Chemistry & Biology, Vol. 11, 1351–1360, October, 2004, ©2004 Elsevier Ltd. All rights reserved. DOI 10.1016/j.chembiol.2004.07.015

PNA-Encoded Protease Substrate Microarrays

Nicolas Winssinger,[1,*] Robert Damoiseaux,[2] David C. Tully,[2] Bernhard H. Geierstanger,[2] Keith Burdick, and Jennifer L. Harris[2,3,*]
[1]Institut de Science et d'Ingénierie Supramoléculaires
Université Louis Pasteur
8 allée Guaspard Monge
67000 Strasbourg
France
[2]Genomics Institute of the Novartis Research Foundation
10675 John Jay Hopkins Drive
San Diego, California 92121
[3]Department of Molecular Biology
The Scripps Research Institute
10550 North Torrey Pines Road
La Jolla, California 92037

Summary

Our current understanding of the role and regulation of protease activity in normal and pathogenic processes is limited by our ability to measure and deconvolute their enzymatic activity. To address this limitation, an approach was developed that utilizes rhodamine-based fluorogenic substrates encoded with PNA tags. The PNA tags address each of the substrates to a predefined location on an oligonucleotide microarray through hybridization, thus allowing the deconvolution of multiple signals from a solution. A library of 192 protease substrates was prepared by split and mix combinatorial synthesis. The methodology and validation of this approach for profiling proteolytic activity from single proteases and from those in crude cell lysates as well as clinical blood samples is described.

Introduction

The precise and limited action of proteases is a mechanism by which cells regulate many vital events. Proteases themselves are nearly exclusively regulated by posttranslational modifications. There are over 500 putative human proteases [1], and a number of these are involved in the regulation of essential cellular processes such as DNA replication, cell cycle progression, differentiation, migration, morphogenesis, immunity, hemostasis, neuronal outgrowth, and apoptosis. Misregulation of proteolytic activity is involved in many pathological situations like neurodegeneration, cardiovascular diseases, arthritis, cancer, and infectious diseases [2]. Furthermore, a number of pathogens require the activity of an endogenous or host protease in their infection/replication cycle. Proteases are thus attractive therapeutic targets, and monitoring their activity can be diagnostic and prognostic of disease states.

A major characteristic of a given protease is its substrate specificity that can range from very broad—as for proteases involved in catabolism—to very narrow—as for proteases involved in the regulation of cellular events. Historically, the substrate specificity of a protease was determined through the identification of cleavage sites in macromolecular substrates or through the screening of individually synthesized peptide substrates [3, 4]. However, these methods do not allow for the complete screening of the substrate specificity of a protease because not all combinations of protein substrate sequences can be considered in these experiments. Several combinatorial peptide synthesis techniques have been employed to identify the substrate specificity of proteases. These techniques include solution screening methods, such as positional scanning substrate libraries [5–7], and solid support screening methods, such as bead-based fluorescence resonance energy transfer (FRET) peptide libraries [8] and peptide microarrays [9, 10]. The advantage of screening substrates in solution is that the interaction between protease and substrate is generally more physiological in nature and avoids nonspecific interactions that can arise with solid support screening strategies [11]. Here, we describe a method for determining the substrate specificity of proteases through the solution-based screening of peptide substrate libraries and microarray detection of the cleavage products.

Results and Discussion

We have previously described the use of microarray-based technology to measure protease activity based on the selective labeling of proteases by using suicide inhibitors tethered to PNA tags [12, 13]. The inhibitors that have not reacted with a protease are removed by size exclusion filtration, and the remaining inhibitors are hybridized to an oligonucleotide array for detection. Herein, we report an alternative and complementary approach that utilizes fluorogenic protease substrates linked to PNA. As shown in Figure 1A, proteolysis of the bond connecting the substrate to the latent fluorophore changes the electronic properties of this fluorophore, resulting in a large increase in fluorescence. An advantage of latent fluorogenic substrates over substrates linking FRET fluorophores is that only proteolysis at the residue adjacent to the fluorophore gives a signal. This property allows for the register of amino acids in the substrate sequence relative to the cleavage site to be clearly defined. Rhodamine exhibits several unique properties that make it particularly well suited for this microarray application. First, the latent fluorophore is 1000-fold less fluorescent than its hydrolyzed counterpart [14] (Figure 1A). Second, rhodamine possesses an absorption/emission spectrum, which allows for the use of an argon-ion laser, a common excitation source in

*Correspondence: winssinger@isis-ulp.org (N.W.); harris@scripps.edu (J.L.H.)

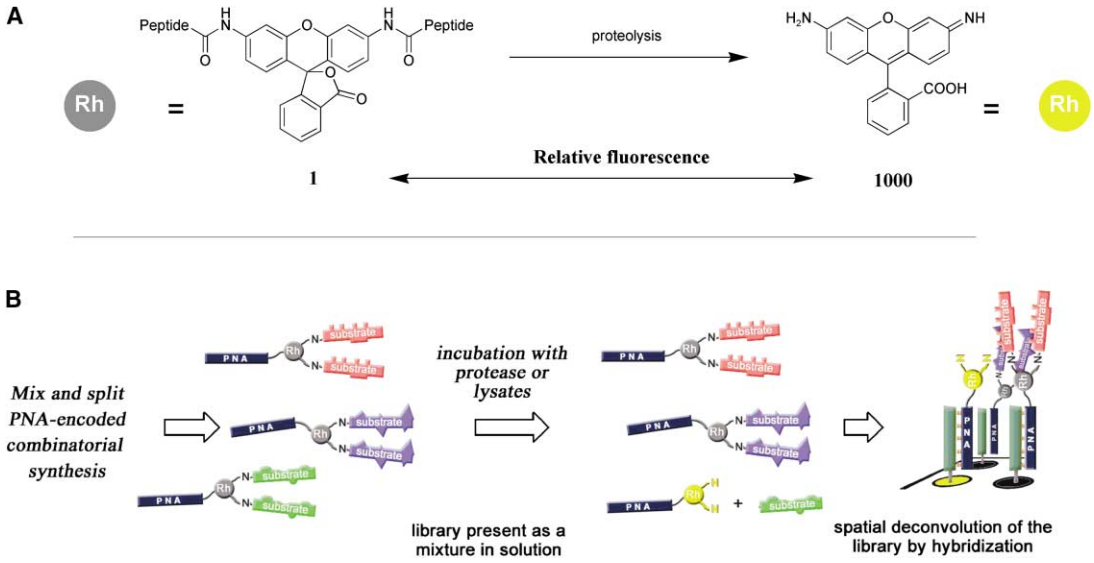


Figure 1. PNA-Encoded Fluorogenic Substrate Libraries

(A) Structure of rhodamine as a latent acylated fluorophore (left) and its fluorescent counterpart (right).
(B) Profiling with the PNA-encoded rhodamine-substrate library.

laboratory instrumentation. Third, the fluorescence of rhodamine is largely independent of the pH, allowing for the adjustment of the pH to the needs of a wide range of conditions. Last but not least, small peptides on the rhodamine have been shown to be accepted as substrates by serine and cysteine proteases [14–16]. It is recognized, however, that aspartyl- and metalloproteinases and some serine and cysteine proteases will require prime-side interactions with the substrate for productive hydrolysis and may not be detected with these fluorogenic substrates.

