Synthesis of positional-scanning libraries of fluorogenic peptide substrates to define the extended substrate specificity of plasmin and thrombin

Bradley J. Backes[1,2†], Jennifer L. Harris[3†], Francesco Leonetti[1], Charles S. Craik[3*], and Jonathan A. Ellman[1*]

*[1]Chemistry Department, University of California Berkeley, Berkeley, CA 94720. [2]Current Address: Chemistry Department, Genomics Institute of the Novartis Research Foundation, 3015 Merryfield Row, San Diego, CA 92121. [3]Department of Pharmaceutical Chemistry, Program in Chemistry and Chemical Biology, University of California San Francisco, San Francisco, CA 94143. [†]These authors contributed equally to this work. *Corresponding authors.*

We have developed a strategy for the synthesis of positional-scanning synthetic combinatorial libraries (PS-SCL) that does not depend on the identity of the P1 substituent. To demonstrate the strategy, we synthesized a tetrapeptide positional library in which the P1 amino acid is held constant as a lysine and the P4-P3-P2 positions are positionally randomized. The 6,859 members of the library were synthesized on solid support with an alkane sulfonamide linker, and then displaced from the solid support by condensation with a fluorogenic 7-amino-4-methylcoumarin-derivatized lysine. This library was used to determine the extended substrate specificities of two trypsin-like enzymes, plasmin and thrombin, which are involved in the blood coagulation pathway. The optimal P4 to P2 substrate specificity for plasmin was P4-Lys/Nle (norleucine)/Val/Ile/Phe, P3-Xaa, and P2-Tyr/Phe/Trp. This cleavage sequence has recently been identified in some of plasmin's physiological substrates. The optimal P4 to P2 extended substrate sequence determined for thrombin was P4-Nle/Leu/Ile/Phe/Val, P3-Xaa, and P2-Pro, a sequence found in many of the physiological substrates of thrombin. Single-substrate kinetic analysis of plasmin and thrombin was used to validate the substrate preferences resulting from the PS-SCL. By three-dimensional structural modeling of the substrates into the active sites of plasmin and thrombin, we identified potential determinants of the defined substrate specificity. This method is amenable to the incorporation of diverse substituents at the P1 position for exploring molecular recognition elements in proteolytic enzymes.

Keywords: Serine protease, proteinase, blood coagulation, molecular recognition, combinatorial libraries

Proteases play essential roles in numerous biological processes. Substrate specificity, or the ability to discriminate among many potential substrates, is central to the function of proteases. Knowledge of a protease's substrate specificity may not only give valuable insights into its biological function but also provide the basis for potent substrate and inhibitor design. Synthetic substrates are typically used to define substrate specificity. However, the synthesis and assay of single substrates is tedious for proteases with specificity beyond P1 and often results in a limited substrate specificity profile.

Combinatorial approaches have recently been used to address the identification of several substrate recognition sites in proteases. All of these combinatorial methods involve the generation of libraries of substrates, proteolysis of the substrates, and identification of the optimal substrate sequence. Substrate libraries can be broken down into two categories: those that are biologically generated and those that are synthetically generated. Biological library methods include the display of peptide libraries on filamentous phage[1,2], the randomization of amino acids at physiological cleavage sites[3], and the identification of macromolecular cleavage sites of in vitro transcription/translation cDNA libraries[4,5]. The diversity of these libraries is often constrained by the transformation efficiency of the host organism and can only contain naturally occurring amino acids. This limitation can be circumvented by use of synthetic substrate libraries. Although combinatorial synthesis allows for the creation of millions of compounds, these methods are only useful when coupled with powerful analytical assays that allow for the identification of the preferred substrate. Discontinuous analysis of the cleavage products through Edman degradation[6,7], mass spectrometry[8,9], and chromatography[10] has proved useful for qualitative assessment of optimal substrates from soluble or support-bound peptides.

Substrate consensus sequences have also been obtained using support-bound fluorescence-quenched substrate libraries prepared by the process of split-synthesis, which results in single substrate sequences on each of the resin beads[11]. Partial proteolysis of the support-bound libraries and subsequent sequence determination of the substrates on the most fluorescent beads provides the consensus sequences[12,13]. Unfortunately, this method suffers from two major limitations: first, the kinetics of support-bound substrates can differ greatly from soluble substrates [14,15], and second, as with the other methods previously mentioned, identification of the substrate occurs after the cleavage event, making the kinetic analysis more cumbersome. A method that avoids these limitations, and gives a quantitative assessment of protease substrate preference, is the use of positional-scanning synthetic combinatorial libraries (PS-SCL)[16].

Positional-scanning [17] synthetic combinatorial libraries (PS-SCL) of fluorogenic peptide substrates are potentially powerful tools for determining protease specificity. In contrast to other combinatorial libraries, this library format provides rapid and continuous information on each of the varied substituents in the substrate. A

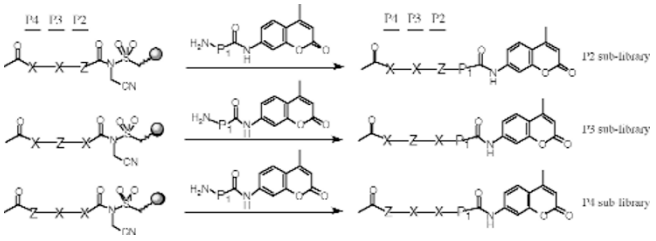

Figure 1. Three sublibraries (P2, P3, P4), each made up of 19 wells containing 361 compounds, are prepared by a segment condensation reaction with a P1 fluorogenic amino acid substrate. Individual proteinogenic amino acids are used to incorporate spatially addressed positions "Z," whereas an isokinetic mixture of proteinogenic amino acids is used to incorporate varied positions "X."

positional-scanning library with the general structure Ac-X-X-X-Asp-AMC was prepared previously by Rano et al. to rapidly and accurately assess the P4-P2 specificity for caspases that require aspartic acid (Asp) in the P1 position[18,19]. (Nomenclature for the substrate amino acid preference is Pn, Pn-1, . . .P2, P1, P1′, P2′, . . . , Pm-1′, Pm′. Amide bond hydrolysis occurs between P1 and P1′. Sn, Sn-1, . . . , S2, S1, S1′, S2′, . . . , Sm-1′, Sm′ denotes the corresponding enzyme binding sites[20]). Specific cleavage of the amide bond after the Asp residue liberates a fluorescent 7-amino-4-methylcoumarin (AMC) leaving group, thus allowing for the simple determination of cleavage rates for a library of substrates. The P1 Asp-coumarin substrate was conveniently linked to an insoluble polymer through the Asp carboxylic acid side chain, which allowed for library synthesis by standard peptide synthesis techniques. However, the method used to synthesize the library was specific for an aspartic acid at P1. The employment of strategies to link P1 amino acid-coumarin derivatives through side chain functionality may prove viable for some residues. However, linkage through hydrophobic side chain functionalities (leucine, phenylalanine, valine, etc.) will prove difficult, limiting the use of this strategy.

By developing a general strategy to incorporate all 20 proteinogenic amino acids at the P1 position of a PS-SCL, the extended specificity of virtually any protease could be rapidly determined. With this aim in mind, we have developed a general method for the preparation and screening of positional-scanning synthetic combinatorial substrate libraries. This design is free from the limitations of the previous approach, which requires linkage to the solid support through the side chain of the P1 substituent and allows for complete randomization at the P1 position.The resulting enzymatic profiles from the library resemble the known physiological cleavage sites of these enzymes and were verified by single-substrate kinetic analysis. Potential substrate recognition determinants on the enzymes were identified through the three-dimensional modeling of the substrates in the active sites of the enzymes.

