHCH$_2$CH_2O, NHCHCH_2), 3.90-3.99 (m, 2H, OCH$_2$CO), 4.74-4.79 (m, 1H, NHCH), 7.24 (s, 1H, 8-H), 8.13 (d, 1H, ^3J = 8.5 Hz, NHCH), 8.51 (s, 1H, 9-H), 8.79 (t, 1H, ^3J = 5.5 Hz, NHCH$_2$CH$_2$), 9.00 (s, 1H, NHCCN); ^{13}C NMR (125 MHz, DMSO-d_6) δ 15.84, 19.73, 20.12, 20.68, 22.45, 22.52, 22.99, 26.94, 38.91, 47.32, 49.15, 49.68, 54.08, 60.14, 69.25, 69.53, 70.01, 70.42, 104.79, 107.53, 107.96, 119.60, 120.70, 127.29, 147.71, 148.18, 152.24, 162.06, 162.64, 169.75, 170.30; LC-MS (ESI) (90% H_2O to 100% MeOH in 10 min, then 100% MeOH over 10 min, DAD 220-450 nm), 94% purity, m/z = 686.3 ([M + H]$^+$).

Inhibition assays

A stock solution of inhibitor **6** was prepared in DMSO. All inhibition experiments were done in duplicates with five different inhibitor concentrations. To determine the activity of cathepsins S and K, fluorometric assays were performed on a Monaco Safas spectrofluorometer flx using Z-Phe-Arg-AMC (40 μM = 0.74 K_m)[40] or Z-Leu-Arg-AMC (40 μM = 13.3 K_m),[23] respectively. Spectrophotometric assays for cathepsins L and B were carried out on a Varian Cary 50 Bio apparatus using Z-Phe-Arg-pNA (100 μM = 5.88 K_m)[23] or Z-Arg-Arg-pNA (500 μM = 0.45 K_m),[23] respectively. The progress curves of the reactions of **6** were analyzed by linear regression.

Molecular docking

The X-ray structure of human cathepsin K (PDB ID 1YK7)[29] covalently linked to a reversible cyanamide inhibitor and of human cathepsin S (PDB ID 2FQ9)[28] in complex with a nitrile inhibitor were obtained from the RCSB Protein Data Bank.[65] The structures were used as templates for covalent ligand docking with AutoDock 4.2.[66] Bound inhibitors and crystallographic water molecules were removed from the template structure and hydrogen atoms were added using the Molecular Operating Environment (MOE 2012.10)[67] prior to the docking calculations. Gasteiger atomic partial charges were calculated with AutoDock Tools.[66] Probe **6** was linked to the active site residue Cys25 of cathepsin K and S by forming a covalent thioimidate bond to its nitrile function. The side chains of the residues Cys25 including the attached inhibitor structures within the active sites of cathepsin K and S were treated flexibly to predict possible binding modes of probe **6** with AutoDock 4.2. All other active site residues were held in their crystallographic conformation for docking. High-scoring docking poses were visually inspected to select putative compound binding modes.

Photophysical characterization

The excitation wavelength for compounds **6** and **7** were determined by recording UV spectra of 10 µM solutions in CH_2Cl_2, MeOH, H_2O and PBS using a Varian spectrophotometer Cary 50 Bio. Into a cuvette, 990 µL of the solvent, 9 µL of DMSO and 1 µL of a 10 mM solution of **6** or **7** in DMSO were added and spectra were recorded from 200 nm to 600 nm. Spectra in PBS were similarly generated with a 5% DMSO content. Fluorescence spectra were recorded with quartz cuvettes using a flx-Xenius, Safas Monaco, spectrofluorometer from 460 nm to 700 nm with the excitation wavelength of 450 nm. Calibration lines for probes **6** and **7** were obtained by measuring the fluorescence of solutions in duplicates. Spectra were recorded with four concentration ranges. For each concentration range, a different PMT value was adjusted, *i.e.* for compound **6**, 920 V (0.1 nM – 1 nM), 690 V (1 nM – 10 nM), 520 V (10 nM – 100 nM), 390 V (100 nM – 1000 nM), and for compound **7**, 690 V (1 nM – 10 nM), 580 V (10 nM – 100 nM), 420 V (100 nM – 1000 nM). For each concentration range, ten different concentrations of **6** or **7** were applied. The peak heights were plotted against the concentration of **6** or **7** to obtain the calibration lines. For the analysis of the cellular uptake, the calibration line for concentrations of **6** and of **7** between 1 nM and 10 nM (with the same PMT value of 690 V), recorded in PBS containing 0.1% DMSO, was taken as a basis.

Cell viability

For the MTT test, a number of 1.4×10^4 HEK cells per well were seeded in a 96-well plate and cultivated for 48 h under a humidified atmosphere of 5% CO_2 at 37 °C in Dulbecco's modified Eagle's medium (DMEM, Life Technologies, Darmstadt, Germany) (200 µL) supplemented with 10% FCS, 100 U/mL penicillin and 100 µg/mL streptomycin (all from Life Technologies, Darmstadt, Germany) prior to treatment with probe **6** or **7**. DMEM was removed and replaced by 200 µL opti-minimal essential medium (Opti-MEM, Life Technologies, Darmstadt, Germany) containing varying concentration of probe **6** or **7** (0 µM,

0.5 µM, 1 µM, 2 µM, 4 µM, 8 µM, 10 µM, and 15 µM) and 0.1 % DMSO. Cells were incubated for 3 h at 37 °C. The medium was removed and replaced by 100 µL of Opti-MEM, free of fetal calf serum (FCS, Life Technologies, Darmstadt, Germany), and 10 µL of a 12 mM MTT solution in Opti-MEM were added. Cells were incubated for 4 h at 37 °C. A volume of 85 µL was removed and DMSO (50 µL) was added to each well. After carefully shaking the plate for 2 min, the formazan absorbance at 600 nm was measured with a microplate reader (Mithras LB940, Berthold). Cell viability was examined in five independent experiments. Statistical analysis was performed using GraphPad Prism and data were analyzed by ANOVA followed by Tukey's post-hoc test.

