Supporting Information

"Activity-Based Fingerprinting and Inhibitor Discovery of Cysteine Proteases in a Microarray"

Mahesh Uttamchandani,[†] Kai Liu,[†] Resmi C. Panicker,[§] Shao Q. Yao[†,§,*]

Departments of Biological Sciences[†] and Chemistry,[§] Medicinal Chemistry Program of the Office of Life Sciences,[*] National University of Singapore, 3 Science Drive 3, Singapore 117543, Republic of Singapore

1. General:

All chemicals were purchased from commercial sources and used directly, unless indicated otherwise. THF was distilled from sodium/benzophenone before use. All chemical reactions were run under N_2, unless otherwise indicated. The ^1H and ^{13}C NMR spectra were taken on a Bruker 300 MHz NMR spectrometer. Chemical shifts are reported in parts per million downfield from $SiMe_4$ (δ 0.0) and relative to the signal of chloroform-d (δ 7.26). LC was performed on a Waters Delta 600 HPLC equipped with a Phenomenex RP-18 (4.6 x 250 mm) column for analytical and Phenomenex RP-18 (21.2 x 250 mm) column for preparative, using an acetonitrile–water gradient (with 0.1% TFA). MS was run on an electrospray mass spectrometer (Finnigan, USA).

The 44 proteins used in Figure 2a (maintext) were purchased commercially from Sigma. They are (numbered according to Figure 2a): 2: Chymopapain (C-8526), 3: Papain (P-4762), 4: Bromelain (B-4882), 5: Ficin (F-6008), 6: α-chymotrypsin (C-4129), 7: β-chymotrypsin(C-4629), 8: α-chymotrypsinogen A (C-4879), 9: Pepsin (77152), 10: Proteainase K(P-8044), 11: Trypsin inhibitor (T-9003), 12: Trypsinogen (T-1143), 13: Thrombin (T-3399), 14: BSA bovine serum albumin (B-4287), 15: lysozyme (L-6876), 16: Acid phosphohydrolase (P-3627), 17: Acid phosphohydrolase (P-3752), 18: Alkaline phosphohydrolase (P-7640), 19: Alkaline phosphohydrolase (P-2265), 20: Alkaline phosphohydrolase (P-4002), 21: Lipase (L-1754), 22: Lipase (L-3001), 23: Lipase(L-9031), 24: Protease (P-6911), 25: Lipase(71548), 26: Thermolysin (T-7902), 27: Renin (83553), 28: Acylase (01821), 29: Penicillin amidase (76427), 30: Esterase (46069), 31: Epoxide hydrolase (71832), 32: Lipoprotein lipase (62336), 33: Acetylcholine esterase (01023), 34: Actinase E (P-8811), 35: Epoxide hydrolase (45299), 36: Alcoholdehydrogenase (05460), 37: Urease (U-1500), 38: Amylose (A-0512), 39: Concannavalin A (L-7647), 40: Achromopeptidase(A-3547), 41: Alcoholdehydrogenase (A-7011), 42: Collagenase (C-7657), 43: Glutathione S-transferase (G-4385), 44: Phosphoenolpyruvate carboxylase (79414), 45: Pyrophosphatase (83205).

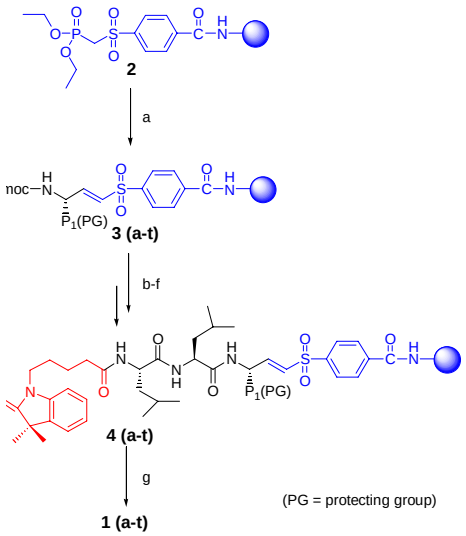
Stock solutions of above proteins were prepared as 5~10 mg/ml solutions in double distilled water, desalted with a NAP-5 column (Amersham, USA) and stored at -20 ^0C until use.

