

Generation of a 100-billion cyclic peptide phage display library having a high skeletal diversity

Vanessa Carle¹, Xu-Dong Kong¹, Alice Comberlato¹, Chelsea Edwards¹, Cristina Díaz-Perlas¹, and Christian Heinis¹

¹Institute of Chemical Sciences and Engineering, School of Basic Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

Correspondence should be addressed to C.H. E-mail: christian.heinis@epfl.ch

Keywords: cyclic peptide, bicyclic peptide, library, phage display, whole plasmid PCR, hyperphage

Abstract

Phage display is a powerful technique routinely used for the generation of peptide- or protein-based ligands. The success of phage display selections critically depends on the size and structural diversity of the libraries, but the generation of large libraries remains challenging. In this work, we have succeeded in developing a phage display library comprising around 100 billion different (bi)cyclic peptides and thus more structures than any previously reported cyclic peptide phage display library. Building such a high diversity was achieved by combining a recently reported library cloning technique, based on whole plasmid PCR, with a small plasmid that facilitated bacterial transformation. The library cloned is based on 273 different peptide backbones and thus has a large skeletal diversity. Panning of the peptide repertoire against the important thrombosis target coagulation factor XI enriched high-affinity peptides with long consensus sequences that can only be found if the library diversity is large. Most of the enriched binders were based on only 3 out of the 273 backbones, underscoring the importance of building libraries with high skeletal diversity.

Introduction

Cyclic peptides are an attractive modality for the development of therapeutics thanks to their ability to bind with high affinity and selectivity to protein targets, their low inherent toxicity, their relatively high stability, and their ease of production by solid-phase peptide synthesis. Currently, more than 40 cyclic peptides derived from nature, or derivatives thereof, are used as drugs, and more than 100 cyclic peptides are in pre-clinical or clinical development (Zorzi et al., 2017). Several of the cyclic peptides in development were isolated by screening large combinatorial libraries of peptides by display techniques such as phage display or mRNA display (Chang et al., 2013; Ricardo et al., 2015; Sahu et al., 1996). In phage display, the peptides are displayed on the surface of phage as fusion with a coat protein and are encoded by DNA packed in the phage particle. The efficient and correct coupling of pheno- and genotype achieved by phage display allows for the identification of binders in only two or three iterative rounds of phage selection. The high stability of filamentous phage (high temperature, detergents, redox reagents, organic solvents) allows for the selection in presence of harsh conditions, such as proteases (to eliminate unstable peptides). Protocols for phage library production and panning are robust and can be applied by non-specialists.

The most critical parameters in the phage display selection of cyclic peptides are the size and structural diversity of the library used. For easy targets such as proteins that bind to short, linear peptides (e.g. integrins to Arg-Gly-Asp peptides, streptavidin to His-Pro-Gln peptides), peptide ligands can be isolated from small libraries of a few thousands to a million different peptides. However, for the development of cyclic peptides to challenging targets, much larger libraries are needed. A large library size is also important for the selection of peptides that fulfill, in addition to binding to a target, additional qualities such as high stability or binding to multiple targets (e.g. binding to the human and mouse homologues of a protein). Since the innovation of phage display in the 1980's, several different cyclic peptide phage display libraries were reported, wherein the size and complexity has gradually increased. The first cyclic peptide libraries reported included peptides containing two flanking cysteines that are cyclized upon disulfide bond formation (Bonnycastle et al., 1996; Fairbrother et al., 1998; Giebel et al., 1995; Koivunen et al., 1995, 1994). Later, libraries of bicyclic peptides were developed by reacting linear peptides containing three cysteines with chemical linkers containing three thiol-reactive groups, such as 1,3,5-tris(bromomethyl)benzene, that could engage with targets through two peptide loops (Angelini et al., 2012; Baeriswyl et al., 2015; Heinis et al., 2009). Most recently, we reported bicyclic peptides

obtained by chemically bridging two pairs of cysteines in peptides containing four cysteines (Kale et al., 2018). To our knowledge, the largest reported library of cyclic peptides contains 1×10^{10} peptides (Bonnycastle et al., 1996), the largest reported bicyclic peptide library 4×10^9 clones (Heinis et al., 2009), and the largest library of double-bridged peptides comprises 8.5×10^9 different peptides, or 2×10^{10} peptides if six related sub-libraries are pooled (Kong et al., 2020).

The difficulty in generating large phage display libraries is stemming from two experimental steps, the first one being the generation of phage DNA vectors coding for a large number of different peptides, and the second one being the transfer of vector DNA into many bacterial cells. Phage DNA coding for random peptides is typically cloned by cassette mutagenesis, ligating double-stranded DNA coding for the peptides into a phage vector opened by restriction enzymes. A challenge with this cloning strategy is the generation of fully opened vectors and the correct ligation of vector and insert. We had recently adapted a strategy based on whole plasmid PCR for the efficient cloning of peptide phage display libraries (Kong et al., 2020). With this approach, the entire phage vector is copied in a long PCR wherein the DNA coding for the peptide library is appended through a degenerated primer. Vector production through plasmid preparation is thus fully bypassed and the ligation is efficient as the vector needs only to be ligated at one site by "self-ligation". With this whole plasmid PCR cloning strategy, we had managed to generate the above mentioned largest phage display double-bridged peptide libraries to date (8.5×10^9 different peptides, or 2×10^{10} peptides if six related sub-libraries are pooled). The second challenging step in the library construction, the transformation of vector DNA into bacterial cells, is difficult because the cell membrane needs to be broken sufficiently to allow entry, but not too excessively to ensure survival of the cells. Even with the most optimized transformation protocols, only a small fraction of the vector DNA is delivered into the cells. For the large phage display libraries of double-bridged peptides that we developed recently, the transformation of DNA into bacterial cells was particularly difficult due to the large size of the phage vector used (around 9 kbases).

In this work, we set out to generate a cyclic peptide phage display library of unprecedented size (measured by the number of different peptide sequences it includes) and structural diversity (given by the number of different cyclic peptide backbones). We speculated that a larger quantity of phage vector may be generated if we apply the recently adopted protocol for whole plasmid PCR to a phagemid vector, which is substantially smaller than a phage vector (around 5 kbases vs.

around 9 kbases). Firstly, PCR amplification of a 5 kbase vector was expected to yield more and better-quality PCR product. Secondly, the circularization of a shorter linear DNA should be more efficient. Another advantage of using a smaller DNA vector is that it can be electroporated more efficiently into bacterial cells than a larger vector.

Results

Design of a cyclic peptide library with high backbone diversity

We designed a peptide library of the format $XC(X)_mC(X)_nC(X)_oCX$ (X = any amino acid, C = cysteine, $m + n + o = 12$), in which all peptides contain four cysteines and have a length of 18 amino acids (Fig. 1a). The first and last cysteines were kept in fixed positions (2nd and second last amino acid; red in Fig. 1) and the positions of the middle two cysteines (orange in Fig. 1) were varied in a combinatorial fashion. In this way, the four cysteines could be positioned in the peptides in 91 different ways (Fig. 1b). Upon bridging two pairs of cysteines by disulfide formation or reaction with two chemical linkers - that is possible in three ways - 273 different bicyclic peptide formats, characterized by different numbers m , n and o of random amino acids X in the three segments spaced by the four cysteines, can theoretically be generated (Fig. 1a). The actual number of possible backbones might be slightly smaller as adjacent cysteines may not easily be connected, in particular in the case of large and rigid chemical linkers. With the 273 theoretically possible backbones, the library has a larger skeletal diversity than the recently developed six peptide phage display libraries XD9 (30 backbones), XD10 (45 backbones), XD11 (63 backbones), XD12 (84 backbones), XD13 (108 backbones) and XD14 (135 backbones), or a slightly smaller skeletal diversity if the six libraries XD9 to XD14 are pooled (465 backbones; Fig. 1c) (Kong et al., 2020). In contrast to the XD9 to XD14 library that contains peptides with 5 to 10 random amino acids, the herein developed repertoire is composed of much longer peptide sequences, all containing 14 random amino acids. The longer sequences ensure the sampling of a much larger sequence space and promise the generation of binders with higher binding affinity. Most of the previously reported cyclic or bicyclic peptide libraries contain peptides with the same number of amino acids in the cyclic region and thus the same skeleton, or are based on only a few different backbones (Fig. 1c). Peptides in which the cysteine pairs Cys1/Cys2, and Cys3/Cys4 are bridged can also be regarded as monocyclic peptides, as the two macrocyclic rings are spaced by a linear, flexible peptide sequence having a length of n amino acids. This would mean that each peptide can be cyclized to two bicyclic configurations and two monocyclic configurations, yielding four structures that can be sampled for binding to protein targets. We chose to display the peptides on filamentous phage as fusion of the coat protein 3 (pIII) which leads to the display of a limited number of peptides per phage and allows for the isolation of high affinity binders (Chen and Heinis, 2013; Rebollo et al., 2013; Rentero Rebollo and Heinis, 2013). The peptides of the library and the pIII were spaced by a Gly-Ser-Gly linker.

Whole plasmid PCR of phagemid vector enables generation of large library

We chose to insert DNA encoding the random peptide sequences into a phagemid vector due to the around 2-fold smaller size (around 5 kbases) in comparison to a phage vector. As described above, this strategy promised a more efficient production and transformation of the DNA library. Phagemid vectors code for only the phage coat protein pIII, which means that bacterial cells need to be co-infected with a helper phage that codes for all other essential phage proteins to enable phage production (Qi et al., 2012). We cloned the DNA library by whole plasmid PCR as previously described (Kong et al., 2020). DNA coding for the random peptides was appended to the phagemid DNA of vector pSEX81 in a PCR reaction using forward degenerate primers that annealed at the start of the pIII gene (Fig. 2a). Both, the forward and the reverse primers contained an identical endonuclease restriction site (*NcoI*) that allowed for circularization of the vector by self-ligating sticky ends, to yield the library vector (Fig. 2b).

