THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 277, No. 29, Issue of July 19, pp. 26623–26631, 2002

Printed in U.S.A.

Computer-assisted Mutagenesis of Ecotin to Engineer Its Secondary Binding Site for Urokinase Inhibition*

Received for publication, March 29, 2002
Published, JBC Papers in Press, April 16, 2002, DOI 10.1074/jbc.M203076200

Martha C. A. Laboissière‡, Malin M. Young§¶, Rilva G. Pinho‡, Stephen Todd§‖, Robert J. Fletterick, Irwin Kuntz§**, and Charles S. Craik§**‡‡**

*From the ‡Departamento de Ciências Farmacêuticas, Faculdade de Ciências da Saúde, Universidade de Brasília, Campus Universitário Darcy Ribeiro, 70910-900 Brasília, DF Brasil, §Departments of Pharmaceutical Chemistry, and **Biochemistry and Biophysics, University of California, San Francisco, California 94143-0446*

Inhibitors of urokinase-type plasminogen activator (uPA) were selected *in vitro* from two ecotin phage-display libraries to study the effect on binding of amino acid substitutions at critical positions 108, 110, 112, and 113 within the 100s loop (RNKL, respectively, in wild type ecotin). The first, a focused library, was the result of a computation-assisted approach using the three-dimensional structure of the ecotin-trypsin complex to guide the modeling of amino acid substitutions predicted to increase affinity for uPA. The second, a complete library, allowed for all substitutions at the above identified positions. The consensus sequences selected from the focused, and complete libraries were RRWS and R(R/N)QL, respectively. Inhibition constant determinations showed ecotin variants containing these sequences to be similarly potent (K_i = 1–2 nM). These substitutions were combined with previously identified substitutions in another critical region of ecotin. One of these combinations (D70R/M84R/RRQL) is the tightest (K_i = 50 pM) ecotin variant inhibitor of uPA. The blending of combinatorial methods and computer algorithms designed to predict stronger binders has allowed us to obtain protein derivatives that exhibit greatly increased affinity for a predetermined target. This technology can be applied to select for enhanced binding interactions at protein-protein interfaces and accelerate the process of protease inhibitor development.

The formation of protein-protein complexes is an essential element of virtually every cellular process. These complexes may be permanent or transient, associating and dissociating as conditions require. Enzyme-protein inhibitor complexes and activator complexes are archetypal examples of transient complexes (1, 2). The biochemical study of numerous enzyme-inhibitor complexes has led to the assembly of structural and mechanistic data bases, providing further impetus for using these complexes as models to investigate fundamental biochemical and biophysical principles of protein-protein recognition.

The interaction of urokinase-type plasminogen activator (uPA)[1] with the macromolecular serine protease inhibitor, ecotin, is the basis for an enzyme-inhibitor complex model system that is particularly amenable to genetic manipulation. uPA is a serine protease originally found to convert plasminogen to plasmin (3), thereby playing an active role in extracellular proteolysis, cell migration, and tissue remodeling (4). As a result of its potential role in cancer metastasis and tumor invasion (5), uPA has become an important target for anti-cancer drug development (6).

Ecotin is a macromolecular serine protease inhibitor found in the periplasm of *Escherichia coli* (7). This highly malleable, dimeric, macromolecular serine protease inhibitor has been modified to enhance inhibition of selected proteases (8–11). Ecotin has four surface loops that participate in tetramer formation through their interaction with two bound protease molecules. The interaction involves two binding sites on each ecotin monomer. The "primary binding site" composes the reactive loop of ecotin (also referred to as the 80s loop (residues 81–86)) and the supporting 50s loop (residues 52–54). These loops cooperate to bind the active site of target proteases. The "secondary binding site" includes two surface loops of ecotin: the 60s loop (residues 67–70) and the 100s loop (residues 108–113), which bind to the C-terminal region of the target protease (Fig.1). Whereas the conformation of the 80s loop is canonical, at least for the central six amino acids around the cleaved site, the conformation of the secondary binding site is less well characterized. The secondary binding site is unique to ecotin and has not been observed in other protease-inhibitor complexes. The combined buried surface area within an ecotin-protease tetramer is ~2,800 Å^2 (10).

We have shown previously (9) that a high affinity inhibitor of uPA can be isolated from a library of phage-displayed ecotin variants harboring randomized amino acid residues at positions 84 and 85. These amino acid residues mimic the corresponding P1 and P1′ residues of an uPA substrate. We have further demonstrated that the side chain of the P1 residue has a cooperative rather than an independent role in determining the protease specificity of an ecotin variant. In addition, the loss of function associated with alanine substitutions in the 60s and 100s loops supports a model in which these loops provide the critical binding energy necessary for formation of the stable tetramer complex and the resulting protease inhibition. Experimental data have established the essential role of the secondary binding site for ecotin-protease complex stability and revealed a dynamic and interdependent relationship between the

* This work was supported by National Institutes of Health Grants DK39304 (to R. J. F.) and GM31497 (to I. D. K.), by National Science Foundation Grant MCB9604379, National Institutes of Health Grant CA72006, and a Daiichi Research Center grant (to C. S. C.). 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.

¶ Present address: Sandia National Laboratories, MS 9951, P. O. Box 969, 7011 East Ave., Livermore, CA 94551-0969.

‖ Present address: Incyte Genomics, Inc., 3160 Porter Dr., Palo Alto, CA 94304.

‡‡ To whom correspondence should be addressed. Tel.: 415-476-8146; Fax: 415-502-8298; E-mail: craik@cgl.ucsf.edu.

[1] The abbreviations used are: uPA, urokinase-type plasminogen activator; LmuPA, low molecular mass uPA; CALD, computer-assisted library design; WT, wild type.

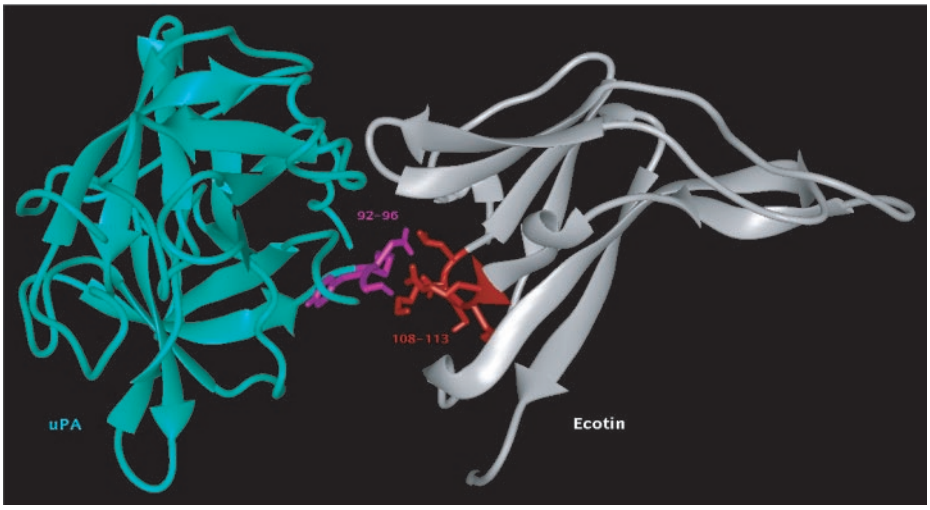

FIG. 1. **Modeled interactions of the interface between the 100s loop of ecotin and a target protease.** The high resolution crystal structure of trypsin-ecotin tetramer (Protein Data Bank code 1SLU) (21) served as the basis for the homology modeling of the ecotin-uPA complexes. The uPA molecule with the backbone modeled as ribbons is colored *green,* and the ecotin molecule is colored *gray*. The side chains of the residues that form protein-protein interactions are shown with the 100s loop in the ecotin molecule colored *red*. The computer graphic model was generated using the program MidasPlus (Computer Graphics Laboratory, University of California, San Francisco).