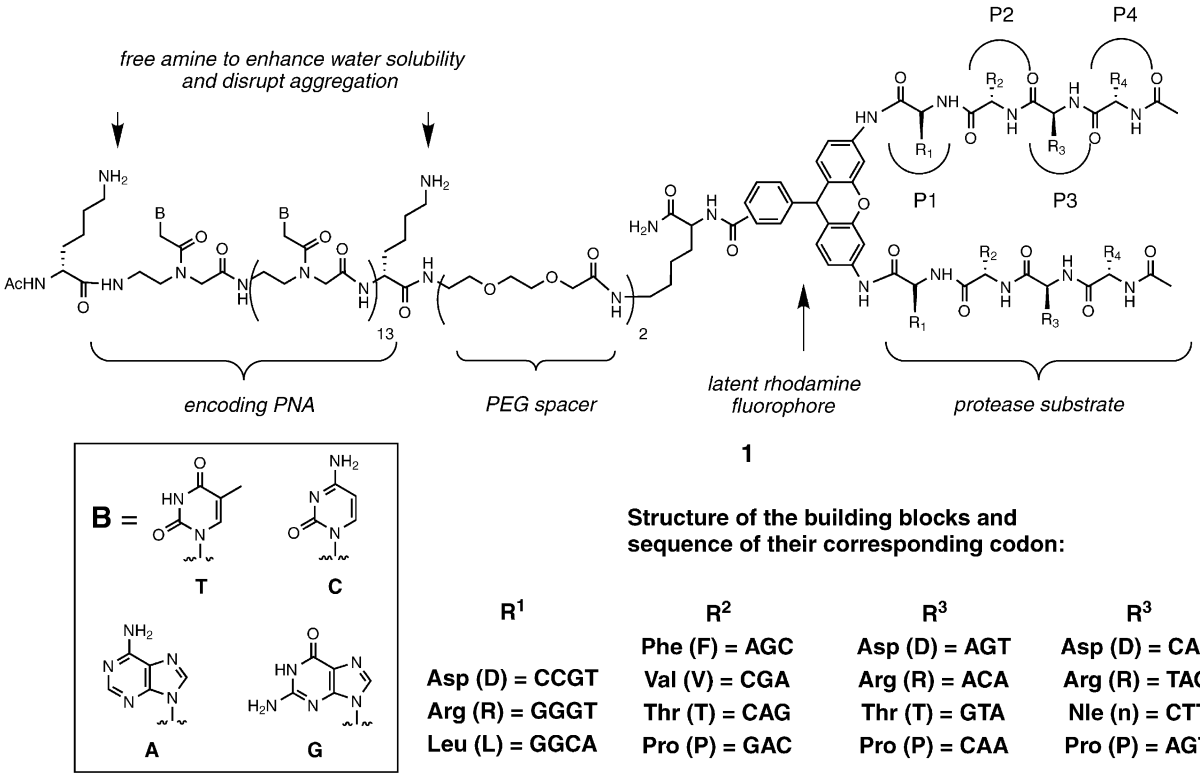
Proteolysis is measured by incubating the PNA-encoded library of substrates with the sample of interest (Figure 1B). The fluorescence signal from the library is resolved by hybridization to an oligonucleotide microarray. The location of hybridization is dictated by the sequence of the PNA, which is linked to the sequence of the substrate peptide. The general structure of the library is shown in Figure 2. Compared to the use of the suicide inhibitor library previously reported, the fluorogenic substrate strategy has the potential of being more sensitive since one protease can turn over many substrates, thus leading to signal amplification. While microarrays of fluorogenic substrates generated by printing substrates directly on a solid surface have been reported [9], the PNA-encoding system has the advantage of allowing proteolysis to be carried out in solution. This is important in order to exclude the effects of nonspecific interactions of the enzymes with the surface and offers better control of substrate/analyte concentration and reaction conditions.

To evaluate the effect of a PNA tag on substrate hydrolysis, two pairs of rhodamine peptide substrates with and without the PNA tag were synthesized (**8** versus **9** and **11** versus **12**, Figure 3). The sequences nTPR (N to C terminus, n represents norleucine) and DEVD were selected based on the fact that they are known to be good substrates for the serine protease thrombin and the cysteine protease caspase-3, respectively [7, 17]. Both substrate pairs were efficiently hydrolyzed only by their cognate enzymes. The Michaelis-Menten parameters of the PNA-tagged substrate were comparable to the substrates lacking the PNA tag (Table 1). This finding is in agreement with our previous observation that the PNA does not interfere with the inhibitory activity of tethered molecules [18].

PNAs are particularly well suited for hybridization to a DNA chip since the DNA-PNA interaction is stronger than a corresponding DNA-DNA interaction [19]. Single base pair mismatchs thus have a more pronounced destabilization on a PNA-DNA duplex than on a DNA-DNA duplex. Furthermore, the PNA-encoding strategy is ideal for the library synthesis, as it allows for the rapid generation of libraries by split and mix synthesis [20]. This strategy requires the use of orthogonal protecting groups for the cosynthesis of the peptide and PNA chains. The library synthesized herein relied on the alternating use of Fmoc-protected PNA monomers (**13**, Figure 4) and Alloc-protected amino acids (**14**, Figure 4) [21].

To evaluate selectivity and quantification of enzymatic activity using PNA-encoded substrates, an equimolar mixture of **9** and **12** was hybridized to Affymetrix GenFlex chips (Figure 5A) at concentrations ranging from 10 pM to 800 pM prior to and after treatment with caspase-3. Only the feature on the chip corresponding to caspase-3 showed a strong linear signal increase (Figure 5).

Our next goal was to prepare a PNA-encoded rhodamine substrate library. A 192 member tetrapeptide library consisting of three different amino acids in the P1 and four different amino acids in the P2–P4 positions was prepared as shown in Figure 4. Each tetrapeptide was encoded by a 14-mer PNA using four codon sets, as shown in Figure 2. The sequences of the codons



Structure of the building blocks and sequence of their corresponding codon:

R^1	R^2	R^3	R^3
	Phe (F) = AGC	Asp (D) = AGT	Asp (D) = CAA
Asp (D) = CCGT	Val (V) = CGA	Arg (R) = ACA	Arg (R) = TAG
Arg (R) = GGGT	Thr (T) = CAG	Thr (T) = GTA	Nle (n) = CTT
Leu (L) = GGCA	Pro (P) = GAC	Pro (P) = CAA	Pro (P) = AGT

Figure 2. General Structure of the 192 Member PNA-Encoded Rhodamine Substrate Library

represent a subset of the previously described codon set [12] and were chosen to balance two primary design goals, fidelity of hybridization and homogeneity of melting temperatures. The specific codons of this subset were selected based on empirical observation of the hybridization properties.