***In vitro* uptake studies**

HEK 293 cells were cultivated under a humidified atmosphere of 5% CO_2 at 37 °C in DMEM supplemented with 10% FCS, 100 U/mL penicillin and 100 µg/mL streptomycin. A number of 1×10^6 cells were seeded in 5 ml fresh DMEM medium onto 25 cm^2 flasks for the experiments. After 48 h, medium was removed and the cells were washed with 2 mL PBS. The stock solution of compounds **6** or **7** in DMSO was diluted with FCS-free Opti-MEM. The resulting mixtures containing 1 µM or 10 µM of **6** or **7** and 0.1% DMSO were added to the cells. After 10 s, 30 min and 3 h, the medium was removed and the cells were carefully rinsed with PBS (3 mL). Then, 3 mL of PBS were added and the adherent cells were harvested by scraping with a cell scraper. The cell suspension was transferred into a 15 mL Falcon tube and centrifuged. Pellets were washed by centrifugation (2 ×, with 500 µL PBS) at 4 °C and 500×*g* for 5 min. The pellets were resuspended in PBS (200 µL) and the cell suspensions were homogenized with a 23-gauge needle (10 ×). Cell debris were separated from the lysate by centrifugation (2000×*g*, 5 min, 4 °C; Heraeus Biofuge primo R). The protein content was determined at a Varian Cary 50 Bio spectrophotometer by using the ROTI-Nanoquant protocol (Roth, Karlsruhe, Germany), a modified method of the Bradford assay. The

corresponding calibration line was obtained by duplicate measurements of bovine serum albumin (BSA) (0 – 20 μg/mL). To measure the fluorescence of the lysates, a fixed protein amount of 30 μg (for experiments with **6**) and 60 μg (for experiments with **7**), respectively, was adjusted. The corresponding volume of the lysate was added to PBS to reach a final volume of 400 μL. Fluorescence was measured in quartz cuvettes at 25 °C using a flx-Xenius, Safas Monaco, spectrofluorometer and applying an excitation-emission set at 450 nm and 492 nm, respectively. Lysates were examined in duplicate determinations from five independent experiments. The statistics were performed using GraphPad Prism and data were analyzed by two-tailed unpaired Student's t-test.

Fluorescence microscopy

HEK 293 cells were cultured in 12-well plates under a humidified atmosphere of 5% CO_2 at 37 °C in DMEM supplemented with 10% FCS, 100 U/mL penicillin and 100 μg/mL streptomycin on poly-D-lysine-coated coverslips. Poly-D-lysine was from Sigma Aldrich, Steinheim, Germany. One day later, DMEM was replaced by Opti-MEM after washing the cells twice with PBS. To visualize lysosomes, cells were incubated for 3 h with 100 nM LysoTracker Deep Red (Life Technologies, Darmstadt, Germany). Uptake of compound **6** was analyzed by incubating the cells with 10 μM of **6** for 10 s or 3 h followed by fixation of cells with 4% (w/v) paraformaldehyde in PBS for 30 min at room temperature. Cells were washed with PBS (3 × 500 μL) and subsequently with distilled water (1 × 500 μL). Coverslips were mounted with Roti-Mount FluorCare (Carl Roth, Karlsruhe, Germany) including 10 ng/ml DAPI (2-(4-amidinophenyl)-1*H*-indole-6-carboxamidine; Merck, Darmstadt, Germany) for nuclear staining. Fluorescence signals were recorded with a microscope (Axio Observer D1, Carl Zeiss, Göttingen, Germany) using different filters. For compound **6**, excitation and emission wavelengths were 498 nm and 516 nm, respectively. For LysoTracker, excitation and emission wavelengths were 649 nm and 670 nm, respectively; it was visualized in magenta.

For DAPI, excitation and emission wavelengths were 359 nm and 461 nm, respectively. The exposure times to measure the fluorescence intensities of compound **6**, LysoTracker and DAPI were adjusted (for compound **6**, 3.5 s; for LysoTracker, 0.89 s; for DAPI, 0.08 s).

HPLC procedure for log $D_{7.4}$ estimation

Estimation of log $D_{7.4}$ for probes **6** and **7** was performed as described.[60,61] Analytical HPLC was performed on a Jasco HPLC 2000 instrument with a Polaris C_{18}-A column (2 mm I.D., 50 mm length, 3 µm particle size) from Agilent Technologies. (Santa Clara, CA, USA). Chromatograms were obtained using ChromPass (version 1.8.6.1, Jasco) chromatography software. A linear mobile phase gradient was used with MeCN as mobile phase A and 10 mM ammonium acetate (adjusted to pH 7.4 with ammonium hydroxide and acetic acid) as mobile phase B. The gradient table was as follows, 0 min/0% A, 2.5 min/95% A, 4.0 min/95% A, 4.1 min/0% A, 5.5 min/0% A. All mobile phases were degassed with a Jasco DG-2080-53 degasser. A flow rate of 0.8 mL/min and a column temperature of 40 °C were applied with a Jasco CO-2060 column oven. UV detection wavelength was fixed at 254 nm using a Jasco UV-2075 device and the excitation and emission wavelengths were adjusted to 450 nm and 492 nm, respectively, using a Jasco FP-2020 device. Detector gain was adjusted to 1. A mixture of five organic compounds was dissolved in MeCN, containing 16.8% DMSO to obtain the following concentrations, atenolol 1.5 mg/mL, metoprolol tartrate 4.0 mg/mL, labetalol hydrochloride 0.12 mg/mL, diltiazem hydrochloride 0.075 mg/mL and triphenylene 0.01 mg/mL. To generate the calibration line, this solution was injected into the HPLC. The mean t_R values of three experiments were plotted versus log $D_{7.4}$ literature values.[60] A correlation coefficient of 0.981 was obtained. Identity of the calibration compounds in the injected mixture was confirmed by comparison with the t_R values of the single compounds. Probe **6** or **7** was dissolved in MeCN, containing 16.8% DMSO and injected into the HPLC and the log $D_{7.4}$ value was obtained from the mean t_R value of three experiments.

Acknowledgement

The authors thank Dorothee Schipper and Stephan Bruns for assistance. J.S. is supported by the Gender Equality Center of the Bonn-Rhein-Sieg University of Applied Sciences, and N.F. by a fellowship from the Jürgen Manchot Foundation, Düsseldorf, Germany. M.S. is supported by the German Research Foundation (STI 660/1-1).