Linear vector synthesis by PCR of the 5 kbase DNA was much more efficient than the PCR amplification of a 9-kbase phage vector that needed optimization and use of a particularly performant DNA polymerase (Kong et al., 2020). For each one of the 91 degenerate primers used (Supplementary Table 1), we performed a separate PCR reaction and purified the products by agarose gel electrophoresis and DNA extraction. At this point, we combined the PCR products of multiple PCR reactions in 10 groups (Supplementary Table 2), digested the ends with *NcoI*, circularized the vector with T4 ligase and transformed the product into electrocompetent TG1 *E. coli* cells. The cells were plated on 2YT/ampicillin agar plates, grown overnight, resuspended, pooled and stored as a single glycerol stock. In parallel, samples of cells transformed with the 10 different DNA pools were serially diluted, plated, and the number of transformed cells calculated. The number of cells transformed with the 10 pools ranged from 1.47×10^9 to 4.25×10^9 (Supplementary Table 3) and the total number was 2.8×10^{10} . Assuming that all these transformed cells can produce phage with a different peptide, and that each peptide can be cyclized into two bicyclic peptides (double bridged) or two monocyclic peptides (each singly bridged), the diversity of (bi)cyclic peptides generated is around 100 billion.

Sequencing of 125 clones, and thus more than 10 for each of the transformed DNA pools, showed that most of the vectors contained DNA sequences that coded for the desired peptide formats

(Supplementary Fig. 1). Of these sequences, 70 peptides and thus 56% contained exactly the desired number of four cysteines, which is close to the statistically expected number (80 peptides, 64.1%). A fraction of 37.6% of the peptides contained five or more cysteines, which was expected given that cysteine is also coded by one of the 32 NNK codons (35.9% expected). Only 6.4% of the peptides contained less than four cysteines. The latter group of peptides derived most likely from errors in the synthesis of the degenerated DNA primers.

Phage need to be produced in large culture volumes to cover the entire library diversity

We produced phage in a large culture volume of one liter in order to ensure that most of the peptides encoded by the library were produced. The library glycerol stock was inoculated in the medium to reach an OD₆₀₀ of 0.1, yielding a total of 2.5×10^{10} viable cells, as determined by plating of 10-fold serial dilutions of the culture. This number was slightly below the library size of 2.8×10^{10} but we considered it as just sufficient to represent a majority of the peptides in the library. We grew the cells, and at an OD₆₀₀ of 0.5 we infected them with 10^{12} particles of hyperphage, a type of helper phage that does not code for the coat protein pIII and leads to production of phage that display multiple copies of the peptides (Rondot et al., 2001). The number of cells infected with hyperphage was around 2.5×10^{10} , which was again at the lower limit that could be tolerated to cover a large fraction of peptides in the library. Incubation of the culture overnight yielded around 10^{13} phage after PEG purification, which meant that each peptide was represented around 400 times on average.

Phage panning experiments reveal importance of the backbone diversity

We divided the 10^{13} phage derived from the 1-liter culture into two halves and cyclized the peptides either by disulfide formation or by reaction with the alkylating reagent 2,6-bis(bromomethyl)pyridine (BBMP). The two phage libraries were individually subjected to three rounds of phage selection against human coagulation factor XIa (FXIa) immobilized on magnetic streptavidin beads, a target that has received much interest due to its role in thrombosis (Gailani and Gruber, 2016; Weitz, J. I., and Fredenburg, 2017). We performed in parallel selections on beads without immobilized target, to assess if FXIa-specific phage peptide were enriched. Indeed, 35-fold and 7-fold more phage were isolated in the third rounds of selection with the library cyclized by disulfide bridges and chemical linker, respectively (Supplementary Fig. 2), indicating successful

selection of FXIa-specific cyclic peptides. After the third round of selection, we sequenced around 100 clones from each of the two libraries and analyzed, based on the positions of the cysteines, which peptide formats were enriched (Fig. 3a). In both the selections, one of the 91 peptide formats was strongly enriched. This finding showed that peptides with specific cysteine spacing patterns, and thus specific peptide backbones, were particularly suited to generate binders to the target. At the same time, this finding underscored the importance of incorporating a large skeletal diversity in the cyclic peptide libraries as it is not known a priori which format will eventually yield binders.

Isolated peptides display long consensus sequences

Comparison of the enriched peptides revealed strong consensus sequences (Fig. 3b). In the selection performed with the disulfide-cyclized peptide library, all of the peptides could be assigned to three slightly different consensus groups wherein the first two groups were rather similar (Fig. 3b, left). In the selection with BBMP-cyclized peptides, two peptides were found particularly often and they shared some similarity in the core sequences but differed in the position of the third cysteine. Compared to phage selections with smaller peptide libraries (Angelini et al., 2012; Baeriswyl et al., 2013; Rentero Rebollo et al., 2016), the identified consensus sequences were significantly longer. Most impressive was the first consensus sequence of the disulfide-cyclized peptides, in which several different peptides shared identical amino acids in as many as seven positions, not including the two cysteines in the middle for which the positions were varied too.

Bicyclic peptides bind FXIa with high affinity

We synthesized two peptides from each selection, purified them by HPLC (Supplementary Fig. 3) and tested if they inhibit the catalytic domain of FXIa, being a trypsin-like serine protease. Both peptides of the selection with the disulfide-cyclized library (F1, F2) inhibited the protease, F1 with a K_i of 150 nM and F2 with a K_i of 117 nM (Fig. 3c). The two peptides pulled out from the BBMB-cyclized library did not inhibit FXIa (results now shown) and we therefore tested if they bind to a site on FXIa that does not affect the proteolytic activity. Toward this end, we synthesized the peptides with fluorescein linked to the N-terminus and tested the binding to the coagulation factor. Both peptides were binding with affinities in the high nanomolar range, F3 with a K_d of 370 ± 100 nM and F4 with a K_d of 310 ± 60 nM (Fig. 3d). Linear analogs of the two peptides did not bind, showing that the cyclic configuration is key for the good binding.

The consensus sequence "PLPDLP" of the BBMB-cyclized peptides resembles the consensus "^F/_γPDFP" of peptides that were previously isolated against FXI from a library of around 3×10^6 random peptides using a one-bead-one-compound approach (e.g. peptide P39; Fig. 3b, right lower corner) (Wong et al., 2016). The group of Emsley and co-workers, who had isolated these peptides, found that the peptides shared a strong sequence similarity with domains of extracellular proteins such as high-molecular-weight kininogen (HMWK), laminin, collagen V and others. X-ray structures of FXI with a peptide from the OBOC library, P39 (YPRHIYPDFPTD^TT), a peptide from HMWK (FNPISDFPD^TT^S) and one from laminin (RLEFPPDFPID^D) revealed that these peptides are binding to the apple 2 domain of FXI. We performed a competition binding experiment with the P39 peptide from the OBOC library and F4, and found that the peptides compete for the same binding site, suggesting that the herein isolated peptides bind to the apple 2 domain of FXI too (Supplementary Fig. 4). Given that the herein identified consensus sequence was slightly different to the one found with the OBOC library, we speculated that the new consensus may allow identifying new proteins that bind FXI via the apple 2 domain. A search for the sequences LPDL, LxDLP (x = any amino acid) and LPDF in extracellular human proteins identified more than 10 proteins that warrant further investigation. The best binder found in our new library, F4 ($K_d = 310 \pm 60$ nM) has an around 1000-fold better binding affinity than the peptide P39 isolated previously from the OBOC library ($K_d = 320$ μM). Peptides with such a much higher affinity were most likely obtained due to the cyclic nature of the new binders and due to the large skeletal diversity and size of the new, 100-billion member library.

Discussion

We managed to generate a peptide phage display library that contains as many as 100 billion different mono- or bicyclic peptides. In previous work, we were never able to generate such a large library, despite enormous efforts in upscaling vector DNA production and electrocompetent cell production. Key for obtaining an even larger library than in our recent work (Kong et al., 2020) was the combination of the therein reported whole plasmid PCR-based approach with a smaller vector, that facilitated the PCR-amplification of the DNA as well as the transformation of the DNA into cells. While we have applied this strategy for cloning peptide libraries, it might equally well be used for the production of antibody phage display libraries, wherein antibody genes may be used as extra-long primers for the whole plasmid PCR.

In addition to increasing the size of the library, we have also achieved a substantial enlargement of the structural diversity by designing the library in a way that it codes for peptides in which the positions of the four cysteines are distributed in 91 different ways. Upon cyclization, this yields the large number of up to 273 different bicyclic peptide formats and thus a large skeletal diversity. In the case in which the two first and two last cysteines are connected by disulfide bridges or linkers, two monocyclic peptides separated by an amino acid linker are generated, allowing for the screening of monocyclic peptide libraries. In the present work, we have cyclized the peptides by forming two disulfide bridges or by alkylation with a chemical linker that contains two thiol-reactive groups. Given the large number of commercial bis-electrophile reagents that can be used to connect pairs of cysteines, even larger libraries can be generated. For example, by cyclizing the library with ten bis-electrophile reagents, a library comprising a diversity of a trillion different cyclic peptides may be generated.

The advantages of a larger and structurally more diverse library became evident when we panned the pool of (bi)cyclic peptides against a first target, coagulation factor XIa. Firstly, peptides with cysteines in specific positions, and thus specific backbones, were strongly enriched, which showed that it was important to cover a large skeletal diversity. Secondly, the consensus sequences identified were longer than those found in selections with smaller libraries. Such long sequences can only be obtained if a sufficiently large sequence diversity is present in the library. And thirdly, the isolated bicyclic peptides bound with nanomolar affinity to the protein target, which

is a 1000-fold higher affinity than linear peptides isolated to the same region of FXI but from a smaller library.

We have recently applied the same 100-billion cyclic peptide library for the phage selection of binders to the catalytic domain of FXIa (Carle et al., 2021). In order to enrich binders to the catalytic site of FXI, we eluted binding phage using a covalent active-site inhibitor of the protease. With this strategy, we were able to isolate low nanomolar inhibitors from the 100-billion peptide phage display library (Carle et al., 2021). In even more recent phage display selections with four additional protein targets, two other human serine proteases, a protein marker of inflammation, and a receptor protein, we were able to isolate nanomolar binders to all of them (unpublished data). Given the large size of the library of around 100 billion bicyclic peptides, its large structural diversity, covering 273 different backbones, and the successful isolation of tight binders with long consensus sequences, it is likely that it will yield good binders to many more protein targets in the future.