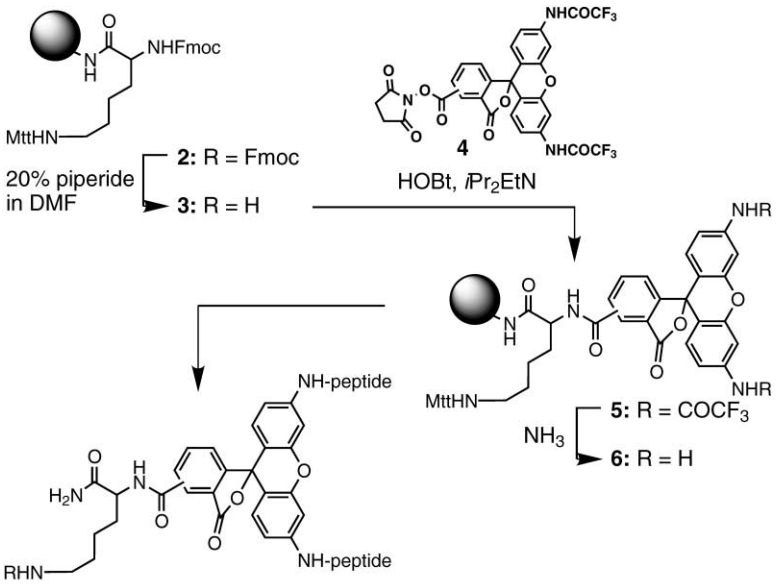


Figure 3. Synthesis of Caspase and Thrombin Substrates with and without PNA Encoding

7: peptide = DVED-Ac, R = H
8: peptide = DVED-Ac, R = Ac
9: peptide = DVED-Ac, R = GGATCTGATAGCGG
10: peptide = RPTn-Ac, R = H
11: peptide = RPTn-Ac, R = Ac
12: peptide = RPTn-Ac, R = TCGCTACGTCGCAG

Table 1. Comparison of K_m, k_{cat}, and k_{cat}/K_m of Rhodamine Protease Substrates with and without PNA Tag for Both Caspase-3 and Thrombin

Substrate	K_M	k_{cat}	k_{cat}/K_M
Caspase-3			
$(DEVD)_2$-Rh-Ac (8)	26 μM	0.083 s^{-1}	3.2×10^{-3} μM^{-1} s^{-1}
$(DEVD)_2$-Rh-PNA (9)	9.4 μM	0.032 s^{-1}	3.2×10^{-3} μM^{-1} s^{-1}
Thrombin			
$(nTPR)_2$-Rh-Ac (11)	71 μM	1.67 s^{-1}	2.3×10^{-2} μM^{-1} s^{-1}
$(nTPR)_2$-Rh-PNA (12)	247 μM	4.40 s^{-1}	1.8×10^{-2} μM^{-1} s^{-1}

We tested our PNA-encoded library by limited proteolysis with thrombin, plasmin, or caspase-3. The product of this proteolysis was then hybridized to an oligonucleotide microarray, and the fluorescence intensity corresponding to each probe was measured (Figures 6A–6C). The observed substrate specificity was indeed consistent with the known specificity for each enzyme. As expected, caspase-3 showed a strong preference for substrates with P1 = Asp, P2 = Val, and P4 = Asp, with DTVD as the optimal substrate. While thrombin and plasmin showed the same strong preference for P1 and P4 residues (Arg and Nle), plasmin was selective for P2 = Phe and had some preference for P3 = Thr, whereas thrombin showed a high specificity for P2 = Pro and broad specificity for P3.

Having established the use of the PNA-encoded substrate library to profile protease specificity, we turned our attention to its use in profiling protease activity in complex mixtures such as crude cell lysates. To this end, we opted to compare profiles of protease activity from apoptotic cell lysates versus nonapoptotic cell lysates. It is important to note that apoptosis does not manifest itself as an enhanced expression of caspase-3 but rather as an increase in enzymatic activity, which is due to a proteolytic conversion of the caspase-3 zymogen. Changes in proteolytic activities that are based on posttranslational modification can not be monitored by protein expression profiling experiments, emphasizing the importance of measuring protein function rather than mRNA message or protein levels [22, 23]. The PNA-encoded substrate library was treated either with apoptotic or nonapoptotic cell lysate and was hybridized to the oligonucleotide arrays. The profile of the nonapoptotic lysate was then subtracted from the apoptotic lysate, resulting in the difference profile. The most intense signals observed in the difference profile are found in P1 = Asp, P2 = Val, and P4 = Asp, with DTVD showing the most activity. This activity matches the specificity fingerprint of caspase-3, clearly indicating the activation of this protease in the apoptotic crude cell lysate (Figure 6D). While less intense than the caspase-3 fingerprint, other prominent activities are found for P1 = Arg and P2 = Val, indicating an increase in the activity of additional protease(s) (Figure 6D). The test library of 192 substrates only represents a small fraction of the 160,000 possible (20^4) tetrapeptide sequences, and while caspase-3 activity is clearly measurable, additional proteolytic activities may also be present that cannot be monitored with the limited set of substrates represented in this test library. Nevertheless, the results demonstrate that it is possible to use the PNA-encoded fluorogenic substrates for screening proteolytic activities in complex biological samples.

An important asset of this technology is its miniaturization, having the ability to measure the activity of hundreds of substrates in only a few microliters. This may prove important in the profiling of proteolytic activity in clinical samples. Comparative activity profiling should allow for a better understanding of protease activity involved in healthy and pathological processes or drug action. To evaluate the potential of PNA-encoded libraries toward diagnostic applications, the coagulation cascade in clinical blood samples appeared to be a relevant example. The activity profile of blood samples from healthy donors was compared to blood samples from an individual treated with warfarin, a frequently used drug for oral anticoagulant therapy. Warfarin decreases the level of active prothrombin, factor X, factor IX, factor VII, and protein C by blocking their vitamin K-dependent carboxylation of glutamic acid moieties that is essential for their synthesis and function [24]. The samples from the patient treated with warfarin were characterized as having an international normalized ratio (INR) of 5.0. The INR is the ratio of the coagulation time of a given sample divided by the coagulation time of a blood sample from a healthy donor. The samples were incubated with the PNA-encoded library and hybridized to the oligonucleotide arrays. Proteolytic activity was only observed for substrates containing P1 = R. The normal blood sample shows a profile in P1 = R, P2 = P (Figure 7) that is similar to thrombin (Figure 6A), a major proteolytic activity in the coagulation cascade. From these data, it is clear that warfarin treatment largely abolishes the thrombin-mediated signal transduction within the coagulation cascade, leading to inhibition of coagulation (Figure 7).

