JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

A Peptide Aldehyde Microarray for High-Throughput Profiling of Cellular Events

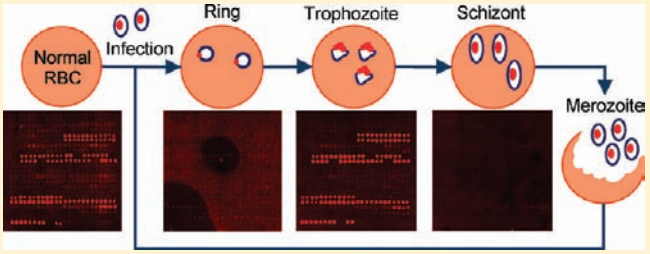
Hao Wu,[†] Jingyan Ge,[†] Peng-Yu Yang,[†,§] Jigang Wang,[†,§] Mahesh Uttamchandani,[†,‡,¶] and Shao Q. Yao*[,†,‡,§]

[†]Department of Chemistry, [‡]Department of Biological Sciences, and [§]NUS MedChem Program of the Life Sciences Institute, 3 Science Drive 3, National University of Singapore, Singapore 117543

[¶]Defence Medical and Environmental Research Institute, DSO National Laboratories, 27 Medical Drive, Singapore 117510

Supporting Information

ABSTRACT: Microarrays provide exciting opportunities in the field of large-scale proteomics. With the aim to elucidate enzymatic activity and profiles within native biological samples, we developed a microarray comprising a focused positional-scanning library of enzyme inhibitors. The library was diversified across P_1-P_4 positions, creating 270 different inhibitor sublibraries which were immobilized onto avidin slides. The peptide aldehyde-based small-molecule microarray (SMM) specifically targeted cysteine proteases, thereby enabling large-scale functional assessment of this subgroup of proteases, within fluorescently labeled samples, including pure proteins, cellular lysates, and infected samples. The arrays were shown to elicit binding fingerprints consistent with those of model proteins, specifically caspases and purified cysteine proteases from parasites (rhodesein and cruzain). When tested against lysates from apoptotic Hela and red blood cells infected with *Plasmodium falciparum*, clear signatures were obtained that were readily attributable to the activity of constituent proteases within these samples. Characteristic binding profiles were further able to distinguish various stages of the parasite infection in erythrocyte lysates. By converting one of our brightest microarray hits into a probe, putative protein markers were identified and pulled down from within apoptotic Hela lysates, demonstrating the potential of target validation and discovery. Taken together, these results demonstrate the utility of targeted SMMs in dissecting cellular biology in complex proteomic samples.



INTRODUCTION

"Omic" methodologies have, in recent years, enabled comprehensive differentiation of complex biological samples and cell states.[1] DNA microarrays, for instance, measure quantitative changes in overall mRNA levels.[2] Mass spectrometry provides sensitive ways in which to elucidate differences in protein and metabolite abundance.[3] These platforms, however, are greatly challenged when measuring functional differences, for example, in detecting the changing activity levels of enzymes *en masse*.[4] Activity-based protein profiling (ABPP), pioneered by Cravatt, has made an attempt to bridge this gap by making use of activity-based small-molecule probes to delineate enzymatic activities.[5] This has, however, been restricted to mostly a single probe (in some cases several related probes), covering only a narrow pool of enzymes that can be screened at a time. Small-molecule microarrays (SMMs) are, in contrast, a highly scalable screening platform. They have predominantly been applied over the past decade for the purpose of screening pure proteins, in a one-to-many format, for ligand identification against known protein targets.[6] We conceived that screening using SMM could be extended to comparative functional assessment of complex proteomes in high throughput, something not easily performed with competing techniques.[2−5] Our goal was thus to apply the SMM platform much more broadly, in a many proteins-to-many ligands format. By taking advantage of the key characteristics of microarray-based technologies (miniaturization, high density, and high throughput) which enable simultaneous measurements of many interactions of component proteins across different cell states, we hoped to dissect complex proteomic samples (like cellular lysates) by rapidly examining their global activity profiles.

Enzymes are involved in every cellular process. Their functional state is tightly regulated across transcriptional and translational levels. SMMs of enzyme ligands (such as substrates or inhibitors) may therefore be used as tools to accurately reflect the activity state. This was demonstrated in previous work with pure proteins and enzymes.[7−9] Several other groups have applied similar techniques in fingerprinting proteins on the basis of their affinity profiles on the microarrays.[10] In order to discern different cell states and to gain insight into the underlying molecular events, we designed our SMMs to target cysteine proteases, which are an important class of enzymes that play vital roles in numerous physiological processes.[11,12] Notably, the overexpression of human cysteine cathepsins and caspases has been implicated in a multitude of pathological conditions, including Alzheimer's disease and cancer.[11] Many cysteine proteases are

Received: October 25, 2010
Published: January 19, 2011

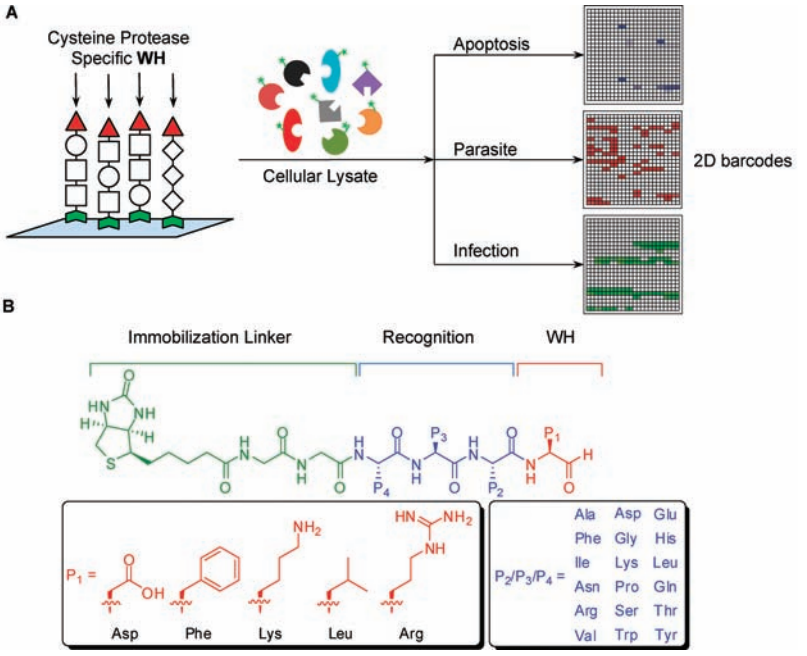


Figure 1. (A) Overall strategy of the SMM platform for comparative profiling of biological events. Unique enzyme/inhibitor interaction profiles were generated upon screening prelabeled cellular lysates with the SMM. (B) Design of the peptide aldehyde SMM. The PSL was designed using five individual P_1 aldehyde warheads (in red), targeting different subclasses of cysteine proteases, and diversity elements from P_2 to P_4 positions (in blue). P_2, P_3, and P_4 were either individual amino acids for mapped positions or an isokinetic mixture of 18 amino acids (which excluded cysteine and methionine).

also essential in the lifecycles of pathogenic protozoa.[12] For example, brucipain (also called rhodesain) is the major cysteine protease in *Trypanosoma brucei*, a parasite that causes African sleeping sickness. Cruzain is another parasitic cysteine protease, in *Trypanosoma cruzi*, which causes Chagas disease. Falcipains are cysteine proteases from *Plasmodium falciparum*, the parasite that causes malaria and infects 300−500 million people annually.[13] We attempt here to monitor and profile the entire subset of functionally active cysteine proteases from different biological samples by using SMMs of tight-binding enzyme inhibitors to detect the corresponding enzyme/inhibitor interactions (Figure 1). Signals on the microarrays would be diminished in the presence of inhibitory cues, down-regulation, or inactivity of any of the enzymes in question.

