J|A|C|S
COMMUNICATIONS

Published on Web 11/22/2002

Peptide Microarrays for the Determination of Protease Substrate Specificity

Cleo M. Salisbury, Dustin J. Maly, and Jonathan A. Ellman*

Center for New Directions in Organic Synthesis, Department of Chemistry, University of California, Berkeley, California 94720

Received June 26, 2002

Proteases, which hydrolyze amide bonds in peptides and proteins, are abundant in nature and essential for cellular function and viability. Many important biological pathways, such as hormone activation and apoptosis, are dependent on the action of proteases. Furthermore, proteases are involved in diverse disease states, including cardiovascular disease, cancer, AIDS, and Alzheimer's disease, making these enzymes important therapeutic targets.[1] Significantly, many proteases have recently been identified through genome sequencing efforts. Only a fraction of the estimated 2% of genes that encode for proteases in organisms from *E. coli* to humans have been studied to date.[2] Establishing the functional roles of these proteases will greatly enhance our understanding of biological systems and will likely provide a number of important new therapeutic targets.

The ability of proteases to selectively cleave specific substrates in the presence of many other potential substrates is essential for their function. Knowledge of the substrate specificity of a protease can therefore be used to facilitate the identification of its physiological substrates, which is essential for defining its biological function. Moreover, determination of substrate specificity greatly aids in the design of potent and selective substrates and inhibitors. Although a number of combinatorial methods have been described for the rapid determination of substrate specificity,[3] they do not provide a full specificity fingerprint because they do not allow for kinetic evaluation of all of the prepared substrate sequences. Herein we report the synthesis and evaluation of microarrays of fluorogenic substrates that enable the rapid and complete characterization of each substrate in the library.

We have recently reported the development of the fluorogenic peptidyl coumarin substrates, 7-amino-4-carbamoylmethyl coumarin (ACC) peptides, which are effective tools for determining the P-side[4] (N-terminal) substrate specificity of serine and cysteine proteases.[5] Proteolytic cleavage at the peptide anilide bond liberates the highly fluorescent coumarin leaving group, allowing for the determination of cleavage rates through the increase in fluorescence. Importantly, the bifunctional nature of ACC enables the fluorophore to be linked to solid support through its C-terminus, and thus libraries of fluorogenic substrates of any amino acid sequence can be rapidly prepared using standard solid-phase peptide synthesis methods. Positional scanning-substrate combinatorial libraries (PS-SCLs) of peptidyl coumarins have proven to be very useful for determining substrate specificity trends;[3a−c,5] however, the mixture-based nature of these libraries does not allow for cooperative interactions between substrate residues to be determined.[3n] To fully map the N-terminal specificity of a protease, we have applied libraries of peptidyl coumarins to microarrays in a spatially addressed fashion. The peptides are linked to glass slide microarray surfaces via a chemoselective oxime forming reaction,[6] using an aldehyde-derivatized surface and alkoxylamine-functionalized substrates (Figure 1). These alkoxylamine-derivatized peptide substrates show kinetic constants comparable to those of the underivatized ACC substrates in solution (see Supporting Information).

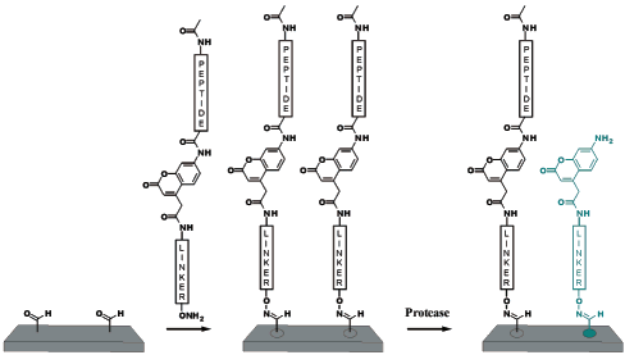


Figure 1. Fluorogenic substrate microarrays.

