

Origins of the specificity of tissue-type plasminogen activator

LI DING*†, GARY S. COOMBS†‡, LEIF STRANDBERG§, MARC NAVRE*, DAVID R. COREY‡¶, AND EDWIN L. MADISON§¶

*Department of Biochemistry, Affymax Research Institute, Santa Clara, CA 95051; ‡Department of Pharmacology and Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, TX 75235; and §Department of Vascular Biology, The Scripps Research Institute, La Jolla, CA 92037

Communicated by Alfred G. Gilman, University of Texas Southwestern Medical Center, Dallas, TX, April 24, 1995

ABSTRACT The role of subsite interactions in defining the stringent substrate specificity of tissue-type plasminogen activator (t-PA) has been examined by using an fd phage library that displayed random hexapeptide sequences and contained 2×10^8 independent recombinants. Forty-four individual hexapeptides were isolated and identified as improved substrates for t-PA. A peptide containing one of the selected amino acid sequences was cleaved by t-PA 5300 times more efficiently than a peptide that contained the primary sequence of the actual cleavage site in plasminogen. These results suggest that small peptides can mimic determinants that mediate specific proteolysis, emphasize the importance of subsite interactions in determining protease specificity, and have important implications for the evolution of protease cascades.

A wide variety of critical biological processes (1–3) depend on specific cleavage of individual target proteins by serine proteases. Enzymes capable of catalyzing this uniquely selective proteolysis have evolved while retaining high homology to related nonselective proteases. For example, tissue-type plasminogen activator (t-PA), a trypsin-like enzyme that catalyzes the rate-limiting step of the endogenous fibrinolytic cascade, has only one known substrate *in vivo*, a single bond (Arg⁵⁶⁰-Val⁵⁶¹) in the proenzyme plasminogen (4). Part of the specificity of t-PA for plasminogen is mediated by formation of a ternary complex between t-PA, plasminogen, and the cofactor fibrin (5). Even in the absence of fibrin, however, t-PA remains specific for cleavage of plasminogen. This stringent specificity is an inherent property of the protease domain of t-PA (6), in spite of its high homology to trypsin, an archetypal nonselective protease (7).

Trypsin activates native plasminogen poorly, exhibiting only 10–30% of the catalytic efficiency of t-PA (6, 8). By contrast, trypsin cleaves small peptides containing the primary sequence of the activation site in plasminogen 29,000–200,000 times more rapidly than t-PA. The relative specificity of these two related enzymes for the same primary sequence, in two distinct structural contexts, therefore, varies by 5–6 orders of magnitude. A detailed understanding of the molecular basis of this dramatic difference between t-PA and trypsin is likely to provide important insights into the evolution of specificity among chymotrypsin-like enzymes and may facilitate development of the ability to tailor enzymes with desired substrate and inhibitor specificity, an advance that should have wide-ranging applications.

The extremely low activity of t-PA toward peptides containing the target sequence within plasminogen may indicate that elements of the active site of t-PA are not optimally formed or aligned to accomplish cleavage of the scissile peptide bond in the target sequence. Alternatively, optimal subsite specificity of t-PA may be matched poorly by the plasminogen target sequence. Both of these hypotheses suggest an important role for a secondary site(s) of interaction between t-PA and

plasminogen during catalysis, either to induce a conformational change in t-PA or to properly position a suboptimal sequence in the active site of t-PA. A third hypothesis, which does not necessarily require secondary sites of interaction between t-PA and plasminogen, is that t-PA recognizes an unusual, constrained conformation of the plasminogen target sequence that cannot be significantly populated by small, linear peptides. Only the second of these three hypotheses unequivocally predicts the existence of small peptides that could be efficiently cleaved by t-PA. Consequently, to distinguish among these hypotheses, we screened a library of random hexamers for peptide sequences that could be hydrolyzed by t-PA. A peptide containing one of the selected amino acid sequences was cleaved by t-PA 5300 times more efficiently than a peptide containing the primary sequence of the actual cleavage site in plasminogen.

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. Poly(vinylidene difluoride) membranes were acquired from Millipore. K91 (F⁺) and MC1061 (F⁻) strains of *E. coli* were provided by Steve Cwirla (Affymax). Monoclonal antibody (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 Gramsch Laboratories (Schwabhausen, Germany). t-PA was obtained from Genentech.