Abbreviations used

BSA, bovine serum albumin; DAPI, 2-(4-amidinophenyl)-1*H*-indole-6-carboxamidine; DIPEA, *N*,*N*-diisopropylethylamine; DMEM, Dulbecco's modified Eagle's medium; FCS, fetal calf serum; HATU, *N*-[(dimethylamino)-1*H*-1,2,3-triazolo-[4,5-*b*]pyridin-1-ylmethylene]-*N*-methylmethanaminiumhexafluoro-phosphate *N*-oxide; HEK, human embryonic kidney; MTT, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide; NBD, 7-nitro-2,1,3-benzoxadiazole; Opti-MEM, opti-minimal essential medium; PBS, phosphate buffered saline; PEG, polyethylene glycol; PMT, photomultiplier tube.

References

1 J. Reiser, B. Adair and T. Reinheckel, *J. Clin. Invest.,* 2010, 120, 3421.

2 V. Turk, V. Stoka, O. Vasiljeva, M. Renko, T. Sun, B. Turk and D. Turk, *Biochim. Biophys. Acta*, 2012, 1824, 68.

3 M. Frizler, M. Stirnberg, T. S. Sisay and M. Gütschow, *Curr. Top. Med. Chem.,* 2010, 10, 294.

4 C. Palermo and J. A. Joyce, *Trends Pharmacol. Sci*., 2008, 29, 22.

5 M. Novinec and B. Lenarčič, *Biol. Chem.,* 2013, 394, 1163.

6 W. C. Black, *Curr. Top. Med. Chem.,* 2010, 10, 745.

7 W. S. Hou, Z. Li, R. E. Gordon, K. Chan, M. J. Klein, R. Levy, M. Keysser, G.

Keyszer and D. Brömme, *Am. J. Pathol.,* 2001, 159, 2167.

8 P. Leung, M. Pickarski, Y. Zhuo, P. J. Masarachia and L. T. Duong, *Bone,* 2011, 49, 623.

9 L. Kjøller, L. H. Engelholm, M. Høyer-Hansen, K. Danø, T. H. Bugge and N. Behrendt, *Exp. Cell Res.,* 2004, 293, 106.

10 T. Reinheckel, C. Peters, A. Krüger, B. Turk and O. Vasiljeva, *Front. Pharmacol.,* 2012, 11, 3:133.

11 J. Kos, A. Mitrović and B. Mirković, *Future Med. Chem.,* 2014, 6, 1355.

12 A. Lee-Dutra, D. K. Wiener and S. Sun, *Expert Opin. Ther. Pat.,* 2011, 21, 311.

13 M. G. Paulick and M. Bogyo, *ACS Chem. Biol.* 2011, 6, 563.

14 M. Verdoes, L. E. Edgington, F. A. Scheeren, M. Leyva, G. Blum, K. Weiskopf, M. H. Bachmann, J. A. Ellman and M. Bogyo, *Chem. Biol.* 2012, 19, 619.

15 M. Verdoes, K. O. Bender, E. Segal, W. A. van der Linden, S. Syed, N. P. Withana, L. E. Sanman and M. Bogyo, *J. Am. Chem. Soc.* 2013, 135, 14726.

16 L. I. Willems, H. S. Overkleeft and S. I. van Kasteren, *Bioconjug. Chem.* 2014, 25, 1181.

17 S. D. Wiedner, L. N. Anderson, N. C. Sadler, W. B. Chrisler, V. K. Kodali, R. D. Smith and A. T. Wright, *Angew. Chem. Int. Ed.* 2014, 53, 2919.

18 R. Frlan and S. Gobec, *Curr. Med. Chem.,* 2006, 13, 2309.

19 R. Vicik, M. Busemann, K. Baumann and T. Schirmeister, *Curr. Top. Med. Chem.,* 2006, 6, 331.

20 M. M. Santos and R. Moreira, *Mini Rev. Med. Chem.,* 2007, 7, 1040.

21 S. Serim, U. Haedke and S. H. Verhelst, *ChemMedChem,* 2012, 7, 1146.

22 R. Ettari, A. Pinto, L. Tamborini, I. C. Angelo, S. Grasso, M. Zappalà, N. Capodicasa, Yzeiraj, E. Gruber, M. N. Aminake, G. Pradel, T. Schirmeister, C. De Micheli and P. Conti, *ChemMedChem,* 2014, 9, 1817.

23 M. Frizler, F. Lohr, N. Furtmann, J. Kläs and M. Gütschow, *J. Med. Chem.,* 2011, 54, 396.

24 M. Frizler, J. Schmitz, A. C. Schulz-Finke and M. Gütschow, *J. Med. Chem.,* 2012, 55, 5982.

25 J. Y. Gauthier, N. Chauret, W. Cromlish, S. Desmarais, T. le Duong, J. P. Falgueyret, D. B. Kimmel, S. Lamontagne, S. Léger, T. LeRiche, C. S. Li, F. Massé, D. J. McKay, D. A. Nicoll-Griffith, R. M. Oballa, J. T. Palmer, M. D. Percival, D. Riendeau, J. Robichaud, G. A. Rodan, S. B. Rodan, C. Seto, M. Thérien, V. L. Truong, M. C. Venuti, G. Wesolowski, R. N. Young, R. Zamboni and W. Black, C. *Bioorg. Med. Chem. Lett.,* 2008, 18, 923.

26 H. Hilpert, H. Mauser, R. Humm, l. Anselm, H. Kuehne, G. Hartmann, S. Gruener, D. W. Banner, J. Benz, B. Gsell, A. Kuglstatter, M. Stihle, R. Thoma, R. A. Sanchez, H. Iding, B. Wirz and W. Haap, *J. Med. Chem.,* 2013, 56, 9789.

27 Y. Zhuo, J. Y. Gauthier, W. C. Black, M. D. Percival and L. T. Duong, *Bone,* 2014, 67, 269.

28 J. R. Somoza, Unpublished, RCSB Protein Data Bank release March 21, 2006; PDB-ID 2FQ9.

29 D. N. Deaton, A. M. Hassell, R. B. McFadyen, A. B. Miller, L. R. Miller, L. M. Shewchuck, F. X. Tavares, D. H. Willard and L. L. Wright, *Bioorg. Med. Chem. Lett.,* 2005, 15, 1815.

30 J. Y. Gauthier, W. C. Black, I. Courchesne, W. Cromlish, S. Desmarais, R. Houle, S. Lamontagne, C. S. Li, F. Massé, D. J. McKay, M. Ouellet, J. Robichaud, J. F. Truchon, V. L. Truong, Q. Wang and M. D. Percival, *Bioorg. Med. Chem. Lett.,* 2007, 17, 4929.