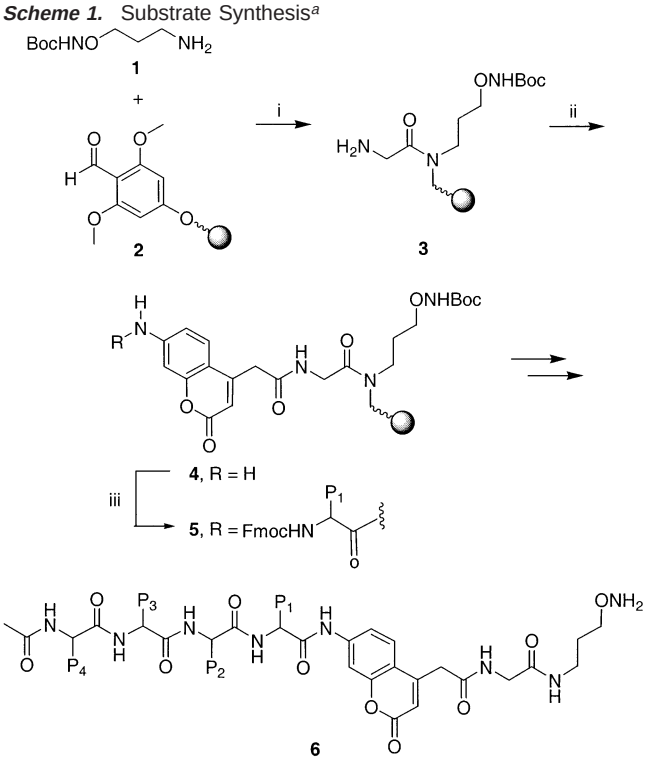
The alkoxylamine-substituted fluorogenic substrates are prepared by first loading the latent alkoxylamine functionality onto solid support by reductive amination of monoprotected alkoxylamine **1** with the backbone amide linker (BAL) aldehyde resin **2** (Scheme 1).[7] The resultant secondary amine is then acylated with Fmoc-Gly-OH activated with HATU. After Fmoc deprotection to provide **3**, Fmoc-protected ACC is then added using standard coupling conditions. The Fmoc protecting group is removed to yield intermediate **4**, and the resulting poorly nucleophilic free aniline is acylated using the symmetric anhydride of the desired P_1 amino acid residue to yield intermediate **5**. Any remaining unreacted aniline is then capped with the nitrotriazole ester of acetic acid. These conditions were found to give efficient loading of the P_1 residue with low background fluorescence due to the efficient capping of the unreacted coumarin. After the loading of the P_1 residue, the desired peptide is constructed using standard Fmoc-based solid-phase synthesis techniques. The desired alkoxylamine fluorogenic peptide substrate **6** is cleaved from solid support using trifluoroacetic acid.

A number of conditions were tested for oxime formation on aldehyde-derivatized glass microscope slides. A concentration of 200−500 μM peptide in a pH 5.2 acetate buffer with 1−5% dimethyl sulfoxide (DMSO) provided the best results. Higher concentrations of peptide and DMSO were found to cause spot spreading, and other buffer combinations did not lead to efficient oxime formation. Fluorogenic ACC-peptides were first printed on commercially available aldehyde-functionalized slides, but were inefficiently cleaved by trypsin. A layer of bovine serum albumin (BSA) has been shown to increase the accessibility of slide-bound peptides to enzymes.[8] We therefore chose to derivatize BSA-covered slides with 4-carboxybenzaldehyde, using standard carbodiimide coupling conditions. Substrates bound to the BSA-aldehyde surface

* To whom correspondence should be addressed. E-mail: jellman@uclink.berkeley.edu.

10.1021/ja027477q CCC: $22.00

Scheme 1. Substrate Synthesis[a]



[a] (i) (a) $NaBH_3CN$, DMF; (b) Fmoc-Gly-OH, DMF, HATU, collidine; (c) 20% piperidine in DMF. (ii) (a) Fmoc-ACC-OH, DIC, HOBt, DMF; (b) AcOH, DIC, HOBt, DMF; (c) 20% piperidine in DMF. (iii) (a) Symmetric anhydride of P_1 amino acid, DMF/dichloroethane (DCE); (b) AcOH, DIC, DMF, nitrotriazole.