Construction of the phage vector fAFF1-tether C (fTC) and the random hexapeptide library fAFF-TC-LIB has been described (9). Control substrate phage fTC-PL and fTC-PAI-II were constructed by hybridizing the single-stranded oligonucleotides 5'-TCGAGCGGTGGATCCGGTACTGGTCTACTGGTCATGCTCTGGTAC-3' and 5'-CGCCACCTAG-GCCAGGACCAGCACAAACCACGAGAC-3' (fTC-PL) or 5'-TCGAGCGGTGGATCCGGTCTGGTCTGGTCTGTTGTTGGTCTCTGGTAC-3' and 5'-CGCCACCTAG-GCCATGACCAGCATGACCAGTACGAGAC-3' (fTC-PAI-II) and then ligating the annealed, double-stranded products into the *Xho* I/*Kpn* I-cut vector fTC. All the constructs were first transformed into MC1061 by electroporation and then transferred into K91. Control nonsubstrate phage fTC-Bad (insert coding for peptide NPVEPA) was provided by M. Smith (9).

Measurement of Enzyme Concentrations. Concentrations of trypsin and t-PA protease domain were measured by active site titration with 4-methylumbelliferyl *p*-guanidinobenzoate (10) using a Perkin-Elmer LS 50B luminescence fluorometer. Titrations of trypsin were performed in 100 mM NaCl/20 mM CaCl₂/50 mM Tris-HCl, pH 8.0. Because CaCl₂ reduces the solubility of t-PA, titrations of t-PA and the isolated protease

domain of t-PA were performed in 150 mM NaCl/10 mM Tris-HCl, pH 7.5.

Phage Selection. Substrate phage panning of monovalent phage was designed and reported by Matthews and Wells (11). We developed an alternative protocol by using polyvalent phage (refs. 9 and 12; Fig. 1) and by performing selection of positive phage as follows: 2×10^{10} phage (20 μ l of the fTC-LIB-N6 library) in assay buffer (50 mM Tris-HCl, pH 7.4/100 mM NaCl/1 mM EDTA/0.01% Tween 20) was digested overnight at 37°C by concentrations of t-PA varying from 10 to 100 μ g/ml. The total volume of the digestion was 250 μ l. The reaction mixture was then placed on ice, bovine serum albumin was added to 0.1%, and 100 μ g of mAb 179 and 10 μ g of mAb 3-E7 were added. After 30 min on ice, 100 μ l of Pansorbin cells was added, and the reaction mixture was rotated at 4°C for 1 hr. The mixture was microcentrifuged for 2 min, and the supernatant was recovered to repeat the Pansorbin adsorption. The final supernatant was amplified overnight in *E. coli* K91 cells. A small aliquot of the final supernatant solution was also used for titrating on K91. Individual clones were selected from the titer plates, grown in 4-ml cultures, and confirmed as "substrate phage" by a dot blot analysis technique (9) that demonstrated their epitopes were lost after digestion with t-PA.

Dot Blot Assay of Phage Proteolysis. To precipitate phage from selected clones, 20 μ l of 20% polyethylene glycol/2.5 M NaCl was added to 100 μ l of culture supernatant. After incubat-