Material and Methods

Library generation

The primers used for library construction are shown in the Supplementary Information (Supplementary Table 1). The library was constructed by inserting DNA sequences encoding the peptides and a Gly–Ser–Gly linker in the phagemid pSEX81 (Progen). The peptide format was Xaa–Cys–(Xaa)_m–Cys–(Xaa)_n–Cys–(Xaa)_o–Cys–Xaa, with $m + n + o = 12$ amino acids and a total peptide length of 18 amino acids. The DNA sequences were incorporated upstream to the pIII gene through a whole plasmid PCR reaction. The DNA encoding the random peptides (in bold), the Gly–Ser–Gly linker (in italic) and the restriction site for the enzyme *NcoI* (underlined) were appended using forward primers with the following format:

5'-AGCGCCCATGGCC**NNKTGT(NNK)_mTGT(NNK)_nTGT(NNK)_oTGTNNKGGTTCTGGCGCTGA**
AACTGTTGAAAGTTGTTTAGC-3'.

The 3' end of forward primers annealed with the 5' end portion of the gene encoding for pIII. The reverse primer had the sequence 5'-CATGCATGGCCGGCTGAGCTGCC-3' and annealed to the pelB leader sequence on pSEX81, which includes a restriction site for the enzyme *NcoI* (underlined). 91 separate PCR reactions were performed (one for each forward primer). The total reaction volume was 30 μ l and contained one of the forward primers (600 nM), the reverse primer (600 nM), dNTP mix (200 μ M) (Axon Lab AG), 30 ng of pSEX81 template (Progen), 6 μ l of 5x Phusion buffer (ThermoFisher Scientific) and four units of Phusion polymerase (ThermoFisher Scientific). The following PCR program was used: 2 min at 95°C, 30 cycles of 30 s at 95°C, 50 s at 60°C and 4 min at 72°C, and final elongation for 7 min at 72°C. The PCR reaction lead to the exclusion of the ScFv gene found on pSEX81 and the incorporation of the peptide library. The PCR products were purified by electrophoresis on a 1% agarose gel in TAE buffer with 1 mM guanosine, gel extracted using a kit (QIAquick gel extraction kit, Qiagen), and the products were pooled in 10 separate groups. DNA from each group was then digested with *NcoI* as follows: 10 μ g of the PCR product and 50 μ l of 10x Tango buffer (ThermoFisher Scientific) were diluted to 490 μ l with ddH₂O. *NcoI* was added (100 units, 10 μ l) (ThermoFisher Scientific), the reaction was incubated for 3 h at 37°C in a water bath. The digested DNA was purified using a commercial kit (QIAquick gel extraction kit, Qiagen) and then self-ligated in a reaction with 50 μ l 10x T4 DNA ligase buffer, 250 Weiss units of T4 ligase (ThermoFisher Scientific) and ddH₂O up to a final

volume 500 μ l. The reaction was incubated at 25°C for 4 h. Ligase inactivation was performed by incubating the tube at 65°C for 10 min. Subsequently, ligated DNA was purified using a commercial kit (QIAquick gel extraction kit, Qiagen) and eluted from the columns using ddH₂O. DNA was electroporated into commercial *E. coli* TG1 cells (Lucigen). 25 μ l of cells were electroporated with 3 μ g of DNA of each group separately. After electroporation, 2 ml of pre-warmed commercial recovery medium (Lucigen) were added to the cells and finally incubated at 37°C for 1 h with shaking (200 rpm). Cells were plated on 20 large (20 cm diameter) 2YT/ampicillin (100 μ g/ml) plates. Plates were incubated at 30°C overnight. The library size was determined by plating cell dilutions on 2xYT/ampicillin agar plates. Electroporation of the DNA into TG1 *E. coli* cells yielded 2.8×10^{10} colonies. Colonies were collected from the plates with 2YT medium/20% v/v glycerol and stored at -80°C. 125 library clones were sequenced by Sanger sequencing (Macrogen) to evaluate the library quality.

Phage production

Library glycerol stocks were inoculated to reach an OD₆₀₀ = 0.1 in 1 L 2YT/ampicillin (100 μ g/ml) culture with 100 mM glucose. Serial dilutions of the starting culture were plated on 2YT/ampicillin (100 μ g/ml) plates in order to assess the number of viable cells in the culture. The culture was grown at 37°C until it reached OD₆₀₀ = 0.5. Serial dilutions of the culture were plated on 2YT/ampicillin (100 μ g/ml) plates. The culture was then infected with hyperphage M13 K07 Δ pIII (Progen Biotechnik GmbH) at a multiplicity of infection (MOI) of 10 and incubated for 20 min at 37°C (no shaking). The culture was then incubated at 37°C for 45 min with 200 rpm shaking. Serial dilutions of the infected culture were plated on 2YT agar plates containing ampicillin (100 μ g/ml) and kanamycin (50 μ g/ml) in order to determine the fraction of infected cells. Cells were pelleted at 2,000 g for 20 min at 4°C and finally resuspended in 1 L of 2YT medium containing ampicillin (100 μ g/ml) and kanamycin (50 μ g/ml). The culture was then grown at 30°C overnight with shaking (250 rpm). Cultures were pelleted at 4,500 g at 4°C for 20 min and the supernatant was kept. A supernatant sample was taken. Phage precipitation was performed by adding 250 ml (25% of the volume) of cold PEG/NaCl solution (20% PEG-6000 (w/v), 2.5 M NaCl) and inverting the bottle, followed by incubation for 45 min on ice. Phage were then centrifuged at 9,000 g for 45 min at 4°C. The phage pellet was re-suspended in 20 ml degassed reaction buffer (20 mM NH₄HCO₃, 5 mM EDTA, pH 8.0). Phage precipitates and remaining cells were removed by centrifugation at 5,000 g at 4°C for 15 min and the supernatant was kept. A sample was stored. The stored phage

samples were used to determine phage titers by performing serial dilutions in 2YT medium and using them to infect 180 μ l of *E. coli* TG1 cells ($OD_{600} = 0.5$). 20 μ l of cells infected with the different phage dilutions were plated on 2YT agar plates containing ampicillin (100 μ g/ml) plates and CFUs were counted. Phage titers were determined at all crucial steps of the process (before and after reduction, cyclization and panning).

Cyclization of cysteines on phage

The purified phage was split into two separate tubes. One of the samples was not further processed (oxidized cysteine residues), while in the other sample the cysteine residues of the peptides were reduced by adding 1 mM TCEP and incubating it for 30 min at 25°C. Phage precipitation was performed by PEG/NaCl precipitation. The tube was incubated 15 min on ice and phage pelleted as described before. Phage were resuspended in 9 ml of reaction buffer. A sample was stored at 4°C for phage titer determination. The chemical linker (2,6-bis(bromomethyl)pyridine) was added at 40 μ M (final conc.) and the reaction incubated at 30°C for 1 h. Phage were again precipitated and finally the phage pellet resuspended in 10 ml binding buffer containing BSA and Tween 20 (10 mM Tris-Cl pH 7.4, 150 mM NaCl, 10 mM MgCl₂, 1 mM CaCl₂, 0.1% v/v Tween 20 and 1% w/v BSA) and stored at 4°C. A sample was stored for phage titer determination.

Biopanning

Human FXIa (Molecular Innovations) was biotinylated by incubating 0.2 mg of protein at 10 μ M with 100 μ M EZ-Link™ Sulfo-NHS-LC-Biotin (10-fold molar excess, ThermoFisher Scientific). The reaction was incubated at room temperature for 1 h. The protein was separated from the unreacted reagent using a PD-10 column (GE Healthcare). 5 μ g of biotinylated target protein were incubated with 50 μ l magnetic streptavidin beads (Dynabeads® M-280 Streptavidin, ThermoFisher Scientific, Waltham, MA, USA) in 500 μ l binding buffer for 10 minutes and non-immobilized protein was removed by washing three times. Beads were resuspended in 300 μ l binding buffer with BSA and Tween 20 and incubated on a rotating wheel at room temperature for 30 min. Beads were then added to each phage library and incubated 30 min at 10 rpm on a rotating wheel. Unbound phage were removed by washing the beads with binding buffer (10 times) wherein the tube was changed once in between the washes. Finally, beads were resuspended in 100 μ l of 20 mM glycine, pH

2.2, and incubated 5 min in order to elute the phage. The solution was neutralized by adding 100 μ l of 1 M Tris/HCl, pH 8.0 and added to 10 ml of *E. coli* TG1 cells at $OD_{600} = 0.5$. After incubation at 37°C for 30 min (no shaking), the freshly infected bacteria were plated on 2YT/ampicillin (100 μ g/ml) plates and grown overnight at 37°C. In addition, phage titers were determined. Bacterial cells of the colonies grown overnight were recovered in 2YT medium and 20% glycerol, flash-frozen and stored at -80°C. Streptavidin- and neutravidin-coated magnetic beads were alternated in the rounds of selection in order to prevent the selection of bead coat protein binders. Neutravidin-coated beads were produced according to the manufacturer's recommended protocol (Dynabeads M280 Tosylactivated, Invitrogen Dynal Biotech AS). After three rounds of phage selection, around 100 clones per condition were sequenced by Sanger sequencing (Macrogen) and resulting sequences were grouped based on sequence similarity.