Results

Library design and synthesis. To incorporate diversity at the P1 position, we condensed fluorogenic AMC P1-amino acid derivatives with a support-bound PS-SCL to provide library compounds in solution (Fig. 1). Three support-bound sublibraries were prepared (P2, P3, P4) using an alkane sulfonamide linker[21] and solid-phase peptide synthesis. The properties of the linker allow for the incorporation of a fluorogenic leaving group through the nucleophilic addition of an AMC-derivatized amino acid. Each sublibrary consisted of 19 resins (one unnatural amino acid, norleucine, was included; cysteine and methionine were excluded) for which a single position was spatially addressed by the coupling of a single amino acid. The two remaining positions of each resin were supplied by the coupling of isokinetic mixtures[22] to give a resin-bound mixture of 361 amino acids. The 57 resins comprising the entire PS-SCL were put in individual wells and cleaved from the resin with a P1-amino acid-coumarin derivative. Filtration, side chain deprotection, and concentration provided a PS-SCL of 57 wells containing 361 tetrapeptide-coumarin derivatives (total of 6,859 peptides). Thus, analysis of the three libraries identifies the enzyme's preferences for amino acids at P4, P3, and P2. A P1-Lys library was prepared and used to elucidate the specificity of plasmin and thrombin.

Profiling of plasmin with the positional scanning P1-Lys library. The preferred tetrapeptide substrate recognition sequence for plasmin was determined to be P4-Lys, P3-Xaa (a nonspecific amino acid), P2-Tyr/Phe/Trp, and P1-Lys (Fig. 2A). To validate the results from the PS-SCL and to quantitate dependence and use of extended interactions, kinetic parameters were determined for several single AMC substrates (Table 1). As indicated from the library, the majority of plasmin's extended substrate specificity resides in P2 and P4. Hydrolysis of the suboptimal P2 substrate, Ac-Lys-Thr-**Ser**-Lys-AMC, is up to 340% disfavored when compared to substrates that possess an aromatic amino acid at P2: Ac-Lys-Thr-**Tyr**-Lys-AMC, Ac-Lys-Thr-**Phe**-Lys-AMC, and Ac-Lys-Thr-**Trp**-Lys-AMC (Table 1), 0.020 $\mu M^{-1}\ s^{-1}$ in k_{cat}/K_M versus 0.544 $\mu M^{-1}\ s^{-1}$, 0.677 $\mu M^{-1}\ s^{-1}$, and 0.601 $\mu M^{-1}\ s^{-1}$, respectively. The subtle preference for lysine over phenylalanine at P4 is also verified by single substrates: Ac-**Phe**-Thr-Tyr-Lys-AMC retains 63% of the activity of Ac-**Lys**-Thr-Tyr-Lys-AMC, 0.342 $\mu M^{-1}\ s^{-1}$ in k_{cat}/K_M and 0.544 $\mu M^{-1}\ s^{-1}$ in

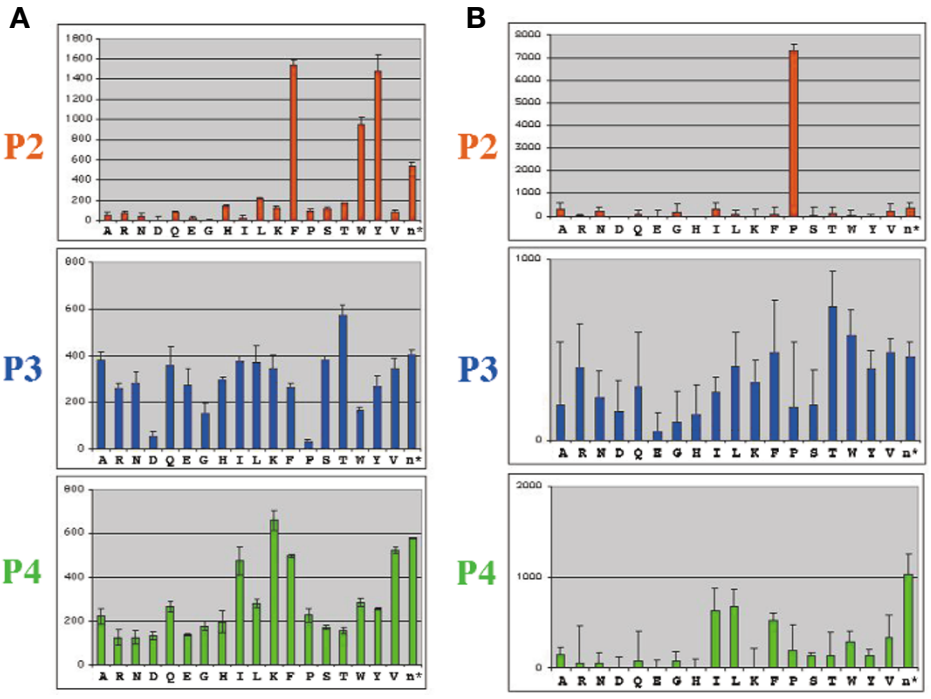

Figure 2. (A) Activity of plasmin in a P1-Lys positional-scanning synthetic combinatorial library. (B) Activity of thrombin in a P1-Lys positional-scanning synthetic combinatorial library. y-axis is picomolar of fluorophore released per second. x-axis indicates the amino acid held constant at each position, designated by the one-letter code (with n representing norleucine).

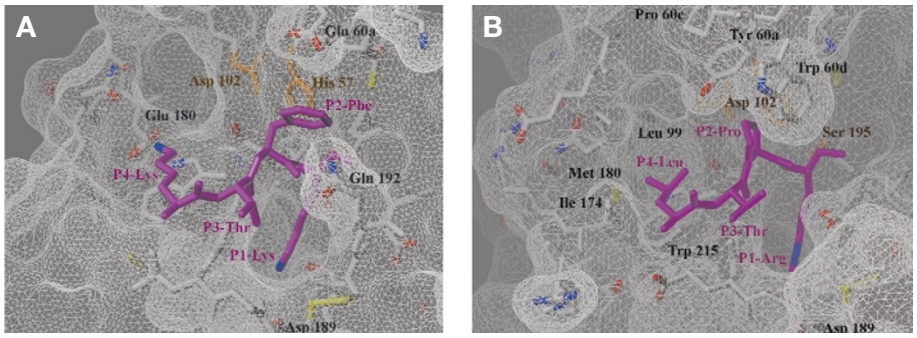

Figure 3. (A) Three-dimensional model of plasmin bound to the tetrapeptide substrate Lys-Thr-Phe-Lys. (B) Three-dimensional model of thrombin bound to the tetrapeptide substrate Leu-Thr-Pro-Arg. The Connolly surface of the enzyme is shown in white mesh. The enzyme backbone is shown in white, with the side chains shown in atom colors (white, carbon; blue, nitrogen; red, oxygen; yellow, sulfur) with the catalytic triad (His57, Asp102, Ser195) in orange. The substrate is shown in magenta.

k_{cat}/K_M, respectively.

Structural determinants of P4-lysine and P2-aromatic substrate specificity in plasmin. The structure of plasmin was solved in the absence of a substrate or inhibitor in the active site[23]. Because of this, analysis of enzyme–substrate interactions required the molecular modeling of the optimal substrate, Lys-Thr-Phe-Lys, into the active site. The resulting model reveals potential structural determinants for substrate recognition. As is appreciated for trypsin, the major determinant for P1-basic specificity lies in Asp189 (according to chymotrypsinogen numbering), a residue at the base of the S1 pocket (Fig. 3A). The (δ+) ring hydrogens from P2-Phe can interact with the carboxylate group of Glu60 that protrudes from above the active site and toward the P2-Phe ring edge. The positively charged (δ+) amino group of Gln192 could then make contact with the (δ-) π-electrons of the P2-Phe face (Fig. 3A). Significant interactions between plasmin and the P3-amino acid side chain are not readily apparent from the structural model. This is a result of the P3-amino acid side chain being directed away from the enzyme and into bulk solvent. A substrate with a P4-Lys could make contact with Glu180, with the aliphatic portion of P4-Lys packing against Trp215 (Fig. 3A). Position 180 is normally occupied by a hydrophobic amino acid in other chymotrypsin-like serine proteases. The additional, though lesser, preference for P4-Nle/Val/Ile/Phe could in part be due to favorable interaction with Trp215.

