nature biotechnology

© 2003 Nature Publishing Group http://www.nature.com/naturebiotechnology

Engineering of a macromolecular scaffold to develop specific protease inhibitors

A Allart Stoop & Charles S Craik

The specific inhibition of serine proteases, which are crucial switches in many physiologically important processes, is of value both for basic research and for therapeutic applications. Ecotin, a potent macromolecular inhibitor of serine proteases of the S1A family, presents an attractive scaffold to engineer specific protease inhibitors because of its large inhibitor-protease interface. Using synthetic shuffling in combination with a restricted tetranomial diversity, we created ecotin libraries that are mutated at all 20 amino acid residues in the binding interface. The efficacy of these libraries was demonstrated against the serine protease plasma kallikrein (Pkal). Competitive phage display selection yielded a Pkal inhibitor with an apparent dissociation equilibrium constant (K_i*) of 11 pM, whereas K_i* values for related proteases (such as Factor Xa (FXa), Factor XIa (FXIa), urokinase-type plasminogen activator (uPA), thrombin, and membrane-type serine protease 1 (MT-SP1)) were four to seven orders of magnitude higher. The adaptability of the scaffold was demonstrated by the isolation of inhibitors to two additional serine proteases, MT-SP1/matriptase and Factor XIIa.

The family of human serine proteases of the trypsin fold, also known as S1A proteases, consists currently of 98 members that share a highly conserved (His, Asp, Ser) catalytic triad[1,2]. About half of these proteases have been thoroughly studied because of their early identification and essential involvement in tightly regulated physiological processes such as blood coagulation, fibrinolysis and host defense. The other half have only recently been discovered and include >20 type II membrane-bound serine proteases[3,4], 15 tissue kallikreins[5] and 15 novel S1A proteases of unknown designation[2]. Although the exact function of the majority of these proteases is unknown, several have been implicated in cancer (including MT-SP1/matriptase/epithin[6–10], hepsin[11], kallikrein 3 (ref. 12) and kallikrein 6 (ref. 13)). Hence, S1A proteases are potential targets for therapeutic intervention as well as for elucidating biological signaling pathways. For the analysis of these pathways, the transient and flexible properties of a chemical-genetic approach would bypass confounding effects often encountered in classical gene-knockout strategies[14,15]. However, the subtle specificity difference of S1A proteases and their highly conserved three-dimensional architecture make it very difficult to target a single protease specifically.

Protease inhibitors can be divided into two classes: small molecules and macromolecules. Small-molecule inhibitors use the chemical nature of the active site and extended substrate-binding pockets to bind and inhibit the protease. Because of their small size and limited recognition determinants, these inhibitors often cannot discriminate sufficiently between closely related serine proteases[16,17]. Most macromolecular inhibitors function through a standard mechanism, in which a single substratelike loop binds the protease active site. Though making greater contacts than small molecules, these inhibitors target only the most conserved elements of the protease[18]. Ecotin, a dimeric fold–specific, macromolecular serine protease inhibitor from *Escherichia coli*[19], with homologs in other microorganisms[20], is an exception to this theme in that it binds not only to the active-site pocket but also to less conserved areas of the protease. The interaction involves four surface loops organized in primary and secondary binding sites (**Fig. 1a**). Analysis of the crystal structure of the ecotin-trypsin complex revealed that these two sites together create a very large (2,850 Å^2) binding interface, helping to explain the broad inhibition of serine proteases by ecotin[20–25].

We postulate that ecotin represents an ideal scaffold to develop highly specific inhibitors to proteases of the S1A family. By mutating all four contact loops simultaneously, it should be possible to sculpt the surface of ecotin so that it specifically recognizes a single protease. As proof of principle, we chose to develop an ecotin variant specific for plasma kallikrein (Pkal) for several reasons. First, this protease is involved in multiple biological pathways of potential therapeutic value[26–29]. Second, both small-molecule[28] and macromolecular inhibitors[30,31] are available for comparison of their specificity profiles. Finally, the resulting inhibitor can be used to discover roles of Pkal *in vivo*[27]. To demonstrate that the adaptability of the ecotin scaffold is not limited to Pkal, we also created inhibitors of MT-SP1 and Factor XIIa.

RESULTS

Design and construction of ecotin libraries

To introduce specificity in the ecotin scaffold, we focused on mutagenizing the four surface loops (50s (positions 51–54), 60s (positions 67–71), 80s (positions 81–86) and 100s (positions 108–112)), while

Department of Pharmaceutical Chemistry, University of California San Francisco, 600 16th Street Suite S512, San Francisco, California 94143-2280, USA. Correspondence should be addressed to C.S.C. (craik@cgl.ucsf.edu).