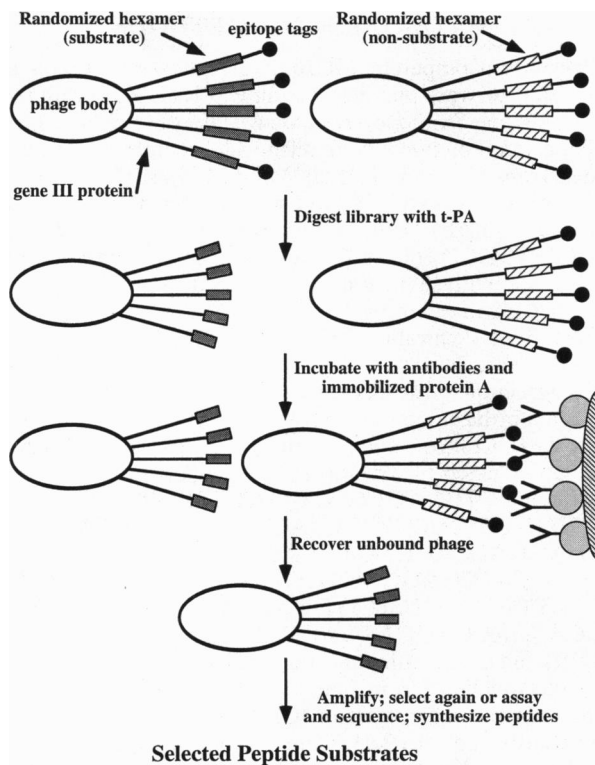


FIG. 1. Outline of the tether phage selection (9). Shown are diagrams of the fTC phage (for clarity, not drawn to scale). The gene III protein extends from phage body from the C to the N terminus. At the N termini are the peptide epitopes and the random hexamer domains that may or may not act as substrates, depending on the encoded peptide sequence. The phage were treated in solution with t-PA, and those carrying substrate sequences (Left) were cleaved in the hexamer site, whereas nonsubstrates (Right) were not. The entire digest was treated with antibodies to the tethers and captured by using a resin that carries protein A. The protein A-antibody-phage complexes were precipitated by centrifugation, and the phage that were not bound by the antibodies remained in solution and could be recovered for amplification. After amplification, the phage could be analyzed or subjected to further rounds of selection.

ing on ice for 30 min, the precipitated phage were microcentrifuged for 5 min. The supernatant was aspirated and the phage were resuspended in 40 μ l of TBS (50 mM Tris-HCl, pH 7.4/150 mM NaCl). In wells of a flexible microtiter plate (Falcon), 12 μ l of the phage suspension was mixed with 18 μ l of a protease/buffer mixture (various amounts of enzyme in assay buffer) and the plate was incubated overnight at 37°C. At the end of each incubation period, 70 μ l of TBS was added. The samples were spotted onto a nitrocellulose filter with a dot blotter (Bio-Rad) and the filter was blocked with 5% nonfat milk in TBS-T (TBS/0.05% Tween 20) for 30 min to 1 hr. The filters were washed three times with TBS-T and then incubated for 1 hr with mAb 179 at 1.9 μ g/ml. The washes were repeated and the filters were probed with a 1:5000 dilution of goat anti-Mouse IgG horseradish peroxidase conjugate. After 1 hr, the washes were repeated and the filter was stained as directed using the Amersham Western enhanced chemiluminescence kit. The results of a typical experiment are displayed in Fig. 2. Phage PL, PAI, and BAD are the control phage described above. Phage 1-30 is a nonsubstrate phage from the first round of selection, while phage 2-9 is a substrate phage isolated during the second round of selection.

Kinetics of Cleavage of Synthetic Peptides by Trypsin and t-PA. Peptides were synthesized by solid-phase synthesis on an Applied Biosystems peptide synthesizer (model 430A) using fluorenylmethoxycarbonyl (Fmoc) chemistry. Each peptide was desalted by reverse-phase HPLC. Kinetic data for cleavage of synthetic peptides was obtained by incubating each peptide at concentrations varying from 0.01 to 10 mM in the presence of 640 nM t-PA or 1–5 nM trypsin. Cleavage of the peptides by trypsin or t-PA was monitored at 220 nm by reverse-phase HPLC (Rainin, Woburn, MA) with a Microsorb 5- μ m 300-Å reverse-phase column (4.6 mm \times 25 cm) (Rainin) using a 0–100% gradient of 0.1% trifluoroacetic acid in doubly distilled water and 0.08% trifluoroacetic acid in 95% acetonitrile/5% doubly distilled water (14). Data were interpreted by Eadie-Hofstee analysis. The identity of hydrolyzed peptide fragments was determined by mass spectral analysis.

RESULTS

A polyvalent fd phage library (fAFF-TC-LIB) that displayed random hexapeptide sequences and contained 2×10^8 independent recombinants was expressed in the vector fAFF-tether C (9). Phage from the peptide library displayed an N-terminal extension from phage coat protein III (pIII) containing a randomized region of 6 amino acids fused to pIII followed by a 6-residue linker sequence (SSGGSG) and the

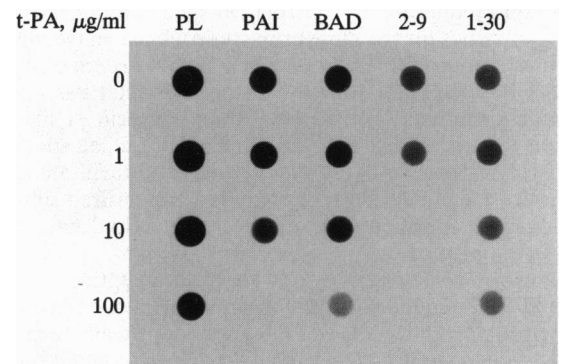


FIG. 2. Dot blot proteolytic analysis of phage clones. Phage from individual clones were prepared, digested with t-PA at the indicated concentrations, dotted onto nitrocellulose, and detected as described. Loss of positive staining indicates cleavage of the epitopes (in the random hexamer region) from the phage gene III protein (9). PL, plasminogen insert; PAI, PAI-I control; BAD, fTC-BAD control; 2-9 and 1-30, positive and negative clones, respectively, from early rounds of screening.

epitopes for mAb 179 and mAb 3-E7. Because t-PA does not digest the PIII sequence, the loss of antibody epitopes from the phage surface upon incubation with t-PA requires cleavage at the random 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 (Fig. 1).