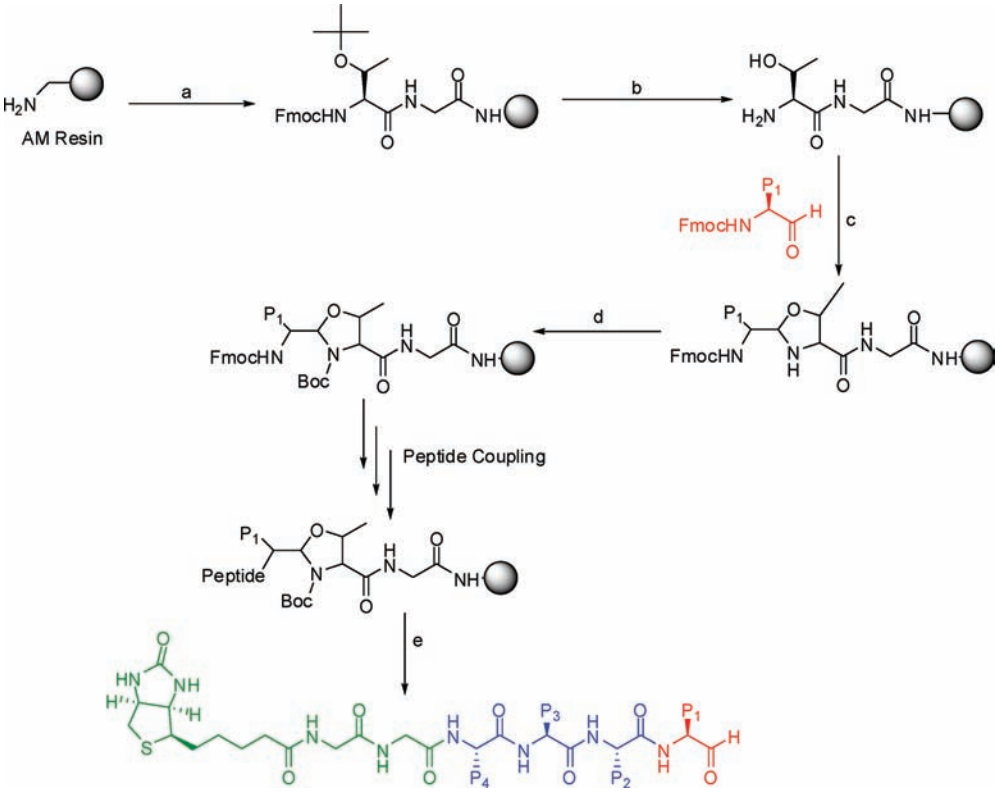
RESULTS AND DISCUSSION

Synthesis of Peptide Aldehydes and Construction of the Corrresponding SMM. Peptide substrate and inhibitor libraries have become an established tool for dissecting the activity of proteases.[14−16] As shown in Figure 1, our SMM comprises a positional-scanning library (PSL)[17] of peptide aldehydes, which are reversible inhibitors of cysteine proteases.[11] They were favored over other known inhibitors, including those that contain reversible (i.e., azidomethylenes and azanitriles) and irreversible (i.e., epoxyl derivatives, peptidyl Michael acceptors, (acyloxy)methyl ketones, and halomethyl ketones) warheads, because they can be synthesized conveniently and are known to bind tightly to active cysteine proteases. The PSL design covers a diverse array of amino acids strategically positioned across P_1, P_2, P_3, and P_4 positions, thus facilitating comprehensive and specific interactions with most known cysteine proteases.[11−13] First, five different amino aldehyde warheads (WH) were synthesized in solution phase from the corresponding Fmoc-protected amino acids (Asp, Phe, Lys, Leu, and Arg) for the P_1 position (see Supporting Information).[18,19] These residues were chosen in order to provide sufficient coverage of P_1 substrate specificity for most cysteine proteases documented in the MEROPS database.[20] These aldehydes were then loaded onto a threonine-functionalized resin by acid-catalyzed oxazolidine formation (Scheme 1). Solid-phase combinatorial library synthesis was then carried out, as previously described,[21] by using standard solid-phase peptide synthesis with Fmoc chemistry. Eighteen different amino acids (excluding methionine and cysteine) were used to permute the P_2 to P_4 positions, generating 270 different sublibraries. In addition, five positive control inhibitors were synthesized, namely DEVD, DEVF, DEVK, DEVL, and DEVR, all of which also contained an aldehyde warhead. A hydrophilic linker and biotin were introduced to the *N*-terminus of each of the 275 peptide aldehydes, which were subsequently immobilized onto avidin-coated glass slides, to generate the corresponding peptide aldehyde SMM.

Inhibitor Fingerprinting with Pure Enzymes on SMMs. We first evaluated our SMM by screening the platform against four recombinant cysteine proteases (each was fluorescently labeled with an amino-reactive Cy5 dye). These included two that belong to clan CA (cruzain from *T. cruzi* and rhodesain from *T. brucei*) and another two from clan CD (caspase-3 and caspase-7 involved in apoptosis). Clan CA proteases, including the papain family (e.g., cathepsins and calpains), have a relatively broader tolerance of amino acid side chains at the P_1 position. This is unlike the CD clan, which exhibits a much stronger specificity at this position. Unique binding fingerprints (or 2D barcodes) were observed for each enzyme on the microarrays (Figure 2A). As expected, Asp consistently appeared in the P_1 position for both caspases (Figure S3 in the Supporting Information), while various different amino acids (including Phe, Lys, and Arg) were preferred by cruzain and rhodesain at this position. Across P_2 to P_4 positions,

dx.doi.org/10.1021/ja109597v |*J. Am. Chem. Soc.* 2011, 133, 1946–1954

Scheme 1. Synthesis of Peptide Aldehyde PSL[a]



[a] Reagents and conditions: (a) i, Fmoc-Gly-OH, HBTU, HOBt, DIEA, DMF, 4 h; ii, 20% piperidine/DMF, 30 min; iii, Fmoc-Thr(O^{t}Bu)-OH, HBTU, HOBt, DIEA, DMF, 4 h; iv, Ac_2O, DIEA, DCM, 2 h. (b) i, 20% piperidine/DMF, 30 min; ii, TFA/DCM (1:1) 1 h; iii, 10% DIEA/DCM. (c) 1% DIEA/MeOH, 2 h, 60 °C. (d) Boc_2O, 1% DIEA/DCM, 2 h. (e) i, TFA/TIS/EDT (95:5:0.1); ii, ACN:H_2O:TFA (60:40:0.1), 30 min, 60 °C.

it was noted that different side chains contributed at varying levels to the overall binding potency of these inhibitors. Caspases-3 and -7 share extremely high homology in both protein sequence and substrate specificity. As a result, their SMM binding fingerprints were almost identical (see top two panels in Figure 2A). Overall binding profiles of the four enzymes were also displayed as position-specific scoring matrixes (Figure 2B) and position-specific bar graphs (Figure S3). K_D measurements were performed by dose-dependent application of proteins on the microarrays. These selectivity profiles for the various inhibitors were further validated using IC_{50} measurements with a solution-phase microplate enzyme assay (Figure 2C). Results with our PSL were in general consistent across microarray and microplate assays, as well as with published reports on the substrate specificity of these enzymes (Table 1). The most potent inhibitor for caspase-3 was DEVD-CHO, as expected, as this inhibitor has been well-reported. This inhibitor was the most potent across both microplate and microarray measurements, with an IC_{50} of 17 nM and an apparent microarray of K_D = 172 nM. When the enzymes were denatured by heat and rendered inactive, no binding was observed on the arrays (data not shown). Taken together, these results correlate very well with established findings against these cysteine proteases.[9,11,15c,22] Our SMM platform hence provides a highly miniaturized, rapid tool to establish the specificity of cysteine proteases.