Published online 17 August 2003; doi:10.1038/nbt860

© 2003 Nature Publishing Group http://www.nature.com/naturebiotechnology

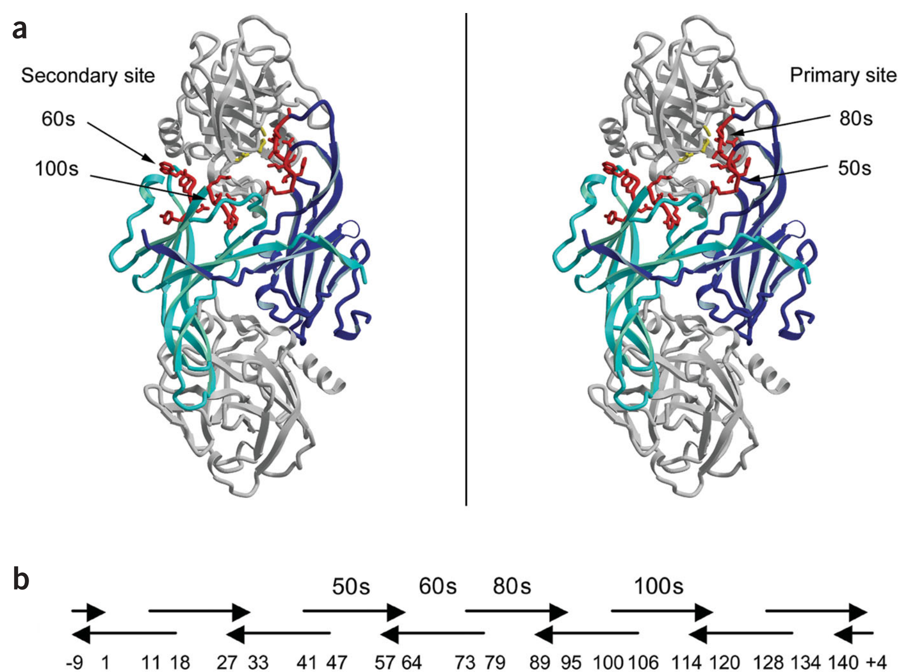

Figure 1 Positioning of ecotin surface loops in crystal structure and during synthetic shuffling. (**a**) Stereo view of the ecotin-trypsin complex, in which the ecotin moiety consists of two ecotin monomers (blue and cyan), which inhibit two trypsin molecules (gray) in a symmetrical, tetrameric structure. The four surface loops of ecotin (red) that form the ecotin-trypsin binding interface are organized in a primary and secondary site. The substrate-mimicking 80s loop interacts directly with the catalytic triad (yellow) of the target protease and is supported by the 50s loop that also has direct contacts with the protease. The 60s and 100s loops of the other ecotin molecule form the secondary binding site with the protease and are located distal from the catalytic triad. The complex shown is ecotin-M84R in complex with trypsin-D102N (personal communication, S. Wang). The figure was prepared using MolScript[48]. (**b**) Schematic representation of surface loop–targeted mutagenesis of ecotin using partially overlapping oligonucleotides. Ten oligonucleotides, spanning ecotin, were used to reconstitute the gene, with the lower numbers indicating the amino acid positions at which the oligonucleotides begin and end. The nucleotide variation incorporated in the surface loops is summarized in **Table 1**. The ecotin libraries were constructed by varying combinations of mutated surface loops.

avoiding destabilizing mutations in the backbone of the scaffold. This was accomplished by reconstituting the entire gene from ten partially overlapping oligonucleotides in a two-step PCR reaction that resembles DNA shuffling[32], with each targeted surface loop encoded by a single oligonucleotide (**Fig. 1b**). This versatile and modular approach is guided by the structural information available for ecotin and makes possible every combination of mutated loops. A similar approach was recently described by Ness *et al.* and termed synthetic DNA shuffling[33].

If all 20 amino acid residues on the surface loops were mutated to all possible substitutions, there would be a total of 20^{20} possibilities, clearly a nonfeasible approach. Therefore, the number of tolerated mutations at a single position was limited to four residues, which were encoded by the same degenerate codon. To preserve the functional integrity of ecotin, we always included the wild-type (wt) amino acid residue. Sequence alignment of ecotin homologs in *Pseudomonas aeruginosa, Yersinia pestis* and *Trypanosoma brucei* illustrates the variation tolerated during evolution[20]. This naturally occurring variation of ecotin was also incorporated into the library. Finally, an analysis by Bogan and Thorn of the contributions of individual amino acid residues in protein-protein binding identified those residues preferentially found in 'hot spots'[34]. These 'hot spots' are enriched in tryptophan, tyrosine and arginine, and surrounded by residues that are energetically less important to binding. Taking this information into account, we chose to include specifically these residues in combination with a smaller residue (for example serine, threonine, alanine, valine and glycine) that may complement a neighboring large residue. The tetranomial diversity incorporated in the library (**Table 1**) should present a chemical bias toward energetically favorable residues and consequently 'stack the deck' for isolation of preferential binders.

Allowing only four mutations at each position within the four loops would result in a number that is still too large ($4^{20} = 1.1 \times 10^{12}$) to make a single, complete library. Therefore, a two-step approach was chosen. In the first step, three ecotin libraries were made, using two surface loop–targeting oligonucleotides in each library. This resulted in a maximum of 11 residues being varied at a time ($4^{11} = 4.2 \times 10^6$). The reconstituted ecotin genes were expressed on phage and the following libraries were made: 50/80 (3.5×10^7 colony forming units, c.f.u.), 60/80 (4.2×10^7 c.f.u.) and 100/80 (3.1×10^7 c.f.u.), respectively. To enrich for surface loop variation that is compatible with Pkal binding, we subsequently subjected all three libraries to two rounds of low-stringency, phage display selection for Pkal binders. The second step consisted of recombining the three previously selected libraries into a single library by DNA shuffling (8.7×10^7 c.f.u.), creating ecotin mutants with varying degrees of mutated loops. DNA sequence analysis of 20 clones confirmed that recombination of the three libraries had occurred: clones containing one (20%), two (55%), three (20%) and four (5%) loops mutated were identified.

Isolation of specific and potent inhibitors of Pkal

To select specific inhibitors from the shuffled ecotin library, we displayed the mutants on phage and subjected them to competitive panning for Pkal binders. Ecotin-displaying phage were preincubated with increasing concentrations of a mixture of five related proteases (FXa, FXIa, FXIIa, uPA and MT-SP1), followed by addition of biotinylated Pkal and pull-down onto neutravidin-coated microtiter plates. By doing the selections with protease in solution, we assure a proper conformation of the protease, and the use of biotin rather than an antibody-mediated pull-down assures selective capture of Pkal from a mixture of highly related proteases. In addition, direct infection of *E. coli* by bound phage abolishes the need for the harsh elution conditions that are necessary to dissociate the most tightly bound phage, a serious problem with tight binders such as ecotin.