Chemical synthesis of peptides

Solid phase peptide synthesis (SPPS) was performed using Fmoc-chemistry, DMF as solvent and rink amide AM resin. Peptides were synthesized in a 50 μ mol scale. Amino acids were coupled twice (4.2 equiv., final conc. 220 mM), using HATU (3.9 equiv.), and NMM (10 equiv.) and each coupling was performed at room temperature for 45 minutes. After the coupling reaction, seven washing cycles with DMF were performed. N-terminal amines that remained free after coupling were capped using acetic anhydride (5% v/v) and lutidine (6% v/v) at RT for 30 minutes. Seven washing cycles were again performed. Fmoc deprotection was performed using piperidine (20% v/v) at RT, two times for 5 min. Seven washing cycles were performed. Cleavage of peptides was performed with a standard cleavage cocktail (90% TFA, 2.5% thioanisole, 2.5% H₂O, 2.5% 1,2-ethanedithiol, 2.5% phenol). 10 ml of cleavage cocktail were added to each peptide and incubated for 3 h while shaking. Peptide-containing solution was collected by vacuum filtration and peptides were then initially purified by cold ether precipitation. 50 ml of ice-cold diethyl ether were added to the peptides, incubated 30 min at -20°C and then centrifuged at 2,700 g for 5 min. Peptide pellets were washed another time with 35 ml of diethyl ether and centrifuged again to remove remaining diethyl ether.

Peptide purification by HPLC

Peptides were resuspended in DMF (0.5 ml), acetonitrile/0.1% TFA v/v (1.5 ml) and ddH₂O/0.1% TFA v/v and purified by RP-HPLC (Prep LC 2535 HPLC, Waters) on a C18 column (Sunfire prep C18 TM ODB, 10 μm, 100 Å, 19 × 250 mm, Waters) at a 20 ml/min flow rate, using ddH₂O/0.1% TFA v/v (solvent A) and acetonitrile/0.1% TFA v/v (solvent B) as solvents and a linear gradient of solvent B over solvent A. The mass of peptides contained in the collected fractions was analyzed by ESI-MS on a single quadrupole LC-MS in positive ion mode (LCMS-2020, Shimadzu). Peptides were then lyophilized.

Peptide cyclization

For cyclization with the alkylating reagent, linear peptides were dissolved in 8 ml reaction buffer (60 mM NH₄HCO₃, pH 8.0) and 2 ml acetonitrile to obtain a concentration of 1 mM. Linker was added to a final concentration of 4 mM and the reaction was incubated for 1 hr at 30°C. The reaction was monitored by ESI-MS as described above and stopped by adding 2% v/v formic acid (final conc.). For cyclization by formation of disulfide bridges, peptides were dissolved in 8 ml of reaction buffer and 2 ml of DMSO. The cyclization reaction was monitored by ESI-MS. The cyclized peptides were purified by RP-HPLC as described above and fractions with pure peptides lyophilized. Peptides were dissolved in ddH₂O at a concentration of 2 mM. Purity was characterized by running the peptide solutions on an analytical HPLC (1260 HPLC system, Agilent) with a C18 column (ZORBAX 300SB-C18, 5 μm, 300 Å, 4.6 × 250 mm, Agilent). A 0–100% gradient of acetonitrile, 0.1% TFA v/v (solvent B) in ddH₂O, 5% acetonitrile v/v, 0.1% TFA v/v (solvent A) in 30 minutes was applied.

Inhibitory activity determination

Inhibitory constants (K_i) of peptides were determined by measuring the residual enzymatic activities of Xla incubated with inhibitor at different concentrations (30 μM to 9 nM final conc.) with the substrate Pyr-Pro-Arg-pNA (Bachem). Activities were measured at 25°C in an activity assay buffer (10 mM Tris-Cl, pH 7.4, 150 mM NaCl, 10 mM MgCl₂, 1 mM CaCl₂, 0.1% (w/v) BSA and 0.01% (v/v) Triton-X100). Reactions were started by adding the substrate (400 μM final conc.) to 0.5 nM FXIa in presence or absence of bicyclic peptides. Substrate hydrolysis led to the release of p-nitroaniline (pNA), which was recorded over at least 20 min using an absorbance microtiter

plate reader (Infinite M200 Pro, Tecan) by measuring the absorbance at 405 nm every minute. The rate of increase in absorbance is a measure of enzyme activity. The following equation was applied to determine the IC_{50} values (using Prism 5, GraphPad):

$$y = \frac{100}{1 + 10^{(\log IC_{50} - x)p}}$$

Wherein y corresponds to the residual activity (%) of protease, x is the logarithm of bicyclic peptide concentration, IC_{50} is the inhibitor concentration necessary to achieve 50% inhibition, and p is the Hill coefficient. The following equation was applied to calculate the K_i values:

$$K_i = \frac{IC_{50}}{1 + \frac{[S]_0}{K_m}}$$

Where $[S]_0$ is the initial concentration of substrate, and K_m is the substrate Michaelis constant. The K_m of FXIa for Pyr-Pro-Arg-pNA was determined to be $260 \pm 10 \mu\text{M}$.

Fluorescence polarization for measurement of binding affinity

FXIa was diluted serially in PBS with 0.01% v/v Tween-20. 15 μl were added to 15 μl of peptide-fluorescein conjugate (25 nM final concentration) in 96-well microtiter plates (black, half-area). An absorption microwell plate reader (Infinite M200Pro, Tecan) was used to measure fluorescence anisotropy, with a filter for excitation at 485 nm and emission at 535 nm. Dissociation constants (K_D) were calculated using the following equation (using Prism 5, GraphPad):

$$A = A_f + (A_b - A_f) \times \left\{ \frac{[L]_T + K_D + [P]_T - \sqrt{([L]_T + K_D + [P]_T)^2 - 4[L]_T[P]_T}}{2[L]_T} \right\}$$

Where A is anisotropy. A_f and A_b are the anisotropy values for free (A_f) and bound (A_b) ligand. $[L]_T$ is the concentration of total fluorescent ligand and $[P]_T$ the concentration of protein.

Acknowledgements

The work was supported by the Swiss National Science Foundation (project grants 169526 and 192368) and the Bridge Proof of Concept program (grant 186382) from the Swiss National Science Foundation and Innosuisse.

Competing interests

The authors have no potential competing interest.

References

- Angelini A, Cendron L, Chen S, Touati J, Winter G, Zanotti G, Heinis C. 2012. Bicyclic Peptide Inhibitor Reveals Large Contact Interface with a Protease Target. *ACS Chem Biol* **7**:817–821. doi:10.1021/cb200478t
- Baeriswyl V, Calzavarini S, Chen S, Zorzi A, Bologna L, Angelillo-Scherrer A, Heinis C. 2015. A Synthetic Factor XIIa Inhibitor Blocks Selectively Intrinsic Coagulation Initiation. *ACS Chem Biol* **10**:1861–1870. doi:10.1021/acscchembio.5b00103
- Baeriswyl V, Calzavarini S, Gerschheimer C, Diderich P, Angelillo-Scherrer A, Heinis C. 2013. Development of a Selective Peptide Macrocycle Inhibitor of Coagulation Factor XII toward the Generation of a Safe Antithrombotic Therapy. *J Med Chem* **56**:3742–3746. doi:10.1021/jm400236j
- Bonnycastle LLC, Mehroke JS, Rashed M, Gong X, Scott JK. 1996. Probing the Basis of Antibody Reactivity with a Panel of Constrained Peptide Libraries Displayed by Filamentous Phage. *J Mol Biol* **258**:747–762. doi:10.1006/jmbi.1996.0284
- Carle V, Wu Y, Mukherjee R, Kong X-D, Rogg C, Laurent Q, Cecere E, Villequey C, Konakalla MS, Maric T, Lamers C, Díaz-Perlas C, Butler K, Goto J, Stegmayr B, Heinis C. 2021. Development of Selective FXIa Inhibitors Based on Cyclic Peptides and Their Application for Safe Anticoagulation. *J Med Chem* **64**:6802–6813. doi:10.1021/acs.jmedchem.1c00056
- Chang YS, Graves B, Guerlavais V, Tovar C, Packman K, To K-H, Olson KA, Kesavan K, Gangurde P, Mukherjee A, Baker T, Darlak K, Elkin C, Filipovic Z, Qureshi FZ, Cai H, Berry P, Feyfant E, Shi XE, Horstick J, Annis DA, Manning AM, Fotouhi N, Nash H, Vassilev T, Sawyer TK, Performed JH. 2013. Stapled α -helical peptide drug development: A potent dual inhibitor of MDM2 and MDMX for p53-dependent cancer therapy. *Proc Natl Acad Sci U S A* **110**:3445–3454. doi:10.1073/pnas.1303002110
- Chen S, Heinis C. 2013. Phage Selection of Mono- and Bicyclic Peptide Ligands. doi:10.1039/9781849737159-00241
- Fairbrother WJ, Christinger HW, Cochran AG, Fuh G, Keenan CJ, Quan C, Shriver SK, Tom JYK, Wells JA, Cunningham BC. 1998. Novel peptides selected to bind vascular endothelial growth factor target the receptor-binding site. *Biochemistry* **37**:17754–17764. doi:10.1021/bi981931e