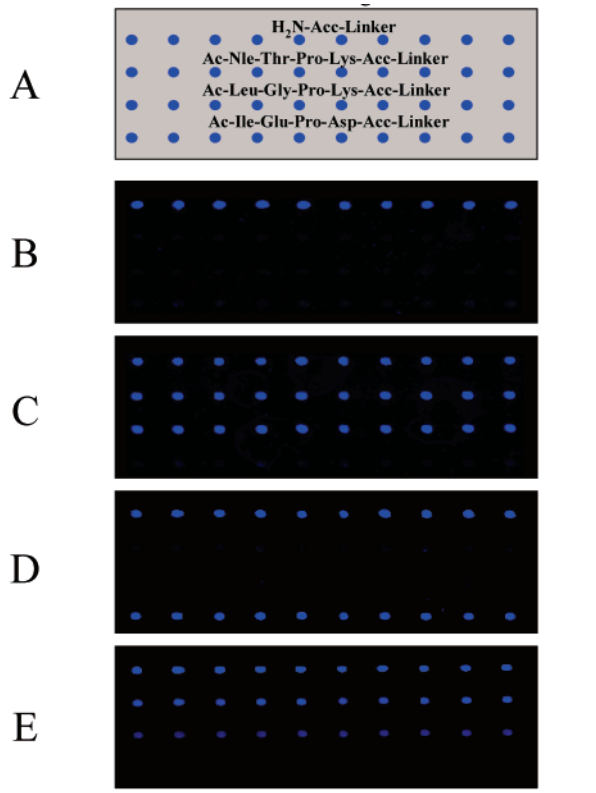


Figure 2. Proteolytic cleavage of array-bound substrates. Panel A shows a schematic of the array. Panels B−E show the fluorescence of the array prior to cleavage (B), and after treatment with trypsin (C), granzyme B (D), and thrombin (E).

demonstrated an enhanced rate of cleavage by trypsin, reflecting the greater accessibility of the substrates presented on this surface and/or the increased stability of the enzyme.

To further demonstrate that array-bound fluorogenic substrates are accessible for proteolytic cleavage, substrate arrays were subjected to proteolysis by a variety of serine proteases. A split pin arrayer was used to spot multiple copies of three different protease substrates on the array surface at a spacing of 314 μm (Figure 2). Specifically, two thrombin substrates with differing specificity constants (k_{cat}/K_m), Ac-Leu-Gly-Pro-Lys-ACC-linker and Ac-Nle-Thr-Pro-Lys-ACC-linker, and a known granzyme B substrate, Ac-Ile-Glu-Pro-Asp-ACC-linker, were printed. To serve as a fluorescence intensity standard, unacylated ACC-linker was also printed. The general accessibility of the array to enzymatic reactions was first demonstrated by exposing the array to trypsin (Figure 2c). As expected, the two thrombin substrates with a P_1-Lys residue showed a comparable rate of cleavage, whereas the granzyme B substrate with a P_1-Asp residue showed no detectable cleavage. Next, the array surface was subjected to the serine protease granzyme B (Figure 2d). The array showed cleavage of the granzyme B substrate with no detectable cleavage of either of the thrombin substrates, demonstrating that enzymes maintain fidelity for their desired substrate when affixed to the surface. To demonstrate that subtle differences in extended substrate binding by a protease could be detected, the array was subjected to cleavage with the serine protease thrombin. The two thrombin substrates were readily cleaved, whereas the granzyme B substrate showed no detectable cleavage. As can be seen from Figure 2e, a significant difference in the extent of cleavage of the two thrombin-susceptible substrates could be detected by the greater increase in fluorescence for the preferred substrate (Ac-Nle-Thr-Pro-Lys-ACC-linker) over the less preferred substrate (Ac-Leu-Gly-Pro-Lys-ACC-linker). These fluorescence differences are readily detected by standard microarray scanning instrumentation. As little as 1% cleaved substrate (unacylated ACC) can be detected, and the amount of unacylated ACC is linearly proportional to the fluorescence (see Supporting Information).

A library of substrates was next synthesized, printed, and assayed to demonstrate the utility of peptide-ACC microarrays for the rapid determination of protease substrate specificity. The 361-member spatially separate library contained a fixed P_1-Lys and P_4-Ala with all combinations of proteinogenic amino acids (except cysteine) at the P_2 and P_3 sites. To serve as standards for quantitation of fluorescence, multiple copies of unacylated ACC and acetyl-capped ACC were also printed on the array surface. Using a spacing of 440 μm, we printed the entire library and the control compounds in duplicate in a 1.8 cm $\times$ 0.9 cm area, representing 800 individual data points in an area of less than 1.7 cm^2. The array was then used to determine the extended substrate specificity of the serine protease thrombin by adding 15 μL of a 250 nM solution of the enzyme to the array surface followed by incubation for 60 min. The quantitative results of the assay are shown in Figure 3. The data show a strong preference for proline at the P_2 position and broad specificity at the P_3 position with basic, polar, and hydrophobic residues all being tolerated. This specificity profile is consistent with prior studies using PS-SCLs and previously known specificity determinants.[9]