2. Chemical synthesis

Detailed synthesis of the resin-bound phosphonate and its precursors, as well as the 20 Fmoc-AA-CHO, was reported previously.[1] Briefly as shown in Scheme S1, starting from the resin-bound phosphonate **2**, a solid-phase Horner-Emmons reaction was carried out smoothly with 20 suitably protected N-Fmoc-α-amino aldehyde using previously optimized conditions, generating exclusively the trans-vinyl sulfones **3 (a-t)** with minimal epimerization at the chiral center. Subsequently, standard solid-phase peptide synthesis protocols with Fmoc chemistry were used to attach a Leu-Leu linker, followed by the Cy_3 dye. The resulting 20 resin-bound probes, **4 (a-t)**, were cleaved off the resin by TFA treatments and subjected to ether precipitation.

Representative Synthesis of **3 (a-t)**. To 200 mg (0.7 mmol/g, 1 mmol) of phosphonate resin **2** was added 5 ml of anhydrous THF. After swelling for 1 h, 2 ml of THF was removed by suction using a syringe. To the remaining solution was added 1 M LHMDS solution (in THF; 0.7 ml, 0.7 mmol, 5 eq) and the reaction was stirred gently for 30 min. The reaction was stopped by removing the solution by suction. About 15 ml of anhydrous THF was added and the mixture further stirred for 5 min. The solution was again removed by suction as described before. This process was repeated once. After the final THF wash step, Fmoc-protected amino aldehyde (0.42 mmol, 3 eq; dissolved in 1 ml of THF) was added to the resin, and the resulting solution was topped up to ~ 5 ml with THF. The reaction was stirred for another 2 h, and the resin was filtered and washed with DMF (3 x), DCM (3 x), DMF (3 x) before drying *in vacuo*.

Representative Synthesis of Probes **1 (a-t)**. To 200 mg of resin **3** was added 3 ml of DBU/HOBT/DMF (2:1:97) solution and the reaction was allowed to proceed for 10 min to cleave off the Fmoc group. The resulting resin was filtered and washed extensively with DMF (3 x), DCM (3 x) and DMD (3 x). A solution containing Fmoc-Leu-COOH (141 mg, 4 eq), HOBT (61.2 mg, 4 eq), DIC (63 μL, 4 eq) and DIEA (137 μL, 8 eq) in 2-3 ml of DMF was then added and the reaction was allowed to proceed for 3-4 h. The Fmoc-deprotection, coupling steps were repeated once more to attach the next amino acid and finally Cy3 dye, before washing with DMF (3 x), DCM (3 x), MeOH (3 x) and dried *in vacuo* to give **4 (a-t)**. The probe was cleaved off the resin with 1.5 ml of TFA/TIS/H_2O (95/2.5/2.5) mixture for 2 h. The filtrate was evaporated *in vacuo* until TFA was removed completely before subjecting to preparative HPLC purification to give the final pure probe **1 (a-t)**. All the probes **1 (a-t)** were characterized by ESI-MS and summarized in **Table S1**. The stock solutions of the 20 probes (200 μM in DMSO) were prepared and stored at -20 ^{0}C until use.



Scheme S1. Solid-phase synthesis of peptide vinyl sulfones. Reagents and Conditions: (a) LHMDS (5 equiv; 1.0 M in THF), 30 min with resin, washed with THF, then AA-CHO (4 equiv) in THF (20 mL/g resin), 2 h at RT; (b) DBU/HOBt/DMF (2:1:97, 10 mL/g resin), 10 min at RT; (c) Fmoc-Leu-OH (4 equiv), DIC (4 equiv), HOBt (4 equiv), DIEA (8 equiv), DMF, 6 h at RT; (d) repeat (b) & (c); (e) repeat (b); (f) Cy3 (4 equiv), DIC (4 equiv), HOBt (4 equiv), DIEA (8 equiv), DMF, 6 h at RT; (g) TFA/DCM/H2O (95:2.5:2.5), 10 mL/g resin, 2-3 h at RT.