After five rounds of selection to enrich and amplify phage that display sequences that are readily cleaved by t-PA, 44 individual phage clones were identified as substrates and analyzed by DNA sequencing. The amino acid sequence of the random hexamer encoded by each phage clone is listed in Table 1. Each of the hexamers contains a different primary sequence, and, as expected from the trypsin-like P1 specificity of t-PA, each hexamer contains at least one arginine residue. Forty-two of the 44 substrate sequences match the 4-amino acid consensus sequence (G/A)RX(A/G), where X can be any amino acid but is most often R, at three of four positions; both exceptions, clones 1-35 and 1-56, were isolated from the initial round of screening before the stringency of the selection was increased and match the consensus sequence at two of four positions. Twenty-seven of the 44 positive clones match the consensus at all four positions. The unique sequences found most often among the substrate phage are ARRA (7 clones), GRRR (4 clones), GRKA (3 clones), GRRG (3 clones), and GRSA (3 clones). Less stringent conservation among the substrate sequences is observed at the P3 position—25 of the 44 clones contain a hydrophobic residue in this position and 12 clones contain a P3 arginine residue.

Four peptides were chosen for detailed kinetic analysis (Table 2, IV–VII) and compared to hydrolysis of a reference peptide (peptide I) containing the P3–P4' sequence of plasminogen, a series of residues that fall within a disulfide-linked loop in the native protein. The peptide containing the plasminogen sequence is a very poor substrate for t-PA. The most active peptide (peptide VII) by contrast is cleaved by t-PA at 5300-fold higher catalytic efficiency. This increase in catalytic efficiency is primarily due to an increase in k_{cat} , suggesting that optimized subsite interactions serve to lower the energy of the transition state rather than the ground state. Because peptide VII exhibited sequence similarity at P2, P1, P1', and P2' with the reactive center loop of the plasminogen activator inhibitors PAI-I and PAI-II (15), we analyzed peptides II and III, consisting of the P8–P7' primary sequence of the two specific inhibitors. These peptides were better substrates than the plasminogen peptide but were cleaved 12–165 times less efficiently by t-PA than the most reactive phage peptide (peptide VII).

The most labile hexapeptide (peptide VII) spans the P4–P2' subsites and differs from the plasminogen-derived

Table 1. Amino acid sequences of the hexapeptide in 44 isolated substrate phage clones

Clone	Amino acid sequence	Clone	Amino acid sequence
1-11	RGRGAG	5-4	SYLRRA
1-17	LGRRRS	5-6	LKGRRR
1-21	IRGRWA	5-7	RGRRAR
1-22	AGRSFA	5-9	SRARKA
1-33	(G)GRTDFW	5-24	FTGRDI
1-35	IKRNIT	5-26	PYSRMA
1-38	PRGRGG	5-28	TFARRA
1-42	NKYARQ (L)	5-31	KAIGRM (L)
1-56	YHIKRS (L)	5-34	WLGRRG
2-1	VLR SAR	5-38	GFARRA
2-3	QYLRYG	5-40	VARRAA
2-5	KDARRA	5-43	(SG)RYARSA
2-8	WLPARRA	5-44	RGRSAG
2-9	ILRAAY	5-52	(SG)RYARRL
2-11	LRGRTA	5-57	MRGRRG
3-1	GFARRA	5-61	TVMRRA
3-7	FGRRRT	5-63	PFGRSA
3-11	TRARRA	5-64	MRLRRA
4-1	QRGRKA	5-66	RRGRRA
4-6	IVRRRAE	5-67	YIGRRG
4-9	VARRAA	5-68	LGRKAT
5-2	IGRRAQ	5-72	QLGRKA

The first digit of the clone number indicates the round of selection in which that clone was isolated. Peptide sequences have been shifted to the right or left in order to align the clones to a consensus sequence. Residues in parentheses are flanking residues from the gene III fusion protein.

