Distinguishing the Specificities of Closely Related Proteases

ROLE OF P3 IN SUBSTRATE AND INHIBITOR DISCRIMINATION BETWEEN TISSUE-TYPE PLASMINOGEN ACTIVATOR AND UROKINASE

(Received for publication, November 19, 1996, and in revised form, April 16, 1997)

Song-Hua Ke†, Gary S. Coombs§, Kathy Tachias‡, Marc Navre†, David R. Corey§‡, and Edwin L. Madison†‡**

From the †Department of Vascular Biology, The Scripps Research Institute, La Jolla, California 92037, the §Department of Pharmacology and Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, Texas 75235, and ‡Affymax Research Institute, Santa Clara, California 95051

Elucidating subtle specificity differences between closely related enzymes is a fundamental challenge for both enzymology and drug design. We have addressed this issue for two intimately related serine proteases, tissue-type plasminogen activator (t-PA) and urokinase-type plasminogen activator (u-PA), by modifying the technique of substrate phage display to create substrate subtraction libraries. Characterization of individual members of the substrate subtraction library accomplished the rapid, direct identification of small, highly selective substrates for t-PA. Comparison of the amino acid sequences of these selective substrates with the consensus sequence for optimal substrates for t-PA, derived using standard substrate phage display protocols, suggested that the P3 and P4 residues are the primary determinants of the ability of a substrate to discriminate between t-PA and u-PA. Mutagenesis of the P3 and P4 residues of plasminogen activator inhibitor type 1, the primary physiological inhibitor of both t-PA and u-PA, confirmed this prediction and indicated a predominant role for the P3 residue. Appropriate replacement of both the P3 and P4 residues enhanced the t-PA specificity of plasminogen activator inhibitor type 1 by a factor of 600, and mutation of the P3 residue alone increased this selectivity by a factor of 170. These results demonstrate that the combination of substrate phage display and substrate subtraction methods can be used to discover specificity differences between very closely related enzymes and that this information can be utilized to create highly selective inhibitors.

The chymotrypsin family of serine proteases has evolved to include members with both widely divergent and intimately related substrate specificities (1). We chose two members of this family, tissue-type plasminogen activator (t-PA) and urokinase (u-PA), to test the hypothesis that small molecule libraries could be used to identify substrates that discriminate between closely related enzymes. This choice of enzymes assured a rigorous test of the hypothesis because t-PA and u-PA possess an extremely high degree of structural similarity (2, 3), share the same primary physiological substrate (plasminogen) and inhibitors (plasminogen activator inhibitor types 1 and 2) (4, 5), and exhibit restricted substrate specificity (6–8).

Despite their striking similarities, the physiological roles of t-PA and u-PA are distinct (9), and many studies (10–16), including several that utilize transgenic mice (9, 11, 16), suggest that selective inhibition of either enzyme might have beneficial therapeutic effects. Mice lacking t-PA, for example, are resistant to specific excitotoxins that cause extensive neurodegeneration in wild type mice (11), and mice lacking u-PA exhibit defects in the proliferation and/or migration of smooth muscle cells in a model of restenosis following vascular injury (9). u-PA-deficient mice are also resistant to the induction and/or progression of several tumor types in a two-stage, chemical carcinogenesis model (16).

Because mice lacking either t-PA or u-PA do not develop thrombotic disorders, selective inhibition of either of these two enzymes seems unlikely to create thrombotic complications in vivo. On the other hand, mice lacking both t-PA and u-PA suffer severe thrombosis in many organs and tissues, resulting in a significantly reduced life expectancy (9). Nonselective inhibition of these two enzymes, therefore, seems almost certain to produce catastrophic consequences in the clinical setting. Consequently, significant interest exists in the development of inhibitors that are stringently specific for either t-PA or u-PA. Rational design of these selective inhibitors is greatly complicated, however, by the absence of obvious "lead compounds"; both their primary physiological substrate and inhibitors fail to discriminate between the two closely related proteases.

We have previously described the use of substrate phage display, a strategy originally developed by Matthews and Wells (17), to elucidate optimal subsite occupancy for substrates of t-PA and to isolate peptide substrates that were cleaved as much as 5300 times more efficiently by t-PA than peptides containing the primary sequence of the actual target site present in plasminogen (18). Four of these selected substrates, including the most labile t-PA substrate, were chosen for detailed characterization; consequently, small peptides containing the four selected amino acid sequences were synthesized, and a kinetic analysis of the cleavage of these peptides by t-PA was performed (18). Subsequent analysis of these four selected substrates demonstrated that they were also efficiently cleaved by u-PA and therefore were not selective for t-PA versus u-PA. Consequently, to facilitate the rapid isolation of peptide sub-

* This study was funded in part by National Institutes of Health Grants RO1 HL52475 and PO1 HL31950 (to E. L. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
† An assistant investigator with the Howard Hughes Medical Institute. To whom correspondence may be addressed: Dept. of Pharmacology and Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, Texas 75235.
‡ To whom correspondence may be addressed: Dept. of Vascular Biology (VB-1), The Scripps Research Institute, La Jolla, CA 92037. Fax: 619-784-7323.
§ To whom correspondence may be addressed: Dept. of Pharmacology and Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, Texas 75235.
strates that were cleaved at greater catalytic efficiencies by t-PA than by u-PA, we developed a novel protocol to prepare phage “substrate subtraction libraries” (Fig. 1). Characterization of individual members of the substrate subtraction library quickly accomplished the direct identification of highly labile, small peptide substrates that were preferentially cleaved by t-PA. In addition, insights gained during the analysis of these selective substrates were used to design a variant of the serpin PAI-1 whose selectivity toward t-PA was enhanced by a factor of 600.