Significance

Proteases play a crucial role in regulating numerous cellular events. The ability to measure their activity on a proteomic scale and correlate such activity to phenotypes of interest would facilitate the understanding of their cellular function and may prove to be a sensitive diagnostic tool. Toward these goals, we have developed a method based on PNA-encoded libraries of fluorogenic substrates to profile protease specificity in a microarray format. We have shown that such a library can be used to quantify protease activity and substrate specificity of both serine and cysteine proteases in minimal volumes. More significantly, we

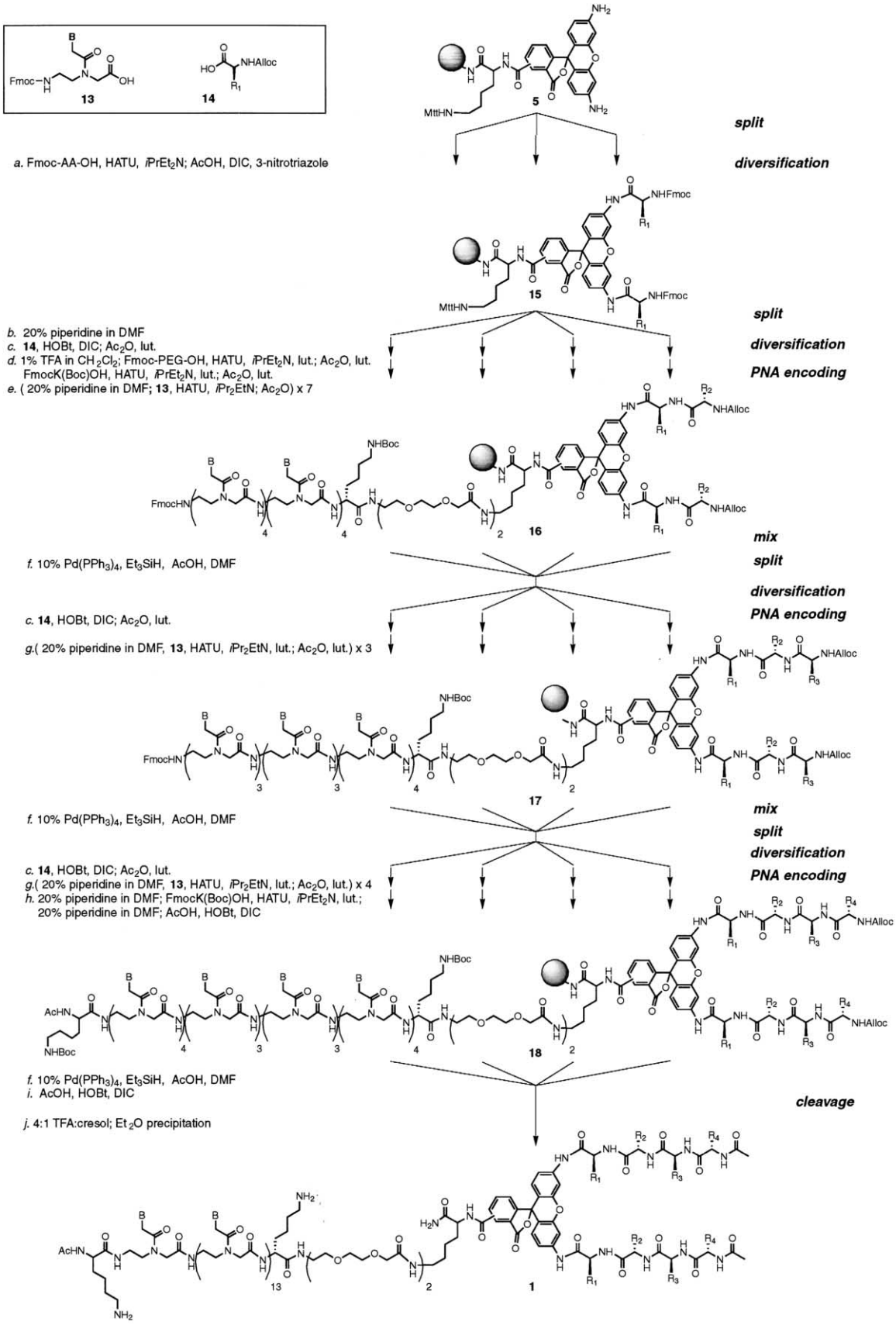
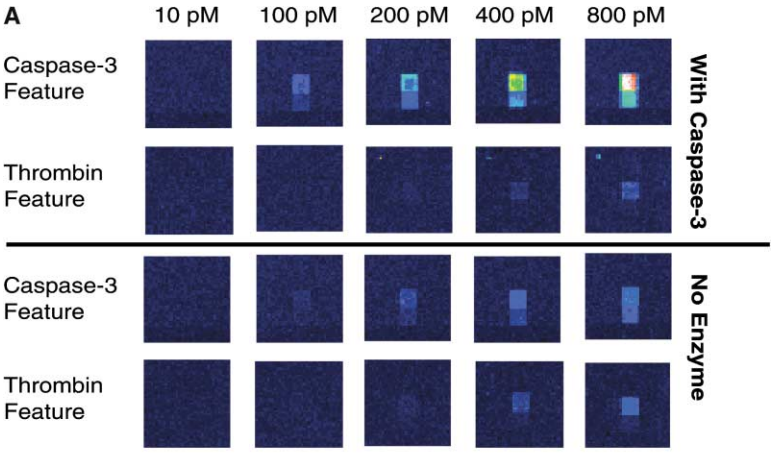


Figure 4. Synthesis of the PNA-Encoded Library



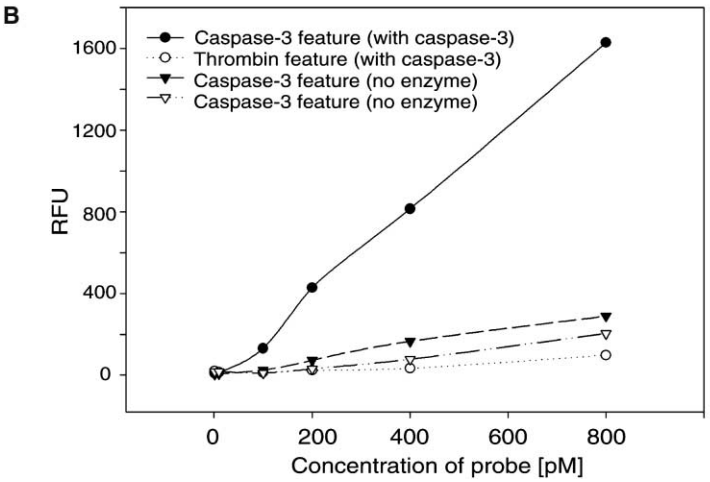


Figure 5. Proteolysis of Test Substrates

(A and B) The complementary oligonucleotide features on an Affymetrix microarray are shown for the caspase-3 and thrombin probes **9** and **12**, respectively. When caspase-3 is incubated with both probes and then hybridized on the microarray, only the caspase-3 feature shows appreciable fluorescence with a linear relationship to the concentration of the substrate shown in (B). The data shown are from a single representative microarray experiment.

demonstrated that the method can be used with complex mixtures such as crude cell lysates and clinical blood samples. Expansion of this technology to larger substrate libraries should allow for even greater versatility in the characterization of proteolytic activity with potential applications as a diagnostic tool. The reported technology should prove valuable in gaining a better understanding of complex proteolytic events involved in the regulation of cellular processes and the pathogenesis of disease. Finally, this report further expands the scope and applications of PNA-encoded libraries and their conversion into a spatially addressable microarray.

Experimental Procedures

All peptide synthesis reagents and resins were purchased from Novabiochem. PNA monomers as well as HATU were purchased from Applied Biosystem, and other chemicals were purchased from Aldrich. Thrombin was purchased from Haematologic Technologies. Plasma samples, pooled normal plasma (catalog number 0010), and the multi-coumadin set (catalog number 9400) were purchased from George King Biomedicals Inc., and activated partial thromboplastin reagents were purchased from Biomérieux and used as recommended by the manufacturer. Caspase-3 was recombinantly expressed and purified by similar methods to those described by Zhou et al. [25]. If not stated otherwise, all solid support reactions were carried out under an inert atmosphere in an Argonaut Quest 210 Organic Synthesizer.