During the sequential rounds of phage display selection, we monitored the development of specificity by expressing the pool of selected clones and analyzing the ecotin protein extracts for inhibition of either Pkal or FXIa (**Fig. 2a**). The phage isolated after the seventh

round of panning were recloned, and 36 individual clones were expressed. These crude protein extracts were screened for inhibition of either Pkal or FXIa, and the ratios between the dilutions necessary to inhibit 50% of each protease activity were plotted (**Fig. 2b**). The ratios ranged from 3.5 times for ecotin wt to 530–1,900 times for the six ecotin variants with the highest ratios.

The six ecotin variants with the highest Pkal/FXIa inhibition ratio were found to be unique by DNA sequence analysis. Purified protein of these six variants was used to profile the inhibition of six proteases (Pkal, FXa, FXIa, FXIIa, MT-SP1 and plasmin) (**Table 2**). The ecotin mutant 7-29, which contained eight amino-acid substitutions in the 50s, 80s and 100s loops (H53P, S82W, T83N, M84R, M85R, A86S, R108S and N110S), was the most specific Pkal inhibitor and was renamed ecotin-Pkal. Detailed analyses of the apparent equilibrium dissociation constants (K_i*) of ecotin wt and ecotin-Pkal were conducted for Pkal and nine related serine proteases: FXa, FXIa, FXIIa, thrombin, MT-SP1, plasmin, uPA, tissue-type plasminogen activator (tPA) and rat trypsin. Comparison of ecotin wt to ecotin-Pkal demonstrated that whereas the K_i* of ecotin-Pkal for Pkal decreased sixfold (from 66 ± 5 pM to 11 ± 2.5 pM), the K_i* of ecotin-Pkal for all other proteases increased by at least two orders of magnitude (FXIIa) and in many cases by four orders of magnitude (for example, FXa, MT-SP1 and rat trypsin) (**Fig. 3** and **Supplementary Table 1** online). The inhibition profile of ecotin-Pkal itself shows that FXIIa is the only other protease inhibited, albeit with a 260-fold lower affinity than Pkal, whereas inhibition of all other proteases tested was four to seven orders of magnitude less potent.

Table 1 Overview of mutations incorporated at indicated position within ecotin

Position	Wt	Mut 1	Mut 2	Mut 3	Codon
51	Asn	Ser	Tyr	Thr	WMY
52	Leu	Arg*	Trp	Met	WKG
53	His	Gln*	Arg	Ser	MRY
54	Arg	Met*	Thr*	Lys	ANG
67	Trp	Ser	Arg	Thr	WSG
68	Gly	Ser	Trp	Ala	KSG
69	Tyr	Phe*	Asp	Val	KWY
70	Asp	Ser*	Tyr	Ala	KMY
71	Tyr	Asn	Phe	Ile	WWY
81	Val	Met*	Arg	Gly	RKG
82	Ser	Thr*	Trp	Arg	WSG
83	Thr	Asn	Tyr	Ser	WMY
84	Met	Arg*	Lys	Thr	ANG
85	Met	Arg*	Lys	Thr	ANG
86	Ala	Ser*	Trp	Gly	KSG
108	Arg	Ser	Trp	Thr	WSG
109	Tyr	Ser	Asp	Ala	KMY
110	Asn	His	Arg	Ser	MRY
111	Ser	Gly	Trp	Ala	KSG
112	Lys	Gln*	Arg*	Ser	ARN

Wt denotes amino acid residue found in ecotin wt. Asterisks denote amino acid residue found in an ecotin homolog (*Pseudomonas aeruginosa, Yersinia pestis or Trypanosoma brucei*). Codon indicates the nucleotide codon used to obtain the indicated mixture of amino acid residues. DNA degeneracies are represented in the IUB code: N, A/C/G/T; W, A/T; M, A/C; Y, C/T; K, G/T; R, G/A; S, C/G.

Adaptability of the ecotin scaffold to other proteases

To demonstrate that the adaptability of the ecotin scaffold is not limited to Pkal, we set out to isolate additional ecotin-based inhibitors for the serine proteases MT-SP1 and FXIIa. To this end, a naive ecotin mutant library was constructed by oligonucleotide synthesis, using the restricted tetranomial diversity described earlier while targeting all four loops simultaneously (50/60/80/100). This library contained 1.5×10^8 c.f.u., which represents <0.01% of the possibilities theoretically available for the diversity used. Within this modest-size library, six rounds of competitive phage display selection, with a subsequent functional screen for either MT-SP1 or FXIIa inhibition, yielded ecotin variants that preferentially inhibit each of these proteases (**Fig. 4** and **Supplementary Table 2** online). MT-6 contained mutations in three loops (W67R, G68W, D70Y, Y71I, V81R, M84K, M85R, A86G, R108S and K112N) and increased affinity fourfold for MT-SP1 (66 pM vs. 250 pM for ecotin wt) while losing affinity for FXa, FXIIa and Pkal by more than two to four orders of magnitude. The variant XII-18, with mutations in two loops (V81M, T83Y, M85R, A86S, R108S and K112S), did lose sevenfold affinity for FXIIa (200 pM vs. 30 pM), but far more affinity was lost for related proteases (29,600-fold against MT-SP1, 13,000-fold against FXa, and 370-fold against Pkal). Although these variants are not as specific for their respective target proteases as ecotin-Pkal, the mere selection of these variants within an incomplete, though chemically biased, library underlines the potential both of the tetranomial diversity allowed in the surface loops and of ecotin as a scaffold for the engineering of specific inhibitors.

DISCUSSION

Previous attempts to re-engineer ecotin's specificity toward uPA have been successful at increasing potency for the target protease, but failed to create specificity[35,36]. An explanation was that in those experiments only a single loop was mutated at a time, albeit to all 20 possible residues, in an approach similar to that used to optimize peptides and Kunitz-type domain inhibitors[37,38]. However, the latter have a far smaller protease-inhibitor interface than ecotin, where all four loops contribute critically to protease inhibition[39]. It is for this reason that we compromised diversity, by incorporating only four different residues at a given position, so as to allow targeting of multiple loops at the same time.