MATERIALS AND METHODS

Reagents—Competent MC1061 (F') Escherichia coli and nitrocellulose were purchased from Bio-Rad. Pansorbin (Protein A-bearing Staphylococcus aureus) cells were obtained from Calbiochem. K91 (F') and MC1061 (F') strains of E. coli were provided by Steve Cwirla (Affymax). mAb 179 which recognizes the epitope (ACLEPYTACD) of the human placental alkaline phosphatase protein with subnanomolar affinity, was provided by Ron Barrett (Affymax). mAb 3-E7 was purchased from Gamsch Laboratories (Schwabhausen, FRG). t-PA was obtained from Bruce Keyt (Genentech), and u-PA was obtained from Jack Henkin (Abbott Laboratories).

Construction of the phage vector fAFF1-tether C (fTC) and the random hexapeptide library fAFF-TC-LIB was subjected to three rounds of high stringency screening with t-PA (18) to prepare an intermediate library containing phage whose randomized hexamer sequences were digested efficiently by t-PA. The intermediate library was then amplified and screened at low stringency, with u-PA. Following digestion of the intermediate library with u-PA, mAb E-7 and immobilized protein A were added to the mixture, and the resulting ternary complexes were pelleted by centrifugation as described previously (18, 19). In contrast to all previous screening steps, however, we retained the precipitated ternary complexes and discarded the supernatant, which contained phage that were digested by u-PA. The precipitated ternary complexes were washed twice with 50 mM Tris pH 7.5, 0.1 mM NaCl, 1 mM EDTA, 0.01% Tween 80, resuspended, and digested with 50 μg/ml t-PA at 37 °C for 4 h. The remaining ternary complexes were precipitated by centrifugation and discarded, and the supernatant, which contained t-PA-selective phage, was retained. Phage in the final supernatant were amplified overnight in E. coli K91 cells. Following amplification, individual phage clones were functionally characterized using a dot blot assay and analyzed by DNA sequencing.

Dot Blot Assay of Phage Proteolysis—Phage precipitation and dot blot analysis were performed as described previously (18, 19). Individual phage stocks were prepared and digested with no enzyme, t-PA, or u-PA in the presence of 1 mM amiloride, a specific inhibitor of u-PA, for periods of time varying from 15 min to 10 h. Individual reaction mixtures were spotted onto a nitrocellulose filter using a dot bloter apparatus (Bio-Rad). The filter was probed with mAb 3E-7 and developed using the Amersham Western ECL kit. Loss of positive staining indicates loss of antibody epitopes from the phage due to proteolytic cleavage of the randomized hexamer region.

Preparation and Sequencing of DNA from Phage Clones—DNA samples were prepared from interesting phage clones as described previously (18). Briefly, phage are precipitated from a 1-ml overnight culture by adding 200 μl of 20% polyethylene glycol in 2.5 mM NaCl. The mixture was incubated on ice for 30 min, and the phage pellet was collected by microcentrifugation for 5 min. The phage were resuspended in 40 μl of lysis buffer (10 mM Tris-HCl, pH 7.6, 0.1 mM EDTA, 0.5% Triton X-100) and heated at 80 °C for 15 min. Single-stranded DNA was purified by phenol extraction and ethanol precipitation. One-third of the single-stranded DNA was used for dideoxy sequencing.

Kinetics of Cleavage of Synthetic Peptides by t-PA and u-PA—Peptides were synthesized and purified as described (7). Kinetic data were obtained by incubating various concentrations of peptide with a constant enzyme concentration to achieve between 5% and 20% cleavage of the peptide in each reaction. For assays with u-PA, enzyme concentration was either 815 or 635 nM. For assays with t-PA, enzyme concentration was 700 nM. Peptide concentrations were chosen where possible to surround Km and in all cases were between 1 and 32 μM. The buffer used in these assays has been described (7). Reactions were stopped by the addition of trifluoracetic acid to 0.33% or by freezing on dry ice. Cleavage of the 13- and 14-residue peptides was monitored by reverse phase HPLC (19) as described (20). The 4–6–residue peptides were acylated at the lysine termini and analyzed at the carboxyl termini. Cleavage of the 4–6–residue peptides was monitored by hydrophilic interaction HPLC chromatography (22) using a polyhydroxyaspartamine column from Polylc (Columbia, MD). Buffer A was 50 mM triethylamine phosphate in 10% acetonitrile, and buffer B was 10 mM triethylamine phosphate in 80% acetonitrile. Peptides were eluted by a gradient that was varied from 100% buffer B to 100% buffer A during a 15-min interval. The percentage of cleaved peptide was calculated by dividing the area under the product peaks by the total area under substrate and product peaks. For all peptides containing multiple basic residues, mass spectral analysis of products confirmed that cleavage occurred at a single site and identified the scissile bond. Data were interpreted by Edie-Hofstee analysis. Errors were determined as described (23) and were ±25%.

Site-directed Mutagenesis and Construction of Expression Vectors Encoding Variants of PAI-1—The expression vector pPAIST7HS was derived from the plasmid pBR322 and contains a full-length cDNA encoding human PAI-1 that is transcribed from a T7 gene 10 promoter (24). The 300-base pair SalI/BamHI fragment of human PAI-1 was subcloned from pPAIST7HS into bacteriophage M13mp18. Single-stranded phage clones were screened with t-PA using reaction conditions reported previously (18). Individual reaction conditions were used to screen the phage library with u-PA except that digestion of the phage was performed using enzyme concentrations varying from 5 to 10 μg/ml and incubation times varying from 1 to 10 h.