Synthesis of TFA-Rhodamine-NHS Ester 4

3-Aminophenol (10.0 g, 0.091 mol, 3.3 eq) was dissolved in H_2SO_4 (65.0 ml, 95%–97%), and 1,2,4-benzenetricarboxilic acid (5.76 g, 0.027 mol, 1.0 eq) was added [26]. The reaction mixture was warmed to 180°C and kept at this temperature for 6 hr. After cooling to room temperature, the reaction mixture was poured onto 70 g ice and stirred. The sulfuric acid was neutralized with solid sodium carbonate, and MeOH (300 ml) was added to precipitate the inorganic salts that were removed by filtration. The filter cake was washed with MeOH (200 ml), and the methanolic solutions were combined. After removal of the solvents, the residues were dissolved in methanol and adsorbed onto silica (30 g). This dry pack product was purified by flash chromatography on silica (200 g) with a step gradient of acetonitrile/MeOH (7:3) to acetonitrile/MeOH/H_2O/Et_3N (20:5:4:1). Evaporation of the solvents under reduced pressure yielded 6.5 g of a mixture of 3′,6′-Di-amino-spiro[phtalan-1,9′-xanthene]-5-carboxylic acid and 3′,6′-Di-amino-spiro[phtalan-1,9′-xanthene]-6-carboxylic acid (0.17 mol, 63%) as dark red crystalline solid. LC/MS characterization indicated about 90% purity. An analytical sample was obtained by recrystallization from acetone/water with a product purity >98%.

The purified (90% purity level) 3′,6′-di-amino-spiro[phtalan-1,9′-xanthene]-5-carboxylic/6-carboxylic acid (188 mg, 0.5 mmol, 1.0 eq, coevaporated three times with dry pyridine) was suspended in dry

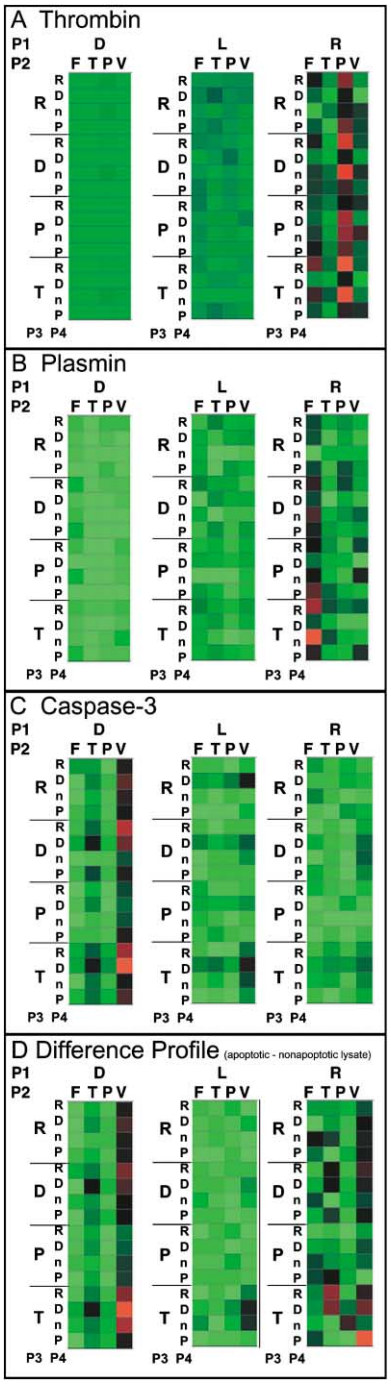


Figure 6. Fluorescence Intensities of the 192 Member PNA-Encoded Substrate Library after Incubation with Different Proteases

(A–C) Map of the spatial deconvolution signals of the 192 member PNA-encoded substrate library after incubation with (A) thrombin, (B) plasmin, or (C) caspase-3.

(D) Difference profile derived by subtracting the fluorescence intensity values of the nonapoptotic cell lysate from the apoptotic cell lysate. The activation of caspase-3 during apoptosis can be monitored by comparison of the signals on the P1 = aspartic acid subarray of the purified enzyme with the differential signal from the two lysates. The data shown represent the normalized mean from six experiments, with green representing the minimum fluorescence and red representing the maximum fluorescence.

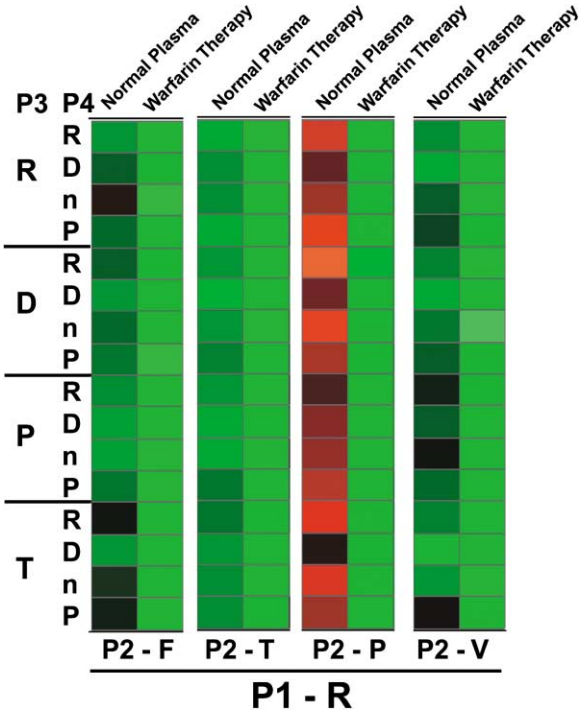


Figure 7. Proteolytic Activity of Plasma from a Normal Individual and an Individual on Warfarin Therapy

Only P1 = R is shown as P1 = L, and P1 = D did not show any significant activity. Treatment with warfarin leads to a dramatic reduction of thrombin proteolytic activity, thereby inhibiting the signal transduction cascade leading to coagulation. The mean of six experiments is represented.

pyridine (4 ml), and trifluoroacetic acid anhydride (190 μl, 1.3 mmol, 2.6 eq) was added dropwise. The reaction mixture was stirred overnight, and the solvents were removed under reduced pressure. The residue was dissolved in 2 ml CH_2Cl_2, and 403 mg N-hydroxysuccinimide (3.5 mmol, 7.0 eq) and 477 mg 1-ethyl-3-(3′-dimethylaminopropyl)carbodiimide hydrochloride (2.5 mmol, 5.0 eq) were added. The reaction mixture was stirred for 35 min, then diluted with CH_2Cl_2 (100 m) and washed with H_2O (100 ml). The organic phase was dried with sodium sulfate, filtered, and concentrated under reduced pressure. Flash chromatography on silica gel (20 g) with a step gradient of hexane/ethyl (3:1 to 1:1) afforded the desired TFA-rhodamine-NHS ester **4** [27] (46 mg, 0.07 mmol, 14% over 2 steps).