31 W. C. Black and M. D. Percival, *ChemBioChem,* 2006, 7, 1525.

32 A. Veilleux, W. C. Black, J. Y. Gauthier, C. Mellon, M. D. Percival, P. Tawa and J. P. Falgueyret, *Anal. Biochem.,* 2011, 411, 43.

33 A. Song, X. Wang and K. S. Lam, *Tetrahedron Lett.,* 2003, 44, 1755.

34 T. Terai and T. Nagano, *Curr. Opin. Chem. Biol.,* 2008, 12, 515.

35 R. S. Agnes, F. Jernigan, J. R. Shell, V. Sharma and D. S. Lawrence, *J. Am. Chem. Soc.,* 2010, 132, 6075.

36 T. Yoshihara, Y. Yamaguchi, M. Hosaka, T. Takeuchi and S. Tobita, *Angew. Chem. Int. Ed.,* 2012, 51, 4148.

37 S. Nizamov, K. I. Willig, M. V. Sednev, V. N. Belov and S. Hell, *Chem. Eur. J.,* 2012, 18, 16339.

38 T. Murase, T. Yoshihara, K. Yamada and S. Tobita, *Bull. Chem. Soc. Jpn.,* 2013, 86, 510.

39 L. G. Meimetis, J. C. Carlson, R. J. Giedt, R. H. Kohler and R. Weissleder, *Angew. Chem. Int. Ed.,* 2014, 53, 7531.

40 M. D. Mertens, J. Schmitz, M. Horn, N. Furtmann, J. Bajorath, M. Mareš and M. Gütschow, *ChemBioChem,* 2014, 15, 955.

41 G. Jones II, S. F. Griffin, C. Y. Choi and W. R. Bergmark, *J. Org. Chem.,* 1984, 49, 2705.

42 M. R. Webb and J. E. Corrie, *Biophys. J.,* 2001, 81, 1562.

43 Mosmann T. *J. Immunol. Methods,* 1983, 65, 55.

44 Y. Liu, D. A. Peterson, H. Kimura and D. Schubert, *J. Neurochem.,* 1997, 69, 581.

45 Y. J. Seo, H. S. Jeong, E. K. Bang, G. T. Hwang, J. H. Jung, S. K. Jang and B. H. Kim, *Bioconjug. Chem.*, 2006, 17, 1151.

46 R. Lalor, H. Baillie-Johnson, C. Redshaw, S. E. Matthews, and A. Mueller, *J. Am. Chem. Soc.,* 2008, 130, 2892.

47 C. Redshaw, M. R. Elsegood, J. A. Wright, H. Baillie-Johnson, T. Yamato, S. De Giovanni and A. Mueller, *Chem. Commun.,* 2012, 48, 1129.

48 M. Nahrwold, C. Weiß, T. Bogner, F. Mertink, J. Conradi, B. Sammet, R. Palmisano,

S. Royo Gracia, T. Preuße and N. Sewald, *J. Med. Chem*., 2013, 56, 1853.

49 J. M. Cheng, S. H. Chee, D. A. Knight, H. Acha-Orbea, I. F. Hermans, M. S. Timmer, and B. L. Stocker, *Carbohydr. Res.,* 2011, 346, 914.

50 C. Abate, J. R. Hornick, D. Spitzer, W. G. Hawkins, M. Niso, R. Perrone and F. Berardi, *J. Med. Chem.,* 2011, 54, 5858.

51 F. Májer, J. J. Salomon, R. Sharma, S. V. Etzbach, M. N. Najib, R. Keaveny, A. Long, J. Wang, C. Ehrhardt and J. F. Gilmer, *Bioorg. Med. Chem.,* 2012, 20, 1767.

52 J. Brugnano, J. McMasters and A. Panitch, *J. Pept. Sci.,* 2013, 19, 629.

53 D. Teupser, J. Thiery, A. K. Walli and D. Seidel, *Biochim. Biophys. Acta,* 1996, 1303, 193.

54 G. M. Consoli, G. Granata, G. Fragassi, M. Grossi, M. Sallese and C. Geraci, *Org. Biomol. Chem.* 2015, 13, 3298.

55 L. H. Jones, D. Beal, M. D. Selby, O. Everson, G. M. Burslem, P. Dodd, J. Millbank, T. D. Tran, F. Wakenhut, E. J. Graham and P. Targett-Adams, *J. Chem. Biol.* 2011, 4, 49.

56 A. R. Vortherms, A. R. Kahkoska, A. E. Rabideau, J. Zubieta, L. L. Andersen, M. Madsen and R. P. Doyle, *Chem. Commun.,* 2011, 47, 9792.

57 A. Mazzaglia, A. Valerio, N. Micali, V. Villari, F. Quaglia, M. A. Castriciano, L. M. Scolaro, M. Giuffrè, G. Siracusano and M. T. Sciortino, *Chem. Commun.,* 2011, 47, 9140.

58 T. Zor and Z. Selinger, *Anal. Biochem.,* **1996**, *236*, 302-308.

59 T. H. Hou, W. Zhang, K. Xia, X. B. Qiao and X. J. Xu, *J. Chem. Inf. Comput. Sci.,* 2004, 44, 1585.

60 E. H. Kerns, L. Di, S. Petusky, T. Kleintop, D. Huryn, O. McConnell and G. Carter, *J. Chromatogr. B, Analyt. Technol. Biomed. Life Sci.,* 2003, 791, 381.

61 S. Blättermann, L. Peters, P. A. Ottersbach, A. Bock, V. Konya, C. D. Weaver, A. Gonzalez, R. Schröder, R. Tyagi, P. Luschnig, J. Gäb, S. Hennen, T. Ulven, L.

Pardo, K. Mohr, M. Gütschow, A. Heinemann and E. Kostenis, *Nat. Chem. Biol.,* 2012, 8, 631.

62 J. Van Gompel and G. B. Schuster, *J. Org. Chem.,* 1987, 52, 1465.

63 Z. Yu, R. M. Schmaltz, T. C. Bozeman, R. Paul, M. J. Rishel, K. F. Tsosie and S. M. Hecht, *J. Am. Chem. Soc.* 2013, 135, 2883.