Preparation of Substrate Subtraction Libraries—The initial random hexamer phage library fAFF-TC-LIB was subjected to three rounds of high stringency screening with t-PA (18) to prepare an intermediate library containing phage whose randomized hexamer sequences were digested efficiently by t-PA. The intermediate library was then amplified and screened, at low stringency, with u-PA. Following digestion of the intermediate library with u-PA, mAb E-7 and immobilized protein A were added to the mixture, and the resulting ternary complexes were pelleted by centrifugation as described previously (18, 19). In contrast to all previous screening steps, however, we retained the precipitated ternary complexes and discarded the supernatant, which contained phage that were digested by u-PA. The precipitated ternary complexes were washed twice with 50 mM Tris pH 7.5, 0.1 mM NaCl, 1 mM EDTA, 0.01% Tween 80, resuspended, and digested with 50 μg/ml t-PA at 37 °C for 4 h. The remaining ternary complexes were precipitated by centrifugation and discarded, and the supernatant, which contained t-PA-selective phage, was retained. Phage in the final supernatant were amplified overnight in E. coli K91 cells. Following amplification, individual phage clones were functionally characterized using a dot blot assay and analyzed by DNA sequencing.

Expression and Purification of Recombinant PAI-1 Variants—Expression of wild type and mutated variants of PAI-1 was accomplished in the E. coli strain BL21 (DE3) lysy+ (Novagen), which synthesizes T7 RNA polymerase in the presence of isopropyl-1-thio-β-D-galactopyranoside. Bacterial cultures were grown at 37 °C with vigorous shaking to an A350 of 1.1–1.3, and isopropyl-1-thio-β-D-galactopyranoside was added to a final concentration of 1 mM to induce the synthesis of T7 RNA polymerase and the production of PAI-1 proteins. Cultures were grown for an additional 1–2 h at 37 °C and then shifted to 30 °C for 2–6 h.

Cells were pelleted by centrifugation at 8000 × g for 20 min at 4 °C and resuspended in 40 ml of cold start buffer (20 mM sodium acetate, 200 mM NaCl, and 0.01% Tween 20, pH 5.6). The cell suspension was disrupted in a French pressure cell (Aminco), and cellular debris was removed by ultracentrifugation for 25 min at 32,000 × g.

Purification of soluble, active PAI-1 was performed as described previously (27). PAI-1-containing supernatants were injected into a XK-26 column (Pharmacia) packed with CM-50 Sephadex (Pharmacia). The column was washed with 5 column volumes of start buffer (20 mM sodium acetate, 200 mM NaCl, and 0.01% Tween 20, pH 5.6), and PAI-1 proteins were eluted using a 0.2–1.8 mM gradient of NaCl in the same buffer. Peak fractions were collected, pooled, and concentrated using a centrisc 30 concentrator (Amicon). Purified preparations were analyzed by activity measurements using standard,
direct assays of t-PA, SDS-polyacrylamide gel electrophoresis, and measurement of optical density at 280 nm.

Measurement of Active PAI-1 in Purified Preparations—A primary standard of trypsin was prepared by active site titration using p-nitrophenyl p′-guanidinoazobenzoate HCl as described previously (28). Concentrations of active molecules in purified preparations of wild type or mutated PAI-1s were determined by titration of standardized trypsin as described by Olson et al. (29) and by titration of standardized t-PA preparations.

Kinetic Analysis of the Inhibition of t-PA and u-PA by Wild Type and Mutated Variants of PAI-1—Second order rate constants ($k_i$) for inhibition of t-PA or u-PA were determined using pseudo-first order ($k_i < 2 \times 10^8$) or second order ($k_i > 2 \times 10^6$) conditions. For each reaction, the concentrations of enzyme and inhibitor were chosen to yield several data points for which the residual enzymatic activity varied between 20 and 80% of the initial activity. Reaction conditions and data analysis for pseudo-first order reactions were as described previously (30–33).

For second order reactions, equimolar concentrations of u-PA and PAI-1 were mixed directly in microtiter plate wells and preincubated at room temperature for periods of time varying from 0 to 30 min. Following preincubation the mixtures were quenched with an excess of neutralizing anti-PAI-1 antibody (generously provided by Dr. David Loskutoff, The Scripps Research Institute), and residual enzymatic activity was measured using a standard, indirect chromogenic assay. These indirect, chromogenic assays were compared with control reactions containing no PAI-1 or with reactions to which PAI-1 was added after preincubation and the addition of anti-PAI-1 antibody, plasminogen, and Spec PL. Data were analyzed by plotting the reciprocal of the residual enzyme concentration versus the time of preincubation.

RESULTS AND DISCUSSION

Construction of Substrate Phage Libraries—A polyvalent fd phage library that displayed random hexapeptide sequences and contained $2 \times 10^9$ independent recombinants was prepared (18, 19). Each member of this library displayed an N-terminal extension from phage coat protein III containing a randomized region of six amino acids followed by a six-residue linker sequence (SSGGSSG) and the epitopes for mAb 179 and mAb 3-E7. Because neither t-PA nor u-PA digests the protein III sequence, the antibody epitopes, or the flexible linker sequence, the loss of antibody epitopes from the phage surface upon incubation with either enzyme requires cleavage of the randomized peptide insert. Incubation of the library with t-PA, followed by removal of phage retaining the antibody epitopes, therefore, accomplishes the enrichment of phage clones whose random hexamer sequence can be cleaved by t-PA.