Table S1 ESI-MS data for the 20 vinyl sulfone probes with P1 variation

No	Compound	$(M-1)^+$ calc.	Found
1a	Cy3-LLA-VS	905.5	905.5
1b	Cy3-LLD-VS	949.5	949.5
1c	Cy3-LLE-VS	963.5	963.5
1d	Cy3-LLF-VS	981.5	981.5
1e	Cy3-LL-*d*F-VS	981.5	981.5
1f	Cy3-LLG-VS	891.5	891.5
1g	Cy3-LLH-VS	971.5	971.5
1h	Cy3-LLI-VS	947.5	947.5
1i	Cy3-LLK-VS	962.5	962.6
1j	Cy3-LLL-VS	947.5	947.5
1k	Cy3-LLN-VS	948.5	948.4
1l	Cy3-LL-Orn-VS	948.5	948.5
1m	Cy3-LLP-VS	931.5	931.5
1n	Cy3-LLQ-VS	962.5	962.5
1o	Cy3-LLR-VS	990.6	990.4
1p	Cy3-LLS-VS	921.5	921.4
1q	Cy3-LLT-VS	935.5	935.5
1r	Cy3-LLV-VS	933.5	933.6
1s	Cy3-LLW-VS	1020.5	1020.5
1t	Cy3-LLY-VS	997.5	997.6

3. General enzyme fingerprinting in a gel format.

2 µL of the protein solution (0.5 mg/ml) was diluted with 18 µL of Reaction Buffer (50 mM CH_3COONa at pH 5, 2 mM DTT, 1 mM EDTA). 0.2 µL of the probe (200 µM) was then added and the reaction mix was incubated at room temperature in the dark for 15 min. The reaction was quenched by boiling the sample at 95℃ for 5 min with 4 µL of 6x SDS-PAGE loading dye. The sample was then resolved on a 12% SDS-PAGE gel and visualized by fluorescence scanning using the Typhoon 9200 scanner (Amersham, USA). The fluorescent bands, corresponding to the labeled protease, were quantified with the ImageQuant software (Amersham, USA) and analyzed by the program Treeview.

4. General labeling experiment in a protein microarray.

The NHS slides were prepared from plain glass slides (Sigma, USA) as previously described.[2] The enzymes were prepared in Na_2CO_3 Buffer (50mM, pH 9) to a final concentration ~ 0.5mg/mL. The samples were spotted in triplicate (or duplicate) using an ESI SMA™ arrayer (Ontario Canada) with a spacing of 300 μm between the spots. The spacing between subarrays was ~3000 μm. After 1-hour incubation in humid conditions, the slides were soaked in the Quench Buffer (0.1 M Na_2CO_3, pH 9 with 0.5 M Glycine) with gentle shaking for 10 min. The slides were subsequently blocked with Blocking Buffer (50 mM CH_3CO_2Na, pH 5 containing 1% w/v BSA) for 10 min, followed by washing with distilled water and dried by centrifugation (200 g, 8 min).

To optimize labeling time, a freshly prepared labeling mix (2 μM probe **1a**, 2 mM DTT, 1 mM EDTA, 50 mM CH_3CO_2Na at pH 5 and 1% w/v BSA) was pipetted to the protein subarray at appropriate time intervals, 2 μL for each subarray. After incubation, the whole slide was subsequently sonicated in hot 1% SDS buffer for 30 min, washed with distilled water dried by centrifugation (200 g, 8 min) and scanned using an ArrayWorx™ microarray scanner (Applied precision, USA) under 548/595 nm channel (Figure S1).[3]

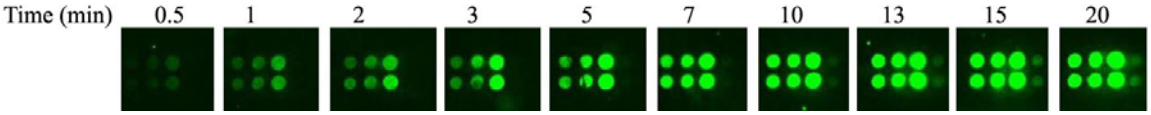


Figure S1. Time-dependent optimization of labeling reactions on the protein microarray. (left to right) in each subarray: Chymopapain, Bromelain, Papain, Ficin.