64 M. Adamczyk, J. R. Fishpaugh and M. Thiruvazhi, *Org. Prep. Proced. Int.* 2002, 34, 326.

65 *RCSB Protein Data Bank.* http://www.rcsb.org.

66 G. M. Morris, R. Hu ey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, and A. J. Olson, *J. Comput. Chem.,* 2009, 16, 2785.

67 Molecular Operating Environment (MOE), 2012.10; Chemical Computing Group Inc., Montreal, Canada.

Table 1 Kinetic parameters of cathepsin inhibition by compound **6**

compd	K_i (μM) ± SEM [a]			
	cat L	cat S	cat K	cat B
6	8.9 ± 0.6	0.87 ± 0.02	0.26 ± 0.03	9.7 ± 1.2
7	71 ± 7	0.093 ± 0.003	0.93 ± 0.06	53 ± 4

[a] IC_{50} values were obtained from duplicate measurements in the presence of five different inhibitor concentrations. Progress curves were analyzed by linear regression. IC_{50} values were determined by nonlinear regression using equation $v_s = v_0/(1+[I]/IC_{50})$, where v_s is the steady-state rate, v_0 is the rate in the absence of the inhibitor, and [I] is the inhibitor concentration. Standard errors of the mean refer to the non-linear regression analysis. K_i values were calculated from IC_{50} values by applying the equation $K_i = IC_{50}/(1+[S]/K_m)$, where [S] is the substrate concentration.

Table 2 Cellular uptake of probes **6** and **7**

probe	concn (μM)	incubation time	fluorescence ± SD (FU)[a]	relative amount in the cell lysate ± SD (pmol/μg protein)
6	1	10 s	3.93 ± 2.82[b]	< 0.013
6	1	30 min	3.44 ± 0.61[b]	< 0.013
6	1	3 h	5.92 ± 3.20[b]	≤ 0.013
6	10	10 s	8.24 ± 4.05	0.018[c]
6	10	30 min	15.84 ± 4.90	0.034[c]
6	10	3 h	40.03 ± 10.77	0.083[c]
7	1	10 s	5.12 ± 0.77[d]	< 0.007
7	1	30 min	3.96 ± 0.46[d]	< 0.007
7	1	3 h	5.83 ± 1.69[d]	< 0.007
7	10	10 s	6.40 ± 2.13[e]	< 0.007
7	10	30 min	11.75 ± 2.58[e]	0.014[c]
7	10	3 h	12.51 ± 1.95[e]	0.015[c]

[a] Lysates were examined in duplicate determinations from five independent experiments. All fluorescent measurements were carried out with a PMT value of 690 V. [b] The differences of the means are not significant. [c] Relative amounts in the cell lysate were calculated from the fluorescence units given. For significant differences, see Fig. 10. [d] The differences of the means are not significant. [e] The differences of the means are not significant.

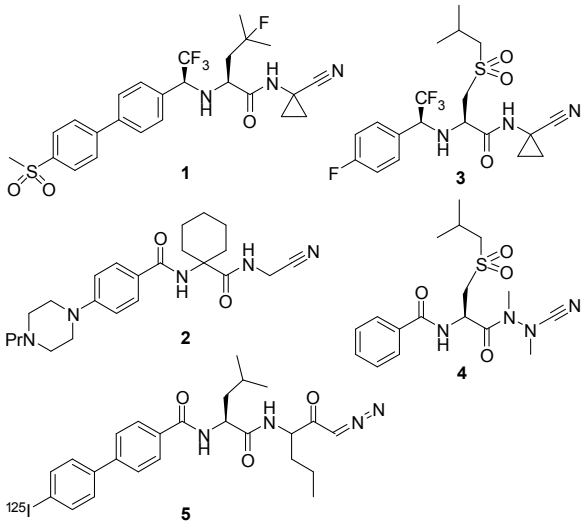


Fig. 1 Structures of peptidomimetic cathepsin inhibitors.

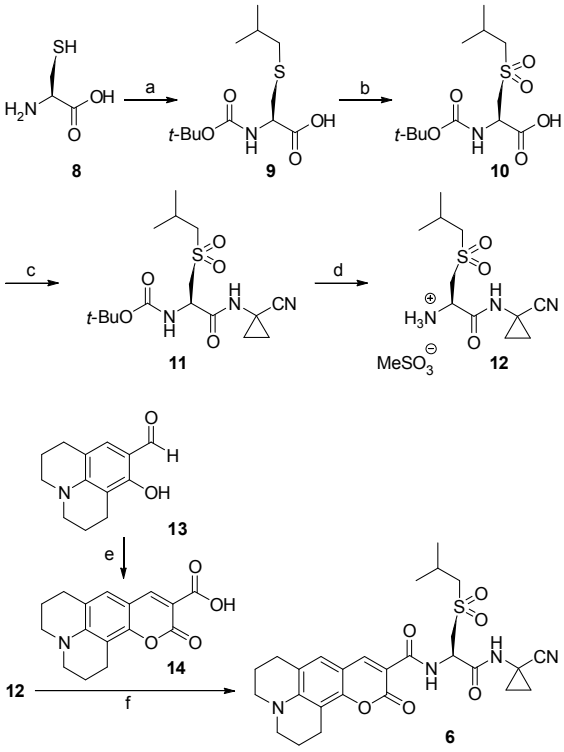


Fig. 2 Synthesis of the probe **6**. (a) 1. isobutyl bromide, tetrabutylammonium iodide, 2N NaOH, EtOH, rt; 2. $(Boc)_2O$, 2N NaOH, EtOH, rt; (b) $KMnO_4$, AcOH, H_2O, rt; (c) 1. *N*-methylmorpholine, $ClCO_2i$-Bu, THF, -25 °C; 2. 1-aminocyclopropanecarbonitrile × HCl, 2N NaOH, H_2O, -25 °C to rt; (d) $MeSO_3H$, THF, rt; (e) 1. isopropylidene malonate, piperidinium acetate, EtOH, 20 min, rt; 2. 2 h, reflux; 3. 30 min, 0 °C; (f) *N*-[(dimethylamino)-1*H*-1,2,3-triazolo-[4,5-*b*]pyridin-1-ylmethylene]-*N*-methylmethanaminiumhexafluoro-phosphate *N*-oxide (HATU), *N*,*N*-diisopropylethylamine (DIPEA), CH_2Cl_2, rt.