To validate our results, we synthesized, purified, and assayed four of these peptidyl coumarin substrates in solution (Table 1). As the data show, the k_{cat}/K_m for the substrates on the array is predictive of the solution-phase results.[10] Of note is that considerably less enzyme was needed to determine this specificity profile on the array relative to prior studies, due to the miniaturized nature of the array and the increased sensitivity of fluorescence detection

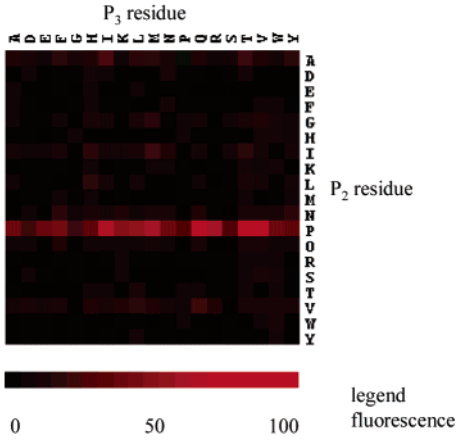


Figure 3. Characterization of the substrate specificity of thrombin using substrates of the general structure Ac-Ala-P_3-P_2-Lys-linker. Each square is colored in proportion to the quantitated fluorescence intensity of the corresponding substrate after treatment with thrombin, corresponding to the relative amount of cleavage (normalized data and 3D bar graph available in Supporting Information).

Table 1. Comparison of the Relative k_{cat}/K_m for Purified Underivatized ACC-Peptides in Solution and the Corresponding Library Substrates on the Array, upon Treatment with Thrombin[a]

substrate	solution phase k_{cat}/K_m	microarray k_{cat}/K_m
Ac-ATPK-ACC	1.00	1.00
Ac-AGPK-ACC	0.27	0.26
Ac-ADAK-ACC	0.02	0.09
Ac-AFSK-ACC	0.02	0.00

[a] In both cases, the data have been normalized to the Ac-ATPK-ACC value.

that results from the high substrate density on the array surface.[11] The minimal enzyme requirements allow for the determination of substrate specificity of proteases that are available in only limited quantities.

In conclusion, we have described an approach to obtain a complete analysis of protease specificity using microarrays of peptidyl coumarin substrates. The ACC substrates are uniquely suited to this application because, unlike fluorescence resonance energy transfer (FRET)-based substrates, ACC peptides do not require further analysis to determine site of cleavage.[3a] The arrays were constructed with standard DNA microarraying equipment, using aldehyde-derivatized glass slides and alkoxylamine-functionalized peptidyl coumarin substrates. The fidelity of proteolytic cleavage of array-bound substrates was demonstrated using a variety of proteases. The speed and efficiency of the methodology were demonstrated by obtaining a proteolytic "fingerprint" of thrombin in a single experiment. This method should enable fingerprints of proteolytic activity to be rapidly obtained for numerous proteases, because thousands of substrates can be printed on a single slide and large numbers of slides can be readily printed using microarray automation. Furthermore, due to the sensitivity and efficiency of this assay, substrate cleavage profiles obtained by this methodology could potentially be used to detect differential proteolytic profiles in cell lysates and, thus, could have diagnostic applications.

Acknowledgment. We thank Dr. Adam Carroll for assistance with the microarray printing and insightful discussions, Patty Holman and Trent Su for the assistance with the microarray scanning, and Sandra M. Waugh for the gift of recombinant granzyme B. This work was supported by the NIH (grant # GM54051). The Center for New Directions in Organic Synthesis is supported by Bristol-Myers Squibb as a Sponsoring Member and Novartis as a Supporting Member.

Supporting Information Available: Complete protocols for single substrate and library syntheses, slide preparation, microarray printing, enzyme assays, and kinetic analyses of microarrays and single substrates, and microarray controls (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Leung, D.; Abbenante, G.; Fairlie, D. P. *J. Med. Chem.* **2000**, *43*, 305.

(2) Rawlings, N. D.; O'Brien, E.; Barrett, A. J. *Nucleic Acids Res.* **2002**, *30*, 343.