To optimizing labeling pH, a labeling mix (2 μM probe **1a**, 2 mM DTT, 1 mM EDTA, 50 mM CH_3CO_2Na at pH 5 and 1% w/v BSA; at different pH) was prepared. 2 μL of the mix was subsequently pipetted to each subarray. After 15 min, the whole slide was sonicated in hot 1% SDS buffer for 30 min. The slide was then washed with distilled water, dried by centrifugation (200 g, 8 min) and scanned (Figure S2)

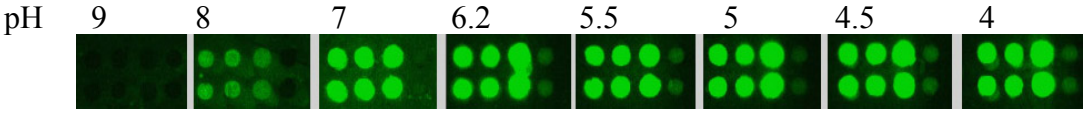


Figure S2. pH-dependent optimization of labeling reactions on the protein microarray. (left to right) in each subarray: Chymopapain, Bromelain, Papain, Ficin.

To prepare a "universal" master mix of labeling solution, equal amounts of the 20 vinyl sulfone probes, **1 (a-t)**, were mixed with an appropriate buffer to make the final "master mix" labeling solution (2 μM of mixed probes, 50 mM CH_3COONa at pH 5, 2 mM DTT, 1 mM EDTA and 1% w/v BSA). 5 μL of this solution was used in each screening.

To ensure the subarray approach generates highly accurate and reproducible results, a protein microarray was generated in which 48 identical subarrays of proteins were spotted (Figure S3). 2 μL of a freshly prepared labeling buffer (2 μM of probe **1a**, 2 mM DTT, 1 mM EDTA, 50 mM CH_3CO_2Na at pH 5, 1% w/v BSA) was pipetted to each of the 48 subarrays. After 15 min of incubation, the slides were sonicated in hot 1% SDS buffer for 30 min, then washed with distilled water, dried by centrifugation (200 g, 8 min) and scanned. Fluorescence signals of spots (from the same proteins across 48 subarrays) were quantified and the typical standard deviation between spots was less than 10%.

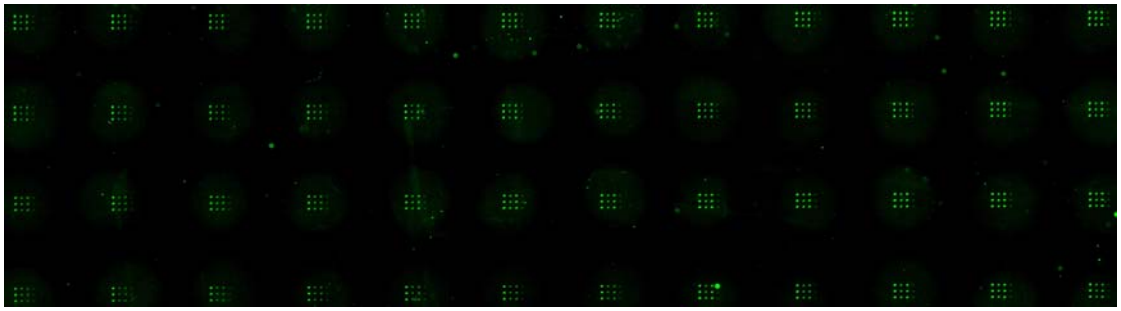


Figure S3. Image of 48 subarrays treated by 2 μM of probe **1a**.