Profiling Mammalian Cell Lysates on SMMs. Having established that the platform provided reproducible fingerprints with pure enzymes, we went on to establish its utility in differentiating cell states. One model pathway we identified was apoptosis, which is an important physiological process in cells and causes cancer when deregulated.[11b] The apoptosis cascade is mediated by cysteine proteases, namely caspases, and we hence sought to profile the changing functional state of these proteins using our SMM (Figure 3A). Human cervical cancer (Hela) cell lines were incubated with an apoptotic trigger, staurosporine (STS).[23] The presence of active caspase-3 in the resulting apoptotic cells was unambiguously confirmed by Western blotting and a caspase enzymatic assay (Figure 3B). Cy5-labeled crude apoptotic and/or non-apoptotic cellular lysates were then applied on the microarrays, producing characteristic binding profiles (Figure 3A and Supporting Information). Apoptotic Hela (ApHela) lysates displayed distinct binding profiles with the peptide aldehyde SMM. Interestingly, it seemed that the binding patterns obtained were highly similar to that of caspase-3/-7 shown in Figure 2. To ensure that these results were indeed attributable to caspase-3/-7, tyramide signal amplification (TSA) was performed on the ApHela-treated arrays with anti-caspase-3 antibodies (Figure 3A, right panel). The Venn diagram shown in Figure 3C clearly displayed that the majority (20 out of 26) of strong binders detected by TSA assay overlapped with those from pure caspase-3. Notably, 12 out of 14 strong binders of the apoptotic Hela samples overlapped with caspase-3, indicating that the profile observed on the SMMs was a result of triggered caspase activation. We further performed pull-down experiments against one of the strongest hits (Biotin-GG-DxxD-CHO) identified on the microarray, and results showed that caspase-3 could indeed be enriched (see Supporting Information). In a separate experiment, it was found that our SMM could be used to generate calcium-dependent binding profiles of Hela cells overexpressing activated calpains (a class of clan

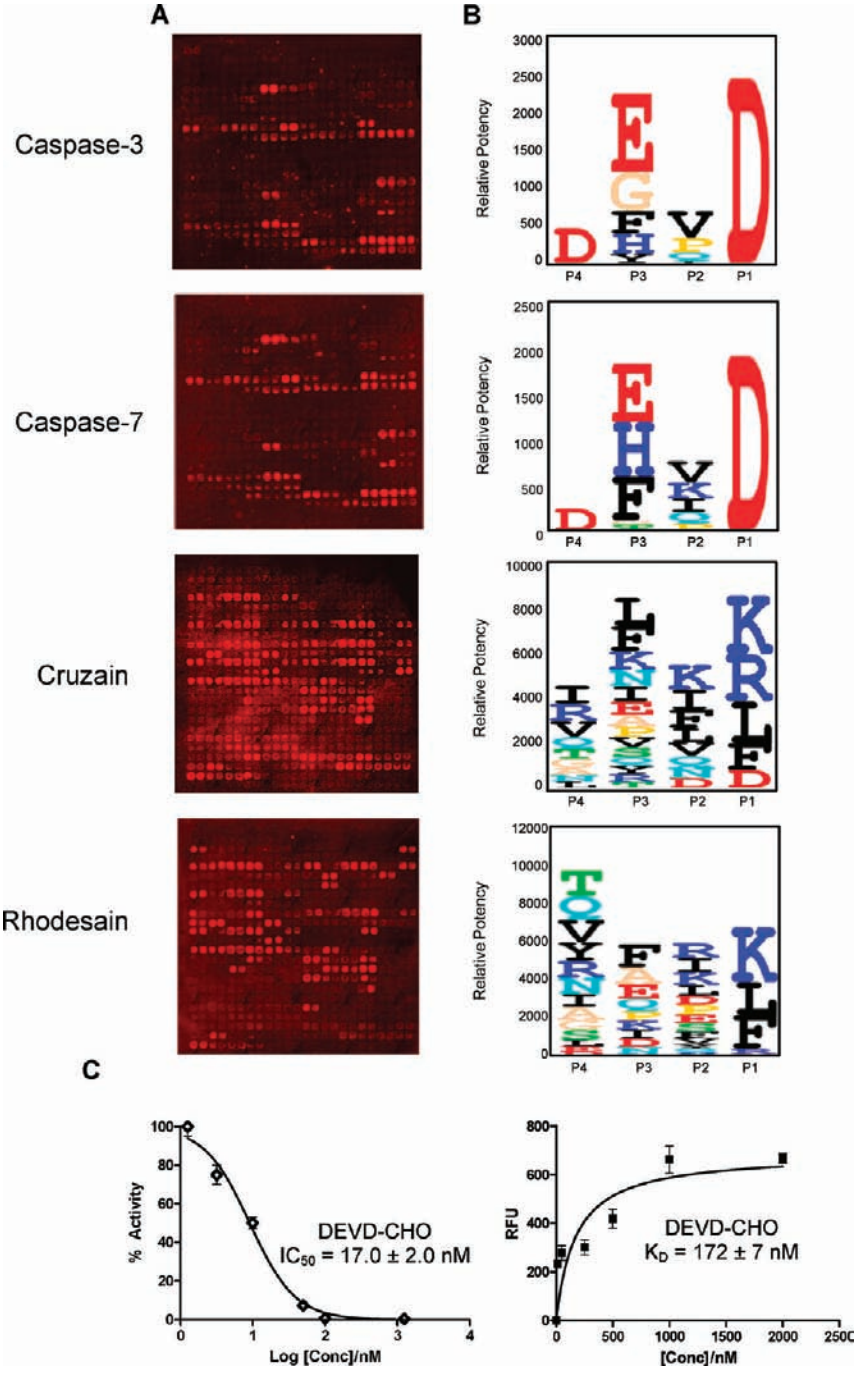


Figure 2. Microarray profiles of the peptide aldehyde SMM with four different recombinant cysteine proteases. All compounds were spotted in duplicate. See Supporting Information for spotting patterns/IDs. (A) Microarray binding profiles of Cy5-labeled cysteine proteases. (B) Position-specific scoring matrix representing binding affinity across P_1 to P_4 positions for each amino acid, in single-letter code. The height of each letter represents the weighted contribution of that residue to overall binding. The side chains are colored according to their properties: hydrophobic/aromatic (black), acidic (red), basic (blue), polar (cyan), hydrophilic (green), and small (beige). (C) A representative example of the quantitative IC_{50} and K_D results against pure caspase-3. Biotin-GG-DEVD-CHO was used.

A calcium-dependent, non-lysosomal cysteine proteases; Figure S7b in Supporting Information), which were readily distinguishable from those of STS-induced apoptotic cells.