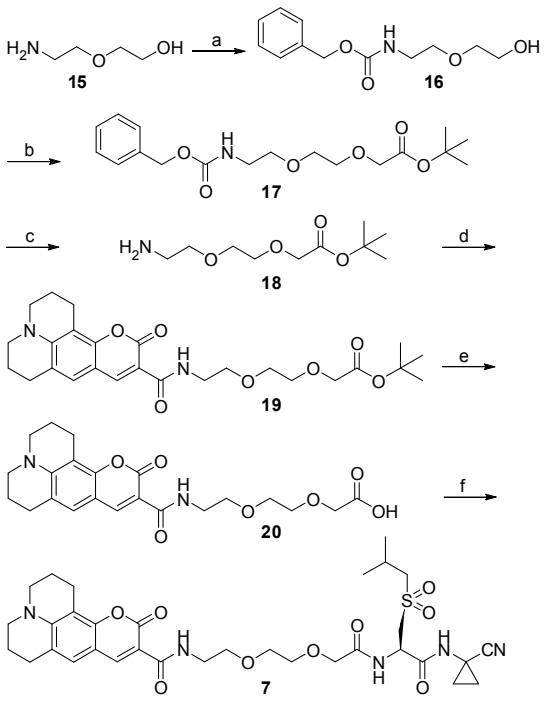


Fig. 3 Synthesis of probe **7**. (a) benzyl chloroformate, triethylamine, CH_2Cl_2, 0 °C to rt; (b) 1. *tert*-butyl bromoacetate, potassium *tert*-butoxide, THF, 3.5 h, 0 °C; 2. 15 h, rt; (c) Pd/C, MeOH, rt; (d) compound **14**, HATU, DIPEA, CH_2Cl_2, rt; (e) trifluoroacetic acid, CH_2Cl_2, 0 °C; (f) compound **12**, HATU, DIPEA, CH_2Cl_2, rt.

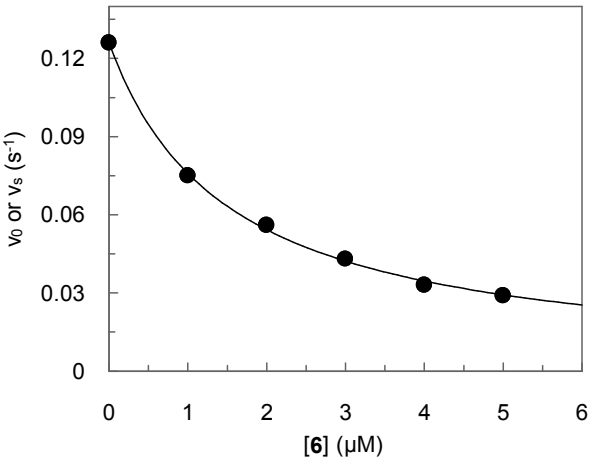


Fig. 4 Inhibition of human cathepsin S by probe **6**. Plot of the rates obtained in duplicate measurements versus concentrations of **6**. Nonlinear regression gave an IC_{50} value of 1.52 (± 0.04) μM.

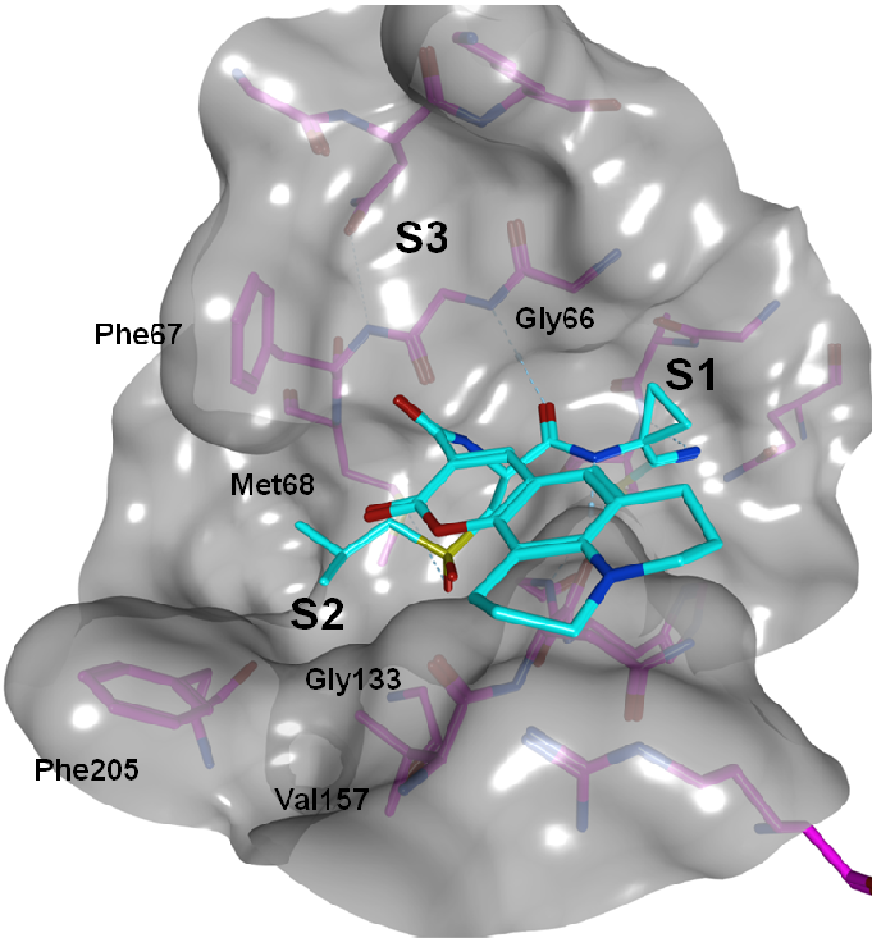


Fig. 5 Shown is the predicted binding mode of probe **6** (cyan) within the active site of cathepsin S. Active site residues are depicted in magenta and the solvent-accessible surface is rendered transparent.