60s and 100s loops (10).

Several biochemical and genetic approaches have been used to detect and study protein-protein interactions, including library-based methods such as the yeast two-hybrid screen (12, 13) and the phage display method (14–16). Phage display is a highly selective combinatorial methodology that has been successfully applied to evaluate the important interfacial amino acids for a wide range of peptides and proteins (17–19). Because a variant polypeptide is expressed on the surface of the phage particle that harbors the library clone, isolation of phage particles based on physical properties (*e.g.* binding to an immobilized protease), followed by propagation in bacteria, leads to the rapid availability of phage DNA for sequence analysis.

Phage display suffers from several notable limitations, including a physical size limitation for the polypeptide to be studied, the requirement for proteins to be secreted in order to bind the immobilized target molecule, and the use of a bacterial host that may preclude the correct folding or modification of some proteins (2). From an experimental perspective, phage display suffers from an additional and potentially more profound constraint, the limited volume of conformational space that can be explored with a single library (20).

We have investigated the use of a combinatorial computer-assisted library design (CALD) algorithm developed to predict amino acid substitutions in ecotin that would increase its affinity for uPA and thus reduce the complexity of a phage display library designed to study macromolecular recognition between uPA and ecotin. The role of specific amino acid substitutions in the 100s loop of ecotin in the inhibition of uPA was evaluated by comparing variants selected from a fully randomized four-position library to those selected from the focused library. The results indicate that, although the focused phage library was 34-fold smaller than the full library, the consensus sequences obtained from the two libraries were similar (although not identical) and that the K_i values of the selected variants were within 2-fold. Moreover, our results confirm that interactions between the 60s, 80s, and 100s loops of ecotin are important for protease binding and inhibition.

EXPERIMENTAL PROCEDURES

Bacterial Strains and Reagents

E. coli strain JM101 XL-1 Blue F′ and VCSM13 helper phage were obtained from Stratagene (La Jolla, CA). *E. coli* ecotin gene deletion strain IMΔecoJ was derived from strain JM101 and was a generous gift from Dr. Iain Murray. Low molecular mass uPA (LMuPA) and carbobenzoyl-L-γ-glutamyl(α-*ortho-tert-butyl*)-glycyl-arginine-*p*-nitroanilide monoacetate salt were obtained from American Diagnostica (Greenwich, CT). Enzymes and other reagents for the manipulation of DNA were purchased from Promega (Madison, WI) or New England Biolabs (Beverly, MA) and were used following the manufacturers' instructions. The Sequenase version 2.0 DNA sequencing kit was obtained from U. S. Biochemical Corp. Oligonucleotides were synthesized with an Applied Biosystems 391 DNA synthesizer (Foster City, CA). Falcon polystyrene Petri dishes were from Fisher. All other chemicals were of reagent grade or better and were used without further purification.

Computer Modeling of uPA-Ecotin Complexes and Prediction of Amino Acid Substitutions to Increase the Affinity of Ecotin for uPA

Homology modeling of ecotin-uPA complexes was based on the published crystal structure of the trypsin-ecotin complex (Protein Data Bank code 1SLU) (21). Briefly, uPA (Protein Data Bank code 1LMW) was aligned to the trypsin molecule with a root mean square standard deviation of 1.6 Å over 216 amino acids (22). Steric clashes in the resulting uPA-ecotin model complex were relaxed with 100 steps of simulated annealing in the 60s, 80s, and 100s loop regions (defined below) using Sybyl (Tripos, Inc.). Prediction of the amino acid substitutions that would increase the affinity of ecotin for uPA was performed as follows. (i) Residues 108–113 in the 100s loop were changed to glycines. (ii) Force field scoring grids were defined for the 100s loop regions of both the uPA-ecotin complexes (*i.e.* bound ecotin) and ecotin in the absence of uPA (*i.e.* free ecotin) using the program CHEMGRID (23). Free ecotin was used as a reference state when performing mutagenesis calculations, and each site was scored independently. (iii) Each glycine residue at the positions listed in (i) were systematically changed to each of 17 other amino acids (ADEFHIKLMNQRSTVWY) using DOCK 4.0 (24). Initially, neighboring loop amino acids were maintained as glycines. In preparation for the docking calculation, energy scoring grids and sphere sets were generated, characterizing both the ecotin surface (in the absence of uPA) and the ecotin-uPA interface. (iv) For each amino acid substitution we calculated the estimated change in binding energy ($\Delta E = E_X - E_{Gly}$) for each conformer in a rotamer library of side chains using the DOCK force field scoring function. The rotamer libraries were generated by systematically iterating through each rotatable bond in all side chains at 15° increments. (v) The estimated change in binding energy/amino acid was calculated for both the complexed and free ecotin structures. These binding energies were then used to calculate a $\Delta\Delta E$, representing the net change in binding energy/amino acid upon uPA-ecotin complex formation (Fig. 2*A*). For the purposes of these calculations, it was assumed that all sites are independent. This assumption is roughly correct for non-neighboring sites but not for spatially proximal sites. (vi) We addressed this latter issue by calculating pairwise interaction energies for overlapping sites. Pairs of sites that we predicted would interact with one another based on the ecotin-uPA homology model were selected. For site pairs 108–110, 110–112, and 112–113, the 1,000 conformers with the lowest individual $\Delta\Delta E$ scores at each site were rebuilt into the bound and free ecotin structures. Thus $1{,}000 \times 1{,}000 = 10^6$ pairs were rescored in the same manner as the independent sites with the addition of a side chain-side chain interaction energy term (Fig. 2*B*). The top 289-amino acid pairs were evaluated, and the lowest (*i.e.* best) energy score for each amino acid in each position in the list was recorded. (vii) The five amino acids with the lowest predicted binding energies at each position, based on the pairwise energies (or individual energies if there were no

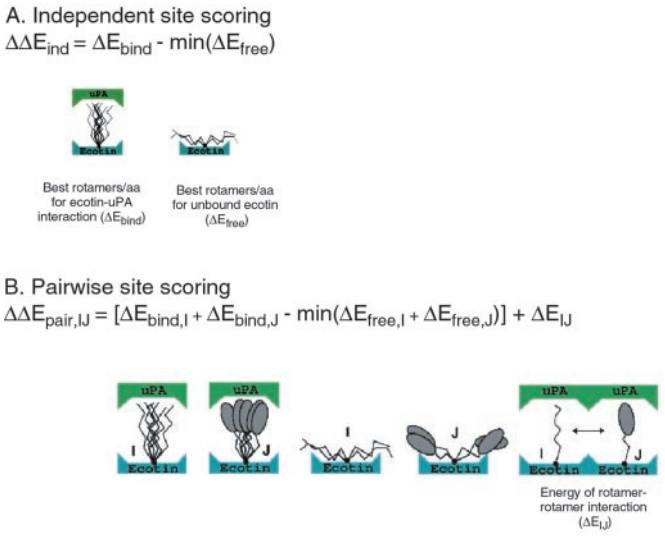

FIG. 2. **Scoring method to predict the amino acid substitutions that might be expected to increase the affinity of ecotin for uPA.** *A,* for each amino acid substitution the estimated change in binding energy per amino acid was calculated for both the ecotin-uPA complexes and free ecotin dimers. These binding energies were then used to calculate a $\Delta\Delta E$, representing the net change in binding energy/amino acid upon formation of the uPA-ecotin complex. *B,* pairwise interaction energies were calculated for overlapping sites by assuming additivity.

neighboring loop residues), were finally selected for the design of a focused phage display library that would be smaller and easier to construct than an exhaustive library. Similar calculations were performed using the ecotin-trypsin complex as a control to determine whether the methodology was able to predict accurately the amino acids in the 60s and 100s loops that contributed to ecotin-trypsin binding.