Parasite Lysates and Infected Erythrocytes Screened on SMMs. To further explore more interesting cellular events, two parasites were selected for studies on the peptide aldehyde SMM (Figure 4). One was *T. brucei*, the etiologic agent of human African sleeping sickness.[12,13] The parasite (bloodstream form, BSF) was lysed and labeled with Cy5. Many positive spots appeared on the microarrays, the majority of which (41 out of 62) were attributable to rhodesain (Figure 4A, left panel). The activity of enzymes in this parasite may thus be potentially tracked using our SMM. The other fluorescent spots identified may be a result other protein binding, the nature of which has yet to be identified. The other parasite we investigated was *P. falciparum*. The ability to track malaria infection and develop new drugs

dx.doi.org/10.1021/ja109597v |*J. Am. Chem. Soc.* 2011, 133, 1946–1954

Table 1. Inhibitor Specificity of Cysteine Proteases Present in This Study[a]

Enzyme	Position				Reference
	P_4	P_3	P_2	P_1	
Caspase-3	D, E, S	E, L, S	V, T, L	D	[11b, 20]
	D, E, T	E, F, A	V, I, L	D	Present Study
	D	E, G, F	V, P, Q	D	Present Study
Caspase-7	D, E, I	E, A	V, T, P	D	[11b, 20]
	D, E, S	E, D, A	V, T, I	D	Present Study
	D	E, H, F	V, K, I	D	Present Study
Cruzain	Broad	Broad	F, W, L	K, R, Q	[15c, 20]
	K, R, L	K, L, R	F, L, Y	F, K, R	Present Study
	Broad	Broad	N, A, E	K, L, F	Present Study
Rhodesain	Broad	K, R, P	F, W L	K, R	[15c, 20]
	Broad	L, R, I	F, Y, L	F, K, R	Present Study
	Broad	Broad	K, I, F	K, R, L	Present Study

[a] The most strongly contributing positions from the averaged profiles are shown for the present study in blue (microplate inhibition assay) and red (microarray inhibitory binding). These positions are displayed in bold if also shown to be preferred consistently with published substrate preferences.[11b,15c,20]

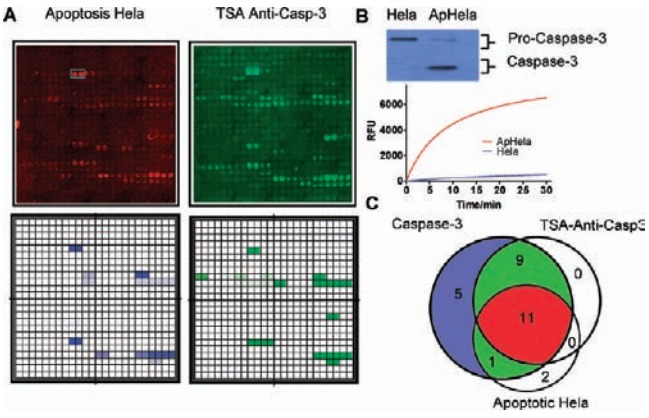


Figure 3. Microarray profiles of apoptotic events. (A) Top panels, microarray binding profiles of Cy5-labeled apoptotic Hela lysate, and TSA assay using anti-caspase-3 antibody; lower panels, corresponding 2D barcodes illustrating the binding profiles from the microarray images. The spots corresponding to DEVD-CHO, a known caspase-3 binding sequence, are boxed in the top panels. (B) Caspase-3 activation of Hela lysate monitored by Western blot and activity assay. Activated caspase-3 appears as a 17 kDa band, while pro-caspase-3 is larger at 35 kDa (Hela). A significant increase in caspase-3 activity was shown after its activation by STS (curve shown in red, monitored by Ac-DEVD-AFC cleavage). (C) Venn diagram illustrating distribution of the top-binders (intensity greater than 500 RFUs). The overlapping regions shown in red indicate 11 common potent binders across the three microarrays (including those of pure caspase-3, obtained from Figure 2A top panel). Regions in white were generally unpopulated, indicating the Hela-binding profiles arose predominantly from caspase-3/-7 activity. The duplicated spots corresponding to DEVD-CHO, a known caspase-3/-7 binding sequence, are boxed in (A).

targeting this deadly disease is still a major healthcare priority in many countries.[12,24] By comparing the two parasites (Figure 4B), we were able to identify a total of 41 binders on the microarrays that uniquely bound either parasite (18 sequences for *T. brucei* and 23 sequences for *P. falciparum*). This enabled reliable

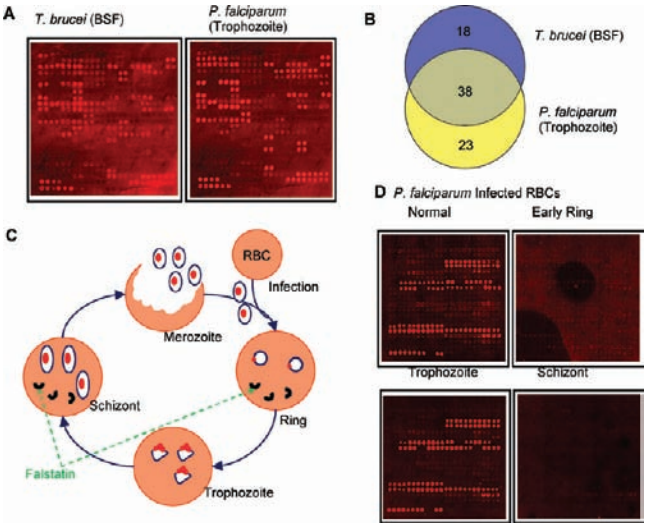
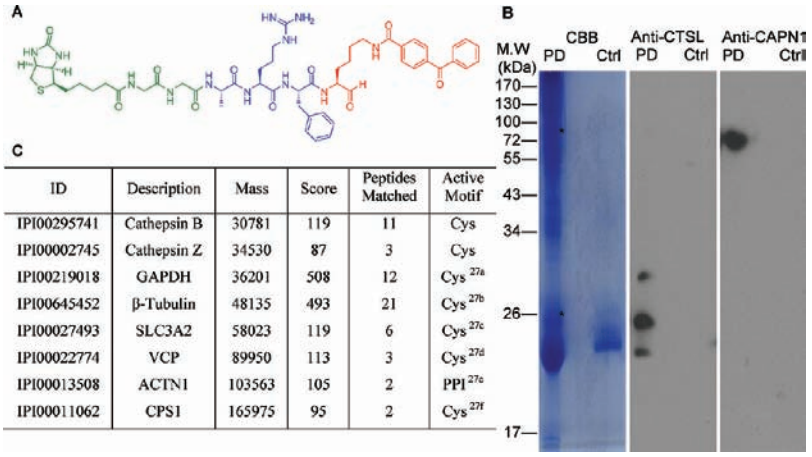


Figure 4. Profiling parasites and infected red blood cell lysates using the peptide aldehyde SMM. (A) Microarray profiles with *T. brucei* (BSF) and *P. falciparum* (trophozoite) lysates. (B) Venn diagram distribution of the top binders (>500 RFUs) in (A). (C) The typical life cycle of *P. falciparum* within the human blood stages. The black arc represents stages where falstatin was secreted. (D) Microarray profiles with *P. falciparum*-infected RBC lysates. The lysates were first activated with Ca^{2+}.