Construction of Substrate Subtraction Libraries—The initial phage library was subjected to three rounds of high stringency selection with t-PA to assure the preparation of an intermediate library that is highly enriched for phage that are efficient substrates of t-PA (Fig. 1). This intermediate library was then digested at low stringency with u-PA to remove phage that are moderate or good substrates for u-PA. Substrate subtraction was accomplished after the protease digestion of phage by adding mAb 3E-7 and immobilized protein A to the reaction mixture and precipitating the ternary complexes that contain the undigested phage. By contrast to all earlier selections, the phage remaining in solution were discarded, and the precipitate containing the ternary complexes was resuspended. Phage that were preferentially cleaved by t-PA were then released from the ternary complexes by digestion with t-PA.

Identification and Kinetic Characterization of t-PA-selective Substrates—Using the protocol outlined in Fig. 1 and a previously described (18, 19) sensitive dot blot assay (Fig. 2), we isolated and functionally verified 37 t-PA-selective clones that contained 32 distinct substrate sequences (Fig. 3). As illustrated by the experiment depicted in Fig. 2, the dot blot assay can rapidly provide information regarding both the activity and specificity of individual substrate phage clones. Based on the results of these assays, three peptide substrates (II–IV) containing hexamer sequences present in individual members of the substrate subtraction library were synthesized and characterized to provide a quantitative analysis of the properties of putative t-PA-selective substrates. These peptides were cleaved 180–1500-fold times more efficiently by t-PA than by a control peptide (I) containing the physiological cleavage site present in plasminogen (PGRVVG) and was not digested by either enzyme under the conditions used in this experiment. Substrate phage 51 was isolated from the intermediate library used to create the t-PA substrate subtraction library, contained the hexamer RIARRA, and was an efficient substrate of both t-PA and u-PA. Phage 7 was a member of the t-PA subtraction library, contained the hexamer FGRGAA, and, as predicted, was a t-PA-selective substrate. Phage 33 was isolated from a conventional u-PA substrate library, contained the hexamer RSNAIL, and was a u-PA-selective substrate.

The t-PA/u-PA selectivity possessed by members of the substrate subtraction library was significantly greater than that displayed by members of an “optimized” substrate library constructed using standard substrate phage display methods. We
Phage Substrate Subtraction Libraries

### Table

| Clone | P5 | P4 | P3 | P2 | P1 | P1 | P2 | P3 |
|-------|----|----|----|----|----|----|----|----|
| 1     | A  | L  | R  | G  | D  |    |    |    |
| 2     | D  | Y  | R  | G  | M  | (L)|    |    |
| 3     | E  | R  | A  | R  | A  |    |    |    |
| 4     | E  | R  | L  |    |    |    |    |    |
| 5     | (G)| F  | G  | R  | H  | A  |    |    |
| 6     | F  | L  | R  |    |    |    |    |    |
| 7     | R  | G  | R  | A  |    |    |    |    |
| 8     | H  | R  | M  |    |    |    |    |    |
| 9     | H  | Y  | G  | R  | A  |    |    |    |
| 10    | (G)| I  | M  | R  | R  | G  | K  |    |
| 11    | I  | V  | G  | R  | R  | (L)|    |    |
| 12    | K  | F  | T  | R  | A  |    |    |    |
| 13    | L  | I  | P  | R  | A  |    |    |    |
| 14    | M  | T  | R  | K  | R  | M  | (L)|    |
| 15    | N  | F  | A  | R  | M  |    |    |    |
| 16    | N  | H  | L  | R  | K  | A  |    |    |
| 17    | N  | V  | G  | R  | M  |    |    |    |
| 18    | N  | V  | S  | R  | R  | G  |    |    |
| 19    | P  | I  | S  | R  | A  |    |    |    |
| 20    | P  | V  | G  | R  | M  |    |    |    |
| 21    | Q  | R  | G  | R  | A  |    |    |    |
| 22    | R  | L  | L  | R  | S  | V  |    |    |
| 23    | S  | F  | G  | R  | H  |    |    |    |
| 24    | S  | L  | R  | G  | S  | (L)|    |    |
| 25    | T  | V  | L  |    |    |    |    |    |
| 26    | (G)| V  | A  | R  | R  | V  | K  |    |
| 27    | V  | I  | A  | R  | S  | N  |    |    |
| 28    | V  | N  | T  | K  | S  | G  |    |    |
| 29    | R  | A  | G  |    |    |    |    |    |
| 30    | V  | R  | R  | G  | S  | (L)|    |    |
| 31    | V  | R  | R  | G  | A  |    |    |    |
| 32    | T  | R  | V  | R  | A  | K  |    |    |

**FIG. 3.** Primary sequence of the randomized hexamer found in phage that were selective substrates of t-PA. Peptide sequences have been shifted to the left or right to align corresponding subsites of each individual hexamer sequence. Amino acids in parenthesis are flanking residues from the gene III fusion protein.