Phage Display Library Design and Construction

Construction of Combinatorial Computer-aided Design Library (Focused Library)—Degenerate ecotin loop sequences were designed that would encode the five top amino acid residues selected at each loop position. In some cases, codons for additional amino acid residues were also possible due to the nature of primer design. The vector pBSeco-gIIIΔ100, consisting of a combined deletion, frameshift, and stop codon mutation at amino acids 108–113, was the generous gift of Dr. C.-I. Wang. The M84R mutation was introduced using oligonucleotide RK84 (5′-GTTCCCCGGTTAGTACTA(GA)GATGGCCTGCC-3′). The resulting vector, pBSeco-gIII(M84RΔ100), was then used to construct a phage library comprising the codon for the M84R mutation and the randomized codons corresponding to the computer-predicted amino acid substitutions for positions 108, 110, 112, and 113 (Table III). Kunkel mutagenesis was employed using oligonucleotide QY/LIB-R10XC (5′-GTATCTGGGCGATGCTGGCATGCTGDDATACMDKAGCDRKRDKCCGATCGTGGTGTATACG-3′), where D = A, G, and T; M = A, C; K = G and T; and R = A and G. All mutations were confirmed by dideoxynucleotide sequencing.

Construction of the Complete Library—A complete library was designed to include all 20 natural amino acids at positions 108, 110, 112, and 113. Vector pBSeco-gIIIΔ(100-gIII), used as the vector for the insertion of the library, was constructed as follows. Plasmid pBSeco-gIII (M84R/N110R/K112W/L113S), which had the *Sph*I site incorporated at the codon for amino acid 106, was digested with *Aat*II and *Afl*II, filled in using T4 DNA polymerase, and religated.

Plasmid pBSeco-gIIIΔ100 was used in a PCR with 5 pmol of oligonucleotide SJ (5′-GATGCTGGCATGCTGNNSTACNNSAGCNNCNNCCCGATCGTGGTGTACACGCCAGACAATGTAGACGTC-3′) and 25 pmol of oligonucleotide ML21 (5′-GTATCGATAAGCTTAAGACTCCTTATTCCGGAGTATGTTAGCAAACGTAGAAAATA-3′) to generate a library of inserts. The amplified region was gel-purified to remove residual template, digested with *Sph*I/*Hin*dIII, and ligated to the *Sph*I/*Hin*dIII fragment of vector pBSeco-gIIIΔ(100-gIII). The ligation mix was cleaned by phenol/chloroform extraction and ethanol precipitation prior to library construction by transformation (through electroporation) into F′ XL-1 Blue cells.

To ensure the randomness of the resulting libraries, individual clones were isolated, and plasmid DNA was purified using standard procedures. An *Sph*I/*Hin*dIII restriction digest of the sample DNA was used to monitor for the presence of a library member. Positive samples were sequenced to confirm the random incorporation of nucleotides.

Ecotin Library Biopanning and Amplification

Ecotin library biopanning and plate amplification procedures were performed as described (8).

Ecotin Mutagenesis

The plasmids pTacTac-ecotin(M84R/RRWS) and pTacTac-ecotin(M84R/RRQL) were constructed by digesting the plasmids pBSeco-gIII(M84R/RRWS) and pBSeco-gIII(M84R/RRQL), respectively, with *Bam*HI and *Aat*II, gel-purifying the small fragment, and incubating it together with the large *Bam*HI/*Aat*II fragment of the pTacTac-ecotin vector in the presence of T4 DNA ligase.

The plasmids pTacTac-ecotin(RRWS) and pTacTac-ecotin(RRQL) were constructed by amplifying a fragment from the plasmid pTacTac-ecotin by PCR using the oligonucleotides ML36 (5′-TCTAGAACTAGTGGATCCATCGATGC-3′) and ML49 (5′-CTGTTGTAACGCAGCATGCCAGCATCGCCCAGATAC-3′). The fragment was cleaned by phenol/chloroform extraction followed by ethanol precipitation, digested with *Bam*HI and *Sph*I, and then incubated with the large *Bam*HI/*Sph*I fragment of the pTacTac-ecotin(M84R/RRWS or pTacTac-ecotin(RRQL) vector, respectively, in the presence of T4 DNA ligase.

The plasmids pTacTac-ecotin(D70R/M84R/RRWS) and pTacTac-ecotin(D70R/M84R/RRQL) were constructed by amplifying a fragment from the plasmid pTacTac-ecotin(D70R/M84R) by PCR using the oligonucleotide primers ML36 and ML49. The fragment was cleaned by phenol/chloroform extraction followed by ethanol precipitation, digested with *Bam*HI and *Sph*I, and incubated with the large *Bam*HI/*Sph*I fragment of the pTacTac-ecotin(M84R/RRWS) or pTacTac-ecotin(M84R/RRQL) vector, respectively, in the presence of T4 DNA ligase.

The plasmids pTacTac-ecotin(D70R/RRWS) and pTacTac-ecotin(D70R/RRQL) were constructed by amplifying a fragment from the plasmid pTacTac-ecotin(M84R/D70R) or pTacTac-ecotin(M84R/RRQL), respectively, by PCR using the primers ML36 and ML50 (5′-GTCAGTTCCCCGGTTTCAACGATGATGGCCTGCCCGGATGGC-3′). A fragment from the plasmid pTacTac-ecotin was amplified by PCR using the primers ML51 (5′-GCCATCCGGGCAGGCCATCATCGTTGAAACCGGGGAACTGAC-3′) and ML49. In a second PCR, 0.5 μl of one of the ML36/ML50 PCR products and 0.5 μl of the ML51/ML49 PCR products were used along with the primers ML36 and ML49. The resulting fragments were cleaned by phenol/chloroform extraction followed by ethanol precipitation, digested with *Bam*HI and *Sph*I, gel-purified, and incubated with the large *Bam*HI/*Sph*I fragment of the pTacTac-ecotin(M84R/RRWS) or pTacTac-ecotin(M84R/RRQL) vector, respectively, in the presence of T4 DNA ligase.

The plasmid pTacTac-ecotin(D70R) was constructed by amplifying a fragment from the plasmid pTacTac-ecotin(M84R/D70R) by PCR using the primers ML36 and ML50. A fragment from plasmid pTacTac-ecotin was amplified by PCR using the primers ML52 (5′-GGTACTTGACGTCTACATTGTCTGG-3′) and ML49. In a second PCR, 0.5 μl of the products from each PCR were used along with primers ML36 and ML52. The resulting fragment was cleaned by phenol/chloroform extraction followed by ethanol precipitation, digested with *Bam*HI and *Aat*II, gel-purified, and incubated with the large *Bam*HI/*Aat*II fragment of the pTacTac-ecotin vector in the presence of T4 DNA ligase.