differentiation of each parasite on the array, indicating that our SMM might be further developed into a platform for high-throughput rapid parasite screening. We next assessed whether our SMM platform could be used to dissect enzymatic profiles of red blood cells (RBCs) infected at different parasitic stages by closely monitoring the up-/down-regulation of endogenous calpains expressed in RBCs. Most of the calpains present in RBCs (both normal and parasite-infected) were pro-calpains and failed to produce any significant binding profile in our SMM (Figure S9 in Supporting Information). Ca^{2+}-activated normal RBCs, on the other hand, produced a very distinctive microarray-binding profile, similar to that of Ca^{2+}-activated Hela cells (Figure 4D, top-left panel). It is well-established that falstatin, a potent cysteine protease inhibitor, is secreted into RBCs by the parasite at the ring and schizont stages but not at the trophozoite stage (Figure 4C),[13] Upon Ca^{2+} activation of parasite-infected RBCs, we observed the corresponding calpain-binding profiles *only* at the trophozoite stage, but *not* at the ring and schizont stages (Figure 4D). We also observed the corresponding protease activity in the lysates (Figure S4). We reasoned that this is due to characteristic biological differences across these different infection stages in controlling falstatin secretion, which consequently affected the binding of active calpains to our SMM. This demonstrates that our platform can be used to discern different parasite-infected erythrocyte cell stages.

Protein Target Identification and Validation. Another advantage of our SMM platform is the ability to conveniently transform hits into chemical probes for target validation and identification of potential new biomarkers in different cellular events. To demonstrate this, we selected one ubiquitous hit sequence, ARFK-CHO, on the basiss of our microarray results against Hela cells and converted it into an affinity-based photo-cross-linking probe, Biotin-GG-ARFK(Bp)-CHO (Figure 5A). Previously, we and others had used similar approaches to successfully convert non-covalent enzyme inhibitors into covalent activity-based probes (ABPs) suitable for large-scale activity-based protein



ID	Description	Mass	Score	Peptides Matched	Active Motif
IPI00295741	Cathepsin B	30781	119	11	Cys
IPI00002745	Cathepsin Z	34530	87	3	Cys
IPI00219018	GAPDH	36201	508	12	Cys^{27a}
IPI00645452	β-Tubulin	48135	493	21	Cys^{27b}
IPI00027493	SLC3A2	58023	119	6	Cys^{27c}
IPI00022774	VCP	89950	113	3	Cys^{27d}
IPI00013508	ACTN1	103563	105	2	PPI^{27e}
IPI00011062	CPS1	165975	95	2	Cys^{27f}

Figure 5. Identification of protein targets by pull-down experiments with a selected probe modified with a photo-cross-linking moiety. (A) The chemical structure of the benzophenone (Bp)-containing probe used for pull-down experiments. Aldehyde warhead coupled with benzophenone is highlight in red, the peptide recognition element is indicated in blue, and the biotin tag is shown in green. (B) Western blot analysis of protein targets pulled down from Hela cell lysates using the probe shown in (A). Proteins were identified by the respective antibodies. Asterisks show the expected locations of cathepsin L and calpain-1. Abbreviations: CBB, Coomassie Brilliant Blue; PD, positive pull-down assay; Ctrl, avidin beads without probe; CTSL, cathepsin L; CAPN1, calpain-1. (C) Pull-down and LCMS/MS results, identifying potential protein targets that bound to the probe (PPI, protein−protein interaction domain). For the complete table, see Table S1 in the Supporting Information.

profiling (ABPP) applications.[25] The introduction of the photoactive moiety benzophenone (Bp) at the P_1 position of the probe enables us to target clan CA proteases (which accept P_1 changes with high tolerance[11,26]) and, at the same time, might reveal other potential binders. As shown in Figure 5B,C, the probe successfully pulled down several proteins. Cathepsin L and calpain-1, both of which are clan CA proteases, were detected by immunoblotting (Figure 5B). Another two members of the same clan, cathepsin B and cathepsin Z, were also identified by mass spectrometry (Figure 5C). Proteins that appeared in the control pull-down experiment (beads only) were excluded from the mass spectrometry hit list. Most other potential targets identified by MS were found to contain cysteine residues in their functional domains (Figure 5C).[27] Several, like β-tubulin and GAPDH, are highly abundant proteins and so might not be true hits. Nevertheless, this demonstrates the potential of our SMM strategy in the development of novel probes and the identification of new protein biomarkers.[7b]

CONCLUSION

We have developed a peptide aldehyde small-molecule microarray (SMM) for profiling the activity of cysteine proteases. Unlike other existing enzyme-detecting microarray platforms, which are mostly substrate-based,[7−10,16] our study herein represents the first example of small-molecule inhibitors being applied *en masse*, in a microarray format, for large-scale functional profiling of cysteine proteases present in native biological samples. We have demonstrated the utility of the platform in dissecting functional cellular signatures at the proteome level to characterize cell states, as well as to develop probes with which to identify potential protein targets and biomarkers. The ability to screen within lysates provides a distinct advantage for SMM-based screening; solution-phase PSL screening, in general, only works with pure proteins.[14−16] This approach will thus open up important ways in which SMMs can be applied to probe biological states, by providing a new-found means to differentiate complex proteomic samples on the basis of differences in activity profiles. Comparative functional proteomics, with the use of SMMs, will strongly supplement existing proteomic techniques and provide a valuable window into biological functions in a rapid and efficient manner.[28]

MATERIALS AND METHODS

Materials. All chemicals were purchased at the highest grade available from commercial vendors and used without further purification, unless otherwise noted. All reactions were carried out under an N_2 atmosphere with HPLC-grade solvents, unless otherwise stated. Analytical HPLC was carried out on Shimadzu LC-IT-TOF system equipped with an autosampler, using reverse-phase Phenomenex Luna 5 μm C_{18} 100 Å 50 × 3.0 mm columns. TFA/H_2O (0.1%) and TFA/acetonitrile (0.1%) were used as eluents for all HPLC experiments. Recombinant caspases were expressed, purified, and fluorescently labeled as previously described.[19]

Synthesis of Amino Aldehyde PSL. Synthesis of all amino aldehydes was based on previously published procedures, details of which are included in the Supporting Information.[18,19] The amino aldehydes were subsequently loaded onto a threonine-functionalized aminomethyl polystyrene (TG) resin by acid-catalyzed oxazolidine formation. Generally, Fmoc-amino-CHO (3 equiv) was dissolved in 1% DIEA/MeOH, added to the TG resin, and shaken for 2 h at 60 °C. The resin was washed with MeOH (3×), DCM (3×), and DMF (3×). Boc_2O (10 equiv) was dissolved in DMF and added to the loaded resin, followed by addition of DIEA (20 equiv). The resin was shaken for 2 h, filtered, and washed with DMF (3×), DCM (3×), and DMF (3×). The resin was then dried and swollen in DMF for 30 min prior to coupling the first Fmoc amino acid. Fmoc amino acids were coupled sequentially onto the resin-bound aldehyde using standard procedures for coupling and Fmoc deprotection for the synthesis of the peptide. Biotin was coupled to the *N*-terminus of each peptide using the same HOBt/DIC/DIEA activation method with an extended coupling cycle (24 h at room temperature). The side chains and Boc-protected oxazolidines were deprotected by shaking the resin in deprotection cocktail of TFA/TIS/EDT (95:5:0.1) for 30 min. The peptide aldehydes were released from the solid support by adding a cleavage cocktail of ACN/H_2O/TFA (60:40:0.1) to the resin and shaking for 30 min at 60 °C. This release procedure was repeated another two times. Subsequently the compounds were concentrated *in vacuo* (Genevac, USA) to remove most of the cleavage cocktail (>80%), followed by addition of cold ether (chilled to −20 °C), and kept

dx.doi.org/10.1021/ja109597v |*J. Am. Chem. Soc.* 2011, 133, 1946–1954

in a −20 °C freezer overnight. The precipitated peptides were collected, washed with more cold ether, and then dried thoroughly *in vacuo*. The resulting peptide solids were dissolved in 1 mL of DMSO and stored at −80 °C in 96-deepwell plates as master stocks for future use.