have previously presented detailed, kinetic analysis of the cleavage by t-PA of four substrates from such an optimized library (18). Subsequent analysis, using the dot blot assay, revealed that these four selected substrates were also cleaved efficiently by u-PA and therefore were not highly selective for t-PA versus u-PA. In fact, one of these substrates, SRARKA, was actually cleaved more rapidly by u-PA than by t-PA. The remaining three selected substrates were cleaved slightly more rapidly in the dot blot assay by t-PA than by u-PA. To determine the precise extent of selectivity toward t-PA exhibited by these substrates, we measured the catalytic efficiency of both t-PA and u-PA for hydrolysis of peptides containing these three selected amino acid sequences (Table I, peptides V–VII). In contrast to the 13–47-fold t-PA/u-PA selectivity exhibited by substrates isolated from the subtraction library, the t-PA/u-PA selectivity of these substrates identified using standard substrate phage display was 2.5–4.9-fold. These data verify the selectivity of these substrates identified using standard subtractive phage display techniques greatly facilitate elucidation of subsite occupancy that actively influence both positive and negative determinants of specificity. While substrate phage display identifies the most labile substrates for a particular enzyme, substrate subtraction identifies the most selective substrates for the enzyme.

With enzymes for which the most labile substrates are also selective substrates (e.g. u-PA), the two protocols may yield similar results. Even in this situation, however, both techniques remain useful because it seems unlikely that the standard protocols will yield the most selective substrate or that substrate subtraction will yield the most labile substrate. In contrast, with enzymes like t-PA, where the most labile substrates are not also highly selective substrates, the two methods produce distinct results. For example, the most labile substrate sequence identified for t-PA, PGGRSA, was isolated by standard substrate display protocols and not by substrate subtraction. This sequence was also cleaved efficiently by u-PA in contrast to the 13–47-fold t-PA/u-PA selectivity. This observation emphasizes the importance of subtle determinants of substrate specificity and the ability of subtraction libraries to identify combinations of amino acids that maximize selectivity. The most selective P4-P1 sequence, FRGR (Table I, peptide X), occurred in the substrate subtraction library but not in the substrate phage library, although the later protocol included one additional round of screening and approximately 40% more positive phage.

For reasons described above, substrate subtraction techniques greatly facilitate elucidation of subsite occupancy that contributes primarily to the specificity, rather than to the catalytic efficiency, of substrate hydrolyses. Data presented in this study suggest that appropriate occupancy of P3 for t-PA substrates is one example of this type of interaction. In physiological contexts, such subtle specificity determinants may be essential. For example, Erythrina trypsin inhibitor inhibits t-PA but not u-PA. Although the molecular basis of this specificity has not been extensively studied, it is likely of significance in these contexts.

---

1. H. K. Ke, G. S. Coombs, K. Tachias, D. R. Corey, and E. L. Madison, unpublished observations.

2. H. K. Ke, G. S. Coombs, K. Tachias, D. R. Corey, and E. L. Madison, unpublished observations.
remains obscure, it is intriguing that the *Erythrina* inhibitor contains an arginine in the P3 position (34).

Although standard substrate phage display protocols can be used successfully to isolate selective substrates, this task will frequently prove arduous due to the necessity of isolating, sequencing, and functionally characterizing an inordinately large number of phage. We find, however, that a straightforward modification of these protocols to include a subtraction step substantially enhances the efficiency of this process and therefore significantly increases the rate at which highly selective substrates can be discovered. For example, we performed two independent substrate phage display experiments with t-PA using standard protocols. Characterization by dot blot analysis of a total of 105 individual phage clones isolated in the final round of selection in the two experiments revealed that either 6.7% (first experiment) or 4.2% (second experiment) of these phage clones were t-PA-selective substrates. In contrast, the second order rate constants for inhibition of t-PA by wild type PAI-1 were 1.6 \times 10^6 \text{ M}^{-1} \text{ s}^{-1} and 1.0 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}, respectively, 1.4 \times 10^6 \text{ M}^{-1} \text{ s}^{-1} and 1.0 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}, respectively, 1.6 \times 10^6 \text{ M}^{-1} \text{ s}^{-1} and 1.8 clones (substrate subtraction) to obtain a single t-PA-selective phage.

**Minimization of the Selective Peptide Substrates**—The kinetic analysis described above was performed using substrate peptides that were 14 amino acids in length. To confirm that the specificity we observed was inherent in the selected hexaepitope sequences, we examined the kinetics of cleavage of short peptides containing only sequences found within selected hexaepitope sequences. Tetrapeptide IX and pentapeptide X, for example, were cleaved 57 or 78 times, respectively, more efficiently by t-PA than by u-PA and were therefore actually more selective than the 14-mer peptides (II–IV). Furthermore, tetrapeptide IX and pentapeptide X were significantly more labile substrates of t-PA than any of the 14-mer peptides (II–IV). Thus, compared with the 14-mers, the tetra- and pentapeptides not only maintained specificity but also acquired increased activity. These data confirm the proposed status of the P3 and P4 residues as specificity determinants for substrates of t-PA and u-PA, suggest a particularly prominent role of the P3 residue in this capacity, and demonstrate that highly selective substrates of t-PA can be created by appropriate occupancy of the P3–P1’ subsites alone. These observations may also provide a firm basis for the rational design of highly selective, small molecule inhibitors of t-PA. Although hydrolysis of the selective, small peptide substrates by t-PA is characterized by $K_m$ values in the millimolar range, it has been routinely observed that the introduction of a transition state bond geometry adjacent to the P1 residue of a protease substrate can create either a reversible inhibitor whose affinity for the target protease is enhanced by 3–6 orders of magnitude or an irreversible inhibitor with an impressive second order rate constant for inhibition of the target protease ($>10^9 \text{ M}^{-1} \text{ s}^{-1}$) (for a review, see Ref. 35). Similar results using the substrates identified in this study would create highly selective t-PA inhibitors, with affinities in the low nanomolar range, that might be further improved by subsequent, systematic chemical modification.