Profiling of thrombin with the positional scanning P1-Lys library. Profiling of thrombin with the PS-SCL revealed that the preferred P4-P2 extended substrate specificity is for large aliphatic amino acids at P4, such as norleucine, leucine, and isoleucine, negligible discrimination at P3, and narrow specificity for proline at P2 (Fig. 2B). These preferences were validated through single-substrate kinetic analysis (Table 1). Replacement of the optimal P2 amino acid proline with the suboptimal amino acid leucine results in a 45-fold decrease in activity, 3.83 $\mu M^{-1}\ s^{-1}$ in k_{cat}/K_M for Ac-Nle-Thr-**Pro**-Arg-AMC versus 0.085 $\mu M^{-1}\ s^{-1}$ in k_{cat}/K_M for Ac-Nle-Thr-**Leu**-Arg-AMC. The requirement for aliphatic amino acids at P4 has a less pronounced effect than the requirement for proline at P2, as reflected by the relative activities in the PS-SCL. However, upon replacement of the preferred leucine at P4 for the suboptimal glycine, there is a 23-fold decrease in specific activity, 0.154 $\mu M^{-1}\ s^{-1}$ in k_{cat}/K_M for Ac-**Leu**-Gly-Val-Arg-AMC versus 0.007 $\mu M^{-1}\ s^{-1}$ in k_{cat}/K_M for Ac-**Gly**-Gly-Val-Arg-AMC.

Structural determinants of P4-aliphatic and P2-proline substrate specificity in thrombin. The coordinates used for enzyme–substrate analysis of thrombin were that of thrombin complexed with the tripeptide inhibitor, D-Phe-Pro-Arg-chloromethylketone[24]. The P3 side chain was converted to a threonine of the L-enantiomer, and a P4-Val was added to the N terminus of the inhibitor. As suggested from the original structural analysis by Bode et al., the preference for P1-basic amino acids is determined by Asp189 and the preference for proline arises from the insertion, relative to the digestive serine proteases, of seven amino acids in the 60's loop[24]. This loop is above the active site Ser195 and creates a rigid pocket for P2-Pro interaction (Fig. 3B). The specificity for P3 amino acids is less well understood from the original structure. The use of P3-D-Phe allows the side chain to point into the enzyme, occupying, in part, the S4 pocket. When the P3-D-Phe is replaced with P3-L-Thr, the side chain is pointed away from the enzyme with significantly fewer interactions (Fig. 3B). The S4 pocket on thrombin is very clearly hydrophobic, with Ile174 making significant interactions with P4-Val modeled into the pocket. Additional hydrophobic determinants for this pocket include Trp215 at the floor of the pocket, Met180 at the end of the pocket, and Leu99 near the top of the pocket (Fig. 3B).

Table 1. Kinetic constants for plasmin and thrombin on single AMC substrates.

Substrate	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($\mu M^{-1}\ s^{-1}$)
Plasmin			
Ac-Lys-Thr-Tyr-Lys-AMC	11.3 ± 0.4	20.8 ± 3.3	0.544 ± 0.071
Ac-Lys-Thr-Phe-Lys-AMC	20.1 ± 0.6	29.7 ± 3.6	0.677 ± 0.069
Ac-Lys-Thr-Trp-Lys-AMC	11.9 ± 0.3	19.9 ± 2.2	0.601 ± 0.054
Ac-Lys-Thr-Ser-Lys-AMC	8.8 ± 1.3	440 ± 100	0.020 ± 0.002
Ac-Phe-Thr-Tyr-Lys-AMC	17.5 ± 0.9	51.0 ± 10.2	0.342 ± 0.054
Ac-Leu-Thr-Phe-Lys-AMC	33.2 ± 3.9	296 ± 70	0.112 ± 0.015
Ac-Leu-Glu-Phe-Lys-AMC	5.5 ± 0.3	74.6 ± 9.9	0.073 ± 0.006
Thrombin			
Ac-Nle-Thr-Pro-Arg-AMC	45.0 ± 1.1	11.3 ± 1.3	3.83 ± 0.35
Ac-Val-Thr-Pro-Arg-AMC	30.8 ± 1.3	29.6 ± 4.5	1.04 ± 0.13
Ac-Nle-Thr-Leu-Arg-AMC	5.8 ± 0.3	67.4 ± 10.7	0.085 ± 0.009
Ac-Leu-Gly-Val-Arg-AMC	15.6 ± 1.8	101.5 ± 29.5	0.154 ± 0.029
Ac-Gly-Gly-Val-Arg-AMC	1.2 ± 0.1	180.7 ± 55.3	0.007 ± 0.001

Discussion

The rapid discovery of new proteases presents the need for generalized assays to aid in the elucidation of their biological functions. The use of synthetic positional-scanning combinatorial libraries offers the ability to rapidly test and evaluate the extended substrate specificity of a protease. The major limitation of previous synthetic meth-

Table 2. AMC substrate mass spectral data.

Substrate	Mass calculated (g/mol)	Mass found (g/mol)
Ac-Lys-Thr-Tyr-Lys-AMC	738.5	738.9
Ac-Lys-Thr-Phe-Lys-AMC	722.5	723.0
Ac-Lys-Thr-Trp-Lys-AMC	761.5	761.3
Ac-Lys-Thr-Ser-Lys-AMC	662.5	662.8
Ac-Phe-Thr-Tyr-Lys-AMC	757.5	758.0
Ac-Leu-Thr-Phe-Lys-AMC	707.5	707.8
Ac-Leu-Glu-Phe-Lys-AMC	735.5	735.9
Ac-Nle-Thr-Pro-Arg-AMC	685.5	685.8
Ac-Val-Thr-Pro-Arg-AMC	671.5	671.8
Ac-Nle-Thr-Leu-Arg-AMC	701.5	702.0
Ac-Leu-Gly-Val-Arg-AMC	643.5	643.8
Ac-Gly-Gly-Val-Arg-AMC	587.4	587.8

	P4	P3	P2	P1
Optimal Plasmin Substrate Specificity from PS-SCL	**K/V/I/F**	**X**	**F/Y/W**	**R/K**
Vitronectin	K	G	Y	R
Osteocalcin	E	A	Y	R
Factor X	I	T	F	R
PAR1 (70)	T	E	Y	R

Figure 4. Cleavage sites in the physiological substrates of plasmin.

	P4	P3	P2	P1
Optimal Thrombin Substrate Specificity from PS-SCL	**L/I/V/F**	**X**	**P**	**R**
PAR1	L	D	P	R
Factor V (709)	L	G	L	R
Factor V (1018)	L	S	P	R
Factor V (1545)	W	Y	L	R
Factor VIII (372)	I	Q	I	R
Factor VIII (740)	I	E	P	R
Factor VIII (1689)	Q	S	P	R
Factor XI	I	K	P	R
Factor XIII	V	V	P	R
Fibrinogen A a	G	G	V	R
Fibrinogen B b	F	S	A	R
Protein C	L	D	P	R
uPA	L	R	P	R

Figure 5. Cleavage sites in the physiological substrates of thrombin.

ods is the inability to permit complete diversity at the P1 position. The development of the synthetic strategy presented in this paper surmounts this limitation and allows for complete diversity of any naturally occurring amino acid at the P1 position. In fact, a variety of nucleophiles, including unnatural amino acid derivatives, could be incorporated in the cleavage step.