Peptide Aldehyde Microarray Preparation and Screening. Procedures for the construction of the peptide aldehyde SMM were based on previously published protocols[7] and are described in detail in the Supporting Information. Protein/proteome samples were minimally labeled with either Cy5 or fluorescein NHS ester (obtained from Cy5, GE Amersham and Pierce, respectively) for 1 h on ice, following the manufacturer's protocols and our previously published procedures.[7] The non-reacted dye was quenched with a 10-fold molar excess of hydroxylamine for a further 1 h. The excess dye was removed by buffer exchange with a Microcon centrifugal filter (Millipore, USA). The labeled protein was reconstituted in a final buffer volume of 80 μL of PBS (pH 7.4) containing 1% bovine serum albumin. In a standard microarray experiment, the labeled protein (50 μL in 1 mM DTT) was applied under coverslip to the array. In a dose-dependent experiment for K_D measurements, various concentrations of the protein (4000 nM to 100 nM) were applied to different subarrays on the same slide, as previously described.[7] The samples were incubated with the array in a humidified chamber for 2 h at room temperature before repeated rinses with PBST (PBS containing 0.05% of Tween 20), typically three 10-min washes with gentle shaking. Slides were scanned using an ArrayWoRx microarray scanner installed with the relevant filters (Cy3, $\lambda_{ex/em}$ = 548/595 nm; Cy5, $\lambda_{ex/em}$ = 633/685 nm; fluorescein, $\lambda_{ex/em}$ = 490/528 nm).

For mammalian lysates, cell pellets were thawed on ice by suspension in 100 μL of lysis buffer (0.1% NP-40 in PBS), incubated further for 15 min, and then spun at 13 000 rpm at 4 °C for 20 min. Cleared supernatant was decanted to an Eppendorf tube and kept at 4 °C. Protein concentration was assessed by Bradford assay (Bio-Rad, USA). *P. falciparum* and RBC lysates were prepared as described below. Briefly, parasites of the *P. falciparum* 3C7 strain were cultured in RPMI medium 1640 (Invitrogen, USA) supplemented with 0.29 g of L-glutamine and 0.05 g of hypoxanthine dissolved in 1 mL of 1 M NaOH. Parasites were synchronized twice, 16 h apart, at ring stage using 2.5% sorbitol. Cultures were stored at 37 °C after gassing with a 5% CO_2, 3% O_2, and 92% N_2 gas mixture, and their hematocrit was maintained at 2.5%. Parasitized red blood cells were collected by centrifugation and treated with 0.1% Saponin in PBS for 15 min at room temperature with shaking. The parasite pellet was collected, resuspended in 50 mM sodium acetate (pH 5.5, containing 1 mM DTT and 0.1% NP-40), and homogenized. The supernatant (infected RBC lysates, without parasites) was kept for further microarray studies. For *T. brucei* (bloodstream form), the parasites were directly lysed by lysis buffer (50 mM sodium acetate at pH 5.5, 1 mM DTT, and 0.1% NP-40). To detect the RBC infection on the microarray, four different stages of RBC lysates were used. The Cy5-labeled lysate was preactivated with $CaCl_2$ (20 μM) at room temperature for 30 min before being applied onto the array. Control experiments were done without $CaCl_2$. The concentration of the lysate was fixed at 0.16 μg/μL during the array screening.

Tyramide Signal Amplification (TSA) Assay. Diluted cellular lysates were applied onto microarray slides. After incubation, slides were quickly rinsed with deionized water. Primary antibody blocking solution was applied onto slides and incubated at room temperature for 1 h, followed by washing and incubation with the HRP-conjugated secondary antibody. Slides were washed three times for 2 min with PBST (PBS containing 0.1% Tween 20). A TSA kit (Invitrogen) was used, following the manufacturer's recommended protocols.

Ca^{2+}-Activated Protease Assay. Protease activity in normal/infected RBC lysates was assayed in black polypropylene flat-bottom 384-well microtiter plates (Greiner, Germany) by monitoring the hydrolysis of the fluorogenic substrate, BODIPY FL casein (EnzChek Protease Assay Kits, Invitrogen). The screening concentration of 0.16 μg/μL was used consistently for all lysate samples in 25 μL assays buffered in PBS (pH 7.4, containing 1 mM DTT with or without 20 μM $CaCl_2$). The lysate was incubated with $CaCl_2$ at room temperature for 30 min before addition of the substrate. Subsequently, the plate was immediately monitored with a BioTek Synergy 4 fluorescence plate reader (BioTek, USA) at λ_{ex} = 480 nm and λ_{em} = 530 nm. For RBC lysates, the lysate preparation was done as below. Normal RBCs were collected by centrifugation and treated with 0.1% Saponin in PBS for 15 min at room temperature with shaking. Parasitized RBCs were similarly collected. The parasites were removed by centrifuge. The supernatant (infected RBC lysates, without parasites) was kept for further studies. Protein concentration was assessed by Bradford assay (Bio-Rad).

Pull-Down/LCMS Analysis. The pull-down assays were carried out with two different probes, the non-covalent probe Biotin-GG-DXXD-CHO (coding **D10** on microarray) and the photo-cross-linking probe Biotin-GG-ARFK(Bp)-CHO (see Supporting Information for details). For the non-covalent pull-down, compound **D10** (5 μM) was incubated with NeutrAvidin agarose beads (Pierce, USA) for 2 h at room temperature. Subsequently, the beads were washed with PBST three times to remove excessive compound. Cellar lysates (apoptotic Hela, 5 mg) were incubated with the beads for 1 h in an acetic buffer (50 mM sodium acetate at pH 5.5 containing 1 mM DTT and 0.1% NP-40). After incubation, the beads were washed with PBST (PBS containing 0.1% Tween 20) three times. Bound proteins were eluted with 0.1 M NaOH and boiled in the SDS loading buffer (50 mM Tris at pH 6.8 containing 100 mM DTT and 1% SDS) before being separated on an SDS−PAGE gel. A negative control (same pull-down procedure but without the probe) was run concurrently. For photo-cross-linking pull-down experiments, cellar lysates (Hela, 5 mg) were incubated with the corresponding probe (5 μM) for 1 h at room temperature in an acetic buffer (50 mM sodium acetate at pH 5.5 containing 1 mM DTT and 0.1% NP-40). Subsequently the sample was irradiated on ice for 30 min using a B100A hand-held UV lamp (UVP, USA) from a distance of 5 cm. The resulting mixture was incubated with NeutrAvidin agarose beads for 2 h at room temperature. Next the supernatant was removed by centrifugation, and the beads were washed with 1% SDS in PBS (4×). After washing, the beads were boiled in the elution buffer (200 mM Tris at pH 6.8 containing 400 mM DTT and 4% SDS). Eluted proteins were separated on an SDS−PAGE gel (together with negative pull-down controls).