Expression of Ecotin and Determination of K_i Values

Wild type and variant ecotin proteins were expressed and purified as described (10). K_i values of wild type and variant ecotin proteins against uPA were determined using methods described previously (9) with the following modifications. (i) The final concentration of ecotin in the reactions was 0–20 nM for ecotin proteins containing the M84R mutation and 0–25 μM for ecotin proteins harboring M84. (ii) The final concentration of LMuPA was 8 nM. (iii) The total volume of the reaction was 990 μl of buffer containing 50 mM NaCl, 50 mM Tris-HCl (pH 8.8), and 0.01% Tween 20. The reaction mixtures were preincubated at room temperature for 5 min to allow them to reach equilibrium. To initiate the reaction, 10 μl of 10 mM (carbobenzoyl-L-γ-glutamyl(α-*ortho*-*tert*-*butyl*)-glycyl-arginine-*p*-nitroanilide monoacetate salt) was then added, and the rate of *p*-nitroanilide formation was measured by monitoring the change of absorption at 410 nm during a 5-min period. The assays were repeated 3–6 times with varying inhibitor concentrations. Data for variants harboring M84 were fitted to the equation derived for traditional Michaelis-Menten inhibitors. Data for variants harboring the M84R mutation were fitted to the equation derived for kinetics of

TABLE I
List of most probable candidates, independent site picks

108	110	112	113
R (−4.48)[a]	K (−4.75)	E (−4.15)	E (−4.27)
Y (−1.73)	M (−4.45)	K (−2.53)	M (−4.14)
W (−1.71)	R (−4.36)	W (−2.52)	K (−3.64)
M (−1.60)	H (−2.12)	D (−2.43)	R (−2.86)
K (−1.43)	N (−1.88)	N (−2.15)	N (−2.84)
[D] (−1.17)	[F] (−1.69)	[F] (−1.85)	[V] (−1.52)

[a] Best interaction energy for residue *X* at position Y assuming independence of sites (for example, in the indicated case *X* = R and Y = 108). The top five ranked residues and the sixth for reference are presented.

TABLE II
List of most probable candidates, pairwise picks

108	110	112	113
R (−7.88)[a]	K (−9.98)	E (−9.92)	E (−9.98)
D (−7.17)	R (−9.59)	D (−8.52)	M (−8.79)
E (−6.49)	M (−8.81)	W (−6.91)	K (−8.01)
M (−5.86)	H (−6.23)	K (−6.83)	N (−7.47)
N (−5.59)	N (−6.12)	N (−6.62)	R (−7.22)
[H] (−5.34)	[F] (−6.00)	[F] (−6.35)	[D] (−7.02)

[a] Lowest interaction energy of a rotamer pair for residue *X* at position Y (for example, in the indicated case *X* = R and Y = 108).

TABLE III
List of most probable candidates, combined results

108	110	112	113
R	K	E	E
D	R	D	M
E	M	W	K
M	H	K	N
N	N	N	R

reversible tight binding inhibitors (25) by non-linear regression analysis. All data analysis was performed using Kaleidagraph (Synergy Software; Reading, PA).

RESULTS

Modeling of Ecotin Amino Acid Substitutions at Positions 108, 110, 112, and 113 Predicted to Increase the Affinity for uPA—Positions 108, 110, 112, and 113 of the ecotin monomer protein were selected for study based on a structure-based approach adopted to dissect and isolate the relative contribution from the distinct binding sites (10).

In this study, the role of the ecotin 100s loop in determining binding specificity was first examined using computer modeling. To predict the preferred substitutions, amino acids were analyzed independently and in pairwise combinations to determine their binding energies based on a homology model of the ecotin-uPA complex. The independent site calculation was performed by iterating through 17 amino acids at one position although keeping the other positions constant as glycines (Table I). The pairwise site calculation was performed for pairs of sites with possible side chain steric overlap. For ecotin, the pairs of sites that were evaluated were 108 and 110, 110 and 112, and 110 and 113 (Table II). The amino acids selected for each position are shown ranked by descending energy score, with the most probable candidates at the top of the list. Both sets of calculations resulted in the same favored amino acids for each position: Arg for 108, Lys for 110, Glu for 112, and Glu for 113; however, there was variation in the subsequent 4 amino acids selected for each position (Table III).

As expected, the amino acids chosen by the independent site calculation tended to have longer or more bulky side chains than those chosen through the pairwise site calculation. In the case of the independent site calculation, the removal of steric constraints imposed by neighboring sites favors the selection of bulky amino acids that maximize van der Waals interactions with the uPA interface. For example, tryptophan and tyrosine were selected through the independent site calculations as good candidates for position 108, whereas they were not selected in the corresponding pairwise site calculations. Results from the independent site calculations indicated that for positions 108 and 112, there was a strong preference for one amino acid, whereas for position 110, three amino acids (Lys, Met, and Arg) appeared to be equally favored. Position 113 showed the most gradual decent in score, indicating that this site was more permissive than others based on our scoring regime.

Upon the inclusion of pairwise rotamer-rotamer interaction energies in the calculation, we observed that neighboring amino acids with complementary charges were favored. Charged amino acids were selected by both the individual and pairwise site calculations as promoting electrostatic interactions with uPA. However, the pairwise site calculation predicted an additional energy bonus resulting from rotamer-rotamer electrostatic interactions in ecotin. An example of this effect can be seen in Table II, row 2. D108, R110, and D112 were considered more promising candidates by the pairwise calculation due to their favorable inter-rotamer electrostatic interactions, whereas in the independent calculation, bulky hydrophobic residues were preferred over these amino acids. As was the case with the independent site calculations, the lowest variation in score is observed at position 113. Positions 110 and 112 exhibit clear preferences for 3 and 2 (Lys, Arg, Met, and Glu, Asp) amino acids, respectively. Unlike the independent site calculation, the pairwise calculation resulted in a weaker selectivity at position 108. This difference is most likely due to the additional energy resulting from rotamer-rotamer electrostatic and van der Waals interactions, which favored smaller, charged side chains. Positions 108 and 110 share a common side chain pocket that is large enough to accommodate amino acids such as tryptophan in the independent site calculation. However, having to share occupancy of this pocket with a side chain from residue 110 in the pairwise calculation restricted the selection of amino acids at 108 to those with less bulky side chains. Because both calculations resulted in the selection of the same 5 amino acids for positions 110, 112, and 113 with variations in the ranking of these residues, merging the results of the two sets of calculations for these three positions was therefore straightforward. To maximize potential side chain packing and minimize steric collisions, the amino acid ranking observed from the pairwise calculation was used for those positions in the design of the focused library. We also opted to use the results of the pairwise calculation to select amino acids for position 108, despite the differences resulting from the independent site calculation.

Design and Construction of Libraries—Preliminary studies suggested that biopanning of the ecotin M84R 100s loop library against uPA would produce a small number of strong binders from the millions of library members (10). To determine the amino acid substitutions required to increase the binding affinity of ecotin for uPA, and to evaluate the ability of computer modeling to predict the substituents, two libraries were designed. Within each library, amino acids 108, 110, 112, and 113 of the ecotin molecule were either changed to a pre-determined subset of amino acids or fully randomized.

The first library (referred to herein as the focused library), based on computer modeling of the most likely strong binders to uPA, incorporated all codons necessary for the translation of ecotins harboring the computer-selected amino acid substitutions. Due to the combination of the nucleotides present at each position of the triplet, coding for additional amino acid residues that were not computer predicted was unavoidable. The second

library (referred to herein as the complete library) was randomized at all four positions.