By limiting the diversity at a given position, one risks not including the optimal residue in the library. On the other hand, by choosing a diversity that is independent of the target protease, the same library can be used to isolate inhibitors to different proteases, as we demonstrate for Pkal, MT-SP1 and FXIIa. In this study, all mutations identified in the MT-SP1-, FXIIa- and Pkal-specific ecotin variants were encoded by the tetranomial diversity allowed, except for a single His→Pro mutation at position 53 in ecotin-Pkal. This position was included in the library, but the variation of a proline was not. A mutation in a single base pair most likely occurred during either gene reconstruction from oligonucleotides or DNA shuffling, processes known to introduce low levels of random mutations[32,40]. The enrichment of such a rare mutation can be explained by the large influence of the H53P mutation on Pkal-specific inhibition (data not shown). Overall, the isolation of ecotin-Pkal and an analysis of the energetic contributions of each loop to specificity (data to be reported elsewhere) underscored the principle that specific inhibition by the ecotin scaffold requires fine-tuning of multiple loops simultaneously, whereas a restricted tetranomial diversity suffices to generate initial specificity for a single protease.

Noncovalent, small-molecule inhibitors are usually the preferred choice when targeting serine proteases[17], because of their pharmacological properties. However, obtaining highly specific and potent inhibitors for therapeutic purposes has proved a serious challenge,

© 2003 Nature Publishing Group http://www.nature.com/naturebiotechnology

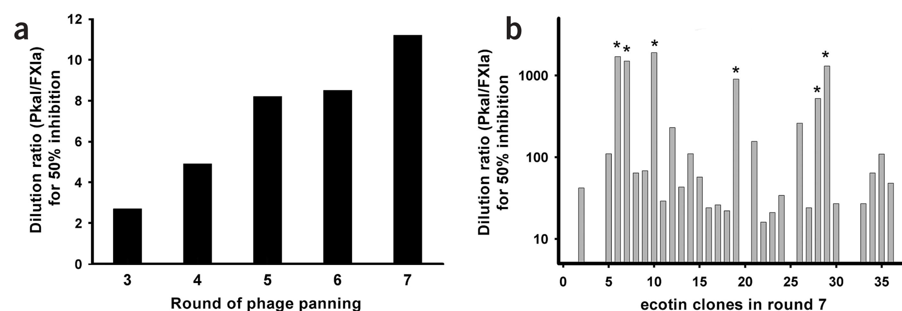

Figure 2 Analysis of the development of ecotin to inhibit Pkal preferentially over FXIa for either pools of variants during sequential rounds of phage display selection or individual variants after seven rounds of selection. (**a**) At the indicated rounds of selection, crude ecotin extracts were made of the entire pool of selected phage and the dilution necessary to obtain 50% inhibition of either Pkal or FXIa was determined. The ratio of these values is interpreted as a measure of specificity toward Pkal. FXIa was chosen because it is structurally the most closely related protease to Pkal (ref. 49), whereas the affinity of ecotin wt toward FXIa is considerably less than for Pkal. (**b**) After the seventh round of competitive, phage display selection, 36 individual clones were expressed and their crude protein extracts were analyzed for the ability to differentiate between Pkal and FXIa. Dilution ratios are plotted on a logarithmic scale; the six variants with the largest differentiation are highlighted.

with limited successes for the coagulation proteases thrombin and FXa[16]. A Pkal inhibitor, P8720 (Bz-Pro-Phe-boroArg-OH), has been described[23], which is potent (150 pM) and reasonably selective (100-fold toward plasmin and 300- to 600-fold for thrombin, FXa and FXIIa). The inhibitor was used to study Pkal-mediated inflammatory reactions in rats[28,29], but was not pursued therapeutically because of its toxicity.

Two Kunitz-type domain scaffolds, the Alzheimer's amyloid β-protein precursor inhibitor (APPI) and the lipoprotein-associated coagulation inhibitor (LACI), were used to develop macromolecular inhibitors of Pkal by phage display[30,31]. Both Pkal inhibitors based on these scaffolds were potent (15–284 pM) and highly specific as judged by cross-reacting with only one other tested protease. Specificity, however, is a relative property of protease inhibitors, because it largely depends on which proteases are tested. The identification of many S1A proteases has further increased the possibilities for cross-reactivity with untested proteases. When examining the cross-reactivity of the macromolecular Pkal inhibitors, the proteases that demonstrate cross-reactivity with the Kunitz-type inhibitors (that is, FXIa and plasmin) are not inhibited by ecotin-Pkal, whereas FXIIa is inhibited by ecotin-Pkal but not by the APPI-based inhibitor. However, it should be taken into consideration that the inhibition profile of the Kunitz-type inhibitors was obtained by searching the complete sequence space available to these inhibitors, whereas the ecotin-based inhibitor was found despite searching only a limited amount of the possible sequence space. This implies ample opportunity for further optimization of the inhibition profile of ecotin-Pkal upon complete mutagenesis at the positions identified in this study. In support of this proposal, MT-SP1 and FXIIa inhibitors were isolated from the same tetranomial diversity as was used for ecotin-Pkal selection. Hence, this further establishes ecotin as an adaptable scaffold for macromolecular inhibitor design of serine proteases.