Synthesis of the Rhodamine Resin 6

Rink amide resin (0.7 mmol/g) was loaded with orthogonally protected Fmoc-Lys(Mtt)-OH to a fixed loading of 0.4 mmol/g then capped with Ac_2O (5.0 eq) and 2,6-lut. (5.0 eq). A portion of this resin (900 mg, 0.36 mmol) was selectively Fmoc deprotected at the α terminus (20% piperidine in DMF), washed, and coupled to the rhodamine **4** (750 mg, 1.13 mmol, 3.1 eq) in the presence of HOBt (178 mg, 1.31 mmol, 3.1 eq) and iPr$_2$EtN (196 μl, 1.13 mmol, 3.1 eq) in DMF (4.5 ml) for 12 hr to obtain resin **5**. The completion of the reaction was verified by LC/MS analysis of the cleavage of a resin aliquot. The trifluoroacetyl protecting groups of the rhodamine were removed by treatment with concentrated aqueous ammonia (10 ml) for 4 hr to obtain polymer bound rhodamine **6**.

Synthesis of Single Substrates 7–12

Two batches of rhodamine resin **6** (250 mg suspended in 1 ml in DMF) were coupled with Fmoc-Asp(O-*t*-Bu)-OH or Fmoc-Arg(Pbf)-OH, respectively, using 4.0 eq of amino acid preactivated with HATU (3.5 eq), iPr$_2$EtN (4.0 eq), and lutidine (6.0 eq) in 1 ml DMF. The couplings were carried out for 12 hr and were repeated twice. Analy-

sis of the cleavage of a resin aliquot by LC/MS indicated a coupling yield of 86% for the Fmoc-Asp(O-*t*-Bu) and 61% for the Fmoc-Arg(Pbf). These yields were confirmed by Fmoc quantification. Any remaining free rhodamine amino functions were acetylated overnight with acetic acid, DIC, and 3-nitrotriazole (1 M each in DMF, 12 ml). The completion of the acetylation was verified by LC/MS analysis of the cleavage of a resin aliquot. Standard Fmoc deprotection (20% piperidine in DMF for 5 min) and amino acid coupling (HOBt, DIC, 5 eq each) were used for the subsequent coupling to yield $(nTPR)_2$-rhodamine and $(DEVD)_2$-rhodamine (where n represents norleucine). The overall coupling yields for the three couplings were >90%, as determined by Fmoc analysis. The N-terminal Fmoc was removed, and the peptide was capped with an acetate (AcOH, DIC, HOBt, 5 eq each). The Mtt was removed under mild acidic conditions (CH_2Cl_2, TFA, and TIS at 94:1:5, 2 ml, 4 × 2 min), and the resulting amine was free based by washing the resin with a 1% solution of iPr_2EtN in CH_2Cl_2 to obtain polymer bound compound **7** and **10**. A portion of each resin was acetylated and cleaved (TFA, water, and TIS, 95: 2.5: 2.5, 2 hr) to afford, after pelleting in Et_2O, compound **8** and **11**. The structure of these two compounds was verified by mass spectroscopy (MALDI-TOF, 1543.6 calcd for **7**, found: 1544.1 $[M+H]^+$, 1567.3 $[M+Na]^+$; 1561.8 calcd for **10**, found: 1562.9$[M+H]^+$). A portion of polymer bound compound **7** and **10** (10 mg) were inserted into a DNA synthesis cartridge and placed on an Applied Biosystems Expedite 8909 PNA synthesizer. The PNA sequence was programmed and carried out according to the manufacturer's recommendations. Cleavage of the resulting resin with TFA:*m*-cresol (4:1, 300 μl), followed by pelleting in Et_2O, afforded PNA-encoded compound **9** and **12**. Again, the structure of these two compounds was verified by mass spectroscopy (MALDI-TOF, 5596.4 calcd for **9**, found: 5595.5 $[M+H]^+$, 1567.3 $[M+Na]^+$; 5553.7calcd for **12**, found: 5553.3 $[M+H]^+$). Compounds **8**, **9**, **11**, and **12** were dissolved into H_2O:MeCN (9:1), frozen, lyophilized, and used for subsequent experiments without further purifications.

Synthesis of the 192 Member PNA-Encoded Protease Substrate Library

The library synthesis was performed by using the previously described general procedures for PNA oligomerization/deprotection and Alloc-protected amino acid coupling/deprotection. The library synthesis is summarized in Figure 4. Briefly, resin **5** was split into three batches, and the first residue was coupled to the rhodamine as described (vide supra) to obtain three pools of resin **15**. Considering the small number of pools in the first step, each of the three pools was directly split into four pools for introduction of the second residue (**14**), which is orthogonally protected with an Alloc (the amount of resin was calculated based on the efficiency of the first coupling). The Mtt was selectively removed, and after appropriate coupling to the polyetnylen spacers and lysine, the respective codons were introduced by reiterative PNA couplings to afford 12 pools of **16**. The pools of **16** were combined, the Alloc group was removed, and the resin was redistributed into four pools for third residue coupling and its PNA encoding, thus yielding four pools of PNA-encoded tripeptide **17**. Repeating the same procedure, followed by the addition of a last lysine residue to cap the encoding PNA afforded four pools of tetrapeptide **18**. The resin was combined and Alloc deprotected, and the peptide chain was capped with an acetyl group. The library was cleaved from the resin by using a solution of TFA:*m*-cresol:H_2O (80:19:1) for 1 hr and precipitated in Et_2O. The precipitates were pelleted, washed with Et_2O, resuspended in 9:1 H_2O:MeCN, and lyophilized.

Enzymatic Activity Monitored with Protease Substrates 1 and 3

Substrates **1** and **3** were used at 250 μM. Thrombin was used at a concentration of 500 pM, and caspase-3 was used at a concentration of 10 nM. For thrombin, a buffer consisting of 50 mM Tris (pH 7.4), 200 mM NaCl, 5 mM $CaCl_2$, and 0.01% (v/v) Tween-20 (THB) was used. The buffer for caspase-3 (CAB) consisted of 20 mM HEPES (pH 7.4), 100 mM NaCl, 1mM EDTA, 0.1% CHAPS, 10% (w/v) sucrose, and 10 mM DTT. A total of 50 μl of the buffer containing the substrate at a concentration of 500 μM was transferred into a well of a black 96-well Microfluor plate (Dynex Technologies). The reaction was initiated by the addition of 50 μl of the buffer containing the enzyme at a concentration of 1 nM for thrombin or 20 nM for caspase-3. The hydrolysis of rhodamine substrates was measured with a Spectramax Gemini XS spectrofluorimeter (Molecular Devices) thermostated at 37°C with an excitation wavelength of 490 nm, an emission wavelength of 530 nm, and a cutoff wavelength of 515 nm.

Single Substrate Kinetics

Thrombin was used at a final concentration of 1.25 nM, and caspase-3 was used at a final concentration of 50 nM. The final concentration of the substrates ranged from 1 μM to 150 μM; the final concentration of DMF in the assay was less than 5%. A total of 10 μl of the buffer containing the substrate was transferred into a well of a black 384-well plate with a clear bottom (Corning). The reaction was initiated by the addition of 10 μl of buffer containing the enzyme at a concentration of 2.5 nM for thrombin or 100 nM for caspase-3 and was monitored by fluorescence as described above.