**Design and Characterization of Variants of PAI-1 That Are Selective for t-PA**—To test the prediction, based on analysis of the cleavage of peptide substrates, that the P3 residue can mediate the ability of an inhibitor to discriminate between t-PA and u-PA, we performed site-specific mutagenesis of PAI-1, the primary physiological inhibitor of both t-PA and u-PA. Three variants of PAI-1 were produced and characterized: a variant in which the P3 serine was converted to an arginine residue, a variant in which the P4 valine was replaced by a glutamine residue, and a double mutant containing both of these substitutions. Kinetic analysis of the inhibition of t-PA and u-PA by these variants of PAI-1 proved consistent with conclusions drawn from the previous experiments utilizing peptide substrates. The second order rate constants for inhibition of t-PA and u-PA by wild type PAI-1 were $1.6 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ and 1.9 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}, respectively (Table II). Thus, wild type PAI-1 exhibits approximately 11.9-fold specificity toward u-PA. By contrast, the second order rate constants for inhibition of t-PA and u-PA by the P3 arginine variant of PAI-1 were, respectively, 1.4 \times 10^6 \text{ M}^{-1} \text{ s}^{-1} and 1.0 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}, an approximately 170-fold reversal in specificity (Table II). This large alteration in specificity was achieved without sacrificing activity toward the target enzyme; the P3 arginine mutation reduced activity of PAI-1 toward u-PA by a factor of approximately 190 without significantly affecting reactivity toward t-PA.

Individual mutation of the P4 valine of wild type PAI-1 to a glutamine residue had no effect on the rate of inhibition of either t-PA or u-PA. As suggested by the predominance in the substrate subtraction libraries of substrates containing both large P3 and large P4 residues, however, the P4 glutamine mutation did not affect the selectivity of the inhibitor.

---

**Table I**

Comparison of $k_{\text{cat}}$, $K_m$, and $k_{\text{cat}}/K_m$ for the hydrolysis by t-PA or u-PA of peptides selected using substrate subtraction or standard substrate phage display protocols

| Substrate | t-PA | u-PA | t-PA/u-PA selectivity |
|-----------|------|------|----------------------|
| (Pn... P2, P1 | P1’... Pn) | | |
| | $k_{\text{cat}}$ | $K_m$ | $k_{\text{cat}}/K_m$ | $k_{\text{cat}}$ | $K_m$ | $k_{\text{cat}}/K_m$ | |
| I | KKSPPGR | VGGSVAR | 0.0043 | 15,000 | 0.29 | 0.003 | 3400 | 0.88 | 0.33 |
| II | LGSSQGRG | KALE | 0.99 | 2390 | 430 | 0.02 | 2180 | 9.2 | 47 |
| III | LGSSQAR | GALE | 0.073 | 1410 | 52 | 0.004 | 970 | 4.0 | 13 |
| IV | LGSSQYR | SGLE | 1.29 | 4010 | 322 | 0.059 | 3800 | 15 | 21 |
| V | GGGWLR | RGLVPE | 2.0 | 10,000 | 200 | 0.32 | 4000 | 80 | 2.5 |
| VI | GGGYR | RGLVPE | 1.6 | 7200 | 220 | 0.20 | 4400 | 45 | 4.9 |
| VII | GGGFR | SGLYP | 3.3 | 2200 | 1500 | 0.71 | 2200 | 320 | 4.7 |
| VIII | YGR | S | 23.7 | 6000 | 3950 | 2.6 | 11,400 | 230 | 17 |
| IX | RGR | K | 15.3 | 16,600 | 922 | 0.76 | 46,500 | 16 | 57 |
| X | FRR | K | 12.2 | 9800 | 1240 | 0.14 | 8600 | 16 | 78 |

*Positional nomenclature of subsite residues. Arrows denote the position of peptide bond hydrolysis.*
increase the t-PA selectivity of the P3 arginine variant of PAI-1. The second order rate constants for the inhibition of t-PA and u-PA by the P4 glutamine, P3 arginine double mutant of PAI-1 were \(1.4 \times 10^6 \text{ M}^{-1} \text{s}^{-1}\) and \(2.9 \times 10^6 \text{ M}^{-1} \text{s}^{-1}\), respectively (Table II). While maintaining full activity toward t-PA, then, the double mutant exhibited an approximately 600-fold enhanced t-PA/u-PA selectivity compared with wild type PAI-1 and a 3.5-fold greater t-PA selectivity than the P3 arginine variant of PAI-1. The absolute t-PA/u-PA selectivity of wild type PAI-1, the P3 Arg single mutant, and the P3 Arg, P4 Gln double mutant was approximately 0.08, 14, and 48, respectively.

### Contribution of Structural Studies to Understanding Restricted Specificity at P2 and the Critical Role of the P3 and P4 Residues in Mediating Specificity

At the time these studies were initiated, no structural information was available for the protease domain of either t-PA or u-PA. During preparation of this manuscript, however, both of these important structures were described (2, 3). The t-PA structure contained a benzamidine molecule, which occupied the S1 pocket, and the u-PA structure included a covalently bound, tripeptide chloromethyl ketone that filled the S1, S2, and S3 sites. These structural reports have provided a clear molecular basis for our observation that both t-PA and u-PA display a strong preference for glycine at the P2 subsite of substrates. In both enzymes, the side chain of residue 99 (chymotrypsin numbering), a tyrosine in t-PA and a histidine in u-PA, severely restricts the size of the S2 pocket. Modeling of even an alanine residue into this subsite produces a slight steric conflict with residue 99 of the enzyme, and larger residues at P2 create extensive clashes.