Here we demonstrate that, in addition to being a potent scaffold for serine protease inhibitors, ecotin can be engineered to be a highly specific active site inhibitor. This approach combines the advantages of ecotin, including very high expression levels in *E. coli* (>100 mg/l) and extreme stability (stable at 100 °C), with specificity comparable to that of monoclonal antibodies. Ecotin-Pkal has a specificity profile that allows its use in biological systems to investigate the functions of Pkal *in vivo*. In this context, the observed cross-reactivity with FXIIa will not obscure any Pkal-mediated effect because FXIIa and Pkal are part of the same proteolytic cascade and are reciprocal activators[41]. Consequently, any inhibitor of Pkal will always lower the concentration of FXIIa, and vice versa. Initial experiments with the inhibitor are under way to assess the functions of Pkal during mammary gland involution and wound healing in mice.

The potential of the ecotin scaffold for therapeutic purposes has yet to be determined. Clearly, the origin of the scaffold is nonhuman, and because no human homolog is present, 'humanizing' the scaffold is not an option. It is not yet known whether the scaffold will elicit an immune response. Ecotin is constantly present in the human gastrointestinal tract. Furthermore, there are examples of nonhuman proteins that are currently used therapeutically, such as hirudin[42], isolated from *Hirudo medicinalis* (European leeches), or that are in clinical trials, such as nematode anticoagulant peptide c2 produced by *Ancylostoma caninum*[43]. Other options include reducing antigenicity by identifying and replacing the responsible residues, as was done for staphylokinase[44]. Future applications of the ecotin-based inhibitors include their use as tools for the dissection of complex biological pathways and possibly as therapeutics in human disease.

Table 2 Analysis of rate of inhibition of six related proteases by ecotin wt and six ecotin mutants

Ecotin variant	53	82	83	84	85	86	108	110	Pkal	FXa	FXIa	FXIIa	MT-SP1	Plasmin
7-6	-	Trp	Tyr	Lys	Arg	Ser	-	-	+++	+	--	+++	++	-
7-7	-	Trp	Asn	Arg	Lys	Ser	-	-	++	+	--	+++	++	---
7-10	-	Trp	Tyr	Arg	Lys	Ser	Pro	-	++	+	--	++	+	---
7-19	-	Trp	Ser	Lys	Met	Ser	-	-	+++	++	-	+++	+	--
7-28	-	Trp	Asn	Arg	Arg	Ser	-	-	++	+	-	+++	+	---
7-29	Pro	Trp	Asn	Arg	Arg	Ser	Ser	Ser	+++	-	--	+	--	---
Wt	His	Ser	Thr	Met	M*et*	*ala*	Arg	Asn	+++	+++	+	+++	+++	---

The six different ecotin mutants were selected after seven rounds of competitive phage display selection. The amino acid mutations present in each of the variants are indicated and compared to the sequence in ecotin wt. Symbols are indicative of approximately the following K_i* values: +++, <0.5 nM; ++, 0.5–5 nM; +, 5–50 nM; -, 50–500 nM; —, 0.5–5 μM; —-, >5 μM.

© 2003 Nature Publishing Group http://www.nature.com/naturebiotechnology

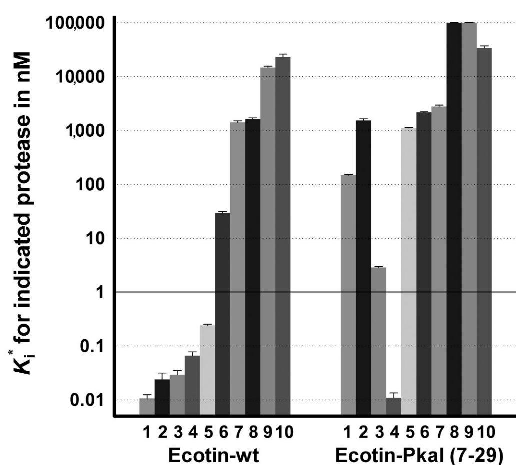

Figure 3 Graphic representation of the K_i* values determined of ecotin wt and ecotin-Pkal for ten proteases. The proteases analyzed are (1) FXa, (2) rat trypsin, (3) FXIIa, (4) Pkal, (5) MT-SP1, (6) FXIa, (7) thrombin, (8) uPA, (9) tPA and (10) plasmin. K_i* values are shown on a logarithmic scale in nM. Values were determined using at least ten different inhibitor concentrations for each K_i* value, as described in Methods. For exact values see **Supplementary Table 1** online.

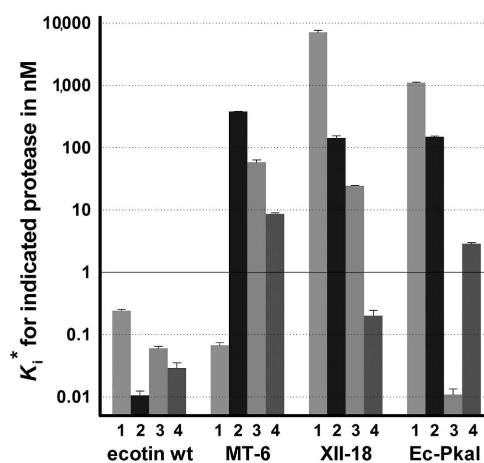

Figure 4 Graphic representation of the K_i* values determined of ecotin wt, MT-6, XII-18 and ecotin-Pkal for four highly related proteases. The proteases analyzed are (1) MT-SP1, (2) FXa, (3) Pkal and (4) FXIIa. K_i* values are shown on a logarithmic scale in nM. Values were determined using at least ten different inhibitor concentrations for each K_i* value, as described in Methods. For exact values see **Supplementary Table 2** online.