The total number of possible codons for the focused library was determined by calculating the number of possible codons for each site resulting in 31,104 possible codons. (For position 108 there were $3 \times 3 \times 2 = 18$ codons; for positions 110, 112, and 113 $3 \times 2 \times 2 = 12$ codons; the product of the 4 positions was $18 \times 12 \times 12 \times 12 = 31{,}104$.) By comparison, the complete library allowed for 32^4 possible combinations or 1.05×10^6 codons. The focused library thus represents a 34-fold decrease in complexity.

The focused library was constructed using Kunkel mutagenesis and contained $\sim 1.43 \times 10^5$ individual clones as determined by evaluating the transformation efficiency. The library was shown to be over 99.9% complete as calculated using the equation $N = \ln(1 - p)/\ln(1 - (1/n))$, where N is the number of total individual clones in the library; n is the number of possible combinations, and p is the probability that any clone can be found in the library given a library size. Libraries were characterized for randomness at the nucleotide level as described under "Experimental Procedures." The results showed that the nucleotide sequence was within the correct reading frame and randomized without significant deviation from the expected frequency distribution (results not shown).

Attempts to construct the complete library using the Kunkel method failed to yield a library that was more than 71% complete. An alternative approach, using degenerate oligonucleotides and the PCR, was then used. A truncated cloning vector was created by deleting the 3′ end of the ecotin gene and all of gene III. This resulting plasmid was digested with *Sph*I/*Hin*dIII to serve as a vector fragment to receive the degenerate PCR product obtained using oligonucleotides SJ and ML21, where the degenerate oligonucleotide SJ encoded the library (this fragment contained the 3′ end of the ecotin gene and all of gene III). The absence of gene III in the cloning vector ensured that all phage-expressed ecotin variants were derived from plasmids harboring the degenerate PCR products. To avoid the preference for a subset of members of the SJ oligonucleotide library, based on specific nucleotide sequences, the concentration of SJ was limited to 5 pmol, thus ensuring complete incorporation of the degenerate oligonucleotides and library randomness. 7.53×10^6 clones were obtained representing a 99.9% complete library. Correct incorporation and randomness were verified as above.

Identifying Stronger Ecotin Binding Sequences Against uPA in the Focused and in the Complete Ecotin Phage Libraries—Each library was individually screened against immobilized LMuPA as described (8). The screening involved *in vitro* selection methods in which the LmuPA was bound to the plate, and the library mixture was added. The plate was washed, and the tight binding phage was eluted. This process is referred to as biopanning (or panning). To ascertain that the biopanning was selecting for complete (full-length ecotin) library members, 10 individual clones were isolated from the eluent pool after each round of biopanning, and plasmid DNA was purified using standard procedures. An *Sph*I/*Hin*dIII restriction digest of the sample DNA was used to confirm the presence of library insert. After the second round of biopanning of the focused library, 60% were positives; after the third round, 70%; after the fourth round, 80%; and after the fifth round, 90%, thus clearly demonstrating enrichment for full-length variants (results not shown). The results of sequence analysis of 15 individual clones, obtained after five rounds of biopanning, are presented in Table IV. All clones contained Arg at position 108. Twelve clones contained Arg at position 110. Ten clones contained the sequence RRWS at positions 108, 110, 112, and 113, respectively, demonstrating this to be the preferred sequence.

TABLE IV
Result of biopanning of the focused library

	108	110	112	113
WT	R	N	K	L
1	R	R	W	S
2	R	R	N	G
3	R	R	W	S
4	R	R	N	G
5	R	R	N	N
6	R	R	W	S
7	R	R	W	S
8	R	R	W	S
9	R	R	W	S
10	R	K	C	N
11	R	R	W	S
12	R	R	N	N
13	R	R	W	S
14	R	R	W	S
15	R	R	W	S
Consensus	R	R	W	S

Consistent with the CALD predictions, the ecotin variant selected during biopanning contained an Arg at position 108. The appearance of Arg at position 110 was also predicted by computer modeling, although Asn was the most favored computer-predicted residue at this position. The appearance of Trp at position 112 was the third most favored, computer-predicted residue. The appearance of Ser at position 113 represents a departure from the computer-predicted substitution. In fact, the presence of Ser codons in the library was a fortuity of the mutagenesis procedure, and the substitution of Ser at position 113 was not predicted to contribute to ecotin-uPA complex formation.

The complete library was panned in a similar manner. Restriction analysis of 10 library samples harvested after each biopanning cycle was performed as above to monitor the phage pool throughout the course of biopanning and to ensure that library members were not accumulating deletions. Twenty individual clones were sequenced after four rounds of biopanning.

Sequence analysis of 19 individual clones obtained after four rounds of biopanning with the full library resulted in the variants shown in Table V. As was the case with results obtained with the computer-generated library, the variant RRWS proved to be a prevalent strong binder (25% of total sample). However, sequencing results demonstrated a trend toward the consensus sequence R(R/N)*X*L. Moreover, four variants differed significantly from the consensus sequence. Three additional rounds of biopanning were performed in an effort to obtain a single strongly bound consensus sequence. The results are presented in Table VI.

Results obtained from the sequence analysis of 21 individual clones after seven rounds of biopanning demonstrated a clear preference for the sequence R(R/N)*X*L with RRQL most represented. The RRWS sequence still appeared, although in a proportionately smaller subset of the clones (14%). The only amino acid allowed at positions 108 was Arg consistent with the computer-generated library prediction for this site. At position 110, there is a *quasi* equal preference for Arg and Asn. Position 112 seems the most adaptable, although neutral or positively charged amino acids were preferred. A notable exception is the appearance of Trp at 112 with the concomitant presence of Ser at 113. Finally, Leu is clearly the preferred amino acid residue at position 113 except when position 112 is occupied by Trp.

Determination of Kinetic Values for the Affinity-selected Ecotin Mutants—To test the ability of the affinity-selected ecotin variants to inhibit the activity of uPA, the phage-selected sequences were cloned into a pTacTac-based vector for expression

TABLE V
Result of biopanning of the complete library after four rounds

	108	110	112	113
WT	R	N	K	L
1	R	R	W	S
2	R	R	A	L
3	R	R	W	S
4	R	R	W	S
5	Q	R	R	I
6	R	R	W	S
7	W	D	C	A
8	G	E	A	A
9	R	L	T	L
10	R	N	Q	S
11	R	K	S	L
12	R	N	R	L
13	R	R	W	S
14	L	W	K	G
15	R	N	R	L
16	R	R	Q	L
17	R	R	I	L
18	R	R	V	L
19	R	R	T	L

TABLE VI
Result of biopanning of the complete library after seven rounds

	108	110	112	113
WT	R	N	K	L
1	R	N	R	L
2	R	R	W	S
3	R	N	Q	L
4	R	R	Q	L
5	R	N	R	L
6	R	N	Q	L
7	R	R	W	S
8	R	R	Q	L
9	R	N	G	L
10	R	R	A	L
11	R	R	Q	L
12	R	N	R	L
13	R	N	G	L
14	R	R	Q	L
15	R	R	W	S
16	R	N	Q	L
17	R	N	Q	L
18	R	R	V	L
19	R	R	Q	L
20	R	R	G	L
21	R	N	L	L
Consensus	R	R/N	Q	L

of the variant ecotin polypeptides. Expression and purification of the ecotin variants were performed as described (10). As the biopanning was performed using a library that contained Arg at position 84 (9), all ecotin variants used for initial kinetic analyses harbored the M84R substitution. The variants analyzed were ecotin M84R, ecotin M84R/RRWS (selected from the focused library), ecotin M84R/RRQL (selected from the complete library), and clones that encoded a D70R substitution in addition to the above affinity-selected mutations (*i.e.* ecotin D70R/M84R, ecotin D70R/M84R/RRWS, and ecotin D70R/M84R/RRQL). The D70R mutation was selected during the biopanning of the 60s loop and was shown to decrease the K_i for uPA (8). The results of the inhibition assays are shown in Table VII.