The use of a positional scanning library with P1 lysine held constant was employed to determine the extended specificity of plasmin and thrombin, proteases involved in the regulation of hemostasis. Plasmin has been traditionally characterized as a protease with broad substrate specificity. Results from the current study show that, on the contrary, plasmin demonstrates a distinct preference for aromatic amino acids at P2 and a moderate preference for lysine and hydrophobic amino acids at P4. Molecular modeling of a substrate bound into the active site of plasmin can aid in the identification of potential structural interactions between enzyme and substrate. Plasmin has an insertion in the 60's loop, relative to the digestive protease trypsin, that could allow for the creation of a S2 pocket. Plasmin's S2 pocket is not simply a hydrophobic pocket, but is specific for aromatic amino acids. Experimental crystallographic evidence exists for structural determinants used to achieve discrimination between aromatic and aliphatic amino acids, including aromatic ring interactions with oxygen[25] and amide nitrogens[26]. Likewise, modeling of a P2-Phe into the putative S2 pocket shows that Glu60a may contribute to ring hydrogen bonding, whereas the polar amino group of Gln192 may interact with the π-electrons of P2-Phe (Fig. 3A). The P3-amino acid side chain is pointed out into solvent and makes few interactions with the enzyme. The moderate P4 preference for lysine may be driven by an electrostatic interaction with Glu180 (Fig. 3A). The aliphatic portion of the lysine side chain as well as other aliphatic amino acids (Nle/Val/Ile) could form favorable interactions with Trp215. Preference for both charged and aromatic amino acids is also known for the S2 pocket of the cysteine protease cruzain, where a glutamate side chain adjusts to accommodate the disparate side chains[27].

Plasmin is an important enzyme in attenuating blood coagulation and restoring blood flow through its degradation of fibrin and procoagulant factors[28,29]. Several nonfibrin substrates of plasmin have recently been demonstrated[30–32]. Although the cleavage sites of the physiological substrates have not all been determined, many of those that have been identified resemble the optimal site determined in the current study, especially in the conservation of a P2-aromatic amino acid (Fig. 4). It has been proposed that plasmin may function in accelerating fibrinolysis and arresting coagulation through the cleavage of the normally procoagulant factor X into an anticoagulant cofactor[33]. It was determined that plasmin cleaves Factor X after the arginine in the site Ile-Thr-Phe-Arg. In a similar vein, plasmin may act in the desensitization of the protease activated receptor PAR1. PAR1 is a transmembrane receptor that when cleaved by thrombin results in the activation of platelets; however plasmin cleaves PAR1 at sites that not only do not result in activation but also remove the thrombin activation site[34]. One of the plasmin cleavage sites in PAR1 was determined as Thr-Glu-Tyr-Arg. Plasmin may attenuate its own production through the cleavage of vitronectin, a protein that binds both plasminogen activator inhibitor 1 (PAI-1) and extracellular matrix (ECM)[35,36]. The cleavage of vitronectin by plasmin may result in the release of PAI-1 from the ECM to inhibit plasminogen activators and subsequent inhibition of plasmin formation. The cleavage site of vitronectin is Lys-Gly-Tyr-Arg, a site resembling the optimal site determined in this study. Plasmin may also play a role in regulating bone resorption through the cleavage of osteocalcin at the site Glu-Ala-Tyr-Arg[37].

Thrombin has been shown to be more restrictive in its substrate cleavage profile than trypsin, a protease with the chief function of digestion rather than regulation[38,39]. This is due, in part, to preferences exhibited in the S3-S2' subsites[40,41]. Here, for the first time, the sequence space for thrombin's P4, P3, and P2 substrate preference has been completely sampled. As with previous results[38,41–43], the PS-SCL shows that thrombin has a pronounced preference for proline at the P2 position of the substrate. The structure of thrombin bound to the tripeptide inhibitor, D-Phe-Pro-Arg-chloromethylketone, reveals that the constraints for proline in the P2 position most probably arise from the insertion in the 60's loop of thrombin[24]. When a P3-L-amino acid was modeled into the active site, rather than the P3-D-Phe solved in the structure, the side chain had minimal interaction with the enzyme and was pointed out into solvent (Fig. 3B). This structural analysis of the P3–S3 interaction supports the lack of defined P3 specificity seen in the substrate library. However, a P4-aliphatic amino acid modeled into the active site can make significant interactions with a hydrophobic pocket of thrombin. The walls of this pocket are formed predominantly by Ile174, other amino acids such as Met180, Leu99, and Trp215 also contribute to this hydrophobic environment (Fig. 3B). The interaction between Ile174 and a P4-aliphatic amino acid was also observed in the structure of thrombin bound to the peptide fragment Leu-Asp-Pro-Arg[44]. The structural analysis was supported by the substrate library analysis that showed thrombin to have a definite preference for P4-aliphatic amino acids.

The results from the library cleavage analysis, structural analysis, and kinetic analysis show a unique specificity fingerprint of thrombin for aliphatic amino acids at P4 and for proline at P2. The importance of thrombin's specificity profile for the recognition of proper substrates in coagulation and vascular integrity is appreciated in the context of its physiological substrate cleavage sites (Fig. 5). Thrombin plays an important role in blood coagulation by activating platelets through the cleavage of PAR1 after arginine in the sequence, Leu-Asp-Pro-Arg[45]. Thrombin further enhances coagulation through the feedback cleavage activation of Factor V and Factor VIII at the sequences Leu-Ser-Pro-Arg (1,018) and Trp-Tyr-Leu-Arg (1,545) in Factor V and Ile-Gln-Ile-Arg (372) and Gln-Ser-Pro-Arg (1,689) in Factor VIII[46,47]. Ultimately, thrombin assists in the formation of the fibrin clot through the cleavage of fibrinogen. The sequence that thrombin cleaves in fibrinogen Aα-chain, Gly-Gly-

Val-Arg, is suboptimal by the current analysis. However, it has been demonstrated that the fibrinogen Aα-chain supplements its specific binding energy through the use of thrombin's distal anion-binding exosite 1 (ref. 48). Fibrinogen Bβ-chain has a more optimal thrombin cleavage site, Phe-Ser-Ala-Arg, and has not been shown to use the anion-binding exosite 1. As thrombin moves away from the wound site, it aids in the attenuation of blood clotting through the activation of protein C at the site, Leu-Asp-Pro-Arg[49]. Although thrombin plays several roles in blood coagulation, the common theme is the requirement for proper recognition elements in the physiological substrates. Information from the P1-Lys PS-SCL facilitates the identification of recognition elements for productive thrombin–substrate interaction.

Use of this method was demonstrated for the serine proteases, plasmin and thrombin, two enzymes that require P1-basic amino acids. This library can have broad application to other enzymes because proteases that cleave P1-basic amino acids are well represented in the serine and cysteine protease families. The defined extended substrate specificity for plasmin and thrombin was in agreement with the cleavage sites in known physiological substrates. Substrate specificity information can aid in the discovery of new physiological substrates and the cleavage sites within substrates. A direct outcome from the library analysis is the creation of sensitive substrates to monitor activity. This information can also be used as a starting point in the design and synthesis of potent and selective inhibitors.

Experimental protocol

Materials. Unless otherwise noted, chemicals were obtained from commercial suppliers and used without further purification. Aminomethyl Merrifield resin was purchased from Novabiochem (San Diego, CA), and the substitution concentration of the resin was determined (0.84 mEq/g) by using a spectrophotometric 9-fluorenylmethoxycarbonyl (Fmoc) quantitation assay[50]. Alkane sulfonamide resin (0.75 mmol/g) was prepared by the method of Backes and Ellman[21] or purchased from Novabiochem. Fmoc-amino acids were purchased from Novabiochem. Iodoacetonitrile and anhydrous NMP (*N*-methylpyrrolidone) were purchased from Aldrich. Iodoacetonitrile was filtered through a small plug of basic alumina immediately before use. Anhydrous, low amine content dimethylformamide (DMF) was purchased from EM Science (Cincinnati, OH). High-loading tosyl chloride resin (PS-TsCl, 1.55 mmol/g) was purchased from Argonaut Technologies (San Carlos, CA). An Argonaut Quest 210 Organic Synthesizer was used for library synthesis. Chromatography was carried out using Merck 60 230–240 mesh silica gel according to the procedure of Still[51]. Thin-layer chromatography was carried out on Merck 60 F_{254} 250 μm silica gel plates. Infrared (IR) spectra were recorded neat (for oils) and as films from CH_2Cl_2 or $CHCl_3$ (for crystalline compounds) and only partial data is reported. NMR chemical shifts are reported in p.p.m. downfield from an internal solvent peak, or trimethylsilane, and J values are in hertz. Elemental analyses were done by M-H-W Labs (Phoenix, AZ). A Savant Speed Vac Plus was used for concentrating single-substrate solutions in vials and library member solutions configured in microtiter plates. The human enzymes, thrombin and plasmin were purchased from Haematologic Technologies. (Essex Junction, VT) and used as received.