Another important observation of the structural studies is that the major distinction between the substrate binding cleft of t-PA and u-PA occurs in the region corresponding to the aryl binding site of thrombin (2, 3, 36). In u-PA this pocket is partially filled by an insertion of two amino acids (threonine 97A, leucine 97B; chymotrypsin numbering) that is absent in t-PA. Consequently, the aryl binding site is significantly larger in t-PA than in u-PA. In addition, unlike Ser\(^{172}\) of u-PA, Arg\(^{174}\) of t-PA extends toward, and partially occupies, the aryl binding site. However, because Arg\(^{174}\) appears highly flexible and mobile (2), the extent to which this residue actually influences the aryl binding site of t-PA in solution remains uncertain.

Depending on the precise binding mode, the aryl binding site can interact with any of several residues of a particular substrate or inhibitor. If the substrate or inhibitor adopts a canonical conformation within the active site cleft, the P4 residue will occupy the aryl binding site (37). However, the most extensively studied aryl binding site, that of thrombin, interacts with the P9 residue of the physiological substrate fibrinogen (38). In addition, the P2 glycine of substrates selected in this study, with its greatly expanded range of allowed conformations compared with other natural amino acids, could allow the P3 residue to interact with the aryl binding site. Consequently, the structural studies indicating that the major differences between the active site cleft of the two proteases occur in the aryl binding site are consistent with our substrate phage studies, which demonstrate that the P3 and P4 residues are the primary determinants of the ability of a substrate to distinguish t-PA and u-PA.

Differences at position 217, a leucine in t-PA and an arginine in u-PA, may also contribute to specificity differences observed in this study for the two enzymes. In the u-PA structure, Arg\(^{217}\) adopts an unusual conformation that allows formation of a salt bridge with the P3 glutamic acid of the bound inhibitor. Clearly, a similar juxtaposition of Arg\(^{217}\) and the P3 arginine residue of a substrate or inhibitor would create an unfavorable electrostatic interaction that would be absent when the same substrate or inhibitor interacted with t-PA. However, if Arg\(^{217}\) of u-PA adopted a more commonly observed conformation, this residue would extend into solvent and could be located at a significant distance from the P3 arginine residue. Thus, the extent to which Arg\(^{217}\) of u-PA and Leu\(^{217}\) of t-PA contribute to specificity distinctions between the two enzymes remains an open question.

As discussed above, insights gained from this study and from very recent structural studies do allow the identification of candidate residues that may mediate important specificity distinctions between t-PA and u-PA. The absence of key structural information regarding the binding mode(s) of a particular selective substrate or inhibitor to both t-PA and u-PA, however, precludes a definitive description at atomic resolution of the mechanisms by which the specificity observed in our studies has been achieved.

### Limitations of Structural Studies and Contribution of Substrate Phage and Substrate Subtraction Techniques to Understanding Protease Specificity

It will not be possible to understand enzyme catalysis and specificity without extensive structural information. On the other hand, it is very unlikely that the recent structural studies alone could have been used to predict the most selective substrates of either t-PA or u-PA. One difficulty described above is that the structural data for the proteases does not necessarily provide information regarding the precise binding mode and backbone conformation of an individual substrate, a critical detail that determines which, if any, region of the substrate will interact with specific regions (e.g. the aryl binding site) of the enzyme.

Another limitation is that the static structures do not necessarily provide information regarding the role of molecular dynamics in enzyme catalysis and specificity. Consequently, contributions to substrate specificity arising from subtle differences in dynamic properties of related enzymes may not be detected by the x-ray structures. The development of a detailed, molecular understanding of enzyme mechanism will therefore require information from studies that utilize a wide variety of techniques of molecular biology, biochemistry, and biophysics. Substrate phage display and substrate subtraction libraries can contribute to these efforts by providing a powerful, combinatorial approach to the identification of key determinants of substrate reactivity and specificity.

### Conclusion

The rational design of small molecule inhibitors as therapeutic agents is often complicated by the necessity of discriminating between closely related enzymes (39). We determined...
onstrate here that appropriate selections of substrate phage libraries can achieve this discrimination. Substrate subtraction phage libraries are likely to provide substrates that can distinguish between any two distinct proteases, and there is no theoretical reason why multiple proteases could not be used in the subtraction step to achieve even greater specificity. Moreover, it should be possible to prepare both substrate and substrate subtraction libraries as described above for any enzymes that can use peptides or proteins as substrates. These techniques might be adapted to protein kinases, for example, by using antibodies against phosphoserine, phosphothreonine, or phosphotyrosine during the selection of substrate phage. Consequently, the construction and characterization of substrate and substrate subtraction libraries may make substantial contributions to the rational design of highly specific, small molecule inhibitors of selected enzymes, a problem of paramount importance during the development of new therapeutic agents. In addition, by revealing specificity determinants that might otherwise remain obscure, these libraries will provide key insights into the molecular basis of specificity for a variety of important enzymes.