Surprisingly, the mutants selected by the biopanning of the 100s library were less effective inhibitors of uPA activity than ecotin M84R. Only upon incorporation of the D70R mutation did the K_i values drop below the values obtained for the single M84R mutant ecotin.

To assess the validity of the data obtained and to better evaluate the independent contributions of the sites, all mutations were cloned into a vector that encoded Met at position 84 (as in WT ecotin). We reasoned that the higher base-line K_i values obtained using ecotin proteins with M84 would allow a more quantitative analysis of the contribution of each loop to binding.

TABLE VII
K_i values of ecotin variants

Variant	μM	Variant	nM
WT	1.19 ± 0.09	M84R	0.26 ± 0.10
D70R	0.03 ± 0.006	M84R/D70R	0.11 ± 0.08
RRWS	6.99 ± 1.20	M84R/RRWS	2.02 ± 0.17
RRQL	10.60 ± 1.58	M84R/RRQL	1.31 ± 0.28
D70R/RRWS	0.20 ± 0.006	M84R/D70R/RRWS	0.18 ± 0.06
D70R/RRQL	0.03 ± 0.004	M84R/D70R/RRQL	0.05 ± 0.03

As was found within the M84R background, the RRWS and RRQL variants were poorer inhibitors of uPA than WT ecotin. The loss of inhibition efficacy by the addition of RRWS and RRQL was ~6–10-fold regardless of the amino acid present at position 84. However, comparison of K_i values of similar combinations of 60s/100s variants on different position 84 scaffolds yielded several interesting results. The D70R mutation in a WT background was a 45-fold better inhibitor, whereas the D70R substitution resulted in only a 2-fold improvement in background of the more potent M84R inhibitor (Table VII). However, whereas D70R/RRQL was similar to D70R alone in the wild type background (M84), D70R/RRQL was a potent combination in the M84R background demonstrating added effect of combining selections obtained independently at each loop. Ultimately, the quadruple mutant M84R/D70R/N110R/K112Q resulted in the lowest inhibition constant (50 pM) yet obtained against uPA.

DISCUSSION

Ecotin binds to proteases through unique combinations of interactions between the primary and secondary site surface loops. These surface loops are adapted to the unique features of target proteases by virtue of their intrinsic flexibility and C-terminal hinge which permits relative adjustments between the primary and secondary binding sites by altering the dimer interface. The importance of the unique secondary binding site of ecotin in recognizing and differentiating between target proteases is just beginning to be understood. Through mutagenesis of the 60s and 100s loops, we have investigated the significance of amino acid residues within these loops in the selective inhibition of several proteases, such as rat trypsin and human uPA (8). In this study we focused on understanding the role of the 100s loop, and its interaction with the 60s and 80s loops, in inhibiting uPA. In addition, we investigated the value of using computational algorithms to predict preferred substitutions that would increase the affinity of ecotin for uPA and decrease the complexity of the phage display library necessary for our studies. A flow diagram of the strategy used to optimize ecotin for inhibition of uPA is depicted in Fig. 4.

Phage display is a powerful technology that has been used to study and dissect protein-protein interactions (17, 18, 26, 27), including those between ecotin and target proteases (8, 9). However, practical considerations impose constraints on randomized polypeptide libraries consisting of more than 4 amino acid residues. These constraints are less problematic with polynucleotide libraries (*e.g.* SELEX (28–30)). Randomizing 30 positions in an RNA molecule results in $4^{30} \sim 1.2 \times 10^{18}$ permutations, requiring as little as 2 μg of RNA. 20 μg of *in vitro* synthesized RNA (statistically representing each combination 10 times) can easily be screened in a functional assay with an appropriate method for the isolation of variants that possess the desired characteristics. Randomizing only four amino acid

residues in a protein (12 nucleotides) results in 32^4 ($\sim 1 \times 10^6$) permutations and five positions results in 32^5 ($\sim 3.3 \times 10^7$) permutations. However, if these coding regions are in the form of plasmids they must be transformed into bacteria. Obtaining these numbers of transformants from a ligation reaction can be challenging, even using high efficiency, electrocompetent cells.

By using molecular modeling to predict optimal amino acid residues at specific positions in the molecule, the number of permutations that must be screened (and obtained as transformants) is drastically reduced. In the work presented herein, two libraries were designed and constructed. The complete library contained randomized codons for all four positions. The focused library, based on computer modeling of the most likely strong binders to uPA, incorporated all codons necessary for the translation of ecotins harboring the computer-selected amino acid substitutions at each of the four positions. The focused library represents a 34-fold reduction in library complexity obviating any problems associated with transformation efficiency. In terms of time, the construction of the focused library involved 1/6th of the time commitment when compared with the complete library.

Independent consensus sequences were generated from the two libraries panned. The focused library resulted in the sequence RRWS, whereas the complete library resulted in the consensus sequence R(R/N)QL (with RRWS representing 15% of the clones sequenced). The consensus sequences generated from the uPA biopanning with the libraries revealed that, as predicted by the combinatorial mutagenesis algorithm described in this work, an Arg residue at position 108 is essential in high affinity binding against uPA. Arg-108 is predicted to form an electrostatic interaction with Asp-93 in uPA and has the potential to form hydrogen bonds with the terminal hydroxyl group in Ser-95 and backbone carbonyl oxygens in residues Glu-92, Asp-93, and Tyr-94 as shown in Fig. 3. The Arg side chain forms close van der Waals contacts with the side chains of Ser-95 and Asp-93 and with the peptide backbone of residues Asp-93 and Tyr-94. Arg or Asn were selected at position 110, as predicted by the computer algorithm. Arg in particular is stabilized by a charge-charge interaction with Asp-93. Asn is predicted to form a hydrogen-bonding interaction with a side chain oxygen from Asp-93. Arg and Asn both have the potential to form a hydrogen bond with the carbonyl oxygen of residue Glu-92.

Unlike positions 108 and 110, positions 112 and 113 showed a codependency pattern among the selected amino acids. Screening of the full library resulted in a preference for a Leu (as in wild type) at position 113. Under these circumstances, a reasonable degree of flexibility is observed at 112. Our results show that whereas position 112 tended to be occupied by positively charged or neutral amino acids, there was a slight preference for Gln. Position 112 is solvent-exposed, therefore favoring charged and polar amino acids. The selection for basic amino acids at this position may be due, in part, to longer range electrostatic interactions with Asp-93 in uPA and Glu-33 in ecotin. Compared with WT ecotin (which has KL at positions 112–113) the substitution of Lys for Gln may therefore promote a tighter interaction between ecotin and uPA while preserving interactions with the solvent.

Screening of the focused library resulted in the selection of the amino acid pair WS at positions 112–113. This amino acid combination also appeared (albeit at a reduced frequency) on the full library screening, thus supporting the uPA-binding potential of this ecotin variant. The codons for Gln (CAA or CAG) and Leu (UUA, UUG, CUU, CUC, CUA, and CUG) were not present at positions 112 and 113 of the focused library, thus explaining the failure of these residues to appear in the variants selected from that biopanning.