peptide I at four of these six positions (i.e., P4, P3, P1', P2'). To assess the effect of altered subsite occupancy at each of these four positions, kinetic analysis of the hydrolysis by t-PA of the 16 peptides listed in Table 3 was performed. Individual replacement of residues at the P4, P3, P1', and P2' positions in plasminogen-derived peptide I with the corresponding residue from selected peptide VIIa increased catalysis by factors of 4.1 (X), 11 (XI), 27 (XII), and 9 (IX) (Table 3). Similarly, individual replacement of residues at the same positions of selected peptide VIIa with the corresponding amino acid from plasminogen decreased the efficiency of t-PA-catalyzed hydrolysis of the selected peptide by factors varying from 1.9 to 16.2 (XVI, XVII, XVIII, and XIX). In addition, replacement of the P1' serine residue in peptide VIIa with a valine significantly reduced both the K_m , by a factor of 10.5, and the k_{cat} , by a factor of 35, of t-PA for the peptide, suggesting that the P1' valine in plasminogen may play a direct role in binding to t-PA. Finally, t-PA displays a

Table 2. Comparison of k_{cat} , K_m , and k_{cat}/K_m values for hydrolysis by t-PA of peptides derived from a substrate phage screen or modeled after the native cleavage sequence in plasminogen or the reactive loops of PAI-I and PAI-II

	Substrate (Pn, P3, P2, P1 ↓ P1', P2', P3', Pn')*	k_{cat} , s ⁻¹	K_m , μM	k_{cat}/K_m , M ⁻¹ s ⁻¹	Ratio rel to I†
I	YKKSPGR ↓ VVGGSKY	0.0043	15,000	0.29	1
II	TAVIVSAR ↓ MAPEEII	0.25	1,900	130	448
III	TGGVMTGR ↓ TGHGGPQ	0.036	3,900	9.2	32
IV	GGSGSRAR ↓ KALVPE	0.061	22,000	2.8	9.6
V	GGSGWLGR ↓ RGLVPE	2.0	10,000	200	700
VI	GGSGYIGR ↓ RGLVPE	1.6	7,300	220	750
VII	GGSGPFGR ↓ SALVPE	3.3	2,200	1500	5300
VIII	TAVIPFGR ↓ SAPEEI	1.4	2,300	610	2100

Error in these determinations was 3–16%. Residues derived from the randomized phage sequences are in boldface. Arrows denote locations for cleavage sites between P1 and P1' used to determine the initial rate of cleavage by t-PA. Identity of the hydrolyzed peptide fragments, and therefore also the cleavage site in the initial peptide, was determined by mass spectral analysis.

*For positional nomenclature of subsite residues, see ref. 15. The peptide bond is cleaved between P1 and P1'.

†Ratio of k_{cat}/K_m relative to k_{cat}/K_m for peptide I.

Table 3. Comparison of k_{cat} , K_m , and k_{cat}/K_m for hydrolysis by t-PA or trypsin of peptides of varying similarity to the most efficient selected substrate (VII) and the native plasminogen sequence (I)