Library synthesis. Fmoc-amino acids (Fmoc-Ala-OH, Fmoc-Arg(Pbf, 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl)-OH, Fmoc-Asn(Trt, trityl)-OH, Fmoc-Asp(O-*t*-Bu, *t*-butyl)-OH, Fmoc-Glu(O-*t*-Bu)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Gly-OH, Fmoc-His(Boc, *tert*-butoxycarbonyl)-OH, Fmoc-Ile-OH, Fmoc-Leu-OH, Fmoc-Lys(Boc)-OH, Fmoc-Nle, Fmoc-Phe-OH, Fmoc-Pro-OH, Fmoc-Ser(O-*t*-Bu)-OH, Fmoc-Thr(O-*t*-Bu)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Tyr(O-*t*-Bu)-OH, Fmoc-Val-OH) were coupled to the alkane sulfonamide resin[21], and the amino acid loading concentrations were determined using a spectrophotometric Fmoc quantitation assay[50]. For the preparation of the P2 sublibrary, 0.10 mmol (~200 mg) of each of the 19 Fmoc-amino acid resins were added to 19 reaction vessels (one Fmoc-amino acid resin/vessel) of the Argonaut Quest 210 Organic Synthesizer and solvated with DMF (2 ml/vessel). After agitating 20 min, the DMF was drained and a solution of 20% piperidine in DMF (2 ml/vessel) was added. The resin was agitated for 25 min, filtered, and washed with DMF (3 × 2 ml/vessel). To install the randomized P3 position, 10 equivalents (~1.0 mmol/well, 19 mmol) of an isokinetic mixture[22] of Fmoc-amino acids (Fmoc-amino acid, mole%: Fmoc-Ala-OH, 3.4; Fmoc-Arg(Pbf)-OH, 6.5; Fmoc-Asn(Trt)-OH, 5.3; Fmoc-Asp(O-*t*-Bu)-OH, 3.5; Fmoc-Glu(O-*t*-Bu)-OH, 3.6; Fmoc-Gln(Trt)-OH, 5.3; Fmoc-Gly-OH, 2.9; Fmoc-His(Boc)-OH, 3.5; Fmoc-Ile-OH, 17.4; Fmoc-Leu-OH, 4.9; Fmoc-Lys(Boc)-OH, 6.2; Fmoc-Nle, 3.8; Fmoc-Phe-OH, 2.5; Fmoc-Pro-OH, 4.3; Fmoc-Ser(O-*t*-Bu)-OH, 2.8; Fmoc-Thr(O-*t*-Bu)-OH, 4.8; Fmoc-Trp(Boc)-OH, 3.8; Fmoc-Tyr(O-*t*-Bu)-OH, 4.1; Fmoc-Val-OH, 11.3) was preactivated with DICI (diisopropylcarbodiimide)(3.0 ml, 19 mmol), and HOBt(1-hydroxybenzotriazole) (2.6 mg, 19 mmol) in DMF (57 ml) in a 100 ml round bottom flask. After the 2 min preactivation period, 3 ml of the solution was added to each of the 19 reaction vessels. The resin was agitated for 3 h, filtered, and washed with DMF (3 × 2 ml/vessel). After Fmoc removal (treatment with 20% piperidine in DMF (2 ml/vessel), agitation for 25 min, filtration, and washing with DMF (3 × 2 ml/vessel)) the randomized P4 position was incorporated in an identical manner. The Fmoc removal step was followed by filtration and washing with DMF (3 × 2 ml/vessel). A capping solution of AcOH (7.6 mmol), DICI (1.2 ml, 7.6 mmol), HOBt (2.3 g, 7.6 mmol), and DMF (38 ml) was premixed in a 100 ml round bottom flask, and 2 ml was added to each of the 19 reaction vessels. After agitating for 3 h, each resin was filtered, washed (DMF, 3 × 2 ml/vessel; THF, 3 × 2 ml/vessel; MeOH, 3 × 2 ml/vessel), and dried overnight under high vacuum with P_2O_5.

To install the P1 residue, the 19 resins with a fixed P2 residue (0.020 mmol, ~40 mg) were added to the reaction vessels (one fixed P2 residue/vessel), swollen with NMP (1 ml/vessel), agitated for 20 min, and filtered. A solution of ICH_2CN (1.4 ml, 19 mmol), *i*-Pr_2EtN (0.65 ml, 3.8 mmol), and NMP (19 ml) was prepared and added to each of the 19 reaction vessels (1 ml/vessel). The vessels were shielded from light with aluminum foil. After agitating 24 h, each resin was filtered, washed with NMP (5 × 2 ml/vessel), and DMF (5 × 2 ml), and filtered again. Resin washes were agitated for 5 min/wash. A solution of coumarin-Lys(Boc)-NH_2 (760 mg, 1.9 mmol) in DMF (9.5 ml) was prepared and 0.5 ml (5 equivalents) of the solution was added to each vessel. The resin was agitated at 80°C for 12 h to liberate the coumarin-tetrapeptide derivatives. The 19 reaction mixtures were brought to room temperature (RT), and filtered into 19 individual scintillation vials, each containing high-loading tosyl chloride resin (125 mg, 0.300 mmol), Et_3N (41 μl, 0.35 mmol), and DMF (1 ml). Each of the 19 resins was washed with DMF (3 × 0.5 ml), and again, the supernatants were filtered into the 19 tosyl chloride resin-containing vials. The vials were agitated with orbital stirring for 3–4 h. The triethylamine salts produced were free-based by adding K_2CO_3 (200 mg, 1.5 mmol) to each vial followed by agitation over 2 h. The 19 reaction mixtures were filtered into 19 scintillation vials and concentrated. Side chain deprotection was accomplished by adding 1 ml of a TFA (trifluoroacetic acid)/H_2O/triisopropylsilane mixture (95:2.5:2.5) to each vial. After aging for 1 h, the reaction mixtures were concentrated, and ethanol (1 ml) was added to each vial followed by concentration. Ethanol (1 ml) was again added to each vial followed by concentration. The contents of each of the 19 vials were lyophilized after the addition of 1:5 acetonitrile/H_2O (1 ml/well). The synthesis of individual substrates prepared by these methods provided products in 50–60% yield based upon the loading of the P2 support-bound Fmoc-amino acid. The yield of coumarin-peptide compounds in each vial was therefore estimated to be ~0.01 mmol.

The P3 sublibrary and P4 sublibrary were prepared in a similar fashion with the exception that the randomized P2 position was incorporated by hand-mixing the preloaded and quantified Fmoc-amino acid resins. The resin was then transferred to the 19 vessels (0.10 mmol/vessel). To supply the fixed positions, each of the 19 Fmoc-amino acids (0.5 mmol, 5 equivalents) were individually premixed with DICI (78 ml, 0.5 mmol) and HOBt (68 mg, 0.5 mmol) in DMF (2 ml) in a vial and added to the designated resin-containing vessel. The resin was agitated for 3 h, filtered, and washed with DMF (3 × 2 ml/vessel).