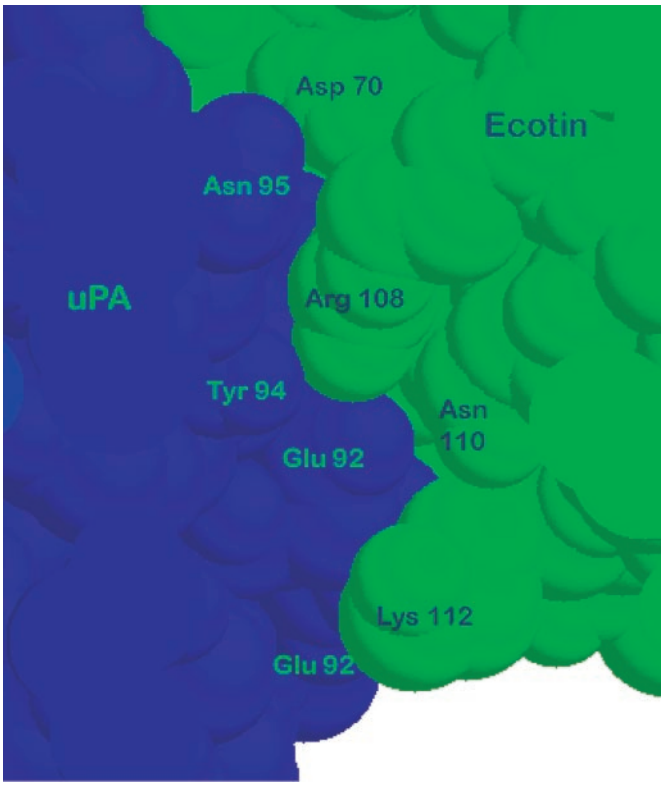

FIG. 3. **A space-filling model drawn using Insight showing the interaction of residues 70, 108, 110, and 112 with the target protease.** The uPA molecule is colored *blue,* and the ecotin molecule is colored *green.* The side chains of the residues involved in protein-protein interaction are shown and labeled. Leu-113 faces the back of the molecule and does not appear. A variant combining mutations at these positions (D70R/N110R/K112Q) and a mutation at residue 84 (M84R) resulted in the tightest inhibitor of uPA activity.

The sequences selected from the biopanning experiments were transferred to expression vectors to evaluate the inhibitory potential of these variants against uPA. The present inhibition study has four major components as shown in Fig. 4. The first is the analysis of proteins that encode the variants selected from the 60s (8) or the 100s loop library biopanning in a protein encoding Met at position 84 (within the primary binding loop). The second is the analysis of ecotin variants (on an M84 scaffold) that combine the variants selected from the 100s loop library biopanning with the variant selected through biopanning of the 60s loop library. The third is the analysis of ecotin variants that combine the variants selected from the 60s or 100s loop with proteins encoding Arg at position 84. Arg was selected upon biopanning of the primary binding site against uPA (9). Finally, the last part is the analysis of ecotin variants that combine the 100s loop selection and the 60s loop selection in a protein encoding Arg at position 84. The inhibition of uPA by ecotins containing M84 is in the micromolar range, as opposed to the nanomolar range for M84R containing ecotin. The purpose of using the less potent inhibitor for these studies was to allow the generation of more widely distributed kinetic data to better dissect the independent contributions of each library-selected amino acid substitution.

The results of the inhibition studies show that the preferred binders selected from each of the 100s loop library biopanning experiments were within 2-fold of each other and 6–10-fold less potent inhibitors of uPA (when compared with the scaffold protein), regardless of the amino acid residue present at position 84. When a Met was present at position 84, RRWS was a better inhibitor than RRQL (consistent with the fact the mod-

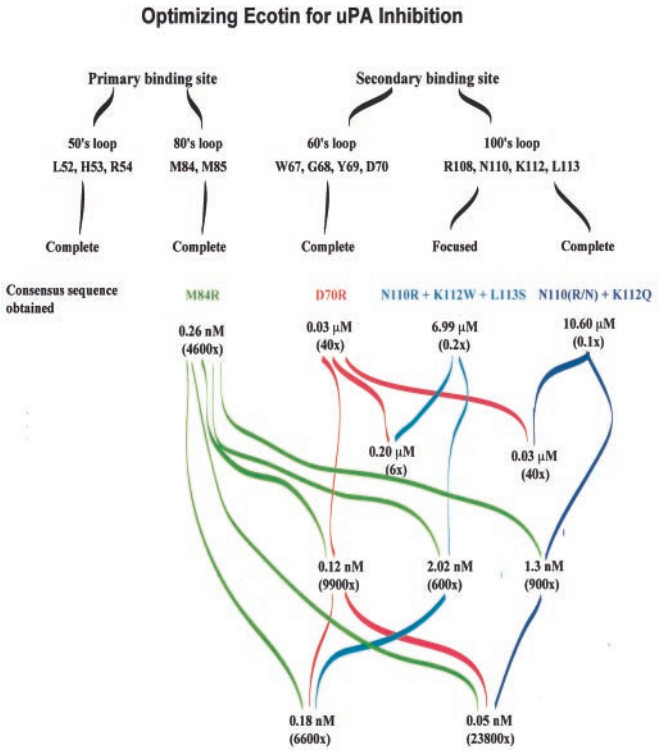

FIG. 4. **Schematic drawing of the procedure developed in our laboratories toward deriving potent and specific uPA inhibitors based on an ecotin scaffold.** Presented are the K_i values obtained for each mutant of ecotin and the fold inhibition improvement over wild type ecotin (K_i = 1.19 μM). The affinities and inhibition for ecotin with mutations in the 50s loop are being evaluated.

eling was done using an ecotin scaffold containing Met at position 84). In contrast, when an Arg was present at position 84, RRQL was a better inhibitor than RRWS (consistent with fact the phage display libraries were constructed on an ecotin scaffold containing an Arg at position 84).

Although we do not clearly understand the relationship between many aspects of binding and catalysis or, in this case, binding and inhibition, the phenomena that better binders are not necessarily better inhibitors can be explained by our understanding of the interactions involved in the binding of the protease to the inhibitor. With respect to the RRWS and RRQL inhibitors, we postulate that as these variants bind more tightly at the 100s loop they distort the complex causing inhibitor misregistration at the substrate-binding site. These results bring out another limitation of the phage display technology when applied to the study and selection of better enzyme inhibitors. The technique in essence selects for better binders that are not degraded upon binding the target enzyme, thus eliminating molecules that are substrates. There is currently no phage display methodology that allows for the immediate selection of molecules that are concomitantly better binders and inhibitors.

Inhibition studies using the variant selected during the biopanning of a 60s loop library and the variants harboring the combined second site selections all resulted in better inhibitors than the starting scaffold ecotins. A pattern was observed in which K_i values for the variants had the decreasing order D70R/RRWS > D70R > D70R/RRQL. Thus, our best inhibitor in either background is that which contains the mutations D70R/N110R/K112Q. The quadruple mutant M84R/D70R/N110R/K112Q resulted in a K_i of 50 pM, the lowest inhibition constant of all mutants analyzed. This result is consistent with previous predictions made by our laboratories which state that the forces involved in ecotin binding to proteases are divided among the four contact sites.