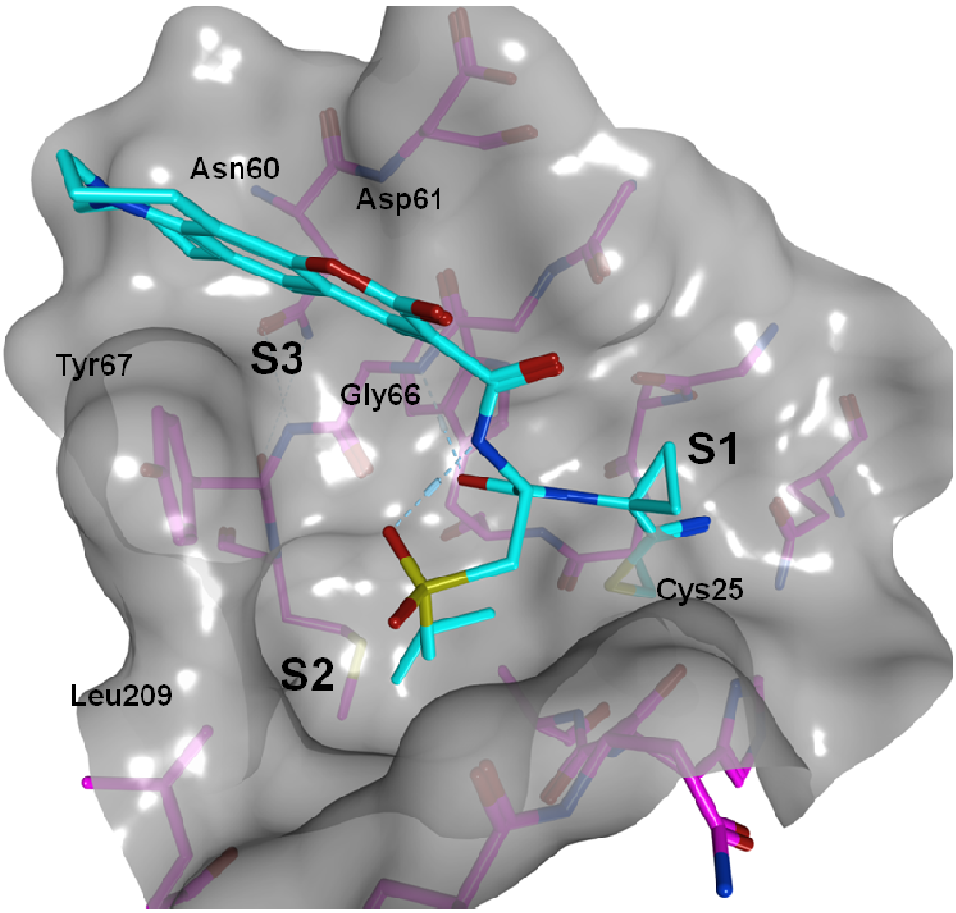


Fig. 6 Shown is the predicted binding mode of probe **6** (cyan) within the active site of cathepsin K. Active site residues are depicted in magenta and the solvent-accessible surface is rendered transparent.

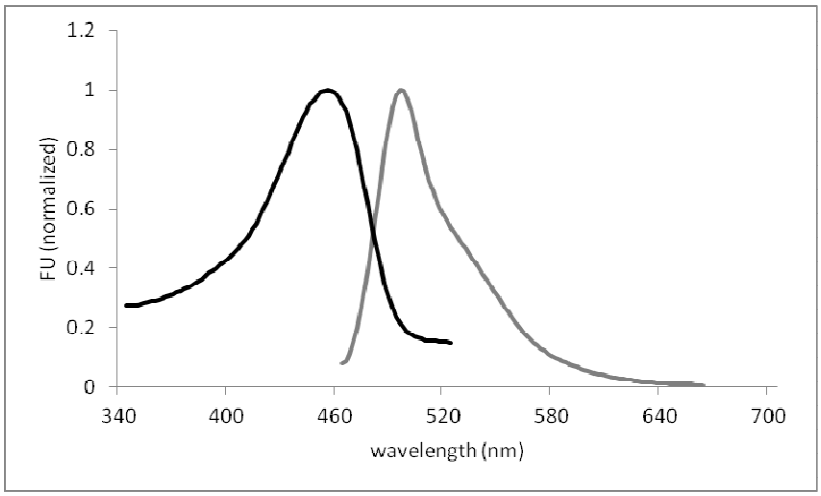


Fig. 7 Normalized absorption (black) and emission (grey) spectra of probe **6**, recorded in PBS (with 5% DMSO) at concentrations of 10 μM (absorption) and 1 μM (emission).

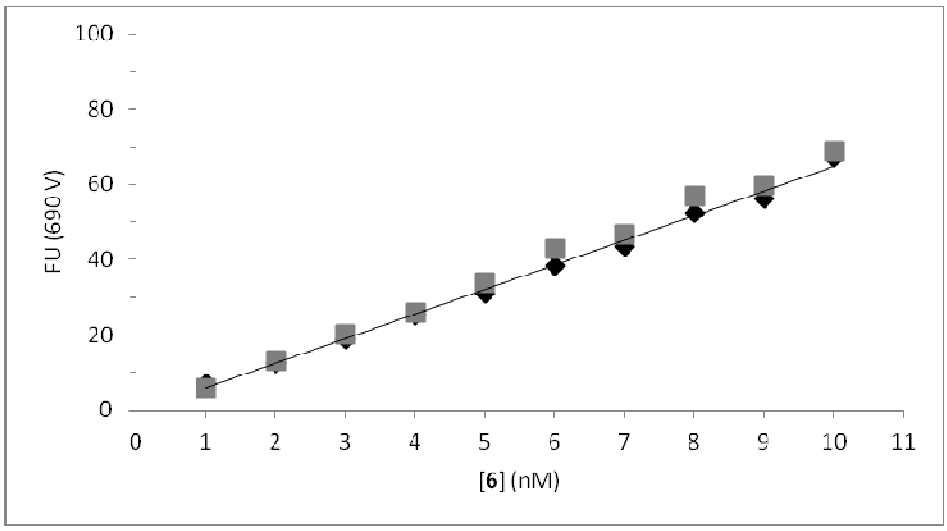


Fig. 8 Calibration line of **6** between 1 nM and 10 nM. Squares and diamonds indicate two measurements. Probe **6** was dissolved in PBS containing 0.1% DMSO. Excitation and emission wavelengths were 450 nm and 492 nm, respectively. The fluorescent units (FU) were determined after adjustment of a PMT value of 690 V. Linear regression gave the equation $y = 6.5275 x - 0.6592$ and $R^2 = 0.996$. Similarly, a calibration line of **7** between 1 nM and 10 nM was generated from solutions of **7** in PBS containing 0.1% DMSO by using the same PMT value of 690 V. Linear regression gave the equation $y = 4.1229 x + 3.0817$ and $R^2 = 0.984$.