Acknowledgments—We thank Drs. Guy Salvesen, Steve Kent, Dave Loskutoff, and Mark Ginsberg for stimulating discussions and critical review of this manuscript. We thank Lynn Mayfield, Bikash C. Pramanik, and Dr. Clive Slaughter for peptide synthesis and mass spectral analysis and Karen Barker for initial kinetic analysis of the peptide substrates.

REFERENCES
1. Perona, J. J., and Craik, C. S. (1995) Protein Sci. 4, 337–360
2. Lamb, R., Bauer, M., Huher, R., Fischer, S., Rudolph, R., Kohnert, U., and Bode, W. (1996) J. Mol. Biol. 258, 117–135
3. Spraggon, G., Phillips, C., Nowak, U. K., Ponting, C. P., Saunders, D., Dobson, C. M., Stuart, D. I., and Jones, E. Y. (1995) Structure 3, 681–691
4. Collen, D., and Lijnen, H. R. (1991) Blood 78, 3114–3124
5. Madson, E. L. (1994) Fibrinolysis Suppl. 1, 221–236
6. Ganu, V. S., and Shaw, E. (1982) Peptide Protein Res. 20, 421–428
7. Madison, E. L., Coombs, G. S., and Corey, D. R. (1995) J. Biol. Chem. 270, 7558–7562
8. Coombs, G. S., Dang, A. T., Madison, E. L., and Corey, D. R. (1996) J. Biol. Chem. 271, 4461–4467
9. Carmeliet, P., Schoonjans, L., Kieckens, L., Ream, B., Degen, J., Bronson, R., De Vos, R., van den Oord, J. J., Collen, D., and Mulligan, R. C. (1994) Nature 368, 419–424
10. Danz, K., Andreason, P. A., Gronenthal-Hansen, J., Kristensen, P., Nielsen, L. S., and Skriver, L. (1985) Adv. Cancer Res. 44, 139–266
11. Tsirka, S. E., Gualandris, A., Amaral, D. G., and Strickland, S. (1995) Nature 377, 340–344
12. Fukudome, K., and Esmon, C. T. (1994) J. Biol. Chem. 269, 26486–26491
13. Min, H. Y., Doyle, L. V., Vitt, C. R., Zandonella, C. L., Stratton-Thomas, J. R., Shuman, M. A., and Rosenberg, S. (1996) Cancer Res. 56, 2428–2433
14. Seeds, N. W., Williams, B. L., and Bickford, P. C. (1995) Science 270, 1992–1994
15. Ossowski, L. (1988) Cell 52, 321–328
16. Shapira, R. L., Duquette, J. G., Roses, D. F., Nunes, I., Harris, M. N., Camingo, H., Wilson, E. L., and Rifkind, D. B. (1996) Cancer Res. 56, 3597–3604
17. Matthews, D. J., and Wells, J. A. (1993) Science 260, 1113–1117
18. Ding, L., Coombs, G. S., Strandberg, L., Navre, M., Corey, D. R., and Madison, E. L. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 7627–7631
19. Smith, M. M., Shi, L., and Navre, M. (1995) J. Biol. Chem. 270, 6440–6449
20. Jameson, G. W., Roberts, D. V., Adams, R. W., Kyle, S. A., and Ellmore, D. T. (1973) Biochem. J. 131, 107–117
21. Urano, T., Urano, S., and Castellino, F. J. (1988) Biochem. Biophys. Res. Commun. 150, 45–51
22. Alpert, A. J. (1990) J. Chromatogr. 499, 177–196
23. Taylor, J. R. (1992) An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements, University Science Books, Mill Valley, CA
24. Tucker, H. M., Mottonen, J., Goldsmith, E. J., and Gerard, R. D. (1995) Nat. Struct. Biol. 2, 442–445
25. Zoller, M. J., and Smith, M. (1984) DNA 3, 479–488
26. Kunkel, T. A. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 488–492
27. Sancho, E., Tonge, D. W., Hockney, R. C., and Booth, N. A. (1994) Eur. J. Biochem. 224, 125–134
28. Chase, T., and Shaw, E. (1987) Biochem. Biophys. Res. Commun. 29, 508–514
29. Olson, S. T., Bock, P. E., Krassman, J., Shore, J. D., Lawrence, D. A., Ginsburg, D., and Bjork, I. (1995) J. Biol. Chem. 270, 30067–30077
30. Holmes, W. E., Lijnen, H. R., and Collen, D. (1987) Biochemistry 26, 5133–5140
31. Beatty, K., Rietz, J., and Travis, J. (1980) J. Biol. Chem. 255, 3931–3934
32. Madison, E. L., Goldsmith, E. J., Gerard, R. D., Gething, M. J., Sambrook, J. F., and Bassel-Duby, R. S. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 5350–5353
33. Madison, E. L., and Sambrook, J. F. (1993) Methods Enzymol. 223, 249–271
34. Teixeira, A. V., Dowdle, E. B. D., and Botes, D. P. (1994) Biochem. Biophys. Acta 1217, 25–28
35. Bagin, R., Shi, Y.-Q., Wu, Y.-Q., and Abeles, R. H. (1996) Biochemistry 35, 3551–3553
36. Bode, W., Turk, D., and Sturzebecher, J. (1990) Eur. J. Biochem. 185, 175–183
37. Bode, W., Turk, D., and Karshikov, A. (1992) Protein Sci. 1, 426–471
38. Ni, F., Meinwald, Y. C., Vasquez, M., and Scheraga, H. A. (1989) Biochemistry 28, 3094–3105
39. Crooke, S. T. (1996) Nat. Biotechnol. 14, 238–241