Synthesis of Lys(Boc)-7-amino-4-methylcoumarin. To a 100 ml round bottom flask were added 7-amino-4-methylcoumarin (2.00 g, 11.4 mmol), Fmoc-Lys-OH (5.62 g, 12.0 mmol), and DMF (40 ml). After stirring for 5 min, HATU (4.56 g, 12.0 mmol) and collidine (3.2 ml, 24 mmol) were added. The reaction mixture was stirred overnight at RT, diluted with ethyl acetate (500 ml), and extracted with 2N HCl (3 × 300 ml). The organic layer was washed with brine (3 × 300 ml), dried (Na_2SO_4), and concentrated. Purification over silica gel (5 × 20 cm eluted with 96:4 $CHCl_3$/MeOH) provided 5.0 g (70%) of Fmoc-Lys(Boc)-7-amino-4-methylcoumarin as a colorless solid: m.p. 189–191°C; IR 3305, 1734, 1686, 1663, 1615; ^{1}H NMR (300 MHz) δ 1.31 (s, 9),

1.35–1.38 (m, 6), 1.62–1.66 (m, 2), 2.36 (s, 3), 2.88–2.90 (m, 2), 6.23 (s, 1), 6.70 (bt, 1), 7.29–7.36 (m, 2), 7.38–7.42 (m, 2), 7.44 (d, 1, J = 8.7) 7.70–7.80 (m, 5), 7.85 (d, 2, J = 7.5), 10.50 (s, 1); ^{13}C (101 MHz) δ 17.9, 23.0, 28.2, 29.2, 31.3, 46.6, 55.6, 65.7, 77.3, 105.6, 112.3, 115.0, 115.2, 120.1, 125.3, 125.9, 127.1, 127.6, 140.7, 142.2, 143.7, 143.8, 153.1, 153.6, 155.6, 156.2, 160.0, 172.0. Anal. Calcd for $C_{36}H_{39}N_3O_7$: C, 69.10; H, 6.23; N, 6.71. Found: C, 69.11; H, 6.21; N, 6.71. To a 100 ml round bottom flask were added Fmoc-Lys(Boc)-7-amino-4-methylcoumarin (5.0 g, 8.0 mmol), DMF (40 ml), and Et_2NH (1.7 ml, 16 mmol). After stirring for 1 h, the reaction mixture was concentrated and purified over silica gel (5 × 20 cm eluted with 95:5 CH_2Cl_2/MeOH) to provide 3.0 g (93%) of NH_2-Lys(boc)-AMC isolated as a solid: m.p. 113–116°C; IR 1698, 1690, 1617; ^{1}H NMR (300 MHz) δ 1.30–1.40 (m, 15), 1.51–1.58 (m, 2), 2.36 (d, 3, J = 1.10), 2.85 (d, 2, J = 5.9), 3.28–3.30 (m, 1), 6.21 (d, 1, J = 1.10), 6.75 (t, 3, J = 5.5), 7.50 (dd, 1, J = 2.0, J = 9.5), 7.66 (d, 1, J = 9.5), 7.80 (d, 1, J = 2.0), 10.50 (s, 1); ^{13}C (101 MHz) δ 18.4, 23.2, 28.7, 29.9, 35.0, 56.2, 72.2, 77.8, 106.1, 112.6, 115.4, 115.7, 126.3, 142.8, 153.6, 154.1, 156.0, 160.5, 175.8. Analysis calculated for $C_{21}H_{29}N_3O_5$ was C, 62.50; H, 7.19; N, 10.41. Analysis found was C, 62.46; H, 7.28; N, 10.31.

To a 100 ml round bottom flask were added 7-amino-4-methylcoumarin (780 mg, 4.5 mmol), Fmoc-Arg(Pbf)-OH (4.42 g, 6.7 mmol), and DMF (10 ml). After stirring for 5 min, HATU (2.0 g, 6.7 mmol) and collidine (1.8 ml, 13 mmol) were added. The reaction mixture was stirred overnight at RT, and then concentrated. The viscous oil was dissolved in hot ethyl acetate (25 ml) and allowed to cool to RT. The precipitate that had formed upon standing was filtered and washed with ethyl acetate (3 × 5 ml) to provide 2.2 g (55%) of Fmoc-Arg(Pbf)-7-amino-4-methylcoumarin as a gray solid: m.p. 224–225°C; IR 3305, 1719, 1692, 1619; ^{1}H NMR (300 MHz) δ 1.10 (t, 2, J = 7.0), 1.35 (s, 6), 170–175 (m, 2), 1.80–1.85 (m, 2), 1.98 (s, 3), 2.34 (s, 6), 2.40 (s, 3), 2.88 (s, 2), 3.01–3.05 (m, 2), 3.99–4.02 (m, 1), 4.17–4.25 (m, 2), 4.30–4.35 (m, 2), 6.3 (s, 1), 6.40 (s, 1), 6.66–6.68 (m, 1), 7.28–7.30 (m, 2), 7.38 (d, 2, J = 7.4), 7.47 (d, 1, J = 8.8), 7.50–7.53 (m, 1), 7.70–7.73 (m, 3), 7.75–7.78 (m, 2), 7.86 (d, 2, J = 7.4), 10.50 (s, 1); ^{13}C (101 MHz) δ 12.2, 14.1, 17.6, 18.0, 18.9, 20.8, 28.2, 30.0, 42.4, 46.7, 55.3, 59.8, 65.7, 86.2, 105.7, 112.3, 115.1, 116.2, 120.1, 124.3, 125.3, 125.9, 127.1, 127.6, 131.4, 137.3, 140.7, 142.2, 143.7, 143.8, 153.1, 153.6, 156.1, 157.4, 160.0, 169.0, 171.8. Analysis calculated for $C_{45}H_{48}N_5SO_8$ was C, 65.56; H, 5.83; N, 8.68. Analysis found was C, 65.28; H, 5.57; N, 8.42. To a 100 ml round bottom flask were added Fmoc-Arg(Pbf)-7-amino-4-methylcoumarin (2.0 g, 2.4 mmol), DMF (12 ml) and Et_2NH (500 µl, 4.8 mmol). After stirring for 1 h, the reaction mixture was concentrated and purified over silica gel (5 × 20 cm eluted with 90:10 $CHCl_3$/MeOH) to provide 1.3 g (90%) of NH_2-Arg(Pbf)-AMC isolated as a colorless solid: m.p. 134–137°C; IR 3305, 1719, 1692, 1619; ^{1}H NMR (300 MHz) δ 1.05 (t, 2, J = 7.0), 1.35 (s, 6), 1.41–1.52 (m, 3), 1.58–1.60 (m, 1), 1.94 (s, 3), 2.36 (s, 3), 2.38 (s, 3), 2.43 (s, 3), 2.89 (s, 2), 3.03–3.06 (m, 2), 3.30–3.38 (m, 1), 3.40–3.43 (m, 2), 4.33 (s, 1), 6.23 (s, 1), 6.36 (s, 1), 6.71 (s, 1), 7.50–7.60 (m,1), 7.68 (d, 1, J = 8.7), 7.81 (d, 1, J = 1.8), 10.08 (s, 1); ^{13}C (101 MHz) δ 12.2, 17.6, 17.9, 18.6, 18.9, 26.0, 28.3, 32.1, 42.4, 55.4, 56.0, 86.3, 105.6, 112.2, 114.9, 115.3, 116.2, 124.3, 125.9, 131.4, 137.3, 142.3, 153.1, 153.6, 156.1, 157.4, 160.1, 175.1.

Synthesis of single substrates. Single substrates for kinetic analysis were prepared by the previously described methods, except that single amino acids were used in place of mixtures. The P1 residue was introduced using Lys(Boc)-7-amino-4-methylcoumarin or Arg(Pbf)-7-amino-4-methylcoumarin. After side chain deprotection, the unpurified products were subjected to C18 reverse-phase high-pressure liquid chromatography (HPLC) with a 10–40% gradient of 0.1% TFA and 0.08% TFA/95% acetonitrile. The purified products were subsequently lyophilized. All coumarin tetrapeptides were ~95% pure and displayed appropriate molecular masses as determined by liquid chromatography-mass spectrometry (LC-MS; Hewlett-Packard 1100, Palo Alto, CA) (Table 2).