Following SDS−PAGE separation, protein bands were visualized by Coomassie blue/silver staining. Gel lanes were each cut into multiple slices. Subsequently, trypsin digestion (using In-Gel Trypsin Digestion Kit, Pierce) and peptide extraction (with 50% acetonitrile and 1% formic acid) were carried out. All samples were dried *in vacuo* and stored at −20 °C. LCMS/MS analysis was performed using an LTQ-FT ultra mass spectrometer (Thermo Electron, Germany) coupled with an online Shimadzu UFLC system utilizing nanospray ionization. Peptides were first enriched with a Zorbax 300SB C_{18} column (5 mm × 0.3 mm, Agilent Technologies), followed by elution into an integrated nanopore column (75 μm × 100 mm) packed with C_{18} material (5 μm particle size, 300 Å pore size). Mobile phase A (0.1% formic acid in H_2O) and mobile phase B (0.1% formic acid in acetonitrile) were used to establish the 90-min gradient, comprising 3 min of 0−5% B, then 52 min of 5−30% B, followed by 12 min of 30−60% B; maintained at 80% B for 8 min before re-equilibrating at 5% B for 15 min. Sample was injected into the MS with an electrospray potential of 1.8 kV without sheath and auxiliary gas flow, ion transfer tube temperature of 180 °C, and collision gas pressure of 0.85 mTorr. A full-survey MS scan (350−2000 m/z range) was acquired in the 7-T FT-ICR cell at a resolution of 100 000 and a maximum ion accumulation time of 1000 ms. Precursor ion charge-state screening was activated. The linear ion trap was used to collect peptides where the 10 most intense ions were selected for collision-induced dissociation (CID) in MS^2, performed concurrently with a maximum ion accumulation time of 200 ms. Dynamic exclusion

was activated for this process, with a repeat count of one and exclusion duration of 30 s. For CID, the activation Q was set at 0.25, isolation width (m/z) 2.0, activation time 30 ms, and normalized collision energy 35%. The Extract_Msn (version 4.0) program found in Bioworks Browser 3.3 (Thermo Electron, Germany) was used to extract tandem MS spectra in the dta format from the raw data of the LTQ-FT ultra. These dta files were then converted into MASCOT generic file format using an in-house program. Intensity values and fragment ion m/z ratios were not manipulated. These data were used to obtain protein identities by searching against the corresponding database by means of an in-house MASCOT server (version 2.2.03, Matrix Science, Boston, MA). The search was limited to a maximum of two missed trypsin cleavages, #13C of 2, mass tolerances of 10 ppm for peptide precursors and 0.8 Da for fragment ions. Fixed modification was carbamidomethyl at the Cys residue, whereas variable modifications were oxidation at the Met residue and phosphorylation at the Ser, Thr, or Tyr residue. Only proteins with a MOWSE score higher than 70, corresponding to $p < 0.05$, were considered significant. The peptide/protein lists obtained were exported to an html file.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information. Experimental details, chemical synthesis and characterizations, complete protocols for array fabrication, K_D binding analysis, IC_{50} analysis, and additional biochemical experiments. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
chmyaosq@nus.edu.sg

■ ACKNOWLEDGMENT

Funding support was provided by the Agency for Science, Technology and Research (R-143-000-391-305) and the Ministry of Education (R-143-000-394-112). We thank Prof. McKerrow (UCSF) for providing proteases cruzain and rhodesain, and the groups of Kevin Tan and Cynthia He (NUS) for donating parasites *P. falciparum* and *T. brucei* (BSF), respectively.

■ REFERENCES

(1) (a) MacBeath, G.; Saghatelian, A. *Curr. Opin. Chem. Biol.* **2009**, *13*, 501–502. (b) Aebersold, R.; Mann, M. *Nature* **2003**, *422*, 198–207.

(2) (a) Brown, P. P.; Botstein, D. *Nat. Genet.* **1999**, *21*, 33–37. (b) Miura, S.; Miura, K.; Masuzaki, H.; Miyake, N.; Yoshiura, K.; Sosonkina, N.; Harada, N.; Shimokawa, O.; Nakayama, D.; Yoshimura, S.; Matsumoto, N.; Niikawa, N.; Ishimaru, T. *J. Hum. Genet.* **2006**, *51*, 412–417.

(3) (a) Gstaiger, M.; Aebersold, R. *Nat. Rev. Genet.* **2009**, *10*, 617–627. (b) Cravatt, B. F.; Simon, G. M.; Yates, J. R., III *Nature* **2007**, *450*, 991–1000.

(4) (a) Cravatt, B. F.; Wright, A. T.; Kozarich, J. W. *Annu. Rev. Biochem.* **2008**, *77*, 383–414. (b) Uttamchandani, M.; Lu, C. H.; Yao, S. Q. *Acc. Chem. Res.* **2009**, *42*, 1183–1192.

(5) (a) Evans, M. J.; Cravatt, B. F. *Chem. Rev.* **2006**, *106*, 3279–3301. (b) Sadaghiani, A. M.; Verhelst, S. H. L.; Bogyo, M. *Curr. Opin. Chem. Biol.* **2007**, *11*, 20–28. (c) Uttamchandani, M.; Li, J.; Sun, H.; Yao, S. Q. *ChemBioChem* **2008**, *9*, 667–675.

(6) (a) Uttamchandani, M.; Yao, S. Q. *Curr. Pharm. Des.* **2008**, *14*, 2428–2438. (b) Duffner, J. L.; Clemons, P. A.; Koehler, A. N. *Curr. Opin. Chem. Biol.* **2007**, *11*, 74–82. (c) Tomizaki, K. Y.; Usui, K.; Mihara, H. *ChemBioChem* **2005**, *6*, 783–799. (d) Kodadek, T. *Trends Biochem. Sci.* **2002**, *27*, 295–300.

(7) (a) Sun, H.; Tan, L. P.; Gao, L.; Yao, S. Q. *Chem. Commun.* **2009**, 677–679. (b) Shi, H.; Liu, K.; Xu, A.; Yao, S. Q. *Chem. Commun.* **2009**, 5030–5032. (c) Lu, C. H. S.; Sun, H.; Bakar, F. B. A.; Uttamchandani, M.; Zhou, W.; Liou, Y.-C.; Yao, S. Q. *Angew. Chem., Int. Ed.* **2008**, *47*, 7438–7441. (d) Sun, H.; Lu, C. H. S.; Uttamchandani, M.; Xia, Y.; Liou, Y.-C.; Yao, S. Q. *Angew. Chem., Int. Ed.* **2008**, *47*, 1698–1702. (e) Sun, H.; Lu, C. H. S.; Shi, H.; Gao, L.; Yao, S. Q. *Nat. Protoc.* **2008**, *3*, 1485–1493. (f) Uttamchandani, M.; Lee, W. L.; Wang, J.; Yao, S. Q. *J. Am. Chem. Soc.* **2007**, *129*, 13110–13117.

(8) (a) Salisbury, C. M.; Maly, D. J.; Ellman, J. A. *J. Am. Chem. Soc.* **2002**, *124*, 14868–14870. (b) Kohn, M.; Gutierrez-Rodriguez, M.; Jonkheijm, P.; Wetzel, S.; Wacker, R.; Schroeder, H.; Prinz, H.; Niemeyer, C. M.; Breinbauer, R.; Szedlacsek, S. E.; Waldmann, H. *Angew. Chem., Int. Ed.* **2007**, *46*, 7700–7703.

(9) Winssinger, N.; Damoiseaux, R.; Tully, D. C.; Geierstanger, B. H.; Burdick, K.; Harris, J. L. *Chem. Biol.* **2004**, *11*, 1351–1360.