© 2003 Nature Publishing Group http://www.nature.com/naturebiotechnology

METHODS

Library construction. Three ecotin mutant libraries (50/80, 60/80, 100/80) and a naive (50/60/80/100) library were made by gene reconstruction from oligonucleotides. To that end, six SDS-PAGE–purified oligonucleotides (Invitrogen) forming the backbone (Ec1, Ec2, Ec3, Ec7, Ec9 and Ec10) and the desired combination of mutated or wt surface loop coding oligonucleotides (Ec50, Ec60, Ec80 and Ec100) (for sequence of oligonucleotides see **Supplementary Table 3** online) were mixed at 5 pmol each in 50 µl, containing 200 µM dNTPs, 1× *Pfx* amplification buffer and 1 unit *Pfx* polymerase. Reassembly was done over 25 cycles (1 min at 95 °C, 1 min at 55 °C, 1 min at 68 °C), and the reaction was then amplified with flanking oligonucleotides XhoF and SpeHisR, using the same PCR cycles just mentioned. Amplification resulted in single bands, which were ligated with *Xho*I-*Spe*I, in phagemid vector pComb3Δ (ref. 45). DNA fragmentation and shuffling of three ecotin mutant libraries (50/80, 60/80 and 100/80) was done using regular protocols for DNA shuffling[32,46]. Briefly, ecotin DNA was digested with DNase I, and 50-bp fragments were isolated and reassembled using the same conditions mentioned earlier, amplifying with XhoF and SpeR oligonucleotides.

Phage display selection of Pka1 binders. Ecotin phage were prepared as described[35]. Phage display selection of in-frame ecotin variants was carried out by incubating 3×10^{11} ecotin-displaying phage with 40 µl Ni-agarose (Qiagen), using washing and elution conditions provided by the manufacturer. The number of eluted phage counted was always $>2 \times 10^7$, ensuring no loss of diversity. Nonstringent selection for Pkal binders was done by incubating 1×10^{11} phage with 5 nM Pkal in 50 µl for 2 h, with subsequent pull-down on microtiter plates coated with 2 µg of sheep anti-human prekallikrein (Enzyme Research Laboratories). Bound phage were washed, eluted at pH 2.2, used to infect 200 µl XL-1 Blue cells, counted and amplified on 2YT plates. Competitive phage pannings of the shuffled ecotin library were done with biotinylated Pkal. Pkal was biotinylated with EZ-Link Sulfo-NHS-LC-LC-biotin (Pierce) according to the manufacturer's instructions, and the level of biotinylation was verified to be ~3 biotin molecules/Pka1 molecule. For these pannings, 1×10^{11} phage were incubated for 1–2 h with a mixture of FXa, FXIa, FXIIa, MT-SP1 and uPA, at individual protease concentrations of 5 nM (rounds 1 and 2), 10 nM (rounds 3, 4 and 5), 25 nM (round 6) and 50 nM (round 7). Subsequently, Pkal-biotin was added at decreasing concentrations (2 nM for round 1 to 0.5 nM for round 7), and incubated for 1 h. The mixture was added to Reacti-Bind NeutraAvidin-coated polystyrene plates (Pierce), incubated for 1 h and washed with TBS + 0.01% Tween-20, after which XL-1 cells were added to the well. After 30 min incubation, cells were removed, plated on 2YT plates and grown overnight. Scraped cells were grown, infected with VCSM13 helper phage (Stratagene) and the procedure repeated for another round of selection. Competitive selections for MT-SP1 and FXIIa inhibitors were carried out similarly to those for Pkal, using biotinylated MT-SP1 and FXIIa instead.

Protein expression and purification. Ecotin variants isolated by phage display were amplified by PCR with the oligonucleotides AatIIlong-F and BsiWI-R, and cloned into the pTacTac vector[21]. Protein expression and purification were carried out essentially as described[39] in an ecotin-deficient strain (IMΔecoJ). SDS-PAGE analysis verified ecotin purity, and ecotin concentrations were determined by A_{280} using 21,860 M/cm (ecotin wt) and 27,550 M/cm (ecotin-Pkal). For crude protein extracts, cells transformed with the ecotin plasmid were grown overnight in the presence of isopropyl-β-D-thiogalactoside, harvested, treated with lysozyme (0.75 mg/ml) in a solution containing 25% sucrose (w/v)–10 mM Tris (pH 8.0)–2 mM EDTA, and the periplasmic fraction isolated.

Screening of ecotin variants. Ecotin variants were screened by incubating 10 nM Pkal or FXIa with dilutions of crude ecotin protein extract in the range of 1:10 to 1:20,000. After 15 min the fluorescent substrate (0.1 mM) H-Pro-Phe-Arg-AMC (Pkal) or Boc-Phe-Ser-Arg-AMC (FXIa) (Bachem AG) was added, and residual protease activity was determined by measuring fluorescence. For each clone the dilution necessary to obtain 50% inhibition of either Pkal or FXIa was determined and these values were divided to obtain a specificity ratio.

Determination of K_i* values against various proteases. Measurements of ecotin inhibition constants were carried out as described[23]. Briefly, fixed protease concentrations in the range of 1–4 nM were incubated for extended periods (as long as 6 h, depending on time needed to reach equilibrium) with at least ten different ecotin concentrations. Subsequently, residual protease activity was determined by adding fluorescent substrates (0.1 mM) for Pkal (H-Pro-Phe-Arg-AMC), FXIa (Boc-Phe-Ser-Arg-AMC) and plasmin (H-D-Ala-Leu-Lys-AMC), all from Bachem AG, or chromogenic substrates for FXa (0.1 mM), FXIIa (0.2 mM), MT-SP1 (0.1 mM), thrombin (0.1 mM) and tPA (0.1 mM) (for all these enzymes Spectrozyme-TPA was used), as well as uPA (0.1 mM, Spectrozyme-UK, American Diagnostica), and change in fluorescence or absorbance was measured. To determine apparent equilibrium dissociation constants (K_i*), plots of the fractional rate versus inhibitor

concentration were fit by nonlinear regression analysis using equation (1) derived for reversible tight-binding inhibitors[47].

$$V_i/V_0 = 1 - ([E_0] + [I_0] + K_i^* - (([E_0] + [I_0] + K_i^*)^2 - 4[E_0][I_0])^{0.5}) / 2[E_0] \quad (1)$$

where V_i/V_0 is the ratio of the inhibited rate versus the uninhibited rate, $[E_0]$ is the total enzyme concentration and $[I_0]$ is the total inhibitor concentration.