The computational methods we used to design the focused library accurately predicted the amino acid preferences for two of the four sites (108 and 110). Limitations in our scoring function and in how we performed the conformational sampling are the most probable explanations for why we did not correctly predict the amino acid preferences for sites 112 and 113. The force field we used to evaluate binding energies consisted of the non-bonded terms from the AMBER force field, which quantify the energy due to non-directional van der Waals and electrostatic interactions. Directional interactions, such as hydrogen bonding, are not explicitly encoded, and solvation effects are not taken into consideration. The force field favors large, charged amino acids due to the absence of a desolvation penalty; indeed we observed a strong selection for charged amino acids at all four positions. Additionally, the only flexibility allowed in our modeling process was for the side chains of residues 108, 110, 112, and 113. Therefore, small van der Waals clashes with the amino acid side chains or backbone of uPA were not readily resolved.

These limitations directly influenced our predictions. At position 113, for example, the algorithm did not predict the appearance of Ser or Leu and led to the observed preference for large and/or charged amino acids. Similarly, both Glu or Gln at position 112 are favored by the computer algorithm. However, the results from the pairwise calculation show a trend toward selecting amino acid pairs with complementary charges, thus explaining the prediction of a Glu at position 112 next to the Arg at position 110. Again, this is a result of the scoring function used, because such pairs maximize the electrostatic term of the force field. This bias also impacted our prediction at position 113 where 3 of the 5 amino acids predicted were charged. A similar ratio of charged to uncharged amino acids was seen at the other 3 positions, but position 113 (unlike positions 108, 110, and 112) is completely buried and close to 112 at the ecotin-uPA interface in our model. We would therefore expect a substantial desolvation penalty for burying a charged residue at that site, a penalty that was not assessed during the course of our modeling. Incorporating the solvation scoring force field developed by Zou *et al.* (31) may improve the accuracy of predicting residues for buried sites.

Eight members of the ecotin family have been identified from genomic sequences, permitting one to observe the evolutionary selection of amino acids at the 100s loop. The RNKL tetrad at positions 108, 110, 112, and 113 of ecotin is favored among a number of ecotin homologues from other bacteria (*e.g. Salmonella typhi*, *Salmonella typhimurium*, *Salmonella paratyphi*, *Salmonella enteritidis*, *Yersinia pestis*, *Pseudomonas aeruginosa*, *Pseudomonas putida*, *Klebsiella pneumoniae*, and *Trypanosoma brucei*). All of these homologues have Arg at 108 with the exception of the *P. aeruginosa* homologue, which has Cys. All have Lys at 112 except the *Y. pestis* homologue, which has Arg, and the *T. brucei* homologue, which has a Gln (a residue selected in this study). Position 110 is usually an Asn, and position 113 is usually a Leu. The ecotin homologue of *K. pneumoniae* is unusual in that it composes the tetrad RKTA, suggesting a different protease target range.

The findings of this study are significant in two ways. First, we have developed a macromolecular inhibitor of uPA that shows 24,000-fold improved inhibition over WT ecotin. This variant, the strongest inhibitor we have developed, is the result of the blending of combinatorial methods and computer algorithms designed to predict the position of important residues at a protein-macromolecular inhibitor interface. Second, by independently screening a complete library and comparing the re-

sulting variant with those isolated from the focused library, we have proven that CALD can be effective in isolating strong binders/inhibitors of a protein. The success of this approach opens up the possibility of mutagenizing a larger number of amino acid residues in a single library by obviating the need to represent all possible codons at each amino acid coding position. Understanding deficiencies in protein structure-function models, based on results obtained from CALD experiments, will lead to successive improvements in the ability to predict the effect of amino acid substitution on protein function. Ultimately, only substitutions likely to yield the desired biological effects need be encoded at each position in the library.

Acknowledgments—We thank Steve Yang for discussion and John Olson for helpful comments.

REFERENCES

1. Jones, S., and Thornton, J. M. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 13–20
2. Phizicky, E. M., and Fields, S. (1995) *Microbiol. Rev.* **59,** 94–123
3. Gulba, D. C., Bode, C., Runge, M. S., and Huber, K. (1996) *Ann. Hematol.* **73** Suppl. 1, 9–27
4. Andreasen, P. A., Kjøller, L., Christensen, L., and Duffy, M. J. (1997) *Int. J. Cancer* **72,** 1–22
5. Reuning, U., Magdolen, V., Wilhelm, O., Fischer, K., Lutz, V., Graeff, H., and Schmitt, M. (1998) *Int. J. Oncol.* **13,** 893–906
6. Fazioli, F., and Blasi, F. (1994) *Trends Pharmacol. Sci.* **15,** 25–29
7. Lee, H. R., Seo, J. H., Kim, O. M., Lee, C. S., Suh, S. W., Hong, Y. M., Tanaka, K., Ichihara, A., Ha, D. B., and Chung, C. H. (1991) *FEBS Lett.* **287,** 53–56
8. Yang, S. Q., and Craik, C. S. (1998) *J. Mol. Biol.* **279,** 1001–1011
9. Wang, C. I., Yang, Q., and Craik, C. S. (1995) *J. Biol. Chem.* **270,** 12250–12256
10. Yang, S. Q., Wang, C. I., Gillmor, S. A., Fletterick, R. J., and Craik, C. S. (1998) *J. Mol. Biol.* **279,** 945–957
11. Sambrano, G. R., Huang, W., Faruqi, T., Mahrus, S., Craik, C. S., and Coughlin, S. R. (2000) *J. Biol. Chem.* **275,** 6819–6823
12. Fields, S., and Sternglanz, R. (1994) *Trends Genet.* **10,** 286–292
13. Bartel, P. L., and Fields, S. (1995) *Methods Enzymol.* **254,** 241–263
14. Smith, G. P. (1985) *Science* **228,** 1315–1317
15. Cortese, R., Felici, F., Galfre, G., Luzzago, A., Monaci, P., and Nicosia, A. (1994) *Trends Biotechnol.* **12,** 262–267
16. Wells, J. A., and Lowman, H. B. (1992) *Curr. Opin. Biotechnol.* **3,** 355–362
17. Adams, G. P., and Schier, R. (1999) *J. Immunol. Methods* **231,** 249–260
18. Hoogenboom, H. R. (1997) *Trends Biotechnol.* **15,** 62–70
19. Hoogenboom, H. R., de Bruïne, A. P., Hufton, S. E., Hoet, R. M., Arends, J. W., and Roovers, R. C. (1998) *Immunotechnology* **4,** 1–20
20. Clackson, T., and Wells, J. A. (1994) *Trends Biotechnol.* **12,** 173–184
21. McGrath, M. E., Erpel, T., Bystroff, C., and Fletterick, R. J. (1994) *EMBO J.* **13,** 1502–1507
22. 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
23. Meng, E. C., Shoichet, B. K., and Kuntz, I. D. (1992) *J. Comput. Chem.* **13,** 505–514
24. Ewing, T. J. A., and Kuntz, I. D. (1997) *J. Comput. Chem.* **18,** 1175–1189
25. Morrison, J. F. (1969) *Biochim. Biophys. Acta* **185,** 269–286
26. Braisted, A. C., and Wells, J. A. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 5688–5692
27. Jamieson, A. C., Kim, S. H., and Wells, J. A. (1994) *Biochemistry* **33,** 5689–5695
28. Ellington, A. D., and Szostak, J. W. (1990) *Nature* **346,** 818–822
29. Joyce, G. (1989) *Gene* (*Amst.*) **82,** 83–87
30. Tuerk, C., and Gold, L. (1990) *Science* **249,** 505–510
31. Zou, X., Sun, Y., and Kuntz, I. D. (1999) *J. Am. Chem. Soc.* **121,** 8033–8043