Determination of the Substrate Concentration by Using Total Hydrolysis

Serial dilutions of the substrates were performed in the corresponding buffers, and thrombin or caspase-3 was added to a final concentration of 5 nM or 100 nM, respectively. The mixture was incubated overnight at room temperature. The endpoint fluorescence was measured as described above. The concentration of the rhodamine substrate for each dilution was determined by comparison of its fluorescence with the fluorescence of solutions containing known amounts of the rhodamine in the same buffer system.

Total hydrolysis of the 192 member library was performed in THB by sequential treatment with the following proteases: protease from *Streptomycis Griseus*, subtilisin *Carlsbad* and trypsin from bovine pancreas. Total hydrolysis was confirmed by MALDI-TOF analysis. The total concentration of the rhodamine library was determined by fluorescence measurement as described above.

Spatial Deconvolution of Single Protease Probes on Affymetrix Arrays

For single substrates, the Affymetrix Geneflex array was used for spatial deconvolution and fluorescence detection. Substrates **2** and **4** were diluted to a concentration range of 1–800 pM in CAB, and caspase-3 was added to a final concentration of 100 nM. The samples were incubated overnight at room temperature. The solutions were diluted with a modified PBS buffer, containing 250 mM NaCl (PBSS).

The Affymetrix GenFlex Arrays were hydrated by applying 180 μl CHB (Chip Hydration Buffer: 100 mM MES [pH 6.5], 1 M NaCl) to the chips, followed by incubation of the chips for 1 hr at 45°C in an Affymetrix hybridization oven. The CHB was removed, and the chips were washed two times with PBSS. A solution (6 μl) of Affymetrix GeneFlex control probes was added to 180 μl of the diluted substrate solutions, and the samples were applied to the GenFlex chips. The samples were hybridized for 4 hr at 45°C to the chips, and the sample solutions were removed. The chips were washed three times with 180 μl PBSS and filled with 180 μl PBSS. The chips were read on an Affymetrix chip reader with the standard argon ion laser as a light source and 530 nm as the detection wavelength. The average intensity of ten randomly picked border probes was used for normalization.

Limited Hydrolysis of the 192 Member PNA-Encoded Protease Substrate Library

The 192 member PNA-encoded library was diluted to a final concentration of 33 μM into 1 ml THB or CAB containing 3% (v/v) DMSO. Caspase-3, thrombin, and plasmin were used at a final concentration of 100, 2.5, and 10 nM, respectively. The mixture was incubated at room temperature, and the fluorescence was monitored over time until the desired percentage of hydrolysis (~3%–5%) was reached. The hydrolysis was monitored by fluorescence as described above. When the desired percentage of hydrolysis was reached, an aliquot of 200 μl was removed, and the enzymatic hydrolysis was quenched by adding 3 μl of a TFA/water (1:5) solution. After the collection of all samples, the solutions were diluted to a final concentration of

2 μM of total substrate into PBSS with 3% (v/v) DMSO (50 μl final volume) and centrifuged (20,000 × g, 4°C, 20 min).

Apoptotic versus Nonapoptotic Cell Lysates Substrate Screening

Whole Jurkat cells (10^7) were incubated for 4 hr with and without 100 ng/ml of a fas-activating antibody, CH-11 (Kaminya Biomedical Co.). The cells were washed twice with PBS, and cytosolic lysates were prepared by treating the cells with 250 μl of buffer containing 10 mM HEPES (pH 7.4), 130 mM NaCl, and 1% (v/v) Triton X-100. The soluble cytosolic fraction was separated from the insoluble membrane fraction by centrifugation.

For single substrates, 150 μl apoptotic or nonapoptotic cell lysate diluted to a protein concentration of 1 mg/ml was mixed with 150 μl CAB containing **2** and **4** (2 μM each) and incubated at 37°C. Aliquots of 40 μl were withdrawn after 0, 1, 2, 3, and 6 hr. A total of 25 μl of the aliquot was diluted into 75 μl PBS, and the endpoint fluorescence was measured as described above. A total of 10 μl of the aliquot was mixed with 1 μl of a TFA/water (1:1) solution, thereby quenching the enzymatic hydrolysis of the substrates. The quenched aliquots were placed on ice. After the collection of all time points, 990 μl PBSS was added to the quenched aliquots and mixed. The samples were centrifuged (20,000 × g, 20°C, 20 min), and the supernatant was applied to the printed oligonucleotide arrays.

For the 192 member library, 100 μl undiluted lysates was added to 100 μl CAB containing 6% (v/v) DMSO and 66 μM library. The lysates were incubated at 37°C until the desired hydrolysis (3%–5% for lysates) was obtained. An aliquot of 100 μl was removed, and the enzymatic hydrolysis was quenched by adding 1.5 μl of a TFA/water (1:5) solution The solutions were diluted to a final concentration of 2 μM library into PBSS with 3% (v/v) DMSO (50 μl final volume) and centrifuged (20,000 × g, 4°C, 20 min). The supernatants were applied to the printed oligonucleotide arrays.

Blood Coagulation Substrate Screening

Fresh frozen plasma samples from pooled normal individuals and an individual on warfarin therapy (George King Bio-medical) were briefly thawed in a 37°C water bath. A total of 10 μl of plasma was combined with 10 μl activated partial thromboplastin reagent (Biomérieux) to initiate the blood coagulation cascade. After 5 min of incubation at 37°C, 180 μl THB containing 6% (v/v) DMSO and 66 μM of the substrate library were added to the sample. The samples were incubated at 37°C until the desired hydrolysis (3%–5%) was obtained. An aliquot of 20 μl was removed and added to 300 μl PBS with 3% DMSO and 0.2 μg/ml BSA and centrifuged (20,000 × g, 4°C, 20 min). The supernatants were applied to the printed oligonucleotide arrays.

Printing of Oligonucleotide Arrays

The 78-mer oligonucleotides printed had the general design $CTCGTN_{13}GGCGTN_{13}GG\ CGTN_{13}GGCCGTN_{13}GGTCC$, where N stands for the complementary sequence to its target PNA. The oligonucleotides were dissolved in 3× SSC at a concentration of 1 mg/ml oligonucleotide and printed in duplicate on Ultra Gap Slides slides from Corning with an Omni Grid Accent contact printer (GeneMachnines) at a spacing of 200 μm. The printing was performed at 22°C and 75% humidity by using an SMP3 Stealth pin from TeleChem. The slides were rehydrated for ca. 10 s over 60°C warm water, snap dried for 20 s on a hot plate at 80°C, and UV-crosslinked with a Stratalinker 2400 (Stratagene) at a total energy of 600 mJ.