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS
We would like to thank C. Eggers for many helpful discussions regarding the data and the manuscript. This work was supported by a TALENT-stipend from the Netherlands Organization for Scientific Research to A.A.S. and grant CA72006 from the National Institutes of Health to C.S.C.

COMPETING INTERESTS STATEMENT
The authors declare that they have no competing financial interests.

Received 12 May; accepted 26 June 2003
Published online at http://www.nature.com/naturebiotechnology/

© 2003 Nature Publishing Group http://www.nature.com/naturebiotechnology

1. Rawlings, N.D. Introduction: family S1 of trypsin. in *Handbook of Proteolytic Enzymes* (eds. Barrett, A.J., Rawlings, N.D. & Woessner, J.F.) 5–12 (Academic Press, London, UK, 1998).
2. Rawlings, N.D., O'Brien, E. A. & Barrett, A.J. MEROPS: the protease database. *Nucleic Acids Res.* **30**, 343–346 (2002).
3. Bhatt, A.S. *et al.* Quantification of membrane type serine protease 1 (MT-SP1) in transformed and normal cells. *Biol. Chem.* **384**, 257–266 (2003).
4. Hooper, J.D., Clements, J.A., Quigley, J.P. & Antalis, T.M. Type II transmembrane serine proteases. Insights into an emerging class of cell surface proteolytic enzymes. *J. Biol. Chem.* 276, 857–860 (2001).
5. Yousef, G.M. & Diamandis, E.P. Human tissue kallikreins: a new enzymatic cascade pathway? *Biol. Chem.* **383**, 1045–1057 (2002).
6. Takeuchi, T., Shuman, M.A. & Craik, C.S. Reverse biochemistry: use of macromolecular protease inhibitors to dissect complex biological processes and identify a membrane-type serine protease in epithelial cancer and normal tissue. *Proc. Natl. Acad. Sci. USA* **96**, 11054–11061 (1999).
7. Lin, C.Y., Anders, J., Johnson, M., Sang, Q.A. & Dickson, R.B. Molecular cloning of cDNA for matriptase, a matrix-degrading serine protease with trypsin-like activity. *J. Biol. Chem.* **274**, 18231–18236 (1999).
8. Kim, M.G. *et al.* Cloning and chromosomal mapping of a gene isolated from thymic stromal cells encoding a new mouse type II membrane serine protease, epithin, containing four LDL receptor modules and two CUB domains. *Immunogenetics* **49**, 420–428 (1999).
9. Oberst, M.D. *et al.* Expression of the serine protease matriptase and its inhibitor HAI-1 in epithelial ovarian cancer: correlation with clinical outcome and tumor clinicopathological parameters. *Clin. Cancer Res.* **8**, 1101–1107 (2002).
10. Kang, J.Y. *et al.* Tissue microarray analysis of hepatocyte growth factor/Met pathway components reveals a role for Met, matriptase, and hepatocyte growth factor activator inhibitor 1 in the progression of node-negative breast cancer. *Cancer Res.* **63**, 1101–1105 (2003).
11. Dhanasekaran, S.M. *et al.* Delineation of prognostic biomarkers in prostate cancer. *Nature* **412**, 822–826 (2001).
12. Catalona, W.J. *et al.* Measurement of prostate-specific antigen in serum as a screening test for prostate cancer. *N. Engl. J. Med.* **324**, 1156–1161 (1991).
13. Hoffman, B.R. *et al.* Immunofluorometric quantitation and histochemical localisation of kallikrein 6 protein in ovarian cancer tissue: a new independent unfavourable prognostic biomarker. *Br. J. Cancer* **87**, 763–771 (2002).
14. List, K. *et al.* Matriptase/MT-SP1 is required for postnatal survival, epidermal barrier function, hair follicle development, and thymic homeostasis. *Oncogene* **21**, 3765–3779 (2002).
15. Ebnet, K. *et al.* Granzyme A-deficient mice retain potent cell-mediated cytotoxicity. *EMBO J.* **14**, 4230–4239 (1995).
16. Sanderson, P.E. Small, noncovalent serine protease inhibitors. *Med. Res. Rev.* **19**, 179–197 (1999).
17. Leung, D., Abbenante, G. & Fairlie, D.P. Protease inhibitors: current status and future prospects. *J. Med. Chem.* **43**, 305–341 (2000).
18. Perona, J.J., Tsu, C.A., Craik, C.S. & Fletterick, R.J. Crystal structures of rat anionic trypsin complexed with the protein inhibitors APPI and BPTI. *J. Mol. Biol.* **230**, 919–933 (1993).
19. Chung, C.H., Ives, H.E., Almeda, S. & Goldberg, A.L. Purification from *Escherichia coli* of a periplasmic protein that is a potent inhibitor of pancreatic proteases. *J. Biol. Chem.* **258**, 11032–11038 (1983).
20. Gillmor, S.A., Takeuchi, T., Yang, S.Q., Craik, C.S. & Fletterick, R.J. Compromise and accommodation in ecotin, a dimeric macromolecular inhibitor of serine proteases. *J. Mol. Biol.* **299**, 993–1003 (2000).
21. McGrath, M.E., Erpel, T., Bystroff, C. & Fletterick, R.J. Macromolecular chelation as an improved mechanism of protease inhibition: structure of the ecotin-trypsin complex. *EMBO J.* **13**, 1502–1507 (1994).
22. Wang, S.X., Esmon, C.T. & Fletterick, R.J. Crystal structure of thrombin-ecotin reveals conformational changes and extended interactions. *Biochemistry* **40**, 10038–10046 (2001).
23. Eggers, C.T., Wang, S.X., Fletterick, R.J. & Craik, C.S. The role of ecotin dimerization in protease inhibition. *J. Mol. Biol.* **308**, 975–991 (2001).