To ensure that our microarray-based, enzyme-detecting method is indeed activity-based, each subarray on the same glass slide was pre-incubated with 2 μL of Normal Reaction Buffer (2 mM DTT, 1 mM EDTA, 50 mM CH_3COONa at pH 5 and 1% w/v BSA), Inhibition Buffer 1 (2 mM DTT, 1 mM EDTA, 50 mM CH_3COONa at pH 5, 1% w/v BSA and 1%SDS), Inhibition Buffer 2 (2 mM DTT, 1 mM EDTA, 50 mM CH_3COONa at pH 5, 1% w/v BSA, and 0.1 mM Leupeptin), or Normal Buffer w/o DTT (1 mM EDTA, 50 mM CH_3COONa at pH 5, and 1% w/v BSA). After 15 min of incubation, 0.5 μL of Labeling buffer (10 μM of probe **1a**, 1 mM EDTA, 50 mM CH_3COONa at pH 5 and 1% w/v BSA) was pipetted onto each subarray. 15 min later, the slide was sonicated in hot 1% SDS buffer for 30 min, washed with distilled water, dried by centrifugation (200 g, 8 min) and scanned.

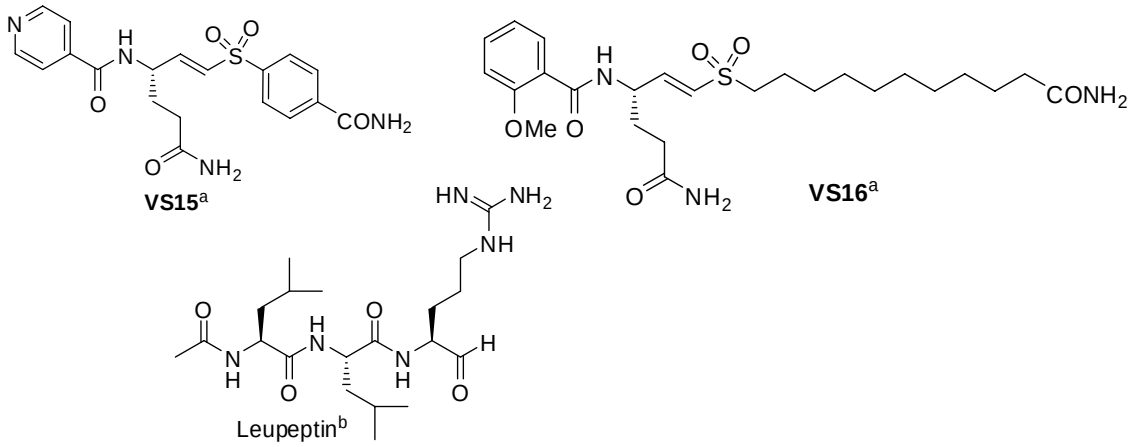
5. Activity-based fingerprinting in a protein microarray.

To generate the activity-based fingerprinting profiles of the 4 cysteine proteases, a freshly prepared labeling buffer contain a different probe (2 μM each of the 20 different probes, **1 (a-t)**, 2 mM DTT, 1 mM EDTA, 50 mM CH_3COONa at pH 5, 1% w/v BSA) was pipetted to the protein microarray, 2 μL for each subarray. After 15 min of incubation, the whole slide was sonicated in hot 1% SDS buffer for 30 min, washed with distilled water, dried by centrifugation (200 g, 8 min) and scanned. The fluorescence signal of each spot was quantified by ArrayWorX^{TM} software and analyzed by the program Treeview.

6. Inhibitor fingerprinting in a protein microarray

2 μL of an inhibitor buffer (50 mM CH_3COONa at pH 5, 2 mM DTT, 1 mM EDTA, 1% w/v BSA, and 3% v/v DMSO) containing a gradient concentration of the inhibitor (from 0.01 μM to 1 mM) was pipetted on each subarray. All inhibition subarrays were prepared on the same glass slide to ensure accuracy and reproducibility. After 1 h incubation, 0.5 μL of the labeling buffer (2 μM of probe **1a**, 2 mM DTT, 1 mM EDTA, 1% w/v BSA, 50 mM CH_3COONa at pH 5) was introduced on each subarray. After incubation at room temperature in the dark for another 15 min, the whole slide was sonicated in hot 1% SDS buffer for 30 min, washed with distilled water, dried by centrifugation (200 g, 8 min) and scanned. The fluorescence signal of each spot (Figure S4) was quantified. The IC_{50} of the inhibitor was calculated by fitting the resulting dose-dependent data with the PRISM software (GraphPad, USA).