	Substrate (Pr, P3, P2, P1 ↓ P1', P2', P3', Pr')*	t-PA				Trypsin			
		k_{cat} , s ⁻¹	K_m , μM	k_{cat}/K_m , M ⁻¹ s ⁻¹	k_{cat}/K_m , rel to I†	k_{cat} , s ⁻¹	K_m , μM	k_{cat}/K_m , M ⁻¹ s ⁻¹	k_{cat}/K_m , rel to I†
I	YKKS <u>P</u> GR ↓ VVGGS <u>K</u> Y	0.0043	15,000	0.29	1.0	25	790	3.2 × 10 ⁴	1.0
IX	YKKS <u>P</u> GR ↓ V <u>A</u> GG <u>S</u> KY	0.058	22,000	2.6	9.0	28	420	6.6 × 10 ⁴	1.0
X	YK <u>K</u> PPGR ↓ VVGGS <u>K</u> Y	0.033	28,000	1.9	6.5	26	600	4.3 × 10 ⁴	1.3
XI	YK <u>K</u> S <u>F</u> GR ↓ VVGGS <u>K</u> Y	0.066	21,000	3.1	11	33	81	4.1 × 10 ⁵	12.8
XII	YKKS <u>P</u> GR ↓ <u>S</u> VGG <u>S</u> KY	0.048	6,200	7.8	27	130	520	2.5 × 10 ⁵	7.8
XIII	YK <u>K</u> P <u>F</u> GR ↓ VVGGS <u>K</u> Y	0.41	8,700	48	170	120	640	1.9 × 10 ⁵	5.9
XIV	YK <u>K</u> P <u>F</u> GR ↓ <u>S</u> A <u>G</u> GG <u>S</u> KY	5.2	9,200	570	2,000	130	200	6.5 × 10 ⁵	15
VIIa	GGSG <u>P</u> GR ↓ SALV <u>P</u> EE	4.2	3,100	1350	4,700	220	103	2.1 × 10 ⁶	66
XVI	GGSG <u>S</u> GR ↓ SALV <u>P</u> EE	1.9	2,700	720	2,500	370	230	1.6 × 10 ⁶	50
XVII	GGSG <u>P</u> GR ↓ <u>S</u> VL <u>V</u> PEE	0.52	1,200	430	1,500	340	160	2.1 × 10 ⁶	66
XVIII	GGSG <u>P</u> GR ↓ <u>V</u> AL <u>V</u> PEE	0.12	290	400	1,400	50	80	6.0 × 10 ⁵	21
XIX	GGSG <u>P</u> GR ↓ SALV <u>P</u> EE	0.26	3,200	82	290	220	320	7.0 × 10 ⁵	22
XX	GGSG <u>S</u> GR ↓ SALV <u>P</u> EE	0.22	3,700	60	210	60	330	1.8 × 10 ⁵	5.6
XXI	GGSG <u>P</u> GR ↓ <u>V</u> VL <u>V</u> PEE	0.04	1,500	27	82	61	97	6.3 × 10 ⁵	20
XXII	GGSG <u>P</u> GR ↓ SALV <u>P</u> EE	0.07	4,400	16	55	111	370	3.0 × 10 ⁵	9.3
XXIII	GG <u>S</u> S <u>P</u> GR ↓ <u>V</u> VGL <u>V</u> PEE	0.0033	1,700	2	6.7	64	1,500	4.2 × 10 ⁴	1.3

Error in these determinations was 4–22%. Underlined residues are amino acid alterations relative to I or VIIa. Peptides contained an additional C-terminal glutamic acid to facilitate synthesis.

*For positional nomenclature of subsite residues, see ref. 15. Arrows denote position of peptide bond hydrolysis. The peptide bond is cleaved between P1 and P1'.

†Ratio of k_{cat}/K_m relative to k_{cat}/K_m for peptide I.

>10-fold enhanced preference for arginine over lysine at the P1 position relative to trypsin. Trypsin cleaves peptide VIIa, containing arginine at P1, ≈7.1 times more efficiently than peptide XXII, containing lysine at P1; this ratio increases to 85 for t-PA.

For comparison, kinetic constants for cleavage by trypsin of the 16 peptides listed in Table 3 were also measured. As expected, trypsin cleaved the peptides much more efficiently than did t-PA. The relative reactivities of the peptides toward trypsin vary, confirming that trypsin also has a preference for specific subsite interactions. The differences in reactivity among the peptides, however, are much more pronounced for catalysis by t-PA than by trypsin (Fig. 3). For example, the most labile selected peptide (VIIa) is cleaved at 60-fold greater efficiency by trypsin than plasminogen peptide I; the corresponding ratio for t-PA is ≈4700. Similarly, peptide

XIV, containing the P4–P2' sequence from the selected peptide and the plasminogen flanking sequence, is cleaved 1965 times more efficiently by t-PA but only 15 times more efficiently by trypsin than the plasminogen peptide I.

To assess the possibility that the enhanced catalysis of peptide VIIa was due to its stable occupation of a labile conformation in solution, circular dichroism spectra were obtained at 260–200 nm. Spectra were examined at 4°C, 16°C, and 37°C. These data were consistent with the existence of the peptide as a random coil and showed no tendency toward secondary structure formation.

DISCUSSION

Trypsin cleaved all of the peptides examined in this study much more efficiently than did t-PA but with a qualitatively similar