24. Seymour, J.L. *et al.* Ecotin is a potent anticoagulant and reversible tight-binding inhibitor of Factor Xa. *Biochemistry* **33**, 3949–3958 (1994).
25. Ulmer, J.S., Lindquist, R.N., Dennis, M.S. & Lazarus, R.A. Ecotin is a potent inhibitor of the contact system proteases Factor XIIa and plasma kallikrein. *FEBS Lett.* **365**, 159–163 (1995).
26. Colman, R.W. & Schmaier, A.H. Contact system: a vascular biology modulator with anticoagulant, profibrinolytic, antiadhesive, and proinflammatory attributes. *Blood* **90**, 3819–3843 (1997).
27. Selvarajan, S., Lund, L.R., Takeuchi, T., Craik, C.S. & Werb, Z. A plasma kallikrein–dependent plasminogen cascade required for adipocyte differentiation. *Nat. Cell Biol.* **3**, 267–275 (2001).
28. Dela Cadena, R.A. *et al.* Inhibition of plasma kallikrein prevents peptidoglycan-induced arthritis in the Lewis rat. *FASEB J.* **9**, 446–452 (1995).
29. Stadnicki, A. *et al.* Specific inhibition of plasma kallikrein modulates chronic granulomatous intestinal and systemic inflammation in genetically susceptible rats. *FASEB J.* **12**, 325–333 (1998).
30. Dennis, M.S., Herzka, A. & Lazarus, R.A. Potent and selective Kunitz domain inhibitors of plasma kallikrein designed by phage display. *J. Biol. Chem.* **270**, 25411–25417 (1995).
31. Markland, W., Ley, A.C. & Ladner, R.C. Iterative optimization of high-affinity protease inhibitors using phage display. 2. Plasma kallikrein and thrombin. *Biochemistry* **35**, 8058–8067 (1996).
32. Stemmer, W.P. Rapid evolution of a protein *in vitro* by DNA shuffling. *Nature* **370**, 389–391 (1994).
33. Ness, J.E. *et al.* Synthetic shuffling expands functional protein diversity by allowing amino acids to recombine independently. *Nat. Biotechnol.* **20**, 1251–1255 (2002).
34. Bogan, A.A. & Thorn, K.S. Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* **280**, 1–9 (1998).
35. Yang, S.Q. & Craik, C.S. Engineering bidentate macromolecular inhibitors for trypsin and urokinase-type plasminogen activator. *J. Mol. Biol.* **279**, 1001–1011 (1998).
36. Laboissiere, M.C. *et al.* Computer-assisted mutagenesis of ecotin to engineer its secondary binding site for urokinase inhibition. *J. Biol. Chem.* **277**, 26623–26631 (2002).
37. McBride, J.D., Freeman, N., Domingo, G.J. & Leatherbarrow, R.J. Selection of chymotrypsin inhibitors from a conformationally-constrained combinatorial peptide library. *J. Mol. Biol.* **259**, 819–827 (1996).
38. Dennis M.S. & Lazarus R.A. Kunitz domain inhibitors of tissue factor-Factor VIIa. I. Potent inhibitors selected from libraries by phage display. *J. Biol. Chem.* **269**, 22129–22136 (1994).
39. Yang, S.Q., Wang, C.I., Gillmor, S.A., Fletterick, R.J. & Craik, C.S. Ecotin: a serine protease inhibitor with two distinct and interacting binding sites. *J. Mol. Biol.* **279**, 945–957 (1998).
40. Hecker, K.H. & Rill, R.L. Error analysis of chemically synthesized polynucleotides. *Biotechniques* **24**, 256–260 (1998).
41. Chung, D.W., Fujikawa, K., McMullen, B.A. & Davie, E.W. Human plasma prekallikrein, a zymogen to a serine protease that contains four tandem repeats. *Biochemistry* **25**, 2410–2417 (1986).
42. OASIS investigators. Effects of recombinant hirudin (lepirudin) compared with heparin on death, myocardial infarction, refractory angina, and revascularisation procedures in patients with acute myocardial ischaemia without ST elevation: a randomised trial. *Lancet* **353**, 429–438 (1999).
43. Cappello, M., Vlasuk, G.P., Bergum, P.W., Huang, S. & Hotez, P.J. *Ancylostoma caninum* anticoagulant peptide: a hookworm-derived inhibitor of human coagulation Factor Xa. *Proc. Natl. Acad. Sci. USA* **92**, 6152–6156 (1995).
44. Laroche, Y. *et al.* Recombinant staphylokinase variants with reduced antigenicity due to elimination of B-lymphocyte epitopes. *Blood* **96**, 1425–1432 (2000).
45. Barbas, C.F., Kang, A.S., Lerner, R.A. & Benkovic, S.J. Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. *Proc. Natl. Acad. Sci. USA* **88**, 7978–7982 (1991).
46. Stoop, A.A., Jespers, L., Lasters, I., Eldering, E. & Pannekoek, H. High-density mutagenesis by combined DNA shuffling and phage display to assign essential amino acid residues in protein-protein interactions: application to study structure-function of plasminogen activation inhibitor 1 (PAI-I). *J. Mol. Biol.* **301**, 1135–1147 (2000).
47. Morrison, J.F. Kinetics of the reversible inhibition of enzyme-catalysed reactions by tight-binding inhibitors. *Biochim. Biophys. Acta* **185**, 269–286 (1969).
48. Kraulis, P.J. MolScript: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
49. Fujikawa, K., Chung, D.W., Hendrickson, L.E. & Davie, E.W. Amino acid sequence of human Factor XI, a blood coagulation factor with four tandem repeats that are highly homologous with plasma prekallikrein. *Biochemistry* **25**, 2417–2424 (1986).