- Gailani D, Gruber A. 2016. Factor XI as a Therapeutic Target. *Arter Thromb Vasc Biol* **36**:1316–1322. doi:10.1161/ATVBAHA.116.306925.
- Giebel LB, Cass RT, Milligan DL, Young DC, Arze R, Johnson CR. 1995. Screening of Cyclic Peptide Phage Libraries Identifies Ligands That Bind Streptavidin with High Affinities. *Biochemistry* **34**:15430–15435. doi:10.1021/bi00047a006
- Heinis C, Rutherford T, Freund S, Winter G. 2009. Phage-encoded combinatorial chemical libraries based on bicyclic peptides. *Nat Chem Biol* **5**:502–507. doi:10.1038/nchembio.184
- Kale SS, Villequey C, Kong XD, Zorzi A, Deyle K, Heinis C. 2018. Cyclization of peptides with two chemical bridges affords large scaffold diversities. *Nat Chem* **10**:715–723. doi:10.1038/s41557-018-0042-7
- Koivunen E, Wang B, Ruoslahti E. 1995. Phage libraries displaying cyclic peptides with different ring sizes: Ligand specificities of the rgd-directed integrine. *Bio/Technology* **13**:265–270. doi:10.1038/nbt0395-265
- Koivunen E, Wang B, Ruoslahti E. 1994. Isolation of a highly specific ligand for the $\alpha 5\beta 1$ integrin from a phage display library. *J Cell Biol* **124**:373–380. doi:10.1083/jcb.124.3.373
- Kong X-D, Carle V, Díaz-Perlas C, Butler K, Heinis C. 2020. Generation of a Large Peptide Phage Display Library by Self-Ligation of Whole-Plasmid PCR Product. *ACS Chem Biol* **15**:2907–2915. doi:10.1021/acscchembio.0c00497
- Qi H, Lu H, Qiu HJ, Petrenko V, Liu A. 2012. Phagemid vectors for phage display: Properties, characteristics and construction. *J Mol Biol* **417**:129–143. doi:10.1016/j.jmb.2012.01.038
- Rebollo IR, Angelini A, Heinis C. 2013. Phage display libraries of differently sized bicyclic peptides. *Medchemcomm* **4**:145–150. doi:10.1039/c2md20171b
- Rentero Rebollo I, Heinis C. 2013. Phage selection of bicyclic peptides. *Methods* **60**:46–54. doi:10.1016/j.ymeth.2012.12.008
- Rentero Rebollo I, McCallin S, Bertoldo D, Entenza JM, Moreillon P, Heinis C. 2016. Development of Potent and Selective *S. aureus* Sortase A Inhibitors Based on Peptide Macrocycles. *ACS Med Chem Lett* **7**:606–11. doi:10.1021/acsmchemlett.6b00045
- Ricardo A, Arata M, DeMarco S, Dhamnaskar K, Hammer R, Fridkis-Hareli M, Rajagopal V, Seyb K, Tang G-Q, Tobe S, Treco D. 2015. Preclinical Evaluation of RA101495, a Potent Cyclic

Peptide Inhibitor of C5 for the Treatment of Paroxysmal Nocturnal Hemoglobinuria. *Blood* **126**:939.

Rondot S, Koch J, Breitling F, Dübel S. 2001. A helper phage to improve single-chain antibody presentation in phage display. *Nat Biotechnol* **19**:75–78. doi:10.1038/83567

Sahu A, Kay BK, Lambris JD. 1996. Inhibition of human complement by a C3-binding peptide isolated from a phage-displayed random peptide library. *J Immunol* **157**:884–91.

Weitz, J. I., and Fredenburg J. 2017. Factor XI and factor XII as targets for new anticoagulants. *Front Med* **4**:1–6. doi:10.1016/S0049-3848(16)30363-2

Wong SS, Østergaard S, Hall G, Li C, Williams PM, Stennicke H, Emsley J. 2016. A novel DFP tripeptide motif interacts with the coagulation factor XI apple 2 domain. *Blood* **127**:2915–2923. doi:10.1182/blood-2015-10-676122

Zorzi A, Deyle K, Heinis C. 2017. Cyclic peptide therapeutics: past, present and future. *Curr Opin Chem Biol* **38**:24–29. doi:10.1016/j.cbpa.2017.02.006

Figures

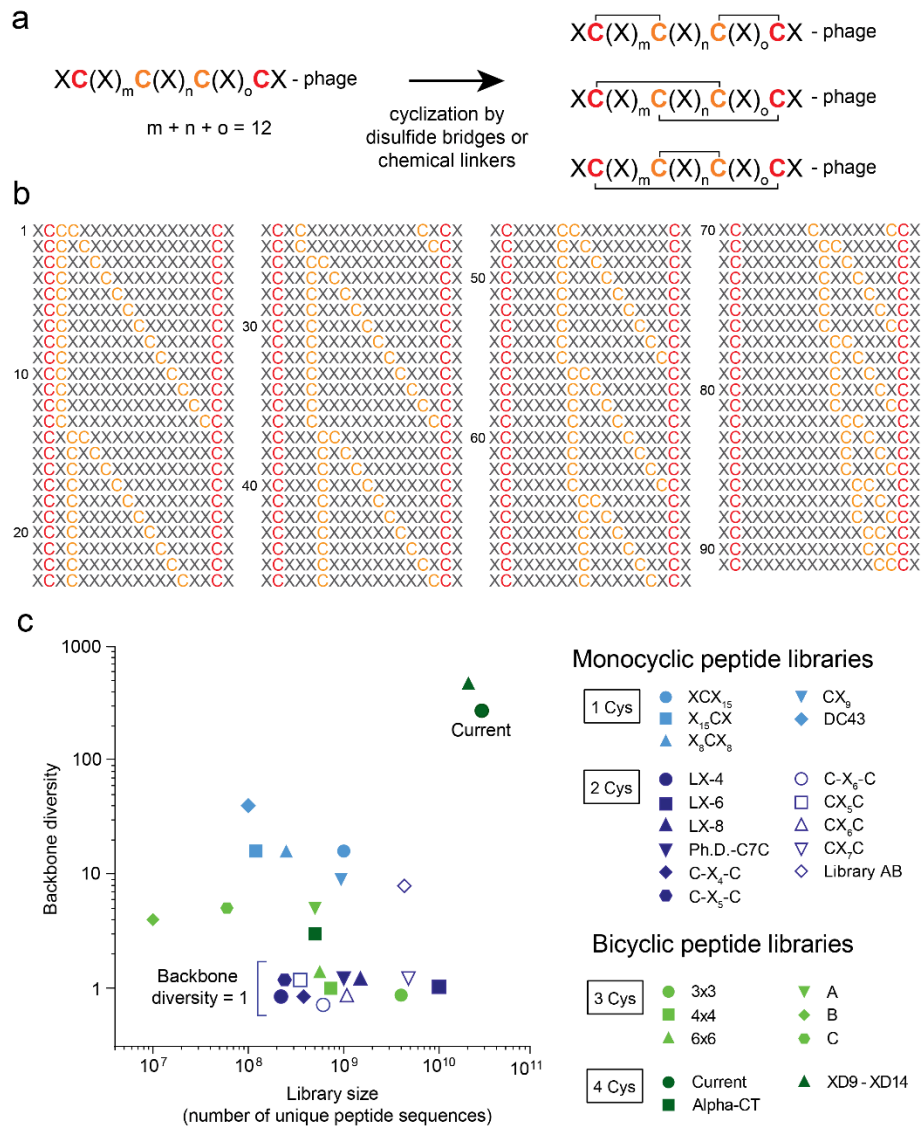


Figure 1. Cyclic peptide phage display library format and size comparison. a) Format of peptide library containing four cysteines (C) and 14 random amino acids (X). Bridging of two pairs of cysteines by disulfide bridges or chemical linkers yields bicyclic peptides. Binders based on the first format may be developed into monocyclic peptide ligands. b) Positions of the four cysteines in the peptide library. The two cysteines at the ends are kept in constant positions (in red) and the positions of the two middle cysteines (in orange) are varied to achieve 91 formats. c) Size and backbone diversity of previously reported cyclic peptide phage display libraries are compared to

the new library. For the herein developed library, the number of different monocyclic and bicyclic peptides is 4-fold larger than the number of unique sequences indicated in the graph, for the following reason: bridging pairs of cysteines yields three configurations as shown in panel a), a first one that comprises two monocyclic peptides and two configurations that are bicyclic peptides.

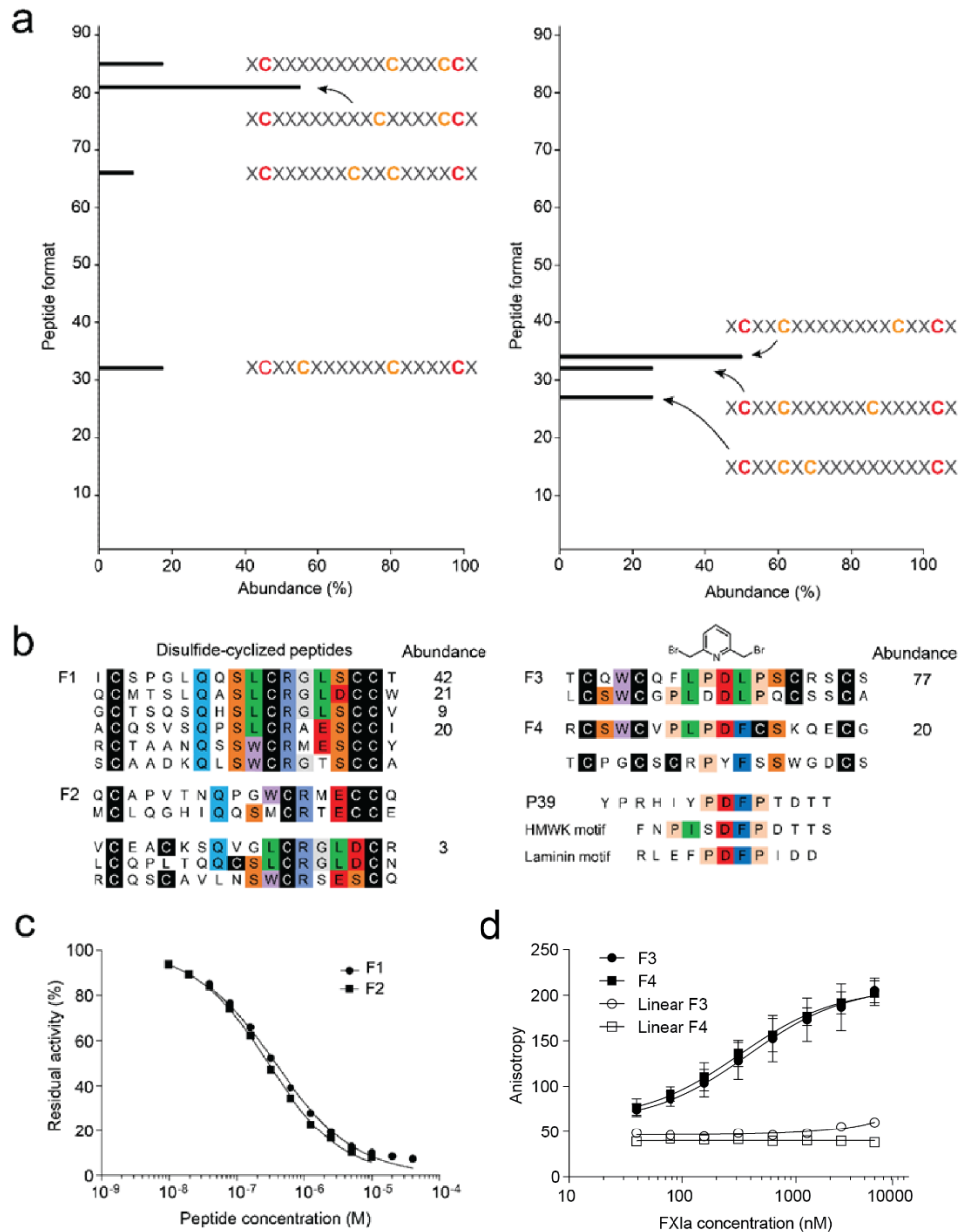


Figure 3. Cyclic peptides isolated in phage selections against human XIa. a) Abundance of peptide formats after three rounds of phage selection. The left graph shows peptides isolated from a library that was cyclized by disulfide bridges and the right graph shows peptides isolated after cyclization with the chemical linker BBMP. The abundance was calculated by dividing the number

of different peptides having the same cysteine positions by the total number of different peptides.