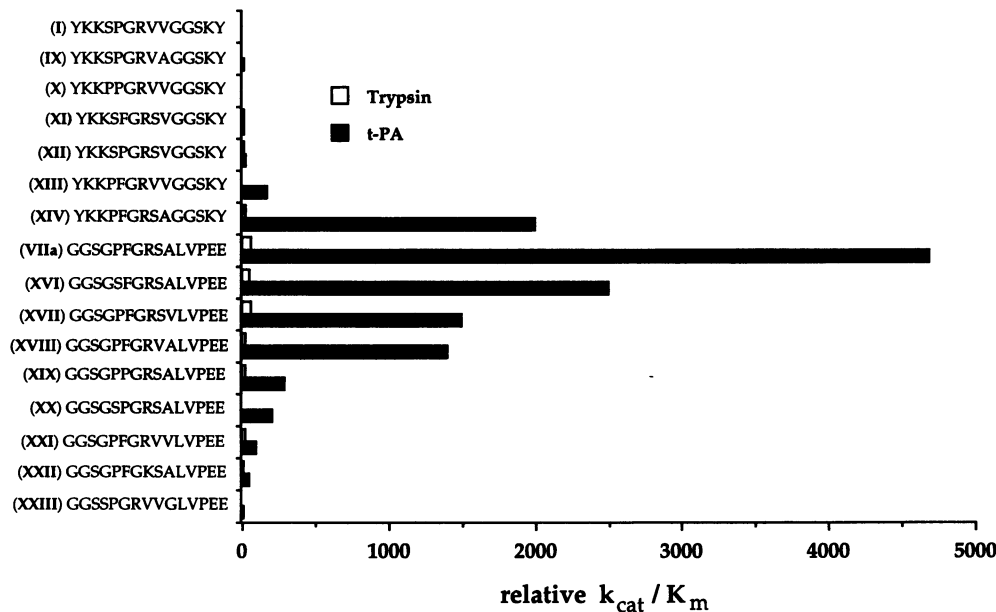


FIG. 3. Catalytic efficiencies for cleavage of peptides, relative to that of peptide I, by t-PA and trypsin. x axis, k_{cat}/K_m relative to peptide I for trypsin (open bars) and t-PA (solid bars); y axis, peptides listed in Table 3.

preference for specific subsite interactions, indicating that trypsin and t-PA possess similar determinants for optimal subsite interactions. The differences in reactivity among the peptides, however, are much more pronounced for catalysis by t-PA than by trypsin (Fig. 3). This observation suggests that, compared to trypsin, both generally poor efficiency for peptide cleavage and high discrimination between peptide substrates contribute to the extremely stringent specificity of t-PA and that t-PA has evolved to restrict substrate specificity by significantly enhancing subsite discrimination already present in trypsin.

The normal physiological substrate of t-PA, Glu-plasminogen, is hydrolyzed by t-PA, in the absence of a cofactor, with a catalytic efficiency of $920 \text{ M}^{-1}\text{s}^{-1}$ (5), while the most labile selected peptide is cleaved with a catalytic efficiency of $1500 \text{ M}^{-1}\text{s}^{-1}$, suggesting that selected small peptides can mimic determinants that mediate specific proteolysis. The divergence between the selected peptide substrates and the plasminogen sequence suggests that the normal target sequence of a highly specific protease is not necessarily a reasonable "lead compound" for the design of small substrates or inhibitors. It is interesting to compare the sequence of the "bait region" of the specific inhibitors of t-PA, PAI-I and PAI-II, with the consensus selected sequence and the target sequence in the substrate plasminogen (Table 2). Similarities at the P2 and P1 positions suggest that optimal subsite occupancy and binding modes for all three classes of compounds—protein substrates, protein inhibitors, and peptide substrates—are likely to be very similar or even identical at these two locations. Divergence of the amino acid sequences at P1' and particularly at P2', on the other hand, suggests that optimal subsite occupancy and precise binding modes at these positions may vary significantly among the three classes of molecules.

Our data suggest a greater similarity in subsite interactions at P1' and P2' between protein inhibitors and selected peptide substrates than between either of these molecules and the protein substrate, possibly reflecting the existence of strict functional constraints on the P1' and P2' residues in plasminogen. After activation cleavage, these two residues move into the "activation pocket" of plasmin and form important contacts that mediate development of an active conformation by the oxyanion hole and the primary specificity pocket. Consequently, unlike the corresponding residues in protease inhibitors or peptide substrates, the P1' and P2' residues of plasminogen are not free to evolve only toward optimal interaction with the S1' and S2' subsites of t-PA. Similar considerations apply to all chymotrypsin-like enzymes that are activated by other proteases. Our data demonstrate that this is not a trivial obstacle to the evolution of efficient protease cascades. Replacement of the P1' and P2' residues of the most active peptide substrate with the P1' and P2' residues of plasminogen reduced the reactivity of the substrate by a factor of ≈ 50 (Table 3, peptides VIIa and XXI, respectively). During the evolution of protease cascades, this obstacle has apparently been overcome, and activation cleavage of zymogens has been accelerated by development of secondary sites of interaction between enzyme and substrate and by involvement of cofactors.