(10) (a) Reddy, M. M.; Kodadek, T. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 12672–12677. (b) Roska, R. L. W.; Lama, T. G. S.; Henne, J. P.; Carlson, R. E. *J. Am. Chem. Soc.* **2009**, *131*, 16660–16662. (c) Usui, K.; Tomizaki, K.; Mihara, H. *Mol. BioSyst.* **2006**, *2*, 417–420.

(11) (a) Lecaille, F.; Kaleta, J.; Brömme, D. *Chem. Rev.* **2002**, *102*, 4459–4488. (b) Denault, J. B.; Salvesen, G. S. *Chem. Rev.* **2002**, *102*, 4489–4500.

(12) Renslo, A. R.; McKerrow, J. H. *Nat. Chem. Biol.* **2006**, *2*, 701–710.

(13) Rosenthal, P. J. *Int. J. Parasitol.* **2004**, *34*, 1489–1499.

(14) (a) Meldal, M.; Svendsen, I.; Breddam, K.; Auzanneau, F.-I. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3314–3318. (b) Thornberry, N. A.; Rano, T. A.; Peterson, E. P.; Rasper, D. M.; Timkey, T.; Garcia-Calvo, M.; Houtzager, V. M.; Nordstrom, P. A.; Roy, S.; Vaillancourt, J. P.; Chapman, K. T.; Nicholson, D. W. *J. Biol. Chem.* **1997**, *272*, 17907–17911. (c) Rano, T. A.; Timkey, T.; Peterson, E. P.; Rotonda, J.; Nicholson, D. W.; Becker, J. W.; Chapman, K. T.; Thornberry, N. A. *Chem. Biol.* **1997**, *4*, 149–155.

(15) (a) Backes, B. J.; Harris, J. L.; Leonetti, F.; Craik, C. S.; Ellman, J. A. *Nat. Biotechnol.* **2000**, *18*, 187–193. (b) Harris, J. L.; Backes, B. J.; Leonetti, F.; Mahrus, S.; Ellman, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 7754–7759. (c) Choe, Y.; Leonetti, F.; Greenbaum, D. C.; Lecaille, F.; Bogyo, M.; Bromme, D.; Ellman, J. A.; Craik, C. S. *J. Biol. Chem.* **2006**, *281*, 12824–12832. (d) Debela, M.; Magdolen, V.; Schechter, N.; Valachova, M.; Lottspeich, F.; Craik, C. S.; Choe, Y.; Bode, W.; Goettig, P. *J. Biol. Chem.* **2006**, *281*, 25678–25688.

(16) (a) Ghosalia, D. N.; Diamond, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 8721–8726. (b) Ghosalia, D. N.; Salisbury, C. M.; Ellman, J. A.; Diamond, S. L. *Mol. Cell Proteomics* **2005**, *4*, 626–636. (c) Gosalia, D. N.; Salisbury, C. M.; Maly, D. J.; Ellman, J. A.; Diamond, S. L. *Proteomics* **2005**, *5*, 1292–1298.

(17) (a) Dooley, C. T.; Houghten, R. A. *Methods Mol. Biol.* **1998**, *87*, 13–24. (b) Pinilla, C.; Appel, J. R.; Blanc, P.; Houghten, R. A. *Biotechniques* **1992**, *13*, 901–905.

(18) Wang, G.; Uttamchandani, M.; Chen, G. Y. J.; Yao, S. Q. *Org. Lett.* **2003**, *5*, 737–740.

(19) Ng, S. L.; Yang, P. -Y.; Chen, K. Y. -T.; Srinivasan, R.; Yao, S. Q. *Org. Biomol. Chem.* **2008**, *6*, 844–847.

(20) Rawlings, N. D.; Barrett, A. J.; Bateman, A. *Nucleic Acids Res.* **2010**, *38*, D227–D233.

(21) Uttamchandani, M.; Wang, J.; Li, J.; Hu, M.; Sun, H.; Chen, K. Y.-T.; Liu, K.; Yao, S. Q. *J. Am. Chem. Soc.* **2007**, *129*, 7848–7858.

(22) Snow, R. W.; Guerra, C. A.; Noor, A. M.; Hay, S. I. *Nature* **2005**, *434*, 214–217.

(23) Tafani, M.; Minchenko, D. A.; Serroni, A.; Farber, J. L. *Cancer Res.* **2001**, *61*, 2459–2466.

(24) Kar, S.; Kar, S. *Nat. Rev. Drug Discov.* **2010**, *9*, 511–512.

(25) (a) Liu, K.; Shi, H.; Xiao, H.; Chong, A. G. L.; Bi, X.; Chang, Y. T.; Tan, K.; Yada, R. Y.; Yao, S. Q. *Angew. Chem., Int. Ed.* **2009**, *48*, 8293–8297. (b) Wang, J.; Uttamchandani, M.; Li, J.; Hu, M.; Yao, S. Q. *Chem. Commun.* **2006**, 3783–3785. (c) Chan, E. W. S.; Chattopadhaya, S.; Panicker, R. C.; Huang, X.; Yao, S. Q. *J. Am. Chem. Soc.* **2004**, *126*,

14435–14446. (d) Saghatelian, A.; Jessani, N.; Joseph, A.; Humphrey, M.; Cravatt, B. F. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 10000–10005. (e) Sieber, S. A.; Niessen, S.; Hoover, H. S.; Cravatt, B. F. *Nat. Chem. Biol.* **2006**, *2*, 274–281.

(26) (a) Blum, G.; Mullins, S. R.; Keren, K.; Fonovic, M.; Jedeszko, C.; Rice, M. J.; Sloane, B. F.; Bogyo, M. *Nat. Chem. Biol.* **2005**, *1*, 203–209. (b) Blum, G.; von Degenfeld, G.; Merchant, M. J.; Blau, H. M.; Bogyo, M. *Nat. Chem. Biol.* **2007**, *3*, 668–677.

(27) (a) Nakajima, H.; Amano, W.; Fujita, A.; Fukuhara, A.; Azuma, Y.-T.; Hata, F.; Inui, T.; Takeuchi, T. *J. Biol. Chem.* **2007**, *282*, 26562–26574. (b) Ludueña, R. F.; Banerjee, A.; Khan, I. A. *Curr. Opin. Cell Biol.* **1992**, *4*, 53–57. (c) Fort, J.; de la Ballina, L. R.; Burghardt, H. E.; Ferrer-Costa, C.; Turnay, J.; Ferrer-Orta, C.; Uson, I.; Zorzano, A.; Fernandez-Recio, J.; Orozco, M.; Lizarbe, M. A.; Fita, I.; Palacín, M. *J. Biol. Chem.* **2007**, *282*, 31444–31452. (d) Noguchi, M.; Takata, T.; Kimura, Y.; Manno, A.; Murakami, K.; Koike, M.; Ohizumi, H.; Hori, S.; Kakizuka, A. *J. Biol. Chem.* **2005**, *280*, 41332–41341. (e) Harper, B. D.; Beckerle, M. C.; Pomies, P. *Biochem. J.* **2000**, *350*, 269–274. (f) Corvi, M. M.; Soltys, C.-L. M.; Berthiaume, L. G. *J. Biol. Chem.* **2001**, *276*, 45704–45712.

(28) Weinrich, D.; Jonkheijm, P.; Niemeyer, C. M.; Waldmann, H. *Angew. Chem., Int. Ed.* **2009**, *48*, 7744–7751.