b) Sequences of peptides isolated against FXIa. Peptides with similar sequences were aligned and identical amino acids are highlighted in color. The abundance of each peptide is indicated.

c) Inhibitory activity of the peptides F1 and F2 shown in b). The peptides were chemically synthesized, cyclized by disulfide formation, HPLC purified and their activity tested by measuring residual activity of FXIa using a chromogenic substrate. Mean values of two independent measurements are shown.

d) Binding affinity of the peptides F3 and F4 shown in b). The peptides were chemically synthesized with fluorescein linked to the N-terminus, cyclized by alkylation with BBMP, HPLC purified and their binding to FXIa measured by fluorescence polarization. For comparison, linear peptides with the same sequences but the cysteines replaced by serines were synthesized and tested in the same assay. Mean values and standard deviations of three independent measurements are shown.

41	AGCGCCATGGCCNNKTGTNNKNNKNNKTGTNNKNNKNNKNNKTGTNNKNNKNNKNNKNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
42	AGCGCCATGGCCNNKTGTNNKNNKNNKTGTNNKNNKNNKNNKNNKTGTNNKNNKNNKNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
43	AGCGCCATGGCCNNKTGTNNKNNKNNKTGTNNKNNKNNKNNKNNKNNKTGTNNKNNKNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
44	AGCGCCATGGCCNNKTGTNNKNNKNNKTGTNNKNNKNNKNNKNNKNNKNNKTGTNNKNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
45	AGCGCCATGGCCNNKTGTNNKNNKNNKTGTNNKNNKNNKNNKNNKNNKNNKNNKTGTNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
46	AGCGCCATGGCCNNKTGTNNKNNKNNKTGTNNKNNKNNKNNKNNKNNKNNKNNKNNKTGTTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
47	AGCGCCATGGCCNNKTGTNNKNNKNNKNNKTGTTGTNNKNNKNNKNNKNNKNNKNNKNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
48	AGCGCCATGGCCNNKTGTNNKNNKNNKNNKTGTNNKTGTNNKNNKNNKNNKNNKNNKNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
49	AGCGCCATGGCCNNKTGTNNKNNKNNKNNKTGTNNKNNKTGTNNKNNKNNKNNKNNKNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
50	AGCGCCATGGCCNNKTGTNNKNNKNNKNNKTGTNNKNNKNNKTGTNNKNNKNNKNNKNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
51	AGCGCCATGGCCNNKTGTNNKNNKNNKNNKTGTNNKNNKNNKNNKNNKTGTNNKNNKNNKNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
52	AGCGCCATGGCCNNKTGTNNKNNKNNKNNKTGTNNKNNKNNKNNKNNKNNKTGTNNKNNKNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
53	AGCGCCATGGCCNNKTGTNNKNNKNNKNNKTGTNNKNNKNNKNNKNNKNNKNNKTGTNNKNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
54	AGCGCCATGGCCNNKTGTNNKNNKNNKNNKTGTNNKNNKNNKNNKNNKNNKNNKNNKTGTNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
55	AGCGCCATGGCCNNKTGTNNKNNKNNKNNKTGTNNKNNKNNKNNKNNKNNKNNKNNKNNKTGTTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
56	AGCGCCATGGCCNNKTGTNNKNNKNNKNNKNNKTGTTGTNNKNNKNNKNNKNNKNNKNNKNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
57	AGCGCCATGGCCNNKTGTNNKNNKNNKNNKNNKTGTNNKTGTNNKNNKNNKNNKNNKNNKNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
58	AGCGCCATGGCCNNKTGTNNKNNKNNKNNKNNKTGTNNKNNKTGTNNKNNKNNKNNKNNKNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
59	AGCGCCATGGCCNNKTGTNNKNNKNNKNNKNNKTGTNNKNNKNNKTGTNNKNNKNNKNNKNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
60	AGCGCCATGGCCNNKTGTNNKNNKNNKNNKNNKTGTNNKNNKNNKNNKNNKTGTNNKNNKNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
61	AGCGCCATGGCCNNKTGTNNKNNKNNKNNKNNKTGTNNKNNKNNKNNKNNKNNKTGTNNKNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
62	AGCGCCATGGCCNNKTGTNNKNNKNNKNNKNNKTGTNNKNNKNNKNNKNNKNNKNNKTGTNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
63	AGCGCCATGGCCNNKTGTNNKNNKNNKNNKNNKTGTNNKNNKNNKNNKNNKNNKNNKNNKTGTTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC
64	AGCGCCATGGCCNNKTGTNNKNNKNNKNNKNNKNNKTGTTGTNNKNNKNNKNNKNNKNNKNNKTGT NNKGGTTCTGGCGCTGAAACTGTTGAAAGTTGTTTAGC

Supplementary Table 2. PCR reactions pooled after purification by electrophoresis and extraction from agarose gel.

Group number	Forward primers	DNA concentration (ng/ μ l)	Total volume (μ l)
1	1 - 9	40	200
2	10 - 18	45	200
3	19 - 27	43	200
4	28 - 36	43	200
5	37 - 45	42	200
6	46 - 54	37	200
7	55 - 63	37	200
8	64 - 72	36	200
9	73 - 81	37	200
10	82 - 91	57	200

Supplementary Table 3. Number of bacterial cells transformed with library vector.

Determined in duplicate by plating dilutions of transformed cells on selective agar plates.

Group number	Titer 1	Titer 2	Average in total volume
1	1.80E+07	1.00E+07	1.47E+09
2	1.70E+07	2.00E+07	1.94E+09
3	2.00E+07	2.00E+07	2.10E+09
4	2.60E+07	2.00E+07	2.42E+09
5	1.90E+07	1.00E+07	1.52E+09
6	2.00E+07	3.00E+07	2.63E+09
7	4.20E+07	3.00E+07	3.78E+09
8	3.80E+07	4.00E+07	4.10E+09
9	2.30E+07	5.00E+07	3.83E+09
10	3.10E+07	5.00E+07	4.25E+09