(3) (a) For a recent review, see: Maly, D. J.; Huang, L.; Ellman, J. A. *ChemBioChem* **2002**, *3*, 16. (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. (c) Harris, J. L.; Niles, A.; Burdick, K.; Maffitt, M.; Backes, B. J.; Ellman, J. A.; Kuntz, I.; Haak-Frendscho, M.; Craik, C. S. *J. Biol. Chem.* **2001**, *276*, 34941. (d) Matthews, D. J.; Wells, J. A. *Science* **1993**, *260*, 1113. (e) Smith, M. M.; Shi, L.; Navre, M. *J. Biol. Chem.* **1995**, *270*, 6440. (f) Turk, B. E.; Huang, L. L.; Piro, E. T.; Cantley, L. C. *Nat. Biotechnol.* **2001**, *19*, 661. (g) Petithory, J. R.; Masiarz, F. R.; Kirsch, J. F.; Santi, D. V.; Malcolm, B. A. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 11510. (h) Birkett, A. J.; Soler, D. F.; Wolz, R. L.; Bond, J. S.; Wiseman, J.; Berman, J.; Harris, R. B. *Anal. Biochem.* **1991**, *196*, 137. (i) Meldal, M.; Svendsen, I.; Breddam, K.; Auzanneau, F.-I. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3314. (j) St. Hilaire, P. M.; Willert, M.; Juliano, M. A.; Juliano, L.; Meldal, M. *J. Comb. Chem.* **1999**, *1*, 509. (k) Rosse, G.; Kueng, E.; Page, M. G.; Schauer-Vukasinovic, V.; Giller, T.; Lahm, H.-W.; Hunziker, P.; Schlatter, D. *J. Comb. Chem.* **2000**, *2*, 461. (l) Singh, J.; Allen, M. P.; Ator, M. A.; Gainor, J. A.; Whipple, D. A.; Solowiej, J. E.; Treasurywala, A. M.; Morgan, B. A.; Gordon, T. D.; Upson, D. A. *J. Med. Chem.* **1995**, *38*, 217. (m) Dekker, N.; Cox, R. C.; Kramer, R. A.; Egmond, M. R. *Biochemistry* **2001**, *40*, 1694. (n) Takeuchi, T.; Harris, J. L.; Huang, W.; Yan, K. W.; Coughlin, S. R.; Craik, C. S. *J. Biol. Chem* **2000**, *275*, 34, 26333.

(4) In the standard notation, substrate residues are referred to as P_n, P_{n-1}...P_2, P_1, $P_{1'}$, $P_{2'}$...$P_{m-1'}$, P_m, where amide bond hydrolysis occurs between P_1 and $P_{1'}$. Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157.

(5) (a) Harris, J. L.; Backes, B. J.; Leonetti, F.; Mahrus, S.; Ellman, J. A.; Craik, C. S. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 7754. (b) Maly, D. J.; Leonetti, F.; Backes, B. J.; Dauber, D. S.; Harris, J. L.; Craik, C. S.; Ellman, J. A. *J. Org. Chem.* **2002**, *67*, 910.

(6) (a) Falsey, J. R.; Renil, M.; Park, S.; Li, S.; Lam, K. S. *Bioconjugate Chem.* **2001**, *12*, 346. (b) Cornish, V. W.; Hahn, K. M.; Schultz, P. G. *J. Am. Chem. Soc.* **1996**, *118*, 8150. (c) Shao, J.; Tam, J. P. *J. Am. Chem. Soc.* **1995**, *117*, 3893. (d) Rodriguez, E. C.; Marcaurelle, L. A.; Bertozzi, C. R. *J. Org. Chem.* **1998**, *63*, 7134. (e) Rose, K. *J. Am. Chem. Soc.* **1994**, *116*, 30.

(7) Jensen, K. J.; Alsina, J.; Songster, M. F.; Vagner, J.; Albericio, F.; Barany, G. *J. Am. Chem. Soc.* **1998**, *120*, 5441.

(8) MacBeath, G.; Schreiber, S. L. *Science* **2000**, *289*, 1760.

(9) (a) Edwards, P. D.; Mauger, R. C.; Cottrell, K.; Morris, F. X.; Pine, K. K.; Sylvester, M. A.; Scott, C. W.; Furlong, S. T. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 2291. (b) Backes, B. J.; Harris, J. L.; Leonetti, F.; Craik, C. S.; Ellman, J. A. *Nat. Biotechnol.* **2000**, *18*, 187.

(10) The assay is run under conditions where the enzyme and substrate concentrations are well below the K_m, and, therefore, the relative rates of hydrolysis reflect k_{cat}/K_m ratios. See: Fersht, A. R. *Enzyme Structure and Mechanism*, 3rd ed.; Freeman: New York, 1985; pp 105−106.

(11) Ekins, R. P.; Berger, H.; Chu, F. W.; Finckh, P.; Krause, F. *Nanobiology* **1998**, *4*, 197.

JA027477Q