Postprocessing of Printed Oligonucleotide Arrays and Spatial Deconvolution of Single Protease Probes and the 192 Member PNA-Encoded Library on Printed Oligonucleotide Arrays

The slides were submerged for 30 min into stirred 50°C solution (3× SSC) containing 0.01% SDS, 0.2 mg/ml BSA, dip rinsed ten times in 250 ml nanopure water. The slides were dip rinsed ten times in 2-propanol and blown dry with nitrogen. The slides were placed into a slide holder that resembled a 384-well microtiter plate. 50 μl sample was applied to each well, and the holder was closed with a tight-sealing lid. After incubation for 1 hr at 37°C, the slides were dip rinsed in 0.05× SSC (50 ml, 3×), centrifuged (1500 × g, 2 min, RT), and scanned on an Applied Precision 4500 scanner with the A488 filter set and an exposure time of 1.0 s. ImaGene 4.2 software (BioDiscovery) was used for data analysis.

Acknowledgments

Human Frontier Science Program is gratefully acknowledged for their financial support (HFSP 0080/2003). We thank Bradley Backes and James Schmeits for advice and Renee Link for technical assistance. We thank Affymetrix for their generous donation of GenFlex tag arrays.

Received: May 7, 2004
Revised: July 20, 2004
Accepted: July 29, 2004
Published: October 15, 2004

References

1. Rawlings, N.D., Tolle, D.P., and Barrett, A.J. (2004). MEROPS: the peptidase database. Nucleic Acids Res. *32*, D160–D164.
2. Barret. A.J., Rawlings, N.D., and Woessner, J.F. (1998). Handbook of Proteolytic Enzymes (London: Academic Press).
3. Krausslich, H.G., Ingraham, R.H., Skoog, M.T., Wimmer, E., Pallai, P.V., and Carter, C.A. (1989). Activity of purified biosynthetic proteinase of human immunodeficiency virus on natural substrates and synthetic peptides. Proc. Natl. Acad. Sci. USA *86*, 807–811.
4. Davie, E.W., Fujikawa, K., and Kisiel, W. (1991). The coagulation cascade: initiation, maintenance, and regulation. Biochemistry *30*, 10363–10370.
5. 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., et al. (1997). A combinatorial approach defines specificities of members of the caspase family and granzyme B. Functional relationships established for key mediators of apoptosis. J. Biol. Chem. *272*, 17907–17911.
6. Backes, B.J., Harris, J.L., Leonetti, F., Craik, C.S., and Ellman, J.A. (2000). Synthesis of positional-scanning libraries of fluorogenic peptide substrates to define the extended substrate specificity of plasmin and thrombin. Nat. Biotechnol. *18*, 187–193.
7. Harris, J.L., Backes, B.J., Leonetti, F., Mahrus, S., Ellman, J.A., and Craik, C.S. (2000). Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries. Proc. Natl. Acad. Sci. USA *97*, 7754–7759.
8. Meldal, M., Svendsen, I., Breddam, K., and Auzanneau, F.-I. (1994). Portion-mixing peptide libraries of quenched fluorogenic substrates for complete subsite mapping of endoprotease specificity. Proc. Natl. Acad. Sci. USA *91*, 3314–3318.
9. Salisbury, C.M., Maly, D.J., and Ellman, J.A. (2002). Peptide microarrays for the determination of protease substrate specificity. J. Am. Chem. Soc. *124*, 14868–14870.
10. Reimer, U., Reineke, U., and Schneider-Mergener, J. (2002). Peptide arrays: from macro to micro. Curr. Opin. Biotechnol. *13*, 315–320.
11. St. Hilaire, P.M., Willert, M., Juliano, M.A., Juliano, L., and Meldal, M. (1999). Fluorescence-quenched solid phase combinatorial libraries in the characterization of cysteine protease substrate specificity. J. Comb. Chem. *1*, 509–523.
12. Harris, J., Mason, D., Burdick, K.W., Backes, B., Chen, T., Van Heeke, G., Gough, L., Ghaemmaghami, A., Shakib, F., Debaene, F., et al. (2004). Activity profile of dust mite allergen extract using substrate libraries and functional proteomic microarrays. Chem. Biol. *11*, this issue, 1361–1372.
13. Winssinger, N., Ficarro, S., Schultz, P.G., and Harris, J.L. (2002). Profiling protein function with small molecule microarrays. Proc. Natl. Acad. Sci. USA *99*, 11139–11144.
14. Leytus, S.P., Patterson, W.L., and Mangel, W.F. (1983). New class of sensitive and selective fluorogenic substrates for serine proteinases. Amino acid and dipeptide derivatives of rhodamine. Biochem. J. *215*, 253–260.
15. Assfalg-Machleidt, I., Rothe, G., Klingel, S., Banati, R., Mangel, W.F., Valet, G., and Machleidt, W. (1992). Membrane permeable

fluorogenic rhodamine substrates for selective determination of cathepsin L. Biol. Chem. Hoppe Seyler *373*, 433–440.
16. Leytus, S.P., Melhado, L.L., and Mangel, W.F. (1983). Rhodamine-based compounds as fluorogenic substrates for serine proteinases. Biochem. J. *209*, 299–307.
17. Cai, S.X., Zhang, H.Z., Guastella, J., Drewe, J., Yang, W., and Weber, E. (2001). Design and synthesis of rhodamine 110 derivative and caspase-3 substrate for enzyme and cell-based fluorescent assay. Bioorg. Med. Chem. Lett. *11*, 39–42.
18. Winssinger, N., Harris, J.L., Backes, B.J., and Schultz, P.G. (2001). From split-pool libraries to spatially addressable microarrays and its application to functional proteomic profiling. Angew. Chem. Int. Ed. Engl. *40*, 3152–3155.
19. Ratilainen, T., Holmen, A., Tuite, E., Haaima, G., Christensen, L., Nielsen, P.E., and Norden, B. (1998). Hybridization of peptide nucleic acid. Biochemistry *37*, 12331–12342.
20. Furka, A., Sebestyen, F., Asgedom, M., and Dibo, G. (1991). General method for rapid synthesis of multicomponent peptide mixtures. Int. J. Pept. Protein Res. *37*, 487–493.
21. Debaene, F., Mejias, L., Harris, J.L., and Winssinger, N. (2004). Synthesis of a PNA-encoded cysteine protease inhibitor library. Tetrahedron, *60*, 8677–8690.
22. Jeffery, D.A., and Bogyo, M. (2003). Chemical proteomics and its application to drug discovery. Curr. Opin. Biotechnol. *14*, 87–95.
23. Jessani, N., and Cravatt, B.F. (2004). The development and application of methods for activity-based protein profiling. Curr. Opin. Chem. Biol. *8*, 54–59.
24. Furie, B., and Furie, B.C. (1988). The molecular basis of blood coagulation. Cell *53*, 505–518.
25. Zhou, Q., Snipas, S., Orth, K., Muzio, M., Dixit, V.M., and Salvesen, G.S. (1997). Target protease specificity of the viral serpin CrmA. Analysis of five caspases. J. Biol. Chem. *272*, 7797–7800.
26. Ioffe, I.S., and Otten, V.F. (1962). Rhodamine dyes and related compounds. III. Reaction of m-aminophenol with phthalic anhydride in hot sulfuric acid. Zh. Obshch. Khimii *32*, 1477–1480.
27. Czerwinski, G., Wank, S.A., Tarasova, N.I., Hudson, E.A., Resau, J.H., and Michejda, C.J. (1995). Synthesis and properties of three fluorescent derivatives of gastrin. Lett. Pept. Sci. *1*, 235–242.