Supplementary Figures

a

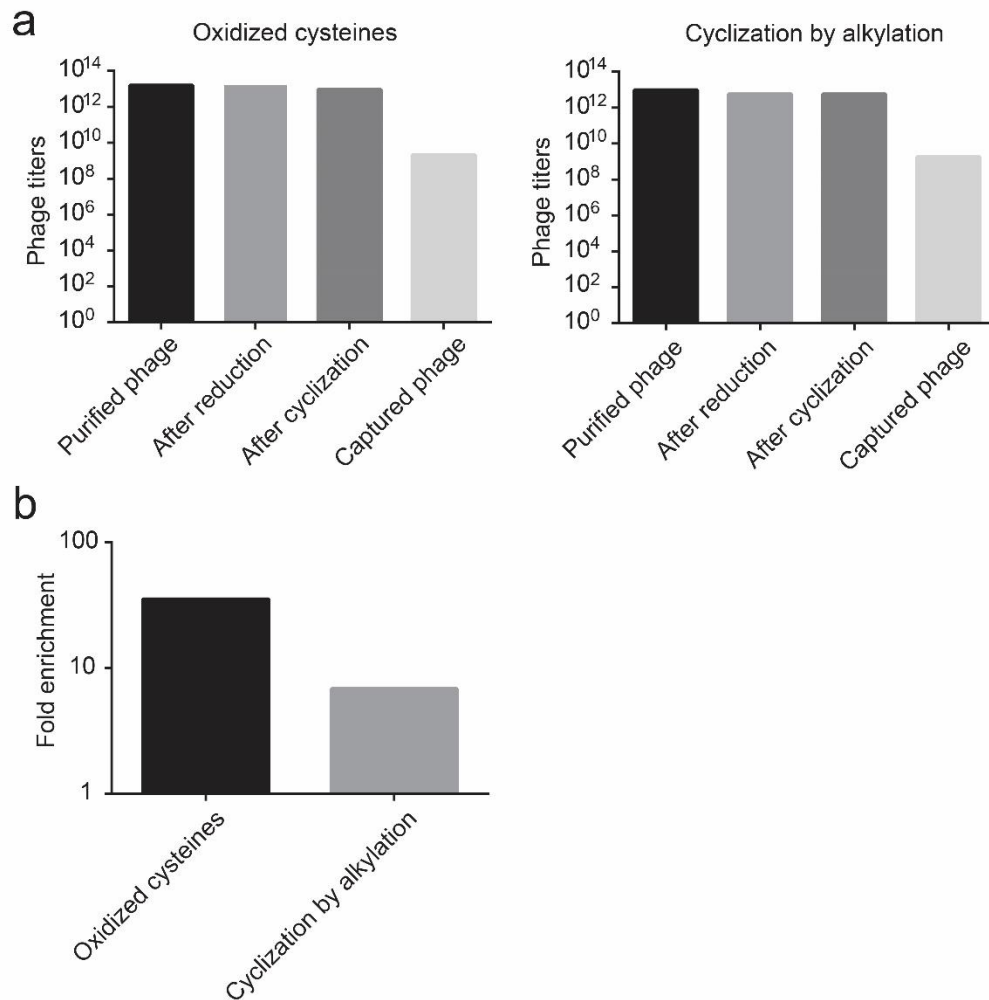
Group	Sequence	Number of Cys	Primer assigned	Group	Sequence	Number of Cys	Primer assigned		
1	M A W C C E S S R Q C P A S G T R E C H G S G	5	6. XCXXXXXXXXXXXXXXXXX	6	M A G C T K Q R C N C K H W N Q K C K G S G	4	48. XCXXXXXXXXXXXXXXXXX		
	M A S C G L V L F P S G A F S G L G S G	4	9. XCXXXXXXXXXXXXXXXXX		M A W C L S S Q C S L T C S G N I S C N G S G	4	50. XCXXXXXXXXXXXXXXXXX		
	M A P C C C L L V L P F S H K V G P C K G S G	4	1. XCXXXXXXXXXXXXXXXXX		M A T C R Y T N C T K H K N C K K T C Q G S G	4	52. XCXXXXXXXXXXXXXXXXX		
	M A C C C V F P Q E N T H R A E A K L G S G	5	2. XCXXXXXXXXXXXXXXXXX		M A F C S T R L C C L Q I I S N S T C L G S G	4	47. XCXXXXXXXXXXXXXXXXX		
	M A P C C T G O P V L I V V E V C C C G S G	6	2. XCXXXXXXXXXXXXXXXXX		M A Y C L D A A C R L V S L C K G G C G S G	4	52. XCXXXXXXXXXXXXXXXXX		
	M A A C C Y L S F P C S T R G Y C C L G S G	4	7. XCXXXXXXXXXXXXXXXXX		M A V C H R L C A R L Y P A I T C C V G S G	5	46. XCXXXXXXXXXXXXXXXXX		
	M A F C C G S R T E R M R D T Y K C Y G S G	5			M A T C E T T R C G C P G T R N L C K G S G	4	49. XCXXXXXXXXXXXXXXXXX		
	M A S C C V C D R C G Q E W F V D C C G S G	6			M A C C A S L C C R R P R A D A S C L G S G	5	47. XCXXXXXXXXXXXXXXXXX		
	M A A C P P S P P F Y S A A P S C C F G S G	3			M A T C P S H G C L S F P T V F S R C R V S G	3			
	M A S C C V M C C M M S P Y L I K R C K G S G	4	4. XCXXXXXXXXXXXXXXXXX		M A M C N S R F C A C R P O L F F G C H G S G	4	48. XCXXXXXXXXXXXXXXXXX		
	M A Q C G M T Q A H R C H L S L C O T G S G	4	9. XCXXXXXXXXXXXXXXXXX		M A A C P M A K C V S L V C P Y D H C Y G S G	4	51. XCXXXXXXXXXXXXXXXXX		
	M A W C G C P A S S W C G G R S R G W G S G	5	1. XCXXXXXXXXXXXXXXXXX		M A E C Q K R Q C G R K K R P N G C P G S G	4	49. XCXXXXXXXXXXXXXXXXX		
	M A P C C S Q K K C S N C G Q V E R C P G S G	5	5. XCXXXXXXXXXXXXXXXXX		M A E C L G D L C S R L L C C C S Q C I G S G	5			
			7 primers out of 9		M A R C N S F S C T T R S Q C N I T C D G S G	4	52. XCXXXXXXXXXXXXXXXXX		
					M A S C H P R P C C S D P S S P H C T G S G	5	48. XCXXXXXXXXXXXXXXXXX		
					M A M C E S L S C P C T R Q N Q S N C K G S G	5	49. XCXXXXXXXXXXXXXXXXX		
2	M A A C C E A K A R W R R R C R R R C R G S G	4	10. XCXXXXXXXXXXXXXXXXX	7	M A S C L I R N L C G L G V S V C L C Q G S G	4	62. XCXXXXXXXXXXXXXXXXX		
	M A P C S G P M S L Q K M F M L G E C S G	3	18. XCXXXXXXXXXXXXXXXXX		M A T C Q P T L T C M C R T R P R C T G S G	4	57. XCXXXXXXXXXXXXXXXXX		
	M A E C C V Y C T S S R H T I G R C R G S G	5	13. XCXXXXXXXXXXXXXXXXX		M A O C P L G K C V T P E C C N A C G S G	4	61. XCXXXXXXXXXXXXXXXXX		
	M A A C C A R T A G R A C S F R C C P G S G	4	12. XCXXXXXXXXXXXXXXXXX		M A Q C A W A R E D P P H C T F H F G S G	4	60. XCXXXXXXXXXXXXXXXXX		
	M A V C L G Y L S L C L G V A G R C C G S G	7	18. XCXXXXXXXXXXXXXXXXX		M A L T P R C G K T A G N L N C S G S G	4	60. XCXXXXXXXXXXXXXXXXX		
	M A C C L C G R L W R E Q P F C A S C R G S G	5	15. XCXXXXXXXXXXXXXXXXX		M A R C Q T P H R C D Q H T P K C H C G S G	4	62. XCXXXXXXXXXXXXXXXXX		
	M A M C V G C G S T S L S L C L W R S V G S G	6			M A S C Q R S D G C P P A C C P P R C G S G	4	69. XCXXXXXXXXXXXXXXXXX		
	M A A C C S G C L D G G R H F Y C C L G S G	6	12. XCXXXXXXXXXXXXXXXXX		M A R C H P E Y P C P P R Q P T C P K G S G	4	62. XCXXXXXXXXXXXXXXXXX		
	M A A C P C F F E L C S S W S Y F R C L G S G	4	18. XCXXXXXXXXXXXXXXXXX		M A V C G H I Y P S L C Q T F C C G S G	5	61. XCXXXXXXXXXXXXXXXXX		
	M A A C G Y P L T I H Q K R S N F L C O I G S G	4	13. XCXXXXXXXXXXXXXXXXX		M A F C T W S T C A N C S S I F G S G	4	61. XCXXXXXXXXXXXXXXXXX		
	M A A C C G S A P C S L V T C C F Y C P G S G	6	17. XCXXXXXXXXXXXXXXXXX		M A Y C Q P Q C C G P C S P P T R C R G S G	5	58. XCXXXXXXXXXXXXXXXXX		
	M A A C C L V S R C C G S F G C S F G S G	7	10. XCXXXXXXXXXXXXXXXXX		M A S C C P L T C G L R Q D L S P C R G S G	5	56. XCXXXXXXXXXXXXXXXXX		
	M A A C R C W S A L C L W Q G L F R C R G S G	6	16. XCXXXXXXXXXXXXXXXXX		M A P C L V P T Y C W L C A N S Q P C D G S G	4	58. XCXXXXXXXXXXXXXXXXX		
			6 primers out of 9				7 primers out of 9		
	3	M A A C C A A C A S R S P A C R P C H G S G	5		26. XCXXXXXXXXXXXXXXXXX	8	M A A C Y L I N L P C S S A Y S L P C Y G S G	4	64. XCXXXXXXXXXXXXXXXXX
		M A S C L C L P Y L C A T P A Q W G S G	4		20. XCXXXXXXXXXXXXXXXXX		M A Q C K P S H Q D C T C K Y K G L C Q G S G	4	65. XCXXXXXXXXXXXXXXXXX
M A E C L C H P T P W H R H C A G Y C G S G		5	22. XCXXXXXXXXXXXXXXXXX	M A A C S C W T P C G S G C L V S H C G S G	5		66. XCXXXXXXXXXXXXXXXXX		
M A L C W C S E S F F C C S S P S C W G S G		6		M A A C S L T H S S C P T T F S P L C C G S G	3				
M A D C R W Q T I T K R S P C P P C H G S G		4	23. XCXXXXXXXXXXXXXXXXX	M A N C R P P P P C S H S L G L C D G S G	4		65. XCXXXXXXXXXXXXXXXXX		
M A A C P L C S S F F V P R S L L F C V V S G		4	26. XCXXXXXXXXXXXXXXXXX	M A H C V S P P I H C G R R R N T C C R G S G	5		65. XCXXXXXXXXXXXXXXXXX		
M A F R L C G T S R S A R T I N W C T G S G		4	28. XCXXXXXXXXXXXXXXXXX	M A I C W W R N G D C P C H S Y G C A G S G	4		65. XCXXXXXXXXXXXXXXXXX		
M A K C G C K N T K N T T K D C N T C K G S G		4	23. XCXXXXXXXXXXXXXXXXX	M A L C R P R P P K C G R R Q G D P C R G S G	4		64. XCXXXXXXXXXXXXXXXXX		
M A P C P L F H C T T K T R N R S L C D G S G		4	27. XCXXXXXXXXXXXXXXXXX	M A A C Q L H N M Y C W T R T C L M C F G S G	4		68. XCXXXXXXXXXXXXXXXXX		
			5 primers out of 9				6 primers out of 9		
4		M A A C C M S G H P R O S S E P S C C P G S G	4	35. XCXXXXXXXXXXXXXXXXX	9		M A P C C S E S A I I Q R C P V C G Y C F G S G	4	79. XCXXXXXXXXXXXXXXXXX
		M A A C G V A G S P C S Y R Y R C C G S G	6				M A P C C Q R V R M T C A S I C W S C L G S G	5	74. XCXXXXXXXXXXXXXXXXX
		M A G Q S G C I A R A T C D V G E C G G S G	3	32. XCXXXXXXXXXXXXXXXXX			M A H C L R W L S W F L G P P D L C C G S G	4	81. XCXXXXXXXXXXXXXXXXX
		M A M C R M C G L F Q F C C G H G T G S G	5				M A R C E Q P G F V G C S A W L C S G S G	4	78. XCXXXXXXXXXXXXXXXXX
		M A A C D C C L A C C C R R C C O L G S G	8				M A L C L R S L Y P S G T E P C R G S G	4	82. XCXXXXXXXXXXXXXXXXX
		M A A C A A C C I K Q P R C E G T L C G S G	5	32. XCXXXXXXXXXXXXXXXXX			M A L C T H L T P L R Y C C Q T N P C C G S G	5	77. XCXXXXXXXXXXXXXXXXX
	M A C C G R C T T F R M A C K S L L V G S G	6	32. XCXXXXXXXXXXXXXXXXX	M A W C P P N T K C L I C C A F F W C P G S G		5	77. XCXXXXXXXXXXXXXXXXX		
	M A A C L L C S V C F F R A C C G F C G S G	8		M A L C S P N T C C S N R C H L I C S C P G S G		6	80. XCXXXXXXXXXXXXXXXXX		
	M A S C L Q R R A D A Q C P R R E C T G S G	4	32. XCXXXXXXXXXXXXXXXXX	M A G C A P R R P L L C N S C T L R C R G S G		4	73. XCXXXXXXXXXXXXXXXXX		
	M A P C H L C Q K C S N P P T T K T C R G S G	4	28. XCXXXXXXXXXXXXXXXXX	M A P C H R R Q P K C R P K T H C P G S G		4	76. XCXXXXXXXXXXXXXXXXX		
			3 primers out of 9	M A T C F L Q Y T A L L C R V S T C C F G S G		4	81. XCXXXXXXXXXXXXXXXXX		
	5	M A S C M Q R C V K W W C F P W C S C G G S G	5			10	M A G C T R P P A L N R L C V C C L G S G	5	87. XCXXXXXXXXXXXXXXXXX
		M A K C Q H F G N I Y A C Q R H C C G G S G	5	42. XCXXXXXXXXXXXXXXXXX			M A G C S T F G V S T F Q P A C R C R G S G	4	87. XCXXXXXXXXXXXXXXXXX
		M A A C P L F C T P S N C G W V L C G G S G	5	42. XCXXXXXXXXXXXXXXXXX			M A L C S L D R F T V S C C D L C F G S G	4	85. XCXXXXXXXXXXXXXXXXX
		M A M C L A M G A Y W C S K P C H G C P G S G	5				M A H F D H S E F R P P L G S C C S G S G	4	
		M A V C P M G C A R M C D T N C R R C L G S G	5				M A A C G S I Q P L S V P G Y A C C C A G S G	4	91. XCXXXXXXXXXXXXXXXXX
M A D C L E V C T H Q W R G F C V C N G S G		4	45. XCXXXXXXXXXXXXXXXXX	M A I C T A R A A R W E A C C F D C L G S G	4		86. XCXXXXXXXXXXXXXXXXX		
M A E C L V N C S L S C W A S L L W R G S G		4		M A M C S K S N T S L S R C T C P P C E G S G	4		83. XCXXXXXXXXXXXXXXXXX		
M A R C R V T C A Y D E G S C R C C T G S G		4	44. XCXXXXXXXXXXXXXXXXX	M A S C S R S R G H S C P G W S L L L C G S G	2				
M A T C A T T C K R M A T A Q C R O C E G S G		4	44. XCXXXXXXXXXXXXXXXXX	M A P C G W M V I I R Q P I A C C G S G	4		91. XCXXXXXXXXXXXXXXXXX		
M A S C L L F C H Q C C W G Y D L C C G S G		6		M A V C M Q T E Q K S M A D C R A L C C G S G	4		86. XCXXXXXXXXXXXXXXXXX		
M A D G L T F E G F F C S Q L F L F L G S G		3		M A R C V P P S W L S Y I C A C M L C P G S G	4		83. XCXXXXXXXXXXXXXXXXX		
M A Y C H A C C S P R Q G C A C S C G S G		6	44. XCXXXXXXXXXXXXXXXXX	M A T C K T P A S P R N L M C S C H G S G	4		87. XCXXXXXXXXXXXXXXXXX		
M A F C R G G C G N K Q R W G F C S C K G S G		6					6 primers out of 9		
M A D C L E V C T H Q W R G F S C C G N G S G		4	44. XCXXXXXXXXXXXXXXXXX						