Table S2. Small molecule inhibitors used in the study



[a] **VS15** & **VS16** were synthesized as previously described.[1]
[b] Leupeptin is commercially available.

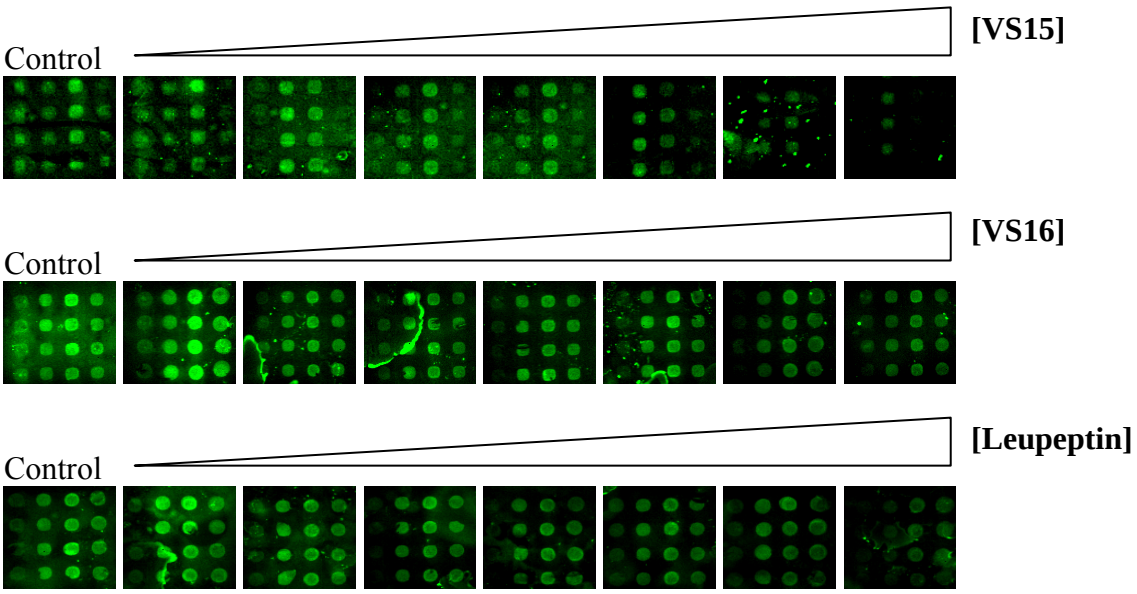
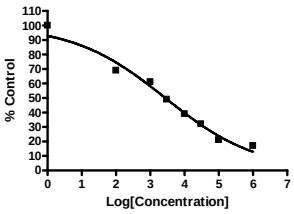
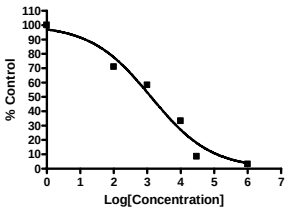
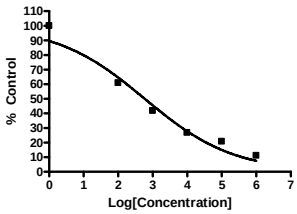
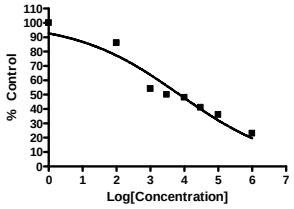
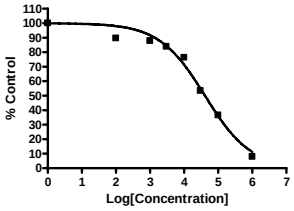
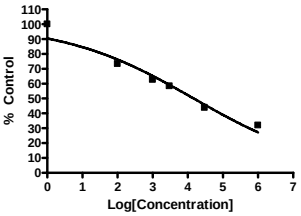
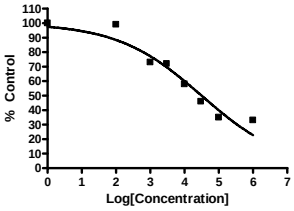
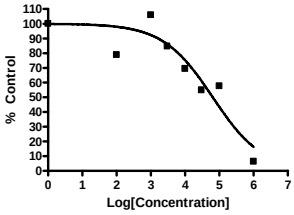
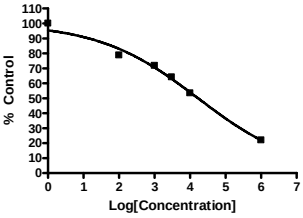
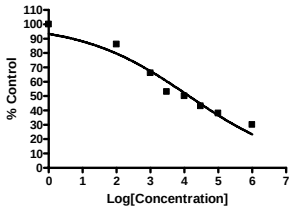
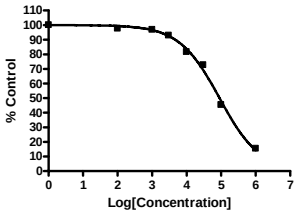
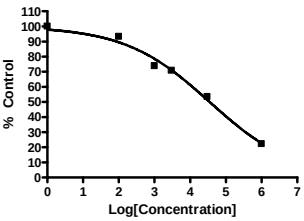