Enzymatic assay of library. The protein concentrations of the enzymes were determined by absorbance measured at 280 nm. Plasmin's and thrombin's extinction coefficients are 1.70 ml mg^{-1} cm^{-1} (ref. 52), 1.83 ml mg^{-1} cm^{-1} (ref. 53), respectively. The proportion of catalytically active protein was quantitated by active site titration with MUGB (4-methylumbelliferyl-*p*-guanidinobenzoate)[54]. Briefly, fluorescence was monitored, with excitation at 360 nm and emission at 450 nm, upon addition of enzyme to MUGB. The concentration of enzyme was determined from the increase in fluorescence based on a standard concentration curve.

Substrates from the PS-SCL were dissolved in DMSO. Approximately 2.5 × 10^{-9} mol of each sublibrary (361 compounds) were added to 57 wells of a 96-well Microfluor White "U" bottom plate (Dynex Technologies, Chantilly, VA). Final substrate concentration was ~0.25 µM, making the hydrolysis of the AMC group directly proportional to the specificity constant, k_{cat}/K_M. Hydrolysis reactions were initiated by the addition of enzyme (0.5–10 nM) and monitored fluorometrically with a Perkin-Elmer LS50B Luminescence Spectrometer 96-well plate reader, with excitation at 380 nm and emission at 460 nm (ref. 55). Assays were done in a buffer containing 50 mM Tris (*tris*-(hydroxymethyl)-amino-methane), pH 8.0, 100 mM NaCl, 5mM $CaCl_2$, 1% DMSO (from substrates), and either 1mg ml^{-1} BSA (bovine serum albumin) or 0.01% Tween-20.

Single-substrate kinetic assays. Enzyme activity was monitored at 25°C in assay buffer containing 50 mM Tris pH 8.0 and 100 mM NaCl, 5 mM $CaCl_2$, and 0.01% Tween-20. Substrate stock solutions were prepared in DMSO. The final concentration of substrate ranged from 0.005 to 2 mM, the concentration of DMSO in the assay was less than 5%. Enzyme concentrations ranged from 5 to 50 nM. Hydrolysis of AMC substrates was monitored fluorometrically with an excitation wavelength of 380 nm and emission wavelength of 460 nm on a Fluoromax-2 spectrofluorimeter.

Molecular modeling of thrombin–substrate and plasmin–substrate complex. The coordinates for thrombin bound to D-Phe-Pro-Arg-chloromethylketone (1PPB) (ref. 24) and plasmin complexed with streptokinase (1BML) (ref. 23) were obtained from the Protein Data Bank[56]. The Biopolymer module of the Insight II (Molecular Simulations, San Diego, CA) molecular modeling package was used to build and model substrates into the active sites of the proteases. Briefly, the P3-D-Phe in the thrombin structure was deleted and P3-L-Thr and P4-L-Leu were added. The preferred side chain rotamers were explored manually to maximize interaction with thrombin. The substrate for plasmin was built into the active site by superposition of the catalytic triad of plasmin (1BML) with the catalytic triad of the peptidyl-chloromethylketone inhibited thrombin (1PPB). The coordinates for the thrombin protein were deleted, leaving the peptidyl-chloromethylketone inhibitor docked to plasmin. The side chains from P3-D-Phe, P2-Pro, and P1-Arg were then deleted from the inhibitor and replaced with P4-Lys, P3-Thr, P2-Phe, and P1-Lys side chains. The side chains were manually rotated to maximize interactions with plasmin.

Acknowledgments

We thank Shaun R. Coughlin and Mark Lipton for insightful discussions. We extend our appreciation to Keith W. Burdick for helpful discussion and assistance with the figures and to Matthew Trammel for assistance with the preparation and analysis of single-substrate kinetics. This work was supported in part by National Science Foundation Grant MCB9604379 and National Institutes of Health Grant CA72006 (to C.S.C.), National Institutes of Health Grant GM 54051 (to J.A.E.) and National Institutes of Health Biotechnology Training Grant Fellowship (to J.L.H.).

1. Matthews, D.J. & Wells, J.A. Substrate phage: selection of protease substrates by monovalent phage display. *Science* **260**, 1113–1117 (1993).
2. Ding, L. et al. Origins of the specificity of tissue-type plasminogen activator. *Proc. Natl. Acad. Sci. USA* **92**, 7627–7631 (1995).
3. Bevan, A., Brenner, C. & Fuller, R.S. Quantitative assessment of enzyme specificity in vivo: P2 recognition by Kex2 protease defined in a genetic system. *Proc. Natl. Acad. Sci.USA* **95**,10384–10389 (1998).
4. Lustig, K.D. et al. Small pool expression screening: identification of genes involved in cell cycle control, apoptosis, and early development. *Methods Enzymol.* **283**, 83–99 (1997).
5. Kothakota, S. et al. Caspase-3-generated fragment of gelsolin: effector of morphological change in apoptosis. *Science* **278**, 294–298 (1997).
6. Petithory, J.R., Masiarz, F.R., Kirsch, J.F., Santi, D.V. & Malcolm, B.A. A rapid method for determination of endoproteinase substrate specificity: specificity of the 3C proteinase from hepatitis A virus. *Proc. Natl. Acad. Sci. USA* **88**, 11510–11514 (1991).
7. Birkett, A.J. et al. Determination of enzyme specificity in a complex mixture of peptide substrates by N-terminal sequence analysis. *Anal. Biochem.* **196**, 137–143 (1991).
8. Berman, J. et al. Rapid optimization of enzyme substrates using defined substrate mixtures. *J. Biol. Chem.* **267**, 1434–1437 (1992).
9. McGeehan, G.M. et al. Characterization of the peptide substrate specificities of interstitial collagenase and 92-kDa gelatinase. Implications for substrate optimization. *J. Biol. Chem.* **269**, 32814–32820 (1994).
10. Schellenberger, V., Turck, C.W., Hedstrom, L. & Rutter, W.J. Mapping the S' subsites of serine proteases using acyl transfer to mixtures of peptide nucleophiles. *Biochemistry* **32**, 4349–4353 (1993).
11. Lam, K.S. & Lebl, M. Synthesis of a one-bead one-compound combinatorial peptide library. *Methods Mol. Biol.* **87**, 1–6 (1998).
12. Meldal, M., Svendsen, I., Breddam, K. & Auzanneau, F.I. Portion-mixing peptide libraries of quenched fluorogenic substrates for complete subsite mapping of endoprotease specificity. *Proc. Natl. Acad. Sci.USA* **91**, 3314–3318 (1994).
13. Meldal, M. et al. Inhibition of cruzipain visualized in a fluorescence quenched solid-phase inhibitor library assay. D-amino acid inhibitors for cruzipain, cathep-