b

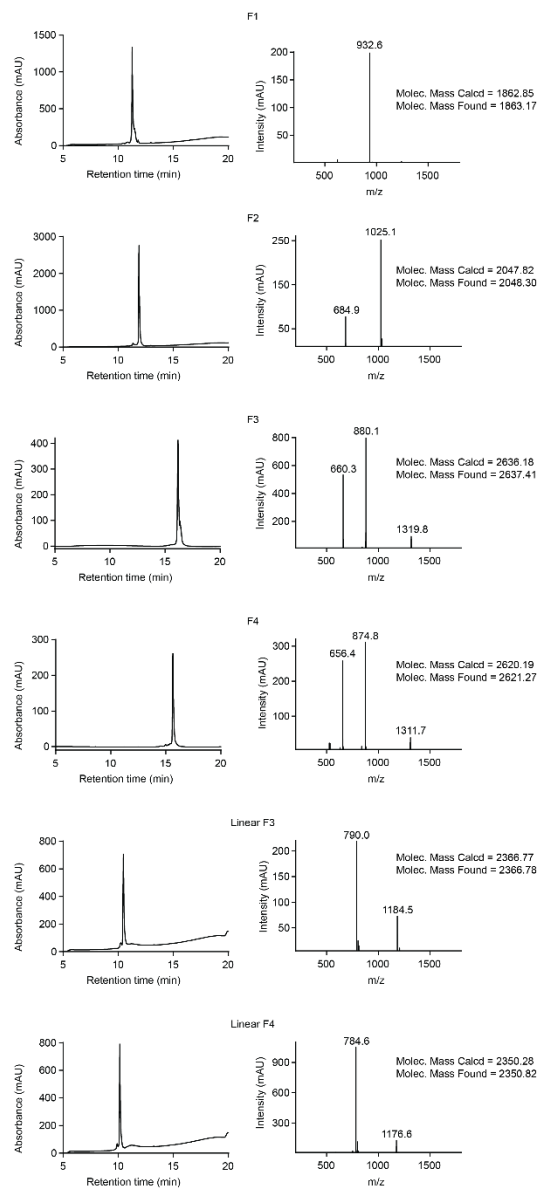
	X	C	X	X	X	X	X	X	X	X	X	X	X	X	C	X		
Number of Cys	2	114	12	20	17	22	17	24	24	15	25	18	26	20	21	18	113	7
Expected	4	125	21	21	21	21	21	21	21	21	21	21	21	21	21	21	125	4

Peptides with 4 Cys: 70, 56 % (exp. 64.1 %)
 Peptides with ≥ 5 Cys: 47, 37.6 % of sequenced clones (exp. 36 %)

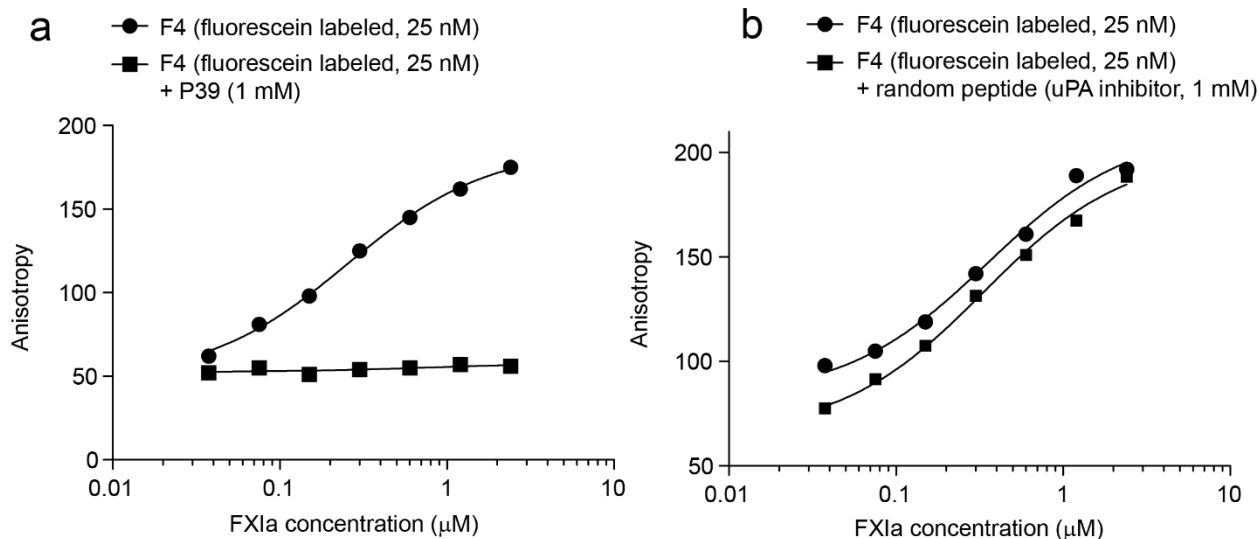
Supplementary Figure 1. Sequences of peptides in phage library. The DNA coding the peptide was sequenced for 125 clones of the library (around 12 clones for each pool of primers). Due to NNK codons used, cysteine were also expected for the first and last randomized position (1/32 × 125 sequenced clones = 4 cysteines expected).



Supplementary Figure 2. Phage titers at different stages of phage selection in round 3. a) Titters at different steps of library production (after PEG purification, after reduction, after cyclization) and after the biopanning with FXIa, shown for disulfide-cyclized peptides (left) and linker-cyclized peptides (right). b) Enrichment of phage defined as the ratio of peptides isolated with streptavidin beads carrying FXIa and peptides isolated to streptavidin beads without target protein.



Supplementary Figure 3. Analytical data of peptides F1 to F4 and linear analogs of F3 and F4. Chromatograms of peptides analyzed by C18 reversed-phase HPLC are shown on the left. Mass spectra of peptides are shown on the right, with the expected and experimentally measured molecular mass indicated.



Supplementary Figure 4. Competition binding experiment with apple 2 domain binding peptide P39. a) Binding curve of F4 (fluorescein labeled) measured by fluorescence polarization in presence or absence of peptide P39 at a concentration of 1 mM. b) Negative control competition experiment with F4 in presence of a random peptide (uPA inhibitor UK18, ACSRYEVDRCRGRGSACG) at a concentration of 1 mM.