The observation that t-PA can efficiently hydrolyze selected peptide substrates suggests that its substrate selectivity is not based primarily on a conformational change of the enzyme, upon binding to plasminogen, that properly aligns residues involved in binding and/or catalysis. Moreover, the failure to detect an ordered structure when studying our most labile selected substrate peptide in solution argues that productive recognition of a substrate by t-PA does not necessarily require an unusual, constrained conformation that is normally inaccessible to small, linear peptides. In addition, as discussed

above, our data indicate a clear discrepancy between the optimal subsite occupancy for t-PA and the primary sequence of the actual cleavage site in plasminogen. This discrepancy explains the low activity of t-PA toward plasminogen-based peptides and strongly suggests that the specificity of t-PA for plasminogen derives from the evolution of optimal secondary interactions of both proteins with the cofactor fibrin and probably also direct, secondary interactions between t-PA and plasminogen that remain uncharacterized.

Our data suggest that t-PA has evolved to limit the catalysis that can be generated through optimal subsite occupation—even the most labile selected substrate is cleaved 500-fold more rapidly by trypsin than by t-PA. In addition, the remaining determinants of productive binding (e.g., arginine at P1) have become much more stringent in t-PA than in trypsin (Fig. 3). The combination of these two properties renders t-PA catalytically inert toward proteins of the bloodstream, an important consideration since, unlike most serine proteases, the circulating proenzyme form of t-PA is enzymatically active.

Most studies of protease specificity have focused on the functional significance of modifications of the S1 pocket (16). Comparison of t-PA with trypsin (Table 3) indicates that t-PA has also evolved to narrow its specificity to arginine through mutations at S1. However, our data demonstrate, in addition, that t-PA has also evolved extremely effective mechanisms to use non-S1 subsite interactions to dramatically increase its discrimination among potential substrates. Understanding this discrimination may facilitate the rational alteration of non-S1 subsites and the design of highly specific and therapeutically useful enzymes.

We wish to thank Lynn DeOgny, Bikash C. Pramanik, and Dr. Clive Slaughter for peptide synthesis and mass spectral analysis; Matt Smith for helpful experimental advice; and Caroline Davies for DNA sequence analysis. We thank Guy Salvesan and Robert Fletterick for many stimulating discussions and also for critical review of this manuscript. This work was supported in part by National Institutes of Health Grants RO1 HL52475 and PO1 HL31950 (E.L.M.). G.S.C. was supported by National Institutes of Health Biophysics Predoctoral Training Program Grant 5P32GMO8287. D.R.C. is an Assistant Investigator with the Howard Hughes Medical Institute.

1. Neurath, H. (1984) *Science* **224**, 350–357.
2. Neurath, H. (1989) *Trends Biochem. Sci.* **14**, 268–271.
3. Neurath, H. (1986) *J. Cell. Biochem.* **32**, 35–49.
4. Madison, E. L. (1994) *Fibrinolysis* **8**, 221–236.
5. Hoylaerts, M., Rijken, D. C., Lijnen, H. R. & Collen, D. (1982) *J. Biol. Chem.* **257**, 2912–2919.
6. Madison, E. L., Coombs, G. S. & Corey, D. R. (1995) *J. Biol. Chem.* **270**, 7558–7562.
7. Pennica, D., Holmes, W. E., Kohr, W. J., Harkins, R. N., Vehar, G. A., Ward, C. A., Bennett, W. F., Yelverton, E., Seeberg, P. H., Heyneker, H. L., Goeddel, D. V. & Collen, D. (1983) *Nature (London)* **301**, 214–221.
8. Wohl, R. C., Summaria, L. & Robbins, K. O. (1980) *J. Biol. Chem.* **255**, 2005–2013.
9. Smith, M., Shi, L. & Navre, M. (1995) *J. Biol. Chem.* **270**, 6440–6449.
10. Jameson, G. W., Roberts, D. V., Adams, R. W., Kyle, S. A. & Ellmore, D. T. (1973) *Biochem. J.* **131**, 107–117.
11. Matthews, D. J. & Wells, J. A. (1993) *Science* **260**, 1113–1117.
12. Cwirla, S. E., Peters, E. A., Barrett, R. W. & Dower, W. J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6378–6382.
13. Sambrook, J. F., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
14. Corey, D. R. & Craik, C. S. (1992) *J. Am. Chem. Soc.* **114**, 1784–1790.
15. Carrell, R. & Boswell, G. (1986) in *Proteinase Inhibitors*, eds. Barrett, A. J. & Salvesan, G. S. (Elsevier, New York), pp. 403–420.
16. Perona, J. J. & Craik, C. S. (1995) *Protein Sci.* **4**, 337–360.