sin B and cathepsin L. *J. Peptide Sci.* **4**, 83–91 (1998).
14. St. Hilaire, P.M., Willert, M., Juliano, M.A., Juliano, L. & Meldal, M. Fluorescence-quenched solid phase combinatorial libraries in the characterization of cysteine protease substrate specificity. *J. Combinatorial Chem.* **VI**, 509–523 (1999).
15. Del Nery, E. et al. Characterization of the substrate specificity of the major cysteine protease (cruzipain) from Trypanosoma cruzi using a portion-mixing combinatorial library and fluorogenic peptides. *Biochem. J.* **323**, 427–433 (1997).
16. Dooley, C.T. & Houghten, R.A. Synthesis and screening of positional scanning combinatorial libraries. *Methods Mol. Biol.* **87**, 13–24 (1998).
17. Pinilla, C., Appel, J.R., Blanc, P. & Houghten, R.A. Rapid identification of high affinity peptide ligands using positional scanning synthetic peptide combinatorial libraries. *Biotechniques* **13**, 901–905 (1992).
18. Rano, T.A. et al. A combinatorial approach for determining protease specificities: application to interleukin-1beta converting enzyme (ICE). *Chem.Biol.* **4**, 149–155 (1997).
19. Thornberry, N.A. et al. 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 (1997).
20. Schechter, I., Berger, A. On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Chem. Commun.* **27**, 157–162 (1968).
21. Backes, B.J. & Ellman, J.A. An alkane sulfonamide "safety-catch" linker for solid-phase synthesis. *J. Org. Chem.* **64**, 2322–2330 (1999).
22. Ostresh, J.M., Winkle, J.H., Hamashin, V.T. & Houghten, R.A. Peptide libraries: determination of relative reaction rates of protected amino acids in competitive couplings. *Biopolymers* **34**, 1681–1689 (1994).
23. Wang, X., Lin, X., Loy, J.A., Tang, J. & Zhang, X.C. Crystal structure of the catalytic domain of human plasmin complexed with streptokinase. *Science* **281**, 1662–1665 (1998).
24. Bode, W. et al. The refined 1.9 Å crystal structure of human α-thrombin: interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment. *EMBO J.* **8**, 3467–3475 (1989).
25. Thomas, K.A., Smith, G.M., Thomas, T.B. & Feldmann, R.J. Electronic distributions within protein phenylalanine aromatic rings are reflected by the three-dimensional oxygen atom environments. *Proc. Natl. Acad. Sci.USA* **79**, 4843–4847 (1982).
26. Burley, S.K. & Petsko, G.A. Amino-aromatic interactions in proteins. *FEBS Lett.* **203**, 139–143 (1986).
27. Gillmor, S.A., Craik, C.S. & Fletterick, R.J. Structural determinants of specificity in the cysteine protease cruzain. *Protein Sci.* **6**, 1603–1611 (1997).
28. Omar, M.N. & Mann, K.G. Inactivation of factor Va by plasmin. *J. Biol. Chem.* **262**, 9750–9755 (1987).
29. McKee, P.A., Andersen, J.C. & Switzer, M.E. Molecular structural studies of human factor VIII. *Ann. NY Acad. Sci.* **240**, 8–33 (1975).
30. Gundersen, D., Traan-Thang, C., Sordat, B., Mourali, F. & Reuegg, C. Plasmin-induced proteolysis of tenascin-C: modulation by T lymphocyte-derived urokinase-type plasminogen activator and effect on T lymphocyte adhesion, activation, and cell clustering. *J.Immunol.* **158**, 1051–1060 (1997).
31. Campbell, P.G. & Andress, D.L. Plasmin degradation of insulin-like growth factor-binding protein-5 (IGFBP-5): regulation by IGFBP-5-(201-218). *Am. J. Physiol.* **273**, E996–1004 (1997).
32. Tsirka, S.E., Bugge, T.H., Degen, J.L. & Strickland, S. Neuronal death in the central nervous system demonstrates a non-fibrin substrate for plasmin (published erratum appears in *Proc Natl Acad Sci USA* **26**, 14976, 1997). *Proc. Natl. Acad. Sci. USA* **94**, 9779–9781 (1997).
33. Pryzdial, E.L., Lavigne, N., Dupuis, N. & Kessler, G.E. Plasmin converts factor X from coagulation zymogen to fibrinolysis cofactor. *J. Biol. Chem.* **274**, 8500–8505 (1999).
34. Kuliopulos, A. et al. Plasmin desensitization of the PAR1 thrombin receptor: kinetics, sites of truncation, and implications for thrombolytic therapy. *Biochemistry* **38**, 4572–4585 (1999).
35. Kost, C., Benner, K., Stockmann, A., Linder, D. & Preissner, K.T. Limited plasmin proteolysis of vitronectin. Characterization of the adhesion protein as morpho-regulatory and angiostatin-binding factor. *Eur. J. Biochem.* **236**, 682–688 (1996).
36. Chain, D., Kreizman, T., Shapira, H. & Shaltiel, S. Plasmin cleavage of vitronectin. Identification of the site and consequent attenuation in binding plasminogen activator inhibitor-1. *FEBS Lett.* **285**, 251–256 (1991).
37. Novak, J.F., Hayes, J.D. & Nishimoto, S.K. Plasmin-mediated proteolysis of osteocalcin. *J.Bone Mineral Res.* **12**, 1035–1042 (1997).
38. Pozsgay, M. et al. Study of the specificity of thrombin with tripeptidyl-*p*-nitroanilide substrates. *Eur. J. Biochem.* **115**, 491–495 (1981).
39. Bode, W., Turk, D. & Karshikov, A. The refined 1.9-Å X-ray crystal structure of D-Phe-Pro-Arg chloromethylketone-inhibited human alpha-thrombin: structure analysis, overall structure, electrostatic properties, detailed active-site geometry, and structure-function relationships. *Protein Sci.* **1**, 426–471 (1992).
40. Le Bonniec, B.F. et al. Characterization of the P2' and P3' specificities of thrombin using fluorescence-quenched substrates and mapping of the subsites by mutagenesis. *Biochemistry* **35**, 7114–7122 (1996).
41. Vindigni, A., Dang, Q.D. & Di Cera, E. Site-specific dissection of substrate recognition by thrombin. *Nat. Biotechnol.* **15**, 891–895 (1997).
42. Lottenberg, R., Hall, J.A., Blinder, M., Binder, E.P. & Jackson, C.M. The action of thrombin on peptide *p*-nitroanilide substrates. Substrate selectivity and examination of hydrolysis under different reaction conditions. *Biochim. Biophys. Acta* **742**, 539–557 (1983).
43. Kawabata, S. et al. Highly sensitive peptide-4-methylcoumaryl-7-amide substrates for blood-clotting proteases and trypsin. *Eur. J. Biochem.* **172**, 17–25 (1988).
44. Mathews, I.I. et al. Crystallographic structures of thrombin complexed with thrombin receptor peptides: existence of expected and novel binding modes. *Biochemistry* **33**, 3266–3279 (1994).
45. Vu, T.K., Wheaton, V.I., Hung, D.T., Charo, I. & Coughlin, S.R. Domains specifying thrombin-receptor interaction. *Nature* **353**, 674–677 (1991).
46. Pittman, D.D., Tomkinson, K.N., Michnick, D., Selighsohn, U. & Kaufman, R.J. Posttranslational sulfation of factor V is required for efficient thrombin cleavage and activation and for full procoagulant activity. *Biochemistry* **33**, 6952–6959 (1994).
47. Keller, F.G., Ortel, T.L., Quinn-Allen, M.A. & Kane, W.H. Thrombin-catalyzed activation of recombinant human factor V. *Biochemistry* **34**, 4118–4124 (1995).
48. Stubbs, M.T. & Bode, W. A model for the specificity of fibrinogen cleavage by thrombin. *Semin. Thrombosis Hemostasis* **19**, 344–351 (1993).
49. Ehrlich, H.J. et al. Recombinant human protein C derivatives: altered response to calcium resulting in enhanced activation by thrombin. *EMBO J.* **9**, 2367–2373 (1990).
50. Bunin, B.A. *The combinatorial index*. xvii, 322 (Academic, San Diego; 1998).
51. Still, W.C., Kahn, M. & Mitra, A. Rapid chromatographic technique for preparative separations with moderate resolution. *J. Org. Chem.* **43**, 2923–2925 (1978).
52. Robbins, K.C., Summaria, L. & Wohl, R.C. Human plasmin. *Methods Enzymol. C* **80**, 379–387 (1981).
53. Fenton, J.W.d., Fasco, M.J. & Stackrow, A.B. Human thrombins. Production, evaluation, and properties of alpha-thrombin. *J. Biol. Chem.* **252**, 3587–3598 (1977).
54. Jameson, G., Roberts, D.V., Adams, R.W., Kyle, W.S., & Elmore, D.T. Determination of the operational molarity of solutions of bovine alpha-chymotrypsin, trypsin, thrombin and factor Xa by spectrofluorimetric titration. *Biochem. J.* **131**, 107–117 (1973).
55. Zimmerman, M., Ashe, B., Yurewicz, E. & Patel, G. Sensitive assay for trypsin, elastase, and chymotrypsin using fluorogenic substrates. *Anal. Biochem.* **78**, 47–51 (1977).
56. Bernstein, F.C. et al. The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535 (1977).