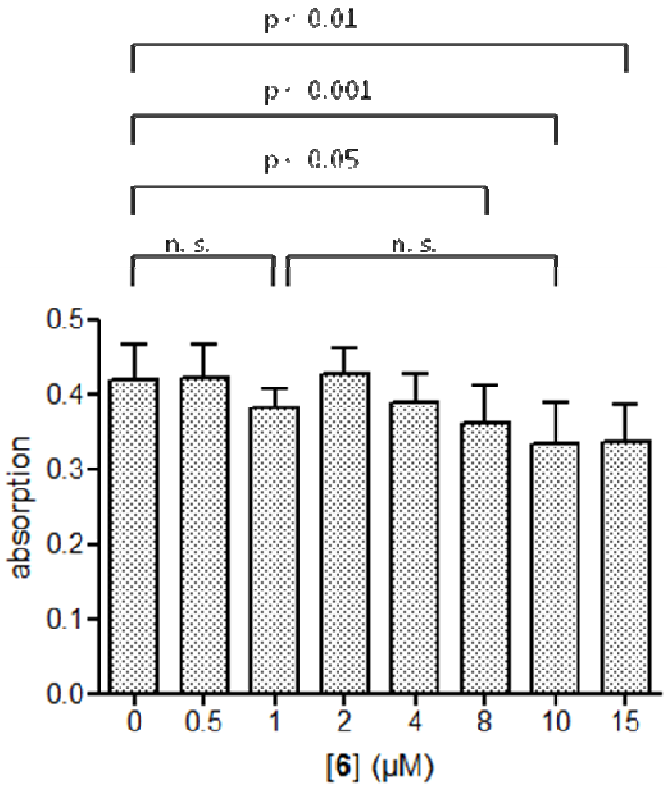


Fig. 9 Results of the MTT assay for different concentrations of probe **6** in HEK 293 cells recorded at 600 nm. Cells were incubated for 3 h with the compound (0 μM, 0.5 μM, 1 μM, 2 μM, 4 μM, 8 μM, 10 μM and 15 μM). Data of five independent experiments are presented as means ± SD. ANOVA revealed significant differences of the means between untreated cells and cells treated with 8, 10 and 15 μM of **6**. The p values were calculated by using post-hoc analysis (Tukey´s test; 'n.s.' means 'not significant'). Only statistically significant differences are indicated except for compound concentrations of 1 μM and 10 μM, which have been used in cell uptake experiments.

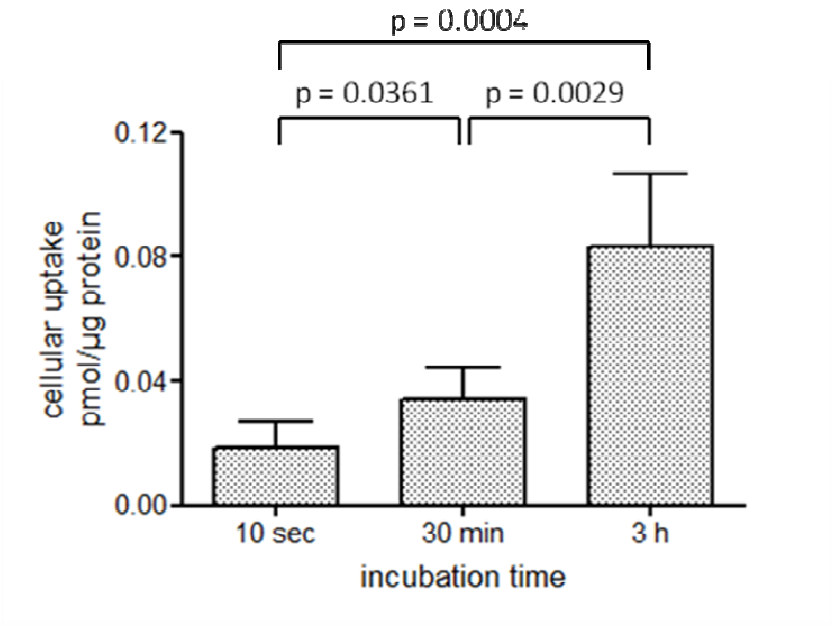


Fig 10 Results of the cell uptake experiments for probe **6** (10 µM). Data of five independent experiments are presented as means ± SD. HEK 293 cells were treated with **6** for 10 s, 30 min or 3 h. The data for 1 µM are not shown, because a significant uptake could not be ascertained.

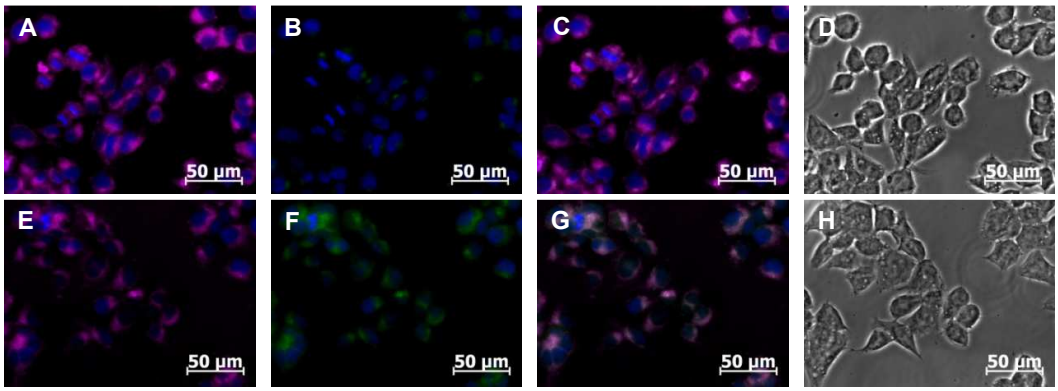


Fig 11 Uptake of probe **6**. HEK 293 cells were incubated with 10 µM of **6** for 10 s (A-D) or 3 h (E-H). A, E, LysoTracker (shown here in magenta) and DAPI (blue); B, F, probe **6** (green) and DAPI (blue); C, G, merged image of LysoTracker, probe **6** and DAPI; D, H, bright field.