Figure S4. (a) Inhibitor fingerprinting of 4 cysteine proteases in a protein microarray. In each subarray, the 4 proteins (left to right: Chymopapin, Bromelain, Papain, Ficin) were spotted in quadruplicate (top to bottom). A different concentration of each inhibitor **VS15, VS16** and **Leupeptin** (left to right: 0, 0.1, 1, 3, 10, 30, 100, 1000 μM) were incubated with each subarray for 1 h, then screened with probe **1a**.

Table S3. IC_{50} graphs of selected inhibitors obtained by protein microarray screening

	VS15	VS16	Leupeptin
Chymopapain	IC_{50} = 2.7 μM 	IC_{50} = 1.3 μM 	IC_{50} = 0.59 μM 
Papain	IC_{50} = 7.1 μM 	IC_{50} = 40.6 μM 	IC_{50} = 15 μM 
Bromelain	IC_{50} = 31 μM 	IC_{50} = 64 μM 	IC_{50} = 17 μM 
Ficin	IC_{50} = 14 μM 	IC_{50} = 90 μM 	IC_{50} = 35 μM 

7. Inhibitor validation using conventional microplate experiments.

Above identified inhibitors were validated by the EnzChek Protease Assay™ (Molecular Probes, USA). The kit employs an intramolecularly quenched bodipy-casein conjugate to generate fluorescence signal upon proteolytic cleavage. Firstly, 0.15 μg of each enzyme, corresponding to around 0.1 μM in a 50 μL reaction buffer (2 mM DTT, 1 mM EDTA, 50 mM CH_3COONa at pH 5, and 3% v/v DMSO) was incubated with gradient concentrations of a desired inhibitor (from 0.01 μM to 300 μM) in a 384-well black plate for 1 h. Then 1 μL of a 0.5 mg/ml bodipy casein solution (provided in the kit) was introduced into each well. The fluorescence signal of each well was recorded over 15 min using a SpectraMAX™ Gemini XS fluorescence plate reader (Molecular Devices, USA). The IC_{50} of the inhibitor was calculated by fitting the resulting dose-dependent data with the PRISM software (GraphPad, USA). Results showed they correlated well with those obtained from our inhibitor fingerprinting experiments carried out on the protein microarray.

References

1. (a) G. Wang, M. Uttamchandani, G. Y. J. Chen and S. Q. Yao, *Org. Lett*. 2003, **5**, 737-740. (b) G. Wang, and S. Q. Yao, *Org. Lett*. 2003, **5**, 4437-4440.
2. M. L. Lesaicherre, M. Uttamchandani, G. Y. J. Chen and S. Q. Yao, *Bioorg. Med. Chem. Lett.* 2002, **12**, 2079-2083.
3. For probe labeling protocol on a microarray format see: G. Y. J. Chen, M. Uttamchandani, Q. Zhu, G. Wang and S. Q. Yao, *ChemBioChem* 2003, **4**, 336-339.